02610 Optimization and Data Fitting 28 September 2020

Homework assignment 1

Hand in on DTU Inside before 4 October 10pm.

1 One-page report on the exercises for week 5 (40%)

2 Stationary points and steepest descent method (20%)

Consider the test function

$$f(\boldsymbol{x}) = f_1(\boldsymbol{x})^2 + f_2(\boldsymbol{x})^2, \qquad \boldsymbol{x} \in \mathbb{R}^2$$

where

$$\begin{aligned} f_1(\boldsymbol{x}) &= x_1 + x_2^2 - 1, \\ f_2(\boldsymbol{x}) &= x_1 + 3. \end{aligned}$$

(5%) Compute the gradient $\nabla f(\boldsymbol{x})$ and Hessian $\nabla^2 f(\boldsymbol{x})$ by hand.

(5%) Find the three stationary points by calculating the first-order optimality condition, and prove that two are global minimizers and one is a saddle point.

(5%) Use Matlab to draw a contour plot of the function f, and mark these three points on the contour plot.

(5%) Use Matlab to apply the steepest descent method with backtracking line search on the problem of minimizing f. Set the stopping criteria as $\|\nabla f(\boldsymbol{x}_k)\|_\infty \leq 10^{-5}$, and the parameters in the backtracking are $\bar{\alpha} = 1$, $\rho = 0.5$ and $c = 0.5$. Test with four different starting points: $[-2, -1]^T$, $[-2, 0]^T$, and $[-2, 1]^T$. For each of the three starting points, show the number of iterations and the point to which the method converged.

3 Local convergence in Newton's method (40%)

Consider the minimization problem

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) = \frac{x_1^4}{4} - x_1^2 + 2x_1 + (x_2 - 1)^2.$$

(5%) Compute the gradient $\nabla f(\boldsymbol{x})$ and Hessian $\nabla^2 f(\boldsymbol{x})$ by hand.

(5%) Make a contour plot of $f(\boldsymbol{x})$ in the interval $-3 \leq x_1 \leq 3$ and $-3 \leq x_2 \leq 3$.

(5%) Run 10 iterations of Newton's method with the starting point $\boldsymbol{x}_0 = [0, 2]^T$ and the step length 1, and output the iterates $\boldsymbol{x}_k$. Did the method converge? What happened and why? (Hint: Pay attention on the Hessian matrix at some iterates, especially at the starting point.)

(10%) Now apply Newton's method with backtracking line search and set the stopping criteria same as $\|\nabla f(\boldsymbol{x}_k)\|_\infty \le 10^{-8}$ or $k \ge 2000$. With the Newton search direction, $\boldsymbol{p}_k = -\nabla^2 f(\boldsymbol{x}_k)^{-1}\nabla f(\boldsymbol{x}_k)$ (note that $\boldsymbol{p}_k$ is not normalized here), in the backtracking line search we set the minimum step length as 0.01, i.e., $\alpha_k \ge 0.01$ for all $k = 1, 2, \cdots$. Plot $f(\boldsymbol{x}_k)$ (with `plot`), $\|\nabla f(\boldsymbol{x}_k)\|_\infty$ (with `semilogy`), and α (with `plot`). Does Newton's method converge now? If yes, how did it converge? Why? How did α_k change? Are the function values $f(\boldsymbol{x}_k)$ monotonically decreasing? Why?

(5%) Apply steepest descent method with backtracking line search and the same starting point. Plot $f(\boldsymbol{x}_k)$, $\|\nabla f(\boldsymbol{x}_k)\|_\infty$, and α. Does it converge? Comparing with Newton's method with line search, which one converges faster?

(10%) Now repeat your test with the starting point $\boldsymbol{x}_0 = [-1, 0]^T$, and compare the convergence results by Newton's method, Newton's with line search, and steepest descent with line search. What are your results and conclusions?