

TD 4: Méthodes itératives

lionel.rieg@ensiie.fr

1 Méthode de Cauchy

Exercice 1

Soit f la fonction de \mathbb{R}^2 dans \mathbb{R} définie par $f(x) = 2x_1^2 + x_2^2 - x_1x_2$. On cherche à minimiser f sur \mathbb{R}^2 . Pour cela, on part de $x^{(0)} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ et on va utiliser la méthode de Cauchy, ou méthode de la plus forte pente.

1. Calculer le gradient de f en $x^{(0)}$. On pose alors $d^{(0)} = -\nabla f(x^{(0)})$.
2. Déterminer $x^{(1)} = x^{(0)} + \lambda d^{(0)}$ avec λ qui minimise $f(x^{(1)})$. Comparer les valeurs de f en $x^{(0)}$ et $x^{(1)}$.
3. Recommencer les deux questions précédentes pour calculer $x^{(2)}$. Est-ce un point critique ?

2 Méthode du gradient conjugué

Exercice 2

Soit f la fonction de \mathbb{R}^2 dans \mathbb{R} définie par $f(x) = 2x_1^2 + x_2^2 - x_1x_2 + x_1 + 2x_2$. On cherche à minimiser f sur \mathbb{R}^2 . Pour cela, on part de $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ et on va utiliser la méthode du gradient conjugué.

1. Écrire f sous la forme $f(x) = \frac{1}{2}x^tQx + b^tx$.
2. Construire $x^{(1)} = x^{(0)} + \lambda d^{(0)}$ avec $d^{(0)}$ la direction opposé au gradient au point $x^{(0)}$ et λ qui minimise $f(x^{(1)})$.
3. On cherche maintenant à déterminer la direction suivante $d^{(1)}$ de descente.
 - (a) Donner la relation que doit satisfaire $d^{(1)}$ pour être une direction de descente de f au point $x^{(1)}$.
 - (b) Donner la relation que doit vérifier $d^{(1)}$ pour être conjuguée a $d^{(0)}$ par rapport à Q .
 - (c) Puisque $d^{(1)}$ n'est qu'une direction, sa norme n'importe pas et on peut donc fixer l'une de ses coordonnées (mais attention au signe!). Donner $d^{(1)}$ tel que sa second composante soit égale à 2 en valeur absolue.
4. Construire de même $x^{(2)} = x^{(1)} + \lambda d^{(1)}$ avec λ qui minimise $f(x^{(2)})$.
5. Est-il nécessaire de faire une étape supplémentaire ?

3 Méthode des directions conjuguées

Exercice 3

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

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

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

sont des directions conjuguées deux à deux par rapport à Q , *i.e.* $d^{(i)} \neq 0$ pour tout $0 \leq i \leq n-1$ et $d^{(j)} \cdot Q d^{(i)} = 0$ pour tout $0 \leq i < j \leq n-1$. Pour $d^{(i)} \neq 0$, on pourra démontrer par récurrence que $d^{(i)}$ appartient à l'espace vectoriel engendré par les $p^{(0)}, \dots, p^{(i)}$.

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

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

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

(c) Pour $k \geq 0$, soit $x^{(k+1)} = x^{(k)} + \alpha^{(k)} d^{(k)}$ où $\alpha^{(k)}$ minimise $\alpha \mapsto f(x^{(k)} + \alpha d^{(k)})$, α étant de signe quelconque. calculer $x^{(1)}$ et $x^{(2)}$ en partant de $x^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

(d) Refaire la question précédente en utilisant les directions $p^{(k)}$ plutôt que $d^{(k)}$.

Solutions

► Exercice 1

1. On a $\nabla f(x) = \begin{pmatrix} 4x_1 - x_2 \\ 2x_2 - x_1 \end{pmatrix}$. En $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$, cela donne $\nabla f(x^{(0)}) = \begin{pmatrix} 0 \\ 7 \end{pmatrix}$.
2. Déterminer le minimum selon une direction revient à dire que la dérivée directionnelle selon cette direction est nulle. On cherche donc $x^{(1)}$ tel que $\nabla f(x^{(1)}) \cdot d^{(0)} = 0$. On résout donc

$$\begin{pmatrix} 4(1) - (4 - 7\lambda) \\ 2(4 - 7\lambda) - (1) \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -7 \end{pmatrix} = 0 - 7(7 - 14\lambda) = 7 - 14\lambda = 0 \iff \lambda = \frac{1}{2}$$

On en tire $x^{(1)} = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}$. On a $f(x^{(0)}) = 2 + 16 - 4 = 14$ et $f(x^{(1)}) = 2 + \frac{1}{4} - \frac{1}{2} = \frac{7}{4}$. On constate que la valeur a effectivement baissée.

3. On calcule $\nabla f(x^{(1)}) = \begin{pmatrix} \frac{7}{2} \\ 0 \end{pmatrix}$. On prend donc $d^{(1)} = \begin{pmatrix} -\frac{7}{2} \\ 0 \end{pmatrix}$. En résolvant $\nabla f(x^{(2)}) \cdot d^{(1)} = 0$ avec $x^{(2)} = x^{(1)} + \lambda d^{(1)}$, on obtient $\lambda = \frac{1}{4}$ donc $x^{(2)} = \begin{pmatrix} \frac{1}{8} \\ \frac{1}{2} \end{pmatrix}$. Comme f est clairement quadratique et coercive, elle admet un unique point critique qui est un minimum global. Avec l'expression du gradient, on voit immédiatement qu'il s'agit de l'origine. En particulier, $x^{(2)} \neq 0$ n'est pas un point critique.

► Exercice 2

1. On a :

$$Q = \begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix} \quad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

2. le calcul du gradient donne : $\nabla f(x) = \begin{pmatrix} 4x_1 - x_2 + 1 \\ -x_1 + 2x_2 + 2 \end{pmatrix}$. En $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, cela donne $\nabla f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ dont l'opposé donne $d^{(0)}$. Déterminons le minimum de $\lambda \mapsto f(x^{(0)} + \lambda d^{(0)})$.

$$f(x^{(0)} + \lambda d^{(0)}) = f\left(\lambda \begin{pmatrix} -1 \\ -2 \end{pmatrix}\right) = 2\lambda^2 + 4\lambda^2 - 2\lambda^2 - \lambda - 4\lambda = 4\lambda^2 - 5\lambda$$

C'est un trinôme de second degré. Pour ces fonctions de la forme $x \mapsto ax^2 + bx + c$, le minimum est atteint en $-\frac{b}{2a}$ et vaut $\frac{-\Delta}{4a}$. Ici, cela donne $\lambda = \frac{5}{8}$ (et la valeur du minimum est $-\frac{25}{16}$). On a donc $x^{(1)} = \begin{pmatrix} -\frac{5}{8} \\ \frac{5}{4} \end{pmatrix}$.

3. (a) On veut que f décroisse, c'est-à-dire que le produit scalaire du gradient et de $d^{(1)}$ soit strictement négatif : $\nabla f(x^{(1)}) \cdot d^{(1)} < 0$. Si on pose $d^{(1)} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$, cela donne $(-\frac{5}{2} + \frac{5}{4} + 1)d_1 + (\frac{5}{8} - \frac{5}{2} + 2)d_2 = -\frac{1}{4}d_1 + \frac{1}{8}d_2 < 0$, i.e. $-2d_1 + d_2 < 0$.
- (b) La relation est $d^{(1)} \cdot Qd^{(0)} = 0$. Elle se développe en :

$$\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \begin{pmatrix} -2 \\ -3 \end{pmatrix} = -2d_1 - 3d_2 = 0 \iff 2d_1 + 3d_2 = 0 \iff d_1 = -\frac{3}{2}d_2$$

- (c) On doit choisir entre $d_2 = 2$ et $d_2 = -2$. Essayons avec $d_2 = 2$. La seconde équation nous donne $d_1 = -3$. L'inégalité devient alors $2 \cdot 3 + 2 < 0$ qui n'est pas vérifiée. On prend donc $d_2 = -2$, qui nous donne $d_1 = 3$ et l'inégalité est vérifiée. On a donc $d^{(1)} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$.

4. On veut minimiser $\lambda \mapsto f(x^{(1)} + \lambda d^{(1)})$. Pour cela, on annule la dérivée directionnelle selon $d^{(1)}$ de f en $x^{(2)}$: $\nabla f(x^{(2)}) \cdot d^{(1)} = 0$. Cela nous donne :

$$\nabla f(x^{(2)}) \cdot d^{(1)} = \begin{pmatrix} 14\lambda - \frac{1}{4} \\ -7\lambda + 1\frac{1}{8} \end{pmatrix} \cdot \begin{pmatrix} 3 \\ -2 \end{pmatrix} = 42\lambda - \frac{3}{4} + 14\lambda - \frac{1}{4} = 56\lambda - 1$$

On a donc $x^{(2)} = \begin{pmatrix} -\frac{5}{8} + \frac{1}{56}3 \\ -\frac{1}{4} - \frac{1}{56}2 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$.

5. On regarde si $\begin{pmatrix} -\frac{4}{7} \\ -\frac{6}{7} \end{pmatrix}$ est un point critique :

$$\nabla f(x^{(1)}) = \begin{pmatrix} -4\frac{4}{7} + \frac{9}{7} + 1 \\ \frac{4}{7} - 2\frac{9}{7} + 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Oui, donc il n'y a pas besoin de faire une étape supplémentaire.

► Exercice 3

1. Montrons tout d'abord que les $d^{(k)}$ sont non nuls. On a $d^{(0)} = p^{(0)} \neq 0$. Supposons que $d^{(k+1)}$ soit nul. Posons $c_i = \frac{p^{(k+1)} \cdot Qd^{(i)}}{d^{(i)} \cdot Qd^{(i)}}$. On a alors $d^{(k+1)} = p^{(k+1)} - \sum_{i=1}^k c_i p^{(i)}$. Ainsi $d^{(k+1)} = 0$ donne $p^{(k+1)} - \sum_{i=1}^k c_i p^{(i)} = 0$, ce qui signifie que les $(p^{(i)})_{0 \leq i \leq k+1}$ sont liés. C'est absurde par hypothèse. Montrons à présent que les directions sont orthogonales par rapport à Q . Pour $d^{(1)}$ (*i.e.* $j = 1$), cela donne

$$d^{(1)} \cdot Qd^{(0)} = \left(p^{(1)} - \frac{p^{(1)} \cdot Qd^{(0)}}{d^{(0)} \cdot Qd^{(0)}} d^{(0)} \right) \cdot Qd^{(0)} = p^{(1)} \cdot Qd^{(0)} - \frac{p^{(1)} \cdot Qd^{(0)}}{d^{(0)} \cdot Qd^{(0)}} d^{(0)} \cdot Qd^{(0)} = p^{(1)} \cdot Qd^{(0)} - p^{(1)} \cdot Qd^{(0)} = 0$$

Pour l'étape inductive, soit $0 \leq i < k + 1$. Par hypothèse d'induction, on sait que $p^{(i)} \cdot Qd^{(j)}$ pour tout $i \leq k$ tel que $i \neq j$ (en utilisant la symétrie du produit scalaire lorsque $i < j$).

$$\begin{aligned} d^{(k+1)} \cdot Qd^{(i)} &= \left(p^{(k+1)} - \sum_{j=0}^k \frac{p^{(k+1)} \cdot Qd^{(j)}}{d^{(j)} \cdot Qd^{(j)}} d^{(j)} \right) \cdot Qd^{(i)} \\ &= p^{(k+1)} \cdot Qd^{(i)} - \sum_{j=0}^k \frac{p^{(k+1)} \cdot Qd^{(j)}}{d^{(j)} \cdot Qd^{(j)}} d^{(j)} \cdot Qd^{(i)} \\ &= p^{(k+1)} \cdot Qd^{(i)} - \frac{p^{(k+1)} \cdot Qd^{(i)}}{d^{(i)} \cdot Qd^{(i)}} d^{(i)} \cdot Qd^{(i)} \quad \text{par HI} \\ &= p^{(k+1)} \cdot Qd^{(i)} - p^{(1)} \cdot Qd^{(i)} = 0 \end{aligned}$$

2. (a) On l'a déjà vu pas mal : $Q = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$.

(b) On a $d^{(0)} = p^{(0)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et

$$d^{(1)} = p^{(1)} - \frac{p^{(1)} \cdot Qd^{(0)}}{d^{(0)} \cdot Qd^{(0)}} d^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{4}{4} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

- (c) Calculons $\alpha^{(0)}$ en étudiant $f_0 : \alpha \mapsto f(x^{(0)} + \alpha d^{(0)}) = 2(1 + \alpha)^2 + 1$. On a $f'_0(\alpha) = 4\alpha + 4$ donc $f'_0(\alpha) = 0 \iff \alpha = -1$. D'où $\alpha^{(0)} = -1$ et donc $x^{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. De même, en étudiant $f_2 : \alpha \mapsto f(x^{(1)} + \alpha d^{(1)}) = (1 + \alpha)^2$, on trouve $\alpha^{(2)} = -1$. D'où $x^{(3)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ qui est le minimum de f .
- (d) Puisque $d^{(0)} = p^{(0)}$, les valeurs de $\alpha^{(0)}$ et $x^{(1)}$ sont les mêmes. Pour la seconde étape, $f_1(\alpha) = \alpha^2 + (1 + \alpha)^2$ d'où l'on tire $\alpha^{(1)} = -\frac{1}{2}$ puis $x^{(3)} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.