Multivariable Calculus
(Week 6)

Topics: Directional Derivatives

Tangent Plane

Critical Points

Extrema of Functions on Restricted Domain

(**Reference Chapters:** Chapter 12.3, 12.7, 13.1-13.2 of Adams and Essex; Chapter 13.6 – 13.9 of Larson and Edwards)

Key References of this file:

Larson, R. and Edwards, B., Multivariable Calculus (Metric Version – 11th Edition), CENGAGE Learning, 2016.

Adams, R.A. and Essex, C., Calculus – A Complete Course, 9th Edition, Pearson, 2018.

Prepared by Dr. Hugo MAK

Directional Derivatives

Recall the terms "gradient" and "partial derivatives":

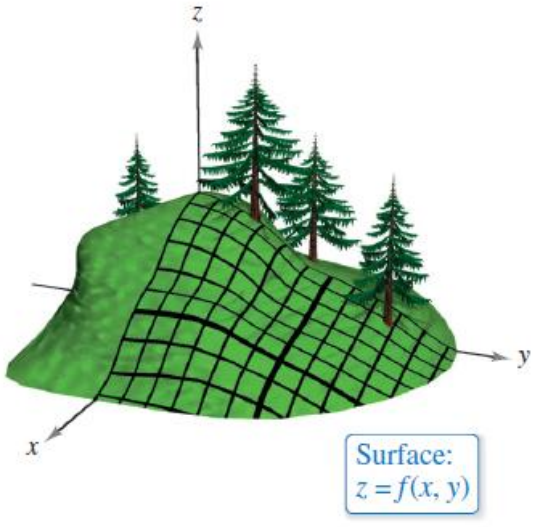
- When we take **first partial derivative** of a function of several variables, we obtain the **rate of change of the function with respect to distance measured in the direction of one of the coordinate axes**.
- We define the gradient vector of a two-variable function $\nabla f(x, y) := f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j}$ and the gradient vector of a three-variable function $\nabla f(x, y, z) := f_x(x, y, z)\mathbf{i} + f_y(x, y, z)\mathbf{j} + f_z(x, y, z)\mathbf{k}$.
- Here **i, j** and **k** are the unit basis vectors from origin to the points $(1,0,0), (0, 1, 0)$ and $(0, 0, 1)$, therefore the symbol ∇ is actually a vector differential operator defined by:

$$\nabla = \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\rangle$$

- Now, we are interested in the rate of change of a function of two or more variables in any direction, where such type of derivative is known as "**Directional Derivative**".

Directional Derivatives

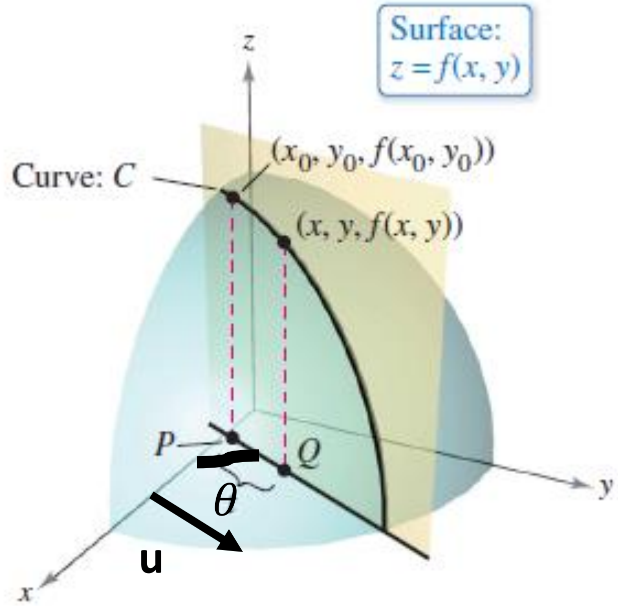
Motivation: You are now standing on the hillside represented by $z = f(x, y)$. You wish to determine the hill's inclination towards the z-axis.
Tool: To determine the slope at a point on a surface, we need "directional derivatives".



Figures due to Larson

Let $z = f(x, y)$ be a surface, and $P = (x_0, y_0)$ be a point in the domain of f.

Direction of interest: $\mathbf{u} = \langle \cos\theta, \sin\theta \rangle$



Goal: Find the slope with respect to this direction.

Strategy: Intersect the surface with a vertical plane that passes through *P* and is parallel to **u**. Such vertical plane will intersect the surface and the intersection will be a curve C.

Outcome: Slope of surface at $(x_0, y_0, f(x_0, y_0))$ in the direction of **u** = slope of C at that point.

Directional Derivatives

Consider the parametrization of the line that passes through P and Q

$$\begin{cases} x = x_0 + t\cos\theta \\ y = y_0 + t\sin\theta \end{cases}$$ (so that $Q(x, y)$ will lie on the line L for any t)

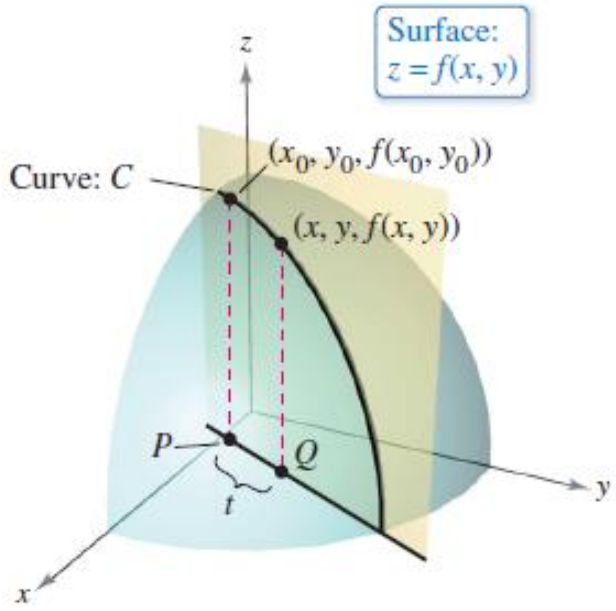


Figure due to Larson

Then we get two corresponding points on the surface:

$\big(x, y, f(x, y)\big)$ (for Q) & $\big(x_0, y_0, f(x_0, y_0)\big)$ (for P)

Distance between P and Q = $\sqrt{(x - x_0)^2 + (y - y_0)^2} = \|t\|$

Slope of the line that passes through the two points on surface

$$= \frac{f(x, y) - f(x_0, y_0)}{t} = \frac{f(x_0 + t\cos\theta, y_0 + t\sin\theta) - f(x_0, y_0)}{t}$$

Let $t \to 0$, we get the definition of directional derivative as follows:

Definition of Directional Derivative

Let f be a function of two variables x and y and let $\mathbf{u} = \cos\theta\mathbf{i} + \sin\theta\mathbf{j}$ be a unit vector. Then the **directional derivative of f in the direction of $\mathbf{u}$,** denoted by $D_{\mathbf{u}}f$, is

$$D_{\mathbf{u}}f(x, y) = \lim_{t \to 0} \frac{f(x + t\cos\theta, y + t\sin\theta) - f(x, y)}{t}$$

provided this limit exists.

General Form of Directional Derivatives

Now, if we let $\mathbf{u} = \langle u, v \rangle$ be a unit vector, then the directional derivative of $f(x, y)$ at the point (a, b) in the direction of **u** is given by the following expression:

$$D_{\mathbf{u}} f(a, b) = \lim_{h \to 0^+} \frac{f(a + hu, b + hv) - f(a, b)}{h}$$

By first principle of limit, we can rewrite the expression as $D_{\mathbf{u}} f(a, b) = \frac{d}{dt} f(a + tu, b + tv) \,|_{t=0}$

(**provided that the limit exists**)

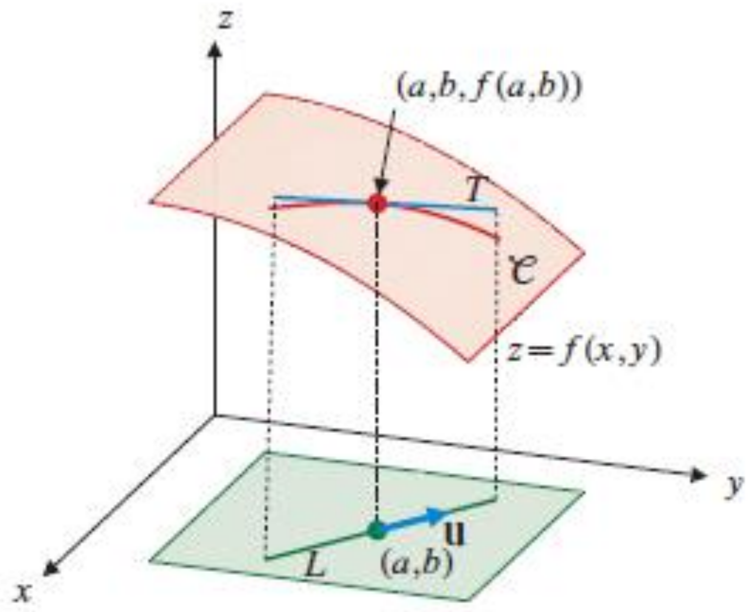
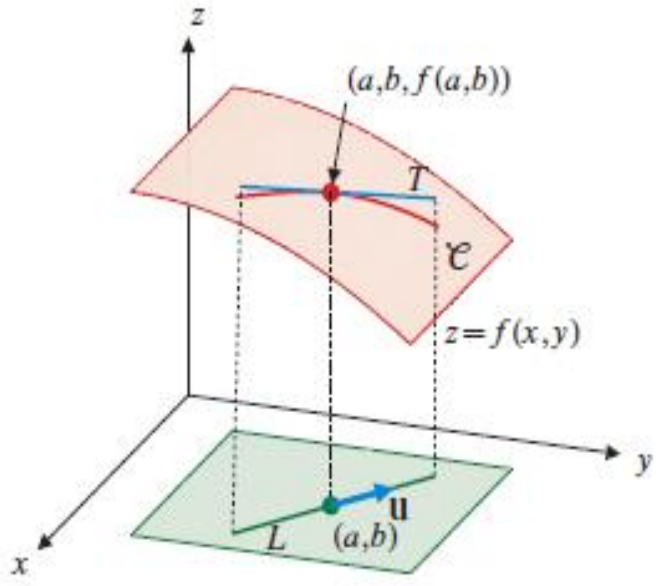


Figure due to Adams and Essex

In the figure, the unit vector **u** will determine a line L that passes through (a, b) in $Dom(f)$. The intersection between the vertical plane that contains L and the graph of f is a curve C.

Then, we can draw a tangent line T, with slope $D_{\mathbf{u}} f(a, b)$ at the point $\left(a, b, f(a, b)\right)$.

General Form of Directional Derivatives



Consider the expression

$$D_{\mathbf{u}}f(a,b) = \frac{d}{dt}f(a+tu, b+tv)\Big|_{t=0} = uf_x(a,b) + vf_y(a,b) = \mathbf{u}\cdot\nabla f(a,b)$$

Hence, we can find directional derivatives with the help of "**gradient**".
If f is differentiable at (a,b), and $\mathbf{u} = \langle u, v\rangle$ is a unit vector, then the directional derivative of f at (a,b) in the direction of **u** is given by the above expression.

Remarks: Given any non-zero vector **v**, we can always normalize **v** to obtain a unit vector.

Explanation of why u must be a unit vector
The line L through (a,b) parallel to **u** is given by the position vector $\mathbf{r}(t) = a\mathbf{i} + b\mathbf{j} + t\mathbf{u}$. We regard L as a single coordinate axis, and all points on it can be expressed in terms of such coordinate, then $f\big(x(t), y(t)\big) = g(t)$ along the line L, hence we have

$$\frac{dg(t)}{dt} = \frac{df\big(x(t), y(t)\big)}{dt} = f_x(x,y)x'(t) + f_y(x,y)y'(t) = \nabla f\cdot\frac{d\mathbf{r}(t)}{dt}$$

However, such derivation is true for all t, the directional derivative along L is the rate of change with respect to arc length s, hence $\|\mathbf{u}\| = \left\|\frac{d\mathbf{r}(t)}{dt}\right\| = \sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2} = \frac{ds}{dt} = 1$

Examples of Directional Derivatives

Formula for "Directional Derivatives"

$$D_{\mathbf{u}} f(x, y) = \nabla f(x, y) \cdot \mathbf{u}$$

$$D_{\mathbf{u}} f(x, y, z) = \nabla f(x, y, z) \cdot \mathbf{u}$$

Example 1. Given that $f(x, y, z) = x \sin yz$.

(a) Find the gradient of f.

(b) Find the directional derivative of f at (1, 3, 0) in the direction of $\mathbf{v} = \mathbf{i} + 2\,\mathbf{j} - \mathbf{k}$.

Solution:

(a) The gradient of f is $\nabla f(x, y, z) = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)\rangle = \langle \sin yz, xz \cos yz, xy \cos yz\rangle$

(b) From (a), all partial derivatives are continuous, thus f is differentiable.

At (1, 3, 0) we have $\nabla f(1, 3, 0) = \langle 0, 0, 3\rangle$.

The unit vector in the direction of $\mathbf{v} = \mathbf{i} + 2\,\mathbf{j} - \mathbf{k}$ is $\mathbf{u} = \left\langle \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}} \right\rangle$

Therefore directional derivative $D_{\mathbf{u}} f(1, 3, 0) = \nabla f(1, 3, 0) \cdot \mathbf{u} = -\frac{3}{\sqrt{6}} = -\frac{\sqrt{6}}{2}$

Examples of Directional Derivatives

Example 2 (due to Larson). Find the directional derivative of $f(x, y) = 3x^2 - 2y^2$ at $\left(-\frac{3}{4}, 0\right)$ in the direction from $P\left(-\frac{3}{4}, 0\right)$ to $Q(0, 1)$.

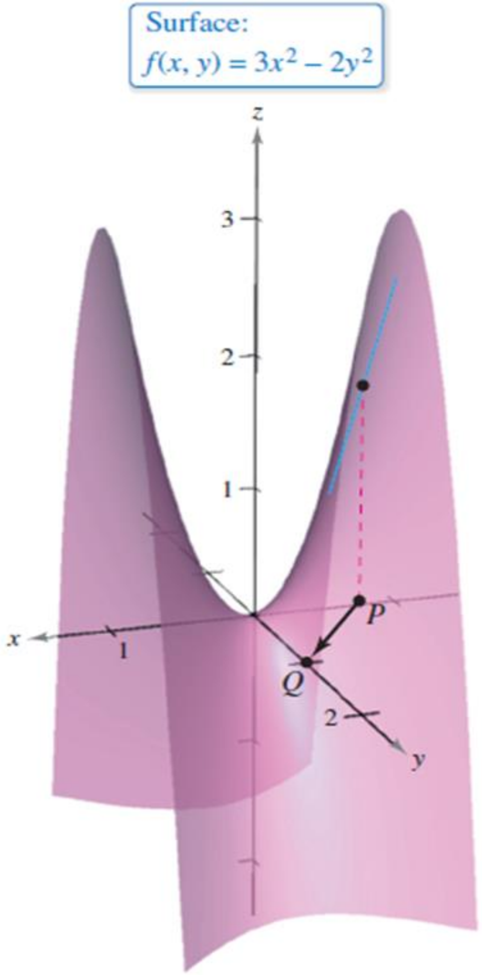


Figure due to Larson

Solution:

Since the partial derivatives of f are continuous, f is differentiable, then we can use the formula to calculate the directional derivative.

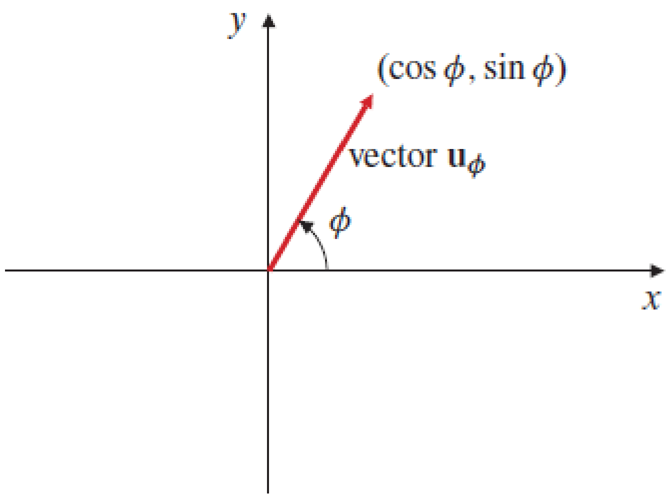
Consider the vector $\overrightarrow{PQ} = \left\langle \frac{3}{4}, 1 \right\rangle$, thus a unit vector in this direction is **u** = $\left\langle \frac{3}{5}, \frac{4}{5} \right\rangle$.

Since gradient of $f(x, y) = \langle 6x, -4y \rangle$, at the point $P\left(-\frac{3}{4}, 0\right)$, the gradient is $\left\langle -\frac{9}{2}, 0 \right\rangle$.

Thus, the directional derivative at *P* is $D_{\mathbf{u}} f\left(-\frac{3}{4}, 0\right) = \left\langle -\frac{9}{2}, 0 \right\rangle \cdot \left\langle \frac{3}{5}, \frac{4}{5} \right\rangle = -\frac{27}{10}$.

Geometric Properties of Gradients

Figure due to Adams and Essex



In the figure, $\mathbf{u}_\phi = \langle \cos\phi, \sin\phi \rangle$.
Directional derivative of f at (x, y) in this direction:

$$D_{\mathbf{u}_\phi} f(x, y) = f_x(x, y)\cos\phi + f_y(x, y)\sin\phi$$

The above notation means the derivative of f with respect to distance measured in the direction ϕ.

Now, for any unit vector $\mathbf{u}$, $D_{\mathbf{u}} f(a, b) = \mathbf{u} \cdot \nabla f(a, b) = \|\nabla f(a, b)\| \cos\theta$, where θ is the angle between the vectors $\mathbf{u}$ and $\nabla f(a, b)$.

Geometric properties of the gradient vector

(i) At (a, b), $f(x, y)$ increases most rapidly in the direction of the gradient vector $\nabla f(a, b)$. The maximum rate of increase is $\|\nabla f(a, b)\|$.

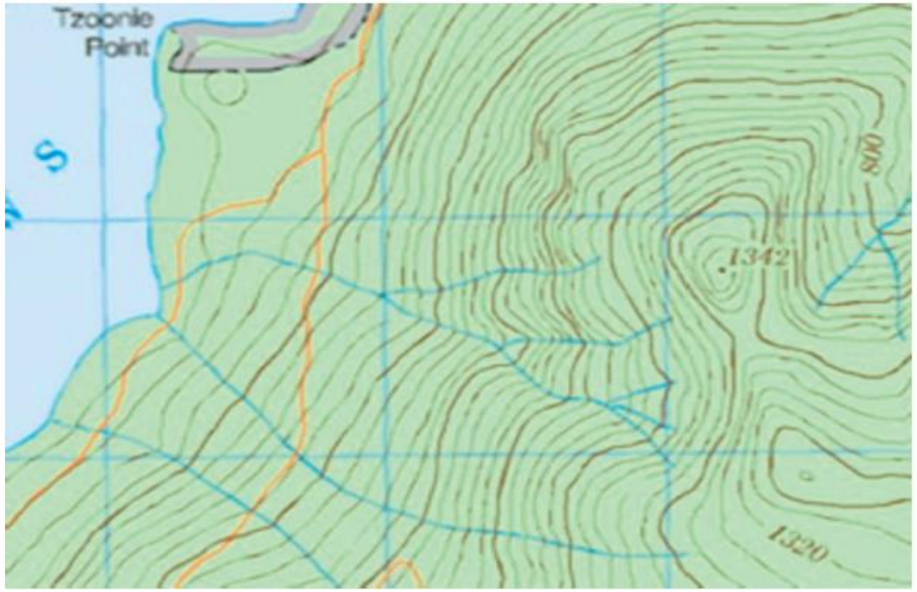
(ii) At (a, b), $f(x, y)$ decreases most rapidly in the direction of $-\nabla f(a, b)$. The maximum rate of decrease is $\|\nabla f(a, b)\|$.

(iii) The rate of change of $f(x, y)$ at (a, b) is zero in directions tangent to the level curve of f that passes through (a, b).

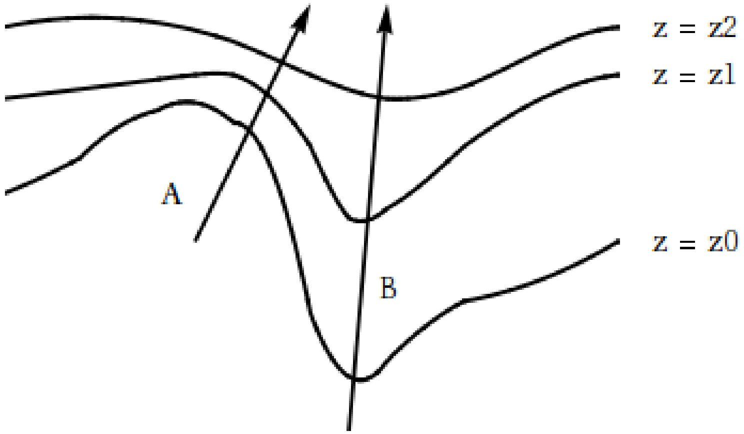
Daily Life Application of Gradients

- Consider the level curves (contour diagram) that represents elevation in a topographic map, where f represents the elevation of land.
- The streams on the map flow in the direction of steepest descent, i.e. in the direction of $-\nabla f$.
- The streams cross the level curves of f at right angles.
- A skier will tend to choose a downhill path that is close to the **direction of the negative gradient**.

Figure due to Adams and Essex



Example 1: The figure shows a contour map of a hill with two paths, A and B. The heights on the contour are $z = z_0 = 200, z = z_1 = 250$, and $z = z_2 = 300$.



(a) On which path, will you have to climb more steeply?

(b) On which path, will you probably have a better view of surrounding countryside (here we assume trees will NOT block your view)?

(c) Alongside which path is there more likely to be a stream?

(a) Path A – The contour lines are much closer together on path A.
(b) Path B – For Path A, if you turn around, you will find hills on your left and on your right, which obscures the view. However, the ground falls away on either side of Path B.
(c) Path A – since water follows the direction of steepest descent.

Daily Life Application of Gradients

Example 2: The temperature on the surface of a metal plate is

$$T(x, y) = 20 - 4x^2 - y^2$$

where x and y are measured in cm, and T is measured in degree Celsius.

(a) In what direction from (2, –3) does the temperature increase most rapidly?

(b) Find the rate of increase.

(c) Suppose a heat-seeking particle is located at (2,–3), find the path of the particle as it continuously moves in the direction of maximum temperature increase.

Solution:

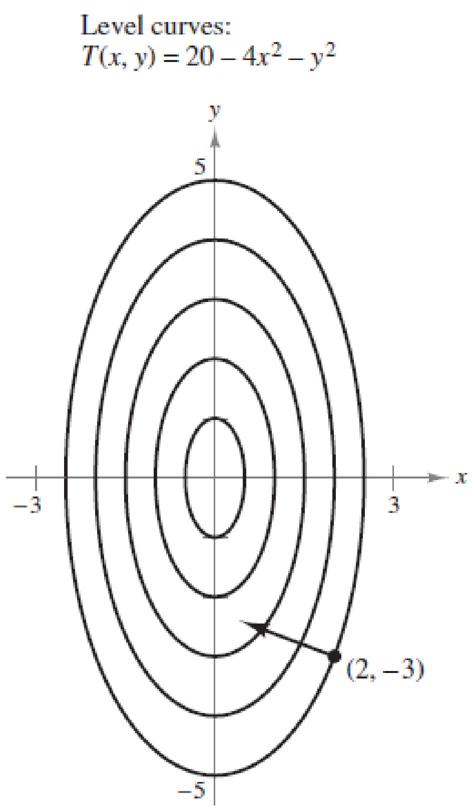
(a) The gradient is

$$\nabla T(x, y) = T_x(x, y)\mathbf{i} + T_y(x, y)\mathbf{j} = -8x\mathbf{i} - 2y\mathbf{j}.$$

It follows that the direction of maximum increase is given by $\nabla T(2, -3) = -16\mathbf{i} + 6\mathbf{j}$.

(b) Rate of increase = norm of gradient vector $= \sqrt{256 + 36} = \sqrt{292}$

≈ 17.09 degree Celsius per cm.



The direction of most rapid increase in temperature at $(2, -3)$ is given by $-16\mathbf{i} + 6\mathbf{j}$.

Figure due to Larson

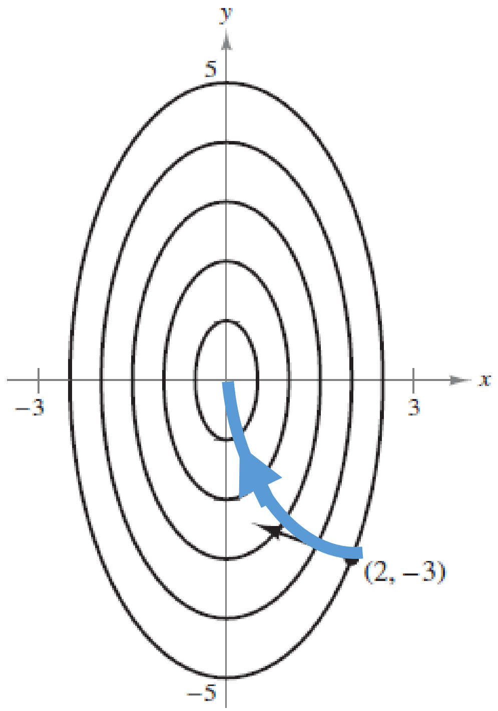
Daily Life Application of Gradients

Example 2: The temperature on the surface of a metal plate is

$$T(x, y) = 20 - 4x^2 - y^2$$

where x and y are measured in cm, and T is measured in degree Celsius.

(c) Suppose a heat-seeking particle is located at (2,–3), find the path of the particle as it continuously moves in the direction of maximum temperature increase.



The direction of most rapid increase in temperature at $(2, -3)$ is given by $-16\mathbf{i} + 6\mathbf{j}$.

Figure due to Larson

Solution:

(c) Let the path be $\mathbf{r}(t) = \langle x(t), y(t)\rangle$, a tangent vector at each point $(x(t), y(t))$ is given by $\mathbf{r}'(t) = \left\langle \frac{dx}{dt}, \frac{dy}{dt} \right\rangle$.

Since the particle seeks maximum temperature increase, the tangent vector is in the same direction as the gradient of $T = \langle -8x, -2y\rangle$ at every point on the required path.

Hence, $\begin{cases} -8x = C\dfrac{dx}{dt} \\ -2y = C\dfrac{dy}{dt} \end{cases}$ (C will depend on t), so $\frac{dx}{-8x} = \frac{dy}{-2y}$, i.e. $x = Ky^4$ (**How**?)

Since the particle starts at the point $(2, -3)$, $K = \frac{2}{81}$.

Therefore, the path of the heat-seeking particle is $x = \frac{2}{81}y^4$.

Directional Derivative and Gradient for 3 variables

Directional Derivative and Gradient for Three Variables

Let f be a function of x, y, and z with continuous first partial derivatives. The **directional derivative of f** in the direction of a unit vector

$$\mathbf{u} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

is given by

$$D_{\mathbf{u}}f(x, y, z) = af_x(x, y, z) + bf_y(x, y, z) + cf_z(x, y, z).$$

The **gradient of f** is defined as

$$\nabla f(x, y, z) = f_x(x, y, z)\mathbf{i} + f_y(x, y, z)\mathbf{j} + f_z(x, y, z)\mathbf{k}.$$

Properties of the gradient are as follows.

1. $D_{\mathbf{u}}f(x, y, z) = \nabla f(x, y, z) \cdot \mathbf{u}$
2. If $\nabla f(x, y, z) = \mathbf{0}$, then $D_{\mathbf{u}}f(x, y, z) = 0$ for all $\mathbf{u}$.
3. The direction of *maximum* increase of f is given by $\nabla f(x, y, z)$. The maximum value of $D_{\mathbf{u}}f(x, y, z)$ is

   $\|\nabla f(x, y, z)\|$. Maximum value of $D_{\mathbf{u}}f(x, y, z)$

4. The direction of *minimum* increase of f is given by $-\nabla f(x, y, z)$. The minimum value of $D_{\mathbf{u}}f(x, y, z)$ is

   $-\|\nabla f(x, y, z)\|$. Minimum value of $D_{\mathbf{u}}f(x, y, z)$

If f is differentiable at (x_0, y_0, z_0) and $\nabla f(x_0, y_0, z_0) \neq \mathbf{0}$, then $\nabla f(x_0, y_0, z_0)$ is normal to the level surface through (x_0, y_0, z_0).

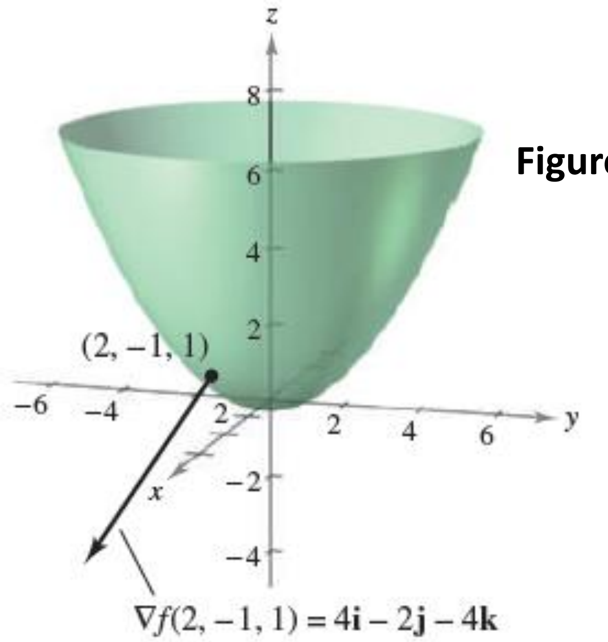


Figure due to Larson

Level surface and gradient at $(2, -1, 1)$ for $f(x, y, z) = x^2 + y^2 - 4z$

Tangent Planes

- Consider the representation of surfaces *S* in space: $z = f(x, y)$.
- Now, we wish to use another representation: $F(x, y, z) = 0$, so we may define
$$F(x, y, z) = f(x, y) - z$$
- Since $f(x, y) - z = 0$, we may consider S to be the level surface of F given by $F(x, y, z) = 0$.

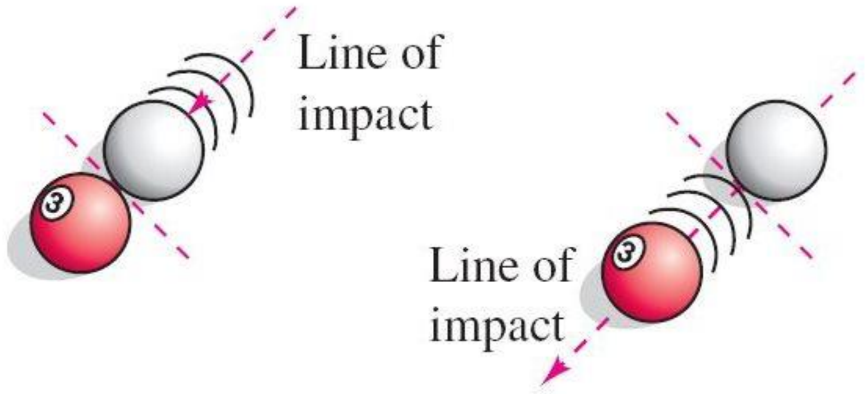
Now, we consider the collision of two billiard balls. When a stationary ball is struck at a point *P* on its surface, it moves along the **line of impact** determined by *P* and the center of the ball.

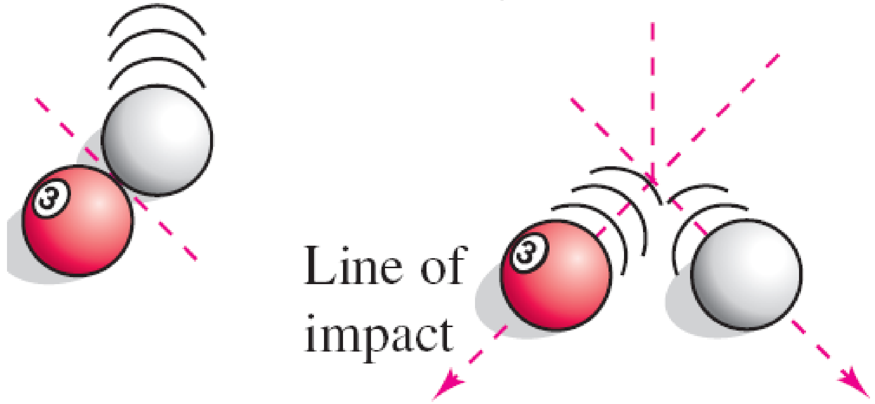
(1) When the cue ball is <u>moving along the line of impact</u>, it stops dead and imparts all of its momentum to the stationary ball.

(2) When the cue ball is <u>not moving along the line of impact</u>, it is deflected to one side or the other and retains part of its momentum. The part of the momentum being transferred to the stationary ball occurs along the line of impact, <u>*regardless*</u> of the direction of the cue ball.

The line of impact is called the "**normal line**" to the surface of the ball at *P*.

Figures due to Larson





Tangent Planes

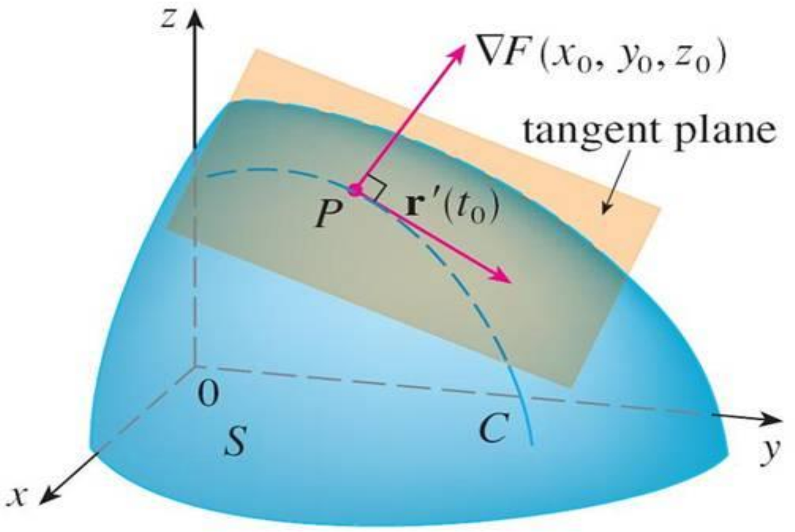
We suppose S is a surface with equation $F(x, y, z) = k$, that is, it is a level surface of a function F of three variables, and let $P(x_0, y_0, z_0)$ be a point on S. Let C be any curve that lies on the surface S and passes through the point P. The curve C is described by a continuous vector function $\mathbf{r}(t) = \langle x(t), y(t), z(t)\rangle$. Let t_0 be the parameter value corresponding to P; that is, $\mathbf{r}(t_0) = \langle x_0, y_0, z_0\rangle$.

Since C lies on S, any point $(x(t), y(t), z(t))$ must satisfy the equation of S, that is, $F(x(t), y(t), z(t)) = k$ ……(*)
Given that x, y, and z are differentiable functions of t and F is differentiable, we can use Chain Rule to differentiate both sides of (*),

$$0 = F'(t) = F_x(x, y, z)x'(t) + F_y(x, y, z)y'(t) + F_z(x, y, z)z'(t)$$

At (x_0, y_0, z_0), the equivalent vector form is

$$\nabla F(x_0, y_0, z_0) \cdot \mathbf{r}'(t_0) = 0$$



(i.e. the dot product of gradient vector and the tangent vector = 0)

Conclusion:

The gradient vector at P, $\nabla F(x_0, y_0, z_0)$, is **perpendicular** to the tangent vector $\mathbf{r}'(t_0)$ to any curve C on S that passes through P.

Tangent Planes

Definitions of Tangent Plane and Normal Line

Let F be differentiable at the point $P(x_0, y_0, z_0)$ on the surface S given by $F(x, y, z) = 0$ such that

$$\nabla F(x_0, y_0, z_0) \neq \mathbf{0}.$$

1. The plane through P that is normal to $\nabla F(x_0, y_0, z_0)$ is called the **tangent plane to S at P.**
2. The line through P having the direction of $\nabla F(x_0, y_0, z_0)$ is called the **normal line to S at P.**

Due to Larson (Page 932)

Using the standard equation of a plane, we can write the equation of this tangent plane as

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

The direction of the normal line is therefore given by the gradient vector $\nabla F(x_0, y_0, z_0)$ and so, its symmetric equations are

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

Tangent Planes

The graph of a function $f(x, y)$ of two variables is the graph of the surface $z = f(x, y)$ in 3-dimensional space. Such surface corresponds to the level surface $g(x, y, z) = 0$, where

$$g(x, y, z) = f(x, y) - z$$

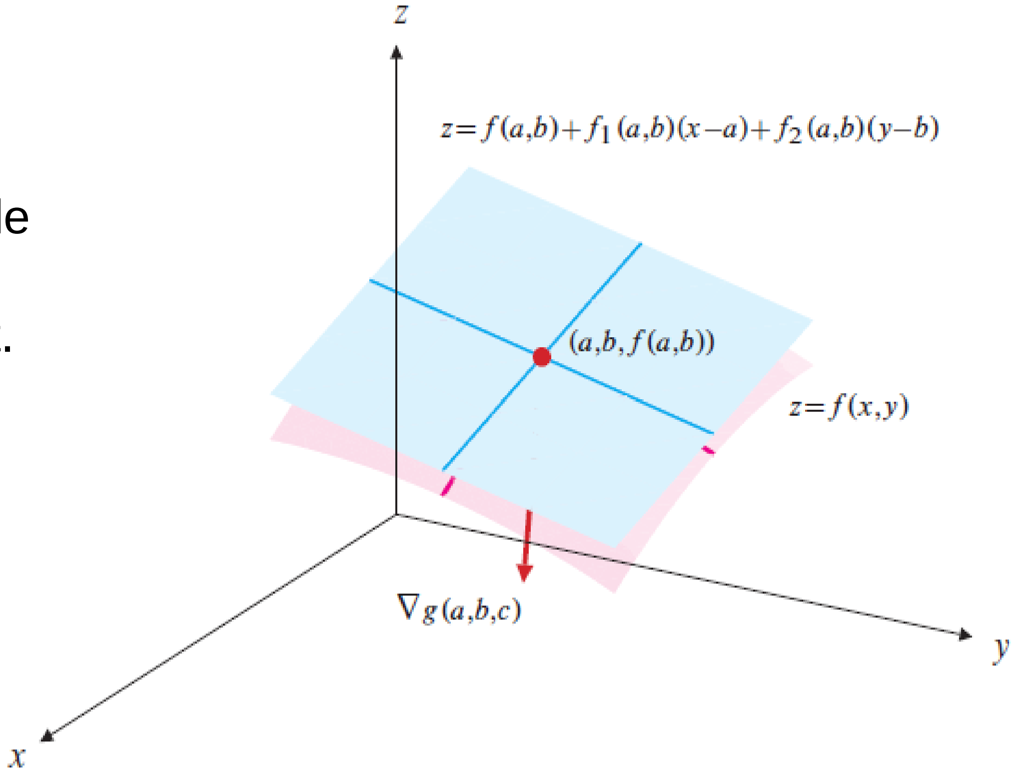
If f is differentiable at (a, b), and $c = f(a, b)$, then g is differentiable at (a, b, c), and gradient of g at the point (a, b, c) is the vector $\langle f_x(a, b), f_y(a, b), -1\rangle$, which is normal to $g(x, y, z) = 0$ at this point.

(Note that $\nabla g(a, b, c) \neq \mathbf{0}$, it follows that the graph of f has a non-vertical tangent plane at (a, b), and is given by:

$$f_x(a, b)(x - a) + f_y(a, b)(y - b) - (z - c) = 0$$

i.e. $z = c + f_x(a, b)(x - a) + f_y(a, b)(y - b)$.

Figure due to Adams and Essex



Tangent Planes

Example 1. Find the equations of the tangent plane and normal line at the point (–2, 1, –3) to the ellipsoid $\frac{x^2}{4} + y^2 + \frac{z^2}{9} = 3$.

Solution

The ellipsoid is the level surface (with $k = 3$) of the function $F(x, y, z) = \frac{x^2}{4} + y^2 + \frac{z^2}{9} = 3$. Therefore we have

$F_x(x, y, z) = \frac{x}{2}$ $\quad$ $F_y(x, y, z) = 2y$ $\quad$ $F_z(x, y, z) = \frac{2z}{9}$

$F_x(-2, 1, -3) = -1$ $\quad$ $F_y(-2, 1, -3) = 2$ $\quad$ $F_z(-2, 1, -3) = -\frac{2}{3}$

Hence, the equation of the tangent plane at (–2, 1, –3) as

$$-1(x + 2) + 2(y - 1) - \frac{2}{3}(z + 3) = 0$$

which simplifies to $3x - 6y + 2z + 18 = 0$.

The symmetric equations of the normal line are $\frac{x+2}{-1} = \frac{y-1}{2} = \frac{z+3}{-\frac{2}{3}}$.

Tangent Planes

Example 2. Find the horizontal plane that is tangent to the surface $z = x^2 - 4xy - 2y^2 + 12x - 12y - 1$. Find also the point of tangency.

<u>Solution</u>

A plane is horizontal if its equation can be expressed in the form $z = K$, i.e. z is independent of both x and y. Hence, at the point of tangency, $z_x = z_y = 0$.

Consider $z_x = 2x - 4y + 12 = 0$ and $z_y = -4x - 4y - 12 = 0$.

Solving the equations, we have $x = -4, y = 1$. Therefore $z = -31$.

The required tangent plane has equation $z = -31$, and the point of tangency is $(-4, 1, -31)$.

Tangent Planes

Example 3. The two surfaces $f(x,y,z) = x^2 + y^2 - 2 = 0$ and $g(x,y,z) = x + z - 4 = 0$ meet in an ellipse E . Find the parametric equations for the line that is tangent to E at $(1,1,3)$.

<u>**Solution (Based on Thomas Calculus)**</u>

Solution The tangent line is orthogonal to both ∇f and ∇g at P_0, and therefore parallel to $\mathbf{v} = \nabla f \times \nabla g$. The components of $\mathbf{v}$ and the coordinates of P_0 give us equations for the line. We have

$$\nabla f|_{(1,1,3)} = (2x\mathbf{i} + 2y\mathbf{j})_{(1,1,3)} = 2\mathbf{i} + 2\mathbf{j}$$

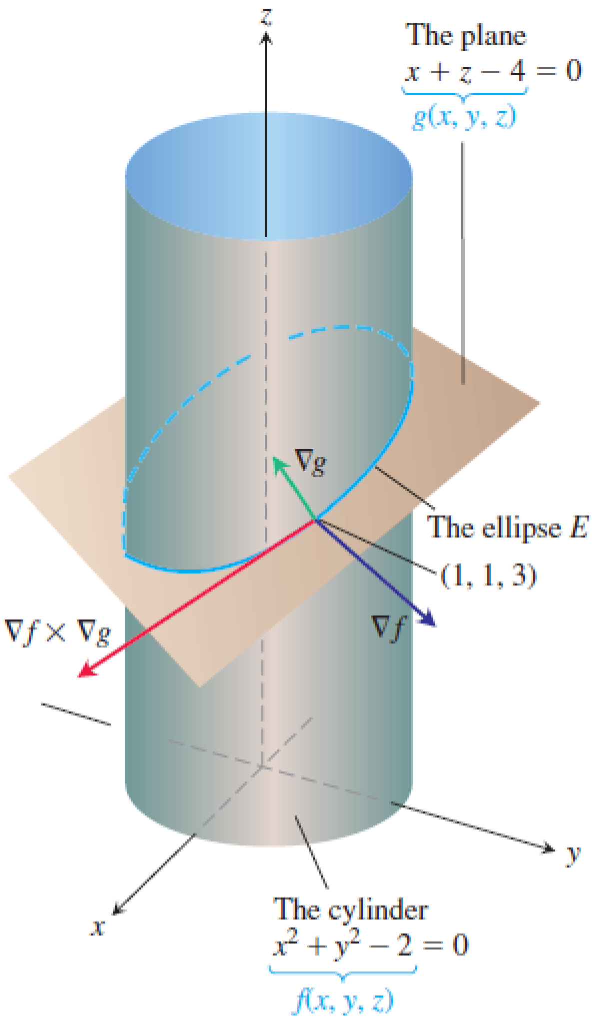
$$\nabla g|_{(1,1,3)} = (\mathbf{i} + \mathbf{k})_{(1,1,3)} = \mathbf{i} + \mathbf{k}$$

$$\mathbf{v} = (2\mathbf{i} + 2\mathbf{j}) \times (\mathbf{i} + \mathbf{k}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{vmatrix} = 2\mathbf{i} - 2\mathbf{j} - 2\mathbf{k}.$$

The tangent line to the ellipse of intersection is

$$x = 1 + 2t, \qquad y = 1 - 2t, \qquad z = 3 - 2t.$$

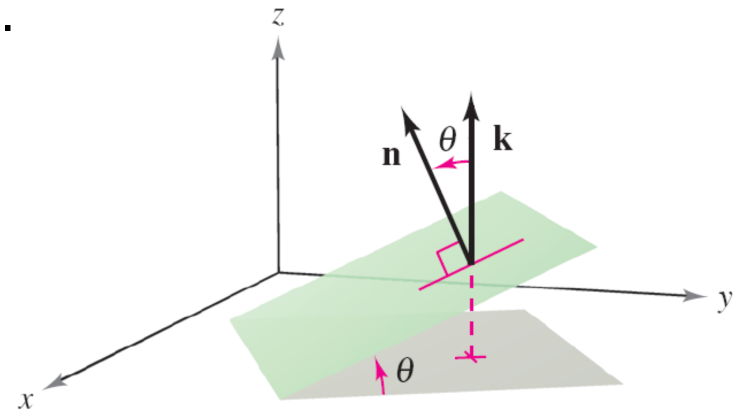
Figure due to Thomas Calculus



Angle of Inclination of a Plane

The **angle of inclination** of a plane is defined as the angle θ ($0 \le \theta \le \pi/2$) between the given plane and the *xy*-plane, as shown in the following figure.

Note: The angle of inclination of a horizontal plane is defined as zero.



The angle of inclination

Figure due to Larson

Since the vector **k** is normal to the *xy*-plane, we can use the formula for the cosine of the angle between two planes.

Then, we can conclude that the angle of inclination of a plane with normal vector **n** is

$$\cos\theta = \frac{|\mathbf{n}\cdot\mathbf{k}|}{\|\mathbf{n}\|\,\|\mathbf{k}\|} = \frac{|\mathbf{n}\cdot\mathbf{k}|}{\|\mathbf{n}\|}.$$

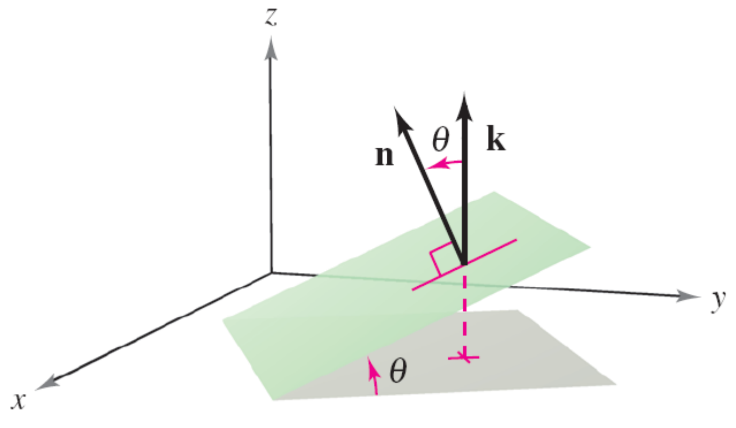
Angle of inclination of a plane

Angle of Inclination of a Plane

Example: Find the angle of inclination of the tangent plane to the ellipsoid

$$\frac{x^2}{12} + \frac{y^2}{12} + \frac{z^2}{3} - 1 = 0$$

at the point (2, 2, 1).



The angle of inclination

Figures due to Larson

Solution

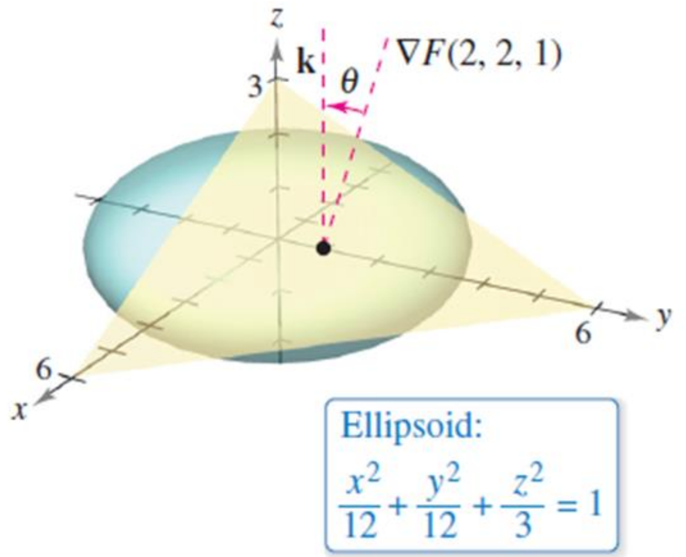
We let $F(x, y, z) = \frac{x^2}{12} + \frac{y^2}{12} + \frac{z^2}{3} - 1 = 0$ be a level surface.

Then the gradient of F at (2, 2, 1) is $\left\langle \frac{1}{3}, \frac{1}{3}, \frac{2}{3} \right\rangle$. (**How?**)

Since the gradient vector at (2, 2, 1) is normal to the tangent plane and **k** is normal to the xy-plane, the angle of inclination of the tangent plane is

$$\cos\theta = \frac{|\nabla F(2, 2, 1) \cdot \mathbf{k}|}{\|\nabla F(2, 2, 1)\|}$$

By computation, $\cos\theta = \frac{\sqrt{6}}{3}$, hence $\theta \approx 35.3^{\circ}$.



Extrema of Multivariable Functions

Recall the **Extreme Value Theorem**:

Let f be a continuous function of two variables x and y defined on a closed bounded region R in the xy-plane.

1. There is at least one point in the region at which f is minimum.
2. There is at least one point in the region at which f is maximum.

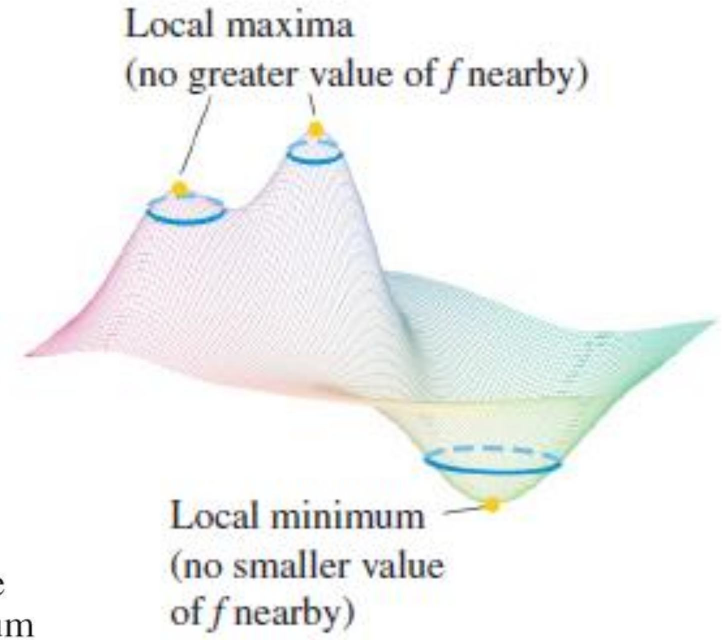


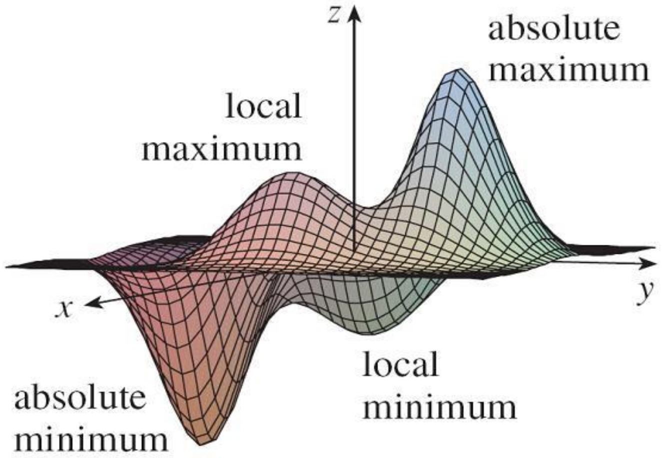
Figure from Thomas Calculus

Local (Relative) Maximum (or Minimum)

DEFINITIONS Let $f(x, y)$ be defined on a region R containing the point (a, b). Then

1. $f(a, b)$ is a **local maximum** value of f if $f(a, b) \geq f(x, y)$ for all domain points (x, y) in an open disk centered at (a, b).
2. $f(a, b)$ is a **local minimum** value of f if $f(a, b) \leq f(x, y)$ for all domain points (x, y) in an open disk centered at (a, b).

A minimum is also called an **absolute minimum** and a maximum is also called an **absolute maximum.**
There is a distinction made between absolute extrema and **relative extrema.**



Extrema of Multivariable Functions

The local or absolute extreme values can occur only at points of one of the following 3 types:

(a) Critical Points (where $f'(x) = 0$ for 1D case, and $f_x(x, y) = f_y(x, y) = 0$ for 2D case and etc.)

(b) Singular points (where $f'(x)$ does **NOT** exist for 1D case, and $\text{gradient of } f$ does **NOT** exist for 2D case and etc.)

(c) Endpoints of the domain of f in 1D, or Boundary Points of the domain of f in 2D and etc.

Proof (due to Adams and Essex)

PROOF Suppose that (a, b) belongs to the domain of f. If (a, b) is not on the boundary of the domain of f, then it must belong to the interior of that domain, and if (a, b) is not a singular point of f, then $\nabla f(a, b)$ exists. Finally, if (a, b) is not a critical point of f, then $\nabla f(a, b) \neq \mathbf{0}$, so f has a positive directional derivative in the direction of $\nabla f(a, b)$ and a negative directional derivative in the direction of $-\nabla f(a, b)$; that is, f is increasing as we move from (a, b) in one direction and decreasing as we move in the opposite direction. Hence, f cannot have either a maximum or a minimum value at (a, b). Therefore, any point where an extreme value occurs must be either a critical point or a singular point of f, or a boundary point of the domain of f.

Extrema of Multivariable Functions

Theorem (First Derivative Test for Local Extreme values)

If $f(x, y)$ has a local (relative) maximum or minimum value at an interior point (a, b) of its domain, and if the first partial derivatives exist there, then $f_x(a, b) = 0$ and $f_y(a, b) = 0$.

Proof. If f has a local (relative) extremum at the point (a, b), then $g(x) = f(x, b)$ has a local (relative) extremum at $x = a$, thus $g'(a) = 0$.

Since $g'(a) = f_x(a, b)$, we conclude that $f_x(a, b) = 0$.

Similar arguments can be applied to the function $h(y) = f(a, y)$, then we conclude that $f_y(a, b) = 0$.

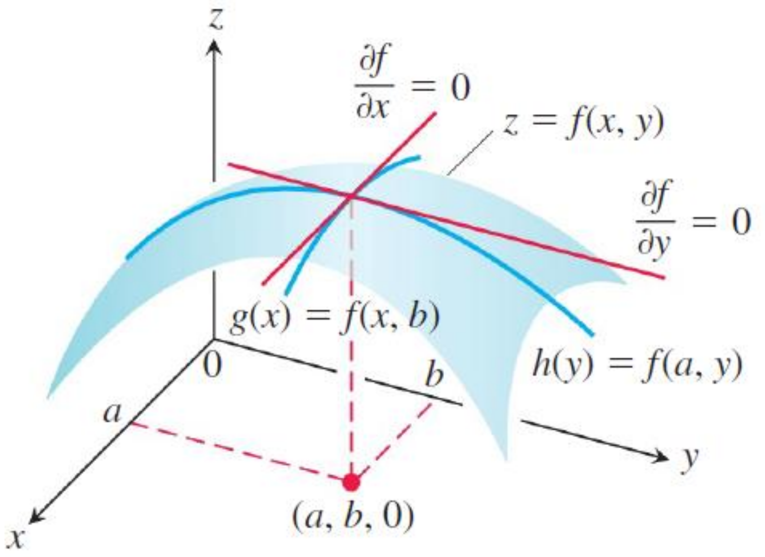


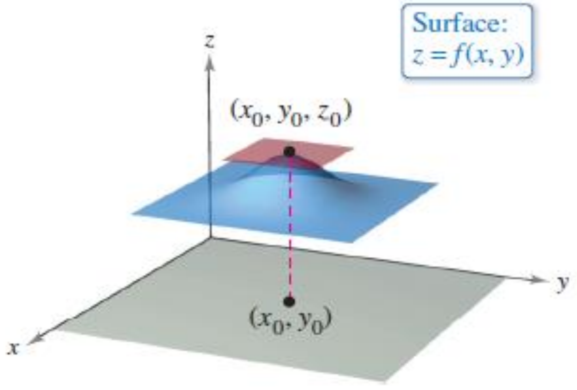
Figure from Thomas Calculus

Note: If we substitute $f_x(a, b) = 0$ and $f_y(a, b) = 0$ into the equation of tangent plane to the surface $z = f(x, y)$ at the point (a, b), the equation can be reduced to $z = f(a, b)$. Hence, the surface has a horizontal tangent plane at local (relative) extremum, provided that there exists a tangent plane there.

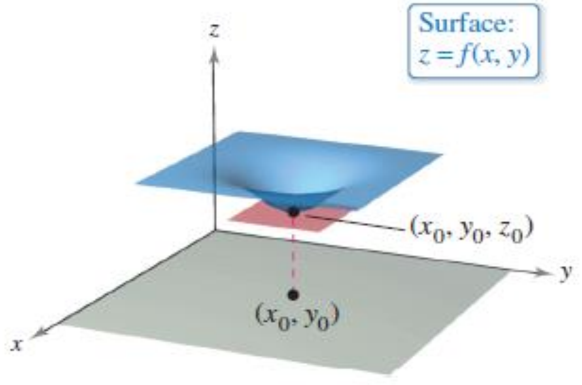
Definition (Critical Point): An interior point of the domain of a function $f(x, y)$ with both f_x and f_y being zero, or where one of the two partial derivatives do not exist is a **critical point** of f.

Extrema of Multivariable Functions

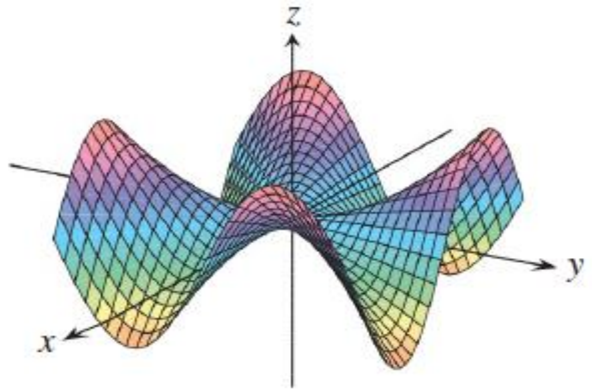
Figures from Thomas Calculus



Relative maximum



Relative minimum

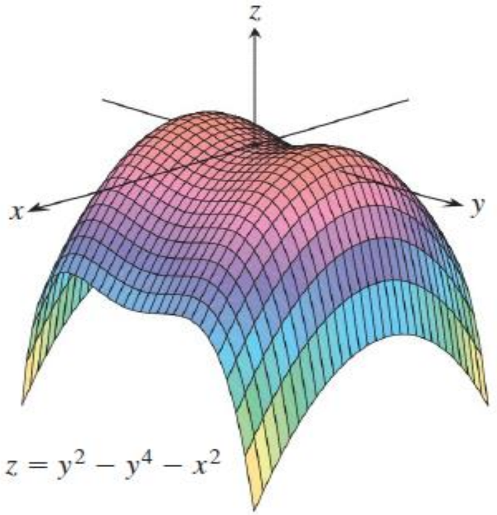


$z = \frac{xy(x^2 - y^2)}{x^2 + y^2}$

The previous theorem implicates that the only points where a function $f(x, y)$ can assume extreme values are either "**critical points**" or "**boundary points**".

<u>Think about 1D case</u>: A differentiable function of a single variable may have point of inflection.

<u>Extend such thought to 2D or higher dimensional case</u>: A differentiable function of two (or more) variables may have a saddle point, with the graph of f crossing the tangent plane defined at that point.



$z = y^2 - y^4 - x^2$

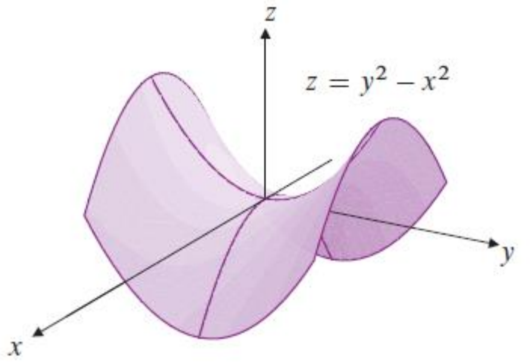
Definition (Saddle Point): A differentiable function $f(x, y)$ has a **saddle point** at a critical point (a, b), if in every open disk centered at (a, b), there exist domain points (x, y) such that $f(x, y) > f(a, b)$, and domain points (x, y) such that $f(x, y) < f(a, b)$. The corresponding point $\big(a, b, f(a, b)\big)$ on the surface $z = f(x, y)$ is called a saddle point of the surface.

Examples

(1) The function $f(x, y) = y^2 - x^2$ has a critical point at (0, 0), but has neither a local maximum nor a local minimum value at that point.

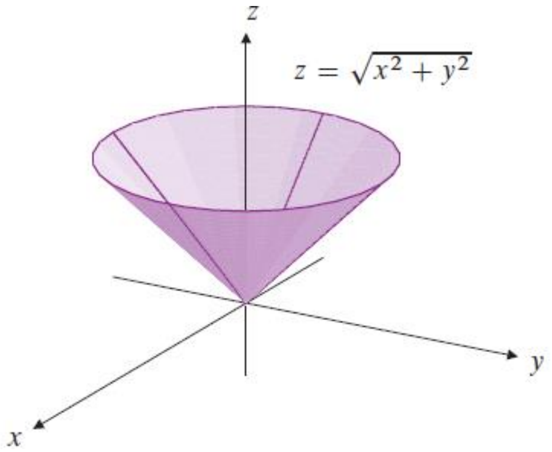
We observe that $f(0,0) = 0$, but $f(0, y) > 0$ and $f(x, 0) < 0$ for all $x, y \neq 0$.

Note that the graph on the right is a hyperbolic paraboloid, i.e. the critical point (0, 0) is a **saddle point** of f.



Figures from Adam and Essex

(2) The function $g(x, y) = \sqrt{x^2 + y^2}$ has no critical points, but has a singular point (0, 0), where it has a **local (relative) minimum value** 0. Such minimum is also a global (absolute) minimum. The graph of g is a circular cone.



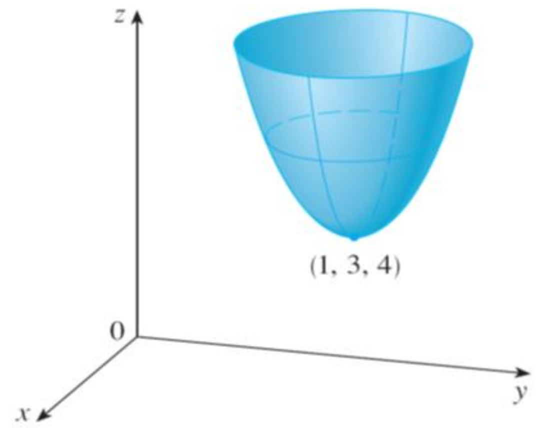
(3) Let $h(x, y) = x^2 + y^2 - 2x - 6y + 14$. Then

$$h_x(x, y) = 2x - 2 \qquad h_y(x, y) = 2y - 6$$

These partial derivatives are equal to 0 when $x = 1$ and $y = 3$, so the only critical point is (1, 3). By completing the square, we find that

$$h(x, y) = 4 + (x - 1)^2 + (y - 3)^2$$

Since $(x - 1)^2 \geq 0$ and $(y - 3)^2 \geq 0$, we have $h(x, y) \geq 4$ for all values of x and y. Therefore $h(1, 3) = 4$ is a **local (relative) minimum**, and in fact it is the **global (absolute) minimum** of h. The graph of h is the elliptic paraboloid with vertex (1, 3, 4).



Second Derivative Test (2 variables)

Theorem (Second Derivative Test for Local Extreme Values) Suppose $f(x, y)$ and its first and second partial derivatives are continuous throughout a disk centered at (a, b), and $f_x(a, b) = f_y(a, b) = 0$.

Then we have the following classification:

(1) f has a **local (relative) maximum** at (a, b) if $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at the point (a, b).

(2) f has a **local (relative) minimum** at (a, b) if $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at the point (a, b).

(3) f has a **saddle point** at (a, b) if $f_{xx}f_{yy} - f_{xy}^2 < 0$ at the point (a, b).

(4) The test is **inconclusive** at (a, b) if $f_{xx}f_{yy} - f_{xy}^2 = 0$ at the point (a, b). In that case, we need some other ways to determine the behavior of f at (a, b).

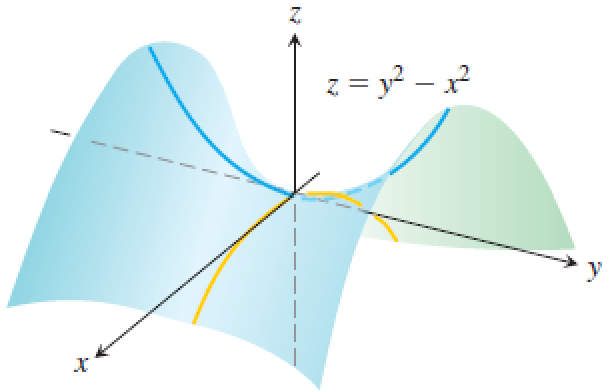
The expression $f_{xx}f_{yy} - f_{xy}^2$ is called the **Hessian** of f. In discriminant form, it is $\begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix}$.

If the **Hessian** is positive at the point (a, b), then the surface curves in the same way in all directions:

Downward if $f_{xx} < 0$ → Local Maximum

Upward if $f_{xx} > 0$ → Local Minimum

If the **Hessian** is negative at the point (a, b), then the surface curves up in some directions and curves down in other directions, then we have a saddle point.



Example 1 (due to Thomas Calculus)

Find the critical points of the function $f(x,y) = 10xye^{-(x^2+y^2)}$, then use the Second Derivative Test to classify each of these critical points.

Solution First we find the partial derivatives f_x and f_y and set them simultaneously to zero in seeking the critical points:

$$f_x = 10ye^{-(x^2+y^2)} - 20x^2ye^{-(x^2+y^2)} = 10y(1-2x^2)e^{-(x^2+y^2)} = 0 \Rightarrow y = 0 \text{ or } 1 - 2x^2 = 0,$$

$$f_y = 10xe^{-(x^2+y^2)} - 20xy^2e^{-(x^2+y^2)} = 10x(1-2y^2)e^{-(x^2+y^2)} = 0 \Rightarrow x = 0 \text{ or } 1 - 2y^2 = 0.$$

Since both partial derivatives are continuous everywhere, the only critical points are

$$(0,0), \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \text{ and } \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right).$$

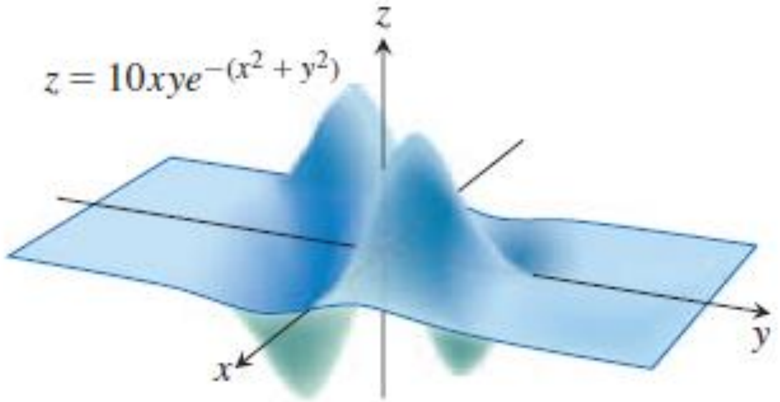
Next we calculate the second partial derivatives in order to evaluate the discriminant at each critical point:

$$f_{xx} = -20xy(1-2x^2)e^{-(x^2+y^2)} - 40xye^{-(x^2+y^2)} = -20xy(3-2x^2)e^{-(x^2+y^2)},$$

$$f_{xy} = f_{yx} = 10(1-2x^2)e^{-(x^2+y^2)} - 20y^2(1-2x^2)e^{-(x^2+y^2)} = 10(1-2x^2)(1-2y^2)e^{-(x^2+y^2)},$$

$$f_{yy} = -20xy(1-2y^2)e^{-(x^2+y^2)} - 40xye^{-(x^2+y^2)} = -20xy(3-2y^2)e^{-(x^2+y^2)}.$$

Critical Point	f_{xx}	f_{xy}	f_{yy}	Discriminant D
$(0, 0)$	0	10	0	-100
$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$	$-\frac{20}{e}$	0	$-\frac{20}{e}$	$\frac{400}{e^2}$
$\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$	$\frac{20}{e}$	0	$\frac{20}{e}$	$\frac{400}{e^2}$
$\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$	$\frac{20}{e}$	0	$\frac{20}{e}$	$\frac{400}{e^2}$
$\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$	$-\frac{20}{e}$	0	$-\frac{20}{e}$	$\frac{400}{e^2}$



Saddle point: $(0, 0)$

Local (relative) maximum points: $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $\left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$

Local (relative) minimum points: $\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$, $\left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$

Example 2

A manufacturer determines that the profit P (in dollars) obtained by producing and selling x units of Product A and y units of Product B is given by the following mathematical model:

$$P(x, y) = 8x + 10y - 10000 - 0.001x^2 - 0.001xy - 0.001y^2$$

Find the production level that will lead to the maximum profit. Find the maximum profit as well.

Solution

$$P_x(x, y) = 8 - 0.002x - 0.001y$$
$$P_y(x, y) = 10 - 0.001x - 0.002y$$

Setting both of these partial derivatives to be 0, we have $x = 2000, y = 4000$.
Consider the 2nd partial derivatives are given by $P_{xx}(2000{,}4000) = -0.002,\ P_{yy}(2000{,}4000) = -0.002,$ $P_{xy}(2000{,}4000) = -0.001$. Since $P_{xx} < 0$, and the Hessian matrix is also greater than 0, we conclude that the production level of $x = 2000$ and $y = 4000$ will yield the maximum profit.

The maximum profit is given by $P(2000{,}4000) = \$18{,}000$

Remarks

In the example, we assume that the manufacturing plant is able to produce the required amount of units of Product A and Product B to obtain a maximum profit. In reality, the production will be restricted by some physical constraints. This is called to be the “**constrained optimization**” problems.

Generalized Second Derivative Test

Refer to Page 617, 755 – 756 of Adams and Essex

Theorem

A second derivative test

Suppose that $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ is a critical point of $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$ and is interior to the domain of f. Also, suppose that all the second partial derivatives of f are continuous throughout a neighbourhood of $\mathbf{a}$, so that the **Hessian matrix**

$$\mathcal{H}(\mathbf{x}) = \begin{pmatrix} f_{11}(\mathbf{x}) & f_{12}(\mathbf{x}) & \cdots & f_{1n}(\mathbf{x}) \\ f_{21}(\mathbf{x}) & f_{22}(\mathbf{x}) & \cdots & f_{2n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1}(\mathbf{x}) & f_{n2}(\mathbf{x}) & \cdots & f_{nn}(\mathbf{x}) \end{pmatrix}$$

is also continuous in that neighbourhood. Note that the continuity of the partials guarantees that $\mathcal{H}$ is a symmetric matrix.

(a) If $\mathcal{H}(\mathbf{a})$ is positive definite, then f has a local minimum at $\mathbf{a}$.

(b) If $\mathcal{H}(\mathbf{a})$ is negative definite, then f has a local maximum at $\mathbf{a}$.

(c) If $\mathcal{H}(\mathbf{a})$ is indefinite, then f has a saddle point at $\mathbf{a}$.

(d) If $\mathcal{H}(\mathbf{a})$ is neither positive nor negative definite, nor indefinite, this test gives no information.

Let $\mathcal{A} = (a_{ij})_{i,j=1}^{n}$ be a real symmetric matrix and consider the determinants

$$D_i = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{21} & a_{22} & \cdots & a_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} \end{vmatrix} \qquad \text{for } 1 \le i \le n.$$

Thus, $D_1 = a_{11}$, $D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} = a_{11}a_{22} - a_{12}^2$, etc.

(a) If $D_i > 0$ for $1 \le i \le n$, then $\mathcal{A}$ is positive definite.

(b) If $D_i > 0$ for even numbers i in $\{1, 2, \ldots, n\}$, and $D_i < 0$ for odd numbers i in $\{1, 2, \ldots, n\}$, then $\mathcal{A}$ is negative definite.

(c) If $\det(\mathcal{A}) = D_n \neq 0$ but neither of the above conditions hold, then $Q(\mathbf{x})$ is indefinite.

(d) If $\det(\mathcal{A}) = 0$, then $\mathcal{A}$ is not positive or negative definite and may be semidefinite or indefinite.

Examples

(1) Let $\mathbf{A} = \begin{pmatrix} 3 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & 1 & 5 \end{pmatrix}$, show that **A** is positive definite.

Consider $D_1 = 3 > 0$, $D_2 = (3)(2) - (-1)(-1) = 5 > 0$, $D_3 = \det(\mathbf{A}) = 10 > 0$. By definition, **A** is positive definite.

(2) For $f(x,y,z) = x^2y + y^2z + z^2 - 2x$. Consider

$$f_x(x,y,z) = 2xy - 2, \qquad f_y(x,y,z) = x^2 + 2yz, \qquad f_z(x,y,z) = y^2 + 2z$$

For critical points, we have $f_x = f_y = f_z = 0$, i.e. $z = -\frac{y^2}{2}$, then the second equation implies $y^3 = x^2$. Plugging all into the first equation, we have $y^{\frac{5}{2}} = 1$. As we assume all x, y, z are real, $y = 1, z = -\frac{1}{2}, x = 1$.

Hence, the only critical point of f is $P = \left(1,1,-\frac{1}{2}\right)$. We write down the **Hessian matrix** as follows:

$$\mathbf{H} = \begin{pmatrix} 2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 2 \end{pmatrix}$$

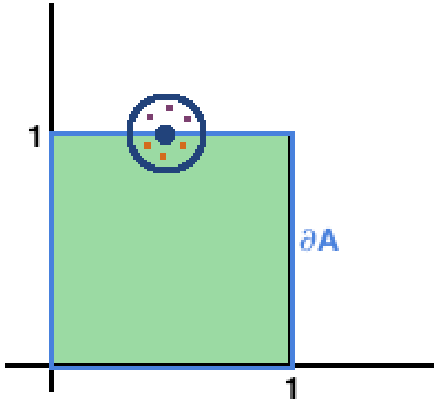
Consider $D_1 = 2 > 0, D_2 = 2(-1) - 2(2) = -6 < 0, \det(\mathbf{H}) = -20 < 0$

By the theorem, *H* is indefinite, therefore *P* is a saddle point of f.

Absolute Maxima and Minima on Closed and Bounded Regions

Steps for searching absolute extrema of a continuous function f on a closed and bounded region R.

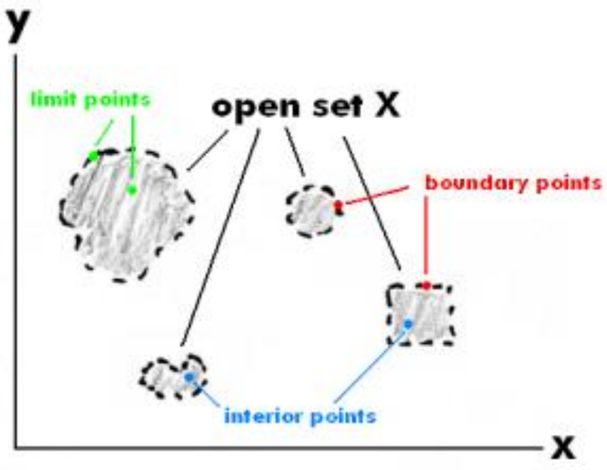
1. *List the interior points of* R where f may have local maxima and minima and evaluate f at these points. These are the critical points of f.
2. *List the boundary points of* R where f has local maxima and minima and evaluate f at these points. We show how to do this in the next example.
3. *Look through the lists* for the maximum and minimum values of f. These will be the absolute maximum and minimum values of f on R. Since absolute maxima and minima are also local maxima and minima, the absolute maximum and minimum values of f appear somewhere in the lists made in Steps 1 and 2.



For every point on the boundary of A, every open neighbourhood of that point contains points in A and points in A^C.

Definition (Closed set): Just as a closed interval contains its endpoints, a **closed set** in $\mathbf{R}^2$ is one that contains all its boundary points.

A **boundary point** of *R* is a point (*a*, *b*) such that every disk with center (*a*, *b*) contains points in *R* and also points not in *R*.



Example: For instance, the disk

$$D = \{(x, y) | \ x^2 + y^2 \leq 1\}$$

which consists of all points on or inside the circle $x^2 + y^2 = 1$, is a **closed set** because it contains all of its boundary points / accumulation points (which are the points on the rim of the circle $x^2 + y^2 = 1$).

Examples

1. Find the absolute maximum and minimum values of the function $f(x, y) = x^2 - 2xy + 2y$ on the rectangle $D = \{(x, y) | 0 \leq x \leq 3, 0 \leq y \leq 2\}$.

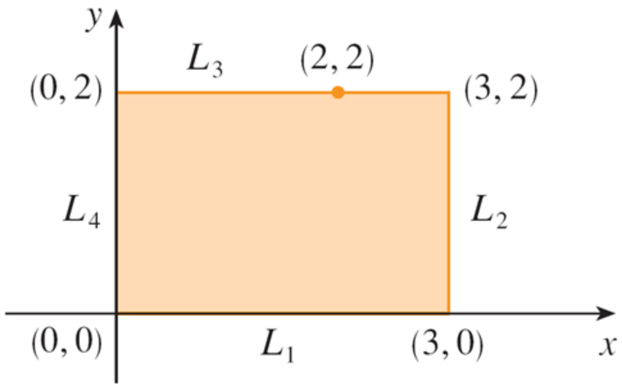
Solution:

Since f is a polynomial, it is continuous on the closed, bounded rectangle D, so there is both an absolute (global) maximum and an absolute (global) minimum within the domain.

Step 1: Consider the critical points: $f_x = 2x - 2y = 0$; $f_y = -2x + 2 = 0$

Therefore, the only critical point is (1, 1), and $f(1, 1) = 1$.

Step 2: We look at the values of f on the boundary of D, which consists of the four line segments L_1, L_2, L_3, L_4 as shown in the following figure.



On L_1 we have $y = 0$ and $f(x, 0) = x^2$,where $0 \leq x \leq 3$.
f is an increasing function of x, so its minimum value is $f(0, 0) = 0$ and its maximum value is $f(3, 0) = 9$.

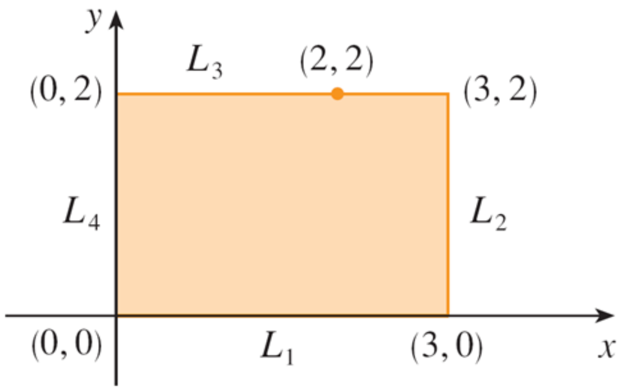
On L_2 we have $x = 3$ and $f(3, y) = 9 - 4y$,where $0 \leq y \leq 2$.

f is a decreasing function of y, so its maximum value is $f(3, 0) = 9$ and its minimum value is $f(3, 2) = 1$.

Examples

1. Find the absolute maximum and minimum values of the function $f(x, y) = x^2 - 2xy + 2y$ on the rectangle $D = \{(x, y) | 0 \le x \le 3, 0 \le y \le 2\}$.

Step 2: We look at the values of f on the boundary of D, which consists of the four line segments L_1, L_2, L_3, L_4 as shown in the following figure.

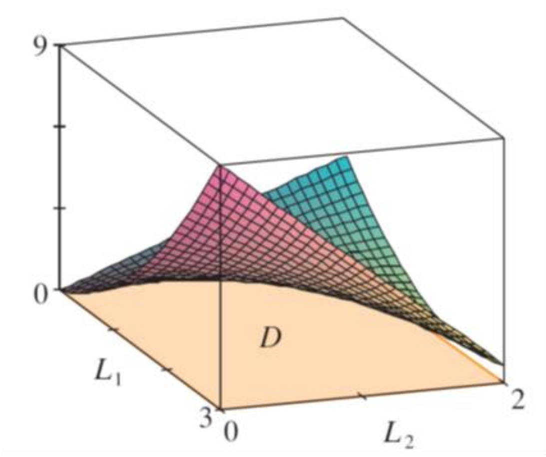


On L_3 we have $y = 2$ and $f(x, 2) = x^2 - 4x + 4$,where $0 \le x \le 3$
By noticing that $f(x, 2) = (x - 2)^2$, the minimum value of this function is $f(2, 2) = 0$ and the maximum value is $f(0, 2) = 4$.

Finally, on L_4 we have $x = 0$ and $f(0, y) = 2y$,where $0 \le y \le 2$
The maximum value is $f(0, 2) = 4$ and the minimum value is $f(0, 0) = 0$.

Figures from Larson

Therefore, on the boundary, the minimum value of f is 0 and the maximum is 9.



Step 3: We compare these values with the value $f(1, 1) = 1$ at the critical point and have the following conclusion:
The absolute (global) maximum value of f on D is $f(3, 0) = 9$.
The absolute (global) minimum value of f on D is $f(0, 0) = f(2, 2) = 0$.

Examples

2. Find the maximum and minimum values of $f(x, y) = 4xy$ on the closed disk $x^2 + y^2 \leq 4$.

<u>**Solution**</u>

Step 1: Since f is continuous and the disk is closed, f must have absolute (global) maximum and minimum values at some points of the disk.

Consider first partial derivatives are $f_x = 4y, f_y = 4x$, hence the only critical point is (0, 0), where $f(0,0) = 0$.

Step 2: Consider values of f on the boundary circle $x^2 + y^2 = 4$.

Use the parametrization of the circle:

$$\begin{cases} x = 2\cos t \\ y = 2\sin t \end{cases}, -\pi \leq t \leq \pi$$

Then, we have $f(2\cos t, 2\sin t) = 16\cos t \sin t = g(t)$

We can re-write $g(t) = 8\sin 2t$,

so maximum value of $g(t) = 8$ (at $t = \frac{\pi}{4}$ and $t = -\frac{3\pi}{4}$)

and the minimum value of $g(t) = -8$ (at $t = \frac{-\pi}{4}$ and $t = \frac{3\pi}{4}$)

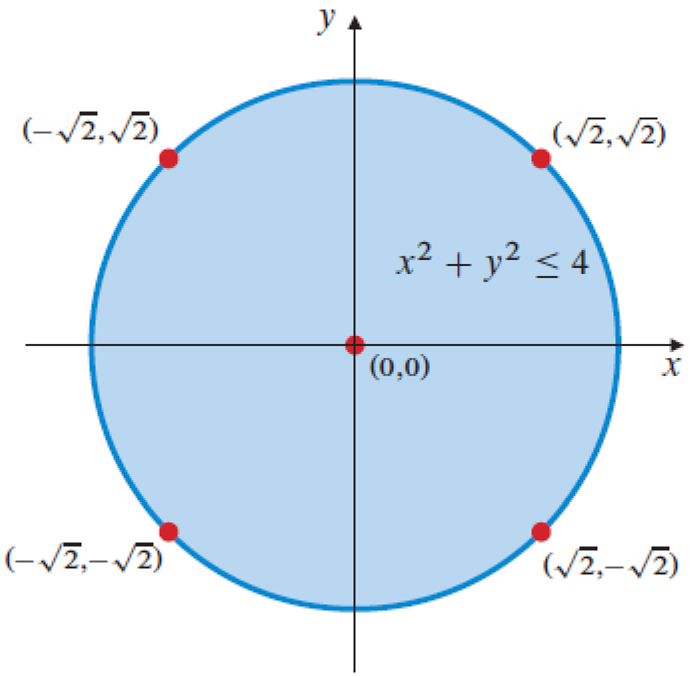


Figure from Adams and Essex

Examples

2. Find the maximum and minimum values of $f(x, y) = 4xy$ on the closed disk $x^2 + y^2 \leq 4$.

Solution

Note: We don't need to check the endpoints, since g is differentiable everywhere and is periodic with period π, any absolute (global) maximum or minimum will occur at a critical point.

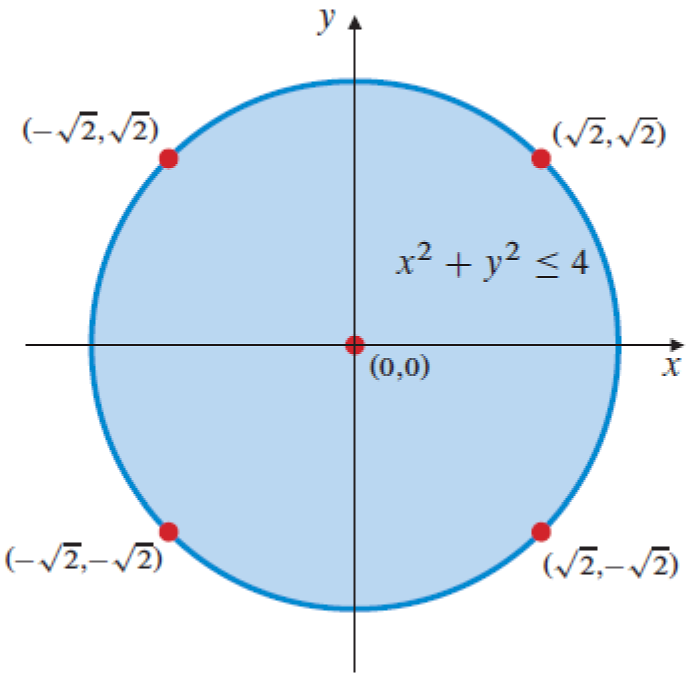


Figure from Adams and Essex

Step 3: f has maximum value 8 at the boundary points $(\sqrt{2}, \sqrt{2})$ and $(-\sqrt{2}, -\sqrt{2})$.

f has minimum value – 8 at the boundary points $(-\sqrt{2}, \sqrt{2})$ and $(\sqrt{2}, -\sqrt{2})$.

(0, 0) is actually a saddle point of f. **(How to check?)**

END OF WEEK 6

Revised by Dr. Hugo MAK