CS726, Fall 2013
Midterm Examination
Monday, October 28, 2013, 7:00pm–9:00pm.

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

1. Consider the unconstrained problem $\min_x f(x)$, where $f : \Re^n \to \Re$ is smooth.
   (a) One form of the Barzilai-Borwein method takes steps of the form $x_{k+1} = x_k - \gamma_k \nabla f(x_k)$, where
   $$\gamma_k := \frac{s_k^T s_k}{s_k^T y_k}, \quad s_k := x_k - x_{k-1}, \;\; y_k := \nabla f(x_k) - \nabla f(x_{k-1}).$$
   Write down an explicit formula for γ_k in terms of s_k and A, for the special case in which f is strictly convex quadratic, that is, $f(x) = (1/2)x^T Ax$, where A is symmetric positive definite.
   (b) Considering the steepest descent method $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$, applied to the strictly convex quadratic, write down an explicit formula for the exact minimizing α_k.
   (c) Show that the steplengths obtained in parts (a) and (b) are related as follows: $\gamma_{k+1} = \alpha_k$.
2. Suppose that $f : \Re^n \to \Re$ be a twice continuously differentiable function and suppose that $\{x_k\}$ is a sequence of iterates in $\mathbf{R}^n$.
   (a) Suppose that $\liminf \|\nabla f(x_k)\| = 0$. Is it true that all accumulation points of $\{x_k\}$ are stationary (that is, satisfy first-order necessary conditions)?
   (b) Suppose that $\lim \nabla f(x_k) = 0$. Is it true that all accumulation points of $\{x_k\}$ are stationary?
   (c) Suppose that the sequence $\{x_k\}$ converges to a point x^*, that the gradients $\nabla f(x_k)$ converge to zero, and that the Hessians $\nabla^2 f(x_k)$ at all these points are positive definite. Show that second-order necessary conditions are satisfied at the limit x^*.
   (d) For the situation described in part (c), can we say that second-order *sufficient* conditions will be satisfied at x^*? Explain.

   (Hint: The eigenvalues of a matrix depend continuously on the elements of the matrix.)
3. (a) The BFGS quasi-Newton updating formula for the approximate inverse Hessian H_k can be written as follows:
   $$H_{k+1} = (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T,$$
   where
   $$\rho_k = \frac{1}{y_k^T s_k}.$$
   Show that if H_k is positive definite and the curvature condition $y_k^T s_k > 0$ holds, then H_{k+1} is also positive definite.
   (b) If $y_k^T s_k \leq 0$, is it still possible for H_{k+1} to be positive definite?

4. Consider the following form of the conjugate gradient method for solving $Ax = b$ (or, equivalently, minimizing $f(x) = (1/2)x^T Ax - b^T x$, where A is symmetric positive definite.

Given x_0;
Set $r_0 \leftarrow Ax_0 - b$, $p_0 \leftarrow -r_0$, $k \leftarrow 0$;
while $r_k \neq 0$

$$\begin{aligned}
\alpha_k &\leftarrow -\frac{r_k^T p_k}{p_k^T A p_k}; \\
x_{k+1} &\leftarrow x_k + \alpha_k p_k; \\
r_{k+1} &\leftarrow A x_{k+1} - b; \\
\beta_{k+1} &\leftarrow \frac{r_{k+1}^T A p_k}{p_k^T A p_k}; \\
p_{k+1} &\leftarrow -r_{k+1} + \beta_{k+1} p_k; \\
k &\leftarrow k + 1;
\end{aligned}$$

end (while)

Show that

$$r_k^T p_j = 0, \quad \text{for all } j = 0, 1, \ldots, k-1. \tag{0.1}$$

You may assume that the vectors p_j are conjugate, that is, $p_j^T A p_i = 0$ when $i \neq j$.
(Hint: Prove by induction. Show first that $r_{k+1}^T p_k = 0$ for all k, which establishes (0.1) for $k = 1$. Then show that if (0.1) holds for some k, it continues to hold for $k + 1$, that is, $r_{k+1}^T p_j = 0$ for all $j = 0, 1, \ldots, k$.)