Université de Grenoble

2009/2010

Optimisation

Soit $J \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R})$ une fonctionnelle dont l'on souhaite trouver le minimum. On supposera que J est coercive, i.e. $\lim_{\|x\| \to +\infty} J(x) = +\infty$. On part d'un point $x_0 \in \mathbb{R}^d$ puis on applique un algorithme donnant une suite $(x_n)_{n\in\mathbb{N}}$ pour laquelle on espère que $\lim_{n\to+\infty} x_n = x_*$ existe et est telle que $J(x_*) = \min J$. On considèrera ici deux classes d'algorithmes.

Méthode du gradient à pas constant
Soit $\tau > 0$ un petit pas à choisir. On part de $x_0 \in \mathbb{R}^d$ quelconque et on pose

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

Méthode de relaxation à pas optimal
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 $J(x_n + t_1 e_1) = \min_{t\in\mathbb{R}} J(x_n + te_1)$,
- $x_n^2 = x_n^1 + t_2 e_2$ où t_2 est tel que $J(x_n^1 + t_2 e_2) = \min_{t\in\mathbb{R}} J(x_n^1 + te_2)$,

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

NB : on peut évidemment aussi faire une méthode de gradient à pas optimal ou une méthode de relaxation à pas constant.

Exercices théoriques

Exercice 1 : Montrer que J est bornée inférieurement et atteint son minimum.

Exercice 2 : On suppose que

- $\overrightarrow{\nabla} J$ est globalement lipschitzien c'est-à-dire qu'il existe C tel que

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

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

$$\forall x, y \in \mathbb{R}^d \ , \quad \langle \overrightarrow{\nabla} J(x) - \overrightarrow{\nabla} J(y) | x - y\rangle \ \ge \ \eta\|x - y\|^2 \ . \tag{1}$$

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

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

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_*. Montrer que $\overrightarrow{\nabla} J$ ne s'annule qu'en au plus un point et en déduire que x_* est unique.
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) donnée par la méthode du gradient à pas constant converge vers x_*.

Exercice 3 : On utilise dans cet exercice la méthode de relaxation à pas optimal.
1) Montrer que la suite (x_n) est bien définie. Cette suite est-elle unique ?
2) Montrer que $(J(x_n))$ est une suite décroissante qui converge vers une valeur $J_0 \in \mathbb{R}$. Montrer qu'il existe une sous-suite $(x_{\varphi(n)})$ qui converge vers un point $x_* \in \mathbb{R}^d$ et que $J(x_*) = J_0$.
3) Montrer que si J est strictement convexe au sens de (1), alors $J(x_*)$ est le minimum de J.

Exercice 4 : 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 $Ax = b$.
1) Montrer que cela équivaut à minimiser sur $\mathbb{R}^d$ la fonctionnelle $J : x \longmapsto \frac{1}{2}\langle Ax|x\rangle - \langle b|x\rangle$.
2) Justifier et appliquer la méthode du gradient à pas constant.
3) 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 J.

Exercices pratiques

Exercice 5 : Programmer la méthode du gradient à pas constant pour trouver le minimum de la fonction $f(x, y) = x^4 - x^3 + 2xy + y^2 - 2y + 1$. Illustrer du mieux possible le chemin parcouru par les différentes étapes de l'algorithme.

Exercice 6 : Programmer une méthode d'optimisation pour trouver le polynôme de degré plus petit que k qui approche le mieux une fonction $f \in \mathcal{C}^0([0, 1], \mathbb{R})$ au sens des moindres carrés.

Exercice 7 : Soit $\Omega =]0, 1[^2$ une plaque carrée. Une certaine température g est imposée au bord de la plaque. Une fois l'équilibre thermique atteint, la température u de la plaque est donnée par

$$\left\{ \begin{array}{ll} \Delta u = 0 & \text{sur } \Omega \\ u = g & \text{sur } \partial\Omega \end{array} \right. \tag{2}$$

Soit $n \in \mathbb{N}^*$. On discrétise le problème en posant remplaçant u par la matrice $n \times n$ schématiquement donnée par $U(i, j) = u(\frac{i-1}{n-1}, \frac{j-1}{n-1})$ $(i = 1, ..., n,\ j = 1, ..., n)$. Donner la discrétisation de l'équation (2) sous la forme

$$\left\{ \begin{array}{l} D(U) = 0 \\ U \in E \end{array} \right.$$

où D est un opérateur linéaire agissant sur les matrices $n \times n$ et E un espace vectoriel à préciser. Montrer que le problème revient à minimiser sur E la fonctionnelle suivante.

$$J(U) = \sum_{\substack{i, j, i', j' = 1 \ldots n \\ (i,j) \text{ voisin de } (i', j')}} |U(i, j) - U(i', j')|^2 \ .$$

Utiliser la méthode de relaxation à pas optimal pour résoudre numériquement ce problème.