Série N°3 (Optimisation non linéaire)

Exercice 1 *Utilisez la méthode de Newton pour trouver un point minimisant de*

$$f(x,y) = 5x^4 + 6y^4 - 6\,x^2 + 2\,x\,y + 5\,y^2 + 15\,x - 7\,y + 13$$

Partez de $(x_0\,,\,y_0) = (1\,,\,1)$.

Solution: Calculant d'abord le gradient g et le hessien h de la fonction f.

$$g(x,y) = \big(20x^3 - 12x + 2y + 15\,,\ 24y^3 + 2x + 10y - 7\big)^T$$
$$h(x,y) = \begin{bmatrix} 60\,x^2 - 12 & 2 \\ 2 & 72\,y^2 + 10 \end{bmatrix}$$

l'inverse de h est donnée formellement par

$$[H(x,y)]^{-1} = \begin{bmatrix} \frac{1}{2}\frac{36\,y^2+5}{1080\,x^2y^2+150\,x^2-216\,y^2-31} & -\frac{1}{2}\left(1080\,x^2y^2 + 150\,x^2 - 216\,y^2 - 31\right)^{-1} \\ -\frac{1}{2}\left(1080\,x^2y^2 + 150\,x^2 - 216\,y^2 - 31\right)^{-1} & 3\,\frac{5\,x^2-1}{1080\,x^2y^2+150\,x^2-216\,y^2-31} \end{bmatrix}$$

La méthode de Newton est donnée par

$$\begin{cases} x_{k+1} & = x_k + d_k \\ d_k & = -\left[h^{-1}(x_k,y_k)\right] g(x_k,y_k) \end{cases} \qquad (1)$$

En substituant les expressions du gradient et du hessien dans (1), on obtient

$$x_{k+1} = x_k - \frac{1}{2}\,\frac{(36\,{y_k}^2+5)\,(20\,{x_k}^3 - 12\,x_k + 2\,y_k + 15)}{1080\,{x_k}^2{y_k}^2 + 150\,{x_k}^2 - 216\,{y_k}^2 - 31} + \frac{1}{2}\,\frac{24\,{y_k}^3 + 2\,x_k + 10\,y_k - 7}{1080\,{x_k}^2{y_k}^2 + 150\,{x_k}^2 - 216\,{y_k}^2 - 31}$$
$$y_{k+1} = y_k + \frac{1}{2}\,\frac{20\,{x_k}^3 - 12\,x_k + 2\,y_k + 15}{1080\,{x_k}^2{y_k}^2 + 150\,{x_k}^2 - 216\,{y_k}^2 - 31} - 3\,\frac{(5\,{x_k}^2 - 1)\,(24\,{y_k}^3 + 2\,x_k + 10\,y_k - 7)}{1080\,{x_k}^2{y_k}^2 + 150\,{x_k}^2 - 216\,{y_k}^2 - 31}$$

Les premières itérations de méthode Newton en partant de $(x_0\,,\,y_0) = (1\,,\,1)$:

— $(x_1, y_1) = (0.4933875890\,, 0.6586978637)$
— $(x_2,\, y_2) = (-4.464349613\,,\, 0.7188967310)$
— $(x_3\,,\, y_3) = (-3.020177683\,,\, 0.6539652503)$
— $(x_4\,,\, y_4) = (-2.088864692\,,\, 0.6031153494)$
— $\cdots\cdots$
— $(x_{10}\,,\, y_{10}) = (-1.142054928\,,\, 0.5433724812)$

avec $g(x_{10}\,,\,y_{10}) = (2.2\,10^{-8}\,,\,1.0\,10^{-9})$. Le minimum est donc $(x^*\,,\,y^*) \approx (-1.142054928\,,\,0.5433724812)$ ■

Exercice 2 *Utilisez la méthode de Newton modifiée (avec une recherche unidimensionnelle) pour la résolution de*

$$\min_{(x,y)\in\mathbb{R}^n} \; 2\,x^2 + y^2 - 2\,x\,y + 2\,x^3 + x^4$$

Effectuez seulement les deux premières itérations et partez de $(2\,,\,-1)$.

Solution: Calculant d'abord le gradient g et le hessien h de la fonction f.

$$g(x,y) = \left(4\,x - 2\,y + 6\,x^2 + 4\,x^3, 2\,y - 2\,x\right)^T$$

$$H(x,y) = \begin{bmatrix} 4 + 12\,x + 12\,x^2 & -2 \\ -2 & 2 \end{bmatrix}$$

l'inverse de h est donnée formellement par

$$[H(x,y)]^{-1} = \begin{bmatrix} 1/2\,\left(1 + 6\,x + 6\,x^2\right)^{-1} & 1/2\,\left(1 + 6\,x + 6\,x^2\right)^{-1} \\ 1/2\,\left(1 + 6\,x + 6\,x^2\right)^{-1} & \frac{1+3\,x+3\,x^2}{1+6\,x+6\,x^2} \end{bmatrix}$$

La méthode de Newton modifiée est donnée par

$$\begin{cases} x_{k+1} & = x_k + t_k\,d_k \\ d_k & = -\left[h^{-1}(x_k, y_k)\right] g(x_k, y_k) \\ t_k & = \arg\min\{f(x_k + td_k),\; t > 0) \end{cases} \tag{2}$$

A chaque itération on calculera t_k comme solution de l'équation

$$\theta'(t) \equiv \frac{d}{dt} f(x_k + td_k) = 0$$

En substituant les expression du gradient et du hessien dans (1), on obtient

$$x_{k+1} = x_k - t_k \left(1/2\,\frac{4\,x_k - 2\,y_k + 6\,{x_k}^2 + 4\,{x_k}^3}{1 + 6\,x_k + 6\,{x_k}^2} + 1/2\,\frac{2\,y_k - 2\,x_k}{1 + 6\,x_k + 6\,{x_k}^2}\right)$$

$$y_{k+1} = y_k - t_k \left(1/2\,\frac{4\,x_k - 2\,y_k + 6\,{x_k}^2 + 4\,{x_k}^3}{1 + 6\,x_k + 6\,{x_k}^2} + \frac{\left(1 + 3\,x_k + 3\,{x_k}^2\right)\left(2\,y_k - 2\,x_k\right)}{1 + 6\,x_k + 6\,{x_k}^2}\right)$$

Les premières itérations de méthode Newton modifiée en partant de $(x_0\,,\,y_0) = (2\,,\,1)$:

0. $\theta'(t) = -\frac{53}{37} + \frac{949}{1369}\,t \;\Rightarrow t_0 = 1961/949 \approx 2.066385669$

1. $(x_1, y_1) = (2.066385669\,, 0.324552160)$, $\theta'(t) = -1.823049784 + 1.638056870\,t \Rightarrow t_1 = 1.112934366$

2. $(x_2,\,y_2) = (0.1041219991\,,\,-0.016309591)$, $\theta'(t) = -0.009182571 + 0.0082186965\,t \;\Rightarrow\; t_2 = 1.1172783$

■

Exercice 3 *Appliquez la méthode de quasi-Newton de rang un à la résolution de*

$$\min \frac{1}{2} x^T A x - b^T x$$

Pour

$$A = \begin{pmatrix} 5 & 2 & 1 \\ 2 & 7 & 3 \\ 1 & 3 & 9 \end{pmatrix}, \quad b = \begin{pmatrix} -9 \\ 0 \\ 8 \end{pmatrix}$$

Initialisez la méthode avec $x_0 = (0,0,0)$ *et* $S_0 = I$. .

Solution: Calculant explicitement le gradient g et le hessien H de la fonction f.

$$g(x) = Ax + b^T = (5\,x_1 + 2\,x_2 + x_3 + 9\,,\, 2\,x_1 + 7\,x_2 + 3\,x_3\,, x_1 + 3\,x_2 + 9\,x_3 - 8)^T$$

$$H(x) = A = \begin{pmatrix} 5 & 2 & 1 \\ 2 & 7 & 3 \\ 1 & 3 & 9 \end{pmatrix}$$

La méthode SR1 est donnée par

$$\begin{cases} x_{k+1} & = x_k + t_k\, d_k \\ d_k & = -S_k g(x_k, y_k) \\ t_k & = \arg\min\{f(x_k + t d_k),\, t > 0) \end{cases} \tag{3}$$

avec

$$S_{k+1} = S_k + \frac{(\delta_k - S_k\gamma_k)(\delta_k - S_k\gamma_k)^\top}{(\delta_k - S_k\gamma_k)^\top \gamma_k}$$

Où

$$\delta_k = x_{k+1} - x_k = t_k d_k$$
$$\gamma_k = g(x_{k+1}) - g(x_k)$$

Puisque $f$est quadratique, on peut calculer exactement t_k à chaque itération. En effet

$$\theta = f(x_k + t d_k) \Rightarrow \theta'(t) = g(x_k + t d_k)^T d_k \Rightarrow t_k = -\frac{g(x_k)^T d_k}{d_k^T A d_k}.$$

Maintenant, si $S_0 = I$ et $x_0 = (0,0,0)$, les premières itérations de méthode SR1 :

Itération 0 : $x_0 = (0,0,0)$, $S_0 = I \;\; \Rightarrow d_0 = -S_0 g(x_0) = (-9,0,8)^T$, $\; t_0 = \dfrac{145}{837}$.

Itération 1 : $x_1 = x_0 + t_0 d_0 = (-1.559139785,\, 0,\, 1.385902031)$,

$$\delta_0 = (-6.409796894\,,\; 1.039426523\,,\, 10.91397850),\;\; \gamma_0 = (-1.559139785\,,\, 0\,,\, 1.385902031)$$

$$S_1 = \begin{bmatrix} 0.3556802439 & -0.3249899334 & 0.0969858512 \\ -0.3249899334 & 0.7141171660 & -0.2588782622 \\ 0.0969858512 & -0.2588782622 & 0.2085991439 \end{bmatrix}$$

$$d_1 = (3.170196781, 0.9151421641, 1.774705204)^T \, , \, t_1 = 0.1224500439$$

Itération 2 : $x_2 = (-1.947330520\, ,\, -.1120591982\, ,\, 1.168589301)\delta_1 = (-0.388190735, -0.1120591982, -0.21731$
$(-2.382384801, -2.212734047, -2.680182895)$

$$S_2 = \begin{bmatrix} 0.8271985895 & 0.03702887367 & 0.3394313422 \\ 0.03702887367 & 0.9920652414 & -0.07273528759 \\ 0.3394313422 & -0.07273528759 & 0.3332598632 \end{bmatrix}$$

$$d_2 = (0.4779048647, -0.9659424995, 0.37266881)^T \, , \, t_2 = 0.2502905830$$

Itération 3 : $x_3 = (-2.066945607\, ,\, 0.1297071131\, ,\, 1.075313807)\delta_2 = (-0.119615087, 0.2417663113, -0.0932754$
$-0.2078183116\, ,\, 1.173307523\, ,\, -0.233795608$

$$S_3 = \begin{bmatrix} 0.8271985895 & 0.03702887367 & 0.3394313422 \\ 0.03702887367 & 0.9920652414 & -0.07273528759 \\ 0.3394313422 & -0.07273528759 & 0.3332598632 \end{bmatrix}$$

$$d_3 = \left(1.485355643\, 10^{-9}\ ,\ -6.903765585\ 10^{-10}\, ,\, 9.539748907\ 10^{-10}\right)^T ,\ \ t_3 = 1.0$$

Itération 4 : $x_4 = (-2.066945606\, ,\, .1297071124\, ,\, 1.075313808)$
On s'arrête à la 4ième itération puisque $g(x_4) \approx 0$.

■

Exercice 4 *Montrer que Toute matrice A symétrique et définie positive admet une racine carré i.e. il existe une matrice $A^{1/2}$ telle que*

$$A = A^{1/2} A^{1/2}$$

Solution: Utiliser le fait que toute matrice symétrique est diagonalisable

■

Exercice 5 *Montrer que si $f : \mathbb{R}^n \to \mathbb{R}$ est une fonction fortement convexe alors*

$$\gamma_k^\top \delta_k > 0$$

pour tout $x_k, x_{k+1} \in \mathbb{R}^n$.

Solution: Faire le développement de f au voisinage de x_k dans la direction de δ_k.

■

Exercice 6 *Montrer que la matrice*

$$S_{k+1} = (I - \frac{\delta_k \gamma_k^\top}{\gamma_k^\top \delta_k}) S_k (I - \frac{\gamma_k \delta_k^\top}{\gamma_k^\top \delta_k}) + \frac{\delta_k \delta_k^\top}{\gamma_k^\top \delta_k}$$

est l'inverse de la matrice

$$B_{k+1} = B_k + \frac{\gamma_k \gamma_k^\top}{\gamma_k^\top \delta_k} - \frac{B_k \delta_k \delta_k^\top B_k}{\delta_k^\top B_k \delta_k}$$

Solution: Par récurrence.

■

Exercice 7 *Utiliser la formule de Sherman–Morrison pour montrer que la matrice*

$$S_{k+1} = S_k + \frac{(\delta_k - S_k \gamma_k)(\delta_k - S_k \gamma_k)^\top}{(\delta_k - S_k \gamma_k)^\top \gamma_k}$$

est l'inverse de la matrice

$$B_{k+1} = B_k + \frac{(\gamma_k - B_k \delta_k)(\gamma_k - B_k \delta_k)^\top}{(\gamma_k - B_k \delta_k)^\top \delta_k}$$

Solution: Application directe.

■

Exercice 8 *Si $h(t) = 1 - t + \ln t$. Montrer que $h(t) < 0$ pour tout $t > 0$.*

Solution: Montrer que $h(t)$ est concave et que $h(1) = 0$ et $h'(1) = 0$.

■

Exercice 9 *Soit B une matrice définie positive et que $\lambda_i \in \sigma(B)$, $i = 1, 2, ..., n$ telle que $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Montrer que la fonction ψ définie par $\psi(B) = Tr(B) - \ln(\det(B))$ peut être écrite sous la forme*

$$\psi(B) = \sum_{i=1}^{i=n} (\lambda_i - \ln \lambda_i).$$

Utiliser cette forme pour montrer que $\psi(B) > 0$.

Solution: Directe.

■

Exercice 10 *En utilisant la définition de la trace d'une matrice Montrer que*

$$Tr(B_{k+1}) = Tr(B_k) + \frac{\|\gamma_k\|^2}{\gamma_k^\top \delta_k} - \frac{\|B_k \delta_k\|^2}{\delta_k^\top B_k \delta_k}$$

où

$$B_{k+1} = B_k + \frac{\gamma_k \gamma_k^\top}{\gamma_k^\top \delta_k} - \frac{B_k \delta_k \delta_k^\top B_k}{\delta_k^\top B_k \delta_k}$$

Solution: Directe.

■