Course on Optimization — University of Mauritius, January 2018
Laurent DUMAS, Versailles University

TD1: descent methods

Exercice 1

On cherche à minimiser sur $\mathbb{R}^3$ la fonction :

$$J(x, y, z) = x^2 + 2y^2 + z^2 - x + y - z$$

1. Donner, en le justifiant, la solution exacte au problème considéré.
2. On veut utiliser la méthode de descente suivant la direction du gradient avec une stratégie de type backtracking avec condition d'Armijo pour approcher la solution à partir du point $X_0 = (0, 1, 1)$. Que vaut la direction de descente à la première itération ?
   On rappelle que la condition d'Armijo pour un point de départ X_0 et une direction de descente d s'écrit :
   $$J(X_0 + \alpha d) \leq J(X_0) + \beta\alpha < d, \nabla J(X_0) >$$
   Calculer explicitement puis représenter graphiquement dans le cas présent la fonction $\alpha \mapsto J(X_0 + \alpha d)$ ainsi que les valeurs de α vérfiant la condition d'Armijo pour $\beta = 0.1$.
3. On prend $\alpha_{init} = 1$ et $\tau = 0.1$ (constante de backtracking).
   Que vaut X_1 ?

Exercice 2

On considère f une fonction C^1 de $\mathbb{R}^n$ dans $\mathbb{R}$, telle que $\lim_{||x|| \to +\infty} f(x) = +\infty$. On note g la fonction gradient de f définie de $\mathbb{R}^n$ dans $\mathbb{R}^n$. On suppose que g est Lipschitzienne sur tout ensemble $S_{x_0} = \{x \in \mathbb{R}^n, f(x) \leq f(x_0)\}$.

1. Montrer que f possède un minimum global x^* pour lequel $g(x^*) = 0$.

2. On cherche à construire une suite $(x_k)_{k\in\mathbb{N}}$ convergeant vers un minimum (local ou global) de f. On suppose qu'il est possible de définir correctement celle-cià partir de la donnée de $x_0 \in \mathbb{R}^n$ quelconque et de la relation :

$$x_{k+1} = x_k + t_k d_k$$

où d_k désigne une direction de descente (c'està dire telle que $(d_k, g(x_k)) < 0$) et t_k le pas dans cette direction supposé satisfaire la conditions suivante :

$$q(t_k) \leq q(0) + m_1 t_k q'(0) \quad \text{et} \quad q'(t_k) \geq m_2 q'(0)$$

où on a noté $q(t) = f(x_k + td_k)$ et où m_1 et m_2 sont deux réels tels que $0 < m_1 < m_2 < 1$.

Donner un exemple graphique des valeurs de t_k admissibles dans le cas d'une fonctionà une variable tracée 'à la main'.

3. On chercheà montrer que la méthode de descente ainsi construite est convergente vers un point critique de f lorsque $d_k = -g(x_k) = -g_k$ (direction du gradient).

3.1 Montrer que $q'(0) = -||g_k||^2$ dans ce cas.

3.2 Montrer que $m_1||g_k|| \cdot ||x_{k+1} - x_k|| \leq f(x_k) - f(x_{k+1})$.

3.3 Montrer que $(1 - m_2)||g_k|| \leq L||x_{k+1} - x_k||$ où L désigne la constante de Lipschitz de g dans S_{x_0}.

3.4 En déduire que $\lim\limits_{k \to +\infty} g_k = 0$.

Question subsidiaire : proposer un algorithme permettant de construire la suite x_k, c'està dire en particulier de déterminer un réel t_k convenable.

Exercice 3

1. Représenter graphiquement quelques lignes de niveau de la fonction
$$f(x, y) = 2(x - 1)^2 + y^2 + 1$$
puis représenter pour un point sur une des lignes de niveau, le gradient en ce point et un exemple de direction de descente.
2. Soit f une fonction différentiable de $\mathbb{R}^n$ dans $\mathbb{R}$ et $x \in \mathbb{R}^n$. On suppose que $d \in \mathbb{R}^n$ est tel que
$$||\nabla f(x) + d|| \leq ||\nabla f(x)||$$
Montrer que d est une direction de descente de f en x.
3. Soit f une fonction différentiable et <u>convexe</u> de $\mathbb{R}^n$ dans $\mathbb{R}$. Soient x et y dans $\mathbb{R}^n$ tels que $f(y) < f(x)$. Montrer que $y - x$ est une direction de descente de f en x.

Exercice 4

On cherche à minimiser sur $\mathbb{R}^3$ la fonction :

$$J(x, y, z) = x^4 + 2y^4 + z^4 - 2x + y - z$$

1. Donner la solution exacte au problème considéré
2. On veut utiliser la méthode de Newton avec un pas $\alpha = 1$ pour approcher la solution à partir du point $X_0 = (1, 1, 1)$. Que vaut la direction de descente à la première itération ? Vérifier qu'il s'agit bien d'une direction de descente.
3. Que vaut X_1 ?

Exercice 5

On considère le programme suivant :

```
function y=J(x)
  y=10*(x(2)-x(1)^2)^2+(x(1)-1)^2;
endfunction
//
function y=nablaJ(x)
  y=[-40*x(1)*(x(2)-x(1)^2)+2*(x(1)-1);20*(x(2)-x(1)^2)];
endfunction
//
function alpha=backtrackA(x,J,Jx,p,gx,betaA,tau,alphainit)
alpha=alphainit;
count=1;
while(J(x+alpha*(p))>Jx+betaA*alpha*(p'*gx))&(count<20)
  count=count+1;
  alpha=tau*alpha;
end
endfunction
//
function p=BFGSdescent(B,gx)
p=linsolve(B,gx); // solve Bp=-gx
endfunction
//
betaA=0.0001; alphainit=1; tau=0.7; epsilon=1E-3;
//
n=2; x=[0;1]; // starting point
//
disp('initial value of x');disp(x);
gx=nablaJ(x);
Jx=J(x);
count=1;
B=eye(n,n);
//
while (norm(gx)>epsilon)&(count<200)
 p=BFGSdescent(B,gx); // p=-gx if steepest descent
 alpha=backtrackA(x,J,Jx,p,gx,betaA,tau,alphainit); // linesearch
 xold=x;//
 gxold=gx;
 x=x+alpha*(p);
  Jx=J(x);
 gx=nablaJ(x);
 count=count+1;
 y=(gx-gxold);s=x-xold;
 B=B+(1/(s'*y))*y*y'-(1/(s'*(B*s)))*B*s*s'*B;
end
disp('final value of x:');disp(x)
disp('iteration number');disp(count)
//
```

1. Expliquer sous forme mathématique le principe de l'algorithme d'optimisation utilisé.
2. Que donne le résultat de la première itération ?
3. On applique l'algorithme précédent pour une fonction de $\mathbb{R}$ dans $\mathbb{R}$ dérivable dont on recherche le minimum. Montrer que la méthode revient à résoudre l'équation $f'(x) = 0$ par la méthode de la sécante.