UNIVERSITÉ DE RENNES 1 **Année 2006-2007**
U.F.R. Mathématiques

TRAVAUX PRATIQUES D'OPTIMISATION

3. Méthode du gradient à pas optimal

On s'intéresse au problème : trouver $u^* \in \mathbf{R}^n$ tel que

$$J(u^*) = \min_{u \in \mathbf{R}^n} J(u),$$

où J est une fonctionnelle continûment différentiable sur $\mathbf{R}^n$.
Une méthode numérique pour résoudre ce problème est la méthode du gradient à pas optimal. Cet algorithme s'écrit:

- On part de u^0, a priori arbitraire.

- À l'étape k, connaissant u^k, on calcule u^{k+1} par

$$u^{k+1} = u^k - \alpha_k \nabla J(u^k)$$

où $\alpha_k \in \mathbf{R}$ réalise le minimum $\min_{\alpha} J(u^k - \alpha \nabla J(u^k))$.

- On arrête l'algorithme lorsqu'un test de convergence est vérifié, ou lorsque le nombre maximum d'itérations est dépassé.

Pour mettre en œuvre la méthode du gradient à pas optimal, on créera trois fichiers Matlab:

- `J.m` qui sert à définir la fonctionnelle J et dont l'en-tête sera

```
function ju = J(u)
% renvoie la valeur de la fonctionnelle J au point u
global numex
```

-`GJ.m` qui contient l'expression du gradient de la fonctionnelle J et dont l'en-tête sera

```
function gu = GJ(u)
% renvoie la valeur du gradient de la fonctionnelle J au point u
global numex
```

-`Gradopt.m` qui contiendra le programme principal mettant en œuvre la méthode du gradient à pas optimal.

Le paramètre `numex` sera utilisé afin de distinguer les différents exemples proposés.
On utilisera pour test d'arrêt la condition suivante:

$$\|\nabla J(u^k)\| \leq \tau \qquad \text{avec} \;\; \tau = 10^{-6}.$$

Méthode du gradient à pas optimal pour une fonctionnelle quadratique elliptique

Dans le cas d'une fonctionnelle quadratique elliptique

$$J(v) = \frac{1}{2}(Av, v) - (b, v) + c$$

on peut calculer facilement le paramètre α_k. On a

$$\alpha_k = \frac{\|w^k\|^2}{(Aw^k, w^k)} \qquad \text{avec} \;\; w^k = Au^k - b = \nabla J(u^k).$$

1 - Écrire un programme Matlab mettant en œuvre la méthode du gradient à pas optimal pour une fonctionnelle quadratique elliptique.

Copier les deux fichiers `/usr/local/ananum/opti/tp3/visiso.m`
et `/usr/local/ananum/opti/tp3/visiter.m`.
Ces fichiers contiennent des fonctions permettant l'affichage graphique des lignes de niveau (`visiso.m`) et des caractéristiques de la méthode du gradient à pas optimal : suite des points (u^k) et directions de descente (`visiter.m`) . Incorporer ces fonctions graphiques au programme.

2- Résoudre par la méthode du gradient à pas optimal les problèmes suivants :

$$(\mathcal{P}_1) \;\; : \;\; \min \quad -2x_1x_2 - 2x_2 + x_1^2 + 2x_2^2,$$

en prenant $u^0 = (0, 0)$ (le minimum est atteint en $u^* = (1, 1)$).

$$(\mathcal{P}_2) \;\; : \;\; \min \quad -2x_1x_2 - 4x_1 + x_1^2 + 2x_2^2,$$

en prenant $u^0 = (0, 0)$ (le minimum est atteint en $u^* = (4, 2)$).

$$(\mathcal{P}_3) \;\; : \;\; \min \quad 5x_1^2 + \frac{5}{2}x_2^2 + 7x_1x_2 + 2x_1 + x_2.$$

en prenant $u^0 = (0, 0)$ (le minimum est atteint en $u^* = (-3, 4)$).

À quelle matrice A correspondent ces exemples ? Observer que deux directions de descente successives sont orthogonales. Justifier cette propriété.

Incidence du conditionnement de la matrice A

3 - Quel est le conditionnement de la matrice A (commande Matlab `cond`) de chacun des exemples précédents.
On souhaite illustrer le résultat suivant ([J.C. Culioli, *Introduction à l'optimisation*, Ellipse 1994.]) :
Si $J(v) = \frac{1}{2}(Av, v) - (b, v) + c$, avec A symétrique définie positive, la méthode du gradient à pas optimal converge vers l'unique optimum $u^* = A^{-1}b$.
De plus

$$\|u^k - u^*\|_A \leq (\frac{r-1}{r+1})^k \|u^0 - u^*\|_A$$

où r est le conditionnement de A relativement à la norme euclidienne.
Tracer sur un même graphique les courbes donnant $\log \|u^k - u^*\|_A$ en fonction de k pour les problèmes $(\mathcal{P}_1)$ et $(\mathcal{P}_3)$. On pourra y superposer les courbes donnant le logarithme de $(\frac{r-1}{r+1})^k \|u^0 - u^*\|_A$. Commenter le graphique obtenu.

Méthode du gradient à pas optimal pour une fonctionnelle quelconque

4 - Écrire un programme Matlab mettant en œuvre la méthode du gradient à pas optimal dans le cas général. On utilisera la fonction Matlab `fminbnd` pour effectuer la recherche du paramètre α_k.

5 - Résoudre les problèmes suivants (on affichera la suite des points (u^k) et les directions de descente successives):

$$(\mathcal{P}_4) \ : \ \min \ \ c(x_1^2 + x_2^2 + x_1x_2 - ax_1 + ax_2)^2 + (1 - bx_1 - bx_2)^2$$

où $a = 4, b = 4$ et $c = 10$. On prendra $u^0 = (-1, 2)$.

$$(\mathcal{P}_5) \ : \ \min \ \ x_1^2 - x_2^2,$$

en prenant $u^0 = (1, 0)$.

$$(\mathcal{P}_6) \ : \ \min \ \ (x_1 - 1)^2 + p(x_1^2 - x_2)^2,$$

en prenant $p = 10$ et $u^0 = (0, 1)$ (le minimum global est atteint en $u^* = (1, 1)$).

On résoudra également ces problèmes pour d'autres valeurs de u^0 que celles proposées.

Le problème $(\mathcal{P}_6)$ est un problème test classique. Les lignes de niveau de la fonctionnelle forment une vallée très étroite en forme de banane conduisant au minimum (cette fonction est appelée "Rosenbrock banana"). Seule une méthode efficace permet de trouver le minimum en un temps raisonnable et en évitant les difficultés numériques dues au mauvais conditionnement du Hessien.

6 - Étudier la convergence de la méthode du gradient à pas optimal (tracer la courbe donnant $\log \|u^k - u^*\|$ en fonction de k pour différents exemples).

Travail à rendre

Il est demandé de rendre un rapport à l'enseignant, au plus tard en fin de sixième séance de TP. Ce rapport doit répondre aux questions du TP. Hors annexes, il n'excèdera pas 3 pages et ne contiendra ni graphe ni script Matlab. Il y sera joint, au plus, les annexes suivantes :

- le script du programme Matlab mettant en œuvre la méthode du gradient à pas optimal pour une fonctionnelle quadratique elliptique, demandé en question 1,
- le graphique demandé en question 3,
- les graphiques affichant les points (u^k) et les directions de descente successives pour deux problèmes au choix parmi $(\mathcal{P}_4)$, $(\mathcal{P}_5)$ et $(\mathcal{P}_6)$ dans le cadre de la question 5,
- la courbe donnant $\log \|u^k - u^*\|$ en fonction de k pour le problème $(\mathcal{P}_6)$ dans le cadre de la question 6, avec u_0 de votre choix.