Optimisation convexe — TDs

Ivan Lejeune

7 février 2025

Table des matières

TD1 — Algorithmes 1D

Exercice 1.1 Minimisation d'une fonction par dichotomie. Soit $f \in \mathcal{C}^0([a,b],\mathbb{R})$. On dit que f est **unimodale** sur l'intervalle $[a,b]$ si il existe un point $\overline{x} \in [a,b]$ tel que f soit strictement décroissante sur $[a,\overline{x}[$ et strictement croissante sur $]\overline{x},b]$.
Pour chercher $\overline{x}$, nous allons générer une suite strictement décroissante d'intervalles dont le diamètre tend vers zéro et qui encadrent le minimum cherché.
Supposons connus cinq points $a = x_1, x_2, x_3, x_4, x_5 = b$. Cinq situations sont possibles :

(i) Si $f(x_1) < f(x_2) < f(x_3) < f(x_4) < f(x_5)$, alors $\overline{x} \in]x_1, x_2[$.

(ii) Si $f(x_1) > f(x_2)$ et $f(x_2) < f(x_3) < f(x_4) < f(x_5)$, alors $\overline{x} \in]x_1, x_3[$.

(iii) Si $f(x_1) > f(x_2) > f(x_3)$ et $f(x_3) < f(x_4) < f(x_5)$, alors $\overline{x} \in]x_2, x_4[$.

(iv) Si $f(x_1) > f(x_2) > f(x_3) > f(x_4)$ et $f(x_4) < f(x_5)$, alors $\overline{x} \in]x_3, x_5[$.

(v) Si $f(x_1) > f(x_2) > f(x_3) > f(x_4) > f(x_5)$, alors $\overline{x} \in]x_4, x_5[$.

(a) Utiliser ces propriétés pour construire un algorithme permettant de génèrer une suite d'intervalles $([a_k,b_k])_{k\in\mathbb{N}}$ telle que
- à chaque étape, $\overline{x} \in [a_k, b_k]$,
- on a $b_k - a_k = \frac{b_{k-1}-a_{k-1}}{2}$,
- à partir de la deuxième étape, 2 évaluations de f sont nécessaires à chaque étape.

(b) Montrer que $a_k \to \overline{x}$ et $b_k \to \overline{x}$ lorsque $k \to \infty$.

Solution. (a) Dans le cas général on prend alors $[a_k,b_k] = [x_i,x_j]$ tel que $\overline{x} \in [x_i,x_j]$. Après la première étape, on évalue f en $x_t = (a_k+b_k)/2$. Si $f(x_t) < f(x_i)$, on choisit $[a_{k+1},b_{k+1}] = [a_k,x_t]$, sinon on choisit $[a_{k+1},b_{k+1}] = [x_t,b_k]$. On répète ainsi de suite. On a bien $b_k - a_k = \frac{b_{k-1}-a_{k-1}}{2}$ et on a besoin de deux évaluations de f à chaque étape. Evidemment, on a toujours $\overline{x} \in [a_k,b_k]$.

- Cas (i) et (ii) : $\overline{x} \in [x_1,x_3]$.
- Cas (iii) : $\overline{x} \in [x_2,x_4]$.
- Cas (iv) et (v) : $\overline{x} \in [x_3,x_5]$.

On peut reprendre 5 points x_1,x_2,x_3,x_4,x_5 comme les bords et les quartiles d'un intervalle $[a,b]$.

(b) Comme l'intervalle $[a_k,b_k]$ est de longueur divisée par 2 à chaque étape, on a $a_k - b_k$ qui tend vers 0 lorsque $k \to \infty$. Comme $\overline{x} \in [a_k,b_k]$, on a bien $a_k \to \overline{x}$ et $b_k \to \overline{x}$ lorsque $k \to \infty$.

Exercice 1.2 Méthode de la section dorée. Nous reprenons le principe de la méthode de la dichotomie précédente mais à chaque itération, nous allons maintenant chercher à diviser l'intervalle d'approximation en 3 parties (au lieu de 4 pour la dichotomie).
Plus précisément, nous allons construire une suite décroissante d'intervalles $([a_k,b_k])_{k\in\mathbb{N}}$ qui contiennent tous le minimum $\overline{x}$ cherché. Pour passer de $[a_k,b_k]$ à $[a_{k+1},b_{k+1}]$, on introduit deux nombres x_2^k et x_3^k tels que $a_k < x_2^k < x_3^k < b_k$. On calcule alors les valeurs de f en x_2^k et x_3^k. On a alors 2 cas :

- Si $f(x_2^k) \le f(x_3^k)$, alors $\overline{x} \in \left[a_k, x_3^k\right]$.
- Si $f(x_2^k) > f(x_3^k)$, alors $\overline{x} \in \left[x_2^k, b_k\right]$.

La question suivante se pose alors : comment choisir x_2^k et x_3^k en pratique ? On privilégie deux aspects :

(i) On souhaite que le facteur de réduction γ, qui représente le ratio de la longueur du nouvel intervalle, noté L_{k+1} par rapport à la longueur du précédent, noté L_k soit constant :

$$\frac{L_{k+1}}{L_k} = \gamma.$$

(ii) On désire, comme pour la méthode de la dichotomie, réutiliser le point qui n'a pas été choisi dans l'itération précédente afin de diminuer les coûts de calcul.

(a) Traduire ces contraintes permettant de choisir $x_2^k, x_3^k, a_{k+1}, b_{k+1}$. Proposer un algorithme et montrer qu'il n'y a qu'une seule valeur possible pour γ.

(b) Montrer que pour tout $k \in \mathbb{N}$, on a

$$b_k - a_k = \gamma^k (b - a).$$

Conclure.

Solution. vide

TD2 — Optimisation sans contraintes

Exercice 2.1 Quelques révisions.

(a) Soit $A \in S_n(\mathbb{R})$, montrer que

$$A \in S_n^+(\mathbb{R}) \iff \mathsf{Sp}(A) \subset \mathbb{R}_+$$

(b) Soit $A \in S_n(\mathbb{R})$, montrer que

$$A^2 \in S_n^+(\mathbb{R})$$

(c) Soit $A \in S_n(\mathbb{R})$. On note $\lambda_1 \leq \cdots \leq \lambda_n$ ses valeurs propres. Montrer que pour tout $x \in \mathbb{R}^n$

$$x^T A x \geq \lambda_1 \|x\|^2$$

(d) Soit $A \in S_n^+(\mathbb{R})$. Montrer qu'il existe $B \in S_n^+(\mathbb{R})$ tel que $A = B^2$.

Solution.

(a) Montrons d'abord le sens direct. On suppose que $A \in S_n^+(\mathbb{R})$. Soit $\lambda \in \mathbb{R}$ une valeur propre de A et v le vecteur propre associé. On a alors

$$v^T A v = \lambda \|v\|^2 \geq 0$$

car $A \in S_n^+(\mathbb{R})$. Donc $\lambda \geq 0$.
Réciproquement, supposons que $\mathsf{Sp}(A) \subset \mathbb{R}_+$, elle est alors diagonalisable dans une base orthonormée. Soit $v_1, \ldots, v_n$ une base orthonormée de vecteurs propres de A. On a alors

$$v = \sum_{i=1}^n \alpha_i v_i$$

et

$$v^T A v = v^T \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \lambda_i \alpha_i^2 \geq 0$$

car $\lambda_i \geq 0$. Donc $A \in S_n^+(\mathbb{R})$.

(b) Soit $A \in S_n(\mathbb{R})$. Alors A est diagonalisable et on note $A = PDP^{-1}$ avec D diagonale et P inversible. On a alors

$$A^2 = PD^2P^{-1}$$

et D^2 est diagonale avec des éléments λ_i^2. Donc tous ses éléments sont positifs $\mathsf{Sp}(A^2) \subset \mathbb{R}^+$ et donc $A^2 \in S_n^+(\mathbb{R})$ (voir question précédente).

(c) Soit $A \in S_n(\mathbb{R})$. On a $A = PDP^{-1}$ avec D diagonale et P inversible. On a alors

$$x^T A x = x^T PDP^{-1} x = (P^{-1}x)^T D (P^{-1}x) = \sum_{i=1}^n \lambda_i \alpha_i^2$$

avec $\alpha = P^{-1}x$. Comme $\lambda_1 \leq \cdots \leq \lambda_n$, on a

$$x^T A x = \sum_{i=1}^n \lambda_i \alpha_i^2 \geq \lambda_1 \sum_{i=1}^n \alpha_i^2 = \lambda_1 \|x\|^2$$

(d) Soit $A \in S_n^+(\mathbb{R})$ qu'on écrit $A = PDP^{-1}$ avec D diagonale et P inversible. On pose

$$B = P \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} P^{-1}$$

qui existe car $\lambda_i \geq 0$ et donc B est bien définie. On a alors

$$B^2 = P \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} P^{-1} P \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} P^{-1} = P \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} P^{-1} = A$$

Exercice 2.2. On définit la fonction

$$\begin{aligned} J : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto y^4 - 3xy^2 + x^2 \end{aligned}$$

(a) Déterminer les points critiques de J.

(b) Soit $d = (d_1, d_2) \in \mathbb{R}^2$. Utiliser l'application

$$\begin{aligned} \mathbb{R} &\to \mathbb{R} \\ t &\mapsto J(td_1, td_2), \end{aligned}$$

pour montrer que $(0, 0)$ est un minimum local le long de toute droite passant par $(0, 0)$.

(c) Le point $(0, 0)$ est-il un minimum local de la restriction de J à la parabole d'équation $x = y^2$?

(d) Calculer J'' et déterminer la nature du point critique $(0, 0)$.

Solution.

(a) On a

$$\begin{aligned} \frac{\partial J}{\partial x}(x, y) &= 2x - 3y^2, \\ \frac{\partial J}{\partial y}(x, y) &= 4y^3 - 6xy. \end{aligned}$$

Les points critiques sont les solutions du système

$$\begin{cases} 2x - 3y^2 &= 0, \\ 4y^3 - 6xy &= 0. \end{cases}$$

On résout le système pour obtenir $x = 0$ et $y = 0$. Donc le seul point critique est $(0, 0)$.

(b) On a

$$\begin{aligned} J(td_1, td_2) &= (td_2)^4 - 3(td_1)(td_2)^2 + (td_1)^2 \\ &= t^4 d_2^4 - 3t^3 d_1 d_2^2 + t^2 d_1^2 \\ &= t^2 (t^2 d_2^4 - 3t d_1 d_2^2 + d_1^2). \end{aligned}$$

On a donc

$$J(td_1, td_2) = t^2 J(d_1, d_2).$$

Donc $J(td_1, td_2)$ est un polynôme de degré 2 en t et donc son minimum est atteint en $t = 0$. Donc $(0, 0)$ est un minimum local le long de toute droite passant par $(0, 0)$. ATTENTION etudier la double dérivée pour montrer sa convexité (ou passer par un equivalent) qui montre que c'est un minimum local

(c) On a $x = y^2$ et donc

$$J(y^2, y) = y^4 - 3y^4 + y^4 = -y^4.$$

Donc $(0, 0)$ est un maximum local de la restriction de J à la parabole d'équation $x = y^2$.

(d) On calcule la matrice hessienne de J :

$$\frac{\partial^2 J}{\partial x^2}(x,y) = 2,$$
$$\frac{\partial^2 J}{\partial y^2}(x,y) = 12y^2 - 6x,$$
$$\frac{\partial^2 J}{\partial xy}(x,y) = -6y.$$

Donc

$$J''(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

On a une valeur propre nulle donc on ne peut pas conclure sur la nature du point critique $(0,0)$ directement avec la matrice hessienne. Mais en utilisant les questions (b) et (c), on peut dire que $(0,0)$ est un point selle.