

## CAPÍTULO 6

Matemática Numérica, 2da Edición

Manuel Álvarez, Alfredo Guerra, Rogelio Lau

## OPTIMIZACIÓN NUMÉRICA

### Objetivos

Al finalizar el estudio y la ejercitación de este capítulo, el lector debe ser capaz de:

- Describir qué se entiende por optimizar una función.
- Decidir cuándo conviene utilizar un método analítico y cuándo uno numérico para realizar el proceso de optimización.
- Describir los conceptos de función unimodal de una variable y función linealmente unimodal de varias variables.
- Describir los métodos de búsqueda simultánea y búsqueda secuencial, uniforme y acelerada.
- Describir los métodos de búsqueda unidimensional en un intervalo: método de bisección y método de la sección áurea.
- Utilizar en forma combinada los métodos de búsqueda abierta y en un intervalo para optimizar funciones de una variable.
- Describir los métodos de optimización para funciones de varias variables: Búsqueda por coordenadas, gradiente, Powell y simplex secuencial, estableciendo las ventajas y desventajas de cada uno de ellos.
- Explicar el algoritmo en pseudo código de cada uno de los métodos de optimización estudiados e ilustrarlos con ejemplos manuales.
- Utilizar los métodos de optimización multidimensional para hallar máximos y mínimos de funciones de varias variables independientes.
- Modelar problemas que conducen a problemas de optimización de una variable con o sin restricciones y problemas de varias variables sin restricciones.

### 6.1 Introducción

#### *Problemas de optimización*

La palabra “optimización” se utiliza en la Matemática para indicar aquellos procedimientos que permiten encontrar valores de las variables independientes para los cuales una cierta función toma su mayor (o menor) valor. En ocasiones el conjunto donde están definidas las variables independientes está restringido en una cierta región del dominio de definición.

Una enorme cantidad de problemas prácticos son problemas de optimización o pueden ser expresados como un problema de optimización, de los cuales existe una gran variedad de acuerdo con las siguientes posibilidades:

- Las características de la función a optimizar, por ejemplo, si es una función lineal o si es cuadrática o si es unimodal o si, por el contrario, no exhibe ninguna de estas propiedades.
- Si se conoce la expresión analítica de la función o si solamente puede ser evaluada de modo experimental.
- El número de variables independientes, que va desde una hasta cientos de variables, sobre todo en problemas relacionados con grandes sistemas económicos.

- Si existen restricciones en los valores de las variables independientes o, por el contrario, se trata de un problema de variables libres.
- En caso de que existan restricciones, el tipo de ellas resulta del mayor interés. Por ejemplo, si todas las restricciones son inecuaciones lineales, existen técnicas especiales para el problema.
- Si las variables solo pueden tomar valores enteros, las técnicas de solución son muy diferentes de las que se emplean cuando se trata de variables reales.

En cursos anteriores se estudiaron métodos analíticos para hallar los extremos de funciones de una y de más variables. Todos estos métodos partían de las siguientes hipótesis:

- Se conoce la expresión analítica de la función que se desea optimizar.
- La función a optimizar es continua y derivable una o más veces.
- Las ecuaciones o sistemas de ecuaciones (en general, no lineales) que se obtiene al igualar las derivadas a cero pueden ser resueltas.

A diferencia de esos casos, aquí se estudiarán técnicas numéricas, las cuales descansarán, fundamentalmente, en la operación de evaluar la función a optimizar.

Primero se analizará el caso más simple de las funciones que dependen de una sola variable independiente. Los procedimientos más generales, para funciones de  $n$  variables, casi siempre consisten en resolver secuencias de problemas unidimensionales; de aquí la importancia del caso unidimensional. Por otra parte, algunos problemas prácticos son a veces unidimensionales. A continuación se muestran varios ejemplos.

### **Ejemplo 1**

En una empresa de producción agrícola desean estimar la cantidad óptima de fertilizante por hectárea que debe emplearse para cierto cultivo. Si la eficiencia en dicho cultivo se mide por  $P$ : relación entre el valor de la producción obtenida y el valor de los recursos gastados, es bastante evidente que  $P$  es una función de la cantidad  $x$  de fertilizante por hectárea. Se supone que otros importantes parámetros no han de cambiar.

Como se ve, se trata de un problema de optimización donde la función  $P(x)$  depende de una sola variable real  $x \geq 0$ . La expresión analítica de  $P(x)$  se desconoce.

### **Ejemplo 2**

La función Gamma se define para  $x > 0$  mediante la integral de Euler:

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt$$

En el intervalo  $[1; 2]$  esta función posee un punto de mínimo pero hallarlo mediante un método analítico no es posible por la dificultad de hallar la derivada respecto a  $x$ .

### **Ejemplo 3**

Se necesita resolver el siguiente sistema de ecuaciones para el diseño de un circuito:

$$\begin{cases} a_1 V_1(T_1) - a_2 V_2(T_1) = V_0 & (1) \\ a_1 V_1(T_2) - a_2 V_2(T_2) = V_0 & (2) \\ a_1 V_1(T_3) - a_2 V_2(T_3) = V_0 & (3) \end{cases}$$

donde  $V_0$  es un voltaje conocido y  $V_1$  y  $V_2$  son ciertos voltajes que dependen de la temperatura  $T$ . A las temperaturas  $T_1, T_2, T_3$  el voltaje  $V_2$  se conoce, pero no  $V_1$ . Tampoco se conocen los coeficientes  $a_1$  y  $a_2$ . Se sabe además que el voltaje  $V_1$  depende implícitamente de la temperatura según la ecuación:

$$\alpha T \left( 1 + \frac{V_1(T)}{V_{ar}} \right) = K T^\eta \exp \left( \frac{V_1(T) - V_{go}}{T^{\frac{k}{q}}} \right) \quad (4)$$

donde  $K, V_{ar}, V_{go}, \eta, k$  y  $q$  son constantes conocidas. Sin embargo la constante  $\alpha$  se desconoce. Se trata pues de un sistema de seis ecuaciones con seis incógnitas.

Ecuaciones: (1), (2), (3) y (4) para  $T_1, T_2$  y  $T_3$ .

Incógnitas:  $V_1(T_1), V_1(T_2), V_1(T_3), a_1, a_2$  y  $\alpha$ .

Aunque es un sistema no lineal de ecuaciones, puede plantearse como un problema de optimización unidimensional. Para ello, se define la función  $f(\alpha)$  de la siguiente forma:

Dado un valor de  $\alpha$ :

1. Se hace  $T = T_1$  en la ecuación (4) y se halla numéricamente su única incógnita  $V_1(T_1)$  (se puede hacer, por ejemplo, mediante el método iterativo simple, despejando la incógnita  $V_1(T_1)$  que aparece en el argumento de la exponencial)
2. Lo mismo para  $T = T_2$ . Se halla  $V_1(T_2)$ .
3. Lo mismo para  $T = T_3$ . Se halla  $V_1(T_3)$ .
4. Se resuelve el sistema formado por las ecuaciones (2) y (3) para hallar  $a_1$  y  $a_2$  (una vez que se han determinado  $V_1(T_2)$  y  $V_1(T_3)$ , este sistema es lineal en  $a_1$  y  $a_2$ ).
5. Se define la función:  $f(\alpha) = [a_1 V_1(T_1) - a_2 V_2(T_1) - V_0]^2$

El valor mínimo de  $f(\alpha)$  es cero y se alcanza para un  $\alpha$  que satisface la ecuación (1). O sea, el valor de  $\alpha$  que minimiza a  $f$  conduce, mediante los pasos 1, 2, 3, 4 y 5, a la obtención de la solución del sistema de ecuaciones. Obsérvese, sin embargo, que a pesar de que existen expresiones analíticas que definen sin ambigüedades la función  $f(\alpha)$ , el tratamiento analítico de dicha función sería muy complicado.

### ***Funciones unimodales de una variable***

De manera no muy formal, puede decirse que una función unimodal de una variable es aquella que posee en su conjunto de definición un solo punto de extremo relativo. En la figura 1 se observan dos funciones a) y b) unimodales y una c) que no lo es.

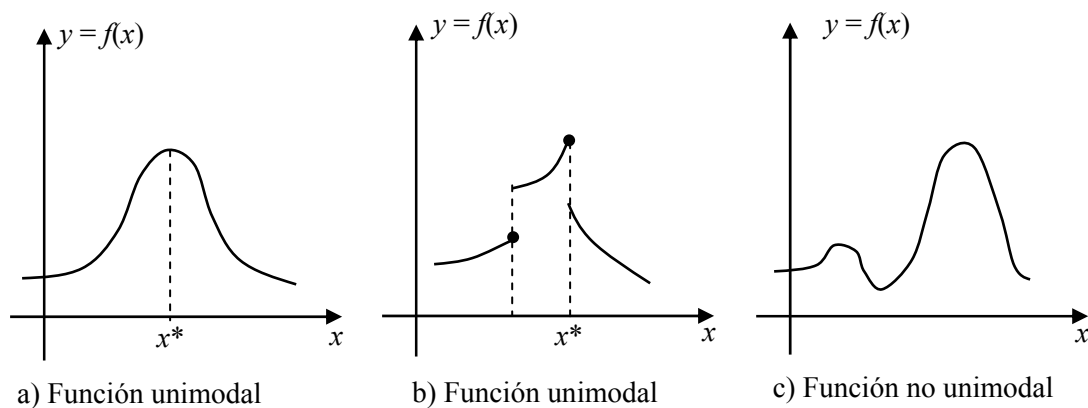


Figura 1

Una definición un poco más precisa es la siguiente:

Definición:

La función  $f(x)$  (vea la figura 2) es unimodal con máximo si existe en su dominio un  $x^*$  tal que, si  $x_1$  y  $x_2$  pertenecen al dominio se cumple que:

$$\text{Si } x_1 < x_2 < x^* \text{ entonces } f(x_1) < f(x_2)$$

$$\text{Si } x^* < x_1 < x_2 \text{ entonces } f(x_1) > f(x_2)$$

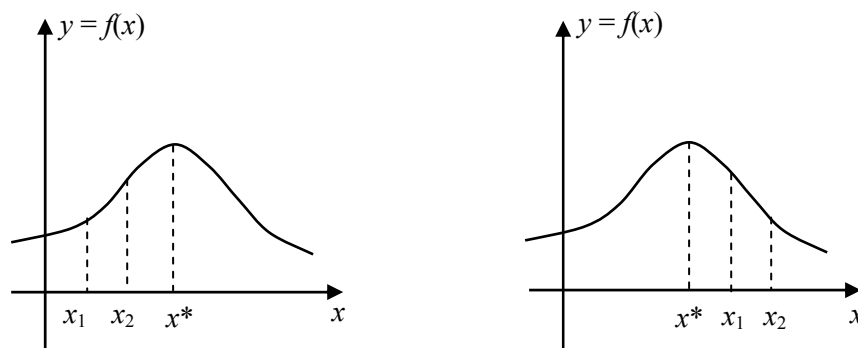


Figura 2

Nótese que la definición anterior simplemente establece que  $f$  es creciente a la izquierda de  $x^*$  y decreciente a la derecha de  $x^*$ . De manera evidente se puede modificar la definición para el caso de una función unimodal con mínimo. Sin embargo, no es necesario estudiar por separado los casos de máximo y de mínimo porque todo problema de minimización se puede reducir a uno de maximización (y viceversa). Por ejemplo, basta tener en cuenta que si  $f(x)$  es unimodal con punto de máximo en  $x^*$  entonces  $-f(x)$  es unimodal con punto de mínimo en  $x^*$ .

En todos los métodos que se estudiarán se aplica el siguiente resultado el cual, por ser tan elemental, no se ha querido llamarlo teorema.

### Propiedad básica de la optimización unidimensional

Sea  $f(x)$  una función unimodal con máximo en  $x^*$  y sean  $x_1$  y  $x_2$  dos valores de su dominio. Sean  $y_1 = f(x_1)$  y  $y_2 = f(x_2)$ . Entonces (vea la figura 3):

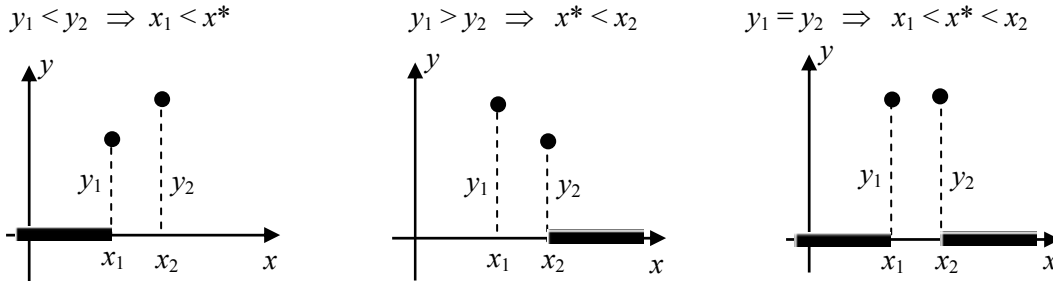


Figura 3

En la figura 3 se ha marcado con líneas gruesas la porción del dominio de  $f(x)$  que puede ser desechada, ya que con seguridad el punto de óptimo no se encuentra en ella.

La justificación es bastante simple. Por ejemplo, en el primer caso, si se admitiera que el punto de máximo  $x^*$  está a la izquierda de  $x_1$  entonces en el intervalo  $[x_1, x_2]$   $f(x)$  tendría que ser decreciente y en ese caso debería ser  $y_1 > y_2$ ; como esto contradice la hipótesis, entonces  $x^*$  no está a la izquierda sino a la derecha de  $x_1$ . De modo similar se prueba para los otros dos casos. Se deja al lector como ejercicio que enuncie y demuestre la propiedad básica para una función unimodal con mínimo.

### Clasificación de los métodos de búsqueda unidimensional

En la optimización numérica se acostumbra llamar experimento al proceso mediante el cual se conoce el valor de la función  $f$  para un determinado valor de  $x$ . Está claro que en muchas ocasiones la evaluación no necesita realizarse experimentalmente, sin embargo, esta forma de hablar es clásica. No obstante, aun cuando los experimentos se realicen evaluando una expresión algebraica en un programa de computadora o mediante un largo proceso práctico, el costo del algoritmo de optimización está directamente relacionado a la cantidad de experimentos que sea necesario realizar pues, en todo caso, los experimentos son la parte del algoritmo en que más recursos se consume.

Una primera clasificación entre los métodos de optimización numérica para funciones de una variable se basa en la simultaneidad o no de los experimentos. Aquellos algoritmos en los que los experimentos se realizan todos a un tiempo, se llaman de *búsqueda simultánea* mientras que los que siguen la estrategia de realizar un experimento después que han sido analizados los resultados de los experimentos anteriores, se llaman *métodos secuenciales*. Los algoritmos secuenciales requieren menos experimentos que los simultáneos, ya que en estos se realizan evaluaciones sin analizar resultados anteriores que podrían haber descartado la necesidad de muchos de estos experimentos. Sin embargo, en muchas ocasiones la naturaleza del problema investigado obliga a utilizar los métodos simultáneos. Por ejemplo, para determinar el punto de un territorio donde la temperatura fue mínima anoche, no queda más alternativa que medir la temperatura simultáneamente en una buena muestra de puntos; para determinar la cantidad óptima de fertilizante a utilizar en un cultivo sería demasiado lento el procedimiento secuencial, pues en este caso cada experimento requiere meses.

En la búsqueda secuencial hay dos situaciones bastante diferentes que requieren también procedimientos diferentes: los casos en que el punto de máximo  $x^*$  se halla dentro de un intervalo conocido (*búsqueda con restricciones*) y los casos en que no se sabe nada acerca de  $x^*$  (*búsqueda sin restricciones*). Es usual que ambos tipos de problemas se combinen y que se comience utilizando un método de búsqueda sin restricciones que termina dando un intervalo donde se halla  $x^*$  y, a continuación, aplicar procedimientos de búsqueda con restricciones para hallar  $x^*$  con una precisión aceptable.

En general, la búsqueda simultánea no tiene mucho que discutir y se le dedicará muy poca atención. La búsqueda secuencial es más rica en resultados y se estudiará con algún detalle posteriormente.

## 6.2 Optimización unidimensional sin restricciones

### *Búsqueda simultánea*

Como ya se dijo, a menos que sea imprescindible, debe evitarse utilizar la búsqueda simultánea. En general los experimentos se sitúan uniformemente espaciados aunque el conocimiento de las características de la función a optimizar podría aconsejar concentrar más experimentos en determinadas regiones donde es más probable la existencia del punto  $x^*$ .

Si los experimentos se realizan en los valores  $x_1 < x_2 < \dots < x_n$  con resultados  $y_1, y_2, \dots, y_n$  y se obtiene  $y_k = \max \{y_i\}$  con  $1 < k < n$ , entonces (vea la figura 1), de acuerdo con la propiedad básica de las funciones unimodales, el punto de máximo se encuentra entre  $x_{k-1}$  y  $x_{k+1}$ , o sea:

$$x_{k-1} < x^* < x_{k+1}$$

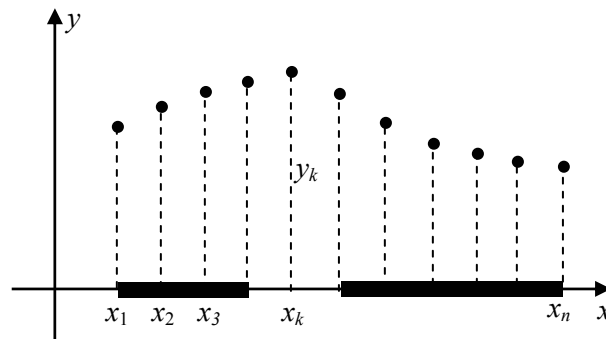


Figura 1

La cantidad de experimentos simultáneos depende del costo de cada experimento y de la exactitud con que se necesite determinar  $x^*$ .

### *Búsqueda secuencial uniforme*

Sea  $y = f(x)$  una función unimodal con un punto de máximo en  $x^*$ , de la cual solo se supone que está definida en un intervalo que incluye a  $x^*$  y que puede ser evaluada en cualquier punto de dicho intervalo. El método más simple para hallar  $x^*$  es la búsqueda secuencial uniforme.

Sea  $x_0$  el punto donde se comenzará la búsqueda. La selección de  $x_0$  requiere, por lo general, algún conocimiento previo de la función que se desea optimizar y a veces algún tipo de “sospecha” basada en análisis realizados a la misma o en la naturaleza del fenómeno con que se está tratando. En cualquier caso, como es natural, se seleccionará  $x_0$  lo más próximo posible de  $x^*$ .

Sea  $s$  un número real distinto de cero (puede ser positivo o negativo) al que se llamará *paso*. El método de búsqueda secuencial uniforme consiste, simplemente, en generar la sucesión de valores:

$$x_i = x_0 + is \quad i = 0, 1, 2, 3, \dots$$

y obtener para cada uno de ellos la imagen de  $f$ , esto es:

$$y_i = f(x_i)$$

El proceso se detiene tan pronto se obtiene un  $y_k$  tal que  $y_{k-1} > y_k$ , es decir:

$$y_0 < y_1 < y_2 < \dots < y_{k-1} > y_k$$

y puede entonces asegurarse, de acuerdo con la propiedad básica, que  $x^*$  se encuentra entre  $x_{k-2}$  y  $x_k$ . Teniendo en cuenta el signo de  $s$ , se puede escribir:

$$\text{si } s > 0, \quad x_{k-2} < x^* < x_k$$

$$\text{si } s < 0, \quad x_k < x^* < x_{k-2}$$

En la figura 2 se muestra geoméricamente la situación en el caso en que  $s > 0$ :

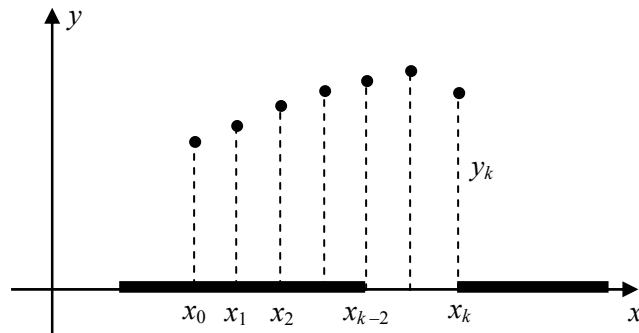


Figura 2

### ***Algoritmo en pseudo código***

El algoritmo que sigue permite determinar un intervalo en el que se encuentra el punto  $x^*$  de máximo de una función  $f(x)$  unimodal. Se supone que la función está definida en todo el intervalo en que se realiza la búsqueda. El algoritmo utiliza como datos: la función  $f(x)$ , el número  $x_0$  a partir del cual se comienza la búsqueda y el paso  $s$ , que puede ser positivo o negativo.

```

 $k := 0$ 
 $y_0 := f(x_0)$ 
repeat
     $k := k + 1$ 
     $x_k := x_{k-1} + s$ 
     $y_k := f(x_k)$ 
until  $y_k < y_{k-1}$ 
 $x^*$  se encuentra entre  $x_{k-2}$  y  $x_k$ 
Terminar

```

Es de notar que, en cualquier caso, la región final en que queda acotado  $x^*$  tiene como amplitud  $2|s|$ ; para obtener  $x^*$  con una exactitud adecuada se debe tomar un paso  $s$  pequeño pero, por otra parte, si  $s$  es pequeño se requiere de muchos experimentos para cubrir la distancia  $|x^* - x_0|$ . Hay varias formas de resolver esta contradicción, que se analizan a continuación.

- Búsqueda secuencial por etapas

Esta estrategia consiste en tomar un paso  $s$  inicial grande (del orden de magnitud de la distancia  $|x^* - x_0|$ ); una vez determinado el intervalo  $x_{k-2} < x^* < x_k$  (aquí y en lo que sigue se está suponiendo  $s > 0$ ; en el caso  $s < 0$ , los cambios a realizar son muy simples) tomar  $x_0 = x_{k-2}$  y comenzar una nueva búsqueda con un paso  $s$  menor. Puede probarse que el valor esperado del número total de experimentos para encontrar el óptimo con una cierta precisión dada, se hace mínimo cuando, en cada etapa, el paso se reduce en  $e$  veces ( $e = 2.71828\dots$ ). Se le propone al lector aventajado que lo demuestre.

- Búsqueda en un intervalo

En este caso, una vez determinado el intervalo, se aplican los métodos de búsqueda con restricciones, que se verán más adelante y que son más eficientes.

- Búsqueda secuencial acelerada

Esta estrategia resulta apropiada cuando no se tiene una idea clara del tamaño del problema, o sea, de la distancia  $|x^* - x_0|$ . Consiste en lo siguiente: En cada paso del algoritmo mientras no se cumpla la condición de parada, el paso se duplica en valor; de esta manera, aun cuando inicialmente  $s$  fuera demasiado pequeño, pronto toma valores suficientemente grandes para alcanzar a  $x^*$  en no muchas iteraciones. A este algoritmo se le llama *búsqueda secuencial acelerada* (en la figura 3 se ve geoméricamente la idea). A continuación se muestra el algoritmo, que solo posee una ligera modificación respecto al anterior:

```

 $k := 0$ 
 $y_0 := f(x_0)$ 
repeat
     $k := k + 1$ 
     $x_k := x_{k-1} + s$ 
     $y_k := f(x_k)$ 
     $s := 2s$ 
until  $y_k < y_{k-1}$ 
 $x^*$  se encuentra entre  $x_{k-2}$  y  $x_k$ 
Terminar

```

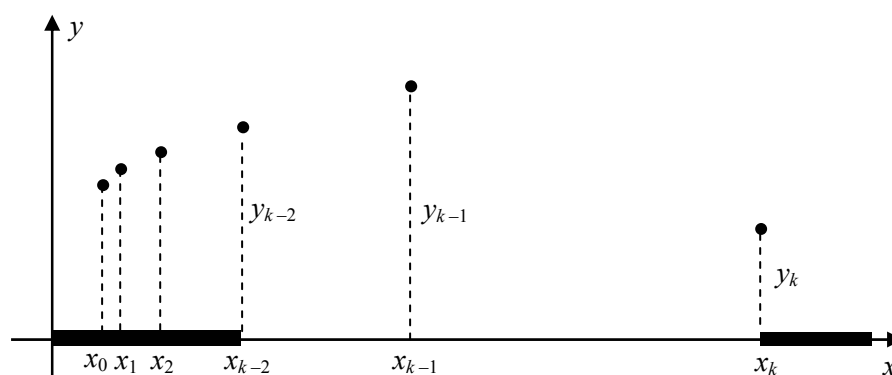


Figura 3

También se aprecia cómo, por lo general, el intervalo final en que queda encerrado  $x^*$  es apreciablemente grande; esto, sin embargo, puede resolverse posteriormente utilizando alguno de los métodos de búsqueda en un intervalo, los cuales son más eficientes.

### Ejemplo 1

Halle el punto de máximo de la función  $f(x) = \sin x$  tomando  $x_0 = 0$  y  $s = 0,1$  mediante búsqueda secuencial a) uniforme, b) acelerada.

Solución:

a) En la tabla 1 se muestran parcialmente los experimentos efectuados. Se obtiene

$$1,5 < x^* < 1,7$$

(el valor de  $x^*$  es  $\pi/2 = 1,5708\dots$ ) y se requieren 17 iteraciones.

| $i$      | $x_i$    | $f(x_i)$ |
|----------|----------|----------|
| 0        | 0,0      | 0,0000   |
| 1        | 0,1      | 0,0998   |
| 2        | 0,2      | 0,1987   |
| $\vdots$ | $\vdots$ | $\vdots$ |
| 14       | 1,4      | 0,9854   |
| 15       | 1,5      | 0,9975   |
| 16       | 1,6      | 0,9996   |
| 17       | 1,7      | 0,9917   |

Tabla 1

b) Utilizando búsqueda secuencial acelerada, en 5 iteraciones se obtiene el resultado:

$$0,7 < x^* < 3,1$$

Los resultados se encuentran en la tabla 2.

| $i$ | $x_i$ | $f(x_i)$ |
|-----|-------|----------|
| 0   | 0,0   | 0,0000   |
| 1   | 0,1   | 0,0998   |
| 2   | 0,3   | 0,2955   |
| 3   | 0,7   | 0,6442   |
| 4   | 1,5   | 0,9975   |
| 5   | 3,1   | 0,0416   |

Tabla 2

Es de destacar el hecho de que el método de búsqueda acelerado llega más rápidamente a la solución pero da como resultado un intervalo final mayor.

### Ejemplo 2

En este ejemplo se ilustra cómo el tamaño inicial del paso  $s_0$  influye en el número de experimentos y en la amplitud del intervalo final  $x_{k-2} < x^* < x_k$ . Considere la función

$$f(x) = 100 - \frac{(x - 1000)^2}{100}$$

que es unimodal y alcanza su valor máximo en  $x^* = 1000$ . A partir de  $x_0 = 0$  utilice el método de búsqueda secuencial acelerada para encontrar el punto de máximo con pasos iniciales de 1, 5, 10, 50, 100 y 900. Determine en cada caso el número de experimentos y la amplitud de la región donde queda acotado  $x^*$ .

Solución:

En la tabla 3 se muestran los resultados obtenidos procediendo de la misma forma que en el ejemplo anterior:

| $s_0$ | Iteraciones | Intervalo final | Amplitud |
|-------|-------------|-----------------|----------|
| 1     | 11          | [511; 2047]     | 1536     |
| 5     | 9           | [635; 2555]     | 1920     |
| 10    | 8           | [630; 2550]     | 1920     |
| 50    | 5           | [350; 1150]     | 800      |
| 100   | 4           | [300; 1500]     | 1200     |
| 900   | 2           | [0; 2700]       | 2700     |

Tabla 3

Como se observa, el hecho de comenzar con un  $x_0$  muy alejado de  $x^*$ , trae como consecuencia que el intervalo final en que  $x^*$  queda encerrado, posee una gran amplitud; sin embargo, la selección del paso inicial no altera en forma significativa esta amplitud. Ahora bien, la cantidad de iteraciones necesarias sí depende fuertemente del valor inicial de  $s$ . Puede probarse que este comportamiento no es casual y que, en general, el valor de  $s$  que minimiza la cantidad de

iteraciones en el algoritmo de búsqueda acelerada es, aproximadamente, igual al *tamaño del problema*, que es la distancia entre  $x_0$  y  $x^*$ .

■

### Selección del sentido de la búsqueda

En algunos casos aparece una dificultad adicional, cuando se desconoce si  $x^*$  es mayor o menor que  $x_0$ . Considérese que este es el caso para la función unimodal con máximo  $f(x)$  y que se toma un paso  $s > 0$ . Si  $f(x_1) > f(x_0)$ , es obvio que puede descartarse la región  $(-\infty, x_0)$ , así que ya se sabe que  $x^* > x_0$ . Sin embargo, si  $f(x_1) < f(x_0)$  la región descartable es  $(x_1, \infty)$ , o sea que  $-\infty < x^* < x_1$ , por tanto, pudiera ser  $x^* > x_0$  o  $x^* < x_0$  (ver figura 4). Observe que el resultado  $f(x_1) < f(x_0)$  puede deberse a que se está buscando  $x^*$  en el sentido equivocado (a) o a que se ha tomado un paso tan grande que se ha sobrepasado  $x^*$  en la primera iteración (b). Hay dos modos de proceder: buscar de nuevo en el mismo sentido con un paso más pequeño a partir de  $x_0$  o comenzar una nueva búsqueda a partir de  $x_1$  pero en el sentido opuesto.

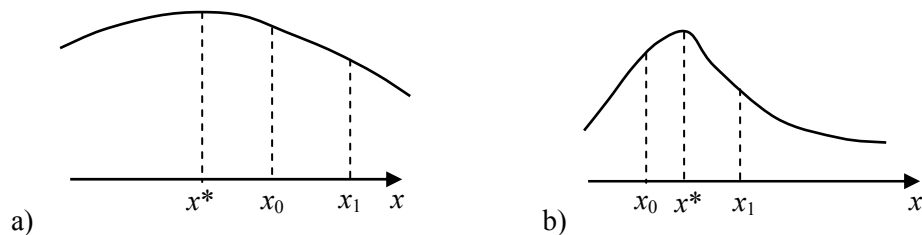


Figura 4

### Ejemplo 3

Halle el punto de máximo de la función  $f(x) = x \sin x$  que se encuentra entre 0 y  $\pi$ . Tome  $x_0 = 2$  y utilice búsqueda secuencia uniforme con paso inicial 0,1. Reduzca el paso a la cuarta parte en cada etapa hasta acotar  $x^*$  en una región de amplitud 0,001.

Solución:

Etapla 1:  $x_0 = 2$   $s = 0,1$

| $x$ | $x \sin x$ |                            |
|-----|------------|----------------------------|
| 2,0 | 1,81859    |                            |
| 2,1 | 1,81273    | Descartado $(2.1, \infty)$ |

Etapla 2:  $x_0 = 2,1$   $s = -0,025$

| $x$   | $x \sin x$ |                       |
|-------|------------|-----------------------|
| 2,1   | 1,81273    |                       |
| 2,075 | 1,81678    |                       |
| 2,050 | 1,81909    |                       |
| 2,025 | 1,81968    |                       |
| 2,000 | 1,81959    | $2,000 < x^* < 2,050$ |

Etapa 3:  $x_0 = 2$   $s = 0,00625$

| $x$     | $x \operatorname{sen} x$ |                        |
|---------|--------------------------|------------------------|
| 2,00000 | 1,81859                  |                        |
| 2,00625 | 1,81902                  |                        |
| 2,01250 | 1,81935                  |                        |
| 2,01875 | 1,81957                  |                        |
| 2,02500 | 1,81968                  |                        |
| 2,03125 | 1,819697                 |                        |
| 2,03750 | 1,819602                 | $2,025 < x^* < 2,0375$ |

Etapa 4:  $x_0 = 2,025$   $s = 0,0015625$

| $x$       | $x \operatorname{sen} x$ |                           |
|-----------|--------------------------|---------------------------|
| 2,025     | 1,819687                 |                           |
| 2,0265625 | 1,819699                 |                           |
| 2,028125  | 1,819705                 |                           |
| 2,0296875 | 1,819704                 | $2,02656 < x^* < 2,02969$ |

Etapa 5:  $x_0 = 2,0265$   $s = 0,0004$

| $x$    | $x \operatorname{sen} x$ |                         |
|--------|--------------------------|-------------------------|
| 2,0265 | 1,8196988                |                         |
| 2,0269 | 1,8197011                |                         |
| 2,0273 | 1,8197029                |                         |
| 2,0277 | 1,8197042                |                         |
| 2,0281 | 1,819705156              |                         |
| 2,0285 | 1,819705651              |                         |
| 2,0289 | 1,819705714              |                         |
| 2,0293 | 1,819705344              | $2,0285 < x^* < 2,0293$ |

### ***Ejercicios***

Algunos de los ejercicios que siguen son muy laboriosos para realizar los cálculos a mano. Se supone que posea programas computacionales, preferiblemente confeccionados por usted, para ayudarle en el trabajo. De no ser así, en los ejercicios que requieran mucho trabajo manual, determine los resultados con menos cifras exactas que las exigidas.

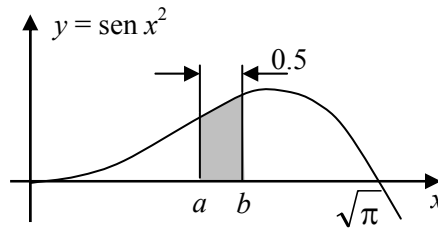
1. Halle el punto de máximo de la función  $f(x) = 2x^2 e^{-3x}$  con un error menor que 0,001. Compruebe su respuesta hallando los ceros de la derivada de  $f(x)$  mediante algún método numérico.

- Halle con tres cifras decimales exactas el punto de mínimo de la función  $f(x) = 3xe^x + e^{-x}$ . Compruebe su respuesta hallando los ceros de la derivada de  $f(x)$  mediante algún método numérico.
- Utilice un asistente matemático, por ejemplo, Derive, para evaluar la función Gamma:

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt \quad x > 0$$

y halle, con error menor que 0,001 el punto de mínimo que posee esta función.

- En el ejemplo 4 de la sección 2.1 fue planteado el problema de hallar las dimensiones de un recipiente cilíndrico de  $1000 \text{ cm}^3$  de capacidad utilizando la mínima cantidad de material y teniendo en cuenta que las piezas que lo conforman deben tener un sobrante de 0,25 cm para poder realizar las soldaduras (vea la figura 4 de ese capítulo). Allí fue resuelto hallando los ceros de la función derivada. Resuélvalo ahora mediante las técnicas estudiadas de optimización numérica con error menor que 0,1 mm.
- Determine los valores de  $a$  y de  $b$ , ambos entre 0 y  $\sqrt{\pi}$ , de manera que el área sombreada de la figura sea máxima. Obtenga la respuesta con dos cifras decimales exactas.



- Elabore un algoritmo en pseudo código que permita resolver el problema anterior.

### 6.3 Optimización en un intervalo

Si en un problema de optimización unidimensional se conoce un intervalo  $[a, b]$  donde se encuentra el punto de máximo  $x^*$  de la función unimodal  $f(x)$ , el problema puede resolverse de modo más efectivo que la búsqueda secuencial. El intervalo  $[a, b]$  puede estar dado desde un inicio a partir de consideraciones físicas, económicas, etc., o puede haberse llegado a él a partir de una etapa previa de búsqueda secuencial. En cualquier caso pueden aplicarse los métodos que siguen.

#### *El método de bisección*

Sea  $f$  una función unimodal definida en  $[a, b]$  y con un punto de máximo  $x^*$  en ese intervalo. Tómense dos puntos experimentales  $x_1$  y  $x_2$  muy próximos entre sí y a ambos lados del centro del intervalo  $[a, b]$ . Para concretar, sea  $\delta$  la distancia entre  $x_1$  y  $x_2$ , entonces, como se aprecia en la figura 1:

$$x_1 = \frac{a+b}{2} - \frac{\delta}{2} \quad \text{y} \quad x_2 = \frac{a+b}{2} + \frac{\delta}{2}$$

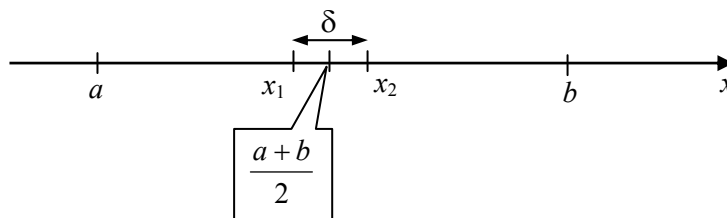


Figura 1

Como se verá de inmediato, debe procurarse que  $\delta$  sea lo menor posible, sin embargo, debe ser razonablemente grande de modo que, al hallar  $f(x_1)$  y  $f(x_2)$  estos valores sean distinguibles uno del otro; esto es aún más importante si  $f(x)$  se evalúa en forma experimental, pues entonces hay que tomar en cuenta los inevitables errores de observación, de medición, etc.

Sean  $y_1 = f(x_1)$  y  $y_2 = f(x_2)$

Como consecuencia inmediata de la propiedad básica de la optimización unidimensional, como muestra la figura 2, se tiene que:

$$y_1 < y_2 \Rightarrow x_1 \leq x^* \leq b$$

$$y_1 > y_2 \Rightarrow a \leq x^* \leq x_2$$

$$y_1 = y_2 \Rightarrow x_1 \leq x^* \leq x_2$$

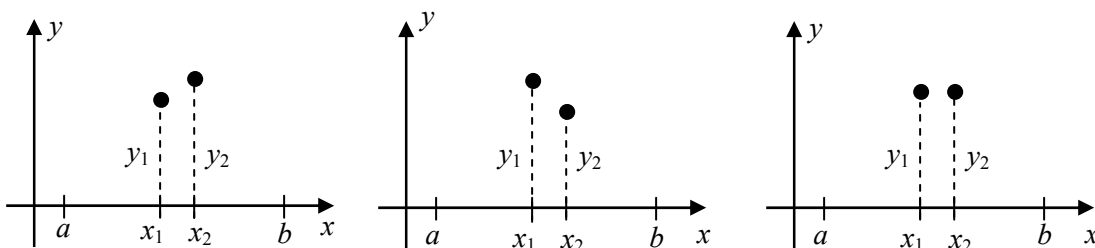


Figura 2

La convergencia del método de bisección es muy fácil de analizar. Si consideramos que  $\delta$  es un número muy pequeño en relación con la amplitud del intervalo, puede suponerse que en cada iteración (que requiere dos evaluaciones de  $f$ ) la longitud del intervalo de búsqueda se reduce a la mitad. Más formalmente, llamando:

$L_n$ : Amplitud del intervalo de búsqueda después de  $n$  evaluaciones ( $n$  par)

$$L_0 = b - a$$

entonces:

$$L_n = \frac{L_0}{2^{n/2}}$$

lo cual prueba que  $L_n$  converge hacia 0 cuando  $n$  tiende hacia infinito. En la práctica, sin embargo,  $L_n$  no puede reducirse a valores menores que  $\delta$ .

La relación  $L_n/L_0$  se utiliza con frecuencia para medir la eficiencia de los métodos de optimización en un intervalo, pues indica cuánto se puede reducir la región de incertidumbre mediante  $n$  experimentos. En el método de bisección esta relación es:

$$\frac{L_n}{L_0} = \frac{1}{2^{n/2}}$$

### **Algoritmo en pseudo código**

El algoritmo que sigue describe el método de optimización de bisección del intervalo. Se supone que la función  $f(x)$  es unimodal con máximo en el intervalo  $[a, b]$  y que está definida en todos los puntos de ese intervalo. El algoritmo da como resultado un intervalo cerrado de amplitud menor que un número especificado  $\varepsilon$  y que contiene al punto de máximo. Los datos que se requiere son: la función  $f(x)$ , el intervalo  $[a, b]$  en que se encuentra inicialmente el punto de máximo de la función, la distancia  $\delta$  que separará a los puntos experimentales y la tolerancia  $\varepsilon$ , que determina la amplitud del intervalo final.

```

repeat
     $x_1 := \frac{a+b}{2} - \frac{\delta}{2}$ 
     $x_2 := \frac{a+b}{2} + \frac{\delta}{2}$ 
     $y_1 := f(x_1)$ 
     $y_2 := f(x_2)$ 
    if  $y_1 < y_2$  then  $a := x_1$  else  $b := x_2$ 
     $L := b - a$ 
until  $L < \varepsilon$ 
El punto de máximo  $x^*$  se halla en  $[a, b]$ 
Terminar

```

### **Ejemplo 1**

La función  $f(x) = x \sin x$  posee un punto  $x^*$  de máximo en el intervalo  $[2; 2,1]$ .

- ¿Cuántos experimentos se necesitará para determinar  $x^*$  en una región de amplitud 0,001 mediante el método de bisección?
- Hállalo.

Solución:

a) En este caso se tiene  $L_0 = 0,1$

y se desea obtener un  $L_n = 0,001$

Como  $L_n = \frac{L_0}{2^{n/2}}$

se tiene:  $2^{n/2} = \frac{L_0}{L_n} = \frac{0,1}{0,001} = 100$

y, aplicando logaritmos:

$$\frac{n}{2} = \frac{\ln 100}{\ln 2}$$

o sea:

$$n = 2 \frac{\ln 100}{\ln 2} = 13.288$$

Luego, se requiere de  $n = 14$  experimentos o, lo que es igual, 7 iteraciones del método. Compare con el ejemplo 3 de la sección anterior, y notará que en las etapas 2, 3, 4 y 5, para lograr la misma reducción mediante búsqueda secuencial uniforme, se requirió de 18 experimentos diferentes.

- b) Como se aspira a obtener un intervalo final de amplitud 0,001, se utilizará una distancia de separación  $\delta = 0,0001$  (diez veces menor que el intervalo final). En la tabla 1 se muestran todos los resultados.

| <i>iteración</i> | <i>a</i> | <i>b</i> | $x_1$   | $x_2$   | $f(x_1)$  | $f(x_2)$    |
|------------------|----------|----------|---------|---------|-----------|-------------|
| 0                | 2.00000  | 2,10000  | 2,04995 | 2,05005 | 1,8190957 | > 1,8190900 |
| 1                | 2.00000  | 2,05005  | 2,02498 | 2,02508 | 1,8196865 | < 1,8196875 |
| 2                | 2,02508  | 2,05005  | 2,03751 | 2,03761 | 1,8196020 | > 1,8195996 |
| 3                | 2,02508  | 2,03761  | 2,03129 | 2,03139 | 1,8196971 | > 1,8196964 |
| 4                | 2,02508  | 2,03139  | 2,02819 | 2,02829 | 1,8197053 | < 1,8197054 |
| 5                | 2,02819  | 2,03139  | 2,02974 | 2,02984 | 1,8197044 | > 1,8197042 |
| 6                | 2,02819  | 2,02984  | 2,02897 | 2,02907 | 1,8197057 | > 1,8197056 |
| 7                | 2,02819  | 2,02907  |         |         |           |             |

Tabla 1

Completadas siete iteraciones (14 experimentos) se ha llegado a un intervalo con amplitud 0,00088 menor que la tolerancia 0,001 requerida. Por tanto, la respuesta buscada es:

$$2,02819 < x^* < 2,02907$$

Nótese que, si se desea resolver un problema de optimización unidimensional (y esto es también válido en el caso multidimensional) con una precisión  $\varepsilon$ , entonces la distancia  $\delta$  debe ser mucho menor que  $\varepsilon$  pues en las iteraciones finales la amplitud de la región de búsqueda se va aproximando a  $\varepsilon$ . En la práctica puede tomarse  $\delta = 0.1\varepsilon$ , lo cual significa que en problemas de optimización se debe tener cuidado de no exigir una precisión  $\varepsilon$  exageradamente pequeña, ya que esto puede conducir a valores de  $\delta$  más pequeños que lo razonable. ■

### ***Ejemplo 2***

Suponga que los especialistas en agua subterránea han determinado que, en una región bajo estudio, el nivel del manto freático solo varía significativamente cuando se mide en puntos alejados entre si por lo menos 100 metros. ¿Puede pedirse entonces que se determine el punto en que el manto freático tiene su mayor altura con una precisión de 10 metros?

Solución:

Obviamente, si los experimentos no pueden hacerse a menos de 100 metros, no tiene mucho sentido pedir una precisión menor de 200 metros.

### *El método de Fibonacci*

Es necesario aclarar que este método no es debido a Fibonacci. Se le llama así porque hace uso de los números de Fibonacci, matemático del siglo XIV que los definió para resolver un problema que nada tuvo que ver con la optimización.

Si se analiza con cuidado el método de optimización mediante bisección, se verá (observe la figura 3) que, en cada iteración, el intervalo de búsqueda que se selecciona contiene un punto experimental que, por estar muy próximo a uno de los extremos, no puede ser aprovechado en la próxima iteración:

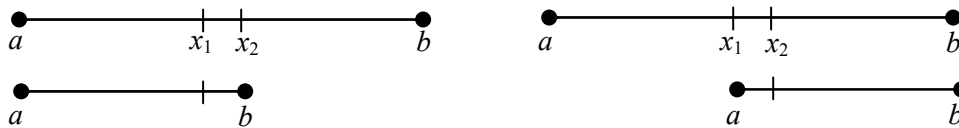


Figura 3

En el método de Fibonacci se trata de tomar los experimentos  $x_1$  y  $x_2$  más alejados entre sí, de modo que en la próxima iteración, el punto que quede dentro del nuevo intervalo de búsqueda, quede suficientemente cercano al centro como para que pueda ser utilizado como un punto experimental. De esta manera, cada iteración solo requerirá de un nuevo experimento, aunque a cambio de ello el intervalo no se reducirá a la mitad en cada paso sino a una proporción mayor.

En la figura 4 se muestra esta idea gráficamente:

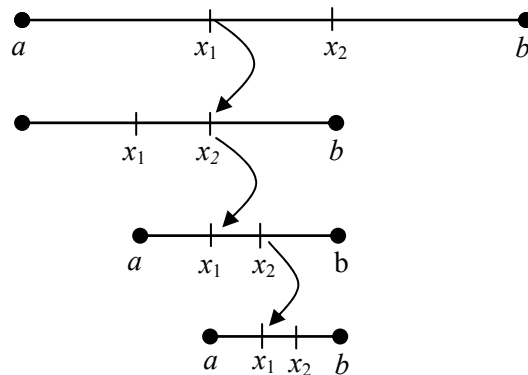


Figura 4

Para que el algoritmo sea sencillo, hay que procurar que las posiciones de  $x_1$  y  $x_2$  en cada iteración se puedan hallar aplicando una regla simple.

Una posible solución se basa en los números de Fibonacci, definidos como:

$$\begin{aligned}
 F_0 &= 1 \\
 F_1 &= 1 \\
 F_n &= F_{n-1} + F_{n-2} \quad \text{para } n = 2, 3, 4, \dots
 \end{aligned}$$

de donde se pueden fácilmente generar sus valores:

|       |   |   |   |   |   |   |    |    |    |     |
|-------|---|---|---|---|---|---|----|----|----|-----|
| $n$   | 0 | 1 | 2 | 3 | 4 | 5 | 6  | 7  | 8  | ... |
| $F_n$ | 1 | 1 | 2 | 3 | 5 | 8 | 13 | 21 | 34 | ... |

Si se supone que la amplitud  $b - a$  del intervalo inicial coincide con el número de Fibonacci  $F_N$  se tiene:

$$L = F_N$$

Como se muestra en la figura 5, el intervalo  $[a, b]$  se puede descomponer en dos subintervalos de longitudes  $F_{N-1}$  y  $F_{N-2}$ :

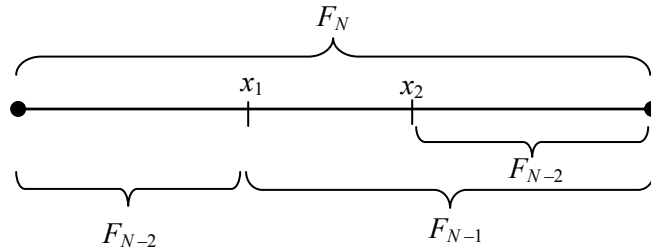


Figura 5

En la próxima iteración, como se observa en la figura 6, al aplicar la propiedad básica, el nuevo intervalo tendrá amplitud  $F_{N-1}$  y poseerá un punto experimental a una distancia  $F_{N-2}$  de uno de sus extremos:

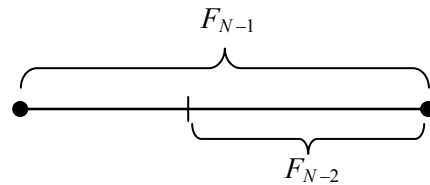


Figura 6

Obviamente (vea la figura 7), la distancia al otro extremo es  $F_{N-3}$ , pues  $F_{N-1} = F_{N-2} + F_{N-3}$ ,

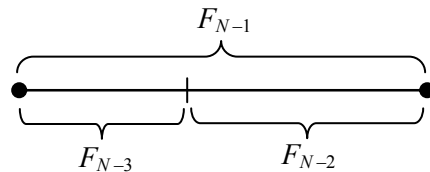


Figura 7

Falta solamente situar otro punto experimental simétricamente, como en la figura 8:

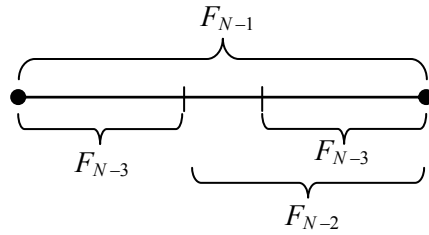


Figura 8

y el procedimiento puede repetirse hasta llegar a un intervalo de longitud  $F_2 = 2$ , como se ve en la figura 9:

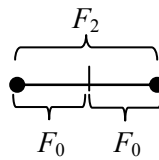


Figura 9

en el cual  $x_1$  y  $x_2$  coinciden. En la práctica, el procedimiento antes descrito tiene el inconveniente de que, desde que se comienza el proceso iterativo hay que seleccionar el número concreto de Fibonacci,  $F_N$  con el que se va a comenzar, del cual depende la reducción que habrá de conseguirse en el proceso, que será exactamente  $F_0/F_N$ . Esto puede evitarse con la siguiente modificación que, además, hace innecesario generar la sucesión de números de Fibonacci. Esta modificación se denomina, por razones que se verá enseguida, método de la *sección de oro*.

### ***El método de la sección de oro***

Considérense los números  $D_n$  definidos como una generalización de los de Fibonacci pero donde la condición  $F_0 = F_1 = 1$  se sustituye por el requerimiento de que la sucesión formada sea geométrica. Esto es:

$$D_n = D_{n-1} + D_{n-2} \quad n = 2, 3, 4, \dots \quad (1)$$

$$D_n = r D_{n-1} \quad n = 1, 2, 3, \dots \quad (2)$$

El valor de  $r$  puede encontrarse fácilmente. Dividiendo en la ecuación (1) por  $D_{n-1}$ :

$$\frac{D_n}{D_{n-1}} = 1 + \frac{D_{n-2}}{D_{n-1}}$$

y como, según (2):  $\frac{D_n}{D_{n-1}} = r$  y  $\frac{D_{n-2}}{D_{n-1}} = \frac{1}{r}$  resulta:

$$r = 1 + \frac{1}{r} \quad (3)$$

de donde

$$r^2 - r - 1 = 0$$

cuya raíz positiva es:

$$r = \frac{1 + \sqrt{5}}{2} = 1,618034$$

$$G = \frac{1}{r} = r - 1 = 0,618034$$

es la famosa *sección de oro* que era conocida desde la antigua Grecia y que posee propiedades sumamente interesantes.

Como los números  $D_n$  satisfacen la condición de que  $D_n = r D_{n-1}$ , sus valores quedan determinados al asignarle a uno cualquiera de ellos un valor deseado. Puesto que los números  $D_n$  poseen la misma propiedad que caracteriza a los de Fibonacci:

$$D_n = D_{n-1} + D_{n-2}$$

ellos pueden usarse en el algoritmo anterior. Por otra parte poseen la ventaja de que cada  $D_n$  se obtiene directamente a partir de su antecedente:

$$D_n = r D_{n-1}$$

o de su sucesor:

$$D_{n-1} = G D_n$$

En la figura 10 se muestra la manera en que se obtienen los puntos experimentales en el algoritmo de la *sección de oro*

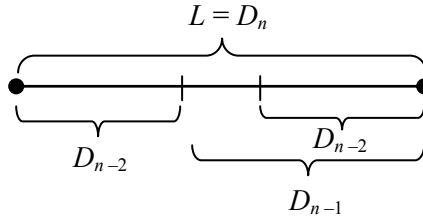


Figura 10

Pero como

$$D_{n-1} = G D_n = G L = 0,618034 L$$

y

$$D_{n-2} = L - D_{n-1} = L - G L = (1 - G)L = 0,381966 L$$

todo lo que se requiere en cada iteración es calcular el valor

$$(b - a) (1 - G)$$

para ubicar el nuevo punto experimental. Entre los números de Fibonacci y los  $D_n$  existe una interesante relación. En efecto, como

$$F_n = F_{n-1} + F_{n-2}$$

se tiene:

$$\frac{F_n}{F_{n-1}} = 1 + \frac{F_{n-2}}{F_{n-1}}$$

y, si  $n$  tiende hacia infinito

$$\lim_{n \rightarrow \infty} \frac{F_n}{F_{n-1}} = 1 + \lim_{n \rightarrow \infty} \frac{F_{n-2}}{F_{n-1}}$$

si se llama  $F$  al límite:  $\lim_{n \rightarrow \infty} \frac{F_n}{F_{n-1}}$  entonces se tiene:

$$F = 1 + \frac{1}{F}$$

Esta ecuación es idéntica a la (3) y, por tanto,  $F = r$ . Esto es, cuando  $n$  tiende hacia infinito, los números de Fibonacci se comportan como los  $D_n$ .

Como el método de Fibonacci y el de la sección de oro poseen idénticas características, en general se prefiere esta segunda variante, más fácil de implementar. Para concretar, se da a continuación el algoritmo de la sección de oro.

### ***Algoritmo en pseudo código del método de la sección de oro***

El algoritmo que sigue describe el método de optimización de la sección de oro. Se supone que la función  $f(x)$  es unimodal con máximo en el intervalo  $[a, b]$  y que está definida en todos los puntos de ese intervalo. El algoritmo da como resultado un intervalo cerrado de amplitud menor que un número especificado  $\varepsilon$  y que contiene al punto de máximo. Los datos que se requiere son: la función  $f(x)$ , el intervalo  $[a, b]$  en que se encuentra inicialmente el punto de máximo de la función y la tolerancia  $\varepsilon$ , que determina la amplitud del intervalo final.

```

L := b - a
Factor := 1 - G = 0,381966
x1 := a + Factor · L
x2 := b - Factor · L
y1 := f(x1)
y2 := f(x2)
repeat
  if y1 < y2 then
    a := x1
    x1 := x2
    y1 := y2
    L := b - a
    x2 := b - Factor · L
    y2 := f(x2)
  else
    b := x2
    x2 := x1
    y2 := y1
    L := b - a
    x1 := a + Factor · L
    y1 := f(x1)
  end
until L < ε

```

El punto de máximo  $x^*$  se halla en  $[a, b]$   
 Terminar

Es fácil determinar la eficiencia del método de la sección áurea. Llamando, como hasta ahora:

Amplitud del intervalo original:  $L_0 = b - a$ :

Amplitud del intervalo después de  $n$  experimentos:  $L_n$

se tiene:  $L_n = G L_{n-1} = G^2 L_{n-2} = \dots = G^{n-2} L_2$

ya que, después de los dos primeros experimentos, con cada experimento nuevo la región de búsqueda queda multiplicada por  $G$ . Como en la primera iteración se requiere de dos experimentos nuevos,

$$L_2 = G L_0$$

de ahí que:

$$L_n = G^{n-1} L_0$$

Esto es:

$$\frac{L_n}{L_0} = G^{n-1}$$

En la tabla 2 se muestra la relación  $L_n/L_0$  para diferentes valores de  $n$  en los métodos de bisección y de la sección áurea. Nótese que con la misma cantidad de experimentos, el método de la sección áurea logra una reducción mucho mayor de la región de búsqueda. La ventaja se ve más clara si se aprecia que en el método de bisección, cada dos nuevos experimentos reducen la región de incertidumbre a la mitad:

$$L_n = 0.5 L_{n-2}$$

pero en la sección áurea, después de la primera iteración

$$L_n = G^2 L_{n-2} = 0.382 L_{n-2}$$

| $n$ | $L_n/L_0$ en bisección | $L_n/L_0$ la sección áurea |
|-----|------------------------|----------------------------|
| 10  | 0.03125                | 0.01316                    |
| 12  | 0.01562                | 0.00502                    |
| 14  | 0.00781                | 0.00192                    |
| 16  | 0.00390                | 0.00073                    |
| 18  | 0.00195                | 0.00028                    |
| 20  | 0.00098                | 0.00011                    |
| 22  | 0.00049                | 0.00004                    |
| 24  | 0.00024                | 0.00002                    |
| 26  | 0.00012                | 0.00001                    |

Tabla 2

### Ejemplo 3

Se sabe que la función  $f(x) = x \sin x$  es unimodal y tiene un punto de máximo  $x^*$  en el intervalo  $[2; 2,1]$ . ¿Cuántos experimentos se necesitará para determinar  $x^*$  en una región de amplitud 0,001 usando el método de la sección de oro?

Solución:

$$\begin{aligned}L_0 &= 0.1 \\ \varepsilon &= 0.001\end{aligned}$$

$$\text{entonces} \quad L_n = G^{n-1}L_0 = G^{n-1}0.1 \leq \varepsilon = 0,001$$

$$\text{de donde:} \quad G^{n-1} \leq \frac{0,001}{0,1} = 0,01$$

$$\text{y, aplicando logaritmos:} \quad (n-1)\ln G \leq \ln 0,01$$

Ahora, al dividir por  $\ln G$ , como es negativo, la desigualdad cambia de sentido:

$$n-1 \geq \frac{\ln(0,01)}{\ln G}$$

$$n-1 \geq 9,57$$

$$n \geq 10,57$$

$$\text{O sea:} \quad n = 11$$

Compare con la solución del ejemplo 1, en el cual se calculó que el método de bisección requeriría 14 experimentos para este mismo problema.

### Refinamiento del resultado mediante interpolación

En general, todos los algoritmos tratados hasta aquí concluyen su trabajo dando un intervalo o región de incertidumbre suficientemente pequeña donde se encuentra el punto de máximo. En muchos casos, sin embargo, es conveniente seleccionar un punto de dicho intervalo que se pueda tomar como una aproximación razonable de  $x^*$ , sobre todo cuando se usa la optimización unidimensional como procedimiento interno de un algoritmo mucho más amplio de optimización multidimensional.

Una solución que a veces se emplea es tomar el punto medio de la región de incertidumbre. Otra posibilidad usada con frecuencia es la que se verá a continuación.

Sea  $f(x)$  una función unimodal con punto de máximo  $x^*$  y se suponen conocidos tres valores de  $x$ :  $x_1 < x_2 < x_3$  tales que  $x_1 < x^* < x_3$ . Sean  $y_1 = f(x_1)$ ,  $y_2 = f(x_2)$  y  $y_3 = f(x_3)$ . Se pretende hallar el número  $\bar{x}$ , abscisa del vértice (vea la figura 11) de la parábola que determinan los tres puntos  $(x_1, y_1)$ ,  $(x_2, y_2)$  y  $(x_3, y_3)$ . Sea  $y = p(x)$  la ecuación de dicha parábola:

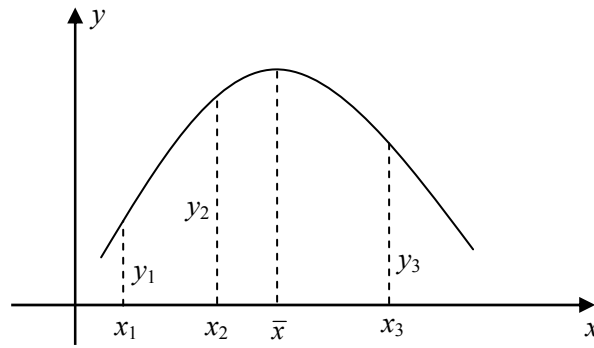


Figura 11

$$p(x) = a x^2 + b x + c$$

El valor  $\bar{x}$  anula a  $p'(x)$  o sea:  $2a \bar{x} + b = 0$

de donde: 
$$\bar{x} = -\frac{b}{2a}$$

Los valores de  $a$  y  $b$  se pueden obtener de diferentes formas, una de ellas es la que sigue. Como la parábola pasa por los tres puntos, se tiene:

$$\begin{cases} ax_1^2 + bx_1 + c = y_1 \\ ax_2^2 + bx_2 + c = y_2 \\ ax_3^2 + bx_3 + c = y_3 \end{cases}$$

Es decir: 
$$\begin{cases} a(x_2^2 - x_1^2) + b(x_2 - x_1) = y_2 - y_1 \\ a(x_3^2 - x_2^2) + b(x_3 - x_2) = y_3 - y_2 \end{cases}$$

Si se divide la primera ecuación por  $(x_2 - x_1)$ , la segunda por  $(x_3 - x_2)$  y se le llama

$$f_{12} = \frac{y_2 - y_1}{x_2 - x_1} \quad \text{y} \quad f_{23} = \frac{y_3 - y_2}{x_3 - x_2}$$

se obtiene:

$$\begin{cases} (x_1 + x_2)a + b = f_{12} \\ (x_2 + x_3)a + b = f_{23} \end{cases}$$

Ahora, utilizando la notación:  $x_{12} = \frac{x_1 + x_2}{2}$  y  $x_{23} = \frac{x_2 + x_3}{2}$  y resolviendo para  $a$  y  $b$ :

$$\bar{x} = \frac{x_{23}f_{12} - x_{12}f_{23}}{f_{12} - f_{23}}$$

$$\text{donde: } x_{12} = \frac{x_1 + x_2}{2} \quad f_{12} = \frac{y_2 - y_1}{x_2 - x_1} \quad x_{23} = \frac{x_2 + x_3}{2} \quad f_{23} = \frac{y_3 - y_2}{x_3 - x_2}$$

#### Ejemplo 4

Se sabe que la función  $f(x) = x \sen x$  es unimodal y posee un punto de máximo en el intervalo  $[2; 2,1]$ . Tome  $x_1 = 2$ ,  $x_2 = 2,05$ ,  $x_3 = 2,1$ , determine  $\bar{x}$  y compare con  $x^* = 2,0287578$ .

Solución:

$$\begin{aligned} y_1 &= f(x_1) = 2 \sen 2 = 1,818595 \\ y_2 &= f(x_2) = 2,05 \sen 2,05 = 1,819093 \\ y_3 &= f(x_3) = 2,1 \sen 2,1 = 1,812740 \end{aligned}$$

$$\begin{aligned} x_{12} &= \frac{x_1 + x_2}{2} = 2,025 & f_{12} &= \frac{y_2 - y_1}{x_2 - x_1} = 0,009960 \\ x_{23} &= \frac{x_2 + x_3}{2} = 2,075 & f_{23} &= \frac{y_3 - y_2}{x_3 - x_2} = -1,127064 \end{aligned}$$

$$\text{De donde: } \bar{x} = \frac{x_{23}f_{12} - x_{12}f_{23}}{f_{12} - f_{23}} = 2,028634$$

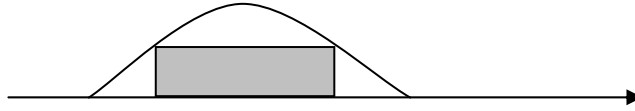
que solamente difiere de  $x^*$  en 0,000124. Como se ha visto en los ejemplos anteriores, para lograr un error unas 10 veces mayor que este, los métodos de bisección y de la sección áurea hubieran requerido respectivamente 14 y 11 evaluaciones de la función  $f(x)$ .

#### Ejercicios

Algunos de los ejercicios que siguen son muy laboriosos para realizar los cálculos a mano. Se supone que posea programas computacionales, preferiblemente confeccionados por usted, para ayudarlo en el trabajo. De no ser así, en los ejercicios que requieran mucho trabajo manual, determine los resultados con menos cifras exactas que las exigidas.

- Halle el menor punto de máximo positivo de la función  $f(x)$  con error menor que 0,001. Primero determine, mediante búsqueda secuencial, un intervalo que contiene al punto buscado y, después, aplique el método de bisección o el de Fibonacci. Compruebe su respuesta hallando los ceros de la derivada de  $f(x)$  mediante alguno de los métodos numéricos para resolver ecuaciones.
  - $f(x) = e^{-x} \sen x$
  - $f(x) = \sen x \cosh x$
  - $f(x) = \ln x - e^x$
  - $f(x) = \sen(\ln x)$
- Halle la mínima distancia desde el punto (3, 2) hasta la gráfica de la senoide  $y = \sen x$ . Obtenga la respuesta con tres cifras decimales exactas.

3. De todos los triángulos rectángulos de área 5, halle las dimensiones del que posee menor perímetro. De la solución con cuatro cifras decimales exactas.
4. Halle las dimensiones (con tres cifras decimales exactas) del rectángulo de mayor área que se puede inscribir en un arco de senoide, como se muestra en la figura:



5. Dados los puntos  $P_1 = (1; 1,2)$ ,  $P_2 = (2; 2,1)$  y  $P_3 = (3; 3)$ , halle la recta  $y = mx$  que minimiza la desviación cuadrática:

$$\sum_{i=1}^3 (y_i - mx_i)^2$$

Obtenga la solución con tres cifras decimales exactas.

6. Para hallar el punto de máximo de una función  $f(x)$  en un intervalo  $[a, b]$  existe un algoritmo que consiste en situar en el intervalo tres puntos distribuidos de manera que dividan al intervalo en cuatro partes iguales; evaluando  $f(x)$  en los tres puntos se puede obtener un nuevo intervalo que contiene al punto de máximo y cuya longitud es la mitad de la del intervalo original. Analice la eficiencia de este método en comparación con el de bisección y elabore un algoritmo en pseudo código para representarlo.

## 6.4 Conceptos básicos para la optimización multidimensional

La mayor parte de los problemas de optimización que surgen en la práctica corresponden a funciones que dependen de más de una variable. El problema general de la optimización multidimensional es mucho más complejo que el unidimensional, de manera que en este texto el análisis estará limitado a algunas técnicas numéricas fundamentales para resolver problemas sin restricciones. Aun así, será necesario dedicar esta sección a introducir algunos conceptos matemáticos imprescindibles para la comprensión de los algoritmos que se estudiarán posteriormente. Antes de acometer esta tarea resulta conveniente ver algunos ejemplos de problemas típicos de optimización multidimensional.

### *Ejemplo 1*

Se desea buscar la mejor ubicación, dentro de una amplia región del país, para una estación de bombeo que alimentará a un acueducto. El criterio que se va a usar para determinar la localización óptima, es que la profundidad del manto freático sea mínima, de manera que las bombas subterráneas que se van a instalar requieran motores de menor potencia. Obviamente, la función  $f(x, y)$  es la profundidad del manto freático y depende de dos variables: las coordenadas del punto donde se mide la profundidad.

■

### *Ejemplo 2*

La eficiencia de una planta industrial depende de una enorme cantidad de parámetros; podrían citarse: los parámetros de funcionamiento (velocidad, presión, temperatura, etc.) de cada uno de

los equipos de la planta, años de experiencia de cada operario, nivel de producción de cada artículo que produce la planta, etc. Un complejo problema de optimización consiste en determinar los valores de todas estas variables que permiten alcanzar una máxima eficiencia. ■

### ***Ejemplo 3***

El campo de ajuste de curvas es una gran fuente de problemas de optimización. Si se desea ajustar un modelo lineal en sus parámetros:

$$g(x) = C_1g_1(x) + C_2g_2(x) + \dots + C_ng_n(x)$$

donde  $g_1, g_2, \dots, g_n$  son funciones conocidas de la variable real  $x$ , a un conjunto  $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$  de datos, de manera que se minimice la desviación cuadrática:

$$D = \sum_{i=1}^m [g(x_i) - y_i]^2$$

entonces, como ya se sabe, el problema posee una solución analítica que se obtienen sin mucha dificultad resolviendo el sistema lineal de ecuaciones obtenido de hacer cero las derivadas parciales

$$\frac{\partial D}{\partial C_j} \quad j = 1, 2, \dots, n$$

Si el modelo es no lineal, o sea, si es del tipo general:

$$g(x) = F(x, C_1, C_2, \dots, C_n)$$

entonces la función a minimizar es muy complicada y sus derivadas parciales igualadas a cero no conducen a un sistema lineal. El problema de minimizar

$$\sum_{i=1}^m [F(x_i, C_1, C_2, \dots, C_n) - y_i]^2$$

se resuelve usualmente de forma numérica. ■

### ***Notación y representación gráfica***

Para representar más fácilmente una función real de  $n$  variables reales:

$$z = f(u_1, u_2, \dots, u_n)$$

se representará por  $\mathbf{X}$  el vector:

$$\mathbf{x} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

con lo cual, la función se escribe:

$$z = f(\mathbf{x})$$

Muchas de las técnicas de optimización multidimensional se comprenden mejor en el caso particular  $n = 2$ , o sea, para funciones de dos variables. La razón es obvia: en este caso la geometría puede ayudar a comprender la estrategia de búsqueda e, incluso, sugerir nuevas ideas y demostraciones. Cuando  $n = 2$ , se tiene:

$$\mathbf{x} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

y la función  $z = f(\mathbf{X})$  puede visualizarse mediante sus curvas de nivel. A modo de ejemplo, la figura 1 muestra la presencia de dos puntos de máximo relativo y un punto de ensilladura en una región del plano  $(u_1, u_2)$ .

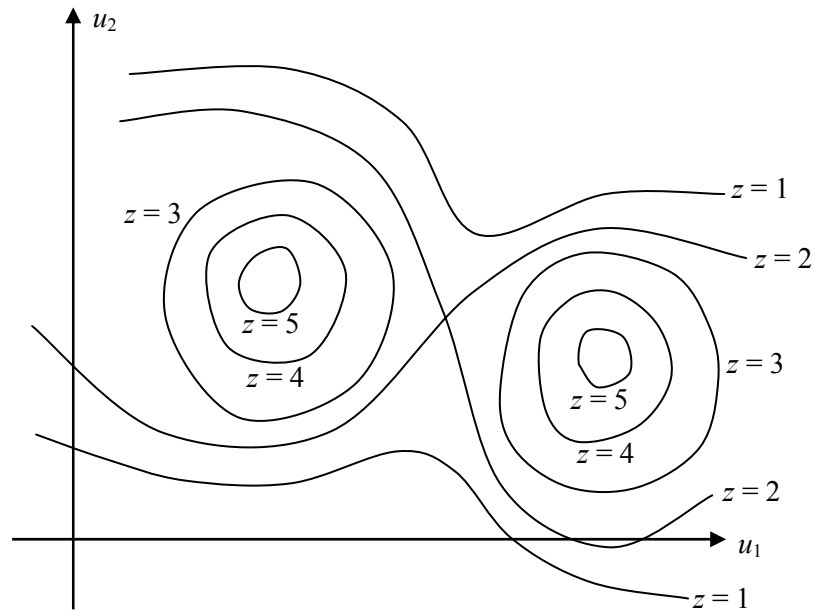


Figura 1

En lo que sigue se usarán con frecuencia funciones de dos variables y su representación mediante curvas de nivel como recurso didáctico para que se comprenda mejor propiedades que resultan necesariamente abstractas para  $n > 2$ .

### ***Trayectoria lineal en $R^n$***

El concepto de unimodalidad, que tan simple resulta en el caso unidimensional, puede hacerse muy complicado para el caso multidimensional. Por esa razón el análisis se restringirá a un tipo sencillo de unimodalidad, la unimodalidad lineal. Serán necesarias algunas definiciones previas.

Sea  $\mathbf{x}_0$  un vector que representa un punto en  $R^n$  y sea  $\mathbf{v}$  otro vector de  $R^n$  con el que se determina una cierta dirección en ese espacio. Se llamará “recta en  $R^n$  por el punto  $\mathbf{x}_0$  y con dirección  $\mathbf{v}$ ”, al conjunto de todos los  $\mathbf{X}$  que se obtienen de la expresión:

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{v}\lambda$$

al dar a  $\lambda$  todos los valores reales.

La figura 2 ilustra el concepto en el caso  $n = 2$ . Se dice también que los puntos  $\mathbf{x}$  forman una trayectoria lineal. Nótese que, aunque  $R^n$  posee dimensión  $n$ , la trayectoria lineal es un objeto unidimensional, en el cual cada punto  $\mathbf{x}$  está determinado por un solo parámetro real  $\lambda$ ; por esto, a veces se escribe:

$$\mathbf{x}(\lambda) = \mathbf{x}_0 + \lambda \mathbf{v}$$

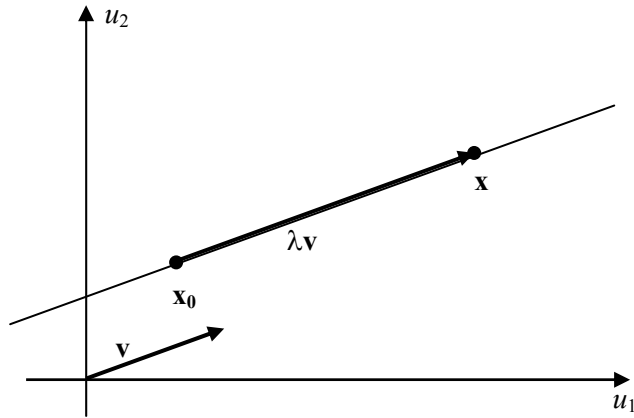


Figura 2

#### Ejemplo 4

Halle la ecuación de la recta de  $R^4$  determinada por:

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix} \quad \mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 1 \\ -2 \end{bmatrix}$$

Solución:

$$\mathbf{x} = \mathbf{x}_0 + \lambda \mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix} + \lambda \begin{bmatrix} 2 \\ 1 \\ 1 \\ -2 \end{bmatrix}$$

que se puede expresar como: 
$$\mathbf{x} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} 1 + 2\lambda \\ -1 + \lambda \\ \lambda \\ 2 - 2\lambda \end{bmatrix}$$

### ***Función linealmente unimodal***

Sea  $z = f(\mathbf{x})$  una función real de  $n$  variables reales, o sea,  $z \in R$ ,  $\mathbf{x} \in R^n$ . Si  $\mathbf{x}$  se restringe a que tome valores en una trayectoria lineal

$$\mathbf{x} = \mathbf{x}_0 + \lambda \mathbf{v}$$

entonces

$$z = f(\mathbf{x}_0 + \lambda \mathbf{v})$$

y como la única variable es  $\lambda$ , podemos decir que

$$z = F(\lambda)$$

es decir, una función unidimensional, en la cual el concepto de unimodalidad es muy sencillo.

#### **Definición:**

Se dice que  $z = f(\mathbf{x})$  es linealmente unimodal con máximo en la región  $R$ , si ella es unimodal con máximo sobre cualquier trayectoria recta contenida en  $R$ . ■

En la figura 3 se muestran las curvas de nivel de una función de dos variables y una trayectoria recta y en la figura 4, el perfil de la función  $z$  a lo largo de la trayectoria recta. Si este comportamiento es así para todas las trayectorias rectas, entonces  $z = f(\mathbf{x})$  es linealmente unimodal con máximo.

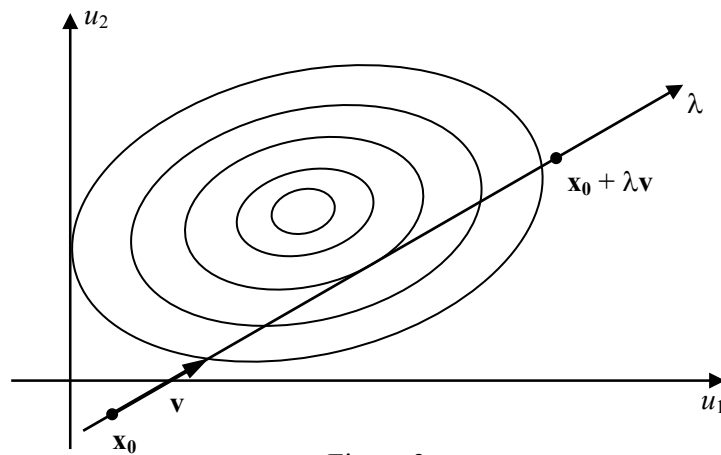


Figura 3

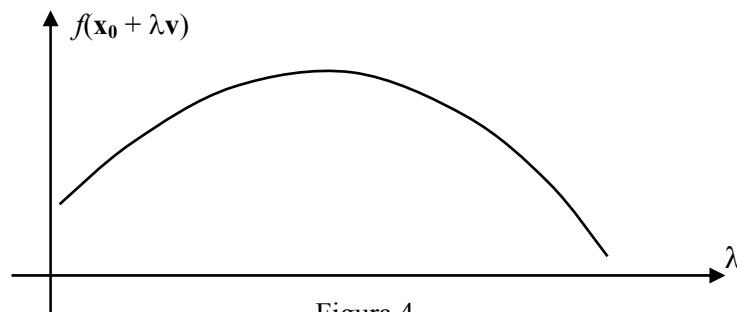


Figura 4

## Funciones cuadráticas de $n$ variables

Una clase importante de funciones linealmente unimodales son las funciones cuadráticas con punto de máximo.

Definición:

Se dice que la función  $f(\mathbf{x})$  es cuadrática si la misma se puede expresar como:

$$\begin{aligned} f(\mathbf{x}) = & \frac{1}{2} [a_{11}u_1^2 + a_{12}u_1u_2 + \dots + a_{1n}u_1u_n + \\ & + a_{21}u_2u_1 + a_{22}u_2^2 + \dots + a_{2n}u_2u_n + \dots \\ & \dots + a_{n1}u_nu_1 + a_{n2}u_nu_2 + \dots + a_{nn}u_n^2] \\ & - [b_1u_1 + b_2u_2 + \dots + b_nu_n] + c \end{aligned}$$

El factor  $\frac{1}{2}$  común a todos los términos cuadráticos, así como el signo menos común a todos los términos lineales, persiguen el objetivo de obtener posteriormente expresiones más sencillas. Nótese que todos los términos rectangulares, esto es, del tipo  $u_iu_j$  ( $i \neq j$ ) aparecen repetidos pues

$$a_{ij}u_iu_j \text{ y } a_{ji}u_ju_i$$

son semejantes y, de hecho, se podrían haber escrito como

$$(a_{ij} + a_{ji}) u_iu_j = k_{ij}u_iu_j$$

Sin embargo, a los efectos de introducir posteriormente la notación matricial, es preferible mantener la duplicidad, así que se conservarán los términos semejantes  $a_{ij}u_iu_j$  y  $a_{ji}u_ju_i$

Como, a los efectos de la optimización de  $f(\mathbf{x})$ , el término  $c$  carece de importancia, ya que

$$f(\mathbf{x}) \text{ y } f(\mathbf{x}) + c$$

poseen idénticos puntos de extremo, en todo lo que sigue, se supondrá que  $c = 0$ .

Las funciones cuadráticas pueden ser representadas matricialmente en forma muy cómoda. Para ello se introduce la matriz  $\mathbf{A}$ : coeficientes de los términos cuadráticos; y  $\mathbf{b}$ : coeficientes de los términos lineales:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

El producto  $\frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x}$  al ser desarrollado resulta:

$$\begin{aligned}
& \frac{1}{2} \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ \vdots \\ a_{n1}u_1 + a_{n2}u_2 + \cdots + a_{nn}u_n \end{bmatrix} \\
&= \frac{1}{2} \left[ a_{11}u_1^2 + a_{12}u_1u_2 + \cdots + a_{1n}u_1u_n + \right. \\
&\quad \left. + a_{21}u_2u_1 + a_{22}u_2^2 + \cdots + a_{2n}u_2u_n + \cdots \right. \\
&\quad \left. \cdots + a_{n1}u_nu_1 + a_{n2}u_nu_2 + \cdots + a_{nn}u_n^2 \right]
\end{aligned}$$

que son los términos cuadráticos de  $f(\mathbf{x})$ . Por otra parte,

$$\mathbf{b}^T \mathbf{x} = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = b_1u_1 + b_2u_2 + \cdots + b_nu_n$$

son los términos lineales de  $f(\mathbf{x})$ . Así que:

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x}$$

Como se recordará, los términos cuadráticos de tipo rectangular aparecen cada uno dos veces:  $a_{ij}u_iu_j$  y  $a_{ji}u_ju_i$ . Puesto que lo importante es que la suma

$$a_{ij} + a_{ji}$$

tenga un valor determinado, no hay pérdida de generalidad si se supone que

$$a_{ij} = a_{ji} \quad \text{para } i \neq j$$

de modo que

$$a_{ij} + a_{ji} = 2a_{ij} = 2a_{ji}$$

El tomar  $a_{ij} = a_{ji}$  hace que la matriz  $\mathbf{A}$  sea simétrica.

A modo de resumen se establecerá como teorema la propiedad que se acaba de probar.

Teorema:

Una función cuadrática cualquiera  $f$  de  $n$  variables independientes se puede expresar como

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x}$$

donde  $\mathbf{A}$  es una matriz  $n \times n$  simétrica  
y  $\mathbf{b}$  es una matriz  $n \times 1$

■

### ***Matrices positivas definidas y negativas definidas***

Dentro de las matrices simétricas, existe un tipo especial de matriz, llamadas *definidas*, que juegan un papel importante en el análisis de las funciones cuadráticas.

Definición:

Sea  $\mathbf{A}_{n \times n}$  una matriz simétrica. Si para todo vector no nulo

$$\mathbf{x} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

se cumple que  $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ , entonces la matriz  $\mathbf{A}$  se llama positiva definida. Si para todo  $\mathbf{x}$  no nulo se tiene  $\mathbf{x}^T \mathbf{A} \mathbf{x} < 0$ , la matriz  $\mathbf{A}$  se llama negativa definida.

■

### ***Ejemplo 5***

Halle qué condiciones deben cumplir  $a$ ,  $b$  y  $c$  para que la matriz simétrica de orden dos:

$$\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

sea positiva definida.

Solución:

Sea  $\mathbf{x} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$  entonces:

$$\begin{aligned} \mathbf{x}^T \mathbf{A} \mathbf{x} &= \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} > 0 \\ &= au_1^2 + 2bu_1u_2 + cu_2^2 > 0 \end{aligned}$$

De aquí resulta que tiene que ser  $a > 0$  pues de lo contrario, bastaría tomar  $u_2 = 0$  y  $u_1 \neq 0$  para obtener un resultado negativo. Por razones similares es obvio que tiene también que ser  $c > 0$ . Pero esto no basta; completando cuadrados con los dos primeros términos de la expresión anterior:

$$\begin{aligned}
&= au_1^2 + 2bu_1u_2 + \frac{b^2u_2^2}{a} - \frac{b^2u_2^2}{a} + cu_2^2 > 0 \\
&= \left( \sqrt{a}u_1 + \frac{bu_2}{\sqrt{a}} \right)^2 - \frac{b^2u_2^2}{a} + cu_2^2 > 0 \\
&= \left( \sqrt{a}u_1 + \frac{bu_2}{\sqrt{a}} \right)^2 + \left( -\frac{b^2}{a} + c \right) u_2^2 > 0 \\
&= \left( \sqrt{a}u_1 + \frac{bu_2}{\sqrt{a}} \right)^2 + \left( \frac{ac - b^2}{a} \right) u_2^2 > 0
\end{aligned}$$

y para que esto sea cierto es necesario y suficiente que, además de la condiciones  $a > 0$  y  $c > 0$ , sea

$$ac - b^2 > 0 \quad \blacksquare$$

Las matrices positivas definidas y negativas definidas poseen una gran importancia a la hora de caracterizar a una función cuadrática. Para ello es fundamental el siguiente teorema.

Teorema:

Si  $\mathbf{A}$  es positiva (negativa) definida la función cuadrática

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x}$$

posee un punto  $\mathbf{x}^*$  de mínimo (máximo) y se cumple que  $\mathbf{A}\mathbf{x}^* = \mathbf{b}$ .

Demostración:

Solo se probará el caso en que  $\mathbf{A}$  es positiva definida; el otro de demuestra de modo similar.

Sea  $\mathbf{x}^*$  tal que

$$\mathbf{A}\mathbf{x}^* = \mathbf{b}$$

(Puede probarse que toda matriz positiva definida es no singular, de ahí que la condición  $\mathbf{A}\mathbf{x}^* = \mathbf{b}$  se satisface para uno y solo un  $\mathbf{x}^*$ ).

Sea  $\mathbf{x} \neq \mathbf{x}^*$

$$f(\mathbf{x}) - f(\mathbf{x}^*) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x} - \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x}^*$$

Pero como  $\mathbf{b}^T = (\mathbf{A}\mathbf{x}^*)^T = \mathbf{x}^{*T} \mathbf{A}^T = \mathbf{x}^{*T} \mathbf{A}$

$$\begin{aligned}
f(\mathbf{x}) - f(\mathbf{x}^*) &= \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^{*T} \mathbf{A} \mathbf{x} - \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} \mathbf{x}^* + \mathbf{x}^{*T} \mathbf{A} \mathbf{x}^* \\
&= \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^{*T} \mathbf{A} \mathbf{x} + \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} \mathbf{x}^*
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} \mathbf{x} - \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} \mathbf{x} + \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} \mathbf{x}^* \\
&= \frac{1}{2} (\mathbf{x}^T - \mathbf{x}^{*T}) \mathbf{A} \mathbf{x} - \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} (\mathbf{x} - \mathbf{x}^*) \\
&= \frac{1}{2} (\mathbf{x} - \mathbf{x}^*)^T \mathbf{A} \mathbf{x} - \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} (\mathbf{x} - \mathbf{x}^*)
\end{aligned}$$

Ambos sumandos de esta expresión son escalares (matrices de 1x1) así que el primer sumando es igual que su traspuesta:

$$\begin{aligned}
f(\mathbf{x}) - f(\mathbf{x}^*) &= \frac{1}{2} (\mathbf{A} \mathbf{x})^T (\mathbf{x} - \mathbf{x}^*) - \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} (\mathbf{x} - \mathbf{x}^*) \\
&= \frac{1}{2} \mathbf{x}^T \mathbf{A} (\mathbf{x} - \mathbf{x}^*) - \frac{1}{2} \mathbf{x}^{*T} \mathbf{A} (\mathbf{x} - \mathbf{x}^*) \\
&= \frac{1}{2} (\mathbf{x}^T - \mathbf{x}^{*T}) \mathbf{A} (\mathbf{x} - \mathbf{x}^*) \\
&= \frac{1}{2} (\mathbf{x} - \mathbf{x}^*)^T \mathbf{A} (\mathbf{x} - \mathbf{x}^*) > 0
\end{aligned}$$

por ser  $\mathbf{A}$  positiva definida y  $\mathbf{x} - \mathbf{x}^*$  no nulo. Como se ha probado que

$$\mathbf{x} \neq \mathbf{x}^* \Rightarrow f(\mathbf{x}) - f(\mathbf{x}^*) > 0$$

queda demostrado que  $\mathbf{x}^*$  es el único punto de mínimo de  $f(\mathbf{x})$

■

Las funciones cuadráticas con mínimo o máximo son un importante ejemplo de funciones linealmente unimodales. Esto se establecerá como un teorema:

Teorema:

Si  $\mathbf{A}$  es positiva definida la función cuadrática con mínimo

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x}$$

es linealmente unimodal.

Demostración:

Sea  $\mathbf{x} = \mathbf{x}_0 + \lambda \mathbf{v}$

Entonces

$$\begin{aligned}
f(\mathbf{x}) &= \frac{1}{2} (\mathbf{x}_0 + \lambda \mathbf{v})^T \mathbf{A} (\mathbf{x}_0 + \lambda \mathbf{v}) - \mathbf{b}^T (\mathbf{x}_0 + \lambda \mathbf{v}) \\
&= \frac{1}{2} (\mathbf{x}_0^T + \lambda \mathbf{v}^T) \mathbf{A} (\mathbf{x}_0 + \lambda \mathbf{v}) - \mathbf{b}^T (\mathbf{x}_0 + \lambda \mathbf{v}) \\
&= \frac{1}{2} \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + \frac{1}{2} \lambda \mathbf{v}^T \mathbf{A} \mathbf{x}_0 + \frac{1}{2} \lambda \mathbf{x}_0^T \mathbf{A} \mathbf{v} + \frac{1}{2} \lambda^2 \mathbf{v}^T \mathbf{A} \mathbf{v} - \mathbf{b}^T \mathbf{x}_0 - \lambda \mathbf{b}^T \mathbf{v}
\end{aligned}$$

$$= \frac{1}{2} (\mathbf{v}^T \mathbf{A} \mathbf{v}) \lambda^2 + (\frac{1}{2} \mathbf{v}^T \mathbf{A} \mathbf{x}_0 + \frac{1}{2} \mathbf{x}_0^T \mathbf{A} \mathbf{v} - \mathbf{b}^T \mathbf{v}) \lambda + (\frac{1}{2} \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 - \mathbf{b}^T \mathbf{x}_0)$$

Es decir,  $f(\mathbf{x})$  evaluada sobre la recta  $\mathbf{x} = \mathbf{x}_0 + \lambda \mathbf{v}$  es una función cuadrática de  $\lambda$ :

$$F(\lambda) = a\lambda^2 + b\lambda + c$$

donde  $a = \frac{1}{2} (\mathbf{v}^T \mathbf{A} \mathbf{v})$  es positivo, pues el vector de dirección  $\mathbf{v}$  no puede ser nulo y  $\mathbf{A}$  es positiva definida. La función  $F(\lambda) = a\lambda^2 + b\lambda + c$  con  $a > 0$  es unimodal con mínimo independientemente de  $b$  y  $c$ , así que  $f(\mathbf{x})$  es linealmente unimodal. ■

### Ejemplo 6

Considere la función cuadrática de dos variables:

$$f(\mathbf{x}) = 100 - 3u_1^2 - 4u_2^2 + 5u_1u_2 + 2u_1$$

- Expresar  $f(\mathbf{x})$  en forma matricial.
- Probar que  $\mathbf{A}$  es definida negativa.
- Halle el punto de máximo de  $f(\mathbf{x})$ .

Solución:

$$a) \quad f(\mathbf{x}) = -3u_1^2 + \frac{5}{2}u_1u_2 + \frac{5}{2}u_1u_2 - 4u_2^2 + 2u_1 + 100$$

$$= \frac{1}{2}(-6u_1^2 + 5u_1u_2 + 5u_1u_2 - 8u_2^2) - (-2u_1) + 100$$

$$= \frac{1}{2} \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} -6 & 5 \\ 5 & -8 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} - \begin{bmatrix} -2 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + 100$$

$$= \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x} + c$$

$$f(\mathbf{X}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x} + c$$

$$\text{donde } \mathbf{A} = \begin{bmatrix} -6 & 5 \\ 5 & -8 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} -2 \\ 0 \end{bmatrix} \quad c = 100$$

- Para probar que  $\mathbf{A}$  es definida negativa basta probar que  $-\mathbf{A}$  es definida positiva. En efecto:

$$-\mathbf{A} = \begin{bmatrix} 6 & -5 \\ -5 & 8 \end{bmatrix}$$

cumple las condiciones deducidas en el ejemplo 5 para una matriz

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

ya que posee a y c positivos y  $ac - b^2 = 48 - 25 > 0$

c) El punto de máximo  $\mathbf{x}^*$  satisface que  $\mathbf{Ax}^* = \mathbf{b}$ . Esto es:

$$\begin{bmatrix} 6 & -5 \\ -5 & 8 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

$$\begin{cases} -6u_1 + 5u_2 = -2 \\ 5u_1 - 8u_2 = 0 \end{cases}$$

cuya solución es:  $u_1^* = \frac{16}{23} \quad u_2^* = \frac{10}{23}$

### ***Algoritmo para hallar el óptimo en una dirección***

En varios de los métodos de optimización multidimensional que se estudiarán a partir de esta sección, se necesita hallar el punto de extremo de una función  $f(\mathbf{X})$  a lo largo de una trayectoria recta. Esto requiere elaborar un algoritmo numérico adecuado a este fin.

Evidentemente el algoritmo necesita como datos:

- $f(\mathbf{x})$ : función de  $u_1, u_2, \dots, u_n$
- $\mathbf{x}_0$ : vector de  $R^n$  a partir del cual se buscará un punto de máximo
- $\mathbf{v}$ : vector unitario de  $R^n$  en cuya dirección se hará la búsqueda
- $\varepsilon$ : tolerancia con que se hallará el punto de máximo

El algoritmo será llamado MaxUnidim y será una función que posee parámetros de entrada  $f(\mathbf{x})$ ,  $\mathbf{x}_0$ ,  $\mathbf{v}$  y  $\varepsilon$  y da como resultado un número real

$$\text{MaxUnidim}(f, \mathbf{x}_0, \mathbf{v}, \varepsilon)$$

Este número real se encuentra de la siguiente forma:

```

s := 1
do while  $f(\mathbf{x}_0 + s\mathbf{v}) < f(\mathbf{x}_0)$  and  $s > \varepsilon/10$ 
    s := s/2
end
if s <  $\varepsilon/10$  then
    No hay máximo en la dirección y sentido de  $\mathbf{v}$ 
    Búsqueda sin éxito
else
    Realizar búsqueda secuencial uniforme con paso s para la función
     $F(\lambda) = f(\mathbf{x}_0 + \lambda\mathbf{v})$  hasta obtener un intervalo  $[\lambda_{\text{inf}}, \lambda_{\text{sup}}]$  o llegar a 1000 pasos
    (Búsqueda sin éxito).
    if (Búsqueda con éxito) then
        En el intervalo  $[\lambda_{\text{inf}}, \lambda_{\text{sup}}]$  realizar búsqueda por bisección con  $\delta = \varepsilon/20$ 
        hasta que  $\lambda_{\text{sup}} - \lambda_{\text{inf}} < \varepsilon$ 
    
```

```

        end
    end
    if (búsqueda con éxito) then
        MaxUnidim =  $\frac{1}{2}(\lambda_{\text{inf}} + \lambda_{\text{sup}})$ 
    else
        MaxUnidim = -1
    end
    Terminar

```

Como se aprecia, no se han tomado las técnicas más eficientes sino las más robustas, teniendo en cuenta que no será posible observar directamente el funcionamiento del algoritmo.

## 6.5 El método de búsqueda por coordenadas

Este es posiblemente el método más elemental para optimizar funciones multidimensionales. En general, es un algoritmo poco eficiente pero, en algunos casos especiales, ofrece buenos resultados. Se basa en la idea de realizar búsquedas unidimensionales en las direcciones de las variables de la función  $f$ . Esto es:

1. Hallar el punto  $\mathbf{x}_1$  de máximo de  $f(\mathbf{x})$  a partir de  $\mathbf{x}_0$  tomando  $u_1$  variable y las demás constantes.
2. Hallar el punto  $\mathbf{x}_2$  de máximo de  $f(\mathbf{x})$  a partir de  $\mathbf{x}_1$  tomando  $u_2$  variable y las demás constantes.
3. Hallar el punto  $\mathbf{x}_3$  de máximo de  $f(\mathbf{x})$  a partir de  $\mathbf{x}_2$  tomando  $u_3$  variable y las demás constantes.

Hasta:

- $n$ . Hallar el punto  $\mathbf{x}_n$  de máximo de  $f(\mathbf{x})$  a partir de  $\mathbf{x}_{n-1}$  tomando  $u_n$  variable y las demás constantes.

Hacer  $\mathbf{x}_0 = \mathbf{x}_n$  y comenzar por el paso 1 de nuevo.

El algoritmo se detiene cuando los puntos  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$  difieren muy poco entre sí.

En la figura 1 se aprecia gráficamente como funciona el algoritmo para una función de dos variables  $u_1$  y  $u_2$  cuyas curvas de nivel se muestran, donde  $z_1 < z_2 < z_3$  y  $\mathbf{x}^*$  es el punto de máximo. En la figura cada flecha representa una búsqueda en una dirección paralela a un eje coordenado.

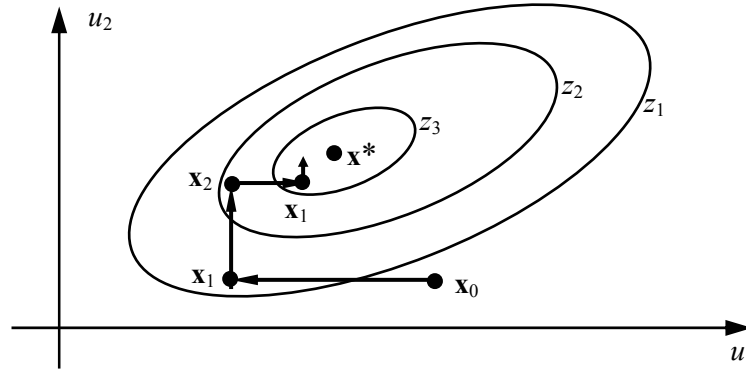


Figura 1

El algoritmo de búsqueda por coordenadas requiere que la función  $f(\mathbf{x})$  sea linealmente unimodal, de modo que cada problema de búsqueda unidimensional tenga una solución.

### *Algoritmo en pseudo código*

El siguiente algoritmo halla el punto de máximo de una función  $f(\mathbf{x})$  linealmente unimodal mediante el método de búsqueda por coordenadas. Utiliza como datos el número  $n$  de variables independientes, la función  $f(\mathbf{x})$ , el vector  $n$  dimensional  $\mathbf{x}_0$  por donde comenzará la búsqueda y la tolerancia  $\varepsilon$  que se permitirá en el algoritmo. Se supone que se cuenta con un algoritmo como MaxUnidim que permite realizar búsquedas unidimensionales en cualquier dirección que se especifique mediante un vector  $n$  dimensional  $\mathbf{v}$  (este algoritmo fue desarrollado en la sección 6.4).

```

repeat
     $\lambda_{\max} := 0$  {Esta variable almacenará la mayor de las distancias recorridas en la
                  secuencia de búsquedas unidimensionales}
    for  $i = 1$  to  $n$ 
         $\mathbf{v} := \mathbf{e}_i$  { $i$ -simo vector canónico: un 1 en la  $i$ -sima coordenada y 0 en las
                     otras}
         $\lambda := \text{MaxUnidim}(f, \mathbf{x}_{i-1}, \mathbf{v}, \varepsilon)$ 
        if  $\lambda < 0$  then {Esto solo sucede si la búsqueda en dirección  $\mathbf{e}_i$  no tuvo
                        éxito}
             $\mathbf{v} := -\mathbf{e}_i$  {Se realiza una búsqueda en el sentido opuesto de  $\mathbf{e}_i$ }
             $\lambda := \text{MaxUnidim}(f, \mathbf{x}_{i-1}, \mathbf{v}, \varepsilon)$ 
            if  $\lambda < 0$  then  $\lambda = 0$  {Se supone que, si en ambos sentidos la
                                búsqueda no tuvo éxito, entonces el punto de
                                máximo se encuentra en el punto inicial}
        end
         $\mathbf{x}_i = \mathbf{x}_{i-1} + \lambda \mathbf{v}$ 
        if  $\lambda > \lambda_{\max}$  then  $\lambda_{\max} := \lambda$ 
    end
     $\mathbf{x}_0 := \mathbf{x}_n$ 
until  $\lambda_{\max} < \varepsilon$ 
El punto de máximo es  $\mathbf{x}_n$ 
Terminar

```

Observe que los vectores  $\mathbf{e}_i$  cuya  $i$ -ésima componente es uno y las demás son ceros, son vectores unitarios. Como se ve, el algoritmo prevé la posibilidad de que en una iteración cualquiera el punto de máximo pueda estar en la dirección  $\mathbf{e}_i$  o  $-\mathbf{e}_i$  (esto es, en la dirección del vector  $\mathbf{e}_i$  pero en su mismo sentido o en el sentido opuesto). Note también cómo, en el caso en que no hay punto de máximo ( $\lambda < 0$ ) en ambos sentidos, se supone  $\lambda = 0$ , lo cual significa que el punto  $\mathbf{X}_{i-1}$  era, casualmente, el punto de máximo en la dirección de  $\mathbf{e}_i$ .

### Ejemplo 1

En el ejemplo 6 de la sección 6.4 se demostró que la función cuadrática

$$f(\mathbf{x}) = 100 - 3u_1^2 - 4u_2^2 + 5u_1u_2 + 2u_1$$

posee un punto de máximo, el cual se calculó analíticamente:

$$\mathbf{x}^* = \begin{bmatrix} \frac{16}{23} \\ \frac{10}{23} \end{bmatrix} = \begin{bmatrix} 0.69565... \\ 0.43478... \end{bmatrix}$$

Determine  $\mathbf{x}^*$  por el método de búsqueda por coordenadas con precisión 0,0005 para:

$$\text{a) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{b) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{c) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{d) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Halle en cada caso la cantidad de búsquedas unidimensionales necesitadas.

Solución:

Nótese que la función  $f(\mathbf{X})$ , por ser cuadrática con máximo, es linealmente unimodal, por lo cual el algoritmo de búsqueda por coordenadas es aplicable. Los cálculos se realizaron con un programa elaborado utilizando el algoritmo anterior. Los resultados obtenidos fueron:

$$\text{a) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,6953 \\ 0,4345 \end{bmatrix} \quad \text{en 24 iteraciones.}$$

$$\text{b) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,6950 \\ 0,4343 \end{bmatrix} \quad \text{en 22 iteraciones.}$$

$$\text{c) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,6961 \\ 0,4351 \end{bmatrix} \quad \text{en 24 iteraciones.}$$

$$\text{d) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,6961 \\ 0,4351 \end{bmatrix} \quad \text{en 24 iteraciones.}$$

■

Como se observa en el ejemplo anterior, el método de búsqueda por coordenadas requiere una buena cantidad de búsquedas unidimensionales para su convergencia. Esto ocurre porque las curvas de nivel tienen forma de elipses inclinadas respecto a los ejes, lo cual sucede en presencia de términos rectangulares (del tipo de  $5u_1u_2$  que posee este ejemplo). Si hay poca interacción entre las variables, lo que corresponde geoméricamente con curvas de nivel de ejes paralelos a los coordenados, la misma interpretación geométrica permite asegurar que la convergencia ha de ser muy rápida.

Note, sin embargo, que la rapidez de convergencia no está influida significativamente por la selección de  $\mathbf{x}_0$  siempre que se tome a una distancia adecuada del punto de óptimo.

## ***Ejemplo 2***

La función cuadrática:  $f(\mathbf{x}) = 100 - 4u_1^2 - 5u_2^2 + 2u_1 + 3u_2$

posee un punto de máximo cuyo valor exacto es:

$$\mathbf{x}^* = \begin{bmatrix} 0,25 \\ 0,3 \end{bmatrix}$$

Utilice el método de búsqueda por coordenadas para hallarlo, con precisión de 0.0005. Tome punto de partida:

$$\text{a) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{b) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{c) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{d) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

y determine el número de búsquedas unidimensionales en cada caso.

Solución:

Utilizando el mismo programa del ejemplo anterior, los resultados fueron:

$$\text{a) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,25 \\ 0,3 \end{bmatrix} \quad \text{en 3 búsquedas unidimensionales.}$$

$$\text{b) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,25 \\ 0,3 \end{bmatrix} \quad \text{en 3 búsquedas unidimensionales.}$$

$$\text{c) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,25 \\ 0,3 \end{bmatrix} \quad \text{en 3 búsquedas unidimensionales.}$$

$$\text{d) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,25 \\ 0,3 \end{bmatrix} \quad \text{en 3 búsquedas unidimensionales.}$$

■

Nótese cómo, al no existir interacción entre  $u_1$  y  $u_2$  (curvas de nivel con ejes no inclinados) el método converge rápidamente.

### Ejemplo 3

La función de Rosembrook ha sido propuesta para poner a prueba los métodos de optimización multidimensional. Es una función muy difícil, pues sus curvas de nivel son, como se muestra en la figura 2, “óvalos” muy alargados y con su eje mayor curvado en forma de parábola. Esta función posee un punto de mínimo en  $(1, 1)$  y se recomienda comenzar la búsqueda en  $(-1, 1)$  para someter a prueba a un algoritmo. Intente usar búsqueda por coordenadas para la función de Rosembrook:

$$f(\mathbf{x}) = 100(u_2 - u_1^2)^2 + (1 - u_1)^2$$

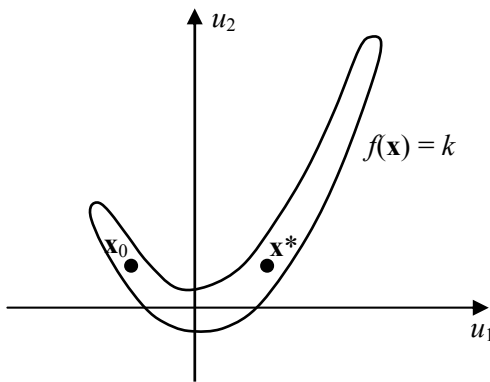


Figura 2

Solución:

La convergencia es sumamente lenta. Después de 1000 búsquedas unidimensionales se tiene:

$$\mathbf{x}^* \approx \begin{bmatrix} 0,922 \\ 0,850 \end{bmatrix}$$

después de 1500 búsquedas unidimensionales:

$$\mathbf{x}^* \approx \begin{bmatrix} 0,961 \\ 0,924 \end{bmatrix}$$

y después de 2000:

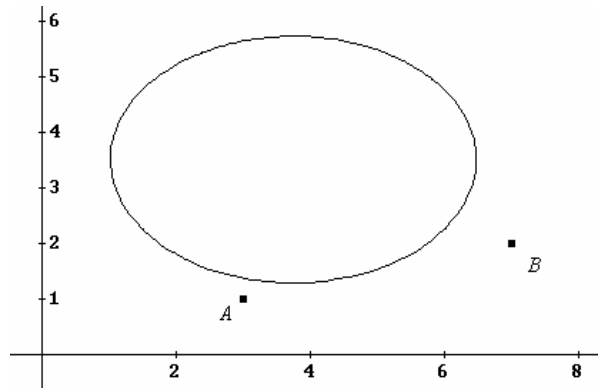
$$\mathbf{x}^* \approx \begin{bmatrix} 0,980 \\ 0,960 \end{bmatrix}$$

## Ejercicios

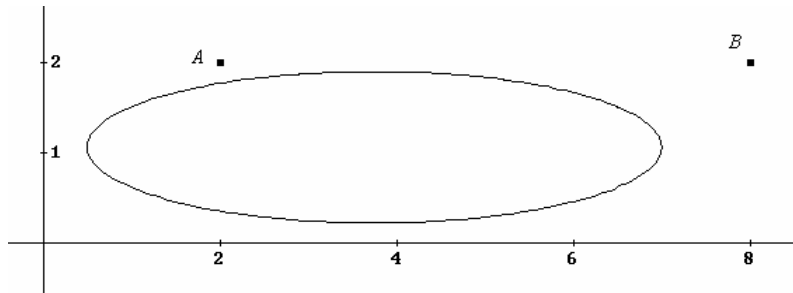
Algunos de los ejercicios que siguen son muy laboriosos para realizar los cálculos a mano. Se supone que posea programas computacionales, preferiblemente confeccionados por usted, para ayudarle en el trabajo. De no ser así, en los ejercicios que requieran mucho trabajo manual, determine los resultados con menos cifras exactas que las exigidas.

1. En cada una de las funciones cuadráticas que siguen se muestra una curva de nivel, cuya forma puede servir para predecir como se comportará el método de búsqueda por coordenadas al determinar su punto de extremo. En cada caso, exprese la función de forma matricial y aplique el método de búsqueda por coordenadas para hallar el punto de extremo utilizando como punto de partida los puntos  $A$  y  $B$  que se dan en la propia figura. Obtenga la solución con tres cifras decimales exactas. ¿Puede dar alguna conclusión general como resultado de este ejercicio?

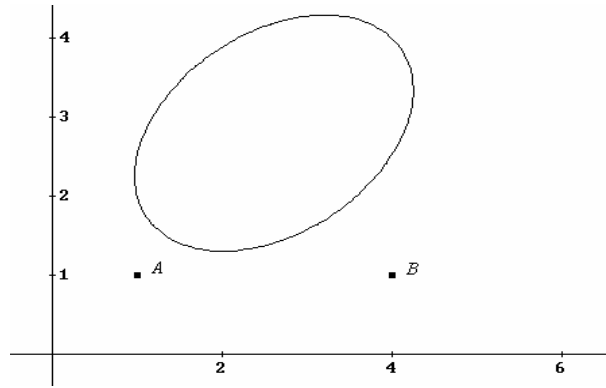
a)  $f(x, y) = 2x^2 - 15x + 3y^2 - 21y$ ;  $A = (3, 1)$ ;  $B = (7, 2)$



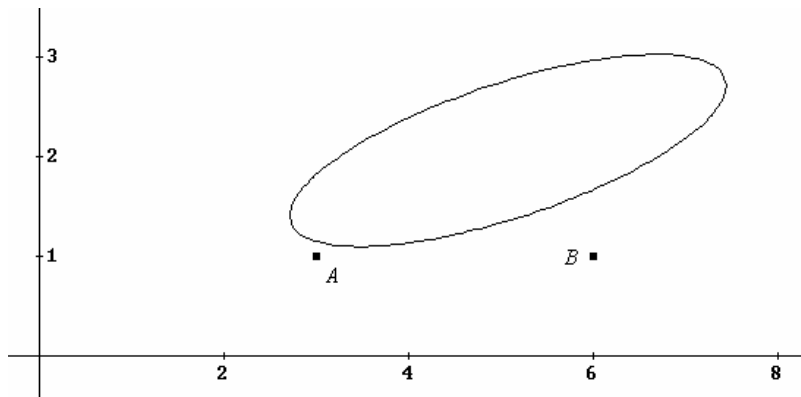
b)  $f(x, y) = 2x^2 - 15x + 30y^2 - 63y$ ;  $A = (2, 2)$ ;  $B = (8, 2)$



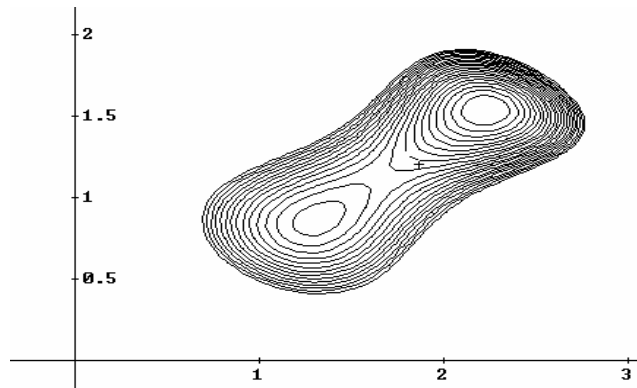
c)  $f(x, y) = 5x^2 - 4xy + 6y^2 - 15x - 23y$ ;  $A = (1, 1)$ ;  $B = (4, 1)$



d)  $f(x, y) = 3x^2 - 10xy + 18y^2 - 10x - 23y$ ;  $A = (3, 1)$ ;  $B = (6, 1)$



2. La función  $f(x, y) = x^2 + \sin(xy) + 2y^2 - 3x - 4y$  no es unimodal. En la figura se observan algunas curvas de nivel que muestran que en esta región ella presenta dos puntos de mínimo. Seleccione adecuadamente el punto inicial  $x_0$  para determinar ambos puntos de mínimo con tres cifras decimales exactas.



3. Halle la recta de mejor ajuste a los cuatro puntos (2, 3), (3, 4), 4, 6) y (5, 5) como un problema de optimización. Obtenga los coeficientes con tres cifras decimales exactas. Compruebe sus resultados utilizando el sistema normal de ecuaciones estudiado en el capítulo 4.

4. Calcule la distancia mínima entre las curvas  $y = \sin x$  y  $y = x^2 - 6x + 10$ . Obtenga el resultado con tres cifras decimales exactas.

5. Halle el plano

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1 \quad (a, b \text{ y } c \text{ son positivos})$$

que contiene al punto  $(2, 1; 0, 9; 1, 1)$  y que determina con los planos coordenados una pirámide de base triangular del menor volumen posible. Obtenga los resultados con error menor que 0,001.

6. A continuación se muestra un algoritmo para hallar un punto de máximo de la función  $f$  mediante el método de búsqueda por coordenadas. Este algoritmo posee un serio problema por el cual no funciona. Analice cual es este problema y repare el algoritmo.

**repeat**

$\lambda_{\max} := 0$  {Esta variable almacenará la mayor de las distancias recorridas en la secuencia de búsquedas unidimensionales}

**for**  $i = 1$  **to**  $n$

$\mathbf{v} := \mathbf{e}_i$  { $i$ -simo vector canónico: un 1 en la  $i$ -sima coordenada y 0 en las otras}

$\lambda := \text{MaxUnidim}(f, \mathbf{X}_{i-1}, \mathbf{V}, \varepsilon)$

$\mathbf{x}_i = \mathbf{x}_{i-1} + \lambda \mathbf{v}$

**if**  $\lambda > \lambda_{\max}$  **then**  $\lambda_{\max} := \lambda$

**end**

$\mathbf{x}_0 = \mathbf{x}_n$

**until**  $\lambda_{\max} < \varepsilon$

El punto de máximo es  $\mathbf{x}_n$

Terminar

## 6.6 El método del gradiente

Cuando se conocen las derivadas parciales de  $f(\mathbf{x})$  respecto a las variables  $u_1, u_2, \dots, u_n$  es posible seleccionar la dirección en el espacio  $R^n$  en que  $f(\mathbf{x})$  aumenta con mayor rapidez. Como se conoce de los cursos de Cálculo, esta dirección viene dada en cada punto  $\mathbf{x}$  por el vector gradiente que corresponde a dicho punto:

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial u_1} \\ \frac{\partial f}{\partial u_2} \\ \vdots \\ \frac{\partial f}{\partial u_n} \end{bmatrix}$$

En la figura 1 se muestran las curvas de nivel de una función  $f(\mathbf{x})$  en  $R^2$  y vectores gradiente en varios puntos. Recuérdese que el vector gradiente en cualquier punto es normal a la curva de nivel que pasa por dicho punto.

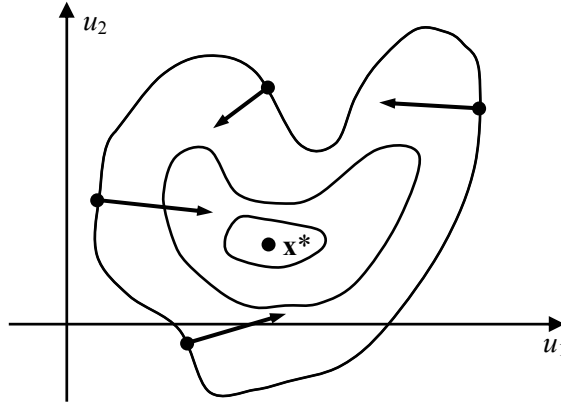


Figura 1

Es evidente que, a menos que las curvas de nivel sean circunferencias concéntricas, la dirección del gradiente en un punto no señala, en general, al punto  $\mathbf{x}^*$  de óptimo.

El método del gradiente, también llamado “steepest ascent” (o “steepest descent”, si se utiliza para minimizar) fue introducido por Cauchy en 1847 y se basa en:

Realizar una búsqueda unidimensional en la dirección de  $\nabla f|_{\mathbf{x}_0}$  hasta encontrar el punto de máximo  $\mathbf{x}_1$  en esa trayectoria.

Realizar una búsqueda unidimensional en la dirección de  $\nabla f|_{\mathbf{x}_1}$  hasta encontrar el punto de máximo  $\mathbf{x}_2$  en esa trayectoria.

Y así sucesivamente hasta que los puntos hallados difieran muy poco entre sí, o el gradiente tome valores muy pequeños o alguna otra condición que indique la proximidad de  $\mathbf{x}^*$ . Obsérvese que se ha supuesto que  $f(\mathbf{x})$  es linealmente unimodal, de manera que en cada trayectoria lineal el problema unidimensional resultante tenga solución. El método puede resumirse en el algoritmo que sigue.

### ***Algoritmo en pseudo código***

El algoritmo que sigue permite hallar el punto de máximo de la función linealmente unimodal  $f(\mathbf{x})$  mediante el método del gradiente. Se supone que  $f$  es diferenciable, de modo que posee gradiente en cada punto. El algoritmo utiliza como datos: el número  $n$  de variables independientes de la función a maximizar, la función  $f(\mathbf{x})$ , las  $n$  derivadas parciales de  $f(\mathbf{x})$ , el vector  $\mathbf{x}_0$  donde se comenzará el proceso de búsqueda y la tolerancia  $\varepsilon$  con que se desea hallar el punto de máximo. El algoritmo ofrece como resultado una aproximación del punto de máximo  $\mathbf{x}^*$  calculado con una tolerancia  $\varepsilon$ .

```

 $i := 0$ 
repeat
     $\mathbf{v} := \nabla f|_{\mathbf{x}_i}$ 
    Normalizar  $\mathbf{v}$ 
     $\lambda := \text{MaxUnidim}(f, \mathbf{x}_i, \mathbf{v}, \varepsilon)$ 
     $\mathbf{x}_{i+1} := \mathbf{x}_i + \lambda \mathbf{v}$ 
     $i := i + 1$ 
until  $\lambda < \varepsilon$ 

```

El método del gradiente posee algunos inconvenientes que deben señalarse. El primero, es la necesidad de conocer las derivadas parciales de  $f(\mathbf{x})$ , lo cual lo limita al caso en que se conoce la expresión analítica de  $f$  y sus derivadas pueden ser calculadas. El método es muy sensible a la geometría de las superficies de nivel. Como se aprecia de las figuras 2 y 3, cuando las curvas de nivel son elipses alargadas, el método posee una lenta convergencia; en cambio, con curvas de nivel circulares y concéntricas, basta una iteración para encontrar el punto de óptimo  $\mathbf{x}^*$ .

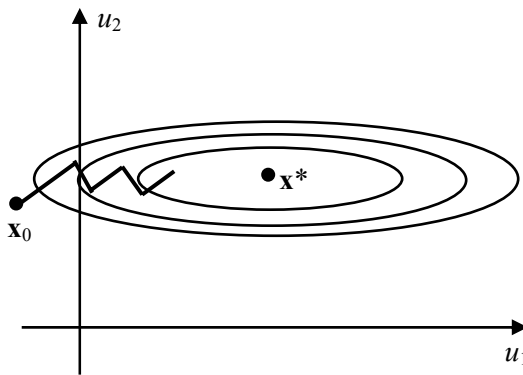


Figura 2

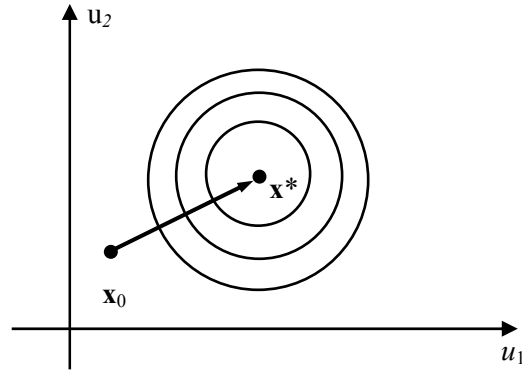


Figura 3

Esta característica trae como consecuencia que un simple cambio de escala, el cual puede ocurrir por utilizar diferentes unidades de medida en una de las variables, puede cambiar sustancialmente la rapidez de convergencia del método.

A diferencia de lo que ocurre en la búsqueda por coordenadas, en el método del gradiente resulta determinante la selección del punto inicial  $\mathbf{x}_0$ . Véase en la figura 4 cómo la rapidez de convergencia puede diferir para distintas selecciones de  $\mathbf{x}_0$ .

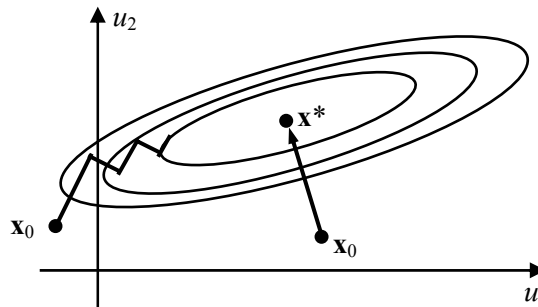


Figura 4

### Ejemplo 1

Como ya se dijo en el ejemplo 6 de la sección 6.4, la función cuadrática

$$f(\mathbf{x}) = 100 - 3u_1^2 - 4u_2^2 + 5u_1u_2 + 2u_1$$

posee un punto de máximo

$$\mathbf{x}^* = \begin{bmatrix} \frac{16}{23} \\ \frac{10}{23} \end{bmatrix} = \begin{bmatrix} 0,69565... \\ 0,43478... \end{bmatrix}$$

Determine  $\mathbf{X}^*$  por el método gradiente con precisión 0,0005 para:

$$\text{a) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{b) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{c) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{d) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

y señale en cada caso la cantidad de búsquedas unidimensionales requeridas.

Solución:

Mediante un programa confeccionado según el pseudo código anterior se obtuvo la solución de cada uno de los incisos:

$$\text{a) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,6951 \\ 0,4342 \end{bmatrix} \quad \text{en 21 búsquedas unidimensionales.}$$

$$\text{b) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,6956 \\ 0,4347 \end{bmatrix} \quad \text{en 3 búsquedas unidimensionales.}$$

$$\text{c) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,6955 \\ 0,4346 \end{bmatrix} \quad \text{en 5 búsquedas unidimensionales.}$$

$$\text{d) } \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,6961 \\ 0,4351 \end{bmatrix} \quad \text{en 13 búsquedas unidimensionales.}$$

■

Nótese cómo varía la rapidez de convergencia de acuerdo con la posición de  $\mathbf{x}_0$ , siendo en muchos casos inferior a la búsqueda por coordenadas. Como se observa en el ejemplo que sigue, este comportamiento persiste aun en casos en que no hay interacción entre las variables (ausencia de términos rectangulares) en los cuales la búsqueda por coordenadas tiene una magnífica eficiencia.

### Ejemplo 2

La función cuadrática:  $f(\mathbf{x}) = 100 - 4u_1^2 - 25u_2^2 + 2u_1 + 3u_2$

tiene curvas de nivel que son elipses con sus ejes horizontal y vertical respectivamente. Su punto de máximo es:

$$\mathbf{x}^* = \begin{bmatrix} 0,25 \\ 0,06 \end{bmatrix}$$

Utilice el método del gradiente para hallarlo con precisión de 0,0005 usando diferentes aproximaciones iniciales:

$$\text{a) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0,4 \end{bmatrix} \quad \text{b) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{c) } \mathbf{x}_0 = \begin{bmatrix} 0,3 \\ 0 \end{bmatrix}$$

y determine el número de iteraciones en cada caso.

Solución:

Mediante el mismo programa del ejemplo anterior fueron obtenidos los siguientes resultados:

$$\text{a) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0,4 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,2499 \\ 0,0600 \end{bmatrix} \quad \text{en 4 búsquedas unidimensionales.}$$

$$\text{b) } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,2494 \\ 0,0601 \end{bmatrix} \quad \text{en 17 búsquedas unidimensionales.}$$

$$\text{c) } \mathbf{x}_0 = \begin{bmatrix} 0,3 \\ 0 \end{bmatrix} \quad \mathbf{x}^* = \begin{bmatrix} 0,2502 \\ 0,0600 \end{bmatrix} \quad \text{en 5 búsquedas unidimensionales.}$$

■

### ***Ejemplo 3***

Halle el punto de mínimo de la función de Rosembrook (vea el ejemplo 3 de la sección 6.5)

$$f(\mathbf{x}) = 100(u_2 - u_1^2)^2 + (1 - u_1)^2$$

Compare con el número de iteraciones obtenido con el algoritmo de búsqueda por coordenadas.

Solución:

$$\text{Tomando } \mathbf{x}_0 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ se obtiene } \mathbf{x}^* = \begin{bmatrix} 0,9202 \\ 0,8468 \end{bmatrix} \text{ después de 1000 búsquedas unidimensionales.}$$

■

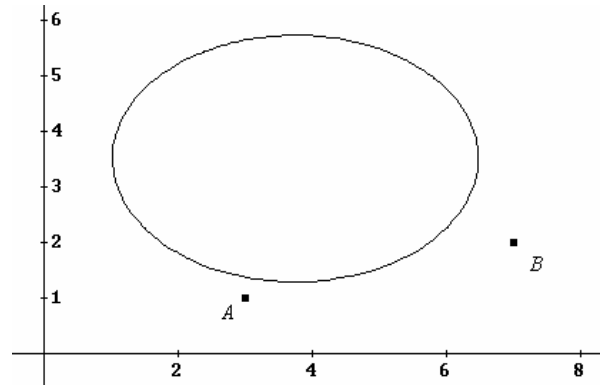
Aunque el resultado es algo menos lento que la búsqueda por coordenadas, el método se comporta con una convergencia sumamente lenta en esta función.

## Ejercicios

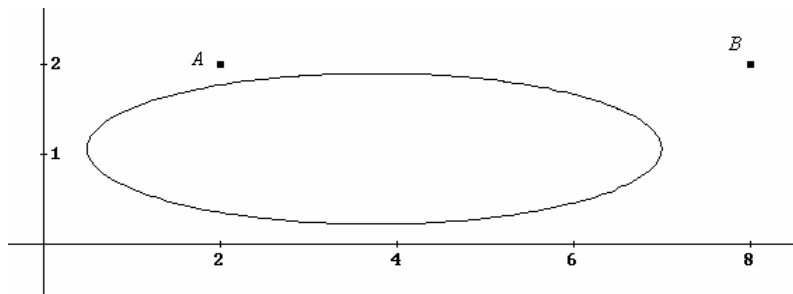
Algunos de los ejercicios que siguen son muy laboriosos para realizar los cálculos a mano. Se supone que posea programas computacionales, preferiblemente confeccionados por usted, para ayudarle en el trabajo. De no ser así, en los ejercicios que requieran mucho trabajo manual, determine los resultados con menos cifras exactas que las exigidas.

1. En cada una de las funciones cuadráticas que siguen se muestra una curva de nivel, cuya forma puede servir para predecir como se comportará el método del gradiente al determinar su punto de extremo. Estos problemas aparecieron también en los ejercicios de la sección anterior. En cada caso, aplique el método del gradiente para hallar el punto de extremo utilizando como punto de partida los puntos  $A$  y  $B$  que se dan en la propia figura. Obtenga la solución con tres cifras decimales exactas. Compare con la solución obtenida utilizando el método de búsqueda por coordenadas ¿Puede dar alguna conclusión general como resultado de este ejercicio?

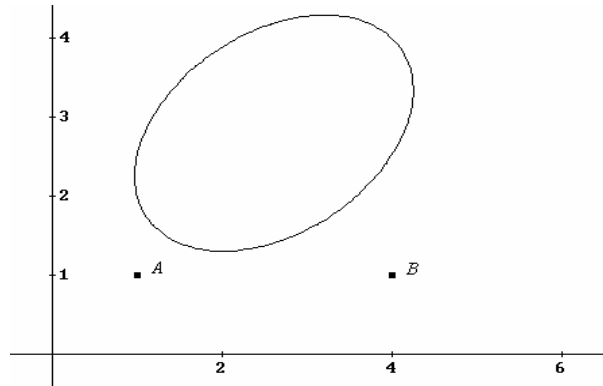
a)  $f(x, y) = 2x^2 - 15x + 3y^2 - 21y$ ;  $A = (3, 1)$ ;  $B = (7, 2)$



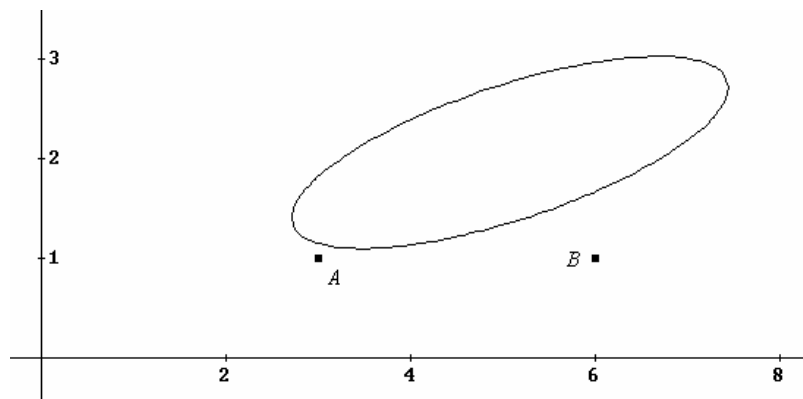
b)  $f(x, y) = 2x^2 - 15x + 30y^2 - 63y$ ;  $A = (2, 2)$ ;  $B = (8, 2)$



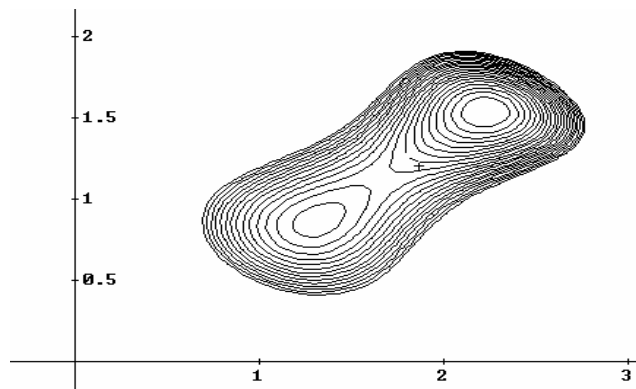
c)  $f(x, y) = 5x^2 - 4xy + 6y^2 - 15x - 23y$ ;  $A = (1, 1)$ ;  $B = (4, 1)$



d)  $f(x, y) = 3x^2 - 10xy + 18y^2 - 10x - 23y$ ;  $A = (3, 1)$ ;  $B = (6, 1)$



2. La función  $f(x, y) = x^2 + \sin(xy) + 2y^2 - 3x - 4y$  no es unimodal. En la figura se observan algunas curvas de nivel que muestran que en esta región ella presenta dos puntos de mínimo. Seleccione adecuadamente el punto inicial  $\mathbf{x}_0$  para determinar ambos puntos de mínimo mediante el método del gradiente. Obtenga los resultados con tres cifras decimales exactas.

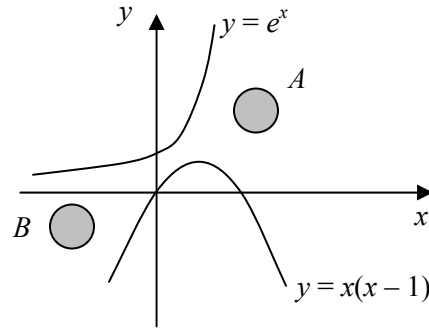


3. Ajuste un modelo no lineal del tipo  $g(x) = ae^{bx}$  a los datos mostrados en la siguiente tabla:

|     |      |      |      |      |      |
|-----|------|------|------|------|------|
| $x$ | 1    | 2    | 3    | 4    | 5    |
| $y$ | 1,18 | 1,26 | 1,36 | 1,46 | 1,57 |

obtenga la solución con tres cifras decimales exactas.

- 4 Halle el diámetro del mayor círculo que puede pasar de la posición A hasta la posición B entre las dos gráficas sin cortar a ninguna de las dos (solamente rozarlas). Obtenga el resultado con tres cifras decimales exactas.



5. En el método del gradiente se puede detectar la proximidad de un punto de extremo por que la norma del vector gradiente se aproxima a cero. Elabore un algoritmo en pseudo código que utilice este criterio para detener el proceso iterativo.

## 6.7 El método de Powell

### *Direcciones conjugadas en una función cuadrática*

Algunos algoritmos de optimización multidimensional están basados en la idea de obtener técnicas que permitan hallar eficientemente el óptimo de una función cuadrática, con la esperanza de que, en el caso de funciones no cuadráticas, el método mantenga su eficiencia si se parte de un punto cercano al óptimo. Varios de los métodos que han logrado mayor éxito utilizan el concepto de direcciones conjugadas, que es una generalización del concepto de direcciones perpendiculares.

Definición:

Sea  $\mathbf{A}$  una matriz cuadrada de orden  $n$  positiva (negativa) definida. Sean  $\mathbf{d}_i$  y  $\mathbf{d}_j$  vectores de  $R^n$  que señalan dos direcciones en ese espacio. Las direcciones  $\mathbf{d}_i$  y  $\mathbf{d}_j$  se llaman conjugadas respecto a  $\mathbf{A}$ , o también  $\mathbf{A}$ -ortogonales, si

$$\mathbf{d}_i^T \mathbf{A} \mathbf{d}_j = 0$$

#### **Ejemplo 1**

Sea la matriz 
$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}$$

la cual es positiva definida (compruébese). Sea  $\mathbf{d}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ . Halle un vector  $\mathbf{d}_2$  conjugado de  $\mathbf{d}_1$  respecto a  $\mathbf{A}$ .

Solución:

Sea un vector

$$\mathbf{d}_2 = \begin{bmatrix} a \\ b \end{bmatrix}$$

Se debe cumplir que

$$\mathbf{d}_1^T \mathbf{A} \mathbf{d}_2 = 0$$

es decir:

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0$$

$$\begin{bmatrix} 5 & 6 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0$$

$$5a + 6b = 0$$

Tomando una solución arbitraria:

$$a = 6; \quad b = -5$$

Por tanto,

$$\mathbf{d}_2 = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$$

es conjugado de  $\mathbf{d}_1$

### ***Interpretación geométrica***

Considérese el caso  $n = 2$ , de modo que la función:

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$$

con  $\mathbf{A}$  positiva definida representa un paraboloide elíptico con vértice (punto de mínimo) en  $\mathbf{x} = \mathbf{0}$ . Las curvas de nivel

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = k$$

con  $k > 0$  son elipses con centro en  $\mathbf{x} = \mathbf{0}$ , como se muestra en la figura 1.

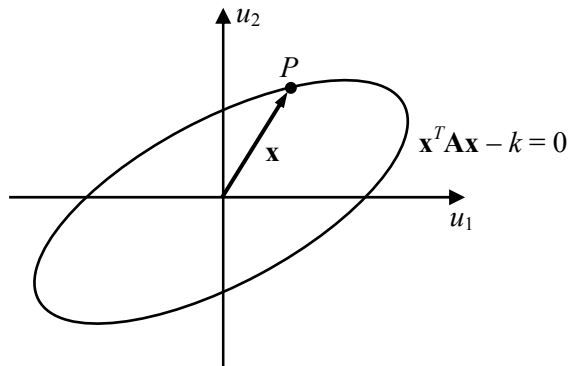


Figura 1

Sea  $\mathbf{x}$  la dirección correspondiente a un punto  $P$  (vector que une el centro de la elipse con el punto  $P$ ). Hallemos el gradiente de  $\mathbf{x}^T \mathbf{A} \mathbf{x} - k$  que, como se sabe, es un vector normal a la elipse en  $P$ :

$$\begin{aligned} \nabla(\mathbf{x}^T \mathbf{A} \mathbf{x} - k) &= \nabla(a_{11}u_1^2 + a_{12}u_1u_2 + a_{21}u_2u_1 + a_{22}u_2^2 - k) \\ &= \begin{bmatrix} 2a_{11}u_1 + 2a_{12}u_2 \\ 2a_{21}u_1 + 2a_{22}u_2 \end{bmatrix} = 2 \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 2 \mathbf{A} \mathbf{x} \end{aligned}$$

Es decir, el vector  $\mathbf{A} \mathbf{x}$  es normal a la elipse en el punto  $P$ . Sea ahora  $\mathbf{y}$  un vector en una dirección conjugada a  $\mathbf{x}$ , esto es:

$$\mathbf{y}^T \mathbf{A} \mathbf{x} = 0$$

lo cual significa que

$$\mathbf{y} \cdot (\mathbf{A} \mathbf{x}) = 0$$

esto es, que  $\mathbf{y}$  es ortogonal a  $\mathbf{A} \mathbf{x}$  y, por tanto, tangente a la elipse en  $P$ . Esto prueba que, si  $\mathbf{x}$  es el vector radial del punto  $P$  de la elipse

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = k$$

entonces su dirección conjugada  $\mathbf{y}$  es tangente a la elipse en el mismo punto  $P$ , tal como se muestra en la figura 2.

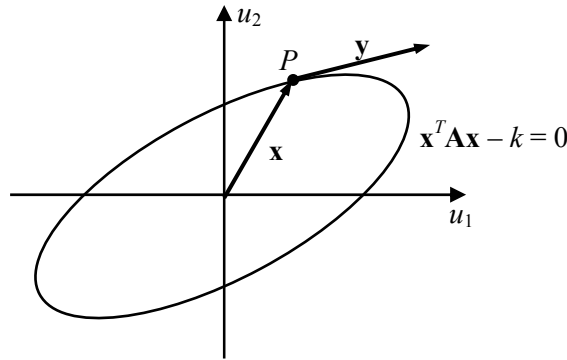


Figura 2

### ***Propiedades fundamentales de las direcciones conjugadas***

Propiedad 1:

Si  $\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_k$  son vectores de  $R^n$  no nulos conjugados entre sí respecto a la matriz  $\mathbf{A}$  positiva (negativa) definida, entonces el conjunto  $\{\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_k\}$  es linealmente independiente.

Demostración:

Supóngase que

$$\alpha_0 \mathbf{d}_0 + \alpha_1 \mathbf{d}_1 + \dots + \alpha_k \mathbf{d}_k = \mathbf{0}$$

Premultiplicando en ambos miembros por  $\mathbf{A}$ :

$$\alpha_0 \mathbf{A} \mathbf{d}_0 + \alpha_1 \mathbf{A} \mathbf{d}_1 + \dots + \alpha_k \mathbf{A} \mathbf{d}_k = \mathbf{0}$$

Si ahora se premultiplica por el transpuesto de cualquiera de los vectores  $\mathbf{d}_i$ :

$$\alpha_0 \mathbf{d}_i^T \mathbf{A} \mathbf{d}_0 + \alpha_1 \mathbf{d}_i^T \mathbf{A} \mathbf{d}_1 + \dots + \alpha_k \mathbf{d}_i^T \mathbf{A} \mathbf{d}_k = 0$$

y, como  $\mathbf{d}_i$  y  $\mathbf{d}_j$  son conjugados para  $i \neq j$ , resulta:

$$\alpha_i \mathbf{d}_i^T \mathbf{A} \mathbf{d}_i = 0$$

Sin embargo, por ser  $\mathbf{A}$  positiva o negativa definida,

$$\mathbf{d}_i^T \mathbf{A} \mathbf{d}_i \neq 0$$

así que  $\alpha_i = 0$ . Como  $\mathbf{d}_i$  es cualquiera de los vectores del conjunto, se ha probado que

$$\alpha_0 = \alpha_1 = \dots = \alpha_k = 0$$

y, por tanto, el conjunto  $\{\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_k\}$  es linealmente independiente.

Propiedad 2:

Sea la función cuadrática de  $n$  variables con punto de mínimo  $\mathbf{x}^*$

$$f(\mathbf{X}) = \frac{1}{2} \mathbf{X}^T \mathbf{A} \mathbf{X} - \mathbf{B}^T \mathbf{X}$$

y sean  $\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_{n-1}$  direcciones conjugadas entre sí respecto a la matriz  $\mathbf{A}$ . Como el conjunto  $\{\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_{n-1}\}$  es linealmente independiente y posee  $n$  vectores, es una base de  $R^n$ , así que  $\mathbf{x}^*$  puede expresarse como combinación de la base conjugada:

$$\mathbf{x}^* = \alpha_0 \mathbf{d}_0 + \alpha_1 \mathbf{d}_1 + \dots + \alpha_{n-1} \mathbf{d}_{n-1}$$

Los  $\alpha_i$  pueden determinarse fácilmente pues:

$$\mathbf{A} \mathbf{x}^* = \alpha_0 \mathbf{A} \mathbf{d}_0 + \alpha_1 \mathbf{A} \mathbf{d}_1 + \dots + \alpha_{n-1} \mathbf{A} \mathbf{d}_{n-1}$$

pero como  $\mathbf{A} \mathbf{x}^* = \mathbf{b}$ :

$$\mathbf{b} = \alpha_0 \mathbf{A} \mathbf{d}_0 + \alpha_1 \mathbf{A} \mathbf{d}_1 + \dots + \alpha_{n-1} \mathbf{A} \mathbf{d}_{n-1}$$

$$\mathbf{d}_i^T \mathbf{b} = \alpha_0 \mathbf{d}_i^T \mathbf{A} \mathbf{d}_0 + \alpha_1 \mathbf{d}_i^T \mathbf{A} \mathbf{d}_1 + \dots + \alpha_{n-1} \mathbf{d}_i^T \mathbf{A} \mathbf{d}_{n-1}$$

$$\mathbf{d}_i^T \mathbf{b} = \alpha_i \mathbf{d}_i^T \mathbf{A} \mathbf{d}_i$$

de donde:

$$\alpha_i = \frac{\mathbf{d}_i^T \mathbf{b}}{\mathbf{d}_i^T \mathbf{A} \mathbf{d}_i} \quad i = 0, 1, 2, \dots, n-1$$

■

Lo notable de la fórmula anterior es que ella permite hallar las coordenadas  $\alpha_0, \alpha_1, \dots, \alpha_{n-1}$  de  $\mathbf{x}$  respecto a una base sin tener que resolver un sistema lineal como  $\mathbf{A} \mathbf{x}^* = \mathbf{b}$ . Sin embargo, como establece la propiedad 3, estas coordenadas  $\alpha_i$  se pueden obtener por otra vía mucho más interesante.

Propiedad 3:

Sea 
$$f(\mathbf{X}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x}$$

una función cuadrática con punto de mínimo  $\mathbf{x}^*$  y  $\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_{n-1}$  direcciones conjugadas entre sí respecto a la matriz  $\mathbf{A}$ . Sea  $\mathbf{x}_0$  un punto arbitrario y:

$$\mathbf{x}^* - \mathbf{x}_0 = \alpha_0 \mathbf{d}_0 + \alpha_1 \mathbf{d}_1 + \dots + \alpha_{n-1} \mathbf{d}_{n-1}$$

de modo que:

$$\mathbf{x}^* = \mathbf{x}_0 + \alpha_0 \mathbf{d}_0 + \alpha_1 \mathbf{d}_1 + \dots + \alpha_{n-1} \mathbf{d}_{n-1}$$

Se define ahora el proceso iterativo:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k \quad k = 0, 1, 2, \dots, n-1$$

esto es:

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{d}_0$$

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha_1 \mathbf{d}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{d}_0 + \alpha_1 \mathbf{d}_1$$

$$\vdots$$

$$\mathbf{x}_n = \mathbf{x}_{n-1} + \alpha_{n-1} \mathbf{d}_{n-1} = \mathbf{x}_0 + \alpha_0 \mathbf{d}_0 + \alpha_1 \mathbf{d}_1 + \dots + \alpha_{n-1} \mathbf{d}_{n-1} = \mathbf{x}^*$$

Entonces, puede demostrarse que, para  $k = 0, 1, 2, \dots, n-1$ :

$\alpha_k$  es el valor de  $\lambda$  para el cual la función unidimensional  $f(\mathbf{x}_k + \lambda \mathbf{d}_k)$  toma su mínimo valor.

■

Esta propiedad permite ir buscando las coordenadas  $\alpha_0, \alpha_1, \dots, \alpha_{n-1}$  mediante sendos problemas de optimización unidimensional. En este proceso se van obteniendo los puntos  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ . Como se ve, el punto  $\mathbf{x}_n$  es ya el óptimo. Observe en la figura 3 cómo funciona el algoritmo para un caso donde  $n = 2$ .

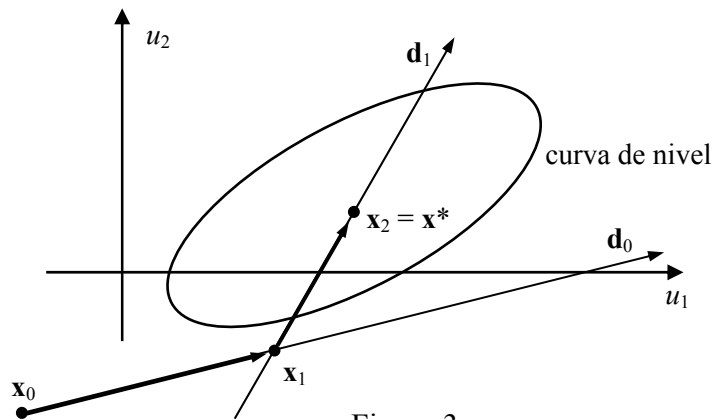


Figura 3

Se puede entonces resumir que:

Si  $f(\mathbf{x})$  es una función cuadrática con punto de extremo  $\mathbf{x}^*$   
 Si  $\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_{n-1}$  son direcciones conjugadas de su matriz  $\mathbf{A}$   
 Si  $\mathbf{x}_0$  es un punto cualquiera

Entonces el punto  $\mathbf{x}^*$  se halla mediante  $n$  búsquedas unidimensionales sucesivas a partir de  $\mathbf{x}_0$  en las direcciones conjugadas.

■

Podría pensarse que hallar un conjunto de  $n$  direcciones conjugadas es muy complicado pero, como se verá a continuación, hay formas muy eficientes de hacerlo.

### ***Método de las tangentes paralelas***

La idea geométrica del método puede verse muy claramente si  $f(\mathbf{x})$  es una función cuadrática de dos variables, como en la figura 4. En ese caso, las curvas de nivel son elipses concéntricas (el centro es  $\mathbf{x}^*$ ) y con la misma excentricidad.

Suponga que, partiendo de puntos diferentes  $P_1$  y  $P_2$  se trazan dos rectas paralelas, ambas en la dirección  $\mathbf{d}_1$ . Sean  $L_1$  y  $L_2$  esas rectas. Obviamente, el mínimo de  $f(\mathbf{x})$  a lo largo de  $L_1$  se obtiene en el punto  $\mathbf{x}_1$  donde  $L_1$  es tangente a la curva de nivel que pasa por  $\mathbf{x}_1$  y el mínimo sobre  $L_2$  estará en un punto  $\mathbf{x}_2$  donde  $L_2$  es tangente a la curva de nivel que pasa por  $\mathbf{x}_2$ .

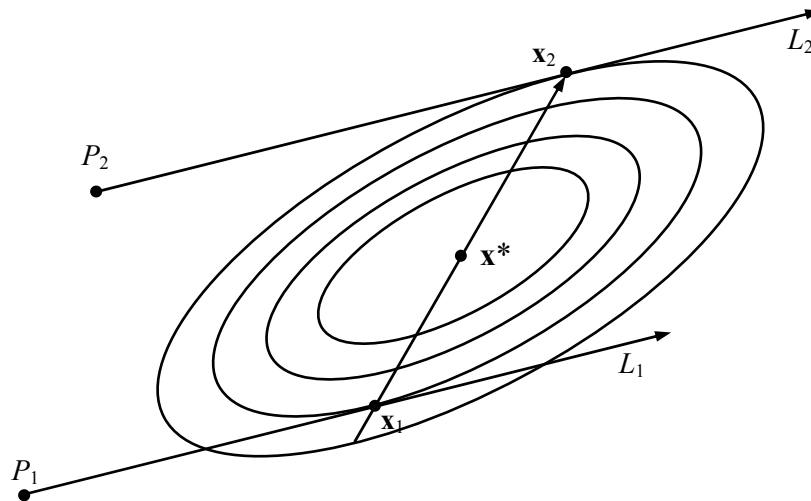


Figura 4

El segmento que une a  $\mathbf{x}_1$  y  $\mathbf{x}_2$  pasa por  $\mathbf{x}^*$ , o sea, la dirección del vector

$$\mathbf{d}_2 = \mathbf{x}_2 - \mathbf{x}_1$$

es radial. De acuerdo con la interpretación geométrica analizada, resulta que  $\mathbf{d}_2$  es una dirección conjugada de  $\mathbf{d}_1$ .

Lo anterior se puede establecer analíticamente.

Sea 
$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x}$$

una función cuadrática de  $n$  variables con punto de mínimo. Sea  $\mathbf{d}$  una dirección cualquiera en  $R^n$  y  $P_1$  y  $P_2$  dos puntos diferentes de ese espacio. Sean:

$\mathbf{x}_1$ : Punto de mínimo de  $f$  sobre la recta  $L_1$  que pasa por  $P_1$  con dirección  $\mathbf{d}$

$\mathbf{x}_2$ : Punto de mínimo de  $f$  sobre la recta  $L_2$  que pasa por  $P_2$  con dirección  $\mathbf{d}$

entonces el vector  $\mathbf{x}_2 - \mathbf{x}_1$  es conjugado de  $\mathbf{d}$ .

Demostración:

La ecuación de  $L_1$  es:  $\mathbf{x} = \mathbf{x}_1 + \lambda \mathbf{d}$

La ecuación de  $L_2$  es:  $\mathbf{x} = \mathbf{x}_2 + \lambda \mathbf{d}$

Sobre  $L_1$ : 
$$f(\mathbf{x}) = \frac{1}{2} (\mathbf{x}_1 + \lambda \mathbf{d})^T \mathbf{A} (\mathbf{x}_1 + \lambda \mathbf{d}) - \mathbf{b}^T (\mathbf{x}_1 + \lambda \mathbf{d})$$

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 + \lambda \mathbf{d}^T \mathbf{A} \mathbf{x}_1 + \frac{1}{2} \lambda^2 \mathbf{d}^T \mathbf{A} \mathbf{d} - \mathbf{b}^T \mathbf{x}_1 - \lambda \mathbf{b}^T \mathbf{d}$$

Como esta función toma su mínimo en  $\mathbf{x}_1$  (o sea, cuando  $\lambda = 0$ ), debe cumplirse que su derivada respecto a  $\lambda$  evaluada en  $\lambda = 0$  se anule, es decir:

$$\mathbf{d}^T \mathbf{A} \mathbf{x}_1 + \lambda \mathbf{d}^T \mathbf{A} \mathbf{d} - \mathbf{b}^T \mathbf{d} \Big|_{\lambda=0} = 0$$

esto es: 
$$\mathbf{d}^T \mathbf{A} \mathbf{x}_1 - \mathbf{b}^T \mathbf{d} = 0$$

Realizando el mismo razonamiento para la recta  $L_2$  se llega a:

$$\mathbf{d}^T \mathbf{A} \mathbf{x}_2 - \mathbf{b}^T \mathbf{d} = 0$$

Restando las dos últimas ecuaciones:

$$\mathbf{d}^T \mathbf{A} (\mathbf{x}_2 - \mathbf{x}_1) = 0$$

y esto significa que  $\mathbf{x}_2 - \mathbf{x}_1$  y  $\mathbf{d}$  son direcciones conjugadas, como se deseaba probar.

### ***El método de Powell***

En las páginas anteriores se han establecido dos importantes propiedades de las direcciones conjugadas de una función cuadrática de  $n$  variables:

- Si se cuenta con un conjunto de  $n$  direcciones conjugadas de la función, entonces realizando  $n$  búsquedas unidimensionales sucesivas en dichas direcciones, se llega al óptimo de la función cuadrática.
- Si se realizan dos búsquedas unidimensionales en una misma dirección a partir de puntos distintos, el vector que une los óptimos alcanzados determina una dirección conjugada de la anterior.

En el método de Powell se combinan estos dos resultados del siguiente modo. Partiendo de un conjunto de  $n$  direcciones de búsqueda  $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_n$  (usualmente, las direcciones de los ejes coordenados) y de un punto inicial  $\mathbf{x}_0$ , se realizan  $n$  procesos de búsqueda unidimensional, obteniéndose sucesivamente:

$\mathbf{x}_1$ : optimizando  $f$  en la dirección  $\mathbf{d}_1$  desde  $\mathbf{x}_0$   
 $\mathbf{x}_2$ : optimizando  $f$  en la dirección  $\mathbf{d}_2$  desde  $\mathbf{x}_1$   
 $\vdots$   
 $\mathbf{x}_n$ : optimizando  $f$  en la dirección  $\mathbf{d}_n$  desde  $\mathbf{x}_{n-1}$

Al llegar al punto  $\mathbf{x}_n$  se realiza una nueva búsqueda unidimensional en la dirección del vector  $\mathbf{x}_n - \mathbf{x}_0$ . La norma de este vector sirve como indicador de la convergencia del proceso. Esta búsqueda suele llamarse un “paso de aceleración” por cuanto se realiza en una dirección que es el resumen de  $n$  búsquedas previas y recoge la “experiencia” obtenida en esos procesos. El vector alcanzado en este paso de aceleración se toma como un nuevo  $\mathbf{x}_0$  para comenzar otra vez todo el proceso pero antes, se hace un cambio en el conjunto de direcciones  $\{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_n\}$ . La dirección  $\mathbf{d}_1$  se elimina, tomándose:

$$\begin{aligned}\mathbf{d}_1 &:= \mathbf{d}_2 \\ \mathbf{d}_2 &:= \mathbf{d}_3 \\ &\vdots \\ \mathbf{d}_{n-1} &:= \mathbf{d}_n\end{aligned}$$

y se introduce una nueva dirección  $\mathbf{d}_n$  según el vector  $\mathbf{x}_n - \mathbf{x}_0$ , previamente normalizado. Nótese que el nuevo  $\mathbf{x}_0$  se obtuvo buscando en la dirección de este  $\mathbf{d}_n$  a partir de  $\mathbf{x}_n$  en el paso de aceleración. En la nueva iteración, el nuevo  $\mathbf{x}_n$  se obtendrá buscando a partir de  $\mathbf{x}_{n-1}$  en la dirección de  $\mathbf{d}_n$ . Como  $\mathbf{x}_0$  y  $\mathbf{x}_n$  se han obtenido con búsquedas según rectas paralelas, el vector  $\mathbf{x}_n - \mathbf{x}_0$  de la segunda iteración será conjugado de  $\mathbf{d}_n$ .

De este modo, el conjunto de direcciones  $\{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_n\}$  se va convirtiendo paulatinamente en un conjunto de direcciones conjugadas, eliminando en cada iteración el primer vector de ese conjunto y añadiendo uno nuevo por el final que es conjugado del último. Después de  $n$  iteraciones, el conjunto está formado por  $n$  direcciones conjugadas entre sí y, por tanto, en la próxima iteración se obtiene el punto de óptimo.

Claro que todo esto es así si  $f$  es una función cuadrática. Si no lo es, ya el concepto mismo de dirección conjugada carece de sentido pero, como se supone que en la región próxima a  $\mathbf{X}^*$  la función  $f$  se comporta similarmente a una función cuadrática, las direcciones que se van obteniendo, van formando una adecuada colección que, además, se va actualizando continuamente.

La figura 5 ilustra la forma en que funciona el algoritmo para el caso de una función cuadrática de dos variables. En este caso cada iteración utiliza tres búsquedas unidimensionales: dos en las direcciones  $\mathbf{d}_1$  y  $\mathbf{d}_2$  y el paso de aceleración, que sirve además para introducir una nueva dirección de búsqueda. Después de la segunda iteración, ya las direcciones  $\mathbf{d}_1$  y  $\mathbf{d}_2$  son conjugadas, de modo que en la segunda iteración (es decir, en la sexta búsqueda unidimensional) se alcanza el punto de máximo  $\mathbf{x}^*$ .

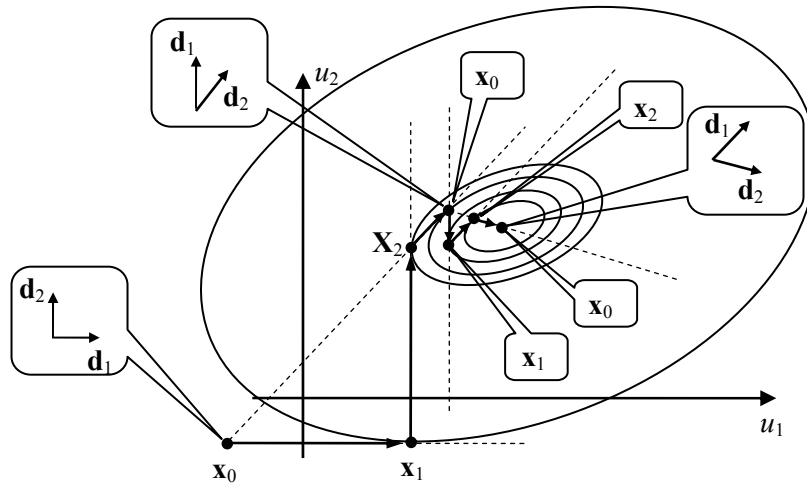


Figura 5

### Algoritmo en pseudo código

En el siguiente algoritmo se describe en detalle el método de Powell para hallar el punto de máximo de una función  $f(\mathbf{x})$  de  $n$  variables independientes. Se supone que  $f$  es linealmente unimodal. Para que el algoritmo resulte más simple, el procedimiento  $\text{MaxBidirec}(f, \mathbf{x}, \mathbf{v})$  empleado aquí realiza la búsqueda a partir de  $\mathbf{x}$  en la dirección del vector  $\mathbf{v}$  (en ambos sentidos) y devuelve como resultado el punto de óptimo obtenido. El algoritmo utiliza como datos, el número  $n$  de variables de la función a optimizar, la función  $f$ , el vector  $\mathbf{x}_0$ , que indica el punto donde se comenzará todo el proceso, y la tolerancia  $\varepsilon$  con que se obtendrá el punto de máximo  $\mathbf{x}^*$ .

```

for  $i = 1$  to  $n$ 
     $\mathbf{d}_i := \mathbf{e}_i$  {Las direcciones de búsqueda se toman inicialmente en la dirección de
                  los vectores canónicos}
end
repeat
    for  $i = 1$  to  $n$ 
         $\mathbf{x}_i := \text{MaxBidirec}(f, \mathbf{x}_{i-1}, \mathbf{d}_i)$ 
    end
    for  $i = 1$  to  $n - 1$ 
         $\mathbf{d}_i := \mathbf{d}_{i+1}$ 
    end
     $\mathbf{d}_n := \mathbf{x}_n - \mathbf{x}_0$ 
     $\text{dist} := \text{norma de } \mathbf{d}_n$ 
     $\mathbf{d}_n := \frac{\mathbf{d}_n}{\text{dist}}$ 
     $\mathbf{x}_0 := \text{MaxBidirec}(f, \mathbf{x}_n, \mathbf{d}_n)$ 
until  $\text{dist} < \varepsilon$ 

```

### Ejemplo 2

Halle el punto de máximo de la función cuadrática

$$f(\mathbf{x}) = 100 - 3u_1^2 - 4u_2^2 + 5u_1u_2 + 2u_1$$

por el método de Powell con precisión de tres cifras decimales exactas tomando

$$\mathbf{x}_0 = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

Solución:

La solución se llevará a cabo con detalle para ayudar a comprender el algoritmo.

Iteración 1:

$$\mathbf{x}_0 = \begin{bmatrix} 5 \\ 5 \end{bmatrix} \quad \mathbf{d}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \mathbf{d}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\text{Búsqueda en dirección } \mathbf{d}_1: \quad \mathbf{x}_1 = \begin{bmatrix} 4,500040 \\ 5,000000 \end{bmatrix}$$

$$\text{Búsqueda en dirección } \mathbf{d}_2: \quad \mathbf{x}_2 = \begin{bmatrix} 4,500040 \\ 2,812548 \end{bmatrix}$$

$$\text{Dirección } \mathbf{x}_2 - \mathbf{x}_0 \text{ (normalizada)} = \begin{bmatrix} -0,222813 \\ -0,974861 \end{bmatrix}$$

$$\text{Búsqueda en dirección } \mathbf{x}_2 - \mathbf{x}_0: \quad \mathbf{x}_0 = \begin{bmatrix} 4,405260 \\ 2,397859 \end{bmatrix} \quad \lambda = 0,425383$$

$$\text{Iteración 2:} \quad \mathbf{x}_0 = \begin{bmatrix} 4,405260 \\ 2,397859 \end{bmatrix} \quad \mathbf{d}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \mathbf{d}_2 = \begin{bmatrix} -0,222813 \\ -0,974861 \end{bmatrix}$$

$$\text{Búsqueda en dirección } \mathbf{d}_1: \quad \mathbf{x}_1 = \begin{bmatrix} 4,405260 \\ 2,753245 \end{bmatrix}$$

$$\text{Búsqueda en dirección } \mathbf{d}_2: \quad \mathbf{x}_2 = \begin{bmatrix} 4,312821 \\ 2,348980 \end{bmatrix}$$

$$\text{Dirección } \mathbf{x}_2 - \mathbf{x}_0 \text{ (normalizada)} = \begin{bmatrix} -0,884022 \\ -0,467445 \end{bmatrix}$$

$$\text{Búsqueda en dirección } \mathbf{x}_2 - \mathbf{x}_0: \quad \mathbf{x}_0 = \begin{bmatrix} 0,695992 \\ 0,435660 \end{bmatrix} \quad \lambda = 4,091764$$

$$\text{Iteración 3: } \mathbf{x}_0 = \begin{bmatrix} 0,695992 \\ 0,435660 \end{bmatrix} \quad \mathbf{d}_1 = \begin{bmatrix} -0,222813 \\ -0,974861 \end{bmatrix} \quad \mathbf{d}_2 = \begin{bmatrix} -0,884022 \\ -0,467445 \end{bmatrix}$$

$$\text{Búsqueda en dirección } \mathbf{d}_1: \quad \mathbf{x}_1 = \begin{bmatrix} 0,695824 \\ 0,434922 \end{bmatrix}$$

$$\text{Búsqueda en dirección } \mathbf{d}_2: \quad \mathbf{x}_2 = \begin{bmatrix} 0,695761 \\ 0,434889 \end{bmatrix}$$

$$\text{Dirección } \mathbf{x}_2 - \mathbf{x}_0 \text{ (normalizada)} = \begin{bmatrix} -0,287006 \\ -0,957929 \end{bmatrix}$$

$$\text{Búsqueda en dirección } \mathbf{x}_2 - \mathbf{x}_0: \quad \mathbf{x}_0 = \begin{bmatrix} 0,695754 \\ 0,434865 \end{bmatrix} \quad \lambda = 0.000025 < \varepsilon$$

$$\text{Respuesta: Con cuatro cifras decimales exactas } \mathbf{x}^* = \begin{bmatrix} 0,695754 \\ 0,434865 \end{bmatrix}$$

Nótese como, sorpresivamente, en la sexta búsqueda (final de la segunda iteración) ya se alcanzó  $\mathbf{x}^*$ . La razón es obvia; como esta es una función cuadrática de dos variables, después de dos iteraciones ya las direcciones de búsqueda forman una base conjugada y el punto de óptimo se encuentra con toda precisión. El método requiere algunas iteraciones más solamente para verificar la satisfacción de algún criterio de parada.

### ***Ejemplo 3***

Pruebe el método de Powell con la función de Rosembrook. Compare el resultado con los obtenidos usando los métodos de búsqueda por coordenadas y del gradiente.

Solución:

$$\text{Tomando } \mathbf{x}_0 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ y } f(\mathbf{x}) = 100(u_2 - u_1^2)^2 + (1 - u_1)^2$$

se obtiene  $\mathbf{x}^*$  con tres cifras decimales exactas:

$$\mathbf{x}^* = \begin{bmatrix} 1.000002 \\ 1.000004 \end{bmatrix}$$

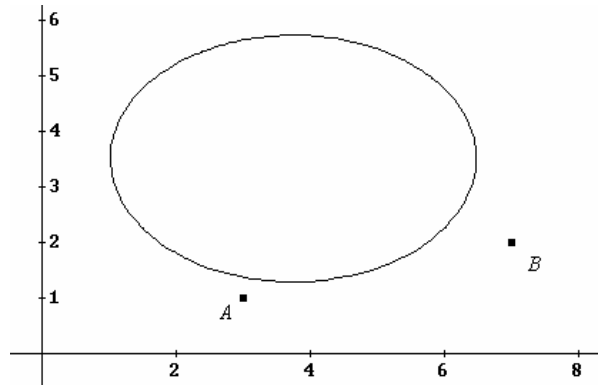
después de 29 búsquedas unidimensionales. En el método de búsqueda por coordenadas, después de 2000 búsquedas unidimensionales, aún no se había obtenido una milésima de precisión. En el método del gradiente, después de 1000 búsquedas unidimensionales, todavía el error es mayor de 0,15.

## Ejercicios

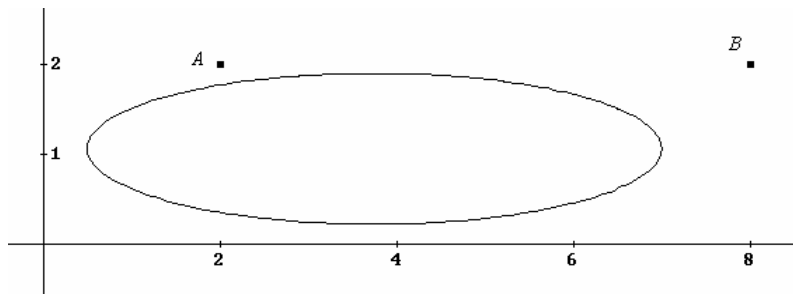
Algunos de los ejercicios que siguen son muy laboriosos para realizar los cálculos a mano. Se supone que posea programas computacionales, preferiblemente confeccionados por usted, para ayudarle en el trabajo. De no ser así, en los ejercicios que requieran mucho trabajo manual, determine los resultados con menos cifras exactas que las exigidas.

1. En cada una de las funciones cuadráticas que siguen se muestra una curva de nivel, cuya forma puede servir para predecir como se comportará el método de Powell al determinar su punto de extremo. Estos problemas aparecieron también en los ejercicios de las dos secciones anteriores. En cada caso, aplique el método de Powell para hallar el punto de extremo utilizando como punto de partida los puntos  $A$  y  $B$  que se dan en la propia figura. Obtenga la solución con tres cifras decimales exactas. Compare con la solución obtenida utilizando los métodos de búsqueda por coordenadas y del gradiente ¿Puede dar alguna conclusión general como resultado de este ejercicio?

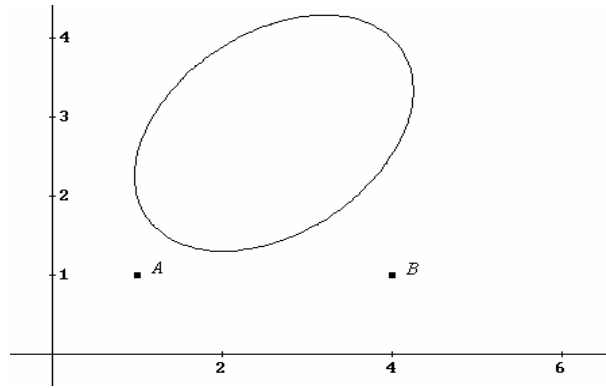
a)  $f(x, y) = 2x^2 - 15x + 3y^2 - 21y$ ;  $A = (3, 1)$ ;  $B = (7, 2)$



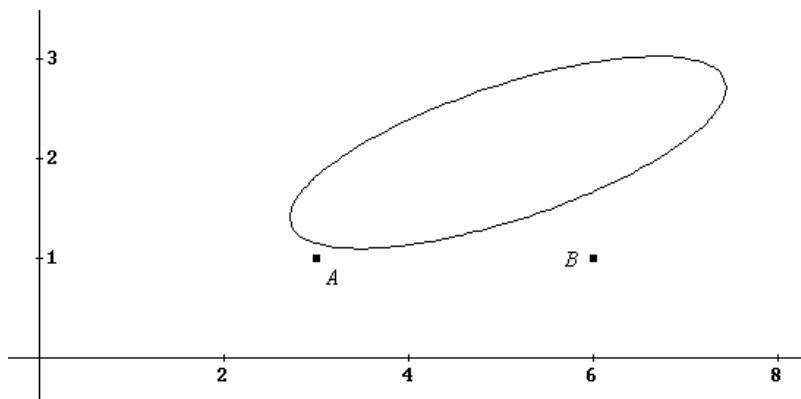
b)  $f(x, y) = 2x^2 - 15x + 30y^2 - 63y$ ;  $A = (2, 2)$ ;  $B = (8, 2)$



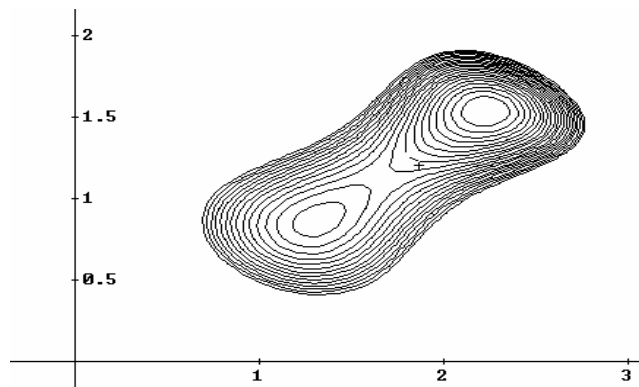
c)  $f(x, y) = 5x^2 - 4xy + 6y^2 - 15x - 23y$ ;  $A = (1, 1)$ ;  $B = (4, 1)$



d)  $f(x, y) = 3x^2 - 10xy + 18y^2 - 10x - 23y$ ;  $A = (3, 1)$ ;  $B = (6, 1)$



2. La función  $f(x, y) = x^2 + \sin(xy) + 2y^2 - 3x - 4y$  no es unimodal. En la figura se observan algunas curvas de nivel que muestran que en esta región ella presenta dos puntos de mínimo. Seleccione adecuadamente el punto inicial  $\mathbf{x}_0$  para determinar ambos puntos de mínimo mediante el método de Powell. Obtenga los resultados con tres cifras decimales exactas. Explique por qué en este caso la convergencia no es tan rápida como en el ejercicio anterior.



3. Halle la mínima distancia entre el punto  $(2, 3, 1)$  y el paraboloide elíptico  $z = x^2 + 2y^2$ . Obtenga el resultado con tres cifras decimales exactas.

4. Dados los puntos  $(1,2; 2)$ ,  $(3,1; 2,5)$ ,  $(3,9; 1,6)$ , halle las coordenadas de un punto  $P$  del plano tal que la suma de sus distancias a los tres puntos dados sea lo menor posible. Obtenga la solución con error menor que 0,001.
5. Cuando un rayo de luz parte del punto  $A$ , se refleja en una superficie y después pasa por un punto  $B$ , lo hace de modo que su trayectoria es mínima. Si un rayo de luz parte del punto  $(1, 1, 6)$ , se refleja en la superficie  $2x^2 + 3y^2 + 4z^2 = 12$  y después pasa por el punto  $(1,5; 1,2; 10)$ , ¿En qué punto de la superficie se reflejó? Determine el resultado con tres cifras decimales exactas.
6. Explique por qué en el algoritmo de Powell el paso de aceleración es fundamental.

## 6.8 El método del simplex secuencial

El método del simplex secuencial es una técnica de búsqueda multidimensional que no se basa en la realización de búsquedas unidimensionales como los otros tres procedimientos estudiados hasta aquí. Además de su eficiencia relativa, se trata de un algoritmo con una lógica muy simple y, por tanto, fácil de programar. Como su fundamento es esencialmente geométrico, es preferible estudiarlo primero en el caso de funciones de dos variables y posteriormente generalizarlo a funciones de  $n$  variables independientes, en las cuales ya no será posible una visualización. Como hasta ahora, se supone que todas las funciones involucradas en el análisis son linealmente unimodales con máximo.

### *El simplex*

En el contexto de este método, se llamará *simplex* de un espacio  $n$  dimensional, al elemento con la misma dimensión del espacio que posea la estructura geométrica más simple posible. Así:

En el espacio unidimensional (la recta), el simplex es un intervalo cerrado, el cual está limitado por dos puntos.

En el espacio bidimensional (el plano), el simplex es un triángulo equilátero, el cual está limitado por tres segmentos iguales y posee tres puntos (vértices) que lo determinan.

En el espacio tridimensional, el simplex es un tetraedro (pirámide de base triangular) regular, que está limitado por cuatro triángulos equiláteros y posee cuatro puntos (vértices) que lo determinan.

En el espacio de dimensión 4, el simplex (que ya no posee un nombre especial y, naturalmente, no se puede representar) sería una “figura” regular delimitada por cinco puntos del espacio, que lo determinan. Esta idea se puede extender a espacios de cualquier dimensión: para un espacio de dimensión  $n$ , el simplex es la figura regular que está limitada por  $n + 1$  puntos de ese espacio que equidistan entre sí. En la figura 1 se muestra el simplex del espacio bidimensional y el del espacio tridimensional.

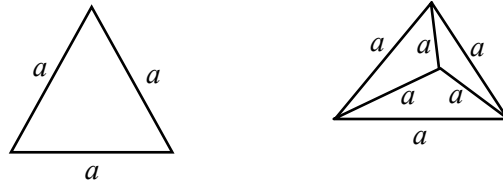


Figura 1

### ***El método del simplex para funciones de dos variables***

Sea  $f$  una función de las variables  $u_1$  y  $u_2$  y sean tres pares ordenados  $\mathbf{x}_1$ ,  $\mathbf{x}_2$  y  $\mathbf{x}_3$  que forman un triángulo equilátero, es decir, un simplex. La idea del algoritmo consiste en ir generando una sucesión de simplex adyacentes uno a otro, que conduzca al punto de máximo de la función  $f$ . La sucesión de simplex se genera siguiendo tres sencillas reglas, de las cuales la primera es muy obvia:

Regla 1: Si  $f_1$ ,  $f_2$  y  $f_3$  son los valores que toma la función  $f(\mathbf{x})$  en los tres vértices del simplex, debe formarse un nuevo simplex en el cual se conserven los dos vértices donde la función toma los dos mayores valores y se sustituya el vértice donde la función toma el menor valor.

Esta es la regla que más frecuentemente se aplica, pero no siempre. En la figura 2, se muestran las curvas de nivel de una función de dos variables con punto de máximo en  $\mathbf{x}^*$  y se aprecian varios simplex, numerados 1, 2, 3, ..., 10 que han sido generados en ese orden aplicando la regla 1. En la figura se ha seguido el convenio de señalar con un punto negro el vértice donde la función toma el mayor valor, con un punto gris el vértice donde el valor de la función es intermedio y con un punto blanco el vértice donde  $f$  toma el valor más pequeño. Obsérvese que, en todos los casos, el vértice señalado con un punto blanco es el que desaparece para formar el simplex que sigue.

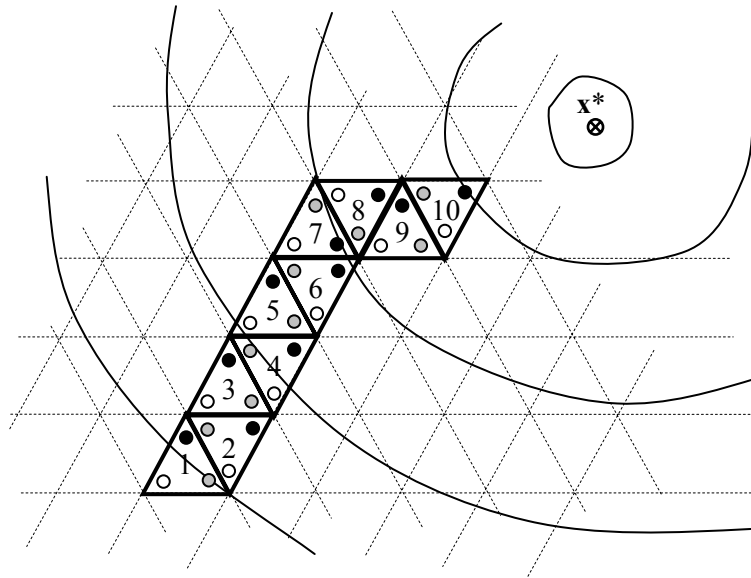


Figura 2

Al generar un nuevo simplex, puede suceder que en el vértice nuevo la función  $f$  tome el valor más pequeño. Si solo existiera la regla 1, el simplex siguiente coincidiría con el anterior y el algoritmo quedaría estancado. En la figura 3 se ilustra una situación en que esto sucede. Nótese

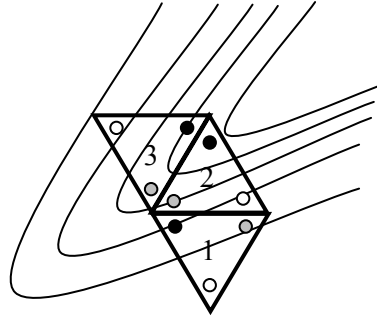


Figura 3

que, según la regla 1, después del simplex 3 el nuevo simplex coincidiría con el 2 y esto detendría el movimiento del simplex. Observe que las curvas de nivel presentan ángulos agudos muy marcados. Esta situación se evita con la regla 2.

Regla 2: El vértice más reciente del simplex no puede eliminarse. Si la función toma el menor valor en el vértice más reciente del simplex, entonces el vértice que se elimina es aquel donde la función toma el siguiente valor en orden creciente.

En la figura 4 se muestra como continuaría el proceso de la figura 3, con la introducción de esta nueva regla. Se muestran en gris los simplex (3 y 4) en que se necesita aplicar la regla 2.

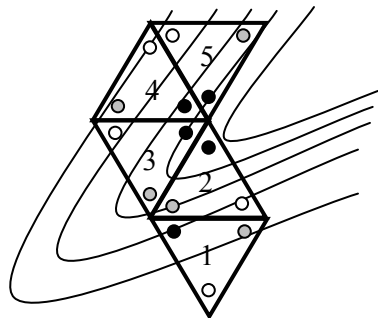


Figura 4

La tercera y última regla es necesaria cuando el simplex se encuentra en las proximidades del punto de máximo. En la figura 5 se aprecia lo que sucede. Los simplex del 4 al 9 van rodeando

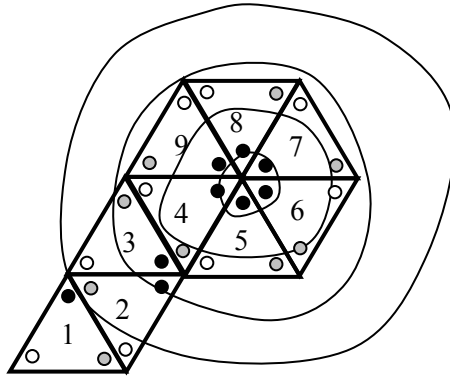


Figura 5

al punto donde ocurre el máximo, de manera que existe un vértice (el que se halla más cerca de  $\mathbf{X}^*$ , que permanece formando parte de los mismos. Evidentemente, el simplex número 10 coincidirá con el 4 y se produce un ciclo: 4, 5, 6, 7, 8, 9, 4,... Cuando se llega a una situación así, que se detecta porque hay un vértice que permanece establemente formando parte del simplex, ya no es posible mejorar la aproximación lograda con un simplex de este tamaño y se procede a tomar una de dos decisiones: Terminar el proceso y aceptar el vértice que se repite como la mejor aproximación a  $\mathbf{x}^*$  o reducir el tamaño del simplex y comenzar de nuevo el proceso a partir de la aproximación que se acaba de obtener. Esta es la regla número 3.

**Regla 3** Cuando un vértice permanece como el de más alto valor en más de  $M$  simplex consecutivos de arista  $a$ , debe reducirse la arista del simplex y comenzar de nuevo a partir del vértice en cuestión, o terminar el proceso y aceptar a dicho vértice como la solución del problema con error absoluto máximo  $a$ . El valor de  $M$  que sugieren los autores del método, depende del número  $n$  de variables independientes y viene dado por:

$$M = 1,65n + 0,05n^2$$

En la tabla 1 se muestra el valor de  $M$  para varias dimensiones  $n$ :

| $n$ | $M$  |
|-----|------|
| 2   | 3,5  |
| 3   | 5,4  |
| 4   | 7,4  |
| 5   | 9,5  |
| 6   | 11,7 |
| 7   | 14,0 |
| 8   | 16,4 |
| 9   | 18,9 |
| 10  | 21,5 |

Tabla 1

así, por ejemplo, para el caso de funciones de dos variables, se sugiere que cuando el vértice óptimo se repita en cuatro simplex consecutivos, debe aplicarse la regla número 3, lo cual significa que, en el caso mostrado en la figura 5, después del simplex número 7 se debió aplicar ya la regla 3.

### ***Cálculo del simplex inicial***

Si el procedimiento secuencial se comenzará a partir del punto  $\mathbf{x}_0$ , es necesario calcular las coordenadas de los vértices del simplex inicial sabiendo que uno de los vértices posee coordenadas  $\mathbf{x}_0$ . Todos los restantes vértices se hallan a una distancia  $a$  de  $\mathbf{x}_0$  y hay una distancia  $a$  entre dos cualesquiera de ellos. Se pueden definir muchos simplex que cumplan estas restricciones y, cuál de ellos se elija, no tiene mucha importancia. Para simplificar los cálculos se supone el simplex colocado en una posición que guarde simetría con todas las variables. En la figura 6 se muestra el simplex inicial para un caso de dos variables y para un caso de tres variables, suponiendo que el vértice inicial se hace corresponder con el origen de un sistema coordenado. En el caso bidimensional, las coordenadas de los tres vértices del simplex inicial son de la forma:

|            | $u_1$ | $u_2$ |
|------------|-------|-------|
| Vértice 1: | 0     | 0     |
| Vértice 2: | $p$   | $q$   |
| Vértice 3: | $q$   | $p$   |

donde  $p$  y  $q$  son parámetros a determinar. En el caso tridimensional, las coordenadas serían:

|            | $u_1$ | $u_2$ | $u_3$ |
|------------|-------|-------|-------|
| Vértice 1: | 0     | 0     | 0     |
| Vértice 2: | $p$   | $q$   | $q$   |
| Vértice 3: | $q$   | $p$   | $q$   |
| Vértice 4: | $q$   | $q$   | $p$   |

donde, de nuevo, los valores de  $p$  y  $q$  deben ser determinados.

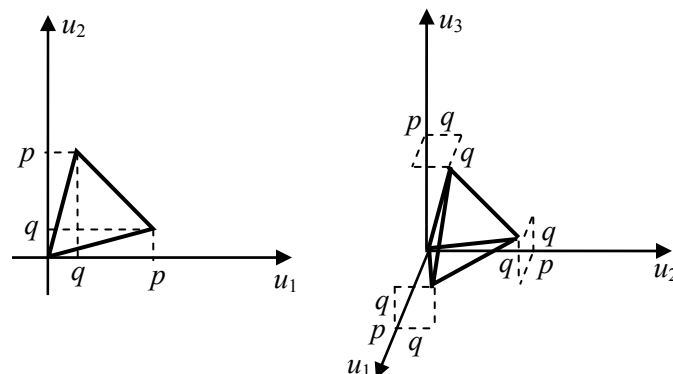


Figura 6

En general, para un caso en el espacio  $n$  dimensional, las coordenadas de los vértices, referidas a un sistema con el origen en el punto  $\mathbf{x}_0$ , serían:

|                   | $u_1$    | $u_2$    | $u_3$    | $\dots$ | $u_n$    |
|-------------------|----------|----------|----------|---------|----------|
| Vértice 1:        | 0        | 0        | 0        | $\dots$ | 0        |
| Vértice 2:        | $p$      | $q$      | $q$      | $\dots$ | $q$      |
| Vértice 3:        | $q$      | $p$      | $q$      | $\dots$ | $q$      |
| Vértice 4:        | $q$      | $q$      | $p$      | $\dots$ | $q$      |
| $\vdots$          | $\vdots$ | $\vdots$ | $\vdots$ |         | $\vdots$ |
| Vértice $n + 1$ : | $q$      | $q$      | $q$      | $\dots$ | $p$      |

Para hallar los valores de  $p$  y  $q$  basta imponer dos condiciones:

La distancia del vértice 1 a cualquiera de los otros debe ser  $a$ , es decir:

$$p^2 + (n-1)q^2 = a^2 \quad (1)$$

y la distancia entre dos cualesquiera de los vértices 2, 3, ...,  $n + 1$  debe ser también igual a  $a$ , esto es:

$$2(p-q)^2 = a^2 \quad (2)$$

De la ecuación (2) resulta:

$$p = \frac{a}{\sqrt{2}} + q \quad (3)$$

Sustituyendo en (1):

$$\left( \frac{a}{\sqrt{2}} + q \right)^2 + (n-1)q^2 = a^2$$

De donde

$$\frac{a^2}{2} + \sqrt{2}aq + q^2 + nq^2 - q^2 = a^2$$

Multiplicando por 2 y ordenando:

$$2nq^2 + 2\sqrt{2}aq - a^2 = 0$$

y, de aquí:

$$q = \frac{-2\sqrt{2}a + \sqrt{8a^2 + 8na^2}}{4n} = \frac{-2\sqrt{2}a + 2\sqrt{2}a\sqrt{n+1}}{4n}$$

Esto es:

$$q = \frac{a(\sqrt{n+1}-1)}{\sqrt{2}n} \quad (4)$$

Sustituyendo en (3) se halla  $p$ :

$$p = \frac{a}{\sqrt{2}} + \frac{a(\sqrt{n+1}-1)}{\sqrt{2}n}$$

$$p = \frac{a}{\sqrt{2}n} (n + \sqrt{n+1} - 1) \quad (5)$$

Nótese que, el hecho de que el sistema formado por las ecuaciones (1) y (2), posea solución, dada por las expresiones (4) y (5), prueba que la forma propuesta para los vértices del simplex es válida para cualquier dimensión  $n$ .

### **Ejemplo 1**

Determine las coordenadas de los vértices de un simplex inicial para un problema de dimensión  $n = 3$  si el punto donde comenzará el proceso es  $\mathbf{x}_0 = (2,734; 1,281; -1,342)$  y la arista del simplex se ha seleccionado con longitud  $a = 0,3$ .

Solución:

A partir de las ecuaciones (4) y (5) se obtienen los valores de  $p$  y  $q$ :

$$p = \frac{a}{\sqrt{2n}}(n + \sqrt{n+1} - 1) = \frac{0,3}{3\sqrt{2}}(3 + \sqrt{3+1} - 1) = 0,2828$$

$$q = \frac{a(\sqrt{n+1} - 1)}{\sqrt{2n}} = \frac{0,3(\sqrt{3+1} - 1)}{3\sqrt{2}} = 0,0707$$

Con estos valores de  $p$  y  $q$  se pueden escribir las coordenadas de los cuatro vértices relativos al punto inicial  $\mathbf{X}_0$ :

Vértice 1:  $(0, 0, 0)$

Vértice 2:  $(p, q, q) = (0,2828; 0,0707; 0,0707)$

Vértice 3:  $(q, p, q) = (0,0707; 0,2828; 0,0707)$

Vértice 4:  $(q, q, p) = (0,0707; 0,0707; 0,2828)$

Para tener las coordenadas absolutas de los cuatro vértices, basta sumar las coordenadas relativas con las del punto inicial  $\mathbf{X}_0$ :

Coordenadas de los vértices:

Vértice 1:  $(0, 0, 0) + (2,734; 1,281; -1,342) = (2,734; 1,281; -1,342)$

Vértice 2:  $(0,2828; 0,0707; 0,0707) + (2,734; 1,281; -1,342) = (3,0168; 1,3517; -1,2713)$

Vértice 3:  $(0,0707; 0,2828; 0,0707) + (2,734; 1,281; -1,342) = (2,8047; 1,5638; -1,2713)$

Vértice 4:  $(0,0707; 0,0707; 0,2828) + (2,734; 1,281; -1,342) = (2,8047; 1,3517; -1,0592)$

### **Determinación de los vértices de un nuevo simplex**

Para completar el algoritmo del simplex secuencial, se necesitan expresiones que permitan calcular analíticamente las coordenadas de los vértices de un nuevo simplex a partir de las del simplex anterior. Evidentemente, solo se necesita calcular las coordenadas de un vértice, ya que dos simplex consecutivos comparten  $n$  de sus  $n + 1$  vértices. Para simplificar la deducción, el análisis que sigue se basa en la geometría del caso de dimensión 3, aunque las fórmulas se obtienen para el caso general de dimensión  $n$ .

La figura 7 muestra dos simplex consecutivos. Se ha llamado  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ , a los vértices que comparten ambos simplex,  $\mathbf{S}$  al vértice que sale, es decir que forma parte del simplex anterior pero no del siguiente, y  $\mathbf{E}$  al vértice que entra, o sea, que forma parte del simplex nuevo pero no del anterior.

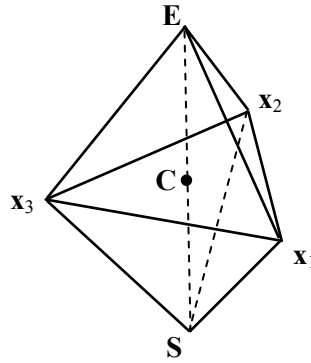


Figura 7

En la figura se muestra, además, el segmento que une a los vértices **E** y **S** y el punto **C** en que este segmento interseca al plano formado por los vértices que comparten ambos simplex. El punto **C** es el centro del segmento **ES** y es también el centroide del conjunto de vértices  $x_1, x_2, \dots, x_n$  que comparten ambos simplex. Por ser **C** el centro del segmento **ES** se cumple que:

$$C = \frac{1}{2}(E + S)$$

y por ser el centroide del conjunto de vértices  $x_1, x_2, \dots, x_n$ , satisface la ecuación:

$$C = \frac{1}{n} \sum_{i=1}^n x_i$$

Igualando los miembros de la derecha de ambas igualdades se obtiene:

$$\frac{1}{2}(E + S) = \frac{1}{n} \sum_{i=1}^n x_i$$

de donde:

$$E = \frac{2}{n} \sum_{i=1}^n x_i - S \quad (6)$$

La ecuación (6) permite determinar las coordenadas del nuevo vértice **E** a partir de las coordenadas de los vértices del simplex actual y de la decisión de cual de ellos será el que dejará de formar parte del simplex, lo cual se decide a partir de las reglas 1 y 2.

### Ejemplo 2

En el ejemplo 1 fueron halladas las coordenadas de los cuatro vértices de un simplex en el espacio tridimensional. Suponga que el vértice 2 será eliminado del simplex y halle las coordenadas del nuevo vértice **E** que formará parte del siguiente simplex.

Solución:

Aplicando la fórmula (6) resulta:  $E = \frac{2}{3}(\text{Vértice 1} + \text{Vértice 2} + \text{Vértice 4}) - \text{Vértice 2}$

ya que ya que los vértices 1, 3 y 4 son los que comparten ambos simplex y el vértice 2 es el que sale. Se obtiene:

$$\mathbf{E} = \frac{2}{3} [(2,734; 1,281; -1,342) + (2,8047; 1,5638; -1,2713) + (2,8047; 1,3517; -1,0592)] - (3,0168; 1,3517; -1,2713)$$

$$\mathbf{E} = (2,5455; 1,4460; -1,1770)$$

### Algoritmo en pseudo código

El siguiente algoritmo permite hallar el punto de máximo de una función de  $n$  variables independientes linealmente unimodal mediante el método del simplex secuencial. Se utiliza la notación  $u_{ij}$  para denotar a la  $j$ -sima ( $j = 1, 2, \dots, n$ ) componente del vector  $\mathbf{x}_i$  ( $i = 0, 1, \dots, n$ ). El algoritmo utiliza como datos la función  $f(\mathbf{x})$ , el número  $n$  de variables independientes de dicha función, el punto  $\mathbf{X}_0 = (u_{01}, u_{02}, \dots, u_{0n})$  donde comenzará el proceso y la longitud  $a$  de la arista del simplex y ofrece como resultado una aproximación al punto de máximo  $\mathbf{X}^*$  con error absoluto menor que  $a$ . Si este error no es lo suficientemente pequeño, todo el algoritmo se repite para un valor menor de  $a$  a partir del resultado anterior.

```

 $p := \frac{a}{\sqrt{2n}}(n + \sqrt{n+1} - 1)$ 
 $q := \frac{a(\sqrt{n+1} - 1)}{\sqrt{2n}}$ 
for  $i = 1$  to  $n$ 
    for  $j = 1$  to  $n$  {En este ciclo se forma el simplex inicial, cuyo vértice  $\mathbf{x}_0$  es dato}
        if  $j = i$  then  $u_{ij} := u_{0j} + p$  else  $u_{ij} := u_{0j} + q$ 
    end
end
for  $i = 1$  to  $n$ 
     $f_i := f(\mathbf{x}_i)$ 
     $R_i := 1$  {La variable  $R_i$  cuenta la cantidad de veces que el vértice  $i$  ha permanecido formando parte del simplex. Inicialmente todos valen 1}
end
 $mejor := i$  tal que  $f_i = \max \{f_0, f_1, \dots, f_n\}$ 
 $peor := i$  tal que  $f_i = \min \{f_0, f_1, \dots, f_n\}$ 
 $casipeor := i$  tal que  $f_i = \min [\{f_0, f_1, \dots, f_n\} - \{f_{peor}\}]$ 
 $\acute{ultimo} := mejor$  {último es el sub índice del último vértice que entró a formar parte del simplex. Al inicio se hace coincidir con el mejor}
 $M := 1,65n + 0,05n^2$ 
repeat
    if  $peor \neq \acute{ultimo}$  then  $k := peor$  else  $k := casipeor$  { $k$  indica el sub-índice del vértice que será cambiado}
     $\acute{ultimo} := k$ 
     $\mathbf{x}_k := \frac{2}{n} \sum_{i \neq k} \mathbf{x}_i - \mathbf{x}_k$  {Se hallan las coordenadas del nuevo vértice}
    for  $i := 0$  to  $n$  {Se actualizan los valores de  $R_i$ }
        if  $i = k$  then  $R_i := 1$  else  $R_i := R_i + 1$ 
    end
     $f_k := f(\mathbf{x}_k)$ 
     $mejor := i$  tal que  $f_i = \max \{f_0, f_1, \dots, f_n\}$ 
     $peor := i$  tal que  $f_i = \min \{f_0, f_1, \dots, f_n\}$ 
     $casipeor := i$  tal que  $f_i = \min [\{f_0, f_1, \dots, f_n\} - \{f_{peor}\}]$ 
until  $R_{mejor} > M$ 

```

El punto de máximo es  $\mathbf{x}_{mejor}$  con error absoluto menor que  $a$ .  
Terminar

### Ejemplo 3

Halle el punto de máximo de la función cuadrática  $f(\mathbf{x}) = 100 - 3u_1^2 - 4u_2^2 + 5u_1u_2 + 2u_1$  por el método del simplex secuencial con precisión de tres cifras decimales exactas tomando  $\mathbf{x}_0 = (5, 5)$ .

Solución:

El problema se resolverá en varias etapas de búsqueda en las cuales se disminuye paulatinamente el valor de  $a$  hasta llegar a 0,0005 para obtener el resultado con tres cifras decimales exactas.

Etapas 1

$\mathbf{x}_0 = (5, 5)$   $a = 1$

La tabla 2 muestra los resultados de esta etapa. En cada simplex se muestra, junto al valor de la función en cada vértice, entre paréntesis las veces que ese vértice ha estado formando parte del simplex y un signo + ó – que señalará el vértice donde la función toma el mayor y el menor valor en este simplex. Por razones de espacio, las entradas de la tabla solo muestran las cifras decimales necesarias para la aplicación de las reglas correspondientes. Del mismo modo, en las coordenadas del mejor vértice (el señalado con +) solo se han incluido las cifras decimales más exactas, de acuerdo con el valor de  $a$  que se esté usando. En cada renglón de la tabla se señala la regla que será aplicada de inmediato. El proceso se detiene cuando el vértice (+) se mantiene en cuatro simplex sucesivos, ya que para el caso de  $n = 2$  el valor de  $M$  es 3,5.

| Simplex | $f_0$      | $f_1$     | $f_2$     | Regla | Coordenadas de (+) |      |
|---------|------------|-----------|-----------|-------|--------------------|------|
| 1       | 60,0(1)+   | 51,4(1)   | 42,1(1)–  | 1     | 5,00               | 5,00 |
| 2       | 60,0(2)    | 51,4(2)–  | 62,5(1)+  | 1     | 5,71               | 4,29 |
| 3       | 60,0(3)–   | 72,6(1)+  | 62,5(2)   | 1     | 4,74               | 4,03 |
| 4       | 68,2(1)    | 72,6(2)+  | 62,5(3)–  | 1     | 4,74               | 4,03 |
| 5       | 68,2(2)–   | 72,6(3)   | 79,8(1)+  | 1     | 4,48               | 3,07 |
| ⋮       | ⋮          | ⋮         | ⋮         | ⋮     | ⋮                  | ⋮    |
| 13      | 99,1(1)+   | 96,9(3)   | 92,8(2)–  | 1     | 1,33               | 1,33 |
| 14      | 99,1(2)+   | 96,9(4)   | 96,4(1)–  | 2     | 1,33               | 1,33 |
| 15      | 99,1(3)    | 100,1(1)+ | 96,4(2)–  | 1     | 1,07               | 0,36 |
| 16      | 99,1(4)    | 100,1(2)+ | 97,7(1)–  | 2     | 1,07               | 0,36 |
| 17      | 100,18(1)+ | 100,12(3) | 97,70(2)– | 1     | 0,10               | 0,10 |
| 18      | 100,18(2)+ | 100,12(4) | 95,74(1)– | 2     | 0,10               | 0,10 |
| 19      | 100,18(3)+ | 97,30(1)  | 95,74(2)– | 1     | 0,10               | 0,10 |
| 20      | 100,18(4)+ | 97,30(2)  | 96,61(2)– | 3     | 0,10               | 0,10 |

Tabla 2

El proceso se detuvo en el punto  $u_1 = 0,101021$ ;  $u_2 = 0,101021$

## Etapa 2

$$\mathbf{x}_0 = (0,101021; 0,101021); a = 0,1$$

La tabla 3 muestra los resultados. Como los valores de la función en los tres vértices son muy similares, solo se muestra la cantidad que excede de 100 y se brindan tres o cuatro cifras decimales, según la necesidad. El proceso se detiene cuando el mejor valor obtenido se mantiene en 4 simplex consecutivos. Nótese como el final de la etapa se caracteriza por el empleo frecuente de la regla 2.

El mejor resultado de esta etapa es el punto  $u_1 = 0,732340$ ;  $u_2 = 0,449498$

| Simplex | $f_0$    | $f_1$    | $f_2$    | Regla | Coordenadas de (+) |       |
|---------|----------|----------|----------|-------|--------------------|-------|
|         | 100,...  | 100,...  | 100,...  |       |                    |       |
| 1       | 182(1)   | 339(1)+  | 175(1)–  | 1     | 0,198              | 0,127 |
| 2       | 182(2)–  | 339(2)+  | 277(1)   | 1     | 0,198              | 0,127 |
| 3       | 383(1)+  | 339(3)   | 277(2)–  | 1     | 0,268              | 0,056 |
| 4       | 383(2)   | 339(4)–  | 460(1)+  | 1     | 0,294              | 0,153 |
| 5       | 383(2)–  | 453(4)   | 460(1)+  | 1     | 0,294              | 0,153 |
| ⋮       | ⋮        | ⋮        | ⋮        | ⋮     | ⋮                  | ⋮     |
| 13      | 6463(4)  | 6733(2)+ | 6457(1)– | 2     | 0,610              | 0,327 |
| 14      | 6877(1)+ | 6733(3)  | 6457(2)– | 1     | 0,636              | 0,424 |
| 15      | 6877(2)+ | 6733(4)  | 6641(1)– | 2     | 0,636              | 0,424 |
| 16      | 6877(3)  | 6934(1)+ | 6641(2)– | 1     | 0,732              | 0,449 |
| 17      | 6877(4)  | 6934(2)+ | 6485(1)– | 2     | 0,732              | 0,449 |
| 18      | 6692(1)  | 6934(3)+ | 6485(2)– | 1     | 0,732              | 0,449 |
| 19      | 6692(2)  | 6934(4)+ | 6628(1)– | 3     | 0,732              | 0,449 |

Tabla 3

## Etapa 3

$$\mathbf{x}_0 = (0,732340; 0,449498), a = 0,05$$

Los resultados se muestran en la tabla 4 en la cual se aprecia que fue necesaria una cantidad de simplex mucho menor, debido a que la reducción de  $a$  fue también menor que en otras etapas. Se obtuvo, como resultado el punto:  $(0,684044; 0,436557)$ .

| Simplex | $f_0$    | $f_1$    | $f_2$    | Regla | Coordenadas de (+) |        |
|---------|----------|----------|----------|-------|--------------------|--------|
|         | 100,6... | 100,6... | 100,6... |       |                    |        |
| 1       | 934(1)+  | 827(1)–  | 880(1)   | 1     | 0,7323             | 0,4495 |
| 2       | 934(2)+  | 860(1)–  | 880(2)   | 2     | 0,7323             | 0,4495 |
| 3       | 934(3)   | 860(2)–  | 951(1)+  | 1     | 0,6840             | 0,4366 |
| 4       | 934(4)   | 855(1)–  | 951(2)+  | 2     | 0,6840             | 0,4366 |
| 5       | 909(1)   | 855(2)–  | 951(3)+  | 1     | 0,6840             | 0,4366 |
| 6       | 909(2)   | 877(1)–  | 951(4)+  | 3     | 0,6840             | 0,4366 |

Tabla 4

#### Etapa 4

$\mathbf{x}_0 = (0,684044; 0,436557)$ ,  $a = 0,005$

Los resultados se muestran en la tabla 5. Recuérdese que, para abreviar, los valores de la función que se muestran en cada fila son las cifras que siguen a 100,69. Se obtuvo como resultado el punto: (0,697239; 0,435610) que pasa a ser el punto de partida de la siguiente y última etapa.

| Simplex | $f_0$     | $f_1$     | $f_2$     | Regla | Coordenadas de (+) |        |
|---------|-----------|-----------|-----------|-------|--------------------|--------|
|         | 100,69... | 100,69... | 100,69... |       |                    |        |
| 1       | 51(1)     | 54(1)+    | 48(1)–    | 1     | 0,6889             | 0,4379 |
| 2       | 5132(2)–  | 5373(2)   | 5515(1)+  | 1     | 0,6876             | 0,4330 |
| 3       | 5627(1)+  | 5373(3)–  | 5515(2)   | 1     | 0,6924             | 0,4343 |
| 4       | 5627(2)+  | 5598(1)   | 5515(3)–  | 1     | 0,6924             | 0,4343 |
| 5       | 5627(3)+  | 5598(2)   | 5582(1)–  | 2     | 0,6924             | 0,4343 |
| 6       | 5627(4)   | 5648(1)+  | 5582(2)–  | 1     | 0,6972             | 0,4356 |
| 7       | 5627(5)   | 5648(2)+  | 5522(1)–  | 2     | 0,6972             | 0,4356 |
| 8       | 5581(1)   | 5648(3)+  | 5522(2)–  | 1     | 0,6972             | 0,4356 |
| 9       | 5581(2)   | 5648(4)+  | 5579(1)–  | 3     | 0,6972             | 0,4356 |

Tabla 5

#### Etapa 5

$\mathbf{x}_0 = (0,697239; 0,435610)$ ,  $a = 0,0005$

Los resultados se muestran en la tabla 6. Recuérdese que, para abreviar, los valores de la función que se muestran en cada fila son las cifras que siguen a 100,6956.

Se obtuvo, como resultado el punto: (0,695661; 0,434739) que es el punto de máximo buscado con tres cifras decimales exactas. Como se recordará, el valor exacto de  $\mathbf{X}^*$  es (0,6956521; 0,4347826) por lo cual, realmente, la respuesta hallada tiene 4 cifras decimales exactas.

| Simplex | $f_0$       | $f_1$       | $f_2$       | Regla | Coordenadas de (+) |        |
|---------|-------------|-------------|-------------|-------|--------------------|--------|
|         | 100,6956... | 100,6956... | 100,6956... |       |                    |        |
| 1       | 4844(1)+    | 4556(1)–    | 4771(1)     | 1     | 0,6972             | 0,4356 |
| 2       | 4844(2)     | 4932(1)+    | 4771(2)–    | 1     | 0,6969             | 0,4360 |
| 3       | 4844(3)–    | 4932(2)     | 5042(1)+    | 1     | 0,6968             | 0,4355 |
| 4       | 5001(1)     | 4932(3)–    | 5042(2)+    | 1     | 0,6968             | 0,4355 |
| 5       | 5001(2)–    | 5149(1)+    | 5042(3)     | 1     | 0,6963             | 0,4354 |
| 6       | 5019(1)–    | 5149(2)+    | 5042(4)     | 2     | 0,6963             | 0,4354 |
| 7       | 5019(2)–    | 5149(3)     | 5163(1)+    | 1     | 0,6961             | 0,4349 |
| 8       | 5165(1)+    | 5149(4)–    | 5163(2)     | 1     | 0,6958             | 0,4352 |
| 9       | 5165(2)     | 5256(1)+    | 5163(3)–    | 1     | 0,6957             | 0,4347 |
| 10      | 5165(3)     | 5256(2)+    | 5090(1)–    | 2     | 0,6957             | 0,4347 |
| 11      | 5179(1)     | 5256(3)+    | 5090(2)–    | 1     | 0,6957             | 0,4347 |
| 12      | 5179(2)     | 5256(4)+    | 5134(1)–    | 3     | 0,6957             | 0,4347 |

Tabla 6

#### Ejemplo 4

Halle el punto de mínimo de la función de Rosembrook (vea el ejemplo 3 de la sección 6.5)

$$f(\mathbf{x}) = 100(u_2 - u_1^2)^2 + (1 - u_1)^2$$

mediante el método del simplex secuencial y compare su efectividad con los métodos anteriormente estudiados.

Solución:

Como se recordará, la función de Rosembrook está especialmente diseñada para que su tratamiento sea sumamente difícil. La presencia de “valles estrechos” hace que todos los métodos confronten serios problemas en la búsqueda del punto de mínimo, el cual se encuentra en

$$\mathbf{x}^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

El método simplex mal interpreta los puntos en el fondo del valle estrecho de esta función como si fuera el punto de mínimo y es necesario utilizar valores de  $a$  sumamente pequeños. A continuación se muestran los resultados finales de las 10 etapas utilizadas para hallar el punto de mínimo con error menor de 0,002. La primera etapa de búsqueda comienza en  $(-1; 1)$ . Las demás lo hacen en el resultado de la etapa precedente.

|                        |                   |                                 |
|------------------------|-------------------|---------------------------------|
| Etapla 1: $a = 0,01$   | Evaluaciones: 511 | Resultado: (0,55296; 0,30436)   |
| Etapla 2: $a = 0,005$  | Evaluaciones: 155 | Resultado: (0,709335; 0,503162) |
| Etapla 3: $a = 0,002$  | Evaluaciones: 452 | Resultado: (0,867905; 0,753650) |
| Etapla 4: $a = 0,001$  | Evaluaciones: 288 | Resultado: (0,917658; 0,842300) |
| Etapla 5: $a = 0,0005$ | Evaluaciones: 545 | Resultado: (0,960028; 0,921793) |
| Etapla 6: $a = 0,0002$ | Evaluaciones: 775 | Resultado: (0,983935; 0,968186) |

|                              |                   |                                 |
|------------------------------|-------------------|---------------------------------|
| Etapa 7: $\alpha = 0,0001$   | Evaluaciones: 563 | Resultado: (0,992300; 0,984683) |
| Etapa 8: $\alpha = 0,00005$  | Evaluaciones: 531 | Resultado: (0,996134; 0,992300) |
| Etapa 9: $\alpha = 0,00002$  | Evaluaciones: 685 | Resultado: (0,998207; 0,996424) |
| Etapa 10: $\alpha = 0,00001$ | Evaluaciones: 607 | Resultado: (0,999111; 0,998226) |

Cantidad de evaluaciones: 5112

La comparación con los métodos anteriormente estudiados resulta difícil, por cuanto todos ellos estaban basados en búsquedas unidimensionales. Si se supone, lo cual no es exagerado, que en cada búsqueda unidimensional se necesite evaluar la función en 20 puntos diferentes, entonces, los métodos anteriores habrían necesitado las siguientes cantidades de evaluaciones:

Búsqueda por coordenadas: 2000 búsquedas, unas 40000 evaluaciones, para llegar a

$$\mathbf{X}^* \approx \begin{bmatrix} 0,980 \\ 0,960 \end{bmatrix}$$

Método del gradiente: 1000 búsquedas unidimensionales, unas 20000 evaluaciones, para obtener:

$$\mathbf{X}^* \approx \begin{bmatrix} 0.9202 \\ 0.8468 \end{bmatrix}$$

Método de Powell: 29 búsquedas unidimensionales, unas 600 evaluaciones, para lograr:

$$\mathbf{X}^* \approx \begin{bmatrix} 1.000002 \\ 1.000004 \end{bmatrix}$$

Método del simplex secuencial: 5112 evaluaciones, para obtener:

$$\mathbf{X}^* \approx \begin{bmatrix} 0,999111 \\ 0,998226 \end{bmatrix}$$

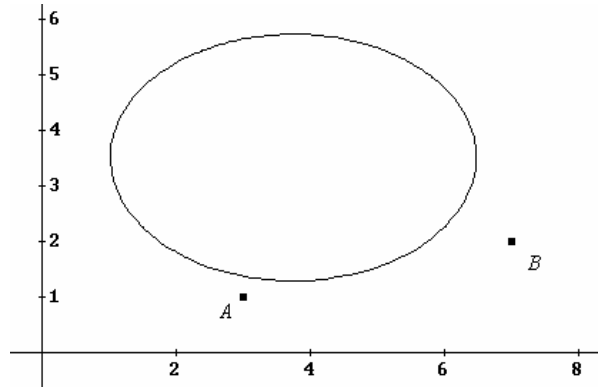
Como se observa, el método de Powell se comporta más eficientemente en el trabajo con la función de Rosembrook. Sin embargo, el método del simplex secuencial aventaja ampliamente a los métodos del gradiente y de búsqueda por coordenadas. Si se tiene en cuenta la dificultad de los algoritmos, el método del simplex secuencial posee la ventaja de no necesitar de algoritmos auxiliares de búsqueda unidimensional, cuya elaboración puede ser bastante compleja. Por otra parte, el método del simplex es, por la simplicidad de su lógica, un algoritmo sumamente robusto y confiable.

## Ejercicios

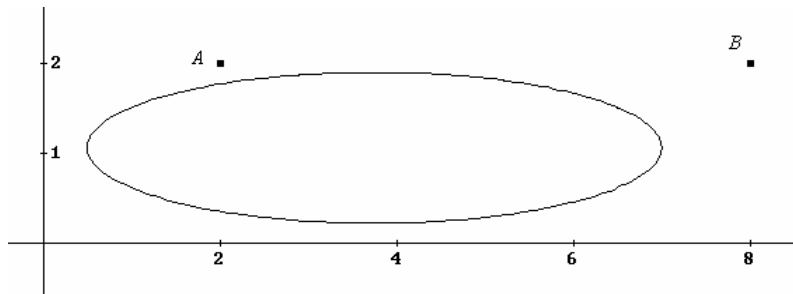
Algunos de los ejercicios que siguen son muy laboriosos para realizar los cálculos a mano. Se supone que posea programas computacionales, preferiblemente confeccionados por usted, para ayudarle en el trabajo. De no ser así, en los ejercicios que requieran mucho trabajo manual, determine los resultados con menos cifras exactas que las exigidas.

1. Cada una de las funciones cuadráticas que siguen aparecieron también en los ejercicios de las tres secciones anteriores. En cada caso aplique el método del simplex secuencial, utilizando como punto de partida los puntos  $A$  y  $B$  que se dan en la propia figura. Obtenga la solución con tres cifras decimales exactas. Compare con la solución obtenida utilizando los métodos anteriores. Observe la curva de nivel que acompaña a cada función y diga si puede sacar alguna conclusión general como resultado de este ejercicio.

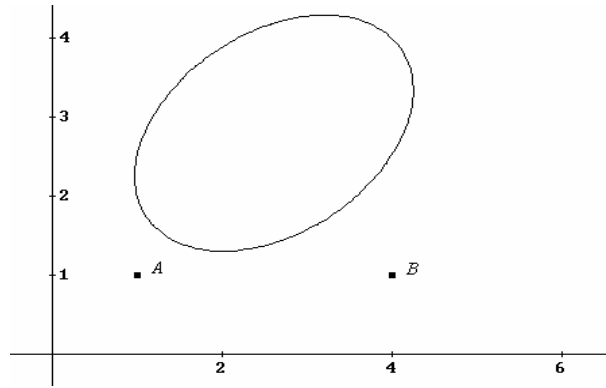
a)  $f(x, y) = 2x^2 - 15x + 3y^2 - 21y$ ;  $A = (3, 1)$ ;  $B = (7, 2)$



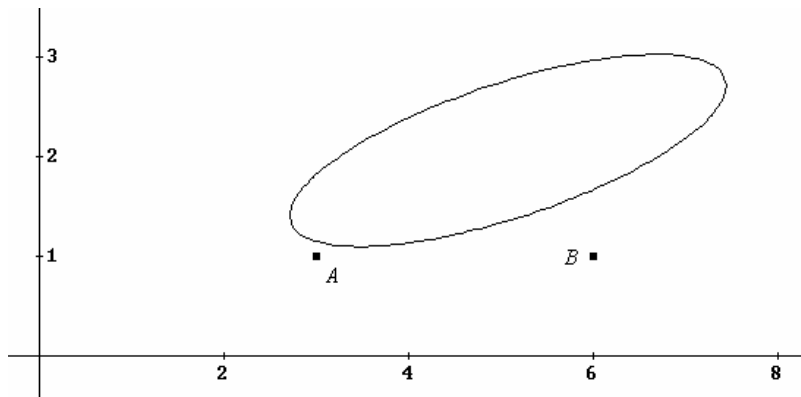
b)  $f(x, y) = 2x^2 - 15x + 30y^2 - 63y$ ;  $A = (2, 2)$ ;  $B = (8, 2)$



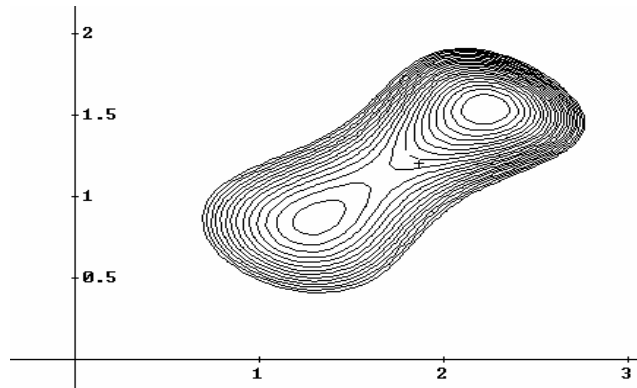
c)  $f(x, y) = 5x^2 - 4xy + 6y^2 - 15x - 23y$ ;  $A = (1, 1)$ ;  $B = (4, 1)$



d)  $f(x, y) = 3x^2 - 10xy + 18y^2 - 10x - 23y$ ;  $A = (3, 1)$ ;  $B = (6, 1)$



2. La función  $f(x, y) = x^2 + \sin(xy) + 2y^2 - 3x - 4y$  no es unimodal. En la figura se observan algunas curvas de nivel que muestran que en esta región ella presenta dos puntos de mínimo. Seleccione adecuadamente el punto inicial  $\mathbf{x}_0$  para determinar ambos puntos de mínimo mediante el método del simplex secuencial. Obtenga los resultados con tres cifras decimales exactas.



3. Ajuste el modelo no lineal  $g(x) = a + be^{px}$  a los puntos que se muestran en la tabla. Obtenga el resultado con error menor que 0,001.

|     |      |      |      |      |      |
|-----|------|------|------|------|------|
| $x$ | 1    | 2    | 3    | 4    | 5    |
| $y$ | 2,10 | 2,60 | 2,82 | 2,92 | 2,96 |

4. Halle la distancia más corta entre las superficies  $z = 4 + 2x^2 + 3y^2$  y  $y = 2 + 3x^2 + z^2$ . Obtenga la solución con tres cifras decimales exactas.
5. En el algoritmo del simplex secuencial aparecen los pasos:

$$\begin{aligned} \text{mejor} &:= i \text{ tal que } f_i = \max \{f_0, f_1, \dots, f_n\} \\ \text{peor} &:= i \text{ tal que } f_i = \min \{f_0, f_1, \dots, f_n\} \end{aligned}$$

Elabore un algoritmo en pseudo código que determine *mejor* y *peor* en un solo lazo.

### Otras lecturas recomendadas

El tema de optimización numérica es sumamente amplio y en el texto solo han sido tocados los métodos más generalmente reconocidos, por lo cual existe una gran cantidad de algoritmos que ni siquiera han sido mencionados. En particular, se ha obviado el problema de la optimización multidimensional con restricciones. Al lector interesado en ampliar sus conocimientos con otros métodos o estudiar otros enfoques, puede consultar la enciclopédica obra “Optimization, Theory and Practice” de Bevrige y Schechter, el cual ha sido reproducido en Cuba. Con un enfoque más moderno, pero de lectura un poco más difícil, puede consultarse el libro “Optimization” de Luemberger, que también ha sido editado en Cuba.

### Principales ideas del capítulo

- La función  $f(x)$  es unimodal con máximo si existe en su dominio un  $x^*$  tal que, si  $x_1$  y  $x_2$  pertenecen al dominio se cumple que: Si  $x_1 < x_2 < x^*$  entonces  $f(x_1) < f(x_2)$  y si  $x^* < x_1 < x_2$  entonces  $f(x_1) > f(x_2)$ .
- La propiedad básica de la optimización unidimensional establece que: Si  $f(x)$  una función unimodal con máximo en  $x^*$  y  $x_1$  y  $x_2$  son dos valores de su dominio y se denomina  $y_1 = f(x_1)$  y  $y_2 = f(x_2)$ . Entonces  $y_1 < y_2 \Rightarrow x_1 < x^*$ ;  $y_1 > y_2 \Rightarrow x^* < x_2$  y  $y_1 = y_2 \Rightarrow x_1 < x^* < x_2$ .
- Aquellos algoritmos en los que los experimentos se realizan todos a un tiempo, se llaman de *búsqueda simultánea* mientras que los que siguen la estrategia de realizar un experimento después que han sido analizados los resultados de los experimentos anteriores, se llaman *métodos secuenciales*.
- Si se realizan  $n$  experimentos simultáneos en los valores  $x_1 < x_2 < \dots < x_n$  con resultados  $y_1, y_2, \dots, y_n$  y se obtiene  $y_k = \max \{y_i\}$  con  $1 < k < n$ , entonces el punto de máximo se encuentra entre  $x_{k-1}$  y  $x_{k+1}$ .
- El método de búsqueda secuencial uniforme consiste en generar la sucesión de valores:  $x_i = x_0 + is$  ( $i = 0, 1, 2, 3, \dots$ ) y obtener para cada uno de ellos la imagen  $y_i = f(x_i)$ . El proceso se detiene tan pronto se obtiene un  $y_k$  tal que  $y_{k-1} > y_k$ , y puede entonces asegurarse, se acuerdo con el principio básico, que  $x^*$  se encuentra entre  $x_{k-2}$  y  $x_k$ .
- La búsqueda secuencial acelerada consiste en lo siguiente: en cada paso del algoritmo mientras no se cumpla la condición de parada, el paso se duplica en valor; de esta manera, aun cuando inicialmente  $s$  fuera demasiado pequeño, pronto toma valores suficientemente grandes para alcanzar a  $x^*$  en no muchas iteraciones.

- En el método de bisección se toman dos puntos experimentales  $x_1$  y  $x_2$  muy próximos entre sí y a ambos lados del centro del intervalo  $[a, b]$ . Si se llama  $y_1 = f(x_1)$  y  $y_2 = f(x_2)$ , entonces se tiene que  $y_1 < y_2 \Rightarrow x_1 \leq x^* \leq b$ ;  $y_1 > y_2 \Rightarrow a \leq x^* \leq x_2$  y  $y_1 = y_2 \Rightarrow x_1 \leq x^* \leq x_2$ .
- En el método de bisección se cumple que:  $\frac{L_n}{L_0} = \frac{1}{2^{n/2}}$ , donde  $L_n$  representa la longitud del intervalo de búsqueda después de haber realizado  $n$  experimentos.
- El método de la sección áurea utiliza los números  $D_n$  definidos mediante las condiciones:  $D_n = D_{n-1} + D_{n-2}$  ( $n = 2, 3, 4, \dots$ ) y  $D_n = r D_{n-1}$  ( $n = 1, 2, 3, \dots$ ). Se demuestra que:  $r = 1,618034\dots$  y  $G = \frac{1}{r} = r - 1 = 0,618034$ . En cada paso del método la longitud del intervalo de búsqueda coincide con uno de los números  $D_n$  y los experimentos se realizan a distancia  $D_{n-2}$  de los extremos. De esta forma se logra que en cada iteración se requiere solamente un nuevo experimento.
- En el método de la sección áurea se cumple que  $\frac{L_n}{L_0} = G^{n-1}$  y resulta mucho menor que para el método de la bisección.
- La función  $z = f(\mathbf{x})$  se denomina linealmente unimodal con máximo en la región  $R$ , si ella es unimodal con máximo sobre cualquier trayectoria recta contenida en  $R$ .
- Una función cuadrática cualquiera  $f$  de  $n$  variables independientes se puede expresar como  $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x}$  donde  $\mathbf{A}$  es una matriz  $n \times n$  simétrica y  $\mathbf{b}$  es una matriz  $n \times 1$ .
- Una matriz simétrica  $\mathbf{A}_{n \times n}$  se llama positiva definida si para todo vector no nulo  $\mathbf{x}$  se cumple que  $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ . Si para todo  $\mathbf{x}$  no nulo se tiene  $\mathbf{x}^T \mathbf{A} \mathbf{x} < 0$ , la matriz  $\mathbf{A}$  se llama negativa definida.
- Si  $\mathbf{A}$  es positiva (negativa) definida la función cuadrática  $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x}$  posee un punto  $\mathbf{x}^*$  de mínimo (máximo) y se cumple que  $\mathbf{A} \mathbf{x}^* = \mathbf{b}$ .
- Si  $\mathbf{A}$  es positiva (negativa) definida la función cuadrática con mínimo (máximo)  $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x}$  es linealmente unimodal.
- El método de búsqueda por coordenadas consiste en realizar búsquedas sucesivas unidimensionales en las direcciones de los ejes de las variables independientes, es decir, haciendo variar una sola variable en cada búsqueda. Se trata de un método solamente recomendable cuando existe poca interacción entre las variables.
- El método del gradiente consiste en realizar sucesivas búsquedas unidimensionales en la dirección del vector gradiente. Presenta una buena rapidez de convergencia cuando las curvas de nivel son aproximadamente circulares pero se hace muy lento en presencia de valles estrechos.
- Si  $\mathbf{A}$  es una matriz cuadrada de orden  $n$  positiva (negativa) definida, las direcciones  $\mathbf{d}_i$  y  $\mathbf{d}_j$  se llaman conjugadas respecto a  $\mathbf{A}$ , o también  $\mathbf{A}$ -ortogonales, si  $\mathbf{d}_i^T \mathbf{A} \mathbf{d}_j = 0$ .
- Si  $f(\mathbf{x})$  es una función cuadrática con punto de extremo  $\mathbf{x}^*$  y  $\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_{n-1}$  son direcciones conjugadas de su matriz  $\mathbf{A}$ , entonces el punto  $\mathbf{x}^*$  se halla mediante  $n$  búsquedas unidimensionales sucesivas en las direcciones conjugadas.
- El método de Powell consiste en realizar búsquedas unidimensionales en ciertas direcciones  $\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_{n-1}$ , que, aunque inicialmente pueden ser arbitrarias, en el transcurso del proceso se van convirtiendo en un conjunto de direcciones conjugadas, si la función que se optimiza es cuadrática. Se trata de un método muy eficiente aunque algo difícil de programar.
- Se llama *simplex* de un espacio  $n$  dimensional, al elemento con la misma dimensión del espacio que posea la estructura geométrica más simple posible. Así, en el espacio bidimensional (el plano), el simplex es un triángulo equilátero, el cual está limitado por tres

segmentos iguales y posee tres puntos (vértices) que lo determinan y en el espacio tridimensional, el simplex es un tetraedro (pirámide de base triangular) regular, que está limitado por cuatro triángulos equiláteros y posee cuatro puntos (vértices) que lo determinan.

- El método del simplex secuencial se basa en evaluar la función a optimizar en los vértices de un simplex y, de acuerdo con los valores de la misma, formar otro simplex, sustituyendo uno de los vértices por un nuevo vértice, siguiendo dos reglas: 1) el vértice donde la función alcanza su peor valor es sustituido por uno nuevo, a menos que, 2) el peor vértice sea precisamente el último que entró a formar parte del simplex, en cuyo caso se elimina el vértice en el que la función alcanza el segundo peor valor. El método se detiene cuando el vértice donde la función alcanza el óptimo se mantiene muchas veces formando parte del simplex.

## Auto examen

1. ¿Qué es un problema de optimización?
2. Cite algunos factores que limitan los procedimientos analíticos para optimizar funciones.
3. ¿Qué se entiende por función unimodal de una variable?
4. Explique cómo se utilizan en el método de la sección áurea las dos propiedades fundamentales de los números  $D_n$ :  $D_n = D_{n-1} + D_{n-2}$  ( $n = 2, 3, 4, \dots$ ) y  $D_n = r D_{n-1}$  ( $n = 1, 2, 3, \dots$ ).
5. Halle el valor de  $x > 0$  más pequeño donde la función  $x^{\text{sen}x}$  posee un máximo.
6. Dos de los métodos más eficientes para optimizar funciones de varias variables son el método de Powell y el del simplex secuencial. Explique el fundamento de cada uno y analice las ventajas y desventajas relativas entre ellos.
7. Halle la mínima distancia entre la recta  $x + y = 4$  y la cardiode  $\rho = 3 - 3\cos\theta$ . Obtenga el resultado con 3 cifras decimales exactas.