Optimisation

mots clef Scilab: *chsolve, deff, function, optim, quapro, spec, plot, plot2d, plot3d, ...*

Soit $J \in C^2(\Omega \subset I\!R^d, I\!R)$, le problème est de trouver X et (ou) $J(X)$ tels que

$$J(X) = \min_{Y \in \Omega} J(Y). \tag{1}$$

Ce problème peut se ramener à la résolution d'un système non linéaire. Par exemple, si Ω est un ouvert, on cherchera à résoudre : $G(X) := \nabla J(X) = 0$. Inversement, il arrive que la résolution d'un sytème peut se ramener à un problème d'optimisation.
On testera des algorithmes sur l'exemple clef suivant, où A est symétrique définie positive et $B \in I\!R^d$,

$$J_0(X) = \frac{1}{2}\langle AX, X\rangle - \langle B, X\rangle, \quad A = \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & \cdots & 0 & -1 & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ \vdots \\ 1 \\ 1 \end{pmatrix}. \tag{2}$$

et un exemple non quadratique $J_1(x, y) := y^2 - x^2 + x^4/4$.

1 Minimisation à une variable

On peut chercher à résoudre $J'(x) = 0$ par des méthodes classiques de résolution approchée d'une équation non linéaires : méthode de la corde, du point fixe, de Newton, ..., voir le TP sur la résolution d'un système nonlinéaire.

1. ☐ **Algorithme de trichotomie** : (voir Chambert Loir et Fermigier p. 215. ou Ciarlet, exo 8.4.4 p. 109 corrigé p.184) L'intérêt de cette méthode est de ne pas calculer de dérivée. Soit f une fonction convexe (et dérivable) sur $[a, d]$. f atteint son minimum en $\xi \in [a, d]$. Soit $a < b < c < d$, $p(x, y) := (f(y) - f(x))/(x - y)$, $p_1 := p(a, b)$, $p_2 := p(b, c)$,$p_3 := p(c, d)$, alors on sait que f' est croissante sur $[a, d]$ et
$$p_1 \leq f'(b) \leq p_2 \leq f'(c) \leq p_3.$$
Donc, si $p_3 \leq 0$, f est décroissante sur $[a, c]$ et $\xi \in [c, d]$, sinon, si $p_2 \leq 0$, $\xi \in [b, d]$, sinon $p_1 \leq 0$, $\xi \in [a, c]$, sinon $p_1 > 0$, et $\xi \in [a, b]$.
Implémenter cette algorithme en prenant $b := (a + d)/2 - t(d - a)/4$, $c := (a + d)/2 + t(d - a)/4$.
Cas test: $f(x) = x \ln(x)$, $[a, d] = [0, 1]$, $t = 0.2$.

2. La minimisation d'une fonctionelle à plusieurs variables peut se ramener à "plusieurs" minimisations de fonctions à une seule variable. C'est le cas pour la méthode du gradient à pas optimal, la méthode du gradient conjugué, la méthode de Relaxation.
La méthode de Relaxation : (Ciarlet p. 201) soit $(e_k)_{1 \leq k \leq d}$ la base canonique de $I\!R^d$, X_0 donné, on définit par récurrence sur n et sur k : $X_{n,k}$ par : $X_{n,0} = X_n$, pour $1 \leq k \leq d$, $X_{n,k} = X_{n,k-1} + \rho_k e_k$ où ρ_k est donné par $J(X_{n,k}) = \min_{\rho \in \boldsymbol{R}} J(X_{n,k-1} + \rho e_k)$, $X_{n+1} = X_{n,d}$.
Ainsi, on résoud à chaque étape d problèmes de minimisation de la forme $j(\rho) = J(X + \rho e)$.
Pour $J = J_0$ on retombe sur la méthode de Gauss-Seidel, voir le TP sur la résolution des systèmes linéaires.

2 Méthodes de gradient

Elles sont de la forme $X_{n+1} = X_n - \rho_n \nabla J(X_n)$, avec un point de départ X_0 et $\rho_n > 0$ pour descendre dans la direction de plus grande pente.

1. ☐ **La méthode du gradient à pas fixe :** $\rho_n = \rho > 0$.
   Tester cette méthode sur J_0, pour $d = 3$, $X_0 := 0_{\mathbb{R}^3}$, $\rho = 0.1, 0.5, 0.7,$ ou 1 .

2. La méthode du gradient à pas optimal : voir Ciarlet p 189.
   $j(\rho) = J(X + \rho \nabla J(X_n))$, $j(\rho_n) = \min_{\rho \in \mathbb{R}} j(\rho)$, $X_{n+1}) = X_n - \rho_n \nabla J(X_n)$. Pour $J = J_0$ (et un choix "optimal" des directions de descente) on obtient la méthode du gradient conjugué : voir Ciarlet p 194-201.

3 Méthodes de Newton $X_{n+1} = X_n - (\nabla^2 J(X_n))^{-1} \nabla J(X_n)$

1. Que donne cette méthode pour la minimisation de J_0 en une itération.

2. ☐ Tester cette méthode sur J_1, $X_0 := [1, 1]$.

3. ☐ Comparer avec la méthode de gradient à pas fixe.

4 Exemples de Minimisation avec contraintes

1. ☐ Utiliser des multiplicateurs de Lagrange pour minimiser J_0 ($B = 0$, $d = 3$), sur $x_1 + 2x_2 + 3x_3 = 10$.

2. ☐ **La méthode du gradient projeté** : voir Ciarlet p.203
   Soit Ω un convexe fermé de $\mathbb{R}^d$ et $Proj$ le projecteur orthogonal de $\mathbb{R}^d$ sur Ω, on a l'algorithme suivant :
   $$X_{n+1} = Proj(X_n - \rho_n \nabla J(X_n)) = X_n - \rho_n Proj(\nabla J(X_n))$$
   Tester cette alghorithme sur $J(X) = -\langle AX, X \rangle$, Ω la boule unité de $\mathbb{R}^d$, $X_0 \neq 0$, $d = 3$. Comparer le résultat obtenu avec les valeurs propres et les vecteurs propres de A.

Références :

Chambert Loir et Fermigier, Exercices pour l'agrégation, 2
Ciarlet, Introduction à l'analyse numérique matricielle et à l'optimisation,
Ciarlet, Miara, Thomas, Exercices 'analyse numérique matricielle et d'optimisations avec solutions,
Dumas, L'épreuve de modélisation à l'agrégation,
Henrici, Elements of numerical analysis,
Lax, Linear Algebra,
Rouvière, Petit guide de calcul différentiel,
Schatzman, analyse numérique : cours et exercices pour la licence,

☐ = Corrigé en TP