Université Claude Bernard - Lyon 1
M2 Statistique, Modélisation et Science des données
9 décembre 2024

Optimisation Convexe : Algorithmes et Applications en Apprentissage

Contrôle terminal. Durée : 2h. Tous les documents sont autorisés, mais pas les objets connectés et les moyens de communication. Le barème dépasse largement 20, il est conseillé de ne pas traiter tous les exercices.
Dans tout le sujet ci-dessous, $||x||$ désigne la norme euclidienne : $||x|| := (\sum_i |x_i|^2)^{1/2}$.

Exercice 1 (8 points)**.** Soit $f : \mathbb{R}^n \to \mathbb{R}$ la fonction définie par

$$f(x) := \begin{cases} ||x|| & \text{si } ||x|| \leq 1, \\ +\infty & \text{sinon.} \end{cases}$$

Calculer f^*, ainsi que ∂f et ∂f^* en tout point.

Solution : On a

$$f^*(y) = \sup_{x\in\mathbb{R}^n} x\cdot y - f(x) = \sup_{x\in B} x\cdot y - ||x||,$$

où B est la boule fermé centrée en l'origine et de rayon 1. On voit tout de suite que le sup dans la définition est un max (parce que la boule est compacte). Si $y = 0$ on doit maximiser $-||x||$ et le max vaut 0. Si $y \neq 0$ on voit que le maximum est atteint par un vecteur x orienté comme y, c'est-à-dire de la forme $x = ty/||y||$ avec $t \leq 1$ parce que changer de direction ne change pas la norme mais la direction optimale pour le produit scalaire est celle de y. On doit donc calculer $\max_{t\in[0,1]} t||y||-t$. Il s'agit donc de maximiser une fonction affine de t et le max est atteint au bord. On a donc $f^*(y) = \max\{||y|| - 1, 0\}$ en considérant les deux possibilité $t = 0, 1$. Donc la formule générale est $f^*(y) = (||y|| - 1)_+$.

La fonction f^* est différentiable en tout point y_0 tel que $||y_0|| \neq 1$ et son gradient vaut $y_0/||y_0||$ si $||y_0|| > 1$ et 0 si $||y_0|| < 1$. Pour calculer le sous-différentiel en y_0 tel que $||y_0|| = 1$ on utilise la définition : on cherche les vecteurs p tels que

$$||y|| - 1 \geq 0 + p\cdot(y - y_0) \text{ pour tout } y \text{ tel que } ||y|| \geq 1, \quad 0 \geq 0 + p\cdot(y - y_0) \text{ pour tout } y \text{ tel que } ||y|| \leq 1.$$

La deuxième condition nous dit que p est vecteur normal sortant de la boule B au point y_0, donc $p = ty_0$ pour $t \geq 0$. La première, en prenant $y = y_0 + p$, et en utilisant $||y|| \leq ||y_0|| + ||p|| = 1 + ||p||$, nous donne $||p|| \geq ||p||^2$, donc $||p|| \leq 1$. On trouve donc un vecteur p de la forme $p = ty_0$ avec $t \in [0, 1]$. On peut vérifier que tous ces vecteurs satisfont les conditions demandées. On a donc

$$\partial f^*(y_0) = \begin{cases} \{0\} & \text{si } ||y_0|| < 1, \\ \{ty_0 \ : t \in [0,1]\} & \text{si } ||y_0|| = 1, \\ \{y_0/||y_0||\} & \text{si } ||y_0|| > 1. \end{cases}$$

La fonction f est différentiable en tout point x_0 tel que $0 < ||x_0|| < 1$ et son gradient vaut $x_0/||x_0||$.
En $x_0 = 0$ son soudifférentiel est l'ensemble des vecteurs p tels que $||x|| \geq 0 + (x - 0)\cdot p$ pour tout $x \in B$, et donc l'ensemble des vecteurs p tels que $||p|| \leq 1$.
Pour x_0 avec $||x_0|| > 1$ la foncion f vaut infini et son sousdifférentiel est vide.
Pour x_0 avec $||x_0|| = 1$ on cherche les vecteurs p tels que $||x|| \geq 1 + (x - x_0)\cdot p$ pour tout $x \in B$. En prenant un autre vecteur x avec $||x|| = 1$ on trouve $p\cdot(x - x_0) \leq 0$ et donc p doit faire un angle obtus avec tout vecteur de la forme $x - x_0$ avec $x \in \partial B$. Cela implique que p est un vecteur normal sortant en x_0 et donc $p = tx_0$ avec $t \geq 0$. D'autre part, en prenant $x = (1 - \varepsilon)x_0$ on trouve $1 - \varepsilon \geq 1 - \varepsilon x_0 \cdot p$ et donc $p\cdot x_0 \geq 1$, ce qui implique $t \geq 1$. On voit bien que tout vecteur de ce type satisfait la condition pour être dans le

sous-différentiel et on trouve donc

$$\partial f(x_0) = \begin{cases} B & \text{si } x_0 = 0, \\ \{x_0/||x_0||\} & \text{si } 0 < ||x_0|| < 1, \\ \{tx_0 \ : \ t \geq 1\} & \text{si } ||x_0|| = 1, \\ \emptyset & \text{si } ||x_0|| > 1. \end{cases}$$

Exercice 2 (5 points)**.** Si $g : \mathbb{R} \to \mathbb{R}$ est une fonction C^2, considérons la fonction $f : \mathbb{R}^n \to \mathbb{R}$ donnée par $f(x) = g(||x||)$. En tout point $x \in \mathbb{R}^n \setminus \{0\}$ écrire le gradient de f et la Hessienne de f. Exprimer la Hessienne dans une base orthonormée où un des vecteurs de la base est parallèle à x. Donner une condition sur g pour que la fonction f soit convexe et pour qu'elle soit elliptique, et calculer en termes de g sa constante α d'ellipticité. De même, donner une condition sur g pour que ∇f soit Lipschitzien et calculer en termes de g la constante L de Lipschitz de ∇f.

Solution : Pour $x \neq 0$ on a bien $\nabla f(x) = g'(||x||)x/||x||$. En dérivant une autre fois on obtient

$$D^2 f(x) = \frac{g'(||x||)}{||x||} I + \frac{g''(||x||)}{||x||} x \otimes \frac{x}{||x||} - \frac{g'(||x||)}{||x||^2} x \otimes \frac{x}{||x||}.$$

Si on choisit une base orthonormée où $e_1 = x/||x||$ (donc $x = ||x||e_1$) on a donc que $D^2 f$ est une matrice diagonale dont la composante $(D^2 f)_{11}$ vaut $\frac{g'(||x||)}{||x||} + g''(||x||) - \frac{g'(||x||)}{||x||} = g''(||x||)$ et dont les autres composantes diagonales valent $\frac{g'(||x||)}{||x||}$. Les valeurs propres de $D^2 f$ sont donc $g''(||x||)$ et $\frac{g'(||x||)}{||x||}$. Pour que f soit convexe il faut qu'elles soient toutes positives ou nulles, donc il faut $g'', g' \geq 0$ (g doit être convexe et croissante). Pour que f soit α-elliptique il faut que toutes ces valeurs propres soient plus grandes d'α, donc $g'' \geq \alpha$ et $g'(s) \geq \alpha s$ pour tout s. On peut aussi dire $g'' \geq \alpha$ et $g' \geq 0$ puisque la condition sur g'', intégrée, donne aussi $g'(s) \geq \alpha s + g'(0) \geq \alpha s$.
Pour le comportement Lipschitzien de ∇f il faut regarder la plus grande valeur propre de $D^2 f$. On trouve alors

$$L := \sup_s \max\{|g''|(s), \frac{|g'|(s)}{s}\}.$$

Le gradient de f est Lipschitzien si et seulement si ce sup est fini. Cela implique en particulier $g'(0) = 0$. Ensuite, par le même argument que précédemment, on peut prendre $L = \sup |g''|$ puisque, en intégrant, on aura aussi $|g'(s)| \leq Ls$.

Exercice 3 (6 points)**.** Étant donné un point $y \in \mathbb{R}^n$, deux nombres $\tau, R > 0$ et une matrice $A \in \mathbb{R}^{n \times n}$ symétrique et définie positive on considère les itérations suivantes : en partant d'un point $x_0 \in \mathbb{R}^n$ on définit

$$\tilde{x}_{k+1} := x_k - \tau A(x_k - y), \quad x_k := \frac{\tilde{x}_{k+1}}{\max\{1, ||\tilde{x}_{k+1}||/R\}}.$$

Soient $0 < \lambda_1 \leq \cdots \leq \lambda_n$ les valeurs propres de A.

1. Prouver que si l'on a $\tau < 2\lambda_n^{-1}$ alors la suite $(x_k)_k$ converge vers un point x_∞ et le caractériser comme solution d'un problème d'optimisation. Sous les mêmes hypothèses, peut-on dire que $\tilde{x}_k$ a une limite aussi ? Dans quels cas a-t-on $x_\infty = y$?
2. Donner une estimation précise de la vitesse de convergence de la suite $(x_k)_k$.
3. Si $\tau = 2\lambda_n^{-1}$ et $||y|| < R$, prouver que la suite $(x_k)_k$ n'admet pas forcément une limite.

Solution : La suite $(x_k)_k$ est la suite donnée par l'algorithme de gradient projeté pour la fonction $f(x) = \frac{1}{2}A(x-y)\cdot(x-y)$ avec la contrainte $x \in \overline{B(0,R)}$. La Hessienne de cette fonction satisfait $\lambda_1 I \leq D^2 f \leq \lambda_n I$ et l'algorithme converge donc dès que $\tau < 2/\lambda_n$. La limite de x_k est donc le point $x_\infty = \operatorname{argmin}_{x \in \overline{B(0,R)}} f(x)$.

Ce point est égal à y si et seulement si $y \in \overline{B(0,R)}$, donc $||y|| \le R$. En ce qui concerne la suite $\tilde{x}_k$, sa limite est donnée par $x_\infty - \tau A(x_\infty - y)$. La suite converge a vitesse exponentielle, c'est-à-dire

$$||x_k - x_\infty|| \le r^k ||x_0 - x_\infty||,$$

où $r = \max\{1 - \tau\lambda_1, \tau\lambda_n - 1\}$.

Si on prend $\tau = 2\lambda_n^{-1}$ et $||y|| < R$, on peut prendre un vecteur v tel que $Av = \lambda_n v$ et $x_0 = y + \varepsilon v$. Pour ε assez petit on a $B(y, \varepsilon||v||) \subset \overline{B(0,R)}$. On a ensuite $\tilde{x}_1 = x_0 - \tau A(\varepsilon v) = x_0 - \tau\lambda_n\varepsilon v = x_0 - 2\varepsilon v = y - \varepsilon v$. Comme $y - \varepsilon v \in \overline{B(0,R)}$ on trouve $x_1 = \tilde{x}_1$. Ensuite on a $\tilde{x}_2 = x_1 - \tau A(-\varepsilon v) = y + \varepsilon v = x_0$ et à nouveau $x_2 = \tilde{x}_2$. Donc, la suite alterne entre les points $y + \varepsilon v$ et $y - \varepsilon v$ et n'a pas de limite.

Exercice 4 (8 points)**.** Étant donnés une matrice $A \in \mathbb{R}^{m\times n}$, un vecteur $y \in \mathbb{R}^m$, et des coefficients strictement positifs $(p_i)_{i=1,\dots,n}$, on s'intéresse au problème d'optimisation suivant :

$$\min\{||Ax - y||^2 + \sum_{i=1}^n p_i|x_i| \quad : \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n\}.$$

1. Prouver que ce problème admet au moins une solution.
2. Prouver que si x et $\tilde{x}$ sont deux solutions de ce problème d'optimisation, alors $Ax = A\tilde{x}$ et que la solution est donc unique si le noyau de A est réduit au seul 0.
3. Si $n = m$ et la matrice A est diagonale de coefficients diagonaux $(\lambda_1, \dots, \lambda_n)$ trouver l'unique solution optimale x en termes de y, $(\lambda_i)_i$ et $(p_i)_i$.
4. Suggérer un algorithme pour résoudre ce problème, en décrivant explicitement ses différentes étapes. Discuter sa vitesse de convergence selon les valeurs propres de la matrice A^tA.

Solution :

1. En définissant $f(x) := ||Ax-y||^2 + \sum_{i=1}^n p_i|x_i|$ on voit que f est coercive, puisque $f(x) \ge c\max\{|x_i| : i = 1, \dots, n\}$ (où $c = \min_i p_i$). Comme f est également continue, elle admet un minimiseur.
2. Si x et $\tilde{x}$ sont deux solutions de ce problème d'optimisation, soit $x_m := (x + \tilde{x})/2$ le milieu entre x et $\tilde{x}$. On écrit $f = g + h$ où $g(x) := ||Ax - y||^2$ et $h(x) := \sum_{i=1}^n p_i|x_i|$. Tant g que h sont convexes, donc $h(x_m) \le (h(x) + h(\tilde{x}))/2$ et $g(x_m) \le (g(x) + g(\tilde{x}))/2$. De plus, comme la fonction $u \mapsto ||u - y||^2$ est strictement convexe, on a l'inégalité stricte $g(x_m) < (g(x) + g(\tilde{x}))/2$ sauf si $Ax = A\tilde{x}$. On en déduit que la solution est unique si le noyau de A est réduit au seul 0.
3. Si $n = m$ et la matrice A est diagonale de coefficients diagonaux $(\lambda_1, \dots, \lambda_n)$ la fonction à minimiser est de la forme

$$f(x) := \sum_{i=1}^n \left(|\lambda_i x_i - y_i|^2 + p_i|x_i|\right).$$

Il s(agit donc d'une fonction séparable, somme de n fonctions chacune portant sur une seule composante x_i. On peut donc les minimiser séparément, et pour cela il faut savoir minimiser la fonction $x \mapsto |\lambda x - y|^2 + p|x|$. Si $\lambda = 0$ on voit bien que le minimiseur est $x = 0$. Si $\lambda \ne 0$ la minimisation est équivalente à minimiser

$$x \mapsto \frac{|x - \lambda^{-1}y|^2}{2 \cdot \frac{1}{2}p|\lambda|^{-2}} + |x|.$$

Il s'agit donc de calculer l'opérateur proximal $\mathrm{Prox}_{\alpha,\ell}$ de la fonction ℓ définie par $\ell(x) := |x|$ pour un parametre $\alpha := \frac{1}{2}p|\lambda|^{-2}$ et l'appliquer au point $\lambda^{-1}y$. On trouve donc

$$x = \mathrm{Prox}_{\alpha,\ell}[\lambda^{-1}y] = \mathrm{sign}(\lambda^{-1}y)\left(|\lambda^{-1}y| - \frac{1}{2}p|\lambda|^{-2}\right)_+ = \frac{\mathrm{sign}(\lambda)\mathrm{sign}(y)}{\lambda^2}\left(|\lambda y| - \frac{1}{2}p\right)_+.$$

On note en passant que pour $\lambda \to 0$ cette expression donne bien $x = 0$

Finalement, le minimiseur serait le vecteur $c = (x_1, \dots, x_n)$ dont les cordonnées sont données par

$$x_i = \begin{cases} \frac{\mathrm{sign}(\lambda_i)\mathrm{sign}(y_i)}{\lambda_i^2}\left(|\lambda_i y_i| - \frac{1}{2}p_i\right)_+, & \text{si } \lambda_i \ne 0, \\ 0 & \text{si } \lambda_i = 0. \end{cases}$$

4. On peut par exemple résoudre ce problème avec l'algorithme de gradient proximal en utilisant g comme partie lisse et h comme partie non-lisse. Le calcul de l'opérateur proximal de h correspond au calcul précédent lorsque l'on doit minimiser $x \mapsto \frac{1}{2\alpha}||x-y||^2+\sum_{i=1}^n p_i|x_i| = \sum_{i=1}^n \left(\frac{1}{2\alpha}|x_i-y_i|^2+p_i|x_i|\right)$. Il suffit de prendre $\lambda_i = 1$ et multiplier les valeurs des p_i par 2α dans la formule calculée précédemment, en obtenant donc
$$(\text{Prox}_{\alpha,h}[y])_i = \text{sign}(y_i)(|y_i| - \alpha p_i)_+.$$
On obtient donc la suite suivante
$$x_{k+1} = \text{Prox}_{\tau,h}[x_k - 2\tau A^t(Ax_k - y)],$$
où on a utilisé le calcul de ∇g qui donne $\nabla g(x) = 2A^t(Ax - y)$. L'algorithme converge à condition de prendre $\tau < 2/L$ où L est la plus grande valeur propre de $D^2 g = 2A^t A$. Il faut donc que τ soit plus petit que l'inverse de la plus grade valeur propre de $A^t A$ (qui est aussi le carré de ce qu'on appelle le rayon spectral de A). Si, de plus, $A^t A$ est définie positive (donc sa plus petite valeur propre est strictement positive) la convergence est expenentielle. Sinon, on obtient juste $f(x_k) - \min f \leq C/k$.
Il est aussi possible d'accélerer l'algorithme en utilisant le gradient accéléré de Nesterov ($x_{k+1} = \text{Prox}_{\tau,h}[\tilde{x}_k - 2\tau A^t(A\tilde{x}_k - y)]$, où $\tilde{x}_k$ est une combinaison linéaire opportune de x_k et x_{k-1}) ou d'utiliser un algorithme de sous-gradient aec des pas $\tau_k \to 0$.