EE/M 520, Spring 2006
Exam 1: Session 18

Solutions (version: March 27, 2006, 20:28) 75 mins.; Total 50 pts.

1. (20 pts.) Consider the problem of minimizing $f : \mathbb{R} \to \mathbb{R}$, $f \in \mathcal{C}^3$, over the constraint set Ω. Suppose that 0 is an *interior* point of Ω.

a. Suppose that 0 is a local minimizer. By the FONC, we know that $f'(0) = 0$ (where f' is the first derivative of f). By the SONC, we know that $f''(0) \geq 0$ (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 interior point 0, but 0 is not a local minimizer of f over Ω. (Show that your example is correct.)

c. Suppose f is a third-order polynomial. If 0 satisfies the FONC, SONC, and TONC (in part a), then is this *sufficient* for 0 to be a local minimizer?

Ans.: a. The TONC is: if $f''(0) = 0$, then $f'''(0) = 0$. To prove this, suppose $f''(0) = 0$. Now, by the FONC, we also have $f'(0) = 0$. Hence, by Taylor's theorem,

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

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

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

Now, if $x > 0$, then

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

which implies that $f'''(0) \geq 0$. On the other hand, if $x < 0$, then

$$f'''(0) \leq 3!\frac{o(x^3)}{x^3},$$

which implies that $f'''(0) \leq 0$. This implies that $f'''(0) = 0$, 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 $x \neq 0$.

c. The answer is yes. To see this, we first write

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f'''(0)}{3!}x^3.$$

Now, if the FONC is satisfied, then

$$f(x) = f(0) + \frac{f''(0)}{2}x^2 + \frac{f'''(0)}{3!}x^3.$$

Moreover, if the SONC is satisfied, then either (i) $f''(0) > 0$ or (ii) $f''(0) = 0$. In the case (i), it is clear from the above equation that $f(x) \geq f(0)$ for all x sufficiently close to 0 (because the third term on the right-hand side is $o(x^2)$). In the case (ii), the TONC implies that $f(x) = f(0)$ for all x. In either case, $f(x) \geq f(0)$ for all x sufficiently close to 0. This shows that 0 is a local minimizer.

2. (12 pts.) Consider a fixed step-size gradient algorithm applied the each of the functions $f : \mathbb{R}^2 \to \mathbb{R}$ in parts a and b below. In each case, find the largest range of values of the step-size α for which the algorithm is globally convergent.

a. $f(\boldsymbol{x}) = 1 + 2x_1 + 3(x_1^2 + x_2^2) + 4x_1x_2$.

b. $f(\boldsymbol{x}) = \boldsymbol{x}^T \begin{bmatrix} 3 & 3 \\ 1 & 3 \end{bmatrix} \boldsymbol{x} + [16, 23]\boldsymbol{x} + \pi^2$.

Ans.: In both cases, we compute the Hessian $\boldsymbol{Q}$ of f, and find its largest eigenvalue $\lambda_{\max}$. Then the range we seek is $0 < \alpha < 2/\lambda_{\max}$.

a. In this case,

$$\boldsymbol{Q} = \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix},$$

with eigenvalues 2 and 10. Hence, the answer is $0 < \alpha < 1/5$.

b. In this case, again we have

$$\boldsymbol{Q} = \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix},$$

with eigenvalues 2 and 10. Hence, the answer is $0 < \alpha < 1/5$.

3. (6 pts.) Is it true that if we initialize Newton's method sufficiently close to a local *maximizer*, then it converges to that maximizer? Explain your answer fully and clearly.

Ans.: Yes, provided $f \in \mathcal{C}^3$ and $\boldsymbol{F}(\boldsymbol{x}^*)$ is invertible, where $\boldsymbol{x}^*$ is the local maximizer. Why? Because the case of a maximizer is covered by the theorem on the convergence of Newton's method (where the assumptions are as given here, and $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$, which is true by the FONC.

4. (12 pts.) Consider the following algorithm:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \boldsymbol{H}\boldsymbol{g}^{(k)},$$

where, as usual, $\boldsymbol{g}^{(k)} = \nabla f(\boldsymbol{x}^{(k)})$ and $\boldsymbol{H}$ is a fixed symmetric matrix.

a. Suppose that $f \in \mathcal{C}^3$ and there is a point $\boldsymbol{x}^*$ such that $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$ and $\boldsymbol{F}(\boldsymbol{x}^*)^{-1}$ exists. Find $\boldsymbol{H}$ such that if $\boldsymbol{x}^{(0)}$ is sufficiently close to $\boldsymbol{x}^*$, then $\boldsymbol{x}^{(k)}$ converges to $\boldsymbol{x}^*$ with order of convergence of at least 2.

b. With the setting of $\boldsymbol{H}$ in part a, is the given algorithm a quasi-Newton method? Explain your answer fully and clearly.

Ans.: a. The appropriate choice is $\boldsymbol{H} = \boldsymbol{F}(\boldsymbol{x}^*)^{-1}$. To show this, we can apply the same argument as in the proof of the theorem on the convergence of Newton's method. (We won't repeat it here.)

b. Yes (provided we incorporate the usual step size). Indeed, if we apply the algorithm with the choice of $\boldsymbol{H}$ in part a, then when applied to a quadratic with Hessian $\boldsymbol{Q}$, the algorithm uses $\boldsymbol{H} = \boldsymbol{Q}^{-1}$, which definitely satisfies the quasi-Newton condition. In fact, the algorithm then behaves just like Newton's algorithm.