MATH/AMATH 516
SECOND HOMEWORK SET

1. Consider minimizing the continuously differentiable function $f : I\!R^n \mapsto I\!R$ on $I\!R^n$. Let $x \in I\!R^n$ and $d \in I\!R^n$ be such that $\nabla f(x)^T d < 0$. These properties of f, x, and d are assumed and used in Parts (a) and (b) of this problem.

   Recall that in the backtracking line–search, we are given parameters $0 < \gamma < 1$ and $0 < c_1 < 1$ and we obtain an update to x, say x_+, of the form $x_+ = x + \lambda d$ where

$$\lambda = \begin{array}{ll} \max \gamma^k & \\ \text{subject to} & k \in \{1, 2, \ldots\},\ \text{and}\ f(x + \gamma^k d) - f(x) \le c_1 \gamma^k \nabla f(x)^T d\ . \end{array}$$

   The key inequality

$$f(x + \lambda d) - f(x) \le c_1 \lambda \nabla f(x)^T d \qquad (1)$$

   is called the Armijo–Goldstein inequality. A shortcoming of this step length is that it is unrelated to the one dimensional problem $\min\{f(x + \lambda d) \,:\, \lambda \ge 0\}$. In this regard, we will study methods that require the step length λ to satisfy both the Armijo–Goldstein inequality and an inequality of the form

$$\nabla f(x + \gamma^k d)^T d \ge c_2 \nabla f(x)^T d \qquad (2)$$

   for a given $c_2 \in (0, 1)$. The conditions (1) and (2) taken together are called the *weak Wolfe conditions*.

   (a) Show that if $0 < c_1 < c_2 < 1$ and the set $\{f(x + \lambda d) \,:\, \lambda \ge 0\}$ is bounded below, then the two conditions (1) and (2) can be satisfied simultaneously. In particular, show that the set

$$\left\{ \lambda \;\middle|\; \begin{array}{c} \lambda > 0, \nabla f(x + \lambda d)^T d \ge c_2 \nabla f(x)^T d,\ \text{and} \\ f(x + \lambda d) - f(x) \le c \lambda \nabla f(x)^T d \end{array} \right\}$$

   has non–empty interior.

   (b) Let $0 < c_1 < c_2 < 1$ and assume that the set $\{f(x + \lambda d) \,:\, \lambda \ge 0\}$ is bounded below. Show that the following bisection method is finitely terminating at a value for t at which the weak Wolfe conditions are satisfied.

   A Bisection Method for the Weak Wolfe Conditions

   INITIALIZATION: Choose $0 < c_1 < c_2 < 1$, and set $\alpha = 0$, $t > 0$, and $\beta = +\infty$.

   REPEAT
   - If $f(x + td) > f(x) + c_1 t f'(x; d)$,
     - set $\beta = t$ and reset $t = \frac{1}{2}(\alpha + \beta)$.
   - Elseif $f'(x + td; d) < c_2 f'(x; d)$,
     - set $\alpha = t$ and reset

$$t = \begin{cases} 2\alpha, & \text{if } \beta = +\infty \\ \frac{1}{2}(\alpha + \beta), & \text{otherwise.} \end{cases}$$

   - Else, STOP.

   END REPEAT

2. Consider the function

$$f(x) = \frac{1}{2}x^T Qx + c^T x,$$

where $Q \in I\!R^{n\times n}$ is symmetric and $c \in I\!R^m$.

(a) Under what condition on the matrix $Q \in I\!R^{n\times n}$ is f convex? Justify your answer.

(b) If Q is symmetric and positive definite, show that there is a nonsingular matrix L such that $Q = LL^T$.

(c) With Q and L as defined in the part (b), show that

$$f(x) = \frac{1}{2}\|L^T x + L^{-1}c\|_2^2 - \frac{1}{2}c^T Q^{-1}c.$$

(d) If f is convex, under what conditions is $\min_{x\in I\!R^n} f(x) = -\infty$?

(e) Assume that Q is symmetric and positive definite and let S be a subspace of $I\!R^n$. Show that $\bar{x}$ solves the problem

$$\min_{x\in S} f(x)$$

if and only if $\nabla f(\bar{x}) \perp S$.

3. Let $F : I\!R^n \mapsto I\!R^m$ be continuously differentiable, and let $\|\cdot\|$ be **any** norm on $I\!R^m$. In this problem we consider the function $f(x) = \|F(x)\|$ and properties of the the associated Gauss-Newton direction for minimizing f. Recall that $\bar{x} \in I\!R^n$ is a first-order stationary point for f if $0 \le f'(\bar{x}; d)$ for all $d \in I\!R^n$.

(a) Given $x, d \in I\!R^n$ and $t > 0$ show that

$$\left|\, \|F(x+td)\| - \|F(x) + tF'(x)d\| \,\right| \le \|F(x+td) - (F(x) + tF'(x)d)\| \;.$$

(b) Why is it true that

$$\lim_{t\to 0} \frac{\|F(x+td) - (F(x) + tF'(x)d)\|}{t} = 0 \;?$$

Hint: What is the definition of $F'(x)$?

(c) Use parts 3a and 3b to show that

$$f'(x;d) = \lim_{t\downarrow 0} \frac{\|F(x) + tF'(x)d\| - \|F(x)\|}{t} \;.$$

(d) Use part 3c and the convexity of the norm to show that

$$f'(x;d) \le \|F(x) + F'(x)d\| - \|F(x)\| \;.$$

Hint: $F(x) + tF'(x)d = (1-t)F(x) \; + \; t(F(x) + F'(x)d)$

(e) Use parts 3c and 3d to show that $0 \le f'(x;d)$ for all d if and only if $\|F(x)\| \le \|F(x) + F'(x)d\|$ for all d.

(f) If x is not a first-order stationary point for f and $\bar{d}$ solves

$$\mathcal{GN} \quad \min_{d\in I\!R^n} \|F(x) + F'(x)d\| \;,$$

show that $\bar{d}$ is a descent direction for f at x by showning that

$$f'(x;\bar{d}) \le \|F(x) + F'(x)\bar{d}\| - \|F(x)\| < 0 \;.$$

(g) Show that the problem

$$\min_{d \in \mathbb{R}^n} \|F(x) + F'(x)d\|_1$$

can be written as a linear program.

(h) Suppose at a given point $x \in \mathbb{R}^n$ one solves the problem $\mathcal{GN}$ in Part 3f above to obtain a direction $\bar{d}$ for which $\|F(x) + F'(x)\bar{d}\| < \|F(x)\|$. Show that the following backtracking line search procedure is finitely terminating in the sense that the solution $\bar{t}$ is not zero.

Line Search: Let $0 < c < 1$ and $0 < \gamma < 1$ and set

$$\begin{array}{rll} \bar{t} := & \max & \gamma^s \\ & \text{s.t.} & s \in \{0, 1, 2, 3, ...\} \text{ and} \\ & & \|F(x + \gamma^s \bar{d})\| \le (1 - \gamma^s c)\|F(x)\| + c\gamma^s \|F(x) + F'(x)\bar{d}\|. \end{array}$$

Hint: $(1 - \gamma^s c)\|F(x)\| + c\gamma^s\|F(x) + F'(x)\bar{d}\| = \|F(x)\| + c\gamma^s[\|F(x) + F'(x)\bar{d}\| - \|F(x)\|]$. Then use Part 3f.