Assignment 1

Given Information
Detailed proof of the Theorem 2.3.1 in the two-dimensional case:

Theorem 2.3.1: If a sufficiently smooth $u(x, y)$ has a local maximum at (x_0, y_0), then $\nabla u(x_0, y_0)$ and the Hessian Matrix $H(x_0, y_0)$ is negative semi-definite

Definition 1: A single-variable function $f(x)$ has a local maximum at x_0 if there exists a $\delta > 0$ such that $f(x) \leq f(x_0)$ for all x, $|x - x_0| < \delta$

Definition 2: A two-variable function $u(x, y)$ has a local maximum at (x_0, y_0) if there exists a $\delta > 0$ such that $u(x, y) \leq u(x_0, y_0)$ for all $\sqrt{x - x_0{}^2 + y - y_0{}^2} < \delta$

Problem 1
Prove for a single-variable function $f(x)$ that, if f has a local maximum at x_0 and $f\prime(x_0)$ exists, then $f\prime(x_0) = 0$ (Fermat's Theorem).

Solution
We begin with the definition of a derivative of a single variable function $f(x)$ at a point $x = c$:

$$f\prime(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

Consider values higher and lower than c, i.e., c^+ and c^-

$$f\prime(c^+) = \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} = \frac{negative}{positive} \leq 0$$

$$f\prime(c^-) = \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} = \frac{negative}{negative} \geq 0$$

For the above expressions to hold simultaneously,

$$f\prime(c) = 0$$

Conclusion: We analyzed the derivative of $f(x)$ around the local maximum point x_0 to prove that the derivative has to be zero at x_0

Problem 2
Prove that, if a function $f(x)$ has a local maximum at x_0 and can be expanded by the Taylor series around x_0, then $f\prime\prime(x_0) \leq 0$

Solution
Taylor series expansion of $f(x)$ about the local maximum x_0, considering an infinitesimal distance Δx

$$f(x_0 + \Delta x) = f(x_0) + f\prime(x_0) * \Delta x + f\prime\prime(x_0) * \frac{\Delta x^2}{2!} + Higher\ Order\ Terms$$

From Problem 1, we know that $f\prime(x_0) = 0$. Rearranging the terms to express $f\prime\prime(x_0)$

$$f\prime\prime(x_0) = \lim_{x \to c^+} \frac{2 * (f(x_0 + \Delta x) - f(x_0))}{\Delta x^2} = \frac{negative}{positive} \leq 0$$

Conclusion: Based on the fact that the difference between the value of function between any other x and the local maximum point x_0 is negative, we could prove that $f\prime\prime(x_0) \leq 0$

Problem 3
Use Fermat's Theorem to prove that, if a differentiable function $u(x, y)$ has a local maximum at (x_0, y_0), then $\nabla u(x_0, y_0) = 0$. Show that the directional derivative $D_v u(x_0, y_0) = 0$ for any v as well.

Solution
Definition of a gradient of a function $u(x, y)$

$$\nabla u(x, y) = \left\langle \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right\rangle$$

$$\frac{\partial u}{\partial x} = \lim_{h \to 0} \frac{u(x_0 + h, y) - u(x_0, y_0)}{h}$$

$$\frac{\partial u}{\partial y} = \lim_{h \to 0} \frac{u(x_0, y + h) - u(x_0, y_0)}{h}$$

Using a treatment similar to Fermat's theorem, we get,

$$\frac{\partial u}{\partial x} = 0$$

$$\frac{\partial u}{\partial y} = 0$$

And hence, the $\nabla u(x_0, y_0) = 0$

In a two-variable function, the gradient considers the change in both x and y. One may change more or less than the other, and this defines the direction in which the change occurs.

Therefore, we can define a unit vector $\vec{v} =< a, b >$ pointing in the direction of change. Next, we can define a new single-variable function $g(z)$ where

$$g(z) = u(x_0 + az, y_0 + bz)$$

In essence, we are parametrizing the x and y change based on a known point (x_0, y_0) and the newly defined direction $< a, b >$

The derivative of $g(z)$ is defined as,

$$g\prime(z) = \lim_{h \to 0} \frac{g(z + h) - g(z)}{h}$$

$$\frac{\mathrm{d}g}{\mathrm{d}x} = \frac{\partial u}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}z} + \frac{\partial u}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}z} = u_x * a + u_y * b = D_{\vec{v}} u(x, y)$$

Conclusion: The above expression is the definition of a Directional Derivative $D_{\vec{v}} u(x, y)$. From the result $\nabla u(x_0, y_0) = 0$, we can conclude that,

$$D_{\vec{v}} u(x_0, y_0) = 0$$

Problem 4
Let H be the Hessian matrix and $\vec{v} = < p, q >$ a unit vector. Show that $\vec{v}^T H\vec{v}$ equals the second derivative of u in the direction of $\vec{v}$, i.e. $\vec{v}^T H\vec{v} = D_{\vec{v}}(D_{\vec{v}}u)$.

Solution
The Hessian matrix H of $u(x, y)$ is defined as,

$$H = \begin{bmatrix} \frac{\partial^2 u(x,y)}{\partial x^2} & \frac{\partial^2 u(x,y)}{\partial x\partial y} \\ \frac{\partial^2 u(x,y)}{\partial x\partial y} & \frac{\partial^2 u(x,y)}{\partial y^2} \end{bmatrix}$$

$$\vec{v}^T H\vec{v} = \begin{bmatrix} p & q \end{bmatrix} . \begin{bmatrix} \frac{\partial^2 u(x,y)}{\partial x^2} & \frac{\partial^2 u(x,y)}{\partial x\partial y} \\ \frac{\partial^2 u(x,y)}{\partial x\partial y} & \frac{\partial^2 u(x,y)}{\partial y^2} \end{bmatrix} . \begin{bmatrix} p \\ q \end{bmatrix}$$

$$\vec{v}^T H\vec{v} = \begin{bmatrix} \left(\frac{p.\partial^2 u(x,y)}{\partial x^2} + \frac{q.\partial^2 u(x,y)}{\partial x\partial y}\right) & \left(\frac{p.\partial^2 u(x,y)}{\partial x\partial y} + \frac{q.\partial^2 u(x,y)}{\partial y^2}\right) \end{bmatrix} . \begin{bmatrix} p \\ q \end{bmatrix}$$

$$\vec{v}^T H\vec{v} = \frac{p^2.\partial^2 u(x,y)}{\partial x^2} + \frac{2pq.\partial^2 u(x,y)}{\partial x\partial y} + \frac{q^2.\partial^2 u(x,y)}{\partial y^2}$$

Now, we compute $D_{\vec{v}}(D_{\vec{v}}u)$ is a dot product of the directional unit vector and the gradient of u,

$$D_{\vec{v}}(D_{\vec{v}}u) = D_{\vec{v}}(\begin{bmatrix} p & q \end{bmatrix} . \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \end{bmatrix})$$

$$D_{\vec{v}}(D_{\vec{v}}u) = D_{\vec{v}}\left(\frac{p.\partial u}{\partial x} + \frac{q.\partial u}{\partial y}\right)$$

$$D_{\vec{v}}(D_{\vec{v}}u) = \begin{bmatrix} p & q \end{bmatrix} . \begin{bmatrix} \frac{d}{dx}\left(\frac{p.\partial u}{\partial x} + \frac{q.\partial u}{\partial y}\right) & \frac{d}{dy}\left(\frac{p.\partial u}{\partial x} + \frac{q.\partial u}{\partial y}\right) \end{bmatrix}$$

$$D_{\vec{v}}(D_{\vec{v}}u) = \begin{bmatrix} p & q \end{bmatrix} . \begin{bmatrix} p.\frac{\partial^2 u}{\partial x^2} + q.\frac{\partial^2 u}{\partial x\partial y} & p.\frac{\partial^2 u}{\partial x\partial y} + q.\frac{\partial^2 u}{\partial y^2} \end{bmatrix} .$$

$$D_{\vec{v}}(D_{\vec{v}}u) = \frac{p^2.\partial^2 u(x,y)}{\partial x^2} + \frac{2pq.\partial^2 u(x,y)}{\partial x\partial y} + \frac{q^2.\partial^2 u(x,y)}{\partial y^2}$$

Conclusion: Therefore, $\vec{v}^T H\vec{v} = D_{\vec{v}}(D_{\vec{v}}u)$

Problem 5
Use the fact that the intersection of a vertical plane going through the point (x_0, y_0) and the graph of $u(x, y)$ is a function of a single variable to finish the proof of Theorem 2.3.1.

Solution
The general equation of a plane is,

$$ax + by + cz + d = 0$$

For a vertical plane, $z = 0$. The intersection of this plane with the function $u(x, y)$ can be found as a function of a single variable as we can express y in terms of x.

$$ax + by + d = 0$$

$$y = \frac{-ax - d}{b} = f(x)$$

The function $u(x, y)$ can now be written as

$$u(x, f(x))$$

The Hessian matrix of the function u will be

$$H = \begin{bmatrix} \frac{\partial^2 u(x,y)}{\partial x^2} & \frac{\partial^2 u(x,y)}{\partial x \partial y} \\ \frac{\partial^2 u(x,y)}{\partial x \partial y} & \frac{\partial^2 u(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 u(x,f(x))}{\partial x^2} & 0 \\ 0 & 0 \end{bmatrix}$$

Deducing the eigenvalue of the matrix H

$$det\,[H - \lambda I] = 0$$

$$\begin{vmatrix} \frac{\partial^2 u(x,f(x))}{\partial x^2} - \lambda & 0 \\ 0 & -\lambda \end{vmatrix} = -\lambda.\left[\frac{\partial^2 u(x, f(x))}{\partial x^2} - \lambda\right]$$

$$\frac{\partial^2 u(x, f(x))}{\partial x^2} = \lambda$$

By the second derivative test we performed in Problem 2, we know that $\frac{\partial^2 u(x,f(x))}{\partial x^2} \leq 0$

Conclusion: Therefore, the eigenvalue $\lambda \leq 0$, which is the condition for the Hessian matrix to be negative semi-definite.