MA 105 : Calculus
Division 3, Lecture 24

Prof. Sudhir R. Ghorpade
IIT Bombay

Recap of the previous lecture

- Necessary condition for differentiability
- Sufficient condition for differentiability
- Examples
- Basic properties of differentiable functions
- Chain Rules (in three avatars). Examples
- Geometric Interpretation of the Gradient
- Review of tangent line of a curve

Tangent Line

Recall the one variable situation:
Let $a < b$ and $x_0 \in (a, b)$. Let a function $f : (a, b) \to \mathbb{R}$ be differentiable at x_0. Then the equation of the **tangent line** to the curve $y = f(x)$ in $\mathbb{R}^2$ at $(x_0, f(x_0))$ is given by

$$y - f(x_0) = f'(x_0)(x - x_0).$$

More generally, suppose $D \subset \mathbb{R}^2$ and $P_0 := (x_0, y_0)$ is an interior point of D. Let a function $F : D \to \mathbb{R}$ have partial derivatives at P_0, and let $(\nabla F)(P_0) \neq (0, 0)$. Suppose F defines a curve C in $\mathbb{R}^2$ (implicitly) by the equation $F(x, y) = 0$ for $(x, y) \in D$, and P_0 lies on C. Then the equation of the **tangent line** to C at P_0 is given by

$$F_x(P_0)(x - x_0) + F_y(P_0)(y - y_0) = 0.$$

Note: If $D := (a, b) \times \mathbb{R}$, $f : (a, b) \to \mathbb{R}$, and for $(x, y) \in D$, $F(x, y) := y - f(x)$, then $y_0 = f(x_0)$, $F_x(P_0) = -f'(x_0)$ and $F_y(P_0) = 1$. We recover the earlier equation of the tangent line.

Tangent Plane

Let us turn to the two variable situation:
Let $D \subset \mathbb{R}^2$, (x_0, y_0) be an interior point of D, and let a function $f : D \to \mathbb{R}$ be differentiable at (x_0, y_0). Then the equation of the **tangent plane** to the surface $z = f(x, y)$ in $\mathbb{R}^3$ at $(x_0, y_0, f(x_0, y_0))$ is given by

$$z - f(x_0, y_0) = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

More generally, suppose $E \subset \mathbb{R}^3$ and $P_0 := (x_0, y_0, z_0)$ is an interior point of E. Let a function $F : E \to \mathbb{R}$ have partial derivatives at P_0, and let $(\nabla F)(P_0) \neq (0, 0, 0)$. Suppose F (implicitly) defines a surface S in $\mathbb{R}^3$ by the equation $F(x, y, z) = 0$ for $(x, y, z) \in E$, and P_0 lies on S. Then the equation of the **tangent plane** to S at P_0 is given by

$$F_x(P_0)(x - x_0) + F_y(P_0)(y - y_0) + F_z(P_0)(z - z_0) = 0.$$

Note: As before, if we let $F(x, y, z) := z - f(x, y)$, then we recover the earlier equation of the tangent plane.

Normal Line

Now $(\nabla F)(P_0) \cdot (x - x_0, y - y_0, z - z_0) = 0$ for all (x, y, z) on the tangent plane to the surface S, that is, $(\nabla F)(P_0)$ is perpendicular to the tangent plane. The line passing through $P_0 = (x_0, y_0, z_0)$ and parallel to the nonzero vector $(\nabla F)(P_0) = (F_x(P_0), F_y(P_0), F_z(P_0))$ is called the **normal line** to the surface defined by $F(x, y, z) = 0$ at P_0. The parametric equations of this normal line are

$$x = x_0 + F_x(P_0)t, \quad y = y_0 + F_y(P_0)t, \quad z = z_0 + F_z(P_0)t, \quad t \in \mathbb{R}.$$

If all $F_x(P_0), F_y(P_0), F_z(P_0)$ are nonzero, then the equations are

$$\frac{x - x_0}{F_x(P_0)} = \frac{y - y_0}{F_y(P_0)} = \frac{z - z_0}{F_z(P_0)}.$$

Also, the parametric equations of the normal line to the surface defined by $z - f(x, y) = 0$ at (x_0, y_0, z_0) are

$$x = x_0 - f_x(x_0, y_0)t, \quad y = y_0 - f_y(x_0, y_0)t, \quad z = f(x_0, y_0) + t, \quad t \in \mathbb{R}.$$

Tangent Vector to a Curve

We now describe a crucial property of the normal line to a surface. Let C be a parametrized smooth curve in $\mathbb{R}^3$ given by

$$t \longmapsto (x(t), y(t), z(t)), \quad t \in [\alpha, \beta].$$

Let $t_0 \in (\alpha, \beta)$ be such that $(x'(t_0), y'(t_0), z'(t_0)) \neq (0, 0, 0)$. Then the vector $\big(x'(t_0), y'(t_0), z'(t_0)\big)$ is called the **tangent vector** to the curve C at the point $\big(x(t_0), y(t_0), z(t_0)\big)$ on it.

Now suppose that the curve C lies on a surface S defined by $F(x, y, z) = 0$, that is, $\phi(t) := F\,(x(t), y(t), z(t)) = 0$ for all $t \in [\alpha, \beta]$. By the Chain Rule,

$$0 = \phi'(t_0) = F_x(P_0)x'(t_0) + F_y(P_0)y'(t_0) + F_z(P_0)z'(t_0).$$

Thus the normal line to the surface S at P_0 is perpendicular to tangent vectors to all curves on S that pass through P_0.

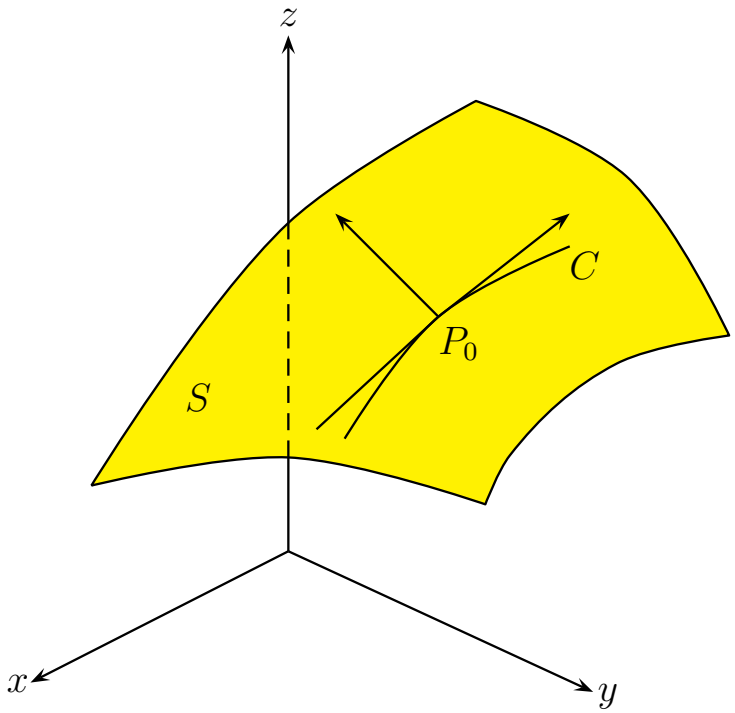


Figure: Normal line to the surface S at P_0 and a curve C on S

Special case: Suppose a surface S is given by $z = f(x, y)$ for $(x, y) \in D$. Let $Q_0 := (x_0, y_0) \in D$, and let a parametrized curve given by $(x(t), y(t))$, $t \in [\alpha, \beta]$, pass through Q_0. Let $z(t) := f(x(t), y(t))$, $t \in [\alpha, \beta]$. Then the curve C given by $(x(t), y(t), z(t))$, $t \in [\alpha, \beta]$, lies on S, and passes through $P_0 := f(Q_0)$.

Hence the normal line

$$x = x_0 - f_x(Q_0)t,\ y = y_0 - f_y(Q_0)t,\ z = f(Q_0) + t,\ t \in \mathbb{R},$$

to the surface S at P_0 is perpendicular to the tangent vector $(x'(t_0), y'(t_0), z'(t_0))$ to the curve C at P_0, that is, $(-f_x(Q_0), -f_y(Q_0), 1) \cdot (x'(t_0), y'(t_0), z'(t_0)) = 0$.
(This can also be independently verified as $z'(t_0) = f_x(Q_0)x'(t_0) + f_y(Q_0)y'(t_0)$ by the chain rule.)

In particular, let $c_0 := f(x_0, y_0)$, and suppose the level curve $\{(x, y) \in D : f(x, y) = c_0\}$ is parametrized by $(x(t), y(t))$ with $t \in [\alpha, \beta]$. Then the contour line C is parametrized by $(x(t), y(t), c_0)$ with $t \in [\alpha, \beta]$, and it lies on the surface S. Hence the normal line to S at $P_0 := (x_0, y_0, c_0)$ is perpendicular to the corresponding contour line C.

Examples:

(i) Let $F(x, y, z) := x^2 + y^2 + z^2 - 1$ for $(x, y, z) \in \mathbb{R}^3$. Then $(\nabla F)(x, y, z) = (2x, 2y, 2z)$ and $(\nabla F)(0, 0, 1) = (0, 0, 2)$. The tangent plane to the surface defined by $F(x, y, z) = 0$ at the point $(0, 0, 1)$ is $0(x-0)+0(y-0)+2(z-1) = 0$, that is, $z = 1$. The normal line at $(0, 0, 1)$ is $x = 0, y = 0, z = 1 + 2t$, where $t \in \mathbb{R}$, that is, the z-axis. Note that the surface is the unit sphere in $\mathbb{R}^3$, and the point $(0, 0, 1)$ is its north pole.

(ii) Let $F(x, y, z) := e^x + \sin y - \cos z$ for $(x, y, z) \in \mathbb{R}^3$. Then $(\nabla F)(x, y, z) = (e^x, \cos y, \sin z)$ for $(x, y, z) \in \mathbb{R}^3$ and $(\nabla F)(0, 0, 0) = (1, 1, 0)$. The tangent plane to the surface defined by $F(x, y, z) = 0$ at $(0, 0, 0)$ is $x + y + 0z = 0$, that is, $x + y = 0$. The normal line at $(0, 0, 0)$ is $x = t, y = t, z = 0$, where $t \in \mathbb{R}$, that is, the intersection of the planes given by $x = y$ and $z = 0$.

(iii) Let $F(x, y, z) := x^2 + y^2 - 1$ for $(x, y, z) \in \mathbb{R}^3$. Then $(\nabla F)(x, y, z) = (2x, 2y, 0)$ for $(x, y, z) \in \mathbb{R}^3$. The tangent plane to the surface defined by $F(x, y, z) = 0$ at the point $P_0 := (\cos\theta_0, \sin\theta_0, z_0)$ on the surface is $2\cos\theta_0(x - \cos\theta_0) + 2\sin\theta_0(y - \sin\theta_0) + 0(z - z_0) = 0$, that is, $\cos\theta_0\ x + \sin\theta_0\ y = 1$. The normal line at P_0 is $x = \cos\theta_0(1 + 2t)$, $y = \sin\theta_0(1 + 2t)$, $z = z_0$, where $t \in \mathbb{R}$. Note that the surface is a cylinder in $\mathbb{R}^3$.

A surface in $\mathbb{R}^3$ can be parametrically given by $\big(x(u, v), y(u, v), z(u, v)\big)$ for $(u, v) \in E$, where $E \subset \mathbb{R}^2$, and $x,\ y,\ z : E \to \mathbb{R}$ are continuous functions. A point $P_0 := (x_0, y_0, z_0)$ on this surface is given by $\big(x(u_0, v_0), y(u_0, v_0), z(u_0, v_0)\big)$ for some $(u_0, v_0) \in E$. In a subsequent lecture, we shall define a normal vector to this parametrically given surface at P_0, and also obtain the equation of the tangent plane to the surface at P_0.

Local Extremum

Let $D \subset \mathbb{R}^2$, and (x_0, y_0) be an interior point of D. We say that a function $f : D \to \mathbb{R}$ has

(i) a **local minimum** at (x_0, y_0) if there is $r > 0$ such that $B((x_0, y_0), r) \subset D$ and $f(x, y) \geq f(x_0, y_0)$ for all $(x, y) \in B((x_0, y_0), r)$.

(ii) a **local maximum** at (x_0, y_0) if there is $r > 0$ such that $B((x_0, y_0), r) \subset D$ and $f(x, y) \leq f(x_0, y_0)$ for all $(x, y) \in B((x_0, y_0), r)$.

(iii) a **local extremum** at (x_0, y_0) if it has a local maximum or a local minimum at (x_0, y_0).

Examples:

Let $f(x, y) := x^2 + y^2$ for $(x, y) \in \mathbb{R}^2$. Then f has a local minimum and $-f$ has a local maximum at $(0, 0)$. In fact, these are the global extrema of f and $-f$.

Proposition (Necessary condition for a local extremum)

Let $D \subset \mathbb{R}^2$, and let (x_0, y_0) be an interior point of D. Suppose $f : D \to \mathbb{R}$ has a local extremum at (x_0, y_0). If $\boldsymbol{u}$ is a unit vector and $(\boldsymbol{D_u} f)(x_0, y_0)$ exists, then $(\boldsymbol{D_u} f)(x_0, y_0) = 0$.

In particular, if $(\nabla f)(x_0, y_0)$ exists, then $(\nabla f)(x_0, y_0) = (0, 0)$.

Proof. Suppose $(\boldsymbol{D_u} f)(x_0, y_0)$ exists. Let E denote the set of all $(x_0 + tu_1, y_0 + tu_2)$ in D, where $t \in \mathbb{R}$. By the Carathéodory lemma for a directional derivative, there is $f_1 : E \to \mathbb{R}$ such that f_1 is continuous at (x_0, y_0), and

$$f(x_0 + tu_1, y_0 + tu_2) - f(x_0, y_0) = t\, f_1(x_0 + tu_1, y_0 + tu_2)$$

whenever $(x_0 + tu_1, y_0 + tu_2) \in D$. Suppose f has a local minimum at (x_0, y_0). Then there is $\delta > 0$ such that $f(x_0 + tu_1, y_0 + tu_2) \geq f(x_0, y_0)$ for all $t \in (-\delta, \delta)$.

Hence $f_1(x_0 + tu_1, y_0 + tu_2) \geq 0$ for all $t \in (0, \delta)$ and $f_1(x_0 + tu_1, y_0 + tu_2) \leq 0$ for all $t \in (-\delta, 0)$.

Thus $0 \leq \lim_{t \to 0^+} f_1(x_0 + tu_1, y_0 + tu_2) = f_1(x_0, y_0)$ as well as $f_1(x_0, y_0) = \lim_{t \to 0^-} f_1(x_0 + tu_1, y_0 + tu_2) \leq 0$ since f_1 is continuous at (x_0, y_0). Hence $(\boldsymbol{D_u} f)(x_0, y_0) = f_1(x_0, y_0) = 0$.

In particular, if $(\nabla f)(x_0, y_0)$ exists, then considering $\boldsymbol{u} := (1, 0)$ and $\boldsymbol{u} := (0, 1)$, we see that $(\nabla f)(x_0, y_0) = (0, 0)$.

A similar argument holds if f has a local maximum at (x_0, y_0).

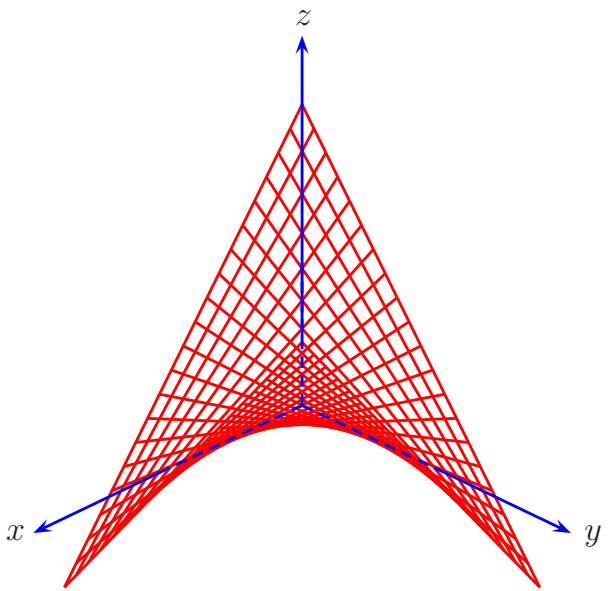
The converse of the above proposition does not hold.

Example:
Let $f(x, y) := x\,y$ for $(x, y) \in \mathbb{R}^2$. Then $(\boldsymbol{D_u} f)(0, 0) = 0$ for every unit vector $\boldsymbol{u}$, but f does not have a local extremum at $(0, 0)$ since $f(0, 0) = 0$, while $f(1/n, 1/n) = 1/n^2 > 0$ and $f(1/n, -1/n) = -1/n^2 < 0$ for all $n \in \mathbb{N}$.

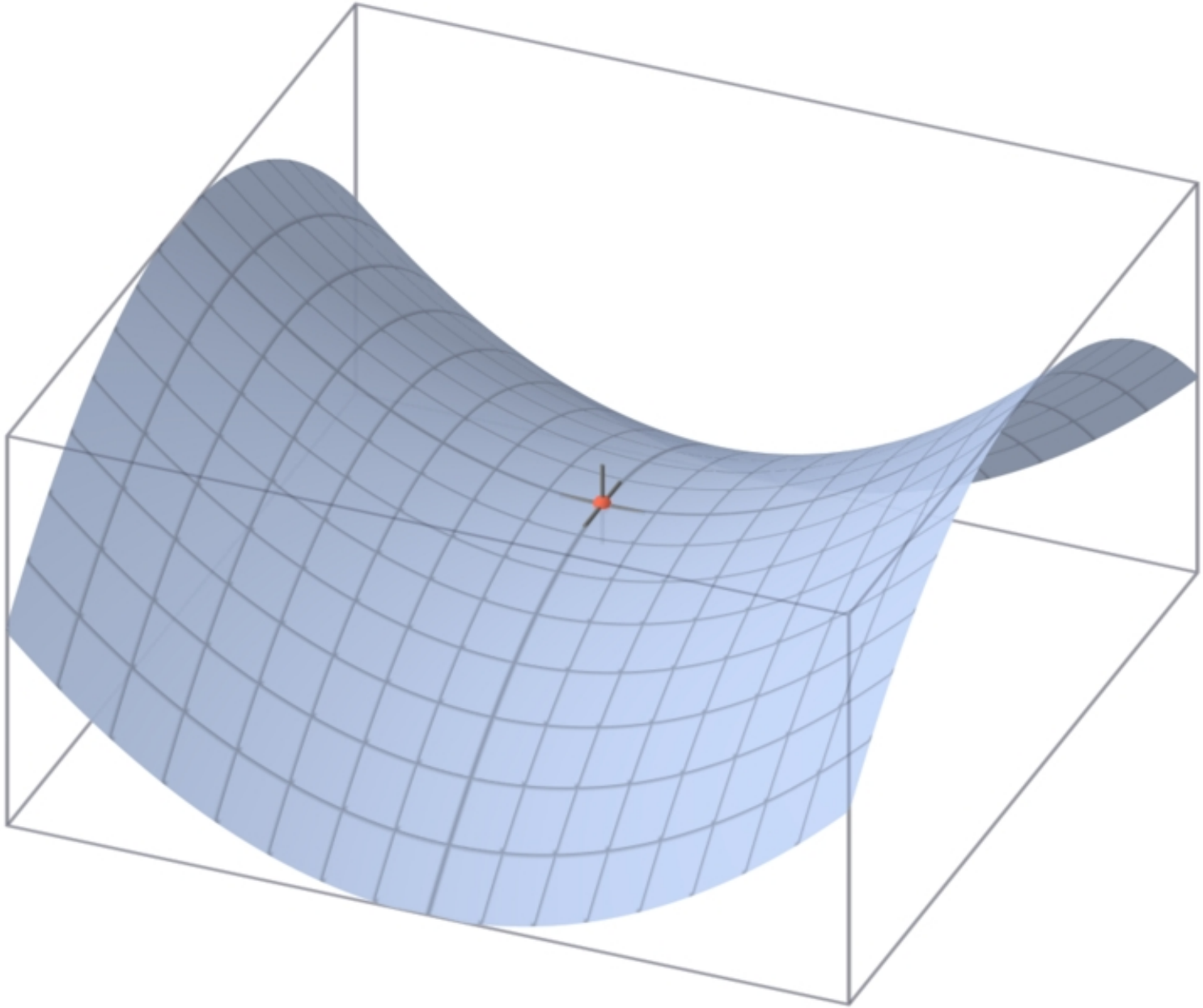
Saddle Point

Let (x_0, y_0) be an interior point of $D \subset \mathbb{R}^2$, and let $f : D \to \mathbb{R}$. We say that the function f has a **saddle point** at (x_0, y_0) if $(\nabla f)(x_0, y_0) = (0, 0)$, but f does not have a local extremum at (x_0, y_0).

Example: $(0, 0)$ is a saddle point of $f(x, y) := x\,y$, $(x, y) \in \mathbb{R}^2$. The surface given by $z = x\,y$ looks as follows.



Picture of a Saddle



Before considering sufficient conditions for a local extremum or for a saddle point, we consider the following useful results.

Let $D \subset \mathbb{R}^2$ be such that every point of D is an interior point, and the line segment joining any two points in D lies in D. Let $(x_0, y_0), (x_1, y_1) \in D$, and consider a function $f : D \to \mathbb{R}$.

(Bivariate Mean Value Theorem)

Suppose f has continuous partial derivatives on D. Then there is $(c, d) \in D$ lying on the open line segment joining (x_0, y_0) and (x_1, y_1) such that

$$f(x_1, y_1) = f(x_0, y_0) + (x_1 - x_0) f_x(c, d) + (y_1 - y_0) f_y(c, d).$$

Proof: Use the one variable MVT for $\phi : [0, 1] \to \mathbb{R}$ defined by $\phi(t) := f(x_0 + t(x_1 - x_0), y_0 + t(y_1 - y_0))$, noting that for $t \in (0, 1)$, $\phi'(t) = (x_1 - x_0) f_x((x_0 + t(x_1 - x_0), y_0 + t(y_1 - y_0)) + (y_1 - y_0) f_y((x_0 + t(x_1 - x_0), y_0 + t(y_1 - y_0))$ by the chain rule. $\square$

To simplify the statements of extensions of the above result, let us introduce a partial differential operator

$$\mathcal{D}_{h,k} := h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}, \quad \text{where } h, k \in \mathbb{R}.$$

It can be applied to a function f which has partial derivatives. Further, by the Mixed Partials Theorem, we obtain

$$\mathcal{D}^2_{h,k} f = \mathcal{D}_{h,k}(\mathcal{D}_{h,k} f) = h^2 f_{xx} + 2hk\ f_{xy} + k^2 f_{yy}$$

if a function f has continuous partial derivatives of order ≤ 2.

(Bivariate Extended Mean Value Theorem)

Suppose f has continuous partial derivatives of order ≤ 2 on D. Let $h := x_1 - x_0$ and $k := y_1 - y_0$. Then there is $(c, d) \in D$ lying on the open line segment joining (x_0, y_0) and (x_1, y_1) with

$$f(x_1, y_1) = f(x_0, y_0) + (\mathcal{D}_{h,k} f)(x_0, y_0) + \frac{1}{2}(\mathcal{D}^2_{h,k} f)(c, d).$$

Thus $f(x_0, y_0) + (\mathcal{D}_{h,k} f)(x_0, y_0)$, or in other words,

$$f(x_0, y_0) + (x_1 - x_0) f_x(x_0, y_0) + (y_1 - y_0) f_y(x_0, y_0)$$

can be considered as an approximation of $f(x_1, y_1)$ with $(1/2)(\mathcal{D}^2_{h,k} f)(c, d)$ as the error.

More generally, we can prove the following result.

(Bivariate Taylor Theorem)

Let $n \in \mathbb{N}$. Suppose f has continuous partial derivatives of order $\leq (n+1)$ on D. Then there is $(c, d) \in D$ lying on the open line segment joining (x_0, y_0) and (x_1, y_1) such that

$$f(x_1, y_1) = \sum_{i=0}^{n} \frac{1}{i!} (\mathcal{D}^i_{h,k} f)(x_0, y_0) + R_n,$$

where $R_n := \dfrac{1}{(n+1)!} (\mathcal{D}^{n+1}_{h,k} f)(c, d)$, $h := x_1 - x_0$, $k := y_1 - y_0$.

The bivariate Taylor theorem can be proved by considering the function $\phi : [0,1] \to \mathbb{R}$ defined by

$$\phi(t) := f(x_0 + t(x_1 - x_0), y_0 + t(y_1 - y_0)),$$

and by using the one variable Taylor theorem for ϕ, as in the case of the bivariate mean value theorem. (Note: The proof involves repeated use of the Mixed Partials Theorem.)

The polynomial

$$P_n(x,y) \;:=\; \sum_{i=0}^{n} \frac{1}{i!} (\mathcal{D}^i_{x-x_0,y-y_0} f)(x_0, y_0)$$

of degree $\leq n$ is an 'approximation' of $f(x,y)$ whenever $(x,y) \in D$. It is called the nth **bivariate Taylor polynomial** of f around (x_0, y_0), and

$$R_n(x,y) := \frac{1}{(n+1)!} (\mathcal{D}^{n+1}_{x-x_0,y-y_0} f)(c,d)$$

is called the nth **Lagrange remainder** of f around (x_0, y_0).