

UNIVERSIDAD DE EL SALVADOR  
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA  
ESCUELA DE MATEMÁTICA



TRABAJO DE GRADUACIÓN TITULADO:  
OPTIMIZACIÓN CONVEXA NO DIFERENCIABLE CON APLICACIONES  
EN CONTROL ÓPTIMO ESTOCÁSTICO

PRESENTADO POR:  
BR. LENNY IVONNE ARIAS RIVERA

PARA OPTAR AL TÍTULO DE:  
LICENCIADA EN MATEMÁTICA

ASESOR:  
MSc. PORFIRIO ARMANDO RODRÍGUEZ

CIUDAD UNIVERSITARIA, 12 DE DICIEMBRE DEL 2019

**UNIVERSIDAD DE EL SALVADOR  
AUTORIDADES UNIVERSITARIAS 2019 – 2023**

**MSc. ROGER ARMANDO ARIAS**  
RECTOR DE LA UNIVERSIDAD DE EL SALVADOR

**PhD. RAÚL ERNESTO AZCÚNAGA LÓPEZ**  
VICERRECTOR ACADÉMICO DE LA UNIVERSIDAD DE EL SALVADOR

**Ing. JUAN ROSA QUINTANILLA**  
VICERRECTOR ADMINISTRATIVO DE LA UNIVERSIDAD DE EL SALVADOR

**Ing. FRANCISCO ALARCÓN**  
SECRETARIO GENERAL DE LA UNIVERSIDAD DE EL SALVADOR

**Lic. RAFAEL HUMBERTO PEÑA MARÍN**  
FISCAL GENERAL

**Lic. LUIS ANTONIO MEJÍA LIPE**  
DEFENSOR DE LOS DERECHOS UNIVERSITARIOS

**FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA  
AUTORIDADES**

**Lic. MAURICIO HERNÁN LOVO  
DECANO**

**MSc. ZOILA VIRGINIA GUERRERO  
VICEDECANA**

**MSc. PORFIRIO ARMANDO RODRÍGUEZ  
DIRECTOR INTERINO DE ESCUELA DE MATEMÁTICA**

**MSc. CARLOS ERNESTO GÁMEZ RODRÍGUEZ  
SECRETARIO INTERINO DE ESCUELA DE MATEMÁTICA**

UNIVERSIDAD DE EL SALVADOR  
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA  
ESCUELA DE MATEMÁTICA



TRABAJO DE GRADUACIÓN TITULADO:  
OPTIMIZACIÓN CONVEXA NO DIFERENCIABLE CON APLICACIONES  
EN CONTROL ÓPTIMO ESTOCÁSTICO

---

Asesor: MSc. Porfirio Armando Rodríguez

CIUDAD UNIVERSITARIA, 12 DE DICIEMBRE DEL 2019

*Dedicado a mí, por tanta paciencia*

# Agradecimientos

A mis padres, por su amor, consejos, comprensión, por cumplir muchos de mis sueños y encaminarme en mi formación académica.

A mi prometido, por apoyarme, consentirme y ser incondicional durante tanto tiempo.

A mi hermana y sobrinos por su alegría y motivarme a ser mejor.

A mis amigos, que siempre me animaban a seguir adelante.

A mi asesor, por tomarse el tiempo y la disposición para guiarme en este trabajo.

Al jurado calificador, por su oportuna colaboración y evaluación.

... a Dios.

# Índice

|   |           |
|---|-----------|
| Índice  | I         |
| Índice de figuras   | III       |
| Resumen   | V         |
| Introducción  | VII       |
| Metodología   | IX        |
| <b>1 Fundamentos de Optimización Convexa</b>                  | <b>1</b>  |
| 1.1 Conjuntos convexos  | 1         |
| 1.1.1 Propiedades topológicas de los conjuntos convexos       | 6         |
| 1.1.2 Teorema de Weierstrass                                  | 10        |
| 1.2 Funciones convexas  | 15        |
| 1.3 Funciones convexas no diferenciables                      | 20        |
| 1.3.1 Subgradiente  | 20        |
| 1.3.2 Funciones convexas diferenciables                       | 25        |
| 1.3.3 Mínimo y máximo de funciones convexas                   | 30        |
| 1.4 Condiciones de optimalidad y dualidad                     | 34        |
| 1.4.1 Optimización con restricciones                          | 36        |
| 1.4.2 Problemas con restricciones de igualdad y desigualdades | 38        |
| 1.4.3 Dualidad Lagrangiana                                    | 41        |
| <b>2 Control Óptimo Estocástico</b>                           | <b>46</b> |
| 2.1 Introducción  | 46        |
| 2.2 Probabilidad y procesos estocásticos                      | 47        |
| 2.2.1 Procesos estocásticos                                   | 49        |
| 2.3 Planteamiento del problema de control óptimo              | 50        |
| 2.4 Resolución del PCO por métodos usuales                    | 53        |
| 2.4.1 El principio de máximo                                  | 53        |
| 2.4.2 Programación dinámica                                   | 55        |
| 2.4.3 Sistema lineal con objetivo cuadrático (LQ)             | 59        |
| 2.5 Problema de control estocástico LQ                        | 63        |
| 2.6 Principio de equivalencia cierta                          | 72        |
| 2.7 Equivalencia cierta para el problema LQR                  | 77        |

|          |  |            |
|----------|--|------------|
| <b>3</b> | <b>Modelamiento en Optimización Convexa No Diferenciable</b>                 | <b>81</b>  |
| 3.1      | El problema del portafolio . . . . .   | 81         |
| 3.1.1    | Formulación matemática del problema . . . . .                                | 83         |
| 3.2      | Maximización de la utilidad . . . . .  | 89         |
| 3.2.1    | Formulación matemática del problema . . . . .                                | 90         |
| 3.3      | Problema de localización de instalaciones capacitado en dos etapas . . . . . | 92         |
| 3.3.1    | Formulación al problema y relajaciones propuestas . . . . .                  | 93         |
| 3.4      | Recuperación de imágenes . . . . .   | 99         |
| 3.4.1    | Otra formulación equivalente . . . . .                                       | 102        |
| 3.5      | Fractales y finanzas . . . . .   | 105        |
| 3.5.1    | Planteamiento del problema . . . . .   | 106        |
| 3.6      | Control de inventario . . . . .  | 108        |
| 3.6.1    | Planteamiento del problema . . . . .   | 109        |
| 3.7      | Sistemas de Control en Red (NCS) . . . . .                                   | 111        |
| 3.7.1    | Formulación del problema . . . . .   | 112        |
| 3.7.2    | Sistemas de operador delta . . . . .   | 114        |
| 3.7.3    | formulación del problema en tiempo discreto . . . . .                        | 114        |
| <b>4</b> | <b>Métodos para optimización no diferenciable</b>                            | <b>116</b> |
| 4.1      | Método del subgradiente . . . . .  | 116        |
| 4.1.1    | Convergencia del subgradiente . . . . .                                      | 118        |
| 4.2      | Método de planos de corte . . . . .  | 120        |
| 4.2.1    | Convergencia del método planos de corte . . . . .                            | 121        |
| 4.3      | Método de haces . . . . .  | 122        |
| 4.3.1    | Regularización de Moreau Yosida . . . . .                                    | 122        |
| 4.3.2    | Convergencia del método de haces . . . . .                                   | 130        |
| 4.4      | Método del gradiente próximo . . . . .                                       | 134        |
| 4.4.1    | Puntos fijos . . . . .   | 136        |
| 4.4.2    | Convergencia del método del gradiente próximo . . . . .                      | 137        |
| <b>5</b> | <b>Ejemplos Numéricos</b>  | <b>140</b> |
| 5.1      | Subgradiente proyectado para el problema dual . . . . .                      | 140        |
| 5.2      | Procesamiento de señal adaptativo . . . . .                                  | 141        |
|          | <b>Recomendaciones</b>   | <b>144</b> |
|          | <b>Conclusiones</b>  | <b>145</b> |
| <b>6</b> | <b>Anexos</b>  | <b>146</b> |
| 6.1      | Algoritmo para subgradiente proyectado del problema dual . . . . .           | 146        |
| 6.2      | Algoritmo para el problema de procesamiento de señal adaptativo . . . . .    | 150        |
|          | <b>Bibliografía</b>  | <b>153</b> |



# Índice de figuras

|      |   |     |
|------|---|-----|
| 1.1  | Conjunto convexo y no convexo . . . . .   | 2   |
| 1.2  | Cápsula convexa . . . . .   | 3   |
| 1.3  | Politopo y simplex . . . . .  | 4   |
| 1.4  | Topología del conjunto $S = \{(x, y) : x^2 + y^2 \leq 1\}$ . . . . .  | 7   |
| 1.5  | Inexistencia de solución de minimización [2] . . . . .  | 10  |
| 1.6  | Ley del paralelogramo . . . . .   | 12  |
| 1.7  | Hiperplanos definidos por $\mathcal{H}$ . . . . .   | 12  |
| 1.8  | $H$ es el hiperplano separador entre $S$ e $y$ . . . . .  | 13  |
| 1.9  | $H_1$ y $H_2$ son dos hiperplanos soporte en el punto $\omega$ . . . . .  | 14  |
| 1.10 | Interpretación geométrica de función convexa en $\mathbb{R}$ [9] . . . . .  | 16  |
| 1.11 | Conjunto de nivel de una variable . . . . .   | 17  |
| 1.12 | Epígrafo de una función $f$ en $\mathbb{R}$ [12] . . . . .  | 21  |
| 1.13 | figura del ejemplo de subgradiente [2] . . . . .  | 23  |
| 1.14 | Interpretación geométrica del subgradiente en $\mathbb{R}$ [14] . . . . .   | 23  |
| 1.15 | Mínimo local y global [2] . . . . .   | 31  |
| 1.16 | El Teorema de KKT no es aplicable en el óptimo . . . . .  | 40  |
| 2.1  | Control óptimo en tiempo discreto . . . . .   | 52  |
| 2.2  | Propiedad de causalidad . . . . .   | 55  |
| 2.3  | Etapas y variables en el problema de control estocástico en tiempo discreto . . . . .   | 64  |
| 2.4  | Períodos y variables en el ejemplo 2.2 . . . . .  | 75  |
| 3.1  | Una combinación convexa de dos portafolios eficientes produce otro portafolio eficiente . . . . .   | 88  |
| 3.2  | Ejemplo de una imagen y su gráfica en escala de grises. . . . .   | 100 |
| 3.3  | Ejemplo de una imagen distorsionada y su escala de grises. . . . .  | 101 |
| 3.4  | Sistemas de control de red con retardo . . . . .  | 112 |
| 3.5  | Diagrama de señales . . . . .   | 113 |
| 4.1  | Operador próximo evaluado en varios puntos . . . . .  | 134 |
| 5.1  | Valores del límite inferior $gval$ y el límite superior $fval$ versus el número de iteración $k$ . Usamos un tamaño de paso fijo con $\alpha = 0.1$ . . . . . | 141 |
| 5.2  | Error de predicción $w^{(k)T}x^{(k+1)} - y^{(k+1)}$ vs. iteración $k$ . . . . .   | 142 |
| 5.3  | Distribución empírica del error de predicción para $w^*$ (más de 1000 muestras) . . . . .   | 143 |

# Resumen

Cuando hablamos en concreto de optimización convexa (cuya base teórica es el análisis convexo) nos referimos a minimizar funciones convexas reales definidas para una variable contenida dentro de un subconjunto convexo de un espacio afín. Pero con el paso del tiempo se encontró que muchos problemas de optimización convexa no eran diferenciables en el punto óptimo y aquí es donde surge la necesidad e importancia de estudiar *Optimización Convexa no Diferenciable* ya que debido a ella no solo se desarrolló nueva teoría matemática sino que también las técnicas abarcadas por este campo de estudio son importantes en aplicaciones de ingeniería, las cuales, también requieren de estudios estadísticos.

Este trabajo inicia haciendo memoria de conocimientos de optimización convexa lo cual nos permite llegar de forma natural a lo que son funciones convexas no diferenciables. Con ello llegamos finalmente a lo que es un subgradiente y las propiedades de éste, del cual se hace uso cuando la diferenciabilidad en el sentido clásico falla. En el primer capítulo se estudian teoremas que brindan condiciones bajo las cuales un problema de optimización convexo no diferenciable pasa a ser diferenciable manteniendo propiedades de convexidad, lo cual facilita en cierto modo la búsqueda de una solución óptima.

Muchos problemas de programación matemática incorporan parámetros que se suponen conocidos en el momento de resolver el problema. Sin embargo, si el problema de optimización es un modelo que representa una situación real, en la que se debe tomar una decisión, es frecuente que se desconozcan los valores de algunos parámetros que intervienen en el modelo. Es aquí donde nos encontramos con lo que es incertidumbre lo cual nos lleva a trabajar con problemas estocásticos. En el capítulo 2 presentamos la teoría necesaria para convertir un problema estocástico a uno determinístico por medio de el *principio de equivalencia cierta* y de esta forma, encontrar un óptimo es menos complicado de lo que era en un inicio.

En muchos problemas, la no diferenciabilidad puede tratarse transformando la parte que provoca la no diferenciabilidad, pero esto no garantiza que ésta desaparezca del todo, por lo que pueden emplearse los métodos presentados, los cuales resuelven en cada iteración un subproblema diferenciable, siempre teniendo en cuenta que las transformaciones hechas no son necesariamente únicas. El contenido del capítulo 3 nos presenta una serie de ejemplos donde ponemos en evidencia como tratar problemas que aún conservan la no diferenciabilidad pero sin dejar de lado las propiedades de convexidad, esto combinado con los conocimientos adquiridos con las propiedades de subgradiente, facilita en gran medida la búsqueda de una solución óptima.

Los métodos para resolver problemas convexos no diferenciables, varían dependiendo de la función que modele el problema de optimización, el capítulo 4 contiene alguno de los algoritmos más conocidos en el área de no diferenciability y las pruebas de convergencia de cada uno de éstos.

Para finalizar este trabajo el capítulo 5 presenta dos ejemplos los cuales se detallan y se presentan resultados de estos mismos, para ello se hizo uso de CVX, que es un sistema de modelado basado en MATLAB para la optimización convexa.

La investigación del tema que ha consistido en la búsqueda bibliográfica de libros y artículos en internet, se han seleccionado de acuerdo a los objetivos del trabajo y se presentan en cada capítulo conteniendo la información principal, conceptos, teoremas y demás resultados importantes, se concluye con la presentación del documento que se espera sea una base para posteriores trabajos que profundicen en el estudio de esta área.

# Introducción

La Matemática, como ciencia, vive y crece con el intercambio de ideas entre los cultivadores de diferentes áreas del conocimiento y el desarrollo del quehacer de la vida del ser humano y su entorno; y por su puesto alimentada por investigaciones que se desarrollan día a día en distintos espacios de la actividad científica (Véase [9])

La programación lineal fue planteada como un modelo matemático desarrollado durante la Segunda Guerra Mundial a fin de reducir los costos del ejército y aumentar las pérdidas del enemigo. Los fundadores de la técnica son George Dantzig, quien publicó el método simplex en 1947, John Von Neumann, que desarrolló la teoría de la dualidad en el mismo año y Leonid Kantoróvich, un matemático ruso que utilizó técnicas similares aplicadas a la economía antes de Dantzig y que ganó el premio Nobel en economía en 1975. (Véase [7])

En el área de investigación operativa la programación lineal estudia la minimización de una función lineal restringida a un conjunto definido por desigualdades lineales. Es una herramienta fundamental gracias a:

- La existencia de algoritmos eficientes de resolución como el simplex (a pesar de su complejidad no polinomial).
- La cantidad de problemas que entran dentro de este paradigma.
- La cantidad de problemas que sin ser lineales pueden ser aproximados o resueltos mediante sucesiones de éstos.

Sin embargo existen numerosos problemas en los que las herramientas lineales no son suficientes (Véase [13]) y que dan lugar al estudio de la programación no lineal. Por ejemplo:

- Problemas de mínimos cuadrados originados en la teoría de redes neuronales.
- Problemas de diseño estructural (mecánico, eléctrico, espacial).
- Problemas de control óptimo (lanzamiento cohete-satélite, planeamiento de producción).
- Problemas de control óptimo estocástico (operación de una represa, de una red de poliductos, portafolio de acciones).
- Problemas de ruteo en redes de transporte y telecomunicaciones.

Las funciones no diferenciables, marcaron un paradigma en la teoría moderna de optimización. Se encontró que muchos problemas de optimización convexa no eran diferenciables en el punto óptimo. Así, un enfoque completamente diferente se desarrolló, en donde fue concebida la noción de subdiferencial.

De esta forma los enfoques modernos de la teoría de optimización deben sus orígenes al cálculo de variaciones que ha sido estudiado por más de tres siglos y que también fue crucial en el desarrollo del análisis funcional.

Los métodos computacionales para la optimización no diferenciable se desarrollaron en dos direcciones:

- a) La investigación dirigida a resolver determinados tipos de problemas de minimización con funciones no diferenciables que tienen una estructura especial y que se define de forma explícita
- b) La investigación sobre la elaboración de métodos para resolver clases más generales de problemas, que no suponen de antemano el conocimiento de la estructura específica de la función a minimizar pero requieren la evaluación de la función y sus gradientes (o sus análogos en el caso no diferenciable) en cualquier punto dado.

Para el primer grupo, se cuenta con numerosos trabajos sobre métodos para resolver problemas de minimización. Para el segundo grupo una serie de obras se dedican a la minimización de funciones convexas lineales a trozos. Para la solución de varios problemas de optimización no diferenciable en los que se dan explícitamente los puntos donde la función no es diferenciable (por ejemplo, las funciones con valores absolutos), se han desarrollado técnicas especiales de suavizado.

En lo que se refiere a métodos generales de optimización no diferenciable, se pueden distinguir dos clases básicas en los cuales se requiere el cálculo del subdiferencial.

1. Métodos de gradiente generalizado. Para minimizar funciones diferenciables, se usan con frecuencia múltiples versiones del método de gradiente, que es natural, ya que la dirección negativa del gradiente en un punto dado es una dirección de descenso. La selección de un tamaño de paso en la mayoría de estos métodos tiene por objeto disminuir de manera significativa el valor de la función objetivo en cada iteración.
2. El método de planos de corte para resolver problemas convexos. El de Kelley por ejemplo se basa en aproximaciones lineales a la gráfica de una función convexa por medio de hiperplanos de soporte; en cada iteración el método resuelve un problema de programación lineal con un número creciente de restricciones en cada iteración.

# Metodología

A continuación se describen los aspectos importantes de la metodología de este trabajo.

1. **Tipo de investigación.** Este proyecto es de carácter bibliográfico aplicado.

1.1 **Bibliográfico:**

Se hará una recopilación de libros impresos, libros obtenidos en internet, tesis de grado y artículos en línea para contar con el material suficiente que cubra las necesidades del estudio y las que puedan surgir más adelante. El objetivo es compilar de forma coherente la información más útil y destacada de los resultados que se aplicarán en el tema en cuestión.

1.2 **aplicado:**

Ya que se pretende estudiar diversas aplicaciones con la teoría que se presentará en este texto y en base a ello hacer uso de las herramientas que nos proporcione ésta para finalmente *hacer uso de software* (MATLAB u Octave) y aplicar esto a los problemas que se pretendan estudiar y de esta manera tener una mejor visión de lo importante que es la optimización convexa y los beneficios que nos genera.

2. **Forma de trabajo:**

Se tendrán reuniones periódicas (una ó dos veces por semana) con el asesor para tratar los aspectos de la investigación tales como analizar y estudiar la teoría, tratar los diferentes aspectos del trabajo escrito y las presentaciones de avances.

3. **Exposiciones.**

Se tendrán dos exposiciones:

- **Primera exposición (pública):** Presentación del Perfil del Proyecto de Investigación.
- **Segunda exposición (pública):** Presentación final del Trabajo de Investigación: resumen de resultados y aplicaciones.

# Capítulo 1

## Fundamentos de Optimización Convexa

### 1.1 Conjuntos convexos

A lo largo de este proyecto,  $\mathbb{R}$  denotará el sistema de números reales y  $\mathbb{R}^n$  es el espacio euclídeo de  $n$ -uplas  $x = (x_1, x_2, \dots, x_n)$ . Se trabajará en  $\mathbb{R}^n$  al menos que se especifique lo contrario. El producto interno de dos vectores  $x, x^* \in \mathbb{R}^n$  está expresado como:

$$\langle x, x^* \rangle = x_1x_1^* + \dots + x_nx_n^*$$

El símbolo  $A$  es usado para denotar una matriz con entradas reales de  $m \times n$  y su correspondiente transformación lineal  $x \mapsto Ax$  de  $\mathbb{R}^n$  a  $\mathbb{R}^m$ . La matriz transpuesta y la correspondiente transformación lineal adjunta de  $\mathbb{R}^n$  a  $\mathbb{R}^m$  está denotada por  $A^*$  de modo que se tiene una identidad:

$$\langle Ax, y^* \rangle = \langle x, A^*x^* \rangle$$

**Definición 1.1.** Si  $x, y$  son puntos distintos en  $\mathbb{R}^n$ , el conjunto de puntos de la forma:

$$(1 - \lambda)x + \lambda y = x + \lambda(y - x), \quad \lambda \in \mathbb{R}$$

es llamado línea a través de  $x$  e  $y$ . Un subconjunto  $M$  de  $\mathbb{R}^n$  es llamado **conjunto afín** si

$$(1 - \lambda)x + \lambda y \in M \quad \forall x, y \in M \quad \lambda \in \mathbb{R}$$

El conjunto vacío  $\emptyset$  y el espacio  $\mathbb{R}^n$  son en sí ejemplos extremos de conjuntos afines. La definición también incluye el caso cuando  $M$  consiste de un único punto.

**Definición 1.2.** : Un conjunto  $C \subset X$  se dice convexo si para todo par de puntos  $x, y \in C$  se tiene

$$[x, y] := \{\alpha x + (1 - \alpha)y : \alpha \in [0, 1]\} \subset C$$

donde  $[x, y]$  es el segmento de recta que une  $x$  e  $y$ .

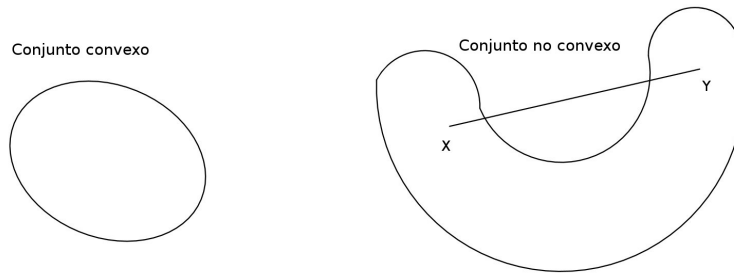


Figura 1.1: Conjunto convexo y no convexo

**Definición 1.3.** : La suma vectorial [2]

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_m x_m$$

es llamada **combinación convexa** de  $x_1, x_2, \dots, x_m$  si los coeficientes  $\lambda_i$  son no negativos y  $\sum_{i=1}^m \lambda_i = 1$

**Definición 1.4.** Sea  $S \in \mathbb{R}^n$  (no necesariamente convexo). Se define la **cápsula convexa** de  $S$  de la siguiente manera:

$$co(S) = \left\{ \sum_{i=1}^k x_i \lambda_i \mid k \in \mathbb{N}, x_1, \dots, x_k \in S, \lambda_1, \dots, \lambda_k \in \mathbb{R}^+, \sum_{i=1}^k \lambda_i = 1 \right\}$$

Es decir, que es el conjunto de todas las posibles combinaciones convexas de puntos de  $S$  y  $\mathbb{R}^+$  se refiere a los reales positivos incluyendo el cero.

**Observación 1.1.** Sean  $S, S_0 \subseteq \mathbb{R}^n$ , entonces:

- $S \subseteq co(S)$ .
- $S$  es convexo si y sólo si  $co(S) = S$ .
- Si  $S \subseteq S_0$  entonces  $co(S) \subseteq co(S_0)$

**Proposición 1.1.**  $co(S)$  es un conjunto convexo

**Demostración**

Sean  $x, y \in co(S)$ , es decir:



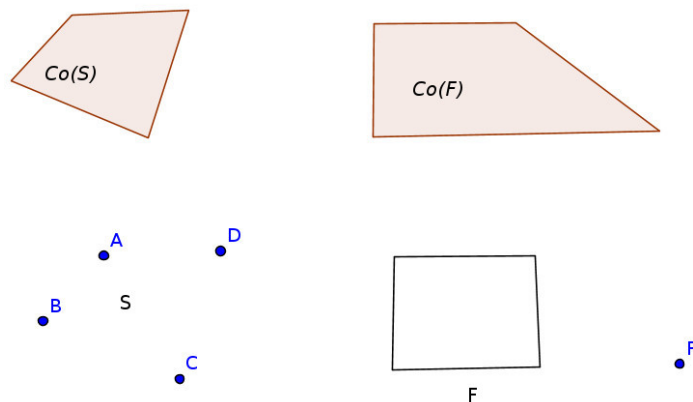


Figura 1.2: Cápsula convexa

$$x = \sum_{i=1}^k v_i x_i \quad y = \sum_{i=1}^m \mu_i y_i \quad (1.1)$$

donde  $x_1, \dots, x_k, y_1, \dots, y_m \in S$  y  $v_1, \dots, v_k, \mu_1, \dots, \mu_m$  son ponderadores de las combinaciones convexas.

$$\text{Sea } \lambda \in [0, 1] \text{ y } \lambda x + (1 - \lambda)y = \lambda \sum_{i=1}^k v_i x_i + (1 - \lambda) \sum_{i=1}^m \mu_i y_i$$

llamando  $z_i = x_i$ ,  $\alpha_i = \lambda v_i$ ,  $\forall i = 1, \dots, k$  y  $z_{k+i} = y_i$ ,  $\alpha_{k+i} = (1 - \lambda)\mu_i$   $\forall i = 1, \dots, m$  se tiene que  $\lambda x + (1 - \lambda)y = \sum_{i=1}^{k+m} \alpha_i z_i$  con

$$z_i \in S, \alpha_i \in [0, 1], \quad \forall i = 1, \dots, k + m \quad \text{y} \quad \sum_{i=1}^{k+m} \alpha_i = 1$$

Luego, por la definición (1.2) se tiene que  $\lambda x + (1 - \lambda)y \in \text{co}(S)$  es convexo.

□

**Proposición 1.2.** *El conjunto  $\text{co}(S)$  es el convexo más pequeño (en el sentido de la inclusión) que contiene a  $S$ , es decir,  $\text{co}(S) = \bigcap \{C \subseteq \mathbb{R}^n \mid C \text{ convexo}, S \subseteq C\}$*

### ***Demostración***

Sea  $x \in \bigcap \{C \subseteq \mathbb{R}^n \mid C \text{ convexo}, S \subseteq C\}$ . Entonces  $x \in C$ ,  $\forall C$  convexo, tal que  $S \subseteq C$ . Luego  $x \in \text{co}(S)$ , que es un convexo particular que contiene a  $S$ .

Sean ahora,  $x \in \text{co}(S)$  y  $C$  un convexo cualquiera que contiene a  $S$ . Entonces,  $\text{co}(S) \subseteq \text{co}(C) = C$ , por lo tanto,  $x \in C$ . Luego,  $x \in \bigcap \{C \subseteq \mathbb{R}^n \mid C \text{ convexo}, S \subseteq C\}$ .

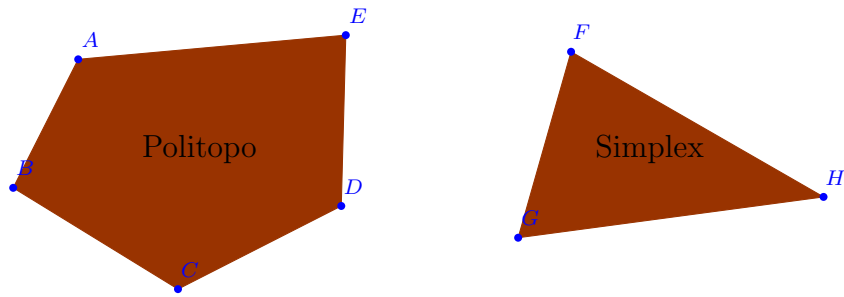


Figura 1.3: Politopo y simplex

□

**Proposición 1.3.** *La cápsula convexa de un conjunto  $C \subset X$ , no vacío, es el conjunto de todas las combinaciones convexas de elementos de  $C$ .*

También podemos definir la *cápsula afín* de  $C$  como la colección de todas las combinaciones afines de puntos en  $C$ . Ésta es la dimensión afín más pequeña de un sub espacio que contiene a  $C$ . Por ejemplo: la cápsula convexa afín de dos puntos distintos es una recta unidimensional conteniendo estos dos puntos. Similarmente, la *cápsula lineal* de  $C$  es la colección de todas las combinaciones lineales de puntos en  $C$  [2].

**Definición 1.5. :**

- *La cápsula convexa de un número finito de puntos  $x_1, x_2, \dots, x_{k+1} \in K$  es llamada politopo.*
- *Si  $x_1, x_2, \dots, x_k$  y  $x_{k+1}$  son afínmente independientes, es decir que  $x_2 - x_1, x_3 - x_1, \dots, x_{k+1} - x_1$  son linealmente independientes, entonces la  $co(x_1, x_2, \dots, x_{k+1})$  es llamada un simplex con vértices  $x_1, x_2, \dots, x_{k+1}$*

La Figura 1.3 nos muestra un ejemplo de un politopo y un simplex en  $\mathbb{R}^n$ . Note que el número máximo de vectores linealmente independientes en  $\mathbb{R}^n$  es  $n$  y por lo tanto no podría haber ningún simplex en  $\mathbb{R}^n$  teniendo más de  $n + 1$  vértices.

**Teorema de Caratheodory [2]**

Por definición, un punto en el conjunto de una cápsula convexa puede ser representado como una combinación convexa de un número finito de puntos en el conjunto. El siguiente teorema muestra que cualquier punto  $x$  en la cápsula convexa de un conjunto  $C$  puede ser representado como una combinación de a lo sumo  $n + 1$  puntos en  $C$ . El teorema es trivialmente cierto para  $x \in C$ .

**Teorema 1.1.** : Sea  $C$  un conjunto arbitrario en  $\mathbb{R}^n$ . Si  $x \in Co(C)$ ,  $x \in Co(x_1, \dots, x_{n+1})$  donde  $x_j \in C$ , para  $j = 1, \dots, n+1$ . En otras palabras  $x$  se puede representar como:

$$x = \sum_{j=1}^{n+1} \lambda_j x_j; \quad x_j \in C \text{ para } j = 1, \dots, n+1$$

$$\sum_{j=1}^{n+1} \lambda_j = 1; \quad \lambda_j \geq 0 \text{ para } j = 1, \dots, n+1$$

**Demostración [2]**

Como  $x \in co(S)$ ,  $x = \sum_{j=1}^k \lambda_j x_j$ , donde  $\lambda_j \geq 0$  para  $j = 1, \dots, k$ ,  $x \in S$  y  $\sum_{j=1}^k \lambda_j = 1$ . Si  $k \leq n+1$  el resultado es evidente.

Procederemos por contradicción.

Con este fin, note que  $x_2 - x_1, x_3 - x_1, \dots, x_k - x_1$  son linealmente independientes, entonces existen escalares  $\mu_2, \mu_3, \dots, \mu_k$  no todos cero tal que  $\sum_{j=2}^k \mu_j (x_j - x_1) = 0$ .

Haciendo  $\mu_1 = -\sum_{j=2}^k \mu_j$  se sigue que  $\sum_{j=1}^k \mu_j x_j = 0$ , por lo que  $\sum_{j=1}^k \mu_j = 0$  donde no todos los  $\mu_j$  son iguales a cero. Note que al menos uno de los  $\mu_j$  es mayor que cero. Entonces:

$$x = \sum_{j=1}^k \lambda_j x_j + 0 = \sum_{j=1}^k \lambda_j x_j - \alpha \sum_{j=1}^k \mu_j x_j = \sum_{j=1}^k (\lambda_j - \alpha \mu_j) x_j$$

para cualquier  $\alpha \in \mathbb{R}$ . Ahora, escogemos  $\alpha$  como sigue:

$$\alpha = \min_{1 \leq j \leq k} \left\{ \frac{\lambda_j}{\mu_j} \mid \mu_j > 0 \right\} = \frac{\lambda_i}{\mu_i}$$

para algún  $i \in \{1, \dots, k\}$

Note que  $\alpha > 0$ . Si  $\mu_j \leq 0$ ,  $\lambda_j - \alpha \mu_j > 0$ , y si  $\frac{\lambda_j}{\mu_j} \geq \frac{\lambda_i}{\mu_i} \alpha$  y por lo tanto  $\lambda_j - \alpha \mu_j \geq 0$ . En otras palabras,  $\lambda_j - \alpha \mu_j \geq 0 \quad \forall j = 1, \dots, k$ . En particular  $\lambda_i - \alpha \mu_i = 0$  por la definición de  $\alpha$ . Por lo tanto

$$x = \sum_{j=1}^k (\lambda_j - \alpha \mu_j) x_j, \text{ donde } \lambda_j - \alpha \mu_j \geq 0, \quad j = 1, \dots, k,$$

$$\sum_{j=1}^k (\lambda_j - \alpha \mu_j) = 1, \text{ y además } \lambda_i - \alpha \mu_i = 0.$$

Es decir que se ha presentado  $x$  como combinación convexa de a lo sumo  $k-1$  puntos de  $S$ . Este proceso se puede repetir hasta que  $x$  sea representado como una combinación convexa de máximo  $n+1$  puntos de  $S$ .

□

### 1.1.1 Propiedades topológicas de los conjuntos convexos

A continuación desarrollaremos algunas propiedades topológicas para conjuntos convexos [2]. Como un preliminar a esta parte tendremos que dado un punto  $x \in \mathbb{R}^n$  un  $\varepsilon$ -vecindario alrededor del conjunto es  $N_\varepsilon(x) = \{y : \|y - x\| < \varepsilon\}$ . Primero revisaremos las definiciones de: clausura, interior y frontera de un conjunto arbitrario en  $\mathbb{R}^n$ , usando el concepto de  $\varepsilon$ -vecindario.

**Definición 1.6.** : Sea  $S$  un conjunto arbitrario en  $\mathbb{R}^n$ .

- **Clausura:** Se dice que un punto  $x \in \mathbb{R}^n$  está en la clausura de  $S$  si  $\forall \varepsilon > 0, S \cap N_\varepsilon(x) \neq \emptyset$
- **Interior:** Un punto  $x$  se dice que está en el interior de  $S$  si  $\exists \varepsilon > 0$  tal que  $N_\varepsilon(x) \subset S$ .
- **Frontera:** Un punto  $x$  se dice que está en la frontera si  $\forall \varepsilon > 0$  se tienen que  $N_\varepsilon(x) \cap S \neq \emptyset$  y  $N_\varepsilon(x) \cap S^c \neq \emptyset$
- **Acumulación:** Un punto  $x$  es de acumulación de  $S$  si  $\forall \varepsilon > 0$  se tiene que  $(N_\varepsilon(x) - \{x\}) \cap S \neq \emptyset$

**Notación:**

- Denotaremos por  $\bar{S}$  a la clausura del conjunto  $S$ . Además se tiene que  $S$  es cerrado si  $\bar{S} = S$ .
- Denotaremos por  $\overset{\circ}{S}$  al interior del conjunto  $S$  y éste es abierto cuando  $\overset{\circ}{S} = S$
- Denotaremos por  $fr(S)$  a la frontera del conjunto  $S$ .
- Denotaremos por  $S'$  al conjunto de puntos de acumulación de  $S$

Finalmente, un conjunto es acotado si puede ser contenido en una bola de radio suficientemente grande. Un conjunto es *compacto* si es cerrado y acotado. Note que el complemento de un conjunto abierto es un conjunto cerrado (y viceversa) y que los puntos de acumulación de cualquier conjunto y su complemento son el mismo.

una manera un poco más fácil de asimilar éstas definiciones es de forma gráfica considerando  $S = \{(x, y) : x^2 + y^2 \leq 1\}$ , este conjunto representa todos los puntos dentro del círculo con centro  $C(0, 0)$  y radio  $r = 1$ . (Véase 1.4)

Fácilmente se puede verificar que  $S$  es cerrado, es decir,  $\bar{S} = S$ .

Además  $\overset{\circ}{S}$  consiste en todos los puntos que están estrictamente dentro del círculo, esto es,  $\overset{\circ}{S} = \{(x, y) : x^2 + y^2 < 1\}$ .

Finalmente,  $S'$  consiste de los puntos en el círculo, esto es,  $S' = \{(x, y) : x^2 + y^2 = 1\}$ .

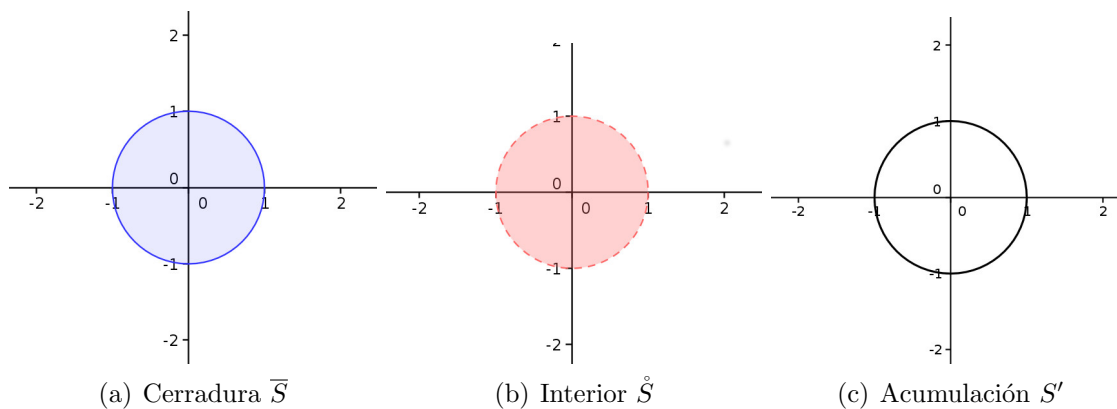


Figura 1.4: Topología del conjunto  $S = \{(x, y) : x^2 + y^2 \leq 1\}$

### Caracterizaciones de conjuntos

- Un conjunto  $S$  es cerrado si y solo si contiene todos sus puntos de acumulación, es decir  $S' \subset S$ .
- $\bar{S} = S \cup S'$  es el cerrado más pequeño que contiene al conjunto  $S$ .
- Un conjunto  $S$  es abierto si y solo si éste no contiene ninguno de sus puntos frontera, mas precisamente  $S' \cap S = \emptyset$ .
- $\mathring{S} \subseteq S$ , por la tanto, tenemos que  $\mathring{S} = S - S'$  mientras que necesariamente  $S' \neq S - \mathring{S}$ .

Un conjunto puede ser ni abierto ni cerrado, los únicos conjuntos que son abiertos y cerrados a la vez son el conjunto vacío ( $\emptyset$ ) y el mismo  $\mathbb{R}^n$ . También consideremos que cualquier punto  $x \in S$  puede ser un punto interior o frontera  $S$ . Sin embargo  $S \neq \mathring{S} \cup S'$ , es decir que  $S$  no necesita contener sus puntos frontera.

Existe otra definición equivalente de conjunto cerrado, desde el punto de vista que es un conjunto cerrado.

Esta definición está basada en las sucesiones de puntos contenidos en  $S$ .

Un conjunto es cerrado si y solo si para cualquier sucesión convergente de puntos  $\{x_k\}_{k \in \mathbb{N}} \in S$  con punto límite  $\bar{x}$ , tenemos que  $\bar{x} \in S$ .

La equivalencia de ésta y la definición previa de cerradura es fácilmente de ver ya que el punto límite  $\bar{x}$  de cualquier sucesión convergente de puntos en  $S$  debe encontrarse en el interior o en la frontera de  $S$ , de lo contrario debería existir un  $\varepsilon > 0$  tal que  $\{x : \|x - \bar{x}\| < \varepsilon\} \cap S = \emptyset$  contradiciendo que  $\bar{x}$  es el punto límite de una sucesión contenida en  $S$ .

**Segmento de línea entre puntos del interior y cerradura de un conjunto [2]**

Dado un conjunto convexo con interior no vacío, el segmento de línea (excluyendo los puntos finales) que une un punto del interior del conjunto con un punto de la clausura de este pertenece al interior del conjunto. Este resultado se muestra a continuación.

**Teorema 1.2.** *Sea  $S$  un conjunto convexo en  $\mathbb{R}^n$  con interior no vacío. Sea  $x_1 \in \bar{S}$  y  $x_2 \in \overset{\circ}{S}$ . Entonces  $\lambda x_1 + (1 - \lambda)x_2 \in \overset{\circ}{S} \quad \forall \lambda \in (0, 1)$*

***Demostración:***

Como  $x_2 \in \overset{\circ}{S}$  existe un  $\varepsilon > 0$  tal que  $\{z : \|z - x_2\| < \varepsilon\} \in S$ . Sea  $y$  tal que

$$y = \lambda x_1 + (1 - \lambda)x_2; \quad \lambda \in (0, 1) \quad (1.2)$$

Para probar que  $y \in \overset{\circ}{S}$  es suficiente construir un vecindario sobre  $y$  que también pertenece a  $S$ . En particular mostraremos que  $\{z : \|z - y\| < (1 - \lambda)\varepsilon\}$ . Sea  $z$  tal que  $\|z - y\| < (1 - \lambda)\varepsilon$  (véase 1.2), ahora bien, como  $x_1 \in \bar{S}$

$$\left\{ x : \|x - x_1\| < \frac{(1 - \lambda)\varepsilon - \|z - y\|}{\lambda} \right\} \cap S$$

es no vacío, en particular, existe  $z_1 \in S$  tal que

$$\|z_1 - x_1\| < \frac{(1 - \lambda)\varepsilon - \|z - y\|}{\lambda} \quad (1.3)$$

Ahora, sea  $z_2 = \frac{z - \lambda z_1}{1 - \lambda}$  de (1.2), la desigualdad de Schwartz y (1.3) tenemos:

$$\begin{aligned} \|z_2 - x_2\| &= \left\| \frac{z - \lambda z_1}{1 - \lambda} - x_2 \right\| = \left\| \frac{(z - \lambda z_1) - (y - \lambda x_1)}{1 - \lambda} \right\| \\ &= \frac{1}{1 - \lambda} \|(z - y) + (x_1 - \lambda x_1)\| \\ &\leq \frac{1}{1 - \lambda} (\|z - y\| + \lambda \|x_1 - z_1\|) \\ &< \varepsilon \end{aligned}$$

Por lo tanto,  $z_2 \in S$ . Dada la definición de  $z_2$  notemos que  $z = \lambda z_1 + (1 - \lambda)z_2$ , como  $z_1$  y  $z_2$  pertenecen a  $S$ , entonces  $z$  también pertenece a  $S$ . Hemos mostrado que para cualquier  $z$  con  $\|z - y\| < (1 - \lambda)\varepsilon$  pertenece a  $S$ . Por lo tanto  $y \in \overset{\circ}{S}$ .

□

**Corolario 1.1.** *Sea  $S$  un conjunto convexo. Entonces  $\overset{\circ}{S}$  es convexo.*

***Demostración***

Sean  $x, y \in \overset{\circ}{S}$  y  $t \in [0, 1]$ . Entonces existen  $\lambda, \mu > 0$  tales que  $B(x, \lambda) \subset S$  y  $B(y, \mu) \subset S$ . No es difícil ver que  $B((1-t)x + ty, (1-t)\lambda + t\mu) \subset S$ , usando la convexidad de  $S$ , lo que nos dice que  $(1-t)x + ty \in \overset{\circ}{S}$  y con esto hemos probado que  $\overset{\circ}{S}$  es convexo.

□

**Corolario 1.2.** *Sea  $S$  un conjunto convexo con interior no vacío. Entonces  $\overline{S}$  es convexa.*

***Demostración***

Asumiendo que  $\overset{\circ}{S} \neq \emptyset$ . Sea  $x_1, x_2 \in \overline{S}$ , por Teorema 1.2 tenemos que

$$\lambda x_2 + (1-\lambda)z \in \overset{\circ}{S} \quad \forall \lambda \in (0, 1), \quad z \in \overset{\circ}{S}$$

Sea  $\mu \in (0, 1)$  fijo. Por Teorema 1.2:

$$\mu x_1 + (1-\mu)[\lambda x_2 + (1-\lambda)z] \in \overset{\circ}{S} \subset S \quad \forall \lambda \in (0, 1)$$

Si tomamos el límite cuando  $\lambda$  se aproxima 1 se tiene que:  $\mu x_1 + (1-\mu)x_2 \in \overline{S}$ .

□

**Corolario 1.3.** *Sea  $S$  un conjunto convexo con interior no vacío. Entonces  $\overline{(\overset{\circ}{S})} = \overline{S}$ .*

***Demostración***

Claramente  $(\overset{\circ}{S}) \subseteq \overline{S}$ . Sea  $x \in \overline{S}$  escogemos  $y \in \overset{\circ}{S}$  (asumiendo que  $\overset{\circ}{S} \neq \emptyset$ ). Entonces:

$$\lambda x + (1-\lambda)y \in \overset{\circ}{S} \quad \forall \lambda \in (0, 1)$$

Haciendo  $\lambda \rightarrow 1^-$  se sigue que  $x \in \overline{(\overset{\circ}{S})}$ .

□

**Corolario 1.4.** *Sea  $S$  un conjunto convexo con interior no vacío. Entonces  $\overline{(\overset{\circ}{S})} = \overset{\circ}{S}$*

***Demostración:***

Note que  $\overset{\circ}{S} \subseteq \overline{(\overset{\circ}{S})}$ . Sea  $x_1 \in \overline{(\overset{\circ}{S})}$ , necesitamos mostrar que  $x_1 \in \overset{\circ}{S}$ . Existe un  $\varepsilon > 0$  tal que  $\|y - x_1\| < \varepsilon$  implica que  $y \in \overline{S}$ . Ahora, sea  $x_1 \neq x_2 \in \overset{\circ}{S}$  y sea  $y = (1+\delta)x_1 - \delta x_2$ , donde  $\delta = \frac{\varepsilon}{2\|x_1 - x_2\|}$ . Como  $\|y - x_1\| = \frac{\varepsilon}{2}$ ,  $y \in \overline{S}$ . Pero  $\lambda y + (1-\lambda)x_2$  donde

$$\lambda = \frac{1}{1+\delta} \in (0, 1)$$

Como  $y \in \overline{S}$  y  $x_2 \in \overset{\circ}{S}$  entonces por Teorema 1.2,  $x_1 \in \overset{\circ}{S}$ .

□

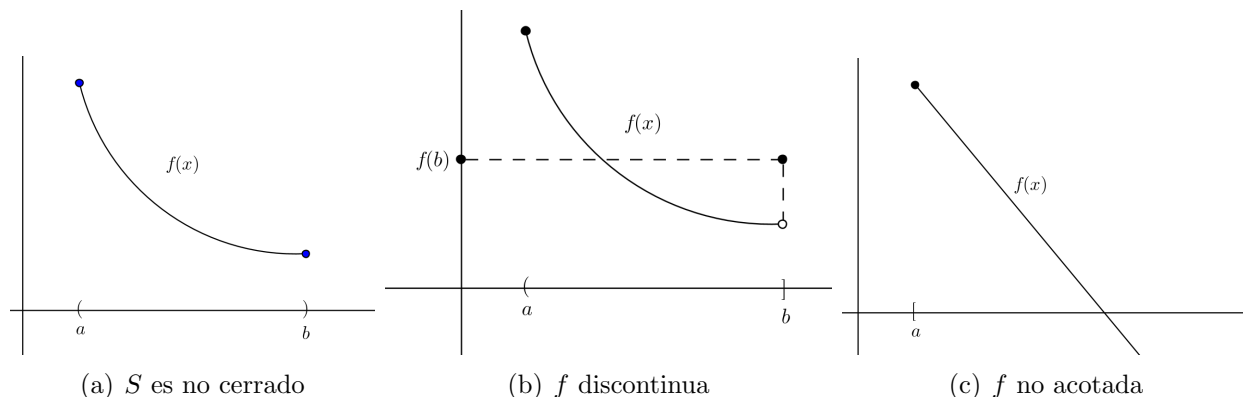


Figura 1.5: Inexistencia de solución de minimización [2]

### 1.1.2 Teorema de Weierstrass

Un resultado muy importante y ampliamente utilizado se basa en los conceptos anteriores. Este resultado relaciona la existencia de una solución de minimización para un problema de optimización [2].

- Podemos decir que  $\bar{x}$  es una solución de minimización al problema  $\min\{f(x)|x \in S\}$ , siempre que  $\bar{x} \in S$  y  $f(\bar{x}) \leq f(x) \forall x \in S$ . En tal caso, decimos que existe un mínimo.
- Por otra parte, decimos que  $\alpha = \inf\{f(x)|x \in S\}$  si  $\alpha$  es la mayor de las cotas inferiores de  $f$  en  $S$ ; esto es:  $\alpha \leq f(x) \forall x \in S$ , es decir que no hay  $\bar{\alpha} > \alpha$  tal que  $\bar{\alpha} \leq f(x) \forall x \in S$ .
- Similarmente,  $\alpha = \max\{f(x)|x \in S\}$  si existe una solución  $\bar{x} \in S$  tal que  $\alpha = f(\bar{x}) \geq f(x) \forall x \in S$ .
- Por otro lado,  $\alpha = \sup\{f(x)|x \in S\}$  si  $\alpha$  es la menor de las cotas superiores de  $f$  en  $S$ ; esto es  $\alpha \geq f(x) \forall x \in S$ , y no hay otro  $\bar{\alpha} < \alpha$  tal que  $\bar{\alpha} \geq f(x) \forall x \in S$ .

En la Figura (1.5) se ilustran tres momentos en los que el mínimo no existe.

En la Figura (1.5a) el ínfimo de  $f$  sobre  $(a, b)$  está dado por  $f(b)$  pero como  $S$  no es cerrado y en particular  $b \notin S$  el ínfimo no existe.

En la Figura (1.5b) tenemos que  $\inf\{f(x)|x \in [a, b]\}$  está dado por el límite de  $f(x)$  cuando  $x$  tiende a  $b$  por la izquierda.  $(\lim_{x \rightarrow b^-} f(x))$ . Como  $f$  es discontinua en  $b$  no existe una solución de minimización.

La figura (1.5c) ilustra cuando  $f$  no es acotada sobre un conjunto  $S = \{x : x \geq a\}$

Ahora probaremos formalmente este resultado cuando  $S$  es no vacío, cerrado y acotado,  $f$  continua en  $S$ , con éstas condiciones el mínimo, a diferencia de las situaciones en la Figura 1.5, si existe.

**Teorema 1.3.** *Sea  $S$  un conjunto no vacío, compacto y sea  $f \mapsto \mathbb{R}$  continua en  $S$  entonces el problema  $\min\{f(x) : x \in S\}$  alcanza su mínimo; esto es, existe una solución de mini-*



mización para este problema.

***Demostración:***

Como  $f$  es continua en  $S$  y  $S$  es cerrado y acotado,  $f$  también está acotada (recordar que  $f$  es continua), como  $S \neq \emptyset$ , existe una cota inferior (la más grande)  $\alpha \equiv \inf\{f(x) : x \in S\}$ . Ahora, sea  $0 < \varepsilon < 1$  y consideremos el conjunto  $S_k = \{x \in S : \alpha \leq f(x) \leq \alpha + \varepsilon^k\} \forall k = 1, 2, \dots$ . Por la definición de ínfimo se tiene que  $S_k \neq \emptyset \forall k$ , entonces, podemos construir una sucesión de puntos  $\{x_k\} \subseteq S$  seleccionando un punto  $x_k \in S_k \forall k = 1, 2, \dots$ . Como  $S$  es acotado, existe una subsucesión convergente  $\{x_k\}_K \rightarrow \bar{x}$  indexada por el conjunto  $K$ . Por la cerradura de  $S$  tenemos  $\bar{x} \in S$ ; y por la continuidad  $f$  se tiene que como  $\alpha \leq f(x) \leq \alpha + \varepsilon^k \forall k$  tenemos que:

$$\alpha = \lim_{k \rightarrow \infty; k \in K} f(x_k) = f(\bar{x})$$

Por lo tanto, hemos demostrado que existe una solución  $\bar{x} \in S$  tal que  $f(\bar{x}) = \alpha = \inf\{f(x) | x \in S\}$ , entonces  $\bar{x}$  es una solución de minimización.

□

### Separación y soporte de conjuntos

Las nociones de hiperplano soporte y separación de conjuntos convexos son muy importantes en optimización. Casi todas las condiciones de optimalidad y dualidad se relacionan con algún tipo de separación o soporte de conjuntos convexos [2]. Los resultados que veremos a continuación están basados en el siguiente hecho geométrico: Dado un conjunto cerrado  $S$  y un punto  $y \in S$  existe un único punto  $\bar{x} \notin S$  con una distancia mínima desde  $y$  y un hiperplano que separa a  $y$  y  $S$ .

### Distancia mínima de un punto a un conjunto convexo [2]

Para establecer el importante resultado anterior seguiremos la *ley del paralelogramo*.

Sean  $a$  y  $b$  dos vectores en  $\mathbb{R}^n$ . Entonces:

$$\begin{aligned} \|a + b\|^2 &= \|a\|^2 + \|b\|^2 + 2a^t b \\ \|a - b\|^2 &= \|a\|^2 + \|b\|^2 - 2a^t b \end{aligned}$$

Sumando ambas se tiene:

$$\|a + b\|^2 + \|a - b\|^2 = 2\|a\|^2 + 2\|b\|^2$$

Este resultado se ilustra en la Figura (1.6) y puede ser representado de la siguiente manera: La suma de las normas al cuadrado de las diagonales de un paralelogramo es igual al doble de la suma de sus normas al cuadrado.

De la teoría básica de geometría en espacios vectoriales normados se tiene que un hiperplano cerrado  $H$  en  $X$  puede ser representado por: [5]

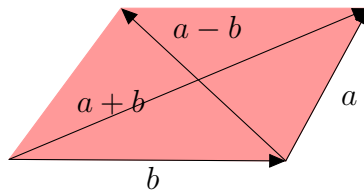


Figura 1.6: Ley del paralelogramo

**Definición 1.7.** Sean  $a \in \mathbb{R}^n$ ,  $\alpha \in \mathbb{R}$  fijos. Se llama hiperplano al conjunto

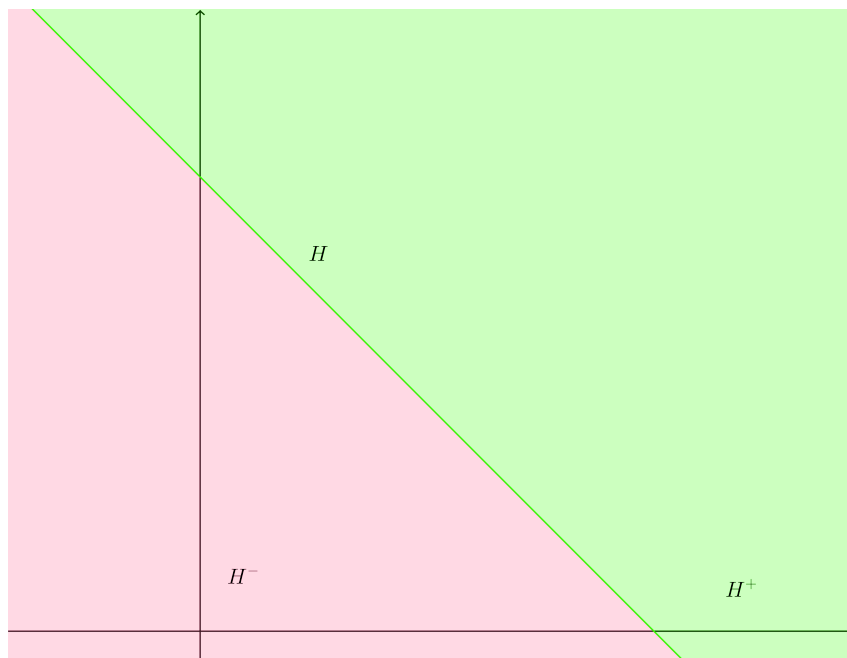
$$\mathcal{H} = \{x \in \mathbb{R}^n / a^T x = \alpha\}$$

Un hiperplano  $\mathcal{H}$  define dos semiespacios:

$$\mathcal{H}^- = \{x \in \mathbb{R}^n / a^T x \leq \alpha\}$$

$$\mathcal{H}^+ = \{x \in \mathbb{R}^n / a^T x \geq \alpha\}$$

Cabe destacar que la unión de estos semiespacios da como resultado todo el espacio en el cual se está trabajando y la intersección de éstos nos da como resultado el hiperplano bajo el cual están definidos (Véase 1.7)

Figura 1.7: Hiperplanos definidos por  $\mathcal{H}$

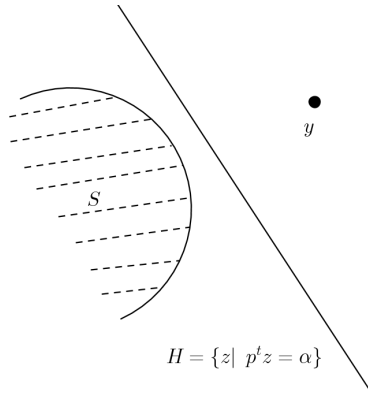


Figura 1.8:  $H$  es el hiperplano separador entre  $S$  e  $y$ .

**Teorema 1.4. Hahn-Banach** Sea  $S$  un convexo cerrado no vacío,  $y \notin S$ . Entonces existen  $p \in \mathbb{R}^n$ ,  $p \neq 0$ ,  $\alpha \in \mathbb{R}$ , tales que

$$\begin{aligned} p^T y &> \alpha \\ p^T x &\leq \alpha, \quad \forall x \in S \end{aligned}$$

**Demostración**

Sea  $\bar{x} \in S$  tal que  $\langle y - \bar{x}, x - \bar{x} \rangle \leq 0$ ,  $\forall x \in S$ .

Sean  $p = y - \bar{x} \neq 0$  y  $\alpha = \langle p, \bar{x} \rangle$ . tenemos que

$$\langle p, x - \bar{x} \rangle \leq 0, \quad \forall x \in S$$

Por lo tanto

$$\langle p, x \rangle \leq \langle p, \bar{x} \rangle = \alpha \quad \forall x \in S \quad (1.4)$$

Por otro lado

$$\langle p, x - \bar{x} \rangle = \langle p, x - y + y - \bar{x} \rangle = \langle p, x - y \rangle + \langle p, y - \bar{x} \rangle = \langle p, x - y \rangle + \|p\|^2 \leq 0, \quad \forall x \in S$$

lo que implica que

$$\langle p, x \rangle + \|p\|^2 \leq \langle p, y \rangle, \quad \forall x \in S$$

Como  $\|p\|^2 \neq 0$ , se tiene que

$$\langle p, x \rangle < \langle p, y \rangle, \quad \forall x \in S \quad (1.5)$$

Por (1.4), se tiene que  $\langle p, x \rangle \leq \alpha \quad \forall x \in S$  y por (1.5), se tiene que  $\alpha < \langle p, y \rangle$ , lo que concluye la demostración.

□

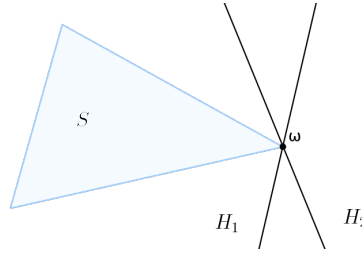


Figura 1.9:  $H_1$  y  $H_2$  son dos hiperplanos soporte en el punto  $\omega$

El teorema anterior introduce la noción de *hiperplano separador* (Ver Figura 1.8)

**Definición 1.8.** Sea  $S$  un convexo cerrado no vacío. Un hiperplano soporte de  $S$  es un hiperplano  $\mathcal{H}$  tal que  $\mathcal{H} \cap S \neq \emptyset$  y  $\{S \subseteq \mathcal{H}^+ \vee S \subseteq \mathcal{H}^-\}$

Para asimilar ésta definición presentamos la Figura (1.9)

El siguiente teorema nos permitirá deducir una caracterización para un conjunto convexo no vacío cualquiera (Ver [11]).

**Teorema 1.5.** Sea  $S \subseteq \mathbb{R}^n$  un conjunto convexo y sea  $\bar{x}$  un punto en la frontera de  $S$ . Entonces  $S$  tiene un hiperplano soporte en  $\bar{x}$ .

### *Separación de dos conjuntos convexos*

Hasta ahora hemos discutido la separación de un conjunto convexo con un punto que no pertenece a este conjunto y hemos discutido el soporte de conjuntos convexos en puntos frontera. Además, si tenemos dos conjuntos convexos disjuntos, ellos pueden ser separados por un hiperplano  $H$  tal que uno de estos conjuntos pertenece a  $H^+$  y el otro pertenece al conjunto  $H^-$ . En efecto, este resultado resulta cierto, incluso si estos dos conjuntos tienen algunos puntos en común siempre y cuando el interior de éstos sea disjunto. Este resultado se hace preciso por el siguiente teorema.

**Teorema 1.6.** Sea  $S_1, S_2$  dos conjuntos convexos no vacíos de  $\mathbb{R}^n$  y suponga que  $S_1 \cap S_2 = \emptyset$ . Entonces, existe un hiperplano que separa  $S_1$  y  $S_2$ , es decir, que existe un vector no nulo  $p \in \mathbb{R}^n$  tal que

$$\inf\{p^t x \mid x \in S_1\} \geq \sup\{p^t x \mid x \in S_2\}$$

### *Demostración ([2] y [11])*

Sea  $S = S_1 \ominus S_2 = \{x_1 - x_2 \mid x_1 \in S_1 \text{ y } x_2 \in S_2\}$ . Note que  $S$  es un conjunto convexo. Además  $0 \in S$  porque de lo contrario  $S_1 \cap S_2$  sería no vacío por Teorema 1.5 y Teorema 1.4 existe un vector no nulo  $p \in \mathbb{R}^n$  tal que

$$p^t x \geq 0 \quad \forall x \in S$$

de donde se concluye

$$p^t(x_2 - x_1) \leq 0 \quad \forall x_1 \in S_1, x_2 \in S_2$$

□

## 1.2 Funciones convexas

Funciones convexas y cóncavas tienen muchas propiedades importantes y especiales. Por ejemplo: cualquier mínimo local de una función convexa sobre un conjunto convexo es también un mínimo global. A continuación introduciremos temas importantes de funciones convexas y desarrollaremos algunas de sus propiedades [2].

**Definición 1.9.** Sea  $f : S \rightarrow \mathbb{R}$ , donde  $S$  es un conjunto convexo no vacío de  $\mathbb{R}^n$ .

- La función  $f$  se dice que es **convexa** en  $S$  si:

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

para cada  $x_1, x_2 \in S$  y para cada  $\lambda \in (0, 1)$ .

- La función  $f$  es **estrictamente convexa** en  $S$  si la desigualdad es estricta:

$$f(\lambda x_1 + (1 - \lambda)x_2) < \lambda f(x_1) + (1 - \lambda)f(x_2)$$

para cada  $x_1, x_2 \in S$  y para cada  $\lambda \in (0, 1)$ .

- La función  $f$  es llamada **cóncava (estrictamente cóncava)** en  $S$  si  $-f$  es convexa (estrictamente convexa) en  $S$ .

Sean  $x_1$  y  $x_2$  dos puntos en el dominio de  $f$ . Entonces una función  $f$  es convexa, si el segmento de recta que une dos puntos  $(x_1, f(x_1))$  y  $(x_2, f(x_2))$  del gráfico de  $f$  está por encima de la gráfica de la función de  $f$  (Ver [9]), como lo muestra la Figura(1.10).

Los siguientes son algunos ejemplos de funciones convexas [2].

1.  $f(x) = 3x + 4$
2.  $f(x) = |x|$
3.  $f(x) = x^2 - 2x$
4.  $f(x) = (-x)^{\frac{1}{2}}$

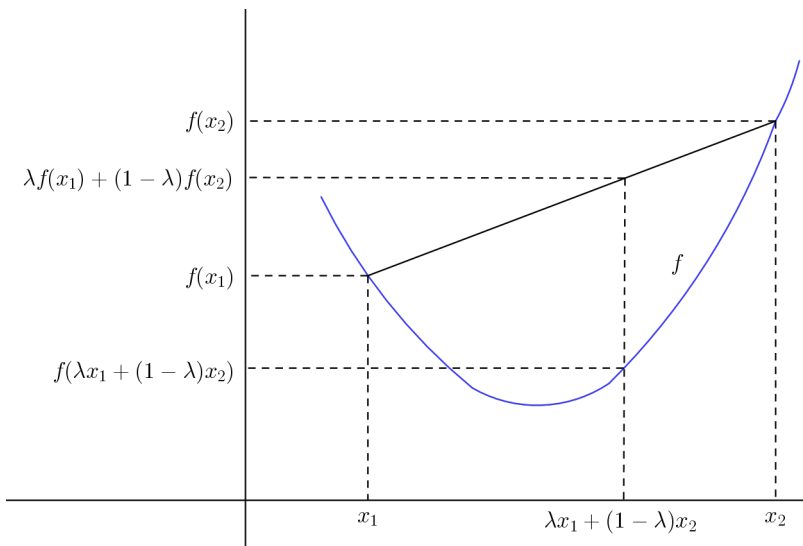


Figura 1.10: Interpretación geométrica de función convexa en  $\mathbb{R}$  [9]

5.  $f(x_1, x_2) = 2x_1^2 + 2x_2^2 - 2x_1x_2$

6.  $f(x_1, x_2, x_3) = x_1^4 + x_2^2 + 3x_3^2 - 4x_1 - 4x_2x_3$

Note que en cada uno de los ejemplos las funciones son convexas sobre  $\mathbb{R}^n$ . Excepto para el ejemplo 4, la función no está definida para  $x < 0$ . Se pueden construir ejemplos de funciones que son convexas en una región pero no sobre todo  $\mathbb{R}^n$ . Por ejemplo  $f(x) = x^3$  no es convexa sobre  $\mathbb{R}$ , pero es convexa sobre  $S = \{x : x \geq 0\}$ .

De ahora en adelante nos concentraremos en funciones convexas, pues el resultado para funciones cóncavas puede ser obtenido fácilmente notando que  $f$  es cóncava si y sólo si  $-f$  es convexa.

Un conjunto asociado a una función convexa  $f$  es  $S_\alpha = \{x \in S : f(x) \leq \alpha\}$ ,  $\alpha \in \mathbb{R}$ , usualmente referido como *conjunto de nivel* (Ver [2]). A veces este conjunto es llamado un *conjunto subnivel*, para diferenciarlo del *conjunto supernivel*  $\{x \in S : f(x) \geq \alpha\}$ . Note que éstos tienen propiedades a las de una función cóncava.

**Teorema 1.7.** *Sea  $S$  un conjunto convexo no vacío en  $\mathbb{R}^n$  y sea  $f : S \rightarrow \mathbb{R}$  una función convexa. Entonces el conjunto de nivel  $S_\alpha = \{x \in S : f(x) \leq \alpha\}$ ;  $\alpha \in \mathbb{R}$ , es un conjunto convexo.*

***Demostración:***

Sean  $x_1, x_2 \in S_\alpha$ . Entonces  $x_1, x_2 \in S$  y  $f(x_1) \leq \alpha$  y  $f(x_2) < \alpha$ . Ahora, sea  $\lambda \in (0, 1)$  y  $x = \lambda x_1 + (1 - \lambda)x_2$ . Por la convexidad de  $S$  se tiene que  $x \in S$ , por otro lado, dada la convexidad de  $f$  se tiene:

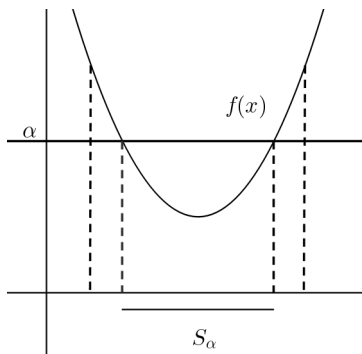


Figura 1.11: Conjunto de nivel de una variable

$$f(x) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \leq \lambda\alpha + (1 - \lambda)\alpha = \alpha$$

Por lo tanto,  $x \in S_\alpha$ , y por consiguiente,  $S_\alpha$  es convexo.

□

### Continuidad de funciones convexas [2]

Una propiedad importante de funciones convexas y cóncavas es que son continuas en el interior de su dominio.

**Teorema 1.8.** *Sea  $S$  un conjunto convexo no vacío en  $\mathbb{R}^n$  y sea  $f : S \rightarrow \mathbb{R}$  una función convexa. Entonces  $f$  es continua en el interior de  $S$ .*

#### **Demostración:**

Sea  $\bar{x} \in \overset{\circ}{S}$ . Para probar la continuidad de  $f$  en  $\bar{x}$ , es necesario mostrar que dado  $\varepsilon > 0$  existe un  $\delta > 0$  tal que  $\|x - \bar{x}\| \leq \delta \Rightarrow |f(x) - f(\bar{x})| \leq \varepsilon$ .

Sea entonces  $\varepsilon > 0$ . Dado que  $\bar{x} \in \overset{\circ}{S} \exists \eta > 0$  tal que  $B(\bar{x}, \eta) \subseteq S$

Claramente  $\bar{x} \pm \eta e_i \in S$  con  $e_i$  vector canónico. Luego, es claro que

$$\begin{aligned} \bar{x} &= \frac{1}{2}(\bar{x} + \eta e_i) + \frac{1}{2}(\bar{x} - \eta e_i) \\ \Rightarrow f(\bar{x}) &= f\left(\frac{1}{2}(\bar{x} + \eta e_i) + \frac{1}{2}(\bar{x} - \eta e_i)\right) \\ &\leq \frac{1}{2}f(\bar{x} + \eta e_i) + \frac{1}{2}f(\bar{x} - \eta e_i); \quad \forall i = 1, \dots, n \end{aligned}$$

de donde

$$\begin{aligned}
0 &\leq \frac{1}{2}f(\bar{x} + \eta e_i) + \frac{1}{2}f(\bar{x} - \eta e_i) - f(\bar{x}) \\
&= \frac{1}{2}f(\bar{x} + \eta e_i) + \frac{1}{2}f(\bar{x} - \eta e_i) - \frac{1}{2}f(\bar{x}) - \frac{1}{2}f(\bar{x}) \\
&= \frac{1}{2}[f(\bar{x} + \eta e_i) - f(\bar{x})] + \frac{1}{2}[f(\bar{x} - \eta e_i) - f(\bar{x})] \text{ es decir} \\
0 &\leq \frac{1}{2}[f(\bar{x} + \eta e_i) - f(\bar{x})] + \frac{1}{2}[f(\bar{x} - \eta e_i) - f(\bar{x})]
\end{aligned}$$

De aquí se desprende que  $\forall i$   $f(\bar{x} + \eta e_i) - f(\bar{x})$  y  $f(\bar{x} - \eta e_i) - f(\bar{x})$  no pueden ser simultáneamente negativos.

Sea  $k = \max\{f(\bar{x} \pm \eta e_i) - f(\bar{x}) \mid \forall i = 1, \dots, n\} \Rightarrow 0 \leq k < \infty$  y definimos  $\delta = \min\{\frac{\eta}{nk}, \frac{\varepsilon\eta}{nk}\}$

Sean  $\alpha_i \geq 0 \quad \forall i = 1, \dots, n$  tales que:  $x - \bar{x} = \sum_{i=1}^n \alpha_i d_i$

$$d_i = \begin{cases} \eta e_i & \text{si } x - \bar{x} \geq 0 \\ -\eta e_i & \text{si } x - \bar{x} < 0 \end{cases}$$

Luego:

$$\begin{aligned}
\|x - \bar{x}\|^2 &= \left\| \sum_{i=1}^n \alpha_i d_i \right\|^2 \\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j d_i d_j \\
&= \sum_{i=1}^n \alpha_i^2 \|d_i\|^2 \\
&= \eta \sum_{i=1}^n \alpha_i^2 \\
&\leq \delta^2
\end{aligned}$$

Así se tiene que  $\sum_{i=1}^n \alpha_i^2 \leq \frac{\delta^2}{\eta^2} = \min\{\frac{1}{n^2}, \frac{\varepsilon^2}{n^2 k^2}\}$  lo que implica en particular que

$$\alpha_i \leq \min\{\frac{1}{n}, \frac{\varepsilon}{nk}\} \quad \forall i = 1, \dots, n$$



$$\begin{aligned}
\Rightarrow f(x) &= f(x - \bar{x} + \bar{x}) \\
&= f\left(\sum_{i=1}^n \alpha_i d_i + \bar{x}\right) \\
&= f\left(\sum_{i=1}^n \frac{1}{n}(n\alpha_i d_i + \bar{x})\right) \\
&\leq \sum_{i=1}^n \frac{1}{n} f(n\alpha_i d_i + \bar{x}) \\
&= \sum_{i=1}^n \frac{1}{n} f(n\alpha_i d_i + \bar{x} + n\alpha_i \bar{x} - n\alpha_i \bar{x}) \\
&= \frac{1}{n} \sum_{i=1}^n f((1 - n\alpha_i)\bar{x} + n\alpha_i(\bar{x} + d_i)) \\
&\leq \frac{1}{n} \sum_{i=1}^n [(1 - n\alpha_i)f(\bar{x}) + n\alpha_i f(\bar{x} + d_i)] \\
&= f(x) + \sum_{i=1}^n \alpha_i [f(\bar{x} + d_i) - f(\bar{x})]
\end{aligned}$$

Luego (de la definición de  $k$  y por la cota establecida para  $\alpha_i$ )

$$\begin{aligned}
f(x) - f(\bar{x}) &\leq \sum_{i=1}^n \alpha_i [f(\bar{x} + d_i) - f(\bar{x})] \\
&\leq k \sum_{i=1}^n \alpha_i \\
&\leq \varepsilon
\end{aligned}$$

Para terminar falta probar que  $f(\bar{x}) + f(x) \leq \varepsilon$

Sea  $y = 2\bar{x} - x$ . Notemos que  $\|y - \bar{x}\| = \|\bar{x} - x\| \leq \delta$ , luego  $f(y) - f(\bar{x}) \leq \varepsilon$ .

Pero  $f(\bar{x}) = f\left(\frac{1}{2}y + \frac{1}{2}x\right) \leq \frac{1}{2}f(y) + \frac{1}{2}f(x)$  implica que

$$\frac{1}{2}[f(\bar{x}) - f(x)] \leq \frac{1}{2}[f(y) - f(\bar{x})] \leq \frac{1}{2}\varepsilon$$

Luego  $|f(x) - f(\bar{x})| \leq \varepsilon$  y se tiene el resultado. □

Note que las funciones convexas y cóncavas pueden no ser continuas en todo lugar. Sin embargo, por el Teorema (1.8), los puntos de discontinuidad solo son permitidos en la frontera de  $S$ , como se ilustra en la siguiente función convexa definida en  $S = \{x : -1 \leq x \leq 1\}$ :

$$f(x) = \begin{cases} x^2 & \text{para } |x| < 1 \\ 2 & \text{para } |x| = 1 \end{cases}$$

### 1.3 Funciones convexas no diferenciables

Las causas de no diferenciables de un problema de optimización son diversas, sin embargo, con un fin explicativo se pueden agrupar en dos: [3]

- La no diferenciables de las funciones del problema, aunque las variables sean continuas.
- La no continuidad de las variables del mismo, aunque las funciones pudieran ser diferenciables en un conjunto más amplio de puntos que el conjunto solución del problema. Se entiende por no continuidad de las variables el hecho de que alguna de ellas sólo pueda tomar valores enteros, o que sea continua por tramos o que sea semicontinua.

#### 1.3.1 Subgradiente

Las funciones convexas poseen algunas propiedades usuales de diferenciables, y una de ellas es la existencia de la derivada direccional lateral (derecha e izquierda) en toda dirección en un punto interior de su dominio. Así como en el caso usual de la derivada direccional de una función diferenciable puede ser escrita en términos del vector gradiente, que está asociado con un hiperplano tangente al gráfico de la función, la derivada lateral derecha de una función convexa, no necesariamente diferenciable puede ser escrita en términos de vectores subgradientes que están asociados con hiperplanos soportes al epígrafo de la función. En lo sucesivo nos referiremos a la derivada direccional derecha como derivada direccional [14].

**Definición 1.10.** Sean  $f : \mathbb{R}^n \mapsto \mathbb{R}$ ,  $x \in \mathbb{R}^n$ ; la derivada direccional de  $f$  en  $x$  con respecto a  $y$  es definida como el siguiente límite:

$$f'(x) = \lim_{t \rightarrow 0^+} \frac{f(x + ty) - f(x)}{t}$$

si este límite existe.

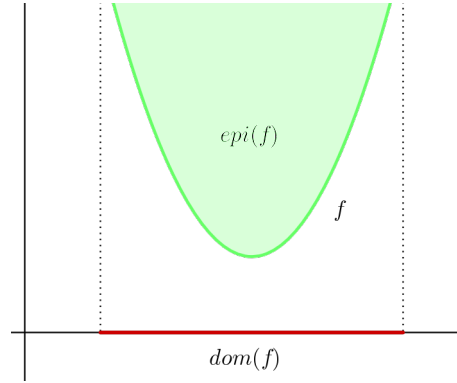


Figura 1.12: Epígrafo de una función  $f$  en  $\mathbb{R}$  [12]

### Hipógrafo y epígrafo de una función

Una función  $f$  en  $S$  puede ser completamente descrita por el conjunto  $\{(x, f(x)) \mid x \in S\} \subset \mathbb{R}^{n+1}$ , denominado **grafo** de la función. Se puede construir uno de dos conjuntos que están relacionados al grafo de la función  $f$ : el **epígrafo** y el **hipógrafo** de  $f$ .

**Definición 1.11.** Sea  $S$  un conjunto no vacío en  $\mathbb{R}^n$ , y sea  $f : S \mapsto \mathbb{R}$ . El epígrafo de  $f$  denotado por  $\text{epi}(f)$ , es un subconjunto de  $\mathbb{R}^{n+1}$  definido por:

$$\{(x, y) \mid x \in S, y \in \mathbb{R}, y \geq f(x)\}$$

El hipógrafo de  $f$  denotado por  $\text{hyp}(f)$  es un subconjunto de  $\mathbb{R}^{n+1}$  definido por:

$$\{(x, y) \mid x \in S, y \in \mathbb{R}, y \leq f(x)\}$$

**Teorema 1.9.** Sea  $S$  un conjunto convexo no vacío en  $\mathbb{R}^n$ , y sea  $f : S \mapsto \mathbb{R}$ . Entonces  $f$  es convexa si y solo si el  $\text{epi}(f)$  es un conjunto convexo.

#### **Demostración** [11]

” $\Rightarrow$ ” Sea  $f$  convexa. Sean  $(x, \alpha), (y, \beta) \in \text{epi}(f)$ ,  $\lambda \in [0, 1]$ . Entonces por convexidad

$$\lambda(x, \alpha) + (1 - \lambda)(y, \beta) = (\lambda x + (1 - \lambda)y, \lambda \alpha + (1 - \lambda)\beta) \in S \times \mathbb{R}$$

Como  $f$  es convexa, entonces

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \leq \lambda \alpha + (1 - \lambda)\beta$$

Puesto que  $f(x) \leq \alpha$  y  $f(y) \leq \beta$ . Luego

$$\lambda(x, \alpha) + (1 - \lambda)(y, \beta) \in \text{epi}(f)$$

Por lo tanto  $\text{epi}(f)$  es convexo.

” $\Leftarrow$ ”

Sea ahora  $\text{epi}(f)$  un conjunto convexo. Claramente se tiene que  $(x, f(x)), (y, f(y)) \in \text{epi}(f)$  y por lo tanto, para  $\lambda \in [0, 1]$  se tiene:

$$\lambda(x, f(x)) + (1 - \lambda)(y, f(y)) \in \text{epi}(f)$$

es decir

$$(\lambda x + (1 - \lambda)y, \lambda f(x) + (1 - \lambda)f(y)) \in \text{epi}(f)$$

de donde se concluye que  $f$  es convexa. □

**Definición 1.12.** Sean  $f : \mathbb{R}^n \mapsto \mathbb{R}$ ,  $x_0 \in \mathbb{R}^n$  Un vector  $s \in \mathbb{R}^n$  tal que:

$$f(x) \geq f(x_0) + \langle s, x - x_0 \rangle, \quad \forall x \in \mathbb{R}^n$$

es llamado un subgradiente de  $f$  en  $x_0$ .

El conjunto de subgradientes de una función  $f$  en el punto  $x_0$  es llamado subdiferencial de  $f$  en  $x_0$  y se denota por  $\partial f(x_0)$ . [14]

**Ejemplo 1.1.** Sea  $f(x) = \min\{f_1(x), f_2(x)\}$ , donde  $f_1$  y  $f_2$  son definidas como sigue:

$$\begin{aligned} f_1(x) &= 4 - |x|, & x \in \mathbb{R} \\ f_2(x) &= 4 - (x - 2)^2, & x \in \mathbb{R} \end{aligned}$$

Como  $f_1(x) \leq f_2(x)$  para  $1 \leq x \leq 4$ ,  $f$  puede ser representada de la siguiente forma:

$$f(x) = \begin{cases} 4 - |x| & \text{si } 1 \leq x \leq 4 \\ 4 - (x - 2)^2 & \text{en otro caso} \end{cases}$$

En la Figura 1.13 se muestra la función cóncava  $f$ . Note que  $\xi = -1$  es la pendiente y por lo tanto el subgradiente de  $f$  en cualquier punto  $x$  en el intervalo abierto  $(1, 4)$ . Si  $x < 1$  o  $x > 4$ ,  $\xi = -2(x - 2)$  es el único subgradiente de  $f$ . En los puntos  $x = 1$  y  $x = 4$ , los subgradientes no son únicos porque muchos hiperplanos soportes existen. En  $x = 1$ , la familia de subgradientes está caracterizada por  $\lambda \nabla f_1(1) + (1 - \lambda) \nabla f_2(1) = \lambda(-1) + (1 - \lambda)(2) = 2 - 3\lambda$  para  $\lambda \in [0, 1]$ . En otras palabras, cualquier  $\xi$  en el intervalo  $[-1, 2]$  es un subgradiente de  $f$  en  $x = 1$ , y éstos corresponden a las pendientes de la familia de hiperplanos soporte de  $f$  en  $x = 1$ . En  $x = 4$ , la familia de subgradientes está caracterizada por  $\lambda \nabla f_1(4) + (1 - \lambda) \nabla f_2(4) = \lambda(-1) + (1 - \lambda)(-4) = -4 + 3\lambda$  para  $\lambda \in [0, 1]$ . En otras palabras, cualquier  $\xi$  en el intervalo  $[-4, -1]$  es un subgradiente de  $f$  en  $x = 4$ .

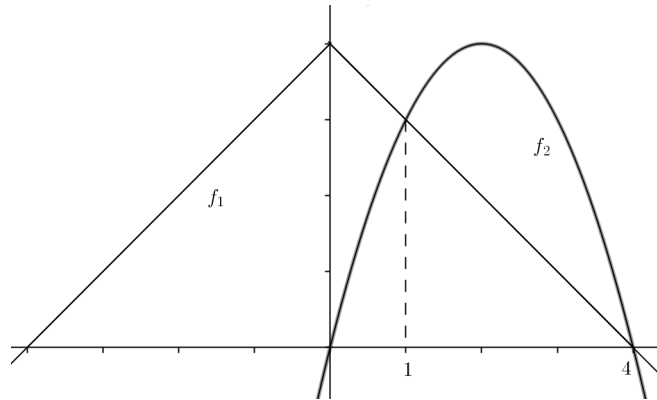
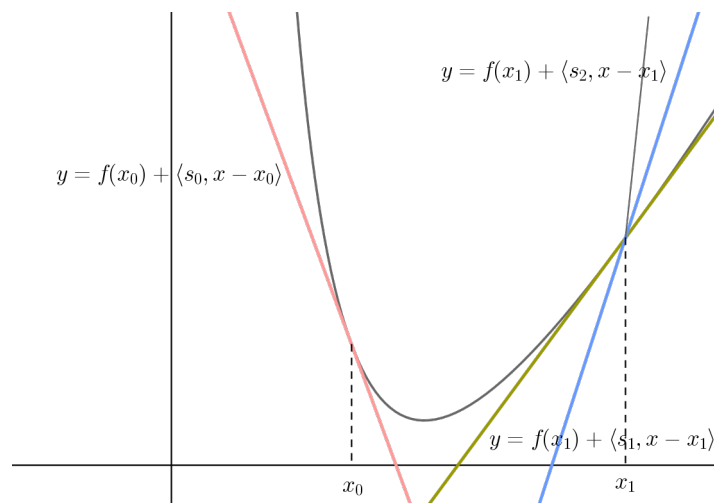


Figura 1.13: figura del ejemplo de subgradiente [2]

Figura 1.14: Interpretación geométrica del subgradiente en  $\mathbb{R}$  [14]

**Definición 1.13.** Si  $\partial f(x_0) \neq \emptyset$  entonces diremos que  $f$  es subdiferenciable en  $x_0$ .

Geoméricamente,  $s \in \mathbb{R}^n$  es un subgradiente de la función convexa  $f$  en  $x_0$  si el grafo de  $f$  en  $\mathbb{R}^{n+1}$  está por encima del gráfico del hiperplano  $y = f(x) + \langle s, x - x_0 \rangle$ . Como  $(x_0, f(x_0))$  está en el hiperplano, este hiperplano constituye un hiperplano soporte a  $\text{epi}(f)$  en  $(x_0, f(x_0))$ . Esto es, la existencia de un vector subgradiente establece la existencia de un hiperplano soporte no vertical a  $\text{epi}(f)$  en  $(x_0, f(x_0))$ .

El siguiente teorema muestra que cada función convexa o cóncava tiene al menos un subgradiente en el interior de su dominio. La prueba se basa en el hecho que un conjunto convexo tiene un hiperplano soporte en sus puntos límite.

**Teorema 1.10.** Sea  $S$  un conjunto no vacío en  $\mathbb{R}^n$ , y sea  $f : S \rightarrow \mathbb{R}$  convexa. Entonces, para  $\bar{x} \in \overset{\circ}{S}$ , existe un vector  $\xi$  tal que el hiperplano

$$H = \{(x, y) : y = f(\bar{x}) + \xi^t(x - \bar{x})\}$$

es soporte para  $\text{epi}(f)$  en  $[\bar{x}, f(\bar{x})]$ . En particular,

$$f(x) \geq f(\bar{x}) + \xi^t(x - \bar{x}) \quad \forall x \in S$$

estos es:  $\xi$  es un subgradiente de  $f$  en  $\bar{x}$

### **Demostración**

Por Teorema (1.9),  $\text{epi}(f)$  es convexo. Note que  $(\bar{x}, f(\bar{x}))$  pertenece a la frontera del  $\text{epi}(f)$ , por el Teorema (1.5) existe un vector no nulo  $(\xi_0, \mu) \in \mathbb{R}^n \times \mathbb{R}$  tal que

$$\xi_0^t(x - \bar{x}) + \mu[y - f(\bar{x})] \leq 0 \quad \forall (x, y) \in \text{epi}(f) \quad (1.6)$$

Note que  $\mu$  no es positivo, porque de otra manera, la desigualdad (1.6) será contradictoria al elegir  $y$  suficientemente grande.

Ahora mostraremos por contradicción que  $\mu < 0$ . Suponga que  $\mu = 0$ , entonces  $\xi_0^t(x - \bar{x}) \leq 0 \quad \forall x \in S$ . Como  $x \in \overset{\circ}{S}$ , existe  $\lambda > 0$  tal que  $\bar{x} + \lambda\xi_0 \in S$  y por lo tanto  $\lambda\xi_0^t\xi_0 \leq 0$ . Esto implica que  $\xi_0 = 0$  y  $(\xi_0, \mu) = 0$ , contradiciendo el hecho que  $(\xi, \mu)$  es un vector no nulo. Por lo tanto,  $\mu < 0$ . Denotando  $\frac{\xi_0}{|\mu|}$  por  $\xi$  y dividiendo la desigualdad (1.6) por  $|\mu|$  tenemos:

$$\xi^t(x - \bar{x}) - y + f(\bar{x}) \leq 0 \quad \forall (x, y) \in \text{epi}(f) \quad (1.7)$$

En particular, el hiperplano  $H = \{(x, y) \mid y = f(\bar{x}) + \xi^t(x - \bar{x})\}$  es soporte de  $\text{epi}(f)$  en  $(\bar{x}, f(\bar{x}))$ . Haciendo  $y = f(\bar{x})$  en (1.7) se obtiene  $f(x) \geq f(\bar{x}) + \xi^t(x - \bar{x}) \quad \forall x \in S$ .

□

El siguiente teorema establece que toda función convexa es subdiferenciable en el interior de su dominio [14].

**Teorema 1.11.** *Sea  $f$  una función convexa definida en un convexo  $S$ , entonces en cada punto  $x_0 \in \overset{\circ}{S}$  se tiene que  $\partial f(x_0) \neq \emptyset$ .*

### **Demostración**

El punto  $(x_0, f(x_0))$  es un punto frontera de  $\text{epi}(f)$  y como  $\text{epi}(f)$  es convexo, podemos aplicar el Teorema (1.5) tomando  $S = \text{epi}(f)$ , entonces existe un vector no nulo  $(a, \beta) \in \mathbb{R}^n \times \mathbb{R}$  tal que

$$\langle (a, \beta), (x, y) \rangle \leq \langle (a, \beta), (x_0, f(x_0)) \rangle$$

Para todo  $(x, y) \in \text{epi}(f)$ , esta relación es equivalente a:

$$\langle a, x \rangle \beta y \leq \langle a, x_0 \rangle + \beta f(x_0). \quad (1.8)$$

como el lado derecho es independiente de  $y$  y  $y \geq f(x)$ , para todo  $x \in S$ , si  $\beta > 0$ , el lado izquierdo se puede hacer tan grande como se quiera haciendo variar  $y$ , lo cual es una contradicción, por lo tanto,  $\beta \leq 0$ .

Si  $\beta = 0$  entonces  $a \neq 0$  luego, para todo  $x \in S$  se tiene:

$$\langle x, a \rangle \leq \langle x_0, a \rangle$$

Como  $x \in \overset{\circ}{S}$  entonces existe  $\varepsilon > 0$  tal que  $x_1 = x_0 + a\varepsilon \in S$ , tomando  $x = x_1$  en la última desigualdad obtenemos:  $\varepsilon \langle a, a \rangle \leq 0$  lo cual es absurdo. Luego  $\beta < 0$ , dividiendo ambos miembros de (1.8) por  $\beta$  y tomando  $y = f(x)$  obtenemos:

$$f(x) \geq f(x_0) + \left\langle -\frac{a}{\beta}, x - x_0 \right\rangle \quad \forall x \in S \quad (1.9)$$

así  $-\frac{a}{\beta} \in \partial f(x_0)$ , esto concluye la prueba.

□

### 1.3.2 Funciones convexas diferenciables

**Definición 1.14.** [2] Sea  $S$  un conjunto no vacío en  $\mathbb{R}^n$ , y sea  $f : S \mapsto \mathbb{R}$ . Entonces  $f$  se dice que es diferenciable en  $\bar{x} \in \overset{\circ}{S}$  si existe un vector  $\nabla f(\bar{x})$ , llamado vector gradiente; y una función  $\alpha : \mathbb{R}^n \mapsto \mathbb{R}$  tal que:

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^t (x - \bar{x}) + \|x - \bar{x}\| \alpha(\bar{x}, x - \bar{x}) \quad \forall x \in S$$

dónde  $\lim_{x \rightarrow \bar{x}} \alpha(\bar{x}, x - \bar{x}) = 0$ .

La función  $f$  se dice que es diferenciable en  $S' \subseteq S$  si ésta es diferenciable en cada punto de  $S'$ . Esta representación de  $f$  es llamada *expansión (serie de Taylor) de primer orden* de  $f$  en el punto  $\bar{x}$ .

Note que si  $f$  es diferenciable en  $\bar{x}$ , solo podría haber un vector gradiente el cual está dado por:

$$\nabla f(\bar{x}) = \left( \frac{\partial f(\bar{x})}{\partial x_1}, \dots, \frac{\partial f(\bar{x})}{\partial x_n} \right)^t \equiv (f_1(\bar{x}), \dots, f_n(\bar{x}))^t$$

dónde  $f_i(\bar{x}) = \frac{\partial f(\bar{x})}{\partial x_i}$  es la derivada parcial de  $f$  respecto a  $x_i$  en  $\bar{x}$ .

**Teorema 1.12.** Sea  $f$  una función convexa definida en un conjunto convexo  $S$  de  $\mathbb{R}^n$  y sea  $x$  un punto interior de  $S$ . Supóngase que  $f$  es diferenciable en  $x$  entonces posee un único subgradiente en  $x$ , es decir  $\partial f(x) = \{\nabla f(x)\}$ .

**Demostración** [2]

(Véase [2]). Por Teorema 1.10 el conjunto de subgradientes de  $f$  en  $\bar{x}$  es no vacío. Ahora,

sea  $\xi$  un subgradiente de  $f$  en  $\bar{x}$  como resultado del Teorema 1.10 y la diferenciabilidad de  $f$  en  $\bar{x}$ , para cualquier vector  $d$  y para  $\lambda$  suficientemente pequeña, se obtiene

$$\begin{aligned} f(\bar{x} + \lambda d) &\geq f(\bar{x}) + \lambda \xi^t d \\ f(\bar{x} + \lambda d) &= f(\bar{x}) + \lambda \nabla f(\bar{x})^t d + \lambda \|d\| \alpha(\bar{x}; \lambda d) \end{aligned}$$

Restando la ecuación de la desigualdad se obtiene:

$$0 \geq \lambda [\xi - \nabla f(\bar{x})]^t d - \lambda \alpha(\bar{x}; \lambda d)$$

Si dividimos por  $\lambda > 0$  haciendo  $\lambda \mapsto 0^+$ , se sigue que  $[\xi - \nabla f(\bar{x})]^t d \leq 0$ . Escogiendo  $d = \xi - \nabla f(\bar{x})$ , la última desigualdad implica que  $\xi = \nabla f(\bar{x})$ .

□

La siguiente proposición garantiza la existencia de la derivada direccional de las funciones convexas (bajo ciertas hipótesis generales).

**Proposición 1.4.** *Sea  $f : S \mapsto \mathbb{R}$ , convexa. Sean  $\bar{x} \in S$  y  $d \neq 0$  tales que existe  $\eta > 0$  que cumple  $\bar{x} + \lambda d \in S$ , para todo  $\lambda \in [0, \eta]$ . Entonces  $f'(\bar{x}; d)$  existe.*

**Demostración**

Sean  $0 < \lambda_1 < \lambda_2 < \eta$ . Entonces

$$f(\bar{x} + \lambda_1 d) = f\left(\frac{\lambda_1}{\lambda_2}(\bar{x} + \lambda_2 d) + \left(1 - \frac{\lambda_1}{\lambda_2}\right)\bar{x}\right) \leq \frac{\lambda_1}{\lambda_2} f(\bar{x} + \lambda_2 d) + \left(1 - \frac{\lambda_1}{\lambda_2}\right) f(\bar{x})$$

implica

$$\frac{f(\bar{x} + \lambda_1 d) - f(\bar{x})}{\lambda_1} \leq \frac{f(\bar{x} + \lambda_2 d) - f(\bar{x})}{\lambda_2}$$

Así, la función  $\varphi(\lambda) = \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda}$  es monótona (no decreciente) y por lo tanto  $\inf_{\lambda > 0} \varphi(\lambda)$  existe (estamos en  $\overline{\mathbb{R}}$ ).

Pero  $\inf_{\lambda > 0} \varphi(\lambda) = \lim_{\lambda \rightarrow 0^+} \varphi(\lambda) = f'(\bar{x}, d)$ , luego la derivada direccional existe.

□

El teorema siguiente permite caracterizar la convexidad de las funciones diferenciables, estableciendo que el hiperplano soporte de la función, en un punto cualquiera de la frontera del epígrafo, acota inferiormente a la función.

**Teorema 1.13. (Caracterización de convexidad en el caso diferenciable)** *Sea  $S$  un conjunto abierto no vacío en  $\mathbb{R}^n$  y sea  $f : S \mapsto \mathbb{R}$  diferenciable en  $\mathbb{R}$ . Entonces  $f$  es convexa si y sólo si para cualquier  $\bar{x} \in S$  se tiene:*

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^t (x - \bar{x}) \quad \forall x \in S$$



De igual forma,  $f$  es estrictamente convexa si y sólo si para cada  $x \in S$  se tiene:

$$f(x) > f(\bar{x}) + \nabla f(\bar{x})^t(x - \bar{x}) \quad \forall x \neq \bar{x} \in S$$

**Demostración [11]**

” $\Rightarrow$ ” Sea  $f$  convexa. Dado  $x, \bar{x} \in S$  se tiene que

$$f(\lambda x + (1 - \lambda)\bar{x}) \leq \lambda f(x) + (1 - \lambda)f(\bar{x}), \quad \forall \lambda \in [0, 1]$$

Reordenando tenemos que

$$\frac{f(\bar{x} + \lambda(x - \bar{x})) - f(\bar{x})}{\lambda} \leq f(x) - f(\bar{x}), \quad \forall \lambda \in [0, 1]$$

por Proposición (1.4) y tomando límite  $\lambda \mapsto 0^+$ :

$$f'(\bar{x}, d) = \nabla f(\bar{x})^t(x - \bar{x}) \leq f(x) - f(\bar{x})$$

que implica

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^t(x - \bar{x}), \quad \forall x, \bar{x} \in S$$

” $\Leftarrow$ ” Sean  $x, \bar{x} \in S$ . Entonces

$$f(x) \geq f(\lambda x + (1 - \lambda)\bar{x}) + \langle \nabla f(\lambda x + (1 - \lambda)\bar{x}), (1 - \lambda)(x - \bar{x}) \rangle \quad \forall \lambda \in [0, 1]$$

$$f(\bar{x}) \geq f(\lambda x + (1 - \lambda)\bar{x}) + \langle \nabla f(\lambda x + (1 - \lambda)\bar{x}), \lambda(\bar{x} - x) \rangle \quad \forall \lambda \in [0, 1]$$

multiplicando la primera desigualdad por  $\lambda$ , la segunda por  $(1 - \lambda)$  y sumando, se tiene

$$\lambda f(x) + (1 - \lambda)f(\bar{x}) \geq f(\lambda x + (1 - \lambda)\bar{x}) \quad \forall \lambda \in [0, 1]$$

es decir,  $f$  es convexa. □

El siguiente teorema da una caracterización suficiente y necesaria para funciones convexas. Para una función  $f$  de una variable, la caracterización se reduce a una pendiente creciente.

**Teorema 1.14.** *Sea  $S$  un conjunto abierto no vacío en  $\mathbb{R}^n$  y sea  $f : S \mapsto \mathbb{R}$  diferenciable en  $S$ . Entonces  $f$  es convexa si y sólo si para cada  $x_1, x_2 \in S$  se tiene:*

$$[\nabla f(x_2) - \nabla f(x_1)]^t(x_2 - x_1) \geq 0.$$

Similarmente,  $f$  es estrictamente convexa para cualquier  $x_1, x_2 \in S$  distintos, entonces se tiene:

$$[\nabla f(x_2) - \nabla f(x_1)]^t(x_2 - x_1) > 0.$$

***Demostración [11]***

De acuerdo al teorema anterior

$$\begin{aligned} f(x_1) &\geq f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle \quad \forall x_1, x_2 \in S \\ f(x_2) &\geq f(x_1) + \langle \nabla f(x_1), x_2 - x_1 \rangle \quad \forall x_1, x_2 \in S \end{aligned}$$

Sumando las desigualdades anteriores, se tiene que

$$0 \geq \langle \nabla f(x_2), x_1 - x_2 \rangle + f(x_1) + \langle \nabla f(x_1), x_2 - x_1 \rangle = \langle -\nabla f(x_2) + \nabla f(x_1), x_2 - x_1 \rangle$$

es decir,  $\langle \nabla f(x_2) - \nabla f(x_1), x_2 - x_1 \rangle \geq 0 \quad \forall x_1, x_2 \in S$ .

□

Aunque los teoremas (1.13) y (1.14) dan una caracterización suficiente y necesaria para funciones convexas, comprobar estas condiciones desde el punto de vista computacional es muy difícil. Una caracterización más simple y manejable, al menos para funciones cuadráticas puede obtenerse siempre que la función sea dos veces diferenciable.

**Funciones dos veces diferenciables [2]**

Una función  $f$  que es diferenciable en  $\bar{x}$  se dice que es dos veces diferenciable en  $\bar{x}$  si la representación de la *expansión de segundo orden (serie de Taylor)* de la siguiente definición existe.

**Definición 1.15.** *Sea  $S$  un conjunto no vacío en  $\mathbb{R}^n$  y sea  $f : S \rightarrow \mathbb{R}$ . Entonces  $f$  se dice que es dos veces diferenciable en  $\bar{x} \in \overset{\circ}{S}$  si existe un vector  $\nabla f(\bar{x})$ , y una matriz simétrica  $H(\bar{x})$  de  $n \times n$ , llamada **matriz hessiana** y una función  $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$  tal que:*

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^t(x - \bar{x}) + \frac{1}{2}(x - \bar{x})H(\bar{x})(x - \bar{x}) + \|x - \bar{x}\|^2 \alpha(\bar{x}, x - \bar{x})$$

para cada  $x \in S$ , donde  $\lim_{x \rightarrow \bar{x}} \alpha(\bar{x}, x - \bar{x}) = 0$ .

La función  $f$  se dice dos veces diferenciable en el conjunto abierto  $S' \subseteq S$  si es dos veces diferenciable en cada punto de  $S'$ .

Cabe señalar que para funciones dos veces diferenciables, la matriz hessiana  $H(\bar{x})$  está compuesta por las derivadas parciales de orden dos  $f_{ij}(\bar{x}) \equiv \frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j}$ ;  $i = 1, \dots, n$ .  $j = 1, \dots, n$  y está dada de la siguiente forma:

$$H(\bar{x}) = \begin{bmatrix} f_{11}(\bar{x}) & f_{12}(\bar{x}) & \dots & f_{1n}(\bar{x}) \\ f_{21}(\bar{x}) & f_{22}(\bar{x}) & \dots & f_{2n}(\bar{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1}(\bar{x}) & f_{n2}(\bar{x}) & \dots & f_{nn}(\bar{x}) \end{bmatrix}$$

**Ejemplo 1.2.** Sea  $f(x_1, x_2) = 2x_1 + 6x_2 - 2x_1^2 - 3x_2^2 + 4x_1x_2$ . Entonces tenemos

$$\nabla f(\bar{x}) = \begin{bmatrix} 2 - 4\bar{x}_1 + 4\bar{x}_2 \\ 6 - 6\bar{x}_2 + 4\bar{x}_1 \end{bmatrix} \quad y \quad H(\bar{x}) = \begin{pmatrix} -4 & 4 \\ 4 & -6 \end{pmatrix}$$

Por ejemplo, tomando  $\bar{x} = (0, 0)^t$ , la expansión de segundo orden de esta función está dada por

$$f(x_1, x_2) = (2, 6) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{1}{2}(x_1, x_2) \begin{pmatrix} -4 & 4 \\ 4 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

El Teorema 1.15 muestra que  $f$  es convexa si y sólo si su matriz hessiana es **semi-definida positiva (PSD)** en todo  $S$ ; esto es:

**Definición 1.16.** Para cualquier  $\bar{x} \in S$  tenemos que  $x^t H(\bar{x})x \geq 0$  para todo  $x \in \mathbb{R}^n$

**Teorema 1.15. (Caracterización de convexidad para una función dos veces diferenciable)** Sea  $S \subseteq \mathbb{R}^n$  un abierto, convexo, no vacío y sea  $f : S \rightarrow \mathbb{R}$  dos veces diferenciable en  $S$ . Entonces,  $f$  es convexa si y sólo si  $H(x)$  es semi-definida positiva  $\forall x \in S$

**Demostración [2]**

Supóngase que  $f$  es convexa y sea  $\bar{x} \in S$ . Necesitamos mostrar que  $x^t H(\bar{x})x \geq 0$  para cualquier  $x \in \mathbb{R}^n$ .

Como  $S$  es abierto, entonces dado cualquier  $x \in \mathbb{R}^n$ ,  $\bar{x} + \lambda x \in S$  para  $|\lambda| \neq 0$  y suficientemente pequeño, por Teorema 1.13 y la doble diferenciable de  $f$ , tenemos las dos siguientes expresiones

$$f(\bar{x} + \lambda x) \geq f(\bar{x}) + \lambda \nabla f(\bar{x})^t x \tag{1.10}$$

$$f(\bar{x} + \lambda x) = f(\bar{x}) + \lambda \nabla f(\bar{x})^t x + \frac{1}{2} \lambda^2 x^t H(\bar{x})x + \lambda^2 \|x\|^2 \alpha(\bar{x}, \lambda x) \tag{1.11}$$

sustrayendo (1.10) de (1.11), obtenemos

$$\frac{1}{2} \lambda^2 x^t H(\bar{x})x + \lambda^2 \|x\|^2 \alpha(\bar{x}, \lambda x) \geq 0$$

Dividiendo por  $\lambda^2 > 0$  y haciendo  $\lambda \rightarrow 0$ , se tiene que  $x^t H(\bar{x})x \geq 0$ .

Inversamente, supongamos que la matriz hessiana es definida semi-positiva en cada punto de  $S$ . Consideremos  $x, \bar{x} \in S$ . Entonces, por el teorema del valor medio tenemos:

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^t (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^t H(\hat{x})(x - \bar{x}) \tag{1.12}$$

donde  $\hat{x} = \lambda \bar{x} + (1 - \lambda)x$  para algún  $\lambda \in (0, 1)$ . Note que  $\hat{x} \in S$  y que  $H(\hat{x})$  es semi-definida positiva, por lo tanto  $(x - \bar{x})^t H(\hat{x})(x - \bar{x}) \geq 0$ , y de (1.12) se concluye que

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^t(x - \bar{x})$$

ya que la desigualdad anterior es cierta para cualquier  $x, \bar{x} \in S$ ,  $f$  es convexa por Teorema 1.13

□

Para entender esta condición, en el caso de una función en  $\mathbb{R}$  se tiene que cumplir que  $f''(x) \geq 0$ , y esto significa que la derivada es creciente (para ser exacto, no decreciente). En general  $\nabla f(x) \geq 0$  indica que la gráfica de la función tendrá curvatura positiva en  $x$

### 1.3.3 Mínimo y máximo de funciones convexas

En esta parte se considerarán problemas de minimizar y maximizar una función convexa sobre un conjunto convexo y se desarrollarán las condiciones necesarias y suficientes para optimalidad [2].

**Definición 1.17.** Sea  $f : \mathbb{R}^n \mapsto \mathbb{R}$  y consideremos el problema a minimizar  $f(x)$  sujeto a  $x \in S$ . Un punto  $x \in S$  es llamado **solución factible** para el problema. Si  $\bar{x} \in S$  y  $f(x) \geq f(\bar{x})$  para cada  $x \in S$ ,  $\bar{x}$  es llamado una **solución óptima**, **solución óptima global** o simplemente una **solución** para el problema.

La colección de soluciones óptimas es llamada **solución óptima alternativa**.

- Si  $\bar{x} \in S$  y si existe un  $\varepsilon$ -vecindario  $N_\varepsilon(\bar{x})$  alrededor de  $\bar{x}$  tal que  $f(x) \geq f(\bar{x})$  para cada  $x \in S \cap N_\varepsilon(\bar{x})$ ,  $\bar{x}$  es llamado **solución óptima local**.
- Si  $\bar{x} \in S$  y  $f(x) > f(\bar{x})$  para todo  $x \in S \cap N_\varepsilon(\bar{x})$ ,  $x \neq \bar{x}$  para algún  $\varepsilon > 0$ ,  $\bar{x}$  es llamada **solución óptima local estricta**.
- Si  $\bar{x} \in S$  es el único mínimo local en  $S \cap N_\varepsilon(\bar{x})$ , para algún vecindario  $N_\varepsilon(\bar{x})$  alrededor de  $\bar{x}$ ,  $x$  es llamada **solución óptima local fuerte o aislada**

Todos estos tipos de óptimo local o mínimo local a veces también se denominan como **mínimo relativo**. La Figura (1.15) ilustra ejemplos de mínimos locales y globales para el problema de minimizar  $f(x)$  sujeto a  $x \in S$

**Teorema 1.16.** Sea  $S$  un conjunto convexo no vacío en  $\mathbb{R}^n$  y sea  $f : S \mapsto \mathbb{R}$  convexa en  $S$ . Considere minimizar el problema  $f(x)$  sujeto a  $x \in S$ . Suponga que  $\bar{x} \in S$  es una solución óptima local para el problema, entonces:

1.  $\bar{x}$  es una solución óptima global.
2. Si  $\bar{x}$  es una solución mínima estricta o  $f$  es estrictamente convexa,  $\bar{x}$  es la única solución óptima global y también es un mínimo local.

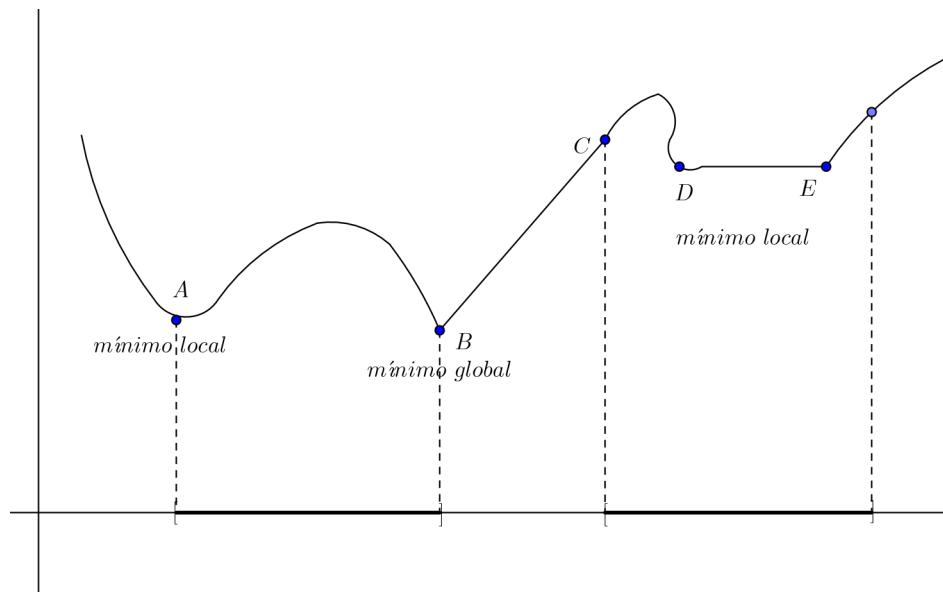


Figura 1.15: Mínimo local y global [2]

**Demostración [2] V**

1. Sea  $\varepsilon > 0$  tal que  $f(\bar{x}) \leq f(x)$ ,  $\forall x \in V_\varepsilon(\bar{x}) = \{x \in S \mid \|\bar{x} - x\| < \varepsilon\}$  y supongamos que  $\bar{x}$  no es óptimo global.

Es decir, existe  $y \in S$  tal que  $f(y) < f(\bar{x})$ . Luego, para  $\lambda \in ]0, 1[$

$$f(\lambda y + (1 - \lambda)\bar{x}) \leq \lambda f(y) + (1 - \lambda)f(\bar{x}) < \lambda f(\bar{x}) + (1 - \lambda)f(\bar{x}) = f(\bar{x})$$

para  $\lambda$  suficientemente pequeño (en realidad basta con elegir  $\lambda < \frac{\varepsilon}{\|y - \bar{x}\|}$ ), se tiene que  $\lambda y + (1 - \lambda)\bar{x} \in V_\varepsilon(\bar{x})$ , lo cual es una contradicción pues  $\bar{x}$  es el mínimo local en  $V_\varepsilon(\bar{x})$

2. Como  $f$  es estrictamente convexa, entonces  $f$  es convexa. Luego, por (1),  $\bar{x}$  es mínimo global.

Supongamos que no es único, esto es, que existe  $y \in S$  con  $y \neq \bar{x}$ , tal que  $f(y) = f(\bar{x})$ . Entonces

$$f\left(\frac{1}{2}y + \frac{1}{2}\bar{x}\right) < \frac{1}{2}f(y) + \frac{1}{2}f(\bar{x}) = f(\bar{x})$$

lo que implica que existe  $z \in S$ ,  $z = \frac{1}{2}y + \frac{1}{2}\bar{x}$ , distinto de  $\bar{x}$ , tal que  $f(z) < f(\bar{x})$ . Esto contradice el hecho que  $\bar{x}$  es mínimo global.

□

A continuación, se presenta una condición necesaria y suficiente para la existencia de una solución global. Si tal solución óptima no existe, entonces  $\inf\{f(x) : x \in S\}$  es finito, pero no se alcanza en ningún punto de  $S$  ó éste es igual a  $-\infty$ .

**Teorema 1.17.** *Sea  $f$  una función convexa en  $\mathbb{R}^n$  y  $S$  un conjunto convexo en  $\mathbb{R}^n$ . Consideremos el problema de optimización*

$$\min_{x \in S} f(x).$$

Entonces  $\bar{x}$  es un mínimo local de  $f$  sobre  $S$  si y sólo si existe  $s \in \partial f(\bar{x})$  tal que

$$\langle s, x - \bar{x} \rangle \geq 0 \quad \text{para toda } x \in S$$

**Demostración [2]**

” $\Rightarrow$ ” Supóngase que  $\langle s, x - \bar{x} \rangle \geq 0$  para toda  $x \in S$ , donde  $s$  es un subgradiente de  $f$  en  $\bar{x}$ . Por la convexidad de  $f$  tenemos:

$$f(x) \geq f(\bar{x}) + \langle s, x - \bar{x} \rangle \geq f(\bar{x}), \quad \forall x \in S$$

y por lo tanto  $\bar{x}$  es una solución óptima del problema dado.

” $\Leftarrow$ ” Supóngase que  $\bar{x}$  es una solución óptima para el problema, ahora definamos los siguientes conjuntos

$$\Lambda_1 = \{(x - \bar{x}, y) \mid x \in \mathbb{R}^n, y > f(x) - f(\bar{x})\}$$

$$\Lambda_2 = \{(x - \bar{x}, y) \mid x \in S, y \leq 0\}$$

donde  $\Lambda_1$  y  $\Lambda_2$  son convexos. Asimismo  $\Lambda_1 \cap \Lambda_2 = \emptyset$  porque sino existiría un punto  $(x, y)$  tal que

$$x \in S \quad 0 \geq y > f(x) - f(\bar{x})$$

contradiciendo la hipótesis de que  $\bar{x}$  es una solución óptima del problema. Por Teorema 1.6 existe un hiperplano que separa  $\Lambda_1$  y  $\Lambda_2$ , es decir, existe un vector no nulo  $(p, \mu)$  y un escalar  $\alpha$  tal que

$$\langle p, x - \bar{x} \rangle + \mu y \leq \alpha \quad \forall x \in \mathbb{R}^n, y > f(x) - \bar{x} \quad (1.13)$$

$$\langle p, x - \bar{x} \rangle + \mu y \geq \alpha \quad \forall x \in S, y \leq 0 \quad (1.14)$$

Si hacemos  $x = \bar{x}$  y  $y = 0$  en (1.14), se sigue que  $\alpha \leq 0$ . Continuando, hacemos  $x = \bar{x}$  y  $y = \varepsilon > 0$  en (1.13), se sigue que  $\mu\varepsilon \leq \alpha$ . Como esto es cierto para cada  $\varepsilon > 0$ ,  $\mu \leq 0$  y  $\alpha \geq 0$ . Para resumir, hemos demostrado que  $\mu \leq 0$  y  $\alpha = 0$ . Si  $\mu = 0$  de (1.13)  $\langle p, x - \bar{x} \rangle \leq 0$  para cada  $x \in \mathbb{R}^n$ . Si hacemos  $x = \bar{x} + p$ , se sigue que

$$0 \geq \langle p, x - \bar{x} \rangle = \|p\|^2$$

y por lo tanto  $p = 0$ . Como  $(p, \mu) \neq (0, 0)$ , se tiene que  $\mu > 0$ . Dividiendo (1.13) y (1.14) por  $-\mu$  y denotando  $\frac{p}{\mu}$  por  $s$ , se tienen las siguientes desigualdades:

$$y \geq \langle s, x - \bar{x} \rangle, \quad \forall x \in \mathbb{R}^n, \quad y > f(x) - f(\bar{x}) \quad (1.15)$$

$$\langle s, x - \bar{x} \rangle \geq 0, \quad \forall x \in S, \quad y \leq 0 \quad (1.16)$$

haciendo  $y = 0$  en (1.16) se obtiene  $\langle s, x - \bar{x} \rangle \geq 0$  para todo  $x \in \mathbb{R}^n$ . De (1.15) es evidente que

$$f(x) \geq f(\bar{x}) + \langle s, x - \bar{x} \rangle \quad \forall x \in \mathbb{R}^n$$

Por lo tanto,  $s$  es un subgradiente de  $f$  en  $\bar{x}$  con la propiedad de  $\langle s, x - \bar{x} \rangle \geq 0$  para todo  $x \in S$

□

### Maximizar una función convexa

Se desarrollará una condición necesaria para un máximo de una función convexa sobre un conjunto convexo. Desafortunadamente, esta condición no es suficiente. Por lo tanto, es improbable, que varios máximos que satisfagan la condición del Teorema 1.18 existan. En el caso de minimizar, no existe información local de tales soluciones que podrían llevar a mejores puntos. Por lo tanto, maximizar una función convexa es una tarea mucho más dura que minimizar una función convexa.

**Teorema 1.18.** *Sea  $f : \mathbb{R}^n \mapsto \mathbb{R}$  una función convexa y sea  $S$  un conjunto convexo no vacío en  $\mathbb{R}^n$ . Considere el problema:*

$$\max_{x \in S} f(x)$$

*Si  $\bar{x} \in S$  es una solución óptima local,*

$$\langle s, x - \bar{x} \rangle \leq 0$$

*para cada  $x \in S$ , donde  $s \in \partial f(\bar{x})$*

### **Demostración [2]**

Supóngase que  $\bar{x} \in S$  es una solución óptima local. Entonces existe un  $\varepsilon$ -vecindario  $N_\varepsilon(\bar{x})$  tal que  $f(x) \leq f(\bar{x}) \quad \forall x \in S \cap N_\varepsilon(\bar{x})$ . Sea  $x \in S$  y note que  $\bar{x} + \lambda(x - \bar{x}) \in S \cap N_\varepsilon(\bar{x})$  para  $\lambda > 0$  y suficientemente pequeño, por lo tanto:

$$f[\bar{x} + \lambda(x - \bar{x})] \leq f(\bar{x}) \quad (1.17)$$

Sea  $s$  un subgradiente de  $f$  en  $\bar{x}$ . Por la convexidad de  $f$  se tiene

$$f[\bar{x} + \lambda(x - \bar{x})] - f(\bar{x}) \geq \lambda \langle s^t, x - \bar{x} \rangle$$

La desigualdad anterior junto con (1.17) implica que  $\lambda \langle s, x - \bar{x} \rangle \leq 0$  y dividiendo por  $\lambda$  se sigue el resultado. □

## 1.4 Condiciones de optimalidad y dualidad

### *Optimización sin restricciones [11]*

En esta sección, obtendremos condiciones necesarias y suficientes para resolver problemas no lineales irrestrictos (que es el caso en que  $S = \mathbb{R}^n$ ), es decir, problemas del tipo:

$$(P) \quad \min_{x \in \mathbb{R}^n} f(x)$$

En general supondremos que  $f$  es una función una o dos veces continuamente diferenciable.

**Teorema 1.19. (Condiciones necesarias de optimalidad)** *Sea  $x^*$  la solución local de (P) y supongamos  $f \in C^2(S)$ , donde  $S$  es un conjunto abierto que contienen a  $x^*$ . Entonces*

a)  $\nabla f(x^*) = 0$

b) *La matriz Hessiana  $H(x^*)$  es semidefinida positiva.*

### *Demostración [11]*

Sea  $d \in \mathbb{R}^n$ , no nulo dado y consideremos la función  $g(\alpha) = f(x^* + \alpha d)$

a) Como  $x^*$  es un mínimo local de  $f$  se tiene que

$$f(x^*) \leq f(x^* + \alpha d) \quad \forall \alpha > 0 \text{ suficientemente pequeño}$$

Es decir

$$0 \leq \frac{f(x^* + \alpha d) - f(x^*)}{\alpha}$$

Tomando límite



$$0 \leq \lim_{\alpha \rightarrow 0^+} \frac{f(x^* + \alpha d) - f(x^*)}{\alpha} = g'(0)$$

Ahora, por la regla de la cadena se tiene que  $g'(\alpha) = \nabla f(x^* + \alpha d)^t d$ , lo que implica que  $g'(0) = \nabla f(x^*)^t d$ , por lo tanto

$$0 \leq \nabla f(x^*)^t d$$

Como  $d$  es arbitrario, tomando  $d = e_i$ , siendo  $e_i$  el  $i$ -ésimo vector de la base canónica, se deduce que

$$\frac{\partial f(x^*)}{\partial x_i} = 0 \Rightarrow \nabla f(x^*) = 0$$

b) Consideremos el desarrollo de Taylor de segundo orden

$$g(\alpha) = g(0) + g'(0)\alpha + \frac{1}{2}g''(0)\alpha^2 + o(\alpha^2)$$

donde la función  $o(t)$  cumple  $\lim_{t \rightarrow 0} \frac{o(t)}{t} = 0$ .

Considerando que  $g'(\alpha) = \nabla f(x^* + \alpha d)^t d$  y  $g''(\alpha) = d^t H(x^* + \alpha d) d$ , la expresión anterior se escribe

$$f(x^* + \alpha d) - f(x^*) = \alpha \nabla f(x^*)^t d + \frac{\alpha^2}{2} d^t H(x^*) d + o(\alpha^2)$$

por lo tanto, como  $\nabla f(x^*) = 0$  y  $f(x^* + \alpha d) - f(x^*) \geq 0$ , para  $\alpha > 0$  suficientemente pequeño, se tiene que

$$0 \leq d^t H(x^*) d + \frac{o(\alpha^2)}{\alpha^2}$$

Tomando  $\alpha \rightarrow 0$  se obtiene que  $H(x^*)$  es semidefinida positiva.

□

**Teorema 1.20.** *Sea  $f \in C^2(S)$ , con  $S$  abierto. Supongamos que  $x^* \in S$  satisface:*

a)  $\nabla f(x^*) = 0$

b) *La matriz  $H(x^*)$  es definida positiva*

Entonces  $x^*$  es un mínimo local estricto.

**Demostración [2]**

Como  $f$  es dos veces diferenciable en  $x^*$ , se debe tener para  $x^* \in \mathbb{R}^n$ ,

$$f(x) = f(x^*) + \nabla f(x^*)^t(x - x^*) + \frac{1}{2}(x - x^*)^t H(x^*)(x - x^*) + \|x - x^*\|^2 \alpha\langle x^*, x - x^* \rangle \quad (1.18)$$

Como  $\alpha\langle x^*, x - x^* \rangle \mapsto 0$  cuando  $x \mapsto x^*$ .

Por contradicción, supongamos que  $x^*$  no es un mínimo local estricto, es decir, suponga que existe una sucesión  $\{x_k\}$  convergiendo a  $x^*$  tal que  $f(x_k) \leq f(x^*)$ ,  $x_k \neq x^*$  y denotando  $\frac{x_k - x^*}{\|x_k - x^*\|}$  por  $d_k$ , por lo tanto, (1.18) implica que

$$\frac{1}{2}d_k^t H(x^*)d_k + \alpha\langle x^*, x - x^* \rangle \quad \forall k \quad (1.19)$$

Pero  $\|d_k\| = 1$  para cada  $k$ ; y por lo tanto existe una familia de índices  $\mathcal{K}$  tal que  $\{d_k\}_{\mathcal{K}}$  converge a  $d$ , donde  $\|d\| = 1$ . Considerando esta subsucesión y el hecho que  $\alpha\langle x^*, x_k - x^* \rangle \mapsto 0$  cuando  $k \in \mathcal{K}$  aproximándose a  $\infty$ , (1.19) implica que  $d^t H(x^*)d \leq 0$ . Esto contradice la hipótesis que  $H(x^*)$  es definida positiva cuando  $\|d\| = 1$ . Por lo tanto,  $x^*$  es efecto un mínimo local estricto.

□

### 1.4.1 Optimización con restricciones

Para el caso con restricciones, las condiciones de optimalidad son algo más complicadas que para el caso irrestricto. Comencemos con algunas definiciones previas.[11]

**Definición 1.18.** Sea  $f : \mathbb{R}^n \mapsto \mathbb{R}$ . Se dice que  $d \in \mathbb{R}^n$  es **dirección de descenso** de  $f$  en  $\bar{x}$  si  $\nabla f(\bar{x})^t d < 0$

**Definición 1.19.** Sea  $S$  un conjunto no vacío en  $\mathbb{R}^n$  y  $\bar{x} \in S$ . Se llama **cono de direcciones admisibles** de  $S$  en  $\bar{x}$  al conjunto:

$$\mathcal{A}(\bar{x}) = \{d \mid d \neq 0, \bar{x} + \lambda d \in S, \forall \lambda \in [0, \eta[ \text{ para algún } \eta > 0\}$$

Denotaremos por  $\mathcal{D}(\bar{x}) = \{d \mid \nabla f(\bar{x})^t d < 0\}$  al conjunto de direcciones de descenso.

Además, por simplicidad notacional y si no hay posible confusión, escribiremos  $\mathcal{A} = \mathcal{A}(\bar{x})$  y  $\mathcal{D} = \mathcal{D}(\bar{x})$ . Ambos conjuntos pueden ser vacíos.

**Teorema 1.21.** Sea  $f : \mathbb{R}^n \mapsto \mathbb{R}$  diferenciable en  $\bar{x} \in S$ . Si  $\bar{x}$  es mínimo local del problema

$$(P) \quad \min f(x) \tag{1.20}$$

$$x \in S \tag{1.21}$$

entonces  $\mathcal{A} \cap \mathcal{D} = \emptyset$

### **Demostración**

Razonando por contradicción, supongamos que existe  $d \in \mathcal{A} \cap \mathcal{D}$ , entonces

$$d \in \mathcal{A} \implies \exists \eta > 0 \text{ tal que } \bar{x} + \lambda d \in S, \lambda \in [0, \eta[$$

$$d \in \mathcal{D} \implies \varepsilon > 0 \text{ tal que } f(\bar{x} + \lambda d) < f(\bar{x}), \lambda \in [0, \varepsilon[$$

Luego,  $f(\bar{x} + \lambda d) < f(\bar{x})$  para todo  $\lambda \in ]0, \min\{\varepsilon, \eta\}[$  lo que contradice la minimalidad local de  $\bar{x}$ .

□

Muy frecuentemente el conjunto factible  $S$  está descrito mediante un sistema de inecuaciones, que denominaremos restricciones. Sean  $g_1, g_2, \dots, g_m$ , funciones de  $\mathbb{R}^n$  en  $\mathbb{R}$ , diferenciables, Consideremos también el conjunto

$$S = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, m\}$$

que es un convexo cuando las funciones  $g_1, \dots, g_m$  son convexas en  $\mathbb{R}^n$ .

**Definición 1.20.** Sea  $\bar{x} \in \mathbb{R}^n$ . Si una restricción  $g_i$  verifica  $g_i(\bar{x}) = 0$  entonces se dice que es **activa** en  $\bar{x}$ . Además  $I = \{i \mid g_i(\bar{x}) = 0\}$  denota el conjunto de índices de restricciones activas en  $\bar{x}$

Sea ahora el conjunto  $\mathcal{G} = \{d \mid \nabla g_i(\bar{x})^t d < 0, \forall i \in I\}$ , el cual denotaremos simplemente por  $\mathcal{G}$ .

**Teorema 1.22.** Sea  $\bar{x}$  un punto factible de (P), es decir,  $g_i(\bar{x}) \leq 0$   $i = 1, \dots, m$ . Entonces, si  $\bar{x}$  es un mínimo local de  $f$ , se cumple que  $\mathcal{G} \cap \mathcal{D} = \emptyset$

### **Demostración**

Basta probar que  $\mathcal{G} \subseteq \mathcal{A}$  y luego concluir usando el Teorema 1.21

Sea  $d \in \mathcal{D}$  y consideremos el caso  $i \in I$ . Entonces  $\nabla g_i(\bar{x})^t d < 0$ , luego existe  $\eta_i > 0$  tal que

$$g_i(\bar{x} + \lambda d) < g_i(\bar{x}) \quad \lambda \in ]0, \eta_i[$$

Notar que el lado derecho de esta última desigualdad es nulo, entonces se tiene que

$$g_i(\bar{x} + \lambda d) < 0 \quad \lambda \in ]0, \eta[$$

para  $\eta = \min_{i \in I} \{\eta_i\}$ .

Tomemos ahora  $i \notin I$ . Como  $\bar{x}$  es factible, entonces cumple  $g_i(\bar{x}) < 0$ . De allí, por continuidad, se deduce que existe  $\varepsilon_i > 0$  tal que  $g_i(\bar{x} + \lambda d) < 0 \quad \lambda \in [0, \varepsilon_i[$

Sea  $\varepsilon = \min_{i \notin I} \{\varepsilon_i\}$  y  $\Delta = \min\{\eta, \varepsilon\}$ . Entonces podemos decir que

$$g_i(\bar{x} + \lambda d) \leq 0 \quad i = 1, \dots, m \quad \forall \lambda \in [0, \Delta[$$

Entonces  $d \in \mathcal{A}$ , lo que completa la demostración. □

## 1.4.2 Problemas con restricciones de igualdad y desigualdades

**Teorema 1.23. (Karush-Huhn-Tucker, desigualdades)** Sea  $\bar{x}$  la solución de (P) y supongamos que  $\{\nabla g_i(\bar{x}) \mid i \in I\}$  es un conjunto linealmente independiente. Entonces existen  $\mu_1, \dots, \mu_m \in \mathbb{R}^n$  tales que

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i=1}^m \mu_i \nabla g_i(\bar{x}) &= 0 \\ \mu_i g_i(\bar{x}) &= 0 \quad i = 1, \dots, m \\ \mu_i &\geq 0 \quad i = 1, \dots, m \end{aligned}$$

### *Demostración* [11]

Sabemos del Teorema 1.21, que  $\mathcal{D} \cap \mathcal{G} = \emptyset$ . Entonces, no existe una dirección  $d \in \mathbb{R}^n$  tal que

$$\begin{aligned} \nabla f(\bar{x})^t d &< 0 \\ \nabla g_i(\bar{x})^t d &< 0 \quad i \in I \end{aligned}$$

Definiendo

$$A = \begin{bmatrix} \nabla f(\bar{x})^t \\ \vdots \\ \nabla g_i(\bar{x})^t \\ \vdots \end{bmatrix}_{i \in I}$$

Podemos interpretar lo anterior como: no existe  $d \in \mathbb{R}^n$  tal que  $Ad < 0$ . Entonces existe  $p \neq 0$  tal que

$$\begin{aligned} A^t p &= 0 \\ p &= 0 \end{aligned}$$

Notemos que

$$p = \begin{pmatrix} \mu_0 \\ \vdots \\ \mu_i \\ \vdots \end{pmatrix}_{i \in I}$$

entonces

$$\mu_0 \nabla f(\bar{x}) + \sum_{i \in I} \mu_i \nabla g_i(\bar{x}) = 0$$

donde los coeficientes  $\mu_0, \mu_i \geq 0$  para  $i \in I$ , no son todos nulos.

Afirmamos que  $\mu_0 > 0$ . En efecto, si  $\mu_0 = 0$ , entonces se tiene que  $\sum_{i \in I} \mu_i \nabla g_i(\bar{x}) = 0$ , donde los coeficientes  $\mu_i$ ,  $i \in I$  no son todos nulos. Eso contradice la hipótesis de independencia lineal.

Así, podemos redefinir los coeficientes  $\mu_i \leftarrow \frac{\mu_i}{\mu_0}$ ,  $\forall i \in I$ , con lo cual

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i \in I} \mu_i \nabla g_i(\bar{x}) &= 0 \\ \mu_i &= 0 \quad i \in I \end{aligned}$$

donde los coeficientes  $\mu_i$   $i \in I$  no son todos nulos.

También sabemos que  $g_i(\bar{x}) = 0$  para  $i \in I$  y que  $g_i(\bar{x}) < 0$  para  $i \notin I$ , entonces, definiendo  $\mu_i = 0 \quad \forall i \notin I$ , se concluye que

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i=1}^m \mu_i \nabla g_i(\bar{x}) &= 0 \\ \mu_i g_i(\bar{x}) &= 0 \quad i = 1, \dots, m \\ \mu_i &\geq 0 \quad i = 1, \dots, m \end{aligned}$$

lo que completa la demostración.

□

**Ejemplo 1.3.** Veremos aquí cuán importantes son las hipótesis del Teorema 1.23. El siguiente problema es un problema clásico, que tiene como solución óptima el punto  $x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

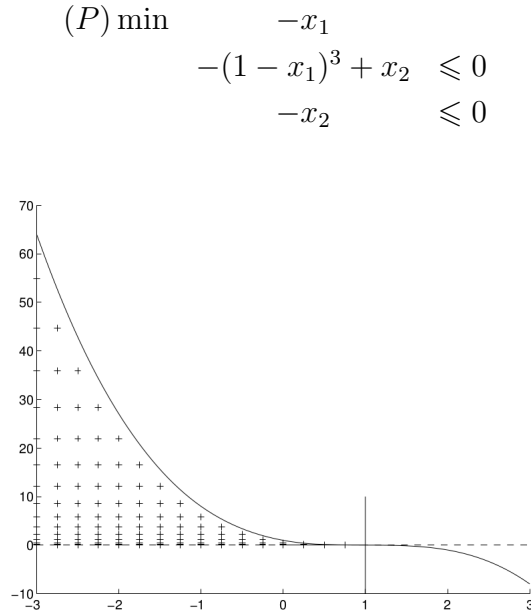


Figura 1.16: El Teorema de KKT no es aplicable en el óptimo

Como puede verse en la Figura 1.16, la versión del teorema de Karush-Kuhn-Tucker (KKT) no es aplicable en este caso, pues los gradientes de  $g_1$  y  $g_2$  no son linealmente independientes en  $x^*$ . En efecto:

$$\nabla g_1(x^*) = \begin{pmatrix} 3(1-x_1^*)^2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\nabla g_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

$$\nabla f(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Se observa que no existen  $\mu_1, \mu_2 \geq 0$  que satisfagan  $\begin{pmatrix} -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mu_1 + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \mu_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

A continuación, se presenta sin demostración el enunciado del teorema de Karush-Kuhn-Tucker, extendido al caso de restricciones de igualdad y desigualdad [11].

**Teorema 1.24.** Sean las funciones  $f : \mathbb{R}^n \mapsto \mathbb{R}$ ,  $g_i : \mathbb{R}^n \mapsto \mathbb{R}$   $i = 1, \dots, m$  y  $g_i : \mathbb{R}^n \mapsto \mathbb{R}$   $j = 1, \dots, l$  diferenciables. Consideremos el problema

$$(P) \min \quad f(x)$$

$$\quad \quad \quad g_i(x) \leq 0 \quad i = 1, \dots, m$$

$$\quad \quad \quad h_j(x) = 0 \quad j = 1, \dots, l$$

Sea  $\bar{x}$  un punto factible. Supongamos que el conjunto  $\{\nabla g_i(\bar{x}), \nabla h_j(\bar{x}) \mid i \in I, j = 1, \dots, l\}$  es linealmente independiente.

Entonces, si  $\bar{x}$  es solución de (P), existen  $\mu_i \in \mathbb{R}$ ,  $i = 1, \dots, m$ ,  $v_j \in \mathbb{R}$ ,  $j = 1, \dots, l$  tales que

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i=1}^m \mu_i \nabla g_i(\bar{x}) + \sum_{j=1}^l v_j \nabla h_j(\bar{x}) &= 0 \\ \mu_i g_i(\bar{x}) &= 0 \quad i = 1, \dots, m \\ \mu_i &\geq 0 \quad i = 1, \dots, m \end{aligned}$$

Tal como en el caso de desigualdades, la hipótesis de independencia lineal de los gradientes de las restricciones activas es demasiado fuerte y el teorema sigue siendo válido aún cuando esta hipótesis no se cumpla.

### 1.4.3 Dualidad Lagrangiana

Comenzaremos con la definición de la función lagrangiana y los multiplicadores de Lagrange, que nos permiten manejar las restricciones incorporándolas a la función objetivo. A continuación veremos qué son la función y el problema dual. Esta teoría hace que terminemos produciendo un problema convexo a partir de cualquier problema de optimización.

Partimos de un problema de optimización en forma estándar

$$\begin{aligned} \text{minimizar} \quad & f_0(x) \\ \text{sujeto a} \quad & f_i(x) \leq 0 \quad \text{para } i = 1, \dots, m \\ & h_i(x) = 0 \quad \text{para } i = 1, \dots, l \\ & x \in \mathbb{R}^n \end{aligned} \tag{1.22}$$

Se define la *función lagrangiana*  $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \mapsto \mathbb{R}$  asociada al problema (1.22) como

$$L(x, \lambda, v) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^l v_i h_i(x)$$

Llamamos a  $\lambda_i$  el multiplicador de Lagrange asociado a la desigualdad  $f_i(x) \leq 0$ , igual que los  $v_i$  son los multiplicadores de Lagrange de sus igualdades correspondientes  $h_i(x) = 0$ . De forma vectorial, consideramos a  $\lambda$  y  $v$  las variables duales o vectores multiplicadores de Lagrange de nuestro problema.

La *función dual de Lagrange* es el mínimo de la función lagrangiana con respecto a  $x$

$$\theta(\lambda, v) = \inf_{x \in D} L(x, \lambda, v)$$

Dado un problema de programación no lineal hay otro problema de programación no lineal asociado con el, el primero es llamado *problema primal* y el otro es llamado *problema dual*

*lagrangiano*. Bajo ciertas condiciones de convexidad y restricciones adecuadas, los problemas dual y primal poseen valores objetivos óptimos iguales, por lo tanto, es posible resolver el problema primal de manera indirecta resolviendo el problema dual [2].

Considere el siguiente problema de programación no lineal, el cual es llamado *problema primal*

**Problema primal P:**

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeto a} && g_i(x) \leq 0 \quad \text{para } i = 1, \dots, m \\ &&& h_i(x) = 0 \quad \text{para } i = 1, \dots, l \\ &&& x \in X \end{aligned}$$

El *problema dual lagrangiano* se indica a continuación.

**Problema dual lagrangiano D**

$$\begin{aligned} &\text{maximizar} && \theta(\lambda, v) \\ &\text{sujeto a} && \lambda, v \geq 0 \\ &\text{donde} && \theta(\lambda, v) = \inf \left\{ f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{i=1}^l v_i h_i(x) : x \in X \right\} \end{aligned}$$

Note que la función dual lagrangiana  $\theta$  puede tomar el valor de  $-\infty$  para algunos vectores  $(\lambda, v)$ . Como el problema dual consiste en maximizar el ínfimo (la mayor de las cotas inferiores) de la función  $f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{i=1}^l v_i h_i(x)$ , a veces se refieren a éste como *dual viable*.

Este problema es convexo sin depender de que el problema principal lo sea o no, para ello enunciamos la siguiente proposición [3].

**Proposición 1.5.** *La función dual de un problema de maximización es una función convexa en cualquier subconjunto convexo de su dominio, siempre que este no sea el vacío.*

**Demostración [3]**

Sea  $D_1$  un subconjunto convexo del dominio  $D$  de la función dual de un problema de maximización. Recordar que una función  $f$  es convexa si cumple con:

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) \quad \forall \alpha \in [0, 1]$$

Ahora procedemos a demostrar que la función dual  $\theta$  es convexa.

$\forall x, y \in D_1 \subseteq D$ . Por ser  $D_1$  un conjunto convexo se tiene que

$$\mu x + (1 - \mu)y \in D_1 \quad \forall \mu \in [0, 1] \quad \forall x, y \in D_1 \subseteq D$$



Por tanto,  $\forall \mu \in [0, 1] \quad \forall x, y \in D_1 \subseteq D$  se tiene que

$$\begin{aligned}
\theta(\lambda, v) &= \inf_{x, y \in D_1} L(\mu x + (1 - \mu)y, \lambda, v) \\
&= \inf_{x, y \in D_1} \{f(\mu x + (1 - \mu)y) + \lambda g(\mu x + (1 - \mu)y) + \\
&\quad v h(\mu x + (1 - \mu)y)\} \\
&= \inf_{x, y \in D_1} \{\mu f(x) + (1 - \mu)f(y) + \lambda(\mu g(x) + (1 - \mu)g(y)) + \\
&\quad v(\mu h(x) + (1 - \mu)h(y))\} \\
&= \inf_{x, y \in D_1} \{\mu(f(x) + \lambda g(x) + v h(x)) + \\
&\quad (1 - \mu)(f(y) + \lambda g(y) + v h(y))\} \\
&= \inf_{x, y \in D_1} \{\mu L(x, \lambda, v) + (1 - \mu)L(y, \lambda, v)\} \\
&\leq \mu \inf_{x, y \in D_1} L(x, \lambda, v) + (1 - \mu) \inf_{x, y \in X} L(x, \lambda, v) \\
&= \mu \theta(\lambda, v) + (1 - \mu) \theta(\lambda, v)
\end{aligned}$$

□

El siguiente teorema muestra que el valor objetivo de cualquier solución factible para el problema dual produce una cota inferior en el valor objetivo de cualquier solución factible para el problema primal [2].

**Teorema 1.25. Dualidad débil** Sea  $x$  una solución factible para el problema  $P$ ; esto es  $x \in X$ ,  $g(x) \leq 0$  y  $h(x) = 0$ . Sea  $(u, v)$  una solución factible para el problema  $D$ ; esto es  $u \geq 0$ . Entonces  $f(x) \geq \theta(u, v)$ .

### **Demostración**

Por definición de  $\theta$ , y como  $x \in X$ , tenemos:

$$\begin{aligned}
\theta(u, v) &= \inf\{f(y) + u^t g(y) + v^t h(y) \mid y \in X\} \\
&\leq f(x) + u^t g(x) + v^t h(x) \\
&\leq f(x)
\end{aligned}$$

ya que  $u \geq 0$ ,  $g(x) \leq 0$  y  $h(x) = 0$ . Esto completa la prueba.

□

Y se cumple sea o no convexo nuestro problema. La diferencia  $f(x) - \theta(x)$  (donde  $f(x)$  es solución del primal y  $\theta(x)$  la solución dual) es la distancia óptima de dualidad. En algunos casos se utiliza este valor como cota inferior para el valor óptimo cuando el problema es difícil de resolver, ya que el problema dual siempre será convexo y permitirá resolver de forma eficiente el valor de  $\theta(x)$ .

Cuando la distancia óptima de dualidad es cero decimos que hay dualidad fuerte, y significa que el problema dual de Lagrange nos da el mejor resultado posible (es decir, que la solución del primal y el dual coinciden). Para que esto se cumpla se tienen que dar unas determinadas condiciones que se conocen como hipótesis de cualificación. Uno de los posibles requisitos es la condición de Slater:

**Condición de Slater [14]**

Diremos que la condición de calificación de restricciones de Slater se satisfacen si

1. Las funciones  $g_i$  (de restricción) son convexas y las funciones  $h_j$  son afines.
2. Existe  $b$  tal que  $g_i(b) = 0$  para  $i = 1, \dots, m$

Es decir, tiene que existir un punto  $b$  estrictamente viable, lo que quiere decir que no sólo cumple las restricciones, sino que cumple las desigualdades con un margen positivo, o de forma estricta.

La condición de Slater también nos garantiza que si se cumple (y el problema es convexo) se puede encontrar una solución para el problema dual.

**Teorema 1.26. Holgura complementaria** Consideremos la pareja primal-dual

$$(P) \min \quad c^t x \\ Ax = b \\ x \geq 0$$

$$(D) \max \quad b^t y \\ A^t y \leq c$$

Sean  $x^*$  e  $y^*$  óptimos respectivos de (P) y (D), si  $s^* = c - A^t y^*$ . Entonces  $x^{*t} s^* = 0$ . Inversamente, si  $x^*$  e  $y^*$  son factibles para los problemas primal y dual respectivamente, y si  $x^{*t} s^* = 0$ , entonces las soluciones  $x^*$  e  $y^*$  son óptimas respectivas para (P) y (D).

**Demostración [11]**

Como  $x^*$  e  $y^*$  son óptimos respectivos de (P) y (D), entonces  $c^t x^* = b^t y^*$ , de donde se tiene

$$c^t x^* = b^t y^* = x^{*t} A^t y^* = x^*(c - s^*) = x^{*t} c - x^{*t} s^*$$

lo que implica que  $x^{*t} s^* = 0$ .

Inversamente,  $x^{*t} s^* = 0$  implica que  $x^{*t} c - x^{*t} A^t y^* = 0$ , es decir  $c^t x^* - b^t y^* = 0$ , que es la condición de optimalidad de  $x^*$  e  $y^*$ .

# Capítulo 2

## Control Óptimo Estocástico

### 2.1 Introducción

Muchos de los problemas de programación matemática incorporan parámetros que se suponen conocidos en el momento de resolver el problema. Sin embargo, si el problema de optimización es un modelo que representa una situación real en la que hemos de tomar una decisión, es frecuente que se desconozcan los valores de algunos de los parámetros que intervienen en el modelo. Este desconocimiento da lugar a que en el momento de adoptar una decisión se desconozcan las posibles consecuencias de la misma [16].

La teoría de control tiene como objetivo determinar decisiones o acciones de control dirigidas a optimizar el comportamiento de un sistema que evoluciona con el tiempo. Dependiendo del fenómeno a estudiar, la dinámica del sistema se puede modelar por medio de:

- Ecuaciones diferenciales, si el sistema se estudia en tiempo continuo.
- Ecuaciones en diferencia, si el sistema se estudia en tiempo discreto.

Además, estas ecuaciones pueden ser estocásticas o determinísticas dependiendo si el sistema incorpora elementos aleatorios o no [18].

En base a la información disponible acerca de los posibles resultados de una acción en cualquier proceso de toma de decisiones, podemos decir que cuando tomamos una decisión estamos ante una situación de [16]:

- **Certidumbre:** Si cada acción da lugar a un resultado conocido e invariable.
- **Riesgo:** Si cada acción lleva a un posible resultado y cada resultado lleva asociada una probabilidad de que ocurra, probabilidad conocida para el decisor (normalmente estudiada mediante técnicas estadísticas), la toma de decisiones está expuesta a la acción que se produzca finalmente en el futuro.

- **Incertidumbre:** Si cada acción tiene una consecuencia de entre un conjunto de posibles resultados, pero se desconocen las probabilidades de estos posibles resultados.

Para resolver problemas en los que se desconocen los valores de algunos de los parámetros que intervienen en el modelo (situaciones de riesgo o de incertidumbre), es posible adoptar distintas "soluciones". En determinadas circunstancias y en base a la información disponible acerca de los parámetros desconocidos, es posible "sustituir" estos valores por una estimación de los mismos, una medición no exacta de su valor esperado o bien tratar estos parámetros como variables aleatorias [16].

Asimismo, se dispone de información sobre el estado del sistema en el tiempo inicial y final, y se conoce la forma de la función objetivo a optimizar. El problema será elegir un control (o decisión) de entre todos los considerados admisibles que maximice la función objetivo para las restricciones impuestas por el sistema. Ese control considerado óptimo tiene la propiedad de determinar automáticamente la trayectoria del estado del sistema [17].

## 2.2 Probabilidad y procesos estocásticos

El propósito de esta sección es revisar tópicos de probabilidad, procesos estocásticos y cálculo estocástico cuyos resultados serán utilizados posteriormente [19].

**Definición 2.1.** Sea  $\Omega$  un conjunto no vacío. Considere una  $\sigma$ -álgebra,  $\mathcal{F}$ , la cual es una familia de subconjuntos de  $\Omega$  tal que:

i)  $\Omega \in \mathcal{F}$

ii) Si  $A \in \mathcal{F}$ , entonces  $A^c \in \mathcal{F}$ , donde  $A^c = \Omega - A$ , es el complemento de  $A$ .

iii) Si  $A_1, A_2, \dots, A_i, \dots$  es una sucesión de conjuntos en  $\mathcal{F}$ , entonces  $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

En particular, si  $\mathcal{F}$  es una  $\sigma$ -álgebra de  $\Omega$ , se dice que el par  $(\Omega, \mathcal{F})$  es un espacio medible. Si  $A$  es el conjunto en  $\mathcal{F}$ , se dice que  $A$  es  $\mathcal{F}$  medible (o simplemente que  $A$  es medible con respecto de  $\mathcal{F}$ ).

**Definición 2.2.** Sea  $(\Omega, \mathcal{F})$  un espacio medible y sea  $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$  el conjunto de los reales extendidos. Se dice que una función  $\mu : \mathcal{F} \rightarrow \overline{\mathbb{R}}$  es una medida sobre  $\mathcal{F}$  si:

i)  $\mu(\emptyset) = 0$

ii)  $\mu(A) \geq 0$  para todo  $A \in \mathcal{F}$

iii)  $\mu$  es  $\sigma$ -aditiva, es decir, si  $\{A_i\}$  es una sucesión de conjuntos disjuntos (dos a dos) en  $\mathcal{F}$  entonces:

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

En este caso se dice que  $(\Omega, \mathcal{F}, \mu)$  **es un espacio de medida**. A  $\mu(A)$  se le llama la medida del conjunto  $A \in \mathcal{F}$ . En particular, si  $\mu(\Omega) = 1$ , se dice que  $\mu$  es una medida de probabilidad. Si  $\mu$  es una medida de probabilidad, se acostumbra a escribir  $\mu \equiv P$  y se dice que  $(\Omega, \mathcal{F}, P)$  es un espacio de probabilidad. Así mismo, se dice que  $P(A)$  es la probabilidad del evento  $A \in \mathcal{F}$ . También se dice que un evento  $A \in \mathcal{F}$  es casi seguro (y se abrevia c.s) si  $P(A) = 1$ .

Cuando se tienen dos conjuntos arbitrarios  $\Omega$  y  $\Omega'$ , y una función  $f : \Omega \rightarrow \Omega'$ , para cualquier conjunto  $B \subset \Omega'$ , se define la función inversa mediante

$$f^{-1}(B) = \{\omega \in \Omega \mid f(\omega) \in B\}$$

**Definición 2.3.** Sea  $(\Omega, \mathcal{F}, \mu)$  un espacio de medida. Se dice que  $X : \Omega \rightarrow \mathbb{R}$  es una función  $\mathcal{F}$ -medible si  $X^{-1}(-\infty, x] = \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F} \quad \forall x \in \mathbb{R}$ .

En particular, sea  $\mathcal{B}(\mathbb{R})$  el conjunto de Borel, es decir, la  $\sigma$ -álgebra más pequeña que contiene todos a los intervalos de  $\mathbb{R}$ . En este caso, se dice que una función  $X : \Omega \rightarrow \mathbb{R}$  es una función  $\mathcal{F}$ -medible si  $X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\} \in \mathcal{F} \quad \forall B \in \mathcal{B}(\mathbb{R})$

Nótese que si  $X^{-1}(B) \in \mathcal{F}$ , entonces su medida  $\mu(X^{-1}(B))$  está bien definida para el caso especial donde  $\mu \equiv P$  es una medida de probabilidad, en cuyo caso se dice que  $X$  es una variable aleatoria.

**Definición 2.4.** Sea  $X$  una variable aleatoria. La función de distribución de  $X$  es la función  $F_X : \mathbb{R} \rightarrow [0, 1]$  dada por

$$F_X = P(X \leq x) \quad \forall x \in \mathbb{R}$$

La medida de probabilidad inducida por  $X$  es la medida de probabilidad sobre el conjunto Boreliano  $\mathcal{B}(\mathbb{R})$  definida  $P_X(B) = P\{X \in B\}$  para todo  $B \in \mathcal{B}(\mathbb{R})$ .

Nótese que  $P_X(B) = P[X^{-1}(B)]$ . Además, la función de distribución de  $X$  y la medida de probabilidad inducida por  $X$  están relacionadas mediante  $F_X(x) = P(-\infty, x]$

Las variables aleatorias que se considerarán,  $\xi$ , serán discretas, es decir, toman un número finito de valores,  $\xi^\omega$ ,  $\omega \in \Omega$ , de manera que cada elemento acontece con probabilidad  $P(\xi = \xi) = w^\omega$  tal que  $\sum_{\omega \in \Omega} w^\omega = 1$ .

Se define la **función de distribución acumulativa** como

$$F(\xi) = P(\{\omega \in \Omega \mid \xi \leq \xi\}) = P(\xi \leq \xi)$$

Además, la **esperanza** de una variable aleatoria se calcula como

$$E[\xi] = \sum_{\omega \in \Omega} w^\omega \xi^\omega$$

y la **varianza** es

$$Var[\xi] = E[\xi - E[\xi]]^2$$

Dado  $\alpha \in (0, 1)$ , un punto  $\eta$  se llama  $\alpha$ -cuantil de  $\xi$  si y sólo si  $\eta = \min\{\xi \in \xi \mid F(\xi) \geq \alpha\}$

### 2.2.1 Procesos estocásticos

En el estudio de las variables aleatorias se han explorado las características aleatorias del fenómeno pero se ha mantenido una premisa por defecto, que esas características aleatorias permanecen constantes a través del tiempo. Al incluir en el estudio la presencia de la variable determinística *tiempo* se está considerando que, de alguna forma, la variable aleatoria depende del tiempo [32].

En otras palabras, la variable aleatoria dependerá del fenómeno probabilístico y del tiempo. En consecuencia, cualquier función que se establezca en términos de la variable aleatoria, serán también dependientes del tiempo.

**Definición 2.5.** *Un proceso estocástico es una colección o familia de variables aleatorias  $\{X_t \mid t \in T\}$ , ordenadas según el subíndice  $t$  que en general se suele identificar con el tiempo.*

Por tanto, para cada instante  $t$  tendremos una variable aleatoria distinta representada por  $X_t$ , con lo que un proceso estocástico puede interpretarse como una sucesión de variables aleatorias cuyas características pueden variar a lo largo del tiempo.

A los posibles valores que puede tomar la variable aleatoria se le denominaran estados, por lo que se puede tener un espacio de *estados* discreto y un espacio de estados continuo. Por otro lado, la variable tiempo puede ser de tipo discreto o de tipo continuo. En el caso del tiempo discreto se podría tomar como ejemplo que los cambios de estado ocurran cada día, cada mes, cada año, etc.. En el caso del tiempo continuo, los cambios de estado se podrían realizar en cualquier instante.

Por tanto, dependiendo de cómo sea el conjunto de subíndices  $T$  y el tipo de variable aleatoria dado por  $X_t$  se puede establecer la siguiente clasificación de los procesos estocásticos:

- *Cadena:* Proceso de estado discreto y tiempo discreto (unidades producidas de un producto).

- *Proceso de saltos puros*: Procesos de estado discreto y tiempo continuo (unidades producidas hasta el instante  $t$ ).
- *Proceso de estado continuo y tiempo discreto*:  $X$  continua y  $t$  discreto (toneladas de producción diaria de un producto).
- *Proceso continuo*: Proceso de estado continuo y tiempo continuo (velocidad de un vehículo en el instante  $t$ ).

Una **Cadena** es un proceso estocástico en el cual el tiempo se mueve en forma discreta y la variable aleatoria sólo toma valores discretos en el espacio de estados. Un Proceso de **Saltos Puros** es un proceso estocástico en el cual los cambios de estados ocurren en forma aislada y aleatoria pero la variable aleatoria sólo toma valores discretos en el espacio de estados. En un **Proceso Continuo** los cambios de estado se producen en cualquier instante y hacia cualquier estado dentro de un espacio continuo de estados [32].

## 2.3 Planteamiento del problema de control óptimo

En un Problema de Control Óptimo (PCO) debemos especificar [21]:

1. Como evoluciona el sistema de interés, es decir, se debe especificar el *modelo dinámico* del sistema.
2. Cómo se va controlar el sistema, es decir, se debe especificar el *conjunto de estrategias admisibles* también llamadas *políticas de control* o simplemente *políticas o estrategias*.
3. *Restricciones adicionales* (por ejemplo, en el estado del sistema o en las estrategias), si es necesario hacerlas.
4. Cómo se va “medir” la respuesta del sistema a cada una de las distintas estrategias admisibles, es decir, se debe especificar la *función objetivo*, también llamada *índice de funcionamiento*.

Dadas estas cuatro componentes, el PCO se puede resumir como sigue: *optimizar la función objetivo sujeta a las restricciones en 1, 2, y 3*.

Se considera un sistema dinámico, formulado en tiempo discreto, para un número dado  $N$  de etapas o períodos, cuya situación inicial viene dada por el vector  $n$ -dimensional  $x_0$  y que evoluciona en el tiempo, dicha evolución depende del valor que se da a ciertas variables, llamadas *variables de control*, que permiten influir en el sistema [4].

- Sea  $a(k)$  el vector  $m$ -dimensional de variables de control en la etapa o período  $k$ , para  $k \in \{0, 1, \dots, N - 1\}$ .
- Representamos por  $x(k)$ , para cada  $k \in \{0, 1, \dots, N - 1\}$ , el vector  $n$ -dimensional llamado *vector de variables de estado*, que nos indica la situación del sistema en la etapa o período  $k$ .



La evolución del sistema en el tiempo viene descrita por un sistema de ecuaciones en diferencias finitas, conocido como **ecuación de estado**:

$$\begin{aligned} x(k+1) &= f(x(k), a(k), k) \quad \text{para } k = 0, 1, \dots, N-1 \\ \text{con: } x(0) &= x_0 \end{aligned}$$

Siendo  $f$  una función:

$$\begin{aligned} f : \mathbb{R}^n \times \mathbb{R}^m \times \{0, 1, \dots, N-1\} &\longmapsto \mathbb{R}^n \\ (x, a, k) &\longmapsto f(x, a, k) \end{aligned}$$

Se supone que para cada  $k$ ,  $\Omega(k)$  es el conjunto de controles admisibles. La función objetivo es del tipo:

$$J = \sum_{k=0}^{N-1} F[x(k), a(k), k] + S[x(N)]$$

siendo  $F$  y  $S$  funciones definidas como:

$$\begin{aligned} F : \mathbb{R}^n \times \mathbb{R}^m \times \{0, 1, \dots, N-1\} &\longmapsto \mathbb{R} \\ (x, a, k) &\longmapsto F(x, a, k) \end{aligned}$$

y

$$\begin{aligned} S : \mathbb{R}^n &\longmapsto \mathbb{R} \\ x &\longmapsto S(x) \end{aligned}$$

respectivamente.

Es decir, el sistema dinámico parte del estado inicial  $x_0$ . En el período o etapa 1 (correspondiente a  $k = 0$ ), hay que elegir un control  $a(0) \in \Omega(0)$ , en dicho período o etapa se realiza una aportación a la función objetivo dada por  $F[x(0), a(0), 0]$  y se inicia el período o etapa 2 (correspondiente a  $k = 1$ ) con el siguiente valor del vector de estado:

$$x(1) = f(x(0), a(0), 0)$$

En dicho período o etapa 2 (correspondiente a  $k = 1$ ), hay que elegir un control  $a(1) \in \Omega(1)$ , se realiza una aportación a la función objetivo dada por  $F(x(1), u(1), 1)$  y se inicia el período o etapa 3 (correspondiente a  $k = 2$ ) con el siguiente valor del vector de estado:

$$x(2) = f(x(1), a(1), 1)$$

Se sigue de esta manera hasta que por último comienza el período o etapa  $N$  (correspondiente a  $k = N-1$ ) con estado inicial  $x(N-1)$ , y hay que elegir un control  $a(N-1) \in \Omega(N-1)$ , se

realiza una aportación a la función objetivo dada por  $F[x(N-1), a(N-1), N-1]$ , alcanzando el sistema un estado final

$$x(N) = f(x(N-1), a(N-1), N-1)$$

y por el hecho de terminar en dicho estado, se realiza una aportación a la función objetivo dada por  $S[x(N)]$ . Véase la Figura 2.1.

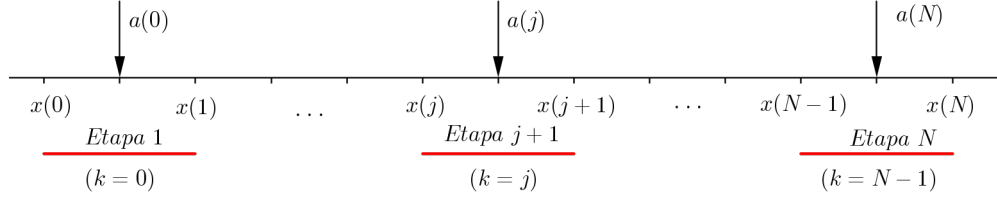


Figura 2.1: Control óptimo en tiempo discreto

Un control óptimo se define como un control

$$a(k) \in \Omega(k) \subset \mathbb{R}^m \quad \text{para } k = 0, 1, \dots, N-1$$

(por tanto admisible), que maximiza la función objetivo.

Por tanto, el problema que nos ocupa es el siguiente:

*Dado un sistema con condición inicial  $x_0$  y que evoluciona en el tiempo de acuerdo a la ecuación de estado  $x(k+1) = f(x(k), a(k), k)$ , se trata de encontrar para cada  $k$  un vector de control que sea admisible y haga que la función objetivo alcance el valor máximo. Expresando en términos matemáticos se tratará de:*

$$\begin{aligned} \max_{\{a(k)\}_{k=0}^{N-1}} \quad & J = \sum_{k=0}^{N-1} F[x(k), a(k), k] + S(N) & (2.1) \\ \text{sujeto a:} \quad & x(k+1) = f(x(k), a(k), k) \quad \text{para } k = 0, 1, \dots, N-1 \\ \text{con:} \quad & x(0) = x_0 \\ & a(k) \in \Omega(k) \end{aligned}$$

La secuencia de controles  $a^* = (a^*(0), a^*(1), \dots, a^*(N-1))'$  que resuelve el problema se llama control óptimo y  $x^* = (x^*(0), x^*(1), \dots, x^*(N))'$ , determinado por la ecuación de estado a partir de  $x(0)$ , se llama *trayectoria de estado óptima* o *camino óptimo*.

Obsérvese que no se ha exigido ninguna condición de diferenciabilidad a ninguna de las funciones en el planteamiento del problema.

## 2.4 Resolución del PCO por métodos usuales

En multitud de ocasiones la ley de control de un determinado sistema de control se decide por criterios de maximización de beneficios o minimización de esfuerzos. A tal fin, se define una función en el espacio de trayectorias del sistema y se aplican técnicas de optimización, especialmente adaptadas a la estructura de nuestro problema. Veremos dos de estas técnicas: la programación dinámica y el principio del máximo de Pontryagin, que proporcionan visiones complementarias del problema.

### 2.4.1 El principio de máximo

Hay varias técnicas generales para estudiar Problemas de Control Óptimo (PCO) como son el análisis convexo y la programación lineal (usualmente en espacios euclidianos de dimensión infinita). Sin embargo, por razones computacionales, en la mayoría de las aplicaciones las técnicas más usadas son el *principio del máximo* (que algunos autores llaman el principio de Pontryagin) y la programación dinámica [21].

Ahora bien, considérese el PCO determinístico (todo esto con el fin de ver como utilizar este principio) con espacio de estados  $X = \mathbb{R}^n$ , espacio de acciones  $A = \mathbb{R}^m$  y modelo dinámico [21]

$$x_{k+1} = F(k, x_k, a_k) \quad \forall k = 0, 1, \dots, N-1 \quad (2.2)$$

con estado inicial  $x = x_0$ . La función objetivo que se desea minimizar es el costo total:

$$V(\pi, x) = \sum_{k=0}^{N-1} L(k, x_k, a_k) + C(x_N) \quad (2.3)$$

sobre el conjunto de estrategias  $\pi = \{a_k\}$ .

A grandes rasgos, la idea del principio del máximo consiste en usar el método de *multiplicadores de Lagrange* para minimizar (2.3) sujeto a la "restricción" (2.2) en la forma

$$x_{k+1} - F(k, x_k, a_k) = 0 \quad \forall k = 0, 1, \dots, N-1 \quad (2.4)$$

Después introducimos los multiplicadores de Lagrange  $p_0, p_1, \dots, p_N$  en  $\mathbb{R}^n$ , a los que llamaremos **vectores adjuntos** (también llamados vectores de co-estado) y definimos el Lagrangiano

$$\tilde{V}(\pi, x, p_*) = V(\pi, x) + \sum_{k=0}^{N-1} p_{k+1} \cdot [x_{k+1} - F(k, x_k, a_k)] \quad (2.5)$$

donde  $p_* = \{p_0, p_1, \dots, p_N\}$ . Por lo tanto, sustituyendo (2.3) en (2.4) y usando el hamiltoniano definido para cada  $k = 0, 1, \dots, N-1$ :

$$H(k, x_k, a_k, p_{k+1}) = p_{k+1} \cdot F(k, x_k, a_k) - L(k, x_k, a_k) \quad (2.6)$$

un poco de álgebra nos permite reescribir (2.4) como:

$$\tilde{V}(\pi, x, p_{\bullet}) = \sum_{k=0}^{N-1} [p_k \cdot x_k - H(k, x_k, a_k, p_{k+1})] + C(x_N) + p_N \cdot x_N - p_0 \cdot x_0$$

Finalmente, bajo las siguientes hipótesis y usando resultados de optimización se obtiene el teorema de **El principio del máximo**

### Hipótesis 2.3

- Para cada  $k = 0, 1, \dots, N - 1$ , las funciones de costo  $L(k, x, a)$  y  $C(x)$  son de clase  $C^1$  en  $x$  y  $a$ .
- Para cada  $k = 0, 1, \dots, N - 1$  y  $a \in A$ , la función  $F(k, x, a)$  es de clase  $C^1$  en  $x$ .
- Para cada  $k = 0, 1, \dots, N - 1$  y  $x \in \mathbb{R}^n$ , el conjunto  $\{F(k, x, a) | a \in A\}$  es convexo

**Teorema 2.1.** (*El principio del máximo - caso determinístico, tiempo discreto*). Suponga que se cumple la **Hipótesis** (2.3). Supóngase también que existe una estrategia óptima  $a_{\bullet}^* = \{a_k^* | k = 0, 1, \dots, N - 1\}$  para el PCO (2.2) - (2.3) y sea  $x_{\bullet}^* = \{x_k^* | k = 0, 1, \dots, N\}$  la trayectoria correspondiente que se obtiene de (2.2) con estado inicial  $x_0^* = x_0$ . Entonces existe un conjunto  $p_{\bullet} = \{p_0, p_1, \dots, p_N\}$  de vectores adjuntos que satisfacen la **ecuación adjunta**

$$p_k = H_x(k, x_k^*, a_k^*, p_{k+1}) \quad \forall k = 0, \dots, N - 1 \quad (2.7)$$

es decir:

$$p_k = F_x(k, x_k^*, a_k^*)' p_{k+1} - L_x(k, x_k^*, a_k^*)$$

con la **condición terminal**

$$p_N = C_x(x_N^*) \quad (2.8)$$

y la **maximización del hamiltoniano**

$$H(k, x_k^*, a_k^*, p_{k+1}) = \max_{a \in A} H(k, x_k^*, a, p_{k+1}) \quad (2.9)$$

para  $k = 0, 1, \dots, N - 1$ .

El nombre principio del máximo para el Teorema 2.1 viene precisamente de la condición (2.9).

**Observación 2.1.** El Teorema 2.1 da **condiciones necesarias de optimalidad**, a saber, la existencia de la sucesión  $p_{\bullet}$  de vectores adjuntos que satisfacen (2.7), (2.8) y (2.9). Bajo hipótesis adecuadas estas condiciones también son suficientes. De hecho, si tales condiciones se satisfacen, entonces la búsqueda de un par óptimo  $(a_{\bullet}^*, x_{\bullet}^*)$  se reduce a resolver un **problema con valores de frontera** que consiste de:

- Las ecuaciones (2.2) y (2.7), que también se conocen como las **ecuaciones canónicas del PCO**.
- las condiciones de frontera de (2.8) y  $x_0^* = x_0$ .

c) La maximización del hamiltoniano. es decir, encontrar  $a_k^*$  tal que:

$$H(k, x_k^*, a_k^*, p_{k+1}) = \max_{a \in A} H(k, x_k^*, a, p_{k+1}).$$

### 2.4.2 Programación dinámica

La programación dinámica, introducida por Bellman (1957), fue creada inicialmente para resolver problemas formulados en tiempo discreto, aunque posteriormente sería adaptada para la resolución de problemas en tiempo continuo. La programación dinámica lo que hace es resolver un problema de  $N$  etapas o períodos, mediante la resolución de  $N$  problemas de una etapa o un período [4].

El problema de control óptimo en tiempo discreto (2.1) (Planteamiento del problema estocástico) del capítulo 2, Verifica una propiedad importante, llamada de *causalidad* que se expresa en los siguientes términos [4]:

**Propiedad de causalidad.** Para cualesquiera  $j, r \in \{0, 1, \dots, N-1\}$ , con  $j < r$ , se verifica que  $x(r)$  depende únicamente de  $x(j)$  y de los controles  $\{a(j), a(j+1), \dots, a(r-1)\}$ .

Es decir, dado el estado de  $x(j)$  en el que se encuentra el sistema dinámico al comienzo de la etapa (o período)  $j+1$ , para cualquier etapa posterior  $r$  se verifica que el estado que se alcanzará al finalizar dicha etapa,  $x(r)$  depende exclusivamente del estado  $x(j)$  y de los controles que se apliquen entre las etapas  $j+1$  y  $r$  tal como se aprecia en la Figura (2.2)

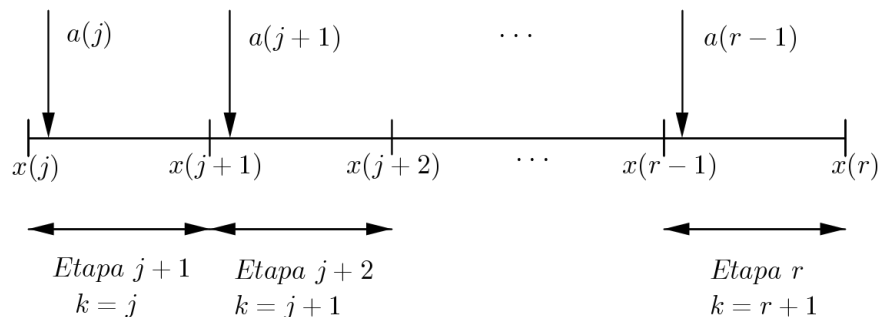


Figura 2.2: Propiedad de causalidad

En otras palabras, dados dos estados cualesquiera del sistema, uno anterior y otro posterior, el valor que tome el estado posterior depende únicamente del valor del estado anterior y de los valores de los controles intermedios entre ambos estados.

Esta propiedad se verifica como consecuencia directa de la estructura del problema. En efecto, teniendo en cuenta la ecuación de estado se tiene que:

$$x(j+1) = f(x(j), a(j), j)$$

y que

$$\begin{aligned} x(j+2) &= f(x(j+1), a(j+1), j+1) \\ &= f(f(x(j), a(j), j), a(j+1), j+1) \end{aligned}$$

y así sucesivamente, hasta obtener:

$$\begin{aligned} x(r) &= f(x(r-1), a(r-1), r-1) \\ &= f(f(x(r-2), a(r-2), r-2), a(r-1), r-1) \\ &\vdots \\ &= \Psi[x(j), a(j), a(j+1), \dots, a(r-1)] \end{aligned}$$

en donde esta última función  $\Psi$  se obtiene al ir sustituyendo de manera recurrente el vector de variables de estado por el valor que resulta al aplicar la ecuación de estado.

Como consecuencia de esta propiedad, el estado inicial  $x(0)$  y el conjunto de controles

$$\{a(0), a(1), \dots, a(N-1)\}$$

determinan la trayectoria del vector de estado  $\{x(0), x(1), \dots, x(N)\}$ . Utilizando la notación:

$$a[0, N-1] = \{a(0), a(1), \dots, a(N-1)\}$$

y en general:

$$a[j, N-1] = \{a(j), a(j+1), \dots, a(N-1)\}$$

la función objetivo de la sección 2.2.1 se puede escribir como

$$J = J_0\{x(0), a[0, N-1]\}$$

Por lo que si  $x(0)$  está dado, para maximizar  $J$  sólo hay que determinar los controles  $a[0, N-1]$ .

La propiedad de causalidad no siempre se cumple en modelos económicos dinámicos. Por ejemplo: una decisión a tomar en cierto momento futuro por una autoridad económica puede estar afectando al estado presente de dicha economía por medio de las expectativas de los agentes. En tales casos, la formulación del modelo de optimización dinámica no corresponde con el problema (2.1) en el que, como se ha comprobado, se cumple el supuesto de causalidad [4].

Consideremos de nuevo el PCO determinístico en (2.2) y (2.3) pero por conveniencia notacional escribiremos las variables  $x_k$  y  $a_k$  como  $x(k)$  y  $a(k)$ , respectivamente. Así pues, tenemos el modelo dinámico [21]:

$$x(k+1) = F(k, x(k), a(k)) \quad \forall k = 0, 1, \dots, N-1; \quad x(0) = x_0 \quad (2.10)$$

con función objetivo

$$V(\pi, x) = \sum_{k=0}^{N-1} L(k, x(k), a(k)) + C(x(N)) \quad (2.11)$$

donde  $\pi = \{a(k)\}$ . El espacio de estados es  $X = \mathbb{R}^n$  y el de acciones de control es un conjunto cerrado  $A \subset \mathbb{R}^m$ .

La programación dinámica se basa en el siguiente principio de optimalidad que introdujo Bellman y cuya demostración es evidente.

**Teorema 2.2. El principio de optimalidad** Sea  $a^*(\cdot) = \{a^*(0), \dots, a^*(N-1)\}$  una estrategia óptima para el problema (2.10) - (2.11) y sea  $x^*(\cdot) = \{x^*(0), \dots, x^*(N)\}$  la trayectoria correspondiente, en particular,  $x^*(0) = x_0$ . Entonces para cualquier tiempo  $s \in \{0, \dots, N-1\}$ , la estrategia "truncada"  $a^*(k)$  para  $s \leq k \leq N-1$ , es la estrategia óptima que lleva el sistema (2.10) del punto  $x^*(s)$  al punto  $x^*(N)$ .

Para ver como se usa el Teorema 2.2, consideremos el PCO (2.10) - (2.11) pero sólo del tiempo  $s$  en adelante (con  $0 \leq s \leq N-1$ ), con estado "inicial"  $x(s) = x$ , es decir, sea

$$V(\pi, s, x) = \sum_{k=s}^{N-1} L(k, x(k), a(k)) + C(x(N)). \quad (2.12)$$

y sea  $v(s, x)$  el correspondiente costo mínimo, es decir:

$$v(s, x) = \inf_{\pi} V(\pi, s, x). \quad (2.13)$$

Además como en el tiempo terminal  $N$  no se aplican acciones de control, definimos

$$v(N, x) = C(x). \quad (2.14)$$

Luego, con el Teorema 2.2 interpretamos  $s$  y  $x^*(s) = x$  como el tiempo y el estado iniciales, se sigue de (2.12) y (2.13) que

$$\begin{aligned} v(s, x) &= V(a^*(\cdot), s, x) \\ &= \sum_{k=s}^{N-1} L(k, x^*(k), a^*(k)) + C(x^*(N)) \\ &= L(s, x, a^*(s)) + V(a^*(\cdot), s+1, x^*(s+1)) \\ &= L(s, x, a^*(s)) + v(s+1, x^*(s+1)). \end{aligned}$$

Por lo tanto, como  $x^*(s+1) = F(s, x^*(s), a^*(s)) = F(s, x, a^*(s))$ , obtenemos

$$v(s, x) = L(s, x, a^*(s)) + v(s+1, F(s, x, a^*(s))). \quad (2.15)$$

Pero por la definición (2.13),  $v(s, x)$  es el costo *mínimo* de operar el sistema del tiempo  $s$  al tiempo  $N$ , de modo que

$$v(s, x) \leq L(s, x, a) + v(s+1, F(s, x, a)) \quad \forall a \in A. \quad (2.16)$$

Finalmente, combinando (2.15) y (2.16) vemos que

$$v(s, x) = \min_{a \in A} [L(s, x, a) + v(s+1, F(s, x, a))] \quad \forall s = 0, \dots, N-1 \quad (2.17)$$

y que el mínimo en el lado derecho de (2.17) se alcanza en  $a^*(s)$ , como en (2.15).

La ecuación (2.17) con la condición terminal (2.14) se llama la **ecuación de programación dinámica** (EPD) o **ecuación de Bellman**, y es la base del algoritmo de programación dinámica en el siguiente teorema

**Teorema 2.3. El teorema de programación dinámica** Sean  $J_0, J_1, \dots, J_N$  las funciones sobre  $X$  definidas "hacia atrás" (de  $s = N$  a  $s = 0$ ) como

$$J_N(x) = C(x) \quad (2.18)$$

y para  $s = N-1, N-2, \dots, 0$

$$J_s(x) = \min_a [L(s, x, a) + J_{s+1}(F(s, x, a))] \quad (2.19)$$

Suponga que para cada  $s = 0, 1, \dots, N-1$  existe una función  $a_s^* : X \mapsto A$  que alcanza el mínimo en el lado derecho de (2.19) para todo  $x \in X$ . Entonces la estrategia markoviana  $\pi^* = \{a_0^*, \dots, a_{N-1}^*\}$  es óptima y la función de valor coincide con  $J_0$ , es decir:

$$\inf_{\pi} V(\pi, x) = V(\pi^*, x) = J_0(x) \quad \forall x \in X \quad (2.20)$$

De hecho, para cada  $s = 0, \dots, N$ ,  $J_s$  coincide con la función en (2.13) y (2.14), es decir,

$$v(s, x) = J_s(x) \quad \forall 0 \leq s \leq N, x \in X \quad (2.21)$$

Es importante observar que (2.21) significa que el algoritmo (2.18) - (2.19) da el *costo óptimo* (o costo mínimo) de PCO (2.10) - (2.11) con tiempo y estado inicial  $0 \leq s \leq N-1$  y  $x(s) = x$ , respectivamente.

Consideremos ahora el sistema estocástico en el que (2.10) y (2.11) se sustituyen por

$$x(k+1) = F(k, x(k), a(k), \xi(k)) \quad \forall k = 0, \dots, N-1; x(0) = x \quad (2.22)$$

$$V(\pi, x) = E \left[ \sum_{k=0}^{N-1} L(k, x(k), a(k)) + C(x(N)) \right] \quad (2.23)$$



con  $\pi = \{a(k)\}$  y las perturbaciones  $\xi(0), \dots, \xi(N-1)$  en (2.22) son variables aleatorias independientes e idénticamente distribuidas (i.i.d) con valores en algún espacio  $S$ . Resulta entonces que, con algunos cambios adecuados, prácticamente todo lo que aparece en los párrafos anteriores sigue siendo válido.

Más precisamente, en las expresiones en las que aparece la función  $F$  debemos escribir  $F(s, x, a, \xi(s))$  en lugar de  $F(s, x, a)$ , además, se debe tomar la esperanza en las expresiones donde aparezcan términos estocásticos, o sea, en el lado derecho de (2.12), (2.15) - (2.17) y (2.19).

### 2.4.3 Sistema lineal con objetivo cuadrático (LQ)

El enunciado del problema, en términos matemáticos, es el siguiente:

$$\begin{aligned} \max_{\{a(k)\}_{k=0}^{N-1}} J &= \frac{1}{2} \sum_{k=0}^{N-1} [x(k) + W_k x(k)^t \Lambda_k u_k] + \frac{1}{2} x(N)^t W_N x(N) \\ \text{sujeto a:} & \quad x(k+1) = A_k x(k) + B_k u(k) \quad k = 0, 1, \dots, N-1 \\ \text{con:} & \quad x(0) = 0 \end{aligned} \quad (2.24)$$

en donde, para cada  $k$  dada:

- $x(k)$  es el vector de estado ( $n$ -dimensional).
- $a(k)$  es el vector de control ( $m$ -dimensional).
- $M_k, B_k, \Lambda_k$  son matrices dadas, siendo  $M_k$  de dimensiones  $n \times n$ ,  $B_k : n \times m$ ,  $W_k : n \times n$ , siendo simétrica y semidefinida negativa,  $\Lambda_k : m \times m$  y definida negativa,  $v^t$  es el vector traspuesto de  $v$

A continuación en la Proposición (2.1), se resuelve el problema, por programación dinámica

**Proposición 2.1.** *Para el Problema (2.24) se obtiene la siguiente solución óptima: para  $k = 0, 1, \dots, N-1$ , el control óptimo y la función valor son, respectivamente:*

$$u^*(k) = G_k x(k), \quad J_k^* = \frac{1}{2} x(k)^t K_k x(k)$$

en donde

$$G_k = -\Theta_k^{-1} \Psi_k^t$$

y

$$K_k = \Phi_k - \Psi_k \Theta_k^{-1} \Psi_k^t \quad k = 0, 1, \dots, N-1, \quad K_N = W_N$$

siendo

$$\begin{aligned} \Phi_k &= W_k + M_k^t K_{k+1} M_k \\ \Theta_k &= \Lambda_k + B_k^t K_{k+1} B_k \\ \Psi_k &= M_k^t K_{k+1} B_k \end{aligned}$$

Además, la evolución del sistema controlado, viene dada por:

$$x^*(k+1) = [M_k + B_k G_k]x^*(k), \quad \text{para } k = 0, 1, \dots, N-1 \quad (2.25)$$

**Demostración.**

Utilizamos el método de inducción sobre  $k$ .

Vamos resolviendo el problema dado, hacia atrás en el tiempo.

- Final: Sea  $x(N)$  dado. Entonces, la aportación a la función objetivo, por el hecho de que el sistema finalice en dicho estado, es

$$J_N^*\{x(N)\} = \frac{1}{2}x(N)^t W_N x(N) = \frac{1}{2}x(N)^t K_N x(N) \quad (2.26)$$

en donde se ha definido la matriz  $K_N = W_N$ , que será útil en los próximos pasos a seguir.

- Período  $N$  : sea  $x(N-1)$  dado. La ecuación de Bellman para este período es:

$$J_{n-1}^*\{x(N-1)\} = \max_{a(N-1)} \left\{ \frac{1}{2}x(N-1)^t W_{N-1} x(N-1) + \frac{1}{2}a(N-1)^t \Lambda_{N-1} a(N-1) + J_N^*\{M_{N-1}x(N-1) + B_{N-1}a(N-1)\} \right\}$$

teniendo en cuenta (2.26), se obtiene que la expresión anterior es igual a:

$$\begin{aligned} J_{N-1}^*\{x(n-1)\} &= \max_{a(N-1)} \left\{ \frac{1}{2}x(N-1)^t W_{N-1} x(N-1) + \frac{1}{2}a(N-1)^t \Lambda_{N-1} a(N-1) \right. \\ &\quad \left. + \frac{1}{2}[M_{N-1}x(N-1) + B_{N-1}a(N-1)]^t + K_N[M_{N-1}x(N-1) + B_{N-1}a(N-1)] \right\} \\ &= \max_{a(n-1)} \left\{ \frac{1}{2}x(n-1)^t [W_{N-1} + M_{N-1}^t K_N M_{N-1}] x(N-1) \right. \\ &\quad \left. + \frac{1}{2}a(N-1)^t [\Lambda_{N-1} + B_{N-1}^t K_N B_{N-1}] a(N-1) \right. \\ &\quad \left. + x(N-1)^t M_{N-1}^t K_N B_{N-1} a(N-1) \right\} \\ &= \max_{a(N-1)} \left\{ \frac{1}{2}x(N-1)^t \Phi_{N-1} x(N-1) \right. \\ &\quad \left. + \frac{1}{2}a(N-1)^t \Theta_{n-1} a(N-1) + x(n-1)^t \Psi_{N-1} a(N-1) \right\} \end{aligned}$$

en donde:

$$\begin{aligned} \Phi_{N-1} &= W_{N-1} + M_{N-1}^t K_N M_{N-1} \\ \Theta_{N-1} &= \Lambda_{N-1} + B_{N-1}^t K_N B_{N-1} \\ \Psi_{N-1} &= M_{N-1}^t K_N B_{N-1} \end{aligned}$$

imponiendo la condición necesaria de mínimo (gradiente igual a cero), se obtiene que:

$$a(N-1)\Theta_{N-1}^{\ominus} + x(N-1)^t\Psi_{N-1} = 0$$

por ser  $\Theta_{N-1}$  matriz simétrica, se tiene que:

$$\Theta_{N-1}a(N-1) + \Psi_{N-1}^t x(N-1) = 0$$

de donde se llega a que:

$$a(N-1) = -\Theta_{N-1}^t \Psi_{N-1}^t x(N-1)$$

Obsérvese que la matriz  $\Theta_{N-1}$  posee inversa, ya que por hipótesis,  $\Lambda_{n-1}$  es definida negativa y  $K_N = W_N$  es semidefinida negativa, por lo que  $\Theta_{N-1}$  es definida negativa y por consiguiente no singular y posee matriz inversa. La condición es también suficiente de optimalidad global, por ser cóncava la función objetivo.

Podemos poner:

$$a^*(N-1) = G_{N-1}x(N-1)$$

en donde:

$$G_{N-1} = -\Theta_{N-1}^{-1} \Psi_{N-1}^t = -[\Lambda_{N-1} + B_{N-1}^t K_N B_{N-1}]^{-1} N_{N-1}^t K_N M_{N-1}$$

Entonces

$$\begin{aligned} J_{N-1}^* &= \frac{1}{2}x(N-1)^t \Phi_{N-1} x(N-1) \\ &+ \frac{1}{2}x(N-1)^t G_{N-1}^t \Theta_{N-1} G_{N-1} x(N-1) \\ &+ x(n-1)^t \Psi_{N-1} G_{N-1} x(N-1) \end{aligned}$$

Se puede poner:

$$J_{N-1}^* = \frac{1}{2}x(N-1)^t K_{N-1} x(N-1)$$

en donde

$$\begin{aligned} K_{N-1} &= \Phi_{N-1} + G_{N-1}^t \Theta_{N-1} G_{N-1} + 2\Psi_{N-1} \\ &= \Phi_{N-1} - \Psi_{N-1} \Theta_{N-1}^{-1} \Psi_{N-1}^t \end{aligned}$$

- Hipótesis de inducción: período  $k+1$ . Sea  $x(x)$  dado. Supongamos que la proposición es cierta para  $k$  (según hipótesis de inducción). Ahora hay que demostrar que también se cumple para  $k-1$  (que corresponde al período  $k$ ).
- Período  $k$ : sea  $x(k-1)$  dado. La ecuación de Bellman para este período es:

$$\begin{aligned} J_{k-1}^* \{x(k-1)\} &= \max_{a(k-1)} \left\{ \frac{1}{2}x(k-1)^t W_{k-1} x(k-1) \right. \\ &= \left. + \frac{1}{2}a(k-1)^t \Lambda_{k-1} a(k-1) + J_k^* \{M_{k-1}x(k-1) + B_{k-1}a(k-1)\} \right\} \end{aligned}$$

Teniendo en cuenta la expresión para  $J_k^*$  de la hipótesis de inducción, se obtiene que la expresión anterior es igual a:

$$\begin{aligned}
J_{k-1}^* &= \max_{a(k-1)} \left\{ \frac{1}{2}x(k-1)^t W_{k-1}x(k-1) + \frac{1}{2}a(k-1)^t \Lambda_{k-1}a(k-1) \right. \\
&\quad \left. + \frac{1}{2}[M_{k-1}x(k-1) + B_{k-1}a(k-1)]^t K_k \right. \\
&\quad \left. [M_{k-1}x(k-1) + B_{k-1}a(k-1)] \right\} \\
&= \max_{a(k-1)} \left\{ \frac{1}{2}x(k-1)^t [W_{k-1} + M_{k-1}^t K_k M_{k-1}]x(k-1) \right. \\
&\quad \left. + \frac{1}{2}a(k-1)^t [\Lambda_{k-1} + B_{k-1}^t K_k B_{k-1}]a(k-1) \right. \\
&\quad \left. + x(k-1)^t M_{k-1}^t K_k B_{k-1}a(k-1) \right\} \\
&= \max_{a(k-1)} \left\{ \frac{1}{2}x(k-1)^t \Phi_{k-1}x(k-1) + \frac{1}{2}u(k-1)^t + \Theta_{k-1}u(k-1) \right. \\
&\quad \left. + x(k-1)^t \Psi_{k-1}a(k-1) \right\}
\end{aligned}$$

en donde:

$$\begin{aligned}
\Phi_{k-1} &= W_{k-1} + M_{k-1}^t K_k M_{k-1} \\
\Theta_{k-1} &= \Lambda_{k-1} + B_{k-1}^t K_k B_{k-1} \\
\Psi_{k-1} &= M_{k-1}^t K_k B_{k-1}
\end{aligned}$$

Imponiendo la condición necesaria de mínimo (gradiente igual a cero), se obtiene que:

$$a(k-1)^t \Theta_{k-1} + x(k-1)^t \Psi_{k-1} = 0$$

por ser  $\Theta_{k-1}$  matriz simétrica, se tiene que:

$$\Theta_{k-1}a(k-1) + \Psi_{k-1}^t x(k-1) = 0$$

de donde se llega a:

$$a(k-1) = -\Theta_{k-1}^{-1} \Psi_{k-1}^t x(k-1)$$

Obsérvese que la matriz  $\Theta_{k-1}$  posee inversa, ya que por hipótesis,  $\Lambda_{k-1}$  es definida negativa y  $K_{k+1}$  es semidefinida negativa, por lo que  $\Theta_{k-1}$  es definida negativa y por consiguiente no singular y posee matriz inversa. La condición es también suficiente de optimalidad global, por ser cóncava la función objetivo.

Podemos poner:

$$a^*(k-1) = G_{k-1}x(k-1)$$

$$\begin{aligned}
G_{k-1} &= -\Theta_{k-1}^{-1} \Psi_{k-1}^t \\
&= -[\Lambda_{k-1} + B_{k-1}^t K_k B_{k-1}]^{-1} B_{k-1}^t K_k M_{k-1}
\end{aligned}$$

Entonces

$$\begin{aligned} J_{k-1}^*\{x(k-1)\} &= \frac{1}{2}x(k-1)^t\Phi_{k-1}x(k-1) \\ &\quad + \frac{1}{2}x(k-1)^tG_{k-1}^tG_{k-1}^t\Theta_{k-1}x(k-1) \\ &\quad + x(k-1)^t\Psi_{k-1}G_{k-1}x(k-1) \end{aligned}$$

Se puede poner:

$$J_{k-1}^*\{x(k-1)\} = \frac{1}{2}x(k-1)^tK_{k-1}x(k-1)$$

en donde

$$\begin{aligned} K_{k-1} &= \Phi_{k-1} + G_{k-1}^t\Theta_{k-1}G_{k-1} + 2\Phi_{k-1}G_{k-1} \\ &= \Phi_{k-1} - \Psi_{k-1}\Theta_{k-1}^{-1}\Psi_{k-1}^t \end{aligned}$$

El sistema encontrado es lineal según expresión (2.25). Ello se comprueba inmediatamente sustituyendo en la ecuación de estado el control óptimo obtenido.

□

## 2.5 Problema de control estocástico LQ

En esta sección se estudia una versión del problema estocástico (2.1) en el caso de que las perturbaciones aleatorias sean de tipo aditivo.

### *Enunciado del problema.*

Consideremos el siguiente sistema dinámico formulado en tiempo discreto:

$$\begin{aligned} x(k+1) &= f(x(k) + a(k), v(k), k) \quad \text{para } k = 0, 1, \dots, N-1 \\ \text{con:} \quad &x(0) = x_0 \end{aligned}$$

en donde para cada  $k$  :

$x(k)$  es el vector de variables de estado, perteneciente al espacio  $S_k$ .

$a(k)$  es el vector de variables de control, pertenecientes al espacio  $C_k$ .

$v(k)$  es el vector de perturbaciones aleatorias, pertenecientes al espacio  $D_k$ .

La secuencia en la que van a ir tomando valores las diferentes variables, a través del tiempo, es la que aparece en la Figura (2.3)

El control  $a(k)$  está restringido a tomar valores pertenecientes al conjunto  $\Omega_k[x(k)]$ , que depende del estado del sistema en ese período  $k$ , es decir:

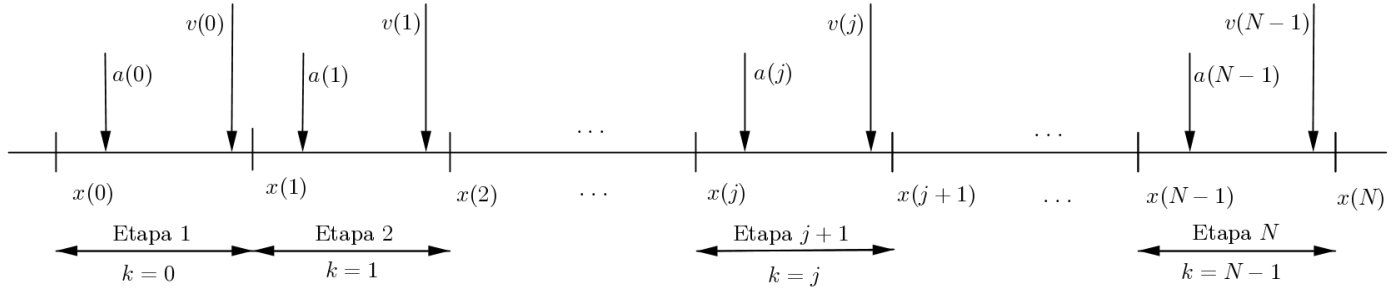


Figura 2.3: Etapas y variables en el problema de control estocástico en tiempo discreto

$$a(k) \in \Omega_k[x(k)] \quad \text{para cada } x(k) \in S_k \quad \forall k = 0, 1, \dots, N-1$$

La perturbación aleatoria  $v(k)$  viene caracterizada por una distribución de probabilidad

$$P_k(\bullet | x(k), a(k))$$

que puede depender explícitamente de  $x(k), a(k)$ , pero es independiente de  $v(k-1), v(k-2), \dots, v(0)$ .

Consideremos las leyes de control (también llamadas políticas), consiste en una secuencia de funciones:

$$\pi = \{\mu_0, \mu_1, \dots, \mu_{N-1}\}$$

en donde cada función  $a_k$  transforma el estado  $x(k)$  en el control  $a(k) = \mu_k[x(k)]$ , de manera que se verifica que

$$\mu_k[x(k)] \in \Omega_k[x(k)], \quad \forall x(k) \in S_k$$

Las leyes de control que cumplen esta condición se llamarán admisibles.

El problema consiste, exactamente, en lo siguiente:

*dato un estado inicial  $x_0$ , se trata de encontrar una ley de control admisible*

$$\pi = \{\mu_0, \mu_1, \dots, \mu_{N-1}\}$$

*que maximice el funcional siguiente:*

$$\begin{aligned} \max \quad & J_\pi(x_0) = E \left\{ \sum_{k=0}^{N-1} F[x(k), \mu_k[x(k)], v(k), k] + S[x(N)] \right\} \\ \text{sujeto a:} \quad & x(k+1) = f[x(k), \mu_k[x(k)], v(k), k] \quad k = 0, 1, \dots, N-1 \\ \text{con:} \quad & x(0) = x_0 \\ & \mu_k[x(k)] \in \Omega_k[x(k)] \quad \forall x(k) \in S_k \end{aligned}$$

en donde  $E\{\bullet\}$  significa esperanza matemática. Por otra parte, las funciones  $F, s$  y  $f$  están dadas.

Se trata de encontrar una ley de control óptimo  $\pi^*$ , para la cual se verifique:

$$J_{\pi^*}(x_0) = \max_{\pi \in \Pi} J_{\pi}(x_0)$$

en donde  $\Pi$  es el conjunto de leyes de control admisibles. Entonces la función  $J^*$ , es definida de la siguiente forma:

$$J^*(x_0) = J_{\pi^*}(x_0)$$

Es la función que asigna a cada estado inicial  $x_0$  el valor objetivo óptimo  $J^*(x_0)$  y se llama **función valor óptimo**.

**Observación 2.2.** Si el problema anterior es de minimizar, en lugar de maximizar, se puede expresar de la forma enunciada anteriormente, sin más que cambiar el signo del funcional objetivo.

**Ejemplo 2.1.** Problema de control estocástico de un sistema lineal con función objetivo cuadrática. Se trata de calcular:

$$\{a(0), a(1), \dots, a(N-1)\}$$

para los que se alcanza el

$$\begin{aligned} \min \quad J &= E \left\{ \sum_{k=0}^{N-1} \left[ \frac{1}{2} x(k)^t W_k x(k) + \frac{1}{2} a(k)^t \Lambda_k a(k) \right] + \frac{1}{2} x(N)^t W_N x(N) \right\} \\ \text{sujeto a:} \quad & x(k+1) = \Lambda_k x(k) + B_k a(k) + v(k) \quad k = 0, 1, \dots, N-1 \\ \text{con:} \quad & x(0) = x_0 \end{aligned}$$

En donde, para cada  $k$  se tiene:

$x(k)$  es el vector de variables de estado  $n$ -dimensional.

$a(k)$  es el vector de variables de control  $m$ -dimensional.

Por otra parte  $\{v(k)\}_{k=0,1,\dots,N}$  son vectores aleatorios normales, tales que:

$$\begin{aligned} E\{v(k)\} &= 0, \quad E\{v(k)v(k)^t\} = Q_k \quad \forall k \\ E\{v(k)v(l)^t\} &= 0, \quad k \neq l \end{aligned}$$

por lo que se trata de vectores aleatorios normales, serialmente incorrelados (por tanto, independientes) y de media cero.

Esto nos lleva al enunciado del problema en términos matemáticos:

$$\begin{aligned} \max_{\{u(k)\}_{k=0}^{N-1}} J &= E \left\{ \frac{1}{2} \sum_{k=0}^{N-1} [x(k)^t W_k x(k) + u(k)^t \Lambda_k x(N)^t u(k)] + \frac{1}{2} x(N)^t W_N x(N) \right\} \\ \text{sujeto a: } x(k+1) &= M_k x(k) + B_k u(k) + v(k), \quad \text{para } k = 0, 1, \dots, N-1 \\ \text{con: } x(0) &= x_0 \end{aligned} \quad (2.27)$$

en donde, para dada  $k$  :

$x(k)$  es el vector de estado ( $n$ -dimensional).

$a(k)$  es el vector de control ( $m$ -dimensional)

$M_k, B_k, \Lambda_k$  son matrices dadas, siendo  $M_k$  de dimensiones  $n \times n$ ,  $B_k : n \times m$ ,  $W_k : n \times n$  siendo simétrica y semidefinida negativa.  $\Lambda_k : m \times m$  y definida negativa,  $z^t$  es el vector traspuesto de  $z$

Por otra parte,  $\{v(k)\}_{k=0}^{N-1}$  son vectores aleatorios independientes, tales que:

$$E\{v(k)\} = 0, \quad E\{v(k)v(k^t)\} = Q_k, \quad E\{v(k)v(l^t)\} = 0, \quad \text{si } k \neq l$$

En la siguiente proposición se resuelve el problema, por programación dinámica

**Proposición 2.2.** *Para el problema (2.27) se obtiene la siguiente solución óptima: para  $k = 0, 1, \dots, N-1$ , el control óptimo y la función valor son, respectivamente*

$$a^*(k) = G_k x(k) \quad J^*\{x(k)\} = \frac{1}{2} x(k)^t \mathcal{K}_k x(k) + \frac{1}{2} \sum_{i=k}^{N-1} E\{v(i)^t \mathcal{K}_{i+1} v(i)\}$$

en donde

$$G_k = -\Theta_k^{-1} \Psi_k^t$$

y

$$\mathcal{K}_k = \Phi_k - \Psi_k \Theta_k^{-1} \Psi_k^t, \quad \text{para } k = 0, 1, \dots, N-1, \quad \text{con } \mathcal{K}_N = W_N$$

siendo

$$\begin{aligned} \Phi_k &= W_k + M_k^r \mathcal{K}_{k+1} M_k \\ \Theta_k &= \Lambda_k + B_k^t \mathcal{K}_{k+1} B_k \\ \Psi_k &= M_k^t \mathcal{K}_{k+1} B_k \end{aligned}$$

Además, la evolución del sistema controlado viene dada por:

$$x^*(k+1) = [M_k + B_k G_k] x^*(k) + v(k), \quad \text{para } k = 0, 1, \dots, N-1 \quad (2.28)$$

### **Demostración**

Utilizamos el método de inducción sobre  $k$ . Vamos resolviendo el problema dado, hacia atrás en el tiempo.



- Final: sea  $x(N)$  dado.  
entonces, la aportación a la función objetivo, por el hecho de que el sistema finalice en dicho estado es:

$$\begin{aligned} J_N^* &= E \left\{ \frac{1}{2} x(N)^t W_N x(N) \right\} = \frac{1}{2} x(N)^t W_N x(N) \\ &= \frac{1}{2} x(N)^t K_N x(N) \end{aligned} \quad (2.29)$$

en donde se ha definido la matriz  $K_N = W_N$ , que será útil en próximos pasos a seguir.

- Período  $N$  : sea  $x(N-1)$  dado.  
La ecuación de Bellman para este período es:

$$\begin{aligned} J_{N-1}^* \{x(N-1)\} &= \max_{u(N-1)} \left\{ \frac{1}{2} x(N-1)^t W_{N-1} x(N-1) \right. \\ &\quad \left. + \frac{1}{2} u(N-1)^t \Lambda_{N-1} u(N-1) + J_N^* \{x(N)\} \right\} \end{aligned}$$

en donde:

$$x(N) = M_{N-1} x(N-1) + B_{N-1} u(N-1) + v(N-1)$$

Teniendo en cuenta (2.29) se obtiene que la expresión anterior es igual a :

$$\begin{aligned} J_{N-1}^* \{x(N-1)\} &= \max_{u(N-1)} \left\{ \frac{1}{2} x(N-1)^t W_{N-1} x(N-1) \right. \\ &\quad \left. + \frac{1}{2} u(N-1)^t \Lambda_{N-1} u(N-1) + \frac{1}{2} x(N)^t K_N x(N) \right\} \end{aligned}$$

siendo

$$x(N) = M_{N-1} x(N-1) + B_{N-1} u(N-1) + v(N-1)$$

Por lo tanto:

$$\begin{aligned}
J_{N-1}^*\{x(N-1)\} &= \max_{u(N-1)} \left\{ \frac{1}{2}x(N-1)^t[W_{N-1} + A_{N-1}^t\mathcal{K}_NA_{N-1}]x(N-1) \right. \\
&\quad + \frac{1}{2}a(N-1)^t[\Lambda_{N-1} + B_{N-1}^t\mathcal{K}_Nb_{N-1}]a(N-1) \\
&\quad + x(N-1)^tA_{N-1}^t\mathcal{K}_NB_{N-1}a(N-1) \\
&\quad + \frac{1}{2}x(N-1)^tA_{N-1}^t\mathcal{K}_Nv(N-1) \\
&\quad + \frac{1}{2}a(N-1)^tB_{N-1}^t\mathcal{K}_Nv(N-1) \\
&\quad + \frac{1}{2}v(N-1)^t\mathcal{K}_NM_{N-1}x(N-1) \\
&\quad \left. + \frac{1}{2}v(N-1)^t\mathcal{K}_NB_{N-1}a(n-1) + \frac{1}{2}v(N-1)^t\mathcal{K}_Nv(N-1) \right\} \\
&= \max_{a(N-1)} \left\{ \frac{1}{2}x(N-1)\Phi_{N-1}x(N-1) + \frac{1}{2}a(N-1)^t\Theta_{N-1}a(N-1) \right. \\
&\quad + x(N-1)^t\Psi_{N-1}a(N-1) \\
&\quad \left. \frac{1}{2}E\{v(N-1)^t\mathcal{K}_Nv(n-1)\} \right\}
\end{aligned}$$

en donde

$$\begin{aligned}
\Phi_{N-1} &= W_{N-1} + A_{N-1}^t\mathcal{K}_NM_{N-1} \\
\Theta_{N-1} &= \Lambda_{N-1} + B_{N-1}^t\mathcal{K}_NB_{N-1} \\
\Psi_{N-1} &= M_{N-1}^t\mathcal{K}_NB_{N-1}
\end{aligned}$$

Imponiendo la condición necesaria de mínimo (gradiente igual a cero), se obtiene que:

$$a(N-1)\Theta_{N-1}x(N-1)^t\Psi_{N-1} = 0$$

por ser  $\Theta_{N-1}$  matriz simétrica, se verifica que:

$$\Theta_{N-1}a(N-1) + \Psi_{N-1}^tx(N-1) = 0$$

de donde se obtiene que:

$$a(N-1) = -\Theta_{N-1}^{-1}\Psi_{N-1}^tx(N-1)$$

Obsérvese que la matriz  $\Theta_{N-1}$  posee inversa, ya que por hipótesis,  $\Lambda_{N-1}$  es definida negativa y  $\mathcal{K}_N = W_N$  es semidefinida negativa, por lo que  $\Theta_{N-1}$  es definida negativa y por consiguiente no singular y posee matriz inversa. La condición es también suficiente

de optimalidad global, por ser cóncava la función objetivo.  
Podemos poner:

$$a^*(N-1) = G_{N-1}x(N-1)$$

en donde:

$$G_{N-1} = -\Theta_{N-1}^{-1}\Psi_{N-1}^t = -[\Lambda_{N-1} + B_{N-1}^t\mathcal{K}_N B_{N-1}]^{-1}B_{N-1}^t\mathcal{K}_N M_{N-1}$$

Entonces,

$$\begin{aligned} J_{N-1}^*\{x(N-1)\} &= \frac{1}{2}x(N-1)^t\Phi_{N-1}x(N-1) \\ &\quad + \frac{1}{2}x(N-1)^tG_{N-1}^t\Theta_{N-1}G_{N-1}x(N-1) \\ &\quad + x(N-1)^t\Psi_{N-1}G_{N-1}x(N-1) \\ &\quad + \frac{1}{2}E\{v(N-1)\mathcal{K}_N v(N-1)\} \end{aligned}$$

Se puede poner:

$$J_{N-1}^*\{x(N-1)\} = \frac{1}{2}x(N-1)^t\mathcal{K}_{N-1}x(N-1) + \frac{1}{2}E\{v(N-1)^t\mathcal{K}_N v(n-1)\}$$

en donde:

$$\begin{aligned} \mathcal{K}_{N-1} &= \Phi_{N-1} + G_{N-1}^t\Theta_{N-1}G_{N-1} + 2\Psi_{N-1}G_{N-1} \\ &\quad \Phi_{N-1} - \Psi_{N-1}\Theta_{N-1}^{-1}\Psi_{N-1}^t \end{aligned}$$

- Hipótesis de inducción: período  $k+1$ . Sea  $x(k)$  dado.  
Supongamos que la proposición es cierta para  $k$  (según hipótesis de inducción). Ahora hay que demostrar que también se cumple para  $k-1$  (que corresponde al período  $k$ ).

Período  $k$ : sea  $x(k-1)$  dado.

La ecuación de Bellman para este período es:

$$\begin{aligned} J_{k-1}^*\{x(k-1)\} &= \max_{a(k-1)} \left\{ \frac{1}{2}x(k-1)^t W_{k-1} x(k-1) + \frac{1}{2}a(k-1)^t \Lambda_{k-1} a(k-1) \right. \\ &\quad \left. + J_{k-1}^*\{x(k)\} \right\} \end{aligned}$$

en donde

$$x(k) = M_{k-1}x(k-1) + B_{k-1}a(k-1) + v(k-1)$$

Teniendo en cuenta la expresión para  $J_k^*$  de la hipótesis de inducción, se obtiene que la expresión anterior es igual a:

$$J_{k-1}^*\{x(k-1)\} = \max_{a(k-1)} E \left\{ \frac{1}{2} x(k-1)^t W_{k-1} x(k-1) + \frac{1}{2} a(k-1)^t \Lambda_{k-1} a(k-1) \right. \\ \left. + \frac{1}{2} x(k)^t \mathcal{K}_k x(k) + \frac{1}{2} \sum_{i=k}^{N-1} [v(i)^t \mathcal{K}_{i+1} v(i)] \right\}$$

siendo

$$x(k) = M_{k-1} x(k-1) + B_{k-1} a(k-1) + v(k-1)$$

Por tanto,

$$J_{k-1}^*\{x(k-1)\} = \max_{a(k-1)} E \left\{ \frac{1}{2} x(k-1)^t [W_{k-1} + \Lambda_{k-1}^t \mathcal{K}_k \Lambda_{k-1}] x(k-1) \right. \\ \left. + \frac{1}{2} a(k-1)^t \left[ \Lambda_{k-1} + B_{k-1}^t \mathcal{K}_k B_{k-1} \right] a(k-1) \right. \\ \left. + x(k-1)^t A_{k-1}^t \mathcal{K}_k B_{k-1} a(k-1) + \frac{1}{2} x(k-1)^t \Lambda_{k-1} \mathcal{K}_k v(k-1) \right. \\ \left. + \frac{1}{2} a(k-1)^t B_{k-1}^t \mathcal{K}_k v(k-1) + \frac{1}{2} v(k-1)^t \mathcal{K}_k \Lambda_{k-1} x(k-1) \right. \\ \left. + \frac{1}{2} v(k-1)^t \mathcal{K}_k B_{k-1} a(k-1) + \frac{1}{2} v(k-1)^t \mathcal{K}_k v(k-1) \right. \\ \left. + \frac{1}{2} \sum_{i=k}^{N-1} [v(i)^t \mathcal{K}_{i+1} v(i)] \right\} \\ = \max_{a(k-1)} \left\{ \frac{1}{2} x(k-1)^t \Phi_{k-1} x(k-1) + \frac{1}{2} a(k-1)^t \Theta_{k-1} a(k-1) \right. \\ \left. + x(k-1)^t \Psi_{k-1} a(k-1) + \frac{1}{2} \sum_{i=k-1}^{N-1} E\{v(i)^t \mathcal{K}_{k+i} v(i)\} \right\}$$

en donde:

$$\begin{aligned} \Phi_{k-1} &= W_{k-1} + A_{k-1}^t \mathcal{K}_k A_{k-1} \\ \Theta_{k-1} &= \Lambda_{k-1} + B_{k-1}^t \mathcal{K}_k B_{k-1} \\ \Psi_{k-1} &= A_{k-1}^t \mathcal{K}_k B_{k-1} \end{aligned}$$

Imponiendo la condición necesaria de mínimo (gradiente igual a cero), se obtiene que:

$$a(k-1)^t \Theta_{k-1} + x(k-1)^t \Psi_{k-1} = 0$$

Por ser  $\Theta_{k-1}$  matriz simétrica, se verifica que:

$$\Theta_{k-1} a(k-1) + \Psi_{k-1}^t x(k-1) = 0$$

de donde se obtiene que:

$$a(k-1) = -\Theta_{k-1}\Psi_{k-1}^t x(k-1)$$

Obsérvese que la matriz  $\Theta_{k-1}$  posee inversa, ya que por hipótesis,  $\Lambda_{k-1}$  es definida negativa y  $\mathcal{K}_{k-1}$  es semidefinida negativa, por lo que  $\Theta_{k-1}$  es definida negativa y por consiguiente no singular y posee matriz inversa. La condición es también suficiente de optimalidad global, por ser cóncava la función objetivo.

Podemos poner:

$$a^*(k-1) = G_{k-1}x(k-1)$$

en donde:

$$G_{k-1} = -\Theta_{k-1}^{-1}\Psi_{k-1}^t = [\Lambda_{k-1} + B_{k-1}^t B_{k-1}]^{-1} B_{k-1}^t \mathcal{K}_k A_{k-1}$$

Entonces,

$$\begin{aligned} J_{k-1}^*\{x(K-1)\} &= \frac{1}{2}x(k-1)^t \Phi_{k-1} \\ &+ \frac{1}{2}x(k-1)^t G_{k-1}^t \Theta_{k-1} G_{k-1} x(k-1) \\ &+ x(k-1)^t \Psi_{k-1} G_{k-1} x(k-1) + \frac{1}{2} \sum_{i=k-1}^{N-1} E\{v(i)\mathcal{K}_{i+1}v(i)\} \end{aligned}$$

Se puede poner:

$$J_{N-1}^*\{x(k-1)\} = \frac{1}{2}x(k-1)^t \mathcal{K}_{k-1} x(k-1) + \frac{1}{2} \sum_{i=k-1}^{N-1} E\{v(i)^t \mathcal{K}_{i+1} v(i)\}$$

en donde:

$$\begin{aligned} \mathcal{K}_{k-1} &= \Phi_{k-1} + G_{k-1}^t \Theta_{k-1} G_{k-1} + 2\Psi_{k-1} G_{k-1} \\ &= \Phi_{k-1} - \Psi_{k-1} \Theta_{k-1}^{-1} \Psi_{k-1}^t \end{aligned}$$

El sistema controlado es lineal. Ello se comprueba inmediatamente sin más que sustituir en la ecuación de estado el control óptimo obtenido.

□

## 2.6 Principio de equivalencia cierta

El trabajo pionero de Bellman con énfasis en los modelos que utilizan variables de estado y el método de programación dinámica ha generado una teoría de la toma de decisiones dinámicas bajo incertidumbre y en particular, en el control estocástico y adaptativo. Una de las ideas en esta área con una gran cantidad de atractivo práctico y conceptual es la de **equivalencia cierta**.

La idea básica aquí es definir una política de decisión poniendo la incertidumbre igual a su valor esperado dadas las observaciones. Si bien a primera vista el contenido intuitivo básico detrás de esta idea parece bastante claro, no es fácil dar formulaciones matemáticas precisas de este principio. [29].

### Notación.

Si  $T \subset \mathbb{R}$ ,  $t \in T$  y  $y : T \mapsto Y$  entonces  $y_t^-$  denota el pasado de  $y$  definido por  $y_t^- : T \cap (-\infty, t) \mapsto Y$  con  $y_t^-(\tau) := y(\tau)$  para  $\tau < t$ . Como de costumbre  $U^T$  denota todos los mapeos de  $T$  a  $U$ , es decir,  $U^T := \{f \mid f : T \mapsto U\}$  como aparece en un contexto diferente, esto no se confundirá con transposición [29].

**Definición 2.6.** Se dice que una solución es **anticipativa**, si los valores de las variables son únicos, implantables e independientes del escenario que acontezca [28].

**Definición 2.7.** Sea  $U \subset U^T$  y  $F : U \mapsto Y^T$  entonces, se dice que  $F$  es [29]

- No anticipativa si  $u, v \in U$  y  $u(t') = v(t')$  para  $t' \leq t$  lo que implica  $Fu(t') = Fv(t')$  para  $t' \leq t$
- Estrictamente no anticipativa si  $u(t') = v(t')$   $t' < t$  implica  $Fu(t') = Fv(t')$  para  $t' \leq t$

La idea básica de esta estrategia consiste en evitar adelantar innecesariamente decisiones a etapas anteriores, si estas se precisan efectuar con posterioridad. De esta forma, la solución que se obtiene considera todos los escenarios posibles en situaciones futuras, pero no se subordina a ninguno de ellos y además proporciona las decisiones a tomar en el momento en el que éstas se necesitan [28].

**Definición 2.8.** Un sistema dinámico de incertidumbre está definido por un mapeo  $\mathbf{G} : U \times \Omega \mapsto Y$  llamado **planta** tal que  $\forall \omega \in \Omega$   $\mathbf{G}(\bullet, \omega)$  es estrictamente no anticipativa.

Denotaremos un sistema dinámico de incertidumbre por su planta  $\mathbf{G}$ .

Veremos una política de control como una receta para decidir el control en base a las observaciones, que respeta las restricciones del flujo de información impuestas por la siguiente definición.

**Definición 2.9.** Una **ley de control** es un mapeo  $\mathbf{F} : Y \mapsto U$  observe que es

*i No anticipativa.*

ii Es tal que la ecuación implícita en  $u$

$$u = \mathbf{FG}(u, \omega) \quad \forall \omega \in \Omega$$

tiene solución única en  $U$  (esta solución depende claramente de  $\mathbf{F}$  y  $\omega$  y se denotarán por  $\mathbf{F}_u(\omega, \mathbf{F})$ ).

iii De forma que el flujo de información de  $\mathbf{G}$  es no anticipativa en el sistema de circuito cerrado, con esto queremos decir que si  $\exists t \in T$  tal que  $\forall u \in U$ ,  $\mathbf{G}(u, w_1)(t') = \mathbf{G}(u, w_2)(t')$  para  $t' \leq t$  entonces  $\mathbf{F}_u(w_1, \mathbf{F})(t') = \mathbf{F}_u(w_2, \mathbf{F})(t')$  para  $t' \leq t$ .

En el caso discreto  $\exists T$  tal que  $T = [t_0, t_0 + 1, \dots, t_1]$ ,  $-\infty < t_0 \leq t_1 \leq \infty$  con  $\mathbf{U} = U^t$  y  $\mathbf{Y} = Y^t$

La idea básica es la siguiente: "Suponemos que nos enfrentamos a la decisión de elegir un control  $a \in A$  para minimizar  $J(a, x)$  [29]

En el caso de tener información perfecta, es decir, si  $a(k) = \mu_k(x(k))$  es conocido, esto lleva a escoger  $a^*(x)$  tal que

$$J(a^*(x), x) \leq J(a, x) \quad \forall a \in A$$

Sin embargo, si sólo los conocemos a través de las observaciones y no podemos resolver el problema de encontrar la ley de control óptimo correspondiente no tiene ningún mérito utilizar la ley de control

$$\hat{a} := E\{a^*(x) | y(x) = y\}$$

donde hemos supuesto que esta expectativa condicional está bien definida" a esta idea se le llama *equivalencia de certidumbre*, ahora se formalizará el marco dinámico.

Sea  $T = \{t_0, t_0 + 1, \dots, t_1\}$  con  $-\infty < t_0 \leq t_1 \leq \infty$  y se asume que la planta  $\mathbf{G}$  y la funcional de costo  $J$  están dadas y que el control de entrada es un subconjunto de un espacio vectorial finito. Por simplicidad se asume que  $\mathbf{U} = U^T$  y  $\mathbf{Y} = Y^T$ .

Sea  $v \in U$ ,  $k \in T$  y considere el siguiente problema de control determinístico

$$\min_{u \in U_{v,k}} J(\mathbf{u}, \omega) \tag{2.30}$$

donde

$$U_{v,k} = \{u \in U | u(k') = v(k') \text{ para } k' < k\}$$

Sea  $u^*(v, k, \omega)$  un control de minimización. Considerando que su valor en el tiempo  $k$  define un mapeo  $\mathbf{R} : \mathbf{U} \times \Omega$

$$\mathbf{R}(v, \omega)(k) = u^*(v, k, \omega)(k)$$

Este  $R(u, o)(k)$  es el control que usaríamos en el tiempo  $k$  si nos encontramos con que la información es perfecta ( $\omega$  es conocido) pero las decisiones previas (el control  $v$  antes del tiempo  $k$ ) ya se han decidido.

Las observaciones hasta el tiempo  $k$  que habrían ocurrido bajo los supuestos anteriores están dadas por  $\mathbf{G}(v, \omega)$  y  $\mathbf{G}(v, \omega)(k')$  para  $k' \leq k$  estando disponible en el momento  $k$ .

La idea de equivalencia de certeza ya explicada anteriormente, sugeriría usar en el tiempo  $k$  el valor de control

$$E\{R(v, \cdot)(k) | G(v, \cdot)(\tau) = y(\tau) \quad \tau \leq k\}$$

(Asumiendo que esta esperanza está bien definida.) Luego se define un mapeo  $\mathbf{S} : U \times Y \mapsto U$  definido por

$$\mathbf{S}(v, y)(k) = E\{R(v, \cdot)(k) | G(v, \cdot)(\tau) = y(\tau) \quad \tau \leq k\}$$

Este mapeo es no anticipativo en  $y$  y es estrictamente no anticipativo en  $v$ . Consideremos ahora ecuaciones simultáneas

$$\mathbf{S}(v, y) = u; \quad u = v$$

que nos dice que la equivalencia de certeza se usa para todo  $k$ , estas ecuaciones definen un mapeo

$$\mathbf{F}_{ce} : Y \mapsto U$$

también define una ley de control según lo explicado. Llamaremos  $\mathbf{F}_{ce}$  *ley de control de equivalencia de certeza*. El uso de esta ley de control es conocido como *principio equivalente de certeza*, y diremos que la propiedad de equivalencia de certidumbre es válida para un problema de control estocástico si la ley de control óptimo  $\mathbf{F}^* = \mathbf{F}_{ce}$  [29].

En el siguiente ejemplo se resuelve un problema de control estocástico lineal–cuadrático, con perturbaciones aleatorias de tipo multiplicativo, para el que no se verifica el principio de equivalencia cierta [4].

**Ejemplo 2.2.** *Se considera el siguiente problema de control estocástico:*

$$\begin{aligned} \min_{a(0), a(1)} \quad & J = E\{[x(0) - 2]^2 + [x(1) - 3]^2 + [x(2) - 3]^2 + a(0)^2 + [a(1) - 1]^2\} \\ \text{sujeto a} \quad & x(1) = \alpha x(0) + 2a(0) + 1 \\ & x(2) = x(1) + ba(1) - 2 \\ \text{con} \quad & x(0) = 1 \end{aligned}$$

en donde  $x$  es la variable de estado,  $a$  es la variable de control,  $\alpha$  es una variable aleatoria normal de media 1 y varianza 1 que se realiza al final del primer período,  $b$  es una variable aleatoria normal de media 3 y varianza 4 que se realiza al final del período 2, siendo  $\alpha$  y  $b$  independientes.

a) Resolver el problema.

b) Comprobar que este problema no cumple el principio de equivalencia cierta



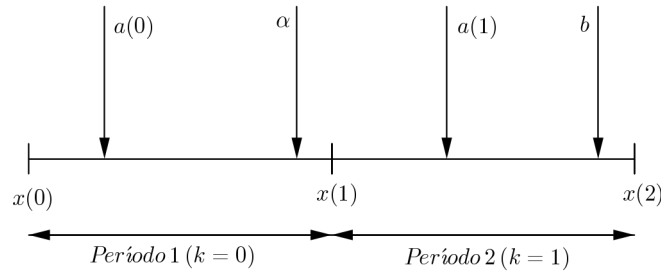


Figura 2.4: Períodos y variables en el ejemplo 2.2

**Solución.** En la figura (2.4) aparece el esquema habitual correspondiente a este problema.

a) Resolvemos por programación dinámica.

- Final: Sea  $x(2)$  dado. Entonces:

$$J_2^*\{x(2)\} = [x(2) - 3]^2$$

.

- Segundo período: sea  $x(1)$  dado.  
La ecuación de Bellman correspondiente al período es:

$$J_1^* = \min_{a(1)} E\{[x(1) - 3]^2 + [a(1) - 1]^2 + J_2^*\{x(2)\}\}$$

en donde

$$x(2) = x(1) + ba(1) - 2$$

Por tanto,

$$J_1^*\{x(1)\} = \min_{a(1)} E\{[x(1) - 3]^2 + [a(1) - 1]^2 + [x(1) + ba(1) - 5]^2\}$$

Efectuando operaciones y teniendo en cuenta que

$$E\{b\} = 3, \quad E\{b^2\} = \text{var}(b) + [E\{b\}]^2 = 13$$

se llega a que

$$J_1^* = \min_{a(1)} [2x(1)^2 - 16x(1) + 35 + 14a(1)^2 - 32a(1) - 32a(1) + 6x(1)a(1)]$$

Aplicando las condiciones de optimalidad se obtiene:

$$a^*(1) = \frac{16 - 3x(1)}{14}$$

$$J_1^*\{x(1)\} = \frac{1}{14}x(1)^2 - \frac{64}{7}x(1) + \frac{117}{7}$$

- Primer período: sea  $x(0) = 1$  dado.  
La ecuación de Bellman para este período es:

$$J_0^* = \min_{a(0)} E\{1 + a(0)^2 + J_1^*\{x(1)\}\}$$

en donde

$$x(1) = \alpha + 2a(0) + 1$$

Por tanto

$$J_0^* = \min_{a(0)} E\left\{1 + a(0)^2 + \frac{19}{14}[\alpha + 2a(0) + 1]^2 - \frac{64}{7}[\alpha + 2a(0) + 1] + \frac{117}{7}\right\}$$

Efectuando operaciones y teniendo en cuenta que

$$E\{\alpha\} = 1, \quad E\{\alpha^2\} = \text{var}(a) + [E\{a\}]^2 = 2$$

se llega a que

$$J_0^* = \min_{a(0)} E\left[\frac{45}{7}a(0)^2 - \frac{52}{7}a(0) + \frac{87}{14}\right]$$

Aplicando las condiciones de optimalidad, se obtiene:

$$a^*(0) = \frac{26}{45}$$

y

$$x(0) = 1, \quad a^*(0) = \frac{26}{45}$$

por lo que

$$x^*(1) = \alpha + \frac{97}{45}$$

y el valor que tome finalmente  $\alpha$  no se conocerá hasta el final del período 1, cuando se realice la variable aleatoria  $\alpha$ .

$$a^*(1) = \frac{16 - 3x^*(1)}{14}$$

$$x^*(2) = \alpha - \frac{3}{14}\alpha b + \frac{143}{210}b + \frac{7}{45}$$

y el valor que tome finalmente no se conocerá hasta el final del período 2, cuando se haya realizado la variable aleatoria  $b$ .

Por otra parte

$$J^* = J_0^* = \frac{2563}{630}$$

- b) En el problema dado, sustituimos las variables aleatorias  $\alpha$  y  $b$  por sus respectivas esperanzas matemáticas, que son 1 y 3 y quitamos el operador esperanza matemática en la función objetivo, obteniendo el siguiente problema de control óptimo determinístico:

$$\begin{aligned} \min_{a(0), a(1)} \quad & J = \{[x(0) - 2]^2 + [x(1) - 3]^2 + [x(2) - 3]^2 + a(0)^2 + [a(1) - 1]^2\} \\ \text{sujeto a} \quad & x(1) = x(0) + 2a(0) + 1 \\ & x(2) = x(1) + 3a(1) - 2 \\ \text{con} \quad & x(0) = 1 \end{aligned}$$

Resolvemos el problema por programación dinámica.

- Sea  $x(2)$  dado. Entonces

$$J_2^*\{x(2)\} = [x(2) - 3]^2.$$

- Segundo período. Sea  $x(1)$  dado.

La ecuación de Bellman correspondiente a este período es:

$$J_1^*\{x(1)\} = \min_{a(1)} \{[x(1) - 3]^2 + [a(1) - 1]^2 + J_2^*\{x(1) + 3a(1) - 2\}\}$$

Se resuelve dicha ecuación de Bellman, obteniendo la siguiente solución:

$$a^*(1) = \frac{16 - 3x(1)}{10}$$

$$J_1^*\{x(1)\} = \frac{11}{10}x(1)^2 - \frac{32}{5}x(1) + \frac{94}{10}$$

- Primer período. Sea  $x(0) = 1$  dado.

La ecuación de Bellman correspondiente a este período es:

$$J_0^* = \min_{a(0)} \{1 + a(0)^2 + J_1^*\{2 + 2a(0)\}\}$$

Se resuelve dicha ecuación de Bellman, obteniéndose la siguiente solución:

$$a^*(0) = \frac{10}{27} \quad J_0^* = \frac{34}{27}$$

Comparando los controles óptimos obtenidos para el problema determinístico definido en este apartado y para el problema estocástico resulto en a), se observa que son diferentes, por lo que no se cumple el principio de equivalencia cierta.

## 2.7 Equivalencia cierta para el problema LQR

El ejemplo anterior muestra claramente que el punto crucial es la pregunta: ¿Qué se entiende por ley de control de equivalencia cierta? Esta es precisamente la pregunta que hemos tratado de responder en la sección anterior.

Debido a lo anterior, en esta sección mostraremos que el problema-LQR es el equivalente cierto.

Consideremos primero el problema LQR en tiempo discreto. Sea  $T = [k_0, k_0+1, \dots, k_1]$ ,  $-\infty < k_0 \leq k_1 < \infty$ ,  $U = \mathbb{R}^m$ ,  $\mathbf{U} = U^T$ ,  $Y = \mathbb{R}^p$ ,  $\mathbf{Y} = Y^T$  y la planta  $\mathbf{G}$  definida análogamente como en 2.8 a través del estado  $x \in \mathbb{R}^n$  por

$$\begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) + w_1(k) \\ y(k) &= C(k)x(k) + w_2(k) \\ x(k_0) &= x_0 \quad y(k_0) = 0 \end{aligned} \tag{2.31}$$

con  $\{x_0, w_1(k_0), \dots, w_1(k_1)w_2(k_0), \dots, w_2(k_1)\}$  vectores aleatorios Gaussianos mutuamente independientes,  $E\{x_0\} := x_0 E\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^t\} := \Sigma(k_0)$ ;  $w_1, w_2$  ruido blanco con media cero y covarianza  $V_1(k), V_2(k)$  respectivamente y  $A(k), B(k)$  y  $C(k)$  matrices de dimensión apropiada para  $k \in T$ . La función de costo está dada por

$$\sum_{k_0}^{k_1-1} [x^t(k+1)R(k+1)x(k+1) + u^t(k)R_2(k)u^t(k)] + x^t(k_1)P_1x(k_1) \quad (2.32)$$

con  $P_1 = P_1^t$ ,  $R_1(k) = R_1^t(k)$  y  $R_2(k) = R_2^t(k)$  matrices de dimensión apropiada,  $R_1(k) \geq 0$ ,  $R_2(k) > 0$  para  $k \in T$  y  $P_1 \geq 0$ .

La ley de control admisible consiste en todos los mapeos no anticipativos de  $\mathbf{Y}$  en  $\mathbf{U}$ . Supongamos, como en toda la teoría de control óptimo LQR, que la ecuación de Riccati en tiempo discreto

$$P(k-1) = [A(k-1) - B(k-1)F(k-1)]^t [R_1(k) + P(k)] A(k-1), \quad P_1(k_1) = P_1 \quad (2.33)$$

$$F(k-1) = [B^t(k-1)(R_1(k) + P(k))B(k-1) + R_2(k-1)]^{-1} B^t(k-1) [R_1(k) + P(k)] A(k-1) \quad (2.34)$$

tiene una solución para  $k_0 \leq k \leq k_1$ .

Ahora, probaremos lo siguiente

**Teorema 2.4.** *El problema LQG en tiempo discreto es el equivalente cierto.*

**Demostración.**

La prueba se basa en dos teoremas, el primero no es más que un ejercicio de teoría LQ (determinístico).

**Teorema 2.5.** *Para el problema en consideración, el control  $\mathbf{u}(v, k, \omega)$  introducido en 2.30 para  $k \in \{k, k+1, \dots, k_1\}$  está dado por*

$$(u)^*(v, k, \omega)(k') = F(k')x(k', \omega) + L(k', \omega) \quad (2.35)$$

donde  $F$  es definida por (2.34) y  $L$  está dada por

$$L = -\tilde{F}w_1 - \hat{F} \quad (2.36)$$

con

$$\tilde{F} = [B^t(k)(R_1(k+1) + P(k+1))B(k) + R_2(k)]^{-1} \cdot B(k)(R_1(k+1) + P(k+1))$$

$$\hat{F} = [B(k)^t(R_1(k+1) + P(k+1))B(k) + R_2(k)]^{-1} \cdot B^t(k)N(k, \omega)$$

donde  $N(k, \omega)$  es la solución de la ecuación en diferencias hacia atrás

$$\begin{aligned} N(k, \omega) &= [A(k) - B(k)F(k)]^T N(k+1, \omega) + [A(k) - B(k)F(k)]^k \cdot [R_1(k+1) + P(k+1)]w_1(k, \omega) \\ N(k_1, \omega) &= 0 \end{aligned} \quad (2.37)$$

finalmente,  $x$  en 2.35 está definida recursivamente por la ecuación de planta 2.31 con

$$\mathbf{u}(k') = \begin{cases} v(k) & k' < k \\ -F(k')x(k', \omega) + L(t', \omega) & k' \geq k \end{cases}$$

Ahora, con el Teorema 2.5 es fácil identificar cuál es el mapeo  $\mathbf{R}$  de 2.30. Por lo tanto

$$\mathbf{R} : (\mathbf{u}, \omega) \longmapsto \tilde{\mathbf{u}}$$

con  $\tilde{\mathbf{u}}$  dado por

$$\tilde{\mathbf{u}} = -Fx + L \quad (2.38)$$

con  $x$  definida recursivamente por

$$\begin{aligned} x(k+1, \omega) &= A(k)x(k, \omega) + B(k)u(k) + w_1(k, \omega) \\ x(k_0, \omega) &= x_0(\omega) \end{aligned}$$

y  $L$  dada por 2.36 para  $k_0 \leq k \leq k_1$ .

Pasemos ahora al cálculo del mapeo  $\mathbf{S}$  del LQG en el caso de que el está definido como en 2.30, en general

**Teorema 2.6.** *Sea  $\tilde{\mathbf{u}}$  definida como se indicó anteriormente y  $y$  está dada por la ecuación de planta 2.31 (consideremos en ambas ecuaciones a  $\mathbf{u}$  como una función determinística). Entonces  $E\{\tilde{\mathbf{u}} | y_k^-\}$  está dada por  $E\{\tilde{\mathbf{u}}(k) | y_k^-\} = -F(k)E\{x(k, \omega) | y_k^-\}$ .*

### **Demostración.**

Resolviendo 2.37 y sustituyendo el resultado en 2.36 se muestra que  $L(k, \omega)$  es linealmente dependiente en  $w_1(k, \omega), w_1(k+1, \omega), \dots, w_1(k_1, \omega)$ . En otras palabras, las mediciones del tiempo hasta  $k$  son funciones históricas  $w_1(k_0, \omega), \dots, w_1(k_1 - 1, \omega)$  y de  $w_2(k, \omega)$ . Por lo tanto,  $L(k, \omega)$  es independiente de  $y_k^-$ .

Del Teorema 2.6 se sigue que  $\mathbf{S} : \mathbf{U} \times \mathbf{Y} \longmapsto \mathbf{U}$  está definida por  $\mathbf{S}(\mathbf{u}, \omega)(k) = -F(k)\hat{x}$  con  $\hat{x}(k) := E\{x(k) | y_k^-\}$ . La evaluación de  $\hat{x}(k)$  es simplemente el filtro de Kalman en tiempo discreto y se define recursivamente por

$$\begin{aligned} \hat{x}(k+1) &= A(k)\hat{x}(k) + B(k)\mathbf{u}(t) + \hat{L}(k) \cdot [y(k) - C(k)\hat{x}(k)] \\ \hat{x}(k_0) &= x_0 \end{aligned}$$

con  $\hat{L}(k)$  la ganancia del filtro de Kalman, obtenida, como es bien sabido por una solución de un tipo de ecuación de Riccati. Es importante observar que  $\hat{L}(k)$  es determinística.

Para calcular la ley de control de la equivalencia cierta basta ahora con resolver las ecuaciones  $\mathbf{S}(\mathbf{u}, \mathbf{y}) = \mathbf{u}$  que después de la sustitución en la expresión anterior para  $\mathbf{S}$ , da como resultado  $\mathbf{F}_{ce} : \mathbf{Y} \mapsto \mathbf{U}$  definida por

$$(\mathbf{F}_{ce}y)(k) = -F(k)\hat{k}$$

con  $\hat{x}(k)$  definida recursivamente por

$$\begin{aligned}\hat{x}(k+1) &= [A(k) - B(k)F(k)]\hat{x}(k) + \hat{L}(k) \cdot [y(k) - C(k)\hat{x}(k)] \\ \hat{x}(k) &= E\{x_0\}\end{aligned}$$

que define como es bien sabido la ley de control de retroalimentación óptima para el problema de LQG. Llegamos a la conclusión de que el problema LQG es el equivalente cierto.

□

Con esto concluimos la prueba del Teorema .

Para el problema de LQG en tiempo continuo la demostración es bastante parecida pesar que se utilizan tres teoremas en lugar de dos, se llega a la misma conclusión. Dado que nuestro tema de estudio consideramos solamente el tiempo discreto, se deja como objeto de estudio a lector la parte de tiempo continuo.

Para finalizar, en esta sección hemos dado un marco general para definir la ley de control de la equivalencia cierta. La idea básica es que la equivalencia cierta elige una política de control, en cada instante de tiempo, como el control de la esperanza condicional (dadas las mediciones hasta ese momento) de lo que sería el control óptimo si no hubiera incertidumbre. Dentro de este marco, no es difícil demostrar que los problemas LQG de tiempo discreto y continuo son equivalentes ciertos. Hemos presentado aquí una prueba explícita del caso de intervalos de tiempo finitos, pero estos resultados se extienden fácilmente al caso de tiempo infinito.

# Capítulo 3

## Modelamiento en Optimización Convexa No Diferenciable

Las causas de la no diferenciabilidad de un problema de optimización son diversas, sin embargo, con un fin explicativo se pueden agrupar en dos:

- La no diferenciabilidad de las funciones del problema, aunque las variables del problema sean continuas.
- La no continuidad de las variables del mismo, aunque las funciones pudieran ser diferenciables en un conjunto más amplio de puntos que el conjunto solución del problema. Se entiende por no continuidad de las variables el hecho de que algunas de ellas sólo pueda tomar valores enteros, o que sea continua por tramos o que sea semicontinua.

En muchos problemas la no diferenciabilidad tiene su origen en las dos causas mencionadas, por lo que no hay que considerarlas excluyentes. Además, hay casos en los que un modelo del primer tipo que, en un principio, no es lineal ni diferenciable, puede pasar a ser diferenciable transformando la parte del modelo que provoca la no diferenciabilidad lineal. En otros, cabe la posibilidad de transformarlo en un problema del segundo tipo, es decir, de programación entera mixta mediante la incorporación de variables enteras y/o binarias. Si bien persiste la no diferenciabilidad, se emplearan los métodos de resolución del capítulo anterior que resuelven en cada iteración un subproblema diferenciable. Aunque con esto no se está suponiendo que sean los únicos casos en que dicha transformación pueda darse ni que no existan transformaciones alternativas.

### 3.1 El problema del portafolio

En finanzas, un portafolio o cartera es una combinación de las inversiones realizadas por una institución o un individuo. La creación de un portafolio es parte de una estrategia de diversificación de la inversión y limitar el riesgo. Al ser propietario de varios activos, determinados tipos de riesgos (en particular, el riesgo específico) puede reducirse. Los activos del portafolio pueden incluir cuentas bancarias, acciones, bonos, opciones, warrants, certificados de oro, materias primas, contratos de futuros, las instalaciones de producción, o cualquier

otro elemento que se espera que conserve su valor.

En la construcción de una cartera de inversiones de una institución financiera generalmente ésta llevará a cabo su propio análisis de la inversión, mientras que un particular puede hacer uso de los servicios de un asesor financiero o una institución financiera que ofrece servicios de gestión de cartera.

### Gestión de un portafolio

La gestión de la cartera implica decidir qué activos incluir en el portafolio, teniendo en cuenta los objetivos del titular de la cartera y las cambiantes condiciones económicas. La selección consiste en decidir cuáles son los bienes a comprar, ¿cuántos comprar, cuándo comprar, y que activos desinvertir. Estas decisiones siempre implican algún tipo de medición del resultado o rendimiento de la cartera, y el riesgo asociado a este cambio (es decir, la desviación estándar del retorno). Normalmente, se compara con el rendimiento esperado de carteras con paquetes de activos diferentes.

Los objetivos y circunstancias de los inversores también deben ser considerados. Algunos inversores son más reacios al riesgo que otros.

### Modelos de gestión de un portafolio

Algunos de los modelos financieros utilizados en el proceso de valoración, selección de valores y gestión de carteras incluyen:

- Maximización el rendimiento, dado un nivel determinado de riesgo.
- La teoría moderna de gestión de portafolios o Modelo de Markowitz.
- Teoría del Arbitraje de Precios

Al tiempo  $t = 0$  se tiene un monto  $M$  que se desea invertir a una semana en un portafolio de inversión, integrado con acciones de  $n$  empresas. Se tiene como datos los precios diarios de cada una de las acciones en los tres meses previos a  $t = 0$ . El número de acciones de cada empresa se deben terminar de tal forma que el riesgo del portafolio sea mínimo y su rendimiento semanal sea mayor o igual a una  $r^*$  dada [39].

Para tener una mejor idea del problema, revisemos algunos conceptos de finanzas.

Un monto  $M^0$  que se invierte en el banco a un interés  $r$  anual, al término de un año se convierte en un monto  $M^1$  igual a

$$M^1 = M^0 + rM^0 = (1 + r)M^0$$

Observemos que  $r = \frac{M^1 - M^0}{M^0}$  es la ganancia relativa, se le conoce como el rendimiento de la inversión y en el caso de los depósitos a plazo fijo coincide con la tasa de interés.

En el caso de las acciones, como de otros activos financieros, el rendimiento durante un período de tiempo se define por las variaciones relativas del precio del activo y está dado por:



$$r = \frac{P^1 - P^0}{P^0} \quad (3.1)$$

con  $P^0$  al tiempo inicial y  $P^1$  al tiempo final. Observemos que  $P^1 = (1 + r)P^0$ , por lo que el concepto de rendimiento coincide con el que definimos para depósitos bancarios.

Los rendimientos de un depósito bancario son deterministas porque al depositar el dinero sabemos de antemano el rendimiento exacto que se recibirá a la fecha de vencimiento; en el caso de las acciones, las variaciones del precio dependen de muchos factores: del desempeño de la empresa, de la situación económica del país, del tipo de cambio, de las tasas de interés e inclusive de que tan optimistas o pesimistas son los participantes en el mercado accionario. En suma, son tantos los factores que intervienen, que es difícil prever de antemano si se incrementarán o se reducirán, y más difícil aún, en cuánto lo harán.

Dado que no podemos determinar con certeza el rendimiento a futuro de cada acción, ésta se comporta como una variable aleatoria. En consecuencia, al tiempo  $t = 0$ , a lo más que podemos aspirar es a calcular el valor esperado del rendimiento de una acción.

Una forma de calcular el valor esperado de una variable aleatoria es a través del cálculo del primer momento de distribución. ¿Qué tipo de distribución tienen los rendimientos de los activos con riesgo? Para tener una idea analicemos el comportamiento histórico de éstos; por ejemplo, a través de un histograma de los rendimientos diarios de cada acción.

### 3.1.1 Formulación matemática del problema

En el mercado financiero hay participantes a quienes se denominan agentes financieros que compran y venden activos financieros. Los activos financieros se clasifican en activos financieros de renta fija, cuyos rendimientos son deterministas, y activos financieros de renta variable, acciones y bienes de consumo cuyos rendimientos tienen incertidumbre.

El modelo de mercado con el que se trabajará tiene las siguientes reglas:

- a) Se trabaja en un mercado ideal en el que los bancos prestan y dan créditos a la misma tasa de interés.
- b) Principio de comparación: entre dos inversiones al mismo plazo, los inversionistas prefieren aquella que tenga rendimientos más altos.
- c) Principio de aversión al riesgo: si una inversión no tiene rendimientos fijos entonces tiene un riesgo. Entre dos inversiones con plazo y rendimiento estimado igual, se prefiere la de menor riesgo. Una inversión con riesgo debe tener rendimientos más altos que otra similar con riesgo menor. La tolerancia al riesgo es personal.
- d) Principio de no arbitraje: no se puede tener un portafolio de inversión cuyo precio inicial  $P^0$  sea cero, pero cuyo precio  $P^0 \geq 0$  y su valor esperado estrictamente mayor

que cero; o sea que sin invertir un centavo se tenga una posibilidad de hacer una ganancia positiva.

- e) Para simplificar el modelo se asume que se pueden comprar fracciones de acciones.
- f) Por realismo se permiten compras en corto; es decir, que se puede pedir prestado para comprar ciertas acciones. En este caso, para distinguir el dinero prestado del propio, a las cantidades prestadas que se inviertan se denotan con signo negativo.

### ***Planteamiento del problema***

Definición de la función objetivo, de las variables y los datos del problema.

*Objetivo:* Dados  $n$  activos con rendimiento esperado  $\bar{r}_i$  y varianza estimada  $\bar{\sigma}_i$ , determinar el portafolio con costo  $M$  y rendimiento esperado  $r^*$  que tenga mínimo riesgo.

El rendimiento relativo de un activo  $A_i$  se denotará por  $r_i$  y se define por la expresión (3.1). Si el precio al tiempo final  $P_1^i$  es una variable aleatoria, también lo es  $r_i$ . Sea  $m_i$  el número de acciones que se compran del activo  $i$ . Entonces

$$\begin{aligned} M &= m_1 P_1^0 + \dots + m_n P_n^0 \\ 1 &= \frac{m_1 P_1^0}{M} + \dots + \frac{m_n P_n^0}{M} \end{aligned}$$

Sean  $w_i = \frac{m_i P_i^0}{M}$  la variable que representa el porcentaje del capital  $M$  invertido en el activo  $A_i$ . Las variables  $w_i$  son las variables del problema de optimización. La ventaja de definir a las variables como  $w_i$  es que éstas no dependen del monto a invertir, por lo que podemos plantear el problema para cualquier monto  $M$ .

Las restricciones que deben satisfacer las  $w_i$  son las siguientes:

- a) Para que se cumpla el requisito de que el costo del portafolio sea igual a  $M$  se debe satisfacer que

$$\sum_{i=1}^n w_i = 1$$

- b) La segunda restricción es que el rendimiento del portafolio sea mayor al de un depósito a plazo fijo, supongamos que se cumple si es igual a  $r^*$ .

Para formular esta restricción en términos de las  $w_i$ , se hace lo siguiente: denotemos por  $V^0$  el valor del portafolio al tiempo  $t_1$ , y como  $r_p$  a los rendimientos del portafolio al tiempo  $t = 1$ . El rendimiento del portafolio es igual a

$$r_p = \frac{V^1 - V^0}{V^0}$$

como  $V^0 = M$  entonces:

$$\begin{aligned} r_p &= \frac{1}{M} \sum_{i=1}^n m_i [P_i^1 - P_i^0] \\ &= \sum_{i=1}^n \frac{m_i P_i^0}{M} \frac{[P_i^1 - P_i^0]}{P_i^0} \\ &= \sum_{i=1}^n w_i r_i \end{aligned}$$

La segunda restricción se formula matemáticamente de la siguiente forma:

$$\begin{aligned} E(r_p) &= \sum_{i=1}^n w_i E(r_i) \\ &\approx \sum_{i=1}^n w_i \bar{r}_i \\ &= r^* \end{aligned}$$

La función a minimizar se llama la función objetivo. La función objetivo es el riesgo del portafolio. El riesgo de un portafolio puede medirse de muchas formas. En el caso que se suponga que los rendimientos son normales, la varianza del portafolio es una buena medida de su riesgo ya que cualquier otra medida de riesgo depende de la varianza. La varianza de un portafolio se calcula de la forma siguiente:

$$\begin{aligned} Var(r_p) &= E[(r_p - E(r_p))^2] \\ &= E[(\sum_{i=1}^n w_i r_i - E(r_p))^2] \\ &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j E[r_i - E(r_i)] E[r_j - E(r_j)] \\ &= \sum_{i=1}^n \sum_{j=1}^n Cov(r_i, r_j) w_i w_j \\ &\approx \sum_{i=1}^n \sum_{j=1}^n \overline{Cov}(r_i, r_j) w_i w_j \end{aligned}$$

La formulación matemática del problema del portafolio óptimo es:

$$\begin{aligned} \min \quad & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \overline{Cov}(r_i, r_j) w_i w_j \\ \text{sujeto a:} \quad & \sum_{i=1}^n w_i \bar{r}_i = r^* \\ & \sum_{i=1}^n w_i = 1 \end{aligned}$$

La función objetivo se divide por un medio por comodidad.

El formular matemáticamente el problema de esta forma es idea de Markowitz, ganador del Premio Nobel de Economía por su teoría de riesgo-rendimiento.

Ahora veamos que el problema planteado es en efecto un *problema de optimización convexa*, para ello se debe probar que tanto la función objetivo y el conjunto definido por las restricciones son convexos.

El problema que se plantea es:

$$\begin{aligned} \min \quad & \frac{1}{2} w^t [\Sigma w] \\ \text{sujeto a:} \quad & w \in \Omega \end{aligned}$$

con  $\Omega = \{e \in \mathbb{R}^n \mid h_1(w) = r \cdot w - r^* = 0, h_2(w) = \vec{1} \cdot w - 1 = 0\}$  donde  $\vec{1}^t = (1, 1, \dots, 1)$ .

La matriz de varianza-covarianza es definida positiva por lo que este problema admite una solución única  $w \in \Omega$  siempre que  $w$  sea un punto regular. Observemos que salvo en el caso de que el portafolio se componga de activos que tienen todos el mismo rendimiento, el resto de los portafolios admisibles son regulares ya que

$$V = \{\vec{r}, \vec{1}\}$$

es un conjunto linealmente independiente. Por otro lado, el espacio de direcciones admisibles está formado por

$$N(x) = \{y \in \mathbb{R}^n \mid \vec{1}y = 0, \vec{1}^t y = 0\}$$

con dimensión  $n - 2$ .

Al aplicar las condiciones de primer orden, se obtiene el siguiente sistema lineal de ecuaciones:

$$[\Sigma]w - \lambda r - \mu \vec{1} = 0 \tag{3.2}$$

$$w^t r = r^* \tag{3.3}$$

$$\vec{1}^t w = 1 \tag{3.4}$$

Por comodidad se puso signo negativo en los términos que tienen los multiplicadores de Lagrange. En el caso de las restricciones de igualdad, el signo de los multiplicadores de Lagrange no es importante. El sistema anterior tiene como solución única a

$$w^* = \lambda[\Sigma]^{-1}r + \mu[\Sigma]^{-1}\vec{1} \quad (3.5)$$

Al sustituir  $w^*$  en (3.3) y (3.4) se obtiene el siguiente sistema de ecuaciones para  $\mu$  y  $\lambda$

$$r^t[\Sigma]^{-1}r\lambda + \vec{1}^t[\Sigma]^{-1}\mu = 1 \quad (3.6)$$

$$\vec{1}^t[\Sigma]^{-1}r\lambda + \vec{1}^t[\Sigma]^{-1}\vec{1}\mu = 1 \quad (3.7)$$

cuya solución es:

$$\lambda = \frac{r^*A - B}{\Delta} \quad \mu = \frac{C - r^*B}{\Delta}$$

con

$$A = \vec{1}^t[\Sigma]^{-1}\vec{1}, \quad B = \vec{1}^t[\Sigma]^{-1}r$$

$$C = r^t[\Sigma]^{-1}r, \quad \Delta = AC - B^2$$

$\Delta \neq 0$  siempre que los rendimientos de las acciones no sean todos iguales. Observemos que por la desigualdad de Cauchy–Schwartz,  $\Delta > 0$ , ya que

$$|\langle r, \vec{1} \rangle_{\Sigma^{-1}}| \leq \|r\|_{\Sigma^{-1}} \|\vec{1}\|_{\Sigma^{-1}}$$

con  $\|w\|_{\Sigma^{-1}} = w^t[\Sigma]^{-1}w$ .

Por otro lado, la varianza del portafolio mínimo es

$$\begin{aligned} (\sigma^*)^2 &= (w^*)^t[\Sigma](w^*) \\ &= (w^*)^t[\Sigma](\lambda[\Sigma]^{-1}r + \mu[\Sigma]^{-1}\vec{1}) \\ &= \lambda(w^*r) + \mu(w^*\vec{1}) \end{aligned}$$

Al sustituir  $\lambda$  y  $\mu$  y al usar el hecho que  $w^*$  está en  $\Omega$  se obtiene que

$$(\sigma^*)^2 = \frac{A(r^*)^2 - 2Br^* + C}{\Delta} \quad (3.8)$$

A cada portafolio admisible  $w$  se le puede asociar un punto  $(\sigma_w, r_w)$  en el plano, cuyo eje de abscisas representa  $\sigma$  y el eje de las ordenadas a  $r$ . A  $\sigma_w$  se le llama *volatilidad del portafolio*  $w$ . El conjunto  $\{(\sigma_w, r_w) | w \in \Omega\}$  se le llama el conjunto de *riesgo–rendimiento*. Los puntos  $(\sigma^*, r^*)$ , correspondiente a los portafolios óptimos  $w^*$ , están sobre la frontera del conjunto de riesgo–rendimiento ya que dado un  $r^*$  tienen la volatilidad mínima. Determinemos la forma de la frontera despejando a  $r^*$  de la expresión (3.8), entonces, dado  $\sigma$ ,  $r$  se obtiene que

$$r(\sigma) = \frac{B}{A} \pm \sqrt{\frac{\Delta}{A}(\sigma^2 - \frac{1}{A})}$$

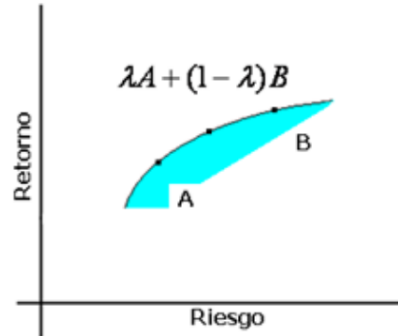


Figura 3.1: Una combinación convexa de dos portafolios eficientes produce otro portafolio eficiente

cuya gráfica es una hipérbola. En la Figura 3.1 se representa la gráfica de un conjunto de riesgo–rendimiento

El portafolio de mínima varianza se obtiene igualando la primera derivada de  $\sigma$  respecto a  $r$  e igualando a cero:

$$\frac{d\sigma}{dr} = \frac{2Ar - 2B}{\Delta} = 0 \Rightarrow R_{MV} = \frac{B}{A}$$

El valor de  $\sigma_{MV} = \frac{1}{\sqrt{A}}$ . Por lo que  $r(\sigma)$  es una hipérbola con vértice en  $\left(\frac{1}{\sqrt{A}}, \frac{B}{A}\right)$  y asíntotas

$$\frac{B}{A} \pm \sqrt{\frac{\Delta}{A}}\sigma$$

Para el portafolio de mínima varianza los multiplicadores de Lagrange correspondientes son  $\lambda_{MV} = 0$  y  $\mu_{MV} = \frac{1}{A}$ . En consecuencia

$$w_{MV} = \frac{[\Sigma]^{-1}\vec{1}}{A}$$

lo que es equivalente a resolver el sistema

$$[\Sigma]w_{MV} = \frac{\vec{1}}{A}$$

Observemos que la curva que va del portafolio de mínima varianza hacia arriba, representa aquellos portafolios que dado un riesgo tienen rendimiento máximo, por lo que a la parte superior de la curva se le llama *frontera eficiente*.

Por otro lado, si escogemos  $\mu = 0$  se obtiene que  $r_d = \frac{C}{B}$ , siempre que  $B \neq 0$ , entonces, si denotamos como  $w_d$  al portafolio correspondiente obtenemos que

$$[\Sigma]w_d = \frac{r}{B}, \quad \sigma_d = \frac{\sqrt{C}}{B}$$

Notemos que cualquier  $w^*$  que satisfaga la ecuación (3.5) se puede escribir en términos del portafolio de mínima varianza  $w_{MV}$  y del portafolio  $w_d$  ya que

$$w^* = (\lambda B)w_d + (\mu A)w_{MV} \quad (3.9)$$

lo que implica que cualquier portafolio sobre la frontera se puede escribir como combinación lineal de los portafolios  $w_d$  y  $w_{MV}$ . Observemos que  $w^*$  satisface también la ecuación (3.4) ya que  $(\lambda B) + (\mu A) = 1$ .

Este resultado no sólo se cumple para los portafolios  $w_d$  y  $w_{MV}$  sino para cualquiera dos portafolio  $w_1$  y  $w_2$  sobre la frontera de la región de riesgo–rendimiento con rendimientos  $r_1$  y  $r_2$  respectivamente. Como  $w_1$  se puede escribir como  $w_1 = (1 - a)w_{MV} + aw_d$  y  $w_2 = (1 - b)w_{MV} + bw_d$ , con  $a, b \geq 0$ , entonces, resolviendo para  $w_1, w_2$  y sustituyendo en (3.9) se obtiene que todo portafolio en la frontera se puede escribir como combinación lineal de dos portafolios sobre dicha frontera.

Como consecuencia, si se desea invertir en la bolsa una cantidad de dinero, en lugar de tratar de integrar uno mismo su portafolio, basta con seleccionar dos fondos de inversión en el mercado que estén sobre la frontera eficiente e invertir en ellos [39].

## 3.2 Maximización de la utilidad

En matemáticas aplicadas, las funciones lineales y no lineales pueden representar algún criterio para tomar decisiones. Por ejemplo, si una función simboliza utilidades, los ingresos, costos o nivel de contaminación de una empresa, esa variable debe seleccionarse adecuadamente por la persona que va a tomar la decisión con el fin de conseguir un objetivo preestablecido [31].

El criterio más común para elegir alternativas tiene como objetivo maximizar algo (la ganancia de una empresa, la utilidad que tiene un producto o servicio para consumidor, etc.) o minimizar (el costo de producción total o de un producto dado). Económicamente, se clasifica a esos problemas de maximización o minimización bajo el título de *optimización*, que debe entenderse como la búsqueda de lo mejor. No obstante, matemáticamente, los términos, máximo y mínimo no implican una optimización. Es por ello que el término general para máximo y mínimo como conceptos matemáticos, se designa más apropiadamente como valores extremos.

Uno de los objetivos fundamentales de una empresa es lograr el máximo beneficio (o utilidad, denotado por la letra  $U$ ). La utilidad o beneficio que espera una empresa es la cantidad monetaria que ha de ganar al vender sus productos y obtener un ingreso  $I(x)$  por esa venta, menos la parte que corresponde a solventar los gastos (costos  $C(x)$ ) necesarios para llevar a cabo la producción.

$$U = I(x) - C(x)$$

Resulta obvio que para determinar la utilidad es un requisito conocer tanto el nivel de ingresos que tiene la empresa como el nivel de costos en el que incurre para producir.

El *ingreso*  $I(x)$  está en función de la cantidad de productos vendidos, es decir, se vende  $x$  cantidad de productos a un determinado precio  $p$ . En el momento en el cual la cantidad de artículos que se vende se multiplica por su precio, se tiene un monto monetario que caracteriza al ingreso de la empresa, es decir:

$$I(x) = (p)(x)$$

La empresa siempre espera vender la mayor cantidad de artículos que le sea posible a fin de obtener el máximo de ganancia, pero también es cierto que si desea vender tales montos, lo hará con el objetivo de reducir al máximo los costos en los que incurre al llevar a cabo su ciclo productivo, es decir, la empresa espera optimizar sus recursos al **maximizar su beneficio y minimizar sus costos**.

Maximizar el beneficio de cualquier empresa implica:

1. Maximizar el ingreso, es decir, vender la mayor cantidad de artículos posibles, con un nivel de costos constante.
2. Maximizar el ingreso y reducir el nivel de costos.
3. Minimizar los costos y mantener constante su nivel de ventas de forma que su ingreso no se vea afectado.

Naturalmente, el punto dos sería el ideal para una empresa, pero es el más difícil de lograr, por lo que es común buscar rutas alternativas para que se cumplan en la medida de lo posible los puntos 1 y 3.

Como puede observarse, cuando se requiere maximizar el beneficio se tienen dos posibilidades para lograrlo: minimizar los costos o maximizar el ingreso. Este punto es importante para recalcar que para maximizar el beneficio es necesaria en la mayoría de las ocasiones una maximización del ingreso, punto clave para obtener las mayores ganancias posibles [31].

### 3.2.1 Formulación matemática del problema

En determinados casos, la función de utilidad no cumple las hipótesis habituales de buen comportamiento, no siendo diferenciable en todo el conjunto de oportunidades de un problema de optimización. Tal es el caso de la *función de utilidad de bienes perfectamente complementarios o función de utilidad de coeficientes fijos*. En esta función, el nivel de utilidad se determina por el consumo de las cantidades de los  $n$  bienes complementarios combinados en una proporción fija, por lo que no se puede conseguir mayores niveles de utilidad con un aumento del consumo de sólo unos pocos bienes, sino que debería crecer el consumo de dichos bienes en forma proporcional para alcanzar mayores niveles de utilidad [3].



**Definición 3.1.** Diremos que una función de utilidad representa las preferencias de una persona cuando se tiene que  $y \succsim x$  si y sólo si  $u(y) \geq u(x)$  [38].

**Teorema 3.1. (Convexidad débil).** Para cualquier  $x \in X$  los conjuntos  $Q_x = \{y \in X \mid y \succsim x\}$  son convexos [38].

Así pues, la utilidad alcanzada viene dada por la menor utilidad proporcionada por uno de los  $n$  bienes complementarios que se recoge en la siguiente función

$$U(x_1, x_2, \dots, x_n) = \min\{a_1x_1, a_2x_2, \dots, a_nx_n\} \quad \text{con } a_j > 0 \quad \forall j$$

El problema de maximizar la utilidad vendrá dado por el siguiente planteamiento:

$$\begin{aligned} \max \quad & U(x_1, x_2, \dots, x_n) = \max\{\min\{a_1x_1, a_2x_2, \dots, a_nx_n\}\} \\ \text{sujeto a} \quad & p_1x_1 + p_2x_2 + \dots + p_nx_n \leq M \\ & x_1, x_2, \dots, x_n \geq 0 \end{aligned}$$

La función de utilidad no es lineal y no es diferenciable en los puntos  $(x_1, \dots, x_n)$  tales que  $a_ix_i = a_jx_j$  con  $i \neq j$ , pues no es derivable en dichos puntos. Sin embargo es una función continua y cóncava.

Alternativamente al tratamiento anterior, este tipo de función admite una transformación que consiste en introducir una nueva variable que representará el valor de la función objetivo, por ejemplo:  $u_{MIN}$ , tal que  $u_{MIN} = \min\{a_1x_1, a_2x_2, \dots, a_nx_n\}$  y dado que  $u_{MIN}$  es igual al menor valor de todos los términos  $a_ix_i$ ,  $i = 1, 2, \dots, n$ , se establecen tantas cotas superiores a esa nueva variable como variables  $x_i$  se contemplen (bienes perfectamente complementarios), de tal manera que el modelo resultante será:

$$\begin{aligned} \max \quad & U(x_1, x_2, \dots, x_n, u_{MIN}) = u_{MIN} \\ \text{sujeto a:} \quad & p_1x_1 + p_2x_2 + \dots + p_nx_n \leq M \\ & u_{MIN} \leq a_ix_i \quad \forall 1, 2, \dots, n \\ & x_1, x_2, \dots, x_n \geq 0 \\ & p_1, p_2, \dots, p_n > 0 \\ & M > 0 \end{aligned}$$

De esta forma, el problema de programación no lineal no diferenciable se ha transformado en uno de programación lineal con una variable y  $n$  restricciones lineales más que la formulación original [3].

Debemos asegurarnos que el problema anterior sea convexo, es evidente que las restricciones de éste sean convexas, por lo que solo falta verificar que la función objetivo también lo sea, para ello presentamos el siguiente teorema [38]

**Teorema 3.2.** Las preferencias son convexas si y sólo si para cualquier par de canastas  $x, y \in X$ ,  $1 \geq \alpha \geq 0$  si  $y \succsim x$ , entonces  $\alpha y + (1 - \alpha)x \succsim x$

***Demostración.***

$\Rightarrow$  Sean  $x, y \in X$  tales que  $x \succsim y$ . Entonces, tanto,  $x$  como  $y$  pertenecen a  $Q_x$  y como este conjunto es convexo  $\alpha y + (1 - \alpha)x \succsim x$

$\Leftarrow$  Sean  $x, y$  canastas en  $Q_x$ . Luego, por condición de convexidad  $\alpha y + (1 - \alpha)x$  está en  $Q_x$ . Luego, por definición de  $Q_x$   $\alpha y + (1 - \alpha)x \succsim x$

□

La convexidad refleja la idea que las personas prefieren la diversidad. Este no es un supuesto muy restrictivo y en términos generales refleja como son las preferencias de las personas.

### 3.3 Problema de localización de instalaciones capacitado en dos etapas

Los problemas de localización de instalaciones han sido estudiados ampliamente en la literatura, ya que surgen en una gran variedad de situaciones reales (telecomunicaciones, escuelas, hospitales). Básicamente, un problema de localización se caracteriza por los siguientes elementos [30]:

- Un conjunto de localizaciones posibles para las instalaciones. Para cada una se tiene información sobre los costos/beneficios de abrir o construir la instalación en dicha localización.
- Un conjunto de clientes que deben ser asignados para su servicio a alguna de las instalaciones. Para los clientes, también se cuenta con información sobre la cantidad de demanda y los costos por ser atendidos desde alguna de las instalaciones.
- Una lista de requerimientos que deben ser cumplidos por las instalaciones abiertas y las asignaciones realizadas a los clientes.
- Una función de costo/beneficio.

El objetivo del problema de localización de instalaciones es encontrar cuáles instalaciones deben ser abiertas de manera que se optimice esa función de costo/beneficio.

Hay una gran variedad de problemas de localización de instalaciones. Si el conjunto de puntos de demanda y localización de las instalaciones es finito, se conoce como *problema de localización de instalaciones discreto*. Si todos los datos son exactos, se refiere al modelo como *determinístico*, si los valores de algunos parámetros son dados por una distribución de probabilidad, el modelo se considera *estocástico*. Podemos clasificar también el modelo como capacitado o no capacitado, esto es, si las instalaciones cuentan con una capacidad limitada o no para satisfacer la demanda de los clientes.

En este problema, un único producto es producido en un conjunto de plantas con capacidades conocidas y enviado a un conjunto de almacenes también con capacidades conocidas, desde

### 3.3. PROBLEMA DE LOCALIZACIÓN DE INSTALACIONES CAPACITADO EN DOS ETAPAS<sup>93</sup>

los cuales se enviará el producto hacia un conjunto de clientes para finalmente abastecer sus demandas. Usar una determinada planta o almacén incurre en un costo fijo, además del costo por transportar cada unidad de producto entre las plantas y los almacenes y entre los almacenes y los clientes.

El objetivo del problema de localización capacitado en dos etapas es determinar cuáles plantas y cuáles almacenes deben utilizarse, así como la cantidad de producto que va de cada planta a cada almacén y de cada almacén a cada cliente.

El manejo de la cadena de suministro implica no solo el movimiento de los bienes, sino también decisiones como: dónde producir, localizar plantas y centros de distribución, así como cuánto producir en cada lugar.

Las decisiones de localización, además de necesarias, son tal vez las más importantes y difíciles para lograr cadena de suministro eficiente.

#### ***Problema de localización de instalaciones no capacitado***

El problema de localización de instalaciones se puede clasificar en distintas categorías dependiendo de las restricciones que se consideran. Entre ellas tenemos los *Problema de localización de instalaciones*:

- *No capacitado*. en éste no se consideran límites de capacidad para ninguna de las instalaciones; cada cliente recibe la demanda requerida desde algunas de las instalaciones.
- *Capacitado*. Cuando se considera un límite de capacidad para cada una de las instalaciones. Este problema puede ser dividido en otras categorías como por ejemplo, el problema de localización de instalaciones capacitado con restricciones de abastecimiento de una sola fuente.

El problema que se considera es el *problema de localización de instalaciones capacitado en dos etapas*, donde la primera etapa se refiere a las fábricas o plantas de producción, en esta etapa la decisión a ser tomada es cuáles fábricas abrir de un conjunto de posibles localizaciones para éstas. La segunda etapa se refiere a los depósitos y la decisión que se busca tomar es cuáles depósitos abrir entre un conjunto de posibles localizaciones para los mismos.

La siguiente decisión está relacionada a los clientes, se busca hacer las asignaciones de cada cliente a los depósitos abiertos y también asignar los depósitos a plantas abiertas y definir el flujo de producto a través de ellos para satisfacer las restricciones de demanda de cada cliente.

#### **3.3.1 Formulación al problema y relajaciones propuestas**

Es conocido que la formulación de un problema es de gran importancia para su resolución. Para formular el problema se describen los siguientes parámetros:

- $I$  conjunto de plantas potenciales.
- $J$  conjunto de almacenes potenciales.

- $K$  conjunto de clientes.
- $b_i$  capacidad de la planta  $i$ .
- $p_j$  capacidad de la planta  $j$ .
- $q_k$  demanda del cliente  $k$ .
- $c_{ij}$  costo por transportar una unidad de producto desde la planta  $i$  hasta el almacén  $j$ .
- $d_{jk}$  costo por transportar una unidad de producto desde el almacén  $j$  hasta el cliente  $k$ .
- $f_i$  costo por emplear la planta  $i$ .
- $g_j$  costo fijo por emplear el almacén  $j$ .

Definimos también las siguientes variables de decisión:

- $y_i \in \{0, 1\}$ , donde  $y_i = 1$  si sólo si la planta  $i$  es empleada y 0 en otro caso.
- $z_j \in \{0, 1\}$ , donde  $z_j = 1$  si y sólo si el almacén  $j$  es empleado y 0 en otro caso.
- $x_{ij} \geq 0$ , cantidad de producto enviado desde la planta  $i$  hasta el almacén  $j$ .
- $s_{jk} \geq 0$ , cantidad de producto enviado desde el almacén  $j$  hasta el cliente  $k$ .

El problema puede ser formulado mediante el siguiente modelo de programación lineal entero mixto.

$$\omega = \min \sum_i f_i y_i + \sum_{(i,j)} c_{ij} x_{ij} + \sum_{(j,k)} d_{jk} s_{jk} \quad (3.10)$$

$$\text{sujeto a:} \quad \sum_j x_{ij} \leq b_i \quad i \in I \quad (3.11)$$

$$\sum_i x_{ij} \leq p_j \quad j \in J \quad (3.12)$$

$$\sum_j s_{jk} \geq q_k \quad k \in K \quad (3.13)$$

$$\sum_i x_{ij} \geq \sum_k s_{jk} \quad j \in J \quad (3.14)$$

$$x_{ij} \leq m_{ij} y_i \quad i \in I, j \in J \quad (3.15)$$

$$s_{jk} \leq l_{jk} z_j \quad j \in J, k \in K \quad (3.16)$$

$$x_{ij}, s_{jk} \in \mathbb{R}^+, y_i, z_j \in \{0, 1\} \quad (3.17)$$

Las restricciones (3.11) y (3.12) aseguran que no sean excedidas las capacidades de las plantas y los almacenes respectivamente. La restricción (3.13) asegura que las demandas de los clientes sean satisfechas. La restricción (3.14) impide que los almacenes envíen producto no recibido mientras que (3.15) y (3.16) impiden que una planta o un almacén envíen producto si no han sido elegidos. Las restricciones (3.17) son conocidas como restricciones lógicas, en ellas  $\mathbb{R}^+$  representa el conjunto de números reales no negativos.

Las cantidades  $m_{ij}$  y  $l_{jk}$  en las restricciones (3.15) y (3.16) son cotas superiores conocidas a las variables  $x_{ij}$  y  $s_{jk}$ , por ejemplo:

$$m_{ij} = \min\{b_i, p_j\}; \quad l_{jk} = \min\{p_j, q_k\}$$

Esta formulación tiene  $i + j$  variables binarias,  $i \times j + j \times k$  variables continuas y  $i + 2 \times j + k + i \times j + j \times k$  restricciones.

### ***Relajaciones propuestas***

Las cotas lagrangianas han sido ampliamente utilizadas como base de muchas técnicas numéricas, por ejemplo en los esquemas de ramificación y acotamiento para problemas enteros y combinatorios. La mayoría de los enfoques de relajaciones lagrangianas para el problema de localización capacitado en dos etapas están basados en dualizar las restricciones de demanda o las restricciones de capacidad de planta.

Se presentan 3 relajaciones lagrangianas para el problema de localización capacitado en dos etapas. Estas relajaciones poseen ciertas propiedades de descomposición de manera que resolver el problema lagrangiano es más fácil que resolver el original y conservando la calidad de la soluciones con respecto al valor óptimo del problema tanto como sea posible. Las relajaciones se describen a continuación:

**Relajación RA1:** Relajamos las restricciones (3.14), que no permiten que un almacén  $j$  envíe producto que no recibió.

$$\omega = \min \sum_i f_i y_i + \sum_j g_j z_j + \sum_{(i,j)} c_{ij} x_{ij} + \sum_{j,k} d_{jk} s_{jk} + \sum_j u_j \left( \sum_k s_{jk} - \sum_i x_{ij} \right)$$

sujeto a:

$$\sum_j x_{ij} \leq b_i \quad i \in I$$

$$\sum_i x_{ij} \leq p_j \quad j \in J$$

$$\sum_j s_{jk} \geq q_k \quad k \in K$$

$$x_{ij} \leq m_{ij} y_i \quad i \in I, j \in J$$

$$s_{jk} \leq l_{jk} z_j \quad j \in J, k \in K$$

$$x_{ij}, s_{jk} \in \mathbb{R}^+, y_i, z_j \in \{0, 1\}$$

El problema lagrangiano se descompone en un subproblema para la primera etapa en  $(x, y)$  con  $i$  variables binarias,  $i \times j$  variables continuas y  $i + j + i \times j$  restricciones:

$$\omega_1 = \min \sum_i f_i y_i + \sum_{(i,j)} c_{i,j} x_{ij} - \sum_{i,j} u_j x_{ij}$$

sujeto a:

$$\sum_j x_{ij} \leq b_i \quad i \in I$$

$$\sum_i x_{ij} \leq p_j \quad j \in J$$

$$x_{ij} \leq m_{ij} y_i \quad i \in I, j \in J$$

$$x_{ij} \in \mathbb{R}^+, y_i \in \{0, 1\}$$

y en  $|K|$  subproblemas para la segunda etapa en  $(s, z)$  con  $j$  variables binarias,  $j \times k$  variables continuas y  $j + 1$  restricciones:

$$\begin{aligned} \omega_2^k &= \min \sum_j g_j z_j + \sum_j d_{jk} s_{jk} + \sum_j u_j s_{jk} \\ \text{sujeto a} \quad & \sum_j s_{jk} \geq q_k \\ & s_{jk} \leq l_{jk} z_j \quad j \in J \\ & s_{jk} \in \mathbb{R}^+, z_j \in \{0, 1\} \end{aligned}$$

Donde la cota  $\omega^{RA1} = \omega_1 + \sum_k \omega_2^k$ .

**Relajación RA2:** Relajamos las restricciones (3.15), que no permiten que una planta  $i$  envíe producto si no ha sido elegida para ser empleada.

$$\begin{aligned} \omega &= \min \sum_i f_i y_i + \sum_j g_j z_j + \sum_{(i,j)} c_{ij} x_{ij} + \sum_{(j,k)} d_{jk} s_{jk} + \sum_{(i,j)} u_{ij} (x_{ij} - m_{ij} y_i) \\ \text{sujeto a:} \quad & \sum_j x_{ij} \leq b_i \quad i \in I \\ & \sum_i x_{ij} \leq p_j \quad j \in J \\ & \sum_j s_{jk} \geq q_k \quad k \in K \\ & \sum_i x_{ij} \geq \sum_k s_{jk} \quad j \in J \\ & s_{jk} \leq l_{jk} z_j \quad j \in J, k \in K \\ & x_{ij}, s_{jk} \in \mathbb{R}^+, y_i, z_j \in \{0, 1\} \end{aligned}$$

El problema lagrangiano se descompone en  $|I|$  subproblemas con solo una variable binaria  $y_i$  :

$$\begin{aligned} \omega_1^i &= \min \left\{ f_i y_i - \sum_j u_{ij} m_{ij} y_i \right\} \\ \text{sujeto a:} \quad & y_i \in \{0, 1\} \end{aligned}$$

Lo cual se puede resolver de manera analítica.

Y un subproblema para  $(x, s, z)$  con  $j$  variables binarias,  $i \times j + j \times k$  variables continuas y con  $i + 2 \times j + k + j \times k$  restricciones:

$$\omega_2 = \min \sum_j g_j z_j + \sum_{(i,j)} c_{ij} x_{ij} + \sum_{(j,k)} d_{jk} s_{jk} + \sum_{(i,j)} u_{ij} x_{ij}$$

sujeto a:

$$\sum_j x_{ij} \leq b_i \quad i \in I$$

$$\sum_i x_{ij} \leq p_j \quad j \in J$$

$$\sum_j s_{jk} \geq q_k \quad k \in K$$

$$\sum_i x_{ij} \geq \sum_k s_{jk} \quad j \in J$$

$$s_{jk} \leq l_{jk} z_j \quad j \in J, k \in K$$

$$x_{ij}, s_{jk} \in \mathbb{R}^+, z_j \in \{0, 1\}$$

Donde la cota  $\omega^{RA2} = \sum_i \omega_1^i + \omega_2$ .

**Relajación RA3:** Relajamos las restricciones (3.16), que no permiten que un almacén  $j$  envíe producto si no ha sido elegido para ser utilizado.

$$\omega = \min \sum_i f_i y_i \sum_j g_j z_j + \sum_{(i,j)} c_{ij} x_{ij} + \sum_{(j,k)} d_{jk} s_{jk} + \sum_{(j,k)} u_{jk} (s_{jk} - l_{jk} z_j)$$

sujeto a:

$$\sum_j x_{ij} \leq b_i \quad i \in I$$

$$\sum_i x_{ij} \leq p_j \quad j \in J$$

$$\sum_j s_{jk} \geq q_k \quad k \in K$$

$$\sum_i x_{ij} \geq \sum_k s_{jk} \quad j \in J$$

$$x_{ij}, s_{jk} \in \mathbb{R}^+, y_i, z_j \in \{0, 1\}$$

El problema lagrangiano se descompone en  $|J|$  subproblemas con solo una variable binaria  $z_j$ .

$$\omega_1^j = \min \left\{ g_j z_j + \sum_k u_{jk} l_{jk} z_j \right\}$$

sujeto a:  $z_j \in \{0, 1\}$



Lo cual se puede resolver de manera analítica.

y con un subproblema para  $(x, s, y)$  con  $i$  variables binarias  $i \times j + j \times k$  variable continuas y con  $i + 2 \times j + k + i \times j$  restricciones:

$$\begin{aligned} \omega_2 = \min \quad & \sum_i f_i y_i + \sum_{(i,j)} d_{jk} s_{jk} + \sum_{(j,k)} u_{jk} s_{jk} \\ \text{sujeto a:} \quad & \sum_j x_{ij} \leq b_i \quad i \in I \\ & \sum_i x_{ij} \leq p_j \quad j \in J \\ & \sum_j s_{jk} \geq q_k \quad k \in K \\ & \sum_i x_{ij} \geq \sum_k s_{jk} \quad j \in J \\ & x_{ij}, s_{jk} \in \mathbb{R}^+, y_i \in \{0, 1\} \end{aligned}$$

Donde la cota  $\omega^{RA3} = \sum_j \omega_1^j + \omega_2$ .

El valor óptimo de la relajación Lagrangiana se usa como cota dual y la solución lagrangiana se usa como punto de partida o referencia para producir una solución factible. Frecuentemente, una relajación lagrangiana es considerada como buena si produce una cota dual cercana al valor óptimo original, sin embargo, la calidad de la cota factible obtenida en base a esa solución lagrangiana puede ser tomada en cuenta también al evaluar la relajación.

Resolviendo las relajaciones propuestas por medio del algoritmo del subgradiente, obtenemos cotas lagrangianas del problema y se aplica un algoritmo simple que recupera factibilidad a partir de dichas soluciones lagrangianas

### 3.4 Recuperación de imágenes

En el desarrollo de este ejemplo se estudia un caso particular en el que la función objetivo es convexa y no diferenciable con restricciones lineales. Más concretamente, se presenta el problema de *programación cuadrática generalizada*.

Por comodidad, se usará la siguiente notación:

$$\begin{aligned} \min \quad F(x) = & \alpha \left( \frac{1}{2} x^t H x + p^t x \right) + \sum_{j=1}^{m_{if}} |\bar{c}_j^t x - \bar{b}_j| + \sum_{j=m_{if}+1}^{m_{if}+m_{df}} \max\{\bar{c}_j^t x - \bar{b}_j, 0\} \\ \text{sujeto a:} \quad & \bar{c}_j^t x = b_j \quad 1 \leq j \leq m_i \\ & \bar{c}_j^t \leq b_j \quad m_i << j m_i + m_d \end{aligned}$$

donde  $\alpha \in \mathbb{R}^+$ ,  $H \in \mathbb{R}^n$  simétrica y definida positiva,  $p, \bar{c}_j, c_j \in \mathbb{R}^n$  y  $\bar{b}_j, b_j \in \mathbb{R}$ .

Antes de presentar el problema en concreto, definiremos lo que es una imagen y a lo que se refiere recuperarla [27].

**Definición 3.2.** Una imagen en escala de grises es una aplicación

$$\begin{aligned} \omega : \Omega = [0, 1] \times [0, 1] &\longmapsto [0, \infty) \\ (x, y) &\longmapsto f(x, y) \end{aligned}$$

donde se asocia a cada punto  $(x, y)$  del cuadrado unidad un número  $\omega(x, y)$  que determinará la intensidad de brillo o escala de gris de la imagen en ese punto.

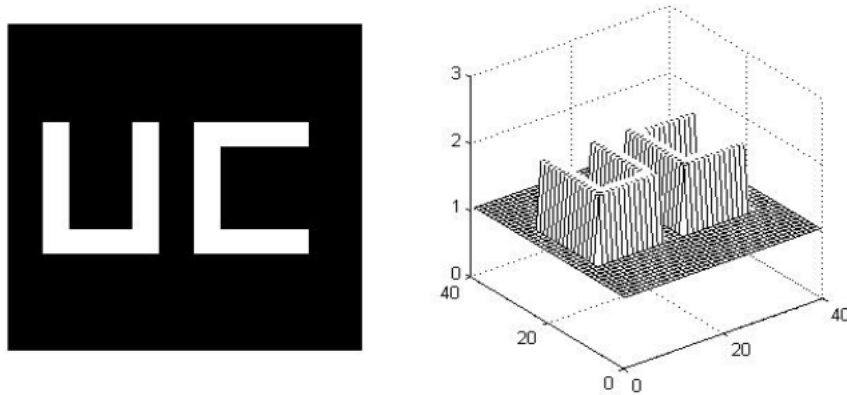


Figura 3.2: Ejemplo de una imagen y su gráfica en escala de grises.

Por diversos motivos, una imagen  $\omega$  puede verse afectada por dos tipos de distorsión, un *difuminado* y una distorsión aditiva o *ruido* de forma que se tiene:

$$Z = H\omega + \eta$$

donde:

- $z$  es la imagen distorsionada.
- $\omega$  es la imagen original.
- $H$  es el difuminado.
- $\eta$  es el ruido.

El objetivo es reconstruir  $\omega$  a partir de  $z$  en el caso especial de ausencia de difuminado, es decir,  $z = \omega + \eta$ .

Por ejemplo, en la Figura 3.2 tenemos una *imagen* a la izquierda y su correspondiente escala de grises o intensidades ( $\omega$ ) a la derecha y en la Figura 3.3 ambas representaciones  $z$  para una distorsión aditiva de la imagen de la Figura 3.2 presentada anteriormente.

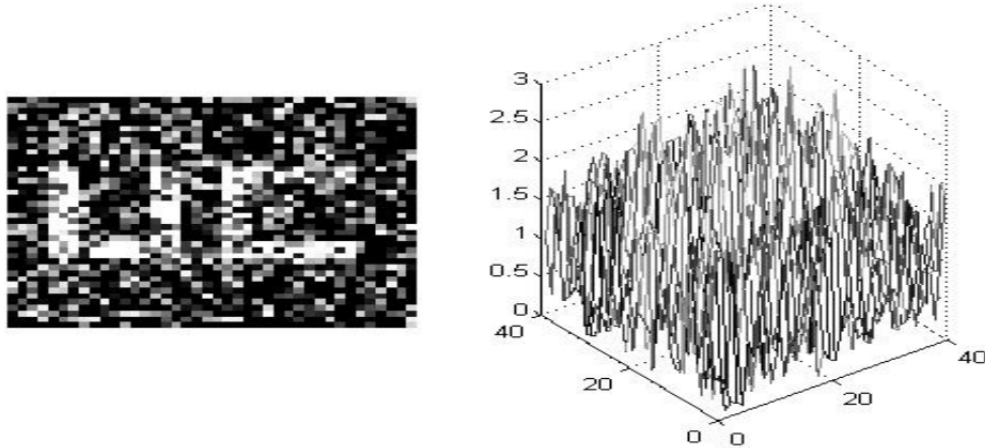


Figura 3.3: Ejemplo de una imagen distorsionada y su escala de grises.

La eliminación o reducción de este tipo de ruido es fundamental en diversos procesos de visión por computador, por ejemplo en imágenes tomadas con una cámara digital u otros dispositivos tales como un escáner. En algunas aplicaciones (por ejemplo médicas) es esencial que el tratamiento de las imágenes conserve los bordes, ya que son los elementos críticos en etapas posteriores para el conteo de objetos o el reconocimiento de estructuras.

Si observamos la imagen distorsionada de la izquierda de la Figura 3.3 fácilmente reconoceremos las siglas UC, lo que nos puede llevar a pensar que la recuperación planteada es un proceso en cierto sentido trivial o innecesario, sin embargo basta un vistazo a su escala de grises para darnos cuenta de que no es así. Este efecto se debe a la visión humana también posee algún tipo de procesamiento de imágenes, de forma que podemos reconocer ciertos patrones.

Para abordar el problema de la recuperación de imágenes, *teselaremos* el dominio de la imagen utilizando  $n = m^2$  cuadrados o *píxeles* y supondremos constante la intensidad en cada uno de los subdominios:

|           |           |         |           |
|-----------|-----------|---------|-----------|
| $x_{1,m}$ | $x_{2,m}$ | $\dots$ | $x_{m,m}$ |
| $\vdots$  | $\vdots$  | $\dots$ | $\vdots$  |
| $x_{1,2}$ | $x_{2,2}$ | $\dots$ | $x_{m,2}$ |
| $x_{1,1}$ | $x_{2,1}$ | $\dots$ | $x_{m,1}$ |

donde  $x_{i,j}$  representa la intensidad en el cuadrado  $(i, j)$ .

Así que la imagen distorsionada vendrá dada por una matriz de datos,  $z$ , que reordenada en forma vectorial da lugar al problema de mínimos cuadrados regularizado:

$$\begin{aligned} \min \quad & F(x) = \frac{\alpha}{2} \|x - z\|_2^2 + TV(x) \\ \text{sujeto a} \quad & x \geq 0 \end{aligned} \tag{3.18}$$

donde el parámetro de penalización,  $\alpha$ , mide la prioridad que damos al término mínimo

cuadrático frente a

$$TV(x) = \frac{1}{m} \left( \sum_{i=1}^m \sum_{j=1}^{m-1} |x_{i,j} - x_{i,j+1}| + \sum_{i=1}^{m-1} \sum_{j=1}^m |x_{i,j} - x_{i+1,j}| \right)$$

Este segundo término,  $TV(x)$ , mide la variación total en niveles de gris de una imagen  $x$  reordenada en forma matricial. Su uso está especialmente indicado para recuperar imágenes con bordes cortantes. Además puede expresarse como la norma de un producto matriz–vector. En efecto, sea  $E = (e_{i,j})$  la matriz  $(m-1) \times m$  dada por

$$e_{i,j} = \begin{cases} 1 & \text{si } i = j \\ -1 & \text{si } i = j - 1 \\ 0 & \text{en otro caso} \end{cases} \quad \text{para } i = 1, \dots, m-1$$

Entonces

$$TV(x) = \|\bar{C}^t x\| \quad \text{con } \bar{C} = \frac{1}{m} \frac{E \otimes_K I}{I \otimes_K E}$$

donde  $\otimes_K$  es el producto de Kronecker y  $\bar{C}$  tiene  $m_{if} = 2m(m-1)$  columnas y  $m^2$  filas. De este modo el problema (3.18) puede reescribirse como

$$\begin{aligned} \min \quad & F(x) = \frac{\alpha}{2} \|x - z\|_2^2 + \|\bar{C}x\|_1 \\ \text{sujeto a:} \quad & x \geq 0 \end{aligned} \quad (3.19)$$

Notemos que en nuestra formulación hemos añadido restricciones de cota sobre las variables, pero podemos encontrar otras formulaciones que han eliminado tales restricciones. Esto se debe fundamentalmente a que la resolución del problema es más sencilla.

### 3.4.1 Otra formulación equivalente

El problema de la recuperación de imágenes, como es usual en matemáticas, puede enfocarse desde varias perspectivas distintas. Ahora consideraremos una formulación equivalente a (3.19), aumentando considerablemente el número de variables, para obtener un problema de programación cuadrática convexa estándar (sin términos no diferenciables). Introduciendo variables auxiliares, podemos convertir el problema (3.19) en el problema:

$$\begin{aligned} \min \quad & G(x, u, v) = \frac{\alpha}{2} \|x - z\|_2^2 + \sum_{i=1}^{m_{if}} (u_i + v_i) \quad (x, u, v) \in \mathbb{R}^n \times \mathbb{R}^{m_{if}} \times \mathbb{R}^{m_{if}} \\ \text{sujeto a:} \quad & \bar{C}^t x - u + v = 0 \\ & x, u, v \geq 0 \end{aligned} \quad (3.20)$$

Veamos que ambos problemas son equivalentes. Antes de comenzar, llamemos  $\hat{p}$  el vector de términos lineales de (3.19) y  $\tilde{p}$  el de (3.20), de forma que se tiene

$$\tilde{p} = \begin{pmatrix} \hat{p} \\ e \\ e \end{pmatrix}$$

siendo  $e = (1, \dots, 1)^t$ . En primer lugar tenemos que  $\bar{x}$  es solución de (3.19) si y sólo si existen  $\bar{\lambda} \in \mathbb{R}^t$  y  $\bar{\mu} \in \mathbb{R}^n$  tales que:

$$\alpha \bar{x} + \hat{p} + \overline{C} \bar{\lambda} - \bar{\mu} = 0 \quad (3.21)$$

$$\bar{\lambda}_i = \begin{cases} 1 & \text{si } \bar{c}_i^t \bar{x} > 0 \\ -1 & \text{si } \bar{c}_i^t \bar{x} < 0 \end{cases} \quad (3.22)$$

$$\bar{\lambda}_i \in [-1, 1] \quad \text{si } \bar{c}_i^t \bar{x} = 0 \quad (3.23)$$

$$\bar{\mu} \geq 0, \quad \bar{\mu} \cdot \bar{x} = 0 \quad (3.24)$$

$$\bar{x} \geq 0 \quad (3.25)$$

Por otra parte  $(\tilde{x}, \tilde{u}, \tilde{v})^t$  es solución de (3.20) si sólo si existen  $\tilde{\lambda} \in \mathbb{R}^{m_{if}}$ ,  $\tilde{\mu}_x \in \mathbb{R}^n$ ,  $\tilde{\mu}_u \in \mathbb{R}^{m_{if}}$  y  $\tilde{\mu}_v \in \mathbb{R}^{m_{if}}$  cumpliendo:

$$\begin{pmatrix} \alpha \tilde{x} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \hat{p} \\ e \\ e \end{pmatrix} + \begin{pmatrix} \overline{C} \\ -I \\ I \end{pmatrix} \tilde{\lambda} - \begin{pmatrix} \tilde{\mu}_x \\ \tilde{\mu}_u \\ \tilde{\mu}_v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (3.26)$$

$$\left. \begin{array}{l} \tilde{\mu}_x \cdot \tilde{x} = 0 \quad \tilde{\mu}_x \geq 0 \\ \tilde{\mu}_u \cdot \tilde{u} = 0 \quad \tilde{\mu}_u \geq 0 \\ \tilde{\mu}_v \cdot \tilde{v} = 0 \quad \tilde{\mu}_v \geq 0 \end{array} \right\} \quad (3.27)$$

$$\left. \begin{array}{l} \overline{C}^t \tilde{x} - \tilde{u} + \tilde{v} = 0 \\ \tilde{x}, \tilde{u}, \tilde{v} \geq 0 \end{array} \right\} \quad (3.28)$$

Ahora veamos la equivalencia de estos dos sistemas de optimalidad. Supongamos que  $(\tilde{x}, \tilde{u}, \tilde{v})$  es solución de (3.20), entonces de la primera igualdad de la condición (3.26) obtenemos la condición (3.21) y de las otras igualdades tenemos que:

$$\tilde{\lambda} = e - \tilde{\mu}_u \quad (3.29)$$

$$\tilde{\lambda} = -e + \tilde{\mu}_v \quad (3.30)$$

de donde usando la positividad de los multiplicadores (ver (3.27)) se deduce que  $\tilde{\lambda}_i \in [-1, 1]$ .

Si  $\bar{c}_i^t \bar{x} > 0$ , por (3.28) tenemos que  $\tilde{u}_i - \tilde{v}_i > 0 \implies \tilde{u}_i > \tilde{v}_i \geq 0$ , lo que por (3.27) implica que  $(\tilde{u}_u)_i = 0$  y por tanto, de (3.30)  $\tilde{\lambda}_i = 1$ .

Si  $\bar{c}_i^t \bar{x} < 0 \implies \tilde{\lambda}_i = -1$ .

Por todo ello, obtenemos las condiciones (3.22) y (3.23). Por último, es claro que tomando  $\bar{x} = \tilde{x}$  y  $\bar{\mu} = \tilde{\mu}_x$  se satisfacen las condiciones (3.24) y (3.25), luego,  $\tilde{x}$  es la solución de (3.19).

Ahora para el recíproco, si  $\bar{x}$  es la solución de (3.19), tomando  $\tilde{u}_i = \max\{\bar{c}_i^t \bar{x}, 0\}$  y  $\tilde{v}_i = \max\{-\bar{c}_i^t \bar{x}, 0\}$  se cumple la condición (3.28). Con  $\tilde{\mu}_x = \bar{\mu}$ ,  $\tilde{\mu}_x = e - \bar{\lambda}$  y  $\tilde{\mu}_v = e + \bar{\lambda}$ , es claro que se cumple la condición (3.26), las condiciones sobre el signo y la primera igualdad de (3.27). Veamos que estos valores también cumplen el resto de la condición (3.27), distinguiendo tres casos:

- Si  $\bar{c}_i^t \bar{x} > 0$ , entonces  $\bar{\lambda}_i = 1$ ,  $(\tilde{\mu})_i = 0$  y  $\tilde{v}_i = 0$ .
- Si  $\bar{c}_i^t \bar{x} < 0$ , se tiene que  $\bar{\lambda}_i = -1$  por tanto  $(\tilde{\mu}_v)_i = 0$  y  $\tilde{u}_i = 0$ .
- Si  $\bar{c}_i^t \bar{x} = 0$  tenemos  $\tilde{u}_i = 0$  y  $\tilde{v}_i = 0$ .

Luego se cumple la condición (3.27) completa, obteniéndose que  $(\bar{x}, \tilde{u}, \tilde{v})$  cumple las condiciones de optimalidad para (3.20) y por tanto es la solución de ese problema.

Hemos logrado obtener una formulación cuadrática estándar equivalente a (3.19), pero aumentando considerablemente el número de variables. En principio es más sencillo abordar la resolución numérica de (3.20).

Veamos que este problema es en efecto convexo, para ello presentamos el siguiente teorema:

**Teorema 3.3.** Sea  $\lambda \geq 0$  y  $\alpha \geq 0$ . Definimos  $F_\alpha(x) : \mathbb{R}^N \mapsto \mathbb{R}$  como

$$F_\alpha(x) = \frac{1}{2} \|y - x\|_2^2 + \lambda \psi_\alpha(x)$$

donde

$$\begin{aligned} \psi_\alpha &= \|Dx\|_1 - S_\alpha(x) \\ S_\alpha(x) &= \min_{v \in \mathbb{R}^N} \{ \|Dv\|_1 + \frac{\alpha}{2} \|x - v\|_2^2 \} \end{aligned}$$

Si  $0 \leq \alpha \leq 1/\lambda$  entonces  $F_\alpha$  es convexa.

Si  $0 \leq \alpha < 1/\lambda$  entonces  $F_\alpha$  es estrictamente convexa.

**Demostración.**

$$\begin{aligned} F_\alpha &= \frac{1}{2} \|y - x\|_2^2 + \lambda \|Dx\|_1 - \lambda S_\alpha(x) \\ &= \frac{1}{2} \|y - x\|_2^2 + \lambda \|Dx\|_1 - \min_{v \in \mathbb{R}^N} \left\{ \|Dv\|_1 + \frac{\alpha}{2} \|x - v\|_2^2 \right\} \\ &= \max_{v \in \mathbb{R}^N} x \left\{ \frac{1}{2} \|y - x\|_2^2 + \lambda \|Dx\|_1 - \lambda \|Dv\|_1 - \frac{\lambda \alpha}{2} \|x - v\|_2^2 \right\} \\ &= \max_{v \in \mathbb{R}^N} \left\{ \frac{1}{2} (1 - \lambda \alpha) \|x\|_2^2 + \|Dx\|_1 + g(x, v) \right\} \\ &= \frac{1}{2} (1 - \lambda \alpha) \|x\|_2^2 + \|Dx\|_1 + \max_{v \in \mathbb{R}^N} g(x, v) \end{aligned}$$

donde  $g(x, v)$  es afín en  $x$ . El último término es convexo ya que es el máximo puntual de un conjunto de funciones convexas, por lo que  $F_\alpha$  es convexa si  $1 - \lambda\alpha \geq 0$  y es estrictamente convexa cuando  $1 - \lambda\alpha > 0$ .

### 3.5 Fractales y finanzas

A lo largo de la historia, la industria bancaria ha sufrido importantes pérdidas por fallos operacionales. Consciente de ello, el *Comité de Basilea* publicó, en 2004, un nuevo acuerdo de capital en el que instaba a las entidades financieras medir, controlar y gestionar su riesgo operacional. En este contexto el *valor en Riesgo* (VaR) se convierte en un elemento crucial para la medición del riesgo operacional.

¿Qué es el valor en riesgo–VAR?. Se trata de un método para cuantificar la exposición al riesgo de mercado, utilizando técnicas estadísticas tradicionales. Partamos de la base que los agentes económicos de hoy enfrentan riesgos de diferente naturaleza, como por ejemplo de crédito, de mercado, de liquidez, operacional, legal, etc. El valor en riesgo vendría a medir la pérdida que se podría sufrir en condiciones normales de mercado en un intervalo de tiempo y con un cierto nivel de probabilidad o de confianza.

Desde entonces han surgido diversas alternativas metodológicas para su cuantificación.

En la gama de alternativas de medición avanzada, el método de distribución de pérdidas (LDA, por sus siglas en inglés), que involucra herramientas del cálculo actuarial, ha sido el más generalmente utilizado en este campo.

**Definición 3.3.** *Un Sistema de Funciones Iteradas (IFS, por sus siglas en inglés) en un espacio métrico completo  $X$ , consiste en un número finito de transformaciones contractivas de  $X$ ,  $w_1, w_2, \dots, w_N$ , denotado por  $\{X | w_1, w_2, \dots, w_N\}$ .*

Bajo el contexto que se está trabajando tenemos la siguiente definición de *fractal*.

**Definición 3.4.** *Al compacto  $A_\infty$ , punto fijo de la transformación  $W$  se le llama Fractal asociado al  $\{X | w_1, w_2, \dots, w_N\}$ , se le conoce también como el "atractor" del IFS.*

Ahora, debido al amplio campo donde se encuentran y aplican los fractales y debido a que es un tema que está en investigación, se pueden encontrar otras definiciones, las más usuales son:

1. Los fractales son los objetos matemáticos que conforman la *Teoría del caos*.
2. Los fractales son objetos cuya dimensión es un número racional.
3. Un fractal es un objeto que su dimensión de **Hausdorff–Besicovich** es mayor que su dimensión topológica.

4. Un objeto fractal es aquel que tiene auto semejanza y su dimensión es racional.

Lo que nos lleva a definir el siguiente concepto.

**Definición 3.5. Dimensión fractal.** *Sentido genérico: Número que sirve para cuantificar el grado de irregularidad y fragmentación de un conjunto geométrico o de un conjunto natural. La dimensión fractal no es necesariamente entera.*

### 3.5.1 Planteamiento del problema

A continuación, se presenta la llamada estimación fractal, basada en sistemas de funciones iteradas, como posible alternativa para la medición del riesgo.

En aplicaciones prácticas un problema crucial es el llamado problema inverso, que se formula de la siguiente manera: dado  $f$  en un espacio métrico  $(S, d)$ , encontrar una contracción  $T : S \rightarrow S$  que admita un único punto fijo  $\hat{f} \in S$  tal que  $d(f, \hat{f})$  sea lo suficientemente pequeño. Si se resuelve el problema inverso con precisión adecuada, se puede identificar  $f$  con el operador  $T$  que tiene a  $\hat{f}$  como punto fijo.

Una solución aproximada del problema inverso es la solución del siguiente problema de optimización restringida, que llamaremos el problema (3.31) :

$$\min_{P \in C} d_{\infty}(T_P F, F) \quad (3.31)$$

donde  $T_P F$  es un operador que aproxima a  $F$  y que depende de un vector de probabilidades  $P$  que es tomado de un conjunto  $C$  y  $d_{\infty}(x, y)$  es la métrica del supremo.

Con un sistema fijo de mapeos  $w_i$  y parámetros  $\delta_i$ , el problema inverso puede ser resuelto, si es posible, usando los parámetros  $p_i$ . Estos deben ser elegidos en el siguiente conjunto convexo:

$$C = \left\{ P \in \mathbb{R}^n \mid p_i \geq 0, i = 1, \dots, N. \sum_{i=1}^N p_i = \sum_{i=1}^{N-1} \delta_i \right\}$$

En efecto, a continuación probamos que efectivamente el conjunto  $C$  es convexo.

**Teorema 3.4.** *El conjunto  $C$  definido por:*

$$C = \left\{ P \in \mathbb{R}^n \mid p_i \geq 0, i = 1, \dots, N. \sum_{i=1}^N p_i = \sum_{i=1}^{N-1} \delta_i \right\}$$

*es convexo.*



***Demostración.***

Sean  $A$ ,  $B$  y  $C$  definimos la recta  $l : A + (B - A)t; t \in \mathbb{R}$ . Consideremos el punto  $x \in \mathbb{R}^N$  que pertenece a la recta  $l$ , así tenemos que:

$$x = A + (B - A)t$$

$$\begin{aligned} \sum_{i=1}^N x_i &= \sum_{i=1}^N (A_i + (B_i - A_i)t) \\ &= \sum_{i=1}^N A_i + \left( \sum_{i=1}^N B_i - \sum_{i=1}^N A_i \right) t \\ &= \sum_{i=1}^N A_i(1 - t) + t \sum_{i=1}^N B_i \\ &= \left( 1 - \sum_{i=1}^N \delta_i \right) (1 - t) + t \left( 1 - \sum_{i=1}^N \delta_i \right) \\ &= \left( 1 - \sum_{i=1}^N \delta_i \right) (1 - t + t) \\ &= \left( 1 - \sum_{i=1}^N \delta_i \right) \end{aligned}$$

es decir:

$$\sum_{i=1}^N x_i = \left( 1 - \sum_{i=1}^{N-1} \delta_i \right)$$

Se ha hecho uso del hecho que:

$$\begin{aligned} \sum_{i=1}^N A_i &= \left( 1 - \sum_{i=1}^N \delta_i \right) \\ \sum_{i=1}^N B_i &= \left( 1 - \sum_{i=1}^N \delta_i \right) \end{aligned}$$

por lo tanto  $C$  es convexo.

□

**Teorema 3.5.** *La función*

$$\begin{aligned} D(P) : \quad \mathbb{R}^N &\mapsto \mathbb{R} \\ &= d_\infty(T_P(f), f) \end{aligned}$$

*es convexa*

***Demostración.***

Si escogemos  $P_1, P_2 \in \mathbb{R}^N$  y  $\lambda \in [0, 1]$  entonces:

$$D(\lambda P_1 + (1 - \lambda)P_2) = \sup_{x \in [0,1]} |T_{\lambda P_1 + (1-\lambda)P_2} f(x) - f(x)|$$

Notar que:  $-f(x) = -\lambda f(x) - (1 - \lambda)f(x)$ .

Tenemos que:

$$\begin{aligned} |T_{\lambda P_1 + (1-\lambda)P_2} f(x) - f(x)| &\leq \left| \lambda(P_1 f(w_i^{-1}(x))) + \sum_{j=1}^{i-1} P_{1j} + \sum_{j=1}^{i-1} \delta_j - \sum_{j=1}^{i-1} \delta_j \right. \\ &\quad \left. + (1 - \lambda)(P_2 f(w_i^{-1}(x))) + \sum_{j=1}^{i-1} P_{2j} - f(x) \right| \\ &\leq \lambda |T_{P_1}(f(x)) - f(x)| + (1 - \lambda) |T_{P_2}(f(x)) - f(x)| \\ &\leq \lambda \sup_{x \in [0,1]} |T_{P_1}(f(x)) - f(x)| + (1 - \lambda) \sup_{x \in [0,1]} |T_{P_2}(f(x)) - f(x)| \\ &= \lambda D(P_1) + (1 - \lambda) D(P_2) \end{aligned}$$

De donde:

$$D(\lambda P_1 + (1 - \lambda)P_2) \leq \lambda D(P_1) + (1 - \lambda) D(P_2)$$

Por lo tanto, la función  $D$  es convexa. □

Estos dos últimos teoremas nos dicen que lo que hemos estado trabajando es un problema de optimización convexa, pues la función objetivo a minimizar  $D(P)$  es convexa y el conjunto solución  $C$ , de dicho problema, es convexo.

## 3.6 Control de inventario

El inventario es el conjunto de suministros, materias primas, productos en proceso y producto terminado, el mantener almacenados todos estos elementos representa un costo para la empresa. La reducción de este costo ha sido siempre una preocupación para los administradores, pero el problema de reducción del inventario implica no solamente la reducción física de las existencias en almacén, sino nuevas estrategias para administrar y controlar esas existencias.

Por otra parte, en la administración del inventario, los objetivos, las políticas y las decisiones que se tomen deben ser congruentes con los objetivos de mercadotecnia, financieros y de

fabricación.

En todo momento, las decisiones referentes al inventario están entrelazadas con las decisiones de planeación de capacidad, con las estructuras de planeación a largo, mediano o corto plazo, así como en las fases de ejecución y control de la administración de las operaciones.

El inventario se clasifica de acuerdo al estado de los insumos requeridos en el proceso de producción, de esta forma existen inventarios de:

- **Materias primas:** materiales utilizados para elaborar los componentes de los productos terminados.
- **Componentes:** partes o sub montajes que se encuentran listos para ir al ensamble final del producto.
- **Materiales en proceso:** aquellos sobre los que se efectúa un trabajo o que se encuentra entre una operación y otra.
- **Productos terminados:** los que se encuentran listos para ser embarcados a un cliente de acuerdo a un pedido.

Existen además otras dos clases de inventario: *los materiales que preceden a la producción y los materiales de mantenimiento*. Ambos requieren de inversiones significativas, son esenciales en la fabricación y pueden administrarse más eficazmente utilizando la información del sistema formal de planeación y control, así como de técnicas apropiadas para el control y reducción de inventarios [36].

### 3.6.1 Planteamiento del problema

Un comerciante se dedica a la venta de un único producto. Para los períodos  $1, 2, \dots, N$ , se consideran las siguientes variables [4]:

$x(k - 1)$  es el stock disponible al principio del período  $k$ .

$u(k - 1)$  es el stock pedido (e inmediatamente suministrado), al comienzo del período  $k$ .

$v(k - 1)$  es la demanda durante el período  $k$ , con distribución de probabilidad dada.

Por otra parte,

$h$  es el coste de tener en stock una unidad no vendida al final del período  $k$ .

$c$  es el coste de cada unidad de stock pedida.

$p$  es el coste por unidad de demanda no atendida por no tener stock disponible.

$B$  es la capacidad de almacén.

Se supone que  $v(0), v(1), \dots, v(N-1)$  son variables aleatorias independientes y que el exceso de demanda no se pierde sino que es atendido en cuanto hay stock adicional disponible (es decir, que puede haber stock negativo).

Formular el problema de optimización dinámica estocástica que tiene ante sí el comerciante, si quiere minimizar costes es el siguiente [4]:

$$\begin{aligned} \min \quad & E\left\{\sum_{k=0}^{N-1} [c \times u(k) + p \times \max(0, v(k) - x(k) - u(k)) + h \times \max(0, x(k) + u(k) - v(k))]\right\} \\ \text{sujeto a:} \quad & x(k+1) = x(k) + u(k) - v(k) \\ \text{con:} \quad & x(0) = x_0 \\ & 0 \leq u(k) \leq B - x(k) \quad k = 0, 1, \dots, N-1 \end{aligned}$$

Ahora bien [37],

$$\begin{aligned} & E\left\{\sum_{k=0}^{N-1} \left[ c \times u(k) + p \times \max(0, v(k) - x(k) - u(k)) + h \times \max(0, x(k) + u(k) - v(k)) \right]\right\} \\ & = c \times u(k) + \sum_{k=0}^{N-1} p \times \max(0, v(k) - x(k) - u(k)) \cdot Pr(x_k) \\ & + \sum_{k=0}^{N-1} h \times \max(0, x(k) + u(k) - v(k)) \cdot Pr(x_k) \end{aligned} \tag{3.32}$$

donde  $Pr(x_k)$  es la probabilidad de la demanda en  $k$  sea igual a  $x_k$ . Considerando que  $\sum_{k=0}^{N-1} Pr(x_k) = 1$ .

Sea

$$\Delta E[J_\pi v(k)] = E[J_\pi v(k+1)] - E[J_\pi v(k)] \tag{3.33}$$

Entonces,  $\Delta E[J_\pi v(k)]$  es el cambio en el costo total esperado cuando cambiamos de  $k$  a  $k+1$ .

Para una **función de costo convexa** la mejor cantidad de demanda  $Q$  será la más baja donde  $\Delta E[J_\pi v(k)]$  es mayor o igual a cero. Por lo tanto, seleccionamos  $k$  la más pequeña para la cual

$$\Delta E[J_\pi v(k)] \geq 0 \tag{3.34}$$

La ecuación anterior es válida si

$$E[J_\pi v(k+1)] - E[J_\pi v(k)] \geq 0 \tag{3.35}$$

sustituyendo la ecuación (3.32) en la ecuación (3.35) nos lleva a

$$\begin{aligned}
c \times u(k+1) &+ \sum_{k=0}^N p \times \max(0, v(k+1) - x(k+1) - u(k+1)) \cdot Pr(x_k) \\
&+ \sum_{k=N+1}^{\infty} h \times \max(0, x(k+1) + u(k+1) - v(k+1)) \cdot Pr(x_k) \tag{3.36}
\end{aligned}$$

$$\begin{aligned}
&- \left[ c \times u(k+1) + \sum_{k=0}^{N-1} p \times \max(0, v(k+1) - x(k+1) - u(k+1)) \cdot Pr(x_k) \right. \\
&+ \left. \sum_{k=N}^{\infty} h \times \max(0, x(k+1) + u(k+1) - v(k+1)) \cdot Pr(x_k) \right] \\
&\geq 0 \tag{3.37}
\end{aligned}$$

$$\begin{aligned}
c &+ \sum_{k=0}^N p \times Pr(x_k) - \sum_{k=N+1}^{\infty} h \times Pr(x_k) \\
&\geq 0
\end{aligned}$$

$$Pr(x(Q)) \geq \frac{h-c}{h+p}$$

donde  $Pr(x(k))$  es la probabilidad que la demandda  $v$  es menor o igual a la cantidad de la orden  $k$ . El costo total esperado  $E[J_{\pi}v(k)]$  será minimizado por el valor más pequeño de  $k$  (llamado  $Q^*$ ) que satisfaga la ecuación anterior.

### 3.7 Sistemas de Control en Red (NCS)

Los recientes avances tecnológicos en comunicación y sistemas de control digital han mostrado un progreso notable en los últimos años, lo que impulsa la aparición y el desarrollo de sistemas de control en red (NCS.: Networked Control System). Los NCS son sistemas de control de retroalimentación completamente distribuidos, en los que se cierra un circuito de control a través de una red de comunicación para conectar los sensores, el controlador y los actuadores. En comparación con los sistemas tradicionales de control punto a punto, los NCS reducen el cableado, facilitan el diagnóstico, el mantenimiento del sistema y aumentan la flexibilidad y la fiabilidad del sistema. Por lo tanto, los NCS son cada vez más populares en diversos contextos como: automóviles, aviones, plantas de fabricación y cirugía remota [35].

Sin embargo, el ancho de banda finito y el servicio limitado en las redes de comunicación causan algunos problemas y desafíos nuevos, tales como retrasos aleatorios, pérdidas de paquetes, transmisión multipack y desorden de paquetes. En general, las pérdidas de paquetes significan que el retraso de la transmisión de paquetes a través de las redes es infinito. Por lo tanto, los retrasos aleatorios son el principal problema y desafío en NCS, y la aleatoriedad

de los retrasos generalmente se ve afectada por muchos factores estocásticos (por ejemplo, carga de red, competencia de nodos y congestión de red).

El retraso de red tiende a ser estocástico ya que se ve afectado por muchos factores estocásticos (por ejemplo, carga de red, competencia de nodos y congestión de red). El modelo de retardo estocástico se puede dividir en dos categorías: *el modelo en el que los retrasos son mutuamente independientes* y *el modelo en el que los retrasos son probabilísticamente dependiente*.

Cuando se desconoce la dependencia probabilística, el modelo de retardo estocástico mutuamente independiente se aplica a menudo al modelado y control de NCS con retardos aleatorios. Se pueden estudiar tres estrategias principales de control (que incluyen control estocástico, control robusto y control predictivo) [35].

### 3.7.1 Formulación del problema

A la hora de diseñar un controlador en tiempo discreto es frecuente partir de un controlador continuo obtenido mediante los métodos clásicos y discretizar éste último, en lugar de realizar el diseño directamente en tiempo discreto.

Se va a considerar los sistemas de control en red ilustrados en la Figura 3.4.

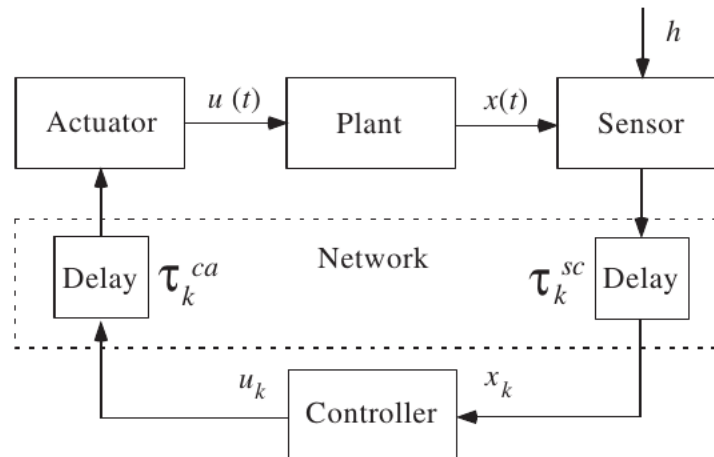


Figura 3.4: Sistemas de control de red con retardo

La comunicación con retardo entre el sensor y el controlador se representa como  $\tau_k^{sc}$ , entre el actuador y controlador como  $\tau_k^{ca}$ . Estos varían al azar. Todos los retardos son independientes en todo el horizonte y sus distribuciones de probabilidad son conocidas a priori. El período de muestreo, que es un valor positivo se denota por  $h$ . La longitud de los retrasos pasados son conocidos por el controlador. La planta que se va a controlar está representada por el sistema:

$$\dot{x}(t) = Ax(t) + Bu(t) + B_w w(t) \quad (3.38)$$

donde  $x \in \mathbb{R}^m$  es el estado,  $u \in \mathbb{R}^p$  es la entrada y  $w$  es el ruido blanco con varianza incremental unitaria  $R_w$ .

La sincronización de las señales en el sistema de control se ilustra en la Figura 3.5. Ésta muestra los retardos  $\tau_k^{sc}$ ,  $\tau_k^{ca}$ . El paso de tiempo es denotado por  $t_k$  y el estado discretizado por  $x_k$ .

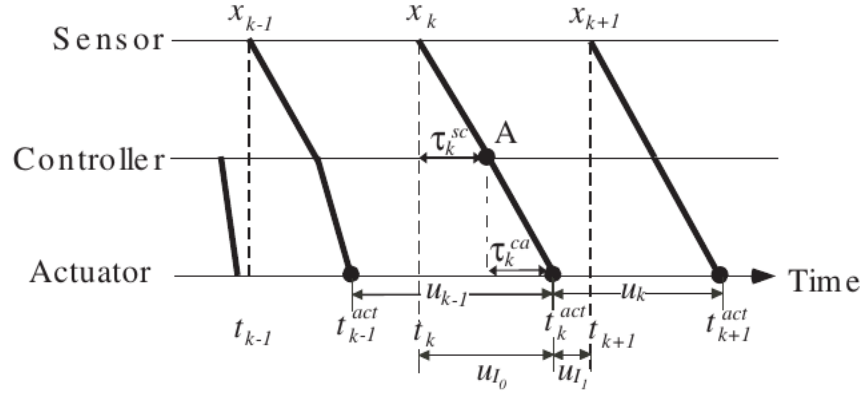


Figura 3.5: Diagrama de señales

El paso de tiempo se denota por  $t_k$  y el estado discretizado se denota por  $x_k$ . Se asume que el retardo total satisface la siguiente condición 3.39,

$$\tau_k^{sc} + \tau_k^{ca} < Zh \quad (3.39)$$

donde  $Z$  es un entero positivo. Usando esta suposición, se puede considerar que el retardo total es más largo que el período de muestreo considerado. El proceso lineal con el actuador de retención de orden cero en la entrada, muestreado periódicamente

$$x_{k+1} = \tilde{A}x_k + \tilde{\Gamma}(\tau_k) \begin{bmatrix} u_{k-Z} \\ \vdots \\ u_k \end{bmatrix} + \tilde{B}_w w_k \quad (3.40)$$

donde

$$\tilde{A} = e^{Ah}, \quad \tilde{B} = \int_0^h e^{As} ds B_w \quad (3.41)$$

$$\tau_k = \{\tau_{k-Z}^{sc} \tau_{k-Z}^{ca} \cdots \tau_k^{sc} \tau_k^{ca}\} \quad (3.42)$$

$$\tilde{\Gamma}(\tau_k) = \sum_{i=0}^{q_k} \Phi(t_{k+1} - t_k^{i+1}, t_{k+1} - t_k^i, I_i + Z - k) \quad (3.43)$$

### 3.7.2 Sistemas de operador delta

Al usar la representación del operador delta, el sistema es calculable incluso si el período de muestreo  $h$  es pequeño. A partir de la relación del operador delta, el sistema (3.40) puede transformarse en

$$\begin{aligned} \frac{\Delta x_k}{\Delta h} &= \frac{x_{k+1} - x_k}{h} \\ &= A^* x_k + \Gamma(\tau_k) \begin{bmatrix} u_{k-Z} \\ \vdots \\ u_k \end{bmatrix} + B_w^* w_k \end{aligned} \quad (3.44)$$

donde

$$A^* = \frac{1}{h}(\tilde{A} - I), \quad \Gamma(\tau_k) = \frac{1}{h}\tilde{\Gamma}(\tau_k), \quad B_w^* = \frac{1}{h}\tilde{B}_w \quad (3.45)$$

introduciendo un vector  $v_k$

$$v_k = [u_{k-Z+1}^t \cdots u_{k-1}^t]^t \quad (3.46)$$

el sistema (3.44) se escribe de la siguiente forma:

$$\frac{\Delta x_k}{\Delta h} = [A^* \quad \Gamma_Z(\tau_k) \quad \Gamma_a(\tau_k)] \begin{bmatrix} x_k \\ u_{k-Z} \\ v_k \end{bmatrix} + \Gamma_0(\tau_k) u_k + B_w^* w_k \quad (3.47)$$

aquí

$$\Gamma_Z \in \mathbb{R}^{m \times p}, \quad \Gamma_a \in \mathbb{R}^{m \times (Z-1)p}, \quad \Gamma_0 \in \mathbb{R}^{m \times p} \quad (3.48)$$

Nótese que las matrices  $[A^* \quad \Gamma_Z(\tau_k) \quad \Gamma_a(\tau_k)]$  y  $\Gamma_0(\tau_k)$  dependen del retardo de tiempo  $\tau_k$ .

Cuando el período de muestreo  $h$  es muy pequeño, la matriz (3.41) puede ser una matriz unitaria. Mientras que mediante el uso del operador delta se hace posible el cálculo numérico que implica que (3.45) no es una matriz de unitaria.

### 3.7.3 formulación del problema en tiempo discreto

Las condiciones del problema para el sistema (3.47) está dada por la función de costo:

$$J = E \left\{ h \sum_{k=0}^N \begin{bmatrix} x_k \\ u_k \end{bmatrix} W \begin{bmatrix} x_k \\ u_k \end{bmatrix} + x_{N+1}^t Q_f x_{N+1} \right\} \quad (3.49)$$

donde  $W$  es semidefinida-positiva y denotada como

$$w = \begin{bmatrix} Q & S \\ S^t & R \end{bmatrix} \quad (3.50)$$



Sea asume que  $R$  es definida positiva y  $Q_f$  es semidefinida positiva.

Falta demostrar que la función de costo  $J^*$  es convexa para  $k \in \{0, 1, \dots, N-1\}$ .

Observe que  $J^*(x_N)$  es convexa porque es cuadrática. Supongamos que para  $k$ ,  $J_k^*(x_k)$  es convexa. Pues  $x_k = Ax_{k-1} + Bu_{k-1} + B_w w_{k-1}$  es un mapeo afín de  $(x_{k-1}, u_{k-1})$  a  $x_k$  para  $w_{k-1}$  fijo, la función  $J_k^*(Ax_{k-1} + Bu_{k-1} + B_w w_{k-1})$  es conjuntamente convexa en  $(x_{k-1}, u_{k-1})$  para  $w_{k-1}$  fijo.

Sea  $\mathcal{U}_k(x) = \{u \in \mathcal{U} : Ax + Bu + B_w w \in \mathcal{X}_{k+1} \forall w \in D_k\}$ . Ahora se mostrará que  $\mathcal{U}_k(x)$  es un mapeo convexo. Sea  $x_1, x_2$  dos estados iniciales y sean  $u_1 \in \mathcal{U}_k(x_1)$  y  $u_2 \in \mathcal{U}_k(x_2)$  debemos mostrar que  $\lambda u_1 + (1 - \lambda)u_2 \in \mathcal{U}_k(\lambda x_1 + (1 - \lambda)x_2)$  para cualquier  $w \in D_k$ , se tiene:

$$A(\lambda x_1 + (1 - \lambda)x_2) + B(\lambda u_1 + (1 - \lambda)u_2) + B_w w_k = \lambda(Ax_1 + Bu_1 + B_w w_k) + (1 - \lambda)(Ax_2 + Bu_2 + B_w w_k)$$

como  $\mathcal{X}_{k+1}$  es convexo, la ecuación anterior muestra que  $\lambda x_1 + (1 - \lambda)x_2 \in \mathcal{X}_{k+1}$  lo que implica que  $\lambda u_1 + (1 - \lambda)u_2 \in \mathcal{U}_k(\lambda x_1 + (1 - \lambda)x_2)$ .

# Capítulo 4

## Métodos para optimización no diferenciable

En este capítulo se describirán los métodos estudiados para la minimización de funciones que son convexas pero no diferenciables. De igual forma se darán las pruebas de la convergencia de cada método [7].

### 4.1 Método del subgradiente

Este es el método más sencillo y es muy parecido al método de gradiente para funciones diferenciables pero cuenta con algunas modificaciones [7].

1. El método de subgradiente se aplica directamente a la función no diferenciable.
2. Los tamaños de paso no son elegidos por búsqueda lineal. En la mayoría de los casos, el tamaño de paso es fijo.
3. A diferencia del método del gradiente, el método de subgradiente no es un método de descenso; el valor de la función puede incrementar.

En este caso solo se tratará el problema de minimizar una función sin restricciones, es decir, minimizar  $f : \mathbb{R}^n \mapsto (0, +\infty]$  donde  $f$  es convexa.

El método de subgradiente usa la iteración:

$$x^{(k+1)} = x^k - \alpha_k g^k$$

$x^k$  : es el punto resultante de la  $k$ -ésima iteración.

$g^k$  : es cualquier subgradiente de  $f$  en  $x^k$ .

$\alpha_k > 0$  : es el  $k$ -ésimo tamaño de paso.

Así, en cada iteración del método de subgradiente, se hace el paso en la dirección del subgradiente negativo.

Cuando la función es diferenciable, la única elección posible de  $g^k$  es  $\nabla f(x^k)$ , y el método del subgradiente se reduce al método de gradiente (excepto por la elección del tamaño de paso). Un problema en el método de subgradiente es que se toma cualquier subgradiente en cada iteración, por lo tanto, no se toma en cuenta cual es la elección de subgradiente que hace decrecer el valor de la función en el nuevo punto.

Puede ocurrir que  $-g^k$  no sea una dirección de descenso para  $f$  en  $x^k$ , es decir, que no se cumple necesariamente

$$\langle -g^k, \xi^k \rangle \leq 0 \quad \forall \xi \in \partial f(x^k)$$

En el método del subgradiente se utilizan varios tamaños de paso, los cuales influyen en la convergencia del método. Los tamaños de paso más usuales son los siguientes [7]:

1. **Tamaño de paso constante.** Se toma  $\alpha_k = \alpha$  una constante positiva que es independiente de  $k$ .
2. **Longitud de paso constante.** Se toma  $\alpha_k = \gamma / \|g^k\|_2$ . esto implica que  $\|x^{k+1} - x^k\|_2 = \gamma$
3. **Cuadrados sumables.** En este caso se toma  $\alpha_k$  de tal manera que

$$\alpha_k \geq 0, \quad \sum_{k=1}^{\infty} \alpha_k^2 < \infty, \quad \sum_{k=1}^{\infty} \alpha_k = \infty$$

4. **Paso decreciente.** Se toma  $\alpha_k$  tal que:

$$\alpha_k \geq 0, \quad \lim_{k \rightarrow \infty} \alpha_k = 0, \quad \sum_{k=1}^{\infty} \alpha_k = \infty$$

El método del subgradiente es un procedimiento iterativo que puede ser usado para resolver el problema de maximizar (minimizar) una función cóncava (convexa) no necesariamente diferenciable, el método se define como sigue [30]:

1. Iniciamos con un punto  $x_0 \in \mathbb{R}^n$  y elegimos una sucesión de pasos dados por números positivos  $\{\alpha_k\}_{k=0}^{\infty}$
2. Construimos la sucesión  $x_{k+1} = x_k - \alpha_k s_k$ , con  $s_k \in \partial f(x_k)$ , que eventualmente converge a la solución óptima del problema.  
Calculamos un subgradiente  $s_k \in \partial f(x_{k+1})$ , si  $s_{k+1} = 0$  paramos,  $x_{k+1}$  es el punto óptimo.  
Caso contrario.
3. Proseguimos con el paso 2 hasta cumplirse una condición de parada establecida.

### 4.1.1 Convergencia del subgradiente

Dado que el método el subgradiente no es un método de descenso se hará en cada iteración lo siguiente

$$f_{best}^k = \min\{f_{best}^{k-1}, f(x^k)\}$$

y se hace  $i_{best}^k = k$  si  $f(x^k) = f_{best}^k$ , es decir, se guardará el mejor punto obtenido hasta la iteración  $k$ .

Ahora, para probar la convergencia del método se supondrá que existe un mínimo de la función  $f$ , digamos  $x^*$ . También se supondrá lo siguiente:

1. La norma de los subgradientes es acotada, es decir, existe  $G \in \mathbb{R}$  tal que  $\|g^k\| \leq G$  para toda  $k$ .
2. Se conoce un número  $R$  tal que  $R \geq \|x^1 - x^*\|_2$ .

Teniendo en cuenta lo anterior y que  $x^*$  es cualquier punto óptimo entonces

$$\begin{aligned} \|x^{k+1} - x^*\|_2^2 &= \|x^k - \alpha_k g^k - x^*\|_2^2 \\ &= \|x^k - x^*\|_2^2 - 2\alpha_k g^{(k)t}(x^k - x^*) + \alpha_k^2 \|g^k\|_2^2 \\ &\leq \|x^k - x^*\|_2^2 - 2\alpha_k (f(x^k) - f(x^*)) + \alpha_k^2 \|g^k\|_2^2 \end{aligned} \quad (4.1)$$

donde la última desigualdad se cumple por la definición de subgradiente, es decir,  $g^k$  cumple que

$$f(x^*) \geq f(x^k) + \langle g^k, x^* - x^k \rangle$$

Aplicando la desigualdad (4.1) recursivamente se obtiene que

$$\|x^{k+1} - x^*\|_2^2 \leq \|x^1 - x^*\|_2^2 - 2 \sum_{i=1}^k \alpha_i (f(x^i) - f(x^*)) + \sum_{i=1}^k \alpha_i^2 \|g^i\|_2^2$$

Usando el hecho de que  $\|x^{k+1} - x^*\|_2^2 \geq 0$  y  $\|x^1 - x^*\|_2^2 \leq R$  se tiene que

$$2 \sum_{i=1}^k \alpha_i (f(x^i) - f(x^*)) \leq R^2 + \sum_{i=1}^k \alpha_i \|g^i\|_2^2 \quad (4.2)$$

por otro lado,

$$\begin{aligned} \sum_{i=1}^k \alpha_i (f(x^i) - f(x^*)) &\geq \left( \sum_{i=1}^k \alpha_i \right) \min_i (f(x^i) - f(x^*)) + \sum_{i=1}^k \alpha_i \|g^i\|_2^2 \\ &= \left( \sum_{i=1}^k \alpha_i \right) (f_{best}^k - f(x^*)) \end{aligned}$$

combinando esta desigualdad con (4.2) se obtiene que

$$f_{best}^k - f(x^*) = \min_i (f(x^i) - f(x^*)) \leq \frac{R^2 + \sum_{i=1}^k \alpha_i^2 \|g^i\|_2^2}{2 \sum_{i=1}^k \alpha_i} \quad (4.3)$$

Finalmente, usando la hipótesis de que  $\|g^k\|_2 \leq G$ , se consigue la siguiente desigualdad

$$f_{best}^k - f(x^*) \leq \frac{R^2 + G^2 \sum_{i=1}^k \alpha_i}{2 \sum_{i=1}^k \alpha_i} \quad (4.4)$$

De esta desigualdad se desprenden varios resultados de convergencia, dependiendo del tamaño de paso que se esté usando se tendrá una mejor aproximación al punto óptimo.

1. *Tamaño de paso constante.* Cuando  $\alpha_k = \alpha$ , se tiene usando (4.4)

$$f_{best}^k - f(x^*) \leq \frac{R^2 + G^2 \alpha^2 k}{2 \alpha k}$$

Cuando  $k \mapsto \infty$  entonces, el lado derecho converge a  $G^2 \alpha / 2$ .

2. *Longitud de paso constante.* Usando  $\alpha_k = \gamma / \|g^k\|_2$  y la desigualdad (4.3) se obtiene

$$f_{best}^k - f(x^*) \leq \frac{R^2 + \gamma k}{2 \sum_{i=1}^k \alpha_i} \leq \frac{R^2 + \gamma k}{2 \gamma k / G}$$

Usando  $\alpha_i \geq \gamma / G$ . cuando  $k \mapsto \infty$  el lado derecho converge a  $G \gamma / 2$ .

3. *Cuadrados sumables.* Ahora supongamos que

$$\|\alpha\|_2^2 = \sum_{i=1}^{\infty} \alpha_k^2 < \infty, \quad \sum_{i=1}^{\infty} \alpha_i = \infty$$

Entonces se tiene que

$$f_{best}^k - f(x^*) \leq \frac{R^2 + G^2 \|\alpha\|_2^2}{2 \sum_{i=1}^{\infty} \alpha_i}$$

Cuando  $k \mapsto \infty$  el lado derecho converge a cero. Así el método del subgradiente converge.

4. *Paso decreciente.* Si la sucesión  $\alpha_k$  converge a cero y  $\sum \alpha_i = \infty$ , entonces el lado derecho de la desigualdad (4.4) converge a cero, lo cual implica que el método del subgradiente converge. Para probar esto, sea  $\epsilon > 0$ . Entonces existe un entero  $N_1$  tal que  $\alpha_i < \epsilon/G^2$  para  $i > N_1$ . También existe un natural  $N_2$  tal que

$$\sum_{i=1}^{N_2} \alpha_i > \frac{1}{\epsilon} (R^2 + G^2 \sum_{i=1}^{N_2-1} \alpha_i^2)$$

Ahora, sea  $N = \max\{N_1, N_2\}$ , entonces para  $k > N$  se tiene que

$$\begin{aligned} \frac{R^2 + G^2 \sum_{i=1}^k \alpha_i^2}{a \sum_{i=1}^k \alpha_i} &\leq \frac{R^2 + G^2 \sum_{i=1}^{N_1} \alpha_i^2}{2 \sum_{i=1}^k \alpha_i} + \frac{G^2 \sum_{i=N_1+1}^k \alpha_i^2}{2 \sum_{i=1}^{N_1} \alpha_i + 2 \sum_{i=N_1+1}^k \alpha_i} \\ &\leq \frac{R^2 + G^2 \sum_{i=1}^{N_1} \alpha_i^2}{\frac{2}{\epsilon} (R^2 + G^2 \sum_{i=1}^{N_1} \alpha_i^2)} + \frac{G^2 \sum_{i=N_1+1}^k (\epsilon \alpha_i / G^2)}{2 \sum_{i=N_1+1}^k \alpha_i} \\ &= \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

Así el método del subgradiente converge.

A partir de la ecuación (4.2) se puede obtener un criterio de paro para nuestro método. De nuevo, suponiendo que  $R \geq \|x^1 - x^*\|_2$  y usando la desigualdad (4.2) se tiene que

$$f(x^*) \geq \frac{2 \sum_{i=1}^k \alpha_i f(x^i) - R^2 - \sum_{i=1}^k \alpha_i^2 \|g^k\|_2^2}{2 \sum_{i=1}^k \alpha_i} = l_k \quad (4.5)$$

donde  $l_k$  puede ser calculado a partir de la  $k$ -ésima iteración. La iteración  $l_k$  no es necesariamente creciente, por lo que se puede guardar la mejor cota superior de  $\{l_k\}$ .

$$l_{best}^k = \max\{l_1, \dots, l_k\}$$

## 4.2 Método de planos de corte

Existen métodos en los cuales toda la información previa es usada para obtener una nueva iteración. Esta información puede ser utilizada para crear un modelo de la función  $f$ , de tal manera que se obtenga una aproximación lineal al problema original. Esto es usado en los métodos de planos de corte [7].

Al tomar en cuenta la información acumulada en las iteraciones anteriores  $\{y^i, f(y^i), s^i \in \partial f(y^i)\}_{i=1}^l$  se puede construir la siguiente aproximación lineal a trozos de  $f$

$$\hat{f}_l(y) = \max_{1, \dots, l} f(y^i) + \langle s^i, y - y^i \rangle$$

Por construcción,  $\hat{f}_l(y) \leq f(y)$  para toda  $y \in \mathbb{R}^n$  y también se tiene que  $\hat{f}_l(y) \leq \hat{f}_{l+1}(y)$  para toda  $y \in \mathbb{R}^n$ . Así se modela la función a minimizar con aproximaciones lineales.

Para que el método esté bien definido, se necesita especificar un conjunto compacto  $C$  y restringir nuestro modelo a  $C$  para asegurar que cada iteración esté bien definida.

El algoritmo del método de planos de corte es el siguiente:

1. Sea  $\bar{\delta} > 0$  y  $C$  un conjunto compacto que contiene el punto mínimo de  $f$ . Sea  $k = 1$ ,  $y^1 \in C$  y  $f_0 = -\infty$ .
2. Calcular  $f(y^k)$  y el subgradiente  $s^k \in \partial f(y^k)$ .
3. Definir  $\delta_k = f(y^k) - \hat{f}_{k-1}(y^k) \geq 0$ .
4. Si  $\delta_k < \hat{\delta}$ .
5. Actualizar el modelo  $\hat{f}_k(y) = \max\{\hat{f}_{k-1}(y), f(y^k) + \langle s^k, y - y^k \rangle\}$ .
6. Calcular  $y^{k+1} \in \arg \min_{y \in C} \hat{f}_k(y)$ .
7. Hacer  $k = k + 1$  e ir al paso 2.

### 4.2.1 Convergencia del método planos de corte

Hasta donde sabemos, la convergencia de este algoritmo en un número finito de pasos no está teóricamente garantizado, lo que podría haber desalentado su implementación práctica hasta la fecha. Sin embargo, hay muchas variantes de este algoritmo propuesto en la literatura técnica que hace que este método sea computacionalmente competitivo y práctico en muchos casos.

Entre las variantes de éste se consideran

- *Algoritmo centro de gravedad (CG)*.  $x^k$  es el centro de gravedad de  $\mathcal{P}_{k-1}$ .
- *Método de plano de corte de elipsoide de volumen máximo (MVE)*.  $x^k$  es el centro del máximo volumen elipsoidal contenido en  $\mathcal{P}_{k-1}$ .
- *Método de plano de corte del centro de Chebyshev*.  $x^k$  es el centro de la bola más larga contenida en  $\mathcal{P}_{k-1}$ .
- *Método de plano de corte del centro analítico (ACCPM)*.  $x^*$  es el centro analítico de (definido por desigualdades)  $\mathcal{P}_{k-1}$ .

A comparación del método del subgradiente, el método de planos de corte brinda un criterio de paro basado en  $\delta_k$  el cual no existe en el método del subgradiente.

Ahora bien, por construcción el modelo satisface que

$$\hat{f}_k(y) \leq f(y)$$

y de aquí se sigue que

$$\min_y \hat{f}_k(y) \leq \min_y f(y)$$

Entonces, si el criterio de paro se satisface se tiene para alguna  $k$  que  $f(y^k) - \hat{y}_{k-1}(y^k) < \bar{\delta}$ , lo cual implica que

$$f(y^k) < \bar{\delta} + \hat{f}_{k-1}(y^k) = \bar{\delta} + \min_y \hat{f}_{k-1}(y) \leq \bar{\delta} + \min_y f(y)$$

### 4.3 Método de haces

Utilizando como motivación el algoritmo anterior, se construye un nuevo método utilizando la regularización de Moreau-Yosida, cuyas propiedades probadas a continuación son importantes para este trabajo.

#### 4.3.1 Regularización de Moreau Yosida

Sea  $f : \mathbb{R}^n \mapsto (-\infty, +\infty]$  una función propia y convexa, además supongamos que  $f$  alcanza su mínimo. Se define la **regularización de Moreau Yosida** de  $f$  para  $\lambda > 0$  como sigue

$$F_\lambda(v) = \min_{x \in \mathbb{R}^n} \left\{ f(x) + \frac{1}{2\lambda} \|x - v\|^2 \right\} \quad (4.6)$$

El punto  $p(v) \in \mathbb{R}^n$  que cumple  $f(p(v)) + \frac{1}{2\lambda} \|p(v) - v\|^2 = F_\lambda(v)$  está bien definido dado que la función que se está minimizando es estrictamente convexa y  $f$  alcanza su mínimo.

**Teorema 4.1.** *Sea  $f : \mathbb{R}^n \mapsto (-\infty, +\infty]$  una función propia, cerrada y convexa, entonces, la función  $F_\lambda(v)$  es cerrada y estrictamente convexa.*

***Demostración.***

Dado que  $f$  es cerrada y la norma es una función continua, entonces se sigue inmediatamente que  $F_\lambda(v)$  es cerrada. Ahora, por la convexidad de  $f$  y dado que la función norma es estrictamente convexa entonces  $F_\lambda$  es estrictamente convexa.

□



**Proposición 4.1.** Si  $f : \mathbb{R}^n \mapsto (-\infty, +\infty]$  es una función propia, cerrada y convexa. Entonces  $F_\lambda$  es finita y diferenciable en cualquier parte con gradiente dado por

$$\nabla F_\lambda(v) = \frac{1}{\lambda}(v - p(v))$$

Más aún,

$$\|\nabla F_\lambda(v) - \nabla F_\lambda(v')\|^2 \leq \frac{1}{\lambda} \langle \nabla F_\lambda(v) - \nabla F_\lambda(v'), v - v' \rangle$$

para toda  $v, v' \in \mathbb{R}^n$ . Además

$$\|\nabla F_\lambda(v) - \nabla F_\lambda(v')\| \leq \frac{1}{\lambda} \|v - v'\|$$

para toda  $v, v' \in \mathbb{R}^n$ .

### **Demostración**

(Finita). Si  $f$  es una función propia, existe  $x \in \mathbb{R}^n$  tal que  $f(x) < \infty$ . Así  $F_\lambda$  es finita dado que

$$F_\lambda(v) \leq f(x) + \frac{1}{2\lambda} \|x - v\|^2 < \infty$$

(Diferenciabilidad). Sea  $d \in \mathbb{R}^n$  unitario,  $v \in \mathbb{R}^n$  y  $t > 0$ . Por definición de  $F_\lambda$

$$\begin{aligned} \frac{F_\lambda(v + td) - F_\lambda(v)}{t} &= \frac{\min_x [f(x) + \frac{1}{2\lambda} \|x - v - td\|^2] - \min_w [f(w) + \frac{1}{2\lambda} \|w - v\|^2]}{t} \\ &\geq \frac{[f(p(v + td)) + \frac{1}{2\lambda} \|p(v + td) - v - td\|^2]}{t} - \frac{[f(p(v + td)) + \frac{1}{2\lambda} \|p(v + td) - v\|^2]}{t} \\ &= \frac{\frac{1}{2\lambda} (\|p(v + td) - v - td\|^2 - \|p(v + td) - v\|^2)}{t} \\ &= \frac{\frac{1}{2\lambda} (\|p(v + td) - p(v) + p(v) - v - td\|^2 - \|p(v + td) - p(v) + p(v) - v\|^2)}{t} \\ &= \frac{\frac{1}{2\lambda} (\|p(v) - v - td\|^2 - \|p(v) - v\|^2)}{t} - \frac{1}{\lambda} \langle p(v + td) - p(v), d \rangle \end{aligned}$$

donde se usó el hecho de que  $F_\lambda(v) \leq f(p(v + td)) + \frac{1}{2\lambda} \|p(v + td) - v\|^2$ . Ahora, al tomar el límite cuando  $t \mapsto 0$  se obtiene

$$\lim_{t \rightarrow 0} \frac{F_\lambda(v + td) - F_\lambda(v)}{t} \geq \frac{1}{\lambda} \langle v - p(v), d \rangle$$

Por otro lado, dado que  $F_\lambda(v + td) \leq f(p(v)) + \frac{1}{2\lambda} \|p(v) - v - td\|^2$

$$\begin{aligned}
\frac{F_\lambda(v+td) - F_\lambda(v)}{t} &= \frac{\min_x [f(x) + \frac{1}{2\lambda} \|x - v - td\|^2] - \min_w [f(w) + \frac{1}{2\lambda} \|w - v\|^2]}{t} \\
&\leq \frac{[f(p(v)) + \frac{1}{2\lambda} \|p(v) - v - td\|^2] - [f(p(v)) + \frac{1}{2\lambda} \|p(v) - v\|^2]}{t} \\
&= \frac{1}{2\lambda} \frac{\|p(v) - v - td\|^2 - \|p(v) - v\|^2}{t}
\end{aligned}$$

De nuevo, al tomar  $t \mapsto 0$

$$\lim_{t \rightarrow 0} \frac{F_\lambda(v+td) - F_\lambda(v)}{t} \leq \frac{1}{\lambda} \langle v - p(v), d \rangle$$

Por lo tanto

$$\nabla F_\lambda(v) = \frac{1}{\lambda} (v - p(v))$$

Para probar las aseveraciones restantes se observa lo siguiente:

$$\begin{aligned}
\|\nabla F_\lambda(v) - \nabla F_\lambda(v')\|^2 &= \frac{1}{\lambda} \langle \nabla F_\lambda(v) - \nabla F_\lambda(v'), v - p(v) - v' + p(v') \rangle \\
&= \frac{1}{\lambda} \langle \nabla F_\lambda(v) - \nabla F_\lambda(v'), v - v' \rangle + \frac{1}{\lambda} \langle \nabla F_\lambda(v) - \nabla F_\lambda(v'), p(v') - p(v) \rangle
\end{aligned}$$

Se mostrará que el último término es negativo. Para hacer esto primero se verificará lo siguiente: sea  $\xi \in \partial f(v)$  y  $\xi' \in \partial f(v')$ , entonces  $\langle \xi - \xi', v - v' \rangle \geq 0$ . Para probar esto, se toman en cuenta las desigualdades siguientes, asociadas a  $\xi$  y  $\xi'$ , y posteriormente aplicadas a  $v'$  y  $v$  respectivamente

$$f(v') \geq f(v) + \langle \xi, v' - v \rangle$$

$$f(v) \geq f(v') + \langle \xi', v - v' \rangle$$

sumando estas dos desigualdades se sigue que

$$\langle \xi - \xi', v - v' \rangle \tag{4.7}$$

Por optimalidad de  $p(v)$  y de  $p(v')$  se tiene que

$$0 \in \partial(f(p(v)) + \frac{1}{2\lambda} \|p(v) - v\|^2)$$

y

$$0 \in \partial(f(p(v')) + \frac{1}{2\lambda} \|p(v') - v'\|^2)$$

Por lo tanto,  $\nabla F_\lambda(v) = \frac{1}{\lambda}(v - p(v)) \in \partial f(p(v))$  y  $\nabla F_\lambda(v') = \frac{1}{\lambda}(v' - p(v')) \in \partial f(p(v'))$  usando (4.7) se obtiene

$$\langle \nabla F_\lambda(v) - \nabla F_\lambda(v'), p(v') - p(v) \rangle \leq 0 \quad (4.8)$$

La prueba de última desigualdad es consecuencia de (4.8)

$$\begin{aligned} \|\nabla F_\lambda(v) - \nabla F_\lambda(v')\|^2 &\leq \frac{1}{\lambda} \langle \nabla F_\lambda(v) - \nabla F_\lambda(v'), v - v' \rangle \\ &\leq \frac{1}{\lambda} \|\nabla F_\lambda(v) - \nabla F_\lambda(v')\| \|v - v'\| \end{aligned}$$

Por lo tanto,  $\|\nabla F_\lambda(v) - \nabla F_\lambda(v')\| \leq \frac{1}{\lambda} \|v - v'\|$

□

La regularización de Moreau-Yosida es importante en nuestro trabajo por la proposición anterior y porque tiene los mismos mínimos que la función  $f$  como se muestra en la siguiente proposición.

**Proposición 4.2.** *Los siguientes enunciados son equivalentes*

1.  $v \in \arg \min\{f(x) \mid x \in \mathbb{R}^n\}$
2.  $v = p(v)$
3.  $\nabla F_\lambda(v) = 0$
4.  $v \in \arg \min\{F_\lambda(x) \mid x \in \mathbb{R}^n\}$
5.  $f(v) = f(p(v))$
6.  $f(v) = F_\lambda(v)$

***Demostración***

1  $\Rightarrow$  2 Dado que  $v \in \arg \min\{f(x) \mid x \in \mathbb{R}^n\}$  se tiene que

$$f(v) = f(v) + \frac{1}{2\lambda} \|v - v\|^2 \leq f(x) + \frac{1}{2\lambda} \|x - v\|^2$$

para toda  $x \in \mathbb{R}^n$ . Entonces,  $p(v) = v$

2  $\Rightarrow$  3 Supongamos que  $p(v) = v$ , entonces  $\nabla F_\lambda(v) = \frac{1}{\lambda}(v - p(v)) = 0$ .

3  $\Rightarrow$  4 Dado que  $F_\lambda$  es convexa, entonces si  $\nabla F_\lambda(v) = 0$  se debe tener que  $v \in \arg \min\{F_\lambda(x) \mid x \in \mathbb{R}^n\}$

4  $\Rightarrow$  5 Usando el inciso 2,  $v = p(v)$  y por lo tanto  $f(v) = f(p(v))$ .

5  $\Rightarrow$  6 Por definición de  $F_\lambda$  e hipótesis se tiene la siguiente desigualdad

$$\begin{aligned} f(v) &= f(p(v)) \\ &\leq f(p(v)) + \frac{1}{2\lambda} \|p(v) - v\|^2 \\ &= F_\lambda(v) \\ &\leq f(v) + \frac{1}{2\lambda} \|v - v\|^2 \end{aligned}$$

Así  $f(v) = F_\lambda(v)$ .

6  $\Rightarrow$  1 Dado que  $F_\lambda(v) = f(v)$  entonces  $p(v) = v$ . Así, por optimalidad de  $v$

$$0 \in \partial f(v) + \frac{1}{\lambda}(v - p(v)) = \partial f(v)$$

□

Tomando como base el método de planos de corte se construye el método de haces. Se empezará agregando un punto  $x^k$  al conjunto de información que se acumula conforme se realizan iteraciones. Se seguirán usando el modelo lineal y se obtendrá un nuevo punto en la iteración usando la regularización de Moreau-Yosida [7].

El algoritmo para el método de haces es el siguiente:

1. Se fija  $\hat{\delta} > 0$ ,  $m \in (0, 1)$ ,  $x^0$  (punto inicial),  $y^0 = x^0$  y  $k = 0$ . Se calcula  $f(x^0)$  y  $s^0 \in \partial f(x^0)$  y se define  $\hat{f}_0(y)$  como sigue

$$\hat{f}_0(y) = f(x^0) + \langle s^0, y - x^0 \rangle \quad (4.9)$$

es decir,  $\hat{f}_0(y)$  es un hiperplano soporte de  $f$  en  $x^0$ .

2. Se calcula la siguiente iteración

$$y^{k+1} \in \arg \min_{y \in \mathbb{R}^n} \hat{f}_k(y) + \frac{\mu_k}{2} \|y - x^k\|^2 \quad (4.10)$$

donde  $\mu_k > 0$  es una sucesión creciente tal que  $\mu \mapsto 0$  y  $\sum_{k=1}^{\infty} \mu_k = \infty$ .

3. Se define  $\delta_k$  como sigue

$$\delta_k = f(x^k) - [\hat{f}_k(y^{k+1}) + \frac{\mu_k}{2} \|y^{k+1} - x^k\|^2]$$

4. Si  $\delta_k \leq \hat{\delta}$  entonces el algoritmo termina.

5. Se calcula  $f(y^{k+1})$  y  $s^{k+1} \in \partial f(y^{k+1})$ .

6. Si  $f(x^k) - f(y^{k+1}) > m\delta_k$  entonces  $x^{k+1} = y^{k+1}$  (a esto se le llama paso serio), si no pasa entonces  $x^{k+1} = x^k$  (paso nulo).

7. Se actualiza el modelo

$$\hat{f}_{k+1}(y) = \max_k \{ \hat{f}_k(y), f(y^{k+1}) + \langle s^{k+1}, y - y^{k+1} \rangle \} \quad (4.11)$$

8.  $k = k + 1$  y se regresa al paso 2.

Esta es una versión básica del método de haces, otras implementaciones se pueden tomar en cuenta tales como búsqueda lineal, actualización del parámetro  $\mu_k$ , etc [7].

Ahora bien, antes de realizar las pruebas e convergencia será más conveniente en teoría y práctica trabajar con el dual del problema (4.10) como se verá a continuación. Primero, se reescribirá el modelo (4.11) de manera más conveniente.

Se define el error de linealización con centro en  $x^k$  como sigue

$$e_i = f(x^k) - [f(y^i) - \langle s^i, y - y^i \rangle]$$

para  $i = 1, \dots, K + 1$ . Utilizando el error de linealización, nuestro modelo (4.11) se vuelve de la siguiente forma

$$\begin{aligned} \hat{f}_k(y) &= \max_{i=1, \dots, k+1} \{ f(y^i) + \langle y - y^i \rangle \} \\ &= \max_{i=1, \dots, k+1} \{ f(x^k) - e_i - \langle s^i, x^k - y^i \rangle + \langle s^i, y - y^i \rangle \} \\ &= f(x^k) + \max_{i=1, \dots, k+1} \{ -e_i + \langle s^i, y - x^k \rangle \} \end{aligned} \quad (4.12)$$

en donde todo lo usado es conocido debido a la acumulación de información que se tiene.

Al considerar el problema (4.10), usando (4.11) se tiene que

$$\begin{aligned} \min_{y \in \mathbb{R}^n} \hat{f}_k(y) + \frac{\mu_k}{2} \|y - x^k\|^2 &= \min_{y, r} \{ r + \frac{\mu_k}{2} \|y - x^k\|^2 \} \\ r &\geq f(x^k) - e_i + \langle s^i, y - x^k \rangle \end{aligned} \quad (4.13)$$

para  $i = 1, \dots, k + 1$ . Introduciendo multiplicadores de Lagrange  $\alpha \in \mathbb{R}_+^{k=1}$ , el lagrangiano se escribe como sigue

$$\begin{aligned}
l(y, r, \lambda) &= r + \frac{\mu_k}{2} \|y - x^k\|^2 + \sum_{i=1}^{k+1} \alpha_i (f(x^k) - e_i + \langle s^i, y - x^k \rangle - r) \\
&= \left(1 - \sum_{i=1}^{k+1} \alpha_i\right) r + \frac{\mu_k}{2} \|y - x^k\|^2 + \sum_{i=1}^{k+1} \alpha_i (f(x^k) - e_i + \langle s^i, y - x^k \rangle - r)
\end{aligned}$$

así, por convexidad

$$\min_{y,r} \max_{\alpha} L(y, r, \lambda) = \max_{\alpha} \min_{y,r} L(y, r, \alpha) \quad (4.14)$$

Dado que ambos lados son finitos se debe tener que  $1 - \sum_{i=1}^{k+1} \alpha_i = 0$ . Así por condiciones de optimalidad de primer orden

$$\nabla L_y(r, \alpha) = \mu_k (y - x^k) + \sum_{i=1}^{k+1} \alpha_i s^i = 0$$

Entonces se tiene que

$$\mu_k (y - x^k) = - \left( \sum_{i=1}^{k+1} \alpha_i s^i \right) \quad (4.15)$$

Al denotar  $\Delta^{k+1} = \{\alpha \in \mathbb{R}_+^{k+1} \mid \sum_{i=1}^{k+1} \alpha_i = 1\}$ , combinando la ecuación (4.14) y (4.15), se obtiene que al resolver el problema (4.9) es equivalente a resolver

$$\begin{aligned}
&\max_{\alpha \in \Delta^{k+1}} \frac{\mu_k}{2} \frac{\| - \sum_{i=1}^{k+1} \alpha_i s^i \|^2}{\mu_k} + \sum_{i=1}^{k+1} \alpha_i \left( f(x^k) - e_i + \left\langle s^i, - \frac{\sum_{i=1}^{k+1} \alpha_i s^i}{\mu_k} \right\rangle \right) \\
&= f(x^k) + \max_{\alpha \in \Delta^{k+1}} \frac{1}{2\mu_k} \left\| \sum_{i=1}^{k+1} \alpha_i s^i \right\|^2 - \sum_{i=1}^{k+1} \alpha_i e_i
\end{aligned} \quad (4.16)$$

**Definición 4.1.** Sea  $\alpha \in \Delta^{k+1}$  una solución óptima del problema (4.16) en la iteración  $k$ , se define el subgradiente agregado y el error de linealización agregado respectivamente como sigue

$$\hat{s}^k = \sum_{i=1}^{k+1} \alpha_i s^i \quad y \quad \hat{e}_k = \sum_{i=1}^{k+1} \alpha_i e_i \quad (4.17)$$

**Lema 4.1.** Sea  $\alpha \in \Delta^{k+1}$  una solución para el problema (4.16), entonces

1.  $\hat{s}^k \in \partial \hat{f}_k(y^{k+1})$ .
2.  $\hat{f}_k(y^{k+1}) = f(x^k) - \frac{1}{\mu_k} \|\hat{s}^k\|^2 - \hat{e}_k$ .
3.  $\delta_k = \frac{1}{2\mu_k} \|\hat{s}^k\|^2 + \hat{e}_k$ .

**Demostración.**

1. De la ecuación (4.15)  $-\mu_k(y^{k+1} - x^k) = \hat{s}^k$ . Así, dado que  $y^{k+1}$  es óptimo para (4.9)

$$0 \in \partial \hat{f}_k(y^{k+1}) + \mu_k(y^{k+1} - x^k)$$

de donde  $\mu_k(y^{k+1} - x^k)$  es la derivada de la parte cuadrática de la regularización. Entonces

$$-\mu_k(y^{k+1} - x^k) \in \partial \hat{f}_k(y^{k+1})$$

2. Por convexidad, no hay brecha de dualidad entre (4.13) y (4.16). Así,

$$\begin{aligned} \hat{f}_k(y^{k+1}) + \frac{\mu_k}{2} \|y^{k+1} - x^k\|^2 &= f(x^k) - \frac{1}{2\mu_k} \|\hat{s}^k\|^2 - \hat{e}_k \\ \hat{f}_k(y^{k+1}) &= f(x^k) - \frac{\mu_k}{2} \left\| \frac{-1}{\mu_k} \hat{s}^k \right\|^2 - \frac{1}{2\mu_k} \|\hat{s}^k\|^2 - \hat{e}_k \\ &= f(x^k) - \frac{1}{\mu_k} \|\hat{s}^k\|^2 - \hat{e}_k \end{aligned}$$

3. Usando el inciso 2 y la definición de  $\delta_k$

$$\begin{aligned} \delta_k &= f(x^k) - \hat{f}_k(y^{k+1}) - \frac{\mu_k}{2} \|y^{k+1} - x^k\|^2 \\ &= f(x^k) - \frac{\mu_k}{2} \|y^{k+1} - x^k\|^2 - f(x^k) + \frac{1}{\mu_k} \|\hat{s}^k\|^2 + \hat{e}_k \\ &= \frac{1}{2\mu_k} \|\hat{s}^k\|^2 + \hat{e}_k \end{aligned}$$

□

**Lema 4.2.** Para el subgradiente agregado y el error de linealización agregado, se tiene que

$$\hat{s}^k \in \partial_{\hat{e}_k} f(x^k)$$

**Demostración.**

Usando el inciso 1 del lema anterior,  $\hat{s}^k \in \partial \hat{f}_k(y^{k+1})$  y por construcción de  $\hat{f}_k \leq f$ . Entonces

$$\begin{aligned} f(y) \geq \hat{f}_k(y) &= \hat{f}_k(y^{k+1}) + \langle \hat{s}^k, y - y^{k+1} \rangle \\ &= f(x^k) - \frac{1}{\mu_k} \|\hat{s}^k\|^2 - \hat{e}_k + \langle \hat{s}^k, y - x^k + x^k - y^{k+1} \rangle \\ &= f(x^k) + \langle \hat{s}^k, y - x^k \rangle - \hat{e}_k + \langle \hat{s}^k, x^k - y^{k+1} \rangle - \frac{1}{\mu_k} \|\hat{s}^k\|^2 \\ &= f(x^k) + \langle \hat{s}^k, y - x^k \rangle - \hat{e}_k \end{aligned}$$

□

Los lemas anteriores son muy importantes para las pruebas de convergencia que se darán adelante. Antes de eso, se define la pieza lineal agregada como sigue

$$f_{\hat{e}_k}(y) = f(x^k) + \langle \hat{s}^k, y - x^k \rangle - \hat{e}_k$$

**Lema 4.3.** Para  $f_{\hat{e}_k}$  se tiene lo siguiente

1.  $f_{\hat{e}_k} = \hat{f}_k(y^{k+1}) + \langle \hat{s}^k, y - y^{k+1} \rangle$
2.  $f_{\hat{e}_k}(y) \leq \hat{f}_k(y)$

### 4.3.2 Convergencia del método de haces

Se supondrá que  $f$  es una función cerrada y finita en  $\mathbb{R}^n$ . Además, sea

$$K_s = \{k \in \mathbb{N} \mid \text{la } k\text{-ésima iteración fue un paso serio}\}$$

Este análisis se dividirá en dos partes, primero se supondrá que  $|K_s| = \infty$ . Además se definirá  $\hat{\delta} = 0$ , es decir, el algoritmo hará iteraciones una infinidad de veces.

**Lema 4.4.** Se considera el método de haces y sea  $\bar{f} = \min_x f(x) > -\infty$ . Entonces,

$$\sum_{k \in K_s} \delta_k \leq \frac{f(x^0) - \bar{f}}{m} < \infty$$

**Lema 4.5.** Supongamos que  $f^* = \lim_{k \in K_s} f(x^k) > -\infty$  y  $|K_s| = \infty$

1. si  $\sum_{k \in K_s} \frac{1}{\mu_k} = \infty$ , entonces el cero es un punto de acumulación de  $\{\hat{s}^k\}_{k \in K_s}$ , esto es  $\liminf \|\hat{s}^k\| = 0$ .
2. si  $\mu_k \geq c > 0$  y  $\arg \min_x f(x) \neq \emptyset$ , entonces la sucesión  $\{x^k\}_{k \in K_s}$  es acotada.

**Demostración.**

1. Por el inciso 3 del lema 4.1, se tiene que  $0 \leq \frac{1}{2\mu_k} \|\hat{s}^k\|^2 = \delta_k - \hat{e}_k \leq \delta_k$ . Así, por el lema anterior, dado que  $f(x^k)$  es una sucesión decreciente cuando  $k \in K_s$

$$\sum_{k \in K_s} \frac{\|\hat{s}^k\|^2}{2\mu_k} \leq \sum_{k \in K_s} \delta_k \leq \frac{f(x^0) - f^*}{m}$$

Así,  $\hat{s}^k \rightarrow 0$  sobre  $k \in K_s$ .



2. Sea  $x^* \in \arg \min_y f(y)$ . Por definición  $f(x^*) \leq f(y)$  para toda  $y \in \mathbb{R}^n$ . Ahora, para  $k \in K_s$

$$\begin{aligned}
\|x^* - x^{k+1}\|^2 &= \|x^* - x^k\|^2 + 2\langle x^* - x^k, x^k - x^{k+1} \rangle + \|x^k - x^{k+1}\|^2 \\
&= \|x^* - x^k\|^2 + \frac{2}{\mu_k} \langle x^* - x^k, \hat{s}^k \rangle + \frac{1}{\mu_k^2} \|\hat{s}^k\|^2 \\
&= \|x^* - x^k\|^2 + \frac{2}{\mu_k} (\langle x^* - x^k, \hat{s}^k \rangle + \frac{1}{2\mu_k} \|\hat{s}^k\|^2) \\
&\leq \|x^* - x^k\|^2 + \frac{2}{\mu_k} (f(x^*) - f(x^k) + \hat{e}_k + \frac{1}{2\mu_k} \|\hat{s}^k\|^2) \\
&\leq \|x^* - x^k\|^2 + \frac{2}{\mu_k} (f(x^*) - f(x^k) + \delta_k) \\
&\leq \|x^* - x^k\|^2 + \frac{2}{\mu_k} \delta_k
\end{aligned}$$

donde se usó el hecho de que  $\hat{s}^k \in \partial_{\hat{e}_k} f(x^k)$  y el inciso 3 del lema 4.1. Haciendo lo anterior  $k$  veces

$$\|x^* - x^{k+1}\|^2 \leq \|x^* - x^0\|^2 + 2 \sum_{i=1}^{k+1} \frac{\delta_i}{\mu_i} \leq \|x^* - x^0\|^2 + \frac{2}{c} \sum_{k \in K_s} \delta_k$$

Usando el lema 4.4 se tiene que el lado derecho es acotado y por tanto, la sucesión  $\{x^k\}_{k \in K_s}$  es acotada.

□

**Teorema 4.2.** *Supongamos que  $f^* = \lim_{k \in K_s} f(x^k) > -\infty$  y que  $|K_s| = \infty$ . Si la sucesión  $\{\mu_k\}$  es acotada y creciente, entonces  $\{x^k\}_{k \in K_s}$  tiene al menos un punto de acumulación que es óptimo.*

**Demostración.**

Por el lema 4.4 y 4.1 se sigue que  $0 \leq \hat{e}_k \leq \delta_k \rightarrow 0$  para  $k \in K_s$ . Luego, por el inciso 1 del lema 4.1

$$\hat{s}^k \in \partial_{\hat{e}_k} f(x^k)$$

para cualquier  $k \in K_s$ . Luego, por el inciso 2 del lema 4.1, existe una subsucesión  $\{\hat{s}^{n_k}\}$  que converge a cero. Dado que  $\{x^k\}_{k \in K_s}$  es acotada, entonces  $\{\hat{x}^{n_k}\}$  también es acotada. Tomando (si es necesario una subsucesión)  $x^{n_k} \rightarrow x^*$ , se tiene que

$$(\hat{x}^{n_k}, \hat{s}^{n_k}, \hat{e}_{n_k}) \rightarrow (x^*, 0, 0)$$

con esto,  $(x, \epsilon) \mapsto \partial_\epsilon f(x)$  es continua, así  $0 \in \partial f(x^*)$

□

En este momento solo se ha tomado en cuenta  $|K_s| = \infty$ , ahora se verá el caso cuando  $|K_s| < \infty$ . Se supondrá que  $k_0$  es la última iteración donde se tiene un paso serio.

**Lema 4.6.** *Sea  $x^{k_0}$  el último paso serio y  $\{y^{k+1}\}_{k \geq k_0}$  la sucesión de pasos nulos. Entonces para toda  $k > k_0$  y  $y \in \mathbb{R}^n$*

$$f(x^{k_0}) - \delta_k + \frac{\mu_k}{2} \|y - y^{k+1}\|^2 = \hat{f}(y^{k+1}) + \langle \hat{s}^k, y - y^{k+1} \rangle + \frac{\mu_k}{2} \|y - x^{k_0}\|^2 \quad (4.18)$$

***Demostración.***

Se verifica lo siguiente

$$\begin{aligned} \|y - x^{k_0}\|^2 &= \|y - y^{k+1} + y^{k+1} - x^{k_0}\|^2 \\ &= \|y - y^{k+1}\|^2 + 2\langle y - y^{k+1}, y^{k+1} - x^{k_0} \rangle + \|y^{k+1} - x^{k_0}\|^2 \\ &= \|y - y^{k+1}\|^2 - \frac{2}{\mu_k} \langle y - y^{k+1}, \hat{s}^k \rangle + \|y^{k+1} - x^{k_0}\|^2 \end{aligned}$$

Usando la definición de  $\delta_k$  se obtiene

$$\begin{aligned} f(x^{k_0}) - \delta_k + \frac{\mu_k}{2} \|y - y^{k+1}\|^2 &= \hat{f}_k(y^{k+1}) + \frac{\mu_k}{2} (\|y^{k+1} - x^{k_0}\|^2 + \|y - y^{k+1}\|^2) \\ &= \hat{f}_k(y^{k+1}) + \langle y - y^{k+1}, \hat{s}^k \rangle + \frac{\mu_k}{2} \|y - x^{k_0}\|^2 \end{aligned}$$

□

**Teorema 4.3.** *Sea  $x^{k_0}$  el último paso serio y sea  $\{y^{k+1}\}_{k \in K_s}$  la sucesión de pasos nulos. Si  $\{\mu_k\}_{k > k_0}$  es no decreciente, entonces  $\delta_k \rightarrow 0$ .*

***Demostración.***

Sea  $y = y^{k+2}$  en lema anterior, entonces

$$\begin{aligned} f(x^{k_0}) - \delta_k + \frac{\mu_k}{2} \|y^{k+2} - y^{k+1}\|^2 &= \hat{f}_k(y^{k+1}) + \langle \hat{s}^k, y^{k+2} - y^{k+1} \rangle + \frac{\mu_k}{2} \|y^{k+2} - x^{k_0}\|^2 \\ &= \hat{f}_{\hat{e}_k}(y^{k+1}) + \frac{\mu_k}{2} \|y^{k+2} - x^{k_0}\|^2 \\ &\leq \hat{f}_{k+1}(y^{k+2}) + \frac{\mu_k}{2} \|y^{k+2} - x^{k_0}\|^2 \\ &\leq \hat{f}_{k+1}(y^{k+2}) + \frac{\mu_{k+1}}{2} \|y^{k+2} - x^{k_0}\|^2 \\ &= f(x^{k_0}) - \delta_{k+1} \end{aligned}$$

donde la última igualdad se deduce de que  $\mu_k \leq \mu_{k+1}$ . Por lo tanto

$$\frac{\mu_k}{2} \|y^{k+2} - y^{k+1}\|^2 + \delta_{k+1} \leq \delta_k \quad (4.19)$$

Ahora, se probará que la sucesión  $\{y^k\}$  es acotada. Usando de nuevo el lema anterior con  $y = x^{k_0}$

$$\begin{aligned} f(x^{k_0}) - \delta_k + \frac{\mu_k}{2} \|x^{k_0} - y^{k+1}\|^2 &= \hat{f}_k(y^{k+1}) + \langle \hat{s}^{k_0}, x^{k_0} - y^{k+1} \rangle \\ &= f_{\hat{e}_k}(x^{k_0}) \leq \hat{f}_k(x^{k_0}) \leq f(x^{k_0}) \end{aligned}$$

Así

$$\|x^{k_0} - y^{k+1}\|^2 \leq \frac{2\delta_k}{\mu_k} \leq \frac{2\delta_{k_0}}{\mu_{k_0}}$$

Dado que  $\{\delta_k\}$  es decreciente y  $\mu_k$  es no decreciente,  $\{y^k\}$  es una sucesión acotada.

Por otro lado, sea  $C$  constante de Lipschitz para  $f$  y  $\hat{f}_k$  en  $B(x^{k_0}, \frac{\delta_{k_0}}{\mu_{k_0}})$ . Combinando

$$-m\delta_k \leq f(y^{k+1}) - f(x^{k_0})$$

y

$$\delta_k \leq f(x^{k_0}) - \hat{f}_k(y^{k+1})$$

y sustituyendo  $y = y^k$  en (4.11) en la iteración  $k - 1$  se tiene que  $f(y^k) + \hat{f}_k(y^k)$ , así

$$\begin{aligned} (1 - m)\delta_k &\leq f(y^{k+1}) - \hat{f}_k(y^{k+1}) \\ &= f(y^{k+1}) + f(y^k) + \hat{f}_k(y^k) - \hat{f}_k(y^{k+1}) \\ &\leq 2C\|y^{k+1} - y^k\|^2 \end{aligned}$$

combinando esta desigualdad con (4.19)

$$\delta_k - \delta_{k+1} \geq \|y^{k+2} - y^{k+1}\|^2 \geq \frac{(1 - m)^2}{8C^2} \mu_k \delta_k^2 \geq \frac{(1 - m)^2}{8C^2} \mu_{k_0} \delta_{k+1}^2$$

Así, sumando sobre  $k \geq k_0$

$$\frac{(1 - m)^2}{8C^2} \mu_{k_0} \sum_{k \geq k_0} \delta_k^2 \leq \sum_{k \geq k_0} (\delta_k - \delta_{k+1}) \leq \delta_{k_0}$$

lo cual implica que  $\delta_k \rightarrow 0$ .

□

**Teorema 4.4.** *Sea  $x^{k_0}$  la última iteración con paso serio del método de haces. Si  $\{\mu_k\}_{k \geq k_0}$  es no decreciente, entonces  $x^{k_0}$  es una solución óptima.*

**Demostración.**

Dado que las suposiciones son las mismas que en el teorema anterior, entonces  $\delta_k \rightarrow 0$  implica que  $\hat{e}_k \rightarrow 0$  y  $\|\hat{s}^k\| \rightarrow 0$  por el inciso 3 del lema 4.1. De nuevo, por el inciso 1 del lema 4.1

$$\hat{s}^k \in \partial_{\hat{e}_k} f(x^{k_0})$$

para todo  $k > k_0$ . Por el teorema ?? que asegura que la correspondencia  $(x, \epsilon) \mapsto \partial_\epsilon f(x)$  es continua, se concluye que  $0 \in \partial f(x^{k_0})$ .

□

## 4.4 Método del gradiente próximo

Sea  $f : \mathbb{R}^n \mapsto (-\infty, +\infty]$  una función propia, convexa y cerrada. Se define el operador próximo como  $prox_f : \mathbb{R}^n \mapsto \mathbb{R}^n$  de  $f$  como [7]

$$prox_f(v) = \arg \min_x \{f(x) + \frac{1}{2}\|x - v\|^2\} \quad (4.20)$$

donde  $\|\cdot\|$  es la norma euclidiana. Es decir, el operador próximo es aquel que cumple

$$f(prox_f(v)) + \frac{1}{2}\|prox_f(v) - v\|^2 = F(v)$$

De aquí que el operador próximo y la regularización de Moreau-Yosida estén relacionadas. El operador próximo es estrictamente convexo dado que la norma euclidiana lo es y tiene valor finito en todas partes. Así, tiene un único minimizador para cualquier  $v \in \mathbb{R}^n$  (incluso cuando  $dom f \subsetneq \mathbb{R}^n$ ).

A menudo se encuentra el operador próximo de una función escalada  $\lambda f$  ( $\lambda > 0$ ) el cual puede ser expresado por

$$prox_{\lambda f}(v) = \arg \min_x \{f(x) + \frac{1}{2\lambda}\|x - v\|^2\} \quad (4.21)$$

En la Figura (4.1) se representa lo que hace el operador próximo. Las líneas delgadas grises representan las curvas de nivel de la función convexa  $f$  y la línea gruesa negra representa la frontera del dominio esencial de  $f$ . Evaluar los puntos azules en el operador próximo tiene como resultado los puntos en rojo respectivamente. Los puntos en el dominio esencial de la función se mantienen en el dominio y se mueven hacia el mínimo de la función, mientras que los puntos fuera del dominio esencial se mueven hacia la frontera de este y hacia el mínimo de la función. El parámetro  $\lambda$  controla la proporción a la cual el operador próximo mapea puntos del dominio esencial de  $f$  hacia el punto mínimo.

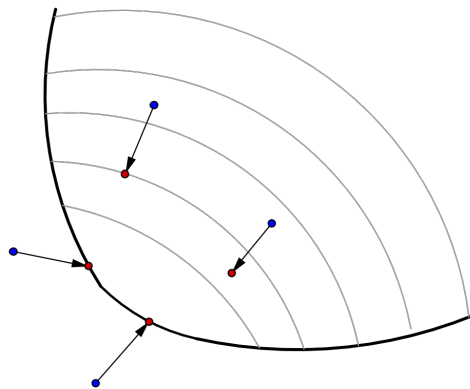


Figura 4.1: Operador próximo evaluado en varios puntos

Por la definición del operador próximo,  $prox_f(v)$  es un punto que se encuentra entre  $x^* = \min f$  y cerca de  $v$ . En el caso de  $prox_{\lambda f}(v)$ , el parámetro  $\lambda$  puede ser interpretado como un

peso relativo entre estos dos términos.

Cuando  $f$  es la función indicadora de  $C \subseteq \mathbb{R}^n$ , es decir

$$\mathbb{I}_C(x) = \begin{cases} 0 & \text{si } x \in C \\ +\infty & \text{si } x \notin C \end{cases}$$

Así el operador próximo se puede ver como una proyección generalizada. Esta perspectiva sugiere varias propiedades que se espera que el operador próximo cumpla.

Si  $f(x, y) = \phi(x) + \eta(y)$ , es decir,  $f$  puede ser separada por dos funciones que dependen de cada una de las variables, entonces:

$$\text{prox}_f(v_1, v_2) = (\text{prox}_\phi(v_1), \text{prox}_\eta(v_2))$$

Esto se observa directamente de la definición del operador próximo y del hecho que  $f$  es separable.

Una generalización de este resultado se tiene si  $f(x) = \sum_{i=1}^n f_i(x_i)$ , entonces, el operador próximo sería

$$(\text{prox}_f(v))_i = \text{prox}_{f_i}(v_i) \quad (4.22)$$

En otras palabras, este caso se reduce a evaluar el operador próximo de funciones escalares. Ahora, si  $f(x) = \alpha\phi(x) + b$ , con  $\alpha > 0$ , entonces

$$\text{prox}_{\lambda f} = \text{prox}_{\alpha\lambda\phi}(v) \quad (4.23)$$

La prueba de esto es inmediata de la definición del operador próximo.

**Proposición 4.3.** *Sea  $f$  una función propia, convexa y cerrada, definida sobre  $\mathbb{R}^n$ . Entonces  $p = \text{prox}_f(x)$  si y sólo si  $x - p \in \partial f(p)$*

### ***Demostración***

Supongamos que  $p = \text{prox}_f(x)$ , esto se cumple si y solo si

$$0 \in \partial f(p) + (p - x)$$

$$\Leftrightarrow x - p \in \partial f(p)$$

□

### 4.4.1 Puntos fijos

Los puntos fijos del operador próximo juegan un papel fundamental cuando se trata de minimizar la función  $f$  como lo enuncia el siguiente resultado.

**Teorema 4.5.**  $x^*$  minimiza a  $f$  si y sólo si  $x^* = \text{prox}_f(x^*)$

**Demostración.**

$\Rightarrow$  Supongamos que  $x^*$  minimiza a  $f$ , es decir,  $f(x) \geq f(x^*)$  para cualquier  $x$ . Entonces

$$f(x) + \frac{1}{2}\|x - x^*\|^2 \geq f(x^*) = f(x^*) + \frac{1}{2}\|x^* - x^*\|^2$$

y por lo tanto  $x^*$  minimiza a  $f(x) + \frac{1}{2}\|x - x^*\|^2$ . Así, tenemos que  $x^* = \text{prox}_f(x^*)$ .

$\Leftarrow$  Como  $\hat{x}$  minimiza  $f(x) + \frac{1}{2}\|x - v\|^2$  si y sólo si

$$0 \in \partial f(\hat{x}) + (\hat{x} - v)$$

entonces si  $x^*$  es punto fijo de  $\text{prox}_f$ ,  $0 \in \partial f(x^*)$ .

□

Dado que los minimizadores de  $f$  son puntos fijos del operador  $\text{prox}_f$ , el problema de minimizar  $f$  se convierte en un problema de punto fijo. Si el operador  $\text{prox}_f$  fuera una contracción, la sucesión definida por  $x_{n+1}(x_n)$  con  $x_0 \in S$  converge al único punto fijo del operador  $\text{prox}_f$ . El operador próximo no necesariamente es una contracción, pero es firmemente no expansivo.

Considerar el siguiente problema

$$\min h(x) = f(x) + g(x)$$

Donde  $f : \mathbb{R}^n \mapsto \mathbb{R}$  y  $g : \mathbb{R}^n \mapsto (-\infty, +\infty]$  son funciones propias, cerradas, convexas y  $f$  es diferenciable. En esta forma se puede partir el problema en dos, en donde una parte es diferenciable. La forma en la cual se parte el problema no es única y puede llevar a diferentes implementaciones del método del gradiente próximo [7].

El método del gradiente próximo es:

$$x_{k+1} = \text{prox}_{\lambda_k g}(x_k - \lambda_k \nabla f(x_k))$$

donde  $\lambda_k > 0$  es el tamaño de paso.

Es decir, es como hacer un descenso por gradiente de la parte diferenciable y luego aplicar el operador proximal de  $g$ .

### 4.4.2 Convergencia del método del gradiente próximo

Primero supongamos que  $\nabla f$  es Lipschitz con constante  $L \in (0, 1)$ , s decir

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$$

para toda  $x, y \in \mathbb{R}^n$ . Además supongamos que  $x^* \in \arg \min h \neq \emptyset$ . Se define

$$G_\lambda(x) = \frac{1}{\lambda}(x - \text{prox}_{\lambda g}(x - \lambda \nabla f(x)))$$

Esta expresión tiene la siguiente propiedad

$$G_\lambda(x) - \nabla f(x) \in \partial g(x - \lambda G_\lambda(x))$$

En efecto, por la proposición 4.3

$$\begin{aligned} \text{prox}_{\lambda g}(x - \lambda \nabla f(x)) = p &\Leftrightarrow 0 \in \partial g(p) + \frac{1}{\lambda}(p - (x - \lambda \nabla f(x))) \\ &\Leftrightarrow \frac{(x - \lambda \nabla f(x)) - p}{\lambda} \in \partial g(p) \\ &\Leftrightarrow G_\lambda(x) - \nabla f(x) \in \partial g(x - \lambda G_\lambda(x)) \end{aligned}$$

ahora se probará que  $h(x^k) - h(x^*)$  decrece con una rapidez  $1/k$  si un tamaño de paso  $\lambda_k = 1/L$  es usado. Del hecho que  $\nabla f$  es Lipschitz se tiene que

$$\begin{aligned} f(y) &= f(x) + \langle \nabla f(x), v \rangle + \int_0^1 \langle \nabla f(x + \lambda v) - \nabla f(x), v \rangle d\lambda \\ &\leq f(x) + \langle \nabla f(x), v \rangle + \int_0^1 \|\nabla f(x + \lambda v) - \nabla f(x)\| \|v\| d\lambda \\ &\leq f(x) + \langle \nabla f(x), v \rangle + \int_0^1 L\lambda \|v\|^2 d\lambda \\ &= f(x) + \langle \nabla f(x), v \rangle + \frac{L}{2} \|v\|^2 \end{aligned}$$

donde  $v = y - x$ . Sustituyendo  $y = x - \lambda G_\lambda(x)$  en la desigualdad anterior se obtiene

$$f(x - \lambda G_\lambda(x)) \leq f(x) - \lambda \langle \nabla f(x), G_\lambda(x) \rangle + \lambda^2 L \|G_\lambda(x)\|^2$$

Dado que  $0 < t \leq 1/L$  se tiene que

$$f(x - \lambda G_\lambda(x)) \leq f(x) - \lambda \langle \nabla f(x), G_\lambda(x) \rangle + \lambda^2 2 \|G_\lambda(x)\|^2$$

Luego, al tomar la función  $h$  y la desigualdad anterior

$$\begin{aligned}
h(x - \lambda G_\lambda(x)) &\leq f(x) - \lambda \langle \nabla f(x), G_\lambda(x) \rangle + \frac{\lambda}{2} \|G_\lambda(x)\|^2 + g(x - \lambda G_\lambda(x)) \\
&\leq f(z) + \langle \nabla f(z), x - z \rangle - \lambda \langle \nabla f(x), G_\lambda(x) \rangle + \frac{\lambda}{2} \|G_\lambda(x)\|^2 + \\
&\quad g(z) + \langle G_\lambda(x) - \nabla f(x), x - z - \lambda G_\lambda(x) \rangle \\
&= f(z) + g(z) + \langle G_\lambda(x), x - z \rangle - \frac{\lambda}{2} \|G_\lambda(x)\|^2 \\
&= h(z) + \langle G_\lambda(x), x - z \rangle - \frac{\lambda}{2} \|G_\lambda(x)\|^2 \tag{4.24}
\end{aligned}$$

donde la segunda desigualdad se tiene del hecho de que  $f$  y  $g$  son convexas y  $G_\delta(x) - \nabla f(x) \in \partial g(x - G_\lambda(x))$ . Ahora, sea  $x^+ = x - \lambda G_\lambda(x)$ , la desigualdad anterior con  $z = x$  implica que

$$h(x^+) \leq h(x) - \frac{\lambda}{2} \|G_\lambda(x)\|^2$$

lo cual nos dice que se tiene un método de descenso.

Igualmente, al tomar la desigualdad (4.24) con  $z = x^* \in \arg \min h$  se tiene que

$$\begin{aligned}
0 \leq h(x^+) - h(x^*) &\leq \langle G_\lambda(x), x - x^* \rangle - \frac{\lambda}{2} \|G_\lambda(x)\|^2 \\
&= \frac{1}{2\lambda} (\|x - x^*\|^2 - \|x - x^* - \lambda G_\lambda(x)\|^2) \\
&= \frac{1}{2\lambda} (\|x - x^*\|^2 - \|x^+ - x^*\|^2)
\end{aligned}$$

Por lo tanto

$$\|x - x^*\| \leq \|x^+ - x^*\|$$

es decir, la distancia al conjunto  $\arg \min h$  decrece.

Al sumar las desigualdades anteriores con  $x = x_{i-1}$ ,  $x_+ = x_i$  y  $t = 1/L$  se obtiene que

$$\begin{aligned}
\sum_{i=1}^k (f(x_i) - f(x^*)) &\leq \frac{1}{2\lambda} \sum_{i=1}^k (\|x_{i-1} - x^*\|^2 - \|x_i - x^*\|^2) \\
&= \frac{1}{2\lambda} (\|x_0 - x^*\|^2 - \|x_k - x^*\|^2) \\
&\leq \frac{1}{2\lambda} \|x_0 - x^*\|^2
\end{aligned}$$

Dado que  $f(x_i)$  es no decreciente

$$f(x_k) - f(x^*) \leq \frac{1}{2\lambda} \sum_{i=1}^k (f(x_i) - f(x^*)) \leq \frac{1}{2k\lambda} \|x_0 - x^*\|^2$$





Así se tiene el resultado deseado.

# Capítulo 5

## Ejemplos Numéricos

### 5.1 Subgradiente proyectado para el problema dual

Consideramos el problema de minimizar una función cuadrática estrictamente convexa sobre el cuadrado unidad donde  $P \succ 0$

$$\begin{aligned} \min \quad & (1/2)x^T P x - q^T x \\ \text{sujeto a:} \quad & x_i^2 \leq 1 \quad i = 1, 2, \dots, n \end{aligned}$$

El lagrangiano es:

$$L(\lambda, x) = (1/2)x^T (P + \mathbf{diag}(2\lambda))x - q^T x + 1^T \lambda$$

Por lo que

$$x^*(\lambda) = (P + \mathbf{diag}(2\lambda))^{-1}q$$

El subgradiente proyectado para el dual es:

$$x^{(k)} = (P + \mathbf{diag}(2\lambda))^{-1}q, \quad \lambda_i^{(k+1)} = (\lambda_i^{(k)} + \alpha_k((x_i^{(k)})^2 - 1))$$

La función dual es diferenciable, por lo que podemos usar un tamaño de paso fijo  $\alpha$  (siempre que sea lo suficientemente pequeño).

Las iteraciones  $x^{(k)}$  no son factibles, pero podemos construir un  $\hat{x}^{x^{(k)}}$  factible cercano como

$$\hat{x}_i^{(k)} = \begin{cases} 1 & x_i^{(k)} > 1 \\ -1 & x_i^{(k)} < -1 \\ x_i^{(k)} & -1 \leq x_i^{(k)} \leq 1 \end{cases}$$

Consideramos  $n = 50$ , salto de paso fijo  $\alpha = 0.1$ ,  $\lambda^{(1)} = 1$

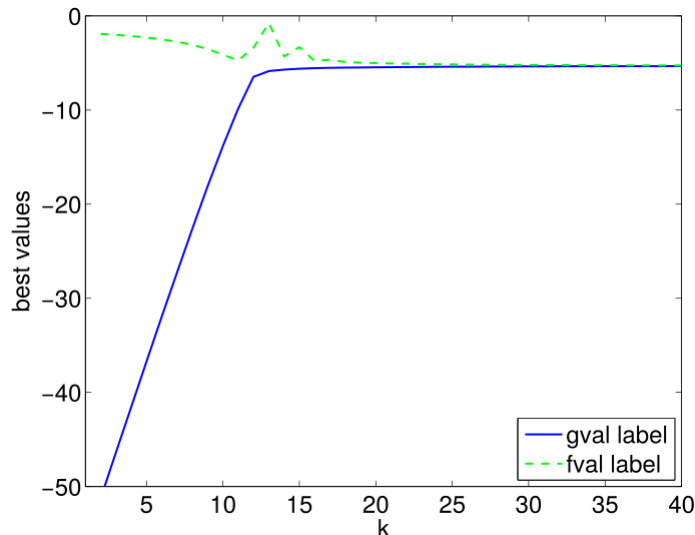


Figura 5.1: Valores del límite inferior  $gval$  y el límite superior  $fval$  versus el número de iteración  $k$ . Usamos un tamaño de paso fijo con  $\alpha = 0.1$

## 5.2 Procesamiento de señal adaptativo

Suponemos que  $(x, y) \in \mathbb{R}^n$  tienen alguna distribución conjunta. Nuestro objetivo es encontrar un vector de peso  $w \in \mathbb{R}^n$  para el cual  $w^T x$  sea un buen estimador de  $y$ . (Esta es una regresión lineal; podemos agregar un componente adicional a  $x$  que siempre es uno para obtener un estimador afín de la forma  $w^T x + v$ ) Nos gustaría elegir un vector de peso que minimice

$$J(w) = \mathbf{E}l(w^T x - y)$$

donde  $l$  es una función de pérdida convexa. Por ejemplo, si:

- $l(u) = u^2$  : error de media cuadrática.
- $l(u) = |u|$  : error media absoluta.

Pero consideraremos aquí el caso más general. Por supuesto,  $J$  es convexo, por lo que minimizar  $J$  sujeto a  $\omega$  es un problema de optimización convexa.

Consideramos una configuración en línea, donde no conocemos la distribución de  $(x, y)$ . En cada paso (que podría corresponder al tiempo, por ejemplo), se nos da una muestra  $(x(i), y(i))$  de la distribución. Después de  $k$  pasos, podríamos usar las  $k$  muestras para estimar la distribución.

Esta aproximación requiere que almacenemos todas las muestras anteriores, y necesitamos resolver un gran problema cada vez que entra una nueva muestra.

En su lugar, utilizaremos el método del subgradiente estocástico, que es realmente simple. En particular, no requiere esencialmente almacenamiento (más allá del valor actual de  $\omega$ ), y muy poco cálculo. La desventaja es que es lento.

En cada paso, se nos da una muestra  $(x^{(k)}, y^{(k)})$  imparcial de la distribución con ruido del subgradiente de  $J$  en  $w^{(k)}$ , basado en la muestra  $x^{(k)}, y^{(k)}$  :

$$g^{(k)} = l'(w^{(k)T}x^{(k+1)} - y^{(k+1)})x^{(k+1)}$$

donde  $l'$  es la derivada de (o subgradiente de  $l$  si esta es no diferenciable) de  $l$ . Este es un subgradiente con ruido imparcial de  $J$ .

El algoritmo en línea sería:

$$w^{(k+1)} = w^{(k)} - \alpha_k l'(w^{(k)T}x^{(k+1)} - y^{(k+1)})x^{(k+1)}$$

donde

- $l(u) = u^2$  da el algoritmo LMS (media de mínimos cuadrados).
- $l(u) = |u|$  da el algoritmo *sign*.
- $w^{(k)T}x^{(k+1)} - y^{(k+1)}$  es el error de predicción.

Datos del problema:  $n = 10$ ,  $(x, y) \sim \mathcal{N}(0, \Sigma)$ ,  $\Sigma$  aleatoria con  $\mathbf{E}(y^2) \approx 12$ ,  $\alpha_k = 1/k$

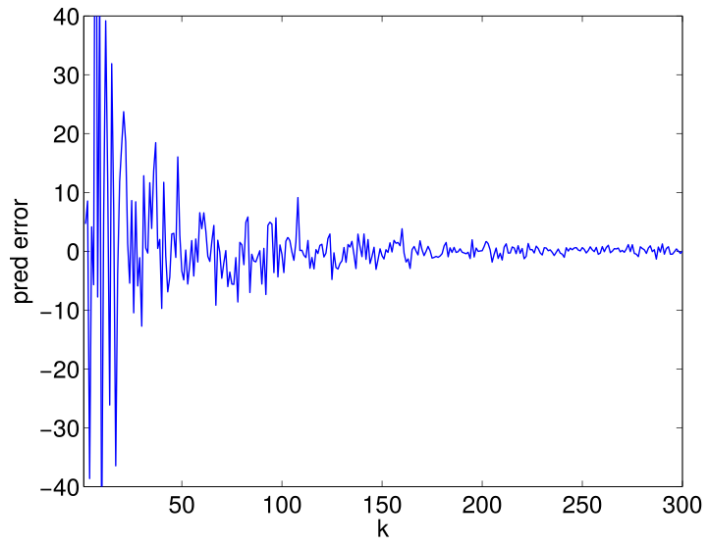


Figura 5.2: Error de predicción  $w^{(k)T}x^{(k+1)} - y^{(k+1)}$  vs. iteración  $k$

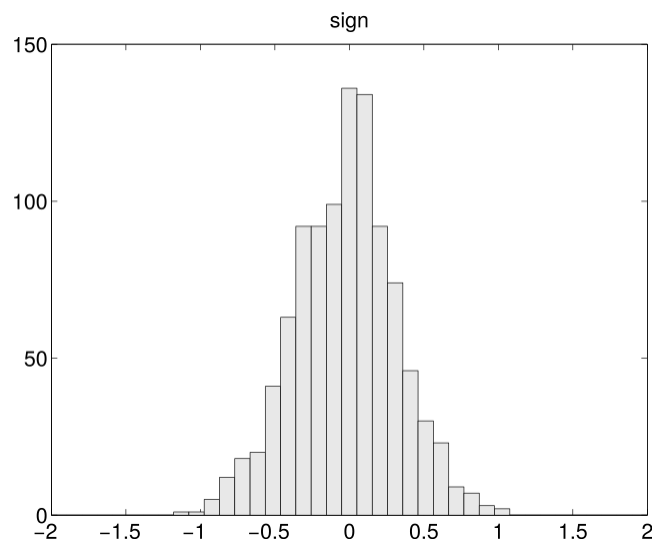


Figura 5.3: Distribución empírica del error de predicción para  $w^*$  (más de 1000 muestras)

# Recomendaciones

A continuación se presentan algunas recomendaciones referentes a distintos aspectos que conlleva hacer un trabajo de este tipo:

- Estudiar optimización convexa no diferenciable requiere de conocimientos vistos en la materia de Métodos de Optimización *I*, por lo que ha sido una buena forma de profundizar y terminar de asimilar muchos conceptos y propiedades, por ello sería bueno que consideren este tema para incluirse en los estudios de licenciatura.
- El estudio de optimización convexa no diferenciable, abre camino para estudiar distintas aplicaciones por lo cual puede considerarse como una opción para trabajos futuros.
- Aunque han habido eventos para impulsar el uso de  $\text{\LaTeX}$  para la elaboración de documentos matemáticos, se recomienda crear talleres de distintos software matemáticos, como MATLAB/Octave, Mathematica, Maple, GAMS, Scilab y otros, ya sea para tener herramientas a la hora de presentar un trabajo con aplicación matemática o ampliar las habilidades de los estudiantes.

# Conclusiones

La optimización no diferenciable produce una serie de dificultades que la mayoría de veces no pueden ser superadas por métodos tradicionales, por lo que es necesario poseer conocimiento de técnicas propias para su desarrollo. El estudio de optimización convexa es indispensable para el tema estudiado ya que proporciona propiedades que permiten resolver problemas de gran dimensión de forma eficiente en términos computacionales. Esto se debe a que la convexidad de éste puede hacer la diferencia al momento de resolver el problema.

Respecto a la parte de control óptimo estocástico (en tiempo discreto) fue bastante difícil encontrar material ya que la mayoría de textos disponibles se limitan a estudiarlo en tiempo continuo. Debo aclarar que lo que hace a un problema estocástico es una o más variables aleatorias definida por una distribución de probabilidad (perturbación, ruido blanco). Por ello, el estudio del *principio de equivalencia cierta* facilitaba en gran medida el proceso de encontrar una solución (siempre y cuando la parte aleatoria de la función sea aditiva) ya que gracias a ello el ruido blanco (perturbación, variable aleatoria) desaparece y nos quedamos con un problema determinístico, el cual es menos complicado de tratar.

Teniendo conocimiento de qué tipo de problema tenemos y la transformación que se ha hecho, podemos llegar a que el problema que se tiene finalmente se haya convertido en diferenciable (raras veces pasa) o se mantenga no diferenciable que es lo usual; sólo queda escoger el algoritmo para encontrar la solución óptima.

Se ha presentado una lista de algoritmos con su respectiva prueba de convergencia, aunque los presentados sean los más conocidos, el algoritmo puede variar dependiendo de las características del problema, mientras se mantenga la convexidad del problema, no surge complicación alguna en las variantes que pueda tener el algoritmo.

Recordar siempre que el método del subgradiente no es un método de descenso, por lo que no es de extrañar que en un principio podamos observar cómo los datos no poseen un comportamiento monótono como se muestra en la Figura (5.1).

Para presentar los ejemplos, se ha utilizado el sistema de modelado CVX, el cual está basado en MATLAB. CVX resultó ser una herramienta bastante efectiva e intuitiva de usar, pues la sintaxis usada es muy similar a la descripción matemática del problema y requiere que el problema a resolver sea convexo.

# Capítulo 6

## Anexos

CVX es un sistema de modelado basado en Matlab para la optimización convexa. CVX convierte a Matlab en un lenguaje de modelado, lo que permite especificar restricciones y funciones objetivos utilizando la sintaxis de expresión estándar de Matlab.

En su modo predeterminado, CVX admite un enfoque particular para la optimización convexa. Bajo este enfoque, las funciones y conjuntos convexos se crean a partir de un pequeño conjunto de reglas del análisis convexo, comenzando desde una biblioteca base de funciones y conjuntos convexos.

Es muy importante tener en cuenta también lo que no es CVX. No es una herramienta de uso general para la optimización no lineal, ni es una herramienta para verificar si el modelo es convexo o no.

El paquete CVX incluye una creciente biblioteca de ejemplos, de los cuales hemos elegidos dos para presentarlos en este trabajo (Véase Capítulo 5).

### 6.1 Algoritmo para subgradiente proyectado del problema dual

A continuación se presenta el código para el problema dual [41] presentado en el capítulo 5.

El código fue escrito por *Almir Mutapcic* y se encuentra disponible en [subgrad\\_method\\_dual\\_prob.m](#)

```
% Subgradient method for the dual problem.
%
% Primal QCQP problem:
% minimize (1/2)x'Px - q'x
% subject to x^2 <= 1
%
% where P is a positive definite matrix (objective is strictly convex)
```



## 6.1. ALGORITMO PARA SUBGRADIENTE PROYECTADO DEL PROBLEMA DUAL147

```

%
% Dual problem is:
% maximize   -(1/2)q'(P + diag(2*lambda))^{-1}q - sum(lambda)
% subject to lambda => 0
%
% where lambda are dual variables
%
% EE364b Convex Optimization II, S. Boyd
% Written by Almir Mutapcic, 01/19/07
%

% generate a problem instance
n = 50; % number of variables n
% generate problem data (P should be positive definite)
randn('state',1); % set state so problem is reproducible
A = randn(n,n);
P = (A'*A);
q = randn(n,1);
eigsP = eig( P );
fprintf(1,'Constructed mtx with min eig = %3.4f and max eig = %3.4f\n',...
        min(eigsP), max(eigsP));

%*****
% optimal solution of the primal QCQP problem
%*****
cvx_begin
    variable x(n,1);
    dual variable lambda_opt;
    minimize( 1/2*x'*P*x - q'*x )
    subject to
        x.^2 <= 1 : lambda_opt;
cvx_end

f_min = cvx_optval;
fprintf(1,'QCQP optimal value is %0.4f.\n\n',f_min);

%*****
% projected subgradient method applied to the primal problem
%*****
fp = [+Inf]; fpbest = [+Inf];
disp('Starting projected subgradient algorithm for the primal problem...')

% initial point
x = zeros(n,1);

```

```

k = 1;
MAX_ITERS = 100;

while k < MAX_ITERS

    % subgradient calculation
    g = P*x - q;

    % primal objective values
    fval = (1/2)*x'*P*x - q'*x;
    fp(end+1) = fval; fpbest(end+1) = min( fval, fpbest(end) );

    % step size selection
    alpha = (fval-f_min)/norm(g)^2;

    % projected subgradient update
    x = x - alpha*g; k = k + 1;

    % projection onto the feasible set (saturation function)
    x = max( min( x, 1 ), -1 );

end

%*****
% subgradient method applied to the dual problem
%*****
f = [+Inf]; fbest = [+Inf];
g = [-Inf]; gbest = [-Inf];
disp('Starting the subgradient algorithm applied to the dual problem...')

% initial point
lambda_1 = ones(n,1);
lambda = lambda_1;

k = 1;

while k < MAX_ITERS

    % subgradient calculation
    x_star = (P + diag(2*lambda))\q;
    h = x_star.^2 - 1;

    % dual objective values
    gval = -(1/2)*q'*x_star - sum(lambda);
    g(end+1) = gval; gbest(end+1) = max( gval, gbest(end) );

```

```

% primal objective values
x_star = max( min( x_star, 1 ), -1 ); % find nearby feasible point
fval = (1/2)*x_star'*P*x_star - q'*x_star;
f(end+1) = fval; fbest(end+1) = min( fval, fbest(end) );

% step size selection
alpha = 0.1;

% projected subgradient update
lambda = max(0, lambda + alpha*h); k = k + 1;

end

%*****
% plot results
%*****
figure(1), clf
set(gca, 'FontSize',18);
semilogy( [1:MAX_ITERS],fbest-gbest, 'b-', 'LineWidth',1.5 ), hold on,
semilogy( [1:MAX_ITERS],fbest-f_min, 'r-', 'LineWidth',1.5 ), hold on,
semilogy( [1:MAX_ITERS],fpbest-f_min,'k-', 'LineWidth',1.5 ), hold on,
xlabel('k');
ylabel('convergence');
legend('duality gap','optim gap dual','optim gap primal',1);

figure(2), clf
set(gca, 'FontSize',18);
plot( [1:MAX_ITERS], g, 'b-', 'LineWidth',1.5 ), hold on,
plot( [1:MAX_ITERS], f, 'g--', 'LineWidth',1.5 ), hold on,
xlabel('k');
ylabel('best values');
axis([1 40 -50 0]);
legend('gval label','fval label',4);
%print -depsc dual_prob_subgrad_conv

figure(3), clf
set(gca, 'FontSize',18);
plot( [1:MAX_ITERS], gbest, 'b-', 'LineWidth',1.5 ), hold on,
plot( [1:MAX_ITERS], fbest, 'g--', 'LineWidth',1.5 ), hold on,
xlabel('k');
ylabel('best values');
axis([1 40 -50 0]);
legend('gbest label','fbest label',4);

```

## 6.2 Algoritmo para el problema de procesamiento de señal adaptativo

A continuación se presenta el código para el problema dual [40] presentado en el capítulo 5.

El ejemplo de procesamiento de señal adaptativo fue sugerido por *Trevor Hastie* en el texto [40], el código fue escrito por *Almir Mutapcic* y está disponible en [stoch\\_subg\\_online\\_reg.m](http://stoch_subg_online_reg.m)

```
% On-line learning and regression example via stochastic subgradients.
%
% EE364b Convex Optimization II, S. Boyd
% Written by Almir Mutapcic, 01/19/07
%

%*****
% generate a problem instance
%*****
n = 10; % number of weights (variables)

% generate a random Sigma_x covariance matrix
randn('state',1); % set state so problem is reproducible
S = randn(n,n);
Sigma_x = S'*S;

% linear relationship
a = randn(n,1)/sqrt(n);
Sigma_y = a'*Sigma_x*a + 0.1;
Sigma_xy = Sigma_x*a;

Sigma = [Sigma_x, Sigma_xy; Sigma_xy', Sigma_y];
R = sqrtm(Sigma);

%*****
% stochastic subgradient method with diminishing step sizes (sign alg)
%*****
f = [+Inf]; fbest = [+Inf]; errors = [+Inf];

MAX_ITERS = 5000;

iter = 1;
```

## 6.2. ALGORITMO PARA EL PROBLEMA DE PROCESAMIENTO DE SEÑAL ADAPTATIVO151

```
w = zeros(n,1); % initial point
whist = [w];

while iter < MAX_ITERS
    if( rem(iter,500) == 0 ), fprintf(1,'iter: %d\n',iter), end

    % generate a new sample from the distribution
    z = R*randn([n+1 1]);
    x = z(1:n); y = z(n+1);

    % noisy subgradient calculation
    e = w'*x - y;
    g = sign(e)*x;

    % step size selection
    alpha = 1/iter;

    % objective values
    errors(end+1) = e;
    fval = abs(e);
    f(end+1) = fval;
    fbest(end+1) = min( fval, fbest(end) );

    % subgradient update
    w = w - alpha*g; iter = iter + 1; whist = [whist, w];
end

w_sign = w;

% compute optimal (lms) weight that minimizes mean-square error
% using analytical formula
w_lms = Sigma(1:n,1:n)\Sigma(1:n,n+1);

%*****
% report and plot problem results
%*****
% generate random samples
NSAMPLES = 1000;
Z = R*randn([n+1 NSAMPLES]);
X = Z(1:n,:); Y = Z(n+1,:);

err_sign = ( w_sign'*X - Y );
abserr_sign = abs( err_sign );

err_lms = ( w_lms'*X - Y );
```

```

abserr_lms = abs( err_lms );

disp(' ')
fprintf(1,'Sign abs predict error %3.4f\n',mean(abs(err_sign)));
fprintf(1,'Sign mean-square error %3.4f\n',mean((err_sign).^2));
disp(' ')
fprintf(1,'LMS abs predict error %3.4f\n',mean(abs(err_lms)));
fprintf(1,'LMS mean-square error %3.4f\n',mean((err_lms).^2));

iters = [1:MAX_ITERS];
figure(1), clf
set(gca, 'FontSize',18);
plot( iters, errors, 'b-', 'LineWidth', 1 )
axis([1 300 -40 40])
xlabel('k');
ylabel('pred error');
%print -depsc online_reg_errors

% histogram plots
figure(2), clf
edges = linspace(-2,2,40);
[count] = histc( err_sign, edges )';
bar( edges, count, 'histc' );
hline = findobj(gca,'Type','line'); delete(hline)
hpatch = findobj(gca,'Type','patch');
set(hpatch,'FaceColor',[0.91,0.91,0.91])
set(gca, 'FontSize',16);
axis([-2 2 0 150]);
title('sign')
%print -depsc online_reg_dist

```

# Bibliografía

- [1] Convex Analysis. R. Tyrrell Rockafellar. Princeton University Press, 1996.
- [2] Nonlinear programming. Theory and algorithms. Bazaraa M., Sherali H., Shetty C. Wiley-Interscience, 2006.
- [3] Fundamentos de optimización matemática en economía. R. Meneu, J.M Pérez-Salamero, M. Ventura. Programacion no lineal. Repro-Exprés, S.L, 1999
- [4] Emilio Cerdá Tena. Optimización dinámica. Pearson educación, S. A. Madrid 2001.
- [5] Introduccion al Analisis Convexo. Pedro Gajardo. Centro de Modelamiento Matemático. Universidad de Chile  
[http://www.dim.uchile.cl/~pgajardo/documentos/A\\_Convexo\\_pgajardo.pdf](http://www.dim.uchile.cl/~pgajardo/documentos/A_Convexo_pgajardo.pdf)
- [6] OPTIMIZACION NO DIFERENCIABLE. Hector Manuel Mora Escobar. Dpto. Matemáticas, U. Nacional, Bogotá  
<http://www.unicauca.edu.co/matematicas/eventos/optimizacion/Cursillos/Hector-Mora.pdf>
- [7] Funciones convexas no diferenciables. Tesis de Rafael Alejandro Nava Manzo. Universidad Autónoma Metropolitana. Unidad Iztapalapa  
[http://mat.izt.uam.mx/mcmai/documentos/tesis/Gen.12-O/Nava\\_Manzo\\_Rafael\\_Alejandro.pdf](http://mat.izt.uam.mx/mcmai/documentos/tesis/Gen.12-O/Nava_Manzo_Rafael_Alejandro.pdf)
- [8] HERRAMIENTAS DE OPTIMIZACIÓN CONVEXA Y APLICACIONES EN TELECOMUNICACIONES. Emilio Mejía Fernández de Velasco. UNIVERSIDAD CARLOS III DE MADRID. ESCUELA POLITÉCNICA SUPERIOR  
[http://e-archivo.uc3m.es/bitstream/handle/10016/12990/PFC\\_Emilio\\_Mejia\\_Fernandez\\_Velasco.pdf?sequence=3](http://e-archivo.uc3m.es/bitstream/handle/10016/12990/PFC_Emilio_Mejia_Fernandez_Velasco.pdf?sequence=3)
- [9] Desarrollo del concepto de función convexa. Nelson J. Merentes D. Universidad Central de Venezuela  
<http://evm.ivic.gob.ve/LibroMerentesRivas.pdf>
- [10] Análisis convexo y dualidad. Felipe Álvarez con la colaboración de Juan Escobar y Juan Peypouquet. Universidad de Chile  
[opteq.dim.uchile.cl/public/?/publicaciones/getUsersDocs/pdf/11f75d0f22278872192b00ab3c5ab428/1](http://opteq.dim.uchile.cl/public/?/publicaciones/getUsersDocs/pdf/11f75d0f22278872192b00ab3c5ab428/1)

- [11] Optimización. Jorge Amaya A. Departamento de Ingeniería Matemática y Centro de Modelamiento Matemático. Universidad de Chile.  
[http://escuela.mate304.org/NotasOptimizaci%C3%B3n\\_Convexa.pdf](http://escuela.mate304.org/NotasOptimizaci%C3%B3n_Convexa.pdf)
- [12] Núcleo Promedio de Funciones convexas y Algoritmos en Computación Convexa. César Jesús Lara Ávila. Universidad Nacional de Ingeniería. Lima-Perú.  
[http://cybertesis.uni.edu.pe/bitstream/uni/1165/1/lara\\_ac.pdf](http://cybertesis.uni.edu.pe/bitstream/uni/1165/1/lara_ac.pdf)
- [13] OPTIMIZACIÓN CONVEXA NO DIFERENCIABLE. Dr. Pablo A. Lotito.
- [14] Algunas aplicaciones y extensión del método del subgradiente. Frank Navarro Rojas. UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS. Lima – Perú.  
[http://cybertesis.unmsm.edu.pe/bitstream/cybertesis/4580/1/Navarro\\_rf.pdf](http://cybertesis.unmsm.edu.pe/bitstream/cybertesis/4580/1/Navarro_rf.pdf)
- [15] Tema 2: Geometría de Convexos. José Miguel Manzano Prego. Universidad de Granada. España  
<http://www.ugr.es/~jmmanzano/material/TallerGeoTop-Tema2.pdf>
- [16] MODELOS DE OPTIMIZACIÓN ESTOCÁSTICA EN LA GENERACIÓN TÉRMICA DE ENERGÍA ELÉCTRICA. José David Escobar Muñoz, Daniel Alejandro Rivas Rivas. Universidad de El Salvador. El Salvador.  
<http://ri.ues.edu.sv/9805/1/19200968.pdf>
- [17] LA MODERNA TEORÍA DE CONTROL ÓPTIMO: UNA APLICACIÓN AL MONOPOLIO ESPAÑOL DE TABACOS. María Isabel Toledo Muñoz. Universidad Autónoma de Madrid. España.  
<https://repositorio.uam.es/handle/10486/5304?show=full>
- [18] Control Óptimo Estocástico en la Enseñanza de la Economía Matemática. María Teresa V. Martínez Palacios, Francisco Venegas-Martínez. Escuela Superior de Economía del Instituto Politécnico Nacional. México.  
<http://www.scielo.org.mx/pdf/ed/v23n3/v23n3a7.pdf>
- [19] VOLATILIDAD, CRECIMIENTO ECONÓMICO Y DESIGUALDAD: UN MODELO DE CONTROL ÓPTIMO ESTOCÁSTICO. Saúl Mendoza Palacios. Universidad Nacional Autónoma de México. México.  
<http://www.ptolomeo.unam.mx:8080/xmlui/bitstream/handle/132.248.52.100/3989/TESIS.pdf?sequence=1>
- [20] <http://tesis.uson.mx/digital/tesis/docs/1608/Capitulo3.pdf>
- [21] Control Óptimo y Juegos Estocásticos. Onésimo Hernández- Lerma. EMALCA. CIMAT. Guanajuato, Agosto 1–12, 2005.  
[https://www.cimat.mx/Eventos/EMALCA/minicurso\\_onesimo.pdf](https://www.cimat.mx/Eventos/EMALCA/minicurso_onesimo.pdf)
- [22] Procesos de Decisión de Markov Descontados: Soluciones Óptimas mediante Problemas de Control Determinista Diferenciables. Hugo Cruz-Suárez. Universidad Autónoma



Metropolitana-Iztapalapa- México

- [23] Optimización de Portafolios. Patricia Saavedra Barrera. Universidad Autónoma Metropolitana–Iztapalapa. México.  
[https://www.google.com/sv/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&uact=8&ved=0ahUKEwjirLmO89DXAhUE2yYKHburDoEQFgg1MAI&url=https%3A%2F%2Fwww.u-cursos.cl%2Fingenieria%2F2012%2F1%2FIN5303%2F2%2Fmaterial\\_docente%2Fbajar%3Fid\\_material%3D418244&usg=AOvVaw2bmoyQap3pYbqz-yIy5s1-](https://www.google.com/sv/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&uact=8&ved=0ahUKEwjirLmO89DXAhUE2yYKHburDoEQFgg1MAI&url=https%3A%2F%2Fwww.u-cursos.cl%2Fingenieria%2F2012%2F1%2FIN5303%2F2%2Fmaterial_docente%2Fbajar%3Fid_material%3D418244&usg=AOvVaw2bmoyQap3pYbqz-yIy5s1-)
- [24] Control Óptimo: tiempo discreto. Eduardo Martínez. Notas sobre teoría de control, curso 2014 – 2015. Universidad de Zaragoza, España.
- [25] Procesos de Control de Markov Descontados Miopes. Raúl MONTES-DE-OCA. Universidad Autónoma Metropolitana-Iztapalapa. México
- [26] Control estocástico. Una visión general. Emilio Cerdá Tena. Madrid, España.  
<http://eprints.ucm.es/26237/1/9125.pdf>
- [27] Técnica de optimización para la recuperación de imágenes. Carlos Laguillo García. Universidad de Cantabria.  
<https://repositorio.unican.es/xmlui/bitstream/handle/10902/7129/Carlos%20Laguillo%20Garcia.pdf?sequence=1>
- [28] Optimización bajo incertidumbre (2008). Ramos, Andrés; Alonso Ayuso, Antonio; Pérez, Gloria. Red. Universidad Pontificia de Comillas.  
<https://docplayer.es/90899670-Optimizacion-bajo-incertidumbre.html>
- [29] The Certainty Equivalence Property in Stochastic Control Theory. HENK VAN DE WATER AND JAN C. WILLEMS.  
<http://homes.esat.kuleuven.be/~jwillems/Articles/JournalArticles/1981.2.pdf>
- [30] HEURÍSTICA LAGRANGIANA PARA EL PROBLEMA DE LOCALIZACIÓN CAPACITADO EN DOS ETAPAS. Edith Lucero Ozuna Espinosa. Universidad Autónoma de Nuevo León. México.  
<http://eprints.uanl.mx/2663/1/1080237524.pdf>
- [31] Aplicación de máximos y mínimos.  
[http://gc.initelabs.com/recursos/files/r157r/w13092w/Mate2\\_Lic\\_4aEd\\_07.pdf](http://gc.initelabs.com/recursos/files/r157r/w13092w/Mate2_Lic_4aEd_07.pdf)
- [32] TEMA 5. PROCESOS ESTOCÁSTICOS  
[http://www.dmae.upct.es/~mcruiz/Telem06/Teoria/apuntes\\_procesos.pdf](http://www.dmae.upct.es/~mcruiz/Telem06/Teoria/apuntes_procesos.pdf)
- [33] ESTUDIO DE LA GEOMETRÍA FRACTAL CON APLICACIONES A FINANZAS Y VULCANOLOGÍA. Milton Arnoldo Rivas Morales. Universidad de El Salvador. El

- Salvador  
<http://ri.ues.edu.sv/16561/1/19201082.pdf>
- [34] Investment portfolio optimization with predictive control. Pablo Andrés Deossa Molina. Universidad Nacional de Colombia. Medellín, Colombia.  
[https://www.researchgate.net/publication/264040421\\_Optimizacion\\_de\\_portafolios\\_de\\_inversion\\_aplicando\\_control\\_predictivoInvestment\\_portfolio\\_optimization\\_with\\_predictive\\_control](https://www.researchgate.net/publication/264040421_Optimizacion_de_portafolios_de_inversion_aplicando_control_predictivoInvestment_portfolio_optimization_with_predictive_control)
- [35] Modeling of Random Delays in Networked Control Systems. Yuan Ge, Qigong Chen, Ming Jiang, and Yiqing Huang. Hindawi Publishing Corporation Journal of Control Science and Engineering.  
<https://www.hindawi.com/journals/jcse/2013/383415/>
- [36] ANALISIS, ADMINISTRACIÓN Y CONTROL DE INVENTARIOS EN EMPRESAS DE MANUFACTURA (UN ENFOQUE CONCEPTUAL). Elvia Corona Medina. Universidad Autónoma de Nuevo León. México.  
<http://cdigital.dgb.uanl.mx/te/1020147068.PDF>
- [37] Single-Period Inventory Models with Discrete Demand Under Fuzzy Environment. HÜLYA BEHRET AND CENGİZ KAHRAMAN. Istanbul Technical University. TURKEY.  
<https://pdfs.semanticscholar.org/e595/0ba25bf0e778e4e6417351aef13daa99605d.pdf>
- [38] Teoría del consumidor. Pablo Serra. 2004.  
[https://www.u-cursos.cl/ingenieria/2005/1/IN701/1/material\\_docente/bajar%3Fid\\_material%3D54457](https://www.u-cursos.cl/ingenieria/2005/1/IN701/1/material_docente/bajar%3Fid_material%3D54457)
- [39] Optimización de Portafolios. Patricia Saavedra Barrera. Departamento de Matemáticas, Universidad Autónoma Metropolitana-Iztapalapa. 2008.  
<https://docplayer.es/6763175-Optimizacion-de-portafolios.html>
- [40] Stochastic Subgradient Methods. Stephen Boyd and Almir Mutapcic, with additions by John Duchi. Notes for EE364b. Stanford University. 2018.  
[https://web.stanford.edu/class/ee364b/lectures/stoch\\_subgrad\\_notes.pdf](https://web.stanford.edu/class/ee364b/lectures/stoch_subgrad_notes.pdf)
- [41] Subgradient Methods. Stephen Boyd (with help from Jaehyun Park). Notes for EE364. Stanford University. 2007.  
[https://web.stanford.edu/class/ee364b/lectures/subgrad\\_method\\_notes.pdf](https://web.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf)