Optimization
Homework 1

Erika Rivadeneira
erika.rivadeneira@cimat.mx

February 9, 2021

1. The directional derivative $\frac{\partial f}{\partial v}(x_0, y_0, z_0)$ of a differentiable function f are $\frac{3}{\sqrt{2}}, \frac{1}{\sqrt{2}}$ and $-\frac{1}{\sqrt{2}}$ in the directions of vectors $\left[0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]^T$, $\left[\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right]^T$ and $\left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right]^T$. Compute $\nabla f(x_0, y_0, z_0)$

   Solution

   Since $D_v f(x_0, y_0, z_0) = \nabla f(x_0, y_0, z_0)^T$ *we can write the following system in matrix form*

$$\begin{bmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x_0} \\ \frac{\partial f}{\partial y_0} \\ \frac{\partial f}{\partial z_0} \end{bmatrix} = \begin{bmatrix} 3/\sqrt{2} \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

   We want to compute $\nabla f(x_0, y_0, z_0)$, *then we have to solve the system of equations*

$$\frac{1}{\sqrt{2}}\frac{\partial f}{\partial y} + \frac{1}{\sqrt{2}}\frac{\partial f}{\partial z} = \frac{3}{\sqrt{2}} \tag{1}$$

$$\frac{1}{\sqrt{2}}\frac{\partial f}{\partial x} + \frac{1}{\sqrt{2}}\frac{\partial f}{\partial z} = \frac{1}{\sqrt{2}} \tag{2}$$

$$\frac{1}{\sqrt{2}}\frac{\partial f}{\partial x} + \frac{1}{\sqrt{2}}\frac{\partial f}{\partial y} = -\frac{1}{\sqrt{2}} \tag{3}$$

   solving for $\frac{\partial f}{\partial z_0}$ *in equation (1) we get*

$$\frac{\partial f}{\partial z_0} = 3 - \frac{\partial f}{\partial x_0} \tag{4}$$

   (4) *in* (2)

$$\frac{\partial f}{\partial x_0} - \frac{\partial f}{\partial y_0} = -2 \tag{5}$$

   Adding (5) and (2)

$$\frac{\partial f}{\partial x_0} = -\frac{3}{2}$$

   then

$$\frac{\partial f}{\partial z_0} = \frac{5}{2} \quad \textit{and} \quad \frac{\partial f}{\partial y_0} = \frac{1}{2}$$

   Therefore,

$$\nabla f(x_0, y_0, z_0) = \left(-\frac{3}{2}, \frac{1}{2}, \frac{5}{2}\right)$$

2. Show that the level curves of the function $f(x, y) = x^2 + y^2$ are orthogonal to the level curves of $g(x, y) = \frac{y}{x}$ for all (x, y).

   Solution
   In order to show that the level curves of the function $f(x, y)$ *are orthogonal to the level curves of* $g(x, y)$ *we have to verify that*

$$\nabla f(x, y) \cdot \nabla g(x, y) = 0$$

Then, we have that

$$\nabla f(x,y) = (2x, 2y)$$
$$\nabla g(x,y) = \left(-\frac{y}{x^2}, \frac{1}{x}\right)$$

So that

$$\nabla f(x,y) \cdot \nabla(x,y) = -\frac{-2xy}{x^2} + \frac{2y}{x} = 0$$

Therefore, the level curves of f(x,y) are orthogonal to the level curves of $g(x,y)$.

3. Compute the stationary points of $f(x,y) = \frac{3x^4-4x^3-12x^2+18}{12(1+4y^2)}$ and determine their corresponding type (ie: minimum, maximum or saddle point)

Solution
First, we compute the first derivatives since the stationary points satisfy $\nabla f(x^*) = 0$

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{x^3-x^2-2x}{1+4y^2} = \frac{x(x-2)(x+1)}{1+4y^2} = 0 \\ \frac{\partial f}{\partial y} &= -\frac{2y\left(3x^4-4x^3-12x^2+18\right)}{3(1+4y^2)^2} = 0 \end{aligned} \tag{6}$$

We find that $\nabla f(x^*) = 0$ when $x = 0, -1, 2$ and $y = 0$. Then, the stationary points are

$$(0,0), (2,0), (-1,0)$$

In order to find their corresponding type we compute the Hessian matrix ($\nabla^2 f(x,y)$)

$$\begin{aligned} \nabla^2 f(x,y) &= \begin{bmatrix} \frac{3x^2-2x-2}{1+4y^2} & -\frac{2y\left(4x^3-4x^2-8x\right)}{(1+4y^2)^2} \\ -\frac{2y\left(4x^3-4x^2-8x\right)}{(1+4y^2)^2} & -\frac{2\left(3x^4-4x^3-12x^2+18\right)\left[\left(1+4y^2\right)^2-16y^2\left(1+4y^2\right)\right]}{3(1+4y^2)^4} \end{bmatrix} \\ &= \begin{bmatrix} \frac{3x^2-2x-2}{1+4y^2} & -\frac{2y\left(4x^3-4x^2-8x\right)}{(1+4y^2)^2} \\ -\frac{2y\left(4x^3-4x^2-8x\right)}{(1+4y^2)^2} & -\frac{2\left(3x^4-4x^3-12x^2+18\right)\left(1-12y^2\right)}{3(1+4y^2)^3} \end{bmatrix} \end{aligned}$$

Now, evaluating the stationary points on the Hessian matrix, looking for the sign of the second derivative with respect to x and the sign of the determinant we get that

$$\frac{\partial^2 f(0,0)}{\partial x^2} = -2 < 0$$
$$det(\nabla^2 f(0,0)) = 24 > 0$$

Then, $(0,0)$ is a maximum.
Now,

$$\frac{\partial^2 f(2,0)}{\partial x^2} = 6 > 0$$
$$det(\nabla^2 f(2,0)) = 56 > 0$$

Then, $(2,0)$ is a minimum.
Finally,

$$\frac{\partial^2 f(-1,0)}{\partial x^2} = 3 > 0$$
$$det(\nabla^2 f(-1,0)) = -26 < 0$$

Then, $(2,0)$ is a saddle point.

4. Compute the gradient $\nabla f(\boldsymbol{x})$ and Hessian $\nabla^2 f(\boldsymbol{x})$ of the Rosenbrock function

$$f(\boldsymbol{x}) = \sum_{i=1}^{N-1} \left[100\left(x_{i+1} - x_i^2\right)^2 + (1 - x_i)^2\right]$$

where $\boldsymbol{x} = [x_1, \ldots, x_N]^T \in \mathbb{R}^N$

Solution

Let's note that

$$
\begin{aligned}
\frac{\partial f(x)}{\partial x_1} &= -400(x_2 - x_1^2) - 2(1 - x_1) \\
\frac{\partial f(x)}{\partial x_2} &= 200(x_2 - x_1^2) - 400(x_3 - x_2^2)x_2 - 2(1 - x_2) \\
&\vdots \\
\frac{\partial f(x)}{\partial x_k} &= 200(x_k - x_{k-1}^2) - 400(x_{k+1} - x_k^2)x_k - 2(1 - x_k) \\
&\vdots \\
\frac{\partial f(x)}{\partial x_{n-1}} &= 200(x_{n-1} - x_{n-2}^2) - 400(x_n - x_{n-1}^2)x_{n-1} - 2(1 - x_{n-1}) \\
\frac{\partial f(x)}{\partial x_n} &= 200(x_n - x_{n-1}^2)
\end{aligned}
$$

Then,

$\nabla f(x) = \left(-400x_1\left(x_2 - x_1^2\right) - 2\left(1 - x_1\right), \cdots, 200\left(x_k - x_{k-1}^2\right) - 400x_k\left(x_{k+1} - x_k^2\right) - 2\left(1 - x_k\right), \cdots, 200\left(x_n - x_{n-1}^2\right)\right)$

Now, for the Hessian matrix we can realize that we get a tridiagonal matrix

$$
\nabla^2 f(x) = \begin{bmatrix}
2 - 400\left(x_2 - 3x_1^2\right) & -400x_1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & & \ddots & & \vdots & \vdots & \vdots & \\
0 & 0 & \cdots & -400x_{k-1} & 202 - 400x_{k+1} + 1200x_k^2 & -400x_k & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & & \ddots & 0 & 0 & \cdots & -400x_{n-1} & 200
\end{bmatrix}
$$

5. Show, without using the optimality conditions, that $f(\boldsymbol{x}) > f(\boldsymbol{x}^*)$ for all $x \neq x^*$ if

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\mathbf{Q}\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x}$$

$\mathbf{Q} = \mathbf{Q}^T \succ 0$ and $\mathbf{Q}x^* = b$

Solution

Since $b^T = x^{*T}Q^T$, *then* $f(x^*) = \frac{1}{2}x^{*T}Qx^* - x^{*T}Q^Tx^* = -\frac{1}{2}x^{*T}Q^Tx^*$, and we want to prove that

$$f(x) > f(x^*) \Rightarrow f(x) - f(x^*) > 0$$

Then,

$$
\begin{aligned}
f(x) - f(x^*) &= \frac{1}{2}x^TQx - b^Tx + \frac{1}{2}x^{*T}Q^Tx^* \\
&= \frac{1}{2}x^TQx - x^{*T}Q^Tx + \frac{1}{2}x^{*T}Q^Tx^* \\
&= \frac{1}{2}x^TQx - \frac{1}{2}x^{*T}Q^Tx - \frac{1}{2}x^{*T}Q^Tx + \frac{1}{2}x^{*T}Q^Tx^* \\
&= \frac{1}{2}\left(x^T - x^{*T}\right)Qx - \frac{1}{2}x^{*T}Q\left(x - x^*\right) \\
&= \frac{1}{2}\left(x - x^*\right)^TQx - \frac{1}{2}x^{*T}Q\left(x - x^*\right)
\end{aligned}
$$

let's define $x = x^* + \delta$, *then we get*

$$
\begin{aligned}
&= \tfrac{1}{2}\delta^TQ\left(x^* + \delta\right) - \tfrac{1}{2}x^{*T}Q\delta \\
&= \tfrac{1}{2}\delta^TQx^* + \tfrac{1}{2}\delta^TQ\delta) - \tfrac{1}{2}x^{*T}Q\delta
\end{aligned}
$$

Since $x^{*T}Q\delta$ *is a real number we can say that* $x^{*T}Q\delta = \left(x^{*T}Q\delta\right)^T = \delta^TQx^*$, *then*

$$f(x) - f(x^*) = \frac{1}{2}\delta^TQ\delta > 0$$

Therefore,

$$f(x) > f(x^*)$$