# CALCULUS REVIEW

TO PREPARE
TO UNDERSTAND
NEURAL NETWORK
LEARNING

J. WUNDERLICH, Ph.D.

CALC. REVIEW ~~HOMEWORK~~ ~~1) CHAIN RULE OF DIFFERENTIATION~~

LOOK UP: ~~2) QUOTIENT RULE OF DIFFERENTIATION~~

~~f(x) AND DO:~~

1) IF $f(x)=2x$ $\quad f'(x)=2$

2) IF $f(x,y)= yx$ $\quad \frac{\partial f}{\partial x}=y$

3) IF $f(x)=\frac{1}{2}(1-x)^2$ $\quad f'(x)=\frac{1}{2}\left(2(1-x)(-1)\right)=-(1-x)$

4) IF $f(x,y)=\frac{1}{2}(y-x)^2$ $\quad \frac{\partial f}{\partial x}=\frac{1}{2}\left(2(y-x)(-1)\right)=-(y-x)$

~~5) IF $f(x)=\frac{1}{1+e^{-x}}$ $\quad f'(x)=?$~~

6) IF $f(x)=2$ $\quad f'(x)=0$

7) IF $f(x)=e^{-2x}$ $\quad f'(x)=\left(e^{-2x}\right)(-2)=-2e^{-2x}$

QUIZ ON FRIDAY

$$f(x)=\frac{x^2}{1-e^x} \quad f'(x)=?$$

USE QUOTIENT RULE

$$f(x)=\frac{u(x)}{v(x)}$$

AND

$$f'(x)=\frac{v(x)u'(x)-u(x)v'(x)}{(v(x))^2}$$

FOR US

$$f(x)=\frac{u(x)}{v(x)}=\frac{x^2}{1-e^x}$$

$u(x)=x^2 \quad u'(x)=2x$

$v(x)=1-e^x \quad v'(x)=0-e^x$

$$f'(x)=\frac{(1-e^x)(2x)-x^2(-e^x)}{(1-e^x)^2}$$

$$=\frac{2x-2xe^x+x^2e^x}{(1-e^x)^2}$$

---

~~$f(x,y)=57$~~

~~$f'(x,y)$~~

GET CHAIN RULE EXAMPLE FROM CS/ENGR 433 NOTES: (I.E. "RIDDLE" CALCULUS BOOK)

GET LEAST MEAN EXAMPLE FROM:

CS 344



CHAIN RULE

*Theorem 4.8*

$$f'(x) = \lim_{t\to x}\frac{f(t)-f(x)}{t-x} = \lim_{t\to x}\frac{f(x)-f(t)}{x-t}.$$

This theorem also can be proved directly from the definition of derivative without recourse to the above geometric argument, by making the substitution $t = x + h$. For some purposes, this new form for the derivative is more convenient than the definition. Some authors use this as the definition and derive the other form from it.

Our next theorem uses the result of Theorem 4.8, but first let us take up a question touched upon earlier in connection with notation. Suppose $y$ is a function of $u$ while $u$ is in turn a function of $x$, in which case $y$ is a function of $x$ as well as a function of $u$. Thus, if $y = f(u)$ and $u = g(x)$, then

$$y = f(g(x)) = F(x).$$

We may now consider two derivatives of $y$—the derivative of $y$ with respect to $u$, which is $dy/du$, and the derivative of $y$ with respect to $x$, which is $dy/dx$. Now when these derivatives are presented in the limit forms, how do we distinguish one from another? Basically it is done merely by noting that there are two different functions, $f(u)$ and $F(x)$. The important thing to note here is not the difference between $u$ and $x$ but the difference between the *functions* $f$ and $F$. Thus

$$\frac{dy}{du} = \lim_{h\to 0}\frac{f(u+h)-f(u)}{h} = \lim_{t\to u}\frac{f(t)-f(u)}{t-u},$$

while

$$\frac{dy}{dx} = \lim_{h\to 0}\frac{F(x+h)-F(x)}{h} = \lim_{t\to x}\frac{F(t)-F(x)}{t-x}.$$

Since $u = g(x)$ and $F(x) = f(g(x))$, both can be written in other forms. Thus we get

$$\frac{dy}{du} = \lim_{t\to u}\frac{f(t)-f(u)}{t-u} = \lim_{g(s)\to g(x)}\frac{f(g(s))-f(g(x))}{g(s)-g(x)}$$

simply by substituting $u = g(x)$ and $t = g(s)$. Note that the letters used here are immaterial; if $t$ had not been used for something else, it might have been used instead of $s$, thus:

$$\frac{dy}{du} = \lim_{g(t)\to g(x)}\frac{f(g(t))-f(g(x))}{g(t)-g(x)}.$$

Similarly,

$$\frac{dy}{dx} = \lim_{t\to x}\frac{F(t)-F(x)}{t-x} = \lim_{t\to x}\frac{f(g(t))-f(g(x))}{t-x}.$$

*Theorem 4.9*

(*Chain rule*) *If $f$ and $g$ are functions such that $g$ is differentiable at $x$ and $f$ is differentiable at $u = g(x)$ and $F$ is a function such that* $F(x) = f(g(x))$, *then*

$$\frac{d}{dx}F(x) = \frac{d}{du}f(u)\cdot\frac{d}{dx}g(x).$$

A simpler, but less accurate statement of this theorem is: If $y = f(u)$ and $u = g(x)$, then

$$\boxed{\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}.}$$

Note here that the differential notation is an aid to the memory. While you are reminded not to think of $dy/du$ and $du/dx$ as fractions, they seem to behave as if they were.

***Proof***

Since $y = f(u)$ and $u = g(x)$, $\quad y = f(g(x)) = F(x)$.

$$\frac{dy}{dx} = \lim_{t \to x} \frac{F(t) - F(x)}{t - x}$$

$$= \lim_{t \to x} \frac{f(g(t)) - f(g(x))}{t - x}.$$

Now let us stop and take stock. Given

$$\lim_{t \to x} \frac{f(g(t)) - f(g(x))}{t - x},$$

we prefer to have

$$\lim_{t \to x} \frac{g(t) - g(x)}{t - x}.$$

Thus we multiply both numerator and denominator by $g(t) - g(x)$ to get

$$\frac{dy}{dx} = \lim_{t \to x} \frac{f(g(t)) - f(g(x))}{g(t) - g(x)} \cdot \frac{g(t) - g(x)}{t - x}$$

$$= \left[\lim_{t \to x} \frac{f(g(t)) - f(g(x))}{g(t) - g(x)}\right] \left[\lim_{t \to x} \frac{g(t) - g(x)}{t - x}\right]$$

$$= \frac{dy}{du} \frac{du}{dx}. \qquad \text{(See Note.)}$$

Actually there is a problem here. We have no assurance that $g(t) - g(x) \neq 0$. We can remedy the situation by splitting the argument into two cases.

*Case I:* $g(t) \neq g(x)$ for any $t$ within a distance $H$ of $x$. By choosing $t$ close enough to $x$, we are assured that $g(t) - g(x) \neq 0$. Thus, for $|t - x|$ small enough ($|t - x| < H$),

$$\frac{dy}{dx} = \lim_{t \to x} \frac{f(g(t)) - f(g(x))}{g(t) - g(x)} \cdot \frac{g(t) - g(x)}{t - x}$$

$$= \frac{df(u)}{du} \cdot \frac{dg(x)}{dx} \qquad \text{(See Note.)}$$

$$= \frac{dy}{du} \cdot \frac{du}{dx}.$$

*Case II:* No matter how small $H$ is, there is a number $t$ within a distance $H$ of $x$ such that $g(t) = g(x)$. Note that this does not imply that $g(t) = g(x)$ for *all* $t$

within a distance $H$ of $x$; it does imply that $g(t) = g(x)$ for infinitely many values of $t$ within a distance $H$ of $x$. In any case

$$\lim_{t \to x} \frac{g(t) - g(x)}{t - x} = 0,$$

since we know that this limit (which is $g'(x)$) exists and

$$\frac{g(t) - g(x)}{t - x} = 0$$

for values of $t$ arbitrarily close to $x$. Now let us consider

$$\lim_{t \to x} \frac{F(t) - F(x)}{t - x}.$$

First let $M$ be the set of all numbers $t$ of the domain of $F$ such that $t$ is within a distance $H$ of $x$ and $g(t) \neq g(x)$; let $N$ be the set of all numbers $t$ of the domain of $F$ such that $t$ is within a distance $H$ of $x$ and $g(t) = g(x)$. If $M$ contains numbers arbitrarily close to $x$, then (by Case I)

$$\lim_{\substack{t \to x \\ t \text{ in } M}} \frac{F(t) - F(x)}{t - x} = \frac{dy}{du} \cdot \frac{du}{dx} = \frac{dy}{du} \cdot 0 = 0.$$

If $M$ does not contain numbers arbitrarily close to $x$, we need not consider this limit. In addition,

$$\lim_{\substack{t \to x \\ t \text{ in } N}} \frac{F(t) - F(x)}{t - x} = \lim_{\substack{t \to x \\ t \text{ in } N}} \frac{f(g(t)) - f(g(x))}{t - x} = \lim_{\substack{t \to x \\ t \text{ in } N}} \frac{f(g(x)) - f(g(x))}{t - x}$$

$$= 0.$$

Thus the combination of these two limits gives

$$\frac{dy}{dx} = \lim_{t \to x} \frac{F(t) - F(x)}{t - x} = 0 = \frac{dy}{du} \cdot \frac{du}{dx}.$$

*Note:* While we had indicated that

$$\frac{dy}{du} = \lim_{g(t) \to g(x)} \frac{f(g(t)) - f(g(x))}{g(t) - g(x)},$$

the above limit is

$$\lim_{t \to x} \frac{f(g(t)) - f(g(x))}{g(t) - g(x)}.$$

We are assuming again that as $t$ approaches $x$, $g(t)$ approaches $g(x)$. Here we are using the fact, to be proved later, that the existence of $g'(x)$ implies the continuity of $g(x)$. Thus these two limits are equivalent.

***Example 2***

If $y = u^2$ and $u = x^2 - 4x + 3$, find $dy/dx$.

By the chain rule, $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$

$$= 2u(2x - 4)$$

$$= 4(x^2 - 4x + 3)(x - 2).$$

**Example 3**

Find $y'$ for $y = \dfrac{1}{(x^2+3x-5)^3}$.

Let us make the substitution $u = x^2 + 3x - 5$. Then $y = u^{-3}$; and, by the chain rule,

$$\begin{aligned}\frac{dy}{dx} &= \frac{dy}{du}\cdot\frac{du}{dx}\\ &= -3u^{-4}(2x+3)\\ &= \frac{-3(2x+3)}{(x^2+3x-5)^4}.\end{aligned}$$

The chain rule is very important and we shall use it to prove the next two theorems, as well as several others.

***Theorem 4.10***

*If $f$ is a function such that $f(x) = x^n$, where $n$ is any rational number, then $y' = nx^{n-1}$.*

***Proof***

First, $n = p/q$, where $p$ and $q$ are integers and $q$ is positive. Now we shall split the argument into two parts. We shall show first that $y' = nx^{n-1}$ if $n = 1/q$ (where $q$ is a positive integer); then we shall use this together with the chain rule to prove the general case.

Suppose that $y = x^{1/q}$, where $q$ is a positive integer. Then

$$y' = \lim_{t\to x}\frac{t^{1/q}-x^{1/q}}{t-x}.$$

Using the substitutions $z = x^{1/q}$ and $s = t^{1/q}$, we get

$$\begin{aligned}y' &= \lim_{s^q\to z^q}\frac{s-z}{s^q-z^q}\\ &= \lim_{s\to z}\frac{s-z}{(s-z)(s^{q-1}+s^{q-2}z+s^{q-3}z^2+\cdots+sz^{q-2}+z^{q-1})}\\ &= \lim_{s\to z}\frac{1}{(s^{q-1}+s^{q-2}z+s^{q-3}z^2+\cdots+sz^{q-2}+z^{q-1})}\\ &= \frac{1}{qz^{q-1}}\\ &= \frac{1}{q}z^{1-q}\\ &= \frac{1}{q}(x^{1/q})^{1-q}\\ &= \frac{1}{q}x^{(1/q)-1}.\end{aligned}$$

Thus, if $n = 1/q$ and $y = x^n$, then $y' = nx^{n-1}$.

Now suppose $y = x^{p/q}$, where $p$ and $q$ are integers and $q$ is positive. Then $y = (x^{1/q})^p$. By substituting $u = x^{1/q}$ we have $y = u^p$.

$$\begin{aligned}
\frac{dy}{dx} &= \frac{dy}{du} \cdot \frac{du}{dx} \\
&= pu^{p-1} \cdot \frac{1}{q} x^{(1/q)-1} \\
&= \frac{p}{q} (x^{1/q})^{p-1} x^{(1/q)-1} \\
&= \frac{p}{q} x^{(p/q)-(1/q)} x^{(1/q)-1} \\
&= \frac{p}{q} x^{(p/q)-1}.
\end{aligned}$$

Thus, the formula which was first stated for positive integers in Theorem 4.2 was extended to all integers in Theorem 4.7 and now applies to all rational numbers. We shall not be in a position to prove it for all real numbers until Chapter 12 (see Theorem 12.6). ~~Nevertheless, the extensions we have made allow us to find derivatives of a wide range of functions.~~

**Example 4**

Differentiate $y = \sqrt{x}$.

$$y = \sqrt{x} = x^{1/2}, \qquad y' = \frac{1}{2} x^{-1/2} = \frac{1}{2\sqrt{x}}.$$

**Example 5**

Differentiate $y = \dfrac{x+2}{\sqrt{x}}$.

There are several possible methods. Two are illustrated.

$$\begin{aligned}
y &= \frac{x+2}{x^{1/2}}, \\
y' &= \frac{x^{1/2} \cdot 1 - (x+2)\frac{1}{2} x^{-1/2}}{x} \\
&= \frac{\sqrt{x} - \dfrac{x+2}{2\sqrt{x}}}{x} \\
&= \frac{2x - (x+2)}{2x^{3/2}} \\
&= \frac{x-2}{2x^{3/2}}.
\end{aligned}$$

$$\begin{aligned}
y &= x^{1/2} + 2x^{-1/2}, \\
y' &= \frac{1}{2} x^{-1/2} - x^{-3/2} \\
&= \frac{1}{2x^{1/2}} - \frac{1}{x^{3/2}} \\
&= \frac{x-2}{2x^{3/2}}.
\end{aligned}$$

In the second method of Example 5 we have avoided the relatively complicated quotient formula by carrying out the division and using negative exponents. While this method is not universally recommended, it sometimes simplifies a problem considerably.

positive. Then

eorem 4.2 was
nal numbers.
apter 12 (see
to find deriv-

complicated
nents. While
problem con-



**Example 6**

Differentiate $y = \frac{1}{x}$.

QUOTIENT RULE

$$y' = \frac{x \cdot 0 - 1 \cdot 1}{x^2} = \frac{-1}{x^2}.$$

However, by writing the original problem as $y = x^{-1}$, we get

SIMPLER:

$$y' = -x^{-2} = -\frac{1}{x^2}.$$

The use of negative exponents makes the problem simple enough to do in your head. This method can be used to advantage when the denominator is very simple or when the numerator is a constant.

### Problems

*A*

In Problems 1–12, differentiate.

1. $y = \frac{1}{x^2}$.

2. $y = \frac{1}{x^5}$.

3. $y = \frac{x^2 + 1}{x^2}$.

4. $y = \frac{x^2 - 4}{x}$.

5. $y = x^{2/3} - a^{2/3}$ ($a$ is a constant).

6. $y = \frac{x}{x^2 + 4}$.

7. $y = \frac{x^3}{x - 1}$.

8. $y = \frac{x(x + 2)}{x + 1}$.

9. $s = t^{2/3} - t^{-1/3}$.

10. $s = \frac{3}{4}t^{4/3} + 3t^{1/3} + \frac{3}{2}t^{-2/3}$.

11. $u = \frac{v^2 + 2v - 2}{v^2 - 2v + 2}$.

12. $u = \frac{v^2 + 4a^2}{v^2 - 4a^2}$ ($a$ is a constant).

In Problems 13–16, use the chain rule to find $dy/dx$.

13. $y = \sqrt{u}, \quad u = 2x^2 - 3$.

14. $y = u^2 + 1, \quad u = 2x + 5$.

15. $y = u^3 + u, \quad u = 3x - 2$.

16. $y = u\sqrt{u + 1}, \quad u = 2x + 3$.

In Problems 17–22, find the derivative at the point indicated.

17. $y = 2x^{1/2} - 3x^{1/3}$, at $(1, -1)$.

18. $y = \frac{x - 1}{x + 2}$, at $(1, 0)$.

19. $y = \frac{x^2 + 1}{4x^2 - 9}$, at $x = 2$.

20. $y = \sqrt[3]{x} - 1$, at $(8, 1)$.

21. $y = \frac{x^3 + 1}{x^2 - 2}$, at $(-1, 0)$.

22. $y = \frac{\sqrt{x} - 1}{\sqrt[3]{x} + 1}$, at $x = 64$.

# 21.3

## The Chain Rule

Recall that in Section 4.4 we considered the chain rule for differentiating a function of a function. For $y = f(u)$ and $u = g(x)$, we had $y = f(g(x))$ and

$$\boxed{\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.}$$

This proved to be very useful, especially when deriving formulas for the derivatives of transcendental functions. In this section we extend the chain rule to functions of several variables.

### Theorem 21.2

*Suppose $z = f(x, y)$, and $x(t)$ and $y(t)$ are differentiable functions in an open interval containing $t$. If $\partial z/\partial x$ and $\partial z/\partial y$ are continuous in a neighborhood of $(x(t), y(t))$, then*

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

***Proof***

Before proving the chain rule, let us derive a useful increment formula for $f(x, y)$. Let $\Delta x$ and $\Delta y$ be increments in $x$ and $y$, respectively. The corresponding increment in $z = f(x, y)$ is

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

Adding and subtracting $f(x, y + \Delta y)$ and applying the mean-value theorem to the resulting differences,

$$\begin{aligned}\Delta z &= f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y) + f(x, y + \Delta y) - f(x, y)\\ &= f_x(x_1, y + \Delta y)\Delta x + f_y(x, y_1)\Delta y,\end{aligned}$$

where $x_1$ is between $x$ and $x + \Delta x$ and $y_1$ is between $y$ and $y + \Delta y$. Since $\partial z/\partial x = f_x$ and $\partial z/\partial y = f_y$ are continuous,

$$\begin{aligned}f_x(x_1, y + \Delta y) &= f_x(x, y) + \epsilon,\\ f_y(x, y_1) &= f_y(x, y) + \eta,\end{aligned}$$

where $\epsilon \to 0$ and $\eta \to 0$ as $(\Delta x, \Delta y) \to (0, 0)$. (The value 0 may be assumed by $\epsilon$ and $\eta$ here.) Thus we have the *fundamental increment formula*,

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y) = f_x(x, y)\Delta x + f_y(x, y)\Delta y + \epsilon\Delta x + \eta\Delta y,$$

where $\epsilon \to 0$ and $\eta \to 0$ as $(\Delta x, \Delta y) \to (0, 0)$.

Suppose now that $z = f(x, y)$, $x = x(t)$, and $y = y(t)$. It then follows that $z$ is a function of $t$, $z = f(x(t), y(t))$, and we can consider the derivative of $z$ with respect to $t$.

$$\frac{dz}{dt} = \lim_{h \to 0} \frac{f(x(t+h), y(t+h)) - f(x(t), y(t))}{h}.$$

Corresponding to the increment $h$ in $t$ there are increments $\Delta x = x(t+h) - x(t)$ and $\Delta y = y(t+h) - y(t)$. Also, since $dx/dt$ and $dy/dt$ exist, $x(t)$ and $y(t)$ are continuous by Theorem 10.10, so both $\Delta x \to 0$ and $\Delta y \to 0$ as $h \to 0$. By the fundamental increment formula,

$$\begin{aligned}
&\frac{1}{h}[f(x(t+h), y(t+h)) - f(x(t), y(t))] \\
&\quad = \frac{1}{h}[f(x(t) + \Delta x, y(t) + \Delta y) - f(x(t), y(t))] \\
&\quad = \frac{1}{h}[f_x(x(t), y(t))\Delta x + f_y(x(t), y(t))\Delta y + \epsilon \Delta x + \eta \Delta y] \\
&\quad = f_x(x(t), y(t)) \cdot \frac{\Delta x}{h} + f_y(x(t), y(t)) \cdot \frac{\Delta y}{h} + \epsilon \cdot \frac{\Delta x}{h} + \eta \cdot \frac{\Delta y}{h}.
\end{aligned}$$

Thus

$$\begin{aligned}
\frac{dz}{dt} &= \lim_{h \to 0} \frac{f(x(t+h), y(t+h)) - f(x(t), y(t))}{h} \\
&= f_x(x(t), y(t)) \cdot \lim_{h \to 0} \frac{x(t+h) - x(t)}{h} \\
&\quad + f_y(x(t), y(t)) \cdot \lim_{h \to 0} \frac{y(t+h) - y(t)}{h} \\
&\quad + \lim_{h \to 0} \epsilon \cdot \lim_{h \to 0} \frac{x(t+h) - x(t)}{h} + \lim_{h \to 0} \eta \cdot \lim_{h \to 0} \frac{y(t+h) - y(t)}{h} \\
&= f_x(x(t), y(t)) \frac{dx}{dt} + f_y(x(t), y(t)) \frac{dy}{dt} + 0 \cdot \frac{dx}{dt} + 0 \cdot \frac{dy}{dt} \\
&= \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}
\end{aligned}$$

**Example 1**

If $z = x^2 + y^2$, $x = \sin t$ and $y = e^t$, find $dz/dt$.

$$\begin{aligned}
\frac{dz}{dt} &= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \\
&= 2x \cdot \cos t + 2y \cdot e^t \\
&= 2(x \cos t + ye^t).
\end{aligned}$$

By substituting $x = \sin t$ and $y = e^t$, we can express the result entirely in terms of $t$.

$$\frac{dz}{dt} = 2(\sin t \cos t + e^{2t}).$$

Of course, this derivative could also be found by substituting first and then differentiating.

$$\begin{aligned} z &= x^2 + y^2 \\ &= \sin^2 t + e^{2t}; \\ \frac{dz}{dt} &= 2 \sin t \cos t + 2e^{2t}. \end{aligned}$$

Theorem 21.2 is especially useful when we do not know what all of the functions are. It can be extended in many ways. Perhaps the most obvious is the case in which $z$ is a function of three or more variables, each of which is a function of $t$.

### *Theorem 21.3*

*If* $z = f(x_1, x_2, \ldots, x_n)$ *and* $x_i = g_i(t)$, $i = 1, 2, \ldots, n$, *and if all partial derivatives of* $z$ *are continuous and* $dx_i/dt$ *exist,* $i = 1, 2, \ldots, n$, *then*

$$\frac{dz}{dt} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{dx_i}{dt}.$$

A special case of Theorem 21.2 follows.

### *Theorem 21.4*

*If* $z = f(x, y)$ *and* $y = g(x)$ *and if* $\partial z/\partial x$ *and* $\partial z/\partial y$ *are continuous and* $dz/dx$ *exists, then*

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

### *Example 2*

If $z = x^2 + xy + y^2$ and $y = \sin x$, find $dz/dx$.

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dx} = (2x + y) + (x + 2y) \cos x.$$

Again, this expression can be put entirely in terms of $x$, or it could be found by substituting $y = \sin x$ first and differentiating in the ordinary way.

Note the distinction between $\partial z/\partial x$ and $dz/dx$. $\partial z/\partial x$ is determined by the original function $z = f(x, y)$, where $x$ and $y$ are assumed to be two independent variables. The fact that $y = g(x)$ puts an additional restriction upon $x$ and $y$ does not enter into consideration when one is finding $\partial z/\partial x$ or $\partial z/\partial y$. This restriction *is* taken into account when finding $dz/dx$.

Just as Theorem 21.2, in which $z$ is a function of two variables, can be extended to give Theorem 21.3, in which $z$ is a function of $n$ variables, it can also be extended from the case in which $x$ and $y$ are functions of a single variable $t$ to the case in which $x$ and $y$ are functions of $m$ variables. Let us consider one special case here.

***Theorem 21.5***

*If* $z = f(x, y)$, $x = F(u, v)$, *and* $y = G(u, v)$ *and if* $\partial z/\partial x$ *and* $\partial z/\partial y$ *are continuous and* $\partial x/\partial u$, $\partial x/\partial v$, $\partial y/\partial u$, *and* $\partial y/\partial v$ *exist, then*

$$\boxed{\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial u}}$$

*and*

$$\boxed{\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial v}.}$$

The proof of this theorem is analogous to the proof of Theorem 21.2.

***Example 3***

Given $z = x^2 - y^3$, $x = u + v$, and $y = u - v$, find $\partial z/\partial u$ and $\partial z/\partial v$.

$$\begin{aligned}\frac{\partial z}{\partial u} &= \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial u}\\ &= 2x\cdot 1 - 3y^2\cdot 1\\ &= 2x - 3y^2.\end{aligned}$$

$$\begin{aligned}\frac{\partial z}{\partial v} &= \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial v}\\ &= 2x\cdot 1 - 3y^2\,(-1)\\ &= 2x + 3y^2.\end{aligned}$$

A question that arises is, "How do we know when to use a partial derivative and when to use an ordinary derivative?" It is simply a matter of noting whether the function in question is a function of one variable or more than one. For instance, in Theorem 21.5, $z$ is a function of the two variables $x$ and $y$. Thus we want the partial derivatives

$$\frac{\partial z}{\partial x} \quad\text{and}\quad \frac{\partial z}{\partial y}.$$

Again, both $x$ and $y$ are functions of the two variables $u$ and $v$. We again want partial derivatives

$$\frac{\partial x}{\partial u},\quad \frac{\partial x}{\partial v},\quad \frac{\partial y}{\partial u},\quad \frac{\partial y}{\partial v}.$$

Finally, if the $x$ and $y$ in $z = f(x, y)$ are replaced by $F(u, v)$ and $G(u, v)$, we have $z = f(F(u, v), G(u, v))$, which is still a function of two variables. Thus, we still want partial derivatives,

$$\frac{\partial z}{\partial u} \quad\text{and}\quad \frac{\partial z}{\partial v}.$$

The chain rule is useful in rate-of-change problems.

**Example 4**

Suppose the radius of a right circular cylinder is increasing at the rate of 2 in./min and the height is decreasing at 4 in./min. At what rate is the volume changing at the moment when the radius is 4 in. and the height is 10 in.?

$V = \pi r^2 h$. Since both $r$ and $h$ are functions of the time $t$ we use the chain rule to differentiate.

$$\frac{dV}{dt} = \frac{\partial V}{\partial r}\frac{dr}{dt} + \frac{\partial V}{\partial h}\frac{dh}{dt}$$

$$= 2\pi rh\frac{dr}{dt} + \pi r^2\frac{dh}{dt}.$$

But $dr/dt = 2$, $dh/dt = -4$, $r = 4$, and $h = 10$. Therefore

$$\frac{dV}{dt} = 2\pi(4)(10)(2) + \pi(4)^2(-4)$$

$$= 96\pi \text{ in.}^3/\text{min}.$$

## Problems

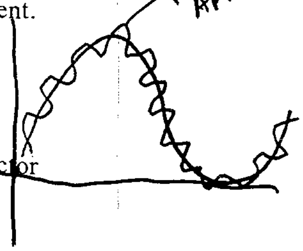
### A

In Problems 1–24, use the theorems of this section to find the required derivatives.

1. $z = x^2y$, $x = t^2 + 1$, $y = e^t$; find $dz/dt$.
2. $z = xy$, $x = \sin t$, $y = 2t + 1$; find $dz/dt$.
3. $z = x^3 - y$, $x = te^t$, $y = \sin t$; find $dz/dt$.
4. $z = x/y$, $x = t \sin t$, $y = \cos t$; find $dz/dt$.
5. $w = \dfrac{x^2 + y^2}{z}$, $x = t - 2$, $y = t^2 - 1$, $z = t^2 + 1$; find $dw/dt$.
6. $w = x^2 + y^2 + z^2$, $x = t - 1$, $y = t$, $z = t + 1$; find $dw/dt$.
7. $w = z(x^2 + y)$, $x = e^t$, $y = e^{-t}$, $z = t$; find $dw/dt$.
8. $w = x^2y + y^2z + z^2x$, $x = \sin t$, $y = \cos t$, $z = \tan t$; find $dw/dt$.
9. $z = x/y$, $x = u \sin v$, $y = v \cos u$; find $\partial z/\partial u$ and $\partial z/\partial v$.
10. $z = x^2 + y^2$, $x = u^2 - v^2$, $y = u^2 + v^2$; find $\partial z/\partial u$ and $\partial z/\partial v$.
11. $z = x^2y$, $x = 2u + v$, $y = 2v - u$; find $\partial z/\partial u$ and $\partial z/\partial v$.
12. $z = xy$, $x = u/v$, $y = u^2 + v^2$; find $\partial z/\partial u$ and $\partial z/\partial v$.
13. $w = xyz$, $x = u + v$, $y = u - v$, $z = 2u + 3v$; find $\partial w/\partial u$ and $\partial w/\partial v$.
14. $w = x^2 + y^2 + z^2$, $x = u^2 - v^2$, $y = u^2 + v^2$, $z = uv$; find $\partial w/\partial u$ and $\partial w/\partial v$.
15. $w = xy + yz + zx$, $x = u + v$, $y = 2u + v$, $z = u - 2v$; find $\partial w/\partial u$ and $\partial w/\partial v$.
16. $w = 3x + 2y - z$, $x = u \sin v$, $y = v \sin u$, $z = \sin u \sin v$; find $\partial w/\partial u$ and $\partial w/\partial v$.
17. $w = 2x + y - z$, $x = t^2 + u^2$, $y = u^2 + v^2$, $z = v^2 + t^2$; find $\partial w/\partial t$, $\partial w/\partial u$, and $\partial w/\partial v$.
18. $w = xyz$, $x = tuv$, $y = t/u$, $z = u/v$; find $\partial w/\partial t$, $\partial w/\partial u$, and $\partial w/\partial v$.
19. $v = 2x + y - z + 3w$, $x = t \sin t$, $y = \dfrac{\sin t}{t}$, $z = t \cos t$, $w = \dfrac{\cos t}{t}$; find $dv/dt$.
20. $v = xy + zw$, $x = st$, $y = s + t$, $z = s/t$, $w = s - t$; find $\partial v/\partial s$ and $\partial v/\partial t$.
21. $z = xy$, $y = e^x \sin x$; find $dz/dx$.
22. $z = x^2 + y^2$, $y = x \sin x$; find $dz/dx$.

# 19.5 Method of Least Squares

POLYNOMIAL APPROX



In $\boxed{\textbf{curve fitting}}$ we are given $n$ points (pairs of numbers)

$$(x_1, y_1), \cdots, (x_n, y_n)$$

and we want to determine a function $f(x)$ such that $f(x_j) \approx y_j$, FOR $j = 1, \cdots, n$. The type of function (for example, polynomials, exponential functions, sine and cosine functions) may be suggested by the nature of the problem (the underlying physical law, for instance), and in many cases a polynomial of a certain degree will be AS appropriate. APPROXIMATION

If we require strict equality $f(x_1) = y_1, \cdots, f(x_n) = y_n$ and use polynomials of sufficiently high degree, we may apply one of the methods discussed in Sec. 18.3 in connection with interpolation. However, in certain situations this would not be the appropriate solution of the actual problem. For instance, to the four points

$$(1) \qquad (-1.0, 1.000), \qquad (-0.1, 1.099), \qquad (0.2, 0.808), \qquad (1.0, 1.000)$$

there corresponds the Lagrange polynomial $f(x) = x^3 - x + 1$ (Fig. 432), but if we graph the points, we see that they lie nearly on a straight line. Hence if these values are obtained in an experiment and thus involve an experimental error, and if the nature of the experiment suggests a linear relation, we better fit a straight line through the points (Fig. 432). Such a line may be useful for predicting values to be expected for other values of $x$. In simple cases a straight line may be fitted by eye, but if the points are scattered, this becomes unreliable and we better use a mathematical principle. A widely used procedure of this type is the **method of least squares** by Gauss. In the present situation it may be formulated as follows.

CURVE FIT A STRAIGHT LINE!

**Method of least squares.** *The straight line*

$$y = b + mx$$

*should be fitted through the given points* $(x_1, y_1), \cdots, (x_n, y_n)$ *so that the*

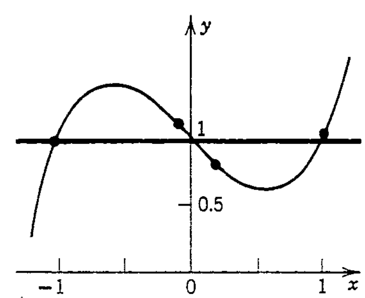



**Fig. 432.** The approximate fitting of a straight line

*sum of the squares of the distances of those points from the straight line is minimum, where the distance is measured in the vertical direction (the y-direction).*

The point on the line with abscissa $x_j$ has the ordinate $b + mx_j$. Hence its distance from $(x_j, y_j)$ is $|y_j - b - mx_j|$ (cf. Fig. 433) and that sum of squares is

$$\boxed{q = \sum_{j=1}^{n} (y_j - b - mx_j)^2.}$$

$q$ depends on $b$ and $m$. A necessary condition for $q$ to be minimum is

$$(2) \qquad \begin{array}{l} \dfrac{\Delta q}{\Delta b} \propto \\[2ex] \dfrac{\Delta q}{\Delta m} \propto \end{array} \boxed{\begin{aligned} \frac{\partial q}{\partial b} &= -2 \sum_{j=1}^{n} (y_j - b - mx_j) = 0 \\ \frac{\partial q}{\partial m} &= -2 \sum_{j=1}^{n} x_j(y_j - b - mx_j) = 0 \end{aligned}}$$

(where we sum over $j$ from 1 to $n$). Writing each sum as three sums and taking one of them to the right, we obtain the result

$$(3) \qquad \boxed{\begin{aligned} bn \quad + m \sum x_j &= \sum y_j \\ b \sum x_j + b \sum x_j^2 &= \sum x_j y_j. \end{aligned}}$$

These equations are called the **normal equations** of our problem.

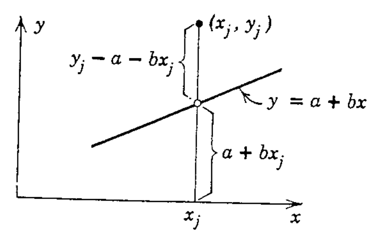



**Fig. 433.** Vertical distance of a point $(x_j, y_j)$ from a straight line $y = a + bx$

**EXAMPLE 1. Straight line**

Using the method of least squares, fit a straight line to the four points given in formula (1).

***Solution.*** We obtain

$$n = 4, \quad \sum x_j = 0.1, \quad \sum x_j^2 = 2.05, \quad \sum y_j = 3.907, \quad \sum x_j y_j = 0.0517.$$

Hence the normal equations are

$$\begin{aligned} 4a + 0.10b &= 3.9070 \\ 0.1a + 2.05b &= 0.0517 \end{aligned}$$

The solution is $a = 0.9773$, $b = -0.0224$, and we obtain the straight line (Fig. 432)

$$y = 0.9773 - 0.0224x. \quad \blacksquare$$

Our approach of curve fitting can be generalized from a polynomial $y = a + bx$ to a polynomial degree $m$

$$p(x) = b_0 + b_1 x + \cdots + b_m x^m$$

where $m \leqq n - 1$. Then $q$ takes the form

$$q = \sum_{j=1}^{n} (y_j - p(x_j))^2$$

and depends on $m + 1$ parameters $b_0, \cdots, b_m$. Instead of (2) we then have $m + 1$ conditions

$$\frac{\partial q}{\partial b_0} = 0, \qquad \cdots, \qquad \frac{\partial q}{\partial b_m} = 0$$

which give a system of $m + 1$ normal equations. The reader may show that in the case of a quadratic polynomial

$$p(x) = b_0 + b_1 x + b_2 x^2 \tag{4}$$

the normal equations are (summation from 1 to $n$)

$$\begin{aligned}
b_0 n \quad &+ b_1 \sum x_j + b_2 \sum x_j^2 = \sum y_j \\
b_0 \sum x_j &+ b_1 \sum x_j^2 + b_2 \sum x_j^3 = \sum x_j y_j \\
b_0 \sum x_j^2 &+ b_1 \sum x_j^3 + b_2 \sum x_j^4 = \sum x_j^2 y_j.
\end{aligned} \tag{5}$$

Note that this system is symmetric. To solve it for the unknowns $b_0$, $b_1$, $b_2$, we may apply one of the methods discussed in Secs. 19.1–19.3.

Beginning in the next section, we shall turn to **matrix eigenvalue problems,** which are of great general interest in engineering and many other fields.

## Problems for Sec. 19.5

In each case plot the given points in the $xy$-plane and fit a straight line (*a*) by eye, (*b*) by the method of least squares.

**1.** (2, 5), (3, 9), (4, 15), (5, 21)

**2.** (0, 2.3), (2, 4.1), (4, 5.7), (6, 6.9)

**3.** (4, 3), (15, 16), (30, 13), (100, 70), (200, 90)

**4.** (5, 5.0), (10, 3.9), (15, 3.2), (20, 2.0)

**5.**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Density of ore $x$ [grams/cm$^3$] | 2.8 | 2.9 | 3.0 | 3.1 | 3.2 | 3.2 | 3.2 | 3.3 | 3.4 |
| Iron content $y$ [percent] | 32 | 28 | 35 | 33 | 35 | 37 | 39 | 38 | 35 |

**6.**

| | | | | | |
|---|---|---|---|---|---|
| Revolutions per minute $x$ | 400 | 500 | 600 | 700 | 750 |
| Power of a Diesel engine $y$ [hp] | 600 | 1050 | 1440 | 1900 | 2120 |