UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
CADERNOS DE MATEMÁTICA E ESTATÍSTICA
SÉRIE B: TRABALHO DE APOIO DIDÁTICO

CÁLCULO APLICADO

FERNANDO AUGUSTO BOEIRA SABINO DA SILVA

SÉRIE B, Nº 75
PORTO ALEGRE, AGOSTO DE 2009

DECISION MODELS FOR BUSINESS

PART II- APPLIED CALCULUS

I Curves and Slopes – Average rates of change

1. The Slope of a Curve
   - Q1: Consider the curve $y = f(x) = x^2$ (draw a picture). What is its slope?
   - A1: Not at all clear! But for sure it's different at different points on the curve, since the curve gets steeper and steeper as x gets larger and larger.

2. The "Average" Slope of a Curve
   - Q2: Let's look at an easier question: What's the average rate of change in y as x changes from 1 to 2? Here we are looking for the slope of the <u>secant</u> of the curve joining the 2 points $(1, 1)$ and $(2, 4)$.
   - A2: We know how to do that! $m = \Delta y/\Delta x = (4 - 1)/(2 - 1) = 3$
   - Some additional notation: We are looking at the two points:
     $(x_1, y_1) = (x_1, f(x_1)) = (1, 1)$ and
     $(x_2, y_2) = (x_1 + \Delta x\,, y_1 + \Delta y) = (x_1 + \Delta x, f(x_1 + \Delta x)) = (x_1 + \Delta x, f(x_1) + \Delta f)$
     Put all these labels on the picture!
   - With this notation, $m = \Delta y/\Delta x = \Delta f/\Delta x$
   - We'll also write $x + h$ instead of Δx, Then the slope of the secant joining the two points $(x, f(x))$ and $(x + h, f(x + h))$ is $m = (f(x + h) - f(x))/h$
   - Q3: What's the average rate of change in y as x changes from 1 to 1.5?
   - A3: $m = (f(x + h) - f(x))/h = (2.25 - 1)/(1.5 - 1) = 2.50$

3. Moving the 2 points closer and closer to each other. Let's fix (x_1, y_1) at (1, 1) and look at the average rates of change of y for various points that get closer and closer to $(1, 1)$:

x_2	$y_2 = x_2^2$	Δx	Δy	$m = \Delta y/\Delta x$
2	4	1	3	3
1.5	2.25	0.5	1.25	2.5
1.1	1.21	0.1	0.21	2.1
1.01	1.0201	0.01	0.0201	2.01
1.001	1.002001	0.001	0.002001	2.001
0.5	0.25	–0.5	–0.75	1.5
0.9	0.81	–0.1	–0.19	1.9
0.99	0.9801	–0.01	–0.0199	1.99

0.999	**0.998001**	**–0.001**	**–0.001999**	**1.999**
$1 + \Delta x$	**$1 + 2\Delta x + (\Delta x)^2$**	**Δx**	**$2\Delta x + (\Delta x)^2$**	**$2 + \Delta x$**

In the last row of the table, we've done the computation symbolically (algebraically), rather than numerically (arithmetically).

4. Tangent Lines and Instantaneous Rates of Change – The Slope of a Curve at a Point

- **Geometrically, as $x_2 \to x_1$ ("as x_2 gets closer and closer to x_1" or "as x_2 approaches x_1" or "as $\Delta x \to 0$"), the secant line gets closer and closer to (approaches) the tangent line to the curve at the point (x_1, y_1).**
- **Thus it makes sense to define the slope of a curve at a point as the slope of its tangent line at that point, and this slope is also called the "instantaneous rate of change in y at that point."**
- **As we said initially, curves thus have different slopes at different points:**

At (1, 1) the slope of the tangent to $y = x^2$ is 2. (As $\Delta x \to 0$, $\Delta y/\Delta x = 2 + \Delta x \to 2$.)

At an arbitrary (x, y) on the curve $y = x^2$, we find that the slope of the secant is:

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

And letting $\Delta x \to 0$, we see that the slope of the curve is $2x$. Hence:

- **At (0, 0), the slope is 0 (does this make sense geometrically?)**
- **At (–1, 1), the slope is –2 (does this make sense geometrically?)**
- **At (2, 4), the slope is 4 (does this make sense geometrically?)**
- **At (–2, 4), the slope is –4 (does this make sense geometrically?)**

5. Examples from section 11.1 from the textbook (Waner/Costenoble)

14. Average rate of change of $f(x) = 2x^2 + 4$ on [–1, 2]: (12 – 6)/(2 – (–1)) = 2

16. Average rate of change of $f(x) = 1/x$ on [1, 4]: (1/4 – 1)/(4 – 1) = –1/4

22. Average rate of change of $f(x) = 2/x$ on [1, 1 +h]. Note that f(1) = 2.

h	f(1 + h) = 2/(1 + h)	Δf	m = Δf/h
1	1	–1	–1
0.1	1.818182	–0.191918	–1.9192
0.01	1.980198	–0.019802	–1.9802
0.001	1.998002	–0.001988	–1.988
0.0001	1.999800	–0.000200	–2.00
		–2h/(1 + h)	–2/(1 + h)

(So, as $h \to 0$, $m \to -2$.)

24 Average rate of change of $f(x) = 3x^2 - 2x$ on [0, 0 +h]. Note that f(0) = 0.

h	$f(h) = 3h^2 - 2h$	Δf	m = Δf/h
1	1	1	1
0.1	–0.17	–0.17	–1.7
0.01	–0.0197	–0.0197	–1.97
0.001	–0.001997	–0.001997	–1.997
0.0001	–0.00019997	–0.00019997	–1.9997
		$3h^2 - 2h$	3h – 2

(So, as $h \to 0$, $m \to -2$.)

47. $R(t) = 95t^2 + 115t + 150$ (millions of $)

year	1997	1998	1999
t	0	1	2
Revenue ($$10^6$)	150	360	760

a. Average rate of change of R from 1997 to 1999: (760 – 150)/(2 – 0) = $305 million/year
b. From 1997 to 1999 revenues
   (A) increased at an increasing rate
   (B) increased at a decreasing rate
   (C) decreased at an increasing rate
   (D) decreased at a decreasing rate
c. Predict average rate of change from 1999 to 2000:
   [R(3) – R(2)]/(3 – 2) = (1350 – 760)/1 = $590 million/year

48. $P(t) = -0.15t^2 + 0.50t + 130$ (thousands of Roman catholic nuns)

year	1975	1985	1995
t	0	10	20
Population (10^3)	130	120	80

a. Average rate of change of P from 1975 to 1995: (80 – 130)/(20 – 0) = –2.5 thousand/year
b. From 1975 to 1995 the number of nuns
   (A) increased at an increasing rate
   (B) increased at a decreasing rate
   (C) decreased at an increasing rate
   (D) decreased at a decreasing rate
c. Predict average rate of change from 1995 to 2005:
   [P(30) – P(20)]/(30 – 20) = (10 – 80)/10 = –7 thousand/year

II Curves and Slopes – Instantaneous rates of change (derivatives)

1. The Slope of the Curve $y = f(x)$
   - Can be different at different points on the curve!
   - Average slope:
     - slope of a secant to the curve
     - $m = \Delta y/\Delta x = \Delta f/\Delta x$
   - Instantaneous slope:
     - slope of a tangent to the curve
     - "limiting value" of $\Delta y/\Delta x$ as $\Delta x \to 0$

2. We did numeric and algebraic examples that all worked nicely. Before we continue formally, let's see what can go wrong.

   Example 1 – what's the instantaneous slope of $y = |x|$ at $x = 0$?
   - A look at the graph shows what the problem is immediately. To the left (right) of (0, 0), the graph is a straight line with slope –1 (+1):
     - For $x > 0$, $|x| = x$, so with $h > 0$, we have $[f(0 + h) – f(0)]/h = h/h = 1$, so the "slope to the right" of $|x|$ at $x = 0$ is +1.
     - For $x < 0$, $|x| = –x$, so with $h < 0$, we have $[f(0 + h) – f(0)]/h = –h/h = –1$, so the "slope to the left" of $|x|$ at $x = 0$ is –1.
     - However, to say that a curve has a particular slope at a given point means that the limiting value is the same regardless of how we approach that point. The "slope to the right" and the "slope to the left" must have the same value!
     - Geometrically what goes wrong in this example is that the curve comes to a point ("has a kink") at the origin. (It isn't "smooth.") In fact, the curve has lots of tangent lines at (0, 0).
     - So $y = |x|$ doesn't have a well defined instantaneous slope at (0, 0)

   Example 2 – what's the instantaneous slope of $y = x^{1/3}$ at $x = 0$?
   - Let's do this one numerically. Note that $f(0) = 0$.

x	–1	–0.1	–0.01	–0.001	0	0.001	0.01	0.1	1
$F(x) = x^{1/3}$	–1	–0.4642	–0.2154	–0.1	0	0.1	0.2154	0.4642	1
$[f(x) – f(0)]/(x – 0)$	1	4.642	21.54	100	–	100	21.54	4.642	1

   - The average rates of change seem to be growing without bound and not reaching a limiting value.
   - Once again, this curve doesn't have a well defined instantaneous slope at (0,0) (Even though the graph is smooth!). The problem here is that the "tangent" line is vertical.

So in order for the slope to be well defined at a point, the curve has to be smooth and can't have a vertical tangent.

3. Some more nice examples.

Example 3 – what's the instantaneous slope of $y = 5x^3$ at the generic point (x, y)?

$$\begin{aligned}\Delta y/\Delta x = \Delta f/\Delta x &= (5(x+h)^3 - 5x^3)/[(x+h) - x] \\ &= (5x^3 + 15x^2h + 15xh^2 + 5h^3 - 5x^3)/h \\ &= (15x^2h + 15xh^2 + 5h^3)/h \\ &= 15x^2 + 15xh + 5h^2 \quad \underline{\text{when } h \neq 0}\end{aligned}$$

and this goes to $15x^2$ as $h \to 0$.

Note that we just can't plug in $h = 0$ at the beginning, because then we'd be dividing by 0, which we know we can't do. Instead we do some algebra to get rid of the h in the denominator, and then let $h \to 0$. Sometimes that algebra can be very messy!

Example 4 – what's the instantaneous slope of $y = \sqrt{x}$ at the generic point (x, y), with $y \geq 0$? Note that in this case the function isn't defined for $y < 0$, so it makes no sense to ask for its slope except when $y \geq 0$.

$$\begin{aligned}\Delta y/\Delta x = \Delta f/\Delta x &= (\sqrt{(x+h)} - \sqrt{x})/h \\ &= \left(\frac{\sqrt{x+h} - \sqrt{x}}{h}\right)\left(\frac{\sqrt{x+h} + \sqrt{x}}{\sqrt{x+h} + \sqrt{x}}\right) \\ &= \frac{x+h-x}{h(\sqrt{x+h} + \sqrt{x})} \\ &= \frac{1}{\sqrt{x+h} + \sqrt{x}} \quad \underline{\text{when } h \neq 0}\end{aligned}$$

and this goes to $1/2\sqrt{x}$ as $h \to 0$.

Note that this is not defined at $x = 0$ (and in fact the curve is vertical at that point!) Hence the slope of $y = \sqrt{x}$ is $1/(2\sqrt{x})$ at all $x > 0$.

4. Continuous functions:

In the last line of the proof, I used the fact that $\sqrt{x+h} \to \sqrt{x}$ as $h \to 0$. When a function satisfies the condition that $f(a + h) \to f(a)$ as $h \to 0$ (or equivalently, that $f(x) \to f(a)$ as $x \to a$), we say that the function is <u>continuous at $x = a$</u>. Loosely, a function is continuous if you can draw its graph without lifting your pen from the paper. Most of the functions we encounter in this class are continuous.

Examples:

- $f(x) = |x|$ is continuous on $(-\infty, +\infty)$ – even though it has a kink at 0!

- $f(x) = 1/x$ is continuous on $(0, +\infty)$ (and on $(-\infty, 0)$) – it's not continuous at $x = 0$, because $f(0)$ is undefined
- $f(x) = \sqrt{x}$ is continuous on $[0, +\infty)$
- $f(x) = (x^2 - 5x + 6)/(x - 3)$ is not continuous at $x = 3$, because even though $f(x) \to 1$ as $x \to 3$ (because $f(x) = x - 2$ when $x \neq 3$), $f(3)$ is not defined

5. **Derivatives and some notation**

Recall that the instantaneous rate of change of a function $f(x)$ at the point $(x, f(x))$ is defined as $\lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}\right)$ if that limit exists. The instantaneous rate of return is called the "derivative of $f(x)$" and is denoted $f'(x)$ (or y'). In our examples, we've seen that if

- $f(x) = 5x^3$, then $f'(x) = 15x^2$
- $f(x) = \sqrt{x}$, then $f'(x) = 1/(2\sqrt{x})$
- $f(x) = x^2$, then $f'(x) = 2x$

If a function has a derivative at a point (is "differentiable" at that point), then necessarily it is continuous there (see the definition – the numerator has to $\to 0$ as $h \to 0$, or else the limit can't exist. However, as $f(x) = |x|$ demonstrates at $x = 0$, continuity alone isn't enough for the derivative to exist (so "differentiable" is a stronger property than "continuous").

Another notation for the derivative:

$$\lim_{\Delta x\to 0}\left(\frac{f(x+\Delta x)-f(x)}{\Delta x}\right)=\lim_{\Delta x\to 0}\frac{\Delta f}{\Delta x}=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}\equiv\frac{df}{dx}\text{ or }\frac{dy}{dx}$$

df/dx or dy/dx is <u>not</u> a fraction (although we shall see later that it often behaves like one!)

Clearly using the definition to find the formulas for a derivative whenever we need one is tedious. Next time we'll look at some general rules for finding derivatives.

III Formulas for derivatives: x^n, $f(x) \pm g(x)$, $cf(x)$

The process of taking the derivative of a function is called *differentiation* or *differentiating the function*. We've done a whole bunch of examples already. Today we want to find differentiation rules, so we don't always have to fall back on the often tedious task of using the limit definition of the derivative.

1. **If $f(x) = c$, where c is any *constant*, clearly changing x has no effect whatsoever on $f(x)$. Hence:**

   Differentiation Rule 1: $(d/dx)c = 0$, for any constant c

2. **We have already seen that the derivatives of x $(= x^1)$, x^2, and x^3 are 1 $(= 1x^0)$, $2x$, and $3x^2$. Do you see a pattern here?**

   Note also that the derivative of $f(x) = 1$ $(=1x^0) = 0$, since 1 is a constant. Does this fit the pattern?

   We also saw that the derivative of $f(x) = \sqrt{x}$ $(= x^{1/2}) = 1/(2\sqrt{x}) = (1/2)x^{-1/2}$. Does this fit the pattern? Of course this holds only for $x > 0$. For $x < 0$, $f(x)$ doesn't exist, so neither does $f'(x)$. And we saw that the slope "blew up" at $x = 0$, which is exactly what our formula $1/(2\sqrt{x})$ does at $x = 0$.

   How about $f(x) = 1/x$ $(= x^{-1})$? Let's see:
   $[f(x + h) - f(x)]/h = [1/(x + h) - 1/x]/h = -h/[hx(x + h)] = -1/[x(x + h)]$ if $h \neq 0$
   So now letting $h \to 0$, we see that the derivative of $x^{-1} = -1/x^2$ $(= (-1)x^{-2})$. Does this fit the pattern? In fact, this result holds for any *constant* power of x, positive or negative, integer or fractional, rational or irrational!

   Differentiation rule 2: $(d/dx)x^n = f'(x^n) = nx^{n-1}$, for any *constant* n

   A warning: the derivative of x^x is *not* $x(x^{x-1}) = x^x$, because here the power to which x is raised isn't a constant – it depends on x. In a few weeks we'll learn how to differentiate functions like x^x.

3. **If $f(x)$ and $g(x)$ are differentiable functions with derivatives $f'(x)$ and $g'(x)$, then $f(x) + g(x)$ and $f(x) - g(x)$ are also differentiable. What are their derivatives?**

$$\lim_{h\to 0}\left(\frac{[f(x+h)\pm g(x+h)]-[f(x)\pm g(x)]}{h}\right)=\lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}\pm\frac{g(x+h)-g(x)}{h}\right)$$

$$=\lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}\right)\pm\lim_{h\to 0}\left(\frac{g(x+h)-g(x)}{h}\right)=f'(x)\pm g'(x)$$

So the derivative of the sum (or difference) of two differentiable functions is the sum or difference) of their derivatives.

Differentiation rule 3: $(d/dx)(f(x) \pm g(x)) = (f + g)'(x) = f'(x) \pm g'(x)$

4. If $f(x)$ is a differentiable function and c is any *constant*, then

Differentiation rule 4: $(d/dx)(cf(x) = (cf)'(x) = cf'(x)$, for any constant c

This one is pretty obvious, because it's really just talking about changing units.

- If $f(t)$ is the total revenue in dollars earned by a new product in the first t days after it is introduced, then $f'(t)$ is the marginal revenue in dollars/day at time t. $100f(t)$ is the total revenue in cents, and clearly $100f'(t)$ is the marginal revenue in cents/day.
- If $f(t)$ is my distance (in miles) from home at time t hours after I start a trip, then $f'(t)$ is my speed in miles/hour at time t. Then $1.6f(t)$ is my distance from home in kilometers after t hours, and its derivative, $1.6f'(t)$, is my speed in kilometers per hour

5. Examples:
- $(d/dx)(3x^4 - 5/\sqrt{x} + x^{\sqrt{2}}) = 12x^3 - 5(-1/2)x^{-1/2} + (\sqrt{2})x^{\sqrt{2}-1}$
- $(d/dx)[(x^2 - 3/\sqrt{x})(2x + \sqrt{x})] = (d/dx)(2x^3 + x^{2.5} - 6x^{0.5} - 3) = 6x^2 + 2.5x^{1.5} - 3x^{0.5}$
- 11.4.67: Find all values of x where the tangent to the graph of $y = 2x^2 + 3x - 1$ is horizontal.
  A: The tangent is horizontal if its slope is 0. $y' = 4x + 3 = 0$ if $x = -3/4$
- 11.4.69: Ditto for $y = 2x + 8$.
  A: The tangent is never horizontal, because $y' = 2$ for all values of x
- 11.4.77: $P(t) = -2.6t^2 + 13t + 19$ is approximately the percent of people in the U.S. who have ever purchased anything on-line t years after January 2000. $P'(t) = -5.2t + 13$ and $P'(2) = 2.6$, which tells us that the instantaneous rate of *increase* in the percent of people who have made on-line purchases was 2.6%/year in January of 2002.

IV Marginal Analysis

As we have seen, the derivative, $f'(x)$, of a function, $f(x)$, tells us the instantaneous rate of change of the function – at the value x.

- **Geometrically, the derivative is the slope of the tangent line to the curve $y = f(x)$ at the point $(x, y) = (x, f(x))$.**
- **What is the equation of the tangent line at the particular point $(x_1, y_1) = (x_1, f(x_1))$? Well, we can just use the point-slope form of the equation for a straight line**
  ($m = (y - y_1)/x - x_1)$) to get: $y = f(x_1) + (x - x_1)f'(x_1) = y_1 + (x - x_1)y_1'$

Examples:

- **11.4.62. Tangent to $y = x^2$ at $(0, 0)$.**
  $f'(x) = 2x$, so $f'(0) = 0$, and the tangent line is $y = 0 + (x - 0)(0) = 0$
- **11.4.63. Tangent to $y = x + 1/x$ at $(2, 2.5)$.**
  $f'(x) = 1 - 1/x^2$, so $f'(2) = ¾$, and the tangent line is $y = 2.5 + 0.75(x - 2) = 1 + 0.75x$

Economists call f' "marginal f." Thus if $C(x)$ (respectively, $R(x)$, $P(x)$) is the (total) cost of (respectively revenue from, profit from) making x units of some product, then:

- **$C'(x)$ is the marginal cost per unit.**
  - **It is <u>approximately</u> the additional cost incurred by making one more unit of the product.**
  - **The <u>exact</u> additional cost is $C(x + 1) - C(x)$.**
  - **The marginal cost would be exactly the additional cost if the tangent line and the cost curve were the same.**
  - **However, since the tangent line is close to the curve near the point of tangency, the marginal cost is approximately the additional cost of another unit.**
- **$R'(x)$ is the marginal revenue per unit. Remarks similar to those for $C'(x)$ also apply here.**
- **$P'(x)$ is the marginal profit per unit. Once again similar remarks hold.**

Economists also talk about "average" cost per unit, $\overline{C}(x) = C(x)/x$ (and average revenue and average profit). Average cost is the cost per unit of all x units made, but marginal cost is (roughly) the cost of the last unit made (or of the next one to be made).

Now profit is related to cost and revenue by the formula $P(x) = R(x) - C(x)$. Accordingly, since the derivative of a difference is the difference of the derivatives, it follows that $P'(x) = R'(x) - C'(x)$, that is, marginal profit = marginal revenue – marginal cost. A similar relationship holds between average profit, average revenue, and average cost.

Examples:

- 11.5.7. (Show slide - or draw graph on board) a. B b. C c. C
- 11.5.19. Demand curve: $p = 20000/q^{1.5}$ $(200 \le q \le 800)$ with p = price in \$/lb and q = quantity sold in lbs/month.
  - a) What price leads to a demand of 400 lbs/month?
    $p = 20000/400^{1.5} = 20000/8000 = \$2.50/\text{lb}$
  - b) What is monthly revenue as a function of q?
    $R(q) = pq = q(20000/q^{1.5}) = 20000/\sqrt{q}$
  - c) What are revenue and marginal revenue at $q = 400$ lbs/month?
    $R(400) = 20000/\sqrt{400} = \$1000/\text{month}$ = monthly revenue when $p = \$2.50/\text{lb}$ (and $q = 400\text{lb/month}$)
    $R'(q) = (-0.5)(20000)/q^{1.5}$, so $R'(400) = -10000/8000 = -\$1.25/\text{lb}$. When $q = 400$ lbs/month, the marginal revenue is $-\$1.25/\text{lb}$
  - d) Increasing q decreases revenue according to (c). So to increase revenue, they should decrease q, which can be accomplished by increasing the price.

- 11.5.17. $C(x) = -0.001x^2 + 0.3x + 500$ (in \$), so
  - marginal cost = $C'(x) = -0.002x + 0.3$ (in \$/unit)
  - average cost = $\overline{C}(x) = C(x)/x = -0.001x + 0.3 + 500/x$ (in \$/unit)

  a) As x increases, $C'(x)$ decreases
  b) As x increases, $\overline{C}(x)$ decreases
  c) At $x = 100$,
  $C'(x) = -0.002(100) + 0.3 = 0.1$
  $\overline{C}(x) = -0.001(100) + 0.3 + 500/100 = 5.2$
  Hence, marginal cost < average cost at $x = 100$ units/day.

- 11.5.29. Marginal product per senior professor (i.e., additional profit in \$ from adding another senior professor) is 50% higher than marginal product per junior professor. Junior professors are paid half what senior professors are paid. University's salary budget is fixed. To maximize profit, how should the university adjust its staffing?
  Answer: If the marginal products apply only at the current staffing levels, they should reduce the number of senior professors and replace each one terminated with 2 junior professors. This will leave salary essentially unchanged, but total productivity will go up by half a junior professor's productivity. If the given relationships hold for all staffing levels, they should replace each and every senior professor by two junior professors.

Let's look ahead a bit. Suppose we want to "optimize" (maximize or minimize) some differentiable function. Clearly, the tangent to the curve will be horizontal at a maximum or minimum point on the graph. The tangent is horizontal if and only if its slope (which, remember is the derivative of the function) equals 0. Such points are called "stationary points." Although maxima and minima of differentiable functions must be stationary points, the converse is not true:

- Consider $f(x) = x^3$ at the origin. Since $f(x) > 0$ if $x > 0$, and $f(x) < 0$ if $x < 0$, clearly $(0,0)$ is neither a maximum nor minimum point of the graph. However, $f'(x) = 3x^2$, so $f'(0) = 0$ and $(0, 0)$ is a stationary point! (it is called a "point of inflection.")

So one way to find maxima and minima is to take derivatives and set them to 0. For now, looking at the graph will enable us to distinguish between maxima, minima, and inflection points. Later we'll see how to do this "analytically," without drawing graphs.

V- Derivatives of Products, Reciprocals, and Quotients

We have already seen several differentiation rules:

1. If c is any constant, then $dc/dx = 0$
2. If n is any constant, then $dx^n/dx = nx^{n-1}$
3. If $f(x)$ and $g(x)$ are differentiable functions, then so are $f(x) \pm g(x)$ and
$$(d/dx)(f(x) \pm g(x)) = (f + g)'(x) = f'(x) \pm g'(x)$$
4. If $f(x)$ is a differentiable function and c is any *constant*, then
$$(d/dx)(cf(x) = (cf)'(x) = cf'(x), \text{ for any constant } c$$

A. Products – how do we differentiate products like $f(x)g(x)$?

Example 1: $h(x) = (2x + 3)(x - 7)$?

One way of course is to multiply out to get $h(x) = 2x^2 - 11x - 21$, and just use rules 3 and 4 to get $h'(x) = dh/dx = 4x - 11$.

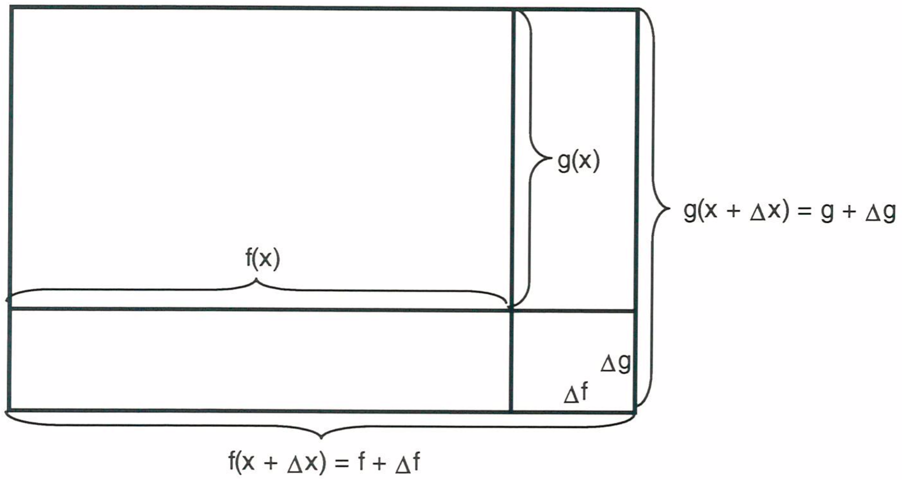
Example 2: $(x^3 - 2x^2 + 7 - 3/x)(x^2 - 6\sqrt{x} + 3x + 4/\sqrt{x})$

Multiplying out would be exceedingly tedious. Is there a formula for $(d/dx)(f(x)g(x))$ in terms of the functions and their derivatives? Looking at rules (3) and (4), you might surmise that

$$(d/dx)(f(x)g(x)) = f'(x)g'(x)$$

but you'd be WRONG! Returning to example 1, $f(x) = 2x + 3$ and $g(x) = x - 7$, so $f'(x)g'(x) = 2x1 = 2$, but $(d/dx)(f(x)g(x)) = 4x - 11$.

So what is the RIGHT way? Consider the picture below:



Think of the product f(x)g(x) as the area of a rectangle. We want to know how that area changes as x changes.

$$\frac{f(x+\Delta x)g(x+\Delta x)-f(x)g(x)}{\Delta x}=\frac{g\Delta f+f\Delta g+\Delta f\Delta g}{\Delta x}=g\frac{\Delta f}{\Delta x}+f\frac{\Delta g}{\Delta x}+\frac{\Delta f\Delta g}{\Delta x}$$

So now we let $\Delta x \to 0$. The first term goes to $f'g$ and the second to fg', but what about the third? It goes to 0, because the numerator – being the product of two small terms goes to 0 faster than the denominator does! This gives us differentiation rule 5:

5. If f(x) and g(x) are differentiable functions, then so is f(x)g(x) and
   $(d/dx)(f(x)g(x)) = (fg)'(x) = f'(x)g(x) + g'(x)f(x) = f'g + fg'$

B. Reciprocals –we want to find out how to differentiate a quotient, f(x)/g(x). A cute way to find out what that rule is applies the product rule twice
   - once to find the derivative of 1/g(x), the reciprocal of a function, and
   - again to find the derivative of f(x)[1/g(x)] = f(x)/g(x)

Notice that $1 = g(x)[1/g(x)]$, provided $g(x) \neq 0$ (in which case 1/g(x) isn't defined, so surely it doesn't have a derivative! So, let's use the product rule and differentiate both sides of that equation:

$$(d/dx)(1) = 0 = \frac{d}{dx}\left(g(x)\left(\frac{1}{g(x)}\right)\right) = g'(x)\left(\frac{1}{g(x)}\right) + \frac{d}{dx}\left(\frac{1}{g(x)}\right)g(x)$$

Now let's solve for $\frac{d}{dx}\left(\frac{1}{g(x)}\right)$: $\frac{d}{dx}\left(\frac{1}{g(x)}\right)g(x) = -g'(x)\left(\frac{1}{g(x)}\right)$, so

$$\frac{d}{dx}\left(\frac{1}{g(x)}\right) = \frac{-g'(x)}{(g(x))^2}$$

So now we have our next differentiation rule:

6. If g(x) is differentiable and $g(x) \neq 0$, then so is $1/g(x)$ differentiable, and

$$\frac{d}{dx}\left(\frac{1}{g(x)}\right) = \frac{-g'(x)}{(g(x))^2}$$

C. Quotients – OK, let's use rules 5 and 6 to see how to differentiate quotients (suppressing some of the x's):

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{d}{dx}\left(f(x)\left(\frac{1}{g(x)}\right)\right) = f'(x)\left(\frac{1}{g(x)}\right) + f(x)\frac{d}{dx}\left(\frac{1}{g(x)}\right) = \frac{f'(x)}{g(x)} - f(x)\left(\frac{g'(x)}{(g(x))^2}\right) = \frac{f'g - fg'}{g^2}$$

7. If f(x) and g(x) are differentiable and $g(x) \neq 0$, then so is $f(x)/g(x)$ differentiable, and

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2} = \frac{f'g - fg'}{g^2}$$

Examples:

- 12.1.27. $y = (x/3.2 + 3.2/x)(x^2 + 1)$, $y' = (1/3.2 - 3.2/x^2)(x^2 + 1) + (x/3.2 + 3.2/x)(2x)$
- 12.1.25. $y = (2x^{0.5} - x^2)^2 = (2x^{0.5} - x^2)(2x^{0.5} - x^2)$, $y' = 2(x^{-0.5} - 2x)(2x^{0.5} - x^2)$
- $y = 1/(x^3 - 3\sqrt{x})$, $y' = -(3x^2 - (3/2\sqrt{x}))/(x^3 - 3\sqrt{x})^2$
- 12.1.43. $y = (\sqrt{x} + 1)/(\sqrt{x} - 1)$, $y' = [(1/2\sqrt{x})(\sqrt{x} - 1) - (\sqrt{x} + 1)(1/2\sqrt{x})]/(\sqrt{x} - 1)^2 = -1/[\sqrt{x}(\sqrt{x} - 1)^2]$
- 12.1.63. x months after introducing a product, monthly sales are $S(x) = 20x - x^2$ hundred units at a price per unit of $p(x) = 1000 - x^2$ dollars, and monthly revenue is $R(x) = S(x)p(x)$. At $x = 5$, find and interpret S', p', and R'
  - $S' = 20 - 2x = 10$ at $x = 5$: after 5 months, sales are increasing by 1000 units/month
  - $p' = -2x = -10$ at $x = 5$: after 5 months, the price is decreasing by \$10/unit/month
  - $R' = S'p + Sp' = 10(975) + 75(-10) = 9000$ at $x = 5$: after 5 months, monthly revenue is increasing by \$900,000
- 12.1.67. t months into the year, a bus company has monthly costs of $C(t) = 10000 + t^2$ dollars and monthly ridership of $P(t) = 1000 + t^2$ passengers. After 6 months how fast is cost/passenger changing?
  - Cost per passenger = C(t)/P(t), so its derivative is

$$(C'P - CP')/P^2 = [2t(1000 + t^2) - (10000 + t^2)(2t)]/[1000 + t^2]^2 = -2t(9000)/[1000 + t^2]^2$$

Plugging in $t = 6$, we get $-108{,}000/1036^2 = -0.1006$, so it's decreasing by about 10¢ per month.

VI- Composite Functions – the "Chain Rule"

We saw last time how to use the product rule to find the derivative of $y = (2x^{0.5} - x^2)^2$. But what about $y = (2x^{0.5} - x^2)^{10}$? Multiplying that out would be exceedingly tedious! And how about $y = \sqrt{2x^{0.5} - x^2}$? That one we can't even multiply out.

What those two examples have in common is that they first compute the function $g(x) = 2x^{0.5} - x^2$, and then they apply some other function to : in the first example $f(g(x)) = [g(x)]^{10}$, and in the second $f(g(x)) = \sqrt{g(x)}$. Such functions of functions are known as "composite functions."

Examples:

- **$f(x) = x^2 + 2x - 3$, $g(x) = x^3$**
  $f(g(x)) = f(x^3) = (x^3)^2 + 2x^3 - 3$
  $g(f(x)) = g(x^2 + 2x - 3) = (x^2 + 2x - 3)^3 \neq f(g(x))$
  As this example shows, you have to pay attention to the order of composition
- **$f(t) = 1/(t + 3)$, $h(s) = \sqrt{(6/s)}$**

  $f(h(s)) = f(\sqrt{(6/s)}) = 1/(\sqrt{(6/s)} + 3)$, $h(f(t) = h(1/(t + 3)) = \sqrt{\dfrac{6}{1/(t+3)}} = \sqrt{6t+18}$
- **$q(r) = r^2 + 3$, $p(w) = (w - 3)^{0.5}$**
  $q(p(x)) = q((x - 3)^{0.5}) = ((x - 3)^{0.5})^2 + 3 = x$
  $p(q(x)) = p(x^2 + 3) = (r^2 + 3 - 3)^{0.5} = x$
  Functions that "undo" each other are called "inverse functions."

We'll only consider cases where both $f(x)$ and $g(x)$ are differentiable functions. Instead of writing $f(g(x))$, we'll simplify things by defining a new variable $u = g(x)$, and looking at $f(u)$.

We know that $f'(u) = df/du$ tells us how f changes when u changes. Loosely speaking, it's the number of units that f changes per unit change in u. But, since u depends on x, when we change x by one unit, u changes by $u' = du/dx$ units. If we put this all together, changing x by one unit changes u by u' units, and for each of those, f changes by f' units. Thus we have differentiation rule #8 – the "chain rule":

$$\frac{d}{dx}(f(u)) = \frac{d}{du}(f(u)) \times \frac{d}{dx}u(x)$$ **or, in a form that's much more intuitive,** $$\frac{df}{dx} = \frac{df}{du} \times \frac{du}{dx}.$$

Although df/du and du/dx are *emphatically NOT* fractions, they sure look like they behave like fractions in the chain rule! On occasion you may also see the chain rule written as

$$\frac{d}{dx}f(g(x)) = f'(g(x))g'(x)$$

That says the same thing, but is nowhere near so easy to remember as $\frac{df}{dx} = \frac{df}{du} \times \frac{du}{dx}$!

Examples:

- $f(x) = x^2 + 2x - 3$, $g(x) = x^3 = u$
  $(d/dx)f(u) = (df/du)(du/dx) = (2u + 2)(3x^2) = (2x^3 + 2)(3x^2) = 6x^5 + 6x^2$
  Let's check that directly: $f(g(x)) = (x^3)^2 + 2x^3 - 3 = x^6 + 2x^3 - 3$, and
  $(d/dx)(x^6 + 2x^3 - 3) = 6x^5 + 6x^2$
- $y = (2x^{0.5} - x^2)^{10}$, so $y' = 10(2x^{0.5} - x^2)^9(x^{-0.5} - 2x)$ [Here $f(u) = u^{10}$ and $u = 2x^{0.5} - x^2$]
- 12.2.15. $f(x) = (2x^2 - 2)^{-1}$, so $f'(x) = -(2x^2 - 2)^{-2}(4x)$.
  Check this via the reciprocal rule: $(d/dx)(1/g(x)) = -g'(x)/[g(x)]^2$
- 12.2.25. $f(x) = (1 - x^2)^{0.5}$, so $f'(x) = 0.5(1 - x^2)^{-0.5}(2x) = x(1 - x^2)^{-0.5}$
- 12.2.37. $f(z) = \left(\frac{z}{1+z^2}\right)^3$, so $f'(z) =$
  $$3\left(\frac{z}{1+z^2}\right)^2\left(\frac{(1+z^2)(1) - z(2z)}{(1+z^2)^2}\right) = 3\left(\frac{z}{1+z^2}\right)^2\left(\frac{1-z^2)}{(1+z^2)^2}\right)$$
- 12.2.55. Average commission per stock trade is $c(u) = 100u^2 - 160u + 110$ (\$/trade), where u = fraction of trades done online. In 1/1/98, the fraction of trades done online was given by $u(t) = 0.42 + 0.02t$, where t = months since 1/1/98. At what rate was the average commission per trade changing on 9/1/98?

  $dc/dt = (dc/du)(du/dt) = (200u - 160)(0.02)$. On 9/1/98, $u = 0.42 + 8(0.02) = 0.58$, so $dc/dt = (200(0.58) - 160)(0.02) = -0.88$, average commission per trade was decreasing by \$0.88/month
- 12.2.65. A circular oil slick's radius is growing at a rate of 2 miles/hour. How fast is the area of the slick changing when the radius is 3 miles?

  $A = \pi r^2$, so $dA/dt = (dA/dr)(dr/dt) = 2\pi r(2) = 12\pi$, i.e., increasing by about 37.7 square miles/hour when $r = 3$ miles.

Inverse functions – we saw that $q(r) = r^2 + 3$ and $p(w) = (w - 3)^{0.5}$ were inverse functions, in the sense that they "undid" each other – $p(q(x)) = x$ and $q(p(x)) = x$. We denote the inverse function of $f(x)$ by $f^{-1}(x)$. <u>NOTE: this is not the same as $[f(x)]^{-1} = 1/f(x)$!</u> How is the derivative of $f^{-1}(x)$ related to the derivative of $f(x)$? Let's use the chain rule to find out:

With $u = f^{-1}(x)$, we have $f(u) = x$, so $1 = f'(u)\frac{du}{dx} = f'(u)\frac{d}{dx}f^{-1}(x)$, and hence

$$\frac{d}{dx}f^{-1}(x) = \frac{1}{f'(u)} = \frac{1}{f'(f^{-1}(x))}$$

Another way to write this is $\frac{du}{dx} = \frac{1}{\frac{dx}{du}}$, so once again derivatives, which are not fractions, nonetheless behave like them!

Applying this to our example, with $f(x) = x^2 + 3$ and $f^{-1}(x) = (x - 3)^{0.5}$, we get $\frac{d}{dx}f(x) = 2x$ and $\frac{d}{dx}f^{-1}(x) = \frac{1}{f'(f^{-1}(x))} = \frac{1}{2(x-3)^{0.5}}$, which you can check by differentiating $(x - 3)^{0.5}$.

VII- Implicit Differentiation

Consider the graph of $x^2 + y^2 = 25$:

- It's a circle of radius 5, centered at the origin.
- What's the equation of the line tangent to the circle at $(x, y) = (3, 4)$? How about at a generic (x_0, y_0) on the circle?
- From elementary geometry, we know that the tangent to a circle at any point on the circle is **perpendicular to the radius** to that point.
- Recall that two lines, with slopes m_1 and m_2 are perpendicular when $m_1 m_2 = -1$.
- The slope of the radius to the point (x_0, y_0) is of course $m = y_0/x_0$.
- Hence the slope of the tangent at that point is $-x_0/y_0$, and the equation of the tangent is

$$\frac{y - y_0}{x - x_0} = -\frac{x_0}{y_0}$$

- So at $(3, 4)$, the tangent line is $(y - 4)/(x - 3) = -3/4$, or $y = -(3/4)x + 25/4$

But what about the slope and equation of the tangent to the curve $y^3 + 2y^2 - 10x = 6$ at $(1, 2)$? Here we don't have geometry to come to our rescue! We'll get back to this example, but let's look at the circle a bit more.

We could solve the equation of the circle for y to get:

$y = +\sqrt{(25 - x^2)}$ on the top half of the circle, and
$y = -\sqrt{(25 - x^2)}$ on the bottom half of the circle

Hence

$(dy/dx) = (1/2)(-2x)/\sqrt{(25 - x^2)} = -x/y$ on the top half of the circle, and
$(dy/dx) = -(1/2)(-2x)/\sqrt{(25 - x^2)} = -x/y$ on the bottom half of the circle

This won't work for $y^3 + 2y^2 - 10x = 6$ at $(1, 2)$, because we can't solve *explicitly* for $y = f(x)$. But note that the equation of the circle (and of our more complicated example) *implicitly* defines y as some function of x. And that's all we need – because now we can apply the chain rule.

- $x^2 + y^2 = 25$, so $2x + 2y(dy/dx) = 0$ and hence $dy/dx = -2x/2y = -x/y$, and surely this was a lot easier than solving for y as an explicit function of x and then differentiating!
- $y^3 + 2y^2 - 10x = 6$, so $3y^2(dy/dx) + 4y(dy/dx) - 10 = 0$ and hence $dy/dx = 10/(3y^2 + 4y)$, provided $y \neq 0$ and $y \neq -4/3$ (where dy/dx blows up). So, at $(1, 2)$:

  the tangent has slope $10/20 = 0.5$
  and equation $(y - 2)/(x - 1) = 0.5$, or $y = 0.5x + 1.5$

  What about at the point $(-0.5, -1)$?
  $dy/dx = 10/(-1) = -10$, so the tangent has equation $y = -10x - 6$

Examples:

- 12.4.1. $2x + 3y = 7$, so $2 + 3dy/dx = 0$ or $dy/dx = -2/3$ (which is obvious from $y = 7/3 - 2/3x$)
- 12.4.5. $2x + 3y = xy$, so $2 + 3dy/dx = x(dy/dx) + y$ or $dy/dx = (y - 2)/(3 - x)$
  Please note that this is a lot easier than solving for $y = 2x/(x - 3)$ and then differentiating: $y' = [(x - 3)(2) - 2x(1)]/(x - 3)^2 = -6/(x - 3)^2$. I leave it to you to verify that $(y - 2)/(3 - x) = -6/(x - 3)^2$.
- 12.4.19. $p^2 - pq = 5p^2q^2$, so $2p(dp/dq) - p - q(dp/dq) = 5(p^2(2q) + 2p(dp/dq)q^2)$, or
  $$dp/dq = (10p^2q + p)/(2p - q - 10pq^2)$$
- 12.4.35. $4x^2 + 2y^2 = 12$; find dy/dx at $(1, -2)$: $8x + 4y(dy/dx) = 0$, so $dy/dx = -2x/y = 1$ at $(1, -2)$
- 12.4.49. C, the cost in \$ of building a house, depends on k and e, the number of carpenters and electricians used. Specifically, $C = 15000 + 50k^2 + 60e^2$. If $C = \$200,000$ and $e = 15$, find and interpret dk/de.

  At $e = 15$, and $C = 200,000$, $k = \sqrt{\dfrac{200000 - 15000 - 60(15^2)}{50}} = \sqrt{3430} = 58.566$
  $100k(dk/de) + 120e = 0$, so $dk/de = -1.2e/k$, and at $e = 15$
  $$dk/de = -1.2(15)/58.566 = -0.307\text{carpenter/electrician}$$
  This means that an electrician can be replaced by 0.307 carpenters and the cost of building the house will still be \$200,000.

APPLICATION: Rates of substitution

- Suppose that the production of widgets is governed by a production function that specifies output in terms of inputs.
- For simplicity, suppose there are only 2 inputs:
  - "labor" (L, measured in, e.g., number of employees) and
  - "capital" (K, measured in, e.g., \$ invested in plant and equipment)
- For any specified level of production (P, in units of output), there may be many possible combinations of L and K that achieve that level of output.
- For a given specified level of production, the production function defines L as an implicit function of K (and vice versa).
- The derivative dL/dK is called the "marginal rate of substitution of labor for capital."

Our last example involved such a marginal rate of substitution (of carpenters for electricians).

Economists often use a "Cobb-Douglas production function"
$$P = cL^{\alpha}K^{1-\alpha}$$
where c and α are constants. In what follows, we'll assume $c = 1$ – this just amounts to redefining the units of output. Note that if we multiply both L and K by a constant factor λ (called a "scale factor"), then the output P gets multiplied by

the same factor (because $\lambda^{\alpha}\lambda^{1-\alpha} = \lambda$). Economists say the "production function has constant returns to scale."

Differentiating with respect to K, we get

$$0 = L^{\alpha}(1-\alpha)K^{-\alpha} + K^{(1-\alpha)}(\alpha L^{\alpha-1})(dL/dK), \text{ so } \frac{dL}{dK} = -\left(\frac{1-\alpha}{\alpha}\right)\left(\frac{L}{K}\right) \left(\text{and } \frac{dK}{dL} = -\left(\frac{\alpha}{1-\alpha}\right)\left(\frac{K}{L}\right)\right)$$

This has the interesting property that the rate of substitution depends only on the proportion of labor to capital and not on their levels, i.e., if we double both the labor and capital inputs, the level of production will be doubled, but the marginal rate of substitution won't change.

Example:

- 12.4.53. $P = L^{0.6}K^{0.4}$, with P = daily production of CD's, L = # of workers, K = annual expenditures ($). Find dK/dL when P = 20000 and L = 100 (and $K = (20000/100^{0.6})^{2.5}$ = $56,568,542)

  dK/dL = –(0.6/0.4)(56,568,542/100) = –$848528/worker

VIII- Maxima and Minima

In all sorts of practical problems, we have some quantity which we wish to optimize, that is, either to maximize (for instance output, profit, grades) or minimize (for instance input, cost, time). We can use derivatives to help us find optima.

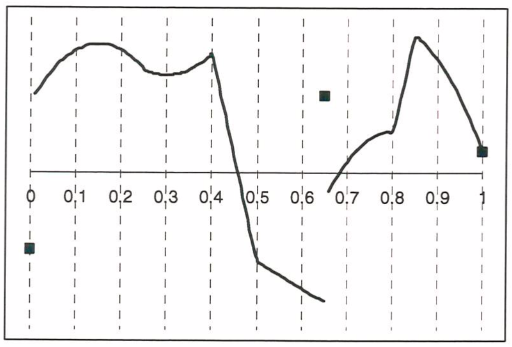
Interval notation and "open" and "closed" ends of intervals

- **(a, b) = {x | a < x < b} – both ends open (end points not included)**
- **[a, b] = {x | a ≤ x ≤ b} – both ends closed (end points included)**
- **[a, b) = {x | a ≤ x < b} – closed at bottom, open at top**
- **(a, b] = {x | a < x ≤ b} – open at bottom, closed at top**
- **(a, +∞) = {x | x > a}**
- **[a, +∞) = {x | x ≥ a}**
- **(–∞, b) = {x | x < b}**
- **(–∞, b] = {x | x ≤ b}**
- **(–∞, +∞) = the set of all possible values of x**

Suppose we have a function, f(x), defined on some interval [a, b], and we wish to find its maxima and minima (optima or extrema) on that interval. We distinguish between:

- **a global (or absolute) maximum, x_0 – where no x in the interval has f(x) > f(x_0), and**
- **a local (or relative) maximum, x_0 – where x_0 is a global maximum if we consider only points "close to x_0"**

The picture below (slide attached) illustrates the various possible ways extrema can occur:



0, 0.3, and 1 are local minima (and there is no global minimum)
0.15, 0.4, 0.65, and 0.85 are local maxima (and 0.85 is the global maximum)

From this picture, we see that extrema can occur at

- **(closed) endpoints of intervals (like 1)**

- points where the derivative exists and equals 0 ("stationary" points, like 0.15 and 0.3)
- points where the derivative fails to exist ("singular" points), either because f(x) is discontinuous (like 0.65) or because f(x) has a kink (like 0.4 or 0.85)
- points where f(x) is discontinuous (like 0 or 0.65)

We saw that f(x) is increasing (respectively, decreasing) instantaneously at points where $f'(x) > 0$ (respectively, $f'(x) < 0$). When $f'(x) = 0$, at that instant f(x) is neither increasing nor decreasing, and we say that f(x) is "stationary" at such a point.

For the most part, we'll be looking at functions that are continuous and differentiable, so we'll look at stationary points and endpoints when looking for extrema. But remember, even apparently nice functions, like $x^{2/3}$ on [-2, 2], can turn out to be kinky!

Examples:

- 13.1.1, 3, 5, 7, 9 – see pictures
- 13.1.23. $g(x) = x^3 - 12x$ on $[-4, 4]$
  $g'(x) = 3x^2 - 12 = 0$ at $x \pm 2$.
  Note: $g'(x) > 0$ on $[-4, -2)$ and $(2, 4]$ and is < 0 on $(-2, 2)$.
  Also $g(x) = -16, 16, -16$, and 16 at $-4, -2, 2$, and 4.
  Hence $g(x)$ has absolute minima at $x = -4$ and $+2$, and absolute maxima at $x = -2$ and $+4$.
- 13.1.31. $g(t) = t^4/4 - 2t^3/3 + t^2/2$ on $(-\infty, +\infty)$
  $g'(t) = t^3 - 2t^2 + t = t(t^2 - 2t + 1) = t(t-1)^2 = 0$ at $t = 0$ and 1
  Note: $g'(t) < 0$ on $(-\infty, 0)$ and is ≥ 0 on $(0, +\infty)$ (and strictly > 0, except at $t = 1$).
  Obviously, g(t) gets big if t is large and negative or large and positive.
  Hence, g(t) has an absolute minimum at $t = 0$.
- 13.1.33. $f(t) = (t^2 + 1)/(t^2 - 1)$ on $[-2, -1) \cup (-1, 1) \cup (1, 2]$
  $f'(t) = [(t^2 - 1)(2t) - (t^2 + 1)(2t)]/(t^2 - 1)^2 = -4t/(t^2 - 1)^2 = 0$ only at $t = 0$.
  Note: $f'(t)$ blows up at $t = \pm 1$, but otherwise is > 0 if $t < 0$, and is < 0 if $t > 0$
  $f(\pm 2) = 5/3$ and $f(0) = -1$
  Hence, f(t) has local minima at $t = \pm 2$, and a local maximum at $t = 0$. It has no global maxima or minima.

Maxima and Minima

To optimize, we look at (DRAW PICTURES!):

A. End points. If $f(x)$ is differentiable, then
   1. $x \geq a$: $f'(a^+) > 0$ for a minimum, $f'(a^+) < 0$ for a maximum
   2. $x \leq b$: $f'(b^-) < 0$ for a minimum, $f'(b^-) > 0$ for a maximum

B. Stationary points. If $f(x)$ is differentiable, then $f'(x) = 0$ and
   1. $f'(x^-) < 0$ and $f'(x^+) > 0$ for a minimum
   2. $f'(x^-) > 0$ and $f'(x^+) < 0$ for a maximum
   3. an "inflection point" if f' doesn't change sign

C. Singular points. If f continuous at x and differentiable elsewhere, then the tests in (B) hold in this case as well.

Examples:

- 13.2.5. minimize $x^2 + y^2$ with $x + 2y = 10$
  Since we don't yet know how to solve optimization problems with 2 variables, solve the constraint to get $x = 10 - 2y$, and plug back into the objective to get a one-variable optimization problem:
  minimize $(10 - 2y)^2 + y^2$
  Take the derivative and set it to 0: $2(10 - 2y)(-2) + 2y = 0 = 10y - 40$
  $(x, y) = (2, 4)$ is the only stationary point
  Check for minimum: $10y - 40 < 0$ at $y = 4^-$ and $10y - 40 > 0$ at $y = 4^+$
  (Could we have eliminated y instead of x? Yes. I leave it to you to verify that we'd have obtained the same results.)

- 13.2.11. Fencing on E&W costs \$4/ft, on N&S costs \$2/ft. Find the largest fenced area with an \$80 budget.

  maximize xy with $2(4x + 2y) = 80$ (or $2x + y = 20$) (and, implicitly: $x, y \geq 0$, $x \leq 10$, $y \leq 20$). Again we solve the constraint for $y = 20 - 2x$, and plug back into the objective to get:
  maximize $x(20 - 2x) = 2(10x - x^2)$
  Take the derivative and set it to 0: $2(10 - 2x) = 0$
  $(x, y) = (5, 10)$ is the only stationary point
  Check for maximum: $10 - 2x > 0$ at $x = 5^-$ and $10 - 2x < 0$ at $x = 5^+$
  So the largest possible area to fence is 50 square feet, with a $5' \text{x} 10'$ plot.

- 13.2.17. $p = 500{,}000/q^{1.5}$ = price in \$/lb when q pounds of tuna are sold in a month.

maximize revenue = $pq = 500{,}000/q^{0.5} = r(q)$ with $q \geq 5000$.
$r'(q) = -250{,}000/q^{1.5} < 0$ for all $q \geq 0$, so as q increases, $r(q)$ decreases.
Hence $r(q)$ is maximized at $q = 5000$ pounds, at which point $p = 500000/5000^{1.5} = \$1.41/\text{lb}$ and $r(q) = pq = \$7050/\text{month}$.

- 13.2.23 $p = 1000/q^{0.3}$, where p = price per headset in \$ and q = weekly sales of headsets;

  total cost per headset is \$100

  a) profit = $\pi(q) = pq - 100q = 1000q^{0.7} - 100q$

  $\pi'(q) = 700q^{-0.3} - 100 = 0 \rightarrow q^{0.3} = 7 \rightarrow q = 7^{3.33333} = 656.14$, or 656 to the nearest whole unit. (Note that $\pi'(q)$ switches from positive to negative at $q = 656$, and there are no other stationary points, so our point is indeed a global maximum. At this level, $\pi(656) = 1000(656^{0.7}) - 65600 = \$\$28{,}120$.

  b) The charge per headset is $p = 1000/656^{0.3} = \$142.86$, or \$143 to the nearest \$

- 13.2.31. maximize LWH, with $H = W$ and $L + W + H = 62$ inches. Solving for $L = 62 - 2H$ and plugging back in, we want to maximize volume = $V(H) = (62 - 2H)H^2 = 62H^2 - 2H^3$, with the physical constraints that $0 \leq H \leq 31$ (because $H = W$).

  $V'(H) = 124H - 6H^2 = H(124 - 6H) = 0 \rightarrow H = 0$ (obviously a local minimum) or $H = 124/6 = 20.66667$ inches (so $W = 20.66667$, and so is L). Clearly V' changes sign from positive to negative at $H = 124/6$, and nowhere else does $V' = 0$, so we have a local maximum which is also a global maximum.

  The resulting maximum volume = $20.66667^3 = 8827$ cubic inches, or about 5.11 cubic feet.

- 13.2.45. If x = the number of copies of a graphing program sold to a customer, then the manufacturer's revenue in dollars is $500\sqrt{x}$. Production cost is $10000 + 2x$ dollars, so profit on an order for x copies is $\pi(x) = 500\sqrt{x} - 10000 - 2x$.

  Average profit (or profit per copy) is $A(x) = \pi(x)/x = 500/\sqrt{x} - 10000/x - 2$. To maximize average profit: $A'(x) = -250/x^{3/2} + 10000/x^2 = (-250\sqrt{x} + 10000)/x^2 = 0 \rightarrow \sqrt{x} = 10000/250 = 40$, or $x = 1600$ copies, the only stationary point. Note that the derivative goes from positive to negative at $x = 1600$, so we do indeed have a local and global maximum.

  At $x = 1600$, $A(x) = 500/40 - 10000/1600 - 2 = \$4.25$ per copy. Marginal cost is $\pi'(x) = 250/\sqrt{x} - 2$, so at $x = 1600$, $\pi'(x) = 250/40 - 2 = \$4.25$ per copy. This is not a coincidence.

- **13.2.53. With $50 + x$ apple trees in an orchard, the annual yield per tree is $100 - x$ pounds. What size orchard maximizes total yield?**

  maximize $(50 + x)(100 - x) = y(x)$
  $y'(x) = (50 + x)(-1) + (100 - x)(1) = 50 - 2x = 0$ → $x = 25$ more trees.
  Since $y'(x)$ goes from $+$ to $-$ as x crosses 25, we do indeed have a maximum.

IX- Second Derivatives & Related Rates

Since $f'(x)$ is itself a function of x, we can ask how it changes as x changes – that is, what is the derivative of $f'(x)$? This quantity is called the "2nd derivative of $f(x)$ and is variously denoted by:

$$f''(x),\ y'',\ \frac{d^2f(x)}{dx^2},\ \text{and}\ \frac{d^2y}{dx^2}.$$

(also, we now call $f'(x)$ "the 1st derivative of $f(x)$.)

What does $f''(x)$ tell us? Does it have any applied interpretation? Consider the example of

$f(t)$ = distance in miles traveled in t hours
$f'(t)$ = rate at which distance is changing = speed (or velocity) at time t
$f''(t)$ = rate at which speed is changing = acceleration at time t

In general. what does $f''(x)$ tell us?

1. If $f''(x) > 0$, then $f'(x)$ (the slope of the curve) is *increasing* as x increases, i.e., the curve is bending upward. DRAW PICTURES of $f'' > 0$ when $f' > 0$ and when $f' < 0$. We say that the curve is "concave up" or "convex."

2. If $f''(x) < 0$, then $f'(x)$ (the slope of the curve) is *decreasing* as x increases, i.e., the curve is bending downward. DRAW PICTURES of $f'' > 0$ when $f' > 0$ and when $f' < 0$. We say that the curve is "concave up" or just simply "concave."

3. A point at which $f''(x)$ changes sign is called an "inflection point."

   a. Consider $f(x) = x^3$ at $x = 0$. (DRAW A PICTURE!)
   $f'(x) = 3x^2 = 0$ at $x = 0, > 0$ at $x \neq 0$ → $f(x)$ is always increasing
   $$f''(x) = 6x \begin{cases} < 0 \text{ at } x = 0^- \text{ (i.e., if } x \text{ is close to 0, but negative)} \\ = 0 \text{ at } x = 0, \text{, and } > 0 \text{ if } x > 0 \\ > 0 \text{ at } x = 0^+ \text{ (i.e., if } x \text{ is close to 0, but positive)} \end{cases}$$
   This curve has a "horizontal point of inflection" at $x = 0$

   b. Consider $f(x) = x^{1/3}$ at $x = 0$. (DRAW A PICTURE!)
   $f'(x) = (1/3)x^{-2/3}$, which blows up at $x = 0$ at $x = 0$, but which is > 0 at $x \neq 0$ → $f(x)$ is always increasing
   $$f''(x) = (-2/9)x^{-5/3} \begin{cases} < 0 \text{ at } x = 0^- \\ = 0 \text{ at } x = 0, \text{, and } > 0 \text{ if } x > 0 \\ > 0 \text{ at } x = 0^+ \end{cases}$$
   This curve has a "vertical point of inflection" at $x = 0$

Second Derivative Test for Optima:

If $f(x)$ is "twice differentiable" (i.e., has a 1st and 2nd derivative), and $f'(x_0) = 0$, then

1. **x_0 is a relative minimum if $f''(x_0) > 0$ (because $f'(x)$ changes from $-$ to $+$ at x_0)**
2. **x_0 is a relative maximum if $f''(x_0) < 0$ (because $f'(x)$ changes from $+$ to $-$ at x_0)**
3. **The test fails if $f''(x_0) = 0$ (Examples: x^3, x^4, and $-x^4$ at $x_0 = 0$)**

This is useful in curve sketching:

Recall that we looked at $f(t) = (t^2 + 1)/(t^2 - 1)$ on $[-2, 2]$

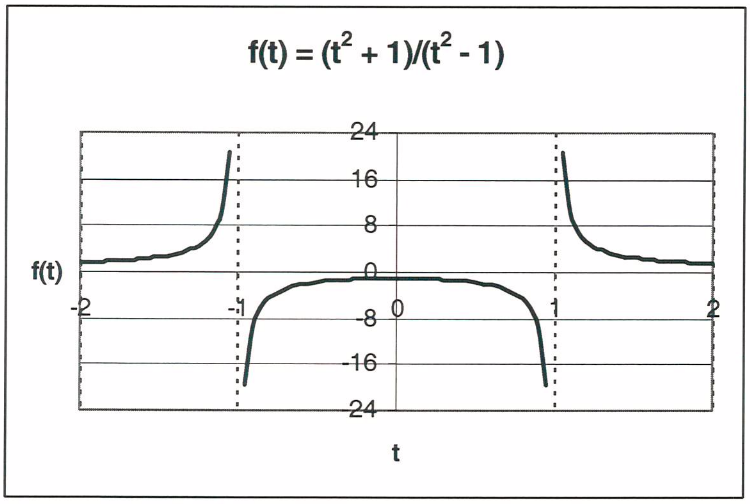
We saw that

- **$f(\pm 2) = 5/3$ and $f(0) = -1$,**
- **$f'(t) = -4t/(t^2 - 1)^2$,**
- **$f'(t)$ blows up at $t = \pm 1$, but otherwise is > 0 if $t < 0$, and < 0 if $t > 0$**
- **$f(t)$ has local minima at $t = \pm 2$, a local maximum at $t = 0$, and no global maxima or minima.**

What about curvature? $f''(t) = \dfrac{(t^2-1)^2(-4) + 4t(t^2-1)(2t)}{(t^2-1)^4} = \dfrac{12t^2+4}{(t^2-1)^3} \begin{cases} > 0 & \text{if } |t| > 1 \\ < 0 & \text{if } |t| < 1 \end{cases}$

(convex)

(concave)



Related Rates: Sometimes we have several quantities which are related to each other by some functional relationship, with each related to some other quantity (frequently time). We know the rate of change of one and we want to find the rate of change of the other. Such questions involve "related rates." They involve no new theory, but just apply what we've seen before, and usually involve implicit differentiation. Let's look at some examples in section 13.4.

9. The area of a circular sunspot is growing at a rate of 1200 $km^2/s = dA/dt$.
   a) How fast is the radius growing when it (the radius) is 10^4 km?
   $A = \pi r^2$, so $dA/dt = 2\pi r(dr/dt)$ → $dr/dt = (1/2\pi r)(dA/dt) = 1.2x10^3/2\pi(10^4) = 0.019$ km/s = 19 m/s

13. The average cost of making x CD players/wk is $\overline{C}(x) = 150000/x + 20 + 0.0001x$.

Suppose x is now 3000 units/week and $\frac{dx}{dt} = 100$ units/week/week. What is $\frac{d\overline{C}}{dt}$?

$$\frac{d\overline{C}}{dt} = -150000x^{-2}\frac{dx}{dt} + 0.0001\frac{dx}{dt} = (-150000/3000^2 + 0.0001)(100) = -1.6567$$

i.e., average cost is decreasing by $1.6567/player/week

17. D(t) = weekly sales of lemonade at time t (currently at 50 cups/week)
P(t) = price of lemonade at time t (currently at 30¢/cup)
dD/dt = –5cups/week/week

If raising price does not affect demand, by how much must you raise your price to keep revenue constant?

R(t) = D(t)P(t) → 0 = D(t)[dP/dt] + [dD/dt]P(t)

$$\rightarrow \frac{dP}{dt} = -\left(\frac{P}{D}\right)\left(\frac{dD}{dt}\right) = -(30/50)(-5) = 3, \text{ so increase prices by}$$

3¢/cup/week

23. One ship is 40 miles north of Montauk and sailing north at 20 mph. A second is 50 miles east of Montauk and sailing east at 30 mph. How fast is the distance between the ships changing? Is it increasing or decreasing?

Let their positions be N(t) and E(t), with D(t) being the distance between them. By the Pythagorean Theorem, $[D(t)]^2 = [N(t)]^2 + [E(t)]^2$. Hence, dropping the common factor of 2:

$$D(t)\frac{dD(t)}{dt} = N(t)\frac{dN(t)}{dt} + E(t)\frac{dE(t)}{dt}$$

So at the current moment, $\frac{dD(t)}{dt} = \frac{40(20)+(50)(30)}{\sqrt{40^2+50^2}}$ **= 35.92 mph > 0; their distance is increasing.**

27. x(t) = number of workers at time t
y(t) = daily operating budget at time t (in $/day)
P(t) = number of automobiles produced per *year* = $10x^{0.3}y^{0.7}$

At the moment, P = 1000 cars/year and x = 150 workers (so y = $84.05/day).
If dx/dt = 10 workers/year, by how much will the daily operating budget change if output is kept constant?

0 = 10[0.3(y/x)(dx/dt) + 0.7(x/y)(dy/dt)]
→ dy/dt = –(3/7)(y/x)(dx/dt) = –(3/7)(84.05/150)(10) = –2.40 (i.e., decrease by $2.40/day/year)

X- Logarithmic & Exponential Functions

Let's go back and look at two very important classes of functions – logarithmic and exponential

Laws of exponents: let b & c be any *positive* real numbers, and x & y be any real numbers (+, 0, or –)

Law	Example
1. $b^x b^y = b^{x+y}$	$10^2 10^3 = 100(1000) = 100{,}000 = 10^5$
2. $b^x/b^y = b^{x-y}$	$2^6/2^2 = 64/4 = 16 = 2^4$
3. Corollary: $b^0 = b^x/b^x = 1$	$10^0 = 1,\ 2^0 = 1$
4. Corollary: $b^{-x} = b^0/b^x = 1/b^{-x}$	$5^{-2} = 1/5^2 = 1/25 = 0.04$
5. $(b^x)^y = b^{xy}$ $= 3^6$	$(3^3)^2 = 27^2 = 729 = 3^6$, $(3^2)^3 = 9^3 = 729$
6. $(bc)^x = b^x c^x$	$[2(3)]^3 = 6^3 = 216 = 8(27) = 2^3 3^3$
7. Corollary: $(b/c)^x = b^x/c^x$	$(12/4)^3 = 3^3 = 27 = 1728/64 = 12^3/4^3$

An *exponential function* is a function of the form $f(x) = Ab^x$, where A is any constant and b is a positive constant. If you experiment a bit with different values of A and b, you quickly discover that (DRAW PICTURES!):

- **$b = 1 \rightarrow f(x) = A$ for all x**
- **$b > 1$ & $A > 0 \rightarrow f(x) > 0$ for all x, and goes to 0 as x goes to $-\infty$, to $+\infty$ as x goes to $+\infty$**
- **$b < 1$ & $A > 0 \rightarrow f(x) > 0$ for all x, and goes to $+\infty$ as x goes to $-\infty$, to 0 as x goes to $+\infty$ (not a surprise: if $b < 1$, then $1/b > 1$, and $Ab^x = A(1/b)^{-x}$)**
- **$A < 0$ just flips the previous two cases around the x-axis**

Let's restrict ourselves to the case where $A > 0$.

- **If $b > 1$, then f(x) increases as x increases. We say that f(x) is *growing exponentially*.**
- **If $b < 1$, then f(x) decreases as x increases. We say that f(x) is *decaying exponentially*.**

Exponential functions are useful in describing all sorts of natural phenomena involving growth and decay. A common business growth example is that of compound interest. If I put \$1000 in a bank account with a 5% annual interest rate, then:

- **after 1 year, I have 1000(1.05) = \$1050 in the bank**
- **after 2 years, I have $1050(1.05) = 1000(1.05)^2$ = \$1102.50 (because I earn interest on the earlier interest as well as on the "principal"**
- **after n years, I have $1000(1.05)^n$ dollars in the bank**

If a battery sits around in an unused portable radio, it loses charge at the rate of 2% per month. If its initial charge will provide 10 continuous hours of playing time, then

- after 1 idle month, it will provide 10(0.98) = 9.8 continuous hours of playing time
- after 2 idle months, it will provide $9.8(0.98) = 10(0.98)^2 = 9.604$ continuous hours of playing time
- after n idle months, it will provide $10(0.98)^n$ continuous hours of playing time

Returning to compound interest, banks typically compound interest more frequently than annually. If a very generous bank pays interest at a nominal 100% annual rate and it is compounded:

- annually, then at the end of a year, \$1 grows to $(1 + 1)^1 = \$2.00$
- semiannually, then at the end of a year, \$1 grows to $(1 + 1/2)^2 = \$2.25$
- quarterly, then at the end of a year, \$1 grows to $(1 + 1/4)^4 = \$2.44140625$
- monthly, then at the end of a year, \$1 grows to $(1 + 1/12)^{12} = \$2.61303529$
- daily, then at the end of a year, \$1 grows to $(1 + 1/365)^{365} = \$2.71456748$
- every second, then at the end of a year, \$1 grows to $(1 + 1/31{,}536{,}000)^{31{,}536{,}000} =$ \$2.718281781

So more frequent compounding is better, but after a while, it really doesn't help in any significant way. In the limit, if they compound *continuously*, you'll have \$2.7182818284590 at the end of one year.

Mathematicians denote $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$ by $e = 2.7182818284590\ldots$. This is one of those incredible numbers like π that just seems to pop up all over the place.

General interest formulas:

- P = initial amount ("principal", or "present value")
- n = number of periods/year
- r = interest rate per year (so r/n = interest rate per period)
- t = number of years
- F = final amount (or "future value") after t years
  $= P(1 + r/n)^{nt}$ with periodic compounding
  $= Pe^{rt}$ with continuous compounding

Examples

- 10.2.49 Find an exponential function, $y = Ab^x$, through (1, 3) and (3, 6)
  $3 = Ab$ and $6 = Ab^3$
  Dividing these gives us $2 = 6/3 = Ab^3/Ab = b^2$
  So $b = \sqrt{2}$ and $A = 3/\sqrt{2}$ → $y = (3/\sqrt{2})(\sqrt{2})^x = 2.1213(1.4142^x)$
- 10.2.53 Find an exponential function, $y = Ab^x$, through (2, 3) and (6, 2)
  $3 = Ab^2$ and $2 = Ab^6$
  Dividing these gives us $2/3 = Ab^6/Ab^2 = b^4$
  So $b = (2/3)^{1/4} = 0.9036$ and $A = 3/0.9036^2 = 3.6742$ → $y = 3.6742(0.9036^x)$
- 10.2.59 r = 4.15%, P = £5000, t = 5, n = 1. Find F to the nearest £.
  $F = P(1 + r/n)^{nt} = 5000(1.0415)^5 = £6127$
  What if interest was compounded continuously?

$F = Pe^{rt} = 5000e^{5(0.0415)} = £6153$

- 10.2.61 How much would have been needed at start in 10.2.59 to get £250,000 in 5 years?

  $F = P(1 + r/n)^{nt}$: $250000 = P(1.0415)^5$, so $P = 250000/(1.0415)^5 =$ £204,006

  What if interest was compounded continuously?

  $F = Pe^{rt}$: $250000 = Pe^{5(0.0415)}$, so $P = 250000/e^{5(0.0145)} = £203{,}153$
- 10.2.69 After taking an aspirin, a patient absorbs 300 mg into her bloodstream. Aspirin has a half life of 2 hours in the bloodstream (i.e., 50% is removed every 2 hours). How much is left after 5 hours?

  $300(0.5)^{5/2} = 53.03$ mg
- 10.2.79 Global Warming: According to a UN report, atmospheric CO_2 content in ppm by volume is given by $C(t) = 277e^{0.00353t}$, where t = years since 1750.

  a. Estimate CO_2 content in 1950, 2000, 2050, and 2100

  1950: $277e^{0.00353(200)} = 561.16$ ppm, 2000: $277e^{0.00353(250)} = 669.49$ ppm

  2050: $277e^{0.00353(300)} = 798.73$ ppm, 2100: $277e^{0.00353(350)} = 952.91$ ppm

  b. To nearest decade, when will it surpass 700 ppm?

  We'll come back to this after we've discussed logarithms

- 10.2.69 After taking an aspirin, a patient absorbs 300 mg into her bloodstream. Aspirin has a half life of 2 hours in the bloodstream (i.e., 50% is removed every 2 hours). How much is left after 5 hours?

  $300(0.5)^{5/2} = 53.03$ mg

Logarithmic & Exponential Functions

Logarithms ("logs" for short), as we shall shortly see are the inverse of exponential functions. Given a *base* b, an exponential function (b^x) tells you what number you get if you raise the base to some power x. A logarithmic function ($\log_b x$) tells you to what power you have to raise b in order to get x. This gives us the important relationships:

$$\log_b b^x = x \text{ and } b^{\log_b x} = x$$

which confirm that logs and exponentials are inverse functions – i.e., they undo each other

Examples:

- $3^4 = 81$, so $4 = \log_3 81$
- $4^3 = 64$, so $3 = \log_4 64$
- $0.001 = 10^{-3}$, $\log_{10} 0.001 = -3$

Laws of logarithms (in what follows all numbers whose logs are being taken are > 0)

Law	Example
0. $\log_b(b) = 1$	$10^1 = 10$, $0.5^1 = 0.5$
1. $\log_b(xy) = \log_b(x) + \log_b(y)$	$100{,}000 = 100(1000)$ and $\log_{10} 100{,}000 = 5 = 2 + 3 = \log_{10} 100 + \log_{10} 1000$
2. $\log_b(x/y) = \log_b(x) - \log_b(y)$	$16 = 64/4$ and $\log_2 16 = 4 = 6 - 2 = \log_2 64 - \log_2 4$
3. Corollary: $\log_b(1) = 0$	$10^0 = 1$, $2^0 = 1$
4. Corollary: $\log_b(1/x) = -\log_b(x)$	$0.04 = 1/25$ and $\log_5 0.04 = -2 = -\log_5 25$
5. $\log_b(x^y) = y\log_b(x)$	$(3^3)^2 = 27^2 = 729 = 3^6$, $(3^2)^3 = 9^3 = 729 = 3^6$

If you think carefully about these, you'll realize that they're just restatements of the laws for exponents – because logs are exponents!

The two most commonly encountered logarithmic functions are:

- "Common" logs ($b = 10$), and we usually write $\log(x)$ rather than $\log_{10}(x)$
- "Natural" logs ($b = e = 2.71828\ldots$), and we usually write $\ln(x)$ rather than $\log_e(x)$

Finding logarithms is not usually so easy as it was in our examples. Fortunately, calculators have single buttons for common and natural logs, and Excel has built-in functions (ln and log10) for calculating logs.

Sometimes you need to convert from logs in one base to another base. Let's see how to do that. Suppose you want $\log_b x$ and you already know $\log_c x$. Let's see how to do that.

$x = b^{\log_b x} = \left(c^{\log_c b}\right)^{\log_b x} = c^{(\log_c b)(\log_b x)} = c^{\log_c x}$, so $\log_c x = (\log_c b)(\log_b x)$, or $\log_b x = \log_c x/\log_c b$

and in particular, $\log_b x = \ln(x)/\ln(b)$

Examples

- 10.2.79 Global Warming: According to a UN report, atmospheric CO_2 content in ppm by volume is given by $C(t) = 277e^{0.00353t}$, where $t =$ years since 1750.
  b. To nearest decade, when will it surpass 700 ppm?
  $277e^{0.00353t} = 700$, so $e^{0.00353t} = 700/277 = 2.5271$. Now take natural logs:
  $0.00353t = \ln(2.5271) = 0.9271$, so $t = 0.9271/0.00353 = 262.63$ (or 260 to the nearest decade → $1750 + 260 = 2010$)
- 10.3.19 Invest \$500 at 10% compounded continuously. How long until you have at least \$700?
  $500e^{0.1t} = 700 \rightarrow e^{0.1t} = 700/500 = 1.4 \rightarrow 0.1t = \ln(1.4) = 0.3365 \rightarrow t =$ 3.36 years
  What about if compounded annually?
  $500(1.1)^t = 700 \rightarrow 1.1^t = 700/500 = 1.4 \rightarrow t\ln(1.1) = \ln(1.4)$

→ $t = \ln(1.4)/\ln(1.1) = 0.3365/0.0953 = 3.53$ → it will take 4 years

- **10.3.23** Carbon-14 Dating: $C(t) = A(0.999879)^t$ = amount of C_{14} left after t years if initial amount of C_{14} was A. If 99.95% of a skull's original C_{14} is gone, how old is the skull?
  $0.0005A = A(0.999879)^t$ → $\ln(0.0005) = t \ln(0.999879)$
  → $t = \ln(0.0005)/\ln(0.999879) = -7.6009/(-0.000121) = 62813.6$
  → it is 62,814 years old

- **10.3.27** \$10400 is invested at 5.2% per year compounded monthly. How many months until you have at least \$20,000?
  $10400(1 + 0.052/12)^x = 20000$ → $1.004333^x = 20000/10400 = 1.9231$
  → $x \ln(1.00433) = \ln(1.9231)$ → $x = \ln(1.9231)/\ln(1.00433) = 0.6539/0.00432 = 151.4$ → it will take 152 months (at the end of 151 months, it will still be a bit short)

- **10.3.39** 300 mg of aspirin in the blood decays exponentially with a half-life of 2 hours. How long before only 100 mg are left?
  $300(0.5)^z = 100$ → $z = \ln(0.3333)/\ln(0.5) = -1.0987/(-0.6931) = 1.585$
  two-hour intervals, or 3.2 hours

- **10.3.45** Richter earthquake scale: $R = (2/3)(\log_{10} E - 4.4)$, with E = released energy in joules.
  a. 1906 SF earthquake had $R = 8.2$. How much energy did it release?
     $8.2 = R = (2/3)(\log_{10} E - 4.4)$ → $\log_{10} E = 1.5(8.2) + 4.4 = 16.7$,
     so $E = 10^{16.7} = 5.012(10^{16})$ joules
     (To give some meaning to this, $4.2(10^{15})$ joules is the energy released by a megaton of TNT, so the earthquake was the equivalent of about 12 one-megaton atomic bombs. The bombs dropped on Hiroshima and Nagasaki were approximately 20-kiloton bombs, making the SF earthquake the equivalent of 600 such bombs!)
  c. What is the energy released ratio, E_2/E_1, of quakes with readings R_2 and R_1?
     $E = 10^{1.5R + 4.4}$, so $E_2/E_1 = 10^{1.5R2 + 4.4}/10^{1.5R1 + 4.4} = 10^{1.5(R2 - R1)}$
  b. 1989 SF quake measured 7.1. What percentage of the 1906 released energy did it release? $10^{1.5(7.1 - 8.2)} = 10^{-1.65} = 0.0224$, or about 2.24%
  d. If quake 2 measures 2 points higher than quake 1, how much more energy does it release? $10^{1.5(R2 - R1)} = 10^3$ → 1000 times as much!

XI- Derivatives of Logarithmic & Exponential Functions

Derivatives of log functions:

1. **Natural logs:** $(d/dx)(\ln(x)) = 1/x$ (if $x > 0$)
2. **Chain rule:** $(d/dx)(\ln(u)) = (1/u)(du/dx)$

   Corollary: If $x < 0$, then $-x > 0$ and $(d/dx)(\ln(-x)) = (1/(-x))(-1) = 1/x$.
   Putting this together with (1) gives us:

1'. $(d/dx)(\ln|x|) = 1/x$ (if $x \neq 0$)

3. **Other bases:** recall that $\log_b x = \ln(x)/\ln(b)$, so $(d/dx)(\log_b(x)) = 1/[x \ln(b)]$

Derivatives of exponential functions: Since $x = \ln(e^x)$, the chain rule tells us: $1 = (1/e^x)(de^x/dx)$, so

4. **Natural exponential function:** $(d/dx)e^x = e^x$
5. **Chain rule:** $(d/dx)e^u = e^u(du/dx)$
6. **Other bases:** recall that $b^x = e^{x\ln(b)}$, so $(d/dx)b^x = e^{x\ln(b)}\ln(b) = b^x\ln(b)$

Examples:

- Recall that the power rule says that $(d/dx)x^n = nx^{n-1}$, when n is a constant, but we said that didn't work for x^x. So what is the derivative of x^x? Well, $x^x = e^{x\ln(x)}$, so let's apply the chain rule: $\left(\frac{d}{dx}\right)(e^{x\ln(x)}) = e^{x\ln(x)}\left(\frac{d}{dx}\right)(x\ln(x)) = e^{x\ln(x)}\left(x\left(\frac{1}{x}\right) + \ln(x)\right) = x^x(1 + \ln(x))$
- 12.3.9. $(d/dx)(e^{-x}) = -e^{-x}$
- 12.3.11. $(d/dx)(4^x) = (d/dx)(e^{x\ln(4)}) = e^{x\ln(4)}\ln(4) = 4^x\ln(4)$
- 12.3.13. $\frac{d}{dx}2^{x^2-1} = \frac{d}{dx}e^{(x^2-1)\ln(2)} = e^{(x^2-1)\ln(2)}(2x\ln(2)) = 2^{x^2-1}(2x\ln(2))$
- 12.3.19. $(d/dx)[(x^2 + 1)^5\ln(x)] = (x^2 + 1)^5/x + 5(x^2 + 1)^4(2x)\ln(x) = (x^2 + 1)^5/x + 10x(x^2 + 1)^4\ln(x)$
- 12.3.27. $(d/dx)\ln|(-2x + 1)(x + 1)| = (d/dx)\{\ln|-2x + 1| + \ln|x + 1|\} = -2/(-2x + 1) + 1/(x + 1)$
- 12.3.31. $(d/dx)\ln|(x + 1)(x - 3)/(-2x - 9)| = (d/dx)\{\ln|x + 1| + \ln|x - 3| - \ln|-2x - 9|\}$
  $= 1/(x + 1) + 1/(x - 3) + 2/(-2x - 9)$

The last example suggests a way to differentiate expressions like $(x^3 + 2x - 3)^2(3x + 5)^4/(2x^2 - 5)^3$
That's a total mess if we use the product and quotient rules. (Try it if you don't believe me!) Instead, we use a technique called "logarithmic differentiation":

$$(d/dx)\ln(f(x)) = f'(x)/f(x) \rightarrow f'(x) = f(x)(d/dx)\ln(f(x))$$

Hence,

$(d/dx)[(x^3 + 2x - 3)^2(3x + 5)^4/(2x^2 - 5)^5]$
$= [(x^3 + 2x - 3)^2(3x + 5)^4/(2x^2 - 5)^5](d/dx)\ln[(x^3 + 2x - 3)^2(3x + 5)^4/(2x^2 - 5)^5]$

$= [(x^3 + 2x - 3)^2(3x + 5)^4/(2x^2 - 5)^5](d/dx)[2\ln(x^3 + 2x - 3) + 4\ln(3x + 5) - 5\ln(2x^2 - 5)]$

$= [(x^3 + 2x - 3)^2(3x + 5)^4/(2x^2 - 5)^5][2(3x^2 + 2)/(x^3 + 2x - 3) + 12/(3x + 5) - 20x/(2x^2 - 5)]$

- **12.3.55.** $(d/dx)(x^2e^{2x-1}) = 2x^2e^{2x-1} + 2xe^{2x-1}$
- **12.3.77.** **\$10,000 at 4%/yr compounded continuously. How fast is it growing after 3 years?**
  $F(t) = 10000e^{0.04t}$, so $F'(t) = 400e^{0.04t}$, and $F'(3) = 400e^{0.12} =$ \$451.00/year
- **12.3.79.** **What if it's compounded semiannually?**
  $F(t) = 10000(1.02)^{2t}$, so $F'(t) = 10000(1.02)^{2t}(2\ln(1.02))$,
  and $F'(3) = 10000(1.02)^6(2\ln(1.02)) =$ \$446.02/year

XII- Applications of Logarithmic & Exponential Functions

- 12.3.89. $N(t) = 39t + 68$ = millions of Chinese cell-phone subscribers, with $t = 0$ in 2000

  Average annual revenue per subscriber in 2000 was $350. Suppose revenue decreases continuously at 10% per year.

  a) Find $R(t)$ = annual revenue in year t: $R(t) = 350(39t + 68)e^{-0.1t}$

  b) Find $R(2)$ and $R'(2)$ to nearest billions

  $R(2) = 350(78 + 68)e^{-0.2} = 41837$, i.e., $42 billion

  $R'(t) = 350[-0.1(39t + 68)e^{-0.1t} + 39e^{-0.1t}] = 350e^{-0.1t}(32.2 - 3.9t)$

  $R'(2) = 350e^{-0.2}(32.2 - 7.8) = 6992$, i.e., revenues are increasing by $7 billion/year

- 12.3.91 $P(t) = 150/(1 + 14999e^{-0.3466t}) = 10^6$ people with flu t weeks after start of an epidemic

  How fast is it growing after 20, 30, and 40 weeks? (Answer to 3 significant figures)

  $P'(t) = 150(14999)(0.3466)e^{-0.3466t}/(1 + 14999e^{-0.3466t})^2$

  $P'(20) = 150(14999)(0.3466)\, e^{-0.3466(20)}/(1 + 14999e^{-0.3466(20)})^2 = 3.11$ million new cases/wk

  $P'(30) = 150(14999)(0.3466)\, e^{-0.3466(30)}/(1 + 14999e^{-0.3466(30)})^2 = 11.2$ million new cases/wk

  $P'(40) = 150(14999)(0.3466)\, e^{-0.3466(40)}/(1 + 14999e^{-0.3466(40)})^2 = 0.722$ million new cases/wk

- 12.4.21. $xe^y - ye^x = 1$. Find dy/dx.

  $xe^y dy/dx + e^y - ye^x - e^x dy/dx = 0$, so $dy/dx = (e^y - ye^x)/(e^x - xe^y)$

- 12.4.29. $\ln(xy + y^2) = e^y$. Find dy/dx.

  $(x dy/dx + y + 2y dy/dx)/(xy + y^2) = e^y dy/dx$, so $dy/dx = y/[(xy + y^2)e^y - x - 2y]$

- 13.1.41. Find local and global exterma of $f(x) = x - \ln(x)$ on $(0, \infty)$

  $f'(x) = 1 - 1/x = 0 \rightarrow x = 1$. $f''(x) = 1/x^2 > 0$ for all $x > 0$.

  Hence, $x = 1$ is a global minimum on $(0, \infty)$

- 13.2.37. Refer to 12.3.89, but now revenue/customer is decreasing at 30% per year.

  Annual revenue is $R(t) = 350(39t + 68)e^{-0.3t}$ $millions

  When will it peak (to the nearest 0.1year) and at what value?

  $R'(t) = 350[(39t + 68)e^{-0.3t}(-0.3) + 39e^{-0.3t}] = 350e^{-0.3t}(-11.7t + 18.6)$

  $R'(t) = 0$ when $-11.7t + 18.6 = 0$, or when $t = 1.6$ years

  $R(1.6) = 350(39(1.6) + 68)e^{-0.3(1.6)} = 28241.3$, i.e., $28,241 million ($28.241 billion)

- 13.2.41. The value t years from now of a classic car collection will be

  $$v(t) = 300000 + 1000t^2 \text{ dollars.}$$

  With general inflation at a continuous 5% per year, the collection's present value is

  $$p(t) = v(t)e^{-0.05t} = (300000 + 1000t^2)e^{-0.05t}$$

  When should you sell to maximize the present value?

$p'(t) = (300000 + 1000t^2)e^{-0.05t}(-0.05) + 2000te^{-0.05t} = 50e^{-0.05t}(-300 + 40t - t^2)$

$p'(t) = 0$ when $-300 + 40t - t^2 = (30 - t)(t - 10) = 0$, or when $t = 10$ or $t = 30$.

Now $p'(t)$ switches from – to + at $t = 10$ and from + to – at $t = 30$, so we have local maxima at $t = 30$ and $t = 0$. (Where did the latter come from?)

$p(0) = \$300,000$ and $p(30) = (300000 + 900000)e^{-1.5} = \$267,756$

Hence they should sell now! (The answer in the back of the book is wrong!)

XIII- Functions of Several Variables, Partial Derivatives

So far we've been looking at functions of a single variable, f(x). But most real-world phenomena are more complex than that, so we want to be able to handle things that depend on several variables. We'll look at:

- modeling
- rates of change ("partial derivatives")
- unconstrained optimization
- constrained optimization

The text (section 16.2) discusses graphs of functions of two variables and has lots of pretty pictures. I can't draw them as nicely, so I'll let you just read that material. If you have any questions about it, please bring them to class.

Examples

- 16.1.39 C(x, y) = 240,000 + 6000x + 4000y = weekly cost in \$ to manufacture x cars and y trucks. Find marginal costs of a car; of a truck

  We're asking how C(x, y) changes as x and y change *individually*. For a linear function, this is easy: \$6000/car and \$4000/truck. These rates of change are clearly related to the derivatives we've been studying.

- 16.1.41 C(v, a) = monthly cost in \$ of v video clips (@ \$0.03/clip) and a audio clips (@ \$0.04/clip), with \$10 set-up cost
  C(v, a) = 10 + 0.03v + 0.04a

- 16.1.45 f(a, c, n) = 3.1a – 0.27c + 0.87n – 36.7 = Fox's prime-time rating, given those of ABC, CBS, and NBC
  a) Which competitor is the most serious threat? As that competitor's rating increases by 1 point, how does Fox's change?
     CBS; decreases by 0.27 points
  b) Who presents the least competition? Why?
     ABC, because Fox goes up 3.1 points for each point ABC increases.
  c) If a = 12.2, c = 11.3, n = 11.3, find f.
     f = 3.1(12.2) – 0.27(11.3) + 0.87(11.3) – 36.7 = 7.9
  d) Find a(f, n, c) = (1/3.1)(f + 0.27c – 0.87n + 36.7) =0.32f + 0.087c – 0.28n + 11.8

- 16.1.51 $U(x, y) = 6x^{0.8}y^{0.2} + x$ = productivity of a newspaper (pages/day) if they use x copies of Macro Publish and y copies of Turbo Publish
  If they use 10 copies of each, how does increasing x by 1 unit affect output?

$U(10, 10) = 6(10^{0.8})(10^{0.2}) + 10 = 70$, $U(11, 10) = 6(11^{0.8})(10^{0.2}) + 11 = 75.75$, so the marginal productivity of another copy of Macro Publish is about 5.75 pages/day

Partial Derivatives – when asking how a function of several variables changes when its variables change, we have to ask about how those variables are changing.

- Is just one changing?
- Are several changing at the same time, and if so, in what direction is the joint change?

The first of these is the simpler question, and let's do it by an example.

How does $f(x, y) = \text{sqrt}(x^2 + y^2)$ change as x changes?

Well, if only x is changing, then y should be treated as a constant, so that just reduces us to taking the derivative of $f(x, y)$ with respect to x. However, to recognize the fact that $f(x, y)$ depends on both variables, we call this derivative the *partial derivative of f(x,y) with respect to x*, and we denote it by $f_x =$

$$\frac{\partial}{\partial x} f(x, y) = \left(\frac{1}{2}\right)\frac{2x}{\sqrt{x^2 + y^2}} = \frac{x}{\sqrt{x^2 + y^2}}.$$

And if only y is changing, we have the *partial derivative of f(x,y) with respect to y*, and we denote it by $f_y = \frac{\partial}{\partial y} f(x, y) = \left(\frac{1}{2}\right)\frac{2y}{\sqrt{x^2 + y^2}} = \frac{y}{\sqrt{x^2 + y^2}}$

So taking partial derivatives involves nothing new. You just treat all the variables but one as constants and take the ordinary derivative with respect to that singled-out variable.

Examples

- 16.3.5 $f(x, y) = 10000 - 40x + 20y + 10xy$
  $f_x = -40 + 10y$, $f_y = 20 + 10x$
- 16.3.9 $f(x, y) = x^2y^3 - x^3y^2 - xy$
  $f_x = 2xy^3 - 3x^2y^2 - y$, $f_y = 3x^2y^2 - 2x^3y - x$
- 16.3.35 $f(x, y, z) = x^{0.1}y^{0.4}z^{0.5}$
  $f_x = 0.1x^{-0.9}y^{0.4}z^{0.5}$, $f_y = 0.4x^{0.1}y^{-0.6}z^{0.5}$, $f_z = 0.5x^{0.1}y^{0.4}z^{-0.5}$
- 16.3.43 $x_3 = 0.66 - 2.2x_1 - 0.02x_2$, with x_1, x_2, and x_3 being domestic market shares for Chrysler, Ford, and GM
  $\partial x_3/\partial x_1 = -2.2$ If C share goes up by 1% (and F share is unchanged), GM share goes down by 2.2%
  To get $\partial x_1/\partial x_3$, you could solve for $x_1 = 0.3 - (0.2/0.22)x_2 - (1/2.2)x_3$. From that, you get $\partial x_1/\partial x_3 = -1/2.2 = 0.455$, meaning that if G share

goes up by 1% (and F share is unchanged), C share goes down by 0.455%.

(Or you could just use the following result, which is pretty obvious when you stop to think about it: $\partial x_1/\partial x_3 = 1/(\partial x_3/\partial x_1)$. Once again, derivatives aren't fractions, but they surely do behave like fractions!)

- 16.3.45 $C(x, y) = 240{,}000 + 6000x + 4000y - 20xy$ = weekly cost in \$ to manufacture x cars and y trucks. At $(x, y) = (10, 20)$, find marginal cost of a car; of a truck.
  $C_x = 6000 - 20y = \$5600/\text{car}$
  $C_y = 4000 - 20x = \$3800/\text{truck}$

- 16.3.49 $z(t, x) = 13000 + 350t + 9900x + 220xt$ = median family incomes in US, with $t = 0$ in 1950 and $x = 0$ or 1 to indicate that the family is black or white.
  a) Estimate black family median income in 1960
  $13000 + 350(10) + 9900(0) + 220(0)(10) = \$16500$
  b) Estimate white family median income in 1960
  $13000 + 350(10) + 9900(1) + 220(1)(10) = \$28600$
  $z_t = 350 + 220x$
  c) How fast is black median income increasing in 1960:
  $z_t = 350 + 220(0) = \$350/\text{year}$
  d) How fast is white median income increasing in 1960:
  $z_t = 350 + 220(1) = \$770/\text{year}$
  e) Is the gap between black and white median incomes increasing or decreasing? by how much?
  Increasing by \$220/year

- Let's revisit 16.1.51.
  $U(x, y) = 6x^{0.8}y^{0.2} + x$ = productivity of a newspaper (pages/day) if they use x copies of Macro Publish and y copies of Turbo Publish
  If they use 10 copies of each, what is the exact marginal productivity of Macro Publish?

  $U_x = 4.8x^{-0.2}y^{0.2} + 1 = 5.8$ pages per day

XIV- Functions of Several Variables – Unconstrained Optimization

Consider the function $z = f(x, y) = x^4 + 8xy^2 + 2y^4$. Its graph is a surface in 3-dimensional space, with x- and y-axes on the floor and the z-axis vertical. Think of the graph as a mountain range. We would like to locate the maxima (mountain tops) and minima (valley bottoms) of the function. How do we find such points?

A bit of reflection suggests that if a point is a relative maximum (mountain top), then in whatever direction we approach it, we have to be climbing. In particular, if we hold x fixed, it has to be a maximum in y, and vice versa. In other words, it must be true that $\partial f/\partial x = \partial f/\partial y = 0$, and the same is true for relative minima. OK, so let's look at those partial derivatives:

$$f_x = 4x^3 + 8y^2 \text{ and } f_y = 16xy + 8y^3 = 8y(2x + y^2)$$

Setting $f_y = 0$, we see that either

- $y = 0$ (in which case $f_x = 0 \rightarrow x = 0$, so (0, 0) is a stationary point), or
- $x = -y^2/2$. Plugging that into $f_x = 0$, we get $0 = 4(-y^2/2)^3 + 8y^2 = -y^6/2 + 8y^2 = (½)y^2(-y^4 + 16)$
  The roots of that equation are $y = 0$ (which we've already examined) and $y = \pm 2$, so $x = -2$. Thus (–2, 2) and (–2, –2) are also stationary points.

What's the nature of these stationary points?

A. (0, 0)
   - Fixing $y = 0$, we see that $f_x = 4x^3$ and this goes from – to + as x crosses 0. Thus the origin is a minimum in x.
   - Fixing $x = 0$, we see that $f_y = 8y^3$, and the same argument says it's a minimum in y.
   - So does it follow that the origin is a local minimum? Well, remember it has to be a valley in <u>*all*</u> directions. Off the top of my head, let's look at the line $x = y = t$, where we now have

     $$h(t) = f(t, t) = 3t^4 + 8t^3, \text{ with } h'(t) = 12t^3 + 24t^2 = 12t^2(t + 2)$$

     This is > 0 on both sides of $t = 0$, so the origin is <u>*not*</u> a local minimum along the line $x = y$.

B. (–2, ±2)
   - Fixing $y = \pm 2$, we see that $f_x = 4x^3 + 16$ and this goes from – to + as x crosses –2. Thus both (–2, –2) and (–2, 2) are minima in x.
   - Fixing $x = -2$, we see that $f_y = 8y(-4 + y^2)$, and this goes from – to + as y crosses –2 or +2. Thus both (–2, –2) and (–2, 2) are minima in y.
   - So does it follow that (–2, –2) and (–2, 2) are local minimum? By now you're getting suspicious, so it's time to look at some pictures! (16.4.19.xls)

The pictures verify what we said about the origin, and they seem to indicate pretty strongly that the other two stationary points are indeed minima, but can we be sure that there isn't some direction floating around along which our points aren't minima? There is a *second* *derivative test* that can often be helpful. For a function of 2 variables there are 4 different second partial derivatives:

- $f_{xx} = \partial^2 f/(\partial x^2) = \partial f_x/\partial x$ $\quad$ $\partial(4x^3 + 8y^2)/\partial x = 12x^2$ in our example
- $f_{yy} = \partial^2 f/(\partial y^2) = \partial f_y/\partial y$ $\quad$ $\partial(16xy + 8y^3)/\partial y = 16x + 24y^2$ in our example
- $f_{xy} = \dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial x}\right) = \partial f_x/\partial y$ $\quad$ $\partial(4x^3 + 8y^2)/\partial y = 16y$ in our example
- $f_{yx} = \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right) = \partial f_y/\partial x$ $\quad$ $\partial(16xy + 8y^3)/\partial x = 16y$ in our example

The first two are called *pure* 2^{nd} partial derivatives, and the last two are called *mixed* 2^{nd} partial derivatives. For the last two, the notation is potentially confusing when you try to keep straight the order of the differentiation. But note the fortunate coincidence that $f_{xy} = f_{yx}$ in our example, so the order doesn't matter! In fact, this is not a coincidence. For every function we'll encounter, the two mixed 2^{nd} partial derivatives will always be equal. (In fact, one has to work hard to construct weird functions for which the mixed 2^{nd} partials are not equal.)

OK, so how do we distinguish maxima, minima, and whatever other weird kinds of stationary points may be out there? Remember the 2^{nd} derivative test for functions of a single variable, say $g(t)$?

- $g' = 0$ and $g'' > 0$ → local minimum
- $g' = 0$ and $g'' < 0$ → local maximum
- $g' = 0$ and $g'' = 0$ → all bets are off

We have similar results for f_{xx} and f_{yy} to determine local maxima and minima with respect to each variable individually. But our example showed us we had to worry about other directions. The mixed partials help us out:

Second Derivative Test for Optima

1. Find a stationary point, where $f_x = f_y = 0$
2. Compute $H = f_{xx}f_{yy} - (f_{xy})^2$
   a. If $H = 0$, the test fails
   b. If $H < 0$, the point is a *saddle point* (i.e., a max in some directions and a min in others)
   c. If $H > 0$ and
      i. $f_{xx} > 0$, the point is a local minimum
      ii. $f_{xx} < 0$, the point is a local maximum

Let's apply this to our example:

A. (0, 0): $f_{xx} = f_{yy} = f_{xy} = 0$, so $H = 0$ and the test fails.
B. (–2, ±2): $f_{xx} = 12x^2 = 48$, $f_{yy} = 16x + 24y^2 = 64$, $f_{xy} = 16y = \pm 32$

Hence, $H = 48(64) - (\pm 32)^2 = 2048 > 0$, so both points are local minima as our pictures suggested.

Examples

- 16.4.23 Find and characterize the stationary point(s) of $f(x, y) = e^{-(x^2+y^2+2x)}$
  $f_x = -(2x + 2)f(x, y) = 0 \rightarrow x = -1$ $f_y = -2yf(x, y) = 0 \rightarrow y = 0$
  Hence $(x, y) = (-1, 0)$ is the only stationary point.

  $f_{xx} = [(2x + 2)^2 - 2]f(x, y)$ $f_{yy} = (4y^2 - 2)f(x, y)$
  $f_{xy} = 2y(2x + 2)f(x, y) = 4y(x + 1)f(x, y)$
  At $(x, y) = (-1, 0)$, $f(x, y)= e$, $f_{xx} = -2e < 0$, $f_{yy} = -2e$, $f_{xy} = 0$, so $H = 2e(2e) - 0^2 = 4e^2 > 0$.

  Hence $(-1, 0)$ is a local (and global) maximum.
  (Global because $x^2 + y^2 + 2x = (x + 1)^2 + y^2 - 1$, so $f(x, y) = e^{-[(x+1)^2+y^2-1]}$.)

- 16.4.27 Likewise for $g(x, y) = x^2 + y^2 + 2/xy$
  $g_x = 2x - 2/x^2y$ $g_y = 2y - 2/xy^2$
  Setting $g_x = 0$ and solving for y gives $y = 1/x^3$.
  Plugging that into $g_y = 0$ gives $2/x^3 - 2x^5 = 0$, or $x^8 = 1 \rightarrow x = \pm 1$
  Hence there are two stationary points, $(1, 1)$ and $(-1, -1)$.

  $g_{xx} = 2 + 4/x^3y = 6 > 0$ at both points $g_{yy} = 2 + 4/xy^3 = 6$ at both points
  $g_{xy} = 2/x^2y^2 = 2$ at both points $H = 6(6) - 2^2 = 32 > 0$ at both points
  Hence, both stationary points are local minima, but not global minima.
  (Not global because $g(x, -x) = 2(x^2 - 1/x^2)$ which $\rightarrow -\infty$ as $x \rightarrow 0$.)

- 16.4.35 (modified a bit) $C(s, p)$ = daily cost (\$) of removing s pounds of sulfur and p pounds of lead from a firm's smokestack gases
  $= 4000 + 100s^2 + 50p^2 - 100sp$
  To help mitigate this expense, the government gives subsidies of \$500 (\$100) per pound of sulfur (lead) removed. How much of the pollutants should be removed to minimize net costs, $N(s, p) = 4000 + 100s^2 + 50p^2 - 100sp - 500s - 100p$?

  $N_s = 200s - 100p - 500$ $N_p = 100p - 100s - 100$
  Setting both to 0 and solving the simultaneous linear equations yields $(s, p) = (6, 7)$.

  $N_{ss} = 200 > 0$, $N_{pp} = 100$, $N_{sp} = -100$, $H = 200(100) - 100^2 > 0 \rightarrow$ we have found a minimum
  They minimize their daily net cost by removing 6 pounds of sulfur and 7 pounds of lead, giving a daily net cost = 4000 + 3600 + 2450 – 4200 – 3000 – 700 = \$2150

XV- Functions of Several Variables – Constrained Optimization

Examples

- 16.4.39 For a checked bag, $L + W + H \le 62$ inches. Find the largest acceptable volume.

  We want to maximize $f(L, W, H) = LWH$, subject to the constraint $L + W + H = 62$. (With implicit additional constraints $L, W, H \ge 0$. Why is the other constraint an equation?)

  Perhaps the easiest way to handle the constraint is to use it to eliminate one variable and then do unconstrained optimization on the remaining two variables:

  $L = 62 - W - H$, so maximize $g(W, H) = (62 - W - H)WH = 62WH - W^2H - WH^2$

  $g_W = 62H - 2WH - H^2 = H(62 - 2W - H)$
  $g_H = 62W - W^2 - 2WH = W(62 - W - 2H)$
  Setting $g_W = 0$, we see that either $H = 0$ (which is clearly a minimum volume bag!), or else $62 - 2W - H = 0$. Similarly, setting $g_H = 0$, we see that either $W = 0$ (which is clearly a minimum volume bag!), or else $62 - W - 2H = 0$. Hence we must solve the two equations $62 - 2W - H = 0$ and $62 - W - 2H = 0$
  The solution to this system is $W = H = 62/3$ (→ $L = 62/3$).

  Is this a maximum? $g_{WW} = -2H$, $g_{HH} = -2W$, $g_{WH} = 62 - 2W - 2H$
  At our stationary point, $g_{WW} = -124/3 < 0$, $g_{HH} = -124/3$, $g_{WH} = -62/3$, so

  $$g_{WW}\, g_{HH} - (g_{WH})^2 = (124/3)^2 - (62/3)^2 > 0$$

  and we do indeed have a maximum! Volume $= (62/3)^3 \approx 8827 \text{ in}^3 \approx 5.1 \text{ ft}^3$

- 16.4.41 For a box shipped via UPS $L + 2W + 2H \le 108$ inches. Find the largest acceptable volume. We want to maximize $f(L, W, H) = LWH$, subject to the constraint $L + 2W + 2H = 108$. (With implicit additional constraints $L, W, H \ge 0$. Why is the other constraint an equation?)

  Once again, we eliminate $L = 108 - 2W - 2H$, and maximize

  $g(W, H) = (108 - 2W - 2H)WH = 108WH - 2W^2H - 2WH^2 = 54WH - W^2H - WH^2$

  Comparing this with 16.4.39, we see (without doing any of the work!) that $W = H = 54/3 = 18$ (→ $L = 36$) yields the maximum volume of $36(18^2) = 11{,}664 \text{ in}^3 = 6.75 \text{ ft}^3$

Lagrange multipliers – to optimize $f(x, y)$ subject to the constraint $g(x, y) = b$, one can proceed formally by introducing a new variable, λ (called a Lagrange multiplier, after the great French mathematician, Joseph Louis Lagrange (1736-1813), whom Frederick the Great called "the greatest mathematician of Europe [of the 18th century]"). We form the "Lagrangian function"

$$\ell(x, y, \lambda) = f(x, y) - \lambda[g(x, y) - b]$$

and then take the partials of $\ell(x, y, \lambda)$ with respect to x, y, and λ. We set all 3 of them to 0, and solve the resulting system of equations. With more than two variables, we proceed in similar fashion. The biggest problem with the method is that we have no simple way to distinguish between maxima, minima, and saddlepoints. However, in many problems, the physical aspects of the system help us decide that.

Example: 16.5.31 (slightly generalized)

The top and bottom of a closed box cost \$0.20/ft^2, but the sides cost only \$0.10/ft^2. What are the dimensions of a least cost box with volume of V ft^3? We want to minimize $20(2)LW + 10(2LH + 2WH) = 20(2LW + LH + WH)$, subject to the constraint $LWH = V$. To make life easier, we'll drop the factor of 20, since it affects only the objective value, but not the location of the optimal point(s).

<u>Solution with a Lagrange Multiplier</u>

1. Form $\ell(L, W, H, \lambda) = 2LW + LH + HW - \lambda(LWH - V)$

2. Take all the partials and set them to 0:

$$\ell_L = 2W + H - \lambda WH = 0 \qquad (1)$$
$$\ell_W = 2L + H - \lambda LH = 0 \qquad (2)$$
$$\ell_H = L + W - \lambda LW = 0 \qquad (3)$$
$$\ell_\lambda = LWH - V = 0 \qquad (4)$$

3. Now solve the resulting system. This may require some ingenuity to avoid getting all twisted up!

Multiply (1) by L: $2WL + HL - \lambda LWH = 0$ (1')

Multiply (2) by W: $2WL + WH - \lambda LWH = 0$ (2')

Subtract (1') from (2'): $H(W - L) = 0 \rightarrow H = 0$ (which is impossible if the box is to have volume of 2 cubic feet) or $W = L$.

Plug $W = L$ into (3): $0 = 2L - \lambda L^2 = L(2 - \lambda L) \rightarrow L = 0$ (again impossible) or $\lambda = 2/L$.

Plug $\lambda = 2/L$ into (2'): $2L + H - 2H = 0 \rightarrow H = 2L$

Plug $W = L$ and $H = 2L$ into (4): $2L^3 - V = 0 \rightarrow L = \sqrt[3]{V/2}$, so $W = L = \sqrt[3]{V/2}$, and $H = 2L = \sqrt[3]{4V}$. While we're at it, $\lambda = 2/L = 2\sqrt[3]{2/V}$.

Thus the box should be $\sqrt[3]{V/2}$ foot square at the base and twice as high. Its cost (in "20-cent pieces) is:

$$C(L, W, H; V) = 2LW + LH + HW = 3\sqrt[3]{2V^2} = 3\sqrt[3]{2}\,V^{2/3}$$

Notice that $\frac{\partial C}{\partial V} = 3(\sqrt[3]{2})(2/3)V^{-1/3} = \lambda$. This is not a coincidence. The value of the Lagrange multiplier is always the derivative of the objective function with respect to the

constraint right-hand-side. Does this remind you of anything? Right – shadow prices in linear programming.

Solution by Elimination

0. **Minimize 2LW + LH + HW, subject to LWH = V**

1. **Solve for H = V/LW, and plug into objective function: minimize 2LW + V/W + V/L = f(L, W).**
2. **Take both partials and set them to 0:**
   $f_L = 2W - V/L^2 = 0 \rightarrow W = V/2L^2$
   $f_W = 2L - V/W^2 = 0$, **and plugging** $W = V/2L^2$, **yields** $2L - 4L^4/V = 2L(1 - 2L^3/V) = 0 \rightarrow$ $L = 0$ **(which we've already dismissed) or** $L = \sqrt[3]{V/2}$, **so** $W = V/2L^2 = \sqrt[3]{V/2}$, **and** $H = V/LW = \sqrt[3]{4V}$.
3. **Looking at the 2nd-order partials:**
   $f_{LL} = 2V/L^3 = 4$ **when** $L = \sqrt[3]{V/2}$ $f_{WW} = 2V/W^3 = 4$ **when** $W = \sqrt[3]{V/2}$ $f_{LW} = 2$
   $f_{LL}f_{WW} - (f_{LW})^2 = 16 - 4 > 0$ **at** $(L, W) = (\sqrt[3]{V/2}, \sqrt[3]{V/2})$, **so we do indeed have a minimum.**