UNIVERSITE DE GENEVE — Section de mathématiques
Faculté des Sciences — Mercredi 10 janvier 2007
À rendre pour le début de la séance du Mercredi 24 janvier 2007

Optimisation – Corrigé 5

Exercice 1 a) Nous avons

$$A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} = (A^{-1}A)^{\mathrm{T}} = \mathrm{I},$$
$$(A^{-1})^{\mathrm{T}}A^{\mathrm{T}} = (AA^{-1})^{\mathrm{T}} = \mathrm{I}.$$

b) D'abord $B^{-1} = (B^{\mathrm{T}})^{-1}$, donc nous n'avons qu'à traiter B^{-1}. Nous avons $(B^{-1})^{\mathrm{T}} = (B^{\mathrm{T}})^{-1} = B^{-1}$.

c) Calculons les dérivées partielles $\frac{\partial f}{\partial x_i} = \frac{1}{2}\sum_{j=1}^{n}(B_{ij} + B_{ji})x_j \ - b_i$. Donc $\nabla f = \frac{1}{2}(B + B^{\mathrm{T}})\boldsymbol{x} + \boldsymbol{b}$. Comme B est symétrique. Démontrer que $\nabla f(\boldsymbol{x}) = B\boldsymbol{x} + b$

d) Comme B est symétrique, B peut être diagonalisé dans une base orthgogonale $B = Q^{\mathrm{T}}\Lambda Q$. Nous avons

$$\{\frac{\boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} \mid \boldsymbol{x} \in \mathbb{R}^n\} = \{\frac{\boldsymbol{y}^{\mathrm{T}}\Lambda\boldsymbol{y}}{\|\boldsymbol{x}\|_2^2} \mid \boldsymbol{y} \in \mathbb{R}^n\}$$

et $\lambda_{min}(B) = \lambda_{min}(\Lambda)$ et $\lambda_{max}(B) = \lambda_{max}(\Lambda)$. Il suffit alors de raisonner sur Λ.

e) Même raisonnement que pour le d).

Exercice 2 a) Rang 1 symétrique :

$$B_{k+1} = B_k + \frac{(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)^{\mathrm{T}}}{\boldsymbol{s}_k^{\mathrm{T}}(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)}$$

i. *Montrer que* $B_{k+1}\boldsymbol{s}_k = \boldsymbol{y}_k$. Nous avons

$$\begin{aligned} B_{k+1}\boldsymbol{s}_k &= B_k\boldsymbol{s}_k + \frac{(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)^{\mathrm{T}}\boldsymbol{s}_k}{\boldsymbol{s}_k^{\mathrm{T}}(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)} \\ &= B_k\boldsymbol{s}_k + \boldsymbol{y}_k - B_k\boldsymbol{s}_k \\ &= \boldsymbol{y}_k \end{aligned}$$

ii. *Montrer que si B_k est symétrique alors B_{k+1} l'est aussi.* Nous avons

$$B_{k+1} = B_k + \frac{(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)^{\mathrm{T}}}{\boldsymbol{s}_k^{\mathrm{T}}(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)}$$

La matrice B_k est symétrique. La matrice $(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)(\boldsymbol{y}_k - B_k\boldsymbol{s}_k)^{\mathrm{T}}$ l'est aussi MM^{T} est toujours symétrique.

b) BFGS : Broyder-Fletcher-Goldbach-Shano

$$B_{k+1} = B_k + \frac{\boldsymbol{y}_k\boldsymbol{y}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{s}_k} - \frac{B_k\boldsymbol{s}_k\boldsymbol{s}_k^{\mathrm{T}}B_k}{\boldsymbol{s}_k^{\mathrm{T}}B_k\boldsymbol{s}_k}$$

i. *Montrer que $B_{k+1}\boldsymbol{s}_k = \boldsymbol{y}_k$.* Nous avons :

$$\begin{aligned}
B_{k+1}\boldsymbol{s}_k &= B_k\boldsymbol{s}_k + \frac{\boldsymbol{y}_k\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{s}_k}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{s}_k} - \frac{B_k\boldsymbol{s}_k\boldsymbol{s}_k^{\mathrm{T}}B_k\boldsymbol{s}_k}{\boldsymbol{s}_k^{\mathrm{T}}B_k\boldsymbol{s}_k} \\
&= B_k\boldsymbol{s}_k + \boldsymbol{y}_k - B_k\boldsymbol{s}_k \\
&= \boldsymbol{y}_k
\end{aligned}$$

ii. *Montrer que si B_k est symétrique alors B_{k+1} l'est aussi.*

$$B_{k+1} = B_k + \frac{\boldsymbol{y}_k\boldsymbol{y}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{s}_k} - \frac{B_k\boldsymbol{s}_k\boldsymbol{s}_k^{\mathrm{T}}B_k}{\boldsymbol{s}_k^{\mathrm{T}}B_k\boldsymbol{s}_k}$$

B_k est symétrique et les autres termes aussi car pour toute matrice M, MM^{T} est toujours symétrique. De plus, la somme de deux matrices symétriques est aussi symétrique.

iii. *Montrer que si B_k est définie positive et si $s_k^{\mathrm{T}}y_k > 0$ alors B_{k+1} est aussi définie positive.* Nous avons

$$\begin{aligned}
\boldsymbol{x}_k^{\mathrm{T}}B_{k+1}\boldsymbol{x}_k &= \boldsymbol{x}_k^{\mathrm{T}}B_k\boldsymbol{x} + \frac{\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{y}_k\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{x}_k}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{s}_k} - \frac{\boldsymbol{x}_k^{\mathrm{T}}B_k\boldsymbol{s}_k\boldsymbol{s}_k^{\mathrm{T}}B_k\boldsymbol{x}_k}{\boldsymbol{s}_k^{\mathrm{T}}B_k\boldsymbol{s}_k} \\
&= \boldsymbol{x}_k^{\mathrm{T}}B_k\boldsymbol{x} + \frac{\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{y}_k\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{x}_k}{\boldsymbol{y}_k^{\mathrm{T}}\boldsymbol{s}_k} - \frac{(\boldsymbol{x}_k^{\mathrm{T}}B_k\boldsymbol{s}_k)^2}{\boldsymbol{s}_k^{\mathrm{T}}B_k\boldsymbol{s}_k}
\end{aligned}$$

Mais, par Cauchy-Schwarz,

$$\frac{(\boldsymbol{x}_k^{\mathrm{T}}B_k\boldsymbol{s}_k)^2}{\boldsymbol{s}_k^{\mathrm{T}}B_k\boldsymbol{s}_k} \leq \boldsymbol{x}_k^{\mathrm{T}}B_k\boldsymbol{x}_k.$$

L'égalité n'a lieu que si $\boldsymbol{s}_k = \lambda\boldsymbol{x}_k$ et donc B_{k+1} est définie positive.

iv. Montrer que

$$B_{k+1}^{-1} = B_k^{-1} - \frac{\boldsymbol{s}_k \boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} + B_k^{-1} \boldsymbol{y}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} + \left(1 + \frac{\boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} \boldsymbol{y}_k}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) \frac{\boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}.$$

Calculons :

$$\begin{aligned}
&\left(B_k + \frac{\boldsymbol{y}_k \boldsymbol{y}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} - \frac{B_k \boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{T}} B_k}{\boldsymbol{s}_k^{\mathrm{T}} B_k \boldsymbol{s}_k}\right) \\
&\left(B_k^{-1} - \frac{\boldsymbol{s}_k \boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} + B_k^{-1} \boldsymbol{y}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} + \left(1 + \frac{\boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} \boldsymbol{y}_k}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) \frac{\boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}.\right) = \\
&= \left(I - \frac{B_k \boldsymbol{s}_k \boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} + \boldsymbol{y}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} + \left(1 + \frac{\boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} \boldsymbol{y}_k}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) \frac{B_k \boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) + \\
&+ \left(\frac{\boldsymbol{y}_k \boldsymbol{y}_k^{\mathrm{T}} B_k^{-1}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} - \frac{\boldsymbol{y}_k \boldsymbol{y}_k^{\mathrm{T}} B_k^{-1}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} - \frac{\boldsymbol{y}_k \boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} \boldsymbol{y}_k \boldsymbol{s}_k^{\mathrm{T}}}{(\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k)^2} + \left(1 + \frac{\boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} \boldsymbol{y}_k}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) \frac{\boldsymbol{y}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}.\right) - \\
&- \left(\frac{B_k \boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{s}_k^{\mathrm{T}} B_k \boldsymbol{s}_k} - \frac{B_k \boldsymbol{s}_k \boldsymbol{y}_k^{\mathrm{T}} B_k^{-1}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} - \frac{B_k \boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{s}_k^{\mathrm{T}} B_k \boldsymbol{s}_k} + \left(1 + \frac{\boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} \boldsymbol{y}_k}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) \frac{B_k \boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) = \\
&= \left(I - \frac{B_k \boldsymbol{s}_k \boldsymbol{y}_k^{\mathrm{T}} B_k^{-1}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} - \frac{\boldsymbol{y}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} + \left(1 + \frac{\boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} \boldsymbol{y}_k}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) \frac{B_k \boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) + \\
&+ \frac{\boldsymbol{y}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} - \left(-\frac{B_k \boldsymbol{s}_k \boldsymbol{y}_k^{\mathrm{T}} B_k^{-1}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k} + \left(1 + \frac{\boldsymbol{y}_k^{\mathrm{T}} B_k^{-1} \boldsymbol{y}_k}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) \frac{B_k \boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{T}}}{\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k}\right) = \\
&= I.
\end{aligned}$$

Exercice 3 a) Implémenter la méthode de Newton.

```
function [x,xk]=Newton(f,fp,fpp,x0,tol,maxiter,tau,be,alinit)
% NEWTON Newton minimum search with Armijo line search
%   [x,xk]=Newton(f,fp,fpp,x0,tol,maxiter,tau,be,alinit) finds an
%   approximate minimum of the function f with gradient fp and
%   Hessian fpp, starting at the initial guess x0. The remaining
%   parameters are optional and default values are used if they are
%    omited. xk contains all the iterates of the method.

if nargin<9, alinit=1; end;
if nargin<8, be=0.1; end;
if nargin<7, tau=0.5; end;
if nargin<6, maxiter=100; end;
if nargin<5, tol=1e-6; end;

x=x0;
xk=x0;
```

```
g=feval(fp,x);
p=-feval(fpp,x)\g;
k=0;
while norm(d)>tol & k<maxiter
  al=alinit;
  while feval(f,x+al*p)>feval(f,x)-al*be*p'*p
    al=tau*al;
  end;
  x=x+al*p;
  g=feval(fp,x);
  p=-feval(fpp,x)\g;
  k=k+1;
  xk(:,k+1)=x;
end;
```

b) *Implémenter la méthode de Quasi-Newton rang 1 symétrique :*

```
function [x,xk]=Rank1(f,fp,x0,tol,maxiter,tau,be,alinit)
% RANK1 Minimization with Quasi Newton Symmetric Rank 1 and Armijo
%  [x,xk]= Rank1 (f,fp,x0,tol,maxiter,tau,be,alinit) finds an
%   approximate minimum of the function f with gradient fp
%  starting at the initial guess x0. The remaining parameters
%  are optional and default values are used if they are omitted.
%   xk contains all the iterates of the method.

if nargin<8, alinit=1; end;
if nargin<7, be=0.1; end;
if nargin<6, tau=0.5; end;
if nargin<5, maxiter=100; end;
if nargin<4, tol=1e-6; end;

x=x0;
xk=x0;
g=feval(fp,x);
B=eyes(length(x));
p=-g;
k=0;
while norm(d)>tol & k<maxiter
  al=alinit;
  while feval(f,x+al*p)>feval(f,x)-al*be*p'*p
```

```
    al=tau*al;
  end;
  x=x+al*p;
  s=xk(:,k+1)-x;
  y=feval(f,xk(:,k+1))-feval(f,x);
  B=B+(y-B*s)*(y-B*s)'/(s'*(y-B*s));
  g=feval(fp,x);
  p=-B\g;
  k=k+1;
  xk(:,k+1)=x;
end;
```

c) *Implémenter la méthode de Quasi-Newton BFGS*

```
function [x,xk]=BFGS(f,fp,x0,tol,maxiter,tau,be,alinit)
% BFGS Minimization with Quasi Newton BFGS and Armijo
%   [x,xk]= BFGS(f,fp,x0,tol,maxiter,tau,be,alinit) finds an
%    approximate minimum of the function f with gradient fp
%   starting at the initial guess x0. The remaining parameters
%   are optional and default values are used if they are omitted.
%    xk contains all the iterates of the method.

if nargin<8, alinit=1; end;
if nargin<7, be=0.1; end;
if nargin<6, tau=0.5; end;
if nargin<5, maxiter=100; end;
if nargin<4, tol=1e-6; end;

x=x0;
xk=x0;
g=feval(fp,x);
B=eyes(length(x));
p=-g;
k=0;
while norm(d)>tol & k<maxiter
  al=alinit;
  while feval(f,x+al*p)>feval(f,x)-al*be*p'*p
    al=tau*al;
  end;
  x=x+al*p;
```

```
    s=xk(:,k+1)-x;
    y=feval(f,xk(:,k+1))-feval(f,x);
    B=B+(y-B*s)*(y-B*s)'/(s'*(y-B*s));
    g=feval(fp,x);
    p=-B\g;
    k=k+1;
    xk(:,k+1)=x;
  end;
```

Indication : Pour les méthodes de quasi Newton, on initialisera la matrice B_0 à l'identité.

Dans les exercices d'implémentation

- Afin de préserver l'uniformité des interfaces et permettre le remplacement transparent d'une implémentation par une autre, l'en-tête fourni doit impérativement être respecté et reproduit dans le code. Vous pouvez copier-coller depuis le PDF disponible sur la page web.
- Les programmes et les résultats doivent être rendus imprimés.

Evaluation du cours d'optimisation

- Les exercices : Les séries d'exercices rendues en retard seront comptées comme non rendues (*i.e.*, note 1 sur 6) dans le calcul de la note finale.
- Un examen oral durant la session d'examens sur le cours.

La note finale est de : $\frac{1}{5}$(exercices.) + $\frac{4}{5}$(note examen oral).

Assistant :	Kévin Santugini
Email :	`Kevin.Santugini@math.unige.ch`
Page web :	`http://www.unige.ch/~santugin/index.php?page=enseignement`