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. In the end of this project, you should finish an ONE-PAGE report to include the required results and comments highlighted in blue color. This report will be the first part of your homework assignment.

1. (by hand) Compute the gradient $\nabla f(\boldsymbol{x})$ and Hessian $\nabla^2 f(\boldsymbol{x})$, and include them in your report.

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 $-1 \le x_1 \le 2$ and $-1 \le x_2 \le 2$. Locate the minimizer in this plot. Include the contour plot in your report.

5. (in Matlab) Make a contour plot of $\log(f(\boldsymbol{x}))$. This is sometimes convenient for including both large and 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 $\|\nabla f(\boldsymbol{x}_k)\|_\infty \le 10^{-10}$ or $k \ge 2000$. 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.

Conclude how the step length affects on the result and convergence? In the report, please explain what you tested and what conclusions you obtained through those tests. You do not need include any numerical results from this part to the report.

(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 $\bar{\alpha} = 1$ and output α_k used in each method at each iteration. Use `semilogy(stat.alpha, '.')` to plot all α values. Observe how α_k changes. Is the solution closer to $\boldsymbol{x}^*$ than the ones with fixed step length?

(c) Compare the results by applying Newton's method with both starting points 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) Make 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 and BFGS as functions of the iteration number with the starting point $\boldsymbol{x}_0 = [1.2, 1.2]^T$;

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

   Include both figures in your report.

(f) Comment on the convergence of all methods that you have tested. Comment on advantages and disadvantages of each method. Include your comments in the report.