C

Name:

Nonlinear Programming, MATP6600/DSES6780
Final Exam, Friday, December 11, 2009.

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		40
Q2		40
Q3		20
Total		100

1. (40 points; each part is worth 10 points.) Consider the unconstrained problem

$$\min\ f(x) := \frac{1}{2}(x-2)^2 + x^4, \qquad (P)$$

where $x \in \mathbb{R}$. We are going to investigate a bundle method for solving this problem. Let $\bar{x} = 0$ be the current point. A bundle subproblem for (P) is

$$\begin{array}{lrcrcr} \min_{x,z} & z & + & \frac{1}{2}(x-\bar{x})^2 & & \\ \text{subject to} & -z & - & 2x & \leq & -2 \end{array} \qquad (BQP)$$

where z and x are variable scalars.

(a) Show that the optimal solution to the bundle subproblem (BQP) is to take $x = 2$.

(b) Show that the objective function value for $x = 2$ predicted by the bundle subproblem is noticeably less than the actual value $f(2)$.

(c) What is the valid subgradient inequality generated at the point $x = 2$? Show that the inequality is violated by the solution to (BQP).

(d) Add the constraint found in part 1c to (BQP) and find the new optimal solution, assuming a null step was taken so we still have $\bar{x} = 0$.

(a) (BQP) is convex, constraints are linear. So suffices to look for KKT point.

KKT conditions: $\left.\begin{array}{l} 1-\lambda = 0 \\ x - \bar{x} - 2\lambda = 0 \\ \lambda(2 - z - 2x) = 0 \end{array}\right\} \Rightarrow \lambda = 1,\ x = 2,\ z = -2$

(b) Predicted value: $z + \frac{1}{2}(x-\bar{x})^2 = -2 + \frac{1}{2}(2-0)^2 = 0$

Actual value: $f(2) = 16$

Gap: 16.

(c) $f(x) \geq f(\bar{x}) + \frac{df}{dx}\Big|_{x=2}(x-2) = 16 + 32(x-2) = 32x - 48.$

So get constraint: $z \geq 32x - 48$, or $-z + 32x \leq 48$

At $x = 2$, this implies $z \geq 16$, violated by our $z = -2$.

(d) $\min\ z + \frac{1}{2}(x)^2$

s.t. $-z - 2x \leq -2$

$-z + 32x \leq +48$

KKT conditions:

$1 - \lambda_1 - \lambda_2 = 0$ (3)

$x - 2\lambda_1 + 32\lambda_2 = 0$ (4)

$\lambda_1(-z - 2x + 2) = 0$

$\lambda_2(-z + 32x - 48) = 0$

$\lambda_1, \lambda_2 \geq 0$, + feasibility

$\lambda_2 = 0 \Rightarrow x = 2, z = -2$ as before. Violates new constraint.

$\lambda_1 = 0 \Rightarrow \lambda_2 = 1 \Rightarrow x = -32$

$\Rightarrow z = -(32)^2 - 48$,

which violates $-z - 2x \leq -2$

So $\lambda_1 > 0$, $\lambda_2 > 0$,

so $-z - 2x + 2 = 0$ (1)

$-z + 32x - 48 = 0$ (2)

(2) − (1) $\Rightarrow x = 50/34 = 25/17$

$\Rightarrow z = 2 - 2x = -16/17$

32(3) + (4) $\Rightarrow 34\lambda_1 = 33\frac{8}{17}$,

so $0 < \lambda_1 < 1$ and $\lambda_2 = 1 - \lambda_1 > 0$.

2. (40 points.) Consider the quadratic programming problem

$$\begin{array}{ll} \min & c^T x + \frac{1}{2} x^T Q x \\ \text{s.t.} & Ax \geq b \end{array} \qquad (QP)$$

where x and c are n-vectors, Q is symmetric and positive semidefinite, b is an m-vector, and A is $m \times n$.

(a) (20 points) Show that the Lagrangian dual of (QP) can be written

$$\begin{array}{lrcl} \max_{x,u} & b^T u - \frac{1}{2} x^T Q x & & \\ \text{subject to} & c + Qx - A^T u & = & 0 \\ & u & \geq & 0 \end{array} \qquad (QD)$$

(b) (10 points) Let $\bar{x}$, $\bar{u}$ be a feasible solution to (QD). Assume there exists an n-vector d_x with $d_x^T Q d_x = 0$. Show that $Qd_x = 0$. Assume there also exists a nonnegative m-vector d_u with $b^T d_u > 0$ and $A^T d_u = 0$. What does this imply about (QD) and about (QP)?

(c) (10 points) Assume there exists a vector d satisfying $c^T d < 0$, $d^T Q d = 0$, and $Ad \geq 0$. What does this imply about (QD) and hence about (QP)?

(a) $\theta(u) = \inf_{x \in \mathbb{R}^n} \; c^T x + \frac{1}{2} x^T Q x + u^T (b - Ax)$

If this convex quadratic function has a minimizer, it satisfies $c + Qx - A^T u = 0$.

Then $c^T x + \frac{1}{2} x^T Q x + u^T (b - Ax) = x^T (c + Qx - A^T u) + b^T u - \frac{1}{2} x^T Q x$

So dual problem is $\max_{u,x} \; b^T u - \frac{1}{2} x^T Q x$ s.t. $c + Qx - A^T u = 0$, $u \geq 0$

(b) We can factor $Q = U \Lambda U^T$, $\underbrace{\text{cols of } U}_{u^i}$ are eigenvectors, Λ diagonal, nonzeroes are eigenvalues.

Then $d^T Q d = \sum_{i=1}^{n} \Lambda_{ii} (d^T u^i)^2$. So $d^T u^i = 0$ if $\Lambda_{ii} > 0$,

and then $Qd = \sum_{i=1}^{n} u_i \Lambda_{ii} d^T u^i = 0$, since $\Lambda_{ii} = 0$ or $d^T u_i = 0$ for each i.

Now $x = \bar{x}$, $u = \bar{u} + \alpha d_u$ is feasible in (QD) for any $\alpha \geq 0$, and value $\to +\infty$ as $\alpha \to \infty$.

So (QD) is unbounded, so (QP) is infeasible.

(c) $d^T Q d = 0 \Rightarrow Qd = 0 \Rightarrow 0 = d^T (c + Qx - A^T u) = c^T d - (Ad)^T u$ for any feasible u.

But $c^T d - (Ad)^T u < 0$. So (QD) is infeasible. Since (QP) is convex and there exists a strictly feasible primal solution, there is no duality gap. So (QP) is unbounded.

3. (20 points; each part is worth 10 points.) The function $f : \mathbb{R}^n \to \mathbb{R}$ is convex and the function $\phi : \mathbb{R} \to \mathbb{R}$ is convex and monotonically strictly increasing. Define $g(x) := \phi(f(x))$, so $g : \mathbb{R}^n \to \mathbb{R}$. The problems of minimizing $f(x)$ and of minimizing $g(x)$ are equivalent, with the same optimal points. Assume f, g, and ϕ are all smooth functions.

   (a) How do the steepest descent directions for $f(x)$ and $g(x)$ compare?

   (b) How do the Newton directions for $f(x)$ and $g(x)$ compare? (Hint: If M is an invertible $n \times n$ matrix and a is an n-vector, then

$$(M + aa^T)^{-1} = M^{-1} - \frac{1}{1 + a^T M^{-1} a} M^{-1} a a^T M^{-1}$$

   provided $1 + a^T M^{-1} a \neq 0$.)

(a) $\nabla g(x) = \left.\frac{d\phi}{dt}\right|_{t=f(x)} \nabla f(x).$

Since ϕ is increasing, $\frac{d\phi}{dt} > 0$, so this is a rescaling of the standard steepest descent direction.

(b) $\nabla^2 g(x) = \underbrace{\left.\frac{d\phi}{dt}\right|_{t=f(x)}}_{\sim \phi'} \nabla^2 f(x) + \underbrace{\left.\frac{d^2\phi}{dt^2}\right|_{t=f(x)}}_{\sim \phi''} \nabla f(x) \nabla f(x)^T$

By hint,

$$(\nabla^2 g(x))^{-1} = \frac{1}{\phi'}(\nabla^2 f(x))^{-1} - \frac{\phi''}{(\phi')^2\left(1 + \frac{\phi''}{\phi'} \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)\right)} \nabla^2 f(x)^{-1} \nabla f(x) \nabla f(x)^T \nabla^2 f(x)^{-1}$$

Newton direction:

$$-(\nabla^2 g(x))^{-1} \nabla g(x) = -\nabla^2 f(x)^{-1} \nabla f(x) + \frac{\phi'' \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)}{\phi' + \phi'' \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)} \nabla^2 f(x)^{-1} \nabla f(x)$$

$$= -\frac{\phi'}{\phi' + \phi'' \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)} \nabla^2 f(x)^{-1} \nabla f(x).$$

So get rescaling of Newton direction. Get exactly Newton direction if ϕ is linear.