UNIVERSITE DE GENEVE Section de mathématiques
Faculté des sciences

Méthodes itératives – Corrigé série 5

Corrigé exercice 1 Il suffit de vérifier que l'application $a : (\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{y}^{\mathrm{T}}\mathbf{A}\boldsymbol{x}$ est un produit scalaire :

– a est bilinéaire.

– a est symétrique : $a(\boldsymbol{x}, \boldsymbol{y}) = a(\boldsymbol{y}, \boldsymbol{x})$ car la matrice $\mathbf{A}$ est symétrique.

– $a(\boldsymbol{x}, \boldsymbol{x} = 0$ si et seulement si $\boldsymbol{x} = 0$ car $\mathbf{A}$ est définie.

Corrigé exercice 2 D'après le cours, si $\mathbf{A} = \mathbf{S\Lambda S}^{-1}$, avec Λ diagonale, on a

$$\frac{\|\boldsymbol{b} - \mathbf{A}\boldsymbol{x}_k\|_2}{\|\boldsymbol{b}\|_2} \leq \kappa(S)\frac{C_k(a/d)}{C_k(c/d)},$$

où les C_k sont les polynômes de Chebyshev complexes. Comme

$$\frac{a}{d} = \frac{1}{2}\left(\frac{a + \sqrt{a^2 - d^2}}{d}\right) + \left(\frac{a + \sqrt{a^2 - d^2}}{d}\right)^{-1}$$
$$\frac{c}{d} = \frac{1}{2}\left(\frac{c + \sqrt{c^2 - d^2}}{d}\right) + \left(\frac{c + \sqrt{c^2 - d^2}}{d}\right)^{-1},$$

nous avons

$$\frac{a}{d} = \frac{1}{2}\left(\frac{a + \sqrt{a^2 - d^2}}{d}\right)^k + \left(\frac{a + \sqrt{a^2 - d^2}}{d}\right)^{-k}$$
$$\frac{c}{d} = \frac{1}{2}\left(\frac{c + \sqrt{c^2 - d^2}}{d}\right)^k + \left(\frac{c + \sqrt{c^2 - d^2}}{d}\right)^{-k}.$$

Par conséquent,

$$\frac{C_k(a/d)}{C_k(c/d)} = \frac{\left(\frac{a+\sqrt{a^2-d^2}}{d}\right)^k + \left(\frac{a+\sqrt{a^2-d^2}}{d}\right)^{-k}}{\left(\frac{c+\sqrt{c^2-d^2}}{d}\right)^k + \left(\frac{c+\sqrt{c^2-d^2}}{d}\right)^{-k}}$$

qui est un équivalent à $\left(\frac{a+\sqrt{a^2-d^2}}{c+\sqrt{c^2-d^2}}\right)^k$ pour k grand.

Corrigé exercice 3 a) On note $F(\boldsymbol{x}) = \boldsymbol{x}\mathbf{A}\boldsymbol{x} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}$ et $\boldsymbol{r}_k = \boldsymbol{b} - \mathbf{A}\boldsymbol{x}_k$. On a

$$F(\boldsymbol{x}_{k+1}) = F(\boldsymbol{x}) - \alpha_k \boldsymbol{r}_k{}^{\mathrm{T}}\boldsymbol{d}_k + \frac{\alpha_k{}^2}{2}\boldsymbol{d}_k{}^{\mathrm{T}}\mathbf{A}\boldsymbol{d}_k.$$

Donc, pour minimiser $F(\boldsymbol{x}_{k+1}$, il faut choisir $\alpha_k = \frac{\boldsymbol{r}_k \mathbf{A}\boldsymbol{d}_k}{\boldsymbol{d}_k{}^{\mathrm{T}}\mathbf{A}\boldsymbol{d}_k}$.

b) Les lignes de niveaux sont des ellipses concentriques de même centre avec un grand axe beaucoup plus grand que le petit axe. Il en résulte que l'algorithme de pente maximale va beaucoup zigzager avant de converger.

c) Pour que l'erreur $\boldsymbol{e}_{k+1}$ soit orthogonale à $\boldsymbol{d}_{\boldsymbol{k}}$, il faut que $\alpha_k = -\frac{\boldsymbol{e}_k \cdot \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|_2^2}$. L'algorithme converge en n itérations au plus car l'erreur $\boldsymbol{e}_k$ est orthogonale à toutes les directions $(\boldsymbol{d}_j)_{0\leq j\leq k-1}$, après n itérations, l'erreur est donc orthogonale à $\mathbb{R}^n$ et est donc forcément nulle. L'algorithme est inimplémentable car il necessite de connaître la solution du système linéaire à l'avance. En effet, si on connait $\boldsymbol{e}_k$, il n'y a plus qu'une addition à faire pour obtenir la solution exacte.

d) Il n'y a rien à calculer $\mathbf{A}$ définit simplement un autre produit scalaire. Il suffit de remplacer le produit scalaire usuel dans la formule précédente par le produit scalaire provenant de $\mathbf{A}$ donc $\alpha_k = -\frac{\mathbf{A}\boldsymbol{e}_k \cdot \boldsymbol{d}_k}{\boldsymbol{d}_k \cdot \mathbf{A}\boldsymbol{d}_k}$. Soit $\alpha_k = \frac{\boldsymbol{r}_k \cdot \boldsymbol{d}_k}{\boldsymbol{d}_k \cdot \mathbf{A}\boldsymbol{d}_k}$. Contrairement au cas précédent, ici α_k est calculable.

e) Tout d'abord, par construction, les directions $\boldsymbol{d}_k$ sont $\mathbf{A}$ orthogonales. et $\mathrm{Vect}(\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}) = \mathrm{Vect}(\boldsymbol{r}_0, \ldots, \boldsymbol{r}_{k-1})$. Encore par construction, l'erreur $\boldsymbol{e}_{k+1}$ est $\mathbf{A}$ orthogonale à $\boldsymbol{d}_k$. Comme $\boldsymbol{e}_{k+1} = \boldsymbol{e}_k + \alpha_k \boldsymbol{d}_k$, et que les $\boldsymbol{d}_k$ sont $\mathbf{A}$ orthogonaux, nous avons par récurrence que l'erreur $\boldsymbol{e}_k$ est $\mathbf{A}$ orthogonale à toutes les directions précédentes. Comme $\boldsymbol{r}_k = -\mathbf{A}\boldsymbol{e}_k$, nous avons que $\boldsymbol{r}_k$ est orthogonale à toutes les directions précédentes. Notez que cela implique que $\boldsymbol{d}_k$ ne peut être nulle à moins que $\boldsymbol{r}_k$ ne le soit aussi. L'égalité entre espaces vectoriels déjà obtenue nous permet alors d'affirmer que que les résidus $\boldsymbol{r}_k$ sont orthogonaux entre eux. Nous avons prouvé les deux premiers point et la première égalité du troisième point. Pour prouver la dernière égalité du trolisième point, on procède par récurence. Pour $k = 1$, l'égalité est trivial. Supposons l'égalité prouvé pour $k \geq 1$. Comme $r_k = r_{k-1} - \alpha_{k-1}\mathbf{A}\boldsymbol{d}_k$ et que $\mathbf{A}\mathcal{K}(\mathbf{A}, \boldsymbol{r}_0, k) \subset \mathcal{K}(\mathbf{A}, \boldsymbol{r}_0, k)$, nous avons forcément l'inclusion $\subset$. L'inclusion inverse est acquise car la dimension de l'espace de Krylov est inférieure à k et que la dimension de l'espace $\mathrm{Vect}(\boldsymbol{r}_0, \ldots, \boldsymbol{r}_{k-1})$ vaut k jusqu'à convergence.

Nous avons $\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k \mathbf{A}\boldsymbol{d}_k$. Si on multiplie scalairement cette égalité par $\boldsymbol{r}_j$ avec j différent de $k, k+1$, on en déduit que le résidu $\boldsymbol{r}_k$ est $\mathbf{A}$ orthogonal avec toutes les directions $\boldsymbol{d}_j$ sauf $\boldsymbol{d}_k$ et $\boldsymbol{d}_{k-1}$. Or, par définition $\boldsymbol{d}_k = \boldsymbol{r}_k - \sum_{j=0}^{k-1} \frac{\boldsymbol{d}_j \mathbf{A} \boldsymbol{r}_k}{\boldsymbol{d}_j \cdot \mathbf{A}\boldsymbol{d}_j} \boldsymbol{d}_j$. Tous les termes dans la somme sont nuls sauf le dernier. Donc $\beta_k = -\frac{\boldsymbol{d}_{k-1} \cdot \mathbf{A}\boldsymbol{r}_k}{\boldsymbol{d}_{k-1} \cdot \mathbf{A}\boldsymbol{d}_{k-1}}$.

f) Il ne reste plus qu'à donner les autres expression de α_k et β_k qui sont utilisés dans la pratique. Nous avons

$$\begin{aligned} \alpha_k &= \frac{\boldsymbol{r}_k \cdot \boldsymbol{d}_k}{\boldsymbol{d}_k \cdot \mathbf{A}\boldsymbol{d}_k} \\ &= \frac{\boldsymbol{r}_k \cdot \boldsymbol{r}_k}{\boldsymbol{d}_k \cdot \mathbf{A}\boldsymbol{d}_k}, \end{aligned}$$

car $\boldsymbol{d}_k - \boldsymbol{r}_{\boldsymbol{k}}$ est orthogonal à $\boldsymbol{r}_k$.

$$\begin{aligned} \beta_k &= -\frac{\mathbf{A}\boldsymbol{d}_{k-1} \cdot \boldsymbol{r}_k}{\boldsymbol{d}_{k-1} \cdot \mathbf{A}\boldsymbol{d}_{k-1}} \\ &= \frac{1}{\alpha_{k-1}} \frac{(\boldsymbol{r}_k - \boldsymbol{r}_{k-1}) \cdot \boldsymbol{r}_k}{\boldsymbol{d}_{k-1} \cdot \mathbf{A}\boldsymbol{d}_{k-1}} \\ &= \frac{\boldsymbol{r}_k \cdot \boldsymbol{r}_k}{\boldsymbol{r}_{k-1} \cdot \boldsymbol{r}_{k-1}} \end{aligned}$$

Corrigé exercice 4 code pour CG

```
function [x,res]=CG(A,b,x0,tol,maxiter)
% CG Résolution de systèmes linéaires par le gradient conjugué
%   [x,res]=CG(A,b,x0,tol,maxiter) résoud le système linéaire Ax=b
%   par la méthode du gradient conjugué ou A est une matrice
%   symétrique. L'algorithme s'arrête quand norm(A*x-b)<=tol*norm(b)
%    ou quand maxiter itérations ont été effectuées.
if(nargin <2)
  error('not enough arguments');
end
if(nargin < 5)
  maxiter=500;
end
if(nargin <4)
  tol=1e-6;
end
if(nargin <3)
   x0=b;
   x0(:)=0;
```

```
end

r=b-A*x;
d=r;

res(1)=norm(b-A*x);
i=1;
while i<=maxiter & res(i)>tol*norm(b)
  alpha=(r'*r)/(d'*(A*d));
  x=x+alpha*d;
  oldr=r;
  r=r-alpha*(A*d));
  beta=(r'*r)/(oldr'*oldr);
  d=r-beta*d;
  res(i+1)=norm(r,2);
  i=i+1;
end
```