CS726, Fall 2013
Final Examination
Friday, December 20, 2013, 12:25p–2:25p

Answer ALL questions below. One handwritten sheet of notes (written front and back) is allowed. EXPLAIN ALL YOUR ANSWERS.

1. Given a line-search function defined by $\phi(t) = 1 - t + \frac{\gamma}{2}t^2$ for some $\gamma > 0$ and scalar $t \geq 0$, consider the Wolfe conditions:

$$\phi(t) \leq \phi(0) + c_1 t \phi'(0) \tag{0.1}$$
$$\phi'(t) \geq c_2 \phi'(0), \tag{0.2}$$

for given parameters c_1 and c_2 in the interval $[0, 1]$.
   (a) Find the subinterval of t values in $[0, \infty)$ for which the first condition (0.1) is satisfied.
   (b) Find the subinterval of t values in $[0, \infty)$ for which the second condition (0.2) is satisfied.
   (c) Find a condition on c_1 and c_2 that ensures that the conditions (0.1) and (0.2) are compatible, that is, the intervals found in parts (a) and (b) have a nonempty intersection. If $c_1 < c_2$, is your condition satisfied?
2. The SRI update formula for an approximate *inverse* Hessian H, given vectors $s := x_{k+1} - x_k$ and $y := \nabla f(x_{k+1}) - \nabla f(x_k)$, is

$$H_+ = H + \frac{(s - Hy)(s - Hy)^T}{(s - Hy)^T y}.$$

   (a) The update formula is not well defined when $s = Hy$. How should you define the updated matrix H_+ in this case, and why?
   (b) The update formula is also not well defined when $s \neq Hy$ but $y^T(s - Hy) = 0$. Show that in this case there does not exist a rank-one update formula (of the form $H_+ = H + aa^T$) such that the secant condition $s = H_+ y$ is satisfied.
   (c) Using the Sherman-Morrison formula, derive an update formula for the SR1 approximation B to the Hessian (not the inverse Hessian). The Sherman-Morrison formula for a nonsingular matrix A and vectors a and b is as follows:

$$(A + ab^T)^{-1} = A^{-1} - \frac{A^{-1}ab^TA^{-1}}{1 + b^TA^{-1}a}.$$

3. (a) Let Ω be a closed convex set and let $P(\cdot)$ denote Euclidean projection onto Ω, that is

$$P(y) := \arg\min_{s \in \Omega} \|s - y\|_2^2.$$

   Show that

$$[y - P(y)]^T [z - P(y)] \leq 0 \text{ for all } z \in \Omega.$$

   (Hint: Note that since $z \in \Omega$ and $P(y) \in \Omega$, we have that $P(y) + \alpha(z - P(y)) \in \Omega$ for all $\alpha \in [0, 1]$.)

(b) Consider the problem

$$\begin{aligned} \min_{(x_1,x_2)\in\mathbb{R}^2} \quad & \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 - x_1x_2 - 3x_2 \\ \text{subject to} \quad & -x_1 - x_2 \geq -1, \\ & x_1 - x_2 \geq -1. \end{aligned}$$

Show that the KKT conditions are satisfied at $x^* = (0, 1)$, and determining the optimal values of the nonnegative multipliers λ_1 and λ_2 for the two constraints.

4. 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 a circle, that is,

$$\|y^i - h\|_2^2 = \gamma, \quad i = 1, 2, 3, 4, 5, 6,$$

for some $h \in \mathbb{R}^2$ and some $\gamma > 0$, then the quadratic m is not well determined by the six points y^i, $i = 1, 2, 3, 4, 5, 6$.