**Contrôle MOM (Avril 2011)**

Documents autorisés. Pas de calculette. 1 heure.

## Programmation géométrique

On considère le programme géométrique (PG) suivant :

(PG) $\min\limits_{x>0, y>0} g(x,y) = x + 2\dfrac{y}{x^2} + \dfrac{2}{y}$

1. Calculer le hessien de $g$ au point de coordonnées $x=1, y=2$. En déduire que $g$ n'est pas convexe.
2. Ecrire (DPG) le dual géométrique de (PG)
3. Résoudre (DPG)
4. En déduire le solution de (PG)
5. Vérifier que la solution trouvée est bien un point critique de $g$.

## Méthode du gradient conjugué

Soit $f(x_1,x_2) = x_1^2 - x_1x_2 + x_2^2 - 3x_2$ avec $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ dans $\mathbf{R}^2$.

On considère le problème : $\min\limits_{(x_1,x_2)\in R^2} f(x_1,x_2)$

Soit $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, la valeur de $f$ en $x^{(0)}$ est $f(x^{(0)}) = 0$.

1. Calculer $\nabla f$ le gradient de $f$.
2. En partant de $x^{(0)}$, appliquer la méthode du gradient conjugué (2 itérations). Pour calculer $\beta^{(k)}$ à l'itération $k$, on pourra utiliser $\beta^{(k)} = \dfrac{\|\nabla f(x^{(k+1)})\|^2}{\|\nabla f(x^{(k)})\|^2}$. Calculer la valeur de $f$ en chaque point parcouru. Vérifier que le point trouvé à la $2^{\text{ème}}$ itération est bien un point critique de $f$.

méthode du gradient conjugué

Soit $f(x_1, x_2) = x_1^2 - x_1 x_2 + x_2^2 - 3x_2$

on part de $x^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

1) calculer $\nabla f$

2) appliquer la méthode du gradient conjugué en partant de $x^{(0)}$

1) $\nabla f = \begin{bmatrix} 2x_1 - x_2 \\ 2x_2 - x_1 - 3 \end{bmatrix}$ $\boxed{2}$

2) $x = x^{(0)} - \lambda g^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \lambda \begin{bmatrix} 0 \\ 3 \end{bmatrix}$

minimisation sur $\lambda$ $\nabla f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \lambda \begin{bmatrix} 0 \\ 3 \end{bmatrix}\right) \cdot \begin{bmatrix} 0 \\ 3 \end{bmatrix} = 0 \Rightarrow \begin{pmatrix} 0 - 3\lambda \\ 6\lambda - 3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 3 \end{pmatrix} = 0 \Rightarrow \lambda = \frac{1}{2}$ $\boxed{2}$

$x^{(1)} = \begin{bmatrix} 0 \\ 3/2 \end{bmatrix}$ , $\nabla f(x^{(1)}) = \begin{bmatrix} -\frac{3}{2} \\ 0 \end{bmatrix}$ , $\beta^{(0)} = \dfrac{\|g^{(1)}\|}{\|g^{(0)}\|} = \dfrac{9/4}{9} = \frac{1}{4}$ $\boxed{2}$ , $f(x^{(1)}) = \frac{9}{4} - \frac{9}{2} = -\frac{9}{4}$

$d^{(1)} = \underbrace{\begin{bmatrix} 3/2 \\ 0 \end{bmatrix}}_{-g^{(1)}} + \frac{1}{4} \underbrace{\begin{bmatrix} 0 \\ 3 \end{bmatrix}}_{d^{(0)}} = \begin{bmatrix} 3/2 \\ 3/4 \end{bmatrix}$ $\boxed{1}$

itération k=1 $x = x^{(1)} + \lambda \begin{bmatrix} 3/2 \\ 3/4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3/2 \end{bmatrix} + \lambda \begin{bmatrix} 3/2 \\ 3/4 \end{bmatrix}$

minimisation sur $\lambda$ $\nabla f\left[\begin{pmatrix} 0 \\ 3/2 \end{pmatrix} + \lambda \begin{pmatrix} 3/2 \\ 3/4 \end{pmatrix}\right] \cdot \begin{pmatrix} 3/2 \\ 3/4 \end{pmatrix} = 0 \Rightarrow \begin{bmatrix} 3\lambda - \frac{3}{2} - \lambda\frac{3}{4} \\ 3 + \lambda\frac{3}{2} - \lambda\frac{3}{2} - 3 \end{bmatrix} \cdot \begin{bmatrix} 3/2 \\ 3/4 \end{bmatrix} = 0 \Rightarrow \lambda\frac{9}{4} = \frac{3}{2} \Rightarrow \lambda = \frac{2}{3}$ $\boxed{2}$

$x^{(2)} = \begin{bmatrix} 0 \\ 3/2 \end{bmatrix} + \frac{2}{3} \begin{bmatrix} 3/2 \\ 3/4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ , $f(x^{(2)}) = 1 - 2 + 4 - 6 = -3$

$\boxed{1}$

$\nabla f(x^{(2)}) = \begin{bmatrix} 2 - 2 \\ 4 - 1 - 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

DFP : méthode quasi-newtonienne de rang 2, approximation inverse du hessien.

Soit $f(x_1, x_2) = x_1^2 - x_1 x_2 + \frac{3}{2} x_2^2$

1) calculer les 2 premiers points de la suite fournie par la méthode DFP en partant de $x^{(0)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $S^{(0)} = I_2$ (identité)

2) calculer $S^{(2)}$ et vérifier que $S^{(2)} = H_f^{-1}$

on calcule le gradient dont on aura besoin tout le long : $\nabla f = \begin{bmatrix} 2x_1 - x_2 \\ -x_1 + 3x_2 \end{bmatrix}$

---

1) <u>itération k = 0.</u>

$x^{(0)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $f(x^{(0)}) = 5$, $\nabla f(x^{(0)}) = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$, $S^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\varphi(t) = f\Big(\underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix} - t \begin{bmatrix} 0 \\ 5 \end{bmatrix}}_{\substack{x_1 = 1 \\ x_2 = 2 - 5t}}\Big)$ , $\varphi'(t) = \nabla f \cdot \left(-\begin{bmatrix} 0 \\ 5 \end{bmatrix}\right) = \begin{bmatrix} 2 - 2 + 5t \\ -1 + 6 - 15t \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 5 \end{bmatrix} = -\begin{bmatrix} 5 & -15t \end{bmatrix} 5 = 0 \Rightarrow t = \frac{1}{3}$

$x^{(1)} = \begin{bmatrix} 1 \\ 1/3 \end{bmatrix}$ , $f(x^{(1)}) = \frac{5}{6}$ (ça a bien diminué) , $\nabla f(x^{(1)}) = \begin{bmatrix} 5/3 \\ 0 \end{bmatrix}$

$d^{(0)} = x^{(1)} - x^{(0)} = \begin{bmatrix} 0 \\ -5/3 \end{bmatrix}$ , $y^{(0)} = \nabla f(x^{(1)}) - \nabla f(x^{(0)}) = \begin{bmatrix} 5/3 \\ -5 \end{bmatrix}$

$d^{(0)} \otimes d^{(0)} = \begin{bmatrix} 0 \\ -5/3 \end{bmatrix} \begin{bmatrix} 0 & -5/3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{25}{9} \end{bmatrix}$ , $y^{(0)} \cdot d^{(0)} = \frac{25}{3}$

$S^{(0)} y^{(0)} \otimes S^{(0)} y^{(0)} = \begin{bmatrix} 5/3 \\ -5 \end{bmatrix} \begin{bmatrix} 5/3 & -5 \end{bmatrix} = \begin{bmatrix} \frac{25}{9} & -\frac{25}{3} \\ -\frac{25}{3} & 25 \end{bmatrix}$ , $S^{(0)} y^{(0)} \cdot y^{(0)} = \frac{25}{9} + 25 = \frac{25 + 225}{9} = \frac{250}{9}$

$S^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{3}{25} \begin{bmatrix} 0 & 0 \\ 0 & \frac{25}{9} \end{bmatrix} - \frac{9}{250} \begin{bmatrix} \frac{25}{9} & -\frac{25}{3} \\ -\frac{25}{3} & 25 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1/3 \end{bmatrix} - \begin{bmatrix} 1/10 & -3/10 \\ -3/10 & 9/10 \end{bmatrix} = \begin{bmatrix} 9/10 & 3/10 \\ 3/10 & 13/30 \end{bmatrix}$

vérifier que $S^{(1)}$ est déf. positif (th.) : $\Delta_1 = \frac{9}{10}$ , $\Delta_2 = \frac{9 \times 13}{10 \times 30} - \frac{9}{10 \times 10} = \frac{9 \times 13 - 3 \times 9}{10 \times 30} > 0$. c'est bon.

<u>itération k = 1</u>

$\varphi(t) = f\left(\begin{pmatrix} 1 \\ 1/3 \end{pmatrix} - t \begin{bmatrix} 9/10 & 3/10 \\ 3/10 & \frac{13}{30} \end{bmatrix} \begin{bmatrix} 5/3 \\ 0 \end{bmatrix}\right) = f\Big(\underbrace{\begin{pmatrix} 1 \\ 1/3 \end{pmatrix} - t \begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix}}_{\substack{x_1 = 1 - \frac{3}{2}t \\ x_2 = \frac{1}{3} - \frac{t}{2}}}\Big)$ , $\varphi'(t) = \begin{bmatrix} 2 - 3t - \frac{1}{3} + \frac{t}{2} \\ -1 + \frac{3}{2}t + 1 - \frac{3}{2}t \end{bmatrix} \cdot \left(-\begin{bmatrix} 3/2 \\ 1/2 \end{bmatrix}\right) = -\begin{bmatrix} 5/3 - \frac{5}{2}t \\ 0 \end{bmatrix} \begin{bmatrix} 3/2 \\ 1/2 \end{bmatrix} = 0$

$\Rightarrow t = \frac{2}{3}$

$x^{(2)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ , $f(x^{(2)}) = 0$ , $\nabla f(x^{(2)}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ c'est un point critique. (STOP)

2) $d^{(1)} = x^{(2)} - x^{(1)} = \begin{bmatrix} -1 \\ -1/3 \end{bmatrix}$ , $y^{(1)} = \nabla f(x^{(2)}) - \nabla f(x^{(1)}) = \begin{bmatrix} -5/3 \\ 0 \end{bmatrix}$

$d^{(1)} \otimes d^{(1)} = \begin{bmatrix} -1 \\ -1/3 \end{bmatrix} \begin{bmatrix} -1 & -1/3 \end{bmatrix} = \begin{bmatrix} 1 & 1/3 \\ 1/3 & 1/9 \end{bmatrix}$ , $y^{(1)} \cdot d^{(1)} = 5/3$

$S^{(1)} y^{(1)} \otimes S^{(1)} y^{(1)} = \begin{bmatrix} 9/10 & 3/10 \\ 3/10 & \frac{13}{30} \end{bmatrix} \begin{bmatrix} -5/3 \\ 0 \end{bmatrix} \otimes$ " $= \begin{bmatrix} -3/2 \\ -1/2 \end{bmatrix} \begin{bmatrix} -3/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 9/4 & 3/4 \\ 3/4 & 1/4 \end{bmatrix}$ , $S^{(1)} y^{(1)} \cdot y^{(1)} = \begin{bmatrix} -3/2 \\ -1/2 \end{bmatrix} \cdot \begin{bmatrix} -5/3 \\ 0 \end{bmatrix} = 5/2$

$S^{(2)} = \begin{bmatrix} 9/10 & 3/10 \\ 3/10 & \frac{13}{30} \end{bmatrix} + \frac{3}{5} \begin{bmatrix} 1 & 1/3 \\ 1/3 & 1/9 \end{bmatrix} - \frac{2}{5} \begin{bmatrix} 9/4 & 3/4 \\ 3/4 & 1/4 \end{bmatrix} = \begin{bmatrix} 9/10 + \frac{3}{5} - \frac{9}{10} & 3/10 + \frac{1}{5} - \frac{3}{10} \\ 3/10 + \frac{1}{5} - \frac{3}{10} & \frac{13}{30} + \frac{1}{15} - \frac{1}{10} \end{bmatrix} = \begin{bmatrix} 3/5 & 1/5 \\ 1/5 & \frac{6}{15} \end{bmatrix} = \begin{bmatrix} 3/5 & 1/5 \\ 1/5 & \frac{2}{5} \end{bmatrix}$

$H_f = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$ $S^{(2)} H_f = \begin{bmatrix} 3/5 & 1/5 \\ 1/5 & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

on retrouve th : si f quadratique strict. convexe c'est à dire $H_f$ déf. positive. alors en 2 itérations on a l'inverse du hessien.

Appliquer méthode gradient conjugué à

$$f = 4x_1^2 + 4x_2^2 - 4x_1x_2 - 12x_2 \quad \text{en partant du point } \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\nabla f = \begin{bmatrix} 8x_1 - 4x_2 \\ 8x_2 - 4x_1 - 12 \end{bmatrix}$$

<u>iter. 1</u> : $\nabla f(0) = \begin{bmatrix} -0 \\ -12 \end{bmatrix}$ $\qquad \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 12\lambda \end{bmatrix} \xrightarrow{\lambda = \frac{1}{8}} \begin{bmatrix} 0 \\ 3/2 \end{bmatrix}$

min $f$ dans direction $\begin{bmatrix} 0 \\ 12 \end{bmatrix}$ : $\nabla f(x + \lambda^* d) \cdot d = 0 \rightarrow \begin{bmatrix} -12\lambda \times 4 \\ 8 \times 12\lambda - 12 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 12 \end{bmatrix} = 0 \rightarrow \lambda = \frac{1}{8}$

$\nabla f \begin{bmatrix} 0 \\ 3/2 \end{bmatrix} = \begin{bmatrix} -6 \\ 0 \end{bmatrix} \qquad \beta = \frac{36}{12 \times 12} = \frac{1}{4}$

<u>iter 2</u> $\qquad d = \begin{bmatrix} 6 \\ 0 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} 0 \\ 12 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$ $\qquad$ vérif directions conjuguées : $H_f = \begin{bmatrix} 8 & -4 \\ -4 & 8 \end{bmatrix}$ $\qquad (0 \;\; 12) \begin{bmatrix} 8 & -4 \\ -4 & 8 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 12 \, [-4 \;\; 8] \begin{bmatrix} 6 \\ 3 \end{bmatrix}$

min $f$ dans direction $\begin{bmatrix} 6 \\ 3 \end{bmatrix}$ : $\begin{bmatrix} 8 \times 6\lambda - 4 \times \frac{3}{2} + -4 \times 3\lambda \\ 8 \times \frac{3}{2} + 8 \times 3\lambda - 4 \times 6\lambda - 12 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \times 4 \left[ \left[ 9\lambda - \frac{3}{2} \right] \times 2 \right] \rightarrow \lambda = \frac{1}{6}$

$$\begin{bmatrix} 0 \\ 3/2 \end{bmatrix} + \lambda \begin{bmatrix} 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 6\lambda \\ \frac{3}{2} + 3\lambda \end{bmatrix}$$

$$\xrightarrow{\lambda = 1/6} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

vérif optimum : $\nabla f \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Sujet A

Directions conjuguées

Soit $f(x)=\frac{1}{2}x^t Ax+b^t x$ avec $A=\begin{pmatrix}4 & -1\\ -1 & 2\end{pmatrix}$, $b=\begin{pmatrix}1\\2\end{pmatrix}$ et $x=\begin{pmatrix}x_1\\x_2\end{pmatrix}\in R^2$.

La transposée d'un vecteur $u=\begin{pmatrix}u_1\\u_2\end{pmatrix}$ est notée $u^t=\begin{pmatrix}u_1 & u_2\end{pmatrix}$.

On veut résoudre le problème :

(P) minimiser $f(x)$ s.c. $x\in R^2$

**1)** On considère le point $x^{(0)}=\begin{pmatrix}0\\0\end{pmatrix}$.

Construire $x^{(1)}=x^{(0)}+td^{(0)}$ avec $t\in R$ qui minimise $f\left(x^{(0)}+td^{(0)}\right)$ où $d^{(0)}=-\begin{pmatrix}1\\2\end{pmatrix}$ est la direction opposée au gradient de $f$ au point $x^{(0)}$.

**2)** On va ici construire la direction suivante.

a) Donner la relation que doit vérifier $d^{(1)}=\begin{pmatrix}d_1\\d_2\end{pmatrix}$ pour être conjuguée à $d^{(0)}$ par rapport à $A$.

b) Donner la condition que doit vérifier $d^{(1)}=\begin{pmatrix}d_1\\d_2\end{pmatrix}$ pour être une direction de descente de $f$ au point $x^{(1)}$.

Tenant compte de la relation trouvée à la question a), en déduire que $d_2$ doit être négatif.

c) Construire $d^{(1)}=\begin{pmatrix}d_1\\d_2\end{pmatrix}$ conjuguée à $d^{(0)}$ par rapport à $A$ et telle que $d_2=-2$.

**3)** Construire $x^{(2)}=x^{(1)}+td^{(1)}$ avec $t\in R$ qui minimise $f\left(x^{(1)}+td^{(1)}\right)$.

Est-il nécessaire de faire une itération supplémentaire ? Vérifier l'optimalité de $x^{(2)}$ pour le problème (P).

## Directions conjuguées

(mai 2007 sujet A)

$$\min f(x) = \frac{1}{2} x^t \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot x \qquad x \in \mathbb{R}^2$$

$$\nabla f = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

1) En partant de $x^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, construire $x^{(1)}$ qui minimise $f$ dans la direction opposée au gradient $d^{(0)} = -\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

2) a) donner la relation que doit vérifier $d^{(1)}$ pour être conjuguée à $d^{(0)}$ par rapport à $A = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$.

b) " " " " " " " " " une direction de descente de $f$ au point $x^{(1)}$

En déduire que $d^{(1)} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$ est t.q. $d_2 < 0$.

c) Construire $d^{(1)} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$ avec $d_2 = -2$

3) En partant de $x^{(1)}$, construire $x^{(2)}$ qui minimise $f$ dans la direction $d^{(1)}$

Est-il nécessaire de continuer l'algo. ? Vérifier l'optimalité de $x^{(2)}$.

---

1) $x^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\quad d^{(0)} = -\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $\quad x^{(1)} = x^{(0)} + t\, d^{(0)} = -\begin{bmatrix} t \\ 2t \end{bmatrix}$

Pour trouver $t$, on annule la dérivée directionnelle : $\nabla f(x^{(1)}) \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0 \Leftrightarrow \left[ -t \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right] \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0$

$$\left[ -t \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right] \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0 \Leftrightarrow (1 - 2t) + 2(2 - 3t) = 0$$

$$\Leftrightarrow \boxed{t = 5/8}$$

$$\Rightarrow x^{(1)} = -\begin{bmatrix} 5/8 \\ 5/4 \end{bmatrix}$$

2) a) $\underbrace{\begin{pmatrix} 1 & 2 \end{pmatrix}}_{d^{(0)}} \underbrace{\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}}_{A} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0 \Leftrightarrow \begin{pmatrix} 2 & 3 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = 0 \Leftrightarrow \boxed{2d_1 + 3d_2 = 0}$

b) $g^{(1)} = -\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 5/8 \\ 5/4 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1/8 \end{bmatrix}$ , $\quad \underset{\uparrow}{g^{(1)} \cdot d} < 0 \Leftrightarrow \begin{bmatrix} -1/4 \\ 1/8 \end{bmatrix} \cdot \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} < 0 \Leftrightarrow \boxed{\frac{1}{4}\left[ -d_1 + \frac{1}{2} d_2 \right] < 0}$

dérivée directionnelle

$$\begin{cases} -d_1 + \frac{1}{2} d_2 < 0 \\ 2d_1 + 3d_2 = 0 \end{cases} \Rightarrow 4d_2 < 0$$

c) $d_2 = -2 \Rightarrow 2d_1 - 6 = 0 \Rightarrow \boxed{d_1 = 3} \quad \Rightarrow d = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$

3) $x^{(2)} = x^{(1)} + t\, d^{(1)} = -\begin{bmatrix} 5/8 \\ 5/4 \end{bmatrix} + t \begin{bmatrix} 3 \\ -2 \end{bmatrix}$

Pour trouver $t$, on annule la dérivée directionnelle : $\nabla f(x^{(2)}) \cdot d = 0 \Leftrightarrow \begin{bmatrix} 4(-5/8 + 3t) - [-5/4 - 2t] + 1 \\ -[-5/8 + 3t] + 2[-5/4 - 2t] + 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -2 \end{bmatrix} = 0$

$$\begin{bmatrix} \overbrace{-\frac{1}{4}}^{g^{(1)}} + 14t \\ \frac{1}{8} - 7t \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -2 \end{bmatrix} = 0 \Leftrightarrow -\frac{3}{4} + 3 \times 14t - \frac{2}{8} + 14t = 0 \Leftrightarrow -1 + 4 \times 14t = 0 \Leftrightarrow \boxed{t = \frac{1}{4 \times 14} = \frac{1}{56}}$$

$$x^{(2)} = \begin{bmatrix} -5/8 + 3/56 \\ -5/4 - 2/56 \end{bmatrix} = \begin{bmatrix} -\frac{35}{56} + \frac{3}{56} \\ -\frac{70}{56} - \frac{2}{56} \end{bmatrix} = \begin{bmatrix} -\frac{32}{56} \\ -\frac{72}{56} \end{bmatrix} = \begin{bmatrix} -\frac{4}{7} \\ -\frac{9}{7} \end{bmatrix}$$

vérifions que $x^{(2)}$ est un point critique :

$$4 \times -\frac{4}{7} + \frac{9}{7} + 1 = \frac{-16 + 9 + 7}{7} = 0$$

$$\frac{4}{7} - \frac{18}{7} + 2 = \frac{4 - 18 + 14}{7} = 0$$

# Méthode de Cauchy, méthode du gradient réduit.

Soit $f(x_1,x_2) = x_1^2 + x_1x_2 + \frac{x_2^2}{2} - 3x_1 - 2x_2 = \frac{1}{2}x^t\begin{pmatrix}2&1\\1&1\end{pmatrix}x + \begin{pmatrix}-3\\-2\end{pmatrix}^t x$ avec $x = \begin{pmatrix}x_1\\x_2\end{pmatrix}$

On considère le pb : minimiser $f(x_1,x_2)$ avec $\begin{pmatrix}x_1\\x_2\end{pmatrix}\in\mathbb{R}^2$

Soit $x^{(0)} = \begin{pmatrix}0\\3\end{pmatrix}$, on a $f(x^{(0)}) = -\frac{3}{2}$ (la valeur de $f$ en $x^{(0)}$)

1) Calculer $\nabla f$ le gradient de $f$

2) En partant de $x^{(0)} = \begin{pmatrix}0\\3\end{pmatrix}$, appliquer la méthode du gradient conjugué (2 itérations)
Calculer les valeurs de $f$ en chaque point parcouru.
Vérifier que le point trouvé à la 2ᵉ itération est bien un point critique de $f$

3) En partant de $x^{(0)}$, appliquer les 2 premières itérations de la méthode de Cauchy (plus forte pente).
En remarquant que la première itération est la même que précédemment, on fera seulement la 2ᵉ itération
Calculer la valeur de $f$ en le point obtenu.

---

1) $\nabla f = \begin{pmatrix}2&1\\1&1\end{pmatrix}x + \begin{pmatrix}-3\\-2\end{pmatrix} = \begin{pmatrix}2x_1 + x_2 - 3\\ x_1 + x_2 - 2\end{pmatrix}$

2) $\nabla f\begin{pmatrix}0\\3\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix}$, $d^{(0)} = -\begin{pmatrix}0\\1\end{pmatrix}$

itération k=0 $f(t) = f\Big(\underbrace{\begin{pmatrix}0\\3\end{pmatrix} - t\begin{pmatrix}0\\1\end{pmatrix}}_{\begin{pmatrix}0\\3-t\end{pmatrix}}\Big)$, $f'(t) = \nabla f\begin{pmatrix}0\\3-t\end{pmatrix}\cdot\left(-\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}x\\1-t\end{pmatrix}\cdot\left(-\begin{pmatrix}0\\1\end{pmatrix}\right) = 0 \Rightarrow \underline{t=1}$

$x^{(1)} = \begin{pmatrix}0\\2\end{pmatrix}$, $\nabla f(x^{(1)}) = \begin{pmatrix}-1\\0\end{pmatrix}$, $f(x^{(1)}) = \frac{4}{2} - 4 = -\frac{4}{2}$

$\beta^{(0)} = \frac{\|\nabla f(x^{(1)})\|^2}{\|\nabla f(x^{(0)})\|^2} = \frac{1}{1} = 1 \Rightarrow d^{(1)} = \beta^{(0)} d^{(0)} - \nabla f(x^{(1)}) = -\begin{pmatrix}0\\1\end{pmatrix} - \begin{pmatrix}-1\\0\end{pmatrix} = \begin{pmatrix}1\\-1\end{pmatrix}$

vérifions que $d^{(1)}\cdot Q\, d^{(0)} = (1\ \ -1)\begin{pmatrix}2&1\\1&1\end{pmatrix}\begin{pmatrix}0\\-1\end{pmatrix} = (1\ \ -1)\begin{pmatrix}-1\\-1\end{pmatrix} = 0$

itération k=1 $f(t) = f\Big(\underbrace{\begin{pmatrix}0\\2\end{pmatrix} + t\begin{pmatrix}1\\-1\end{pmatrix}}_{\begin{pmatrix}t\\2-t\end{pmatrix}}\Big)$, $f'(t) = \nabla f\begin{pmatrix}t\\2-t\end{pmatrix}\cdot\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}2t+2-t-3\\t+2-t-2\end{pmatrix}\cdot\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}t-1\\0\end{pmatrix}\cdot\begin{pmatrix}1\\-1\end{pmatrix} = t-1 = 0 \Rightarrow \underline{t=1}$

$x^{(2)} = \begin{pmatrix}1\\1\end{pmatrix}$ on a bien $\nabla f(x^{(2)}) = 0$, $f\begin{pmatrix}1\\1\end{pmatrix} = 1 + 1 + \frac{1}{2} - 5 = -\frac{6}{2} + \frac{1}{2} = -\frac{5}{2}$

3) 2ᵉ itération de la méthode de Cauchy. $k=1$ : $d^{(1)} = -\nabla f(x^{(1)})$

$f(t) = f\Big(\underbrace{\begin{pmatrix}0\\2\end{pmatrix} - t\begin{pmatrix}-1\\0\end{pmatrix}}_{\begin{pmatrix}t\\2\end{pmatrix}}\Big)$, $f'(t) = \nabla f\begin{pmatrix}t\\2\end{pmatrix}\cdot\left(-\begin{pmatrix}-1\\0\end{pmatrix}\right) = \begin{pmatrix}2t+2-3\\x\end{pmatrix}\cdot\left(-\begin{pmatrix}-1\\0\end{pmatrix}\right) = 2t-1 = 0 \Rightarrow \underline{t=\frac{1}{2}}$

$x^{(2)} = \begin{pmatrix}1/2\\2\end{pmatrix}$ ; $f\begin{pmatrix}1/2\\2\end{pmatrix} = \frac{1}{4} + 1 + 2 - \frac{3}{2} - 4 = \frac{1}{4} - \frac{6}{4} - \frac{4}{4} = -\frac{9}{4} = -\frac{4,5}{2}$

**Contrôle MOM**
Septembre 2008

Documents autorisés. Calculettes interdites.

## Problème n°1 : Méthode des directions conjuguées

Notations :

- Etant donné $v=\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ un vecteur colonne de $R^2$, la transposée de $v$ est notée $v^t$ et $v^t=\begin{pmatrix} v_1 & v_2 \end{pmatrix}$.

- Etant donnés $v=\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, u=\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ deux vecteurs colonne de $R^2$, $v^t u = v \bullet u = v_1 u_1 + v_2 u_2$.

Soit la fonction $f(x)=\frac{1}{2}x^t Q x + b^t x$ avec $x=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in R^2$, $Q=\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}, b=\begin{pmatrix} -4 \\ -3 \end{pmatrix}$.

1) Montrer que $d^{(1)}=\begin{pmatrix} 1 \\ -1 \end{pmatrix}, d^{(2)}=\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ sont conjuguées par rapport à $Q$.

2) On exprime $x$ dans la base $d^{(1)}$, $d^{(2)}$ en fonction de $y_1$, $y_2$ : $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=y_1\begin{pmatrix} 1 \\ -1 \end{pmatrix}+y_2\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, et on note $g(y)=f\left(y_1 d^{(1)}+y_2 d^{(2)}\right)$ avec $y=\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$.

Montrer que $g(y)=g_1(y_1)+g_2(y_2)$ avec $g_1(y_1)=\frac{3}{2}y_1^2-y_1$, $g_2(y_2)=\frac{15}{2}y_2^2-10y_2$.

3) Le problème $\min\limits_{y} g(y)$ se décompose en deux problèmes indépendants $\min\limits_{y_1} g_1(y_1)$ et $\min\limits_{y_2} g_2(y_2)$. Résoudre ces deux problèmes et en déduire la solution du problème $\min\limits_{x} f(x)$. Vérifier que la solution est bien un point critique de $f$.

directions conjuguées. 8

$f(x) = \frac{1}{2}x^tQx + b^tx$ avec $Q = \begin{bmatrix}3&1\\1&2\end{bmatrix}$, $b = \begin{bmatrix}-4\\-3\end{bmatrix}$

1) $d^{(1)} = \begin{bmatrix}1\\-1\end{bmatrix}$, $d^{(2)} = \begin{bmatrix}1\\2\end{bmatrix}$

$d^{(1)}\cdot Q d^{(2)} = (1\ \ -1)\begin{bmatrix}3&1\\1&2\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix} = (1\ \ -1)\begin{bmatrix}5\\5\end{bmatrix} = 5-5 = 0$

1 $= (2\ \ -1)\begin{bmatrix}1\\2\end{bmatrix} = 2-2 = 0$

2) $x = y_1 d^{(1)} + y_2 d^{(2)} \Rightarrow f(x) = \frac{1}{2}\left[y_1d^{(1)} + y_2d^{(2)}\right]^t Q \left[y_1d^{(1)} + y_2d^{(2)}\right] + b^t\left[y_1d^{(1)} + y_2d^{(2)}\right]$

$= \frac{1}{2}y_1^2(1\ \ -1)\begin{bmatrix}3&1\\1&2\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} + \frac{1}{2}y_2^2(1\ \ 2)\begin{bmatrix}3&1\\1&2\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix} + y_1[-4\ \ -3]\begin{bmatrix}1\\-1\end{bmatrix} + y_2[-4\ \ -3]\begin{bmatrix}1\\2\end{bmatrix}$

[ les termes $d^{(1)}\cdot Q d^{(2)} = 0$

$d^{(2)}\cdot Q d^{(1)} = 0$

d'après 1) ]

3

$= \frac{1}{2}y_1^2(1\ \ -1)\begin{bmatrix}2\\-1\end{bmatrix} + \frac{1}{2}y_2^2[1\ \ 2]\begin{bmatrix}5\\5\end{bmatrix} + y_1(-4+3) + y_2(-4-6)$

$= \frac{1}{2}y_1^2\, 3 + \frac{1}{2}y_2^2\, 15 + y_1(-1) + y_2(-10)$

$= \underbrace{\frac{3}{2}y_1^2 - y_1}_{g_1(y_1)} + \underbrace{\frac{15}{2}y_2^2 - 10y_2}_{g_2(y_2)}$

3) $\min_x f(x) \Leftrightarrow \min_y g(y) = \min_{y_1} g_1(y_1) + \min_{y_2} g_2(y_2)$

2 $g_1(y_1) \to g_1'(y_1) = 3y_1 - 1 = 0 \to y_1 = 1/3$ $\quad \to g_1(y_1) = \frac{3}{2}\times\left(\frac{1}{3}\right)^2 - \frac{1}{3} = \frac{1}{6} - \frac{1}{3} = -\frac{1}{6}$

$g_2(y_2) \to g_2'(y_2) = 15y_2 - 10 = 0 \to y_2 = \frac{10}{15} = \frac{2}{3}$ $\quad \to g_2(y_2) = \frac{15}{2}\times\left(\frac{2}{3}\right)^2 - 10\times\frac{2}{3} = \frac{5\times 2}{3} - 10\times\frac{2}{3} = \frac{2}{3}(-5)$

$\Big\} -\frac{1}{6} - \frac{20}{6} = -\frac{21}{6} = -\frac{7}{2}$

1 $x = \frac{1}{3}\begin{bmatrix}1\\-1\end{bmatrix} + \frac{2}{3}\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix}$ $\quad \to f(x) = \frac{1}{2}[1\ \ 1]\begin{bmatrix}3&1\\1&2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} + [-4\ \ -3]\begin{bmatrix}1\\1\end{bmatrix} = \frac{1}{2}[1\ \ 1]\begin{bmatrix}4\\3\end{bmatrix} + (-4-3) = \frac{1}{2}\times 7 - 7 = -\frac{7}{2}$

on retrouve bien $g_1(y_1) + g_2(y_2)$

1 $\nabla f(x) = Qx + b = \begin{bmatrix}3&1\\1&2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}-4\\-3\end{bmatrix} = \begin{bmatrix}4\\3\end{bmatrix} + \begin{bmatrix}-4\\-3\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$ $\quad x = \begin{bmatrix}1\\1\end{bmatrix}$ est bien un point critique de $f$

donc c'est le minimiseur de $f$

**Contrôle Optimisation Mai 2010 – groupes 1 et 2**

NOM :

Calculettes interdites – Documents autorisés – 1h - Répondre sur la feuille.

Soit $f(x) = x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 - 4x_1 - 3x_2$ définie sur $\mathbb{R}^2$.

I Généralités

1) Mettre $f(x)$ sous la forme $f(x) = \frac{1}{2}x \bullet Ax - b \bullet x$ avec $A$ symétrique et donner l'expression de $A$, $b$. (1 point)

2) Donner l'expression du gradient de $f$, $\nabla f$, en fonction de $A$, $b$. Donner l'expression du hessien de $f$, $H_f$, en fonction de $A$. (1 point)

3) $f$ est-elle convexe ? (1 point)

4) Donner le point $x^*$ minimisant $f$. (2 points)

II Algorithme du gradient conjugué

On va maintenant appliquer la méthode du gradient conjugué à $f$ en partant du point $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

1) Calculer $d^{(0)}$, $\lambda^{(0)}$, $x^{(1)}$. Donner $\lambda^{(0)}$ sous forme d'une fraction irréductible. (5 points)

2) Calculer $\beta^{(0)}$ et $d^{(1)}$. On pourra utiliser $\beta^{(0)} = \frac{\|g^{(1)}\|^2}{\|g^{(0)}\|^2}$. Vérifier que $d^{(1)}$ est conjuguée à $d^{(0)}$ par rapport à $A$. (5 points)

3) Calculer $\lambda^{(1)}$, $x^{(2)}$. Donner $\lambda^{(1)}$ sous forme d'une fraction irréductible. Quel point retrouve-t-on ? (5 points)

**Contrôle Optimisation Mai 2010 – groupes 3 et 4**

NOM :

Calculettes interdites – Documents autorisés – 1h - Répondre sur la feuille.

Soit $f(x) = x_1^2 + x_1x_2 + x_2^2 - x_1 - 2x_2$ définie sur $\mathbb{R}^2$.

I Généralités

1) Mettre $f(x)$ sous la forme $f(x) = \frac{1}{2}x \bullet Ax - b \bullet x$ avec $A$ symétrique et donner l'expression de $A$, $b$. (1 point)

2) Donner l'expression du gradient de $f$, $\nabla f$, en fonction de $A$, $b$. Donner l'expression du hessien de $f$, $H_f$, en fonction de $A$. (1 point)

3) $f$ est-elle convexe ? (1 point)

4) Donner le point $x^*$ minimisant $f$. (2 points)

II Algorithme du gradient conjugué

On va maintenant appliquer la méthode du gradient conjugué à $f$ en partant du point $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

1) Calculer $d^{(0)}$, $\lambda^{(0)}$, $x^{(1)}$. Donner $\lambda^{(0)}$ sous forme d'une fraction irréductible. (5 points)

2) Calculer $\beta^{(0)}$ et $d^{(1)}$. On pourra utiliser $\beta^{(0)} = \frac{\|g^{(1)}\|^2}{\|g^{(0)}\|^2}$. Vérifier que $d^{(1)}$ est conjuguée à $d^{(0)}$ par rapport à $A$. (5 points)

3) Calculer $\lambda^{(1)}$, $x^{(2)}$. Donner $\lambda^{(1)}$ sous forme d'une fraction irréductible. Quel point retrouve-t-on ? (5 points)

**Contrôle Optimisation Mai 2010 – groupe Strasbourg**

NOM :

Calculettes interdites – Documents autorisés – 1h - Répondre sur la feuille.

Soit $f(x) = \frac{1}{2}x_1^2 + x_1x_2 + x_2^2 - 3x_1 - 4x_2$ définie sur $\mathbb{R}^2$.

I <u>Généralités</u>

1) Mettre $f(x)$ sous la forme $f(x) = \frac{1}{2}x \bullet Ax - b \bullet x$ avec $A$ symétrique et donner l'expression de $A$, $b$. (1 point)

2) Donner l'expression du gradient de $f$, $\nabla f$, en fonction de $A$, $b$. Donner l'expression du hessien de $f$, $H_f$, en fonction de $A$. (1 point)

3) $f$ est-elle convexe ? (1 point)

4) Donner le point $x^*$ minimisant $f$. (2 points)

II <u>Algorithme du gradient conjugué</u>

On va maintenant appliquer la méthode du gradient conjugué à $f$ en partant du point $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

1) Calculer $d^{(0)}$, $\lambda^{(0)}$, $x^{(1)}$. Donner $\lambda^{(0)}$ sous forme d'une fraction irréductible. (5 points)

2) Calculer $\beta^{(0)}$ et $d^{(1)}$. On pourra utiliser $\beta^{(0)} = \frac{\|g^{(1)}\|^2}{\|g^{(0)}\|^2}$. Vérifier que $d^{(1)}$ est conjuguée à $d^{(0)}$ par rapport à $A$. (5 points)

3) Calculer $\lambda^{(1)}$, $x^{(2)}$. Donner $\lambda^{(1)}$ sous forme d'une fraction irréductible. Quel point retrouve-t-on ? (5 points)

Correction mai 2010

qu. 1 et 2. $f(x) = x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 - 4x_1 - 3x_2$

$f(x) = \frac{1}{2}\, x \cdot A x - b \cdot x \qquad A = H_f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \qquad b = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ avec A symétrique. A déf positive : $\Delta_1 = 2 > 0$, $\Delta_2 = 1 > 0$

$\Rightarrow f$ convexe (même strictement)

$$\nabla f(x) = Ax - b = \begin{bmatrix} 2x_1 + x_2 - 4 \\ x_1 + x_2 - 3 \end{bmatrix}$$

$$\nabla f(x^*) = 0 \rightarrow \begin{cases} 2x_1 + x_2 = 4 \\ x_1 + x_2 = 3 \end{cases} \rightarrow \begin{cases} x_1 = 1 \\ x_2 = 2 \end{cases}$$

gradient conjugué en partant de $x^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$g^{(0)} = \begin{bmatrix} -4 \\ -3 \end{bmatrix}, \quad d^{(0)} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$

détermination de $\lambda^{(0)}$ : on minimise $f(\lambda d^{(0)})$ $\rightarrow$ dérivée directionnelle $= 0$ $\rightarrow \nabla f\begin{bmatrix} 4\lambda \\ 3\lambda \end{bmatrix} \cdot d^{(0)} = \begin{bmatrix} 8\lambda + 3\lambda - 4 \\ 4\lambda + 3\lambda - 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 3 \end{bmatrix} = 44\lambda + 21\lambda - 25 = 0 \rightarrow \boxed{\lambda^{(0)} = \frac{25}{65} = \frac{5}{13}}$

$x^{(1)} = \begin{bmatrix} \frac{20}{13} \\ \frac{15}{13} \end{bmatrix}, \quad g^{(1)} = \begin{bmatrix} \frac{40}{13} + \frac{15}{13} - 4 \\ \frac{20}{13} + \frac{15}{13} - 3 \end{bmatrix} = \begin{bmatrix} 3/13 \\ -4/13 \end{bmatrix}, \quad \beta^{(0)} = \frac{g^{(1)} \cdot Ad^{(0)}}{d^{(0)} \cdot Ad^{(0)}} = \frac{\|g^{(1)}\|^2}{\|g^{(0)}\|^2} = \boxed{\frac{1}{13^2}}, \quad d^{(1)} = \begin{bmatrix} -3/13 \\ +4/13 \end{bmatrix} + \frac{1}{13^2}\begin{bmatrix} 4 \\ 3 \end{bmatrix} = \frac{1}{13^2}\begin{bmatrix} -39 + 4 \\ 52 + 3 \end{bmatrix} = \frac{1}{13^2}\begin{bmatrix} -35 \\ 55 \end{bmatrix} = \boxed{\frac{5}{13^2}\begin{bmatrix} -7 \\ 11 \end{bmatrix}}$

$$d^{(1)} \cdot Ad^{(0)} = \frac{5}{13^2}\begin{bmatrix} -7 & 11 \end{bmatrix}\begin{bmatrix} 11 \\ 7 \end{bmatrix} = \frac{5}{13^2}(-7 \times 11 + 11 \times 7) = 0$$

détermination de $\lambda^{(1)}$ : on minimise $f(x^{(1)} + \lambda d^{(1)})$ $\rightarrow$ dérivée directionnelle $= 0$ $\rightarrow \nabla f\left(\frac{5}{13}\begin{bmatrix} 4 - \frac{7\lambda}{13} \\ 3 + \frac{11\lambda}{13} \end{bmatrix}\right) \cdot d^{(1)} = \cancel{\nabla f\left(\frac{5}{13}\begin{bmatrix} \frac{45}{13} \\ \frac{14}{13} \end{bmatrix}\right) \cdot d^{(1)} = \frac{16}{13}\,\frac{45}{13} + \frac{5 \times 11}{13}/13}$

$$\begin{bmatrix} \frac{10}{13}\left(4 - \frac{7\lambda}{13}\right) + \frac{5}{13}\left(3 + \frac{11\lambda}{13}\right) - 4 \\ \frac{5}{13}\left(3 + \frac{11\lambda}{13} + 4 - \frac{7\lambda}{13}\right) - 3 \end{bmatrix} \cdot \frac{5}{13^2}\begin{bmatrix} -7 \\ 11 \end{bmatrix} = \begin{bmatrix} \frac{3}{13} - \frac{15\lambda}{13^2} \\ -\frac{4}{13} + \frac{20\lambda}{13^2} \end{bmatrix} \cdot \begin{bmatrix} -7 \\ 11 \end{bmatrix}\frac{5}{13^2} = 0$$

$$\Rightarrow -65 + \frac{325\lambda}{13} = 0 \Rightarrow \boxed{\lambda^{(1)} = \frac{65 \times 13}{325} = \frac{13}{5}}$$

$x^{(2)} = \frac{5}{13}\begin{bmatrix} 4 - \frac{7}{5} \\ 3 + \frac{11}{5} \end{bmatrix} = \frac{5}{13}\begin{bmatrix} \frac{13}{5} \\ \frac{26}{5} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ on retrouve bien $x^*$ le point critique de $f$

# Corrigé mai 2010

**gr. 3 et 4** : $f(x) = x_1^2 + x_1x_2 + x_2^2 - x_1 - 2x_2$

$f(x) = \frac{1}{2} x \cdot Ax - b \cdot x$ avec $A$ symétrique. $A = H_f = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $b = \begin{bmatrix} +1 \\ 2 \end{bmatrix}$, $A$ déf positive : $\Delta_1 = 2 > 0$, $\Delta_2 = 3 > 0$

$\Rightarrow f$ convexe (même strictement)

$$\nabla f = Ax - b = \begin{bmatrix} 2x_1 + x_2 - 1 \\ x_1 + 2x_2 - 2 \end{bmatrix}$$

$$\nabla f(x^*) = 0 \to \begin{cases} 2x_1 + x_2 = 1 \\ x_1 + 2x_2 = 2 \end{cases} \to \begin{cases} x_1 = 0 \\ x_2 = 1 \end{cases}$$

méthode du gradient conjugué en partant de $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$g^{(0)} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, \quad d^{(0)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

déterminer $\rho^{(0)} \to$ minimiser $f(x^{(0)} + \rho d^{(0)}) = f\begin{bmatrix} \rho \\ 2\rho \end{bmatrix} \to$ dérivée directionnelle $= 0 \to \nabla f \begin{bmatrix} \rho \\ 2\rho \end{bmatrix} \cdot d^{(0)} = \begin{bmatrix} 2\rho + 2\rho - 1 \\ \rho + 4\rho - 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 14\rho - 5 = 0 \Rightarrow \boxed{\rho^{(0)} = \frac{5}{14}}$

$$x^{(1)} = \frac{5}{14}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$g^{(1)} = \begin{bmatrix} \frac{10}{14} + \frac{10}{14} - 1 \\ \frac{5}{14} + \frac{20}{14} - 2 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 6 \\ -3 \end{bmatrix}, \quad d^{(1)} = -g^{(1)} + \beta^{(0)} d^{(0)}, \quad \beta^{(0)} = \frac{g^{(1)} \cdot Ad^{(0)}}{d^{(0)} \cdot Ad^{(0)}} = \frac{\|g^{(1)}\|^2}{\|g^{(0)}\|^2} = \boxed{\frac{9}{14^2}}$$

$$d^{(1)} = \frac{1}{14}\begin{bmatrix} -6 \\ 3 \end{bmatrix} + \frac{9}{14^2}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{3}{14}\begin{bmatrix} -2 + 3/14 \\ 1 + 6/14 \end{bmatrix} = \frac{3}{14^2}\begin{bmatrix} -25 \\ 20 \end{bmatrix} = \underline{\frac{15}{14^2}\begin{bmatrix} -5 \\ 4 \end{bmatrix}}$$

$$d^{(1)} \cdot Ad^{(0)} = \frac{15}{14^2}\begin{bmatrix} -5 & 4 \end{bmatrix}\begin{bmatrix} 4 \\ 5 \end{bmatrix} = \frac{15}{14^2}\left[-5 \times 4 + 4 \times 5\right] = 0$$

déterminer $\rho^{(1)} \to$ minimiser $f(x^{(1)} + \rho d^{(1)}) = f\left(\frac{5}{14}\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \rho \frac{15}{14^2}\begin{bmatrix} -5 \\ 4 \end{bmatrix}\right) \to$ dérivée directionnelle $= 0 \to \nabla f\left[\frac{5}{14}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \rho \frac{3}{14}\begin{bmatrix} -5 \\ 4 \end{bmatrix}\right] \cdot d^{(1)} = 0$

$$\begin{bmatrix} \frac{10}{14} + \rho \frac{6}{14} \times (-5) \frac{5}{14} + \frac{10}{14} + \frac{5}{14} \rho \frac{3}{14} \times 4 - 1 \\ \frac{5}{14} + \rho \frac{3}{14}(-5)\frac{5}{14} + \frac{20}{14} + \frac{5}{14}\frac{3}{14} \rho \times 8 - 2 \end{bmatrix} \cdot \frac{15}{14^2}\begin{bmatrix} -5 \\ 4 \end{bmatrix} = 0$$

$$-100 + 70 + 100 - 112 + \frac{5}{14} 90\rho + \frac{5}{14} 36\rho = 0 \Rightarrow \boxed{\rho^{(1)} = \frac{42}{126} \times \frac{14}{5} = \frac{14}{15}}$$

$$x^{(2)} = \frac{5}{14}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{14}{15} \times \frac{15}{14^2}\begin{pmatrix} -5 \\ 4 \end{pmatrix} = \frac{5}{14}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{1}{14}\begin{pmatrix} -5 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

on retrouve bien $x^*$ le point critique.

# Correction mai 2010

ex. Strasbourg : $f(x) = \frac{1}{2}x_1^2 + x_1x_2 + x_2^2 - 3x_1 - 4x_2$

$f(x) = \frac{1}{2}x.Ax - b.x$ avec A symétrique $A = Hf = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ $b = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ A déf positive : $\Delta_1 = 1 > 0, \Delta_2 = 1 > 0$ donc f convexe (même strictement)

$$\nabla f = Ax - b = \begin{pmatrix} x_1 + x_2 - 3 \\ x_1 + 2x_2 - 4 \end{pmatrix}$$

$$\nabla f(x^*) = 0 \rightarrow \begin{cases} x_1 + x_2 = 3 \\ x_1 + 2x_2 = 4 \end{cases} \rightarrow \begin{cases} x_1 = 2 \\ x_2 = 1 \end{cases}$$

gradient conjugué en partant de $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$g^{(0)} = \begin{bmatrix} -3 \\ -4 \end{bmatrix}$ , $d^{(0)} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$

déterminer $\lambda^{(0)} \rightarrow$ minimiser $f(x^{(0)} + \lambda d^{(0)}) = f\begin{pmatrix} 3\lambda \\ 4\lambda \end{pmatrix} \rightarrow$ dérivée directionnelle $= 0 \rightarrow \nabla f\begin{pmatrix} 3\lambda \\ 4\lambda \end{pmatrix} \cdot d^{(0)} = 0 \rightarrow \begin{pmatrix} 7\lambda - 3 \\ 11\lambda - 4 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 4 \end{pmatrix} = 0 \rightarrow \boxed{\lambda^{(0)} = \frac{25}{65} = \frac{5}{13}}$

$x^{(1)} = \frac{5}{13}\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ , $g^{(1)} = \frac{1}{13}\begin{bmatrix} -4 \\ 3 \end{bmatrix}$ , $d^{(1)} = -g^{(1)} + \beta^{(0)} d^{(0)}$, $\beta^{(0)} = \frac{g^{(1)} \cdot Ad^{(0)}}{d^{(0)} \cdot Ad^{(0)}} = \frac{\|g^{(1)}\|^2}{\|g^{(0)}\|^2} = \boxed{\frac{1}{13^2}}$ , $d^{(1)} = \begin{bmatrix} 55 \\ -35 \end{bmatrix} \frac{1}{13^2} = \boxed{\begin{bmatrix} 11 \\ -7 \end{bmatrix} \frac{5}{13^2}}$

$$d^{(1)} \cdot Ad^{(0)} = \frac{1}{13^2}[11 \ \ -7]\begin{bmatrix} 7 \\ 11 \end{bmatrix} = \frac{1}{13^2}(11 \times 7 - 7 \times 11) = 0$$

déterminer $\lambda^{(1)} \rightarrow$ minimiser $f(x^{(1)} + \lambda d^{(1)}) = f\left(\frac{5}{13}\begin{pmatrix} 3 \\ 4 \end{pmatrix} + \lambda \frac{5}{13^2}\begin{pmatrix} 11 \\ -7 \end{pmatrix}\right) \rightarrow$ dérivée directionnelle $= 0 \rightarrow \nabla f\left(\frac{5}{13}\begin{pmatrix} 3 \\ 4 \end{pmatrix} + \frac{5}{13^2}\lambda\begin{bmatrix} 11 \\ -7 \end{bmatrix}\right) \cdot \frac{5}{13^2}\begin{bmatrix} 11 \\ -7 \end{bmatrix} = 0$

$$\begin{bmatrix} \frac{5}{13} \times 3 + \frac{5}{13^2}\lambda 11 + \frac{5}{13} \times 4 - \frac{5}{13^2}\lambda 7 - 3 \\ \frac{5}{13} \times 3 + \frac{5}{13^2}\lambda 11 + \frac{5}{13} \times 8 + \frac{5}{13^2}\lambda(-14) - 4 \end{bmatrix} \cdot \begin{bmatrix} 11 \\ -7 \end{bmatrix} = 0 \rightarrow \boxed{\lambda^{(1)} = \frac{13}{5}}$$

$x^{(2)} = \frac{5}{13}\begin{pmatrix} 3 \\ 4 \end{pmatrix} + \frac{13}{5} \times \frac{5}{13^2}\begin{pmatrix} 11 \\ -7 \end{pmatrix} = \frac{5}{13}\begin{pmatrix} 3 \\ 4 \end{pmatrix} + \frac{1}{13}\begin{pmatrix} 11 \\ -7 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ on retrouve $x^*$ le point critique.

# Méthode de Newton

Soit $f(x_1, x_2) = \frac{x_1^4}{4} + x_2^2 \qquad (x_1, x_2) \in \mathbb{R}^2$

1) Appliquer les 2 premières itérations de la méthode de Newton en partant de $X^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
Calculer les valeurs de $f$ aux différents points parcourus

2) Appliquer les 2 premières itérations de la méthode de Newton en prenant successivement comme pas :
$t^{(0)} = 1$ (question 1) et $t^{(1)} = 3$.
Vérifier que dans ce cas le dernier point est le minimum de $f$.

---

1) $\nabla f = \begin{bmatrix} x_1^3 \\ 2x_2 \end{bmatrix} \qquad H_f = \begin{bmatrix} 3x_1^2 & 0 \\ 0 & 2 \end{bmatrix} \qquad H_f^{-1} = \frac{1}{6x_1^2} \begin{bmatrix} 2 & 0 \\ 0 & 3x_1^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3x_1^2} & 0 \\ 0 & 1/2 \end{bmatrix}$

itération 0
$\nabla f(X^{(0)}) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $H_f^{-1}(X^{(0)}) = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/2 \end{bmatrix}$, $X^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{bmatrix} 1/3 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 0 \end{pmatrix}$

itération 1
$\nabla f(X^{(1)}) = \begin{bmatrix} 8/27 \\ 0 \end{bmatrix}$, $H_f^{-1}(X^{(1)}) = \begin{bmatrix} 3/4 & 0 \\ 0 & 1/2 \end{bmatrix}$, $X^{(2)} = \begin{bmatrix} 2/3 \\ 0 \end{bmatrix} - \begin{bmatrix} 3/4 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 8/27 \\ 0 \end{bmatrix} = \begin{pmatrix} 2/3 \\ 0 \end{pmatrix} - \begin{pmatrix} 2/9 \\ 0 \end{pmatrix} = \begin{pmatrix} 4/9 \\ 0 \end{pmatrix}$

$f(X^{(0)}) = \frac{1}{4} + 1 = \frac{5}{4}$, $f(X^{(1)}) = \left(\frac{2}{3}\right)^4 \times \frac{1}{4} + 0 = \frac{4}{81}$, $f(X^{(2)}) = \left(\frac{4}{9}\right)^4 \times \frac{1}{4} = \left(\frac{2}{3}\right)^{\overbrace{2\times 4}^{4+4}} \times \frac{1}{4} = \frac{4}{81} \times \left(\frac{2}{3}\right)^4$

on voit que les valeurs sont décroissantes

2) itération $k=0$ identique à la précédente
itération $k=1$
avec $t^{(1)} = 3$
$X^{(2)} = \begin{bmatrix} 2/3 \\ 0 \end{bmatrix} - 3 \begin{bmatrix} 3/4 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 8/27 \\ 0 \end{bmatrix} = \begin{pmatrix} 2/3 \\ 0 \end{pmatrix} - 3\begin{pmatrix} 2/9 \\ 0 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 0 \end{pmatrix} - \begin{pmatrix} 2/3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$f(X^{(2)}) = 0$ on ne peut faire mieux.

Exercice

M.q. la fonction $f(x) = x^{4/3}$, $x \in \mathbb{R}$ a un min. global unique mais que pour tout point initial $\neq 0$, la méthode de Newton donne une suite divergente.

$x^{4/3} \geq 0$ et s'annule en $x = 0$

méthode de Newton : $f'(x) = \frac{4}{3} x^{1/3}$ $\quad f''(x) = \frac{4}{9} x^{-\frac{2}{3}}$ $\quad f''^{-1}(x) = \frac{9}{4} x^{2/3}$ $\quad f''^{-1}(x) f'(x) = 3x$

$x^{(1)} = x^{(0)} - 3x^{(0)} = -2x^{(0)} \Rightarrow x^{(k)} = (-2)^k x^{(0)}$ suite divergente sauf si $x^{(0)} = 0$

Exercice

Soit $f(x_1, x_2) = \frac{x_1^4}{4} + x_2^2$

1) M.q. en partant du point $X^{(0)} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$ ($\alpha \neq 0$ quelconque), la méthode de Newton converge vers le minimum de $f$

2) On considère maintenant la méthode de Newton modifiée dans laquelle on minimise dans la direction de Newton $-Hf(x^{(k)}) \nabla f(x^{(k)})$.

Appliquer la méthode de Newton modifiée en partant de $X^{(0)} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$. Que constate-t-on ?

) $\nabla f(x) = \begin{bmatrix} x_1^3 \\ 2x_2 \end{bmatrix}$ $\quad Hf(x) = \begin{bmatrix} 3x_1^2 & 0 \\ 0 & 2 \end{bmatrix}$ $\quad Hf(x)^{-1} = \frac{1}{6x_1^2} \begin{bmatrix} 2 & 0 \\ 0 & 3x_1^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3x_1^2} & 0 \\ 0 & 1/2 \end{bmatrix}$

$$X^{(1)} = \begin{bmatrix} \alpha \\ 0 \end{bmatrix} - \begin{bmatrix} 1/3\alpha^2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} \alpha^3 \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \end{bmatrix} - \begin{bmatrix} 1/3\,\alpha \\ 0 \end{bmatrix} = \begin{bmatrix} 2/3\,\alpha \\ 0 \end{bmatrix}$$

$$X^{(2)} = \begin{bmatrix} (2/3)^2 \alpha \\ 0 \end{bmatrix} \qquad X^{(k)} = \left(\frac{2}{3}\right)^k \begin{bmatrix} \alpha \\ 0 \end{bmatrix} = \left(\frac{2}{3}\right)^k X^{(0)} \qquad X^{(k)} \xrightarrow[k]{} 0$$

2) direction de Newton en $X^{(0)}$ : $-\begin{bmatrix} 1/3\,\alpha \\ 0 \end{bmatrix}$

$X(t) = X^{(0)} - t \begin{bmatrix} \frac{1}{3}\alpha \\ 0 \end{bmatrix} \Rightarrow f(X(t)) = g(t) = \dfrac{(\alpha - \frac{t}{3}\alpha)^4}{4} \geq 0$ $\quad g(t)$ atteint son minimum en $t = 3$

$$X^{(1)} = X(3) = \begin{bmatrix} \alpha \\ 0 \end{bmatrix} - 3 \begin{bmatrix} 1/3\,\alpha \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \end{bmatrix} - \begin{bmatrix} \alpha \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$ minimum atteint.

$f(x) = 5x_1^2 + 2x_1x_2 + x_2^2 - x_1 + x_2$

<u>Appliquer la méthode des gradients conjugués en partant du point</u> $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$\nabla f = \begin{bmatrix} 10x_1 + 2x_2 - 1 \\ 2x_1 + 2x_2 + 1 \end{bmatrix}$$

$$H_f = \begin{bmatrix} 10 & 2 \\ 2 & 2 \end{bmatrix} = A$$

méthode directions gradients conjugués : $f$ est quadratique, on peut appliquer les formules

$$\lambda^{(k)} = -\frac{g^{(k)} \cdot d^{(k)}}{d^{(k)} \cdot A d^{(k)}} \qquad \beta^{(k)} = \frac{\|g^{(k+1)}\|^2}{\|g^{(k)}\|^2}$$

point initial $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $\quad g^{(0)} = \begin{bmatrix} -1 \\ +1 \end{bmatrix}$ $\quad d^{(0)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ $\quad k = 0$

$$\lambda^{(k)} = \frac{\|g^{(0)}\|^2}{d^{(0)} \cdot A d^{(0)}} = \frac{2}{\begin{bmatrix} +1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 8 \\ 0 \end{bmatrix}} = \frac{1}{4} \quad \to \quad x^{(1)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$g^{(1)} = \begin{bmatrix} 10/4 - 2/4 - 1 \\ 2/4 - 2/4 + 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \beta^{(1)} = \frac{2}{2} = 1 \qquad d^{(1)} = \underbrace{-\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{-g^{(1)}} + \underbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{+\beta^{(1)} d^{(0)}} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

vérifier que $d^{(1)} \cdot A d^{(0)} = \begin{bmatrix} 0 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 8 \\ 0 \end{bmatrix} = 0$

itération $k = 1$

$$\lambda^{(1)} = \frac{-g^{(1)} \cdot d^{(1)}}{d^{(1)} \cdot A d^{(1)}} = \frac{-\begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -2 \end{bmatrix}}{\begin{bmatrix} 0 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} -4 \\ -4 \end{bmatrix}} = \frac{2}{8} = \frac{1}{4} \quad \to \quad x^{(2)} = \begin{bmatrix} 1/4 \\ -1/4 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/4 \\ -3/4 \end{bmatrix}$$

$$g^{(2)} = \begin{bmatrix} 10/4 - 2 \times 3/4 - 1 \\ 2/4 - 2 \times 3/4 + 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

on atteint un point critique. STOP.

# Méthode de Cauchy - Méthode du gradient conjugué.

Soit $f$ la fonction définie sur $\mathbb{R}^2$ par $f(x_1,x_2) = 2x_1^2 + x_2^2 - x_1x_2$

On note $x^{(0)}$ le point $(1,4)$

1) En partant de $x^{(0)}$, calculer les 2 premiers termes $x^{(1)}, x^{(2)}$ de la suite de la méthode de Cauchy (plus forte pente). Vérifier que $x^{(2)}-x^{(1)}$ est orthogonal à $x^{(1)}-x^{(0)}$.

2) En partant de $x^{(0)}$, calculer les 2 premiers termes $x^{(1)}, x^{(2)}$ de la suite de la méthode du gradient conjugué. Le minimum est-il atteint ?

$$\nabla f(x_1,x_2) = \begin{bmatrix} 4x_1 - x_2 \\ -x_1 + 2x_2 \end{bmatrix}$$

1) $$\nabla f(\underbrace{1,4}_{x^{(0)}}) = \begin{pmatrix} 0 \\ 7 \end{pmatrix} \quad , \quad x^{(0)} - t\nabla f(x^{(0)}) = \begin{bmatrix} 1 \\ 4-7t \end{bmatrix}$$

minimisation dans la direction $-\nabla f(x^{(0)})$ $\Rightarrow$ annuler la dérivée dans la direction $-\nabla f(x^{(0)})$

$$\nabla f(\overbrace{x^{(0)} - t\nabla f(x^{(0)})}^{x^{(1)}}) \cdot \nabla f(x^{(0)}) = \begin{bmatrix} 7t \\ 7-14t \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 7 \end{bmatrix} = 7(7-14t) = 0 \Rightarrow t = \frac{1}{2}$$

$$x^{(1)} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 0 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}$$

$$\nabla f(x^{(1)}) = \begin{pmatrix} 7/2 \\ 0 \end{pmatrix} \quad , \quad x^{(1)} - t\nabla f(x^{(1)}) = \begin{bmatrix} 1 - \frac{7}{2}t \\ 1/2 \end{bmatrix}$$

minimisation (sur $t$) dans la direction $-\nabla f(x^{(1)})$ $\Rightarrow$ annuler la dérivée dans la direction $-\nabla f(x^{(1)})$

$$\nabla f(x^{(1)} - t\nabla f(x^{(1)})) \cdot \nabla f(x^{(1)}) = \begin{bmatrix} 7/2 - 14t \\ \frac{7}{2}t \end{bmatrix} \cdot \begin{bmatrix} +7/2 \\ 0 \end{bmatrix} = \frac{7}{2}\left(\frac{7}{2} - 14t\right) = 0 \Rightarrow t = \frac{1}{4}$$

$$x^{(2)} = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} - \frac{1}{4}\begin{pmatrix} 7/2 \\ 0 \end{pmatrix} = \begin{bmatrix} 1/8 \\ 1/2 \end{bmatrix}$$

$x^{(1)} - x^{(0)} = -\frac{1}{2}\nabla f(x^{(0)})$ et $x^{(2)} - x^{(1)} = -\frac{1}{4}\nabla f(x^{(1)})$ et on vérifie immédiatement que $\nabla f(x^{(0)}) \perp \nabla f(x^{(1)})$ (annulation de la dérivée dans la direction $-\nabla f(x^{(0)})$ au point $x^{(1)}$)

2) $d^{(0)} = -\nabla f(x^{(0)})$ donc $x^{(1)} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}$ identique à la première itération de méthode de Cauchy.

$$\beta_0 = \frac{\|\nabla f(x^{(1)})\|^2}{\|\nabla f(x^{(0)})\|^2} = \frac{49/4}{49} = \frac{1}{4} \quad , \quad d^{(1)} = -\nabla f(x^{(1)}) + \beta_0 d^{(0)} = -\begin{bmatrix} 7/2 \\ 0 \end{bmatrix} + \frac{1}{4}\begin{bmatrix} 0 \\ -7 \end{bmatrix} = -\begin{bmatrix} 7/2 \\ 7/4 \end{bmatrix}$$

on vérifie que $d^{(0)}$ conjugué par rapport à $A$ à $d^{(1)}$ : $\frac{1}{2}A = \begin{bmatrix} 2 & -1/2 \\ -1/2 & 1 \end{bmatrix}$, $\frac{1}{2}Ad^{(1)} = \begin{bmatrix} -7 + 7/8 \\ 0 \end{bmatrix}$, $\frac{1}{2}d^{(0)} \cdot Ad^{(1)} = 0$

$$x^{(1)} + t\,d^{(1)} = \begin{bmatrix} 1 - \frac{7}{2}t \\ 1/2 - \frac{7}{4}t \end{bmatrix}$$

minimisation (sur $t$) dans la direction $d^{(1)}$ $\Rightarrow$ annuler la dérivée dans cette direction

$$\nabla f(x^{(1)} + t\,d^{(1)}) \cdot d^{(1)} = \begin{bmatrix} 7/2 + \frac{t}{4}(7/4 - 14) \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 7/2 \\ 7/4 \end{bmatrix} = 0 \Rightarrow \frac{7}{2} + t(7/4 - 14) = 0 \rightarrow t = \frac{2}{7}$$

$$x^{(2)} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} + \frac{2}{7}\begin{bmatrix} -7/2 \\ -7/4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$n = 2$ donc normalement le minimum est atteint. Vérifions : $f$ est une forme quadratique définie positive (cf $A$) donc la plus petite valeur qu'elle peut atteindre est 0 pour le vecteur $(0,0)$.

suite de la question 1

construire $d^{(1)}$ conjuguée à $d^{(0)} = \begin{pmatrix} 0 \\ 7 \end{pmatrix}$ par rapport au hessien de $f$.
puis minimiser $f$ selon cette direction en partant de $x^{(1)}$

$$\begin{bmatrix} 0 & 7 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} -7 & 14 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0 \rightarrow -d_1 + 2d_2 = 0 \rightarrow d_1 = 2d_2$$

prenons $d^{(1)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$$x^{(1)} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} \qquad x^{(1)} + t\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 + 2t \\ 1/2 + t \end{bmatrix}$$

méthode l'angle droit pour trouver $t$ le minimiseur.

$$\nabla f\left(x^{(1)} + t\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4(1+2t) - 1/2 - t \\ -1 - 2t + 1 + 2t \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 8(1 + 2t) - 1 - 2t = 0 \rightarrow 7 + 14t = 0 \rightarrow t = -\frac{1}{2}$$

$$x^{(2)} = \begin{bmatrix} 1 - 1 \\ 1/2 - 1/2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

qui est bien point critique.

Sujet B

Directions conjuguées

Soit $f(x)=\frac{1}{2}x^t Ax+b^t x$ avec $A=\begin{pmatrix}4 & -1\\ -1 & 2\end{pmatrix}$, $b=\begin{pmatrix}1\\2\end{pmatrix}$ et $x=\begin{pmatrix}x_1\\x_2\end{pmatrix}\in R^2$.

La transposée d'un vecteur $u=\begin{pmatrix}u_1\\u_2\end{pmatrix}$ est notée $u^t=(u_1 \quad u_2)$.

On veut résoudre le problème :

(P) minimiser $f(x)$ s.c. $x\in R^2$

**1)** On considère le point $x^{(0)}=\begin{pmatrix}\frac{1}{5}\\0\end{pmatrix}$.

Construire $x^{(1)}=x^{(0)}+td^{(0)}$ avec $t\in R$ qui minimise $f\left(x^{(0)}+td^{(0)}\right)$ où $d^{(0)}=-\begin{pmatrix}\frac{9}{5}\\\frac{9}{5}\end{pmatrix}$ est la direction opposée au gradient de $f$ au point $x^{(0)}$.

**2)** On va ici construire la direction suivante.

a) Donner la relation que doit vérifier $d^{(1)}=\begin{pmatrix}d_1\\d_2\end{pmatrix}$ pour être conjuguée à $d^{(0)}$ par rapport à $A$.

b) Donner la condition que doit vérifier $d^{(1)}=\begin{pmatrix}d_1\\d_2\end{pmatrix}$ pour être une direction de descente de $f$ au point $x^{(1)}$.

Tenant compte de la relation trouvée à la question a), en déduire que $d_2$ doit être négatif.

c) Construire $d^{(1)}=\begin{pmatrix}d_1\\d_2\end{pmatrix}$ conjuguée à $d^{(0)}$ par rapport à $A$ et telle que $d_2=-3$.

**3)** Construire $x^{(2)}=x^{(1)}+td^{(1)}$ avec $t\in R$ qui minimise $f\left(x^{(1)}+td^{(1)}\right)$.

Est-il nécessaire de faire une itération supplémentaire ? Vérifier l'optimalité de $x^{(2)}$ pour le problème (P).

## Directions conjuguées (mai 2007 Sujet B)

$$\min f(x) = \frac{1}{2} x^t \underbrace{\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}}_{A} x + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot x \qquad x \in \mathbb{R}^2$$

$$\nabla f = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

mêmes questions que sujet A avec $x^{(0)} = \begin{bmatrix} 1/5 \\ 0 \end{bmatrix}$ et $d_2 = -3$

$d^{(0)} = -\begin{bmatrix} 9/5 \\ 9/5 \end{bmatrix}$ la direction opposée au gradient

---

1) $x^{(1)} = x^{(0)} - t \begin{bmatrix} 9/5 \\ 9/5 \end{bmatrix} = \begin{bmatrix} 1/5 - t\,9/5 \\ -t\,9/5 \end{bmatrix}$ pour trouver $t$ on annule la dérivée directionnelle $\to \nabla f(x^{(1)}) \cdot d^{(0)} = 0$

$$\to \begin{bmatrix} 4/5 - t\,4\times 9/5 + t\,9/5 + 1 \\ -1/5 + t\,9/5 - t\,2\times 9/5 + 2 \end{bmatrix} \cdot \begin{bmatrix} 9/5 \\ 9/5 \end{bmatrix} = 0 \iff \frac{3}{5} - t\left[4\times\frac{9}{5} - 2\times\frac{9}{5} + 2\times\frac{9}{5}\right] + 3 = 0 \iff \frac{\cancel{18}^{2}}{\cancel{9}} - t\frac{4\times\cancel{9}^{2}}{\cancel{5}} = 0 \iff \boxed{t = \frac{1}{2}}$$

$$x^{(1)} = \begin{bmatrix} 1/5 - \frac{9}{5\times 2} \\ -\frac{9}{5\times 2} \end{bmatrix} = \begin{bmatrix} -7/10 \\ -9/10 \end{bmatrix}$$

2) a) $\underbrace{\begin{bmatrix} 9/5 & 9/5 \end{bmatrix}}_{d^{(0)}} \underbrace{\begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}}_{A} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0 \iff \begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0 \iff \boxed{3d_1 + d_2 = 0}$

b) $g^{(1)} = \begin{bmatrix} 4\times(-7/10) + 9/10 + 1 \\ 7/10 - \frac{18}{10} + 2 \end{bmatrix} = \begin{bmatrix} -9/10 \\ 9/10 \end{bmatrix}$ $\quad \underbrace{g^{(1)} \cdot d < 0}_{\text{dérivée directionnelle}} \iff \boxed{-d_1 + d_2 < 0}$

c) $d_2 = -3 \to 3d_1 - 3 = 0 \to \boxed{d_1 = 1}$

3) $x^{(2)} = x^{(1)} + t \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \begin{bmatrix} -7/10 + t \\ -9/10 - 3t \end{bmatrix}$ , pour trouver $t$ on annule la dérivée directionnelle $\to \nabla f(x^{(2)}) \cdot \begin{bmatrix} 1 \\ -3 \end{bmatrix} = 0$

$$\iff \begin{bmatrix} 4(-7/10 + t) + \frac{9}{10} + 3t + 1 \\ \frac{7}{10} - t + -\frac{2\times 9}{10} - 2\times 3t + 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -3 \end{bmatrix} = 0 \iff \begin{bmatrix} -\frac{9}{10} + 7t \\ \frac{9}{10} - 7t \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -3 \end{bmatrix} = 0 \iff -\frac{9}{10} + 7t - \frac{3\times 9}{10} + 3\times 7t = 0 \iff 4\times 7t = 4\times\frac{9}{10}$$

$\begin{bmatrix} -\frac{9}{10} \\ \frac{9}{10} \end{bmatrix} \downarrow g^{(1)}$

$$\boxed{t = \frac{9}{70}}$$

$$x^{(2)} = \begin{bmatrix} -7/10 + 9/70 \\ -9/10 - 3\times 9/70 \end{bmatrix} = \begin{bmatrix} \frac{-49+9}{70} \\ \frac{-63-27}{70} \end{bmatrix} = \begin{bmatrix} -\frac{40}{70} \\ -\frac{90}{70} \end{bmatrix} = \begin{bmatrix} -4/7 \\ -9/7 \end{bmatrix}$$

on vérifie que $x^{(2)}$ est un point critique :

$$4\times(-4/7) + 9/7 + 1 = \frac{-16 + 9 + 7}{7} = 0$$

$$4/7 - \frac{2\times 9}{7} + 2 = \frac{4 - 18 + 14}{7} = 0$$

Méthode de plus forte pente

On considère le pb : minimiser $f(x) = \frac{1}{2} x \cdot Qx - b \cdot x \quad x \in \mathbb{R}^m$ Q symétrique, définie positive.

Soient $e_1, e_2, \dots, e_m$ les vecteurs propres de Q $\left( \|e_i\| = 1 \ , \ e_i \cdot e_j = 0 \right)$

Soit $x_{(0)}$ choisi de sorte que $g_{(0)} \overset{\text{déf.}}{=} \nabla f(x^{(0)})$ appartient à M un sous-espace engendré par un sous-ensemble des $e_i$.

1) M.q. pour la méthode de plus forte pente $g_{(k)} \in M \ \forall k$.

2) Comment évolue l'erreur $E(x^{(k)})$ de l'itération $k$ à $k+1$ si $M = \text{Lin}\{e_i\}$ pour un $e_i$ quelconque

3) M.q. si $x^{(0)}$ est choisi de sorte que $x^{(0)} - x^*$ soit un vecteur propre, la méthode converge en une itération.

1) On note I les indices des vecteurs propres qui engendrent M.

" $d_i$ la valeur propre de $e_i$

par récurrence : supposons $g^{(k)} \in M$ c'est-à-dire $g^{(k)} = \sum_{i \in I} \beta_i e_i$

$$x^{(k+1)} = x^{(k)} - \alpha^{(k)} g^{(k)} \to Qx^{(k+1)} = Qx^{(k)} - \alpha^{(k)} Qg^{(k)} \to g^{(k+1)} = g^{(k)} - \alpha^{(k)} Qg^{(k)}$$

$$\Rightarrow g^{(k+1)} = g^{(k)} - \alpha^{(k)} \sum_{i \in I} \beta_i d_i e_i \in M$$

2) D'après cours
$$\frac{E(x^{(k)}) - E(x^{(k+1)})}{E(x^{(k)})} = \frac{\left(g^{(k)} \cdot g^{(k)}\right)^2}{\left(g^{(k)} \cdot Qg^{(k)}\right)\left(g^{(k)} \cdot Q^{-1} g^{(k)}\right)} = \frac{\beta_i^4}{\left(d_i \beta_i^2\right)\left(\frac{\beta_i^2}{d_i}\right)} = 1$$

$\Rightarrow E(x^{(k+1)}) = 0$. $\Rightarrow$ la méthode converge en au plus une itération.

3) $x^{(0)} - x^* = e_i$ pour un certain $i$

$$\Rightarrow Q(x^{(0)}) - \underbrace{Qx^*}_{b} = Qe_i = d_i e_i \Rightarrow g^{(0)} = d_i e_i$$ c'est-à-dire $g^{(0)} = \text{Lin}\{e_i\}$

d'après 2) dans ce cas la méthode converge en au plus une itération.

minimiser $f(x) = \frac{1}{2} x \cdot Qx - b \cdot x$ [illegible]
$x \in \mathbb{R}^n$

[illegible] nite
$n \times n$ matrix $Q$. For the [illegible] pose
$\mathbf{x}_0$ is chosen so that $\mathbf{g}_0$ belongs to a subspace $M$ spanned by a subset of the $\mathbf{e}_i$'s.
[illegible] Find the rate of
[illegible]

[illegible]
valeur propre [illegible]

Soit $M = \text{Lin}\{e_i : i \in I\}$  $I$ = indices d'un sous-ensemble des vecteurs propres.
on note $\lambda_i$ valeur propre de $e_i$

1) Par récurrence : $g^{(0)} \in M$

hyp. de récurrence : $g^{(k)} \in M$

$$x^{(k+1)} = x^{(k)} - \alpha^k g^{(k)} \Rightarrow Qx^{(k+1)} = Qx^{(k)} - \alpha^{(k)} Qg^{(k)} \rightarrow g^{(k+1)} = g^{(k)} - \alpha^{(k)} Q\left(\sum_{i \in I} \beta_i e_i\right)$$

$$\Rightarrow g^{(k+1)} = g^{(k)} - \alpha^{(k)} \sum_{i \in I} \beta_i \lambda_i e_i \in M$$

2) Les vecteurs propres sont orthonormés $(P^t P = I)$ i.e. $e_i \cdot e_i = 1$ $e_i \cdot e_j = 0$ $i \neq j$.

$$\frac{E(x^{(k)}) - E(x^{(k+1)})}{E(x^{(k)})} = \frac{(g^{(k)} \cdot g^{(k)})^2}{(g^{(k)} \cdot Qg^{(k)})(g^{(k)} \cdot Q^{-1} g^{(k)})} = \frac{\left(\sum_{i \in I} \beta_i^2\right)^2}{\left(\sum_{i \in I} \lambda_i \beta_i^2\right)\left(\sum_{i \in I} \frac{\beta_i^2}{\lambda_i}\right)} \geq \frac{4\lambda_{max}\lambda_{min}}{(\lambda_{min} + \lambda_{max})^2}$$

cf. cours

$e_i$ orthonormés.

cf. démonstration de l'inégalité de Kantorovitch

où $\lambda_{min} = \min_{i \in I} \lambda_i$

$\lambda_{max} = \max_{i \in I} \lambda_i$

$$g^{(k)} = \sum_{i \in I} \beta_i e_i$$

$$\Rightarrow E(x^{(k+1)}) \leq E(x^{(k)}) \left(\frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} - \lambda_{min}}\right)^2$$

3) $x^{(0)} - x^* = e_i$ pour un certain $i$.

dans ce cas $\lambda_{max} = \lambda_{min} = \lambda$

$$\Rightarrow \underbrace{Q(x^{(0)}) - Q\underbrace{x^*}_{b}}_{g^{(0)}} = Qe_i = \lambda_i e_i \rightarrow g^{(0)} \in \text{Lin}\{e_i\}$$

$$\Rightarrow E(x^{(1)}) \leq E(x^{(0)}) \left(\frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}}\right)^2 = 0 \Rightarrow x^{(1)} = x^*$$

méthode quasi-newtonienne : <u>méthode de rang 1</u>

Appliquer la méthode de rang 1 à $f(x_1, x_2) = x_1^2 - x_1 x_2 + \frac{3}{2} x_2^2$
en partant du point $x^{(0)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ et de la matrice $S^{(0)} = I$.
On a $\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 - x_2 \\ -x_1 + 3x_2 \end{pmatrix}$

<u>itération 1</u>

$\nabla f(x^{(0)}) = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$ $\quad S^{(0)} = I$ $\quad \Rightarrow$ on minimise dans direction opposée au gradient.
i.e min $f\left[\begin{pmatrix} 1 \\ 2 \end{pmatrix} - d \begin{pmatrix} 0 \\ 5 \end{pmatrix}\right]$

annulation dérivée directionnelle : $\nabla f\left[\begin{pmatrix} 1 \\ 2 \end{pmatrix} - d\begin{pmatrix} 0 \\ 5 \end{pmatrix}\right] \cdot \begin{pmatrix} 0 \\ 5 \end{pmatrix} = (-1 + 6 - 15d) \cdot 5 = 0 \rightarrow d = \frac{1}{3}$

$$x^{(1)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1/3 \end{bmatrix}$$

par curiosité regardons le gain obtenu : $f(x^{(0)}) = 5$, $f(x^{(1)}) = \frac{5}{6}$

itération 1 (suite) : calcul de $S^{(1)}$

$\left(\nabla f(x^{(1)}) - \nabla f(x^{(0)})\right) = y^{(0)} = \begin{pmatrix} 5/3 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 5/3 \\ -5 \end{pmatrix}$ $\qquad d^{(0)} = x^{(1)} - x^{(0)} = -\begin{bmatrix} 0 \\ 5/3 \end{bmatrix}$

$$y^{(0)} \cdot \left(d^{(0)} - S^{(0)} y^{(0)}\right) = \begin{bmatrix} 5/3 \\ -5 \end{bmatrix} \cdot \left[\begin{bmatrix} 0 \\ -5/3 \end{bmatrix} - \begin{bmatrix} 5/3 \\ -5 \end{bmatrix}\right] = -\frac{175}{9}$$

$$\left(d^{0} - S^{(0)} y^{(0)}\right)\left(d^{(0)} - S^{(0)} y^{(0)}\right)^t = \begin{bmatrix} -5/3 \\ 10/3 \end{bmatrix}\begin{bmatrix} -5/3 & 10/3 \end{bmatrix} = \begin{bmatrix} 25/9 & -50/9 \\ -50/9 & \frac{100}{9} \end{bmatrix}$$

$$S^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{9}{\underbrace{175}_{25\times7}} \begin{bmatrix} 25/9 & -50/9 \\ -50/9 & 100/9 \end{bmatrix} = \begin{bmatrix} 6/7 & 2/7 \\ 2/7 & 3/7 \end{bmatrix}$$

<u>itération 2</u> : direction $S^{(1)} \nabla f(x^{(1)}) = \begin{bmatrix} 6/7 & 2/7 \\ 2/7 & 3/7 \end{bmatrix}\begin{bmatrix} 5/3 \\ 0 \end{bmatrix} = \frac{10}{7}\begin{bmatrix} 1 \\ 1/3 \end{bmatrix}$

min $f\left[\begin{bmatrix} 1 \\ 1/3 \end{bmatrix} - d\frac{10}{7}\begin{bmatrix} 1 \\ 1/3 \end{bmatrix}\right]$

annulation dérivée directionnelle : $\begin{pmatrix} 2 - d\frac{10}{7}\times 2 - \frac{1}{3} + d\frac{10}{7\times3} \\ -1 + d\frac{10}{7} + 1 - d\frac{10}{7} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1/3 \end{pmatrix} = \frac{5}{3} + d\left[\frac{10}{7\times3} - \frac{10}{7}\times 2\right] = \frac{5}{3} + d\frac{10}{7}\times\left(-\frac{5}{3}\right) = 0 \rightarrow d = \frac{7}{10}$

$$x^{(2)} = \begin{bmatrix} 1 \\ 1/3 \end{bmatrix} - \begin{bmatrix} 1 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$f$ minimum atteint ? $\quad f(x_1, x_2) = \left(x_1 - \frac{x_2}{2}\right)^2 + \frac{5}{4} x_2^2$ $\quad$ définie positive.

Sept. 99 

## Méthode du gradient conjugué

9 pts

Soit $f(x) = x_1^2 + x_2^2 - x_1 x_2 - 3x_2$

Trouver le minimum de $f(x)$ en appliquant la méthode du gradient conjugué avec comme point initial $\begin{pmatrix}0\\0\end{pmatrix}$

$$\nabla f = \begin{bmatrix} 2x_1 - x_2 \\ 2x_2 - x_1 - 3 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \qquad x^{(0)} = \begin{pmatrix}0\\0\end{pmatrix}$$

itération (0).

$$\nabla f(x^{(0)}) = \begin{bmatrix}0\\-3\end{bmatrix} \qquad d^{(0)} = \begin{bmatrix}0\\3\end{bmatrix} \to 1\,pt \qquad x = \begin{bmatrix}0\\0\end{bmatrix} + d\begin{bmatrix}0\\3\end{bmatrix} = \begin{bmatrix}0\\3d\end{bmatrix}$$

minimiser $9d^2 - 9d$ dérivé $\to 18d - 9 = 0 \Rightarrow d = \frac{1}{2}$ (2 pts) $\quad x^{(1)} = \begin{bmatrix}0\\3/2\end{bmatrix} \to 1\,pt$

itération (1)

$$\nabla f(x^{(1)}) = \begin{bmatrix}-3/2\\0\end{bmatrix} \qquad \beta^{(1)} = \frac{\begin{bmatrix}-3/2\\0\end{bmatrix} \cdot \begin{bmatrix}2 & -1\\-1 & 2\end{bmatrix}\begin{bmatrix}0\\3\end{bmatrix}}{\begin{bmatrix}0\\3\end{bmatrix} \cdot \begin{bmatrix}2 & -1\\-1 & 2\end{bmatrix}\begin{bmatrix}0\\3\end{bmatrix}} = \frac{1}{4} \to 1\,pt$$

$$d^{(1)} = \underbrace{\begin{bmatrix}3/2\\0\end{bmatrix}}_{-\nabla f(x^{(1)})} + \underbrace{\frac{1}{4}}_{\beta^{(1)}}\underbrace{\begin{bmatrix}0\\3\end{bmatrix}}_{d^{(0)}} = \begin{bmatrix}3/2\\3/4\end{bmatrix} \quad 1\,pt$$

$$\beta^{(1)} = \frac{\|g^{(1)}\|}{\|g^{(0)}\|} = \frac{1}{4}$$

$$x = \begin{bmatrix}0\\3/2\end{bmatrix} + d\begin{bmatrix}3/2\\3/4\end{bmatrix} = \begin{bmatrix}3/2\,d\\ \frac{3}{2} + \frac{3}{4}d\end{bmatrix}$$

minimiser $\left(\frac{3}{2}\right)^2 d^2 + \left[\frac{3}{2} + \frac{3}{4}d\right]^2 - \frac{3}{2}d\left[\frac{3}{2} + \frac{3}{4}d\right] - 3\left[\frac{3}{2} + \frac{3}{4}d\right]$

dérivé $\to d\left(2\times\left(\frac{3}{2}\right)^2 + \frac{9}{4\times 2} - \frac{9}{4}\right) + \frac{9}{4} - \frac{9}{4} - \frac{9}{4} = 0 \Rightarrow d = \frac{2}{3} \to 2\,pts$

$x^{(2)} = \begin{bmatrix}1\\2\end{bmatrix} \to 1\,pt$

dérivée dans direction $d^{(1)}$ $\quad \nabla f(x) \cdot d^{(1)} = 0 \Leftrightarrow \begin{bmatrix}3d - \frac{3}{4}d - \frac{3}{2}\\0\end{bmatrix} \cdot \begin{bmatrix}3/2\\3/4\end{bmatrix} = 0 \Leftrightarrow 9d = 6$

on peut vérifier que $\nabla f(x^{(2)}) = \begin{bmatrix}0\\0\end{bmatrix}$

Contrôle 99

directions conjuguées

On considère le pb (P) : minimiser $f(x) = \frac{1}{2} x \cdot Hx + b \cdot x$ , $x \in \mathbb{R}^n$

où H matrice $(n \times n)$ définie positive.

Soient $d_1, \dots, d_n$ n vecteurs non nuls, conjugués 2 à 2 par rapport à H i.e $d_j \cdot H d_i = 0 \quad \forall j \neq i$

On sait que $d_1, \dots, d_n$ forment une base de $\mathbb{R}^n$

1) [En écrivant $x = \sum_{i=1}^{n} \alpha_i d_i \quad \alpha_i \in \mathbb{R}$,] montrer que [(P) se décompose en les n problèmes de de minimisation unidimensionnelle suivants : $\underset{\alpha_i}{\text{minimiser}}\ f(\alpha_i d_i)$ et que par conséquent] la solution de (P) est donnée par $x^* = \sum_{i=1}^{n} \alpha_i^* d_i$ où $\alpha_i^*$ $\underset{\alpha_i}{\text{minimise}}\ f(\alpha_i d_i)$

2) Résoudre le problème minimiser $f(x) = 4x_1^2 + 4x_2^2 - 4x_1x_2 - 12x_2$ par la méthode précédente avec $d_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ , $d_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

# Correction

1) $f(x) = f\left(\sum_i \alpha_i d_i\right) = \frac{1}{2}\left(\sum_i \alpha_i d_i\right) \cdot H \left(\sum_i \alpha_i d_i\right) + b \cdot \left(\sum_i \alpha_i d_i\right)$

conjuguées $\Rightarrow d_j \cdot H d_i = 0 \Rightarrow \frac{1}{2} \sum_i \left(\alpha_i d_i \cdot H \alpha_i d_i\right) + \sum_i \left(b \cdot \alpha_i d_i\right) = \sum_i f(\alpha_i d_i)$

donc $\min_{x \in \mathbb{R}^n} f(x) = \min_{\alpha \in \mathbb{R}^n} \sum_i f(\alpha_i d_i) = \sum_i \min_{\alpha_i} f(\alpha_i d_i)$

soient $\alpha_i^*$ solution de $\min_{\alpha_i} f(\alpha_i d_i)$ alors $x^* = \sum \alpha_i^* d_i$ est solution de $(P)$

2) $f(x_1, x_2) = 4x_1^2 + 4x_2^2 - 4x_1 x_2 - 12 x_2$

$d_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow x_1 = 1,\ x_2 = 0 \Rightarrow f(\alpha d_1) = 4\alpha^2 \Rightarrow \alpha_1^* = 0$

$d_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow x_1 = 1,\ x_2 = 2 \Rightarrow f(\alpha d_2) = 4\alpha^2 + 4 \times 4\alpha^2 - 4 \times 2\alpha^2 - 12 \times 2\alpha$

$= 12\alpha^2 - 24\alpha$

$f'(\alpha^* d_2) = 24\alpha^* - 24 = 0 \Rightarrow \alpha_2^* = 1$

$x^* = \alpha_1^* d_1 + \alpha_2^* d_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

vérification : $H x^* + b = 0 \qquad \begin{bmatrix} 8 & -4 \\ -4 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ -12 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Méthode des directions conjuguées

Soit la fonction quadratique $f(x_1, x_2) = 2x_1 - 4x_2 + x_1^2 + 2x_1x_2 + 3x_2^2 \quad (x_1, x_2) \in \mathbb{R}^2$.

1) Calculer H le hessien de f

Montrer que $d^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ et $d^{(1)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ sont conjuguées par rapport à H

Partant d'un point $x^{(0)}$ de $\mathbb{R}^2$, si on applique la méthode des directions conjuguées, avec les directions $d^{(0)}$ et $d^{(1)}$, va-t-on atteindre le minimum de f et en combien d'itérations ? 2

2) Appliquer à f la méthode des directions conjuguées, avec les directions $d^{(0)}$ et $d^{(1)}$, en partant de $x^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

2

1) $H = \begin{bmatrix} 2 & 2 \\ 2 & 6 \end{bmatrix} \qquad (1 \quad 0) \begin{bmatrix} 2 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 0$

f se met sous la forme $\frac{1}{2} x \cdot Hx + c \cdot x$ avec H déf. positive et $d^{(0)}, d^{(1)}$ conjuguées par rapport à H alors d'après le th. des directions conjuguées on atteint le point critique de f (donc son minimum) en au plus 2 itérations. $(n=2)$

2) itération 0. $\begin{bmatrix} 0 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda \\ 0 \end{bmatrix} \qquad \min_{\lambda} 2\lambda + \lambda^2 \qquad$ point critique $2\lambda + 2 = 0 \rightarrow \lambda = -1$

itération 1. $\begin{bmatrix} -1 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\lambda - 1 \\ \lambda \end{bmatrix}$

$$\min_{\lambda} -2(\lambda+1) - 4\lambda + (\lambda+1)^2 - 2(\lambda+1)\lambda + 3\lambda^2$$

$$'' \qquad (\lambda+1)(\lambda+1-2\lambda) + \quad ''$$

$$'' \qquad (\lambda+1)(1-\lambda) + \quad ''$$

$$'' \qquad -\lambda^2 + 1 \quad + \quad '' \quad = -1 - 6\lambda + 2\lambda^2$$

pt critique : $4\lambda = 6 \rightarrow \lambda = \frac{3}{2}$

$$x^{(2)} = \begin{bmatrix} -5/2 \\ 3/2 \end{bmatrix}$$

# Méthode des directions conjuguées.

Soit $f$ de $\mathbb{R}^n \to \mathbb{R}$ définie par $f(x) = c^t x + \frac{1}{2} x^t H x$

où $H$ est une matrice carrée, symétrique, d'ordre $n$

$c$ un vecteur colonne de $n$ lignes.

$t$ l'opérateur "transposé" est noté à droite de son argument.

Soit $D$ une matrice carrée d'ordre $n$ dont les colonnes $d_1, \dots, d_n$ sont conjuguées deux à deux par rapport à $H$.

Soit $u$ un vecteur (colonne) de $\mathbb{R}^n$ fixé.

On supposera que les colonnes de $D$ sont lin$^t$ ind$^t$ de sorte que tout $x$ de $\mathbb{R}^n$ peut s'écrire $x = Dy + u$ (où $y \in \mathbb{R}^n$)

1) On pose $\beta = D^t H D$ et on remarque que $\beta_{ij} = d_i^t H d_j$.

a) M.q. $f(x)$ peut s'écrire $g(y) = \sum_{i=1}^{n} \alpha_i y_i + \frac{1}{2} \sum_{i=1}^{n} \beta_{ii} y_i^2 + c^t u + \frac{1}{2} u^t H u$

avec $(\alpha_1 \dots \alpha_n) = (c^t + u^t H) D$

(On rappelle que $(A+B)^t = A^t + B^t$ et $(AB)^t = B^t A^t$)

b) En déduire que le minimum de $f$ se décompose en une somme de minima de $n$ fonctions indépendantes les unes des autres, plus une constante.

c) Appliquer le résultat précédent pour calculer le minimum de $2x_1 - 4x_2 + x_1^2 + 2x_1x_2 + 3x_2^2$ ainsi que le point $x$ réalisant ce minimum, en prenant $D = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ et $u = 0$.

2) On considère maintenant $u$ point critique de $f$.

a) M.q. $f(x)$ s'écrit $g(y) = c^t u + \frac{1}{2} u^t H u + \frac{1}{2} \sum_{i=1}^{n} \beta_{ii} y_i^2$

b) Quel résultat relatif au minimum de $f$, cette écriture permet-elle de retrouver lorsque $H$ est semi-définie positive ?

c) Appliquer cette réécriture à $f(x) = 2x_1 - 4x_2 + x_1^2 + 2x_1x_2 + 3x_2^2$ avec $D = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ et donner l'allure, dans le plan $x_1, x_2$ (muni d'un repère orthonormé), des courbes de niveau $f(x) = -\frac{11}{2} + c^{te}$ où $c^{te}$ est une constante $\geq 0$

d) Appliquer la réécriture du a) à $f(x) = \ldots$

On note $S$ la courbe de niveau $f(x) = -\frac{11}{2} + c^{te}$ c'est-à-dire

$S = \{ x : f(x) + \frac{11}{2} - c^{te} = 0 \}$ où $c^{te}$ est une constante positive ou nulle.

En $x \in S$ t.q. $\nabla f(x) \neq 0$, peut-on ~~donner~~ expliciter $\underbrace{\text{le cône tangent}}_{T(S,x)}$ à $S$ en $x$ ?

En utilisant le resultat de la question 2c), ~~pour~~ tracer dans le plan $x_1, x_2$, (muni d'un repère orthonormé) le cône tangent à $S$ en les points

$$x = \begin{bmatrix} -5/2 \\ 3/2 \end{bmatrix} \pm \sqrt{c^{te}}\, d_1 \quad \text{et} \quad x = \begin{bmatrix} -5/2 \\ 3/2 \end{bmatrix} \pm \sqrt{\frac{c^{te}}{2}}\, d_2 \qquad \text{pour } c^{te} > 0$$

puis donner l'allure de $S$ pour $c^{te} > 0$ et $c^{te} = 0$.

# Méthode des directions conjuguées

Soit $f$ de $\mathbb{R}^n \to \mathbb{R}$ définie par $f(x) = c^t x + \frac{1}{2} x^t H x$

où $H$ est une matrice carrée, symétrique, d'ordre $n$
$c$ est un vecteur colonne de $n$ lignes.

$t$ désigne l'opérateur "transposé" et est noté à droite de son argument.

Soit $D$ une matrice carrée, d'ordre $n$ dont les colonnes $d_1, \dots, d_n$ sont conjuguées 2 à 2 par rapport à $H$.
Soit $u$ un vecteur colonne de $\mathbb{R}^n$ fixé.
On supposera que les colonnes de $D$ sont linéairement indépendantes de sorte que tout $x$ de $\mathbb{R}^n$ peut s'écrire $x = Dy + u$ $(y \in \mathbb{R}^n)$

1) On pose $\beta = D^t H D$ et on remarque que $\beta_{ij} = d_i^t H d_j$

a) M.q. $f(x)$ peut s'écrire $g(y) = \sum_{i=1}^{n} \alpha_i y_i + \frac{1}{2} \sum_{i=1}^{n} \beta_{ii} y_i^2 + c^t u + \frac{1}{2} u^t H u$
avec $(\alpha_1 \cdots \alpha_n) = (c^t + u^t H) D$

(On rappelle que $(A+B)^t = A^t + B^t$ et $(AB)^t = B^t A^t$ et que si $A$ est symétrique $A^t = A$ )

b) En déduire que le minimum de $f$ se décompose en une somme de minima de $n$ fonctions de une variable plus une constante.

c) Appliquer le résultat précédent pour calculer le minimum de $2x_1 - 4x_2 + x_1^2 + 2x_1x_2 + 3x_2^2$ ainsi que le point $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ réalisant ce minimum, en prenant $D = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ et $u = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

2) On suppose que $f$ admet un point critique et
On considère maintenant $u$ point critique de $f$

a) M.q $f(x)$ s'écrit $g(y) = c^t u + \frac{1}{2} u^t H u + \frac{1}{2} \sum_{i=1}^{n} \beta_{ii} y_i^2$

b) Appliquer la réécriture précédente à $q(x) = \frac{1}{2} x^t H x$ (avec $x = Dy + v$ où $v$ est un point critique de $q$)
m.q. si $\exists\, z$ t.q. $q(z) < 0$ alors $\exists\, \beta_{ii} < 0$

En déduire que $f$ admet un minimum si et seulement si $H$ est semi-défini positif. (i.e $f$ est convexe)

c) Soit $N$ un sous-ensemble non vide de $\{1, \dots, n\}$.
M.q. en un point $x = Dy + u$ t.q. $y_i = 0$ $\forall i \in N$, $y'^t D^t \nabla f(x) = 0$ $\forall y'$ t.q. $y'_i = 0$ pour $i \notin N$
(autrement dit $\nabla f(x)$ est orthogonal à l'espace engendré par $\{d_i ;\ i \in N\}$)
M.q. si $H$ est défini positive, ce résultat n'est autre que le théorème des directions conjuguées.

d) Appliquer la réécriture du a) à $f(x) = 2x_1 - 4x_2 + x_1^2 + 2x_1x_2 + 3x_2^2$ avec $D = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$
Dans le plan $x_1, x_2$ (muni d'un repère orthonormé), donner l'allure des courbes de niveau $\{x : f(x) = -\frac{11}{2} + c^{te}\}$ où $c^{te}$ est une constante positive ou nulle.
Sachant qu'en tout point d'une courbe de niveau, le gradient de $f$ est orthogonal à cette courbe de niveau, on s'aidera du résultat du c) (indiqué entre parenthèses) pour représenter la courbe dans le voisinage des points $x = \begin{bmatrix} -5/2 \\ 3/2 \end{bmatrix} \pm \sqrt{c^{te}}\, d_1$ et $x = \begin{bmatrix} -5/2 \\ 3/2 \end{bmatrix} \pm \sqrt{\frac{c^{te}}{2}}\, d_2$

# Correction

1) a) $f(x) = c^t x + \frac{1}{2} x^t H x$ et $x = Dy + u$

$$= c^t(Dy+u) + \frac{1}{2}(y^t D^t + u^t) H (Dy+u)$$

$$= c^t Dy + c^t u + \frac{1}{2} y^t D^t H D y + \frac{1}{2} \underbrace{y^t D^t H u}_{} + \frac{1}{2} u^t H D y + \frac{1}{2} u^t H u$$

$$= \quad " \qquad " \qquad u^t H D y \quad (H \text{ symétrique}) \qquad "$$ 1 pt

$$= (c^t + u^t H) D y + \frac{1}{2} y^t \underbrace{D^t H D}_{} y + c^t u + \frac{1}{2} u^t H u$$

$\beta$ qui est diagonale (directions $d_i, d_j$ conjuguées) 1 pt

$$= \sum_{i=1}^{n} \alpha_i y_i + \frac{1}{2} \sum_{i=1}^{n} \beta_{ii} y_i^2 + c^t u + \frac{1}{2} u^t H u = g(y)$$

avec $(\alpha_1 \ldots \alpha_n) = (c^t + u^t H) D$

b) $f(x) = g(y) = \sum_{i=1}^{n} g_i(y_i) + c^{te}$ avec $g_i(y_i) = \alpha_i y_i + \frac{1}{2}\beta_{ii} y_i^2$ et $c^{te} = c^t u + \frac{1}{2} u^t H u$

$$\to \min_x f(x) = \min_y g(y) = \sum_{i=1}^{n} \min_{y_i} g_i(y_i) + c^{te}$$

c) $f(x) = 2x_1 - 4x_2 + x_1^2 + 2x_1 x_2 + 3x_2^2 \to c = (2 \quad -4)$ $\frac{1}{2} H = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$

$c^t D = (2 \quad -4)\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = (2 \quad -6)$ $\frac{1}{2}\beta_{11} = [1 \quad 0]\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1$ $\frac{1}{2}\beta_{22} = [-1 \quad 1]\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = 2$

$g(y) = 2y_1 - 6y_2 + y_1^2 + 2y_2^2$ 1 pt

$g_1(y_1) = 2y_1 + y_1^2 \to g_1'(y_1) = 2 + 2y_1 = 0 \to y_1 = -1 \to \min g_1(y_1) = -1$

$g_2(y_2) = -6y_2 + 2y_2^2 \to g_2'(y_2) = -6 + 4y_2 = 0 \to y_2 = \frac{3}{2} \to \min g_2(y_2) = -\frac{9}{2}$ 1 pt

$\min g(y) = \min g_1(y_1) + \min g_2(y_2) = -\frac{11}{2} = \min f(x)$

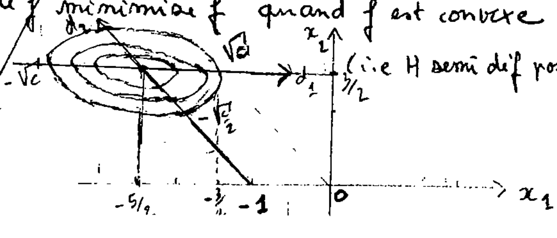
solution en $x$ : $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 3/2 \end{bmatrix} = \begin{bmatrix} -5/2 \\ 3/2 \end{bmatrix}$ 1 pt

2) a) $u$ point critique $\to Hu + c = 0 \xrightarrow{t} u^t H + c^t = 0$ (H symétrique) $\to \alpha = 0$

b) H semi-déf. positive $\to \beta_{ii} \geq 0 \to y = 0$ minimise $g(y) = c^t u + \frac{1}{2} u^t H u + \frac{1}{2} \sum_{i=1}^{n} \underbrace{\beta_{ii} y_i^2}_{\geq 0}$

$y = 0 \to x = u$ (point critique). Un point critique de $f$ minimise $f$ quand $f$ est convexe (i.e H semi déf positif)

c) $f(x) = 2x_1 - 4x_2 + x_1^2 + 2x_1 x_2 + 3x_2^2 = \underset{\substack{\uparrow \\ \min f}}{-\frac{11}{2}} + y_1^2 + 2y_2^2$



b) pour $q(x)$ on a $c=0$ et $Hv=0$ point critique $v$

$\rightarrow q(x) = \frac{1}{2}\sum_{i=1}^{n} \beta_{ii} y_i^2$

1 pt

supposons $q(z)<0$. Si $\beta_{ii} \geq 0 \ \forall i=1,\dots,m$ alors $\underbrace{\sum_{i=1}^{m} \beta_{ii} y_i^2 \geq 0}_{\forall y}$ et donc $q(z) \geq 0 \rightarrow$ impossible.

si $H$ semi-défini positif alors $\beta_{ii} \geq 0$ et le min. de $f$ est atteint en prenant $y=0$.

(On remarque alors $x=u$ est point critique et donc le point critique minimise $f$.)

1 pt

si $H$ pas semi-défini positif alors $\exists z$ t.q. $q(z)<0$ et donc $\exists \beta_{ii}<0$.

En prenant $y = \begin{bmatrix} 0 \\ \vdots \\ y_i \\ \vdots \\ 0 \end{bmatrix}$ et $y_i \rightarrow \infty$ alors $f(x) = g(y) \longrightarrow -\infty$

c) $\nabla f(x) = Hx + c \rightarrow HDy + \underbrace{Hu+c}_{0 \text{ u point critique}}$

$y'^t D^t H D y = \sum_{i,j} y'_i \beta_{ij} y_j \underset{\uparrow}{=} \sum_{i\neq j} y'_i \beta_{ij} y_j \underset{\uparrow}{=} 0$ 1 pt.

$y'_i y_i = 0 \ \forall i$ car $\begin{cases} \text{si } i\in N \ y_i = 0 \\ \text{si } i \notin N \ y'_i = 0 \end{cases}$

$\beta$ est diagonale

A une itération $k$ de la méthode des directions conjuguées, on minimise $f(x+\alpha_k d_k)$ selon $\alpha_k$.
Soit $x = Dy+u$ alors $x+\alpha_k d_k = Dy'+u$ avec $y' = \begin{bmatrix} y_1 \\ \vdots \\ y_k+\alpha_k \\ \vdots \\ y_m \end{bmatrix}$

et $f(x+\alpha_k d_k) = c^t u + \frac{1}{2} u^t H u + \frac{1}{2}\sum_{i\neq k} \beta_{ii} y_i^2 + \frac{1}{2}\beta_{kk}(y_k+\alpha_k)^2$

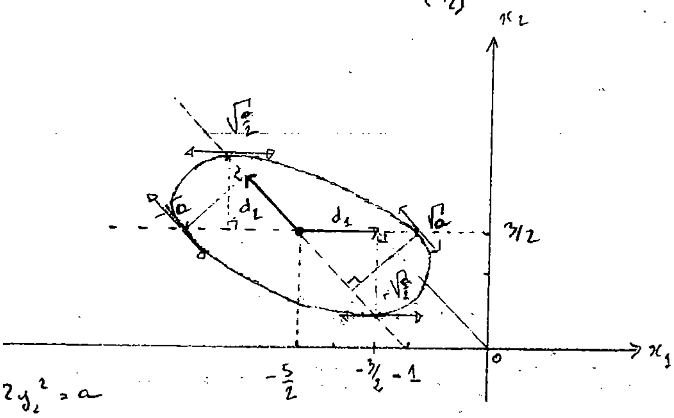
$H$ déf. positive $\Rightarrow \beta_{kk}>0$. Pour minimiser il faut $\alpha_k = -y_k$ c'est à dire $y_k + \alpha_k = 0$.
Après avoir minimisé suivant les directions $d_i$ $i\in N$, on obtient $x$ t.q. $y_i = 0 \ \forall i \in N$
et donc t.q. $d_i^t \nabla f(x) = 0 \ \forall i\in N$ (c'est le théorème des directions conjuguées)

d) $f(x) = 2x_1 - 4x_2 + x_1^2 + 2x_1x_2 + 3x_2^2 \underset{\uparrow}{=} -\frac{11}{2} + y_1^2 + 2y_2^2$

minimum de $f$ atteint en $u = \begin{bmatrix} -5/2 \\ 3/2 \end{bmatrix}$

soit $h(x) = f(x) + \frac{11}{2} - c^t c$
alors $\nabla h(x) = \nabla f(x) \ \forall x$
si $\nabla h(x) \neq 0$ (i.e $\nabla f(x) \neq 0$)
alors $x$ est qualifié et donc
$T(S,x) = \{ y : \nabla f(x) \cdot y = 0 \}$
Ici $\nabla f(x) \neq 0$ ssi $x \neq \begin{bmatrix} -5/2 \\ 3/2 \end{bmatrix}$

1 pt




2 pt.

$f(x) = -\frac{11}{2} + y_1^2 + 2y_2^2 = -\frac{11}{2} + a \rightarrow y_1^2 + 2y_2^2 = a$

$x$ t.q. $y_1 = 0 \rightarrow y_2 = \pm\sqrt{\frac{a}{2}} \rightarrow x-u = \pm\sqrt{\frac{a}{2}}\, d_2$ et $\|x-u\| = \sqrt{a}$

$x$ t.q. $y_2 = 0 \rightarrow y_1 = \pm\sqrt{a} \rightarrow x-u = \pm\sqrt{a}\, d_1$ et $\|x-u\| = \sqrt{a}$

distances à $u$ égales

## Algorithme des Directions conjuguées

On considère le problème (P) : minimiser $f(x) = \frac{1}{2} x \cdot Hx + b \cdot x$ où $H$ matrice $n \times n$ définie positive et $x \in \mathbb{R}^n$

Soient $d_0, d_1, \dots, d_{n-1}$ $n$ vecteurs non nuls, conjugués 2 à 2 par rapport à $H$ i.e $d_j \cdot H d_i = 0 \quad j \neq i$

On sait que $d_0, \dots, d_{n-1}$ forment une base de $\mathbb{R}^n$.

1) En écrivant $x$ dans la base $d_0, \dots, d_{n-1}$ i.e. $x = \sum_{i=0}^{n-1} \alpha_i d_i$, montrer que (P) se décompose en les $n$ problèmes de minimisation à une dimension suivants $\underset{\alpha_i \in \mathbb{R}}{\text{minimiser}} f(\alpha_i d_i)$ et que la solution de (P) est donnée par $x^* = \sum_{i=0}^{n-1} \alpha_i^* d_i$ où $\alpha_i^*$ minimise $f(\alpha_i d_i)$

2) Montrer que lorsque le point initial est $x^{(0)} = 0$ l'algorithme des directions conjuguées revient à résoudre (P) selon la méthode de la question 1).

3) A.N. Résoudre le problème minimiser $4x_1^2 + 4x_2^2 - 4x_1x_2 - 12x_2$ par la méthode de la question 1) ou 2) (ce sont en fait les mêmes) avec $d_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $d_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

---

1) $$f(x) = f\left(\sum_i \alpha_i d_i\right) = \frac{1}{2}\left(\sum_i \alpha_i d_i\right) \cdot H\left(\sum_i \alpha_i d_i\right) + b \cdot \left(\sum_i \alpha_i d_i\right)$$

conjuguées $\Rightarrow d_j \cdot H d_i = 0 \Rightarrow = \frac{1}{2} \sum_i \left(\alpha_i d_i \cdot H \alpha_i d_i\right) + \sum_i \left(b \cdot \alpha_i d_i\right) = \sum_i f(\alpha_i d_i)$

$$\Rightarrow \min_{x \in \mathbb{R}^n} f(x) = \min_{\alpha_{i=0,\dots,n-1}} \sum_{i=0}^{n-1} f(\alpha_i d_i) = \sum_{i=0}^{n-1} \min_{\alpha_i} f(\alpha_i d_i)$$

Soit $\alpha_i^*$ solution de $\min_{\alpha_i} f(\alpha_i d_i)$ alors $x^* = \sum_i \alpha_i^* d_i$ est solution de (P).

2) algo : $k = 0, x^{(0)}$ $\left[\begin{array}{l} \Rightarrow \text{minimiser } f(x^{(k)} + \alpha d^{(k)}) \text{ ; soit } \alpha_k^* \text{ solution.} \\ x^{(k+1)} = x^{(k)} + \alpha_k^* d_k \end{array}\right.$

itération $k=0$ $\quad x^{(0)} = 0 \Rightarrow$ on minimise $f(\alpha d_0) \xrightarrow{\text{solution}} \alpha_0^*$

itération $k$ $\quad$ on minimise $f(x^{(k)} + \alpha d_k) = \frac{1}{2}(x^{(k)} + \alpha d_k) \cdot H(x^{(k)} + \alpha d_k) + b \cdot (x^{(k)} + \alpha d_k)$

$\left.\begin{array}{l} x^{(k)} = \sum_{i=0}^{k-1} \alpha_i d_i \\ + \text{ directions conjuguées} \end{array}\right\} \Rightarrow \quad = \frac{1}{2} x^{(k)} \cdot H x^{(k)} + \frac{1}{2}(\alpha d_k) \cdot H(\alpha d_k) + b \cdot x^{(k)} + b \cdot (\alpha d_k)$

$$= \underbrace{f(x^{(k)})}_{\text{constant}} + f(\alpha d_k)$$

constant $\Rightarrow$ revient à $\underset{\alpha}{\text{minimiser}} f(\alpha d_k) \xrightarrow{\text{solution}} \alpha_k^*$

donc $x^{(n)} = \alpha_0^* d_0 + \dots + \alpha_{n-1}^* d_{n-1} = \alpha_0^* d_0 + \dots + \alpha_{n-1}^* d_{n-1} = x^*$

3) Les 2 méthodes sont équivalentes $\qquad f = 4x_1^2 + 4x_2^2 - 4x_1 x_2 - 12 x_2$

$$d_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \rightarrow \begin{cases} x_1 = 1 \\ x_2 = 0 \end{cases} \rightarrow f(\alpha_0 d_0) = 4\alpha_0^2 \rightarrow \text{minimum en } \alpha_0^* = 0$$

$$d_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \rightarrow \begin{cases} x_1 = 1 \\ x_2 = 2 \end{cases} \rightarrow f(\alpha_1 d_1) = 4\alpha_1^2 + 4\alpha_1^2 4 - 4\alpha_1^2 2 - 12\alpha_1 2$$

$$= 12\alpha_1^2 - 24\alpha_1 \Rightarrow f' = 24\alpha_1^* - 24 = 0 \rightarrow \alpha_1^* = 1$$

$$x^* = \alpha_0^* \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_1^* \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

vérification : $$\underbrace{\begin{bmatrix} 8 & -4 \\ -4 & 8 \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{x^*} + \underbrace{\begin{bmatrix} 0 \\ -12 \end{bmatrix}}_{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$H \; x^* + b = 0$$

# Méthode des coordonnées.

a) <u>méthode cyclique</u>

- minimiser $f(u_1,u_2) = u_1^2 + u_2^2$ en partant de $(u_1,u_2)$ quelconque.
  - minimiser selon la coordonnée $u_1$

$$\frac{\partial f}{\partial u_1} = 2u_1 = 0 \Rightarrow u_1^* = 0$$

nouveau point $(0, u_2)$

  - minimiser selon la coordonnée $u_2$

$$\frac{\partial f}{\partial u_2}(0,u_2) = 2u_2 = 0 \Rightarrow u_2^* = 0 \Rightarrow \text{nouveau point} = (0,0)$$

On converge en 2 itérations : c'est normal car on retrouve ici la méthode des directions conjuguées
En effet $f(u_1,u_2) = (u_1\ u_2)\, I \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ où $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

et les directions $d^{(0)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $d^{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sont conjuguées par rapport à $I$.

- minimiser $f(u_1,u_2) = u_1^2 + u_2^2 + u_1u_2$ en partant de $(u_1^{(0)}, u_2^{(0)})$ quelconque
établir la suite $U^{(2k)}$ des points obtenus.
  - minimiser selon la variable $u_1$

$$\frac{\partial f}{\partial u_1} = 2u_1 + u_2^{(0)} = 0 \qquad u_1^* = -\frac{u_2^{(0)}}{2} \Rightarrow \text{nouveau point } (u_1^*, u_2^{(0)})$$

  - minimiser selon la variable $u_2$

$$\frac{\partial f}{\partial u_2}(u_1^*, u_2) = 2u_2 + u_1^* = 0 \Rightarrow u_2^* = -\frac{u_1^*}{2} = \frac{u_2^{(0)}}{4} \Rightarrow \text{nouveau point } (u_1^*, u_2^*) = \left(-\frac{u_2^{(0)}}{2}, \frac{u_2^{(0)}}{4}\right)$$

après 1er cycle : $(u_1^{(2)}, u_2^{(2)}) = \left(-\frac{u_2^{(0)}}{2}, \frac{u_2^{(0)}}{4}\right)$

après 2e cycle : $(u_1^{(4)}, u_2^{(4)}) = \left(-\frac{u_2^{(2)}}{2}, \frac{u_2^{(2)}}{4}\right) = \left(-\frac{u_2^{(0)}}{8}, \frac{u_2^{(0)}}{16}\right)$

après 3e cycle : $(u_1^{(6)}, u_2^{(6)}) = \left(-\frac{u_2^{(4)}}{2}, \frac{u_2^{(4)}}{4}\right) = -\frac{u_2^{(0)}}{32}, \frac{u_2^{(0)}}{64}$

après $n$e cycle : $(u_1^{(2n)}, u_2^{(2n)}) = \left(-\frac{u_2^{(0)}}{2\times 4^{n-1}}, \frac{u_2^{(0)}}{4^n}\right) = \frac{1}{4^n} \times \left(-2u_2^{(0)}, u_2^{(0)}\right) \xrightarrow{n} 0$

cette suite est ind$^t$ de $u_1$. En particulier si on démarre avec $u_1^{(0)} = 0$ cela ne change rien.

b) méthode de Gauss-Southwell

- minimiser $f(u_1,u_2) = u_1^2 + u_2^2 + u_1 u_2$ en partant de $(0, u_2)$

- choix de la coordonnée selon laquelle on va minimiser

$$\left|\frac{\partial f}{\partial u_1}(0,u_2)\right| = |u_2| \qquad \left|\frac{\partial f}{\partial u_2}(0,u_2)\right| = |2u_2|$$

on choisit 2.

- minimiser selon $u_2$

$$\frac{\partial f^{(0,u_2)}}{\partial u_2} = 2u_2 = 0 \Rightarrow u_2^* = 0 \Rightarrow \text{nouveau point } (0,0).$$

$$(u_1\ u_2)\underbrace{\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}}_{A}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$\Delta_1 = 1$, $\Delta_2 = 1 - \frac{1}{4} > 0$ déf positive $\Rightarrow$ minimum $= (0, 0)$

valeur propre : $\begin{vmatrix} 1-\lambda & 1/2 \\ 1/2 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - \frac{1}{4} = 0 \Rightarrow 1-\lambda = \pm\frac{1}{2} \Rightarrow \lambda = \frac{3}{2}, \frac{1}{2}$.

$$E^k \left(1 - \frac{1/2}{3/2}\right) \geq E^{k+1}$$

$$\frac{2}{3}E^k \geq E^{k+1}$$

optimisation sans contraintes - Méthode itérative

min. $f(x,y) = 4x^2 - 4xy + 2y^2$

gradient conjugué en partant de $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

$$\nabla f = \begin{bmatrix} 8x - 4y \\ -4x + 4y \end{bmatrix}$$

itération 1

$\nabla f(2,3) = \begin{bmatrix} 4 \\ 4 \end{bmatrix} \Rightarrow d^{(0)} = -\begin{bmatrix} 4 \\ 4 \end{bmatrix}$ dérivée directionnelle $= 0 \rightarrow \nabla f\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix} - t\begin{bmatrix} 4 \\ 4 \end{bmatrix}\right) \cdot \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 0$

$\Rightarrow 4x = 0 \rightarrow 2 - 4t = 0 \Rightarrow t = \frac{1}{2}$

$x^{(1)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $\quad \nabla f(0,1) = \begin{bmatrix} -4 \\ 4 \end{bmatrix}$ $\quad \beta = \dfrac{\|\nabla f(x^{(1)})\|^2}{\|\nabla f(x^{(0)})\|^2} = 1$

$$d^{(1)} = \underbrace{-\begin{bmatrix} -4 \\ 4 \end{bmatrix}}_{-\nabla f(x^{(1)})} \underbrace{- \begin{bmatrix} 4 \\ 4 \end{bmatrix}}_{+\beta d^{(0)}} = \begin{bmatrix} 0 \\ -8 \end{bmatrix}$$

itération 2

dérivée directionnelle $= 0 \rightarrow \nabla f\left[\begin{bmatrix} 0 \\ 1 \end{bmatrix} - t\begin{bmatrix} 0 \\ 8 \end{bmatrix}\right] \cdot \begin{bmatrix} 0 \\ 8 \end{bmatrix} = 0 \Rightarrow 1 - 8t = 0 \rightarrow t = \frac{1}{8}$

$$x^{(2)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

# Optimisation Mathématique - 1ère année

le 7/4/95

**durée 2h.**

**Les <u>seuls</u> documents autorisés sont les notes de cours et de TD manuscrites**

**I)** Soit f(x,y,z) de $R^3$ dans R définie par :

$f(x,y,z) = e^x + e^y + e^z + 2e^{-x-y-z}$

1) Montrer que Hf(x,y,z) est défini positif en tout point de $R^3$

2) Déterminer le minimum global de f(x,y,z) sur $R^3$

**II)** Soit f(x) de classe $C^2$ de $R^n$ dans R.

1) Montrer que la matrice $\nabla f(x) \otimes \nabla f(x)$ est toujours semi définie positive

2) Soit $\alpha>0$. Montrer que si $Hf(x) + \alpha(\nabla f(x) \otimes \nabla f(x))$ est semi définie positive alors $e^{\alpha f(x)}$ est convexe sur $R^n$ .

**III)** minimiser $\frac{x}{y} + 2\frac{y}{z} + 4\frac{z}{x} + 12$ sur $(R_*^+)^3$

(indication: utiliser l'inégalité arithmétique-géométrique)

**IV)** Soit $f(x_1,x_2) = 2x_1^2 + x_2^2$ . En partant de $x^{(0)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ , calculer $x^{(1)}$, $x^{(2)}$ les 2 premiers points de la méthode de Cauchy

**V)** Soit $C = \left\{ x \in R^n : \sum_{i=1}^{n} x_i = 0 \text{ et } |x_i| \le 1 \ (i = 1,\ldots,n) \right\}$

1) Montrer que C est convexe.

2) a) Exhiber n points affinement indépendants de C pour les cas n=6 et n=7.

b) Quelle est la dimension de C dans chacun de ces 2 cas ?

7

VI) On considère le problème minimiser $f(x)$ $x \in R^n$ où f est de classe $C^2$.
On s'intéresse à la méthode quasi-Newtonnienne de rang 1 suivante:

- initialisation: soit $x^{(0)}$ un point et $D_0$ une matrice nxn
- itérations: on construit les points et matrices suivantes:

$$k \geq 0 \quad \begin{cases} x^{(k+1)} = x^{(k)} - D_k^{-1} \nabla f\left(x^{(k)}\right) \\ D_{k+1} = D_k + \dfrac{\left(y^{(k)} - D_k d^{(k)}\right) \otimes d^{(k)}}{d^{(k)} \bullet d^{(k)}} \end{cases}$$

$$\text{où} \quad \begin{cases} d^{(k)} = x^{(k+1)} - x^{(k)} \\ y^{(k)} = \nabla f\left(x^{(k+1)}\right) - \nabla f\left(x^{(k)}\right) \end{cases}$$

$(b \otimes a)a = (a \cdot a)b$

1) Montrer que $D_{k+1}$ vérifie la condition de la sécante $D_{k+1} d^{(k)} = y^{(k)}$

2) Soient $l_{k+1}(x) = \nabla f\left(x^{(k+1)}\right) + D_{k+1}\left(x - x^{(k+1)}\right)$ et $l_k(x) = \nabla f\left(x^{(k)}\right) + D_k\left(x - x^{(k)}\right)$, montrer que $l_k(x) = l_{k+1}(x)$ lorsque x est tel que $\left(x - x^{(k)}\right) \bullet d^{(k)} = 0$

3) On considère la norme matricielle définie par $\|A\| = \sup\{\|Ax\| : \|x\| \leq 1\}$ où A est une matrice nxn

a) Si a et b sont 2 vecteurs de $R^n$, exprimer $\|a \otimes b\|$ en fonction de $\|a\|$ et $\|b\|$

On définit la distance entre deux matrices A et B par $d(A,B) = \|A - B\|$ et on rappelle que $\|AB\| \leq \|A\|\|B\|$. $\|Ax\| \leq \|A\| \|x\|$

b) Soit D une matrice qui satisfait la condition de la sécante $D\, d^{(k)} = y^{(k)}$, montrer que $d\left(D_{k+1}, D_k\right) \leq d\left(D, D_k\right)$

4) Montrer que $y^{(k)} - D_k d^{(k)} = \nabla f\left(x^{(k+1)}\right)$

5) Soit $f(x_1, x_2) = 2x_1^2 + x_2^2 - x_1 x_2$ de $R^2$ dans R et $x^{(0)} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$

a) Déterminer a priori le minimum de f en étudiant ses dérivées premières et secondes.

b) i) Calculer les points $x^{(1)}$ et $x^{(2)}$ avec $D_0$=I la matrice identité.
ii) Les directions $-D_k^{-1} \nabla f\left(x^{(k)}\right)$ (k=0 et 1) sont-elles des directions de descente?

iii) Comparer $f\left(x^{(0)}\right)$, $f\left(x^{(1)}\right)$, $f\left(x^{(2)}\right)$ . Le résultat obtenu est-il contradictoire avec le point précédent? Pourquoi ?

c) i) Calculer les points $x^{(1)}$ et $x^{(2)}$ avec $D_0 = Hf\left(x^{(0)}\right)$

ii) Que constate-t-on ? Etait-ce prévisible ?

<u>Barême:</u>

I) 2 points
II) 2 points
III) 2 points
IV) 2 points
V) 4 points : 1) 1 point 2) a) 2 points 2) b) 1 point
VI) 10 points : 1) 1 point 2) 1 point 3) a) 1 point b) 2 points
4) 1 point 5) a) 1 point b) 2 points c) 1 point

I) $\nabla f = \begin{bmatrix} e^x - 2e^{-x-y-z} \\ e^y - \quad " \\ e^z - \quad " \end{bmatrix}$ $\qquad Hf = \underbrace{\begin{bmatrix} e^x & 0 & 0 \\ 0 & e^y & 0 \\ 0 & 0 & e^z \end{bmatrix}}_{A} + \underbrace{2e^{-x-y-z}}_{\alpha>0} \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}}_{B}$

1) $Hf = A + \alpha B \quad \alpha > 0$

A définie positive. $\qquad XBX = (x\ y\ z)\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x+y+z)^2 \geq 0$

donc B semi-définie positive.

2) $\nabla f = 0 \Rightarrow e^x = e^y = e^z = 2e^{-x-y-z} \Rightarrow x = y = z$ et $e^x = 2e^{-3x} \Rightarrow x = \mathrm{Log}\,2 - 3x$

$\Rightarrow x = \frac{1}{4}\mathrm{Log}\,2 = y = z$

II) 1) $y\cdot(\nabla f(x) \otimes \nabla f(x))y = y \cdot \left[(y\cdot\nabla f(x))\nabla f(x)\right] = (y\cdot\nabla f(x))^2 \geq 0.$

2) $g(x) = e^{\alpha f(x)}$

$\frac{\partial g}{\partial x_i} = \alpha e^{\alpha f(x)} \frac{\partial f}{\partial x_i}(x) \qquad \frac{\partial^2 g}{\partial x_i \partial x_j}(x) = \alpha^2 e^{\alpha f(x)} \frac{\partial f}{\partial x_j}\frac{\partial f}{\partial x_i} + \alpha e^{\alpha f(x)} \frac{\partial^2 f}{\partial x_i \partial x_j}$

$= \alpha e^{\alpha f(x)}\left[\alpha \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} + \frac{\partial^2 f}{\partial x_i \partial x_j}\right]$

et $Hg(x) = \underbrace{\alpha e^{\alpha f(x)}}_{>0} \underbrace{\left[\alpha \nabla f(x) \otimes \nabla f(x) + Hf(x)\right]}_{A}$

si A semi-déf. positive $\Rightarrow$ $Hg(x)$ semi-déf positif $\Rightarrow$ g convexe.

III $\frac{x}{y} + \frac{2y}{z} + 4\frac{z}{x} = 3\left(\frac{x}{3y} + \frac{2y}{3z} + \frac{4z}{3x}\right) \overset{AG}{\geq} 3\left(\frac{x}{y} \times \frac{2y}{z} \times 4\frac{z}{x}\right)^{1/3} = 3 \times 8^{1/3}$
$= 6$

minimum quand égalité dans AG : $\frac{x}{y} = \frac{2y}{z} = \frac{4z}{x} = 2$

$\Rightarrow x = 2y \qquad y = z$

$\delta_1 - \delta_3 = 0 \qquad \delta_1 + \delta_2 + \delta_3 = 1$
$-\delta_1 + \delta_2 = 0$
$-\delta_2 + \delta_3 = 0$
$\Rightarrow \delta_i = \frac{1}{3} \qquad v(\delta) = 3^{1/3} \times (2\times 3)^{1/3} \times (4\times 3)^{1/3}$
$= 6$
$\frac{1}{3} = \frac{x}{6y} \quad \frac{1}{3} = \frac{2y}{6z} \quad \frac{1}{3} = 4\frac{z}{6x}$
$\Rightarrow x = 2y,\ y = z$

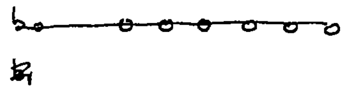
IV $t_1 = \frac{1}{3} \quad x^{(1)} = \begin{pmatrix} -1/3 \\ 2/3 \end{pmatrix} \qquad t_2 = \frac{1}{3} \quad x^{(2)} = \frac{1}{9}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$

V 1) $x$ et $y \in C$ $\quad (1-\lambda)x+\lambda y \in C$ ?

$$\sum_{i=1}^{n} x_i = f(x) \text{ linéaire} \Rightarrow f((1-\lambda)x+\lambda y) = (1-\lambda)f(x)+\lambda f(y) = 0$$

$$|(1-\lambda)x_i+\lambda y_i| \leq (1-\lambda)|x_i|+\lambda|y_i| \leq (1-\lambda)+\lambda = 1$$

2) a) $n = 6$



$$\begin{array}{cccccc} b_0 & b_1 & b_2 & b_3 & b_4 & b_5 \\ \begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix} & \begin{bmatrix}-1\\+1\\0\\0\\0\\0\end{bmatrix} & \begin{bmatrix}-1\\0\\+1\\0\\0\\0\end{bmatrix} & \begin{bmatrix}-1\\0\\0\\+1\\0\\0\end{bmatrix} & \begin{bmatrix}-1\\0\\0\\0\\+1\\0\end{bmatrix} & \begin{bmatrix}-1\\0\\0\\0\\0\\+1\end{bmatrix} \end{array}$$

$b_1-b_0, b_2-b_0, \dots, b_5-b_0$ sont lin$^t$ ind$^t$

$n = 7$ identique

b) $C$ contient un $(n-1)$-simplexe donc $\dim C \geq n-1$

mais $C \subset H = \{x : \sum_{i=1}^{n} x_i = 0\}$ $\quad H$ espace affine : hyperplan.

donc $\mathrm{Aff}\, C \subset H$

et $\dim C \leq n-1$

VI 1)
$$D_{k+1} d^{(k)} = D_k d^{(k)} + \frac{d^{(k)}\cdot d^{(k)}\left(y^{(k)} - D_k d^{(k)}\right)}{d^{(k)}\cdot d^{(k)}} = y^{(k)}$$

2)
$$\ell_{k+1}(x) = \nabla f(x^{(k+1)}) + D_{k+1}(x - x^{(k+1)})$$
$$= \nabla f(x^{k+1}) + \left[D_k + \frac{(y^k - D_k d^k)\otimes d^k}{d^k\cdot d^k}\right]\left(x - x^{k} + \underbrace{x^k - x^{k+1}}_{-d^k}\right)$$
$$= \quad " \quad + D_k(x - x^{k+1}) - \frac{d^k \cancel{d^k}\left(y^k - D_k d^k\right)}{d^k\cdot\cancel{d^k}} \qquad \text{car } d^k\cdot(x - x^k) = 0$$
$$= \nabla f(x^k) + D_k(x - x^k) = \ell_k(x)$$

interprétation : approximation du gradient à l'ordre 1 à l'itération $k = \ell_k(x)$
si on se déplace dans une direction $\perp (x^{k+1} - x^k)$ l'approximation du gradient à l'itération $k+1$ est la même qu'à l'itération $k$.

$a\otimes b)x = ab^t x = (b\cdot x)a$

3) a) $\|a\otimes b\| = \sup\{\|a\otimes b\, x\| : \|x\|\leq 1\}$

$a\otimes b\, x = (b\cdot x)\, a \Rightarrow \|a\otimes b\, x\| = \|a\|\,|b\cdot x|$

$|b\cdot x| \leq \|b\|\underbrace{\|x\|}_{1}$ et quand $x$ colinéaire à $b$ et de norme 1 on $|b\cdot x| = \|b\|\,\|x\| = \|b\|$

donc $\|a\otimes b\| = \|a\|\,\|b\|$

b) $\|D_{k+1} - D_k\| = \left\| \dfrac{(y^k - D_k d^k)\otimes d^k}{d^k\cdot d^k} \right\| = \dfrac{1}{d^k\cdot d^k}\|(y^k - D_k d^k)\otimes d^k\|$

$\|Ax\| \leq \|A\|\|x\|$

produit de matrice est associatif

$= \dfrac{1}{d^k\cdot d^k}\|(D d^k - D_k d^k)\otimes d^k\| = \dfrac{1}{d^k\cdot d^k}\|(D - D_k)d^k\otimes d^k\|$

$\dfrac{1}{d^k\cdot d^k}\|Dd^k - D_k d^k\|\,\|d^k\|$

$\dfrac{1}{d^k\cdot d^k}\|(D-D_k)d^k\|\,\|d^k\|$

$\dfrac{1}{d^k\cdot d^k}\|D - D_k\|\,\|d^k\|\,\|d^k\|$

$\leq \dfrac{1}{d^k\cdot d^k}\|D - D_k\|\,\|d^k\otimes d^k\| = \|D - D_k\|\,\dfrac{\|d^k\|^2}{d^k\cdot d^k}$

4) $d^k = -D_k^{-1}\nabla f(x^{(k)}) \Rightarrow D_k d^k = -\nabla f(x^{(k)})$

$\Rightarrow y^k - D_k d^k = \nabla f(x^{(k+1)})$

5) a) $\nabla f = \begin{pmatrix} 4x_1 - x_2 \\ 2x_2 - x_1 \end{pmatrix}$ $\quad Hf = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$

$Hf$ déf positif. $\Rightarrow f$ convexe strict $\Rightarrow$ un minimum global qui est le point critique $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

b) $D_0 = I \quad x^0 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$

$\rightarrow x^{(1)} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} - I\begin{pmatrix} 0 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$ $\quad \nabla f(x^1) = \begin{bmatrix} 7 \\ -7 \end{bmatrix}$ $\quad \nabla f(x^1)\otimes d^0 = \begin{bmatrix} 7 \\ -7 \end{bmatrix}\begin{bmatrix} 0 & -7 \end{bmatrix} = \begin{bmatrix} 0 & -49 \\ 0 & 49 \end{bmatrix}$

$d^0\cdot d^0 = (0\ \ -7)\begin{pmatrix} 0 \\ -7 \end{pmatrix} = 49$

$D_1 = D_0 + \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$

$\rightarrow x^{(2)} = \begin{pmatrix} 1 \\ -3 \end{pmatrix} - D_1^{-1}\begin{bmatrix} 7 \\ -7 \end{bmatrix} = \begin{pmatrix} 1 \\ -3 \end{pmatrix} - \begin{bmatrix} 1 & 1/2 \\ 0 & 1/2 \end{bmatrix}\begin{bmatrix} 7 \\ -7 \end{bmatrix} = \begin{bmatrix} -5/2 \\ 1/2 \end{bmatrix}$

ii) $D_0$ et $D_1$ sont définies positives $\Rightarrow$ $-D_0\nabla f(x^0)$ et $-D_1\nabla f(x^{(1)})$ sont directions de descente.

Rappel cours $\longrightarrow$ $f_t'(x^0 - tD^0\nabla f(x^0)) = \nabla f(x^0 - tD_0\nabla f(x_0))\cdot(-D^0\nabla f(x^0))$

en $t=0$ $\quad f_t'(x^0) = -\nabla f(x^0)\cdot D^0\nabla f(x^0) < 0$

iii) $f(x^{(0)}) = f(x^{(1)}) = f(x^{(2)}) = 14$

ceci n'est pas contradictoire car dans l'algo. on se déplace dans la direction de descente avec un pas de 1 et il se peut qu'on se retrouve sur une courbe de même niveau. Par contre si on minimisait $f$ dans la direction de descente alors dans ce cas il y aurait une stricte décroissance $f(x^{(0)}) > f(x^{(1)}) > f(x^{(2)})$

c) i) $D_0 = Hf(x^0)$ $\quad D_0^{-1} = \frac{1}{7}\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$

$$x^{(1)} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} - \begin{bmatrix} 2/7 & 1/7 \\ 1/7 & 4/7 \end{bmatrix}\begin{bmatrix} 0 \\ 7 \end{bmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ii) on arrive sur le minimum en 1 étape.
c'est normal : avec $D_0 = Hf(x^0)$ on a, à la première itération, la méthode de Newton qui converge en une étape dans le cas quadratique strict. convexe. (cf. théorème cours)

# Méthode de Newton

Soit $f$ définie de $\mathbb{R}^n$ dans $\mathbb{R}$ par $f(x) = a + b \cdot x + \frac{1}{2} x \cdot A x$.

où $a \in \mathbb{R}$, $b \in \mathbb{R}^n$, $A$ matrice réelle $n \times n$ définie positive et symétrique.

Montrer que la méthode de Newton permet d'atteindre le minimum de $f(x)$ en une étape quelque soit le point $x_0$ initial.

dém :

$$\nabla f(x) = b + Ax$$

$$H_f = A$$

$$\Rightarrow \quad x_1 = x_0 - H_f^{-1} \nabla_f(x_0)$$

$$= x_0 - A^{-1}(b + A x_0)$$

A est inversible car définie positive $\Rightarrow \Delta_n > 0$.

$$= \cancel{x_0} - \cancel{x_0} - A^{-1}(b)$$

$$= -A^{-1}(-A x^*)$$

$$= x^*$$

Le minimum $x^*$ de $f$ est unique et global car $f$ strict convexe. Il vérifie $\nabla_f(x^*) = 0$ i.e $b + Ax^* = 0$

# Steepest descent méthod.

## Exercice 12.

Soit $f(x)$ une fonction quadratique de $n$ variables : $f(x) = a + b\cdot X + \frac{1}{2} x\cdot Ax$

où $a \in \mathbb{R}$, $b \in \mathbb{R}^n$ et $A$ est une $\underbrace{\text{matrice}}_{n\times n,}$ symétrique, et définie positive.

a/ Montrer que $f(x)$ admet un minimum global unique.

b/ Montrer que si le point initial $X^0$ est choisi t.q. $X^0 - X^*$ est un vecteur propre de $A$, alors la suite de la "Steepest Descent" $\{x^k\}$ avec pour point initial $X^0$ atteint $X^*$ (le point minimum) en une étape, cad $X^1 = X^*$.

a/ $\nabla f = b + Ax$ voir exercice 3 sur fonction quadratique convexe.

$Hf = A$

$A$ définie positive $\Rightarrow$ il y a un minimum strict $x^*$ (strict $\Rightarrow$ unique)

b/ $X' = X^0 - t \times (b + AX^0)$ où $t$ minimise la fonction $f(t)$ où $f(t)$ est la fonction $f(X^0 - t\times(b + AX^0))$ cad la fonction $f$ astreinte à varier dans la direction $-\nabla f_{(X^0)} = -(b + AX^0)$ à partir du point $X^0$.

$f(t) = f(X^0 - t \times (b + AX^0))$

$f'(t) = \nabla f(X^0 - t\times(b + AX^0)) \times (-(b + AX^0)) = 0$ dérivée nulle au point minimum. (point critique)

$\Rightarrow (b + A(X^0 - t(b + AX^0))) \times (b + AX^0) = 0$ d'autre part $\nabla f(X^*) = 0 = b + AX^* \Rightarrow b = -AX^*$

$\Rightarrow (A(X^0 - X^* - t(b + AX^0))) \times A(X^0 - X^*) = 0$

$\Rightarrow$ ———————— $\times \cancel{\lambda}(X^0 - X^*) = 0$ on peut simplifier par $\lambda$ qui est $\neq 0$ car $A$ définie positive $\Rightarrow$ toutes les valeurs propres sont $> 0$.

$\Rightarrow (X^0 - X^*)^T A(X^0 - X^*) - t(X^0 - X^*)^T A(b + AX^0) = 0$

$>0 \qquad - t(X^0 - X^*)^T A(A(X^0 - X^*)) = 0$

$\Rightarrow \quad >0 \qquad -\lambda t \underbrace{(X^0 - X^*)^T A(X^0 - X^*)}_{>0 \text{ car } A \text{ définie positive et } X^0 - X^* \neq 0 \text{ car vecteur propre}} = 0$

$\Rightarrow \quad 1 - \lambda t = 0 \Rightarrow t = 1/\lambda$ on a vu que $\lambda \neq 0$.

$X' = X^0 - \frac{1}{\lambda}(b + AX^0) = X^0 - \frac{1}{\lambda} A(X^0 - X^*) = X^0 - \frac{\lambda}{\lambda}(X^0 - X^*) = X^*$

# Controle MAC

Durée 1 h 45
Sans document

I - Pour chacune de ces fonctions,déterminer la nature des points critiques;c'est-à-dire si ce sont des minimums ou maximums locaux ou globaux.

3 1°) f définie sur $R^2$ par $f(x_1,x_2)= x_1^3 + x_2^3 - 3x_1 - 12x_2 + 20$

2 2°) g définie sur $R^3$ par $g(x_1,x_2,x_3) = x_1^2 + x_2^2 + x_3^2 - 4x_1 x_2$

2 II - Soit f définie sur R par $f(x) = x^4$. En utilisant la méthode de Newton,trouver le minimum de f.

III - Soit f la fonction définie sur $R^2$ par $f(x_1,x_2)= 2x_1^2 + x_2^2 - x_1 x_2$

5 En partant de $x_0 = (1,4)$,calculer les deux premiers termes $x_1$ et $x_2$ de la suite de descente de Cauchy. Vérifier que $x_2 - x_1$ est orthogonal à $x_1 - x_0$

2 IV - 1°) Trouver le vecteur du plan d'équation $x + 2y + 3z = 6$ le plus proche de l'origine.

2°) Trouver le vecteur de norme minimale satisfaisant à

3 $$\left|\begin{array}{l} x_1 + x_2 + 5x_3 - 7x_4 = 1 \\ x_1 - 3x_2 - x_3 + x_4 = 2 \end{array}\right.$$

V - Soit f une fonction convexe de R dans R. Soit $a \in R^n$ ,$\alpha \in R$ fixés. On définit la fonction g de $R^n$ dans R par $g(x) = f(a.x + \alpha)$ pour tout $x \in R^n$.

2 Montrer que g est convexe. En déduire que la fonction h de $R^3$ dans R définie par

1 $h(x,y,z) = (4x + 5y - 8z + 7)^8$ est convexe.

VI - Montrer l'inégalité $\frac{x}{4} + \frac{3y}{4} \leq \sqrt{\ln\left(\frac{e^{x^2}}{4} + \frac{3e^{y^2}}{4}\right)}$ pour $x,y \geq 0$, en utilisant la convexité d'une

3 fonction bien choisie.

VII - En utilisant l'inégalité arithmético-géométrique (inégalité A-G), résoudre le problème:

5 Minimiser $x^2 + y + z$
sous la contrainte $xyz = 1$ et $x,y,z>0$

VIII - Soit C un convexe fermé de $R^n$. Si $x \in R^n$ et $x^* \in R^n$,montrer que $x^* = P_C(x)$ ssi

2 $(x^* - x).(y - x) \geq \| x^* - x \|^2$ pour tout $y \in C$

IX - Soit f une fonction convexe de $R^n$ dans R admettant des dérivées partielles continues. Soit

3 $x^* \in R^n$.

Montrer que $x^*$ est un minimum global de f ssi $\nabla f(x^*).(x - x^*) \geq 0$ pour tout $x \in R^n$.

(1)

Correction

I - 1) $f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1 - 12x_2 + 20$

$$\nabla f(x_1, x_2) = \begin{pmatrix} 3x_1^2 - 3 \\ 3x_2^2 - 12 \end{pmatrix}$$

pas de minimum global ni maximum global.

$f(0, x_2) \underset{x_2 \to \infty}{\longrightarrow} +\infty$ et $f(0, x_2) \longrightarrow -\infty$ si $x_2 \to -\infty$

points critiques $x_1 = \pm 1$ $x_2 = \pm 2$

$$Hf(x_1, x_2) = \begin{pmatrix} 6x_1 & 0 \\ 0 & 6x_2 \end{pmatrix}$$

- pour $(1, 2)$ on a $Hf(1,2) = \begin{pmatrix} 6 & 0 \\ 0 & 12 \end{pmatrix}$ définie positive

Donc $(1, 2)$ minimum local.

Or $f(1, x_2) = x_2^3 - 12x_2 + 18$

Si $x_2 \to -\infty$ $f(1, x_2) \to -\infty$

donc $(1, 2)$ n'est pas minimum global

- $(-1, 2)$ $Hf(-1,2) = \begin{pmatrix} -6 & 0 \\ 0 & 12 \end{pmatrix}$ indéfinie ~~négative~~.

$\Rightarrow$ point selle

- $(1, -2)$ $Hf(1,-2) = \begin{pmatrix} 6 & 0 \\ 0 & -12 \end{pmatrix}$. indéfinie

$\Rightarrow$ point selle

- $(-1, -2)$ $Hf(-1,-2) = \begin{pmatrix} -6 & 0 \\ 0 & -12 \end{pmatrix}$ définie négative

Donc $(-1, -2)$ maximum local. Or $f(-1, x_2) \to +\infty$ donc pas maximum global qd $x_2 \to +\infty$.

2) $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 4x_1x_2$ ②

$$Df(x_1,x_2,x_3) = \begin{pmatrix} 2x_1 - 4x_2 \\ 2x_2 - 4x_1 \\ 2x_3 \end{pmatrix} \qquad Hf(x_1,x_2,x_3) = \begin{pmatrix} 2 & -4 & 0 \\ -4 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Unique point critique $(0,0,0)$

$Hf(0,0,0)$ indéfini.

$\Rightarrow$ point selle

recherche valeurs propres :

$$\begin{vmatrix} 2-\lambda & -4 & 0 \\ -4 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda)[(2-\lambda)^2-16] = (2-\lambda)(6-\lambda)(-2-\lambda)$$

$\lambda_1 = 2,\ \lambda_2 = 6,\ \lambda_3 = -2$

$\Rightarrow$ indéfini.

II .. Si $x_0$ qcq, $x_0 \neq 0$.

~~$Hf(x,y) = 2\,Id\ .(x_0,y_0)$~~

~~$(x,y) = (x,y)$~~

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)} = x_0 - \frac{3x_0^2}{6x_0} = \frac{1}{2}x_0$$

$$x_0 - \frac{4x^3}{12x^2} = x_0 - \frac{1}{3}x_0 = \frac{2}{3}x_0$$

Donc $x_k = \left(\frac{2}{3}\right)^n x_0 \longrightarrow 0$.

III $f(x_1,x_2) = 2x_1^2 + x_2^2 - x_1x_2$ 1.

$x_0 = (1,4)$

$$\nabla f(x_1,x_2) = \begin{pmatrix} 4x_1 - x_2 \\ 2x_2 - x_1 \end{pmatrix} = \begin{pmatrix} 4x_1 - x_2 \\ -x_1 + 2x_2 \end{pmatrix}.$$

$$\nabla f(1,4) = \begin{pmatrix} 0 \\ 7 \end{pmatrix} \qquad x_0 - t\nabla f(x_0) = \begin{pmatrix} 1 \\ 4-7t \end{pmatrix}$$

$f_0(t) = 2 + (4-7t)^2 - 4 + 7t = 14 - 49t + 49t^2$

$\varphi_0'(t) = -49(1-2t)$. $\rightarrow$ $t = \frac{1}{2}$.

$\varphi_0''(\frac{1}{2}) > 0 \rightarrow \frac{1}{2}$ minimum de $\varphi$

$f_1'(t) = \nabla f(\vec{x}_1 + t\vec{u}) \cdot \vec{u}$

$$\text{Donc } x_1 = \begin{pmatrix} 1 \\ 4 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 0 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}.$$

$-\nabla f(1, 1/2) = -\begin{pmatrix} 7/2 \\ 0 \end{pmatrix}$

$$f_1'(t) = \nabla f \cdot \begin{pmatrix} 7/2 \\ 0 \end{pmatrix} = \begin{pmatrix} 4(1 - \frac{7t}{2}) - (1/2) \\ -(1 - 7t/2) + 1 \end{pmatrix} \cdot \begin{pmatrix} 7/2 \\ 0 \end{pmatrix} = -\frac{49}{4} + 49t$$

$\varphi_1'(t) = -\left(\frac{49}{4} - 49t\right)$

D'où $t = \frac{1}{4}$ $\quad \varphi_1''(\frac{1}{4}) > 0 \Rightarrow t = \frac{1}{4}$

$x_2 = x_1 - \frac{1}{4} \nabla f(x_1)$

$$= \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} - \frac{1}{4} \begin{pmatrix} 7/2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/8 \\ 1/2 \end{pmatrix}.$$

IV

1) $A = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$ d'où $A\,{}^tA = 14$.

$A\,{}^tA\, w = b \Rightarrow w = \frac{3}{7}$

D'où $x^* = \frac{3}{7} \times \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3/7 \\ 6/7 \\ 9/7 \end{pmatrix}$

2)

$$A = \begin{pmatrix} 1 & 1 & 5 & -7 \\ 1 & -3 & -1 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

$$A\,{}^tA = \begin{pmatrix} 76 & -14 \\ -14 & 12 \end{pmatrix}$$

$$w = (w_1, w_2) = \left( \frac{44}{5}, \frac{313}{30} \right)$$

D'où $x^* = \frac{44}{5} \times \begin{pmatrix} 1 \\ 1 \\ 5 \\ -7 \end{pmatrix} + \frac{313}{30} \begin{pmatrix} 1 \\ -3 \\ -1 \\ 1 \end{pmatrix}$

$$x^* = \begin{pmatrix} \frac{577}{30} \\ -\frac{675}{30} \\ \frac{1007}{30} \\ -\frac{1535}{30} \end{pmatrix}$$

V Si $x^* = P_C(x)$

$$(x^* - x \mid x^* - y) \leq 0 \quad \forall y \in C$$

D'où $(x^* - x \mid y - x) = (x^* - x \mid y - x^* + x^* - x)$

$$= (x^* - x \mid y - x^*) + \|x^* - x\|^2$$

Donc évident.

VI On a $f(x) \geq f(x^*) + \nabla f(x^*) \cdot (x - x^*)$.

$$\Rightarrow \quad f(x) - f(x^*) \geq \nabla f(x^*) \cdot (x - x^*) \quad \forall x$$

Si $\nabla f(x^*) \cdot (x - x^*) \geq 0 \Rightarrow x^*$ minimum global.

Si $x^*$ minimum global $\Rightarrow \nabla f(x^*) = 0 \Rightarrow \nabla f(x^*) \cdot (x - x^*) = 0$

VII Minimiser $x^2 + y + z$ s.c. $xyz = 1$ $x, y, z > 0$

On a

$$x^2 + y + z = 5 \left( \frac{x^2 + \frac{y}{2} + \frac{y}{2} + \frac{z}{2} + \frac{z}{2}}{5} \right).$$

$$\underset{A.G.}{\geq} 5 \, (x^2)^{1/5} \left(\frac{y}{2}\right)^{1/5} \left(\frac{y}{2}\right)^{1/5} \left(\frac{z}{2}\right)^{1/5} \left(\frac{z}{2}\right)^{1/5}$$

$$= 5 \left(\frac{1}{2}\right)^{4/5} x^{2/5} y^{2/5} z^{2/5} = 5 \left(\frac{1}{2}\right)^{4/5}$$

Donc égalité $x^2 = \frac{y}{2} = \frac{z}{2} = \left(\frac{1}{2}\right)^{4/5}$

$$\Rightarrow y = z = 2^{1/5} \qquad x = \left(\frac{1}{2}\right)^{2/5}$$

VIII

$$g((1-\lambda)x_1+\lambda x_2) = f(\cancel{(1-\lambda)}\, a.x_1 + \lambda\, a.x_2 + (1-\lambda)\alpha + \lambda\alpha)$$

$$\leq (1-\lambda)\, f(a.x_1+\alpha) + \lambda\, f(a.x_2+\alpha)$$

$$= (1-\lambda)\, g(x_1) + \lambda\, g(x_2).$$

$$h(x,y,z) = f((4,5,-8)\cdot(x,y,z) + 7).$$

$f(z) = z^8$ convexe.

IX Comme $x, y \geq 0$ on a

$$\frac{x}{4} + \frac{3y}{4} \leq \sqrt{\ln\left(\frac{e^{x^2}}{4} + \frac{3}{4}e^{y^2}\right)}$$ ssi

$$\left(\frac{x}{4} + \frac{3y}{4}\right)^2 \leq \ln\left(\frac{e^{x^2}}{4} + \frac{3}{4}e^{y^2}\right)$$

ssi $$e^{\left(\frac{x}{4}+\frac{3y}{4}\right)^2} \leq \frac{e^{x^2}}{4} + \frac{3}{4}e^{y^2}$$

d'où convexité de $e^{z^2}$

**Contrôle MOM (Optimisation Mathématique)**

12 Avril 96. Durée=2H.

Notes de cours autorisées

Exercice 1

a) Résoudre le problème:

minimiser $\frac{x_1x_2x_3}{2}+\frac{1}{x_1}+\frac{2}{x_2}+\frac{16}{x_3}$ sous les contraintes $x_1>0$, $x_2>0$, $x_3>0$

b) Résoudre le problème:

minimiser $\frac{1}{2}\left(xyz-xz^2-zy^2+yz^2\right)+\frac{1}{x-y}+\frac{2}{y-z}+\frac{16}{z}$ sous les contraintes $x>y>z>0$

Exercice 2

Soit $Q$ une matrice symétrique, définie positive, soient $p^{(0)}, p^{(1)},\dots,p^{(n-1)}$ $n$ vecteurs linéairement indépendants de $R^n$.

1) Montrer par récurrence que $d^{(0}, d^{(1)},\dots,d^{(n-1)}$ définis par

$$\begin{cases} d^{(0)}=p^{(0)} \\ d^{(k+1)}=p^{(k+1)}-\sum_{i=0}^{k}\frac{p^{(k+1)}\bullet Qd^{(i)}}{d^{(i)}\bullet Qd^{(i)}}d^{(i)} \end{cases}$$

sont des directions conjuguées 2 à 2 par rapport à $Q$ c'est-à-dire $d^{(i)}\neq 0$, $i=0,\dots,n-1$ et $d^{(i)}\bullet Qd^{(j)}=0$, $0\leq i<j\leq n-1$.

2) Soit $f(x_1,x_2)=2x_1^2+x_2^2$ de $R^2$ dans $R$.

a) Déterminer $Q$ symétrique telle que $f(x)=\frac{1}{2}x\bullet Qx$ où $x=\begin{pmatrix}x_1\\x_2\end{pmatrix}$.

b) Soient $p^{(0)}=\begin{pmatrix}1\\0\end{pmatrix}$, $p^{(1)}=\begin{pmatrix}1\\1\end{pmatrix}$. Calculer $d^{(0)}$, $d^{(1)}$.

c) Soit $x^{(k+1)}=x^{(k)}+\alpha^{(k)}d^{(k)}$ , $k\geq 0$ , où $\alpha^{(k)}$ minimise $f(x^{(k)}+\alpha d^{(k)})$, $\alpha$ étant de signe quelconque.

Calculer $x^{(1)}, x^{(2)}$ en partant de $x^{(0)}=\begin{pmatrix}1\\1\end{pmatrix}$.

d) Même question que c) en utilisant cette fois les directions $p^{(k)}$ à la place de $d^{(k)}$.

## Exercice 3

Soient un ensemble $d^{(k)}$ de vecteurs non nuls de $R^n$. Soit $f(x)=\frac{1}{2}x \bullet Ax - b \bullet x$ de $R^n$ dans $R$, où $A$ est symétrique, définie positive.

Soit l'algorithme itératif: $x^{(k+1)}=x^{(k)}+\alpha^{(k)}d^{(k)}$, où $\alpha^{(k)}$ minimise $f(x^{(k)}+\alpha d^{(k)})$, $\alpha$ étant de signe quelconque.

On note $x^*$ le minimum de $f(x)$, $x \in R^n$ et $E(x)=f(x)-f(x^*)$. On note $g^{(k)}=\nabla f\left(x^{(k)}\right)$ le gradient de $f$ au point $x^{(k)}$.

1) Donner la condition nécessaire que doit vérifier $x^*$.

2) Montrer que:

a) $\alpha^{(k)}=-\dfrac{g^{(k)} \bullet d^{(k)}}{d^{(k)} \bullet Ad^{(k)}}$

b) $E(x)=\frac{1}{2}(x-x^*) \bullet A(x-x^*)$

c) $\dfrac{E\left(x^{(k)}\right)-E\left(x^{(k+1)}\right)}{E\left(x^{(k)}\right)}=\dfrac{\left(g^{(k)} \bullet d^{(k)}\right)^2}{\left(d^{(k)} \bullet Ad^{(k)}\right)\left(g^{(k)} \bullet A^{-1}g^{(k)}\right)}$

3) Supposons que $g^{(k)} \bullet d^{(k)}<0$ et $\left(g^{(k)} \bullet d^{(k)}\right)^2 \geq \beta\left(d^{(k)} \bullet Ad^{(k)}\right)\left(g^{(k)} \bullet A^{-1}g^{(k)}\right)$ où $0<\beta \leq 1$.

a) $d^{(k)}$ est-elle une direction de descente pour $f$?

b) Pour tout $k$, montrer que $\dfrac{E\left(x^{(k+1)}\right)}{E\left(x^{(k)}\right)} \leq 1-\beta$.

4) Supposons que $g^{(k)} \bullet d^{(k)}<0$ et $\left(g^{(k)} \bullet d^{(k)}\right)^2 \geq \beta\left(d^{(k)} \bullet d^{(k)}\right)\left(g^{(k)} \bullet g^{(k)}\right)$ où $0<\beta \leq 1$.

a) Pour tout $k$, donner une majoration du rapport $\dfrac{E\left(x^{(k+1)}\right)}{E\left(x^{(k)}\right)}$ en fonction de $\beta$ et des valeurs propres de $A$.

b) Des questions 3 et 4, quelles $d^{(k)}$ semblent les meilleures pour la convergence des $x^{(k)}$ vers $x^*$?

5) On s'intéresse au cas de la question 4.
Montrer qu'il est possible d'avoir $\beta$=1. Dans ce cas, quelle méthode retrouve-t-on?

6) On s'intéresse au cas de la question 3.
Montrer qu'il est possible d'avoir $\beta$=1. Dans ce cas, en combien d'itérations $x^*$ est-il atteint?

Barême

Exercice 1: a) 3 pts b) 2 pts
Exercice 2: 1) 2 pts 2) a) 1 pt -b) 1 pt -c) 1 pt -d) 1pt
Exercice 3: 1) 1 pt 2) a) 1 pt -b) 1 pt -c) 2 pts 3) a) 0,5 pts -b) 1 pts 4) a) 2 pts -b) 0,5 pts 5) 1 pt 6) 2 pts

Exercice 2.

1) a) $k=0 \quad d^{(0)} = p^{(0)}$

$k \to k+1 \quad$ hyp: $\quad d^{(k)} = \mathrm{Lin}\{p^{(0)}, p^{(1)}, \dots, p^{(k)}\}$

$$d^{(k+1)} = p^{(k+1)} - \sum_{i=0}^{k} \underbrace{\frac{p^{(k+1)} \cdot Q d^{(i)}}{d^{(i)} \cdot Q d^{(i)}}}_{\alpha_i} d^{(i)} = p^{(k+1)} - \sum_{i=0}^{k} \alpha_i d^{(i)} \in \mathrm{Lin}\{p^{(0)}, p^{(1)}, \dots, p^{(k+1)}\}$$

$\forall k \geq 0, \quad d^{(k)} = 0 \Rightarrow \quad 0 = p^{(k)} - \sum_{i=0}^{k-1} \alpha_i d^{(i)} = p^{(k)} - \sum_{i=0}^{k-1} \alpha'_i p^{(i)}$

$\Rightarrow p^{(k)}$ est lié avec $p^{(0)}, \dots, p^{(k-1)}$ faux par hypothèse du pb.

b) $k=1 \quad d^{(1)} \cdot Q d^{(0)} = p^{(1)} \cdot Q d^{(0)} - \dfrac{p^{(1)} Q d^{(0)}}{\cancel{d^{(0)} Q d^{(0)}}} \cancel{d^{(0)} \cdot Q d^{(0)}} = 0$

$k \to k+1 \quad$ hyp: $0 \leq i < j \leq k \quad d^{(i)} \cdot Q d^{(j)} = 0$

$\leq j < k+1, \quad d^{(k+1)} \cdot Q d^{(j)} = p^{(k+1)} \cdot Q d^{(j)} - \sum_{i=0}^{k} \dfrac{p^{(k+1)} \cdot Q d^{(i)}}{d^{(i)} \cdot Q d^{(i)}} d^{(i)} \cdot Q d^{(j)}$

$= p^{(k+1)} \cdot Q d^{(j)} - \dfrac{p^{(k+1)} \cdot Q d^{(j)}}{\cancel{d^{(j)} \cdot Q d^{(j)}}} \cancel{d^{(j)} \cdot Q d^{(j)}} = 0.$

2) a) $Q = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \qquad f(x) = \frac{1}{2} x \cdot Q x.$

b) $d^{(0)} = p^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad d^{(1)} = p^{(1)} - \dfrac{p^{(1)} Q d^{(0)}}{d^{(0)} Q d^{(0)}} d^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \dfrac{[1 \ 1]\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}}{[1 \ 0]\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$= \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{4}{4} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

c) $x^{(0)} + \lambda d^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1+\lambda \\ 1 \end{bmatrix} \qquad 4(1+\lambda^{(0)}) = 0 \Rightarrow \lambda^{(0)} = -1$

$x^{(1)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad x^{(1)} + \lambda d^{(1)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1+\lambda \end{bmatrix} \qquad 2(1+\lambda^{(1)}) = 0 \Rightarrow \lambda^{(1)} = -1$

$x^{(2)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (c'est le minimum de $f$)

d) $x^{(0)}, x^{(1)}$ identiques au c). $\quad x^{(2)} = x^{(1)} + \lambda p^{(1)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda \\ 1+\lambda \end{bmatrix}$

$4\lambda^{(1)} + 2(1+\lambda^{(1)}) = 0 \Rightarrow \lambda^{(1)} = -\frac{1}{3} \Rightarrow x^{(2)} = \begin{bmatrix} -1/3 \\ 2/3 \end{bmatrix}$

Si on ne prend pas des directions conjuguées, on n'atteind pas le minimum.

Exercice 3

1) $\nabla f(x^*) = 0 \rightarrow Ax^* = b$

2) a) La dérivée directionnelle au point $x^{(k+1)}$ est nulle : $\nabla f(x^{(k)} + \alpha^{(k)} d^{(k)}) \cdot d^{(k)} = 0$

$[A(x^{(k)} + \alpha^{(k)} d^{(k)}) - b] \cdot d^{(k)} = 0$

$[g^{(k)} + \alpha^{(k)} A d^{(k)}] \cdot d^{(k)} = 0 \rightarrow \alpha^{(k)} = -\dfrac{g^{(k)} \cdot d^{(k)}}{d^{(k)} \cdot A d^{(k)}}$

b) $E(x) = f(x) - f(x^*) = \frac{1}{2} x \cdot Ax - b \cdot x - \frac{1}{2} x^* \cdot Ax^* + b \cdot x^*$

$= \frac{1}{2} x \cdot Ax - x \cdot Ax^* - \frac{1}{2} x^* \cdot Ax^* + x^* \cdot Ax^*$

$= \frac{1}{2}(x - x^*) \cdot A(x - x^*)$

Taylor : $E(x) = f(x) - f(x^*) = \overbrace{\nabla f(x^*)}^{0} \cdot (x - x^*) + \frac{1}{2}(x - x^*) \underbrace{Hf(x^*)}_{\downarrow A} (x - x^*)$

c) $E(x^{(k)}) - E(x^{(k+1)}) = \frac{1}{2}(x^{(k)} - x^*) \cdot A(x^{(k)} - x^*) - \frac{1}{2}(\underbrace{x^{(k)} + \alpha^{(k)} d^{(k)} - x^*}) \cdot A(\underbrace{x^{(k)} + \alpha^{(k)} d^{(k)} - x^*})$

$= -(x^{(k)} - x^*) \cdot A \alpha^{(k)} d^{(k)} - \frac{1}{2} \alpha^{2(k)} d^{(k)} \cdot A d^{(k)}$

on remplace $\alpha^{(k)}$ et on remarque que $A x^{(k)} - Ax^* = A(x^{(k)} - x^*) = g^{(k)} - \cancel{g^*} \to 0$

$\Rightarrow x^{(k)} - x^* = A^{-1} g^{(k)}$

$= + \dfrac{g^{(k)} \cdot d^{(k)}}{d^{(k)} \cdot A d^{(k)}} (A^{-1} g^{(k)}) \cdot A d^{(k)} - \frac{1}{2} \left[\dfrac{g^{(k)} \cdot d^{(k)}}{d^{(k)} \cdot A d^{(k)}}\right]^2 d^{(k)} \cdot A d^{(k)}$

$= \dfrac{1}{2} \dfrac{(g^{(k)} \cdot d^{(k)})^2}{d^{(k)} \cdot A d^{(k)}}$

$$\frac{E(x^{(k)}) - E(x^{(k+1)})}{E(x^{(k)})} = \frac{[g^{(k)} \cdot d^{(k)}]^2}{[d^{(k)} \cdot A d^{(k)}][(A^{-1} g^{(k)}) \cdot A A^{-1} g^{(k)}]} = \frac{N}{(d^{(k)} \cdot A d^{(k)})(g^{(k)} \cdot A^{-1} g^{(k)})}$$

3) a) $f'(x^{(k)}; d^{(k)}) = \nabla f(x^{(k)}) \cdot d^{(k)} = g^{(k)} \cdot d^{(k)} < 0$ donc $f$ décroît dans direction $d^{(k)}$ à partir de $x^{(k)}$

remarquer que $\alpha^{(k)} > 0$ (c'est cohérent)

b) $\dfrac{E(x^{(k)}) - E(x^{(k+1)})}{E(x^{(k)})} \geqslant \beta \Rightarrow E(x^{(k)})(1 - \beta) \geqslant E(x^{(k+1)})$

4) a) on appelle $\lambda_{min}$ la valeur propre la plus petite de $A$, $\lambda_{max}$ la plus grande. $\lambda_{max} \geqslant \lambda_{min}$

on a : $\lambda_{max}\, d^{(k)} \cdot d^{(k)} \geqslant d^{(k)} \cdot A d^{(k)} \geqslant \lambda_{min}\, d^{(k)} \cdot d^{(k)} \geqslant 0$

de même $\dfrac{1}{\lambda_{min}} g^{(k)} \cdot g^{(k)} \geqslant g^{(k)} \cdot A^{-1} g^{(k)} \geqslant \dfrac{1}{\lambda_{max}} g^{(k)} \cdot g^{(k)} \geqslant 0$

$\Rightarrow (d^{(k)} \cdot A d^{(k)})(g^{(k)} \cdot A^{-1} g^{(k)}) \leqslant \dfrac{\lambda_{max}}{\lambda_{min}} (d^{(k)} \cdot d^{(k)})(g^{(k)} \cdot g^{(k)})$

$$\frac{E(x^{(k)}) - E(x^{(k+1)})}{E(x^{(k)})} \geqslant \frac{\left[d^{(k)} \cdot g^{(k)}\right]^2}{\frac{d_{max}}{d_{min}} \left[d^{(k)} \cdot d^{(k)}\right]\left[g^{(k)} \cdot g^{(k)}\right]} \geqslant \frac{d_{min}}{d_{max}} \beta$$

$$\Rightarrow E(x^{(k)}) \left(1 - \frac{d_{min}}{d_{max}} \beta\right) \geqslant E(x^{(k+1)})$$

b] $1 - \frac{d_{min}}{d_{max}} \beta \geqslant 1 - \beta$ il y a des chances que la méthode dans le cas 3) converge plus vite.

5) Prendre $d^{(k)} = \alpha\, g^{(k)}$ $\alpha < 0$

en effet $(d^{(k)} \cdot g^{(k)})^2 = \cancel{\alpha}^2 (g^{(k)} \cdot g^{(k)})^2 \geqslant \beta \cancel{\alpha}^2 (g^k \cdot g^k)(g^{(k)} \cdot g^{(k)}) \Rightarrow \beta = 1$ est possible

on se déplace, à chaque itération k, dans la direction opposée au gradient $\Rightarrow$ méthode de plus forte pente.

6) Soit $e$ un vecteur propre associé à une valeur $\lambda$ de $A$ (rappelons que $\lambda > 0$)

prendre $d^{(k)} = e$

et $g^{(k)} = \alpha e$ $\alpha < 0$ possible si $g^{(k)} = A x^{(k)} - b = A x^{(k)} - A x^* = A(x^{(k)} - x^*) = \alpha e$

c'est-à-dire si $x^{(k)} - x^* = \frac{\alpha e}{\lambda} \rightarrow x^{(k)} = \frac{\alpha e}{\lambda} + x^*$

Dans ce cas la condition de 3) devient :

$$(d^{(k)} \cdot g^{(k)})^2 = \cancel{\alpha}^2 (e \cdot e)^2 \geqslant \beta\, \underbrace{\cancel{\lambda}\, e \cdot e}_{d^{(k)} \cdot A d^{(k)}} \;\underbrace{\frac{1}{\cancel{\lambda}} \cancel{\alpha}^2 (e \cdot e)}_{g^{(k)} \cdot A^{-1} g^{(k)}} = \beta\, (e \cdot e)^2$$

$\beta = 1$ est possible.

Dans ce cas $x^*$ est atteint en 1 itération : $x^{(1)} = x^*$.

Exercice 1

a) minimiser $\underset{\underset{\delta_1}{\uparrow}}{\dfrac{x_1 x_2 x_3}{2}} + \underset{\underset{\delta_2}{\uparrow}}{\dfrac{1}{x_1}} + \underset{\underset{\delta_3}{\uparrow}}{\dfrac{2}{x_2}} + \underset{\underset{\delta_4}{\uparrow}}{\dfrac{16}{x_3}}$ $\quad x_1 > 0,\ x_2 > 0,\ x_3 > 0$

Le dual :

$$\left.\begin{array}{ll} (x_1) & \delta_1 - \delta_2 = 0 \\ (x_2) & \delta_1 - \delta_3 = 0 \\ (x_3) & \delta_1 - \delta_4 = 0 \end{array}\right\} \Rightarrow \delta_1 = \delta_2 = \delta_3 = \delta_4 \qquad \delta_i = \frac{1}{4} \quad i = 1, 2, 3, 4$$

$$\delta_1 + \delta_2 + \delta_3 + \delta_4 = 1, \quad \delta_i > 0$$

$$v(\delta) = \prod \left(\frac{c_i}{\delta_i}\right)^{\delta_i} = (2)^{1/4} (4)^{1/4} (8)^{1/4} (16 \times 4)^{1/4} = (4 \times 4 \times 4 \times 4)^{1/4} \big(\underbrace{2 \times 2 \times 2 \times 2}_{16}\big)^{1/4} = 4 \times 2 = 8$$

$$\delta_i = \frac{u_i(x)}{v(\delta)} \Rightarrow \quad \frac{x_1 x_2 x_3}{2} = 2, \quad \frac{1}{x_1} = 2, \quad \frac{2}{x_2} = 2, \quad \frac{16}{x_3} = 2$$

$$\Rightarrow x_1 = \frac{1}{2}, \quad x_2 = 1, \quad x_3 = 8$$

b) minimiser $\dfrac{1}{2}(xyz - xz^2 - zy^2 + yz^2) + \dfrac{1}{x-y} + \dfrac{2}{y-z} + \dfrac{16}{z}$ $\quad x > y > z > 0$

les contraintes : $x - y > 0$, $y - z > 0$, $z > 0$

l'objectif : $xyz - xz^2 - zy^2 + yz^2 = (xy - xz - y^2 + yz)z = (x-y)(y-z)z$

Le problème est équivalent au pb du a) en posant : $x - y = x_1$, $y - z = x_2$, $z = x_3$

$\Rightarrow \quad z = x_3, \quad y = x_2 + x_3, \quad x = x_1 + x_2 + x_3$

$\quad z = 8, \quad y = 9, \quad x = 9{,}5$

Exercice

$$f(x_1, x_2) = x_1^2 - x_1 x_2 + \frac{3}{2} x_2^2$$

Calculer les 2 premiers termes de la suite de pts fournis par (DFP) puis par (BFGS) avec $x^0 = \binom{1}{2}$, $H^0 = I$.

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 - x_2 \\ -x_1 + 3x_2 \end{pmatrix}$$

$x^0 = (1, 2)$, $f(x^0) = 5$, $\nabla f(x^0) = \binom{0}{5}$

$\boxed{DFP}$

$$g_0(\lambda) = f\left(\binom{1}{2} - \lambda \binom{0}{5}\right) = f\binom{1}{2 - 5\lambda} \qquad (5\lambda \quad 5 - 15\lambda)\binom{0}{-5} = -25 + 75\lambda$$

$g_0'(\lambda) = -25 + 75\lambda$ ; $g_0'(\lambda_0) = 0 \Rightarrow \lambda_0 = 1/3$

$$x^1 = \binom{1}{2} - \frac{1}{3}\binom{0}{5} = \binom{1}{1/3} \qquad \boxed{x_1 = (1, \tfrac{1}{3})}$$

$$f(x^1) = 1 - \frac{1}{3} + \frac{3}{2} \times \frac{1}{9 \times 3} = 1 - \frac{1}{3} + \frac{1}{6} = \frac{5}{6}$$

$$\nabla f(x^1) = \binom{5/3}{0}$$

$$\delta_0 = x^1 - x^0 = \binom{0}{-5/3} \quad ; \quad \gamma_0 = \binom{5/3}{-5} \qquad (\gamma_0 = \nabla f(x^1) - \nabla f(x^0))$$

$$\delta_0 \delta_0^T = \binom{0}{-5/3}\begin{pmatrix} 0 & -5/3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \frac{25}{9} \end{pmatrix} \qquad \gamma_0 \gamma_0^T = \binom{5/3}{-5}\begin{pmatrix} 5/3 & -5 \end{pmatrix} = \begin{pmatrix} 25/9 & -\frac{25}{3} \\ -\frac{25}{3} & 25 \end{pmatrix}$$

$$\delta_0^T \cdot \gamma_0 = \begin{pmatrix} 0 & -\frac{5}{3} \end{pmatrix}\binom{5/3}{-5} = \frac{25}{3} \qquad \gamma_0^T \cdot \gamma_0 = \left(\frac{5}{3}\right)^2 + 25 = \frac{250}{9}$$

$$H_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{3} \end{pmatrix} - \begin{pmatrix} \frac{1}{10} & -\frac{3}{10} \\ -\frac{3}{10} & \frac{9}{10} \end{pmatrix} = \begin{pmatrix} \frac{9}{10} & +\frac{3}{10} \\ +\frac{3}{10} & \frac{13}{30} \end{pmatrix}$$

$k=1$

$$g_1(\lambda) = f\left( \begin{pmatrix} 1 \\ 1/3 \end{pmatrix} - \lambda \begin{pmatrix} \frac{9}{10} & \frac{3}{10} \\ \frac{3}{10} & \frac{13}{30} \end{pmatrix} \begin{pmatrix} 5/3 \\ 0 \end{pmatrix} \right)$$

$$= f\left( \begin{pmatrix} 1 \\ 1/3 \end{pmatrix} - \lambda \begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix} \right)$$

$$= f\begin{pmatrix} 1-\frac{3}{2}\lambda \\ \frac{1}{3}-\frac{1}{2}\lambda \end{pmatrix}$$

$2-3\lambda-\frac{1}{3}+\frac{1}{2}\lambda \quad -1+\frac{3}{2}\lambda+1-\frac{3}{2}\lambda$

$\begin{pmatrix} \frac{5}{3}-\frac{5}{2}\lambda & 0 \end{pmatrix} \begin{pmatrix} -3/2 \\ -1/2 \end{pmatrix}$

$$g_1'(\lambda) = -\frac{3}{2}\left[2-3\lambda-\frac{1}{3}+\frac{1}{2}\lambda\right] - \frac{1}{2}\left[\cancel{-1}+\cancel{\frac{3}{2}\lambda}+\cancel{1}-\cancel{\frac{3}{2}\lambda}\right]$$

$$= -\frac{3}{2}\left[\frac{5}{3}-\frac{5}{2}\lambda\right]$$

$$g_1'(\lambda_1) = 0 \Rightarrow \boxed{\lambda_1 = \frac{2}{3}}$$

$x^{(2)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ; $\nabla f(x^2) = 0$ $\nabla f(x^2) = 0$

$\Rightarrow$ opt local.

$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$ déf pos

BFGS

calcul de $D_{(1)}$ puis $D_{(1)}^{-1}$

$$\gamma^0 \otimes \gamma^0 = \begin{bmatrix} 25/9 & -\frac{25}{3} \\ -\frac{25}{3} & 25 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{3}{25}\begin{bmatrix} 25/9 & -\frac{25}{3} \\ -\frac{25}{3} & 25 \end{bmatrix} - \frac{9}{25}\begin{bmatrix} 0 & 0 \\ 0 & \frac{25}{9} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1/3 & -1 \\ -1 & 3 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4/3 & -1 \\ -1 & 3 \end{bmatrix}$$

$$D_1^{-1} = \frac{1}{3}\begin{bmatrix} 3 & +1 \\ +1 & 4/3 \end{bmatrix} = \begin{bmatrix} 1 & 1/3 \\ 1/3 & 4/9 \end{bmatrix}$$

$$d^{(1)} = \begin{bmatrix} -1 \\ -1/3 \end{bmatrix} \qquad y^{(1)} = \begin{bmatrix} -5/3 \\ 0 \end{bmatrix} \qquad \boxed{H_2 = A^{-1}}$$

$$S^{(1)} = \begin{bmatrix} 9/10 & 3/10 \\ 3/10 & 13/30 \end{bmatrix} \qquad \begin{bmatrix} 9/10 & 3/10 \\ 3/10 & 13/30 \end{bmatrix}\begin{bmatrix} -5/3 \\ 0 \end{bmatrix} = \begin{bmatrix} -3/2 \\ -1/2 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 1/3 \end{bmatrix}\begin{bmatrix} 1 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 1/3 \\ 1/3 & 1/9 \end{bmatrix} \qquad \begin{bmatrix} -\frac{5}{3} & 0 \end{bmatrix}\begin{bmatrix} -1 \\ -1/3 \end{bmatrix} = \frac{5}{3}$$

$$\begin{bmatrix} 3/2 \\ 1/2 \end{bmatrix}\begin{bmatrix} 3/2 & 1/2 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix} \qquad \begin{bmatrix} -5/3 & 0 \end{bmatrix}\begin{bmatrix} -3/2 \\ -1/2 \end{bmatrix} = 5/2$$

$$\begin{bmatrix} \frac{9}{10} & \frac{3}{10} \\ \frac{3}{10} & \frac{13}{30} \end{bmatrix} + \frac{3}{5}\begin{bmatrix} 1 & 1/3 \\ 1/3 & 1/9 \end{bmatrix} - \frac{1}{4}\times\frac{2}{5}\begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \frac{9}{10} & \frac{3}{10} \\ \frac{3}{10} & \frac{13}{30} \end{bmatrix} + \begin{bmatrix} \frac{3}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{15} \end{bmatrix} - \begin{bmatrix} 9/10 & 3/10 \\ 3/10 & 1/10 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 \\ 0 & 8/30 \end{bmatrix} + \begin{bmatrix} \frac{3}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{15} \end{bmatrix} = \begin{bmatrix} 3/5 & 1/5 \\ 1/5 & \frac{10}{15} \end{bmatrix}$$

$$\underset{H_2}{\begin{bmatrix} 3/5 & 1/5 \\ 1/5 & 6/30 \end{bmatrix}} \underset{A}{\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}} \underset{=}{=} \underset{I}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}$$

**BFGS**

$x^1$ : m̃ point

$H_1 = ?$

$\gamma_0^T H_0 \gamma_0 = \gamma_0^T \gamma_0 = \left(\frac{5}{3} \quad -5\right)\begin{pmatrix} 5/3 \\ -5 \end{pmatrix} = \frac{250}{9}$

$$\delta_0 \gamma_0^T = \begin{pmatrix} 0 \\ 5/3 \end{pmatrix} \left(5/3 \quad -5\right) = \begin{pmatrix} 0 & 0 \\ \frac{25}{9} & -\frac{25}{3} \end{pmatrix}$$

$$\gamma_0 \cdot \delta_0^T = \begin{pmatrix} 5/3 \\ -5 \end{pmatrix} \left(0 \quad 5/3\right) = \begin{pmatrix} 0 & 25/9 \\ 0 & -25/3 \end{pmatrix}$$

$$H_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left(1 + \frac{250}{9} \times \frac{3}{25}\right) \begin{pmatrix} 0 & 0 \\ 0 & \frac{25}{9} \times \frac{3}{25} \end{pmatrix}$$

$$- \frac{3}{25} \left[ \begin{pmatrix} 0 & 0 \\ \frac{25}{9} & -\frac{25}{3} \end{pmatrix} + \begin{pmatrix} 0 & 25/9 \\ 0 & -25/3 \end{pmatrix} \right]$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{3} \times \frac{13}{3} \end{pmatrix} - \begin{pmatrix} 0 & \frac{1}{3} \\ \frac{1}{3} & -2 \end{pmatrix}$$

$\frac{13}{9}$

$$= \begin{pmatrix} 1 & -1/3 \\ -\frac{1}{3} & 4/9 \end{pmatrix}$$

44

$$g_1(\lambda) = f\left(\begin{pmatrix}1\\1/3\end{pmatrix} - \lambda\begin{pmatrix}1 & 1/3\\1/3 & 4/9\end{pmatrix}\begin{pmatrix}5/3\\0\end{pmatrix}\right)$$

$-\frac{10}{3} + \frac{5}{9}$

$5-30$

$$= f\left(\begin{pmatrix}1\\1/3\end{pmatrix} - \lambda\begin{pmatrix}5/3\\5/9\end{pmatrix}\right) = f\begin{pmatrix}1 - \frac{5}{3}\lambda\\ \frac{1}{3} - \frac{5}{9}\lambda\end{pmatrix}$$

$$g_1'(\lambda) = -\frac{5}{3}\left[\frac{5}{3} - \frac{25}{9}\lambda\right] - \frac{5}{9}\left[0\right]$$

$$g_1'(\lambda_1) = 0 \Rightarrow \lambda_1 = \frac{5}{3} \times \frac{9}{25} \Rightarrow \boxed{\lambda_1 = \frac{3}{5}}$$

$$x^2 = \begin{pmatrix}0\\0\end{pmatrix}$$

Rq $\nabla^2 f(x) = \begin{pmatrix}2 & -1\\ & 3\end{pmatrix} = A$

$$Ax = \begin{pmatrix}2x_1 - x_2\\ -x_1 + 3x_2\end{pmatrix} = \nabla f(x) \Rightarrow b = 0$$

$$f(x) = \frac{1}{2}x^T A x$$
$$= \frac{1}{2}(x_1 \;\; x_2)\begin{pmatrix}2x_1 - x_2\\ -x_1 + 3x_2\end{pmatrix} = \frac{1}{2}\left[2x_1^2 - 2x_1x_2 + 3x_2^2\right]$$

$$f(x) = \frac{1}{2}x^T\begin{pmatrix}2 & -1\\ -1 & 3\end{pmatrix}x$$

vérifions que $(x^1 - x^0)^T A (x^2 - x^1) = 0$

$$x^1 - x^0 = S_0 = \begin{pmatrix}0\\-5/3\end{pmatrix} \qquad x^2 - x^1 = \begin{pmatrix}-1\\-1/3\end{pmatrix}$$

$$\begin{pmatrix}0 & -5/3\end{pmatrix}\begin{pmatrix}2 & -1\\ -1 & 3\end{pmatrix}\begin{pmatrix}-1\\-1/3\end{pmatrix} = \begin{pmatrix}\frac{5}{3} & -5\end{pmatrix}\begin{pmatrix}-1\\-1/3\end{pmatrix} = -\frac{5}{3} + \frac{5}{3} = 0$$

$$\left(A_2\right)^1 = A$$