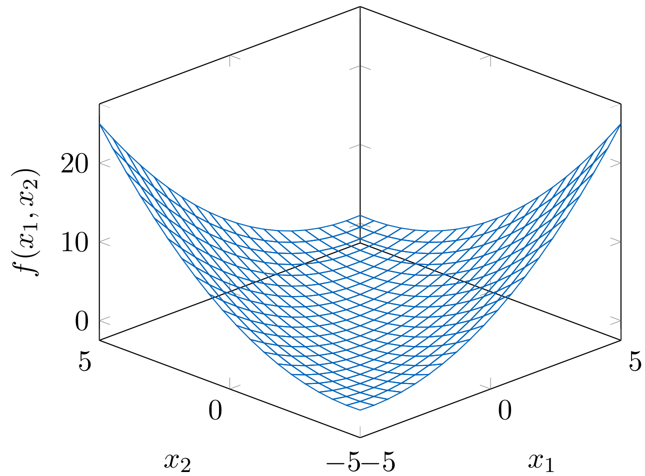
CK0031/CK0248: AP-02 (17 de novembro de 2017)

Questão 01. You are given the objective function $f(x_1, x_2) = 0.26(x_1^2 + x_2^2) - 0.48x_1x_2$ (see figure). You are requested to find its minimiser $\mathbf{x}^*$ using a descent-direction method.



Let $\mathbf{x}^{(0)} = (1, 1)'$ be the initial solution and let $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha\mathbf{d}$ $(k = 0, 1, \dots)$ be the general structure of line-search methods, with $\mathbf{d}$ the search direction and $\alpha = 1$ the fixed step-length.

- $\rightsquigarrow$ (20%) Calculate expressions for the gradient vector $\nabla f(\mathbf{x})$ and the Hessian matrix $\nabla^2 f(\mathbf{x})$
- $\rightsquigarrow$ (40%) Calculate the first 3 iterates ($\mathbf{x}^{(k)}$ and $f\left[\mathbf{x}^{(k)}\right]$, $k = 1, 2, 3$) using the gradient method

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

- $\rightsquigarrow$ (40%) Calculate the first 3 iterates ($\mathbf{x}^{(k)}$ and $f\left[\mathbf{x}^{(k)}\right]$, $k = 1, 2, 3$) using the Newton method[1]

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

[1] For calculating the inverse of a 2×2 matrix A

$$\text{Let } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \text{ then } A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Solution:

$$f(\mathbf{x}) = 0.26(x_1^2 + x_2^2) - 0.48x_1x_2 = 0.26x_1^2 + 0.26x_2^2 - 0.48x_1x_2$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f(x_1,x_2)}{\partial x_1} \\ \dfrac{\partial f(x_1,x_2)}{\partial x_2} \end{bmatrix}_{(x_1,x_2)} = \begin{bmatrix} +0.52x_1 - 0.48x_2 \\ -0.48x_1 + 0.52x_2 \end{bmatrix}_{(x_1,x_2)}$$

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial^2 f(\mathbf{x})}{\partial {x_1}^2} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial {x_2}^2} \end{bmatrix}_{(x_1,x_2)} = \begin{bmatrix} +0.52 & -0.48 \\ -0.48 & +0.52 \end{bmatrix}$$

Gradient method: $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha\mathbf{d}$, with $\mathbf{d} = -\nabla f\left[\mathbf{x}^{(k)}\right]$ and $\alpha = 1$

0|

$$\mathbf{x}^{(0)} = (1,1)'$$
$$f\left[\mathbf{x}^{(0)}\right] = 0.04$$

1|

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - \alpha \cdot \nabla f\left[\mathbf{x}^{(0)}\right] = \begin{bmatrix} x_1^{(0)} = 1 \\ x_2^{(0)} = 1 \end{bmatrix} - 1 \cdot \begin{bmatrix} +0.52 \cdot (x_1^{(0)} = 1) - 0.48 \cdot (x_2^{(0)} = 1) \\ -0.48 \cdot (x_1^{(0)} = 1) + 0.52 \cdot (x_2^{(0)} = 1) \end{bmatrix}$$
$$= \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0.04 \\ 0.04 \end{bmatrix} = \begin{bmatrix} 0.96 \\ 0.96 \end{bmatrix}$$
$$f\left[\mathbf{x}^{(1)}\right] = 0.036864 < f\left[\mathbf{x}^{(0)}\right]$$

2|

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} - \alpha \cdot \nabla f\left[\mathbf{x}^{(1)}\right] = \begin{bmatrix} x_1^{(1)} = 0.96 \\ x_2^{(1)} = 0.96 \end{bmatrix} - 1 \cdot \begin{bmatrix} +0.52 \cdot (x_1^{(1)} = 0.96) - 0.48 \cdot (x_2^{(1)} = 0.96) \\ -0.48 \cdot (x_1^{(1)} = 0.96) + 0.52 \cdot (x_2^{(1)} = 0.96) \end{bmatrix}$$
$$= \begin{bmatrix} 0.96 \\ 0.96 \end{bmatrix} - \begin{bmatrix} 0.96 \cdot 0.04 \\ 0.96 \cdot 0.04 \end{bmatrix} = \begin{bmatrix} 0.9216 \\ 0.9216 \end{bmatrix}$$
$$f\left[\mathbf{x}^{(2)}\right] = 0.0339738624 < f\left[\mathbf{x}^{(1)}\right]$$

3|

$$\mathbf{x}^{(3)} = \mathbf{x}^{(2)} - \alpha \cdot \nabla f\left[\mathbf{x}^{(2)}\right] = \begin{bmatrix} x_1^{(2)} = 0.9216 \\ x_2^{(2)} = 0.92126 \end{bmatrix} - 1 \cdot \begin{bmatrix} +0.52 \cdot (x_1^{(2)} = 0.9216) - 0.48 \cdot (x_2^{(2)} = 0.9216) \\ -0.48 \cdot (x_1^{(2)} = 0.9216) + 0.52 \cdot (x_2^{(2)} = 0.9216) \end{bmatrix}$$
$$= \begin{bmatrix} 0.9216 \\ 0.9216 \end{bmatrix} - \begin{bmatrix} 0.9216 \cdot 0.04 \\ 0.9216 \cdot 0.04 \end{bmatrix} = \begin{bmatrix} 0.884736 \\ 0.884736 \end{bmatrix}$$
$$f\left[\mathbf{x}^{(3)}\right] = 0.03131031159 < f\left[\mathbf{x}^{(2)}\right]$$

Newton's method: $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha\mathbf{d}$, with $\mathbf{d} = -\left[\nabla^2 f\left[\mathbf{x}^{(k)}\right]\right]^{-1} \nabla f\left[\mathbf{x}^{(k)}\right]$ and $\alpha = 1$

$$\left[\nabla^2 f(\mathbf{x})\right]^{-1} = \frac{1}{\frac{\partial^2 f(\mathbf{x})}{\partial x_1{}^2}\frac{\partial^2 f(\mathbf{x})}{\partial x_2{}^2} - \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2}\frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1}} \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_2{}^2} & -\frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} \\ -\frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1{}^2} \end{bmatrix}_{(x_1,x_2)}$$
$$= \frac{1}{(0.52)^2 - (-0.48)^2} \begin{bmatrix} 0.52 & 0.48 \\ 0.48 & 0.52 \end{bmatrix} = \begin{bmatrix} 13 & 12 \\ 12 & 13 \end{bmatrix}$$

$$\mathbf{d} = -\left[\nabla^2 f\left[\mathbf{x}^{(k)}\right]\right]^{-1} \nabla f\left[\mathbf{x}^{(k)}\right] = -\begin{bmatrix} 13 & 12 \\ 12 & 13 \end{bmatrix} \begin{bmatrix} +0.52x_1^{(k)} - 0.48x_2^{(k)} \\ -0.48x_1^{(k)} + 0.52x_2^{(k)} \end{bmatrix}_{\left[x_1^{(k)}, x_2^{(k)}\right]} = -\begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} = -\mathbf{x}^{(k)}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \mathbf{x}^{(k)} = \mathbf{0}, \text{ for } k = 0, 1, \dots$$

0|

$$\mathbf{x}^{(0)} = (1, 1)'$$
$$f\left[\mathbf{x}^{(0)}\right] = 0.04$$

1|

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - \alpha\left[\nabla^2 f\left[\mathbf{x}^{(0)}\right]\right]^{-1} \nabla f\left[\mathbf{x}^{(0)}\right] = \mathbf{x}^{(0)} - \mathbf{x}^{(0)} = \begin{bmatrix} x_1^{(0)} = 1 \\ x_2^{(0)} = 1 \end{bmatrix} - \begin{bmatrix} x_1^{(0)} = 1 \\ x_2^{(0)} = 1 \end{bmatrix} = \mathbf{0}$$
$$f\left[\mathbf{x}^{(1)}\right] = 0 < f\left[\mathbf{x}^{(0)}\right]$$

2|

$$\mathbf{x}^{(1)} = \mathbf{x}^{(1)} - \alpha\left[\nabla^2 f\left[\mathbf{x}^{(1)}\right]\right]^{-1} \nabla f\left[\mathbf{x}^{(1)}\right] = \mathbf{x}^{(1)} - \mathbf{x}^{(1)} = \begin{bmatrix} x_1^{(1)} = 0 \\ x_2^{(1)} = 0 \end{bmatrix} - \begin{bmatrix} x_1^{(1)} = 0 \\ x_2^{(1)} = 0 \end{bmatrix} = \mathbf{0}$$
$$f\left[\mathbf{x}^{(2)}\right] = 0 = f\left[\mathbf{x}^{(1)}\right]$$

3|

$$\mathbf{x}^{(3)} = \cdots$$
$$f\left[\mathbf{x}^{(3)}\right] = \cdots$$

At $\mathbf{x}^* = \mathbf{0}$ first- and second-order necessary conditions for $\mathbf{x}^*$ to be a local minimiser are satisfied

- $\nabla f(\mathbf{x}) = \mathbf{0}$
- $\nabla^2 f(\mathbf{x}) \succeq 0$

Thus $\mathbf{x}^*$ is a local minimiser. As function is convex, the local minimiser is also a global minimiser.