MATH 516 SPRING 2007
THIRD HOMEWORK SET

Homework Due Tuesday 05/08/07

1. The dog-leg search direction.

   (a) Given vectors $u, v \in \mathbb{R}^n$ and a scalar $\delta > 0$ satisfying $\|u\|_2 \le \delta$ and $\|v\|_2 > \delta$, write a Matlab script m-file to compute the point of intersection of the line segment connecting u and v and the surface of the δ–ball given by $\delta I\!\!B = \{\delta u \,:\, \|u\|_2 \le 1\}$, that is, find a $\lambda \in [0,1]$ satisfying $\|(1-\lambda)u + \lambda v\|_2 = \delta$. **Caution:** There is only one *correct numerical* procedure to perform this computation. This procedure results from a single closed form expression for the solution. Run the script on the test $u = (1,0)^T,\ v = (4,4)^T,\ \delta = 4$.

   (b) Use the results of part (a) to write a Matlab subroutine that computes the dog-leg search direction.

   (c) Apply the dog-leg search direction with $H_\nu = \nabla^2 f(x^\nu)$ along with the backtracking line search to Rosenbrock's function in the previous homework set. Use the starting point $x = (-2, 2)$ and the stopping criteria outlined in the previous homework set. Choose c and γ appropriately.

2. As you well know, Newton's method for minimizing a function $f : \mathbb{R}^n \to \mathbb{R}$ is just Newton's method applied to the equation $\nabla f(x) = 0$. Thus, the Newton direction ignores the underlying minimization problem. For this reason the Newton direction, if it exists, may not be a direction of descent for the function, or even if it is a direction of descent, it may require an excessive number of steps in a line search procedure to induce descent. Even so, the full Newton step may still be a very good choice of direction even if it may occasionally increase the value of the objective function. For this reason a number of authors have propose *non–monotone* line search procedures in order to increase the likelihood of accepting a Newton (or Newton like) step even if it may occasionally increase the value of the objective. In this problem, we consider one such proposal.

   Let p be some positive integer (usually $p = 3$ or $p = 4$). Suppose that $x^0 \in \mathbb{R}^n$ is given and define $x^{-\ell} = x^0$ for $\ell = 1, 2, \ldots, p-1$. Consider the following backtracking procedure for choosing a step length.

   <u>Non–Monotone Backtracking:</u> $0 < c < 1$ and $0 < \gamma < 1$

   Let $d^k \in \mathbb{R}^n$ be such that $\nabla f(x^k)^T d^k < 0$. Set

$$\begin{array}{rll} \lambda_k = & \max & \gamma^s \\ & \text{subject to} & s = 0, 1, 2, \ldots \\ & & f(x^k + \gamma^s d^k) \le \left[\displaystyle\max_{\ell = 0,1,\ldots,p-1} f(x^{k-\ell}) \right] + c\gamma^s \nabla f(x^k)^T d^k \ . \end{array}$$

   Observe that this line search procedure does not guarantee that $x^{k+1} = x^k + \lambda_k d^k$ satisfies $f(x^{k+1}) < f(x^k)$.

   (a) Show that the sequence $\{f_k^{\max}\}$ given by

$$f_k^{\max} = \max_{\ell = 0,1,\ldots,p-1} f(x^{k-\ell})$$

   for $k = 0, 1, \ldots$ is non-increasing.

   (b) Show by induction that the sequence $\{f_k^{\max}\}$ is p–**step strictly** decreasing sequence, i.e. $f_{k+p}^{\max} < f_k^{\max}$ for $k = 0, 1, \ldots$.

   (c) For each $k = 0, 1, \ldots$, define $m(k)$ to be the smallest integer in the set $\{k, k-1, \ldots, k-p+1\}$ such that

$$f(x^{m(k)}) = \max_{\ell = 0,1,\ldots,p-1} f(x^{k-\ell}),$$

   so that $f_k^{\max} = f(x^{m(k)})$. By paralleling the proof of Theorem 2.1.1, prove the following result.

Theorem Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable with ∇f Lipschitz continuous on the closed convex hull of the set $\{x : f(x) \le f(x^0)\} + \delta I\!B$ where $\delta > 0$ and $x^0 \in \mathbb{R}^n$ are given. Let $\{x^k\}$ be a sequence of iterates given by the following algorithm:

Caution: I haven't proven this result yet, so the statement may need tweaking. Once I get the time to prove it, I'll give you an update.

```
Non-Monotone Descent Algorithm
Step 1: If $D_k \subset \{d : f'(x^k; d) < 0\}$ is empty, STOP; otherwise choose $d^k \in D_k$.
Step 2: Choose the step length $\lambda_k$ by the non-monotone backtracking procedure specified
        above.
Step 3: Set $x^{k+1} = x^k + \lambda_k d^k,\ k = k + 1$, and return to Step 1.
```

If $\|d^k\| \le \delta$ for all $k = 0, 1, \ldots$, then one of the following must occur:

i. There is a k_0 such that $D_{k_0} = \emptyset$.

ii. $f_k^{\max} \downarrow -\infty$.

iii. $f'(x^{m(k)-1}; d^{m(k)-1}) \to 0$.

(d) Use the dog-leg search direction with $H_\nu = \nabla^2 f(x^\nu)$ in conjunction with this line search procedure to minimize the Rosenbrock function. Use the starting point and stopping criteria given in the second homework set. Try the values $p = 3$ and $p = 4$ as suggested above. Compare the outcome of this numerical experiment with your previous algorithms for solving this problem, specifically, problem 4. in homework set 2 and problem 1 above.

3. This problem concerns numerical methods for solving the nonlinear least squares problem

$$\mathcal{P} \qquad \min_{\mathbb{R}^7} f(x) := \frac{1}{2}\|g(x)\|_2^2$$

where matlab function files for the function $g : \mathbb{R}^7 \to \mathbb{R}^8$, its Jacobian, and its Hessians are available through the course webpage.

Each of the algorithms described below is to be initiated at the point $x^0 = zeros(7, 1)$ with a stopping criteria of $\|\nabla f(x)\| \le 10^{-12}$. As part of your output, you need to include

(i) a graph of the norm of the gradient (consider log scale),

(ii) a graph of the function values,

(iii) a graph of the magnitude of the steps taken at each iteration, and

(iv) a table listing the total number of function, gradient, and Hessian evaluations (listed separately).

Write a matlab m-file to implement the *dog-leg* method ($\delta = 3$) with

(a) backtracking line-search ($\gamma = 0.8,\ c = 0.01$), and

(b) non–monotone backtracking line–search ($\gamma = 0.8,\ c = 0.01,\ p = 5$).

Compare the results of these methods when applied with (a) $H_\nu = \nabla^2 f(x^\nu)$ and (b) H_ν equal to the BFGS update with $H_0 = I$.

4. As observed in class, Newton's method for equation solving may fail if it is not initiated sufficiently close to a solution. To overcome this difficulty a number of line search procedures have been proposed that use the Newton step $s = -F'(x)^{-1}F(x)$ as a search direction for solving the equation $F(x) = 0$. These are called *damped Newton steps.* Given $F : \mathbb{R}^n \to \mathbb{R}^n$ smooth with $F'(x)^{-1}$ assumed to exist on all of $\mathbb{R}^n$, propose a line search procedure that uses the Newton step as the search direction and explore its convergence properties. In particular, show that your approach is implementable. Then state and prove a *global* convergence result for your proposed method. By a global convergence result, I mean one that makes no assumption on the location of the initial point.