CS726, Fall 2006
Final Examination
Tuesday, December 19, 2006, 10:05am-12:05pm

Answer all FOUR questions below. One handwritten sheet of notes (written front and back) is allowed. Explain all your answers.

1. The DFP quasi-Newton updating formula for the approximate Hessian B_k can be written as follows:

$$B_{k+1} = (I - \rho_k y_k s_k^T) B_k (I - \rho_k s_k y_k^T) + \rho_k y_k y_k^T,$$

where

$$\rho_k = \frac{1}{y_k^T s_k}.$$

Show that if B_k is positive definite and the curvature condition $y_k^T s_k > 0$ holds, then B_{k+1} is also positive definite.

2. (a) Consider the problem

$$\min_x f(x) \ \text{ subject to } \ l \le x \le u,$$

where $x \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, and the lower- and upper-bound vectors l and u are also in $\mathbb{R}^n$. Write down the first-order necessary (KKT) conditions for this problem, eliminating Lagrange multipliers to obtain a simplified form.

(b) Consider the problem

$$\begin{aligned} \min_{(x_1,x_2)\in\mathbb{R}^2} \ & \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 - x_1x_2 - 2x_2 \\ \text{subject to } \ & 1 - x_1 - x_2 \ge 0, \\ & 1 + x_1 - x_2 \ge 0. \end{aligned}$$

Show that the KKT conditions are satisfied at $x^* = (0,1)$, and determine the optimal values of the Lagrange multipliers for the two constraints.

3. Consider the determination of a quadratic function of two variables using function value information. That is, we seek the values of the scalars a_{11}, a_{12}, a_{22}, b_1, b_2, and c such that for the model function $m(x)$ defined by

$$m(x) = \frac{1}{2}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + c,$$

we have $m(y^i) = f(y^i)$, for the chosen values y^i, $i = 1,2,3,4,5,6$ and the given function f. Show that if the points y^i all lie on an ellipse satisfying

$$(y^i)^T W y^i = \gamma, \quad i = 1,2,3,4,5,6,$$

where W is a 2×2 positive definite matrix and γ is a positive scalar, then the quadratic m is not in general determined by the six points y^i, $i = 1,2,3,4,5,6$.

4. (a) Suppose that an algorithm for minimizing the continuously differentiable function f generates a sequence $\{x_k\}$ lying in a bounded set $\mathcal{B}$, such that the sequence of gradient norms $\{\|\nabla f(x_k)\|\}$ has an accumulation point at zero. Show that there exists an accumulation point x_∞ of $\{x_k\}$ such that $\nabla f(x_\infty) = 0$.

   (b) Suppose that an algorithm for minimizing the twice continuously differentiable function f generates a sequence $\{x_k\}$ for which

$$\lim_{k \to \infty} \nabla f(x_k) = 0$$

   and

$$\nabla^2 f(x_k) \text{ are positive definite for all } k.$$

   Show that all accumulation points of $\{x_k\}$ satisfy second-order necessary conditions to be a minimizer of f.

   (c) Can we claim that all accumulation points for the sequence in part (b) satisfy second-order *sufficient* conditions? Why or why not?