Mathematics for AI

Adams Wai Kin Kong

School of Computer Science and Engineering

Nanyang Technological University, Singapore

adamskong@ntu.edu.sg

The Chain Rule

The Chain Rule

We know that the Chain Rule for functions of a single variable gives the rule for differentiating a composite function: If $y = f(x)$ and $x = g(t)$, where f and g are differentiable functions, then y is indirectly a differentiable function of t and

$$\boxed{1} \qquad \frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$$

For functions of more than one variable, the Chain Rule has several versions, each of them giving a rule for differentiating a composite function.

The Chain Rule

The first version (Theorem 2) deals with the case where $z = f(x, y)$ and each of the variables x and y is, in turn, a function of a variable t.

This means that z is indirectly a function of t, $z = f(g(t), h(t))$, and the Chain Rule gives a formula for differentiating z as a function of t. We assume that f is differentiable.

The Chain Rule

We know that this is the case when f_x and f_y are continuous.

2 The Chain Rule (Case 1) Suppose that $z = f(x, y)$ is a differentiable function of x and y, where $x = g(t)$ and $y = h(t)$ are both differentiable functions of t. Then z is a differentiable function of t and

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

Since we often write $\partial z/\partial x$ in place of $\partial f/\partial x$, we can rewrite the Chain Rule in the form

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

Example 1

If $z = x^2y + 3xy^4$, where $x = \sin 2t$ and $y = \cos t$, find dz/dt when $t = 0$.

Solution:

The Chain Rule gives

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

$$= (2xy + 3y^4)(2\cos 2t) + (x^2 + 12xy^3)(-\sin t)$$

It's not necessary to substitute the expressions for x and y in terms of t.

Example 1 – *Solution*

We simply observe that when $t = 0$, we have $x = \sin 0 = 0$ and $y = \cos 0 = 1$.

Therefore

$$\left.\frac{dz}{dt}\right|_{t=0} = (0 + 3)(2\cos 0) + (0 + 0)(-\sin 0) = 6$$

The Chain Rule

We now consider the situation where $z = f(x, y)$ but each of x and y is a function of two variables s and t:
$x = g(s, t), y = h(s, t)$.

Then z is indirectly a function of s and t and we wish to find $\partial z/\partial s$ and $\partial z/\partial t$.

We know that in computing $\partial z/\partial t$ we hold s fixed and compute the ordinary derivative of z with respect to t.

Therefore we can apply Theorem 2 to obtain

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}$$

The Chain Rule

A similar argument holds for $\partial z/\partial s$ and so we have proved the following version of the Chain Rule.

3 The Chain Rule (Case 2) Suppose that $z = f(x, y)$ is a differentiable function of x and y, where $x = g(s, t)$ and $y = h(s, t)$ are differentiable functions of s and t. Then

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \qquad \frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}$$

Case 2 of the Chain Rule contains three types of variables: s and t are **independent** variables, x and y are called **intermediate** variables, and z is the **dependent** variable.

The Chain Rule

- Notice that Theorem 3 has one term for each intermediate variable and each of these terms resembles the one-dimensional Chain Rule in Equation 1.

- To remember the Chain Rule, it's helpful to draw the **tree diagram** in Figure 2.

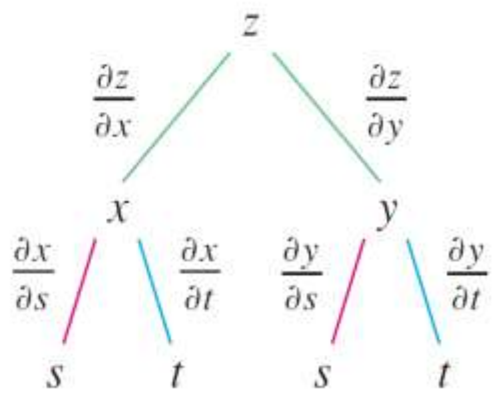


Figure 2

The Chain Rule

We draw branches from the dependent variable *z* to the intermediate variables *x* and *y* to indicate that *z* is a function of *x* and *y*. Then we draw branches from *x* and *y* to the independent variables *s* and *t*.

On each branch we write the corresponding partial derivative. To find $\partial z/\partial s$, we find the product of the partial derivatives along each path from *z* to *s* and then add these products:

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}$$

The Chain Rule

4 **The Chain Rule (General Version)** Suppose that u is a differentiable function of the n variables $x_1, x_2, \ldots, x_n$ and each x_j is a differentiable function of the m variables $t_1, t_2, \ldots, t_m$. Then u is a function of $t_1, t_2, \ldots, t_m$ and

$$\frac{\partial u}{\partial t_i} = \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial u}{\partial x_n}\frac{\partial x_n}{\partial t_i}$$

for each $i = 1, 2, \ldots, m$.

Implicit Differentiation

The Chain Rule can be used to give a more complete description of the process of implicit differentiation.

We suppose that an equation of the form $F(x, y) = 0$ defines y implicitly as a differentiable function of x, that is, $y = f(x)$, where $F(x, f(x)) = 0$ for all x in the domain of f.

If F is differentiable, we can apply Case 1 of the Chain Rule to differentiate both sides of the equation $F(x, y) = 0$ with respect to x.

Since both x and y are functions of x, we obtain

$$\frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0$$

Implicit Differentiation

But $dx/dx = 1$, so if $\partial F/\partial x \neq 0$ we solve for dy/dx and obtain

$$\boxed{6} \qquad \boxed{\frac{dy}{dx} = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial y}} = -\frac{F_x}{F_y}}$$

To derive this equation we assumed that $F(x, y) = 0$ defines y implicitly as a function of x.

Implicit Differentiation

The **Implicit Function Theorem**, proved in advanced calculus, gives conditions under which this assumption is valid: it states that if F is defined on a disk containing (a, b), where $F(a, b) = 0$, $F_y(a, b) \neq 0$, and F_x and F_y are continuous on the disk, then the equation $F(x, y) = 0$ defines y as a function of x near the point (a, b) and the derivative of this function is given by Equation 6.

Example 8

Find y' if $x^3 + y^3 = 6xy$.

Solution:

The given equation can be written as

$$F(x, y) = x^3 + y^3 - 6xy = 0$$

so Equation 6 gives

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{3x^2 - 6y}{3y^2 - 6x} = -\frac{x^2 - 2y}{y^2 - 2x}$$

Implicit Differentiation

Now we suppose that z is given implicitly as a function $z = f(x, y)$ by an equation of the form $F(x, y, z) = 0$.

This means that $F(x, y, f(x, y)) = 0$ for all (x, y) in the domain of f. If F and f are differentiable, then we can use the Chain Rule to differentiate the equation $F(x, y, z) = 0$ as follows:

$$\frac{\partial F}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0$$

Implicit Differentiation

But $\frac{\partial}{\partial x}(x) = 1$ and $\frac{\partial}{\partial x}(y) = 0$

so this equation becomes

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0$$

If $\partial F/\partial z \neq 0$, we solve for $\partial z/\partial x$ and obtain the first formula in Equations 7.

The formula for $\partial z/\partial y$ is obtained in a similar manner.

Implicit Differentiation

$$\boxed{7} \qquad \frac{\partial z}{\partial x} = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial z}} \qquad \frac{\partial z}{\partial y} = -\frac{\dfrac{\partial F}{\partial y}}{\dfrac{\partial F}{\partial z}}$$

Again, a version of the **Implicit Function Theorem** stipulates conditions under which our assumption is valid: if F is defined within a sphere containing (a, b, c), where $F(a, b, c) = 0$, $F_z(a, b, c) \neq 0$, and F_x, F_y, and F_z are continuous inside the sphere, then the equation $F(x, y, z) = 0$ defines z as a function of x and y near the point (a, b, c) and this function is differentiable, with partial derivatives given by (7).

Directional Derivatives

Directional Derivatives

Recall that if $z = f(x, y)$, then the partial derivatives f_x and f_y are defined as

$$f_x(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$$

1

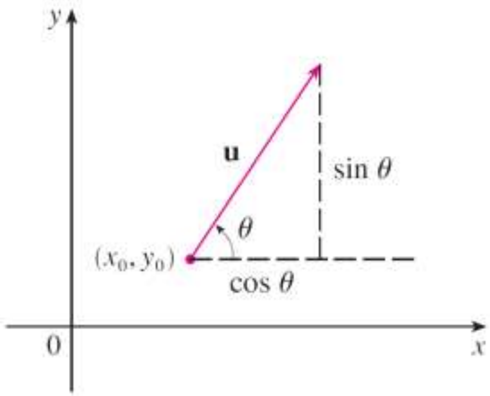
$$f_y(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}$$

and represent the rates of change of z in the x- and y-directions, that is, in the directions of the unit vectors **i** and **j**.

Directional Derivatives

Suppose that we now wish to find the rate of change of z at (x_0, y_0) in the direction of an arbitrary unit vector $\mathbf{u} = \langle a, b \rangle$. (See Figure 1.)

To do this we consider the surface S with the equation $z = f(x, y)$ (the graph of f) and we let $z_0 = f(x_0, y_0)$. Then the point $P(x_0, y_0, z_0)$ lies on S.



A unit vector $\mathbf{u} = \langle a, b \rangle = \langle \cos u, \sin u \rangle$

Figure 1

Directional Derivatives

The vertical plane that passes through *P* in the direction of **u** intersects *S* in a curve *C*. (See Figure 2.)

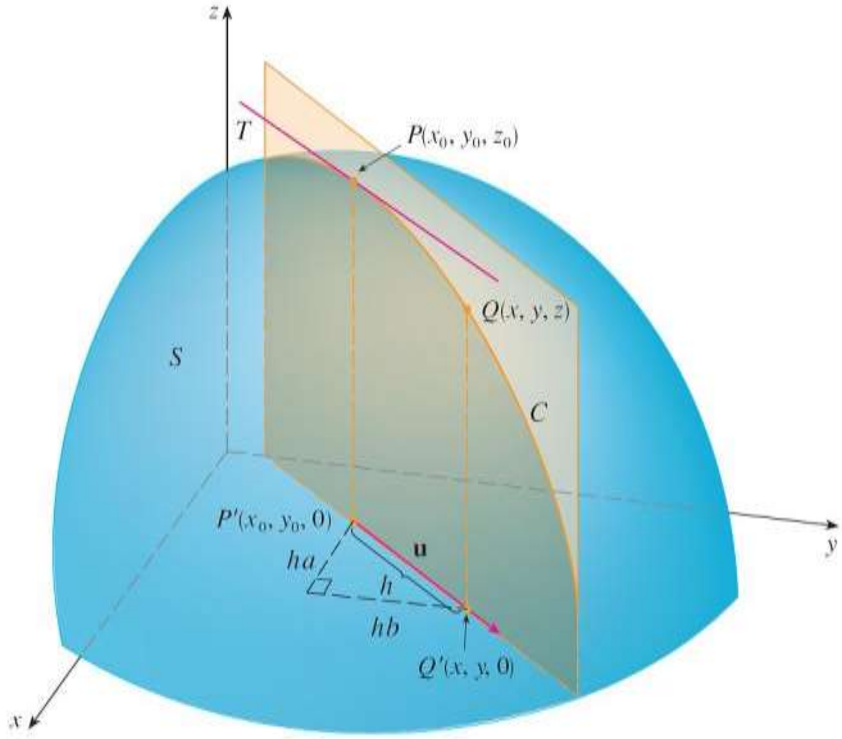


Figure 2

Directional Derivatives

The slope of the tangent line T to C at the point P is the rate of change of z in the direction of **u**.

If $Q(x, y, z)$ is another point on C and P', Q' are the projections of P, Q onto the xy-plane, then the vector is parallel to **u** and so

$$= h\mathbf{u} = \langle ha, hb \rangle$$

for some scalar h. Therefore $x - x_0 = ha$, $y - y_0 = hb$, so $x = x_0 + ha$, $y = y_0 + hb$, and

$$\frac{\Delta z}{h} = \frac{z - z_0}{h} = \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

Directional Derivatives

If we take the limit as $h \to 0$, we obtain the rate of change of z (with respect to distance) in the direction of **u**, which is called the directional derivative of *f* in the direction of **u**.

2 Definition The **directional derivative** of f at (x_0, y_0) in the direction of a unit vector $\mathbf{u} = \langle a, b \rangle$ is

$$D_{\mathbf{u}} f(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

if this limit exists.

Directional Derivatives

By comparing Definition 2 with Equations 1, we see that if $\mathbf{u} = \mathbf{i} = \langle 1, 0 \rangle$, then $D_{\mathbf{i}}f = f_x$ and if $\mathbf{u} = \mathbf{j} = \langle 0, 1 \rangle$, then $D_{\mathbf{j}}f = f_y$.

In other words, the partial derivatives of f with respect to x and y are just special cases of the directional derivative.

Example 1

Use the weather map in Figure 3 to estimate the value of the directional derivative of the temperature function at Reno in the southeasterly direction.

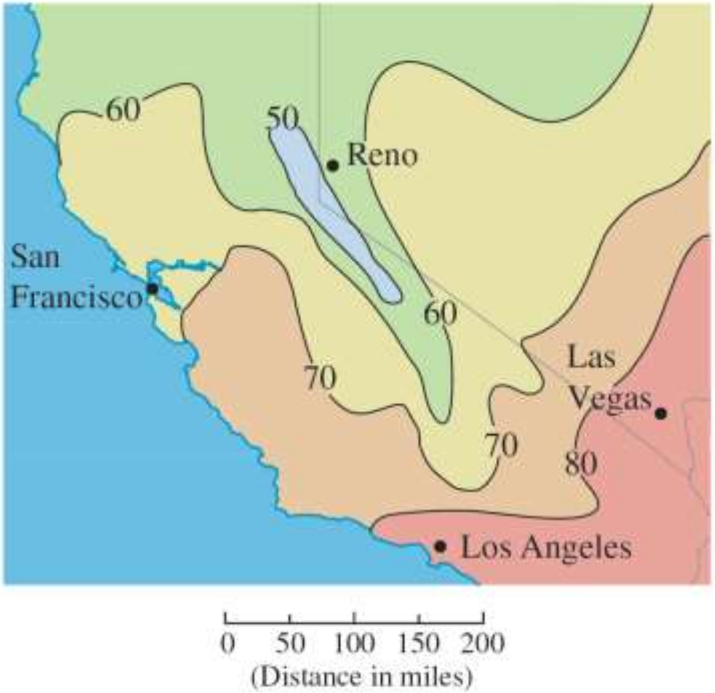


Figure 3

Example 1 – *Solution*

The unit vector directed toward the southeast is

$\mathbf{u} = (\mathbf{i} - \mathbf{j})/\sqrt{2}$, but we won't need to use this expression.

We start by drawing a line through Reno toward the southeast (see Figure 4).

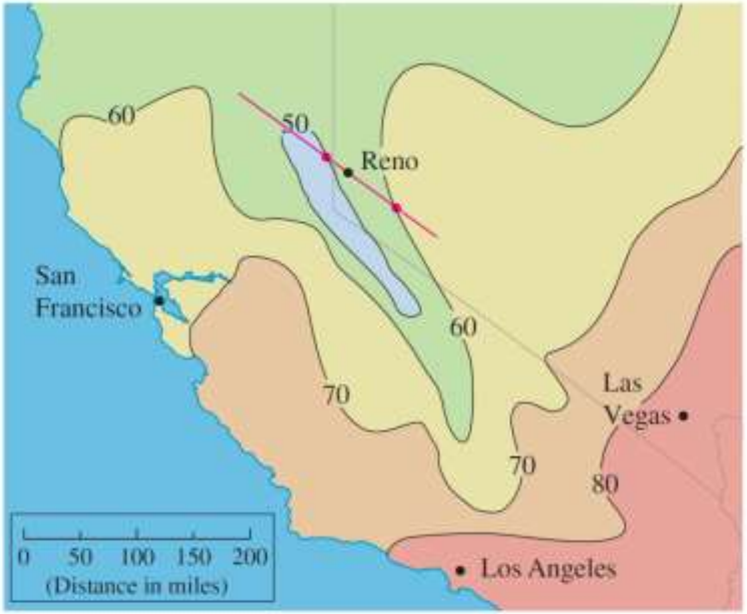


Figure 4

Example 1 – *Solution*

We approximate the directional derivative $D_{\mathbf{u}}T$ by the average rate of change of the temperature between the points where this line intersects the isothermals $T = 50$ and $T = 60$.

The temperature at the point southeast of Reno is $T = 60°F$ and the temperature at the point northwest of Reno is $T = 50°F$.

The distance between these points looks to be about 75 miles. So the rate of change of the temperature in the southeasterly direction is

$$D_{\mathbf{u}}T \approx \frac{60 - 50}{75} = \frac{10}{75} \approx 0.13°\text{F/mi}$$

Directional Derivatives

When we compute the directional derivative of a function defined by a formula, we generally use the following theorem.

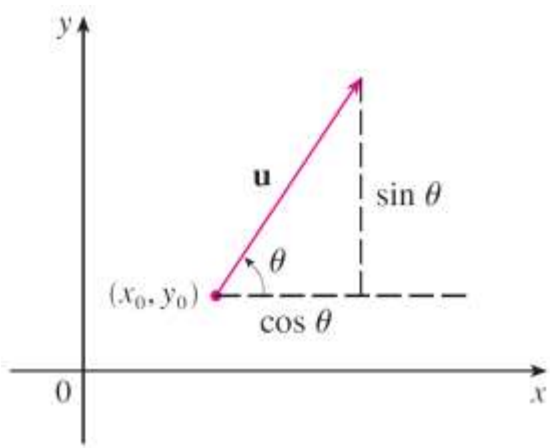
3 Theorem If f is a differentiable function of x and y, then f has a directional derivative in the direction of any unit vector $\mathbf{u} = \langle a, b \rangle$ and

$$D_{\mathbf{u}} f(x, y) = f_x(x, y)\, a + f_y(x, y)\, b$$

Directional Derivatives

If the unit vector **u** makes an angle θ with the positive x-axis (as in Figure 2), then we can write $\mathbf{u} = \langle \cos\theta, \sin\theta \rangle$ and the formula in Theorem 3 becomes

$$\boxed{6} \qquad D_{\mathbf{u}}f(x, y) = f_x(x, y)\cos\theta + f_y(x, y)\sin\theta$$



A unit vector **u** = ⟨*a*, *b*⟩ = ⟨cos *u*, sin *u*⟩

The Gradient Vector

Notice from Theorem 3 that the directional derivative of a differentiable function can be written as the dot product of two vectors:

$$\boxed{7} \qquad \begin{aligned} D_{\mathbf{u}}f(x, y) &= f_x(x, y)a + f_y(x, y)b \\ &= \langle f_x(x, y), f_y(x, y) \rangle \cdot \langle a, b \rangle \\ &= \langle f_x(x, y), f_y(x, y) \rangle \cdot \mathbf{u} \end{aligned}$$

The first vector in this dot product occurs not only in computing directional derivatives but in many other contexts as well.

So we give it a special name (the *gradient* of f) and a special notation (**grad** f or ∇f, which is read "del f").

The Gradient Vector

8 Definition If f is a function of two variables x and y, then the **gradient** of f is the vector function ∇f defined by

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$$

Example 3: If $f(x, y) = \sin x + e^{xy}$, then

$$\nabla f(x, y) = \langle f_x, f_y \rangle = \langle \cos x + ye^{xy}, xe^{xy} \rangle$$

and $$\nabla f(0, 1) = \langle 2, 0 \rangle$$

The Gradient Vector

With the notation for the gradient vector, we can rewrite Equation 7 for the directional derivative of a differentiable function as

9 $$\boxed{D_{\mathbf{u}} f(x, y) = \nabla f(x, y) \cdot \mathbf{u}}$$

This expresses the directional derivative in the direction of a unit vector **u** as the scalar projection of the gradient vector onto **u**.

Functions of Three Variables

For functions of three variables we can define directional derivatives in a similar manner.

Again $D_{\mathbf{u}}f(x, y, z)$ can be interpreted as the rate of change of the function in the direction of a unit vector **u**.

10 Definition The **directional derivative** of f at (x_0, y_0, z_0) in the direction of a unit vector $\mathbf{u} = \langle a, b, c \rangle$ is

$$D_{\mathbf{u}}f(x_0, y_0, z_0) = \lim_{h \to 0} \frac{f(x_0 + ha, y_0 + hb, z_0 + hc) - f(x_0, y_0, z_0)}{h}$$

if this limit exists.

Functions of Three Variables

If we use vector notation, then we can write both definitions (2 and 10) of the directional derivative in the compact form

11

$$D_{\mathbf{u}} f(\mathbf{x}_0) = \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\mathbf{u}) - f(\mathbf{x}_0)}{h}$$

where $\mathbf{x}_0 = \langle x_0, y_0 \rangle$ if $n = 2$ and $\mathbf{x}_0 = \langle x_0, y_0, z_0 \rangle$ if $n = 3$.

This is reasonable because the vector equation of the line through $\mathbf{x}_0$ in the direction of the vector $\mathbf{u}$ is given by $\mathbf{x} = \mathbf{x}_0 + t\,\mathbf{u}$ and so $f(\mathbf{x}_0 + h\mathbf{u})$ represents the value of f at a point on this line.

Functions of Three Variables

If $f(x, y, z)$ is differentiable and $\mathbf{u} = \langle a, b, c \rangle$, then

12 $$D_{\mathbf{u}}f(x, y, z) = f_x(x, y, z)a + f_y(x, y, z)b + f_z(x, y, z)c$$

For a function f of three variables, the **gradient vector**, denoted by ∇f or **grad** f, is

$$\nabla f(x, y, z) = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z) \rangle$$

or, for short,

13 $$\boxed{\nabla f = \langle f_x, f_y, f_z \rangle = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}}$$

Functions of Three Variables

Then, just as with functions of two variables, Formula 12 for the directional derivative can be rewritten as

14

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

Example 5

If $f(x, y, z) = x \sin(yz)$, (a) find the gradient of f and (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

$$\nabla f(x, y, z) = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)\rangle$$

$$= \langle \sin yz, xz \cos yz, xy \cos yz\rangle$$

Example 5 – *Solution*

(b) 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

$$\mathbf{u} = \frac{1}{\sqrt{6}}\mathbf{i} + \frac{2}{\sqrt{6}}\mathbf{j} - \frac{1}{\sqrt{6}}\mathbf{k}$$

Therefore Equation 14 gives

$$D_{\mathbf{u}} f(1, 3, 0) = \nabla f(1, 3, 0) \cdot \mathbf{u} = 3\mathbf{k} \cdot \left(\frac{1}{\sqrt{6}}\mathbf{i} + \frac{2}{\sqrt{6}}\mathbf{j} - \frac{1}{\sqrt{6}}\mathbf{k} \right)$$

$$= 3\left(-\frac{1}{\sqrt{6}} \right) = -\sqrt{\frac{3}{2}}$$

Maximizing the Directional Derivatives

Suppose we have a function *f* of two or three variables and we consider all possible directional derivatives of *f* at a given point.

These give the rates of change of *f* in all possible directions.

We can then ask the questions: In which of these directions does *f* change fastest and what is the maximum rate of change? The answers are provided by the following theorem.

15 Theorem Suppose f is a differentiable function of two or three variables. The maximum value of the directional derivative $D_{\mathbf{u}} f(\mathbf{x})$ is $|\nabla f(\mathbf{x})|$ and it occurs when $\mathbf{u}$ has the same direction as the gradient vector $\nabla f(\mathbf{x})$.

Example 6

(a) If $f(x, y) = xe^y$, find the rate of change of f at the point $P(2, 0)$ in the direction from P to $Q(\frac{1}{2}, 2)$.

(b) In what direction does f have the maximum rate of change? What is this maximum rate of change?

Solution:

(a) We first compute the gradient vector:

$$\nabla f(x, y) = \langle f_x, f_y \rangle = \langle e^y,\ xe^y \rangle$$

$$\nabla f(2, 0) = \langle 1, 2 \rangle$$

Example 6 – *Solution*

The unit vector in the direction $\overrightarrow{PQ} = \langle -\frac{3}{2}, 2 \rangle$ is $\mathbf{u} = \langle -\frac{3}{5}, \frac{4}{5} \rangle$, so the rate of change of *f* in the direction from *P* to *Q* is

$$\begin{aligned} D_{\mathbf{u}}f(2, 0) &= \nabla f(2, 0) \cdot \mathbf{u} \\ &= \langle 1, 2 \rangle \cdot \langle -\tfrac{3}{5}, \tfrac{4}{5} \rangle \\ &= 1\left(-\tfrac{3}{5}\right) + 2\left(\tfrac{4}{5}\right) = 1 \end{aligned}$$

(b) According to Theorem 15, *f* increases fastest in the direction of the gradient vector $\nabla f(2, 0) = \langle 1, 2 \rangle$. The maximum rate of change is

$$|\nabla f(2, 0)| = |\langle 1, 2 \rangle| = \sqrt{5}$$

Tangent Planes to Level Surfaces

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. For example, $x^2 + y^2 + z^2 = 25$. It is a surface of a ball.

Let C be any curve that lies on the surface S and passes through the point P. Recall that 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,

$$F(x(t), y(t), z(t)) = k \tag{16}$$

NANYANG TECHNOLOGICAL UNIVERSITY

Tangent Planes to Level Surfaces

If x, y, and z are differentiable functions of t and F is also differentiable, then we can use the Chain Rule to differentiate both sides of Equation 16 as follows:

$$\boxed{17} \qquad \frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt} + \frac{\partial F}{\partial z}\frac{dz}{dt} = 0$$

But, since $\nabla F = \langle F_x, F_y, F_z \rangle$ and $\mathbf{r}'(t) = \langle x'(t), y'(t), z'(t) \rangle$, Equation 17 can be written in terms of a dot product as

$$\nabla F \cdot \mathbf{r}'(t) = 0$$

Tangent Planes to Level Surfaces

In particular, when $t = t_0$ we have $\mathbf{r}(t_0) = \langle x_0, y_0, z_0 \rangle$, so

$$\boxed{18} \qquad \nabla F(x_0, y_0, z_0) \cdot \mathbf{r}'(t_0) = 0$$

Equation 18 says that *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*. (See Figure 9.)

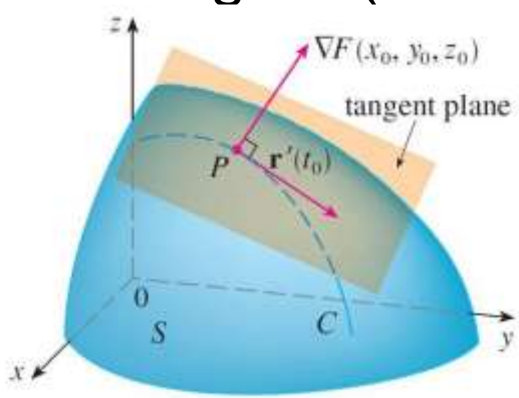


Figure 9

Tangent Planes to Level Surfaces

If $\nabla F(x_0, y_0, z_0) \neq \mathbf{0}$, it is therefore natural to define the **tangent plane to the level surface** $F(x, y, z) = k$ **at** $P(x_0, y_0, z_0)$ as the plane that passes through P and has normal vector $\nabla F(x_0, y_0, z_0)$.

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

$$\boxed{19} \quad F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0$$

Tangent Planes to Level Surfaces

The **normal line** to S at P is the line passing through P and perpendicular to the tangent plane.

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

20
$$\frac{x - x_0}{F_x(x_0, y_0, z_0)} = \frac{y - y_0}{F_y(x_0, y_0, z_0)} = \frac{z - z_0}{F_z(x_0, y_0, z_0)}$$

Tangent Planes to Level Surfaces

In the special case in which the equation of a surface S is of the form $z = f(x, y)$ (that is, S is the graph of a function f of two variables), we can rewrite the equation as

$$F(x, y, z) = f(x, y) - z = 0$$

and regard S as a level surface (with $k = 0$) of F. Then

$$F_x(x_0, y_0, z_0) = f_x(x_0, y_0)$$
$$F_y(x_0, y_0, z_0) = f_y(x_0, y_0)$$
$$F_z(x_0, y_0, z_0) = -1$$

so Equation 19 becomes

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - z_0) = 0$$

Example 8

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

$$F(x, y, z) = \frac{x^2}{4} + y^2 + \frac{z^2}{9}$$

Example 8 – *Solution*

Therefore we have

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

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

Then Equation 19 gives the equation of the tangent plane at (–2, 1, –3) as

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

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

By Equation 20, symmetric equations of the normal line are

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

Significance of the Gradient Vector

We now summarize the ways in which the gradient vector is significant.

We first consider a function f of three variables and a point $P(x_0, y_0, z_0)$ in its domain.

On the one hand, we know from Theorem 15 that the gradient vector $\nabla f(x_0, y_0, z_0)$ gives the direction of fastest increase of f.

> **15 Theorem** Suppose f is a differentiable function of two or three variables. The maximum value of the directional derivative $D_{\mathbf{u}} f(\mathbf{x})$ is $|\nabla f(\mathbf{x})|$ and it occurs when $\mathbf{u}$ has the same direction as the gradient vector $\nabla f(\mathbf{x})$.

Significance of the Gradient Vector

On the other hand, we know that $\nabla f(x_0, y_0, z_0)$ is orthogonal to the level surface S of f through P. (Refer to Figure 9.)

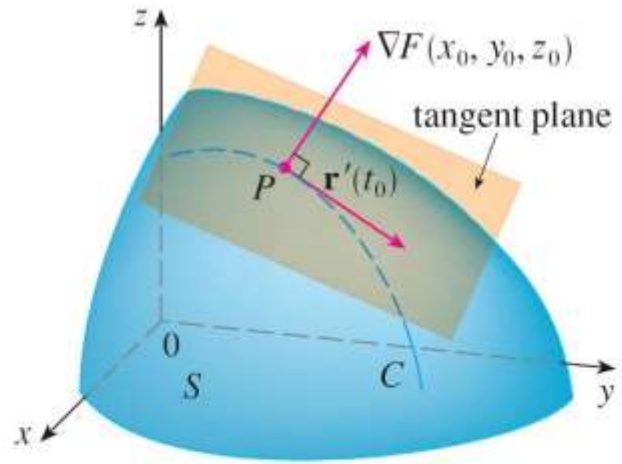


Figure 9

These two properties are quite compatible intuitively because as we move away from P on the level surface S, the value of f does not change at all.

Significance of the Gradient Vector

So it seems reasonable that if we move in the perpendicular direction, we get the maximum increase.

In like manner we consider a function f of two variables and a point $P(x_0, y_0)$ in its domain.

Again the gradient vector $\nabla f(x_0, y_0)$ gives the direction of fastest increase of f. Also, by considerations similar to our discussion of tangent planes, it can be shown that $\nabla f(x_0, y_0)$ is perpendicular to the level curve $f(x, y) = k$ that passes through P.

Significance of the Gradient Vector

Again this is intuitively plausible because the values of f remain constant as we move along the curve.
(See Figure 11.)

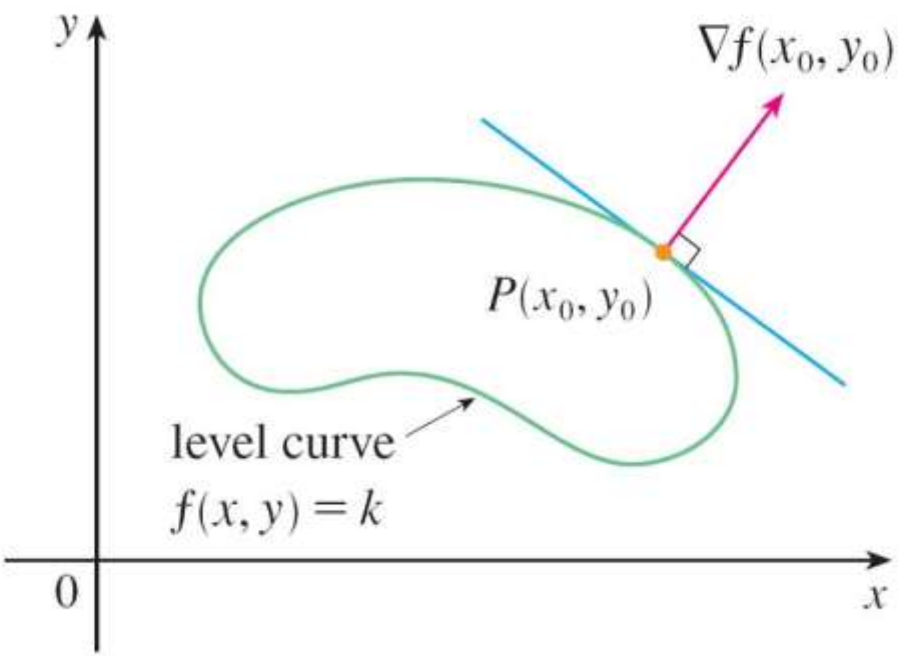


Figure 11

Significance of the Gradient Vector

Computer algebra systems have commands that plot sample gradient vectors.

Each gradient vector $\nabla f(a, b)$ is plotted starting at the point (a, b). Figure 13 shows such a plot (called a *gradient vector field*) for the function $f(x, y) = x^2 - y^2$ superimposed on a contour map of f.

As expected, the gradient vectors point "uphill" and are perpendicular to the level curves.

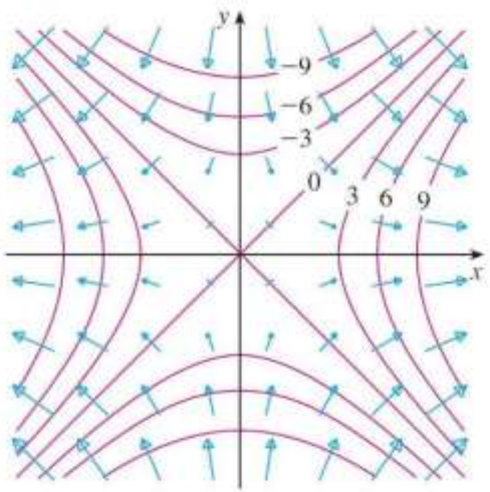


Figure 13