OTIMIZAÇÃO

Método do Gradiente (Máxima Descida)

Equação de Busca

Seja $f : \mathbb{R}^n \to \mathbb{R}$, uma função a minimizar; i. e., deseja-se encontrar $x^* \in \mathbb{R}^n$, tal que $f(x^*)$ seja mínima.

Supondo que x^p é um vetor próximo de x^*, obtém-se, através do desenvolvimento de $f(x)$ em série de Taylor:

$$f(x^p + \delta x) = f(x^p) + [f'(x^p)]^T \delta x + \frac{1}{2} H^T \delta x^2 H + \cdots$$

Assim:

$$f(x^p + \delta x) - f(x^p) = \nabla f(x^p) \cdot \delta x$$

ou:

$$[-\nabla f(x^p)] \cdot [dx] = f(x^p) - f(x^p + \delta x)$$

Para que a redução em f seja a máxima possível, é necessário que o produto escalar do 1º membro seja máximo, ou seja:

$$\delta x = -c \nabla f(x^p) \text{ (colineares)}$$

Considerando $\delta x = x^{(p+1)} - x^{(p)}$, fica:

$$x^{(t+1)} = x^{(t)} - c \nabla f(x^{(t)})$$

onde:

c → Passo do método

Ex.: O presente exemplo tem por objetivo mostrar a influência do valor do passo c na convergência do método. Para tanto, será analisado o caso simples de uma função de duas variáveis, sem restrições, definida por:

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

Calculando as derivadas parciais para formação do vetor gradiente, tem-se:

$$\frac{\partial f}{\partial x} = 2x - 2 \qquad\qquad \frac{\partial f}{\partial y} = 2y$$

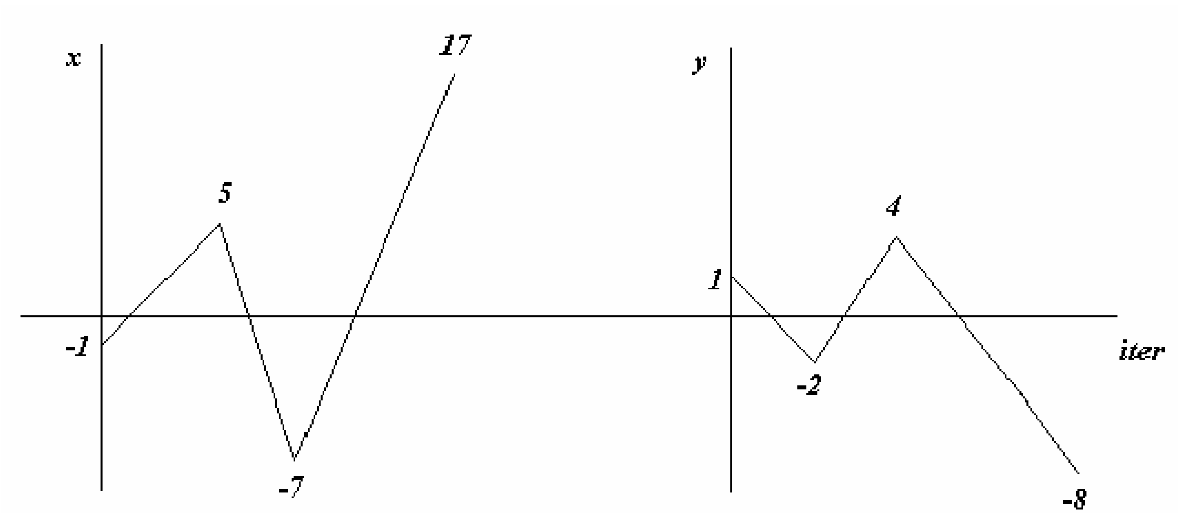
Aplicando a equação determinada anteriormente:

$$\begin{bmatrix} x \\ y \end{bmatrix}^{(t+1)} = \begin{bmatrix} x \\ y \end{bmatrix}^{(t)} - c \cdot \begin{bmatrix} 2x-2 \\ 2y \end{bmatrix}^{(t)}$$

A fim de comprovar a dependência da convergência com o passo, adotou-se como iniciação o ponto $(x_0; y_0) = (-1;1)$ e testam-se 3 valores distintos para c: c = 1,50; c = 0,75 e c = 0,50.

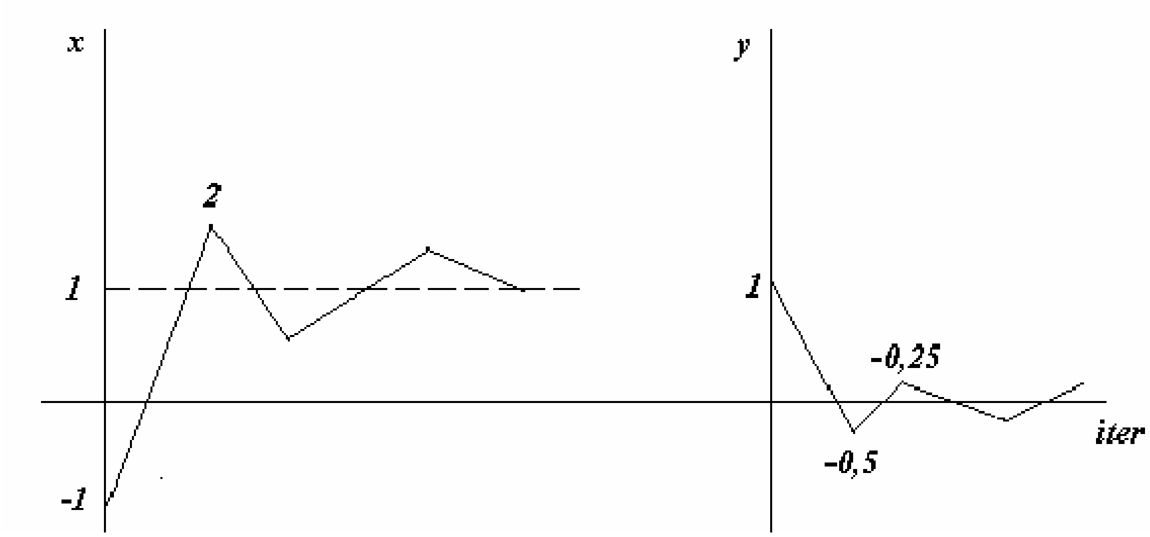
Para c = 1,50:

iter	0	1	2	3	...	...
x	-1	5	-7	17	...	...
y	1	-2	4	-8	...	...



Para c = 0,75:

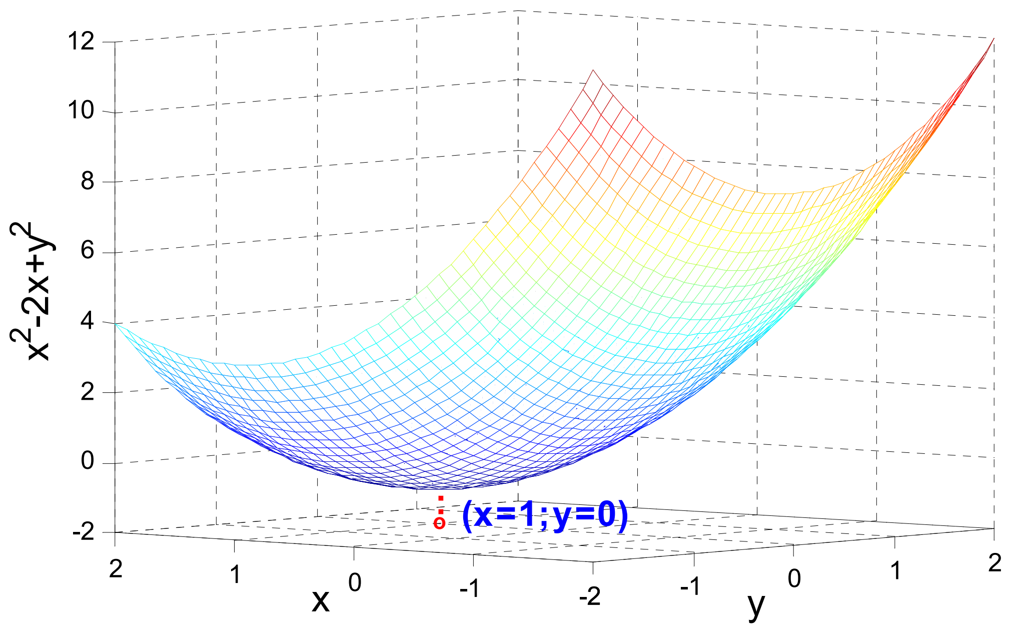
$$\left[\begin{array}{c|c|c|c|c|c|c} iter & 0 & 1 & 2 & 3 & 4 & \cdots \\ \hline x & -1 & 2 & 0.50 & 1.25 & 0.875 & \cdots \\ \hline y & 1 & -0.5 & 0.25 & -0.125 & 0.0625 & \cdots \end{array}\right]$$



Para c = 0,50:

$$\left[\begin{array}{c|c|c|c} iter & 0 & 1 & \cdots \\ \hline x & -1 & 1 & \cdots \\ \hline y & 1 & 0 & \cdots \end{array}\right]$$

Notar que, para $(x; y) = (1; 0)$ o vetor gradiente se anula. Portanto, esse é o ponto de ótimo da função, que, para c = 0,50, é alcançado em uma única iteração.



Método do Gradiente Ótimo

O fator c determina o tamanho do passo na direção de máxima descida.

Aqui se estabelece um sub-problema de otimização: determinar o melhor valor para c.

Isso pode ser resolvido através de uma busca unidirecional, tomando como variável o parâmetro c e aproximando a função $f(x - c\nabla f(x))$ por uma parábola, definida pelos seguintes parâmetros:

- Valor da função objetivo na iteração mais recente:

$$f_0 = f(c = 0)$$

- Inclinação da função para $c = 0$:

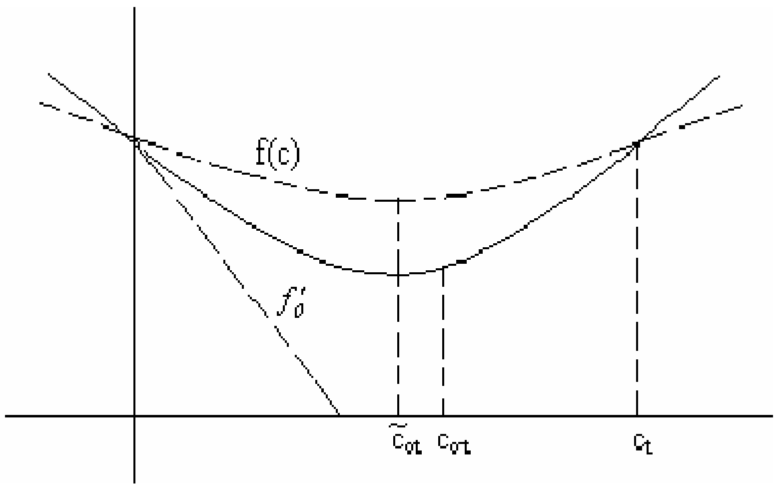
$$\frac{df(x - c\nabla f)}{dc} = \frac{df(x)}{dc} \cdot (-\nabla f) = \nabla^T f \ \cdot (-\nabla f)$$

$$f_0' = \frac{df}{dc} = -\nabla^T f \ \cdot \nabla f$$

- Valor da função para um passo-teste c_T:

$$f_T = f(c = c_T)$$

A parábola assim definida é mostrada na figura a seguir:



A equação da parábola fica:

$$p = f_0 + f'_c + \frac{\left[f_T - \left(f_0 + f' c_T\right)\right]}{c_T^2} \cdot c^2$$

e o valor ótimo de c fica:

$$\tilde{c}_{ot} = \frac{1}{2}\left[\left(c_T^2 f_0'\right) / \left(f_0 + c_T f_0' - f_T\right)\right]$$

Supondo ainda uma função objetivo quadrática:

$$f(x) = \frac{1}{2} x^T H\, x + b^T x + a$$

então, $\nabla f(x)$ é dado por:

$$\nabla f(x) = H\, x + b$$

A busca linear exata para $x^{(a)} = x + c\, d$ é:

$$f(x + c\, d) = \frac{1}{2}(x + c\, d)^T H(x + c\, d) + b^T (x + c\, d) + a$$

$$= \frac{1}{2} x^T H\, x + \left(c\, d^T H + b^T\right) \cdot x + \frac{1}{2} c^2 d^T H\, d + c\, b^T d + a$$

Então:

$$\frac{df(x + c\, d)}{dc} = d^T H\, x + c\, d^T H\, d + b^T d = 0$$

$$= d^T H\, x + c\, d^T H\, d + d^T b = 0$$

$$c\, d^T H\, d = -d^T (H\, x + b)$$

$$c = -\frac{d^T \nabla f}{d^T H\, d}$$

Como $d = -\nabla f$

$$c = \frac{\nabla^T f \cdot \nabla f}{\nabla^T f \cdot H \cdot \nabla f}$$

Ex.: Realizar uma busca linear exata para a função:

$$f(x,y) = x^2 - 2x - y + y^2 \qquad \nabla f = \begin{bmatrix} 2x-2 \\ 2y-1 \end{bmatrix}$$

$$f\left(x_0 - c(2x_0-2), y_0 - c(2y_0-1)\right) = \left[x_0 - c(2x_0-2)\right]^2 -$$
$$-2\left[x_0 - c(2x_0-2)\right] - \left[y_0 - c(2y_0-1)\right] + \left[y_0 - c(2y_0-1)\right]^2$$

$$= x_0^2 - 2cx_0(2x_0-2) + c^2(2x_0-2)^2 - 2\left[x_0 - c(2x_0-2)\right] +$$
$$+ y_0^2 - 2c\,y_0(2y_0-1) + c^2(2y_0-1)^2 - \left[y_0 - c(2y_0-1)\right]$$

$$= c^2\left[(2x_0-2)^2 + (2y_0-1)^2\right] - c\left[(2x_0-2)(2x_0) - 2(2x_0-2) +\right.$$
$$\left. + (2y_0)(2y_0-1) - (2y_0-1)\right] + x_0^2 + y_0^2 - 2x_0 - y_0$$

$$= c^2\left[\underbrace{(2x_0-2)^2 + (2y_0-1)^2}_{z}\right] - c\left[\underbrace{(2x_0-2)^2 + (2y_0-1)^2}_{z}\right] +$$
$$+ x_0^2 + y_0^2 - 2x_0 - y_0$$

$$\frac{\partial f}{\partial c} = 0 \Rightarrow 2c\,z - z = 0 \Rightarrow \text{ para } z \neq 0 \text{ tem-se: } c_{ot} = 1/2 = 0.5$$

para $x_0 = 5$, $y_0 = -5$ $\nabla f = [8,-11]^T \Rightarrow \begin{cases} x_{ot} = 5 - 0.5 \times 8 = 1 \\ y_{ot} = -5 - 0.5 \times (-11) = 0.5 \end{cases}$

para $x_0 = -3$, $y_0 = 2$ $\nabla f = [-8,3]^T \Rightarrow \begin{cases} x_{ot} = -3 - 0.5 \times (-8) = 1 \\ y_{ot} = 2 - 0.5 \times (3) = 0.5 \end{cases}$

Observa-se que, como a função objetivo é quadrática, adoção do passo ótimo para qualquer ponto de partida leva à obtenção do ótimo em uma única iteração.