MATH 4/5420 - Optimization
Midterm Exam - Spring 2017

In Class Portion

1. (a) Given the linear function $f(x) = b^T x$, where $b \in \mathbb{R}^n$, compute the gradient $\bigtriangledown f$, using the directional derivative calculation. [Avoid the use of coordinates in your proof]

   (b) Given the quadratic function $f(x) = x^T Ax$, where A is a symmetric matrix, positive definite matrix, show that $x^* = 0$ is the unique global minimizer for f.

2. Using Taylor's Theorem

$$f(x+p) = f(x) + \bigtriangledown f(x)^T p + \frac{1}{2} p^T \bigtriangledown^2 f(x+tp)p, \text{ for some } t \in (0,1)$$

   prove that if x^* is such that $\bigtriangledown f(x^*) = 0$, then: [Choose only one of the two]
   (a) [Necessary condition] If x^* is a (local) minimizer, then $\bigtriangledown^2 f(x^*)$ is positive semidefinite.

   (b) [Sufficient condition] If $\bigtriangledown^2 f(x^*)$ is positive definite, then x^* is a (local) minimizer.

   (c) What can happen in (b) if the Hessian is not positive definite but only positive semi-definite? Give an example in 1-dim and 2-dim.

3. Consider the optimization problem

$$\min \|Ax - b\|^2$$

   where A is a (possibly nonsquare) $m \times n$ matrix with $m \geq n$, and $b \in \mathbb{R}^m$.
   (a) Rewrite the objective function as a quadratic function of $x \in \mathbb{R}^n$ and write down (without computing) the gradient and Hessian of this quadratic function.

   (b) Show that if A has maximum rank (all n columns are linearly independent), then the quadratic function is strictly convex and hence the minimization problem has an unique minimizer.

   (c) Write down the steepest descent method with exact line search. Then perform the first iteration, that is, given a starting point x^0, compute x^1.

4. Consider the least-square approximation problem

$$\min_{x_1, x_2} \sum_j |y_j - (x_1 + x_2 t_j)|^2$$

   of finding the line that best fits the data set $\{(t_j, y_j)\} = \{(-1, 1), (0, -1), (1, 0), (2, 2)\}$, which amounts to finding the solution to the minimization problem

$$\min \|Ax - b\|^2$$

   where

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ and } b = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix}$$

   Following the steps in exercise 3, find the minimizer x^* and give a reason for why the steepest descent method converges.

Take Home Portion (Due Wed, March 22, before class)

5. (a) Find (if any), all 'critical' points (where $\bigtriangledown f(x) = 0$) for the function

$$f(x) = 5x_1^2 + 2x_2 - x_1^2 x_2^2$$

. Solve analytically, i.e. do not use an iterative method.

(b) Classify each point as a local min, local max or saddle.

(c) Apply the Newton method with $x^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, to obtain $x^{(k)}$, for $k = 1 \ldots 10$. Does this method descend at each iteration? Does it converge?

6. Suppose in Newton's method at iteration k, the Hessian $\bigtriangledown^2 f(x^{(k)})$ is not invertible. One approach to get around this is to switch to a quasi-Newton iteration

$$x^{(k+1)} = x^{(k)} - \alpha_k H_k \bigtriangledown f(x^{(k)}).$$

where H_k is computed via say the BFGS method from the previous iteration.
Implement this method and apply it to the function $f(x_1, x_2) = x_1^2 + 4 * x_1 * x_2 + 4x_2^2$.

7. Consider the following fixed-step-size modification of the steepest descent method:

$$x^{k+1} = x^k + \alpha p_k, \quad p_k = - \bigtriangledown f_k, \quad \alpha > 0 \text{ (fixed)}$$

.

Implement it for the quadratic function

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

where $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$, and $b = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.
Determine the range of values of $\alpha > 0$ for which the iterations x^k converge to the minimizer x^*. You may pick several initial starting points x^0.

8. Find the solution of the system $Ax = b$

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 3 & 2 & 1 \\ 1 & 2 & 2 & -1 \\ 0 & 1 & -1 & 5 \end{bmatrix}, \quad \text{and} b \begin{bmatrix} 1 \\ 3 \\ 4 \\ 1 \end{bmatrix}$$

by minimizing vis the Conjugate gradient method the quadratic function

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

(b) Should you check that the matrix A is positive definite when applying this method?

(c) Consider the Conjugate Gradient method applied to solving $Ax = 0$ with A is invertible, but not positive definite:

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 3 \\ -1 & 2 & 2 \end{bmatrix} \quad \text{and} b = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Run the standard CG method (can use the FR version of the code) with starting point at $x^0 = [140]^T$ (or pick your own) and show that the CG iterates do not converge to 0 (the unique solution $x^* = 0$ of the system $Ax = 0$), the initial starting point $x^0 = [140]^T$. Finally, modify the CG method by reseting β_k to 0 periodically to see if the convergence is improved.