CS726, Fall 2007
Final Examination
Tuesday, December 21, 2007, 5:05pm–7:05pm.

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

1. Define the function f to be the following strictly convex quadratic:

$$f(x) = \frac{1}{2}x^T Ax + b^T x + c,$$

where $x \in \mathbf{R}^n$ and A is an $n \times n$ positive definite matrix.

(a) Find an explicit formula for the exact minimizing α of the function

$$t(\alpha) \stackrel{\text{def}}{=} f(x + \alpha p),$$

where x and p are vectors such that $p \neq 0$ and x is not a minimizer of f.

Solution (5 points):

$$t(\alpha) = \frac{1}{2}\alpha^2 p^T Ap + \alpha p^T (Ax + b) + (\frac{1}{2}x^T Ax + b^T x + c),$$

so

$$t'(\alpha) = \alpha p^T Ap + p^T (Ax + b)$$

so the minimizing α satisfies

$$\alpha = -\frac{p^T (Ax + b)}{p^T Ap}.$$

(Clearly $t''(\alpha) = p^T Ap > 0$ by positive definiteness of A.)

(b) For what values of c_1 is the first Wolfe condition satisfied by the minimizing α from part (a)? (The first Wolfe condition is that $f(x + \alpha p) \leq f(x) + c_1 \alpha \nabla f(x)^T p$.)

Solution (5 points): Letting α^* be the minimizing α, we have

$$f(x + \alpha^* p) - f(x) = -\tfrac{1}{2}\frac{(p^T (Ax + b))^2}{p^T Ap},$$

whereas

$$c_1 \alpha^* \nabla f(x)^T p = -c_1 \frac{(p^T (Ax + b))^2}{p^T Ap}.$$

Hence, the Wolfe condition is satisfied if $c_1 \leq 1/2$.

2. Let $\{x_k\}$ be a sequence of vectors in $\mathbf{R}^n$ and let f be a twice continuously differentiable function.

   (a) If $\{\nabla f(x_k)\}$ has an accumulation point at 0, does it follow that the sequence $\{x_k\}$ must have a stationary accumulation point?

   Solution (5 points): No. The sequence $\{x_k\}$ may not be bounded, so may not have an accumulation point.

   (b) Suppose that $\lim_{k\to\infty} x_k = x^*$ for some x^*, that $\lim_{k\to\infty} \nabla f(x_k) = 0$, and that there is a constant $\beta > 0$ such that matrices $\nabla^2 f(x_k)$ are positive definite with

$$\|\nabla^2 f(x_k)\| \|(\nabla^2 f(x_k))^{-1}\| \leq \beta, \quad \text{for all } k > 0.$$

   Are the second-order sufficient conditions for x^* to be a local minimizer of f satisfied at x^*?

Solution (5 points): No. Clearly by the smoothness conditions on f, we have that $\nabla f(x^*) = \lim_{k\to\infty} \nabla f(x_k) = 0$, and

$$\nabla^2 f(x^*) = \lim_{k\to\infty} \nabla^2 f(x_k).$$

Since all $\nabla^2 f(x_k)$ are positive definite, it follows that $\nabla^2 f(x^*)$ is at least positive semidefinite. However, the limit does not have to be positive definite. Consider the case of $f(x) = x^3$ and $x_k = k^{-1}$, for which the condition numbers of the Hessians are all 1.

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.

Solution (6 points): Note that the curvature condition implies that $\rho_k > 0$. Defining $E_k = I - \rho_k y_k s_k^T$, we have for any $z \in \mathbf{R}^n$ that

$$z^T H_{k+1} z = z^T E_k^T H_k E_k z + \rho_k (s_k^T z)^2.$$

Both terms on the right-hand side are nonnegative, by positive definiteness of H_k, so that H_{k+1} is at least positive semidefinite.
If $z^T H_{k+1} z = 0$, we must have $s_k^T z$, which implies that $E_k z = z$. Hence,

$$0 = z^T H_{k+1} z = z^T H_k z,$$

which by positive definiteness of H_k implies that $z = 0$. Hence $z^T H_{k+1} z = 0 \Rightarrow z = 0$, so H_{k+1} is positive definite.

(b) If $y_k^T s_k \leq 0$, is it still possible for H_{k+1} to be positive definite?

Solution (4 points): No. Setting $z = y_k$ in the discussion above, we have $E_k z = 0$ and therefore

$$z^T H_{k+1} z = \rho_k (s_k^T y_k)^2 = s_k^T y_k \leq 0,$$

so H_{k+1} has at least one nonpositive eigenvalue.

4. (a) Consider the function $r : \mathbf{R} \to \mathbf{R}$ defined by $r(x) = x^q$, where q is an integer greater than 2. (Note that $x^* = 0$ is the sole root of this function and that it is degenerate, that is, $r'(x^*)$ is singular.) Show that Newton's method converges Q-linearly, and find the value of the convergence ratio (the limiting bound on $\|x_{k+1} - x^*\|/\|x_k - x^*\|$).

Solution (4 points):

$$x_{k+1} = x_k - \frac{r(x_k)}{r'(x_k)} = x_k - \frac{x_k^q}{q x_k^{q-1}} = x_k \left(1 - \frac{1}{q}\right).$$

Hence defining $e_k = x_k - x^* = x_k$, we have

$$\frac{\|e_{k+1}\|}{\|e_k\|} = 1 - \frac{1}{q} = \frac{q-1}{q}.$$

(b) Show that Newton's method applied to the function $r(x) = -x^5 + x^3 + 4x$ starting from $x_0 = 1$ generates a sequence of iterates that alternates between $+1$ and -1.

Solution (3 points): We have

$$x_{k+1} = x_k - \frac{-x_k^5 + x_k^3 + 4x_k}{-5x_k^4 + 3x_k^2 + 4}.$$

Hence, when $x_k = 1$, we have

$$x_{k+1} = 1 - \frac{-1 + 1 + 4}{-5 + 3 + 4} = -1,$$

while when $x_k = -1$, we have

$$x_{k=1} = -1 - \frac{1 - 1 - 4}{-5 + 3 + 4} = -1 + 2 = 1.$$

(c) Find the roots of the function in (b), and check that they are nondegenerate.

Solution (3 points): We have $r(x) = x(-x^4 + x^2 + 4)$ which besides the degenerate root at $x = 0$ has roots when $-x^4 + x^2 + 4 = 0$, that is, when

$$x^2 = \frac{-1 \pm \sqrt{1 + 16}}{-2} = \frac{1 \pm \sqrt{17}}{2}.$$

Hence the other roots are at

$$x = \pm \sqrt{\frac{1 + \sqrt{17}}{2}}.$$

To verify nondegeneracy we need to check that $r'(x) \neq 0$, that is,

$$0 \neq -5x^4 + 3x^2 + 4 = -5 \left(\frac{1 + \sqrt{17}}{2}\right)^2 + 3\frac{1 + \sqrt{17}}{2} + 4 = -17 - \sqrt{17}.$$