Solutionnaire 1

Frederic Boileau

28th February 2019

J'introduis quelques notations. En analyse convexe on dénote l'ensemble des fonctions convexes *propres* sur un espace euclidien $\mathbb{E}$ (dans notre cas simplement $\mathbb{R}$) par $\Gamma(\mathbb{E})$.

$$\Gamma \triangleq \{f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}\}$$

Une fonction *propre* est une fonction pour laquelle il existe au moins un x t.q. $f(x) < +\infty$ et pour laquelle pour tout x $f(x) > -\infty$. On écarte ainsi certains cas pathologiques.

J'utilise la notation de Evans pour la dérivation. Ainsi $D_x f$ est la dérivée de f par rapport à x, si c'est un nombre, un vecteur (comme pour le gradient) ou une matrice (jacobien) dépend du contexte (de f et de x).

Soit A une matrice. Si A est définie positive (resp. semi-définies positives) on écrit $A \succ 0$ (resp $A \succeq 0$).

1 Direction de Descente

Soit f une fonction convexe différentiable sur $\mathbb{R}^n$ et x et y deux points de $\mathbb{R}^n$ t.q. $f(y) < f(x)$. Montrer que $y - x$ est une direction de descente de f en x. Notons qu'aucun ordre n'est imposé sur x et y.

La preuve découle directement de la définition de la convexité donnée en cours:

$$\begin{aligned}
f(\alpha y + (1-\alpha)x) &\leq \alpha f(y) + (1-\alpha) f(x) \\
f(x + \alpha(y - x)) &\leq \alpha f(y) + (1-\alpha) f(x) \\
f(x + \alpha(y - x)) &< \alpha f(y) + (1-\alpha) f(x) \\
f(x + \alpha(y - x)) &< f(x)
\end{aligned}$$

2 Système linéaire et optimisation

Prouvez que tout système linéaire $Ax = b$, avec A définie positive, peut se reformuler comme un problème d'optimisation sans contrainte.

Pour la reformulation nous avons plusieurs critères; premièrement on veut que le minimum global corresponde à la solution, dans un second temps si le problème est convexe nous avons des garanties quant à la recherche de ce minimiseur global. Il y a deux formulations simples, je traite les deux.

2.1 Forme Quadratique

Le reformulation la plus simple est de prendre la forme quadratique suivante:

$$f(x) = \frac{1}{2}x^{\mathsf{T}}Ax - x^{\mathsf{T}}b \tag{1}$$

Le calcul de dérivés par rapport à des vecteurs ou des matrices peut être laborieux mais l'identité suivante est très importante:

$$D_x(x^{\mathsf{T}}Ax) = x^{\mathsf{T}}(A + A^{\mathsf{T}}) \tag{2}$$

Ainsi nous avons:

$$D_x(f) = Ax - b \qquad D_x^2(f) = A \succ 0 \tag{3}$$

La première équation montre que le gradient de la fonction est nul ssi le x est une solution du système linéaire. La deuxième équation montre que la deuxième dérivée (le hessien) de f par rapport est *définie positive* (par supposition sur A). La fonction est donc strictement convexe et atteint son minimum ssi le gradient est nulle, ce qui n'arrive que lorsque $Ax = b$.

Notons que pour une fonction à une variable $D_x^2(f) > 0$ implique que f est strictement convexe en x. Pour une fonction à plusieures variables, $f : \mathbb{R}^n \to \mathbb{R}$, nous avons que $D_x^2(f) = \nabla_x^2(f) \succ 0$ implique f strictement convexe. Nous voyons donc que la notation est appropriée puisque l'on généralise la notion de 'plus grand que zéro.'