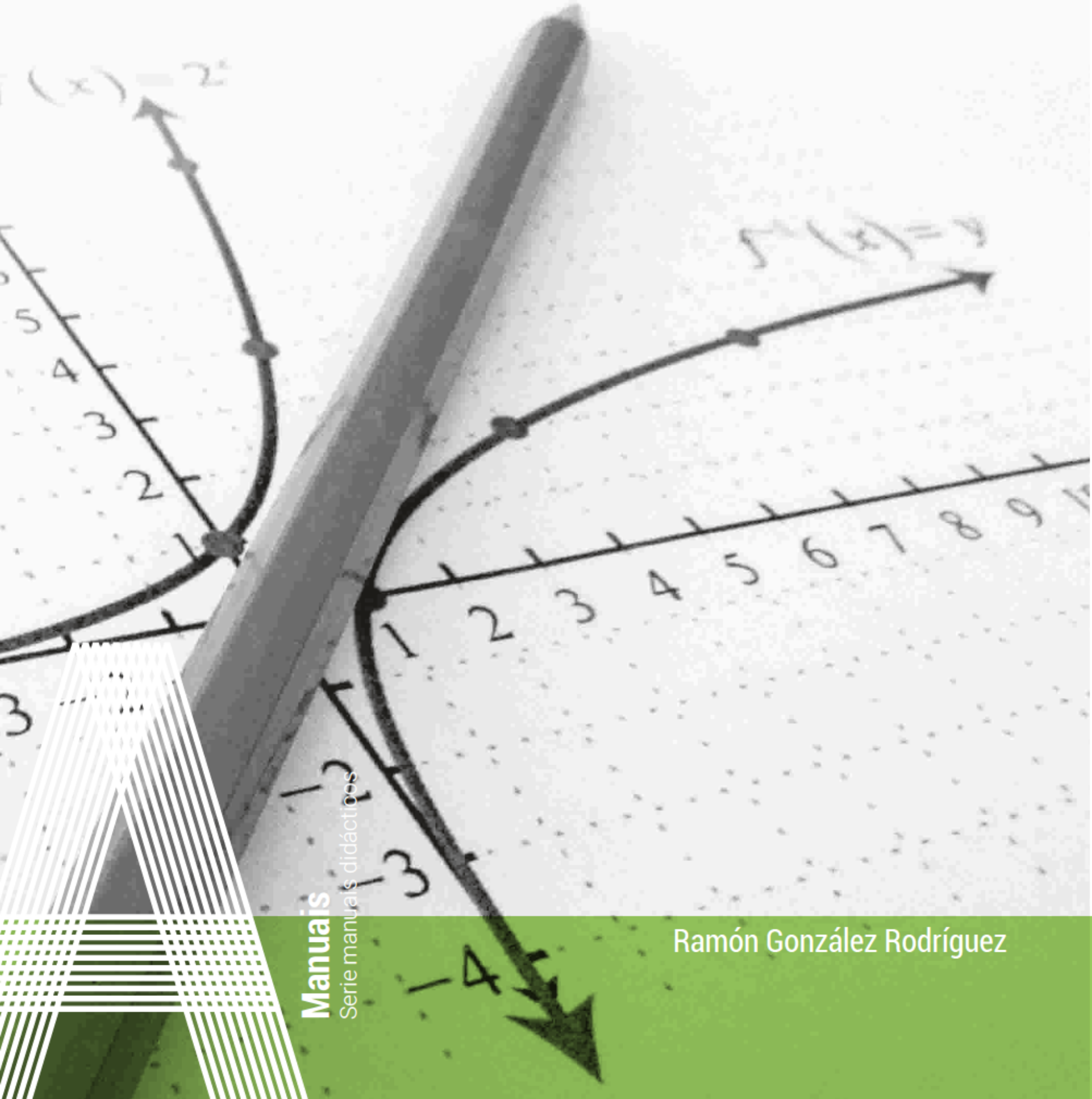
Álxebra linear

Historia, teoría e práctica



Manuais
Serie manuais didácticos

Ramón González Rodríguez

Ramón González Rodríguez

Ramón González Rodríguez, nado en 1965 en O Porriño, é Licenciado en Matemáticas e doutor en Matemáticas pola Universidade de Santiago de Compostela.
É membro do Departamento de Matemática Aplicada II da Universidade de Vigo e desde 2019 pertence ao corpo de Catedráticos de Universidade.
A súa liña de investigación encádrase dentro da álxebra non conmutativa e en particular no estudo de estruturas alxébricas con interese en física, como as álxebras de Hopf e as súas xeneralizacións, utilizando técnicas categóricas.
Foi membro do Comité Científico da Rede Española de Álxebra non Conmutativa así como o responsable do nodo galego da devandita rede.
Desde 1996 dirixiu nove proxectos de investigación e participou en catorce sendo, na actualidade, membro do equipo do proxecto PID2020-115155 GB-I00 (Homoloxía, homotopía e invariantes categóricos en grupos e álxebras non asociativas).
É autor de 68 publicacións (3 en prensa e 61 delas en revistas do JCR) e nos últimos anos asistiu como conferenciante invitado a numerosos congresos relevantes dentro da súa área de investigación.

Servizo de Publicacións
Universida de Vigo

Manuais
Serie de manuais didácticos

n.º 079

Edición
Universidade de Vigo
Servizo de Publicacións
Rúa de Leonardo da Vinci, s/n
36310 Vigo

Deseño da portada
Tania Sueiro Graña
Área de Imaxe
Vicerreitoría de Comunicacións e Relacións Institucionais

Maquetación
Tórculo Comunicación Gráfica, S. A.

Fotografía da portada
Adobe Stock

Impresión
Tórculo Comunicación Gráfica, S. A.

ISBN (Libro impreso)
978-84-8158-919-1

Depósito legal
VG 722-2021

Ao ser esta editorial membro da une, garántense a difusión e a comercialización das súas publicacións no ámbito nacional e internacional.

Servizo de Publicacións
UniversidadeVigo

Álxebra linear

Historia, teoría e práctica

Ramón González Rodríguez

Índice xeral

CAPíTULO 1

Un pouco de historia

A álxebra linear moderna ten a súa base na teoría dos espazos vectoriais ou, de forma máis xeral, na teoría de módulos sobre un anel. Ao redor do ano 1930 esta formulación da álxebra linear unificou o suxeito e fixo del unha parte da álxebra abstracta ou moderna. O estudo de sistemas de ecuacións lineares e a investigación dun cálculo xeométrico intrínseco foron historicamente as principais fontes do desenvolvemento da teoría da linearidade. A superación da dimensión 3 en xeometría, a mediados do século XIX, así como o desenvolvemento dialéctico entre a álxebra e a xeometría, despois da creación da xeometría analítica, levaron á unificación dos aspectos lineares ao redor da noción de determinante. Este cadro foi xeneralizado á dimensión infinita enumerábel cos traballos de análise funcional. A axiomatización realizada a finais do século XIX, pero realmente utilizada despois de 1920, é un proceso máis longo que se inscribe nun desenvolvemento xeral da matemática do século XX. Hoxe en día a álxebra linear é unha parte esencial das ferramentas requiridas no estudo de moitas áreas nas ciencias do comportamento, ciencias naturais, físicas e sociais, en enxeñería, en economía, en ciencia dos computadores e desde logo en matemáticas puras e aplicadas. Neste capítulo preténdese analizar de forma resumida tanto as orixes históricas como o desenvolvemento desta disciplina da que podemos afirmar que é á vez unha das ramas máis antigas e máis modernas da matemática.

Comezaremos esta sección cun breve resumo sobre a historia da álxebra, para pasar a continuación á historia propiamente dita da álxebra linear.

1.1. Evolución histórica da álxebra

As operacións alxébricas máis simples (operacións aritméticas con enteiros positivos e con números racionais positivos) pódense atopar nos textos matemáticos máis antigos. Isto indica que as propiedades fundamentais desas operacións eran coñecidas desde épocas temperás polas principais civilizacións como, por exemplo, a babilónica, a exipcia ou a grega. Os primeiros documentos que se conservan sobre a matemática exipcia e babilónica demóstrannos que estas antigas civilizacións xa posuían un sistema de regras de cálculo para os naturais, os racionais positivos, as lonxitudes e as áreas. A pesar de que a información que nos chegou se refire exclusivamente a problemas con datos numéricos concretos, non existen grandes dúbidas sobre o carácter xeral das regras empregadas e o dominio alcanzado no manexo das ecuacións de primeiro e segundo grao. Doutra banda, nestes tempos remotos, non se atopa a menor preocupación por xustificar as regras empregadas e, tampouco, de dar unha definición precisa das operacións que aparecen.

É na matemática grega da época clásica cando empeza a manifestarse de forma clara unha preocupación pola precisión e o rigor en canto ás definicións e as técnicas empregadas. A pesar de non aparecer aínda un tratamento axiomático dos enteiros naturais (tratamento que non aparecerá ata finais do século XIX), hai numerosas pasaxes dos *Elementos* de Euclides de Alexandría (365–275 a.C.) nos que se inclúen demostracións formais de regras de cálculo tan evidentes intuitivamente como as de cálculo con enteiros. As demostracións máis relevantes son as que se refiren á teoría das magnitudes, a creación máis importante da matemática grega equivalente á nosa teoría de números naturais positivos. Naquela, por exemplo, Euclides considera o produto de dúas razóns de magnitudes e demostra que é independente da forma en que aparezan estas razóns (primeiro exemplo de cociente dunha lei de composición por unha relación de equivalencia) e que é conmutativa.

Con todo, non se debe ocultar que este camiño a prol do rigor vai acompañado nos traballos de Euclides de Alexandría dun certo estancamento e, ás veces, dun retroceso na técnica de cálculo alxébrico. O predominio avasalador da xeometría deu lugar a unha parálise no crecemento autónomo das notacións alxébricas xa que os elementos que aparecen nos cálculos debíanse representar sempre de forma xeométrica. Ademais, doutra banda, as leis de composición que interveñen non están definidas sobre o mesmo conxunto (a suma de magnitudes non está sempre definida, e o produto de lonxitudes non é unha lonxitude senón unha área); todo isto xerou unha serie de dificultades que fixeron moi problemático o manexo de relacións alxébricas de grao superior a dous.

Xa no declinar da matemática grega, a *Aritmética* de Diofanto de Alexandría (século III d. C. [200/204–284/298]) é un dos textos que maior influencia tivo no desenvolvemento das ideas alxébricas e dos seus símbolos. A *Aritmética* está dedicada case exclusivamente á resolución exacta de ecuacións, agora ben, non é un texto de álxebra, senón unha colección de problemas sobre aplicacións da álxebra. De feito a concepción axiomática das leis de composición parece tan afastada da mente deste matemático como da dos seus continuadores inmediatos. Desde este punto de vista, Diofanto de Alexandría parécese máis aos seus colegas alxebristas babilónicos, coa condición de que os seus números son completamente abstractos e non se refiren a medidas de gran, dimensións de campos ou unidades monetarias, como era o caso da álxebra exipcia ou mesopotámica.

Aínda que Diofanto é tratado nalgúns lugares como pai da álxebra, o termo álxebra derivouse do título do traballo máis importante do matemático árabe M. Al-Khuwärizmi (780–850). No *Al-jabr al-muqabala* M. Al-Khuwärizmi describe métodos xerais para resolver problemas que se poden reducir a ecuacións alxébricas de primeiro e de segundo grao.

A pesar de que algúns problemas alxébricos foron expostos e resoltos desde os babilonios, o desenvolvemento da álxebra ata finais da Idade Media foi lento debido sobre todo á ausencia de notacións adecuadas e ao concepto restrinxido de número que se manexaba. A introdución do cero e dos números negativos por parte dos matemáticos indios, así como a creación dos números imaxinarios por parte dos alxebristas italianos do século XVI supuxeron avances fundamentais.

Se deixamos o cero á marxe, introducido como símbolo de numeración antes de ser considerado como número, o carácter común destas extensións é o de ser, na súa orixe, puramente formais. Isto débese entender de forma de que os novos números apareceron como o resultado de operacións realizadas en condicións nas cales non terían sentido (por exemplo, a diferenza $a - b$ cando $a < b$): de aí a denominación de números falsos, ficticios, imaxinarios, etc. Para os matemáticos gregos da época clásica este tipo de extensións eran difíciles de concibir e, como é natural, só podían proceder de mentes máis dispostas a depositar unha confianza case mística nos seus métodos e, doutra banda, coa confianza de deixarse levar polo mecanismo dos seus cálculos sen pararse a comprobar o lícito de cada paso. Esta confianza quedaba xustificada e reafirmada posteriormente pola obtención de resultados exactos. Por exemplo, en canto a isto último, os indios xa eran conscientes da interpretación que debe darse aos números negativos nalgúns casos (por exemplo, unha débeda nun problema comercial). Parte destes avances xa quedan patentes na obra de Aryabhata (476–550) onde o sistema de notación posicional, visto por primeira vez no Manuscrito Bakhshali do século III, estaba claramente contido sendo o cero un marcador de posición para as potencias de dez con coeficientes nulos.

Nos séculos seguintes, na medida en que os árabes difunden os métodos da matemática grega e india, comeza a familiaridade con estes números e empézanse a dar novas interpretacións deles. Isto é, xunto coa mellora continuada das notacións empregadas, o único progreso notábel da álxebra na Idade Media onde merece especial mención o matemático italiano, educado no norte de África, Leonardo de Pisa Fibonacci (1170–1240) quen pasou á posteridade pola publicación, en 1202, do seu soado *Liber Abacci* onde entre outras cousas introduciu a numeración hindú-árabe. Esta obra non foi superada ata varios séculos despois e xunto co *Liber quadratorum*, escrito en 1225, foron fonte de inspiración matemática ata a actualidade.

Coa chegada de F. Viète (1540–1603) e R. Descartes (1596–1650) a álxebra recibe un novo impulso, sendo o principal acontecemento alxébrico da primeira metade do século XVI a resolución das ecuacións cúbica e cuártica. Estes matemáticos proporcionan notacións moi parecidas ás que actualmente están vixentes tanto para os datos e as incógnitas que figuran nas ecuacións, como para as operacións alxébricas e os expoñentes (os índices non son case utilizados ata G. W. Leibniz (1646–1716) e I. Newton (1642–1727)).

Durante este século publicáronse libros de aritmética e de álxebra en distintos países de Europa. Así en Alemaña, a álxebra tomou o nome de *Die Coss*, é dicir a cousa, nome que en Italia designaba á incógnita. A primeira obra publicada en alemán vulgar, en 1525, débese a J. C. Rudolff (1499–1545). Alí aparece, por primeira vez, o símbolo $\surd$ para indicar a raíz cadrada. O signo $=$ aparece por primeira vez en *The Whetstone of Witte (O aguzador do enxeño)* publicado en 1557 por R. Recorde (1510–1558), que é o primeiro tratado inglés de álxebra, onde o autor afirma que elixiu ese símbolo porque dúas cousas non poden ser máis iguais que dúas rectas paralelas. Con R. Bombelli (1530–1573), a álxebra italiana alcanza a súa cima máis alta. O autor de *L'algebra, parte maggiore dell'aritmetica, divisa in tre libri*, en 1572, foi o primeiro matemático, e único durante moito tempo, que tivo a audacia de aceptar a existencia dos números imaxinarios, achegando así una certa claridade ao enigma do caso irredutíbel das ecuacións de terceiro grao, operando coas raíces cadradas das magnitudes negativas aplicándolles as regras elaboradas para o cálculo das raíces dos números positivos.

Antes de F. Viète, os coeficientes das ecuacións eran números dados explicitamente, sendo un dos maiores méritos deste matemático, o utilizar nas cuestións alxébricas cantidades calquera e, por tanto, a introdución do uso sistemático das letras. Ao inglés T. Harriot (1560–1621), a quen se debe a importante innovación no simbolismo de indicar as potencias mediante factores repetidos, tamén se lle debe a introdución dos símbolos maior e menor, utilizando nalgunha ocasión o punto como símbolo de multiplicación, aínda que como tal, o punto non se difundiu ata o século XVIII grazas á obra de G. W. Leibniz. O símbolo $\times$ para a multiplicación débese a W. Oughtred (1574–1660), quen achegou entre propios e non propios 150 signos matemáticos.

A álxebra renacentista enriqueceu prodixiosamente os coñecementos alxébricos, elaborou unha notación moi compacta e nada incómoda, pero foi incapaz de elevarse ata a noción abstracta de operación alxébrica e tomar esta como centro das súas reflexións.

Nos séculos XVII e XVIII a álxebra entendeuse como a ciencia dos cálculos que involucraban símbolos alxébricos (transformacións de fórmulas consistentes en letras, solución de ecuacións alxébricas) distinguíndose da aritmética que quedaba relegada aos cálculos con números concretos. Un dos tratados máis importantes desta época titulado *Algebra* débese a L. Euler (1707–1783). A influencia desta obra na determinación da problemática científica da álxebra e na estrutura do curso de álxebra nas universidades foi moi grande. O carácter monográfico deste libro e os obxectivos que expuña permitían xulgar o estado da álxebra na segunda metade do século XVIII. Constaba de dúas partes: na primeira, tratábase fundamentalmente de xeneralizar as regras de resolución de problemas aritméticos e de desenvolver o aparello simbólico da álxebra. Así, na primeira sección, acláranse as operacións sobre números e monomios, sobre radicais e números complexos e introdúcense os logaritmos. Ademais danse regras de extracción de raíces dos números e as expresións alxébricas (polinomiais). A segunda parte estaba dedicada aos métodos de resolución de ecuacións alxébricas dos primeiros catro graos e á súa teoría xeneral. Tamén, trátanse os métodos de cálculo aproximado das raíces das ecuacións alxébricas.

As letras e, π, i, cos significados actuais, débense tamén a L. Euler, que se relacionan cos enteiros 0 e 1, mediante a igualdade

$$e^{\pi i} + 1 = 0,$$

na que figuran os cinco números e as máis importantes operacións das matemáticas.

L. Euler
(Emanuel Handmann/Wikimedia Commons)

Un dos campos fundamentais de investigación dentro da álxebra durante os séculos anteriormente citados, e que se estenderá ao século XIX, foi o estudo dos polinomios e as súas raíces. Historicamente, o problema orixinario consistiu no cálculo de raíces de polinomios nunha incógnita, isto é, expresións da forma

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0.$$

Nun principio o propósito foi o de atopar fórmulas que expresasen as raíces dos polinomios en función dos seus coeficientes. Nestas fórmulas podían intervir a adición, a multiplicación, a subtracción, a división e a extracción de raíces (solución por radicais). A pesar de que a solución para os casos de grao un e dous eran coñecidos desde os tempos máis temperáns, non é ata a chegada do século XVI cando se conseguen logros substanciais. Estes importantes avances son debidos a matemáticos italianos como G. Cardano (1501–1576), que atopou unha fórmula para a ecuación de terceiro grao, ou como L. Ferrari (1522–1565) que conseguiu un método de solución para as ecuacións de cuarto grao. A partir de aquí fixéronse grandes esforzos para atopar fórmulas similares para resolver ecuacións de graos maiores. En conexión con isto, a utilización dos números complexos marca un fito no posterior desenvolvemento da álxebra.

C. F. Gauss
(Deutsche Bundesbank/Wikimedia Commons)

Os números imaxinarios apareceron, como xa o citamos, no século XVI con motivo da resolución das ecuacións de terceiro grao e o seu estudo foi avanzando baixo a presión das necesidades da análise matemática. O primeiro que elaborou (ao parecer, en interese da práctica xeodésica e cartográfica) un procedemento de interpretación xeométrica dos números complexos como puntos do plano foi o agrimensor danés K. Wessel (1745–1818) en 1797. Con todo, o seu traballo resultou inadvertido, así como unha interpretación análoga de J. R. Argand (1768–1812) en 1806. Na segunda década do século XIX, C. F. Gauss (1777–1855) e A. L. Cauchy (1789–1857) introduciron e fundamentaron as operacións cos números de forma $a+bi$ (o símbolo i para designar a $\sqrt{-1}$ foi introducido por L. Euler en 1777 e C. F. Gauss foi o primeiro que sistematizou o seu uso), a noción de módulo, a de norma e a de conxugado dun número complexo. Isto trouxo consigo a súa consolidación e por extensión a súa entrada dentro do mundo da álxebra.

En 1629, A. Girard (1590–1633) afirmou que toda ecuación alxébrica de grao n posúe n raíces. Este enunciado foi recollido de modo cada vez máis preciso por R. Descartes en 1637, I. Newton en 1685 e L. Euler en 1742. Aínda que a demostración do chamado *Teorema Fundamental da Álxebra* foi acometida, entre outros, por J. D'Alambert (1717–1783) en 1746 e L. Euler en 1751. A primeira demostración rigorosa débese a C. F. Gauss o cal presentou posteriormente outras. O feito de que toda ecuación alxébrica posúa polo menos unha raíz, real ou imaxinaria, áchase na base das demostracións de C. F. Gauss.

Doutra banda, en 1707 viu a luz a *Arithmetica universalis* de I. Newton. Nesta obra atópanse métodos diversos para determinar un límite superior das raíces reais dunha ecuación alxébrica, así como unha regra para determinar o límite inferior do número das raíces imaxinarias e o límite superior das raíces positivas e negativas. I. Newton tamén prestou interese ao importante problema da determinación das raíces dunha ecuación $f(x) = 0$. O seu procedemento, utilizado desde 1685 na *Algebra* de J. Wallis (1616–1703), foi lixeiramente modificado en 1690 por J. Raphson (1648–1715) e recollido ulteriormente por J. L. de Lagrange (1736–1813) e por J-B. J. Fourier (1768–1830).

J-B. J. Fourier dedicou boa parte da súa actividade científica ao estudo de ecuacións e os seus resultados apareceron publicados en 1831 nun traballo póstumo titulado *Analyse de equations determiées.* Neste tratado, entre outras cuestións de aritmética e álxebra, perfecciónase o método de Newton para aproximar raíces reais e introdúcese un método de separación de raíces reais aproximado, chamado de Budan-Fourier, pois o médico francés F. D. Budan (1761–1840) o enunciara sen demostración en 1807, época na cal J-B. J. Fourier xa o ensinaba aos seus alumnos na Politécnica.

Cando apareceu o libro de J-B. J. Fourier o problema da separación das raíces reais estaba resolto grazas ao teorema de J. Ch. F. Sturm (1803-1885), publicado en 1829 pero demostrado en 1835, onde se precisa o número exacto de raíces reais comprendidas entre dous límites dados. Este último manifesta que o seu descubrimento é resultado das investigacións de J-B. J. Fourier sobre o tema. No caso das raíces complexas, A. L. Cauchy anunciou en 1831 un teorema sobre o número de raíces reais ou complexas comprendidas no interior dun contorno pechado.

A pesar de todos estes avances o problema fundamental da álxebra continuaba sendo o exposto pola resolución de ecuacións de grao superior ao cuarto. O estudo deste problema foi abordado por P. Ruffini (1765–1822) e N. H. Abel (1802–1829) quen en 1824 estableceu que as ecuacións de grao maior ou igual que cinco non se poden resolver en xeral por radicais. Outro dos resultados fundamentais desta época débese a E. Galois (1811–1832), cuxas ideas non se coñeceron ata catorce anos despois da súa temperá morte cando J. Liouville (1809–1882) os publicou. O problema estudado por E. Galois foi o de determinar cando unha ecuación polinómica é resolúbel por radicais. Inspirado pola demostración de N. H. Abel da insolubilidade das ecuacións de grao maior ou igual que cinco, descubriu que unha ecuación alxébrica pódese resolver por radicais se, e só se, o seu grupo, é dicir, o grupo simétrico do conxunto das súas raíces, é resolúbel . A obra de E. Galois non é só importante por facer do concepto abstracto de grupo o centro da teoría de ecuacións, senón por conducir, a través dos descubrimentos doutros matemáticos dese século como J. W. R. Dedekind (1831–1916), L. Kronecker (1823–1891) e E. Kummer (1810–1893), ao que poderiamos chamar o enfoque aritmético da álxebra.

Estes estudos de ecuacións nunha variábel foron acompañados dos de ecuacións en varias variábeis, en particular sistemas de ecuacións lineares. Disto resultou a introdución dos conceptos de matriz e de determinante. As matrices deron lugar co paso do tempo a unha teoría independente, a álxebra de matrices, e o seu radio de aplicación foi estendido máis aló do campo dos sistemas de ecuacións lineares.

Tamén é importante salientar que a finais do século pasado, por influencia das ideas de F. Klein (1849–1925) e pola obra de C. D. T. Runge (1856–1927), creouse unha rama da matemática con métodos e caracteres propios, que tomou o nome de cálculo numérico (con este nome a Grande Encicopledia das Ciencias Matemáticas de Leipzig dedícalle no seu primeiro tomo de 1898–1904 un artigo de case cento cincuenta páxinas) e de matemática da aproximación, o nome sen dúbida máis adecuado, pois diso se trata. Partindo do suposto de que en moitas aplicacións prácticas da matemática, o obxectivo final é un resultado numérico e que este por esencia é aproximado na maior parte dos casos, ten sentido un corpo de doutrina e un campo propio de investigacións que tende a crear e estudar os métodos numéricos, gráficos e mecánicos que permiten obter eses resultados dunha forma aproximada.

A partir de mediados do século XIX, os estudos alxébricos movéronse gradualmente desde a teoría de ecuacións ao ámbito máis xeral das operacións alxébricas. Como vimos, os primeiros intentos dun estudo axiomático de operacións alxébricas datan da época de Euclides coa súa teoría das relacións; agora ben, non é ata finais do século XIX cando se conseguen avances significativos. Os progresos decisivos viñeron da man dunha análise intensiva do concepto de número e sobre todo da aparición de operacións aritméticas entre obxectos totalmente distintos aos números. Os primeiros exemplos podémolos atopar na composición de formas cuadráticas binarias de C. F. Gauss e na multiplicación de permutacións debida a P. Ruffini e A. L. Cauchy. Desde ese momento desenvólvense os estudos sobre os números complexos, aparece a álxebra da lóxica de G. Boole (1815–1864), a álxebra exterior de H. Grassmann (1809–1877), os cuaternións (cuaternios) de W. R. Hamilton (1805–1865) e a álxebra matricial de A. Cayley (1821–1895). Ademais é nesta época cando M. C. Jordan (1838–1922) publica o seu maior tratado sobre grupos de permutacións e E. H. Moore (1862–1932) e L. E. Dickson (1874–1954) a súa teoría de corpos finitos. Tamén nestes anos aparece o concepto de álxebra de Lie, introducido por M. S. Lie (1842–1899) e descuberto de forma independente por W. Killing (1847–1923). Doutra banda, en 1878, G. F. Frobenius (1849–1918) demostrou que os cuaternións son a única extensión asociativa posíbel dos números complexos con división. Este matemático tamén comprobou que os cuaternións son a única extensión non conmutativa posíbel dos números reais con división.

Estes estudos prepararon o camiño de transición da álxebra do século XIX cara a un estado moderno de desenvolvemento, o cal está caracterizado pola combinación de ideas, antes dispersas, nunha base axiomática común e pola considerábel extensión dos seus campos de aplicación. A moderna visión da álxebra, e da teoría xeral de operacións alxébricas, cristalizou na primeira metade do século vinte baixo a influencia de D. Hilbert (1862–1943), E. Steinitz (1871–1928), E. Artin (1893–1962), E. Noether (1882–1935), J. H. M. Wedderburn (1882–1948) e foi plenamente establecida coa aparición do libro de B. L. van der Waerden (1903–1996) *Modern algebra* en 1930.

O principal obxecto de interese para a álxebra moderna son os conxuntos e as operacións alxébricas definidas neles. Estes conxuntos dotados de operacións, coñecidos de forma xenérica co nome de álxebras, foron aparecendo co desenvolvemento das matemáticas e as súas aplicacións. Un dos tipos máis importantes de tales álxebras, e por tanto dos que foi estudado dunha maneira máis profunda, son os grupos. Un grupo é unha álxebra cunha operación binaria asociativa, un elemento unidade e onde todo elemento ten inverso. O concepto de grupo foi, historicamente, o primeiro exemplo de álxebra e de feito serviu, en moitos aspectos, como modelo para a construción da álxebra e das matemáticas durante o século XIX.

Os aneis e os corpos son outros tipos importantes de álxebras pero a diferenza dos grupos en vez dunha operación binaria débense considerar dúas. Estas operacións son comunmente chamadas adición e multiplicación. Un anel é un grupo abeliano para a operación de adición que cumpre unhas leis distributivas da multiplicación con respecto á adición. Orixinalmente só se estudaron os aneis con multiplicación asociativa e de feito o requirimento de que esta operación sexa asociativa en moitos casos pasou a ser parte da definición de anel. O estudo dos aneis non asociativos é tamén nos nosos días unha rama de gran importancia dentro da álxebra. Doutra banda, un corpo é un anel asociativo onde o conxunto dos elementos distintos de cero é grupo multiplicativo. Os corpos numéricos estaban implicitamente incluídos nos primeiros estudos sobre ecuacións alxébricas. Así mesmo, os aneis e corpos conmutativos son obxectos fundamentais dentro da álxebra conmutativa e da xeometría alxébrica.

Outro tipo importante de álxebra con dúas operacións binarias é un retículo. Exemplos típicos de retículos son as partes dun conxunto coas operacións de intersección e de unión ou o conxunto dos enteiros positivos tomando como operacións o mínimo común múltiplo e o máximo común divisor. Tamén é importante resaltar que unha álxebra de Boole é un tipo especial de retículo, de feito, é un retículo distributivo e complementario.

Os espazos vectoriais e os módulos poden ser tratados como álxebras cunha operación binaria (adición), con respecto á cal son grupos conmutativos, e cunha operación de multiplicación por escalares que pertencen a un corpo no caso dos espazos vectoriais e a un anel no caso dos módulos. Unha parte importante da álxebra, que trataremos con máis detemento na seguinte sección, ocúpase do estudo dos espazos lineares, dos módulos, das súas transformacións lineares, así como dos problemas relacionados con eles. Esta parte da álxebra recibe o nome de álxebra linear e, como veremos, ten un dos seus puntos de orixe na teoría de ecuacións lineares e ao longo do tempo deu lugar, entre outras ramas de interese, á teoría de matrices ou a xeneralizacións como a álxebra multilinear.

Os estudos iniciais sobre a teoría xeral de álxebras arbitrarias (esta teoría foi denominada álxebra universal e é unha das orixes da teoría de categorías) datan dos anos trinta e foron levados a cabo por G. Birkhoff (1884–1944). Ao mesmo tempo A. I. Maltsev (1909–1967) e A. Tarski (1901–1983) sentaron as bases da teoría de modelos, i.e., conxuntos cunhas relacións determinadas entre eles. A unión da teoría de álxebras universais e da teoría de modelos deu lugar a unha nova disciplina intermedia entre a álxebra e a lóxica matemática, chamada a teoría de sistemas alxébricos.

Un grande número de disciplinas intermedias entre a álxebra e outros campos das matemáticas foron creadas pola introdución de estruturas compatíbeis coas operacións alxébricas definidas nunha álxebra. Como exemplos de tales casos podemos citar a álxebra topolóxica (incluíndo a teoría de grupos topolóxicos e grupos de Lie), a teoría de aneis normados, a álxebra diferencial, etc. A álxebra homolóxica, que se orixina grazas ás influencias entre a álxebra e a topoloxía, aparece nos anos cincuenta como unha disciplina totalmente definida e é un dos xermolos da teoría de categorías.

O rol da álxebra na matemática moderna é extremadamente importante e como consecuencia existe unha tendencia bastante forte cara a unha alxebrización das matemáticas. Neste sentido paga a pena sinalar que un camiño habitual para estudar obxectos matemáticos pasa pola construción de sistemas alxébricos que os modelicen, como ocorre, por exemplo, en topoloxía ao asignar de forma canónica a cada espazo topolóxico unhas series infinitas de homoloxía de grupos que permiten avaliar dunha forma moi precisa as propiedades dos espazos. De feito, os maiores descubrimentos en topoloxía foron alcanzados utilizando a álxebra como ferramenta. O poderoso aparello de cálculo formal alxébrico foi empregado tamén en campos tan dispares das matemáticas como a teoría de números, a análise funcional, a teoría de ecuacións diferenciais, en xeometría (por exemplo, en teoría de invariantes, xeometría proxectiva, álxebra tensorial, etc...), en física (a teoría de representación de grupos finitos e álxebras de Hopf é fundamental na mecánica cuántica, os grupos discretos xogan un importante lugar na cristalografía), en cibernética (teoría de autómatas), en economía (desigualdades lineares) e noutras disciplinas.

1.2. Ecuacións lineares

A álxebra linear naceu para satisfacer as necesidades dos calculadores prácticos. De feito, a regra de tres e a regra de falsa posición enunciadas de forma máis ou menos confusa[1], desempeñan un importante papel en todos os manuais de aritmética práctica, desde o *Papiro Rhind* dos exipcios (foi comprado en 1858 polo anticuario escocés H. Rhind e desde 1865 atópase depositado no British Museum) ata os das nosas escolas primarias, pasando por Aryabhata, os árabes, Fibonacci e a grande cantidade de tratados de cálculo elaborados durante a Idade Media e o Renacemento.

Na matemática grega antiga, tal como aparece recollido nos *Elementos* de Euclides, podemos recoñecer dúas teorías abstractas de carácter linear: a das magnitudes e a dos enteiros. Nos babilonios, atopamos métodos moito máis próximos a nosa álxebra elemental; sabían resolver sistemas de ecuacións de primeiro grao.

O máis importante libro chinés da antigüidade, o chamado *Chiu Chang Suan Su*, datado a mediados do século segundo antes de Cristo, trata no seu capítulo oitavo da resolución de sistemas de ecuacións lineares e a súa xeneralización a sistemas con maior número de incógnitas segundo se expón na regra *fan-chen*[2], sendo no proceso de transformación da matriz do sistema onde os matemáticos chineses introduciron os números negativos.

Durante moito tempo, con todo, os progresos tiveron lugar no ámbito do cálculo alxébrico; en efecto, para reducir un sistema linear a unha ecuación do tipo $ax = b$ é suficiente coñecer as regras, xa enunciadas por Diofanto, para pasar os termos dun lado a outro e combinar os termos semellantes; e se se trata de varias incógnitas, basta ademais saber eliminalas sucesivamente ata que quede só unha. Tamén os tratados de álxebra ata o século XVIII danse por satisfeitos, no referente ao primeiro grao, unha vez que expuxeron estas regras.

G. Cramer
(Biblioteca de Xenebra/Wikimedia Commons)

O primeiro exemplo de cálculo explícito da solución dun sistema de ecuacións lineares con n incógnitas, onde os coeficientes son indeterminados, foi realizado por C. Maclaurin (1698–1746) en 1748 para $n = 2$ e para $n = 3$. Dous anos máis tarde, aparecen dous traballos que van xogar un papel fundamental no desenvolvemento do concepto de espazo vectorial e, por extensión, no da teoría da linearidade. O primeiro é o famoso tratado *Introduction à l'analyse des courbes algébriques* de G. Cramer (1704–1752), no cal se sentan as bases para a teoría de

[1] O procedemento de falsa posición seguido polos exipcios, para resolver ecuacións do tipo $x + ax = b$ ou do tipo $x+ax+bx = c$, consistía no seguinte: daban en primeiro lugar como solución un número ao azar, comparaban o resultado co resultado que figuraba no enunciado do problema, axustaban a solución que lles daba coa correcta mediante unha proporción e finalmente obtiñan a solución correcta.

[2] Esta regra é en esencia o método de eliminación gaussiana dos nosos días.

determinantes[3]. O segundo titúlase *Sur unha contradiction apparente dans a doctrine deas lignes courbes* e foi escrito por L. Euler. Neste último traballo o autor analiza o paradoxo de G. Cramer, xa identificado anteriormente por C. MacLaurin, para curvas alxébricas, sendo un dos primeiros en tratar a cuestión de dependencia de ecuacións lineares. Cando L. Euler estuda o caso $n = 4$ pódense recoñecer argumentos nos cales se utiliza unha intuición empírica da noción de rango.

A partir do traballo de G. Cramer a teoría de determinantes converteuse nun dos campos máis florecentes da actividade matemática, sendo despois de 1770 cando nos traballos de A. T. Vandermonde (1735–1796) e P. S. Laplace (1749–1827) aparece a idea de definir os determinantes de orde n por recorrencia sobre n (desenvolvemento por filas ou por columnas), así como as súas primeiras propiedades xerais: o ser funcións multilineares alternadas de filas e de columnas, e a invariabilidade do determinante con respecto á transposición Doutra banda, nunha memoria presentada na Academia de París en 1764, e dunha maneira máis detallada no seu tratado titulado *Théorie générale des équations algébriques*, de 1779, E. Bézout (1730–1783) deu un sistema de regras para resolver sistemas de ecuacións de n ecuacións lineares con n incógnitas.

E. Bézout
(Wikimedia Commons)

A pesar de todo este desenvolvemento, as cuestións referentes á indeterminacións e sistemas inconsistentes de ecuacións lineares foron descoidadas, xa que só é posíbel tratalas desde un achegamento baseado nas nocións de dependencia e rango. O concepto de rango, cuxo proceso de formación podemos situalo entre 1840 e 1879, é un invariante que determina o tamaño do conxunto de solucións dun sistema de ecuacións lineares (mínimo número de xeradores/máximo número de solucións independentes) e por dualidade o número de relacións de dependencia entre as ecuacións (mínimo número de ecuacións describindo o sistema de solucións/máximo número de ecuacións independentes).

Para crear o concepto de rango os matemáticos tiveron que vencer variados obstáculos e cambiar o seu punto de vista sobre certas nocións. As orixes do concepto de rango están moi ligados ao de determinante e máis concretamente ao de menor, de feito, na primeira metade do século XIX, era ben coñecido o método de solución de sistemas de ecuacións lineares consistente en illar a parte correspondente a un menor distinto de cero e de orde maximal para utilizar despois a *regra de Cramer* coas outras incógnitas como parámetros nos segundos membros.

En 1861 H. J. S. Smith (1826–1883) publicou un traballo titulado *On systems of linear indeterminates* no que demostraba que a orde dun menor maximal distinto de cero estaba relacionado co número maximal de solucións independentes. Isto non foi unha grande axuda para describir mellor o conxunto de solucións, pero a súa importancia radica en que a aproximación aos sistemas de ecuacións lineares faise por primeira vez desde un punto de vista teórico e non só co interese de atopar camiños para resolvelos.

[3]É A. L. Cauchy quen en 1815 introduce o nome de determinante ademais de probar completamente as súas propiedades fundamentais.

É cos traballos de G. L. Frobenius cando o estudo da independencia de ecuacións se libera da utilización dos determinantes. G. L. Frobenius tivo a idea orixinal de definir a independencia de ecuacións e n-uplas sen o uso de determinantes e introduciu a noción de sistema asociado. Dado un sistema de n ecuacións lineares e p incógnitas, cun menor maximal distinto de cero de orde r, demostrou que se poden atopar un número máximo de $p - r$ solucións independentes. Dado un destes conxuntos de solucións (unha base), construíu un sistema asociado co mesmo conxunto de solucións do sistema inicial. O seu método utiliza resultados técnicos da teoría de determinantes, pero moitos dos seus logros poden ser expresados sen o uso destes. Nunha publicación do ano 1879, *Uber homogene totale differentialgleichungen*, é onde finalmente introduce o termo rango e noutra de 1905, *Zur theorie der linearen gleichungen*, dá unha completa relación dos resultados referentes ao estudo teórico dos sistemas de ecuacións lineares.

G. L. Frobenius
(Universidade de Hamburgo/Wikimedia Commons)

Outros resultados interesantes e complementarios foron os establecidos por A. Capelli (1855–1910) entre 1886 e 1891. Este matemático italiano probou que un sistema de rango r é equivalente a un sistema triangular de r ecuacións e entón deduciu que o rango por filas dunha matriz é igual ao rango por columnas. Tamén se debe a matemáticos da escola italiana o resultado que garante que un sistema de ecuacións é consistente se o rango da matriz formada polos coeficientes é o mesmo que o rango da matriz ampliada pola columna dos segundos membros.

O coñecido como *Teorema de Rouché-Frobenius* enunciouse nesta época e é un dos grandes fitos da historia da álxebra linear xa que permite calcular o número de solucións dun sistema de ecuacións lineares en función do rango da matriz de coeficientes e do rango da matriz ampliada. E. Rouché (1832–1910) foi un matemático francés, coñecido, sobre todo, por este teorema e polo seu teorema sobre funcións holomorfas publicado en 1862. O *Teorema de Rouché-Frobenius* apareceu, primeiro, nun artigo de dúas páxinas en 1875 *Sur a discussion deas equations du premier degré* e despois, en 1890, foi publicada unha versión máis completa no *Journal de l'Ecole Polytechnique*. Posteriormente G. Fontené (1848-1923) e outros reclamaron a autoría da demostración. É por isto que o teorema se coñece como *Teorema de Rouché-Fontené* (isto último en Francia) ou como *Teorema de Kronecker-Capelli*.

Por tanto, podemos concluír que o estudo dos sistemas de ecuacións lineares e dos determinantes foi o contexto no cal apareceron por primeira vez conceptos como dependencia, rango, dualidade, que máis adiante xogarán un papel fundamental na moderna teoría dos espazos vectoriais.

Entre 1750 e os comezos do século XX, os determinantes foron omnipresentes en todos os problemas de tipo linear (excepto nalgúns traballos relacionados coa xeometría que trataremos máis adiante). Este feito ten unha influencia na natureza dos conceptos xa que, aínda que a noción de rango aparece asociada á de dimensión na moderna teoría axiomática dos espazos vectoriais, non debemos esquecer que adquiriu o seu sentido durante case dous séculos nos que tivo aos determinantes como soporte.

1.3. Vectores e xeometría

A relación entre a teoría de espazos vectoriais e a xeometría parece non xerar dúbidas para moitas persoas por varios feitos como, por exemplo, o uso da representación xeométrica para ilustrar ideas vectoriais. A vella representación xeométrica da suma de vectores, chamada paralelogramo de velocidades, non foi suficiente para a creación do concepto de segmento dirixido ou vector[4]. Na antigüidade a linearidade en xeometría referiuse á liña recta, que é unha das figuras básicas. Pero, seguindo con este tipo de argumentación, tamén é claro que o círculo é básico.

O método analítico introducido independentemente por R. Descartes no seu *Géométrie* (1637) e por P. de Fermat (1601–1665) en *Ad locos planos et solidos isagoge* (1643) organizou a xeometría dun xeito diferente seguindo novos criterios e ideas. As rectas corresponderanse coas ecuacións de primeiro grao e ocuparán o primeiro nivel. As ecuacións de segundo grao representarán as cónicas. O cambio de coordenadas, útil para a análise de invariantes de curvas, deu lugar a un marco para o estudo das transformacións lineares. Por tanto, co uso dos métodos analíticos en xeometría, a linearidade converteuse no punto de partida en moitos problemas relevantes dentro da matemática da época [5].

En 1679, nunha carta a C. Huygens (1629–1695), G. W. Leibniz critica o método analítico e tenta, de forma infrutuosa, crear unha análise xeométrica intrínseca. G. W. Leibniz consideraba que era difícil o estudo de propiedades xeométricas utilizando cálculos alxébricos, xa que segundo el os números non expresaban conceptos como situación, ángulo ou dirección. Esta corrente crítica tivo moitos simpatizantes, de feito, ata comezos do século XIX, a investigación orientada a conseguir unha análise xeométrica intrínseca foi un campo no que traballaron numerosos matemáticos. Agora ben, as críticas de G. W. Leibniz acharon unha resposta na representación xeométrica dos números complexos. Esta resposta pode definirse como indirecta, xa que o estudo da representación xeométrica dos números complexos estivo fundamentalmente motivada polo intento de lexitimar o seu uso ante aqueles que os consideraban inadecuados para a realidade matemática. O estudo da representación xeométrica dos números complexos foi iniciado por J. Wallis[6] en 1693 e continuada ao longo do tempo por diversos matemáticos como K. Wessel en 1799, J. R. Argand en 1806 e C. V. Mourey (1791–1830) en 1828. Con todo só cos traballos de C. F. Gauss, aproximadamente en 1831, e os de A. L. Cauchy en 1848, estas teorías alcanzan o seu grao de madurez e a aceptación pola comunidade matemática. Os números complexos deron lugar a un modelo para a análise da xeometría bidimensional. Nalgúns dos traballos citados anteriormente (especialmente os de Wessel) tentouse xeneralizar estas ideas á dimensión tres, pero sempre se topou coa dificultade do problema da multiplicación.

Durante o mesmo período de tempo, A. F. Möbius (1790–1868) e G. Bellavitis (1803–1880), deron a coñecer dous modelos de análise xeométrica, válidos tanto para a dimensión tres como para a dimensión dous, que sentaron as bases para a xeometría vectorial.

A. F. Möbius foi un dos primeiros en falar da noción de segmento orientado. En 1818, no primeiro capítulo do seu *Barycentrische calcul*, foron concibidos os principios do mesmo, designando o segmento unindo os puntos A e B por AB e establecendo a suma de segmentos colineares de forma que se chega á conclusión de que $AB = -BA$. Desde un punto de vista máis moderno, pódese afirmar que a teoría de A. F. Möbius é unha álxebra de puntos e non unha estrutura

[4]O nome de vector foi establecido por W. R. Hamilton en 1845, quen o utiliza para distinguir o que el chama a *parte vector* da *parte escalar* dun cuaternión.

[5]Por exemplo o traballo de G. Cramer *Introduction à l'analyse des courbes algébriques* citado na sección anterior.

[6]Este autor foi incapaz de ilustrar a multiplicación de números complexos.

alxébrica con todos os seus detalles. Con todo, as súas aplicacións son convincentes, numerosas e tiveron unha grande influencia en traballos de matemáticos posteriores. Baste citar por exemplo que serviu de inspiración para que K. G. C. Von Staudt (1798–1867) inventase as coordenadas proxectivas. Finalmente, en 1862 A. F. Möbius escribiu *Uber geometrische addition und multiplication*. Neste traballo, publicado en 1867, definiu a suma de segmentos non colineares, a multiplicación por un número de segmentos e dous tipos de produtos entre segmentos.

Grazas aos seus traballos, A. F. Möbius gozou do recoñecemento de matemáticos da súa época como C. F. Gauss, A. L. Cauchy, H. Grassmann, C. G. K. Jacobi (1804–1851) e P. E. G. Dirichlet (1805–1859). A súa contribución fundamental consistiu na creación dun método práctico e eficiente para resolver problemas xeométricos que só ten o defecto de estar baseado na percepción intuitiva do espazo, o que dalgunha forma invalida a súa xeneralización a un concepto máis amplo de espazo de vectores.

A. F. Möbius
(Adolf Neumann/Wikimedia Commons)

G. Bellavitis pode ser considerado o primeiro matemático en definir, en 1833, a suma de vectores no espazo no seu *Calcolo delle equipollence*. A súa representación foi orixinal por dous aspectos: os obxectos cos que traballa son puramente entidades xeométricas (non como os números complexos) e a primeira parte do seu cálculo pode ser aplicado en xeometría espacial. Con todo, G. Bellavitis, como moitos outros, fallou ao tratar de xeneralizar o produto de vectores no espazo. Esta xeneralización foi unha das principais contribucións de W. R. Hamilton.

W. R. Hamilton (Wikimedia Commons)

O matemático irlandés W. R. Hamilton levaba moito tempo interesado na xeneralización á dimensión tres da representación xeométrica dos números complexos, cando finalmente descubriu os cuaternións en 1843. Como os seus predecesores, e dunha forma natural, considerou ternas para as cales definiu unha suma e un produto que cumprían unhas propiedades equivalentes ás modernas de corpo. Despois de varios intentos, todos eles errados, W. R. Hamilton, ao darse conta, no famoso paseo pola ponte de Brongham, de que a multiplicación de dous vectores en dimensión dous estaba baseada nas súas lonxitudes e no ángulo que forman, cambiou o seu punto de partida e tivo en conta estes factores, o que lle permitiu construír o corpo dos cuaternións. En 1848 W. R. Hamilton conseguiu por fin publicar os seus descubrimentos. Os cuaternións $\mathbb{H} = \mathbb{R} \oplus \mathbb{H}_0$, onde $\mathbb{H}_0 = \mathbb{R}i + \mathbb{R}j + \mathbb{R}k$ é o subespazo dos cuaternións imaxinarios puros, son números alxébricos que teñen representación xeométrica no espazo. A multiplicación, que non é conmutativa, representa ao mesmo tempo o produto escalar e o produto vectorial mediante a fórmula

$$u \cdot v = -u \bullet v + u \times v,$$

onde $u, v \in \mathbb{H}_0$ e $\bullet$, $\times$ denotan respectivamente o produto escalar e vectorial en $\mathbb{R}^3$.

O cambio introducido na álxebra polos cuaternións tivo unha forte influencia na emerxente álxebra linear. De feito, a posibilidade de dar unha interpretación xeométrica de resultados alxébricos supuxo un camiño de enriquecemento para esta, xa que lle deu un fondo intuitivo e unha maior consistencia. O uso de termos xeométricos na teoría xeral de espazos vectoriais é unha proba das fondas relacións que existen entre a xeometría e a álxebra linear.

1.4. Álxebra e xeometría

Os intentos descritos nas seccións anteriores para crear unha análise xeométrica intrínseca poden ser vistos tanto como un desexo de liberar á xeometría dunha invasión externa por parte da aritmética como a vontade de introducir algúns aspectos da álxebra na xeometría. En calquera caso, é claro que despois do descubrimento do método analítico chegaron novos tempos e marcouse unha nova relación entre a álxebra e a xeometría. O uso do método analítico en xeometría deu lugar á creación de moitas das ferramentas da álxebra de matrices grazas ao estudo das substitucións lineares (transformacións lineares nunha linguaxe máis moderna). Moitos dos problemas de xeometría analítica conducen á utilización de cambios de coordenadas e, por tanto, ao transformacións lineares. É máis, as transformacións lineares aparecen noutros campos como a aritmética ou a solución de sistemas de ecuacións diferenciais.

En 1770 L. Euler, no seu traballo *Problema algebricum ob affectiones prorsus singulares memorabile*, trata cuestións que poden ser interpretadas en termos de transformacións lineares ortogonais. Entre 1773 e 1775 J. L. Lagrange tamén analiza o efecto de transformacións de tipo linear en formas cuadráticas de dúas variábeis a raíz do estudo das propiedades dos números que son suma de cadrados. Sobre 1798, C. F. Gauss na obra *Disquisitiones arithmeticae* trata a mesma cuestión en dúas e tres variábeis. É de resaltar que este último autor introduce unha notación similar a unha matriz para caracterizar unha transformación linear e establece a fórmula para a composición de dúas transformacións lineares.

De todos os xeitos o concepto de matriz non aparece claramente separado do concepto de determinante ata mediados do século XIX. En 1853, W. R. Hamilton introdúceo máis claramente nos seus *Lectures on quaternions* tendo a partir deste punto un rápido desenvolvemento. Desde esta perspectiva, a escola inglesa, cuxas máis representativas figuras do momento son A. Cayley e J. J. Sylvester (1814–1897), é un dos centros máis activos nesta materia. É de resaltar tamén que se poden atopar ideas similares en Alemaña. De feito, un dos primeiros intentos de dar unha lista sistemática das propiedades das matrices é debido a G. M. Eisenstein (1823–1852). Nos traballos deste autor resáltase a non conmutatividade do produto de matrices utilizándose tamén unha única letra para referirse a unha matriz e para describir as operacións entre elas.

O estudo das operacións entre matrices (cadradas ou rectangulares) ten como fundamental punto de partida a publicación en 1858 do famoso traballo de A. Cayley *Memoir on the theory of matrices*, no cal o autor reuniu, de forma detallada e organizada, todos os resultados descubertos nas dúas décadas precedentes.

A mediados o século XIX prodúcese outro dos feitos importantes para o desenvolvemento da teoría da linearidade. Tamén obra de A. Cayley, consistiu en dar os primeiros pasos para a superación do espazo de tres dimensións. No seu traballo *Sur quelques résultats de géométrie de position*, publicado en 1846, demostra como se poden obter resultados na xeometría tridimensional utilizando un espazo de máis de tres dimensións.

A. Cayley
(Herbert Beraud/Wikimedia Commons)

Doutra banda, o desenvolvemento das matemáticas no século XIX facilita e xustifica o uso de espazos de máis de tres dimensións. Neste sentido, dous tipos de eventos son fundamentais. En primeiro lugar as discusións para fundamentar a xeometría, xa que o descubrimento da xeometría non euclidiana e o desenvolvemento da xeometría proxectiva modificaron o campo tradicional da investigación xeométrica. En segundo lugar, o descubrimento dos cuaternións por W. R. Hamilton acabou co principio de formas equivalentes de G. Peacock (1791–1858), o cal establecía que a álxebra debe de ter só as leis aritméticas como fundamento. Isto abriu o camiño para realizar descubrimentos de novos tipos de operacións alxébricas, dando como resultado que nos nosos días o desenvolvemento desta rama das matemáticas é independente da aritmética.

Na segunda metade do século XIX a álxebra linear aínda non existía como un campo unificado, pero a xeometría de n dimensións estaba desenvolvida tomando como bases a xeometría analítica e a teoría de determinantes e matrices. De feito, o estudo das formas cuadráticas, iniciado por J. L. Lagrange e C. F. Gauss, converteuse nun dos eidos da actividade de numerosos investigadores.

1.5. A orixe da axiomatización da álxebra linear

En 1844, H. Grassmann publicou a primeira versión da súa *Lineale ausdehnunslhre* (literalmente teoría linear da extensión). Este traballo foi un dos máis brillantes e interesantes do seu tempo e aínda en tempos recentes foi a fonte de inspiración para teorías como a álxebra exterior de E. Cartan (1869–1951) ou o cálculo exterior de G. C. Rota (1932–1999).

A teoría de H. Grassmann contén as bases para unificar a teoría da linearidade e introduce con grande orixinalidade e xeneralidade conceptos tales como dependencia, base e dimensión.

Ademais, entre outros resultados, podemos atopar o teorema de cambio e unha elegante demostración dunha fórmula, coñecida como *Fórmula de Grassmann* na súa honra, que, utilizando unha linguaxe actual, liga a dimensión da suma e a intersección de subespazos do xeito seguinte:

$$\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

En moitos puntos a teoría de H. Grassmann adiantouse ao seu tempo dun xeito singular. Cada un dos seus resultados correspóndese a conceptos e teorías modernas, contribuíndo á creación de moitas delas. Para a teoría de espazos vectoriais as ideas deste matemático prusiano xogaron un papel fundamental no camiño da súa axiomatización, de feito a primeira definición próxima á de espazo vectorial debida a G. Peano (1858–1932) está inspirada na lectura dos traballos de H. Grassmann.

En 1888, G. Peano publicou unha versión condensada das súas lecturas sobre os traballos de H. Grassmann titulada *Calcolo geometrico*. Neste tratado introduce a definición de sistema linear, que é a primeira definición axiomática de espazo vectorial realmente próxima a que coñecemos nos nosos días. Aínda que os axiomas de G. Peano son moi similares ás propiedades fundamentais enunciadas por H. Grassmann, a principal contribución deste matemático italiano é enunciar as propiedades dunha operación para definir unha estrutura, mentres que H. Grassmann tenas que deducir da definición das operacións nas coordenadas.

H. Grassmann
(Wikimedia Commons)

Despois de G. Peano, xa no século XX, son de destacar as contribucións de C. C. H. Weyl (1885–1955) e J. W. R. Dedekind. A pesar de que nos traballos destes últimos xa atopamos un achegamento moi xeral ás estruturas lineares[7], é con E. Steinitz (1871–1928) cando se dá un paso decisivo no estudo e axiomatización das estruturas lineares.

En 1910, E. Steinitz publicou *Algebraische theorie der körper*, que marcou un fito relevante na historia da álxebra linear moderna ademais de ser unha referencia fundamental durante case un cuarto de século. Neste traballo, E. Steinitz introduce unha definición precisa de dependencia linear sobre un corpo K, e define unha extensión finita de orde n da seguinte maneira:

> *Sexa K un subcorpo de L. Dise que L é finito de orde n con respecto a K, se existen n elementos de L linearmente independentes sobre K, mentres que calquera outro sistema de máis de n elementos de L é linearmente dependente sobre K.*

Esta definición é idéntica á dada por G. Peano para o número de dimensións dun espazo linear. Aínda que na obra de E. Steinitz non atopamos referencias explícitas aos traballos de J. W. R. Dedekind, é importante resaltar que existe unha proximidade moi grande entre o pensamento de ambos científicos.

[7]No caso de J. W. R. Dedekind este prodúcese a partir do estudo das extensións de corpos.

E. Steinitz probou tamén resultados que traducidos ao contexto linear dan o teorema sobre a posibilidade de completar un sistema independente a unha base, a invariabilidade do número de elementos nas bases dun espazo linear e propiedades da dimensión dun subespazo dun espazo linear de dimensión finita.

1.6. Ecuacións diferenciais e análise funcional

É ben coñecido que as ecuacións diferenciais foron un importante punto de interese no mundo das matemáticas desde o século XVIII. De feito, o seu estudo deu lugar á análise funcional nos comezos do século XX. É máis, o estudo das ecuacións diferenciais lineares xogou un importante papel na teorización da noción de linearidade xa que aportou contextos máis amplos de dimensión non finita nos que esta apareceu dun xeito natural.

A mediados do século XVIII, J. L. D'Alembert, J. L. Lagrange e L. Euler xa coñecían que a solución xeral dun sistema linear de ecuacións diferenciais homoxéneas de orde n pódese expresar como combinación linear dun conxunto de n solucións fundamentais. Por suposto, isto non se expresaba nestes termos modernos e tampouco se tiña unha demostración rigorosa. De feito, a solución era considerada como series de potencias na contorna de cada punto, e as súas derivadas obtíñanse derivando cada termo das series. Entón, sumando termo a termo obtense unha ecuación linear recorrente que determina os coeficientes das series como funcións das n primeiras. Nese mesmo tempo, tamén era coñecido que a solución xeral dunha ecuación diferencial podíase obter como a suma dunha solución particular e a solución xeral da ecuación homoxénea. En 1777, J. L. Lagrange presentou o método de variación de constantes para atopar solucións particulares de ecuacións diferenciais, aínda que non é ata a chegada de A. L. Cauchy, cando todas estas nocións quedan claras e, ao mesmo tempo, obtéñense demostracións rigorosas.

O desenvolvemento do estudo das ecuacións diferenciais e as ecuacións diferenciais en derivadas parciais expuxo sofisticadas e complexas cuestións relacionadas coa linearidade. De feito, algúns dos conceptos esenciais da álxebra linear naceron de problemas deste tipo. Xa en 1770, J. L. Lagrange introduce un método chamado do operador adxunto no que se pode recoñecer a noción de dualidade. Así mesmo, o concepto de valor propio ou autovalor aparece co desenvolvemento dos sistemas de ecuacións diferenciais homoxéneas con coeficientes constantes.

En 1762, J. L. Lagrange, examinando os pequenos movementos dun sistema de n parámetros nas proximidades da posición de equilibrio, chega a un sistema de ecuacións diferenciais da forma:

$$x_j'' = \sum_{k=1}^{n} a_{jk} x_k \quad 1 \leq j \leq n$$

onde os a_{jk} son constantes. Entón seguindo o sistema dado por L. Euler para as ecuacións escalares de orde calquera, unha solución da forma $x_j(t) = y_j e^{\rho t}$, onde os y_j son constantes a determinar e ρ un número complexo, conduce ao sistema linear

$$\rho^2 y_j = \sum_{k=1}^{n} a_{jk} y_k \quad 1 \leq j \leq n$$

onde, como vemos, ρ^2 debe de ser, utilizando a terminoloxía moderna, un autovalor da matriz (a_{jk}). En 1774, atópase un sistema análogo, na teoría das desigualdades seculares dos planetas e P. S. Laplace, en 1784, estuda tamén este tipo de cuestións. Nestes problemas, as partes imaxinarias dos expoñentes ρ aparecen como as frecuencias dos fenómenos estudados. Aínda que a idea xeral de valor propio dun operador non aparecerá en análise ata a chegada da teoría de Sturm–Liouville, D. Bernoulli (1700–1782) xa a intúe cando chega á ecuación da corda vibrante, como un paso finito a partir do movemento dun número finito de masas repartidas pola corda en intervalos equidistantes.

Doutra banda, na década dos anos trinta do século pasado, matemáticos da escola soviética desenvolveron con éxito a teoría de funcións de matrices, sendo os primeiros en aplicala á investigación de sistemas de ecuacións diferenciais lineares. Os seus resultados atópanse entre as realizacións máis brillantes dos últimos cincuenta anos. Dentro da teoría de funcións de matrices, unha das propiedades importantes é a que garante que toda matriz cadrada satisfai a súa ecuación característica. Este resultado tivo a súa orixe no estudo dos cuaternións e coñécese como o *Teorema de Cayley-Hamilton*.

J-B. J. Fourier, para resolver ecuacións diferenciais empregando series de potencias, foi un dos pioneiros en estudar sistemas lineares infinitos na súa obra *Théorie analytique da chaleur* publicada en 1822. Un dos primeiros textos en presentar resultados consistentes sobre o tema foi publicado en 1886 e débese a H. Poincaré (1854–1912). Despois do traballo de H. Poincaré, moitos matemáticos como D. Hilbert, F. Riesz (1880–1956), J. Hadamard (1865–1963) e R. M. Fréchet (1878–1973) estudaron sistemas infinitos de ecuacións lineares. De todos os xeitos os métodos de aproximación a estes problemas eran excesivamente técnicos e difíciles de manipular. Por exemplo, en 1909 O. Toeplitz (1881–1940) probou certos teoremas sen o uso de determinantes, pero cun sofisticado método de eliminación, para obter sistemas triangulares equivalentes. Por este motivo, produciuse un gradual cambio de orientación ata chegar aos espazos vectoriais de funcións. Un dos pasos máis decisivos neste sentido foi dado por F. Riesz, nun traballo publicado en 1916. Neste traballo dáse a definición de norma dun subespazo vectorial pechado de funcións. En 1921, E. Helly (1884–1943) considera un espazo vectorial xeral normado marcando un novo avance no camiño cara a axiomatización da análise funcional. Finalmente, a etapa decisiva neste proceso é obra de S. Banach (1892–1945) e H. Hahn (1879–1934). A estrutura básica utilizada por S. Banach recibe aínda hoxe o nome de *Espazo de Banach* (isto é, un espazo vectorial normado completo). En 1932 S. Banach publicou a obra *Théorie des opérateurs linéaires* onde aparecen gran parte dos teoremas fundamentais da análise funcional e da álxebra linear de dimensión infinita. Este tratado tivo unha enorme popularidade e abriu unha nova era nestes dous campos das matemáticas.

1.7. Módulos e álxebras sobre un corpo

O termo módulo así como o de ideal, aínda que non co seu significado moderno, aparecen por primeira vez nun traballo de J. W. R. Dedekind publicado en 1871 sobre a teoría alxébrica de números. Coa palabra módulo J. W. R. Dedekind designou aos subconxuntos de números complexos pechados para a adición e a subtracción[8].

Nos traballos deste autor, atópanse tamén os primeiros intentos de relación entre as formas lineares e a teoría de números. De feito, J. W. R. Dedekind centrou a súa atención nos corpos Ω de enteiros alxébricos que teñen unha base, isto é, tales que consisten en todas as combinacións lineares de n elementos independentes con coeficientes racionais. Como vemos, desde o punto de vista moderno, Ω non é máis que un espazo vectorial de dimensión n sobre os racionais. A noción de módulo finitamente xerado tamén foi introducida por J. W. R. Dedekind para referirse ás combinacións lineares dun número finito de elementos alxébricos con coeficientes enteiros, ou o que é o mesmo, na terminoloxía dos nosos días, un Z-módulo finitamente xerado. É de resaltar que na obra de J. W. R. Dedekind aparece un concepto restrinxido de anel, chamado *Ordnung*, consistente nun módulo contendo ao elemento neutro e pechado para a multiplicación.

As contribucións de J. W. R. Dedekind non só se restrinxiron á teoría alxébrica de números. En 1882, publicou un artigo con H. Weber (1842–1913) onde se utilizaban conceptos similares aos anteriormente citados no contexto de funcións alxébricas de variábel complexa co obxecto de sentar unhas bases sólidas para a teoría de Riemann. Neste traballo, J. W. R. Dedekind

[8] J. W. R. Dedekind utilizou a notación $a \equiv b(Mod\ M)$ para designar que $a-b \in M$, inspirado pola empregada por C. F. Gauss na obra *Disquisitiones arithmeticae* para representar a congruencia en $\mathbb{Z}$, $a \equiv b(mod\ m)$. De aí é de onde provén o termo módulo.

e H. Weber estenden o temperán concepto de módulo. Un módulo de funcións é para eles un subconxunto pechado para a adición, a subtracción e para a multiplicación por calquera polinomio complexo (hoxe en día falariamos dun módulo sobre o anel dos polinomios $\mathbb{C}[z]$).

Unha década despois, no contexto da teoría de Galois, H. Weber unificou as nocións de corpo existentes, (por exemplo, corpo de números alxébricos, corpo de funcións alxébricas, corpo finito) introducindo o concepto abstracto e moderno de corpo. Este concepto, en contraste cos corpos utilizados por J. W. R. Dedekind, non tiña sempre característica cero. En 1910, os corpos abstractos de H. Weber foron analizados en profundidade por E. Steinitz, quen, como sabemos, no curso do seu traballo utilizou a idea de independencia linear. De feito, E. Steinitz comezou chamando a un elemento x transcendente sobre un corpo $\mathbb{F}$ si 1, x, $x^2, \ldots$ eran linearmente independentes sobre $\mathbb{F}$; en caso contrario, o elemento era chamado alxébrico sobre $\mathbb{F}$.

Mentres tanto, a noción de módulo de Dedekind foi utilizada por D. Hilbert en 1897 nun traballo sobre a teoría de números alxébricos para a *Deutsche Mathematiker-Verinigung*. Neste traballo, utilizou o termo *Zahlring* ou *Ring* para denotar os *Ordnung* de J. W. R. Dedekind. Nestas estruturas, D. Hilbert puxo a énfase no paralelismo creado pola clausura: un corpo de números alxébricos $\mathbb{F}$ era pechado para a adición, subtracción, multiplicación e división; un anel era un conxunto de enteiros alxébricos de $\mathbb{F}$ pechado para a adición, subtracción e multiplicación e un módulo era un conxunto pechado para a adición e subtracción. Un ideal era algo intermedio, pechado para a adición, subtracción e para a multiplicación por escalares de $\mathbb{F}$.

En 1914, A. A. H. Fraenkel (1891–1965), estimulado polas ideas de D. Hilbert, así como polos sistemas de números hipercomplexos e os aneis de matrices, definiu un concepto abstracto de anel moi próximo ao que coñecemos hoxe en día, de feito só difire deste en que ten condicións especiais que concirnen aos divisores de cero. A. A. H. Fraenkel, demostrou como os aneis non só xogan un papel importante na teoría de números senón que tamén o teñen noutros campos das matemáticas.

Os modernos conceptos de anel, ideal e módulo sobre un anel aparecen por primeira vez nun traballo publicado por E. Noether en 1921 e titulado *Idealtheorie in ring-bereichen*. A definición orixinal de módulo dada por esta autora consideraba un par (Σ, T), onde Σ era un anel e T era o que nós coñecemos hoxe en día por un Σ-módulo. Un módulo en (Σ, T) quedaba definido como un subconxunto de T pechado para a subtracción así como para a multiplicación pola esquerda para escalares de Σ. Máis adiante, nos traballos publicados despois dos anos vinte, E. Noether puxo máis a énfase na teoría de ideais que na de módulos, desenvolvendo a teoría de aneis non conmutativos e ademais reformulou a definición de álxebra sobre un corpo incorporando os conceptos de anel e de módulo.

A noción de álxebra sobre un corpo ten as súas orixes nas álxebras lineares asociativas ou sistemas de números hipercomplexos utilizados por W. R. Hamilton e Ch. S. Peirce (1839–1914). Esta noción contemporánea difire das máis antigas en varios aspectos. Durante o século XIX, tales álxebras eran consideradas sobre os reais ou os complexos, no canto de sobre un corpo arbitrario, e sempre se traballaba en espazos de dimensión finita. A noción moderna foi establecida en termos de espazos vectoriais e aneis. Máis especificamente, A é unha álxebra sobre un corpo $\mathbb{F}$ se A é un espazo vectorial sobre $\mathbb{F}$ e un anel con unidade onde a adición de vectores e a multiplicación por escalares coincide coas correspondentes operacións do anel.

L. E. Dickson (1874–1939), en 1903, foi o primeiro en propoñer un concepto de álxebra sobre un corpo arbitrario, aínda que sempre dentro do caso de dimensión finita, que modificado sería estendido á dimensión infinita por J. H. M. Wedderburn en 1924. L. E. Dickson deu ao longo da súa vida varias definicións de álxebras sobre un corpo, aínda que a máis próxima á que coñecemos actualmente aparece nun libro que publicou en 1923. Unha álxebra A sobre un corpo $\mathbb{F}$ era definida como un sistema con dúas operacións sobre A, adición e multiplicación, e unha terceira operación de $\mathbb{F}$ sobre A; isto é, a multiplicación por escalares. A multiplicación, a adición e a multiplicación por escalares supoñíanse asociativas e as dúas últimas conmutativas.

A multiplicación por escalares debía ser ademais distributiva con respecto á adición de A e $\mathbb{F}$. O requisito final é que A debía de ser finita. É de resaltar que nesta definición, aínda que non existe mención algunha dos conceptos de anel, módulo e vector, a multiplicación de matrices xoga un importante papel.

E. Noether
(Wikimedia Commons)

A definición actual de álxebra sobre un corpo débese a E. Noether e aparece nun traballo desta autora publicado en 1929. Neste traballo, establécense as relacións entre a teoría de grupos e a teoría de módulos sobre un anel, sendo tamén importante por mor de que nel faise mención a unhas conexións atopadas por B. L. Van der Waerden (1903–1996) entre as transformacións lineares de módulos e as matrices. De feito para B. L. Van der Waerden unha transformación linear é un homomorfismo de dous módulos de formas lineares, unha matriz é unha expresión (a representación) dese homomorfismo unha vez feita unha elección de bases. Esta é a conexión moderna esencial entre as nocións de transformación linear, matriz e módulo (ou espazo vectorial). Dous anos despois B. L. Van der Waerden publicaría un libro onde expuxo todas estas relacións a un público máis amplo sendo este o comezo do que coñecemos como álxebra moderna.

1.8. A álxebra moderna

Como vimos J. W. R. Dedekind deu as primeiras definicións de tipo axiomático de anel, ideal, corpo e módulo. H. Weber, na súa obra *Lehrbuch der algebra*, publicada en 1894, deu a primeira definición axiomática de grupo. No primeiros trinta anos do século XX, atopamos varios matemáticos, especialmente en Alemaña, que coas súas investigacións sobre teoría de grupos e extensións de corpos sentarían as bases do que hoxe coñecemos como álxebra linear.

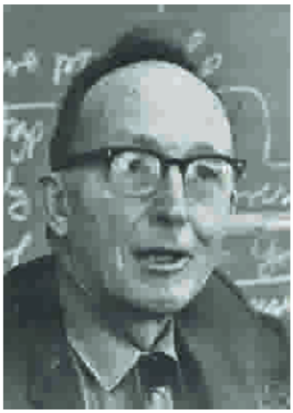
En 1930, B. L. Van der Waerden publicou o primeiro volume da primeira edición da súa obra *Moderne algebra*; o volume segundo foi publicado ao ano seguinte. Nesta primeira edición, a álxebra linear aparece centrada fundamentalmente sobre a estrutura de módulo, aínda que os primeiros conceptos e resultados tales como combinacións lineares, dependencia, bases e dimensión son, de feito, establecidos nun capítulo previo sobre extensións de corpos. Este autor define os módulos como grupos abelianos aditivos cun dominio de operadores, que forman un anel, satisfacendo determinados axiomas. Para un anel R con unidade, os módulos, cumprindo que $1x = x$, para todo x, son chamados unitarios. Os módulos unitarios finitamente xerados teñen unha importancia especial, denominándose como módulos de formas lineares sobre o anel R. Demóstrase tamén que en tales módulos os elementos poden ser tomados como n-uplas, que serán chamados vectores.

Aínda que o termo álxebra linear xa fora utilizado por H. Weyl, é no libro xa mencionado de B. L. Van der Waerden onde aparece asociado por primeira vez coa teoría de módulos sobre un anel e os seus homomorfismos, isto é, as transformacións lineares entre eles. Neste contexto, as matrices aparecen como representacións de tales homomorfismos cando se fixan unhas bases determinadas.

B. L. Van der Waerden
(ETH Zürich/Wikimedia Commons)

En edicións posteriores, o lugar ocupado pola álxebra linear e os espazos vectoriais pasa a ser central. O estudo dos sistemas de ecuacións lineares é presentado como unha aplicación da teoría de espazos vectoriais e o papel dos determinantes queda considerabelmente reducido. Rapidamente, esta forma de entender a álxebra linear cambiou a forma de achegamento a moitos problemas xa que a teoría axiomática permitiu unificar os campos de dimensión finita e infinita.

G. D. Birkhoff
(Wikimedia Commons)



S. Mac Lane
(Konrad Jacobs/Wikimedia Commons)

En 1941, G. D. Birkhoff e S. Mac Lane (1909–2005) publicaron *Survey of modern algebra*, e no ano 1942, aparece o libro de P. R. Halmos (1916–2006) titulado *Finite-dimensional spaces*. Así mesmo, en 1947, en Francia, o colectivo N. Bourbaki publicou o segundo capítulo do libro II dos seus *Éléments de mathématique* baixo o título de *Algèbre linéaire*. Estes tres libros tiveron unha influencia longa e notábel na teoría axiomática dos espazos vectoriais así como no seu uso e no seu ensino. É a partir de aquí, cando podemos afirmar que a álxebra linear alcanza a súa maioría de idade.

CAPíTULO 2

Preliminares. Números reais e complexos

2.1. Propiedades básicas dos números reais

O obxectivo deste capítulo é lembrar as propiedades básicas dos conxuntos de números que utilizaremos ao longo desta obra. Por tanto, centraremos a nosa atención no conxunto dos números reais, que denotaremos por $\mathbb{R}$, e no conxunto dos números complexos que denotaremos por $\mathbb{C}$. Tamén manexaremos outros conxuntos numéricos como os números naturais

$$\mathbb{N} = \{1, 2, 3, \dots\},$$

os números enteiros

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$$

e os números racionais

$$\mathbb{Q} = \{ \frac{p}{q} \ \mid p, q \in \mathbb{Z},\ q \neq 0\}.$$

Para os anteriores conxuntos de números temos os seguintes contidos.

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

Os contidos previos son estritos xa que, por exemplo, o número -3 é enteiro e non é natural. O número $\frac{1}{2}$ é racional pero non é enteiro, o número $\sqrt{2}$ é un número real que non se pode representar como cociente de dous números enteiros e, finalmente, a unidade imaxinaria i é un número complexo que non é real xa que cumpre que

$$i^2 = -1.$$

Todo número real admite unha representación decimal. Se o número é racional, o decimal correspondente é periódico; por exemplo,

$$\frac{1}{2} = 0{,}5\overline{0}, \quad \frac{2}{3} = 0, \overline{6}, \quad \frac{157}{495} = 0{,}3\overline{17}, \quad \frac{9}{7} = 1, \overline{285714}$$

onde a liña sobre os números indica que a secuencia de díxitos se repite sucesivamente. Doutra banda, se a representación decimal non é periódica o número chámase irracional e o conxunto destes números denotarase por $\mathbb{I}$. Por exemplo

$$\sqrt{2} = 1, 414213562373095\dots,$$

$$\pi = 3, 141592653589793\dots,$$

son números irracionais. Daquela, os números reais son a unión dos números racionais e os irracionais:

$$\mathbb{R} = \mathbb{Q} \cup \mathbb{I}.$$

Para cada par de números reais a e b existe un único número real chamado suma de a e b, denotado por $a + b$, e un único número real chamado produto de a e b, denotado por $a \cdot b$, tales que se cumpren as seguintes propiedades:

1) $a + b = b + a$ e $a \cdot b = b \cdot a$, $\forall\, a, b \in \mathbb{R}$.
2) $a + (b + c) = (a + b) + c$ e $a \cdot (b \cdot c) = (a \cdot b) \cdot c$, $\forall\, a, b, c \in \mathbb{R}$.
3) $a \cdot (b + c) = a \cdot b + a \cdot c$, $\forall\, a, b, c \in \mathbb{R}$.
4) Existen dous números reais 0 e 1 tales que $0 + a = a$ e $1 \cdot a = a$, $\ \forall\, a \in \mathbb{R}$.

5) Para todo número real a existe un número real $-a$, chamado o oposto de a, tal que $a + (-a) = 0$
6) Para todo número real non nulo a existe un número real a^{-1}, chamado o inverso de a, tal que $a \cdot a^{-1} = 1$.

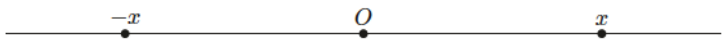
A primeira das propiedades indica que a suma e o produto son operacións conmutativas, a segunda propiedade afirma que suma e produto son asociativas e a terceira expresa o carácter distributivo do produto respecto da suma. Na cuarta propiedade temos que o número 0 é o elemento neutro da suma e que o número 1 o é do produto. Finalmente, a quinta e a sexta propiedades indican que todo número ten oposto para a suma e que todo número real non nulo ten inverso para o produto. Este inverso tamén denótase por

$$a^{-1} = \frac{1}{a}.$$

No que segue, por comodidade, representaremos o produto de dous números reais a e b como ab. Das propiedades anteriores dedúcese de forma fácil que 0 e 1 son únicos e que para todo a real non nulo $-a$ e a^{-1} tamén son únicos. Obviamente $-0 = 0$.

Un conxunto con dúas operacións satisfacendo as propiedades anteriores chámase corpo. Por tanto, os números reais son un corpo. Os números racionais tamén o son se utilizamos como operacións a suma e o produto de números racionais. Os números enteiros non o son xa que, por exemplo, o número 2 non ten inverso para o produto de números enteiros.

Os números reais pódense representar mediante puntos nunha recta. Nela elíxese un punto arbitrario O, que se chama orixe, e corresponde ao número real 0. Cada número positivo x represéntase mediante o punto nesa recta que está a unha distancia de x unidades cara á dereita da orixe, e cada número negativo $-x$, represéntase mediante o punto que está x unidades á esquerda da orixe. Así, cada número real represéntase como un punto da recta, e cada punto da recta corresponde a un número real.



Os números reais están ordenados. Dise que a é menor ou igual que b e escríbese $a \leq b$ se $b - a$ é un número non negativo. Isto significa, xeometricamente, que se a é distinto de b, a está á esquerda de b na recta numérica anterior. Cando $a \leq b$ e a e b son distintos diremos que a é menor que b e representarémolo como $a < b$. A relación $\leq$ é de orde xa que se cumpren as propiedades reflexiva

$$a \leq a \; \forall a \in \mathbb{R},$$

antisimétrica

$$a \leq b \text{ e } b \leq a \;\Rightarrow\; a = b$$

e transitiva

$$a \leq b \text{ e } b \leq c \;\Rightarrow\; a \leq c.$$

A orde é total xa que para cada par de números reais a e b tense que $a \leq b$ ou $b \leq a$. Doutra banda, cúmprense as propiedades:

1) Para todo número real c, se $a \leq b$, tense que

$$a + c \leq b + c.$$

2) Para todo número real $c > 0$, se $a \leq b$, tense que

$$ac \leq bc.$$

Por tanto $\mathbb{R}$ é un corpo ordenado.

2.2. Os números complexos

Definición 2.2.1. Un número complexo é unha expresión da forma $z = a + bi$ onde a, b son números reais e i é un elemento chamado unidade imaxinaria.

O número real a chamarase parte real de z, representado como $\mathrm{Re}(z)$, e o número real b denominarase parte imaxinaria de z, representado como $\mathrm{Im}(z)$. Como xa vimos na sección anterior, o conxunto dos números complexos denotarase como $\mathbb{C}$ e obviamente dous números complexos son iguais se teñen as súas partes reais e imaxinarias respectivas coincidentes. Da propia definición de número complexo dedúcese que os números reais son exactamente os números complexos con parte imaxinaria nula. Cando un número complexo teña parte real nula chamarase imaxinario puro.

Se $z = a + bi$ é un número complexo defínese o seu conxugado como

$$\overline{z} = a - bi.$$

Doutra banda, defínese o módulo de z como o número real representado por $|z|$ onde

$$|z| = +\sqrt{a^2 + b^2}.$$

Do mesmo xeito que no caso dos números reais, no conxunto dos números complexos podemos definir unha operación de suma e outra de produto da seguinte forma: dados dous números complexos $z = a + bi$, $z' = a' + b'i$ defínese a suma de z e z' como

$$z + z' = (a + a') + (b + b')i$$

e o seu produto está dado por

$$z \cdot z' = (aa' - bb') + (ab' + a'b)i.$$

Tendo en conta a definición do produto de números complexos obtemos

$$i^2 = (0 + 1i) \cdot (0 + 1i) = -1,$$

pódese demostrar facilmente a identidade

$$z \cdot \overline{z} = a^2 + b^2 = |z|^2$$

e ademais cúmprense as seguintes propiedades:

1) $z + z' = z' + z$ e $z \cdot z' = z' \cdot z$, $\forall\, z, z' \in \mathbb{C}$.
2) $z + (z' + z'') = (z + z') + z''$ e $z \cdot (z' \cdot z'') = (z \cdot z') \cdot z''$, $\forall\, z, z', z'' \in \mathbb{C}$.
3) $z \cdot (z' + z'') = z \cdot z' + z \cdot z''$, $\forall\, z, z', z'' \in \mathbb{C}$.
4) Existen dous números complexos $0 = 0 + 0i$ e $1 = 1 + 0i$ tales que $0 + z = z$ e $1 \cdot z = z$ $\forall\, z \in \mathbb{C}$.
5) Para todo número complexo $z = a + bi$ existe un número complexo $-z = -a - bi$, chamado oposto de z, tal que $z + (-z) = 0$.
6) Para todo número complexo non nulo $z = a + bi$ existe un número complexo z^{-1} (ou $\dfrac{1}{z}$) tal que $z \cdot z^{-1} = 1$. Este complexo chámase inverso de z e está dado por
$$z^{-1} = \frac{\overline{z}}{z.\overline{z}} = \frac{\overline{z}}{|z|^2} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i.$$

Do mesmo xeito que pasaba cos números reais, a primeira propiedade indica que a suma e o produto son operacións conmutativas, a segunda propiedade afirma que suma e produto son asociativas e a terceira expresa o carácter distributivo do produto respecto da suma. Na cuarta propiedade temos que o número 0 é o elemento neutro da suma e que o número 1 o é do produto. Finalmente, a quinta e a sexta propiedades indican que todo número complexo ten oposto para a suma e que todo número complexo z non nulo ten inverso para o produto.

Das propiedades anteriores dedúcese de forma sinxela que 0 e 1 son únicos e que para todo z complexo non nulo $-z$ e z^{-1} tamén o son. Obviamente, do mesmo xeito que no caso real,

$-0 = 0$. Ademais $\mathbb{R} \subseteq \mathbb{C}$ xa que cada número real a pódese pensar como o número complexo $z = a+0i$ e, dado que se satisfán as propiedades da lista anterior, temos que $\mathbb{C}$ tamén é un corpo. A gran diferenza con $\mathbb{R}$ é que $\mathbb{C}$ non é ordenado. No que segue, por comodidade, denotaremos o produto de dous números complexos z e z' como zz'.

Exemplos 2.2.2. Por exemplo, dados os números complexos $1-2i$, $2+3i$ y $-1+i$ tense que

$$(1-2i)-(2+3i)(-1+i) = (1-2i)-(-5-i) = 6-i.$$

Calculemos, en segundo lugar, o valor de $\left[(3+2i)^3+(1-i)\right]^2$. Como

$$(3+2i)^3 = 3^3+3\cdot(3)^2\cdot(2i)+3\cdot 3\cdot(2i)^2+(2i)^3 = 27+54i+36i^2+8i^3 =$$

$$= 27-36+54i-8i = -9+46i$$

obtemos que

$$\left[(3+2i)^3+(1-i)\right]^2 = [(-9+46i)+(1-i)]^2 = [-8+45i]^2 =$$

$$(-8)^2-720i+(45i)^2 = -1961-720i.$$

A existencia de inversos para o produto tamén permite calcular cocientes da seguinte forma

$$\frac{z}{z'} = zz'^{-1} = \frac{z\overline{z'}}{|z'|^2}$$

sendo z un complexo cualquera e z' un complexo non nulo.

Exemplo 2.2.3. Por exemplo, dados os complexos $z = 1-2i$, $z' = 2+3i$ o cociente $\dfrac{z}{z'}$ é

$$\frac{1-2i}{2+3i} = \frac{(1-2i)(2-3i)}{13} = \frac{-4-7i}{13} = -\frac{4}{13}-\frac{7}{13}i.$$

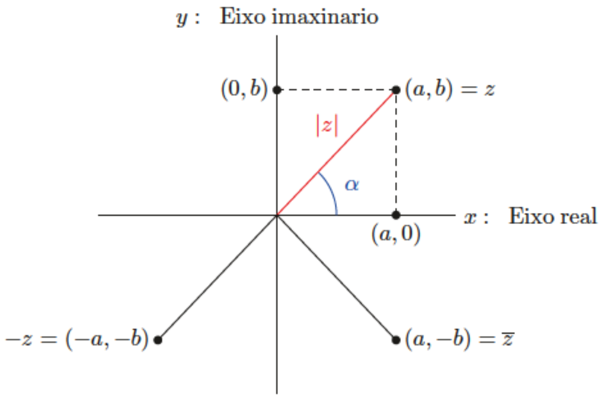
Outras propiedades interesantes que teñen unha fácil demostración son as seguintes:

1) $z+\overline{z} = 2\mathrm{Re}(z)$ e $z-\overline{z} = 2\mathrm{Im}(z)i$, $\forall\, z \in \mathbb{C}$.
2) $\overline{z+z'} = \overline{z}+\overline{z'}$, $\forall\, z, z' \in \mathbb{C}$.
3) $\overline{z-z'} = \overline{z}-\overline{z'}$, $\forall\, z, z' \in \mathbb{C}$.
4) $\overline{zz'} = \overline{z}\,\overline{z'}$, $\forall\, z, z' \in \mathbb{C}$.
5) $\overline{\left(\dfrac{z}{z'}\right)} = \dfrac{\overline{z}}{\overline{z'}}$, $\forall\, z, z' \in \mathbb{C}$ sendo z' non nulo.

Todo número complexo $z = a+bi$ pódese identificar cun único punto do plano real $\mathbb{R}^2$ dado por (a,b). Reciprocamente, calquera punto (x,y) do plano real $\mathbb{R}^2$ pódese identificar cun único número complexo dado por $z = x+iy$. Desta forma o eixo OX chámase eixo real e o eixo OY eixo imaxinario. O número complexo i correspóndese co punto $(0,1)$ de $\mathbb{R}^2$ e o número complexo 1 co punto $(1,0)$ de $\mathbb{R}^2$. Tendo en conta todo isto, de agora en diante, poderemos pensar os números complexos como puntos de $\mathbb{R}^2$. Máis concretamente, se identificamos $z = a+bi$ con (a,b) tense que:

$$z = (a,b) = a(1,0)+b(0,1) = a+bi.$$

O oposto dun complexo $z = a+bi$ é a reflexión de z con respecto á orixe e o conxugado é a reflexión con respecto ao eixo real. A seguinte figura ilustra estas afirmacións e as do parágrafo previo.



Na figura anterior pódese observar que o módulo de z correspóndese coa lonxitude do segmento de extremos $(0, 0)$, (a, b). O ángulo α que forma a parte positiva do eixo x co segmento anterior chámase argumento de z. Como

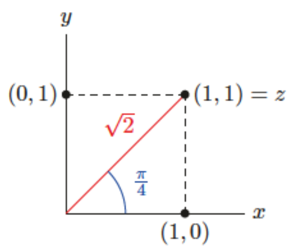
$$a = |z|\cos(\alpha), \quad b = |z|\text{sen}(\alpha)$$

obtense que z pódese representar na seguinte forma, chamada forma polar ou trigonométrica,

$$z = |z|(\cos(\alpha) + i\text{sen}(\alpha)).$$

Exemplo 2.2.4. Consideremos o número complexo $z = 1 + i$. Tense que $|z| = \sqrt{2}$ e que $\tan(\alpha) = \frac{1}{1} = 1$. Daquela $\alpha = \frac{\pi}{4}$ e

$$z = \sqrt{2}(\cos(\frac{\pi}{4}) + i\text{sen}(\frac{\pi}{4})).$$



Por tanto todo número complexo está completamente determinado polo seu módulo e o seu argumento. A representación polar dos números complexos é especialmente útil para calcular produtos e cocientes de complexos xa que se

$$z = |z|(\cos(\alpha) + i\text{sen}(\alpha)), \quad z' = |z'|(\cos(\beta) + i\text{sen}(\beta))$$

podemos demostrar que se cumpren as seguintes igualdades sen máis que usar as identidades trigonométricas que expresan o coseno e o seno da suma ou a diferenza de dous ángulos:

$$zz' = |z||z'|(\cos(\alpha + \beta) + i\text{sen}(\alpha + \beta)),$$
$$\frac{z}{z'} = \frac{|z|}{|z'|}(\cos(\alpha - \beta) + i\text{sen}(\alpha - \beta)).$$

Como consecuencia do anterior, obtense que a potencia n-ésima de z pódese calcular como

$$z^n = |z|^n(\cos(n\alpha) + i\text{sen}(n\alpha)).$$

Se temos en conta a fórmula de Euler

$$e^{i\alpha} = \cos(\alpha) + i\mathrm{sen}(\alpha)$$

un número complexo $z = |z|(\cos(\alpha) + i\mathrm{sen}(\alpha))$ pódese expresar como

$$z = |z|e^{i\alpha}.$$

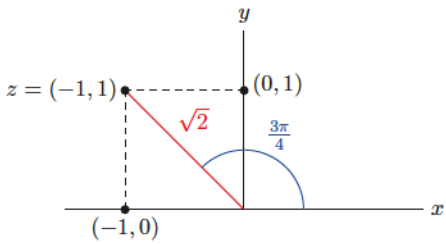
Esta forma de representación coñécese como forma de Euler ou forma polar compacta. Logo as fórmulas anteriores quedan reducidas a

$$zz' = |z||z'|e^{i(\alpha+\beta)}, \quad \frac{z}{z'} = \frac{|z|}{|z'|}e^{i(\alpha-\beta)}, \quad z^n = |z^n|e^{in\alpha}.$$

Exemplo 2.2.5. Consideremos o número complexo $z = -1+i$. Neste exemplo calcularemos z^{10}. Temos que $|z| = \sqrt{2}$ e que $\tan(\alpha) = \frac{-1}{1} = -1$. Entón $\alpha = \frac{3\pi}{4}$ e

$$z = \sqrt{2}(\cos(\frac{3\pi}{4}) + i\mathrm{sen}(\frac{3\pi}{4})) = \sqrt{2}e^{i\frac{3\pi}{4}}.$$

Graficamente



Entón,

$$z^{10} = (\sqrt{2})^{10}e^{i\frac{15\pi}{2}} = 2^5e^{i\frac{3\pi}{2}} = -32i.$$

Para finalizar este capítulo veremos como se poden calcular raíces n-ésimas de números complexos. Supoñamos que temos un número complexo $w = |w|e^{i\beta}$. Dado $n \in \mathbb{N}$, trátase de atopar todas as solucións da ecuación

$$z^n = w.$$

Se $z = |z|e^{i\alpha}$, a anterior igualdade é equivalente a

$$|z|^n(\cos(n\alpha) + i\mathrm{sen}(n\alpha)) = |w|(\cos(\beta) + i\mathrm{sen}(\beta))$$

e, por tanto,

$$|z| = |w|^{\frac{1}{n}}$$

e

$$\cos(n\alpha) = \cos(\beta), \quad \mathrm{sen}(n\alpha) = \mathrm{sen}(\beta).$$

Xa que as funcións seno e coseno teñen período 2π, as dúas últimas igualdades implican que $n\alpha$ e β difiren nun múltiplo enteiro de 2π. Entón,

$$n\alpha = \beta + 2k\pi$$

ou, equivalentemente,

$$\alpha = \frac{\beta + 2k\pi}{n}.$$

Deste xeito obtense que as raíces n-ésimas de w son

$$z_k = |w|^{\frac{1}{n}}(\cos(\frac{\beta + 2k\pi}{n}) + i\mathrm{sen}(\frac{\beta + 2k\pi}{n})) = |w|^{\frac{1}{n}}e^{i(\frac{\beta+2k\pi}{n})}$$

con $k \in \{0, 1, \ldots, n-1\}$.

En particular, se queremos calcular as raíces n-ésimas da unidade ou, o que é o mesmo, resolver a ecuación $z^n = 1$, teremos que as solucións son

$$z_k = \cos(\frac{2k\pi}{n}) + i\mathrm{sen}(\frac{2k\pi}{n}) = e^{i\frac{2k\pi}{n}}$$

con $k \in \{0, 1, \ldots, n-1\}$.

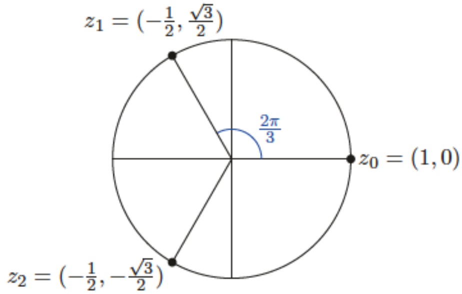
Exemplos 2.2.6. En primeiro lugar calcularemos as raíces cúbicas e cuartas da unidade.

As raíces cúbicas da unidade serán

$$z_k = \mathrm{e}^{i\frac{2k\pi}{3}}, \quad k = 0, 1, 2.$$

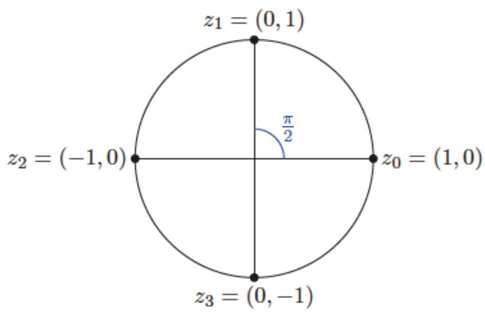
Isto é,

$$\begin{array}{lllllll} z_0 & = & \mathrm{e}^{i\frac{0}{3}} & = & \cos(0) + i\mathrm{sen}(0) & = & 1 \\ z_1 & = & \mathrm{e}^{i\frac{2\pi}{3}} & = & \cos(2\pi/3) + i\mathrm{sen}(2\pi/3) & = & -\frac{1}{2} + \frac{\sqrt{3}}{2}i \\ z_2 & = & \mathrm{e}^{i\frac{4\pi}{3}} & = & \cos(4\pi/3) + i\mathrm{sen}(4\pi/3) & = & -\frac{1}{2} - \frac{\sqrt{3}}{2}i \end{array}$$



As raíces cuartas da unidade serán $z_k = \mathrm{e}^{i\frac{2k\pi}{4}}$, $k = 0, 1, 2, 3$. Isto é,

$$\begin{array}{lllllll} z_0 & = & \mathrm{e}^{i\frac{0}{4}} & = & \cos(0) + i\mathrm{sen}(0) & = & 1 \\ z_1 & = & \mathrm{e}^{i\frac{2\pi}{4}} & = & \cos(\pi/2) + i\mathrm{sen}(\pi/2) & = & i \\ z_2 & = & \mathrm{e}^{i\frac{4\pi}{4}} & = & \cos(\pi) + i\mathrm{sen}(\pi) & = & -1 \\ z_3 & = & \mathrm{e}^{i\frac{6\pi}{4}} & = & \cos(3\pi/2) + i\mathrm{sen}(3\pi/2) & = & -i \end{array}$$



Finalmente calculamos as raíces sextas de $\omega = -1$. Dado que 6 é un número par o problema non admite solucións reais. Como

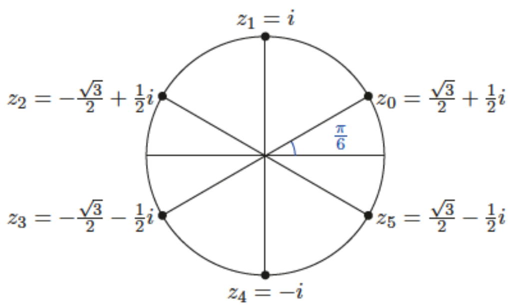
$$\omega = e^{i\pi},$$

tense que o argumento de ω é π e doutra banda $|\omega| = 1$. Daquela, as raíces sextas de -1 serán

$$z_k = \mathrm{e}^{i\frac{\pi+2k\pi}{6}},$$

con $k = 0, 1, 2, 3, 4, 5$. Isto é,

$$\begin{array}{rclcl}
z_0 & = & \mathrm{e}^{i\frac{\pi}{6}} & = & \cos(\frac{\pi}{6}) + i\,\mathrm{sen}(\frac{\pi}{6}) = \dfrac{\sqrt{3}}{2} + \dfrac{1}{2}i \\
z_1 & = & \mathrm{e}^{i\frac{3\pi}{6}} & = & \cos(\frac{3\pi}{6}) + i\,\mathrm{sen}(\frac{3\pi}{6}) = \cos(\frac{\pi}{2}) + i\,\mathrm{sen}(\frac{\pi}{2}) = i \\
z_2 & = & \mathrm{e}^{i\frac{5\pi}{6}} & = & \cos(\frac{5\pi}{6}) + i\,\mathrm{sen}(\frac{5\pi}{6}) = -\dfrac{\sqrt{3}}{2} + \dfrac{1}{2}i \\
z_3 & = & \mathrm{e}^{i\frac{7\pi}{6}} & = & \cos(\frac{7\pi}{6}) + i\,\mathrm{sen}(\frac{7\pi}{6}) = -\dfrac{\sqrt{3}}{2} - \dfrac{1}{2}i \\
z_4 & = & \mathrm{e}^{i\frac{9\pi}{6}} & = & \cos(\frac{9\pi}{6}) + i\,\mathrm{sen}(\frac{9\pi}{6}) = \cos(\frac{3\pi}{2}) + i\,\mathrm{sen}(\frac{3\pi}{2}) = -i \\
z_5 & = & \mathrm{e}^{i\frac{11\pi}{6}} & = & \cos(\frac{11\pi}{6}) + i\,\mathrm{sen}(\frac{11\pi}{6}) = \dfrac{\sqrt{3}}{2} - \dfrac{1}{2}i
\end{array}$$



2.3. Problemas propostos

1. Comprobe que:
   a) $(\sqrt{2} - i) - i(1 - \sqrt{2}i) = -2i$.
   b) $(3+i)(3-i)\left(\dfrac{1}{5} + \dfrac{1}{10}i\right) = 2 + i$.
   c) $\dfrac{5}{(1-i)(2-i)(3-i)} = \dfrac{i}{2}$.
   d) $(2-3i)(-2+i) = -1 + 8i$.
   e) $\dfrac{1+2i}{3-4i} + \dfrac{2-i}{5i} = -\dfrac{2}{5}$.
   f) $(1-i)^4 = -4$.
2. Comprobe que os complexos $z = 1 \pm i$ satisfán a ecuación $z^2 - 2z + 2 = 0$.
3. Resolva a ecuación $z^2 + z + 1 = 0$.
4. Probe que:
   a) $\mathrm{Im}(iz) = \mathrm{Re}(z)$.
   b) $\dfrac{1}{1/z} = z \ (z \neq 0)$.
   c) $\mathrm{Re}(iz) = -\mathrm{Im}(z)$.
   d) $(-1)z = -z$.

5. Probe que:
   a) $\overline{z+3i} = z - 3i$.
   b) $\overline{iz} = -i\overline{z}$.
   c) $\overline{(2+i)^2} = 3 - 4i$.
   d) $|(2\overline{z}+5)(\sqrt{2}-i)| = \sqrt{3}|2z+5|$.
6. Probe que:
   a) z é real, se, e só se, $\overline{z} = z$.
   b) z é real ou imaxinario puro, se, e só se, $\overline{z}^2 = z^2$.
7. En cada caso, caracterize o conxunto de puntos determinado pola condición exposta:
   a) $|z-1+i| = 1$.
   b) $|z+i| \leq 3$.
   c) $\mathrm{Re}(\overline{z}-i) = 2$.
   d) $|2z-i| = 4$.
8. Ache o argumento de:
   a) $z = \dfrac{-2}{1+\sqrt{3}i}$.
   b) $z = \dfrac{i}{-2-2i}$.
   c) $z = (\sqrt{3}-i)^6$.
9. Probe que:
   a) $i(1-\sqrt{3}i)(\sqrt{3}+i) = 2(1+\sqrt{3}i)$
   b) $5i/(2+i) = 1+2i$.
   c) $(-1+i)^7 = -8(1+i)$.
   d) $(1+\sqrt{3}i)^{-10} = 2^{-11}(-1+\sqrt{3}i)$.
10. Calcule:
    a) $(2i)^{1/2}$.
    b) $(1-\sqrt{3}i)^{1/2}$.
    c) $(-1)^{1/3}$.
    d) $(-16)^{1/4}$.
    e) $8^{1/6}$.
    f) $(-8-8\sqrt{3}i)^{1/4}$.
11. Ache as catro solucións da ecuación $z^4 + 4 = 0$.

CAPÍTULO 3

Matrices e determinantes

Na primeira parte do capítulo introducirase o concepto de matriz

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

con coeficientes a_{ij} nun corpo $\mathbb{K}$, onde $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$, e describiranse os principais tipos de matrices, como, por exemplo, as matrices diagonais, as triangulares superiores (inferiores), as matrices fila e columna, etc.

Posteriormente se definirán as operacións básicas que podemos realizar con matrices: suma de matrices, produto dun escalar por unha matriz e produto de matrices. Presentaranse as principais propiedades destas operacións, facendo especial fincapé nas dificultades que presenta o produto de matrices como, por exemplo, a falla de conmutatividade.

A continuación introducirase o operador trasposición e estudaranse as súas principais propiedades. O coñecemento deste operador permitiranos definir as nocións de matriz simétrica e de matriz antisimétrica e comprobar que toda matriz cadrada pódese escribir de forma única como suma dunha matriz simétrica e outra matriz antisimétrica. Dunha forma similar, cando os coeficientes da matriz son complexos, definiranse a trasposta conxugada dunha matriz e introduciranse as nocións de matriz hermitiana e antihermitiana.

O seguinte punto de interese do capítulo pasa pola definición de matriz invertíbel e o estudo das súas propiedades fundamentais. Como caso particular deste tipo de matrices introduciranse as nocións de matriz ortogonal e de matriz unitaria.

Manexar con soltura as operacións elementais por filas e por columnas, así como a súa interpretación matricial, forma parte do corazón deste capítulo xa que estas serán utilizadas para calcular o rango dunha matriz, calcular a súa inversa (se a ten), resolver sistemas de ecuacións, etc. É por iso que prestaremos unha grande atención a este tipo de operacións e sobre todo ás matrices, chamadas matrices de operacións elementais, que nos permiten mediante produtos, ben pola dereita ben pola esquerda, realizar as operacións elementais por filas e columnas. Finalmente, neste apartado comprobarase que as matrices de operacións elementais son todas invertíbeis.

Outro dos puntos centrais do capítulo é o cálculo de rangos de matrices. Para iso introduciranse as nocións de forma graduada reducida por filas ou por columnas e a de matrices equivalentes por filas ou por columnas. Mostrarase que cada matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ é equivalente por filas (columnas) a unha matriz graduada reducida por filas (columnas). Isto é, existe unha matriz P (Q) cadrada de orde m (n) que é invertíbel e produto de matrices de operacións elementais por filas (columnas), tal que PA (AQ) é graduada reducida por filas (columnas). Como consecuencia destes feitos obterase que, dada unha matriz cadrada A con coeficientes reais ou complexos:

1) A matriz A é equivalente por filas á matriz identidade I_n, se, e só se, A é invertíbel.
2) Se existe unha matriz $X \in \mathcal{M}_{n\times n}(\mathbb{K})$ tal que

$$AX = I_n \text{ ou } XA = I_n,$$

a matriz A é invertíbel.

Grazas ás operacións elementais tamén veremos que, dada unha matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ tal que ten unha forma graduada reducida por filas B con r entradas principais, existen matrices $P \in \mathcal{M}_{m\times m}(\mathbb{K})$ e $Q \in \mathcal{M}_{n\times n}(\mathbb{K})$ tales que:

1) P e Q son produto de matrices de operacións elementais por filas e por columnas respectivamente.
2) Cúmprese que
$$PAQ = I_r^{m\times n}$$
onde
$$I_r^{m\times n} = \left(\begin{array}{c|c} I_r & \Theta_{r,n-r} \\ --- & ----- \\ \Theta_{m-r,r} & \Theta_{m-r,n-r} \end{array} \right).$$

Esta última propiedade indica que as matrices A e $I_r^{m\times n}$ son equivalentes. A noción de equivalencia de matrices tamén a introduciremos neste capítulo, demostrándose que toda matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ é equivalente a unha única matriz da forma $I_r^{m\times n}$. A matriz $I_r^{m\times n}$ chamarase forma reducida completa de A e o rango de A, representado como rang (A), definirase como o número r asociado a $I_r^{m\times n}$. Este número é un invariante da matriz A xa que $I_r^{m\times n}$ é única para cada A. Doutra banda, resaltarase que o número r coincide co número de filas non nulas da forma graduada reducida por filas de A a partir da cal se calcula $I_r^{m\times n}$ ou, equivalentemente, co número de entradas principais desa forma.

Como consecuencia de todas estas propiedades, obterase que dúas matrices $A, B \in \mathcal{M}_{m\times n}(\mathbb{K})$ son equivalentes, se, e só se, teñen o mesmo rango. Tamén se comprobará que o rango non varía cando multiplicamos unha matriz pola esquerda ou pola dereita por unha matriz invertíbel e que o rango dunha matriz e o da súa trasposta son iguais. Finalmente, comprobarase que unha matriz cadrada de orde n é invertíbel se, e só se, é equivalente á matriz identidade I_n. Ademais indicarase como calcular a súa matriz inversa utilizando operacións elementais por filas e/ou columnas.

Na parte final do capítulo introducirase a noción de determinante de forma indutiva para chegar ao que se coñece como desenvolvemento de Laplace do determinante pola primeira columna da matriz. Posteriormente, veremos as propiedades fundamentais dos determinantes resaltando en especial que, grazas a elas, garántese que o determinante dunha matriz cadrada pódese desenvolver por calquera das súas filas ou columnas.

Doutra banda, indicarase como se utilizan as operacións elementais á hora de calcular determinantes, as relacións existentes entre o determinante dun produto e o produto dos determinantes e canto vale o determinante da inversa dunha matriz invertíbel A coñecido o determinante de A. Máis concretamente, se A e B son dúas matrices cadradas, tense que
$$\det(AB) = \det(A)\det(B),$$
e ademais se A é invertíbel
$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

3.1. Matrices. Operacións con matrices

Definición 3.1.1. Defínese unha matriz de orde m por n, ou de m filas e n columnas, con coeficientes no corpo $\mathbb{K}$ como unha táboa rectangular
$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

onde a_{ij} (o elemento colocado na fila i e na columna j) pertence a $\mathbb{K}$ para todo $i \in \{1, \dots, m\}$ e $j \in \{1, \dots, n\}$.

O conxunto de todas as matrices de m filas e n columnas con coeficientes en $\mathbb{K}$ denótase por $\mathcal{M}_{m\times n}(\mathbb{K})$ e para os elementos deste conxunto utilizaremos a notación abreviada $A = (a_{ij})$. Dúas matrices $A = (a_{ij})$ e $B = (b_{ij})$ pertencentes a $\mathcal{M}_{m\times n}(\mathbb{K})$ son iguais, se $a_{ij} = b_{ij}$ para todos $i \in \{1, \dots, m\}$ e $j \in \{1, \dots, n\}$. No caso de que $m = n$, a matriz chamarase cadrada de orde n.

Para unha matriz A chamaremos submatriz de A a toda matriz obtida dela suprimindo algunha das súas filas ou columnas. Un bloque é unha submatriz de A onde os índices de filas e columnas son consecutivos.

Exemplo 3.1.2. A matriz

$$A = \begin{pmatrix} 1 & 2-i & -1+2i \\ 1 & 2 & -2+i \\ -1 & -2+i & 2-2i \end{pmatrix}$$

é cadrada de tres filas e tres columnas con coeficientes en $\mathbb{C}$. Por tanto, pertence ao conxunto $\mathcal{M}_{3\times 3}(\mathbb{C})$. Se suprimimos a primeira fila, obtemos a submatriz

$$B = \begin{pmatrix} 1 & 2 & -2+i \\ -1 & -2+i & 2-2i \end{pmatrix}$$

que non é cadrada, pertence ao conxunto $\mathcal{M}_{2\times 3}(\mathbb{C})$ e é un bloque de A.

A matriz

$$C = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

pertence ao conxunto $\mathcal{M}_{4\times 3}(\mathbb{R})$. Se eliminamos as dúas últimas columnas e a terceira fila obtemos a submatriz de C dada por

$$C = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$$

que pertence a $\mathcal{M}_{3\times 1}(\mathbb{R})$ e non é un bloque de A.

Definición 3.1.3. Dada unha matriz cadrada $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{K})$, os coeficientes da forma a_{ii} constitúen a diagonal principal da matriz. A este conxunto de elementos denotarémolo como $\operatorname{diag}(A)$. Dise que unha matriz cadrada A é diagonal se todos os coeficientes fóra da diagonal principal son nulos, isto é, $a_{ij} = 0$ para todo $i \neq j$. Por tanto, este tipo de matrices son da forma:

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

Cando unha matriz A diagonal cumpre que $a_{ii} = a$ para todo $i \in \{1, \dots, n\}$ chamarase escalar. Dentro das matrices escalares debemos resaltar o caso particular de $a = 1$. Esta matriz chámase matriz identidade de orde n e representarémola como I_n. Daquela,

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Definición 3.1.4. Unha matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ é triangular superior se $a_{ij} = 0$ para todo $i > j$. No caso de que $a_{ij} = 0$ para todo $i < j$, a matriz é triangular inferior.

Exemplo 3.1.5. A matriz

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

é triangular superior e a matriz

$$B = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 & 0 \\ 3-i & 9 & 9 & 0 & 0 \end{pmatrix}$$

é triangular inferior.

Definición 3.1.6. Unha matriz $A \in \mathcal{M}_{1\times n}(\mathbb{K})$ chámase matriz fila e unha matriz $B \in \mathcal{M}_{m\times 1}(\mathbb{K})$ chámase matriz columna. Por conseguinte, estas matrices son da forma

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{pmatrix}, \qquad B = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix},$$

respectivamente.

Despois destas definicións preliminares, pasemos a continuación a estudar as principais operacións que podemos realizar con matrices. A primeira delas é a suma.

Definición 3.1.7. Dadas dúas matrices $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ e $B = (b_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$, defínese a suma $A + B$ como a matriz pertencente a $\mathcal{M}_{m\times n}(\mathbb{K})$ dada por

$$A + B = (a_{ij} + b_{ij}).$$

Por tanto,

$$A + B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

Obsérvese que a suma de matrices só está definida para matrices de igual orde. É relativamente sinxelo comprobar que o conxunto $\mathcal{M}_{m\times n}(\mathbb{K})$ xunto coa operación de suma de matrices é un grupo conmutativo. Isto é, a suma de matrices cumpre as seguintes propiedades

1) ASOCIATIVA: $A + (B + C) = (A + B) + C,\ \forall\, A, B, C \in \mathcal{M}_{m\times n}(\mathbb{K})$.
2) CONMUTATIVA: $A + B = B + A,\ \forall\, A, B \in \mathcal{M}_{m\times n}(\mathbb{K})$.
3) ELEMENTO NEUTRO: $\exists\, \Theta_{m,n} \in \mathcal{M}_{m\times n}(\mathbb{K}) \mid A + \Theta_{m,n} = A,\ \forall\, A \in \mathcal{M}_{m\times n}(\mathbb{K})$.
4) ELEMENTO OPOSTO: $\forall\, A \in \mathcal{M}_{m\times n}(\mathbb{K})\ \exists\, -A \in \mathcal{M}_{m\times n}(\mathbb{K}) \mid A + (-A) = \Theta_{m,n}$.

Obviamente a matriz $\Theta_{m,n}$ que xoga o papel do elemento neutro é a matriz

$$\Theta_{m,n} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

e, se $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$, a oposta de A é

$$-A = \begin{pmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{m1} & -a_{m2} & \cdots & -a_{mn} \end{pmatrix}$$

Exemplo 3.1.8. Dadas as matrices pertencentes a $\mathcal{M}_{3\times 5}(\mathbb{C})$

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 & 0 \\ 3-i & 9 & 9 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 1 & 1 & 1 & i \\ 3 & -5 & i & 2-i & 0 \\ -i & 0 & 5+4i & 1 & 1 \end{pmatrix}$$

a súa suma é

$$A+B = \begin{pmatrix} 3+3 & 0+1 & 0+1 & 0+1 & 0+i \\ 3+3 & 5+(-5) & 0+i & 0+2-i & 0+0 \\ (3-i)+(-i) & 9+0 & 9+(5+4i) & 0+1 & 0+1 \end{pmatrix} =$$

$$\begin{pmatrix} 6 & 1 & 1 & 1 & i \\ 6 & 0 & i & 2-i & 0 \\ 3-2i & 9 & 14+4i & 1 & 1 \end{pmatrix}$$

A seguinte operación interesante que podemos facer coas matrices é o produto por escalares.

Definición 3.1.9. Dada unha matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ e un escalar $\alpha \in \mathbb{K}$, defínese o produto de α por A como a matriz $\alpha A \in \mathcal{M}_{m\times n}(\mathbb{K})$ dada por

$$\alpha A = (\alpha a_{ij}).$$

Por tanto,

$$\alpha A = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{pmatrix}$$

Da mesma forma que coa suma de matrices, non é difícil comprobar que o conxunto $\mathcal{M}_{m\times n}(\mathbb{K})$ xunto coa operación produto por escalares cumpre as seguintes propiedades:

1) DISTRIBUTIVA RESPECTO Á SUMA DE ESCALARES:
$$(\alpha+\beta)A = \alpha A + \beta A \ \forall\, A \in \mathcal{M}_{m\times n}(\mathbb{K}),\ \forall\, \alpha, \beta \in \mathbb{K}.$$
2) DISTRIBUTIVA RESPECTO Á SUMA DE MATRICES:
$$\alpha(A+B) = \alpha A + \alpha B \ \forall\, A, B \in \mathcal{M}_{m\times n}(\mathbb{K}),\ \forall\, \alpha \in \mathbb{K}.$$
3) ASOCIATIVA DO PRODUTO DE ESCALARES CO PRODUTO DE ESCALARES CON MATRICES:
$$(\alpha\beta)A = \alpha(\beta A) \ \forall\, A \in \mathcal{M}_{m\times n}(\mathbb{K}),\ \forall\, \alpha, \beta \in \mathbb{K}.$$
4) LEY DA IDENTIDADE: $1A = A$, $\forall\, A \in \mathcal{M}_{m\times n}(\mathbb{K})$.

Exemplo 3.1.10. Dada a matriz

$$A = \begin{pmatrix} 3-i & 1 & 5+i & 1 & 0 \\ 3 & 5 & 3 & 3 & \sqrt{2} \end{pmatrix}$$

o seu produto polo escalar $1+2i$ é

$$(1+2i)A = \begin{pmatrix} (1+2i)(3-i) & 1(1+2i) & (1+2i)(5+i) & 1(1+2i) & 0(1+2i) \\ 3(1+2i) & 5(1+2i) & 3(1+2i) & 3(1+2i) & \sqrt{2}(1+2i) \end{pmatrix} =$$

$$\begin{pmatrix} 5+5i & 1+2i & 3+11i & 1+2i & 0 \\ 3+6i & 5+10i & 3+6i & 3+6i & \sqrt{2}+2\sqrt{2}i \end{pmatrix}$$

Finalmente, neste apartado introducimos a operación produto de matrices, que, como veremos, é a que presenta maior número de diferenzas respecto ao produto de números reais ou complexos.

Definición 3.1.11. Dadas dúas matrices $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ e $B = (b_{jl}) \in \mathcal{M}_{n\times p}(\mathbb{K})$, defínese o produto AB como a matriz pertencente a $\mathcal{M}_{m\times p}(\mathbb{K})$ dada por

$$AB = (c_{il}), \quad c_{il} = \sum_{j=1}^{n} a_{ij}b_{jl} = a_{i1}b_{1l} + a_{i2}b_{2l} + \cdots + a_{in}b_{nl}.$$

En consecuencia, o produto de matrices só está definido se o número de columnas de A é coincidente co número de filas de B. Isto implica que, se A e B son matrices cadradas de orde n, sempre se poden multiplicar en calquera orde e a operación produto é interna en $\mathcal{M}_{n\times n}(\mathbb{K})$, isto é, se $A, B \in \mathcal{M}_{n\times n}(\mathbb{K})$, as matrices AB e BA pertencen a $\mathcal{M}_{n\times n}(\mathbb{K})$. O que acabamos de afirmar non implica que AB e BA sexan iguais xa que, por exemplo, se

$$A = \begin{pmatrix} 6 & 1 & 1 \\ 6 & 0 & 0 \\ 3 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

temos que

$$AB = \begin{pmatrix} 6\cdot 0 + 1\cdot 0 + 1\cdot 1 & 6\cdot 0 + 1\cdot 1 + 1\cdot 0 & 6\cdot 1 + 1\cdot 0 + 1\cdot 0 \\ 6\cdot 0 + 0\cdot 0 + 1\cdot 1 & 6\cdot 0 + 0\cdot 1 + 0\cdot 0 & 6\cdot 1 + 0\cdot 0 + 0\cdot 0 \\ 3\cdot 0 + 1\cdot 0 + 1\cdot 1 & 3\cdot 0 + 1\cdot 1 + 1\cdot 0 & 3\cdot 1 + 1\cdot 0 + 1\cdot 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 6 \\ 0 & 0 & 6 \\ 1 & 1 & 3 \end{pmatrix}$$

e, doutra banda,

$$BA = \begin{pmatrix} 0\cdot 6 + 0\cdot 6 + 1\cdot 3 & 0\cdot 1 + 0\cdot 0 + 1\cdot 1 & 0\cdot 1 + 0\cdot 0 + 1\cdot 1 \\ 0\cdot 6 + 1\cdot 6 + 0\cdot 3 & 0\cdot 1 + 1\cdot 0 + 0\cdot 1 & 0\cdot 1 + 1\cdot 0 + 0\cdot 1 \\ 1\cdot 6 + 0\cdot 6 + 0\cdot 3 & 1\cdot 1 + 0\cdot 0 + 0\cdot 1 & 1\cdot 1 + 0\cdot 0 + 0\cdot 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 1 \\ 6 & 0 & 0 \\ 6 & 1 & 1 \end{pmatrix}$$

Por tanto, o produto de matrices non é conmutativo.

A pesar da falta de conmutatividade o produto de matrices cumpre algunhas propiedades interesantes como, por exemplo:

1) ASOCIATIVA: $A(BC) = (AB)C$, $\forall\, A \in \mathcal{M}_{m\times n}(\mathbb{K}), B \in \mathcal{M}_{n\times p}(\mathbb{K}), C \in \mathcal{M}_{p\times q}(\mathbb{K})$.
2) ELEMENTO NEUTRO: $AI_n = A = I_nA$, $\forall\, A \in \mathcal{M}_{n\times n}(\mathbb{K})$.
3) DISTRIBUTIVA RESPECTO Á SUMA DE MATRICES:
$$A(B + C) = AB + AC, \ \forall\, A \in \mathcal{M}_{m\times n}(\mathbb{K}), \ \forall\ B, C \in \mathcal{M}_{n\times p}(\mathbb{K}).$$
4) ASOCIATIVIDADE RESPECTO AO PRODUTO DE ESCALARES:
$$\alpha(AB) = (\alpha A)B, \ \forall\ A \in \mathcal{M}_{m\times n}(\mathbb{K}), \ \forall\, B \in \mathcal{M}_{n\times p}(\mathbb{K}), \ \forall\, \alpha \in \mathbb{K}.$$

Nota 3.1.12. Débese resaltar tamén que, cando se ten que dous produtos de matrices AB e AC coinciden, isto non implica que $B = C$. Por exemplo, se

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix}$$

temos que

$$AB = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} = AC$$

e A e B non son iguais.

Nota 3.1.13. É interesante ver que no caso particular de multiplicar unha matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ por unha matriz columna $b = (b_{j1}) \in \mathcal{M}_{n\times 1}(\mathbb{K})$ temos que $Ab \in \mathcal{M}_{m\times 1}(\mathbb{K})$ e ademais é combinación linear das columnas de A xa que:

$$Ab = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix} = b_{11} \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + b_{21} \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + b_{n1} \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Doutra banda, se $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ e $B = (b_{jl}) \in \mathcal{M}_{n\times p}(\mathbb{K})$, entón denotando a columna l-ésima de B por

$$b_l = \begin{pmatrix} b_{1l} \\ b_{2l} \\ \vdots \\ b_{nl} \end{pmatrix}$$

temos que a columna l-ésima de AB é Ab_l e, como consecuencia, podemos pensar o produto das matrices A e B como

$$AB = (Ab_1|Ab_2|\cdots|Ab_p).$$

Por tanto, cada columna de AB é combinación linear das columnas de A.

Definición 3.1.14. Dada unha matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$, unha descomposición por bloques de A está dada por dous conxuntos de números naturais

$$\{m_1, \dots, m_M\}, \ \{n_1, \dots, n_N\},$$

cumprindo as igualdades

$$\sum_{I=1}^{M} m_I = m, \ \sum_{J=1}^{N} n_J = n,$$

de forma que determinan MN bloques A_{IJ} de A onde os coeficientes de A que pertencen ao bloque A_{IJ} tómanse do xeito seguinte:

- Filas:
  $I = 1$: Desde a fila 1 ata a fila m_1.
  $I > 1$: Desde a fila $m_1 + \cdots + m_{I-1} + 1$ ata a fila $m_1 + \cdots + m_{I-1} + m_I$.
- Columnas:
  $J = 1$: Desde a columna 1 ata a columna n_1.
  $J > 1$: Desde a columna $n_1 + \cdots + n_{J-1} + 1$ ata a columna $n_1 + \cdots + n_{J-1} + n_J$.

Daquela, podemos escribir A como

$$A = \left(\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1N} \\ -- & -- & \cdots & -- \\ A_{21} & A_{22} & \cdots & A_{2N} \\ -- & -- & \cdots & -- \\ \vdots & \vdots & \ddots & \vdots \\ -- & -- & \cdots & -- \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{array}\right)$$

onde $A_{IJ} \in \mathcal{M}_{m_I\times n_J}(\mathbb{K})$.

Exemplo 3.1.15. Sexa a matriz

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 \\ 3-i & 9 & 9 & 0 \\ 1 & 1 & 9 & 0 \end{pmatrix}$$

e tomemos os naturais

$$m_1 = 2,\ m_2 = 1, ; m_3 = 1,\ \ n_1 = 2,\ n_2 = 1,\ n_3 = 1.$$

Entón, estes números determinan a seguinte descomposición por bloques de A:

$$A = \begin{pmatrix} 3 & 0 & | & 0 & | & 0 \\ 3 & 5 & | & 0 & | & 0 \\ - & - & | & - & | & - \\ 3-i & 9 & | & 9 & | & 0 \\ - & - & | & - & | & - \\ 1 & 1 & | & 9 & | & 0 \end{pmatrix}.$$

Se tomásemos

$$m_1 = 4,\ \ n_1 = 1,\ n_2 = 1,\ n_3 = 1,\ n_4 = 1,$$

obteriamos unha descomposición en bloques para A onde cada bloque coincide coas columnas da matriz.

Finalmente, se tomamos

$$m_1 = 2,\ m_2 = 2,\ \ n_1 = 2,\ n_2 = 2$$

a descomposición por bloques será

$$A = \begin{pmatrix} 3 & 0 & | & 0 & 0 \\ 3 & 5 & | & 0 & 0 \\ - & - & | & - & - \\ 3-i & 9 & | & 9 & 0 \\ 1 & 1 & | & 9 & 0 \end{pmatrix}.$$

O produto de matrices tamén admite unha formulación por bloques.

3.1.16. Dadas dúas matrices $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ e $B = (b_{jl}) \in \mathcal{M}_{n\times p}(\mathbb{K})$, con descomposicións por bloques

$$\{m_1, \dots, m_M\},\ \ \{n_1, \dots, n_N\},\ \ \sum_{I=1}^{M} m_I = m,\ \ \sum_{J=1}^{N} n_J = n$$

e

$$\{n_1, \dots, n_N\},\ \ \{p_1, \dots, p_P\},\ \ \sum_{J=1}^{N} n_J = n,\ \ \sum_{L=1}^{P} p_L = p$$

respectivamente, temos que o produto AB tamén admite unha descomposición por bloques determinada polos números naturais

$$\{m_1, \dots, m_M\},\ \ \{p_1, \dots, p_P\},\ \ \sum_{I=1}^{M} m_I = m,\ \ \sum_{L=1}^{P} p_L = p$$

onde

$$AB = (C_{IL}), \quad C_{IL} = \sum_{J=1}^{N} A_{IJ}B_{JL} = A_{I1}B_{1L} + A_{I2}B_{2L} + \cdots + A_{IN}B_{NL}.$$

Exemplo 3.1.17. Sexa a matriz

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$$

e a descomposición dada por

$$m_1 = 2,\ m_2 = 1,\ \ n_1 = 2,\ n_2 = 1.$$

Sexa a matriz

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 1 & 0 \end{pmatrix}$$

e a descomposición dada por

$$n_1 = 2,\ n_2 = 1,\ \ p_1 = 1,\ p_2 = 1.$$

Entón, podemos facer o produto AB por bloques obtendo unha descomposición dada por

$$m_1 = 2,\ m_2 = 1, \quad p_1 = 1,\ p_2 = 1$$

onde

$$AB = \left(\begin{array}{c|c} C_{11} & C_{12} \\ - & - \\ C_{21} & C_{22} \end{array} \right) = \left(\begin{array}{cc|c} 2 & 3 & 4 \\ 1 & 2 & 3 \\ - & - & - \\ 1 & 1 & 1 \end{array} \right) \left(\begin{array}{c|c} 1 & 2 \\ 2 & 3 \\ - & - \\ 1 & 0 \end{array} \right)$$

$$\left(\begin{array}{c|c} \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 4 \\ 3 \end{pmatrix} \cdot 1 & \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 3 \end{pmatrix} \cdot 0 \\ - - - - - - - - - - - - - & - - - - - - - - - - - - - \\ \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 1 \cdot 1 & \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} + 1 \cdot 0 \end{array} \right)$$

$$= \left(\begin{array}{c|c} 12 & 13 \\ 8 & 8 \\ - & - \\ 4 & 5 \end{array} \right).$$

Definición 3.1.18. Dada unha matriz cadrada A defínese, para cada $k \in \mathbb{N}$, a potencia k-ésima de A como

$$A^k = AA^{k-1}$$

sendo $A^1 = A$.

De forma directa dedúcense as seguintes propiedades para toda matriz cadrada A

1) $A^p A^q = A^{p+q}$; $\forall\, p, q \in \mathbb{N}$.
2) $(A^p)^q = A^{pq}$; $\forall\, p, q \in \mathbb{N}$.
3) Se A é unha matriz diagonal con diagonal principal $a_{11}, \dots, a_{nn}$, entón $\forall\, p \in \mathbb{N}$, A^p é unha matriz diagonal con diagonal principal $a_{11}^p, \dots, a_{nn}^p$.

Se A é unha matriz cadrada de orde n, dise que é nilpotente se existe $k \in \mathbb{N}$ tal que $A^k = \Theta_{n,n}$. Ao menor $k \in \mathbb{N}$ tal que $A^k = \Theta_{n,n}$ chámaselle índice de nilpotencia de A.

Por exemplo, a matriz

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

é nilpotente de índice dous xa que $A^2 = \Theta_{2,2}$.

Definición 3.1.19. Dada unha matriz cadrada A, dise que é idempotente se $A^2 = A$.

Por exemplo a matriz

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

é idempotente xa que $A^2 = A$.

Por tanto, se unha matriz A é idempotente $A^k = A$ para todo número natural k.

Definición 3.1.20. Dada unha matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$, defínese a súa matriz trasposta, denotada como A^t, como a que obtemos intercambiando en A filas e columnas. Logo, $A^t \in \mathcal{M}_{n\times m}(\mathbb{K})$ e

$$A^t = (b_{ji});\ b_{ji} = a_{ij}.$$

Por exemplo, a trasposta da matriz

$$A = \begin{pmatrix} 3-i & 1 & 5+i & 1 & 0 \\ 3 & 5 & 3 & 3 & \sqrt{2} \end{pmatrix} \in \mathcal{M}_{2\times 5}(\mathbb{C})$$

é a matriz

$$A^t = \begin{pmatrix} 3-i & 3 \\ 1 & 5 \\ 5+i & 3 \\ 1 & 3 \\ 0 & \sqrt{2} \end{pmatrix} \in \mathcal{M}_{5\times 2}(\mathbb{C})$$

De forma inmediata (salvo a relativa ao produto que non o é, sendo en calquera caso fácil de demostrar) pódense probar as seguintes propiedades:

1) $(A^t)^t = A$, $\forall\, A \in \mathcal{M}_{m\times n}(\mathbb{K})$.
2) $(A+B)^t = A^t + B^t$, $\forall\, A, B \in \mathcal{M}_{m\times n}(\mathbb{K})$.
3) $(\alpha A)^t = \alpha A^t$, $\forall\, A \in \mathcal{M}_{m\times n}(\mathbb{K})$, $\forall\, \alpha \in \mathbb{K}$.
4) $(AB)^t = B^t A^t$, $\forall\, A \in \mathcal{M}_{m\times n}(\mathbb{K})$, $\forall\, B \in \mathcal{M}_{n\times m}(\mathbb{K})$.

Definición 3.1.21. Unha matriz cadrada $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{K})$ dise que é simétrica se $A^t = A$. Por tanto, A é simétrica se, e só se, $a_{ij} = a_{ji}$ para todo $i \neq j$.

Unha matriz cadrada $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{K})$ dise que é antisimétrica se $A^t = -A$. Por tanto, A é antisimétrica se, e só se, $a_{ij} = -a_{ji}$ para todo $i \neq j$ e, ademais, $a_{ii} = 0$ para todo $i \in \{1, \dots, n\}$.

Exemplo 3.1.22. Por exemplo a matriz

$$A = \begin{pmatrix} 6 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

é simétrica e a matriz

$$B = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & -i \\ -1 & i & 0 \end{pmatrix}$$

é antisimétrica.

Nota 3.1.23. A importancia das matrices simétricas e antisimétricas radica, entre outras cousas, en que toda matriz cadrada pódese poñer de forma única como suma dunha matriz simétrica e outra antisimétrica. Se a matriz de partida é A, temos que

$$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t)$$

onde $\frac{1}{2}(A + A^t)$ é a matriz simétrica e $\frac{1}{2}(A - A^t)$ a antisimétrica.

Definición 3.1.24. Sexa $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{C})$. Chámase matriz conxugada de A e denótase por $\overline{A}$ á matriz resultante de conxugar todos os coeficientes de A. Daquela,

$$\overline{A} = (b_{ij});\ b_{ij} = \overline{a_{ij}}.$$

Con A^* denotaremos á matriz $\overline{A}^t \in \mathcal{M}_{n\times m}(\mathbb{C})$. Para o operador $*$ temos as seguintes propiedades:

1) $(A^*)^* = A \ \ \forall\, A \in \mathcal{M}_{m\times n}(\mathbb{C})$.
2) $(A+B)^* = A^* + B^* \ \ \forall\, A, B \in \mathcal{M}_{m\times n}(\mathbb{C})$.
3) $(\alpha A)^* = \overline{\alpha} A^* \ \ \forall\, A \in \mathcal{M}_{m\times n}(\mathbb{C}),\ \forall\, \alpha \in \mathbb{C}$.
4) $(AB)^* = B^* A^* \ \ \forall\, A \in \mathcal{M}_{m\times n}(\mathbb{C}),\ \forall\, B \in \mathcal{M}_{n\times m}(\mathbb{C})$.

Exemplo 3.1.25. Se, por exemplo, tomamos a matriz

$$A = \begin{pmatrix} 1-i & 1 & 1 \\ -1 & 0 & -i \\ -1 & i & 0 \end{pmatrix}$$

temos que

$$\overline{A} = \begin{pmatrix} 1+i & 1 & 1 \\ -1 & 0 & i \\ -1 & -i & 0 \end{pmatrix}$$

e, daquela,

$$A^* = \begin{pmatrix} 1+i & -1 & -1 \\ 1 & 0 & -i \\ 1 & i & 0 \end{pmatrix}$$

Definición 3.1.26. Unha matriz cadrada $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{C})$ dise que é hermitiana se $A^* = A$. Por tanto, A é hermitiana se, e só se, $a_{ij} = \overline{a_{ji}}$. En particular isto implica que $a_{ii} = \overline{a_{ii}}$ ou, o que é o mesmo, todos os elementos da diagonal principal son números reais.

Unha matriz cadrada $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{C})$ dise que é antihermitiana se $A^* = -A$. Por tanto, A é antihermitiana, se, e só se, $-a_{ij} = \overline{a_{ji}}$. En particular isto implica que $-a_{ii} = \overline{a_{ii}}$ ou, o que é o mesmo, todos os elementos da diagonal principal teñen a parte real nula.

Exemplo 3.1.27. Se, por exemplo, tomamos a matriz

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & i \\ 1 & -i & 0 \end{pmatrix}$$

temos que A é hermitiana e, se B é a matriz

$$B = \begin{pmatrix} i & 1 & 1 \\ -1 & 0 & i \\ -1 & i & 0 \end{pmatrix},$$

obtemos un exemplo de matriz antihermitiana.

Definición 3.1.28. Sexa A unha matriz cadrada de orde n. Diremos que A é unha matriz invertíbel, ou que ten inversa, se existe outra matriz B da mesma orde, tal que

$$AB = BA = I_n.$$

Xeralmente, a matriz B coñécese co nome de inversa de A e denótase como A^{-1}. As matrices invertíbeis chámanse tamén matrices regulares ou matrices non singulares.

Cando unha matriz cadrada é invertíbel a súa inversa é única. En efecto, sexa A unha matriz invertíbel e supoñamos que existen B e C tales que

$$AB = BA = AC = CA = I_n.$$

Entón,

$$B = BI_n = B(AC) = (BA)C = I_nC = C.$$

Utilizando unicidade da inversa pódense probar de forma inmediata as seguintes propiedades:

1) Se A e B son matrices cadradas invertíbeis, tamén o é o seu produto e cúmprese que

$$(AB)^{-1} = B^{-1}A^{-1}.$$

2) Se unha matriz A é invertíbel, tamén o é A^{-1} e cúmprese que

$$(A^{-1})^{-1} = A.$$

3) Se α é un escalar non nulo e A é unha matriz invertíbel, cúmprese que αA é invertíbel e

$$(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}.$$

4) Se A é invertíbel, tense que A^k é invertíbel para todo $k \in \mathbb{N}$. Ademais

$$(A^k)^{-1} = (A^{-1})^k.$$

5) Se A é invertíbel, entón A^t tamén o é e

$$(A^t)^{-1} = (A^{-1})^t.$$

6) Se $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ é invertíbel, entón A^* tamén o é e

$$(A^*)^{-1} = (A^{-1})^*.$$

Exemplo 3.1.29. Sexa A unha matriz cadrada de orde dúas dada por

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

e tal que $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Daquela, A é invertíbel e a súa inversa está dada por

$$A^{-1} = \begin{pmatrix} \dfrac{a_{22}}{a_{11}a_{22} - a_{12}a_{21}} & \dfrac{-a_{12}}{a_{11}a_{22} - a_{12}a_{21}} \\ \dfrac{-a_{21}}{a_{11}a_{22} - a_{12}a_{21}} & \dfrac{a_{11}}{a_{11}a_{22} - a_{12}a_{21}} \end{pmatrix}$$

Esta fórmula é un caso particular das que podemos obter de forma máis xeral seguindo o método de Cramer. Ten o grave defecto de que, cando a matriz A é grande, o número de operacións necesarias para o cálculo da inversa fano inviábel, mesmo dispoñendo de grandes recursos computacionais.

Nota 3.1.30. É importante resaltar que toda matriz que teña unha columna ou unha fila de ceros non pode ser invertíbel. En efecto, se $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ ten unha columna de ceros, entón $\forall\ B \in \mathcal{M}_{n\times n}(\mathbb{K})$ tense que BA ten unha columna de ceros e non pode ser a identidade. Se A ten unha fila de ceros, a súa trasposta ten unha columna de ceros e, por tanto, non ten inversa. Se a trasposta non é invertíbel, A tampouco o é.

Entre as matrices invertíbeis debemos resaltar dous tipos.

Definición 3.1.31. Diremos que unha matriz cadrada A é ortogonal se é invertíbel e a súa inversa é A^t.

Unha matriz cadrada $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ dise que é unitaria se é invertíbel e a súa inversa coincide con A^*.

Exemplo 3.1.32. A matriz

$$A = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

é ortogonal xa que $A^tA = AA^t = I_2$.

A matriz

$$A = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

é unitaria xa que $AA^* = A^*A = I_2$.

3.2. Operacións elementais. Forma graduada reducida. Rango dunha matriz

Existen uns tipos especiais de matrices que serán de grande utilidade á hora de calcular inversas e rangos. Estas matrices permitirán interpretar desde o punto de vista do produto de matrices as operacións elementais sobre as filas ou as columnas dunha matriz A calquera.

Definición 3.2.1. Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Denomínanse operacións elementais sobre as filas da matriz A a calquera das seguintes transformacións:

- Permutar dúas filas de A.
- Multiplicar unha fila por un escalar.
- Sumar a unha fila un múltiplo doutra fila.

Do mesmo xeito, chámanse operacións elementais sobre as columnas de A ás transformacións:

- Permutar dúas columnas de A.
- Multiplicar unha columna por un escalar.
- Sumar a unha columna un múltiplo doutra columna.

Definición 3.2.2. Toda matriz obtida despois de facer unha operación elemental nunha matriz identidade chamarase matriz de operacións elementais.

A continuación enumeraremos todas as matrices de operacións elementais que podemos construír así como o significado das operacións elementais desde un punto de vista matricial.

As matrices de operacións elementais son as seguintes:

1) Con F_{ij} denotamos a matriz resultante de permutar nunha matriz identidade a fila i e a fila j.

   Por exemplo, para $n = 3$

$$F_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

   é o resultado de intercambiar en I_3 as filas primeira e terceira.

2) Con $F_i(\alpha)$ denotamos a matriz obtida ao multiplicar a fila i dunha matriz identidade por un escalar α non nulo.

   Por exemplo, para $n = 4$

$$F_4(4) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

   é o resultado de multiplicar en I_4 a fila cuarta por catro.

3) Con $F_{ij}(\alpha)$ denotamos a matriz obtida ao sumar á fila i dunha matriz identidade a súa fila j multiplicada por un escalar α.

Por exemplo, para $n=4$

$$F_{23}(\pi)=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \pi & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

é o resultado de multiplicar en I_4 a fila terceira por π e sumar o resultado á segunda.

Obsérvese que no exemplo anterior a matriz $F_{23}(\pi)$ é triangular superior con uns na diagonal principal xa que estamos a sumar a unha fila, a segunda, un múltiplo doutra que está debaixo, isto é, un múltiplo da terceira fila. Se por exemplo, sumamos á fila terceira a segunda multiplicada por π, obtemos a matriz.

$$F_{32}(\pi)=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \pi & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

que resulta ser triangular inferior con uns na diagonal principal.

4) Con C_{ij} denotamos a matriz resultante de permutar nunha matriz identidade a columna i e a columna j.

Por exemplo, para $n=3$

$$C_{12}=\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

é o resultado de intercambiar en I_3 as columnas primeira e segunda.

5) Con $C_i(\alpha)$ denotamos a matriz obtida ao multiplicar a fila i dunha matriz identidade por un escalar α non nulo.

Por exemplo, para $n=4$

$$C_3(4)=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

é o resultado de multiplicar en I_4 a columna terceira por catro.

6) Con $C_{ij}(\alpha)$ denotamos a matriz obtida ao sumar á columna i dunha matriz identidade a súa columna j multiplicada por un escalar α.

Por exemplo, para $n=4$

$$C_{23}(\pi)=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \pi & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

é o resultado de multiplicar en I_4 a columna terceira por π e sumar o resultado á segunda.

Obsérvese que no exemplo anterior a matriz $C_{23}(\pi)$ é triangular inferior con uns na diagonal principal xa que estamos a sumar a unha columna, a segunda, un múltiplo doutra que está despois, isto é, un múltiplo da terceira columna. Se por exemplo, sumamos á columna terceira a segunda multiplicada por π, obtemos a matriz

$$C_{32}(\pi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \pi & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

que resulta ser triangular superior con uns na diagonal principal.

Con estas matrices elementais, dada $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, temos o seguinte:

1) $F_{ij}A$ é a matriz que se obtén ao permutar en A as filas i e j.
2) $F_i(\alpha)A$ é a matriz que se obtén ao multiplicar a fila i de A polo escalar α.
3) $F_{ij}(\alpha)A$ é a matriz que se obtén ao sumar á fila i de A a fila j de A multiplicada polo escalar α.
4) AC_{ij} é a matriz que se obtén ao permutar en A as columnas i e j.
5) $AC_i(\alpha)$ é a matriz que se obtén ao multiplicar a columna i de A polo escalar α.
6) $AC_{ij}(\alpha)$ é a matriz que se obtén ao sumar á columna i de A a columna j de A multiplicada polo escalar α.

Por tanto, todas as operacións elementais pódense realizar multiplicando a matriz A, ben pola dereita ou ben pola esquerda, pola matriz de operacións elementais axeitada. É importante resaltar que, se $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, entón F_{ij}, $F_i(\alpha)$, $F_{ij}(\alpha)$ son matrices pertencentes ao conxunto $\mathcal{M}_{m\times m}(\mathbb{K})$ e, doutra banda, C_{ij}, $C_i(\alpha)$, $C_{ij}(\alpha)$ son matrices pertencentes ao conxunto $\mathcal{M}_{n\times n}(\mathbb{K})$.

Exemplo 3.2.3. Supoñamos que temos a matriz

$$A = \begin{pmatrix} 6 & 2 & 2 \\ 6 & 1 & 3 \\ 3 & 1 & 1 \end{pmatrix}$$

daquela

$$F_{13}A = \begin{pmatrix} 3 & 1 & 1 \\ 6 & 1 & 3 \\ 6 & 2 & 2 \end{pmatrix}, \quad F_2(3)A = \begin{pmatrix} 6 & 2 & 2 \\ 18 & 3 & 9 \\ 3 & 1 & 1 \end{pmatrix}, \quad F_{32}(-1)A = \begin{pmatrix} 6 & 2 & 2 \\ 6 & 1 & 3 \\ -3 & 0 & -2 \end{pmatrix}$$

e

$$AC_{12} = \begin{pmatrix} 2 & 6 & 1 \\ 1 & 6 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad AC_1(-1) = \begin{pmatrix} -6 & 2 & 2 \\ -6 & 1 & 3 \\ -3 & 1 & 1 \end{pmatrix}, \quad AC_{23}(-1) = \begin{pmatrix} 6 & 0 & 2 \\ 6 & -2 & 3 \\ 3 & 0 & 1 \end{pmatrix}$$

Da propia definición obtemos o seguinte resultado.

Proposición 3.2.4. *As matrices de operacións elementais son todas invertíbeis e cúmprese que*

$$F_{ij}^{-1} = F_{ij},\ F_i(\alpha)^{-1} = F_i(\alpha^{-1}),\ F_{ij}(\alpha)^{-1} = F_{ij}(-\alpha),$$

$$C_{ij}^{-1} = C_{ij},\ C_i(\alpha)^{-1} = C_i(\alpha^{-1}),\ C_{ij}(\alpha)^{-1} = C_{ij}(-\alpha).$$

Definición 3.2.5. Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Supoñamos que a fila i da matriz A non ten todos os seus elementos iguais a cero. Chámase entrada principal (ou pivote) da fila i ao primeiro elemento da devandita fila distinto de cero. Se ese elemento é o a_{ij} diremos, ademais, que a fila i ten $j-1$ ceros principais.

Dise que a matriz A está en forma graduada por filas (ou abreviadamente en forma graduada) se se cumpren as seguintes condicións:

1) Se existe algunha fila de ceros está ao final.
2) Se hai varias filas distintas de cero, entón cada fila ten máis ceros principais que a anterior.

Exemplo 3.2.6. Se tomamos as matrices

$$A = \begin{pmatrix} \mathbf{2} & 0 & 0 & 5 \\ 0 & \mathbf{1} & 0 & -2 \\ 0 & 0 & \mathbf{1} & 4 \end{pmatrix};\ B = \begin{pmatrix} \mathbf{1} & 0 & 0 & 5 \\ 0 & \mathbf{1} & 0 & -2 \\ 0 & \mathbf{1} & 1 & 4 \end{pmatrix},\ C = \begin{pmatrix} \mathbf{1} & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 2 \\ 0 & 0 & 0 & \mathbf{1} \end{pmatrix}$$

temos que en groso están resaltadas as entradas principais ou pivotes. A matriz A está en forma graduada, a matriz B non o está, xa que a fila tres ten tantos ceros principais como a segunda. Finalmente, C tampouco está en forma graduada xa que ten unha fila de ceros entre dúas que non son nulas.

Definición 3.2.7. Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Dise que a matriz A está en forma graduada reducida por filas se se cumpren as seguintes condicións:

1) A está en forma graduada por filas.
2) Todas as entradas principais son iguais a 1.
3) Todos os elementos que están por enriba das entradas principais son ceros.

Exemplo 3.2.8. Dada a matriz

$$A = \begin{pmatrix} \mathbf{1} & 0 & 0 & 5 \\ 0 & \mathbf{1} & 0 & -2 \\ 0 & 0 & \mathbf{1} & 4 \end{pmatrix}$$

temos que en groso están resaltadas as entradas principais e a matriz A está en forma graduada reducida.

Definición 3.2.9. Diremos que dúas matrices A, $B \in \mathcal{M}_{m\times n}(\mathbb{K})$ son equivalentes por filas, e denotarémolo como $A \sim_f B$, se se pode pasar dunha á outra mediante unha sucesión de operacións elementais por filas. Isto é, existen matrices $M_1, \cdots, M_r \in \mathcal{M}_{m\times m}(\mathbb{K})$ de operacións elementais tales que

$$B = M_r \cdots M_1 A.$$

Para pasar de A a B seguiremos o seguinte esquema

$$\left(\begin{array}{c|c} A & I_m \end{array} \right) \overset{\text{O.E.F}}{\Longrightarrow} \left(\begin{array}{c|c} B & P \end{array} \right)$$

onde P representa o produto de todas as operacións elementais por filas $P = M_r \cdots M_1$ empregadas para pasar de A a B e, por tanto, $B = PA$. As siglas O.E.F. significan operacións elementais por filas.

Nótese que se $PA = B$ tense que

$$A = P^{-1}B = (M_r \cdots M_1)^{-1}B = M_1^{-1} \cdots M_r^{-1}B$$

e entón $B \sim_f A$. Obviamente, $A \sim_f A$ e ademais, se $A \sim_f B$ e $B \sim_f C$, isto implica que $A \sim_f C$.

Exemplo 3.2.10. Dadas as matrices

$$A = \begin{pmatrix} 4 & 4 & 2 \\ 6 & 1 & -2 \\ 2 & 3 & 2 \end{pmatrix};\ B = \begin{pmatrix} 2 & 3 & 2 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

temos que

$$\left(\begin{array}{c|c} A & I_3 \end{array} \right) \overset{F_{13}}{\Longrightarrow} \left(\begin{array}{ccc|ccc} 2 & 3 & 2 & 0 & 0 & 1 \\ 6 & 1 & -2 & 0 & 1 & 0 \\ 4 & 4 & 2 & 1 & 0 & 0 \end{array} \right) \overset{F_{31}(-2),\, F_{21}(-3)}{\Longrightarrow} \left(\begin{array}{ccc|ccc} 2 & 3 & 2 & 0 & 0 & 1 \\ 0 & -8 & -8 & 0 & 1 & -3 \\ 0 & -2 & -2 & 1 & 0 & -2 \end{array} \right)$$

$$\overset{F_{23}}{\Longrightarrow} \left(\begin{array}{ccc|ccc} 2 & 3 & 2 & 0 & 0 & 1 \\ 0 & -2 & -2 & 1 & 0 & -2 \\ 0 & -8 & -8 & 0 & 1 & -3 \end{array} \right) \overset{F_{32}(-4)}{\Longrightarrow} \left(\begin{array}{ccc|ccc} 2 & 3 & 2 & 0 & 0 & 1 \\ 0 & -2 & -2 & 1 & 0 & -2 \\ 0 & 0 & 0 & -4 & 1 & 5 \end{array} \right)$$

Por tanto, $P = F_{32}(-4)F_{23}F_{31}(-2)F_{21}(-3)F_{13}$ é tal que $PA = B$. Nótese que

$$\begin{aligned} P^{-1} &= (F_{32}(-4)F_{23}F_{31}(-2)F_{21}(-3)F_{13})^{-1} = F_{13}^{-1}F_{21}(-3)^{-1}F_{31}(-2)^{-1}F_{23}^{-1}F_{32}(-4)^{-1} \\ &= F_{13}F_{21}(3)F_{31}(2)F_{23}F_{32}(4). \end{aligned}$$

Proposición 3.2.11. *Cada matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ é equivalente por filas a unha matriz graduada reducida por filas. Isto é, existe unha matriz P cadrada de orde n que é invertíbel e produto de matrices de operacións elementais por filas, tal que PA é graduada reducida por filas.*

Demostración. O método para obter a matriz graduada reducida por filas equivalente por filas á matriz A pódese resumir nos seguintes pasos:

1) Ponse como primeira fila unha que teña o primeiro coeficiente non nulo. Se todos fosen ceros, pasamos ao segundo coeficiente e seguiriamos así de forma sucesiva.
2) Se a entrada principal da primeira fila é a, multiplicámola por a^{-1} para obter unha entrada principal igual a 1.
3) Convertemos en cero os coeficientes que están debaixo do un do paso anterior sumándolle a cada fila a primeira multiplicada polo escalar axeitado.
4) Repetimos os pasos anteriores coas restantes filas ata conseguir unha matriz graduada.
5) Finalmente, usando as operacións elementais convenientes facemos cero todos os coeficientes situados enriba de cada entrada principal. Se existe unha fila de ceros pásase ao final.

A matriz P é o produto de todas as matrices de operacións elementais usadas no proceso.

$\square$

Exemplo 3.2.12. Dada a matriz

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 4 & 2 \\ 3 & 1 & 1 & 3 \end{pmatrix}$$

calculamos unha matriz graduada reducida por filas equivalente por filas a A seguindo o método proposto na proposición anterior.

$$\left(\begin{array}{c|c} A & I_3 \end{array} \right)$$

$$\overset{F_{31}(-3),\, F_{21}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|ccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & -1 & 1 & 0 \\ 0 & -2 & -2 & 0 & -3 & 0 & 1 \end{array} \right) \overset{F_2(1/2)}{\Longrightarrow} \left(\begin{array}{cccc|ccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3/2 & 1/2 & -1/2 & 1/2 & 0 \\ 0 & -2 & -2 & 0 & -3 & 0 & 1 \end{array} \right)$$

$$\overset{F_{32}(2)}{\Longrightarrow} \left(\begin{array}{cccc|ccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3/2 & 1/2 & -1/2 & 1/2 & 0 \\ 0 & 0 & 1 & 1 & -4 & 1 & 1 \end{array} \right)$$

$$\overset{F_{13}(-1),\, F_{23}(-3/2)}{\Longrightarrow} \left(\begin{array}{cccc|ccc} 1 & 1 & 0 & 0 & 5 & -1 & -1 \\ 0 & 1 & 0 & -1 & 11/2 & -1 & -3/2 \\ 0 & 0 & 1 & 1 & -4 & 1 & 1 \end{array} \right)$$

$$\overset{F_{12}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|ccc} 1 & 0 & 0 & 1 & -1/2 & 0 & 1/2 \\ 0 & 1 & 0 & -1 & 11/2 & -1 & -3/2 \\ 0 & 0 & 1 & 1 & -4 & 1 & 1 \end{array} \right)$$

Por tanto, temos que

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = F_{12}(-1)F_{13}(-1)F_{23}(-3/2)F_{32}(2)F_{2}(1/2)F_{31}(-3)F_{21}(-1)A$$

e

$$P = F_{12}(-1)F_{13}(-1)F_{23}(-3/2)F_{32}(2)F_{2}(1/2)F_{31}(-3)F_{21}(-1) = \begin{pmatrix} -1/2 & 0 & 1/2 \\ 11/2 & -1 & -3/2 \\ -4 & 1 & 1 \end{pmatrix}.$$

Proposición 3.2.13. *Sexa A unha matriz cadrada de orde n con coeficientes en $\mathbb{K}$. Entón:*

1) *A matriz A é equivalente por filas a I_n, se, e só se, A é invertíbel.*
2) *Se existe unha matriz X tal que $AX = I_n$ ou $XA = I_n$, a matriz A é invertíbel.*

Demostración. A equivalencia dada en 1) é consecuencia do seguinte: en primeiro lugar, se A é equivalente por filas a I_n, temos que existe unha matriz invertíbel P, produto de matrices de operacións elementais por filas, tal que $PA = I_n$. Entón, $A = P^{-1}$ e, por tanto, A é invertíbel. Inversamente, se A é invertíbel, existe unha matriz A^{-1}, tal que $AA^{-1} = I_n$. Se A fose equivalente por filas a unha matriz graduada reducida B distinta da matriz identidade, esta matriz B tería unha fila de ceros e, entón, existe unha matriz invertíbel P, produto de matrices de operacións elementais por filas, tal que $PA = B$. Así pois, $P = PAA^{-1} = BA^{-1}$. Como B ten polo menos unha fila de ceros, $P = BA^{-1}$ ten unha fila de ceros e isto é un absurdo, xa que P é invertíbel. Por tanto, $A \sim_f I_n$.

Para probar 2) teñamos en conta que 1) garante que, se A non é invertíbel, A é equivalente por filas a unha matriz graduada reducida B distinta da identidade e, por tanto, B ten unha fila de ceros. Se P é a matriz invertíbel, produto de matrices de operacións elementais, tal que $PA = B$, obtemos que $P = PAX = BX$. Como B ten polo menos unha fila de ceros, $P = BX$ ten unha fila de ceros e isto é un absurdo, xa que P é invertíbel. Finalmente, a propiedade para o produto $XA = I_n$ obtense traballando coa trasposta. . □

Exemplo 3.2.14. Supoñamos que temos unha matriz cadrada A de orde n coa seguinte descomposición por bloques

$$A = \left(\begin{array}{c|c} H & \Theta_{r,n-r} \\ \hline C & D \end{array}\right).$$

Logo, A é invertíbel, se, e só se, o son H e D. Ademais cúmprese que, no caso de ter inversa, esta é

$$A^{-1} = \left(\begin{array}{c|c} H^{-1} & \Theta_{r,n-r} \\ \hline -D^{-1}CH^{-1} & D^{-1} \end{array}\right).$$

En efecto, se H e D teñen inversa, multiplicando por bloques temos que

$$\left(\begin{array}{c|c} H & \Theta_{r,n-r} \\ \hline C & D \end{array}\right)\left(\begin{array}{c|c} H^{-1} & \Theta_{r,n-r} \\ \hline -D^{-1}CH^{-1} & D^{-1} \end{array}\right)$$

$$= \left(\begin{array}{c|c} HH^{-1} & \Theta_{r,n-r} \\ \hline CH^{-1} - DD^{-1}CH^{-1} & DD^{-1} \end{array}\right) = \left(\begin{array}{c|c} I_r & \Theta_{r,n-r} \\ \hline \Theta_{n-r,r} & I_{n-r} \end{array}\right) = I_n$$

Daquela, atopamos unha matriz que multiplicada por A dá a identidade. Aplicando agora a propiedade 2) do anterior resultado obtemos que A é invertíbel.

Inversamente, se A ten inversa existe unha matriz A^{-1} tal que $AA^{-1} = I_n$. Supoñamos que descompoñemos ambas as dúas por bloques e facemos o seu produto. Logo

$$\left(\begin{array}{c|c} H & \Theta_{r,n-r} \\ -- & --- \\ C & D \end{array}\right)\left(\begin{array}{c|c} X_{11} & X_{12} \\ -- & --- \\ X_{21} & X_{22} \end{array}\right) = \left(\begin{array}{c|c} I_r & \Theta_{r,n-r} \\ --- & --- \\ \Theta_{n-r,r} & I_{n-r} \end{array}\right) = I_n.$$

Por tanto,

$$HX_{11} = I_r, \;\; HX_{12} = \Theta_{r,n-r}, \;\; CX_{11} + DX_{21} = \Theta_{n-r,r}, \;\; CX_{12} + DX_{22} = I_{n-r}.$$

Disto deducimos, outra vez polo segundo apartado do anterior resultado, que H é invertíbel, o que implica que $X_{12} = \Theta_{r,n-r}$ e que $X_{11} = H^{-1}$. Disto séguese que $DX_{22} = I_{n-r}$, o que implica tamén que D é invertíbel. Finalmente, obsérvese que

$$CX_{11} + DX_{21} = \Theta_{n-r,r} \Leftrightarrow CH^{-1} = -DX_{21} \Leftrightarrow X_{21} = -D^{-1}CH^{-1}.$$

Proposición 3.2.15. *Sexa $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ tal que ten unha forma graduada reducida por filas B con r entradas principais. Entón, existen matrices $P \in \mathcal{M}_{m\times m}(\mathbb{K})$ e $Q \in \mathcal{M}_{n\times n}(\mathbb{K})$ tales que*

1) *P e Q son produto de matrices de operacións elementais por filas e por columnas respectivamente.*
2) *Cúmprese que*

$$PAQ = \left(\begin{array}{c|c} I_r & \Theta_{r,n-r} \\ --- & ----- \\ \Theta_{m-r,r} & \Theta_{m-r,n-r} \end{array}\right).$$

Demostración. Mediante o proceso dado na Proposición 3.2.11 obtemos unha matriz invertíbel $P \in \mathcal{M}_{m\times m}(\mathbb{K})$, produto de matrices de operacións elementais por filas, tal que $PA = B$ sendo B unha matriz graduada reducida por filas con r entradas principais. En cada fila con entrada principal, facendo operacións elementais coas columnas, podemos converter en ceros todos os elementos non nulos distintos da entrada principal ou pivote. Unha vez feito isto reordenamos as columnas para obter a matriz I_r. A matriz Q será o produto de todas as matrices de operacións elementais por columnas que usamos no proceso. Daquela, Q é invertíbel e

$$PAQ = \left(\begin{array}{c|c} I_r & \Theta_{r,n-r} \\ --- & ----- \\ \Theta_{m-r,r} & \Theta_{m-r,n-r} \end{array}\right).$$

□

Exemplo 3.2.16. Dada a matriz

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 4 & 2 \\ 1 & 1 & 1 & 3 \end{pmatrix}$$

temos que

$$\left(\begin{array}{c|c} A & I_3 \end{array}\right)$$

$$\stackrel{F_{31}(-1),\, F_{21}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|ccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 2 & -1 & 0 & 1 \end{array}\right)$$

$$\stackrel{F_{3}(1/2),\, F_{2}(1/3)}{\Longrightarrow} \left(\begin{array}{cccc|ccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/3 & -1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 1 & -1/2 & 0 & 1/2 \end{array}\right)$$

$$\stackrel{F_{13}(-1),\, F_{23}(-1/3)}{\Longrightarrow} \left(\begin{array}{cccc|ccc} 1 & 1 & 1 & 0 & 3/2 & 0 & -1/2 \\ 0 & 0 & 1 & 0 & -1/6 & 1/3 & -1/6 \\ 0 & 0 & 0 & 1 & -1/2 & 0 & 1/2 \end{array}\right)$$

$$\stackrel{F_{12}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|ccc} 1 & 1 & 0 & 0 & 5/3 & -1/3 & -1/3 \\ 0 & 0 & 1 & 0 & -1/6 & 1/3 & -1/6 \\ 0 & 0 & 0 & 1 & -1/2 & 0 & 1/2 \end{array}\right)$$

Logo, tomando

$$P = F_{12}(-1)F_{13}(-1)F_{23}(-1/3)F_3(1/2)F_2(1/3)F_{31}(-1)F_{21}(-1)$$

temos que PA é graduada reducida por filas e $PA = B$ onde

$$B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Fagamos a continuación as operacións por columnas. Neste caso partimos de

$$\begin{pmatrix} B \\ - \\ I_4 \end{pmatrix}$$

$$\stackrel{C_{21}(-1)}{\Longrightarrow} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ - & - & - & - \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \stackrel{C_{23},\, C_{34}}{\Longrightarrow} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ - & - & - & - \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Por tanto,

$$Q = C_{21}(-1)C_{23}C_{34} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

e temos que

$$PAQ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \left(\begin{array}{c|c} I_3 & \Theta_{3,1} \end{array} \right).$$

Notación 3.2.17. A partir de agora denotaremos as matrices

$$\left(\begin{array}{c|c} I_r & \Theta_{r,n-r} \\ --- & ----- \\ \Theta_{m-r,r} & \Theta_{m-r,n-r} \end{array}\right)$$

como $I_r^{m\times n}$.

Definición 3.2.18. Sexan $A,\ B \in \mathcal{M}_{m\times n}(\mathbb{K})$. Diremos que A e B son matrices equivalentes se existen matrices $P \in \mathcal{M}_{m\times m}(\mathbb{K})$ e $Q \in \mathcal{M}_{n\times n}(\mathbb{K})$ tales que:

1) P e Q son produto de matrices de operacións elementais por filas e por columnas respectivamente.
2) Cúmprese que $PAQ = B$.

Cando dúas matrices A e B sexan equivalentes denotarémolo por $A \sim B$. Obviamente se dúas matrices son equivalentes por filas son tamén equivalentes sen máis que tomar $Q = I_n$. Doutra banda, $A \sim A$ e, se ademais $A \sim B$, entón $B \sim A$. Finalmente, se $A \sim B$ e $B \sim C$ isto implica que $A \sim C$.

Proposición 3.2.19. *Toda matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ é equivalente a unha única matriz da forma $I_r^{m\times n}$.*

Demostración. Grazas á Proposición 3.2.15 sabemos que existen matrices $P \in \mathcal{M}_{m\times m}(\mathbb{K})$ e $Q \in \mathcal{M}_{n\times n}(\mathbb{K})$ tales que:

- P e Q son produto de matrices de operacións elementais por filas e por columnas respectivamente.
- Verifícase que $PAQ = I_r^{m\times n}$.

Entón,

$$A \sim I_r^{m\times n}.$$

Se existise un natural $s > r$ tal que $A \sim I_s^{m\times n}$, obteriamos que $I_r^{m\times n} \sim I_s^{m\times n}$. Por tanto, existen matrices $T \in \mathcal{M}_{m\times m}(\mathbb{K})$ e $H \in \mathcal{M}_{n\times n}(\mathbb{K})$ tales que:

- T e H son produto de matrices de operacións elementais por filas e por columnas respectivamente.
- Verifícase que $TI_r^{m\times n}H = I_s^{m\times n}$ e, o que é o mesmo, $TI_r^{m\times n} = I_s^{m\times n}R$, sendo $R = H^{-1}$.

Descompoñendo convenientemente por bloques temos que

$$I_s^{m\times n}R = \left(\begin{array}{c|c|c} I_r & \Theta_{r,s-r} & \Theta_{r,n-s} \\ \hline \Theta_{s-r,r} & I_{s-r} & \Theta_{s-r,n-s} \\ \hline \Theta_{n-s,r} & \Theta_{n-s,s-r} & \Theta_{n-s,n-s} \end{array}\right)\left(\begin{array}{c|c|c} R_{11} & R_{12} & R_{13} \\ \hline R_{21} & R_{22} & R_{23} \\ \hline R_{31} & R_{32} & R_{33} \end{array}\right)$$

$$= \left(\begin{array}{c|c|c} R_{11} & R_{12} & R_{13} \\ \hline R_{21} & R_{22} & R_{23} \\ \hline \Theta_{n-s,r} & \Theta_{n-s,s-r} & \Theta_{n-s,n-s} \end{array}\right)$$

e que

$$TI_r^{m\times n} = \left(\begin{array}{c|c} T_{11} & T_{12} \\ \hline T_{21} & T_{22} \end{array}\right)\left(\begin{array}{c|c} I_r & \Theta_{r,n-r} \\ \hline \Theta_{n-r,r} & \Theta_{n-r,n-r} \end{array}\right) = \left(\begin{array}{c|c} T_{11} & \Theta_{r,n-r} \\ \hline T_{12} & \Theta_{n-r,n-r} \end{array}\right).$$

Por tanto,

$$R = \left(\begin{array}{c|c|c} R_{11} & \Theta_{r,s-r} & \Theta_{r,n-s} \\ \hline R_{12} & \Theta_{s-r,s-r} & \Theta_{s-r,n-s} \\ \hline R_{31} & R_{32} & R_{33} \end{array}\right)$$

e entón, aplicando o visto no exemplo 3.2.14, temos que R non é invertíbel. Isto é un absurdo e, daquela, a matriz

$$I_s^{m\times n}$$

é única. □

Definición 3.2.20. Dada unha matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ chamaremos forma reducida completa (forma normal) de A á única matriz $I_r^{m\times n}$ que é equivalente con A.

Chámase rango de A e represéntase como

$$\text{rang}\,(A)$$

ao número r asociado a $I_r^{m\times n}$. Este número é un invariante da matriz A xa que $I_r^{m\times n}$ é única para cada A.

Obsérvese que ese número r coincide co número de filas non nulas da forma graduada reducida de A a partir da cal se calcula $I_r^{m\times n}$ ou, equivalentemente, co número de entradas principais desa forma. Daquela, se existisen dúas formas graduadas reducidas de A distintas terían o mesmo número de entradas principais. De feito, pódese probar, mediante unha demostración que omitiremos, que a forma graduada reducida de A é única e utilizar esta propiedade para definir o rango.

Finalmente, é evidente que

$$\text{rang}\,(A) \leq min\{m, n\}.$$

Teorema 3.2.21. *Dúas matrices $A, B \in \mathcal{M}_{m\times n}(\mathbb{K})$ son equivalentes, se, e só se, teñen o mesmo rango.*

Demostración. Se A e B son equivalentes, as súas formas reducidas completas tamén o son e, por tanto, son iguais. Isto implica que o rango de A coincide co de B. Ao revés, se teñen o mesmo rango, teñen a mesma forma reducida completa e isto implica que son equivalentes.

□

Como consecuencia inmediata deste teorema obtemos o corolario seguinte:

Corolario 3.2.22. *Dada unha matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ cúmprese que:*

1) *Se $P \in \mathcal{M}_{m\times m}(\mathbb{K})$ é unha matriz invertíbel,* $\text{rang}\,(PA) = \text{rang}\,(A)$.
2) *Se $Q \in \mathcal{M}_{n\times n}(\mathbb{K})$ é unha matriz invertíbel,* $\text{rang}\,(AQ) = \text{rang}\,(A)$.
3) $\text{rang}\,(A) = \text{rang}\,(A^t)$.

3.3. Matrices invertíbeis. Cálculo da matriz inversa

Grazas aos resultados probados na sección anterior podemos obter un teorema que caracteriza totalmente ás matrices invertíbeis. A demostración das equivalencias é sinxela e déixase como exercicio.

Teorema 3.3.1. *Sexa A unha matriz cadrada de orden n. Son equivalentes:*

1) *A é invertíbel.*
2) *$A \sim_f I_n$.*
3) *$A \sim I_n$.*
4) *A forma graduada reducida de A ten n entradas principais.*
5) *O rango de A é n.*
6) *A matriz A é produto de matrices de operacións elementais por filas.*
7) *A matriz A é produto de matrices de operacións elementais por columnas.*
8) *A matriz A é produto de matrices de operacións elementais por filas e columnas.*
9) *Existe unha matriz invertíbel P tal que $PA = I_n$.*
10) *Existe unha matriz invertíbel Q tal que $AQ = I_n$.*

Tendo en conta que toda matriz cadrada A invertíbel é equivalente por filas á matriz identidade, para calcular a inversa de A, calcularemos a matriz P, tal que $PA = I_n$ e, por tanto, $A^{-1} = P$. Neste proceso só utilizamos operacións por filas. Vexamos un exemplo.

Exemplo 3.3.2. Sexa a matriz

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

Logo:

$$\left(A \mid I_4 \right)$$

$$\overset{F_{41}(-1),\, F_{31}(-1),\, F_{21}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 3 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\overset{F_{42}(-1),\, F_{32}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & -1 & 0 & 1 \end{array}\right)$$

$$\overset{F_{43}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 1 \end{array}\right)$$

$$\overset{F_{14}(-1),\, F_{24}(-1),\, F_{34}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 & -1 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 1 \end{array}\right)$$

$$\overset{F_{13}(-1),\, F_{23}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 1 \end{array}\right)$$

$$\overset{F_{12}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 1 \end{array}\right)$$

Daquela,

$$A^{-1} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} =$$

$$F_{12}(-1)F_{13}(-1)F_{23}(-1)F_{14}(-1)F_{24}(-1)F_{34}(-1)$$
$$F_{43}(-1)F_{42}(-1)F_{32}(-1)F_{41}(-1)F_{31}(-1)F_{21}(-1).$$

Obsérvese que tamén podemos combinar operacións con filas con operacións con columnas. Grazas aos cálculos anteriores sabemos que existe unha matriz

$$P = F_{43}(-1)F_{42}(-1)F_{32}(-1)F_{41}(-1)F_{31}(-1)F_{21}(-1),$$

tal que

$$PA = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Fagamos a continuación operacións por columnas para converter a matriz PA na matriz identidade.

$$\begin{pmatrix} PA \\ - \\ I_4 \end{pmatrix}$$

$$\xrightarrow{C_{21}(-1),\, C_{31}(-1),\, C_{41}(-1)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ - & - & - & - \\ 1 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_{32}(-1),\, C_{42}(-1)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ - & - & - & - \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{C_{43}(-1)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ - & - & - & - \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Por tanto, $PAQ = I_n$, onde

$$Q = C_{21}(-1)C_{31}(-1)C_{41}(-1)C_{32}(-1)C_{42}(-1)C_{43}(-1) = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

e, ademais,

$$A^{-1} = (P^{-1}Q^{-1})^{-1} = QP.$$

Nota 3.3.3. É moi importante que, no caso de usar operacións por filas e operacións por columnas, non as mesturemos. Para as primeiras utilizaremos a matriz P e para as segundas, a matriz Q.

3.4. Determinante dunha matriz cadrada. Propiedades e cálculo

Os determinantes introducirémolos mediante un proceso indutivo da forma seguinte:

Definición 3.4.1. Para unha matriz cadrada $A = (a_{11})$ de orde 1 defínese o determinante de A, denotado por $\det(A)$ ou $|A|$, como

$$\det(A) = a_{11}.$$

Coñecido o valor do determinante de calquera matriz cadrada de orde $(n-1)$, para unha matriz cadrada de orde n

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

chámase cofactor do coeficiente a_{ij}, denotado por A_{ij}, a

$$A_{ij} = (-1)^{i+j} \det(M_{ij})$$

sendo M_{ij} a matriz matriz cadrada de orde $n-1$ que resulta ao suprimir en A a fila i-ésima e a columna j-ésima.

O determinante de A defínese como

$$\det(A) = a_{11}A_{11} + \cdots + a_{n1}A_{1n}.$$

Esta definición que acabamos de dar de determinante coñécese como desenvolvemento de Laplace do determinante pola primeira columna da matriz A. Como consecuencia inmediata desta definición tense que

$$\det(I_n) = 1.$$

Exemplo 3.4.2. Para unha matriz $A = (a_{ij})$ cadrada de orde dous temos que

$$\det(A) = det \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}A_{11} + a_{21}A_{21} = a_{11}a_{22} - a_{12}a_{21}.$$

Desta forma, como sabemos calcular determinantes de matrices cadradas de orde dous, podemos calcular determinantes de matrices cadradas de orde tres do seguinte xeito:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31} =$$

$$a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} =$$

$$a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}) =$$

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

obtendo a regra de Sarrus. A continuación calculariamos determinantes de orde catro e así, sucesivamente. Sexa, por exemplo,

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 0 & 2 & 3 & 4 \end{pmatrix}.$$

Entón,

$$\det(A) = 0.A_{11} + 1.A_{21} + 1.A_{31} + 0.A_{41} = (-1)^{1+2}\det(M_{21}) + (-1)^{1+3}\det(M_{31}) =$$

$$-\begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 2 & 3 & 4 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 2 & 3 & 4 \end{vmatrix} =$$

$$-(12 + 6 + 6 - 6 - 9 - 8) + (8 + 6 + 4 - 4 - 6 - 8) = -1.$$

A fórmula do desenvolvemento de Laplace descrita na definición pódese optimizar se probamos algunhas propiedades interesantes dos determinantes. Enunciarémolas e probarémolas nas proposicións seguintes nas cales a fila i dunha matriz cadrada A denotarase como

$$A_i = (a_{i1}\ a_{i2}\ \cdots\ a_{in}).$$

Proposición 3.4.3. *Sexan tres matrices cadradas da mesma orde, A, A' e A'', tales que teñen todas as filas iguais, salvo a fila i. Se a fila i de A é a suma da filas i de A' e A'', entón*

$$\det(A) = \det(A') + \det(A'').$$

Demostración. A propiedade probarémola por indución. Para $n = 1$ é evidente. Supoñámola certa para $n-1$ e probemos a propiedade para n. Temos que

$$\det(A) = \det(A) = a_{11}A_{11} + \cdots + a_{i1}A_{i1} + \cdots + a_{n1}A_{n1} =$$
$$a_{11}(-1)^{1+1}\det(M_{11}) + \cdots + a_{i1}(-1)^{1+i}\det(M_{i1}) + \cdots + a_{1n}(-1)^{1+n}\det(M_{1n}).$$

Para $k \neq i$ temos que M_{k1}, M'_{k1} e M''_{k1} son tales que teñen todas as filas iguais salvo unha fila. Esa fila distinta de M_{k1} é a suma das respectivas filas de M'_{k1} y M''_{k1}, entón

$$\det(M_{k1}) = \det(M'_{k1}) + \det(M''_{k1})$$

por hipótese de indución. Desta forma,

$$\det(A) = a_{11}(-1)^{1+1}\det(M_{11}) + \cdots + a_{i1}(-1)^{1+i}\det(M_{i1}) + \cdots + a_{1n}(-1)^{1+n}\det(M_{1n}) =$$
$$a_{11}(-1)^{1+1}(\det(M'_{11}) + \det(M''_{11})) + \cdots + (a'_{i1} + a''_{i1})(-1)^{1+i}\det(M_{i1}) + \cdots$$
$$\cdots + a_{n1}(-1)^{1+n}(\det(M'_{n1}) + \det(M''_{n1})) =$$
$$(a_{11}(-1)^{1+1}\det(M'_{11}) + \cdots + a'_{i1}(-1)^{1+i}\det(M_{i1}) + \cdots + a_{n1}(-1)^{1+n}\det(M'_{n1}))$$
$$+(a_{11}(-1)^{1+1}\det(M''_{11}) + \cdots + a''_{i1}(-1)^{1+i}\det(M_{i1}) + \cdots + a_{1n}(-1)^{1+n}\det(M''_{1n})) =$$
$$\det(A') + \det(A'').$$

□

Proposición 3.4.4. *Se unha matriz cadrada ten dúas filas consecutivas iguais, o seu determinante vale cero. Como consecuencia disto, se se intercambian dúas filas dunha matriz cadrada A, o seu determinante cambia de signo.*

Demostración. Vexamos en primeiro lugar que, se unha matriz cadrada ten dúas filas consecutivas iguais, o seu determinante vale cero. Farémolo outra vez por indución sobre o número de filas n. Para $n = 2$ é evidente. Supoñamos que é certo para $n-1$ e probemos a propiedade para n. Se as filas consecutivas son a fila i e a fila $i+1$, temos que:

$$\det(A) = a_{11}A_{11} + \cdots + a_{i1}A_{i1} + a_{i1}A_{i+1\,1} + \cdots + a_{n1}A_{n1}.$$

Agora ben, para $k \neq i,\ i+1$ cúmprese a igualdade $A_{k1} = (-1)^{1+k}\det(M_{k1}) = 0$ xa que M_{k1} ten dúas filas iguais e é cadrada de orde $n-1$. Doutra banda,

$$A_{i1} = (-1)^{1+i}\det(M_{i1}), \quad A_{i+1,1} = (-1)^{2+i}\det(M_{i+1,1})$$

e, como $M_{i1} = M_{i+1,1}$, obtense que $A_{i1} = -A_{i+1,1}$ e isto implica que $\det(A) = 0$.

Utilizando esta propiedade que acabamos de probar xunto coa Proposición 3.4.3, temos que

$$0 = \left|\begin{pmatrix} A_1 \\ \cdots \\ A_i + A_{i+1} \\ A_i + A_{i+1} \\ \cdots \\ A_n \end{pmatrix}\right| = \left|\begin{pmatrix} A_1 \\ \cdots \\ A_i \\ A_i + A_{i+1} \\ \cdots \\ A_n \end{pmatrix}\right| + \left|\begin{pmatrix} A_1 \\ \cdots \\ A_{i+1} \\ A_i + A_{i+1} \\ \cdots \\ A_n \end{pmatrix}\right| =$$

$$\left|\begin{pmatrix} A_1 \\ \cdots \\ A_i \\ A_i \\ \cdots \\ A_n \end{pmatrix}\right| + \left|\begin{pmatrix} A_1 \\ \cdots \\ A_i \\ A_{i+1} \\ \cdots \\ A_n \end{pmatrix}\right| + \left|\begin{pmatrix} A_1 \\ \cdots \\ A_{i+1} \\ A_i \\ \cdots \\ A_n \end{pmatrix}\right| + \left|\begin{pmatrix} A_1 \\ \cdots \\ A_{i+1} \\ A_{i+1} \\ \cdots \\ A_n \end{pmatrix}\right| =$$

$$\left|\begin{pmatrix} A_1 \\ \dots \\ A_i \\ A_{i+1} \\ \dots \\ A_n \end{pmatrix}\right| + \left|\begin{pmatrix} A_1 \\ \dots \\ A_{i+1} \\ A_i \\ \dots \\ A_n \end{pmatrix}\right|.$$

Por tanto,

$$\left|\begin{pmatrix} A_1 \\ \dots \\ A_i \\ A_{i+1} \\ \dots \\ A_n \end{pmatrix}\right| = -\left|\begin{pmatrix} A_1 \\ \dots \\ A_{i+1} \\ A_i \\ \dots \\ A_n \end{pmatrix}\right|.$$

□

Proposición 3.4.5. *Se unha matriz cadrada ten dúas filas iguais, o seu determinante vale cero.*

Demostración. Este resultado é consecuencia trivial do anterior.

□

Proposición 3.4.6. *Se unha matriz cadrada ten unha fila de ceros o seu determinante é nulo.*

Demostración. Como en casos anteriores, o resultado próbase por indución. Para $n = 1$ é evidente. Supoñamos que é certo para $n-1$ e probemos a propiedade para n. Se a fila nula é a fila i temos que

$$\det(A) = a_{11}A_{11} + \dots + 0.A_{i1} + \dots + a_{n1}A_{n1}.$$

Agora ben, para $k \neq i$ cúmprese a igualdade $A_{k1} = (-1)^{1+k}\det(M_{k1}) = 0$, xa que M_{k1} ten unha fila de ceros e é cadrada de orde $n-1$. Isto implica que $\det(A) = 0$. □

Proposición 3.4.7. *Se se multiplica unha fila dunha matriz cadrada A por un escalar, o determinante queda multiplicado por ese escalar.*

Demostración. Outra vez utilizamos o método de indución. Para $n = 1$ é evidente. Supoñamos que é certo para $n-1$ e probemos o resultado para n. Se a fila multiplicada polo escalar α é a fila i e A' é a matriz resultante tras esa operación, temos que:

$$\det(A') = a_{11}A'_{11} + \dots + \alpha a_{i1}A'_{i1} + \dots + a_{n1}A'_{n1}.$$

Entón, para $k \neq i$, cúmprese a igualdade $A'_{k1} = \alpha A_{k1}$ por hipótese de indución e, ademais, $A'_{i1} = A_{i1}$. Por tanto,

$$\det(A') = \alpha a_{11}A_{11} + \dots + \alpha a_{i1}A_{i1} + \dots + \alpha a_{n1}A_{n1} = \alpha\det(A).$$

□

Proposición 3.4.8. *Se a unha fila dunha matriz cadrada sumámoslle un múltiplo doutra fila, o valor do determinante non varía.*

Demostración. Aplicando que unha matriz con dúas filas iguais ten determinante nulo e as Proposicións 3.4.3 e 3.4.7, tense que

$$\left|\begin{pmatrix} A_1 \\ \cdots \\ A_i \\ \cdots \\ A_j + \alpha A_i \\ \cdots \\ A_n \end{pmatrix}\right| = \left|\begin{pmatrix} A_1 \\ \cdots \\ A_i \\ \cdots \\ A_j \\ \cdots \\ A_n \end{pmatrix}\right| + \alpha \left|\begin{pmatrix} A_1 \\ \cdots \\ A_i \\ \cdots \\ A_i \\ \cdots \\ A_n \end{pmatrix}\right| = \det(A).$$

□

Nota 3.4.9. Dos resultados anteriores séguese de forma trivial que

$$\det(F_{ij}) = \det(F_{ij}^t) = -1, \quad \det(F_i(\alpha)) = \det(F_i(\alpha)^t) = \alpha, \quad \det(F_{ij}(\alpha)) = \det(F_{ij}(\alpha)^t) = 1.$$

Ademais, para toda matriz cadrada B de orde n tense que

$$\det(F_{ij}B) = -\det(B) = \det(F_{ij})\det(B),$$
$$\det(F_i(\alpha)B) = \alpha\det(B) = \det(F_i(\alpha))\det(B),$$

e

$$\det(F_{ij}(\alpha)B) = \det(B) = \det(F_{ij}(\alpha))\det(B).$$

Proposición 3.4.10. *Unha matriz A é invertíbel, se, e só se,* $\det(A) \neq 0$.

Demostración. Toda matriz A invertíbel é produto de matrices de operacións elementais por filas

$$F_r F_{r-1} \cdots F_1 = A.$$

Polo visto na nota anterior, temos que, se B é unha matriz cadrada de orde n, entón $\det(FB) = \det(F)\det(B)$ para toda matriz de operacións elementais por filas F. Como consecuencia,

$$\det(A) = \det(F_r)\det(F_{r-1}\cdots F_1) = \det(F_r)\det(F_{r-1})\det(F_{r-2}\cdots F_1) = \cdots =$$
$$\det(F_r)\det(F_{r-1})\det(F_{r-2})\cdots\det(F_1) \neq 0.$$

Supoñamos que A non é invertíbel. Entón é equivalente por filas a unha matriz cadrada de orde n graduada reducida por filas con algunha delas nula, xa que o rango de A ten que ser menor que n. Sexa H_A a forma graduada reducida con polo menos unha fila nula. Temos que $A \sim_f H_A$ e isto implica que existen matrices de operacións elementais por filas, tales que

$$F_s F_{s-1} \cdots F_1 A = H_A$$

Isto implica que

$$\det(F_s)\det(F_{s-1})\cdots\det(F_1)\det(A) = \det(H_A) = 0$$

e, por tanto, o determinante de A é nulo.

□

Proposición 3.4.11. *Se A e B son dúas matrices cadradas, tense que*

$$\det(AB) = \det(A)\det(B).$$

Como consecuencia, se A é invertíbel

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Demostración. Se $\det(A)$ ou $\det(B)$ son nulos, entón $\det(AB)$ tamén o é. Isto é así xa que, se existe $(AB)^{-1}$, tense que

$$AB(AB)^{-1} = I_n, \quad (AB)^{-1}AB = I_n$$

o que implica que A e B son invertíbeis, o cal é un absurdo. Se A é invertíbel, é produto de matrices de operacións elementais por filas

$$F_r F_{r-1} \cdots F_1 = A.$$

Daquela,

$$\det(F_r)\det(F_{r-1})\det(F_{r-2})\cdots\det(F_1) = \det(A)$$

e

$$\det(AB) = \det(F_r)\det(F_{r-1}\cdots F_1 B) = \det(F_r)\det(F_{r-1})\det(F_{r-2}\cdots F_1 B) = \cdots =$$
$$\det(F_r)\det(F_{r-1})\det(F_{r-2})\cdots\det(F_1)\det(B) = \det(A)\det(B).$$

A igualdade

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

é consecuencia de que $A^{-1}A = I_n$. □

Proposición 3.4.12. *O determinante dunha matriz cadrada coincide co da súa trasposta.*

Demostración. Se A non é invertíbel, entón tampouco o é A^t e, entón, o resultado é certo xa que $\det(A) = \det(A^t) = 0$. Se A é invertíbel, é produto de matrices de operacións elementais por filas.

$$F_r F_{r-1} \cdots F_1 = A.$$

Daquela,

$$\det(A^t) = \det(F_1^t)\cdots\det(F_{r-2}^t)\det(F_{r-1}^t)\det(F_r^t) =$$
$$\det(F_1)\cdots\det(F_{r-2})\det(F_{r-1})\det(F_r) =$$
$$\det(F_r)\det(F_{r-1})\det(F_{r-2})\cdots\det(F_1) = \det(A).$$

□

Utilizando a propiedade de que $\det(A) = \det(A^t)$, podemos demostrar de forma sinxela que todos os resultados anteriores que involucran filas tamén as temos para columnas.

Corolario 3.4.13. *Cúmprese o seguinte:*

1) *Sexan tres matrices cadradas da mesma orde, A, A' e A'', tales que teñen todas as columnas iguais, salvo a columna j. Se a columna j de A é a suma das columnas j de A' e A'', entón*
$$\det(A) = \det(A') + \det(A'').$$
2) *Se unha matriz cadrada A ten dúas columnas iguais, o seu determinante vale cero.*
3) *Se se intercambian dúas columnas dunha matriz cadrada A, o seu determinante cambia de signo.*
4) *Se unha matriz cadrada A ten unha columna de ceros, o seu determinante é nulo.*
5) *Se se multiplica unha columna dunha matriz cadrada A por un escalar, o determinante queda multiplicado por ese escalar.*
6) *Se a unha columna dunha matriz cadrada A sumámoslle un múltiplo doutra columna, o valor do determinante non varía.*

Proposición 3.4.14. *O determinante dunha matriz cadrada pódese desenvolver por calquera das súas filas ou columnas.*

Demostración. A demostración é unha consecuencia das proposicións e corolarios anteriores.

$\square$

Grazas ás propiedades anteriores podemos realizar os cálculos de determinantes utilizando as operacións elementais. Vexamos un exemplo:

Exemplo 3.4.15. Sexa

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 0 & 2 & 3 & 4 \end{pmatrix}.$$

Entón, restando a ultima columna á terceira e á segunda

$$\det(A) = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 1 & -1 & 0 & 3 \\ 0 & -2 & -1 & 4 \end{vmatrix}.$$

Desenvolvendo o determinante pola primeira fila, temos que:

$$\begin{vmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 1 & -1 & 0 & 3 \\ 0 & -2 & -1 & 4 \end{vmatrix} = (-1)^{1+4} \begin{vmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & -2 & -1 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & -2 & -1 \end{vmatrix}.$$

Volvendo desenvolver pola primeira fila, obtéñense as igualdades

$$- \begin{vmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & -2 & -1 \end{vmatrix} = -(-1)^{1+1} \begin{vmatrix} -1 & 0 \\ -2 & -1 \end{vmatrix} = -1.$$

Como o determinante de A é distinto de cero a matriz é invertíbel e

$$\det(A^{-1}) = \frac{1}{-1} = -1.$$

3.5. Problemas propostos

1. Considéranse as matrices reais

$$A = \begin{pmatrix} -1 & 2 & 3 \\ 3 & -4 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & -2 \\ -1 & 1 \\ 3 & 0 \end{pmatrix},$$

$$D = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad E = (\ 1\ ,\ 2\ ,\ 3\ ), \quad F = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}.$$

Calcule as seguintes matrices:

a) $3A + 2B$.	*b*) $B - A$.
c) AC e CA.	*d*) ED, DF, EF e FE.
e) CBD e EDE^t.	*f*) $D - (4CB)^t$.
g) $AC + (AC)^t$.	*h*) $(A + B)D$.
i) D^2 e D^3.	*j*) $FEDC$.

2. Sexa $A = (a_{ij}) \in \mathcal{M}_{4\times 4}(\mathbb{R})$ a matriz definida por $a_{ij} = |i-j|$ para $1 \le i, j \le 4$ e sexa $B = (b_{ij}) \in \mathcal{M}_{4\times 4}(\mathbb{R})$ onde
$$b_{ij} = \begin{cases} ij & \text{se } i \neq j \\ 0 & \text{se } i = j. \end{cases}$$
Calcule os produtos AB e BA.
3. Sexa $A = (a_{ij}) \in \mathcal{M}_{3\times 3}(\mathbb{R})$ unha matriz estritamente triangular superior tal que $a_{12} \neq 0$ e consideremos a matriz
$$B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & -a_{13}/a_{12} & 0 \\ 0 & 1 & 0 \end{pmatrix} \in \mathcal{M}_{3\times 3}(\mathbb{R}).$$
Calcule a matriz AB. É AB diagonal? É AB escalar?
4. Considéranse as matrices complexas
$$A = \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 8 & 2-3i \\ 2+3i & 0 \end{pmatrix}, \quad C = \begin{pmatrix} i \,,\, 1 \end{pmatrix},$$
$$D = \begin{pmatrix} -2i & 1+i \\ -4 & 2-i \end{pmatrix}, \quad E = \begin{pmatrix} -2 & 2+3i \\ 1 & -i \\ i & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 2i \\ i \end{pmatrix}.$$
Calcule as matrices:
a) $(1+i)B$. *b*) A^*, B^* e E^*.
c) CC^t e CC^* . *d*) $A+B+D$, $FC+D$ e $EB+E$.
e) $ABDF$. *f*) $((1-3i)D)^*$ e $(1+3i)D^*$.
g) $(FC)^*$ e $(FC)^t$. *h*) EAE^t e EAE^*.
i) $A^2 - B^2$ e $(A-B)(A+B)$. *j*) $(A+D)^2$ e $A^2 + 2AD + D^2$.
5. Sexan as matrices
$$A = \begin{pmatrix} i & 1 & 1-2i \\ 2 & -1 & -3-i \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 5 & 1-2i & -1 \\ -2 & 0 & -3+i \end{pmatrix}.$$
a) Determine a matriz X nas seguintes ecuacións:
*a*1) $A + X = B$.
*a*2) $A - iX = 2B$.
*a*3) $(1+i)X^* - A = 0$.
*a*4) $A^t + 3X = 0$.
b) Atope as solucións $X, Y \in \mathcal{M}_{2\times 3}(\mathbb{C})$ dos seguintes sistemas de ecuacións:
*b*1) $\left.\begin{array}{l} 3X + Y = B \\ 5X + 2Y = -A \end{array}\right\}$
*b*2) $\left.\begin{array}{l} iX - Y = -B \\ (1-i)X + Y = -(1+i)A \end{array}\right\}$
6. Sexan $A \in \mathcal{M}_{2\times 3}(\mathbb{R}), B \in \mathcal{M}_{3\times 2}(\mathbb{R})$ e $C \in \mathcal{M}_{3\times 2}(\mathbb{R})$ as matrices
$$A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & -2 \\ 3 & -3 \\ -1 & 1 \end{pmatrix}.$$
Comprobe que $AB = AC = I_2$. É A invertíbel?
7. Probe que, se $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ verifican $B \neq \Theta_{n,n}$ e $AB = \Theta_{n,n}$, entón A é singular.
8. Sexan $A, B \in \mathcal{M}_{2\times 4}(\mathbb{R})$ as matrices:
$$A = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & 2 & -2 & -1 \end{pmatrix}.$$

Estude cales das seguintes afirmacións son verdadeiras e cales falsas:

a) A^tB é simétrica.

b) $AB^t - 2I$ é antisimétrica.

c) $AA^tB = 2B$.

d) $AB^tB = -A$.

9. Sexan $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ tales que A é simétrica e $B + I$ é antisimétrica. Probe que AB é simétrica, se, e só se, $AB + BA = -2A$.
10. Ache o valor de $\alpha \in \mathbb{R}$ para que a matriz

$$A = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -1/2 & -1/2 & 1/2 & 1/2 \\ -1/2 & 1/2 & -1/2 & 1/2 \\ 1/2 & \alpha & -1/2 & 1/2 \end{pmatrix}$$

sexa ortogonal.

11. Se $u \in \mathbb{R}^n$ é un vector tal que $u^tu = 1$ entón a matriz

$$H(u) = I_n - 2uu^t \in \mathcal{M}_{n\times n}(\mathbb{R})$$

chámase *matriz de Householder* correspondente ao vector u.

a) Probe que toda matriz de Houlseholder $H(u)$ é simétrica e ortogonal.

b) Demostre que $H(u)v = -v$ se existe $\lambda \in \mathbb{R}$ tal que $v = \lambda u$.

c) Comprobe que, se $v \in \mathbb{R}^n$ verifica que $v^tu = 0$, entón $H(u)v = v$.

d) Se $u = (\sqrt{3}/3, -\sqrt{3}/3, \sqrt{3}/3)^t \in \mathbb{R}^3$, verifique que $u^tu = 1$ e calcule a correspondente matriz de Householder $H(u)$.

12. Sexan $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ e sexa $M = A + iB \in \mathcal{M}_{n\times n}(\mathbb{C})$. Probe que M é hermitiana, se, e só se, A é simétrica e B é antisimétrica.
13. Sexan $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ tales que A é ortogonal e $AB = BA$. Probe:

a) $A^tB = BA^t$.

b) Se $H = A + iB \in \mathcal{M}_{n\times n}(\mathbb{C})$ entón $HH^* = H^*H \Longleftrightarrow BB^t = B^tB$.

14. Dada unha matriz cadrada A de orde n, defínese a súa traza como a suma de todos os elementos da diagonal principal, isto é:

$$\mathrm{tr}\,(A) = a_{11} + \cdots + a_{nn}.$$

Demostre que para todas A e B matrices cadradas de orde n tense o seguinte.

a) $\mathrm{tr}\,(A + B) = \mathrm{tr}\,(A) + \mathrm{tr}\,(B)$.

b) $\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA)$.

c) Para cada escalar α cúmprese que $\mathrm{tr}\,(\alpha A) = \alpha \mathrm{tr}\,(A)$.

d) $\mathrm{tr}\,(A^t) = \mathrm{tr}\,(A)$.

15. Sexan $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ verificando que:

a) $\mathrm{tr}\,(B) = 0$,

b) $A - 2B + 2A^t - B^t = 6I$.

Atope todas as matrices $X \in \mathcal{M}_{n\times n}(\mathbb{R})$ satisfacendo a ecuación matricial

$$X = \mathrm{tr}\,(X)A + B.$$

16. Probe que, se $A \in \mathcal{M}_{p\times q}(\mathbb{C})$ verifica $\mathrm{tr}\,(AA^*) = 0$, entón $A = \Theta_{p,q}$.
17. Sexan $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ dúas matrices ortogonais.

a) Calcule a traza da matriz $N = B^tAB - A^t$.

b) Demostre que $(A + B)(A^t - B^t) = \Theta_{n,n}$, se, e só se, BA^t é unha matriz simétrica.

18. Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, $A \neq \Theta_{n.n}$, e sexa $B \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que $(A + B)^2 = (A - B)^2$. Demostre:

a) $\mathrm{tr}\,(BA) = 0$.

b) $(AB)^t = -A^tB^t$.

c) Se n é impar e A é invertíbel, entón B non é invertíbel.

19. Sexan $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$. Razoe cales das seguintes afirmacións son verdadeiras e cales falsas.
   a) $\operatorname{rang}(AB) = \operatorname{rang}(BA)$.
   b) $\operatorname{rang}(A+B) = \operatorname{rang}(A) + \operatorname{rang}(B)$.
   c) $\operatorname{rang}(\lambda A) = \operatorname{rang}(A)\,,\ \forall\,\lambda \in \mathbb{R}\setminus\{0\}$.
   d) Se $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ e $B = \frac{1}{2}(A + A^t)$ entón $\operatorname{rang}(A) = \operatorname{rang}(B)$.

20. Calcule, segundo os distintos valores de $\alpha, \beta \in \mathbb{R}$, o rango das matrices:

$$A = \begin{pmatrix} 2\alpha & \beta & 1 \\ 2 & \alpha\beta & 1 \\ 2 & \beta & \alpha \end{pmatrix}, \quad B = \begin{pmatrix} \alpha & 3 & 4 \\ 1 & -1 & 1 \\ \alpha & 4 & 3 \\ 1 & \alpha & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} \alpha & 1 & 1 & 1 \\ 1 & \alpha & 1 & \alpha \\ 1 & 1 & \alpha & \alpha^2 \end{pmatrix},$$

$$D = \begin{pmatrix} \alpha & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 2 \\ 1 & 0 & -\alpha \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 5 & -1 & 2 \\ 2 & 1 & 4 & 1 \\ 1 & 2 & 1 & \alpha \end{pmatrix}, \quad F = \begin{pmatrix} 2 & 0 & 1 & 2 \\ 1 & -1 & 2 & \beta \\ 0 & 2 & -3 & \alpha \end{pmatrix}.$$

21. Sexa $S = \{A \in \mathcal{M}_{3\times 4}(\mathbb{R}) \,/\, \operatorname{rang}(A) = 2\}$. Cales das seguintes afirmacións son verdadeiras e cales falsas? Razoe a resposta.
   a) Se $A \in S$ e $B \in \mathcal{M}_{4\times 4}(\mathbb{R})$, entón $AB \in S$.
   b) Se $A, B \in S$, entón $A + B \in S$.
   c) Se $A \in S$, entón $\lambda A \in S$, $\forall \lambda \in \mathbb{R}\setminus\{0\}$.

22. Calcule a forma graduada reducida de B e determine $\alpha, \beta \in \mathbb{R}$ para que as matrices

$$A = \begin{pmatrix} 1 & 0 & -2 & 1 \\ -2 & 1 & 1 & \alpha \\ 0 & 2 & \beta & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 4 & -6 & -7 \\ -1 & 1 & 1 & -3 \\ 2 & 2 & -6 & -12 \end{pmatrix}$$

teñan o mesmo rango.

23. a) Determine os valores dos parámetros α e β para que as matrices reais

$$A = \begin{pmatrix} 1 & 0 & -2 & 1 \\ -2 & 1 & 1 & \alpha \\ 0 & 2 & \beta & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 4 & -6 & -7 \\ -1 & 1 & 1 & -3 \\ 2 & 2 & -6 & -12 \end{pmatrix}$$

   sexan equivalentes.
   b) Calcule a forma reducida completa ou normal

$$I_r^{3\times 4} = \left(\begin{array}{c|c} I_r & \Theta_{r,4-r} \\ --- & ----- \\ \Theta_{3-r,r} & \Theta_{3-r,4-r} \end{array}\right)$$

   da matriz B.
   c) Para $\alpha = \beta = 0$, ache as matrices regulares P e Q tales que $Q^{-1}AP = B$.

24. Ache o rango das matrices:
   a)

$$A = \begin{pmatrix} 1 & \alpha & -\alpha^2 \\ 1 & \beta & -\beta^2 \\ 1 & \delta & -\delta^2 \end{pmatrix}.$$

   b)

$$B = \begin{pmatrix} 1 & i & 1+i \\ 1+2i & 2i & 0 \\ 1+i & 1+2i & 2+i \end{pmatrix}.$$

c)

$$C = \begin{pmatrix} 1 & -1 & i \\ 1+i & i-1 & 1-i \\ 2+i & i-2 & \alpha \end{pmatrix}.$$

d)

$$D = \begin{pmatrix} \alpha & 3i-2+\alpha & i-2 \\ 1-i-\alpha & -\alpha & i-1 \\ i & -1+i & -1 \end{pmatrix}.$$

e)

$$E = \begin{pmatrix} \alpha & 1 & 1 & 1 \\ 1 & \alpha & 1 & 1 \\ 1 & 1 & \alpha & 1 \\ 1 & 1 & 1 & \alpha \end{pmatrix}.$$

f)

$$F = \begin{pmatrix} 1 & \beta & \alpha & \beta \\ 2 & 3\beta & \alpha & 2\beta \\ 1 & \beta & 2\alpha & 2\beta \\ 1 & 2\beta & 0 & 2\beta \end{pmatrix}.$$

25. Considéranse as matrices

$$A = \begin{pmatrix} 4 & -5 \\ -3 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 4 & -1 \\ -3 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

a) Probe que A e B son matrices regulares e calcule A^{-1} y B^{-1}.
b) Resolva, se é posíbel, as seguintes ecuacións matriciais:
*b*1) $BX = C$.
*b*2) $3BX - A^t = 0$.
*b*3) $AX - BX = I$.
*b*4) $AXB = A + B$.
*b*5) $A^tBX^t = C$.
*b*6) $A^2X = I$.

26. Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Probe que:
a) Se A é ortogonal, entón $\det(A) = \pm 1$.
b) Se n é impar e A é antisimétrica, entón $\det(A) = 0$.
c) Se $P \in \mathcal{M}_{n\times n}(\mathbb{C})$ é unha matriz invertíbel, entón

$$\det(P^{-1}AP) = \det(A).$$

27. Sexan $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ matrices ortogonais.
a) Probe que $A + B = A(A^t + B^t)B$.
b) Como consecuencia, probe que, se $\det(A) + \det(B) = 0$, entón $A + B$ é singular.

28. Probe que

$$\begin{vmatrix} 1 & \alpha & \beta+\gamma \\ 1 & \beta & \alpha+\gamma \\ 1 & \gamma & \alpha+\beta \end{vmatrix} = 0.$$

29. Dados $\alpha, \beta \in \mathbb{R}$, considérase a matriz $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ definida por

$$a_{ij} = \begin{cases} \alpha & \text{si } i = j \\ \beta & \text{si } i \neq j \end{cases}$$

Probe que $\det(A) = (\alpha-\beta)^{n-1}[\alpha + (n-1)\beta]$.

30. Calcule o determinante da matriz

$$A = \begin{pmatrix} \alpha+1 & 1 & 1 & 3 \\ 2\alpha & \alpha+\beta & \beta & \beta \\ \alpha & \beta & \beta & \beta \\ \alpha+1 & 1+\beta & 1+\beta & 3+\beta \end{pmatrix}$$

segundo os valores de $\alpha, \beta \in \mathbb{R}$. ¿Para que valores de α e β é a matriz A regular?

31. Calcule o determinante e, se é posíbel, as inversas das seguintes matrices:

$$A = \begin{pmatrix} -1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{pmatrix},\ B = \begin{pmatrix} 1 & -1 & 2 & 1 & 1 \\ 2 & -1 & 3 & 2 & 0 \\ 1 & 0 & 2 & 2 & -2 \\ 3 & -1 & 5 & 5 & 4 \\ 1 & -1 & 2 & 0 & -4 \end{pmatrix},\ C = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix},$$

$$D = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 1 & 2 & 4 \end{pmatrix},\quad E = \begin{pmatrix} 1 & 1+i & -i \\ 0 & i & 1-2i \\ 1 & 1 & i \end{pmatrix},\quad F = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 1 & 1 & 0 & 1 \\ -1 & 0 & 1 & -1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

32. Calcule os determinantes das seguintes matrices:

$$A = \begin{pmatrix} 1 & 2 & -1 & -0{,}002 \\ 3 & 8 & 0 & -0{,}004 \\ 2 & 2 & -4 & -0{,}003 \\ 3000 & 8000 & -1000 & -6 \end{pmatrix},\quad B = \begin{pmatrix} 1 & 3 & 5 & 2 & 2 \\ 3 & 6 & 9 & 5 & 1 \\ 6 & 5 & 7 & 4 & 2 \\ 1 & 6 & 8 & 5 & 0 \\ 1 & 4 & 6 & 3 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 128 & 256 & 384 & 512 \\ \frac{1}{4} & \frac{3}{8} & \frac{1}{8} & \frac{1}{4} \\ \frac{1}{64} & \frac{1}{64} & \frac{1}{64} & \frac{-1}{64} \\ 2 & 0 & -4 & -12 \end{pmatrix},\quad D = \begin{pmatrix} 5 & 7 & 6 & 8 & 5 \\ 3 & 5 & 4 & 6 & 4 \\ 2 & 3 & 2 & 4 & 3 \\ 1 & 2 & 0 & 2 & 2 \\ 2 & 4 & 2 & 4 & 4 \end{pmatrix}.$$

33. Sabendo que o determinante de Vandermonde admite a expresión

$$det \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ x_1 & x_2 & x_3 & \ldots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \ldots & x_n^2 \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \ldots & x_n^{n-1} \end{pmatrix} = \prod_{i>j}(x_i - x_j)$$

demostre que para todo número $k \geq 1$ o determinante da matriz

$$A = \begin{pmatrix} 1 & 2 & 3 & \ldots & k \\ 1 & 4 & 9 & \ldots & k^2 \\ 1 & 8 & 27 & \ldots & k^3 \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ 1 & 2^k & 3^k & \ldots & k^k \end{pmatrix}$$

é non nulo.

34. Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ tal que $A - A^* = \operatorname{tr}(A)A^*$. Demostre que $\operatorname{tr}(A) = 0$ e $\det(A) \in \mathbb{R}$.
35. Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ antisimétrica e sexa $B = (I_n + A)(I_n - A)^{-1}$. Probe que B é unha matriz ortogonal.
36. Razoe se as seguintes afirmacións son verdadeiras ou falsas:
    a) Se $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ verifica que $(A + I_n)(A - I_n) = \Theta_{n,n}$, entón $A = I_n$ ou $A = -I_n$.
    b) Se $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ é unha matriz nilpotente de índice 2, entón
    $$A(M + A)^2 = AM^2\,,\ \forall\, M \in \mathcal{M}_{n\times n}(\mathbb{R})\,.$$
    c) Se A é nilpotente, entón A é singular.
    d) Sexa $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ definida por $a_{ij} = \alpha$, se $i + j = n + 1$, e $a_{ij} = 0$ noutro caso, onde $\alpha \in \mathbb{R}$, $\alpha \neq 0$. Entón cúmprese que $\det(A) = -\det(\alpha I_n)$.
    e) Sexan $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ dúas matrices invertíbeis. Daquela cúmprese que
    $$\operatorname{rang}(B^t A^{-1} B) = \operatorname{rang}(A^t A).$$

CAPíTULO 4

Sistemas de ecuacións lineares

Os contidos deste capítulo son dos máis relevantes dentro da álxebra linear xa que a solución de sistemas de ecuacións lineares foi un dos problemas, por non dicir o problema, que deu lugar á súa orixe.

Comezaremos o capítulo introducindo as nocións de ecuación linear e de sistema de ecuacións lineares, para pasar a continuación a dar a interpretación matricial destes últimos. No seguinte punto clasificaremos os sistemas de ecuacións lineares en función de se teñen solución (compatíbeis) ou se non a teñen (incompatíbeis). Dentro dos compatíbeis, distinguiremos entre os que teñen infinitas solucións ou unha única solución. Os primeiros serán os compatíbeis indeterminados, mentres que os segundos chamaranse compatíbeis determinados.

En segundo lugar, definiremos a importante noción de sistemas equivalentes, isto é, sistemas co mesmo conxunto de solucións. Neste capítulo recalcarase ao lector que, se nun sistema $Ax = b$ multiplicamos por unha matriz invertíbel P, o novo sistema $PAx = Pb$ é equivalente ao primeiro. Por tanto, tendo en conta a interpretación matricial das operacións elementais por filas, poderemos garantir que, se nun sistema de ecuacións lineares $Ax = b$ intercambiamos dúas ecuacións, multiplicamos unha ecuación por un escalar non nulo ou sumamos a unha ecuación un múltiplo doutra ecuación, obtemos un sistema de ecuacións equivalente. Como consecuencia destes feitos, poderase usar a seguinte propiedade para discutir e resolver sistemas de ecuacións lineares: sexa $Ax = b$ un sistema de m ecuacións lineares con n incógnitas e sexa $(A|b)$ a súa matriz ampliada. Se $C = (A'|b')$ é a forma graduada reducida de $(A|b)$, tense que o sistema $A'x = b'$ é equivalente ao sistema $Ax = b$.

A primeira consecuencia directa do resultado enunciado nas últimas liñas do parágrafo precedente é o que se coñece na literatura como Teorema de Rouché-Frobenius. Este teorema é o que permitirá discutir os sistemas de ecuacións lineares sen máis que estudar os rangos da matriz do sistema e da matriz ampliada.

Finalizaremos esta parte do capítulo introducindo as nocións de sistema homoxéneo e non homoxéneo recalcando que, se $Ax = b$ é un sistema de ecuacións lineares compatíbel, o conxunto de solucións do sistema obtense como a suma dunha solución particular e as solucións do sistema homoxéneo asociado $Ax = \theta_{\mathbb{K}^n}$ sendo $\theta_{\mathbb{K}^n}$ a matriz columna con todas as súas compoñentes nulas.

O primeiro método de solución de sistemas de ecuacións lineares que se introducirá é o derivado da eliminación gaussiana, que, á súa vez, ten o seu fundamento na utilización das operacións elementais para converter un sistema de ecuacións lineares noutro máis sinxelo de resolver; neste caso pasaríase a outro sistema equivalente con matriz asociada triangular superior. Neste capítulo explicarase con rigor o procedemento de eliminación gaussiana e, como consecuencia, verase que se no proceso de Gauss para resolver o sistema $Ax = b$ non temos necesidade de intercambiar filas ou de multiplicar unha fila por un escalar non nulo, existe unha matriz P, triangular inferior cos coeficientes da diagonal principal iguais a un, tal que $PA = U$ sendo U unha matriz triangular superior. Como a inversa dunha matriz triangular inferior cos coeficientes da diagonal principal iguais a un é triangular inferior cos coeficientes da diagonal principal iguais a un, teremos o seguinte:

$$A = LU, \;\; L = P^{-1}$$

onde L é unha matriz triangular inferior cos coeficientes da diagonal principal iguais a un e U é unha matriz triangular superior; esta é a factorización LU da matriz A. Agora ben, nestas condicións xerais non podemos garantir a unicidade de L e de U. Para garantir a unicidade da factorización é necesario que todos os pivotes que aparezan no desenvolvemento do procedemento de eliminación gaussiana deben ser non nulos ou, equivalentemente, as submatrices diagonais de A,

$$A(r,r) = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}$$

deben ser invertíbeis para todo $r \in \{1, \cdots, n\}$.

Finalmente, neste capítulo tamén se verá como se pode aplicar a factorización LU á resolución de sistemas de ecuacións lineares e comprobarase que, se A admite factorización LU, entón

$$\det(A) = \det(U)$$

e, como $U = (u_{ij})$ é triangular superior, tense que

$$\det(A) = u_{11}u_{22}\cdots u_{nn}.$$

4.1. Sistemas homoxéneos e non homoxéneos. Existencia de solucións

Definición 4.1.1. Unha ecuación linear con coeficientes no corpo $\mathbb{K}$ é unha expresión da forma:

$$a_1x_1 + \cdots + a_nx_n = b$$

onde os a_i son elementos coñecidos do corpo que chamaremos coeficientes, as x_i son símbolos que chamaremos incógnitas e b é outro escalar coñecido que chamaremos termo independente. É importante resaltar que nunha ecuación linear non poden aparecer as incógnitas elevadas ao cadrado, nin produtos de incógnitas, nin funcións de tipo trigonométrico, logarítmico, etc.

Unha solución da ecuación é unha asignación de escalares v_i para as incógnitas x_i de forma que se cumpra a igualdade. Normalmente unha solución representarémola como unha matriz columna

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

cumpríndose que

$$\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = a_1v_1 + \cdots + a_nv_n = b.$$

Cando temos poucas incógnitas adóitanse representar polas letras x, y, z, t, ...

Exemplo 4.1.2. As seguintes ecuacións

$$x + 20y - 3t = 23, \quad ix - (4+i)z - 6t = 0$$

son lineares, mentres que

$$x^2 + 3y + 4z = 6, \quad xyz + 3t = 9, \quad sen(x) + 4t + 5r = e^x$$

non o son.

Por exemplo, $x = 6$, $y = 1$, $z = 0$, $t = 1$ é unha solución da ecuación $x + 20y - z - 3t = 23$. Neste caso,

$$\begin{pmatrix} 1 & 20 & -1 & -3 \end{pmatrix} \begin{pmatrix} 6 \\ 1 \\ 0 \\ 1 \end{pmatrix} = 1 \cdot 6 + 20 \cdot 1 + 0 \cdot 0 + (-3) \cdot 1 = 23.$$

Definición 4.1.3. A un conxunto de m ecuacións lineares coas mesmas incógnitas

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{i1}x_1 + \cdots + a_{in}x_n = b_i \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

chámaselle sistema de m ecuacións lineares con n incógnitas.

Nótese que o anterior sistema admite a seguinte formulación matricial:

$$\begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ a_{i1} & \cdots & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{pmatrix}.$$

Se chamamos A á matriz

$$\begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ a_{i1} & \cdots & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{pmatrix}$$

e x, b a

$$\begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix}, \quad \begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{pmatrix},$$

respectivamente, o sistema está dado polo produto matricial

$$Ax = b.$$

Este é o formato máis simple para representar de forma abstracta un sistema. A matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ chamarase matriz do sistema, x será coñecido como vector de incógnitas (ao darlle valores escalares a cada incógnita temos un elemento de $\mathcal{M}_{n\times 1}(\mathbb{K})$) e $b \in \mathcal{M}_{m\times 1}(\mathbb{K})$ denominarase termo independente.

Definición 4.1.4. Dado un sistema de m ecuacións lineares con n incógnitas $Ax = b$, chámase matriz ampliada do sistema á que resulta cando se engade a A o termo independente como columna final. Representarémola como

$$(A|b).$$

Definición 4.1.5. Dado un sistema de m ecuacións lineares con n incógnitas $Ax = b$, chamarase homoxéneo se $b = \theta_{\mathbb{K}^m}$. Dado un sistema non homoxéneo $Ax = b$, chamarase sistema homoxéneo asociado ao sistema $Ax = \theta_{\mathbb{K}^m}$.

Definición 4.1.6. Dado un sistema de m ecuacións lineares con n incógnitas $Ax = b$, chamarase solución do sistema a todo $v \in \mathcal{M}_{n\times 1}(\mathbb{K})$ tal que $Av = b$. Nótese que se as columnas de A son $A^1, \cdots, A^n$, entón $v \in \mathcal{M}_{n\times 1}(\mathbb{K})$ é solución, se, e só se,

$$\begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ a_{i1} & \cdots & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ \vdots \\ v_n \end{pmatrix} = v_1 A^1 + \cdots + v_n A^n = b.$$

Cando isto ocorre, dise que b é combinación linear das columnas de A. Máis adiante veremos o sentido que ten este nome no contexto da teoría de espazos vectoriais.

Un sistema de ecuacións lineares chamarase compatíbel se admite solución. Cando é única, chamarase compatíbel determinado. Se temos máis dunha solución, o sistema denominarase compatíbel indeterminado. Os sistemas incompatíbeis son aqueles que non teñen solución.

Exemplos 4.1.7. Se temos un sistema de ecuacións lineares $Ax = b$ onde A é unha matriz invertíbel, verifícase que $Ax = b$ é compatíbel determinado con solución $v = A^{-1}b$. Doutra banda, todo sistema homoxéneo de n incógnitas, é compatíbel xa que $\theta_{\mathbb{K}^n}$ é unha solución.

Vexamos agora exemplos concretos. Tomemos en primeiro lugar o sistema:

$$\left\{ \begin{array}{l} x+y+z=6 \\ x+y-z=0 \\ x+z=4 \\ x+y=3 \end{array} \right. \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 4 \\ 3 \end{pmatrix}$$

onde da última ecuación obtemos que $x = 3 - y$. Ademais

$$\left\{ \begin{array}{l} x+y-z=0 \\ x+z=4 \end{array} \right. \quad \Rightarrow \quad 2x+y=4 \quad \Rightarrow \quad 2(3-y)+y=4 \quad \Rightarrow \quad 6-y=4 \quad \Rightarrow y=2.$$

Entón, $x = 1$ e $z = 3$. Daquela, o sistema resulta compatíbel determinado.

Sexa a continuación o sistema

$$\left\{ \begin{array}{l} x+y+z=6 \\ x+y-z=0 \end{array} \right. \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix}.$$

Neste caso, sumando as dúas ecuacións, obtemos que $2x + 2y = 6$ ou, o que é o mesmo, $x + y = 3$. Por tanto, $y = 3 - x$ e $z = x + y = 3$. Isto implica que calquera vector v tal que

$$v = \begin{pmatrix} \alpha \\ 3-\alpha \\ 3 \end{pmatrix}, \quad \alpha \in \mathbb{R}$$

é solución do sistema. Como temos infinitas solucións o sistema é compatíbel indeterminado.

Finalmente, consideremos o sistema

$$\left\{ \begin{array}{l} x+y+z=6 \\ x+y-z=0 \\ -2x-2y=1 \end{array} \right. \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ -2 & -2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 1 \end{pmatrix}.$$

Sumando as dúas primeiras ecuacións obtemos que $2x+2y = 6$, equivalentemente, $x+y = 3$. Isto é imposíbel que se cumpra ao mesmo tempo que $-2x - 2y = 1 \Leftrightarrow x + y = -1/2$. Por tanto, non existe solución e o sistema resulta incompatíbel.

Definición 4.1.8. Diremos que dous sistemas de ecuacións lineares co mesmo número de incógnitas son equivalentes, se teñen o mesmo conxunto de solucións.

Proposición 4.1.9. *Sexa $P \in \mathcal{M}_{m\times m}(\mathbb{R})$ unha matriz invertíbel. Tense que todo sistema de m ecuacións lineares con n incógnitas $Ax = b$ é equivalente a $PAx = Pb$.*

Demostración. Supoñamos que v é solución de $Ax = b$. Entón é solución de $PAx = Pb$ xa que $PAv = Pb$ por ser $Av = b$. De xeito inverso, se v é solución de $PAx = Pb$, temos que $PAv = Pb$ e multiplicando por P^{-1} obtemos que

$$Av = P^{-1}PAv = P^{-1}Pb = b.$$

Por tanto, v é solución de $Ax = b$. □

Como consecuencia deste resultado obtemos unhas importantes propiedades que nos serán de utilidade á hora de resolver sistemas.

Proposición 4.1.10. *Se nun sistema de ecuacións lineares $Ax = b$ intercámbianse dúas ecuacións, multiplícase unha ecuación por un escalar non nulo ou se suma a unha ecuación un múltiplo doutra ecuación, obtense un sistema de ecuacións equivalente.*

Demostración. Intercambiar dúas ecuacións, por exemplo a ecuación i e a ecuación j, faise sen máis que multiplicar o sistema pola matriz de operacións elementais F_{ij}. Como é invertíbel, o novo sistema é equivalente a $Ax = b$. Se queremos multiplicar unha ecuación por un escalar α non nulo, temos que multiplicar o sistema pola matriz de operacións elementais $F_i(\alpha)$ que tamén é invertíbel. Desta forma, neste caso, temos un novo sistema equivalente ao de partida. Finalmente, a última operación con ecuacións faise multiplicando por unha matriz de operacións elementais do tipo $F_{ij}(\alpha)$ que, como é invertíbel, proporciona un sistema equivalente a $Ax = b$

□

Corolario 4.1.11. *Sexa $Ax = b$ un sistema de m ecuacións lineares con n incógnitas. Sexa $(A|b)$ a súa matriz ampliada. Se $(A'|b')$ é a forma graduada reducida de $(A|b)$, tense que o sistema $A'x = b'$ é equivalente a $Ax = b$.*

Demostración. É consecuencia directa da proposición anterior. □

Corolario 4.1.12. *Sexa $Ax = b$ un sistema de m ecuacións lineares con n incógnitas. Sexa $(A|b)$ a súa matriz ampliada. Entón:*

1) *Se $\operatorname{rang}(A) = \operatorname{rang}(A|b)$, o sistema é compatíbel. Cando $rang(A) = n$, o sistema é determinado e, se $\operatorname{rang}(A) < n$, é indeterminado.*
2) *Se $\operatorname{rang}(A) < \operatorname{rang}(A|b)$, o sistema é incompatíbel.*

Demostración. Calculemos a forma graduada reducida $(A'|b')$ da matriz ampliada do sistema $(A|b)$. Se eliminamos a última columna de $(A'|b')$, temos unha forma graduada reducida para a matriz A. Como podemos deducir de forma fácil, o sistema é compatíbel cando no sistema $A'x = b'$ non aparece a ecuación $0 = \alpha$ con α un escalar non nulo, é dicir, se A' e $(A'|b')$ teñen o mesmo número de filas non nulas, equivalentemente, $\operatorname{rang}(A) = \operatorname{rang}(A|b)$.

Se $\operatorname{rang}(A) = \operatorname{rang}(A|b) = r$, o sistema resólvese despexando as r incógnitas correspondentes ás r entradas principais en función das $n-r$ restantes. O sistema será determinado se $n = r$. □

Estes últimos corolarios proporcionan un método, coñecido como método de Gauss-Jordan, baseado na utilización das operacións elementais para simplificar os procesos de caracterización do sistema e obtención do seu conxunto de solucións no caso de que as teña. Vexamos uns exemplos:

Exemplos 4.1.13. Tomemos o sistema

$$\left\{\begin{array}{l} 3x+6y-5z=0 \\ x+y+2z=0 \\ 2x+4y-3z=1 \end{array}\right. \Leftrightarrow \begin{pmatrix} 3 & 6 & -5 \\ 1 & 1 & 2 \\ 2 & 4 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Sexa a matriz ampliada do sistema

$$\left(\begin{array}{ccc|c} 3 & 6 & -5 & 0 \\ 1 & 1 & 2 & 0 \\ 2 & 4 & -3 & 1 \end{array}\right)$$

e calculemos a forma graduada reducida (neste caso, a matriz onde se acumulan as operacións elementais non nos interesa polo que non a escribimos).

$$(A|b) \overset{F_{12}}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 3 & 6 & -5 & 0 \\ 2 & 4 & -3 & 1 \end{array}\right) \overset{F_{31}(-2),F_{21}(-3)}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 3 & -11 & 0 \\ 0 & 2 & -7 & 1 \end{array}\right) \overset{F_{23}}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & -7 & 1 \\ 0 & 3 & -11 & 0 \end{array}\right)$$

$$\overset{F_{2}(1/2)}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 1 & -7/2 & 1/2 \\ 0 & 3 & -11 & 0 \end{array}\right) \overset{F_{32}(-3)}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 1 & -7/2 & 1/2 \\ 0 & 0 & -1/2 & -3/2 \end{array}\right)$$

$$\overset{F_{3}(-2)}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 1 & -7/2 & 1/2 \\ 0 & 0 & 1 & 3 \end{array}\right) \overset{F_{12}(-1)}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 0 & 11/2 & -1/2 \\ 0 & 1 & -7/2 & 1/2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$$\overset{F_{13}(-11/2),F_{23}(7/2)}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -17 \\ 0 & 1 & 0 & 11 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

Por tanto, $\operatorname{rang}(A) = \operatorname{rang}(A|b) = 3$. Daquela, o sistema orixinal é compatíbel determinado e a solución está dada por

$$\left\{\begin{array}{l} x=-17 \\ y=11 \\ z=3 \end{array}\right.$$

Tomenos agora o sistema

$$\left\{\begin{array}{l} x+y+z=1 \\ x+2y=1 \\ x+2z=1 \end{array}\right. \Leftrightarrow \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

e calculemos a forma graduada reducida da matriz ampliada

$$(A|b) \overset{F_{31}(-1),F_{21}(-1)}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right) \overset{F_{32}(1)}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\overset{F_{12}(-1)}{\Longrightarrow} \left(\begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Entón, $\operatorname{rang}(A) = \operatorname{rang}(A|b) = 2 < 3$ e o sistema orixinal é compatíbel indeterminado e equivalente ao sistema

$$\left\{\begin{array}{l} x+2z=1 \\ y-z=0 \end{array}\right.$$

que ten como solucións aos u dados por

$$u = \begin{pmatrix} 1-2\alpha \\ \alpha \\ \alpha \end{pmatrix}, \ \alpha \in \mathbb{R}.$$

Finalmente, sexa o sistema

$$\begin{cases} 2x+2y+2z=0 \\ 4x+7y+7z=24 \\ 6x+18y+18z=70 \end{cases} \Leftrightarrow \begin{pmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 18 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 24 \\ 70 \end{pmatrix}$$

e, como nos casos anteriores, empecemos o cálculo da forma graduada reducida da matriz ampliada.

$$(A|b) \overset{F_{31}(-3),F_{21}(-2)}{\Longrightarrow} \left(\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ 0 & 3 & 3 & 24 \\ 0 & 12 & 12 & 70 \end{array}\right) \overset{F_{32}(-4)}{\Longrightarrow} \left(\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ 0 & 3 & 3 & 24 \\ 0 & 0 & 0 & -26 \end{array}\right).$$

Chegados a este punto non temos que facer nada máis, xa que se ve claramente que temos $0=-26$. O sistema é incompatíbel e $\operatorname{rang}(A) = 2 < 3 = \operatorname{rang}(A|b)$.

Proposición 4.1.14. *Sexa $Ax = b$ un sistema de ecuacións lineares de m ecuacións con n incógnitas tal que é compatíbel. O conxunto de solucións do sistema obtense como a suma dunha solución particular e as solucións do sistema homoxéneo asociado $Ax = \theta_{\mathbb{K}^m}$.*

Demostración. Supoñamos que v é unha solución de $Ax = b$ e que w o é de $Ax = \theta_{\mathbb{K}^m}$. Entón,

$$A(v+w) = Av + Aw = b + \theta_{\mathbb{K}^m} = b$$

e isto implica que $v+w$ é unha solución de $Ax = b$. Inversamente, se temos dúas solucións de $Ax = b$, v e u, sexa $w = u - v$. É claro que $u = v + w$ e, ademais, w é solución do sistema homoxéneo asociado, xa que $Aw = A(u-v) = Au - Av = b - b = \theta_{\mathbb{K}^m}$. □

Exemplo 4.1.15. Por exemplo, nun caso do exemplo anterior estudamos o sistema

$$\begin{cases} x+y+z=1 \\ x+2y=1 \\ x+2z=1 \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

que resultou ser compatíbel indeterminado. Ten como solucións aos u, tales que

$$u = \begin{pmatrix} 1-2\alpha \\ \alpha \\ \alpha \end{pmatrix}, \ \alpha \in \mathbb{R}.$$

Se tomamos

$$v = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

temos que v é unha solución do sistema $Ax = b$ (é unha solución particular) e que

$$\left\{ v + \begin{pmatrix} -2\alpha \\ \alpha \\ \alpha \end{pmatrix}, \ \alpha \in \mathbb{R} \right\}$$

é o conxunto total de solucións, onde

$$\left\{ \begin{pmatrix} -2\alpha \\ \alpha \\ \alpha \end{pmatrix}, \ \alpha \in \mathbb{R} \right\}$$

son as solucións do sistema homoxéneo asociado.

Nota 4.1.16. Obviamente, se $Ax = b$ é compatíbel e o seu sistema homoxéneo asociado só ten a solución trivial, obtemos que o sistema é determinado.

4.2. Eliminación Gaussiana. Factorización LU

Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e sexan A_1, A_2 dúas matrices tales que $A = A_1A_2$. A solución dun sistema de ecuacións lineares $Ax = b$ pódese abordar resolvendo os sistemas $A_1z = b$, $A_2x = z$, xa que

$$Ax = A_1A_2x = A_1z = b.$$

Este procedemento é especialmente útil cando os dous sistemas $A_1z = b$ e $A_2x = z$ son moito máis fáciles de resolver que o orixinal. Existen varios métodos baseados nesta idea. Na parte final do tema veremos un dos máis clásicos.

Definición 4.2.1. Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Diremos que A admite unha factorización LU se podemos factorizar A como produto de dúas matrices L e U onde L é unha matriz triangular inferior cos elementos da diagonal principal iguais a un e U é unha matriz triangular superior.

$$A = LU = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ l_{21} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ l_{n-1\,1} & l_{n-1\,2} & \cdots & 1 & 0 \\ l_{n1} & l_{n2} & \cdots & l_{n\,n-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1\,n-1} & u_{nn} \\ 0 & u_{22} & \cdots & u_{2\,n-1} & u_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & u_{n-1\,n-1} & u_{n-1\,n} \\ 0 & 0 & \cdots & 0 & u_{nn} \end{pmatrix}$$

Se existe tal descomposición, temos que

$$\det(A) = \det(LU) = \det(L)\det(U) = \det(U) = u_{11}u_{22}\cdots u_{n-1\,n-1}u_{nn}.$$

O cálculo da factorización LU dunha matriz A está baseado no procedemento de eliminación gaussiana que consiste en converter a matriz de partida A nunha matriz triangular superior mediante a utilización de operacións elementais con filas. O método de Gauss realizarase en $n-1$ etapas, se a matriz A ten n filas.

<u>ETAPA 1 DO MÉTODO DE GAUSS.</u>

Na primeira etapa do método faranse ceros debaixo do primeiro elemento da diagonal principal de A. Para iso distinguimos dous pasos:

<u>Paso I:</u> Compróbase se o elemento a_{11} é non nulo. Se é distinto de cero, pásase ao paso II; se é nulo, débese intercambiar a primeira fila con outra fila, por exemplo, a fila i-ésima ($i > 1$), tal que $a_{i1} \neq 0$, e pásase ao paso II. Se non existise unha fila en tales condicións, pásase á etapa II.

<u>Paso II:</u> Anúlanse todos os elementos colocados na primeira columna debaixo do que ocupa a primeira posición da diagonal principal.

A interpretación matricial destes pasos é a seguinte:

<u>Paso I:</u> Multiplicamos A pola esquerda pola matriz:

$$P_1 = \begin{cases} I_n & \text{se } a_{11} \neq 0 \\ F_{1i} & \text{se } a_{11} = 0 \text{ e } \exists\, i > 1\ a_{i1} \neq 0 \\ I_n & \text{se } a_{i1} = 0\ \forall\, i > 1 \end{cases}$$

<u>Paso II:</u> Sexa $P_1A = B_1 = (b^1_{ij})$. Se debaixo do elemento b^1_{11} todos os b^1_{i1} son nulos tomamos $E_1 = I_n$. En caso contrario, multiplicamos B_1 pola esquerda polas matrices:

$$F_{j1}(-\frac{b^1_{j1}}{b^1_{11}}) \ \ j \in \{2, \cdots, n\}$$

para anulalos e tomamos

$$E_1 = F_{n1}(-\frac{b^1_{n1}}{b^1_{11}})F_{n-1\,1}(-\frac{b^1_{n-1\,1}}{b^1_{11}}) \cdots F_{21}(-\frac{b^1_{21}}{b^1_{11}})$$

En calquera das dúas posibilidades, obtemos que E_1 é unha matriz triangular inferior con uns na diagonal principal, xa que $E_1 = I_n$ ou

$$E_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -\frac{b^1_{21}}{b^1_{11}} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{b^1_{n-1\,1}}{b^1_{11}} & 0 & \cdots & 1 & 0 \\ -\frac{b^1_{n1}}{b^1_{11}} & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Ademais, se $A_1 = (a^1_{ij}) = E_1B_1$, temos que

$$A_1 = E_1P_1A = \begin{pmatrix} a^1_{11} & a^1_{12} & \cdots & a^1_{1\,n-1} & a^1_{1n} \\ 0 & a^1_{22} & \cdots & a^1_{2\,n-1} & a^1_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a^1_{n-1\,2} & \cdots & a^1_{n-1\,n-1} & a^1_{n-1\,n} \\ 0 & a^1_{n1} & \cdots & a^1_{n\,n-1} & a^1_{nn} \end{pmatrix}$$

<u>Etapa r do método de Gauss.</u>

Nesta etapa, atopámonos cunha matriz

$$A_{r-1} = (a^{r-1}_{ij}) = \begin{pmatrix} a^{r-1}_{11} & a^{r-1}_{12} & \cdots & a^{r-1}_{1\,r} & \cdots & a^{r-1}_{1\,n-1} & a^{r-1}_{1n} \\ 0 & 0 & \cdots & a^{r-1}_{2r} & \cdots & a^{r-1}_{2\,n-1} & a^{r-1}_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a^{r-1}_{rr} & \cdots & a^{r-1}_{r\,n-1} & a^{r-1}_{rn} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a^{r-1}_{n-1\,r} & \cdots & a^{r-1}_{n-1\,n-1} & a^{r-1}_{n-1\,n} \\ 0 & 0 & \cdots & a^{r-1}_{nr} & \cdots & a^{r-1}_{n\,n-1} & a^{r-1}_{nn} \end{pmatrix}$$

e temos que facer ceros debaixo do coeficiente $a^r_{r\,r}$. Do mesmo xeito que nas outras etapas debemos dar dous pasos:

<u>Paso I:</u> Compróbase se o elemento a^{r-1}_{rr} é non nulo. Se é distinto de cero, pásase ao paso II; se é nulo, débese intercambiar a fila r con outra fila, por exemplo, a fila i-ésima $(i > r)$ tal que $a^{r-1}_{ir} \neq 0$, e pásase ao paso II. Se non existise unha fila en tales condicións, pásase á etapa $r + 1$.

<u>Paso II:</u> Anúlanse todos os elementos colocados na columna r-ésima debaixo do que ocupa a posición r da diagonal principal.

A interpretación matricial destes pasos é a seguinte:

Paso I: Multiplicamos A_{r-1} pola esquerda pola matriz:

$$P_r = \begin{cases} I_n & \text{se } a_{rr}^{r-1} \neq 0 \\ F_{ir} & \text{se } a_{rr}^{r-1} = 0 \text{ e } \exists\, i > r\ a_{ir}^{r-1} \neq 0 \\ I_n & \text{se } a_{ir}^{r-1} = 0\ \forall\, i > r \end{cases}$$

Paso II: Sexa $P_r A_{r-1} = B_r = (b_{ij}^r)$. Se debaixo do elemento b_{rr}^r todos os b_{ir}^r son nulos tomamos $E_r = I_n$. En caso contrario, multiplicamos B_r pola esquerda polas matrices:

$$F_{jr}(-\frac{b_{jr}^r}{b_{rr}^r}) \quad j \in \{r+1, \cdots, n\}$$

para anulalos e tomamos

$$E_r = F_{nr}(-\frac{b_{nr}^r}{b_{rr}^r}) F_{n-1\,r}(-\frac{b_{n-1\,r}^r}{b_{rr}^r}) \cdots F_{r+1\,r}(-\frac{b_{r+1\,1}^r}{b_{rr}^r})$$

En calquera das dúas posibilidades, obtemos que E_r é unha matriz triangular inferior con uns na diagonal principal, xa que $E_r = I_n$ ou

$$E_r = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & -\frac{b_{r+1\,1}^r}{b_{rr}^r} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\frac{b_{n-1\,r}^r}{b_{rr}^r} & \cdots & 1 & 0 \\ 0 & 0 & \cdots & -\frac{b_{nr}^r}{b_{rr}^r} & \cdots & 0 & 1 \end{pmatrix}$$

Ademais, se $A_r = (a_{ij}^r) = E_r B_r$, temos que

$$A_r = E_r P_r A_{r-1} = \begin{pmatrix} a_{11}^r & a_{12}^r & \cdots & a_{1r}^r & a_{1r+1}^r & \cdots & a_{1\,n-1}^r & a_{1n}^r \\ 0 & a_{22}^r & \cdots & a_{2r}^r & a_{2\,r+1}^r & \cdots & a_{2\,n-1}^r & a_{2n}^r \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{rr}^r & a_{r\,r+1}^r & \cdots & a_{r\,n-1}^r & a_{rn}^r \\ 0 & 0 & \cdots & 0 & a_{r+1\,r+1}^r & \cdots & a_{r+1\,n-1}^r & a_{r+1\,n}^r \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n-1\,r+1}^r & \cdots & a_{n-1\,n-1}^r & a_{n-1\,n}^r \\ 0 & 0 & \cdots & 0 & a_{n\,r+1}^r & \cdots & a_{n\,n-1}^r & a_{nn}^r \end{pmatrix}$$

Desta maneira, tras $n-1$ etapas obtense que

$$E_{n-1}P_{n-1} \cdots E_1 P_1 A = A_{n-1} = (a_{ij}^{n-1}) =$$

$$\begin{pmatrix} a_{11}^{n-1} & a_{12}^{n-1} & \cdots & a_{1r}^{n-1} & a_{1\,r+1}^{n-1} & \cdots & a_{1\,n-1}^{n-1} & a_{1n}^{n-1} \\ 0 & a_{22}^{n-1} & \cdots & a_{2r}^{n-1} & a_{2\,r+1}^{n-1} & \cdots & a_{2\,n-1}^{n-1} & a_{2n}^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{rr}^{n-1} & a_{r\,r+1}^{n-1} & \cdots & a_{r\,n-1}^{n-1} & a_{rn}^{n-1} \\ \vdots & \vdots & \vdots & \vdots & a_{r+1\,r+1}^{n-1} & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & a_{n-1\,n-1}^{n-1} & a_{n-1\,n}^{n-1} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & a_{nn}^{n-1} \end{pmatrix}$$

é triangular superior. Por tanto, se chamamos $P = E_{n-1}P_{n-1}\cdots E_1P_1$, temos que

$$PA = U, \;\; U = A_{n-1}.$$

Exemplo 4.2.2. Tomemos a matriz

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 3 \\ -1 & 3 & 1 \end{pmatrix}.$$

Vexamos como sería cada etapa do método de Gauss.

<u>Etapa 1 do método de Gauss.</u>
<u>Paso I:</u> Como $a_{11} = 2 \neq 0$, tomamos $P_1 = I_3$.
<u>Paso II:</u> $B_1 = P_1A = A$.
Fagamos os ceros a continuación. Tomamos $E_1 = F_{31}(1/2)F_{21}(-1)$ e desta forma

$$A_1 = E_1B_1 = E_1A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 7/2 & 3/2 \end{pmatrix}$$

<u>Etapa 2 do método de Gauss.</u>
<u>Paso I:</u> Como $a^1_{22} = 0$ e $a^1_{32} = 7/2 \neq 0$, tomamos $P_2 = F_{23}$.
<u>Paso II:</u> $B_2 = P_2A_1$.
Fagamos os ceros a continuación. Tomamos $E_2 = I_3$ xa que debaixo do coeficiente b^2_{22} todos os elementos son nulos. Desta forma

$$A_2 = E_2B_2 = P_2A_1 = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 7/2 & 3/2 \\ 0 & 0 & 2 \end{pmatrix}$$

Daquela $U = A_2$, e, ademais, a matriz P, tal que $PA = U$, está dada por

$$P = E_2P_2E_1P_1 = I_3F_{23}F_{31}(1/2)F_{21}(-1)I_3 = F_{23}F_{31}(1/2)F_{21}(-1) = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}.$$

Esta matriz non é triangular inferior, xa que no produto que a define figura a matriz F_{23} que non o é.

Nota 4.2.3. Dos tres tipos de matrices de operacións elementais só as matrices $F_{ij}(\alpha)$ con $i > j$ son triangulares inferiores con uns na diagonal principal. As matrices F_{ij} non son triangulares inferiores nin superiores e as matrices $F_i(\alpha)$ son diagonais, pero o elemento colocado na posición i da diagonal principal é α. Por tanto, se no proceso de Gauss non temos necesidade de intercambiar filas ou de multiplicar unha fila por un escalar non nulo (obviamente tamén distinto de un), a matriz P é igual a

$$P = E_{n-1}E_{n-2}\cdots E_1$$

e isto implica que P é triangular inferior con uns na diagonal principal. Como toda inversa dunha matriz triangular inferior con uns na diagonal principal é triangular inferior con uns na diagonal principal, temos o seguinte:

$$A = LU, \;\; L = P^{-1}$$

onde L é unha matriz triangular inferior con uns na diagonal principal e U é unha matriz triangular superior. Agora ben, nestas condicións xerais non podemos garantir a unicidade de L e de U.

Exemplo 4.2.4. Sexa a matriz

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

e apliquemos o proceso de eliminación de Gauss. Neste exemplo iremos directos ao gran sen especificar con tanto detalle o que se fai en cada etapa. Entón:

$$\overset{F_{41}(-1),\, F_{31}(-1),\, F_{21}(-1)}{\Longrightarrow} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix} \overset{F_{42}(-1),\, F_{32}(-1)}{\Longrightarrow} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$$\overset{F_{43}(-1)}{\Longrightarrow} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Deste xeito, temos que

$$U = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

e

$$P = F_{43}(-1)F_{42}(-1)F_{32}(-1)F_{41}(-1)F_{31}(-1)F_{21}(-1)$$

que é un produto de matrices triangulares inferiores con uns na diagonal principal. Nótese que en ningún momento do proceso tivemos necesidade de intercambiar dúas filas ou de multiplicar unha fila por un escalar non nulo distinto do un.

Entón,

$$\begin{gathered} L = P^{-1} \\ = (F_{43}(-1)F_{42}(-1)F_{32}(-1)F_{41}(-1)F_{31}(-1)F_{21}(-1))^{-1} = \\ F_{21}(-1)^{-1}F_{31}(-1)^{-1}F_{41}(-1)^{-1}F_{32}(-1)^{-1}F_{42}(-1)^{-1}F_{43}(-1)^{-1} = \\ F_{21}(1)F_{31}(1)F_{41}(1)F_{32}(1)F_{42}(1)F_{43}(1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \end{gathered}$$

é unha matriz triangular inferior con uns na diagonal principal tal que

$$A = LU.$$

O teorema que baixo certas condicións garante a existencia dunha única factorización LU é o seguinte:

Teorema 4.2.5. *Sexa $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ e supoñamos que as súas submatrices diagonais*

$$A(r,r) = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}$$

son invertíbeis para todo $r \in \{1, \ldots, n\}$. Entón A admite unha única factorización LU con L unha matriz triangular inferior con uns na diagonal principal e U unha matriz triangular superior invertíbel.

Demostración. Polo visto anteriormente, para garantir a existencia da factorización LU, debemos demostrar que en cada etapa do método non é necesario intercambiar filas. A demostración deste feito farémola por indución. Xa que $a_{11} \neq 0$ entón na primeira etapa da eliminación gaussiana non temos necesidade de intercambiar filas. Supoñamos que isto é así ata a etapa r e comprobemos que tamén é certo para a etapa $r+1$. Nesa etapa temos que

$$A_r = \begin{pmatrix} a_{11}^r & a_{12}^r & \cdots & a_{1r}^r & a_{1r+1}^r & \cdots & a_{1\,n-1}^r & a_{1n}^r \\ 0 & a_{22}^r & \cdots & a_{2r}^r & a_{2\,r+1}^r & \cdots & a_{2\,n-1}^r & a_{2n}^r \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{rr}^r & a_{r\,r+1}^r & \cdots & a_{r\,n-1}^r & a_{rn}^r \\ 0 & 0 & \cdots & 0 & a_{r+1\,r+1}^r & \cdots & a_{r+1\,n-1}^r & a_{r+1\,n}^r \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n-1\,r+1}^r & \cdots & a_{n-1\,n-1}^r & a_{n-1\,n}^r \\ 0 & 0 & \cdots & 0 & a_{n\,r+1}^r & \cdots & a_{n\,n-1}^r & a_{nn}^r \end{pmatrix}$$

e ademais $A_r = E_r P_r E_{r-1} P_{r-1} \cdots E_1 P_1 A$, onde por hipótese de indución non fixemos intercambios de filas ou, o que é o mesmo, $P_r = P_{r-1} = \cdots P_1 = I_n$, o que implica

$$A_r = E_r E_{r-1} \cdots E_1 A$$

con $E_r E_{r-1} \cdots E_1 = (c_{ij})$ matriz triangular inferior con uns na diagonal principal. Daquela:

$$E_r E_{r-1} \cdots E_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ c_{21} & 1 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{r1} & c_{r2} & \cdots & 1 & 0 & \cdots & 0 & 0 \\ c_{r+1\,1} & c_{r+1\,1} & \cdots & c_{r+1\,r} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{n+1\,1} & c_{n+1\,2} & \cdots & c_{n-1\,r} & c_{n-1\,r+1} & \cdots & 1 & 0 \\ c_{n\,1} & c_{n\,2} & \cdots & c_{n\,r} & c_{n\,r+1} & \cdots & c_{n\,n-1} & 1 \end{pmatrix}$$

Descompoñendo por bloques temos que a igualdade

$$A_r = E_r E_{r-1} \cdots E_1 A$$

podémola expresar como

$$\left(\begin{array}{ccccc|cccc} a_{1;1}^r & a_{1\,2}^r & \cdots & a_{1\,r}^r & a_{1\,r+1}^r & a_{1\,r+2}^r & \cdots & a_{1\,n-1}^r & a_{1n}^r \\ 0 & a_{2\,2}^r & \cdots & a_{2\,r}^r & a_{2\,r+1}^r & a_{2\,r+2}^r & \cdots & a_{2\,n-1}^r & a_{2\,n}^r \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{r\,r}^r & a_{r\,r+1}^r & a_{r\,r+2}^r & \cdots & a_{r\,n-1}^r & a_{r\,n}^r \\ 0 & 0 & \cdots & 0 & a_{r+1\,r+1}^r & a_{r+1\,r+2}^r & \cdots & a_{r+1\,n-1}^r & a_{r+1\,n}^r \\ \hline 0 & 0 & \cdots & 0 & a_{r+2\,r+1}^r & a_{r+2\,r+2}^r & \cdots & a_{r+2\,n-1}^r & a_{r+2\,n}^r \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n-1\,r+1}^r & a_{n-1,r+2}^r & \cdots & a_{n-1\,n-1}^r & a_{n-1\,n}^r \\ 0 & 0 & \cdots & 0 & a_{n\,r+1}^r & a_{n,r+2}^r & \cdots & a_{n\,n-1}^r & a_{n\,n}^r \end{array}\right) =$$

$$
\left(\begin{array}{ccccc|cccc}
1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
c_{2\,1} & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
c_{r\,1} & c_{r\,2} & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\
c_{r+1\,1} & c_{r+1\,2} & \cdots & c_{r+1\,r} & 1 & 0 & \cdots & 0 & 0 \\
\hline
c_{r+2\,1} & c_{r+2\,2} & \cdots & c_{r+2\,r} & c_{r+2;r+1} & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
c_{n+1\,1} & c_{n+1\,2} & \cdots & c_{n-1\,r} & c_{n-1\,r+1} & c_{n-1\,r+2} & \cdots & 1 & 0 \\
c_{n\,1} & c_{n\,2} & \cdots & c_{nr} & c_{n\,r+1} & c_{n\,r+2} & \cdots & c_{n\,n-1} & 1
\end{array}\right) \times
$$

$$
\left(\begin{array}{ccccc|cccc}
a_{1\,1} & a_{1\,2} & \cdots & a_{1\,r} & a_{1\,r+1} & a_{1\,r+2} & \cdots & a_{1\,n-1} & a_{1\,n} \\
a_{2\,1} & a_{2\,2} & \cdots & a_{2\,r} & a_{2\,r+1} & a_{2\,r+2} & \cdots & a_{2\,n-1} & a_{2\,n} \\
a_{r\,1} & a_{r\,2} & \cdots & a_{r\,r} & a_{r\,r+1} & a_{r\,r+2} & \cdots & a_{r\,n-1} & a_{r\,n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
a_{r\,1} & a_{r\,2} & \cdots & a_{r\,r} & a_{r\,r+1} & a_{r\,r+2} & \cdots & a_{r\,n-1} & a_{r\,n} \\
\hline
a_{r+1\,1} & a_{r+1\,2} & \cdots & a_{r+1\,r} & a_{r+1\,r+1} & a_{r+1\,r+2} & \cdots & a_{r+1\,n-1} & a_{r+1\,n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
a_{n-1\,1} & a_{n-1\,2} & \cdots & a_{n-1\,r} & a_{n-1\,r+1} & a_{n-1\,r+2} & \cdots & a_{n-1\,n-1} & a_{n-1\,n} \\
a_{n\,1} & a_{n\,2} & \cdots & a_{n\,r} & a_{n\,r+1} & a_{n\,r+2} & \cdots & a_{n\,n-1} & a_{n\,n}
\end{array}\right)
$$

e, por tanto,

$$
\begin{pmatrix}
a^r_{1\,1} & a^r_{1\,2} & \cdots & a^r_{1\,r} & a_{1\,r+1} \\
0 & a^r_{2\,2} & \cdots & a^r_{2\,r} & a^r_{2\,r+1} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & a^r_{r\,r} & a^r_{r\,r+1} \\
0 & 0 & \cdots & 0 & a_{r+1\,r+1}
\end{pmatrix} =
$$

$$
\begin{pmatrix}
1 & 0 & \cdots & 0 & 0 \\
c_{2\,1} & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
c_{r\,1} & c_{r\,2} & \cdots & 1 & 0 \\
c_{r+1\,1} & c_{r+1\,2} & \cdots & c_{r+1\,r} & 1
\end{pmatrix}
\begin{pmatrix}
a_{1\,1} & a_{1\,2} & \cdots & a_{1\,r} & a_{1\,r+1} \\
a_{2\,1} & a_{2\,2} & \cdots & a_{2\,r} & a_{2\,r+1} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
a_{r\,1} & a_{r\,2} & \cdots & a_{r\,r} & a_{r\,r+1} \\
a_{r+1\,1} & a_{r+1\,2} & \cdots & a_{r+1\,r} & a_{r+1\,r+1}
\end{pmatrix}
$$

Aplicando determinantes á anterior igualdade, obtemos que

$$
\begin{vmatrix}
a^r_{1\,1} & a^r_{1\,2} & \cdots & a^r_{1\,r} & a_{1\,r+1} \\
0 & a^r_{2\,2} & \cdots & a^r_{2\,r} & a^r_{2\,r+1} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & a^r_{r\,r} & a^r_{r\,r+1} \\
0 & 0 & \cdots & 0 & a_{r+1\,r+1}
\end{vmatrix} =
$$

$$
\begin{vmatrix}
1 & 0 & \cdots & 0 & 0 \\
c_{2\,1} & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
c_{r\,1} & c_{r\,2} & \cdots & 1 & 0 \\
c_{r+1\,1} & c_{r+1\,2} & \cdots & c_{r+1\,r} & 1
\end{vmatrix}
\begin{vmatrix}
a_{1\,1} & a_{1\,2} & \cdots & a_{1\,r} & a_{1\,r+1} \\
a_{2\,1} & a_{2\,2} & \cdots & a_{2\,r} & a_{2\,r+1} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
a_{r\,1} & a_{r\,2} & \cdots & a_{r\,r} & a_{r\,r+1} \\
a_{r+1\,1} & a_{r+1\,2} & \cdots & a_{r+1\,r} & a_{r+1\,r+1}
\end{vmatrix}
$$

e, entón,

$$
a^r_{1\,1}a^r_{2\,2}\cdots a^r_{r\,r}a^r_{r+1\,r+1} \neq 0
$$

xa que por hipótese a submatriz diagonal $A(r+1, r+1)$ é invertíbel. O verdadeiramente importante de que o produto anterior sexa non nulo é que isto implica que

$$a^r_{r+1\, r+1} \neq 0$$

e, entón, non temos necesidade de intercambiar filas na etapa $r+1$ do método de Gauss.

Probemos agora a unicidade da descomposición. Supoñamos que A admite dúas factorizacións

$$A = L_1 U_1 = L_2 U_2$$

onde L_1 e L_2 son triangulares inferiores con uns na diagonal principal e U_1, U_2 son triangulares superiores invertíbeis. Entón,

$$L_2^{-1} L_1 = U_2 U_1^{-1}.$$

Agora ben, como $L_2^{-1} L_1$ é triangular inferior con uns na diagonal principal e $U_2 U_1^{-1}$ é triangular superior invertíbel, temos que

$$L_2^{-1} L_1 = U_2 U_1^{-1} = I_n$$

ou, o que é o mesmo,

$$L_1 = L_2, \;\; U_1 = U_2.$$

□

Exemplo 4.2.6. Sexa a matriz

$$A = \begin{pmatrix} 2 & -1 & 0 & 1 \\ 4 & -4 & 1 & 5 \\ -2 & 1 & -1 & 0 \\ -2 & 5 & -4 & -1 \end{pmatrix}$$

e apliquemos o proceso de eliminación de Gauss. Entón:

$$\overset{F_{41}(1),\, F_{31}(1),\, F_{21}(-2)}{\Longrightarrow} \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & -2 & 1 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 4 & -4 & 0 \end{pmatrix} \overset{F_{42}(2)}{\Longrightarrow} \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & -2 & 1 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -2 & 6 \end{pmatrix}$$

$$\overset{F_{43}(-2)}{\Longrightarrow} \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & -2 & 1 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

Entón

$$U = \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & -2 & 1 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

e

$$P = F_{43}(-2) F_{42}(2) F_{41}(1) F_{31}(1) F_{21}(-2)$$

é un produto de matrices triangulares inferiores con uns na diagonal principal. Nótese que en ningún momento do proceso tivemos necesidade de intercambiar dúas filas ou de multiplicar unha fila por un escalar non nulo distinto do un, o que garante a existencia e unicidade da factorización LU como a citada no teorema anterior.

Daquela,

$$L = P^{-1} =$$
$$(F_{43}(-2) F_{42}(2) F_{41}(1) F_{31}(1) F_{21}(-2))^{-1} =$$

$$F_{21}(-2)^{-1}F_{31}(1)^{-1}F_{41}(1)^{-1}F_{42}(2)^{-1}F_{43}(-2)^{-1} =$$

$$F_{21}(2)F_{31}(-1)F_{41}(-1)F_{42}(-2)F_{43}(2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & -2 & 2 & 1 \end{pmatrix}.$$

4.3. Problemas propostos

1. Determine a solución xeral dos sistemas homoxéneos seguintes:

$$a) \left\{ \begin{array}{l} 2x + y + z = 0 \\ 4x + 2y + z = 0 \\ 6x + 3y + z = 0 \\ 8x + 4y + z = 0 \end{array} \right. . \qquad b) \left\{ \begin{array}{l} 2x + y + z = 0 \\ 4x + 2y + z = 0 \\ 6x + 3y + z = 0 \\ 8x + 5y + z = 0 \end{array} \right. .$$

$$c) \left\{ \begin{array}{l} x + 2y + z + 2t = 0 \\ 2x + 4y + z + 3t = 0 \\ 3x + 6y + z + 4t = 0 \end{array} \right. . \qquad d) \left\{ \begin{array}{l} 2x + 2y + 2z = 0 \\ 2x + 5y + 7z = 0 \\ 3x + 6y + 6z = 0 \end{array} \right. .$$

2. Determine a solución xeral dos sistemas seguintes:

$$a) \left\{ \begin{array}{l} x + 2y + z + 2t = 3 \\ 2x + 4y + z + 3t = 4 \\ 3x + 6y + z + 4t = 5 \end{array} \right. . \qquad b) \left\{ \begin{array}{l} x + y + 2z + 2t + u = 1 \\ 2x + 2y + 4z + 4t + 3u = 1 \\ 2x + 2y + 4z + 4t + 2u = 2 \\ 3x + 5y + 8z + 6t + 5u = 3 \end{array} \right. .$$

$$c) \left\{ \begin{array}{l} x + y - z = 1 \\ 3x + 4y - z = 1 \\ x + 2y + z = -1 \end{array} \right. . \qquad d) \left\{ \begin{array}{l} x - y + z = 1 \\ x + 2z = 0 \\ 3x - 2y + 4z + t + u = 0 \end{array} \right. .$$

3. Resolva a ecuación matricial

$$\begin{pmatrix} 3 & 1 \\ -1 & 2 \end{pmatrix} X - X \begin{pmatrix} 1 & 4 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ 5 & 4 \end{pmatrix}.$$

4. Estude o sistema de ecuacións lineares $Ax = b$ onde:

   a)

$$A = \begin{pmatrix} 1 & 2 & -1 & -1 \\ -1 & \alpha - 4 & \alpha - 1 & 5 - 2\alpha \\ 2 & 6 - \alpha & \beta - \alpha & 2\alpha + \beta - 6 \\ -1 & \alpha - 4 & \alpha - 2\beta - 1 & 2 - \beta - 2\alpha \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \end{pmatrix}.$$

   b)

$$A = \begin{pmatrix} 1 & \beta & \alpha & \beta \\ 2 & 3\beta & \alpha & 2\beta \\ 1 & \beta & 2\alpha & 2\beta \\ 1 & 2\beta & 0 & 2\beta \end{pmatrix} \quad b = \begin{pmatrix} \alpha + \beta + 1 \\ 3\alpha + 2\beta + 1 \\ 2\beta + 2 \\ \alpha + 2\beta \end{pmatrix}.$$

5. Ache o conxunto de solucións dos sistemas $Ax = b$ e $Bx = b$, onde

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

6. Discuta e resolva, nos casos en que sexa posíbel, o seguinte sistema de ecuacións lineares homoxéneo segundo os distintos valores de $\alpha \in \mathbb{R}$.
$$\begin{aligned} (\alpha+1)x + (\alpha+2)z &= 0, \\ (\alpha+1)y + (\alpha+2)t &= 0, \\ \alpha x + (\alpha-1)z &= 0, \\ \alpha y + (\alpha-1)t &= 0. \end{aligned}$$
7. Discuta segundo os valores de α, $\beta \in \mathbb{R}$ o seguinte sistema e calcule a solución nos casos en que sexa posíbel.
$$\begin{pmatrix} \alpha & 1 & 1 \\ 1 & \beta & 0 \\ 1 & 0 & \beta \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$
8. Discuta segundo os valores de α, $\beta \in \mathbb{R}$ e calcule a solución, nos casos en que sexa posíbel, dos seguintes sistemas de ecuacións lineares.
$$a) \left\{ \begin{array}{l} \alpha x + \beta y + z = 1 \\ x + \alpha\beta y + z = \beta \\ x + \beta y + z = 1 \end{array} \right. . \qquad b) \left\{ \begin{array}{l} \alpha x + y + z = 1 \\ x + \beta y = 1 \\ x + \beta z = 1 \end{array} \right. .$$
9. Ache o conxunto de solucións do sistema
$$\begin{pmatrix} i & 1+i & 1 \\ 2i & -2-i & i \\ -1 & i & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1+i \\ 2 \\ -2-i \end{pmatrix}.$$
10. Considéranse as matrices
$$A = \begin{pmatrix} 2 & -1 & 0 & 1 \\ 4 & -4 & 1 & 5 \\ -2 & 1 & -1 & 0 \\ -2 & 5 & -4 & -1 \end{pmatrix}, \; B = \begin{pmatrix} 2 & 1 & 0 & 1 \\ -2 & -2 & 1 & -2 \\ 2 & 2 & 1 & 4 \\ -2 & -2 & -1 & -5 \end{pmatrix}.$$
   a) Demostre que A e B admiten factorización LU e calcule ditas factorizacións.
   b) Utilice o apartado anterior para resolver os sistemas $Ax = b$, $Bx = c$ onde $b = (-5, -14, 1, 1)^t$ e $c = (0, 3, -3, 4)^t$.
11. Sexa o sistema de ecuacións lineares $Ax = b$ onde
$$A = \begin{pmatrix} 2 & 0 & -1 & 1 \\ -4 & -1 & 2 & 0 \\ -2 & -2 & -1 & 1 \\ -6 & 3 & 3 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ -5 \\ 2 \\ -9 \end{pmatrix}.$$
   a) Discuta e resolva, se é posíbel, o sistema utilizando o método de Gauss.
   b) Demostre que A admite factorización LU e calcule dita factorización.
   c) Resolva, se é posíbel, o sistema utilizando a factorización LU.
12. Considérase a matriz real
$$A = \begin{pmatrix} -1 & 0 & -2 & 0 \\ -1 & -1 & -2 & 1 \\ -1 & -2 & -1 & 4 \\ -1 & -2 & 1 & 7 \end{pmatrix}.$$
   a) Estude se existe a factorización LU da matriz A.
   b) En caso afirmativo, ache dita factorización e ademais:
      *b*1) Calcule o determinante da matriz A.
      *b*2) Atope a matriz inversa de A.

$b3)$ Resolva o sistema de ecuacións lineares $Ax = b$ sendo $b^t = (-5, -6, -3, 4)$.

13. Sexa $A = (a_{ij}) \in \mathcal{M}_{4\times 4}(\mathbb{R})$ a matriz definida por

$$a_{ij} = (-1)^{\text{máx}\{i,j\}},\ 1 \leq i, j \leq 4.$$

Probe que A admite factorización LU e calcule dita factorización.

14. Calcule, se é posíbel, a factorización LU das seguintes matrices:

$$A = \begin{pmatrix} 2 & 1 & 4 \\ 8 & -3 & 2 \\ 4 & 11 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & -2 & 6 & -5 \\ 24 & -12 & 41 & -39 \\ -27 & 18 & -62 & 54 \\ 9 & 14 & 15 & -47 \end{pmatrix},$$

$$C = \begin{pmatrix} 2 & 4 & -2 \\ 1 & 3 & 1 \\ -2 & -1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} -2 & -2 & -2 & 5 \\ -2 & -1 & -2 & 4 \\ -2 & -1 & -1 & 4 \\ -3 & -2 & -2 & 6 \end{pmatrix}.$$

15. Sexan $A, B \in \mathcal{M}_{2\times 3}(\mathbb{R})$, $C \in \mathcal{M}_{2\times 2}(\mathbb{R})$ as matrices

$$A = \begin{pmatrix} -2 & -1 & -1 \\ 3 & 2 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} -2 & -1 \\ 6 & 3 \end{pmatrix}.$$

$a)$ Ache todas as matrices X que son solucións da ecuación matricial

$$CX + AB^t = BB^t.$$

$b)$ Estude se é posíbel atopar unha solución da ecuación anterior que verifique

$$X + X^t = 0.$$

CAPíTULO 5

Espazos vectoriais e aplicacións lineares

Neste capítulo introduciremos o concepto de espazo vectorial o cal dará lugar ao ecosistema no que vivirán de forma natural gran parte das nocións estudadas nos capítulos anteriores.

Comezaremos introducindo as definicións de espazo e subespazo vectorial sobre un corpo $\mathbb{K}$. Despois de ver varios exemplos significativos, pasaremos a definir as nocións de combinación linear dun conxunto finito de vectores e de clausura linear dun sistema de vectores. Neste punto quedará patente a estreita relación que existe entre ambos conceptos grazas ao seguinte resultado: se V é un $\mathbb{K}$-espazo vectorial e $T \subset V$ non é baleiro, a clausura linear de T coincide co conxunto de todas as combinacións lineares de vectores de T. Doutra banda, grazas á clausura linear, introduciremos a definición de sistema xerador, ou sistema de xeradores, dun subespazo vectorial e veremos as súas propiedades fundamentais.

Na seguinte parte do capítulo estudaranse as nocións de dependencia e independencia linear e de sistema libre ou ligado. Tamén neste punto introducirase a definición de rango dun sistema de vectores T dun espazo vectorial V, e verase que, se $V = \mathbb{K}^n$ e $T = \{v_1, \cdots, v_r\}$, o rango de T é o rango da matriz que resulta ao escribir os vectores de T ben en fila, ben en columna.

Grazas ao mencionado no parágrafo anterior poderemos falar de bases e dimensión para un espazo vectorial. A partir deste intre asumirase que os espazos vectoriais cos que traballamos son de tipo finito. Agora ben, é importante resaltar que tamén existen espazos vectoriais que, aínda que non son deste tipo, teñen o seu interese matemático (por exemplo, os espazos de funcións definidas nun intervalo).

Unha vez fixadas as definicións de base e dimensión e estudadas as súas principais propiedades, pasaremos á noción de coordenadas dun vector respecto a unha base e demostraremos que o seu cálculo redúcese á solución dun sistema de ecuacións lineares. Veremos que dada unha base dun espazo vectorial e un vector do devandito espazo, sempre se poden calcular as súas coordenadas respecto da base e que estas, xunto coa base, determinan completamente ao vector. Dado que as coordenadas dan lugar ao que se chama vector de coordenadas e este é un vector dun espazo $\mathbb{K}^n$, recalcarase ao lector que esencialmente, cando traballamos con espazos vectoriais de tipo finito, unha vez fixada unha base, é coma se trasladásemos todos os problemas a un espazo do tipo $\mathbb{K}^n$. Na parte final deste apartado verase o cambio de coordenadas asociado a un cambio de base e a súa interpretación matricial. Isto dará lugar á aparición da matriz de cambio de base que posteriormente poderemos utilizar como exemplo de matriz asociada a unha aplicación linear.

A parte final do capítulo dedícase ao estudo das aplicacións lineares e as súas principais propiedades. Comezaremos introducindo a noción de aplicación linear para posteriormente ver unha serie de exemplos dos que destacaremos especialmente aqueles que relacionan matrices e aplicacións lineares. Despois disto pasaremos a ver cal é a información mínima que necesitamos para ter completamente caracterizada unha aplicación linear. A continuación afondaremos na relación existente entre matrices e aplicacións lineares, grazas á introdución da noción de matriz asociada a unha aplicación linear. Estudaremos as principais propiedades das matrices asociadas comprobando que dúas matrices asociadas á mesma aplicación linear son equivalentes e, por tanto, teñen o mesmo rango. Tamén se clarificará a relación inversa, chegando a enunciar o seguinte resultado: dúas matrices son equivalentes, se, e só se, son matrices asociadas a unha mesma aplicación linear.

Neste capítulo os últimos conceptos importantes que se introducirán son os de núcleo $\mathrm{Ker}(f)$ e imaxe $\mathrm{Im}(f)$ dunha aplicación linear f. Demostrarase que calcular o núcleo redúcese a resolver un sistema homoxéneo e que calcular a dimensión da imaxe dunha aplicación linear redúcese a calcular o rango de calquera das súas matrices asociadas. Finalmente comprobaremos que para o núcleo e a imaxe dunha aplicación linear f con dominio V cúmprese a seguinte fórmula que liga as súas dimensións:

$$\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(\mathrm{Ker}(f)) + \dim_{\mathbb{K}}(\mathrm{Im}(f)).$$

5.1. Espazos e subespazos vectoriais. Sistemas de xeradores

Definición 5.1.1. Un espazo vectorial sobre o corpo $\mathbb{K}$ (ou un $\mathbb{K}$-espazo vectorial) é un conxunto non baleiro V no cal se teñen definidas dúas operacións chamadas suma (denotada por $+$) e produto por escalares (denotado por $\cdot$)

$$\begin{array}{rcl} V \times V & \to & V\,, \\ (u,v) & \mapsto & u+v \end{array} \qquad \begin{array}{rcl} \mathbb{K} \times V & \to & V\,, \\ (\alpha,v) & \mapsto & \alpha \cdot v \end{array}$$

satisfacendo as seguintes propiedades:

1) $u + (v + w) = (u + v) + w,\ \forall\, u, v, w \in V$.
2) $u + v = v + u,\ \forall\, u, v \in V$.
3) $\exists\, \theta_V \in V$ tal que $\theta_V + u = u,\ \forall\, u \in V$.
4) $\forall\, v \in V, \exists\, -v \in V$ tal que $v + (-v) = \theta_V$.
5) $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v,\ \forall\, u, v \in V,\ \alpha \in \mathbb{K}$.
6) $(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u,\ \forall\, \alpha, \beta \in \mathbb{K},\ u \in V$.
7) $1 \cdot u = u,\ \forall\, u \in V$.
8) $(\alpha\beta) \cdot u = \alpha \cdot (\beta \cdot u),\ \forall\, \alpha, \beta \in \mathbb{K},\ u \in V$.

Os elementos de V adóitanse chamar vectores e os elementos de $\mathbb{K}$, escalares. Ao vector θ_V chámaselle vector nulo ou vector cero de V. Dada a amplitude de exemplos existentes de espazo vectorial, veremos que a palabra vector acabará designando obxectos distintos dos que usualmente chamamos por ese nome. Finalmente, por comodidade, o produto por escalares denotarémolo a partir de agora sen o punto, isto é, $\alpha \cdot v = \alpha v$.

Vexamos agora os exemplos de espazos vectoriais que manexaremos ao longo do libro.

Exemplo 5.1.2. 1) Un corpo $\mathbb{K}$ é un $\mathbb{K}$-espazo vectorial coa suma e o produto de escalares. Neste caso $\theta_{\mathbb{K}} = 0$. Por tanto, $\mathbb{R}$ é un $\mathbb{R}$-espazo vectorial e $\mathbb{C}$ é un $\mathbb{C}$-espazo vectorial.

2) O corpo $\mathbb{C}$ é un $\mathbb{R}$-espazo vectorial coa suma de números complexos e o produto por escalares dado por

$$\alpha(a + bi) = \alpha a + \alpha bi$$

para todo α número real e $a + bi$ número complexo. Obviamente, neste caso, $\theta_{\mathbb{C}} = 0 = 0 + 0i$.

3) O conxunto

$$\mathbb{K}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \ \middle|\ x_i \in \mathbb{K}\ \forall\, i \in \{1, \cdots, n\} \right\}$$

é un $\mathbb{K}$ espazo vectorial coas operacións de suma de vectores

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

e produto por escalares

$$\alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

Neste espazo vectorial $\theta_{\mathbb{K}^n}$ é o vector que ten todas as súas compoñentes nulas. Logo, $\mathbb{R}^n$ é un $\mathbb{R}$-espazo vectorial e $\mathbb{C}^n$ é un $\mathbb{C}$-espazo vectorial. Tamén se cumpre que $\mathbb{C}^n$ é un $\mathbb{R}$-espazo vectorial.

4) O conxunto $\mathcal{M}_{m\times n}(\mathbb{K})$ das matrices de m filas e n columnas con coeficientes en $\mathbb{K}$ é un exemplo de $\mathbb{K}$-espazo vectorial, coas operacións de suma de matrices e produto de escalares por matrices definidas no capítulo anterior. Neste caso, $\theta_{\mathcal{M}_{m\times n}(\mathbb{K})} = \Theta_{m,n}$.

5) Se denotamos por $\Pi_n(\mathbb{K})$ o conxunto de polinomios de grao menor ou igual que n con coeficientes en $\mathbb{K}$, obtemos un novo exemplo de $\mathbb{K}$-espazo vectorial onde, se cada elemento de $\Pi_n(\mathbb{K})$ escríbese como

$$p(x) = a_0 + a_1x + \cdots + a_nx^n,$$

a suma defínese por

$$p(x) + q(x) = (a_0 + a_1x + \cdots + a_nx^n) + (b_0 + b_1x + \cdots + b_nx^n) =$$
$$(a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n,$$

e o produto por escalares por

$$\alpha p(x) = \alpha a_0 + \alpha a_1x + \cdots + \alpha a_nx^n.$$

Finalmente, o vector cero $\theta_{\Pi_n(\mathbb{K})}$ é o polinomio con todos os coeficientes nulos.

6) Neste exemplo veremos unha estrutura non común de espazo vectorial. Sexa

$$V = \{x \in \mathbb{R} : x > 0\}.$$

Entón, V é un $\mathbb{R}$-espazo vectorial coas operacións de suma

$$x \boxplus y = xy$$

e produto por escalares

$$\alpha \boxdot x = x^{\alpha}.$$

Neste caso, o vector nulo é $\theta_V = 1$.

Proposición 5.1.3. *Supoñamos que V é un $\mathbb{K}$-espazo vectorial. Cúmprese o seguinte:*

1) *O vector θ_V é único.*
2) *Para todo $v \in V$ o vector $-v$ é único.*
3) *Para todo $\alpha \in \mathbb{K}$, $\alpha\theta_V = \theta_V$.*
4) *Para todo $v \in V$, $0v = \theta_V$.*
5) *Se $\alpha v = \theta_V$, entón $\alpha = 0$ ou $v = \theta_V$.*
6) *Para todo $\alpha \in \mathbb{K}$ e $v \in V$, $(-\alpha)v = -\alpha v = \alpha(-v)$.*
7) *Se $\alpha v = \alpha u$ e $\alpha \neq 0$, entón $v = u$.*
8) *Se $\alpha u = \beta u$ e $u \neq \theta_V$, entón $\alpha = \beta$.*

Demostración. A demostración destas propiedades déixase como exercicio. □

Definición 5.1.4. Sexa V un $\mathbb{K}$-espazo vectorial e U un subconxunto non baleiro de V. Diremos que U é un subespazo vectorial de V, se se cumpre:

1) Para todos $u, v \in U$ tense que $u - v \in U$.
2) Para todo $\alpha \in \mathbb{K}$ e $u \in U$ tense que $\alpha u \in U$.

Nesta definición podemos cambiar a primeira condición por

3) Para todos $u, v \in U$ tense que $u + v \in U$

xa que, como se pode comprobar de forma sinxela, 1) e 2) son equivalentes a 2) e 3).

As dúas condicións que aparecen na definición tamén son equivalentes a

4) Para todo $\alpha, \beta \in \mathbb{K}$ e $u, v \in U$ tense que $\alpha u + \beta v \in U$.

Nota 5.1.5. A anterior definición implica que U é un subespazo vectorial de V se ao restrinxir a suma de vectores e o produto por escalares a U, este conxunto é un $\mathbb{K}$-espazo vectorial. Isto implica que $\theta_V \in U$. Daquela, todo subconxunto de V que non conteña ao vector θ_V non poderá ser un subespazo de V.

Exemplos 5.1.6. 1) Tomemos $V = \mathbb{R}^2$ e $\mathbb{K} = \mathbb{R}$. O subconxunto

$$U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in V \ \mid \ x + y = 3 \right\}$$

non é un subespazo de V, xa que

$$\theta_V = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \notin U.$$

2) Tomemos $V = \mathbb{R}^2$ e $\mathbb{K} = \mathbb{R}$. O subconxunto

$$U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in V \ \mid \ x \leq 0 \right\}$$

non é un subespazo de V, xa que

$$-\begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix} \notin U.$$

3) Tomemos $V = \mathbb{R}^3$ e $\mathbb{K} = \mathbb{R}$. O subconxunto

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V \ \mid \ xy = 0 \right\}$$

non é un subespazo de V, xa que

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

son vectores de U e

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \notin U.$$

4) Tomemos $V = \mathbb{C}^2$ e $\mathbb{K} = \mathbb{R}$. O subconxunto

$$U = \left\{ \begin{pmatrix} z \\ z' \end{pmatrix} \in V \ \mid \ \mathrm{Im}(z) = \mathrm{Im}(z') \right\}$$

é un subespazo de V.

5) Tomemos $V = \mathbb{C}^2$ e $\mathbb{K} = \mathbb{C}$. O subconxunto

$$U = \left\{ \begin{pmatrix} z \\ z' \end{pmatrix} \in V \ \mid \ \mathrm{Im}(z) = \mathrm{Im}(z') \right\}$$

non é un subespazo de V, xa que

$$\begin{pmatrix} 1+i \\ 3+i \end{pmatrix} \in U$$

e

$$i\begin{pmatrix} 1+i \\ 3+i \end{pmatrix} = \begin{pmatrix} -1+i \\ -1+3i \end{pmatrix} \notin U.$$

6) Tomemos $V = \mathbb{R}^4$ e $\mathbb{K} = \mathbb{R}$. O subconxunto

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in V \ \mid \ x + 3y + z + t = 0 \right\}$$

é un subespazo de V.

7) Tomemos $V = \mathcal{M}_{2\times 2}(\mathbb{R})$ e $\mathbb{K} = \mathbb{R}$. O subconxunto $U = \{A \in V \ \mid \ \det(A) = 0\}$ non é un subespazo de V, xa que

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

son elementos de U e $A + B = I_2 \notin U$.

Do mesmo xeito, $H = \{A \in V \ \mid \ \operatorname{rang}(A) = 1\}$ non é un subespazo de V, xa que para as matrices anteriores temos que $\operatorname{rang}(A+B) = \operatorname{rang}(I_2) = 2$ e $\operatorname{rang}(A) = \operatorname{rang}(B) = 1$.

O subconxunto $D = \{A \in V \ \mid \ \det(A) \neq 0\}$ tampouco é un subespazo de V, xa que $\theta_V = \Theta_{2,2} \notin D$.

8) Tomemos $V = \mathcal{M}_{n\times n}(\mathbb{K})$ e $\mathbb{K}$. Os subconxuntos das matrices simétricas e antisimétricas son dous exemplos de subespazos.

9) Tomemos $V = \mathbb{K}^n$ e $\mathbb{K}$. Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Consideremos o sistema homoxéneo de m ecuacións lineares e n incógnitas, $Ax = \theta_{\mathbb{K}^n}$. Entón o seu conxunto de solucións sempre é un subespazo de V (o exemplo 6) e un caso particular desta situación máis xeral).

10) Tomemos $V = \Pi_n(\mathbb{R})$ e $\mathbb{K} = \mathbb{R}$. O subconxunto $U = \{p(x) \in V \ \mid \ p'(x) = 0\}$ é un subespazo de V.

11) Todo $\mathbb{K}$-espazo vectorial V sempre ten dous subespazos triviais que son o propio V e o subespazo nulo $U = \{\theta_V\}$.

Notación 5.1.7. A partir deste intre, dado un $\mathbb{K}$-espazo vectorial V, con

$$\mathcal{S}(V)$$

denotaremos ao conxunto de todos os subespazos de V.

Proposición 5.1.8. *Sexa V un $\mathbb{K}$-espazo vectorial e sexan U_1 e $U_2 \in \mathcal{S}(V)$. Cúmprese que $U_1 \cap U_2 \in \mathcal{S}(V)$.*

Demostración. Sexan $\alpha, \beta \in \mathbb{K}$ e $u, v \in U_1 \cap U_2$. Entón, $\alpha u + \beta v \in U_1$, xa que $u, v \in U_1$ e U_1 é un subespazo de V. Do mesmo xeito, $\alpha u + \beta v \in U_2$, xa que $u, v \in U_2$ e U_2 é un subespazo de V. Por tanto, $\alpha u + \beta v \in U_1 \cap U_2$ e, como consecuencia, $U_1 \cap U_2 \in \mathcal{S}(V)$. □

Nota 5.1.9. Aplicando o razoamento anterior, obtemos que a intersección de calquera familia de subespazos de V é sempre un subespazo de V que é distinto do baleiro (lémbrese que o vector nulo está en todo subespazo). Coa unión non temos tanta sorte xa que a unión de subespazos non sempre é un subespazo. Por exemplo, sexan os subespazos

$$U_1 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in V \ / \ x = 0 \right\}, \quad U_2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in V \ \mid \ y = 0 \right\}$$

do $\mathbb{R}$-espazo vectorial $V = \mathbb{R}^2$. Logo,

$$v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in U_1 \subset U_1 \cup U_2, \quad v_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in U_2 \subset U_1 \cup U_2$$

e
$$v = v_1 + v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin U_1 \cup U_2$$
xa que $v \notin U_1$ e $v \notin U_2$.

Definición 5.1.10. Sexa V un $\mathbb{K}$- espazo vectorial. Un sistema de vectores de V é calquera subconxunto T de V.

Definición 5.1.11. Sexa V un $\mathbb{K}$-espazo vectorial e sexa $T \subset V$ non baleiro. Unha combinación linear de vectores de T é una expresión da forma
$$\alpha_1 v_1 + \cdots + \alpha_s v_s$$
onde $\alpha_i \in \mathbb{K}$ e $v_i \in T$ para todo $i \in \{1, \ldots, s\}$.

Nótese que o vector cero de V é sempre combinación linear de calquera sistema non baleiro de vectores T (só temos que tomar todos escalares nulos)

Exemplo 5.1.12. Tomemos o $\mathbb{R}$-espazo vectorial $V = \mathbb{R}^3$. Sexa o sistema de vectores de V dado por
$$T = \left\{ \begin{pmatrix} 1 \\ 6 \\ -1 \end{pmatrix}, \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix} \right\}.$$

Daquela,
$$v = 3 \begin{pmatrix} 1 \\ 6 \\ -1 \end{pmatrix} + (-1) \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 14 \\ -5 \end{pmatrix}$$
é unha combinación linear de vectores de T.

Se, por exemplo, tomamos o vector
$$u = \begin{pmatrix} 4 \\ -1 \\ 8 \end{pmatrix},$$
poderemos garantir que é combinación linear dos vectores de T, se existen escalares x e y, tales que
$$u = x \begin{pmatrix} 1 \\ 6 \\ -1 \end{pmatrix} + y \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix}$$
ou, equivalentemente, o sistema
$$\begin{pmatrix} 1 & 6 \\ 6 & 4 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ 8 \end{pmatrix}$$
ten solución.

Tomemos a matriz ampliada do sistema
$$(A|u) = \begin{pmatrix} 1 & 6 & 4 \\ 6 & 4 & -1 \\ -1 & 2 & 8 \end{pmatrix}$$
e fagamos operacións elementais para determinar o seu rango xunto co rango da matriz do sistema.
$$(A|u) \overset{F_{31}(1),\, F_{21}(-6)}{\Longrightarrow} \begin{pmatrix} 1 & 6 & 4 \\ 0 & -32 & -25 \\ 0 & 8 & 12 \end{pmatrix} \overset{F_{23}}{\Longrightarrow} \begin{pmatrix} 1 & 6 & 4 \\ 0 & 8 & 12 \\ 0 & -32 & -25 \end{pmatrix}$$

$$\stackrel{F_{32}(4)}{\Longrightarrow}\begin{pmatrix} 1 & 6 & 4 \\ 0 & 8 & 12 \\ 0 & 0 & 23 \end{pmatrix}.$$

Logo, como $\operatorname{rang}(A|u) = 3 > \operatorname{rang}(A) = 2$, temos un sistema incompatíbel e u non é combinación linear dos vectores de T.

Definición 5.1.13. Sexa V un $\mathbb{K}$- espazo vectorial e sexa $T \subset V$. Chámase clausura linear de T ao subespazo máis pequeno de V que contén a T. Denotarémola como

$$\langle T \rangle.$$

Nota 5.1.14. En primeiro lugar, debemos aclarar que por subespazo máis pequeno de V que contén a T entendemos o seguinte:

Se U é un subespazo de V tal que $T \subset U$, entón $\langle T \rangle \subset U$.

É importante resaltar que a clausura linear sempre existe xa que $T \subset V$ e V é un subespazo.

Doutra banda, se $T = \emptyset$, entón $\langle T \rangle = \{\theta_V\}$. Ademais, se U é un subespazo $U = \langle U \rangle$, é máis, U é un subespazo, se, e só se, $U = \langle U \rangle$. Finalmente, é obvio que, se $S \subset T$, entón $\langle S \rangle \subset \langle T \rangle$.

A forma práctica de calcular a clausura linear dun conxunto de vectores dedúcese do seguinte resultado.

Proposición 5.1.15. *Sexa V un $\mathbb{K}$-espazo vectorial e sexa $T \subset V$ non baleiro. A clausura linear de T coincide co conxunto de todas as combinacións lineares de vectores de T.*

Demostración. É evidente que, se restamos dúas combinacións lineares de vectores de T, obtemos unha nova combinación linear de vectores de T. Tamén, se multiplicamos unha combinación linear de vectores de T por un escalar, obtemos unha nova combinación linear de vectores de T. Entón, o conxunto de todas as combinacións lineares de vectores de T é un subespazo. Agora ben, se U é un subespazo que contén a T, por ser subespazo, contén a todas as combinacións lineares de vectores de T. Isto último implica que o conxunto de todas as combinacións lineares de vectores de T é o subespazo máis pequeno de V que contén a T ou, o que é o mesmo, coincide con $\langle T \rangle$. □

Exemplos 5.1.16. 1) Tomemos o $\mathbb{R}$-espazo vectorial $V = \mathbb{R}^3$. Sexa T o sistema de vectores

$$T = \left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Como todo vector $v \in \mathbb{R}^3$ cumpre que

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = v_1 e_1 + v_2 e_2 + v_3 e_3,$$

tense a igualdade $\langle T \rangle = \mathbb{R}^3$.

Neste mesmo espazo, se tomamos

$$S = \left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

temos que

$$\langle S \rangle = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V \;\middle|\; \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha e_1 + \beta e_3, \alpha, \beta \in \mathbb{R} \right\} =$$

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V \ / \ \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \\ \beta \end{pmatrix}, \alpha, \beta \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V \ \mid \ y = 0 \right\}.$$

2) Tomemos agora o $\mathbb{R}$-espazo vectorial $V = \mathcal{M}_{2\times 2}(\mathbb{R})$. Sexa U o subespazo de V dado polas matrices de traza nula. Unha matriz

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

pertence a U, se $\operatorname{tr}(A) = a_{11} + a_{22} = 0$. Logo,

$$A \in U \Leftrightarrow A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & -a_{11} \end{pmatrix} \Leftrightarrow$$

$$A = a_{11} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + a_{12} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + a_{21} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Por tanto,

$$U = \langle \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\} \rangle.$$

3) Finalmente, tomemos o $\mathbb{R}$-espazo vectorial $V = \Pi_3(\mathbb{R})$. Sexa o subespazo

$$U = \{p(x) \in V \ \mid \ p(1) = p''(-1) = 0\}.$$

Se eliximos un polinomio xenérico $p(x)$ de grao menor ou igual que tres, será da forma

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3.$$

Entón,

$$p(1) = 0 \Leftrightarrow a_0 + a_1 + a_2 + a_3 = 0$$

e

$$p''(-1) = 0 \Leftrightarrow 2a_2 - 6a_3 = 0 \Leftrightarrow a_2 = 3a_3.$$

Dada esta última igualdade, se a aplicamos en $a_0 + a_1 + a_2 + a_3 = 0$, implica que

$$a_0 = -a_1 - 4a_3.$$

Por tanto,

$$p(x) \in U \Leftrightarrow p(x) = -a_1 - 4a_3 + a_1 x + 3a_3 x^2 + a_3 x^3 = a_1(-1 + x) + a_3(-4 + 3x^2 + x^3)$$

e isto garante que

$$U = \langle \{-1 + x, \ -4 + 3x^2 + x^3\} \rangle.$$

Definición 5.1.17. Sexa V un $\mathbb{K}$- espazo vectorial e sexa U un subespazo de V. Diremos que $T \subset V$ é un sistema de xeradores para U se $\langle T \rangle = U$.

Como se pode deducir da definición de clausura linear, $\langle T \rangle = U$, se, e só se, todo vector de U é combinación linear de vectores de T.

Dous sistemas de vectores S e T de V chámanse equivalentes se $\langle S \rangle = \langle T \rangle$. Isto significa que todo vector de T é combinación linear dos vectores de S e todo vector de S é combinación linear dos vectores de T.

Exemplo 5.1.18. Se tomamos o $\mathbb{R}$-espazo vectorial $V = \mathcal{M}_{2\times 2}(\mathbb{R})$ e U o subespazo de V dado polas matrices de traza nula, probamos nos exemplos anteriores que

$$U = \langle \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\} \rangle.$$

Por tanto, o conxunto

$$T = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$$

é un sistema de xeradores para U.

Finalmente, é doado comprobar que

$$S = \left\{ \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \right\},$$

cumpre que $\langle S \rangle = \langle T \rangle$ e, como consecuencia, S e T son equivalentes.

Definición 5.1.19. Sexa V un $\mathbb{K}$-espazo vectorial e sexan U_1 e U_2 dous subespazos de V. Defínese o subespazo vectorial suma de U_1 e U_2 como o subespazo máis pequeno que contén a ambos.

É sinxelo comprobar que:

$$U_1 + U_2 = \{u_1 + u_2 \ / \ u_1 \in U_1, \ u_2 \in U_2\}.$$

Ademais, se T_1 é un sistema de xeradores de U_1 e T_2 é un sistema de xeradores de U_2, tense que $T_1 \cup T_2$ é un sistema de xeradores de $U_1 + U_2$.

Exemplo 5.1.20. Se tomamos o $\mathbb{R}$-espazo vectorial $V = \mathcal{M}_{2\times 2}(\mathbb{R})$ e

$$U_1 = \langle \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\} \rangle,$$

$$U_2 = \langle \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right\} \rangle,$$

tense que

$$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\}$$

é un sistema de xeradores de $U_1 + U_2$. Agora ben, S é equivalente a

$$T = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$

e isto implica que $U_1 + U_2 = \mathcal{M}_{2\times 2}(\mathbb{R})$.

5.2. Independencia linear. Bases e dimensión

Definición 5.2.1. Sexan V un $\mathbb{K}$-espazo vectorial e $S = \{v_1, \dots, v_s\}$ un sistema de vectores de V. Diremos que S é libre ou linearmente independente, se, dados $\alpha_1, \cdots, \alpha_s \in \mathbb{K}$ tales que

$$\alpha_1 v_1 + \cdots + \alpha_s v_s = \theta_V,$$

cúmprese que $\alpha_1 = \cdots = \alpha_s = 0$.

En caso contrario, dirase que o sistema é ligado ou linearmente dependente. Por tanto, cando o sistema é ligado, existen escalares non todos nulos $\alpha_1, \cdots, \alpha_s \in \mathbb{K}$ para os cales

$$\alpha_1 v_1 + \cdots + \alpha_s v_s = \theta_V.$$

Isto último implica que alomenos un dos vectores de S é combinación linear dos restantes.

Exemplos 5.2.2. 1) Tomemos o $\mathbb{R}$-espazo vectorial $V = \mathbb{R}^3$. Sexa S o sistema de vectores

$$S = \left\{ v_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, v_3 = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} \right\}.$$

Tense que

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = \theta_V \Leftrightarrow \alpha_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \alpha_3 \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow$$

$$\left\{\begin{array}{l}\alpha_1+\alpha_2+5\alpha_3=0\\ 2\alpha_1+6\alpha_3=0\\ \alpha_1+2\alpha_2+7\alpha_3=0\end{array}\right. \Leftrightarrow \begin{pmatrix}1&1&5\\2&0&6\\1&2&7\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

e, por tanto, se o anterior sistema homoxéneo é compatíbel determinado, o sistema de vectores S será libre. Facendo operacións elementais na matriz do sistema temos que

$$A=\begin{pmatrix}1&1&5\\2&0&6\\1&2&7\end{pmatrix}\overset{F_{31}(-1),\,F_{21}(-2)}{\Longrightarrow}\begin{pmatrix}1&1&5\\0&-2&-4\\0&1&2\end{pmatrix}$$

$$\overset{F_{32}(1/2)}{\Longrightarrow}\begin{pmatrix}1&1&5\\0&-2&-4\\0&0&0\end{pmatrix}$$

e isto implica que $\operatorname{rang}(A)=2$. Daquela, o sistema admite infinitas solucións (é compatíbel indeterminado) e, como consecuencia, o sistema é ligado.

2) Tomemos o $\mathbb{R}$-espazo vectorial $V=\mathcal{M}_{2\times 2}(\mathbb{R})$. Sexa S o sistema dado por

$$S=\left\{A_1=\begin{pmatrix}1&0\\0&0\end{pmatrix}, A_2=\begin{pmatrix}1&1\\0&0\end{pmatrix}, A_3=\begin{pmatrix}1&1\\1&0\end{pmatrix}, A_4=\begin{pmatrix}1&1\\1&1\end{pmatrix}\right\}.$$

Neste caso cúmprese que

$$\alpha_1A_1+\alpha_2A_2+\alpha_3A_3+\alpha_4A_4=\theta_V \Leftrightarrow$$

$$\alpha_1\begin{pmatrix}1&0\\0&0\end{pmatrix}+\alpha_2\begin{pmatrix}1&1\\0&0\end{pmatrix}+\alpha_3\begin{pmatrix}1&1\\1&0\end{pmatrix}+\alpha_4\begin{pmatrix}1&1\\1&1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}\Leftrightarrow$$

$$\left\{\begin{array}{l}\alpha_1+\alpha_2+\alpha_3+\alpha_4=0\\ \alpha_2+\alpha_3+\alpha_4=0\\ \alpha_3+\alpha_4=0\\ \alpha_4=0\end{array}\right. \Leftrightarrow \begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}.$$

Como a matriz

$$A=\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{pmatrix}$$

ten rango máximo, este resulta compatíbel determinado e, como consecuencia, S é libre.

Algunhas propiedades interesantes dos sistemas libres e ligados enúncianse no seguinte resultado:

Proposición 5.2.3. *Sexa V un $\mathbb{K}$-espazo vectorial. Verifícase o seguinte:*

1) *Todo sistema que conteña ao vector nulo é ligado.*
2) *Se $v\in V$ é un vector non nulo, o sistema $T=\{v\}$ é libre.*
3) *Se T é libre e $v\notin\langle T\rangle$, o sistema $S=T\cup\{v\}$ é libre.*
4) *Se T_1 é ligado e $T_1\subset T_2$, ó sistema T_2 tamén é ligado.*
5) *Todo sistema de vectores contido nun sistema libre é libre.*

Demostración. A demostración é moi sinxela (déixase como exercicio ao lector). □

Nota 5.2.4. A propiedade 3) da proposición anterior é interesante, xa que nos indica o procedemento para introducir vectores nun sistema libre T de forma que o novo sistema resultante sexa libre.

Definición 5.2.5. Sexa V un $\mathbb{K}$-espazo vectorial e sexa $T = \{v_1, \dots, v_r\}$ un sistema de vectores de V. Chámase rango de T, denotado por rang (T), ao máximo número de vectores independentes contidos en T.

Daquela, o sistema $T = \{v_1, \dots, v_r\}$ é libre, se, e só se, rang $(T) = r$.

Proposición 5.2.6. *Sexa V un $\mathbb{K}$-espazo vectorial e sexa $T = \{v_1, \dots, v_r\}$ un sistema libre de vectores de V. Verifícanse as seguintes afirmacións.*

1) *O sistema resultante ao intercambiar dous vectores en T é libre.*
2) *Se multiplicamos un vector de T por un escalar non nulo, o novo sistema segue sendo libre.*
3) *Se a un vector de T lle sumamos un múltiplo doutro vector de T, o sistema resultante é libre.*

Demostración. As dúas primeiras propiedades son evidentes. Para probar a terceira, supoñamos que ao vector v_j de T lle sumamos αv_i con $i \neq j$. Entón,

$$\alpha_1 v_1 + \cdots + \alpha_i v_i + \cdots + \alpha_j(v_j + \alpha v_i) + \cdots + \alpha_r v_r = \theta_V \Leftrightarrow$$

$$\alpha_1 v_1 + \cdots + (\alpha_i + \alpha_j \alpha) v_i + \cdots + \alpha_j v_j + \cdots + \alpha_r v_r = \theta_V.$$

Como o sistema T é libre, tense que

$$\alpha_1 = \cdots = \alpha_i + \alpha_j \alpha = \cdots = \alpha_j = \cdots = \alpha_r = 0$$

e, como consecuencia,

$$\alpha_1 = \cdots = \alpha_i = \cdots = \alpha_j = \cdots = \alpha_r = 0.$$

Por tanto, $T' = \{v_1, \dots, v_i, \dots, v_j + \alpha v_i, \dots, v_r\}$ é libre. □

A proposición anterior implica que en $V = \mathbb{K}^n$ podemos utilizar as operacións elementais para calcular o rango dun sistema de vectores. Isto é:

Corolario 5.2.7. *Sexa $V = \mathbb{K}^n$ e sexa $T = \{v_1, \dots, v_r\}$ un sistema de vectores de V. O rango de T é o rango da matriz que resulta cando escribimos os vectores de T ben en fila, ben en columna.*

Demostración. Consecuencia evidente da Proposición 5.2.6. □

Exemplo 5.2.8. Consideremos o $\mathbb{R}$-espazo vectorial $V = \mathbb{R}^4$. Sexa T o sistema de vectores

$$S = \left\{ v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 3 \end{pmatrix} \right\}.$$

Tomemos a matriz

$$\begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix}$$

e fagamos operacións elementais para calcular o seu rango. Tense que

$$A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix} \overset{F_{41}(-1),\, F_{21}(1)}{\Longrightarrow} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 4 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \overset{F_{23}}{\Longrightarrow} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \\ 0 & 1 & 2 \end{pmatrix}$$

$$\overset{F_{42}(-1),\, F_{32}(-2)}{\Longrightarrow} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

e isto implica que rang $(A) = 2$. Entón, o sistema de vectores T ten rango dous.

Definición 5.2.9. Sexa V un $\mathbb{K}$-espazo vectorial. Diremos que V é de tipo finito se existe un sistema finito de vectores $T = \{v_1, \ldots, v_r\}$, tal que $< T >= V$.

Nota 5.2.10. A partir deste intre todos os espazos vectoriais serán de tipo finito. Por tanto, cando escribamos *sexa V un $\mathbb{K}$-espazo vectorial* estaremos a entender que V é un $\mathbb{K}$-espazo vectorial de tipo finito.

Definición 5.2.11. Sexa V un $\mathbb{K}$-espazo vectorial. Diremos que un sistema de vectores B é unha base de V se é libre e un sistema de xeradores de V.

Teorema 5.2.12. *Todo $\mathbb{K}$-espazo vectorial V admite unha base.*

Demostración. Sexa $T = \{v_1, \ldots, v_r\}$, tal que $< T >= V$. Se T é libre, entón T é unha base e acabamos a proba. Se non o é, existe $v_i^1 \in T$ tal que é combinación linear dos restantes vectores de T. Tomemos $T_1 = T - \{v_i^1\}$. Cúmprese que $< T_1 >= V$ e isto implica que, se T_1 é libre, T_1 é unha base, e a demostración está feita. En caso contrario, repetimos con T_1 o mesmo razoamento que fixemos con T. No peor dos casos chegaremos a $T_{r-1} = \{v_i^r\}$, con v_i^r non nulo, e tal que $< T_{r-1} >= V$. Obviamente, neste caso a base de V sería T_{r-1} xa que, como $v_i^r \neq \theta_V$, o sistema é libre.

□

Exemplos 5.2.13.

1) Se $V = \mathbb{R}$ e $\mathbb{K} = \mathbb{R}$, unha base de V é $B_a = \{a\}$ sendo a un escalar real non nulo.
2) Se $V = \mathbb{C}$ e $\mathbb{K} = \mathbb{C}$, unha base de V é $B_z = \{z\}$ sendo z un escalar complexo non nulo.
3) Se $V = \mathbb{C}$ e $\mathbb{K} = \mathbb{R}$, unha base de V é $B = \{1, i\}$.
4) Sexa o $\mathbb{K}$-espazo vectorial $V = \mathbb{K}^n$. Unha base de V é $C_n = \{e_1, \ldots, e_n\}$, sendo e_i o vector que ten todas as súas compoñentes nulas salvo a compoñente i-ésima que vale 1. Esta base chámase base canónica de $\mathbb{K}^n$.
5) Sexa o $\mathbb{K}$-espazo vectorial $V = \mathcal{M}_{m\times n}(\mathbb{K})$. Unha base de V é
$$B = \{E_{ij} \mid i \in \{1, \ldots, m\},\ j \in \{1, \ldots, n\}\}$$
sendo E_{ij} a matriz que ten todos os seus coeficientes nulos salvo o que ocupa o lugar de subíndices ij que vale 1.
6) Sexa o $\mathbb{K}$-espazo vectorial $V = \Pi_n(\mathbb{K})$. Unha base de V é
$$B = \{1, x, \ldots, x^n\}.$$
7) Sexa o $\mathbb{R}$-espazo vectorial $V = \mathbb{R}^3$. Unha base do subespazo
$$U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V \mid y = 0 \right\}$$
está dada por
$$B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Proposición 5.2.14. *Sexa V un $\mathbb{K}$-espazo vectorial. Se $L = \{u_1, \dots, u_p\}$ é un sistema libre e $B = \{e_1, \dots, e_n\}$ é unha base de V, cúmprese que $p \leq n$.*

Demostración. Supoñamos que $p > n$ e consideremos a igualdade

$$x_1 u_1 + \cdots + x_p u_p = \theta_V.$$

Como B é un sistema xerador de V, temos que para todo $j \in \{1, \cdots, p\}$

$$u_j = \alpha_{1j} e_1 + \cdots + \alpha_{nj} e_n = \sum_{i=1}^{n} \alpha_{ij} e_i.$$

Por tanto,

$$x_1 \sum_{i=1}^{n} \alpha_{i1} e_i + \cdots + x_p \sum_{i=1}^{n} \alpha_{ip} e_i = \theta_V,$$

ou, equivalentemente,

$$(\sum_{j=1}^{p} \alpha_{1j} x_j) e_1 + \cdots + (\sum_{j=1}^{p} \alpha_{nj} x_j) e_n = \theta_V.$$

Doutra banda, como B é libre, todos os coeficientes da igualdade anterior serán nulos. Isto implica que

$$\sum_{j=1}^{p} \alpha_{1j} x_j = \cdots = \sum_{j=1}^{p} \alpha_{nj} x_j = 0$$

e tamén

$$\begin{cases} \alpha_{11} x_1 + \cdots + \alpha_{1p} x_p = 0 \\ \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \alpha_{n1} x_1 + \cdots + \alpha_{np} x_p = 0 \end{cases}$$

O anterior sistema é homoxéneo cun número de incógnitas p maior que o número de ecuacións n. Daquela, é compatíbel indeterminado e isto implica que existen escalares non todos nulos $x_1, \dots, x_n$, tales que

$$x_1 u_1 + \cdots + x_p u_p = \theta_V.$$

Por tanto, chegamos a que L é ligado, o que é un absurdo que provén de supoñer que $p > n$. □

Corolario 5.2.15. *Todas as bases dun $\mathbb{K}$-espazo vectorial V teñen o mesmo número de elementos.*

Demostración. Sexan B e B' dúas bases de V con n e n' elementos respectivamente. Como B é libre e B' é unha base, aplicando a proposición anterior, obtemos que $n \leq n'$. Intercambiando os papeis de B e B' e aplicando o mesmo razoamento chegamos a que $n' \leq n$. Por tanto, $n = n'$. □

Definición 5.2.16. Dado un $\mathbb{K}$-espazo vectorial V non nulo, chamaremos dimensión de V, denotada por $\dim_{\mathbb{K}}(V)$, ao número de elementos dunha das súas bases.

Xa que polo corolario anterior sabemos que todas as bases de V teñen o mesmo número de elementos, a dimensión de V está ben definida e resulta ser un invariante de V. Se $V = \{\theta_V\}$ defínese a súa dimensión como $\dim_{\mathbb{K}}(V) = 0$.

Como primeira consecuencia da definición de dimensión tense que, se $\dim_{\mathbb{K}}(V) = n$, todo sistema xerador con n elementos resulta ser unha base de V e todo sistema libre con n elementos tamén é unha base.

Exemplos 5.2.17. Para os casos considerados en Exemplos 5.2.13 temos que:

1) Se $V = \mathbb{R}$ e $\mathbb{K} = \mathbb{R}$, $\dim_{\mathbb{K}}(V) = 1$.

2) Se $V = \mathbb{C}$ e $\mathbb{K} = \mathbb{C}$, $\dim_{\mathbb{K}}(V) = 1$.
3) Se $V = \mathbb{C}$ e $\mathbb{K} = \mathbb{R}$, $\dim_{\mathbb{K}}(V) = 2$.
4) Se $V = \mathbb{K}^n$, $\dim_{\mathbb{K}}(V) = n$.
5) Se $V = \mathcal{M}_{m\times n}(\mathbb{K})$, $\dim_{\mathbb{K}}(V) = mn$.
6) Se $V = \Pi_n(\mathbb{K})$, $\dim_{\mathbb{K}}(V) = n + 1$.
7) Sexan o $\mathbb{K}$-espazo vectorial $V = \mathbb{R}^3$ e o subespazo
$$U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V \mid y = 0 \right\}.$$
Entón $\dim_{\mathbb{R}}(U) = 2$.
8) Sexan $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ e U o subespazo de $\mathbb{K}^n$ dado polas solucións do sistema homoxéneo $Ax = \theta_{\mathbb{K}^n}$, isto é
$$U = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n \mid A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \right\}.$$
Entón,
$$\dim_{\mathbb{K}}(U) = n - \operatorname{rang}(A).$$
Por exemplo, se U é o subespazo de $\mathbb{R}^3$ dado por
$$U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\},$$
como $\operatorname{rang}(A) = 2$, tense que
$$\dim_{\mathbb{R}}(U) = 3 - \operatorname{rang}(A) = 3 - 2 = 1.$$

Teorema 5.2.18. *Supoñamos que V é un $\mathbb{K}$-espazo vectorial de dimensión n. Se $L = \{u_1, \dots, u_p\}$ é un sistema libre con $p < n$, existen $u_{p+1}, \dots, u_n$ vectores de V tal que*
$$B = \{u_1, \dots, u_p, u_{p+1}, \dots, u_n\}$$
é unha base de V.

Demostración. Como $p < n$, tense que $\langle L\rangle \subsetneq V$. Por tanto, existe $u_{p+1} \in V$ tal que $u_{p+1} \notin \langle L\rangle$. Sexa $L_1 = L \cup \{u_{p+1}\}$. Como L_1 é libre, se $p+1 = n$, xa atopamos a base. Se $p+1 < n$, repítese o mesmo proceso con L_1. Despois de varios pasos finalizaremos cando se chegue á dimensión de V. Nese momento teriamos engadido a L $n-p$ vectores de forma que o sistema
$$L_{n-p} = \{u_1, \dots, u_p, u_{p+1}, \dots, u_n\}$$
é libre e, por tanto, a base buscada. □

Teorema 5.2.19. **(Fórmula de Grassmann)** *Supoñamos que V é un $\mathbb{K}$-espazo vectorial de dimensión n. Sexan U_1, U_2 subespazos de V. Entón*
$$\dim_{\mathbb{K}}(U_1 + U_2) = \dim_{\mathbb{K}}(U_1) + \dim_{\mathbb{K}}(U_2) - \dim_{\mathbb{K}}(U_1 \cap U_2).$$

Demostración. Supoñamos que $\dim_{\mathbb{K}}(U_1 \cap U_2) = k$, que $\dim_{\mathbb{K}}(U_1) = r$ e que $\dim_{\mathbb{K}}(U_2) = s$. Sexa $\{e_1, \dots, e_k\}$ unha base de $U_1 \cap U_2$. Grazas ao teorema anterior existen vectores $\{u_{k+1}, \dots, u_r\}$ e $\{v_{k+1}, \dots, v_r\}$ tales que
$$\{e_1, \dots, e_k, u_{k+1}, \dots u_r\}$$
é unha base de U_1 e
$$\{e_1, \dots, e_k, v_{k+1}, \dots v_s\}$$

é unha base de U_2. Tendo en conta isto, o sistema de vectores

$$T = \{e_1, \dots, e_k, u_{k+1}, \dots u_r, v_{k+1}, \dots v_s\}$$

é unha base de $U_1 + U_2$. En efecto, claramente é un sistema de xeradores para a suma e é libre polo seguinte. Se

$$\alpha_1 e_1 + \cdots + \alpha_k e_k + \alpha_{k+1} u_{k+1} + \cdots + \alpha_r u_r + \beta_{k+1} v_{k+1} + \cdots + \beta_s v_s = \theta_V$$

tense que

$$\alpha_1 e_1 + \cdots + \alpha_k e_k + \alpha_{k+1} u_{k+1} + \cdots + \alpha_r u_r = -\beta_{k+1} v_{k+1} - \cdots - \beta_s v_s$$

e, por tanto, $\beta_{k+1} v_{k+1} + \cdots + \beta_s v_s \in U_1 \cap U_2$. Logo $\beta_{k+1} v_{k+1} + \cdots + \beta_s v_s = \beta_1 e_1 + \cdots + \beta_k e_k$ e isto implica que $\beta_i = 0$ para todo $i \in \{1, \dots, k, k+1, \dots, s\}$. Como consecuencia, $\alpha_j = 0$ para todo $j \in \{1, \dots, k, k+1, \dots, r\}$ e o sistema T é libre.

Como consecuencia, obtemos a fórmula de Grassmann xa que

$$\dim_{\mathbb{K}}(U_1 + U_2) = k + (r - k) + (s - k) = r + s - k = \dim_{\mathbb{K}}(U_1) + \dim_{\mathbb{K}}(U_2) - \dim_{\mathbb{K}}(U_1 \cap U_2).$$

□

5.3. Sistemas de coordenadas. Cambio de base

Teorema 5.3.1. *Sexa $B = \{e_1, \dots, e_n\}$ unha base dun $\mathbb{K}$-espazo vectorial V. Para todo $v \in V$ existe un único vector*

$$v_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

tal que $v = x_1 e_1 + \cdots + x_n e_n$.

Demostración. Como B é un conxunto xerador de V é evidente que para todo $v \in V$ existen escalares $x_1, \cdots, x_n$ tales que $v = x_1 e_1 + \cdots + x_n e_n$. Se existisen outros escalares $y_1, \dots, y_n$ para os que $v = y_1 e_1 + \cdots + y_n e_n$ teriamos o seguinte:

$$\theta_V = v - v = (x_1 e_1 + \cdots + x_n e_n) - (y_1 e_1 + \cdots + y_n e_n) = (x_1 - y_1) e_1 + \cdots + (x_n - y_n) e_n.$$

Agora ben, como B é libre obtense que $x_i - y_i = 0$ para todo $i \in \{1, \cdots, n\}$ ou, equivalentemente, $x_i = y_i$ para todo $i \in \{1, \cdots, n\}$. □

Definición 5.3.2. Tomemos un $\mathbb{K}$-espazo vectorial V e $v \in V$. Chámase vector de coordenadas de v na base $B = \{e_1, \dots, e_n\}$ ao único vector

$$v_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

tal que $v = x_1 e_1 + \cdots + x_n e_n$.

É claro que, se coñecemos o vector v e a base B, podemos calcular o vector de coordenadas v_B resolvendo a ecuación $v = x_1 e_1 + \cdots + x_n e_n$. De xeito inverso, se coñecemos a base B e o vector de coordenadas v_B, recuperamos v como

$$v = x_1 e_1 + \cdots + x_n e_n.$$

Obsérvese tamén que, se cambiamos a base, o vector de coordenadas pode cambiar. No exemplo seguinte aclaramos este comentario.

Exemplos 5.3.3. 1) Sexa o $\mathbb{R}$-espazo vectorial $V = \Pi_2(\mathbb{R})$ e

$$B = \{p_1(x) = 1 + x + x^2, p_2(x) = 1 + x, p_3(x) = 1\}$$

unha base de V. Tomemos $q(x) = 1 + 2x + 3x^2$ e calculemos o seu vector de coordenadas respecto á base B. En primeiro lugar, tense que

$$q(x) = y_1 p_1(x) + y_2 p_2(x) + y_3 p_3(x) \Leftrightarrow$$

$$1 + 2x + 3x^2 = (y_1 + y_2 + y_3) + (y_1 + y_2)x + y_1 x^2$$

ou, equivalentemente,

$$\begin{cases} y_1 + y_2 + y_3 = 1 \\ \quad\;\; y_1 + y_2 = 2 \\ \qquad\quad\; y_1 = 3 \end{cases}$$

Se solucionamos o sistema anterior, obtemos que $y_1 = 3$, $y_2 = -1$ e $y_3 = -1$. Por tanto, o vector de coordenadas de $q(x)$ na base B é

$$q(x)_B = \begin{pmatrix} 3 \\ -1 \\ -1 \end{pmatrix} \in \mathbb{R}^3.$$

Se neste mesmo exemplo cambiamos a base B pola base

$$C = \{t_1(x) = 1, t_2(x) = x, t_3(x) = x^2\}$$

obtemos que

$$q(x)_C = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^3.$$

Desta forma queda claro que se varía a base, as coordenadas poden cambiar.

2) Sexa o $\mathbb{R}$-espazo vectorial $V = \mathbb{R}^n$ e

$$C_n = \{e_1, \dots, e_n\}$$

a base canónica de $\mathbb{R}^n$ introducida no cuarto exemplo de Exemplos 5.2.13. É moi fácil ver que para todo $v \in \mathbb{R}^n$ tense que $v_{C_n} = v$. Neste caso as coordenadas de v coinciden coas compoñentes de v.

Para finalizar esta parte falaremos do cambio de base. Supoñamos que

$$B = \{e_1, \dots, e_n\}, \quad D = \{u_1, \dots, u_n\}$$

son dúas bases dun $\mathbb{K}$-espazo vectorial V. Sexan $v \in V$ e

$$v_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad v_D = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$

os vectores de coordenadas de v nas bases B e D respectivamente. Probaremos que é posíbel relacionar os dous vectores de coordenadas mediante unha matriz que será chamada a matriz de cambio de base de B a D. En primeiro lugar calculamos os vectores de coordenadas dos vectores da base B na base D. Entón,

$$e_{jD} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}, \;\; j \in \{1, \dots, n\} \Leftrightarrow e_j = a_{1j}u_1 + \cdots + a_{nj}u_n = \sum_{i=1}^{n} a_{ij}u_i.$$

Doutra banda

$$v = z_1u_1 + \cdots + z_nu_n = x_1e_1 + \cdots + x_ne_n = x_1(\sum_{i=1}^{n} a_{i1}u_i) + \cdots + x_n(\sum_{i=1}^{n} a_{in}u_i) =$$

$$(\sum_{j=1}^{n} a_{1j}x_j)u_1 + \cdots + (\sum_{j=1}^{n} a_{nj}x_j)u_n.$$

Por tanto,

$$\begin{cases} z_1 = & \sum_{j=1}^{n} a_{1j}x_j = & a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots & \vdots & \vdots \quad \vdots \\ z_n = & \sum_{j=1}^{n} a_{nj}x_j = & a_{n1}x_1 + \cdots + a_{nn}x_n \end{cases}$$

ou, equivalentemente,

$$\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Daquela, $v_D = P_{B,D}v_B$ sendo

$$P_{B,D} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

a matriz de cambio de base de B a D. As columnas da matriz anterior son os vectores de coordenadas na base D dos vectores da base B, isto é:

$$P_{B,D} = \left(\begin{array}{c|c|c} e_{1D} & \cdots & e_{nD} \end{array} \right).$$

Esta matriz e invertíbel e a súa inversa é a matriz $P_{D,B}$ de cambio de base de D a B. Finalmente, nótese que, se $V = \mathbb{K}^n$ e $D = C_n$, para toda base $B = \{e_1, \ldots, e_n\}$, tense que

$$P_{B,C_n} = \left(\begin{array}{c|c|c} e_1 & \cdots & e_n \end{array} \right).$$

Desta forma, as columnas da matriz de cambio de base son os propios vectores de B, xa que $e_{jC_n} = e_j$.

Exemplo 5.3.4. Sexa $V = \mathbb{R}^4$. Tomemos as bases

$$B = \left\{ e_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, e_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\},$$

$$D = \left\{ u_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, u_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

e calculemos a matriz de cambio de base $P_{B,D}$.

En primeiro lugar obtemos os vectores de coordenadas e_{1D}, e_{2D}, e_{3D} y e_{4D}. Para iso, resolvemos as ecuacións

$$e_j = x_1u_1 + x_2u_2 + x_3u_3 + x_4u_4, \;\; j \in \{1, 2, 3, 4\}$$

Equivalentemente, debemos resolver os catro sistemas:

$$\begin{cases} x_4 = 1 \\ x_3 + x_4 = 1 \\ x_2 + x_3 + x_4 = 1 \\ x_1 + x_2 + x_3 + x_4 = 1 \end{cases} \qquad \begin{cases} x_4 = 1 \\ x_3 + x_4 = 1 \\ x_2 + x_3 + x_4 = 1 \\ x_1 + x_2 + x_3 + x_4 = 0 \end{cases}$$

$$\begin{cases} x_4 = 1 \\ x_3 + x_4 = 1 \\ x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 + x_4 = 0 \end{cases} \qquad \begin{cases} x_4 = 1 \\ x_3 + x_4 = 0 \\ x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 + x_4 = 0 \end{cases}.$$

A solución do primeiro é

$$x_1 = 0, \ \ x_2 = 0, \ \ x_3 = 0, \ \ x_4 = 1$$

e a do segundo ven dada por

$$x_1 = -1, \ \ x_2 = 0, \ \ x_3 = 0, \ \ x_4 = 1.$$

Para o terceiro temos que

$$x_1 = 0, \ \ x_2 = -1, \ \ x_3 = 0, \ \ x_4 = 1$$

e o cuarto ten como solución

$$x_1 = 0, \ \ x_2 = 0, \ \ x_3 = -1, \ \ x_4 = 1.$$

Por tanto,

$$e_{1D} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, e_{2D} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, e_{3D} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}, e_{4D} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

e, como consecuencia, a matriz de cambio de base é

$$P_{BD} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Se, por exemplo, tomamos o vector

$$v = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

o seu vector de coordenadas na base B é

$$v_B = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

xa que $v = e_2 - e_3 + e_4$. Se queremos calcular o vector de coordenadas de v en D, como temos a matriz de cambio de base, só temos que facer o produto:

$$P_{BD} v_B = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}.$$

Logo,

$$v_D = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}.$$

5.4. Aplicacións lineares. Matriz asociada. Núcleo, imaxe e rango

Definición 5.4.1. Sexan V e W dous $\mathbb{K}$-espazos vectoriais. Unha aplicación $f : V \to W$ chamarase linear se cumpre que

1) $f(u+v) = f(u) + f(v)$, $\forall\, u, v \in V$,
2) $f(\alpha v) = \alpha f(v)$, $\forall\, \alpha \in \mathbb{K}$, $v \in V$.

Exemplos 5.4.2. 1) A aplicación identidade $id_V : V \to V$ é linear.

2) O segundo exemplo de aplicación linear é o que empregaremos máis. Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ e tomemos a aplicación

$$f_A : \mathbb{K}^n \to \mathbb{K}^m$$

definida por

$$f_A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Entón f_A é unha aplicación linear xa que

$$f_A\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = A\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + A\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} =$$

$$f_A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + f_A\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

e

$$f_A\left(\alpha\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = A\left(\alpha\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \alpha A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \alpha f_A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

3) A aplicación $f : \mathbb{R}^4 \to \mathbb{R}^3$ definida por

$$f\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x \\ x+y \\ x+y+z+t \end{pmatrix}$$

é linear.

4) A aplicación $f : \mathcal{M}_{m\times n}(\mathbb{K}) \to \mathcal{M}_{n\times m}(\mathbb{K})$ dada por $f(A) = A^t$ tamén é linear.
5) A aplicación $f : \mathcal{M}_{m\times n}(\mathbb{C}) \to \mathcal{M}_{n\times m}(\mathbb{C})$ dada por $f(A) = A^*$ non é linear xa que $f(\alpha A) = \overline{\alpha} A^*$.
6) A aplicación $f : \Pi_n(\mathbb{R}) \to \Pi_{n-1}(\mathbb{R})$ dada por $f(p(x)) = p'(x)$ é linear.
7) A aplicación $f : \Pi_n(\mathbb{R}) \to \mathbb{R}$ dada por $f(p(x)) = \int_0^1 p(x)\mathrm{d(x)}$ é linear.
8) A aplicación tr $: \mathcal{M}_{n\times n}(\mathbb{K}) \to \mathbb{K}$ definida no problema proposto número 14 do terceiro capítulo é linear.

9) Tomemos un $\mathbb{K}$-espazo vectorial V e sexa $B = \{e_1, \cdots, e_n\}$ unha base de V. A aplicación

$$coord_B : V \to \mathbb{K}^n$$

dada por

$$coord_B(v) = v_B$$

é unha aplicación linear. En efecto, se u e v teñen como vectores de coordenadas a

$$u_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad v_B = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

respectivamente, entón

$$u + v = \sum_{i=1}^{n} x_i e_i + \sum_{i=1}^{n} y_i e_i = \sum_{i=1}^{n} (x_i + y_i) e_i$$

e isto implica que

$$(u+v)_B = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} = u_B + v_B.$$

Por conseguinte, $coord_B(u + v) = coord_B(u) + coord_B(v)$.
Doutra banda, $coord_B(\alpha u) = \alpha coord_B(u)$, xa que

$$\alpha v = \sum_{i=1}^{n} \alpha x_i e_i$$

e, por tanto, $(\alpha u)_B = \alpha u_B$.

Nota 5.4.3. Como se pode comprobar de forma sinxela, as dúas condicións dadas na definición de aplicación linear equivalen a

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v), \;\; \forall\, \alpha, \beta \in \mathbb{K}, \; u, v \in V.$$

Obsérvese que como consecuencia tense que

$$f(\sum_{i=1}^{r} \alpha_i v_i) = \sum_{i=1}^{r} \alpha_i f(v_i)$$

para toda combinación linear $\sum_{i=1}^{r} \alpha_i v_i$ de vectores de V.

Proposición 5.4.4. *Supoñamos que $f : V \to W$ é unha aplicación linear entre dous $\mathbb{K}$-espazos vectoriais V e W. Cúmprese o seguinte:*

1) $f(\theta_V) = \theta_W$.
2) $f(-v) = -f(v)$.
3) *Se U é un subespazo de V, o conxunto*

$$f(U) = \{f(u) \;\mid\; u \in U\}$$

é un subespazo de W.

4) *Se $f : V \to W$ e $g : W \to X$ son aplicacións lineares, a súa composición $g \circ f : V \to X$ tamén é linear.*

Demostración. A demostración é moi sinxela. Déixase como exercicio ao lector.

No seguinte resultado veremos que toda aplicación linear $f : V \to W$ queda determinada se coñecemos as imaxes dos vectores dunha base de V.

Proposición 5.4.5. *Supoñamos que V e W son dous $\mathbb{K}$-espazos vectoriais. Tomemos unha base $B = \{e_1, \dots, e_n\}$ de V e n vectores $w_1, \dots, w_n$ de W. Cúmprese que existe unha única aplicación linear $f : V \to W$, tal que $f(e_i) = w_i$, $i \in \{1, \dots, n\}$.*

Demostración. Dado $u \in V$, como B é unha base de V, sabemos que existen uns escalares únicos $x_1, \dots, x_n$ tales que $u = \sum_{i=1}^{n} x_i e_i$. Tendo en conta este feito, definimos $f : V \to W$ como

$$f(u) = \sum_{i=1}^{n} x_i w_i.$$

A aplicación definida deste xeito é linear xa que

$$f(u+v) = f(\sum_{i=1}^{n} x_i e_i + \sum_{i=1}^{n} y_i e_i) = f(\sum_{i=1}^{n} (x_i + y_i) e_i) = \sum_{i=1}^{n} (x_i + y_i) w_i =$$

$$\sum_{i=1}^{n} x_i w_i + \sum_{i=1}^{n} y_i w_i = f(u) + f(v)$$

e

$$f(\alpha u) = f(\alpha \sum_{i=1}^{n} x_i e_i) = f(\sum_{i=1}^{n} \alpha x_i e_i) = \sum_{i=1}^{n} \alpha x_i w_i = \alpha \sum_{i=1}^{n} x_i w_i = \alpha f(u).$$

Doutra banda, se existise unha aplicación linear $g : V \to W$ tal que $g(e_i) = w_i$, $i \in \{1, \dots, n\}$, para todo vector $u = \sum_{i=1}^{r} x_i e_i \in V$ temos que

$$f(u) = \sum_{i=1}^{n} x_i w_i = \sum_{i=1}^{n} x_i g(e_i) = g(\sum_{i=1}^{n} x_i e_i) = g(u)$$

e, como consecuencia, $f = g$. □

Exemplo 5.4.6. Sexan os $\mathbb{R}$-espazos vectoriais $V = \mathbb{R}^4$ e $W = \mathbb{R}^3$. Tomemos a base

$$C_4 = \left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, e_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

canónica de V e catro vectores de W

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \ w_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \ w_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \ w_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Entón, a única aplicación linear tal que $f(e_i) = w_i$ para $i \in \{1, 2, 3, 4\}$ está dada por:

$$f \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = f(xe_1 + ye_2 + ze_3 + te_4) = xw_1 + yw_2 + zw_3 + tw_4 =$$

$$\begin{pmatrix} x \\ x \\ x \end{pmatrix} + \begin{pmatrix} y \\ y \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ t \\ 0 \end{pmatrix} = \begin{pmatrix} x+y \\ x+y+t \\ x+y+z \end{pmatrix}.$$

5.4.7. Anteriormente vimos que toda matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ determina unha aplicación linear $f_A : \mathbb{K}^n \to \mathbb{K}^m$ definida por

$$f_A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Por tanto, se $C_n = \{e_1, \cdots, e_n\}$ é a base canónica de $\mathbb{K}^n$, temos que

$$f_A(e_j) = Ae_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

e, o que é o mesmo, $f_A(e_j)$ é a columna j-ésima de A.

No seguinte resultado veremos que a relación entre aplicacións lineares e matrices é máis profunda.

Proposición 5.4.8. *Sexan V, W dous $\mathbb{K}$-espazos vectoriais e sexan $B = \{e_1, \dots, e_n\}$, $D = \{w_1, \dots, w_m\}$ bases de V e W respectivamente. Sexan $f : V \to W$ unha aplicación linear e $v \in V$. Se o vector de coordenadas de v respecto á base B é v_B, o vector de coordenadas de $f(v)$ respecto á base D obtense como*

$$f(v)_D = M(f)_{B,D} v_B$$

onde $M(f)_{B,D} \in \mathcal{M}_{m\times n}(\mathbb{K})$ é a matriz cuxas columnas son

$$M(f)_{B,D} = (f(e_1)_D \mid \cdots \mid f(e_n)_D).$$

Demostración. Tomemos $v \in V$. Para cada $j \in \{1, \dots, n\}$ tense que

$$f(e_j) = a_{1j}w_1 + \cdots + a_{mj}w_m$$

e, entón,

$$f(e_j)_D = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Doutra banda, como B é unha base de V, podemos expresar o vector v como

$$v = x_1e_1 + \cdots + x_ne_n$$

e, aplicando f, obtense a seguinte igualdade:

$$f(v) = x_1f(e_1) + \cdots + x_nf(e_n).$$

Daquela, tendo en conta que o cálculo de coordenadas é unha aplicación linear, obtemos as seguintes igualdades:

$$f(v)_D = x_1f(e_1)_D + \cdots + x_nf(e_n)_D =$$

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \; \vdots \; \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} =$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \left(\; f(e_1)_D \;\mid\; \cdots \;\mid\; f(e_n)_D \; \right) v_B$$

Deste xeito, se $M(f)_{B,D} \in \mathcal{M}_{m\times n}(\mathbb{K})$ é a matriz cuxas columnas son $f(e_1)_D, \dots, f(e_n)_D$, acabamos de demostrar que

$$f(v)_D = M(f)_{B,D} v_B.$$

$\square$

Definición 5.4.9. Sexan V, W dous $\mathbb{K}$-espazos vectoriais e sexan B, D bases de V e W respectivamente. Sexa $f : V \to W$ unha aplicación linear. A matriz $M(f)_{B,D}$ calculada no apartado anterior chámase matriz asociada a f respecto ás bases B e D.

Nota 5.4.10. Un caso coñecido de matriz asociada a unha aplicación linear é a matriz de cambio de base. Sexan V un $\mathbb{K}$-espazo vectorial e $B = \{e_1, \dots, e_n\}$, $D = \{u_1, \dots, u_n\}$ dúas bases de V. Se $f = id_V$, repetindo a demostración do teorema anterior, obtemos que

$$M(id_V)_{B,D} = P_{B,D}$$

e, se $B = D$, temos a igualdade

$$M(id_V)_{B,B} = P_{B,B} = I_n.$$

Outro caso no cal é sinxelo o cálculo da matriz asociada é o seguinte: sexa $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ e tomemos a aplicación linear

$$f_A : \mathbb{K}^n \to \mathbb{K}^m.$$

Entón, se C_n, C_m son as bases canónicas de $\mathbb{K}^n$ e $\mathbb{K}^m$,

$$M(f_A)_{C_n,C_m} = A.$$

Exemplo 5.4.11. Sexa a aplicación linear $f : \mathbb{R}^4 \to \mathbb{R}^3$ dada por

$$f\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x+y \\ x+y+t \\ x+y+z \end{pmatrix}.$$

Neste exemplo calcularemos $M(f)_{B,D}$ sendo $B = C_4$ e

$$D = \left\{ w_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, w_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

En primeiro lugar, debemos obter as imaxes dos vectores da base C_4:

$$f(e_1) = f\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},\ f(e_2) = f\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

$$f(e_3) = f\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\ f(e_4) = f\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

A continuación calcúlanse os vectores de coordenadas das imaxes anteriores na base D. Para obtelos debemos resolver os seguintes sistemas

$$\begin{cases} a_{11} + a_{21} + a_{31} = 1 \\ a_{11} + a_{21} = 1 \\ a_{11} = 1 \end{cases}, \quad \begin{cases} a_{12} + a_{22} + a_{32} = 1 \\ a_{12} + a_{22} = 1 \\ a_{12} = 1 \end{cases},$$

$$\begin{cases} a_{13} + a_{23} + a_{33} = 0 \\ a_{13} + a_{23} = 0 \\ a_{13} = 1 \end{cases}, \quad \begin{cases} a_{14} + a_{24} + a_{34} = 0 \\ a_{14} + a_{24} = 1 \\ a_{14} = 0 \end{cases},$$

inducidos polas igualdades

$$f(e_j) = a_{1j}w_1 + a_{2j}w_2 + a_{3j}w_3,\ j \in \{1,2,3,4\}.$$

Por tanto,

$$\begin{gathered} a_{11}=1,\ a_{21}=0,\ a_{31}=0, \\ a_{12}=1,\ a_{22}=0,\ a_{32}=0, \\ a_{13}=1,\ a_{23}=-1,\ a_{33}=0, \\ a_{14}=0,\ a_{24}=1,\ a_{34}=-1, \end{gathered}$$

e isto implica que:

$$M(f)_{B,D} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Se

$$v = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = v_{C_4},$$

o vector de coordenadas de $f(v)$ na base D é

$$f(v)_D = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}.$$

Logo,

$$f\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 3w_1 - w_3 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}.$$

5.4.12. A continuación enunciamos as principais propiedades que cumpre a matriz asociada a unha aplicación linear. As probas obtéñense de xeito sinxelo.

1) Sexan $f, g : V \to W$ dúas aplicacións lineares e sexan $B = \{e_1, \dots, e_n\}$ e $D = \{w_1, \dots, w_m\}$ bases de V e W respectivamente. Entón

$$M(f+g)_{B,D} = M(f)_{B,D} + M(g)_{B,D}, \quad M(\alpha f)_{B,D} = \alpha M(f)_{B,D},\ \forall \alpha \in \mathbb{K}.$$

2) Sexan $f : V \to W$, $g : W \to X$ dúas aplicacións lineares e sexan $B = \{e_1, \dots, e_n\}$, $D = \{w_1, \dots, w_m\}$ e $E = \{u_1, \dots, u_p\}$ bases de V, W y X respectivamente. Entón

$$M(g \circ f)_{B,E} = M(g)_{D,E} M(f)_{B,D}.$$

5.4.13. Neste punto estudaremos como varía a matriz asociada a unha aplicación linear cando cambiamos as bases.

Sexan $f : V \to W$ unha aplicación linear e B_1, B_2 dúas bases de V e D_1, D_2 dúas bases de W. Tense o seguinte diagrama conmutativo:

$$\begin{tikzcd} V_{B_1} \arrow[r, "f"] & W_{D_1} \arrow[d, "id_W"] \\ V_{B_2} \arrow[u, "id_V"] \arrow[r, "f"'] & W_{D_2} \end{tikzcd}$$

onde o subíndice indica respecto a que base estamos a realizar os cálculos.

Aplicando as propiedades das matrices asociadas respecto da composición temos que

$$M(f)_{B_2,D_2} = M(id_W \circ f \circ id_V)_{B_2,D_2} = M(id_W \circ f)_{B_1D_2}M(id_V)_{B_2B_1} =$$

$$M(id_W)_{D_1D_2}M(f)_{B_1,D_1}M(id_V)_{B_2B_1}.$$

Por tanto,

$$M(f)_{B_2,D_2} = P_{D_1D_2}M(f)_{B_1,D_1}P_{B_2B_1}.$$

Daquela, dúas matrices asociadas á mesma aplicación linear son equivalentes e disto dedúcese, aplicando o Teorema 3.2.21, que teñen o mesmo rango:

$$\operatorname{rang}(M(f)_{B_1,D_1}) = \operatorname{rang}(M(f)_{B_2,D_2}).$$

De xeito inverso, se dúas matrices $A, B \in \mathcal{M}_{m\times n}(\mathbb{K})$ son equivalentes, existen matrices $P \in \mathcal{M}_{m\times m}(\mathbb{K})$ e $Q \in \mathcal{M}_{n\times n}(\mathbb{K})$ invertíbeis tales que $PAQ = B$. Como Q é invertíbel, as súas columnas forman unha base B_2 de $\mathbb{K}^n$ e, analogamente, as columnas de P^{-1} forman unha base D_2 de $\mathbb{K}^m$. Sexa a aplicación linear

$$f_A : \mathbb{K}^n \to \mathbb{K}^m.$$

Tense que:

$$\begin{tikzcd}
\mathbb{K}^n_{C_n} \arrow[r, "f_A"] & \mathbb{K}^m_{C_m} \arrow[d, "id_{\mathbb{K}^m}"] \\
\mathbb{K}^n_{B_2} \arrow[u, "id_{\mathbb{K}^n}"] \arrow[r, "f_A"'] & \mathbb{K}^m_{D_2}
\end{tikzcd}$$

é conmutativo e, logo,

$$M(f_A)_{B_2,D_2} = P_{C_mD_2}M(f_A)_{C_n,C_m}P_{B_2C_n} =$$

$$P^{-1}_{D_2C_m}M(f_A)_{C_n,C_m}P_{B_2C_n} = (P^{-1})^{-1}AQ = PAQ = B.$$

Como consecuencia destes feitos podemos enunciar o seguinte teorema.

Teorema 5.4.14. *Dúas matrices son equivalentes, se, e só se, son matrices asociadas a unha mesma aplicación linear.*

Definición 5.4.15. Dada unha aplicación linear $f : V \to W$ entre dous $\mathbb{K}$-espazos vectoriais V e W defínense o núcleo e a imaxe de f do xeito seguinte:

$$\operatorname{Ker}(f) = \{v \in V \mid f(v) = \theta_W\},$$

$$\operatorname{Im}(f) = f(V).$$

Con esta definición xa obtemos que $\operatorname{Im}(f)$ é un subespazo de W sen máis que aplicar o apartado terceiro da Proposición 5.4.4. De feito a imaxe de f está dada por

$$\operatorname{Im}(f) = \{f(v) \mid v \in V\}.$$

Doutra banda, o núcleo tamén é un subespazo, xa que, se $u, v \in \operatorname{Ker}(f)$, temos que

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v) = \alpha\theta_W + \beta\theta_W = \theta_W$$

e, como consecuencia, $\alpha u + \beta v \in \operatorname{Ker}(f)$.

Exemplo 5.4.16. Sexa $f : \mathbb{R}^4 \to \mathbb{R}^4$ a aplicación linear definida por

$$f\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x \\ x+y \\ x+y \\ x+y+z+t \end{pmatrix}.$$

Temos que

$$f\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} x = 0 \\ x+y=0 \\ x+y=0 \\ x+y+z+t=0 \end{cases} \Leftrightarrow \begin{cases} x=0 \\ y=0 \\ t=-z \end{cases}$$

ou, equivalentemente,

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \operatorname{Ker}(f) \Leftrightarrow \begin{cases} x=0 \\ y=0 \\ t=-z \end{cases}.$$

Logo,

$$\operatorname{Ker}(f) = \left\{ \begin{pmatrix} 0 \\ 0 \\ z \\ -z \end{pmatrix} \in \mathbb{R}^4 \;/\; z \in \mathbb{R} \right\} = \left\{ z\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{R}^4 \;/\; z \in \mathbb{R} \right\} = \langle \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right\} \rangle.$$

e isto implica que $\dim_{\mathbb{R}}(\operatorname{Ker}(f)) = 1$.

A imaxe de f está dada por

$$\left\{ f\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \;/\; \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \right\} = \left\{ \begin{pmatrix} x \\ x+y \\ x+y \\ x+y+z+t \end{pmatrix} \;/\; x,y,z,t \in \mathbb{R} \right\} =$$

$$\left\{ x\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} + (z+t)\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \;/\; x,y,z,t \in \mathbb{R} \right\} = \langle \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\} \rangle$$

e, como os vectores que xeran a imaxe forman un sistema libre, temos que

$$\dim_{\mathbb{R}}(\operatorname{Im}(f)) = 3.$$

Proposición 5.4.17. *Dada unha aplicación linear $f : V \to W$ entre dous $\mathbb{K}$-espazos vectoriais V e W tense que*

$$\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(\operatorname{Ker}(f)) + \dim_{\mathbb{K}}(\operatorname{Im}(f)).$$

Demostración. Supoñamos que $\dim_{\mathbb{K}}(V) = n$ e que $\dim_{\mathbb{K}}(\operatorname{Ker}(f)) = p < n$. Se

$$\dim_{\mathbb{K}}(\operatorname{Ker}(f)) = n$$

temos que $\operatorname{Im}(f) = \{\theta_W\}$ e a fórmula cúmprese trivialmente. Sexa $C = \{e_1, \dots, e_p\}$ unha base de $\operatorname{Ker}(f)$ e fagamos unha ampliación ata obter unha base de V dada por

$$B = \{e_1, \dots, e_p, e_{p+1}, \dots, e_n\}.$$

Tense que $I = \{f(e_{p+1}), \dots, f(e_n)\}$ é unha base de $\operatorname{Im}(f)$. En efecto, o sistema é libre xa que, se

$$\alpha_{p+1} f(e_{p+1}) + \dots + \alpha_n f(e_n) = \theta_W,$$

pola linearidade de f, cúmprese que

$$f(\alpha_{p+1}e_{p+1} + \cdots + \alpha_n e_n) = \theta_V \Leftrightarrow \alpha_{p+1}e_{p+1} + \cdots + \alpha_n e_n \in \mathrm{Ker}(f) \Leftrightarrow$$

$$\exists\, \alpha_1, \cdots, \alpha_p \in \mathbb{K} \,/\, \alpha_1 e_1 + \cdots + \alpha_p e_p = \alpha_{p+1}e_{p+1} + \cdots + \alpha_n e_n \Leftrightarrow$$

$$\alpha_1 e_1 + \cdots + \alpha_p e_p - \alpha_{p+1}e_{p+1} - \cdots - \alpha_n e_n = \theta_V.$$

Como B é unha base de V, $\alpha_1 = \cdots = \alpha_p = \alpha_{p+1} = \cdots = \alpha_n = 0$. Por tanto, I é libre.

Finalmente, o sistema de vectores I é xerador. Se $w \in \mathrm{Im}(f)$, existe un vector $v \in V$ tal que $f(v) = w$ e, como B é unha base de V, o vector v é combinación linear dos vectores de B

$$v = \alpha_1 e_1 + \cdots + \alpha_p e_p + \alpha_{p+1}e_{p+1} + \cdots + \alpha_n e_n.$$

Entón,

$$w = f(v) = f(\alpha_1 e_1 + \cdots + \alpha_p e_p + \alpha_{p+1}e_{p+1} + \cdots + \alpha_n e_n) =$$

$$\alpha_1 f(e_1) + \cdots + \alpha_p f(e_p) + \alpha_{p+1} f(e_{p+1}) + \cdots + \alpha_n f(e_n) =$$

$$\alpha_1 \theta_W + \cdots + \alpha_p \theta_W + \alpha_{p+1} f(e_{p+1}) + \cdots + \alpha_n f(e_n) = \alpha_{p+1} f(e_{p+1}) + \cdots + \alpha_n f(e_n) \Leftrightarrow$$

$$w \in \langle \{f(e_{p+1}), \cdots, f(e_n)\} \rangle.$$

□

Definición 5.4.18. Dada unha aplicación linear $f : V \to W$ entre dous $\mathbb{K}$-espazos vectoriais V e W, diremos que é un monomorfismo se é inxectiva. Cando sexa sobrexectiva, chamarase epimorfismo e, se é inxectiva e sobrexectiva, chamarémola isomorfismo.

Estes tipos de aplicacións lineares caracterízanse totalmente utilizando o núcleo e a imaxe.

Proposición 5.4.19. *Dada unha aplicación linear $f : V \to W$ entre dous $\mathbb{K}$-espazos vectoriais V e W verifícase o seguinte:*

1) *A aplicación linear f é un monomorfismo, se, e só se,* $\mathrm{Ker}(f) = \theta_V$.
2) *A aplicación linear f é un epimorfismo, se, e só se,* $\mathrm{Im}(f) = W$.
3) *A aplicación linear f é un isomorfismo, se, e só se,* $\mathrm{Ker}(f) = \theta_V$ *e* $\mathrm{Im}(f) = W$.

Demostración. A demostración da segunda afirmación é trivial. A terceira é consecuencia da segunda e da primeira. Por tanto, só probaremos a primeira.

Se f é un monomorfismo e $f(v) = \theta_W$, tense que $v = \theta_V$ e isto implica que $\mathrm{Ker}(f) = \{\theta_V\}$. Inversamente, supoñamos que $f(v) = f(u)$. Entón, $f(v-u) = \theta_W$ ou, equivalentemente, $v - u \in \mathrm{Ker}(f)$. Se $\mathrm{Ker}(f) = \theta_V$, obtemos $v - u = \theta_V$ e isto implica que os dous vectores son iguais. Por tanto, f é un monomorfismo. □

Nota 5.4.20. Se unha aplicación linear $f : V \to W$, entre dous $\mathbb{K}$-espazos vectoriais V e W, é un isomorfismo, pódese construír a súa inversa $f^{-1} : W \to V$ como

$$f^{-1}(w) = v \Leftrightarrow w = f(v).$$

Esta aplicación tamén é linear (a comprobación déixase como exercicio) e cumpre que

$$f^{-1} \circ f = id_V, \quad f \circ f^{-1} = id_W.$$

Por tanto, se $B = \{e_1, \dots, e_n\}$ é unha base de V e $D = \{w_1, \dots, w_n\}$ é unha base de W, temos que

$$I_n = M(id_V)_{B,B} = M(f^{-1} \circ f)_{B,B} = M(f^{-1})_{D,B} M(f)_{B,D}$$

e

$$I_n = M(id_W)_{D,D} = M(f \circ f^{-1})_{D,D} = M(f)_{B,D} M(f^{-1})_{D,B}.$$

Por tanto, $M(f)_{B,D}$ é invertíbel e

$$M(f)^{-1}_{B,D} = M(f^{-1})_{D,B}.$$

Exemplo 5.4.21. Tomemos un $\mathbb{K}$-espazo vectorial V e sexa $B = \{e_1, \dots, e_n\}$ unha base de V. Sexa a aplicación linear considerada no exemplo noveno de Exemplos 5.4.2

$$coord_B : V \to \mathbb{K}^n$$

dada por

$$coord_B(v) = v_B.$$

Esta aplicación é un monomorfismo e tamén un epimorfismo. Por tanto, é un isomorfismo cuxa inversa é

$$coord_B^{-1} : \mathbb{K}^n \to V$$

$$coord_B^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \dots + x_n e_n.$$

Proposición 5.4.22. *Dada unha aplicación linear $f : V \to W$ entre dous $\mathbb{K}$-espazos vectoriais V e W cúmprese o seguinte.*

1) *A aplicación linear f é un monomorfismo, se, e só se, transforma sistemas libres de V en sistemas libres de W.*
2) *A aplicación linear f é un epimorfismo, se, e só se, transforma sistemas xeradores de V en sistemas xeradores de W.*
3) *A aplicación linear f é un isomorfismo, se, e só se, transforma bases de V en bases de W.*

Demostración. A terceira afirmación é consecuencia das dúas primeiras. Probaremos en primeiro lugar 1). Supoñamos que v é un vector non nulo de V. O sistema $T = \{v\}$ é libre e, se f transforma sistemas libres de V en sistemas libres de W, temos que $f(T) = \{f(v)\}$ é libre ou, equivalentemente, $f(v) \neq \theta_W$. Por tanto, $\mathrm{Ker}(f) = \theta_V$ e f é un monomorfismo.

Inversamente, supoñamos que f é un monomorfismo e que $T = \{e_1, \dots, e_p\}$ é un sistema libre de V. O sistema $f(T)$ tamén é libre, xa que

$$\alpha_1 f(e_1) + \dots + \alpha_p f(e_p) = \theta_W \Leftrightarrow f(\alpha_1 e_1 + \dots + \alpha_p e_p) = f(\theta_V) \Leftrightarrow$$

$$\alpha_1 e_1 + \dots + \alpha_p e_p = \theta_V$$

e isto implica que $\alpha_1 = \dots = \alpha_p = 0$ grazas a que T é libre.

Supoñamos agora que $T = \{e_1, \dots, e_p\}$ é un sistema xerador de V. Entón $\mathrm{Im}(f) = \langle f(T) \rangle$ e, por tanto, se $f(T)$ é un sistema xerador de W, temos que $W = \langle f(T) \rangle = \mathrm{Im}(f)$. Daquela, f resulta ser un epimorfismo. Doutra banda, é evidente que, se a aplicación linear é sobrexectiva, entón os sistemas xeradores de V transfórmanse en sistemas xeradores de W. $\square$

Nota 5.4.23. Volvamos ao exemplo da aplicación linear de coordenadas $coord_B$ respecto a unha base $B = \{e_1, \dots, e_n\}$ dun $\mathbb{K}$-espazo vectorial V. Esta aplicación é un isomorfismo e como consecuencia, leva bases de V en bases de $\mathbb{K}^n$. En particular transforma sistemas libres de V en sistemas libres de $\mathbb{K}^n$. Por tanto, temos que, se $T = \{v_1, \dots, v_r\}$ é un sistema de vectores de V, cúmprese a igualdade

$$\operatorname{rang}(T) = s \Leftrightarrow \operatorname{rang}(\{v_{1B}, \dots, v_{rB}\}) = s.$$

Supoñamos que $f : V \to W$ é unha aplicación linear entre dous $\mathbb{K}$-espazos vectoriais e que $\dim_{\mathbb{K}}(\operatorname{Im}(f)) = r$. Isto implica que $\dim_{\mathbb{K}}(\operatorname{Ker}(f)) = n - r$, se n é a dimensión de V. Sexa $C = \{e_1, \dots, e_{n-r}\}$ unha base de $\operatorname{Ker}(f)$ e fagamos unha ampliación ata obter unha base de V dada por

$$B = \{e_1, \dots, e_{n-r}, e_{n-r+1}, \dots, e_n\}.$$

Sexa $\dim_{\mathbb{K}}(W) = m$ e tomemos agora unha base de W dada por $D = \{w_1, \dots, w_m\}$. Como $C = \{e_1, \dots, e_{n-r}\}$ é unha base de $\operatorname{Ker}(f)$ temos que

$$M(f)_{B,D} = \left(\theta_{\mathbb{K}^m} \mid \cdots \mid \theta_{\mathbb{K}^m} \mid f(e_{n-r+1})_D \mid \cdots \mid f(e_n)_D \right)$$

e, entón, $\operatorname{rang}(M(f)_{B,D}) = r$, xa que $\{f(e_{n-r+1}), \dots, f(e_n)\}$ é unha base da imaxe de f e, en particular, un sistema libre cuxo rango coincide co rango do sistema

$$\{f(e_{n-r+1})_D, \dots, f(e_n)_D\}$$

formado polos vectores de coordenadas na base D.

Como todas as matrices asociadas a unha mesma aplicación linear teñen o mesmo rango por ser equivalentes, podemos enunciar o seguinte resultado cuxa demostración atópase nas liñas que preceden a este parágrafo.

Teorema 5.4.24. *Sexa $f : V \to W$ unha aplicación linear entre dous $\mathbb{K}$-espazos vectoriais. Cúmprese que a dimensión da imaxe de f é igual ao rango de calquera das súas matrices asociadas.*

Nota 5.4.25. No exemplo oitavo de Exemplos 5.2.17 vimos que se $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ e U é o subespazo de $\mathbb{K}^n$ dado polas solucións do sistema homoxéneo $Ax = \theta_{\mathbb{K}^n}$, isto é

$$U = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n \mid A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \right\},$$

entón,

$$\dim_{\mathbb{K}}(U) = n - \operatorname{rang}(A).$$

Utilizando o visto anteriormente podemos xustificar a fórmula anterior do seguinte xeito: tomemos a aplicación linear $f_A : \mathbb{K}^n \to \mathbb{K}^m$. Daquela, como

$$U = \operatorname{Ker}(f_A)$$

e $A = M(f_A)_{C_n, C_m}$, pódese afirmar que

$$n = \dim_{\mathbb{K}}(\mathbb{K}^n) = \dim_{\mathbb{K}}(\operatorname{Ker}(f_A)) + \dim_{\mathbb{K}}(\operatorname{Im}(f_A))$$

é equivalente a

$$\dim_{\mathbb{K}}(U) = n - \operatorname{rang}(A).$$

5.5. Problemas propostos

1. Estude cales dos seguintes conxuntos son subespazos de $\mathbb{R}^4$:

$$U_1 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \mid x \in \mathbb{Z} \right\},$$

$$U_2 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \mid x - z - t = 1 \right\},$$

$$U_3 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \mid x - z + t = y \right\},$$

$$U_4 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \mid |y| = |z| \right\},$$

$$U_5 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \mid x^2 = y^2 + t^2 \right\}.$$

2. Considérase o conxunto $U = \{A \in \mathcal{M}_{n\times n}(\mathbb{R}) \mid A - A^t = \mathrm{tr}\,(A) I_n\}$.
   a) Probe que U é un subespazo vectorial de $\mathcal{M}_{n\times n}(\mathbb{R})$.
   b) Para $n = 2$, atope unha base e a dimensión de U.
3. Considéranse os seguintes subconxuntos de $\mathcal{M}_{2\times 2}(\mathbb{R})$:

$$\begin{aligned}
U_1 &= \{A \in \mathcal{M}_{2\times 2}(\mathbb{R}) \mid \mathrm{rang}\,(A) = 1\},\\
U_2 &= \{A \in \mathcal{M}_{2\times 2}(\mathbb{R}) \mid \det(A) = 0\},\\
U_3 &= \{A \in \mathcal{M}_{2\times 2}(\mathbb{R}) \mid A + A^t = 0\},\\
U_4 &= \left\{A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R}) \mid \mathrm{tr}\,(A) = b = 0\right\},\\
U_5 &= \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R}) \mid a + c = a - c = 0\right\}.
\end{aligned}$$

   a) Estude cales son subespazos vectoriais de $\mathcal{M}_{2\times 2}(\mathbb{R})$.
   b) Para cada un dos subespazos obtidos no apartado anterior, ache unha base e determine a dimensión de cada un deles.
4. Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Demostre que o conxunto de solucións da ecuación matricial

$$AXA^t = \theta_{\mathcal{M}_{n\times n}(\mathbb{R})}$$

é un subespazo vectorial de $\mathcal{M}_{n\times n}(\mathbb{R})$.

5. Indique cales das seguintes matrices

$$A = \begin{pmatrix} 6 & -8 \\ -1 & -8 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 6 & 0 \\ 3 & 8 \end{pmatrix} \quad , \quad C = \begin{pmatrix} -1 & 5 \\ 7 & 1 \end{pmatrix},$$

son combinación linear dos vectores do sistema

$$S = \left\{ \begin{pmatrix} 4 & 0 \\ -2 & -2 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 1 & 4 \end{pmatrix} \right\}.$$

6. En $\mathbb{R}^5$ considéranse os subespazos

$$U = \langle \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \\ 7 \\ 2 \end{pmatrix} \right\} \rangle,$$

$$W = \langle \left\{ \begin{pmatrix} 0 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \\ 1 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \\ 1 \\ -4 \end{pmatrix} \right\} \rangle.$$

Estude se U é un subespazo de W. Son U e W iguais?

7. No espazo vectorial $\mathbb{R}^4$, considérase o subespazo vectorial

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \mid y - z - t = y - z + t = 0 \right\}.$$

a) Ache unha base e a dimensión de U.

b) Estude se o conxunto

$$S = \left\{ \begin{pmatrix} 1 \\ 2 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

é un sistema xerador de U.

c) Atope un vector $v \in \mathbb{R}^4$ que non pertenza ao subespazo xerado polo sistema S.

d) Calcule unha base C de U que conteña ao vector $u = \begin{pmatrix} 3 \\ 4 \\ 4 \\ 0 \end{pmatrix}$.

e) É posíbel escribir o vector

$$v = \begin{pmatrix} 2 \\ -1 \\ -1 \\ 0 \end{pmatrix}$$

como combinación linear de elementos de S de dúas formas distintas? En caso de resposta afirmativa, escriba as dúas formas.

8. En $\Pi_2(\mathbb{R})$ consideramos o subespazo $U = \langle \{p_1(x), p_2(x), p_3(x)\} \rangle$ onde

$$p_1(x) = \alpha + 5x + x^2, \quad p_2(x) = 2x - 2x^2, \quad p_3(x) = 2 + \beta x^2.$$

a) Calcule, segundo os distintos valores de $\alpha, \beta \in \mathbb{R}$, a dimensión de U.

b) Para que valores de α e β se pode expresar o polinomio $p(x) = 2 + 9x + 3x^2$ como combinación linear de $p_1(x), p_2(x)$ e $p_3(x)$?

c) Obteña os valores de α e β para os que é única a expresión do apartado anterior.

9. Sexa

$$B = \left\{ \begin{pmatrix} 3 \\ 6 \\ 3 \\ -6 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -8 \\ -12 \\ -4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix} \right\}.$$

a) Probe que B é unha base de $\mathbb{R}^4$.

b) Atope as coordenadas do vector

$$w = \begin{pmatrix} 0 \\ -1 \\ -1 \\ 0 \end{pmatrix} \in \mathbb{R}^4$$

respecto á base B.

c) Calcule o vector de coordenadas respecto á base B de $v = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$.

d) Sexa o subespazo

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \mid t - x + y - z = 0 \right\}.$$

Ache a dimensión e unha base C de U.

e) É $w = (0, -1, -1, 0)^t$ un vector de U? En caso de resposta afirmativa, atope as coordenadas do vector w respecto da base C de U.

10. Determine se os seguintes sistemas xeran $\Pi_2(\mathbb{R})$:

a) $S = \{3 + x, 1 - x + 2x^2, -2 - 2x + 2x^2\}$.

b) $T = \{2 + x + 4x^2, 1 - x + 3x^2, 3 + 2x + 5x^2\}$.

11. Ache a dimensión e unha base dos subespazos

a) $$U_1 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{array}{r} x + y + z = 0 \\ 3x + 2y - 2z = 0 \\ 4x + 3y - z = 0 \\ 6x + 5y + z = 0 \end{array} \right\},$$

b) $$U_2 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \;\middle|\; \begin{array}{r} 3x + y + z + t = 0 \\ 5x - y + z - t = 0 \end{array} \right\},$$

c) $$U_3 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \mid x - 4y + 3z - t = 0 \right\}.$$

12. Ache a dimensión e unha base do subespazo xerado polos vectores:

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -3 \\ 3 \\ 7 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} -1 \\ 3 \\ 9 \\ 3 \end{pmatrix}, v_4 = \begin{pmatrix} -5 \\ 3 \\ 5 \\ -1 \end{pmatrix}.$$

Calcule as coordenadas de cada un deles respecto á base calculada.

13. Sexa $V = \{A \in \mathcal{M}_{2\times 3}(\mathbb{R}) \mid MA = A\}$ onde $M = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R})$.
   a) Probe que V é un subespazo vectorial de $\mathcal{M}_{2\times 3}(\mathbb{R})$ de dimensión 3.
   b) Determine os valores de $\alpha, \beta \in \mathbb{R}$ para os que o conxunto

$$B_{\alpha,\beta} = \left\{ \begin{pmatrix} -2 & 0 & 2 \\ 1 & 0 & -1 \end{pmatrix}, \begin{pmatrix} -2 & \alpha & 0 \\ 1 & \alpha & 0 \end{pmatrix}, \begin{pmatrix} 0 & \beta & 0 \\ 0 & \alpha+1 & 0 \end{pmatrix} \right\}$$

   é unha base de V.

14. En $\Pi_3(\ R)$ considéranse as bases $B = \{1+x^3, 1+x^2, 1+x, 1\}$ e $C = \{1, x, x^2, x^3\}$.
   a) Determine as matrices de cambio de base de C a B e de B a C.
   b) Ache as coordenadas do polinomio $p(x) = x + x^2 + x^3$ respecto da base B.

15. En $\mathbb{R}^3$ considérase a base

$$B = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Sexa B' outra base de $\mathbb{R}^3$ tal que a matriz de cambio de coordenadas de B' a B é

$$P_{B',B} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix}.$$

   a) Determine a base B'.
   b) Calcule $P_{B,B'}$.
   c) Ache as coordenadas de $v = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$ respecto das bases B y B'.
   d) Atope as coordenadas do vector w na base B' sabendo que $w_B = (2, 0, -3)^t$.

16. Sexa $f : \mathbb{R}^4 \to \mathbb{R}^4$ a aplicación linear definida por

$$f\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} y - 2z - t \\ -2x + 3y - 4z - 2t \\ -x + y - z - t \\ t \end{pmatrix}.$$

   a) Ache a matriz asociada a f respecto da base canónica de $\mathbb{R}^4$.
   b) Determine o rango de f e estude se f é un isomorfismo.
   c) Calcule unha base do núcleo de f e outra da imaxe de f.

17. Sexa $f : \mathcal{M}_{2\times 2}(\mathbb{R}) \to \mathcal{M}_{2\times 2}(\mathbb{R})$ a aplicación definida por $f(X) = X + \operatorname{tr}(AX)A$, onde

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

   a) Probe que f é linear.
   b) Ache unha base de $f(S)$, onde

$$S = \left\{ \begin{pmatrix} x & y \\ z & t \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R}) \mid x + y + t = x - y + t = 0 \right\}.$$

18. Sexa $f : V \to V$ unha aplicación linear. Dadas as aplicacións lineares g e h definidas por $g = f - id_V$ e $h = f + id_V$, demostre que se $v_1 \in \text{Ker}(g)$ e $v_2 \in \text{Ker}(h)$, entón $f(v_1 + v_2) = v_1 - v_2$.
19. Sexa
$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \mid x = z + t \right\}.$$
Considéranse as aplicacións lineares $g : \mathbb{R}^4 \to \mathbb{R}^2$, dada por
$$g \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x - y + z \\ 2x + y + t \end{pmatrix},$$
e $f : U \to \mathbb{R}^4$ tal que
$$M(f)_{B,C_4} = \begin{pmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$
onde $B = \{(0,1,0,0)^t, (1,0,1,0)^t, (1,0,0,1)^t\}$ é unha base de U e C_4 é a base canónica de $\mathbb{R}^4$.
   a) Ache $M(g)_{C_4,D}$ e $M(g \circ f)_{B,D}$, onde $D = \{(1,0), (0,1)\}$.
   b) Determine unha base do subespazo $f(W)$, onde
   $$W = \langle \{(1,1,1,0)^t, (0,0,1,-1)^t\} \rangle.$$
   c) Atope unha base de $\text{Ker}(g \circ f)$.
20. Sexa $f : \mathbb{R}^4 \to \mathbb{R}^4$ a aplicación linear definida por
$$f \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 3x + 4t \\ 2x - y + 2t \\ 2x - z + 2t \\ -2x - 3t \end{pmatrix}.$$
   a) Calcule a matriz asociada a f respecto da base canónica C_4 de $\mathbb{R}^4$.
   b) Probe que f é un isomorfismo e determine $M(f^{-1})_{C_4,C_4}$.
21. Sexan $U = \{A \in \mathcal{M}_{2\times2}(\mathbb{R}) \mid A = A^t\}$ e $f : U \longrightarrow \mathcal{M}_{2\times2}(\mathbb{R})$ a aplicación linear definida por
$$f \begin{pmatrix} x & y \\ y & z \end{pmatrix} = \begin{pmatrix} x - y + z & x + y + z \\ 2x + y + 2z & z + x \end{pmatrix}, \forall \begin{pmatrix} x & y \\ y & z \end{pmatrix} \in U.$$
   a) Ache a matriz asociada a f respecto das bases B e C, sendo
   $$B = \left\{ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\},$$
   $$C = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$
   b) Calcule unha base da imaxe e outra do núcleo de f. É f inxectiva? É sobrexectiva? Razoe a resposta.

22. Considéranse o subespazo $V = \{p(x) \in \Pi_3(\mathbb{R}) \mid p'(-1) = 0\}$ e a aplicación linear
$$f : \mathbb{R}^3 \to V$$
definida por
$$f\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a + (b - 3a + 2c)x + (c - b)x^2 + (a - b)x^3.$$

   a) Probe que f está ben definida, é dicir, que $f\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in V$ para todo $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$.

   b) Calcule a matriz asociada a f respecto das bases C_3 e B, onde C_3 é a base canónica de $\mathbb{R}^3$ e $B = \{1, 2x + x^2, -3x + x^3\}$.

   c) Sexa $g : \mathbb{R}^3 \to V$ outra aplicación linear tal que
$$M(g)_{C_3,B} = \begin{pmatrix} 0 & 1 & -1 \\ 2 & 2 & -1 \\ 4 & 1 & -1 \end{pmatrix}.$$
   Calcule unha base de $\operatorname{Im}(f + g)$ e outra de $(f + g)(U)$, onde
$$U = \left\langle \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} \right\} \right\rangle.$$

23. Considéranse o subespazo vectorial de $\mathcal{M}_{3\times3}(\mathbb{R})$
$$W = \{A = (a_{ij}) \in \mathcal{M}_{3\times3}(\mathbb{R}) \mid A = A^t,\, a_{11} = a_{12} = 0\}$$
e a aplicación $f : W \to \mathcal{M}_{2\times2}(\mathbb{R})$ dada por:
$$f(A) = NAN^t, \quad \forall A \in W$$
onde $N = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in \mathcal{M}_{2\times3}(\mathbb{R})$.

   a) Ache a dimensión e unha base B de W.

   b) Probe que f é linear.

   c) Calcule a matriz asociada a f respecto das bases B e C, sendo B a base obtida no primeiro apartado e
$$C = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

   d) Sexa $U = \{A = (a_{ij}) \in W \mid a_{13} + a_{23} + a_{33} = 0\}$. Ache unha base de $f(U)$.

24. Sexa $f : \Pi_2(\mathbb{R}) \to \Pi_2(\mathbb{R})$ a aplicación linear definida por
$$f(p(x)) = [2p'(0) - 2p(1)]x^2 + p(-1)x - p(1), \quad \forall\, p(x) \in \Pi_2(\mathbb{R}).$$

   a) Calcule a matriz asociada a f respecto da base $B = \{1, x, x^2\}$.

   b) Atope unha base da imaxe de f. É f inxectiva?

25. Sexa $B = \{v_1, v_2, v_3\}$ unha base de $\mathbb{R}^3$ e sexa
$$U = \{A \in M_{2\times2}(\mathbb{R}) \mid A = A^t\}.$$
   Considérase a aplicación linear $f : \mathbb{R}^3 \to U$ tal que
$$f(v_1) = \begin{pmatrix} -2 & 1 \\ 1 & -3 \end{pmatrix}, \; f(v_2) = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}, \; f(v_3) = \begin{pmatrix} 4 & -1 \\ -1 & -1 \end{pmatrix}.$$

a) Calcule a matriz asociada a f respecto das bases B e D, onde

$$D = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

b) Ache unha base de $f(W)$, onde

$$W = \langle \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}_B, \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}_B, \begin{pmatrix} 1 \\ 3 \\ -3 \end{pmatrix}_B \right\} \rangle.$$

CAPÍTULO 6

Diagonalización e funcións de matrices

Os contidos dos capítulos anteriores estiveron relacionados principalmente coas nocións básicas da teoría de matrices, a solución de sistemas de ecuacións lineares, a teoría de espazos vectoriais e a de aplicacións lineares. A partir de agora centrarémonos noutro tipo de problemas. Máis concretamente, prestaremos a nosa atención a aqueles relacionados con saber cando una matriz é semellante a unha matriz triangular superior ou a unha matriz diagonal. O estudo deste tipo de cuestións comporta a introdución das nocións de autovalor e autovector que, como se puxo de manifesto no esbozo histórico, apareceron asociadas a certos problemas relacionados con sistemas de ecuacións diferenciais.

En primeiro lugar, e despois de introducir as definicións de autovector e autovalor, veremos que, para toda matriz cadrada $A \in \mathcal{M}_{n\times n}(\mathbb{K})$, os seus autovalores son xustamente as raíces dun polinomio de grao n que se chama polinomio característico e que denotaremos como $p_A(x)$. Presentarase como se expresa este polinomio en función dun determinante e comprobaremos que A é invertíbel se, e só se, o número cero non é autovalor de A. Un feito relevante que se debe salientar é que, aínda que unha matriz teña todos os seus coeficientes reais, pode ocorrer que aparezan autovalores complexos xa que estes non son máis nada que raíces dun polinomio. Tendo en conta isto, definirase a multiplicidade alxébrica de cada autovalor λ, denotada por $\mathrm{ma}(\lambda)$, e se $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ é o conxunto de autovalores de A, onde cada un aparece tantas veces como indica a súa multiplicidade alxébrica, utilizando as propiedades básicas das raíces de polinomios, téñense as identidades:

$$\det(A) = \lambda_1\lambda_2\cdots\lambda_n, \qquad \mathrm{tr}\,(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

A continuación introducirase a noción de subespazo propio asociado a un autovalor λ e probarase que o seu cálculo redúcese á solución dun sistema de ecuacións lineares homoxéneo. A dimensión deste subespazo, denotado por $V(\lambda)$, é o que se coñece como multiplicidade xeométrica do autovalor λ, denotada por $\mathrm{mg}(\lambda)$, e esta calcúlase como

$$\mathrm{mg}(\lambda) = n - \mathrm{rang}\,(A - \lambda I_n).$$

Cando os coeficientes de A son todos reais e $\lambda \in \mathbb{R}$, tense que $V(\lambda)$ é un subespazo do $\mathbb{R}$-espazo vectorial $\mathbb{R}^n$. Agora ben, se aparecen autovalores complexos, $V(\lambda)$ é un subespazo do $\mathbb{C}$-espazo vectorial $\mathbb{C}^n$. Finalmente, neste apartado probarase que a multiplicidade xeométrica é sempre maior ou igual que un e menor ou igual que a multiplicidade alxébrica.

No seguinte punto do capítulo introduciremos en primeiro lugar a noción de semellanza de matrices e probaremos que dúas matrices semellantes teñen o mesmo polinomio característico e, por tanto, os mesmos autovalores. O resultado importante desta sección afirma que para toda matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ as condicións seguintes son equivalentes:

1) O polinomio característico de A descomponse como

$$p_A(x) = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x)$$

onde os escalares $\lambda_i \in \mathbb{K}$, $i \in \{1, \dots, n\}$ non son necesariamente distintos.

2) A matriz A é semellante a unha matriz triangular superior.

Se se cumpre algunha das anteriores condicións, os autovalores de A son os elementos da diagonal principal da matriz triangular superior semellante a A.

Todo o anterior prepáranos para comezar a estudar o problema de diagonalización de matrices. Unha matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ dirase que é diagonalizábel se é semellante a unha matriz diagonal, é dicir, se existen dúas matrices P e D tales que P é invertíbel, D é diagonal e

$$P^{-1}AP = D,$$

equivalentemente, $AP = PD$ ou tamén $A = PDP^{-1}$.

Se P se escribe en columnas como $P = \left(u_1 \mid u_2 \mid \cdots \mid u_n \right)$ e

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

tense que a igualdade $AP = PD$ é equivalente a

$$\left(Au_1 \mid Au_2 \mid \cdots \mid Au_n \right) = \left(\lambda_1 u_1 \mid \lambda_2 u_2 \mid \cdots \mid \lambda_n u_n \right).$$

Como consecuencia, os autovalores de A son $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ e $B = \{u_1, u_2, \dots, u_n\}$ é unha base de $\mathbb{K}^n$ formada por autovectores de A xa que $Au_i = \lambda_i u_i$ para $i \in \{1, 2, \dots, n\}$.

Inversamente, se existe unha base $B = \{u_1, u_2, \dots, u_n\}$ de $\mathbb{K}^n$ formada por autovectores de A, tense que, tomando P a matriz cuxas columnas son os vectores da base B, obtemos unha matriz invertíbel tal que

$$AP = \left(Au_1 \mid Au_2 \mid \cdots \mid Au_n \right) = \left(\lambda_1 u_1 \mid \lambda_2 u_2 \mid \cdots \mid \lambda_n u_n \right) = PD$$

onde

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

ou, equivalentemente,

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Polo tanto, temos que A é diagonalizábel, se, e só se, o $\mathbb{K}$-espazo vectorial $\mathbb{K}^n$ admite unha base formada por autovectores da matriz A.

Para obter unha condición que nos permita verificar de forma máis fácil cando una matriz é diagonalizábel, en primeiro lugar, demostraremos ás seguintes propiedades que se cumpren para toda matriz cadrada A.

1) Autovectores asociados a autovalores distintos son independentes.
2) Se con $\mathrm{Sp}(A)$ (espectro de A) denotamos o conxunto dos autovalores de A, tense que

$$\bigcap_{\lambda \in \mathrm{Sp}(A)} V(\lambda) = \{\theta_{\mathbb{K}^n}\}.$$

3) Se $\{\lambda_1, \dots, \lambda_r\}$ son autovalores distintos de A e B_i é unha base de $V(\lambda_i)$, tense que o sistema $B = B_1 \cup \cdots \cup B_r$ é libre.

Tendo en conta o anterior, poderase deducir un criterio sinxelo para saber cando una matriz é diagonalizábel ou non. Este criterio é o seguinte: para toda matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$, son equivalentes:

1) O polinomio característico de A descomponse como

$$p_A(x) = (\lambda_1 - x)^{\alpha_1}(\lambda_2 - x)^{\alpha_2} \cdots (\lambda_r - x)^{\alpha_r}$$

e $\mathrm{mg}(\lambda_i) = \alpha_i = \mathrm{ma}(\lambda_i)$.

2) A matriz A é diagonalizábel.

Na parte final do capítulo centrarémonos na teoría de polinomios de matrices, polinomios anuladores e funcións de matrices. De forma máis concreta, en primeiro lugar indicarase como se pode calcular a inversa dunha matriz cadrada invertíbel se coñecemos un polinomio anulador para ela e, en segundo lugar, introducirase o Teorema de Cayley–Hamilton o cal garante que toda matriz cadrada admite como polinomio anulador o seu polinomio característico. Como consecuencia, se $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ e $p(x)$ é un polinomio de grao $k \geq n$, existe un polinomio $r(x)$ de grao menor ou igual que n tal que $p(A) = r(A)$. Para calcular $r(x)$ podemos proceder da seguinte forma: se λ é un autovalor de A tense que $p(\lambda) = p_A(\lambda)t(\lambda) + r(\lambda) = r(\lambda)$ xa que $p_A(\lambda) = 0$. Isto implica que os polinomios $p(x)$ e $r(x)$ deben tomar o mesmo valor sobre todos os autovalores de A. Da mesma forma probaríase que, se a multiplicidade alxébrica de λ é m, tense que

$$p^{(k)}(\lambda) = r^{(k)}(\lambda)$$

para todo k pertencente ao conxunto $\{1, \dots, m-1\}$. Isto podémolo facer para cada autovalor e, deste xeito, obtemos un sistema de ecuacións lineares compatíbel determinado cuxas incógnitas son os coeficientes de $r(x)$.

A idea anterior poñeranos sobre a pista de como calcular funcións de matrices para funcións reais de variábel real que cumpran a condición de ser analíticas, isto é, funcións que se poden obter como límites de polinomios. Entre elas están as funcións racionais, as raíces k-ésimas, as funcións trigonométricas seno e coseno, a función exponencial, a función logaritmo, etc. A forma de facer isto pasa por definir a noción de función definida sobre o espectro dunha matriz.

Se A é unha matriz cadrada de n filas e $f : I \subset \mathbb{R} \to \mathbb{R}$ é unha función como as anteriores, cando

$$p_A(x) = (\lambda_1 - x)^{\alpha_1}(\lambda_2 - x)^{\alpha_2} \cdots (\lambda_r - x)^{\alpha_r}$$

con $\lambda_i \in \mathbb{R}$, $i \in \{1, \dots, r\}$ e $\mathrm{ma}(\lambda_i) = \alpha_i$, diremos que f está definida sobre o espectro de A, se para cada λ_i existen

$$f(\lambda_i),\ f'(\lambda_i),\ \dots\ , f^{(\alpha_i - 1)}(\lambda_i).$$

Se f está definida sobre o espectro de A, con

$$V_{f,A} = \{f^{(k)}(\lambda_i),\ |\ \ i \in \{1, \dots, r\},\ k \in \{1, \dots, \alpha_i - 1\}\}$$

denotaremos o conxunto dos valores de f sobre o espectro de A.

Tendo en conta o anterior, sempre é posíbel atopar un único polinomio, chamado polinomio interpolador de Lagrange- Sylvester,

$$r(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1},$$

de grao menor que n, tal que

$$V_{f,A} = V_{r,A},$$

sen máis que resolver o sistema de ecuacións lineares

$$f^{(k)}(\lambda_i) = r^{(k)}(\lambda_i),\ |\ \ i \in \{1, \dots, r\},\ k \in \{1, \dots, \alpha_i - 1\}$$

cuxas incógnitas serán os coeficientes a_k do polinomio $r(x)$. Por tanto, definirase $f(A) = r(A)$. En particular, esta definición conduciranos á forma que permite introducir a exponencial e^{tA} dunha matriz A.

Finalízase o capítulo indicando que, se

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

é unha matriz diagonal e f é unha función definida sobre o espectro de A, tense que

$$f(D) = \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix}.$$

Por tanto, se A é diagonalizábel e P é a matriz invertíbel tal que $P^{-1}AP = D$ ou, equivalentemente, que

$$A = PDP^{-1},$$

obtemos as igualdades

$$f(A) = Pf(D)P^{-1} = P \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix} P^{-1}.$$

6.1. Autovalores e autovectores. Polinomio característico

Definición 6.1.1. Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{K})$. Un escalar $\lambda \in \mathbb{K}$ chámase autovalor ou valor propio de A, se existe un vector $v \in \mathbb{K}^n$ non nulo tal que

$$Av = \lambda v.$$

O vector $v \neq \theta_{\mathbb{K}^n}$ cumprindo a anterior igualdade chámase autovector, ou vector propio, de A asociado ao autovalor λ.

Destas definicións séguese facilmente que un mesmo autovector v non pode estar asociado a dous autovalores distintos xa que, se existisen λ_1, $\lambda_2 \in \mathbb{K}$ tales que

$$\lambda_1 v = Av = \lambda_2 v,$$

obteriamos a igualdade $(\lambda_1 - \lambda_2)v = \theta_{\mathbb{K}^n}$, e como v é non nulo, tense que $\lambda_1 - \lambda_2 = 0$, ou equivalentemente, $\lambda_1 = \lambda_2$.

Doutra banda é claro que, se λ é un autovector de A e v un autovector asociado a λ, o vector αv tamén é un autovector asociado a λ para todo $\alpha \in \mathbb{K}$. Así pois, un autovalor ten asociados infinitos autovectores distintos.

O conxunto de todos os autovalores da matriz A chámase espectro de A e represéntase por

$$\mathrm{Sp}(A).$$

Definición 6.1.2. Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{K})$. O polinomio de grao n dado por

$$p_A(x) = \det(A - xI_n)$$

chámase polinomio característico de A.

Teorema 6.1.3. *Para toda matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ o conxunto $\mathrm{Sp}(A)$ coincide coas raíces de $p_A(x)$. Dito doutra forma, os autovalores de A son as raíces do seu polinomio característico.*

Demostración. Se $\lambda \in \mathrm{Sp}(A)$, existe un $v \neq \theta_{\mathbb{K}^n}$ tal que $Av = \lambda v$, ou ben, tal que $(A - \lambda I_n)v = \theta_{\mathbb{K}^n}$. Por tanto, $v \in \mathrm{Ker}(f_{A-\lambda I_n})$, co cal, $\dim_{\mathbb{K}}(\mathrm{Im}(f_{A-\lambda I_n})) = \mathrm{rang}\,(A - \lambda I_n) < n$. Por tanto tense que $A - \lambda I_n$ non é invertíbel, ou equivalentemente, $\det(A - \lambda I_n) = 0$. Así pois, λ é unha raíz do polinomio característico.

Reciprocamente, se $\lambda \in \mathbb{K}$ é tal que $\det(A - \lambda I_n) = 0$, obtemos que $\mathrm{rang}\,(A - \lambda I_n) < n$ e, por tanto, $\dim_{\mathbb{K}}(\mathrm{Ker}(f_{A-\lambda I_n})) \geq 1$. Isto implica que existe $v \neq \theta_{\mathbb{K}^n}$ pertencente ao subespazo $\mathrm{Ker}(f_{A-\lambda I_n})$. Equivalentemente v é un autovector asociado a λ xa que

$$f_{A-\lambda I_n}(v) = (A - \lambda I_n)v = \theta_{\mathbb{K}^n} \Leftrightarrow Av = \lambda v.$$

$\square$

Nota 6.1.4. Como acabamos de comprobar no anterior resultado o cálculo dos autovalores dunha matriz está ligado ao cálculo das raíces do polinomio característico. Se a matriz ten coeficientes reais, o seu polinomio característico tamén; agora ben, pode non ter ningunha raíz real e, por tanto, se queremos calcular o seu espectro, temos que contar cos números complexos. Por exemplo, isto ocorre, se

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

xa que

$$p_A(x) = \begin{vmatrix} -x & -1 \\ 1 & -x \end{vmatrix} = x^2 + 1.$$

Lémbrese que, se $p(x)$ é un polinomio de grao n con coeficientes reais ou complexos, $\lambda \in \mathbb{K}$ é unha raíz de $p(x)$, se $p(\lambda) = 0$. Isto é equivalente a que $p(x)$ sexa divisíbel por $x - \lambda$, ou ben, $p(x) = (x - \lambda)q(x)$, onde $q(x)$ é un polinomio de grao $n - 1$.

Se os coeficientes de $p(x)$ son complexos calquera, o Teorema Fundamental da Álxebra afirma que

$$p(x) = a(x - \lambda_1)^{\alpha_1}(\lambda_2 - x)^{\alpha_2} \cdots (x - \lambda_r)^{\alpha_r}$$

onde $\lambda_i \in \mathbb{C}$, $i \in \{1, \dots, r\}$, son as raíces de $p(x)$ e $\alpha_i \in \mathbb{N}$, $i \in \{1, \dots, r\}$, cumpren que

$$\alpha_1 + \cdots + \alpha_r = n.$$

Para cada raíz λ_i o número natural α_i denota a súa multiplicidade alxébrica, isto é, o maior número natural tal que

$$p(x) = (x - \lambda_i)^{\alpha_i} t(x).$$

Por tanto, un polinomio de grao n con coeficientes reais ou complexos ten como máximo n raíces distintas que poden ser reais ou complexas. Isto implica que toda matriz cadrada de orde n, con coeficientes reais ou complexos, ten como máximo n autovalores distintos que poden ser reais ou complexos.

Tamén se debe resaltar a seguinte propiedade das raíces dos polinomios: se $p(x)$ ten todos os seus coeficientes reais, cúmprese que

$$\lambda = a + bi \in \mathrm{Sp}(A) \Leftrightarrow \overline{\lambda} = a - bi \in \mathrm{Sp}(A).$$

Desta forma, se unha matriz A cuxos coeficientes son todos reais, ten un autovalor complexo, o seu conxugado tamén é autovalor de A.

Finalmente, utilizando propiedades básicas das raíces de polinomios téñense as identidades:

$$\det(A) = \lambda_1\lambda_2 \cdots \lambda_n, \qquad \mathrm{tr}\,(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n,$$

onde $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ é o conxunto de autovalores de A contados segundo a súa multiplicidade alxébrica.

Exemplos 6.1.5. 1) Dada a matriz

$$A = \begin{pmatrix} 0 & -1 & -2 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

o polinomio característico é

$$p_A(x) = \det(A - xI_3) = \begin{vmatrix} -x & -1 & -2 \\ 0 & -x & 2 \\ 1 & 1 & 1-x \end{vmatrix} = -x^3 + x^2 - 2.$$

Logo, como $p_A(-1) = 0$, tense que

$$p_A(x) = (x+1)(-x^2 + 2x - 2).$$

As raíces de $-x^2 + 2x - 2$, que son complexas e conxugadas, veñen dadas por

$$x = \frac{-2 \pm \sqrt{4-8}}{-2} = \frac{-2 \pm 2i}{-2} = 1 \pm i.$$

Por tanto, $p_A(x) = (x+1)(x-(1-i))(x-(1+i))$ e isto implica que o espectro de A é

$$\mathrm{Sp}(A) = \{-1, 1-i, 1+i\}$$

sendo

$$\mathrm{ma}(-1) = \mathrm{ma}(1-i) = \mathrm{ma}(1+i) = 1.$$

2) Se a matriz A é triangular, ben superior, ben inferior, tense que os elementos que forman a súa diagonal principal son o seu espectro. Por exemplo, se

$$A = \begin{pmatrix} -1 & 2 & 1 & 3 & 4 \\ 0 & -1 & 1 & 1 & 1 \\ 0 & 0 & 3-i & -2 & -3 \\ 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix},$$

o polinomio característico é

$$p_A(x) = \det(A - xI_3) = \begin{vmatrix} -1-x & 2 & 1 & 3 & 4 \\ 0 & -1-x & 1 & 1 & 1 \\ 0 & 0 & (3-i)-x & -2 & -3 \\ 0 & 0 & 0 & -x & -3 \\ 0 & 0 & 0 & 0 & -3-x \end{vmatrix} =$$

$$x(3+x)(1+x)^2((3-i)-x).$$

e, por tanto, $\mathrm{Sp}(A) = \{0, -3, -1, 3-i\}$ sendo a multiplicidade alxébrica de cada autovalor

$$\mathrm{ma}(0) = \mathrm{ma}(-3) = \mathrm{ma}(3-i) = 1, \;\; \mathrm{ma}(-1) = 2.$$

Proposición 6.1.6. *Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{K})$. A matriz A é invertíbel, se, e só se, 0 non pertence ao espectro de A. Ademais, se A é invertíbel, cúmprese que $\lambda \in \mathrm{Sp}(A)$, se, e só se, $\lambda^{-1} \in \mathrm{Sp}(A^{-1})$.*

Demostración. Se $0 \in \mathrm{Sp}(A)$, temos que existe un vector non nulo en $\mathbb{K}^n$ tal que $Av = 0v = \theta_{\mathbb{K}^n}$. Isto implica que o sistema $Ax = \theta_{\mathbb{K}^n}$ ten solucións distintas da trivial e, por tanto, $\mathrm{rang}\,(A) < n$ ou, equivalentemente, A non é invertíbel. Se A non é invertíbel, tense que $\mathrm{rang}\,(A) < n$. Por tanto o sistema $Ax = \theta_{\mathbb{K}^n}$ é compatíbel indeterminado e isto implica que existe un vector non nulo en $\mathbb{K}^n$ tal que $Av = \theta_{\mathbb{K}^n} = 0v$. Como consecuencia, o 0 é un autovalor de A.

A segunda equivalencia séguese dos seguintes feitos:

$$\lambda \in \mathrm{Sp}(A) \;\Leftrightarrow\; \exists\, v \neq \theta_{\mathbb{K}^n} \;\mid\; Av = \lambda v \;\Leftrightarrow\; \exists\, v \neq \theta_{\mathbb{K}^n} \;\mid\; v = A^{-1}\lambda v \;\Leftrightarrow$$

$$\exists\, v \neq \theta_{\mathbb{K}^n} \ \mid \ A^{-1}v = \frac{1}{\lambda}v \ \Leftrightarrow \ \lambda^{-1} \in \mathrm{Sp}(A^{-1}).$$

□

Definición 6.1.7. Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ e $\lambda \in \mathbb{K}$ un autovalor de A. Chámase subespazo vectorial propio asociado a λ ao subespazo de $\mathbb{K}^n$ dado por

$$V(\lambda) = \mathrm{Ker}(f_{A-\lambda I_n}).$$

Tendo en conta a definición, temos que

$$V(\lambda) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n \ \mid \ (A - \lambda I_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \right\}$$

e

$$\dim_{\mathbb{K}}(V(\lambda)) = n - \mathrm{rang}\,(A - \lambda I_n).$$

Cando os coeficientes de A son todos reais e $\lambda \in \mathbb{R}$, tense que $V(\lambda)$ é un subespazo do $\mathbb{R}$-espazo vectorial $\mathbb{R}^n$.

Chámase multiplicidade xeométrica dun autovalor λ, denotada por $\mathrm{mg}(\lambda)$, á dimensión do subespazo propio $V(\lambda)$.

Exemplos 6.1.8. 1) Tomemos a matriz

$$A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}.$$

O polinomio característico de A é

$$p_A(x) = \det(A - xI_3) = \begin{vmatrix} 1-x & 3 & 3 \\ -3 & -5-x & -3 \\ 3 & 3 & 1-x \end{vmatrix} = \begin{vmatrix} 1-x & 3 & 3 \\ 0 & -2-x & -2-x \\ 3 & 3 & 1-x \end{vmatrix} =$$

$$= \begin{vmatrix} 1-x & 3 & 0 \\ 0 & -2-x & 0 \\ 3 & 3 & -2-x \end{vmatrix} = -(2+x) \begin{vmatrix} 1-x & 3 \\ 0 & -2-x \end{vmatrix} = (2+x)^2(1-x),$$

e isto implica que

$$\mathrm{Sp}(A) = \{-2, 1\}$$

sendo

$$\mathrm{ma}(-2) = 2, \ \ \mathrm{ma}(1) = 1.$$

Os subespazos propios calcúlanse da seguinte forma:

$$V(-2) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \ \mid \ (A + 2I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \ \mid \ \begin{pmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \ \mid \ x + y + z = 0 \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ -x-y \end{pmatrix} \in \mathbb{R}^3 \ \mid \ x, y \in \mathbb{R} \right\} = \langle \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\} \rangle,$$

$$V(1)=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{R}^3 \mid (A-I_3)\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\right\}$$

$$=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{R}^3 \mid \begin{pmatrix}0&3&3\\-3&-6&-3\\3&3&0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\right\}$$

$$=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{R}^3 \mid y+z=0=x+y\right\}=\left\{\begin{pmatrix}-y\\y\\-y\end{pmatrix}\in\mathbb{R}^3 \mid y\in\mathbb{R}\right\}=\langle\left\{\begin{pmatrix}-1\\1\\-1\end{pmatrix}\right\}\rangle.$$

Por tanto, $\mathrm{mg}(-2)=2$ e $\mathrm{mg}(1)=1$.

2) Dada a matriz

$$A=\begin{pmatrix}0&-1&-2\\0&0&2\\1&1&1\end{pmatrix},$$

no primeiro caso de Exemplos 6.1.5 demostramos que $\mathrm{Sp}(A)=\{-1,1-i,1+i\}$ xa que $p_A(x)=(x+1)(x-(1-i))(x-(1+i))$. Neste caso os subespazos propios asociados son:

$$V(-1)=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{C}^3 \mid (A+I_3)\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\right\}$$

$$=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{C}^3 \mid \begin{pmatrix}1&-1&-2\\0&1&2\\1&1&2\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\right\}$$

$$=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{C}^3 \mid \begin{cases}x-y-2z=0\\y+2z=0\\x+y+z=0\end{cases}\right\}=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{C}^3 \mid \begin{matrix}x=0\\y=-2z\end{matrix}\right\}$$

$$=\left\{\begin{pmatrix}0\\-2z\\z\end{pmatrix}\in\mathbb{C}^3 \mid z\in\mathbb{C}\right\}=\langle\left\{\begin{pmatrix}0\\-2\\1\end{pmatrix}\right\}\rangle,$$

$$V(1-i)=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{C}^3 \mid (A-(1-i)I_3)\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\right\}$$

$$=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{C}^3 \mid \begin{pmatrix}-1+i&-1&-2\\0&-1+i&2\\1&1&i\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\right\}$$

$$=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{C}^3 \mid \begin{matrix}(-1+i)x-y-2z=0\\(-1+i)y+2z=0\\x+y+iz=0\end{matrix}\right\}$$

$$=\left\{\begin{pmatrix}(-1-2i)z\\(1+i)z\\z\end{pmatrix}\in\mathbb{C}^3 \mid z\in\mathbb{C}\right\}=\langle\left\{\begin{pmatrix}-1-2i\\1+i\\1\end{pmatrix}\right\}\rangle$$

e facendo uns cálculos similares obtemos que

$$V(1+i)=\langle\left\{\begin{pmatrix}-1+2i\\1-i\\1\end{pmatrix}\right\}\rangle.$$

Neste exemplo $\mathrm{mg}(-1)=\mathrm{mg}(1-i)=\mathrm{mg}(1+i)=1$.

Definición 6.1.9. Sexan $A, B \in \mathcal{M}_{n\times n}(\mathbb{K})$. Dirase que A e B son matrices semellantes, se existe unha matriz invertíbel $P \in \mathcal{M}_{n\times n}(\mathbb{K})$ tal que

$$P^{-1}AP = B.$$

Evidentemente, se dúas matrices son semellantes, son equivalentes e disto séguese que teñen o mesmo rango.

Proposición 6.1.10. *Se dúas matrices son semellantes, teñen o mesmo polinomio característico. Por tanto teñen o mesmo espectro.*

Demostración. Sexan $A, B \in \mathcal{M}_{n\times n}(\mathbb{K})$ matrices semellantes e sexa $P \in \mathcal{M}_{n\times n}(\mathbb{K})$ invertíbel tal que $P^{-1}AP = B$. Por tanto:

$$p_B(x) = \det(B - xI_n) = \det(P^{-1}AP - xP^{-1}P) = \det(P^{-1}(A - xI_n)P) =$$
$$\det(P^{-1})\det(A - xI_n)\det(P) = \det(P)^{-1}\det(A - xI_n)\det(P) = \det(P)^{-1}\det(P)\det(A - xI_n)$$
$$= \det(A - xI_n) = p_A(x).$$

□

Nota 6.1.11. Na anterior proposición demostramos que, se dúas matrices son semellantes, teñen o mesmo polinomio característico, o que implica que teñen o mesmo espectro. O recíproco non é certo. As matrices $A = I_2$ e

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

teñen o mesmo polinomio característico e non son semellantes. Nótese que se o fosen, existiría $P \in \mathcal{M}_{2\times 2}(\mathbb{R})$ invertíbel tal que $P^{-1}I_2P = B$. Isto implica que $B = I_2$ o que é un absurdo.

Nota 6.1.12. Existe unha interesante relación entre os autovectores de A e os de $P^{-1}AP$. Se v é un autovector de A asociado ao autovalor λ, tense que

$$Av = \lambda v \Leftrightarrow P^{-1}Av = \lambda P^{-1}v \Leftrightarrow P^{-1}APP^{-1}v = \lambda P^{-1}v.$$

Por tanto, λ é un autovalor de $P^{-1}AP$ e $P^{-1}v$ é un autovector da mesma matriz asociado a λ.

Proposición 6.1.13. *Sexa* $\lambda \in \mathrm{Sp}(A)$ *sendo* $A \in \mathcal{M}_{n\times n}(\mathbb{K})$. *Entón,* $\mathrm{mg}(\lambda) \leqslant \mathrm{ma}(\lambda)$.

Demostración. Sexa $\lambda \in \mathrm{Sp}(A)$ e sexa $V(\lambda)$ o subespazo propio asociado a λ. Supoñamos que $\dim_{\mathbb{K}}(V(\lambda)) = r$. Tomemos unha base $\{e_1, \cdots, e_r\}$ de $V(\lambda)$ e ampliemos dita base a unha base

$$B = \{e_1, \dots, e_r, e_{r+1}, \dots, e_n\}$$

de $\mathbb{K}^n$. Por tanto, como para $i \in \{1, \dots, r\}$, $f_A(e_i) = Ae_i = \lambda e_i$, tense que

$$M(f_A)_{B,B} = \begin{pmatrix} \lambda I_r & | & U \\ --- & - & - \\ \Theta_{n-r,r} & | & Q \end{pmatrix}$$

con $U \in \mathcal{M}_{r\times(n-r)}(\mathbb{K})$ e $Q \in \mathcal{M}_{(n-r)\times(n-r)}(\mathbb{K})$.

Daquela, como A e $M(f_A)_{B,B}$ son semellantes, obtemos as igualdades

$$p_A(x) = p_{M(f_A)_{B,B}}(x) = (\lambda - x)^r P_Q(x)$$

e, por tanto,

$$\mathrm{mg}(\lambda) = r \leq \mathrm{ma}(\lambda).$$

□

Teorema 6.1.14. *Para toda matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ as condicións seguintes son equivalentes:*

1) *O polinomio característico de A descomponse como*
$$p_A(x) = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x)$$
onde os $\lambda_i \in \mathbb{K}$, $i \in \{1, \dots, n\}$, son escalares non necesariamente distintos.

2) *A matriz A é semellante a unha matriz triangular superior.*

Se se cumpre algunha das anteriores condicións, os autovalores de A son os elementos da diagonal principal da matriz triangular superior semellante a A.

Demostración. Supoñamos que a matriz A é semellante a unha matriz triangular superior
$$T = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Por tanto, usando que dúas matrices semellantes teñen o mesmo polinomio característico, obtemos as igualdades $p_A(x) = p_T(x) = (a_{11} - x)(a_{22} - x)\cdots(a_{nn} - x)$.

Para demostrar a outra implicación usaremos un razoamento indutivo. Se $n = 1$, podemos tomar $P = I_1$, xa que toda matriz cadrada de orde un é triangular superior. Supoñamos que o enunciado é certo para toda matriz cadrada de orde $n-1$ e fagamos a proba para matrices cadradas de orde n. Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ tal que
$$p_A(x) = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x)$$
cos λ_i non necesariamente distintos. Como $p_A(x)$ descomponse da forma anterior, admite todas as súas raíces en $\mathbb{K}$. Tomemos a primeira, λ_1, e un autovector v_1 asociado a este autovalor. Como v_1 é non nulo, tense que $\{v_1\}$ é un sistema libre de $\mathbb{K}^n$ e, por tanto, podemos completalo a unha base
$$B_1 = \{v_1, u_2, \dots, u_n\}$$
de $\mathbb{K}^n$. Sexa P_1 a matriz cadrada invertíbel de orde n cuxas columnas son os vectores da base B_1:
$$P_1 = \left(\, v_1 \;\middle|\; u_2 \;\middle|\; \cdots \;\middle|\; u_n \, \right).$$

Daquela,
$$\begin{aligned} P_1^{-1}AP &= \left(\, P_1^{-1}Av_1 \;\middle|\; P_1^{-1}Au_2 \;\middle|\; \cdots \;\middle|\; P_1^{-1}Au_n \, \right) = \\ &\left(\, P_1^{-1}\lambda_1 v_1 \;\middle|\; P_1^{-1}Au_2 \;\middle|\; \cdots \;\middle|\; P_1^{-1}Au_n \, \right). \end{aligned}$$

Como v_1 é a primeira columna de P_1, tense que
$$P_1^{-1}\lambda_1 v_1 = \lambda_1 P_1^{-1} v_1 = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$
e disto deducimos a igualdade
$$P_1^{-1}AP_1 = \left(\begin{array}{c|c} \lambda_1 & u \\ - & --- \\ 0 & \\ \vdots & A_1 \\ 0 & \end{array} \right)$$
onde A_1 é unha matriz cadrada de orde $n-1$. Como consecuencia,
$$p_A(x) = (\lambda_1 - x)p_{A_1}(x)$$

e, por tanto,

$$p_{A_1}(x) = (\lambda_2 - x)\cdots(\lambda_n - x).$$

Grazas á última igualdade podemos aplicar a hipótese de indución para a matriz A_1 obtendo que é semellante a unha matriz diagonal, isto é, existe $P_2 \in \mathcal{M}_{(n-1)\times(n-1)}(\mathbb{K})$ invertíbel tal que

$$P_2^{-1}A_1P_2 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} \\ 0 & a_{22} & \cdots & a_{2\,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n-1\,n-1} \end{pmatrix}.$$

Tomemos agora

$$P = \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ - & --- \\ 0 & \\ \vdots & P_2 \\ 0 & \end{array}\right).$$

Logo, P é invertíbel, xa que

$$\left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ - & --- \\ 0 & \\ \vdots & P_2 \\ 0 & \end{array}\right)\left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ - & --- \\ 0 & \\ \vdots & P_2^{-1} \\ 0 & \end{array}\right) = \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ - & --- \\ 0 & \\ \vdots & P_2P_2^{-1} \\ 0 & \end{array}\right) = I_n.$$

Ademais

$$P^{-1}P_1^{-1}AP_1P = \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ - & --- \\ 0 & \\ \vdots & P_2^{-1} \\ 0 & \end{array}\right)\left(\begin{array}{c|c} \lambda_1 & u \\ - & --- \\ 0 & \\ \vdots & A_1 \\ 0 & \end{array}\right)\left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ - & --- \\ 0 & \\ \vdots & P_2 \\ 0 & \end{array}\right) =$$

$$\left(\begin{array}{c|c} \lambda_1 & uP_2 \\ - & --- \\ 0 & \\ \vdots & P_2^{-1}A_1P_2 \\ 0 & \end{array}\right).$$

Deste xeito, como $P_2^{-1}A_1P_2$ é triangular superior, a matriz $P^{-1}P_1^{-1}AP_1P$ tamén o é e obtemos que A é semellante a unha matriz triangular superior.

□

Corolario 6.1.15. *Toda matriz $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ é semellante a unha matriz triangular superior.*

Demostración. O resultado é consecuencia do teorema anterior e do Teorema Fundamental da Álxebra. □

6.2. Matrices diagonalizábeis

Definición 6.2.1. Unha matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ dirase que é diagonalizábel se é semellante a unha matriz diagonal, é dicir, se existen dúas matrices P, $D \in \mathcal{M}_{n\times n}(\mathbb{K})$ tales que P é invertíbel, D é diagonal e

$$P^{-1}AP = D$$

ou, equivalentemente,

$$AP = PD.$$

Se P admite a seguinte expresión por columnas

$$P = \left(\begin{array}{c|c|c} u_1 & \cdots & u_n \end{array} \right)$$

e

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

tense que a igualdade $AP = PD$ é equivalente a

$$\left(\begin{array}{c|c|c|c} Au_1 & Au_2 & \cdots & Au_n \end{array} \right) = \left(\begin{array}{c|c|c|c} \lambda_1 u_1 & \lambda_2 u_2 & \cdots & \lambda_n u_n \end{array} \right).$$

Por tanto, $B = \{u_1, \dots, u_n\}$ é unha base de $\mathbb{K}^n$ formada por autovectores de A xa que $Au_i = \lambda_i u_i$ para $i \in \{1, \dots, n\}$.

De xeito inverso, se existe unha base $B = \{u_1, u_2, \dots, u_n\}$ de $\mathbb{K}^n$ formada por autovectores de A, tense que, tomando P a matriz cuxas columnas son os vectores da base B, obtemos unha matriz invertíbel tal que

$$AP = \left(\begin{array}{c|c|c|c} Au_1 & Au_2 & \cdots & Au_n \end{array} \right) = \left(\begin{array}{c|c|c|c} \lambda_1 u_1 & \lambda_2 u_2 & \cdots & \lambda_n u_n \end{array} \right) = PD,$$

onde

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

ou, equivalentemente,

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Como consecuencia do anterior, temos o seguinte resultado.

Teorema 6.2.2. *Unha matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ é diagonalizábel, se, e só se, o $\mathbb{K}$-espazo vectorial $\mathbb{K}^n$ admite unha base formada por autovectores de A.*

Para obter unha condición que nos permita verificar de forma fácil cando una matriz é diagonalizábel, en primeiro lugar, probaremos uns resultados técnicos.

Proposición 6.2.3. *Para toda matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ cúmprese que autovectores asociados a autovalores distintos son independentes.*

Demostración. A demostración farase por indución. Sexan $\lambda_1, \dots, \lambda_r$ autovalores distintos de A e sexan $v_1, \dots, v_r$ vectores non nulos de $\mathbb{K}^n$ tales que

$$Av_1 = \lambda_1 v_1, \dots, Av_r = \lambda_r v_r.$$

Se $r = 2$, o resultado é certo xa que, se

$$\alpha_1 v_1 + \alpha_2 v_2 = \theta_{\mathbb{K}^n},$$

multiplicando por $A - \lambda_1 I_n$, obtemos

$$\theta_{\mathbb{K}^n} = (A - \lambda_1 I_n)\theta_{\mathbb{K}^n} = (A - \lambda_1 I_n)(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 (A - \lambda_1 I_n) v_1 + \alpha_2 (A - \lambda_1 I_n) v_2 =$$
$$\alpha_1(\lambda_1 - \lambda_1) v_1 + \alpha_2(\lambda_2 - \lambda_1) v_2 \Leftrightarrow$$
$$\alpha_2(\lambda_2 - \lambda_1) v_2 = \theta_{\mathbb{K}^n} \Leftrightarrow \alpha_2 = 0,$$

tendo en conta que $\lambda_2 \neq \lambda_1$ e que $v_2 \neq \theta_{\mathbb{K}^n}$. Como consecuencia, $\alpha_1 = 0$ e o sistema é libre. Supoñamos que o resultado é certo para $r-1$ e tratemos de probar o mesmo para r.

Poñamos

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r = \theta_{\mathbb{K}^n}$$

e multipliquemos por $A - \lambda_r I_n$. Por tanto,

$$\theta_{\mathbb{K}^n} = (A - \lambda_r I_n)\theta_{\mathbb{K}^n} = (A - \lambda_r I_n)(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r) =$$
$$\alpha_1(A - \lambda_r I_n) v_1 + \alpha_2 (A - \lambda_r I_n) v_2 + \cdots + \alpha_r (A - \lambda_r I_n) v_r =$$
$$\alpha_1(\lambda_1 - \lambda_r) v_1 + \alpha_2(\lambda_2 - \lambda_r) v_2 + \cdots + \alpha_r(\lambda_r - \lambda_r) v_r =$$
$$\alpha_1(\lambda_1 - \lambda_r) v_1 + \alpha_2(\lambda_2 - \lambda_r) v_2 + \cdots + \alpha_r(\lambda_{r-1} - \lambda_r) v_{r-1}.$$

Como, por hipótese de indución, $\{v_1, \cdots, v_{r-1}\}$ é un sistema libre, tense que $\alpha_i(\lambda_i - \lambda_r) = 0$, $i \in \{1, \cdots, r-1\}$. Agora ben, como $\lambda_i \neq \lambda_r$, $i \in \{1, \cdots, r-1\}$, tense que $\alpha_i = 0$, $i \in \{1, \cdots, r-1\}$. Logo, tendo en conta que v_r é non nulo, o escalar $\alpha_r = 0$ e a demostración queda finalizada. □

Proposición 6.2.4. *Para toda matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ cúmprese que* $\displaystyle\bigcap_{\lambda \in \mathrm{Sp}(A)} V(\lambda) = \{\theta_{\mathbb{K}^n}\}$.

Demostración. A demostración é consecuencia directa da propiedade que afirma que un mesmo autovector non pode estar asociado a dous autovalores distintos. □

Proposición 6.2.5. *Para toda matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$, se $\{\lambda_1, \dots, \lambda_r\}$ son os seus autovalores distintos e B_i é unha base de $V(\lambda_i)$, $i \in \{1, \dots, r\}$, tense que o sistema $B = B_1 \cup \cdots \cup B_r$ é libre.*

Demostración. Nesta demostración tamén procederemos por indución. Sexa

$$B_i = \{v_1^i, v_2^i, \dots, v_{s_i}^i\}$$

unha base de $V(\lambda_i)$, $i \in \{1, \dots, r\}$. Obsérvese que estamos a tomar $s_i = \mathrm{mg}(\lambda_i)$.

Se $r = 1$, o resultado é certo xa que B_1 é unha base de $V(\lambda_1)$. Supoñamos que o resultado é certo para $r-1$ e tratemos de probalo para r. Poñamos

$$\alpha_1^1 v_1^1 + \alpha_2^1 v_2^1 + \cdots + \alpha_{s_1}^1 v_{s_1}^1 + \cdots \cdots + \alpha_1^{r-1} v_1^{r-1} + \alpha_2^{r-1} v_2^{r-1} + \cdots + \alpha_{s_{r-1}}^{r} v_{s_{r-1}}^{r-1} +$$
$$\alpha_1^r v_1^r + \alpha_2^r v_2^r + \cdots + \alpha_{s_r}^r v_{s_r}^r = \theta_{\mathbb{K}^n}$$
$$\Leftrightarrow \sum_{i=1}^{s_1} \alpha_i^1 v_i^1 + \cdots + \sum_{i=1}^{s_{r-1}} \alpha_i^{r-1} v_i^{r-1} + \sum_{i=1}^{s_r} \alpha_i^r v_i^r = \theta_{\mathbb{K}^n}$$

e multipliquemos por $A - \lambda_r I_n$. Por tanto,

$$\theta_{\mathbb{K}^n} = (A - \lambda_r I_n)\theta_{\mathbb{K}^n} = (A - \lambda_r I_n)\left(\sum_{i=1}^{s_1} \alpha_i^1 v_i^1 + \cdots + \sum_{i=1}^{s_{r-1}} \alpha_i^{r-1} v_i^{r-1} + \sum_{i=1}^{s_r} \alpha_i^r v_i^r\right) =$$

$$\sum_{i=1}^{s_1} \alpha_i^1 (A - \lambda_r I_n) v_i^1 + \cdots + \sum_{i=1}^{s_{r-1}} \alpha_i^{r-1} (A - \lambda_r I_n) v_i^{r-1} + \sum_{i=1}^{s_r} \alpha_i^r (A - \lambda_r I_n) v_i^r =$$

$$\sum_{i=1}^{s_1}\alpha_i^1(\lambda_1-\lambda_r)v_i^1+\cdots+\sum_{i=1}^{s_{r-1}}\alpha_i^{r-1}(\lambda_{r-1}-\lambda_r)v_i^{r-1}+\sum_{i=1}^{s_r}\alpha_i^r(\lambda_r-\lambda_r)v_i^r=$$
$$\sum_{i=1}^{s_1}\alpha_i^1(\lambda_1-\lambda_r)v_i^1+\cdots+\sum_{i=1}^{s_{r-1}}\alpha_i^{r-1}(\lambda_{r-1}-\lambda_r)v_i^{r-1}.$$

Como, por hipótese de indución, $B_1\cup\cdots\cup B_{r-1}$ é libre, tense que

$$\begin{cases}\alpha_i^1(\lambda_1-\lambda_r)=0,\ i\in\{1,\ldots,s_1\}\\ \alpha_i^2(\lambda_1-\lambda_r)=0,\ i\in\{1,\ldots,s_2\}\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \alpha_i^{r-1}(\lambda_1-\lambda_r)=0,\ i\in\{1,\ldots,s_{r-1}\}\end{cases}$$

e, ademais, tendo en conta que os autovalores son distintos, obtemos as igualdades

$$\begin{cases}\alpha_i^1=0,\ i\in\{1,\ldots,s_1\}\\ \alpha_i^2=0,\ i\in\{1,\ldots,s_2\}\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \alpha_i^{r-1}=0,\ i\in\{1,\ldots,s_{r-1}\}\end{cases}.$$

Como consecuencia, $\alpha_1^rv_1^r+\alpha_2^rv_2^r+\ \cdots\ +\alpha_{s_r}^rv_{s_r}^r=\theta_{\mathbb{K}^n}$ e, como B_r é un sistema libre, podemos asegurar que

$$\alpha_i^r=0,\ i\in\{1,\ldots,s_r\}.$$

Logo, $B=B_1\cup\cdots\cup B_r$ é un sistema libre. □

Tendo en conta os resultados anteriores podemos enunciar un teorema que dá un criterio de diagonalización en función das multiplicidades alxébricas e xeométricas dos autovalores

Teorema 6.2.6. *Para toda matriz $A\in\mathcal{M}_{n\times n}(\mathbb{K})$ as condicións seguintes son equivalentes.*

1) *O polinomio característico de A descomponse como*
$$p_A(x)=(\lambda_1-x)^{\alpha_1}\cdots(\lambda_r-x)^{\alpha_r},$$
$\mathrm{mg}(\lambda_i)=\alpha_i=\mathrm{ma}(\lambda_i)$, *$i\in\{1,\ldots,r\}$, e $\alpha_1+\cdots+\alpha_r=n$.*
2) *A matriz A é diagonalizábel, isto é, existen dúas matrices P, $D\in\mathcal{M}_{n\times n}(\mathbb{K})$ tales que P é invertíbel, D é diagonal e*
$$P^{-1}AP=D.$$

Demostración. Supoñamos que A é diagonalizábel e

$$D=\begin{pmatrix}\lambda_1&0&\cdots&0\\0&\lambda_2&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&\lambda_n\end{pmatrix}$$

a matriz diagonal á que é semellante. Daquela

$$p_A(x)=p_D(x)=(\lambda_1-x)(\lambda_2-x)\cdots(\lambda_n-x).$$

Se no anterior polinomio agrupamos os factores que se repiten, obtemos que

$$p_A(x)=(\lambda_1-x)^{\alpha_1}(\lambda_2-x)^{\alpha_2}\cdots(\lambda_r-x)^{\alpha_r}$$

con $\alpha_1+\cdots+\alpha_r=n$.

Ademais, se P é a matriz tal que $P^{-1}AP=D$, para un autovalor λ_i existen α_i columnas de P, $\{P_i^1,\ldots,P_i^{\alpha_i}\}$ tales que $AP_i^1=\lambda_iP_i^1,\ldots,AP_i^{\alpha_i}=\lambda_iP_i^{\alpha_i}$. Como está formado por columnas dunha matriz invertíbel, o sistema $\{P_i^1,\ldots,P_i^{\alpha_i}\}$ é libre e logo $\alpha_i\geq\mathrm{mg}(\lambda_i)\geq\alpha_i$. Por tanto, $\mathrm{mg}(\lambda_i)=\mathrm{ma}(\lambda_i)$.

Doutra banda, supoñamos que o polinomio característico de A descomponse como

$$p_A(x) = (\lambda_1 - x)^{\alpha_1}(\lambda_2 - x)^{\alpha_2} \cdots (\lambda_r - x)^{\alpha_r}$$

$\mathrm{mg}(\lambda_i) = \alpha_i = \mathrm{ma}(\lambda_i)$, $i \in \{1, \dots, r\}$, e $\alpha_1 + \cdots + \alpha_r = n$. Se temos que B_i é unha base de $V(\lambda_i)$, o sistema $B = B_1 \cup \cdots \cup B_r$ é libre e está formado por

$$\mathrm{mg}(\lambda_1) + \cdots + \mathrm{mg}(\lambda_r) = \mathrm{ma}(\lambda_1) + \cdots + \mathrm{ma}(\lambda_r) = n$$

vectores. Por tanto, é unha base de $\mathbb{K}^n$ formada por autovectores e, polo visto ao principio da sección, A é diagonalizábel. □

Corolario 6.2.7. *Toda matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ con n autovalores distintos é diagonalizábel.*

Demostración. Se A ten n autovalores distintos $\lambda_1, \dots, \lambda_n$, o polinomio característico é $p_A(x) = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x)$ e $\alpha_i = 1$ para todo i. Entón, $\mathrm{mg}(\lambda_i) = 1 = \mathrm{ma}(\lambda_i)$ e, aplicando o resultado anterior, podemos afirmar que a matriz A é diagonalizábel. □

Exemplos 6.2.8. 1) Dada a matriz

$$A = \begin{pmatrix} 0 & -1 & -2 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix},$$

considerada no primeiro caso de Exemplos 6.1.5, tense que non é diagonalizábel como matriz real xa que $p_A(x) = (x+1)(x-(1-i))(x-(1+i))$. Como matriz complexa o espectro de A ven dado por

$$\mathrm{Sp}(A) = \{-1, 1-i, 1+i\}$$

e, ademais,

$$\mathrm{ma}(-1) = \mathrm{ma}a(1-i) = \mathrm{ma}(1+i) = 1.$$

Por tanto, A é diagonalizábel como matriz de coeficientes complexos. A matriz P tal que

$$P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1-i & 0 \\ 0 & 0 & 1+i \end{pmatrix}$$

ven dada por

$$P = \begin{pmatrix} 0 & -1-2i & -1+2i \\ -2 & 1+i & 1-i \\ 1 & 1 & 1 \end{pmatrix}$$

onde

$$V(-1) = \langle \left\{ \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\} \rangle, \; V(1-i) = \langle \left\{ \begin{pmatrix} -1-2i \\ 1+i \\ 1 \end{pmatrix} \right\} \rangle,$$

$$V(1+i) = \langle \left\{ \begin{pmatrix} -1+2i \\ 1-i \\ 1 \end{pmatrix} \right\} \rangle.$$

2) Sexa a matriz

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

O seu polinomio característico é

$$p_A(x) = \det(A - xI_4) =$$

$$\begin{vmatrix} 2-x & 0 & 0 & 0 \\ 0 & 1-x & 1 & -1 \\ 0 & -1 & 2-x & -1 \\ 0 & 0 & -1 & 2-x \end{vmatrix} = (2-x) \begin{vmatrix} 1-x & 1 & -1 \\ -1 & 2-x & -1 \\ 0 & -1 & 2-x \end{vmatrix} =$$

$$(2-x) \begin{vmatrix} 2-x & 1 & -1 \\ 0 & 2-x & -1 \\ x-2 & -1 & 2-x \end{vmatrix} = (2-x) \begin{vmatrix} 2-x & 1 & -1 \\ 0 & 2-x & -1 \\ 0 & 0 & 1-x \end{vmatrix} = (2-x)^3(1-x).$$

Por tanto, podemos afirmar que

$$\mathrm{Sp}(A) = \{1, 2\}$$

e que $\mathrm{ma}(1) = 1$, $\mathrm{ma}(2) = 3$.

Neste caso tense que

$$\mathrm{mg}(2) = \dim_{\mathbb{R}}(V(2)) = 4 - \mathrm{rang}\,(A - 2I_4) = 4 - \mathrm{rang} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix} =$$

$$4 - 2 = 2 < 3 = \mathrm{ma}(2)$$

e, en consecuencia, A non é diagonalizábel (nin como matriz real, nin como matriz complexa).

3) Tomemos a matriz

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

O seu polinomio característico é

$$p_A(x) = \det(A - xI_4) =$$

$$\begin{vmatrix} 1-x & 1 & 1 & 1 \\ 1 & 1-x & 1 & 1 \\ 1 & 1 & 1-x & 1 \\ 1 & 1 & 1 & 1-x \end{vmatrix} = \begin{vmatrix} 4-x & 1 & 1 & 1 \\ 4-x & 1-x & 1 & 1 \\ 4-x & 1 & 1-x & 1 \\ 4-x & 1 & 1 & 1-x \end{vmatrix} =$$

$$\begin{vmatrix} 4-x & 1 & 1 & 1 \\ 0 & -x & 0 & 0 \\ 0 & 0 & -x & 0 \\ 0 & 0 & 0 & -x \end{vmatrix} = -(4-x)x^3.$$

Por tanto, $\mathrm{Sp}(A) = \{0, 4\}$ e $\mathrm{ma}(4) = 1$, $\mathrm{ma}(0) = 3$. Neste caso tense que

$$\mathrm{mg}(0) = \dim_{\mathbb{R}}(V(0)) = 4 - \mathrm{rang}\,(A - 0I_4) = 4 - \mathrm{rang}\,(A) = 4 - 1 = 3 = \mathrm{ma}(0)$$

e, como $\mathrm{mg}(4) = \mathrm{ma}(4) = 1$, A é diagonalizábel como matriz real. Isto quere dicir que existe unha matriz invertíbel $P \in \mathcal{M}_{4\times 4}(\mathbb{R})$ tal que

$$P^{-1}AP = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

A matriz P obtense buscando unha base de $\mathbb{R}^4$ formada por autovectores de A. Como

$$V(0) = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \;/\; (A - 0I_4) \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \;/\; \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^3 \;\mid\; x + y + z + t = 0 \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \\ -x-y-z \end{pmatrix} \in \mathbb{R}^4 \;\mid\; x, y, z \in \mathbb{R} \right\} = \langle \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right\} \rangle$$

e

$$V(4) = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \;\mid\; (A - 4I_4) \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \;\mid\; \begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \;\mid\; x = y = z = t \right\}$$

$$= \left\{ \begin{pmatrix} x \\ x \\ x \\ x \end{pmatrix} \in \mathbb{R}^4 \;\mid\; x \in \mathbb{R} \right\} = \langle \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\} \rangle,$$

a matriz invertíbel P está dada por:

$$P = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{pmatrix}.$$

Nota 6.2.9. Supoñamos que $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ é diagonalizábel e que

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

é a matriz diagonal á que é semellante. Logo, se P é a matriz invertíbel tal que $P^{-1}AP = D$, tense que $A = PDP^{-1}$ e

$$A^n = PDP^{-1}PDP^{-1} \overset{n-2)}{\cdots} PDP^{-1}PDP^{-1} = PD^nP^{-1}$$

onde

$$D^n = \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^n \end{pmatrix}.$$

Entón, $\mathrm{Sp}(A^n) = \{\lambda_1^n, \lambda_2^n, \dots, \lambda_r^n\}$ e $V(\lambda_i) = V(\lambda_i^n)$ para $i = \{1, \dots, r\}$.

Doutra banda, é importante salientar que, se A é diagonalizábel e invertíbel, a súa inversa tamén é diagonalizábel xa que

$$A^{-1} = PD^{-1}P^{-1}$$

sendo

$$D^{-1} = \begin{pmatrix} \lambda_1^{-1} & 0 & \cdots & 0 \\ 0 & \lambda_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^{-1} \end{pmatrix}.$$

6.3. Polinomios anuladores. Teorema de Cayley-Hamilton

Definición 6.3.1. Sexa $p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n$ un polinomio de grao n con coeficientes en $\mathbb{K}$. Dada $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ defínese a matriz $p(A)$ como:

$$p(A) = a_0I_n + a_1A + \cdots + a_{n-1}A^{n-1} + a_nA^n.$$

Supoñamos que $\lambda \in \mathrm{Sp}(A)$ e que $v \in \mathbb{K}^n$ é un autovector asociado a λ. Tense que

$$A^rv = A^{r-1}\lambda v = \lambda A^{r-1}v = \cdots = \lambda^r v$$

e, como consecuencia,

$$p(A)v = (a_0I_n + a_1A + \cdots + a_{n-1}A^{n-1} + a_nA^n)v = a_0I_nv + a_1Av + \cdots + a_{n-1}A^{n-1}v + a_nA^nv =$$

$$a_0v + a_1\lambda v + \cdots + a_{n-1}\lambda^{n-1}v + a_n\lambda^nv = (a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + a_n\lambda^n)v = p(\lambda)v.$$

Logo, se $\lambda \in \mathrm{Sp}(A)$, obtemos que $p(\lambda) \in \mathrm{Sp}(p(A))$.

Nota 6.3.2. Supoñamos que $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ é diagonalizábel e que

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

é a matriz diagonal á que é semellante. Sexa P a matriz invertíbel tal que $P^{-1}AP = D$ ou, equivalentemente, tal que $A = PDP^{-1}$. Polo visto no apartado de diagonalización sabemos que

$$A^r = PD^rP^{-1}$$

onde

$$D^r = \begin{pmatrix} \lambda_1^r & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^r \end{pmatrix}.$$

Sexa $p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n$ un polinomio de grao n con coeficientes en $\mathbb{K}$. Por tanto,

$$p(A) = a_0I_n + a_1A + \cdots + a_{n-1}A^{n-1} + a_nA^n =$$

$$a_0I_n + a_1PDP^{-1} + \cdots + a_{n-1}PD^{n-1}P^{-1} + a_nPD^nP^{-1} =$$

$$a_0PP^{-1} + a_1PDP^{-1} + \cdots + a_{n-1}PD^{n-1}P^{-1} + a_nPD^nP^{-1} =$$

$$P(a_0I_n + a_1D + \cdots + a_{n-1}D^{n-1} + a_nD^n)P^{-1} = Pp(D)P^{-1} =$$

$$P\begin{pmatrix} p(\lambda_1) & 0 & \cdots & 0 \\ 0 & p(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p(\lambda_n) \end{pmatrix}P^{-1}.$$

Deste xeito obtemos que $p(A)$ tamén é diagonalizábel e que $V(\lambda_i) = V(p(\lambda_i))$ para cada λ_i no espectro da matriz A.

Definición 6.3.3. Dada $A \in \mathcal{M}_{n\times n}(\mathbb{K})$, dise que un polinomio

$$p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n,$$

con coeficientes en $\mathbb{K}$, é un polinomio anulador de A, se

$$p(A) = a_0I_n + a_1A + \cdots + a_{n-1}A^{n-1} + a_nA^n = \Theta_{n,n}.$$

Anteriormente demostramos que, se $\lambda \in \mathrm{Sp}(A)$, tense que $p(\lambda) \in \mathrm{Sp}(p(A))$ para todo polinomio $p(x)$ con coeficientes en $\mathbb{K}$. Entón, se $p(x)$ é un polinomio anulador, tense que $p(\lambda) = 0$ ou, o que é o mesmo, λ é raíz de calquera polinomio anulador de A.

Nota 6.3.4. Supoñamos que $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ e que

$$p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n$$

é un polinomio anulador de A. Logo,

$$p(A) = a_0I_n + a_1A + \cdots + a_{n-1}A^{n-1} + a_nA^n = \Theta_{n,n} \Leftrightarrow$$

$$a_0I_n + (a_1I_n + \cdots + a_{n-1}A^{n-2} + a_nA^{n-1})A = \Theta_{n,n}.$$

Se $a_0 \neq 0$, a última igualdade pódese escribir como:

$$-a_0I_n = (a_1I_n + \cdots + a_{n-1}A^{n-2} + a_nA^{n-1})A \Leftrightarrow$$

$$\frac{1}{-a_0}(a_1 I_n + \cdots + a_{n-1}A^{n-2} + a_n A^{n-1})A = I_n.$$

Isto é equivalente a dicir que, se $a_0 \neq 0$, a matriz A é invertíbel e a súa inversa está dada por

$$A^{-1} = \frac{1}{-a_0}(a_1 I_n + \cdots + a_{n-1}A^{n-2} + a_n A^{n-1}).$$

Nota 6.3.5. Supoñamos que A e $B \in \mathcal{M}_{n\times n}(\mathbb{K})$ son matrices semellantes e que $P \in \mathcal{M}_{n\times n}(\mathbb{K})$ é a matriz invertíbel tal que $B = P^{-1}AP$. Entón, para todo polinomio

$$p(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + a_n x^n$$

de grao n con coeficientes en $\mathbb{K}$ tense que

$$p(A) = a_0 I_n + a_1 A + \cdots + a_{n-1}A^{n-1} + a_n A^n =$$

$$a_0 I_n + a_1 PBP^{-1} + \cdots + a_{n-1}PB^{n-1}P^{-1} + a_n PB^n P^{-1} =$$

$$a_0 PP^{-1} + a_1 PBP^{-1} + \cdots + a_{n-1}PB^{n-1}P^{-1} + a_n PB^n P^{-1} =$$

$$P(a_0 I_n + a_1 B + \cdots + a_{n-1}B^{n-1} + a_n B^n)P^{-1} = Pp(B)P^{-1}.$$

Por tanto, $p(A)$ e $p(B)$ son semellantes e, como consecuencia, $p(x)$ é un polinomio anulador para A, se, e só se, o é para B.

O seguinte teorema indícanos como obter polinomios anuladores.

Teorema 6.3.6. (Teorema de Cayley-Hamilton) *Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ e supoñamos que o polinomio característico de A descomponse como*

$$p_A(x) = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x),$$

onde os $\lambda_i \in \mathbb{K}$, $i \in \{1, \ldots, n\}$ son escalares non necesariamente distintos. Baixo estas condicións tense que o polinomio característico é un polinomio anulador para A.

Demostración. Grazas ao Teorema 6.1.14 e a nota 6.3.5 é suficiente probar o resultado para matrices triangulares superiores. A demostración farémola por indución na orde de A. Para $n = 1$ a propiedade é evidente. Supoñamos que o resultado é certo para $n-1$ e tratemos de probalo para n.

Sexa A unha matriz cadrada triangular superior de orde n que escribiremos como

$$A = \left(\begin{array}{c|c} \lambda_1 & u \\ - & --- \\ 0 & \\ \vdots & A_1 \\ 0 & \end{array}\right),$$

onde A_1 é unha matriz cadrada triangular superior de orde $n-1$. Para a matriz A tense que

$$p_A(x) = (\lambda_1 - x)p_{A_1}(x).$$

Logo,

$$p_A(A) = (\lambda_1 I_n - A)p_{A_1}(A) =$$

$$\left(\begin{array}{c|c} 0 & -u \\ - & --- \\ 0 & \\ \vdots & \lambda_1 I_{n-1} - A_1 \\ 0 & \end{array}\right)\left(\begin{array}{c|c} p_{A_1}(\lambda_1) & u' \\ - & --- \\ 0 & \\ \vdots & p_{A_1}(A_1) \\ 0 & \end{array}\right).$$

Como por hipótese de indución $p_{A_1}(A_1) = \Theta_{n-1,n-1}$, obtemos

$$p_A(A) = \left(\begin{array}{c|c} 0 & -u \\ - & --- \\ 0 & \\ \vdots & \lambda_1 I_{n-1} - A_1 \\ 0 & \end{array}\right)\left(\begin{array}{c|c} p_{A_1}(\lambda_1) & u' \\ - & --- \\ 0 & \\ \vdots & \Theta_{n-1,n-1} \\ 0 & \end{array}\right) = \Theta_{n,n},$$

co que a proba queda finalizada. □

Como consecuencia do Teorema Fundamental da Álxebra tense o seguinte corolario.

Corolario 6.3.7. *Para toda matriz $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ o seu polinomio característico é un polinomio anulador.*

Exemplo 6.3.8. Dada a matriz

$$A = \begin{pmatrix} 1 & -2 & 1 \\ 0 & -1 & 2 \\ 2 & 0 & 3 \end{pmatrix},$$

temos que

$$p_A(x) = \begin{vmatrix} 1-x & -2 & 1 \\ 0 & -1-x & 2 \\ 2 & 0 & 3-x \end{vmatrix} = -x^3 - 3x^2 + 9x - 3.$$

Como o coeficiente de grao cero de $p_A(x)$ é non nulo, obtemos que a matriz é invertíbel. A súa inversa pódese calcular, utilizando que $p_A(x)$ é un polinomio anulador para A, da seguinte forma:

$$\Theta_{3,3} = p_A(A) = -A^3 - 3A^2 + 9A - 3I_3 \Leftrightarrow$$

$$-A^3 - 3A^2 + 9A = 3I_3 \Leftrightarrow$$

$$-\frac{1}{3}(A^2 - 3A + 9I_3)A = I_3 \Leftrightarrow A^{-1} = -\frac{1}{3}(A^2 - 3A + 9I_3) = -\frac{1}{3}A^2 + A - 3I_3.$$

6.4. Funcións de matrices. Matriz exponencial dunha matriz cadrada

Definición 6.4.1. Sexan $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ e $f : I \subset \mathbb{R} \to \mathbb{R}$ unha función. Supoñamos que $\lambda_1, \cdots, \lambda_r$ son os seus autovalores distintos e que

$$p_A(x) = (\lambda_1 - x)^{\alpha_1}(\lambda_2 - x)^{\alpha_2} \cdots (\lambda_r - x)^{\alpha_r},$$

onde $\mathrm{ma}(\lambda_i) = \alpha_i$. Diremos que f está definida sobre o espectro de A, se para cada λ_i existen

$$f(\lambda_i),\ f'(\lambda_i),\ \dots\ , f^{(\alpha_i - 1)}(\lambda_i).$$

Se f está definida sobre o espectro de A, con

$$V_{f,A} = \{f^k(\lambda_i),\ |\ i \in \{1, \dots, r\},\ k \in \{1, \dots, \alpha_i - 1\}\}$$

denotaremos o conxunto dos valores de f sobre o espectro de A.

Nota 6.4.2. Supoñamos que nos atopamos nas condicións da definición anterior. Sempre é posíbel atopar un único polinomio, chamado polinomio interpolador de Lagrange–Sylvester,

$$r(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1},$$

de grao menor que n tal que

$$V_{f,A} = V_{r,A}.$$

Para atopar $r(x)$ non temos máis que resolver o sistema de ecuacións lineares

$$f^k(\lambda_i) = r^k(\lambda_i), \;\; i \in \{1, \dots, r\}, \; k \in \{1, \dots, \alpha_i - 1\}$$

cuxas incógnitas serán os coeficientes a_k do polinomio $r(x)$.

Definición 6.4.3. Sexan $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ e $f : I \subset \mathbb{R} \to \mathbb{R}$ unha función. Supoñamos que $\lambda_1, \cdots, \lambda_r$ son os seus autovalores distintos e que

$$p_A(x) = (\lambda_1 - x)^{\alpha_1}(\lambda_2 - x)^{\alpha_2} \cdots (\lambda_r - x)^{\alpha_r},$$

onde $\mathrm{ma}(\lambda_i) = \alpha_i$. Se f está definida sobre o espectro de A e $r(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$ é o polinomio interpolador de Lagrange–Sylvester, defínese $f(A)$ como

$$f(A) = a_0 I_n + a_1 A + \cdots + a_{n-1}A^{n-1}.$$

Nota 6.4.4. De todas as funcións posíbeis que podemos definir sobre unha matriz A, é particularmente importante a exponencial de A debido a que na teoría de ecuacións diferenciais lineares é necesario calcular e^{tA}, sendo t un parámetro real. Se definimos a función $f : \mathbb{R} \to \mathbb{R}$ como $f(x) = e^{tx}$, observamos que f e todas as súas derivadas sucesivas están definidas para calquera número real e grazas a isto podemos definir $f(A) = e^{tA}$ para toda $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que

$$p_A(x) = (\lambda_1 - x)^{\alpha_1}(\lambda_2 - x)^{\alpha_2} \cdots (\lambda_r - x)^{\alpha_r},$$

con $\lambda_i \in \mathbb{R}$, $i \in \{1, \dots, r\}$, e $\mathrm{ma}(\lambda_i) = \alpha_i$.

É interesante salientar que $e^{\Theta_{n,n}} = I_n$ e que a matriz e^{tA} é sempre invertíbel sendo a súa inversa $(e^{tA})^{-1} = e^{-tA}$.

O devandito para a función exponencial tamén se ten para outras funcións como $f(x) = \mathrm{sen}(x)$ ou $f(x) = \cos(x)$ xa que elas e todas as súas derivadas sucesivas están definidas para calquera número real. Por tanto, sempre poderemos calcular $f(A) = \mathrm{sen}(A)$ ou $f(A) = \cos(A)$, etc.

A exponencial dunha matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tamén se pode calcular no caso de que aparezan autovalores complexos aplicando o mesmo método dado para os autovalores reais e tendo en conta que

$$e^{a+ib} = e^a e^{ib} = e^a(\cos(b) + i\mathrm{sen}(b)).$$

Exemplos 6.4.5. 1) Sexa a matriz

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{pmatrix}.$$

O polinomio característico da matriz A está dado por

$$p_A(x) = \det(A - xI_3) = \begin{vmatrix} -x & 1 & 0 \\ 0 & -x & 1 \\ 1 & -3 & 3-x \end{vmatrix} = \begin{vmatrix} 1-x & 1 & 0 \\ 1-x & -x & 1 \\ 1-x & -3 & 3-x \end{vmatrix} =$$

$$\begin{vmatrix} 1-x & 1 & 0 \\ 0 & -x-1 & 1 \\ 0 & -4 & 3-x \end{vmatrix} = (1-x)\begin{vmatrix} -x-1 & 1 \\ -4 & 3-x \end{vmatrix} = (1-x)\begin{vmatrix} -x-1 & 1 \\ -4 & 3-x \end{vmatrix} =$$

$$(1-x)((x-3)(1+x)+4) = (1-x)^3.$$

Por tanto, se tomamos a función $f(x) = e^{tx}$, sabemos que existe a matriz $f(A) = e^{tA}$. Para calculala vexamos primeiro que valores contén $V_{f,A}$. Neste caso tense que

$$V_{f,A} = \{f(1) = e^t, f'(1) = te^t, f''(1) = t^2e^t\}$$

e, se o polinomio interpolador ten a forma $r(x) = a_0 + a_1x + a_2x^2$, resolvendo o sistema

$$\begin{cases} f(1) = e^t = r(1) = a_0 + a_1 + a_2 \\ f'(1) = te^t = r'(1) = a_1 + 2a_2 \\ f''(1) = t^2e^t = r''(1) = 2a_2 \end{cases}$$

obtemos que

$$a_0 = (1 - t - \frac{t^2}{2})e^t,\ a_1 = (t - t^2)e^t,\ a_2 = \frac{t^2}{2}e^t.$$

Logo,

$$e^{tA} = (1 - t - \frac{t^2}{2})e^tI_3 + (t - t^2)e^tA + \frac{t^2}{2}e^tA^2.$$

2) Tomemos a matriz

$$A = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 2 & 3 & 5 \\ -1 & -2 & -3 & -4 \end{pmatrix}$$

e a función $f(x) = \ln(x)$. Vexamos se existe $\ln(A)$. O polinomio característico da matriz A esta dado por

$$p_A(x) = \det(A - xI_3) = \begin{vmatrix} 1-x & 2 & 1 & 1 \\ 0 & -x & 1 & 0 \\ 1 & 2 & 3-x & 5 \\ -1 & -2 & -3 & -4-x \end{vmatrix} = \begin{vmatrix} 1-x & 2 & 1 & 1 \\ 0 & -x & 1 & 0 \\ 1 & 2 & 3-x & 5 \\ 0 & 0 & -x & 1-x \end{vmatrix} =$$

$$\begin{vmatrix} 1-x & 2x & 1 & 1 \\ 0 & -x & 1 & 0 \\ 1 & 0 & 3-x & 5 \\ 0 & 0 & -x & 1-x \end{vmatrix} = \begin{vmatrix} 1-x & 0 & 3 & 1 \\ 0 & -x & 1 & 0 \\ 1 & 0 & 3-x & 5 \\ 0 & 0 & -x & 1-x \end{vmatrix} = -x \begin{vmatrix} 1-x & 3 & 1 \\ 1 & 3-x & 5 \\ 0 & -x & 1-x \end{vmatrix} =$$

$$-x((1-x)^2(3-x) - x + 5x(1-x) - 3(1-x)) = -x(1-x)((1-x)(3-x) + 5x - 3) - x =$$

$$-x((1-x)(x^2+x) - x) = x^4$$

e, por tanto, $\mathrm{Sp}(A) = \{0\}$ sendo $\mathrm{ma}(0) = 4$. Como $f(0) = \ln(0)$ non existe, a función non está definida sobre o espectro de A e como consecuencia non podemos calcular $\ln(A)$.

Para a función $f(x) = e^{tx}$ temos que

$$V_{f,A} = \{f(0) = 1, f'(0) = t, f''(0) = t^2,\ f'''(0) = t^3\}$$

e, se o polinomio interpolador ten a forma $r(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, resolvendo o sistema

$$\begin{cases} f(0) = 1 = r(0) = a_0 \\ f'(0) = t = r'(0) = a_1 \\ f''(0) = t^2 = r''(0) = 2a_2 \\ f'''(0) = t^3 = r'''(0) = 6a_3 \end{cases}$$

obtemos que

$$a_0 = 1,\ a_1 = t,\ a_2 = \frac{t^2}{2},\ a_3 = \frac{t^3}{6}.$$

Deste xeito, $r(x) = 1 + tx + \frac{t^2}{2}x^2 + \frac{t^3}{6}x^3$ e logo

$$e^{tA} = I_4 + tA + \frac{t^2}{2}A^2 + \frac{t^3}{6}A^3.$$

3) Sexan a matriz

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

e a función $f(x) = e^{tx}$. Vexamos se existe e^{tA}. O polinomio característico da matriz A está dado por

$$p_A(x) = \det(A - xI_3) = \begin{vmatrix} 1-x & 2 \\ -2 & 1-x \end{vmatrix} = (1-x)^2 + 4 = x^2 - 2x + 5$$

e, por tanto, $\mathrm{Sp}(A) = \{1+2i, 1-2i\}$, sendo $\mathrm{ma}(1+2i) = \mathrm{ma}(1-2i) = 1$. Neste caso temos raíces complexas polo que, se

$$r(x) = a_0 + a_1 x,$$

debemos considerar o sistema

$$\begin{cases} f(1+2i) = e^{t(1+2i)} = e^t((\cos(2t) + i\mathrm{sen}(2t)) = r(1+2i) = a_0 + a_1(1+2i) \\ f(1-2i) = e^{t(1-2i)} = e^t((\cos(2t) - i\mathrm{sen}(2t)) = r(1-2i) = a_0 + a_1(1-2i) \end{cases} \Leftrightarrow$$

$$\begin{cases} a_0 + a_1(1+2i) = e^t((\cos(2t) + i\mathrm{sen}(2t)) \\ a_0 + a_1(1-2i) = e^t((\cos(2t) - i\mathrm{sen}(2t)) \end{cases} \Leftrightarrow \begin{cases} a_0 + a_1 = e^t\cos(2t) \\ 2a_1 = e^t\mathrm{sen}(2t) \end{cases} \Leftrightarrow$$

$$a_0 = e^t\cos(2t) - \frac{1}{2}e^t\mathrm{sen}(2t),\ \ a_1 = \frac{1}{2}e^t\mathrm{sen}(2t).$$

Entón,

$$r(x) = e^t\cos(2t) - \frac{1}{2}e^t\mathrm{sen}(2t) + \frac{1}{2}e^t\mathrm{sen}(2t)x$$

e

$$e^{tA} = (e^t\cos(2t) - \frac{1}{2}e^t\mathrm{sen}(2t))I_2 + \frac{1}{2}e^t\mathrm{sen}(2t)A = \begin{pmatrix} e^t\cos(2t) & e^t\mathrm{sen}(2t) \\ -e^t\mathrm{sen}(2t) & e^t\cos(2t) \end{pmatrix}.$$

Nota 6.4.6. Se

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

é unha matriz diagonal e f é unha función definida sobre o espectro de A, tense que

$$f(D) = \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix}.$$

Como consecuencia, se A é diagonalizábel e P é a matriz invertíbel tal que $P^{-1}AP = D$ ou, equivalentemente, que $A = PDP^{-1}$, obtemos a igualdade

$$f(A) = Pf(D)P^{-1} = P \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix} P^{-1}.$$

Isto implica que, cando a matriz é diagonalizábel, o cálculo de funcións de matrices é particularmente sinxelo.

6.5. Problemas propostos

1. Ache os autovalores e os correspondentes subespazos propios das seguintes matrices. Estude ademais se son diagonalizábeis.

   a)
   $$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 2 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

   b)
   $$B = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -8 & 3 & -4 & 4 \\ 8 & -4 & 3 & -4 \\ 20 & -10 & 10 & -11 \end{pmatrix}.$$

   c)
   $$C = \begin{pmatrix} -3 & -5 \\ 2 & 3 \end{pmatrix}.$$

2. Considérase a matriz
   $$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

   a) Probe que A é diagonalizábel.
   b) Atope unha matriz diagonal D e unha matriz invertíbel P tales que $A = PDP^{-1}$.
   c) Calcule a expresión de A^n para cada número natural n.
   d) Encontre unha raíz cadrada de A.

3. Estude para que valores de $\alpha \in \mathbb{R}$ é diagonalizábel a matriz
   $$A = \begin{pmatrix} 1 & 0 & 2+\alpha \\ 0 & -1 & 4+\alpha \\ 2+\alpha & 0 & 1 \end{pmatrix}.$$

4. Para cada $\alpha \in \mathbb{R}$ considérase a matriz
   $$A = \begin{pmatrix} 0 & 0 & \alpha & \alpha \\ 0 & 0 & \alpha & \alpha \\ 4\alpha & 4\alpha & 0 & 0 \\ 4\alpha & 4\alpha & 0 & 0 \end{pmatrix}.$$

   a) Estude para que valores de α a matriz A é diagonalizábel.
   b) Para $\alpha = 1/4$, determine unha matriz regular P e unha matriz diagonal D tales que $AP = PD$

5. Considérase a matriz

$$A = \begin{pmatrix} -1 & 2 & 0 & -2 \\ 0 & 4 & -2 & -3 \\ 0 & 2 & -1 & -2 \\ 0 & 3 & -2 & -2 \end{pmatrix}.$$

   a) Determine o espectro de A.
   b) Estude se A é diagonalizábel.
   c) Probe que A é invertíbel e determine un polinomio $p(x)$ de grao menor que 4 tal que $p(A) = A^{-1}$.

6. Considéranse as seguintes matrices:

$$A_1 = \begin{pmatrix} -7 & 0 & 0 \\ 0 & 4 & -12 \\ 0 & 11 & 7 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 0 & -2 \\ 1 & 0 & 1 \\ -2 & 0 & 2 \end{pmatrix},$$
$$A_3 = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1 & 1 & 1 \\ 1/2 & 0 & 1/2 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

   Sexa f a función definida por

$$f(x) = \frac{1+x}{2-x} + 2\sqrt{x}.$$

   Estude se existen os valores de f sobre cada unha das matrices dadas. Nos casos afirmativos, calcule $f(A_i)$.

7. Considérase a matriz

$$A = \begin{pmatrix} 2 & 1 & 0 \\ -2 & 0 & 2 \\ 0 & -1 & -2 \end{pmatrix}.$$

   a) Calcule a matriz e^{tA}.
   b) Sexa $p(x) = 2 + 3x^3 + 4x^9$. Estude se $p(A)$ é invertíbel.

8. Considérase a matriz real

$$A = \begin{pmatrix} \beta+1 & -1 & 1 \\ \alpha & \beta+1 & 0 \\ 1 & -\alpha & \alpha+\beta+1 \end{pmatrix}.$$

   a) Calcule o espectro de A.
   b) Determine os valores de α e β para os cales a matriz A é singular.
   c) Estude para que valores de α e β existe unha base de $\mathbb{R}^3$ formada por autovectores de A.
   d) Para $\alpha = 0$,
      d1) Atope, se é posíbel, unha matriz regular P e unha matriz diagonal D tales que $AP = PD$.
      d2) Calcule o determinante da matriz $B = 3\,[A - (\beta+1)I_3]^n$, onde $n \in \mathbb{N}$.
   e) Para $\alpha = 1$ e $\beta = -1$, calcule a matriz e^{tA}.
   f) Para $\alpha = 3$ e $\beta = 0$, determine números reais a, b e c tales que

$$A^3 + a\,A^2 + b\,A + c\,I_3 = 0$$

   e calcule, se é posíbel, un polinomio $p(x)$ tal que $A^{-1} = p(A)$.

9. Sexa

$$A = \begin{pmatrix} -3 & -2 & 1 \\ 5 & 3 & -2 \\ -1 & -1 & 0 \end{pmatrix}.$$

a) Ache o espectro de A. É a matriz $I_3 - A$ invertíbel?
b) Estude se a matriz A é diagonalizábel.
c) Sexa $p(x) = a + bx + cx^2 + dx^3$, onde $a, b, c, d \in \mathbb{R}$. Calcule o determinante da matriz $p(A)$.
d) Determine a matriz e^{tA}.

10. Considérase a matriz real

$$A = \begin{pmatrix} -1 & -1 & -2 \\ 8 & -11 & -8 \\ -10 & 11 & 7 \end{pmatrix}.$$

a) Atope o espectro de A.
b) Probe que $A^3 = -5\,A^2 - 3\,A + 9\,I$.
c) Calcule unha base do subespazo propio asociado ao maior autovalor de A.
d) Estude se a matriz A é diagonalizábel.
e) Considérase a función escalar $f(x) = \sqrt{x}$. Determine, se é posíbel, un polinomio $r(x)$ tal que $r(A) = f(A)$.

CAPíTULO 7

Espazos vectoriais con produto escalar

Neste capítulo comeza unha nova parte relacionada con certos conceptos ben coñecidos en $\mathbb{R}^2$ e $\mathbb{R}^3$, tales como norma, perpendicularidade e ángulo. Algunhas destas nocións están na base das ferramentas que nos permitirán, entre outras cousas, abordar a aproximación de solucións de sistemas incompatíbeis mediante técnicas de mínimos cadrados no último capítulo deste libro.

Comezaremos o capítulo introducindo a noción de forma bilinear

$$f : V \times V \to \mathbb{R}$$

nun $\mathbb{R}$-espazo vectorial V, para pasar a ver que, unha vez fixada unha base B en V, a forma bilinear pódese expresar en función dunha matriz chamada matriz de Gram, denotada por G_B, da seguinte forma

$$f(u, v) = u_B^t G_B v_B,$$

onde u_B e v_B son os vectores de coordenadas de u e v respecto á base B. A continuación mostrarase ao lector o que ocorre coa matriz de Gram cando cambiamos a base do espazo vectorial onde temos definida a forma bilinear. Máis concretamente, comprobarase que, dadas B e B', dúas bases de V, tense que

$$G_{B'} = P_{B',B}^t G_B P_{B',B}$$

sendo $P_{B',B}$ a matriz de cambio de base de B' a B. Logo, as matrices de Gram dunha forma bilinear respecto a diferentes bases son congruentes.

Os produtos escalares non son máis que tipos especiais de formas bilineares que cumpren a condición de ser simétricas, isto é, $f(u, v) = f(v, u) \ \forall \ u, v \in V$ e ademais definidas positivas: $f(u, u) > 0 \ \forall \ u \in V, u \neq \theta_V$. A notación clásica que se utiliza para o produto escalar é

$$f(u, v) =< u, v >$$

e neste libro é a que usaremos. Tamén con $(V, < \ , \ >)$ denotaremos ao par dado polo $\mathbb{R}$-espazo vectorial V xunto co produto escalar definido sobre V e chamarémolo espazo vectorial euclídeo. É importante recalcar que, se nun mesmo $\mathbb{R}$-espazo vectorial V temos definidos varios produtos escalares, obteremos diferentes exemplos de espazos vectoriais euclídeos. Doutra banda a matriz de Gram asociada a un produto escalar será obviamente simétrica.

Unha vez aclarada a noción de produto escalar introduciranse as de norma e norma inducida por un produto escalar. Neste punto probarase a desigualdade de Schwartz e indicarase como a partir dela pódese dar unha definición de ángulo entre dous vectores de V. Finalmente, tamén comprobaremos que a existencia dunha norma permitiranos introducir unha distancia en V.

Na segunda parte do capítulo traballarase coas nocións de ortogonalidade e ortonormalidade respecto a un produto escalar. Veranse entre outras cousas que, se $(V, < \ , \ >)$ é un espazo vectorial euclídeo e $T = \{v_1, \dots, v_r\}$ un sistema de vectores non nulos e ortogonal, cúmprese que T é libre. É máis, se $B = \{v_1, \dots, v_n\}$ é unha base ortogonal de V, entón o vector de coordenadas de calquera vector v escríbese como:

$$v_B = \begin{pmatrix} \dfrac{< v_1, v >}{\|v_1\|^2} \\ \vdots \\ \dfrac{< v_n, v >}{\|v_n\|^2} \end{pmatrix}.$$

Os coeficientes dados na igualdade anterior chámanse coeficientes de Fourier de v na base ortogonal B. Se B fose ortonormal, os coeficientes de Fourier de v quedan como $< v_j, v >$.

J-B. J. Fourier
(Amédée Félix Barthélemy Geille /Wikimedia Commons)

Dentro deste punto, tamén será importante resaltar que, se B é unha base de V, tense o seguinte:

1) A base B é un sistema ortogonal de vectores, se, e só se, G_B é diagonal.
2) A base B é un sistema ortonormal de vectores, se, e só se, $G_B = I_n$.

Como consecuencia destes resultados poderemos garantir que a matriz de cambio de base entre dúas bases ortonormais de V é unha matriz ortogonal. Doutra banda, se traballamos co produto escalar usual de $\mathbb{R}^n$ e tendo en conta que as columnas ou as filas dunha matriz invertíbel $Q \in \mathcal{M}_{n\times n}(\mathbb{R})$ forman unha base de $\mathbb{R}^n$, é doado probar que as seguintes afirmacións son equivalentes:

1) A matriz Q é ortogonal.
2) As columnas de Q forman unha base ortonormal de $\mathbb{R}^n$ respecto ao produto escalar usual.

Finalizaremos esta parte introducindo o proceso de ortonormalización de Gram-Schmidt grazas ao cal se obtén que nun espazo vectorial euclídeo todo subespazo admite unha base ortonormal. En particular terase que o propio espazo admite unha base ortonormal.

J. P. Gram
(Johannes Hauerslev/Wikimedia Commons)

E. Schmidt
(Konrad Jacobs/Wikimedia Commons)

O terceiro bloque do capítulo dedicarase ao estudo de certas propiedades das matrices simétricas reais. Unha destas propiedades é a que afirma o seguinte: se $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ é unha matriz simétrica, cúmprese que todos os autovalores de A son reais. Ademais, tense que A é simétrica, se e só se, existe unha matriz ortogonal Q tal que Q^tAQ é diagonal. Por tanto, todas as matrices simétricas son diagonalizábeis. A demostración desta afirmación conduciranos ao método que debemos seguir para realizar o cálculo da diagonalización ortogonal dunha matriz simétrica real. Indicaremos como facelo paso a paso do seguinte xeito:

- Calcúlase o espectro de A, $\mathrm{Sp}(A) = \{\lambda_1, \dots, \lambda_r\}$.
- Calcúlase unha base B_{λ_i} de $V(\lambda_i)$, $i \in \{1, \dots, r\}$.
- Aplícase o procedemento de ortonormalización de Gram-Schmidt a cada base B_{λ_i}, obténdose unha base ortonormal C_{λ_i} de $V(\lambda_i)$, $i \in \{1, \dots, r\}$.
- Únense as bases C_{λ_i}. Esta unión dá lugar a unha base ortonormal de $\mathbb{R}^n$ formada por autovectores.
- Defínese a matriz P como aquela cuxas columnas son os vectores da base ortonormal calculada no apartado anterior.
- A matriz diagonal D resulta de facer o produto P^tAP.

Na parte final do capítulo, veremos que, se $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ ten rango r, existen dúas matrices $Q \in \mathcal{M}_{m\times r}(\mathbb{R})$ e $R \in \mathcal{M}_{r\times n}(\mathbb{R})$, tales que $A = QR$, $Q^tQ = I_r$ e $r_{lj} = 0$ se $l > j$. Isto dá lugar á chamada factorización QR da matriz A que será de especial utilidade á hora de resolver problemas de mínimos cadrados.

Para realizar o cálculo da factorización QR dunha matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ con rango r e columnas $\{v_1, \dots, v_n\}$, podemos seguir os seguintes pasos:

- Aplicamos o procedemento de ortonormalización de Gram-Schmidt ao sistema de vectores $T = \{v_1, \dots, v_n\}$ formado polas columnas de A no espazo vectorial euclídeo n- dimensional $\mathbb{R}^n$ co produto escalar usual. Se algún dos vectores e_i que resultan nos cálculos é nulo, elimínase e non se ten en conta no proceso.
- Tomamos $Q \in \mathcal{M}_{m\times r}(\mathbb{R})$ cuxas columnas son os vectores do sistema ortonormal obtido no punto anterior.
- Calcúlase R como $R = Q^tA$.

Tamén podemos proceder partindo do sistema de vectores dado polas primeiras r columnas linearmente independentes da matriz A. A matriz $Q \in \mathcal{M}_{m\times r}(\mathbb{R})$ é aquela cuxas columnas son os vectores do sistema ortonormal obtido ao aplicar o procedemento de ortonormalización de Gram-Schmidt a este sistema. Finalmente, $R = Q^tA$.

7.1. Espazos vectoriais con produto escalar

Definición 7.1.1. Sexa V un $\mathbb{R}$-espazo vectorial. Unha forma bilinear é unha aplicación

$$f : V \times V \to \mathbb{R}$$

satisfacendo:

1) $f(\alpha u + \beta v, w) = \alpha f(u, w) + \beta f(v, w) \; \forall \, \alpha, \beta \in \mathbb{R}, \; u, v \in V.$
2) $f(w, \alpha u + \beta v) = \alpha f(w, u) + \beta f(w, v) \; \forall \, \alpha, \beta \in \mathbb{R}, \; u, v \in V.$

Noutras palabras, f é bilinear se é linear nas dúas compoñentes. É por isto que se deducen de forma trivial as seguintes propiedades:

3) $f(\theta_V, u) = f(u, \theta_V) = 0 \; \forall \, u \in V.$
4) $f(\sum_{i=1}^{n} \alpha_i u_i, \sum_{j=1}^{n} \beta_j v_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j f(u_i, v_j) \; \forall \, \alpha_i, \beta_j \in \mathbb{R}, \; u_i, v_j \in V, \; i, j \in \{1, \ldots, n\}.$

Definición 7.1.2. Sexa V un $\mathbb{R}$-espazo vectorial. Unha forma bilinear $f : V \times V \to \mathbb{R}$ dirase simétrica se $f(u, v) = f(v, u) \; \forall \, u, v \in V$.

Definición 7.1.3. Sexa V un $\mathbb{R}$-espazo vectorial. Unha forma bilinear $f : V \times V \to \mathbb{R}$ dirase definida positiva se $f(u, u) > 0 \; \forall \, u \in V, u \neq \theta_V$.

Definición 7.1.4. Sexa V un $\mathbb{R}$-espazo vectorial. Diremos que $f : V \times V \to \mathbb{R}$ é un produto escalar en V se é unha forma bilinear simétrica definida positiva.

A notación clásica que se utiliza para o produto escalar é a seguinte:

$$f(u, v) = < u, v > .$$

Con $(V, < \, , \, >)$ denotaremos ao par dado polo $\mathbb{R}$-espazo vectorial V xunto co produto escalar definido sobre V e chamarémolo espazo vectorial euclídeo. Nótese que, se nun mesmo $\mathbb{R}$-espazo vectorial V temos definidos varios produtos escalares, obteremos diferentes exemplos de espazos vectoriais euclídeos.

Exemplos 7.1.5. 1) Sexa o $\mathbb{R}$-espazo vectorial $V = \mathbb{R}^n$. Para dous vectores $u, v \in V$ defínese a aplicación bilinear

$$< u, v >= u^t v = (u_1, \ldots, u_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^{n} u_i v_i.$$

Este produto resulta ser un produto escalar chamado o produto escalar usual do $\mathbb{R}$-espazo vectorial $\mathbb{R}^n$.

2) Sexa o $\mathbb{R}$-espazo vectorial $V = \Pi_n(\mathbb{R})$. Se para dous polinomios en V definimos

$$< p(x), q(x) >= \int_0^1 p(x) q(x) \mathrm{d}(\mathrm{x}),$$

temos un novo exemplo de produto escalar.

3) No $\mathbb{R}$-espazo vectorial $V = \mathcal{M}_{n\times n}(\mathbb{R})$ pódese definir un produto escalar por:

$$< A, B >= \mathrm{tr}\,(AB^t).$$

Definición 7.1.6. Sexa $(V, < \, , \, >)$ un espazo vectorial euclídeo e sexa $B = \{e_1, \ldots, e_n\}$ unha base de V. Chámase matriz de Gram do produto escalar $< \, , \, >$ respecto á base B á matriz G_B dada por

$$G_B = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix}$$

onde $g_{i,j} =< e_i, e_j >$, $i, j \in \{1, \dots, n\}$.

Esta matriz é simétrica xa que $g_{i,j} =< e_i, e_j >=< e_j, e_i >= g_{ji}$. Ademais, se

$$u = \sum_{i=1}^{n} x_i e_i, \quad v = \sum_{i=1}^{n} y_j e_j,$$

tense que

$$< u, v >=< \sum_{i=1}^{n} x_i e_i, \sum_{i=1}^{n} y_j e_j >= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j < e_i, e_j >= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j g_{ij} = u_B^t G_B v_B,$$

sendo u_B e v_B os vectores de coordenadas de u e v respecto á base B.

Exemplo 7.1.7. Por exemplo, se $V = \mathbb{R}^n$ e $< , >$ é o produto escalar usual, a matriz de Gram de $< , >$ respecto á base canónica $C_n = \{e_1, \dots, e_n\}$ de $\mathbb{R}^n$ ten como coeficientes a

$$g_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

e, como consecuencia, $G_{C_n} = I_n$.

Se en $\mathbb{R}^4$ tomásemos a base

$$B = \left\{ e_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, e_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\},$$

e o produto escalar usual, teriamos que

$$\begin{array}{llll} < e_1, e_1 >= 4 & < e_1, e_2 >= 3 & < e_1, e_3 >= 2 & < e_1, e_4 >= 1 \\ < e_2, e_1 >= 3 & < e_2, e_2 >= 3 & < e_2, e_3 >= 2 & < e_2, e_4 >= 1 \\ < e_3, e_1 >= 2 & < e_3, e_2 >= 2 & < e_3, e_3 >= 2 & < e_3, e_4 >= 1 \\ < e_4, e_1 >= 1 & < e_4, e_2 >= 1 & < e_4, e_3 >= 1 & < e_1, e_4 >= 1 \end{array}$$

e, por tanto, a matriz de Gran respecto á base B é:

$$G_B = \begin{pmatrix} 4 & 3 & 2 & 1 \\ 3 & 3 & 2 & 1 \\ 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Deste xeito, se

$$u = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 1 \end{pmatrix}, \; v = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

obtemos que

$$< u, v >= (2, 2, 2, 1) \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} = 6.$$

Equivalentemente,

$$< u, v >= u_B^t G_B v_B = (1, 1, 0, 0) \begin{pmatrix} 4 & 3 & 2 & 1 \\ 3 & 3 & 2 & 1 \\ 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = 6.$$

7.1.8. Como puidemos ver no exemplo anterior

$$G_{C_4} = I_4, \quad G_B = \begin{pmatrix} 4 & 3 & 2 & 1 \\ 3 & 3 & 2 & 1 \\ 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Por tanto, se cambiamos a base, a matriz de Gram tamén se modifica. En xeral existe unha relación entre as diferentes matrices de Gram asociadas a un mesmo produto escalar. Esta relación é a seguinte: sexan $B = \{e_1, \dots, e_n\}$ e $D = \{u_1, \dots, u_n\}$ dúas bases de V e $P_{B,D}$ a matriz de cambio de base de B a D. Entón, se $<\ ,\ >$ é o produto escalar definido en V, temos que $\forall\, u, v \in V$

$$u_D = P_{B,D} u_B, \;\; v_D = P_{B,D} v_B,$$

$$u_B^t G_B v_B = < u, v > = u_D^t G_D v_D.$$

Logo:

$$u_B^t G_B v_B = (P_{B,D}^{-1} u_D)^t G_B P_{B,D}^{-1} v_D = u_D^t (P_{B,D}^{-1})^t G_B P_{B,D}^{-1} v_D = u_D^t P_{D,B}^t G_B P_{D,B} v_D = u_D^t G_D v_D$$

e isto implica que

$$G_D = P_{D,B}^t G_B P_{D,B}.$$

Se dúas matrices cadradas A e B cumpren que existe unha matriz P invertíbel tal que $P^t A P = B$, dise que son congruentes. Tendo en conta esta definición, acabamos de probar que as matrices de Gram dun mesmo produto escalar respecto a diferentes bases son congruentes.

Por exemplo, se tomamos as bases de $\mathbb{R}^4$ dadas por

$$B = C_4, \;\; D = \left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, e_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\},$$

tense que

$$G_D = P_{D,C_4}^t G_{C_4} P_{D,C_4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} I_4 \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

Definición 7.1.9. Sexa V un $\mathbb{R}$-espazo vectorial. Unha norma sobre V é unha aplicación

$$\|\ \| : V \to \mathbb{R}$$

tal que:

1) $\forall\, u \in V$ $\|u\| \geq 0$. $\|u\| = 0 \Leftrightarrow u = \theta_V$.
2) $\|\alpha u\| = |\alpha| \|u\|$ $\forall\, u \in V$ $\alpha \in \mathbb{R}$.
3) $\|u + v\| \leq \|u\| + \|v\|$ $\forall\, u, v \in V$.

Proposición 7.1.10. *Sexa* $(V, <\ ,\ >)$ *un espazo vectorial euclídeo. A aplicación*

$$\|\ \| : V \to \mathbb{R}$$

dada por

$$\|u\| = +\sqrt{< u, u >}$$

é unha norma.

Demostración. Pola propia definición de aplicación $\| \ \|$, as probas das dúas primeiras condicións de norma son triviais. Vexamos como se obtén a terceira (coñecida como desigualdade triangular ou de Minkowski).

Sexan $u, v \in V$. En primeiro lugar probaremos a desigualdade

$$|<u,v>| \leq \|u\|\|v\|,$$

coñecida como desigualdade de Schwartz. En efecto, se algún dos dous vectores u, v son nulos, a desigualdade anterior é trivial. Por iso podemos supoñer que os dous son non nulos. Neste caso, para todo $\alpha \in \mathbb{R}$, tense que

$$<u-\alpha v, u-\alpha v> \geq \ 0.$$

Equivalentemente,

$$<u,u> -2\alpha <u,v> +\alpha^2 <v,v> \geq \ 0.$$

Tomando $\alpha = \dfrac{<u,v>}{<v,v>}$, a anterior desigualdade convértese en

$$<u,u> -2\frac{<u,v>^2}{<v,v>} + \frac{<u,v>^2}{<v,v>} \geq 0.$$

Logo,

$$\|u\|^2 - \frac{<u,v>^2}{\|v\|^2} \geq 0.$$

Multiplicando a anterior desigualdade por $\|v\|^2$ tense que

$$\|u\|^2\|v\|^2 - <u,v>^2 \geq 0 \Leftrightarrow \|u\|^2\|v\|^2 \geq <u,v>^2$$

e, por tanto,

$$|<u,v>| \leq \|u\|\|v\|.$$

Desta forma, aplicando a desigualdade de Schwartz, para os vectores u e v obtemos que

$$\|u+v\|^2 = <u+v, u+v> = <u,u> +2<u,v> + <v,v> =$$

$$\|u\|^2 + 2<u,v> + \|v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$$

e, como os dous termos da desigualdade son positivos, conséguese a terceira condición de norma

$$\|u+v\| \leq \|u\| + \|v\|.$$

□

Exemplos 7.1.11. Vexamos cales son as normas dadas polos produtos escalares de Exemplos 7.1.5.

1) No caso do produto escalar usual de $\mathbb{R}^n$ a norma asociada será

$$\|v\| = +\sqrt{v_1^2 + \cdots + v_n^2}.$$

2) Se $V = \Pi_n(\mathbb{R})$, tense que a norma dada polo produto escalar

$$<p(x), q(x)> = \int_0^1 p(x)q(x)\mathrm{d(x)}$$

é

$$\|p(x)\| = +\sqrt{\int_0^1 p(x)^2\mathrm{d(x)}}.$$

3) Finalmente, se $V = \mathcal{M}_{n\times n}(\mathbb{R})$, tense que a norma dada polo produto escalar

$$< A, B >= \mathrm{tr}\,(AB^t)$$

é

$$\|A\| = +\sqrt{\mathrm{tr}\,(AA^t)}.$$

Nota 7.1.12. A desigualdade de Schwartz indica que, se u, v son dous vectores non nulos do espazo euclideo V, cúmprese que

$$\frac{|< u, v >|}{\|u\|\|v\|} \leq 1$$

e, como consecuencia,

$$-1 \leq \frac{< u, v >}{\|u\|\|v\|} \leq 1.$$

Tendo en conta esta nota, xa estamos en condicións de introducir a definición de ángulo entre dous vectores.

Definición 7.1.13. Chamarase ángulo entre os vectores u e v dun espazo euclideo V ao único número real $\alpha \in [0, \pi]$ tal que

$$\cos(\alpha) = \frac{< u, v >}{\|u\|\|v\|}.$$

A existencia dunha norma tamén nos permite introducir un concepto de distancia. Definimos a distancia entre u e v como

$$d(u, v) = \|u - v\| = +\sqrt{< u - v, u - v >}.$$

7.2. Ortogonalidade. Bases ortonormais. Procedemento de Gram-Schmidt

Definición 7.2.1. Sexa $(V, < \,,\, >)$ un espazo vectorial euclídeo. Dirase que $T = \{v_1, \dots, v_r\}$ é un sistema ortogonal de vectores respecto do produto escalar $< \,,\, >$, se $< v_i, v_j >= 0$ para todo $i \neq j$. Se ademais se ten que $\|v_i\| = 1$, para todo $i \in \{1, \dots, r\}$, o sistema chamarase ortonormal.

Por exemplo, en $\mathbb{R}^n$ a base canónica é un exemplo de base ortonormal para o produto escalar usual.

Teorema 7.2.2. (Teorema de Pitágoras) *Sexa $(V, < \,,\, >)$ un espazo vectorial euclídeo. Entón, se u, v son vectores ortogonais de V, tense que*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Demostración. A demostración séguese das seguintes igualdades:

$$\|u+v\|^2 =< u+v, u+v >=< u, u > +2 < u, v > + < v, v >=< u, u > + < v, v >= \|u\|^2+\|v\|^2.$$

□

Proposición 7.2.3. *Sexa $(V, < \,,\, >)$ un espazo vectorial euclídeo e sexa $T = \{v_1, \dots, v_r\}$ un sistema de vectores non nulos e ortogonal respecto do produto escalar $< \,,\, >$. Cúmprese que T é libre.*

Demostración. Tomemos a combinación linear

$$\alpha_1 v_1 + \cdots + \alpha_j v_j + \cdots + \alpha_r v_r = \theta_V.$$

Entón,

$$0 =< v_j, \alpha_1 v_1 + \cdots + \alpha_j v_j + \cdots + \alpha_r v_r >= \alpha_1 < v_j, v_1 > + \cdots + \alpha_j < v_j, v_j > + \cdots + \alpha_r < v_j, v_r >=$$
$$\alpha_j < v_j, v_j >$$

e, como $< v_j, v_j >\neq 0$, tense que $\alpha_j = 0$. Por tanto, o sistema é libre.

□

Como consecuencia das definicións de sistema ortogonal e de sistema ortonormal de vectores, podemos obter o resultado seguinte.

Proposición 7.2.4. *Sexa $(V, <\ ,\ >)$ un espazo vectorial euclídeo e sexa $B = \{e_1, \ldots, e_n\}$ unha base de V. Sexa G_B a matriz de Gram do produto escalar $<\ ,\ >$ respecto da base B. Cúmprese o seguinte:*

1) *A base B é un sistema ortogonal de vectores, se, e só se, G_B é diagonal.*
2) *A base B é un sistema ortonormal de vectores, se, e só se, $G_B = I_n$.*

Nota 7.2.5. A vantaxe que ten traballar con bases ortonormais é que, como $G_B = I_n$, obtense que

$$< u, v >= u_B^t v_B,\ \forall u, v \in V$$

ou, dito doutra forma, tomando coordenadas respecto á base B o produto escalar obtense como o produto escalar usual de $\mathbb{R}^n$.

É evidente que, se $B = \{v_1, \cdots, v_n\}$ é unha base ortogonal entón

$$C = \left\{ \frac{v_1}{\|v_1\|}, \cdots, \frac{v_n}{\|v_n\|} \right\}$$

é unha base ortonormal.

Proposición 7.2.6. *Sexa $(V, <\ ,\ >)$ un espazo vectorial euclídeo. A matriz de cambio de base entre dúas bases ortonormais de V é unha matriz ortogonal.*

Demostración. Polo visto en 7.1.8, sabemos que, se B e D son dúas bases de V,

$$G_D = P_{D,B}^t G_B P_{D,B}.$$

Agora ben, se B y D son ortonormais, $G_B = G_D = I_n$ e, por tanto,

$$P_{D,B}^t P_{D,B} = I_n.$$

Como consecuencia, a matriz de cambio de base $P_{D,B}$ é ortogonal. □

Nota 7.2.7. Se temos en conta que as columnas ou as filas dunha matriz invertíbel Q, cadrada de orde n, forman unha base de $\mathbb{R}^n$, é trivial probar que as seguintes afirmacións son equivalentes:

1) A matriz Q é ortogonal.
2) A matriz Q^t é ortogonal.
3) As columnas de Q forman unha base ortonormal de $\mathbb{R}^n$ respecto ao produto escalar usual.
4) As columnas de Q^t forman unha base ortonormal de $\mathbb{R}^n$ respecto ao produto escalar usual.

Proposición 7.2.8. *Sexa $(V, <\ ,\ >)$ un espazo vectorial euclídeo e sexa $B = \{v_1, \ldots, v_n\}$ unha base ortogonal de V. Entón*

$$v_B = \begin{pmatrix} \dfrac{< v_1, v >}{\|v_1\|^2} \\ \vdots \\ \dfrac{< v_n, v >}{\|v_n\|^2} \end{pmatrix}.$$

Demostración. Se $v_B = (x_1, \dots, x_n)^t$, tense que

$$v = x_1 v_1 + \cdots + x_j v_j + \cdots + x_n v_n$$

e, por tanto,

$$< v_j, v >=$$
$$< v_j, x_1 v_1 + \cdots + x_j v_j + \cdots + x_n v_n >= x_1 < v_j, v_1 > + \cdots + x_j < v_j, v_j > + \cdots + x_n < v_j, v_n >=$$
$$x_j < v_j, v_j >= x_j \|v_j\|^2.$$

Como consecuencia, $x_j = \dfrac{< v_j, v >}{\|v_j\|^2}$, para todo $j \in \{1, \cdots, n\}$.

□

Nota 7.2.9. Os coeficientes

$$x_j = \frac{< v_j, v >}{\|v_j\|^2}$$

dados na anterior proposición chámanse coeficientes de Fourier de v na base ortogonal B. Se B fose ortonormal, os coeficientes de Fourier de v son

$$x_j =< v_j, v > .$$

A continuación estudaremos o procedemento de ortonormalización de Gram-Schmidt que permite obter unha base ortonormal a partir dunha base calquera dun espazo vectorial euclídeo V.

Teorema 7.2.10. (Procedemento de ortonormalización de Gram-Schmidt) *Sexa un espazo vectorial euclídeo $(V, < \ , \ >)$ de dimensión n e sexa $B = \{v_1, \dots, v_n\}$ unha base de V. Existe unha base ortonormal $C = \{u_1, \dots, u_n\}$ tal que*

$$\forall\, i \in \{1, \dots, n\} \ \ \langle \{u_1, \dots, u_i\} \rangle = \langle \{v_1, \cdots, v_i\} \rangle.$$

Demostración. O sistema de vectores C calcúlase do xeito seguinte:

$$\begin{cases} e_1 = v_1, \quad u_1 = \dfrac{e_1}{\|e_1\|}, \\ e_2 = v_2 - < v_2, u_1 > u_1, \quad u_2 = \dfrac{e_2}{\|e_2\|}, \\ \qquad \cdots \\ e_i = v_i - \displaystyle\sum_{j=1}^{i-1} < v_i, u_j > u_j, \quad u_i = \dfrac{e_i}{\|e_i\|}, \\ \qquad \cdots \\ e_n = v_n - \displaystyle\sum_{j=1}^{n-1} < v_i, u_j > u_j, \quad u_n = \dfrac{e_n}{\|e_n\|}. \end{cases}$$

Probaremos que $\forall\, i \in \{1, \dots, n\}$ $\langle \{u_1, \dots, u_i\} \rangle = \langle \{v_1, \dots, v_i\} \rangle$ e que $\{u_1, \dots, u_i\}$ é un sistema ortonormal. A proba farémola por indución. Para $k = 1$ é evidente. Supoñamos que se cumpre que $\langle \{u_1, \dots, u_{i-1}\} \rangle = \langle \{v_1, \dots, v_{i-1}\} \rangle$ e que $\{u_1, \dots, u_{i-1}\}$ é un sistema ortonormal. Tense que probar que $\langle \{u_1, \dots, u_{i-1}, u_i\} \rangle = \langle \{v_1, \dots, v_{i-1}, v_i\} \rangle$ e que $\{u_1, \dots, u_{i-1}, u_i\}$ é un sistema ortonormal.

Por construción de e_i e pola hipótese de indución tense que

$$u_i \in \langle \{u_1, \dots, u_{i-1}, v_i\} \rangle = \langle \{v_1, \dots, v_{i-1}, v_i\} \rangle.$$

Por tanto,

$$\langle \{u_1, \dots, u_{i-1}, u_i\} \rangle = \langle \{v_1, \dots, v_{i-1}, v_i\} \rangle.$$

Por hipótese de indución sabemos que $\{u_1, \dots, u_{i-1}\}$ é un sistema ortonormal e, por construción, temos que

$$u_i = \frac{v_i}{\|e_i\|} - \sum_{j=1}^{i-1} \frac{< v_i, u_j >}{\|e_i\|} u_j.$$

Entón, $\forall\, k < i$,

$$< u_i, u_k > = < \frac{v_i}{\|e_i\|} - \sum_{j=1}^{i-1} \frac{< v_i, u_j >}{\|e_i\|} u_j, u_k > =$$

$$\frac{< v_i, u_k >}{\|e_i\|} - \sum_{j=1}^{i-1} \frac{< v_i, u_j >}{\|e_i\|} < u_j, u_k > = \frac{< v_i, u_k >}{\|e_i\|} - \frac{< v_i, u_k >}{\|e_i\|} < u_k, u_k > = 0$$

xa que $< u_k, u_k > = 1$. □

Exemplo 7.2.11. Sexa U o subespazo do $\mathbb{R}$-espazo vectorial $\mathbb{R}^4$ xerado polos vectores

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}, v_3 = \begin{pmatrix} 3 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

Supoñamos que en $\mathbb{R}^4$ consideramos o produto escalar usual e a norma dada por este. Se aplicamos o procedemento de ortonormalización de Gram-Schmidt a v_1, v_2 e v_3 obteremos unha base ortonormal de U. En efecto:

$$e_1 = v_1, \quad u_1 = \frac{e_1}{\|e_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix},$$

$$e_2 = v_2 - < v_2, u_1 > u_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix} - (1, 2, 0, -1) \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} =$$

$$\begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix} - \frac{2}{\sqrt{2}} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$

$$u_2 = \frac{e_2}{\|e_2\|} = \frac{1}{2} \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

$$e_3 = v_3 - < v_3, u_1 > u_1 - < v_3, u_2 > u_2 =$$

$$\begin{pmatrix} 3 \\ 1 \\ 1 \\ -1 \end{pmatrix} - (3, 1, 1, -1) \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} - (3, 1, 1, -1) \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} =$$

$$\begin{pmatrix} 3 \\ 1 \\ 1 \\ -1 \end{pmatrix} - \frac{4}{\sqrt{2}} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

$$u_3 = \frac{e_3}{\|e_3\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ 0 \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Logo, a base ortonormal de U é

$$C = \left\{ u_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix},\ u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},\ u_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ 0 \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right\}.$$

7.3. Diagonalización ortogonal de matrices simétricas

Neste apartado veremos que toda matriz simétrica é diagonalizábel dunha forma especial.

Teorema 7.3.1. *Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ unha matriz simétrica. Cúmprese que todos os autovalores de A son reais.*

Demostración. Supoñamos que $\lambda = a+bi$ é un autovalor de A e que $w = u+iv$ é un autovector asociado con $u, v \in \mathbb{R}^n$. Temos que

$$Aw = Au + iAv = \lambda(u+iv) = (a+bi)(u+iv) = (au-bv) + i(bu+av) \ \Leftrightarrow$$
$$Au = au - bv, \quad Av = bu + av.$$

Como A é simétrica, os produtos escalares usuais $< Au, v >$ e $< u, Av >$ son iguais e tense que

$$< Au, v > = < au - bv, v > = a < u, v > -b < v, v >,$$
$$< u, Av > = < u, bu + av > = b < u, u > +a < u, v > .$$

Se restamos as expresións anteriores, obtemos

$$b(< u, u > + < v, v >) = 0.$$

Como $< u, u > + < v, v > = \|u\|^2 + \|v\|^2 \neq 0$, tense que $b = 0$ e, por tanto, λ é real. □

Teorema 7.3.2. *Sexa $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. A matriz A é simétrica se, e só se, existe unha matriz ortogonal Q tal que Q^tAQ é diagonal.*

Demostración. Se existe unha matriz ortogonal Q tal que $Q^tAQ = D$ é diagonal, entón

$$Q^tAQ = D = D^t = Q^tA^tQ$$

e, como consecuencia, $A = A^t$.

Se A é simétrica, polo teorema anterior, sabemos que todos os seus autovalores $\lambda_1, \dots, \lambda_r$ son reais. Sexa B_{λ_i} unha base de $V(\lambda_i)$, $i \in \{1, \cdots, r\}$. Se o sistema libre $B = B_{\lambda_1} \cup \cdots \cup B_{\lambda_r}$ non é unha base de V (i.e. A non é diagonalizábel), temos que $U = \langle B_{\lambda_1} \cup \cdots \cup B_{\lambda_r} \rangle \neq \mathbb{R}^n$. Tomando $U^\perp = \{v \in \mathbb{R}^n \ \mid < v, u > = 0,\ \forall\, u \in U\}$ temos que U é un subespazo de $\mathbb{R}^n$ (chamado complemento ortogonal de U) e que $U \cap U^\perp = \{\theta_{\mathbb{R}^n}\}$. Ademais, todo $w \in \mathbb{R}^n$ escríbese de forma única como

$$w = u + v,\ u \in U, v \in U^\perp.$$

Se restrinximos a aplicación linear $f_A : \mathbb{R}^n \to \mathbb{R}^n$ a $U^\perp$, tense que $f_{A_{|U^\perp}} : U^\perp \to \mathbb{R}^n$ cumpre que non é nula e ademais $f_A(U^\perp) \subseteq U^\perp$. Por tanto, debe ter alomenos un autovector $v \neq \theta_{\mathbb{R}^n}$. Logo, $Av = \lambda v$, para un certo λ real, e isto implica que $v \in U \cap U^\perp$. Entón, $v = \theta^n_{\mathbb{R}}$, o que é un absurdo que provén de supoñer que $B = B_{\lambda_1} \cup \cdots \cup B_{\lambda_r}$ non é unha base de V. Por tanto, A é diagonalizábel.

Supoñamos $\theta_{\mathbb{R}^n} \neq v \in V_{\lambda_i}$ e que $\theta_{\mathbb{R}^n} \neq u \in V_{\lambda_j}$ con $i \neq j$. Como A é unha matriz simétrica

$$\lambda_i < u, v >=< Au, v >=< u, Av >= \lambda_j < u, v > \Leftrightarrow (\lambda_i - \lambda_j) < u, v >= 0 \Leftrightarrow < u, v >= 0.$$

Logo, autovectores asociados a autovalores distintos son ortogonais. Tendo en conta este feito, se aplicamos o procedemento de ortonormalización de Gram-Schmidt á base B_{λ_i} de cada subespazo propio e unimos os sistemas resultantes, obtemos unha base ortonormal de $\mathbb{R}^n$ formada por vectores propios. Se P é a matriz cuxas columnas son os vectores desta base ortonormal, P resulta ortogonal e P^tAP diagonal. □

Nota 7.3.3. A demostración do anterior teorema indícanos como debemos realizar o cálculo da diagonalización ortogonal dunha matriz simétrica real. Séguense os seguintes pasos:

- Calcúlase o espectro de A, $\mathrm{Sp}(A) = \{\lambda_1, \ldots, \lambda_r\}$.
- Calcúlase unha base B_{λ_i} de $V(\lambda_i)$, $i \in \{1, \ldots, r\}$.
- Aplícase o procedemento de ortonormalización de Gram-Schmidt a cada base B_{λ_i}, obténdose unha base ortonormal C_{λ_i} de $V(\lambda_i)$, $i \in \{1, \ldots, r\}$.
- Únense as bases C_{λ_i}. Esta unión dá lugar a unha base ortonormal de $\mathbb{R}^n$ formada por autovectores.
- Defínese a matriz P como aquela cuxas columnas son os vectores da base ortonormal calculada no apartado anterior.
- A matriz diagonal D resulta de facer o produto P^tAP.

Exemplo 7.3.4. Sexa a matriz

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Como A é simétrica pódese diagonalizar de forma ortogonal. Calculemos en primeiro lugar os autovalores e unha base de cada subespazo propio de A.

$$p_A(x) = \det(A - xI_3) = \begin{vmatrix} -x & 1 & 0 \\ 1 & -x & 1 \\ 1 & 1 & -x \end{vmatrix} = \begin{vmatrix} -x & 1 & 0 \\ 0 & -(x+1) & 1+x \\ 1 & 1 & -x \end{vmatrix} = \begin{vmatrix} -x & 1 & 2 \\ 0 & -(x+1) & 0 \\ 1 & 1 & 1-x \end{vmatrix}$$

$$= -(x+1)\begin{vmatrix} -x & 2 \\ 1 & 1-x \end{vmatrix} = -(x+1)(x^2 - x - 2) = -(x+1)^2(x-2).$$

Por tanto, os autovalores de A son $\lambda_1 = -1$ y $\lambda_2 = 2$, con $\mathrm{ma}(-1) = 2$, $\mathrm{ma}(2) = 1$. Ademais,

$$V(-1) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid (A + I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x + y + z = 0 \right\} = \langle \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\} \rangle$$

$$V(2) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid (A - 2I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid -2x + y + z = 0 = x - 2y + z \right\} = \langle \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \rangle.$$

Logo, as bases de $V(-1)$ e de $V(2)$ son, respectivamente,

$$B_{-1} = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}, \quad B_2 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Aplicando o procedemento de ortonormalización de Gram-Schmidt ás bases B_{-1} e B_2, obtemos unha base ortonormal para $V(-1)$ e outra para $V(2)$:

$$C_{-1} = \left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{pmatrix} \right\}, \quad C_2 = \left\{ \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right\}.$$

En consecuencia, a base ortonormal de $\mathbb{R}^3$ é

$$C = \left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right\}$$

e a matriz P, dada por

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix},$$

é ortogonal e

$$P^t A P = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = D.$$

Nota 7.3.5. Como aplicación da diagonalización ortogonal dunha matriz simétrica real A de rango r, tense a súa descomposición espectral e unha aplicación ao cálculo de potencias de A. Neste caso o que se observa é que, se $u_1, u_2, \dots, u_n$ son as columnas de P, sendo P a matriz ortogonal tal que $P^t A P = D$, tense que

$$A = \lambda_1 u_1 u_1^t + \lambda_2 u_2 u_2^t + \cdots + \lambda_r u_r u_r^t,$$

onde $\lambda_1, \lambda_2, \cdots, \lambda_r$ son os autovalores non nulos de A. Logo, as potencias de A calcúlanse doadamente como:

$$A^k = \lambda_1^k u_1 u_1^t + \lambda_2^k u_2 u_2^t + \cdots + \lambda_r^k u_r u_r^t.$$

7.4. Factorización QR

Sexa $(V, <\ ,\ >)$ un espazo vectorial euclídeo e sexa $T = \{v_1, \dots, v_{r-1}\}$ un sistema libre de vectores de V. Se aplicamos o procedemento de ortonormalización de Gram-Schmidt ao sistema de vectores T, obtemos un sistema ortonormal $C = \{u_1, \dots, u_{r-1}\}$ de vectores de V tal que $\langle T\rangle = \langle C\rangle$ e, ademais, téñense as igualdades:

$$\begin{cases} v_1 = \|e_1\| u_1, \\ v_2 =< v_2, u_1 > u_1 + \|e_1\| u_2, \\ \qquad \cdots \\ v_i = \displaystyle\sum_{j=1}^{i-1} < v_i, u_j > u_j + \|e_i\| u_i, \\ \qquad \cdots \\ v_{r-1} = \displaystyle\sum_{j=1}^{r-2} < v_i, u_j > u_j + \|e_{r-1}\| u_{r-1}. \end{cases}$$

Supoñamos que tomamos $v_r \in \langle T\rangle = \langle C\rangle$. Ao aplicar o procedemento de ortonormalización de Gram-Schmidt ao sistema de vectores $T \cup \{v_r\}$ obteriamos o sistema C e

$$e_r = v_r - \sum_{j=1}^{r-1} < v_i, u_j > u_j = \theta_V,$$

xa que, se $v_r \in \langle C\rangle$, $v_r = \alpha_1 u_1 + \cdots + \alpha_{r-1} u_{r-1}$ con

$$\alpha_j =< v_i, u_j >, \quad j \in \{1, \dots, r-1\}.$$

Logo, cando aplicamos o procedemento de ortonormalización de Gram-Schmidt a un sistema onde existen vectores que son combinación linear doutros vectores do sistema, para cada un destes vectores, obtemos o vector nulo θ_V.

Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Supoñamos que as súas columnas son

$$\left(\ v_1 \ \middle|\ v_2 \ \middle|\ \cdots \ \middle|\ v_n\ \right)$$

e que rang $(A) = r \leq n$. Tomemos o espazo euclídeo $(\mathbb{R}^n, <\ ,\ >)$ co produto escalar usual. Se aplicamos o procedemento de ortonormalización de Gram-Schmidt ao sistema de vectores $T = \{v_1, \dots, v_n\}$ formado polas columnas de A, obtemos os mesmos vectores non nulos que forman o sistema ortonormal $C = \{u_{k1}, \dots, u_{kr}\}$ que resultaría ao aplicar o mesmo procedemento de ortonormalización ao sistema $T = \{v_{k1}, \dots, v_{kr}\}$ formado unicamente polas r primeiras columnas independentes de A.

Así temos que, cando $v_{ki} \in T$,

$$v_{ki} = \sum_{j=1}^{i-1} < v_{ki}, u_{kj} > u_{kj} + \|e_{ki}\| u_{ki}, \quad e_{ki} \neq \theta_{\mathbb{R}^n}$$

e, se $v_{ki} \notin T$, $e_{ki} = \theta_{\mathbb{R}^n}$, sendo ademais certa a igualdade

$$v_{ki} = \sum_{kj<ki} < v_{ki}, u_{kj} > u_{kj}.$$

Por tanto, dada a matriz $Q \in \mathcal{M}_{m\times r}(\mathbb{R})$ cuxas columnas son

$$\left(\ u_{k1} \ \middle|\ u_{k2} \ \middle|\ \cdots \ \middle|\ u_{kr}\ \right),$$

grazas a que o sistema $C = \{u_{k1}, \dots, u_{kr}\}$ é ortonormal, temos que $Q^t Q = I_r$ e, ademais,

$$A = QR,$$

sendo $R = Q^t A = (r_{lj}) \in \mathcal{M}_{r\times n}(\mathbb{R})$ unha matriz tal que $r_{lj} = 0$, se $l > j$.

Finalmente, é evidente que $\operatorname{rang}(A) = \operatorname{rang}(Q) = \operatorname{rang}(R) = r$.

Como consecuencia do anterior temos o seguinte resultado.

Teorema 7.4.1. **(Factorización QR)** *Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ con rango r. Existen dúas matrices $Q \in \mathcal{M}_{m\times r}(\mathbb{R})$ e $R = (r_{lj}) \in \mathcal{M}_{r\times n}(\mathbb{R})$, tales que $A = QR$, $Q^tQ = I_r$ e $r_{lj} = 0$ se $l > j$.*

Nota 7.4.2. Para realizar o cálculo efectivo da factorización QR dunha matriz $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ con rango r e columnas $\{v_1, \dots, v_n\}$, temos os seguintes pasos:

- Aplicamos o procedemento de ortonormalización de Gram-Schmidt ao sistema de vectores $T = \{v_1, \dots, v_n\}$ formado polas columnas de A no espazo vectorial euclídeo n-dimensional $(\mathbb{R}^n, <\ ,\ >)$ co produto escalar usual. Se algún dos vectores e_i que resultan nos cálculos é nulo, elimínase e non se ten en conta no proceso.
- Tomamos $Q \in \mathcal{M}_{m\times r}(\mathbb{R})$ cuxas columnas son os vectores do sistema ortonormal obtido no punto anterior.
- Calcúlase R como $R = Q^tA$.

Tamén podemos proceder partindo do sistema de vectores dado polas primeiras r columnas linearmente independentes da matriz A. A matriz $Q \in \mathcal{M}_{m\times r}(\mathbb{R})$ é aquela cuxas columnas son os vectores do sistema ortonormal obtido ao aplicar o procedemento de ortonormalización de Gram-Schmidt a este sistema. Finalmente, $R = Q^tA$.

Exemplos 7.4.3. 1) Sexa a matriz A cuxas columnas son os vectores do sistema considerado no Exemplo 7.2.11. Entón

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{pmatrix}.$$

Supoñamos que en $\mathbb{R}^4$ consideramos o produto escalar usual. Se aplicamos o procedemento de ortonormalización de Gram-Schmidt ás columnas de A, obtemos un sistema ortonormal

$$C = \left\{ u_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix},\ u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},\ u_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ 0 \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right\}$$

e, por tanto, $A = QR$ onde

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{3}} \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

e

$$R = Q^tA = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 & 0 \\ \frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} 1 & 1 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{pmatrix} =$$

$$\begin{pmatrix} \frac{2}{\sqrt{2}} & \frac{2}{\sqrt{2}} & \frac{4}{\sqrt{2}} \\ 0 & 2 & 1 \\ 0 & 0 & \frac{3}{\sqrt{3}} \end{pmatrix}.$$

Deste xeito, $\operatorname{rang}(A) = \operatorname{rang}(Q) = \operatorname{rang}(R) = 3$.

2) Se, por exemplo,

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

considerando en $\mathbb{R}^4$ o produto escalar usual, ao aplicar o procedemento de ortonormalización de Gram-Schmidt ás columnas de A, obtemos o seguinte:

$$e_1 = v_1, \quad u_1 = \frac{e_1}{\|e_1\|} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{pmatrix}.$$

$$e_2 = v_2 - < v_2, u_1 > u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - (1,0,1) \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{pmatrix} =$$

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/2 \\ 1 \end{pmatrix}$$

$$u_2 = \frac{e_2}{\|e_2\|} = \frac{\sqrt{6}}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{pmatrix},$$

$$e_3 = v_3 - < v_3, u_1 > u_1 - < v_3, u_2 > u_2 =$$

$$\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - (2,1,1) \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix} - (2,1,1) \begin{pmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{pmatrix} \begin{pmatrix} \sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{pmatrix} =$$

$$\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \frac{3}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Logo, $A = QR$, onde

$$Q = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ 0 & \frac{\sqrt{6}}{3} \end{pmatrix} \in \mathcal{M}_{3\times 2}(\mathbb{R})$$

e

$$R = Q^t A = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ \sqrt{6}/6 & -\sqrt{6}/6 & \sqrt{6}/3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \sqrt{2}/2 & 3\sqrt{2}/2 \\ 0 & \sqrt{6}/2 & \sqrt{6}/2 \end{pmatrix}.$$

Por tanto, $\operatorname{rang}(A) = \operatorname{rang}(Q) = \operatorname{rang}(R) = 2$.

Nota 7.4.4. Aplicando a factorización QR temos que, se $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ con rango n, $Q \in \mathcal{M}_{m\times n}(\mathbb{R})$ e $R = Q^t A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Como o $\operatorname{rang}(R) = n$, obtemos que R é invertíbel. Doutra banda, se $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ con rango n (A é invertíbel), $Q \in \mathcal{M}_{n\times n}(\mathbb{R})$ é tal que $Q^t Q = QQ^t = I_n$ e, por tanto, Q resulta unha matriz ortogonal. Ademais $R = Q^t A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Como o $\operatorname{rang}(R) = n$, obtemos que R é invertíbel.

7.5. Problemas propostos

1. Utilice o método de Gram-Schmidt para achar unha base ortonormal do subespazo de $\mathbb{R}^3$ xerado polo seguinte sistema de vectores
   a) $S = \{(1,1,0)^t, (1,0,0)^t, (1,1,1)^t\}$.
   b) $T = \{(1,0,1)^t, (2,-1,1)^t, (1,-1,0)^t\}$.
2. Determine os valores dos parámetros a, b, x, y, z, t para que a matriz real
$$A = \begin{pmatrix} 1/2 & 1/6 & a & x \\ 1/2 & 1/6 & b & y \\ 1/2 & 1/2 & 0 & z \\ 1/2 & -5/6 & 0 & t \end{pmatrix}$$
   sexa ortogonal, sendo a e x positivos.
3. Considérase a matriz $A \in \mathcal{M}_{4\times 4}(\mathbb{R})$ dada por
$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$
   Se é posíbel, calcule unha matriz ortogonal P e unha matriz diagonal D tales que $P^tAP = D$.
4. Considérase a matriz $A \in \mathcal{M}_{4\times 4}(\mathbb{R})$ dada por
$$A = \begin{pmatrix} -2 & -1 & 0 & 2 \\ -1 & -2 & 0 & -2 \\ 0 & 0 & -3 & 0 \\ 2 & -2 & 0 & 1 \end{pmatrix}.$$
   Se é posíbel, encontre unha matriz ortogonal P e unha matriz diagonal D tales que $P^tAP = D$.
5. Considérase a matriz
$$A = \begin{pmatrix} -1 & -1 & 2 \\ -1 & -1 & -2 \\ 2 & -2 & 2 \end{pmatrix}.$$
   Determine unha matriz ortogonal P e unha matriz diagonal D tales que $AP = PD$.
6. Atope a factorización QR das seguintes matrices
   a)
$$A = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}.$$
   b)
$$B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$
   c)
$$C = \begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}.$$

7. Considérase a matriz real

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix}.$$

   a) Ache a factorización QR de A.
   b) Utilice o cálculo do apartado anterior para resolver o sistema $Ax = b$ sendo $b^t = (0, 1, 3, -2)$.

CAPíTULO 8

Formas cuadráticas

Neste capítulo abordarase o estudo das formas cuadráticas e verase que parte do traballo realizado coas matrices simétricas, especialmente todo o que ten que ver coa diagonalización ortogonal, será de gran utilidade á hora de clasificalas. As formas cuadráticas aparecen con frecuencia en certas aplicacións da álxebra linear á enxeñería, no procesado de sinais e en problemas de optimización ligados ao cálculo de extremos de funcións de varias variábeis. Tamén as podemos atopar en física, xeometría diferencial, economía e estatística.

Comezaremos introducindo a noción de forma cuadrática como unha aplicación

$$\omega : \mathbb{R}^n \to \mathbb{R}$$

dada por

$$\omega(v) = v^t A v, \ \forall\, v \in \mathbb{R}^n$$

onde $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. A continuación veremos que se $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ é unha forma bilinear e C_n é a base canónica de $\mathbb{R}^n$, tense que a f podemos asociarlle unha forma cuadrática definida por

$$\omega_f(v) = v^t G_{\mathcal{C}_n} v, \ \forall\, v \in \mathbb{R}^n$$

onde G_{C_n} é a matriz de Gram de f respecto á base C_n. De xeito inverso demostrarase que dada $\omega : \mathbb{R}^n \to \mathbb{R}$, $\omega(v) = v^t A v$, unha forma cuadrática en $\mathbb{R}^n$, existe unha forma bilinear

$$f_\omega(u, v) = \frac{1}{2}(\omega(u + v) - w(u) - w(v))$$

tal que

$$f_\omega(u, v) = u^t(\frac{1}{2}(A + A^t))v$$

e, por tanto, $f_\omega : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ é unha forma bilinear simétrica xa que a matriz $\frac{1}{2}(A + A^t)$ é simétrica. Tamén se cumpre que f_ω é a única forma bilinear simétrica para a cal $\omega_{f_\omega} = \omega$. Esta forma bilinear simétrica única chamarase forma polar de ω. Grazas a isto último, poderemos garantir que, dada $\omega : \mathbb{R}^n \to \mathbb{R}$ unha forma cuadrática, existe unha única matriz simétrica $M(\omega)$ tal que

$$\omega(v) = v^t M(\omega) v, \ \forall\, v \in \mathbb{R}^n$$

sendo $M(\omega) = \frac{1}{2}(A + A^t)$, se $\omega(v) = v^t A v$.

En todas as definicións anteriores tomamos coordenadas respecto á base canónica de $\mathbb{R}^n$. Como en casos anteriores é importante resaltar que, se cambiamos a base, a expresión da forma cuadrática tamén o fai aínda que, como sempre, o cambio está controlado por unha matriz de cambio de base. Sexa B outra base de $\mathbb{R}^n$ e chamemos P_{B,C_n} á matriz de cambio de base de B á base canónica de $\mathbb{R}^n$. Como sabemos, $\forall\, v \in \mathbb{R}^n$,

$$v = P_{B,C_n} v_B$$

sendo v_B o vector de coordenadas de v na base B. Por tanto:

$$\omega(v) = v^t M(\omega) v = v_B^t P_{B,C_n}^t M(\omega) P_{B,C_n} v_B.$$

Á matriz $M(\omega)_B = P_{B,C_n}^t M(\omega) P_{B,C_n}$ chámaselle matriz asociada a ω respecto á base B e é congruente con $M(\omega)$. Evidentemente tamén é simétrica e tense que $M(\omega)_{C_n} = M(\omega)$.

Logo, a expresión da forma cuadrática na base B é

$$\omega(v) = v_B^t M(\omega)_B v_B.$$

Cando a matriz $M(\omega)_B$ é diagonal diremos que a base B é ortogonal respecto a ω. Neste caso tense que, se $v_B^t = (x_1, \dots, x_n)$,

$$\omega(v) = v_B^t M(\omega)_B v_B =$$

$$(x_1, x_2, \dots, x_n) \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1^2 + a_2 x_2^2 + \cdots + a_n x_n^2.$$

Se a base B ten como vectores a $\{e_1, \dots, e_n\}$, o vector de coordenadas de e_i na base será

$$e_{iB} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

onde o 1 atópase na compoñente i-ésima. Logo, $w(e_i) = a_i$ e, como consecuencia,

$$M(\omega)_B = \begin{pmatrix} \omega(e_1) & 0 & \cdots & 0 \\ 0 & \omega(e_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega(e_n) \end{pmatrix}.$$

No seguinte apartado veremos que dada unha forma cuadrática ω podemos atopar unha base ortogonal para ela. Isto é equivalente a atopar unha matriz invertíbel P tal que $P^t M(\omega) P$ sexa diagonal. Obviamente, as columnas de P formarán a base buscada. Indicaremos dous camiños para calcular P. O primeiro deles baseado na diagonalización ortogonal e o segundo na diagonalización por congruencia.

Para a clasificación das formas cuadráticas, estes dous procesos converxen no mesmo grazas á lei de inercia de Sylvester que asegura o seguinte: se $\omega : \mathbb{R}^n \to \mathbb{R}$ é unha forma cuadrática e D_1, D_2 son matrices diagonais asociadas a ω respecto a distintas bases, entón o número de elementos positivos e negativos das súas diagonais principais é o mesmo.

Tendo en conta todo isto poderemos clasificar as formas cuadráticas como (semi)definidas positivas, (semi)definidas negativas ou indefinidas, contando o número de elementos negativos e positivos que temos en calquera das matrices diagonais asociadas á forma cuadrática que esteamos a estudar. Máis concretamente, sexan $\omega : \mathbb{R}^n \to \mathbb{R}$ unha forma cuadrática e B unha base ortogonal respecto de ω. Se chamamos sinatura de ω ao par:

$$\operatorname{sg}(\omega) = (p, q),$$

onde p, q denotan, respectivamente, o número de elementos positivos e negativos da matriz diagonal asociada a ω na base B, tense que:

1) A forma cuadrática ω é definida positiva ($\omega(v) > 0,\ \forall\, v \in \mathbb{R}^n,\ v \neq \theta_{\mathbb{R}^n}$), se, e só se, $\operatorname{sg}(\omega) = (n, 0)$.
2) A forma cuadrática ω é definida negativa ($\omega(v) < 0,\ \forall\, v \in \mathbb{R}^n,\ v \neq \theta_{\mathbb{R}^n}$), se, e só se, $\operatorname{sg}(\omega) = (0, n)$.
3) A forma cuadrática ω é semidefinida positiva ($\omega(v) \geq 0,\ \forall\, v \in \mathbb{R}^n$), se, e só se, $\operatorname{sg}(\omega) = (r, 0)$ con $r < n$.

4) A forma cuadrática ω é semidefinida negativa ($\omega(v) \leq 0,\ \forall\, v \in \mathbb{R}^n$) se, e só se, $\operatorname{sg}(\omega) = (0, s)$ con $s < n$.
5) A forma cuadrática ω é indefinida (existen dous vectores non nulos u, v de $\mathbb{R}^n$ tales que $\omega(u) < 0$ e $\omega(v) > 0$), se, e só se, $\operatorname{sg}(\omega) = (r, s)$ con r, s non nulos.

Equivalentemente, se utilizamos a diagonalización ortogonal, teremos que dada ω unha forma cuadrática e $M(\omega)$ a súa matriz asociada, cúmprese o seguinte:

1) A forma cuadrática ω é definida positiva, se, e só se, os autovalores de $M(\omega)$ son todos positivos.
2) A forma cuadrática ω é definida negativa, se, e só se, os autovalores de $M(\omega)$ son todos negativos.
3) A forma cuadrática ω é semidefinida positiva, se, e só se, os autovalores de $M(\omega)$ son todos non negativos sendo algún deles nulo.
4) A forma cuadrática ω é semidefinida negativa, se, e só se, os autovalores de $M(\omega)$ son todos non positivos sendo algún deles nulo.
5) A forma cuadrática ω é indefinida, se, e só se, existen autovalores de $M(\omega)$ positivos e negativos.

Finalmente acabaremos o capítulo expoñendo o coñecido como criterio de Sylvester para clasificar formas cuadráticas. Este criterio clasifica as formas cuadráticas mediante o uso de determinantes do seguinte xeito: se $\omega : \mathbb{R}^n \to \mathbb{R}$ é unha forma cuadrática dada por $\omega(v) = v^t A v$ con $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ unha matriz simétrica e denotamos por $A_{r,r}$ as súas submatrices diagonais

$$A(r,r) = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix},$$

cúmprese o seguinte:

1) A forma cuadrática ω é definida positiva, se, e só se, $\det(A_{r,r}) > 0$ para todo $r \in \{1, \dots, n\}$.
2) A forma cuadrática ω é definida negativa, se, e só se, $(-1)^r \det(A_{r,r}) > 0$ para todo $r \in \{1, \dots, n\}$.

8.1. Formas cuadráticas e formas bilineares. Forma polar

Definición 8.1.1. Chámase forma cuadrática sobre $\mathbb{R}^n$ a calquera aplicación $\omega : \mathbb{R}^n \to \mathbb{R}$ definida por

$$\omega(v) = v^t A v,\ \forall\, v \in \mathbb{R}^n$$

onde $A \in \mathcal{M}_{n\times n}(\mathbb{R})$.

Exemplo 8.1.2. A aplicación $\omega : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$\omega\begin{pmatrix} x \\ y \end{pmatrix} = 2xy,\ \forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$$

é unha forma cuadrática xa que

$$\omega\begin{pmatrix} x \\ y \end{pmatrix} = 2xy = (x, y)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Se tomamos as matrices

$$B = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix},\ C = \begin{pmatrix} 1 & 3 \\ -1 & -1 \end{pmatrix}$$

tense que

$$(x, y)\begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + 2xy - y^2 =$$

$$(x, y)\begin{pmatrix} 1 & 3 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Por tanto, é posíbel que dúas matrices distintas dean lugar á mesma forma cuadrática.

Nota 8.1.3. 1) Se $\omega : \mathbb{R}^n \to \mathbb{R}$ é unha forma cuadrática, tense que $\omega(\theta_{\mathbb{R}^n}) = 0$.

2) En xeral, se $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$, a expresión da forma cuadrática está dada por

$$\omega(v) = v^t Av = (v_1, \dots, v_n)A\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i,j=1}^{n} a_{ij}v_iv_j.$$

3) Se $\omega : \mathbb{R}^n \to \mathbb{R}$ é unha forma cuadrática,

$$\omega(\alpha v) = \alpha^2\omega(v)$$

para todo escalar α e todo vector v.

Definición 8.1.4. Sexa unha forma bilinear $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Chámase forma cuadrática asociada a f á aplicación $\omega : \mathbb{R}^n \to \mathbb{R}$ definida por

$$\omega_f(v) = v^t G_{C_n} v, \ \forall\, v \in \mathbb{R}^n$$

onde G_{C_n} é a matriz de Gram de f respecto á base canónica de $\mathbb{R}^n$.

Proposición 8.1.5. *Dada $\omega : \mathbb{R}^n \to \mathbb{R}$ unha forma cuadrática en $\mathbb{R}^n$, existe unha única forma bilinear simétrica $f_\omega : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ tal que $\omega_{f_\omega} = \omega$.*

Demostración. Definamos $f_\omega : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ como

$$f_\omega(u, v) = \frac{1}{2}(\omega(u + v) - w(u) - w(v)).$$

Entón, se $\forall\, u \in \mathbb{R}^n$ $w(u) = u^t Au$, tense que

$$f_\omega(u, v) = \frac{1}{2}((u + v)^t A(u + v) - u^t Au - v^t Av) = \frac{1}{2}(u^t Av + v^t Au) = \frac{1}{2}(u^t Av + u^t A^t v) =$$

$$u^t(\frac{1}{2}(A + A^t))v.$$

Logo, $f_\omega : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ é unha forma bilinear simétrica xa que a matriz $\frac{1}{2}(A + A^t)$ é simétrica.

A forma cuadrática asociada a f_ω é

$$\omega_{f_\omega}(v) = v^t G_{C_n} v, \ \forall\, v \in \mathbb{R}^n.$$

Como $G_{C_n} = \frac{1}{2}(A + A^t)$, temos que $\omega_{f_\omega} = \omega$.

Por outra banda, se existise unha forma bilinear simétrica $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ tal que $\omega_g = \omega$, temos que $f_\omega = g$, xa que, se G_{C_n} é a matriz de Gram de g respecto á base canónica de $\mathbb{R}^n$, usando o feito de que G_{C_n} é simétrica, obtemos as igualdades

$$f_\omega(u, v) = \frac{1}{2}(\omega(u + v) - w(u) - w(v)) = \frac{1}{2}(\omega_g(u + v) - w_g(u) - w_g(v)) =$$

$$\frac{1}{2}(u^t G_{C_n} v + v^t G_{C_n} u) = u^t G_{C_n} v = g(u, v).$$

□

Definición 8.1.6. A forma bilinear simétrica introducida na proposición anterior chámase forma polar da forma cuadrática $\omega : \mathbb{R}^n \to \mathbb{R}$.

Exemplo 8.1.7. Para a forma cuadrática $\omega : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$\omega\begin{pmatrix} x \\ y \end{pmatrix} = 2xy,\ \forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$$

ou, o que é o mesmo,

$$\omega\begin{pmatrix} x \\ y \end{pmatrix} = (x,y)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},\ \forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2,$$

tense que $f_\omega : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ está dada por

$$f_\omega(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}) = (x_1, x_2)(\frac{1}{2}(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^t))\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} =$$

$$(x_1, x_2)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

8.2. Matriz asociada a unha forma cuadrática. Diagonalización por congruencia

Teorema 8.2.1. *Sexa $\omega : \mathbb{R}^n \to \mathbb{R}$ unha forma cuadrática. Existe unha única matriz simétrica $M(\omega)$ tal que*

$$\omega(v) = v^t M(\omega) v,\ \forall\, v \in \mathbb{R}^n$$

Demostración. Se A é unha matriz cadrada de orde n tal que

$$\omega(v) = v^t A v,\ \forall\, v \in \mathbb{R}^n$$

é suficiente tomar

$$M(\omega) = \frac{1}{2}(A + A^t).$$

Entón, como $v^t A v = v^t A^t v\ \forall\, v \in \mathbb{R}^n$,

$$v^t M(\omega) v = v^t(\frac{1}{2}(A + A^t))v = \frac{1}{2}(v^t A v + v^t A^t v) = v^t A v = \omega(v).$$

A forma polar asociada a ω será

$$f_\omega(u, v) = u^t M(\omega) v.$$

Daquela, se existise outra matriz B simétrica tal que $\omega(v) = v^t B v,\ \forall\, v \in \mathbb{R}^n$, obteriamos a igualdade

$$f_\omega(u, v) = u^t M(\omega) v = u^t B v,\ \forall\, u, v \in \mathbb{R}^n.$$

Logo, como consecuencia, $M(\omega) = B$.

□

Definición 8.2.2. Dada unha forma cuadrática $\omega : \mathbb{R}^n \to \mathbb{R}$ chamaremos matriz asociada a ω, respecto á base canónica, á única matriz simétrica $M(\omega)$ tal que $\omega(u) = u^t M(\omega) u,\ \forall\, v \in \mathbb{R}^n$.

Exemplo 8.2.3. Sexa a forma cuadrática $\omega : \mathbb{R}^4 \to \mathbb{R}$ dada por

$$\omega\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = (x,y,z,t)\begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix},\quad \forall \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4.$$

Como neste caso

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

a matriz asociada a ω é

$$M(\omega) = \frac{1}{2}(A + A^t) = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{2} & 1 & 1 & 1 \\ \frac{1}{2} & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

8.2.4. En todas as definicións anteriores tomamos coordenadas respecto á base canónica de $\mathbb{R}^n$. Se cambiamos a base, a expresión da forma cuadrática tamén o fai. Sexa B outra base de $\mathbb{R}^n$ e chamemos P_{B,C_n} á matriz de cambio de base de B á base canónica de $\mathbb{R}^n$. Como sabemos, $\forall\, v \in \mathbb{R}^n$,

$$v = P_{B,C_n} v_B$$

sendo v_B o vector de coordenadas de v na base B. Entón:

$$\omega(v) = v^t M(\omega) v = v_B^t P_{B,C_n}^t M(\omega) P_{B,C_n} v_B.$$

A matriz $P_{B,C_n}^t M(\omega) P_{B,C_n}$ chámase matriz asociada a ω respecto á base B e denótase por $M(\omega)_B$. Evidentemente é simétrica e

$$M(\omega)_{C_n} = M(\omega).$$

Logo, a expresión da forma cuadrática na base B é

$$\omega(v) = v_B^t M(\omega)_B v_B.$$

Cando a matriz $M(\omega)_B$ é diagonal dirase que a base B é ortogonal respecto a ω. Neste caso tense que, se $v_B^t = (x_1, \dots, x_n)$,

$$\omega(v) = v_B^t M(\omega)_B v_B =$$

$$(x_1, x_2, \dots, x_n) \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & . & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1^2 + a_2 x_2^2 + \cdots + a_n x_n^2.$$

Se a base B ten como vectores a

$$B = \{e_1, \dots, e_n\}$$

e

$$e_{iB} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

onde o 1 atópase na compoñente i-ésima, podemos concluír que

$$w(e_i) = (0, \dots, 1, \dots, 0) \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = a_i.$$

Logo, se $M(\omega)_B$ é diagonal, pódese expresar da seguinte forma:

$$M(\omega)_B = \begin{pmatrix} \omega(e_1) & 0 & \cdots & 0 \\ 0 & \omega(e_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega(e_n) \end{pmatrix}.$$

O resultado que se presenta a continuación axudaranos a clasificar as formas cuadráticas.

Teorema 8.2.5. *Toda matriz simétrica $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ é congruente cunha matriz diagonal. Como consecuencia, toda forma cuadrática $\omega : \mathbb{R}^n \to \mathbb{R}$ admite unha base ortogonal.*

Demostración. A demostración consiste en usar que podemos transformar A nunha matriz diagonal D realizando certas operacións elementais por filas e as mesmas operacións elementais por columnas. Se $P = C_1 \cdots C_r$ é o produto de todas as operacións elementais por columnas usadas no proceso, tense que $P^t = C_r^t \cdots C_1^t$ é o produto de todas as operacións elementais realizadas por filas e grazas a isto:

$$P^tAP = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} = D.$$

Como a matriz P é invertíbel, as súas columnas forman unha base B de $\mathbb{R}^n$ e tense que $P_{B,C_n} = P$. Por tanto, se A fose a matriz asociada á forma cuadrática $\omega : \mathbb{R}^n \to \mathbb{R}$, B sería unha base ortogonal para ω.

□

Nota 8.2.6. No resultado anterior comprobouse que dada unha forma cuadrática ω podemos atopar unha base ortogonal para ela. Isto é equivalente a atopar unha matriz invertíbel P tal que $P^tM(\omega)P$ sexa diagonal. Obviamente, as columnas de P forman a base buscada.

Desde o punto de vista práctico temos dous camiños. O primeiro está baseado no uso da diagonalización ortogonal da matriz $M(\omega)$. Xa que $M(\omega)$ é simétrica, polo visto no capítulo anterior, existen unha matriz ortogonal P e unha matriz diagonal D tales que $P^tM(\omega)P = D$. Equivalentemente,

$$M(\omega) = PDP^t,$$

sendo os elementos da diagonal de D os autovalores $\lambda_1, \lambda_2, \cdots, \lambda_n$ de $M(\omega)$, contados cada un segundo a súa multiplicidade. Por tanto,

$$\omega(v) = (P^tv)^tD(P^tv),$$

e, se denotamos como u o vector P^tv, tense que

$$\omega(v) = \lambda_1u_1^2 + \lambda_2u_2^2 + \cdots + \lambda_nu_n^2$$

sendo u_i as compoñentes do vector u.

A outra forma de buscar a base ortogonal coñécese como diagonalización por congruencia e está baseada na utilización das mesmas operacións elementais por filas e por columnas (ver a demostración da última proposición). O esquema que se utilizará para efectuar dita diagonalización é o seguinte:

- Achamos a matriz asociada a ω:

$$M(\omega) = \frac{1}{2}(A + A^t).$$

- Construímos a matriz ampliada $(M(\omega)|I_n)$.
- Realizamos operacións elementais por filas en $(M(\omega)|I_n)$ e, de forma simultánea, as mesmas operacións elementais por columnas en $(M(\omega)|I_n)$, ata obter a matriz diagonal D. Ao final do proceso a matriz $(M(\omega)|I_n)$ converteuse en $(D|Q)$ con D diagonal e Q a matriz na que almacenamos as operacións por filas do proceso.
- Poñemos $P^t = Q$. Deste xeito $P^t M(\omega) P = D$ e as columnas de P forman a base ortogonal respecto a ω.

Obviamente, neste caso a matriz P é invertíbel pero en xeral non é ortogonal. Doutra banda, os elementos da diagonal de D non son os autovalores de $M(\omega)$. De todos os xeitos tense que

$$\omega(v) = (P^{-1}v)^t D(P^{-1}v),$$

e, se denotamos como u o vector $P^{-1}v$, obtemos a igualdade

$$\omega(v) = d_1 u_1^2 + d_2 u_2^2 + \cdots + d_n u_n^2$$

onde $d_1, d_2, \cdots, d_n$ son os elementos da diagonal de D, contados cada un segundo a súa multiplicidade, e u_i son as compoñentes do vector u.

Exemplo 8.2.7. Sexa a forma cuadrática $\omega : \mathbb{R}^4 \to \mathbb{R}$ dada por

$$\omega\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = (x, y, z, t)\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \quad \forall \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4.$$

Para calcular unha base ortogonal respecto de ω procederemos do seguinte xeito:

$$M(\omega) = \frac{1}{2}(A + A^t) = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

xa que

$$A = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

é simétrica.

Realicemos agora as operacións elementais que nos conduzan á matriz diagonal D:

$$\left(\begin{array}{c|c} A & I_4 \end{array} \right)$$

$$\xrightarrow{F_{14}(-1)} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{C_{14}(-1)} \left(\begin{array}{cccc|cccc} 2 & 0 & 0 & -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{F_{41}(\frac{1}{2})} \left(\begin{array}{cccc|cccc} 2 & 0 & 0 & -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{array}\right)$$

$$\overset{C_{41}(\frac{1}{2})}{\Longrightarrow}\left(\begin{array}{cccc|cccc} 2 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{array}\right)$$

$$\overset{F_{23}(1)}{\Longrightarrow}\left(\begin{array}{cccc|cccc} 2 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{array}\right)$$

$$\overset{C_{23}(1)}{\Longrightarrow}\left(\begin{array}{cccc|cccc} 2 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 2 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{array}\right)$$

$$\overset{F_{32}(-\frac{1}{2})}{\Longrightarrow}\left(\begin{array}{cccc|cccc} 2 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 2 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{array}\right)$$

$$\overset{C_{32}(-\frac{1}{2})}{\Longrightarrow}\left(\begin{array}{cccc|cccc} 2 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 2 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{array}\right).$$

Logo

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

e

$$P^tAP = D$$

sendo

$$P^t = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

8.3. Clasificación de formas cuadráticas

Definición 8.3.1. Sexa a forma cuadrática $\omega : \mathbb{R}^n \to \mathbb{R}$. Diremos que

1) ω é definida positiva se $\omega(v) > 0,\ \forall\, v \in \mathbb{R}^n,\ v \neq \theta_{\mathbb{R}^n}$.
2) ω é definida negativa se $\omega(v) < 0,\ \forall\, v \in \mathbb{R}^n,\ v \neq \theta_{\mathbb{R}^n}$.
3) ω é semidefinida positiva se $\omega(v) \geq 0,\ \forall\, v \in \mathbb{R}^n$.
4) ω é semidefinida negativa se $\omega(v) \leq 0,\ \forall\, v \in \mathbb{R}^n$.
5) ω é indefinida se existen dous vectores non nulos u, v de $\mathbb{R}^n$ tales que $\omega(u) < 0$ e $\omega(v) > 0$.

Exemplo 8.3.2. 1) Supoñamos que $(\mathbb{R}^n, <\ >)$ é un espacio vectorial euclídeo. O produto escalar definido en $\mathbb{R}^n$ é unha forma bilinear simétrica e a súa forma cuadrática asociada $\omega_{<\ >}$ está dada por

$$\omega_{<\ >}(v) = < v, v > .$$

Como para o produto escalar cúmprese que $< v, v > > 0$, se v é non nulo, temos que a forma cuadrática $\omega_{<\ >}$ é definida positiva.

2) A forma cuadrática

$$\omega\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = x^2+y^2+z^2+t^2$$

é definida positiva e

$$\omega\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = -x^2-y^2-z^2-t^2$$

é definida negativa.

3) A forma cuadrática

$$\omega\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = -x^2-y^2+z^2+t^2$$

é non definida xa que

$$\omega\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = -1, \ \ \omega\begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} = 1.$$

Definición 8.3.3. Sexa $\omega : \mathbb{R}^n \to \mathbb{R}$ unha forma cuadrática e f_ω a súa forma polar asociada. Como sabemos, se $M(\omega)$ é a matriz asociada a ω, temos que

$$f_\omega(u, v) = u^t M(\omega) v.$$

Chámase radical da forma cuadrática ω ao subespazo

$$N(\omega) = \{v \in \mathbb{R}^n, \mid \ f_\omega(v, u) = 0, \forall\, u \in \mathbb{R}^n\}.$$

Chámase rango de ω ao rango de $M(\omega)$.

Dise que ω é non dexenerada se $N(\omega) = \{\theta_{\mathbb{R}^n}\}$. En caso contrario dirase que é dexenerada. Nótese que é moi sinxelo probar que ω é non dexenerada, se, e só se, $\operatorname{rang}(M(\omega)) = n$, isto é, se, e só se, $M(\omega)$ é invertíbel

Teorema 8.3.4. (Lei de inercia de Sylvester) *Sexa $\omega : \mathbb{R}^n \to \mathbb{R}$ unha forma cuadrática. Se D_1, D_2 son matrices diagonais asociadas a ω respecto a distintas bases, entón o número de elementos positivos e negativos das súas diagonais principais é o mesmo.*

Demostración. É evidente que o número de elementos positivos máis o de negativos da diagonal principal de D_1 coincide co número de elementos positivos máis o de negativos da diagonal principal de D_2 e co $\operatorname{rang}(\omega)$. Por tanto, só temos que demostrar que o número de positivos é coincidente. Sexan B_1 e B_2 dúas bases de $\mathbb{R}^n$ para as cales as matrices asociadas a ω son diagonais. Supoñamos que teñen na diagonal r e s elementos positivos, respectivamente. Probaremos que $s = r$.

Se

$$B_1 = \{e_1, \dots, e_r, e_{r+1}, \dots, e_n\}$$

e

$$B_2 = \{u_1, \dots, u_s, u_{s+1}, \dots, u_n\}$$

podemos supoñer que as matrices asociadas ás bases B_1 y B_2 son da forma (en caso contrario, só temos que reordenar os vectores das bases)

$$D_1 = \begin{pmatrix} \omega(e_1) & 0 & \cdots & 0 \\ 0 & . & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega(e_n) \end{pmatrix}, \; D_2 = \begin{pmatrix} \omega(u_1) & 0 & \cdots & 0 \\ 0 & . & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega(u_n) \end{pmatrix},$$

onde

$$\omega(e_1) > 0, \dots, \omega(e_r) > 0, \omega(e_{r+1}) \leq 0, \dots, \omega(e_n) \leq 0$$

$$\omega(u_1) > 0, \dots, \omega(u_s) > 0, \omega(u_{s+1}) \leq 0, \dots, \omega(u_n) \leq 0.$$

Consideremos os subespazos

$$U = \langle\{e_1, \dots, e_r\}\rangle, \quad W = \langle\{u_{s+1}, \dots, u_n\}\rangle.$$

Cúmprese que $U \cap W = \{\theta_{\mathbb{R}^n}\}$ e, por tanto, o sistema

$$L = \{e_1, \dots, e_r, u_{s+1}, \dots, u_n\}$$

é libre. O sistema L ten $r+n-s$ vectores e ademais $r+n-s \leq n$. Entón $r \leq s$. Razoando dunha forma simétrica, obteriamos que $s \leq r$ e isto implica que $s = r$. □

O teorema anterior permite introducir un invariante das formas cuadráticas coñecido co nome de sinatura.

Definición 8.3.5. Sexa $\omega : \mathbb{R}^n \to \mathbb{R}$ unha forma cuadrática. Sexa B unha base ortogonal respecto de ω. Chámase sinatura de ω ao par:

$$\mathrm{sg}(\omega) = (p, q),$$

onde p, q denotan, respectivamente, o número de elementos positivos e negativos da matriz diagonal asociada a ω na base B.

Segundo o visto na demostración do teorema anterior, cúmprese que

$$\mathrm{rang}\,(\omega) = p + q.$$

Logo, ω é dexenerada, se, e só se, $p + q < n$ e non dexenerada, se, e só se, $p + q = n$.

Exemplo 8.3.6. Dada a forma cuadrática do Exemplo 8.2.7

$$\omega\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = (x, y, z, t)\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \quad \forall \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4,$$

demostramos que existe unha base ortogonal para ω dada por

$$B = \{e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ -1/2 \\ 1/2 \\ 0 \end{pmatrix}, e_4 = \begin{pmatrix} 1/2 \\ 0 \\ 0 \\ 1/2 \end{pmatrix}\}$$

para a cal

$$M(\omega)_B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \end{pmatrix}.$$

Por tanto, $\mathrm{sg}(\omega) = (2, 2)$ e a forma cuadrática é non dexenerada, xa que $2 + 2 = 4$.

Como consecuencia inmediata da lei de inercia de Sylvester, tense o seguinte resultado.

Corolario 8.3.7. *Sea $\omega : \mathbb{R}^n \to \mathbb{R}$ una forma cuadrática. Verifícase o seguinte:*

1) *A forma cuadrática ω é definida positiva, se, e só se, $\mathrm{sg}(\omega) = (n, 0)$.*
2) *A forma cuadrática ω é definida negativa, se, e só se, $\mathrm{sg}(\omega) = (0, n)$.*
3) *A forma cuadrática ω é semidefinida positiva, se, e só se, $\mathrm{sg}(\omega) = (r, 0)$ con $r < n$.*
4) *A forma cuadrática ω é semidefinida negativa, se, e só se, $\mathrm{sg}(\omega) = (0, s)$ con $s < n$.*
5) *A forma cuadrática ω é indefinida, se, e só se, $\mathrm{sg}(\omega) = (r, s)$ con r, s non nulos.*

Exemplo 8.3.8. Sexa a forma cuadrática $\omega : \mathbb{R}^4 \to \mathbb{R}$ dada por

$$\omega\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = (x, y, z, t)\begin{pmatrix} 1 & 1 & 2 & 0 \\ 1 & 3 & 2 & 0 \\ 2 & 2 & 7 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \quad \forall \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4.$$

No noso caso

$$M(\omega) = \frac{1}{2}(A + A^t) = \begin{pmatrix} 1 & 1 & 2 & 0 \\ 1 & 3 & 2 & 0 \\ 2 & 2 & 7 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

xa que

$$A = \begin{pmatrix} 1 & 1 & 2 & 0 \\ 1 & 3 & 2 & 0 \\ 2 & 2 & 7 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

é simétrica.

Realicemos agora as operacións elementais que nos conduzan á matriz diagonal D:

$$\left(\begin{array}{c|c} A & I_4 \end{array}\right)$$

$$\overset{F_{13}(-2),\, F_{12}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|cccc} 1 & 1 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\overset{C_{13}(-2),\, C_{12}(-1)}{\Longrightarrow} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Polo tanto,

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

e

$$P^t A P = D$$

sendo

$$P^t = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Como consecuencia, ω é definida positiva, xa que $\mathrm{sg}(\omega) = (4, 0)$.

Nota 8.3.9. Sexa $\omega : \mathbb{R}^n \to \mathbb{R}$ unha forma cuadrática. Sexa $M(\omega)$ a súa matriz asociada. Como esta matriz é simétrica, polo Teorema 7.3.2, sabemos que existe unha matriz ortogonal Q tal que $Q^t M(\omega) Q$ é diagonal. As columnas de Q dan lugar a unha base ortonormal do espazo $\mathbb{R}^n$, respecto ao produto escalar usual, formada por autovectores e, dado que $Q^t M(\omega) Q$ é diagonal, tamén dan lugar a unha base ortogonal respecto de ω. Os elementos da diagonal principal de $D = Q^t M(\omega) Q$ son os autovalores de $M(\omega)$.

Logo temos o seguinte resultado:

Corolario 8.3.10. *Sexa $\omega : \mathbb{R}^n \to \mathbb{R}$ unha forma cuadrática. Sexa $M(\omega)$ a súa matriz asociada. Cúmprese o seguinte:*

1) *A forma cuadrática ω é definida positiva, se, e só se, os autovalores de $M(\omega)$ son todos positivos.*
2) *A forma cuadrática ω é definida negativa, se, e só se, os autovalores de $M(\omega)$ son todos negativos.*
3) *A forma cuadrática ω é semidefinida positiva, se, e só se, os autovalores de $M(\omega)$ son todos non negativos sendo algún deles nulo.*
4) *A forma cuadrática ω é semidefinida negativa, se, e só se, os autovalores de $M(\omega)$ son todos non positivos sendo algún deles nulo.*
5) *A forma cuadrática ω é indefinida se existen autovalores de $M(\omega)$ positivos e negativos.*

Finalmente, mostraremos un criterio que tamén pode ser útil á hora de caracterizar as formas cuadráticas definidas positivas e negativas.

Teorema 8.3.11. (Criterio de Sylvester) *Sexa $\omega : \mathbb{R}^n \to \mathbb{R}$ unha forma cuadrática dada por $\omega(v) = v^t A v$ sendo $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ unha matriz simétrica. Denotemos por $A_{r,r}$ as súas submatrices diagonais*

$$A(r,r) = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}.$$

Verifícase o seguinte:

1) *A forma cuadrática ω é definida positiva, se, e só se, $\det(A_{r,r}) > 0$ para todo $r \in \{1, \dots, n\}$.*
2) *A forma cuadrática ω é definida negativa, se, e só se, $(-1)^r \det(A_{r,r}) > 0$ para todo $r \in \{1, \dots, n\}$.*

8.4. Problemas propostos

1. Considérase a forma $w : \mathbb{R}^3 \to \mathbb{R}$ definida por

$$\omega \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + 2xy + 4yz.$$

   a) Atope a matriz asociada a w respecto da base canónica de $\mathbb{R}^3$.

   b) Atope a matriz matriz asociada a w respecto da base

$$B = \{(1,0,1)^t, (0,1,2)^t, (1,-1,1)^t\}.$$

2. Considérase a matriz $A \in \mathcal{M}_{4\times 4}(\mathbb{R})$ dada por

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

   Se é posíbel, calcule unha matriz ortogonal P e unha matriz diagonal D tales que $P^t A P = D$.

3. Considérase a matriz $A \in \mathcal{M}_{4\times 4}(\mathbb{R})$ dada por

$$A = \begin{pmatrix} -2 & -1 & 0 & 2 \\ -1 & -2 & 0 & -2 \\ 0 & 0 & -3 & 0 \\ 2 & -2 & 0 & 1 \end{pmatrix}.$$

   a) Calcule unha matriz ortogonal P e unha matriz diagonal D tales que $P^tAP = D$.
   b) Calcule o rango e a signatura da forma cuadrática $\omega : \mathbb{R}^4 \to \mathbb{R}$ definida por $\omega(v) = v^tAv$, $\forall x \in \mathbb{R}^4$.

4. Considérase a matriz

$$A = \begin{pmatrix} -1 & -1 & 2 \\ -1 & -1 & -2 \\ 2 & -2 & 2 \end{pmatrix}.$$

   a) Determine unha matriz ortogonal P e unha matriz diagonal D tales que $AP = PD$.
   b) Sexa $\omega : \mathbb{R}^3 \to \mathbb{R}$ a forma cuadrática dada por

$$\omega(v) = v^t\, e^A\, v\,.$$

   Calcule o rango e a sinatura de ω.

5. Considérase a forma cuadrática $w : \mathbb{R}^4 \to \mathbb{R}$ definida por

$$\omega \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = (x, y, z, t) \begin{pmatrix} 1 & 2 & -2 & -3 \\ 0 & 1 & 2 & -2 \\ 2 & 0 & 1 & 2 \\ 3 & 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}.$$

   a) Atope a matriz asociada a w respecto da base canónica de $\mathbb{R}^4$.
   b) Atope o rango de w.

6. Sexa $w : \mathbb{R}^4 \to \mathbb{R}$ a forma cuadrática definida por

$$\omega \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = x^2 + z^2 + t^2 + yz + 2xt + yt.$$

   a) Atope a matriz asociada a w respecto da base canónica de $\mathbb{R}^4$.
   b) Calcule o rango e a sinatura de w.
   c) Atope unha base de $\mathbb{R}^4$ ortogonal respecto de w.

7. Diagonalize e clasifique a forma cuadrática real $w(v) = v^tAv$ sendo
   a)

$$A = \begin{pmatrix} 1 & 0 & 2 \\ -2 & 1 & 1 \\ 0 & 1 & 5 \end{pmatrix}.$$

   b)

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 2 \\ 2 & 2 & 7 \end{pmatrix}.$$

   c)

$$A = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

d)
$$A = \begin{pmatrix} -4 & -2 & 0 \\ -2 & -4 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

e)
$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ -1 & -2 & 5 \end{pmatrix}.$$

f)
$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

g)
$$A = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 8 & -6 & 4 \\ -1 & -6 & 6 & -6 \\ 1 & 4 & -6 & 15 \end{pmatrix}.$$

Ademais, calcule unha matriz invertíbel P tal que a matriz P^tAP sexa una matriz diagonal.

8. Para cada $\alpha \in \mathbb{R}$, considérase a forma cuadrática real $w_\alpha : \mathbb{R}^3 \to \mathbb{R}$ definida por
$$w_\alpha \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x, y, z) \begin{pmatrix} 2\alpha & 0 & 0 \\ 2\alpha & \alpha^2 & 0 \\ 0 & 2\alpha & 2\alpha \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$
   a) Atope a matriz asociada a w_α respecto da base canónica de $\mathbb{R}^3$.
   b) Clasifique w_α segundo os distintos valores do parámetro α, indicando o seu rango e sinatura.
9. Sexa $w_\alpha : \mathbb{R}^4 \to \mathbb{R}$ a forma cuadrática definida por
$$w_\alpha \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = -x^2 - 3y^2 - (\alpha^2 + 6)z^2 + (\alpha - 3)t^2 + 2xy + 4xz - 2xt - 8yz + 2yt + 4zt.$$
   a) Clasifique w_α segundo os distintos valores do parámetro α, calculando o seu rango e sinatura.
   b) Para $\alpha = 1$, ache unha base B de $\mathbb{R}^4$ respecto da cal a matriz asociada a w_1 sexa diagonal.

CAPíTULO 9

Valores singulares, pseudoinversas e mínimos cadrados

Este capítulo final está dedicado á introdución da teoría de descomposición en valores singulares (SVD), pseudoinversas e as súas aplicacións á solución de problemas de mínimos cadrados. A descomposición en valores singulares proporciona un método para calcular pseudoinversas e, como se verá máis adiante, para resolver problemas de mínimos cadrados de xeito óptimo. Este tipo de descomposición móstrase especialmente útil para calcular o rango dunha matriz debido, sobre todo, aos problemas que presenta o método de Gauss con respecto á coma flotante. De feito, o método baseado na SVD é moito máis efectivo para calcular o rango dunha matriz que os derivados da utilización da eliminación gaussiana e, por suposto, de calquera outro que pase polo uso de determinantes. Hoxe en día a SVD converteuse nunha importante ferramenta computacional.

Comezaremos introducindo a teoría de descomposición en valores singulares para unha matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Como a matriz A^tA é simétrica, podemos afirmar que todos os seus autovalores son reais. Agora ben, neste caso pódese demostrar que ademais de ser reais son todos maiores ou iguais que cero. Tendo en conta isto último, chámanse valores singulares da matriz A ás raíces cadradas positivas dos autovalores de A^tA. Logo, se

$$\lambda_1, \lambda_2, \dots, \lambda_n$$

son os autovalores de A^tA, contados cada un segundo a súa multiplicidade, tense que os valores singulares de A son:

$$\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}.$$

Denotaremos os valores singulares como $\sigma_1, \sigma_2, \cdots, \sigma_n$ e os ordenaremos de xeito que

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

Unha vez feito isto enunciarase o teorema de descomposición en valores singulares, o cal garante que para toda matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ existen matrices ortogonais $U \in \mathcal{M}_{p\times p}(\mathbb{R})$ e $V \in \mathcal{M}_{n\times n}(\mathbb{R})$ e unha matriz

$$\Sigma = \left(\begin{array}{ccc} D_r & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{array}\right) \in \mathcal{M}_{n\times p}(\mathbb{R}),$$

con

$$D_r = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{pmatrix} \in \mathcal{M}_{r\times r}(\mathbb{R}),$$

tales que

$$A = U\Sigma V^t,$$

sendo $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ os valores singulares non nulos. Así pois o rango de A coincide co número de valores singulares non nulos, contados cada un segundo a súa multiplicidade, e tamén coincide co rango de A^tA.

De xeito similar ao obtido ao traballar coa descomposición espectral da matriz A, pódese probar que, se o rango de A é r, podemos escribir A como suma de r matrices de rango un como segue:

$$A = U\Sigma V^t = \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t + \cdots + \sigma_r u_r v_r^t,$$

onde $u_1, u_2, \ldots, u_r$ e $v_1, v_2, \ldots, v_r$ son as r primeiras columnas das matrices U e V respectivamente.

O procedemento que se utilizará para calcular a SVD dunha matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ con rango r será o seguinte:

- As columnas de V tómanse como unha base ortonormal de $\mathbb{R}^n$ formada por autovectores asociados aos autovalores de A^tA ordenados de maior a menor.
- Se $V = (v_1|v_2|\cdots|v_n)$ e $U = (u_1|u_2|\cdots|u_p)$, como $A = U\Sigma V^t$, obtense que

  $$AV = U\Sigma,$$

  e, por tanto,

  $$Av_i = \sigma_i u_i,\ i \in \{1, \ldots, r\},$$

  ou, o que é o mesmo,

  $$u_i = \frac{1}{\sigma_i} A v_i,\ i \in \{1, \ldots, r\}.$$

  Se $p > r$, complétanse as columnas de U de xeito que

  $$\{u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_p\}$$

  sexa unha base ortonormal de $\mathbb{R}^p$. Os vectores

  $$\{u_{r+1}, \cdots, u_p\}$$

  pódense tomar como unha base ortonormal do subespazo dado polo conxunto de solucións do sistema linear homoxéneo $A^t x = \theta$. Isto pódese facer así xa que a dimensión dese subespazo de solucións é $p - r$ e todos os seus vectores son ortogonais aos vectores $\{u_1, u_2, \ldots, u_r\}$.

Se temos a descomposición en valores singulares dunha matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$, no segundo punto do capítulo verase que existe unha norma $||\ ||$ no espazo de matrices $\mathcal{M}_{p\times n}(\mathbb{R})$, tal que, se

$$A_k = \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t + \cdots + \sigma_k u_k v_k^t,$$

entón

$$||A - A_k|| = min\{||A - B||\ ;\ B \in \mathcal{M}_{p\times n}(\mathbb{R}),\ \text{rang}\,(B) = k\} = \sigma_{k+1}.$$

Logo, en certo modo A_k é a matriz que máis preto está a A de entre todas as de rango k que pertencen ao espazo vectorial $\mathcal{M}_{p\times n}(\mathbb{R})$. Esta matriz chamarase aproximación de rango k da matriz A.

A terceira parte do capítulo está dedicada a estudar a noción de pseudoinversa, ou inversa xeneralizada de Moore-Penrose, dunha matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Denotarase como A^+ e, cando a matriz A é cadrada e invertíbel, a pseudoinversa de A é exactamente a matriz inversa de A. Tamén veremos que, cando a matriz A^tA é invertíbel, entón

$$A^+ = (A^tA)^{-1}A^t.$$

Neste punto presentaranse dous métodos xerais para calcular a pseudoinversa. O primeiro baseado na factorización QR e o segundo na descomposición en valores singulares. Para o primeiro caso tense que, se $A = QR$, entón a pseudoinversa pódese calcular como:

$$A^+ = R^t(RR^t)^{-1}Q^t.$$

No segundo caso, tense que, se a SVD de A é

$$A = U\Sigma V^t$$

con valores singulares

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0,$$

pódese demostrar que

$$A^+ = V\Omega U^t$$

con

$$\Omega = \left(\begin{array}{ccc} T_r & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{array} \right) \in \mathcal{M}_{n\times p}(\mathbb{R})$$

onde

$$T_r = \begin{pmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_r} \end{pmatrix} = D_r^{-1} \in \mathcal{M}_{r\times r}(\mathbb{R}).$$

Logo,

$$A^+ = V\Omega U^t = \frac{1}{\sigma_1}v_1u_1^t + \frac{1}{\sigma_2}v_2u_2^t + \cdots + \frac{1}{\sigma_r}v_ru_r^t.$$

E. H. Moore
(Wikipedia Commons)

R. Penrose
(Cirone-Musi, Festival della Scienza/Wikipedia Commons)

A parte final do capítulo dedicarase aos problemas de mínimos cadrados e a resolución destes utilizando tanto a factorización QR como a teoría de pseudoinversas.

Para motivar o problema comezaremos expoñendo o que ocorre cando traballamos con sistemas de ecuacións lineares. Neste caso, dados $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ e $b \in \mathbb{R}^p$, supoñamos que temos o sistema $Ax = b$. Como sabemos, o anterior sistema é incompatíbel se b non pertence ao subespazo de $\mathbb{R}^n$ dado por

$$\{Aw \ | \ w \in \mathbb{R}^n\}.$$

Malia ser $Ax = b$ un sistema incompatíbel, a álxebra linear permite atopar unha solución aproximada óptima ou, o que é o mesmo, un vector $w \in \mathbb{R}^n$ tal que a distancia entre Aw e b sexa mínima.

Por outra banda, dado que o conxunto $\{Aw \ | \ w \in \mathbb{R}^n\}$ é un subespazo de $\mathbb{R}^p$, podemos dicir de xeito máis xenérico que, se U é un subespazo do espazo vectorial euclídeo $\mathbb{R}^p$ co produto escalar usual, $v \in U$ é a mellor aproximación dun vector $b \in \mathbb{R}^p$ por vectores de U non sentido dos mínimos cadrados, se

$$\|v - b\| = min\{\|u - b\|, \ u \in U\}.$$

Tendo en conta a última igualdade, enténdese que o nome de mínimos cadrados débese a que minimizar $\|u-b\|$ é equivalente a minimizar

$$\|u-b\|^2 = \sum_{i=1}^{p}(u_i - b_i)^2$$

sendo u_i e b_i as compoñentes dos vectores de coordenadas de u e b respecto da base canónica de $\mathbb{R}^p$.

Logo, se $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ e $b \in \mathbb{R}^p$ e temos o sistema de ecuacións lineares $Ax = b$, diremos que o vector $w \in \mathbb{R}^n$ é unha solución mínimo-cuadrática do sistema $Ax = b$, se Aw é a mellor aproximación de $b \in \mathbb{R}^n$ non sentido dos mínimos cadrados.

O primeiro resultado importante que se presentará é o que garante que, se U é un subespazo do espazo vectorial euclídeo $\mathbb{R}^p$ co produto escalar usual, entón $v \in U$ é a mellor aproximación de $b \in \mathbb{R}^p$ por vectores de U, se $< v-b, u >= 0$, $\forall\, u \in U$ ($v - b$ é ortogonal a todos os vectores de U). Este vector v chámase proxección ortogonal de b sobre o subespazo U e pódese garantir que existe sempre e que é único.

Das propiedades anteriores deducimos que, se $w, w' \in \mathbb{R}^n$ son solucións mínimo-cuadráticas do sistema $Ax = b$, tense a igualdade $Aw = Aw'$. Logo, $A(w - w') = \theta$, equivalentemente, $w - w' \in \mathrm{Ker}(f_A)$. Entón, se w é unha solución mínimo-cuadrática de $Ax = b$, o conxunto global de solucións do problema de mínimos cadrados tamén ven dado por

$$S = w + \mathrm{Ker}(f_A) = w + \mathrm{Ker}(f_{A^tA}).$$

Tamén, se $A \in \mathcal{M}_{p\times n}(\mathbb{R})$, $b \in \mathbb{R}^p$ e temos o sistema de ecuacións lineares $Ax = b$, pódese construír un novo sistema

$$A^tAx = A^tb$$

que se coñece co nome de sistema de ecuacións normais do sistema $Ax = b$. O realmente interesante deste sistema é que o conxunto de solucións mínimo-cuadráticas do sistema $Ax = b$ coincide co conxunto de solucións do sistema de ecuacións normais $A^tAx = A^tb$. Logo, resolver o problema de mínimos cadrados para $Ax = b$ é equivalente a resolver o sistema de ecuacións lineares dado polas súas ecuacións normais.

Evidentemente, se A^tA é invertíbel, S só ten un elemento w dado por

$$w = (A^tA)^{-1}A^tb = A^+b$$

xa que neste caso $A^+ = (A^tA)^{-1}A^t$.

Ademais verase que a factorización QR proporciona un método para resolver este problema xa que, se $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ é unha matriz, tense que w é solución mínimo-cuadrática para $Ax = b$, se, e só se, w é solución do sistema

$$Rx = Q^tb$$

onde R e Q son as matrices que dan lugar á factorización QR de A.

De xeito máis xeral demostrarase que, nas condicións do parágrafo anterior, pódese afirmar que o conxunto de solucións do problema de mínimos cadrados está dado por

$$S = A^+b + \mathrm{Ker}(f_A) = A^+b + \mathrm{Ker}(f_{A^tA}).$$

Este último método de solución ten como vantaxe, sobre o baseado na utilización das ecuacións normais, que é menos sensíbel aos problemas de estabilidade e que en xeral dá lugar a unha técnica computacional mellor. Como exemplo, finalizarase o capítulo aplicando os métodos de mínimos cadrados á solución de problemas de axuste polinómico.

En canto á notación, neste capítulo $< \, , \, >$ denotará o produto escalar usual de $\mathbb{R}^m$ e $\|v\|$ o valor da súa norma inducida nun vector de $\mathbb{R}^m$.

9.1. Descomposición en valores singulares

Supoñamos que $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Como a matriz A^tA é simétrica, grazas ao Teorema 7.3.1, podemos afirmar que todos os seus autovalores son reais. Ademais cúmprense outras propiedades interesantes que debemos ter en conta non que segue.

Proposición 9.1.1. *Se* $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ *entón,*

$$\mathrm{Ker}(f_A) = \mathrm{Ker}(f_{A^tA}), \quad \mathrm{rang}\,(A) = \mathrm{rang}\,(A^tA) = \mathrm{rang}\,(AA^t).$$

Demostración. Se $v \in \mathbb{R}^n$ é tal que $v \in \mathrm{Ker}(f_A)$ tense que $Av = \theta_{\mathbb{R}^p}$ e, como consecuencia, $A^tAv = \theta_{\mathbb{R}^n}$. Isto implica que $v \in \mathrm{Ker}(f_{A^tA})$. No caso de que $v \in \mathrm{Ker}(f_{A^tA})$ tense que $A^tAv = \theta_{\mathbb{R}^n}$. Entón

$$v^tA^tAv = 0 \Longleftrightarrow (Av)^tAv = 0 \Longleftrightarrow < Av, Av >= 0 \Longleftrightarrow Av = \theta_{\mathbb{R}^p}$$

Logo, $v \in \mathrm{Ker}(f_A)$. Finalmente,

$$\mathrm{rang}\,(A) = n - \dim_{\mathbb{R}}(\mathrm{Ker}(f_A)) = n - \dim_{\mathbb{R}}(\mathrm{Ker}(f_{A^tA})) = \mathrm{rang}\,(A^tA)$$

e

$$\mathrm{rang}\,(A) = \mathrm{rang}\,(A^t) = \mathrm{rang}\,((A^t)^tA^t)) = \mathrm{rang}\,(AA^t).$$

□

Proposición 9.1.2. *Sexa* $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. *Se* $\lambda \in \mathrm{Sp}(A^tA)$, $\lambda \geq 0$.

Demostración. Sexa $\lambda \in \mathrm{Sp}(A^tA)$ e $v \in \mathbb{R}^n$ un autovector asociado a λ. Entón

$$\|Av\|^2 = (Av)^tAv = v^tA^tAv = \lambda v^tv = \lambda\|v\|^2,$$

e, como consecuencia,

$$\lambda = \frac{\|Av\|^2}{\|v\|^2} \geq 0.$$

□

Definición 9.1.3. Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Chámanse valores singulares da matriz A ás raíces cadradas positivas dos autovalores de A^tA. Logo, se

$$\lambda_1, \lambda_2, \ldots, \lambda_n$$

son os autovalores de A^tA, contados cada un segundo a súa multiplicidade, tense que os valores singulares de A son:

$$\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}.$$

Denotaremos os valores singulares como $\sigma_1, \sigma_2, \cdots, \sigma_n$ e escribirémolos en orde decrecente

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

Exemplo 9.1.4. se

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

tense que

$$A^tA = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

entón

$$p_{A^tA}(x) = \det(A^tA - xI_3) = \begin{vmatrix} 1-x & 0 & 1 \\ 0 & 4-x & 0 \\ 1 & 0 & 1-x \end{vmatrix} = (4-x)\begin{vmatrix} 1-x & 1 \\ 1 & 1-x \end{vmatrix}$$

$$= (4-x)\begin{vmatrix} 2-x & 1 \\ 2-x & 1-x \end{vmatrix} = (4-x)\begin{vmatrix} 2-x & 1 \\ 0 & -x \end{vmatrix} = -(4-x)(2-x)x$$

e isto implica que $\mathrm{Sp}(A^tA) = \{0,2,4\}$. Como consecuencia, obtemos que os valores singulares son

$$\sigma_1 = \sqrt{4} = 2 \geq \sigma_2 = \sqrt{2} \geq \sigma_3 = \sqrt{0} = 0 \geq 0.$$

Proposición 9.1.5. *Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Cúmprese que A e A^t teñen os mesmos valores singulares non nulos.*

Demostración. Supoñamos que λ é un autovalor non nulo de A^tA e $v \in \mathbb{R}^n$ un autovector asociado a λ. Entón $A^tAv = \lambda v$ e isto implica que

$$(AA^t)Av = A(A^tAv) = A(\lambda v) = \lambda Av$$

ou, o que é o mesmo, λ é un autovalor non nulo de AA^t e $Av \in \mathbb{R}^p$ un autovector asociado a λ. Nótese que se $Av = \theta_{\mathbb{R}^p}$ tense que

$$\theta_{\mathbb{R}^n} = A^tAv = \lambda v$$

e, como consecuencia, λ sería nulo (cousa imposíbel xa que supuxemos que non o era).

Deste xeito temos probado que se λ é un autovalor non nulo de A^tA entón λ é un autovalor de AA^t. De xeito similar pódese comprobar que se λ é un autovalor non nulo de AA^t tamén o é de A^tA. □

Nota 9.1.6. Grazas ao resultado anterior está claro que, se $p \neq n$, unha das matrices A, A^t ten máis valores singulares nulos que a outra xa que $A^tA \in \mathcal{M}_{p\times p}(\mathbb{R})$ e $A^tA \in \mathcal{M}_{n\times n}(\mathbb{R})$.

Tamén é importante resaltar que, se $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ é unha matriz simétrica, tense que $A^tA = A^2$ e, por tanto, os valores singulares de A coinciden cos valores absolutos dos seus autovalores.

Teorema 9.1.7. (Descomposición en valores singulares) *Para toda matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ existen matrices ortogonais $U \in \mathcal{M}_{p\times p}(\mathbb{R})$ e $V \in \mathcal{M}_{n\times n}(\mathbb{R})$ e unha matriz*

$$\Sigma = \left(\begin{array}{c|c} D_r & \Theta \\ \hline \Theta & \Theta \end{array}\right) \in \mathcal{M}_{p\times n}(\mathbb{R}),$$

con

$$D_r = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{pmatrix} \in \mathcal{M}_{r\times r}(\mathbb{R}),$$

tales que

$$A = U\Sigma V^t,$$

sendo $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ os valores singulares non nulos de A.

Demostración. Supoñamos que o rango de A é r (isto implica que ten r valores singulares non nulos contados cada un segundo a súa multiplicidade alxébrica). O procedemento que se utilizará para calcular a SVD de A será o seguinte:

- As columnas de V tómanse como unha base ortonormal de $\mathbb{R}^n$ formada por autovectores asociados aos autovalores de A^tA ordenados de maior a menor.

- Se $V = (v_1|v_2|\cdots|v_n)$ e $U = (u_1|u_2|\cdots|u_p)$, como se terá que cumprir que $A = U\Sigma V^t$, obtense que
$$AV = U\Sigma,$$
e, por tanto,
$$Av_i = \sigma_i u_i,\ i \in \{1, \dots, r\},$$
ou, o que é o mesmo,
$$u_i = \frac{1}{\sigma_i} Av_i,\ i \in \{1, \dots, r\}.$$
Se $p > r$, complétanse as columnas de U de xeito que
$$\{u_1, \dots, u_r, u_{r+1}, \dots, u_p\}$$
sexa unha base ortonormal de $\mathbb{R}^p$. Os vectores
$$\{u_{r+1}, \dots, u_p\}$$
pódense tomar como unha base ortonormal do subespazo dado polo conxunto de solucións do sistema linear homoxéneo $A^t x = \theta_{\mathbb{R}^n}$. Isto pódese facer así xa que a dimensión dese subespazo de solucións é $p - r$ e todos os seus vectores son ortogonais aos vectores $\{u_1, u_2, \dots, u_r\}$.
En efecto: se tomamos $u_i \in \{u_1, u_2, \dots, u_r\}$ e $u_k \in \{u_{r+1}, \dots, u_p\}$ tense que
$$< u_i, u_k > = < \frac{1}{\sigma_i} Av_i, u_k > = \frac{1}{\sigma_i} < Av_i, u_k > = \frac{1}{\sigma_i} < v_i, A^t u_k > = \frac{1}{\sigma_i} < v_i, \theta_{\mathbb{R}^n} > = 0,$$
e, como consecuencia, u_i e u_k son ortogonais.
Por outra banda, $\{u_1, u_2, \dots, u_r\}$ forman un sistema ortonormal xa que se tomamos $u_i \in \{u_1, u_2, \dots, u_r\}$
$$\begin{aligned} < u_i, u_i > &= < \frac{1}{\sigma_i} Av_i, \frac{1}{\sigma_i} Av_i > = \frac{1}{\sigma_i^2} < Av_i, Av_i > = \frac{1}{\lambda_i} < v_i, A^t A v_i > \\ &= \frac{1}{\lambda_i} < v_i, \lambda_i v_i > = \frac{\lambda_i}{\lambda_i} < v_i, v_i > = \|v_i\|^2 = 1, \end{aligned}$$
e, se tomamos dous vectores distintos $u_i, u_k \in \{u_1, u_2, \dots, u_r\}$,
$$\begin{aligned} < u_i, u_k > &= < \frac{1}{\sigma_i} Av_i, \frac{1}{\sigma_k} Av_k > = \frac{1}{\sigma_i \sigma_k} < Av_i, Av_k > = \frac{1}{\sigma_i \sigma_k} < v_i, A^t A v_k > \\ &= \frac{1}{\sigma_i \sigma_k} < v_i, \lambda_k v_k > = \frac{\lambda_k}{\sigma_i \sigma_k} < v_i, v_k > = 0. \end{aligned}$$

□

Definición 9.1.8. Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. A descomposición $A = U\Sigma V^t$ obtida no teorema anterior chámase descomposición en valores singulares da matriz A. Denótase como SVD(A) (SVD son as iniciais da súa tradución ao inglés: Singular Value Decomposition).

Exemplo 9.1.9. Neste exemplo calcularemos SVD(A) para a matriz
$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \sqrt{3} & 0 \end{pmatrix}.$$
A matriz $A^t A$ está dada por
$$A^t A = \begin{pmatrix} 1 & 0 & \sqrt{3} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \sqrt{3} & 0 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}.$$

Entón, $\mathrm{Sp}(A^tA) = \{4, 1\}$ e isto implica que os valores singulares de A son

$$\sigma_1 = \sqrt{4} = 2 > \sigma_2 = \sqrt{1} = 1.$$

Como consecuencia do anterior obtense que a matriz Σ será

$$\Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Pasemos a calcular a matriz V. Os subespazos propios asociados aos autovalores $\lambda_1 = 4$ e $\lambda_2 = 1$ son:

$$V(4) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid (A^tA - 4I_2) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid \begin{pmatrix} 0 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = 0 \right\} = \langle \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \rangle,$$

$$V(1) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid (A^tA - I_2) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x = 0 \right\} = \langle \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \rangle.$$

Deste xeito,

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

é unha base de $\mathbb{R}^2$ formada por autovectores. Como os dous vectores son unitarios e ortogonais entre si, tamén forman unha base ortonormal de $\mathbb{R}^2$ (non é necesario aplicar o procedemento de ortonormalización de Gram-Schmidt) o que implica que podemos tomar

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

e $V = I_2$.

A matriz $U \in \mathcal{M}_{3\times 3}(\mathbb{R})$ constrúese do seguinte xeito. A primeira columna está dada por

$$u_1 = \frac{1}{2} A v_1 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \sqrt{3} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ \sqrt{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{\sqrt{3}}{2} \end{pmatrix}.$$

Similarmente, a segunda columna de U calcúlase como

$$u_2 = \frac{1}{2} A v_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \sqrt{3} & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Como o rango da matriz A é $r = 2$ e $2 < p = 3$, o vector u_3 debe obterse grazas a determinación dunha base do conxunto de solucións S do sistema linear homoxéneo

$$A^t \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \theta_{\mathbb{R}^3}.$$

Agora ben, como

$$A^t \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \left. \begin{aligned} x + \sqrt{3}z = 0 \\ y = 0 \end{aligned} \right\},$$

tense que

$$S = \left\{ \begin{pmatrix} x \\ 0 \\ \frac{-x}{\sqrt{3}} \end{pmatrix} \mid x \in \mathbb{R} \right\}$$

e, desta forma, unha base de S está dada polo vector

$$w = \begin{pmatrix} \sqrt{3} \\ 0 \\ -1 \end{pmatrix}.$$

Tendo en conta isto, tomamos a terceira columna de U como

$$u_3 = \frac{w}{\|w\|} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix}$$

e U é a matriz dada por

$$U = \begin{pmatrix} \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 1 & 0 \\ \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \end{pmatrix}.$$

Obviamente se calculamos $U\Sigma V^t$ obtemos A xa que

$$U\Sigma V^t = \begin{pmatrix} \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 1 & 0 \\ \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \sqrt{3} & 0 \end{pmatrix}.$$

Nota 9.1.10. Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. De xeito similar ao obtido coa descomposición espectral da matriz A, pode probarse que, se o rango de A é r, utilizando a descomposición en valores singulares de A podemos escribir A como suma de r matrices de rango un como segue:

$$A = U\Sigma V^t = \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t + \cdots + \sigma_r u_r v_r^t,$$

onde $u_1, u_2, \ldots, u_r$ e $v_1, v_2, \ldots, v_r$ son as r primeiras columnas das matrices U e V respectivamente.

9.2. Aproximacións de rango k

O resultado principal desta sección dinos que se aproximamos unha matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ polas primeiras k compoñentes da súa descomposición en valores singulares, pérdese unha cantidade de información da orde do valor singular colocado no posto $k+1$. Esta aproximación é de especial utilidade no tratamento de matrices onde pequenos cambios nos valores provocan alteración do rango, na redución de ruído no procesamento dixital de sinais, na restauración de imaxes, na análise de series temporais, na extracción de información de bases de datos e, finalmente, na compresión de imaxes.

Definición 9.2.1. Unha norma matricial é unha norma no espazo vectorial $V = \mathcal{M}_{p\times n}(\mathbb{R})$.

As normas usuais definidas en $\mathbb{R}^p$ e $\mathbb{R}^n$ dan lugar ao que se coñece como norma matricial inducida. Defínese do seguinte xeito para cada $A \in \mathcal{M}_{p\times n}(\mathbb{R})$:

$$\|A\| = max\{\|Av\| \mid \|v\| = 1\}.$$

Nótese que $\|A\|$ está ben definida xa que a función $f_A : \mathbb{R}^n \to \mathbb{R}^p$ é continua e acada un máximo no compacto $C_1 = \{v \in \mathbb{R}^n \mid \|v\| = 1\}$.

Proposición 9.2.2. *Para toda $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ e $u \in \mathbb{R}^n$ cúmprese que*

$$\|Au\| \leq \|A\|\|u\|.$$

Demostración. Obviamente se u é nulo a desigualdade anterior cúmprese de xeito trivial. Supoñamos que u é non nulo e consideremos o vector unitario $w = \dfrac{u}{\|u\|}$. Entón,

$$\|A\| = max\{\|Av\| \ | \ \|v\| = 1\} \geq \|Aw\| = \frac{\|Au\|}{\|u\|},$$

e isto implica que $\|Au\| \leq \|A\|\|u\|$. □

Proposición 9.2.3. *Para todo par de matrices $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ e $B \in \mathcal{M}_{n\times m}(\mathbb{R})$ cúmprese que*

$$\|AB\| \leq \|A\|\|B\|.$$

Demostración. Grazas á Proposición 9.2.2 obtemos a desigualdade desexada xa que

$$\|AB\| = max\{\|ABv\| \ | \ \|v\| = 1\} \leq max\{|A\|\|Bv\| \ | \ \|v\| = 1\} \leq max\{\|A\|\|B\|\|v\| \ | \ \|v\| = 1\}$$
$$= \|A\|\|B\|.$$

□

Nota 9.2.4. Nalgúns libros da literatura sobre o tema, cando se traballa con matrices cadradas, a desigualdade anterior (chamada propiedade multiplicativa) inclúese na definición de norma matricial. Outra opción é definir unha norma matricial como unha aplicación do conxunto de matrices sobre un corpo, denotado por $\mathcal{M}(\mathbb{R})$, en $\mathbb{R}$, tal que se cumpren as seguintes propiedades:

1) $\forall\, A \in \mathcal{M}(\mathbb{R})$ $\|A\| \geq 0$. $\|A\| = 0 \Leftrightarrow A = \Theta$.
2) $\|\alpha A\| = |\alpha|\|A\|$ $\forall\, A \in \mathcal{M}(\mathbb{R})$, $\alpha \in \mathbb{R}$.
3) Dadas $A, B \in \mathcal{M}(\mathbb{R})$, se existe $A + B$, $\|A + B\| \leq \|A\| + \|B\|$.
4) Dadas $A, B \in \mathcal{M}(\mathbb{R})$, se existe AB, $\|AB\| \leq \|A\|\|B\|$.

Definición 9.2.5. Dada unha matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ defínese o radio espectral de A como

$$\rho(A) = max\{|\lambda| \ | \ \lambda \in \mathrm{Sp}(A)\}.$$

Proposición 9.2.6. *Dada unha matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ tense que*

$$\|A\| = \sqrt{\rho(A^tA)}.$$

Logo, $\|A\| = \sigma_1$, sendo σ_1 o valor singular máis grande de A.

Demostración. Como a matriz A^tA é simétrica pódese diagonalizar de xeito ortogonal. Grazas a isto podemos garantir que existen unha matriz ortogonal $U \in \mathcal{M}_{n\times n}(\mathbb{R})$ e unha matriz diagonal $D \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que

$$U^tA^tAU = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

sendo $\{\lambda_1, \dots, \lambda_n\}$ os autovalores de A^tA. Equivalentemente,

$$A^tA = UDU^t.$$

Entón, para todo $v \in \mathbb{R}^n$

$$\|Av\|^2 = v^tA^tAv = v^tUDU^tv = (U^tv)^tD(U^tv) = \sum_{i=1}^{n} \lambda_i|w_i|^2,$$

onde $w = U^t v$. Disto último dedúcese que se v é un vector unitario

$$\|Av\|^2 \leq \rho(A^tA)\sum_{i=1}^{n}|w_i|^2 = \rho(A^tA)\|U^tv\|^2 = \rho(A^tA)\|v\|^2 = \rho(A^tA)$$

e, por tanto, $\|Av\| \leq \sqrt{\rho(A^tA)}$. En consecuencia,

$$\|A\| \leq \sqrt{\rho(A^tA)}.$$

Por outra banda, sexa λ o autovalor máis grande de A^tA. Entón, se v é un autovector unitario asociado a λ, tense que

$$\|Av\|^2 = v^tA^tAv = \lambda\|v\|^2 = \lambda$$

e deste xeito $\|A\| \geq \sqrt{\rho(A^tA)}$. □

Nota 9.2.7. Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Do resultado anterior séguese que se λ é o autovalor máis grande de A^tA e v é un autovector unitario asociado a λ,

$$\|Av\| = \sqrt{\rho(A^tA)} = \|A\|.$$

Proposición 9.2.8. *Tense o seguinte:*

1) *Dada* $A \in \mathcal{M}_{p\times n}(\mathbb{R})$,

$$\|A\| = max\{|v^tAu| \ \mid \ \|v\| = \|u\| = 1\}.$$

2) *Sexan* $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ *e* $U \in \mathcal{M}_{p\times p}(\mathbb{R})$, $V \in \mathcal{M}_{n\times n}(\mathbb{R})$ *matrices ortogonais. Entón*

$$\|A\| = \|UA\| = \|AV\|.$$

Como consecuencia tense que $\|A\| = \|U^tAV\|$.

3) *Se* $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ *é unha matriz simétrica,* $\|A\| = \rho(A)$.
4) *Se* $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ *é unha matriz ortogonal,* $\|A\| = 1$.
5) *Dada* $A \in \mathcal{M}_{p\times n}(\mathbb{R})$,

$$\|A^tA\| = \|A\|^2.$$

Demostración. Supoñamos en primeiro lugar que temos dous vectores unitarios $v \in \mathbb{R}^p$ e $u \in \mathbb{R}^n$. Entón

$$|v^tAu| = |<v, Au>| \leq \|Au\|\|v\| = \|Au\| \leq \|A\|\|u\| = \|A\|.$$

Así pois, $\|A\| \geq max\{|v^tAu| \ \mid \ \|v\| = \|u\| = 1\}$. Vexamos que se acada o valor máximo. Sexa $u \in \mathbb{R}^n$ un vector unitario tal que $\|Au\| = \|A\|$ (sabemos que existe pola nota 9.2.7) e consideremos $v = \frac{1}{\|A\|}Au$. Entón

$$v^tAu = \frac{1}{\|A\|}\|Au\|^2 = \frac{1}{\|A\|}\|A\|^2 = \|A\|$$

e isto implica que se acada o máximo.

Se $U \in \mathcal{M}_{p\times p}(\mathbb{R})$ é unha matriz ortogonal

$$\|A\|^2 = \rho(A^tA) = \rho(A^tU^tUA) = \|UA\|^2.$$

De xeito análogo obteriamos que $\|A\| = \|AV\|$ para toda matriz ortogonal $V \in \mathcal{M}_{n\times n}(\mathbb{R})$. Como consecuencia, $\|A\| = \|U^tAV\|$.

Se A é simétrica pódese diagonalizar de xeito ortogonal. Así pois existe unha matriz ortogonal $U \in \mathcal{M}_{n\times n}(\mathbb{R})$ e unha matriz diagonal $D \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que

$$U^tAU = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

sendo $\{\lambda_1, \dots, \lambda_n\}$ o seu conxunto de autovalores. Entón,

$$\|A\|^2 = \|UDU^t\|^2 = \|D\|^2 = \rho(D^tD) = max\{\lambda_1^2, \cdots, \lambda_n^2\} = \rho(A)^2,$$

e, como consecuencia, $\|A\| = \rho(A)$.

Obviamente, se U é ortogonal $\|U\| = \|UU^t\| = \|I_n\| = 1$. Finalmente, como A^tA é simétrica tense que $\|A^tA\| = \rho(A^tA) = \|A\|^2$. □

Definición 9.2.9. Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ tal que $\operatorname{rang}(A) = r$. Supoñamos que

$$A = U\Sigma V^t = \sigma_1 u_1 v_1^t + \cdots + \sigma_r u_r v_r^t$$

é a súa descomposición en valores singulares. Se k é un número natural positivo menor que r, chámase aproximación de rango k de A á matriz

$$A_k = \sigma_1 u_1 v_1^t + \cdots + \sigma_k u_k v_k^t.$$

Nota 9.2.10. Se A_k é a aproximación de rango k de $A \in \mathcal{M}_{p\times n}(\mathbb{R})$, $\operatorname{rang}(A) = r$ e

$$A = U\Sigma V^t = \sigma_1 u_1 v_1^t + \cdots + \sigma_r u_r v_r^t$$

é a súa descomposición en valores singulares, tense que

$$A_k = U\Sigma_k V^t,$$

onde

$$\Sigma_k = \begin{pmatrix} D_k & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{pmatrix} \in \mathcal{M}_{p\times n}(\mathbb{R}),$$

con

$$D_k = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_k \end{pmatrix} \in \mathcal{M}_{k\times k}(\mathbb{R}).$$

Entón, $A_k = \sigma_1 u_1 v_1^t + \cdots + \sigma_k u_k v_k^t$ é a descomposición en valores singulares de A_k e o seu rango é k. Logo,

$$A - A_k = UT_kV^t,$$

onde

$$T_k = \begin{pmatrix} \Theta & | & \Theta & | & \Theta \\ - & - & - & - & - \\ \Theta & | & R_k & | & \Theta \\ - & - & - & - & - \\ \Theta & | & \Theta & | & \Theta \end{pmatrix} \in \mathcal{M}_{p\times n}(\mathbb{R}),$$

con

$$R_k = \begin{pmatrix} \sigma_{k+1} & 0 & \cdots & 0 \\ 0 & \sigma_{k+2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{pmatrix} \in \mathcal{M}_{r-k\times r-k}(\mathbb{R}).$$

Tendo en conta isto e reordenando as columnas de U e V, obtemos dúas novas matrices ortogonais W e H tales que

$$A - A_k = W P_k H^t,$$

onde

$$P_k = \begin{pmatrix} S_k & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{pmatrix} \in \mathcal{M}_{p\times n}(\mathbb{R}),$$

e

$$S_k = \begin{pmatrix} \sigma_{k+1} & 0 & \cdots & 0 \\ 0 & \sigma_{k+2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{pmatrix} \in \mathcal{M}_{r-k\times r-k}(\mathbb{R}).$$

Esta é a descomposición en valores singulares de $A - A_k$, ou dito doutro xeito,

$$A - A_k = \sigma_{k+1} w_1 h_1^t + \sigma_{k+2} w_2 h_2^t + \cdots + \sigma_r w_{r-k} h_{r-k}^t.$$

Logo, pola Proposición 9.2.6, obtemos que

$$\|A - A_k\| = \sigma_{k+1}.$$

Teorema 9.2.11. *Sexa* $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ *tal que* $\operatorname{rang}(A) = r$. *Supoñamos que*

$$A = U\Sigma V^t = \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t + \cdots + \sigma_r u_r v_r^t$$

é a súa descomposición en valores singulares e que A_k *é a aproximación de rango* k *de* A *para* k *tal que* $0 < k < r$. *Entón*

$$||A - A_k|| = min\{||A - B||, \mid B \in \mathcal{M}_{p\times n}(\mathbb{R}), \operatorname{rang}(B) = k\}.$$

Demostración. Sexa k un número natural tal que $0 < k < r$. Sexa $B \in \mathcal{M}_{p\times n}(\mathbb{R})$ de rango k. Se $\|A - B\| < \|A - A_k\|$, teriamos que

$$\|A - B\| < \sigma_{k+1}.$$

O conxunto de solucións do sistema $Bx = \theta_{\mathbb{R}^n}$ é un subespazo S de dimensión $n - k$. Se $w \in S$ tense que

$$\|Aw\| = \|(A - B)w\| \le \|(A - B)\|\|w\| < \sigma_{k+1}\|w\|.$$

Sexa L o subespazo xerado polos vectores $\{v_1, \ldots, v_k, v_{k+1}\}$. Para estes vectores temos que $Av_i = \sigma_i u_i$ e como $\dim_{\mathbb{R}}(L) = k+1$ e $\dim_{\mathbb{R}}(S) = n-k$ obtense, grazas á fórmula de Grassmann, que $S \cap L \neq \emptyset$. Daquela existe w non nulo en $S \cap L$ e isto implica que se

$$w = \sum_{i=1}^{k+1} \alpha_i v_i$$

obtemos que

$$\|Aw\| = \|\sum_{i=1}^{k+1} \alpha_i A v_i\| = \|\sum_{i=1}^{k+1} \alpha_i \sigma_i u_i\| = (\sum_{i=1}^{k+1} (\alpha_i \sigma_i)^2)^{\frac{1}{2}} \ge \sigma_{k+1} (\sum_{i=1}^{k+1} \alpha_i^2)^{\frac{1}{2}} = \sigma_{k+1}\|w\|,$$

o que é un absurdo polo probado anteriormente. Deste xeito cúmprese que

$$||A - A_k|| = min\{||A - B||, \mid B \in \mathcal{M}_{p\times n}(\mathbb{R}), \operatorname{rang}(B) = k\}.$$

$\square$

Nota 9.2.12. Nótese que en certo modo A_k é a matriz que máis preto está de A segundo a norma $\|\ \|$ de entre todas as de rango k que pertencen ao espazo vectorial $\mathcal{M}_{p\times n}(\mathbb{R})$.

Exemplo 9.2.13. Neste exemplo calcularemos a mellor aproximación de rango dous da matriz

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \in \mathcal{M}_{4\times 3}(\mathbb{R}).$$

Dado que

$$A^tA = \begin{pmatrix} 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

os autovalores de A^tA son 3, 2 e 1, sendo os seus valores singulares

$$\sigma_1 = \sqrt{3} > \sigma_2 = \sqrt{2} > \sigma_3 = \sqrt{1} = 1.$$

Entón,

$$\Sigma = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

e $\operatorname{rang}(A) = 3$.

Os subespazos propios de cada un dos autovalores de A^tA son:

$$V(3) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid (A^tA - 3I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid y = z = 0 \right\}$$

$$= \langle \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\} \rangle,$$

$$V(2) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid (A^tA - 2I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x = z = 0 \right\}$$

$$= \langle \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \rangle,$$

$$V(1) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid (A^tA - 1I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x = y = 0 \right\}$$

$$= \langle \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \rangle,$$

e deste xeito, tendo en conta que son vectores ortogonais e unitarios, tómase

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\ v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

e $V = I_3$.

Os vectores que dan lugar ás columnas de U virán dados por:

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{3}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \sqrt{3}/3 \\ -\sqrt{3}/3 \\ \sqrt{3}/3 \end{pmatrix},$$

$$u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \sqrt{2}/2 \\ -\sqrt{2}/2 \\ 0 \end{pmatrix},$$

e

$$u_3 = \frac{1}{\sigma_3} A v_3 = \frac{1}{1} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Logo, a descomposición en valores singulares de A é

$$A = \sqrt{3} u_1 v_1^t + \sqrt{2} u_2 v_2^t + u_3 v_3^t$$

e a aproximación de rango dous de A é

$$A_2 = \sqrt{3} u_1 v_1^t + \sqrt{2} u_2 v_2^t = \sqrt{3} \begin{pmatrix} 0 \\ \sqrt{3}/3 \\ -\sqrt{3}/3 \\ \sqrt{3}/3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} + \sqrt{2} \begin{pmatrix} 0 \\ \sqrt{2}/2 \\ -\sqrt{2}/2 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

9.3. Pseudoinversas ou inversas xeneralizadas de Moore-Penrose

Como outra aplicación da descomposición en valores singulares nesta sección estudaremos unha extensión da noción de invertibilidade para matrices debida a E.H. Moore (1862–1932) e R. Penrose (1931–).

Teorema 9.3.1. *Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Existe unha única matriz $A^+ \in \mathcal{M}_{n\times p}(\mathbb{R})$, chamada matriz pseudoinversa de A, ou tamén inversa xeneralizada de Moore-Penrose de A, tal que cumpre as seguintes propiedades:*

1) $AA^+A = A$.
2) $A^+AA^+ = A^+$.
3) AA^+ *é simétrica.*
4) A^+A *é simétrica.*

Demostración. Veremos en primeiro lugar que existe unha matriz cumprindo as catro propiedades do enunciado do teorema. Supoñamos que a descomposición en valores singulares de A é

$$A = U\Sigma V^t,$$

onde $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ son os valores singulares non nulos de A. Tomemos

$$A^+ = V\Omega U^t$$

con

$$\Omega = \begin{pmatrix} T_r & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{pmatrix} \in \mathcal{M}_{n\times p}(\mathbb{R})$$

onde

$$T_r = \begin{pmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_r} \end{pmatrix} = D_r^{-1} \in \mathcal{M}_{r\times r}(\mathbb{R}).$$

A matriz A^+ cumpre as catro propiedades do enunciado. En efecto, a primeira igualdade cúmprese xa que

$$AA^+A = U\Sigma V^t V\Omega U^t U\Sigma V^t = U\Sigma\Omega\Sigma V^t = U\Sigma V^t = A,$$

e de xeito similar obtense que $A^+AA^+ = A^+$. Por outra banda,

$$(AA^+)^t = (A^+)^t A^t = (V\Omega U^t)^t (U\Sigma V^t)^t = U\Omega^t V^t V\Sigma^t U^t = U\Omega^t\Sigma^t U^t$$

$$= U(\Sigma\Omega)^t U^t = U\Sigma V^t V\Omega U^t = AA^+$$

e, como consecuencia, AA^+ é unha matriz simétrica. A proba do carácter simétrico da matriz A^+A faise de xeito análogo.

Para probar a unicidade da matriz A^+ supoñamos que existe outra matriz B satisfacendo as condicións $ABA = A$, $BAB = B$, AB é simétrica e BA é simétrica. Entón:

$$(A^+)^t A^t = (AA^+)^t = AA^+, \; B^t A^t = (AB)^t = AB,$$

e, similarmente,

$$A^t (A^+)^t = (A^+A)^t = A^+A, \; A^t B^t = (BA)^t = BA.$$

Ademais cúmprese que

$$AA^+ = (AA^+)^t = (A^+)^t A^t = (A^+)^t (ABA)^t = (A^+)^t A^t (AB)^t = (AA^+)^t AB = AA^+AB = AB,$$

e, de xeito análogo,

$$A^+A = BA.$$

Logo,

$$A^+ = A^+AA^+ = A^+AB = BAB = B,$$

quedando a demostración finalizada. □

Nota 9.3.2. Nótese que das propiedades que definen á matriz pesudoinversa dedúcese que

$$A^+ = (A^+)^+.$$

Exemplo 9.3.3. Neste exemplo calcularemos a matriz pseudoinversa de

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

Grazas ao feito no Exemplo 9.1.4 sabemos que

$$A^t A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

que $\mathrm{Sp}(A^t A) = \{0, 2, 4\}$ e que os valores singulares de A son

$$\sigma_1 = \sqrt{4} = 2 \geq \sigma_2 = \sqrt{2} \geq \sigma_3 = \sqrt{0} = 0 \geq 0.$$

Como consecuencia do anterior, obtense que a matriz Σ será

$$\Sigma = \begin{pmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Pasemos a calcular a matriz V. Os subespazos propios asociados aos autovalores $\lambda_1 = 4$, $\lambda_2 = 2$ e $\lambda_3 = 0$ son:

$$V(4) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid (A^t A - 4I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} -3 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x = z = 0 \right\}$$

$$= \langle \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \rangle,$$

$$V(2) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid (A^t A - 2I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} -1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{array}{r} x = z \\ y = 0 \end{array} \right\}$$

$$= \langle \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\} \rangle,$$

$$V(0) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid (A^t A - 0I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} 1 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{array}{r} z = -x \\ y = 0 \end{array} \right\}$$

$$= \langle \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\} \rangle.$$

Deste xeito

$$\left\{ w_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\ w_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},\ w_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}$$

é unha base de $\mathbb{R}^3$ formada por autovectores. Como son ortogonais entre si, para calcular v_1, v_2 e v_3, só temos que normalizar. Entón:

$$v_1 = w_1,\ v_2 = \frac{w_2}{\|w_2\|} = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix},\ v_3 = \frac{w_3}{\|w_3\|} = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

Por tanto,

$$V = \begin{pmatrix} 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

Pasemos agora a construír a matriz $U \in \mathcal{M}_{3\times 3}(\mathbb{R})$. A primeira columna está dada por

$$u_1 = \frac{1}{2} A v_1 = \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Similarmente, a segunda columna de U calcúlase como

$$u_2 = \frac{1}{\sqrt{2}} A v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Como o rango da matriz A é $r = 2$ e $2 < p = 3$, o vector u_3 debe obterse grazas á determinación dunha base do conxunto de solucións S do sistema linear homoxéneo

$$A^t \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \theta_{\mathbb{R}^3}.$$

Agora ben, como

$$A^t \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \left. \begin{array}{r} x = 0 \\ z = 0 \end{array} \right\},$$

tense que

$$S = \langle \left\{ \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\} \rangle,$$

e, como consecuencia, unha base de S pode ser a dada polo vector

$$w = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

que xa ten norma un. Tendo en conta isto último, tomamos a terceira columna de U como $u_3 = w$ e, deste xeito

$$U = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Dado que

$$\Omega = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

a matriz pseudoinversa de A será:

$$\begin{aligned} A^+ &= V\Omega U^t \\ &= \begin{pmatrix} 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}. \end{aligned}$$

É doado comprobar que $AA^+A = A$ e $A^+AA^+ = A^+$. Finalmente, AA^+ e AA^+ son simétricas xa que

$$AA^+ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

e

$$AA^+ = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

Nota 9.3.4. Do Teorema 9.3.1 dedúcese que se $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ e A^+ é a súa pseudoinversa cúmprese a seguinte igualdade:

$$A^+ = V\Omega U^t = \frac{1}{\sigma_1} v_1 u_1^t + \frac{1}{\sigma_2} v_2 u_2^t + \cdots + \frac{1}{\sigma_r} v_r u_r^t.$$

Nota 9.3.5. Tamén do Teorema 9.3.1 dedúcese que, se $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ é invertíbel, a pseudoinversa de A é exactamente a matriz inversa de A. Cando a matriz A^tA é invertíbel, entón

$$A^+ = (A^tA)^{-1}A^t.$$

Finalmente, existe un método alternativo baseado na factorización QR para o cálculo da pseudoinversa. Neste caso tense que, se $A = QR$, entón a pseudoinversa pódese calcular como:

$$A^+ = R^t(RR^t)^{-1}Q^t.$$

Exemplo 9.3.6. Supoñamos que tomamos a matriz

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

entón

$$A^tA = \begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix}$$

resulta ser unha matriz invertíbel con inversa

$$(A^tA)^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{pmatrix}.$$

Como consecuencia, a pseudoinversa de A está dada por:

$$A^+ = (A^tA)^{-1}A^t = \begin{pmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & -1 & \frac{1}{2} \end{pmatrix}.$$

Exemplo 9.3.7. Supoñamos que tomamos a matriz

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in \mathcal{M}_{4\times 3}(\mathbb{R}).$$

No segundo caso de Exemplos 7.4.3 vimos que a factorización QR de A está dada por

$$Q = \begin{pmatrix} \sqrt{2}/2 & \sqrt{6}/6 \\ \sqrt{2}/2 & -\sqrt{6}/6 \\ 0 & \sqrt{6}/3 \end{pmatrix} \in \mathcal{M}_{3\times 2}(\mathbb{R})$$

e

$$R = Q^tA = \begin{pmatrix} \sqrt{2} & \sqrt{2}/2 & 3\sqrt{2}/2 \\ 0 & \sqrt{6}/2 & \sqrt{6}/2 \end{pmatrix} \in \mathcal{M}_{2\times 3}(\mathbb{R}).$$

Entón

$$A^+ = R^t(RR^t)^{-1}Q^t = \begin{pmatrix} 1/9 & 5/9 & -4/9 \\ 1/9 & -4/9 & -5/9 \\ 2/9 & 1/9 & -1/9 \end{pmatrix}.$$

9.4. Problemas de mínimos cadrados

Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ e $b \in \mathbb{R}^p$. Supoñamos que temos o sistema $Ax = b$. Como sabemos o anterior sistema é incompatíbel se b non pertence ao subespazo de $\mathbb{R}^n$ dado por

$$\mathrm{Im}(f_A) = \{Aw \in \mathbb{R}^n \ \mid \ w \in \mathbb{R}^n\}.$$

A continuación veremos que, a pesar de ser un sistema incompatíbel, podemos atopar unha solución aproximada óptima ou, o que é o mesmo, un vector $w \in \mathbb{R}^n$ tal que a distancia en $\mathbb{R}^p$ entre Aw e b sexa mínima.

Definición 9.4.1. Sexa U un subespazo vectorial de $\mathbb{R}^p$. Diremos que $v \in U$ é a mellor aproximación dun vector $b \in \mathbb{R}^p$ por vectores de U no sentido dos mínimos cadrados, se

$$\|v - b\| = min\{\|u - b\|,\ u \in U\}.$$

Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ e $b \in \mathbb{R}^p$. Supoñamos que temos o sistema de ecuacións lineares $Ax = b$. Diremos que o vector $w \in \mathbb{R}^n$ é unha solución mínimo-cuadrática do sistema $Ax = b$, se Aw é a mellor aproximación de $b \in \mathbb{R}^p$ no sentido dos mínimos cadrados por vectores de $U = \mathrm{Im}(f_A)$.

O nome de mínimos cadrados débese a que minimizar $\|u - b\|$ é equivalente a minimizar

$$\|u - b\|^2 = \sum_{i=1}^{p}(u_i - b_i)^2,$$

sendo u_i e b_i as compoñentes dos vectores de coordenadas de u e b respecto da base canónica de $\mathbb{R}^p$.

Proposición 9.4.2. *Sexa U un subespazo do espazo vectorial $\mathbb{R}^p$. Un vector $v \in U$ é a mellor aproximación de $b \in \mathbb{R}^p$ por vectores de U, se $< v - b, u >= 0$, $\forall u \in U$ ($v - b$ é ortogonal a todos os vectores de U).*

Demostración. Sexa $v \in U$ tal que $< v - b, u >= 0$, $\forall u \in U$. Entón, dado $u \in U$,

$$\|u - b\|^2 = \|(u - v) + (v - b)\|^2 =< (u - v) + (v - b), (u - v) + (v - b) >=$$

$$\|u - v\|^2 + \|v - b\|^2 + 2 < u - v, v - b >=$$

$$\|u - v\|^2 + \|v - b\|^2,$$

xa que, como $u - v \in U$, tense que $< u - v, v - b >= 0$.

Como $\|u-v\|^2 \geq 0$, deducimos que $\|u-b\|^2 \geq \|v-b\|^2$ para todo $u \in U$ e, como consecuencia,

$$\|v-b\| = min\{\|u-b\|,\ u \in U\}.$$

□

Definición 9.4.3. Sexa U un subespazo do espazo vectorial $\mathbb{R}^p$. Ao vector $v \in U$ tal que $v-b$ é ortogonal a todos os vectores de U chámaselle proxección ortogonal de b sobre o subespazo U.

Teorema 9.4.4. *Sexa U un subespazo do espazo vectorial $\mathbb{R}^p$. A proxección ortogonal de b sobre o subespazo U sempre existe e é única.*

Demostración. Vexamos en primeiro lugar que sempre é posíbel calcular unha proxección ortogonal de b sobre o subespazo U. Supoñamos que $\dim_{\mathbb{R}}(U) = r$ e que $B' = \{v_1, \dots, v_r\}$ é unha base de U. Se ampliamos B' a unha base $B = \{v_1, \dots, v_r, v_{r+1}, \dots, v_p\}$ do espazo $\mathbb{R}^p$ e aplicamos o procedemento de ortonormalización de Gram-Schmidt a B, obtemos unha base ortonormal $C = \{u_1, \dots, u_r, u_{r+1}, \dots, u_p\}$ de $\mathbb{R}^p$ tal que $C' = \{u_1, \dots, u_r\}$ é unha base ortonormal de U. Así,

$$b = \alpha_1 u_1 + \cdots + \alpha_r u_r + \alpha_{r+1} u_{r+1} + \cdots + \alpha_p u_p$$

e, se tomamos $v = \alpha_1 u_1 + \cdots + \alpha_r u_r$, temos que $v \in U$ e $< v-b, u >= 0$, $\forall u \in U$. Logo, v é unha proxección ortogonal de b sobre o subespazo U.

Comprobemos a continuación a unicidade de dita proxección. Se existise $v' \in U$ tal que

$$< v'-b, u >= 0,\ \forall u \in U,$$

teriamos que

$$v - v' = (v-b) - (v'-b)$$

é un vector que pertence a U e que ademais é ortogonal a todos os vectores de U xa que

$$\forall u \in U \ \ < v-v', u >=< (v-b)-(v'-b), u >=< v-b, u > - < v'-b, u >= 0.$$

Entón, $v-v'$ é ortogonal a si mesmo. Isto implica que

$$0 =< v-v', v-v' >= \|v-v'\|^2,$$

e entón $v = v'$.

□

Nota 9.4.5. Tendo en conta o resultado anterior, se U é un subespazo do espazo vectorial $\mathbb{R}^p$, a proxección ortogonal de b sobre o subespazo U está dada por $v = \alpha_1 u_1 + \cdots + \alpha_r u_r$ sendo $b = \alpha_1 u_1 + \cdots + \alpha_r u_r + \alpha_{r+1} u_{r+1} + \cdots + \alpha_p u_p$. Como $C = \{u_1, \dots, u_r, u_{r+1}, \dots, u_p\}$ é unha base ortonormal de $\mathbb{R}^p$ e $\{u_1, \dots, u_r\}$ unha base ortonormal de U, tense que

$$v =< b, u_1 > u_1 + \cdots + < b, u_r > u_r$$

é a expresión para a proxección ortogonal. Ademais, a anterior igualdade implica que

$$v = Qb,$$

sendo Q a matriz dada por

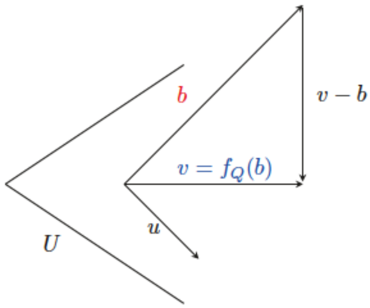
$$Q = AA^t = u_1 u_1^t + \cdots + u_r u_r^t,$$

con

$$A = \left(\begin{array}{c|c|c} u_1 & \cdots & u_r \end{array} \right) \in \mathcal{M}_{p\times r}(\mathbb{R}).$$

A matriz Q chámase matriz de proxección ortogonal sobre U. Deste xeito a proxección ortogonal sobre U pódese interpretar como unha aplicación linear dada por:

$$f_Q : \mathbb{R}^p \to \mathbb{R}^p.$$



Exemplo 9.4.6. Sexa o subespazo de $\mathbb{R}^3$ dado por

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \ \mid \ x + y - z = 0 \right\}.$$

A dimensión de U é dous e unha base deste subespazo é

$$B' = \left\{ v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \ v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}.$$

Se aplicamos o procedemento de ortonormalización de Gram-Schmidt á base anterior obtemos os vectores

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix},$$

$$u_2' = v_2 - < v_2, u_1 > u_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} - < \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix} > \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} - \frac{\sqrt{2}}{2} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} - \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -1 \\ -\frac{1}{2} \end{pmatrix},$$

$$u_2 = \frac{u_2'}{\|u_2'\|} = \frac{2}{\sqrt{6}} \begin{pmatrix} \frac{1}{2} \\ -1 \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \sqrt{6}/6 \\ -\sqrt{6}/3 \\ -\sqrt{6}/6 \end{pmatrix}$$

e, como consecuencia,

$$C' = \left\{ u_1 = \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{pmatrix}, \ u_2 = \begin{pmatrix} \sqrt{6}/6 \\ -\sqrt{6}/3 \\ -\sqrt{6}/6 \end{pmatrix} \right\}$$

é unha base ortonormal de U.

Deste xeito, dado

$$b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

a súa proxección ortogonal sobre o subespazo U está dada por

$$v = < b, u_1 > u_1 + < b, u_2 > u_2$$

sendo

$$< b, u_1 > = \sqrt{2}, \ \ < b, u_2 > = -\frac{\sqrt{6}}{3}.$$

Por tanto,

$$v = \sqrt{2}u_1 - \frac{\sqrt{6}}{3}u_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -1/3 \\ 2/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 2/3 \\ 4/3 \end{pmatrix}.$$

Por outra banda, a matriz de proxección ortogonal é $Q = AA^t$ onde

$$A = \left(\ u_1 \ \ | \ \ u_2 \ \right) = \begin{pmatrix} \sqrt{2}/2 & \sqrt{6}/6 \\ 0 & -\sqrt{6}/3 \\ \sqrt{2}/2 & -\sqrt{6}/6 \end{pmatrix}.$$

Deste xeito,

$$Q = \begin{pmatrix} 2/3 & -1/3 & 1/3 \\ -1/3 & 2/3 & 1/3 \\ 1/3 & 1/3 & 2/3 \end{pmatrix},$$

e a proxección ortogonal de b sobre U obtense tamén como

$$v = Qb = \begin{pmatrix} 2/3 \\ 2/3 \\ 4/3 \end{pmatrix}.$$

Definición 9.4.7. Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$, $b \in \mathbb{R}^p$ e supoñamos que temos o sistema de ecuacións lineares $Ax = b$. O sistema $A^tAx = A^tb$ coñécese co nome de sistema de ecuacións normais do sistema $Ax = b$.

Teorema 9.4.8. *Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$, $b \in \mathbb{R}^p$ e supoñamos que temos o sistema de ecuacións lineares $Ax = b$. O conxunto de solucións mínimo-cuadráticas do sistema $Ax = b$ coincide co conxunto de solucións do sistema de ecuacións normais $A^tAx = A^tb$.*

Demostración. Sexa $w \in \mathbb{R}^n$ unha solución mínimo-cuadrática do sistema $Ax = b$. Entón, Aw é a proxección ortogonal de b sobre $U = \mathrm{Im}(f_A)$ e, como consecuencia, $Aw - b$ é ortogonal a todos os vectores do subespazo U. Utilizando este feito, tense que

$$0 = < Aw - b, Av > = (Av)^t(Aw - b) = v^tA^tAw - v^tA^tb = v^t(A^tAw - A^tb)$$

$$= < A^tAw - A^tb, v >, \ \forall v \in \mathbb{R}^n.$$

Isto implica que $A^tAw - A^tb = \theta_{\mathbb{R}^n}$ ou, o que é o mesmo, $A^tAw = A^tb$. Así pois w é unha solución do sistema $A^tAx = A^tb$.

Inversamente, se w é unha solución do sistema $A^tAx = A^tb$, temos que $A^tAw - A^tb = \theta_{\mathbb{R}^n}$ e w é a solución mínimo-cuadrática do sistema $Ax = b$, xa que $\forall v \in \mathbb{R}^n$

$$0 = < A^tAw - A^tb, v >$$

$$= v^t(A^tAw - A^tb) = v^tA^tAw - v^tA^tb = (Av)^t(Aw - b) = < Aw - b, Av > .$$

□

Nota 9.4.9. Do teorema anterior deducimos que, se $w, w' \in \mathbb{R}^n$ son solucións mínimo-cuadráticas do sistema $Ax = b$, tense a igualdade $Aw = Aw'$. Logo, $A(w - w') = \theta_{\mathbb{R}^p}$, equivalentemente, $w - w' \in \mathrm{Ker}(f_A)$. Entón, se w é unha solución mínimo-cuadrática de $Ax = b$, o conxunto global de solucións do problema de mínimos cadrados tamén ven dado por

$$S = w + \mathrm{Ker}(f_A) = \{w + u \mid u \in \mathrm{Ker}(f_A)\}.$$

Evidentemente, se A^tA é invertíbel, S só ten un elemento e é

$$w = (A^tA)^{-1}A^tb = A^+b$$

Exemplo 9.4.10. Sexa A a matriz dada por

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & -2 \\ 1 & 2 & -1 \end{pmatrix} \in \mathcal{M}_{4\times 3}(\mathbb{R}).$$

Consideremos o sistema $Ax = b$ sendo

$$b = \begin{pmatrix} 4 \\ 6 \\ 1 \\ 2 \end{pmatrix}.$$

Como $\mathrm{rang}\,(A) = 2 \neq 3 = \mathrm{rang}\,(A|b)$, tense que o sistema é incompatíbel. As solucións mínimo-cuadráticas as atoparemos resolvendo o sistema $A^tAx = A^tb$ que neste caso está dado por:

$$A = \begin{pmatrix} 4 & 8 & 0 \\ 8 & 16 & 0 \\ 0 & 0 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 13 \\ 26 \\ 12 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{l} z = \frac{6}{5} \\ 4x + 8y = 13 \end{array} \right. .$$

Entón,

$$S = \left\{ \begin{pmatrix} \frac{13}{4} - 2y \\ y \\ \frac{6}{5} \end{pmatrix} \mid y \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} \frac{13}{4} \\ 0 \\ \frac{6}{5} \end{pmatrix} + y \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \mid y \in \mathbb{R} \right\} =$$

$$\begin{pmatrix} \frac{13}{4} \\ 0 \\ \frac{6}{5} \end{pmatrix} + \langle \left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \right\} \rangle,$$

sendo

$$\mathrm{Ker}(f_A) = \langle \left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \right\} \rangle.$$

Teorema 9.4.11. *Sexa $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ e supoñamos que a súa factorización QR é $A = QR$. Cúmprese que w é solución mínimo-cuadrática para $Ax = b$, se, e só se, w é solución do sistema $Rx = Q^tb$.*

Demostración. O vector w é solución mínimo-cuadrática para $Ax = b$, se, e só se, w é solución do sistema $A^tAx = A^tb$. Se $A = QR$, temos que

$$A^tAw = A^tb \Leftrightarrow R^tQ^tQRw = R^tQ^tb \Leftrightarrow R^tRw = R^tQ^tb \Leftrightarrow$$
$$RR^tRw = RR^tQ^tb$$

e, como $RR^t \in \mathcal{M}_{r\times r}(\mathbb{R})$ é invertíbel, obtemos que $Rw = Q^tb$.

Inversamente, se $Rw = Q^tb$, tense que

$$R^tRw = R^tQ^tb \Leftrightarrow R^tQ^tQRw = R^tQ^tb \Leftrightarrow A^tAw = A^tb.$$

□

Exemplo 9.4.12. Tomemos o sistema dado no exemplo 9.4.10. Entón, como

$$A = QR = \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{10}} \\ \frac{1}{2} & \frac{2}{\sqrt{10}} \\ \frac{1}{2} & \frac{-2}{\sqrt{10}} \\ \frac{1}{2} & \frac{-1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 2 & 4 & 0 \\ 0 & 0 & \sqrt{10} \end{pmatrix}$$

e

$$Q^t b = \begin{pmatrix} 13/2 \\ \frac{12}{\sqrt{10}} \end{pmatrix},$$

obtemos

$$R \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{13}{2} \\ \frac{12}{\sqrt{10}} \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{l} z = \frac{12}{10} \\ 2x + 4y = \frac{13}{2} \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} z = \frac{6}{5} \\ 4x + 8y = 13 \end{array} \right. .$$

De xeito máis xeral pódese afirmar que o conxunto de solucións do problema de mínimos cadrados está dado por

$$S = A^+ b + \mathrm{Ker}(f_A) = A^+ b + \mathrm{Ker}(f_{A^t A}).$$

Este último método de solución ten como vantaxe sobre o baseado na utilización das ecuacións normais, que é menos sensíbel aos problemas de estabilidade e que en xeral dá lugar a unha técnica computacional mellor.

Teorema 9.4.13. *Sexa $A \in \mathcal{M}_{p \times n}(\mathbb{R})$. Cúmprese que $A^+ b$ é a solución mínimo-cuadrática de norma mínima do sistema $Ax = b$.*

Demostración. Supoñamos en primeiro lugar que o sistema é compatíbel indeterminado xa que a proba do caso determinado é trivial. Neste caso a solución mínimo-cuadrática coincide coa solución do sistema e o que teremos que probar é que $A^+ b$ é a solución de norma mínima do sistema $Ax = b$. Supoñamos que $u \in \mathbb{R}^n$ é unha solución do sistema $Ax = b$. Entón $Au = b$. Logo,

$$b = Au = AA^+ Au = AA^+ b,$$

e así próbase que $A^+ b$ é unha solución do sistema $Ax = b$.

Como sabemos pola teoría xeral de sistemas de ecuacións lineares a solución xeral do sistema $Ax = b$ escríbese como unha solución particular sumada con todas as solucións do sistema homoxéneo asociado. Entón

$$A^+ b + \mathrm{Ker}(f_A)$$

é a solución xeral neste caso. Sexa $w \in \mathrm{Ker}(f_A)$ e sexa $A = QR$ a descomposición QR de A. Como $Aw = \theta_{\mathbb{R}^p}$ se e só se $Rw = \theta_{\mathbb{R}^p}$,

$$< A^+ b, w > = < R^t (RR^t)^{-1} Q^t b, w > = < (RR^t)^{-1} Q^t b, Rw > = 0$$

e, como consecuencia, $A^+ b$ e w son ortogonais. Aplicando o Teorema de Pitágoras a $A^+ b$ e w obtemos que

$$\|A^+ b + w\|^2 = \|A^+ b\|^2 + \|w\|^2 \geq \|A^+ b\|^2,$$

co que a igualdade dáse se e só se $w = \theta_{\mathbb{R}^n}$. Por tanto, deste xeito próbase que $A^+ b$ é a solución de norma mínima do sistema $Ax = b$.

Supoñamos a continuación que o sistema é incompatíbel. As solucións mínimo-cuadráticas son as solucións do sistema de ecuacións normais $A^t Ax = A^t b$ que é compatíbel. Entón

$(A^tA)^+A^tb$ é a solución de norma mínima do sistema $A^tAx = A^tb$. Agora ben, se a descomposición en valores singulares de A é

$$A = U \left(\begin{array}{ccc} D_r & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{array} \right) V^t,$$

tense que

$$A^tA = V \left(\begin{array}{ccc} D_r^t & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{array} \right) U^tU \left(\begin{array}{ccc} D_r & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{array} \right) V^t = V \left(\begin{array}{ccc} D_r^2 & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{array} \right) V^t.$$

Logo,

$$(A^tA)^+ = V \left(\begin{array}{ccc} D_r^{-2} & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{array} \right) V^t$$

e

$$(A^tA)^+A^t = V \left(\begin{array}{ccc} D_r^{-1} & | & \Theta \\ - & - & - \\ \Theta & | & \Theta \end{array} \right) U^t = A^+.$$

Deste xeito obtense que $(A^tA)^+A^t = A^+$ e o resultado queda probado.

$\square$

Exemplo 9.4.14. Sexa o sistema

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

O sistema anterior é incompatíbel xa que $\operatorname{rang}(A) = 2 < \operatorname{rang}(A|b) = 3$. Como temos comprobado no teorema anterior A^+b é a solución mínimo-cuadrática de norma mínima do sistema $Ax = b$. Como

$$A^+ = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

a calculamos no exemplo 9.3.3, tense que

$$A^+b = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \frac{1}{2} \end{pmatrix}.$$

Por outra banda,

$$\operatorname{Ker}(f_A) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{array}{l} x + z = 0 \\ y = 0 \end{array} \right\} = \langle \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\} \rangle$$

e, como consecuencia, o conxunto de solucións do problema de mínimos cadrados está dado por

$$S = A^+b + \operatorname{Ker}(f_A) = \begin{pmatrix} 1 \\ 0 \\ \frac{1}{2} \end{pmatrix} + \langle \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\} \rangle.$$

9.5. Axuste polinómico de datos mediante mínimos cadrados

markright9.5. AXUSTE POLINÓMICO DE DATOS MEDIANTE MÍNIMOS CADRADOS

Supoñamos que temos un conxunto de pares $\{(x_i, y_i)\}_{i=1}^n$ e que estamos interesados en atopar o polinomio $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ cuxa gráfica se axuste mellor aos datos dados polos pares anteriores. Isto é, queremos atopar os coeficientes dun polinomio como o anterior de xeito que se minimice o seguinte funcional:

$$\begin{aligned} P(a_0, a_1, \dots, a_n) &= \sum_{i=1}^n (p(x_i) - y_i)^2 \\ &= \sum_{i=1}^n \left(a_0 + a_1x_i + a_2x_i^2 + \cdots + a_nx_i^n - y_i\right)^2 \\ &= \sum_{i=1}^n \left((1, x_i, x_i^2, \dots, x_i^n) \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} - y_i \right)^2. \end{aligned}$$

Isto último é equivalente a atopar as solucións mínimo-cuadráticas do sistema linear $Ax = b$, onde:

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix}, \quad x = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad e \quad b = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}.$$

Exemplo 9.5.1. Neste exemplo calcularemos a parábola que axuste no sentido de mínimos cadrados os seguintes datos

x_i	0	0	1	-1
y_i	0	2	4	0

Dados dos conxuntos de puntos $\{x_1, x_2, x_3, x_4\}$ e $\{y_1, y_2, y_3, y_4\}$, trátase de axustar os coeficientes a_0, a_1 e a_2 da parábola de ecuación $y(x) = a_0 + a_1x + a_2x^2$. Daquela, debemos atopar as solucións mínimo-cuadráticas do sistema linear:

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix},$$

isto é,

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 4 \\ 0 \end{pmatrix}.$$

O sistema de ecuacións normais ($A^tAx = A^tb$) estará dado por:

$$\begin{pmatrix} 4 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 4 \end{pmatrix}.$$

Realizando operacións elementais por filas sobre a matriz ampliada do sistema obtemos que o sistema anterior é equivalente a

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

e deste xeito $a_0 = 1$, $a_1 = 2$ e $a_2 = 2$. Logo, a parábola buscada é

$$y = 1 + 2x + x^2.$$

9.6. Problemas propostos

1. Dada a matriz
$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$
calcule a súa descomposición en valores singulares, a súa aproximación de rango dous e a súa pseudoinversa.
2. Dada a matriz
$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & -2 \\ 1 & -1 & 2 \end{pmatrix},$$
calcule a súa descomposición en valores singulares, a súa aproximación de rango dos e a súa pseudoinversa.
3. Dada a matriz
$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \\ 1 & 1 \end{pmatrix},$$
calcule a súa pseudoinversa.
4. Dada a matriz
$$A = \begin{pmatrix} 2 & 1 \\ 0 & 3 \\ 1 & -2 \end{pmatrix},$$
calcule a súa pseudoinversa.
5. Dada a matriz
$$A = \begin{pmatrix} 2 & -1 \\ -2 & 1 \\ 4 & -2 \end{pmatrix},$$
calcule a súa pseudoinversa.
6. Calcule as pseudoinversas de A e A^t sendo
$$A = \begin{pmatrix} -2 & 0 \\ 0 & 0 \\ 0 & 5 \end{pmatrix},$$
7. Considérase o sistema de ecuacións lineares $Ax = b$, onde
$$A = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$
e $b = (1, 1, 1)^t$.
   a) Probe que o sistema é incompatíbel.
   b) Calcule o conxunto de solucións do sistema no sentido de mínimos cadrados.

8. Considérase a matriz real

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

Determine o conxunto de solucións no sentido de mínimos cadrados do sistema $Ax = b$, con $b = (3, 1, 3, 1)^t$.

9. Dada a matriz

$$A = \begin{pmatrix} 1 & 2 \\ -2 & -4 \\ 2 & 4 \end{pmatrix},$$

determine, utilizando a pseudoinversa de A, o conxunto de solucións no sentido de mínimos cadrados do sistema $Ax = b$, onde $b = (4, -5, 5)^t$.

10. Dada a matriz

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 2 \end{pmatrix},$$

determine, utilizando a pseudoinversa de A, o conxunto de solucións no sentido de mínimos cadrados do sistema $Ax = b$ onde $b = (0, 1, 0)^t$.

11. Atope a parábola que axuste no sentido de mínimos cadrados os seguintes datos

a)

x	0	0	1	-1
y	0	2	4	0

b)

x	-2	-1	1	2
y	3	1	1	5

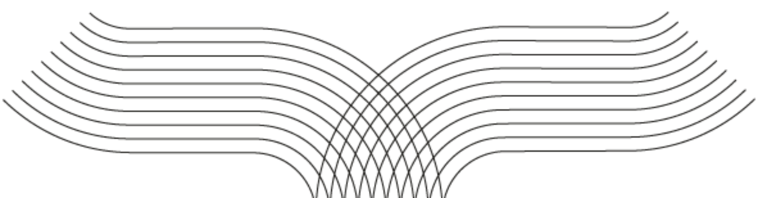
12. Determine $a, b \in \mathbb{R}$ para que a recta $ax + by = 1$ sexa a máis próxima no sentido de mínimos cadrados aos puntos $(0, 0), (1, 1), (-1, 1)$ e $(1, -1)$.

Bibliografía

[1] J. Arvesú, R. Nodarse, F. Marcellán. *Álgebra lineal y aplicaciones*, Síntesis, 1999.
[2] S. Axler. *Linear algebra done right*, 2ª ed., Springer-Verlag, 2001.
[3] R. Bathia. *Matrix analysis*, Springer, 1997.
[4] E. T. Bell. *Historia de las matemáticas*, Fondo de Cultura Económica, 1992.
[5] E. T. Bell. *Los grandes matemáticos*, Losada, 2010.
[6] R. Bellman. *Introduction to matrix analysis*, Society for Industrial and Applied Mathematics, 1995.
[7] G. Birkhoff, S. Mac Lane. *Álgebra moderna*, 4ª ed., Vicens-Vives, 1970.
[8] N. Bourbaki. *Elementos de historia de las matemáticas*, Alianza Universidad, 1976.
[9] C. B. Boyer. *Historia de la matemática*, Alianza Universidad, 1999.
[10] J. de Burgos. *Álgebra lineal y geometría cartesiana*, 3ª ed., McGraw Hill, 2006.
[11] S. L. Campbell, C. D. Meyer. *Generalized inverses of linear transformations*, Dover, 1991.
[12] J-L. Dorier. *A general outline of the gemesis of vector space theory*, Historia Mathematica 22, 227-261, 1995.
[13] P. Fernández. *El secreto de Google y el álgebra lineal*, Boletín de la Sociedad Española de Matemática Aplicada, 30 (2004), 115-141.
[14] F. R. Gantmacher. *The theory of matrices. 2 Tomos*, AMS Chelsea Publishing, 1998.
[15] S. Grossman. *Álgebra lineal*, 7ª ed., McGraw-Hill, 2012.
[16] G. Golub, C. V. Loan. *Matrix computations*, 3ª ed., Johns Hopkins Univ. Press, 1996.
[17] N. J. Higham. *Functions of matrices: theory and computation*, Society for Industrial and Applied Mathematics, 2008.
[18] V. J. Katz. *A history of mathematics*, Addison-Wesley, 1998.
[19] I. Kleiner. *A history of abstract algebra*, Birkhäuser, 2007.
[20] B. Kolman, D. Hill. *Álgebra lineal con aplicaciones y Matlab*, 8ª ed., Pearson, 2006.
[21] A. I. Kostrikin. *Introducción al álgebra lineal*, McGraw-Hill, 1992.
[22] P. Lancaster, M. Tismenetsky. *The theory of matrices*, 2ª ed., Academic Press, 1985.
[23] R. Larson. *Elementary linear algebra*, 7ª ed., Brooks/Cole-Cengage Learning, 2013.
[24] R. Larson, D. Falvo. *Fundamentos del álgebra lineal*, 7ª ed., Cengage Learning, 2015.
[25] D. Lay, *Álgebra lineal y sus aplicaciones*, 5ª ed., Pearson, 2016.
[26] D. Luzardo, A. J. Peña. *Historia del álgebra lineal hasta los albores del siglo XX,* Divulgaciones Matemáticas 14 (2006), 153-170.
[27] S. Mac Lane. *History of abstract algebra: Origin, rise, and decline of a movement,* in American Mathematical Heritage: Algebra and Applied Mathematics, Texas Tech University, Mathematics Series, 13 (1981), 3-35.
[28] J. J. Martínez. *La descomposición en valores singulares (SVD) y algunas aplicaciones*, La Gaceta de la RSME, 8 (2005), 795-810.
[29] L. Merino, E. Santos. *Álgebra lineal con métodos elementales*, Thomson, 2006.
[30] C. D. Meyer. *Matrix analysis and applied linear algebra*, Society for Industrial and Applied Mathematics, 2000.
[31] G. Nakos, D. Joyner. *Álgebra lineal con aplicaciones*, Thomson, 1999.
[32] W. K. Nicholson. *Álgebra lineal con aplicaciones*, 4ª ed., McGraw-Hill, 2003.
[33] J. Paradís, A. Malet. *Los orígenes del álgebra: de los árabes al renacimiento*, PPU, 1989.
[34] J. Paradís, J. Miralles, A. Malet. *El álgebra en el periodo renacentista. La recuperación de los clásicos griegos*, PPU, 1989.
[35] D. Poole. *Álgebra lineal. Una introducción moderna*, 3ª ed., Cengage Learning, 2011.
[36] V. V. Prasolov. *Problems and theorems in linear algebra*, American Mathematical Society, 1994.
[37] J. Rojo, *Álgebra lineal.* 2ª ed., McGraw-Hill, 2007.
[38] S. Roman. *Advanced linear algebra*, GTM 135, 3ª ed., Springer, 2008.
[39] G. Schay. *A concise introduction to linear algebra*, Birkhäuser, 2012.
[40] I. Stewart. *Historia de las matemáticas*, Crítica, 2008.

[41] J. Stillwell. *Mathematics and its history*, Springer, 2010.
[42] G. Strang. *Álgebra lineal y sus aplicaciones*, 4ª ed., Paraninfo, 2007.
[43] G. Strang. *Introduction to linear algebra*, 4ª ed., Wellesley Cambridge Press, 2009.
[44] S. Treil. *Linear algebra done wrong*, http://www. math.brown.edu/~treil/papers/LADW/LADW.html, 2017.
[45] B. L. van der Waerden. *On the sources of my book Modern Algebra*, Hist. Math. 2 (1975), 31-40.
[46] B. L. van der Waerden. *A history of algebra: From al-Khwärizmi to Emmy Noether*, Springer-Verlag, 1985.

Índice alfabético



Manuais

Serie de manuais didácticos

Últimas publicacions na colección

Manual de programación en Ensamblador: Unha achega teórico-práctica (2021)
Manuel José Fernández Iglesias, Martín Llamas Nistal, Luis Eulogio Anido Rifón, Juan Manuel Santos Gago e Fernando Ariel Mikic Fonte

Elaboración de TFG, TFM e Teses: Claves para o éxito (2021)
Laura Novelle López

Gestión del circulante. Una aplicación práctica para la PYME (2021)
Javier Lorenzo Paniagua, Pablo Cabanelas Lorenzo e Pedro González Santamaría

Las ecuaciones del océano: Teoría y problemas resueltos (2020)
Gabriel Rosón Porto

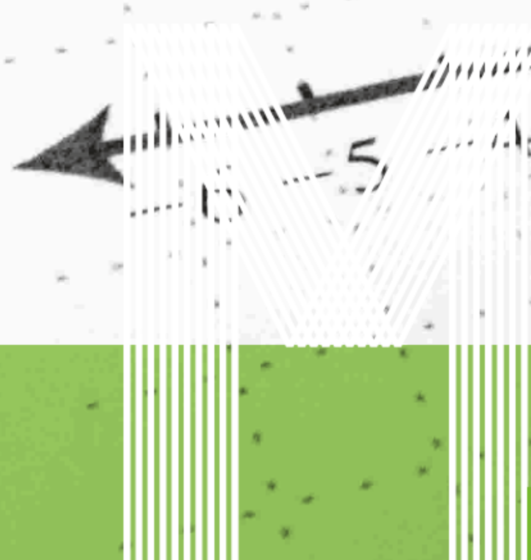
Design Thinking: Guía de iniciación (2020)
Manuel José Fernández Iglesias, Manuel Caeiro Rodríguez, Íñigo Cuiñas Gómez, Enrique Costa Montenegro, Francisco Javier Díaz Otero e Perfecto Mariño Espiñeira

9 7 8 8 4 8 1 5 8 9 1 9 1

Álxebra linear

Historia, teoría e práctica

É ben coñecido que a álxebra linear é unha materia que aparece nas memorias da maior parte dos graos de contido científico-técnico. Tendo en conta isto, esta obra está pensada para que o lector interesado poda afondar no seu estudo tendo a man una exposición rigorosa da teoría, unha boa serie de exemplos ilustrativos e unha ampla colección de problemas para practicar e afirmar o estudado. Partindo dunha breve introdución histórica, no segundo capítulo resúmense as propiedades básicas dos números reais e complexos. A continuación, no terceiro capítulo, defínense as nocións de matriz e rango que permitirán abordar no capítulo cuarto o estudo dos sistemas de ecuacións lineares. O quinto capítulo está adicado aos espazos vectoriais e ás aplicacións lineares deixando clara a íntima relación existente entre este tipo de aplicacións e as matrices. No sexto capítulo estúdase o problema de diagonalización, é dicir, saber cando una matriz cadrada é semellante a unha matriz diagonal. O capítulo sétimo está adicado ás formas bilineares e aos espazos vectoriais con produto escalar. Nel introdúcense a noción de ortogonalidade e finalízase examinando certas propiedades das matrices simétricas con coeficientes reais. Este tipo de matrices xunto coa diagonalización ortogonal utilízanse no capítulo oitavo para clasificar as formas cuadráticas reais. Finalmente, o último capítulo da obra adícase a expor a teoría de descomposición en valores singulares, á construcción da pseudoinversa dunha matriz e a súa aplicación á resolución de problemas de mínimos cadrados.



Servizo de Publicacións
Universidade de Vigo