The University of Texas at Austin
Department of Electrical and Computer Engineering

EE381V-11: Large Scale Optimization — Fall 2012

Problem Set Four

Caramanis/Sanghavi | Due: Thursday, October 4, 2012.

Matlab and Computational Assignments. Please provide a printout of the Matlab code you wrote to generate the solutions to the problems below.

1. **Conjugate Gradient Algorithm.** Recall the linear conjugate gradient algorithm. Download the file `http://users.ece.utexas.edu/~cmcaram/EE381V_2012F/ConjugateGradient.mat`. There you will find matrices and vectors defining two equations: $M_1 x = b_1$, and $M_2 x = b_2$. The solution, x^*, is there as well, although this is easy to find since both M_1 and M_2 are invertible. Use conjugate gradient to solve these two linear systems, and plot the error, $\log(\|x^{(k)} - x^*\|^2_{M_i})$ vs. iteration k for both.

2. **BFGS for Quadratic Problems.** Now use BFGS for the functions above. Note that for non-quadratic problems, as with Newton's method, BFGS should use an approximate line search. In this quadratic case, however, *use an exact line search.* In addition to solving and plotting the error, show that BFGS generates conjugate directions. For quadratic problems, BFGS is also guaranteed to terminate in at most n steps.

3. **Newton's Method**. This problem will demonstrate the two convergence behaviors of Newton's method, damped and quadratic, by matlab simulation.

   Consider $f_m(x) = \|x\|^3 + \frac{m}{2}\|x\|^2$ for $m \in \{0, 0.0001, 0.001, 0.1\}$ and $x \in \mathbb{R}^5$.

   (a) For each m, implement Newton's method on $f_m(x)$ and provide the convergence plots, i.e $\log(\|x^{(k)} - x^*\|^2)$ vs. iteration k. Use the constant step size $t = 1$.

   (b) Using the condition for quadratic convergence, explain how and why your result changes according to m.

4. **Central Path**. Consider the linear optimization problem:

$$\begin{aligned} \min: \quad & 2x_1 + 4x_2 + x_3 + x_4 \\ \text{s.t.}: \quad & x_1 + 3x_2 + x_4 \leq 4 \\ & 2x_1 + x_2 \leq 3 \\ & x_2 + 4x_3 + x_4 \leq 3 \\ & x_i \geq 0, \quad i = 1, 2, 3, 4. \end{aligned}$$

   (a) Find a function F that is a self-concordant-barrier function, such that the closure of its domain is equal to the feasible set of the problem. (Recall that $-\log(a^\top x - b)$ is a self-concordant-barrier function, as you show in an exercise below.)

(b) Find the analytic center x_F^* using Newton's method. You can initialize at any point in the domain (e.g., $(1/2, 1/2, 1/2, 1/2)$ or any other point you like).

$$x_F^* = \arg \min_{x \in \mathrm{dom} F} F(x).$$

(c) Now you will generate the central path:

$$x^*(t) = \arg \min_{x \in \mathrm{dom} F} f(t; x),$$

where recall:

$$f(t; x) = tc^\top x + F(x).$$

For $t = 0$, the solution, and first point of the central path, is the analytic center. At each iteration, you will compute $t_{k+1} = t_k(1 + \alpha)$. Experiment with different values of α. If α is too small, progress may not be that fast as t will grow slowly. If α is too big, we might move outside the region of quadratic convergence, and although t will grow more quickly, each individual step of the central path will take longer to compute.

(d) Plot the error, $\log(\|x^{(k)} - x^*\|)$ as a function of number of iterations, for different values of α.

Note: Chapter 11 in Boyd & Vandenberghe has much information about central path and barrier methods, although the chapter also contains a lot of information, definitions and ideas we have not yet discussed.

5. Do the same for the (slightly larger) LP contained in `http://users.ece.utexas.edu/~cmcaram/EE381V_2012F/LP_centralpath.mat`. In that file you will find specified: c, A, and b, thus defining the problem:

$$\begin{aligned} \min : \quad & c^\top x \\ \text{s.t.} : \quad & Ax \leq b \\ & x \geq 0. \end{aligned}$$

Note that you can use CVX to quickly solve both this LP and the previous problem, in order to have the solution.

6. **Gradient, Conjugate Gradient, Newton and BFGS**. Consider the function (called the Rosenbrock function)

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

This function is not convex, however it not hard to see that it has a unique minimizer $x^* = (1, 1)$, and that in a neighborhood of this point, the Hessian is positive definite. Initializing at $x_{\text{init}} = (-1.2, 1)$, implement (a) gradient descent, (b) (non-linear) Conjugate Gradient, (c) BFGS, and (d) Newton, using a back-tracking line search for all four. Plot the error in each as a function of the iteration.

Written Problems

1. We proved quadratic convergence of Newton's method (locally), using the assumption that f is strongly convex, smooth, and with L-Lipschitz Hessian. Locally (so, only after BTLS is taking full steps) derive a rate of convergence in the case where the Hessian is α-Hölder continuous. (Recall that a function F is called α-Hölder continuous if $\|F(x) - F(y)\| \leq H\|x - y\|^\alpha$ for all x, y.)

2. **Self Concordant Barriers.** Recall that a self-concordant function f is called a *self-concordant barrier* if in addition to the properties for self-concordance, it satisfies the property that $\lambda_f(x)^2$ is uniformly bounded by some constant ν. That is, f is called a ν-self-concordant-barrier if

$$\lambda_f(x)^2 = \|\nabla f(x)\|^2_{\nabla^2 f(x)^{-1}} \leq \nu, \qquad x \in \mathrm{dom}(f).$$

   (a) Explain why we need this property (in connection to minimizing a linear function over a convex set).

   (b) For A a positive semidefinite matrix, consider the *concave* quadratic $\phi(x) = c + b^\top x - \frac{1}{2}x^\top Ax$. Show that $f(x) = -\ln\phi(x)$ is a ν-self-concordant barrier function, with $\nu = 1$.