MATH60005/70005: Optimization (Autumn 23-24)

Week 4: Exercises

Dr Dante Kalise & Dr Estefanía Loayza-Romero
Department of Mathematics
Imperial College London, United Kingdom
{dkaliseb,kloayzar}@imperial.ac.uk

1. Find the exact linesearch stepsize when $f(\mathbf{x})$ is a quadratic function $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} + 2\mathbf{b}^\top \mathbf{x} + \mathbf{c}$ where $\mathbf{A}$ is an $n \times n$ positive definite matrix, $\mathbf{b} \in \mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}$.

2. Let $\mathbf{A}$ be a symmetric $n \times n$ matrix, $\mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then the function $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} + 2\mathbf{b}^\top \mathbf{x} + \mathbf{c}$ is a $C^{1,1}$ function. The smallest Lipschitz constant of f is $2\|\mathbf{A}\|_2$

3. Show that $f(\mathbf{x}) = \sqrt{1+x^2} \in C_L^{1,1}$.

4. Give an example of a function $f \in C_L^{1,1}(\mathbb{R})$ and a starting point $x_0 \in \mathbb{R}$ such that the problem $\min f(x)$ has an optimal solution and the gradient method with constant stepsize $t = \frac{2}{L}$ diverges.

5. Consider the localization problem where we are given m locations of sensors $\mathscr{A} := \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\}$, with each sensor in $\mathbb{R}^n$, and approximate distances between the sensors and an unknown source located at $\mathbf{x} \in \mathbb{R}^n$: $d_i \approx \|\mathbf{x} - \mathbf{a}_i\|$. We try to find the source location $\mathbf{x}$ given the sensor locations $\mathscr{A}$ and the approximate distances $d_1, d_2, \dots, d_m$. For this, we write the optimization problem:

$$\min_{\mathbf{x}} \left\{ f(\mathbf{x}) \equiv \sum_{i=1}^{m} (\|\mathbf{x} - \mathbf{a}_i\| - d_i)^2 \right\}.$$

   a) State the first-order optimality condition for this problem, and show that for $\mathbf{x} \notin \mathscr{A}$ it is equivalent to

$$\mathbf{x} = \frac{1}{m} \left\{ \sum_{i=1}^{m} \mathbf{a}_i + \sum_{i=1}^{m} d_i \frac{\mathbf{x} - \mathbf{a}_i}{\|\mathbf{x} - \mathbf{a}_i\|} \right\}$$

   b) Show that the iteration:

$$\mathbf{x}^{k+1} = \frac{1}{m} \left\{ \sum_{i=1}^{m} \mathbf{a}_i + \sum_{i=1}^{m} d_i \frac{\mathbf{x}^k - \mathbf{a}_i}{\|\mathbf{x}^k - \mathbf{a}_i\|} \right\}$$

   is a gradient method, assuming that $\mathbf{x}^k \notin \mathscr{A}$ for all $k \geq 0$. What is the stepsize?

c) Write an explicit Gauss-Newton iteration of the form

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mathbf{d}^k\,,$$

giving an expression for $\mathbf{d}^k$ in terms of the Jacobian and vectorized cost for this problem, without computing the inverse.

6. Consider the quadratic function $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top Q\mathbf{x}$$

where Q is a symmetric matrix of size 2×2 with eigenvalues $0 < \lambda_{\min} < \lambda_{\max}$. Suppose we apply the gradient descent method to the problem of minimizing f, with exact line search and initial point

$$\mathbf{x}_0 = \frac{1}{\lambda_{\min}}\mathbf{u}_{\min} + \frac{1}{\lambda_{\max}}\mathbf{u}_{\max}$$

where $\mathbf{u}_{\min}$ and $\mathbf{u}_{\max}$ are the norm one eigenvectors associated with $\lambda_{\min}$ and $\lambda_{\max}$, respectively.

a) Show that after 1 iteration

$$\mathbf{x}_1 = \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)\left(\frac{1}{\lambda_{\min}}\mathbf{u}_{\min} - \frac{1}{\lambda_{\max}}\mathbf{u}_{\max}\right)\,.$$

b) Assuming that

$$\mathbf{x}_k = \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^k\left(\frac{1}{\lambda_{\min}}\mathbf{u}_{\min} + \frac{(-1)^k}{\lambda_{\max}}\mathbf{u}_{\max}\right) \quad \text{for} \;\; k = 0, 1, \ldots,$$

show that

$$\frac{f(\mathbf{x}_{k+1})}{f(\mathbf{x}_k)} = \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^2\,.$$

Using this, what can be said about the convergence of this method based on the ratio $\kappa = \frac{\lambda_{\max}}{\lambda_{\min}}$?

Quadratic Optimization Benchmark

Consider the quadratic minimization problem

$$\min_{\mathbf{x}}\left\{\mathbf{x}^\top\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{R}^5\right\}$$

where $\mathbf{A}$ is the 5×5 Hilbert matrix defined by

$$\mathbf{A}_{i,j} = \frac{1}{i + j - 1}, \qquad i, j = 1, 2, 3, 4, 5$$

The matrix can be constructed via the MATLAB command `A=hilb(5)`. Run the following methods and compare the number of iterations required by each of the methods when the initial vector is $\mathbf{x}^0 = (1, 2, 3, 4, 5)^\top$ to obtain a solution $\mathbf{x}^*$ with $\|\nabla f(\mathbf{x})\| \leq 10^{-4}$:

- Gradient method with backtracking stepsize rule and parameters $\alpha = 0.5, \beta = 0.5, s = 1$
- Gradient method with backtracking stepsize rule and parameters $\alpha = 0.1, \beta = 0.5, s = 1$
- Diagonally scaled gradient method with diagonal elements $\mathbf{D}_{i,i} = \frac{1}{\mathbf{A}_{i,i}}, i = 1,2,3,4,5$ and exact line search;
- Diagonally scaled gradient method with diagonal elements $\mathbf{D}_{i,i} = \frac{1}{\mathbf{A}_{i,i}}, i = 1,2,3,4,5$ and backtracking line search with parameters $\alpha = 0.1, \beta = 0.5$ $s = 1$.