Numerical Optimization, 2020 Fall

Homework 8

Due 14:59 (CST), Dec. 10, 2020

(NOTE: Homework will not be accepted after this due for any reason.)

Throughout this assignment, we focus on the following trust region subproblem, which reads

$$\begin{aligned} \min_{\boldsymbol{d}\in\mathbb{R}^n} \quad & m_k(\boldsymbol{d}) := f(\boldsymbol{x}_k) + \nabla f(\boldsymbol{x}_k)^T \boldsymbol{d}_k + \tfrac{1}{2}\boldsymbol{d}_k^T H_k \boldsymbol{d}_k \\ \text{s.t.} \quad & \|\boldsymbol{d}\| \le \Delta_k, \end{aligned} \tag{1}$$

where $\Delta_k > 0$ is the trust-region radius.

Note: Throughout this assignment, the notion of positive definiteness applies exclusively to symmetric matrices. Thus whenever we say that a matrix is positive (semi)definite, we implicitly assume that the matrix is symmetric.

1 Cauchy point calculation

[20pts] Please write down a closed-form expression of the Cauchy point. (Make sure you provided detailed proof; otherwise you won't earn marks.)

Specifically, first solve the a linear version of (1) to obtain vector $\boldsymbol{d}_k^s$, that is,

$$\boldsymbol{d}_k^s = \arg\min_{\boldsymbol{d}\in\mathbb{R}^n} f(\boldsymbol{x}_k) + \nabla f(\boldsymbol{x}_k)^T \boldsymbol{d}_k \qquad \text{s.t.} \quad \|\boldsymbol{d}\| \le \Delta_k. \tag{2}$$

Then, calculate the scalar $\tau_k > 0$ that minimizes $m_k(\tau \boldsymbol{d}_k^s)$ subject to the trust region bound, that is

$$\tau_k = \arg\min_{\tau\ge 0} \; m_k(\tau \boldsymbol{d}_k^s) \qquad \text{s.t.} \quad \|\tau \boldsymbol{d}_k^s\| \le \Delta_k. \tag{3}$$

Set $\boldsymbol{d}_k^c = \tau_k \boldsymbol{d}_k^s$.

2 Local convergence for trust region methods

[20pts] Given a step $\boldsymbol{d}_k$, consider the ratio (with positive denominator):

$$\rho_k := \frac{f(\boldsymbol{x}_k) - f(\boldsymbol{x}_k + \boldsymbol{d}_k)}{m_k(\boldsymbol{0}) - m_k(\boldsymbol{d}_k)}. \tag{4}$$

Show that if $\Delta_k \to 0$, then $\rho_k \to 1$. (This proves that for Δ_k sufficiently small, $m_k(\boldsymbol{d})$ approximates $f(\boldsymbol{x}_k + \boldsymbol{d}_k)$ well.)

3 Exact line search

[20pts] Consider minimizing the following quadratic function

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \quad f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x}, \tag{5}$$

where $Q \in \mathbb{R}^{n\times n}$ is positive definite and $\boldsymbol{b} \in \mathbb{R}^n$.

Let $\boldsymbol{d}_k$ be a descent direction at the kth iterate. Suppose that we search along this direction from $\boldsymbol{x}^k$ for a new iterate, and the line search are exact. Please find the stepsize α. This can be achieved exactly solving the following one-dimensional minimization problem

$$\min_{\alpha>0} \quad f(\boldsymbol{x}_k + \alpha\boldsymbol{d}_k). \tag{6}$$

4 The conjugate gradient algorithm

[20pts] Let $A \in \mathbb{R}^{n\times n}$ be a positive definite matrix. Show that if the directions $\boldsymbol{d}_0, \ldots, \boldsymbol{d}_k \in \mathbb{R}^n$, $k \le n-1$, are A-conjugate, then they are linearly independent. (Hint: We say that a set of nonzero vectors $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_m \in \mathbb{R}^n$ are A-conjugate if $\boldsymbol{d}_i^T A\boldsymbol{d}_i = 0, \quad \forall i, j, \quad i \ne j$.)

5 Trust region subproblems

Consider the trust region subproblem (1), and H_k is positive definite. Let θ_k denote the angle between $\boldsymbol{d}_k$ and $-\nabla f(\boldsymbol{x}_k)$, defined by

$$\cos\theta_k = \frac{-\nabla f(\boldsymbol{x}_k)^T\boldsymbol{d}_k}{\|\nabla f(\boldsymbol{x}_k)\|\|\boldsymbol{d}_k\|}.$$

Show that

(i) [10pts] For sufficiently large Δ_k, the trust region subproblem (1) will be solved by the Newton step.

(ii) [10pts] When Δ_k approaches 0, the angle $\theta_k \to 0$.