Optimisation Continue
ISTIL 2ème année
Corrigé de la feuille 3[1]

Exercice 1 : Gradient à pas variable

1.a La fonctionnelle J est elliptique, donc elle est strictement convexe. On en déduit que le minimum, s'il existe, est unique. De plus, toujours du fait de l'ellipticité de J, cette fonctionnelle est coercive. En outre, elle est continue. On en déduit qu'elle admet un minimum. Ainsi

$$\boxed{\text{J admet un unique minimum}}$$

Comme J est de classe $\mathscr{C}^1$, le minimum u de J sur un espace vectoriel est caractérisé par l'équation d'Euler

$$\boxed{\mathrm{J}'(u) = 0}$$

1.b D'après la définition de la suite $(u_k)_{k\in\mathbb{N}}$, nous avons

$$u_{k+1} - u = u_k - \rho_k \mathrm{J}'(u_k) - u$$

d'où
$$\begin{aligned}\|u_{k+1} - u\|^2 &= \|u_k - u - \rho_k \mathrm{J}'(u_k)\|^2 \\ &= \|u_k - u\|^2 + \|\rho_k \mathrm{J}'(u_k)\|^2 - 2\rho_k\left(u_k - u, \mathrm{J}'(u_k)\right) \\ &= \|u_k - u\|^2 + \rho_k^2\|\mathrm{J}'(u_k)\|^2 - 2\rho_k\left(u_k - u, \mathrm{J}'(u_k)\right)\end{aligned}$$

La condition Lipschitzienne sur J donne

$$\|\mathrm{J}'(u_k)\| = \|\mathrm{J}'(u_k) - \mathrm{J}'(u)\| \leqslant \mathrm{M}\|u_k - u\|$$

d'où
$$\rho_k^2\|\mathrm{J}'(u_k)\|^2 \leqslant \rho_k^2\mathrm{M}^2\|u_k - u\|^2$$

Utilisons maintenant la condition d'ellipticité de J

$$(u_k - u, \mathrm{J}'(u_k)) = (u_k - u, \mathrm{J}'(u_k) - \mathrm{J}'(u)) \geqslant \alpha\|u_k - u\|^2$$

soit
$$-2\rho_k\left(u_k - u, \mathrm{J}'(u_k)\right) \leqslant -2\rho_k\alpha\|u_k - u\|^2$$

En utilisant les deux inégalités que l'on vient de trouver, on obtient

$$\begin{aligned}\|u_{k+1} - u\|^2 &\leqslant \|u_k - u\|^2 + \rho_k^2\mathrm{M}^2\|u_k - u\|^2 - 2\rho_k\alpha\|u_k - u\|^2 \\ &\leqslant \left(\mathrm{M}^2\rho_k^2 - 2\alpha\rho_k + 1\right)\|u_k - u\|^2\end{aligned}$$

On en déduit que
$$\boxed{\begin{array}{c}\|u_{k+1} - u\|^2 \leqslant \tau(\rho_k)\|u_k - u\|^2 \\ \text{avec} \\ \tau(\rho_k) = \mathrm{M}^2\rho_k^2 - 2\alpha\rho_k + 1\end{array}}$$

[1] généré avec $\LaTeX\,2_\varepsilon$. Tous les commentaires, compléments, insultes et remarques désobligeantes sont les bienvenus à `perrier@math.u-bordeaux1.fr`

1.c On a $$\tau(\rho) = \mathrm{M}^2\rho^2 - 2\alpha\rho + 1 = \mathrm{M}^2\rho\left(\rho - \frac{2\alpha}{\mathrm{M}^2}\right) + 1$$

On en déduit que si $$\rho \leqslant b < \frac{2\alpha}{\mathrm{M}^2}$$

alors $$\rho - \frac{2\alpha}{\mathrm{M}^2} \leqslant b - \frac{2\alpha}{\mathrm{M}^2} < 0$$

d'où $$\mathrm{M}^2\rho\left(\rho - \frac{2\alpha}{\mathrm{M}^2}\right) \leqslant \mathrm{M}^2 a\left(b - \frac{2\alpha}{\mathrm{M}^2}\right) < 0$$

soit $$\tau(\rho) \leqslant \mathrm{M}^2 a\left(b - \frac{2\alpha}{\mathrm{M}^2}\right) + 1 < 1$$

Posons $$\beta = \sqrt{\mathrm{M}^2 a\left(b - \frac{2\alpha}{\mathrm{M}^2}\right) + 1} < 1$$

On a alors $$\|u_{k+1} - u\| \leqslant \beta\|u_k - u\|$$

Montrons à présent par récurrence la propriété suivante

$$\mathscr{P}(k) : \text{« } \|u_k - u\| \leqslant \beta^k\|u_0 - u\| \text{ »}$$

- $\underline{\mathscr{P}(0)}$ La propriété est évidente.
- $\underline{\mathscr{P}(k) \Longrightarrow \mathscr{P}(k+1)}$ Supposons que $\mathscr{P}(k)$ soit vraie. Alors

$$\|u_k - u\| \leqslant \beta^k\|u_0 - u\|$$

D'après ce que l'on a montré auparavant dans la question

$$\|u_{k+1} - u\| \leqslant \beta\|u_k - u\| \leqslant \beta^{k+1}\|u_0 - u\|$$

Donc $\mathscr{P}(k+1)$ est vraie.

- Conclusion pour tout entier k, $\mathscr{P}(k)$ est vraie.

D'où $$\boxed{\exists\beta \in \,]\,0\,;1\,[\qquad \|u_k - u\| \leqslant \beta^k\|u_0 - u\|}$$

1.d On s'intéresse au cas de la fonctionnelle elliptique

$$\mathrm{J} : v \in \mathbb{R}^n \longmapsto \frac{1}{2}(\mathrm{A}v, v) - (b, v)$$

Dans ce cas, on a $\qquad \forall u, v \in \mathrm{V} \qquad \mathrm{J}'(u)v = (\mathrm{A}u - b, v)$

On en déduit que pour tout vecteur x de norme inférieure à 1

$$|(\mathrm{J}'(u) - \mathrm{J}'(v), x)| = |(\mathrm{A}(u-v), x)| \leqslant \|\mathrm{A}(u-v)\|\|x\| \leqslant \|\mathrm{A}(u-v)\|$$

Par ailleurs, on a

$$\begin{aligned}(\mathrm{J}'(v) - \mathrm{J}'(u), v - u) &= (\mathrm{J}'(v), v-u) - (\mathrm{J}'(u), v-u)\\ &= (\mathrm{A}v - b, v - u) - (\mathrm{A}u - b, v-u)\\ &= (\mathrm{A}(v-u), v-u)\end{aligned}$$

La matrice A est symétrique définie positive, donc elle est diagonalisable en base orthonormée $(e_1, \dots, e_n)$, et ses valeurs propres $\lambda_1 \leqslant \lambda_2 \dots \leqslant \lambda_n$ sont positives. On a donc

$$\begin{aligned}
(\mathrm{A}x, x) &= \left(\mathrm{A}\left(\sum_{i=1}^n x_i e_i\right), \sum_{i=1}^n x_j e_j\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n x_i x_j (\mathrm{A}e_i, e_j) \\
&= \sum_{i=1}^n \sum_{j=1}^n x_i x_j \lambda_i (e_i, e_j) \\
&= \sum_{i=1}^n \lambda_i x_i^2 \\
&\geqslant \lambda_1 \sum_{i=1}^n x_i^2 \\
(\mathrm{A}x, x) &\geqslant \lambda_1 \|x\|^2
\end{aligned}$$

d'où $$\boxed{(\mathrm{J}'(v) - \mathrm{J}'(u), v - u) \geqslant \left(\min_{\lambda \in \mathrm{Sp}(\mathrm{A})} \lambda\right) \|u - v\|^2}$$

de même, on a
$$\begin{aligned}
\|\mathrm{A}x\|^2 &= (\mathrm{A}x, \mathrm{A}x) \\
&= \sum_{i=1}^n \lambda_i^2 x_i^2 \\
&\leqslant \lambda_n^2 \sum_{i=1}^n x_i^2 \\
\|\mathrm{A}x\|^2 &\leqslant \lambda_n^2 \|x\|^2
\end{aligned}$$

soit $$\|\mathrm{A}x\| \leqslant \lambda_n \|x\|$$

En conclusion $$\boxed{\|\mathrm{J}'(u) - \mathrm{J}'(v)\| \leqslant \left(\max_{\lambda \in \mathrm{Sp}(\mathrm{A})} \lambda\right) \|u - v\|}$$

Exercice 2 : Méthode B de quasi Newton

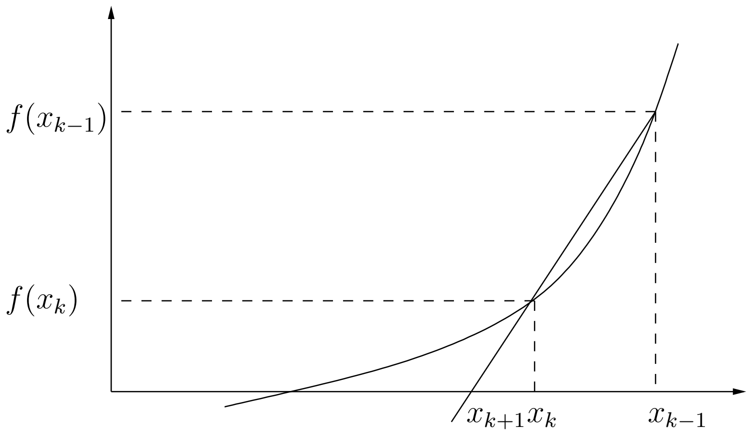
$\boxed{\textbf{2.a}}$ Dans le cas $n = 1$, la méthode s'écrit

$$f(x_k) + \frac{f(x_{k-1}) - f(x_k)}{x_{k-1} - x_k}(x_{k+1} - x_k) = 0$$

Le point x_{k+1} est donc l'intersection de la droite passant par les points

$$(x_k, f(x_k)) \qquad \text{et} \qquad (x_{k-1}, f(x_{k-1})),$$

avec l'axe des abscisses. Cette méthode peut se représenter graphiquement de la manière suivante



Cette méthode est connue, en dimension 1 sous le nom de *méthode de la sécante.*

$\boxed{\textbf{2.b}}$ Grâce à la relation (**2**), on connaît l'image $\mathbb{R}s$ par la matrice $\bar{\mathrm{B}}$. La relation (**3**) permet de connaître l'image de $s^\perp$ par $\bar{\mathrm{B}}$. On connaît donc l'image de $s \oplus s^\perp = \mathbb{R}^n$ par la matrice $\bar{\mathrm{B}}$. Si on suppose qu'il existe $\bar{\mathrm{B}}_1$ et $\bar{\mathrm{B}}_2$ vérifiant (**2**) et (**3**), alors elles coïncident sur s et $s^\perp$, donc sur $\mathbb{R}^n$. On en déduit que la matrice $\bar{\mathrm{B}}$ est unique si elle existe.

De plus, si on pose, comme suggéré par l'énoncé

$$\bar{\mathrm{B}} = \mathrm{B} + \frac{(y - \mathrm{B}s)\,{}^t s}{{}^t s\, s},$$

alors
$$\begin{aligned}\bar{\mathrm{B}}s &= \mathrm{B}s + \frac{(y - \mathrm{B}s)\,{}^t s\, s}{{}^t s\, s}\\ &= \mathrm{B}s + y - \mathrm{B}s\\ &= y\end{aligned}$$

et, pour tout $z \in s^\perp$
$$\begin{aligned}\bar{\mathrm{B}}z &= \mathrm{B}z + \frac{(y - \mathrm{B}s)\,{}^t s\, z}{{}^t s\, s}\\ &= \mathrm{B}z\end{aligned}$$

Donc la matrice donnée par l'énoncé vérifie bien les conditions (3) et (4).

2.c Commençons par montrer que la norme de Frobenius est une norme d'opérateur.

$$\begin{aligned}\|\mathrm{AB}\|_{\mathrm{F}}^2 &= \sum_{i,j=1}^{n} (\mathrm{AB})_{i,j}{}^2 \\ &= \sum_{i,j=1}^{n} \left(\sum_{k=1}^{n} a_{i,k} b_{k,j}\right)^2\end{aligned}$$

or $$\forall i,j \qquad \left(\sum_{k=1}^{n} a_{i,k} b_{k,j}\right)^2 \leqslant \left(\sum_{k=1}^{n} a_{i,k}^2\right)\left(\sum_{l=1}^{n} b_{l,j}^2\right)$$

d'où $$\|\mathrm{AB}\|_{\mathrm{F}}^2 \leqslant \sum_{i,j,k,l=1}^{n} a_{i,k}^2 b_{l,j}^2 = \|\mathrm{A}\|_{\mathrm{F}}^2 \|\mathrm{B}\|_{\mathrm{F}}^2$$

Soit $\widehat{\mathrm{B}}$ une matrice telle que $\widehat{\mathrm{B}}s = y$. Alors

$$\bar{\mathrm{B}} - \mathrm{B} = \frac{(y - \mathrm{B}s)\,{}^t s}{{}^t s\, s} = \frac{(\widehat{\mathrm{B}}s - \mathrm{B}s)\,{}^t s}{{}^t s\, s} = \frac{(\widehat{\mathrm{B}} - \mathrm{B})s\,{}^t s}{{}^t s\, s}$$

d'où $$\|\bar{\mathrm{B}} - \mathrm{B}\|_{\mathrm{F}} \leqslant \frac{\|\widehat{\mathrm{B}} - \mathrm{B}\|_{\mathrm{F}} \|s\,{}^t s\,\|_{\mathrm{F}}}{\left|{}^t s\, s\right|}$$

Si s a pour coordonnées $(s_1, s_2 \ldots, s_n)$, alors la matrice $s\,{}^t s$ a pour terme général $s_i s_j$. On a donc

$$\|s\,{}^t s\,\|_{\mathrm{F}}^2 = \sum_{i,j=1}^{n} s_i^2 s_j^2 = \sum_{i=1}^{n} s_i^2 \sum_{j=1}^{n} s_j^2 = \|s\|^4$$

On en déduit en particulier que $\|s\,{}^t s\,\|_{\mathrm{F}} = {}^t s\, s$. On trouve donc que

$$\forall \widehat{\mathrm{B}} \quad \widehat{\mathrm{B}} = y \qquad \|\bar{\mathrm{B}} - \mathrm{B}\|_{\mathrm{F}} \leqslant \|\widehat{\mathrm{B}} - \mathrm{B}\|$$

La matrice $\bar{\mathrm{B}}$ est donc une solution du problème d'optimisation.

2.d Supposons que $\sigma = 1 + {}^t v\, \mathrm{A}^{-1} u \neq 0$. Alors on a

$$\begin{aligned}\left(\mathrm{A} + u\,{}^t v\right)\left(\mathrm{A}^{-1} - \frac{1}{\sigma}\,\mathrm{A}^{-1} u\,{}^t v\,\mathrm{A}^{-1}\right) &= \mathrm{A}\left(\mathrm{id} + \mathrm{A}^{-1} u\,{}^t v\right)\left(\mathrm{id} - \frac{1}{\sigma}\,\mathrm{A}^{-1} u\,{}^t v\right)\mathrm{A}^{-1} \\ &= \mathrm{A}\Bigg(\mathrm{id} - \frac{1}{\sigma}\,\mathrm{A}^{-1} u\,{}^t v \\ &\qquad + \mathrm{A}^{-1} u\,{}^t v - \frac{1}{\sigma}\mathrm{A}^{-1} u\,{}^t v\,\mathrm{A}^{-1} u\,{}^t v\Bigg)\mathrm{A}^{-1} \\ &= \mathrm{A}\Bigg(\mathrm{id} - \frac{1}{\sigma}\,\mathrm{A}^{-1} u\,{}^t v \\ &\qquad + \mathrm{A}^{-1} u\,{}^t v - \frac{{}^t v\,\mathrm{A}^{-1} u}{\sigma}\mathrm{A}^{-1} u\,{}^t v\Bigg)\mathrm{A}^{-1} \\ &= \mathrm{A}\left(\mathrm{id} + \frac{-\,{}^t v\,\mathrm{A}^{-1} u - 1 + \sigma}{\sigma}\right)\mathrm{A}^{-1} \\ &= \mathrm{A}\mathrm{A}^{-1} \\ &= \mathrm{id}\end{aligned}$$

On en déduit que la matrice $\mathrm{A} + u\,{}^t v$ est une matrice régulière, dont l'inverse est

$$\mathrm{A}^{-1} - \frac{1}{\sigma}\,\mathrm{A}^{-1} u\,{}^t v\,\mathrm{A}^{-1}$$

Afin de montrer la réciproque, nous allons montrer la contraposée : supposons que $\sigma = 0$. Montrons que la matrice $\mathrm{A} + u\,{}^t v$ n'est alors pas inversible. On a

$$\mathrm{A}^{-1} u\,{}^t v\,\mathrm{A}^{-1} u\,{}^t v = \left({}^t v\,\mathrm{A}^{-1} u\right) \mathrm{A}^{-1} u\,{}^t v = -\mathrm{A}^{-1} u\,{}^t v$$

d'où
$$\mathrm{A}^{-1} u\,{}^t v \left(\mathrm{A}^{-1} u\,{}^t v + \mathrm{id}\right) = 0$$

donc le couple $\left(\mathrm{A}^{-1} u\,{}^t v, \left(\mathrm{A}^{-1} u\,{}^t v + \mathrm{id}\right)\right)$ est un couple diviseur de zéro. On en déduit qu'aucune de ces deux matrices n'est inversible.

$\boxed{\textbf{2.e}}$ La méthode (**1**), (**6**) s'écrit

$$\begin{cases} f(x_k) + \mathrm{B}_k(x_{k+1} - x_k) = 0 \\ \mathrm{B}_{k+1} = \mathrm{B}_k + \dfrac{(y_k - \mathrm{B}_k s_k)\,{}^t s_k}{{}^t s_k\, s_k} \end{cases}$$

L'itération de Newton s'écrit en fait

$$x_{k+1} = x_k - \mathrm{H}_k f(x_k)$$

avec
$$\mathrm{H}_k = \mathrm{B}_k^{-1}$$

Appliquons la question 2.d à B_k, avec

$$\begin{cases} u = \dfrac{y_k - \mathrm{B}_k s_k}{{}^t s_k\, s_k} \\ v = s_k \end{cases}$$

On a alors

$$\begin{aligned}
\mathrm{H}_{k+1} &= \mathrm{B}_{k+1}^{-1} \\
&= \left(\mathrm{B}_k + u\,{}^t v\right)^{-1} \\
&= \mathrm{B}_k^{-1} - \frac{1}{1 + {}^t u\, v}\,\mathrm{B}_k^{-1} u\,{}^t v\,\mathrm{B}_k^{-1} \\
&= \mathrm{B}_k^{-1} - \frac{1}{1 + {}^t s_k\,\mathrm{B}_k^{-1}\,\dfrac{(y_k - \mathrm{B}_k s_k)}{{}^t s_k\, s_k}}\,\mathrm{B}_k^{-1}\,\frac{y_k - \mathrm{B}_k s_k}{{}^t s_k\, s_k}\;{}^t s_k\,\mathrm{B}_k^{-1} \\
&= \mathrm{B}_k^{-1} - \frac{{}^t s_k\, s_k}{{}^t s_k\,\mathrm{B}_k^{-1} y_k}\,\mathrm{B}_k^{-1}\,\frac{y_k - \mathrm{B}_k s_k}{{}^t s_k\, s_k}\;{}^t s_k\,\mathrm{B}_k^{-1} \\
&= \mathrm{B}_k^{-1} - \frac{1}{{}^t s_k\,\mathrm{B}_k^{-1} y_k}\,\mathrm{B}_k^{-1}(y_k - \mathrm{B}_k s_k)\;{}^t s_k\,\mathrm{B}_k^{-1} \\
&= \mathrm{B}_k^{-1} - \frac{1}{{}^t s_k\,\mathrm{B}_k^{-1} y_k}\,(\mathrm{B}_k^{-1} y_k - s_k)\;{}^t s_k\,\mathrm{B}_k^{-1} \\
\mathrm{H}_{k+1} &= \mathrm{H}_k - \frac{1}{{}^t s_k\,\mathrm{H}_k y_k}\,(\mathrm{H}_k y_k - s_k)\;{}^t s_k\,\mathrm{H}_k
\end{aligned}$$

soit
$$\boxed{\mathrm{H}_{k+1} = \mathrm{H}_k - \frac{1}{{}^t s_k\,\mathrm{H}_k y_k}\,(\mathrm{H}_k y_k - s_k)\;{}^t s_k\,\mathrm{H}_k}$$