Name:
Code:

Nonlinear Programming, MATP6600/DSES6780
Final Exam, Thursday, December 9, 2004.

Please do all three problems. You must show all work to obtain full credit. Results from class or the text may be used if properly stated. No books or calculators allowed. The exam lasts one hour and 50 minutes.

Solutions

Q1	
Q2	
Q3	
Total	
Grade	

1. (25 points)
   Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function.

   (a) (5 points) Let $\bar{x} \in \mathbb{R}^n$ Since f is convex, it has a subgradient $\xi(\bar{x})$ at $\bar{x}$. Give the subgradient inequality satisfied by $f(x)$ for any $x \in \mathbb{R}^n$.

   (b) (20 points) Define the function $g(x,t) = tf(\frac{1}{t}x)$, where $x \in \mathbb{R}^n$ and t is a positive scalar. Show that g is a convex function of x and t.

(a) $f(x) \geq f(\bar{x}) + \xi(\bar{x})^T(x - \bar{x})$

(b) Pick $\bar{x}, \bar{t}$, and try to find a subgradient there.

$$g(x,t) = t f\left(\tfrac{1}{t}x\right) \geq t f\left(\tfrac{1}{\bar{t}}\bar{x}\right) + t\xi\left(\tfrac{1}{\bar{t}}\bar{x}\right)^T\left(\tfrac{1}{t}x - \tfrac{1}{\bar{t}}\bar{x}\right)$$

$$= \bar{t} f\left(\tfrac{1}{\bar{t}}\bar{x}\right) + \xi\left(\tfrac{1}{\bar{t}}\bar{x}\right)^T(x - \bar{x})$$

$$+ f\left(\tfrac{1}{\bar{t}}\bar{x}\right)(t - \bar{t}) - (t - \bar{t})\,\xi\left(\tfrac{1}{\bar{t}}\bar{x}\right)^T\left(\tfrac{1}{\bar{t}}\bar{x}\right)$$

$$= g(\bar{x},\bar{t}) + \xi\left(\tfrac{1}{\bar{t}}\bar{x}\right)^T(x - \bar{x})$$

$$+ \left(f\left(\tfrac{1}{\bar{t}}\bar{x}\right) - \tfrac{1}{\bar{t}}\,\xi\left(\tfrac{1}{\bar{t}}\bar{x}\right)^T\bar{x}\right)(t - \bar{t})$$

So $g(x,t)$ has a subgradient at $(\bar{x},\bar{t})$,

and so it is convex.

2. (35 points) Consider the nonlinear programming problem

$$\begin{array}{lrcll} \min & f(x) & & & \\ \text{subject to} & h_i(x) & = & 0 \quad \text{for } i = 1, \ldots, p & (NLP) \end{array}$$

where f and each h_i are functions from $I\!R^n$ to $I\!R$, and $x \in I\!R^n$.

(a) (5 points) Let $\bar{x} \in I\!R^n$. Assume $\{\nabla h_i(\bar{x}), i = 1, \ldots, p\}$ is a linearly independent set of vectors. What are the first order conditions that must be satisfied if $\bar{x}$ is a local minimum?

(b) (15 points) The SQP subproblem to find a direction from a point $\bar{x}$ is

$$\begin{array}{lrcll} \min & \nabla f(\bar{x})^T d & + & 0.5 d^T Q d & \\ \text{subject to} & \nabla h_i(\bar{x})^T d & = & -h_i(\bar{x}) \quad \text{for } i = 1, \ldots, p & (SQP) \\ & d^T d & \leq & r^2 & \end{array}$$

where Q is the Hessian of the Lagrangian and r is the radius of a trust region.

i. (15 points) Assume the feasible point $\bar{x}$ satisfies the first order conditions with multipliers $v \in I\!R^p$. Assume further that $\bar{x}$ is a local minimum and satisfies the second order sufficient conditions for (NLP). Show that ~~the~~ an optimal solution to (SQP) is $d = 0$.

ii. (15 points) Let $n = 2$, $p = 1$, and $r = 5$. Let $f(x) = 20e^{0.3x_1 + 0.4x_2}$ and $h(x) = -x_1^2 - x_2^2 + 25$. The feasible point $x = (3, 4)$ satisfies the first order KKT conditions with multiplier $v = e^{2.5}$. Show that the optimal solution to (SQP) is nonzero.

(a) CQ holds, since $\nabla h_i(\bar{x})$ are linearly independent.

Need: $$\nabla f(\bar{x}) + \sum_{i=1}^{p} \bar{v}_i \nabla h_i(\bar{x}) = 0$$
$$h_i(\bar{x}) = 0 \quad i = 1, \ldots, p.$$

for some $\bar{v} \in I\!R^p$.

(this page deliberately left blank)

(b) (i) $h_i(\bar{x}) = 0$

So any d feasible in (SQP) satisfies

$$d^T \nabla h_i(\bar{x}) = 0 \quad \forall i.$$

Thus, $d^T \nabla f(\bar{x}) = -\sum \bar{v}_i d^T \nabla h_i(\bar{x}) = 0$

Also, since $\bar{x}$ satisfies the second order conditions, we have that any d satisfying $\nabla h_i(\bar{x})^T d = 0$ also satisfies $d^T Q d \geq 0$.

Thus, any d feasible in (SQP) has value ≥ 0.

$d = 0$ has value 0, so it is optimal.

(this page deliberately left blank)

$$\min \ 20e^{0.3x_1+0.4x_2}$$
$$\text{s.t.} \ -x_1^2-x_2^2+25=0.$$

2 (b) (ii) $\nabla f = 20e^{0.3x_1+0.4x_2}\begin{bmatrix}0.3\\0.4\end{bmatrix} = 2e^{0.3x_1+0.4x_2}\begin{bmatrix}3\\4\end{bmatrix}$

$$\nabla h = \begin{bmatrix}-2x_1\\-2x_2\end{bmatrix}$$

At $\bar{x}=(3,4)$: $\nabla f = 2e^{2.5}\begin{bmatrix}3\\4\end{bmatrix}$. $\nabla h = \begin{bmatrix}-6\\-8\end{bmatrix}$

$\nabla^2 f = 2e^{0.3x_1+0.4x_2}\begin{bmatrix}9 & 12\\12 & 16\end{bmatrix}$ $\nabla^2 f(\bar{x}) = 2e^{2.5}\begin{bmatrix}9 & 12\\12 & 16\end{bmatrix}$

$\nabla^2 h = -2\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$

(SQP) $\min \ 2e^{2.5}(3d_1+4d_2)+\frac{1}{2}d^T\nabla^2 f(\bar{x})d + \frac{1}{2}v\,d^T\nabla^2 h(\bar{x})d$

s.t. $-6d_1-8d_2=0$

$d^Td \leq 25$

Now, $\nabla h(\bar{x})^T d = 0 \Rightarrow d^T\begin{bmatrix}-6\\-8\end{bmatrix}=0$, so $d = \alpha\begin{bmatrix}4\\-3\end{bmatrix}$ $\begin{bmatrix}4\\+3\end{bmatrix}$

5

Then $d^T\nabla f(\bar{x}) = 0$ and $d^T\nabla^2 f(\bar{x}) = [0 \ \ 0]$

5

Thus, $d^TQd = -2v\,d^Td = -50v\alpha^2$

So the optimal solution is to take α as large in magnitude as possible

5

3. (40 points) We want to solve the unconstrained problem

$$\min f(x) := \frac{1}{4}x_1^4 - \frac{7}{3}x_1^3 + 6x_1^2 + \frac{1}{2}x_2^2. \qquad (UNLP)$$

The gradient and Hessian of f are

$$\nabla f(x) = \begin{bmatrix} x_1^3 - 7x_1^2 + 12x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad \nabla^2 f(x) = \begin{bmatrix} 3x_1^2 - 14x_1 + 12 & 0 \\ 0 & 1 \end{bmatrix}.$$

Let $\bar{x} = (2, 1)$.

(a) (10 points) Show that the Newton direction is not a descent direction at $\bar{x}$. Consider using a matrix Q instead of $\nabla^2 f(x)$ in the definition of the Newton direction. What must Q satisfy in order to ensure that the resulting direction is a descent direction?

$$\nabla f(\bar{x}) = \begin{bmatrix} 8 - 28 + 24 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \qquad 2$$

$$\nabla^2 f(\bar{x}) = \begin{bmatrix} 12 - 28 + 12 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix} \qquad 2$$

$$d = -\nabla^2 f(x)^{-1} \nabla f(\bar{x}) = -\begin{bmatrix} -\frac{1}{4} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad 2$$

$\nabla f(\bar{x})^T d = 3 > 0$, so not a descent direction. 2

Sufficient to have Q positive definite, since then 2

~~Ax~~

$$d^T \nabla f(\bar{x}) \cancel{= d^T Q^{-1} d < 0}$$
$$= -\nabla f(\bar{x})^T Q^{-1} \nabla f(\bar{x}) < 0.$$

(b) (10 points) A direction can also be found by using a trust region scheme:

$$\begin{array}{llcl} \min & \nabla f(\bar{x})^T d & + & \frac{1}{2} d^T \nabla^2 f(\bar{x}) d \\ \text{subject to} & d^T d & \leq & r^2 \end{array} \qquad (TRP)$$

Consider the general case, for any $f(x)$. Let $\bar{d}$ be the optimal solution to this problem. Show that $\bar{d}^T \nabla f(\bar{x}) \leq 0$.

Assume $\bar{d}^T \nabla f(\bar{x}) > 0$.

Let $\tilde{d} = -\bar{d}$.

Then $\tilde{d}^T \tilde{d} = \bar{d}^T \bar{d}$ so $\tilde{d}$ is feasible

Also, $\tilde{d}^T \nabla^2 f(\bar{x}) \tilde{d} = \bar{d}^T \nabla^2 f(\bar{x}) \bar{d}$

and $\tilde{d}^T \nabla f(\bar{x}) = -\bar{d}^T \nabla f(\bar{x}) < 0$.

So $\tilde{d}$ is better than $\bar{d}$, so $\bar{d}$ cannot be optimal.

(c) (10 points)
Return now to problem $(UNLP)$. Let $r = 10$ in (TRP) and let $\bar{d}$ solve (TRP). Show that $f(\bar{x} + \bar{d}) > f(\bar{x})$. How would you suggest using $\bar{d}$ to find a new point? (Hint: It is not necessary to solve (TRP). Note that $f(2, 1) = 9\frac{5}{6}$, and that if $x_1 > 7$ or $x_1 < -3$ then $f(x) > 20$.)

From (a), $\nabla f(\bar{x}) = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ and $\nabla^2 f(\bar{x}) = \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix}$.

So (TRP) is: $\min \; 4d_1 + d_2 - 2d_1^2 + \frac{1}{2} d_2^2$ 4
$\quad \text{s.t.} \; d_1^2 + d_2^2 \leq 100.$

Let $g(d_1, d_2) = 4d_1 + d_2 - 2d_1^2 + \frac{1}{2} d_2^2$.

If $|d_1| \leq 5$: 1
$g(d_1, d_2) \geq -20 - 10 - 50 = -80$

If $d_1 = -10$, $d_2 = 0$: 2
$g(-10, 0) = -40 - 200 = -240.$

So no point with $|d_1| \leq 5$ can be optimal.

So $\bar{x}_1 + \bar{d}_1$ is either >7 or <-3, 1
and thus $f(\bar{x} + \bar{d}) > 20 > 9\frac{5}{6} = f(\bar{x})$.

Use a line search to find α and update to $\bar{x} + \alpha \bar{d}$. 2

(d) (10 points)
Now consider choosing a step in $(UNLP)$ from $\bar{x} = (2, 1)$ using a proximal point approach, so we find d by solving

$$\min \quad \nabla f(\bar{x})^T d + \frac{1}{2} d^T \nabla^2 f(\bar{x}) d + \frac{1}{2}\mu d^T d. \qquad (PP)$$

Let $\mu = 4.1$ and let $\bar{d}$ solve (PP). Show that $f(\bar{x} + \bar{d}) > f(\bar{x})$.

404

(Hint: most of part (c) hint.)

~~Let~~ g

For (UNLP) get (PP) to be:

$$\min \quad 4d_1 + d_2 - 2d_1^2 + \tfrac{1}{2}d_2^2 + \cancel{2.05}\,2.2\, d_1^2 + \cancel{2.05}\,2.2\, d_2^2$$

or $$\min \quad 4d_1 + d_2 + \cancel{0.05}\,0.2\, d_1^2 + \cancel{2.55}\,2.7\, d_2^2 =: q(d)$$ 4

This is a convex quadratic function.

So differentiate and set derivative to zero

$$\frac{dq}{dd_1} = 4 + 0.4 d_1 \quad \text{so} \quad \bar{d}_1 = -10$$ 4

$$\frac{dq}{dd_2} = 1 + 5.4 d_2 \quad \text{so} \quad \bar{d}_2 = -\frac{1}{5.4}$$

From argument in part (c), $f(\bar{x} + \bar{d}) > f(\bar{x})$.

2