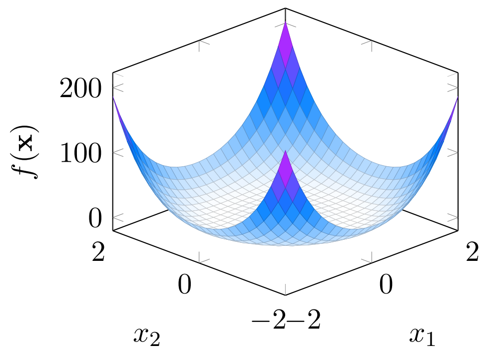
CK0031/CK0248: AP-01 (10 de outubro de 2018)

Questão 01. You are given the objective function $f(\mathbf{x}) = (x_1 + x_2)^2 + [2(x_1^2 + x_2^2 - 1) - 1/3]^2$ (see figure)[1]. You are requested to find its minimiser $\mathbf{x}^*$ using a descent-direction method.



Let $\mathbf{x}^{(0)} = (\sqrt{7/6}, 0)^\top$ be the initial solution and let $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$ $(k = 0, 1, \dots)$ be the general line-search method, with $\mathbf{d}^{(k)}$ the search direction and $\alpha_k = 1$ the fixed step-length.

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

$$\mathbf{d}^{(k)} = -\Big[\nabla^2 f\big[\mathbf{x}^{(k)}\big]\Big]^{-1} \nabla f\big[\mathbf{x}^{(k)}\big], \quad (k = 0, 1, 2).$$

Solution: The objective function, its gradient and its Hesse matrix:

$$f(\mathbf{x}) = f(x_1, x_2, \dots, x_n) = f(x_1, x_2) = (x_1 + x_2)^2 + [2(x_1^2 + x_2^2 - 1) - 1/3]^2;$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f(x_1, x_2)}{\partial x_1} \\ \dfrac{\partial f(x_1, x_2)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 + 2x_2 + 8x_1(2x_1^2 + 2x_2^2 - 7/3) \\ 2x_1 + 2x_2 + 8x_2(2x_1^2 + 2x_2^2 - 7/3) \end{bmatrix};$$

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_n} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f(x_1, x_2)}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(x_1, x_2)}{\partial x_2 \partial x_2} \end{bmatrix}$$

$$= \begin{bmatrix} 48x_1^2 + 16x_2^2 - 50/3 & 32x_1x_2 + 2 \\ 32x_1x_2 + 2 & 16x_1^2 + 48x_2^2 - 50/3 \end{bmatrix}.$$

[1] $(a + b + c)^2 = a^2 + b^2 + c^2 + 2ab + 2bc + 2ca$.

[2] For calculating the inverse of a 2×2 matrix A: Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $A^{-1} = \dfrac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Iterations ($\mathbf{x}^{(0)} = (\sqrt{7/6})^\top.0$, $\alpha_k = \alpha = 1$) with Newton's directions $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \{\nabla^2 f[\mathbf{x}^{(k)}]\}^{-1} \nabla f[\mathbf{x}^{(k)}]$

1. $k = 0$

$$
\begin{aligned}
\mathbf{x}^{(1)} &= \mathbf{x}^{(0)} - \{\nabla^2 f[\mathbf{x}^{(0)}]\}^{-1} \nabla f[\mathbf{x}^{(0)}] \\
&= \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} - \begin{bmatrix} 48x_1^{(0)^2} + 16x_2^{(0)^2} - 50/3 & 32x_1^{(0)}x_2^{(0)} + 2 \\ 32x_1^{(0)}x_2^{(0)} + 2 & 16x_1^{(0)^2} + 48x_2^{(0)^2} - 50/3 \end{bmatrix}^{-1} \\
&\times \begin{bmatrix} 2x_1^{(0)} + 2x_2^{(0)} + 8x_1[2x_1^{(0)^2} + 2x_2^{(0)^2} - 7/3] \\ 2x_1^{(0)} + 2x_2^{(0)} + 8x_2[2x_1^{(0)^2} + 2x_2^{(0)^2} - 7/3] \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{7/6} \\ 0 \end{bmatrix} - \begin{bmatrix} 48\sqrt{7/6}^2 + 16 \cdot 0^2 - 50/3 & 32\sqrt{7/6} \cdot 0 + 2 \\ 32\sqrt{7/6} \cdot 0 + 2 & 16\sqrt{7/6}^2 + 48 \cdot 0^2 - 50/3 \end{bmatrix}^{-1} \\
&\times \begin{bmatrix} 2\sqrt{7/6} + 2 \cdot 0 + 8\sqrt{7/6}[2\sqrt{7/6}^2 + 2 \cdot 0^2 - 7/3] \\ 2\sqrt{7/6} + 2 \cdot 0 + 8 \cdot 0 \cdot [2\sqrt{7/6}^2 + 2 \cdot 0^2 - 7/3] \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{7/6} \\ 0 \end{bmatrix} - \begin{bmatrix} 48(7/6) - 50/3 & 2 \\ 2 & 16(7/6) - 50/3 \end{bmatrix}^{-1} \begin{bmatrix} 2\sqrt{7/6} + 8\sqrt{7/6}[2(7/6) - 7/3] \\ 2\sqrt{7/6} \end{bmatrix} \\
&= \begin{bmatrix} 1.08 \\ 0 \end{bmatrix} - \begin{bmatrix} 39.33 & 2 \\ 2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 2.16 \\ 2.16 \end{bmatrix} = \begin{bmatrix} 1.08 \\ 0 \end{bmatrix} - \frac{1}{74.66} \begin{bmatrix} 2 & -2 \\ -2 & 39.33 \end{bmatrix} \begin{bmatrix} 2.16 \\ 2.16 \end{bmatrix} \\
&= \begin{bmatrix} 1.17 \\ 0 \end{bmatrix} - \frac{1}{74.66} \begin{bmatrix} 2 \cdot 2.16 - 2 \cdot 2.16 \\ -2 \cdot 2.16 + 39.33 \cdot 2.16 \end{bmatrix} = \begin{bmatrix} 1.08 \\ 0 \end{bmatrix} - \frac{1}{74.66} \begin{bmatrix} 0 \\ 80.63 \end{bmatrix} = \begin{bmatrix} 1.08 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1.08 \end{bmatrix} \\
&\simeq \begin{bmatrix} 1.08 \\ -1.08 \end{bmatrix} \quad \rightsquigarrow \quad f[\mathbf{x}^{(1)}] \simeq 5.44
\end{aligned}
$$

2. $k = 1$

$$
\begin{aligned}
\mathbf{x}^{(2)} &= \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} - \begin{bmatrix} 48x_1^{(1)^2} + 16x_2^{(1)^2} - 50/3 & 32x_1^{(1)}x_2^{(1)} + 2 \\ 32x_1^{(1)}x_2^{(1)} + 2 & 16x_1^{(1)^2} + 48x_2^{(1)^2} - 50/3 \end{bmatrix}^{-1} \\
&\times \begin{bmatrix} 2x_1^{(1)} + 2x_2^{(1)} + 8x_1[2x_1^{(1)^2} + 2x_2^{(1)^2} - 7/3] \\ 2x_1^{(1)} + 2x_2^{(1)} + 8x_2[2x_1^{(1)^2} + 2x_2^{(1)^2} - 7/3] \end{bmatrix} \\
&= \begin{bmatrix} 1.08 \\ -1.08 \end{bmatrix} - \begin{bmatrix} 48(1.08)^2 + 16(-1.08)^2 - 50/3 & 32(1.08)(-1.08) + 2 \\ 32(1.08)(-1.08) + 2 & 16(1.08)^2 + 48(-1.08)^2 - 50/3 \end{bmatrix}^{-1} \\
&\times \begin{bmatrix} 2(1.08) + 2(-1.08) + 8(1.08)[2(1.08)^2 + 2(-1.08)^2 - 7/3] \\ 2(1.08) + 2(-1.08) + 8(-1.08)[2(1.08)^2 + 2(-1.08)^2 - 7/3] \end{bmatrix} \\
&= \begin{bmatrix} 1.08 \\ -1.08 \end{bmatrix} - \begin{bmatrix} 58 & -35.33 \\ -35.33 & 58 \end{bmatrix}^{-1} \begin{bmatrix} 20.16 \\ -20.16 \end{bmatrix} = \begin{bmatrix} 1.08 \\ -1.08 \end{bmatrix} - \frac{1}{2116} \begin{bmatrix} 58 & 35.33 \\ 35.33 & 58 \end{bmatrix} \begin{bmatrix} 20.16 \\ -20.16 \end{bmatrix} \\
&= \begin{bmatrix} 1.08 \\ -1.08 \end{bmatrix} - \frac{1}{2116} \begin{bmatrix} 58 \cdot 20.16 - 35.33 \cdot 20.16 \\ 35.33 \cdot 20.16 - 58 \cdot 20.16 \end{bmatrix} = \begin{bmatrix} 1.08 \\ -1.08 \end{bmatrix} - \begin{bmatrix} 0.22 \\ -0.22 \end{bmatrix} \\
&\simeq \begin{bmatrix} 0.86 \\ -0.86 \end{bmatrix} \quad \rightsquigarrow \quad f[\mathbf{x}^{(2)}] \simeq 0.43
\end{aligned}
$$

3. $k = 2$

$$\mathbf{x}^{(3)} = \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \end{bmatrix} - \begin{bmatrix} 48x_1^{(2)^2} + 16x_2^{(2)^2} - 50/3 & 32x_1^{(2)}x_2^{(2)} + 2 \\ 32x_1^{(2)}x_2^{(2)} + 2 & 16x_1^{(2)^2} + 48x_2^{(2)^2} - 50/3 \end{bmatrix}^{-1}$$

$$\times \begin{bmatrix} 2x_1^{(1)} + 2x_2^{(1)} + 8x_1[2x_1^{(1)^2} + 2x_2^{(1)^2} - 7/3] \\ 2x_1^{(1)} + 2x_2^{(1)} + 8x_2[2x_1^{(1)^2} + 2x_2^{(1)^2} - 7/3] \end{bmatrix}$$

$$= \begin{bmatrix} 0.86 \\ -0.86 \end{bmatrix} - \begin{bmatrix} 48(0.86)^2 + 16(-0.86)^2 - 50/3 & 32(0.86)(-0.86) + 2 \\ 32(0.86)(-0.86) + 2 & 16(0.86)^2 + 48(-0.86)^2 - 50/3 \end{bmatrix}^{-1}$$

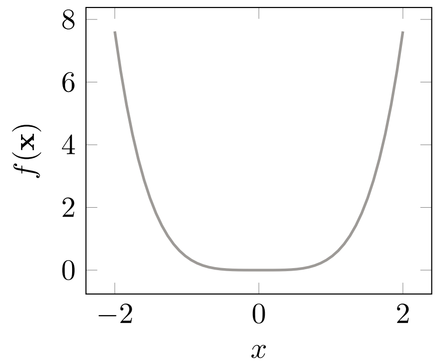
$$\times \begin{bmatrix} 2(0.86) + 2(-0.86) + 8(0.86)\left[2(0.86)^2 + 2(-0.86)^2 - 7/3\right] \\ 2(0.86) + 2(-0.86) + 8(-0.86)\left[2(0.86)^2 + 2(-0.86)^2 - 7/3\right] \end{bmatrix}$$

$$= \begin{bmatrix} 0.86 \\ -0.86 \end{bmatrix} - \begin{bmatrix} 30.67 & -21.67 \\ -21.67 & 30.67 \end{bmatrix}^{-1} \begin{bmatrix} 4.30 \\ -4.30 \end{bmatrix} = \begin{bmatrix} 0.86 \\ -0.86 \end{bmatrix} - \frac{1}{471} \begin{bmatrix} 30.67 & 21.67 \\ 21.67 & 30.67 \end{bmatrix} \begin{bmatrix} 4.30 \\ -4.30 \end{bmatrix}$$

$$= \begin{bmatrix} 0.86 \\ -0.86 \end{bmatrix} - \frac{1}{471} \begin{bmatrix} 30.67 \cdot 4.30 - 21.67 \cdot 4.30 \\ 21.67 \cdot 4.30 - 30.67 \cdot 4.30 \end{bmatrix} = \begin{bmatrix} 0.86 \\ -0.86 \end{bmatrix} - \frac{1}{471} \begin{bmatrix} 38.7 \\ -38.7 \end{bmatrix} = \begin{bmatrix} 0.86 \\ -0.86 \end{bmatrix} - \begin{bmatrix} 0.082 \\ -0.082 \end{bmatrix}$$

$$\simeq \begin{bmatrix} 0.779 \\ -0.779 \end{bmatrix} \quad \rightsquigarrow \quad f[\mathbf{x}^{()}] \simeq 0.009$$

Questão 02. You are given the objective function $f(\mathbf{x}) = (11/546)x^6 - (38/364)x^4 + (1/2)x^2$ (see figure). You are requested to find its minimiser $\mathbf{x}^*$ using a descent-direction method.



Let $\mathbf{x}^{(0)} = 1.01$ be the initial solution and let $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$ $(k = 0, 1, \dots)$ be the general line-search method, with $\mathbf{d}^{(k)}$ the search direction and $\alpha_k = 0.1$ the fixed step-length.

- ⇝ (10%) Calculate expressions for the gradient vector $\nabla f(\mathbf{x})$ and the Hessian matrix $\nabla^2 f(\mathbf{x})$;
- ⇝ (30%) Calculate the first 3 iterates ($\mathbf{x}^{(k)}$ and $f\big[\mathbf{x}^{(k)}\big]$, $k = 1, 2, 3$) using the Newton method

$$\mathbf{d}^{(k)} = -\Big[\nabla^2 f\big[\mathbf{x}^{(k)}\big]\Big]^{-1} \nabla f\big[\mathbf{x}^{(k)}\big], \quad (k = 0, 1, 2).$$

Solution: The objective function, its gradient and its Hesse matrix:

$$f(\mathbf{x}) = f(x_1, x_2, \dots, x_n) = f(x) = (11/546)x^6 - (38/364)x^4 + (1/2)x^2;$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix} = \frac{\mathrm{d} f(x)}{\mathrm{d} x} = (11/91)x^5 - (38/91)x^4 + x;$$

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_n} \end{bmatrix} = \frac{\mathrm{d}^2 f(x)}{\mathrm{d} x^2} = (55/91)x^4 - (114/91)x^2 + 1.$$

Iterations ($\mathbf{x}^{(0)} =$, $\alpha_k = \alpha =$) with Newton's directions $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \big\{\nabla^2 f\big[\mathbf{x}^{(k)}\big]\big\}^{-1} \nabla f\big[\mathbf{x}^{(k)}\big]$

1. $k = 0$

$$\begin{aligned} x^{(1)} &= x^{(0)} - \alpha\big\{\mathrm{d}^2 f[x^{(0)}]/\mathrm{d}x^2\big\}^{-1}\big\{\mathrm{d}f[x^{(0)}]/\mathrm{d}x\big\} \\ &= x^{(0)} - \alpha\big[(55/91)x^{(0)^4} - (114/91)x^{(0)^2} + 1\big]^{-1}\big[(11/91)x^{(0)^5} - (38/91)x^{(0)^4} + x^{(0)}\big] \\ &= 1.01 - 0.1[(55/91)1.01^4 - (114/91)1.01^2 + 1]^{-1}[(11/91)1.01^5 - (38/91)1.01^4 + 1.01] \\ &\simeq 0.81 \quad \rightsquigarrow \quad f[x^{(1)}] \simeq 0.29 \end{aligned}$$

2. $k = 1$

$$
\begin{aligned}
x^{(2)} &= x^{(1)} - \alpha\big[(55/91)x^{(1)^4} - (114/91)x^{(1)^2} + 1\big]^{-1}\big[(11/91)x^{(1)^5} - (38/91)x^{(1)^4} + x^{(1)}\big] \\
&= 1.01 - 0.1[(55/91)0.81^4 - (114/91)0.81^2 + 1]^{-1}[(11/91)0.81^5 - (38/91)0.81^4 + 0.81] \\
&\simeq 0.66 \quad \rightsquigarrow \quad f[x^{(2)}] \simeq 0.20
\end{aligned}
$$

3. $k = 2$

$$
\begin{aligned}
x^{(3)} &= x^{(2)} - \alpha\big[(55/91)x^{(2)^4} - (114/91)x^{(2)^2} + 1\big]^{-1}\big[(11/91)x^{(2)^5} - (38/91)x^{(2)^4} + x^{(2)}\big] \\
&= 1.01 - 0.1[(55/91)0.66^4 - (114/91)0.66^2 + 1]^{-1}[(11/91)0.66^5 - (38/91)0.66^4 + 0.66] \\
&\simeq 0.57 \quad \rightsquigarrow \quad f[x^{(3)}] \simeq 0.15
\end{aligned}
$$