ISyE 6663 Optimization III

Spring 2011

Assignment 4

Issued: March 29, 2011

Due: April 7, 2011

Problem 1
Program a trust region algorithm that uses the dogleg method. Use your program to minimize the Rosenbrock function

$$f(x_1, x_2) \quad := \quad 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

For the second order term, choose $B^k = \nabla^2 f(x^k)$. Choose $\eta_1 = 0.1$, $\eta_2 = 0.2$, $\eta_3 = 0.9$, $\beta = 1/2$, $\gamma = 1$, $c_1 = 1/2$. Use $\|\cdot\| = \|\cdot\|_2$, and another version with $\|\cdot\| = \|\cdot\|_Q$, with $Q = \nabla^2 f(x^*)$. First try the initial point $x^0 = (1.2, 1.2)$, and then the initial point $x^0 = (-1.2, 1.0)$. Plot a graph of the distance $\|x^k - x^*\|_2$ between the iterate x^k and the optimal solution x^* versus iteration index k for each norm and each initial point. Interpret the results.

Problem 2
Suppose that the objective function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is bounded below and continuously differentiable. A trust region algorithm is used, with matrices B^k for the second order term of the model functions satisfying $\|B^k\| \leq \beta$ for all k. The approximate solutions of the subproblem at each iteration satisfy the two conditions of class. Show that if the iterates remain in a bounded set, then there is a limit point x^∞ of the sequence $\{x^k\}$ such that $\nabla f(x^\infty) = 0$.

Problem 3
The Cauchy-Schwartz inequality states that for any $u, v \in \mathbb{R}^n$, it holds that

$$\left|u^T v\right|^2 \quad \leq \quad \left(u^T u\right)\left(v^T v\right)$$

with equality only when u and v are parallel. Use the Cauchy-Schwartz inequality to show that, for any $Q \in \mathbb{R}^{n \times n}$ positive definite and for any $w \in \mathbb{R}^n$, it holds that

$$|w|_2^4 \quad \leq \quad \left(w^T Q w\right)\left(w^T Q^{-1} w\right)$$

with equality only when w and Qw (and thus also $Q^{-1}w$) are parallel.

Problem 4
Show that if $B \in \mathbb{R}^{n \times n}$ is symmetric, then $B + \lambda I$ is positive definite for all λ sufficiently large.

Problem 5
Equivalent trust-region methods:
Consider a trust region method with subproblems given by

$$\begin{array}{ll} \min\limits_{p \in \mathbb{R}^n} & \left\{m_k^1(p) \;:=\; f(x_k) + \nabla f(x_k)^T p + \tfrac{1}{2} p^T H_k p\right\} \\ \text{subject to} & \|p\|_2 \;\leq\; \Delta_k \end{array}$$

Consider another trust region method with subproblems given by

$$\min_{p \in \mathbb{R}^n} \left\{m_k^2(p) \;:=\; f(x_k) + \nabla f(x_k)^T p + \frac{1}{2} p^T (H_k + \lambda_k I) p\right\}$$

(Note that the subproblem above is an unconstrained optimization problem.)

1. Show that the two trust region methods are equivalent, in the sense that for every $\Delta_k > 0$, there is a $\lambda_k \geq 0$ such that $H_k + \lambda_k I$ is positive semidefinite, $\nabla f(x_k)$ is in the range of $H_k + \lambda_k I$ (which automatically holds if $H_k + \lambda_k I$ is positive definite and thus nonsingular), and every optimal solution of the first subproblem is an optimal solution of the second subproblem (in particular, if $H_k + \lambda_k I$ is positive definite then the two subproblems have the same optimal solution); and for every $\lambda_k \geq 0$ and every optimal solution of the second subproblem (which implies that $H_k + \lambda_k I$ is positive semidefinite and $\nabla f(x_k)$ is in the range of $H_k + \lambda_k I$), there is a $\Delta_k > 0$ such that the optimal solution of the second subproblem is also optimal for the first subproblem (again, if $H_k + \lambda_k I$ is positive definite then the two subproblems have the same optimal solution).

2. List some potential advantages/disadvantages of each of the two trust region methods.

Problem 6

Consider the following two-dimensional subspace minimization problem (which gives a solution to the subproblem at least as good as that of the dogleg method).

$$\begin{array}{rrcl} \min\limits_{p \in \mathbb{R}^n} & \{m(p) & := & a + b^T p + \frac{1}{2} p^T B p\} \\ \text{subject to} & \|p\|_2 & \leq & \Delta \\ & p & \in & \text{span}[b, B^{-1} b] \end{array}$$

Suppose that B is positive definite. Describe (in a precise way) a procedure to solve the two-dimensional subspace minimization problem above.