Université Joseph Fourier
Licence de Mathématiques

Méthodes Numériques
$2^{\text{ème}}$ semestre 2007/2008

Feuille d'exercices n° 4 : un peu d'optimisation.

$1^{\text{ère}}$ méthode : méthode du gradient à pas constant

Soit $d \geq 1$ et soit $J : \mathbb{R}^d \longrightarrow \mathbb{R}$ une fonctionnelle différentiable. On suppose que
• $\overrightarrow{\nabla} J$ est globalement lipschitzien c'est-à-dire qu'il existe C tel que

$$\forall x, y \in \mathbb{R}^d \ , \quad \|\overrightarrow{\nabla} J(x) - \overrightarrow{\nabla} J(y)\| \leq C\|x - y\| \ ,$$

• J est strictement convexe c'est-à-dire qu'il existe $\eta > 0$ tel que

$$\forall x, y \in \mathbb{R}^d \ , \quad < \overrightarrow{\nabla} J(x) - \overrightarrow{\nabla} J(y)|x - y > \quad \geq \quad \eta\|x - y\|^2 \ .$$

1) Etablir que pour tous $x, y \in \mathbb{R}^d$ et $t \in \mathbb{R}$,

$$\begin{aligned} J(y + t(x - y)) &= J(y) + \int_0^t < \overrightarrow{\nabla} J(y + s(x - y))|x - y > ds \\ &\geq \frac{\eta}{2} t^2 \|x - y\|^2 + t < \overrightarrow{\nabla} J(y)|x - y > + J(y) \ . \end{aligned}$$

Donner une interprétation géométrique de l'inégalité précédente justifiant l'appellation “strictement convexe” pour une telle fonction.
2) En déduire que J est bornée inférieurement et atteint son minimum en un point x_{min}. Montrer que $\overrightarrow{\nabla} J$ ne s'annule qu'en au plus un point et en déduire que x_{min} est unique.

On souhaite trouver le minimum de J ainsi que x_{min}. Pour cela, on utilise la méthode du gradient à pas constant. Soit $\tau > 0$ un petit pas. On part de $x_0 \in \mathbb{R}^d$ quelconque et on pose

$$x_{n+1} = x_n - \tau \overrightarrow{\nabla} J(x_n) \ .$$

3) En utilisant un théorème de point fixe, montrer que si le pas $\tau > 0$ est choisi assez petit, alors pour tout x_0, la suite (x_n) converge vers x_{min}. Quelle est la vitesse de convergence ?

Application aux fonctions réelles : soit $f \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$. On suppose que $f''(x) \geq \alpha > 0$ pour tout $x \in \mathbb{R}$. Justifier et appliquer la méthode du gradient à pas constant.

Application à la résolution de systèmes linéaires : soit $A \in \mathcal{M}_d(\mathbb{R})$ une matrice symétrique définie positive et soit b un vecteur de $\mathbb{R}^d$. On souhaite résoudre le système $Ax = b$. Montrer que cela est équivalent à minimiser sur $\mathbb{R}^d$ la fonctionnelle $J : x \longmapsto \frac{1}{2} < Ax|x > - < b|x >$. Justifier et appliquer la méthode du gradient à pas constant.

Application aux moindres carrés : soit $\{(x_i, y_i)\}_{i=1..n}$ une famille de points de $\mathbb{R}^2$. On souhaite trouver la droite $y = \alpha x + \beta$ approchant au mieux les points selon les moindres carrés, c'est-à-dire minimisant $J(\alpha, \beta) = \sum_{i=1}^n |y_i - (\alpha x_i + \beta)|^2$. Montrer que sous des hypothèses naturelles, le problème se ramène à l'application précédente.

$2^{\text{ème}}$ méthode : méthode de relaxation à pas optimal

Soit $J : \mathbb{R}^d \longrightarrow \mathbb{R}$ une fonctionnelle continue. On suppose que J est coercive, c'est-à-dire que $J(x)$ tend vers $+\infty$ quand $\|x\|$ tend vers $+\infty$.
1) Montrer que J est bornée inférieurement et atteint son minimum.

On se propose de trouver un point où le minimum de J est atteint à l'aide de l'algorithme suivant. Soit $(e_i)_{i=1..d}$ une base de $\mathbb{R}^d$. Soit x_0 un point de $\mathbb{R}^d$. On construit par récurrence une suite (x_n) en posant :
- $x_n^1 = x_n + t_1 e_1$ où t_1 est tel que $J(x_n + t_1 e_1) = \min_{t \in \mathbb{R}} J(x_n + t e_1)$,
- $x_n^2 = x_n^1 + t_2 e_2$ où t_2 est tel que $J(x_n^1 + t_2 e_2) = \min_{t \in \mathbb{R}} J(x_n^1 + t e_2)$,

  ...
- $x_{n+1} = x_n^d = x_n^{d-1} + t_d e_d$ où t_d est tel que $J(x_n^{d-1} + t_d e_d) = \min_{t \in \mathbb{R}} J(x_n^{d-1} + t e_d)$.

2) Montrer que l'on peut toujours construire une telle suite (x_n). Cette suite est-elle unique ?
3) Montrer que $(J(x_n))$ est une suite décroissante qui converge vers une valeur $J_0 \in \mathbb{R}$.
4) Montrer qu'il existe une sous-suite $(x_{\varphi(n)})$ qui converge vers un point $x_* \in \mathbb{R}^d$ et que $J(x_*) = J_0$.
5) Montrer que si J est différentiable et strictement convexe, alors $J(x_*)$ est le minimum de J (indication : on commencera par montrer que pour une telle fonction, le choix de t_i est unique).

Application à la résolution de systèmes linéaires : soit $A \in \mathcal{M}_d(\mathbb{R})$ une matrice symétrique définie positive et soit b un vecteur de $\mathbb{R}^d$. Montrer que la méthode itérative de Gauss-Seidel pour résoudre le système $Ax = b$ est en fait une méthode de relaxation à pas optimal pour trouver le minimum de la fonctionnelle $J : x \longmapsto \frac{1}{2} < Ax|x > - < b|x >$.