EE/M 520, Spring 2005
Exam 1: Due start of Session 14

Solutions (version: March 8, 2005, 11:7) 75 mins.; Total 50 pts.

1. (14 pts.) Consider the problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega, \end{array}$$

where $\Omega = \{\boldsymbol{x} \in \mathbb{R}^2 : x_1^2 + x_2^2 \geq 1\}$ and $f(\boldsymbol{x}) = x_2$.

a. Find all point(s) satisfying the FONC.

b. Which of the point(s) in part a satisfy the SONC?

c. Which of the point(s) in part a are local minimizers? Explain fully.

Ans.: a. We have $\nabla f(\boldsymbol{x}) = [0, 1]$, which is nonzero everywhere. Hence, no interior point satisfies the FONC. Moreover, any boundary point with a feasible direction $\boldsymbol{d}$ such that $d_2 < 0$ cannot be satisfy the FONC, because for such a $\boldsymbol{d}$, $\boldsymbol{d}^T \nabla f(\boldsymbol{x}) = d_2 < 0$. By drawing a picture, it is easy to see that the only boundary point remaining is $\boldsymbol{x}^* = [0, 1]^T$. For this point, any feasible direction satisfies $d_2 \geq 0$. Hence, for any feasible direction, $\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) = d_2 \geq 0$. Hence, $\boldsymbol{x}^* = [0, 1]^T$ satisfies the FONC, and is the only such point.

b. We have $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{O}$. So any point (and in particular $\boldsymbol{x}^* = [0, 1]^T$) satisfies the SONC.

c. The point $\boldsymbol{x}^* = [0, 1]^T$ is not a local minimizer. To see this, consider points of the form $\boldsymbol{x} = [\sqrt{1 - x_2^2}, x_2]^T$ where $x_2 \in [1/2, 1)$. It is clear that such points are feasible, and are arbitrarily close to $\boldsymbol{x}^* = [0, 1]^T$. However, for such points, $f(\boldsymbol{x}) = x_2 < 1 = f(\boldsymbol{x}^*)$.

2. (13 pts.) Consider the problem of minimizing $f : \mathbb{R} \to \mathbb{R}$, $f \in \mathcal{C}^3$, over the constraint set $\Omega = [0, 1]$. Suppose that $x^* = 0$ is a local minimizer.

a. By the FONC, we know that $f'(0) \geq 0$ (where f' is the first derivative of f). By the SONC, we know that if $f'(0) = 0$, then $f''(0) \geq 0$ (where f'' is the second derivative of f). State and prove a *third-order necessary condition (TONC)* involving the third derivative at 0, $f'''(0)$.

b. Give an example of f such that the FONC, SONC, and TONC (in part a) holds at the point 0, but 0 is not a local minimizer of f over $\Omega = [0, 1]$. (Show that your example is correct.)

Ans.: a. The TONC is: if $f'(0) = 0$ and $f''(0) = 0$, then $f'''(0) \geq 0$. To prove this, suppose $f'(0) = 0$ and $f''(0) = 0$. By Taylor's theorem, for $x \geq 0$,

$$f(x) = f(0) + \frac{x^3}{3!} f'''(0) + o(x^3).$$

Since 0 is a local minimizer, $f(x) \geq f(0)$ for sufficiently small $x \geq 0$. Hence, for all $x \geq 0$ sufficiently small,

$$f'''(0) \geq 3! \frac{o(x^3)}{x^3}.$$

This implies that $f'''(0) \geq 0$, as required.

b. Let $f(x) = -x^4$. Then, $f'(0) = 0$, $f''(0) = 0$, and $f'''(0) = 0$, which means that the FONC, SONC, and TONC are all satisfied. However, 0 is not a local minimizer: $f(x) < 0$ for all $x > 0$.

3. (10 pts.) Suppose we apply the steepest descent algorithm $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha_k \boldsymbol{g}^{(k)}$ to a quadratic function f with Hessian $\boldsymbol{Q} > 0$. Let $\lambda_{\max}$ and $\lambda_{\min}$ be the largest and smallest eigenvalue of $\boldsymbol{Q}$, respectively. Which of the following two inequalities are possibly true? (When we say here that an inequality is "possibly" true, we mean that there exists a choice of f and $\boldsymbol{x}^{(0)}$ such that the inequality holds.) Explain your answer fully.

a. $\alpha_0 \geq 2/\lambda_{\max}$

b. $\alpha_0 > 1/\lambda_{\min}$

Ans.: a. Possible. Pick f such that $\lambda_{\max} \geq 2\lambda_{\min}$ and $\boldsymbol{x}^{(0)}$ such that $\boldsymbol{g}^{(0)}$ is an eigenvector of $\boldsymbol{Q}$ with eigenvalue $\lambda_{\min}$. Then,

$$\alpha_0 = \frac{\boldsymbol{g}^{(0)T}\boldsymbol{g}^{(0)}}{\boldsymbol{g}^{(0)T}\boldsymbol{Q}\boldsymbol{g}^{(0)}} = \frac{1}{\lambda_{\min}} \geq \frac{2}{\lambda_{\max}}.$$

b. Not possible. Indeed, using Rayleigh's inequality,

$$\alpha_0 = \frac{\boldsymbol{g}^{(0)T}\boldsymbol{g}^{(0)}}{\boldsymbol{g}^{(0)T}\boldsymbol{Q}\boldsymbol{g}^{(0)}} \leq \frac{1}{\lambda_{\min}}.$$

4. (14 pts.)

a. Consider a function $f : \mathbb{R} \to \mathbb{R}$, $f \in \mathcal{C}^2$, such that x^* is a local minimizer and $f''(x^*) \neq 0$. Suppose we apply the algorithm $x_{k+1} = x_k - \alpha_k f'(x_k)$ such that $\{\alpha_k\}$ is a positive step-size sequence that converges to $1/f''(x^*)$. Show that if $x_k \to x^*$, then the order of convergence of the algorithm is *superlinear* (i.e., strictly greater than 1).

b. Given part a, what can you say about the order of convergence of the secant algorithm?

Ans.: a. We have

$$|x_{k+1} - x^*| = |x_k - x^* - \alpha_k f'(x_k)|.$$

By Taylor's theorem applied to f',

$$f'(x_k) = f'(x^*) + f''(x^*)(x_k - x^*) + o(|x_k - x^*|).$$

Since $f'(x^*) = 0$ by the FONC, we get

$$\begin{aligned} x_k - x^* - \alpha_k f'(x_k) &= (1 - \alpha_k f''(x^*))(x_k - x^*) + \alpha_k o(|x_k - x^*|) \\ &= o(|x_k - x^*|) + \alpha_k o(|x_k - x^*|) \\ &= (1 + \alpha_k) o(|x_k - x^*|). \end{aligned}$$

Because $\{\alpha_k\}$ converges, it is bounded, and so $(1 + \alpha_k)o(|x_k - x^*|) = o(|x_k - x^*|)$. Combining the above with the first equation, we get

$$|x_{k+1} - x^*| = o(|x_k - x^*|),$$

which implies that the order of convergence is superlinear.

b. In the secant algorithm, if $x_k \to x^*$, then $(f'(x_k) - f'(x_{k-1}))/(x_k - x_{k-1}) \to f''(x^*)$. Since the secant algorithm has the form $x_{k+1} = x_k - \alpha_k f'(x_k)$ with $\alpha_k = (x_k - x_{k-1})/(f'(x_k) - f'(x_{k-1}))$, we deduce that $\alpha_k \to 1/f''(x^*)$. Hence, if we apply the secant algorithm to a function $f \in \mathcal{C}^2$, and it converges to a local minimizer x^* such that $f''(x^*) \neq 0$, then the order of convergence is superlinear.