

Análisis aplicado. Descenso suficiente.

José Luis Morales

<http://allman.rhon.itam.mx/~jmorales>

Departamento de Matemáticas. ITAM. 2009.

El problema por resolver.

El problema por resolver.

- Problemas *reales* que se pueden formular como:

$$\text{minimizar } f(x), \quad f : \mathbb{R}^n \rightarrow \mathbb{R},$$

en donde f es una función con *derivadas continuas de orden 2*.

El problema por resolver.

- Problemas *reales* que se pueden formular como:

$$\text{minimizar } f(x), \quad f : \mathbb{R}^n \rightarrow \mathbb{R},$$

en donde f es una función con *derivadas continuas de orden 2*.

- Optimización diferenciable.

Cuantificar descenso.

$$f(x_k + \alpha p_k) - f(x_k) = \boxed{\alpha p_k^T \nabla f(x_k)} + \mathcal{O}(\alpha^2), \quad \alpha > 0.$$

Supongamos que p_k es una dirección de descenso: $\nabla f_k p_k < 0$.
Entonces es posible encontrar $\alpha_k > 0$ tal que

$$f(x_{k+1}) = f(x_k + \alpha_k p_k) < f(x_k).$$

Notación: al escalar α_k se le llama el **paso**

Riesgos

Cuantificar descenso.

$$f(x_k + \alpha p_k) - f(x_k) = \boxed{\alpha p_k^T \nabla f(x_k)} + \mathcal{O}(\alpha^2), \quad \alpha > 0.$$

Supongamos que p_k es una dirección de descenso: $\nabla f_k^T p_k < 0$.
Entonces es posible encontrar $\alpha_k > 0$ tal que

$$f(x_{k+1}) = f(x_k + \alpha_k p_k) < f(x_k).$$

Notación: al escalar α_k se le llama el **paso**

Riesgos

- Pasos *cortos*:

$$\alpha_k \rightarrow 0 \implies f(x_k) - f(x_{k+1}) \rightarrow 0.$$

$$f(x_k + \alpha p_k) - f(x_k) = \boxed{\alpha p_k^T \nabla f(x_k)} + \mathcal{O}(\alpha^2), \quad \alpha > 0.$$

Supongamos que p_k es una dirección de descenso: $\nabla f_k^T p_k < 0$.
Entonces es posible encontrar $\alpha_k > 0$ tal que

$$f(x_{k+1}) = f(x_k + \alpha_k p_k) < f(x_k).$$

Notación: al escalar α_k se le llama el **paso**

Riesgos

- Pasos *cortos*:

$$\alpha_k \rightarrow 0 \implies f(x_k) - f(x_{k+1}) \rightarrow 0.$$

- Direcciones *poco productivas*. Ejemplo $p_k^T \nabla f(x_k) \rightarrow 0$

$$\implies f(x_k) - f(x_{k+1}) \rightarrow 0$$

Contraejemplo A. Descenso insuficiente.

(Pasos cortos.) Sea $f(x) = x^2$ con punto inicial $x_0 = 2$. La dirección de descenso siempre es $p_k = -1$ con pasos $\alpha_k = 2^{-k+1}$

$$\{x_k\} = \left\{2, \frac{3}{2}, \frac{5}{4}, \frac{9}{8}, \dots\right\} = \{1 + 2^{-k}\}.$$

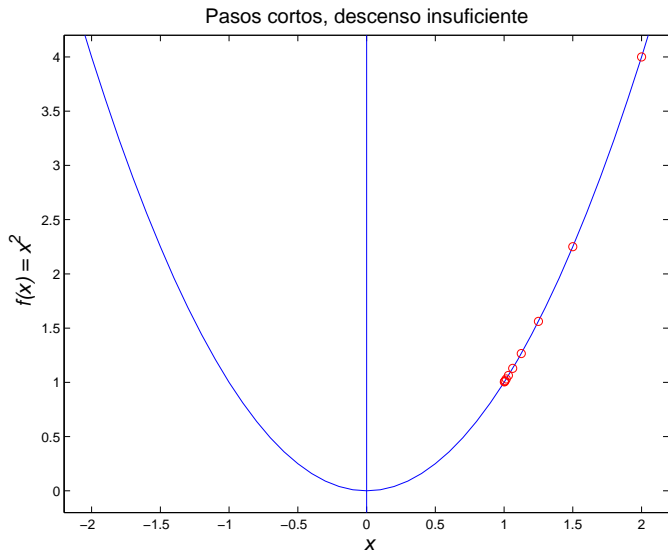
$f(x_k)$ decrece monótonamente y converge a 1, que no es un mínimo de f .

Contraejemplo B. Descenso insuficiente.

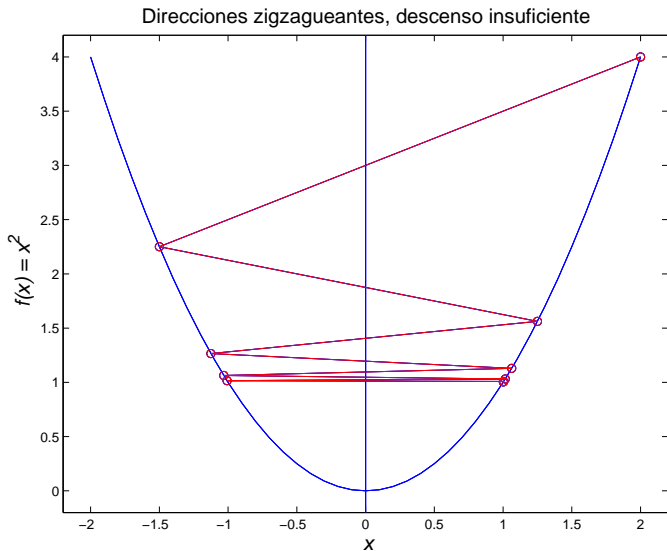
(Direcciones zigzagueantes.) Sea $f(x) = x^2$ con punto inicial $x_0 = 2$. Las direcciones de descenso son $p_k = (-1)^{k+1}$ con pasos $\alpha_k = 2 + 3(2^{-k-1})$

$$\{x_k\} = \left\{2, -\frac{3}{2}, \frac{5}{4}, -\frac{9}{8}, \dots\right\} = \{(-1)^k(1 + 2^{-k})\}$$

$f(x_k)$ decrece monótonamente y converge a 1, que no es un mínimo de f . La sucesión $\{x_k\}$ diverge y tiene dos puntos de acumulación.



Direcciones improductivas.



Métodología para generar direcciones de descenso.

Modelos locales:

$$\text{minimizar } m(p) = f(x_k) + p^T \nabla f(x_k) + \frac{1}{2} p^T B_k p,$$

Métodología para generar direcciones de descenso.

Modelos locales:

$$\text{minimizar } m(p) = f(x_k) + p^T \nabla f(x_k) + \frac{1}{2} p^T B_k p,$$

- Método del gradiente

$$B_k = I, \quad p_k = -\nabla f(x_k)$$

Métodología para generar direcciones de descenso.

Modelos locales:

$$\text{minimizar } m(p) = f(x_k) + p^T \nabla f(x_k) + \frac{1}{2} p^T B_k p,$$

- Método del gradiente

$$B_k = I, \quad p_k = -\nabla f(x_k)$$

- Método de Newton ($\nabla^2 f(x_k)$ spd)

$$B_k = \nabla^2 f(x_k), \quad p_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$$

Métodología para generar direcciones de descenso.

Modelos locales:

$$\text{minimizar } m(p) = f(x_k) + p^T \nabla f(x_k) + \frac{1}{2} p^T B_k p,$$

- Método del gradiente

$$B_k = I, \quad p_k = -\nabla f(x_k)$$

- Método de Newton ($\nabla^2 f(x_k)$ spd)

$$B_k = \nabla^2 f(x_k), \quad p_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$$

- Métodos cuasi-Newton (B_k spd)

$$B_k \approx \nabla^2 f(x_k), \quad p_k = -B_k^{-1} \nabla f(x_k)$$

Búsqueda lineal:

Si p_k es una dirección de descenso, entonces

$$\alpha_k = \arg \min f(x_k + \alpha p_k), \quad \alpha \in (0, 1].$$

En la práctica, **búsqueda lineal inexacta**, α_k es una aproximación al primer minimizador local de

$$f(x_k + \alpha p_k).$$

Método de máximo descenso

EJEMPLO: Método de máximo descenso para cuadráticas estrictamente convexas:

$$f(x) = a + b^T x + \frac{1}{2} x^T A x, \quad A \text{ spd.}$$

Escoger x_0 una aproximación inicial.

Desde $k = 0$ hasta **convergencia**

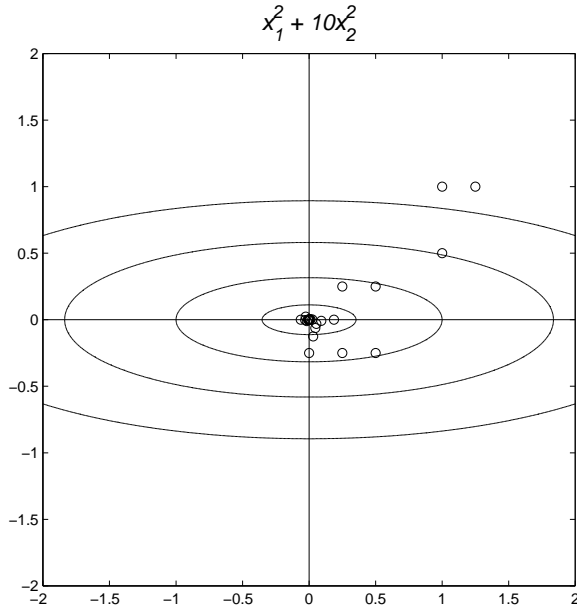
$$p_k = -\nabla f_k$$

$$\alpha_k = \arg \min_{\alpha} f(x_k + \alpha p_k)$$

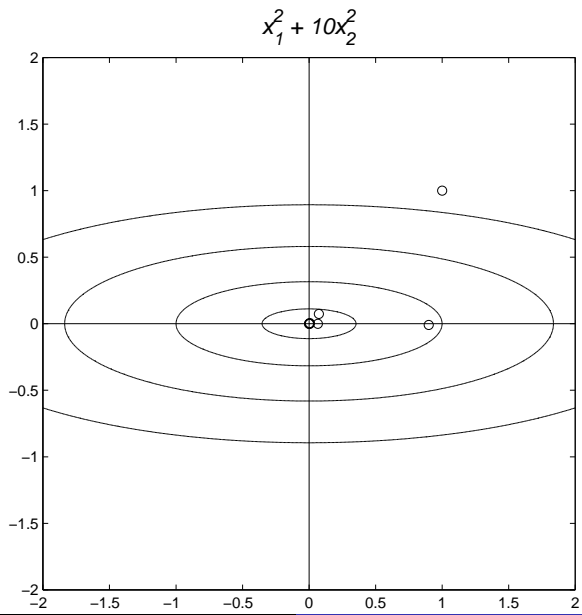
$$x_{k+1} = x_k + \alpha_k p_k$$

$$\alpha_k = \frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T A \nabla f_k}$$

Método de Nelder-Mead. $x_0 = (1, 1)^T$



Método de gradiente. $x_0 = (1, 1)^T$



Método de gradiente. $x_0 = (1.8, 0.1)^T$

