Universidad Rey Juan Carlos

Métodos Operativos y Estadísticos de Gestión

César Beltrán Royo

Versión 2.4

Índice general

Capítulo 1

Introducción

Este capítulo está estructurado en la siguientes secciones:

- La empresa y sus fines.
- Organización y estructura de la empresa.
- El papel de la investigación de operaciones en la organización de empresas.

1.1. La empresa y sus fines

- Definición: Una empresa es una organización o entidad en la que se unen una serie de factores o recursos, materiales, técnicos, financieros y humanos, para realizar actividades y que persigue unos fines que suelen contemplar la obtención de beneficio económico.
  - ▷ Existen distintas definiciones desde los puntos de vista legal, financiero, sociológico, comercial, sus objetivos o de sus actividades
  - ▷ Por un lado existen gran cantidad de formas legales para una empresa, que van desde las grandes multinacionales (por ejemplo Movistar) hasta las microempresas donde el propietario es el único trabajador.
- Objetivo: La mayoría de las empresas tienen como objetivo la obtención de beneficio económico, pero existen organizaciones, como las empresas públicas, cuyo objetivo es prestar un servicio público y donde el beneficio es algo secundario, por ejemplo la corporación RTVE, la Sociedad Estatal de Correos, etc.
- Actividades: Comerciales, financieras, industriales o de servicios, pudiendo mezclar varios tipos de actividades en muchas ocasiones.
- Grupos de interés (stakeholders): Se refiere a los grupos que tienen interés en una empresa, dado que pueden afectar o ser afectados por las actividades de la empresa. Dentro de los grupos de interés podemos encontrar:
  - ▷ Propietarios: pueden tener varios objetivos como la rentabilidad de la empresa, su expansión, la sostenibilidad de la empresa a largo plazo, etc.
  - ▷ Inversores: su objetivo es rentabilizar su inversión.

- ▷ Trabajadores: su objetivo es el mantenimiento de su puesto de trabajo y el cobro de sus remuneraciones.
- ▷ Acreedores: su objetivo es el cobro de las cantidades adeudadas por la empresa (dentro de éstos podemos incluir a los Bancos).
- ▷ Administración: su objetivo es el cobro de impuestos y el mantenimiento de la empresa (por su generación de riqueza económica, puestos de trabajo y el servicio prestado a la sociedad).
- ▷ Directivos: su objetivo es el crecimiento de la empresa y su beneficio (gran parte de sus ingresos dependen de los bonus que obtienen por ello).
- ▷ Clientes: su objetivo es el mantenimiento y mejora de los productos que compra a la empresa.
- ▷ Ciudadanos en general: sus objetivos pasan por la corrección medioambiental de las empresas y la generación de puestos de trabajo / riqueza.

- El objetivo de una empresa depende de los objetivos de sus grupos de interés en proporción directa a la fuerza que cada grupo tenga en la empresa.

- Ejemplo: En el caso de una empresa cooperativa, donde los propietarios son los trabajadores (por lo que son el principal grupo de poder), el objetivo fundamental es la permanencia de los puestos de trabajo y el pago de las nóminas.

- Ejemplo: En una empresa cuya titularidad es pública (propiedad de una Administración) el fin es la prestación de un servicio público.

- Ejemplo: En una empresa donde los directivos tienen mucho poder (no tienen porqué ser propietarios, como ocurriría en una empresa que cotice en Bolsa) el fin es el crecimiento, etc.

1.2. Organización y estructura de la empresa

- La empresa debe organizar los recursos materiales, humanos, financieros y tecnológicos de los que dispone para la realización de sus actividades.

- Esta organización se realizará a través de una estructura espacio-temporal que será diferente para cada entidad y que por supuesto irá evolucionando con el paso del tiempo.

- Normalmente las empresas suelen organizarse por funciones, a través de departamentos, que se encargan de una serie de tareas definidas previamente. Sin embargo, hay ocasiones en las que las empresas se organizan departamentalmente por ubicaciones geográficas (por Ej. en multinacionales), por tipología de clientes (por Ej. las constructoras), por producto, por operaciones, etc.

- La escala organizativa de una empresa se suele representar a través de un esquema en forma de pirámide donde se indica los tres niveles organizativos, el flujo de información y las decisiones a tomar (Figura. 1.1). Se organiza en los siguientes niveles:

  - ▷ Nivel inferior: nivel Operativo, formado por aquellos que desarrollan las actividades del negocio de la empresa.
  - ▷ Nivel medio: nivel Ejecutivo, que toma decisiones a corto plazo y conforme con las directrices y planificación a medio y largo plazo establecidas en la compañía.
  - ▷ Nivel superior: nivel Directivo o Político, que marca las directrices de la compañía y realiza la planificación a largo plazo.

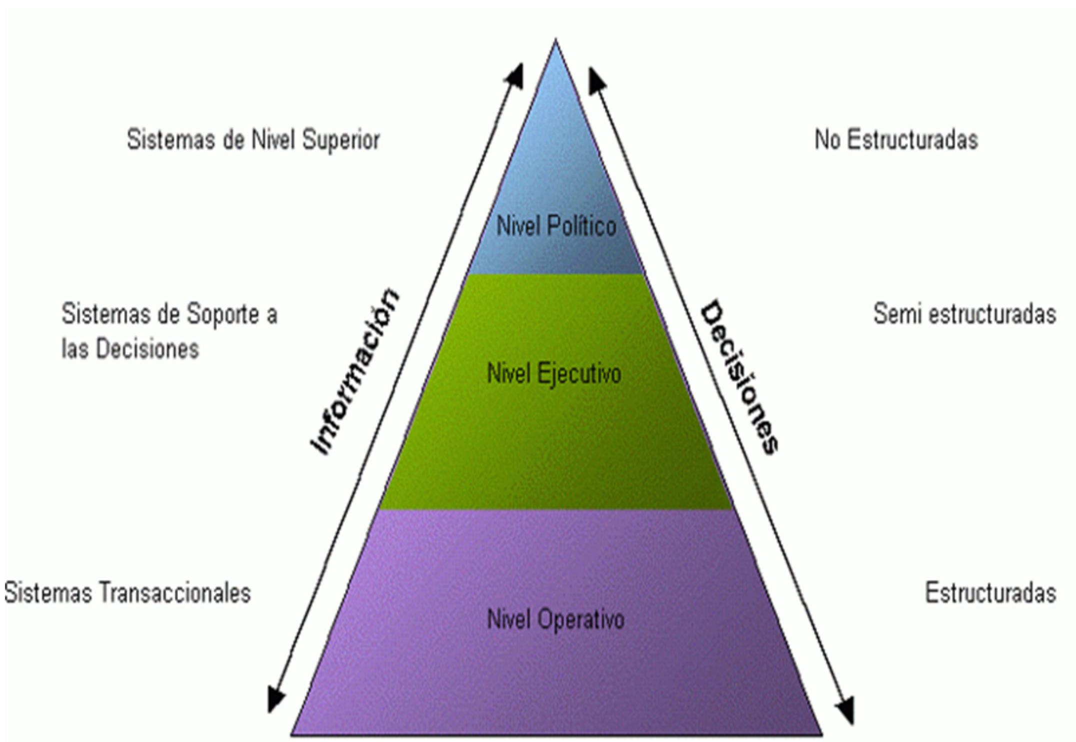


Figura 1.1: Pirámide organizativa.

- La información deberá fluir de cada uno de estos niveles tanto en dirección ascendente como descendente, aunque las necesidades de datos no son las mismas en todos los niveles. De la misma forma también en todos los niveles se deberán tomar decisiones, pero no serán del mismo tipo, ni con los mismos efectos en la empresa.

- Otro tipo de organización de empresas es a través de Departamentos funcionales, que se encargan de una serie de tareas definidas previamente.

- Tradicionalmente los Departamentos funcionales más comunes son: Aprovisionamiento, Producción, Marketing y Administración, aunque hay empresas que también cuentan con Departamentos de Control de Calidad, I+D, Personal, etc.

- El número y tamaño de los Departamentos funcionales dependerá de cada empresa, condicionados por el tamaño, sector e importancia de las funciones de la compañía y sobre todo de la estructura planteada por el nivel Político.

- En muchos casos se mezclan estructuras departamentales funcionales con estructuras de otro tipo, como las de ubicación geográfica, las de tipo de producto o de cliente. Estas situaciones suelen ocurrir en empresas de tamaño muy grande, en las que es necesaria una especialización importante y en las que un control centralizado (en muy pocas manos) sería imposible.

1.3. El papel de la investigación de operaciones en la organización de empresas.

- Los métodos operativos de gestión que se explican en esta asignatura pertenecen al área científico-técnica denominada 'Investigación de Operaciones'.

- La investigación de operaciones, también llamada, Investigación Operativa (IO) se encarga de la aplicación de herramientas y métodos matemáticos y estadísticos para la resolución de problemas sobre organización, planificación y control en cualquier tipo de sistema, sea artificial como una empresa o natural como pueda ser un río.

- El objeto fundamental de la IO es el de servir como ayuda en la toma de decisiones complejas o muy complejas, de cara a tomar la mejor decisión posible en cada momento.

- En el caso de las empresas, el papel de la IO es fundamental en la toma de decisiones a corto plazo, pero sobre todo, en la toma de decisiones a medio y largo plazo a través de la planificación estratégica de las compañías.

- Las principales cuestiones sobre las que utilizaremos las técnicas de IO dentro de la organización y gestión en las empresas serán:

  ▷ Toma de decisiones, ya sea a través de la optimización como a través de los procesos decisorios.

  ▷ Planificación y gestión a corto, medio y largo plazo

  ▷ Gestión de la Calidad en la empresa

- A continuación pasaremos a describir en qué consiste cada uno de estos temas que posteriormente iremos desarrollando a lo largo del curso.

- Toma de Decisiones:

- ▷ Son fundamentales para la correcta gestión de la empresa, ya que una decisión mala pone en riesgo la solvencia de la compañía y disminuye su rentabilidad.
- ▷ Los procesos de Toma de Decisiones se pueden producir a través de la utilización de herramientas de optimización, que buscan maximizar los conceptos positivos (ingresos, ventas, beneficios, etc.) y minimizar los negativos (costes, sanciones, etc.)
- ▷ También pueden apoyarse en herramientas de Análisis de Decisiones, ya sea para casos en los que se considera un único criterio o para aquellos otros en los que se consideran varios criterios.

- Planificación:
  - ▷ El realizar planes a corto, medio o largo plazo es algo fundamental en una empresa, ya que muchas inversiones son productivas al termino de plazos medios o largos.
  - ▷ En nuestro caso vamos a estudiar con mayor profundidad la planificación de la gestión de proyectos.
  - ▷ La planificación de la gestión de proyectos es una herramienta fundamental tanto para el desarrollo de proyectos de cara a clientes, como para el desarrollo de proyectos de inversión.
  - ▷ En general la planificación es una herramienta fundamental para el desarrollo futuro y presente de cualquier compañía, sobre todo de cara al crecimiento y rentabilidad futura.
- Gestión de la Calidad:
  - ▷ Hoy en día nadie puede dudar de la importancia de la existencia de sistemas de control y gestión de calidad en cualquier empresa. Han representado uno de los grandes avances industriales del siglo XX y deberán desarrollarse durante el actual siglo.
  - ▷ Por un lado aumentan la rentabilidad, aumentando las ventas y disminuyendo los costes (por las menores devoluciones y demandas).
  - ▷ Por otro lado garantizan la competitividad con el entorno, haciendo el producto más atractivo hacia el cliente y dejando la puerta siempre abierta de la ingeniería de procesos y de la gestión del cambio.

Capítulo 2

Programación lineal

2.1. Introducción

- *Optimizar:* Dado un conjunto de dos o más opciones, elegir la *mejor.*
- La posibilidad de optimizar es *ubicua:*
  - ▷ *Número de cajeros/as* en un supermercado: Cuantos más cajeros mayor satisfacción de los clientes, pero mayor coste para el supermercado.
  - ▷ *Precio de un producto:* A mayor precio más ganancias por producto, pero menos ventas.
  - ▷ *Vida útil de una infraestructura informática:* El reducir los intervalos para actualizar la infraestructura informática aumenta los costes, pero también aumenta la productividad.
  - ▷ Etc.
- *Filosofía de este curso:* Prácticamente todo se puede optimizar'.
- En este capítulo está dedicado a la denominada *'Optimización lineal'.*
- A menudo se utiliza el término *'Programación lineal'* como *sinónimo* de 'Optimización lineal'. Se refiere a 'Programación/Planificación de operaciones' (no a la 'Programación informática').

Este capítulo está estructurado en la siguientes *secciones:*

- Modelos de la programación lineal.
- Introducción a las técnicas de resolución.
- Postoptimización.

Nota. Este tema también se puede consultar y/o ampliar en los Capítulos 4 y 5 del siguiente libro: Rardin, R. L., 'Optimization in operations research', Editorial Prentice Hall, 1998.

2.2. Modelos de la programación lineal

Objetivo:

- El objetivo de esta sección es aprender a *formular problemas* de 'programación lineal' (PL), también llamada 'optimización lineal'.
- Notar que plantearemos los problemas pero *no los resolveremos.* En la sección siguiente veremos cómo se pueden resolver.
- A menudo simplificaremos la expresión 'tenemos un problema de PL' por 'tenemos un PL'.

Apartados:

- Modelos de asignación de recursos
- Modelos de mezclas
- Modelos de planificación de operaciones
- Modelos de gestión de personal
- Modelos temporales (multiperiodo)

2.2.1. Modelos de asignación de recursos

Ejemplo 1 (Asignación de horas de estudio)

Datos:

- Juan, estudiante de ingeniería, quiere maximizar sus *resultados académicos.*
- Número total de *horas disponibles* para estudiar: 30 h.
- La estimación del *incremento de la nota* de cada asignatura aportado por cada hora de estudio viene dado en la siguiente tabla:

Investigación Operativa (IO)	2 %
Ingeniería Económica (IE)	3 %
Estadística (ES)	1 %
Programación (PR)	5 %

- Juan quiere:
  - ▷ Que el número de *horas de estudio dedicadas a IO* sea igual o superior al resto.
  - ▷ Un máximo de *10 h de estudio* por asignatura.

Objetivo: *Formula* este problema de optimización como un problema de Programación Lineal (PL) (no hay que resolverlo).

Operaciones 1:

Solución:

- De forma compacta podemos escribir:

$$\begin{array}{lll} \max & z = 2\,x_{IO} + 3\,x_{IE} + 1\,x_{ES} + 5\,x_{PR}, & \textit{(incremento total)} \\ \text{s. a.} & x_{IO} + x_{IE} + x_{ES} + x_{PR} = 30, & \textit{(asignación)} \\ & x_{IO} \geq x_j, \quad j \in \{\text{IE, ES, PR}\} & \textit{(IO la que más)} \\ & x_j \leq 10, \quad j \in \{\text{IO, IE, ES, PR}\} & \textit{(máximo 10 h)} \\ & x_j \geq 0 \quad j \in \{\text{IO, IE, ES, PR}\} & \textit{(mínimo 0 h),} \end{array}$$

  donde *'max'* significa 'Maximizar' y *'s. a.'* significa 'Sujeto a (las siguientes restricciones)'.

- Más adelante veremos cómo se puede *resolver* este tipo de problemas de optimización.

- De momento nos tenemos que conformar con saber que una *solución óptima* vendría dada por el siguiente número de horas:

$$x^*_{IO} = 10, \quad x^*_{IE} = 10, \quad x^*_{ES} = 0, \quad x^*_{PR} = 10.$$

- Con esta asignación óptima el *incremento total acumulado* para todas las notas de Juan sería $z^* = 100\,\%$:
$$z^* = 2 \cdot 10 + 3 \cdot 10 + 1 \cdot 0 + 5 \cdot 10 = 100.$$
- Notar que para indicar valores óptimos usamos los símbolos x^* y z^*.

General

- *Modelos de asignación:*
  - ▷ *Repartimos* un recurso entre varias actividades.
  - ▷ En el ejemplo anterior hemos repartido 30 horas de estudio (el recurso) entre varias asignaturas (las actividades).
- *Variables de decisión:*
  - ▷ Determinan qué *cantidad de recurso* asignamos a cada actividad.
- *Problema de Programación Lineal (PL):*
  - ▷ El modelo o problema de asignación es un caso particular de PL.
  - ▷ Un problema de optimización es un PL si su función objetivo $z(x)$ y sus restricciones son funciones *lineales.*

$$\begin{array}{lll} \min & z(x) = c_1 x_1 + \ldots + c_n x_n & \textit{(función objetivo)} \\ \text{s. a.} & a_{11} x_1 + \ldots + a_{1n} x_n = b_1 & \textit{(restricciones)} \\ & \vdots & \\ & a_{m1} x_1 + \ldots + a_{mn} x_n = b_m & \\ & x_i \geq 0 \qquad \forall i \in \{1, \ldots, n\} & \end{array}$$

  - ▷ En este PL tenemos sólo restricciones de igualdad. En general un PL puede tener restricciones tanto de *igualdad* como de *desigualdad.*
  - ▷ En un PL nos puede interesar maximizar un *beneficio* (dinero, fiabilidad, seguridad, etc.) o minimizar un *coste* (dinero, tiempo, distancia, etc.). Escribiremos 'min' o 'max' según cada caso.
- *Expresión matricial de un PL:*
  - ▷ Notar que en el anterior PL usando el *producto escalar* podemos escribir la función objetivo como
$$z(x) = c^\intercal x.$$
  - ▷ Usando el producto de *matriz por vector* podemos escribir las anteriores restricciones en formato matricial
$$Ax = b,$$
donde la matriz A y el vector b están definidos a partir de los *coeficientes* a_{ij} y b_i, respectivamente.
  - ▷ De esta forma podemos escribir el anterior PL en formato matricial como sigue:

$$\begin{array}{llr} \min & c^\intercal x & \textit{(función objetivo)} \\ \text{s. a.} & Ax = b & \textit{(restricciones)} \\ & x \geq 0. & \textit{(restricciones de no negatividad)} \end{array}$$

2.2.2. Modelos de mezclas

Ejemplo 2 (Problema de la dieta)

Objetivo:

- Una compañía quiere decidir la *composición* de mínimo coste de un pienso para gallinas.
- *Formula* este problema de optimización como un PL (no hay que resolverlo).

Datos:

- Aunque la compañía debe fabricar 10.000 kg de pienso, es suficiente con calcular la composición de *1 kg* y luego aplicar la composición resultante a todo el proceso de fabricación.
- Cada kg de *pienso* debe contener:
  - ▷ Como máximo 10 mg de *sodio.*
  - ▷ Al menos 50 g de *fibra.*
  - ▷ Al menos 400 g de *maíz.*
- El pienso se va a fabricar a partir de tres ingredientes: maíz, trigo y avena.
- La composición y precio por kg de cada *ingrediente* es:

Ingrediente	g *Fibra* / kg	mg de *Sodio* / kg	*Euros* / kg
Maíz (1)	40	10	0,15
Trigo (2)	90	15	0,12
Avena (3)	30	5	0,20

Operaciones 2:

Solución: El PL para calcular la *composición* de 1 kg de pienso de mínimo coste corresponde a:

$$
\begin{array}{lll}
\min & z = 0,15\,x_1 + 0,12\,x_2 + 0,20\,x_3 & \textit{(coste)} \\
\text{s. a.} & x_1 + x_2 + x_3 = 1 & \textit{(composición de 1 kg)} \\
& 10x_1 + 15x_2 + 5x_3 \leq 10 & \textit{(sodio)} \\
& 40x_1 + 90x_2 + 30x_3 \geq 50 & \textit{(fibra)} \\
& x_1 \geq 0,4 & \textit{(maiz)} \\
& x_i \geq 0 \quad i \in \{1,2,3\} & \textit{(no negatividad)}
\end{array}
$$

General

- *Modelos de mezclas:* En los modelos de asignación *repartimos* un recurso. En los modelos de mezclas los *combinamos.*
- *Variables de decisión:* Determinan qué *cantidad de cada recurso* incluimos en la mezcla.
- *Restricciones de composición:* Determinan cotas superiores y/o inferiores a las *propiedades de la mezcla* resultante.

2.2.3. Modelos de planificación de operaciones

Ejemplo 3

Objetivo:

- Una compañía quiere *planificar su producción* de zumo de naranja para maximizar su beneficio.
- *Formula* este problema de optimización como un PL (no hay que resolverlo).

Datos:

- El *precio de venta al público* del zumo fabricado es de 1500 Euros/Tonelada.
- La compañía estima que tiene una *demanda* de 15000 toneladas (T) de zumo.
- La compañía fabrica su zumo a partir de *naranjas* y/o *concentrado* de zumo de naranja (CZN).

Producto	*Precio* (Euros/ T)	*Efectividad* (T. de zumo / T. de producto)	*Existencias* (T)
Naranjas	200	0,2	10000
CZN	1600	2	Sin límite

Operaciones 3:

Solución: Ver apartado 'Operaciones'.

General

- *Modelos de planificación de operaciones:* Debemos decidir *qué* hacer, *cuándo* y *dónde.*
- *Variables de decisión:* Las variables de decisión *enteras* y de gran magnitud se suelen tratar como variables *continuas* para simplificar la resolución del problemas.
- *Restricciones de balance:* El *flujo entrante* de materias primas debe ser igual al *flujo saliente* de productos manufacturados *(ecuación de balance).* Intervienen *factores de conversión.*

2.2.4. Modelos de gestión de personal

Ejemplo 4

Objetivo:

- Una agencia estatal quiere *optimizar* el número de operarios de su plantilla, pero cubriendo sus necesidades operativas.
- *Formula* este problema de optimización como un PL (no hay que resolverlo).

Datos:

- De los 5 días laborables, los empleados *trabajan 4 días* a razón de 10 h/día y tienen un *día libre,* según su turno.

Turno	*Día libre*
1	Martes
2	Miércoles
3	Jueves

- El *número de empleados* requeridos cada día de la semana se detalla a continuación.

L	M	Mx	J	V
10	7	7	7	9

Operaciones 4:

Solución: Ver apartado ‘Operaciones’.

General

- *Modelos de gestión de personal:*
  - ▷ Los modelos de *gestión de operaciones* deciden qué tarea realizar en cada momento de forma que los recursos sean usados eficientemente.
  - ▷ En la gestión de personal decidimos qué *tipo* de empleado y *cuántos* de cada tipo deben realizar cada tarea.
- *Restricciones de cubrimiento:*

▷ En la planificación de los turnos debemos asegurar que el número de operarios asignados cubre las *necesidades de cada periodo.*

▷ Para ello imponemos la siguiente *restricción de cubrimiento:*

$$\sum_{\text{turnos}} (\text{Producción / Operario}) \times (\text{Operarios en servicio})$$

$\geq$ *Necesidades* del periodo.

2.2.5. Modelos de planificación multiperiodo

Ejemplo 5

Objetivo:

- Una compañía de fabricación de palas para la nieve quiere *optimizar su producción* y almacenaje.
- *Formula* este problema de optimización como un PL (no hay que resolverlo).

Datos:

- El *coste* de fabricación y almacenamiento en euros/pala para cada trimestre se detalla en la siguiente tabla:

T1	T2	T3	T4
13	12	11	10

- El *stock* de palas al principio del primer trimestre es de 0 palas.
- La *demanda* en palas para cada trimestre se detalla en la siguiente tabla:

T1	T2	T3	T4
11.000	48.000	64.000	15.000

Operaciones 5:

Solución: El PL para optimizar la *producción* y almacenamiento de las palas corresponde a:

$$
\begin{array}{lll}
\min & z = 13\,x_1 + 12\,x_2 + 11\,x_3 + 10\,x_4 & \textit{(euros)} \\
\text{s. a.} & s_0 = 0 & \textit{(stock inicial)} \\
& s_0 + x_1 = 11000 + s_1 & \textit{(primer trimestre)} \\
& s_1 + x_2 = 48000 + s_2 & \textit{(segundo trimestre)} \\
& s_2 + x_3 = 64000 + s_3 & \textit{(tercer trimestre)} \\
& s_3 + x_4 = 15000 + s_4 & \textit{(cuarto trimestre)} \\
& s_t \geq 0 \quad t \in T & \textit{(no negatividad)} \\
& x_t \geq 0 \quad t \in T & \textit{(no negatividad)}
\end{array}
$$

General

- *Modelos de planificación multiperiodo:*
  - ▷ Los modelos estudiados hasta ahora se han limitado a un solo periodo *(modelos estáticos).*
  - ▷ En muchas ocasiones hay que planificar en un horizonte de tiempo que abarca varios periodos *(modelos dinámicos* o *multiperiodo).*
- *Variables de decisión:* En estos problemas, muchas variables de decisión tiene una *versión para cada periodo.*
- *Restricciones de balance:*
  - ▷ Describen la *evolución temporal* de una magnitud.
  - ▷ A menudo es suficiente describir la evolución experimentada entre dos *periodos consecutivos.*

$$\text{Situación}_t + \text{Cambios}_t = \text{Situación}_{t+1}.$$

2.3. Método para formular un PL

Para formular un PL asociado a un *problema de decisión,* seguiremos los siguientes pasos:

- *Datos:* Escribir en *lenguaje natural:*

  - ▷ Los datos relevantes (con *unidades).*
  - ▷ Las condiciones que debe satisfacer toda solución.

- *Objetivo:*
  - ▷ Escribir el objetivo del problema en lenguaje natural.
  - ▷ Especificar: minimizar un coste o maximizar un beneficio.

- *Operaciones:*
  1) Definir las *variables de decisión* (con unidades) x_j, mencionando j en la definición.
  2) *Familiarizarse* con el problema: Buscar un vector solución $\widehat{x}$ por tanteo, que cumpla todas las condiciones y calcular su coste o beneficio $\widehat{z}$.
  3) Escribir en *lenguaje matemático:*
     a) $z(x)$, es decir, la *función objetivo* a maximizar o minimizar (con unidades).
     b) Plasmar cada condición con una *restricción* (ecuación o inecuación).
     c) En cada restricción, las unidades de los lados derecho e izquierdo deben ser *iguales.*
     d) Dos restricciones cualesquiera pueden tener unidades diferentes.
     e) *Verificar* si $\widehat{x}$ cumple las restricciones y si $z(\widehat{x}) = \widehat{z}$.

- *Solución:* Resumir en un *PL* todo lo anterior (normalmente hay que añadir restricciones de no negatividad, $x \geq 0$).

2.4. Introducción a las técnicas de resolución de un PL

En esta sección veremos dos apartados:

- Resolución *geométrica* de un problema PL.
- Problema PL en *formato estandar.*

2.4.1. Resolución geométrica de un PL

Ejemplo 6 (Representación de rectas)

Datos:

- Consideramos la recta R_k definida por la siguiente *ecuación:*

$$2x_1 + 3x_2 = k.$$

- Una notación *alternativa* consiste en decir: Consideramos la recta R_k

$$R_k = \{x \in \mathbb{R}^2 \mid 2x_1 + 3x_2 = k\}.$$

Objetivo:

1) Representa las rectas R_k para $k = 6, 12, 18, 0$.
2) Supongamos que en un PL su *función objetivo* es

$$z(x) = 2x_1 + 3x_2$$

   ¿Cuánto vale la función objetivo para todos los puntos que están sobre la *recta* R_6? ¿Y los que están en R_{12}? ¿Y en R_{18}? ¿Y en R_0?
3) Supongamos que en el anterior PL estamos *maximizando* un beneficio. ¿En cuál de las anteriores rectas la función objetivo alcanza un mejor valor?
4) Escribe la función objetivo como *producto escalar* $c^\intercal x$.
5) ¿Qué *ángulo* forma el vector c con las rectas anteriores?

Operaciones 6:

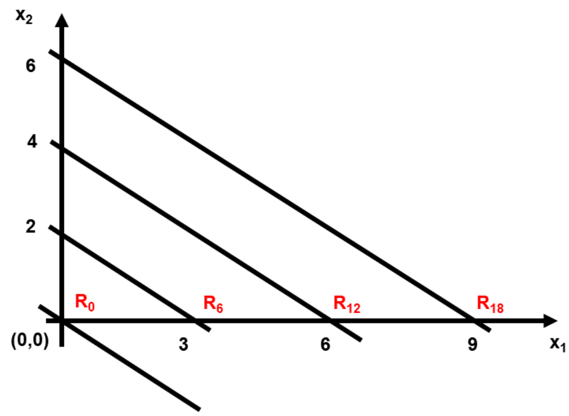


Figura 2.1: Las cuatro rectas son *paralelas.*

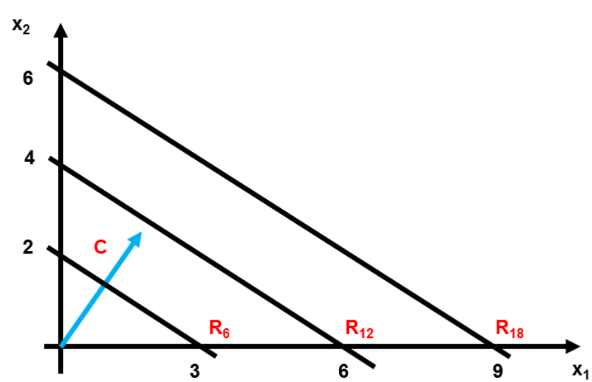


Figura 2.2: c es perpendicular a las curvas de nivel de $z(x) = c^\intercal x$.

Solución:

1) Ver apartado 'Operaciones' (cuestión 1).
2) Para todos los puntos que están sobre $R_0,\ R_6,\ R_{12}$ o R_{18} la función objetivo vale $0,\ 6,\ 12$ ó 18, respectivamente.
3) La función objetivo alcanza mejor valor en los puntos que están en R_{18}.
4) $z(x) = c^\intercal x$. donde $c^\intercal = (2,\ 3)$ y $x = \binom{x_1}{x_2}$.
5) 90º.

General (Curvas de nivel en un PL)

- Consideramos un PL de *dimensión dos,* cuya función objetivo es

$$z(x) = c^\intercal x,$$

  donde $c^\intercal = (c_1,\ c_2)$ y $x = \binom{x_1}{x_2}$.

- La *'curva de nivel k'* de $z(x)$ es el lugar geométrico donde $z(x)$ alcanza un valor constante igual a k :

$$\{x \in \mathbb{R}^2 \mid c^\intercal x = k\}.$$

- Es decir, la *'curva de nivel k'* de $z(x)$ es la recta

$$R_k = \{x \in \mathbb{R}^2 \mid c^\intercal x = k\}.$$

- Así, las curvas de nivel de $z(x) = c^\intercal x$ son todas las rectas perpendiculares a c y, por tanto, paralelas entre ellas.

- En estas curvas el valor de k aumenta en la dirección de c.

Ejemplo 7 (Representación de semiplanos)

Datos: Consideramos los *semiplanos* S_1 y S_2 definidos por la siguientes *inecuaciones*

$$S_1 \equiv 2x_1 + x_2 \leq 6$$

g

$$S_2 \equiv x_1 + 2x_2 \leq 6.$$

Objetivo:

1) Representa la región del *primer cuadrante* del plano que cumple la primera inecuación. Nota: el primer cuadrante es la región del plano con $x_1 \geq 0$ y $x_2 \geq 0$.
2) Representa la región del primer cuadrante del plano que cumple las dos inecuaciones anteriores.

Operaciones 7:

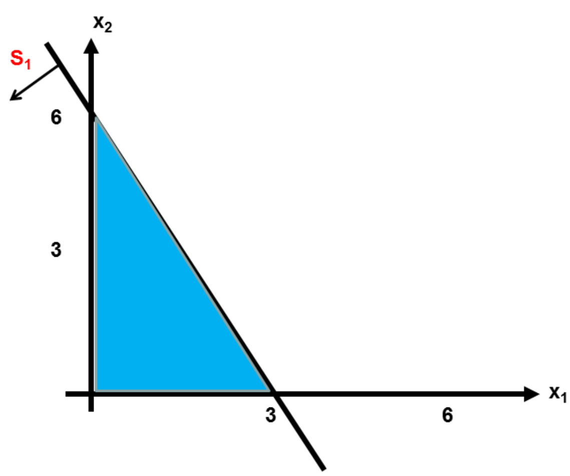


Figura 2.3: Región determinada por S_1.

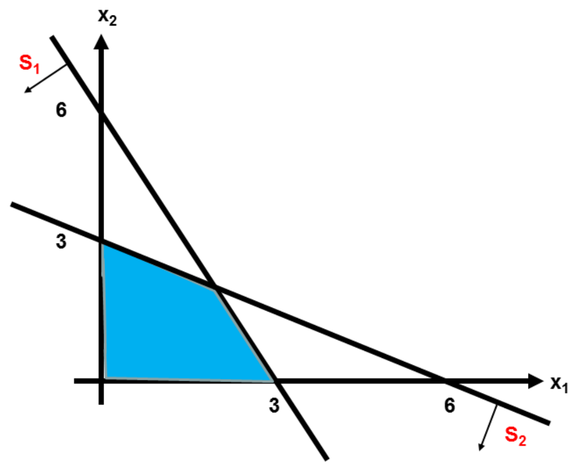


Figura 2.4: Región determinada por S_1 y S_2.

Solución: Ver apartado 'Operaciones'.

General (Semiplanos en un PL)

- Consideramos un PL de *dimensión dos* cuyas restricciones son todas de desigualdad y con variables positivas.
- Cada una de dichas desigualdades corresponde a una inecuación que determina un *semiplano.*
  - ▷ Nota: El punto p elegido para determinar el semiplano no tiene que pertenecer a la *recta frontera* del semiplano (p debe cuplir la inecuación de forma estricta).
- La *región factible* del PL corresponde a la región determinada por los puntos que cumplen todas las restricciones.
- En este caso, la región factible corresponde a la región del primer cuadrante donde *intersectan* todos los semiplanos asociados a las restricciones del PL.

Ejemplo 8 (Resolución gráfica de un PL)

Datos: Consideramos el siguiente PL

$$\begin{array}{rll} \max & z = 2x_1 + 3x_2 & \\ \text{s.a.} & 2x_1 + x_2 \leq 6 & \text{(Semiplano } S_1) \\ & x_1 + 2x_2 \leq 6 & \text{(Semiplano } S_2) \\ & x \geq 0. & \end{array}$$

Objetivo:

1) Resuelve este PL mediante el método de las *curvas de nivel.*
2) ¿Qué solución obtendríamos si en vez de maximizar, quisiéramos minimizar $z(x)$?
3) Resuelve el apartado 1 pero con $2x_1 + x_2 \geq 6$ en vez de $2x_1 + x_2 \leq 6$ (Semiplano S_1).

Operaciones 8:

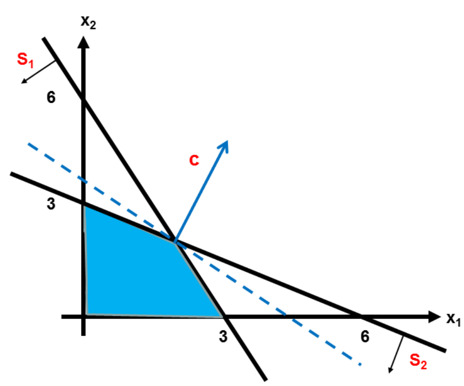


Figura 2.5: Maximización: Método de las curvas de nivel.

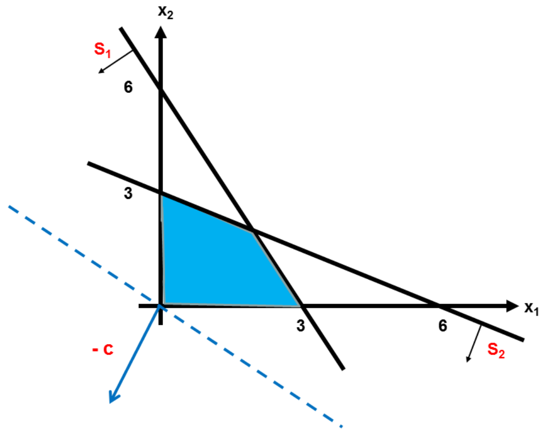


Figura 2.6: Minimización: Método de las curvas de nivel.

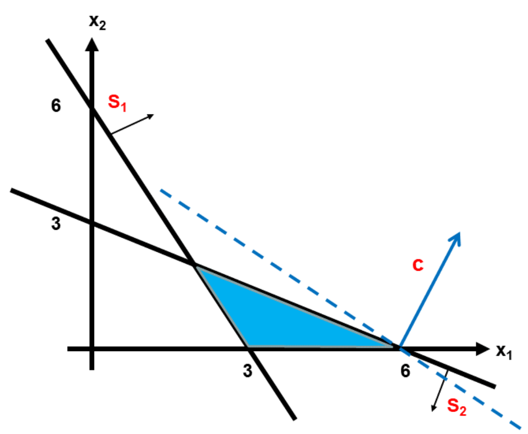


Figura 2.7: El punto óptimo depende de la región factible.

Solución:

1) El punto óptimo es $x^* = (2 \;\; 2)^\intercal$ con valor de la función objetivo $z^* = 10$.
2) El punto óptimo es $x^* = (0 \;\; 0)^\intercal$ con valor de la función objetivo $z^* = 0$.
3) El punto óptimo es $x^* = (6 \;\; 0)^\intercal$ con valor de la función objetivo $z^* = 12$.

General (Método de las curvas de nivel)

- *Objetivo:* Resolver un PL.
- Consideramos un PL de *dimensión dos,* donde se quiere *maximizar* $z(x) = c^\intercal x$.
- El método de las curvas de nivel consiste en los siguientes pasos:
  1) Representar el *conjunto factible.*
  2) Representar el *vector* c.
  3) Representar la mejor *curva de nivel* (recta)

  $$R_{k^*} \equiv c^\intercal x = k^*$$

  tal que:
     a) Es perpendicular a c.
     b) Intersecta con el conjunto factible,
     c) k^* tiene el valor máximo posible.
  4) Elegir un *vértice* factible que pertenezca a R_{k^*} y que llamamos x^*.
  5) Determinar las coordenadas exactas de x^* resolviendo el *sistema lineal* asociado a las restricciones que determinan x^*.
  6) El *punto óptimo* del PL es x^* y el *beneficio* óptimo es $z(x^*)$.
- En caso de *minimizar,* hacemos lo mismo salvo que en el anterior algoritmo debemos sustituir c por $-c$.

General (Otros métodos para resolver un PL)

- El método de las curvas de nivel tiene una finalidad didáctica.
- El *software* destinado a resolver PLs normalmente implementan el método del símplex y/o el método de punto interior.
- *Método del simplex:* Se parte de $x^{(1)}$, un vértice del conjunto factible, y si éste es óptimo paramos. En caso contrario, pasamos a $x^{(2)}$, un vértice adyacente donde se mejore la función objetivo, y si este punto es óptimo paramos. En caso contrario seguimos moviéndonos de vértice a vértice, mejorando cada vez la función objetivo, hasta que finalmente lleguemos a un vértice óptimo o declaremos que el problema es no acotado.
- *Método de punto interior:* Se parte de $x^{(1)}$, un punto interior del conjunto factible. De ahí pasamos a $x^{(2)}$, otro punto interior donde se mejore la función objetivo, y si este punto es óptimo paramos. En caso contrario seguimos moviéndonos de punto interior a punto interior, mejorando cada vez la función objetivo, hasta que finalmente lleguemos a un punto óptimo o declaremos que el problema es no acotado.

- En el libro de Rardin, Capítulos 5 y 6, se pueden encontrar más detalles.

Ejemplo 9 (Fabricación de trofeos)

Datos:

- Una compañía produce *trofeos* para fútbol y trofeos para tenis.

Tipo de trofeo	*Beneficio* (Euros/trofeo)	*Madera* (pies/trofeo)
Fútbol	12	4
Tenis	9	2

- Un *pie maderero* corresponde a una pieza cuadrada de 1 pie x 1 pie x 1 pulgada (aproximadamente 30 x 30 x 2,5 cm).
- En el almacén tenemos:
  - ▷ 1000 *adornos* para el trofeo de fútbol,
  - ▷ 1500 *adornos* para el trofeo de tenis,
  - ▷ 1750 *placas* para la inscripción de cada trofeo,
  - ▷ 4800 pies de *madera*
- Se supone que se vende toda la producción.

Objetivo:

- Esta compañía quiere *maximizar sus beneficios.* ¿Cuántos trofeos de cada tipo debería fabricar? ¿Cuánto sería su beneficio?
- Buscar una buena *solución por tanteo.*

Operaciones: Se deja al lector.

Ejemplo 10 (Fabricación de trofeos - continuación)

Objetivo:

1) *Formular* este problema de optimización como un PL.

2) *Resolverlo* mediante el ‘método de las curvas de nivel’.

Datos: Ver ejemplo anterior.

Operaciones 10:

Solución:

1) Ver apartado 1.
2) Ver apartado 2.

General (Tipos de puntos en un PL)

- *Tipos de puntos de un PL:*

  ▷ *Infactible:* Un punto es infactible si incumple al menos una restricción.

  ▷ *Factible:* Un punto es factible si cumple todas las restricciones. Tipos:

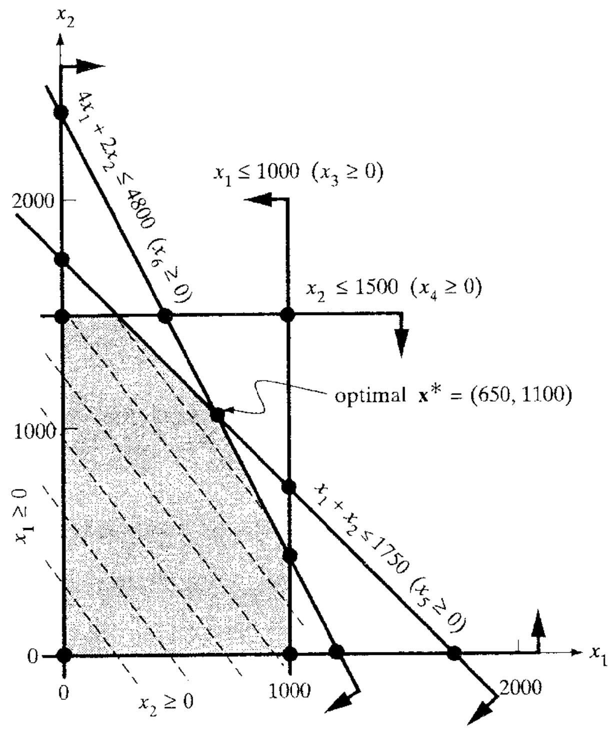


Figura 2.8: *Región factible.*

- *Frontera:* Al menos en una desigualdad ($\le$ ó $\ge$) el punto cumple la igualdad ($=$).
- *Interior:* Todas las desigualdades ($\le$ ó $\ge$) se cumplen de forma estricta ($<$ ó $>$) .
- *Extremo:* Son los *vértices* del conjunto factible.
- *Óptimo:* Punto factible con el mejor valor de la función objetivo.

• *Tipos de conjuntos:* Con los puntos anteriores, podemos formar:

  ▷ El conjunto factible.
  ▷ La frontera del conjunto factible.
  ▷ El interior del conjunto factible.
  ▷ El conjunto óptimo.

• *Propiedades de los puntos óptimos:*

  ▷ Todo punto óptimo de un PL está en la *frontera.*
  ▷ Un PL puede tener 0, 1 ó infinitos puntos óptimos:
  - *0 puntos óptimos:* Ocurre cuando el PL es infactible o con la función objetivo no acotada.
  - *1 punto óptimo:* Necesariamente es un punto extremo del conjunto factible (un vértice).
  - *Infinitos puntos óptimos:* Algunos de ellos son puntos extremos y el resto no.

2.4.2. Problema PL en formato estandar

Ejemplo 11 (Rardin, pag. 181, Ejercicio 5.3 b)

Objetivo: Pasar a *formato estándar* el siguiente PL.

Datos:

$$\begin{array}{lll} \min & z = 15(2x_1 + 8x_2) - 4x_3 & \\ \text{s. a.} & 2(10 - x_1) + x_2 + 5(9 - x_3) & \geq 10 \\ & x_1 - 2x_3 & \leq x_3 \\ & 2x_2 + 18x_3 & = 50 \\ & x_1, x_2, x_3 \geq 0 & \end{array}$$

Operaciones 11:

Solución: Ver apartado 'Operaciones'.

General

- *PL en formato estándar (versión expandida):*

$$\begin{array}{llll} \min & c_1x_1 + \ldots + c_nx_n & & \textit{(función objetivo)} \\ \text{s. a.} & a_{11}x_1 + \ldots + a_{1n}x_n & = b_1 & \textit{(m restricciones de igualdad)} \\ & \vdots & & \\ & a_{m1}x_1 + \ldots + a_{mn}x_n & = b_m & \\ & x_j \geq 0 \quad j = 1, 2, \ldots, n & & \textit{(restricciones de no negatividad)} \end{array}$$

donde:

x_j *variable de decisión,*

c_j *coste* de x_j,

a_{ij} *coeficiente* de la i-ésima restricción correspondiente a x_j,

b_i *término derecho* de la restricción i (en inglés: Right-hand side, RHS).

m número de restricciones,

n número de variables de decisión.

- *PL en formato estándar (versión sumatorio):*

$$
\begin{array}{lll}
\min & \displaystyle\sum_{j=1}^{n} c_j x_j & \\
\text{s. a.} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j = b_i & \text{para todo } i = 1, 2, \ldots, m \\
& x_j \geq 0 & \text{para todo } j = 1, 2, \ldots, n
\end{array}
$$

- *PL en formato estándar (versión matricial):* (Repaso de matrices: Rardin, pag. 185.)

$$
\begin{array}{ll}
\min & c'x \\
\text{s. a.} & Ax = b \\
& x \geq 0
\end{array}
$$

- *PL en formato estándar (con cotas superiores):* Un PL con variables acotadas superiormente por el *vector* u, tiene el siguiente formato:

$$
\begin{array}{ll}
\min & c'x \\
\text{s. a.} & Ax = b \\
& 0 \leq x \leq u
\end{array}
$$

2.5. Postoptimización

Las ideas principales de la postoptimización son:

- *Datos estimados:* Algunos de los parámetros que definen un PL más que datos reales son *estimaciones:*
  - ▷ Demanda *futura,* precio futuro, generación eólica, etc.
  - ▷ ¿Qué *confianza* nos merece el óptimo del PL calculado ante posibles errores de estimación de nuestros parámetros?
- *Análisis de sensibilidad:* Estudia cómo y cuánto varía el óptimo frente a *cambios* en los parámetros de un PL.
- *Problema dual:*
  - ▷ Para todo PL podemos definir su problema dual o *PL dual* que puede usarse como herramienta para la postoptimización.
  - ▷ El concreto, si v^* es el vector óptimo del PL dual, entonces v^* se puede usar para realizar el *análisis de sensibilidad* de la solución óptima del PL.
  - ▷ Los *paquetes informáticos* para resolver PLs además de x^*, el vector óptimo del PL, también facilitan v^* (son las llamadas *variables duales).*
  - ▷ *Nota:* Para más detalles se puede consultar el Capítulo 7 del siguiente libro: Rardin, R. L., 'Optimization in operations research', Editorial Prentice Hall, 1998.

Capítulo 3

Programación no lineal sin restricciones

3.1. Introducción

- El término *'programación no lineal'* es sinónimo de 'optimización no lineal'(ONL).
- Este capítulo está dedicado a la optimización no lineal *sin restricciones.*
- Una introducción a la optimización no lineal *con restricciones,* puede encontrarse en el Capítulo 14 del Rardin.
- Los modelos de optimización no lineal *sin restricciones,* tienen dos diferencias fundamentales comparados con la *programación lineal:*
  - ▷ Se optimiza una función *no lineal.*
  - ▷ No se consideran restricciones: el *conjunto factible es* $\mathbb{R}^n$.
  - ▷ El punto x^* puede ser óptimo local u óptimo global.
- Este capítulo está estructurado en la siguientes secciones:
  - ▷ *Modelos* de optimización no lineal.
  - ▷ *Métodos* de optimización unidimensional.
  - ▷ *Derivadas* y matrices definidas.
  - ▷ *Condiciones* de óptimo local.
  - ▷ *Métodos* de optimización multidimensional.

3.2. Modelos de optimización no lineal

Ejemplo 12 (Cercado de área máxima)

Objetivo: Un ganadero quiere *diseñar* el cercado rectangular de área máxima.

Datos:

- Dispone de un rollo de *valla* metálica de 100 m (suponemos que el ganadero quiere utilizar todo el rollo).

- Para hacer el cercado quiere aprovechar (total o parcialmente) un *muro* que mide 100 m.

Operaciones 12:

Solución: Ver apartado 'Operaciones'.

General

Problemas de Optimización No Lineal sin restricciones (ONL): El problema del cercado es un ejemplo de *problema ONL:*

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

es decir, 'mínimo' de $f(x)$ para cualquier x de $\mathbb{R}^n$.

Solución óptima: En un problema ONL, una solución x^* es óptima si

$$f(x^*) \leq f(x) \qquad \text{para cualquier } x \in \mathbb{R}^n.$$

¿Minimización o maximización?: Optimizar una función corresponderá a minimizarla o maximizarla según estemos en un contexto de *costes* o *beneficios,* respectivamente.

Equivalencia entre minimización y maximización: Todo problema de maximización es equivalente a uno de minimización (y viceversa):

$$\max_{x\in\mathbb{R}^n} f(x) = -\min_{x\in\mathbb{R}^n} \big(-f(x) \big).$$

Problemas de decisión: La mayoría de problemas de decisión pueden verse como un *problema de optimización* (tomamos una decisión óptima).

Criterio de decisión: Una decisión se toma en base a un criterio (costes, beneficios, etc.). Este criterio viene recogido en la *función objetivo* $f(x)$.

Variables de decisión:

- Tomar una decisión equivale a elegir el valor del *vector de decisión* x (vector que parametriza una decisión).
- Las componentes x_j del vector de decisión se denominan *variables de decisión.*

3.3. Métodos de optimización unidimensional

Objetivo: Aprender a *resolver* problemas ONL unidimensionales:

$$\min_{x\in\mathbb{R}} f(x).$$

Apartados:

- Método *analítico.*
- Método de *Newton.*

3.3.1. Método analítico

Ejemplo 13 (Cercado de área máxima - Cont.)

Objetivo: *Maximizar* la función área $A(h)$ analíticamente.

Datos: $A(h) = -2h^2 + 100h$ m^2

Operaciones 13:

Solución: $h^* = 25$m maximiza $A(h)$ y $A^* = 1250$ m^2.

General (Método de ONL analítico - unidimensional)

Objetivo: Resolución analítica de

$$\min_{x \in \mathbb{R}} f(x).$$

Pasos:

1) Resolvemos la *ecuación* $f'(x) = 0$ y obtenemos x^*.
2) Si $f''(x^*) > 0$, x^* es un *minimizador* (local) de f.
3) Si $f''(x^*) < 0$, x^* es un *maximizador* (local) de f.
4) Si $f''(x^*) = 0$, *repetir el proceso* anterior resolviendo $f'''(x) = 0$ y analizando el signo de la cuarta derivada....

3.3.2. Ejercicios

1) Queremos fabricar latas de Coca-Cola de 333 cm^3 minimizando la cantidad de aluminio utilizado, es decir, queremos minimizar el área de la lata. Ayuda: área del círculo $= \pi r^2$, longitud de la circunferencia $= 2\pi r$, volumen del cilindro $= \pi r^2\ h$.

   a) Calcula el área de la anterior lata $A(r, h)$ en función de su altura h y de su radio r.
   b) Calcula el área de la anterior lata $A(r)$ en función sólo de su radio r.
   c) Representa $A(r)$.
   d) Calcula las dimensiones (h^*, r^*) de la lata de menor área. ¿Que área tiene?

3.3.3. Método de Newton (unidimensional)

General (Métodos de ONL numéricos)

- Cuando la optimización analítica es *difícil o imposible,* recurrimos a métodos de optimización *numéricos* (ejem. método de Newton).
- Además del método de Newton existen otros métodos numéricos, como por ejemplo el 'Método de la *sección áurea'* (Rardin, pag. 726).
- El método de Newton es uno de los más *rápidos.*

General (Método de Newton - unidimensional)

Método de Newton (idea): Es un método iterativo que para calcular x_{k+1}, aproxima $f(x)$ en x_k mediante una *parábola* (ver Fig. 3.1).

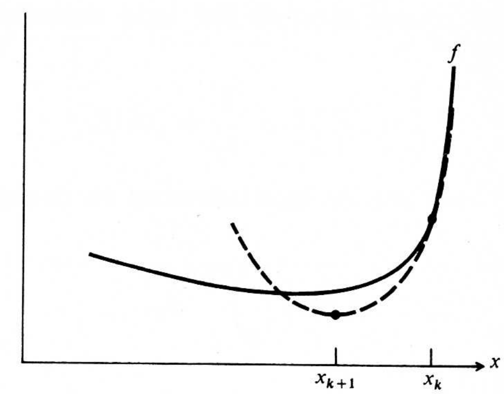


Figura 3.1: Idea intuitiva del *método de Newton:* x_{k+1} es el punto mínimo de la parábola que mejor aproxima $f(x)$ en x_k. El punto x_{k+2} se obtiene de forma análoga.

Método de Newton: Resolución numérica de

$$\min_{x\in\mathbb{R}} f(x),$$

donde $f(x)$ es *dos veces derivable.* Pasos:

1) Seleccionamos:
   - Un *punto inicial* x_0.
   - Una *tolerancia* $\epsilon > 0$ para el criterio de parada.
2) Mientras se cumpla $\mid f'(x_k) \mid > \epsilon$, *iterar* según:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

3) Nota: Este mismo método nos serviría para resolver un problema de *maximización* (sin hacer ningún cambio), pues el método de Newton lo que obtiene es un punto de derivada 0. Hay que calcular $f''(x^*)$ para saber si x^* es un máximo o un mínimo.

Ejemplo 14 (Método de Newton)

Objetivo: *Minimizar* la función $f(x)$ mediante el método de Newton.

$$f(x) = x^4 + 4x^3 - 11x^2 - 36x + 118$$

Datos:

- Tomamos arbitrariamente $x_0 = 1$.
- Tomamos arbitrariamente $\epsilon = 0{,}0001 = 10^{-4}$.

Operaciones 14:

Solución:

- $x^* = 2,0000$ es un *minimizador (local)* de $f(x)$.
- $f(2,0000) = 50$ es un *mínimo (local)* de $f(x)$.

3.3.4. Ejercicios

1) (Ref. Rardin 13-17). Consideramos la función

$$f(x) = 10x + \frac{70}{x}.$$

Minimiza $f(x)$ para $x > 0$, utilizando los siguientes métodos:

a) Método analítico.

b) Método de Newton. Tomar el punto inicial $x_0 = 1$ y una tolerancia $\epsilon = 10^{-4}$ para el criterio de parada. Se recomienda el uso de una hoja de cálculo.

3.4. Derivadas y matrices definidas

En esta sección veremos los apartados:

- Primera y segunda derivada.
- Matrices definidas.

3.4.1. Primera y segunda derivada

Ejemplo 15

Objetivo: *Interpretar* la primera y segunda derivada de $f(x)$ en $\overline{x} = 4$ (Fig. 3.2).

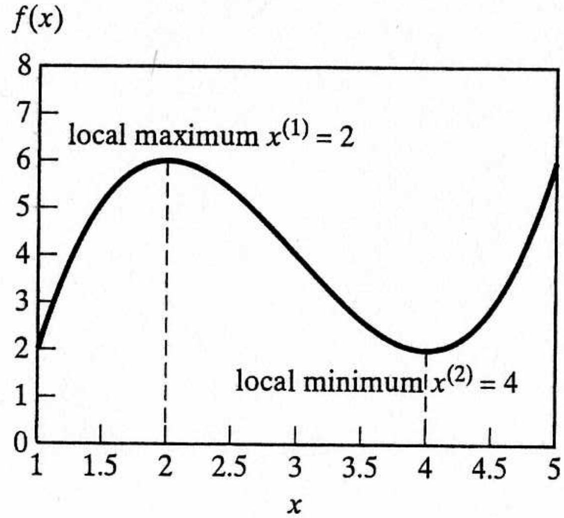


Figura 3.2: $f(x) = x^3 - 9x^2 + 24x - 14$.

Datos:

$$f(x) = x^3 - 9x^2 + 24x - 14.$$

Operaciones 15:

Solución: La *pendiente* de la recta tangente a $f(x)$ en $\overline{x} = 4$, es 0 y su *curvatura* es positiva (el signo de la curvatura viene dado por el signo de la segunda derivada).

General (Curvatura de una función unidimensional)

- *Curvatura y segunda derivada coinciden en signo:* Dada la función $y = f(x)$, su *curvatura* k viene dada por:

$$k = \frac{y''}{(1 + (y')^2)^{2/3}}.$$

- Notar que cuando $y' = 0$, la curvatura coincide con y''.
- Para optimizar lo que realmente nos interesa es el *signo* de la curvatura (siempre coincide con el signo de y'').
- Por ejemplo $f(x) = x^2$ es una función de *curvatura positiva* en todo $\mathbb{R}$ pues $f''(x) = 2$.

Ejemplo 16

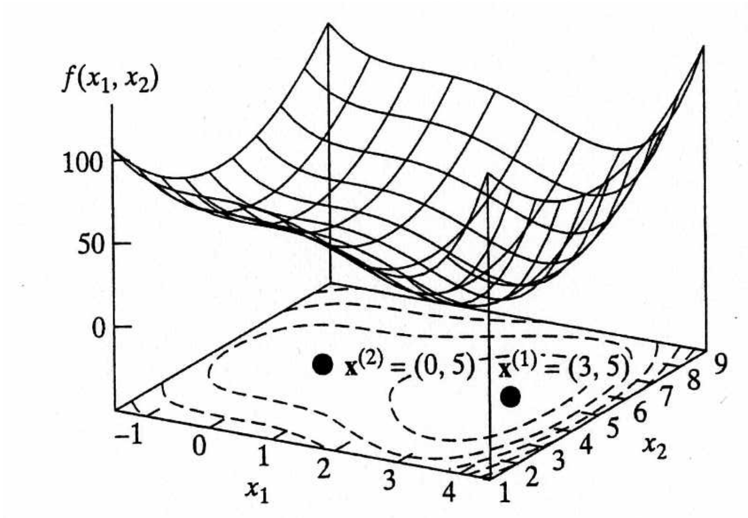


Figura 3.3: $f(x) = 40 + x_1^3(x_1 - 4) + 3(x_2 - 5)^2$.

Datos: En la Fig. 3.3 tenemos la *representación* de $f(x)$.

$$f(x) = 40 + x_1^3(x_1 - 4) + 3(x_2 - 5)^2.$$

Objetivo: Calcular el vector *gradiente* y la matriz *hessiana* de $f(x)$ en

$$\overline{x} = (3, 5)^{\mathsf{T}}$$

Operaciones 16:

Solución: Ver apartado 'Operaciones'.

General (Gradiente y matriz hessiana)

Gradiente de $f : \mathbb{R}^n \to \mathbb{R}$:

- *Generaliza* el concepto de derivada.
- Se *calcula:*

$$\nabla f(\overline{x}) = \left(\frac{\partial f}{\partial x_1} \quad \dots \quad \frac{\partial f}{\partial x_n} \right)^{\mathsf{T}}_{x=\overline{x}}.$$

- Describe la *pendiente* o razón de cambio de f en las direcciones principales (ejes coordenados).
- En el caso unidimensional ($n = 1$) el gradiente corresponde a la *derivada:*

$$\nabla f(\overline{x}) = \left(\frac{df}{dx} \right)_{x=\overline{x}} = f'(\overline{x}).$$

Matriz hessiana de $f : \mathbb{R}^n \to \mathbb{R}$:

- Generaliza el concepto de *segunda derivada.*
- Se calcula:

$$H(\overline{x}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial^2 x_1} & \dots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dots & \dfrac{\partial^2 f}{\partial^2 x_n} \end{pmatrix}_{x=\overline{x}}.$$

- Nota: Otra *notación* muy usada para $H(x)$ es $\nabla^2 f(x)$.
- En el caso unidimensional ($n = 1$) la matriz hessiana corresponde a la *segunda derivada:*

$$H(\overline{x}) = \left(\frac{d^2 f}{d^2 x} \right)_{x=\overline{x}} = f''(\overline{x}).$$

- La hessiana es una matriz *simétrica.*

3.4.2. Matrices definidas positivas o negativas

- En optimización unidimensional, las condiciones de óptimo local requieren analizar el signo de la *segunda derivada.*

- De forma análoga, en optimización multidimensional, las condiciones de óptimo local requieren analizar la *hessiana.*

- Analizaremos si la hessiana es *definida* positiva, definida negativa o indefinida.

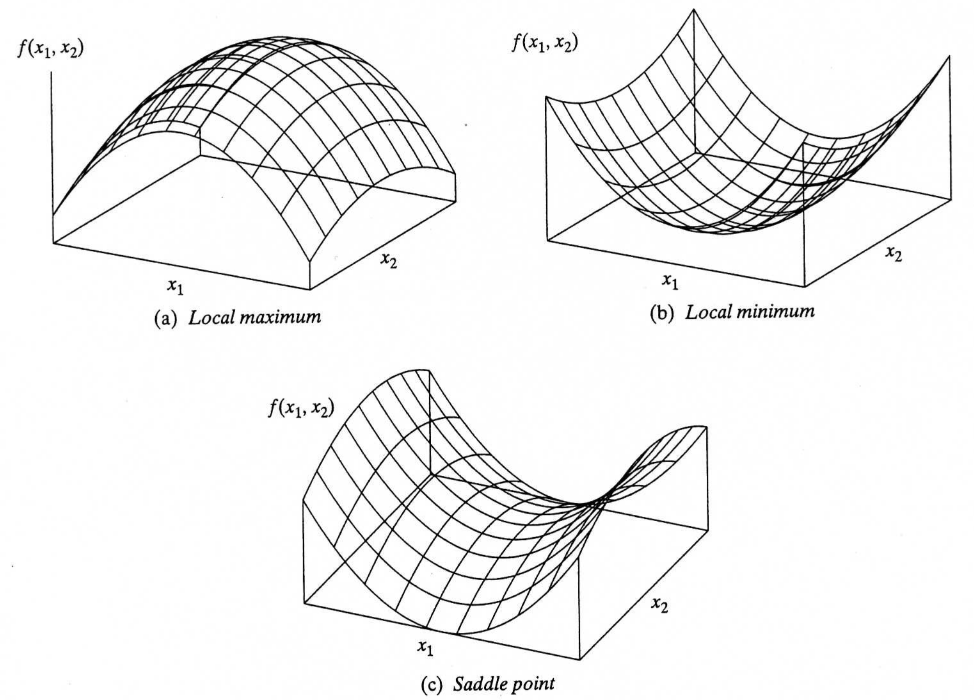


Figura 3.4: Paraboloides asociados a una matriz A: (a) definida *negativa,* (b) definida *positiva,* (c) *indefinida.*

General (Geometría de las funciones cuadráticas)

- *Geometría:* Dada la matriz A *simétrica* respecto la diagonal principal y la función cuadrática $f(x) = x^\mathsf{T} A\, x$:
  - ▷ Si A es *definida positiva,* la gráfica de f es un paraboloide elíptico con mínimo *('bol').*
  - ▷ Si A es *definida negativa,* la gráfica de f es un paraboloide elíptico con máximo *('cúpula').*
  - ▷ Si A es *indefinida,* la gráfica de f es un paraboloide hiperbólico *('silla de montar').* No tiene ni máximo ni mínimo.
  - ▷ Ver Fig. 3.4.

Ejemplo 17

Objetivo: *Analizar* si la matriz A es definida positiva.

Datos:

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

Una matriz simétrica A es definida *positiva* si para cualquier vector x no nulo,

$$x^\mathsf{T} A\, x > 0.$$

Operaciones 17:

Solución: A es *definida positiva* ($A \succ 0$).

General (Matrices definidas (definición))

Matriz diagonal: Dada una matrix $A_{n\times n}$, decimos que es *diagonal* si los elementos no diagonales, a_{ij} con $i \neq j$, son cero. Ejemplo:

$$M = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Matriz simétrica: Dada una matrix $A_{n\times n}$, decimos que es *simétrica* si:

$$A = A^{\intercal}.$$

Ejemplo: La anterior matriz M es simétrica.

Matriz definida positiva $A \succ 0$: Dada una matrix $A_{n\times n}$ *simétrica,* decimos que es *definida positiva* si para cualquier vector x no nulo:

$$x^{\intercal} A\, x > 0.$$

Matriz semidefinida positiva $A \succeq 0$: Análogamente, $x^{\intercal} A\, x \geq 0$.

Matriz definida negativa $A \prec 0$: Análogamente, $x^{\intercal} A\, x < 0$.

Matriz semidefinida negativa $A \preceq 0$: Análogamente, $x^{\intercal} A\, x \leq 0$.

Nota: Toda matriz definida positiva (negativa) es también, semidefinida positiva (negativa).

Matriz indefinida $A \gtrless 0$: Dada una matrix $A_{n\times n}$ *simétrica,* decimos que es *indefinida* si no es semidefinida.

Nota: Toadas las anteriores definiciones solo se aplican a matrices *simétricas.*

Ejemplo 18

Objetivo: *Analizar* si la matriz A es definida positiva.

Datos:

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

Operaciones 18:

Solución: Por lo tango, A es *definida positiva* ($A \succ 0$).

General (Clasificación de matrices diagonales)

- Una matriz A, *simétrica* y diagonal, es:

1) *Definida positiva,* si y sólo si, sus elementos diagonales son estrictamente positivos.

$$A \succ 0 \Leftrightarrow a_{ii} > 0, \qquad i = 1, \ldots, n.$$

2) *Semidefinida positiva*

$$A \succeq 0 \Leftrightarrow a_{ii} \geq 0, \qquad i = 1, \ldots, n.$$

3) *Definida negativa,* si y sólo si, sus elementos diagonales son estrictamente negativos.

$$A \prec 0 \Leftrightarrow a_{ii} < 0, \qquad i = 1, \ldots, n.$$

4) *Semidefinida negativa*

$$A \preceq 0 \Leftrightarrow a_{ii} \leq 0, \qquad i = 1, \ldots, n.$$

5) *Indefinida,* en el resto de casos.

Ejemplo 19

Objetivo: *Analizar* si la matriz A es definida positiva.

Datos:

$$A = \begin{pmatrix} 5 & -2 \\ -2 & 3 \end{pmatrix}.$$

- *Determinante* de orden 2:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Operaciones 19:

Solución: A es *definida positiva* pues todos sus determinantes principales son estrictamente positivos.

General (Clasificación de matrices (caso general))

Determinantes de orden 2:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Determinante principal de orden k de A (escribiremos Δ_k): Es el *determinante* de las primeras k filas y k columnas de A, empezando a contar por el elemento 'arriba-izquierda' (a_{11}).

Clasificación de matrices (caso general): Una matriz *simétrica* $A_{n\times n}$ es:

1) *Definida positiva,* si y sólo si, sus determinantes principales son estrictamente positivos.

$$A \succ 0 \Leftrightarrow \Delta_k > 0, \qquad k = 1, \ldots, n.$$

2) *Semidefinida positiva*

$$A \succeq 0 \Leftrightarrow \Delta_k \geq 0, \qquad k = 1, \ldots, n.$$

3) *Definida negativa,* si y sólo si, sus determinantes principales son no nulos, alternados en signo y, además,
$\Delta_1 < 0$.

4) *Semidefinida negativa,* si y sólo si, sus determinantes principales son alternados en signo (pueden ser nulos) y, además, $\Delta_1 \leq 0$.

5) *Indefinida,* en el resto de casos.

6) *Nota:* Este método también sirve para matrices *diagonales,* aunque es mejor usar su método especializado.

3.5. Condiciones de óptimo local

Objetivo:

- Estudiamos las condiciones de *óptimo local* para el problema ONL (Rardin, pag. 737):

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

- Solo tenemos garantías que un óptimo local sea *óptimo global* en algunos casos especiales (ejem. si $f(x)$ es convexa).
- En esta sección nos limitamos a estudiar las condiciones de *óptimo local.*
- Nota: En un PL, z^* es un óptimo global, pues un PL es un problema de optimización convexa.

Apartados:

- Condiciones de optimalidad de *primer* orden.
- Condiciones de optimalidad de *segundo* orden.

3.5.1. Condiciones de primer orden

Ejemplo 20 (Puntos estacionarios)

Objetivo: *Calcular* los puntos estacionarios de

$$f(x) = 40 + x_1^3(x_1 - 4) + 3(x_2 - 5)^2,$$

representada en la Fig. 3.5.

Datos:

- Un punto $\overline{x}$ es *estacionario* si

$$\nabla f(\overline{x}) = (0, \ldots, 0)^\mathsf{T}.$$

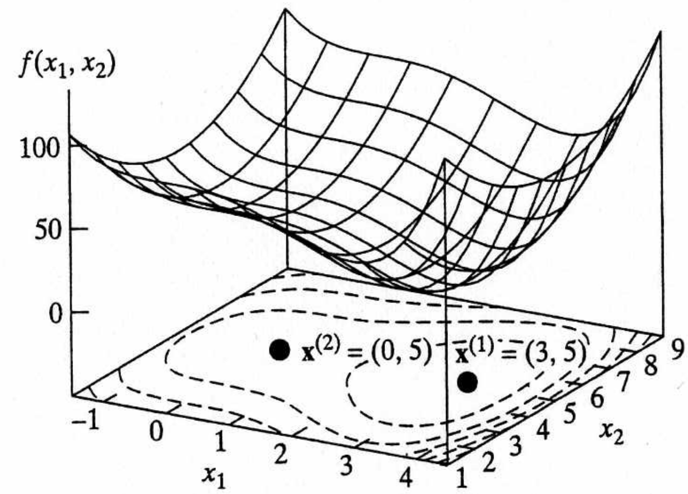


Figura 3.5: $f(x) = 40 + x_1^3(x_1 - 4) + 3(x_2 - 5)^2$.

- Ya vimos que

$$\nabla f(x) = \begin{pmatrix} 4x_1^3 - 12x_1^2 & 6(x_2 - 5) \end{pmatrix}^{\mathsf{T}}.$$

Operaciones 20:

Solución: $f(x)$ tiene dos puntos *estacionarios:*

$$\overline{x}^a = (0,\ 5) \quad \text{y} \quad \overline{x}^b = (3,\ 5).$$

General (Condición de optimalidad de primer orden)

Punto estacionario (definición): Dada la función diferenciable $f(x)$, un punto $\overline{x}$ es *estacionario,* si y sólo si,

$$\nabla f(\overline{x}) = 0.$$

Condición necesaria de optimalidad de primer orden (CNO1): Si x^* es un *óptimo local* irrestringido de una función diferenciable, entonces (necesariamente) x^* es un *punto estacionario.*

3.5.2. Condiciones de segundo orden

Ejemplo 21 (Máximos, mínimos y puntos de silla - I)

Objetivo: Verificar si el punto *estacionario*

$$\overline{x}^b = (3,\ 5)$$

corresponde a un máximo local, mínimo local o a un punto de silla de $f(x)$.

Datos:

- $f(x) = 40 + x_1^3(x_1 - 4) + 3(x_2 - 5)^2$.
- Dado un punto *estacionario* $\overline{x}$:
  - ▷ Si $H(\overline{x}) \succ 0$, entonces $\overline{x}$ es un *mínimo* local.
  - ▷ Si $H(\overline{x}) \prec 0$, entonces $\overline{x}$ es un *máximo* local.
  - ▷ Si $H(\overline{x})$ es indefinida, entonces $\overline{x}$ es un *punto de silla.*
- Ya vimos que
$$H(x) = \begin{pmatrix} 12x_1^2 - 24x_1 & 0 \\ 0 & 6 \end{pmatrix}$$

Operaciones 21:

Solución: $\overline{x}^a = (3,\ 5)$ es un *mínimo* local.

General (Condiciones de optimalidad de segundo orden)

Condición necesaria de optimalidad de segundo orden (CNO2): Si x^* es un *óptimo local* irrestringido de una función dos veces diferenciable, entonces (necesariamente) x^* es un punto de *hessiana semidefinida:* positiva si mínimo y negativa si máximo.

Condición suficiente de optimalidad de segundo orden (CSO2): Es suficiente que x^* sea un punto *estacionario* con *hessiana definida* para garantizar que es un *óptimo local* irrestringido: mínimo si definida positiva y máximo si definida negativa.

Punto de silla (definición): $\overline{x}$, punto *estacionario* de $f(x)$, es un punto de silla si no es ni un máximo local ni un mínimo local.

Punto de silla (condición suficiente de segundo orden): Es suficiente que $\bar{x}$ sea un punto *estacionario* con *hessiana indefinida* para garantizar que es un *punto de silla.*

Ejemplo 22 (Máximos, mínimos y puntos de silla - II)

Objetivo: *Verificar* si el punto

$$\overline{x}^a = (0,\ 5)$$

corresponden a un máximo local, mínimo local o a un punto de silla de $f(x)$.

Datos:

- $f(x) = 40 + x_1^3(x_1 - 4) + 3(x_2 - 5)^2$.
- Tenemos que

$$H(x) = \begin{pmatrix} 12x_1^2 - 24x_1 & 0 \\ 0 & 6 \end{pmatrix}$$

Operaciones 22:

Solución: Con la información de primer y segundo orden, podemos afirmar que $\overline{x}^a = (0,\ 5)$ puede ser un *mínimo* local, un *máximo* local o un *punto de silla.*

General (Métodos para la resolución del problema ONL)

Método analítico: Resolución analítica del *problema ONL*

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

1) Resolvemos la *ecuación* vectorial $\nabla f(x) = (0, \ldots, 0)^\intercal$ para obtener x^*.

2) Analizamos la hessiana:

   a) Si $H(x^*) \succ 0$, x^* es un *minimizador* de f.
   b) Si $H(x^*) \prec 0$, x^* es un *maximizador* de f.
   c) Si $H(x^*)$ es indefinida, x^* es un *punto de silla* de f.
   d) Si $H(x^*)$ es *semidefinida,* tendremos que recurrir a otros argumentos (convexidad de f, etc.) para poder determinar el carácter de x^*.

Métodos numéricos:

- Cuando la optimización analítica es *difícil o imposible,* recurrimos a métodos de optimización numéricos.
- En este capítulo estudiaremos el método del *gradiente* y el método de *Newton* (caso multidimensional).

3.6. Métodos de optimización multidimensional

Objetivo: Resolver el problema ONL multidimensional

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

mediante métodos *numéricos* (Rardin, pag. 757).

Apartados:

- Método del *gradiente.*
- Método de *Newton* (multidimensional).

3.6.1. Método del gradiente

Idea: Para un problema ONL mejoramos el punto actual x^k siguiendo:

- La dirección de *máximo descenso* (si minimizamos):

$$d^k = -\nabla f(x^k).$$

- La dirección de *máximo ascenso* (si maximizamos):

$$d^k = +\nabla f(x^k).$$

ALGORITHM 13D: GRADIENT SEARCH

Step 0: Initialization. Choose any starting solution $\mathbf{x}^{(0)}$, pick stopping tolerance $\epsilon > 0$, and set solution index $t \leftarrow 0$.

Step 1: Gradient. Compute objective function gradient $\nabla f(\mathbf{x}^{(t)})$ at current point $\mathbf{x}^{(t)}$.

Step 2: Stationary Point. If gradient norm $||\nabla f(\mathbf{x}^{(t)})|| < \epsilon$, stop. Point $\mathbf{x}^{(t)}$ is sufficiently close to a stationary point.

Step 3: Direction. Choose gradient move direction

$$\Delta\mathbf{x}^{(t+1)} \leftarrow \pm\nabla f(\mathbf{x}^{(t)})$$

(+ for maximize and − for minimize).

Step 4: Line Search. Solve (at least approximately) corresponding one-dimensional line search

$$\text{max or min } f(\mathbf{x}^{(t)} + \lambda\Delta\mathbf{x}^{(t+1)})$$

to compute λ_{t+1}.

Step 5: New Point. Update

$$\mathbf{x}^{(t+1)} \leftarrow \mathbf{x}^{(t)} + \lambda_{t+1}\Delta\mathbf{x}^{(t+1)}$$

Step 6: Advance. Increment $t \leftarrow t+1$, and return to Step 1.

Nota 1: En el anterior algoritmo, también podemos usar una longitud de paso *constante* y pequeña $\lambda_{t+1} > 0$ (por ejemplo $\lambda_{t+1} = 0,1$).

Nota 2: El paso 4 de método del gradiente se denomina 'búsqueda lineal exacta' cuando se calcula un λ^* óptimo.

Ejemplo 23

Objetivo: Realizar *una iteración* del método del gradiente para minimizar $f(x)$, tomando el punto inicial $x^0 = (0\ 0)^\intercal$ y $\epsilon = 0,1$.

Datos:

$$f(x) = (x_1 - 1)^2 + (x_2 - 3)^2.$$

Operaciones 23:

Solución:

- El nuevo *punto* dado por el método del gradiente es

$$x^1 = (1, 3)^\intercal.$$

- La función *objetivo* ha pasado de valer $f(x^0) = 10$ a valer $f(x^1) = 0$.
- Es fácil ver que x^1 es ya un punto *óptimo*.
- Esto no es lo habitual. Normalmente, el método del gradiente requerirá varias *iteraciones* para obtener un punto óptimo.

3.6.2. Método de Newton (multidimensional)

- *Idea intuitiva:* x^{k+1} es el punto mínimo de la *aproximación cuadrática* que mejor aproxima $f(x)$ en x^k. El punto x^{k+2} se obtiene de forma análoga.

- Notar que en el método de Newton en *contraste* con el método del Gradiente:

  ▷ Resolvemos un sistema de ecuaciones donde aparece

  $$-\nabla f(x^k),$$

  tanto para *minimizar* como para *maximizar* (ver Paso 3)

  ▷ La *longitud de paso* siempre es 1 (ver Paso 4).

ALGORITHM 13E: NEWTON'S METHOD

Step 0: Initialization. Choose any starting solution $\mathbf{x}^{(0)}$, pick stopping tolerance $\epsilon > 0$, and set solution index $t \leftarrow 0$.

Step 1: Derivatives. Compute objective function gradient $\nabla f(\mathbf{x}^{(t)})$ and Hessian matrix $\mathbf{H}(\mathbf{x}^{(t)})$ at current point $\mathbf{x}^{(t)}$.

Step 2: Stationary Point. If $||\nabla f(\mathbf{x}^{(t)})|| < \epsilon$, stop. Point $\mathbf{x}^{(t)}$ is sufficiently close to a stationary point.

Step 3: Newton Move. Solve the linear system

$$\mathbf{H}(\mathbf{x}^{(t)})\Delta\mathbf{x} = -\nabla f(\mathbf{x}^{(t)})$$

for Newton move $\Delta\mathbf{x}^{(t+1)}$.

Step 4: New Point. Update

$$\mathbf{x}^{(t+1)} \leftarrow \mathbf{x}^{(t)} + \Delta\mathbf{x}^{(t+1)}$$

Step 5: Advance. Increment $t \leftarrow t+1$, and return to Step 1.

Ejemplo 24

Objetivo: Realizar una *iteración* del método de Newton para minimizar $f(x)$, tomando el punto inicial $x^0 = (0\ 1)^\intercal$ y $\epsilon = 0,1$.

Datos:

$$f(x_1, x_2) = (x_1 + 1)^4 + x_1 x_2 + (x_2 + 1)^4.$$

Operaciones 24:

Solución:

- El nuevo *punto* dado por el método de Newton es

$$(-0{,}36, +0{,}34).$$

- La función *objetivo* ha pasado de valer $f(x^0) = 17$ a valer $f(x^1) = 3{,}28$.
- Notar que en el método de Newton no hay que calcular la *longitud de paso* λ^k pues siempre vale 1.

i	Number, p_i	Cost, q_i	i	Number, p_i	Cost, q_i	i	Number, p_i	Cost, q_i
1	19	7.9	5	5	19.5	9	14	9.2
2	2	25.0	6	6	13.0	10	17	6.3
3	9	13.1	7	3	17.8	11	1	42.0
4	4	17.4	8	11	8.0	12	20	6.6

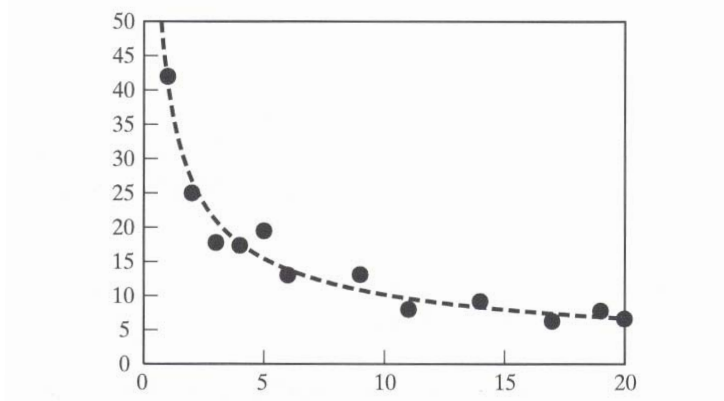


Figura 3.6: Nube de puntos $(p_i, q_i) =$ (Nº de estaciones, Precio por estación (miles de euros)).

3.6.3. M. del gradiente versus M. de Newton

Ejemplo 25 (Política de precios para las estaciones de trabajo IBM (Rardin, pag. 718))

Objetivo: Ajustar la curva $q = r(p)$, es decir, 'precio por estación' en función del 'número de estaciones', para facilitar futuras ofertas de venta (ver Fig. 3.6).

Datos:

- Este es un ejemplo de 'problema de regresión no lineal'.
- Notación:
  - ▷ i es el *índice* para el cliente $i \in \mathcal{I} = \{1, \ldots, 12\}$.
  - ▷ p_i es el *número de estaciones* que compró el cliente $i \in \mathcal{I}$.
  - ▷ q_i (miles de euros), es el *precio* por estación de trabajo IBM que se cobró al cliente $i \in \mathcal{I}$.
  - ▷ $\|v\|$ es la *norma o módulo* al cuadrado del vector $v = (v_1, \ldots, v_I)^\intercal$:
  $$\|v\|^2 = \sum_{i \in \mathcal{I}} v_i^2.$$
  - ▷ *Propiedad:* Si $g(x) = a^x$ entonces $g'(x) = a^x \ln(a)$.

- Consideramos la forma de la nube de puntos en la Fig. 3.6:
  - ▷ Es una función no lineal, decreciente y con asíntotas en los ejes OX y OY.
  - ▷ Así, la *función exponencial*
  $$r(p_i) = x_1 \cdot p_i^{x_2}$$
  parece una buena familia de funciones candidata ((x_1, x_2) son datos desconocidos y x_2 es negativo).

▷ Explicación: Para $(x_1, x_2) = (1,\ -1)$ tenemos

$$r(1, -1 \mid p_i) = 1 \cdot p_i^{-1} = \frac{1}{p_i}.$$

- Dicho de otra manera, $r(p_i)$ debe cumplir las 12 condiciones de regresión:

$$q_i - r(p_i) \approx 0 \qquad\qquad i \in \mathcal{I} \tag{3.1}$$

Operaciones 25:

Solución: En el siguiente ejemplo veremos que, mediante el método del gradiente, se obtiene el siguiente *punto óptimo*

$$x_1^* = 40{,}66 \quad \text{y} \quad x_2^* = -0{,}6020,$$

y por lo tanto la *función de precios óptima* es

$$r(x_1^*, x_2^* \mid p_i) \quad = \quad 40{,}66 p_i^{-0{,}6020}.$$

Ejemplo 26 (Resolución por el método del gradiente)

t	$\mathbf{x}^{(t)}$	$f(\mathbf{x}^{(t)})$	$\nabla f(\mathbf{x}^{(t)})$	$\|\nabla f(\mathbf{x}^{(t)})\|$	λ_{t+1}
0	(32.00, –0.4000)	174.746	(–6.24, 1053.37)	1053.39	0.00007
1	(32.00, –0.4687)	141.138	(–23.06, –0.14)	23.06	0.10558
2	(34.44, –0.4540)	112.599	(–4.57, 759.60)	759.61	0.00007
3	(34.44, –0.5078)	93.297	(–16.28, –0.10)	16.28	0.11303
4	(36.28, –0.4962)	78.123	(–3.34, 530.01)	530.02	0.00008
5	(36.28, –0.5365)	67.897	(–11.34, –0.07)	11.34	0.11970
6	(37.63, –0.5278)	60.133	(–2.33, 365.74)	365.75	0.00008
7	(37.63, –0.5571)	54.932	(–7.78, –0.05)	7.78	0.12905
8	(38.64, –0.5512)	51.006	(–1.48, 251.17)	251.17	0.00008
9	(38.64, –0.5722)	48.428	(–5.19, –0.03)	5.19	0.13926
10	(39.36, –0.5684)	46.548	(0.88, 168.22)	168.22	0.00009
11	(39.36, –0.5829)	45.348	(–3.34, –0.02)	3.34	0.14737
12	(39.85, –0.5806)	44.522	(–0.51, 108.66)	108.66	0.00009
13	(39.85, –0.5902)	44.007	(–2.10, –0.01)	2.10	0.15220
14	(40.17, –0.5888)	43.672	(–0.30, 68.07)	68.07	0.00009
15	(40.17, –0.5949)	43.466	(–1.29, 0.00)	1.29	0.16435
16	(40.38, –0.5941)	43.329	(–0.14, 42. 34)	42.34	0.00009
17	(40.38, –0.5980)	43.248	(–0.76, 0.00)	0.76	0.15652
18	(40.50, –0.5975)	43.203	(–0.10, 24.60)	24.60	0.00009
19	(40.50, –0.5997)	43.175	(–0.46, 0.00)	0.46	0.14805
20	(40.57, –0.5994)	43.160	(–0.08, 14.77)	14.77	0.00009
21	(40.57, –0.6007)	43.150	(–0.29, 0.00)	0.29	0.15527
22	(40.62, –0.6005)	43.143	(–0.04, 9.37)	9.37	0.00009
23	(40.62, –0.6014)	43.139	(–0.18, 0.00)	0.18	0.16376
24	(40.65, –0.6013)	43.137	(–0.02, 5.78)	5.78	0.00009
25	(40.65, –0.6018)	43.135	(–0.11, 0.00)	0.11	0.15266
26	(40.66, –0.6017)	43.134	(–0.02, 3.37)	3.37	0.00009
27	(40.66, –0.6020)	43.134	(–0.06, 0.00)	0.06	Stop

Tabla 3.1: *Iteraciones* del método del gradiente.

Objetivo: *Resolver* mediante el método del gradiente el problem ONL del ejemplo anterior:

$$\min_{x \in S} f(x_1, x_2), \tag{3.2}$$

donde

$$f(x_1, x_2) = \sum_{x \in \mathcal{I}} \big(q_i - r(p_i) \big)^2,$$

(q_i, p_i son datos conocidos).

Datos:

- Nota: En realidad resolveremos (3.2) con $x \in \mathbb{R}^2$ en vez de $x \in \mathcal{S}$.
- Tomamos el punto inicial $x^0 = (32, -0, 4)^\intercal$.
- Tomamos la tolerancia $\epsilon = 10^{-1}$.

Operaciones 26:

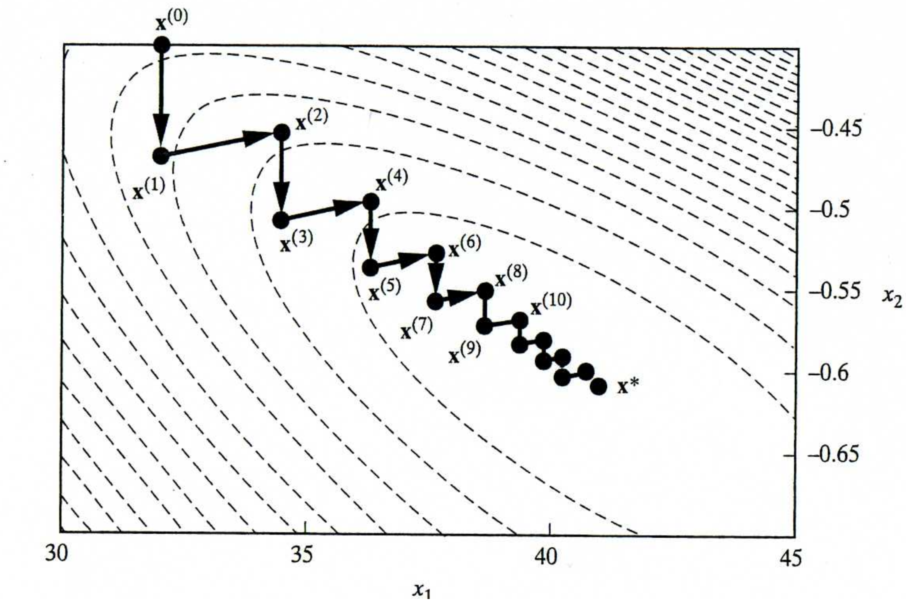


Figura 3.7: Representación de las *iteraciones* del método del gradiente.

Solución: Mediante el método del gradiente, hemos obtenido el siguiente *punto óptimo*

$$x_1^* = 40{,}66 \qquad x_2^* = -0{,}60, \qquad f(x^*) = 43,13$$

y por lo tanto la *función de precios óptima* $r^*(p_i)$ es:

$$r^*(p_i) \quad = \quad r(x_1^*, x_2^* \mid p_i) \quad = \quad 40{,}66 p_i^{-0{,}60}.$$

Ejemplo 27 (Resolución por el método de Newton)

Objetivo: *Resolver* mediante el método de Newton el problem ONL del ejemplo anterior:

$$\min_{x\in\mathcal{S}} f(x_1, x_2),$$

donde

$$f(x_1, x_2) = \sum_{x\in\mathcal{I}} \left(q_i - x_1\, p_i^{x_2} \right)^2,$$

(q_i, p_i son datos conocidos).

Datos:

- Tomamos el punto inicial $x^0 = (32, -0,4)^\intercal$.
- Tomamos la tolerancia $\epsilon = 10^{-2}$.

Operaciones 27:

t	$\mathbf{x}^{(t)}$	$f(\mathbf{x}^{(t)})$	$\nabla f(\mathbf{x}^{(t)})$	$\mathbf{H}(\mathbf{x}^{(t)})$	$\lVert\nabla f(\mathbf{x}^{(t)})\rVert$	$\Delta\mathbf{x}^{(t+1)}$
0	(32.00, –0.4000)	174.746	(–6.240, 1053)	$\begin{pmatrix} 7.13 & 293.99 \\ 293.99 & 18{,}817 \end{pmatrix}$	1053.4	(8.956, –0.1959)
1	(40.96, –0.5959)	43.820	(2.347, 116.3)	$\begin{pmatrix} 4.86 & 166.21 \\ 166.21 & 11{,}564 \end{pmatrix}$	116.36	(–0.2733, –0.0061)
2	(40.68, –0.6020)	43.133	(0.0289, 2.989)	$\begin{pmatrix} 4.81 & 158.97 \\ 158.97 & 10{,}918 \end{pmatrix}$	2.9895	(0.0058, –0.0004)
3	(40.69, –0.6024)	43.133	(0.0000, 0.0027)	$\begin{pmatrix} 4.81 & 158.73 \\ 158.73 & 10{,}899 \end{pmatrix}$	0.00272	Stop

Tabla 3.2: *Iteraciones* del método de Newton.

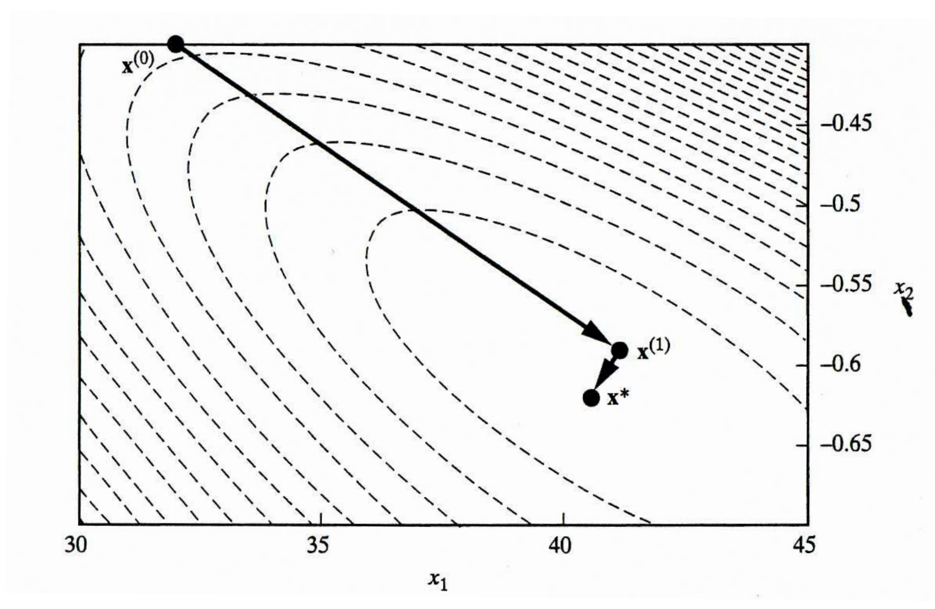


Figura 3.8: Representación de las *iteraciones* del método de Newton.

Solución: Mediante el método de Newton, hemos obtenido el siguiente *punto óptimo*

$$x_1^* = 40{,}69 \qquad x_2^* = -0{,}6024, \qquad f(x^*) = 43,133$$

y por lo tanto la *función de precios óptima* $r^*(p_i)$ es:

$$r^*(p_i) \quad = \quad r(x_1^*, x_2^* \mid p_i) \quad = \quad 40{,}69 p_i^{-0{,}6024}.$$

General (Ventajas (+) y desventajas ($-$))

Método del gradiente:

+ *Convergencia* a un óptimo (local) garantizada.
+ En las *primeras iteraciones* el progreso hacia el óptimo es buena.
– Después de las primeras iteraciones, la convergencia pude ser lenta: puede avanzar hacia el óptimo siguiendo una trayectoria en forma de *zig-zag* (muchas iteraciones).
+ Cada iteración es computacionalmente *ligera* (en tiempo de CPU y en requerimientos de memoria RAM): en cada iteración calculamos el gradiente (dimensión n) y la longitud de paso.

Método de Newton:

– *Convergencia* a un óptimo (local) garantizada sólo si empezamos 'suficientemente' cerca del óptimo.
+ La convergencia es *rápida* (pocas iteraciones).
– Cada iteración es computacionalmente *cara* (en tiempo de CPU y en requerimientos de memoria RAM): en cada iteración calculamos la matriz hessiana (dimensión n^2) y resolvemos un sistema lineal.

Estrategia mixta: Usar un método u otro dependerá del problema. Una buena alternativa es *combinar* los dos métodos:

- Usar el método del gradiente en una *primera fase* para 'acercarnos' al óptimo.
- Cuando el progreso del método del gradiente sea lento, empezar una *segunda fase* con el método de Newton que nos lleve en pocas iteraciones al óptimo.

Capítulo 4

Teoría de la decisión: Decisiones multiobjetivo

4.1. Introducción al análisis de decisiones

- El 'Análisis de decisiones' y la 'Optimización' están íntimamente *relacionadas.*
- En general se desea tomar la 'mejor' decisión según uno o varios criterios, es decir, se desea tomar la *decisión óptima* según uno o varios criterios.
- Por lo tanto, los *métodos de optimización* de los capítulos anteriores (programación lineal (PL) y programación no lineal (PNL)) pueden ser de gran ayuda a la hora de tomar decisiones.
- Tanto la PL como la PNL son métodos de optimización *monoobjetivo,* pues consideran sólo un criterio u objetivo (coste, beneficio, etc.).
- Sin embargo, tal como veremos en el próximo ejemplo, en muchos casos la toma de decisiones implica considerar varios criterios u objetivos (decisiones *multiobjetivo).*

Nota. Este tema también se puede consultar y/o ampliar en:

- El Capítulo 8 del siguiente libro: Rardin, R. L., 'Optimization in operations research', Editorial Prentice Hall, 1998.
- Los Capítulos 2 y 4 del siguiente libro: C. Beltran-Royo. 'Decisiones Óptimas: Cómo Optimizar las Cuestiones de la Vida Cotidiana', CreateSpace (Amazon), EEUU, 2013.

Ejemplo 28

Datos:

- *María* es representante comercial de una prestigiosa marca de relojes suizos.
- Su trabajo consiste en *visitar* a los diferentes clientes que la firma posee en las principales ciudades de España y Portugal.
- Por ese motivo María debe recorrer alrededor de *30000 Km* de autovía o autopista cada año.
- María, por un lado estima que cada hora de conducción tiene un valor de *40 euros.*

- Por otro lado, sabe que el *consumo* de su coche es de 8,0, 11,5 ó 15,7 litros cada 100 Km viajando a una velocidad de 100, 120 ó 140 Km/h, respectivamente.
- El precio del gasoil está en *1,50 euros/litro.*

Objetivo: El objetivo de María es decidir qué *velocidad* le conviene al hacer sus desplazamientos por autovía o autopista, teniendo en cuenta el consumo de *gasoil* y el *tiempo* de desplazamiento.

Operaciones 28:

Velocidad	**F1 Coste del gasoil**	**F2 Coste del tiempo**	**F3 Coste global**
100	3600	12000	15600
120	5175	10000	*15175*
140	7065	8571	15636

Tabla 4.1: Coste anual global (euros) en función de la velocidad (Km/h).

Solución: Con los datos de María, la velocidad media óptima para viajar por autovía, es de *120 Km/h,* ocasionándole un gasto de *15.175 euros/año.*

General (Decisiones multiobjetivo)

- Como ya hemos mencionado, en muchos casos la toma de decisiones implica considerar varios objetivos simultáneamente (decisiones *multiobjetivo).*

- En el ejemplo anterior ha sido posible transformar un problema con *dos criterios* (consumo de gasoil y consumo de tiempo) en un problema con un único criterio (consumo de euros).

- La idea ha sido expresar en la misma *unidad* (euros) el coste del gasoil y el coste del tiempo de trabajo de María.

- Por lo tanto, en algunos casos es posible transformar los objetivos involucrados en un único objetivo y utilizar la optimización *monoobjetivo.*

- En otros, tal como veremos en el siguiente ejemplo, es *difícil* o imposible transformar todos los criterios involucrados en un único criterio.

Ejemplo 29 (Banco Trébol)

Datos:

- El Banco Trébol tiene los siguientes *recursos* (en millones de euros):
  - ▷ capital *propio:* 20,
  - ▷ capital en *cuentas* corrientes: 150,
  - ▷ capital en *depósitos* a plazo fijo: 80.
- Está planificando su estrategia de *inversión* y en la Tabla 4.2 ha recopilado los datos de los posibles tipos de inversión:

Tipo de Inversión j	**Tasa de Rendimiento (%)**	**Tasa de Liquidez (%)**	**Tasa de Reservas (%)**	**¿Riesgo?**
1: Dinero en caja	0,0	100,0	0,0	No
2: Inversiones a corto plazo	4,0	99,5	0,5	No
3: Bonos del Estado (1 a 5 años)	4,5	96,0	4,0	No
4: Bonos del Estado (6 a 10 años)	5,5	90,0	5,0	No
5: Bonos del Estado (más de 10 años)	7,0	85,0	7,5	No
6: Préstamos a plazo	10,5	0,0	10,0	Sí
7: Préstamos hipotecarios	8,5	0,0	10,0	Sí
8: Préstamos comerciales	9,2	0,0	10,0	Sí

Tabla 4.2: Posibles *inversiones* para el Banco Trébol.

  - ▷ La 'Tasa de *Rendimiento'* corresponde al rendimiento anual.
  - ▷ La 'Tasa de *Liquidez'* refleja el grado de disponibilidad del dinero. Por ejemplo, una tasa del 90 % significa que sólo el 90 % del capital invertido en ese producto está disponible de forma inmediata.
  - ▷ La 'Tasa de *Reserva'* sirve para constituir un fondo de reserva. Por ejemplo, una tasa del 10 % significa que el Banco debe dejar en el fondo de Reservas un capital igual al 10 % del capital invertido en ese producto.
  - ▷ En la columna *'¿Riesgo?'* se indica si la inversión tiene cierto nivel de riesgo.

Objetivo: Define la *función objetivo* para cada uno de los siguientes tres objetivos:

1) Maximizar el *beneficio* $z_1(x)$.

2) Minimizar el coeficiente de *riesgo* (CRi):

$$z_2(x) = \text{CRi} = \frac{\text{Inversiones con riesgo}}{\text{Capital propio}}$$

3) Minimizar el coeficiente de *reservas* (CRe):

$$z_3(x) = \text{CRe} = \frac{\text{Fondo de reserva}}{\text{Capital propio}}$$

Operaciones 29:

Solución:

- Ver los anteriores apartados.
- Notar que maximizar el beneficio entra en conflicto con minimizar el coeficiente de riesgo.

General (Decisiones multiobjetivo II)

- En muchos problemas de decisión, como en el ejemplo anterior, sería muy complicado si no imposible *transformar* todos los objetivos involucrados en un único objetivo.
- En este caso podemos usar las técnicas de optimización *multiobjetivo* que presentaremos en este capítulo.

Ejemplo 30 (Banco Trébol - continuación)

Datos:

- El Banco Trébol, además de los tres objetivos del ejemplo anterior, tiene que cumplir las siguientes *restricciones:*
  1) El banco desea *invertir* todo el capital disponible.
  2) El *dinero en caja* deben ser al menos el 14 % del dinero actualmente depositado en las cuentas corrientes más el 4 % del dinero actualmente depositado a plazo fijo.
  3) La parte de la inversión considerada *líquida* debería ser al menos el 47 % del dinero actualmente depositado en las cuentas corrientes más el 36 % del dinero actualmente depositado a plazo fijo.
  4) El Banco desea invertir al menos un 5 % en cada uno de los 8 productos de cara a *diversificar* su inversión.
  5) Al menos el 30 % debería destinarse a *préstamos* comerciales.

Objetivo: Formula un problema de programación lineal multiobjetivo para *optimizar* las inversiones del Banco Trébol.

Operaciones 30:

Solución:

- Notar que la restricción de 'Liquidez' está escrita en dos líneas debido a su excesiva longitud.
- Ver el apartado anterior.

General (Programación lineal multiobjetivo)

- Un problema de programación lineal *multiobjetivo* se caracteriza por:
  - ▷ Tener varias *funciones objetivo* que han de optimizarse (maximizar o minimizar según convenga).
  - ▷ La *región factible* es idéntica a la región factible de un problema PL monoobjetivo, es decir, definido a partir de restricciones lineales de igualdad o de desigualdad junto con cotas para las variables.

4.2. Puntos eficientes y frontera eficiente

General

- *Ambigüedad del concepto 'óptimo multiobjetivo':*
  - ▷ En la mayoría de PLMs no existe ningún punto que sea óptimo respecto a todos los criterios.
  - ▷ Por ejemplo, un punto x^* puede ser óptimo respecto a un criterio z_1 y subóptimo respecto a otro criterio z_2.
  - ▷ Por ese motivo en los PLMs, buscaremos puntos eficientes en vez de puntos que sean óptimos en todos los criterios.
- *Punto eficiente:*
  - ▷ Dado un PLM, un punto factible es *eficiente* si para mejorarlo en algún criterio tenemos que empeorarlo en algún otro criterio.
  - ▷ Un punto eficiente se denomina también punto *óptimo de Pareto.*
  - ▷ Análogamente, un punto no eficiente se denomina *ineficiente.*
- *Frontera eficiente:*
  - ▷ Es el conjunto formado por todos los puntos eficientes.
  - ▷ Se denomina también *frontera de Pareto.*

Ejemplo 31

Datos:

- Consideramos el *PL multiobjetivo* siguiente:

$$\begin{array}{rl} \max & z_1(x) = +3x_1 + 1x_2 \\ \max & z_2(x) = -1x_1 + 2x_2 \\ \text{s. a} & x_1 + x_2 \leq 4 \\ & 0 \leq x_1 \leq 3 \\ & 0 \leq x_2 \leq 3 \end{array}$$

Objetivo:

1) Analiza si el *punto* $x = (3, 0)$ es eficiente.
2) Analiza si el *punto* $x = (2, 2)$ es eficiente.

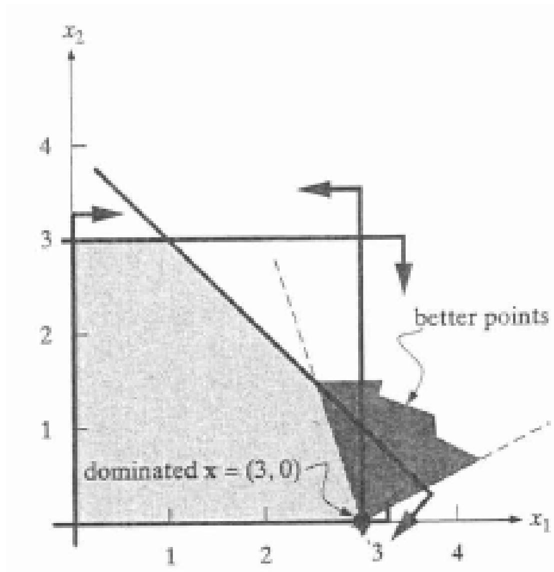


Figura 4.1: El punto $x = (3,\ 0)$ es *ineficiente.*

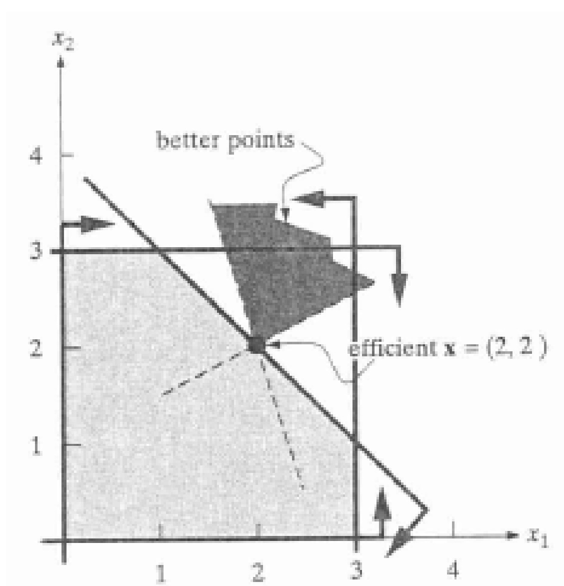


Figura 4.2: El punto $x = (2,\ 2)$ es *eficiente.*

3) Determina la *frontera eficiente* para este PL multiobjetivo.

Operaciones 31:

Solución: El punto $x = (3, 0)$ es *ineficiente,* el punto $x = (2, 2)$ es *eficiente* y la frontera eficiente corresponde a la línea poligonal descrita por los puntos:

$$(3, 1) \to (1, 3) \to (0, 3).$$

(ver Figura 4.2)

General (Análisis gráfico de la optimalidad de Pareto)

- Dado un programa lineal multiobjetivo (PLM) con dos variables, podemos determinar si un punto factible x es *eficiente* mediante los siguientes pasos:

  1) Representamos el *conjunto factible* y el punto x.
  2) Representamos las *curvas de nivel* de las funciones objetivo y que pasan por el punto x.
  3) Determinamos los *semiplanos* de mejora de cada función objetivo, a partir de las anteriores curvas de nivel.
  4) Consideramos el *cono* K definido por la intersección los anteriores semiplanos de mejora (su vértice es el punto x).
  5) Si la intersección de K y la región factible es $\{x\}$, entonces x es un punto *eficiente* del PLM.
  6) En caso contrario x es *ineficiente.*

4.3. Optimización multiobjetivo por suma ponderada de objetivos

Ejemplo 32 (Suma ponderada de objetivos)

Datos: Consideramos el siguiente *PLM:*

$$
\begin{array}{lll}
\text{min} & z_1(x) = +3x_1 + 1x_2 & \text{(horas)} \\
\text{max} & z_2(x) = -1x_1 + 2x_2 & \text{(euros)} \\
\text{max} & z_3(x) = +2x_1 & \text{(kg)} \\
\text{s. a.} & x_1 + x_2 \leq 4 & \\
 & 0 \leq x_1 \leq 3 & \\
 & 0 \leq x_2 \leq 3. &
\end{array}
$$

Objetivo: Transforma el anterior PLM en un PL mediante la *suma ponderada* de objetivos, usando los pesos $\gamma_1 = 2$ PC / hora, $\gamma_2 = 3$ PC / euro y $\gamma_3 = 10$ PC / kg, donde PC significa *'Puntos de Coste'*.

Operaciones 32:

Solución:

- El correspondiente PL es

$$\begin{array}{lll} \min & z(x) = -11x_1 - 4x_2 & \text{(PC)} \\ \text{s. a.} & x_1 + x_2 \leq 4 & \\ & 0 \leq x_1 \leq 3 & \\ & 0 \leq x_2 \leq 3 & \end{array}$$

- Notar que la región factible no cambia.

General (Optimización multiobjetivo por suma ponderada de objetivos)

- Consideramos el siguiente *PL multiobjetivo* (PLM)

$$\begin{array}{lll} \min & z_1(x) & (\text{Unidad}_1) \\ & \vdots & \\ \min & z_m(x) & (\text{Unidad}_m) \\ \text{s. a} & Ax \leq b & \\ & x \geq 0, & \end{array}$$

  donde $z_i(x) = c_{i1}x_1 + \ldots + c_{in}x_n$ para cada $i \in \mathcal{I} = \{1, \ldots, m\}$.

- Una forma de *resolver* este problema es usando la técnica de la *suma ponderada* de objetivos:

$$\begin{array}{ll} \min & z(x) = \sum_{i \in \mathcal{I}} \gamma_i z_i(x) \\ \text{s. a} & Ax \leq b \\ & x \geq 0, \end{array}$$

  donde cada peso γ_i es un escalar estrictamente *positivo* para todo $i \in \mathcal{I}$.

- Los pesos γ_i sirven para dar mayor o menor *importancia* a su correspondiente objetivo $z_i(x)$.
- Las unidades de cada γ_i son: 'Puntos de Coste por Unidad$_i$', es decir, PC / Unidad$_i$.
- Por lo tanto, las unidades de la función objetivo son 'Puntos de Coste' (PC).
- El decisor debe seleccionar estos pesos dependiendo de su *criterio* y de las *circunstancias.*
- En este algoritmo debemos tener o bien solo min, o bien solo max en todos los criterios. Si hace falta podemos usar la equivalencia

$$\text{máx}\, z(x) \equiv -\text{mín}[-z(x)].$$

General (Suma ponderada de objetivos y optimalidad de Pareto)

Propiedad: Toda solución óptima obtenidas mediante la suma ponderada de objetivos es solución óptima de *Pareto.*

Ejemplo 33 (Eliminación de residuos peligrosos)

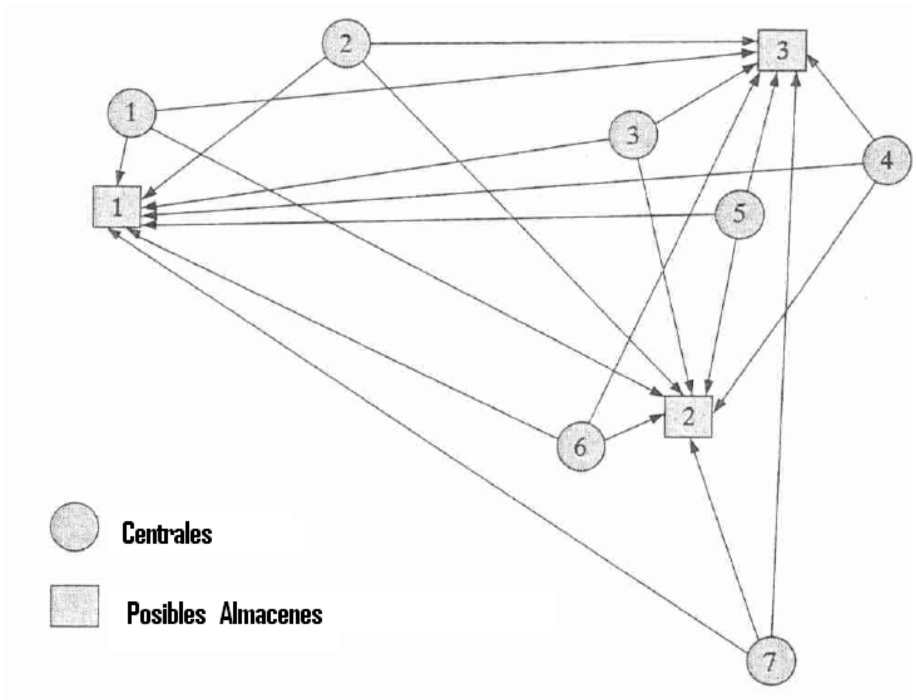


Figura 4.3: *Centrales* y posibles *almacenes.*

TABLE 8.2 Hazardous Waste Example

Source i		Site $j = 1$		Site $j = 2$		Site $j = 3$		Supply
		$k = 1$	$k = 2$	$k = 1$	$k = 2$	$k = 1$	$k = 2$	
1	Distance	200	280	850	1090	900	1100	1.2
	Population	50	15	300	80	400	190	
2	Distance	400	530	730	860	450	600	0.5
	Population	105	60	380	210	350	160	
3	Distance	600	735	550	600	210	240	0.3
	Population	300	130	520	220	270	140	
4	Distance	900	1060	450	570	180	360	0.7
	Population	620	410	700	430	800	280	
5	Distance	600	640	390	440	360	510	0.6
	Population	205	180	440	370	680	330	
6	Distance	900	1240	100	120	640	800	0.1
	Population	390	125	80	30	800	410	
7	Distance	1230	1410	400	460	1305	1500	0.2
	Population	465	310	180	105	1245	790	

Figura 4.4: Datos sobre *distancia* (millas) y *población* (miles de personas).

Datos:

- Una compañía propietaria de siete *centrales* nucleares está decidiendo en coordinación con el gobierno la localización de dos *almacenes* de residuos nucleares.
- En la Figura 4.3 tenemos representadas la *localización* de las siete centrales y las tres *posibles localizaciones* para los dos almacenes nucleares.
- Cuando los almacenes estén construidos, los *residuos* serán transportados desde las centrales a los almacenes mediante camiones utilizando la red publica de *carreteras.*
- Se persiguen dos *objetivos:*
  - ▷ Minimizar la *distancia* de las centrales a los almacenes (coste transporte: combustible y tiempo empleado).
  - ▷ Minimizar el tránsito por áreas con mayor *población* (gente expuesta al riesgo de mercancías peligrosas y riesgo de accidentes de tráfico en zonas con mayor densidad de tráfico).
- En la tabla de la Figura 4.4 tenemos los siguientes *parámetros:*
  - ▷ $i =$ índice para las *centrales,* $i \in \mathcal{I} = \{1, \ldots, 7\}$.
  - ▷ $j =$ índice para las *localizaciones,* $j \in \mathcal{J} = \{1, 2, 3\}$.
  - ▷ $k =$ índice para las *rutas,* $k \in \mathcal{K} = \{1, 2\}$.
  - ▷ $s_i =$ ('Supply') cantidad (toneladas) esperada de *residuos* en la central $i \in \mathcal{I}$.
  - ▷ $d_{ijk} =$ *Distancia* (millas) de la central i a la localización j a través de la ruta k, para todos los $i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}$.
  - ▷ $p_{ijk} =$ *Población* (en miles de personas) que encontramos de la central i a la localización j a lo largo de la ruta k, para todos los $i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}$.
  - ▷ *Notar* que para cada caso la ruta $k = 1$, comparada con la ruta $k = 2$, es más corta pero involucra más población.

Objetivo: *Formular* el PL multiobjetivo para ayudar a decidir la *localización* de los dos almacenes de residuos nucleares.

Operaciones 33:

Solución: Ver el apartado anterior.

Ejemplo 34 (Eliminación de residuos peligrosos - continuación)

Datos: Continuamos con el ejemplo anterior.

Objetivo:

1) *Formular* el anterior problema de optimización multiobjetivo como un problema monoobjetivo mediante la *suma ponderada* de sus objetivos.
2) *Resolver* el anterior problema comparando varias ponderaciones de las funciones objetivo $z_1(x)$ y $z_2(x)$.

Operaciones 34:

Caso	γ_1	γ_2	**Millas×Tm** $k=1$	**Millas×Tm** $k=2$	**Total Millas×Tm** z_1^*	**Total 10^3 Personas ×Tm** z_2^*	**Loc. Óptimas**
1	10	1	1155	0	1155	1334	1 3
2	10	5	754	591	1345	782	1 3
3	10	10	1046	404	1450	607	1 2
4	5	10	440	1114	1554	533	1 3
5	1	10	0	1715	1715	468	1 3

Tabla 4.3: *Localización óptima* de los dos almacenes en función de los pesos γ_1 y γ_2.

Solución:

- Ver en el apartado ‘Operaciones’.
- Ver Tabla 4.3. Con estos datos la compañía, en coordinación con el gobierno, puede tomar una decisión.
- En la Tabla 4.3 tenemos los resultados obtenidos para diferentes pesos (estos resultados se han obtenido mediante un solver para PL).
- La etiqueta *‘Millas×Tm’* corresponde a multiplicar ‘Millas’ por ‘Toneladas’ en el término $d_{ijk}x_{ijk}$ que aparece en $z_1(x)$.
- La etiqueta ‘10^3 *Personas×Tm’* corresponde a multiplicar ‘10^3 Personas’ por ‘Toneladas’ en el término $p_{ijk}x_{ijk}$ que aparece en $z_2(x)$.
- En el caso 1, donde el criterio *‘distancia’* tiene un peso 10 veces superior al criterio *‘población’,* todos los residuos son enviados por las rutas con índice $k = 1$ *(rutas más cortas,* pero con más población).
- En el caso 5, ocurre lo contrario.
- En todos los casos la *localización óptima* de los dos almacenes es en las localizaciones candidatas 1 y 3, excepto para el caso 3.

4.4. Optimización multiobjetivo por metas

Ejemplo 35 (Optimización por metas)

Datos: Consideramos el siguiente problema multiobjetivo por *metas:*

$$\begin{array}{rl}
\text{meta} & z_1(x) = +3x_1 + 1x_2 \leq 10 \\
\text{meta} & z_2(x) = -1x_1 + 2x_2 \geq 5 \\
\text{meta} & z_3(x) = +2x_1 = 8 \\
\text{s. a} & x_1 + x_2 \leq 4 \\
 & 0 \leq x_1 \leq 3 \\
 & 0 \leq x_2 \leq 3.
\end{array}$$

Objetivo: Formular el anterior problema mediante un PL introduciendo *variables de deficiencia* ponderadas según los siguientes pesos $\gamma_1 = 2, \gamma_2 = 3$ y $\gamma_3^+ = 10, \gamma_3^- = 5$.

Operaciones 35:

Solución: El correspondiente PL es

$$\begin{array}{rl} \min & 2d_1 + 3d_2 + 10d_3^+ + 5d_3^- \\ \text{s. a} & 3x_1 + 1x_2 - d_1 \leq 10 \\ & 1x_1 - 2x_2 - d_2 \leq -5 \\ & +2x_1 + d_3^+ - d_3^- = 8 \\ & x_1 + x_2 \leq 4 \\ & 0 \leq x_1 \leq 3 \\ & 0 \leq x_2 \leq 3 \\ & d \geq 0. \end{array}$$

General (Optimización multiobjetivo por metas)

- Las metas pueden considerarse *restricciones suaves* (no es imprescindible que se cumplan), es decir, intentamos calcular una solución que las cumpla en el mayor grado posible.
- Consideramos el siguiente problema multiobjetivo por *metas:*

$$\begin{array}{rl} \text{meta} & z_1(x) \leq m_1 \\ & \vdots \\ \text{meta} & z_m(x) \leq m_m \\ \text{s. a} & Ax \leq b \\ & x \geq 0, \end{array}$$

donde $z_i(x) = c_{i1}x_1 + \ldots + c_{in}x_n$ para cada $i \in \mathcal{I} = \{1, \ldots, m\}$.

- Podemos resolver este problema multiobjetivo transformándolo en un problema PL:
  - ▷ Introducimos *variables de deficiencia* d_i mediante sendas restricciones.
  - ▷ Minimizar la *suma ponderada* de la variables de deficiencia.

$$\begin{aligned} \min \quad & z(d) = \sum_{i \in \mathcal{I}} \gamma_i d_i \\ \text{s. a} \quad & z_1(x) - d_1 \leq m_1 \\ & \vdots \\ & z_m(x) - d_m \leq m_m \\ & Ax \leq b \\ & x \geq 0, \ d \geq 0 \end{aligned} \tag{4.6}$$

donde cada peso γ_i es un escalar estrictamente *positivo* para todo $i \in \mathcal{I}$.

- Los pesos γ_i sirven para dar mayor o menor *importancia* a su correspondiente meta.
- El decisor debe seleccionar los pesos y las metas dependiendo de su *criterio* y de las *circunstancias.*
- Si en el problema tenemos una meta con desigualdad de $\geq$, tenemos que transformarla en una desigualdad de $\leq$ multiplicando los coeficientes de la desigualdad por -1.
- Si una meta es de igualdad, por ejemplo $z_0(x) = m_0$, tendremos que asignarle dos variables de deficiencia positivas: d_0^+ y d_0^- :
$$z_0(x) + d_0^+ - d_0^- = m_0.$$
- Notar que las unidades de los pesos γ_i son 'Puntos de Coste por Unidad$_i$', es decir, PC / Unidad$_i$.
- Por tanto, las unidades de la función objetivo $z(d)$ son PC.

General (Optimización por metas y optimalidad de Pareto)

- La optimización multiobjetivo por metas es un método *popular* entre los decisores, pues el trabajar con metas resulta intuitivo.
- Pero, en la optimización multiobjetivo por metas *no tenemos garantías* de que la solución óptima obtenida sea una solución óptima de Pareto.
- Tenemos dos casos:
  - ▷ Si $d_j^* \neq 0$ para todo j, entonces x^* es un punto de Pareto.
  - ▷ Si $d_j^* = 0$ para algún j, entonces *no tenemos garantías* de que x^* se un punto de Pareto. Explicación: una vez que una meta j está satisfecha ($d_j^* = 0$), el algoritmo no busca puntos mejores con relación a esa meta.

Ejemplo 36 (Banco Trébol - continuación)

Datos:

- Consideramos de nuevo el ejemplo del Banco Trébol.

- Lo que pretendemos ahora es resolver este problema multiobjetivo de forma que alcancemos las siguientes *'metas':*
  - ▷ Un *beneficio* de al menos 18,5 millones de euros.
  - ▷ Un coeficiente de *riesgo* menor o iqual que 7,0.
  - ▷ Un coeficiente de *reservas* menor o iqual que 0,8.

Objetivo: Introducir estas tres *metas* en la formulación del problema.

Operaciones 36:

Solución: Ver el apartado 'Operaciones'.

Ejemplo 37 (Banco Trébol - continuación)

Datos:

- Vamos a concluir el ejemplo del Banco Trébol.
- Recordamos que lo que pretendemos es resolver este problema multiobjetivo de forma que alcancemos las siguientes *metas:*
  - ▷ Un *beneficio* de al menos 18,5 millones de euros.
  - ▷ Un coeficiente de *riesgo* menor que 7,0 (sin unidades).
  - ▷ Un coeficiente de *reservas* menor que 0,8 (sin unidades).

Objetivo:

1) Formular el problema PL para resolver este problema multiobjetivo, dando *igual importancia* a todas las metas.
2) Calcular una *solución óptima* para el problema PL.
3) Repetir los dos apartados anteriores, pero dando 10 veces *más importancia* a la meta 'reservas' que a las otras dos metas.

Operaciones 37:

Solución:

1) Ver apartado ‘Operaciones’.
2) Ver columna ‘Caso 1’ de la Tabla 4.4.
3) Ver columna ‘Caso 2’ de la Tabla 4.4.

	Caso 1	Caso 2
γ_1 : Peso para el beneficio	1	1
γ_2 : Peso para el coeficiente riesgo	1	1
γ_3 : Peso para el coeficiente de reservas	1	10
$z_1(x^*)$: Beneficio ($\geq 18,50$)	18,50	17,53
d_1^*	0,00	0,97
$z_2(x^*)$: Riesgo ($\leq 7,00$)	7,00	7,00
d_2^*	0,00	0,00
$z_3(x^*)$: Reservas ($\leq 0,80$)	0,93	0,82
d_3^*	0,13	0,02
x_1^*: Dinero en caja	24,20	24,20
x_2^*: Inversiones a corto plazo	16,03	48,30
x_3^*: Bonos del Estado (1 a 5 años)	12,50	12,50
x_4^*: Bonos del Estado (6 a 10 años)	12,50	12,50
x_5^*: Bonos del Estado (más de 10 años)	44,77	12,50
x_6^*: Préstamos a plazo	52,50	52,50
x_7^*: Préstamos hipotecarios	12,50	12,50
x_8^*: Préstamos comerciales	75,00	75,00

Tabla 4.4: *Resultados* para dos conjuntos de pesos: $\gamma = (1, 1, 1)$ y $\gamma = (1, 1, 10)$.

Capítulo 5

Gestión de proyectos

5.1. Método del camino crítico

Nota: Este método, en inglés, se denomina *'Critical Path Method'* (CPM).

Ejemplo 38 (Construcción de un almacén)

Datos:

- Consideramos el *proyecto* para la construcción de un almacén.
- Dicho proyecto consta de *9 actividades* que siguen cierto orden de precedencia (Tabla 5.1).

k	**Actividad**	**Duración** a_k **(días)**	**Actividades Precedentes**
1	Cimientos	15	–
2	Alcantarillas	5	–
3	Base de hormigón	4	1, 2
4	Estructura y paredes	3	3
5	Tejado	7	4
6	Instalación eléctrica primaria	10	4
7	Calefacción y aire acondicionado	13	2, 4
8	Tabiques	18	4, 6, 7
9	Acabados interiores	20	5, 8

Tabla 5.1: *Duración* de las actividades para la construcción del almacén.

- En la Tabla 5.1 tenemos a_k, la *duración estimada* de cada actividad (días).

Objetivo: Representa la *red* asociada a este proyecto.

Operaciones 38:

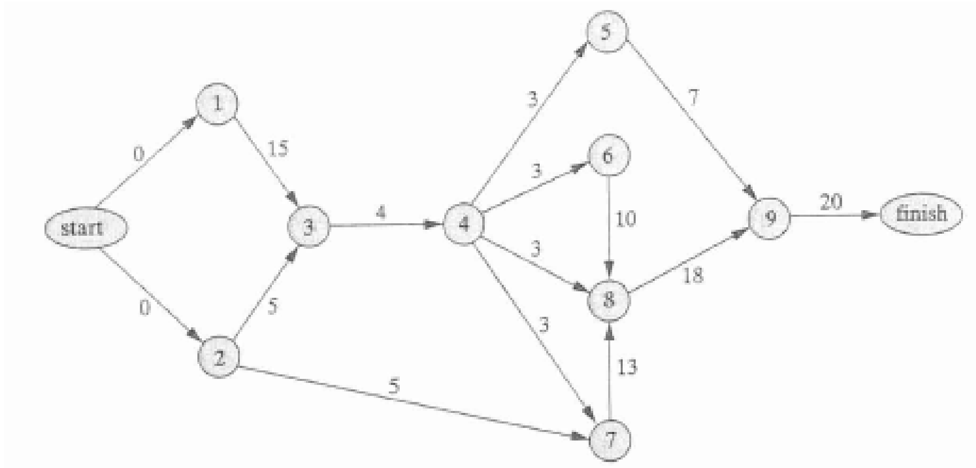


Figura 5.1: *Red* del proyecto 'Construcción de un almacén'.

Solución: En la Figura 5.1 tenemos representada la *red* del proyecto.

General (Red de un proyecto)

- De cara a *gestionar* un proyecto, nos será muy útil representar la denominada *red* del proyecto donde:
  - ▷ Cada *nodo* representa una actividad.
  - ▷ De cada *actividad* k sale uno o varios arcos de longitud a_k dirigidos a cada actividad sucesora.
  - ▷ Lo nodos que no tienen predecesores se unen al denominado *'nodo inicial'* (nodo 0) mediante un arco de longitud (duración) cero.
  - ▷ Los nodos que no tienen sucesores se unen al denominado *'nodo final'*.

General (Tiempos de un proyecto I)

- A la hora de *gestionar* un proyecto nos puede interesar:
  - ▷ Estimar la *duración* total del proyecto.
  - ▷ Planificar la *compra* de materiales.
  - ▷ Planificar la *contratación* de los correspondientes especialistas.

- ▷ Planificar las correspondientes *licencias* y estudios.
- ▷ Etc.

- Por estos motivos nos interesa *conocer:*
  - ▷ El *tiempo total* estimado que durará la ejecución del proyecto.
  - ▷ Para cada actividad del proyecto, el tiempo mínimo que debe pasar antes de que podamos iniciarla *(tiempo de inicio temprano).*
  - ▷ Notar que si el 'tiempo de inicio temprano' de una actividad es t días significa que podemos empezarla al inicio del día $t + 1$.
  - ▷ En este capítulo, cada día corresponde a un intervalo [I, F], donde I es el inicio y F es el final.
  - ▷ Empezamos a ejecutar cada tarea al inicio I del correspondiente día.

- Dada la red de un proyecto, se denomina *'camino crítico'* para la actividad p al camino más largo del nodo 0 al nodo p.
  - ▷ El *tiempo de inicio temprano* para la actividad p corresponde a la longitud del su camino crítico.
  - ▷ Por lo tanto, la *duración total* de un proyecto corresponde a la longitud del camino crítico hasta el 'nodo final'.

Nota. Este tema también se puede consultar y/o ampliar en el Capítulo 9 del siguiente libro: Rardin, R. L., 'Optimization in Operations Research', Editorial Prentice Hall, 1998.

Ejemplo 39 (Construcción de un almacén - continuación)

Datos:

- En la Tabla 5.1 tenemos a_k la *duración* estimada de cada actividad (días) junto con las respectivas actividades precedentes.
- El *tiempo de inicio temprano* para una actividad p, que denotamos por $v[p]$, podemos calcularlo con la siguiente fórmula recursiva:

$$v[p] = \max\{v[i] + a_i \mid i \text{ es un predecesor de } p\}$$

- Denotamos por $d[p]$ el nodo i para el que se alcanza el anterior *máximo.*
- d[p] es el *predecesor* de p en su camino crítico.

Objetivo: Calcula el *tiempo* de inicio temprano y el *camino* crítico para cada una de las actividades requeridas para construir el almacén.

Operaciones 39:

k	**Actividad**	**Inicio Temprano**	**Camino Crítico**
1	Cimientos	0	0 - 1
2	Alcantarillas	0	0 - 2
3	Base de hormigón	15	0 - 1 - 3
4	Estructura y paredes	19	0 - 1 - 3 - 4
5	Tejado	22	0 - 1 - 3 - 4 - 5
6	Instalación eléctrica primaria	22	0 - 1 - 3 - 4 - 6
7	Calefacción y aire acondicionado	22	0 - 1 - 3 - 4 - 7
8	Tabiques	35	0 - 1 - 3 - 4 - 7 - 8
9	Acabados interiores	53	0 - 1 - 3 - 4 - 7 - 8 - 9

Tabla 5.2: *Inicio* temprano (días) de las actividades para la construcción del almacén.

Solución: El *tiempo* de inicio temprano y el *camino* crítico para cada una de las actividades puede encontrarse en la Tabla 5.2.

General (Método del camino crítico - tiempo de inicio temprano)

- Notar que el método del camino crítico equivale a calcular el camino más largo desde el nodo 0 a cada nodo p mediante el método de optimización denominado *'Programación dinámica'*.
- *Input:*
  1) a_k : *duración* estimada de la actividad $k \in \mathcal{K} = \{1, \ldots, K\}$.
  2) Conjunto de actividades *predecesoras* de cada actividad $k \in \mathcal{K}$.
- *Output:*
  1) $v[k]$: *Tiempo de inicio temprano* de la actividad $k \in \mathcal{K}$.
  2) $d[k]$: *Predecesor* de la actividad k en el camino crítico de $k \in \mathcal{K}$.
  3) *Caminos críticos:* Se reconstruyen a partir de los $d[k]$.
  4) *Duración* estimada de todo el proyecto: $v[K+1]$, donde $K+1$ es el 'nodo final'.
- *Paso 0:* Inicialización.
  1) *Numerar* los nodos de forma que los arcos (i, j) de la red del proyecto cumpla $i < j$.
  2) *Inicializar* $v[0] = 0$.
- *Paso 1:* Criterio de parada.
  1) Si ya hemos calculado el tiempo de inicio temprano para el 'nodo final' *paramos.*
  2) En caso contrario, sea p el *nodo* no procesado con la etiqueta más pequeña.
- *Paso 2:* Procesado.
  1) Calcular el *tiempo de inicio temprano* para p

$$v[p] \quad = \quad \max\{v[i] + a_i \mid i \text{ es un predecesor de } p\} \tag{5.1}$$

  2) Guardar en $d[p]$ el nodo i para el que se alcanza el anterior *máximo.*
  3) d[p] es el *predecesor* de p en su camino crítico.

4) Volver al *Paso 1*.

General (Tiempos de un proyecto II)

- A la hora de *gestionar* un proyecto también nos interesa conocer:
  - ▷ La fecha límite para *finalizar* el proyecto.
  - ▷ Para cada actividad del proyecto, el tiempo máximo que puede pasar antes de iniciarla sin retrasar el proyecto *(tiempo de inicio tardío).*
  - ▷ *Notar* que, dada una actividad con tiempos de inicio temprano y tardío de t_0 y t_1 días, respectivamente, el momento en que podemos iniciar dicha actividad sin retrasar el proyecto puede ser en cualquier día del *intervalo* $[t_0 + 1,\ t_1 + 1]$.
- *Cálculo del tiempo de inicio tardío:*
  - ▷ Podemos calcular el tiempo de inicio tardío para la actividad p, a partir de la longitud del *camino más largo* desde el nodo p hasta el nodo final.
  - ▷ Para calcular el *tiempo de inicio tardío* para la actividad p, que denotamos por $\widehat{v}[p]$, podemos usar la siguiente fórmula recursiva:

$$\widehat{v}[p] = \min\{\widehat{v}[i] \mid i \text{ es un sucesor de } p\} - a_p$$

    Notar que a_p está fuera de la minimización y no depende de i (comparar con ecuación (5.1)).
  - ▷ Denotamos por $\widehat{d}[p]$ el nodo i donde se alcanza el anterior *mínimo.*
- *Notación:*
  - ▷ $\widehat{v}[k]$: *Tiempo de inicio tardío* de la actividad $k \in \mathcal{K}$.
  - ▷ $\widehat{d}[k]$: *Sucesor* de k en el camino más largo de k al nodo final, $k \in \mathcal{K}$.
- *Holgura del tiempo de inicio:*
  - ▷ Denotamos por $h[k]$ esta holgura.
  - ▷ Es la *diferencia* entre el tiempo de inicio tardío y el tiempo de inicio temprano: $h[k] = \widehat{v}[k] - v[k]$.
  - ▷ Es el margen que tenemos para *retrasar* una actividad para terminar el proyecto dentro de plazo.
- *Actividades críticas:*
  - ▷ Son las actividades con una holgura de tiempo de inicio igual a *cero.*
  - ▷ El retraso en una actividad crítica conlleva no poder terminar el proyecto dentro de *plazo.*

Ejemplo 40 (Construcción de un almacén - continuación)

Datos:

- Continuamos con el *proyecto* para construir un almacén.
- Suponemos que el *plazo* límite para ejecutar el proyecto es de 80 días.

Objetivo: Para cada actividad, calcula el *tiempo* de inicio tardío, la *holgura* del tiempo de inicio y analiza si alguna actividad es *crítica.*

Operaciones 40:

k	**Actividad**	**Inicio Temprano**	**Inicio Tardío**	**Holgura**
1	Cimientos	0	7	7
2	Alcantarillas	0	17	17
3	Base de hormigón	15	22	7
4	Estructura y paredes	19	26	7
5	Tejado	22	53	31
6	Instalación eléctrica primaria	22	32	10
7	Calefacción y aire acondicionado	22	29	7
8	Tabiques	35	42	7
9	Acabados interiores	53	60	7

Tabla 5.3: *Holgura* en días del momento para iniciar las actividades.

Solución: El *tiempo* de inicio tardío y la *holgura* del tiempo de inicio para cada una de las actividades, puede encontrarse en la Tabla 5.3, donde puede verse que no hay actividades críticas.

General (Método del camino más largo - tiempo de inicio tardío)

- Notar que este método calcula el camino más largo desde el nodo $K+1$ (nodo final) a cada nodo p mediante el método de optimización denominado *'Programación dinámica'*.
- Este método es análogo al método del camino crítico pero para *distinguirlo,* lo hemos denominado método del camino más largo.
- *Input:*
  1) a_k : *duración* estimada de la actividad $k \in \mathcal{K} = \{1, \ldots, K\}$.
  2) Conjunto de actividades *sucesoras* de cada actividad $k \in \mathcal{K}$.
  3) Nodo final: $K+1$.
- *Output:*
  1) $\widehat{v}[k]$: *Tiempo de inicio tardío* de la actividad $k \in \mathcal{K}$.
  2) $\widehat{d}[k]$: *Sucesor* de k en el camino más largo de k al nodo final, $k \in \mathcal{K}$.
  3) *Caminos más largos:* Se reconstruyen a partir de los $\widehat{d}[k]$.
- *Paso 0:* Inicialización.
  1) *Numerar* los nodos de forma que los arcos (i,j) de la red del proyecto cumpla $i < j$.
  2) *Inicializar* $\widehat{v}[K+1] =$ 'Plazo para acabar el proyecto'.
- *Paso 1:* Criterio de parada.
  1) Si ya hemos calculado el tiempo de inicio tardío para el nodo 1 *paramos.*
  2) En caso contrario, sea p el *nodo* no procesado con la etiqueta más grande.

- *Paso 2:* Procesado.

    1) Calcular el *tiempo de inicio tardío* para p :

    $$\widehat{v}[p] = \min\{\widehat{v}[i] \mid i \text{ es un sucesor de } p\} - a_p.$$

    2) Guardar en $\widehat{d}[p]$ el nodo i para el que se alcanza el anterior *mínimo*.
    3) $\widehat{d}[p]$ es el *sucesor* de p en su camino más largo.
    4) Volver al *Paso 1*.

5.2. Gestión de proyectos mediante programación lineal

- Numerosos aspectos de la gestión de proyectos pueden ser abordados mediante la *programación lineal* (PL).

- En particular, el cálculo de los *tiempos* asociados a la gestión de un proyecto pueden ser abordados mediante PL tal como veremos en la siguiente sección.

- Otros aspectos de la gestión de proyectos que también pueden ser abordados mediante PL son:

    ▷ Asignación de *recursos* (operarios, máquinas, etc.) a las actividades de un proyecto.
    ▷ *Selección* de los proyectos más interesantes de un conjunto de proyectos candidatos.
    ▷ *Planificación* de la ejecución de un proyecto.
    ▷ Etc.

5.2.1. Gestión de los tiempos de un proyecto mediante programación lineal

Ejemplo 41 (Construcción de un almacén - continuación)

Datos: Consideramos el proyecto para la *construcción* de un almacén, visto en las secciones anteriores.

Objetivo: *Modeliza* mediante un problema de programación lineal (PL):

1) El cálculo del *tiempo de inicio temprano* para la actividad 4.
2) El cálculo del *tiempo de inicio tardío* para la actividad 4.
3) El cálculo de la *duración mínima* del proyecto.

Operaciones 41:

Solución: Ver apartado 'Operaciones'.

General (Gestión de tiempos mediante PL)

- En las secciones anteriores hemos supuesto que la duración de una actividad era un *dato*.
- Sin embargo, la duración de una actividad dependerá de los *recursos* que destinemos para ejecutarla (número de máquinas, número de operarios, etc.)
- Al mismo tiempo, el *coste* de una actividad en general crece al acortar su duración (hay que pagar horas extras, etc.)
- Una forma sencilla de *modelizar* el coste para realizar una actividad en función del tiempo de ejecución x_k es
$$c_k(x_k) = n_k - m_k x_k \qquad k \in \mathcal{K} = \{1, \ldots, K+1\},$$
donde n_k y m_k son positivos y x_k es una variable de decisión (cuando era un dato lo hemos llamado a_k).

- Sea t_k el *tiempo* transcurrido antes de iniciar la actividad $k \in \mathcal{K}$ ($K+1$ corresponde al 'nodo final').
- Sea $\mathcal{J}(k)$ el conjunto de actividades *predecesoras* de la actividad $k \in \mathcal{K}$.
- De cara a calcular el *coste* mínimo de ejecución para un proyecto podemos definir el siguiente *PL*

$$
\begin{array}{lll}
\min & \displaystyle\sum_{k \in \mathcal{K}} c_k(x_k) & \\
s.a & t_1 \geq 0 & \\
 & t_k \geq t_j + x_j & k \in \mathcal{K}, j \in \mathcal{J}(k) \\
 & \underline{x}_k \leq x_k \leq \overline{x}_k & k \in \mathcal{K} \\
 & t_{K+1} \leq \bar{t}_{K+1} & \\
 & t \geq 0. &
\end{array}
$$

5.3. Diagrama de Gantt

General

- El diagrama de Gantt es una herramienta *gráfica* para la gestión de proyectos (ver Figura 5.2).
- Su objetivo es mostrar el *tiempo* de dedicación previsto para las actividades de un proyecto.
- Su formato básico es:
  - ▷ Representar en el eje X el *tiempo.*
  - ▷ Representar en el eje Y las *actividades* del proyecto.
  - ▷ Representar por medio de una barra horizontal el *comienzo* y la *duración* de cada actividad.
- Para la gestión de proyectos complejos (superiores a *25* actividades) se requiere además el uso de técnicas basadas en *redes* de precedencia (como el método del camino crítico).

5.4. Contexto aleatorio

- Todos los cálculos hechos hasta ahora se han basado en datos *estimados.*
- Por ejemplo, a_k corresponde a la duración estimada de la actividad k.
- Sin embargo, la duración de una actividad está sujeta a *incertidumbre* por lo que puede ser modelizada como una variable aleatoria.
- En este contexto aleatorio, de cara a gestionar un proyecto podemos usar el método denominado *PERT* (Project Evaluation and Review Techniques)
  - ▷ El método PERT tiene los mismos *objetivos* que el método del camino crítico (tiempos de inicio temprano y tardío, duración mínima del proyecto, etc.)
  - ▷ La diferencia fundamental es que el método PERT tiene en cuenta que la duración de cada actividad es *aleatoria.*

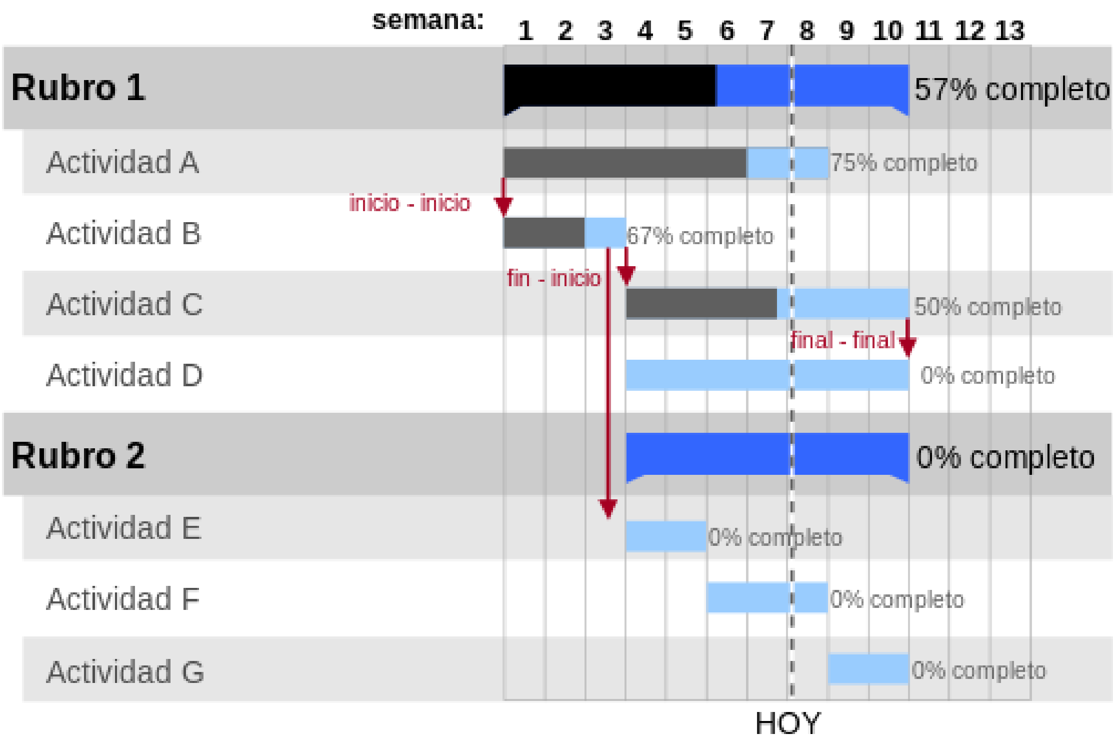


Figura 5.2: *Diagrama de Gantt.*

▷ El método PERT puede consultarse, por ejemplo, en el libro 'Investigación de Operaciones', F. S. Hillier, G. J. Lieberman, Ed. McGraw-Hill (2002).

Capítulo 6

Control estadístico de la calidad

6.1. Introducción

General (Relevancia de la calidad)

- A partir de ahora hay que traer a clase los *enunciados impresos* de los ejercicios y la *calculadora* para poder hacer ejercicios en clase. Se asignará una *nota* al grupo.
- Entendemos 'calidad' como sinónimo de *'aptitud para el uso'* presente y futuro.
- Muchos de nosotros hemos tenido experiencias de *mala calidad* al adquirir un producto o un servicio.
- Nuestra sensación es peor cuando los empleados no están autorizados o dispuestos a corregir problemas de calidad.
- Las consecuencia principal de tal actitud es la *perdida de clientes* (a corto y medio plazo).
- Las empresas con *éxito* entienden que la calidad tiene gran *impacto* en sus negocios y por esta razón fomentan políticas para controlar/mejorar la calidad de sus productos.

Nota. Este tema también se puede consultar y/o ampliar en el Capítulo 8 del siguiente libro: Montgomery, D. C., Runger, G. C., and Hubele, N. F. 'Engineering Statistics', Editorial Wiley, 2006.

General (Mejora de la calidad y Estadística)

- Para muchos consumidores dos de los factores más importantes de cara valorar un producto o servicio son:
  - ▷ *Precio:* los clientes buscamos el 'low cost'.
  - ▷ *Calidad:* tanto en el diseño como en la fabricación.
- Para lograr estos dos objetivos los métodos estadísticos son de gran importancia. La Estadística aplicada a la mejora de la calidad se basa en dos *herramientas* fundamentales:
  - ▷ *'Diseño de experimentos en ingeniería',* útil en el proceso de diseño y desarrollo de los productos (lo veremos en el próximo capítulo).

▷ *'Control estadístico de procesos',* útil para monitorizar los procesos de producción (lo veremos en ese capítulo).

General (Control estadístico de procesos)

- El 'control estadístico de procesos' es sinónimo de *'control estadístico de la calidad'.*

- Un proceso es una combinación única de herramientas, métodos, materiales y personal dedicados a la labor de producir un *resultado medible.*

- Por ejemplo, una *línea de producción* para ensamblar los componentes de un PC es un proceso.

- Todos los procesos tienen una *variabilidad* inherente que puede evaluarse por medio de métodos estadísticos. Por ejemplo:

  ▷ Los sucesivos *productos* fabricados en un proceso de producción no son exactamente idénticos.

  ▷ Los *tiempos* que tardan los usuarios en recibir un determinado servicio tampoco son idénticos,

  ▷ etc.

- Se considera que este tipo de variación, llamada *variación aleatoria,* es inherente al sistema.

- Hay otro tipo de variación que no es inherente al sistema. Por ejemplo:

  ▷ La *avería* de una máquina de la cadena de producción pude producir productos defectuosos.

  ▷ Una materia prima *defectuosa* puede originar también productos defectuosos.

- Este tipo de variación, llamada *variación asignable,* debe ser asignada a una causa para ser corregida.

- Cuando la única variación presente es aleatoria, se dice que el proceso está *bajo control.*

- En caso contrario se dice que el proceso está *fuera de control.*

- El objetivo del control estadístico de procesos consiste en *monitorizar* si un proceso está bajo control.

- El control estadístico de procesos utiliza diferentes *herramientas estadísticas,* como el histograma, y especialmente los gráficos de control.

General (Gráficos de control)

- En la Figura 6.1 tenemos un *gráfico de control* típico.

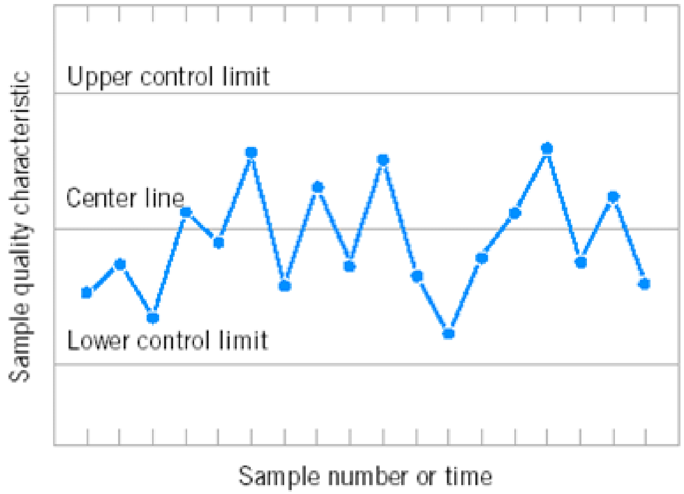


Figura 6.1: Gráfico de control típico.

- El objetivo de los gráficos de control es reconocer si un proceso particular está *bajo control.*
- Los tipos de gráficos de control que se consideran vienen determinados por tres valores:
  - ▷ Límite de control *superior* (UCL, 'upper control limit').
  - ▷ Límite de control *inferior* (LCL, 'lower control limit').
  - ▷ Línea *central* (CL, 'center line').
- Si para controlar un proceso industrial usamos, por ejemplo, la media muestral $\bar{X}$, entonces tenemos el *gráfico de control* $\bar{X}$, que se representa como sigue:
  - ▷ Primero se toman diferentes muestras de la *magnitud* o característica del proceso industrial que se quiere estudiar.
  - ▷ Después se calcula la *media* para cada muestra y se representa en el gráfico de control, usando los siguientes *límites de control:*

$$\begin{aligned} UCL &= \mu_{\bar{X}} + k\,\sigma_{\bar{X}} \\ CL &= \mu_{\bar{X}} \\ LCL &= \mu_{\bar{X}} - k\,\sigma_{\bar{X}} \end{aligned}$$

    donde habitualmente se toma $k = 3$, $\mu_{\bar{X}}$ es el valor esperado de $\bar{X}$ y $\sigma_{\bar{X}}$ su desviación típica.
  - ▷ Cuando todas las medias muestrales caen dentro de los límites de control se concluye que el proceso está *bajo control.*
- En las siguientes secciones veremos los siguientes tipos de gráfico de control:
  - ▷ *Gráficos de control* $\bar{X}$ (para controlar la media).
  - ▷ *Gráficos de control R* (para controlar la varianza).

6.2. Gráficos de control X y R

Ejemplo 42 (Fabricación de aros para pistón)

Datos:

- Consideramos la fabricación de *aros para pistón* de motor.
- Una característica crítica en la calidad de un aro es su *diámetro interno.*
- Debido a las *variaciones* en su proceso de producción, podemos modelizar este diámetro como una variable aleatoria.
- Sea X la *variable aleatoria* 'diámetro interno' (en mm).
- Suponemos que la distribución de X es *Normal* de media $\mu = 74$ mm y desviación típica $\sigma = 0,01$ mm. Es decir:
$$X \sim N(\mu = 74, \sigma = 0,01).$$
- En un control de calidad se toman 16 *muestras* de tamaño $n = 5$ y se calcula las correspondientes medias muestrales (ver Tabla 6.1).

Muestra i	**Media Muestral** $\bar{x}_i$
1	73,9930
2	74,0080
3	73,9980
4	74,0060
5	73,9990
6	74,0102
7	74,0045
8	73,9870
9	74,0011
10	74,0062
11	73,9957
12	73,9946
13	74,0071
14	74,0090
15	73,9968
16	74,0081

Tabla 6.1: *Media* de X para las 16 muestras (en mm).

Objetivo:

1) *Calcula* el límite de control superior *(UCL),* el límite de control inferior *(LCL),* y el valor de la línea central *(CL)* asociados a $\bar{X}$.
2) *Representa* el gráfico de control $\bar{X}$.

Operaciones 42:

Solución:

1) Los límites pedidos son $UCL = 74,0135$, $LCL = 73,9865$ y $LC = 74$ (mm).
2) El gráfico de control $\bar{X}$ lo tenemos en la Figura 6.2.

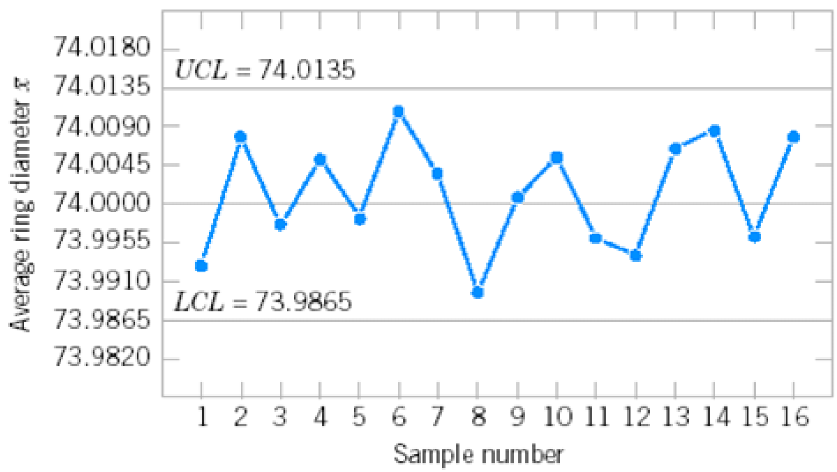


Figura 6.2: *Gráfico de control* $\bar{X}$.

General (Gráficos de control $\bar{X}$ con media y desviación conocidas)

- Los gráficos de control $\bar{X}$ se denominan también 'gráficos de control *Shewhart*' (pronunciado como 'Shu-jart') en honor al ingeniero que introdujo ente tipo de gráficos.

- Se considera un *proceso* (fabricación, servicio, etc.) y se estudia una de sus magnitudes modelizada mediante una variable aleatoria X con media μ y con desviación típica σ.

- El gráfico de control $\bar{X}$ se utiliza para detectar posibles *deslizamientos* de la media μ.

- En este apartado suponemos que *conocemos* μ y σ.

- Para *representar* el gráfico de control $\bar{X}$ necesitamos los siguientes parámetros:

$$\begin{aligned} UCL &= \mu_{\bar{X}} + 3\,\sigma_{\bar{X}} \\ CL &= \mu_{\bar{X}} \\ LCL &= \mu_{\bar{X}} - 3\,\sigma_{\bar{X}} \end{aligned}$$

  donde $\mu_{\bar{X}}$ es el valor esperado de $\bar{X}$ y $\sigma_{\bar{X}}$ su desviación típica.

- Sabemos que $\mu_{\bar{X}} = \mu$ y que

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

donde n es el *tamaño muestral.*

Ejemplo 43 (Fabricación de aros para pistón - continuación)

Datos:

- Continuamos con el ejemplo anterior.
- En este ejemplo usamos la variable aleatoria *normal* estándar que denotamos por Z.
- También usamos la función de distribución de probabilidad de Z que denotamos por

$$\Phi(z) = P(Z \leq z).$$

- La *tabla* de $\Phi(z)$ se puede consultar en el apéndice del final de este dosier.

Objetivo:

1) ¿Podemos decir que el proceso de fabricación de aros de pistón está *'bajo control'?*
2) Suponiendo que el proceso de fabricación de aros esté bajo control ¿cuál es la *probabilidad* de que la media muestral caiga fuera de los límites de control?

Operaciones 43:

Solución:

1) *Sí,* pues la media muestral de cada muestra cae dentro de los límites de control y de forma aleatoria alrededor de la media.
2) La *probabilidad* de que la media muestral caiga fuera de los límites de control es del $0,26\,\%$ (1 de cada 385) suponiendo que el proceso de fabricación de aros está bajo control.

General (Gráficos de control $\bar{X}$ y contraste de hipótesis)

- En el ejemplo anterior hemos visto que aunque el sistema esté bajo control tenemos una probabilidad del $0,26\,\%$ de declararlo fuera de control *(error del tipo I).*

- En esencia los gráficos de control equivalen a realizar una secuencia de *contrastes de hipótesis* (tantos como muestras).

- Para cada muestra, implícitamente realizamos el siguiente contraste de hipótesis:

$$\begin{aligned} H_0 : &\quad \text{media} = \mu_0 \quad \text{(sistema bajo control)} \\ H_1 : &\quad \text{media} \neq \mu_0 \quad \text{(sistema fuera de control),} \end{aligned}$$

al nivel de significación $\alpha = 0,0026$.

- Recordamos que el *nivel de significación* α mide el riesgo de tipo I, es decir,

$$\begin{aligned}
\alpha &= P\Big(\text{Equivocarnos al aceptar } H_1\Big) \\
&= P\Big(\text{Declarar que el sistema está} \\
&\qquad \textit{fuera de control} \quad | \quad \text{el sistema está bajo control}\Big) \\
&= 0,0026.
\end{aligned}$$

- La barra vertical anterior '|' se lee: 'dado que'.

General (Gráficos de control $\bar{X}$ con media y desviación desconocidas)

- Hasta ahora hemos supuesto que la magnitud estudiada X, tenía un distribución *Normal* con media μ y desviación típica σ *conocidas.*

- En general, por el *teorema central del límite,* podemos suponer la hipótesis de normalidad (además de poder hacer los correspondientes test de normalidad).

- Sin embargo, los parámetros μ y σ normalmente son *desconocidos* y tendremos que estimarlos.

- Podemos estimar μ mediante la denominada *'gran media'* o *'media de medias':*

$$\bar{\bar{x}} = \frac{1}{m}\sum_{i=1}^{m} \bar{x}_i,$$

  donde m es el número de muestras y $\bar{x}_i$ es la media para la muestra i.

- Para cuantificar la *variabilidad* del proceso, aunque se puede estimar la desviación típica, en este contexto se utiliza más el *rango promedio:*

$$\bar{r} = \frac{1}{m}\sum_{i=1}^{m} r_i,$$

  donde r_i es el rango de la muestra i (valor mayor menos valor menor de la muestra).

- En caso de desconocer μ y σ, para representar el *gráfico de control* $\bar{X}$ necesitamos los siguientes parámetros:

$$\begin{aligned}
UCL &= \bar{\bar{x}} + A_2(n)\,\bar{r} \\
CL &= \bar{\bar{x}} \\
LCL &= \bar{\bar{x}} - A_2(n)\,\bar{r}.
\end{aligned}$$

- n es el *tamaño* de cada muestra, $\bar{\bar{x}}$ y $\bar{r}$ son los valores observados de la *gran media* y del *rango promedio,* respectivamente.

- La *tabla* para los valores de $A_2(n)$ puede encontrarse en el apéndice, al final de este dosier.

General (Gráficos de control R)

- Los gráficos de control $\bar{X}$ están diseñados para detectar posibles deslizamientos de la *media* poblacional μ.

- Por otro lado, nos puede interesar recoger los cambios producidos en la *desviación típica* poblacional σ (para analizar la variabilidad del proceso).

  ▷ En estos casos se puede usar el *gráfico de control S* que se basa en la desviación típica muestral S (implícitamente estamos suponiendo que desconocemos σ).

  ▷ Sin embargo, en la práctica se usa más frecuentemente el *gráfico de control R* basado en el rango, que también permite monitorizar la variabilidad.

- Para representar el *gráfico de control R* necesitamos los siguientes parámetros:

$$\begin{aligned} UCL &= D_4(n)\,\bar{r} \\ CL &= \bar{r} \\ LCL &= D_3(n)\,\bar{r}. \end{aligned}$$

- n es el *tamaño* de cada muestra y $\bar{r}$ es el valor observado del *rango promedio.*

- La *tabla* para los valores de $D_3(n)$ y $D_4(n)$ puede encontrarse en el apéndice, al final de este dosier.

- Representamos los rangos observados r_i en el gráfico de control.

- Cuando todos los rangos caen dentro de los límites de control del gráfico R, se concluye que el proceso tiene *variabilidad constante.*

Ejemplo 44 (Fabricación de turbinas)

Datos:

- En la fabricación de motores de avión la *turbina* es un elemento fundamental.
- Éstas se fabrican mediante un proceso de *fundición* de precisión.
- Una magnitud importante es la *apertura de las paletas* de las turbinas, que denotamos por X.
- La finalidad del siguiente control de calidad es verificar la *estabilidad* estadística de este proceso de fabricación de turbinas.
- En la tabla de la Figura 6.3 tenemos los datos de 20 *muestras* aleatorias de tamaño 5 para X.

Sample Number	x_1	x_2	x_3	x_4	x_5	x	r
1	33	29	31	32	33	31.6	4
2	33	31	35	37	31	33.4	6
3	35	37	33	34	36	35.0	4
4	30	31	33	34	33	32.2	4
5	33	34	35	33	34	33.8	2
6	38	37	39	40	38	38.4	3
7	30	31	32	34	31	31.6	4
8	29	39	38	39	39	36.8	10
9	28	33	35	36	43	35.0	15
10	38	33	32	35	32	34.0	6
11	28	30	28	32	31	29.8	4
12	31	35	35	35	34	34.0	4
13	27	32	34	35	37	33.0	10
14	33	33	35	37	36	34.8	4
15	35	37	32	35	39	35.6	7
16	33	33	27	31	30	30.8	6
17	35	34	34	30	32	33.0	5
18	32	33	30	30	33	31.6	3
19	25	27	34	27	28	28.2	9
20	35	35	36	33	30	33.8	6
						$\bar{\bar{x}} = 33.32$	$\bar{r} = 5.8$

Figura 6.3: *Apertura* de las paletas de las turbinas.

- Los datos de la tabla se han *codificado* usando los últimos dos dígitos de la correspondiente observación. Los dos primeros dígitos son siempre 0,50.
- Por *ejemplo* el primer dato es 33 y representa el valor observado 0,5033 pulgadas.
- *Unidades:* Las unidades de '33' son 10^{-4} *pulgadas,* pues:

$$\text{'33' corresponde a} \quad 0{,}0033 \text{ pulgadas} \quad = \quad 33 \cdot 10^{-4} \text{ pulgadas.}$$

Objetivo:

1) Representa el gráfico de control $\bar{X}$.
2) Representa el gráfico de control R.

Operaciones 44:

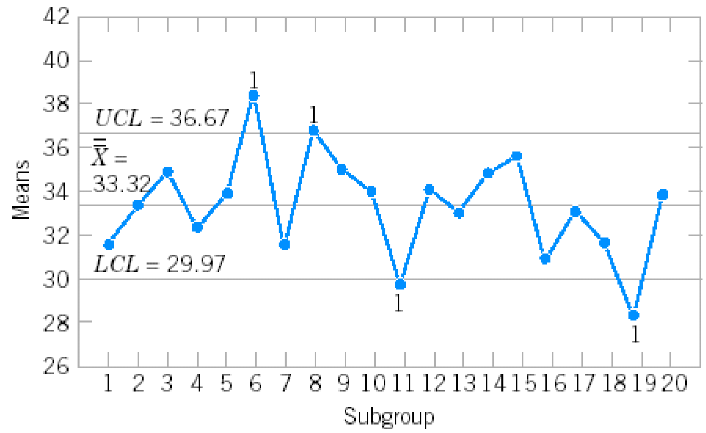


Figura 6.4: *Gráfico de control* $\bar{X}$.

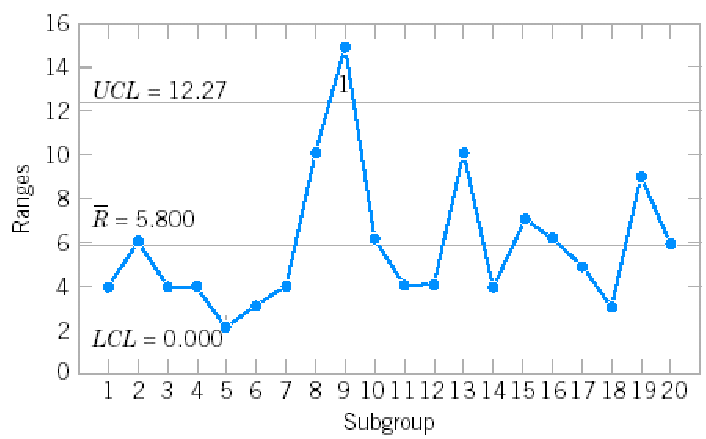


Figura 6.5: *Gráfico de control R.*

Solución: Ver Figuras 6.4 y 6.5.

Ejemplo 45 (Fabricación de turbinas - continuación)

Datos:

- En las Figuras 6.4 y 6.5 podemos observar que:
  - ▷ En el gráfico de control $\bar{X}$ las muestras 6, 8, 11 y 19 están *fuera de los límites.*
  - ▷ En el gráfico de control R la muestra 9 está fuera de los límites.
  - ▷ En total tenemos *cinco* muestras patológicas entre los dos gráficos.
  - ▷ *Eliminamos* las cinco muestras patológicas y *actualizamos* $\bar{\bar{x}} = 33,21$ y $\bar{r} = 5,0$ (unidades: 10^{-4} pulgadas).
- Supongamos que se han podido determinar las causas que han producido estas variaciones y que se han *subsanado.*

Objetivo:

1) Representa el nuevo gráfico de control $\bar{X}$ que no tenga en cuenta estas cinco *muestras patológicas.*
2) Idem para el nuevo gráfico de control R.

Operaciones 45:

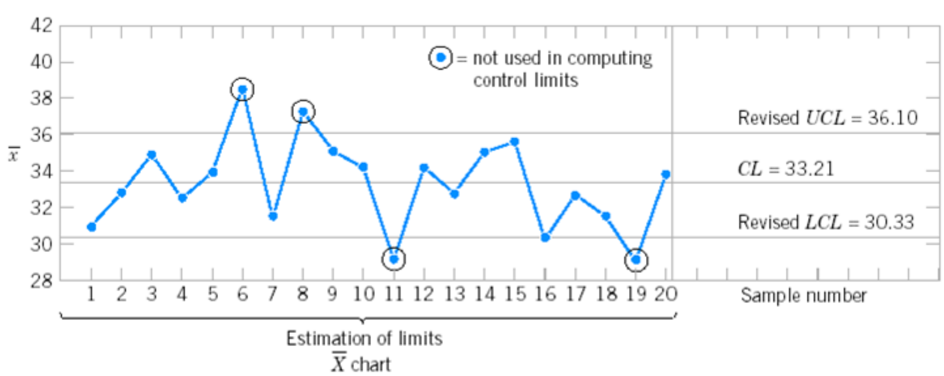


Figura 6.6: Gráfico de control $\bar{X}$ *actualizado*

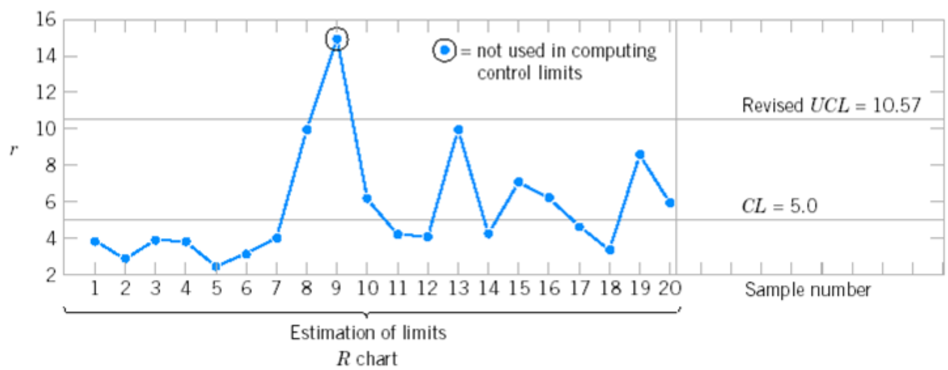


Figura 6.7: Gráficos de control R *actualizado*

Solución: Ver Figuras 6.6 y 6.7.

General (Protocolo para usar los gráficos de control $\bar{X}$ y R)

- Al usar estos gráficos de control normalmente necesitaremos *dos etapas.*
- *Etapa 1:* En esta etapa se *calibran* sus parámetros.
  1) Para ello usamos m *muestras* de tamaño n. Se recomienda tomar m entre 20 y 25, y n entre 4 y 6.
  2) Se analiza mediante un gráfico de control R si la variabilidad del proceso es *constante.*
  3) Se analiza mediante un gráfico de control $\bar{X}$ si el proceso está *bajo control.*
  4) Si hay muestras *patológicas* en cualquiera de los dos gráficos:
     a) Se subsanan las posibles causas que las han producido.
     b) A continuación se eliminan estas muestras y se recalculan los parámetros de los dos gráficos de control.
     c) Se vuelve a la Etapa 1.2.
  5) Si no hay muestras patológicas, se considera que el proceso está *bajo control* y que el nivel de variabilidad es *constante.*
- *Etapa 2:* En esta etapa se usan los gráfico de control $\bar{X}$ y R para *monitorizar* el estado del proceso de fabricación.

1) Para cada *nueva muestra* analizada, el valor de $\bar{x}$ y r se representa en su correspondiente gráfico.
2) De forma periódica se revisan los *límites de control* de los dos gráficos.

6.3. Capacidad de un proceso

Veremos dos casos:

- Caso con la media *centrada.*
- Caso con la media *no centrada.*

6.3.1. Capacidad de un proceso con la media centrada

Ejemplo 46 (Fabricación de turbinas - continuación)

Datos:

- Estamos considerando el proceso de fabricación de *turbinas* para motor de avión.
- En este proceso se consideran *aceptables* las turbinas con una apertura de paletas nominal de 0,5030 pulgadas y una tolerancia de $\pm$ 0,0010 pulgadas.
- Por lo tanto se considera aceptable una apertura de paletas en el *intervalo* [0,5020 , 0,5040] pulgadas.
- *Codificando* los datos respecto a los dos últimos decimales (los que realmente varían), se considera aceptable una apertura de paletas en el intervalo [20, 40].
- Estos dos valores se denominan 'límite *inferior* de la especificación' y 'límite *superior* de la especificación':
  - $\triangleright$ LSL = 20 ('Lower Specification Limit')
  - $\triangleright$ USL = 40 ('Upper Specification Limit')
- Recordamos que el *tamaño* de cada muestra era $n = 5$,
- y que el *rango* promedio era $\bar{r} = 5,0 \cdot 10^{-4}$ pulgadas (una vez eliminadas la cinco muestras patológicas).

Objetivo:

1) Dados los datos sobre la apertura de las paletas recogidos en la Tabla 6.3, representa el correspondiente *diagrama de tolerancia* e histograma.
2) Estima la *desviación típica* de la variable aleatoria X (apertura de paletas).
3) Estima el *coeficiente de capacidad* de este proceso.
4) Los resultados del proceso de fabricación ¿qué *proporción* ocupan del intervalo de especificaciones? (proporción estimada).

Operaciones 46:

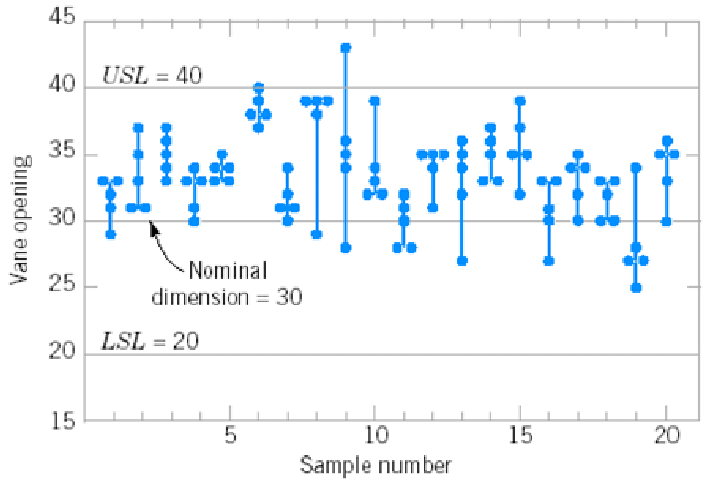


Figura 6.8: *Diagrama de tolerancia* de la apertura de las paletas.

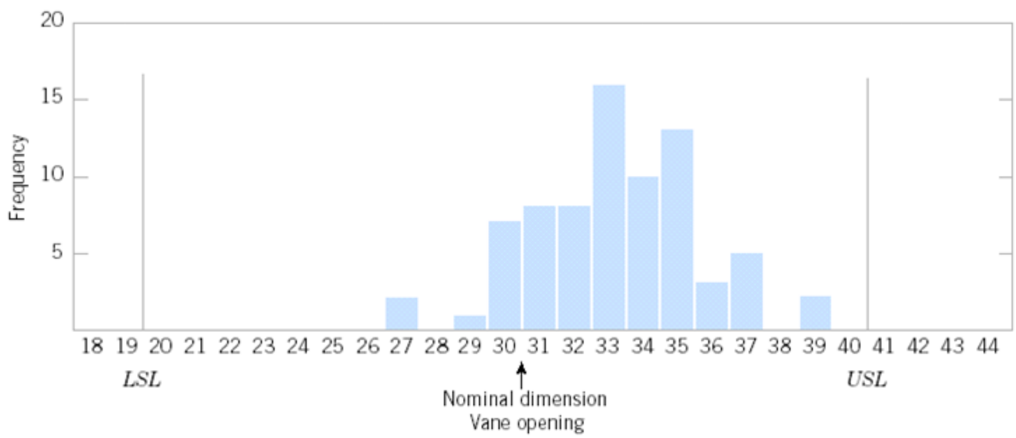


Figura 6.9: *Histograma* de la apertura de las paletas. Podemos asumir que esta variable tiene una *distribución Gaussiana.*

Solución:

1) Ver Figuras 6.8 y 6.9.
2) La estimación para la *desviación típica* es $\hat{\sigma} = 2,15 \cdot 10^{-4}$ pulgadas.
3) La estimación para el *coeficiente de capacidad* es $\hat{C}_p = 1,55$ (sin unidades).
4) El proceso de fabricación ocupa un 64,5 % del intervalo de especificaciones (este valor es una *estimación).*

General (Capacidad de un proceso)

- Analizar la 'capacidad de un proceso' significa analizar su *rendimiento* (suponiendo que está operando bajo control).

- Uno de sus objetivos principales es *predecir* la proporción de productos fabricados que *cumplirá* con las especificaciones.

- Las *especificaciones* vienen dadas por un límite inferior y un límite superior: LSL y USL, respectivamente.

- Como hemos dicho, analizar la capacidad de un proceso sólo tiene sentido si el proceso está *bajo control.*

- Para analizar la capacidad de un proceso tenemos *herramientas* gráficas y numéricas.

- Entre las herramientas gráficas destacamos:

  ▷ El *histograma.*

  ▷ El *gráfico de tolerancia,* donde representamos cada muestra en un eje vertical (ver Figura 6.8).

  ▷ El gráfico de tolerancia no tiene que *confundirse* ni mezclarse con los gráficos de control $\bar{X}$ y R.

- Entre las herramientas numéricas tenemos:

▷ Una estimación de la *media* μ mediante la denominada 'gran media' o 'media de medias':

$$\bar{\bar{X}} = \frac{1}{m}\sum_{i=1}^{m}\bar{X}_i,$$

donde m es el *número de muestras* y $\bar{X}_i$ es la *media* de la muestra i.

▷ Una estimación de *desviación típica* σ mediante:

$$\hat{\sigma} = \frac{\bar{r}}{d_2(n)}$$

donde n es el *tamaño* de cada muestra y el *parámetro* $d_2(n)$ puede obtenerse en la correspondiente *tabla* del apéndice, al final de este dosier.

▷ El *'coeficiente de capacidad del proceso':*

$$C_p = \frac{USL - LSL}{6\sigma},$$

que en caso de desconocer σ puede ser estimado por

$$\hat{C}_p = \frac{USL - LSL}{6\hat{\sigma}}.$$

- En inglés C_p se denomina 'Process Capability Ratio' (PCR) - ver Figura 6.10.

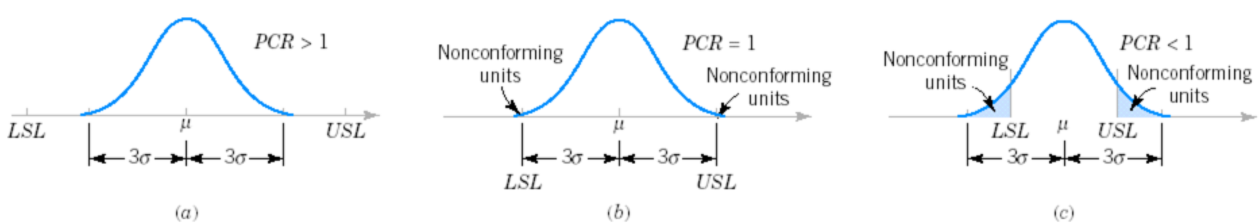


Figura 6.10: Productos *defectuosos* y coeficiente de capacidad C_p (PCR).

- $C_p \geq 1$ indica que prácticamente todos los productos fabricados *cumplirán* con las especificaciones (ver Figura 6.10, gráficas a) y b)).
- $C_p < 1$ indica que un número relevante de productos fabricados *no cumplirán* las especificaciones(ver Figura 6.10 gráfica c)).
- $1/C_p$ tiene una *interpretación* natural, pues representa la proporción del intervalo de especificaciones ocupado por los valores del proceso.

6.3.2. Capacidad de un proceso con la media no centrada

Ejemplo 47 (Fabricación de turbinas - continuación)

Datos:

- En la Figura 6.9 podemos observar que el proceso no está centrado en 30, el *valor nominal* para la apertura de paletas.
- En este caso la *capacidad real* del proceso es menor que la capacidad indicada por C_p por lo que es recomendable usar el 'coeficiente de capacidad *ajustado*' C_{pk}.

Objetivo:

1) *Estimar* el ‘coeficiente de capacidad ajustado’ de este proceso.

2) *Comparar* el ‘coeficiente de capacidad ajustado’ con el ‘coeficiente de capacidad’ y sacar conclusiones.

3) Estimar qué *proporción* de productos fabricados no cumplirán las especificaciones.

Operaciones 47:

Solución:

1) Una estimación del coeficiente de capacidad *ajustado* es $\hat{C}_{pk} = 1,05$.
2) El proceso está *desplazado* del centro del intervalo de especificaciones.
3) Se estima que un 0,08 % de las turbinas *no cumplirá* las especificaciones.

General (coeficiente de capacidad ajustado)

- El coeficiente de capacidad C_p presupone que μ está en el *centro* del intervalo de especificaciones (es decir μ coincide con la especificación nominal).

- Cuando no se da ese caso, como en el ejemplo anterior, es mejor usar el coeficiente de capacidad *ajustado* C_{pk} :

$$C_{pk} = \min\left[\frac{USL-\mu}{3\sigma},\ \frac{\mu-LSL}{3\sigma}\right],$$

que en caso de desconocer μ y σ pueden ser reemplazadas por sus estimaciones $\bar{\bar{x}}$ y $\hat{\sigma}$ para así obtener la *estimación* $\hat{C}_{pk}$.

General (Proporción de productos defectuosos)

- Como ya mencionamos, uno de los objetivos principales al analizar la capacidad de un proceso es *predecir* la proporción de productos defectuosos.

- Un producto es *defectuoso* cuando la variable de control X no cumple con las especificaciones, es decir, no está en el intervalo $[LSL\ ,\ USL]$.

- La *proporción* de productos defectuosos puede obtenerse sumando las dos probabilidades siguientes:

$$P(X < LSL) = P\left(Z < \frac{LSL-\mu}{\sigma}\right) \tag{6.1}$$

$$P(X > USL) = P\left(Z > \frac{USL-\mu}{\sigma}\right), \tag{6.2}$$

donde Z tiene una distribución normal estándar.

- Si en las anteriores fórmulas *desconocemos* μ y σ podemos realizar un cálculo aproximando usando su estimación $\bar{\bar{x}}$ y $\hat{\sigma}$, respectivamente.

- Por lo tanto, para controlar si la *'variación aleatoria'* de un proceso es aceptable o no con respecto a las especificaciones tenemos dos opciones:

  1) Calcular la *proporción* de productos defectuosos mediante las fórmulas (6.1-6.2).
  2) Calcular el *coeficiente de capacidad* C_p o C_{pk}.

6.4. Metodología Seis Sigma

Ejemplo 48 (Fabricación de turbinas - continuación)

Datos: En el ejemplo anterior hemos visto que el proceso de fabricación de turbinas tiene un *coeficiente de capacidad* ajustado $C_{pk} = 1,05$.

Objetivo: Analizar si este proceso de fabricación de turbinas tiene una eficiencia de *6 sigma:*

1) *Numéricamente:* Un proceso tiene una eficiencia 6 sigma cuando tiene un *coeficiente de capacidad* ajustado $C_{pk} = 2,0$.
2) *Gráficamente:* Un proceso tiene una eficiencia 6 sigma cuando la *media* del proceso está a una distancia 6σ del límite de especificación más cercano.

Operaciones 48:

Solución: El proceso de fabricación de turbinas no tiene una eficiencia de 6 sigma.

General (Seis Sigma)

- Como ya hemos mencionado, para medir si la *'variación aleatoria'* de un proceso bajo control es aceptable o no con respecto a las especificaciones (LSL y USL), una opción es usar el *coeficiente de capacidad* C_p o C_{pk}.

- Si μ está en el *centro* del intervalo de especificaciones entonces podemos usar C_p e imponer $C_p \geq 1$. Muchas compañías añaden un *margen de seguridad* y usan:

- ▷ $C_p = 1,33$ como valor *mínimo* aceptable.
- ▷ $C_p = 1,66$ como valor mínimo aceptable para características *críticas* del producto (seguridad, resistencia, etc.)

- Si μ *no está* en el centro del intervalo de especificaciones entonces es mejor usar C_{pk}:
  - ▷ En general debe cumplirse que $C_{pk} \geq 1$.
  - ▷ Algunas compañías imponen $C_{pk} = 2,0$ para los procesos internos y para los procesos de los proveedores.

- Proceso con eficiencia *6 sigma.*
  - ▷ Un proceso con $C_{pk} = 2,0$ se dice que tiene una eficiencia 6 sigma.
  - ▷ En ese caso la *distancia* de μ, la media del proceso, al límite de especificación más cercano (LSL o USL) es de 6 veces σ (ver Figura 6.11).

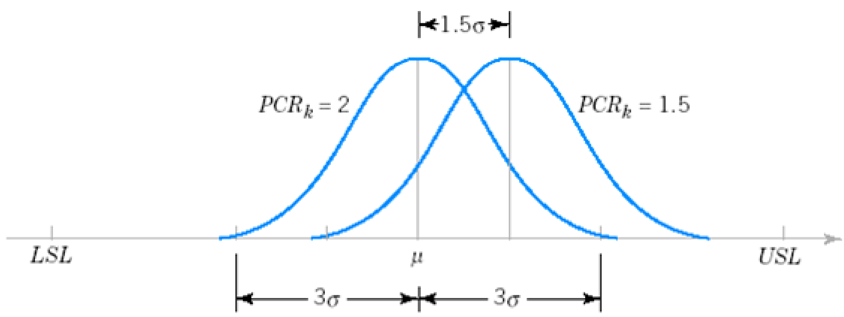


Figura 6.11: Media de un proceso *Seis Sigma* desplazada 1,5 desviaciones típicas.

  - ▷ Bajo estas condiciones la proporción de productos *defectuosos* es prácticamente 0 %.
  - ▷ Sin embargo, es difícil que la media del proceso no sufra *desplazamientos.*
  - ▷ La *ventaja* de la eficiencia 6 sigma radica en que incluso si la media del proceso se desplaza 1,5σ hacia el límite de especificación más cercano:
    - ◦ La proporción de productos defectuosos es *muy baja* (3,4 productos por millon)
    - ◦ C_{pk} pasa de 2,0 a 1,5:

$$C_{pk} = \frac{6\sigma - 1,5\sigma}{3\sigma} = 1,5.$$

- *Metodología 'Seis Sigma':*
  - ▷ En inglés se denomina *'Six Sigma'*.
  - ▷ Es una metodología de *mejora de procesos,* centrada en la reducción de la variabilidad de los mismos de cara a eliminar los defectos en productos o servicios al cliente.
  - ▷ Su *meta* es llegar a procesos con una eficiencia 6 sigma.
  - ▷ La metodología Seis Sigma se caracteriza por *5 etapas* concretas:
    - ◦ Definir el *problema* o el defecto
    - ◦ Medir y recopilar *datos*
    - ◦ Analizar datos
    - ◦ Mejorar

- ◦ Controlar
- ▷ La metodología Seis Sigma se ha vuelto muy popular entre las *grandes empresas* debido a sus buenos resultados.

General (Otros gráficos de control)

- Hemos visto los gráficos de control $\bar{X}$ y R.
- Existen otros tipos de gráficos de control como por ejemplo:
  - ▷ *Gráficos de control P* para analizar la proporción de productos defectuosos.
  - ▷ *Gráficos de control U* para analizar la cantidad de defectos por unidad de producto.
- Para más detalles se puede consultar por ejemplo el Capítulo 8 del siguiente libro: Montgomery, D. C., Runger, G. C., and Hubele, N. F. 'Engineering Statistics', Editorial Wiley, 2006.

Capítulo 7

Diseño de experimentos en ingeniería

Nota: En este capítulo hay que traer a clase los *enunciados impresos* de los ejercicios y la *calculadora* para poder hacer ejercicios en clase. Se asignará una *nota* al grupo.

7.1. Diseño de experimentos y control de la calidad

Ejemplo 49 (Resistencia de una fibra sintética)

Datos:

- Consideramos un proceso de fabricación de una *fibra sintética.*
- En particular estamos interesados en mejorar su *resistencia* a la rotura.
- Conjeturamos las *hipótesis* de que dicha resistencia depende de dos factores: *temperatura* (T) y *presión* (P) a la que se fabrica.

Objetivo: *Diseñar un experimento* en el que manipulando diferentes niveles de T y P:

- *Experimentemos* sobre el efecto que tienen estos dos factores en la resistencia.
- *Analicemos* los resultados con técnicas estadísticas.
- Saquemos *conclusiones* sobre la validez de la hipótesis inicial.

Solución: En este tema veremos una introducción a cómo abordar este tipo de cuestiones.

General (Diseño de experimentos y control de la calidad)

- Tal como hemos visto en el ejemplo anterior, muchas de las *variables* que intervienen en un proceso de fabricación se pueden controlar.
- Son las denominadas *variables de control,* como por ejemplo la presión, la temperatura, la velocidad de la cadena de montaje, etc.
- Controlando estas variables se pretende por un lado controlar y mejorar la *calidad* del producto:
  - ▷ Productos más *adecuados* para realizar su función.
  - ▷ Productos más *fiables* y duraderos.

- ▷ Productos con menos *variabilidad* respecto a las especificaciones nominales de fabricación (productos más *homogéneos).*

- Por otro lado se pretende mejorar la *productividad:*
  - ▷ Reducir el *tiempo para diseñar y desarrollar* nuevos productos.
  - ▷ Reducir los *costes* de producción:
    - ◦ Reducir los *tiempos de producción* (más productos en menos tiempo).
    - ◦ Reducir el coste de las *materias primas.*

- Todo *proceso de experimentación* en ingeniería se compone de *cuatro etapas:*
  1) *Conjeturar una hipótesis* sobre el proceso de producción.
  2) *Experimentar* para obtener información sobre la conjetura.
  3) *Analizar* estadísticamente la información del experimento para ver si es significativa.
  4) *Sacar conclusiones* sobre si la conjetura inicial es cierta o hay que adaptarla a los resultados obtenidos y quizás realizar un nuevo experimento.

Nota. Este tema también se puede consultar y/o ampliar en el Capítulo 7 del siguiente libro: Montgomery, D. C., Runger, G. C., and Hubele, N. F. 'Engineering Statistics', Editorial Wiley, 2006.

7.2. Introducción a los experimentos factoriales

Ejemplo 50 (Fabricación de circuitos integrados)

Datos:

- Consideramos un proceso de fabricación de *circuitos integrados.*
- En una de las caras del circuito integrado, o substrato, se hace crecer una *capa uniforme* de material semiconductor y de poco grosor, medido en micrómetros μm (1 μm = 10^{-6} m).
- Aplicando un *vapor rico en arsénico* se consigue esta capa, cuyo *grosor* conjeturamos que depende de dos factores:
  1) El *tiempo* de aplicación del vapor.
  2) El *nivel de arsénico* en el vapor aplicado.
- Para estudiar esta *conjetura* se diseña el siguiente *experimento:*
  - ▷ Que considera los *factores:* A (tiempo de aplicación) y B (nivel de arsénico).
  - ▷ Para cada factor, se consideran *dos niveles* etiquetados con $+$ (nivel alto) y con $-$ (nivel bajo):
    - ◦ *Factor A:*

$$\begin{aligned} + &= \text{tiempo de aplicación largo} \\ - &= \text{tiempo de aplicación corto.} \end{aligned}$$

    - ◦ *Factor B:*

$$\begin{aligned} + &= \text{nivel de arsénico del } 59\,\% \\ - &= \text{nivel de arsénico del } 55\,\%. \end{aligned}$$

▷ Para que el experimento sea significativo se decide tomar $n = 4$ *muestras* para cada caso (hay que repetir el experimento 4 veces para cada caso).

▷ En la Tabla 7.1 tenemos los *resultados* del experimento.

Combinación de Niveles	**Grosor Muestra 1**	**Grosor Muestra 2**	**Grosor Muestra 3**	**Grosor Muestra 4**	**Suma Total**	**Media** (μm)	**Varianza** (μm^2)
(1)	14,037	14,165	13,972	13,907	56,081	14,020	0,0121
a	14,821	14,757	14,843	14,878	59,299	14,825	0,0026
b	13,880	13,860	14,032	13,914	55,686	13,922	0,0059
ab	14,888	14,921	14,415	14,932	59,156	14,789	0,0625

Tabla 7.1: *Grosor* de la capa uniforme (en μm).

▷ En esta tabla los 4 *símbolos* (1), a, b y ab tienen la siguiente interpretación:

- (1) corresponde al caso en que los dos factores están en su nivel *bajo.*
- a corresponde al caso en que sólo el factor A está en su nivel *alto* (b análogamente).
- ab corresponde al caso en que los dos factores A y B están en su nivel *alto.*

Objetivo:

1) Estimar el *efecto* que el factor A (tiempo de aplicación) produce en el grosor alcanzado por la capa del circuito integrado.
2) Hacer lo mismo con el factor B (nivel de arsénico).

Operaciones 50:

Solución:

- El efecto del *tiempo de aplicación* del gas es considerable y positivo ($A = 0,836\ \mu$m).
- *Interpretación:* Cambiando el tiempo de aplicación de corto a largo, aumenta el grosor promedio de la capa uniforme en $0,836\ \mu$m.
- Sin embargo, el efecto del *nivel de arsénico* en el gas aplicado es bajo y negativo ($B = -0,067\ \mu$m).
- *Interpretación:* Cambiando el nivel de arsénico de bajo a alto, disminuye el grosor promedio de la capa uniforme en $0,067\ \mu$m.

General (Experimentos factoriales $C = 2^2$)

- En el ejemplo anterior hemos visto un experimento de dos factores con dos niveles *(experimento factorial* 2^2).
- En todo experimento factorial se define una *variable respuesta* Y como la magnitud que queremos analizar.
- El objetivo es analizar:
  - ▷ Si la variable respuesta es *afectada* por los factores.
  - ▷ Si existe *interacción* (dependencia) entre los factores.
- En estos experimentos se consideran *dos factores* A y B y *dos niveles* para cada factor que se etiquetan con '+' (nivel alto) y con ' $-$ ' (nivel bajo).
- Debemos analizar las $C = 2^2 = 4$ *combinaciones* de los 2 niveles de cada factor que representamos por los símbolos $(1), a, b$ y ab :
  - ▷ *(1)* Representa que los dos factores están en su nivel *bajo.*
  - ▷ a representa que sólo el factor A está en su nivel *alto.*
  - ▷ b representa que sólo el factor B está en su nivel *alto.*
  - ▷ ab representa los dos factores A y B están en su nivel *alto.*
- El experimento se *repite* n veces (cuanto mayor es n los resultados del experimento son más robustos, pero más caros...).
- En cada repetición del experimento deben hacerse *C=4 pruebas:* una para cada una de las combinaciones (1), a, b y ab.

- En los cálculos, los símbolos (1), a, b y ab representan la *suma total* de los resultados obtenidos para cada combinación de niveles, después de haber repetido el experimento n veces (columna 'Suma Total' de la Tabla 7.1).
- Podemos estimar el efecto promedio del *factor A* mediante la siguiente fórmula:

$$\begin{aligned} A &= \bar{y}_{A+} - \bar{y}_{A-} \\ &= \frac{1}{n}\left[\frac{a+ab}{2} - \frac{b+(1)}{2}\right] \\ &= \frac{1}{2n}\left[a+ab-b-(1)\right], \end{aligned}$$

  donde $\bar{y}_{A+}$ es el *valor promedio* de la magnitud bajo estudio cuando el factor A está en su nivel alto ($\bar{y}_{A-}$ análogamente).
- *Interpretación* del efecto del factor A: Cambiar el factor A del nivel bajo al nivel alto produce un cambio en la variable respuesta Y de A unidades (en promedio).
- Podemos estimar el efecto promedio del *factor B* mediante la siguiente fórmula:

$$\begin{aligned} B &= \bar{y}_{B+} - \bar{y}_{B-} \\ &= \frac{1}{n}\left[\frac{b+ab}{2} - \frac{a+(1)}{2}\right] \\ &= \frac{1}{2n}\left[b+ab-a-(1)\right]. \end{aligned}$$

- *Interpretación* del efecto del factor B: Análogamente al factor A.
- Normalmente al hablar del *efecto* del factor A, nos estaremos refiriendo al *efecto promedio* del factor A, pues los cálculos se realizan con valores promedio (análogamente para B).

Ejemplo 51 (Fabricación de circuitos integrados - cont)

Datos:

- *Interacción* entre dos factores: Existe interacción cuando el efecto de un factor depende del nivel del otro factor.
- *Ejemplo* de dos factores con interacción:
  - ▷ Añadir *azúcar* al café y *remover* el café (factores A y B).
  - ▷ Ninguno de estos dos factores, por *separado,* tiene mucho efecto en el sabor más o menos dulce del café, sin embargo la *combinación* de los dos sí.
- Continuamos con el ejemplo de la fabricación de *circuitos integrados.*
- Estimamos el efecto del factor B con 'A en su nivel alto':

$$\Delta_{(b \mid a)} = ab - a.$$

- Estimamos el efecto del factor B con 'A en su nivel bajo':

$$\Delta_{(b \mid \text{no } a)} = b - (1).$$

Objetivo: Estimar la *interacción* entre los factores A (tiempo de aplicación) y B (nivel de arsénico).

Operaciones 51:

Solución:

- La *interacción* entre los dos factores es $AB = 0,0315$ μm.
- La *interacción* entre los dos factores ($AB = 0,0315$ μm), es relativamente baja, pues es 26 veces menor que el efecto del factor A ($A = 0,8360$ μm).

General (Interacción entre dos factores)

- *Interacción* entre dos factores A y B: Existe interacción cuando el efecto del factor B depende del nivel del factor A (y viceversa).
- *Ejemplo* de dos factores con interacción:
  - ▷ *Fumar* e inhalar polvo de *amianto.*
  - ▷ Los dos factores, *por separado,* aumentan el riesgo de padecer cáncer de pulmón.
  - ▷ Sin embargo, el *efecto conjunto* de fumar e inhalar polvo de amianto es mayor que la suma de los dos efectos.
- La *interacción* entre los factores A y B se puede estimar mediante la siguiente fórmula:

$$\begin{aligned} AB &= \frac{1}{n}\left[\frac{\Delta_{(b \mid a)} \quad - \quad \Delta_{(b \mid \text{no } a)}}{2}\right] \qquad (7.1) \\ &= \frac{1}{n}\left[\frac{ab - a \quad - \quad (b-(1))}{2}\right] \\ &= \frac{1}{2n}\left[ab + (1) - a - b\right]. \end{aligned}$$

  donde n es el tamaño muestral, $\Delta_{(b \mid a)}$ es el efecto de B con A en su nivel alto y $\Delta_{(b \mid \text{no } a)}$ es el efecto de B con A en su nivel bajo.

- Notar que si $\Delta_{(b \mid a)} = \Delta_{(b \mid \text{no } a)}$ entonces $AB = 0$.
- En (7.1) dividimos por 2 para que la magnitud de la interacción sea coherente con la magnitud de los efectos principales A y B, donde también dividimos por 2.

- Al hablar de la *interacción* entre dos factores, nos estaremos refiriendo a la *interacción promedio,* pues los cálculos se realizan con valores promedio (repetimos n veces el experimento).

- *Interpretación de AB:* En promedio, pasando el factor A del nivel bajo al nivel alto, cambia el efecto del factor B en $2AB$ unidades. Es decir:

$$\frac{\Delta_{(b \mid a)}}{n} - \frac{\Delta_{(b \mid \text{no } a)}}{n} = 2AB.$$

  Esto se deduce fácilmente de la Ecuación (7.1).

- *Ejemplo anterior:* Cambiando el tiempo de aplicación de corto a largo (factor A), cambia el efecto del nivel de arsénico (factor B) en $2AB = 2 \cdot 0,0315 = 0,0630\mu$m.
  El siguiente cálculo no es imprescindible (sólo para verificar):

$$\begin{aligned}\frac{\Delta_{(b \mid a)}}{n} - \frac{\Delta_{(b \mid \text{no } a)}}{n} &= \frac{59,156 - 59,299}{4} - \frac{55,686 - 56,081}{4} \\ &= -0,03575 - (-0,09875) \\ &= 0,0630.\end{aligned}$$

- *Relevancia:*

  ▷ El nivel de interacción es un parámetro *importante* a tener en cuenta.

  ▷ Cuando la interacción AB es *grande,* los correspondientes efectos individuales A y B, tienen muy poco significado práctico.

- Se verifica fácilmente que $AB = BA$, pues

$$\begin{aligned}BA &= \frac{1}{n}\left[\frac{\Delta_{(a \mid b)} - \Delta_{(a \mid \text{no } b)}}{2}\right] \\ &= \frac{1}{n}\left[\frac{ab - b - (a - (1))}{2}\right] \\ &= \frac{1}{2n}\left[ab + (1) - a - b\right] = AB.\end{aligned}$$

General (Experimentos factoriales 2^k)

- Los experimentos factoriales 2^2 son un *caso particular* de los experimentos factoriales 2^k donde se consideran k factores y 2 niveles por factor.

- En el caso de tener sólo dos niveles el diseño del experimento y su análisis es básicamente el mismo tanto para niveles *cuantitativos* como para niveles *cualitativos.*

- En un experimento factorial 2^k en cada realización del experimento se analizan las $C = 2^k$ posibles *combinaciones* de los 2 niveles de cada factor.

- El diseño 2^k es particularmente útil al *inicio* del trabajo experimental cuando habiendo muchos posibles factores conviene descartar los que no son relevantes.

- Si se consideran $r > 2$ *niveles* por factor, entonces tenemos los experimentos factoriales r^k.

7.3. Experimentos factoriales y regresión

Veremos los siguientes apartados:

- *Modelo* de regresión (no lineal).
- *Inferencia* sobre el modelo de regresión.

7.3.1. Modelo de regresión (no lineal)

Ejemplo 52 (Fabricación de circuitos integrados - cont.)

Datos:

- Continuamos con el ejemplo anterior con la finalidad de asociarle un modelo de regresión.
- Consideramos el siguiente *modelo de regresión* (no lineal)

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon.$$

- En este modelo:
  - ▷ Y es la *variable respuesta* o dependiente y las variables x_j son las *variables explicativas* o independientes.
  - ▷ Y representa el *grosor* de la capa uniforme de material semiconductor del circuito integrado.
  - ▷ x_1 representa el *tiempo de aplicación* del vapor rico en arsénico (factor A).
  - ▷ x_2 representa el *nivel de arsénico* en el vapor aplicado (factor B).
  - ▷ En este contexto los niveles alto (+) y bajo ($-$) de cada factor se *codifican* mediante los valores $x_i = +1$ y $x_i = -1$ para $i = 1, 2$.
  - ▷ $x_1 x_2$ representa la *interacción* entre los dos factores A y B.
- La correspondiente *función de regresión* tiene la siguiente expresión:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_1 x_2.$$

Objetivo:

1) Calcular los *coeficientes* de la función de regresión.
2) Calcular el valor de $\hat{y}$ para $(x_1, x_2) = (1,\ 1)$ e *interpretar* el resultado con relación al experimento factorial del ejemplo anterior donde las posibles combinaciones eran $(1), a, b$ y ab.

Operaciones 52:

Solución:

1) La *función de regresión* tiene la siguiente expresión:

$$\hat{y}(x) = 14,389 + 0,418\, x_1 - 0,034\, x_2 + 0,016\, x_1 x_2.$$

2) $\hat{y}(x)$ es una estimación del *grosor promedio* asociado a los factores codificados en x. Así:
   - $\hat{y}(1,\ 1)$ es una estimación del grosor promedio correspondiente a la combinación ab,
   - $\hat{y}(1, -1)$ es una estimación del grosor promedio correspondiente a la combinación a, etc.

General (Modelo de regresión no lineal)

- Asociado a todo experimento factorial 2^2 tenemos un *modelo de regresión* (no lineal):

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon.$$

donde

▷ Y es la variable respuesta o dependiente y las variables x_j son las variables explicativas o independientes.

▷ Y representa la *magnitud* bajo estudio.

▷ x_1 representa el *factor A*.

▷ x_2 representa el *factor B*.

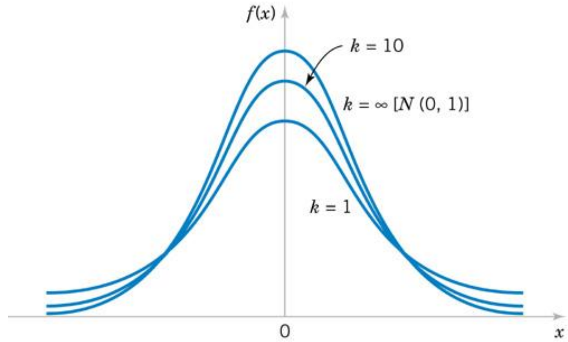


Figura 7.1: La función de densidad de una variable aleatoria t de Student con k grados de libertad es *simétrica.*

- ▷ x_1x_2 representa la *interacción* entre los dos factores A y B.
- ▷ ϵ es el *error* aleatorio cuya distribución se supone $N(0, \sigma)$.

- En este contexto los niveles alto (+) y bajo ($-$) de cada factor se *codifican* mediante los valores $x_i = +1$ y $x_i = -1$ para $i = 1, 2$.
- Por lo tanto tenemos $C = 2^2 = 4$ combinaciones de niveles: $(1), a, b, ab$.
- La correspondiente *función de regresión* tiene la siguiente expresión

$$\hat{y}(x) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_1 x_2,$$

  donde $\hat{y}(x)$ es una estimación del valor esperado de Y en función de x.

- Se puede demostrar que en este contexto los coeficientes $\hat{\beta}$ se pueden calcular como sigue:
  - ▷ $\hat{\beta}_0$ es la *gran media* de las C medias (tenemos $C = 4$ combinaciones de niveles):

$$\hat{\beta}_0 = \bar{\bar{y}} = \frac{1}{C}\sum_{c=1}^{C} \bar{y}_c.$$

  - ▷ El resto de los *coeficientes* de regresión corresponden a los efectos de los factores y a la interacción, divididos entre 2:

$$\begin{aligned}
\hat{\beta}_1 &= \frac{A}{2} = \frac{1}{4n}\left[a + ab - b - (1)\right] \\
\hat{\beta}_2 &= \frac{B}{2} = \frac{1}{4n}\left[b + ab - a - (1)\right] \\
\hat{\beta}_3 &= \frac{AB}{2} = \frac{1}{4n}\left[ab + (1) - a - b\right].
\end{aligned}$$

7.3.2. Inferencia sobre el modelo de regresión

General (Distribución t de Student)

- $X \sim T_k$, se lee, X tiene una distribución t de Student con k *grados de libertad.*

- En la Fig. 7.1 tenemos representada la función de *densidad de probabilidad* de una variable aleatoria t de Student para varios valores de k.
- La gráfica de la función de densidad de probabilidad (fdp) de la distribución t es parecida a la gráfica de la fdp de la distribución normal, pero con *'colas más pesadas'*.
- $T_\infty \equiv N(0,1)$.
- *Media* y *varianza:* Si $X \sim T_k$ entonces:

$$\mu = E(X) = 0 \quad \text{y} \quad \sigma^2 = V(X) = \frac{k}{k-2}.$$

- Función de *densidad de probabilidad.* Si $X \sim T_k$ entonces:

$$f(x) = \frac{C_0}{[(x^2/k)+1]^{(k+1)/2}} \qquad -\infty < x < \infty$$

donde

$$C_0 = \frac{1}{\sqrt{\pi k}} \frac{\Gamma[(k+1)/2]}{\Gamma[k/2]}$$

y

$$\Gamma(k) = \int_0^\infty e^{-z} z^{k-1} dz.$$

- Para hacer los cálculos asociados a esta distribución normalmente usaremos una *tabla* (tabla del Apéndice.2)

Ejemplo 53 (Cálculos con la distribución t de Student)

Datos: Consideramos T, una variable aleatoria t de Student con 12 grados de libertad.

Objetivo: Calcula:

1) $P(T \leq 3,055)$.
2) $P(T > 3,055)$.

Operaciones 53:

Solución:

1) $P\big(T \leq 3,055\big) \;=\; 0,995.$
2) $P\big(T > 3,055\big) \;=\; 0,005.$

Ejemplo 54 (Cálculos con la distribución t de Student)

Datos: Consideramos T, una variable aleatoria t de Student con 12 grados de libertad.

Objetivo: Aproxima $P\big(T > t_0\big)$ para los siguientes casos:

1) $t_0 = 399$.
2) $t_0 = 11,60$.
3) $t_0 = 0,93$.
4) $t_0 = 0,44$.

Operaciones 54:

Solución:

1) $P\big(T > 399\big) \quad \approx \quad 0.$
2) $P\big(T > 11,60\big) \approx \quad 0.$
3) $P\big(T > 0,93\big) \quad > \quad 0{,}15.$
4) $P\big(T > 0,44\big) \quad > \quad 0{,}25.$

Ejemplo 55 (Fabricación de circuitos integrados - cont.)

Datos: En el primer ejemplo de la Sección 7.3.1 hemos visto que podemos estimar el promedio del *grosor* Y de la capa uniforme en el circuito integrado mediante la siguiente función no lineal:

$$\hat{y}(x) \quad = \quad 14,389 \; + 0,418\, x_1 - 0,0335\, x_2 + 0,016\, x_1 x_2,$$

donde

- x_1 representa el *factor A* $(x_1 \in \{-1, +1\})$.
- x_2 representa el *factor B* $(x_2 \in \{-1, +1\})$.

Recordamos que en este contexto:

$C = 4$	Número total de *combinaciones* de niveles.
$n = 4$	Número total de *repeticiones* del experimento.
	(número de muestras para cada combinación de niveles).

Objetivo: Analizar si en la función $\hat{y}(x)$ todos los coeficientes son *relevantes* (distintos de 0) o por el contrario podemos prescindir de alguno de ellos (tomar $\alpha = 0{,}05$).

Operaciones 55:

Coef.	Estimación	Error estándar	t_0	P-valor	$H_1 : \beta_i \neq 0$
β_0	14,389	0,03605	399	0,00	Aceptamos
β_1	0,4180	0,03605	11,60	0,00	Aceptamos
β_2	-0,0335	0,03605	-0,93	$>$ 0,30	Rechazamos
β_3	0,0160	0,03605	0,44	$>$ 0,50	Rechazamos

Tabla 7.2: *P-valor* de los coeficientes de regresión.

Solución: Aceptamos que los coeficientes distintos de 0 son β_0 y β_1, por lo que podemos *simplificar* nuestro modelo de regresión a

$$\hat{y}(x) \quad = \quad 14,389 \; + 0,418 \; x_1.$$

Notar que en el contexto de los experimentos factoriales no nos hace falta volver a calcular β_0 y β_1 una vez que hemos descartado los otros coeficientes (esto en general no es cierto en todos los modelos de regresión).

General (Relevancia de los coeficientes β)

- Dado el *modelo de regresión* (no lineal)

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon,$$

  nos puede interesar analizar si todos los coeficientes β_i son *relevantes* (distintos de 0) o por el contrario podemos prescindir de alguno de ellos (iguales a 0).

- Podemos abordar esta cuestión mediante un *contraste de hipótesis* por cada coeficiente:

$$\begin{aligned} H_0 : & \quad \beta_i = 0 \\ H_1 : & \quad \beta_i \neq 0 \end{aligned}$$

  para $i \in \{0, 1, 2, 3\}$.

- Para resolver este contraste podemos proceder como sigue:

  1) Calculamos $\hat{\sigma}^2$ (una estimación de la *varianza* de Y):

  $$\begin{aligned}\hat{\sigma}_c^2 &= \frac{1}{n-1}\sum_{j=1}^{n}(y_{cj}-\bar{y}_c)^2 \qquad\qquad c = 1,\cdots,C \\ \hat{\sigma}^2 &= \frac{1}{C}\sum_{c=1}^{C}\hat{\sigma}_c^2,\end{aligned}$$

  donde:

c	es el índice para la *combinación* de niveles de los factores
n	es el número de *repeticiones* del experimento (número de muestras para cada combinación de niveles)
j	es el índice para cada *muestra*
C	Número total de combinaciones de niveles (en este capitulo $C = 2^2 = 4$).

  2) Calculamos una estimación del *error estándar* (desviación típica) del coeficiente $\hat{\beta}_i$ (el mismo error estándar para todos los coeficientes):

  $$se(\text{coeficiente}) = \frac{1}{2}\sqrt{\frac{\hat{\sigma}^2}{n}}$$

  donde se son las siglas de 'Standard Error'.

  3) Calculamos el estadístico

  $$t_0 = \frac{\text{coeficiente}}{se(\text{coeficiente})}.$$

  4) Se calcula el *P-valor* asociado a t_0 mediante la siguiente fórmula:

  $$\begin{aligned}\text{P-valor } &= P\big(T_0 < -\mid t_0 \mid\big) + P\big(T_0 > \mid t_0 \mid\big) \\ &= 2\cdot P\big(T_0 > \mid t_0 \mid\big)\end{aligned}$$

  donde T_0 es una variable aleatoria t de Student con $C(n-1)$ grados de libertad.

  5) Para realizar el anterior cálculo se consulta la tabla de la distribución t de Student (ver el Apéndice al final de estos apuntes).

  6) Si el P-valor es menor o igual que el *nivel de significación* α, aceptamos H_1.

  7) En este caso, tenemos evidencia estadística de que el correspondiente coeficiente es distinto de 0 y por tanto es relevante (valores usuales para α son 0,01, 0,05 y 0,10).

  8) Si el P-valor es mayor que α, rechazamos H_1, y consideramos que el correspondiente coeficiente no es relevante (asumimos que es 0).

- Recordamos que:

  ▷ El *P-valor* corresponde a la probabilidad de equivocarnos al aceptar H_1.

  ▷ El *nivel de significación* α es una cota arbitraria a la probabilidad de equivocarnos al aceptar H_1.

  ▷ *Regla de decisión:* Si el P-valor no supera α, aceptamos H_1. En caso contrario, rechazamos H_1.

Capítulo 8

Apéndice

En este apéndice tenemos las siguientes tablas:

- Tabla Apéndice.1. Distribución Normal estándar:

$$\Phi(z) = P(Z \leq z).$$

- Tabla Apéndice.2. Distribución t de Student: La tabla da el valor de c, tal que

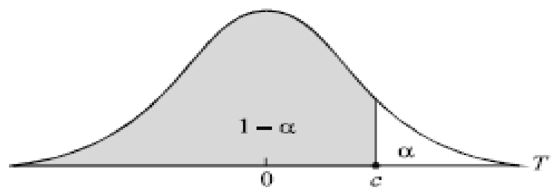
$$P(T \leq c) = 1 - \alpha.$$

- Tabla Apéndice.3. Factores para construir tablas de control $\bar{X}$ y R.

The Cumulative Distribution Function for the
Standard Normal Distribution: Values of $\Phi(z)$ for nonnegative z

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

TABLA DE LA DISTRIBUCION *t*-Student con *n* grados de libertad..



	$1-\alpha$							
n	0.75	0.80	0.85	0.90	0.95	0.975	0.99	0.995
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704
60	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660
120	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617
¥	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576

	$\overline{X}$ Chart			R Chart		
	Factors for Control Limits			Factors for Control Limits		
n^a	A_1	A_2	d_2	D_3	D_4	n
2	3.760	1.880	1.128	0	3.267	2
3	2.394	1.023	1.693	0	2.575	3
4	1.880	.729	2.059	0	2.282	4
5	1.596	.577	2.326	0	2.115	5
6	1.410	.483	2.534	0	2.004	6
7	1.277	.419	2.704	.076	1.924	7
8	1.175	.373	2.847	.136	1.864	8
9	1.094	.337	2.970	.184	1.816	9
10	1.028	.308	3.078	.223	1.777	10
11	.973	.285	3.173	.256	1.744	11
12	.925	.266	3.258	.284	1.716	12
13	.884	.249	3.336	.308	1.692	13
14	.848	.235	3.407	.329	1.671	14
15	.816	.223	3.472	.348	1.652	15
16	.788	.212	3.532	.364	1.636	16
17	.762	.203	3.588	.379	1.621	17
18	.738	.194	3.640	.392	1.608	18
19	.717	.187	3.689	.404	1.596	19
20	.697	.180	3.735	.414	1.586	20
21	.679	.173	3.778	.425	1.575	21
22	.662	.167	3.819	.434	1.566	22
23	.647	.162	3.858	.443	1.557	23
24	.632	.157	3.895	.452	1.548	24
25	.619	.153	3.931	.459	1.541	25

[a] $n > 25$: $A_1 = 3/\sqrt{n}$ where n = number of observations in sample.