Optimización Multidimensional no restringida 2022

Prof.: Dr. Juan Ignacio Manassaldi

JTP: Ing. Amalia Rueda

Repaso

- Decimos que $\underline{x}^*$ es un mínimo local si:

$$\underline{x}^* \in S \ tal \ que \ f\left(\underline{x}^*\right) \langle f\left(\underline{x}^* + \underline{\delta}\right)$$

- $\underline{x}^*$ es un mínimo global si:

$$\underline{x}^* \in S \ tal \ que \ f\left(\underline{x}^*\right) \langle f\left(\underline{x}\right) \forall \underline{x} \in S$$

- Decimos que $\underline{x}^*$ es un máximo local si:

$$\underline{x}^* \in S \ tal \ que \ f\left(\underline{x}^*\right) \rangle f\left(\underline{x}^* + \underline{\delta}\right)$$

- $\underline{x}^*$ es un máximo global si:

$$\underline{x}^* \in S \ tal \ que \ f\left(\underline{x}^*\right) \rangle f\left(\underline{x}\right) \forall \underline{x} \in S$$

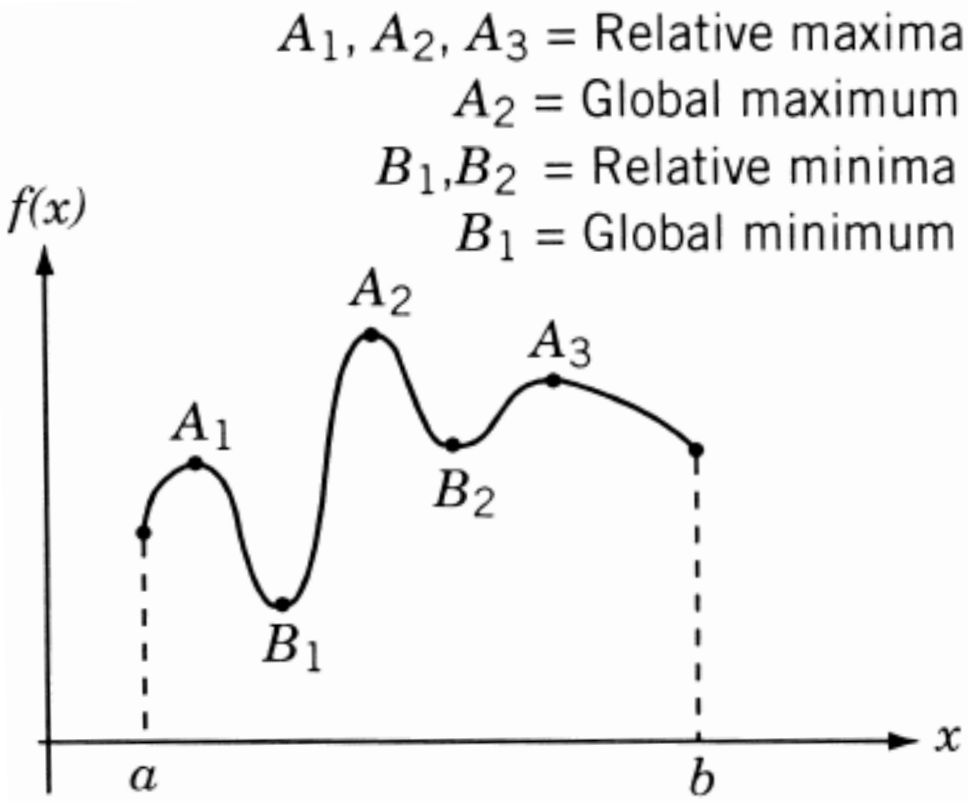
Repaso

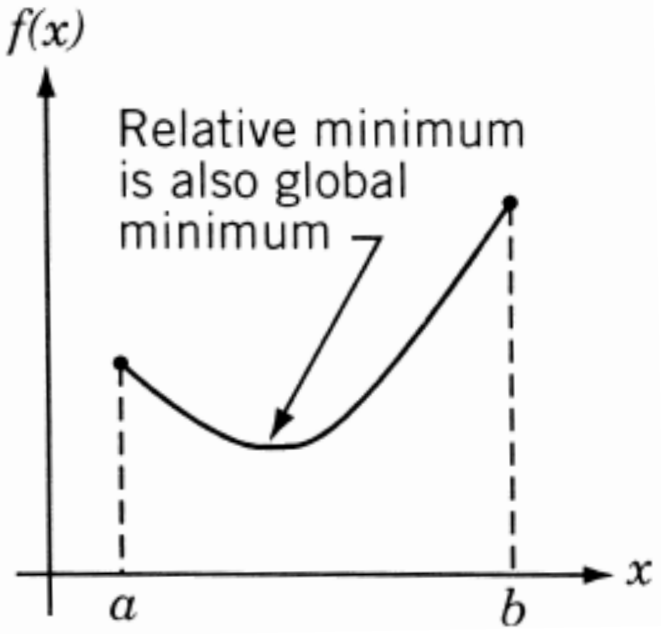
- Si tenemos algún punto $\underline{x}^*$ del dominio tal que: $\underline{\nabla} f\left(\underline{x}^*\right) = \underline{0}$

  Entonces *$\underline{x}^*$ es un punto crítico*, esto es, puede ser:

  1) Un mínimo local
  2) Un máximo local
  3) Un punto de ensilladura

Repaso


A_1, A_2, A_3 = Relative maxima
A_2 = Global maximum
B_1, B_2 = Relative minima
B_1 = Global minimum
f(x)
A_1
A_2
A_3
B_1
B_2
a
b
x


f(x)
Relative minimum is also global minimum
a
b
x

Repaso

Serie de Taylor en una variable: $f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^k(x_0)(x - x_0)^k$

Aproximación de primer orden:

$$f(x, x_0) = f(x_0) + f'(x_0)(x - x_0)$$

Aproximación de segundo orden:

$$f(x, x_0) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2$$

Repaso

Serie de Taylor en multivariable: $f\left(\underline{x}\right) = f\left(\underline{x}_0\right) + \sum_{k=1}^{\infty} \frac{1}{k!} d^k f\left(\underline{x}\right)$

Se define el diferencial *r* de una función *f* como:

$$d^r f\left(\underline{x}\right) = \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n}\cdots\sum_{k=1}^{n}}_{r \text{ sumatorias}} \left(x - x_0\right)_i \left(x - x_0\right)_j \cdots \left(x - x_0\right)_k \frac{\partial^r f\left(\underline{x}_0\right)}{\partial x_i \partial x_j \cdots \partial x_k}$$

Para $r=1$:

$$df\left(\underline{x}\right) = \sum_{i=1}^{n} \left(x - x_0\right)_i \frac{\partial f\left(\underline{x}_0\right)}{\partial x_i}$$

$$df\left(\underline{x}\right) = \left(x - x_0\right)_1 \frac{\partial f\left(\underline{x}_0\right)}{\partial x_1} + \left(x - x_0\right)_2 \frac{\partial f\left(\underline{x}_0\right)}{\partial x_2} + \cdots + \left(x - x_0\right)_n \frac{\partial f\left(\underline{x}_0\right)}{\partial x_n}$$

Repaso

$$df\left(\underline{x}\right)=\left(x-x_0\right)_1\frac{\partial f\left(\underline{x}_0\right)}{\partial x_1}+\left(x-x_0\right)_2\frac{\partial f\left(\underline{x}_0\right)}{\partial x_2}+\cdots+\left(x-x_0\right)_n\frac{\partial f\left(\underline{x}_0\right)}{\partial x_n}$$

$$df\left(\underline{x}\right)=\left(\underline{x}-\underline{x}_0\right)^T\nabla f\left(\underline{x}_0\right)$$

$$\left(x-x_0\right)_1 \quad \left(x-x_0\right)_2 \quad \cdots \quad \left(x-x_0\right)_n$$

$$\frac{\partial f\left(\underline{x}_0\right)}{\partial x_1}$$

$$\frac{\partial f\left(\underline{x}_0\right)}{\partial x_2}$$

$$\vdots$$

$$\frac{\partial f\left(\underline{x}_0\right)}{\partial x_n}$$

Repaso

Para $r=2$: $$d^2 f\left(\underline{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(x-x_0\right)_i \left(x-x_0\right)_j \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_i \partial x_j}$$

$$d^2 f\left(\underline{x}\right) =$$

$$\left(x-x_0\right)_1\left(x-x_0\right)_1 \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_1 \partial x_1} + \left(x-x_0\right)_1\left(x-x_0\right)_2 \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_1 \partial x_2} + \cdots + \left(x-x_0\right)_1\left(x-x_0\right)_n \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_1 \partial x_n} +$$

$$+\left(x-x_0\right)_2\left(x-x_0\right)_1 \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_2 \partial x_1} + \left(x-x_0\right)_2\left(x-x_0\right)_2 \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_2 \partial x_2} + \cdots + \left(x-x_0\right)_2\left(x-x_0\right)_n \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_1 \partial x_n} +$$

$$\vdots$$

$$+\left(x-x_0\right)_n\left(x-x_0\right)_1 \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_n \partial x_1} + \left(x-x_0\right)_n\left(x-x_0\right)_2 \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_n \partial x_2} + \cdots + \left(x-x_0\right)_n\left(x-x_0\right)_n \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_n \partial x_n}$$

Repaso

$$d^2 f(\underline{x}) = (x-x_0)_1 (x-x_0)_1 \frac{\partial^2 f(\underline{x}_0)}{\partial x_1 \partial x_1} + (x-x_0)_1 (x-x_0)_2 \frac{\partial^2 f(\underline{x}_0)}{\partial x_1 \partial x_2} + \cdots + (x-x_0)_1 (x-x_0)_n \frac{\partial^2 f(\underline{x}_0)}{\partial x_1 \partial x_n} +$$

$$+ (x-x_0)_2 (x-x_0)_1 \frac{\partial^2 f(\underline{x}_0)}{\partial x_2 \partial x_1} + (x-x_0)_2 (x-x_0)_2 \frac{\partial^2 f(\underline{x}_0)}{\partial x_2 \partial x_2} + \cdots + (x-x_0)_2 (x-x_0)_n \frac{\partial^2 f(\underline{x}_0)}{\partial x_1 \partial x_n} +$$

$$\vdots$$

$$+ (x-x_0)_n (x-x_0)_1 \frac{\partial^2 f(\underline{x}_0)}{\partial x_n \partial x_1} + (x-x_0)_n (x-x_0)_2 \frac{\partial^2 f(\underline{x}_0)}{\partial x_n \partial x_2} + \cdots + (x-x_0)_n (x-x_0)_n \frac{\partial^2 f(\underline{x}_0)}{\partial x_n \partial x_n}$$

Matriz Hessiana

$$\begin{array}{c|ccc|c} \frac{\partial^2 f(\underline{x}_0)}{\partial x_1 \partial x_1} & \frac{\partial^2 f(\underline{x}_0)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\underline{x}_0)}{\partial x_1 \partial x_n} & (x-x_0)_1 \\ \frac{\partial^2 f(\underline{x}_0)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\underline{x}_0)}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f(\underline{x}_0)}{\partial x_2 \partial x_n} & (x-x_0)_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial^2 f(\underline{x}_0)}{\partial x_n \partial x_1} & \frac{\partial^2 f(\underline{x}_0)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\underline{x}_0)}{\partial x_n \partial x_n} & (x-x_0)_n \\ \hline (x-x_0)_1 & (x-x_0)_2 & \cdots & (x-x_0)_n & \end{array}$$

$$d^2 f(\underline{x}) = (\underline{x} - \underline{x}_0)^T \nabla^2 f(\underline{x}_0)(\underline{x} - \underline{x}_0)$$

Repaso

Aproximación de primer orden:

$$f\left(\underline{x},\underline{x}_0\right)=f\left(\underline{x}_0\right)+\left(\underline{x}-\underline{x}_0\right)^T\nabla f\left(\underline{x}_0\right)$$

Aproximación de segundo orden:

$$f\left(\underline{x},\underline{x}_0\right)=f\left(\underline{x}_0\right)+\left(\underline{x}-\underline{x}_0\right)^T\nabla f\left(\underline{x}_0\right)+\frac{1}{2}\left(\underline{x}-\underline{x}_0\right)^T\nabla^2 f\left(\underline{x}_0\right)\left(\underline{x}-\underline{x}_0\right)$$

Introducción

Los métodos que se han ideado se pueden clasificar en tres amplias categorías según el tipo de información que debe proporcionar el usuario:

- Métodos de búsqueda directa, que utilizan solo valores de función.

  Método de una variable a la vez

- Métodos de gradiente (primer orden), que requieren estimaciones de la primera derivada de $f(x)$.

  Descenso mas pronunciado

- Métodos de segundo orden, que requieren estimaciones de la primera y segunda derivada de $f(x)$.

  Método de Newton

Introducción - Búsqueda de línea

- La búsqueda de línea nos permite encontrar el valor mínimo (o máximo) que toma una función sobre una dirección determinada de búsqueda.
- Para realizar la búsqueda de línea sobre $f(\underline{x})$, necesitamos un valor perteneciente al dominio de la función $\underline{x}^0$ y una dirección de búsqueda $\underline{d}$.
- Partiendo del punto $\underline{x}^0$ y utilizando la dirección en el espacio $\underline{d}$ podemos encontrar el valor mínimo que toma la función mediante una optimización unidimensional.

$$\underline{x} = \underline{x}^0 + \alpha \underline{d}$$

Ecuación de todos los puntos sobre la dirección $\underline{d}$ desde $\underline{x}^0$

Búsqueda de línea

Ecuación de todos los puntos sobre la dirección $\underline{d}$ desde $\underline{x}^0$:

$$\underline{x} = \underline{x}^0 + \alpha \underline{d}$$

Para encontrar el valor mínimo sobre esta dirección se debe resolver el siguiente problema:

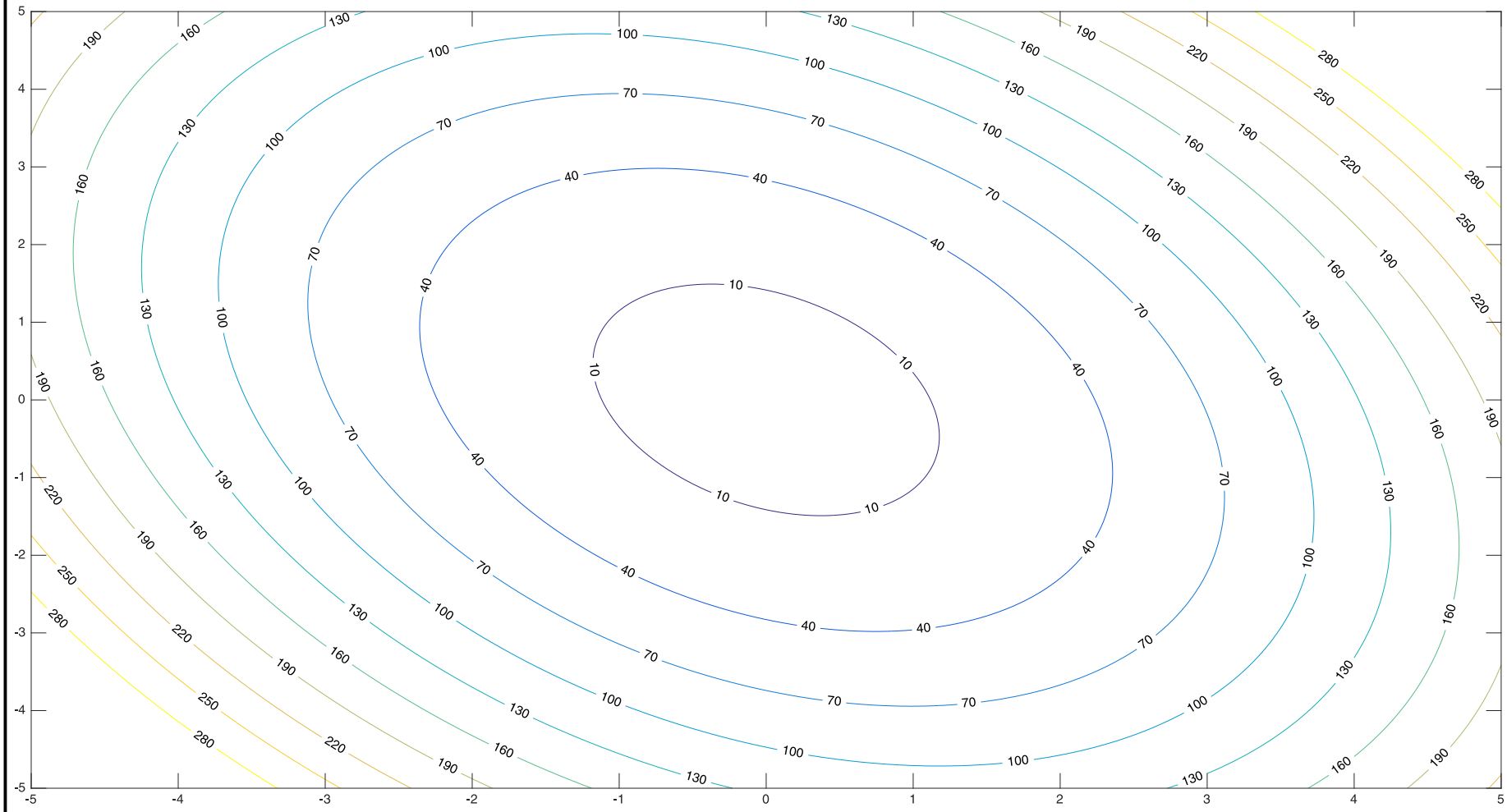
$$Min.\ f\left(\underline{x}^0 + \alpha \underline{d}\right)$$

¡Es un problema de optimización unidimensional!

Solo debo encontrar el α que minimiza f

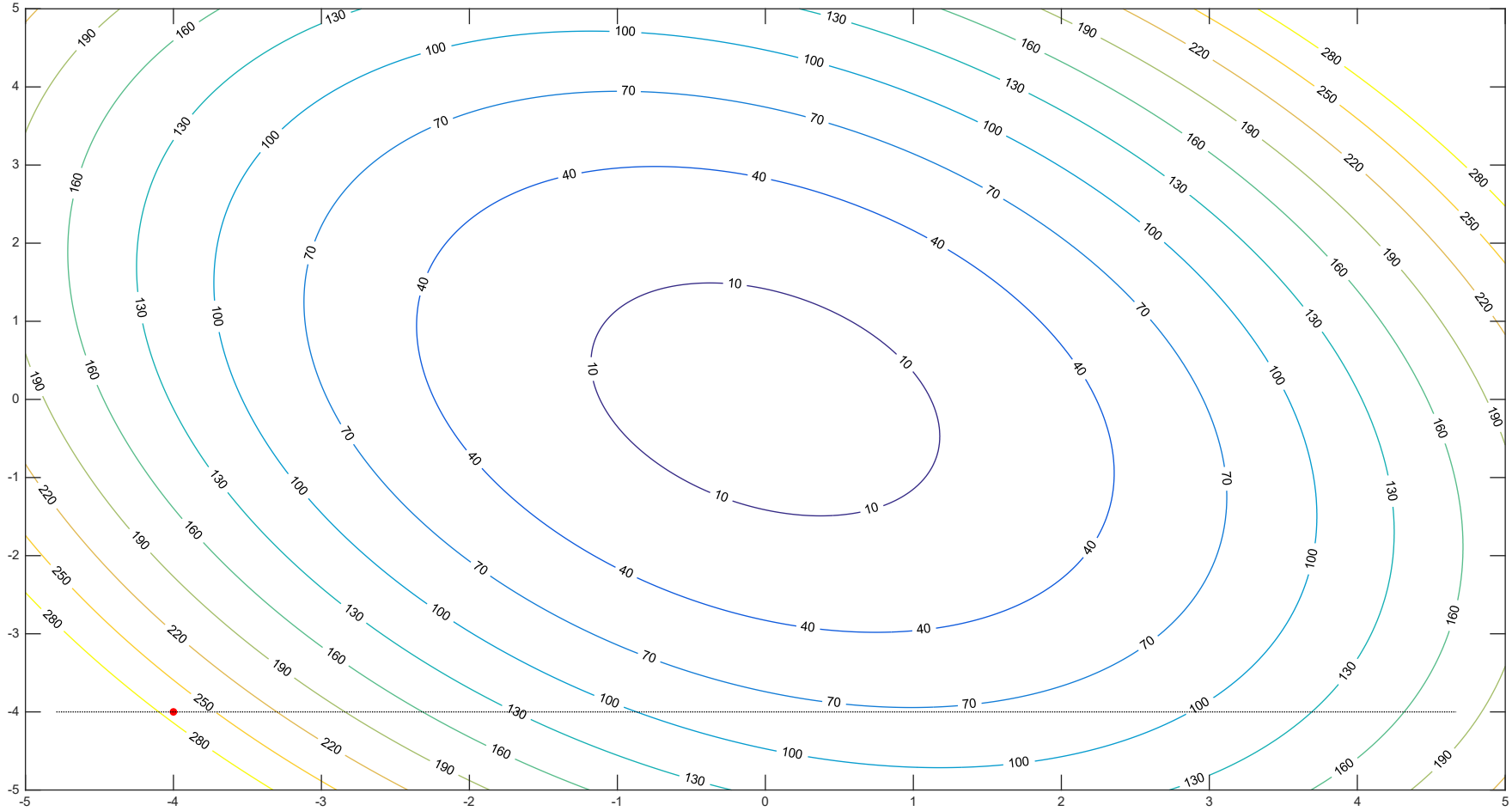
Búsqueda de línea - Ejemplo

$$f\left(\underline{x}\right) = 8x_1^{\,2} + 4x_1x_2 + 5x_2^{\,2}$$



Búsqueda de línea - Ejemplo

$$f(\underline{x}) = 8x_1^2 + 4x_1x_2 + 5x_2^2 \qquad \underline{x}^0 = \begin{bmatrix} -4 \\ -4 \end{bmatrix} \qquad \underline{d} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



Búsqueda de línea - Ejemplo

$$\underline{x}^0 = \begin{bmatrix} -4 \\ -4 \end{bmatrix} \quad \underline{d} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \rightarrow \underline{X} = \underline{X}^0 + \alpha \underline{d}$$

$$\underline{X} = \begin{bmatrix} -4 \\ -4 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} \rightarrow \underline{X} = \begin{bmatrix} -4+\alpha \\ -4 \end{bmatrix}$$

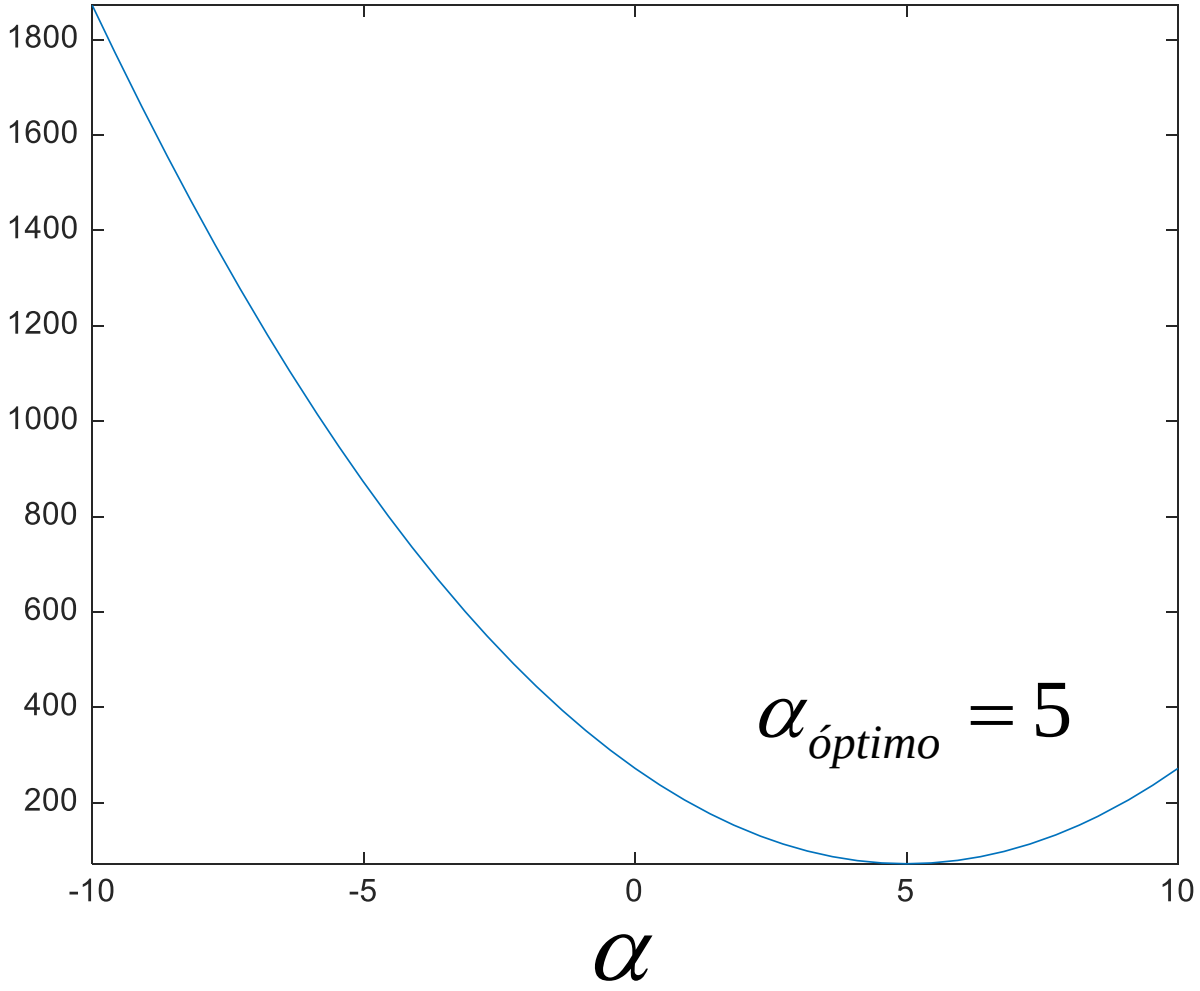
$$Min.\ f\left(\underline{x}^0 + \alpha \underline{d}\right) \qquad f\left(\underline{x}\right) = 8x_1^{\,2} + 4x_1x_2 + 5x_2^{\,2}$$

$$Min\ 8\left(-4+\alpha\right)^2 + 4\left(-4+\alpha\right)\left(-4\right) + 5\left(-4\right)^2$$

Búsqueda de línea - Ejemplo

$$Min\ 8(-4+\alpha)^2+4(-4+\alpha)(-4)+5(-4)^2$$

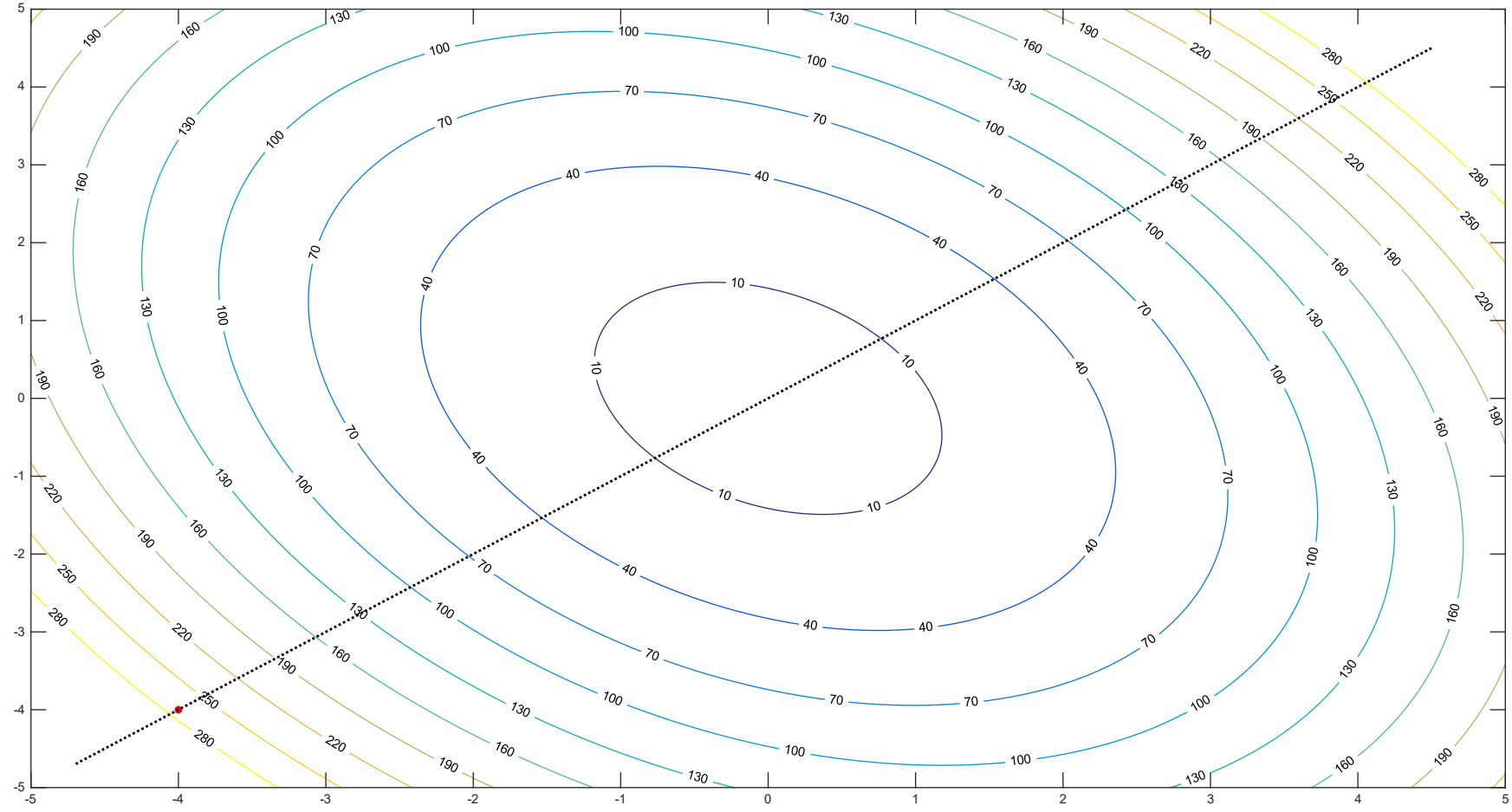
$$f\left(\underline{x}^0+\alpha\underline{d}\right)$$



Búsqueda de línea - Propuesta

$$\underline{x}^0 = \begin{bmatrix} -4 \\ -4 \end{bmatrix} \quad \underline{d} = \begin{bmatrix} \cos(\pi/4) \\ sen(\pi/4) \end{bmatrix}$$

$$Min.\ f(\underline{x}) = 8x_1^2 + 4x_1x_2 + 5x_2^2$$



Método directo: “Una variable a la vez”

- En el método de una variable a la vez, buscamos el mínimo a lo largo de las direcciones paralelas a los ejes.
- Este proceso se realiza de manera iterativa hasta que no sea posible una mejora en la FO.
- El método de optimización unidimensional elegido influye en el desempeño del método.
- Este método puede no converger y en algunos casos su convergencia resultar muy lenta.
- Estos problemas se pueden evitar cambiando las direcciones de búsqueda de una manera favorable en lugar de mantenerlas siempre paralelas a los ejes de coordenadas.

Una variable a la vez - Algoritmo

$$\underline{x}^{(0)} \in D \subseteq \mathbb{R}^n;\ \underline{d}_1, \underline{d}_2, \ldots \underline{d}_n;\ k = 1$$

$$\underline{x}^{(k)} = \underline{x}^{(k-1)}$$

$i = 1\ a\ n$

$$\alpha^* = Min.\ f\left(\underline{x}^{(k)} + \alpha \underline{d}_i\right)$$

$$\underline{x}^{(k)} = \underline{x}^{(k)} + \alpha^* \underline{d}_i$$

$$\left\| \underline{x}^{(k)} - \underline{x}^{(k-1)} \right\| \leq tol$$

si

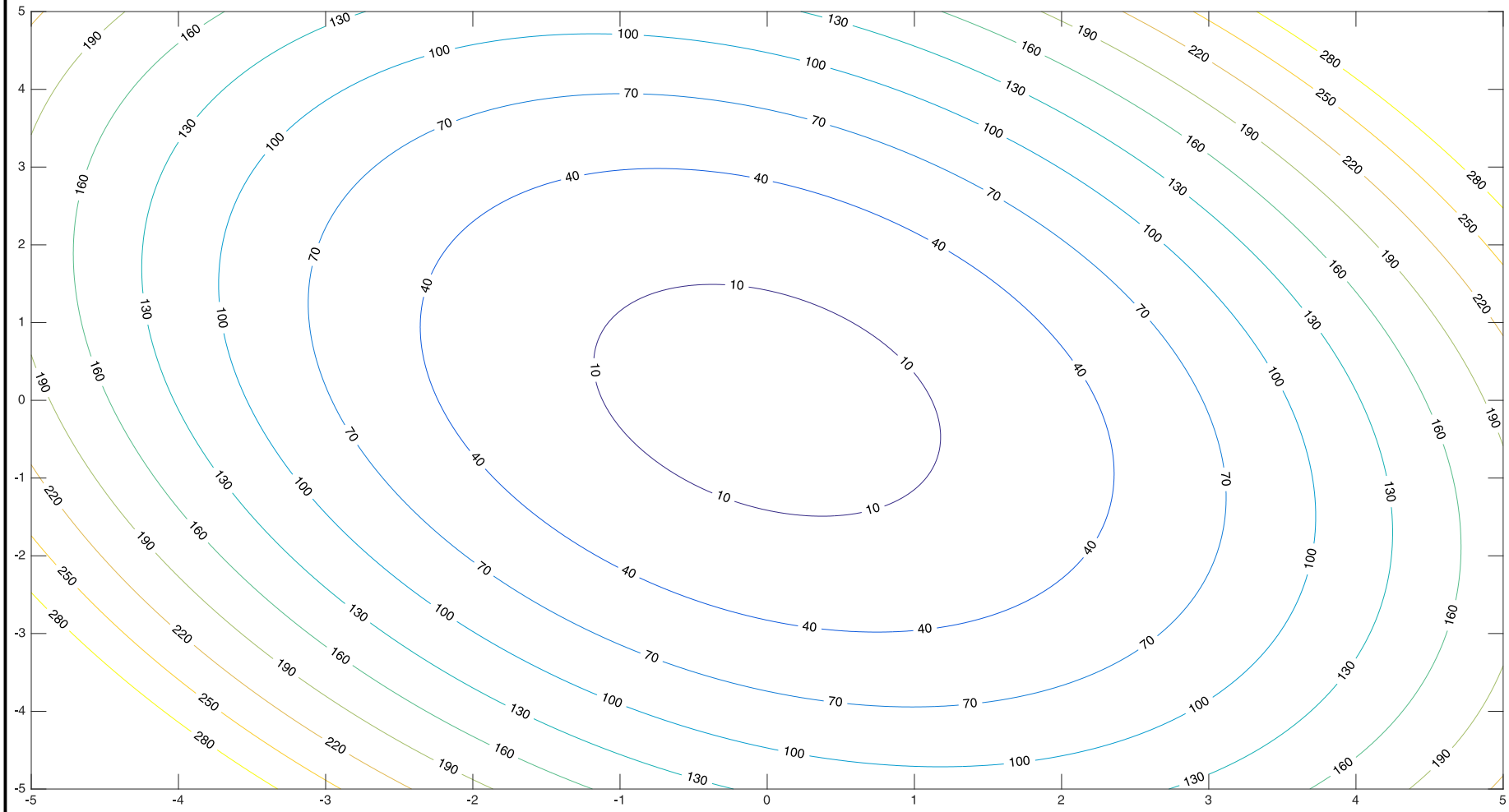
$$\underline{x}^* = \underline{x}^{(k)}$$

no

$$k = k + 1$$

Una variable a la vez - Ejemplo

$$Min.\ f(\underline{x}) = 8x_1^{\,2} + 4x_1x_2 + 5x_2^{\,2} \qquad \underline{x}^{(0)} = \begin{bmatrix} -4 \\ -4 \end{bmatrix}$$



Una variable a la vez - Ejemplo

$$Min.\ \ f(\underline{x}) = 8x_1^{\,2} + 4x_1x_2 + 5x_2^{\,2}$$

$k = 1$

$$\underline{x}^{(0)} = \begin{bmatrix} -4 \\ -4 \end{bmatrix}$$

$$\underline{d}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\underline{d}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\underline{x}^{(1)} = \underline{x}^{(0)} = \begin{bmatrix} -4 \\ -4 \end{bmatrix}$$

$i = 1$

$$\alpha = Min.\ f\left(\underline{x}^{(1)} + \alpha \underline{d}_1\right) \rightarrow \alpha = 5$$

$$\underline{x}^{(1)} = \underline{x}^{(1)} + \alpha \underline{d}_1 \rightarrow \underline{x}^{(1)} = \begin{bmatrix} -4 \\ -4 \end{bmatrix} + 5 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$$

$i = 2$

$$\alpha = Min.\ f\left(\underline{x}^{(1)} + \alpha \underline{d}_2\right) \rightarrow \alpha = 3.6$$

$$\underline{x}^{(1)} = \underline{x}^{(1)} + \alpha \underline{d}_2 \rightarrow \underline{x}^{(1)} = \begin{bmatrix} 1 \\ -4 \end{bmatrix} + 3.6 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.4 \end{bmatrix}$$

$$\underline{x}^{(0)} = \begin{bmatrix} -4 \\ -4 \end{bmatrix};\ \underline{x}^{(1)} = \begin{bmatrix} 1 \\ -0.4 \end{bmatrix} \rightarrow \begin{cases} f\left(\underline{x}^{(0)}\right) = 272 \\ f\left(\underline{x}^{(1)}\right) = 7.2 \end{cases}$$

¿? ¡Seguimos!

Una variable a la vez - Ejemplo

$k = 2$

$$\underline{x}^{(1)} = \begin{bmatrix} 1 \\ -0.4 \end{bmatrix}$$

$$\underline{d}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\underline{d}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\underline{x}^{(2)} = \underline{x}^{(1)} = \begin{bmatrix} 1 \\ -0.4 \end{bmatrix}$$

$i = 1$

$$\alpha = Min.\ f\left(\underline{x}^{(2)} + \alpha \underline{d}_1\right) \rightarrow \alpha = -0.9$$

$$\underline{x}^{(2)} = \underline{x}^{(2)} + \alpha \underline{d}_1 \rightarrow \underline{x}^{(2)} = \begin{bmatrix} 1 \\ -0.4 \end{bmatrix} - 0.9 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.1 \\ -0.4 \end{bmatrix}$$

$i = 2$

$$\alpha = Min.\ f\left(\underline{x}^{(2)} + \alpha \underline{d}_2\right) \rightarrow \alpha = 0.36$$

$$\underline{x}^{(2)} = \underline{x}^{(2)} + \alpha \underline{d}_2 \rightarrow \underline{x}^{(2)} = \begin{bmatrix} 0.1 \\ -0.4 \end{bmatrix} + 0.36 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.1 \\ -0.04 \end{bmatrix}$$

$$\underline{x}^{(1)} = \begin{bmatrix} 1 \\ -0.4 \end{bmatrix};\ \underline{x}^{(2)} = \begin{bmatrix} 0.1 \\ -0.04 \end{bmatrix} \rightarrow \begin{cases} f\left(\underline{x}^{(1)}\right) = 7.2 \\ f\left(\underline{x}^{(2)}\right) = 0.072 \end{cases}$$

¿ ? ¡Seguimos!

Una variable a la vez - Ejemplo

$k = 3$

$$\underline{x}^{(2)} = \begin{bmatrix} 0.1 \\ -0.04 \end{bmatrix}$$

$$\underline{d}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\underline{d}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\underline{x}^{(3)} = \underline{x}^{(2)} = \begin{bmatrix} 0.1 \\ -0.04 \end{bmatrix}$$

$i = 1$

$$\alpha = Min.\ f\left(\underline{x}^{(3)} + \alpha \underline{d}_1\right) \rightarrow \alpha = -0.09$$

$$\underline{x}^{(3)} = \underline{x}^{(3)} + \alpha \underline{d}_1 \rightarrow \underline{x}^{(3)} = \begin{bmatrix} 0.1 \\ -0.04 \end{bmatrix} - 0.09 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.01 \\ -0.04 \end{bmatrix}$$

$i = 2$

$$\alpha = Min.\ f\left(\underline{x}^{(3)} + \alpha \underline{d}_2\right) \rightarrow \alpha = 0.036$$

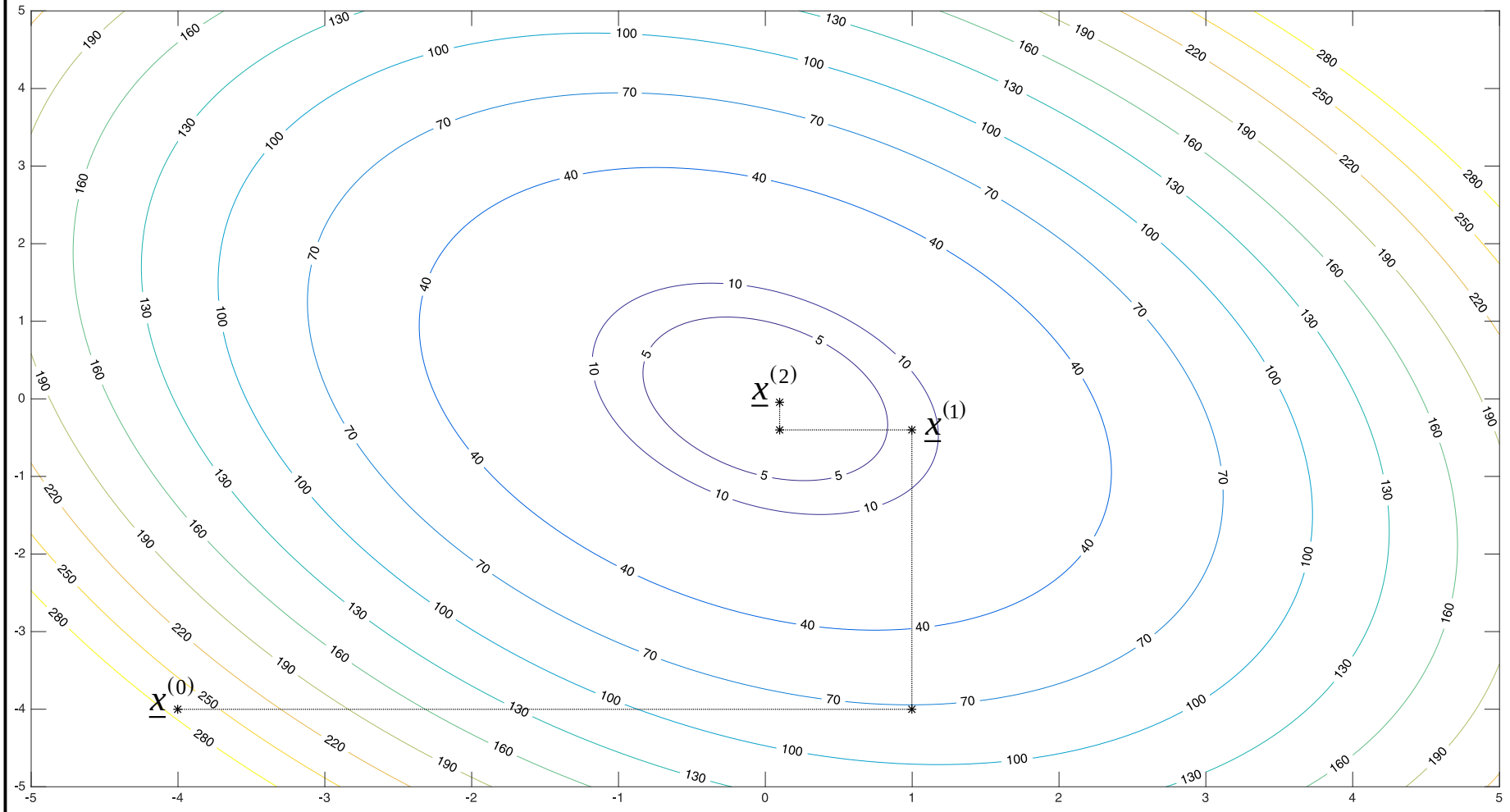
$$\underline{x}^{(3)} = \underline{x}^{(3)} + \alpha \underline{d}_2 \rightarrow \underline{x}^{(3)} = \begin{bmatrix} 0.01 \\ -0.04 \end{bmatrix} + 0.036 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.01 \\ -0.004 \end{bmatrix}$$

$$\underline{x}^{(2)} = \begin{bmatrix} 0.1 \\ -0.04 \end{bmatrix};\ \underline{x}^{(3)} = \begin{bmatrix} 0.01 \\ -0.004 \end{bmatrix} \rightarrow \begin{cases} f\left(\underline{x}^{(2)}\right) = 0.072 \\ f\left(\underline{x}^{(3)}\right) = 0.0072 \end{cases}$$

¿?

¡Sigan!

Una variable a la vez - Ejemplo



Métodos basados en el gradiente

- Los métodos directo requieren solo valores de la función objetivo para avanzar hacia la solución.
- Los métodos directos son importantes, porque muy a menudo, en un problema práctico de ingeniería, esta es la única información confiable disponible.
- Por otro lado, incluso los mejores métodos directos pueden requerir un número excesivo de evaluaciones de la función para localizar la solución.
- El ítem anterior, combinado con el deseo natural de buscar puntos críticos, motiva a considerar métodos que emplean información del gradiente.
- Se supone en todo momento que $f(\underline{x}), \nabla f(\underline{x})$ y $\nabla^2 f(\underline{x})$ existen y son continuas.

Métodos basados en el gradiente

- Todos los métodos que se presentaremos emplean un procedimiento de iteración similar, el nuevo punto se obtiene a partir de una dirección de búsqueda:

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha^k \, \underline{s}\left(\underline{x}^{(k)}\right)$$

$$\underline{x}^{(k)} : \textit{Estimación actual de } \underline{x}^*$$

$$\alpha^k : \textit{Tamaño del salto}$$

$$\underline{s}\left(\underline{x}^{(k)}\right) : \textit{Dirección de busqueda en el espacio } \mathbb{R}^n$$

- En general se busca entonces encontrar el tamaño del salto en la dirección de $\underline{s}$ que minimiza el valor de la función objetivo (búsqueda de línea).

Método del descenso mas pronunciado

- Supongamos que estamos en el punto $\underline{x}^{(k)}$ y deseamos determinar la dirección de “el mayor descenso local".
- A partir de la serie de Taylor, formamos una aproximación lineal a $f(x)$ a partir $\underline{x}^{(k)}$ eliminando los términos de orden 2 y superiores:

$$f\left(\underline{x}, \underline{x}^{(k)}\right) = f\left(\underline{x}^{(k)}\right) + \left(\underline{x} - \underline{x}^{(k)}\right)^T \nabla f\left(\underline{x}^{(k)}\right)$$

- Evaluando la aproximación de la función objetivo en el siguiente punto $(k+1)$ del proceso iterativo se llega a:

$$f\left(\underline{x}^{(k+1)}, \underline{x}^{(k)}\right) = f\left(\underline{x}^{(k)}\right) + \left(\underline{x}^{(k+1)} - \underline{x}^{(k)}\right)^T \nabla f\left(\underline{x}^{(k)}\right)$$

- Suponemos que la aproximación lineal representa a la función, por lo que:

$$f\left(\underline{x}^{(k+1)}\right) = f\left(\underline{x}^{(k)}\right) + \left(\underline{x}^{(k+1)} - \underline{x}^{(k)}\right)^T \nabla f\left(\underline{x}^{(k)}\right)$$

Método del descenso mas pronunciado

- Por lo tanto, el cambio en la función objetivo corresponde a:

$$f\left(\underline{x}^{(k+1)}\right) - f\left(\underline{x}^{(k)}\right) = \left(\underline{x}^{(k+1)} - \underline{x}^{(k)}\right)^T \nabla f\left(\underline{x}^{(k)}\right)$$

- Siguiendo la estrategia de resolución propuesta:

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha^k \underline{s}\left(\underline{x}^{(k)}\right) \Rightarrow \left(\underline{x}^{(k+1)} - \underline{x}^{(k)}\right)^T = \alpha^k \underline{s}\left(\underline{x}^{(k)}\right)^T$$

- Reemplazando en la expresión anterior:

$$f\left(\underline{x}^{(k+1)}\right) - f\left(\underline{x}^{(k)}\right) = \alpha^k \underline{s}\left(\underline{x}^{(k)}\right)^T \nabla f\left(\underline{x}^{(k)}\right) \rightarrow \left\|\underline{s}\left(\underline{x}^{(k)}\right)\right\| \left\|\nabla f\left(\underline{x}^{(k)}\right)\right\| \cos\alpha$$

- Analizando, proponemos la siguiente dirección:

$$s\left(\underline{x}^{(k)}\right) = -\nabla f\left(\underline{x}^{(k)}\right)$$

Método del descenso mas pronunciado

- Luego, reemplazando la dirección propuesta:

$$f\left(\underline{x}^{(k+1)}\right) - f\left(\underline{x}^{(k)}\right) = -\alpha^k \nabla f\left(\underline{x}^{(k)}\right)^T \nabla f\left(\underline{x}^{(k)}\right)$$

$$f\left(\underline{x}^{(k+1)}\right) - f\left(\underline{x}^{(k)}\right) = -\alpha^k \left\|\nabla f\left(\underline{x}^{(k)}\right)\right\|^2$$

- El método mas simple basado en el gradiente consiste en tomar la dirección opuesta del gradiente y fijar α como un parámetro fijo positivo:

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} - \alpha \nabla f\left(\underline{x}^{(k)}\right)$$

- Esto tiene dos desventajas: la necesidad de hacer una elección adecuada para α y la lentitud inherente cerca del mínimo debido a que en esta zona el gradiente tiende a cero.

Método del descenso mas pronunciado

- El Método del descenso mas pronunciado propone realizar una optimización unidimensional para la elección de α en cada iteración:

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} - \alpha^{(k)} \nabla f\left(\underline{x}^{(k)}\right)$$

- Por lo tanto, en cada iteración se debe resolver el siguiente problema de optimización unidimensional:

$$\alpha^* = \min f\left(\underline{x}^{(k)} - \alpha \nabla f\left(\underline{x}^{(k)}\right)\right) \Rightarrow \underline{x}^{(k+1)} = \underline{x}^{(k)} - \alpha^* \nabla f\left(\underline{x}^{(k)}\right)$$

- Este método es más confiable que el método de gradiente simple, pero la convergencia sigue siendo lenta para la mayoría de los problemas prácticos.

Método simple (λ=cte)

Ejemplo: $Min.\ f(\underline{x}) = \left(x_1^{\,2} + x_2 - 11\right)^2 + \left(x_1 + x_2^{\,2} - 7\right)^2$

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} - \alpha \nabla f\left(\underline{x}^{(k)}\right) \qquad \nabla f(\underline{x}) = \begin{bmatrix} 2\left(x_1^{\,2} + x_2 - 11\right)(2x_1) + 2\left(x_1 + x_2^{\,2} - 7\right) \\ 2\left(x_1^{\,2} + x_2 - 11\right) + 2\left(x_1 + x_2^{\,2} - 7\right)(2x_2) \end{bmatrix}$$

$$\left.\begin{array}{l} \alpha = 1 \\ \underline{x}^{(0)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \end{array}\right\} \Rightarrow \underline{x}^{(1)} = \begin{bmatrix} 44 \\ 20 \end{bmatrix} \qquad \Rightarrow \underline{x}^{(2)} = \begin{bmatrix} -343150 \\ -38830 \end{bmatrix}$$

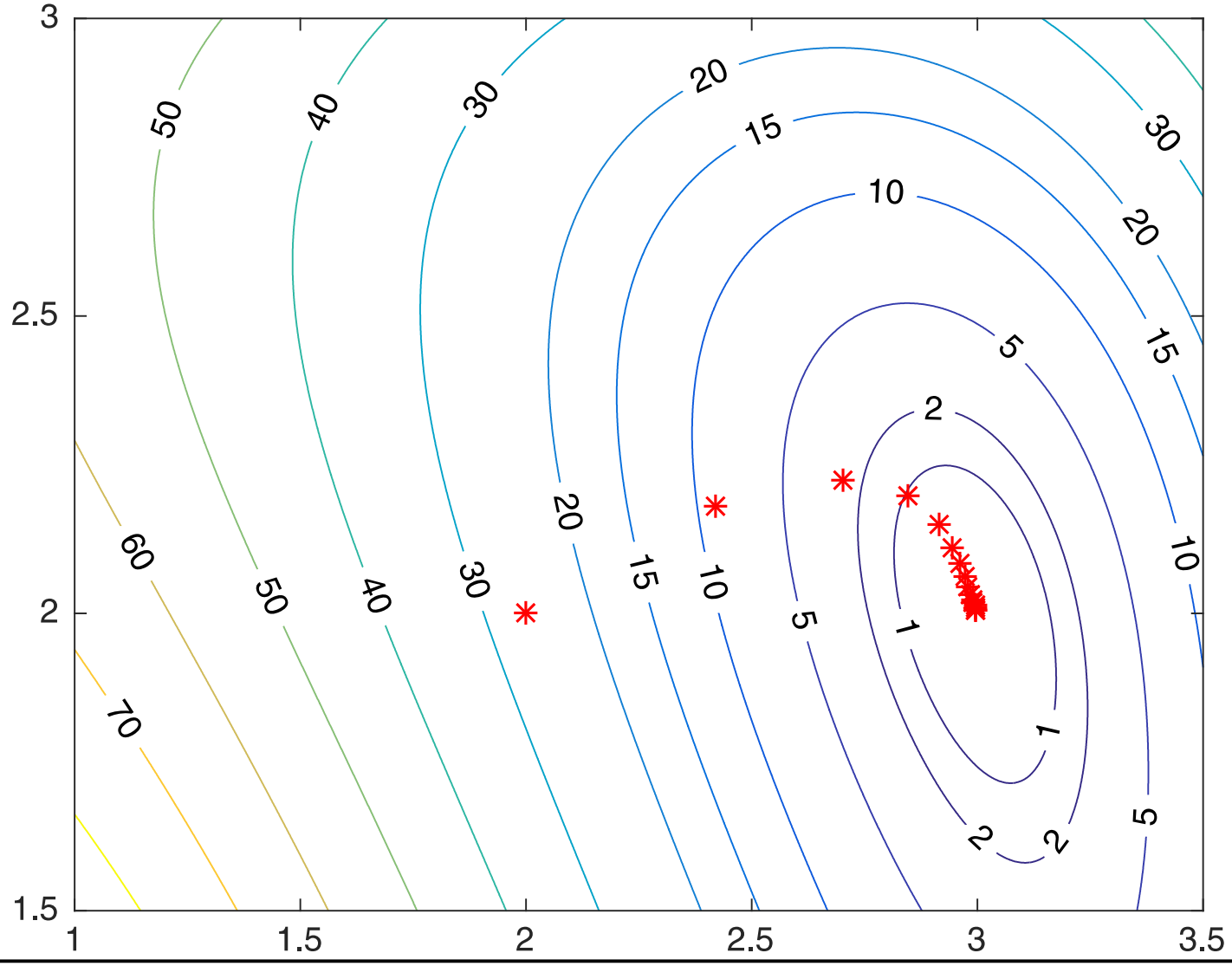
$$\left.\begin{array}{l} \alpha = 10^{-1} \\ \underline{x}^{(0)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \end{array}\right\} \Rightarrow \underline{x}^{(1)} = \begin{bmatrix} 6.2 \\ 3.8 \end{bmatrix} \qquad \Rightarrow \underline{x}^{(2)} = \begin{bmatrix} -74.003 \\ -23.181 \end{bmatrix}$$

$$\left.\begin{array}{l} \alpha = 10^{-2} \\ \underline{x}^{(0)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \end{array}\right\} \Rightarrow \underline{x}^{(1)} = \begin{bmatrix} 2.42 \\ 2.18 \end{bmatrix} \qquad \Rightarrow \underline{x}^{(2)} = \begin{bmatrix} 2.703 \\ 2.224 \end{bmatrix}$$

Matemática Superior Aplicada

$$Min.\ f(\underline{x}) = \left(x_1^{\,2} + x_2 - 11\right)^2 + \left(x_1 + x_2^{\,2} - 7\right)^2$$

UTN



$$\alpha = 10^{-2}$$

$$\underline{x}^{(0)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$\underline{x}^{(33)} = \begin{bmatrix} 2.999989 \\ 2.000025 \end{bmatrix}$$

$$tol = 10^{-5}$$

Método del descenso mas pronunciado

$$Min.\ \ f\left(\underline{x}\right)=\left(x_1{}^2+x_2-11\right)^2+\left(x_1+x_2{}^2-7\right)^2$$

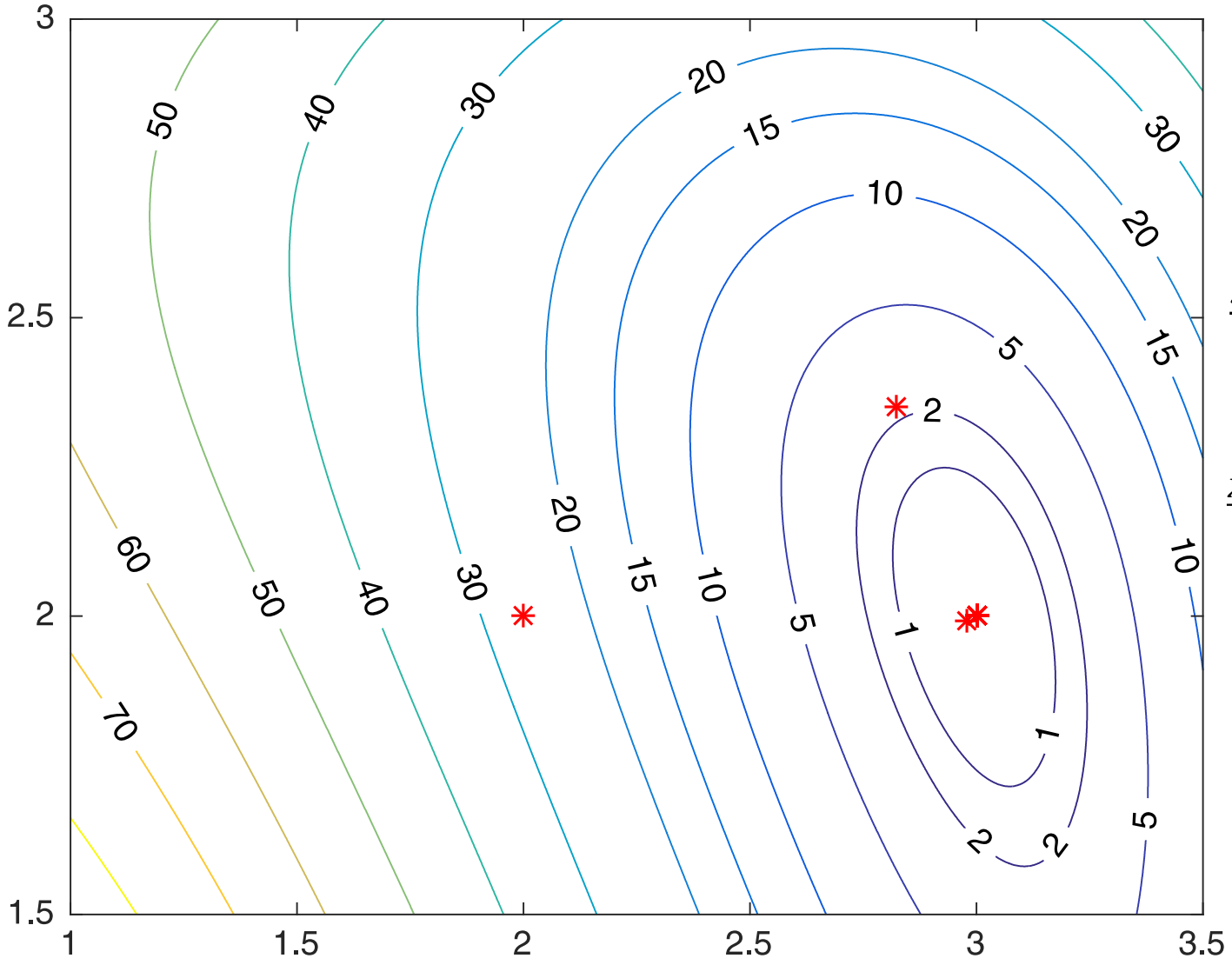
$$\nabla f\left(\underline{x}\right)=\begin{bmatrix}2\left(x_1{}^2+x_2-11\right)\left(2x_1\right)+2\left(x_1+x_2{}^2-7\right)\\2\left(x_1{}^2+x_2-11\right)+2\left(x_1+x_2{}^2-7\right)\left(2x_2\right)\end{bmatrix}$$

$$\underline{x}^{(0)}=\begin{bmatrix}2\\2\end{bmatrix}$$

$$\alpha^*=\min f\left(\underline{x}^{(0)}-\alpha\nabla f\left(\underline{x}^{(0)}\right)\right)$$

$$\underline{x}^{(1)}=\underline{x}^{(0)}-\alpha^*\nabla f\left(\underline{x}^{(0)}\right)$$

Método del descenso mas pronunciado



$$\underline{x}^{(0)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$\underline{x}^{(13)} = \begin{bmatrix} 2.999999 \\ 2.000002 \end{bmatrix}$$

$$tol = 10^{-5}$$

Método del descenso mas pronunciado - Algoritmo

$$\underline{x}^{(0)} \in D \subseteq \mathbb{R}^n;\ \nabla f\left(\underline{x}\right);\ k = 0$$

$$\underline{d} = -\nabla f\left(\underline{x}^{(k)}\right)$$

$$\alpha^* = Min.\ f\left(\underline{x}^{(k)} + \alpha \underline{d}\right)$$

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha^* \underline{d}$$

$$\left\|\alpha^* \underline{d}\right\| \leq tol$$

si

$$\underline{x}^* = \underline{x}^{(k+1)}$$

no

$$k = k + 1$$

Método de Newton

- El gradiente negativo apunta directamente hacia el mínimo solo cuando los contornos de $f(x)$ son circulares, por lo tanto el gradiente negativo no es una buena dirección global (en general) para funciones no lineales.
- El método del descenso mas pronunciado emplea aproximaciones lineales sucesivas al objetivo y requiere valores de la función objetivo y su gradiente en cada iteración.
- Un avance es considerar el uso de información de orden superior (derivadas segundas) en un esfuerzo por construir una estrategia de iteración más global.

Método de Newton

- A partir de la serie de Taylor, formamos una aproximación cuadrática a $f(x)$ eliminando los términos de orden 3 y superiores:

$$f\left(\underline{x}, \underline{x}^{(k)}\right) = f\left(\underline{x}^{(k)}\right) + \left(\underline{x} - \underline{x}^{(k)}\right)^T \nabla f\left(\underline{x}^{(k)}\right) + \tfrac{1}{2}\left(\underline{x} - \underline{x}^{(k)}\right)^T \nabla^2 f\left(\underline{x}^{(k)}\right)\left(\underline{x} - \underline{x}^{(k)}\right)$$

- Se propone una secuencia de iteración en la que el siguiente punto, es un punto en donde el gradiente de la aproximación es cero.

$$\nabla f\left(\underline{x}, \underline{x}^{(k)}\right) = \nabla f\left(\underline{x}^{(k)}\right) + \nabla^2 f\left(\underline{x}^{(k)}\right)\left(\underline{x} - \underline{x}^{(k)}\right)$$

$$\nabla f\left(\underline{x}^{(k)}\right) + \nabla^2 f\left(\underline{x}^{(k)}\right)\left(\underline{x}^{(k+1)} - \underline{x}^{(k)}\right) = 0$$

- Siguiendo la estrategia de resolución propuesta:

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha^k s\left(\underline{x}^{(k)}\right) \Rightarrow \left(\underline{x}^{(k+1)} - \underline{x}^{(k)}\right) = \alpha^k s\left(\underline{x}^{(k)}\right)$$

Método de Newton - I

- Reemplazando y asumimos α=1:

$$\nabla f\left(\underline{x}^{(k)}\right)+\nabla^2 f\left(\underline{x}^{(k)}\right)\alpha^k s\left(\underline{x}^{(k)}\right)=0 \Rightarrow s\left(\underline{x}^{(k)}\right)=-\left(\nabla^2 f\left(\underline{x}^{(k)}\right)\right)^{-1}\nabla f\left(\underline{x}^{(k)}\right)$$

- Por lo que finalmente:

$$\underline{x}^{(k+1)}=\underline{x}^{(k)}-\left(\nabla^2 f\left(\underline{x}^{(k)}\right)\right)^{-1}\nabla f\left(\underline{x}^{(k)}\right)$$

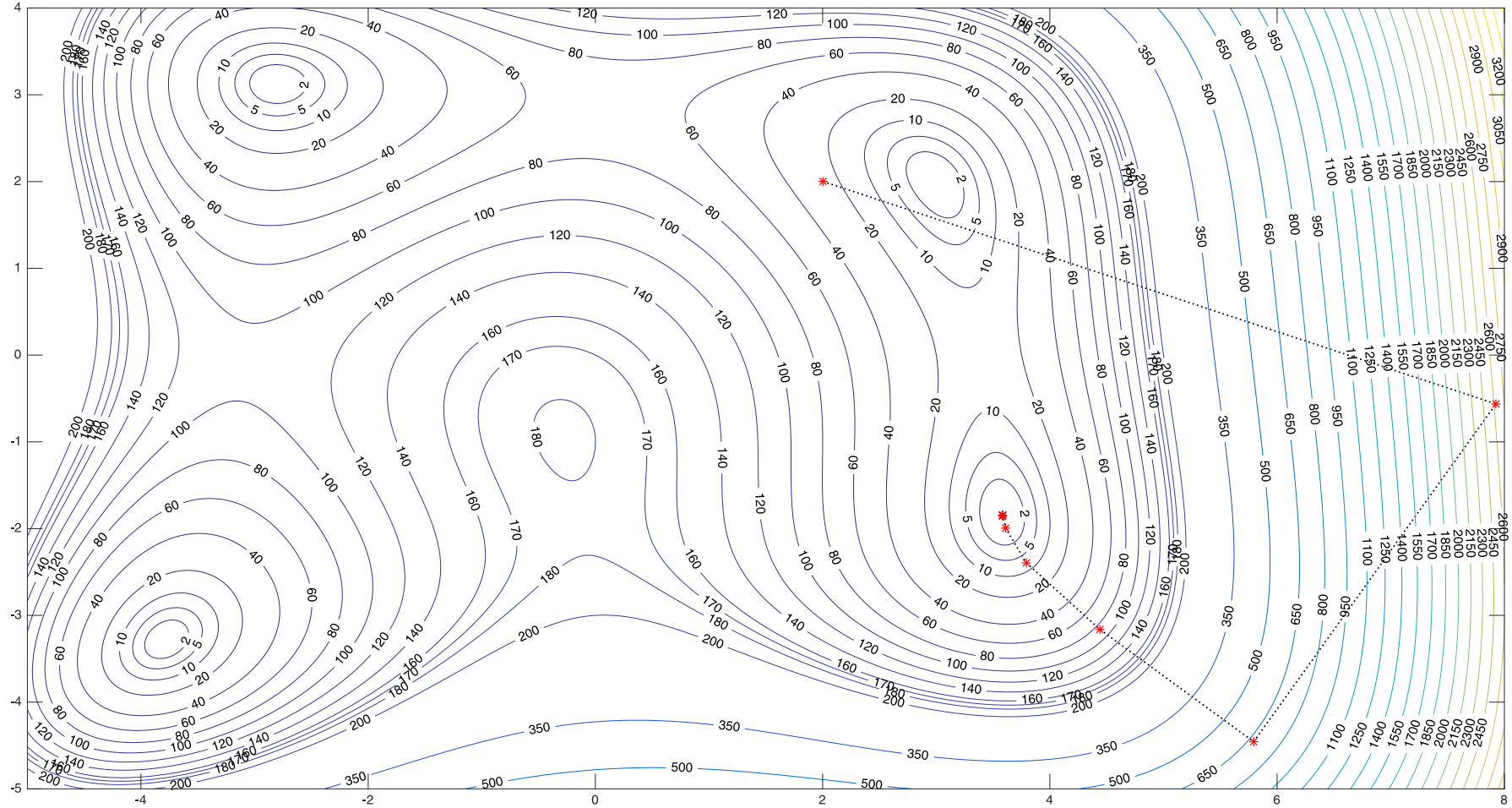
Método de Newton - I - Ejemplo

Ejemplo: $Min.\ f(\underline{x}) = \left(x_1^{\,2} + x_2 - 11\right)^2 + \left(x_1 + x_2^{\,2} - 7\right)^2$

$$\nabla f(\underline{x}) = \begin{bmatrix} 2\left(x_1^{\,2} + x_2 - 11\right)2x_1 + 2\left(x_1 + x_2^{\,2} - 7\right) \\ 2\left(x_1^{\,2} + x_2 - 11\right) + 2\left(x_1 + x_2^{\,2} - 7\right)2x_2 \end{bmatrix}$$

$$\nabla^2 f(\underline{x}) = \begin{bmatrix} 8x_1^{\,2} + 4\left(x_1^{\,2} + x_2 - 11\right) + 2 & 4x_1 + 4x_2 \\ 4x_1 + 4x_2 & 2 + 8x_2^{\,2} + 4\left(x_1 + x_2^{\,2} - 7\right) \end{bmatrix}$$

Método de Newton - I - Ejemplo



$$\underline{x}^{(0)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \Rightarrow \underline{x}^{(9)} = \begin{bmatrix} 3.584428 & -1.848126 \end{bmatrix}^T$$

$$tol = 10^{-5}$$

Método de Newton - II

- El método de Newton minimizará una función cuadrática (desde cualquier punto de partida) en exactamente un paso.
- La experiencia ha demostrado que el método de Newton puede ser poco confiable para funciones no cuadráticas.
- Es decir, el paso de Newton a menudo será grande cuando $x^{(0)}$ está lejos de x*, y existe la posibilidad real de divergencia.
- Es posible modificar el método de una manera lógica y simple para asegurar el descenso agregando una búsqueda de línea como en el método del descenso mas pronunciado.

$$\alpha^* = \min f\left(\underline{x}^{(k)} - \alpha^k \left(\nabla^2 f\left(\underline{x}^{(k)}\right)\right)^{-1} \nabla f\left(\underline{x}^{(k)}\right)\right)$$

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} - \alpha^* \left(\nabla^2 f\left(\underline{x}^{(k)}\right)\right)^{-1} \nabla f\left(\underline{x}^{(k)}\right)$$

Método de Newton - II

- El método de Newton II (o modificado) es confiable y eficiente cuando la primera y la segunda derivada se calculan de manera precisa y económica (desde el punto de vista computacional).
- La mayor dificultad es la necesidad de calcular y resolver en cada iteración el conjunto de ecuaciones lineales que involucra el termino de la matriz Hessiana.

Método de Newton - II

$$\underline{x}^{(0)} \in D \subseteq \mathbb{R}^n;\ \nabla f\left(\underline{x}\right);\ \nabla^2 f\left(\underline{x}^{(k)}\right);\ k = 0$$

$$\underline{d} = -\left(\nabla^2 f\left(\underline{x}^{(k)}\right)\right)^{-1} \nabla f\left(\underline{x}^{(k)}\right)$$

También puede resolverse como un SEL

$$\alpha^* = Min.\ f\left(\underline{x}^{(k)} + \alpha \underline{d}\right)$$

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha^* \underline{d}$$

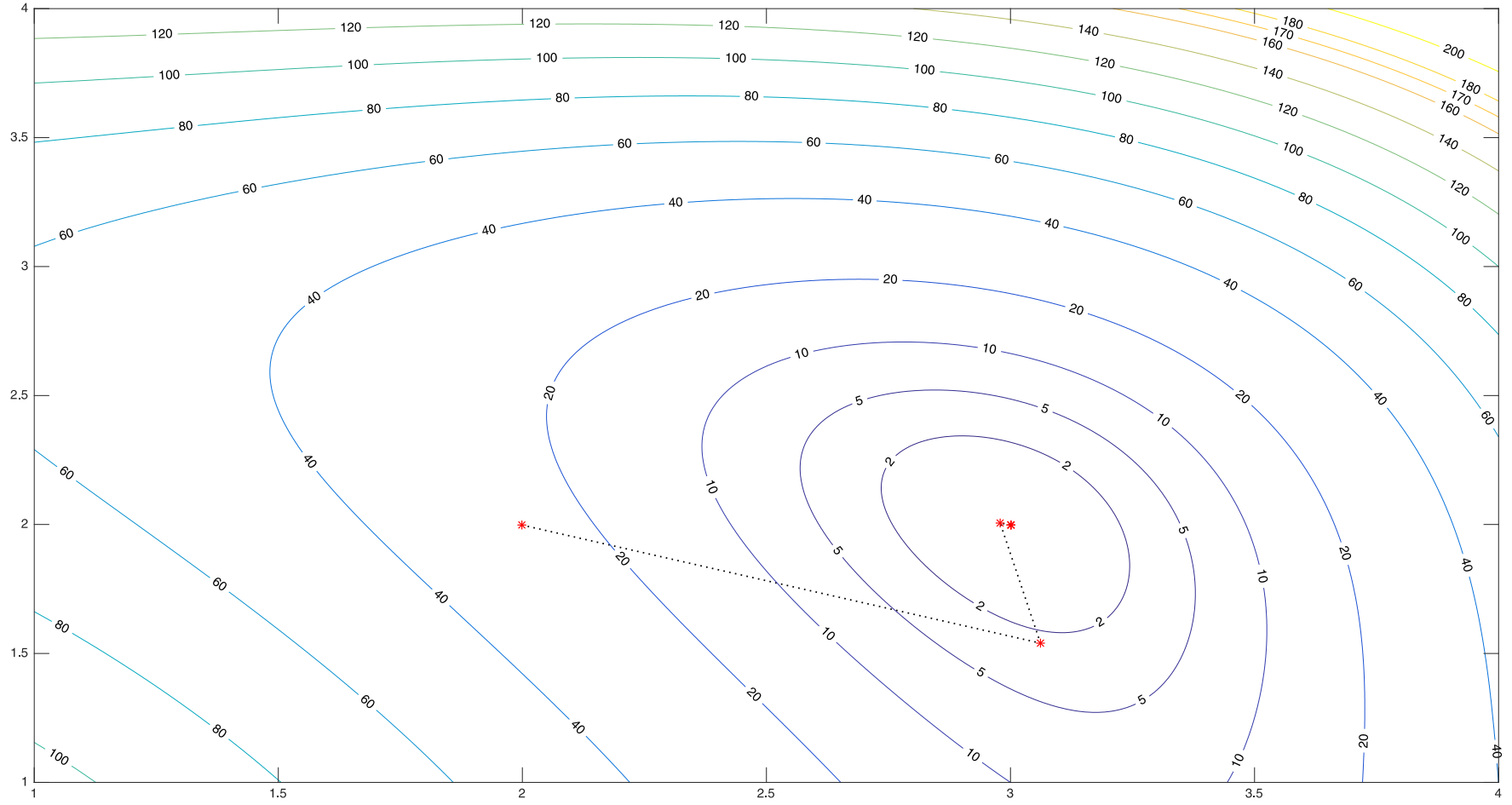
$$\left\|\alpha^* \underline{d}\right\| \leq tol$$

si

$$\underline{x}^* = \underline{x}^{(k+1)}$$

no

$$k = k + 1$$



$$\underline{x}^{(0)} = \begin{bmatrix} 2 & 2 \end{bmatrix}^T \Rightarrow \underline{x}^{(5)} = \begin{bmatrix} 3 & 2 \end{bmatrix}^T$$

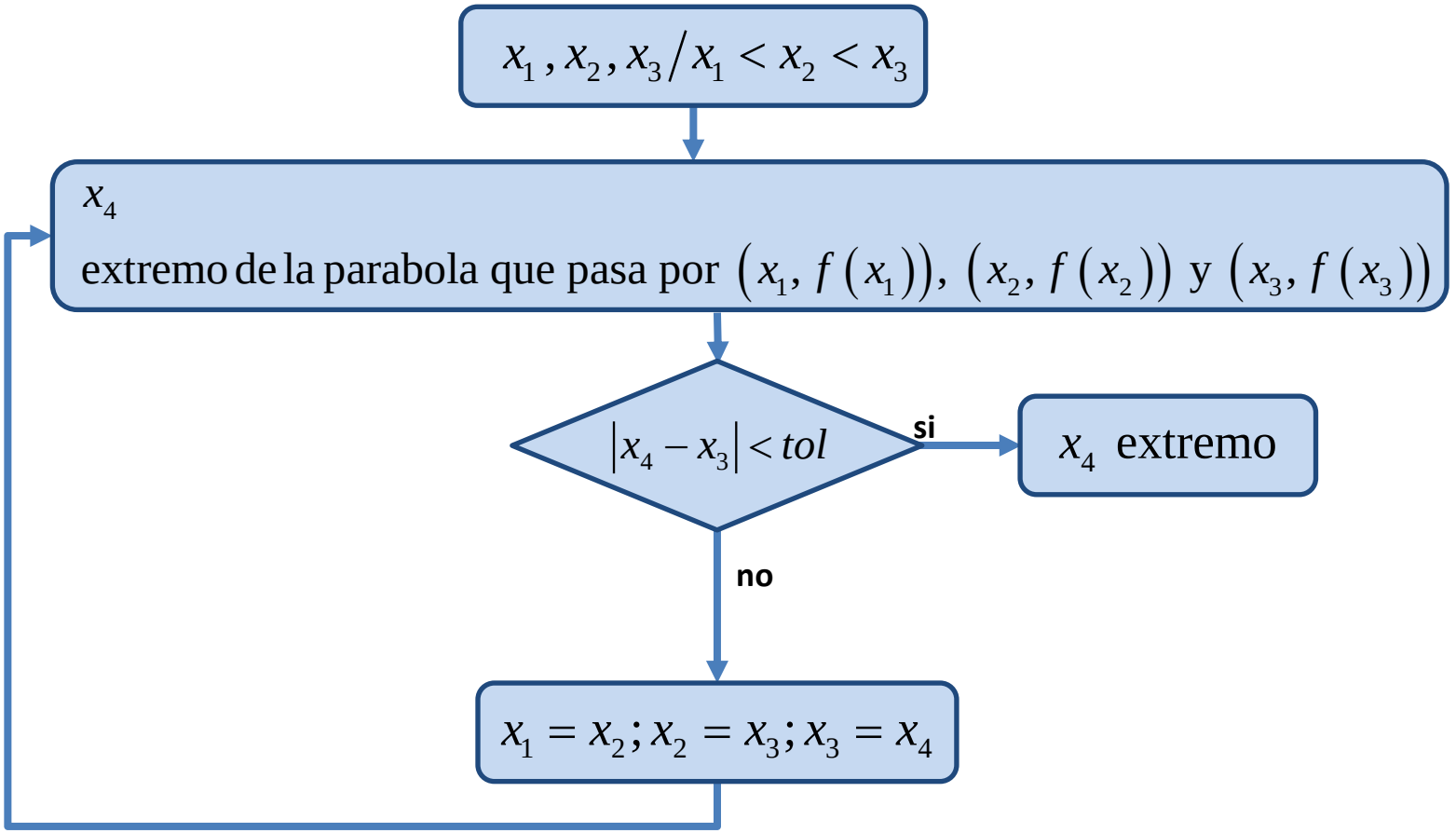
$$tol = 10^{-5}$$

Claves para implementar en Scilab

- Programar las “function” necesarias en las dimensiones correctas.
- Manejar las incógnitas como un único vector.
- Tener una buena implementación de los algoritmos de búsqueda unidimensional.
- Una buena herramienta es programar una “function” que realice la búsqueda de línea.

Volvemos a Optimización Unidimensional…

Método de interpolación parabólica sucesivas



Búsqueda de línea con MIPS

$$Min.\ f\left(\underline{x}^{(i)} + \alpha \underline{d}\right)$$

Función a minimizar

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

Solución actual

$\underline{x}^{(i)} \in \mathbb{R}^n$

Tamaño del salto

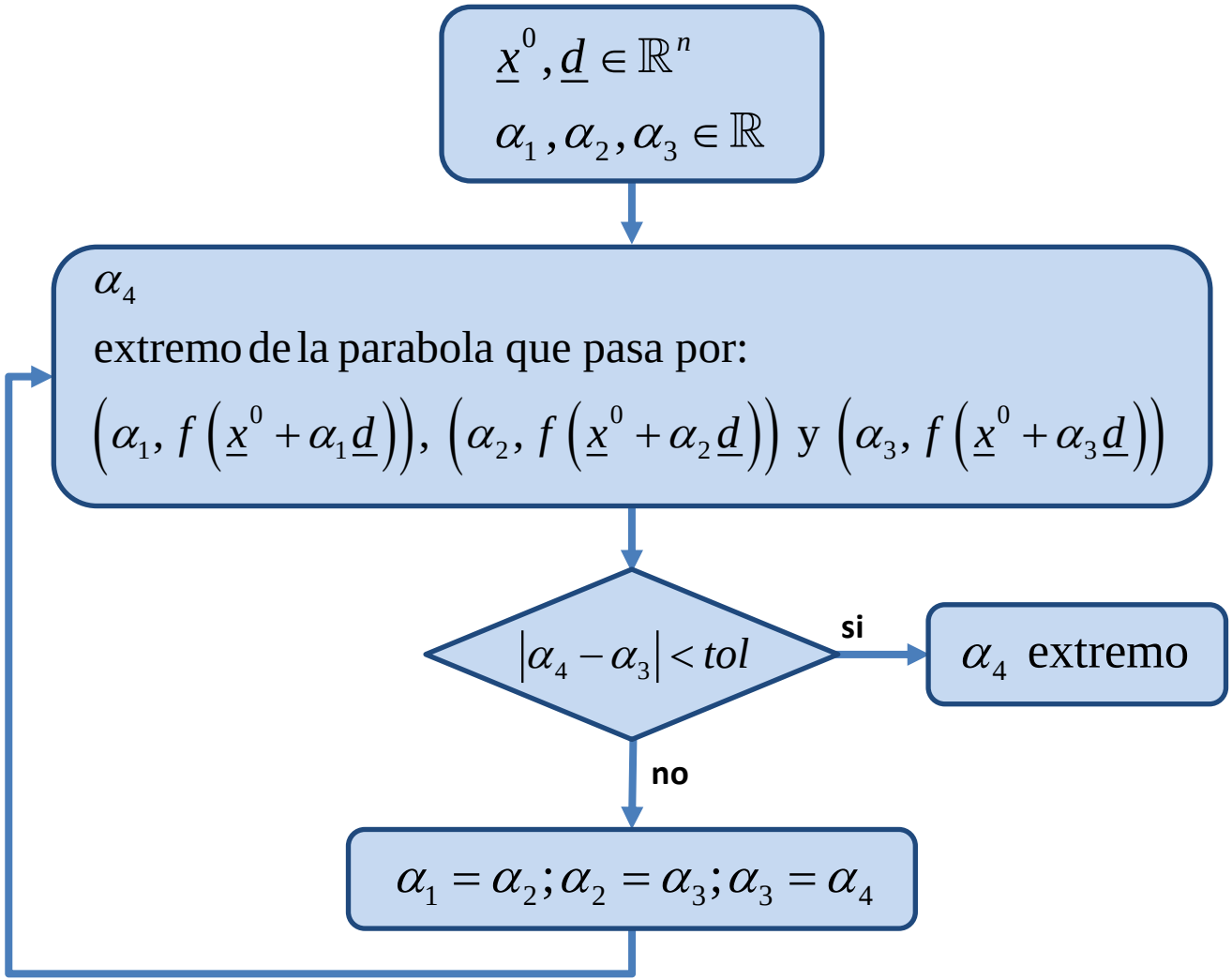
$\alpha \in \mathbb{R}$

Dirección de búsqueda

$\underline{d} \in \mathbb{R}^n$

Objetivo de la búsqueda de línea

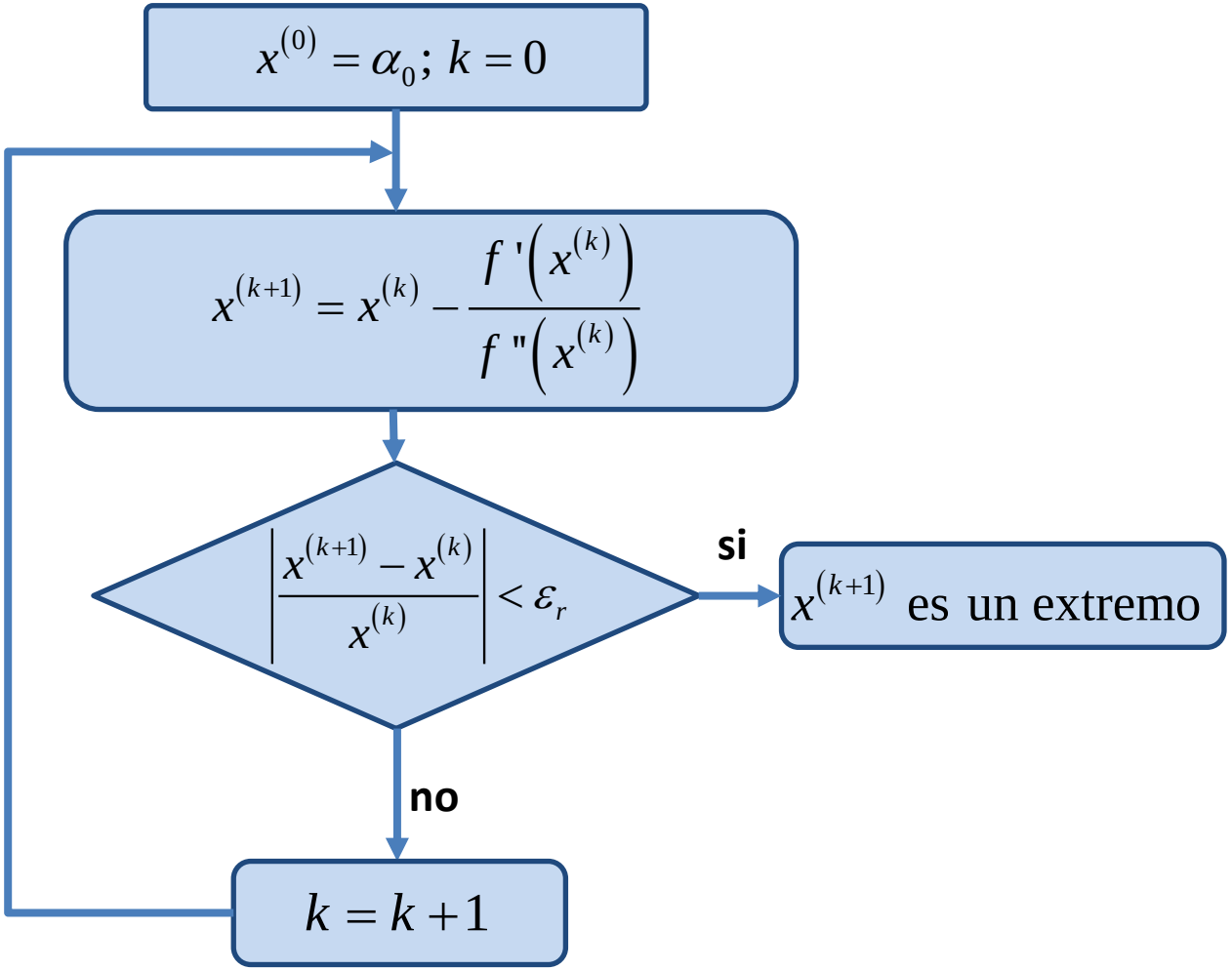
Búsqueda de línea con MIPS



Búsqueda de línea en Scilab con MIPS

```
function alfa=linesearchMIPS(fun, xk,d,a_0)
x1 = a_0*0.98;
x2 = a_0;
x3 = a_0*1.02;
for i= 1:100
        A=[x1^2 x1 1
           x2^2 x2 1
           x3^2 x3 1 ];
        b=[fun(xk + x1*d)
           fun(xk + x2*d)
           fun(xk + x3*d)];
        xx=A\b;
        x4=-xx(2)/(2*xx(1));
        if norm(x4-x3) < 1e-6
        alfa=x4;
        break
        end
        x1=x2; x2=x3; x3=x4;
end
if i==100
alfa=[];
end
endfunction
```

Método de Newton



Búsqueda de línea con Newton

$$Min.\ f\left(\underline{x}^{(k)} + \alpha \underline{d}\right)$$

Función a minimizar

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

Solución actual

$\underline{x}^{(i)} \in \mathbb{R}^n$

Tamaño del salto

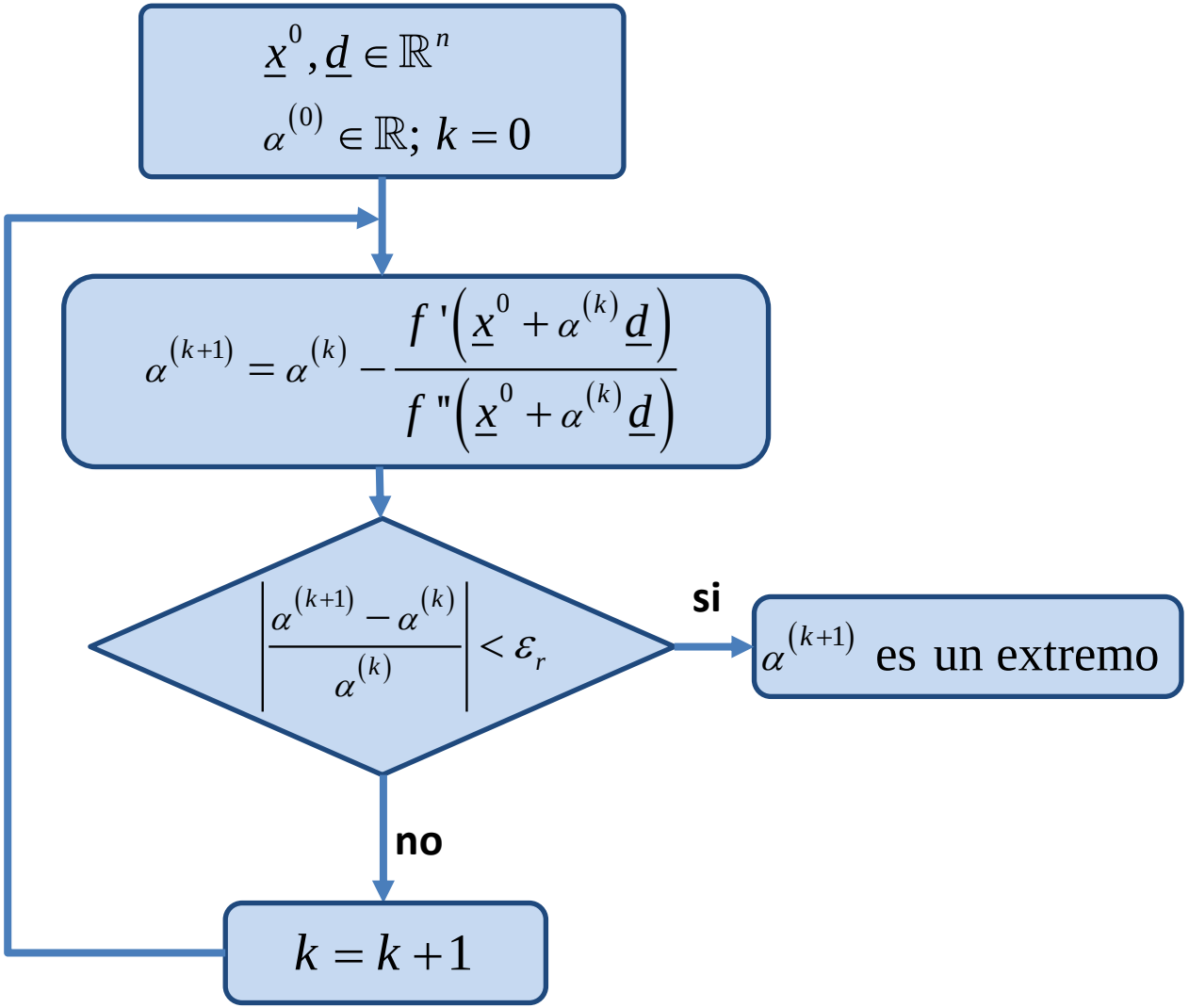
$\alpha \in \mathbb{R}$

Dirección de búsqueda

$\underline{d} \in \mathbb{R}^n$

Objetivo de la búsqueda de línea

Método de Newton



Búsqueda Lineal en Scilab con Newton

$$f'\left(\underline{x}^{(k)}+\alpha\underline{d}\right)=\frac{df\left(\underline{x}^{(k)}+\alpha\underline{d}\right)}{d\alpha}$$

$$\frac{df\left(\underline{x}^{(k)}+\alpha\underline{d}\right)}{d\alpha}=\frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial \alpha}+\frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial \alpha}+\ldots+\frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial \alpha}$$

$$\frac{df\left(\underline{x}^{(k)}+\alpha\underline{d}\right)}{d\alpha}=\frac{\partial f}{\partial x_1}d_1+\frac{\partial f}{\partial x_2}d_2+\ldots+\frac{\partial f}{\partial x_n}d_n$$

$$\frac{df\left(\underline{x}^{(k)}+\alpha\underline{d}\right)}{d\alpha}=\nabla f\left(\underline{x}^{(k)}+\alpha\underline{d}\right)^T\underline{d}$$

Búsqueda Lineal en Scilab con Newton

$$\frac{df\left(\underline{x}^{(k)}+\alpha\underline{d}\right)}{d\alpha} = \frac{\partial f}{\partial x_1}d_1 + \frac{\partial f}{\partial x_2}d_2 + \ldots + \frac{\partial f}{\partial x_n}d_n$$

$$\frac{d^2 f\left(\underline{x}^{(k)}+\alpha\underline{d}\right)}{d\alpha^2} = \left(\frac{\partial^2 f}{\partial x_1{}^2}\frac{\partial x_1}{\partial\alpha} + \frac{\partial^2 f}{\partial x_1\partial x_2}\frac{\partial x_2}{\partial\alpha} + \ldots + \frac{\partial^2 f}{\partial x_1\partial x_n}\frac{\partial x_n}{\partial\alpha}\right)d_1$$

$$+\left(\frac{\partial^2 f}{\partial x_2\partial x_1}\frac{\partial x_1}{\partial\alpha} + \frac{\partial^2 f}{\partial x_2{}^2}\frac{\partial x_2}{\partial\alpha} + \ldots + \frac{\partial^2 f}{\partial x_2\partial x_n}\frac{\partial x_n}{\partial\alpha}\right)d_2$$

$$+\cdots$$

$$+\left(\frac{\partial^2 f}{\partial x_n\partial x_1}\frac{\partial x_1}{\partial\alpha} + \frac{\partial^2 f}{\partial x_n\partial x_2}\frac{\partial x_2}{\partial\alpha} + \ldots + \frac{\partial^2 f}{\partial x_n{}^2}\frac{\partial x_n}{\partial\alpha}\right)d_n$$

Búsqueda Lineal en Scilab con Newton

$$\frac{d^2 f\left(\underline{x}^{(k)}+\alpha \underline{d}\right)}{d\alpha^2} = \left(\frac{\partial^2 f}{\partial x_1^{\,2}} d_1 + \frac{\partial^2 f}{\partial x_1 \partial x_2} d_2 + \ldots + \frac{\partial^2 f}{\partial x_1 \partial x_n} d_n \right) d_1$$

$$+ \left(\frac{\partial^2 f}{\partial x_2 \partial x_1} d_1 + \frac{\partial^2 f}{\partial x_2^{\,2}} d_2 + \ldots + \frac{\partial^2 f}{\partial x_2 \partial x_n} d_n \right) d_2$$

$$+ \cdots$$

$$+ \left(\frac{\partial^2 f}{\partial x_n \partial x_1} d_1 + \frac{\partial^2 f}{\partial x_n \partial x_2} d_2 + \ldots + \frac{\partial^2 f}{\partial x_n^{\,2}} d_n \right) d_n$$

Búsqueda Lineal en Scilab con Newton

$$\frac{d^2 f\left(\underline{x}^{(k)}+\alpha \underline{d}\right)}{d\alpha^2} = \left(\frac{\partial^2 f}{\partial x_1^{\,2}} d_1 + \frac{\partial^2 f}{\partial x_1 \partial x_2} d_2 + \ldots + \frac{\partial^2 f}{\partial x_1 \partial x_n} d_n\right) d_1 + \left(\frac{\partial^2 f}{\partial x_2 \partial x_1} d_1 + \frac{\partial^2 f}{\partial x_2^{\,2}} d_2 + \ldots + \frac{\partial^2 f}{\partial x_2 \partial x_n} d_n\right) d_2$$

$$+ \cdots + \left(\frac{\partial^2 f}{\partial x_n \partial x_1} d_1 + \frac{\partial^2 f}{\partial x_n \partial x_2} d_2 + \ldots + \frac{\partial^2 f}{\partial x_n^{\,2}} d_n\right) d_n$$

$$\begin{array}{cccc|c}
\frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_1 \partial x_1} & \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_1 \partial x_n} & d_1 \\
\frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_2 \partial x_1} & \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_2 \partial x_n} & d_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_n \partial x_1} & \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f\left(\underline{x}_0\right)}{\partial x_n \partial x_n} & d_n \\
\hline
d_1 & d_2 & \cdots & d_n &
\end{array}$$

$$\underline{d}^T \nabla^2 f\left(\underline{x}^{(k)}+\alpha \underline{d}\right) \underline{d}$$

Búsqueda Lineal en Scilab con Newton

```
function alfa=linesearchMN(Fun, Grad, Hess, xk, d, a_0)
    a=a_0;
    for i=1:100
        fp = Grad(xk + a*d)'*d;
        fpp= d'*Hess(xk + a*d)*d;
        a=a-fp/fpp;
        if norm(fp/fpp) < 1e-5
            alfa=a;
            break
        end
    end
    if i==100
        alfa=[];
    end
endfunction
```

Ejemplo

$$Min.\ \ f\left(\underline{x}\right)=\left(x_1{}^2+x_2-11\right)^2+\left(x_1+x_2{}^2-7\right)^2$$

$$\nabla f\left(\underline{x}\right)=\begin{bmatrix}2\left(x_1{}^2+x_2-11\right)2x_1+2\left(x_1+x_2{}^2-7\right)\\2\left(x_1{}^2+x_2-11\right)+2\left(x_1+x_2{}^2-7\right)2x_2\end{bmatrix}$$

$$\nabla^2 f\left(\underline{x}\right)=\begin{bmatrix}8x_1{}^2+4\left(x_1{}^2+x_2-11\right)+2 & 4x_1+4x_2\\4x_1+4x_2 & 2+8x_2{}^2+4\left(x_1+x_2{}^2-7\right)\end{bmatrix}$$

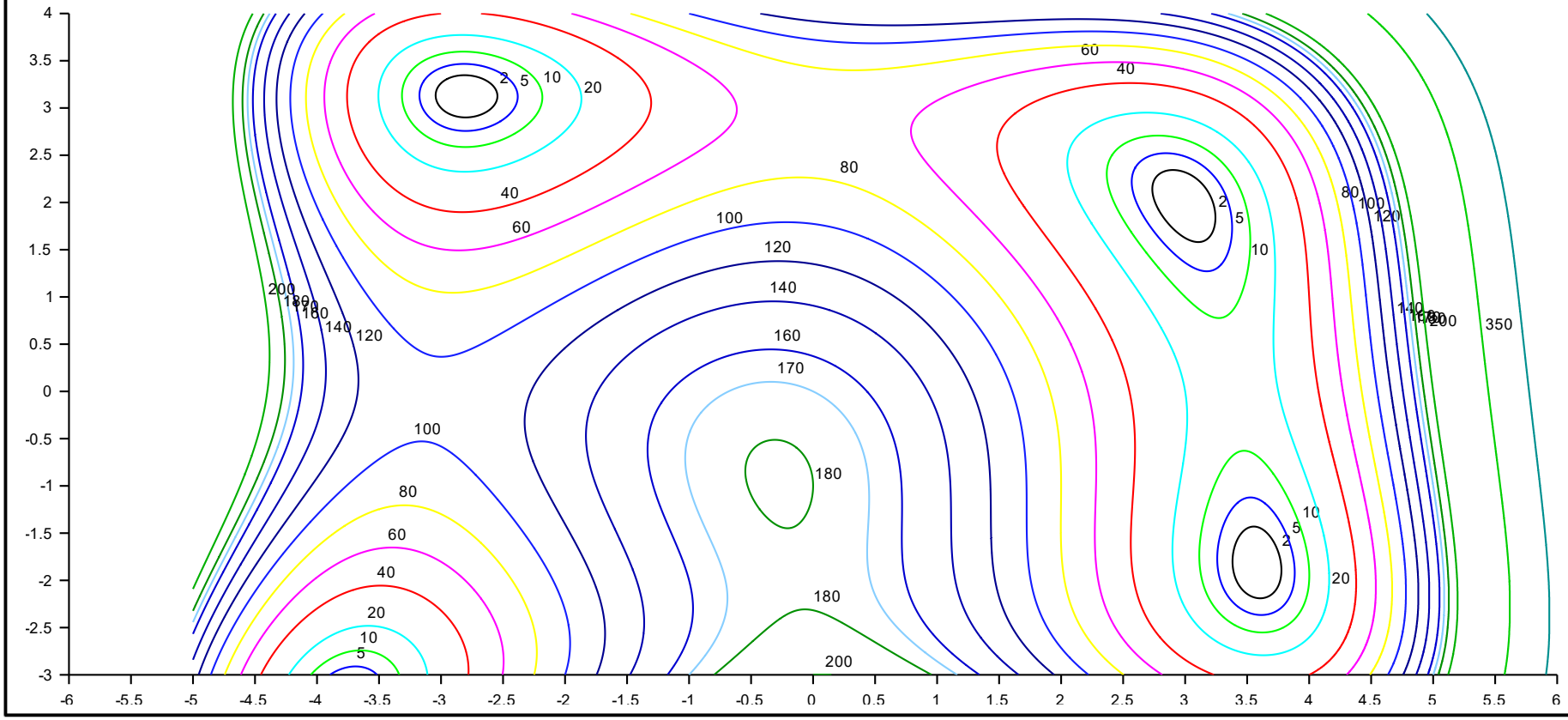
Ejemplo

```
function f=fun(x)
    f=(x(1)^2 + x(2) - 11)^2 + (x(1)+x(2)^2-7)^2
endfunction

function Gf=gradf(x)
    Gf(1,1) = 2*(x(1)^2 + x(2) - 11)*(2*x(1)) + 2*(x(1)+x(2)^2-7)
    Gf(2,1) = 2*(x(1)^2 + x(2) - 11) + 2*(x(1)+x(2)^2-7)*(2*x(2))
endfunction

function Hf=hessf(x)
    Hf(1,1) = 2*(2*x(1))*(2*x(1)) + 2*(x(1)^2 + x(2) - 11)*(2) + 2
    Hf(1,2) = 2*(2*x(1)) + 2*(2*x(2))
    Hf(2,1) = 2*(2*x(1)) + 2*(2*x(2))
    Hf(2,2) = 2 + 2*(2*x(2))*(2*x(2)) + 2*(x(1)+x(2)^2-7)*(2)
endfunction
```

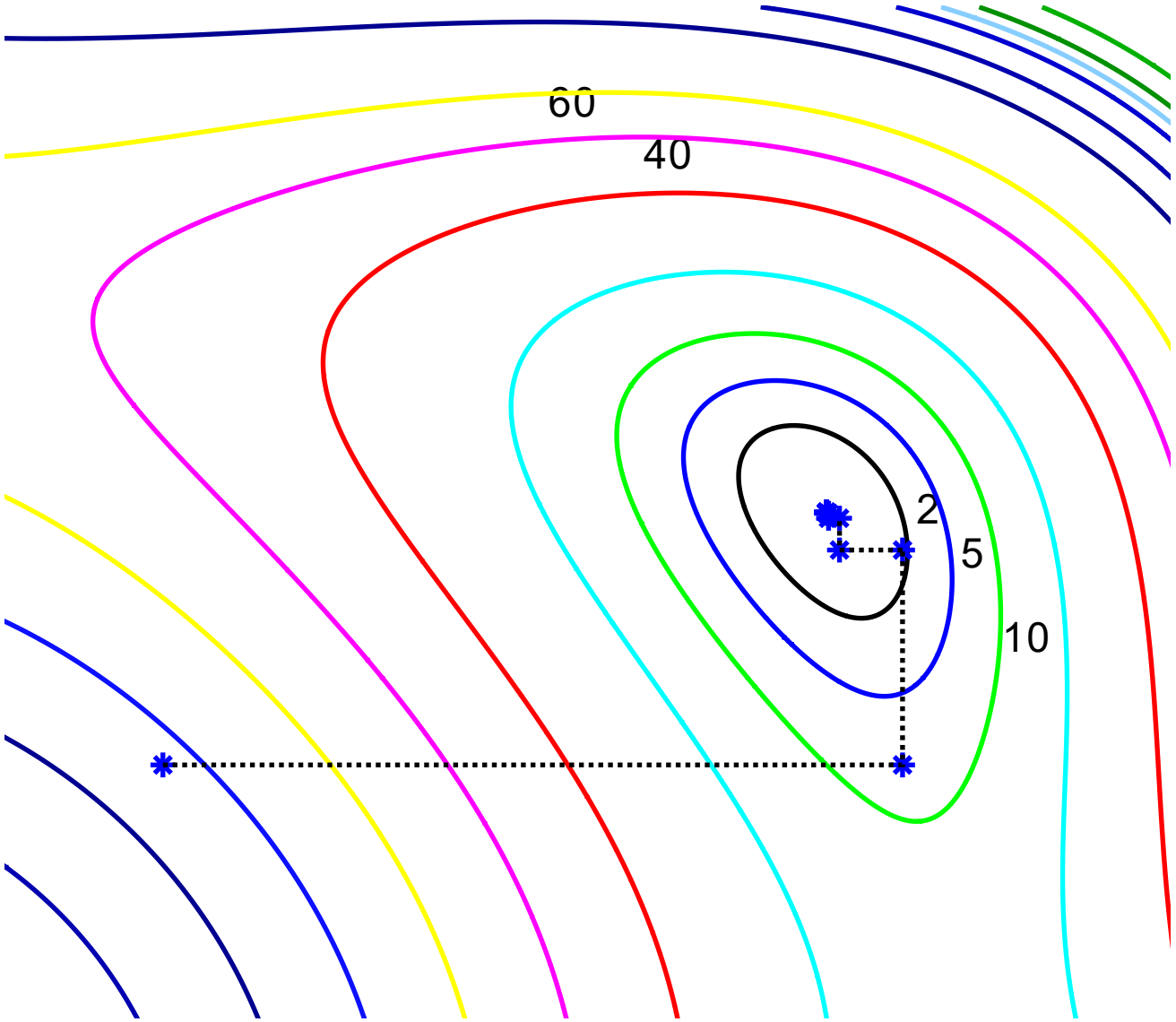
Ejemplo

```
x1=linspace(-5,6,1000);
x2=linspace(-3,4,1000);
for i=1:length(x1)
   for j=1:length(x2)
      z(i,j)=fun([x1(i);x2(j)]);
   end
end
```

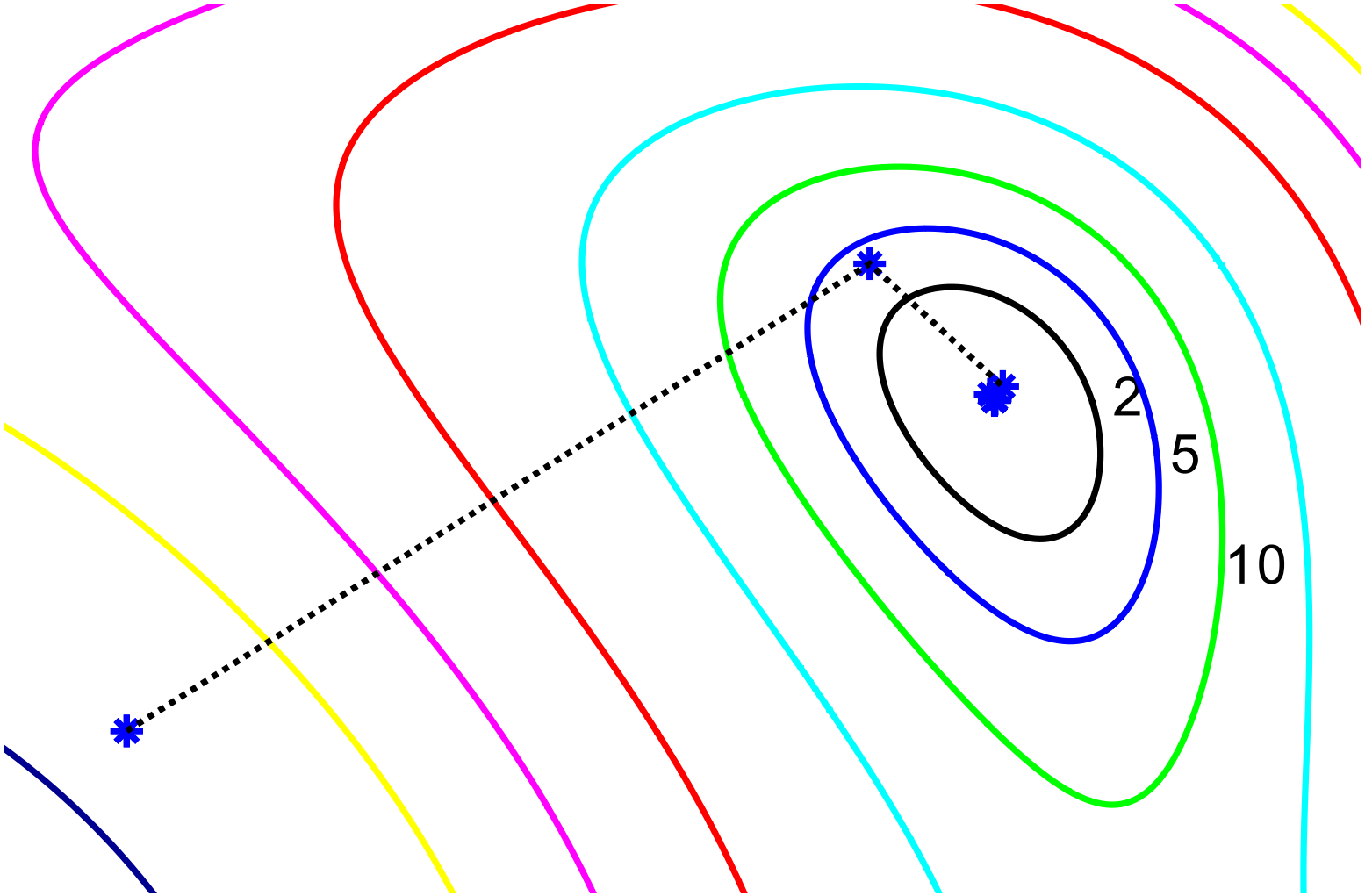
```
contour(x1,x2,z,[2 5 10 20:20:160 170 180 200 350 500])
```



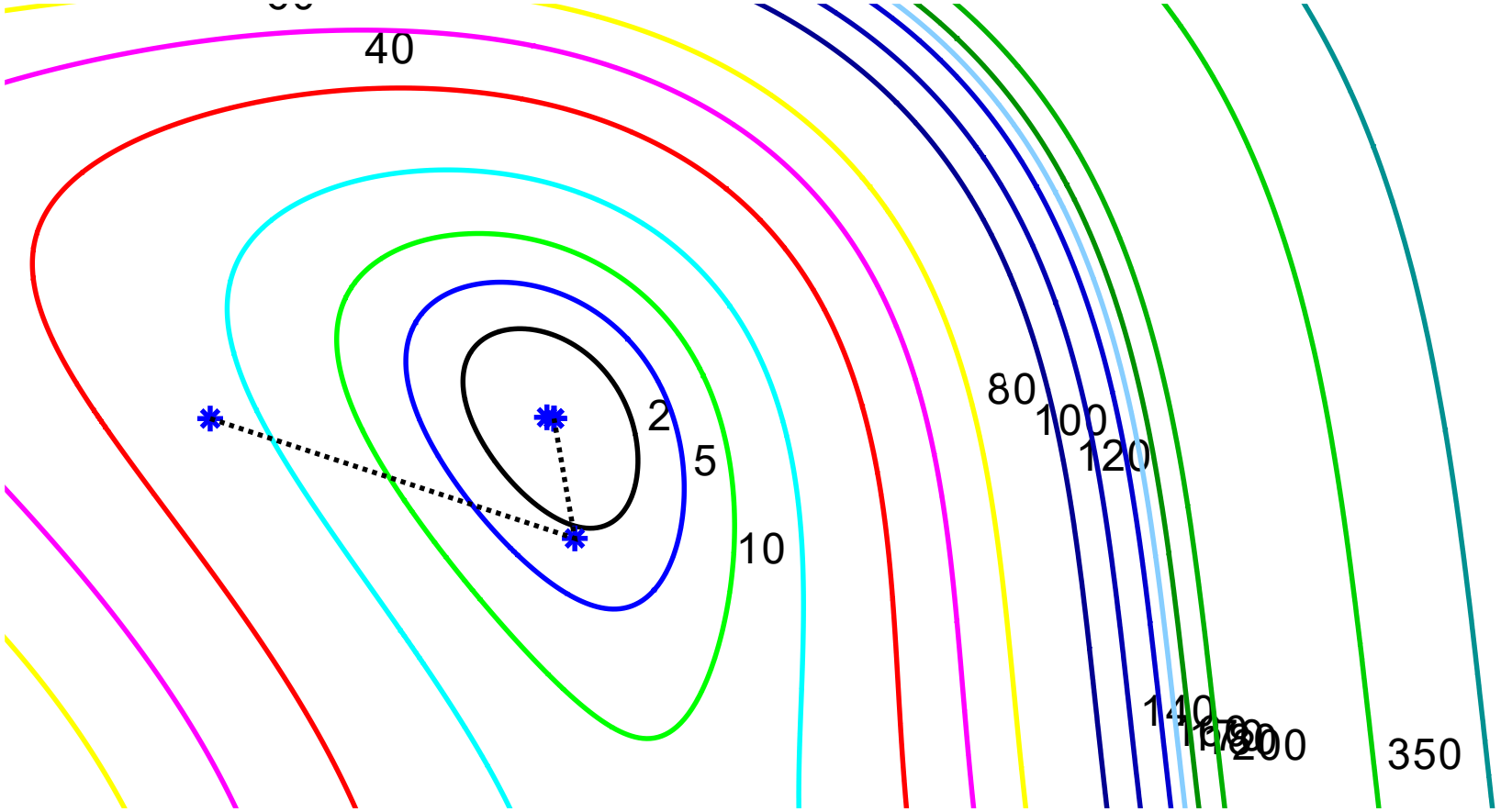
Ejemplo – 1 variable a la vez



Ejemplo – Descenso mas pronunciado



Ejemplo – Newton



Ejercicio

$$Min.\quad f\left(\underline{x}\right)=\left(1-x_1\right)^2+100\left(x_2-x_1{}^2\right)^2$$