Numerical Optimization 2024 - Homework 3

Aral Cimcim, k11720457

Artificial Intelligence, JKU Linz

April 12, 2024

Problem 1.

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

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

Show that $x^* = (1,1)^T$ is the only local minimizer of this function, and that the Hessian matrix at that point is positive definite.

$$\frac{\partial f}{\partial x_1} = 100\ 2(x_2 - x_1^2)(-2x_1) + 2(1 - x_1)(-1)$$

$$= -400x_1(x_2 - x_1^2) - 2(1 - x_1)$$

$$\frac{\partial f}{\partial x_2} = 100\ 2(x_2 - x_1^2)$$

$$\nabla f(x) = \begin{bmatrix} -400x_1(x_2 - x_1^2) - 2(1 - x_1) \\ 200(x_2 - x_1^2) \end{bmatrix}$$

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

$$\frac{\partial^2 f}{\partial x_1^2} = -400[(x_2 - x_1^2) + (-2x_1 x_1)] + 2$$

$$= -400(x_2 - 3x_1^2) + 2$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = -400x_1$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = -400x_1$$

$$\frac{\partial^2 f}{\partial x_2^2} = 200$$

$$\nabla^2 f(x) = \begin{bmatrix} -400(x_2 - 3x_1^2) + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

Optimality conditions: $\nabla f(x^*) = 0$, $\nabla^2 f(x^*)$ is positive definite.

For $x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\nabla f(x^*) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\nabla^2 f(x^*) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$

Testing for determinants:

The determinant of $\nabla^2 f(x^*)$ is $160400 - 160000 = 400 > 0$ and determinant of $802 > 0$ meaning that the matrix is positive definite.

Problem 2.

Show that the function $f(x) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2$ has only one stationary point, and that it is neither a maximum or minimum, but a saddle point. Sketch the contour lines of f.

$$\frac{\partial f}{\partial x_1} = 8 + 2x_1$$

$$\frac{\partial f}{\partial x_2} = 12 - 4x_2$$

$$\nabla f(x) = \begin{bmatrix} 8 + 2x_1 \\ 12 - 4x_2 \end{bmatrix}$$

Solving for $\nabla f(x) = 0$:

$$8 + 2x_1 = 0, x_1 = -4$$

$$12 - 4x_2 = 0, x_2 = 3$$

The stationary point is $(x_1, x_2) = (-4, 3)$

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

$$\frac{\partial^2 f}{\partial x_1^2} = \frac{\partial(8 + 2x_1)}{\partial x_1} = 2$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial(8 + 2x_1)}{\partial x_2} = 0$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial(12 - 4x_2)}{\partial x_1} = 0$$

$$\frac{\partial^2 f}{\partial x_2^2} = \frac{\partial(12 - 4x_2)}{\partial x_2} = -4$$

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$$

Testing for determinants:
The determinant of $\nabla^2 f(x^*)$ is $-8 < 0$ meaning that x^* is a saddle point.

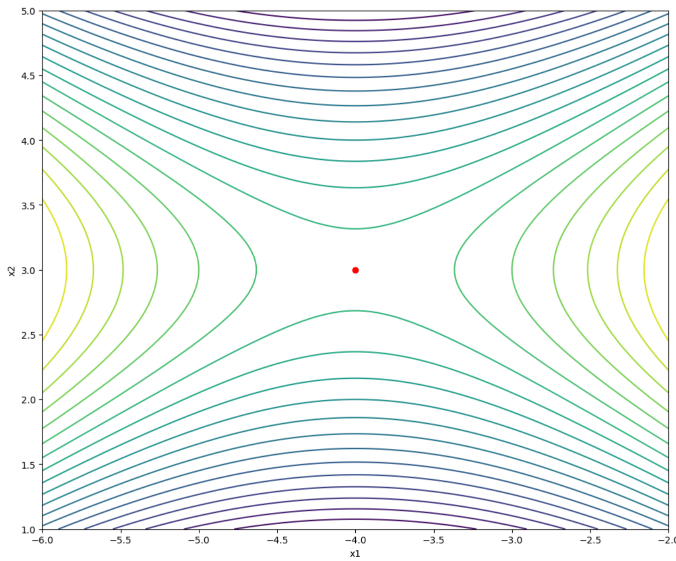

Figure 1: Contour plot of $f(x) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2$

Problem 3.

Consider the function $f(x_1, x_2) = (x_1 + x_2^2)^2$. At the point $x^T = (1, 0)$ we consider the search direction $p^T = (-1, 1)$. Show that p is a descent direction and find all minimizers of the problem (2.10).

$$f(x_1, x_2) = (x_1 + x_2^2)^2$$

$$\frac{\partial f}{\partial x_1} = 2(x_1 + x_2^2)$$

$$\frac{\partial f}{\partial x_2} = 4x_2(x_1 + x_2^2)$$

$$\nabla f(x) = \begin{bmatrix} 2(x_1 + x_2^2) \\ 4x_2(x_1 + x_2^2) \end{bmatrix}$$

$$p = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

plug in point $x^T = (1, 0)$

$$\nabla f(x^T) = \begin{bmatrix} 2(1+0) \\ 4 \cdot 0 \cdot (1+0) \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$\nabla f(x^T) \cdot p = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -2 + 0 = -2$$

$-2 < 0$ suggests that p is a descent direction.

Minimizers:
From equation 2.10:

$$\min f(x_k + \alpha p_k) = x + \alpha p = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1-\alpha \\ \alpha \end{bmatrix}$$

$$f(x + \alpha p) = (1 - \alpha + \alpha^2)^2$$

Substitute $x_1 = 1 - \alpha$, $x_2 = \alpha$ into $f(x_1, x_2)$:

$$f(x_1, x_2) = (1 - \alpha + \alpha^2)^2 = (1 - \alpha + \alpha^2)^2$$

$$\frac{\partial f}{\partial \alpha} = 2(1 - \alpha + \alpha^2)(-1 + 2\alpha) = 0$$

Step length is $\alpha = 1/2$

$$\frac{\partial^2 f}{\partial \alpha^2} = 2(1 - \alpha + \alpha^2)(-1 + 2\alpha) = (2 - 2\alpha + 2\alpha^2)(2\alpha - 1)$$

$$= (4\alpha - 2)(2\alpha - 1) + 4 - 4\alpha + 4\alpha^2$$

$$= 8\alpha^2 - 8\alpha + 2 + 4 - 4\alpha + 4\alpha^2$$

$$= 12\alpha^2 - 12\alpha + 6$$

$$= 6(2\alpha^2 - 2\alpha + 1)$$

$$= 6(2(1/2)^2 - 2(1/2) + 1) = 3$$

$$\frac{\partial^2 f}{\partial \alpha^2} = 3 > 0 \text{ holds for step length} = 1/2$$