Numerical Analysis [WI4014TU]
Assignment 1

October 2, 2024

Chirag Bansal
Student number: 6137776
MS Aerospace Engineering

---

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

Solution: According to Fermat's Theorem, if f has a local maximum at c, and if $f'(c)$ exists, then $f'(c) = 0$.

To prove: $f'(c) = 0$

Let $f(x)$ be a single-variable function having a local maximum at point $x = c$. Therefore, in the range where c exists, $f(c) \geq f(x)$.

Let h be any positive or negative integer with $h \to 0$ such that

$$f(c) \geq f(c+h)$$

$$\implies f(c) - f(c+h) \geq 0 \tag{1}$$

Since h can be either positive or negative. We have two cases at hand — $h > 0$ or $h < 0$

For $h > 0$: Dividing (1) by h

$$\frac{f(c+h) - f(c)}{h} \leq 0$$

Taking limits from the right-hand side

$$\lim_{h\to 0^+} \frac{f(c+h)-f(c)}{h} \leq \lim_{h\to 0^+} 0 = 0$$

Since $f'(c)$ exists

$$\therefore f'(c) = \lim_{h\to 0} \frac{f(c+h)-f(c)}{h} = \lim_{h\to 0^+} \frac{f(c+h)-f(c)}{h} \leq 0$$

Hence, $f'(c) \leq 0$

Similarly, for $h < 0$: Dividing (1) by h [Note: here $h < 0$]

$$\frac{f(c+h)-f(c)}{h} \geq 0$$

Taking limits from the left-hand side

$$\lim_{h\to 0^-} \frac{f(c+h)-f(c)}{h} \geq \lim_{h\to 0^-} 0 = 0$$

Since $f'(c)$ exists

$$\therefore f'(c) = \lim_{h\to 0} \frac{f(c+h)-f(c)}{h} = \lim_{h\to 0^-} \frac{f(c+h)-f(c)}{h} \geq 0$$

Hence, $f'(c) \geq 0$

Since, $f'(c) \leq 0$ and $f'(c) \geq 0$, Therefore, $f'(c) = 0$.

2. Prove that, if a function $f(x)$ has a local maximum at x_0 and can be expanded in the Taylor series around x_0, then $f''(x_0) \leq 0$ (*Converse of the Second Derivative Test*).

Solution: Given that $f(x)$ has a local maximum at x_0, then $f'(x_0) = 0$ in a range where $|x - x_0| < \delta$.
Since, $f(x)$ can be expanded in a Taylor series,
Taking the forward Taylor series expansion

$$\implies f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \frac{f'''(x_0)}{3!}(x - x_0)^3 + \cdots \quad (2)$$

Similarly, taking the backward Taylor series expansion

$$\implies f(x) = f(x_0) - f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 - \frac{f'''(x_0)}{3!}(x - x_0)^3 + \cdots \quad (3)$$

Adding (2) and (3), we get

$$2f(x) = 2f(x_0) + 2\frac{f''(x_0)}{2!}(x - x_0)^2 + 2\frac{f''''(x_0)}{2!}(x - x_0)^4 + \cdots$$

$$\implies f(x) - f(x_0) - \frac{f''''(x_0)}{4!}(x - x_0)^4 - \cdots = \frac{f''(x_0)}{2!}(x - x_0)^2$$

Since x_0 is the local maximum, then $f(x) \leq f(x_0)$. Therefore all the terms in the LHS of the equation will be ≤ 0

$$\implies 0 \geq f(x) - f(x_0) - \frac{f''''(x_0)}{4!}(x - x_0)^4 - \cdots = \frac{f''(x_0)}{2!}(x - x_0)^2$$

$$\implies \frac{f''(x_0)}{2!}(x - x_0)^2 \leq 0$$

$$\therefore f''(x_0) \leq 0$$

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 $\nabla u(x_0, y_0) = \mathbf{0}$. Show that the directional derivative $D_{\mathbf{v}} u(x_0, y_0) = 0$ for any $\mathbf{v}$ as well.

Solution: Given: $u(x, y)$ is differentiable and has a local maximum at (x_0, y_0), Therefore, there exists a neighbourhood, N, where $u(x, y) \leq u(x_0, y_0)$, for all $(x, y) \in N \subset \mathbb{R}^2$

Since Fermat's Theorem is defined for single variable functions, we cannot use it directly to prove $\nabla u(x_0, y_0) = \mathbf{0}$. Therefore, we need to approach it in a different way.

We know that:

$$\nabla u(x_0, y_0) = \begin{bmatrix} \frac{\partial}{\partial x} u(x_0, y_0) \\ \frac{\partial}{\partial y} u(x_0, y_0) \end{bmatrix} \tag{4}$$

Therefore, if we can prove the individual elements to be equal to 0, then inherently we will prove that

$$\nabla u(x_0, y_0) = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Assume two differentiable, single-variable functions $f(x)$ and $g(y)$ defined as follows:

$$f(x) = u(x, y_0)$$

This makes $f(x)$ a single line function defined over the surface of $u(x, y)$ where $y = y_0$

By similar logic, we can define $g(y)$ as

$$g(y) = u(x_0, y)$$

Now, for $f(x)$:
Since, x_0 is the local maximum for $u(x, y)$ at $y = y_0$, and $f(x)$ is defined for a constant value of $y = y_0$, then x_0 is the local maximum for $f(x)$ as well. Therefore, applying Fermat's Theorem to $f(x)$ at point x_0

$$\implies \frac{d}{dx} f(x_0) = f'(x_0) = 0$$

$$\implies \frac{\partial}{\partial x} u(x_0, y_0) = 0 \tag{5}$$

Similarly, for $g(y)$:
Since, y_0 is the local maximum for $u(x, y)$ at $x = x_0$, and $g(y)$ is defined for a constant value of $x = x_0$, then y_0 is the local maximum for $g(y)$ as well. Therefore, applying Fermat's Theorem to $g(y)$ at point y_0

$$\implies \frac{d}{dy} g(y_0) = g'(y_0) = 0$$

$$\implies \frac{\partial}{\partial y} u(x_0, y_0) = 0 \tag{6}$$

Plugging (5), and (6) in (4), we get

$$\nabla u(x_0, y_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0}$$

Additionally, from the definition of directional derivative, for any direction vector, $\mathbf{v}$, we know that

$$D_{\mathbf{v}} u = \nabla u \cdot \mathbf{v} = |\nabla u||\mathbf{v}| \cos \theta$$

Since, at point (x_0, y_0), we proved that $\nabla u(x_0, y_0) = \mathbf{0}$. Therefore $|\nabla u| = 0$ at (x_0, y_0).

$$\therefore D_{\mathbf{v}} u = |\nabla u||\mathbf{v}| \cos \theta = |0||\mathbf{v}| \cos \theta = 0$$

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

Solution: To prove:

$$\mathbf{v}^T H \mathbf{v} = D_{\mathbf{v}}(D_{\mathbf{v}}u) \tag{7}$$

Let $\mathbf{v}$ be a column matrix, such that

$$\mathbf{v} = \begin{bmatrix} p \\ q \end{bmatrix}$$

Taking LHS of equation (7):

$$\mathbf{v}^T H \mathbf{v} = \begin{bmatrix} p & q \end{bmatrix} \begin{bmatrix} u_{xx} & u_{xy} \\ u_{xy} & u_{yy} \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix}$$

$$\implies \mathbf{v}^T H \mathbf{v} = \begin{bmatrix} pu_{xx} + qu_{xy} & pu_{xy} + qu_{yy} \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix}$$

$$\implies \mathbf{v}^T H \mathbf{v} = \begin{bmatrix} p^2 u_{xx} + 2pqu_{xy} + q^2 u_{yy} \end{bmatrix} \tag{8}$$

Now, taking RHS of equation (7) and using the definition of directional derivative:

$$D_{\mathbf{v}}u = \nabla u \cdot \mathbf{v}$$

we get

$$D_{\mathbf{v}}u = p\frac{\partial u}{\partial x} + q\frac{\partial u}{\partial y}$$

This quantity is a scalar quantity. Now,

$$D_{\mathbf{v}}(D_{\mathbf{v}}u) = D_{\mathbf{v}}(p\frac{\partial u}{\partial x} + q\frac{\partial u}{\partial y})$$

$$\implies D_{\mathbf{v}}(D_{\mathbf{v}}u) = \nabla(p\frac{\partial u}{\partial x} + q\frac{\partial u}{\partial y}) \cdot \mathbf{v}$$

$$\implies D_{\mathbf{v}}(D_{\mathbf{v}}u) = \langle \frac{\partial}{\partial x}(p\frac{\partial u}{\partial x} + q\frac{\partial u}{\partial y}), \frac{\partial}{\partial y}(p\frac{\partial u}{\partial x} + q\frac{\partial u}{\partial y})\rangle \cdot \langle p, q \rangle$$

$$\implies D_{\mathbf{v}}(D_{\mathbf{v}}u) = \langle pu_{xx} + qu_{xy}, pu_{xy} + qu_{yy} \rangle \cdot \langle p, q \rangle$$

$$\implies D_{\mathbf{v}}(D_{\mathbf{v}}u) = p^2 u_{xx} + 2pq u_{xy} + q^2 u_{yy} \tag{9}$$

Since (8) = (9), Therefore, we have proved that $\mathbf{v}^T H \mathbf{v} = D_{\mathbf{v}}(D_{\mathbf{v}}u)$.

5. Use the fact that the intersection of any 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:
Theorem 2.3.1: If a sufficiently smooth $u(x, y)$ has a local maximum at (x_0, y_0), then $\nabla u(x_0, y_0) = 0$ and the Hessian matrix $H(x_0, y_0)$ is negative semi-definite.

Consider a vertical plane passing through the point (x_0, y_0) in any arbitrary direction given by $\mathbf{v} = \langle p, q \rangle$ where p and q are some constant values. Then, the intersection of this plane with the graph of $u(x, y)$ will yield a single variable function $\tilde{u}$ which, say, is a function of a variable t given by

$$\tilde{u}(t) = u(x_0 + tp, y_0 + tq)$$

This new function $\tilde{u}$ gives the values of the function $u(x, y)$ along a line in the direction of $\mathbf{v}$, and the spatial coordinates are given by $x = x_0 + tp$ and $y = y_0 + tq$
Now, the first derivative of $\tilde{u}(t)$:

$$\frac{d\tilde{u}}{dt} = \frac{d}{dt}u(x_0 + tp, y_0 + tq)$$

Using chain rule (since u is not directly dependent on t)

$$\frac{d\tilde{u}}{dt} = \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt}$$

$$\implies \frac{d\tilde{u}}{dt} = u_x p + u_y q$$

Since, from equations (4), (5), (6) [see question 3], we know that:

$$\nabla u(x_0, y_0) = \begin{bmatrix} \frac{\partial}{\partial x}u(x_0, y_0) \\ \frac{\partial}{\partial y}u(x_0, y_0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0}$$

At, $t = 0$, we have $x = x_0$ and $y = y_0$ which are the local maximum coordinates for u, and inherently for $\tilde{u}$ as well (since $\tilde{u}$ exists along the line at the intersection of the vertical plane and the graph of u)

$$\implies \frac{d}{dt}\tilde{u}(0) = u_x p + u_y q = 0$$

Therefore, $\tilde{u}$ has a critical point at $t = 0$.
Now, assuming that a local maximum exists at the location $t = 0$ for the function $\tilde{u}$. Therefore it's second derivative must be ≤ 0

$$\implies \frac{d^2\tilde{u}}{dt^2} \leq 0$$

$$\implies \frac{d}{dt}(u_x p + u_y q) \leq 0$$

$$\implies \frac{d}{dt}(u_x p + u_y q) = p\frac{du_x}{dt} + q\frac{du_y}{dt} \leq 0$$

Applying the chain rule to this equation, we get

$$\implies \frac{d}{dt}(u_x p + u_y q) = p(\frac{\partial u_x}{\partial x}\frac{dx}{dt} + \frac{\partial u_x}{\partial y}\frac{dy}{dt}) + q(\frac{\partial u_y}{\partial x}\frac{dx}{dt} + \frac{\partial u_y}{\partial y}\frac{dy}{dt}) \leq 0$$

$$\implies \frac{d}{dt}(u_x p + u_y q) = p(u_{xx}p + u_{xy}q) + q(u_{xy}p + u_{yy}q) \leq 0$$

$$\implies \frac{d}{dt}(u_x p + u_y q) = p^2 u_{xx} + 2pq u_{xy} + q^2 u_{yy} \leq 0$$

If we observe, this equation resembles (8) [see question 4], and therefore

$$\implies \frac{d^2\tilde{u}}{dt^2} = \mathbf{v}^T H(x, y)\mathbf{v} \leq 0$$

Since, t is a critical point and $(x, y) = (x_0, y_0)$ at $t = 0$

$$\implies \frac{d^2}{dt^2}\tilde{u}(0) = \mathbf{v}^T H(x_0, y_0)\mathbf{v} \leq 0$$

The result $\mathbf{v}^T H(x_0, y_0)\mathbf{v} \leq 0$ for any directional vector $\mathbf{v}$ means that the Hessian matrix satisfies the condition for it to be negative semi-definite.