

MA15010H: Multi-variable Calculus

(Lecture note 3: Directional derivatives and differentiability)

July - November, 2025

Directional derivative: Let $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ and let $\mathbf{x}_0 \in S^0$. If $\mathbf{u} \in \mathbb{R}^m$ with $\|\mathbf{u}\| = 1$, then the directional derivative of f in the direction of $\mathbf{u}$ at $\mathbf{x}_0$ is defined as

$$D_{\mathbf{u}}f(\mathbf{x}_0) = f'_{\mathbf{u}}(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{u}) - f(\mathbf{x}_0)}{t},$$

provided this limit exists (in $\mathbb{R}$). Thus if $r > 0$ is such that $B_r(\mathbf{x}_0) \subseteq S$ and if $\varphi(t) = f(\mathbf{x}_0 + t\mathbf{u})$ for all $t \in (-r, r)$, then

$$D_{\mathbf{u}}f(\mathbf{x}_0) = \varphi'(0).$$

We note that $D_{-\mathbf{u}}f(\mathbf{x}_0) = -D_{\mathbf{u}}f(\mathbf{x}_0)$ and if $\mathbf{u} = \mathbf{e}_j$ for some $j \in \{1, \dots, m\}$, then

$$D_{\mathbf{u}}f(\mathbf{x}_0) = f_{x_j}(\mathbf{x}_0).$$

Example 0.1. (a) Let $f(x, y) = x^2 + xy + 2y$ for all $(x, y) \in \mathbb{R}^2$. Then

$$f_x(x_0, y_0) = 2x_0 + y_0 \quad \text{and} \quad f_y(x_0, y_0) = x_0 + 2$$

for all $(x_0, y_0) \in \mathbb{R}^2$. If $(x_0, y_0) \in \mathbb{R}^2$ and $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$ with $\|\mathbf{u}\| = 1$, then

$$D_{\mathbf{u}}f(x_0, y_0) = (2x_0 + y_0)u_1 + (x_0 + 2)u_2.$$

(b) Let $f(x, y) = \sqrt{x^2 + y^2}$ for all $(x, y) \in \mathbb{R}^2$. Let $(x_0, y_0) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ and let $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$ with $\|\mathbf{u}\| = 1$. Then

$$f_x(x_0, y_0) = \frac{x_0}{\sqrt{x_0^2 + y_0^2}}, \quad f_y(x_0, y_0) = \frac{y_0}{\sqrt{x_0^2 + y_0^2}} \quad \text{and} \quad D_{\mathbf{u}}f(x_0, y_0) = \frac{x_0 u_1 + y_0 u_2}{\sqrt{x_0^2 + y_0^2}}.$$

However, $f_x(0, 0)$ and $f_y(0, 0)$ do not exist (in $\mathbb{R}$). In fact, $D_{\mathbf{u}}f(0, 0)$ does not exist (in $\mathbb{R}$). We note that f is continuous at $(0, 0)$. Thus the continuity of a function $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ at a point $\mathbf{x}_0 \in S^0$ does not guarantee the existence of the partial derivatives or the directional derivatives of f at $\mathbf{x}_0$.

(c) Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be defined by

$$f(x, y) = \begin{cases} 0 & \text{if } xy = 0, \\ 1 & \text{if } xy \neq 0. \end{cases}$$

Then $f_x(0, 0) = f_y(0, 0) = 0$, but if $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$ with $\|\mathbf{u}\| = 1$ and $u_1 u_2 \neq 0$, then $D_{\mathbf{u}}f(0, 0)$ does not exist. Thus the existence of all the partial derivatives of a function $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ at a point $\mathbf{x}_0 \in S^0$ does not guarantee the existence of the other directional derivatives of f at $\mathbf{x}_0$.

(d) Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be defined by

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

If $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$ with $\|\mathbf{u}\| = 1$, then

$$D_{\mathbf{u}}f(0,0) = \begin{cases} \frac{u_1^2}{u_2} & \text{if } u_2 \neq 0, \\ 0 & \text{if } u_2 = 0. \end{cases}$$

So all the directional derivatives of f at $(0,0)$ exist but we know that f is not continuous at $(0,0)$. Thus the existence of all the directional derivatives of a function $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ at a point $\mathbf{x}_0 \in S^0$ does not guarantee the continuity of f at $\mathbf{x}_0$.

Higher order partial derivatives: Let the first order partial derivative f_x of $f : S \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ exist at every point of S^0 . Then the partial derivatives of the function $f_x : S^0 \rightarrow \mathbb{R}$ with respect to x and y at $(x_0, y_0) \in S^0$ are denoted respectively by $f_{xx}(x_0, y_0)$ or $\frac{\partial^2 f}{\partial x^2}(x_0, y_0)$ and $f_{xy}(x_0, y_0)$ or $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$, provided these exist (in $\mathbb{R}$).

Similarly, we define $f_{yy}(x_0, y_0)$ or $\frac{\partial^2 f}{\partial y^2}(x_0, y_0)$ and $f_{yx}(x_0, y_0)$ or $\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$.

In general, the mixed partial derivatives $f_{xy}(x_0, y_0)$ and $f_{yx}(x_0, y_0)$ need not be equal. Further, the second and higher order partial derivatives can be defined in an analogous way for a general function $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$.

Example 0.2. Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be defined by

$$f(x, y) = \begin{cases} xy \frac{y^2 - x^2}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

Then $f_{xy}(0, 0) = -1 \neq 1 = f_{yx}(0, 0)$.

Equality of mixed partial derivatives: Let $f : S \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ and let $(x_0, y_0) \in S^0$. Let f_{xy}, f_{yx} exist on $B_\delta((x_0, y_0)) \subseteq S$ for some $\delta > 0$ and let f_{xy}, f_{yx} be continuous at (x_0, y_0) . Then

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0).$$

Proof. Let $(h, k) \in B_\delta((0, 0))$ so that $(x_0 + h, y_0 + k) \in B_\delta((x_0, y_0))$ and let

$$\Delta = f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - f(x_0, y_0 + k) + f(x_0, y_0).$$

Also, let $\varphi(x) = f(x, y_0 + k) - f(x, y_0)$ for all $x \in [x_0, x_0 + h]$ and $\psi(y) = f(x_0 + h, y) - f(x_0, y)$ for all $y \in [y_0, y_0 + k]$. Then

$$\Delta = \varphi(x_0 + h) - \varphi(x_0) = h\varphi'(x_0 + \theta_1 h)$$

for some $\theta_1 \in (0, 1)$ by Lagrange's MVT of one real variable. Hence

$$\Delta = h[f_y(x_0 + \theta_1 h, y_0 + k) - f_y(x_0 + \theta_1 h, y_0)] = hk f_{xy}(x_0 + \theta_1 h, y_0 + \theta_2 k)$$

for some $\theta_2 \in (0, 1)$ by Lagrange's MVT of one real variable. Again,

$$\Delta = \psi(y_0 + k) - \psi(y_0) = k\psi'(y_0 + \eta_2 k) \text{ for some } \eta_2 \in (0, 1)$$

by Lagrange's MVT of one real variable. Hence

$$\Delta = k[f_y(x_0 + h, y_0 + \eta_2 k) - f_y(x_0, y_0 + \eta_2 k)] = kh f_{yx}(x_0 + \eta_1 h, y_0 + \eta_2 k) \text{ for some } \eta_1 \in (0, 1)$$

by Lagrange's MVT of one real variable. Thus

$$f_{xy}(x_0 + \theta_1 h, y_0 + \theta_2 k) = f_{yx}(x_0 + \eta_1 h, y_0 + \eta_2 k) = \frac{\Delta}{hk}$$

for all $(h, k) \in B_\delta((0, 0))$ with $h \neq 0, k \neq 0$. Since f_{xy} and f_{yx} are continuous at (x_0, y_0) , taking limits as $(h, k) \rightarrow (0, 0)$, we obtain

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0).$$

□

Motivation for the definition of differentiability: If $g : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$, then we know that g is called differentiable at $x_0 \in A^0$ if there exists $\alpha \in \mathbb{R}$ such that

$$\lim_{x \rightarrow x_0} \frac{g(x) - g(x_0) - \alpha(x - x_0)}{x - x_0} = 0.$$

Now, let $f : S \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ and let $(x_0, y_0) \in S^0$. A similar definition of differentiability of f at (x_0, y_0) is not meaningful as $\frac{f(x, y) - f(x_0, y_0)}{(x, y) - (x_0, y_0)}$, where $(x, y) \in S \setminus \{(x_0, y_0)\}$, is not defined. However, we can rewrite the above definition of differentiability of g at x_0 in the following equivalent way:

The function g is differentiable at x_0 if there exists an $\alpha \in \mathbb{R}$ such that

$$\lim_{x \rightarrow x_0} \frac{|g(x) - g(x_0) - \alpha(x - x_0)|}{|x - x_0|} = 0.$$

We may interpret this geometrically as follows:

For each $\alpha \in \mathbb{R}$, if $L_\alpha : \mathbb{R} \rightarrow \mathbb{R}$ is defined by $L_\alpha(x) = g(x_0) + \alpha(x - x_0)$ for all $x \in \mathbb{R}$, then $y = L_\alpha(x)$ represents a straight line in $\mathbb{R}^2$ passing through the point $(x_0, g(x_0))$. If g is assumed to be only continuous at x_0 , then every such line approximates the curve $y = g(x)$ around x_0 in the sense that

$$\lim_{x \rightarrow x_0} (g(x) - L_\alpha(x)) = 0$$

for each $\alpha \in \mathbb{R}$.

The differentiability of g at x_0 is equivalent to the existence of an $\alpha \in \mathbb{R}$ (i.e., the existence of a straight line $y = L_\alpha(x)$) such that among all the straight lines in $\mathbb{R}^2$ passing through $(x_0, g(x_0))$, the straight line $y = L_\alpha(x)$ approximates the curve $y = g(x)$ around x_0 in the 'best' possible way in the sense that

$$\lim_{x \rightarrow x_0} \frac{|g(x) - L_\alpha(x)|}{|x - x_0|} = 0.$$

Also, such an α and hence such a straight line, when it exists, is unique.

Differentiability and derivative: A function $f : S \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ is said to be differentiable at $(x_0, y_0) \in S^0$ if there exist $\alpha, \beta \in \mathbb{R}$ such that

$$\lim_{(x, y) \rightarrow (x_0, y_0)} \frac{|f(x, y) - f(x_0, y_0) - \alpha(x - x_0) - \beta(y - y_0)|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0,$$

or equivalently,

$$\lim_{(h,k) \rightarrow (0,0)} \frac{|f(x_0 + h, y_0 + k) - f(x_0, y_0) - \alpha h - \beta k|}{\sqrt{h^2 + k^2}} = 0.$$

Geometrically speaking, among all the planes in $\mathbb{R}^3$ passing through the point $(x_0, y_0, f(x_0, y_0))$, there exists a plane

$$z = f(x_0, y_0) + \alpha(x - x_0) + \beta(y - y_0)$$

which approximates the surface $z = f(x, y)$ around (x_0, y_0) in the ‘best’ possible way in the sense given above.

Although this geometrical interpretation may no longer be possible while defining differentiability in a similar manner for a function like $f : S \subseteq \mathbb{R}^5 \rightarrow \mathbb{R}$, the above definition has another equivalent interpretation in terms of linearization (given at the end of these notes), which is valid and useful for functions $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$.

Thus more generally, $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ is said to be differentiable at $x_0 \in S^0$ if there exists $\alpha \in \mathbb{R}^m$ such that

$$\lim_{h \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) - \alpha \cdot h|}{\|h\|} = 0.$$

Such an $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$, when exists, is unique and in that case we define the (total) derivative of f at x_0 to be $f'(x_0) = Df(x_0) = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_m]$, which is a $1 \times m$ matrix.

If S is a nonempty open subset of $\mathbb{R}^m$, then $f : S \rightarrow \mathbb{R}$ is said to be differentiable on (S) if f is differentiable at each $x_0 \in S$.

Example 0.3. If $f(x, y) = 2x^2 + y^3$ for all $(x, y) \in \mathbb{R}^2$, then $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is differentiable at $(1, 1)$ and

$$f'(1, 1) = [4 \quad 3].$$

Two questions: Let $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ and let $x_0 \in S^0$.

- (a) How do we examine the differentiability of f at x_0 ?
- (b) Given that f is differentiable at x_0 , how do we find $f'(x_0)$?

Although the definitions of differentiability and derivative can be used to answer these questions directly (as we have seen in the example given above), they are difficult to apply in most cases. Hence we provide alternative answers below.

Answer to question (b): If $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ is differentiable at $x_0 \in S^0$, then for each $j \in \{1, \dots, m\}$, $\frac{\partial f}{\partial x_j} \Big|_{x_0}$ exists (in $\mathbb{R}$) and

$$f'(x_0) = \left[\frac{\partial f}{\partial x_1} \Big|_{x_0} \quad \dots \quad \frac{\partial f}{\partial x_m} \Big|_{x_0} \right].$$

Proof. Since f is differentiable at x_0 , there exists $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$ such that

$$\lim_{h \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) - \alpha \cdot h|}{\|h\|} = 0.$$

If $j \in \{1, \dots, m\}$, then from above, we get

$$\lim_{t \rightarrow 0} \frac{|f(x_0 + te_j) - f(x_0) - \alpha_j t|}{|t|} = 0,$$

which gives

$$\lim_{t \rightarrow 0} \frac{f(x_0 + te_j) - f(x_0)}{t} - \alpha_j = 0.$$

Hence,

$$\left. \frac{\partial f}{\partial x_j} \right|_{x_0} = \lim_{t \rightarrow 0} \frac{f(x_0 + te_j) - f(x_0)}{t} = \alpha_j,$$

and consequently

$$f'(x_0) = \begin{bmatrix} \left. \frac{\partial f}{\partial x_1} \right|_{x_0} & \cdots & \left. \frac{\partial f}{\partial x_m} \right|_{x_0} \end{bmatrix}.$$

□

Answer to question (a):

The answer to question (b) given above provides a partial answer to question (a) as well because it says that if at least one of the partial derivatives of $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ at $x_0 \in S^0$ does not exist (in $\mathbb{R}$), then f is not differentiable at x_0 .

For example, if $f(x, y) = |x| + |y|$ for all $(x, y) \in \mathbb{R}^2$, then $f_x(0, 0)$ does not exist (in $\mathbb{R}$) and hence f is not differentiable at $(0, 0)$.

However, as we shall see below, f need not be differentiable at x_0 even if all the partial derivatives of f at x_0 exist (in $\mathbb{R}$).

The next four theorems contain further results which serve to answer question (a).

Theorem 0.4. *If $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ is differentiable at $x_0 \in S^0$, then f is continuous at x_0 .*

Proof. Since f is differentiable at x_0 , there exists $\alpha \in \mathbb{R}^m$ such that

$$\lim_{h \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) - \alpha \cdot h|}{\|h\|} = 0.$$

Hence

$$\lim_{h \rightarrow 0} |f(x_0 + h) - f(x_0)| = \lim_{h \rightarrow 0} |f(x_0 + h) - f(x_0) - \alpha \cdot h + \alpha \cdot h| \leq \lim_{h \rightarrow 0} |f(x_0 + h) - f(x_0) - \alpha \cdot h| + |\alpha \cdot h|.$$

Since $|\alpha \cdot h| \leq \|\alpha\| \|h\| \rightarrow 0$ as $h \rightarrow 0$, it follows that $\lim_{h \rightarrow 0} f(x_0 + h) = f(x_0)$. Therefore f is continuous at x_0 . □

Example 0.5. The function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, defined by

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0), \end{cases}$$

is not continuous at $(0, 0)$ and hence it is not differentiable at $(0, 0)$. However, we note that both the partial derivatives $f_x(0, 0)$ and $f_y(0, 0)$ exist (in $\mathbb{R}$)

Remark 0.6. The converse of the above theorem is not true, in general. For example, if $f(x, y) = |x| + |y|$ for all $(x, y) \in \mathbb{R}^2$, then $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is continuous at $(0, 0)$ but not differentiable at $(0, 0)$.

We need the following definition for the next theorem.

Gradient: If $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ and $x_0 \in S^0$, then the gradient (vector) of f at x_0 is defined as

$$\nabla f(x_0) = \text{grad } f(x_0) = \left(\left. \frac{\partial f}{\partial x_1} \right|_{x_0}, \dots, \left. \frac{\partial f}{\partial x_m} \right|_{x_0} \right),$$

provided $\left. \frac{\partial f}{\partial x_j} \right|_{x_0}$ exists (in $\mathbb{R}$) for each $j \in \{1, \dots, m\}$. For example, if $f(x, y, z) = x^3y + 4xyz + 3$ for all $(x, y, z) \in \mathbb{R}^3$, then

$$\nabla f(1, 1, 0) = (3, 1, 4).$$

Theorem 0.7. If $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ is differentiable at $x_0 \in S^0$, then for each $u \in \mathbb{R}^m$ with $\|u\| = 1$,

$$D_u f(x_0) \text{ exists (in } \mathbb{R}) \text{ and } D_u f(x_0) = \nabla f(x_0) \cdot u.$$

Proof. This proof is exactly similar to the proof of the existence of partial derivatives, which is proved above. Here we need to consider $h = tu$, where $u \in \mathbb{R}^m$ with $\|u\| = 1$. $\square$

Remark 0.8. (a) The above theorem implies that if at least one of the directional derivatives of $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ at $x_0 \in S^0$ does not exist (in $\mathbb{R}$), then f cannot be differentiable at x_0 . However, the existence (in $\mathbb{R}$) of all the directional derivatives of f at x_0 does not ensure the differentiability of f at x_0 .

For example, consider the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, defined by

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

We have seen earlier that all the directional derivatives of f at $(0, 0)$ exist (in $\mathbb{R}$) but since f is not continuous at $(0, 0)$, f is not differentiable at $(0, 0)$.

(b) If $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ is differentiable at $x_0 \in S^0$, then the above theorem provides a way to calculate the directional derivative of f at x_0 if all the partial derivatives of f at x_0 are known.

For example, if

$$f(x, y, z) = 4x^2z + 3y^3z + 7y^2$$

for all $(x, y, z) \in \mathbb{R}^3$ and

$$u = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right),$$

then using the fact that f is differentiable at $x_0 = (1, 1, 1)$ (which will follow from the next theorem), we find that

$$D_u f(x_0) = (f_x(x_0), f_y(x_0), f_z(x_0)) \cdot u = (4, 23, 5) \cdot u = \frac{32}{\sqrt{3}}.$$

- (c) The above formula for calculating $D_u f(x_0)$ is not correct, in general, if $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ is not differentiable at $x_0 \in S^0$ even if $\nabla f(x_0)$ is defined, i.e., even if all the partial derivatives of f at x_0 exist (in $\mathbb{R}$).

For example, consider the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, defined by

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

We have seen earlier that $f_x(0, 0) = 0 = f_y(0, 0)$ and if

$$u = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right),$$

then

$$D_u f(0, 0) = \frac{1}{\sqrt{2}} \neq 0 = \nabla f(0, 0) \cdot u.$$

Theorem 0.9. *Let $f : S \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ and let $x_0 \in S^0$. If there exists $\delta > 0$ such that for each $j \in \{1, \dots, m\}$, $\frac{\partial f}{\partial x_j}$ exists on $B_\delta(x_0) \subseteq S$ and is continuous at x_0 , then f is differentiable at x_0 .*

Proof. We prove this result for $m = 2$. The general case can be proved similarly.

Let $x_0 = (x_0, y_0)$. For all $(h, k) \in B_\delta((0, 0))$, we have

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = f(x_0 + h, y_0 + k) - f(x_0, y_0 + k) + f(x_0, y_0 + k) - f(x_0, y_0).$$

Now, by the Lagrange's MVT of one real variable, there exist $\theta_1, \theta_2 \in (0, 1)$ such that

$$f(x_0 + h, y_0 + k) - f(x_0, y_0 + k) = h f_x(x_0 + \theta_1 h, y_0 + k) \quad \text{and} \quad f(x_0, y_0 + k) - f(x_0, y_0) = k f_y(x_0, y_0 + \theta_2 k).$$

Hence for all $(h, k) \in B_\delta((0, 0)) \setminus \{(0, 0)\}$, we have

$$\begin{aligned} & \frac{|f(x_0 + h, y_0 + k) - f(x_0, y_0) - h f_x(x_0, y_0) - k f_y(x_0, y_0)|}{\sqrt{h^2 + k^2}} \\ & \leq \frac{|h|}{\sqrt{h^2 + k^2}} |f_x(x_0 + \theta_1 h, y_0 + k) - f_x(x_0, y_0)| + \frac{|k|}{\sqrt{h^2 + k^2}} |f_y(x_0, y_0 + \theta_2 k) - f_y(x_0, y_0)|. \end{aligned}$$

Since f_x and f_y are continuous at (x_0, y_0) ,

$$\lim_{(h, k) \rightarrow (0, 0)} (|f_x(x_0 + \theta_1 h, y_0 + k) - f_x(x_0, y_0)| + |f_y(x_0, y_0 + \theta_2 k) - f_y(x_0, y_0)|) = 0$$

and consequently

$$\lim_{(h, k) \rightarrow (0, 0)} \frac{f(x_0 + h, y_0 + k) - f(x_0, y_0) - h f_x(x_0, y_0) - k f_y(x_0, y_0)}{\sqrt{h^2 + k^2}} = 0.$$

Therefore f is differentiable at (x_0, y_0) . □

Example 0.10. If $f(x, y) = e^{xy} + x^2 \sin y + 4x$ for all $(x, y) \in \mathbb{R}^2$, then

$$f_x(x, y) = y e^{xy} + 2x \sin y + 4$$

and

$$f_y(x, y) = x e^{xy} + x^2 \cos y$$

for all $(x, y) \in \mathbb{R}^2$. Since $f_x, f_y : \mathbb{R}^2 \rightarrow \mathbb{R}$ are continuous at $(0, 0)$, $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is differentiable at $(0, 0)$ and $f'(0, 0) = \begin{bmatrix} 4 & 0 \end{bmatrix}$.

Remark 0.11. (a) $f : S \subset \mathbb{R}^m \rightarrow \mathbb{R}$ can be differentiable at $x_0 \in S^0$ even though none of the partial derivatives continuous at x_0 .

For example, the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, defined by

$$f(x, y) = \begin{cases} \frac{(x^2 + y^2) \sin\left(\frac{1}{x^2 + y^2}\right)}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0), \end{cases}$$

is differentiable at $(0, 0)$ (using the next theorem) although neither $f_x : \mathbb{R}^2 \rightarrow \mathbb{R}$ nor $f_y : \mathbb{R}^2 \rightarrow \mathbb{R}$ is continuous at $(0, 0)$.

(b) If S is a nonempty open subset of $\mathbb{R}^m$, then a function $f : S \subset \mathbb{R}^m \rightarrow \mathbb{R}$ is said to be continuously differentiable (on S) if all the partial derivatives of f are continuous on S .

Similarly, f is called twice continuously differentiable (on S) if all the second order partial derivatives of f are continuous on S .

Theorem 0.12. Let $f : S \subset \mathbb{R}^m \rightarrow \mathbb{R}$ and let $x_0 \in S^0$. Also, let $\frac{\partial f}{\partial x_j} \Big|_{x_0}$ exist for each $j \in \{1, \dots, m\}$. Then f is differentiable at x_0 iff

$$\lim_{h \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) - \alpha \cdot h|}{\|h\|} = 0,$$

where $\alpha = \left(\frac{\partial f}{\partial x_1} \Big|_{x_0}, \dots, \frac{\partial f}{\partial x_m} \Big|_{x_0} \right)$.

Proof. If

$$\lim_{h \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) - \alpha \cdot h|}{\|h\|} = 0,$$

then by definition, f is differentiable at x_0 . Conversely, let f be differentiable at x_0 . Then there exists $\beta = (\beta_1, \dots, \beta_m) \in \mathbb{R}^m$ such that

$$\lim_{h \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) - \beta \cdot h|}{\|h\|} = 0.$$

We have seen earlier that in this case $\beta_j = \frac{\partial f}{\partial x_j} \Big|_{x_0}$ for all $j \in \{1, \dots, m\}$ and hence $\beta = \alpha$.

Therefore

$$\lim_{h \rightarrow 0} \frac{|f(x_0 + h) - f(x_0) - \alpha \cdot h|}{\|h\|} = 0.$$

□

Example 0.13. (a) The function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, defined by $f(x, y) = \begin{cases} xy \frac{y^2 - x^2}{x^2 + y^2}, & \text{if } (x, y) \neq (0, 0), \\ 0, & \text{if } (x, y) = (0, 0), \end{cases}$ is differentiable at $(0, 0)$ and $f'(0, 0) = \begin{bmatrix} 0 & 0 \end{bmatrix}$.

(b) The function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, defined by $f(x, y) = \begin{cases} \frac{xy}{\sqrt{x^2+y^2}}, & \text{if } (x, y) \neq (0, 0), \\ 0, & \text{if } (x, y) = (0, 0), \end{cases}$ is not differentiable at $(0, 0)$.

Maximum and minimum values of $D_u f(x_0)$: Let $f : S \subset \mathbb{R}^m \rightarrow \mathbb{R}$ be differentiable at $x_0 \in S^0$ and let $\nabla f(x_0) \neq 0$. If $u \in \mathbb{R}^m$ such that $\|u\| = 1$, then $D_u f(x_0) = \nabla f(x_0) \cdot u$ and since $|\nabla f(x_0) \cdot u| \leq \|\nabla f(x_0)\| \|u\| = \|\nabla f(x_0)\|$, we have $-\|\nabla f(x_0)\| \leq D_u f(x_0) \leq \|\nabla f(x_0)\|$. Also, by the equality condition of Cauchy–Schwarz inequality, the function $u \mapsto D_u f(x_0)$ attains its maximum value $\|\nabla f(x_0)\|$ for $u = \frac{1}{\|\nabla f(x_0)\|} \nabla f(x_0)$ and the minimum value $-\|\nabla f(x_0)\|$ for $u = -\frac{1}{\|\nabla f(x_0)\|} \nabla f(x_0)$.

Example 0.14. Let $f(x, y) = 4x - x^2 - y^2$ for all $(x, y) \in \mathbb{R}^2$. Then $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is differentiable and $\nabla f(1, 1) = (-2, -2)$. Hence the maximum value of $D_u f(1, 1)$ is $2\sqrt{2}$ which occurs for $u = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ and the minimum value of $D_u f(1, 1)$ is $-2\sqrt{2}$ which occurs for $u = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.

Equivalent conditions of differentiability (Increment theorem): A function $f : S \subset \mathbb{R}^m \rightarrow \mathbb{R}$ is differentiable at $x_0 \in S^0$ if either of the following two conditions holds.

(a) There exist $\delta > 0$ and a function $e : B_\delta(0) \rightarrow \mathbb{R}$ such that $\lim_{h \rightarrow 0} e(h) = 0$ and

$$f(x_0 + h) = f(x_0) + \nabla f(x_0) \cdot h + \|h\|e(h) \quad \text{for all } h \in B_\delta(0).$$

(b) There exist $\delta > 0$ and functions $\varepsilon_j : B_\delta(0) \rightarrow \mathbb{R}$ for $j = 1, \dots, m$ such that $\lim_{h \rightarrow 0} \varepsilon_j(h) = 0$ for each $j \in \{1, \dots, m\}$ and

$$f(x_0 + h) = f(x_0) + \nabla f(x_0) \cdot h + h_1 \varepsilon_1(h) + \dots + h_m \varepsilon_m(h) \quad \text{for all } h \in B_\delta(0).$$

Proof. Let us first assume that f is differentiable at x_0 . Since $x_0 \in S^0$, there exists $\delta > 0$ such that $B_\delta(x_0) \subset S$. Let $\varepsilon : B_\delta(0) \rightarrow \mathbb{R}$ be defined by

$$e(h) = \begin{cases} \frac{f(x_0+h) - f(x_0) - \nabla f(x_0) \cdot h}{\|h\|} & \text{if } h \neq 0, \\ 0 & \text{if } h = 0. \end{cases}$$

Since f is differentiable at x_0 , $\lim_{h \rightarrow 0} e(h) = 0$. Also, the equation in (a) is satisfied for all $h \in B_\delta(0)$. Again, let $j \in \{1, \dots, m\}$ and let $\varepsilon_j : B_\delta(0) \rightarrow \mathbb{R}$ be defined by

$$\varepsilon_j(h) = \begin{cases} \frac{|h_j|}{\|h\|} e(h) & \text{if } h = (h_1, \dots, h_m) \neq 0, \\ 0 & \text{if } h = 0. \end{cases}$$

Since $\frac{|h_j|}{\|h\|} \leq 1$ for all $h = (h_1, \dots, h_m) \in B_\delta(0) \setminus \{0\}$ and since $\lim_{h \rightarrow 0} e(h) = 0$, we get $\lim_{h \rightarrow 0} \varepsilon_j(h) = 0$. Also, since $h_1 \varepsilon_1(h) + \dots + h_m \varepsilon_m(h) = \sum_{j=1}^m h_j \varepsilon_j(h) = \|h\| e(h)$ for all $h = (h_1, \dots, h_m) \in B_\delta(0)$, the equation in (b) is satisfied for all $h \in B_\delta(0)$ as we have already shown that the equation in (a) is satisfied for all $h \in B_\delta(0)$.

Conversely, let the condition in (a) be satisfied. Then

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - \nabla f(x_0) \cdot h}{\|h\|} = \lim_{h \rightarrow 0} e(h) = 0$$

and so f is differentiable at x_0 . Again, let the condition in (b) be satisfied. Then

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - \nabla f(x_0) \cdot h}{\|h\|} &= \lim_{h \rightarrow 0} \frac{h_1 \varepsilon_1(h) + \cdots + h_m \varepsilon_m(h)}{\|h\|} \\ &= \lim_{h \rightarrow 0} \frac{\|h\|(\varepsilon_1(h) \frac{h_1}{\|h\|} + \cdots + \varepsilon_m(h) \frac{h_m}{\|h\|})}{\|h\|} \\ &= \lim_{h \rightarrow 0} (\varepsilon_1(h) \frac{h_1}{\|h\|} + \cdots + \varepsilon_m(h) \frac{h_m}{\|h\|}). \end{aligned}$$

But since $\frac{|h_j|}{\|h\|} \leq 1$ and $\varepsilon_j(h) \rightarrow 0$ as $h \rightarrow 0$, this limit is 0, and so f is differentiable at x_0 . $\square$

Linearization: Let $f : S \subset \mathbb{R}^m \rightarrow \mathbb{R}$ be differentiable at $x_0 \in S^0$. The function $L : \mathbb{R}^m \rightarrow \mathbb{R}$, defined by $L(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0)$ for all $x \in \mathbb{R}^m$, is called the linearization of f at x_0 . We note that L is the unique polynomial function of degree one satisfying $L(x_0) = f(x_0)$ and

$$\lim_{x \rightarrow x_0} \frac{|f(x) - L(x)|}{\|x - x_0\|} = 0.$$

Example 0.15. If $f(x, y, z) = x^3 + y^2 + zx$ for all $(x, y, z) \in \mathbb{R}^3$, then $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ is differentiable at

$$(1, 1, 1) \text{ and } \nabla f(1, 1, 1) = (4, 2, 3).$$

Hence the linearization $L : \mathbb{R}^3 \rightarrow \mathbb{R}$ of f at $(1, 1, 1)$ is defined by

$$L(x, y, z) = 3 + (4, 2, 3) \cdot (x - 1, y - 1, z - 1) = 4x + 2y + 3z - 6 \quad \text{for all } (x, y, z) \in \mathbb{R}^3.$$