Informática para la Optimización

Grado en Economía

J. Manuel Cascón, <casbar@usal.es>

Facultad de Economía y Empresa
Dpto Economía e Historia Económica
Perfil Matemáticas

15 de enero de 2018

Contenidos

Introducción

Tema 1: Introducción a la programación
Tema 2: Optimización convexa diferenciable
Tema 3: Programación Lineal
Tema 4: Programación Cuadrática
Tema 5: Optimización Dinámica

Bibliografía

Introducción

Presentación

Objetivos

Contenido

Introducción

Presentación

Objetivos

Contenido

Informática para la Optimización

- Optativa. 6 ECTS.
- Profesor:
  - J Manuel Cascón.
    `casbar@usal.es`, Despacho 226.
- Aula 105A
- Martes de $18:30$ a $21:00$

Introducción

Presentación

Objetivos

Contenido

Objetivos

El objetivo de la asignatura es presentar un enfoque conjunto de la optimización matemática y las aplicaciones económicas, orientado al uso de las nuevas tecnologías.

En ella se introduce al alumno en la algoritmia, la programación y en la teoría clásica de la optimización.

Modelo Económico $\rightarrow$ Modelo Matemático $\rightarrow$

$\rightarrow$ Implementación $\rightarrow$ Resolución

Por ejemplo,

Modelo Económico: Valoración de opciones.

Modelo Matemático: Ecuación en derivadas parciales.

Implementación: Método de diferencias finitas.

Resolución: Aplicación a un caso real.

Modelo Económico $\rightarrow$ Modelo Matemático $\rightarrow$

$\rightarrow$ Implementación $\rightarrow$ Resolución

Por ejemplo,

Modelo Económico: Valoración de opciones.

Modelo Matemático: Ecuación en derivadas parciales.

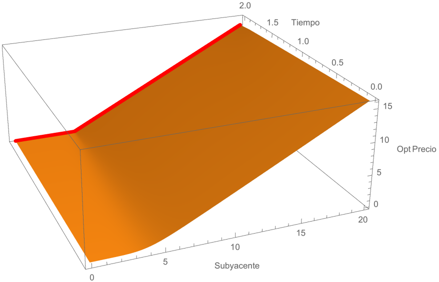
Implementación: Método de diferencias finitas.

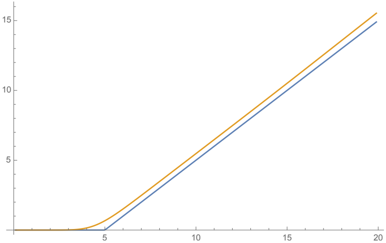
Resolución: Aplicación a un caso real.

Ejemplo: Valoración de una opción de tipo Call

Opción(Europea) es un contrato (derivado) que da al titular el derecho, pero no la obligación, a comprar (Call) o vencer (Put), el activo subyacente a un precio determinado (strike) en un fecha de vencimiento.

Valorar una opción es determinar su precio justo.





Problema Modelo

Problema Económico

Optimizar recursos

Formulación Matematica

$$
\begin{array}{ll}
\text{mín} & f(x) \\
\text{s.a} & g_1(x) \leqslant 0 \\
 & g_2(x) = 0 \\
 & x \geqslant 0
\end{array}
$$

Ejemplo: Elección de Cartera

Dada una cantidad T, se desea invertir en tres activos A, B y C: ¿cúal debería ser la distribución de la inversión para obtener el mayor rendimiento con el mínimio riesgo?

Formulación Matemática

- Modelo de Markowitz (1952)
- $\mathbf{x} = (x_A, x_B, x_C)$ vector de porcentajes.
- $\mu = (\mu_A, \mu_B, \mu_C)$ vector de rendimientos (o retornos).
- $(\sigma_A, \sigma_B, \sigma_C)$ vector de riesgos.
- V matriz de covarianzas (Riesgo).
- $\bar{\mu}$ rendimiento esperado
- Si se admiten operaciones en corto no hay restricción de no negatividad

$$
\begin{array}{ll}
\text{mín} & \frac{1}{2}\mathbf{x}'V\mathbf{x} \\
\text{s.a} & \mathbf{x}'\mu = \bar{\mu}, \\
 & \mathbf{x}'\mathbf{1} = 1, \\
 & \mathbf{x} \geqslant 0
\end{array}
$$

Ejemplo: Elección de cartera

Datos

$$\mu_A = 0.10 \quad \sigma_A = 0.07$$
$$\mu_B = 0.25 \quad \sigma_B = 0.14$$
$$\mu_C = 0.33 \quad \sigma_C = 0.28$$

$$V = \begin{pmatrix} 0.0050 & -0.0060 & 0.0020 \\ -0.0060 & 0.0199 & -0.0158 \\ 0.0020 & -0.0158 & 0.0790 \end{pmatrix}$$

Composición de la *mejor cartera* con rendimiento 0.2:

$$\begin{array}{ll} \text{mín} & \frac{1}{2}(x_A, x_B, x_C)' V (x_A, x_B, x_C) \\ \text{s.a} & x_A\mu_A + x_B\mu_B + x_C\mu_C = 0.2, \\ & x_A + x_B + x_C = 1, \\ & \mathbf{x} \geqslant 0 \end{array} \Rightarrow \begin{array}{l} x_A = 0.41 \\ x_B = 0.44, \\ x_C = 0.15 \\ \sigma_C = 0.05 \end{array}$$

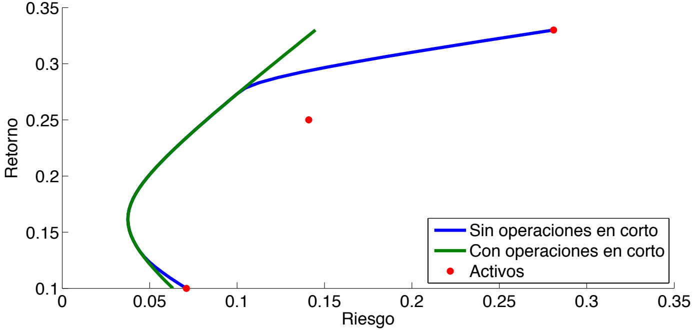
Ejemplo: Elección de cartera. Frontera Eficiente

Datos

$$\mu_A = 0.10 \quad \sigma_A = 0.07$$
$$\mu_B = 0.25 \quad \sigma_B = 0.14$$
$$\mu_C = 0.33 \quad \sigma_C = 0.28$$

$$V = \begin{pmatrix} 0.0050 & -0.0060 & 0.0020 \\ -0.0060 & 0.0199 & -0.0158 \\ 0.0020 & -0.0158 & 0.0790 \end{pmatrix}$$

Frontera eficiente



Ejemplo: Control óptimo en tiempo continuo

Un consumidor tiene una función de utilidad $U(c(t))$ donde $c(t)$ representa el consumo. La persona recibe una dotación inicial de K_0 stock de capital. Los ingresos de la persona vienen dados por iK donde i es el interés. Además la persona puede consumir el stock de capital en cualquier momento. Suponer que el consumidor es impaciente siendo δ su tasa de descuento. Se debe cumplir además $K(t) \geqslant 0$.

Matematicamente:

$$\max_{c(t)} \quad \int_0^T e^{-\delta t} U(c(t))dt$$

s.a.

$$\begin{aligned}
&K'(t) = iK(t) - c(t) && t \in (0, T) \\
&K(0) = K_0, \quad K(t) \geqslant 0 && t \in (0, T) \\
&c(t) \geqslant 0 && t \in (0, T)
\end{aligned}$$

Ejemplo: Control óptimo en tiempo continuo

Un consumidor tiene una función de utilidad $U(c(t))$ donde $c(t)$ representa el consumo. La persona recibe una dotación inicial de K_0 stock de capital. Los ingresos de la persona vienen dados por iK donde i es el interés. Además la persona puede consumir el stock de capital en cualquier momento. Suponer que el consumidor es impaciente siendo δ su tasa de descuento. Se debe cumplir además $K(t) \geqslant 0$.

Matematicamente:

$$\max_{c(t)} \quad \int_0^T e^{-\delta t} U(c(t)) dt$$

s.a.

$$\begin{aligned}
&K'(t) = iK(t) - c(t) && t \in (0, T)\\
&K(0) = K_0, \quad K(t) \geqslant 0 && t \in (0, T)\\
&c(t) \geqslant 0 && t \in (0, T)
\end{aligned}$$

Introducción

Presentación

Objetivos

Contenido

Contenido

TEMA 1 Introducción a al programación Informática

TEMA 2 Optimización Convexa Diferenciable

TEMA 3 Programación Lineal

TEMA 4 Programación Cuadrática

TEMA 5 Optimización Dinámica

Contenido

TEMA 1 Introducción a al programación Informática

Introducción a la informática. Algoritmos. Programación. Pseudocódigo. Eficiencia y Complejidad. Introducción a Matlab/Mathematica.

TEMA 2 Optimización Convexa Diferenciable

TEMA 3 Programación Lineal

TEMA 4 Programación Cuadrática

TEMA 5 Optimización Dinámica

Contenido

TEMA 1 Introducción a al programación Informática

TEMA 2 Optimización Convexa Diferenciable
Conjuntos convexos. Funciones convexas. Teorema de Lagrange. Teorema de Karush-Kuhn-Tucker. Condiciones de segundo orden. Introducción a Mathematica

TEMA 3 Programación Lineal

TEMA 4 Programación Cuadrática

TEMA 5 Optimización Dinámica

Contenido

TEMA 1 Introducción a al programación Informática

TEMA 2 Optimización Convexa Diferenciable

TEMA 3 Programación Lineal

Programación Lineal. Método Simplex. Dualidad. Precios sombra. Análisis de sensibilidad. Juegos matriciales. Excel/Solver.

TEMA 4 Programación Cuadrática

TEMA 5 Optimización Dinámica

Contenido

TEMA 1 Introducción a al programación Informática

TEMA 2 Optimización Convexa Diferenciable

TEMA 3 Programación Lineal

TEMA 4 Programación Cuadrática
Condiciones KKT en problemas cuadrátticos. Algoritmo de Newton. Métodos de punto interior (predicción corrección). Problemas paramétricos. Gestión de cartera. Resolución con Matlab.

TEMA 5 Optimización Dinámica

Contenido

TEMA 1 Introducción a al programación Informática

TEMA 2 Optimización Convexa Diferenciable

TEMA 3 Programación Lineal

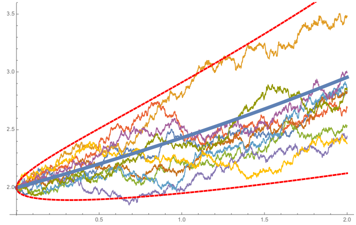
TEMA 4 Programación Cuadrática

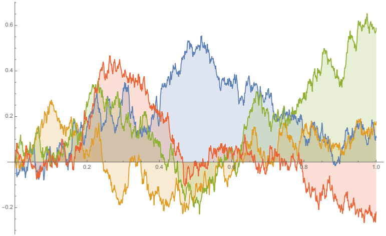
TEMA 5 Optimización Dinámica
Introducción a la optimización dinámica.
Control óptimo en tiempo discreto.
Principio de optimalidad de Bellman.
Control ótimo en tiempo continuo.
Principio del máximo de Pontryagin.
Aplicaciones económicas. Resolución con Mathematica/Matlab.

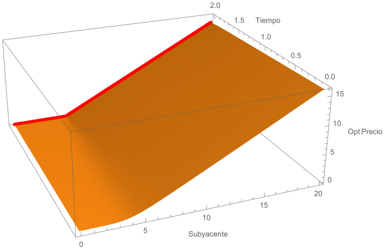
Por qué Mathematica

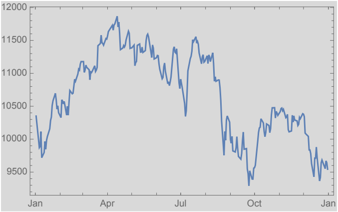
- Porque es una herramienta muy potente, pero al mismo tiempo muy fácil de manejar
- Porque se adapta muy bien a las características del curso: cálculo simbólico
- Porque dispone de un gran soporte gráfico
- Porque es una herramienta de amplia difusión: empresa, banca, investigación, etc
- Porque la USAL proporcional licencia para estudiantes

Ejemplos con Mathematica









Por qué Excel

- Porque es una herramienta muy sencilla y de gran potencia (VisualBasic)
- Porque está disponible en todos los equipos
- Porque lo usan todas las empresas
- Porque con toda seguridad lo usaréis después de vuestro paso por la Universidad.
- Porque incorpora el paquete Solver para la resolución de problemas de optimización

Ejemplos con Excel

4	A) Se desea invertir en los productos financieros que se muestran en la tabla. En la misma tabla se incluyen las rentabilidades anuales ofrecidas. Una prestigiosa agencia de valoración califica la solvencia de cada producto de 1 a 5 (1 poca solvencia, 5 máxima solvencia). Se desea obtener la máxima rentabilidad garantizando que la solvencia ponderada será de al menos 3,8. Del total a invertir se está dispuesto a invertir para cada producto los porcentajes máximos y mínimos indicados en la tabla				
5	(Nota: Solvencia ponderada = ∑ Fracción invertida x Calificación solvencia).				
6	Producto	*Rentabilidad anual*	*Calificación solvencia*	*Máximo a invertir*	*Mínimo a invertir*
7	Fondo Asiatico	12%	*1*	10%	5%
8	Fondo Emergente	10%	2	10%	5%
9	Fondo Mixto	8%	3	20%	10%
10	Bonos USA	5%	4	40%	25%
11	Bonos EURO	3%	5	40%	25%
12					

Ejemplos con Excel

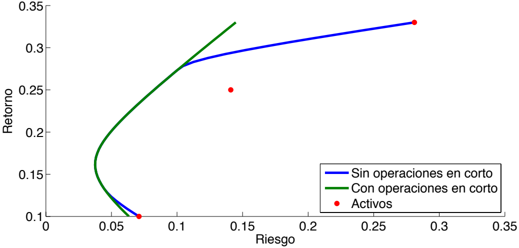
5	Solucion						
6	*Xi*	Producto	Rentabilidad anual	Calificación solvencia	Máximo a invertir	Mínimo a invertir	Xi
7	X1	Fondo Asiatico	12,00%	1	10%	5%	6,25%
8	X2	Fondo	10,00%	2	10%	5%	10,00%
9	X3	Fondo Mixto	8,00%	3	20%	10%	20,00%
10	X4	Bonos USA	5,00%	4	40%	25%	25,00%
11	X5	Bonos EURO	3,00%	5	40%	25%	38,75%
12				3,80			**1,00**
13		Rentabilidad anual:	5,76%	3,80			1,00

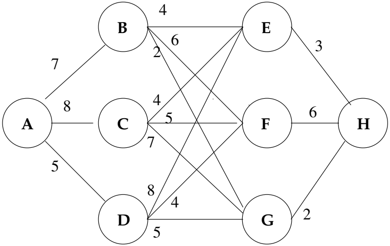
Por qué Matlab

MATLAB
The Language of Technical Computing

- Porque trabaja de forma muy eficiente con datos matriciales
- Porque iniciarse en la programación es muy sencillo
- Porque es una herramienta de gran difusión: ingeniería, simulación, etc.

Ejemplos con Matlab





Tema 1

Introducción a la programación

Introducción a la Informática

Algoritmo y Pseudocódigo

Eficiencia y Complejidad

Tema 1

Introducción a la programación

Introducción a la Informática
- Hardware
- Software
- Sistemas Operativos
- Computación en la nube

Algoritmo y Pseudocódigo

Eficiencia y Complejidad

Introducción

Un sistema informático es una máquina o conjunto de máquinas formadas por elementos físicos (*hardware*) y elementos lógicos (*software*) capaces de almacenar información y de tratarla de forma automática siguiendo un conjunto de instrucciones que pueden ser previamente programadas.

Información $\Rightarrow$ Procesamiento $\Rightarrow$ Resultado

Introducción

En el estudio de los sistemas informáticos se distinguen dos grandes apartados:

Hardware: conjunto de elementos mecánicos y electrónicos que conforman la estructura física de un equipo informático

Software: conjunto de programas de diferentes niveles que controlan la acción del ordenador. Estructura Lógica de un sistema informático.

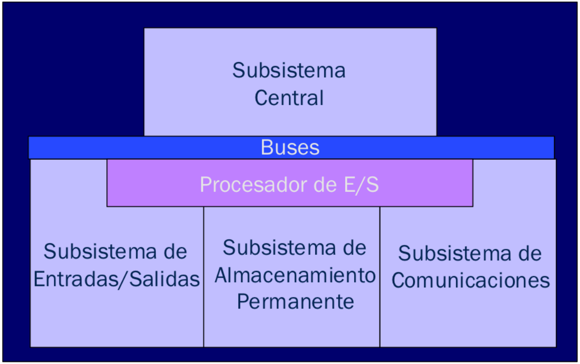
Arquitectura de un ordenador

John von Neumann divide el hardware de un ordenador en cinco grupos principales:

1. CPU (Unidad Central de Proceso).
2. Perifericos de entrada.
3. Perifericos de salida.
4. Memoria de trabajo.
5. Memoria permanente.

Arquitectura de un ordenador

En la actualidad es común distinguir 5 subsistemas:



Dispositivos del Sistema

Subsistema Central

- **CPU**: chip de silicio con varios millones de transistores. La velocidad de operación se mide en hertzios. (Velocidad actual ~ 4Ghz)

- **Memoria RAM** (*Random Access Memory*): chips que funciona como un bloque de celdas de almacenamiento. Memoria volátil.

- **Otras Memorias**: ROM-BIOS, CMOS, CACHE

Dispositivos del Sistema

Subsistema de Entradas/Salidas

Formado por todos aquellos dispositivos que permiten para intoducir o/y obtener información. Entre otros:

- teclado
- joystick
- escáner
- lector de códigos de barras
- lápiz óptico
- cámara de video
- ratón
- monitor
- impresora
- altavoces
- plotter
- tarjeta gráfica
- etc. . .

Dispositivos del Sistema

Subsistema de Almacenamiento Permanente

Dispositivos que permiten el almacenamiento y recuperación de información. Almacenamiento másivo

Tipos de dispositivos:

- **Dispositivos magnéticos**: Se basan en propiedades magnéticas de ciertos materiales para el registro de datos. Discos duros ($\sim$ 1Tb)
- **Dispositivos ópticos**: Se basan en el uso de un láser para el almacenamiento y la recuperación de la información. CD ($\sim 640 - 700$ Mg), DVD ($4,7 - 17$ Gb), Blue-Ray ($27 - 54$Gb)
- **Memoria flash**: memoria no volátil EEPROM. *Electrically-Erasable Programmable Read-Only Memory*. Dispositivos de estados sólido, Memoria USB. Capacidad máxima ~ 512 GB

Dispositivos del Sistema

Subsistema de Comunicaciones

- **Tarjetas de Red**: Son dispositivos que permiten conectar el ordenador a una red de área local (LAN). El estándar más popular es Ethernet con velocidades de 100Mb/s, 1Tb/s.

- **Router**: dispositivos que permiten la interconexión de redes y el intercambio de datos entre ellas.
- **Modem**: convierte las senãles digitales para poder ser transmitidas a través de líneas analógicas. No

Dispositivos del Sistema

Subsistema de Comunicaciones

- **Tarjetas de Red**: Son dispositivos que permiten conectar el ordenador a una red de área local (LAN). El estándar más popular es Ethernet con velocidades de 100Mb/s, 1Tb/s.
- **Router**: dispositivos que permiten la interconexión de redes y el intercambio de datos entre ellas.
- **Modem**: convierte las senãles digitales para poder ser transmitidas a través de líneas analógicas. No se debe confundir con el router aunque habitualmente forma parte de él.

Dispositivos del Sistema

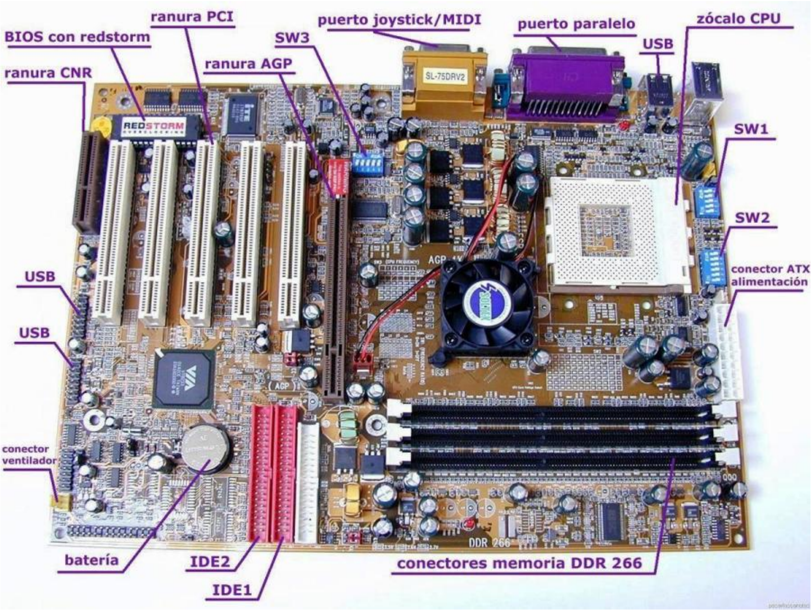
Buses y Procesador de entradas/salidas

Bus: Soporte físico para la transmisión de datos, instrucciones y controles.

El sistema de buses conecta todo el sistema informático.

La CPU y la memoria RAM están conectados por un bus particularmente potente denominado FSB (*Front Side Bus*).

Todos los buses están regulados por una serie de controladores, chips que manejan la transferencia de datos.


ranura PCI
BIOS con redstorm
ranura CNR
puerto joystick/MIDI
SW3
ranura AGP
puerto paralelo
USB
zócalo CPU
SW1
SW2
conector ATX alimentación
USB
USB
conector ventilador
batería
IDE2
IDE1
conectores memoria DDR 266

Tipos de ordenadores

- **Microordenador**: Ordenadores de pequeño tamaño para uso personal (PC).
- **Miniordenador**: Ordenadores de tamaño medio que dan servicios a múltiples usuarios. Servidores de red, de correo, aplicaciones compartidas.
- **MainFrame**: Grandes ordenadores con gran capacidad de almacenamiento y con altas prestaciones. Gestión de grandes bases de datos. Ordenadores centrales de grandes corporaciones.
- **Superordenador** Sistemas de gran potencia y elevadísimas prestaciones destinados a la investigación y la realización de cálculos científicos.

Software

Clasificación del software:

- **Sistemas Operativos**: Conjunto de programas que gestionan el hardware y dan soporte al resto de programas.
- **Lenguajes de Programación**: Programas que permiten el desarrollo de nuevas aplicaciones.
- **Aplicaciones Informáticas** : Programas que permiten resolver un proceso concreto y especializado.

Sistemas Operativos

Características Importantes de un SO

- **Seguridad:** Integridad, disponibilidad y confidenciabilidad de los datos
- **Eficiencia:** Rendimiento incluso con grandes cantidades de información
- **Estabilidad:** Tolerancia a fallos de componentes *hardware* o *software*
- **Administración:** Políticas de uso de los recursos
- **Flexibilidad:** Platamormas múltiples
- **Concurrencia:** Acceso simultáneo de múltiples usuarios y ejecución de múltiples tareas

Sistemas Operativos

- **Microsoft Windows**: surege como entorno gráfico para MS-DOS. Multitarea y monousuario (excepto versiones de red).
- **Unix**: diseñado en los laboratorios Bell (AT&T, 1971). Multitarea y multusuario. Robusto.
- **Linux**: Creado por Linus Tordvals y Richard Stallman (1991). Multitarea y multiusuario. Licencia GNU GPL. Multitud de distribuciones: Ubuntu, Debian, Suse, RedHat,
- **Mac OS**: desarrollado y comercializado por Apple. Las versiones actuales se basan en Unix.
- **Telefonía móvil**: Android, iOS

Lenguajes de Programación

- **Algoritmo**: Secuencia finita y ordenada de reglas bien definidas que permiten resolver un problema.

- **Programa**: Algoritmo escrito en un lenguaje inteligible por un ordenador

- **Lenguaje de programación**: lenguaje (albafeto, vocabulario, reglas sintácticas y semánticas) que permite la escritura de un algoritmo.

- **Clasificación de los lenguajes de programación**
  - Bajo nivel
  - Alto nivel

Lenguajes de Programación

Algoritmo

Programa

```
if not  funciona (lampara)
   if not  enchufada (lamapara)
               enchufar (lampara)
   else if  quemado (lampara.foco)
               reemplazar (lampara
   else
               comprar (lampara)
```

Lenguajes de Programación

Tipos de Lenguajes

- **Bajo Nivel**: Lenguaje relacionado con el diseño del hardware.
  - Lenguaje Máquina: codificado en binario. Depende del procesador.
  - Lenguaje Ensamblador: Lenguaje simbólico. Uso en situaciones muy concretas.

- **Alto Nivel**: Independiente del hardware. Lenguaje natural (cercano al humano).
  - Compiladores: Traduce el programa a lenguaje máquina antes de ejecutarlo. Generan código más rápido y eficiente.
  - Interpretes: Ejecuta directamemente el programa fuente. La detección de errores es más sencilla. En general, producen códigos más lentos.

Lenguajes de Programación

Algunos lenguajes de alto nivel

- **FORTRAN**: 1956, surge como un lenguaje de propósito general diseñado para aplicaciones científicas y técnicas.
- **COBOL**: 1957-60 lenguaje para aplicaciones comerciales y económicas
- **BASIC**: 1964. Fines didácticos. Ha sufrido múltiples evoluciones.
- **PASCAL**: 1970. Aprendizaje de estructuras de programación concretas y su uso en diferentes metodologías. Admite recursividad
- **C**: Década de los 70. Potente y rápido, además de modular e independiente de la máquina. Admitiendo la recursividad. Su versión para la programación orientada a objetos es C++ (80s).
- **JAVA**: 1990-94. Evolución de los LOO. Uso reorientado hacia la web.
- **PHP**: 1994. Programación web con bases de datos.
- **JAVASCRIPT**: 1995. Lenguaje orientado a la programación web.

Aplicaciones Informáticas

El software de aplicación está formado por el conjunto de programas que realizan un trabajo específico para responder a las necesidades de los usuarios.

- **Sw específico**: CAD/CAM, diseño gráfico, creación multimedia,. . .
- **Sw de productividad**: Planificación personal, procesadores de textos, hojas de cálculo, bases de datos, presentaciones, . . .
- **Sw de comunicaciones**: correo electrónico, mensajería instántanea, videoconferencia, . . .
- **Sw de gestión**: gestión de clientes, contabilidad, facturación, nóminas, . . .
- **Sw de seguridad**: control de accesos, antivirus, firewalls, . . .

Computación en la nube

- La nube es un nuevo paradigma en el que el software y el hardware se convierten en un servicio
- Bajo esta denominación se engloba al uso remoto de aplicaciones e infraestructuras no instaladas físicamente en el ordenador del usuario
- El acceso al servicio es más flexible ya que se puede realizar desde cualquier dispositivo y aporta, además, un valor añadido: en el uso colaborativo
- Las ventajas de su uso (seguridad, acceso flexible,. . .) suponen, al mismo tiempo, una serie de inconvenientes (falta de seguridad y de control)

Tema 1

Introducción a la programación

Introducción a la Informática

Algoritmo y Pseudocódigo
- Introducción
- Algoritmo
- Pseudocódigo
- Ejemplos
- Ejercicios

Eficiencia y Complejidad

Introducción

Problema vs Instancia

Fases para la Resolución de un Problema:

1. Fase de Identificación
   - Reconocimiento del problema

2. Fase de Resolución
   - Análisis del Problema
   - Diseño del Algoritmo
   - Verificación

3. Fase de Implementación
   - Codificación
   - Verificación

Introducción

Problema vs Instancia

Fases para la Resolución de un Problema:

1. Fase de Identificación
   - Reconocimiento del problema
2. Fase de Resolución
   - Análisis del Problema
   - Diseño del Algoritmo
   - Verificación
3. Fase de Implementación
   - Codificación
   - Verificación

Introducción

Problema vs Instancia

Fases para la Resolución de un Problema:

1. Fase de Identificación
   - Reconocimiento del problema

2. Fase de Resolución
   - Análisis del Problema
   - Diseño del Algoritmo
   - Verificación

3. Fase de Implementación
   - Codificación
   - Verificación

Introducción

Problema vs Instancia

Fases para la Resolución de un Problema:

1. Fase de Identificación
   - Reconocimiento del problema

2. Fase de Resolución
   - Análisis del Problema
   - Diseño del Algoritmo
   - Verificación

3. Fase de Implementación
   - Codificación
   - Verificación

Algoritmo

Definición (Algoritmo): Secuencia ordenada, bien definida y finita de instrucciones que permite resolver un problema.

- Ordenada : Las instrucciones ocupan un lugar preciso en el código.
- Bien Definida : Las instrucciones son unívocas. Dos ejecuciones distintas del algoritmo producen el mismo resultado.
- Finita : número finito de instrucciones. El algoritmo siempre termina.

Diseño un Algoritmo

- **Entrada**: Datos proporcionados
- **Proceso**: Operaciones o cálculos para la resolución del problema
- **Salida**: Resultados

Ejemplo : Cálculo del área de un rectángulo.

- Entrada: base y altura
- Proceso: Área = base $\times$ altura
- Salida: Área

Diseño un Algoritmo

- **Entrada**: Datos proporcionados
- **Proceso**: Operaciones o cálculos para la resolución del problema
- **Salida**: Resultados

Ejemplo : Cálculo del área de un rectángulo.

- Entrada: base y altura
- Proceso: Área = base $\times$ altura
- Salida: Área

Pseudocódigo

Definición (Pseudocódigo): Es un lenguaje semi-formalizado con reglas semejantes a las de un lenguaje de programación, pero sin su rígidez que permite codificar un algoritmo.

Cada una de las unidades elementales en las que se divide un pseudocódigo se denomina sentencia.

En una sentencia pueden aparecer:

- Datos (constantes, variables).
- Operadores (aritméticos, relacionales, lógicos, asignación).
- Expresiones.
- Estructuras de control o instrucciones (secuenciación, repetitivas, alternativas, entrada-salida, otras funciones).

Datos

- **Constantes:** Son datos cuyo valor no cambia durante la ejecución del programa. Pueden ser literales o con nombre.

- **Variables:** Son datos (posición de memoria) cuyo valor puede cambiar durante la ejecución del programa.

Las variables o las constantes con nombre se identifican por medio de un nombre o identificador

Un identificador debe verificar las siguientes reglas:

- Es una secuencia de letras o dígitos. (No se admiten: $+, -, *, /$, &, $\|$, etc...). Se admite _
- Debe comenzar con una letra.
- No puede ser el nombre de una *palabra reservada*

Tipos de Datos

- Básicos:
  - Numéricos: entero, real, etc,
  - Lógicos: booleano.
  - Carácter: carácter, cadena.
- Derivados:
  - Vectores, matrices.
  - Estructuras.

Operadores

Un **operador** es un carácter o grupo de caracteres que actúa sobre una o más variables para realizar una operación y producir un resultado.

Tipos de Operadores

- Aritméticos: suma (+), resta(-), multiplicación (*), división (/).
- Asignación: (=) asigna el valor de una expresión a una variable.

```
nombre_variable = expresión
```

- Relacionales: mayor (¿), menor(¡), mayor o igual (¿=), menor o igual (¡=), igual (==), distinto (~=)
- Lógicos: conjunción (AND, &), diyunción (OR, |), negación (NEG, ~)

Expresiones

Una **expresión** es una combinación de variables y/o constantes y operadores evaluadas durante la ejecución y que produce un resultado.

Ejemplos:

```
beneficio = precioVenta - costeProduccion

a ~= 2;

(nOfIter <10) & (error>1.e-6)
```

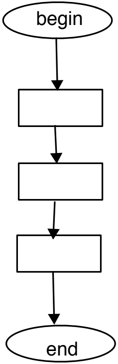
Estructuras de Control o Instrucciones

Las **estructuras de control** o **instrucciones** son una serie de reglas que definen el flujo de ejecución del programa y permiten su modificación.

Tipos de instrucciones

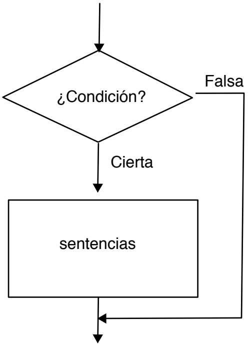
- Secuenciación: **begin**, **;**, **end**
- Alternativas: **if them else**, **select case**
- Repetitivas: **for to do**, **while**, **do while**

Instrucciones de secuenciación



- **begin** marca el incio del pseudocódigo.
- **;** permite separar sentencias.
- **end** marca el final del pseudocódigo.

Instrucciones alternativas



if them

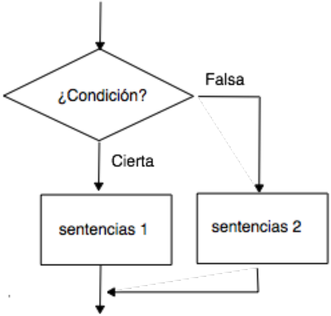
- Sintaxis:

  if *(condición)* **them**
  sentencias;
  endif

- Descripción
  Si la condición es cierta se evalúa el grupo de sentencias, en otro caso se continúa la ejecución del programa.

Instrucciones alternativas

if them else



- <u>Sintaxis</u>:

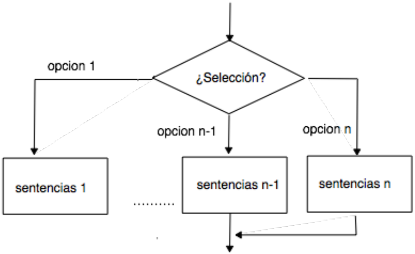
  if *(condición)* **them**
  sentencias1;
  else
  sentencias2;
  endif

- <u>Descripción</u>
  Si la condición es cierta se evalúa el grupo de sentencias 1 en otro caso se evalúa el grupo de sentencias 2.

Nota: Es posible construir una estructura con bucles **if them else** anidados.

Instrucciones alternativas

select case



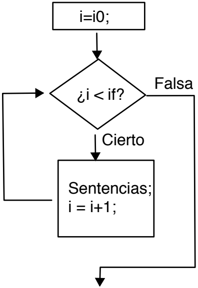
- Sintaxis:

  select *(opción)*
  case *1:*
  sentencias1;
  . . . . . .
  case *n-1:*
  sentencias n-1;
  case *n:*
  sentencias n;
  endselect

- Descripción
  Según el valor de *opción* se evalúa el grupo de sentencias correspondiente

Instrucciones repetitivas

for to do



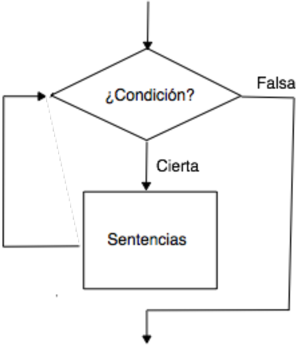
- Sintaxis:

  for $i = i_0$ **to** i_f **do**
  sentencias;
  endfor

- Descripción

  1. Se inicia la variable i a i_0
  2. Si $i > i_f$ el bucle termina.
  3. Se ejecuta el grupo de sentencias.
  4. Se incrementa en una unidad el valor de i y se vuelve al paso 1.

Instrucciones repetitivas



while

- Sintaxis:

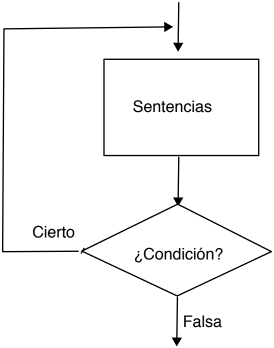
  while *(condición)*
  sentencias;
  endwhile

- Descripción
  Mientras sea cierto el contenido de *condición* se evalúa repetidamente el grupo de sentencias.

Nota: El valor de la condición debe modificarse en el grupo de sentencias.

Instrucciones repetitivas

while



- Sintaxis:

  do

  sentencias;

  while *(condición);*

- Descripción

  Se repite la evaluación del grupo de sentencias mientras la condición sea cierta.

Nota:

- El valor de la condición debe modificarse en el grupo de sentencias.
- En un bucle **do while** el grupo de sentencias al menos se evalúa una vez.

Entrada -Salida y otras funciones

- Las instrucciones
  de entrada-salida se escriben en lenguaje *ordinario*:

  read();
  write();

- En pseudocódigo se asume el conocimiento de funciones básicas:

  cos();
  sin();
  tan();
  sqrt();
  abs();
  log();
  exp();
  ...

Notas finales

- Al inicio del pseudocódigo se suele añadir una línea especificando los datos de entrada:

  INPUT: . . .

- Es recomendable utilizar programación modular: dividir el programa en acciones elementales, que se agrupan formando funciones o subrutinas.
- En pseudócodigo el acceso a los elementos de un vector o matriz se hace a través de subíndices, al igual que en matemáticas.

Ejemplo 1

Ejemplo: Escribir el pseudocódigo de un programa que calcule la suma de los n primeros número naturales:

INPUT: $n \in \mathbb{N}$
begin
read(n);
$s = 0$;
for *i=1* **to** n **do**
    $s = s + i$;
endfor
write(s);
end

Ejemplo 2

Ejemplo: Escribir el pseudocódigo de un programa que calcule la media de un vector:

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$
begin
read(n,a);
$s = 0$;
for *i=1* **to** n **do**
    $s = s + a_i$;
endfor
$m = s/n$;
write(m);
end

Ejemplo 3

Ejemplo: Escribir el pseudocódigo de un programa que calcule el mínimo de n números dados:

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$
begin
read(n,a);
$m = a_1$;
for $i = 2$ **to** n **do**
  if $m > a_i$ **them**
    $m = a_i$;
  endif
endfor
write(m);
end

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo

8	2	5	1	4

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 1 cambio!

8	2	5	1	4

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo: Etapa 1

2	8	5	1	4

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 1 cambio!

2	8	5	1	4

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo: Etapa 1

2	5	8	1	4

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n - 1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n - i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 1 cambio!

2	5	8	1	4

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n - 1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n - i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 1

2	5	1	8	4

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 1 cambio!

2	5	1	8	4

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 1

2	5	1	4	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo: Etapa 2 correcto!

2	5	1	4	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 2 cambio!

2	5	1	4	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n - 1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n - i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 2

2	1	5	4	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo: Etapa 2 cambio!

2	1	5	4	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 2

2	1	4	5	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo: Etapa 3 cambio

2	1	4	5	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 3

1	2	4	5	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n - 1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n - i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 3 correcto!

1	2	4	5	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 3 correcto!

1	2	4	5	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n-1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n-i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 4 correcto!

1	2	4	5	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

Método de la burbuja (Bubble Sort)

- El algorimto realiza $n - 1$ etapas (barridas).
- Al finalizar la etapa i se garantiza que están bien ordenados los $n - i$-últimos términos.
- En la etapa i-ésima se recorre la lista desde la posición i hasta la posición n, comparando 2 términos consecutivos, si su orden no es correcto, se invierten sus posiciones.

Ejemplo : Etapa 4 correcto!

1	2	4	5	8

Ejemplo 4

Ejemplo: Escribir el pseudocódigo de un programa que ordene de menor a mayor n números dados:

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$
begin
read(n,a);
for $i = 1$ **to** $n - 1$ **do**
  for $j = 1$ **to** $n - i$ **do**
    if $a_j > a_{j+1}$ **them**
      $temp = a_j;$
      $a_j = a_{j+1};$
      $a_{j+1} = temp;$
    endif
  endfor
endfor
write(a);
end

Ejercicios

1. Escribir el pseudocódigo de un algoritmo que calcule los cuadrados de todos los números naturales impares menores o iguales que $n \in \mathbb{N}$.
2. Escribir el pseudocódigo de un algoritmo que calcule todos los números naturales de tres cifras tales que sean iguales a la suma de los cubos de sus tres cifras.
3. Escribir el pseudocódigo de un algoritmo que calcule el factorial de un número natural.
4. La raíz cuadrada de un número positivo a, puede ser aproximada por la sucesión:

$$x_0 = 1; \qquad x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right);$$

Se considera que un término es una buena aproximación de $\sqrt{a}$, cuando $|x_n - x_{n-1}| < tol$. Escribir un algoritmo que use esta sucesión para calcular la raíz de a.

Tema 1

Introducción a la programación

Introducción a la Informática

Algoritmo y Pseudocódigo

Eficiencia y Complejidad
- Introducción
- Eficiencia
- Complejidad

Introducción

Cálculo del cuadrado de un número $n \in N$

INPUT: $n \in \mathbb{N}$,
begin
read(n);
$c = 0$;
for $i = 1$ **to** n **do**
    $c = c + n$;
endfor
write(c);
end

INPUT: $n \in \mathbb{N}$,
begin
read(n);
$c = n * n$;
write(c);
end

¿Qué algoritmo es *mejor* ?

Introducción

- El objetivo de está sessión es proporcionar una herramienta que permita analizar con rigor la *bondad* de un algoritmo.
- Se estudiarán los recursos que consume un algoritmo: memoria y tiempo.
- El cáculo de la eficiencia de un algoritmo permite establecer un orden entre los diferentes algoritmos que resuelven un mismo problema.
- La complejidad de un problema mide su dificultad. Esta relacionada con la eficiencia del mejor algoritmo que es capaz de resolverlo.

Noción de Eficiencia

Definición (Algoritmo): Secuencia precisa, bien definida y finita de instrucciones que permite resolver un problema.

Definición (Eficiencia): Es la medida de los recursos que consume un algoritmo en el peor de los casos en función del tamaño de la instancia.

- Recursos: tiempo (eficiencia temporal) y memoria (eficiencia espacial).
- Peor de los Casos: se mide el caso más desfavorable. (En ocasiones también se habla de eficiencia media).
- Tamaño de la instancia: La eficiencia se parametriza en función del tamaño de la entrada.

Noción de eficiencia

Factores de los que depende la ejecución

- Externos: Máquina, Compilador
- Internos: Número de instrucciones asociadas al algoritmo. Independiente de la máquina.

Estudio del tiempo de ejecución

- Análisis empírico (*a posterior*): Se generan ejecuciones del algoritmos para distintos valores de entrada y se mide el tiempo de ejecución. Se obtiene una medida real, pero depende de factores internos y externos.
- Análisis análitico (*a priori*): se determina una función que represente el tiempo de ejecución del algoritmo. Sólo depende de factores internos.

Función de Eficiencia Temporal

$$\begin{aligned} f : \mathbb{N}^* &\longrightarrow \mathbb{N}^* \\ n &\longrightarrow f(n) \end{aligned}$$

Cálculo de número de operaciones

- Se consideran las operaciones aritméticas, lógicas y relacionales.
- No se consideran las operaciones de asignación (son rápidas)
- No se consideran las operaciones de entrada-salida (son muy pocas).
- En estructuras de selección

$$f(s_{if}) = f(c) + \max\{f(s1), f(s2)\}$$

- En estructuras repetitivas

$$f(s_{while}) = f(c) + (n^o iter) * (f(c) + f(s)),$$
$$f(s_{for}) = (n^o iter) * (f(c) + f(s))$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$	–
begin	–
read(n);	–
$s = 0$;	–
for i=1 **to** n **do**	i=1 — n
$\quad s = s + i$;	1 op.
	2 op. téc.
endfor	
write(s);	–
end	–

$$f(n) = n(1+2) = 3n$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$	–
begin	–
read(n);	–
$s = 0$;	–
for i=1 **to** n **do**	i=1 — n
$s = s + i$;	1 op.
	2 op. téc.
endfor	
write(s);	–
end	–

$$f(n) = n(1+2) = 3n$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$	–
begin	–
read(n);	–
$s = 0$;	–
for i=1 **to** n **do**	i=1 — n
$\quad s = s + i$;	1 op.
	2 op. téc.
endfor	–
write(s);	–
end	–

$$f(n) = n(1 + 2) = 3n$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$	–
begin	–
read(n,a);	–
$s = 0$;	–
for i=1 **to** n **do**	i=1 — n
$\quad s = s + a_i$;	1 op.
endfor	2 op. téc.
m = s/n;	1 op.
write(m);	–
end	–

$$f(n) = n(1+2) + 1 = 3n + 1$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$	–
begin	–
read(n,a);	–
$s = 0$;	–
for i=1 **to** n **do**	i=1 — n
$\quad s = s + a_i$;	1 op.
endfor	2 op. téc.
m = s/n;	1 op.
write(m);	–
end	–

$$f(n) = n(1+2) + 1 = 3n + 1$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$	–
begin	–
read(n,a);	–
$s = 0$;	–
for i=1 **to** n **do**	i=1 — n
$\quad s = s + a_i$;	1 op.
	2 op. téc.
endfor	
m = s/n;	1 op.
write(m);	–
end	–

$$f(n) = n(1+2) + 1 = 3n + 1$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$ —
begin —
read(n,a); —
$m = a_1$; —
for $i = 2$ **to** n **do** i=2 — n
    if $m > a_i$ **them** 1 op.
        $m = a_i$; —
    endif —
endfor 2 op. téc.
write(m); —
end —

$$f(n) = (n-1)(1+2) = 3n-3$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$	–
begin	–
read(n,a);	–
$m = a_1$;	–
for $i = 2$ **to** n **do**	i=2 — n
if $m > a_i$ **them**	1 op.
$m = a_i$;	–
endif	–
	2 op. téc.
endfor	–
write(m);	–
end	

$$f(n) = (n-1)(1+2) = 3n-3$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$	–
begin	–
read(n,a);	–
$m = a_1$;	–
for $i = 2$ **to** n **do**	i=2 — n
if $m > a_i$ **them**	1 op.
$m = a_i$;	–
endif	–
endfor	2 op. téc.
write(m);	–
end	–

$$f(n) = (n-1)(1+2) = 3n-3$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$ –
begin –
read(n,a); –
for $i = 1$ **to** $n - 1$ **do** i=1 – n-1
  for $j = 1$ **to** $n - i$ **do** j=1 – n-i
    if $a_j > a_{j+1}$ **them** 2 op.
      $temp = a_j$; –
      $a_j = a_{j+1}$; 1 op.
      $a_{j+1} = temp$; 1 op.
    endif –
  endfor 2op. téc.
endfor 2 op. téc.
write(a); –
end –

$$f(n) = \sum_{i=1}^{n-1}\left(2 + \sum_{j=1}^{n-i}(2+1+1+2)\right) = \sum_{i=1}^{n-1} 2 + 6(n-i) =$$

$$= 2(n-1) + 6\sum_{i=1}^{n-1}(n-i) = 2(n-1) + 6\frac{(n-1+1)(n-1)}{2} = 3n^2 - n - 2$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$	–
begin	–
read(n,a);	–
for $i = 1$ **to** $n - 1$ **do**	i=1 – n-1
for $j = 1$ **to** $n - i$ **do**	j=1 – n-i
if $a_j > a_{j+1}$ **them**	2 op.
$temp = a_j$;	–
$a_j = a_{j+1}$;	1 op.
$a_{j+1} = temp$;	1 op.
endif	–
endfor	2op. téc.
endfor	2 op. téc.
write(a);	–
end	–

$$f(n) = \sum_{i=1}^{n-1}\left(2 + \sum_{j=1}^{n-i}(2+1+1+2)\right) = \sum_{i=1}^{n-1} 2 + 6(n-i) =$$

$$= 2(n-1) + 6\sum_{i=1}^{n-1}(n-i) = 2(n-1) + 6\frac{(n-1+1)(n-1)}{2} = 3n^2 - n - 2$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$	–
begin	–
read(n,a);	–
for $i = 1$ **to** $n - 1$ **do**	i=1 – n-1
for $j = 1$ **to** $n - i$ **do**	j=1 – n-i
if $a_j > a_{j+1}$ **them**	2 op.
$temp = a_j$;	–
$a_j = a_{j+1}$;	1 op.
$a_{j+1} = temp$;	1 op.
endif	–
endfor	2op. téc.
endfor	2 op. téc.
write(a);	–
end	–

$$f(n) = \sum_{i=1}^{n-1}\left(2 + \sum_{j=1}^{n-i}(2+1+1+2)\right) = \sum_{i=1}^{n-1} 2 + 6(n-i) =$$

$$= 2(n-1) + 6\sum_{i=1}^{n-1}(n-i) = 2(n-1) + 6\frac{(n-1+1)(n-1)}{2} = 3n^2 - n - 2$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $A, B \in \mathbb{R}^{n\times n}$ –
begin –
read(n,$\{a\}_{ij}$,$\{b\}_{ij}$); –
for $i = 1$ **to** n **do** i=1 – n
  for $j = 1$ **to** n **do** j=1 – n
    $c_{ij} = 0$ –
    for $k = 1$ **to** n **do** k=1 – n
      $c_{ij} = c_{ij} + a_{ik} * b_{kj}$ 2 op.
    endfor 2 op. téc.
  endfor 2 op. téc.
endfor 2 op. téc.
write($\{c\}_{ij}$); –
end –

$$f(n) = \sum_{i}^{n}\left(2 + \sum_{j=1}^{n}\left(2 + \sum_{k=1}^{n} 2 + 2\right)\right) = \sum_{i}^{n}\left(2 + \sum_{j=1}^{n}(2+4n)\right)$$

$$= \sum_{i}^{n}\left(2 + 2n + 4n^2\right) = 2n + 2n^2 + 4n^3$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $A, B \in \mathbb{R}^{n \times n}$	–
begin	–
read(n,$\{a\}_{ij}$,$\{b\}_{ij}$);	–
for $i = 1$ **to** n **do**	i=1 – n
for $j = 1$ **to** n **do**	j=1 – n
$c_{ij} = 0$	–
for $k = 1$ **to** n **do**	k=1 – n
$c_{ij} = c_{ij} + a_{ik} * b_{kj}$	2 op.
endfor	2 op. téc.
endfor	2 op. téc.
endfor	2 op. téc.
write($\{c\}_{ij}$);	–
end	–

$$f(n) = \sum_{i}^{n} \left(2 + \sum_{j=1}^{n} \left(2 + \sum_{k=1}^{n} 2 + 2 \right) \right) = \sum_{i}^{n} \left(2 + \sum_{j=1}^{n} (2 + 4n) \right)$$

$$= \sum_{i}^{n} \left(2 + 2n + 4n^2 \right) = 2n + 2n^2 + 4n^3$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $A, B \in \mathbb{R}^{n\times n}$
begin
read(n,$\{a\}_{ij}$,$\{b\}_{ij}$);
for $i = 1$ **to** n **do**
    for $j = 1$ **to** n **do**
        $c_{ij} = 0$
        for $k = 1$ **to** n **do**
            $c_{ij} = c_{ij} + a_{ik} * b_{kj}$
        endfor
    endfor
endfor
write($\{c\}_{ij}$);
end

–
–
–
i=1 – n
    j=1 – n
        –
        k=1 – n
            2 op.
            2 op. téc.
        2 op. téc.
    2 op. téc.
–
–

$$f(n) = \sum_{i}^{n}\left(2 + \sum_{j=1}^{n}\left(2 + \sum_{k=1}^{n} 2 + 2\right)\right) = \sum_{i}^{n}\left(2 + \sum_{j=1}^{n}(2+4n)\right)$$

$$= \sum_{i}^{n}\left(2 + 2n + 4n^2\right) = 2n + 2n^2 + 4n^3$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$, $z \in \mathbb{R}$	–
begin	–
read(n,$\{a\}_i$,z);	–
$i = 1$;	–
while $(i <= n)$ && $(a_i \sim= z)$	i — n
i=i+1;	1 op.
	3 op. téc.
endwhile	1 op.
if $(i \sim= n+1)$ **them**	–
write("Error");	–
else	–
write(i);	–
end	–

$$f(n) = n(1+3) + 1 = 4n + 1$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$, $z \in \mathbb{R}$	–
begin	–
read(n,$\{a\}_i$,z);	–
$i = 1$;	–
while ($i <= n$) && ($a_i \sim= z$)	i — n
i=i+1;	1 op.
	3 op. téc.
endwhile	1 op.
if ($i \sim= n + 1$) **them**	–
write("Error");	–
else	–
write(i);	–
end	–

$$f(n) = n(1 + 3) + 1 = 4n + 1$$

Ejemplos

Ejemplo: Calcular la eficiencia del siguiente algoritmo

INPUT: $n \in \mathbb{N}$, $a \in \mathbb{R}^n$, $z \in \mathbb{R}$	–
begin	–
read(n,$\{a\}_i$,z);	–
$i = 1$;	–
while $(i <= n)$ && $(a_i \sim= z)$	i — n
i=i+1;	1 op.
endwhile	3 op. téc.
if $(i \sim= n+1)$ **them**	1 op.
write("Error");	–
else	–
write(i);	–
end	–

$$f(n) = n(1+3) + 1 = 4n + 1$$

Análisis Asintótico

- En Ciencias de la Computación resulta especialmente interesante estudiar el comportamiento de los algoritmos para tamaños grandes de la instancia.
- Este estudio se denomina Análisis Asintótico
- Se definen clases de equivalencia que permiten clasificar las funciones de eficiencia.

Notación O

$$O(f) = \{g : \mathbb{N}^* \to \mathbb{N}^* : \exists c > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0 \Rightarrow g(n) \leqslant cf(n)\}$$

Si $g(n)$ es la eficiencia de un algoritmo, entocnes se dice que su eficiencia asintótica es $O(f)$.

Clases asintóticas más utilizadas

$O(1)$	constantes	
$O(\log(\log n))$	sublogarítimicas	
$O(\log(n))$	logarítimicas	
$O(\sqrt{n})$	sublineales	
$O(n)$	lineales	
$O(n^2) \subset O(n^{a>2})$	polinómicas	
$O(2^n) \subset O(n!) \subset (n^n)$	exponenciales	

Eficiencia Asintótica

Ejemplo : En la siguiente tabla aparece el máximo tamaño de instancia que puede resolver por hora un algoritmo con una determinada eficiencia asintótica

iCore 7 - 2 núcleos ~ 2.000.000.000 op./seg. = 2.10^9
iCore 7 - 8 núcleos ~ 8.000.000.000 op./seg. = 8.10^9

Ef. Asint.	iCore 7 - 2 núcleos	iCore 7 - 8 núcleos
$O(\log(n))$	$\sim 10^{7.2*10^{12}}$	$\sim 10^{2.8*10^{13}}$
$O(n)$	$\sim 7.2 * 10^{12}$	$\sim 2.8 * 10^{13}$
$O(n^2)$	$\sim 2.8 * 10^6$	$\sim 5.4 * 10^6$
$O(n^3)$	$\sim 1.9 * 10^4$	$\sim 3.1 * 10^4$
$O(2^n)$	~ 43	~ 44
$O(n!)$	~ 16	~ 16
$O(n^n)$	~ 12	~ 12

Eficiencia Asintótica

Ejemplo : En la siguiente tabla aparece tiempo que requiere un algoritmo de una determinada eficiencia asintótica para resolver una instancia de tamaño 10, 100, 1.000 y 1.000.000

iCore 7 - 8 núcleos ~ 8.000.000.000 op./seg. = 8.10^9

Ef. Asint.	$n = 10$	$n = 10^2$	$n = 10^3$	$n = 10^6$
$O(\log(n))$	$\sim 1.2 * 10^{-10}$ seg.	$\sim 2.5 * 10^{-10}$ seg.	$\sim 3.7 * 10^{-10}$ seg.	$\sim 7.5 * 10^{-10}$ seg.
$O(n)$	$\sim 1.2 * 10^{-9}$ seg.	$\sim 1.2 * 10^{-8}$ seg.	$\sim 1.2 * 10^{-7}$ seg.	$\sim 1.2 * 10^{-4}$ seg.
$O(n^2)$	$\sim 1.2 * 10^{-8}$ seg.	$\sim 1.2 * 10^{-6}$ seg.	$\sim 1.2 * 10^{-4}$ seg.	~ 125 seg.
$O(n^3)$	$\sim 1.2 * 10^{-7}$ seg.	$\sim 1.2 * 10^{-4}$ seg.	~ 0.12 seg.	~ 3.96 años
$O(2^n)$	$\sim 1.2 * 10^{-7}$ seg.	$\sim 5. * 10^{10}$ siglos	∞	∞
$O(n!)$	$\sim 4.5 * 10^{-4}$ seg.	$\sim 3.7 * 10^{138}$ siglos	∞	∞
$O(n^n)$	~ 1.25 seg.	$\sim 3.9 * 10^{180}$ siglos	∞	∞

Complejidad

La teoría de complejidad clasifica los problemas atendiendo a su *dificultad*.

El estudio de la complejidad es matemáticamente muy exigente. En estas notas daremos sólo una versión informal.

Se deben distinguir dos tipos de problemas:

- **Problemas decisionales**: Son aquellos problemas cuya solución es un SI o NO. (ej. ¿Es p primo?, ¿Existe un camino hamiltoniano de longitud k?)
- **Problemas no decisionales**: Resto de problemas (optimización, evaluación).

La teoría de la complejidad se ocupa de los problemas decisionales. Si bien puede ser extendida a problemas de optimización y evaluación.

Complejidad de problemas decisionales

Dentro de los problemas decisionales se distinguen las siguientes clases:

- **Clase P** (*Polynomial time*): Un problema es de la clase P si existe un algoritmo de eficiencia polinomial que lo resuelve.

  Ejemplo: ¿Es la suma de n enteros par?, ¿Es p un número primo?

- **Clase NP**:(*Non-deterministic polynomial time*). Un problema es de la clase NP si dada una posible solución (certificado) existe un algorimto de eficiencia polinomial que permite probar si es solución del problema.

  Ejemplo: TSA. Un agente de viaje debe visitar n ciudades interconectadas. ¿Es posible visitar todas recorriendo una distancia inferior a d?

Desde la definición se tiene $P \subseteq NP$

Complejidad de problemas decisionales

Dentro de los problemas decisionales se distinguen las siguientes clases:

- **Clase P** (*Polynomial time*): Un problema es de la clase P si existe un algoritmo de eficiencia polinomial que lo resuelve.

  Ejemplo: ¿Es la suma de n enteros par?, ¿Es p un número primo?

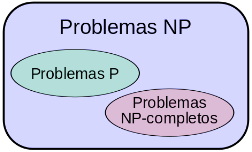
- **Clase NP**:(*Non-deterministic polynomial time*). Un problema es de la clase NP si dada una posible solución (certificado) existe un algorimto de eficiencia polinomial que permite probar si es solución del problema.

  Ejemplo: TSA. Un agente de viaje debe visitar n ciudades interconectadas. ¿Es posible visitar todas recorriendo una distancia inferior a d?

Desde la definición se tiene $P \subseteq NP$

Uno de los problemas del milenio

¿Es $P \subsetneq NP$ ó $P = NP$?



El Clay Mathematics Institute (CMI), Cambridge, Massachusetts, (USA) premia con $1,000,000$ $ a la primera persona que de respuesta a este problema

Problemas NP-Completos: Son problemas de la clase NP que se caracterizan por la siguiente propiedad:

> *Si existe un algoritmo polinomial que resuelve este problema, entonces todos los problemas de la clase NP son resolubles en tiempo polinómico.*

Es decir, resolver un problema NP con algoritmo polinómico tiene como premio un millón de dolares. (Ej. TSP, Satisfacibilidad)

Tema 2

Optimización convexa diferenciable

Introducción

Conjuntos y funciones convexos

Resultados generales de optimización

Optimización convexa (conceptos básicos)

Optimización diferenciable

Optimización convexa diferenciable

Resolución con Mathematica

Tema 2

Optimización convexa diferenciable

Introducción

Conjuntos y funciones convexos

Resultados generales de optimización

Optimización convexa (conceptos básicos)

Optimización diferenciable

Optimización convexa diferenciable

Resolución con Mathematica

Problema de optimización

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & x \in S \\ & [\, x \in X] \end{array}$$

donde,

$f : X \longrightarrow \mathbb{R}$

$S \subseteq X$

X : Conjunto de referencia ($X = \mathbb{R}^n$)

f : Función objetivo.

S : Conjunto factible. $x \in S$ *solución* factible. S suele venir dado como una serie de condiciones impuestas a los elementos de X:

$$S := \{x \in X \,:\, g_i(x) \leqslant 0, i = 1, \ldots, m\}$$

Las soluciones (óptimas) son los mínimos de la función f restringida al conjunto S: $f_{|S}$:

$$\mathcal{O} = \{z \in S \,:\, f(z) \leqslant f(x) \; \forall x \in S\}$$

Si $\mathcal{O} = \emptyset$ problema no factible (sin solución) , en otro caso

$\mathcal{O} \neq \emptyset$ problema factible (con solución).

Problema de optimización

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & x \in S \\ & [\, x \in X] \end{array}$$

donde,

$$f : X \longrightarrow \mathbb{R}$$
$$S \subseteq X$$

X : Conjunto de referencia ($X = \mathbb{R}^n$)

f : Función objetivo.

S : Conjunto factible. $x \in S$ *solución* factible. S suele venir dado como una serie de condiciones impuestas a los elementos de X:

$$S := \{x \in X \,:\, g_i(x) \leqslant 0, i = 1, \ldots, m\}$$

Las soluciones (óptimas) son los mínimos de la función f restringida al conjunto S: $f_{|S}$:

$$\mathcal{O} = \{z \in S \,:\, f(z) \leqslant f(x) \ \forall x \in S\}$$

Si $\mathcal{O} = \emptyset$ problema no factible (sin solución) , en otro caso

$\mathcal{O} \neq \emptyset$ problema factible (con solución).

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones. Se distinguen:

1. Optimización lineal
2. Optimización entera
3. Optimización mixta
4. Optimización combinatorial
5. Optimización no-lineal:
   - Optimización cuadrática
   - Optimización convexa , diferenciable, convexa diferenciable
   - Otros casos.
6. Optimización dinámica
7. Optimización estocástica

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones.Se distinguen:

1. Optimización lineal (Función objetivo y restricciones lineales)
2. Optimización entera
3. Optimización mixta
4. Optimización combinatorial
5. Optimización no-lineal:
   - Optimización cuadrática
   - Optimización convexa, diferenciable, convexa diferenciable
   - Otros casos.
6. Optimización dinámica
7. Optimización estocástica

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones.Se distinguen:

1. Optimización lineal
2. Optimización entera (Función objetivo y restricciones lineales. Variables enteras)
3. Optimización mixta
4. Optimización combinatorial
5. Optimización no-lineal:
   - Optimización cuadrática
   - Optimización convexa , diferenciable , convexa diferenciable
   - Otros casos.
6. Optimización dinámica
7. Optimización estocástica

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones.Se distinguen:

1. Optimización lineal
2. Optimización entera
3. Optimización mixta (Función objetivo y restricciones lineales. Variables enteras y continuas)
4. Optimización combinatorial
5. Optimización no-lineal:
    - Optimización cuadrática
    - Optimización convexa , diferenciable , convexa diferenciable
    - Otros casos.
6. Optimización dinámica
7. Optimización estocástica

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones.Se distinguen:

1. Optimización lineal
2. Optimización entera
3. Optimización mixta
4. Optimización combinatorial (El conjunto factible es finito)
5. Optimización no-lineal:
   - Optimización cuadrática
   - Optimización convexa , diferenciable , convexa diferenciable
   - Otros casos.
6. Optimización dinámica
7. Optimización estocástica

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones.Se distinguen:

1. Optimización lineal
2. Optimización entera
3. Optimización mixta
4. Optimización combinatorial
5. Optimización no-lineal:
   - Optimización cuadrática (Función objetivo y/o restricciones cuadráticas)
   - Optimización convexa, diferenciable, convexa diferenciable
   - Otros casos.
6. Optimización dinámica
7. Optimización estocástica

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones.Se distinguen:

1. Optimización lineal
2. Optimización entera
3. Optimización mixta
4. Optimización combinatorial
5. Optimización no-lineal:
   - Optimización cuadrática
   - Optimización convexa, diferenciable, convexa diferenciable (Función objetivo y/o restricciones convexas y/o diferenciables)
   - Otros casos.
6. Optimización dinámica
7. Optimización estocástica

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones.Se distinguen:

1. Optimización lineal
2. Optimización entera
3. Optimización mixta
4. Optimización combinatorial
5. Optimización no-lineal:
   - Optimización cuadrática
   - Optimización convexa, diferenciable, convexa diferenciable
   - Otros casos. cónica, semidefinida, geométrica
6. Optimización dinámica
7. Optimización estocástica

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones.Se distinguen:

1. Optimización lineal
2. Optimización entera
3. Optimización mixta
4. Optimización combinatorial
5. Optimización no-lineal:
   - Optimización cuadrática
   - Optimización convexa, diferenciable, convexa diferenciable
   - Otros casos.
6. Optimización dinámica (Incluye dependencia temporal)
7. Optimización estocástica

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones.Se distinguen:

1. Optimización lineal
2. Optimización entera
3. Optimización mixta
4. Optimización combinatorial
5. Optimización no-lineal:
   - Optimización cuadrática
   - Optimización convexa, diferenciable, convexa diferenciable
   - Otros casos.
6. Optimización dinámica
7. Optimización estocástica (Incluye variables estocásticas (no deterministas))

Clasificación de los problemas de optimización

Los problemas de optimización (programación) según características de la función objetivo y su restricciones.Se distinguen:

1. Optimización lineal
2. Optimización entera
3. Optimización mixta
4. Optimización combinatorial
5. Optimización no-lineal:
   - Optimización cuadrática
   - Optimización convexa , diferenciable , convexa diferenciable
   - Otros casos.
6. Optimización dinámica
7. Optimización estocástica

Objetivos de tema

En este tema nos centraremos en la optimización convexa, diferenciable, y convexa diferenciable. Los principales objetivos son:

- Revisar/Introducir los conceptos de convexidad (conjuntos, funciones)
- Revisar los conceptos básicos de optimización (máximos/mínimos) , Th Weirstrass.
- Introducir los resultados básicos de la optimización convexa: Th Local-Global, poliedros convexos.
- Revisar los resultados de optimización bajo supuestos de diferenciabilidad: abiertos, Lagrange, KKT.
- Introducir los resultados de optimización convexa-diferenciable.
- Resolver problemas del tipo anterior con *Mathematica*

Nota: Los conceptos desarrollados en este tema son de gran utilidad en los temas siguientes.: programación lineal, cuadrática, dinámica

Ejemplo

Un consumidor que dispone de una renta $M > 0$, puede consumir tres bienes con precios dados por p_1, p_2 y p_3 (p_i, $i > 0$) respectivamente y su función de utilidad es:

$$U(x_1, x_2, x_3) = (x_1 + \bar{x}_1)x_2^m x_3^n$$

donde x_i, $i = 1, 2, 3$ representa la cantidad consumida del bien i-ésima y $\bar{x}_1$, m y n son constantes positivas conocidas. Se pide

1. Calcular la combinación de bienes que ha de elegir el consumidor para maximizar su utilidad.
2. Estudiar el signo de la variación de la utilidad óptima como consecuencia de un incremento de valor de cada uno de los parámetros $\bar{\alpha} = (m, n, p_1, p_2, p_3, M)$ permaneciendo las demás constantes.

Notación

- Las variables escalares utilizan tipografía clásica: $x, y, z, x_i, \lambda, \ldots$
- La negrita en variables se utiliza para denotar vectores: $\mathbf{x}, \mathbf{y}, \mathbf{z}, \boldsymbol{\lambda}, \ldots$:

$$\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$$

- $S \subset \mathbb{R}^n, \quad f : S \to \mathbb{R}$ denota una función vectorial a valores escalares: $f(\mathbf{x}) = f(x_1, \ldots, x_n) \in \mathbb{R}$.
- $S \subset \mathbb{R}^m, \quad \mathbf{g} : S \to \mathbb{R}^m$ denota una función vectorial a valores escalares: $\mathbf{g}(\mathbf{x}) = \mathbf{g}(x_1, \ldots, x_n) \in \mathbb{R}^n$. La función $\mathbf{g}$ tiene m funciones componentes:

$$\mathbf{g} = (g_1, \ldots, g_m), \qquad g_i : S \to \mathbb{R}, \forall i = 1, \ldots, n$$

- El gradiente de f se denota: $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$
- El hessiano de f se denota $H_f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1,\ldots,n}$

Notación

- Las variables escalares utilizan tipografía clásica: $x, y, z, x_i, \lambda, \ldots$
- La negrita en variables se utiliza para denotar vectores: $\mathbf{x}, \mathbf{y}, \mathbf{z}, \boldsymbol{\lambda}, \ldots$:
$$\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$$
- $S \subset \mathbb{R}^n$, $f : S \to \mathbb{R}$ denota una función vectorial a valores escalares: $f(\mathbf{x}) = f(x_1, \ldots, x_n) \in \mathbb{R}$.
- $S \subset \mathbb{R}^m$, $\mathbf{g} : S \to \mathbb{R}^m$ denota una función vectorial a valores escalares: $\mathbf{g}(\mathbf{x}) = \mathbf{g}(x_1, \ldots, x_n) \in \mathbb{R}^n$. La función $\mathbf{g}$ tiene m funciones componentes:
$$\mathbf{g} = (g_1, \ldots, g_m), \qquad g_i : S \to \mathbb{R}, \forall i = 1, \ldots, n$$
- El gradiente de f se denota: $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$
- El hessiano de f se denota $H_f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1,\ldots,n}$

Notación

- Las variables escalares utilizan tipografía clásica: $x, y, z, x_i, \lambda, \ldots$
- La negrita en variables se utiliza para denotar vectores: $\mathbf{x}, \mathbf{y}, \mathbf{z}, \boldsymbol{\lambda}, \ldots$:

$$\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$$

- $S \subset \mathbb{R}^n$, $f : S \to \mathbb{R}$ denota una función vectorial a valores escalares: $f(\mathbf{x}) = f(x_1, \ldots, x_n) \in \mathbb{R}$.
- $S \subset \mathbb{R}^m$, $\mathbf{g} : S \to \mathbb{R}^m$ denota una función vectorial a valores escalares: $\mathbf{g}(\mathbf{x}) = \mathbf{g}(x_1, \ldots, x_n) \in \mathbb{R}^n$. La función $\mathbf{g}$ tiene m funciones componentes:

$$\mathbf{g} = (g_1, \ldots, g_m), \qquad g_i : S \to \mathbb{R}, \forall i = 1, \ldots, n$$

- El gradiente de f se denota: $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$
- El hessiano de f se denota $H_f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1,\ldots,n}$

Notación

- Las variables escalares utilizan tipografía clásica: $x, y, z, x_i, \lambda, \ldots$
- La negrita en variables se utiliza para denotar vectores: $\mathbf{x}, \mathbf{y}, \mathbf{z}, \boldsymbol{\lambda}, \ldots$:

$$\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$$

- $S \subset \mathbb{R}^n, \;\; f : S \to \mathbb{R}$ denota una función vectorial a valores escalares: $f(\mathbf{x}) = f(x_1, \ldots, x_n) \in \mathbb{R}$.
- $S \subset \mathbb{R}^m, \;\; \mathbf{g} : S \to \mathbb{R}^m$ denota una función vectorial a valores escalares: $\mathbf{g}(\mathbf{x}) = \mathbf{g}(x_1, \ldots, x_n) \in \mathbb{R}^n$. La función $\mathbf{g}$ tiene m funciones componentes:

$$\mathbf{g} = (g_1, \ldots, g_m), \qquad g_i : S \to \mathbb{R}, \forall i = 1, \ldots, n$$

- El gradiente de f se denota: $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$
- El hessiano de f se denota $H_f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1,\ldots,n}$

Notación

- Las variables escalares utilizan tipografía clásica: $x, y, z, x_i, \lambda, \ldots$
- La negrita en variables se utiliza para denotar vectores: $\mathbf{x}, \mathbf{y}, \mathbf{z}, \boldsymbol{\lambda}, \ldots$:
$$\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$$
- $S \subset \mathbb{R}^n$, $f : S \to \mathbb{R}$ denota una función vectorial a valores escalares: $f(\mathbf{x}) = f(x_1, \ldots, x_n) \in \mathbb{R}$.
- $S \subset \mathbb{R}^m$, $\mathbf{g} : S \to \mathbb{R}^m$ denota una función vectorial a valores escalares: $\mathbf{g}(\mathbf{x}) = \mathbf{g}(x_1, \ldots, x_n) \in \mathbb{R}^n$. La función $\mathbf{g}$ tiene m funciones componentes:
$$\mathbf{g} = (g_1, \ldots, g_m), \qquad g_i : S \to \mathbb{R}, \forall i = 1, \ldots, n$$
- El gradiente de f se denota: $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$
- El hessiano de f se denota $H_f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1,\ldots,n}$

Notación

- Las variables escalares utilizan tipografía clásica: $x, y, z, x_i, \lambda, \ldots$
- La negrita en variables se utiliza para denotar vectores: $\mathbf{x}, \mathbf{y}, \mathbf{z}, \boldsymbol{\lambda}, \ldots$:
$$\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$$
- $S \subset \mathbb{R}^n$, $f : S \to \mathbb{R}$ denota una función vectorial a valores escalares: $f(\mathbf{x}) = f(x_1, \ldots, x_n) \in \mathbb{R}$.
- $S \subset \mathbb{R}^m$, $\mathbf{g} : S \to \mathbb{R}^m$ denota una función vectorial a valores escalares: $\mathbf{g}(\mathbf{x}) = \mathbf{g}(x_1, \ldots, x_n) \in \mathbb{R}^n$. La función $\mathbf{g}$ tiene m funciones componentes:
$$\mathbf{g} = (g_1, \ldots, g_m), \qquad g_i : S \to \mathbb{R}, \forall i = 1, \ldots, n$$
- El gradiente de f se denota: $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$
- El hessiano de f se denota $H_f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1,\ldots,n}$

Tema 2

Optimización convexa diferenciable

Introducción

Conjuntos y funciones convexos

Resultados generales de optimización

Optimización convexa (conceptos básicos)

Optimización diferenciable

Optimización convexa diferenciable

Resolución con Mathematica

Conjuntos convexos

Definición Sean $x, y \in \mathbb{R}^n$. Llamaremos segmento cerrado que une $\mathbf{x}$ e $\mathbf{y}$ al siguiente subconjunto de $\mathbb{R}^n$:

$$[\mathbf{x}, \mathbf{y}] := \{\mathbf{z} \in \mathbb{R}^n ; \mathbf{z} = \lambda \mathbf{x} + (1 - \lambda)\mathbf{y} ; \lambda \in \mathbb{R}, 0 \leqslant \lambda \leqslant 1\}$$

Definición Sean $x, y \in \mathbb{R}^n$. Llamaremos segmento abierto que une $\mathbf{x}$ e $\mathbf{y}$ al siguiente subconjunto de $\mathbb{R}^n$:

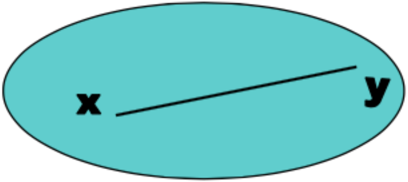
$$(\mathbf{x}, \mathbf{y}) = \{\mathbf{z} \in \mathbb{R}^n ; \mathbf{z} = \lambda \mathbf{x} + (1 - \lambda)\mathbf{y} ; \lambda \in \mathbb{R}, 0 < \lambda < 1\}$$

Conjuntos convexos

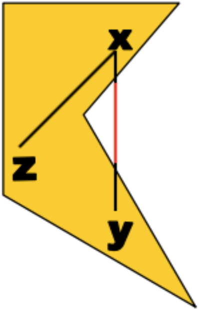
Definición Sea $A \subseteq \mathbb{R}^n$, $A \neq \emptyset$.

$$A \text{ es } \text{convexo} \quad \Leftrightarrow \quad \forall\, \mathbf{x}, \mathbf{y} \in A, \quad [\mathbf{x}, \mathbf{y}] \subseteq A$$

Nota Por convenio el conjunto vacío $(\emptyset)$ se considera convexo. Si $A = \{\mathbf{x}\}$ es convexo.



convexo



no convexo

Ejemplos de conjuntos convexos

Decidir cuál de los siguientes conjuntos son convexos:

1. $(-\infty, 3)$.
2. $(-\infty, -1) \cup [0, 1]$.
3. $A = \{x, y) \in \mathbb{R}^2; x^2 + y^2 \leqslant 1\}$.
4. $A = \{x, y) \in \mathbb{R}^2; 1 < x^2 + y^2 \leqslant 4\}$. $\Rightarrow$ No convexo
5. $A = \{(x, y, z) \in \mathbb{R}^3; 0 < x < 1; 0 < y < 1; 0 < z < 1; \}$.
6. $B = \{(x, y, z) \in \mathbb{R}^3; \quad z^2 = x^2 + y^2\}$.

Ejemplos de conjuntos convexos

Decidir cuál de los siguientes conjuntos son convexos:

1. $(-\infty, 3)$. $\Rightarrow$ Convexo
2. $(-\infty, -1) \cup [0, 1]$.
3. $A = \{x, y) \in \mathbb{R}^2; x^2 + y^2 \leqslant 1\}$.
4. $A = \{x, y) \in \mathbb{R}^2; 1 < x^2 + y^2 \leqslant 4\}$.
5. $A = \{(x, y, z) \in \mathbb{R}^3; 0 < x < 1; 0 < y < 1; 0 < z < 1;\}$.
6. $B = \{(x, y, z) \in \mathbb{R}^3; \quad z^2 = x^2 + y^2\}$.

Ejemplos de conjuntos convexos

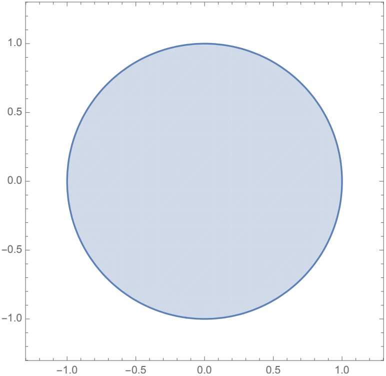
Decidir cuál de los siguientes conjuntos son convexos:

1. $(-\infty, 3)$. $\Rightarrow$ Convexo
2. $(-\infty, -1) \cup [0, 1]$. $\Rightarrow$ No convexo
3. $A = \{x, y) \in \mathbb{R}^2; x^2 + y^2 \leqslant 1\}$.
4. $A = \{x, y) \in \mathbb{R}^2; 1 < x^2 + y^2 \leqslant 4\}$.
5. $A = \{(x, y, z) \in \mathbb{R}^3; 0 < x < 1; 0 < y < 1; 0 < z < 1;\}$.
6. $B = \{(x, y, z) \in \mathbb{R}^3; \quad z^2 = x^2 + y^2\}$.

Ejemplos de conjuntos convexos

Decidir cuál de los siguientes conjuntos son convexos:

1. $(-\infty, 3)$. $\Rightarrow$ Convexo
2. $(-\infty, -1) \cup [0, 1]$. $\Rightarrow$ No convexo
3. $A = \{x, y) \in \mathbb{R}^2; x^2 + y^2 \leqslant 1\}$. $\Rightarrow$ Convexo

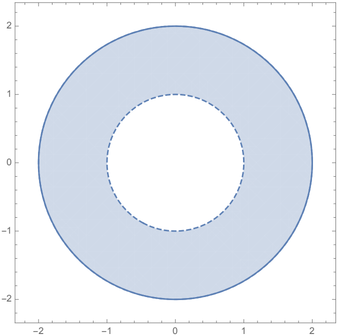


4. $A = \{x, y) \in \mathbb{R}^2; 1 < x^2 + y^2 \leqslant 4\}$.
5. $A = \{(x, y, z) \in \mathbb{R}^3, 0 < x < 1, 0 < y < 1, 0 < z < 1; \}$.

Ejemplos de conjuntos convexos

Decidir cuál de los siguientes conjuntos son convexos:

1. $(-\infty, 3)$. $\Rightarrow$ Convexo
2. $(-\infty, -1) \cup [0, 1]$. $\Rightarrow$ No convexo
3. $A = \{x, y) \in \mathbb{R}^2; x^2 + y^2 \leqslant 1\}$. $\Rightarrow$ Convexo
4. $A = \{x, y) \in \mathbb{R}^2; 1 < x^2 + y^2 \leqslant 4\}$. $\Rightarrow$ No convexo

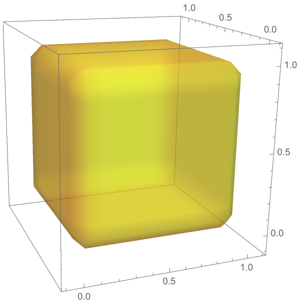


5. $A = \{(x, y, z) \in \mathbb{R}^3; 0 < x < 1; 0 < y < 1; 0 < z < 1;\}$.
6. $B = \{(x, y, z) \in \mathbb{R}^3; \; z^2 = x^2 + y^2\}$.

Ejemplos de conjuntos convexos

Decidir cuál de los siguientes conjuntos son convexos:

1. $(-\infty, 3)$. $\Rightarrow$ Convexo
2. $(-\infty, -1) \cup [0, 1]$. $\Rightarrow$ No convexo
3. $A = \{x, y) \in \mathbb{R}^2; x^2 + y^2 \leqslant 1\}$. $\Rightarrow$ Convexo
4. $A = \{x, y) \in \mathbb{R}^2; 1 < x^2 + y^2 \leqslant 4\}$. $\Rightarrow$ No convexo
5. $A = \{(x, y, z) \in \mathbb{R}^3; 0 < x < 1; 0 < y < 1; 0 < z < 1; \}$.
   $\Rightarrow$ Convexo

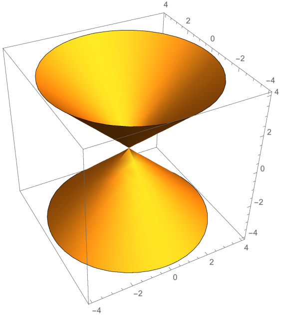


6. $B = \{(x, y, z) \in \mathbb{R}^3; \quad z^2 = x^2 + y^2\}$.

Ejemplos de conjuntos convexos

Decidir cuál de los siguientes conjuntos son convexos:

1. $(-\infty, 3)$. $\Rightarrow$ Convexo
2. $(-\infty, -1) \cup [0, 1]$. $\Rightarrow$ No convexo
3. $A = \{x, y) \in \mathbb{R}^2; x^2 + y^2 \leqslant 1\}$. $\Rightarrow$ Convexo
4. $A = \{x, y) \in \mathbb{R}^2; 1 < x^2 + y^2 \leqslant 4\}$. $\Rightarrow$ No convexo
5. $A = \{(x, y, z) \in \mathbb{R}^3; 0 < x < 1; 0 < y < 1; 0 < z < 1; \}$.
   $\Rightarrow$ Convexo
6. $B = \{(x, y, z) \in \mathbb{R}^3; \quad z^2 = x^2 + y^2\}$. $\Rightarrow$ No convexo



Combinaciones convexas

Definición Sean $\mathbf{x}_1, \ldots, \mathbf{x}_s \in \mathbb{R}^n$.
Llamaremos combinación lineal convexa de $\mathbf{x}_1, \ldots, \mathbf{x}_s$ a todo $\mathbf{z} \in \mathbb{R}$ tal que:

$$\mathbf{z} = \lambda_1 \mathbf{x}_1 + \ldots + \lambda_s \mathbf{x}_s, \quad \text{donde} \quad \begin{array}{l} \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+ y \\ \lambda_1 + \ldots + \lambda_s = 1 \end{array}$$

Proposición Sea $S \in \mathbb{R}^n$ un convexo no vacío y sean $\mathbf{x}_1, \ldots, \mathbf{x}_s \in S$. Cualquier combinación lineal convexa de $\mathbf{x}_1, \ldots, \mathbf{x}_s$ pertenece a S.

Proposición El conjunto de combinaciones lineales convexas de un conjunto de puntos $\mathbf{x}_1, \ldots, \mathbf{x}_s \in \mathbb{R}^n$ es un conjunto convexo.

Propiedades conjuntos convexos

Proposición Sean S y S' subconjuntos de $\mathbb{R}^n$ convexos. Entonces:

- $S \cap S'$ es convexo
- $S + S' = \{\mathbf{z} \in \mathbb{R}^n; \mathbf{z} = \mathbf{s} + \mathbf{s}', \mathbf{s} \in S, \mathbf{s}' \in S'\}$ es convexo
- $S - S' = \{\mathbf{z} \in \mathbb{R}^n; \mathbf{z} = \mathbf{s} - \mathbf{s}', \mathbf{s} \in S, \mathbf{s}' \in S'\}$ es convexo
- $\lambda S = \{\mathbf{z} \in \mathbb{R}^n; \mathbf{z} = \lambda \mathbf{s}, \mathbf{s} \in S, \lambda \in \mathbb{R}\}$ es convexo $\forall \lambda \in \mathbb{R}$

Proposición Sea

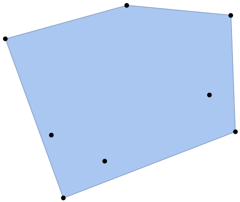
$$f : \mathbb{R}^n \to \mathbb{R}^m$$

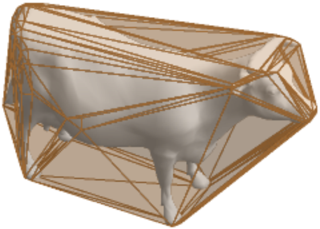
una aplicación lineal y sean $S \in \mathbb{R}^n, S' \in \mathbb{R}^m$ convexos. Entonces:

- $f(S)$ es un convexo de $\mathbb{R}^m$
- $f^{-1}(S')$ es un convexo de $\mathbb{R}^n$.

Envolvente conexa

Definición Sea S un subconjunto de $\mathbb{R}^n$. Se denomina envolvente convexa de S, $C_o(S)$, a la intersección de todos los convexos que contienen a S.





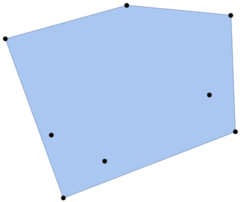
Notas

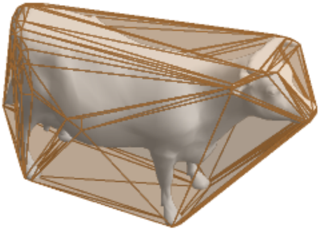
- Si S es convexo entonces $C_o(S) = S$
- Si S está formado por un número finito de puntos, $C_o(S)$ es precisamente el conjunto de combinaciones lineales convexas de esos puntos.

Proposición Sea $A \subset \mathbb{R}^n$. Sea

Envolvente conexa

Definición Sea S un subconjunto de $\mathbb{R}^n$. Se denomina envolvente convexa de S, $C_o(S)$, a la intersección de todos los convexos que contienen a S.





Notas

- Si S es convexo entonces $C_o(S) = S$
- Si S está formado por un número finito de puntos, $C_o(S)$ es precisamente el conjunto de combinaciones lineales convexas de esos puntos.

Proposición Sea $A \subset \mathbb{R}^n$. Sea

Envolvente conexa

Definición Sea S un subconjunto de $\mathbb{R}^n$. Se denomina envolvente convexa de S, $C_o(S)$, a la intersección de todos los convexos que contienen a S.

Notas

- Si S es convexo entonces $C_o(S) = S$
- Si S está formado por un número finito de puntos, $C_o(S)$ es precisamente el conjunto de combinaciones lineales convexas de esos puntos.

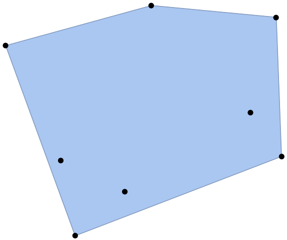
Proposición Sea $A \subset \mathbb{R}^n$. Sea

$$T = \left\{ \mathbf{x} = \lambda_1 \mathbf{x}_1 + \ldots + \lambda_s \mathbf{x}_s : \begin{array}{l} \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+ \\ \lambda_1 + \ldots + \lambda_s = 1 \\ \mathbf{x}_1, \ldots, \mathbf{x}_s \in A \end{array}, \forall s \in \mathbb{N} \right\}$$

Entonces $\boxed{C_o(A) = T}$.

Poliedros y puntos extremos

Definición Llamaremos poliedro convexo a la envolvente convexa de un conjunto finito de puntos.

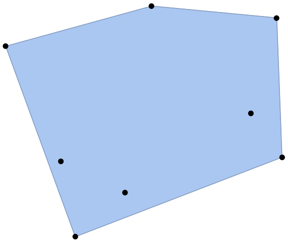


Definición Sea S un convexo de $\mathbb{R}^n$. Sea $\mathbf{x} \in S$. Diremos que $\mathbf{x}$ es un punto extremo o vértice de S si no puede expresarse como combinación lineal convexa de dos puntos distintos de S

Nota Los puntos extremos no está en el interior de ningún segmento que une puntos de S.

Poliedros y puntos extremos

Definición Llamaremos poliedro convexo a la envolvente convexa de un conjunto finito de puntos.



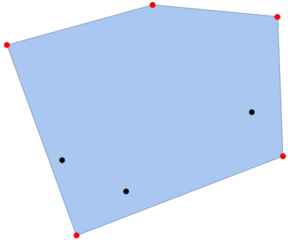
Definición Sea S un convexo de $\mathbb{R}^n$. Sea $\mathbf{x} \in S$. Diremos que $\mathbf{x}$ es un punto extremo o vértice de S si no puede expresarse como combinación lineal convexa de dos puntos distintos de S

Nota Los puntos extremos no está en el interior de ningún segmento que une puntos de S.

Poliedros y puntos extremos

Definición Llamaremos poliedro convexo a la envolvente convexa de un conjunto finito de puntos.

Definición Sea S un convexo de $\mathbb{R}^n$. Sea $\mathbf{x} \in S$. Diremos que $\mathbf{x}$ es un punto extremo o vértice de S si no puede expresarse como combinación lineal convexa de dos puntos distintos de S



En rojo figuran los extremos.

Nota Los puntos extremos no está en el interior de ningún segmento que une puntos de S.

Puntos extremos. Teorema de Krein-Milnan

Proposición Sea A convexo

$$\mathbf{x} \text{ un punto extremo} \quad \Rightarrow \quad \mathbf{x} \text{ punto frontera de } A$$

Nota El recíproco no es cierto.

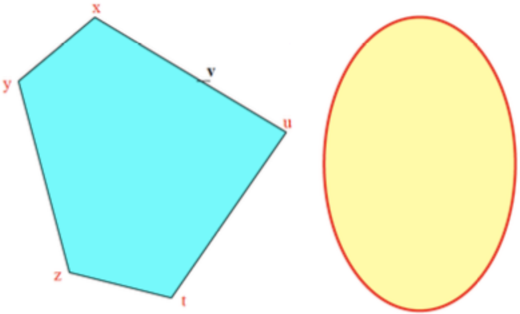
Teorema de Krein-Milman Sea $A \subset \mathbb{R}^n$ un conjunto convexo, cerrado y acotado.
Entonces, A contiene a sus puntos extremos y además es la envolvente convexa de los mismos.

Puntos extremos. Teorema de Krein-Milnan

Proposición Sea A convexo

$\mathbf{x}$ un punto extremo $\Rightarrow$ $\mathbf{x}$ punto frontera de A

Nota El recíproco no es cierto.



Puntos extremos. Teorema de Krein-Milnan

Proposición Sea A convexo

$\mathbf{x}$ un punto extremo $\Rightarrow$ $\mathbf{x}$ punto frontera de A

Nota El recíproco no es cierto.

Teorema de Krein-Milman Sea $A \subset \mathbb{R}^n$ un conjunto convexo, cerrado y acotado.
Entonces, A contiene a sus puntos extremos y además es la envolvente convexa de los mismos.

Teorema de Carátheodory

Teorema de Carátheodory Sea S un subconjunto de $\mathbb{R}^n$ y $C_o(S)$ su envolvente convexa.
Entonces todo punto de $C_o(S)$ puede expresarse como combinación lineal convexa de $n + 1$ puntos de S.

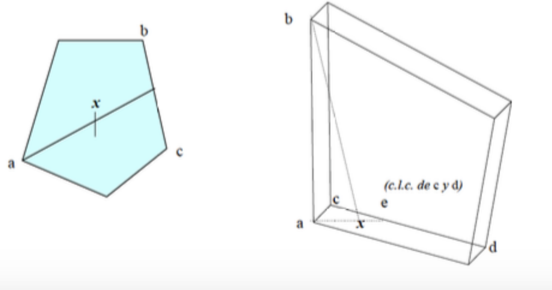
Nota A partir de este Teorema se deduce que todo punto de un poliedro convexo de $\mathbb{R}^n$ puede expresarse como combinación lineal convexa de $n + 1$ de los vértices del poliedro.

Teorema de Carátheodory

Teorema de Carátheodory Sea S un subconjunto de $\mathbb{R}^n$ y $C_o(S)$ su envolvente convexa.
Entonces todo punto de $C_o(S)$ puede expresarse como combinación lineal convexa de $n+1$ puntos de S.

Nota A partir de este Teorema se deduce que todo punto de un poliedro convexo de $\mathbb{R}^n$ puede expresarse como combinación lineal convexa de $n+1$ de los vértices del poliedro.

Ejemplo



Funciones convexas/concavas

Definición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$.

- f es convexa en S si

$$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leqslant \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

- f es estrictamente convexa en S si

$$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) < \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

- f es cóncava en S si

$$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \geqslant \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

- f es estrictamente cóncava en S si

$$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) > \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

Funciones convexas/concavas

Definición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$.

- f es convexa en S si

  $$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leqslant \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

- f es estrictamente convexa en S si

  $$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) < \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

- f es cóncava en S si

  $$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \geqslant \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

- f es estrictamente cóncava en S si

  $$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) > \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

Funciones convexas/concavas

Definición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$.

- f es convexa en S si

$$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leqslant \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

- f es estrictamente convexa en S si

$$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) < \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

- f es cóncava en S si

$$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \geqslant \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

- f es estrictamente cóncava en S si

$$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) > \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

Funciones convexas/concavas

Definición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$.

- f es convexa en S si

  $$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] \ : f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leqslant \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$$

- f es estrictamente convexa en S si

  $$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] : f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) < \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$$

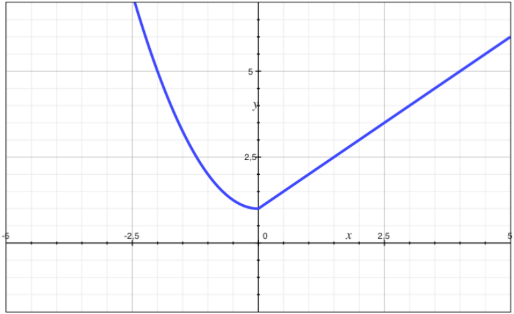
- f es cóncava en S si

  $$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] \ : f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \geqslant \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$$

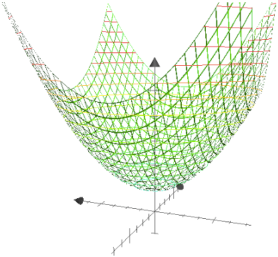
- f es estrictamente cóncava en S si

  $$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] \ : f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) > \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$$

Funciones convexas. Ejemplos



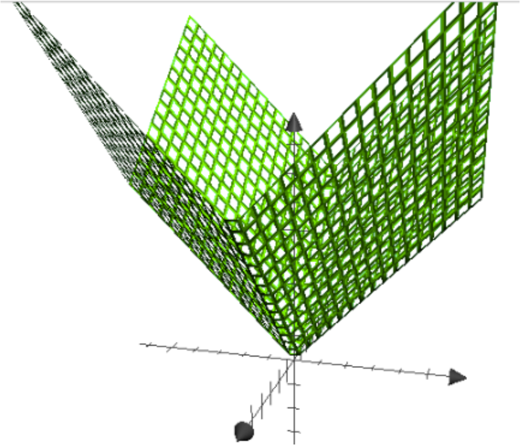
Función convexa



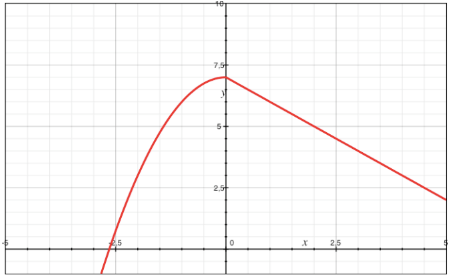
Función estrict. convexa

Funciones convexas. Ejemplos

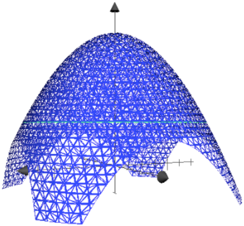
La función $f(x, y) = |x| + |y|$ es convexa



Funciones concavas. Ejemplos



Función concava



Función estrict. cóncava

Funciones convexas/cóncavas. Propiedades I

Proposición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$ no vacio.

- Toda función lineal es cóncava y convexa
- f es convexa $\Leftrightarrow -f$ es cóncava.
- f es estrict. convexa $\Leftrightarrow -f$ es estrict. cóncava.
- f es convexa $\Leftrightarrow \forall c \in \mathbb{R}^+$ la función cf es convexa.
- Si f es convexa entonces es continua.

Proposición Sean $f_1, \ldots, f_s$ funciones reales convexas definidas en un convexo S de $\mathbb{R}^n$. Entonces la función

$$\lambda_1 f_1 + \ldots + \lambda_s f_s$$

es una función convexa en C, $\forall \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+$

¿Cómo se interpreta este resultado en el caso de funciones cóncavas?

Funciones convexas/cóncavas. Propiedades I

Proposición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$ no vacio.

- Toda función lineal es cóncava y convexa
- f es convexa $\Leftrightarrow -f$ es cóncava.
- f es estrict. convexa $\Leftrightarrow -f$ es estrict. cóncava.
- f es convexa $\Leftrightarrow \forall c \in \mathbb{R}^+$ la función cf es convexa.
- Si f es convexa entonces es continua.

Proposición Sean $f_1, \ldots, f_s$ funciones reales convexas definidas en un convexo S de $\mathbb{R}^n$. Entonces la función

$$\lambda_1 f_1 + \ldots + \lambda_s f_s$$

es una función convexa en C, $\forall \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+$

¿Cómo se interpreta este resultado en el caso de funciones cóncavas?

Funciones convexas/cóncavas. Propiedades I

Proposición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$ no vacio.

- Toda función lineal es cóncava y convexa
- f es convexa $\Leftrightarrow -f$ es cóncava.
- f es estrict. convexa $\Leftrightarrow -f$ es estrict. cóncava.
- f es convexa $\Leftrightarrow \forall c \in \mathbb{R}^+$ la función cf es convexa.
- Si f es convexa entonces es continua.

Proposición Sean $f_1, \ldots, f_s$ funciones reales convexas definidas en un convexo S de $\mathbb{R}^n$. Entonces la función

$$\lambda_1 f_1 + \ldots + \lambda_s f_s$$

es una función convexa en C, $\forall \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+$

¿Cómo se interpreta este resultado en el caso de funciones cóncavas?

Funciones convexas/cóncavas. Propiedades I

Proposición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$ no vacio.

- Toda función lineal es cóncava y convexa
- f es convexa $\Leftrightarrow -f$ es cóncava.
- f es estrict. convexa $\Leftrightarrow -f$ es estrict. cóncava.
- f es convexa $\Leftrightarrow \forall c \in \mathbb{R}^+$ la función cf es convexa.
- Si f es convexa entonces es continua.

Proposición Sean $f_1, \ldots, f_s$ funciones reales convexas definidas en un convexo S de $\mathbb{R}^n$. Entonces la función

$$\lambda_1 f_1 + \ldots + \lambda_s f_s$$

es una función convexa en C, $\forall \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+$

¿Cómo se interpreta este resultado en el caso de funciones cóncavas?

Funciones convexas/cóncavas. Propiedades I

Proposición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$ no vacio.

- Toda función lineal es cóncava y convexa
- f es convexa $\Leftrightarrow -f$ es cóncava.
- f es estrict. convexa $\Leftrightarrow -f$ es estrict. cóncava.
- f es convexa $\Leftrightarrow \forall c \in \mathbb{R}^+$ la función cf es convexa.
- Si f es convexa entonces es continua.

Proposición Sean $f_1, \ldots, f_s$ funciones reales convexas definidas en un convexo S de $\mathbb{R}^n$. Entonces la función

$$\lambda_1 f_1 + \ldots + \lambda_s f_s$$

es una función convexa en C, $\forall \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+$

¿Cómo se interpreta este resultado en el caso de funciones cóncavas?

Funciones convexas/cóncavas. Propiedades I

Proposición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$ no vacio.

- Toda función lineal es cóncava y convexa
- f es convexa $\Leftrightarrow -f$ es cóncava.
- f es estrict. convexa $\Leftrightarrow -f$ es estrict. cóncava.
- f es convexa $\Leftrightarrow \forall c \in \mathbb{R}^+$ la función cf es convexa.
- Si f es convexa entonces es continua.

Nota: Todo resultado para funciones convexas tiene una lectura paralela en el caso de funciones cóncavas

Proposición Sean $f_1, \ldots, f_s$ funciones reales convexas definidas en un convexo S de $\mathbb{R}^n$. Entonces la función

$$\lambda_1 f_1 + \ldots + \lambda_s f_s$$

es una función convexa en C, $\forall \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+$

¿Cómo se interpreta este resultado en el caso de funciones cóncavas?

Funciones convexas/cóncavas. Propiedades I

Proposición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$ no vacio.

- Toda función lineal es cóncava y convexa
- f es convexa $\Leftrightarrow -f$ es cóncava.
- f es estrict. convexa $\Leftrightarrow -f$ es estrict. cóncava.
- f es convexa $\Leftrightarrow \forall c \in \mathbb{R}^+$ la función cf es convexa.
- Si f es convexa entonces es continua.

Proposición Sean $f_1, \ldots, f_s$ funciones reales convexas definidas en un convexo S de $\mathbb{R}^n$. Entonces la función

$$\lambda_1 f_1 + \ldots + \lambda_s f_s$$

es una función convexa en C, $\forall\, \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+$

¿Cómo se interpreta este resultado en el caso de funciones cóncavas?

Funciones convexas/cóncavas. Propiedades II

Proposición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$.

- f es convexa en S si y solo si el conjunto

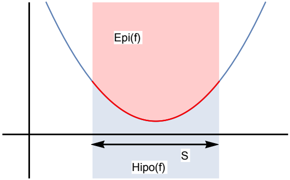
$$\text{epi}\,(f) = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1}, x \in S; y \geqslant f(\mathbf{x})\}$$

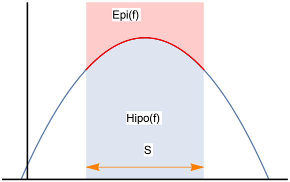
  es convexo

- f es cóncava en S si y sólo si el conjunto

$$\text{hipo}\,(f) = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1}, x \in S; y \leqslant f(\mathbf{x})\}$$

  es convexo.



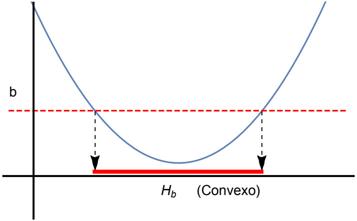


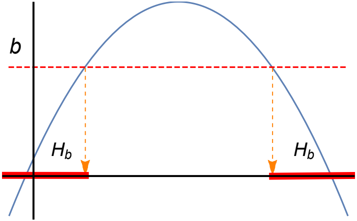
Funciones convexas/cóncavas. Propiedades II

Proposición Sea una función $f : S \to \mathbb{R}$ donde S es un convexo de $\mathbb{R}^n$. Si f es convexa entonces el conjunto

$$H_b = \{\mathbf{x} \in S; f(\mathbf{x}) \leqslant b\}$$

es convexo, $\forall b \in \mathbb{R}$.





Funciones convexas/cóncavas. Propiedades III

Proposición Sea una función $f : S \to \mathbb{R}$ continua con S es un convexo de $\mathbb{R}^n$.

$$f \text{ es convexa} \quad \Leftrightarrow \quad f\left(\frac{\mathbf{x}+\mathbf{y}}{2}\right) \leqslant \frac{f(\mathbf{x})+f(\mathbf{y})}{2}, \quad \forall\, \mathbf{x}, \mathbf{y} \in S$$

Ejemplo Probar que $f(x) = x^2 + 5x$ es convexa y que $f(x) = \sqrt{x}$ es cóncava.

Proposición Sea $S \subset \mathbb{R}^n$ convexo y $f : S \to \mathbb{R}$. Entonces f es convexa en S si y solo si

- $\forall \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+, \lambda_1 + \ldots + \lambda_s = 1$
- $\forall \mathbf{x}_1, \ldots, \mathbf{x}_s \in S$

se verifica

$$f(\lambda_1 \mathbf{x}_1 + \ldots + \lambda_s \mathbf{x}_s) \leqslant \lambda_1 f(\mathbf{x}_1) + \ldots + \lambda_s f(\mathbf{x}_s)$$

Funciones convexas/cóncavas. Propiedades III

Proposición Sea una función $f : S \to \mathbb{R}$ continua con S es un convexo de $\mathbb{R}^n$.

$$f \text{ es convexa} \quad \Leftrightarrow \quad f\left(\frac{\mathbf{x}+\mathbf{y}}{2}\right) \leqslant \frac{f(\mathbf{x})+f(\mathbf{y})}{2}, \quad \forall\, \mathbf{x},\mathbf{y} \in S$$

Ejemplo Probar que $f(x) = x^2 + 5x$ es convexa y que $f(x) = \sqrt{x}$ es cóncava.

Proposición Sea $S \subset \mathbb{R}^n$ convexo y $f : S \to \mathbb{R}$. Entonces f es convexa en S si y solo si

- $\forall\, \lambda_1, \ldots, \lambda_s \in \mathbb{R}^+, \lambda_1 + \ldots + \lambda_s = 1$
- $\forall\, \mathbf{x}_1, \ldots, \mathbf{x}_s \in S$

se verifica

$$f(\lambda_1 \mathbf{x}_1 + \ldots + \lambda_s \mathbf{x}_s) \leqslant \lambda_1 f(\mathbf{x}_1) + \ldots + \lambda_s f(\mathbf{x}_s)$$

Funciones convexas/concavas y diferenciablidad

Proposición Sea una función $f : S \to \mathbb{R}$ una función de clase $\mathcal{C}^2$, donde S es convexo y abierto en $\mathbb{R}^n$. Entonces:

- f es convexa $\Leftrightarrow$ $H_f(\mathbf{x})$ es semidefinida positiva en S
- f es cóncava $\Leftrightarrow$ $H_f(\mathbf{x})$ es semidefinida negativa en S
- f es estr. convexa $\Leftrightarrow$ $H_f(\mathbf{x})$ es definida positiva en S
- f es estr. cóncava $\Leftrightarrow$ $H_f(\mathbf{x})$ es definida negativa en S

Nota Recordar que:

$$H_f(\mathbf{x}) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{1 \leqslant i,j \leqslant n}$$

Tema 2

Optimización convexa diferenciable

Introducción

Conjuntos y funciones convexos

Resultados generales de optimización
- Conceptos previos
- Teorema Weierstrass

Optimización convexa (conceptos básicos)

Optimización diferenciable

Optimización convexa diferenciable

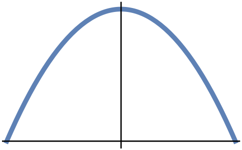
Resolución con Mathematica

Conceptos previos

Definición: (Máximo y mínimos) (Globales)
Sea S un conjunto, y sea $f : S \to \mathbb{R}$ una función.

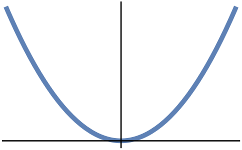
Se dice que $\mathbf{x}^* \in S$ es un mínimo de f si:

$$f(\mathbf{x}^*) \leqslant f(\mathbf{x}),\ \forall \mathbf{x} \in S$$



Se dice que $\mathbf{x}^* \in S$ es un máximo de f si:

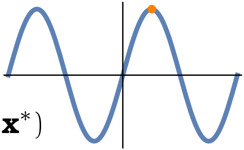
$$f(\mathbf{x}^*) \geqslant f(\mathbf{x}),\ \forall \mathbf{x} \in S$$



Conceptos previos

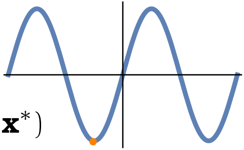
Definición: (Máximo y mínimos Locales)
Sea S un conjunto, y sea $f : S \to \mathbb{R}$ una función.

Se dice que $\mathbf{x}^* \in S$ es un
mínimo local de f si:



$\exists\epsilon > 0$ tal que $f(\mathbf{x}^*) \leqslant f(\mathbf{x}),\ \forall\mathbf{x} \in S \cap B_\epsilon(\mathbf{x}^*)$

Se dice que $\mathbf{x}^* \in S$ es un
máximo local de f si:



$\exists\epsilon > 0$ tal que $f(\mathbf{x}^*) \geqslant f(\mathbf{x}),\ \forall\mathbf{x} \in S \cap B_\epsilon(\mathbf{x}^*)$

Teorema de Weierstrass

Teorema Sean

- S un conjunto compacto y no vacío de $\mathbb{R}$
- $f : S \to \mathbb{R}$ continua.

Entonces existen mínimo y máximo de f.

Nota:

- Dado que $S \subset \mathbb{R}^n$, compacto equivale a cerrado y acotado
- El resultado anterior solo garantiza la existencia pero no proporciona ninguna información acerca de su cáclulo

Tema 2

Optimización convexa diferenciable

Introducción

Conjuntos y funciones convexos

Resultados generales de optimización

Optimización convexa (conceptos básicos)
- Teorema Local-Global
- Optimización en poliedros convexos

Optimización diferenciable

Optimización convexa diferenciable

Resolución con Mathematica

Optimización de funciones convexas

Teorema (Local Global) Sea $S \subset \mathbb{R}^n$ convexo, y sea $f : S \to \mathbb{R}$ convexa. Dado $\mathbf{x}^* \in S$, se tiene:

$$\boxed{\mathbf{x}^* \text{ es mínimo de } f \;\Leftrightarrow\; \mathbf{x}^* \text{ es mínimo local de } f}$$

Teorema Sea $S \subset \mathbb{R}^n$ convexo y sea $f : S \to \mathbb{R}$ estríctamente convexa. Entonces, si existe mínimo de f es $\boxed{\text{único}}$.

Optimización de funciones convexas

Teorema (Local Global) Sea $S \subset \mathbb{R}^n$ convexo, y sea $f : S \to \mathbb{R}$ convexa. Dado $\mathbf{x}^* \in S$, se tiene:

$$\mathbf{x}^* \text{ es mínimo de } f \quad \Leftrightarrow \quad \mathbf{x}^* \text{ es mínimo local de } f$$

Demo.

1. Como $\mathbf{x}^*$ es un mínimo local existe un entorno $V_{\mathbf{x}^*}$ tal que si $\mathbf{y} \in V_{\mathbf{x}^*} \cap S$ entonces $f(\mathbf{y}) \geqslant f(\bar{\mathbf{x}})$.
2. Si $\mathbf{x}^*$ no fuera mínimo global existiría $\mathbf{y} \in S$ tal que $f(\mathbf{y}) < f(\mathbf{x}^*)$. Dado que S es convexo existirá un $\lambda \in [0, 1]$ tal que $\lambda\mathbf{x}^* + (1 - \lambda)\mathbf{y} \in V_{\mathbf{x}^*}$.
3. Utilizando la convexidad deducimos:

$$f(\mathbf{x}^*) \leqslant f(\lambda\mathbf{x}^* + (1-\lambda)\mathbf{y}) \leqslant \lambda f(\mathbf{x}^*) + (1-\lambda)f(\mathbf{y}) < f(\mathbf{x}^*)$$

Lo cual es una contradicción.

Teorema Sea $S \subset \mathbb{R}^n$ convexo y sea $f : S \to \mathbb{R}$ estríctamente convexa. Entonces, si existe mínimo de f es único.

Optimización de funciones convexas

Teorema (Local Global) Sea $S \subset \mathbb{R}^n$ convexo, y sea $f : S \to \mathbb{R}$ convexa. Dado $\mathbf{x}^* \in S$, se tiene:

$$\boxed{\mathbf{x}^* \text{ es mínimo de } f \quad \Leftrightarrow \quad \mathbf{x}^* \text{ es mínimo local de } f}$$

Teorema Sea $S \subset \mathbb{R}^n$ convexo y sea $f : S \to \mathbb{R}$ estríctamente convexa. Entonces, si existe mínimo de f es $\boxed{\text{único}}$.

Optimización de funciones convexas

Teorema (Local Global) Sea $S \subset \mathbb{R}^n$ convexo, y sea $f : S \to \mathbb{R}$ convexa. Dado $\mathbf{x}^* \in S$, se tiene:

$$\boxed{\mathbf{x}^* \text{ es mínimo de } f \quad \Leftrightarrow \quad \mathbf{x}^* \text{ es mínimo local de } f}$$

Teorema Sea $S \subset \mathbb{R}^n$ convexo y sea $f : S \to \mathbb{R}$ estríctamente convexa. Entonces, si existe mínimo de f es $\boxed{\text{único}}$.

Demo. Sean $\mathbf{x}, \mathbf{y} \in S$, con $\mathbf{x} \neq \mathbf{y}$ mínimos de f (i.e. $f(\mathbf{x}) = f(\mathbf{y}) = m$).
Usando la convexidad estricta:

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) < \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}) = m$$

Lo cual es una contradicción

Optimización de funciones convexas

Teorema (Local Global) Sea $S \subset \mathbb{R}^n$ convexo, y sea $f : S \to \mathbb{R}$ convexa. Dado $\mathbf{x}^* \in S$, se tiene:

$$\boxed{\mathbf{x}^* \text{ es mínimo de } f \quad \Leftrightarrow \quad \mathbf{x}^* \text{ es mínimo local de } f}$$

Teorema Sea $S \subset \mathbb{R}^n$ convexo y sea $f : S \to \mathbb{R}$ estríctamente convexa. Entonces, si existe mínimo de f es $\boxed{\text{único}}$.

Ejercicio Enunciar y demostrar los resultados análogos para el caso de funciones cóncavas.

Optimización en poliedros convexos

Teorema Sea S un conjunto cerrado, acotado y convexo de $\mathbb{R}^n$ con un número finito de puntos extremos y $f : S \to \mathbb{R}$ una función convexa.

Entonces el máximo de f en S se alcanza en uno de los puntos extremos.

Ejercicio Enunciar y demostar el resultado análogo para funciones concavas.

Nota Este resultado permite demostar que la solución de un problema de programación lineal cuyo conjunto factible es acotado, está en uno de los vértices (¿Por qué?)

Optimización en poliedros convexos

Teorema Sea S un conjunto cerrado, acotado y convexo de $\mathbb{R}^n$ con un número finito de puntos extremos y $f : S \to \mathbb{R}$ una función convexa.

Entonces el máximo de f en S se alcanza en uno de los puntos extremos.

Demo. Sean $\{\mathbf{x}_i\}_{i=1,\ldots,r}$ los puntos extremos. Por el Th de Krein-Milman, para todo $\mathbf{x} \in S$ exiten $\lambda_i \in [0, 1]$ $(i = 1 \ldots r)$ tal que:

$$\mathbf{x} = \lambda_1 \mathbf{x}_1 + \cdots + \lambda_r \mathbf{x}_r, \qquad \lambda_1 + \cdots + \lambda_r = 1$$

Sea $\mathbf{x}_k$ un vértice donde se alcance $\max\limits_{i=1\ldots r} f(\mathbf{x}_i)$.
Entonces, utilizando la convexidad:

$$f(\mathbf{x}) \leqslant \lambda_1 f(\mathbf{x}_1) + \cdots + \lambda_r f(\mathbf{x}_r) \leqslant (\lambda_1 + \cdots + \lambda_r) f(\mathbf{x}_k) = f(\mathbf{x}_k)$$

concluimos que $\mathbf{x}_k$ es máximo.

Ejercicio Enunciar y demostar el resultado análogo para funciones concavas.

Nota Este resultado permite demostar que la solución de un problema de programación lineal cuyo conjunto factible

Optimización en poliedros convexos

Teorema Sea S un conjunto cerrado, acotado y convexo de $\mathbb{R}^n$ con un número finito de puntos extremos y $f : S \to \mathbb{R}$ una función convexa.

Entonces el máximo de f en S se alcanza en uno de los puntos extremos.

Ejercicio Enunciar y demostar el resultado análogo para funciones concavas.

Nota Este resultado permite demostar que la solución de un problema de programación lineal cuyo conjunto factible es acotado, está en uno de los vértices (¿Por qué?)

Tema 2

Optimización convexa diferenciable

Introducción

Conjuntos y funciones convexos

Resultados generales de optimización

Optimización convexa (conceptos básicos)

Optimización diferenciable
- Abiertos
- Restricciones de igualdad (Lagrange)
- Restricciones de desigualdad. KKT

Optimización convexa diferenciable

Resultados generales de optimización

Presentaremos en esta sección los resultados clásicos de optimización, muchos de ellos ya fueron abordados en el primer curso de la titulación.

Estos resultados se refieren al problema:

$$\boxed{\min_{\mathbf{x}\in S} / \max_{\mathbf{x}\in S} \quad f(\mathbf{x})}$$

siendo $S \subset \mathbb{R}^n$ conjunto no vacio y $f : S \to \mathbb{R}$ una función (dos veces) diferencialbe.

Se estudiará el problema anterior en los casos:

1. S abierto
2. S determinado por una restricción de igualdad:

$$S = \{\mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{g}(\mathbf{x}) = \mathbf{b}\}$$

3. S determinado por restricción de desigualdad:

$$S = \{\mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{g}(\mathbf{x}) \leqslant \mathbf{b}\}$$

Resultados generales de optimización

Presentaremos en esta sección los resultados clásicos de optimización, muchos de ellos ya fueron abordados en el primer curso de la titulación.

Estos resultados se refieren al problema:

$$\boxed{\min_{\mathbf{x}\in S} / \max_{\mathbf{x}\in S} \quad f(\mathbf{x})}$$

siendo $S \subset \mathbb{R}^n$ conjunto no vacio y $f : S \to \mathbb{R}$ una función (dos veces) diferencialbe.

Se estudiará el problema anterior en los casos:

1. S abierto
2. S determinado por una restricción de igualdad:

$$S = \{\mathbf{x} \in \mathbb{R}^n \, : \, \mathbf{g}(\mathbf{x}) = \mathbf{b}\}$$

3. S determinado por restricción de desigualdad:

$$S = \{\mathbf{x} \in \mathbb{R}^n \, : \, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{b}\}$$

Resultados generales de optimización

Presentaremos en esta sección los resultados clásicos de optimización, muchos de ellos ya fueron abordados en el primer curso de la titulación.

Estos resultados se refieren al problema:

$$\boxed{\min_{\mathbf{x}\in S} / \max_{\mathbf{x}\in S} \quad f(\mathbf{x})}$$

siendo $S \subset \mathbb{R}^n$ conjunto no vacio y $f : S \to \mathbb{R}$ una función (dos veces) diferencialbe.

Se estudiará el problema anterior en los casos:

1. S abierto
2. S determinado por una restricción de igualdad:

$$S = \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{g}(\mathbf{x}) = \mathbf{b}\}$$

3. S determinado por restricción de desigualdad:

$$S = \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{b}\}$$

Optimización en abiertos. Condición de primer orden

Hipótesis En esta sección consideraremos

- $S \subseteq \mathbb{R}^n$ abierto
- $f : S \to \mathbb{R}$ función diferenciable (clase $\mathcal{C}^1$)

Definición Se denomina puntos críticos o estacionarios de f a aquellos puntos $\mathbf{x}^* \in \mathbb{R}^n$ tales que $\nabla f(\mathbf{x}^*) = \mathbf{0}$

Teorema (Condición de primer orden)
Si $\mathbf{x}^* \in S$ es un óptimo (máximo o mínimo) local de f entonces $\mathbf{x}^*$ es un punto crítico de f.

$$\boxed{\mathbf{x}^* \text{ óptimo local} \quad \Rightarrow \quad \nabla f(\mathbf{x}^*) = \mathbf{0}}$$

Optimización en abiertos. Condición de primer orden

Hipótesis En esta sección consideraremos

- $S \subseteq \mathbb{R}^n$ abierto
- $f : S \to \mathbb{R}$ función diferenciable (clase $\mathcal{C}^1$)

Definición Se denomina puntos críticos o estacionarios de f a aquellos puntos $\mathbf{x}^* \in \mathbb{R}^n$ tales que $\boxed{\nabla f(\mathbf{x}^*) = \mathbf{0}}$

Teorema (Condición de primer orden)
Si $\mathbf{x}^* \in S$ es un óptimo (máximo o mínimo) local de f entonces $\mathbf{x}^*$ es un punto crítico de f.

$$\boxed{\mathbf{x}^* \text{ óptimo local} \quad \Rightarrow \quad \nabla f(\mathbf{x}^*) = \mathbf{0}}$$

Optimización en abiertos. Condición de primer orden

Hipótesis En esta sección consideraremos

- $S \subseteq \mathbb{R}^n$ abierto
- $f : S \to \mathbb{R}$ función diferenciable (clase $\mathcal{C}^1$)

Definición Se denomina puntos críticos o estacionarios de f a aquellos puntos $\mathbf{x}^* \in \mathbb{R}^n$ tales que $\nabla f(\mathbf{x}^*) = \mathbf{0}$

Teorema (Condición de primer orden)
Si $\mathbf{x}^* \in S$ es un óptimo (máximo o mínimo) local de f entonces $\mathbf{x}^*$ es un punto crítico de f.

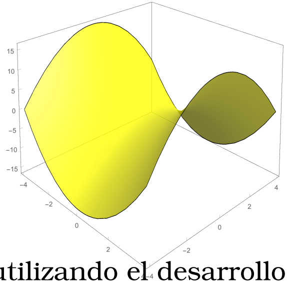
$$\boxed{\mathbf{x}^* \text{ óptimo local} \quad \Rightarrow \quad \nabla f(\mathbf{x}^*) = \mathbf{0}}$$

Condición de primer orden. Observaciones

1. La condición anterior es un condición necesaria de óptimo. Permite seleccionar un conjunto de puntos (críticos) entre los cuales estarán los óptimos.

2. El recíproco de la proposición anterior no es cierto (no todo punto crítico es un óptimo local). A los puntos críticos que no son óptimos se les denomina puntos silla

$f(x, y) = x^2 - y^2$
tiene un punto silla en $(0, 0)$



3. El teorema se puede deducir utilizando el desarrollo de Taylor.

Optimización en abiertos. Condición de segundo orden

Hipótesis

- $S \subseteq \mathbb{R}^n$ abierto
- $f : S \to \mathbb{R}$ función (dos veces) diferenciable (clase $\mathcal{C}^2$)

Teorema Sea $\mathbf{x}^* \in S$ un punto crítico de f. Se cumple:

- Si $H_f(\mathbf{x}^*)$ es una matriz definida positiva entonces $\mathbf{x}^*$ es un mínimo local (estricto) de f en S.
- Si $H_f(\mathbf{x}^*)$ es una matriz definida negativa entonces $\mathbf{x}^*$ es un máximo local (estricto) de f en S.
- Si $H_f(\mathbf{x}^*)$ es una matriz indefinida entonces $\mathbf{x}^*$ no es un óptimo local de f en S.

Optimización en abiertos. Condición de segundo orden

Hipótesis

- $S \subseteq \mathbb{R}^n$ abierto
- $f : S \to \mathbb{R}$ función (dos veces) diferenciable (clase $\mathcal{C}^2$)

Teorema Sea $\mathbf{x}^* \in S$ un punto crítico de f. Se cumple:

- Si $H_f(\mathbf{x}^*)$ es una matriz definida positiva entonces $\mathbf{x}^*$ es un mínimo local (estricto) de f en S.
- Si $H_f(\mathbf{x}^*)$ es una matriz definida negativa entonces $\mathbf{x}^*$ es un máximo local (estricto) de f en S.
- Si $H_f(\mathbf{x}^*)$ es una matriz indefinida entonces $\mathbf{x}^*$ no es un óptimo local de f en S.

Optimización en abiertos. Condición de segundo orden

Hipótesis

- $S \subseteq \mathbb{R}^n$ abierto
- $f : S \to \mathbb{R}$ función (dos veces) diferenciable (clase $\mathcal{C}^2$)

Teorema Sea $\mathbf{x}^* \in S$ un punto crítico de f. Se cumple:

- Si $H_f(\mathbf{x}^*)$ es una matriz definida positiva entonces $\mathbf{x}^*$ es un mínimo local (estricto) de f en S.
- Si $H_f(\mathbf{x}^*)$ es una matriz definida negativa entonces $\mathbf{x}^*$ es un máximo local (estricto) de f en S.
- Si $H_f(\mathbf{x}^*)$ es una matriz indefinida entonces $\mathbf{x}^*$ no es un óptimo local de f en S.

Optimización en abiertos. Condición de segundo orden

Hipótesis

- $S \subseteq \mathbb{R}^n$ abierto
- $f : S \to \mathbb{R}$ función (dos veces) diferenciable (clase $\mathcal{C}^2$)

Teorema Sea $\mathbf{x}^* \in S$ un punto crítico de f. Se cumple:

- Si $H_f(\mathbf{x}^*)$ es una matriz definida positiva entonces $\mathbf{x}^*$ es un mínimo local (estricto) de f en S.
- Si $H_f(\mathbf{x}^*)$ es una matriz definida negativa entonces $\mathbf{x}^*$ es un máximo local (estricto) de f en S.
- Si $H_f(\mathbf{x}^*)$ es una matriz indefinida entonces $\mathbf{x}^*$ no es un óptimo local de f en S.

Condición de segundo orden. Observaciones

- El resultado anterior no clasifica todos los puntos críticos. Notar que los casos en los cuales el hessiano es semidefinido no han sido clasificados.
- El teorema anterior se demuestra via el desarrollo de Taylor.

Ejemplo Clasificar los puntos críticos de

$$f(x,y) = xy - x^2 - y^2$$

Los puntos críticos de f son las soluciones del sistema

$$\left.\begin{array}{l} \frac{\partial f}{\partial x}(x,y) = y - 2x = 0 \\ \frac{\partial f}{\partial y}(x,y) = x - 2y = 0 \end{array}\right\} \qquad \Rightarrow \qquad (x,y) = (0,0)$$

Dado que su matriz hessiana es $H_f(x,y) = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$

que es definida negativa, deducimos que $(0,0)$ es un máximo local (después veremos que es global).

Condición de segundo orden. Observaciones

- El resultado anterior no clasifica todos los puntos críticos. Notar que los casos en los cuales el hessiano es semidefinido no han sido clasificados.
- El teorema anterior se demuestra via el desarrollo de Taylor.

Ejemplo Clasificar los puntos críticos de

$$f(x,y) = xy - x^2 - y^2$$

Los puntos críticos de f son las soluciones del sistema

$$\left.\begin{array}{l} \frac{\partial f}{\partial x}(x,y) = y - 2x = 0 \\ \frac{\partial f}{\partial y}(x,y) = x - 2y = 0 \end{array}\right\} \qquad \Rightarrow \qquad (x,y) = (0,0)$$

Dado que su matriz hessiana es $H_f(x,y) = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$

que es definida negativa, deducimos que $(0,0)$ es un máximo local (después veremos que es global).

Optimización con restricciones de igualdad (Lagrange)

Hipótesis

- $S = \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{g}(\mathbf{x}) = \mathbf{b} \,(\equiv g_i(\mathbf{x}) = b_i,\ \forall i = 1, \ldots, m\}$
- $f : S \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, funciónes diferenciables

Problema modelo

$$\begin{array}{ll} \text{mín / máx} & f(\mathbf{x}) \\ \text{s.a} & \mathbf{g}(\mathbf{x}) = \mathbf{b},\ (\ g_i(\mathbf{x}) = b_i, \quad i = 1, \ldots, m) \end{array}$$

Lagrangiano asociado

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda^T(\mathbf{g}(\mathbf{x}) - \mathbf{b}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i (g_i(\mathbf{x}) - b_i)$$

Optimización con restricciones de igualdad (Lagrange)

Hipótesis

- $S = \{\mathbf{x} \in \mathbb{R}^n \, : \, \mathbf{g}(\mathbf{x}) = \mathbf{b} \, (\equiv g_i(\mathbf{x}) = b_i, \, \forall i = 1, \ldots, m\}$
- $f : S \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, funciónes diferenciables

Problema modelo

$$\begin{array}{ll} \text{mín / máx} & f(\mathbf{x}) \\ \text{s.a} & \mathbf{g}(\mathbf{x}) = \mathbf{b}, \; (\; g_i(\mathbf{x}) = b_i, \quad i = 1, \ldots, m) \end{array}$$

Lagrangiano asociado

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda^T(\mathbf{g}(\mathbf{x}) - \mathbf{b}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i (g_i(\mathbf{x}) - b_i)$$

Optimización con restricciones de igualdad (Lagrange)

Hipótesis

- $S = \{\mathbf{x} \in \mathbb{R}^n \, : \, \mathbf{g}(\mathbf{x}) = \mathbf{b} \, (\equiv g_i(\mathbf{x}) = b_i, \, \forall i = 1, \ldots, m\}$
- $f : S \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, funciónes diferenciables

Problema modelo

$$\begin{array}{ll} \text{mín / máx} & f(\mathbf{x}) \\ \text{s.a} & \mathbf{g}(\mathbf{x}) = \mathbf{b}, \; (\, g_i(\mathbf{x}) = b_i, \quad i = 1, \ldots, m) \end{array}$$

Lagrangiano asociado

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \boldsymbol{\lambda}^T(\mathbf{g}(\mathbf{x}) - \mathbf{b}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i (g_i(\mathbf{x}) - b_i)$$

Condición de primer orden. Lagrange

Definición: $\mathbf{x}$ se dice un punto regular si se verifica $\text{rang}(J\mathbf{g}) = m$, $\left(J\mathbf{g} = \frac{\partial(g_1,\ldots,g_m)}{\partial(x_1,\ldots x_n)},\ \text{Jacobiano de } \mathbf{g}\right)$

Teorema(Condición de primer orden Lagrange): Si $\mathbf{x}^*$ es un punto regular y óptimo local del problema anterior, existe un único vector de multiplicadores $\lambda^* = (\lambda_i^*)_{i=1,\ldots,m}$ que satisface:

$$\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}^*, \lambda^*) &= 0, \\ \mathbf{g}(\mathbf{x}^*) &= \mathbf{b}, \\ (g_i(\mathbf{x}^*) &= b_i,\ \text{para todo } i = 1, \ldots, m) \end{aligned}$$

Condición de primer orden. Lagrange

Definición: $\mathbf{x}$ se dice un punto regular si se verifica $\text{rang}(J\mathbf{g}) = m$, $\left(J\mathbf{g} = \frac{\partial(g_1,\ldots,g_m)}{\partial(x_1,\ldots x_n)},\ \text{Jacobiano de } \mathbf{g}\right)$

Teorema(Condición de primer orden Lagrange): Si $\mathbf{x}^*$ es un punto regular y óptimo local del problema anterior, existe un único vector de multiplicadores $\boldsymbol{\lambda}^* = (\lambda_i^*)_{i=1,\ldots,m}$ que satisface:

$$\boxed{\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*) &= 0, \\ \mathbf{g}(\mathbf{x}^*) &= \mathbf{b}, \\ (g_i(\mathbf{x}^*) &= b_i,\ \text{para todo } i = 1, \ldots, m) \end{aligned}}$$

Condición de primer orden. Lagrange. Observaciones

1. Al igual que el caso anterior, esta condición es un condición necesaria de óptimo. Permite seleccionar un conjunto de puntos entre los cuales estarán los óptimos.
2. El recíproco de la proposición anterior no es cierto, (es decir, no todo punto crítico del lagrangiano es óptimo)
3. Una demostración de este resultado puede encontrarse en R.K. Sundaram *A First Course in Optimization Theory* pág. 135,
4. El signo con el que se introducen los multiplicadores en el Lagrangiano (negativo en nuestro caso), depende del autor. No tiene mayor importancia, pero afecta al la interpretación de los mismos como precio sombra.
5. Si no se cumple la condición de regularidad los óptimos pueden no cumplir la condición de Lagrange.

Interpretación geométrica. Lagrange

Consideremos el siguiente problema: $\left\{ \begin{array}{ll} \text{opt} & x+y \\ & x^2+y^2=2 \end{array} \right.$

Dibujamos:

- Restricción ($g(x,y)=2$)
- Líneas de nivel de $f(x,y)$
- $\nabla f(x,y)=(1,1)$ (azul)
- $\nabla g(x,y)=(2x,2y)$ (rojo)
- Máximo $(1,1)$
- Mínimo $(-1,-1)$

En los óptimos ∇f y ∇g son paralelos (condición de Lagrange).

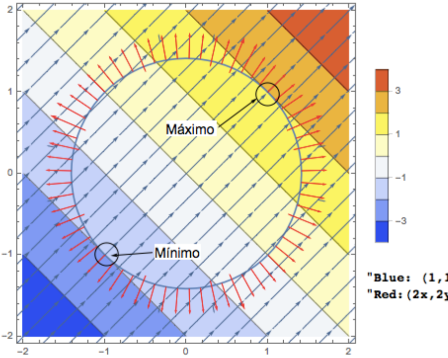
Si ∇f y ∇g no son paralelos, siempre es posible desplazarse a lo largo de la curva que determina la restricción para mejorar/empeorar el valor óptimo.

Interpretación geométrica. Lagrange

Consideremos el siguiente problema: $\begin{cases} \text{opt} & x + y \\ & x^2 + y^2 = 2 \end{cases}$

Dibujamos:

- Restricción ($g(x, y) = 2$)
- Líneas de nivel de $f(x, y)$
- $\nabla f(x, y) = (1, 1)$ (azul)
- $\nabla g(x, y) = (2x, 2y)$ (rojo)
- Máximo $(1, 1)$
- Mínimo $(-1, -1)$

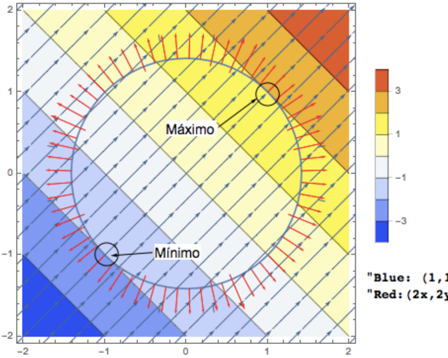


En los óptimos ∇f y ∇g son paralelos (condición de Lagrange).

Si ∇f y ∇g no son paralelos, siempre es posible desplazarse a lo largo de la curva que determina la restricción para mejorar/empeorar el valor óptimo.

Interpretación geométrica. Lagrange

Consideremos el siguiente problema: $\begin{cases} \text{opt} & x+y \\ & x^2+y^2=2 \end{cases}$



Dibujamos:

- Restricción ($g(x,y)=2$)
- Líneas de nivel de $f(x,y)$
- $\nabla f(x,y) = (1,1)$ (azul)
- $\nabla g(x,y) = (2x,2y)$ (rojo)
- Máximo $(1,1)$
- Mínimo $(-1,-1)$

En los óptimos ∇f y ∇g son paralelos (condición de Lagrange).

Si ∇f y ∇g no son paralelos, siempre es posible desplazarse a lo largo de la curva que determina la restricción para mejorar/empeorar el valor óptimo.

Condiciónes de segundo orden. Lagrange

Hipótesis

- $S = \{\mathbf{x} \in \mathbb{R}^n \, : \, \mathbf{g}(\mathbf{x}) = \mathbf{b} \, (\equiv g_i(\mathbf{x}) = b_i, \, \forall i = 1, \ldots, m\}$
- $f : S \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, funciónes de clase $\mathcal{C}^2$

Teorema (Condición de segundo orden): Sea $(\mathbf{x}^*, \lambda^*)$ un punto regular que satisfacen las condiciones de primer orden y $H = L_{\mathbf{xx}}(\mathbf{x}^*, \lambda^*)$. Entonces:

- Si H es definida positiva para todo $\mathbf{z}$ que verifique $J\mathbf{g}(\mathbf{z}) = 0$, $(\mathbf{z} \neq \mathbf{0})$ entonces $\mathbf{x}^*$ es un mínimo local
- Si H es definida negativa para todo $\mathbf{z}$ que verifique $J\mathbf{g}(\mathbf{z}) = 0$, $(\mathbf{z} \neq \mathbf{0})$ entonces $\mathbf{x}^*$ es un máximo local
- Si H es indefinida para todo $\mathbf{z}$ que verifique $J\mathbf{g}(\mathbf{z}) = 0$, entonces $\mathbf{x}^*$ no es un óptimo

Condición de segundo orden. Lagrange. Observaciones

- De nuevo, el resultado no clasifica todos los puntos críticos del Lagrangiano.
- Una demostración de este resultado puede encontrarse en R.K. Sundaram *A First Course in Optimization Theory* pág. 137,
- El caracter del hessiano H debe estudiarse sobre el núcleo del jacobiano de las restricciones, es decir $J\mathbf{g}(z) = 0$, $(\mathbf{z} \neq \mathbf{0})$. No obstante, si el hessiano es definido positivo (o negativo) sobre todo el espacio $\mathbb{R}^n$ no es necesario realizar la restricción al núcleo.

Opt. restricciones igualdad. Ejemplo

Ejercicio

$$\begin{array}{ll} \text{opt} & f(x,y,z) = 2x^2 - 2xy + y^2 + 2z^2 \\ \text{s.a.} & x + y - z = 11 \end{array}$$

Solución

Opt. restricciones igualdad. Ejemplo

Ejercicio

$$\begin{array}{ll} \text{opt} & f(x,y,z) = 2x^2 - 2xy + y^2 + 2z^2 \\ \text{s.a.} & x + y - z = 11 \end{array}$$

Solución

El lagrangiano viene dado por:

$$L(x,y,z,\lambda) = 2x^2 - 2xy + y^2 + 2z^2 - \lambda(x + y - z - 11)$$

y el punto crítico:

$$\left.\begin{array}{l} \dfrac{\partial L}{\partial x} = 4x - 2y - \lambda = 0 \\ \dfrac{\partial L}{\partial y} = -2x + 2y - \lambda = 0 \\ \dfrac{\partial L}{\partial z} = 4z + \lambda = 0 \\ \quad x + y - z = 11 \end{array}\right. \quad \Rightarrow \quad (x,y,z,\lambda) = \boxed{(4, 6, -1, 4)}$$

Opt. restricciones igualdad. Ejemplo

Ejercicio

$$\begin{array}{ll} \text{opt} & f(x,y,z) = 2x^2 - 2xy + y^2 + 2z^2 \\ \text{s.a.} & x + y - z = 11 \end{array}$$

Solución

Por último, dado que HL es definido positivo, concluimos que $\boxed{(4,6,-1,4)}$ es un mínimo local.

$$HL = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Interpretación económica multiplicadores

Si definimos la función:

$$F(\mathbf{b}) := f(\mathbf{x}^*(\mathbf{b}))$$

y asumimos suficiente suavidad de las funciones, no es difícil probar (Inténtalo!), que:

$$\boxed{\frac{\partial F}{\partial b_i} = \frac{\partial f(\mathbf{x}^*(\mathbf{b}))}{\partial b_i} = \lambda_i}$$

Lo que significa que λ_i indica cómo cambia la función objetivo en el óptimo al incrementar una unidad la cantidad disponible en el recurso b_i.

Este valor es un precio sombra o valor marginal del recurso i-ésimo: la cantidad máxima que se puede estar dispuesto a pagar por incrementar la disponibilidad del recurso i-ésimo.

Optimización con restricciones de desigualdad. KKT

Hipótesis

- $S = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) = b_i,\ i \in \mathcal{E},\ g_i(\mathbf{x}) \leqslant b_i,\ i \in \mathcal{I}\}$
- $f : S \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, funciónes diferenciables

Problema modelo

$$\begin{array}{lll} \text{mín} & f(\mathbf{x}) & \\ \text{s.a} & g_i(\mathbf{x}) = b_i, & i \in \mathcal{E} \\ & g_i(\mathbf{x}) \leqslant b_i, & i \in \mathcal{I} \end{array}$$

Lagrangiano asociado

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda^T(\mathbf{g}(\mathbf{x}) - \mathbf{b}) = f(\mathbf{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i (g_i(\mathbf{x}) - b_i)$$

Optimización con restricciones de desigualdad. KKT

Hipótesis

- $S = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) = b_i,\ i \in \mathcal{E},\ g_i(\mathbf{x}) \leqslant b_i,\ i \in \mathcal{I}\}$
- $f : S \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, funciónes diferenciables

Problema modelo

$$\begin{array}{lll} \text{mín} & f(\mathbf{x}) & \\ \text{s.a} & g_i(\mathbf{x}) = b_i, & i \in \mathcal{E} \\ & g_i(\mathbf{x}) \leqslant b_i, & i \in \mathcal{I} \end{array}$$

Lagrangiano asociado

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda^T(\mathbf{g}(\mathbf{x}) - \mathbf{b}) = f(\mathbf{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i (g_i(\mathbf{x}) - b_i)$$

Optimización con restricciones de desigualdad. KKT

Hipótesis

- $S = \{\mathbf{x} \in \mathbb{R}^n \; : g_i(\mathbf{x}) = b_i, \; i \in \mathcal{E}, \; g_i(\mathbf{x}) \leqslant b_i, \; i \in \mathcal{I}\}$
- $f : S \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, funciónes diferenciables

Problema modelo

$$
\begin{array}{lll}
\text{mín} & f(\mathbf{x}) & \\
\text{s.a} & g_i(\mathbf{x}) = b_i, & i \in \mathcal{E} \\
 & g_i(\mathbf{x}) \leqslant b_i, & i \in \mathcal{I}
\end{array}
$$

Lagrangiano asociado

$$
L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda^T(\mathbf{g}(\mathbf{x}) - \mathbf{b}) = f(\mathbf{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i (g_i(\mathbf{x}) - b_i)
$$

Condición de primer orden. KKT

Defnción Una restricción $g_i(x)$ se dice *activa* en el punto $\mathbf{x}$ si se verifica $\boxed{g_i(\mathbf{x}) = b_i}$

Defnción Un punto $\mathbf{x}$ se dice un punto regular si el rango del jacobiano de las restricciones activas en el punto $\mathbf{x}$ es máximo, es decir:

$$A = \{i \,:\, g_i(\mathbf{x}) = b_i\}, \quad \mathbf{g}_A = \{g_i \,:\, i \in A\} \Rightarrow \operatorname{rang}(J\mathbf{g}_A) = |A|$$

Condición de primer orden. KKT

Teorema(Condición de primer order KKT) Si $\mathbf{x}^*$ es un punto regular y óptimo del problema anterior, existe un vector de multiplicadores $\boldsymbol{\lambda}^* = (\lambda_i^*)_{i\in\mathcal{E}\cup\mathcal{I}}$ que satisface:

Máximo local

$$\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*) &= 0, \\ g_i(\mathbf{x}^*) &= b_i, \ \forall\ i \in \mathcal{E}, \\ g_i(\mathbf{x}^*) &\leqslant b_i, \ \forall\ i \in \mathcal{I}\ , \\ \lambda_i^* &\geqslant 0, \ \forall\ i \in \mathcal{I}\ , \\ \lambda_i^* g_i(\mathbf{x}^*) &= 0, \ \forall\ i \in \mathcal{E} \cup \mathcal{I} \end{aligned}$$

Mínimo local

$$\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*) &= 0, \\ g_i(\mathbf{x}^*) &= b_i, \ \forall\ i \in \mathcal{E}, \\ g_i(\mathbf{x}^*) &\leqslant b_i, \ \forall\ i \in \mathcal{I}\ , \\ \lambda_i^* &\leqslant 0, \ \forall\ i \in \mathcal{I}\ , \\ \lambda_i^* g_i(\mathbf{x}^*) &= 0, \ \forall\ i \in \mathcal{E} \cup \mathcal{I} \end{aligned}$$

Condición de primer orden KKT. Observaciones

- Como los casos anteriores, se trata de una condición necesaria.
- Una demostración del resultado se puede encontrar en R.K. Sundaram *A First Course in Optimization Theory* pág. 137
- En sistema resultante es una inecuación lo cual complica su resolución.
- Notar que se tiene condiciones diferentes para máximos y mínimos
- El *signo* de multiplicadores ($-$ en nuestro caso) y restricciones ($\leqslant$ en nuestro caso), varia según bibliografía, lo cual ocasiona un caos, pues ambos alteran la condición 4 de la pág. anterior.
- En ocasiones, es más sencillo tratar el problema como uno de optimización en abiertos y otros con restricciones de igualdad.
- Observar que desde la última condición es fácil deducir:
  - Si las restricciones NO son *activas* $\quad\Rightarrow\quad \lambda_i = 0$
  - Si el multiplicador $\lambda_i > 0 \quad\Rightarrow\quad$ La restricción es *activa*

Condiciónes de segundo orden. KKT

Hipótesis

- $S = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) = b_i,\ i \in \mathcal{E},\ g_i(\mathbf{x}) \leqslant b_i,\ i \in \mathcal{I}\}$
- $f : S \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$ de clase $\mathcal{C}^2$

Teorema (Condición de segundo orden KKT) Sea $(\mathbf{x}^*, \lambda^*)$ tal que:

1. Es un punto regular
2. las restricciones NO activas no lo son en una vecindad de $\mathbf{x}^*$
3. satisface las condiciones de primer orden de KKT

Sea, además, λ_A^* el conjunto de multiplicadores asociado a las restricciones activas. Entonces:

Si $(\mathbf{x}^*, \lambda_A^*)$ satisface las condiciones de segundo orden de óptimo de Lagrange sobre el conjunto de restricciones activas, entonces $\mathbf{x}^*$ es un óptimo local.

Condición de segundo orden. Observaciones

- La demostración de este teorema se reduce al caso de Lagrange.
- Los multiplicadores tienen la misma interpretación económica que en el caso de Lagrange.

$$\frac{\partial f(\mathbf{x}^*(\mathbf{b}))}{\partial b_i} = \lambda_i^*$$

Opt. con restricciones de desigualdad. Ejemplo (I)

Ejercicio

$$\begin{aligned} \text{opt} \quad & x + 2y \\ s.a. \quad & g_1(x,y) = x - 2y - 2 \leqslant 0 \\ & g_2(x,y) = x^2 + y^2 - 4 \leqslant 0 \end{aligned}$$

Solución El lagrangiano viene dado por:

$$L(x,y,\lambda,\mu) = x + 2y - \lambda(x - 2y - 2) - \mu(x^2 + y^2 - 4)$$

las condiciones de KKT:

$$\begin{aligned} \frac{\partial L}{\partial x} &= 1 - \lambda - 2\mu x = 0 \\ \frac{\partial L}{\partial y} &= 2 + 2\lambda - 2\mu y = 0 \\ & x - 2y - 2 \leqslant 0 \\ & x^2 + y^2 - 4 \leqslant 0 \\ & \lambda(x - 2y - 2) = 0 \\ & \mu(x^2 + y^2 - 4) = 0 \\ & \text{Mínimo:} \quad \lambda \leqslant 0,\ \mu \leqslant 0 \\ & \text{Máximo:} \quad \lambda \geqslant 0,\ \mu \geqslant 0 \end{aligned}$$

Opt. con restricciones de desigualdad. Ejemplo (II)

Distinguimos cuatro casos:

1. Caso: $g_1(x,y) < 0$, $g_2(x,y) < 0$ $\Rightarrow$ $\lambda = \mu = 0$. Conduce a un sistema incompatible
2. Caso: $g_1(x,y) = 0$, $g_2(x,y) < 0$ $\Rightarrow$ $\mu = 0$. Conduce a un sistema incompatible
3. Caso: $g_1(x,y) < 0$, $g_2(x,y) = 0$ $\Rightarrow$ $\lambda = 0$
4. Caso: $g_1(x,y) = 0$, $g_2(x,y) = 0$

Opt. con restricciones de desigualdad. Ejemplo (II)

Distinguimos cuatro casos:

1. Caso: $g_1(x,y) < 0$, $g_2(x,y) < 0 \quad \Rightarrow \quad \lambda = \mu = 0$. Conduce a un sistema incompatible
2. Caso: $g_1(x,y) = 0$, $g_2(x,y) < 0 \quad \Rightarrow \quad \mu = 0$. Conduce a un sistema incompatible
3. Caso: $g_1(x,y) < 0$, $g_2(x,y) = 0 \quad \Rightarrow \quad \lambda = 0$
4. Caso: $g_1(x,y) = 0$, $g_2(x,y) = 0$

Opt. con restricciones de desigualdad. Ejemplo (II)

Distinguimos cuatro casos:

1. Caso: $g_1(x, y) < 0$, $g_2(x, y) < 0$ $\Rightarrow$ $\lambda = \mu = 0$. Conduce a un sistema incompatible
2. Caso: $g_1(x, y) = 0$, $g_2(x, y) < 0$ $\Rightarrow$ $\mu = 0$. Conduce a un sistema incompatible
3. Caso: $g_1(x, y) < 0$, $g_2(x, y) = 0$ $\Rightarrow$ $\lambda = 0$ Resulta el sistema:

$$\begin{aligned} 1 - 2\mu x &= 0 \\ 2 - 2\mu y &= 0 \\ x^2 + y^2 - 4 &= 0 \\ x - 2y - 2 &< 0 \end{aligned}$$

De las tres primeras ecuaciones resultan: $p_1 = \left(\frac{2}{\sqrt{5}}, \frac{4}{\sqrt{5}}\right)$ y $p_2 = \left(-\frac{2}{\sqrt{5}}, -\frac{4}{\sqrt{5}}\right)$. p_2 debe descartarse pues no cumple la 4a ecuación. El punto p_1 conduce a $\mu > 0$, luego es un posible máximo local

4. Caso: $g_1(x, y) = 0$, $g_2(x, y) = 0$

Opt. con restricciones de desigualdad. Ejemplo (II)

Distinguimos cuatro casos:

1. Caso: $g_1(x,y) < 0$, $g_2(x,y) < 0 \quad \Rightarrow \quad \lambda = \mu = 0$. Conduce a un sistema incompatible
2. Caso: $g_1(x,y) = 0$, $g_2(x,y) < 0 \quad \Rightarrow \quad \mu = 0$. Conduce a un sistema incompatible
3. Caso: $g_1(x,y) < 0$, $g_2(x,y) = 0 \quad \Rightarrow \quad \lambda = 0 \quad \mu = \frac{\sqrt{5}}{4}$, $(x,y) = \left(\frac{2}{\sqrt{5}}, \frac{4}{\sqrt{5}}\right)$ candidato a máximo local
4. Caso: $g_1(x,y) = 0$, $g_2(x,y) = 0$

Opt. con restricciones de desigualdad. Ejemplo (II)

Distinguimos cuatro casos:

1. Caso: $g_1(x,y) < 0$, $g_2(x,y) < 0 \quad \Rightarrow \quad \lambda = \mu = 0$. Conduce a un sistema incompatible
2. Caso: $g_1(x,y) = 0$, $g_2(x,y) < 0 \quad \Rightarrow \quad \mu = 0$. Conduce a un sistema incompatible
3. Caso: $g_1(x,y) < 0$, $g_2(x,y) = 0 \quad \Rightarrow \quad \lambda = 0 \quad \mu = \frac{\sqrt{5}}{4}$, $(x,y) = \left(\frac{2}{\sqrt{5}}, \frac{4}{\sqrt{5}}\right)$ candidato a máximo local
4. Caso: $g_1(x,y) = 0$, $g_2(x,y) = 0$ Resulta el sistema:

$$
\begin{aligned}
1 - 2\mu x &= 0 \\
2 - 2\mu y &= 0 \\
x^2 + y^2 - 4 &= 0 \\
x - 2y - 2 &= 0
\end{aligned}
$$

La 3a y 4a ecuaciones conducen a : $p_3 = (2,0)$ y $p_4 = \left(-\frac{6}{5}, -\frac{8}{5}\right)$. p_3 implica multiplicadores de distinto signo, luego se descarta. Desde el p_4 resulta $\lambda = -\frac{1}{5}$ y $\mu = -\frac{1}{2}$ luego es un posible mínimo local.

Opt. con restricciones de desigualdad. Ejemplo (II)

Distinguimos cuatro casos:

1. Caso: $g_1(x,y) < 0$, $g_2(x,y) < 0 \quad \Rightarrow \quad \lambda = \mu = 0$. Conduce a un sistema incompatible
2. Caso: $g_1(x,y) = 0$, $g_2(x,y) < 0 \quad \Rightarrow \quad \mu = 0$. Conduce a un sistema incompatible
3. Caso: $g_1(x,y) < 0$, $g_2(x,y) = 0 \quad \Rightarrow \quad \lambda = 0 \quad \mu = \frac{\sqrt{5}}{4}$, $(x,y) = \left(\frac{2}{\sqrt{5}}, \frac{4}{\sqrt{5}}\right)$ candidato a máximo local
4. Caso: $g_1(x,y) = 0$, $g_2(x,y) = 0 \quad \Rightarrow \quad \lambda = -\frac{1}{5}$, $\mu = -\frac{1}{2}$, $(x,y) = \left(-\frac{6}{5}, -\frac{8}{5}\right)$ candidato a mínimo local

Opt. con restricciones de desigualdad. Ejemplo (II)

Distinguimos cuatro casos:

1. Caso: $g_1(x,y) < 0$, $g_2(x,y) < 0 \quad \Rightarrow \quad \lambda = \mu = 0$. Conduce a un sistema incompatible
2. Caso: $g_1(x,y) = 0$, $g_2(x,y) < 0 \quad \Rightarrow \quad \mu = 0$. Conduce a un sistema incompatible
3. Caso: $g_1(x,y) < 0$, $g_2(x,y) = 0 \quad \Rightarrow \quad \lambda = 0 \quad \mu = \frac{\sqrt{5}}{4}$, $(x,y) = \left(\frac{2}{\sqrt{5}}, \frac{4}{\sqrt{5}}\right)$ candidato a máximo local
4. Caso: $g_1(x,y) = 0$, $g_2(x,y) = 0 \quad \Rightarrow \quad \lambda = -\frac{1}{5}$, $\mu = -\frac{1}{2}$, $(x,y) = \left(-\frac{6}{5}, -\frac{8}{5}\right)$ candidato a mínimo local

Veamos ahora, que ambos puntos verifican la condición de regularidad:

$$J\mathbf{g} = \begin{pmatrix} 1 & -2 \\ 2x & 2y \end{pmatrix} \quad \Rightarrow \quad \operatorname{rang}\left[J\mathbf{g}\left(\frac{2}{\sqrt{5}}, \frac{4}{\sqrt{5}}\right)\right] = \operatorname{rang}\left[J\mathbf{g}\left(-\frac{6}{5}, -\frac{8}{5}\right)\right] = 2$$

Opt. con restricciones de desigualdad. Ejemplo (III)

Por último, analicemos la condición de segundo orden:

1. Punto: $\left(\frac{2}{\sqrt{5}}, \frac{4}{\sqrt{5}}\right)$, $\mu = \frac{\sqrt{5}}{4}$ Caso: $g_1(x,y) < 0$, $g_2(x,y) = 0$.
   Restricción activa g_2.

   $$L(x,y,\mu) = x + 2y - \mu(x^2 + y^2 - 4)$$

   $$HL = \begin{pmatrix} -2\mu & 0 \\ 0 & -2\mu \end{pmatrix} \Rightarrow \text{Def. Neg. en} \left(x = \frac{2}{\sqrt{5}}, y = \frac{4}{\sqrt{5}}, \mu = \frac{\sqrt{5}}{4}\right)$$

   $\Rightarrow$ Máximo local

2. Punto $\left(-\frac{6}{5}, -\frac{8}{5}\right)$ $\lambda = -\frac{1}{5}$, $\mu = -\frac{1}{2}$. Caso: $g_1(x,y) = 0$, $g_2(x,y) = 0$.
   Restricciones activas g_1, g_2.

   $$L(x,y,\mu) = x + 2y - \lambda(x - 2y - 2) - \mu(x^2 + y^2 - 4)$$

   $$HL = \begin{pmatrix} -2\mu & 0 \\ 0 & -2\mu \end{pmatrix} \Rightarrow \text{Def. Pos.} \left(x = -\frac{6}{5}, y = -\frac{8}{5}, \lambda = -\frac{1}{5}, \mu = -\frac{1}{2}\right)$$

   $\Rightarrow$ Mínimo local

Opt. con restricciones de desigualdad. Ejemplo (III)

Por último, analicemos la condición de segundo orden:

1. Punto: $\left(\frac{2}{\sqrt{5}}, \frac{4}{\sqrt{5}}\right)$, $\mu = \frac{\sqrt{5}}{4}$ Caso: $g_1(x, y) < 0$, $g_2(x, y) = 0$.
   Restricción activa g_2.

$$L(x, y, \mu) = x + 2y - \mu(x^2 + y^2 - 4)$$

$$HL = \begin{pmatrix} -2\mu & 0 \\ 0 & -2\mu \end{pmatrix} \Rightarrow \text{Def. Neg. en} \left(x = \frac{2}{\sqrt{5}}, y = \frac{4}{\sqrt{5}}, \mu = \frac{\sqrt{5}}{4}\right)$$

$\Rightarrow$ Máximo local

2. Punto $\left(-\frac{6}{5}, -\frac{8}{5}\right)$ $\lambda = -\frac{1}{5}$, $\mu = -\frac{1}{2}$. Caso: $g_1(x, y) = 0$, $g_2(x, y) = 0$.
   Restricciones activas g_1, g_2.

$$L(x, y, \mu) = x + 2y - \lambda(x - 2y - 2) - \mu(x^2 + y^2 - 4)$$

$$HL = \begin{pmatrix} -2\mu & 0 \\ 0 & -2\mu \end{pmatrix} \Rightarrow \text{Def. Pos.} \left(x = -\frac{6}{5}, y = -\frac{8}{5}, \lambda = -\frac{1}{5}, \mu = -\frac{1}{2}\right)$$

$\Rightarrow$ Mínimo local

Optimización convexa diferenciable

Introducción

Conjuntos y funciones convexos

Resultados generales de optimización

Optimización convexa (conceptos básicos)

Optimización diferenciable

Optimización convexa diferenciable
- Abiertos
- Restricciones de igualdad (Lagrange)
- Restricciones de desigualdad (KKT)

Optimización convexa diferenciable

Estudiaremos los problemas de optimización de la sección anterior bajo hipótesis de convexidad y diferenciabiliad. Con esta nueva hipótesis:

- Los extremos locales serán además globales.
- La convexidad sustituirá a la condición de segundo orden.

Como en la sección anterior distinguiremos:

1. S abierto y convexo
2. S determinado por una restricción de igualdad y convexo:

$$S = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) = \mathbf{b}\}$$

3. S determinado por restricción de desigualdad y convexo:

$$S = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \leqslant \mathbf{b}\}$$

Optimización convexa diferenciable

Estudiaremos los problemas de optimización de la sección anterior bajo hipótesis de convexidad y diferenciabiliad. Con esta nueva hipótesis:

- Los extremos locales serán además globales.
- La convexidad sustituirá a la condición de segundo orden.

Como en la sección anterior distinguiremos:

1. S abierto y convexo
2. S determinado por una restricción de igualdad y convexo:
$$S = \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{g}(\mathbf{x}) = \mathbf{b}\}$$
3. S determinado por restricción de desigualdad y convexo:
$$S = \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{b}\}$$

Optimización en abiertos

Hipótesis En esta sección consideraremos

- $S \subseteq \mathbb{R}^n$ abierto y convexo
- $f : S \to \mathbb{R}$ función diferenciable (clase $\mathcal{C}^1$) y convexa / o cóncava

Teorema (Mínimo) Sea $f : S \to \mathbb{R}$ convexa.

$$\mathbf{x}^* \text{ es punto crítico} \quad \Leftrightarrow \quad \mathbf{x}^* \text{ es mínimo}$$

Teorema (Máximo) Sea $f : S \to \mathbb{R}$ cóncava.

$$\mathbf{x}^* \text{ es punto crítico} \quad \Leftrightarrow \quad \mathbf{x}^* \text{ es máximo}$$

Optimización en abiertos

Hipótesis En esta sección consideraremos

- $S \subseteq \mathbb{R}^n$ abierto y convexo
- $f : S \to \mathbb{R}$ función diferenciable (clase $\mathcal{C}^1$) y convexa / o cóncava

Teorema (Mínimo) Sea $f : S \to \mathbb{R}$ convexa.

$$\mathbf{x}^* \text{ es punto crítico} \quad \Leftrightarrow \quad \mathbf{x}^* \text{ es mínimo}$$

Teorema (Máximo) Sea $f : S \to \mathbb{R}$ cóncava.

$$\mathbf{x}^* \text{ es punto crítico} \quad \Leftrightarrow \quad \mathbf{x}^* \text{ es máximo}$$

Optimización en abiertos

Hipótesis En esta sección consideraremos

- $S \subseteq \mathbb{R}^n$ abierto y convexo
- $f : S \to \mathbb{R}$ función diferenciable (clase $\mathcal{C}^1$) y convexa / o cóncava

Teorema (Mínimo) Sea $f : S \to \mathbb{R}$ convexa.

$$\mathbf{x}^* \text{ es punto crítico} \quad \Leftrightarrow \quad \mathbf{x}^* \text{ es mínimo}$$

Teorema (Máximo) Sea $f : S \to \mathbb{R}$ cóncava.

$$\mathbf{x}^* \text{ es punto crítico} \quad \Leftrightarrow \quad \mathbf{x}^* \text{ es máximo}$$

Optimización en abiertos. Observaciones

- Los resultados anteriores son una consecuencia directa del teorema local-global y la condición de primer orden para funciones diferenciables.
- Si además la función f es estrictamente convexa / cóncava se puede asegurar que el mínimo / máximo es único
- La convexidad evita analizar la condición de segundo orden.

Ejemplo

$$\max_{(x,y)\in\mathbb{R}^2} f(x,y) = xy - x^2 - y^2$$

Notar que f es cóncava, pues su hessiano es definidio negativo:

$$H_f(x,y) = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

De modo que el único punto crítico será máximo: $(x,y) = (0,0)$

Optimización con restricciones de igualdad (Lagrange)

Hipótesis

- $S = \{\mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{g}(\mathbf{x}) = \mathbf{b}\}$ y convexo
  $\Rightarrow g_i$ funciones lineales .
- $f : S \to \mathbb{R}$, función diferenciable y convexa / cóncava

Teorema(Mínimo) Si $f : S \to \mathbb{R}$ es convexa

Todo punto que cumpla la condición de primer orden de Lagrange es mínimo

Teorema(Máximo) Si $f : S \to \mathbb{R}$ es cóncaxa

Todo punto que cumpla la condición de primer orden de Lagrange es máximo

Opt. restricciones igualdad. Ejemplo

Ejercicio Calcular los extremos globales de
$f(x, y, z) = 2x^2 - 2xy + y^2 + 2z^2$ sobre $x + y - z = 11$.

Solución Observar que dado que Hf es definido positivo:

$$Hf = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

podemos concluir que f es convexa y por tanto que estamos en las condiciones de Lagrange para el caso convexo. Calculamos el punto crítico del Lagrangiano:

$$\left.\begin{aligned} \frac{\partial L}{\partial x} &= 4x - 2y - \lambda = 0 \\ \frac{\partial L}{\partial y} &= -2x + 2y - \lambda = 0 \\ \frac{\partial L}{\partial z} &= 4z + \lambda = 0 \\ x + y - z &= 11 \end{aligned}\right. \qquad \Rightarrow \qquad (x, y, z, \lambda) = \boxed{(4, 6, -1, 4)}$$

y concluimos que $\boxed{(4, 6, -1, 4)}$ es un mínimo.

Optimización con restricciones de desigualdad. KKT

Hipótesis

- $S = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) = b_i,\ i \in \mathcal{E},\ g_i(\mathbf{x}) \leqslant b_i,\ i \in \mathcal{I}\}$ y convexo

$$\Rightarrow \quad \begin{array}{ll} g_i\ i \in \mathcal{E} & \text{lineales} \\ g_i\ i \in \mathcal{I} & \text{convexas} \end{array}$$

- $f : S \to \mathbb{R}$, función diferenciable y convexa / cóncava

Teorema(Mínimo) Si $f : S \to \mathbb{R}$ es convexa

Todo punto que cumpla la condición de primer orden de mínimo de KKT es mínimo

Teorema(Máximo) Si $f : S \to \mathbb{R}$ es cóncaxa

Todo punto que cumpla la condición de primer orden de máximo de KKT es máximo

Optimización con restricciones de desigualdad. KKT. Observaciones

- En ocasiones en el conjunto S, puede venir definido como:
$$S = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \geqslant \mathbf{b}\}$$
en ese caso, cada componente de $\mathbf{g}$, debe ser cóncava. Recordar que si f es convexa entonces $-f$ es cóncava.
- Es conveniente recordar que la hipótesis es la convexidad de S y no tanto la naturaleza de las función $\mathbf{g}$, pues en general podrían aparecer con distinto signo de desigualdad según el problema.
- En este caso es posible probar una condición necesaria y suficiente de óptimo evitando incluso la condición de regularidad (rang$(J\mathbf{g}_A) = |A|$), por la denominada *slater's condition*:
$$\exists\, \tilde{\mathbf{x}} \text{ tal que } \mathbf{g}(\tilde{\mathbf{x}}) < \mathbf{b} \quad (g_i(\tilde{x}_i) < b_i \quad \forall i)$$

Opt. con restricciones de desigualdad. Ejemplo

Ejercicio Determinar los extremos globales

$$
\begin{aligned}
\text{opt} \quad & x + 2y \\
s.a. \quad & g_1(x,y) = x - 2y - 2 \leqslant 0 \\
& g_2(x,y) = x^2 + y^2 - 4 \leqslant 0
\end{aligned}
$$

Solución Observar que la función objetivo es convexa y también cóncava (es lineal). Además las dos restricciones son convexas, y por el sentido de la desigualdad podemos concluir que el conjunto factible es convexo.

Estamos por tanto en las hipótesis del th de KKT bajo condiciones de convexidad.

Utilizando los cálculos que hicimos en el capítulo anterior podemos concluir:

- $\left(\frac{2}{\sqrt{5}}, \frac{4}{\sqrt{5}}\right) \quad \Rightarrow$ máximo global
- $\left(-\frac{6}{5}, -\frac{8}{5}\right) \quad \Rightarrow$ mínimo global

Nota El mismo resultado se podría haber obtenido via el teorema de Weierstrass.

Tema 2

Optimización convexa diferenciable

Introducción

Conjuntos y funciones convexos

Resultados generales de optimización

Optimización convexa (conceptos básicos)

Optimización diferenciable

Optimización convexa diferenciable

Resolución con Mathematica

Resolución con Mathematica

Comandos útiles para la resolución con Mathematica:

- Cálculo de derivadas: `D`, `Grad`
- Resolución de sistemas: `Solve`, `NSolve`
- Cálculo de valores y vectores propios: `Eigenvalues`, `Eigenvectors`
- Representación de regiones en $\mathbb{R}^2$ y $\mathbb{R}^3$: `RegionPlot`, `RegionPlot3D`
- Optimización directa: `Maximize`, `Minimize`, `NMaximize`, `NMinimize`

Tema 3

Programación Lineal

Introducción

Programación Lineal

El método símplex

Análisis de Sensibilidad

Dualidad

Tema 3

Programación Lineal

Introducción
- Ejemplos
- Métodos de mejora

Programación Lineal

El método símplex

Análisis de Sensibilidad

Dualidad

Introducción

- La Programación Lineal es una rama de la Optimización, Programación Matemática o Management Science, que estudia la resolución de problemas de optimización lineales.
- En términos matemáticos

$$\boxed{\begin{array}{ll} \min & f(x) \\ \text{s.a} & g(x) = 0 \\ & h(x) \leqslant 0 \end{array}}$$

  donde $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ y $h : \mathbb{R}^n \to R^p$ son lineales.
- Juega un papel primordial en muchas facetas de la economía, la técnica, la planificación, la gestión, . . . , como instrumento de ayuda a la toma de decisiones que tienen que ver con la asignación óptima de recursos con diversas funciones de utilidad.

Un poco de Historia

- Orígenes en Newton, Leibnitz, Lagrange y Fourier
- 1930s Kantorovich; primeras aplicaciones en economía
- 1940s Se crea por G. Dantzig, John von Neumann y L. Kantorovich el área de conocimiento Programación Lineal
- 1947 G. Dantzig formula el Método Símplex para resolver problemas logísticos militares
- 1950-1970 Se disparan las aplicaciones en múltiples áreas: control, estructuras, transporte, mecánica,. . .
- 1979 L. Khachiyan desarrolla el algoritmo del elipsoide
- 1984 N. Karmarkar desarrolla un método basado en puntos interiores.
- 1995-hoy Desarrollo de variantes de punto interior muy eficientes para problemas de grandes dimensiones

Problema de optimización

$$\boxed{\begin{array}{ll} \min & f(x) \\ \text{s.a} & x \in S \\ & [\, x \in X] \end{array}} \quad (1)$$

donde,

$$f : X \longrightarrow \mathbb{R}$$
$$S \subseteq X$$

X : Conjunto de referencia ($X = \mathbb{R}^n$)

f : Función objetivo.

S : Conjunto factible. $x \in S$ *solución* factible. S suele venir dado como una serie de condiciones impuestas a los elementos de X:

$$S := \{x \in X \,:\, g_i(x) \leqslant 0, i = 1, \ldots, m\}$$

Las soluciones (óptimas) son los mínimos de la función f restringida al conjunto S: $f_{|S}$:

$$\mathcal{O} = \{z \in S : f(z) \leqslant f(x)\ \forall x \in S\}$$

Si $\mathcal{O} = \emptyset$ problema no factible (sin solución) , en otro caso $\mathcal{O} \neq \emptyset$ problema factible (con solución).

Problema de optimización

$$
\boxed{\begin{array}{ll} \text{mín} & f(x) \\ \text{s.a} & x \in S \\ & [\, x \in X] \end{array}} \quad (1)
$$

donde,

$$f : X \longrightarrow \mathbb{R}$$
$$S \subseteq X$$

X : Conjunto de referencia ($X = \mathbb{R}^n$)

f : Función objetivo.

S : Conjunto factible. $x \in S$ *solución* factible. S suele venir dado como una serie de condiciones impuestas a los elementos de X:

$$S := \{x \in X \, : \, g_i(x) \leqslant 0, i = 1, \ldots, m\}$$

Las soluciones (óptimas) son los mínimos de la función f restringida al conjunto S: $f_{|S}$:

$$\mathcal{O} = \{z \in S \, : \, f(z) \leqslant f(x) \; \forall x \in S\}$$

Si $\mathcal{O} = \emptyset$ problema no factible (sin solución) , en otro caso

$\mathcal{O} \neq \emptyset$ problema factible (con solución).

Programación lineal

Un problema de programación lineal tiene las siguientes características:

- f lineal:

$$f(x) = f(x_1, \ldots, x_n) = c_1 x_1 + \ldots c_n x_m = \mathbf{c}'\mathbf{x}$$

- g_i son funciones afines:

$$g_i(x) = g(x_1, \ldots, x_n) = a_{i1} x_1 + \ldots a_{in} x_n + b_i, \quad i = 1, \ldots, m.$$

- Es habitual imponer condiciones de positividad sobre las variables.

$$x_i \geqslant 0, \ i = 1, \ldots n$$

Con todo ello un problema de programación lineal se suele escribir de la siguiente forma:

Programación lineal

$$\begin{array}{ll} \min & \mathbf{c}'\mathbf{x} \\ \text{s.a} & A\mathbf{x} \begin{pmatrix} = \\ \geqslant \\ \leqslant \end{pmatrix} b \\ & x_i \geqslant 0,\ i = 1, \dots n \end{array}$$

- n número de variables o incógnitas.
- m número de restricciones.

donde,

$$\mathbf{c} = (c_1, \dots, c_n)$$

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

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Emisión de deuda pública

El ayuntamiento de una capital de provincia tiene comprometido invertir en proyectos de infraestructura en cuatro años 2.000, 4.000, 8.000 y 5.000 millones de euros, respectivamente. Se supone que todo ese dinero está disponible el 1 de enero del año en que se gasta.

Para financiar estas inversiones el ayuntamiento planea emitir unos bonos a 20 años con un interés del $7\,\%$ para la deuda emitida el primer año, del $6\,\%$ para la del segundo año, $6,5\,\%$ para la del tercer año y del $7,5\,\%$ para la del cuarto año. Los intereses se empiezan a pagar inmediatamente.

Si parte del dinero recaudado se deposita en cuentas a plazo fijo, el ayuntamiento podría obtener el $6\,\%$ de interés el primer año, el $5,5\,\%$ el segundo y el $4,5\,\%$ el tercero.

¿Cuál es el plan óptimo de financiación?

Solución

Variables

x_i cantidad emitida en deuda el año i (m. mill.)
y_i cantidad depositada en cc el año i (m. mill.)

Datos

	Inversión	Interés bono	Interés cc
1er año	2	0.07	0.06
2o año	4	0.06	0.055
3er año	8	0.065	0.045
4o año	5	0.075	–

Solución

Planteamiento

$$
\begin{array}{llllllll}
\text{mín} & 20 * (0.07x_1 + 0.06x_2 + 0.065x_3 + 0.075x_4) \\
s.a. \\
x_1 & & & & -y_1 & & & = 2 \\
& x_2 & & & +1.06y_1 & -y_2 & & = 4 \\
& & x_3 & & & +1.055y_2 & -y_3 & = 8 \\
& & & x_4 & & & 1.045y_3 & = 5 \\
& x_1, x_2, x_3, x_4, y_1, y_2, y_3 \geqslant 0
\end{array}
$$

Solución

$x_1 = 2,\quad x_2 = 16.1182,\quad x_3 = 0,\quad x_4 = 0$

$y_1 = 0\quad y_2 = 12.1182\quad y_3 = 4.7847$

Solución

Planteamiento

$$
\begin{array}{ll}
\text{mín} & 20 * (0.07x_1 + 0.06x_2 + 0.065x_3 + 0.075x_4) \\
s.a. &
\end{array}
$$

$$
\begin{array}{cccccccccc}
x_1 & & & & -y_1 & & & = & 2 \\
& x_2 & & & +1.06y_1 & -y_2 & & = & 4 \\
& & x_3 & & & +1.055y_2 & -y_3 & = & 8 \\
& & & x_4 & & & 1.045y_3 & = & 5 \\
\end{array}
$$

$$x_1, x_2, x_3, x_4, y_1, y_2, y_3 \geqslant 0$$

Solución

$$
\boxed{
\begin{array}{llll}
x_1 = 2, & x_2 = 16.1182, & x_3 = 0, & x_4 = 0 \\
y_1 = 0 & y_2 = 12.1182 & y_3 = 4.7847 &
\end{array}
}
$$

Cartera de valores

Un inversor es propietario b_i participaciones de varios valores bursátiles A_i, $i = 1, 2, \ldots, m$. Los precios actuales de estos valores son v_i. Considérese que se pueden predecir los dividendos que se pagarán al final del año. Es decir, A_i pagará d_i y tendrá un nuevo precio w_i.

El objetivo es ajustar la cartera, es decir, determinar el número de participaciones en cada valor, de modo que se maximicen los dividendos.

Las incógnitas son x_i, el cambio en el número de participaciones que ya se tienen. Como en un principio se tenían b_i participaciones del valor bursátil i, después del ajuste se tendrán $b_i + x_i$

Cartera de valores

Datos

- m : no. de valores bursátiles
- b_i : no. actual de participaciones del valor bursátil i
- v_i : precio actual del valor i por participación
- d_i : dividendo que paga al final del año b_i
- w_i : nuevo precio del valor bursátil i
- r : porcentaje mínimo del valor actual de la cartera que no debe superarse en el ajuste
- s : porcentaje mínimo del valor actual de la cartera que no debe superarse por el valor futuro de la cartera

Variables

- x_i : cambio en el no. de participaciones del valor bursátil i.

Cartera de valores

1. Participaciones no negativas

$$x_i \geqslant -b_i$$

2. Independencia de un valor de la cartera. El capital asociado a todo valor, después del ajuste, representa al menos una cierta fracción r del capital total de la cartera.

$$r(\sum_{i=1}^{m} v_i(b_i + x_i) \leqslant v_j(b_j + x_j), j = 1, \dots, m$$

3. El capital total de cartera no cambia.

$$\sum_{i=1}^{m} v_i x_i = 0$$

4. Inflacción. El captital fúturo debe ser al menos un cierto porcentaje s mayor que el capital invertido actualmente.

$$\sum_{i=1}^{m} w_i(b_i + x_i) \geqslant (1 + s) \sum_{i=1}^{m} v_i b_i$$

Cartera de valores

Función objetivo

$$f(x) = \sum_{i=1}^{m} d_i(b_i + x_i) \Rightarrow \bar{f}(x) = \sum_{i=1}^{m} d_i x_i = d'x$$

Ejemplo:
Estúdiese el caso particular en que se tienen participaciones de tres valores bursátiles, 75 de A, 100 de B, y 35 de C, con precios 20, 20 y 100 euros, respectivamente. Además se dispone de la siguiente información: A no pagará dividendos y alcanzará una nueva cotización de 18 euros, B pagará 3 euros por participación y la nueva cotización será 23 euros, y C pagará 5 euros por participación con una nueva cotización de 102 euros. Además se toman $r = 0.25$ y $s = 0.30$.

Cartera de valores

Planteamiento

$$
\begin{aligned}
&\text{máx} \quad 0(75+x_A)+3(100+x_B)+5(35+x_C)\\
&s.a.\\
&\quad x_A \geqslant -75\\
&\quad x_B \geqslant -100\\
&\quad x_C \geqslant -35\\
&\quad 0.25[20(75+x_A)+20(100+x_B)+100(35+x_C)] \leqslant 20(75+x_A)\\
&\quad 0.25[20(75+x_A)+20(100+x_B)+100(35+x_C)] \leqslant 20(100+x_B)\\
&\quad 0.25[20(75+x_A)+20(100+x_B)+100(35+x_C)] \leqslant 100(35+x_C)\\
&\quad 20x_A+20x_B+100x_C=0\\
&\quad 18(75+x_A)+23(100+x_B)+102(35+x_C) \geqslant 1.03 * 7000
\end{aligned}
$$

Cartera de valores

Planteamiento

$$
\begin{aligned}
&\text{máx} \qquad 475 + 3x_B + 5x_C \\
&s.a. \\
&\quad x_A \geqslant -75 \\
&\quad x_B \geqslant -100 \\
&\quad x_C \geqslant -35 \\
&\quad -15x_A + 5x_B + 25x_C \leqslant -250 \\
&\quad 5x_A - 15x_B + 25x_C \leqslant 250 \\
&\quad 5x_A + 5x_B - 75x_C \leqslant 1750 \\
&\quad 20x_A + 20x_B + 100x_C = 0 \\
&\quad 18x_A + 23x_B + 102x_C \geqslant -10
\end{aligned}
$$

Solución

$$\boxed{x_A = 12.5, \qquad x_B = 75.0, \qquad x_C = -17.5 \qquad f(x) = 612.5}$$

Métodos de Mejora

- Son métodos iterativos. En cada iteración se dispone de una solución factible (óptimo provisional): $y^n \in S$
- El método dispone de un criterio de optimalidad, es decir una condición suficiente de optimalidad. En cada iteración se somete a y^n a dicho criterio. Si se satisface, y^n es solución óptima, y se finaliza; en caso contario se continúa.
- El método dispone de un procedimiento de mejora, esto es un procedimiento que permite calcular $y^{n+1} \in S$ tal que
$$f(y^{n+1}) \leqslant f(y^n)$$

Métodos de Mejora

Algoritmo de un Método de Mejora

PASO 1 **Inicialización**.

- Se elige $y^1 \in S$, si es posible en otro caso FIN (msg:Problema no factible)
- $n \leftarrow 1$

PASO 2 **Criterio de optimalidad**.

- Si se satisface el criterio de optimalidad y^n es solución óptima. FIN
- Si el problema no tiene solución óptima. FIN (msg: Problema sin solución)

PASO 3 **Mejora**

- Se elige $y^{n+1} \in S$ tal que $f(y^{n+1}) \leqslant f(y^n)$
- $n \leftarrow n + 1$
- Volver a PASO 2

Métodos de Mejora

Del algoritmo anterior, se deduce que para construir un método de mejora se precisa determinar:

1. Procedimiento para determinar si existe o no, una solución factible inicial.
2. Criterio de optimalidad.
3. Procedimiento de mejora.
4. Detector de problemas sin solución.

Método de direcciones de mejora Un método de direcciones de mejora es aquel en el cuál la iteración de mejora se realiza del modo: $y^{n+1} := y^n + \alpha d^n$

- d^n es un vector no nulo (dirección)
- $\alpha \geqslant 0$ indica el avance en la dirección d, su cálculo se basa en la resolución del problema unidimensional:

$$\begin{aligned} & \min f(y^n + \alpha d^n) \\ \text{s.a.} \quad & y^n + \alpha d^n \in S \\ & \alpha \geqslant 0 \end{aligned}$$

Tema 3

Programación Lineal

Introducción

Programación Lineal
- Formulación
- Interpretación geométrica

El método símplex

Análisis de Sensibilidad

Dualidad

Forma general de un problema de optimización lineal

$$\boxed{\begin{array}{ll} \min & c^T x \\ \text{s.a} & Ax \leqslant b \\ & [x \in \mathbb{R}^n] \end{array}}$$

- $c \in \mathbb{R}^n$ vector de costes.
- $A \in \mathbb{R}^{m \times n}$ matriz del sistema de desigualdades.
- $b \in \mathbb{R}^m$ vector de términos independientes.

donde,

$$c = (c_1, \ldots, c_n)$$

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

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Formas de un problema de optimización lineal

Forma General

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax \leqslant b \\ & [x \in \mathbb{R}^n] \end{array}$$

Forma Canónica

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & \bar{A}x \leqslant \bar{b} \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}$$

Forma Estándar

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & \tilde{A}x = \tilde{b} \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}$$

Nota: Observar que la forma estándar es un caso particular de la forma canónica, y que la forma canónica es un caso particular de la general (basta considerar la restricción $-\mathrm{Id}\, x \leqslant 0$, siendo Id la matriz identidad).

Paso a forma general a canónica y estándar

- Problemas de maximización, todo problema de maximización se puede tansformar en uno de minimización, teniendo en cuenta:

$$\arg\max c^T x = \arg\min -c^T x$$

- No negatividad de las variables, si alguna variable x_i no está condicionada a ser no negativa, se reemplaza por otras dos x_i', x_i'' tales que,

$$x_i = x_i' - x_i''$$

  donde $x_i' \geqslant 0$ y $x_i'' \geqslant 0$.

- Restricciones = Se introducen variables de holgura $s_i \geqslant 0$, de modo, que:

$$a_{i1}x_1+\ldots+a_{in}x_n \leqslant b_i \qquad \Rightarrow \qquad a_{i1}x_1+\ldots+a_{in}x_n+s_i = b_i$$

Ejemplo

Ejemplo Escribir el siguiente problema de optimización lineal en forma estándar:

$$\begin{aligned} \text{máx} \quad & x_1 - 4x_2 \\ s.a. \quad & -3x_1 + x_2 \leqslant 9 \\ & x_1 + 2x_2 \leqslant -1 \\ & x_2 \geqslant 0 \end{aligned}$$

Ejemplo

Ejemplo Escribir el siguiente problema de optimización lineal en forma estándar:

$$\begin{aligned} \text{máx} \quad & x_1 - 4x_2 \\ s.a. \quad & -3x_1 + x_2 \leqslant 9 \\ & x_1 + 2x_2 \leqslant -1 \\ & x_2 \geqslant 0 \end{aligned}$$

Forma general

$$\begin{aligned} \text{mín} \quad & -x_1 + 4x_2 \\ s.a. \quad & -3x_1 + x_2 \leqslant 9 \\ & x_1 + 2x_2 \leqslant -1 \\ & x_2 \geqslant 0 \end{aligned}$$

Ejemplo

Ejemplo Escribir el siguiente problema de optimización lineal en forma estándar:

$$\begin{array}{ll} \text{máx} & x_1 - 4x_2 \\ s.a. & -3x_1 + x_2 \leqslant 9 \\ & x_1 + 2x_2 \leqslant -1 \\ & x_2 \geqslant 0 \end{array}$$

Forma general

$$\begin{array}{ll} \text{mín} & -x_1 + 4x_2 \\ s.a. & -3x_1 + x_2 \leqslant 9 \\ & x_1 + 2x_2 \leqslant -1 \\ & x_2 \geqslant 0 \end{array}$$

Ejemplo

Ejemplo Escribir el siguiente problema de optimización lineal en forma estándar:

Forma general

$$\begin{aligned} \text{mín} \quad & -x_1 + 4x_2 \\ s.a. \quad & -3x_1 + x_2 \leqslant 9 \\ & x_1 + 2x_2 \leqslant -1 \\ & x_2 \geqslant 0 \end{aligned}$$

Forma canónica

$$\begin{aligned} \text{mín} \quad & -x_1 + 4x_2 \\ s.a. \quad & -3(x_1' - x_1'') + x_2 \leqslant 9 \\ & (x_1' - x_1'') + 2x_2 \leqslant -1 \\ & x_1', x_1'', x_2 \geqslant 0 \end{aligned}$$

Ejemplo

Ejemplo Escribir el siguiente problema de optimización lineal en forma estándar:

Forma general

$$\begin{aligned} \text{mín} \quad & -x_1 + 4x_2 \\ s.a. \quad & -3x_1 + x_2 \leqslant 9 \\ & x_1 + 2x_2 \leqslant -1 \\ & x_2 \geqslant 0 \end{aligned}$$

Forma canónica

$$\begin{aligned} \text{mín} \quad & -x_1 + 4x_2 \\ s.a. \quad & -3(x_1' - x_1'') + x_2 \leqslant 9 \\ & (x_1' - x_1'') + 2x_2 \leqslant -1 \\ & x_1', x_1'', x_2 \geqslant 0 \end{aligned}$$

Ejemplo

Ejemplo Escribir el siguiente problema de optimización lineal en forma estándar:

Forma canónica

$$
\begin{aligned}
\text{mín} \quad & -x_1 + 4x_2 \\
s.a. \quad & -3(x_1' - x_1'') + x_2 \leqslant 9 \\
& (x_1' - x_1'') + 2x_2 \leqslant -1 \\
& x_1', x_1'', x_2 \geqslant 0
\end{aligned}
$$

Forma estándar

$$
\begin{aligned}
\text{mín} \quad & -x_1 + 4x_2 \\
s.a. \quad & -3x_1' + 3x_1'' + x_2 + s_1 = 9 \\
& x_1' - x_1'' + 2x_2 + s_2 = -1 \\
& x_1', x_1'', x_2, s_1, s_2 \geqslant 0
\end{aligned}
$$

Ejemplo

Ejercicio Escribir el siguiente problema de optimización lineal en forma estándar:

$$\begin{array}{rl} \text{máx} & 2x_1 + 5x_2 \\ s.a. & x_1 + x_2 \geqslant 6 \\ & -x_1 - 2x_2 \leqslant -18 \end{array}$$

Interpretación geométrica

Consideremos el siguiente problema:

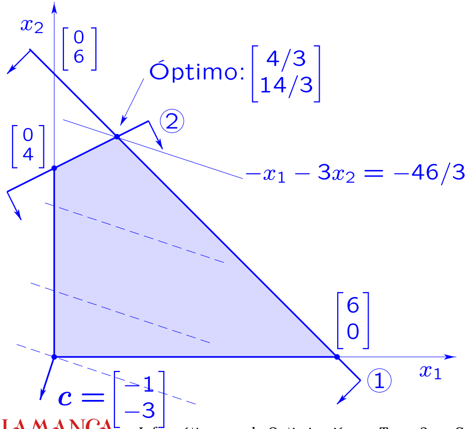
$$\begin{array}{ll} \text{mín} & -x_1 - 3x_2 \\ s.a. & x_1 + x_2 \leqslant 6 \\ & x_1 - 2x_2 \geqslant -8 \\ & x_1, x_2 \geqslant 0 \end{array}$$

- La región factible esta determinada por la intersección de 4 semiplanos.
- El óptimo será el punto de la región factible que minimiza $z = -x_1 - 3x_2$ (una recta).
- Minimizar z es equivalente a desplazar el subespacio de referencia, $-x_1 - 3x_2 = 0$ en la dirección que minimiza más dicho objetivo, esto es $-\nabla z = -c = (1, 3)^T$
- El punto óptimo será aquel el que no sea posible mover más $z = -x_1 - 3x_2$ en la dirección $-c$.

Interpretación geométrica

Consideremos el siguiente problema:

$$\begin{array}{ll} \text{mín} & -x_1 - 3x_2 \\ s.a. & x_1 + x_2 \leqslant 6 \\ & x_1 - 2x_2 \geqslant -8 \\ & x_1, x_2 \geqslant 0 \end{array}$$



Ejercicios

Utilizar el método gráfico para resolver los siguientes problemas:

1. Se dispone de 210.000 euros para invertir en bolsa. Se recomiendan dos tipos de acciones. Las del tipo A, ofrecen un rendimiento del 10 % y las del tipo B, uno del 8 %. Se decide invertir un máximo de 130.000 euros en las del tipo A y como mínimo 60.000 en las del tipo B. Además se desea que la inversión en las del tipo A sea menor que el doble de la inversión en B. ¿Cuál debe ser la distribución de la inversión para obtener el máximo interés anual?
2. Para recorrer un determinado trayecto, una compañía aérea desea ofertar, a lo sumo, 5000 plazas de dos tipos: T(turista) y P(primera). La ganancia correspondiente a cada plaza de tipo T es de 30 euros, mientras que la ganancia del tipo P es de 40 euros. El número de plazas tipo T no puede exceder de 4500 y el del tipo P, debe ser, como máximo, la tercera parte de las del tipo T que se oferten. Calcular cuántas tienen que ofertarse de cada clase para que las ganancias sean máximas.

Tipos de soluciones en un problema de optimización lineal

Tipos de soluciones

- Solución óptima única.
- Múltiples soluciones óptimas.
- Problema no acotado.
- Problema no factible ($\mathcal{O} = \emptyset$).

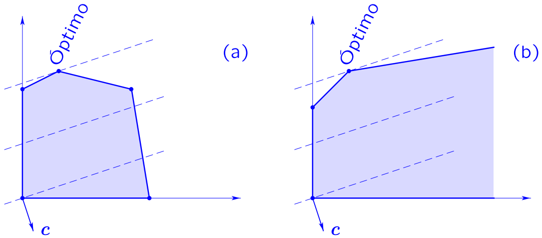
Observaciones para $d > 2$

- La región factible de un problema de optimización lineal es un politopo convexo (intersección de un número finito de semiespacios cerrados).
- Dado un politopo no vacio, el valor mínimo de $c^T x$, si existe, se alcanza en un vértice del politopo.

Tipos de soluciones en un problema de optimización lineal

Tipos de soluciones

- Solución óptima única.



(a) Región factible acotada.
(b) Región factible no acotada.

- Múltiples soluciones óptimas.
- Problema no acotado.
- Problema no factible ($\mathcal{O} = \emptyset$).

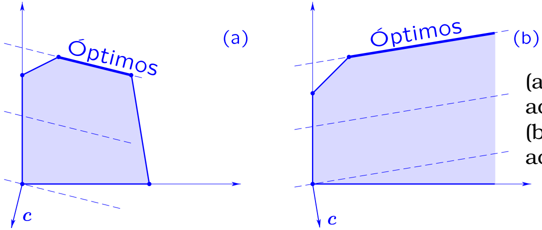
Observaciones para $d > 2$

- La región factible de un problema de optimización lineal es un politopo convexo (intersección de un número finito de semiespacios cerrados).

Tipos de soluciones en un problema de optimización lineal

Tipos de soluciones

- Solución óptima única.
- Múltiples soluciones óptimas.



(a) Región factible acotada.
(b) Región factible no acotada.

- Problema no acotado.
- Problema no factible ($\mathcal{O} = \emptyset$).

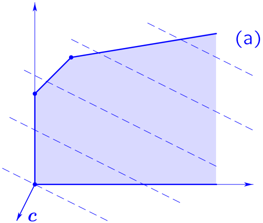
Observaciones para $d > 2$

- La región factible de un problema de optimización lineal es un politopo convexo (intersección de un número finito de semiespacios cerrados).

Tipos de soluciones en un problema de optimización lineal

Tipos de soluciones

- Solución óptima única.
- Múltiples soluciones óptimas.
- Problema no acotado.



- Problema no factible ($\mathcal{O} = \emptyset$).

Observaciones para $d > 2$

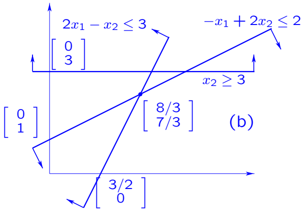
- La región factible de un problema de optimización lineal es un politopo convexo (intersección de un número finito de semiespacios cerrados).

Tipos de soluciones en un problema de optimización lineal

Tipos de soluciones

- Solución óptima única.
- Múltiples soluciones óptimas.
- Problema no acotado.
- Problema no factible ($\mathcal{O} = \emptyset$).
  Por ejemplo:

$$
\begin{array}{ll}
\text{mín} & -2x_1 + 3x_2 \\
s.a. & -x_1 + 2x_2 \leqslant 2 \\
 & 2x_1 - x_2 \leqslant 3 \\
 & x_1 \geqslant 0, \\
 & x_2 \geqslant 3
\end{array}
$$



Observaciones para $d > 2$

- La región factible de un problema de optimización lineal es un politopo convexo (intersección de un

Tipos de soluciones en un problema de optimización lineal

Tipos de soluciones

- Solución óptima única.
- Múltiples soluciones óptimas.
- Problema no acotado.
- Problema no factible ($\mathcal{O} = \emptyset$).

Observaciones para $d > 2$

- La región factible de un problema de optimización lineal es un politopo convexo (intersección de un número finito de semiespacios cerrados).
- Dado un politopo no vacio, el valor mínimo de $c^T x$, si existe, se alcanza en un vértice del politopo.

Tema 3

Programación Lineal

Introducción

Programación Lineal

El método símplex
- Algoritmo
- Inicio del símplex
- Convergencia y Complejidad

Análisis de Sensibilidad

Dualidad

Preliminares

En lo que sigue consideraremos un problema de optimización lineal en forma estándar:

$$\boxed{\begin{array}{ll} \min & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}}$$

donde,

- $c \in \mathbb{R}^n$
- $A \in \mathbb{R}^{m \times n}$
- $b \in \mathbb{R}^m$.

Además se asumen las siguientes hipótesis,

1. $\operatorname{rang}(A) = m$
2. $m < n$

Estas hipótesis son equivalentes a:

1. No existen ecuaciones redundantes.
2. El sistema es compatible indeterminado.

Definiciones y notación

Base: se denomina base (de A), a una base de $\mathbb{R}^m$ formada por columnas de A.

Dada una base B de A está induce una partición en c, A y x y permiten introducir la siguiente notación:

B : Columnas de A que forman la base

N : Columnas de A que estan fuera de la base

I : Índices de las columnas de A que forman la base

J :Índices de las columnas de B que están fuera de la base

c^B : Coeficientes de c asociados a la base

c^N : Coeficientes de c asociados a la no base

x^B : Coeficientes de x asociadas a la base

x^N : Coeficientes de x asociados a la no base

Ejemplo

$$
\begin{array}{llrrrrcr}
\text{mín} & \multicolumn{7}{l}{x_1 + x_2 + 13x_3 + 10x_4} \\
\text{s.a.} & & x_1 & +x_2 & +x_3 & +x_4 & = & 5 \\
& & -x_1 & +x_2 & & & = & 2 \\
& & & -x_2 & +x_3 & -x_4 & = & -5 \\
& \multicolumn{7}{l}{x \geqslant 0} \\
& \multicolumn{7}{l}{[x \in \mathbb{R}^5]}
\end{array}
$$

Ejemplo

$$
\begin{array}{llllllll}
\text{mín} & \multicolumn{7}{l}{x_1 + x_2 + 13x_3 + 10x_4} \\
\text{s.a.} & x_1 & +x_2 & +x_3 & +x_4 & & = & 5 \\
 & -x_1 & +x_2 & & & & = & 2 \\
 & & -x_2 & +x_3 & -x_4 & & = & -5 \\
 & \multicolumn{7}{l}{x \geqslant 0} \\
 & \multicolumn{7}{l}{[x \in \mathbb{R}^5]}
\end{array}
$$

Datos

$$
c = (1, 1, 13, 10, 0),\ A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & 0 \end{pmatrix}\ b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}
$$

Ejemplo

Datos

$$c = (1, 1, 13, 10, 0),\ A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & 0 \end{pmatrix}\ b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

Dada la base $I = (1, 4, 5)$ y la no base $J = (2, 3)$, se tiene:

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

$$c^B = (1, 10, 0) \qquad c^N = (1, 13)$$

$$x^B = \begin{pmatrix} x_1 \\ x_4 \\ x_5 \end{pmatrix} \qquad x^N = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}$$

Ejemplo

Datos

$$c = (1, 1, 13, 10, 0),\ A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & 0 \end{pmatrix} \ b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

Dada la base $I = (1, 4, 5)$ y la no base $J = (2, 3)$, se tiene:

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

$$c^B = (1, 10, 0) \qquad c^N = (1, 13)$$

$$x^B = \begin{pmatrix} x_1 \\ x_4 \\ x_5 \end{pmatrix} \qquad x^N = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}$$

Ejemplo

Datos

$$c = (1, 1, 13, 10, 0),\ A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & 0 \end{pmatrix}\ b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

Dada la base $I = (1, 4, 5)$ y la no base $J = (2, 3)$, se tiene:

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

$$c^B = (1, 10, 0) \qquad c^N = (1, 13)$$

$$x^B = \begin{pmatrix} x_1 \\ x_4 \\ x_5 \end{pmatrix} \qquad x^N = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}$$

Ejemplo

Datos

$$c = (1, 1, 13, 10, 0), \quad A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & 0 \end{pmatrix} \quad b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

Dada la base $I = (1, 4, 5)$ y la no base $J = (2, 3)$, se tiene:

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

$$c^B = (1, 10, 0) \qquad c^N = (1, 13)$$

$$x^B = \begin{pmatrix} x_1 \\ x_4 \\ x_5 \end{pmatrix} \qquad x^N = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}$$

Ejemplo

Datos

$$c = (1, 1, 13, 10, 0), \; A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & 0 \end{pmatrix} \; b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

Dada la base $I = (1, 4, 5)$ y la no base $J = (2, 3)$, se tiene:

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

$$c^B = (1, 10, 0) \qquad c^N = (1, 13)$$

$$x^B = \begin{pmatrix} x_1 \\ x_4 \\ x_5 \end{pmatrix} \qquad x^N = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}$$

Reformulación del problema en términos de una base

Dada una base $\boxed{B}$ se puede deducir:

$$\left.\begin{array}{l} Ax = Bx^B + Nx^N \\ Ax = b \end{array}\right\} \Rightarrow \quad x^B = B^{-1}b - B^{-1}Nx^N$$

$$cx = c^B x^B + c^N X^n = c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N$$

$$\boxed{\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}} \Rightarrow \boxed{\begin{array}{ll} \text{mín} & c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N \\ \text{s.a} & x^B = B^{-1}b - B^{-1}Nx^N \\ & B^{-1}b - B^{-1}Nx^N \geqslant 0 \\ & x^N \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}}$$

Reformulación del problema en términos de una base

Dada una base B se puede deducir:

$$\begin{aligned} Ax &= Bx^B + Nx^N \\ Ax &= b \end{aligned} \quad \Rightarrow \quad x^B = B^{-1}b - B^{-1}Nx^N$$

$$cx = c^B x^B + c^N X^n = c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N$$

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array} \quad \Rightarrow \quad \begin{array}{ll} \text{mín} & c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N \\ \text{s.a} & x^B = B^{-1}b - B^{-1}Nx^N \\ & B^{-1}b - B^{-1}Nx^N \geqslant 0 \\ & x^N \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}$$

Reformulación del problema en términos de una base

Dada una base B se puede deducir:

$$\left.\begin{aligned} Ax &= Bx^B + Nx^N \\ Ax &= b \end{aligned}\right\} \Rightarrow \quad x^B = B^{-1}b - B^{-1}Nx^N$$

$$cx = c^B x^B + c^N X^n = c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N$$

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array} \quad \Rightarrow \quad \begin{array}{ll} \text{mín} & c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N \\ \text{s.a} & x^B = B^{-1} b - B^{-1} N x^N \\ & B^{-1} b - B^{-1} N x^N \geqslant 0 \\ & x^N \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}$$

Reformulación del problema en términos de una base

Dada una base B se puede deducir:

$$\left.\begin{aligned} Ax &= Bx^B + Nx^N \\ Ax &= b \end{aligned}\right. \quad \Rightarrow \quad x^B = B^{-1}b - B^{-1}Nx^N$$

$$cx = c^B x^B + c^N X^n = c^B B^{-1} b + (c^N - c^B B^{-1} N)x^N$$

$$\begin{aligned} \min \quad & c^T x \\ \text{s.a} \quad & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{aligned} \quad \Rightarrow \quad \begin{aligned} \min \quad & c^B B^{-1} b + (c^N - c^B B^{-1} N)x^N \\ \text{s.a} \quad & x^B = B^{-1}b - B^{-1}Nx^N \\ & B^{-1}b - B^{-1}Nx^N \geqslant 0 \\ & x^N \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{aligned}$$

Reformulación del problema en términos de una base

Dada una base B se puede deducir:

$$\left.\begin{array}{l} Ax = Bx^B + Nx^N \\ Ax = b \end{array}\right. \quad \Rightarrow \quad x^B = B^{-1}b - B^{-1}Nx^N$$

$$cx = c^B x^B + c^N X^n = c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N$$

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array} \quad \Rightarrow \quad \begin{array}{ll} \text{mín} & c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N \\ \text{s.a} & x^B = B^{-1}b - B^{-1}Nx^N \\ & B^{-1}b - B^{-1}Nx^N \geqslant 0 \\ & x^N \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}$$

Reformulación del problema en términos de una base

Dada una base $\boxed{B}$ se puede deducir:

$$\left.\begin{array}{l} Ax = Bx^B + Nx^N \\ Ax = b \end{array}\right. \quad \Rightarrow \quad x^B = B^{-1}b - B^{-1}Nx^N$$

$$cx = c^B x^B + c^N X^n = c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N$$

$$\boxed{\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}} \quad \Rightarrow \quad \boxed{\begin{array}{ll} \text{mín} & c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N \\ \text{s.a} & x^B = B^{-1} b - B^{-1} N x^N \\ & B^{-1} b - B^{-1} N x^N \geqslant 0 \\ & x^N \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}}$$

Reformulación del problema en términos de una base

Dada una base $\boxed{B}$ se puede deducir:

$$\left.\begin{aligned} Ax &= Bx^B + Nx^N \\ Ax &= b \end{aligned}\right. \quad \Rightarrow \quad x^B = B^{-1}b - B^{-1}Nx^N$$

$$cx = c^Bx^B + c^NX^n = c^BB^{-1}b + (c^N - c^BB^{-1}N)x^N$$

$$\boxed{\begin{array}{ll} \text{mín} & c^Tx \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}} \quad \Rightarrow \quad \boxed{\begin{array}{ll} \text{mín} & c^BB^{-1}b + (c^N - c^BB^{-1}N)x^N \\ \text{s.a} & x^B = B^{-1}b - B^{-1}Nx^N \\ & B^{-1}b - B^{-1}Nx^N \geqslant 0 \\ & x^N \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}}$$

Reformulación del problema en términos de una base

Dada una base B se puede deducir:

$$\left.\begin{array}{l} Ax = Bx^B + Nx^N \\ Ax = b \end{array}\right. \quad \Rightarrow \quad x^B = B^{-1}b - B^{-1}Nx^N$$

$$cx = c^B x^B + c^N X^n = c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N$$

$$\boxed{\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}} \quad \Rightarrow \quad \boxed{\begin{array}{ll} \text{mín} & c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N \\ \text{s.a} & x^B = B^{-1}b - B^{-1}Nx^N \\ & B^{-1}b - B^{-1}Nx^N \geqslant 0 \\ & x^N \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}}$$

Reformulación del problema en términos de una base

Dada una base B se puede deducir:

$$\begin{aligned} Ax &= Bx^B + Nx^N \\ Ax &= b \end{aligned} \quad \Rightarrow \quad x^B = B^{-1}b - B^{-1}Nx^N$$

$$cx = c^B x^B + c^N X^n = c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N$$

$$\boxed{\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}} \quad \Rightarrow \quad \boxed{\begin{array}{ll} \text{mín} & c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N \\ \text{s.a} & x^B = B^{-1}b - B^{-1}Nx^N \\ & B^{-1}b - B^{-1}Nx^N \geqslant 0 \\ & x^N \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}}$$

Solución básica

$$\begin{array}{ll} \text{mín} & c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N \\ \text{s.a} & x^B = B^{-1} b - B^{-1} N x^N \\ & B^{-1} b - B^{-1} N x^N \geqslant 0 \\ & x^N \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}$$

Definición: (Solución básica). Una solución básica respecto a la base B es un vector $\mathbf{x} \in \mathbb{R}^n$ tal que:

$$(1)\ A\mathbf{x} = b \qquad (2)\ \mathbf{x}^N = 0$$

Nota: La defincián anterior es equivalente a:

$$(1')\ \mathbf{x}^B = B^{-1} b \qquad (2)\ \mathbf{x}^N = 0$$

Definición: (Solución factible básica). Es aquella que es factible y a la vez básica, esto es equivalente a:

$$(1')\ \mathbf{x}^B = B^{-1} b \geqslant 0 \qquad (2)\ \mathbf{x}^N = 0$$

Ejemplo

$$
\begin{array}{lrrrrcr}
\text{mín} & \multicolumn{6}{l}{x_1 + x_2 + 13x_3 + 10x_4} \\
\text{s.a.} & x_1 & +x_2 & +x_3 & +x_4 & = & 5 \\
 & -x_1 & +x_2 & & & = & 2 \\
 & & -x_2 & +x_3 & -x_4 & = & -5 \\
 & \multicolumn{6}{l}{x \geqslant 0} \\
 & \multicolumn{6}{l}{[x \in \mathbb{R}^5]}
\end{array}
$$

La solución básica asociada a la base $I = (1, 4, 5)$ es:

$$
x^B = B^{-1} b = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ 2 \end{pmatrix} \quad x^N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
$$

Dado que $x^B > 0$ la solución también es factible básica.

Características del método símplex

- Es un método de direcciones de mejora
- En cada iteración se pasa de una solución factible básica a otra solución factible básica, hasta alcanzar el óptimo.
- El criterio de optimalidad del símplex se basa en el análisis de un número fínito de direcciones candidatas, si ninguna de ellas es de mejora, se ha alcanzado el óptimo.
- Geométricamente, el conjunto factible es un politopo convexo. Las soluciones factibles básicas son vertices del politopo. Las direcciones de mejora, son las aristas del politopo. En cada iteración, se pasa de un vértice a otro hasta alcanzar el óptimo.

Símplex como métod de mejora

Para desarrollar el método símplex como un método de mejora se requiere:

- Inicialización: Cálculo de la primera solución factible incial, si existe ⇒ FASE I , (siguiente sección)
- Criterio de Optimalidad ⇒ Teorema de optimalidad
- Cálculo de la dirección de mejora ⇒ Teorema de direcciones de mejora
- Cálculo del avance en la dirección de mejora ⇒ Resolución de un problema unidimensional lineal (Detectará además problemas no acotados)

Teorema de optimalidad del Símplex

Teorema: Considerar un problema de programación lineal en forma estándar. Sean B una base, N una no base, y $\mathbf{x}$ la corresondiente solución factible básica.
Si

$$\{j \,:\, (c^N - c^B B^{-1} N)_j < 0\} = \emptyset$$

Entonces $\mathbf{x}$ es solución óptima

Demostración:
Veamos que $c^T x \geqslant c^T \mathbf{x}, \quad \forall x \in S$.
Usando la base B:

$$c^T x = c^B B^{-1} b + (c^N - c^B B^{-1} N) x^N \geqslant c^B B^{-1} b = c^T \mathbf{x}$$

donde se ha usado $(c^N - c^B B^{-1} N) \geqslant 0$ y $x^N \geqslant 0$.

Teorema de direcciones de Mejora del Símplex

Teorema Considerar un problema de programación lineal en forma estándar. Sean B una base, N una no base, y $\mathbf{x}$ la corresondiente solución factible básica.
Sea

$$k \in \{j \;:\; (c^N - c^B B^{-1} N)_j < 0\}$$

y sea $d \in \mathbb{R}^n$ una dirección definida:

$$d^B = -B^{-1} N_k, \qquad d^N = \mathbf{e}_k = (0, \ldots, \overbrace{1}^{k}, \ldots, 0)^T$$

Entonces:

1. d es una dirección de mejora en $\mathbf{x}$
2. $(\mathbf{x} + \alpha d)$ satisface $Ax = b$, $\forall \alpha \geqslant 0$.

Teorema de direcciones de Mejora del Símplex

Demostración:

1. Veamos que $c^T(\mathbf{x} + \alpha d) < c^T\mathbf{x}$. Usando la expresión para $c^T x$ en términos e la base B, y la definición de d:

$$\begin{aligned}
c^T(\mathbf{x} + \alpha d) &= c^B B^{-1} b + (c^N - c^B B^{-1} N)(\mathbf{x} + \alpha d)^N \\
&= c^B B^{-1} b + \alpha (c^N - c^B B^{-1} N) d^N \\
&= c^B B^{-1} b + \alpha (c^N - c^B B^{-1} N) \mathbf{e}_k \\
&= c^B B^{-1} b + \alpha \underbrace{(c^N - c^B B^{-1} N)_k}_{<0} < c^B B^{-1} b = c^T \mathbf{x}
\end{aligned}$$

2. Usando la defінción de d y que $\mathbf{x}$ es solución factible básica:

$$\begin{aligned}
A(\mathbf{x} + \alpha d) = A\mathbf{x} + \alpha A d &= B\mathbf{x}^B + N\mathbf{x}^N + \alpha(Bd^B + Nd^N) \\
&= b + \alpha(-BB^{-1}N_k + N\mathbf{e}_k) \\
&= b + \alpha(-N_k + N_k) = b
\end{aligned}$$

Teorema de direcciones de Mejora del Símplex

Demostración:

1. Veamos que $c^T(\mathbf{x}+\alpha d) < c^T\mathbf{x}$. Usando la expresión para $c^T x$ en términos e la base B, y la definición de d:

$$\begin{aligned}
c^T(\mathbf{x}+\alpha d) &= c^B B^{-1} b + (c^N - c^B B^{-1} N)(\mathbf{x}+\alpha d)^N \\
&= c^B B^{-1} b + \alpha (c^N - c^B B^{-1} N) d^N \\
&= c^B B^{-1} b + \alpha (c^N - c^B B^{-1} N) \mathbf{e}_k \\
&= c^B B^{-1} b + \alpha \underbrace{(c^N - c^B B^{-1} N)_k}_{<0} < c^B B^{-1} b = c^T \mathbf{x}
\end{aligned}$$

2. Usando la definción de d y que $\mathbf{x}$ es solución factible básica:

$$\begin{aligned}
A(\mathbf{x}+\alpha d) &= A\mathbf{x} + \alpha A d = B\mathbf{x}^B + N\mathbf{x}^N + \alpha(Bd^B + Nd^N) \\
&= b + \alpha(-BB^{-1}N_k + N\mathbf{e}_k) \\
&= b + \alpha(-N_k + N_k) = b
\end{aligned}$$

Avance en la dirección de mejora

El problema unidimensional asociado se puede simplicar del siguiente modo:

$$\begin{array}{ll} \min & c^T(\mathbf{x}+\alpha d) \\ \text{s.a} & A(\mathbf{x}+\alpha d) = b \\ & (\mathbf{x}+\alpha d) \geqslant 0 \\ & \alpha \geqslant 0 \end{array}$$

$$\begin{array}{ll} \max & \alpha \\ \text{s.a} & \mathbf{x}^B + \alpha d^B \geqslant 0 \\ & \alpha \geqslant 0 \end{array}$$

$$\begin{array}{ll} \max & \alpha \\ \text{s.a} & 0 \leqslant \alpha \leqslant \Delta \end{array}$$

$$\Delta := \min\{\frac{\mathbf{x}_i^B}{-d_i^B} \ : \ d_i^B < 0\}$$

Nota: Si el conjunto $\{\frac{\mathbf{x}_i^B}{-d_i^B} \ : \ d_i^B < 0\} = \emptyset$ entonces $\Delta := \infty$.

Se distinguen dos casos:

$\Delta < \infty$ La solución óptima del problema unidimensional es $\alpha = \Delta$

$\Delta = \infty$ El problema unidimensional no tiene solución óptima, y por tanto el problema original tampco. Se dice que es un Problema no acotado.

Avance en la dirección de mejora

El problema unidimensional asociado se puede simplicar del siguiente modo:

$$\begin{aligned} \text{mín} \;\; & c^T(\mathbf{x}+\alpha d) \\ \text{s.a} \;\; & A(\mathbf{x}+\alpha d) = b \\ & (\mathbf{x}+\alpha d) \geqslant 0 \\ & \alpha \geqslant 0 \end{aligned}$$

$$\begin{aligned} \text{máx} \;\; & \alpha \\ \text{s.a} \;\; & \mathbf{x}^B + \alpha d^B \geqslant 0 \\ & \alpha \geqslant 0 \end{aligned}$$

$$\begin{aligned} & \text{máx} \;\; \alpha \\ & \text{s.a} \;\; 0 \leqslant \alpha \leqslant \Delta \\ & \Delta := \text{mín}\{\frac{\mathbf{x}_i^B}{-d_i^B} \;:\; d_i^B < 0\} \end{aligned}$$

Nota: Si el conjunto $\{\frac{\mathbf{x}_i^B}{-d_i^B} \;:\; d_i^B < 0\} = \emptyset$ entonces $\Delta := \infty$.

Se distinguen dos casos:

$\Delta < \infty$ La solución óptima del problema unidimensional es $\alpha = \Delta$

$\Delta = \infty$ El problema unidimensional no tiene solución óptima, y por tanto el problema original tampco. Se dice que es un Problema no acotado.

Carácter básico de la nueva solución básica

- Se puede demostar que la nueva solucón óbtenida con la iteración:

$$x = \mathbf{x} + \Delta d$$

  es también básica.

- La base asociado a la nueva solucón resulta de intercambiar columnas entre las matrices B y N. En concreto, sean:

  $k =$ Índice seleccionado en la selección de dirección de mejora

  $l = \text{argmin}\{\frac{\mathbf{x}_i^B}{-d_i^B} \ : \ d_i^B < 0\}$

  Entonces la columna k deja de ser básica, y la columna l pasa a ser básica.

- Si $\Delta = 0$, la solución no cambia, pero si la base asociada.

Algoritmo del Símplex

Datos : n, m, c, A, b

Paso 1 : (Inicialización) Elegir base B, no base N, solución factible básica $\mathbf{x}$, e iniciar vectores de índices I, J)

Paso 2 : (Criterio de optimalidad)]

Si $\{j : c^N - c^B B^{-1} N)_j < 0\} = \emptyset \Rightarrow \mathbf{x}$ es óptimo, FIN

Paso 3 : (Mejora)

Elegir $k \in \{j : c^N - c^B B^{-1} N)_j < 0\}$

$d^B \leftarrow -B^{-1} N_k$

$\Delta \leftarrow \min\{-\frac{x_i^B}{d_i^B} : d_i^B < 0\} \qquad l = \operatorname{argmin}\{-\frac{x_i^B}{d_i^B} : d_i^B < 0\}$

Si $\Delta = \infty$: Problema no acotado; FIN

Si $\Delta < \infty$: troncar J_k, I_l;

Calcular nueva solución básica;

Volver al Paso 2

Ejemplo

Dado el problema:

$$
\begin{array}{lrrrrcr}
\text{mín} & \multicolumn{6}{l}{x_1 + x_2 + 13x_3 + 10x_4} \\
\text{s.a.} & x_1 & +x_2 & +x_3 & +x_4 & = & 5 \\
 & -x_1 & +x_2 & & & = & 2 \\
 & & -x_2 & +x_3 & -x_4 & = & -5 \\
 & \multicolumn{6}{l}{x \geqslant 0} \\
 & \multicolumn{6}{l}{[x \in \mathbb{R}^5]}
\end{array}
$$

Realizar una iteración completa del algoritmo del símplex, siendo $I = (1, 4, 5)$ y $J = (2, 3)$

Ejemplo

Dado el problema:

$$
\begin{array}{lrrrrcr}
\text{mín} & \multicolumn{6}{l}{x_1 + x_2 + 13x_3 + 10x_4} \\
\text{s.a.} & x_1 & +x_2 & +x_3 & +x_4 & = & 5 \\
 & -x_1 & +x_2 & & & = & 2 \\
 & & -x_2 & +x_3 & -x_4 & = & -5 \\
 & \multicolumn{6}{l}{x \geqslant 0} \\
 & \multicolumn{6}{l}{[x \in \mathbb{R}^5]}
\end{array}
$$

Realizar una iteración completa del algoritmo del símplex, siendo $I = (1, 4, 5)$ y $J = (2, 3)$

Datos

$$
c = (1, 1, 13, 10, 0),\ A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & 0 \end{pmatrix}\ b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}
$$

Ejemplo

Datos

$$c = (1, 1, 13, 10, 0), \; A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & 0 \end{pmatrix} \; b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

Partición según base

$$I = (1, 4, 5) \quad J = (2, 3) \quad c^B = (1, 10, 0) \quad c^N = (1, 13)$$

$$B = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad N = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

Ejemplo

Variables

$$I = (1, 4, 5) \quad J = (2, 3) \quad c^B = (1, 10, 0) \quad c^N = (1, 13)$$

$$B = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad N = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

Inicialización

$$x^B = B^{-1} b = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ 2 \end{pmatrix} \quad x^N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Ejemplo

Variables

$$I = (1,4,5) \quad J = (2,3) \quad c^B = (1,10,0) \quad c^N = (1,13)$$

$$B = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad N = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

$$x^B = (0,5,2)^T, \qquad x^N = (0,0)^T$$

Criterio de optimalidad

$$(c^N - c^B B^{-1} N) = (1,13) - (1,10,0) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}$$

$$= (-9,3) \Rightarrow \text{No es óptimo } k = 1$$

Ejemplo

Variables

$$I=(1,4,5)\quad J=(2,3)\quad c^B=(1,10,0)\quad c^N=(1,13)$$

$$B=\begin{pmatrix} 1 & 1 & 0\\ -1 & 0 & 1\\ 0 & -1 & 0 \end{pmatrix}\quad N=\begin{pmatrix} 1 & 1\\ 1 & 0\\ -1 & 1 \end{pmatrix}\quad b=\begin{pmatrix} 5\\ 2\\ -5 \end{pmatrix}$$

$$x^B=(0,5,2)^T,\qquad x^N=(0,0)^T,\qquad k=1$$

Mejora

$$d^B=-B^{-1}N_1=-\begin{pmatrix} 1 & 0 & 1\\ 0 & 0 & -1\\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1\\ 1\\ -1 \end{pmatrix}=\begin{pmatrix} 0\\ -1\\ -1 \end{pmatrix}\quad d^N=\begin{pmatrix} 1\\ 0 \end{pmatrix}$$

Ejemplo

Variables

$$I = (1, 4, 5) \quad J = (2, 3) \quad c^B = (1, 10, 0) \quad c^N = (1, 13)$$

$$B = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad N = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

$$x^B = (0, 5, 2)^T, \quad x^N = (0, 0)^T, \quad k = 1, \quad d^B = (0, -1, -1)^T$$

Mejora

$$\Delta = \min \left\{ -\frac{\mathbf{x}_i^B}{d_i^B} \ : \ d_i^B < 0 \right\} = \min \left\{ \frac{5}{1}, \frac{2}{1} \right\} = 2, \Rightarrow \quad l = 3$$

Ejemplo

Variables

$$I = (1,4,5) \quad J = (2,3) \quad c^B = (1,10,0) \quad c^N = (1,13)$$

$$B = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad N = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

$x^B = (0,5,2)^T,\ x^N = (0,0)^T, k = 1,\ d^B = (0,-1,-1)^T,\ \Delta = 2,\ l = 3$

Mejora $\quad k = 1,\ l = 3$

ANTES	DESPUÉS
$I = (1,4,5)$	$I = (1,4,2)$
$J = (2,3)$	$J = (5,3)$

Ejemplo

Variables actualizadas

$$I = (1,4,2) \quad J = (5,3) \quad c^B = (1, 10, 1) \quad c^N = (0, 13)$$

$$B = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & -1 \end{pmatrix} \quad N = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix}$$

Mejora

$$x^B = B^{-1}b = \begin{pmatrix} 1 & 0 & 1 \\ -1 & -1 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} \quad x^N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\mathbf{x} = (0, 2, 0, 3, 0)^T$$

Ejercicio

Dado el problema:

$$
\begin{array}{lrrrrrcl}
\text{mín} & -45x_1 + 30x_2 & & & & & & \\
\text{s.a.} & 2x_1 & +10x_2 & +x_3 & & & = & 60 \\
 & 6x_1 & +6x_2 & & +x_4 & & = & 60 \\
 & 10x_1 & +5x_2 & & & +x_5 & = & 85 \\
 & x \geqslant 0 & & & & & & \\
 & [x \in \mathbb{R}^5] & & & & & &
\end{array}
$$

Realizar una iteración completa del algoritmo del símplex, siendo $I = (3, 4, 1)$ y $J = (5, 2)$

Inicio del Símplex: Fase I

Para seleccionar la primera base existen diferentes métodos. Entre ellos esta el denominado método de dos fases :

- Fase I: Calculo de la base inicial.
- Fase II: Algoritmo del símplex.

Para el cáculo de la base inicial se plantea un problema auxiliar introduciendo una serie de variables artificiales (tantas como restricciones).
Se asume que $b_i \geqslant 0$, $\forall i$ en otro caso se transforma para que así sea (multiplicando por -1 la ecuación).

$$\boxed{\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \\ & [x \in \mathbb{R}^n] \end{array}} \quad \Rightarrow \quad \boxed{\begin{array}{ll} \text{mín} & \sum_{i=1}^{n} x_i^a \\ \text{s.a} & Ax + x_a = b \\ & x, x_a \geqslant 0 \\ & [x, x_a \in \mathbb{R}^n] \end{array}}$$

Fase I

$$
\begin{array}{ll}
\text{mín} & \displaystyle\sum_{i=1}^{n} x_i^a \\
\text{s.a} & Ax + x_a = b \\
& x, x_a \geqslant 0 \\
& [x, x_a \in \mathbb{R}^n]
\end{array}
$$

- Una solución factible de este problema es:

$$\mathbf{x}^B = x_a, \qquad \mathbf{x}^N = 0$$

- Se aplica el algorimto del símplex, calculado el óptimo existen dos casos:
  1. Si en la solución óptima $x_a \neq 0$ el problema original es NO FACTIBLE
  2. Si en la solución óptima $x_a = 0$, esta solución es una solución factible básica del problema original.

Método de dos Fases

Fase I : Partiendo de $\mathbf{x}^B = x_a = b$, $\mathbf{x}^N = 0$, resolver

$$\begin{array}{ll} \min & \sum_{i=1}^{n} x_i^a \\ \text{s.a} & Ax + x_a = b \\ & x, x_a \geqslant 0 \end{array}$$

- Si $x_a \neq 0$:
- Si $x_a = 0$ ir a la Fase II , tomando como base inicial, la última de esta fase.

Fase II : Partiendo de la base B, definir $\mathbf{x}^B = B^{-1}b$ y $\mathbf{x}^N = 0$, y resolver:

$$\begin{array}{ll} \min & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \end{array}$$

Ejercicio

Utilizar el método de las dos fases para resolver el problema

$$
\begin{array}{llrrrrcl}
\text{mín} & x_1 - x_2 & & & & & & \\
\text{s.a.} & & x_1 & +x_2 & -x_3 & & = & 2 \\
& & -x_1 & +x_2 & +x_3 & -x_4 & = & 1 \\
& x_1, x_2, x_3, x_4 \geqslant 0 & & & & & &
\end{array}
$$

Convergencia del Símplex

- El método simple tal y como se ha descrito no es convergente
- La razón es que se pueden construir casos patológicos (Beale, 1955), donde el método quede atrapado en un ciclo infinito.
- Este fenómeno se denomina ciclaje, y las causas son las elecciones de los índices k y l que gestionan el cambio de base.
- Existen reglas que permiten evitar el ciclaje, como por ejemplo la regla de Bland (1977):
  - De entre las posibles variables que pueden entran en la base se escoge la de menor índice.
  - Si existen variables variables que pueden salir de la base (mismo ratio) se escoge la de menor índice
- Un algoritmo del símplex que usa la regla de Bland es convergente.

Eficiencia y Complejidad

- El número de operaciones en cada iteración del símplex es $\mathcal{O}(nm)$.
- El número de soluciones factibles básicas en un problema en forma estándar es del orden:

$$C_{n,m} = \frac{n!}{(n-m)!} \geqslant \left(\frac{n}{m}\right)^m \geqslant 2^m, \quad (\text{si } n \geqslant 2m)$$

- En el peor de los casos, el número de operaciones sería de orden exponencial:

$$\mathcal{O}(nm) * \mathcal{O}(2^m) = \mathcal{O}(nm2^m) \geqslant \mathcal{O}(2^n)$$

- Klee y Minty (1972) proponen un problema con 2^n soluciones básicas, en el cual el método símplex recorre todos ellas. Probando que el peor caso es posible.

Eficiencia y Complejidad

- La práctica ha constatado que casos extremos, como el ejemplo de Klee y Minty, ocurren rara vez.
- En la práctica:
  - En problemas de tamaño moderado, el símplex requiere de un número de iteraciones comprendido entre $4m$ y $6m$.
  - En problemas muy grandes el número esperado de iteraciones es αm donde α satisface

$$e^{\alpha} \log_2(2 + n/m)$$

- De lo dicho anteriormente, se deduce que la eficiencia práctica es $\mathcal{O}(nm^2)$.

Eficiencia y Complejidad

- A pesar de lo dicho anteriormente, es importante destacar que el problema de optimización lineal es de complejidad polinómica. Es decir, existen algoritmos más eficientes que el símplex.
- L. Khachiyan (1979) desarrolla el algoritmo del elipsoide, que es de tipo polinómico. Pero en la práctica se comporta peor que el símplex.
- N. Karmarkar (1984) desarrolla un método basado en puntos interiores, que también es de tipo polinómico. Este procedimiento se comporta mejor que el símplex en algunos casos, sobre todo en problemas muy grandes.

Tema 3

Programación Lineal

Introducción

Programación Lineal

El método símplex

Análisis de Sensibilidad
- Condiciones de Punto Óptimo
- Precios Sombra
- Análisis de Sensibilidad
- Cambios en el vector de costes
- Cambios en el término independiente

Dualidad

Objetivos

- Interpretación económica de los multiplicadores simplex (Precios Sombra)
- Obtener información de la solución de un problema lineal si se realizan modificaciones en alguno de los coeficientes sin necesidad de resolver el nuevo problema. (Análisis de Sensibilidad).
- Establecer la relaciones entre un problema y su dual.
- Interpretación económica del problema dual.

Condiciones de Karush-Kuhn-Tucker

$$\begin{array}{lll} \min & f(x) & \\ \text{s.a} & g_i(x) = 0, & i \in \mathcal{E} \\ & g_i(x) \geqslant 0, & i \in \mathcal{I} \end{array}$$

Lagrangiano asociado

$$\begin{aligned} L(x,\lambda) &= f(x) - \lambda^T G(x) \\ &= f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i g_i(x) \end{aligned}$$

Teorema : Si x^* es un punto regular y mínimo local del problema anterior, existe un vector de multiplicadores $\lambda^* = (\lambda_i^*)_{i \in \mathcal{E} \cup \mathcal{I}}$ que satisface:

$$\begin{aligned} \nabla_x L(x^*, \lambda^*) = 0,& \\ g_i(x^*) = 0,&\ \text{para todo } i \in \mathcal{E}, \\ g_i(x^*) \geqslant 0,&\ \text{para todo } i \in \mathcal{I}, \\ \lambda_i^* \geqslant 0,&\ \text{para todo } i \in \mathcal{I}, \\ \lambda_i^* g_i(x^*) = 0,&\ \text{para todo } i \in \mathcal{E} \cup \mathcal{I} \end{aligned}$$

Condiciones KKT en un problema lineal

$$\begin{aligned} \text{mín} \quad & c^T x \\ \text{s.a} \quad & Ax = b, \\ & x \geqslant 0, \end{aligned}$$

Lagrangiano asociado

$$L(x, \lambda, \mu) = c^T x - \lambda^T (Ax - b) - \mu^T x$$
$$= \sum_{j=1}^{n} c_j x_j - \sum_{i=1}^{m} \lambda_i \left(\sum_{j=1}^{n} a_i j x_j - b_i \right) - \sum_{j=1}^{n} \mu_j x_j$$

Condiciones de KKT

$$\begin{aligned} & \nabla_x L(x, \lambda, \mu) = c - A^T \lambda - \mu = 0, \\ & Ax = b \\ & x \geqslant 0 \\ & \mu \geqslant 0 \\ & \mu_j x_j = 0, \quad j = 1, \ldots, n \end{aligned}$$

Condiciones KKT en un problema lineal

Reordenando, estas condiciones se pueden expresar:

$$\begin{array}{l} \text{i) } Ax = b, \quad x \geqslant 0, \\ \text{ii) } c = A^T\lambda + \mu, \quad \mu \geqslant 0, \\ \text{iii)} \mu_j x_j = 0, \quad i = 1, \ldots, n. \end{array}$$

donde

i) Condiciones de factibilidad del problema

ii) Criterio de optimalidad (Costes reducidos positivos)

iii) Complementariedad de holguras. Los multiplicadores μ_j correspondientes a variables x_j no negativas son cero (condiciones no activas)

Condiciones KKT en un problema lineal

Veamos que la condición *ii*) es en realidad el Criterio de Optimalidad que se estableció en la clase anterior:

Sea B una base de A, es decir $A = [B, N]$ y $c^T = [c_B^T, c_N^T]$.

- La condición iii), implica $\mu_B = 0$
- Sustituyendo en la condición ii) se tiene:

$$\begin{bmatrix} c_B \\ c_N \end{bmatrix} = \begin{bmatrix} B^T \\ N^T \end{bmatrix} \lambda + \begin{bmatrix} 0 \\ \mu_N \end{bmatrix}$$

$$\Rightarrow \quad \begin{array}{l} c_B = B^T\lambda \\ c_N = N^T\lambda + \mu_N \end{array} \quad \Rightarrow \quad \begin{array}{r} \lambda = B^{-T}c_B \\ c_N = N^T B^{-T} c_B + \mu_N \end{array}$$

$$\xRightarrow{\mu_N \geqslant 0} c_N - N^T B^{-T} c_B \geqslant 0 \quad \Rightarrow \quad c_N^T - c_B^T B^{-1} N \geqslant 0^T$$

Condiciones KKT en un problema lineal

Resumen

$$\boxed{\begin{aligned} &\mu_N^T = c_N^T - c_B^T B^{-1} N \\ &\lambda = B^{-T} c_B \\ &c_N^T - c_B^T B^{-1} N \geqslant 0 \end{aligned}}$$

1. A μ_N, vector de multiplicadores de Lagrange de las condiciones activas $x \geqslant 0$, se le denomina vector de costes reducidos
2. A λ, vector de multiplicadores de Lagrange de las condiciones $Ax = b$ se le denomina vector de multiplicadores simplex
3. La condición $c_N^T - c_B^T B^{-1} N \geqslant 0$ es la condición de optimalidad

Interpretación económica de los multiplicadores simplex

Sea x^* una solución óptima de un problema de optimización lineal,

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b, \\ & x \geqslant 0, \end{array}$$

$$x^* = \begin{bmatrix} x_B^* \\ 0 \end{bmatrix} = \begin{bmatrix} B^{-1} b \\ 0 \end{bmatrix}$$

Si suponemos $x_B^* > 0$ (Solución No Degenerada) un cambio suficientemente pequeño δb en b no debe modificar la base óptima B; la solución óptima sería:

$$\hat{x}^* = \begin{bmatrix} \hat{x}_B^* \\ 0 \end{bmatrix} = \begin{bmatrix} B^{-1}(b + \delta b) \\ 0 \end{bmatrix}$$

El nuevo valor de la función objetivo será:

$$c^T \hat{x}_B^* = c_B^T B^{-1} b + c_B^T B^{-1} \delta b = c^T x_B^* + \lambda^{*T} \delta b$$

Interpretación económica de los multiplicadores simplex

Sea x^* una solución óptima de un problema de optimización lineal,

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b, \\ & x \geqslant 0, \end{array}$$

$$x^* = \begin{bmatrix} x_B^* \\ 0 \end{bmatrix} = \begin{bmatrix} B^{-1} b \\ 0 \end{bmatrix}$$

Si suponemos $x_B^* > 0$ (Solución No Degenerada) un cambio suficientemente pequeño δb en b no debe modificar la base óptima B; la solución óptima sería:

$$\hat{x}^* = \begin{bmatrix} \hat{x}_B^* \\ 0 \end{bmatrix} = \begin{bmatrix} B^{-1}(b + \delta b) \\ 0 \end{bmatrix}$$

El nuevo valor de la función objetivo será:

$$c^T \hat{x}_B^* = c_B^T B^{-1} b + c_B^T B^{-1} \delta b = c^T x_B^* + \lambda^{*T} \delta b$$

Interpretación económica de los multiplicadores

Si denotamos:

$$z^* = c^T x_B^* \qquad \hat{z}^* = z^* + \lambda^T \delta b$$

Obtenemos:

$$\left(\frac{\partial z}{\partial b}\right)_{x=x^*} = \lim_{\delta b \to 0} \frac{\hat{z}^* - z^*}{\delta b} = \lim_{\delta b \to 0} \frac{\lambda^T \delta b}{\delta b} = {\lambda^*}^T$$

Es decir:

λ_i^* indica cómo cambia la función objetivo al incrementar una unidad la cantidad disponible en el recurso b_i. Su valor es un precio sombra o valor marginal del recurso i-ésimo en el óptimo del problema: la cantidad máxima que se puede estar dispuesto a pagar por incrementar la disponibilidad del recurso i-ésimo.

Análisis de Sensibilidad

Objetivo: Obtener información de la solucón de un problema de optimización lineal si se realizan modificaciones en alguno de los coeficientes sin necesidad de resolver el problema modificado.

Este análisis es de gran importancia cuando existe incertidumbre en los parámetros del problema y se trabaja con estimaciones que oscilan entre ciertos márgenes.

- Cambios en el vector de costes c
- Cambios en el término independiente b
- Cambios en los coeficientes de A
- Incorporación de una nueva variable.
- Incorporación de una nueva restricción.

Cambios en el vector de costes c

Sea x^* solución del problema de programación lineal, en la base B, ($A = [B, N]$), $x^* = (x_B^*, x_N^*)$, y cuyo valor óptimo es $z = c^T x^* = c_B^T x_B^*$.

Consideremos la modificación al problema, $\theta_k \in \mathbb{R}$:

$$\begin{array}{ll} \text{mín} & \bar{c}^T x = (c + \theta_k)^T x \\ \text{s.a} & Ax = b, \\ & x \geqslant 0, \end{array}$$

Se debe dar respuesta a:

1. ¿Sigue siendo x^* solución óptima?
2. ¿Cúal es el nuevo valor de la función objetivo?

Se distinguen dos casos:

i) c_k corresponde a una variable básica ($k \in I$).

ii) c_k no corresponde a una variable básica ($k \in J$).

c_k corresponde a una variable básica ($k \in I$)

1. x^* sigue siendo solución óptima si los costes reducidos del problema modificado siguen siendo positivos, es decir:

$$(c_N^T - \bar{c}_B^T B^{-1} N) \geqslant 0 \quad \Rightarrow \quad (c_N^T - (c_B^T + \theta_k) B^{-1} N) \geqslant 0$$
$$\Rightarrow \quad (c_N^T - c_B^T B^{-1} N) - \theta_k B^{-1} N \geqslant 0$$

2. El nuevo valor de la función objetivo es:

$$\bar{z} = \bar{c}^T x^* = (c + \theta_k)^T x^* = c_B^T x_B^* + \theta_k x_k^* = z + \theta_k x_k^*$$

c_k no corresponde a una variable básica ($k \in J$)

1. x^* sigue siendo solución óptima si los costes reducidos del problema modificado siguen siendo positivos, es decir:

$$\bar{c}_N^T - c_B^T B^{-1} N \geqslant 0 \quad \Rightarrow \quad (c_N^T + \theta_k) - c_B^T B^{-1} N \geqslant 0$$
$$\Rightarrow (c_N^T - c_B^T B^{-1} N)_k + \theta_k \geqslant 0$$

   Sólo es necesario comprobar la componente k, pues el resto no varían.

2. El valor de la función objetivo no cambia

$$\bar{z} = \bar{c}^T x^* = c_B^T x_B^* = c^T x^* = z$$

Nota: Al igual que el caso anterior la condición anterior también puede ser interpretada en términos vectoriales.

Cambios en el término independiente b

Sea x^* solución del problema de programación lineal, en la base B, $(A = [B, N]$, $x^* = (x_B^*, x_N^*)$, y cuyo valor óptimo es $z = c^T x^* = c_B^T x_B^*$.

Consideremos la modificación al problema, $\delta_i \in \mathbb{R}$:

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = \bar{b} = b + \delta_i \\ & x \geqslant 0, \end{array}$$

Se debe dar respuesta a:

1. ¿Sigue siendo B base?
2. ¿Cúal es la nueva solución óptima?
3. ¿Cúal es el nuevo valor de la función objetivo?

Cambios en el término independiente b

1. La base no se modificará si la nueva solución básica es factible, es decir:

$$B^{-1}\bar{b} \geqslant 0 \Rightarrow B^{-1}(b + \delta_i) \geqslant 0 \Rightarrow B^{-1}b + \delta_i B^{-1}\mathbf{e}_i \geqslant 0$$

donde $\mathbf{e}_i = (0, \ldots, \overbrace{1}^{i}, \ldots, 0)$

2. La nueva solución óptima vendrá dada por:

$$\bar{x}^* = \begin{bmatrix} \bar{x}_B^* \\ \bar{x}_N^* \end{bmatrix}, \qquad \begin{array}{l} \bar{x}_B^* = B^{-1}(b + \delta_i) = x_B^* + \delta_i B^{-1}\mathbf{e}_i \\ \bar{x}_N^* = 0 \end{array}$$

3. El nuevo valor de la función objetivo es

$$\begin{aligned} \bar{z} = c^T \bar{x}^* = c_B^T \bar{x}_B^* = c_B^T B^{-1}(b + \delta_i) &= c_B^T B^{-1} b + \delta_i c_B^T B^{-1}\mathbf{e}_i \\ &= z + \delta_i \lambda_i \end{aligned}$$

donde $\lambda_i = [c_B^T B^{-1}]_i$ es el multiplicador simplex de la restricción i.

Ejemplo

$$\text{Sea}\left\{\begin{array}{ll} \text{mín} & 40x_1+53x_2 \\ s.a. & 60x_1+80x_2-x_3=360 \\ & 160x_1+120x_2-x_4=680 \\ & x_1,x_2,x_3,x_4\geqslant 0 \end{array}\right\}\text{, con solución:}$$

$$c^Tx^*=239, \quad x^*=\begin{bmatrix} 2\\3\\0\\0 \end{bmatrix}, \quad B=\begin{bmatrix} 60 & 80 \\ 160 & 120 \end{bmatrix}, \quad B^{-1}=\frac{1}{5600}\begin{bmatrix} -120 & 80 \\ 160 & -60 \end{bmatrix}$$

$$N=\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \lambda=c_B^TB^{-1}=\frac{1}{560}\begin{bmatrix} 368 \\ 2 \end{bmatrix}, \quad c_N^T-c_B^TB^{-1}N=\frac{1}{560}[368,2]$$

a) Determinar para qué rango de c_1 la solución x^* sigue siendo óptima. ¿Cuál es el valor de la función objetivo en ese caso?.

b) ¿Cuánto puede variar el coeficiente 360 de la primera restricción sin que cambie la base óptima?, ¿Cómo varía la función objetivo en ese caso?

c) ¿Cuál es la nueva solución del problema si la última restricción 680 se incrementa en 10 unidades?, ¿y el valor de la nueva función objetivo?

Ejemplo: solución

a) Como la variable x_1 está en la base, se debe verificar:

$$(c_N^T - c_B^T B^{-1} N) - \theta_1 B^{-1} N \geqslant 0 \Rightarrow \frac{1}{560}[368, 2] - \theta_1 \frac{1}{5600}[120, -80] \geqslant 0$$

$$\Rightarrow \begin{cases} 3680 - 120\theta_1 \geqslant 0 \Rightarrow \theta_1 \leqslant \frac{3680}{120} = \frac{92}{3} \\ 20 + 80\theta_1 \geqslant 0 \Rightarrow \theta_1 \geqslant -\frac{20}{80} = -\frac{1}{4} \end{cases} \Rightarrow -\frac{1}{4} \leqslant \theta_1 \leqslant \frac{92}{3}$$

Por tanto, el coeficiente c_1 puede variar en el rango:

$$\left[c_1 - \frac{1}{4}, c_1 + \frac{92}{3}\right] = [39.75, 70.67]$$

El valor de la función objetivo varía en el intervalo:

$$\bar{z} = z + \theta_1 x_1^* \in \left[239 - \frac{1}{4}2, 239 + \frac{92}{3}2\right] = [237, 300.33]$$

Ejemplo: solución

b) La base sigue siendo óptima si:

$$B^{-1}b + \delta_1 B^{-1}\mathbf{e}_1 \geqslant 0 \Rightarrow x_B^* + \delta_1 B^{-1}\mathbf{e}_1 \geqslant 0$$

$$\Rightarrow \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \delta_1 \frac{1}{5600} \begin{bmatrix} -120 \\ 160 \end{bmatrix} \geqslant 0 \Rightarrow \begin{cases} 2 - \frac{120}{5600}\delta_1 \geqslant 0 \\ 3 + \frac{160}{5600}\delta_1 \geqslant 0 \end{cases}$$

$$\Rightarrow \begin{cases} \delta_1 \leqslant \frac{11200}{120} = \frac{280}{3} \\ \delta_1 \geqslant -\frac{16800}{160} = -105 \end{cases} \Rightarrow -105 \leqslant \delta_1 \leqslant \frac{280}{3}$$

Luego, el coeficiente b_1 puede variar en el intervalo:

$$\left[b_1 - 105, b_1 + \frac{280}{3}\right] = [255, 453.33]$$

Para determinar la variación de la función objetivo, utilizamos el multiplicador simplex $\lambda_1 = \frac{368}{560}$:

$$\bar{z} = z + \delta_1\lambda_1 \in \left[239 - 105\frac{368}{560}, 239 + \frac{280}{3}\frac{368}{560}\right] = [170, 300.33]$$

Ejemplo: solución

c) Veamos primero si la base sigue siendo óptima, si no lo fuera sería necesario resolver de nuevo el problema. La condición en este caso es:

$$B^{-1}b + \delta_2 B^{-1}\mathbf{e}_2 \geqslant 0 \Rightarrow x_B^* + \delta_2 B^{-1}\mathbf{e}_2 \geqslant 0$$

$$\Rightarrow \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \delta_2 \frac{1}{5600} \begin{bmatrix} 80 \\ -60 \end{bmatrix} \geqslant 0 \Rightarrow \begin{cases} 2 + \frac{80}{5600}\delta_2 \geqslant 0 \\ 3 - \frac{60}{5600}\delta_2 \geqslant 0 \end{cases}$$

$$\Rightarrow \begin{cases} \delta_2 \geqslant -\frac{11200}{80} = -140 \\ \delta_2 \leqslant \frac{16800}{60} = 280 \end{cases} \Rightarrow -140 \leqslant \delta_2 \leqslant 280$$

Dado que $690 \in [680 - 140, 680 + 280]$ la base sigue siendo óptima. La nueva solución es:

$$\bar{x}_B^* = \frac{1}{5600} \begin{bmatrix} -120 & 80 \\ 160 & -60 \end{bmatrix} \begin{bmatrix} 360 \\ 690 \end{bmatrix} = \begin{bmatrix} \frac{15}{7} \\ \frac{81}{28} \end{bmatrix} \Rightarrow \hat{x}^* = [\frac{15}{7}, \frac{81}{28}, 0, 0]^T$$

El nuevo valor de la función objetivo:

$$\bar{z} = z + \delta_2\lambda_2 = 239 + 10\frac{2}{560} = 239.036$$

Tema 3

Programación Lineal

Introducción

Programación Lineal

El método símplex

Análisis de Sensibilidad

Dualidad
- Problema Lagrangiano Dual
- Teoremas de Dualidad
- Ejemplo primal-dual asociado

Problema Lagrangiano Dual

Problema Primal

$$
\begin{array}{ll}
\text{mín} & f(x) \\
\text{s.a} & g(x) = 0, \\
 & h(x) \geqslant 0, \\
 & x \in X
\end{array}
$$

Problema Dual (Lagrangiano)

$$
\begin{array}{ll}
\text{máx} & \theta(\lambda, \mu) \\
\text{s.a} & \mu \geqslant 0
\end{array}
$$

$$\theta(\lambda, \mu) := \inf_{x \in X} L(x, \lambda, \mu)$$

$$L(x, \lambda, \mu) := f(x) - \lambda^T g(x) - \mu^T h(x)$$

Teorema (Dualidad débil) Para todo $x \in X$, verificando $g(x) = 0,\ h(x) \geqslant 0$, y $\mu \geqslant 0$, se tiene

$$f(x) \geqslant \theta(\lambda, \mu)$$

Teorema (Dualidad fuerte) Si $X \subset \mathbb{R}^n$ es convexo, f, g convexas, h afín y

$$\exists\, \hat{x} \in X \,:\, h(\hat{x}) < 0,\ g(\hat{x}) = 0, \quad 0 \in g(\mathring{X}),\ g(X) := \{g(x), x \in X\}$$

entonces

$$\inf\{f(x) \,:\, g(x) = 0,\ h(x) \geqslant 0\} = \sup\{\theta(\lambda, \mu) \,:\, \mu \geqslant 0\}$$

Dualidad. Caso lineal

Aplicamos la teoría anterior al caso lineal

$$\boxed{\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \end{array}} \xRightarrow{\text{DUAL}: X=\{x\,:\,x\geqslant 0\}} \left\{ \begin{array}{ll} \text{máx} & \theta(\lambda) \\ \theta(\lambda) = \inf_{x\in X}\{c^T x - \lambda^T (Ax - b)\} \end{array} \right.$$

$$\theta(\lambda) = \lambda^T b + \inf_x (c^T x - \lambda^T Ax) = \lambda^T b + \inf_x (c^T - \lambda^T A)x$$

$$= \left\{ \begin{array}{ll} \lambda^T b & \text{si } c^T - \lambda^T A \geqslant 0 \\ -\infty & \text{otro caso} \end{array} \right.$$

Por tanto:

$$\text{máx}\,\theta(\lambda) \quad \Rightarrow \quad \boxed{\begin{array}{ll} \text{máx} & \lambda^T b \\ \text{s.a} & \lambda^T A \leqslant c^T \end{array}}$$

Dualidad. Caso lineal

Forma Estándar

PRIMAL $\Rightarrow$ *DUAL*

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax = b \\ & x \geqslant 0 \end{array} \quad \Rightarrow \quad \begin{array}{ll} \text{máx} & b^T \lambda \\ \text{s.a} & A^T \lambda \leqslant c \\ & \lambda \in \mathbb{R}^m \end{array}$$

Forma Simétrica de la Dualidad

PRIMAL $\Rightarrow$ *DUAL*

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax \geqslant b \\ & x \geqslant 0 \end{array} \quad \Rightarrow \quad \begin{array}{ll} \text{máx} & b^T \lambda \\ \text{s.a} & A^T \lambda \leqslant c \\ & \lambda \geqslant 0 \end{array}$$

Relación primal-dual

	Prob. min.		Prob. max.	
Variables	$\geqslant$	$\Longleftrightarrow$	$\leqslant$	Restricciones
	$\leqslant$	$\Longleftrightarrow$	$\geqslant$	
	Sin restric.	$\Longleftrightarrow$	$=$	
Restricciones	$\geqslant$	$\Longleftrightarrow$	$\geqslant$	Variables
	$\leqslant$	$\Longleftrightarrow$	$\leqslant$	
	$=$	$\Longleftrightarrow$	Sin restric.	

Ejemplo

Calcular el problema dual de:

$$\begin{array}{ll} \text{mín} & x_1 + x_2 + 5x_3 \\ s.a. & \begin{array}{cccc} x_1 & & +x_3 & \geqslant 1 \\ & x_2 & +x_3 & \geqslant 2 \end{array} \\ & x_1, x_2, x_3 \geqslant 0 \end{array}$$

$$\begin{array}{ll} \text{máx} & \lambda_1 + 2\lambda_2 \\ s.a. & \begin{array}{cccc} \lambda_1 & & \leqslant 1 \\ & \lambda_2 & \leqslant 1 \\ \lambda_1 & +\lambda_2 & \leqslant 5 \end{array} \\ & \lambda_1, \lambda_2 \geqslant 0 \end{array}$$

Ejemplo

Calcular el problema dual de:

$$
\begin{array}{ll}
\text{mín} & x_1 + x_2 + 5x_3 \\
\textit{s.a.} & \begin{array}{ccc} x_1 & & +x_3 & \geqslant 1 \\ & x_2 & +x_3 & \geqslant 2 \end{array} \\
 & x_1, x_2, x_3 \geqslant 0
\end{array}
$$

$$
\begin{array}{ll}
\text{máx} & \lambda_1 + 2\lambda_2 \\
\textit{s.a.} & \begin{array}{ccc} \lambda_1 & & \leqslant 1 \\ & \lambda_2 & \leqslant 1 \\ \lambda_1 & +\lambda_2 & \leqslant 5 \end{array} \\
 & \lambda_1, \lambda_2 \geqslant 0
\end{array}
$$

Ejemplo

Calcular el problema dual de:

$$
\begin{array}{ll}
\text{máx} & 8x_1 + 3x_2 - 2x_3 \\
s.a. & \begin{array}{rrrl} x_1 & -6x_2 & +x_3 & \geqslant 2 \\ 5x_1 & +7x_2 & -2x_3 & = -4 \end{array} \\
 & x_1 \leqslant 0, x_2 \geqslant 0, x_3 \in \mathbb{R}
\end{array}
$$

$$
\begin{array}{ll}
\text{mín} & 2\lambda_1 - 4\lambda_2 \\
s.a. & \begin{array}{rrl} \lambda_1 & +5\lambda_2 & \leqslant 8 \\ -6\lambda_1 & +7\lambda_2 & \geqslant 3 \\ \lambda_1 & -\lambda_2 & = -2 \end{array} \\
 & \lambda_1 \leqslant 0, \lambda_2 \in \mathbb{R}
\end{array}
$$

Ejemplo

Calcular el problema dual de:

$$\begin{array}{ll} \text{máx} & 8x_1 + 3x_2 - 2x_3 \\ s.a. & \begin{array}{rrrl} x_1 & -6x_2 & +x_3 & \geqslant 2 \\ 5x_1 & +7x_2 & -2x_3 & = -4 \end{array} \\ & x_1 \leqslant 0, x_2 \geqslant 0, x_3 \in \mathbb{R} \end{array}$$

$$\begin{array}{ll} \text{mín} & 2\lambda_1 - 4\lambda_2 \\ s.a. & \begin{array}{rrl} \lambda_1 & +5\lambda_2 & \leqslant 8 \\ -6\lambda_1 & +7\lambda_2 & \geqslant 3 \\ \lambda_1 & -\lambda_2 & = -2 \end{array} \\ & \lambda_1 \leqslant 0, \lambda_2 \in \mathbb{R} \end{array}$$

Ejemplo

Calcular el problema dual de:

$$
\begin{array}{ll}
\text{máx} & 8x_1 + 3x_2 - 2x_3 \\
s.a. & \begin{array}{rrrl} x_1 & -6x_2 & +x_3 & \geqslant 2 \\ 5x_1 & +7x_2 & -2x_3 & = -4 \end{array} \\
 & x_1 \leqslant 0, x_2 \geqslant 0, x_3 \in \mathbb{R}
\end{array}
$$

$$
\begin{array}{ll}
\text{mín} & 2\lambda_1 - 4\lambda_2 \\
s.a. & \begin{array}{rrl} \lambda_1 & +5\lambda_2 & \leqslant 8 \\ -6\lambda_1 & +7\lambda_2 & \geqslant 3 \\ \lambda_1 & -\lambda_2 & = -2 \end{array} \\
 & \lambda_1 \leqslant 0, \lambda_2 \in \mathbb{R}
\end{array}
$$

Teoremas de dualidad

PRIMAL(P)

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax \geqslant b \\ & x \geqslant 0 \end{array}$$

$\Rightarrow$

DUAL(D)

$$\begin{array}{ll} \text{máx} & b^T \lambda \\ \text{s.a} & A^T\lambda \leqslant c \\ & \lambda \geqslant 0 \end{array}$$

Teorema (Dualidad débil)

$$\left\{ \begin{array}{lll} \bar{x} \text{ satisface} & Ax \geqslant b, & x \geqslant 0 \\ \bar{\lambda} \text{ satisface} & A^T\lambda \leqslant c, & \lambda \geqslant 0 \end{array} \right\} \Rightarrow \quad c^T\bar{x} \geqslant b^T\bar{\lambda}$$

Demo:

$$c^T\bar{x} \geqslant \bar{\lambda}^T A\bar{x} \geqslant \bar{\lambda}^T b$$

Teoremas de dualidad

$PRIMAL(P)$

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax \geqslant b \\ & x \geqslant 0 \end{array}$$

$\Rightarrow$

$DUAL(D)$

$$\begin{array}{ll} \text{máx} & b^T \lambda \\ \text{s.a} & A^T\lambda \leqslant c \\ & \lambda \geqslant 0 \end{array}$$

Teorema (Dualidad débil)

$$\left\{ \begin{array}{lll} \bar{x} \text{ satisface} & Ax \geqslant b, & x \geqslant 0 \\ \bar{\lambda} \text{ satisface} & A^T\lambda \leqslant c, & \lambda \geqslant 0 \end{array} \right\} \quad \Rightarrow \quad c^T\bar{x} \geqslant b^T\bar{\lambda}$$

Demo:

$$c^T\bar{x} \geqslant \bar{\lambda}^T A\bar{x} \geqslant \bar{\lambda}^T b$$

Teoremas de dualidad

$PRIMAL(P)$

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax \geqslant b \\ & x \geqslant 0 \end{array}$$

$\Rightarrow$

$DUAL(D)$

$$\begin{array}{ll} \text{máx} & b^T \lambda \\ \text{s.a} & A^T\lambda \leqslant c \\ & \lambda \geqslant 0 \end{array}$$

Teorema (Dualidad débil)

$$\left\{ \begin{array}{lll} \bar{x} \text{ satisface} & Ax \geqslant b, & x \geqslant 0 \\ \bar{\lambda} \text{ satisface} & A^T\lambda \leqslant c, & \lambda \geqslant 0 \end{array} \right\} \Rightarrow \quad c^T\bar{x} \geqslant b^T\bar{\lambda}$$

Demo:

$$c^T\bar{x} \geqslant \bar{\lambda}^T A\bar{x} \geqslant \bar{\lambda}^T b$$

Teoremas de dualidad

$$PRIMAL(P) \qquad \boxed{\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a} & Ax \geqslant b \\ & x \geqslant 0 \end{array}} \quad \Rightarrow \quad DUAL(D) \qquad \boxed{\begin{array}{ll} \text{máx} & b^T \lambda \\ \text{s.a} & A^T\lambda \leqslant c \\ & \lambda \geqslant 0 \end{array}}$$

Teorema (Dualidad fuerte) Una de las siguientes alternativas es cierta:

- Primal y dual poseen soluciones óptimas: x^* y λ^*, y satisfacen $\boxed{c^T x^* = b^T \lambda^*}$.
- Uno de los problemas es no acotado y el otro no factible.
- Ambos problemas son no factibles

Relación Problema Primal - Dual

Condiciones de Holgura Complementaria
Sean x^* y λ^* soluciones óptimas de (P) y (D) en forma simétrica de dualidad, entonces:

$$(c - A^T\lambda)^T x^* = 0$$
$$\lambda^T(Ax^* - b) = 0$$

Solución del problema dual Si $x^* = (x_B^*, x_N^*)$ es la solución óptima del primal, entonces el problema dual tiene por solución:

$$\lambda^* = B^{-T} c_B$$

Nota: La solución del problema dual son los denominados multiplicadores simplex o precios sombra (ver pág.12) y satisfacen:

$$\left(\frac{\partial z}{\partial b}\right)_{x=x^*} = \lambda^*$$

Ejemplo primal-dual asociado

Problema Primal
Un empresario quiere vender $b_1, \ldots, b_m$ cantidades de m productos.

Para ello se pone en contacto con un fabricante que se los puede elaborar realizando para ello n actividades distintas en sus fábricas. El coste unitario de cada actividad j lo fija el fabricante en c_j.

Si a_{ij} representa la cantidad de producto i producido por una unidad de actividad j, $\sum_{j=1}^{n} a_{ij}x_j$ representa las unidades que se producen del producto i en todas las actividades.

Las unidades requeridas del producto i son b_i.

El problema que debe resolver el empresario es:

$$\begin{array}{ll} \text{mín} & \sum_{j=1}^{n} c_j x_j \\ \text{s.a} & \sum_{j=1}^{n} a_{ij}x_j \geqslant b_i,\ i = 1, \ldots, m \\ & x_j \geqslant 0,\ j = 1, \ldots, n \end{array}$$

Ejemplo primal-dual asociado

Problema Primal
Supóngase ahora que el empresario acuerda pagar al fabricante precios por unidad, $y_1, \ldots, y_m$ de cada uno de los m productos ya fabricados.

Si a_{ij} sigue siendo el número de unidades i producidas en cada actividad j la suma $\sum_{i=1}^{m} a_{ij}$ es el ingreso del fabricante por cada unidad de actividad j con los precios $y_1, \ldots, y_m$. Ese ingreso no debe exceder al coste prefijado c_j (precio justo)

El frabricante deberá fijar los niveles de produción b_i de cada i que maximicen su ganancia $\sum_{i=1}^{m} y_i b_i$.

El problema que se plantea fabricante es:

$$
\begin{array}{ll}
\text{máx} & \sum_{i=1}^{m} y_i b_i \\
\text{s.a} & \sum_{i=1}^{m} a_{ij} y_i \leqslant c_j,\ j = 1, \ldots, n \\
& y_i \geqslant 0,\ i = 1, \ldots, m
\end{array}
$$

Ejemplo primal-dual asociado

El programa primal trata en este caso sobre valores económicos. Garantizadas unas ventas mínimas de productos a unos precios, busca el esquema concreto de producción que minimiza el coste total, o maximiza el valor económico de la producción.

El programa dual de cantidades físicas. Con un esquema de costes unitario dado, busca qué cantidades de productos se deben producir a unos precios conocidos para maximizar el beneficio de su venta.

Tema 4

Programación Cuadrática

Introducción/Motivación

Preliminares

Elección de cartera. Restricciones de Igualdad

Elección de cartera. Restricciones de Desigualdad

Elección de cartera. Maximización de utilizad y línea del mercado de capitales

Ejercicio

Tema 4

Programación Cuadrática

Introducción/Motivación

Preliminares

Elección de cartera. Restricciones de Igualdad

Elección de cartera. Restricciones de Desigualdad

Elección de cartera. Maximización de utilizad y línea del mercado de capitales

Ejercicio

Motivación: Elección de cartera

Objetivo Dada una cantidad C, se desea invertir en tres activos A, B y C: ¿cúal debería ser la distribución de la inversión para obtener el mayor rendimiento con el mínimio riesgo?

Algunos conceptos

- **Activo**: se modeliza mediante una variable aleatoria

$$(A, p_A(t))$$

- **Rendimiento/Retorno**: está determinado por la diferencia del precio del tírulo al final y principio de la inversión y por sus dividendos:

$$r_A(t) = \frac{p_A(t) - p_A(t-1) + d_A(t)}{p_A(t-1)}$$

  El rendimiento se estima mediante su esperanza matemática, $E(r_A) = \mu_A$.

- **Riesgo**: se modeliza como la varianza o desviación típica del rendimiento: $\sigma_A^2 = E((r_A - E(r_A))^2)$

Motivación: Elección de cartera

Objetivo Dada una cantidad C, se desea invertir en tres activos A, B y C: ¿cúal debería ser la distribución de la inversión para obtener el mayor rendimiento con el mínimio riesgo?

Algunos conceptos

- **Activo**: se modeliza mediante una variable aleatoria

$$(A, p_A(t))$$

- **Rendimiento/Retorno**: está determinado por la diferencia del precio del tírulo al final y principio de la inversión y por sus dividendos:

$$r_A(t) = \frac{p_A(t) - p_A(t-1) + d_A(t)}{p_A(t-1)}$$

  El rendimiento se estima mediante su esperanza matemática, $E(r_A) = \mu_A$.

- **Riesgo**: se modeliza como la varianza o desviación típica del rendimiento: $\sigma_A^2 = E((r_A - E(r_A))^2)$

Acceso a datos:

‣ https://es.finance.yahoo.com/

Banco Bilbao Vizcaya Argentaria, S.A. (BBVA.MC) - MCE

5,8420 ↓0,1070(1.80%) 17:38

Cotizaciones históricas

Consulta Cotizaciones históricas

ELEGIR RANGO DE FECHAS

Fecha de inicio: 3 Ene 2000 Por ejemplo: Ene 1, 2010

Fecha de finalización: 31 Mar 2016

- (•) Diario
- ( ) Semanal
- ( ) Mensual
- ( ) Sólo dividendos

Consulta la cotización

Primera | Anterior | Siguiente | Última

PRECIOS

Fecha	Apertura	Máximo	Mínimo	Cerrar	Volumen	Ajustes de Cierre*
30 de mar de 2016	6,01	6,07	5,95	5,95	52.624.900	5,82
29 de mar de 2016	6,08	6,13	5,90	5,95	33.692.300	5,82
28 de mar de 2016	6,01	6,01	6,01	6,01	0	5,88
25 de mar de 2016	6,01	6,01	6,01	6,01	0	5,88
24 de mar de 2016	6,09	6,14	5,98	6,01	46.231.200	5,88
23 de mar de 2016	6,29	6,32	6,10	6,14	53.492.300	6,01
22 de mar de 2016	6,24	6,26	6,12	6,26	119.298.600	6,13
21 de mar de 2016	6,32	6,46	6,28	6,30	32.057.300	6,16
18 de mar de 2016	6,34	6,47	6,23	6,38	57.456.500	6,24
17 de mar de 2016	6,44	6,45	6,12	6,33	61.612.500	6,19
16 de mar de 2016	6,35	6,45	6,22	6,31	56.819.500	6,17

Algunos conceptos

- **Cartera (Portfolio)**: activo formado por la combinación de varios activos: $C = (A_1, A_2, \ldots, A_n)$

  Vector de pesos: $\mathbf{x} = (x_1, x_2, \ldots, x_n)$
  - x_i porcentaje de participación en A_i
  - $\sum_{i=1}^{n} x_i = 1$
  - Si $x_i \geqslant 0, \forall i$ no se admiten ventas en corto
- **Rendimiento/Retorno de la Cartera**: se estima como la esperanza matemática de su valor:

$$E(r_C) = \bar{\mu} = \sum_{i=1}^{n} x_i \mu_i = \mathbf{x}' \boldsymbol{\mu}$$

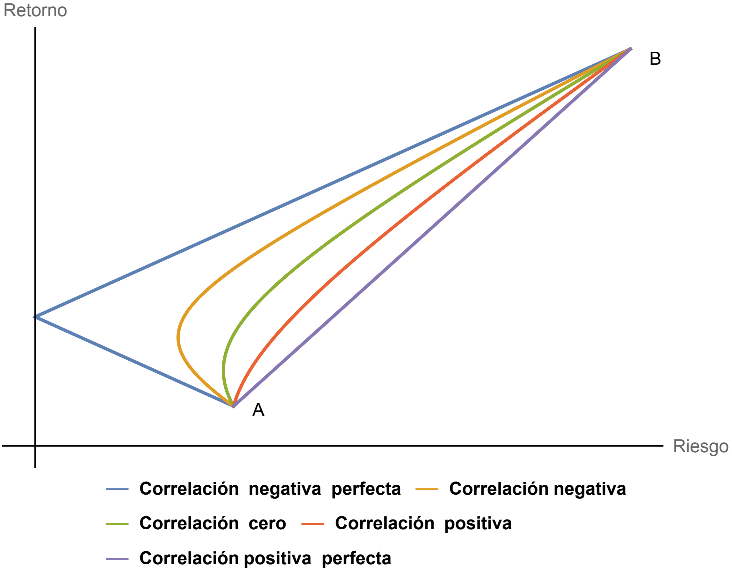
  siendo $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)$
- **Riesgo de la Cartera**: viene dado por su varianza o desviación típica:

$$\sigma_C^2 = \sum_{i,j=1}^{n} x_i x_j V_{ij} = \mathbf{x}' V \mathbf{x}$$

  donde V es la matriz de varianzas-covarianzas:

$$V = E((\mathbf{r} - E(\mathbf{r}))(\mathbf{r} - E(\mathbf{r}))'), \qquad \text{con } \mathbf{r} = (r_1, r_2, \ldots, r_n)$$

Gráfico Riesgo-Rendimiento para dos activos



Objetivo Dada una cantidad C, se desea invertir en tres activos A, B y C: ¿cúal debería ser la distribución de la inversión para obtener el mayor rendimiento con el mínimio riesgo?

Formulación Matemática

- Modelo de Markowitz (1952)
- $\bar{\mu}$ rendimiento esperado
- Si se admiten operaciones en corto no hay restricción de no negatividad

$$
\begin{array}{ll}
\text{mín} & \frac{1}{2}\mathbf{x}'V\mathbf{x} \\
\text{s.a} & \mathbf{x}'\mu = \bar{\mu}, \\
 & \mathbf{x}'\mathbf{1} = 1, \\
 & \mathbf{x} \geqslant 0
\end{array}
$$

Ejemplo Elección de cartera:

Datos

$$\mu_A = 0.10 \quad \sigma_A = 0.71 \qquad\qquad V = \begin{pmatrix} 0.0050 & -0.0060 & 0.0020 \\ -0.0060 & 0.0199 & -0.0158 \\ 0.0020 & -0.0158 & 0.0790 \end{pmatrix}$$
$$\mu_B = 0.25 \quad \sigma_B = 0.14$$
$$\mu_C = 0.33 \quad \sigma_C = 0.28$$

Composición de la *mejor cartera* con rendimiento 0.2:

$$\begin{array}{ll} \text{mín} & \frac{1}{2}(x_A, x_B, x_C)'V(x_A, x_B, x_C) \\ \text{s.a} & x_A\mu_A + x_B\mu_B + x_C\mu_C = 0.2, \\ & x_A + x_B + x_C = 1, \\ & \mathbf{x} \geqslant 0 \end{array} \quad \Rightarrow \quad \begin{array}{l} x_A = 0.41 \\ x_B = 0.44, \\ x_C = 0.15 \\ \sigma_C = 0.05 \end{array}$$

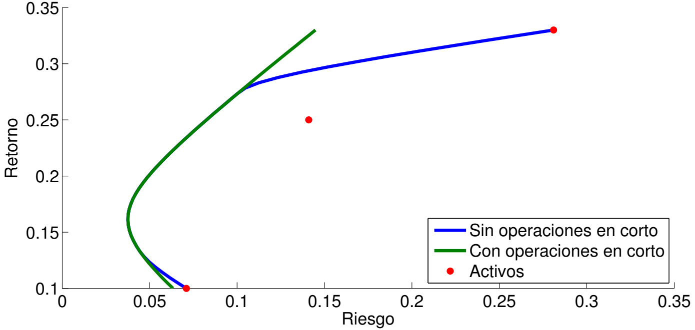
Ejemplo Elección de cartera: Frontera Eficiente

Datos

$$\mu_A = 0.10 \quad \sigma_A = 0.07$$
$$\mu_B = 0.25 \quad \sigma_B = 0.14$$
$$\mu_C = 0.33 \quad \sigma_C = 0.28$$

$$V = \begin{pmatrix} 0.0050 & -0.0060 & 0.0020 \\ -0.0060 & 0.0199 & -0.0158 \\ 0.0020 & -0.0158 & 0.0790 \end{pmatrix}$$

Frontera eficiente



Tema 4

Programación Cuadrática

Introducción/Motivación

Preliminares
- Optimización con Restricciones de Igualdad
- Optimización con Restricciones de Desigualdad
- Método de Newton

Elección de cartera. Restricciones de Igualdad

Elección de cartera. Restricciones de Desigualdad

Elección de cartera. Maximización de utilizad y línea del mercado de capitales

Ejercicio

Condición necesaria de Extremo Local (Lagrange)

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.a} & g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m \end{array}$$

Lagrangiano asociado

$$\begin{aligned} L(\mathbf{x}, \lambda) &= f(\mathbf{x}) - \boldsymbol{\lambda}^T G(\mathbf{x}) \\ &= f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) \end{aligned}$$

Definición: $\mathbf{x}$ se dice un punto regular si se verifica $\text{rang}(DG) = m$, $\left(DG = \frac{\partial(g_1,\dots,g_m)}{\partial(x_1,\dots x_n)}\right)$

Teorema (Condición Necesaria): Si $\mathbf{x}^*$ es un punto regular y mínimo local del problema anterior, existe un único vector de multiplicadores $\boldsymbol{\lambda}^* = (\lambda_i^*)_{i=1,\dots,m}$ que satisface:

$$\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*) &= 0, \\ g_i(\mathbf{x}^*) &= 0, \text{ para todo } i = 1, \dots, n \end{aligned}$$

Nota: Si $\mathbf{x}^*$ es máximo verfica la misma condición.

Condicióndes de óptimo de Segundo Orden

Necesaria: Sea $\mathbf{x}^*$ un punto regular y mínimo local del problema anterior, y $\boldsymbol{\lambda}^*$ el vector de multiplicadores de Lagrange asociado, entonces la matriz $H = \mathcal{L}_{\mathbf{xx}}(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ es semi-definida positiva para todo $\mathbf{z}$ que verifique $DG\,\mathbf{z} = 0$, *i.e.*

$$\mathbf{z}^T H \mathbf{z} \geqslant 0, \qquad \forall\, \mathbf{z}, \quad DG\,\mathbf{z} = 0$$

Suficiente: Si $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ satisfacen las condiciones de primer orden y $H = \mathcal{L}_{\mathbf{xx}}(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ es definida positiva para todo $\mathbf{z}$ que verifique $DG\,\mathbf{z} = 0$, entonces $\mathbf{x}^*$ es un mínimo local del problema.

Caso convexo: Si f es convexa y $g_i(x)$ lineales todo punto que cumpla las condiciones necesarias de mínimo local es mínimo global.

Condición necesaria de Extremo Local (KKT)

$$
\begin{array}{lll}
\text{mín} & f(\mathbf{x}) & \\
\text{s.a} & g_i(\mathbf{x}) = 0, & i \in \mathcal{E} \\
 & g_i(\mathbf{x}) \leqslant 0, & i \in \mathcal{I}
\end{array}
$$

Lagrangiano asociado

$$
\begin{aligned}
L(\mathbf{x}, \lambda) &= f(\mathbf{x}) - \lambda^T G(\mathbf{x}) \\
&= f(\mathbf{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i g_i(\mathbf{x})
\end{aligned}
$$

Defnción: Una restricción $g_i(x)$ se dice *activa* en el punto $\mathbf{x}$ si se verifica $g_i(\mathbf{x}) = 0$

Defnción: Un punto $\mathbf{x}$ se dice un punto regular si el rango del jacobiano de las restricciones activas en el punto $\mathbf{x}$ es máximo, es decir:

$$A = \{i : g_i(\mathbf{x}) = 0\}, \quad G_a = \{g_i : i \in A\} \quad \Rightarrow \quad \text{rang}(DG_a) = |A|$$

Condición necesaria de Extremo Local (KKT)

$$\begin{array}{lll} \text{mín} & f(\mathbf{x}) & \\ \text{s.a} & g_i(\mathbf{x}) = 0, & i \in \mathcal{E} \\ & g_i(\mathbf{x}) \leqslant 0, & i \in \mathcal{I} \end{array}$$

Lagrangiano asociado

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\lambda}) &= f(\mathbf{x}) - \boldsymbol{\lambda}^T G(\mathbf{x}) \\ &= f(\mathbf{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i g_i(\mathbf{x}) \end{aligned}$$

Teorema (Condición Necesaria): Si $\mathbf{x}^*$ es un punto regular y mínimo local del problema anterior, existe un vector de multiplicadores $\boldsymbol{\lambda}^* = (\lambda_i^*)_{i \in \mathcal{E} \cup \mathcal{I}}$ que satisface:

$$\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*) &= 0, \\ g_i(\mathbf{x}^*) &= 0, \ \text{para todo } i \in \mathcal{E}, \\ g_i(\mathbf{x}^*) &\leqslant 0, \ \text{para todo } i \in \mathcal{I}, \\ \lambda_i^* &\leqslant 0, \ \text{para todo } i \in \mathcal{I}, \\ \lambda_i^* g_i(\mathbf{x}^*) &= 0, \ \text{para todo } i \in \mathcal{E} \cup \mathcal{I} \end{aligned}$$

Condición necesaria de Extremo Local (KKT)

$$
\begin{array}{ll}
\text{mín} & f(\mathbf{x}) \\
\text{s.a} & g_i(\mathbf{x}) = 0, \quad i \in \mathcal{E} \\
 & g_i(\mathbf{x}) \leqslant 0, \quad i \in \mathcal{I}
\end{array}
$$

Lagrangiano asociado

$$
\begin{aligned}
L(\mathbf{x}, \boldsymbol{\lambda}) &= f(\mathbf{x}) - \boldsymbol{\lambda}^T G(\mathbf{x}) \\
&= f(\mathbf{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i g_i(\mathbf{x})
\end{aligned}
$$

Nota:

Si las restricciones NO son *activas* $\Rightarrow$ $\lambda_i = 0$

Si el multiplicador $\lambda_i < 0$ $\Rightarrow$ La restricción es *activa*

Condiciónes de óptimo de Segundo Orden

Suficiente: Si $(\mathbf{x}^*, \lambda^*)$ satisfacen las condiciones de KKT, $\mathbf{x}^*$ es un punto regular, y las restricciones NO activas no lo son en una vecindad de $\mathbf{x}^*$.
Sea λ_A^* el conjunto de multiplicadores asociado a las restricciones activas.
Si $(\mathbf{x}^*, \lambda_A^*)$ satisface las condiciones de suficientes de mínimio de Lagrange, sobre el conjunto de restricciones activas, entonces $\mathbf{x}^*$ es un mínimo local del problema.

Caso convexo: Si f es convexa y g_i son convexas, todo punto que cumpla las condiciones necesarias de mínimo local para KKT es mínimo global.

Método de Newton (una variable)

Objetivo. Dada $f : [a,b] \to \mathbb{R}$ calcular $x^* \in [a,b]$ tal que $f(x^*) = 0$.

Construcción del método de Newton. Si $x^* \in [a,b]$ una raíz de la ecuación $f(x)$, y f es suficientemente regular:

$$0 = f(x^*) = f(x) + f'(x)(x - x^*) + \frac{1}{2} f''(\xi)(x - x^*)^2, \qquad \xi \in (a,b)$$

Cerca de x^*, $(x - x^*)^2$ es *pequeño*, y se puede despreciar. Entonces x^* puede ser aproximado por:

$$\boxed{x^* \approx x - \frac{f(x)}{f'(x)}}$$

Método de Newton (una variable)

Objetivo. Dada $f:[a,b]\to\mathbb{R}$ calcular $x^*\in[a,b]$ tal que $f(x^*)=0$.

Construcción del método de Newton. Si $x^*\in[a,b]$ una raíz de la ecuación $f(x)$, y f es suficientemente regular:

$$0=f(x^*)=f(x)+f'(x)(x-x^*)+\frac{1}{2}f''(\xi)(x-x^*)^2,\qquad \xi\in(a,b)$$

Cerca de x^*, $(x-x^*)^2$ es *pequeño*, y se puede despreciar. Entonces x^* puede ser aproximado por:

$$\boxed{x^*\approx x-\frac{f(x)}{f'(x)}}$$

Método de Newton (una variable)

Objetivo. Dada $f:[a,b]\to\mathbb{R}$ calcular $x^*\in[a,b]$ tal que $f(x^*)=0$.

Construcción del método de Newton. Si $x^*\in[a,b]$ una raíz de la ecuación $f(x)$, y f es suficientemente regular:

$$0=f(x^*)=f(x)+f'(x)(x-x^*)+\frac{1}{2}f''(\xi)(x-x^*)^2, \qquad \xi\in(a,b)$$

Cerca de x^*, $(x-x^*)^2$ es *pequeño*, y se puede despreciar. Entonces x^* puede ser aproximado por:

$$\boxed{x^*\approx x-\frac{f(x)}{f'(x)}}$$

Algoritmo y Pseudocódigo de Newton (una variable)

Algoritmo de Newton

- INICIO: $x_0 \in (a, b)$ *cercano* a x^*
- Dado x_n se calcula x_{n+1}:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Pseudocódigo de Newton

INPUT: x_0, M, δ, ϵ
begin
read(x_0, M, δ, ϵ*);*
$v = f(x_0)$;
if $|v| < \epsilon$ **then return** x_0; **endif**
for $i = 1$ **to** M **do**
    $x_1 = x_0 - v/f'(x_0)$;
    $v = f(x_1)$;
    if $|x_1 - x_0| \leqslant \delta$ *OR* $|v| < \epsilon$
        then return x_1; **endif**
    $x_0 = x_1$
endfor
write*("Error: No se alcanzo convergencia");*
end

Ejercicio. Implementar el método de Newton en matlab

Interpretación geométrica (una variable)

Es fácil demostrar que el punto x_{n+1} de la iteración de Newton, resulta de la intersección de la recta tangente a $f(x)$ en el punto $(x_n, f(x_n))$ con el eje OX.

Ejemplo

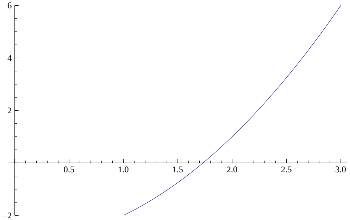
$$f(x) = x^2 - 3$$

- $x_0 = 3$

Interpretación geométrica (una variable)

Es fácil demostrar que el punto x_{n+1} de la iteración de Newton, resulta de la intersección de la recta tangente a $f(x)$ en el punto $(x_n, f(x_n))$ con el eje OX.

Ejemplo



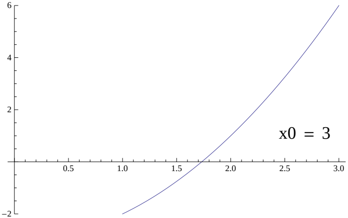
$$f(x) = x^2 - 3$$

- $x_0 = 3$

Interpretación geométrica (una variable)

Es fácil demostrar que el punto x_{n+1} de la iteración de Newton, resulta de la intersección de la recta tangente a $f(x)$ en el punto $(x_n, f(x_n))$ con el eje OX.

Ejemplo



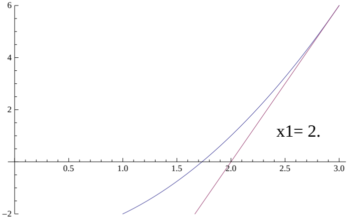
$$f(x) = x^2 - 3$$

- $x_0 = 3$

Interpretación geométrica (una variable)

Es fácil demostrar que el punto x_{n+1} de la iteración de Newton, resulta de la intersección de la recta tangente a $f(x)$ en el punto $(x_n, f(x_n))$ con el eje OX.

Ejemplo



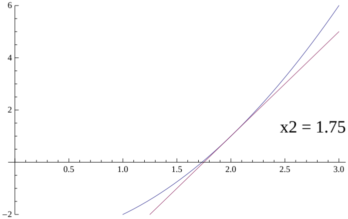
$$f(x) = x^2 - 3$$

- $x_0 = 3$
- $x_1 \equiv \begin{cases} y = 0 \\ y - f(x_0) = f'(x_0)(x - x_0) \end{cases}$

Interpretación geométrica (una variable)

Es fácil demostrar que el punto x_{n+1} de la iteración de Newton, resulta de la intersección de la recta tangente a $f(x)$ en el punto $(x_n, f(x_n))$ con el eje OX.

Ejemplo



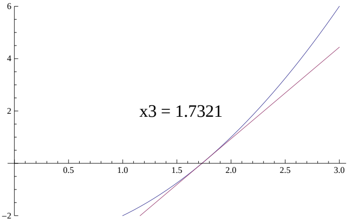
$$f(x) = x^2 - 3$$

- $x_0 = 3$
- $x_1 \equiv \begin{cases} y = 0 \\ y - f(x_0) = f'(x_0)(x - x_0) \end{cases}$
- $x_2 \equiv \begin{cases} y = 0 \\ y - f(x_1) = f'(x_1)(x - x_1) \end{cases}$

Interpretación geométrica (una variable)

Es fácil demostrar que el punto x_{n+1} de la iteración de Newton, resulta de la intersección de la recta tangente a $f(x)$ en el punto $(x_n, f(x_n))$ con el eje OX.

Ejemplo



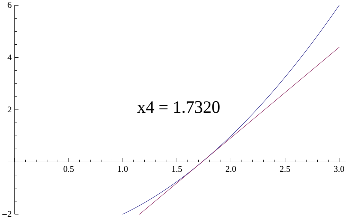
$$f(x) = x^2 - 3$$

- $x_0 = 3 \ldots$
- $x_2 \equiv \begin{cases} y = 0 \\ y - f(x_1) = f'(x_1)(x - x_1) \end{cases}$
- $x_3 \equiv \begin{cases} y = 0 \\ y - f(x_2) = f'(x_2)(x - x_2) \end{cases}$

Interpretación geométrica (una variable)

Es fácil demostrar que el punto x_{n+1} de la iteración de Newton, resulta de la intersección de la recta tangente a $f(x)$ en el punto $(x_n, f(x_n))$ con el eje OX.

Ejemplo



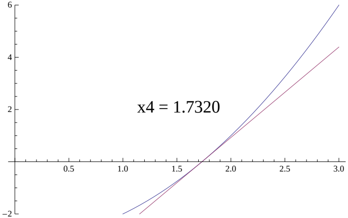
$$\boxed{f(x) = x^2 - 3}$$

- $x_0 = 3 \ldots$
- $x_3 \equiv \begin{cases} y = 0 \\ y - f(x_2) = f'(x_2)(x - x_2) \end{cases}$
- $x_4 \equiv \begin{cases} y = 0 \\ y - f(x_3) = f'(x_3)(x - x_3) \end{cases}$

Interpretación geométrica (una variable)

Es fácil demostrar que el punto x_{n+1} de la iteración de Newton, resulta de la intersección de la recta tangente a $f(x)$ en el punto $(x_n, f(x_n))$ con el eje OX.

Ejemplo



$$f(x) = x^2 - 3$$

- $x_0 = 3 \ldots$
- $x_4 \equiv \begin{cases} y = 0 \\ y - f(x_3) = f'(x_3)(x - x_3) \end{cases}$

$$x_4 = 1.73205 \quad \longrightarrow \quad \sqrt{3} \approx 1.732050807568\ldots$$

Convergencia del Método de Newton (una variable)

Teorema.(LOCAL) Sea $f : (a, b) \to \mathbb{R}$ verificando:

1. f' es Lipchitz en (a, b):

$$\exists \gamma \ : \ |f'(x) - f'(y)| \leqslant \gamma |x - y| \qquad \forall x, y \in (a, b)$$

2. $f'(x)$ no se anula en (a, b):

$$\exists \rho > 0 \ : \ |f'(x)| \geqslant \rho, \qquad \forall x \in (a, b)$$

3. $f(x) = 0$ tiene una solución $x^* \in (a, b)$

Entonces: **exige algún** μ tal que si $|x_0 - x^*| \leqslant \mu$ la sucesión generada por el algoritmo de Newton **converge hacia** x^* y además lo hace con orden cuadrático, es decir:

$$\boxed{|x_{n+1} - x_n| \leqslant \frac{\gamma}{2\rho} |x_n - x^*|^2}$$

Convergencia del Método de Newton (una variable)

Observaciones

- En general, se require que la aproximación initial x_0 este suficientemente cerca de la solución x^*, para que el método converja.
- Existe una versión global del teorema anterior, que garantiza convergencia de la iteración de Newton para cualquier aproximación inicial x_0, si la función $f : \mathbb{R} \to \mathbb{R}$ es estrictamente convexa (o concava).
- Se puede probar un resultado análogos para el caso $f'(x^*) = 0$ (raíz doble), sin embargo, en este caso se pierde la convergencia cuadrática.
- En la práctica, el método de Newton se combina con otros métodos *más lentos*, aunque más robustos, que permiten determinar una aproximación incial.

Método de Newton (varias variables)

Objetivo. Sea $\Omega \subset \mathbb{R}^d$. Dada $\mathbf{F} : \Omega \to \mathbb{R}^d$ calcular $\mathbf{x}^* \in \Omega$ tal que $\mathbf{F}(\mathbf{x}^*) = 0$.

Observación. El método anterior es aplicable en problemas de optimización $f : \Omega \to \mathbb{R}$, donde f es diferenciable:

$$\mathbf{x}^* \text{ óptimo } \Rightarrow \nabla f(\mathbf{x}) = \mathbf{0}$$

Construcción del método de Newton. Si $\mathbf{x}^* \in \Omega$ tal que $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$, y $\mathbf{F}$ es suficientemente regular:

$$0 = \mathbf{F}(\mathbf{x}^*) = \mathbf{F}(\mathbf{x}) + d\mathbf{F}(\mathbf{x})(\mathbf{h}) + \mathcal{O}(\|\mathbf{h}\|^2)$$

siendo $\mathbf{h} := \mathbf{x}^* - \mathbf{x}$. Si $\|\mathbf{h}\|$ es *pequeño*, resulta

$$d\mathbf{F}(x)\mathbf{h} \approx -d\mathbf{F}(x) \Rightarrow \begin{cases} \mathbf{h} \approx -d\mathbf{F}^{-1}(\mathbf{x})\mathbf{F}(\mathbf{x}) \\ \mathbf{x}^* \approx \mathbf{x} + \mathbf{h} \end{cases}$$

Método de Newton (varias variables)

Objetivo. Sea $\Omega \subset \mathbb{R}^d$. Dada $\mathbf{F} : \Omega \to \mathbb{R}^d$ calcular $\mathbf{x}^* \in \Omega$ tal que $\boxed{\mathbf{F}(\mathbf{x}^*) = 0}$.

Observación. El método anterior es aplicable en problemas de optimización $f : \Omega \to \mathbb{R}$, donde f es diferenciable:

$$\mathbf{x}^* \text{ óptimo } \Rightarrow \nabla f(\mathbf{x}) = \mathbf{0}$$

Construcción del método de Newton. Si $\mathbf{x}^* \in \Omega$ tal que $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$, y $\mathbf{F}$ es suficientemente regular:

$$0 = \mathbf{F}(\mathbf{x}^*) = \mathbf{F}(\mathbf{x}) + d\mathbf{F}(\mathbf{x})(\mathbf{h}) + \mathcal{O}(\|\mathbf{h}\|^2)$$

siendo $\boxed{\mathbf{h} := \mathbf{x}^* - \mathbf{x}}$. Si $\|\mathbf{h}\|$ es *pequeño*, resulta

$$\boxed{d\mathbf{F}(x)\mathbf{h} \approx -d\mathbf{F}(x) \Rightarrow \begin{cases} \mathbf{h} \approx -d\mathbf{F}^{-1}(\mathbf{x})\mathbf{F}(\mathbf{x}) \\ \mathbf{x}^* \approx \mathbf{x} + \mathbf{h} \end{cases}}$$

Método de Newton (varias variables)

Objetivo. Sea $\Omega \subset \mathbb{R}^d$. Dada $\mathbf{F} : \Omega \to \mathbb{R}^d$ calcular $\mathbf{x}^* \in \Omega$ tal que $\boxed{\mathbf{F}(\mathbf{x}^*) = 0}$.

Observación. El método anterior es aplicable en problemas de optimización $f : \Omega \to \mathbb{R}$, donde f es diferenciable:

$$\mathbf{x}^* \text{ óptimo } \Rightarrow \nabla f(\mathbf{x}) = \mathbf{0}$$

Construcción del método de Newton. Si $\mathbf{x}^* \in \Omega$ tal que $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$, y $\mathbf{F}$ es suficientemente regular:

$$0 = \mathbf{F}(\mathbf{x}^*) = \mathbf{F}(\mathbf{x}) + d\mathbf{F}(\mathbf{x})(\mathbf{h}) + \mathcal{O}(\|\mathbf{h}\|^2)$$

siendo $\boxed{\mathbf{h} := \mathbf{x}^* - \mathbf{x}}$. Si $\|\mathbf{h}\|$ es *pequeño*, resulta

$$d\mathbf{F}(x)\mathbf{h} \approx -d\mathbf{F}(x) \Rightarrow \begin{cases} \mathbf{h} \approx -d\mathbf{F}^{-1}(\mathbf{x})\mathbf{F}(\mathbf{x}) \\ \mathbf{x}^* \approx \mathbf{x} + \mathbf{h} \end{cases}$$

Método de Newton (varias variables)

Objetivo. Sea $\Omega \subset \mathbb{R}^d$. Dada $\mathbf{F} : \Omega \to \mathbb{R}^d$ calcular $\mathbf{x}^* \in \Omega$ tal que $\mathbf{F}(\mathbf{x}^*) = 0$.

Observación. El método anterior es aplicable en problemas de optimización $f : \Omega \to \mathbb{R}$, donde f es diferenciable:

$$\mathbf{x}^* \text{ óptimo } \Rightarrow \nabla f(\mathbf{x}) = \mathbf{0}$$

Construcción del método de Newton. Si $\mathbf{x}^* \in \Omega$ tal que $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$, y $\mathbf{F}$ es suficientemente regular:

$$0 = \mathbf{F}(\mathbf{x}^*) = \mathbf{F}(\mathbf{x}) + d\mathbf{F}(\mathbf{x})(\mathbf{h}) + \mathcal{O}(\|\mathbf{h}\|^2)$$

siendo $\mathbf{h} := \mathbf{x}^* - \mathbf{x}$. Si $\|\mathbf{h}\|$ es *pequeño*, resulta

$$\boxed{d\mathbf{F}(x)\mathbf{h} \approx -d\mathbf{F}(x) \Rightarrow \begin{cases} \mathbf{h} \approx -d\mathbf{F}^{-1}(\mathbf{x})\mathbf{F}(\mathbf{x}) \\ \mathbf{x}^* \approx \mathbf{x} + \mathbf{h} \end{cases}}$$

Algoritmo de Newton (varias variables)

Algoritmo:

- INICIO: $\mathbf{x}_0 \in \Omega$ *cercano* a $\mathbf{x}^*$
- Dado $\mathbf{x}_n$ se calcula $\mathbf{x}_{n+1}$:

$$\begin{cases} d\mathbf{F}(\mathbf{x}^n)\mathbf{h} = -\mathbf{F}(\mathbf{x}_n) \\ \mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{h} \end{cases}$$

Ejercicio

1. Escribir el pseudocódigo.
2. Implementar el método de Newton en matlab

Convergencia de método de Newton (varias variables)

Teorema.(LOCAL) Sea $\Omega \subset \mathbb{R}^d$ abierto y convexo, y $\mathbf{F} : \Omega \to \mathbb{R}^d$ verificando:

1. $d\mathbf{F}'$ es Lipchitz en $B(\mathbf{x}^*, r) \subset \Omega$:

$$\exists \gamma \ : \ \|d\mathbf{F}(\mathbf{x}) - d\mathbf{F}(\mathbf{y})| \leqslant \gamma \|\mathbf{x} - \mathbf{y}\| \qquad \forall \mathbf{x}, \mathbf{y} \in \Omega$$

2. $d\mathbf{F}'(\mathbf{x})$ es regular y $\|d\mathbf{F}(\mathbf{x}^{-1})\| \geqslant \beta$
3. $\mathbf{F}(\mathbf{x}) = 0$ tiene una solución $\mathbf{x}^* \in \Omega$

Entonces para $\epsilon = \text{mín}\left\{r, \frac{1}{2\beta\gamma}\right\}$ y todo $\mathbf{x} \in B(\mathbf{x}^*, \epsilon)$ la sucesión generada por el algoritmo de Newton **converge hacia $\mathbf{x}^*$** y además lo hace con orden cuadrático, es decir:

$$\boxed{|\mathbf{x}_{n+1} - \mathbf{x}_n| \leqslant \beta\gamma |\mathbf{x}_n - \mathbf{x}^*|^2}$$

Método de Newton (varias variables)

Observaciones:

- Todas las observaciones reflejadas en el caso de una variable son, también aplicables ahora:
  - Se requiere una buena aproximación inicial $\mathbf{x}_0$
  - La convergencia para cualquier aproximación inicial se garantiza si $\mathbf{F}$ es estrictamente concava/convexa
  - La convergencia cuadrática degenera con raíces multiples.
- El método requiere resolver un sistema de ecuaciones lineales en cada paso (cálculo de $\mathbf{h}$).
- El método de Newton puede ser interpretado como un método de méjora. Donde la dirección de mejora se cálcula como:

$$\mathbf{d}_n = -d\mathbf{F}(\mathbf{x}_n)^{-1}\mathbf{F}(\mathbf{x}_n)$$

- Si el método de Newton es empleado para optimizar una función $f(\mathbf{x}) = 0$, se requieren segundas derivadas de f (Notar $\mathbf{F} = df$, $d\mathbf{F} = d^2 f$)
- Existen otros métodos que requieren menos regularidad (QuasiNewton, Secante, Gradiente), aunque son más *lentos*.
- Los métodos de Gradiente, se emplean para minimzar (resp. maximizar) una función $f : \Omega \to \mathbb{R}$, y se basan en la iteración:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \alpha_n \nabla f(\mathbf{x}_n)$$

Método de Newton (varias variables)

Observaciones:

- Todas las observaciones reflejadas en el caso de una variable son, también aplicables ahora:
- El método requiere resolver un sistema de ecuaciones lineales en cada paso (cálculo de $\mathbf{h}$).
- El método de Newton puede ser interpretado como un método de méjora. Donde la dirección de mejora se cálcula como:

$$\mathbf{d}_n = -d\mathbf{F}(\mathbf{x}_n)^{-1}\mathbf{F}(\mathbf{x}_n)$$

- Si el método de Newton es empleado para optimizar una función $f(\mathbf{x}) = 0$, se requieren segundas derivadas de f (Notar $\mathbf{F} = df$, $d\mathbf{F} = d^2 f$)
- Existen otros métodos que requieren menos regularidad (QuasiNewton, Secante, Gradiente), aunque son más *lentos*.
- Los métodos de Gradiente, se emplean para minimzar (resp. maximizar) una función $f : \Omega \to \mathbb{R}$, y se basan en la iteración:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \rho_n \nabla f(\mathbf{x}_n)$$

donde $\rho_n \in \mathbb{R}^n$ es el paso, y $\mathbf{d} = -\nabla f(\mathbf{x}_n) \in \mathbb{R}^d$ es el gradiente de f y es la dirección de máximo descenso local (Probar)

Método de Newton (varias variables)

Observaciones:

- Todas las observaciones reflejadas en el caso de una variable son, también aplicables ahora:
- El método requiere resolver un sistema de ecuaciones lineales en cada paso (cálculo de $\mathbf{h}$).
- El método de Newton puede ser interpretado como un método de méjora. Donde la dirección de mejora se cálcula como:

$$\mathbf{d}_n = -d\mathbf{F}(\mathbf{x}_n)^{-1}\mathbf{F}(\mathbf{x}_n)$$

- Si el método de Newton es empleado para optimizar una función $f(\mathbf{x}) = 0$, se requieren segundas derivadas de f (Notar $\mathbf{F} = df$, $d\mathbf{F} = d^2 f$)
- Existen otros métodos que requieren menos regularidad (QuasiNewton, Secante, Gradiente), aunque son más *lentos*.
- Los métodos de Gradiente, se emplean para minimzar (resp. maximizar) una función $f : \Omega \to \mathbb{R}$, y se basan en la iteración:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \rho_n \nabla f(\mathbf{x}_n)$$

donde $\rho_n \in \mathbb{R}^n$ es el paso, y $\mathbf{d} = -\nabla f(\mathbf{x}_n) \in \mathbb{R}^d$ es el gradiente de f y es la dirección de máximo descenso local (Probar)

Método de Newton (varias variables)

Observaciones:

- Todas las observaciones reflejadas en el caso de una variable son, también aplicables ahora:
- El método requiere resolver un sistema de ecuaciones lineales en cada paso (cálculo de $\mathbf{h}$).
- El método de Newton puede ser interpretado como un método de méjora. Donde la dirección de mejora se cálcula como:

$$\mathbf{d}_n = -d\mathbf{F}(\mathbf{x}_n)^{-1}\mathbf{F}(\mathbf{x}_n)$$

- Si el método de Newton es empleado para optimizar una función $f(\mathbf{x}) = 0$, se requieren segundas derivadas de f (Notar $\mathbf{F} = df$, $d\mathbf{F} = d^2 f$)
- Existen otros métodos que requieren menos regularidad (QuasiNewton, Secante, Gradiente), aunque son más *lentos*.
- Los métodos de Gradiente, se emplean para minimzar (resp. maximizar) una función $f : \Omega \to \mathbb{R}$, y se basan en la iteración:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \rho_n \nabla f(\mathbf{x}_n)$$

donde $\rho_n \in \mathbb{R}^n$ es el paso, y $\mathbf{d} = -\nabla f(\mathbf{x}_n) \in \mathbb{R}^d$ es el gradiente de f y es la dirección de máximo descenso local (Probar)

Método de Newton (varias variables)

Observaciones:

- Todas las observaciones reflejadas en el caso de una variable son, también aplicables ahora:
- El método requiere resolver un sistema de ecuaciones lineales en cada paso (cálculo de $\mathbf{h}$).
- El método de Newton puede ser interpretado como un método de méjora. Donde la dirección de mejora se cálcula como:

$$\mathbf{d}_n = -d\mathbf{F}(\mathbf{x}_n)^{-1}\mathbf{F}(\mathbf{x}_n)$$

- Si el método de Newton es empleado para optimizar una función $f(\mathbf{x}) = 0$, se requieren segundas derivadas de f (Notar $\mathbf{F} = df$, $d\mathbf{F} = d^2 f$)
- Existen otros métodos que requieren menos regularidad (QuasiNewton, Secante, Gradiente), aunque son más *lentos*.
- Los métodos de Gradiente, se emplean para minimzar (resp. maximizar) una función $f : \Omega \to \mathbb{R}$, y se basan en la iteración:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \rho_n \nabla f(\mathbf{x}_n)$$

donde $\rho_n \in \mathbb{R}^n$ es el paso, y $\mathbf{d} = -\nabla f(\mathbf{x}_n) \in \mathbb{R}^d$ es el gradiente de f y es la dirección de máximo descenso local (Probar)

Método de Newton (varias variables)

Observaciones:

- Todas las observaciones reflejadas en el caso de una variable son, también aplicables ahora:
- El método requiere resolver un sistema de ecuaciones lineales en cada paso (cálculo de $\mathbf{h}$).
- El método de Newton puede ser interpretado como un método de méjora. Donde la dirección de mejora se cálcula como:

$$\mathbf{d}_n = -d\mathbf{F}(\mathbf{x}_n)^{-1}\mathbf{F}(\mathbf{x}_n)$$

- Si el método de Newton es empleado para optimizar una función $f(\mathbf{x}) = 0$, se requieren segundas derivadas de f (Notar $\mathbf{F} = df$, $d\mathbf{F} = d^2 f$)
- Existen otros métodos que requieren menos regularidad (QuasiNewton, Secante, Gradiente), aunque son más *lentos*.
- Los métodos de Gradiente, se emplean para minimzar (resp. maximizar) una función $f : \Omega \to \mathbb{R}$, y se basan en la iteración:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \rho_n \nabla f(\mathbf{x}_n)$$

donde $\rho_n \in \mathbb{R}^n$ es el paso, y $\mathbf{d} = -\nabla f(\mathbf{x}_n) \in \mathbb{R}^d$ es el gradiente de f y es la dirección de máximo descenso local (Probar)

Tema 4

Programación Cuadrática

Introducción/Motivación

Preliminares

Elección de cartera. Restricciones de Igualdad

Elección de cartera. Restricciones de Desigualdad

Elección de cartera. Maximización de utilizad y línea del mercado de capitales

Ejercicio

Optimización con restricciones de igualdad(1)

Problema modelo:

$$\begin{array}{ll} \text{mín} & \frac{1}{2}\mathbf{x}'V\mathbf{x} \\ \text{s.a} & \mathbf{x}'\mathbf{1} = 1, \\ & \mathbf{x}'\mu = \bar{\mu} \end{array}$$

El lagrangiano asociado a este problema es:

$$L(\mathbf{x}, \lambda, \mu) := \frac{1}{2}\mathbf{x}'V\mathbf{x} - \lambda(\mathbf{x}'\mathbf{1} - 1) - \delta(\mathbf{x}'\mu - \bar{\mu})$$

En el óptimo se deben verificar las condiciones
Teorema de Lagrange :

$$\begin{array}{ll} (i) & \nabla_{\mathbf{x}} L = V\mathbf{x} - \lambda\mathbf{1} - \delta\mu = 0 \\ (ii) & \mathbf{x}'\mathbf{1} - 1 = 0 \\ (iii) & \mathbf{x}\mu - \bar{\mu} = 0 \end{array}$$

Optimización con restricciones de igualdad(2)

Desde (i) y dado que λ y δ son reales:

$$\mathbf{x}^* = \lambda V^{-1}\mathbf{1} + \delta V^{-1}\boldsymbol{\mu} \tag{2}$$

Calculando la traspuesta de la expresión anterior, multiplicándola por $\mathbf{1}$ y usando (ii), y el hecho de que V es simétrica se obtiene:

$$\begin{aligned}\mathbf{x}^{*\prime}\mathbf{1} = 1 &= (\lambda V^{-1}\mathbf{1} + \delta V^{-1}\boldsymbol{\mu})'\mathbf{1} \\ &= \lambda \mathbf{1}' V^{-1}\mathbf{1} + \delta \boldsymbol{\mu}' V^{-1}\mathbf{1}\end{aligned}$$

Repitiendo la misma operación con la segunda restricción:

$$\begin{aligned}\mathbf{x}^{*\prime}\boldsymbol{\mu} = \bar{\mu} &= (\lambda V^{-1}\mathbf{1} + \delta V^{-1}\boldsymbol{\mu})'\boldsymbol{\mu} \\ &= \lambda \mathbf{1}' V^{-1}\boldsymbol{\mu} + \delta \boldsymbol{\mu}' V^{-1}\boldsymbol{\mu}\end{aligned}$$

Optimización con restricciones de igualdad(3)

Introduciendo la notación, $A, B, C \in \mathbb{R}$:

$$A = \mathbf{1}'V^{-1}\mathbf{1}, \qquad B = \mathbf{1}'V^{-1}\boldsymbol{\mu}, \qquad C = \boldsymbol{\mu}'V^{-1}\boldsymbol{\mu}$$

El sistema anterior puede escribirse en forma matricial como:

$$\underbrace{\begin{pmatrix} A & B \\ B & C \end{pmatrix}}_{M} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 1 \\ \bar{\mu} \end{pmatrix}$$

Nota: Es posible demostrar que si V es definida positiva entonces M es invertible.

Despejando obtenemos:

$$\begin{pmatrix} \lambda \\ \delta \end{pmatrix} = M^{-1} \begin{pmatrix} 1 \\ \bar{\mu} \end{pmatrix} = \frac{1}{|M|} \begin{pmatrix} C & -B \\ -B & A \end{pmatrix} \begin{pmatrix} 1 \\ \bar{\mu} \end{pmatrix} \Rightarrow \boxed{\begin{array}{l} \lambda = \frac{C - \bar{\mu}B}{|M|} \\ \delta = \frac{\bar{\mu}A - B}{|M|} \end{array}}$$

Optimización con restricciones de igualdad(4)

Susitituyendo el valor de los multiplicadores en (2), obtenemos el vector de pesos:

$$\begin{aligned}\mathbf{x}^* &= \frac{C-\bar{\mu}B}{|M|}V^{-1}\mathbf{1} + \frac{\bar{\mu}A-B}{|M|}V^{-1}\boldsymbol{\mu} \\ &= \underbrace{\frac{1}{|M|}\left(CV^{-1}\mathbf{1}-BV^{-1}\boldsymbol{\mu}\right)}_{\Lambda_1} + \underbrace{\frac{1}{|M|}\left(AV^{-1}\boldsymbol{\mu}-BV^{-1}\mathbf{1}\right)}_{\Lambda_2}\bar{\mu} \qquad (3) \\ &= \boxed{\Lambda_1+\Lambda_2\bar{\mu}}\end{aligned}$$

y la varianza de la cartera viene dada por la expresión:

$$\begin{aligned}\sigma_C^2 &= \mathbf{x}^{*\prime}V\mathbf{x} =^{(2)} \mathbf{x}^{*\prime}V(\lambda V^{-1}\mathbf{1}+\delta V^{-1}\boldsymbol{\mu}) \\ &= \lambda\mathbf{x}^{*\prime}VV^{-1}\mathbf{1}+\delta\mathbf{x}^{*\prime}VV^{-1}\boldsymbol{\mu} = \lambda\mathbf{x}^{*\prime}\mathbf{1}+\delta\mathbf{x}^{*\prime}\boldsymbol{\mu} =^{(ii),(iii)} \lambda+\delta\bar{\mu} \\ &= \frac{1}{|M|}\left(C-\bar{\mu}B+\bar{\mu}^2A-\bar{\mu}B\right) = \boxed{\tfrac{1}{|M|}\left(\bar{\mu}^2A-2\bar{\mu}B+C\right)}\end{aligned}$$

Optimización con restricciones de igualdad(5)

Resolución con Matlab

1 Construir una función MATLAB que determine la composición de una cartera de mínimo riesgo para un determinado retorno si se admiten operaciones en corto (**Nota:** Usar los datos de la introducción):

Prototipo:

```
optimPorfolioI[returns, V, expectedReturn]
```

Función

```
function [x,sigma] = optimPortfolioI(returns,V,expectedReturn)

n = size(V,1);

invV = inv(V);

A = ones(1,n)*invV*ones(n,1);
B = ones(1,n)*invV*returns;
C = returns' *invV*returns;

M = [A B ; B C];

Lambda1 = 1/det(M)*(C*invV*ones(n,1) - B*invV*returns);
Lambda2 = 1/det(M)*(A*invV*returns - B*invV*ones(n,1));

x= Lambda1 + Lambda2*expectedReturn;

sigma2 = 1/det(M)*(expectedReturn^2*A - 2*expectedReturn*B + C);
sigma = sqrt(sigma2);

end
```

Ejemplo de Uso

```
>> returns=[0.1; 0.25 ; 0.33];
>>
>> V = [ 0.0050   -0.0060    0.0020
       -0.0060    0.0199   -0.0158
        0.0020   -0.0158    0.0790];
>>
>> [x,sigma] = optimPortfolioI(returns,V,0.2)

x =

    0.4140
    0.4347
    0.1513


sigma =

    0.0494
```

Optimización con restricciones de igualdad(6)

Resolución con matlab. Cálculo de la frontera eficiente

1. Construir una función MATLAB que determine frontera eficiente de una cartera de mínimo riesgo para un rango de retornos determinado si se admiten operaciones en corto:

Prototipo:

```
effPorfolioIFrontier[returns, V, range]
```

Función

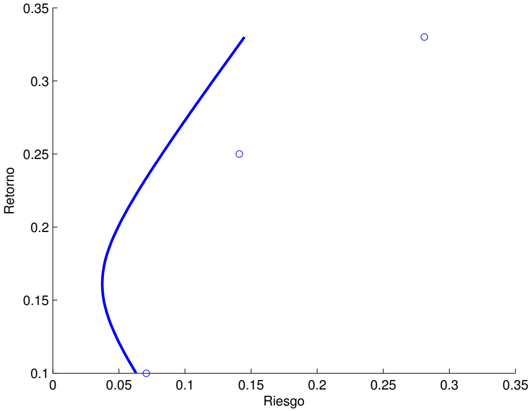
```
function [pX, pSigma] = effPortfolioIFrontier(returns,V, range)

expectedReturns = linspace(range(1), range(2), 100);

n       = size(V,1);
pX      = zeros(n,100);
pSigma  = zeros(1,100);

for i=1:100
    [pX(:,i),pSigma(i)] =...
        optimPorfolioI(returns,V, expectedReturns(i));
end

figure;
hold on;
plot(sqrt(diag(V)), returns, 'o', 'MarkerSize',5);
plot(pSigma, expectedReturns, 'LineWidth',2);
xlabel('Riesgo')
yLabel('Retorno')
```

Ejemplo de Uso

```
>> returns=[0.1; 0.25 ; 0.33];
>>
>> V = [ 0.0050   -0.0060    0.0020
-0.0060    0.0199   -0.0158
0.0020   -0.0158    0.0790];
>>
>> returns=[0.1; 0.25 ; 0.33];
>> V = [ 0.0050   -0.0060    0.0020
        -0.0060    0.0199   -0.0158
         0.0020   -0.0158    0.0790];
>>
>> effPortfolioIFrontier(returns, V, [0.1,0.33]);
```

Optimización con restricciones de igualdad(6)

Resolución con matlab. Cálculo de la frontera eficiente



Optimización con restricciones de igualdad(7)

Con las expresiones que se han obtenido para la varianza de la cartera, es sencillo determinar la cartera de varianza mínima:

$$\frac{d\sigma_C^2}{d\bar{\mu}} = \frac{d}{d\bar{\mu}}\left(\frac{1}{|M|}\left(\bar{\mu}^2 A - 2\bar{\mu}B + C\right)\right) = \frac{1}{|M|}\left(2\bar{\mu}A - 2B\right) = 0$$

$$\Rightarrow \quad \boxed{\bar{\mu} = \tfrac{B}{A}}$$

Sustituyendo obtenemos que la cartera de varianza mínima toma:

$$\mathbf{x}^* = \Lambda_1 + \Lambda_2 \frac{B}{A} \qquad \sigma^* = \sqrt{\frac{1}{|M|}\left(C - \frac{B^2}{A}\right)}$$

Observación: El resultado anterior tambien se puede obtener argumentando que en la cartera de varianza mínima se debe verificar $\delta = \frac{\bar{\mu}A - B}{|M|} = 0$, i.e. el precio sobra de la restricción asociada a $\bar{\mu}$ debe ser cero en el óptimo. (Probarlo)

Ejercicio: Construir una función Mathematica que calcule la cartera de varianza mínima. (En el ejemplo anterior la cartera de varianza mínima es $\mathbf{x}^* = (0.63, 0.30, 0.07) \quad \sigma_C^* = 0.037$

Tema 4

Programación Cuadrática

Introducción/Motivación

Preliminares

Elección de cartera. Restricciones de Igualdad

Elección de cartera. Restricciones de Desigualdad

Elección de cartera. Maximización de utilizad y línea del mercado de capitales

Ejercicio

Optimización restricciones de desigualdad(1)

Problema modelo:

$$\begin{array}{ll} \text{mín} & \frac{1}{2}\mathbf{x}'V\mathbf{x} \\ \text{s.a} & \mathbf{x}'\mathbf{1} = 1, \\ & \mathbf{x}'\mu = \bar{\mu}, \\ & \mathbf{x} \geqslant \mathbf{0} \end{array}$$

El lagrangiano asociado a este problema es:

$$L(\mathbf{x}, \lambda, \mu, \gamma) := \frac{1}{2}\mathbf{x}'V\mathbf{x} - \lambda(\mathbf{x}'\mathbf{1} - 1) - \delta(\mathbf{x}'\mu - \bar{\mu}) - \gamma^T\mathbf{x}$$

En el óptimo se deben verificar las condiciones ▸ Teorema de KKT:

$$\begin{array}{ll} (i) & V\mathbf{x} - \lambda\mathbf{1} - \delta\mu - \gamma = \mathbf{0} \\ (ii) & \mathbf{x}'\mathbf{1} - 1 = 0, \qquad \mathbf{x}\mu - \bar{\mu} = 0, \\ (iii) & \mathbf{x} \geqslant \mathbf{0}, \\ (iv) & \gamma \geqslant \mathbf{0}, \\ (v) & \gamma_i x_i = 0, \qquad i = 1, \ldots, n \end{array}$$

Optimización restricciones de desigualdad(2)

El problema anterior se puede reescribir del siguiente modo:

$$
\boxed{\begin{array}{l} \text{Hallar} \quad (\mathbf{x}, \mathbf{y}, \mathbf{z}) \quad \text{tal que:} \\ \qquad F(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 0, \\ \qquad\qquad\qquad \mathbf{z} \geqslant \mathbf{0} \end{array}}
$$

donde:

$$
F(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \begin{pmatrix} V\mathbf{x} - A'\mathbf{y} - \mathbf{z} \\ \mathbf{b} - A\mathbf{x} \\ \mathbf{x}'\mathbf{z} \end{pmatrix}, \quad A = \begin{pmatrix} \mathbf{1}' \\ \mu' \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ \bar{\mu} \end{pmatrix}
$$

$$
\mathbf{y} = \begin{pmatrix} \lambda \\ \mu \end{pmatrix}, \qquad \mathbf{z} = \gamma
$$

Optimización restricciones de desigualdad(3)

Algoritmo de punto interior

Hallar X tal que:

$$F(X) = 0,$$

con algún $X_i \geqslant 0$

1. Inicio: X^0, $\epsilon \geqslant 0$ (tolerancia), $\eta \approx 1$ (amortiguamiento).
2. Predicción: Iteración de Newton:

$$DF(X^n)\Delta X^{n,p} = -F(X^n)$$

3. Cálculo de parámetro de centrado (μ, σ):

$$X^n + \alpha \Delta X^{n,p} \geqslant \mathbf{0} \quad \Rightarrow \quad \alpha \rightarrow \mu,\ \sigma$$

4. Corrección: Iteración de Newton de segundo orden:

$$DF(X^n)\Delta X^{n,c} = -F(X^n, \mu) - \frac{1}{2}\Delta X^{n,p\,\prime} D^2 F(X_n)\Delta X^{n,p}$$

5. Dirección final

$$X^n + \alpha \Delta X^{n,c} \geqslant \mathbf{0} \quad \Rightarrow \quad X^{n+1} = X^n + \eta\alpha\Delta X^{n,c}$$

6. Evaluación del residuo

$$\text{Si} \quad \|F(X^{n+1})\| \geqslant \epsilon \qquad \text{volver a Paso 2}$$

Optimización restricciones de desigualdad(3)

En el nuestro caso el **Algorimto de punto interior** será:

Predicción Resolver:

$$\begin{bmatrix} V & -A' & -I \\ -A & 0 & 0 \\ \mathbb{Z} & 0 & \mathbb{X} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x}^p \\ \Delta \mathbf{y}^p \\ \Delta \mathbf{z}^p \end{bmatrix} = - \begin{bmatrix} V\mathbf{x} - A'\mathbf{y} - \mathbf{z} \\ \mathbf{b} - A\mathbf{x} \\ \mathbb{X}\mathbb{Z}\mathbf{1} \end{bmatrix}$$

donde $\mathbb{X}$, $\mathbb{Z}$ denotan matrices diagonales que en la fila iésima contienen a x_i, respectivamente z_i

Cálculo parámetro centrado Para las variables con restricción de positividad $(\mathbf{x}, \mathbf{z})$ calcular el máximo α que verifica:

$$\begin{aligned} \mathbf{x} + \alpha \Delta \mathbf{x}^p \geqslant \mathbf{0} \\ \mathbf{z} + \alpha \Delta \mathbf{z}^p \geqslant \mathbf{0} \end{aligned} \quad \Rightarrow \quad \alpha$$

El parámetro de centrado se calcula como:

$$\begin{aligned} \mu &= (\mathbf{x}'\mathbf{z})/n \\ \mu^a &= (\mathbf{x} + \alpha \Delta \mathbf{x}^p)'(\mathbf{z} + \alpha \Delta \mathbf{z}^p)/n \end{aligned} \quad \Rightarrow \quad \sigma = \left(\frac{\mu}{\mu^a} \right)^3$$

Optimización restricciones de desigualdad(4)

Corrección Resolver:

$$\begin{bmatrix} V & -A' & -I \\ -A & 0 & 0 \\ \mathbb{Z} & 0 & \mathbb{X} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{x}^c \\ \Delta\mathbf{y}^c \\ \Delta\mathbf{z}^c \end{bmatrix} = - \begin{bmatrix} V\mathbf{x} - A'\mathbf{y} - \mathbf{z} \\ \mathbf{b} - A\mathbf{x} \\ \mathbb{X}\mathbb{Z}\mathbf{1} + \Delta\mathbb{X}^c\Delta\mathbb{Z}^c\mathbf{1} - \sigma\mu\mathbf{1} \end{bmatrix}$$

donde $\Delta\mathbb{X}$, $\Delta\mathbb{Z}$ denotan matrices diagonales que en la fila iésima contienen a Δx_i, respectivamente Δz_i

Dirección final Calcular el máximo α que verifica:

$$\begin{array}{l} \mathbf{x} + \alpha\Delta\mathbf{x}^c \geqslant \mathbf{0} \\ \mathbf{z} + \alpha\Delta\mathbf{z}^c \geqslant \mathbf{0} \end{array} \quad \Rightarrow \quad \alpha$$

Actualizar $\mathbf{x}, \mathbf{y}, \mathbf{z}$ según $(\tau \approx 1, \tau < 1)$:

$$\mathbf{x} = \mathbf{x} + \tau\alpha\Delta\mathbf{x}^c$$
$$\mathbf{y} = \mathbf{y} + \tau\alpha\Delta\mathbf{y}^c$$
$$\mathbf{z} = \mathbf{z} + \tau\alpha\Delta\mathbf{z}^c$$

Optimización restricciones de desigualdad(5)

Cálculo del residuo

$$r = \|F(\mathbf{x}, \mathbf{y}, \mathbf{z})\| = \sqrt{\begin{array}{r}(V\mathbf{x} - A'\mathbf{y} - z)'(V\mathbf{x} - A'\mathbf{y} - z) \\ +(\mathbf{b} - A\mathbf{x})'(\mathbf{b} - A\mathbf{x}) + (\mathbf{x}'\mathbf{z})'(\mathbf{x}'\mathbf{z})\end{array}}$$

Si $r \geqslant \epsilon$ se inicia el nuevo el proceso.

El algoritmo términa cuando el residuo está por debajo de la tolerancia prefijada.

Optimización restricciones de desigualdad(6)

Resolución con Matlab

Para la resolución de problemas de optimización cuadráticos, Matlab incorpora la función `quadprog`

```
x = quadprog(H,f,A,b,Aeq,beq,lb,ub)
```

```
x = quadprog(H,f,A,b,Aeq,beq,lb,ub,x0, options)
```

halla una solución del problema

$$min_x \quad \frac{1}{2}\mathbf{x}'H\mathbf{x}+f' \text{ tal que } \begin{cases} A\mathbf{x} \leqslant \mathbf{b}, \\ A_{eq}\mathbf{x} = \mathbf{b}_{eq}, \\ \mathbf{lb} \leqslant \mathbf{x} \leqslant \mathbf{ub} \end{cases}$$

Optimización restricciones de desigualdad(7)

Resolución con Matlab

1. Construir una función MATLAB que determine la composición de una cartera de mínimo riesgo para un determinado retorno si no se admiten operaciones en corto:

   Prototipo:

   ```
   optimPorfolioII[returns, V, expectedReturn)
   ```

2. Construir una función MATLAB que determine frontera eficiente de una cartera de mínimo riesgo para un rango de retornos determinado si no se admiten operaciones en corto:

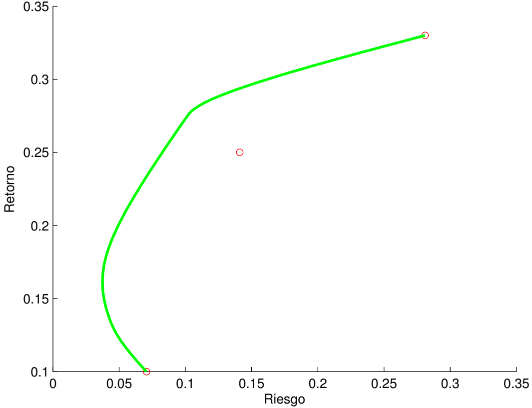
   Prototipo:

   ```
   effPorfolioIIFrontier[returns, V, range]
   ```

Nota: Utilizar los datos que figuran en la introducción.

Optimización restricciones de desigualdad(7)

Resolución con Matlab



Ejemplo. Construir una función Matlab que calcule la cartera de varianza mínima.

Programación Cuadrática

Introducción/Motivación

Preliminares

Elección de cartera. Restricciones de Igualdad

Elección de cartera. Restricciones de Desigualdad

Elección de cartera. Maximización de utilizad y línea del mercado de capitales

Ejercicio

Elección de cartera. Maximización de utilidad

Objetivo: Maximizar la utilidad de un inversor.

Hipótesis: Consideremos que el inversor tiene una función de utilidad media-varianza. Es decir, para un retorno r su función de utilidad viene dada por:

$$\boxed{u(r_C) = E(r_C) - \frac{1}{2}\gamma \mathrm{Var}(r_C) = \bar{\mu} - \frac{\gamma}{2}\sigma_C^2}$$

donde γ es el coeficiente de aversión al riesgo.

Si se admiten operaciones en corto, y el inversor es racional resulta:

$$u = \bar{\mu} - \frac{\gamma}{2}\frac{A\bar{\mu}^2 - 2\bar{\mu}B + C}{|M|}$$

Imponiendo la condición de primer orden:

$$\frac{du}{d\bar{\mu}} = 0 \Rightarrow 1 - \frac{\gamma}{|M|}(A\bar{\mu} - B) = 0 \Rightarrow \bar{\mu}^* = \frac{|M|}{\gamma A} + \frac{B}{A}$$

Elección de cartera. Maximización de utilidad

Objetivo: Maximizar la utilidad de un inversor.

Hipótesis: Consideremos que el inversor tiene una función de utilidad media-varianza. Es decir, para un retorno r su función de utilidad viene dada por:

$$\boxed{u(r_C) = E(r_C) - \frac{1}{2}\gamma \mathrm{Var}(r_C) = \bar{\mu} - \frac{\gamma}{2}\sigma_C^2}$$

donde γ es el coeficiente de aversión al riesgo.

Si se admiten operaciones en corto, y el inversor es racional resulta:

$$u = \bar{\mu} - \frac{\gamma}{2}\frac{A\bar{\mu}^2 - 2\bar{\mu}B + C}{|M|}$$

Imponiendo la condición de primer orden:

$$\frac{du}{d\bar{\mu}} = 0 \Rightarrow 1 - \frac{\gamma}{|M|}(A\bar{\mu} - B) = 0 \Rightarrow \bar{\mu}^* = \frac{|M|}{\gamma A} + \frac{B}{A}$$

Elección de cartera. Maximización de utilidad

Objetivo: Maximizar la utilidad de un inversor.

Hipótesis: Consideremos que el inversor tiene una función de utilidad media-varianza. Es decir, para un retorno r su función de utilidad viene dada por:

$$\boxed{u(r_C) = E(r_C) - \frac{1}{2}\gamma \mathrm{Var}(r_C) = \bar{\mu} - \frac{\gamma}{2}\sigma_C^2}$$

donde γ es el coeficiente de aversión al riesgo.

Si se admiten operaciones en corto, y el inversor es racional resulta:

$$u = \bar{\mu} - \frac{\gamma}{2}\frac{A\bar{\mu}^2 - 2\bar{\mu}B + C}{|M|}$$

Imponiendo la condición de primer orden:

$$\frac{du}{d\bar{\mu}} = 0 \Rightarrow 1 - \frac{\gamma}{|M|}(A\bar{\mu} - B) = 0 \Rightarrow \bar{\mu}^* = \frac{|M|}{\gamma A} + \frac{B}{A}$$

Elección de cartera. Maximización de utilidad

Objetivo: Maximizar la utilidad de un inversor.

Hipótesis: Consideremos que el inversor tiene una función de utilidad media-varianza. Es decir, para un retorno r su función de utilidad viene dada por:

$$\boxed{u(r_C) = E(r_C) - \frac{1}{2}\gamma \mathrm{Var}(r_C) = \bar{\mu} - \frac{\gamma}{2}\sigma_C^2}$$

donde γ es el coeficiente de aversión al riesgo.

Si se admiten operaciones en corto, y el inversor es racional resulta:

$$u = \bar{\mu} - \frac{\gamma}{2}\frac{A\bar{\mu}^2 - 2\bar{\mu}B + C}{|M|}$$

Imponiendo la condición de primer orden:

$$\frac{du}{d\bar{\mu}} = 0 \Rightarrow 1 - \frac{\gamma}{|M|}(A\bar{\mu} - B) = 0 \Rightarrow \bar{\mu}^* = \frac{|M|}{\gamma A} + \frac{B}{A}$$

Elección de cartera. Maximización de utilidad

Obsevación: Notar que en el óptimo, de acuerdo a la definición de δ, se tiene: $\boxed{\delta\gamma = 1}$

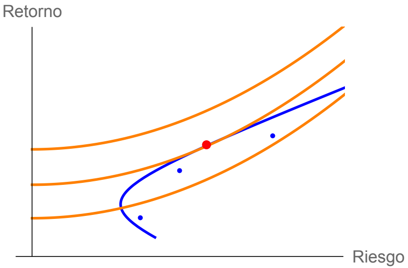
Ejemplos Maximización de la utilidad para dos niveles γ de aversión al riesgo

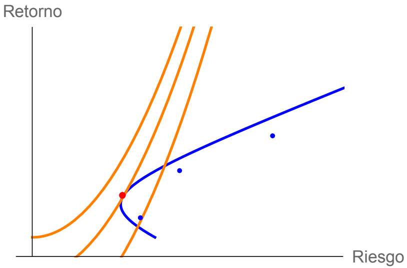
El inversor con mayor aversión al riesgo (derecha) selecciona una cartera con menor varianza.

Elección de cartera. Maximización de utilidad

Obsevación: Notar que en el óptimo, de acuerdo a la definición de δ, se tiene: $\boxed{\delta\gamma = 1}$

Ejemplos Maximización de la utilidad para dos niveles γ de aversión al riesgo





El inversor con mayor aversión al riesgo (derecha) selecciona una cartera con menor varianza.

Línea del mercado de capitales (1)

Objetivo: Consideremos ahora, que el inversor tiene la posibilidad de dividir su inversión en un activo libre de riesgo (C_0) y otro con riesgo (C_1). El último construido con los procedimientos estudiados. Estos activos viene caracterizados:

$$C_0 \Leftrightarrow (\mu_0, \sigma_0 = 0) \qquad C_1 \Leftrightarrow (\mu_1, \sigma_1)$$

¿Cuál es la distribución óptima?

La cartera resultante (C) queda determinada si se conoce:

- ω, proporción de la inversión en C_1.
- $\mathbf{x}$ pesos que determinan la composición del activo C_1.

Asumamos por el momento $\mu_0 = 0$, la cartera C esta caracterizada por:

$$\mu_C = (1-\omega)\mu_0 + \omega\mu_1 = \omega\mu_1, \qquad \sigma_C^2 = (1-\omega)^2\sigma_0^2 + \omega^2\sigma_1^2 = \omega^2\sigma_1^2$$

Línea del mercado de capitales (1)

Objetivo: Consideremos ahora, que el inversor tiene la posibilidad de dividir su inversión en un activo libre de riesgo (C_0) y otro con riesgo (C_1). El último construido con los procedimientos estudiados. Estos activos viene caracterizados:

$$C_0 \Leftrightarrow (\mu_0, \sigma_0 = 0) \qquad C_1 \Leftrightarrow (\mu_1, \sigma_1)$$

¿Cuál es la distribución óptima?

La cartera resultante (C) queda determinada si se conoce:

- ω, proporción de la inversión en C_1.
- $\mathbf{x}$ pesos que determinan la composición del activo C_1.

Asumamos por el momento $\mu_0 = 0$, la cartera C esta caracterizada por:

$$\mu_C = (1-\omega)\mu_0 + \omega\mu_1 = \omega\mu_1, \qquad \sigma_C^2 = (1-\omega)^2\sigma_0^2 + \omega^2\sigma_1^2 = \omega^2\sigma_1^2$$

Línea del mercado de capitales (1)

Objetivo: Consideremos ahora, que el inversor tiene la posibilidad de dividir su inversión en un activo libre de riesgo (C_0) y otro con riesgo (C_1). El último construido con los procedimientos estudiados. Estos activos viene caracterizados:

$$C_0 \Leftrightarrow (\mu_0, \sigma_0 = 0) \qquad C_1 \Leftrightarrow (\mu_1, \sigma_1)$$

¿Cuál es la distribución óptima?

La cartera resultante (C) queda determinada si se conoce:

- ω, proporción de la inversión en C_1.
- $\mathbf{x}$ pesos que determinan la composición del activo C_1.

Asumamos por el momento $\mu_0 = 0$, la cartera C esta caracterizada por:

$$\mu_C = (1-\omega)\mu_0 + \omega\mu_1 = \omega\mu_1, \qquad \sigma_C^2 = (1-\omega)^2\sigma_0^2 + \omega^2\sigma_1^2 = \omega^2\sigma_1^2$$

Línea de mercado de capitales (2)

Antes de resolver el problema, estudiaremos, cuál es el peso óptimo ω para un nivel de retorno μ_1 prefijado.

Si asumimos un modelo media-varianza para utilidad del inversor,i.e.:

$$u(r_C) = \omega\mu_1 - \frac{\gamma}{2}\omega^2\sigma_1^2(\mu_1)$$

Tras aplicar la condición de primer orden y un simple cáculo se obtiene:

$$\frac{du}{d\omega} = 0 \quad \Rightarrow \quad \omega^* = \frac{\mu_1}{\gamma\sigma^2(\mu_1)}$$

En la siguiente página interpretamos gráficamente este resultado.

Línea de mercado de capitales (2)

Antes de resolver el problema, estudiaremos, cuál es el peso óptimo ω para un nivel de retorno μ_1 prefijado.

Si asumimos un modelo media-varianza para utilidad del inversor,i.e.:

$$u(r_C) = \omega\mu_1 - \frac{\gamma}{2}\omega^2\sigma_1^2(\mu_1)$$

Tras aplicar la condición de primer orden y un simple cáculo se obtiene:

$$\frac{du}{d\omega} = 0 \quad \Rightarrow \quad \omega^* = \frac{\mu_1}{\gamma\sigma^2(\mu_1)}$$

En la siguiente página interpretamos gráficamente este resultado.

Línea de mercado de capitales (2)

Antes de resolver el problema, estudiaremos, cuál es el peso óptimo ω para un nivel de retorno μ_1 prefijado.

Si asumimos un modelo media-varianza para utilidad del inversor,i.e.:

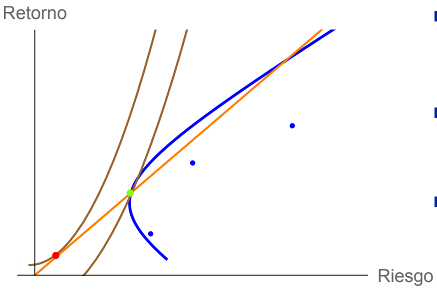
$$u(r_C) = \omega\mu_1 - \frac{\gamma}{2}\omega^2\sigma_1^2(\mu_1)$$

Tras aplicar la condición de primer orden y un simple cáculo se obtiene:

$$\frac{du}{d\omega} = 0 \quad \Rightarrow \quad \omega^* = \frac{\mu_1}{\gamma\sigma^2(\mu_1)}$$

En la siguiente página interpretamos gráficamente este resultado.

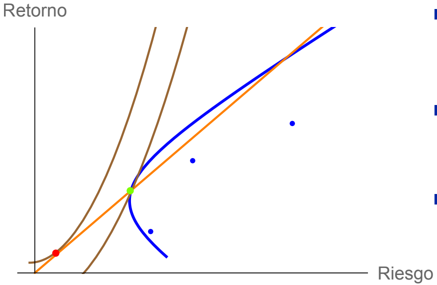
Línea del mercado de capitales (3)



Observaciones

- Sobre la recta se sitúan las carteras que combinan C_0 y C_1
- El punto rojo representa una cartera para un peso $\omega \in (0, 1)$ para un nivel de aversión γ.
- El punto verde con coordenadas $(\sigma_1(\mu_1), \mu_1)$ representa la cartera para $\omega = 1$.
- Del gráfico se deduce que la elección no es óptima. Modificando localmente μ_1 es posible obtener carteras con mayor rentabilidad y el mismo nivel de riesgo.

Línea del mercado de capitales (3)

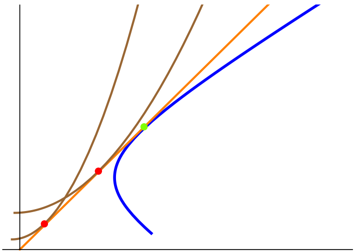


Observaciones

- Sobre la recta se sitúan las carteras que combinan C_0 y C_1
- El punto rojo representa una cartera para un peso $\omega \in (0, 1)$ para un nivel de aversión γ.
- El punto verde con coordenadas $(\sigma_1(\mu_1), \mu_1)$ representa la cartera para $\omega = 1$.
- Del gráfico se deduce que la elección no es óptima. Modificando localmente μ_1 es posible obtener carteras con mayor rentabilidad y el mismo nivel de riesgo.

¿Cuál es la situación óptima?

Línea del mercado de capitales (3)



Observaciones

- Sobre la recta se sitúan las carteras que combinan C_0 y C_1
- El punto rojo representa una cartera para un peso $\omega \in (0, 1)$ para un nivel de aversión γ.
- El punto verde con coordenadas $(\sigma_1(\mu_1), \mu_1)$ representa la cartera para $\omega = 1$.
- Del gráfico se deduce que la elección no es óptima. Modificando localmente μ_1 es posible obtener carteras con mayor rentabilidad y el mismo nivel de riesgo.

¿Cuál es la situación óptima?

$\Rightarrow$ La recta debe ser tangente a la frontera eficiente

Sólo existe una cartera (cartera de mercado). El nivel de aversión fija la posición del inversión en la recta (recta del mercado de capitales).

Línea del mercado de capitales (4)

Matématicamente, la cartera que está sobre el punto de tangencia está caracterizada por (¿Por qué?):

$$\lambda = 0 \quad \Rightarrow \quad \frac{C - \mu_m^* B}{|M|} = 0 \quad \Rightarrow \quad \mu_m^* = \frac{C}{B}$$

Utilizando está relación resulta, que la cartera del mercado está caracterizada:

$$\mathbf{x}_m^* = \frac{1}{B} V^{-1} \mu, \qquad \mu_m^* = \frac{C}{B} \qquad (\sigma_m^*)^2 = \frac{1}{B^2} \mu' V^{-1} \mu$$

El nivel de aversión al riesgo asociado a esta cartera, que coincide con la pendiente de la línea del mercado de capitales es:

$$\gamma = B$$

Ejercicio 1: Probar las equivalencias anteriores.
Ejercicio 2: Rehacer los cálculos si la cartera de riesgo cero tiene asociado un rendimiento $\mu_0 \neq 0$.

Línea del mercado de capitales (4)

Matématicamente, la cartera que está sobre el punto de tangencia está caracterizada por (¿Por qué?):

$$\lambda = 0 \quad \Rightarrow \quad \frac{C - \mu_m^* B}{|M|} = 0 \quad \Rightarrow \quad \mu_m^* = \frac{C}{B}$$

Utilizando está relación resulta, que la cartera del mercado está caracterizada:

$$\boxed{\mathbf{x}_m^* = \frac{1}{B} V^{-1} \mu, \qquad \mu_m^* = \frac{C}{B} \qquad (\sigma_m^*)^2 = \frac{1}{B^2} \mu' V^{-1} \mu}$$

El nivel de aversión al riesgo asociado a esta cartera, que coincide con la pendiente de la línea del mercado de capitales es:

$$\gamma = B$$

Ejercicio 1: Probar las equivalencias anteriores.
Ejercicio 2: Rehacer los cálculos si la cartera de riesgo cero tiene asociado un rendimiento $\mu_0 \neq 0$.

Línea del mercado de capitales (4)

Matématicamente, la cartera que está sobre el punto de tangencia está caracterizada por (¿Por qué?):

$$\lambda = 0 \quad \Rightarrow \quad \frac{C - \mu_m^* B}{|M|} = 0 \quad \Rightarrow \quad \mu_m^* = \frac{C}{B}$$

Utilizando está relación resulta, que la cartera del mercado está caracterizada:

$$\boxed{\mathbf{x}_m^* = \frac{1}{B} V^{-1} \mu, \qquad \mu_m^* = \frac{C}{B} \qquad (\sigma_m^*)^2 = \frac{1}{B^2} \mu' V^{-1} \mu}$$

El nivel de aversión al riesgo asociado a esta cartera, que coincide con la pendiente de la línea del mercado de capitales es:

$$\gamma = B$$

Ejercicio 1: Probar las equivalencias anteriores.
Ejercicio 2: Rehacer los cálculos si la cartera de riesgo cero tiene asociado un rendimiento $\mu_0 \neq 0$.

Línea del mercado de capitales (4)

Matématicamente, la cartera que está sobre el punto de tangencia está caracterizada por (¿Por qué?):

$$\lambda = 0 \quad \Rightarrow \quad \frac{C - \mu_m^* B}{|M|} = 0 \quad \Rightarrow \quad \mu_m^* = \frac{C}{B}$$

Utilizando está relación resulta, que la cartera del mercado está caracterizada:

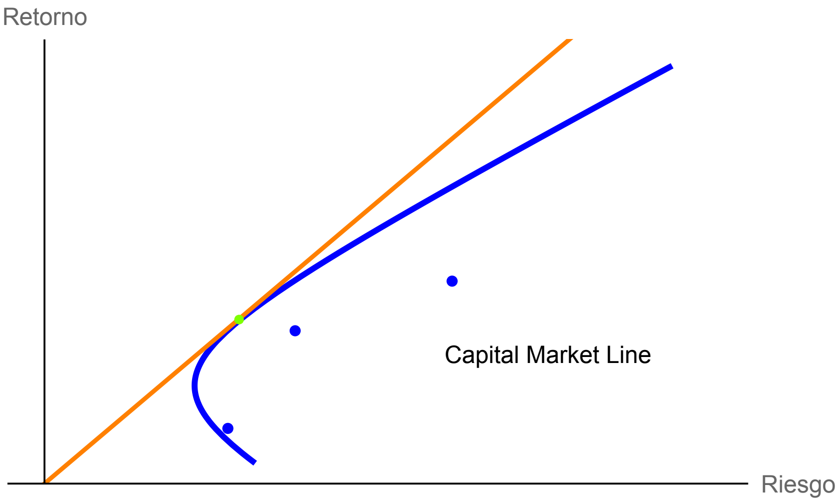
$$\boxed{\mathbf{x}_m^* = \frac{1}{B} V^{-1} \boldsymbol{\mu}, \qquad \mu_m^* = \frac{C}{B} \qquad (\sigma_m^*)^2 = \frac{1}{B^2} \boldsymbol{\mu}' V^{-1} \boldsymbol{\mu}}$$

El nivel de aversión al riesgo asociado a esta cartera, que coincide con la pendiente de la línea del mercado de capitales es:

$$\gamma = B$$

Ejercicio 1: Probar las equivalencias anteriores.
Ejercicio 2: Rehacer los cálculos si la cartera de riesgo cero tiene asociado un rendimiento $\mu_0 \neq 0$.

Línea del mercado de capitales (4)



Tema 4

Programación Cuadrática

Introducción/Motivación

Preliminares

Elección de cartera. Restricciones de Igualdad

Elección de cartera. Restricciones de Desigualdad

Elección de cartera. Maximización de utilizad y línea del mercado de capitales

Ejercicio

Ejercicio

Construir una cartera formada por al menos 5 activos y calcular:

- Frontera eficiente, cartera de menor varianza, cartera de mercado y línea del mercado de capitales si se admiten operaciones en corto.
- Frontera eficiente, cartera de menor varianza, cartera de mercado y línea del mercado de capitales si no se admiten operaciones en corto.

Nota.

- Para la resolución del ejercicio anterior se pueden emplear las funciones matlab a que se han presentado durante el capítulo.
- La descarga de la información (retornos, matriz de varianzas) se hara desde https://es.finance.yahoo.com/ o bien usando las herramientas de Mathematica.
- El cálculo de la cartera de menor varianza, cartera de mercado y línea de mercado, para el caso en el que no se admiten operaciones en corto, se debe realizar numéricamente. La estrategia correspondiente debe ser desarrollada por el alumno.

Tema 5

Optimización Dinámica

Introducción/Motivación

Control óptimo en tiempo discreto

Control óptimo en tiempo continuo

Tema 5

Optimización Dinámica

Introducción/Motivación

Control óptimo en tiempo discreto

Control óptimo en tiempo continuo

Introducción: Optimización Dinámica

Sistema dinámico $\Leftrightarrow$ Proceso que evoluciona en tiempo

Objetivo (Opt. Dinámica): Guiar o controlar de manera óptima un sistema dinámico a lo largo del horizonte temporal.

En un sistema dinámico se distinguen:

- Variables de estado (x).
- Variables de control (u).
- Ecuaciones de transición ($f(x, u)$).
- Función objetivo ($J, F(x, u), S(x)$),

Ejemplos: cuerpo humano, viaje en un coche, economía de un país.

Introducción: Ejemplo 1

Un consumidor tiene una función de utilidad $U(c(t))$ donde $c(t)$ representa el consumo. Se supone $U' > 0$ y $U'' < 0$. La persona recibe una dotación inicial de K_0 stock de capital. Los ingresos de la persona vienen dados por iK donde i es el interés. Además la persona puede consumir el stock de capital en cualquier momento. Suponer que el consumidor es impaciente siendo δ su tasa de descuento. Se debe cumplir además $K(t) \geqslant 0$.

El problema puede ser modelado

$$\max_{c(t)} \int_0^T e^{-\delta t} U(c(t))dt$$

s.a.

$$\begin{aligned} & K'(t) = iK(t) - c(t) && t \in (0, T) \\ & K(0) = K_0, \quad K(t) \geqslant 0 && t \in (0, T) \\ & c(t) \geqslant 0 && t \in (0, T) \end{aligned}$$

Ejercicio: Identificar variable de estado, variable de control, ecuación de estado y función objetivo.

Otros Ejemplos: CERDÁ Ejemplo 4.1 (pág. 111), CERDÁ Ejemplo 4.2 (Modelo neoclásido de crecimiento) (pág. 112)

Introducción: Ejemplo 1

Un consumidor tiene una función de utilidad $U(c(t))$ donde $c(t)$ representa el consumo. Se supone $U' > 0$ y $U'' < 0$. La persona recibe una dotación inicial de K_0 stock de capital. Los ingresos de la persona vienen dados por iK donde i es el interés. Además la persona puede consumir el stock de capital en cualquier momento. Suponer que el consumidor es impaciente siendo δ su tasa de descuento. Se debe cumplir además $K(t) \geqslant 0$.

El problema puede ser modelado

$$\max_{c(t)} \quad \int_0^T e^{-\delta t} U(c(t))dt$$

s.a.

$$\begin{array}{ll} K'(t) = iK(t) - c(t) & t \in (0,T) \\ K(0) = K_0, \quad K(t) \geqslant 0 & t \in (0,T) \\ c(t) \geqslant 0 & t \in (0,T) \end{array}$$

Ejercicio: Identificar variable de estado, variable de control, ecuación de estado y función objetivo.

Otros Ejemplos: CERDÁ Ejemplo 4.1 (pág. 111), CERDÁ Ejemplo 4.2 (Modelo neoclásido de crecimiento) (pág. 112)

Introducción: Ejemplo 2

Una mina tiene unas reservas iniciales $R_0 = 250$ y puede estar activa durante 20 anõs. En $t = 20$, la mina será expropiada. El beneficio neto en cada periodo viene dado por:

$$\Pi_k = \left[p - \frac{cq_k}{R_k}\right] q_k$$

donde q_k es la cantidad de mineral extraída, R_k es el stock de mineral en la mina al comienzo del periodo, p es el precio unitario, y c el coste. Se supone un factor de descuento β. Determinar la cantidad de mineral a extraer para maximizar el valor presente de los beneficios descontados.

El problema puede ser modelado:

$$\begin{array}{lll} \max\limits_{\{q_k\}_{k=0}^{19}} & \sum\limits_{t=0}^{19} \beta^k \Pi_k & \\ \text{s.a.} & R_{k+1} = R_k - q_k & k = 0, \dots, 19 \\ & R_0 = 250, \quad R_k \geqslant 0 & k = 0, \dots, 19 \end{array}$$

Ejercicio: Identificar variable de estado, variable de control, ecuación de transición y función objetivo.

Otros Ejemplos: CERDÁ sección 6.3 (pág. 225)

Introducción: Ejemplo 2

Una mina tiene unas reservas iniciales $R_0 = 250$ y puede estar activa durante 20 anõs. En $t = 20$, la mina será expropiada. El beneficio neto en cada periodo viene dado por:

$$\Pi_k = \left[p - \frac{cq_k}{R_k}\right] q_k$$

donde q_k es la cantidad de mineral extraída, R_k es el stock de mineral en la mina al comienzo del periodo, p es el precio unitario, y c el coste. Se supone un factor de descuento β. Determinar la cantidad de mineral a extraer para maximizar el valor presente de los beneficios descontados.

El problema puede ser modelado:

$$\begin{array}{lll} \max\limits_{\{q_k\}_{k=0}^{19}} & \displaystyle\sum_{t=0}^{19} \beta^k \Pi_k & \\ \text{s.a.} & R_{k+1} = R_k - q_k & k = 0, \ldots, 19 \\ & R_0 = 250, \quad R_k \geqslant 0 & k = 0, \ldots, 19 \end{array}$$

Ejercicio: Identificar variable de estado, variable de control, ecuación de transición y función objetivo.

Otros Ejemplos: CERDÁ sección 6.3 (pág. 225)

Introducción

Control Tiempo Continuo

$$\max_{u(t)} \quad J = \int_{t_0}^{t_1} F(x,u,t)dt + S(x(t_1))$$

s.a.

$$\dot{x}(t) = f(x,u,t)$$

$$x(t_0) = x_0$$

$$u(t) \in \Omega(t)$$

Control Tiempo Discreto

$$\max_{\{u_k\}_{k=0}^{N-1}} \quad J = \sum_{k=0}^{N-1} F(x_k,u_k,k) + S(x_N)$$

s.a.

$$x_{k+1} = f(x_k,u_k,k)$$

$$x_0 = \alpha_0$$

$$u_k \in \Omega_k$$

Ejercicio: Identificar variable de estado, variable de control, ecuación de transición y función objetivo.

Tema 5

Optimización Dinámica

Introducción/Motivación

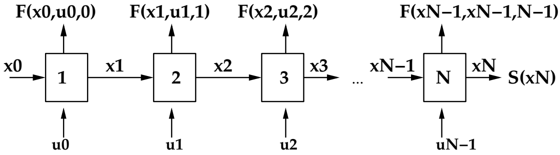
Control óptimo en tiempo discreto

- Programación dinámica
- Principio de optimalidad de Bellman
- Ejemplo
- Resolución con Matlab

Control óptimo en tiempo continuo

Planteamiento de un problema de control óptimo en tiempo discreto

- N: número de etapas
- u_k: variable de control en la etapa $k \in \{0, 1, \ldots, N-1\}$
- x_k: variable de estado en la etapa $k \in \{0, 1, \ldots, N-1\}$
- $x_{k+1} = f(x_k, u_k, k)$: ec. de estado $k \in \{0, 1, \ldots, N-1\}$
- $J = \sum_{k=0}^{N-1} F(x_k, u_k, k) + S(x_N)$: funcional objetivo



Planteamiento de un problema de control óptimo en tiempo discreto

Problema modelo: Encontrar $u_k \in \Omega_k$ para $k = 0, 1, \ldots, N-1$ que maximiza (ó minimiza) el funcional:

$$\max_{\{u_k\}_{k=0}^{N-1}} \quad J = \sum_{k=0}^{N-1} F(x_k, u_k, k) + S(x_N)$$

s.a.

$$\begin{aligned} x_{k+1} &= f(x_k, u_k, k) & \qquad k &= 0, 1, \ldots, N-1 \\ x_0 &= \alpha_0 & & \\ u_k &\in \Omega_k & k &= 0, 1, \ldots, N-1 \end{aligned}$$

Definiciónes Se denomina **control óptimo** a la secuencia de controles $\{u_k^*\}_{k=0}^{N-1}$ que resuelve el problema. Se denomina **trayectoria óptima** a la secuencia $\{x_k^*\}_{k=0}^{N}$ determinada por la ecuación de estado y $\{u_k^*\}_{k=0}^{N-1}$. El **valor óptimo** del funcional se denota J^*.

Planteamiento de un problema de control óptimo en tiempo discreto

Problema modelo: Encontrar $u_k \in \Omega_k$ para $k = 0, 1, \ldots, N-1$ que maximiza (ó minimiza) el funcional:

$$\max_{\{u_k\}_{k=0}^{N-1}} \quad J = \sum_{k=0}^{N-1} F(x_k, u_k, k) + S(x_N)$$

s.a.

$$\begin{aligned} & x_{k+1} = f(x_k, u_k, k) && k = 0, 1, \ldots, N-1 \\ & x_0 = \alpha_0 && \\ & u_k \in \Omega_k && k = 0, 1, \ldots, N-1 \end{aligned}$$

Definiciónes Se denomina **control óptimo** a la secuencia

de controles $\{u_k^*\}_{k=0}^{N-1}$ que resuelve el problema. Se denomina **trayectoria óptima** a la secuencia $\{x_k^*\}_{k=0}^{N}$ determinada por la ecuación de estado y $\{u_k^*\}_{k=0}^{N-1}$. El **valor óptimo** del funcional se denota J^*.

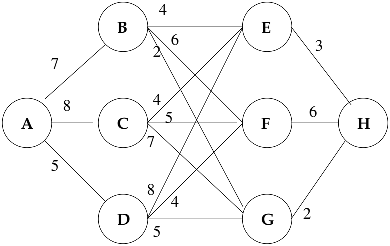
Programación dinámica

- Es un procedimiento para la resolución de un problema de control óptimo en tiempo discreto.
- Fue introducida por Bellman (1957).
- Resuelve el problema de N etapas mediante la resolución de N problemas de una etapa.

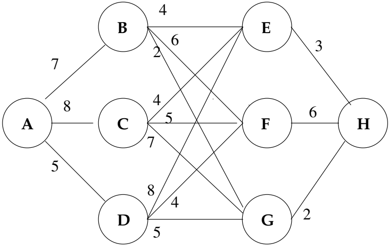
Programación dinámica

- Es un procedimiento para la resolución de un problema de control óptimo en tiempo discreto.
- Fue introducida por Bellman (1957).
- Resuelve el problema de N etapas mediante la resolución de N problemas de una etapa.

Ilustraremos el procedimiento con un ejemplo:
Determinar la ruta mínima entre las ciudades A y H.



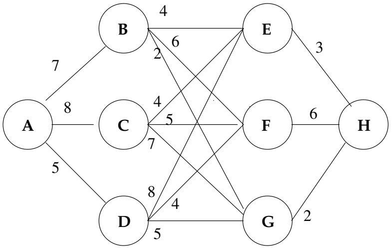
Ejemplo: Ruta mínima



- Descomponemos el problema en tres etapas:

  1. A $\rightarrow$ B, C ó D   2. B,C ó D $\rightarrow$ E, F ó G   3. E, F ó G $\rightarrow$ H

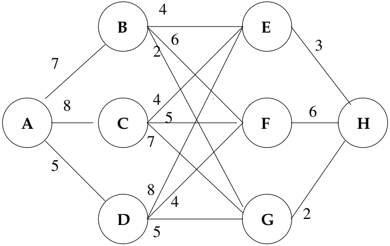
Ejemplo: Ruta mínima



- **Etapa 3**: E, F ó G $\rightarrow$ H

Ciudad de partida	Ir a:	Distancia a H
E	H	3
F	H	6
G	H	2

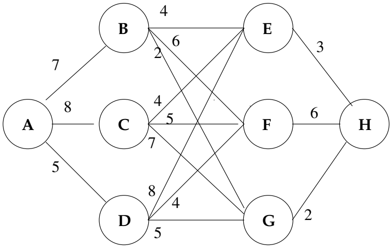
Ejemplo: Ruta mínima



- **Etapa 2**: B, C ó D $\rightarrow$ E, F ó G.

Ciudad de partida	Ir a:	Distancia a H
B	E	4+3 =7
	F	6+6 =12
	G	2+2 = 4
C	E	4+3 = 7
	F	5 +6 =11
	G	7 +2 = 9
D	E	8 +3 =11
	F	4 +6 =10
	G	5 +2 =7

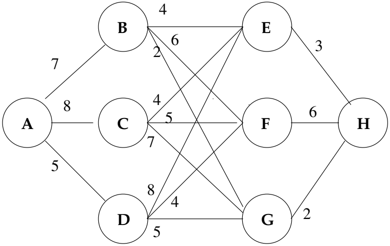
Ejemplo: Ruta mínima



- **Etapa 2**: B, C ó D → E, F ó G.

Ciudad de partida	Ir a:	Distancia a H
B	E	4+3 =7
	F	6+6 =12
	G	2+2 = 4
C	E	4+3 = 7
	F	5 +6 =11
	G	7 +2 = 9
D	E	8 +3 =11
	F	4 +6 =10
	G	5 +2 =7

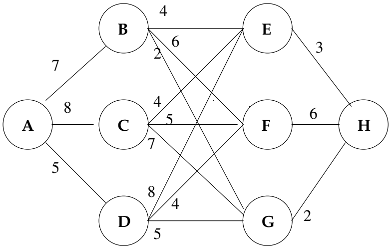
Ejemplo: Ruta mínima



- **Etapa 2**: B, C ó D $\rightarrow$ E, F ó G. Información relevante

Ciudad de partida	Ir a:	Distancia a H
B	G	4
C	E	7
D	G	7

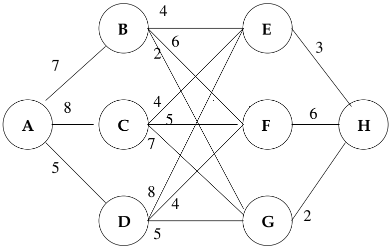
Ejemplo: Ruta mínima



- **Etapa 1**: A → B, C ó D

Ciudad de partida	Ir a:	Distancia a H
A	B	7 + 4 = 11
A	C	8 + 7 = 15
A	D	7 + 5 = 12

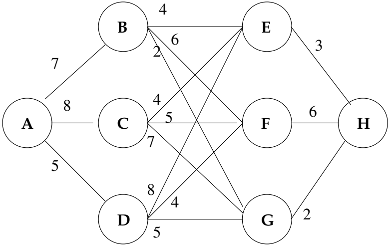
Ejemplo: Ruta mínima



- **Etapa 1**: A $\rightarrow$ B, C ó D

Ciudad de partida	Ir a:	Distancia a H
A	B	7 + 4 = 11
A	C	8 + 7 = 15
A	D	7 + 5 = 12

Ejemplo: Ruta mínima

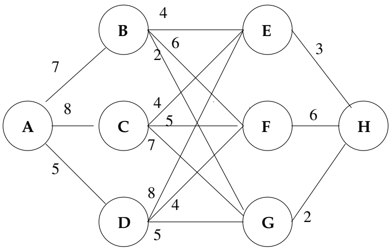


- Recuperación de la ruta

$$A \to B \to G \to H$$

Distancia : 11

Ejemplo: Ruta mínima



Ejercicios

1. Resolver el problema estudiando todas las rutas posibles y compararlo con la alternativa estudiada.
2. Interpretar el problema anterior como un problema de control óptimo en tiempo discreto, determinado variables de estados, variables de control, funciones de transición y función objetivo.

Causalidad

Hipótesis de Causalidad. Para cualesquiera $j,\ r \in \{0, 1, \ldots, N-1\}$ con $j < r$ se verifica que x_r depende únicamente de x_j y de los controles $\{u_j, u_{j+1}, \ldots, u_{r-1}\}$

Nota.

- Esta hipótesis no siempre se cumple en los modelos económicos. Una decisión política futura puede influir en la actitud presente de los agentes.
- Una consecuencia de esta hipótesis es:

$$\left.\begin{array}{l} u_0, u_1, \ldots, u_{N-1} \\ \alpha_0 \end{array}\right\} \quad \Rightarrow \quad \{x_0, x_1, \ldots, x_N\}$$

- Bajo esta hipótesis el valor óptimo del funcional sólo depende de la condición inicial:

$$\boxed{J^* = J^*(\alpha_0)}$$

Ecuación de Bellman

Consideremos la función $J_k^*\{x_k\}$ (función valor) definido por recurrencia:

- Para $k = N$,

$$\boxed{J_N^*\{x_N\} = S(x_N)}$$

- Para $k = \{N-1, N-2, \ldots, 0\}$:

$$\boxed{J_k^*\{x_k\} = \max_{u_k \in \Omega_k} \{F(x_k, u_k, k) + J_{k+1}^*\{(f(x_k, u_k, k)\}}$$

Proposición. Se verifica:

- $J^*(\alpha_0) = J_0^*\{\alpha_0\}$
- Si $u^*[0, N-1] = \{u_0^*, u_1^*, \ldots, u_{N-1}^*\}$ es la secuencia que maximiza las ecuaciones de Bellman, entonces es el control óptimo

Nota. La proposición anterior proporciona un método para resolver el problema de N etapas a través de la resolución de N problemas de una etapa.

Ecuación de Bellman

Consideremos la función $J_k^*\{x_k\}$ (función valor) definido por recurrencia:

- Para $k = N$,

$$\boxed{J_N^*\{x_N\} = S(x_N)}$$

- Para $k = \{N-1, N-2, \ldots, 0\}$:

$$\boxed{J_k^*\{x_k\} = \max_{u_k \in \Omega_k} \{F(x_k, u_k, k) + J_{k+1}^*\{(f(x_k, u_k, k)\}}$$

Proposición. Se verifica:

- $J^*(\alpha_0) = J_0^*\{\alpha_0\}$
- Si $u^*[0, N-1] = \{u_0^*, u_1^*, \ldots, u_{N-1}^*\}$ es la secuencia que maximiza las ecuaciones de Bellman, entonces es el control óptimo

Nota. La proposición anterior proporciona un método para resolver el problema de N etapas a través de la resolución de N problemas de una etapa.

Ecuación de Bellman

Consideremos la función $J_k^*\{x_k\}$ (función valor) definido por recurrencia:

- Para $k = N$,

$$\boxed{J_N^*\{x_N\} = S(x_N)}$$

- Para $k = \{N-1, N-2, \ldots, 0\}$:

$$\boxed{J_k^*\{x_k\} = \max_{u_k \in \Omega_k}\{F(x_k, u_k, k) + J_{k+1}^*\{(f(x_k, u_k, k)\}}$$

Proposición. Se verifica:

- $J^*(\alpha_0) = J_0^*\{\alpha_0\}$
- Si $u^*[0, N-1] = \{u_0^*, u_1^*, \ldots, u_{N-1}^*\}$ es la secuencia que maximiza las ecuaciones de Bellman, entonces es el control óptimo

Nota. La proposición anterior proporciona un método para resolver el problema de N etapas a través de la resolución de N problemas de una etapa.

Principio de optimalidad de Bellman

Sean

$$u^*[0,N-1]=\{u_0^*,u_1^*,\dots,u_{N-1}^*\},\quad x^*[0,N-1]=\{x_0^*,x_1^*,\dots,x_N^*\}$$

el control óptimo y la trayectoria óptima, el **subproblema**:

$$\begin{array}{lll}
\max\limits_{\{u_k\}_{k=j}^{N-1}} & \displaystyle\sum_{k=j}^{N-1} F(x_k,u_k,k)dt+S(x_N) & \\
\text{s.a.} & x_{k+1}=f(x_k,u_k,k) & k=j,\dots,N-1\\
 & x_j=x_j^* & \\
 & u_k\in\Omega_k & k=j,\dots,N-1
\end{array}$$

tiene por control óptimo $u^*[j,N-1]=\{u_j^*,\dots,u_{N-1}^*\}$

Observación: El principio anterior se puede resumir en: dada una secuencia óptima de decisiones, toda subsecuencia de ella es, a su vez, óptima

Ejemplo

Un inversor dispone de §5 m. y está considerando invertir su capital en tres compañias C_1, C_2 y C_3 (en m.). Dispone de una estimación del retorno de su inversión reflejada en la siguiente tabla:

	$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$	$x = 5$
C_1	0	1	4	6	7	8
C_2	0	3	5	8	10	11
C_3	0	0	2	10	11	11

Cómo debería distribuir su inversión para obtener el retorno máximo. Resolver el problema utilizando programación dinámica.

Identificación de variables

- Etapas: $N = 3$ (Inversión en cada compañía)
- Variable de Estados: $x_0 = 5$, $x_k \in \{0, 1, 2, 3, 4, 5\}$ (Cantidad por invertir)
- Variable de Control: $u_k = \{0, 1, 2, 3, 4, 5\}$ (Cantidad invertida en cada compañía)
- Función de transición: $f(x_k, u_k, k) = x_k - u_k$
- Función objetivo:
  - $F(x_k, u_k, k) =$ Posición $(k, u_k + 1)$ de la tabla.
  - $S(x_3) = 0$

Ejemplo

Un inversor dispone de §5 m. y está considerando invertir su capital en tres compañias C_1, C_2 y C_3 (en m.). Dispone de una estimación del retorno de su inversión reflejada en la siguiente tabla:

	$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$	$x = 5$
C_1	0	1	4	6	7	8
C_2	0	3	5	8	10	11
C_3	0	0	2	10	11	11

Cómo debería distribuir su inversión para obtener el retorno máximo. Resolver el problema utilizando programación dinámica.

Identificación de variables

- Etapas: $N = 3$ (Inversión en cada compañia)
- Variable de Estados: $x_0 = 5$, $x_k \in \{0, 1, 2, 3, 4, 5\}$ (Cantidad por invertir)
- Variable de Control: $u_k = \{0, 1, 2, 3, 4, 5\}$ (Cantidad invertida en cada compañia)
- Función de transición: $f(x_k, u_k, k) = x_k - u_k$
- Función objetivo:
  - $F(x_k, u_k, k) =$ Posición $(k, u_k + 1)$ de la tabla.
  - $S(x_3) = 0$

Ejemplo. Función valor J_3^*

x_3	$J_3^*\{x_3\}$
0	0
1	0
2	0
3	0
4	0
5	0

$x_3 = 0$: $J_3^*\{0\} = S(0) = 0.$

$x_3 = 1$: $J_3^*\{1\} = S(1) = 0.$

$x_3 = 2$: $J_3^*\{2\} = S(2) = 0.$

$x_3 = 3$: $J_3^*\{3\} = S(3) = 0.$

$x_3 = 4$: $J_3^*\{4\} = S(4) = 0.$

$x_3 = 5$: $J_3^*\{5\} = S(5) = 0.$

Ejemplo. Función valor J_2^*

$x_2 = 0$:

x_2	$J_2^*\{x_2\}$	u_2
0		
1		
2		
3		
4		
5		

$$\begin{aligned}
J_2^*\{0\} &= \max_{u_2 \in \{0\}} \{F(0, u_2, 2) + J_3^*\{f(0, u_2, 2)\}\} \\
&= \max\{F(0, 0, 2) + J_3^*\{f(0, 0, 2)\}\} \\
&= \max\{F(0, 0, 2) + J_3^*\{0\}\} \\
&= \max\{0 + 0\} \\
&= 0 \quad \Rightarrow \quad u_2 = 0
\end{aligned}$$

Ejemplo. Función valor J_2^*

$x_2 = 0$:

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1		
2		
3		
4		
5		

$$\begin{aligned} J_2^*\{0\} &= \max_{u_2 \in \{0\}} \{F(0, u_2, 2) + J_3^*\{f(0, u_2, 2)\}\} \\ &= \max\{F(0, 0, 2) + J_3^*\{f(0, 0, 2)\}\} \\ &= \max\{F(0, 0, 2) + J_3^*\{0\}\} \\ &= \max\{0 + 0\} \\ &= 0 \quad \Rightarrow \quad u_2 = 0 \end{aligned}$$

Ejemplo. Función valor J_2^*

$x_2 = 1$:

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1		
2		
3		
4		
5		

$$\begin{aligned}
J_2^*\{1\} &= \max_{u_2 \in \{0,1\}} \{F(1, u_2, 2) + J_3^*\{f(1, u_2, 2)\}\} \\
&= \max\{F(1,0,2) + J_3^*\{f(1,0,2)\}, \\
&\qquad\quad F(1,1,2) + J_3^*\{f(1,1,2)\}\} \\
&= \max\{F(1,0,2) + J_3^*\{1\}, F(1,1,2) + J_3^*\{0\}\} \\
&= \max\{0+0, 0+0\} \\
&= 0 \quad \Rightarrow \quad u_2 = \{0, 1\}
\end{aligned}$$

Ejemplo. Función valor J_2^*

$x_2 = 1$:

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1	0	$\{0, 1\}$
2		
3		
4		
5		

$$
\begin{aligned}
J_2^*\{1\} &= \max_{u_2 \in \{0,1\}} \{F(1, u_2, 2) + J_3^*\{f(1, u_2, 2)\}\} \\
&= \max\{F(1,0,2) + J_3^*\{f(1,0,2)\}, \\
&\qquad F(1,1,2) + J_3^*\{f(1,1,2)\}\} \\
&= \max\{F(1,0,2) + J_3^*\{1\}, F(1,1,2) + J_3^*\{0\}\} \\
&= \max\{0+0, 0+0\} \\
&= 0 \quad \Rightarrow \quad u_2 = \{0, 1\}
\end{aligned}
$$

Ejemplo. Función valor J_2^*

$x_2 = 2$:

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1	0	$\{0,1\}$
2		
3		
4		
5		

$$
\begin{aligned}
J_2^*\{2\} &= \max_{u_2 \in \{0,1,2\}} \{F(2,u_2,2) + J_3^*\{f(2,u_2,2)\}\} \\
&= \text{máx}\{F(2,0,2) + J_3^*\{f(2,0,2)\}, \\
&\qquad F(2,1,2) + J_3^*\{f(2,1,2)\}, \\
&\qquad F(2,2,2) + J_3^*\{f(2,2,2)\}\} \\
&= \text{máx}\{F(2,0,2) + J_3^*\{2\}, \\
&\qquad F(2,1,2) + J_3^*\{1\}, \\
&\qquad F(2,2,2) + J_3^*\{0\}\} \\
&= \text{máx}\{0+0, 0+0, 2+0\} \\
&= 2 \quad \Rightarrow \quad u_2 = 2
\end{aligned}
$$

Ejemplo. Función valor J_2^*

$x_2 = 2$:

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1	0	$\{0,1\}$
2	2	2
3		
4		
5		

$$
\begin{aligned}
J_2^*\{2\} &= \max_{u_2\in\{0,1,2\}}\{F(2,u_2,2)+J_3^*\{f(2,u_2,2)\}\} \\
&= \text{máx}\{F(2,0,2)+J_3^*\{f(2,0,2)\}, \\
&\qquad F(2,1,2)+J_3^*\{f(2,1,2)\}, \\
&\qquad F(2,2,2)+J_3^*\{f(2,2,2)\}\} \\
&= \text{máx}\{F(2,0,2)+J_3^*\{2\}, \\
&\qquad F(2,1,2)+J_3^*\{1\}, \\
&\qquad F(2,2,2)+J_3^*\{0\}\} \\
&= \text{máx}\{0+0, 0+0, 2+0\} \\
&= 2 \quad \Rightarrow \quad u_2 = 2
\end{aligned}
$$

Ejemplo. Función valor J_2^*

$x_2 = 3$:

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1	0	$\{0, 1\}$
2	2	2
3		
4		
5		

$$
\begin{aligned}
J_2^*\{3\} &= \max_{u_2 \in \{0,1,2,3\}} \{F(3, u_2, 2) + J_3^*\{f(3, u_2, 2)\}\} \\
&= \text{máx}\{F(3, 0, 2) + J_3^*\{f(3, 0, 2)\}, \\
&\qquad F(3, 1, 2) + J_3^*\{f(3, 1, 2)\}, \\
&\qquad F(3, 2, 2) + J_3^*\{f(3, 2, 2)\}, \\
&\qquad F(3, 3, 2) + J_3^*\{f(3, 3, 2)\}\} \\
&= \text{máx}\{F(3, 0, 2) + J_3^*\{3\}, F(3, 1, 2) + J_3^*\{2\}, \\
&\qquad F(3, 2, 2) + J_3^*\{1\}, F(3, 3, 2) + J_3^*\{0\}\} \\
&= \text{máx}\{0 + 0, 0 + 0, 2 + 0, 10 + 0\} \\
&= 10 \quad \Rightarrow \quad u_2 = 3
\end{aligned}
$$

Ejemplo. Función valor J_2^*

$x_2 = 3$:

$$
\begin{aligned}
J_2^*\{3\} &= \max_{u_2\in\{0,1,2,3\}} \{F(3,u_2,2) + J_3^*\{f(3,u_2,2)\}\} \\
&= \max\{F(3,0,2) + J_3^*\{f(3,0,2)\}, \\
&\qquad F(3,1,2) + J_3^*\{f(3,1,2)\}, \\
&\qquad F(3,2,2) + J_3^*\{f(3,2,2)\}, \\
&\qquad F(3,3,2) + J_3^*\{f(3,3,2)\}\} \\
&= \max\{F(3,0,2) + J_3^*\{3\}, F(3,1,2) + J_3^*\{2\}, \\
&\qquad F(3,2,2) + J_3^*\{1\}, F(3,3,2) + J_3^*\{0\}\} \\
&= \max\{0+0, 0+0, 2+0, 10+0\} \\
&= 10 \quad \Rightarrow \quad u_2 = 3
\end{aligned}
$$

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1	0	$\{0,1\}$
2	2	2
3	10	3
4		
5		

Ejemplo. Función valor J_2^*

$x_2 = 4$:

$$
\begin{aligned}
J_2^*\{4\} &= \max_{u_2 \in \{0,1,\ldots,4\}} \{F(4,u_2,2) + J_3^*\{f(4,u_2,2)\}\} \\
&= \text{máx}\{F(4,0,2) + J_3^*\{f(4,0,2)\}, \\
&\qquad F(4,1,2) + J_3^*\{f(4,1,2)\}, \\
&\qquad F(4,2,2) + J_3^*\{f(4,2,2)\}, \\
&\qquad F(4,3,2) + J_3^*\{f(4,3,2)\}, \\
&\qquad F(4,4,2) + J_3^*\{f(4,4,2)\}\} \\
&= \text{máx}\{F(4,0,2) + J_3^*\{4\}, F(4,1,2) + J_3^*\{3\}, \\
&\qquad F(4,2,2) + J_3^*\{2\}, F(4,3,2) + J_3^*\{1\}, \\
&\qquad F(4,4,2) + J_3^*\{0\}\} \\
&= \text{máx}\{0+0, 0+0, 2+0, 10+0, \\
&\qquad 11+0\} \\
&= 11 \quad \Rightarrow \quad u_2 = 4
\end{aligned}
$$

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1	0	$\{0,1\}$
2	2	2
3	10	3
4		
5		

Ejemplo. Función valor J_2^*

$x_2 = 4$:

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1	0	$\{0, 1\}$
2	2	2
3	10	3
4	11	4
5		

$$
\begin{aligned}
J_2^*\{4\} &= \max_{u_2\in\{0,1,\ldots,4\}} \{F(4,u_2,2) + J_3^*\{f(4,u_2,2)\}\} \\
&= \text{máx}\{F(4,0,2) + J_3^*\{f(4,0,2)\}, \\
&\qquad F(4,1,2) + J_3^*\{f(4,1,2)\}, \\
&\qquad F(4,2,2) + J_3^*\{f(4,2,2)\}, \\
&\qquad F(4,3,2) + J_3^*\{f(4,3,2)\}, \\
&\qquad F(4,4,2) + J_3^*\{f(4,4,2)\}\} \\
&= \text{máx}\{F(4,0,2) + J_3^*\{4\}, F(4,1,2) + J_3^*\{3\}, \\
&\qquad F(4,2,2) + J_3^*\{2\}, F(4,3,2) + J_3^*\{1\}, \\
&\qquad F(4,4,2) + J_3^*\{0\}\} \\
&= \text{máx}\{0+0, 0+0, 2+0, 10+0, \\
&\qquad 11+0\} \\
&= 11 \quad \Rightarrow \quad u_2 = 4
\end{aligned}
$$

Ejemplo. Función valor J_2^*

$x_2 = 5$:

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1	0	$\{0,1\}$
2	2	2
3	10	3
4	11	4
5		

$$
\begin{aligned}
J_2^*\{5\} &= \max_{u_2\in\{0,1,\ldots,5\}} \{F(5,u_2,2) + J_3^*\{f(5,u_2,2)\}\} \\
&= \max\{F(5,0,2) + J_3^*\{f(5,0,2)\}, \\
&\qquad F(5,1,2) + J_3^*\{f(5,1,2)\}, \\
&\qquad F(5,2,2) + J_3^*\{f(5,2,2)\}, \\
&\qquad F(5,3,2) + J_3^*\{f(5,3,2)\}, \\
&\qquad F(5,4,2) + J_3^*\{f(5,4,2)\}, \\
&\qquad F(5,5,2) + J_3^*\{f(5,5,2)\}\} \\
&= \max\{F(5,0,2) + J_3^*\{5\}, F(5,1,2) + J_3^*\{4\}, \\
&\qquad F(5,2,2) + J_3^*\{3\}, F(5,3,2) + J_3^*\{2\}, \\
&\qquad F(5,4,2) + J_3^*\{1\}, F(5,5,2) + J_3^*\{0\}\} \\
&= \max\{0+0, 0+0, 2+0, 10+0, \\
&\qquad 11+0, 11+0\} \\
&= 11 \quad \Rightarrow \quad u_2 = \{4,5\}
\end{aligned}
$$

Ejemplo. Función valor J_2^*

$x_2 = 5$:

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1	0	$\{0,1\}$
2	2	2
3	10	3
4	11	4
5	11	$\{4,5\}$

$$
\begin{aligned}
J_2^*\{5\} &= \max_{u_2\in\{0,1,\ldots,5\}}\{F(5,u_2,2)+J_3^*\{f(5,u_2,2)\}\}\\
&= \max\{F(5,0,2)+J_3^*\{f(5,0,2)\},\\
&\qquad F(5,1,2)+J_3^*\{f(5,1,2)\},\\
&\qquad F(5,2,2)+J_3^*\{f(5,2,2)\},\\
&\qquad F(5,3,2)+J_3^*\{f(5,3,2)\},\\
&\qquad F(5,4,2)+J_3^*\{f(5,4,2)\},\\
&\qquad F(5,5,2)+J_3^*\{f(5,5,2)\}\}\\
&= \max\{F(5,0,2)+J_3^*\{5\},F(5,1,2)+J_3^*\{4\},\\
&\qquad F(5,2,2)+J_3^*\{3\},F(5,3,2)+J_3^*\{2\},\\
&\qquad F(5,4,2)+J_3^*\{1\},F(5,5,2)+J_3^*\{0\}\}\\
&= \max\{0+0,0+0,2+0,10+0,\\
&\qquad 11+0,11+0\}\\
&= 11 \quad\Rightarrow\quad u_2=\{4,5\}
\end{aligned}
$$

Ejemplo. Función valor J_1^*

$x_1 = 0$:

x_1	$J_1^*\{x_1\}$	u_1
0		
1		
2		
3		
4		
5		

$$\begin{aligned} J_1^*\{0\} &= \max_{u_1 \in \{0\}} \{F(0, u_1, 2) + J_2^*\{f(0, u_1, 1)\}\} \\ &= \max\{F(0, 0, 1) + J_2^*\{f(0, 0, 1)\}\} \\ &= \max\{F(0, 0, 1) + J_2^*\{0\}\} \\ &= \max\{0 + 0\} \\ &= 0 \quad \Rightarrow \quad u_2 = 0 \end{aligned}$$

Ejemplo. Función valor J_1^*

$\boxed{x_1 = 0}$:

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1		
2		
3		
4		
5		

$$
\begin{aligned}
J_1^*\{0\} &= \max_{u_1 \in \{0\}} \{F(0, u_1, 2) + J_2^*\{f(0, u_1, 1)\}\} \\
&= \max\{F(0,0,1) + J_2^*\{f(0,0,1)\}\} \\
&= \max\{F(0,0,1) + J_2^*\{0\}\} \\
&= \max\{0 + 0\} \\
&= 0 \quad \Rightarrow \quad u_2 = 0
\end{aligned}
$$

Ejemplo. Función valor J_1^*

$x_1 = 1$:

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1		
2		
3		
4		
5		

$$
\begin{aligned}
J_1^*\{1\} &= \max_{u_1 \in \{0,1\}} \{F(1, u_1, 1) + J_2^*\{f(1, u_1, 1)\}\} \\
&= \max\{F(1,0,1) + J_2^*\{f(1,0,1)\}, \\
&\qquad\quad F(1,1,1) + J_2^*\{f(1,1,1)\}\} \\
&= \max\{F(1,0,1) + J_2^*\{1\}, F(1,1,1) + J_2^*\{0\}\} \\
&= \max\{0 + 0, 3 + 0\} \\
&= 3 \quad \Rightarrow \quad u_1 = 1
\end{aligned}
$$

Ejemplo. Función valor J_1^*

$x_1 = 1$:

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1	3	1
2		
3		
4		
5		

$$
\begin{aligned}
J_1^*\{1\} &= \max_{u_1 \in \{0,1\}} \{F(1, u_1, 1) + J_2^*\{f(1, u_1, 1)\}\} \\
&= \max\{F(1, 0, 1) + J_2^*\{f(1, 0, 1)\}, \\
&\qquad F(1, 1, 1) + J_2^*\{f(1, 1, 1)\}\} \\
&= \max\{F(1, 0, 1) + J_2^*\{1\}, F(1, 1, 1) + J_2^*\{0\}\} \\
&= \max\{0 + 0, 3 + 0\} \\
&= 3 \quad \Rightarrow \quad u_1 = 1
\end{aligned}
$$

Ejemplo. Función valor J_1^*

$x_1 = 2$:

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1	3	1
2		
3		
4		
5		

$$
\begin{aligned}
J_1^*\{2\} &= \max_{u_1 \in \{0,1,2\}} \{F(2, u_1, 1) + J_2^*\{f(2, u_1, 1)\}\} \\
&= \text{máx}\{F(2,0,1) + J_2^*\{f(2,0,1)\}, \\
&\qquad F(2,1,1) + J_2^*\{f(2,1,1)\}, \\
&\qquad F(2,2,1) + J_2^*\{f(2,2,1)\}\} \\
&= \text{máx}\{F(2,0,1) + J_2^*\{2\}, \\
&\qquad F(2,1,1) + J_2^*\{1\}, \\
&\qquad F(2,2,1) + J_2^*\{0\}\} \\
&= \text{máx}\{5+0, 3+0, 0+2\} \\
&= 5 \quad \Rightarrow \quad u_1 = 2
\end{aligned}
$$

Ejemplo. Función valor J_1^*

$x_1 = 2$:

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1	3	1
2	5	2
3		
4		
5		

$$
\begin{aligned}
J_1^*\{2\} &= \max_{u_1 \in \{0,1,2\}} \{F(2, u_1, 1) + J_2^*\{f(2, u_1, 1)\}\} \\
&= \max\{F(2,0,1) + J_2^*\{f(2,0,1)\}, \\
&\qquad\quad F(2,1,1) + J_2^*\{f(2,1,1)\}, \\
&\qquad\quad F(2,2,1) + J_2^*\{f(2,2,1)\}\} \\
&= \max\{F(2,0,1) + J_2^*\{2\}, \\
&\qquad\quad F(2,1,1) + J_2^*\{1\}, \\
&\qquad\quad F(2,2,1) + J_2^*\{0\}\} \\
&= \max\{5+0, 3+0, 0+2\} \\
&= 5 \quad \Rightarrow \quad u_1 = 2
\end{aligned}
$$

Ejemplo. Función valor J_1^*

$x_1 = 3$:

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1	3	1
2	5	2
3		
4		
5		

$$
\begin{aligned}
J_1^*\{3\} &= \max_{u_1 \in \{0,1,2,3\}} \{F(3, u_1, 1) + J_2^*\{f(3, u_1, 1)\}\} \\
&= \text{máx}\{F(3,0,1) + J_2^*\{f(3,0,1)\}, \\
&\qquad F(3,1,1) + J_2^*\{f(3,1,1)\}, \\
&\qquad F(3,2,1) + J_2^*\{f(3,2,1)\}, \\
&\qquad F(3,3,1) + J_2^*\{f(3,3,1)\}\} \\
&= \text{máx}\{F(3,0,1) + J_2^*\{3\}, F(3,1,1) + J_2^*\{2\}, \\
&\qquad F(3,2,1) + J_2^*\{1\}, F(3,3,1) + J_2^*\{0\}\} \\
&= \text{máx}\{0 + 10, 3 + 2, 5 + 0, 8 + 0\} \\
&= 10 \quad \Rightarrow \quad u_1 = 0
\end{aligned}
$$

Ejemplo. Función valor J_1^*

$x_1 = 3$:

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1	3	1
2	5	2
3	10	0
4		
5		

$$
\begin{aligned}
J_1^*\{3\} &= \max_{u_1 \in \{0,1,2,3\}} \{F(3,u_1,1) + J_2^*\{f(3,u_1,1)\}\} \\
&= \max\{F(3,0,1) + J_2^*\{f(3,0,1)\}, \\
&\qquad F(3,1,1) + J_2^*\{f(3,1,1)\}, \\
&\qquad F(3,2,1) + J_2^*\{f(3,2,1)\}, \\
&\qquad F(3,3,1) + J_2^*\{f(3,3,1)\}\} \\
&= \max\{F(3,0,1) + J_2^*\{3\}, F(3,1,1) + J_2^*\{2\}, \\
&\qquad F(3,2,1) + J_2^*\{1\}, F(3,3,1) + J_2^*\{0\}\} \\
&= \max\{0+10, 3+2, 5+0, 8+0\} \\
&= 10 \quad \Rightarrow \quad u_1 = 0
\end{aligned}
$$

Ejemplo. Función valor J_1^*

$x_1 = 4$:

$$
\begin{aligned}
J_1^*\{4\} &= \max_{u_1 \in \{0,1,\dots,4\}} \{F(4, u_1, 1) + J_2^*\{f(4, u_1, 1)\}\} \\
&= \max\{F(4,0,1) + J_2^*\{f(4,0,1)\}, \\
&\qquad F(4,1,1) + J_2^*\{f(4,1,1)\}, \\
&\qquad F(4,2,1) + J_2^*\{f(4,2,1)\}, \\
&\qquad F(4,3,1) + J_2^*\{f(4,3,1)\}, \\
&\qquad F(4,4,1) + J_2^*\{f(4,4,1)\}\} \\
&= \max\{F(4,0,1) + J_2^*\{4\}, F(4,1,1) + J_2^*\{3\}, \\
&\qquad F(4,2,1) + J_2^*\{2\}, F(4,3,1) + J_2^*\{1\}, \\
&\qquad F(4,4,1) + J_2^*\{0\}\} \\
&= \max\{0 + 11, 3 + 10, 5 + 2, 8 + 0, \\
&\qquad 10 + 0\} \\
&= 13 \quad \Rightarrow \quad u_1 = 1
\end{aligned}
$$

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1	3	1
2	5	2
3	10	0
4		
5		

Ejemplo. Función valor J_1^*

$x_1 = 4$:

$$
\begin{aligned}
J_1^*\{4\} &= \max_{u_1 \in \{0,1,\ldots,4\}} \{F(4, u_1, 1) + J_2^*\{f(4, u_1, 1)\}\} \\
&= \max\{F(4,0,1) + J_2^*\{f(4,0,1)\}, \\
&\qquad F(4,1,1) + J_2^*\{f(4,1,1)\}, \\
&\qquad F(4,2,1) + J_2^*\{f(4,2,1)\}, \\
&\qquad F(4,3,1) + J_2^*\{f(4,3,1)\}, \\
&\qquad F(4,4,1) + J_2^*\{f(4,4,1)\}\} \\
&= \max\{F(4,0,1) + J_2^*\{4\}, F(4,1,1) + J_2^*\{3\}, \\
&\qquad F(4,2,1) + J_2^*\{2\}, F(4,3,1) + J_2^*\{1\}, \\
&\qquad F(4,4,1) + J_2^*\{0\}\} \\
&= \max\{0 + 11, 3 + 10, 5 + 2, 8 + 0, \\
&\qquad 10 + 0\} \\
&= 13 \quad \Rightarrow \quad u_1 = 1
\end{aligned}
$$

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1	3	1
2	5	2
3	10	0
4	13	1
5		

Ejemplo. Función valor J_1^*

$x_1 = 5$:

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1	3	1
2	5	2
3	10	0
4	13	1
5		

$$
\begin{aligned}
J_1^*\{5\} &= \max_{u_1 \in \{0,1,\ldots,5\}} \{F(5,u_1,1) + J_2^*\{f(5,u_1,1)\}\} \\
&= \text{máx}\{F(5,0,1) + J_2^*\{f(5,0,1)\}, \\
&\qquad F(5,1,1) + J_2^*\{f(5,1,1)\}, \\
&\qquad F(5,2,1) + J_2^*\{f(5,2,1)\}, \\
&\qquad F(5,3,1) + J_2^*\{f(5,3,1)\}, \\
&\qquad F(5,4,1) + J_2^*\{f(5,4,1)\}, \\
&\qquad F(5,5,1) + J_2^*\{f(5,5,1)\}\} \\
&= \text{máx}\{F(5,0,1) + J_2^*\{5\}, F(5,1,1) + J_2^*\{4\}, \\
&\qquad F(5,2,1) + J_2^*\{3\}, F(5,3,1) + J_2^*\{2\}, \\
&\qquad F(5,4,1) + J_2^*\{1\}, F(5,5,1) + J_2^*\{0\}\} \\
&= \text{máx}\{0+11, 3+11, 5+10, 8+2, \\
&\qquad 10+0, 11+0\} \\
&= 15 \quad \Rightarrow \quad u_1 = 2
\end{aligned}
$$

Ejemplo. Función valor J_1^*

$x_1 = 5$:

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1	3	1
2	5	2
3	10	0
4	13	1
5	15	2

$$
\begin{aligned}
J_1^*\{5\} &= \max_{u_1 \in \{0,1,\ldots,5\}} \{F(5,u_1,1) + J_2^*\{f(5,u_1,1)\}\} \\
&= \max\{F(5,0,1) + J_2^*\{f(5,0,1)\}, \\
&\qquad F(5,1,1) + J_2^*\{f(5,1,1)\}, \\
&\qquad F(5,2,1) + J_2^*\{f(5,2,1)\}, \\
&\qquad F(5,3,1) + J_2^*\{f(5,3,1)\}, \\
&\qquad F(5,4,1) + J_2^*\{f(5,4,1)\}, \\
&\qquad F(5,5,1) + J_2^*\{f(5,5,1)\}\} \\
&= \max\{F(5,0,1) + J_2^*\{5\}, F(5,1,1) + J_2^*\{4\}, \\
&\qquad F(5,2,1) + J_2^*\{3\}, F(5,3,1) + J_2^*\{2\}, \\
&\qquad F(5,4,1) + J_2^*\{1\}, F(5,5,1) + J_2^*\{0\}\} \\
&= \max\{0+11, 3+11, 5+10, 8+2, \\
&\qquad 10+0, 11+0\} \\
&= 15 \quad \Rightarrow \quad u_1 = 2
\end{aligned}
$$

Ejemplo. Función valor J_0^*

$x_0 = 5$:

$$
\begin{aligned}
J_0^*\{5\} &= \max_{u_0 \in \{0,1,\ldots,5\}} \{F(5, u_0, 0) + J_1^*\{f(5, u_0, 0)\}\} \\
&= \text{máx}\{F(5,0,0) + J_2^*\{f(5,0,1)\}, \\
&\qquad F(5,1,0) + J_1^*\{f(5,1,0)\}, \\
&\qquad F(5,2,0) + J_1^*\{f(5,2,0)\}, \\
&\qquad F(5,3,0) + J_1^*\{f(5,3,0)\}, \\
&\qquad F(5,4,0) + J_1^*\{f(5,4,0)\}, \\
&\qquad F(5,5,0) + J_1^*\{f(5,5,0)\}\} \\
&= \text{máx}\{F(5,0,0) + J_1^*\{5\}, F(5,1,0) + J_1^*\{4\}, \\
&\qquad F(5,2,0) + J_1^*\{3\}, F(5,3,0) + J_1^*\{2\}, \\
&\qquad F(5,4,0) + J_1^*\{1\}, F(5,5,0) + J_1^*\{0\}\} \\
&= \text{máx}\{8+0, 7+3, 6+4, 4+10, \\
&\qquad 1+13, 0+15\} \\
&= 15 \quad \Rightarrow \quad u_0 = 0
\end{aligned}
$$

x_0	$J_0^*\{x_0\}$	u_0
5		

Ejemplo. Función valor J_0^*

$x_0 = 5$:

$$
\begin{aligned}
J_0^*\{5\} &= \max_{u_0 \in \{0,1,\ldots,5\}} \{F(5,u_0,0) + J_1^*\{f(5,u_0,0)\}\} \\
&= \max\{F(5,0,0) + J_2^*\{f(5,0,1)\}, \\
&\qquad F(5,1,0) + J_1^*\{f(5,1,0)\}, \\
&\qquad F(5,2,0) + J_1^*\{f(5,2,0)\}, \\
&\qquad F(5,3,0) + J_1^*\{f(5,3,0)\}, \\
&\qquad F(5,4,0) + J_1^*\{f(5,4,0)\}, \\
&\qquad F(5,5,0) + J_1^*\{f(5,5,0)\}\} \\
&= \max\{F(5,0,0) + J_1^*\{5\}, F(5,1,0) + J_1^*\{4\}, \\
&\qquad F(5,2,0) + J_1^*\{3\}, F(5,3,0) + J_1^*\{2\}, \\
&\qquad F(5,4,0) + J_1^*\{1\}, F(5,5,0) + J_1^*\{0\}\} \\
&= \max\{8+0, 7+3, 6+4, 4+10, \\
&\qquad 1+13, 0+15\} \\
&= 15 \quad \Rightarrow \quad u_0 = 0
\end{aligned}
$$

x_0	$J_0^*\{x_0\}$	u_0
5	15	0

Ejemplo. Resolución

x_0	$J_0^*\{x_0\}$	u_0
5	15	0

x_1	$J_1^*\{x_1\}$	u_1
0	0	0
1	3	1
2	5	2
3	10	0
4	13	1
5	15	2

x_2	$J_2^*\{x_2\}$	u_2
0	0	0
1	0	$\{0,1\}$
2	2	2
3	10	3
4	11	4
5	11	$\{4,5\}$

x_3	$J_3^*\{x_3\}$
0	0
1	0
2	0
3	0
4	0
5	0

- $J^*(x_0) = J_0^*\{5\} = 15$
- Control óptimo: $0, 2, 3$
- Trayectoria óptima: $5, 5, 3, 0$

Resolución con Matlab

- Las siguientes funciones implementan en MATLAB el principio de Bellman para la resolución de problemas de control óptimo en tiempo discreto.
- La primera función resuelve un problema de control óptimo general, mientras que la segunda se ha particularizado al caso del cáculo de camino mínimo.
- La información del sistema se pasa a través de la estructura data.

{*Js, xs, us*} = *Bellman[data]*

{*Js, xs, us*}= *MinPath[data]*

Tema 5

Optimización Dinámica

Introducción/Motivación

Control óptimo en tiempo discreto

Control óptimo en tiempo continuo
- Planteamiento
- Principio del máximo de Pontryagin
- Condiciones suficientes
- Ejemplo
- Interpretación económica de multiplicadores
- Resolución con Mathematica

Planteamiento de un problema de control óptimo en tiempo continuo

- $[t_0, t_1]$: intervalo de tiempo en el que se estudia el sistema
- $u(t)$: variable de control, $u : \Omega_u \to \mathbb{R}$, continua a trozos.
- $x(t)$: variable de estado, $x : \Omega_x \to \mathbb{R}$, derivable.
- $\frac{dx}{dt} = \dot{x} = f(x, u, t)$: ec. de estado, $t \in [t_0, t1]$, $x(t_0) = x_0$, donde
$$f : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$
y se asume que f tiene derivadas parciales primeras continuas.
- $J = \int_{t_0}^{t_1} F(x(t), u(t), t)dt + S(x(t_1))$, funcional objetivo, donde
$$F : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad S : \mathbb{R} \to \mathbb{R}$$
y F y S con derivadas parciales primeras continuas.

Nota: Las funciones x, u y f podrían ser de carácter vectorial.

Planteamiento de un problema de control óptimo en tiempo continuo

Problema modelo: Encontrar el control admisible $u : \Omega_u \to \mathbb{R}$ que maximiza (ó minimiza) el funcional:

$$\max_{u(t)} \quad J = \int_{t_0}^{t_1} F(x,u,t)dt + S(x(t_1))$$

s.a.

$$\dot{x}(t) = f(x,u,t)$$

$$x(t_0) = x_0$$

$$u(t) \in \Omega(t)$$

Definiciónes. Se denomina **control óptimo** a la función u^* que resuelve el problema. Se denomina **trayectoria de estado óptima** o **camino óptimo** a la función x^* determinada por la ecuación de estado.

Planteamiento de un problema de control óptimo en tiempo continuo

Problema modelo: Encontrar el control admisible $u : \Omega_u \to \mathbb{R}$ que maximiza (ó minimiza) el funcional:

$$\max_{u(t)} \quad J = \int_{t_0}^{t_1} F(x,u,t)dt + S(x(t_1))$$

s.a.

$$\dot{x}(t) = f(x,u,t)$$

$$x(t_0) = x_0$$

$$u(t) \in \Omega(t)$$

Definiciónes. Se denomina **control óptimo** a la función u^* que resuelve el problema. Se denomina **trayectoria de estado óptima** o **camino óptimo** a la función x^* determinada por la ecuación de estado.

Formas del funcional objetivo

Bolza	Mayer	Lagrange
$\int_{t_0}^{t_1} F(x,u,t)dt + S(x(t_1))$	$S(x(t_1))$	$\int_{t_0}^{t_1} F(x,u,t)dt$

Propiedad. Las tres formas son equivalentes. (CERDÁ. Pág. 116). En lo que sigue consideraremos la forma de Bolza.

Formas del funcional objetivo

Bolza	Mayer	Lagrange
$\int_{t_0}^{t_1} F(x,u,t)dt + S(x(t_1))$	$S(x(t_1))$	$\int_{t_0}^{t_1} F(x,u,t)dt$

Propiedad. Las tres formas son equivalentes. (CERDÁ. Pág. 116). En lo que sigue consideraremos la forma de Bolza.

Ejercicio 1: Escribir el siguiente problema en forma de Mayer.

$$\begin{aligned} \max_{u} \quad & J = \int_0^5 \left(3x - 2u - \frac{1}{2}u^2\right) dt \\ s.a. \quad & \dot{x} = x + u \\ & x(0) = 4 \end{aligned}$$

Formas del funcional objetivo

Bolza	Mayer	Lagrange
$\int_{t_0}^{t_1} F(x,u,t)dt + S(x(t_1))$	$S(x(t_1))$	$\int_{t_0}^{t_1} F(x,u,t)dt$

Propiedad. Las tres formas son equivalentes. (CERDÁ. Pág. 116). En lo que sigue consideraremos la forma de Bolza.

Ejercicio 2: Escribir el siguiente problema en forma de Lagrange.

$$
\begin{aligned}
\max_u \quad & J = \int_0^2 u^2 dt + 4x^2(2) \\
s.a. \quad & \dot{x} = 2x + 3u \\
& x(0) = 5 \\
& 0 \leqslant u \leqslant 1
\end{aligned}
$$

Principio del máximo de Pontryagin (Condiciones necesarias de optimalidad)

Defineión. Se denomina Hamiltoniano asociado a un problema de control a:

$$\boxed{H = H(x, u, \lambda, t) = F(x, u, t) + \lambda f(x, u, t)}$$

Nota. Recordar que f, y por tanto λ, podrían ser vectores.

Teorema. (Principio del máximo de Pontryagin) Sean $u^*(t)$ y $x^*(t)$ control y trayectoria óptima, entonces **existe** λ tal que:

$$\begin{aligned}
&1) \quad \dot{\lambda}^* = \frac{\partial \lambda^*}{\partial t} = -\frac{\partial H(x^*, u^*, \lambda^*, t)}{\partial x} \quad \text{con} \quad \lambda^*(t_1) = \frac{\partial S(x^*(t1))}{\partial x} \\
&2) \quad H(x^*, u^*, \lambda^*, t) \geqslant H(x^*, u, \lambda^*, t) \qquad \forall u \in \Omega_u \\
&3) \quad \dot{x}^* = f(x^*, u^*, t) \qquad \text{con } x^*(t_0) = x_0
\end{aligned}$$

Principio del máximo de Pontryagin (Condiciones necesarias de optimalidad)

Defınción. Se denomina Hamiltoniano asociado a un problema de control a:

$$\boxed{H = H(x, u, \lambda, t) = F(x, u, t) + \lambda f(x, u, t)}$$

Teorema. (Principio del máximo de Pontryagin) Sean $u^*(t)$ y $x^*(t)$ control y trayectoria óptima, entonces **existe** λ tal que:

$$\begin{array}{ll}
1) & \dot{\lambda}^* = \dfrac{\partial \lambda^*}{\partial t} = -\dfrac{\partial H(x^*, u^*, \lambda^*, t)}{\partial x} \quad \text{con} \quad \lambda^*(t_1) = \dfrac{\partial S(x^*(t1))}{\partial x} \\
2) & H(x^*, u^*, \lambda^*, t) \geqslant H(x^*, u, \lambda^*, t) \qquad \forall u \in \Omega_u \\
3) & \dot{x}^* = f(x^*, u^*, t) \qquad \text{con } x^*(t_0) = x_0
\end{array}$$

Principio del máximo de Pontryagin

Observaciones

- El problema anterior requiere resolver dos ecuaciones diferenciales (condiciones 1 y 3) y un problema de optimización (condición 2)
- Las condiciones del teorema son para el caso de un problema de maximización, si se require minimizar el funcional objetivo las condiciones son las mismas salvo que cambia el sentido de la desigualdad en la condición 2.
- El resultado anterior es válido con carácter vectorial, $f = (f_i)_{i=1}^n$, $x = (x_i)_{i=1}^n$ resulta para cada $i = 1, \ldots, n$:

$$\lambda_i = -\frac{\partial H}{\partial x_i}, \qquad \lambda_i(t_1) = \frac{\partial S(x^*(t_1))}{\partial x_i}$$
$$\dot{x}_i{}^* = f_i(x^*, u^*, t), \qquad x_i^*(t_0) = x_{i0}$$

- El teorema anterior establece condiciones necesarias de optimalidad.

Condiciones suficientes

Teorema (Mangasarian). Sean x^*, u^* y λ^* determinados por el Principio del Máximo. Entonces:

$$\left.\begin{array}{l} \text{a) } F \text{ y } f \text{ cóncavas en } x \text{ y } u \\ \text{b) } S \text{ cóncava en } x \\ \text{c) } \lambda^* \geqslant 0 \text{ (si } f \text{ no lineal)} \end{array}\right\} \Rightarrow \begin{array}{l} u^* \text{es el control óptimo} \\ x^* \text{es la trayectoria óptima} \end{array}$$

Teorema (Arrow) Sea $H^0(x, \lambda, t) = \text{máx}_u\, H(x, u, \lambda, t)$ el Hamiltoniano derivado. Sean x^*, u^* y λ^* determinados por el Principio del Máximo. Entonces:

$$\left.\begin{array}{l} \text{a) } H^0(x, \lambda, t) \text{ es cóncava en } x \\ \text{b) } S \text{ cóncava en } x \end{array}\right\} \Rightarrow \begin{array}{l} u^* \text{es el control óptimo} \\ x^* \text{es la trayectoria óptima} \end{array}$$

Ejemplo

Resolver

$$\begin{aligned} \text{máx} \qquad & J = \int_0^1 (x+u)dt, \\ s.a. \qquad & \dot{x} = 1 - u^2 \\ & x(0) = 1 \end{aligned}$$

<u>Solución</u>: $H = (x+u) + \lambda(1-u^2)$

Ejemplo

Resolver

$$\begin{aligned} \text{máx} \qquad & J = \int_0^1 (x+u)dt, \\ s.a. \qquad & \dot{x} = 1 - u^2 \\ & x(0) = 1 \end{aligned}$$

Solución: $H = (x+u) + \lambda(1-u^2)$ Solución:
$H = (x+u) + \lambda(1-u^2)$
Condición 1:

$$\left.\begin{array}{l} \dot{\lambda} = -\frac{\partial H}{\partial x} = -1 \\ \lambda(1) = 0 \end{array}\right\} \Rightarrow \lambda(t) = -t + C \xrightarrow{\lambda(1)=0} \lambda(t) = 1 - t$$

Por tanto: $\boxed{\lambda^* = 1 - t}$

Ejemplo

Resolver

$$\begin{aligned} \text{máx} \qquad & J = \int_0^1 (x+u)dt, \\ s.a. \qquad & \dot{x} = 1 - u^2 \\ & x(0) = 1 \end{aligned}$$

<u>Solución</u>: $H = (x+u) + \lambda(1-u^2)$
Condición 2: Debemos resolver:

$$\max_u \qquad H(x,u,\lambda^*,t) = x + u + (1-t)(1-u^2)$$

$$\Rightarrow \frac{\partial H}{\partial u} = 1 - 2u(1-t) = 0 \Rightarrow u = \frac{1}{2(1-t)}, \quad t \neq 1$$

$$\Rightarrow \frac{\partial^2 H}{\partial u^2} = 2 - 2t \geqslant 0, \quad (0 \leqslant t \leqslant 1) \Rightarrow \text{Es máximo}$$

Por tanto $\boxed{u^* = \frac{1}{2(1-t)}}$

Ejemplo

Resolver

$$\begin{aligned} \text{máx} \qquad & J = \int_0^1 (x+u)dt, \\ s.a. \qquad & \dot{x} = 1 - u^2 \\ & x(0) = 1 \end{aligned}$$

Solución: $H = (x+u) + \lambda(1-u^2)$
Condición 3: Resolvemos la ecuación de estado:

$$\left.\begin{array}{l} \dot{x} = 1 - \frac{1}{4(1-t)^2} \\ x(0) = 1 \end{array}\right\} \Rightarrow x(t) = \int 1 - \frac{1}{4(1-t)^2} dt = t - \frac{1}{4(1-t)} + C$$

$$\xrightarrow[x(0)=1]{} \boxed{x(t) = t - \frac{1}{4(1-t)} + \frac{5}{4}.}$$

Ejemplo

Resolver

$$\begin{aligned} \text{máx} \qquad & J = \int_0^1 (x+u)dt, \\ s.a. \qquad & \dot{x} = 1 - u^2 \\ & x(0) = 1 \end{aligned}$$

<u>Solución</u>: $H = (x+u) + \lambda(1-u^2)$
Por tanto:

$$\boxed{\begin{aligned} & u^*(t) = \frac{1}{2(1-t)} \\ & x^*(t) = t - \frac{1}{4(1-t)} + \frac{5}{4} \\ & \lambda^*(t) = 1 - t \end{aligned}}$$

Ejemplo

Resolver

$$\begin{array}{ll} \text{máx} & J = \int_0^1 (x+u)dt, \\ s.a. & \dot{x} = 1 - u^2 \\ & x(0) = 1 \end{array}$$

Solución: $H = (x+u) + \lambda(1-u^2)$
Veamos que es un máximo
Mangasarian.

1) $F(x,u,t) = x + u$ es lineal, en particular, cóncava

2) $f(x,u,t) = 1 - u^2$ tiene por hessiano con respecto a x, u
$\begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}$, por tanto cóncava

3) $\lambda^*(t) = 1 - t \geqslant 0$ en $[0, 1]$.

Ejemplo

Resolver

$$\begin{aligned} \max \qquad & J = \int_0^1 (x+u)dt, \\ s.a. \qquad & \dot{x} = 1 - u^2 \\ & x(0) = 1 \end{aligned}$$

Solución: $H = (x+u) + \lambda(1-u^2)$
Veamos que es un máximo
Arrow.
Es fácil determinar que $H^0(x,\lambda,t) = x + \lambda + \frac{1}{4\lambda}$

1) H^0 es lineal en x, en particular, cóncava para cualquier λ

2) $S \equiv 0$, en particular cóncava.

Interpretación económica de los multiplicadores

Se define la función valor como:

$$V^*(x,t) = \max_u \int_{t_0}^{t_1} F(x,u,t)dt + S(x(t_1))$$

s.a.

$$\dot{x}(t) = f(x,u,t)$$

$$x(t_0) = x_0$$

$$u(t) \in \Omega(t)$$

Se puede demostrar que:

$$\boxed{\lambda(x,t) = \frac{\partial V^*(x,t)}{\partial x}}$$

es decir, es la variación de la función valor frente a los cambios de la variable de estado x en el tiempo t.

Resolución con Mathematica

El software Mathematica ofrece herramientas para la resolución de problemas de control en tiempo continuo. Algunos comandos útiles son:

- **DSolve[. . .]**
- **Solve[. . .]**
- **D[. . .]**
- **/.**

Bibliografía

Bibliografía I

- M. S. Bazaraa, J. J. Jarvis, H. D. Sherali. *Linear Programming and Network Flows*. John Wiley & Sons, 1977.
- M. S. Bazaraa, C. M. Shetty. *Nonlinear Programming. Theory and Algorithms*. John Wiley & Sons, 1979.
- B. Beavis. *Optimization and stability theory for economic analysis*. Cambridge University Press, 1990.
- N. L. Biggs. *Discrete Mathematics*. Oxford U.P., 2002
- E. Cerdá. *Optimización dinámica*. Prentice Hall, 2001.
- B. García-Bernalt, M. D. García Sanz. *Modelos de Decisión*. Apuntes, 2015.
- M. A. Goberna, V. Jornet, R.O. Puente. *Optimización lineal. Teoría, métodos y modelos*. McGraw-Hill, 2004.

Bibliografía II

- L. Ljungqvist, T. L. Sargent. *Recursive Macroeconomic Theory*. MIT Press, 2012.
- G. Sánchez. *Mathematica: Más allá de las Matemáticas*. Addlink Media, 2013.
- N. L. Stokey, R. E. Lucas. *Recursive Methods in Economic Dynamics*. Harvard U. P., 1989.
- R.K. Sundaram. *A First Course in Optimization Theory*. Cambridge University Press. 1996.