06-720 **Advanced Process Systems Engineering** **Spring 2011**
Homework 1 **Due: 1/19/11**

1. Consider the minimization problem: min $|x| + |y|$ s.t. $x^2 + y^2 = 1$. By introducing binary decision variables to handle the absolute value terms, reformulate this problem as a mixed integer programming problem of the form.

2. Examine whether the following functions are convex or not.

   $x^2 + ax + b;\ x \in R$
   $x^3;\ x \in R$
   $x^4;\ x \in R$
   $log(x);\ x \in (0;\ 1]$

3. Consider the quadratic function:

$$f(x) = 3x_1 + x_2 + 2x_3 + 4x_1^2 + 3x_2^2 + 2x_3^2 + (M-2)x_1x_2 + 2x_2x_3$$

For M= 0 find the eigenvalues and eigenvectors and any stationary points. Are the stationary points local optima? global optima? Find the path of optimal solutions as M increases.

4. Prove that a Hessian matrix which is positive definite has positive eigenvalues.

5. Show that if B^k is positive definite, then $\cos \theta^k > 1/\kappa(B^k)$ where $\kappa(B^k)$ is the condition number of B^k, based on the 2-norm.

6. Derive a stepsize rule for the Armijo linesearch that minimizes the quadratic interpolant from the Armijo inequality.

Homework 1
Solution

1) Min $|x|+|y|$ s.t. $x^2+y^2=1$

Let $x = x_+ - x_-$
$y = y_+ - y_-$
and add $\beta_x \in \{0,1\}$, $\beta_y \in \{0,1\}$
~~with~~
(a) $0 \le x_+ \le \beta_x$, $0 \le x_- \le 1-\beta_x$
(b) $0 \le y_+ \le \beta_y$, $0 \le y_- \le 1-\beta_y$

Reformulate to: Min $(x_+ + x_-) + (y_+ + y_-)$
s.t. (a), (b)
$(x_+ - x_-)^2 + (y_+ - y_-)^2 = 1$

2) $x^2 + ax + b$ - convex due to p.d. Hessian
x^3 - nonconvex, Hessian not psd
x^4 - convex, Hessian psd
$\log(x)$ - nonconvex $\in (0,1]$, Hessian not psd.

3) Min $[3 \ 1 \ 2]x + x^T \begin{bmatrix} 4 & \frac{M-2}{2} & 0 \\ \frac{M-2}{2} & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix} x$

$$\det(H-\lambda I) = (4-\lambda)(3-\lambda)(2-\lambda) - (4-\lambda) - (M-2)^2(1/2 - \lambda/4)$$

$$= \left(20 - \frac{(M-2)^2}{2}\right) + \left(\frac{(M-2)^2}{4} - 25\right)\lambda + 9\lambda^2 - \lambda^3 = 0$$

for $M=2$, $\lambda = 1.27, 3.0, 4.72$ (p.d)

3.

μ	λ
0	1.27, 3, 4.72
1	1.34, 3.36, 4.3
2	1.38, 3.6, 4.0 — (symmetric about $\mu = 2$)
3	1.34, 3.36, 4.3
4	1.27, 3, 4.72
5	1.10, 2.70, 5.20
6	0.85, 2.48, 5.67

positive definite for $\mu \in (-4.324, 8.324)$

Solutions

$$\begin{bmatrix} 4 & \mu-2/2 & 0 \\ \mu-2/2 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix} x = - \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

$$B x^* = -a$$

Solutions for x^* for various values of $\mu \in [0, 6]$ are given on next page.

4. Hessian is symmetric $(H = H^T)$ and can be decomposed as:

$$H = V \Lambda V^T \text{ where } V^{-1} = V^T$$

Therefore if $p^T H p > 0$, then

$$0 < p^T H p = p^T V \Lambda V^T p = \bar{p}^T \Lambda \bar{p}$$
$$= \sum \lambda_i \bar{p}_i^2 \text{ where } \bar{p} = V^T p,$$

Now if $\lambda_i \leq 0$, then choosing $\bar{p} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = e_i$ leads to $0 < \lambda_i e_i^T e_i = \lambda_i \leq 0$ which is a contradiction.

```
m = 0
:        x          a     :=
1   -0.833333    3
2   -0.333333    1
3   -0.833333    2

b :=
1 1    4
1 2   -1
1 3    0
2 1   -1
2 2    3
2 3    1
3 1    0
3 2    1
3 3    2

m = 1
:        x          a     :=
1   -0.769231    3
2   -0.153846    1
3   -0.923077    2
;

b :=
1 1    4
1 2   -0.5
1 3    0
2 1   -0.5
2 2    3
2 3    1
3 1    0
3 2    1
3 3    2

m = 2
:        x              a     :=
1   -0.75            3
2    2.77556e-17     1
3   -1               2
;

b :=
1 1    4
1 2    0
1 3    0
2 1    0
2 2    3
2 3    1
3 1    0
3 2    1
3 3    2

m = 3
:        x          a     :=
1   -0.769231    3
2    0.153846    1
3   -1.07692     2
;

b :=
1 1    4
1 2    0.5
1 3    0
2 1    0.5
2 2    3
2 3    1
3 1    0
3 2    1
3 3    2

m = 4
:        x          a     :=
1   -0.833333    3
2    0.333333    1
3   -1.16667     2
;

b :=
1 1    4
1 2    1
1 3    0
2 1    1
2 2    3
2 3    1
3 1    0
3 2    1
3 3    2

m = 5
:        x          a     :=
1   -0.967742    3
2    0.580645    1
3   -1.29032     2
;

b :=
1 1    4
1 2    1.5
1 3    0
2 1    1.5
2 2    3
2 3    1
3 1    0
3 2    1
3 3    2

m = 6
:     x       a     :=
1   -1.25    3
2    1       1
3   -1.5     2
;

b :=
1 1    4
1 2    2
1 3    0
2 1    2
2 2    3
2 3    1
3 1    0
3 2    1
3 3    2
```

5. B^k is positive def.

$$\cos\theta^k = \frac{-\nabla f(x^k)^T p}{\|\nabla f(x^k)\|\,\|p\|}$$

– a few identities

$$p^T B^k p = \|B^{1/2} p\|^2, \quad \|B^{1/2}\|^2 = \|B^k\|$$

$$B^k p = -\nabla f(x^k)$$

$$-\nabla f(x)^T p = p^T B^k p = \|B^{1/2} p\|^2$$

– norm inequalities

$$p = (B^{-1/2}) B^{1/2} p$$

$$\|p\| \le \|B^{-1/2}\|\,\|B^{1/2} p\| \implies \|B^{1/2} p\| \ge \frac{\|p\|}{\|B^{-1/2}\|}$$

so $$\cos\theta^k = \frac{p^T B^k p}{\|B^k p\|\,\|p\|} \ge \frac{\|B^{1/2} p\|^2}{\|B^k\|\,\|p\|^2}$$

$$\ge \frac{\|p\|^2}{\|B^k\|\,\|B^{-1/2}\|^2\|p\|^2} = \frac{1}{\|B^k\|\,\|B^{-1}\|}$$

$$= 1/\kappa(B^k)$$

6. Quadratic interpolation

Armijo inequality not satisfied

$$f(x^k + \bar\alpha p) > f(x^k) + \nabla f(x^k)^T p\,(\eta\bar\alpha)$$

$a\alpha^2 + b\alpha + c$ for $\alpha \in (0, \bar\alpha]$

$\alpha = 0$, $c = f(x^k)$

$\alpha = 0$, $b = \nabla f(x^k)^T p$

$\alpha = \bar\alpha$, $a\bar\alpha^2 + b\bar\alpha + c = f(x^k + \bar\alpha p)$

$$\alpha_q = \frac{-b}{2a} = \frac{-\nabla f(x^k)^T p\,\bar\alpha^2}{2\left(f(x^k+\bar\alpha p) - f(x^k) - \bar\alpha\nabla f(x^k)^T p\right)}$$