Lliçó Inaugural de la Llicenciatura de Matemàtiques de la FME

"Sistemes Lineals: una aproximació a algun dels seus conceptes i problemes rellevants"

Dr. Ferran Puerta Sales

Curs 1997-98

UPC
Facultat de Matemàtiques i Estadística

SISTEMES LINEALS:

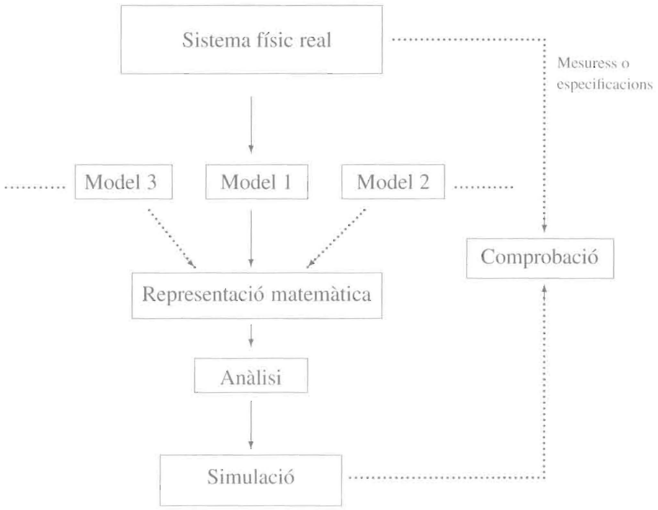
una aproximació a algun dels seus conceptes i problemes relevants

Voldria en primer lloc manifestar el meu agraïment i alhora la meva satisfacció, per l'oportunitat que em dóna la Facultat de Matemàtiques i Estadística "d'impartir" aquesta lliçó inaugural del curs 1997-98.
L'objectiu que m'he proposat al preparar aquesta exposició ha estat el de mostrar, per una banda, com alguns dels problemes bàsics de la teoria de Sistemes Dinàmics Lineals tenen una resposta senzilla en termes de l'Àlgebra lineal i, d'altra, com alguns problemes importants d'aquesta teoria s'entronquen amb conceptes i tècniques bàsiques de la geometria i la topologia.

Ferran Puerta

1 Introducció

D'una forma simple l'estudi dels sistemes físics reals pot esquematitzar-se a través del procés descrit en el diagrama adjunt



Si apareixen diversos models d'un sistema real es perquè, efectivament, un mateix sistema físic pot admetre diverses modelitzacions. Per exemple:

- Si volem calcular l'impuls i el combustible necessaris per posar un satèl·lit en òrbita podem identificar (modelitzar) el satèl·li com una *partícula* (especificada per la seva massa, posició i velocitat).
- Una vegada està en òrbita, per estudiar l'orientació de la seva antena, és necessari modelitzar el satèl·lit com un *cos rígid* movent-se al voltant del seu centre de masses.
- Si el satèl·lit és grand i si les maniobres es fan molt ràpidament potser s'hauria de modelitzar com una *estructura flexible* i tenir en compte els seus nodes de vibració elàstica.

Per tant, no existeix cap requeriment de que el model s'assembli, d'alguna manera, al sistema real: l'únic que es requereix és que, *per al problema que es consideri*, el model permeti fer prediccions útils amb un cost raonable.
Un model d'un sistema físic real s'anomenarà simplement un *sistema*.
Associat a un sistema existeix un conjunt de variables: voltatges elèctrics, corrents, forces mecàniques, desplaçaments, fluxes, temperatures,... que, en general, poden canviar amb el temps. Algunes d'aquestes variables poden ser modificades per tal de produir un determinat efecte sobre d'altres: les primeres són les *variables d'entrada* o estímuls o excitacions i les

segones *variables de sortida* o resposta. Sempre suposarem que hi ha un número finit d'aquestes variables.

El conjunt de variables d'entrada i sortida no estan definides de manera única. Tanmateix, un cop s'han escollit, tenim una situació definida, segons la qual el sistema rep uns estímuls o entrades i els transforma en unes respostes o sortides.

Les variables d'entrada les notarem habitualment per $u_1, \ldots, u_m$ i les de sortida per $y_1, \ldots, y_p$.

Les *variables d'estat* són un conjunt de variables que determinen de forma única l'estat del sistema, una vegada s'ha fixat un valor per a elles, en un instant t_0, així com les variables d'entrada $u_1(t), \ldots, u_m(t)$, per a $t \geq t_0$. Suposarem també que són un número finit i les designarem habitualment per $x_1, \ldots, x_n$.

Els sistemes que considerem són aquells que admeten una representació de la forma

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \qquad \text{(EE)} \\ y(t) &= Cx(t) \qquad \text{(ES)} \end{aligned}$$

on A, B i C són matrius reals de tipus $n \times n$, $n \times m$ i $p \times n$, respectivament. En aquesta representació $x(t)$, $u(t)$ i $y(t)$ són funcions vectorials que representen l'*estat* del sistema, l'*entrada* o *control* i la *sortida* o *resposta*, respectivament. Observis que la primera igualtat (EE) és una equació diferencial matricial que descriu l'evolució de l'estat del sistema $x(t)$ sotmès a excitacions externes $u(t)$ i s'anomena l'*equació d'estat del sistema*, en tant que la segona (ES) és una simple igualtat algebraica que diu quina és la sortida en funció de l'estat i que conseqüentment s'anomena *equació entrada/sortida* (per simplificar i perquè en la majoria de casos és així, estem suposant que la sortida no depèn del control $u(t)$). Suposarem que l'aplicació $u(t)$ verifica condicions suficients de continuïtat perquè es pugui aplicar a (EE) el teorema d'existència i unicitat de solucions, fixades les condicions inicials. De forma més explícita, la representació anterior ens diu que fixat un estat inicial x_0 en $t = 0$ per cada funció $u(t)$, l'equació EE ens determina una única funció $x(t)$ tal que $x(0) = x_0$ i l'equació ES ens dóna per cada $x(t)$ la resposta o sortida $y(t)$ del sistema. Naturalment, si variem $u(t)$, les funcions $x(t)$ i $y(t)$ variaran de forma corresponent. És ben sabut que en aquest cas en què les matrius són constants, les solucions estan definides en qualsevol interval real i venen donades per,

$$x(t) = e^t A x_0 + \int_0^t e^{t-\tau A} Bu(\tau)\, d\tau \,.$$

Per brevetat designem aquesta solució per $\varphi(t, 0, x_0, u(t))$. Tal com hem dit anteriorment anem a considerar alguns problemes qualitatius de la representació matemàtica anterior. Denotarem per Σ el sistema representat per les equacions anteriors. El fet que Σ vingui representat per aquest tipus d'equacions on les matrius A, B i C són constants es tradueix dient que Σ és un *sistema lineal invariant en el temps*.

2 Controlabilitat

El problema que ens plantegem és el següent: Fixats, no solament l'estat inicial x_0, sinó també un estat final x_1 al cap d'un temps t_1, ens preguntem si existeix algun control $u(t)$ tal que la solució corresponent $x(t)$ de EE amb aquest control que verifiqui $x(t_0) = x_0$, *també verifiqui que* $x(t_1) = x_1$. Si això succeeix, per a qualsevol estat inicial i per a qualsevol estat final es diu que el sistema Σ o que el parell (A, B) és *controlable*. (Observis que en aquest problema l'equació ES no té cap intervenció).

En el cas que estem suposant on A i B són matrius constants, es tracta doncs de veure si l'equació

$$x_1 - e^{t_1 A} x_0 = \int_0^{t_1} e^{(t_1 - \tau)} Bu(\tau)\, d\tau$$

té solució en $u(\tau)$, qualssevol que siguin $x_0, x_1 \in \mathbb{R}^n$ i $t_1 > 0$. La resposta és particularment simple i de *naturalesa algebraica*. En efecte, es té el teorema següent.

Teorema 2.1 *El sistema Σ és controlable si, i només si,*

$$\operatorname{rang}\begin{pmatrix} B & AB & \ldots & A^{n-1} & B \end{pmatrix} = n\,.$$

Aleshores, la funció de control fent $t_0 = 0$ ve donada per

$$u(t) = B^T e^{-A^T} z$$

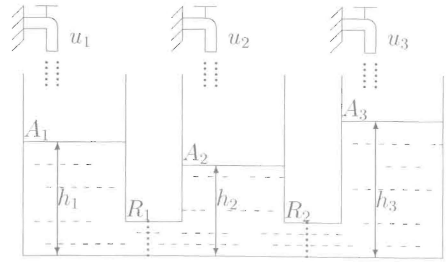
on z és l'única solució de l'equació

$$\left(\int_0^{t_1} e^{A\tau} B B^T e^{-A^T\tau} d\tau \right) z = e^{-t_1 A x_1 - x_0}\,.$$

La matriu $\begin{pmatrix} B & AB & \ldots & A^{n-1} & B \end{pmatrix}$ *s'anomena* matriu de controlabilitat de Σ.

Observis que per a aquests sistemes en què A i B són matrius constants, la controlabilitat del sistema no depèn de t_0 ni de t_1.

Ejemplo 2.2 Mostrem una aplicació d'aquest teorema.
Considerem el sistema físic (recordis el model!) de la figura



on u_i són els cabals, A_i l'àrea de les seccions transversals dels dipòsits, h_i les altures de líquid, $1 \le i \le 3$ i R_j, $1 \le j \le 2$ és una constant de resistència al pas de líquid entre dos dipòsits. Amb hipòtesis físiques convenients, les equacions de continuïtat queden en la forma

$$\left.\begin{aligned} A_1\frac{dh_1}{dt} &= -\frac{h_1-h_2}{R_1} + u_1(t) \\ A_2\frac{dh_2}{dt} &= \frac{h_1-h_2}{R_1} - \frac{h_2-h_3}{R_2} + u_2(t) \\ A_3\frac{dh_3}{dt} &= \frac{h_2-h_3}{R_2} + u_3(t) \end{aligned}\right\}$$

Si prenem com variables d'estat les altures, és a dir, si fem $h_i = x_i$, $1 \le i \le 3$ i suposem, per exemple, $A_2 = 1$, $A_1 = A_3 = 2/3$ i $R_1 = R_2 = 1/2$ les equacions anteriors queden

$$\dot{x} = \underbrace{\begin{pmatrix} -3 & 3 & 0 \\ 2 & -4 & 2 \\ 0 & 3 & -3 \end{pmatrix}}_{A} x + \underbrace{\begin{pmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{pmatrix}}_{B} u$$

on $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ i α_1, $1 \le i \le 3$ són constants positives. El problema de la controlabilitat d'aquest sistema consisteix en determinar si fixades les altures en $t = 0$, $x(0)$ i en $t = T > 0$, $x(T)$, existeixen cabals u_i tals que al cap del temps T les altures que inicialment eren $x(0)$, al final siguin $x(T)$. La resposta la dóna el teorema anterior: Es comprova fàcilment que rang $(B \quad AB \quad A^2B) = 3$ i per tant la resposta és sí. (Naturalment, en un plantejament *teòric* en què les altures finals fossin *més petites* que les inicials, les funcions $u_i(t)$ podrien prendre valors negatius amb una evident interpretació física).
En aquest mateix exemple, suposem que només funciona una de les aixetes i ens fem la mateixa pregunta que abans. Si, per exemple, només funcions u_1, la matriu B queda $\begin{pmatrix} \alpha_1 \\ 0 \\ 0 \end{pmatrix}$ i la matriu de controlabilitat continua rang 3, per tant *és possible controlar el sistema només amb l'aixeta* u_1.
Per contra, es comprova fàcilment que no és possible fer-ho només amb l'aixeta u_2 ja que si $B = \begin{pmatrix} 0 \\ \alpha_2 \\ 0 \end{pmatrix}$, la matriu de controlabilitat té rang 2.

Si el sistema no és controlable ja no és possible controlar las trajectòries $x(t)$ perquè puguin assolir un estat final arbitrari al cap d'un temps fixat, també de forma arbitrària. Ens podem preguntar si existeix un subespai de l'espai d'estats $\mathbb{R}^n$ on aquella propietat segueixi sent certa. La primera condició que cal exigir és que si l'estat inicial pertany al subespai es pugui controlar

la trajectòria de forma que tota ella romangui en el subespai. Això condueix a la definició de *subespai* (A, B)*-invariant* que donem a continuació:
Un subespai S de $\mathbb{R}^n$ es diu (A, B)-invariant si per a tot $x_0 \in S$ existeix un control $u(t)$ tal que si $x(t) = \varphi(t, 0, x_0, u(t))$ aleshores $x(t) \in S$ per a tot $t \ge 0$.

Aquests subespais tenen, també, una caracterització algebraica simple que recollim en la proposició següent.

Proposición 2.3 *Un subespai S de $\mathbb{R}^n$ és (A, B)-invariant si, i només si, existeix $F \in M_{m\times n}(\mathbb{R})$ tal que*

$$(A + BF)S \subset S\,.$$

Observis que, en particular, els subespais A-invariants són (A, B)-invariants ($F = 0!$) Designem Inv (A, B) el conjunt de subespais (A, B)-invariants.
Un d'aquests subespais A-invariants té la propietat que requeríem al principi d'aquesta discussió. Es té, en efecte, el resultat següent.

Proposición 2.4 *Sigui K la matriu de controlabilitat de (A, B). Aleshores* Im K *és un subespai A-invariant tal que per a tot $x_0, x_1 \in$* Im K *i $T > 0$ existeix un control $u(t)$ tal que si $x(t) = \varphi(t, 0; x_0, u(t))$ aleshores $x(t) \in$* Im K *per a tot $t \ge 0$ i $x(T) = x_1$.*

K rep el nom de *subespai de controlabilitat* de Σ.

Més generalment existeix un subconjunt $Ctr(A, B)$ de Inv (A, B) tal que cadascun dels seus elements gaudeix d'aquesta propietat; és a dir: $S \in Ctr(A, B)$ *si per a tot $x_0, x_1 \in S$ i $T > 0$ existeix un control $u(t)$* tal que si $x(t) = \varphi(t, 0, x_0, u(t))$ *aleshores $x(t) \in S$ per a tot $t \ge 0$ i $x(T) = x_1$. El subespai S es diu que és un subespai de controlabilitat de Σ o de (A, B)* .

Com en el cas anterior existeix una caracterització algebraica d'aquests subespais l'enunciat de la qual és una mica més sofisticat que l'anterior i l'ometem. De fet en la literatura habitual de la teoria de control se sol donar com definició pel que fa als subespais (A, B)-invariants com als de controlabilitat, la caracterització algebraica esmentada, ja que és a partir d'aquesta que s'estudien més còmodament les seves propietats.
Aquesta discussió obra un línia d'estudi interessant dintre de la teoria de sistemes lineals. En efecte, per estudiar Inv (A, B) i $Ctr(A, B)$ el primer que fem és considerar la seva partició natural per dimensió dels subespais. És a dir, considerem els subconjunts respectius Inv $_d(A, B)$ i $Ctr_d(A, B)$ formats pels seus subespais de dimensió d. Es tenen aleshores les inclusions

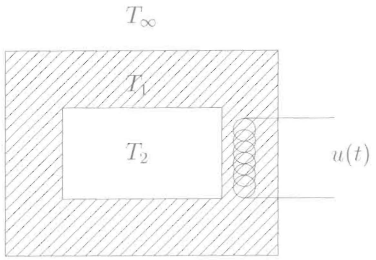
$$Ctr_d(A, B) \subset \text{Inv}\,_d(A, B) \subset Gr_d(\mathbb{R}^n)$$

on $Gr_d(\mathbb{R}^n)$ és la varietat grassmaniana de subespais d- dimensionals de $\mathbb{R}^n$. Sorgeixen aleshores diverses preguntes relatives a la seva estructura diferenciable, algebraica, topològica,... D'algunes d'elles es coneix la resposta, d'altres encara no.

Per exemple, Inv $_d(A,B)$ i $Ctr_d(A,B)$ són subvarietats estratificades de $Gr_d(\mathbb{R}^n)$, de les quals es coneix explícitament la dimensió de cada estrat. Això és interessant ja que, per exemple, permet saber en quines condicions la dimensió és zero. Se sap que si substituïm $\mathbb{R}$ per $\mathbb{C}$, tant Inv $_d(A,B)$ com cadascun dels seus estrats són connexos, però s'ignora si $Ctr_d(A,B)$ té aquesta propietat. Així mateix s'ignora la naturalesa algebraica d'aquestes varietats.

3 Observabilitat

La determinació directa dels estats d'un sistema pot resultar molt complicada o impossible. Per exemple, considerem un forn representat per la figura adjunta



on la temperatura interior T_2 es controla variant la potència tèrmica $u(t)$ aportada a través de les parets del forn, mitjançant una resistència elèctrica. Suposant una dificultat o impossibilitat de mesurar directament T_2, ens pregunten si és possible conèixer T_2 en un instant t_0, a partir del coneixement de $u(t)$ i T_1 durant un interval $[t_0, t_1]$. Com veurem, amb hipòtesis físiques convenients això és possible, és a dir, podem determinar l'estat del sistema en un instant si coneixem el control i la sortida al llarg d'un interval. Direm aleshores que el sistema és observable (en t_0). De forma general, es diu que el sistema Σ és *observable* en t_0 si per qualsevol estat inicial $x_0 = x(t_0)$, existeix $t_1 > t_0$ tal que x_0 queda *determinat de forma única* pel coneixement de $u(t)$ i $y(t)$ per tot $t \in [t_0, t_1]$.

Novament aquest problema té una solució senzilla per als sistemes que estem considerant.
Es té, en efecte, el resultat següent.

Teorema 3.1 *El sistema Σ és* observable *en t_0 si, i només si,*

$$\operatorname{rang}\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} = n$$

i, en aquest cas, l'estat inicial x_0 ve donat per (suposant $t_0 = 0$)

$$x_0 = M(t_1)^{-1}\int_0^{t_1} e^{-\tau A^T} C^T y(\tau)\, d\tau$$

on $M(t_1) = \displaystyle\int_0^{t_1} e^{\tau A^T} C^T C e^{\tau A} d\tau.$

La matriu $\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix}$ s'anomena *matriu d'observabilitat* del sistema. Observis que la condició d'observabilitat només afecta les matrius A_i. És per això que el parell (C, A) es diu *observable* si verifica la condició del teorema.
Anàlogament al que hem vist anteriorment, observis que per als sistemes que estem estudiant la condició d'observabilitat és independent del temps.
Si tornem a l'exemple del forn el primer que cal fer és obtenir la representació matemàtica del model. Amb hipòtesis físiques convenients es demostra que si considerem com variables d'estat les diferències de temperatures $x_1 = T_1 - T_\infty$, $x_2 = T_2 - T_\infty$, aquest sistema es representa per

$$\begin{aligned} \dot{x}(t) &= \begin{pmatrix} -\dfrac{a_1h_1 + a_2h_2}{c_1} & \dfrac{a_1h_1}{c_1} \\ \dfrac{a_1h_1}{c_2} & -\dfrac{a_1h_1}{c_2} \end{pmatrix} x(t) + \begin{pmatrix} \dfrac{1}{c_1} \\ 0 \end{pmatrix} u(t) \\ y(t) &= (1 \quad 0)x(t) \end{aligned}$$

on c_1 és la capacitat calorífica de les parets del forn, c_2 la del seu interior, a_1 i a_2 les àrees de les superfícies interior i exterior i h_1, h_2 els coeficients respectius de transferència tèrmica d'aquestes superfícies.
La matriu d'observabilitat d'aquest sistema és

$$\begin{pmatrix} 1 & 0 \\ -\dfrac{a_1h_1 + a_2h_2}{c_1} & \dfrac{a_1h_1}{c_1} \end{pmatrix}$$

que, clarament té rang 2. Per tant el sistema és observable, és a dir, mesurant $u(t)$ i $y(t) = x_1(t)$ al llarg d'un interval $[t_0, t_1]$ podem deduir l'estat en t_0 i per tant la temperatura T_2 en aquest instant.

Malgrat que entre les definicions geomètriques donades de controlabilitat i observabilitat d'un sistema lineal no és aparent que existeixi cap relació anem a veure que, pel contrari, es poden reduir una a l'altre a través del sistema dual que tot seguit introduïm.

S'anomena *sistema dual* de Σ el sistema lineal Σ_d definit per

$$\begin{aligned} \dot{x}(t) &= A^T x(t) + C^T u(t) \\ y(t) &= B^T x(t) \end{aligned}$$

(A^T és la matriu transposada de A).
Es té aleshores, el interessant resultat següent.

Proposición 3.2 *Un sistema lineal Σ és observable si, i només si, el seu dual Σ_d és controlable.*

Observis que amb aquest resultat el teorema 3.1 resulta immediatament del teorema 2.1.

Pel que fa a l'observabilitat, acabem dient que quan el sistema no és observable es poden plantejar qüestions similars a les que hem introduït al parlar de sistemes no controlables. En particular, es defineixen els subespais (C, A)-invariants simplement per dualitat: Un subespai S de $\mathbb{R}^n$ és (C, A)-invariant si $S^\perp$ és $(A^T\ C^T)$-invariant. Aleshores s'introdueixen les varietats Inv (C, A) (resp. Inv $_d(C, A)$) formada pels subespais (C, A)-invariants (resp. (C, A)-invariants de dimensió d) que condueixen a resultats i problemes anàlegs als que hem comentat pels subespais (A, B)-invariants.

Un problema particularment interessant on la varietat Inv $_d(C, A)$ juga un paper important és el conegut com "Disturbance Decoupling Problem (DDP)", la formulació del qual és senzilla i donem a continuació.
Suposem que l'entrada u del sistema es descompon en dos parts que representem per $\binom{u}{q}$ on q és una pertorbació no desitjable però que no podem evitar. Les equacions del sistema prenen aleshores la forma

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + Qq(t) \\ y(t) &= Cx(t) \end{aligned}$$

El DDP consisteix en trobar, si és possible, una realimentació d'estat, és a dir una nova entrada $\bar{u} = u - Fx$ de tal manera que la pertorbació $q(t)$ no tingui cap influència a la sortida $y(t)$. És a dir es tracta de trobar F de manera que la sortida $y(t)$ del sistema representat per

$$\begin{aligned} \dot{x}(t) &= (A + BF)x(t) + Bu(t) + Qq(t) \\ y(t) &= Cx(t) \end{aligned}$$

no depengui de $q(t)$. La solució d'aquest problema la proporciona el resultat següent.

Proposición 3.3 *Denotem per $\mathcal{J}(A, B; \ker C)$ l'únic subespai (A, B)-invariant maximal contingut a $\ker C$. El DDP té solució si, i només si*

$$(*) \qquad \mathcal{J}(A, B; \ker C) \supset \text{Im } Q\,.$$

Aleshores té interès considerar per una Q *genèrica* la màxima dimensió de Im C de forma que es verifiqui la inclusió $(*)$. Per dualitat aquest problema és equivalent al de trobar la mínima dimensió de $\ker C$ tal que $(\ker C)^\perp$ sigui un subespai (B^T, A^T)-invariant de màxima dimensió contingut en $(\text{Im } Q)^\perp$. I això a la seva vegada és equivalent a determinar el màxim enter d tal que la intersecció

$$\text{Gr}_d(F) \cap \text{Inv}_\mu(B^T, A^T)$$

no sigui buida, on F=(Im $Q)^\perp$ i Inv $_\mu(B^T, A^T)$ és l'estrat de màxima dimensió de Inv $_\mu(B^T, A^T)$. Clarament apareix aquí la conveniència de saber si els estrats d'aquesta varietat són varietats algebraiques. Com que això no és sabut en l'actualitat, cal abordar aquest problema des d'un altre punt de vista. Per exemple, a través de l'obtenció de sistemes de coordenades de Inv $_\mu(B^T, A^T)$. Això és possible i condueix a una formulació de la solució més complexa, encara que més explícita ja que permet obtenir no solament la dimensió d sinó també una parametrització del conjunt de varietats lineals Im C.

4 Realització

En la representació del sistema Σ a través de les equacions

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned}$$

la sortida del sistema $y(t)$ s'obté a partir del coneixement de l'estat $x(t)$. És possible no obstant obtenir una relació directa entre la entrada u i la sortida y. En efecte, si suposem el sistema inicialment en repòs, és a dir $x(0) = 0$, aplicant la transformació de Laplace a Σ obtenim fàcilment la relació

$$\widehat{y}(s) = C(sI_n - A)^{-1}B\widehat{u}(s)$$

on $\widehat{u}(s)$ i $\widehat{y}(s)$ són les corresponents transformades de Laplace de $u(t)$ i $y(t)$.

La matriu $p \times m$, $H(s) = C(sI_n - A)^{-1}B$ s'anomena *matriu de transferència* del sistema. Es tracta d'una matriu amb coeficients funcions racionals en s, el numerador de les quals té grau estrictament inferior al denominador; això s'expressa escrivint $H(\infty) = 0$. En la pràctica, succeeix sovint que la descripció matemàtica del sistema en termes d'equacions diferencials no és coneguda, però en canvi $H(s)$ por determinar-se a partir de mesures experimentals o per altres consideracions. Aleshores resulta útil trobar un sistema en la forma habitual que hem vingut considerant, la matriu de transferència del qual sigui precisament $H(s)$. Això ens condueix a la definició de realització.
Si $H(s) \in M_{p\times n}(\mathbb{R}(s))$ és tal que $H(\infty) = 0$, direm que $H(s)$ és una *matriu racional pròpia*. Aleshores s'anomena *realització* de $H(s)$ tota terna (A, B, C) que verifiqui la igualtat

$$C(sI_n - A)^{-1}B = H(s)\,.$$

Observis que C ha de tenir p files i B m columnes, però que el nombre de files i columnes de A, n no queda determinat.
Es demostra que aquest problema té solució i que aquesta no és única. Naturalment, es planteja aleshores el problema de determinar *una realització que tingui el menor nombre de variables d'estat*. Una tal realització s'anomena *realització minimal* i ve caracteritzada per la proposició següent

Proposición 4.1 *Una realització (A, B, C) de $H(s)$ és minimal si, i només si (A, B) és controlable i (C, A) és observable.*
El nombre de variables d'estat d'una realització minimal de $H(s)$ s'anomena grau de MacMillan *de $H(s)$.*

Les realitzacions minimals no són úniques, però estan relacionades d'una forma molt interessant, tal com es posa de manifest en el resultat següent

Proposición 4.2 *Dues ternes (A, B, C) i (A', B', C') són realitzacions minimals de $H(s)$ si, i només si, existeix una matriu invertible S tal que*

$$A' = S^{-1}AS\,, \qquad B' = S^{-1}B \qquad C' = CS\,.$$

En general, dos sistemes lineals Σ i Σ' definits per ternes (A, B, C) i (A', B', C'), respectivament, es diuen *algebraicament equivalents* si les matrius que els defineixen verifiquen les igualtats anteriors. Aquestes igualtats són, de fet, la traducció d'una important relació entre els sistemes de la que ens ocuparem en la secció següent. En efecte, és senzill demostrar que Σ *i* Σ' són *algebraicament equivalents si, i només si, Σ' s'obté a partir de Σ mitjançant un canvi lineal de variables d'estat*, és a dir, si existeix $S \in \mathrm{G}\ell(n, \mathbb{R})$ tal que $x = Sx'$.

Sembla raonable "identificar"dos sistemes lineals Σ i Σ' que siguin algebraicament equivalents. De forma precisa això es pot formular de la manera següent. Denotem per

$$\widetilde{\Sigma}_{n,m,p} = \{(A, B, C);\ (A, B) \text{ és controlable i } (C, A) \text{ és observable}\}$$

i fem operar el grup lineal $\mathrm{G}\ell(n)$ sobre $\widetilde{\Sigma}_{n,m,p}$ mitjançant

$$\alpha : \mathrm{G}\ell(n) \times \widetilde{\Sigma}_{n,m,p} \longrightarrow \widetilde{\Sigma}_{n,m,p}$$

$$\alpha(S, (A, B, C)) = (S^{-1}AS, S^{-1}B, CS)\,.$$

L'espai d'òrbites $\{\alpha(S, (A, B, C));\ S \in \mathrm{G}\ell(n)\} := \widetilde{(A, B, C)}$ el denotem $\Sigma_{n,m,p}$ i cada òrbita representa un sistema lineal, mòdul un canvi lineal de les variables d'estat. Aquest espai d'òrbites $\Sigma_{n,m,p}$ es pot dotar d'una estructura de varietat diferenciable. Un fet remarcable d'aquesta estructura ens el dóna la proposició següent.

Proposición 4.3 *La projecció natural $\pi : \widetilde{\Sigma}_{n,m,p} \longrightarrow \Sigma_{n,m,p}$ és un $\mathrm{G}\ell(n)$-fibrat principal que és trivial si, i només si,* $\min(m, p) = 1$.

Aleshores si designem per $\mathrm{Rat}_{\,n,m,p}(\mathbb{R})$ el conjunt de matrius racionals pròpies, el grau de McMillan de les quals és n, les proposicions 4.1 i 4.2 ens permeten concloure el resultat següent.

Proposición 4.4 *L'aplicació*

$$\rho : \mathrm{Rat}_{\,n,m,p}(\mathbb{R}) \longrightarrow \Sigma_{n,m,p}$$

definida per $\rho(H(s)) = \widetilde{(A, B, C)}$ on (A, B, C) és una realització minimal de $H(s)$ està ben definida i és un homeomorfisme.

És clar, aleshores, que aquesta bijecció permet traslladar propietats topològiques d'una varietat a l'altre.

5 Formes canòniques

L'acció de grups juga un paper important en la teoria de sistemes lineals. Per exemple, en diversos problemes de control i estimació apareixen transformacions lineals de coordenades en els espais d'estats, d'entrada o de sortida, realimentacions lineals d'estat i injeccions de sortida. Aquestes transformacions es poden descriure com l'acció de determinats grups de matrius sobre la terna de matrius (A, B, C) que defineix el sistema lineal Σ.

Per exemple, denotem per conveniència les ternes Σ que defineixen els sistemes lineals en forma de matrius per blocs, $\Sigma = \left(\begin{smallmatrix} A & B \\ C & O \end{smallmatrix}\right)$ i per $\widetilde{\Sigma}_{n,m,p}$ el conjunt de totes elles (sense hipòtesis de controlabilitat o observabilitat). Considerem el conjunt $\mathcal{G}$ de parelles de matrius per blocs (M, N) on $M = \left(\begin{smallmatrix} T & J \\ O & W \end{smallmatrix}\right)$, $N = \left(\begin{smallmatrix} T & O \\ K & V \end{smallmatrix}\right)$ amb T, W i V matrius invertibles $n \times n$, $p \times p$ i $m \times m$, respectivament. Aleshores, aquest conjunt $\mathcal{G}$ amb el producte definit per

$$(M, N)(M', N') = (MM', NN')$$

és un grup $\mathcal{G}$ i si definim una aplicació $\alpha : \mathcal{G} \times \widetilde{\Sigma}_{n,m,p} \longrightarrow \widetilde{\Sigma}_{n,m,p}$ mitjançant

$$g * \Sigma := \alpha(g, \Sigma) = M\Sigma N^{-1}\,,$$

on $g = (M, N)$, resulta que la imatge de Σ per la composició de les transformacions anteriors és precisament $M\Sigma N^{-1}$.
L'acció d'aquest grup produeix una partició de $\widetilde{\Sigma}_{n,m,p}$ en classes d'equivalència o *òrbites*: Σ_1 i Σ_2 pertanyen a la mateixa òrbita si $\Sigma_2 = g * \Sigma_1$ per a algun $g \in \mathcal{G}$.
Un dels problemes que es presenta aleshores és caracteritzar aquestes òrbites a través d'un conjunt d'invariants. Més precisament, es tracta de trobar una funció (no té perquè ser única) $\Gamma : \widetilde{\Sigma}_{n,m,p} \longrightarrow \widetilde{\Sigma}_{n,m,p}$ tal que

(i) Σ i $\Gamma(\Sigma)$ pertanyen a la mateixa òrbita, és a dir, $\Gamma(\Sigma) = g * \Sigma$ per a algun $g \in \mathcal{G}$.

(ii) $\Gamma(\Sigma_1) = \Gamma(\Sigma_2)$ si, i només si, $\Sigma_2 = g * \Sigma_1$ per a algun $g \in \mathcal{G}$.

Aleshores $\Gamma(\Sigma)$ és un element de $\widetilde{\Sigma}_{n,m,p}$ que caracteritza l'òrbita definida per Σ i que s'anomena *forma canònica* de Σ i els elements que defineixen $\Gamma(\Sigma)$ constitueixen un *conjunt complet d'invariants* de l'òrbita.
La determinació efectiva d'aquesta forma canònica, o equivalentment del seu conjunt complet d'invariants, és un problema important de la teoria de sistemes lineals (per exemple, per a l'aplicació de la teoria d'Arnold per a l'estudi de deformacions locals). Problema que, naturalment, es planteja de forma anàloga per a d'altres accions de grups que resulten de considerar

les diverses transformacions que hem considerat al començament d'aquesta secció i que constitueixen casos particulars del cas general que estem comentant.
Continuant amb aquest cas notem que si designem per $\Sigma_{n,m,p}$ l'espai d'òrbites de $\widetilde{\Sigma}_{n,m,p}$ per l'acció de $\mathcal{G}$, l'aplicació Γ indueix una aplicació injectiva, que seguim denotant Γ,

$$\Gamma : \Sigma_{n,m,p} \longrightarrow \widetilde{\Sigma}_{n,m,p}$$

i, naturalment, tenim també una projecció natural

$$\pi : \widetilde{\Sigma}_{n,m,p} \longrightarrow \Sigma_{n,m,p}$$

a través de la qual es dota $\widetilde{\Sigma}_{n,m,p}$ de la corresponent estructura topològica inicial.
Notis que Γ és una secció de π, és a dir, és una aplicació (no necessàriament contínua!) tal que $\pi \circ \Gamma$=Id.

Es plantegen aleshores diversos problemes d'interès en teoria de sistemes relatius a Γ i π. Citem com més importants l'existència de Γ contínua i l'estudi de l'estructura topològica i geomètrica de $\Sigma_{n,m,p}$ (connexió, cohomologia, estructura algebraica, diferenciable, analítica,...)

Poc és conegut en el cas general considerat anteriorment, i força més en el cas particular, en què $J = 0$, $K = 0$, $W = I_p$ i $V = I_m$. Observis que, en aquest cas l'acció de $\mathcal{G}$ es redueix a canvis lineals en les variables d'estat ja que com es comprova fàcilment

$$g * (A, B, C) = (TAT^{-1}, TB, CT^{-1})$$

acció que ja hem considerat en l'apartat anterior.
Indiquem i comentem alguns d'aquests resultats. En tot el que segueix suposarem que el cos d'esclars és $\mathbb{C}$.

(i) El problema d'existència d'una forma canònica Γ pel cas general és equivalent al problema següent de classificació de parelles d'aplicacions lineals. Considerem dos espais vectorials de dimensió finita $\mathcal{Z}$, $\mathcal{X}$ i en el conjunt $\mathcal{L}(\mathcal{Z}, \mathcal{X})$ definim la relació d'equivalència $(f, g) \sim (f', g')$ si existeixen automorfismes φ de $\mathcal{Z}$ i ψ de $\mathcal{X}$ tals que $f' = \psi \circ f \circ \varphi^{-1}$ i $g' = \psi \circ g \circ \varphi^{-1}$. Aleshores el problema d'existència d'una forma canònica és equivalent al problema de classificació de parelles d'aplicacions lineals (f, g) per la relació d'equivalència anterior.

Aquest problema té una solució explicita que, gràcies a l'equivalència anterior, es pot obtenir per mètodes geomètrics. Observis que si $\mathcal{X} = \{0\}$, l'equivalència considerada es redueix a la relació de conjugació. Una forma canònica és en aquest cas ben coneguda: la forma de Jordan. Diguem que el cas $\mathcal{X} \neq \{0\}$ és notablement més complicat i afegim que si $\mathcal{Z} = \mathcal{X}$ i $\psi = \varphi$, la classificació de parelles d'endomorfismes per aquesta relació és un problema obert.

En resum: *Existeix una forma canònica per l'acció general considerada en aquest apartat.* Aquesta forma però *no és contínua.*

Poc o res es coneix relatiu a l'estructura topològica i geomètrica en aquest cas general.

(ii) Considerem finalment el cas particular comentat anteriorment. Aleshores i sota la hipòtesi de que el parell (A, B) *és controlable* es tenen els resultats següents.

Proposición 5.1 *1) $\Sigma_{n,m,p}$ és una varietat algebraica llisa.*

2) La projecció $\pi : \widetilde{\Sigma}_{n,m,p} \longrightarrow \Sigma_{n,m,p}$ és un $\mathrm{G}\ell(n, \mathbb{C})$-fibrat principal que és no trivial si $m > 1$.

3) No existeixen formes canòniques globals per a $m > 1$.

Si a més, suposem $C = 0$, és coneguda la homologia de $\Sigma_{n,n}$ $(:= \Sigma_{n,m,0})$. Més concretament es té la proposició següent.

Proposición 5.2 *Els grups d'homologia singular de $\Sigma_{m,n}$ són isomorfs als de la grassmaniana $\mathrm{Gr}_n(\mathbb{C}^{n+m-1})$, però no ho són els corresponents anells de cohomologia.*

6 Famílies de sistemes lineals

Si Λ denota un entorn obert de $O \in \mathbb{R}^\ell$ i $\{\Sigma\}$ denota el conjunt de sistemes lineals definits per ternes de matrius (A, B, C) de tipus $n \times n$, $n \times m$ i $p \times m$, respectivament, una *família de sistemes* parametritzada per Λ és una aplicació $F : \Lambda \longrightarrow \{\Sigma\}$. La família es diu contínua, diferenciable, analítica,... segons que ho sigui l'aplicació F.
Famílies de sistemes apareixen de forma natural en la teoria de sistemes lineals. Per exemple, en l'estudi de sistemes sotmesos a pertorbacions o a incertesa de tots o part dels paràmetres que defineixen les matrius del sistema, en l'estudi de l'existència de formes canòniques contínues, etc....
Una de les branques de les matemàtiques que més interaccionen amb l'estudi de famílies de sistemes és la geometria. Ho il·lustrarem amb un exemple.
Suposem que volem estudiar les pertorbacions locals d'una família de sistemes $\dot{x}(t) = A(\lambda)x(t) + B(\lambda)u(t)$ (ens limitem a considerar l'equació d'estat) és a dir d'una aplicació (que suposarem C^∞ o analítica) $F : \Lambda \longrightarrow M_n(\mathbb{C}) \times M_{n,m}(\mathbb{C})$ $(= \mathcal{M})$.
El mètode utilitzat és l'introduït per Arnold en l'estudi de pertorbacions de matrius quadrades. Així, es considera la partició de $\mathcal{M}$ en òrbites per la restricció a $\mathcal{M}$ del grup $\mathcal{G}$ introduït a la secció 4. És a dir, si seguim denotant per $\mathcal{G}$ aquesta restricció, l'acció sobre $\mathcal{M}$ ve definida per

$$g * (A, B) = P^{-1}(A, B)\begin{pmatrix} P & O \\ R & Q \end{pmatrix}$$

on hem escrit $P = T^{-1}$, $Q = V^{-1}$, $R = -V^{-1}KT^{-1}$. Com ja hem dit a la secció anterior cada òrbita té una forma canònica que anomenarem forma de Brunovsky-Kronecker.
Aquestes òrbites constitueixen una estratificació de l'espai $\mathcal{M}$ i l'estudi de les pertorbacions lineals del sistema donat es podria realitzar a partir de la descripció local d'aquesta estratificació que a la vegada pot obtenir-se a partir de la *deformació versal* del parell $(A(0), B(0))$. No obstant això, i com succeeix en el cas de matrius quadrades, la partició anterior de $\mathcal{M}$ en òrbites no és localment finita. Cal doncs considerar una nova estratificació de $\mathcal{M}$, en què cada estrat està format per la no numerable unió d'òrbites que tenen els mateixos invariants discrets però no necessàriament els continus. Aleshores, és important fer notar que la descripció local donada per la deformació versal permet provar que cada estrat és una varietat diferenciable.
Aquesta estratificació, que designarem per BK, indueix de forma natural una partició en l'espai de paràmetres Λ, que és coneguda amb el nom de *diagrama de bifurcació* de la família. El diagrama de bifurcació dóna informació precisa respecte de les propietats qualitatives dels sistemes de la família així com de l'efecte de pertorbacions locals dels paràmetres. Tanmateix, cal afegir alguna condició per poder assegurar que el diagrama de bifurcació és realment una estratificació. Això ens condueix a considerar només famílies *transverses* a l'estratificació BK. Per aquestes famílies es té una importat propietat: la co-dimensió d'un estrat en l'espai de paràmetres Λ és la mateixa que la co-dimensió del corresponent estrat de l'estratificació BK. Així doncs, per a una família transversa tenim una limitació precisa sobre les possibilitats de canvi de l'estructura local de la família.
Aleshores ens preguntem: Hi ha "moltes" famílies transverses a l'estratificació BK? És aquí on juguen un paper important les condicions de regularitat de Whitney. En efecte, és sabut que si es donen aquestes condicions "quasi" totes les famílies són transverses. És important doncs saber si l'estratificació BK satisfà aquestes condicions. Només es coneix una resposta parcial a aquest problema que és la que dóna el teorema següent.

Teorema 6.1 *Si $m = 1$, l'estratificació BK verifica les condicions de regularitat de Whitney.*

Amb aquest exemple finalitzem aquesta "lliçó" en què, tal com hem dit al començament, hem volgut posar de manifest l'entroncament de diversos problemes de la teoria de sistemes lineals amb conceptes i tècniques bàsiques de l'àlgebra lineal, de la geometria i la topologia. Moltes gràcies per la vostra atenció.