02610 Optimization and Data Fitting

Exercises for Week 5

1 Rosenbrock function

In this project, we consider Rosenbrock function

$$f(\boldsymbol{x}) = f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

We will start with studying some properties of this function, then apply the methods that we have learnt until now to find its minimizer.

Please first complete all the following steps, then write a report to address the listed questions based on the theories that you learned in the course and your results from this project. This report will be the first part of your homework assignment.

You can work in groups with maximum 3 members. Each group only need hand in ONE individualized report.

1. (by hand) Compute the gradient $\nabla f(\boldsymbol{x})$ and Hessian $\nabla^2 f(\boldsymbol{x})$.
2. (by hand) Show that $\boldsymbol{x}^* = [1, 1]^T$ is the only local minimizer of this function.
3. (by hand) Verify that the minimizer satisfies the sufficient optimality conditions. (Hint: you can compute eigenvalues of a matrix in Matlab using the command `eig`).
4. (in Matlab) Make a contour plot of $f(\boldsymbol{x})$ in the interval $-1 \leq x_1 \leq 2$ and $-1 \leq x_2 \leq 2$. Locate the minimizer in this plot.
5. (in Matlab) Make a contour plot of $\log(f(\boldsymbol{x}))$. This is sometimes convenient for the case that $f(\boldsymbol{x})$ has very large or small values. Compare this to your previous contour plot. Do they give the same information?
6. (in Matlab) Write a Matlab function `rosenbrock` to return the function value $f(\boldsymbol{x})$, its gradient $\nabla f(\boldsymbol{x})$ and Hessian $\nabla^2 f(\boldsymbol{x})$.
7. (in Matlab) In the following questions, you will apply the steepest descent method, Newton's method and BFGS method to find the minimizer of $f(\boldsymbol{x})$. Set the stopping criteria as $\|\nabla f(\boldsymbol{x}_k)\|_\infty \leq 10^{-10}$ or $k \geq 2000$. For each method, you need save $\boldsymbol{x}_k$, $e_k = \|\boldsymbol{x}_k - \boldsymbol{x}^*\|_2$ and $f(\boldsymbol{x}_k)$.
   (a) Apply the steepest descent method with a fixed step length to minimize $f(\boldsymbol{x})$. First, use the starting point $\boldsymbol{x}_0 = [1.2, 1.2]^T$ and try different choices of the step length. Use `semilogy` to plot e_k and $f(\boldsymbol{x}_k)$. How do the values change? Does it always converge?

Then, use a starting point $\boldsymbol{x}_0 = [-1.2, 1]^T$ a bit farther away from $\boldsymbol{x}^*$. Try again with the same choices of the step length. Does it always converge? According to the contour plot of $f(\boldsymbol{x})$ explain how the iterate $\boldsymbol{x}_k$ moves with large and small step length.

(b) Apply the steepest descent with backtracking line search on finding the minimizer of $f(\boldsymbol{x})$ with both choices of starting points. Set the initial step length $\alpha_0 = 1$ and output $\{\alpha_k : k = 0, 1, \cdots\}$ from both choices of $\boldsymbol{x}_0$. Use `semilogy(stat.alpha, '.')` to plot all $\{\alpha_k\}$ in both cases. Observe how α_k changes.

(c) Compare the results by applying Newton's method with both starting points, $[1.2, 1.2]^T$ and $[-1.2, 1]^T$, and fixed step length 1 on minimizing $f(\boldsymbol{x})$. Does Newton's method converge in both cases? How many iterations does it need? You can plot the iterates in the contour figure to see how Newton's method converged. Which starting point need more iterations? Why? Are function values $f(\boldsymbol{x}_k)$ monotonically decreasing? Why?

(d) Use the BFGS method with backtracking line search to minimizing $f(\boldsymbol{x})$. Output $\boldsymbol{x}_k$, $e_k = \|\boldsymbol{x}_k - \boldsymbol{x}^*\|_2$ and $f(\boldsymbol{x}_k)$.

(e) Plot two figures by using `loglog` function:

   i. One figure to show the plots of e_k from the steepest descent (with line search), Newton's (with step size 1) and BFGS (with line search) as functions of the iteration number with the starting point $\boldsymbol{x}_0 = [1.2, 1.2]^T$;

   ii. the other figure to show the plots of $f(\boldsymbol{x}_k)$ from all three methods as functions of the iteration number with the starting point $\boldsymbol{x}_0 = [1.2, 1.2]^T$;

8. (**Report**) Each group only need write ONE report. At the beginning of the report, please indicate the names and student numbers of all group members.

   - **Maximum 3 members in each group.**
   - **The report must be subjected to individualization.** Each member's individual contributions must be clearly distinguishable in regard to the sections in the assignment.
   - **The page limit of the report is 3.**
   - **The report should include the figures and the discussions according to the following list.** Your discussions can be based on the theories learnt in the course and the results that you got in the project.

   i. Discuss how a fixed step length for the steepest descent method can affect on the convergences? In the report, please explain what you tested and what conclusions you obtained through those tests.

   ii. Include the two figures obtained in step (e).

iii. Based on your tests on applying the steepest descent with line search, Newton's and BFGS for finding Rosenbrock function's minimizer, compare and comment on the progress of these three methods, e.g. how the iterates move; numerically check and compare the convergence rate; how the computational costs are; CPU time; etc.

iv. Based on your tests and the theories learnt in the course, comment on advantages and disadvantages of these three methods.