Exercício sobre o método de quasi-Newton

Isabel Espírito Santo

Departamento de Produção e Sistemas

Escola de Engenharia

Universidade do Minho

iapinho@dps.uminho.pt

O lucro, em milhares de euros, da colocação de um sistema elétrico é dado por

$$\mathcal{L}(x_1, x_2) = 20x_1 + 26x_2 + 4x_1x_2 - 4x_1^2 - 3x_2^2$$

em que x_1 e x_2 designam, respetivamente, o custo da mão de obra e do material. Calcule o lucro máximo usando o método quasi-Newton baseado na fórmula DFP, considerando na paragem do processo iterativo $\varepsilon = 0.0001$. Tome a seguinte aproximação inicial $(0, 0)$. No critério de Armijo use $\mu = 0.001$.

Resolução:

$$\max \mathcal{L}(x_1, x_2) = 20x_1 + 26x_2 + 4x_1x_2 - 4x_1^2 - 3x_2^2$$

$$\min -\mathcal{L}(x_1, x_2) = f(x_1, x_2) = -20x_1 - 26x_2 - 4x_1x_2 + 4x_1^2 + 3x_2^2$$

$$\nabla f(x) = \begin{pmatrix} -20 - 4x_2 + 8x_1 \\ -26 - 4x_1 + 6x_2 \end{pmatrix}$$

Iniciar o algoritmo de quasi-Newton:
$x^1 = (0, 0), \mu = 0.001, \varepsilon = 0.001$

- **1ª iteração**

$$x^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \nabla f(x^1) = \begin{pmatrix} -20 \\ -26 \end{pmatrix}$$

$$H^1 = I$$

Cálculo da direção d^1_{QN}

$$d^1_{QN} = -H^1 \nabla f(x^1) = \begin{pmatrix} 20 \\ 26 \end{pmatrix}$$

$\nabla f(x^1)^T d^1_{QN} = \begin{pmatrix} -20 & -26 \end{pmatrix} \begin{pmatrix} 20 \\ 26 \end{pmatrix} = -1076 < 0$, logo d^1_{QN} é descendente.

Cálculo de α

$\alpha = 1$

$x^{\text{aux}} = x^1 + \alpha d_{QN}^1 = \begin{pmatrix} 20 \\ 26 \end{pmatrix}$

$\begin{cases} f(x^1) = 0 \\ f(x^{\text{aux}}) = 472 \quad \uparrow \end{cases}$

$\alpha = 0.5$

$x^{\text{aux}} = x^1 + \alpha d_{QN}^1 = \begin{pmatrix} 10 \\ 13 \end{pmatrix}$

$\begin{cases} f(x^1) = 0 \\ f(x^{\text{aux}}) = -151 \quad \downarrow \end{cases}$

Critério de Armijo

$f(x^{\text{aux}}) \leq f(x^1) + \mu\alpha\nabla f(x^1)^T d_{QN}^1 \Leftrightarrow -151 \leq$
$0 + 0.001 \times 0.5 \times (-1076)$
$\Leftrightarrow -151 \leq -0.538$ (verdadeiro) logo a descida é significativa.

$$x^2 = \begin{pmatrix} 10 \\ 13 \end{pmatrix}$$

- *Critério de Paragem*

$$\|\nabla f(x^2)\|_2 = \left\| \begin{pmatrix} 8 \\ 12 \end{pmatrix} \right\|_2 = 14.4222 \leq \varepsilon \quad \text{(falso)}$$

- **2ª iteração**

$$x^2 = \begin{pmatrix} 10 \\ 13 \end{pmatrix} \qquad \nabla f(x^2) = \begin{pmatrix} 8 \\ 12 \end{pmatrix}$$

$$H^2 = H^1 - \frac{H^1 y^1 {y^1}^T H^1}{{y^1}^T H^1 y^1} + \frac{s^1 {s^1}^T}{{s^1}^T y^1}$$

$$s^1 = x^2 - x^1 = \begin{pmatrix} 10 \\ 13 \end{pmatrix}$$

$$y^1 = \nabla f(x^2) - \nabla f(x^1) = \begin{pmatrix} 28 \\ 38 \end{pmatrix}$$

$$y^1 {y^1}^T = \begin{pmatrix} 784 & 1064 \\ 1064 & 1444 \end{pmatrix}$$

$${y^1}^T y^1 = 2228$$

$$s^1 {s^1}^T = \begin{pmatrix} 100 & 130 \\ 130 & 169 \end{pmatrix}$$

$${s^1}^T y^1 = 774$$

$$H^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{\begin{pmatrix} 784 & 1064 \\ 1064 & 1444 \end{pmatrix}}{2228} + \frac{\begin{pmatrix} 100 & 130 \\ 130 & 169 \end{pmatrix}}{774} =$$
$$\begin{pmatrix} 0.7773 & -0.3096 \\ -0.3096 & 0.5702 \end{pmatrix}$$

Cálculo da direção d_{QN}^2

$$d_{QN}^2 = -H^2 \nabla f(x^2) = \begin{pmatrix} -2.5032 \\ -4.3656 \end{pmatrix}$$

$\nabla f(x^2)^T d_{QN}^2 = -72.4128 < 0$, logo d_{QN}^2 é descendente.

Cálculo de α

$\alpha = 1$

$x^{\text{aux}} = x^2 + \alpha d_{QN}^2 = \begin{pmatrix} 7.4968 \\ 8.6344 \end{pmatrix}$

$$\begin{cases} f(x^2) = -151 \\ f(x^{\text{aux}}) = -184.8852 \quad \downarrow \end{cases}$$

Critério de Armijo

$f(x^{\text{aux}}) \leq f(x^2) + \mu\alpha\nabla f(x^2)^T d_{QN}^2 \Leftrightarrow -184.8852 \leq$
$-151 + 0.001 \times 1 \times (-72.4128)$
$\Leftrightarrow -184.8852 \leq -151.0724$ (verdadeiro) logo a descida é significativa.

$x^3 = \begin{pmatrix} 7.4968 \\ 8.6344 \end{pmatrix}$

- *Critério de Paragem*

$$\|\nabla f(x^3)\|_2 = \left\| \begin{pmatrix} 5.4368 \\ -4.1808 \end{pmatrix} \right\|_2 = 6.8584 \leq \varepsilon \quad \text{(falso)}$$

- **3ª iteração**

$$x^3 = \begin{pmatrix} 7.4968 \\ 8.6344 \end{pmatrix} \qquad \nabla f(x^3) = \begin{pmatrix} 5.4368 \\ -4.1808 \end{pmatrix}$$

$$H^3 = H^2 - \frac{H^2 y^2 y^{2^T} H^2}{y^{2^T} H^2 y^2} + \frac{s^2 s^{2^T}}{s^{2^T} y^2}$$

$$s^2 = x^3 - x^2 = \begin{pmatrix} -2.5032 \\ -4.3656 \end{pmatrix}$$

$$y^2 = \nabla f(x^3) - \nabla f(x^2) = \begin{pmatrix} -2.5632 \\ -16.1808 \end{pmatrix}$$

$$H^2 y^2 = \begin{pmatrix} 3.0172 \\ -8.4327 \end{pmatrix}$$

$$y^{2^T} H^2 = \begin{pmatrix} 3.0172 & -8.4327 \end{pmatrix}$$

$$H^2 y^2 y^{2^T} H^2 = \begin{pmatrix} 9.1035 & -25.4431 \\ -25.4431 & 71.1104 \end{pmatrix}$$

$$y^{2^T} H^2 y^2 = 128.7141$$

$$s^2 s^{2^T} = \begin{pmatrix} 6.2660 & 10.9280 \\ 10.9280 & 19.0585 \end{pmatrix}$$

$$s^{2^T} y^2 = 77.0551$$

$$H^3 = \begin{pmatrix} 0.7773 & -0.3096 \\ -0.3096 & 0.5702 \end{pmatrix} - \frac{\begin{pmatrix} 9.1035 & -25.4431 \\ -25.4431 & 71.1104 \end{pmatrix}}{128.7141} + \frac{\begin{pmatrix} 6.2660 & 10.9280 \\ 10.9280 & 19.0585 \end{pmatrix}}{77.0551} = \begin{pmatrix} 0.7879 & 0.0299 \\ 0.0299 & 0.2651 \end{pmatrix}$$

Cálculo da direção d_{QN}^3

$$d_{QN}^3 = -H^3 \nabla f(x^3) = \begin{pmatrix} -4.1586 \\ 0.9458 \end{pmatrix}$$

$\nabla f(x^3)^T d_{QN}^3 = -26.5637 < 0$, logo d_{QN}^3 é descendente.

Cálculo de α

$\alpha = 1$

$$x^{\text{aux}} = x^3 + \alpha d_{QN}^3 = \begin{pmatrix} 3.3382 \\ 9.5802 \end{pmatrix}$$

$$\begin{cases} f(x^3) = -184.8852 \\ f(x^{\text{aux}}) = -123.8567 \quad \uparrow \end{cases}$$

$\alpha = 0.5$

$$x^{\text{aux}} = x^3 + \alpha d_{QN}^3 = \begin{pmatrix} 5.4175 \\ 9.1073 \end{pmatrix}$$

$$\begin{cases} f(x^3) = -184.8852 \\ f(x^{\text{aux}}) = -176.2690 \quad \uparrow \end{cases}$$

$\alpha = 0.25$

$$x^{\text{aux}} = x^3 + \alpha d_{QN}^3 = \begin{pmatrix} 6.4572 \\ 8.8708 \end{pmatrix}$$

$$\begin{cases} f(x^3) = -184.8852 \\ f(x^{\text{aux}}) = -186.0518 \quad \downarrow \end{cases}$$

Critério de Armijo

$f(x^{\text{aux}}) \leq f(x^3) + \mu\alpha\nabla f(x^3)^T d_{QN}^3 \Leftrightarrow -186.0517 \leq -184.8852 + 0.001 \times 0.25 \times (-26.5637)$

$\Leftrightarrow -186.0517 \leq -184.8918$ (verdadeiro) logo a descida é significativa.

$$x^4 = \begin{pmatrix} 6.4572 \\ 8.8709 \end{pmatrix}$$

- *Critério de Paragem*

$$\|\nabla f(x^3)\|_2 = \left\| \begin{pmatrix} -3.8262 \\ 1.3965 \end{pmatrix} \right\|_2 = 4.0731 \leq \varepsilon \quad \text{(falso)}$$

- **4ª iteração**

$$x^4 = \begin{pmatrix} 6.4572 \\ 8.8709 \end{pmatrix} \qquad \nabla f(x^4) = \begin{pmatrix} -3.8262 \\ 1.3965 \end{pmatrix}$$

$$H^4 = H^3 - \frac{H^3 y^3 {y^3}^T H^3}{{y^3}^T H^3 y^3} + \frac{s^3 {s^3}^T}{{s^3}^T y^3}$$

$$s^3 = x^4 - x^3 = \begin{pmatrix} -1.0397 \\ 0.2364 \end{pmatrix}$$

$$y^3 = \nabla f(x^4) - \nabla f(x^3) = \begin{pmatrix} -9.2630 \\ 5.5773 \end{pmatrix}$$

$$H^3 y^3 = \begin{pmatrix} -7.1316 \\ 1.2016 \end{pmatrix}$$

$$y^{3^T} H^3 = \begin{pmatrix} -7.1316 & 1.2016 \end{pmatrix}$$

$$H^3 y^3 y^{3^T} H^3 = \begin{pmatrix} 50.8597 & -8.5693 \\ -8.5693 & 1.4438 \end{pmatrix}$$

$$y^{3^T} H^3 y^3 = 72.7617$$

$$s^3 s^{3^T} = \begin{pmatrix} 1.0809 & -0.2458 \\ -0.2458 & 0.0559 \end{pmatrix}$$

$$s^{3^T} y^3 = 10.9490$$

$$H^4 = \begin{pmatrix} 0.7879 & 0.0299 \\ 0.0299 & 0.2651 \end{pmatrix} - \frac{\begin{pmatrix} 50.8597 & -8.5693 \\ -8.5693 & 1.4438 \end{pmatrix}}{72.7617} + \frac{\begin{pmatrix} 1.0809 & -0.2458 \\ -0.2458 & 0.0559 \end{pmatrix}}{10.9490} = \begin{pmatrix} 0.1876 & 0.1252 \\ 0.1252 & 0.2504 \end{pmatrix}$$

Cálculo da direção d_{QN}^4

$$d_{QN}^4 = -H^4 \nabla f(x^4) = \begin{pmatrix} 0.5430 \\ 0.1294 \end{pmatrix}$$

$\nabla f(x^4)^T d_{QN}^4 = -1.8968 < 0$, logo d_{QN}^4 é descendente.

Cálculo de α

$\alpha = 1$

$x^{\text{aux}} = x^4 + \alpha d_{QN}^4 = \begin{pmatrix} 7.0001 \\ 9.0002 \end{pmatrix}$

$$\begin{cases} f(x^4) = -186.0517 \\ f(x^{\text{aux}}) = -187.0000 \quad \downarrow \end{cases}$$

Critério de Armijo

$f(x^{\text{aux}}) \leq f(x^4) + \mu\alpha\nabla f(x^4)^T d_{QN}^4 \Leftrightarrow -187.0000 \leq -186.0517 + 0.001 \times 1 \times (-1.8968)$

$\Leftrightarrow -187.0000 \leq -186.0536$ (verdadeiro) logo a descida é significativa.

$x^5 = \begin{pmatrix} 7.0001 \\ 9.0002 \end{pmatrix}$

- *Critério de Paragem*

$$\|\nabla f(x^4)\|_2 = \left\| \begin{pmatrix} 0.0000 \\ 0.0008 \end{pmatrix} \right\|_2 = 0.0008 \leq \varepsilon \quad \text{(verdadeiro)}$$

$x_{\max} \approx \begin{pmatrix} 7.0001 \\ 9.0002 \end{pmatrix}$ e $\mathcal{L}_{\max} \approx 187$.