Série N°4 (Optimisation non linéaire)

Master 1 COTA-MA (2020–2021)

Exercice 1

1) Considérons la mise à jour de rang un

$$H_1 = I + uv^T$$

où $u, v \in \mathbb{R}^n$ et $u \neq 0$. Trouver les valeurs propres de H_1, puis montrer que

$$\det(H_1) = 1 + u^T v.$$

2) Soit $u_1,\, u_2,\, v_1,\, v_2 \in \mathbb{R}^n$ et considérons la mise à jour de rang deux

$$H_2 = I + u_1 u_2^T + v_1 v_2^T.$$

Montrer que

$$H_2 = \left(I + u_1 u_2^T\right)\left[I + \left(I + u_1 u_2^T\right)^{-1} v_1 v_2^T\right],$$

puis, en utilisant la question précédente et le théorème de Sherman-Morrison-Woodburg, montrer que

$$\det(H_2) = \left(1 + u_1^T u_2\right)\left(1 + v_1^T v_2\right) - \left(u_1^T v_2\right)\left(u_2^T v_2\right).$$

Solution: .

1) Soit $x \in \mathbb{R}^n$ un vecteur propre associé à la valeur propre λ de H_1, on a

$$\left(I + uv^T\right)x = \lambda x \Longleftrightarrow uv^T x = (\lambda - 1)x,$$

donc x est soit parallel à $u \neq 0$ ou orthogonal à v. Si x est orthogonal à v, alors la valeur propre correspondante est égale à 1 ; sinon la valeur propre correspondante est égale à $1 + u^T v$. Puisque le determinant et le produit des valeurs propre d'une matrice, alors

$$\det(H_1) = 1 + u^T v.$$

2) Soit $u_1,\, u_2,\, v_1,\, v_2 \in \mathbb{R}^n$ et considérons la mise à jour de rang deux

$$H_2 = I + u_1 u_2^T + v_1 v_2^T.$$

Il est facile de remarquer que

$$\left(I + u_1 u_2^T\right)\left[I + \left(I + u_1 u_2^T\right)^{-1} v_1 v_2^T\right] = I + u_1 u_2^T + v_1 v_2^T.$$

En utilisant la question précédente et le théorème de Sherman-Morrison-Woodburg, on a

$$\begin{aligned}\det(I+u_1u_2^T+v_1v_2^T) &= \left(1+u_1^Tu_2\right)\left[1+v_2^T\left(I+u_1^Tv_2\right)^{-1}v_1^T\right],\\ &= \left(1+u_1^Tu_2\right)\left[1+v_2^T\left(I-\frac{u_1u_2^T}{1+u_1^Tu_2}\right)v_1^T\right],\\ &= \left(1+u_1^Tu_2\right)\left(1+v_1^Tv_2\right)-\left(u_1^Tv_2\right)\left(u_2^Tv_2\right).\end{aligned}$$

■

Théorème (de Sherman-Morrison-Woodburg) : *Soit A une matrice $n\times n$ non singulière et $u,v\in\mathbb{R}^n$. Si $1+v^TA^{-1}u\neq 0$, alors la mise à jour de rang un $A+uv^T$ est inversible et*

$$(A+uv^T)^{-1}=A^{-1}-\frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}.$$

Exercice 2 *Soit A une matrice $n\times n$ non singulière et $u,v\in\mathbb{R}^n$. Si $1+v^TA^{-1}u\neq 0$, alors la mise à jour de rang un $A+uv^T$ est inversible et*

$$(A+uv^T)^{-1}=A^{-1}-\frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}.$$

Exercice 3 *On veut minimiser la fonction donnée par*

$$f(x)=\frac{1}{2}x^TAx-x^Tb;\quad x\in\mathbb{R}^2.$$

Où $A=\begin{pmatrix}1&0\\0&2\end{pmatrix},\quad b=\begin{pmatrix}-1\\1\end{pmatrix}$ *en utilisant la méthode SR1.*

1) *Montrer que si $\delta_k=(\delta_k^1,\delta_k^2)^T$ et $\gamma_k=(\gamma_k^1,\gamma_k^2)^T$, alors*

$$S_{k+1}=S_k+\begin{bmatrix}\frac{\left(\delta_k^1-\gamma_k^1\right)^2}{\left(\delta_k^1-\gamma_k^1\right)\gamma_k^1+\left(\delta_k^2-\gamma_k^2\right)\gamma_k^2} & \frac{\left(\delta_k^1-\gamma_k^1\right)\left(\delta_k^2-\gamma_k^2\right)}{\left(\delta_k^1-\gamma_k^1\right)\gamma_k^1+\left(\delta_k^2-\gamma_k^2\right)\gamma_k^2}\\ \frac{\left(\delta_k^1-\gamma_k^1\right)\left(\delta_k^2-\gamma_k^2\right)}{\left(\delta_k^1-\gamma_k^1\right)\gamma_k^1+\left(\delta_k^2-\gamma_k^2\right)\gamma_k^2} & \frac{\left(\delta_k^2-\gamma_k^2\right)^2}{\left(\delta_k^1-\gamma_k^1\right)\gamma_k^1+\left(\delta_k^2-\gamma_k^2\right)\gamma_k^2}\end{bmatrix}.$$

2) *Montrer que* $t_k=\arg\min\{f(x_k+td_k),\ t>0)=-\dfrac{g(x_k)^\top d_k}{d_k^\top Ad_k},\ k=0,\ 1,\ 2,\ ....$

3) *Calculer les trois premières itérations de l'algorithme SR1 pour minimiser f sur $\mathbb{R}^2$ avec $x_0=(2,2)^T$ et $S_0=I$. Que peut on déduire.*

Solution:

1)

$$S_{k+1}=S_k+\begin{bmatrix}\frac{\left(\delta_k^1-\gamma_k^1\right)^2}{\left(\delta_k^1-\gamma_k^1\right)\gamma_k^1+\left(\delta_k^2-\gamma_k^2\right)\gamma_k^2} & \frac{\left(\delta_k^1-\gamma_k^1\right)\left(\delta_k^2-\gamma_k^2\right)}{\left(\delta_k^1-\gamma_k^1\right)\gamma_k^1+\left(\delta_k^2-\gamma_k^2\right)\gamma_k^2}\\ \frac{\left(\delta_k^1-\gamma_k^1\right)\left(\delta_k^2-\gamma_k^2\right)}{\left(\delta_k^1-\gamma_k^1\right)\gamma_k^1+\left(\delta_k^2-\gamma_k^2\right)\gamma_k^2} & \frac{\left(\delta_k^2-\gamma_k^2\right)^2}{\left(\delta_k^1-\gamma_k^1\right)\gamma_k^1+\left(\delta_k^2-\gamma_k^2\right)\gamma_k^2}\end{bmatrix}.$$

2) Si f est une fonction quadratique convexe,

$$\theta(t) = f(x_k + td_k)$$

alors

$$\theta'(t_k) = g(x_k+td_k)^T d_k = \left((x_k + td_k)^\top A + b^\top\right) d_k = \left({x_k}^\top A + b^\top\right) d_k + td_k^T A d_k = g(x_k+td_k)^T d_k$$

t_k est un minimum local pour θ sur $]0\,,\,\infty[\Rightarrow \theta'(t_k) = g(x_k + td_k)^T d_k = 0 \Rightarrow t_k = -\dfrac{g(x_k)^\top d_k}{d_k^\top A d_k}$.

3) On a

$$g(x) = Ax - b = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} x - \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Puisque $x_0 = (2\,,\,2)$ et $S_1 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, alors

$$g(x_0) = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$
$$d_0 = -S_0 g(x_0) = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 3 \end{pmatrix} = -\begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

La fonction objective est quadratique ce qui donne

$$t_0 = \arg\min\{f(x_0 + td_0),\, t > 0) = -\frac{g(x_0)^\top d_0}{d_0^\top A d_0} = \frac{2}{3}$$

donc

$$x_1 = x_0 + t_0 d_0 = (0\,,\,0)^T$$

et on

$$\delta_0 = (-2\,,\,-2)^T$$
$$\gamma_0 = (-2\,,\,-4)^T,$$

alors

$$S_1 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

aussi

$$g(x_1) = (1\,,\,-1)^T$$

$$d_1 = -S_1 g(x_1) = (-1\,,\,\frac{1}{2})^T$$

et

$$t_1 = -\frac{g(x_1)^\top d_1}{d_1^\top A d_1} = 1.$$

En suite, on a

$$x_2 = x_1 + t_1 d_1 = (-1\,,\,\frac{1}{2})^T = x^*,$$

puisque $f$est une fonction à deux variables.

■

Exercice 4 *On veut minimiser la fonction donnée par*

$$f(x) = \frac{1}{2}x^T Ax - x^T b;\ \ x \in \mathbb{R}^2.$$

Où $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, $b = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ *et en utilisant la méthode DFP.*

3) *Calculer les trois premières itérations de l'algorithme DFP pour minimiser* f *sur* $\mathbb{R}^2$ *avec* $x_0 = (2,2)^T$ *et* $S_0 = I$*. Que peut on déduire.*

Solution: Si $\delta_k = (\delta_k^1, \delta_k^2)^T$ et $\gamma_k = (\gamma_k^1, \gamma_k^2)^T$, alors

$$S_{k+1} = S_k + \left[\begin{array}{cc} \frac{\left(\delta_k^1\right)^2}{\delta_k^1\gamma_k^1+\delta_k^2\gamma_k^2} - \frac{(\gamma_k^1)^2}{\left(\gamma_k^1\right)^2+\left(\gamma_k^2\right)^2} & \frac{\delta_k^1\delta_k^2}{\left(\gamma_k^1\right)^2+\left(\gamma_k^2\right)^2} - \frac{\gamma_k^1\gamma_k^2}{\left(\gamma_k^1\right)^2+\left(\gamma_k^2\right)^2} \\ \frac{\delta_k^1\delta_k^2}{\left(\gamma_k^1\right)^2+\left(\gamma_k^2\right)^2} - \frac{\gamma_k^1\gamma_k^2}{\left(\gamma_k^1\right)^2+\left(\gamma_k^2\right)^2} & \frac{\left(\delta_k^2-\gamma_k^2\right)^2}{\left(\delta_k^1-\gamma_k^1\right)\gamma_k^1+\left(\delta_k^2-\gamma_k^2\right)\gamma_k^2} \end{array}\right].$$

$$g(x) = Ax - b = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} x - \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Puisque $x_0 = (2\,,\,2)$ et $S_1 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, alors

$$g(x_0) = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

$$d_0 = -S_0 g(x_0) = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 3 \end{pmatrix} = -\begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

La fonction objective est quadratique ce qui donne

$$t_0 = \arg\min\{f(x_0 + td_0),\, t > 0) = -\frac{g(x_0)^\top d_0}{d_0^\top Ad_0} = \frac{2}{3}$$

donc

$$x_1 = x_0 + t_0 d_0 = (0\,,\,0)^T$$

et on a

$$\delta_0 = (-2\,,\,-2)^T$$
$$\gamma_0 = (-2\,,\,-4)^T,$$

alors

$$S_1 = \begin{pmatrix} \frac{17}{15} & -\frac{1}{15} \\ -\frac{1}{15} & \frac{8}{15} \end{pmatrix}$$

aussi

$$g(x_1) = (1\,,\,-1)^T$$

$$d_1 = -S_1 g(x_1) = (-\frac{6}{5}\,,\,\frac{3}{5})^T$$

et

$$t_1 = -\frac{g(x_1)^\top d_1}{d_1^\top A d_1} = \frac{5}{6}.$$

En suite, on a

$$x_2 = x_1 + t_1 d_1 = (-1\,,\,\frac{1}{2x^2x^{22}})^T = x^*,$$

et

$$\delta_1 = (-1\,,\,\frac{1}{2})^T$$

$$\gamma_0 = (-1\,,\,1)^T,$$

alors

$$S_2 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = A^{-1}$$

puisque $f$est une fonction à deux variables. $\boxed{1P^{ts}}$

■

Exercice 5 *Soit la fonction donnée par*

$$f(x) = \frac{1}{2}x^T A x - x^T b;\;\; x \in \mathbb{R}^2.$$

Où $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix},\;\; b = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

— *Calculer les quatre premières itérations de l'algorithme SR1 pour minimiser f sur $\mathbb{R}^2$ avec $x_0 = (-1, 1)^T$ et $S_0 = I$. Que peut on déduire ?*

Solution:

Pour faciliter les calcules, il est possible de calculer la mise à jour de la matrice S_k par la méthode SR1 :

$$S_{k+1} = S_k + \begin{bmatrix} \frac{\left(\delta_k^1 - \gamma_k^1\right)^2}{\left(\delta_k^1 - \gamma_k^1\right)\gamma_k^1 + \left(\delta_k^2 - \gamma_k^2\right)\gamma_k^2} & \frac{\left(\delta_k^1 - \gamma_k^1\right)\left(\delta_k^2 - \gamma_k^2\right)}{\left(\delta_k^1 - \gamma_k^1\right)\gamma_k^1 + \left(\delta_k^2 - \gamma_k^2\right)\gamma_k^2} \\ \frac{\left(\delta_k^1 - \gamma_k^1\right)\left(\delta_k^2 - \gamma_k^2\right)}{\left(\delta_k^1 - \gamma_k^1\right)\gamma_k^1 + \left(\delta_k^2 - \gamma_k^2\right)\gamma_k^2} & \frac{\left(\delta_k^2 - \gamma_k^2\right)^2}{\left(\delta_k^1 - \gamma_k^1\right)\gamma_k^1 + \left(\delta_k^2 - \gamma_k^2\right)\gamma_k^2} \end{bmatrix}.$$

— D'autre part, si f est une fonction quadratique convexe,

$$\theta(t) = f(x_k + td_k)$$

alors

$$\theta'(t_k) = g(x_k+td_k)^T d_k = \left((x_k + td_k)^\top A + b^\top\right) d_k = \left({x_k}^\top A + b^\top\right) d_k + td_k^T A d_k = g(x_k+td_k)^T d_k$$

t_k est un minimum local pour θ sur $]0\,,\,\infty[\ \Rightarrow \theta'(t_k) = g(x_k + td_k)^T d_k = 0 \Rightarrow t_k = -\dfrac{g(x_k)^\top d_k}{d_k^\top A d_k}$.

— On a

$$g(x) = Ax - b = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} x - \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Puisque $x_0 = (-1\,,\,1)$ et $S_1 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, alors

$$g(x_0) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

$$d_0 = -S_0 g(x_0) = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

La fonction objective est quadratique ce qui donne

$$t_0 = \arg\min\{f(x_0 + td_0),\, t > 0) = -\frac{g(x_0)^\top d_0}{d_0^\top A d_0} = \frac{5}{14}$$

donc

$$x_1 = x_0 + t_0 d_0 = (-\frac{2}{7}\,,\,\frac{9}{14})^T$$

et on

$$\delta_0 = (\frac{5}{7}\,,\,-\frac{5}{14})^T,$$

$$\gamma_0 = (\frac{15}{7}\,,\,-\frac{5}{7})^T,$$

alors

$$S_1 = \begin{pmatrix} \frac{5}{13} & \frac{2}{13} \\ \frac{2}{13} & \frac{25}{26} \end{pmatrix}$$

aussi

$$g(x_1) = (\frac{1}{7}\,,\,\frac{2}{7})^T$$

$$d_1 = -S_1 g(x_1) = (\frac{9}{91}\,,\,\frac{27}{91})^T$$

et

$$t_1 = -\frac{g(x_1)^\top d_1}{d_1^\top A d_1} = \frac{13}{27}.$$

En suite, on a

$$x_2 = x_1 + t_1 d_1 = (-\frac{1}{3}\,,\,\frac{1}{2})^T,$$

puisque $g(x_2) = (0,0)^T$ et f est une fonction strictement convexe alors $x^* = (-\frac{1}{3}\,,\,\frac{1}{2})^T$ est un minimum global stricte de plus $S_2 = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = A^{-1}$.

■