

---

## FEUILLE DE TD NUMÉRO 6

Optimisation L3 - Deuxième semestre 2009-2010 - Université Denis-Diderot

---

### *Gradient conjugué; Vitesse de convergence*

Dans les exercices suivants,  $A$  désigne une matrice symétrique définie positive  $N \times N$ , de valeurs propres  $0 < \lambda_1 \leq \dots \leq \lambda_N$ , et  $b \in \mathbb{R}^N$ . Pour  $v \in \mathbb{R}^N$ , on pose

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

On notera  $u$  le point de minimum de  $J$  sur  $\mathbb{R}^N$ .

Le conditionnement de  $A$  est défini par  $\kappa(A) := \|A\|_2 \|A^{-1}\|_2$  et vaut  $\frac{\lambda_N}{\lambda_1}$ . On note aussi  $\|v\|_A^2 = (Av, v)$  pour  $v \in \mathbb{R}^N$ . On notera enfin  $u$  le minimum de  $J$  sur  $\mathbb{R}^N$ .

### **Exercice 1. (Algorithme du gradient conjugué)**

Il s'agit de l'algorithme de descente à pas optimal :

$$u^{n+1} = u^n + \rho^n d^n \tag{1}$$

avec

$$d^0 = -\nabla J(u^0), \quad d^n = -\nabla J(u^n) + \frac{\|\nabla J(u^n)\|^2}{\|\nabla J(u^{n-1})\|^2} d^{n-1}, \quad \rho^n = \frac{\|\nabla J(u^n)\|^2}{(Ad^n, d^n)}.$$

Rappelons que le principe de cette méthode est le suivant:

On part de  $u^0 \in \mathbb{R}^N$ , fixé.

Puis pour  $n \geq 0$  on construit récursivement une suite  $(u^n)$  telle que :

$$u^{n+1} \in u^n + G^n \quad \text{et} \quad J(u^{n+1}) = \inf_{v \in u^n + G^n} J(v) \tag{2}$$

avec  $G^n = \text{vect}(\nabla J(u^0), \dots, \nabla J(u^n))$ .

1. Vérifier que

$$J(v) - J(u) = \frac{1}{2} \|v - u\|_A^2 \quad \forall v \in \mathbb{R}^N. \tag{3}$$

2. Montrer que l'espace  $G^n$  est inclu dans l'espace  $H^n$  défini par :

$$H^n := \text{Vect}(\nabla J(u^0), A\nabla J(u^0), \dots, A^n \nabla J(u^0)).$$

3. Réciproquement, supposons que les  $\nabla J(u^k)$  sont non nuls pour  $0 \leq k \leq n$  (on notera que si  $\nabla J(u^n) = 0$  alors  $u^n$  est la solution cherchée et on peut arrêter l'algorithme). Montrer que les  $\nabla J(u^k)$  sont deux à deux orthogonaux. En déduire que  $G^n = H^n$ .

4. Pour  $k \geq 0$ , on pose  $e^k = u^k - u$ .

(a) Soit  $v \in u^n + G^n$ . Montrer qu'il existe un polynôme  $Q$  de degré  $n$  tel que:  $v = u^0 + Q(A)\nabla J(u^0)$ .

(b) En utilisant (2) et (3), montrer que:

$$\|e^{n+1}\|_A = \min\{\|P(A)e^0\|_A, P \in \mathcal{P}_{n+1}\} \quad (4)$$

où  $\mathcal{P}_n$  désigne l'ensemble des polynômes  $P$  de degré  $n$  tels que  $P(0) = 1$ .

5. Montrer que

$$\|P(A)e^0\|_A^2 \leq \|e^0\|_A^2 \max_i P^2(\lambda_i) \quad \forall P \in \mathcal{P}_k.$$

6. En déduire que

$$\|e^n\|_A \leq 2 \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^n \|e^0\|_A, \quad \text{avec } \kappa = \frac{\lambda_N}{\lambda_1}.$$

(On pourra utiliser sans le démontrer le fait que  $\min\{\|p\|_{L^\infty([a,b])} : p \in \mathcal{P}_k\} = 1/T_k\left(\frac{b+a}{b-a}\right)$ , où  $T_k$  désigne le polynôme de Tchebychev de degré  $k$ , ainsi que l'inégalité suivante, valable lorsque  $b > a$  :

$$T_k\left(\frac{b+a}{b-a}\right) = T_k\left(\frac{\frac{b}{a}+1}{\frac{b}{a}-1}\right) > \frac{1}{2} \left( \frac{\sqrt{\frac{b}{a}+1}}{\sqrt{\frac{b}{a}-1}} \right)^k.$$

**Exercice 2. (Comparaison de méthodes itératives)** On suppose maintenant que la matrice  $A$  est donnée par :

$$A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \dots & \ddots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix},$$

où  $h = \frac{1}{N+1}$ . Le but de cet exercice est de comparer sur cet exemple, les méthodes de descente de gradient et de gradient conjugué.

1. Montrer que les valeurs propres de la matrice  $A$  sont les:

$$\lambda^k = \frac{4}{h^2} \sin^2\left(\frac{k\pi h}{2}\right) \quad 1 \leq k \leq N,$$

associées aux vecteurs propres  $\varphi^k$

$$\varphi^k = [\sin(k\pi h) \quad \sin(2k\pi h) \quad \dots \quad \sin((N-1)k\pi h) \quad \sin(Nk\pi h)]^T.$$

(Indication : on pourra calculer directement  $(A\varphi^k)_j$ .)

2. En déduire un équivalent de  $\lambda_1$ ,  $\lambda_N$  puis de  $\kappa(A)$  pour  $h$  petit (c'est dire pour  $N$  grand).

3. Evaluer la vitesse de convergence du gradient à pas optimal pour  $h$  petit (on rappelle l'estimation d'erreur  $\|e^{n+1}\|_A \leq \frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} \|e^n\|_A$ .)

4. Qu'en est-t-il pour le gradient à pas fixe, avec le choix optimal du pas ?

5. Evaluer la vitesse de convergence du gradient conjugué pour  $h$  petit.

6. Estimer le nombre d'itérations nécessaires pour obtenir la convergence à une précision  $\varepsilon$  donnée pour la méthode du gradient à pas optimal (par exemple).

Application:  $n = 99$ ,  $\varepsilon = 10^{-3}$ .