HOMEWORK
David F. Gleich
February 15, 2012

Homework 4

Please answer the following questions in complete sentences in a typed manuscript and submit the solution to me in class on February 16th, 2012.

Problem 1: Covering the basics.

1. **Decrease alone is insufficient!** Construct a sequence of iterates such that $f(x_{k+1}) < f(x_k)$, but that do not converge to a minimizer of f for a convex function f. (Hint: think $f(x) = x^2$.)

2. **Back to calculus.** Let

$$L_k(\alpha) = f(\mathbf{x}_k + \alpha \mathbf{p}_k)$$

be the line search function at the kth iteration of an optimization algorithm. Use the definition of the directional derivative to show that $L_k'(\alpha) = g(\mathbf{x}_k + \alpha \mathbf{p}_k)^T \mathbf{p}_k$.

Problem 2: Finally! An optimization algorithm

Non-negative least squares is an important variant. Formally, it is:

$$\begin{array}{ll} \text{minimize} & \|\boldsymbol{A}\mathbf{x} - \mathbf{b}\| \\ \text{subject to} & \mathbf{x} \geq 0. \end{array}$$

We'll see this problem again when we study constrained optimization. Here, we'll investigate a log-barrier function to approximate it in an unconstrained manner:

$$\text{minimize} \quad \|\boldsymbol{A}\mathbf{x} - \mathbf{b}\| - \mu \mathbf{e}^T \log(\mathbf{x}),$$

where log is an elementwise function. Use the definition where: $\log(x) = -\infty$ if $x \leq 0$.

1. Determine the gradient and Hessian.

2. What did you learn in class that you should always do after step 1? Do it.

3. Implement a backtracking line search routine that satisfies sufficient decrease. Convince your professor and TA that your implementation does not have any flaws. Discuss any flaws you known.

4. Modify the `gradient_descent_1.m` and `newtons_method_1.m` functions to use your backtracking line search. Read page 59 of Nocedal and Wright and follow the advice.
(Use (3.60) if applicable.)

5. Suppose that $\mathbf{x}^{(0)}$ is strictly positive.
Does $\mathbf{x}^{(k)}$ stay strictly positive with these two algorithms? Discuss or prove.

6. Show plots of the convergence in terms of function values and of infinity norms of the gradients of the methods for the matrices:

```
A = [
    0.0372    0.2869
    0.6861    0.7071
    0.6233    0.6245
    0.6344    0.6170];
b = [
    0.8587
    0.1781
    0.0747
    0.8405];
mu = 0.1;
```

1. Suppose we say that a method converges if the infinity norm of the gradient is less than 10^{-4}. Plot the number of function evaluations of your new steepest descent method and the Newton method for the values of $\mu = \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$.
   Show your solutions for each of these cases. Compare the solutions to Matlab's `fminunc` routine.