Mathematics for AI

Adams Wai Kin Kong
School of Computer Science and Engineering
Nanyang Technological University, Singapore
adamskong@ntu.edu.sg

Limits and Continuity

Let's compare the behavior of the functions

$$f(x, y) = \frac{\sin(x^2 + y^2)}{x^2 + y^2}$$

and

$$g(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$$

as x and y both approach 0 [and therefore the point (x, y) approaches the origin].

Limits and Continuity

Tables 1 and 2 show values of $f(x, y)$ and $g(x, y)$, correct to three decimal places, for points (x, y) near the origin. (Notice that neither function is defined at the origin.)

$x \backslash y$	−1.0	−0.5	−0.2	0	0.2	0.5	1.0
−1.0	0.455	0.759	0.829	0.841	0.829	0.759	0.455
−0.5	0.759	0.959	0.986	0.990	0.986	0.959	0.759
−0.2	0.829	0.986	0.999	1.000	0.999	0.986	0.829
0	0.841	0.990	1.000		1.000	0.990	0.841
0.2	0.829	0.986	0.999	1.000	0.999	0.986	0.829
0.5	0.759	0.959	0.986	0.990	0.986	0.959	0.759
1.0	0.455	0.759	0.829	0.841	0.829	0.759	0.455

Values of $f(x, y)$

Table 1

$x \backslash y$	−1.0	−0.5	−0.2	0	0.2	0.5	1.0
−1.0	0.000	0.600	0.923	1.000	0.923	0.600	0.000
−0.5	−0.600	0.000	0.724	1.000	0.724	0.000	−0.600
−0.2	−0.923	−0.724	0.000	1.000	0.000	−0.724	−0.923
0	−1.000	−1.000	−1.000		−1.000	−1.000	−1.000
0.2	−0.923	−0.724	0.000	1.000	0.000	−0.724	−0.923
0.5	−0.600	0.000	0.724	1.000	0.724	0.000	−0.600
1.0	0.000	0.600	0.923	1.000	0.923	0.600	0.000

Values of $g(x, y)$

Table 2

Limits and Continuity

It appears that as (x, y) approaches $(0, 0)$, the values of $f(x, y)$ are approaching 1 whereas the values of $g(x, y)$ aren't approaching any number. It turns out that these guesses based on numerical evidence are correct, and we write

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^2+y^2)}{x^2+y^2} = 1 \quad \text{and}$$

$$\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2} \quad \text{does not exist}$$

Limits and Continuity

In general, we use the notation

$$\lim_{(x, y) \to (a, b)} f(x, y) = L$$

to indicate that the values of *f*(x, y) approach the number *L* as the point (x, y) approaches the point (a, b) along any path that stays within the domain of *f*.

Limits and Continuity

In other words, we can make the values of $f(x, y)$ as close to L as we like by taking the point (x, y) sufficiently close to the point (a, b), but not equal to (a, b). A more precise definition follows.

1 Definition Let f be a function of two variables whose domain D includes points arbitrarily close to (a, b). Then we say that the **limit of $f(x, y)$ as (x, y) approaches (a, b)** is L and we write

$$\lim_{(x, y) \to (a, b)} f(x, y) = L$$

if for every number $\varepsilon > 0$ there is a corresponding number $\delta > 0$ such that

$$\text{if} \quad (x, y) \in D \quad \text{and} \quad 0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta \quad \text{then} \quad |f(x, y) - L| < \varepsilon$$

Limits and Continuity

Other notations for the limit in Definition 1 are

$$\lim_{\substack{x \to a \\ y \to b}} f(x, y) = L \qquad \text{and}$$

$$f(x, y) \to L \text{ as } (x, y) \to (a, b)$$

For functions of a single variable, when we let x approach a, there are only two possible directions of approach, from the left or from the right.

We know that if $\lim_{x\to a^-} f(x) \neq \lim_{x\to a^+} f(x)$, then $\lim_{x\to a} f(x)$ does not exist.

Limits and Continuity

For functions of two variables the situation is not as simple because we can let (x, y) approach (a, b) from an infinite number of directions in any manner whatsoever (see Figure 3) as long as (x, y) stays within the domain of f.

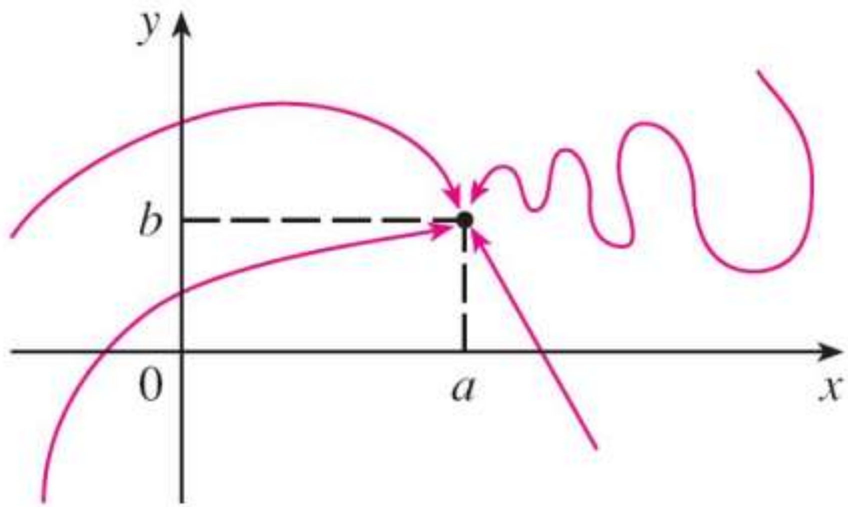


Figure 3

Limits and Continuity

Definition 1 says that the distance between $f(x, y)$ and L can be made arbitrarily small by making the distance from (x, y) to (a, b) sufficiently small (but not 0).

The definition refers only to the *distance* between (x, y) and (a, b). It does not refer to the direction of approach.

Therefore, if the limit exists, then $f(x, y)$ must approach the same limit no matter how (x, y) approaches (a, b).

Limits and Continuity

Thus, if we can find two different paths of approach along which the function $f(x, y)$ has different limits, then it follows that $\lim_{(x, y) \to (a, b)} f(x, y)$ does not exist.

If $f(x, y) \to L_1$ as $(x, y) \to (a, b)$ along a path C_1 and $f(x, y) \to L_2$ as $(x, y) \to (a, b)$ along a path C_2, where $L_1 \neq L_2$, then $\lim_{(x, y) \to (a, b)} f(x, y)$ does not exist.

Example 1

Show that $\lim_{(x,y)\to(0,0)} \frac{x^2 - y^2}{x^2 + y^2}$ does not exist.

Solution:

Let $f(x, y) = (x^2 - y^2)/(x^2 + y^2)$.

First let's approach (0, 0) along the x-axis.

Then $y = 0$ gives $f(x, 0) = x^2/x^2 = 1$ for all $x \neq 0$, so

$$f(x, y) \to 1 \quad \text{as} \quad (x, y) \to (0, 0) \text{ along the } x\text{-axis}$$

Example 1 – *Solution*

We now approach along the y-axis by putting $x = 0$.

Then $f(0, y) = \dfrac{-y^2}{y^2} = -1$ for all $y \neq 0$, so

$$f(x, y) \to -1 \quad \text{as} \quad (x, y) \to (0, 0) \text{ along the } y\text{-axis}$$

Thus limit doesn't exist.

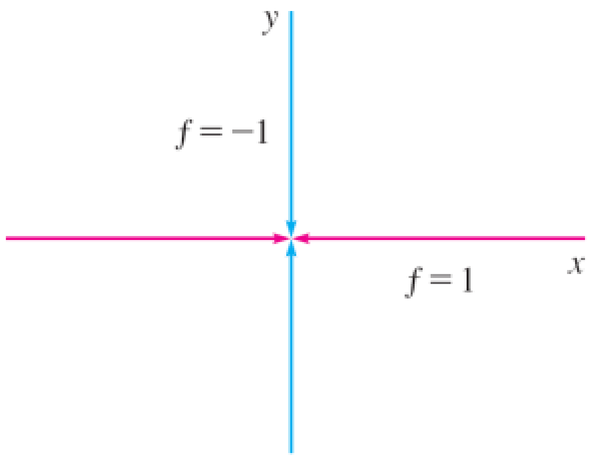


Figure 4

Limits and Continuity

Now let's look at limits that *do* exist. Just as for functions of one variable, the calculation of limits for functions of two variables can be greatly simplified by the use of properties of limits.

The Limit Laws can be extended to functions of two variables: the limit of a sum is the sum of the limits, the limit of a product is the product of the limits, and so on.

In particular, the following equations are true.

$$\boxed{2} \quad \lim_{(x, y) \to (a, b)} x = a \qquad \lim_{(x, y) \to (a, b)} y = b \qquad \lim_{(x, y) \to (a, b)} c = c$$

The Squeeze Theorem also holds.

Continuity

We know that evaluating limits of *continuous* functions of a single variable is easy.

It can be accomplished by direct substitution because the defining property of a continuous function is $\lim_{x \to a} f(x) = f(a)$.

Continuous functions of two variables are also defined by the direct substitution property.

4 Definition A function f of two variables is called **continuous at** (a, b) if

$$\lim_{(x, y) \to (a, b)} f(x, y) = f(a, b)$$

We say f is **continuous on** D if f is continuous at every point (a, b) in D.

Continuity

The intuitive meaning of continuity is that if the point (x, y) changes by a small amount, then the value of $f(x, y)$ changes by a small amount.

This means that a surface that is the graph of a continuous function has no hole or break.

Using the properties of limits, you can see that sums, differences, products, and quotients of continuous functions are continuous on their domains.

Let's use this fact to give examples of continuous functions.

Continuity

A **polynomial function of two variables** (or polynomial, for short) is a sum of terms of the form $cx^m y^n$, where c is a constant and m and n are nonnegative integers.

A **rational function** is a ratio of polynomials.

For instance,

$$f(x, y) = x^4 + 5x^3y^2 + 6xy^4 - 7y + 6$$

is a polynomial, whereas

$$g(x, y) = \frac{2xy + 1}{x^2 + y^2}$$

is a rational function.

Continuity

The limits in (2) show that the functions
$f(x, y) = x$, $g(x, y) = y$, and $h(x, y) = c$ are continuous.

Since any polynomial can be built up out of the simple functions f, g, and h by multiplication and addition, it follows that *all polynomials are continuous on* $\mathbb{R}^2$.

Likewise, any rational function is continuous on its domain because it is a quotient of continuous functions.

Example 5

Evaluate $\lim_{(x,y)\to(1,2)} (x^2y^3 - x^3y^2 + 3x + 2y)$.

Solution:

Since $f(x, y) = x^2y^3 - x^3y^2 + 3x + 2y$ is a polynomial, it is continuous everywhere, so we can find the limit by direct substitution:

$$\lim_{(x,y)\to(1,2)} (x^2y^3 - x^3y^2 + 3x + 2y) = 1^2 \cdot 2^3 - 1^3 \cdot 2^2 + 3 \cdot 1 + 2 \cdot 2$$

$$= 11$$

Continuity

Just as for functions of one variable, composition is another way of combining two continuous functions to get a third.

In fact, it can be shown that if f is a continuous function of two variables and g is a continuous function of a single variable that is defined on the range of f, then the composite function $h = g \circ f$ defined by $h(x, y) = g(f(x, y))$ is also a continuous function.

How about deep network, continue or not?

Functions of Three or More Variables

Everything that we have done in this section can be extended to functions of three or more variables.

The notation

$$\lim_{(x, y, z) \to (a, b, c)} f(x, y, z) = L$$

means that the values of *f*(*x*, *y*, *z*) approach the number *L* as the point (*x*, *y*, *z*) approaches the point (*a*, *b*, *c*) along any path in the domain of *f*.

Functions of Three or More Variables

Because the distance between two points (x, y, z) and (a, b, c) in $\mathbb{R}^3$ is given by $\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$, we can write the precise definition as follows: for every number $\varepsilon > 0$ there is a corresponding number $\delta > 0$ such that

if (x, y, z) is in the domain of f and

$$0 < \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} < \delta$$

then $|f(x, y, z) - L| < \varepsilon$

Functions of Three or More Variables

The function f is **continuous** at (a, b, c) if

$$\lim_{(x, y, z) \to (a, b, c)} f(x, y, z) = f(a, b, c)$$

For instance, the function

$$f(x, y, z) = \frac{1}{x^2 + y^2 + z^2 - 1}$$

is a rational function of three variables and so is continuous at every point in $\mathbb{R}^3$ except where $x^2 + y^2 + z^2 = 1$. In other words, it is discontinuous on the sphere with center the origin and radius 1.

Functions of Three or More Variables

We can write the definitions of a limit for functions of two or three variables in a single compact form as follows.

5 If f is defined on a subset D of $\mathbb{R}^n$, then $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = L$ means that for every number $\varepsilon > 0$ there is a corresponding number $\delta > 0$ such that

$$\text{if} \quad \mathbf{x} \in D \quad \text{and} \quad 0 < |\mathbf{x} - \mathbf{a}| < \delta \quad \text{then} \quad |f(\mathbf{x}) - L| < \varepsilon$$

Partial Derivatives

Partial Derivatives

On a hot day, extreme humidity makes us think the temperature is higher than it really is, whereas in very dry air we perceive the temperature to be lower than the thermometer indicates.

The National Weather Service has devised the *heat index* (also called the temperature-humidity index, or humidex, in some countries) to describe the combined effects of temperature and humidity.

The heat index I is the perceived air temperature when the actual temperature is T and the relative humidity is H.
So I is a function of T and H and we can write $I = f(T, H)$.

Partial Derivatives

The following table of values of *I* is an excerpt from a table compiled by the National Weather Service.

Relative humidity (%)

Actual temperature (°F) T \ H	50	55	60	65	70	75	80	85	90
90	96	98	100	103	106	109	112	115	119
92	100	103	105	108	112	115	119	123	128
94	104	107	111	114	118	122	127	132	137
96	109	113	116	121	125	130	135	141	146
98	114	118	123	127	133	138	144	150	157
100	119	124	129	135	141	147	154	161	168

Heat index *I* as a function of temperature and humidity

Partial Derivatives

If we concentrate on the highlighted column of the table, which corresponds to a relative humidity of $H = 70\%$, we are considering the heat index as a function of the single variable T for a fixed value of H. Let's write $g(T) = f(T, 70)$.

Then $g(T)$ describes how the heat index I increases as the actual temperature T increases when the relative humidity is 70%.

The derivative of g when $T = 96°F$ is the rate of change of I with respect to T when $T = 96°F$:

$$g'(96) = \lim_{h \to 0} \frac{g(96 + h) - g(96)}{h} = \lim_{h \to 0} \frac{f(96 + h, 70) - f(96, 70)}{h}$$

Partial Derivatives

We can approximate $g'(96)$ using the values in Table 1 by taking h = 2 and –2:

$$g'(96) \approx \frac{g(98) - g(96)}{2} = \frac{f(98, 70) - f(96, 70)}{2} = \frac{133 - 125}{2} = 4$$

$$g'(96) \approx \frac{g(94) - g(96)}{-2} = \frac{f(94, 70) - f(96, 70)}{-2} = \frac{118 - 125}{-2} = 3.5$$

Averaging these values, we can say that the derivative $g'(96)$ is approximately 3.75.

Partial Derivatives

This means that, when the actual temperature is 96°F and the relative humidity is 70%, the apparent temperature (heat index) rises by about 3.75°F for every degree that the actual temperature rises!

Partial Derivatives

Now let's look at the highlighted row in Table 1, which corresponds to a fixed temperature of T = 96°F.

Relative humidity (%)

Actual temperature (°F) T \ H	50	55	60	65	70	75	80	85	90
90	96	98	100	103	106	109	112	115	119
92	100	103	105	108	112	115	119	123	128
94	104	107	111	114	118	122	127	132	137
96	109	113	116	121	125	130	135	141	146
98	114	118	123	127	133	138	144	150	157
100	119	124	129	135	141	147	154	161	168

Heat index I as a function of temperature and humidity

Table 1

Partial Derivatives

The numbers in this row are values of the function $G(H) = f(96, H)$, which describes how the heat index increases as the relative humidity H increases when the actual temperature is $T = 96°F$.

The derivative of this function when $H = 70\%$ is the rate of change of I with respect to H when $H = 70\%$:

$$G'(70) = \lim_{h \to 0} \frac{G(70 + h) - G(70)}{h} = \lim_{h \to 0} \frac{f(96, 70 + h) - f(96, 70)}{h}$$

Partial Derivatives

By taking h = 5 and –5, we approximate $G'(70)$ using the tabular values:

$$G'(70) \approx \frac{G(75) - G(70)}{5} = \frac{f(96, 75) - f(96, 70)}{5} = \frac{130 - 125}{5} = 1$$

$$G'(70) \approx \frac{G(65) - G(70)}{-5} = \frac{f(96, 65) - f(96, 70)}{-5} = \frac{121 - 125}{-5} = 0.8$$

By averaging these values we get the estimate $G'(70) \approx 0.9$. This says that, when the temperature is 96°F and the relative humidity is 70%, the heat index rises about 0.9°F for every percent that the relative humidity rises.

Partial Derivatives

In general, if f is a function of two variables x and y, suppose we let only x vary while keeping y fixed, say $y = b$, where b is a constant.

Then we are really considering a function of a single variable x, namely, $g(x) = f(x, b)$. If g has a derivative at a, then we call it the **partial derivative of *f* with respect to *x* at (*a*, *b*)** and denote it by $f_x(a, b)$. Thus

1 $$\boxed{f_x(a, b) = g'(a) \qquad \text{where} \qquad g(x) = f(x, b)}$$

NANYANG TECHNOLOGICAL UNIVERSITY

Partial Derivatives

By the definition of a derivative, we have

$$g'(a) = \lim_{h \to 0} \frac{g(a + h) - g(a)}{h}$$

and so Equation 1 becomes

2

$$\boxed{f_x(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}}$$

Partial Derivatives

Similarly, the **partial derivative of *f* with respect to *y* at (*a*, *b*)**, denoted by $f_y(a, b)$, is obtained by keeping x fixed ($x = a$) and finding the ordinary derivative at b of the function $G(y) = f(a, y)$:

3

$$\boxed{f_y(a, b) = \lim_{h \to 0} \frac{f(a, b + h) - f(a, b)}{h}}$$

With this notation for partial derivatives, we can write the rates of change of the heat index I with respect to the actual temperature T and relative humidity H when $T = 96°F$ and $H = 70\%$ as follows:

$$f_T(96, 70) \approx 3.75 \qquad f_H(96, 70) \approx 0.9$$

Partial Derivatives

If we now let the point (a, b) vary in Equations 2 and 3, f_x and f_y become functions of two variables.

$\boxed{4}$ If f is a function of two variables, its **partial derivatives** are the functions f_x and f_y defined by

$$f_x(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}$$

$$f_y(x, y) = \lim_{h \to 0} \frac{f(x, y + h) - f(x, y)}{h}$$

Partial Derivatives

There are many alternative notations for partial derivatives.

For instance, instead of f_x we can write f_1 or $D_1 f$ (to indicate differentiation with respect to the *first* variable) or $\partial f/\partial x$.

But here $\partial f/\partial x$ can't be interpreted as a ratio of differentials.

Notations for Partial Derivatives If $z = f(x, y)$, we write

$$f_x(x, y) = f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} f(x, y) = \frac{\partial z}{\partial x} = f_1 = D_1 f = D_x f$$

$$f_y(x, y) = f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y} f(x, y) = \frac{\partial z}{\partial y} = f_2 = D_2 f = D_y f$$

Partial Derivatives

To compute partial derivatives, all we have to do is remember from Equation 1 that the partial derivative with respect to x is just the *ordinary* derivative of the function g of a single variable that we get by keeping y fixed.

Thus we have the following rule.

Rule for Finding Partial Derivatives of $z = f(x, y)$

1. To find f_x, regard y as a constant and differentiate $f(x, y)$ with respect to x.
2. To find f_y, regard x as a constant and differentiate $f(x, y)$ with respect to y.

Example 1

If $f(x, y) = x^3 + x^2y^3 - 2y^2$, find $f_x(2, 1)$ and $f_y(2, 1)$.

Solution:

Holding y constant and differentiating with respect to x, we get

$$f_x(x, y) = 3x^2 + 2xy^3$$

and so $$f_x(2, 1) = 3 \cdot 2^2 + 2 \cdot 2 \cdot 1^3 = 16$$

Holding x constant and differentiating with respect to y, we get

$$f_y(x, y) = 3x^2y^2 - 4y$$

$$f_y(2, 1) = 3 \cdot 2^2 \cdot 1^2 - 4 \cdot 1 = 8$$

Interpretations of Partial Derivatives

To give a geometric interpretation of partial derivatives, we recall that the equation $z = f(x, y)$ represents a surface S (the graph of f). If $f(a, b) = c$, then the point $P(a, b, c)$ lies on S.

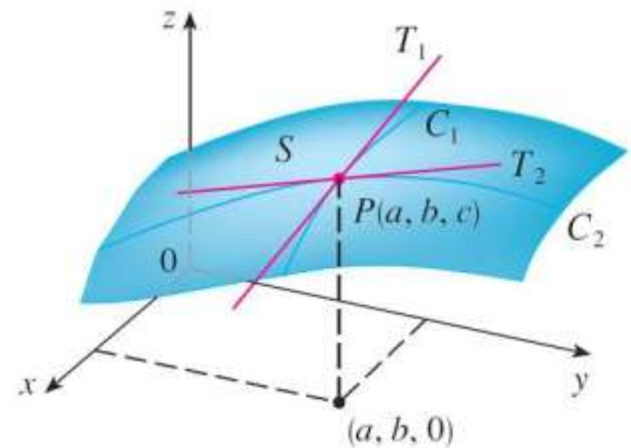
By fixing $y = b$, we are restricting our attention to the curve C_1 in which the vertical plane $y = b$ intersects S. (In other words, C_1 is the trace of S in the plane $y = b$.)

Interpretations of Partial Derivatives

Likewise, the vertical plane $x = a$ intersects S in a curve C_2. Both of the curves C_1 and C_2 pass through the point P. (See Figure 1.)

Note that the curve C_1 is the graph of the function $g(x) = f(x, b)$, so the slope of its tangent T_1 at P is $g'(a) = f_x(a, b)$.

The curve C_2 is the graph of the function $G(y) = f(a, y)$, so the slope of its tangent T_2 at P is $G'(b) = f_y(a, b)$.



The partial derivatives of f at (a, b) are the slopes of the tangents to C_1 and C_2.

Figure 1

Interpretations of Partial Derivatives

Thus the partial derivatives $f_x(a, b)$ and $f_y(a, b)$ can be interpreted geometrically as the slopes of the tangent lines at $P(a, b, c)$ to the traces C_1 and C_2 of S in the planes $y = b$ and $x = a$.

As we have seen in the case of the heat index function, partial derivatives can also be interpreted as *rates of change*.

If $z = f(x, y)$, then $\partial z/\partial x$ represents the rate of change of z with respect to x when y is fixed. Similarly, $\partial z/\partial y$ represents the rate of change of z with respect to y when x is fixed.

Example 2

If $f(x, y) = 4 - x^2 - 2y^2$, find $f_x(1, 1)$ and $f_y(1, 1)$ and interpret these numbers as slopes.

Solution:

We have

$$f_x(x, y) = -2x \qquad f_y(x, y) = -4y$$

$$f_x(1, 1) = -2 \qquad f_y(1, 1) = -4$$

Example 2 – *Solution*

The graph of f is the paraboloid $z = 4 - x^2 - 2y^2$ and the vertical plane $y = 1$ intersects it in the parabola $z = 2 - x^2$, $y = 1$. (As in the preceding discussion, we label it C_1 in Figure 2.)

The slope of the tangent line to this parabola at the point $(1, 1, 1)$ is $f_x(1, 1) = -2$.

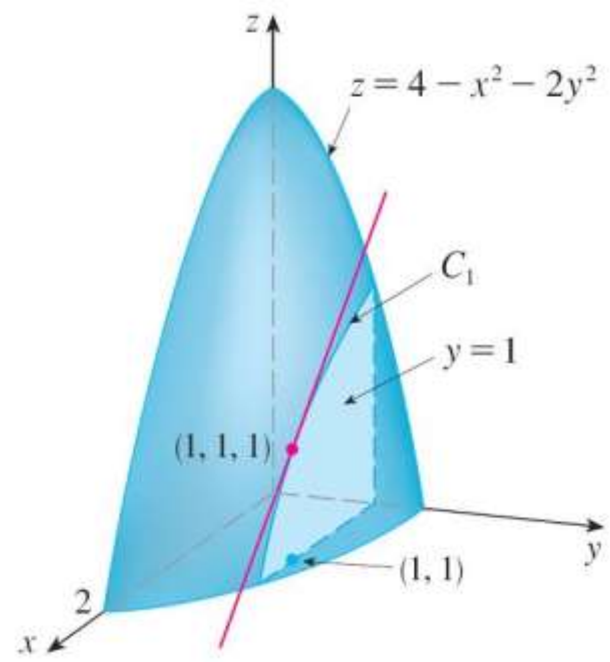


Figure 2

NANYANG TECHNOLOGICAL UNIVERSITY

Example 2 – *Solution*

Similarly, the curve C_2 in which the plane $x = 1$ intersects the paraboloid is the parabola $z = 3 - 2y^2$, $x = 1$, and the slope of the tangent line at (1, 1, 1) is $f_y(1, 1) = -4$. (See Figure 3.)

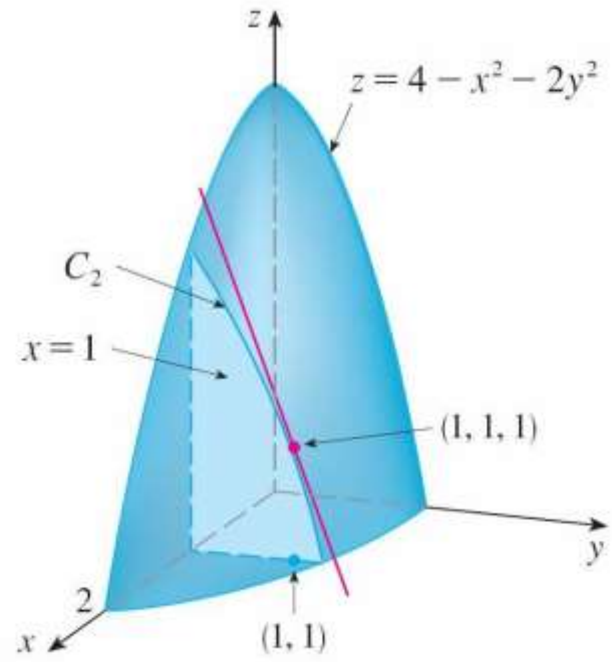


Figure 3

Functions of More Than Two Variables

Partial derivatives can also be defined for functions of three or more variables. For example, if *f* is a function of three variables *x*, *y*, and *z*, then its partial derivative with respect to *x* is defined as

$$f_x(x, y, z) = \lim_{h \to 0} \frac{f(x + h, y, z) - f(x, y, z)}{h}$$

and it is found by regarding *y* and *z* as constants and differentiating *f*(*x*, *y*, *z*) with respect to *x*.

Functions of More Than Two Variables

If $w = f(x, y, z)$, then $f_x = \partial w/\partial x$ can be interpreted as the rate of change of w with respect to x when y and z are held fixed. But we can't interpret it geometrically because the graph of f lies in four-dimensional space.

In general, if u is a function of n variables, $u = f(x_1, x_2, \ldots, x_n)$, its partial derivative with respect to the ith variable x_i is

$$\frac{\partial u}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{h}$$

and we also write $$\frac{\partial u}{\partial x_i} = \frac{\partial f}{\partial x_i} = f_{x_i} = f_i = D_i f$$

Example 6

Find f_x, f_y, and f_z if $f(x, y, z) = e^{xy} \ln z$.

Solution:

Holding y and z constant and differentiating with respect to x, we have

$$f_x = ye^{xy} \ln z$$

Similarly,

$$f_y = xe^{xy} \ln z \qquad \text{and} \qquad f_z = \frac{e^{xy}}{z}$$

Higher Derivatives

If f is a function of two variables, then its partial derivatives f_x and f_y are also functions of two variables, so we can consider their partial derivatives $(f_x)_x$, $(f_x)_y$, $(f_y)_x$, and $(f_y)_y$, which are called the **second partial derivatives** of f.

If $z = f(x, y)$, we use the following notation:

$$(f_x)_x = f_{xx} = f_{11} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 z}{\partial x^2}$$

$$(f_x)_y = f_{xy} = f_{12} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y\, \partial x} = \frac{\partial^2 z}{\partial y\, \partial x}$$

Higher Derivatives

$$(f_y)_x = f_{yx} = f_{21} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x\, \partial y} = \frac{\partial^2 z}{\partial x\, \partial y}$$

$$(f_y)_y = f_{yy} = f_{22} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 z}{\partial y^2}$$

Thus the notation f_{xy} (or $\partial^2 f/\partial y\, \partial x$) means that we first differentiate with respect to x and then with respect to y, whereas in computing f_{yx} the order is reversed.

Example 7

Find the second partial derivatives of

$$f(x, y) = x^3 + x^2y^3 - 2y^2$$

Solution:

In Example 1 we found that

$$f_x(x, y) = 3x^2 + 2xy^3 \qquad f_y(x, y) = 3x^2y^2 - 4y$$

Therefore

$$f_{xx} = \frac{\partial}{\partial x}(3x^2 + 2xy^3)$$

$$= 6x + 2y^3$$

Example 7 – *Solution*

$$f_{xy} = \frac{\partial}{\partial y}(3x^2 + 2xy^3)$$

$$= 6xy^2$$

$$f_{yx} = \frac{\partial}{\partial x}(3x^2y^2 - 4y)$$

$$= 6xy^2$$

$$f_{yy} = \frac{\partial}{\partial y}(3x^2y^2 - 4y)$$

$$= 6x^2y - 4$$

Higher Derivatives

Notice that $f_{xy} = f_{yx}$ in Example 7. This is not just a coincidence. It turns out that the mixed partial derivatives f_{xy} and f_{yx} are equal for most functions that one meets in practice.

The following theorem, which was discovered by the French mathematician Alexis Clairaut (1713–1765), gives conditions under which we can assert that $f_{xy} = f_{yx}$.

Clairaut's Theorem Suppose f is defined on a disk D that contains the point (a, b). If the functions f_{xy} and f_{yx} are both continuous on D, then

$$f_{xy}(a, b) = f_{yx}(a, b)$$

https://en.wikipedia.org/wiki/Disk_(mathematics)

Higher Derivatives

Partial derivatives of order 3 or higher can also be defined. For instance,

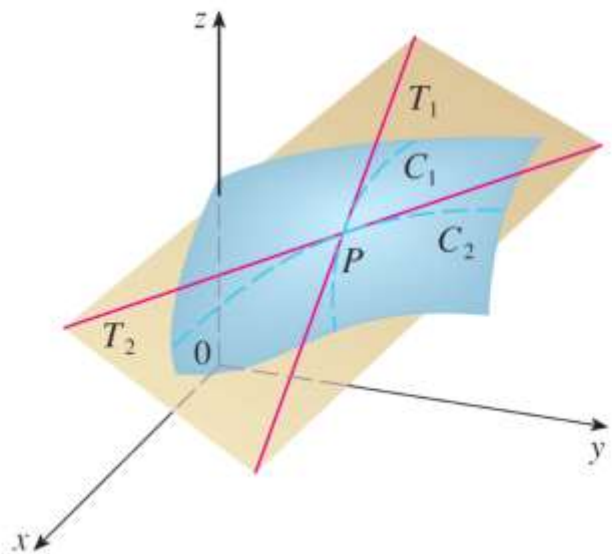
$$f_{xyy} = (f_{xy})_y = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial y\, \partial x}\right) = \frac{\partial^3 f}{\partial y^2\, \partial x}$$

and using Clairaut's Theorem it can be shown that $f_{xyy} = f_{yxy} = f_{yyx}$ if these functions are continuous.

Tangent Planes and Approximations

Tangent Planes

The **tangent plane** to the surface S at the point P is defined to be the plane that contains both tangent lines T_1 and T_2. (See Figure 1.)



The tangent plane contains the tangent lines T_1 and T_2.

Figure 1

Tangent Planes

If C is any other curve that lies on the surface S and passes through P, then its tangent line at P also lies in the tangent plane.

Therefore you can think of the tangent plane to S at P as consisting of all possible tangent lines at P to curves that lie on S and pass through P. The tangent plane at P is the plane that most closely approximates the surface S near the point P. We know that any plane passing through the point $P(x_0, y_0, z_0)$ has an equation of the form

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$$

Tangent Planes

By dividing this equation by C and letting $a = -A/C$ and $b = -B/C$, we can write it in the form

$$\boxed{1} \qquad z - z_0 = a(x - x_0) + b(y - y_0)$$

If Equation 1 represents the tangent plane at P, then its intersection with the plane $y = y_0$ must be the tangent line T_1. Setting $y = y_0$ in Equation 1 gives

$$z - z_0 = a(x - x_0) \qquad \text{where } y = y_0$$

and we recognize this as the equation (in point-slope form) of a line with slope a.

Tangent Planes

But we know that the slope of the tangent T_1 is $f_x(x_0, y_0)$.

Therefore $a = f_x(x_0, y_0)$.

Similarly, putting $x = x_0$ in Equation 1, we get $z - z_0 = b(y - y_0)$, which must represent the tangent line T_2, so $b = f_y(x_0, y_0)$.

$\boxed{2}$ Suppose f has continuous partial derivatives. An equation of the tangent plane to the surface $z = f(x, y)$ at the point $P(x_0, y_0, z_0)$ is

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

Example 1

Find the tangent plane to the elliptic paraboloid $z = 2x^2 + y^2$ at the point (1, 1, 3).

Solution:

Let $f(x, y) = 2x^2 + y^2$.

Then

$$f_x(x, y) = 4x \qquad f_y(x, y) = 2y$$
$$f_x(1, 1) = 4 \qquad f_y(1, 1) = 2$$

Then (2) gives the equation of the tangent plane at (1, 1, 3) as

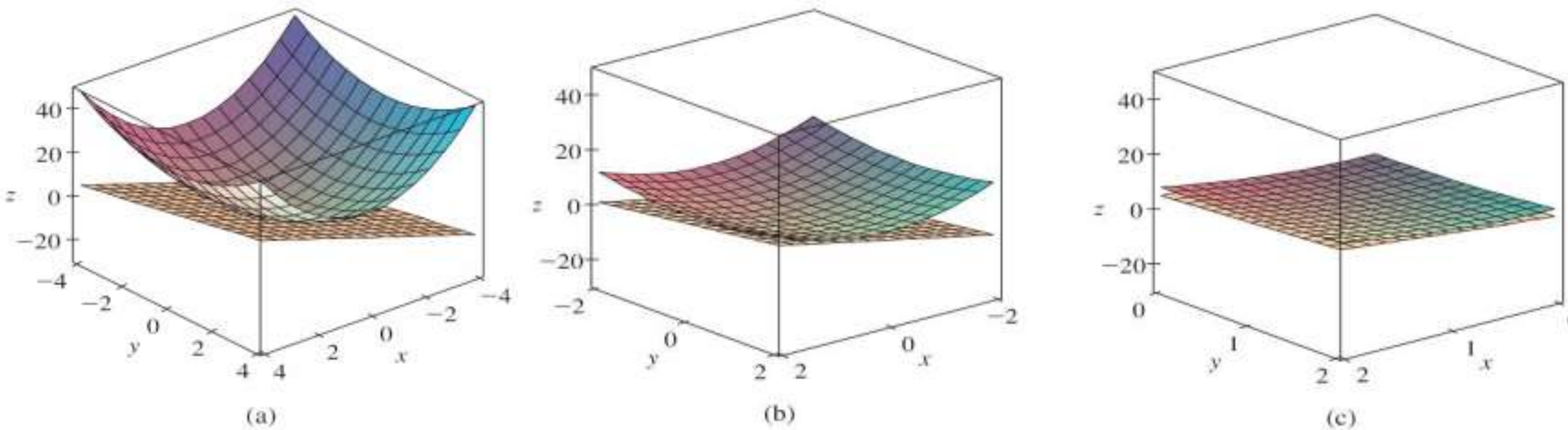
$$z - 3 = 4(x - 1) + 2(y - 1)$$

or

$$z = 4x + 2y - 3$$

Tangent Planes

Figure 2(a) shows the elliptic paraboloid and its tangent plane at (1, 1, 3) that we found in Example 1. In parts (b) and (c) we zoom in toward the point (1, 1, 3) by restricting the domain of the function $f(x, y) = 2x^2 + y^2$.



The elliptic paraboloid $z = 2x^2 + y^2$ appears to coincide with its tangent plane as we zoom in toward (1, 1, 3).

Figure 2

Linear Approximations

In Example 1 we found that an equation of the tangent plane to the graph of the function $f(x, y) = 2x^2 + y^2$ at the point (1, 1, 3) is $z = 4x + 2y - 3$. Therefore, the linear function of two variables

$$L(x, y) = 4x + 2y - 3$$

is a good approximation to $f(x, y)$ when (x, y) is near (1, 1). The function L is called the *linearization* of f at (1, 1) and the approximation

$$f(x, y) \approx 4x + 2y - 3$$

is called the *linear approximation* or *tangent plane approximation* of f at (1, 1).

Linear Approximations

For instance, at the point (1.1, 0.95) the linear approximation gives

$$f(1.1, 0.95) \approx 4(1.1) + 2(0.95) - 3 = 3.3$$

which is quite close to the true value of

$$f(1.1, 0.95) = 2(1.1)^2 + (0.95)^2 = 3.3225.$$

But if we take a point farther away from (1, 1), such as (2, 3), we no longer get a good approximation.

In fact, $L(2, 3) = 11$ whereas $f(2, 3) = 17$.

Linear Approximations

In general, we know from (2) that an equation of the tangent plane to the graph of a function f of two variables at the point $(a, b, f(a, b))$ is

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

The linear function whose graph is this tangent plane, namely

$$L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

is called the **linearization** of f at (a, b).

Linear Approximations

The approximation

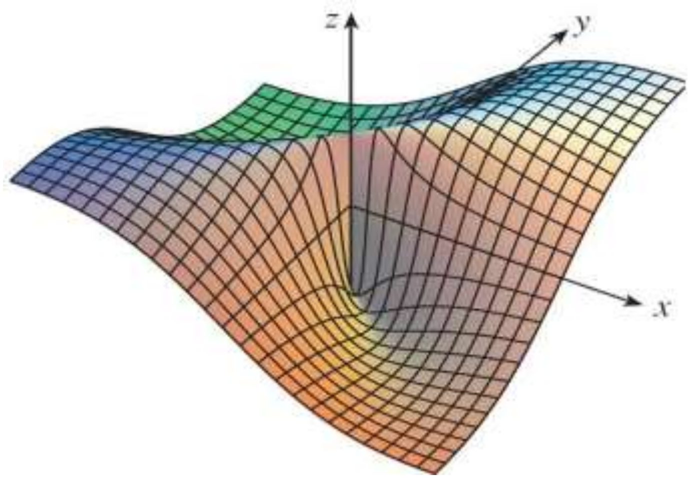
$$\boxed{4} \qquad f(x, y) \approx f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

is called the **linear approximation** or the **tangent plane approximation** of f at (a, b).

Linear Approximations

We have defined tangent planes for surfaces $z = f(x, y)$, where f has continuous first partial derivatives. What happens if f_x and f_y are not continuous? Figure 4 pictures such a function; its equation is

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



$f(x, y) = \dfrac{xy}{x^2 + y^2}$ if $(x, y) \neq (0, 0)$,
$f(0, 0) = 0$

Figure 4

You can verify that its partial derivatives exist at the origin and, in fact, $f_x(0, 0) = 0$ and $f_y(0, 0) = 0$, but f_x and f_y are not continuous.

Linear Approximations

Considering (0, 0) and by definition

$$f_x(0,0) = \lim_{h\to 0} \frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{1}{h}\left(\frac{(0+h)\times 0}{h^2+0^2} - 0\right) = 0$$

Similarly $f_y(0,0) = \lim_{h\to 0} \frac{f(0,0+h)-f(0,0)}{h} = \lim_{h\to 0} \frac{1}{h}\left(\frac{0\times(0+h)}{0^2+h^2} - 0\right) = 0$. Thus, $f_x(0,0)$ and $f_y(0,0)$ exist and are equal to zero.

Considering $(x, y) \neq (0, 0)$

$$f_x(x,y) = \frac{y^3 - x^2 y}{(x^2 + y^2)^2} \text{ and } f_y(x,y) = \frac{x^3 - xy^2}{(x^2+y^2)^2}$$

$f_x(0,h) = \frac{h^3 - 0^2\times h}{(0^2+h^2)^2} = \frac{h^3}{h^4} = \frac{1}{h}$. When $h \to 0$, $f_x(0,h)$ does not exist. Thus, f_x is not continuous. Similarly for $f_y(0,h)$.

Linear Approximations

The linear approximation would be $f(x, y) \approx 0$, but $f(x, y) = \frac{1}{2}$ at all points on the line $y = x$.

So a function of two variables can behave badly even though both of its partial derivatives exist.

To rule out such behavior, we formulate the idea of a differentiable function of two variables.

Recall that for a function of one variable, $y = f(x)$, if x changes from a to $a + \Delta x$, we defined the increment of y as

$$\Delta y = f(a + \Delta x) - f(a)$$

Linear Approximations

If f is differentiable at a, then

$$\boxed{5} \qquad \Delta y = f'(a)\,\Delta x + \varepsilon\,\Delta x \quad \text{where } \varepsilon \to 0 \text{ as } \Delta x \to 0$$

Now consider a function of two variables, $z = f(x, y)$, and suppose x changes from a to $a + \Delta x$ and y changes from b to $b + \Delta y$. Then the corresponding **increment** of z is

$$\boxed{6} \qquad \Delta z = f(a + \Delta x, b + \Delta y) - f(a, b)$$

Thus the increment Δz represents the change in the value of f when (x, y) changes from (a, b) to $(a + \Delta x, b + \Delta y)$.

Linear Approximations

By analogy with (5) we define the differentiability of a function of two variables as follows.

7 Definition If $z = f(x, y)$, then f is **differentiable** at (a, b) if Δz can be expressed in the form

$$\Delta z = f_x(a, b)\,\Delta x + f_y(a, b)\,\Delta y + \varepsilon_1\,\Delta x + \varepsilon_2\,\Delta y$$

where ε_1 and $\varepsilon_2 \to 0$ as $(\Delta x,\ \Delta y) \to (0, 0)$.

Definition 7 says that a differentiable function is one for which the linear approximation (4) is a good approximation when (x, y) is near (a, b). In other words, the tangent plane approximates the graph of f well near the point of tangency

Linear Approximations

It's sometimes hard to use Definition 7 directly to check the differentiability of a function, but the next theorem provides a convenient sufficient condition for differentiability.

8 Theorem If the partial derivatives f_x and f_y exist near (a, b) and are continuous at (a, b), then f is differentiable at (a, b).

Example 2

Show that $f(x, y) = xe^{xy}$ is differentiable at (1, 0) and find its linearization there. Then use it to approximate $f(1.1, -0.1)$.

Solution:

The partial derivatives are

$$f_x(x, y) = e^{xy} + xye^{xy} \qquad f_y(x, y) = x^2e^{xy}$$
$$f_x(1, 0) = 1 \qquad f_y(1, 0) = 1$$

Both f_x *and* f_y are continuous functions, so f is differentiable by Theorem 8. The linearization is

$$L(x, y) = f(1, 0) + f_x(1, 0)(x - 1) + f_y(1, 0)(y - 0)$$
$$= 1 + 1(x - 1) + 1 \cdot y = x + y$$

NANYANG TECHNOLOGICAL UNIVERSITY

Example 2 – *Solution*

The corresponding linear approximation is

$$xe^{xy} \approx x + y$$

so $$f(1.1, -0.1) \approx 1.1 - 0.1 = 1$$

Compare this with the actual value of

$$f(1.1, -0.1) = 1.1e^{-0.11}$$

$$\approx 0.98542.$$

Differentials

For a differentiable function of one variable, $y = f(x)$, we define the differential dx to be an independent variable; that is, dx can be given the value of any real number.

The differential of y is then defined as

$$\boxed{9} \qquad dy = f'(x)\,dx$$

Differentials

Figure 6 shows the relationship between the increment Δy and the differential dy: Δy represents the change in height of the curve $y = f(x)$ and dy represents the change in height of the tangent line when x changes by an amount $dx = \Delta x$.

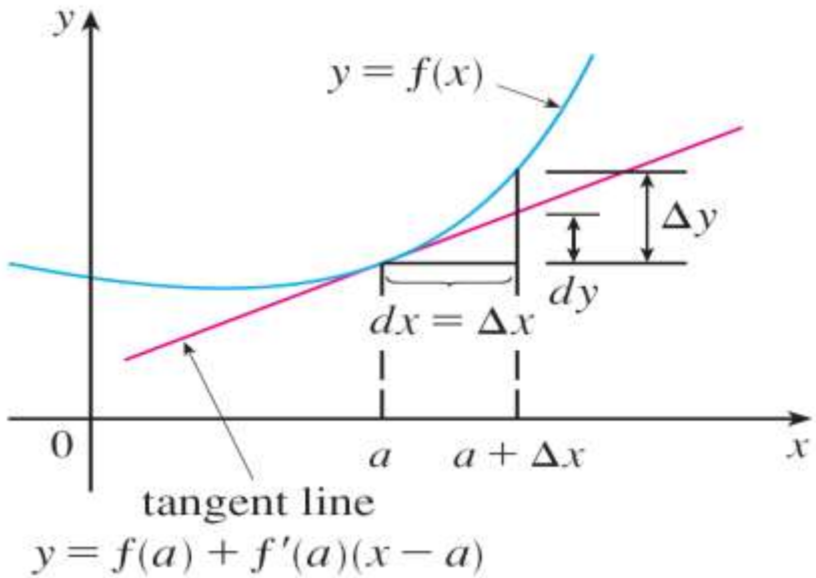


Figure 6

Differentials

For a differentiable function of two variables, $z = f(x, y)$, we define the **differentials** dx and dy to be independent variables; that is, they can be given any values. Then the **differential** dz, also called the **total differential**, is defined by

$$dz = f_x(x, y)\,dx + f_y(x, y)\,dy = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy \tag{10}$$

Sometimes the notation df is used in place of dz.

Differentials

If we take $dx = \Delta x = x - a$ and $dy = \Delta y = y - b$ in Equation 10, then the differential of z is

$$dz = f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

So, in the notation of differentials, the linear approximation (4) can be written as

$$f(x, y) \approx f(a, b) + dz$$

Differentials

Figure 7 is the three-dimensional counterpart of Figure 6 and shows the geometric interpretation of the differential *dz* and the increment Δz: *dz* represents the change in height of the tangent plane, whereas Δz represents the change in height of the surface $z = f(x, y)$ when (x, y) changes from (a, b) to $(a + \Delta x, b + \Delta y)$.

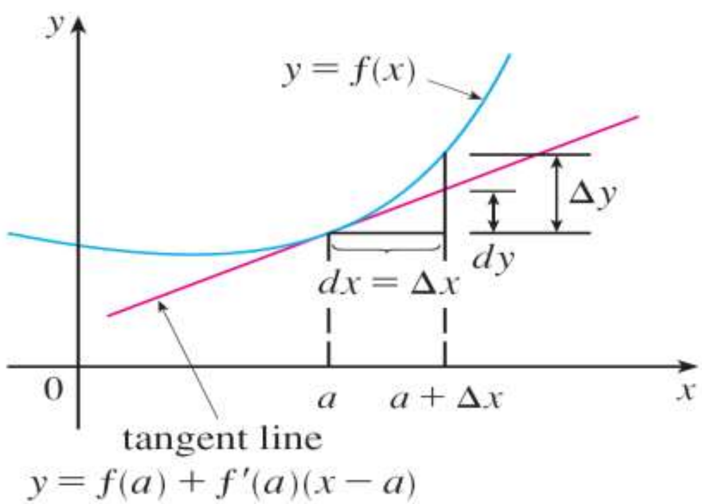


Figure 6

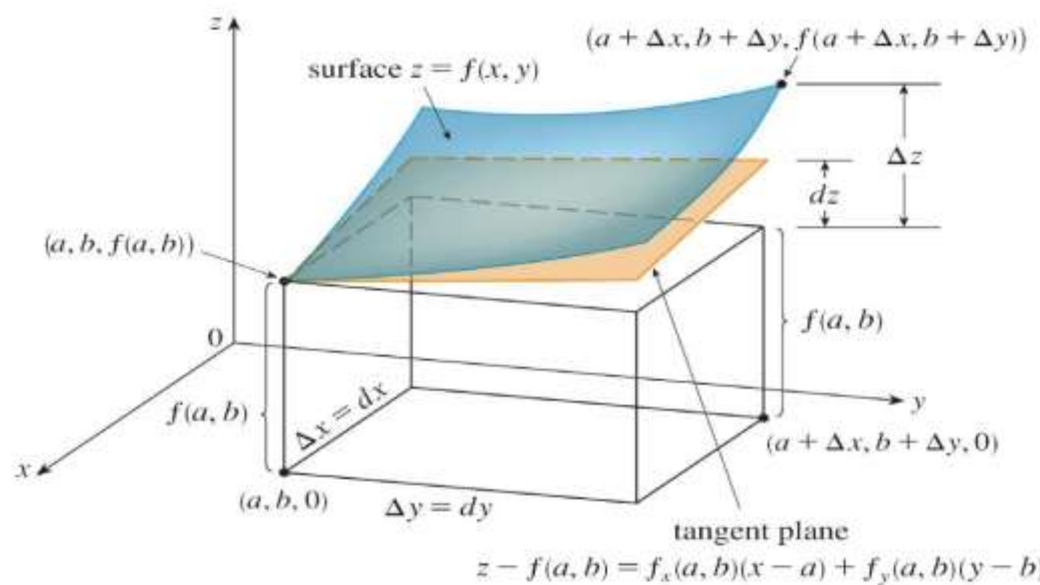


Figure 7

Example 4

(a) If $z = f(x, y) = x^2 + 3xy - y^2$, find the differential dz.

(b) If x changes from 2 to 2.05 and y changes from 3 to 2.96, compare the values of Δz and dz.

Solution:

(a) Definition 10 gives

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy$$

$$= (2x + 3y)\,dx + (3x - 2y)\,dy$$

Example 4 – *Solution*

(b) Putting $x = 2$, $dx = \Delta x = 0.05$, $y = 3$, and $dy = \Delta y = -0.04$, we get

$$dz = [2(2) + 3(3)]0.05 + [3(2) - 2(3)](-0.04) = 0.65$$

The increment of z is

$$\begin{aligned}\Delta z &= f(2.05, 2.96) - f(2, 3)\\ &= [(2.05)^2 + 3(2.05)(2.96) - (2.96)^2] \\ &\quad - [2^2 + 3(2)(3) - 3^2]\\ &= 0.6449\end{aligned}$$

Notice that $\Delta z \approx dz$ but dz is easier to compute.

Functions of Three or More Variables

Linear approximations, differentiability, and differentials can be defined in a similar manner for functions of more than two variables. A differentiable function is defined by an expression similar to the one in Definition 7. For such functions the **linear approximation** is

$$f(x, y, z) \approx f(a, b, c) + f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c)$$

and the linearization $L(x, y, z)$ is the right side of this expression.

Functions of Three or More Variables

If $w = f(x, y, z)$, then the **increment** of w is

$$\Delta w = f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z)$$

The **differential** dw is defined in terms of the differentials dx, dy, and dz of the independent variables by

$$dw = \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial y}dy + \frac{\partial w}{\partial z}dz$$

Example 6

The dimensions of a rectangular box are measured to be 75 cm, 60 cm, and 40 cm, and each measurement is correct to within 0.2 cm. Use differentials to estimate the largest possible error when the volume of the box is calculated from these measurements.

Solution:

If the dimensions of the box are *x*, *y*, and *z*, its volume is $V = xyz$ and so

$$dV = \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz$$

$$= yz\,dx + xz\,dy + xy\,dz$$

Example 6 – *Solution*

We are given that $|\Delta x| \le 0.2$, $|\Delta y| \le 0.2$, and $|\Delta z| \le 0.2$.

To estimate the largest error in the volume, we therefore use $dx = 0.2$, $dy = 0.2$, and $dz = 0.2$ together with $x = 75$, $y = 60$, and $z = 40$:

$$\Delta V \approx dV = (60)(40)(0.2) + (75)(40)(0.2) + (75)(60)(0.2)$$

$$= 1980$$

Example 6 – *Solution*

Thus an error of only 0.2 cm in measuring each dimension could lead to an error of approximately 1980 cm^3 in the calculated volume! This may seem like a large error, but it's only about 1% of the volume of the box.