AMS230 – Homework #1

Zayd Hammoudeh

April 13, 2018

Name: Zayd Hammoudeh
Course Name: AMS230
Assignment Name: Homework #1
Due Date: April 18, 2018
Student Discussions: I discussed the problems with the following students. All write-ups were prepared separately and independently.

-

Exercise #1
Exercise 2.1 in Nocedal and Wright.

Compute the gradient, $\nabla f(x)$, and Hessian, $\nabla^2 f(x)$ of the Rosenbrock function

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

The gradient $\nabla f(x)$ is:

$$\begin{aligned}
\nabla f(x) &= \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \end{bmatrix} \\
&= \begin{bmatrix} -400x_1(x_2 - x_1^2) - 2(1 - x_1) \\ 200(x_2 - x_1^2) \end{bmatrix} \\
&= \boxed{\begin{bmatrix} 400x_1^3 - 400x_1x_2 + 2x_1 - 2 \\ 200x_2 - 200x_1^2 \end{bmatrix}}.
\end{aligned}$$

The Hessian $\nabla^2 f(x)$ equals:

$$\begin{aligned}
\nabla^2 f(x) &= \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} \end{bmatrix} \\
&= \boxed{\begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}}
\end{aligned}$$

(a) **Show that $x^* = (1, 1)^{\mathbf{T}}$ is the only local minimizer of the function, and that the Hessian at that point is positive definite.**

To determine the minimizer(s) (if any), set the gradient equal to 0. Therefore,

$$\begin{bmatrix} 400x_1^3 - 400x_1x_2 + 2x_1 - 2 \\ 200x_2 - 200x_1^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Therefore, $x_2 = x_1^2$. If we substitute this into the first equation, we get:

$$\begin{aligned}
400x_1^3 - 400x_1x_2 + 2x_1 - 2 &= 0 \\
400x_1^3 - 400x_1x_1^2 + 2x_1 - 2 &= 0 \\
2x_1 - 2 &= 0 \\
x_1 &= 1
\end{aligned}$$

It is clear then that the only root is $\boxed{(1, 1)^{\mathrm{T}}}$.

The Hessian matrix $\nabla^2 f(x^*)$ of the Rosenbrock function equals:

$$\nabla^2 f(x^*) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}.$$

Hessian matrix, $\nabla^2 f(x^*)$, is positive definite (PD) if all of its eigenvalues are positive. These are found via:

$$\begin{aligned} |A - \lambda I| = \begin{vmatrix} 802 - \lambda & -400 \\ -400 & 200 - \lambda \end{vmatrix} &= 0 \\ 160400 - 1002\lambda + \lambda^2 - 160000 &= 0 \\ 400 - 1002\lambda + \lambda^2 &= 0 \end{aligned}$$

Solving for λ (in Matlab), the eigenvalues are $\boxed{\lambda_1 \approx 0.3994}$ and $\boxed{\lambda_2 \approx 1001.6}$. Therefore, $\nabla^2 f(x^*)$ is PD.

Exercise #2
Exercise 2.9 in Nocedal and Wright.

Consider the function $f(x_1, x_2) = (x_1 + x_2^2)^2$**. At the point** $x^{\mathbf{T}} = (1, 0)$**, we consider the search direction** $p^{\mathbf{T}} = (-1, 1)$**. Show that** p **is a descent direction.**

Any vector p is a descent direction if $p \cdot \nabla f(x_1, x_2) < 0$. The gradient of $f(x_1, x_2)$ equals:

$$\begin{aligned}\nabla f(x_1, x_2) &= \begin{bmatrix} \frac{\partial f(x_1,x_2)}{\partial x_1} \\ \frac{\partial f(x_1,x_2)}{\partial x_2} \end{bmatrix} \\ &= \begin{bmatrix} 2(x_1 + x_2^2) \\ 4x_2(x_1 + x_2^2) \end{bmatrix}\end{aligned}$$

Therefore, the gradient at $(1, 0)^{\mathrm{T}}$ equals $(2, 0)^{\mathrm{T}}$. The inner product is in turn:

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 0 \end{bmatrix} = -1 \cdot 2 + 1 \cdot 0 = \boxed{-2 < 0},$$

making p a descent direction.

(a) Find all minimizers of the problem (2.10).

Exercise #3
Exercise 3.6 in Nocedal and Wright.

Consider the steepest descent method with exact line searches applied to the convex quadratic function (3.24). Using the properties given in this chapter, show that if the initial point is such that $x_0 - x^*$ is parallel to an eigenvector of Q, then the steepest descent method will find the solution in one step.

Exercise #4

Consider the steepest descent method with exact line searches pplied to the convex quadratic function $f(x) = \frac{1}{2}x^{\mathbf{T}}Qx$, where Q is symmetric and positive definite. Show that the search direction at step $k+1$ is always orthogonal to the search direction at step k, i.e., $p_k p_{k+1} = 0$ for all k.

Exercise #5

Program Algorithm 3.5 (Line-Search) and Algorithm 3.6 (zoom) in the text book. For the selection of trial step length α_i in Algorithm 3.5 and α_j in Algorithm 3.6, you can use bisection.

Exercise #6

Consider the problem of minimizing:

$$f(x_1, x_2) = (cx_1 - 2)^4 + x_2^2(cx_1 - 2)^2 - (x_2 + 1)^2,$$

where c is a nonzero parameter.

(a) Compute $\nabla f(x)$ and $\nabla f^2(x)$, and find the optimal solution.

$$\begin{aligned}\nabla f(x) &= \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \end{bmatrix} \\ &= \boxed{\begin{bmatrix} 4c(cx_1 - 2)^3 + 2cx_2^2(cx_1 - 2) \\ 2x_2(cx_1 - 2)^2 - 2(x_2 + 1) \end{bmatrix}}\end{aligned}$$

$$\begin{aligned}\nabla^2 f(x) &= \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} \end{bmatrix} \\ &= \boxed{\begin{bmatrix} 12c^2(cx_1 - 2)^2 + 2c^2x_2^2 & 4cx_2(cx_1 - 2) \\ 4cx_2(cx_1 - 2) & 2(cx_1 - 2)^2 - 2 \end{bmatrix}}.\end{aligned}$$

The optimal solution occurs when the gradient equals $\vec{0}$. We find then:

$$\begin{aligned} 2x_2(cx_1 - 2)^2 - 2(x_2 + 1) &= 0 \\ x_2(cx_1 - 2)^2 - x_2 &= 1 \\ x_2 &= \frac{1}{(cx_1 - 2)^2 - 1} \end{aligned}$$

$$\begin{aligned} 4c(cx_1 - 2)^3 + 2cx_2^2(cx_1 - 2) &= 0 \\ 2(cx_1 - 2)^2 + x_2^2 &= 0 \\ x_2^2 &= -2(cx_1 - 2)^2 \end{aligned}$$

Substituting that into the other function, we get

(b) Program the steepest descent method (using your program from Problem 5 to find the step length).

(c) Use your program to numerically solve this problem under two cases: i) $c = 1$, ii) $c = 10$. Compare the convergence in these two cases.

(d) **Using Theorem 3.4, explain why as c increases the convergence of the steepest descent method deteriorates.**