MATH 4/5420 - Optimization
Sample Midterm Exam - Spring 2017

1. (a) Show that the rank-one matrix aa^T, where $a \in \mathbb{R}^n$, is symmetric and positive semidefinite.

   (b) Given a *nonsymmetric* matrix B, show that the gradient and Hessian for the function $g(x) = x^T Bx$ are:

   $$\nabla g(x) = (B^T + B)x \text{ and } \nabla^2 g(x) = B^T + B.$$

   [You may consider an example first, say $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$].

2. Consider the unconstrained minimization problem for the quadratic function

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

   where A is positive definite symmetric $n \times n$ matrix.

   (a) Write down the gradient $\nabla f(x)$ and Hessian $\nabla^2 f(x)$

   (b) Determine the minimizer x^* for f and explain whether it is global or not.

   (b) Show that for any starting point x^0, the Newton's iteration computes the minimizer x^* in one step.

3. Consider the steepest descent method with exact line search for the quadratic function

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

   (with A is positive definite):

   $$x^{k+1} = x^k + \alpha_k p^k, \quad p^k = -\nabla f(x^k), \alpha_k = \arg\min_{\alpha > 0} f(x^k + \alpha p^k)$$

   (a) Show that

   $$\alpha_k = -\frac{\nabla f(x^k) p^k}{(p^k)^T A p^k}$$

   (b) Show that $x^{k+1} - x^k$ is orthogonal to $x^k - x^{k-1}$.

4. (a) Sketch contour lines for the function

   $$f(x_1, x_2) = x_1^2 + 4x_2^2 - 4x_1 - 8x_2$$

   and determine $\mathbf{x}^*$ which minimizes f.

   (b) Apply the BFGS algorithm (page 140 in Nocedal & Wright) choosing $\mathbf{x}^0 = (0,0)^T$ and $H^0 = I$, and show that the line $\mathbf{x} = \mathbf{x}^1 + \alpha \mathbf{p}^1$ passes through $\mathbf{x}^*$, thus guaranteeing termination in two steps.

   (c) Show that if the method of steepest descent is started at $\mathbf{x}^0 = (0,0)^T$ it cannot converge to $\mathbf{x}^*$ in a finite number of steps. Are there any values of $\mathbf{x}^0$ for which there will be convergence in a finite number of steps?

5. Consider the following method for unconstraint optimization, in which an exact line search is used, and the search direction is calculated from

$$Bp^k = -\nabla f(x^k)$$

where B is a fixed matrix. Apply this method to the problem

$$\min \frac{1}{2}\|\mathbf{x}\|^2$$

from the initial point $\mathbf{x}^0 = (a, ta)^T$, $(a > 0, t > 0)$ with $B = \begin{bmatrix} 1 & 0 \\ 0 & b \end{bmatrix}$, $(0 < b < 1$. Find the values of t such that $|x_2^k| = t|x_1^k|$ for all k. What kind of convergence does this example exhibit?

6. One can solve the system of equations

$$r_1(x) = x_1^3 - x_2 - 1 = 0,$$
$$r_2(x) = x_1^2 - x_2 = 0$$

by reformulating it as a nonlinear least square problem

$$\min_x [r_1(x)^2 + r_2(x)^2]$$

(a) Show that $f(x) = \|r(x)\|^2$ has a global minimum at $x = (1.46557, 2.14790)^T$ which solves the above system, and a local minimizer at $x = (0, -\frac{1}{2})^T$, which does not. Show that there is also a saddle point at $x = (\frac{2}{3}, -\frac{7}{54})^T$.

(b) Show that the Gauss-Newton method converges rapidly from the initial point $x^0 = (1.5, 2.25)^T$ and verify the order of convergence is quadratic. If x^0 is chosen close to the local minimizer, observe that the Gauss-Newton method behaves wildly, but ultimately converges to the global solution.

7. The linear conjugate gradient (CG) method (Algotirhm 5.1 or 5.2 in Nocedal& Wright) is designed to solve $Ax = b$ when A is symmetric and positive definite matrix. This CG method can be used to solve a general system $Ax = b$, when A is not necessarily symmetric (or even square), by solving instead (via the CG method) the normal equation

$$A^T A x = A^T b$$

(a) Write an implementation of the CG method for this normal equation. [You will get bonus points if you write the code such that the matrix $A^T A$ is never precomputed. For example, to compute $A^T Ap$, multiply first A to p to get Ap then multiply the result to A^T] This is important for large matrices.]

(b) Apply this procedure to the (overdetermined) 4×3 system $Ax = b$, where

$$A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & 5 \\ 0 & 4 & -1 \\ 3 & 1 & 1 \end{bmatrix}, \qquad , b = \begin{bmatrix} 0 \\ 3 \\ -5 \\ 6 \end{bmatrix}$$

(c) [Math 5420 only!] What precautions need to be put in place so the same procedure be used for an underdetermined system $Ax = b$, where A is $m \times n$ with $m < n$, which in this case has infinitely many solutions? Can you tell if the *CG* method converges to a solution? Which one? Does it converge in finitely many steps? [You can test this on a concrete example, say 2×3]