Finding Maxima and Minima

When you were learning about derivatives about functions of one variable, you learned some techniques for finding the maximum and minimum values of functions of one variable. We'll now extend those techniques to functions of more than one variable. We'll concentrate on functions of two variables, though many of the techniques work more generally.

Local Maxima and Minima

One of the first things you did when you were developing the techniques used to find the maximum and minimum values of $f(x)$ was you asked yourself

> Suppose that the largest (or smallest) value of $f(x)$ is $f(a)$. What does that tell us about a?

After a little thought you answered

> If the largest (or smallest) value of $f(x)$ is $f(a)$ and f is differentiable at a, then $f'(a) = 0$.

Let's recall what that's true. Suppose that the largest value of $f(x)$ is $f(a)$. Then for all $h > 0$,

$$f(a+h) \geq f(a) \implies f(a+h) - f(a) \geq 0 \implies \frac{f(a+h) - f(a)}{h} \geq 0 \quad \text{if } h > 0$$

Taking the limit $h \to 0$ tells us that $f'(a) \geq 0$. Similarly, for all $h < 0$,

$$f(a+h) \geq f(a) \implies f(a+h) - f(a) \geq 0 \implies \frac{f(a+h) - f(a)}{h} \leq 0 \quad \text{if } h < 0$$

Taking the limit $h \to 0$ now tells us that $f'(a) \leq 0$. So we have both $f'(a) \geq 0$ and $f'(a) \leq 0$ which forces $f'(a) = 0$. You also observed at the time that for this argument to work, you only need $f(x) \leq f(a)$ for all x's close to a, not necessarily for all x's in the whole world. (In the above inequalities, we only used $f(a+h)$ with h small.) So you said

> If $f(a)$ is a local maximum or minimum for $f(x)$ and f is differentiable at a, then $f'(a) = 0$.

Exactly the same discussion applies to functions of more than one variable. Here are the corresponding definitions and statements.

Definition 1 (Local Max and Min).

The point (a, b) is a local maximum of the function $f(x, y)$ if there is an $r > 0$ such that $f(x, y) \leq f(a, b)$ for all points (x, y) within a distance r of (a, b).

Similarly, (a, b) is a local minimum of the function $f(x, y)$ if there is an $r > 0$ such that $f(x, y) \geq f(a, b)$ for all points (x, y) within a distance r of (a, b).

Local maximum and minimum values are also called extremal values.

© Joel Feldman. 2014. All rights reserved.

Definition 2 (Critical Point).

The point (a, b) is a critical point of the function $f(x, y)$

- if $\frac{\partial f}{\partial x}(a, b) = \frac{\partial f}{\partial y}(a, b) = 0$
- or if at least one of the derivatives $\frac{\partial f}{\partial x}(a, b)$, $\frac{\partial f}{\partial y}(a, b)$ does not exist.

Theorem 3.

If the function $f(x, y)$ has local maximum or minimum at (a, b) and the partial derivatives $\frac{\partial f}{\partial x}(a, b)$, $\frac{\partial f}{\partial y}(a, b)$ exist, then

$$\frac{\partial f}{\partial x}(a, b) = \frac{\partial f}{\partial y}(a, b) = 0$$

Proof. It easy to see that this theorem follows from what we already know about functions of one variable. Suppose that $f(x, y)$ has a local maximum or minimum at (a, b). Define the single variable functions

$$F(x) = f(x, b) \qquad G(y) = f(a, y)$$

Then a is a local maximum or minimum for the function $F(x)$, so that $F'(a) = 0$, and b is a local maximum or minimum for the function $G(y)$, so that $G'(b) = 0$. Now we just have to observe that $F'(x)$ is the rate of change of $F(x) = f(x, y)\big|_{y=b}$ with respect to x when y is held fixed at $y = b$, which is exactly $\frac{\partial f}{\partial x}(x, b)$. Similarly $G'(y)$ is the rate of change of $G(y) = f(x, y)\big|_{x=a}$ with respect to y when x is held fixed at $x = a$, which is exactly $\frac{\partial f}{\partial y}(a, y)$. Thus

$$\frac{\partial f}{\partial x}(a, b) = F'(a) = 0 \qquad \frac{\partial f}{\partial y}(a, b) = G'(b) = 0$$

$\square$

Theorem 3 tells us that every local maximum or minimum is a critical point. Beware that it does *not* tell us that every critical point is either a local maximum or a local minimum. In fact, we shall see later, in Example 10, a critical point that is neither a local maximum nor a local minimum. None–the–less, Theorem 3 is very useful because often functions have only a small number of critical points. To find local maxima and minima of such functions, we only need to consider its critical points. We'll return later to the question of how to tell if a critical point is a local maximum, local minimum or neither. For now, we'll just practice finding critical points.

© Joel Feldman. 2014. All rights reserved.

Example 4 ($f(x,y) = x^2 - 2xy + 2y^2 + 2x - 6y + 12$)

Find all critical points of $f(x,y) = x^2 - 2xy + 2y^2 + 2x - 6y + 12$.

Solution. As a preliminary calculation, we find the two first order partial derivatives of $f(x,y)$.

$$\begin{aligned} f_x(x,y) &= 2x - 2y + 2 \\ f_y(x,y) &= -2x + 4y - 6 \end{aligned}$$

So the critical points are the solutions of the pair of equations $2x-2y+2=0$, $-2x+4y-6$, or equivalently (dividing by two and moving the constants to the right hand side)

$$x - y = -1 \tag{1a}$$

$$-x + 2y = 3 \tag{1b}$$

One strategy for solving a system of two equations in two unknowns (x and y) like this is to

- First use one of the equations to solve for one of the unkowns in terms of the other unknown. For example (1a) tells us that $y = x+1$. This expresses y in terms of x. We say that we have solved for y in terms of x.
- Then substitute the result, $y = x+1$ in our case, into the other equation, (1b). In our case, this gives

  $$-x + 2(x+1) = 3 \iff x + 2 = 3 \iff x = 1$$

- We have now found that $x = 1$, $y = x+1 = 2$ is the only solution. So the only critical point is $(1,2)$.

An alternative strategy for solving a system of two equations in two unknowns like (1) is to

- add equations (1a) and (1b) together. This gives

  $$(1a) + (1b): \quad (1-1)x + (-1+2)y = -1 + 3 \iff y = 2$$

  The point here is that adding equations (1a) and (1b) together eliminates the unknown x, leaving us with one equation in the unknown y, which is easily solved. For other systems of equations you might have multiply the equations by some numbers before adding them together.

- We now know that $y = 2$. Substituting it into (1a) gives us

  $$x - 2 = -1 \implies x = 1$$

- Once again we have found that the only critical point is $(1,2)$.

Example 4

© Joel Feldman. 2014. All rights reserved.

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

Find all critical points of $f(x,y) = 2x^3 - 6xy + y^2 + 4y$.

Solution. The first order partial derivatives are

$$f_x = 6x^2 - 6y \qquad f_y = -6x + 2y + 4$$

So the critical points are the solutions of

$$6x^2 - 6y = 0 \qquad -6x + 2y + 4 = 0$$

We can rewrite the first equation as $y = x^2$, which expresses y as a function of x. We can then substitute $y = x^2$ into the second equation, giving

$$\begin{aligned} -6x + 2y + 4 = 0 &\iff -6x + 2x^2 + 4 = 0 \iff x^2 - 3x + 2 = 0 \iff (x-1)(x-2) = 0 \\ &\iff x = 1 \text{ or } 2 \end{aligned}$$

When $x = 1$, $y = 1^2 = 1$ and when $x = 2$, $y = 2^2 = 4$. So, there are two critical points: $(1,1)$, $(2,4)$.

Example 5

Example 6 ($f(x,y) = xy(5x + y - 15)$)

Find all critical points of $f(x,y) = xy(5x + y - 15)$.

Solution. The first order partial derivatives of $f(x,y) = xy(5x + y - 15)$ are

$$\begin{aligned} f_x(x,y) &= y(5x + y - 15) + xy(5) = y(5x + y - 15) + y(5x) = y(10x + y - 15) \\ f_y(x,y) &= x(5x + y - 15) + xy(1) = x(5x + y - 15) + x(y) = x(5x + 2y - 15) \end{aligned}$$

The critical points are the solutions of $f_x(x,y) = f_y(x,y) = 0$ or

$$y(10x + y - 15) = 0 \qquad \text{and} \qquad x(5x + 2y - 15) = 0 \tag{2}$$

The first equation, $y(10x+y-15) = 0$, is satisfied if either of the two factors y, $(10x+y-15)$ is zero. So the first equation is satisfied if either of the two equations

$$y = 0 \tag{3a}$$

$$10x + y = 15 \tag{3b}$$

is satisfied. The second equation, $x(5x + 2y - 15) = 0$, is satisfied if either of the two factors x, $(5x + 2y - 15)$ is zero. So the first equation is satisfied if either of the two equations

$$x = 0 \tag{4a}$$

$$5x + 2y = 15 \tag{4b}$$

is satisfied. So both critical point equations (2) are satisfied if one of (3a), (3b) is satisfied and in addition one of (4a), (4b) is satisfied. There are four possibilities:

© Joel Feldman. 2014. All rights reserved.

- (3a) and (4a) are satisfied if and only if $x = y = 0$
- (3a) and (4b) are satisfied if and only if $y = 0,\ 5x + 2y = 15 \iff y = 0,\ 3x = 15$
- (3b) and (4a) are satisfied if and only if $10x + y = 15,\ x = 0 \iff y = 15,\ x = 0$
- (3b) and (4b) are satisfied if and only if $10x + y = 15,\ 5x + 2y = 15$. We can use, for example, the second of these equations to solve for x in terms of y: $x = \frac{1}{5}(15 - 2y)$. When we substitute this into the first equation we get $2(15 - 2y) + y = 15$, which we can solve for y. This gives $-3y = 15 - 30$ or $y = 5$ and then $x = \frac{1}{5}(15 - 2 \times 5) = 1$.

In conclusion, the critical points are $(0, 0)$, $(3, 0)$, $(0, 15)$ and $(1, 5)$.

A more compact way to write what we have just done is

$$\begin{aligned}
& & f_x(x, y) = 0 \qquad & \text{and} & f_y(x, y) = 0 \\
&\iff & y(10x + y - 15) = 0 \qquad & \text{and} & x(5x + 2y - 15) = 0 \\
&\iff & \big\{y = 0 \text{ or } 10x + y = 15\big\} \qquad & \text{and} & \big\{x = 0 \text{ or } 5x + 2y = 15\big\} \\
&\iff & \big\{x = y = 0\big\} \text{ or } \big\{y = 0,\ x = 3\big\} & \text{ or } \big\{x = 0,\ y = 15\big\} & \text{ or } \big\{x = 1,\ y = 5\big\}
\end{aligned}$$

Example 6

Example 7

In a certain community, there are two breweries in competition, so that sales of each negatively affect the profits of the other. If brewery A produces x litres of beer per month and brewery B produces y litres per month, then the profits of the two breweries are given by

$$P = 2x - \frac{2x^2 + y^2}{10^6} \qquad Q = 2y - \frac{4y^2 + x^2}{2 \times 10^6}$$

respectively. Find the sum of the two profits if each brewery independently sets its own production level to maximize its own profit and assumes that its competitor does likewise. Find the sum of the two profits if the two breweries cooperate so as to maximize that sum.

Solution. If A adjusts x to maximize P (for y held fixed) and B adjusts y to maximize Q (for x held fixed) then x and y are determined by

$$\begin{aligned}
P_x = 2 - \tfrac{4x}{10^6} = 0 \quad &\implies & x &= \tfrac{1}{2}10^6 \\
Q_y = 2 - \tfrac{8y}{2\times 10^6} = 0 \quad &\implies & y &= \tfrac{1}{2}10^6 \\
&\implies & P + Q &= 2(x + y) - \tfrac{1}{10^6}\big(\tfrac{5}{2}x^2 + 3y^2\big) \\
&\implies & &= 10^6\big(1 + 1 - \tfrac{5}{8} - \tfrac{3}{4}\big) = \tfrac{5}{8}10^6
\end{aligned}$$

On the other hand if (A, B) adjust (x, y) to maximize $P + Q = 2(x + y) - \frac{1}{10^6}\big(\frac{5}{2}x^2 + 3y^2\big)$, then x and y are determined by

© Joel Feldman. 2014. All rights reserved.

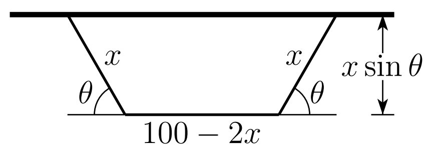
$$
\begin{aligned}
(P+Q)_x &= 2 - \tfrac{5x}{10^6} = 0 &\implies&& x &= \tfrac{2}{5}10^6\\
(P+Q)_y &= 2 - \tfrac{6y}{10^6} = 0 &\implies&& y &= \tfrac{1}{3}10^6\\
&&\implies&& P+Q &= 2(x+y) - \tfrac{1}{10^6}\big(\tfrac{5}{2}x^2 + 3y^2\big)\\
&&\implies&& &= 10^6\big(\tfrac{4}{5} + \tfrac{2}{3} - \tfrac{2}{5} - \tfrac{1}{3}\big) = \tfrac{11}{15}10^6
\end{aligned}
$$

Example 7

Example 8

Equal angle bends are made at equal distances from the two ends of a 100 metre long fence so the resulting three segment fence can be placed along an existing wall to make an enclosure of trapezoidal shape. What is the largest possible area for such an enclosure?

Solution. Here is a figure of the fence.



The area that it encloses is

$$
\begin{aligned}
A(x,\theta) &= (100-2x)x\sin\theta + 2\cdot\tfrac{1}{2}\cdot x\sin\theta\cdot x\cos\theta\\
&= (100x - 2x^2)\sin\theta + \tfrac{1}{2}x^2\sin(2\theta)
\end{aligned}
$$

The maximize the area, we need to solve

$$
\begin{aligned}
0 = A_x &= (100-4x)\sin\theta + x\sin(2\theta) &\implies&& (100-4x) + 2x\cos\theta &= 0\\
0 = A_\theta &= (100x-2x^2)\cos\theta + x^2\cos(2\theta) &\implies&& (100-2x)\cos\theta + x\cos(2\theta) &= 0
\end{aligned}
$$

Here we have used that the fence of maximum area cannot have $\sin\theta = 0$ or $x = 0$. The first equation forces $\cos\theta = -\frac{100-4x}{2x}$ and hence $\cos(2\theta) = 2\cos^2\theta - 1 = \frac{(100-4x)^2}{2x^2} - 1$. Substituting these into the second equation gives

$$
\begin{aligned}
&& -(100-2x)\frac{100-4x}{2x} + x\Big[\frac{(100-4x)^2}{2x^2} - 1\Big] &= 0\\
&\implies& -(100-2x)(100-4x) + (100-4x)^2 - 2x^2 &= 0\\
&\implies& 6x^2 - 200x &= 0\\
&\implies& x = \frac{100}{3} \quad \cos\theta = -\frac{-100/3}{200/3} = \frac{1}{2} \quad \theta &= 60^\circ
\end{aligned}
$$

© Joel Feldman. 2014. All rights reserved.

and the maximum area enclosed is

$$A = \Big(100\frac{100}{3} - 2\frac{100^2}{3^2}\Big)\frac{\sqrt{3}}{2} \;+\; \frac{1}{2}\frac{100^2}{3^2}\frac{\sqrt{3}}{2} \;=\; \frac{2500}{\sqrt{3}}$$

Example 8

Example 9

An experiment yields data points (x_i, y_i), $i = 1, 2, \cdots, n$. We wish to find the straight line $y = mx + b$ which "best" fits the data. The definition of "best" is "minimizes the root mean square error", i.e. minimizes

$$E(m, b) = \sum_{i=1}^{n} (mx_i + b - y_i)^2$$

Find m and b.

Solution. We wish to choose m and b so that

$$\begin{aligned} 0 = \frac{\partial E}{\partial m} &= \sum_{i=1}^{n} 2(mx_i + b - y_i)x_i = m\Big[\sum_{i=1}^{n} 2x_i^2\Big] + b\Big[\sum_{i=1}^{n} 2x_i\Big] - \Big[\sum_{i=1}^{n} 2x_iy_i\Big] \\ 0 = \frac{\partial E}{\partial b} &= \sum_{i=1}^{n} 2(mx_i + b - y_i) \quad = m\Big[\sum_{i=1}^{n} 2x_i\Big] + b\Big[\sum_{i=1}^{n} 2\Big] - \Big[\sum_{i=1}^{n} 2y_i\Big] \end{aligned}$$

There are a lot of symbols here. But remember that all of the x_i's and y_i's are given constants. The only unknowns are m and b. To emphasize this, and to save some writing, define the constants

$$S_x = \sum_{i=1}^{n} x_i \qquad S_y = \sum_{i=1}^{n} y_i \qquad S_{x^2} = \sum_{i=1}^{n} x_i^2 \qquad S_{xy} = \sum_{i=1}^{n} x_iy_i$$

The equations are (after dividing by two)

$$S_{x^2}\, m + S_x\, b = S_{xy} \tag{5a}$$

$$S_x\, m + n\, b = S_y \tag{5b}$$

These are two linear equations on the unknowns m and b. They may be solved in any of the usual ways. One is to use (5b) to solve for b in terms of m

$$b = \frac{1}{n}\big(S_y - S_x m\big) \tag{6}$$

and then substitute this into (5a) to get the equation

$$S_{x^2}\, m + \frac{1}{n} S_x \big(S_y - S_x m\big) = S_{xy} \qquad \Longrightarrow \qquad \big(nS_{x^2} - S_x^2\big)m = nS_{xy} - S_xS_y$$

for m. We can then solve this equation for m and substitute back into (6) to get b. This gives

$$m = \frac{nS_{xy} - S_xS_y}{nS_{x^2} - S_x^2} \qquad b = -\frac{S_xS_{xy} - S_yS_{x^2}}{nS_{x^2} - S_x^2}$$

© Joel Feldman. 2014. All rights reserved.

Another way to solve the equations is

$$\begin{aligned} n(5\text{a}) - S_x(5\text{b}) :\quad & \Big[nS_{x^2} - S_x^2\Big] m = nS_{xy} - S_x S_y \\ -S_x(5\text{a}) + S_{x^2}(5\text{b}) :\quad & \Big[nS_{x^2} - S_x^2\Big] b \ = -S_x S_{xy} + S_y S_{x^2} \end{aligned}$$

which gives the same solution.

Example 9

The Second Derivative Test

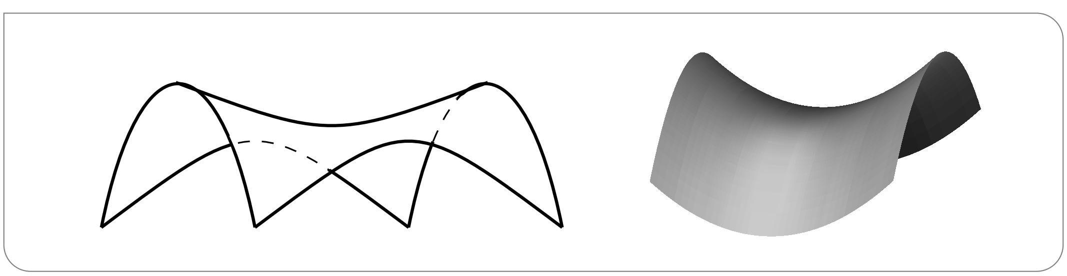
Now let's start thinking about how to tell if a critical point is a local minimum or maximum. First here is an example which shows that sometimes critical points are neither local minima or maxima.

Example 10 ($f(x, y) = x^2 - y^2$)

The first partial derivatives of $f(x, y) = x^2 - y^2$ are $f_x(x, y) = 2x$ and $f_y(x, y) = -2y$. So the only critical point of this function is $(0, 0)$. Is this a local minimum or maximum? Well let's start with (x, y) at $(0, 0)$ and then move (x, y) away from $(0, 0)$ and see if $f(x, y)$ gets bigger or smaller. At the origin $f(0, 0) = 0$. Of course we can move (x, y) away from $(0, 0)$ in many different directions.

- Let's start by moving (x, y) along the x–axis. Then $(x, y) = (x, 0)$ and $f(x, y) = f(x, 0) = x^2$. So when we start with $x = 0$ and then increase x, the value of the function f increases — which means that $(0, 0)$ cannot be a local maximum for f.
- Now let's move (x, y) away from $(0, 0)$ along the y–axis. Then $(x, y) = (0, y)$ and $f(x, y) = f(0, y) = -y^2$. So when we start with $y = 0$ and then increase y, the value of the function f decreases — which means that $(0, 0)$ cannot be a local minimum for f.

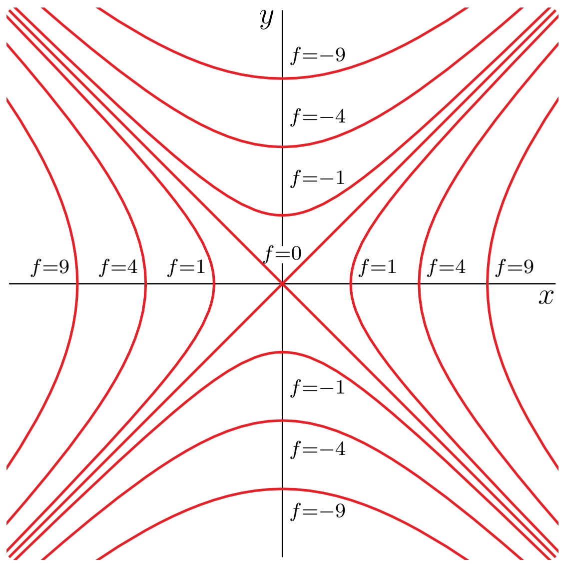
So $(0, 0)$ is neither a local minimum or maximum for f. It is called a saddle point, because the graph of f looks like a saddle. (The full definition of "saddle point" is given immediately after this example.) Here are some figures showing the graph of f.



and the level curves of f. Observe from the level curves that

- f increases as you leave $(0, 0)$ walking along the x axis
- f decreases as you leave $(0, 0)$ walking along the y axis

© Joel Feldman. 2014. All rights reserved.



Example 10

Definition 11.

The point (a,b) is called a saddle point for the function $f(x,y)$ if, for each $r>0$,

- there is at least one point (x,y), within a distance r of (a,b), for which $f(x,y)>f(a,b)$ and
- there is at least one point (x,y), within a distance r of (a,b), for which $f(x,y)<f(a,b)$.

So how do you tell if a critical point is a local minimum, local maximum or saddle point? Well let's remember what happens for functions of one variable. Suppose that a is a critical point of the function $f(x)$. Any (sufficiently smooth) function is well approximated, when x is close to a, by

$$f(x) = f(a) + f'(a)\,(x-a) + \tfrac{1}{2}f''(a)\,(x-a)^2 + \tfrac{1}{3!}f^{(3)}(a)\,(x-a)^3 + \cdots$$

As a is a critical point, $f'(a)=0$ and

$$f(x) = f(a) + \tfrac{1}{2}f''(a)\,(x-a)^2 + \tfrac{1}{3!}f^{(3)}(a)\,(x-a)^3 + \cdots$$

If $f''(a)\neq 0$, $f(x)$ is going to look a lot like $f(a)+\tfrac{1}{2}f''(a)\,(x-a)^2$ when x is really close to a. In particular

- if $f''(a)>0$, then we will have $f(x)>f(a)$ when x is close to (but not equal to) a, so that a will be a local minimum and

© Joel Feldman. 2014. All rights reserved.

- if $f''(a) < 0$, then we will have $f(x) < f(a)$ when x is close to (but not equal to) a, so that a will be a local maximum, but
- if $f''(a) = 0$, then we cannot draw any conclusions without more work.

A similiar, but messier, analysis is possible for functions of two variables. Define

$$D(x,y) = f_{xx}(x,y)\, f_{yy}(x,y) - f_{xy}(x,y)^2$$

It is called the discriminant of f. Then the second derivative test for functions of two variables is

Theorem 12 (Second Derivative Test)**.**

Let $r > 0$ and assume that all second order derivatives of the function $f(x,y)$ are continuous at all points (x,y) that are within a distance r of (a,b). Assume that $f_x(a,b) = f_y(a,b) = 0$. Then

- if $D(a,b) > 0$ and $f_{xx}(a,b) > 0$, then $f(x,y)$ has a local minimum at (a,b),
- if $D(a,b) > 0$ and $f_{xx}(a,b) < 0$, then $f(x,y)$ has a local maximum at (a,b),
- if $D(a,b) < 0$, then $f(x,y)$ has a saddle point at (a,b), but
- if $D(a,b) = 0$, then we cannot draw any conclusions without more work.

You might wonder why, in the local maximum/local minimum cases of this theorem, $f_{xx}(a,b)$ appears rather than $f_{yy}(a,b)$. The answer is only that x is before y in the alphabet. You can use $f_{yy}(a,b)$ just as well as $f_{xx}(a,b)$. The reason is that if $D(a,b) > 0$ (as in the first two bullets of the theorem), then because $D(a,b) = f_{xx}(a,b)\, f_{yy}(a,b) - f_{xy}(a,b)^2 > 0$, we necessarily have $f_{xx}(a,b)\, f_{yy}(a,b) > 0$ so that $f_{xx}(a,b)$ and $f_{yy}(a,b)$ must have the same sign — either both are positive or both are negative.

You might also wonder why we cannot draw any conclusions when $D(a,b) = 0$ and what happens then. The second derivative test for functions of two variables is derived in precisely the same way as the second derivative test for functions of one variable is derived — you approximate the function by a polynomial that is of degree two in $(x-a)$, $(y-b)$ and then you analyze the behaviour of the quadratic polynomial near (a,b). For this to work, the contributions to $f(x,y)$ from terms that are of degree two in $(x-a)$, $(y-b)$ had better be bigger than the contributions to $f(x,y)$ from terms that are of degree three and higher in $(x-a)$, $(y-b)$ when $(x-a)$, $(y-b)$ are really small. If this is not the case, for example when the terms in $f(x,y)$ that are of degree two in $(x-a)$, $(y-b)$ all have coefficients that are exactly zero, the analysis will certainly break down. That's exactly what happens when $D(a,b) = 0$. Here are some examples. The functions

$$f_1(x,y) = x^4 + y^4 \qquad f_2(x,y) = -x^4 - y^4 \qquad f_3(x,y) = x^3 + y^3 \qquad f_4(x,y) = x^4 - y^4$$

all have $(0,0)$ as the only critical point. The first, f_1 has its minimum there. The second, f_2, has its maximum there. The third and fourth have a saddle point there.

© Joel Feldman. 2014. All rights reserved.

Example 13 ($f(x,y) = 2x^3 - 6xy + y^2 + 4y$)

Find and classify all critical points of $f(x,y) = 2x^3 - 6xy + y^2 + 4y$.

Solution. The partial derivatives, of order up to two, are

$$\begin{aligned} f &= 2x^3 - 6xy + y^2 + 4y \\ f_x &= 6x^2 - 6y & f_{xx} &= 12x & f_{xy} &= -6 \\ f_y &= -6x + 2y + 4 & f_{yy} &= 2 & f_{yx} &= -6 \end{aligned}$$

(Of course, f_{xy} and f_{yx} have to be the same. It is still useful to compute both, as a way to catch some mechanical errors.) We have already found, in Example 5, that the critical points are $(1,1)$, $(2,4)$. The classification is

critical point	$f_{xx}f_{yy} - f_{xy}^2$	f_{xx}	type
$(1,1)$	$12 \times 2 - (-6)^2 < 0$		saddle point
$(2,4)$	$24 \times 2 - (-6)^2 > 0$	24	local min

Example 13

Example 14 ($f(x,y) = xy(5x + y - 15)$)

Find and classify all critical points of $f(x,y) = xy(5x + y - 15)$.

Solution. We have already computed the first order partial derivatives

$$f_x(x,y) = y(10x + y - 15) \qquad f_y(x,y) = x(5x + 2y - 15)$$

of $f(x,y)$ in Example 6. The second order derivatives are

$$\begin{aligned} f_{xx}(x,y) &= 10y \\ f_{yy}(x,y) &= 2x \\ f_{xy}(x,y) &= (1)(10x + y - 15) + y(1) = 10x + 2y - 15 \\ f_{yx}(x,y) &= (1)(5x + 2y - 15) + x(5) = 10x + 2y - 15 \end{aligned}$$

(Once again, we have computed both f_{xy} and f_{yx} to guard against mechanical errors.) We have already found, in Example 6, that the critical points are $(0,0)$, $(0,15)$, $(3,0)$ and $(1,5)$. The classification is

critical point	$f_{xx}f_{yy} - f_{xy}^2$	f_{xx}	type
$(0,0)$	$0 \times 0 - (-15)^2 < 0$		saddle point
$(0,15)$	$150 \times 0 - 15^2 < 0$		saddle point
$(3,0)$	$0 \times 6 - 15^2 < 0$		saddle pt
$(1,5)$	$50 \times 2 - 5^2 > 0$	50	local min

Example 14

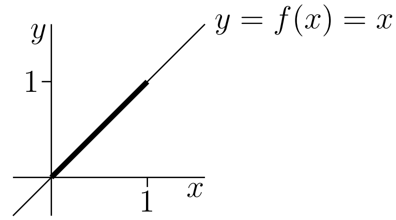
© Joel Feldman. 2014. All rights reserved.

Absolute Maxima and Minima

Definition 15.

Let D be a subset of $\mathbb{R}^2$ and let the function $f(x, y)$ be defined on D. Then f has an absolute maximum at the point (a, b) of D if $f(x, y) \le f(a, b)$ for all (x, y) in D. Similarly, f has an absolute minimum at (a, b) if $f(x, y) \ge f(a, b)$ for all (x, y) in D.

Let's review how one finds the absolute maximum and minimum of a function of one variable on an interval. For concreteness, let's suppose that we want to find the extremal values of a function $f(x)$ on the interval $0 \le x \le 1$. If an extremal value is attained at some $x = a$ which is in the interior of the interval, i.e. if $0 < a < 1$, then a is also a local maximum or minimum and so has to be a critical point of f. But if an extremal value is attained at a boundary point a of the interval, i.e. if $a = 0$ or $a = 1$, then a need not be a critical point of f. This happens, for example, when $f(x) = x$. The largest value of $f(x)$ on the interval $0 \le x \le 1$ is 1 and is attained at $x = 1$, but $f'(x) = 1$ is never zero, so that f has no critical points.



So to find the maximum and minumum of the function $f(x)$ on the interval $[0, 1]$, you

1. build up a list of all candidate points $0 \le a \le 1$ at which the maximum or miminum could be attained, by finding all a's for which either

   (a) $0 < a < 1$ and $f'(a) = 0$ or

   (b) $0 < a < 1$ and $f'(a)$ does not exist or

   (c) a is a boundary point, i.e. $a = 0$ or $a = 1$,

2. and then you evaluate $f(a)$ at each a on the list of candidates. The biggest of these candidate values of $f(a)$ is the absolute maximum and the smallest of these candidate values is the absolute minimum.

The procedure for finding the maximum and minimum of a function of two variables, $f(x, y)$ in a set like $x^2 + y^2 \le 1$, for example, is similar. You again

1. build up a list of all candidate points (a, b) in the set at which the maximum or miminum could be attained, by finding all (a, b)'s for which either

© Joel Feldman. 2014. All rights reserved.

(a) (a,b) is in the interior of the set (for example $x^2+y^2<1$) and $f_x(a,b)=f_y(a,b)=0$ or

(b) (a,b) is in the interior of the set and $f_x(a,b)$ or $f_y(a,b)$ does not exist or

(c) (a,b) is a boundary point, (for example $a^2+b^2=1$), and could give the maximum or minimum on the boundary (more about this shortly)

2. and then you evaluate $f(a,b)$ at each (a,b) on the list of candidates. The biggest of these candidate values of $f(a,b)$ is the absolute maximum and the smallest of these candidate values is the absolute minimum.

The boundary of a set, like $x^2+y^2\le 1$, in $\mathbb{R}^2$ is a curve, like $x^2+y^2=1$. This curve is a one dimensional set, meaning that it is like a deformed x–axis. We can find the maximum and minimum of $f(x,y)$ on this curve by converting $f(x,y)$ into a function of one variable (on the curve) and using the standard function of one variable techniques. This is best explained by some examples.

Example 16

Find the maximum and minimum of $T(x,y)=(x+y)e^{-x^2-y^2}$ on $x^2+y^2\le 1$.

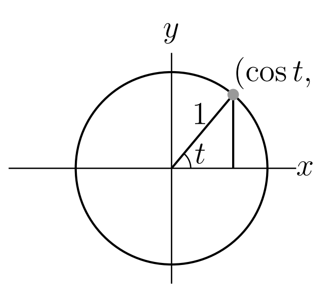
Solution.
Interior: If T takes its maximum or minimum value at a point in the interior, $x^2+y^2<1$, then that point must be a critical point of T. To find the critical points we compute the first order derivatives.

$$T_x(x,y)=(1-2x^2-2xy)e^{-x^2-y^2}\qquad T_y(x,y)=(1-2xy-2y^2)e^{-x^2-y^2}$$

So the critical points are the solutions of

$$\begin{aligned} T_x=0 &\iff 2x(x+y)=1\\ T_y=0 &\iff 2y(x+y)=1 \end{aligned}$$

As both $2x(x+y)$ and $2y(x+y)$ are nonzero, dividing the two equations gives $\frac{x}{y}=1$ which forces $x=y$. Substituting this into either equation gives $2x(2x)=1$ so that $x=y=\pm\frac{1}{2}$. So the only critical points are $(\frac{1}{2},\frac{1}{2})$ and $(-\frac{1}{2},-\frac{1}{2})$. Both are in $x^2+y^2<1$.
Boundary: On the boundary, $x^2+y^2=1$, we may use the figure below to write $x=\cos t$ and $y=\sin t$, so that $T=(\cos t+\sin t)e^{-1}$. As all t's are allowed, this function takes its max and min at zeroes of $\frac{dT}{dt}=\big(-\sin t+\cos t\big)e^{-1}$. That is, when $\sin t=\cos t$, or $x=y$ and

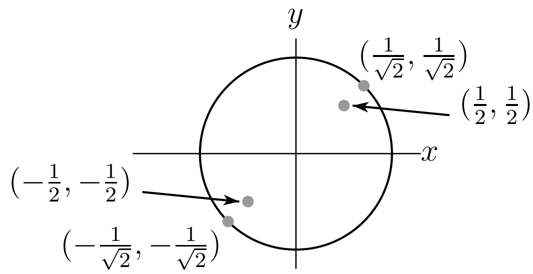


© Joel Feldman. 2014. All rights reserved.

$x^2+y^2=1$, which forces $x=y=\pm\frac{1}{\sqrt{2}}$. All together, we have the following candidates for max and min, with the max and min indicated.

point	$(\frac{1}{2},\frac{1}{2})$	$(-\frac{1}{2},-\frac{1}{2})$	$(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})$	$(-\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}})$
value of T	$\frac{1}{\sqrt{e}}\approx 0.61$	$-\frac{1}{\sqrt{e}}$	$\frac{\sqrt{2}}{e}\approx 0.52$	$-\frac{\sqrt{2}}{e}$
	max	min		

The following sketch shows all of the critical points. It is a good idea to make such a sketch so that you don't accidentally include a critical point that is outside of the allowed region.



Example 16

Example 17

Find the maximum and minimum values of $f(x,y)=x^3+xy^2-3x^2-4y^2+4$ on $x^2+y^2\le 1$.

Solution.
Interior: If f takes its maximum or minimum value at a point in the interior, $x^2+y^2<1$, then that point must be a critical point of f. To find the critical points we compute the first order derivatives.

$$f_x = 3x^2+y^2-6x \qquad f_y = 2xy-8y$$

The critical points are the solutions of

$$\begin{aligned} & f_x = 0 && \text{and} \quad f_y=0 \\ \iff\quad & 3x^2+y^2-6x=0 && \text{and} \quad 2y(x-4)=0 \\ \iff\quad & 3x^2+y^2-6x=0 && \text{and} \quad \big\{\ y=0 \text{ or } x=4\ \big\} \end{aligned}$$

- When $y=0$, x must obey $0=3x^2-6x=3x(x-2)$ so that $x=0$ or $x=2$.
- When $x=4$, y must obey $0=3\times 4^2+y^2-6\times 4=24+y^2$, which is impossible,

So, there are two critical points: $(0,0)$, $(2,0)$.
Boundary: On the boundary, $x^2+y^2=1$, we could again write $x=\cos t$ and $y=\sin t$. But, for practice, we'll use another method. When $x^2+y^2=1$, $y^2=1-x^2$ and

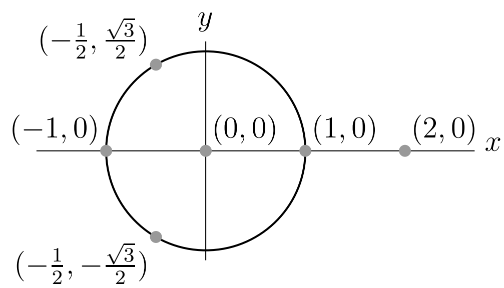
$$f=x^3+x(1-x^2)-3x^2-4(1-x^2)+4=x+x^2$$

The max and min of $x+x^2$ for $-1\le x\le 1$ must occur either

© Joel Feldman. 2014. All rights reserved.

- when $x = -1$ ($\Rightarrow y = f = 0$) or
- when $x = +1$ ($\Rightarrow y = 0,\ f = 2$) or
- when $0 = \frac{d}{dx}(x + x^2) = 1 + 2x$ ($\Rightarrow x = -\frac{1}{2},\ y = \pm\sqrt{\frac{3}{4}},\ f = -\frac{1}{4}$).

Here is a sketch showing all of the points that we have identified.



Note that the point $(2, 0)$ is outside the allowed region. So all together, we have the following candidates for max and min, with the max and min indicated.

point	$(0, 0)$	$(-1, 0)$	$(1, 0)$	$\left(-\frac{1}{2}, \pm\frac{\sqrt{3}}{2}\right)$
value of f	4	2	0	$-\frac{1}{4}$
	max			min

Example 17

Example 18

Find the maximum and minimum values of $f(x, y) = xy - x^3y^2$ when (x, y) runs over the square $0 \le x \le 1,\ 0 \le y \le 1$.

Solution.

Interior: If f takes its maximum or minimum value at a point in the interior, $0 < x < 1$, $0 < y < 1$, then that point must be a critical point of f. To find the critical points we compute the first order derivatives.

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

The critical points are the solutions of

$$\begin{aligned} f_x = 0 \quad &\Longleftrightarrow \quad y(1 - 3x^2y) = 0 \quad \Longleftrightarrow \quad y = 0 \ \text{ or } \ 3x^2y = 1 \\ f_y = 0 \quad &\Longleftrightarrow \quad x(1 - 2x^2y) = 0 \quad \Longleftrightarrow \quad x = 0 \ \text{ or } \ 2x^2y = 1 \end{aligned}$$

- If $y = 0$, we cannot have $2x^2y = 1$, so we must have $x = 0$.
- If $3x^2y = 1$, we cannot have $x = 0$, so we must have $2x^2y = 1$. Dividing gives $1 = \frac{3x^2y}{2x^2y} = \frac{3}{2}$ which is impossible.

© Joel Feldman. 2014. All rights reserved.

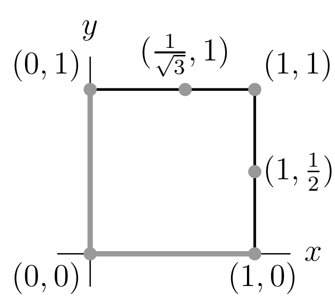
So the only critical point in the square is $(0,0)$. There $f = 0$.

Boundary: The region is a square, so its boundary consists of its four sides.

- First, we look at the part of the boundary with $x = 0$. There $f = 0$.
- Next, we look at the part of the boundary with $y = 0$. There $f = 0$.
- Next, we look at the part of the boundary with $y = 1$. There $f = f(x,1) = x - x^3$. To find the maximum and minimum of $f(x,y)$ on the part of the boundary with $y = 1$, we must find the maximum and minimum of $x - x^3$ when $0 \le x \le 1$. Recall that, in general, the maximum and minimum of a function $h(x)$ on the interval $a \le x \le b$, must occur either at $x = a$ or at $x = b$ or at a critical point of h, i.e. an x for which either $h'(x) = 0$ or $h'(x)$ does not exist. In this case, $\frac{d}{dx}(x - x^3) = 1 - 3x^2$, so the max and min of $x - x^3$ for $0 \le x \le 1$ must occur either at $x = 0$, where $f = 0$, or at $x = \frac{1}{\sqrt{3}}$, where $f = \frac{2}{3\sqrt{3}}$, or at $x = 1$, where $f = 0$.
- Finally, we look at the part of the boundary with $x = 1$. There $f = f(1,y) = y - y^2$. As $\frac{d}{dy}(y - y^2) = 1 - 2y$, the only critical point of $y - y^2$ is at $y = \frac{1}{2}$. So the the max and min of $y - y^2$ for $0 \le y \le 1$ must occur either at $y = 0$, where $f = 0$, or at $y = \frac{1}{2}$, where $f = \frac{1}{4}$, or at $y = 1$, where $f = 0$.

All together, we have the following candidates for max and min, with the max and min indicated.

point	$(0,0)$	$(0, 0{\le}y{\le}1)$	$(0{\le}x{\le}1, 0)$	$(1,0)$	$(1,\frac{1}{2})$	$(1,1)$	$(0,1)$	$(\frac{1}{\sqrt{3}},1)$
value of f	0	0	0	0	$\frac{1}{4}$	0	0	$\frac{2}{3\sqrt{3}} \approx 0.385$
	min	min	min	min		min	min	max



Example 18

© Joel Feldman. 2014. All rights reserved.

Example 19

Find the high and low points of the surface $z = \sqrt{x^2+y^2}$ with (x,y) varying over the square $|x| \le 1,\ |y| \le 1$. Discuss the values of $z_x,\ z_y$ there.

Solution. The surface is a cone. The minimum height is at $(0,0,0)$. The cone has a point there and the derivatives z_x and z_y do not exist. The maximum height is achieved when (x,y) is as far as possible from $(0,0)$. The highest points are at $(\pm 1, \pm 1, \sqrt{2})$. There z_x and z_y exist but are not zero. These points would not be the highest points if it were not for the restriction $|x|, |y| \le 1$.

Example 19

Lagrange Multipliers

A problem of the form

"Find the maximum and minimum values of the function $f(x,y)$ on the curve $g(x,y) = 0$."

is one type of *constrained optimization* problem. The function being mazimized or minimized, $f(x,y)$, is called the *objective function*. The function, $g(x,y)$, whose zero set is the curve of interest, is called the *constraint function*. Such problems are quite common. We have already encountered them in the last section on absolute maxima and minima, when we were looking for the extreme values of a function on the boundary of a region. In economics "utility functions" are used to model the relative "usefulness" or "desirability" or "preference" of various economic choices. For example, a utility function $U(w,\kappa)$ might specify the relative level of satisfaction a consumer would get from purchasing a quantity w of wine and κ of coffee. If the consumer wants to spend \$100 and wine costs \$20 per unit and coffee costs \$5 per unit, then the consumer would like to mazimize $U(w,\kappa)$ subjet to the constraint that $20w + 5\kappa = 100$.

To this point we have always solved such constrained optimization problems either by

- solving $g(x,y) = 0$ for y as a function of x (or for x as a function of y) or by
- parametrizing the curve $g(x,y) = 0$. This means writing all points of the curve in the form $\big(x(t), y(t)\big)$ for some functions $x(t)$ and $y(t)$. For example $x(t) = \cos t$, $y(t) = \sin t$ is a parametrization of the circle $x^2 + y^2 = 1$.

However quite often the function $g(x,y)$ is so complicated that one cannot explicitly solve $g(x,y) = 0$ for y as a function of x or for x as a function of y and one also cannot explicitly parametrize $g(x,y) = 0$. Or sometimes you can, for example, solve $g(x,y) = 0$ for y as a function of x, but the resulting solution is so complicated that it is really hard, or even virtually impossible, to work with. There is another procedure called the method of "Lagrange multipliers" that comes to our rescue in these scenarios.

In this section the method of Lagange mutlipliers will be described and then applied in some examples. The method will be derived in the next, optional, section. It is convenient

© Joel Feldman. 2014. All rights reserved.

to state the method of Lagrange multipliers using a new piece of notation. The *gradient* of a function of two variables $f(x, y)$ is the (two component) vector

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

Theorem 20 (Lagrange Multipliers).

Let $f(x, y)$ and $g(x, y)$ have continuous first partial derivatives in a region of $\mathbb{R}^2$ that contains the curve C given by the equation $g(x, y) = 0$. Assume that $\nabla g(x, y) \neq \mathbf{0}$ there. If f, restricted to the curve C, has a local extreme value at the point (a, b) on C, then there is a real number λ (called a *Lagrange multiplier*) such that

$$\nabla f(a, b) = \lambda \nabla g(a, b) \qquad \text{i.e.} \qquad f_x(a, b) = \lambda g_x(a, b) \quad f_y(a, b) = \lambda g_y(a, b)$$

So to find the maximum and minimum values of $f(x, y)$ on a curve $g(x, y) = 0$, assuming that both the objective function $f(x, y)$ and constraint function $g(x, y)$ have continuous first partial derivatives and that $\nabla g(x, y) \neq \mathbf{0}$, you

1. build up a list of candidate points (x, y) by finding all solutions to the equations

$$f_x(x, y) = \lambda g_x(x, y) \qquad f_y(x, y) = \lambda g_y(x, y) \qquad g(x, y) = 0$$

2. and then you evaluate $f(x, y)$ at each (x, y) on the list of candidates. The biggest of these candidate values is the absolute maximum and the smallest of these candidate values is the absolute minimum.

Example 21

Find the maximum and minimum of $x^2 - 10x - y^2$ on the ellipse $x^2 + 4y^2 = 16$.

Solution. For this problem the objective function is $f(x, y) = x^2 - 10x - y^2$ and the constraint function is $g(x, y) = x^2 + 4y^2 - 16$. The first order derivatives of these functions are

$$f_x = 2x - 10 \qquad f_y = -2y \qquad g_x = 2x \qquad g_y = 8y$$

So, according to the method of Lagrange multipliers, we need to find all solutions to

$$2x - 10 = \lambda(2x) \qquad \iff \qquad (\lambda - 1)x = -5 \tag{7a}$$
$$-2y = \lambda(8y) \qquad \iff \qquad (4\lambda + 1)y = 0 \tag{7b}$$
$$0 = x^2 + 4y^2 - 16 \tag{7c}$$

From (7b), we see that we must have either $\lambda = -1/4$ or $y = 0$.

- If $\lambda = -1/4$, (7a) gives $-\frac{5}{4}x = -5$, i.e. $x = 4$, and then (7c) gives $y = 0$.

© Joel Feldman. 2014. All rights reserved.

- If $y = 0$, then (7c) gives $x = \pm 4$.

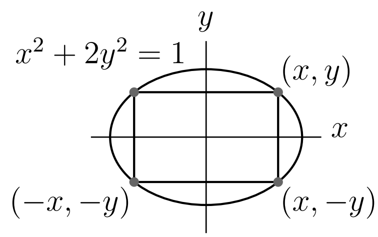
So we have the following table of candidates.

point	$(4,0)$	$(-4,0)$
value of f	-24	56
	min	max

Example 21

Example 22

Find the rectangle of largest area (with sides parallel to the coordinates axes) that can be inscribed in the ellipse $x^2 + 2y^2 = 1$.



Solution. Call the coordinates of the upper right corner of the rectangle (x, y), as in the figure above. The four corners of the rectangle are $(\pm x, \pm y)$ so the rectangle has width $2x$ and height $2y$ and the objective function is $f(x, y) = 4xy$. The constraint function for this problem is $g(x, y) = x^2 + 2y^2 - 1$. The first order derivatives of these functions are

$$f_x = 4y \qquad f_y = 4x \qquad g_x = 2x \qquad g_y = 4y$$

So, according to the method of Lagrange multipliers, we need to find all solutions to

$$4y = \lambda(2x) \qquad \iff \qquad y = \frac{1}{2}\lambda x \tag{8a}$$

$$4x = \lambda(4y) \qquad \implies \qquad x = \lambda y = \frac{1}{2}\lambda^2 x \qquad \implies \qquad x\Big(1 - \frac{\lambda^2}{2}\Big) = 0 \tag{8b}$$

$$0 = x^2 + 2y^2 - 1 \tag{8c}$$

So (8b) is satisfied if either $x = 0$ or $\lambda = \sqrt{2}$ or $\lambda = -\sqrt{2}$.

- If $x = 0$, then (8a) gives $y = 0$ too. But $(0, 0)$ violates the constraint.
- If $\lambda = \sqrt{2}$, then (8a) gives $x = \sqrt{2}y$ and then (8c) gives $2y^2 + 2y^2 = 1$ so that $y = \pm 1/2$ and $x = \pm 1/\sqrt{2}$.
- If $\lambda = -\sqrt{2}$, then (8a) gives $x = -\sqrt{2}y$ and then (8c) gives $2y^2 + 2y^2 = 1$ so that $y = \pm 1/2$ and $x = \mp 1/\sqrt{2}$.

© Joel Feldman. 2014. All rights reserved.

The rectangle of largest area has the vertex $\left(1/\sqrt{2}, 1/2\right)$ in the first quadrant.

Example 22

Example 23

Find the ends of the major and minor axes of the ellipse $3x^2 - 2xy + 3y^2 = 4$. They are the points on the ellipse that are farthest from and nearest to the origin.

Solution. Let (x, y) be a point on $3x^2 - 2xy + 3y^2 = 4$. This point is at the end of a major axis when it maximizes its distance from the centre, $(0, 0)$ of the ellipse. It is at the end of a minor axis when it minimizes its distance from $(0, 0)$. So we wish to maximize and minimize the distance $\sqrt{x^2 + y^2}$ subject to the constraint $g(x, y) = 3x^2 - 2xy + 3y^2 - 4 = 0$. Now maximizing/minmizing $\sqrt{x^2 + y^2}$ is equivalent to maximizing/minmizing $\left(\sqrt{x^2 + y^2}\right)^2 = x^2 + y^2$. So we are free to choose the objective function $f(x, y) = x^2 + y^2$, which we will do, because it makes the derivatives cleaner. Since

$$f_x(x, y) = 2x \qquad f_y(x, y) = 2y \qquad g_x(x, y) = 6x - 2y \qquad g_y(x, y) = -2x + 6y$$

we need to find all solutions to

$$2x = \lambda(6x - 2y) \qquad \iff \qquad (1 - 3\lambda)x + \lambda y = 0 \tag{9a}$$

$$2y = \lambda(-2x + 6y) \qquad \iff \qquad \lambda x + (1 - 3\lambda)y = 0 \tag{9b}$$

$$0 = 3x^2 - 2xy + 3y^2 - 4 \tag{9c}$$

To start, let's concentrate on the first two equations. Pretend, for a couple of minutes, that we already know the value of λ and are trying to find x and y. Note that λ cannot be zero because if it is, (9a) forces $x = 0$ and (9b) forces $y = 0$ and $(0, 0)$ is not on the ellipse. So we may divide by λ and (9a) gives $y = -\frac{1-3\lambda}{\lambda}x$. Subbing this into (9b) gives $\lambda x - \frac{(1-3\lambda)^2}{\lambda}x = 0$. Again, x cannot be zero, since then $y = -\frac{1-3\lambda}{\lambda}x$ would give $y = 0$ and $(0, 0)$ is still not on the ellipse. So we may divide $\lambda x - \frac{(1-3\lambda)^2}{\lambda}x = 0$ by x, giving

$$\lambda - \frac{(1 - 3\lambda)^2}{\lambda} = 0 \iff (1 - 3\lambda)^2 - \lambda^2 = 0 \iff 8\lambda^2 - 6\lambda + 1 = (2\lambda - 1)(4\lambda - 1) = 0$$

We now know that λ must be either $\frac{1}{2}$ or $\frac{1}{4}$. Subbing these into either (9a) or (9b) gives

$$\lambda = \frac{1}{2} \implies -\frac{1}{2}x + \frac{1}{2}y = 0 \implies x = y \quad \overset{(9c)}{\implies} 3x^2 - 2x^2 + 3x^2 = 4 \implies x = \pm 1$$

$$\lambda = \frac{1}{4} \quad \implies \frac{1}{4}x + \frac{1}{4}y = 0 \implies x = -y \overset{(9c)}{\implies} 3x^2 + 2x^2 + 3x^2 = 4 \implies x = \pm\frac{1}{\sqrt{2}}$$

Here " $\overset{(9c)}{\implies}$ " indicates that we have just used (9c). The ends of the minor axes are $\pm\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$. The ends of the major axes are $\pm(1, 1)$.

Example 23

© Joel Feldman. 2014. All rights reserved.

Example 24

Find the values of $w \geq 0$ and $\kappa \geq 0$ that maximize the utility function

$$U(w,\kappa) = 6w^{2/3}\kappa^{1/3} \qquad \text{subject to the constraint} \qquad 4w + 2\kappa = 12$$

Solution. For this problem the objective function is $U(w,\kappa) = 6w^{2/3}\kappa^{1/3}$ and the constraint function is $g(w,\kappa) = 4w + 2\kappa - 12$. The first order derivatives of these functions are

$$U_w = 4w^{-1/3}\kappa^{1/3} \qquad U_\kappa = 2w^{2/3}\kappa^{-2/3} \qquad g_w = 4 \qquad g_\kappa = 2$$

The boundary values $w = 0$ and $\kappa = 0$ give utility 0, which is obviously not going to be the maximum utility. So it suffices to consider only local maxima. According to the method of Lagrange multipliers, we need to find all solutions to

$$4w^{-1/3}\kappa^{1/3} = 4\lambda \quad \Longrightarrow \quad \lambda = w^{-1/3}\kappa^{1/3} \tag{10a}$$

$$2w^{2/3}\kappa^{-2/3} = 2\lambda \quad \Longrightarrow \quad w^{2/3}\kappa^{-2/3} = \lambda = w^{-1/3}\kappa^{1/3} \quad \Longrightarrow \quad w = \kappa \tag{10b}$$

$$0 = 4w + 2\kappa - 12 \tag{10c}$$

Substituting $w = \kappa$, from (10b), into (10c) gives $6\kappa = 12$. So $w = \kappa = 2$ and the maximum utility is $U(2,2) = 12$.

Example 24

Derivation of the Method of Lagrange Multipliers (Optional)

We'll now develop the method of Lagrange multipliers. Let's use C to denote the curve $g(x,y) = 0$. Suppose that (a,b) is a point of C and that $f(x,y) \geq f(a,b)$ for all points (x,y) on C that are close to (a,b). That is (a,b) is a local minimum for f on C. Of course the argument for a local maximum would be virtually identical.

Imagine that we go for a walk on C, with the time t running, say, from $t = -1$ to $t = +1$ and that at time $t = 0$ we happen to be at (a,b). Let's say that our position is $\big(x(t), y(t)\big)$ at time t. We are always on C, so $g\big(x(t), y(t)\big) = 0$ for all t. Write

$$F(t) = f\big(x(t), y(t)\big)$$

That's the value of f we see at time t. Then

$$F(0) = f\big(x(0), y(0)\big) = f(a,b) \leq f\big(x(t), y(t)\big) = F(t)$$

for all t close to zero (so that $\big(x(t)\,,\, y(t)\big)$ is close to (a,b)). So $F(t)$ has a local minimum at $t = 0$ and consequently $F'(0) = 0$.

Now we need to figure out what

$$F'(0) = \frac{\mathrm{d}}{\mathrm{d}t} f\big(x(t), y(t)\big)\Big|_{t=0}$$

© Joel Feldman. 2014. All rights reserved.

is, in terms of the partial derivatives of f. We'll start with a simpler problem of the same type. Define $u(x) = f(x,b)$ and $U(t) = u\big(x(t)\big) = f\big(x(t),b\big)$. Note that $u'(x)$ is the rate of change of f with respect to x when y is held fixed at b. That is, $u'(x) = f_x(x,b)$. By the ordinary chain rule for functions of one variable,

$$\frac{\mathrm{d}}{\mathrm{d}t}U(t) = u'\big(x(t)\big)\, x'(t)$$

Putting in what U and u' are in terms of f,

$$\frac{\mathrm{d}}{\mathrm{d}t}f\big(x(t),b\big) = f_x\big(x(t),b\big)\, x'(t) \implies \frac{\mathrm{d}}{\mathrm{d}t}f\big(x(t),b\big)\Big|_{t=0} = f_x\big(x(0),b\big)\, x'(0) = f_x\big(a,b\big)\, x'(0)$$

By a similar argument with $v(y) = f(a,y)$ and $V(t) = v\big(y(t)\big) = f\big(a,y(t)\big)$,

$$\frac{\mathrm{d}}{\mathrm{d}t}f\big(a,y(t)\big) = f_y\big(a,y(t)\big)\, y'(t) \implies \frac{\mathrm{d}}{\mathrm{d}t}f\big(a,y(t)\big)\Big|_{t=0} = f_y\big(a,y(0)\big)\, y'(0) = f_y\big(a,b)\big)\, y'(0)$$

So the contribution to the rate of change (at $t = 0$) of f from the motion in the x direction is $f_x\big(a,b\big)x'(0)$ and the contribution to the rate of change of f from the motion in the y direction is $f_y\big(a,b\big)y'(0)$. All together the rate of change of f from the full motion $\big(x(t),y(t)\big)$ is

$$\frac{\mathrm{d}}{\mathrm{d}t}f\big(x(t),y(t)\big)\Big|_{t=0} = f_x\big(a,b\big)x'(0) + f_y\big(a,b\big)y'(0) \tag{11}$$

We will not justify this statement, but it is true. It is the chain rule for functions of two variables. Recalling the definition of the gradient, and recalling that $F'(0) = 0$, we may rewrite this as

$$0 = F'(0) = \nabla f(a,b)\cdot\langle x'(0)\,,\, y'(0)\rangle \implies \nabla f(a,b) \perp \langle x'(0)\,,\, y'(0)\rangle$$

Replacing f by g in (11), we also have

$$\frac{\mathrm{d}}{\mathrm{d}t}g\big(x(t),y(t)\big)\Big|_{t=0} = g_x\big(a,b\big)x'(0) + g_y\big(a,b\big)y'(0)$$

Since $\big(x(t),y(t)\big)$ is on C for all t, we have that $g\big(x(t),y(t)\big) = 0$ for all t, so

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}g\big(x(t),y(t)\big)\Big|_{t=0} = \nabla g(a,b)\cdot\langle x'(0)\,,\, y'(0)\rangle \implies \nabla g(a,b) \perp \langle x'(0)\,,\, y'(0)\rangle$$

Now both the vectors $\nabla f(a,b)$ and $\nabla g(a,b)$ are perpendicular to the same vector $\langle x'(0)\,,\, y'(0)\rangle$ (which we can always choose to be nonzero), so $\nabla f(a,b)$ and $\nabla g(a,b)$ have to be parallel vectors. That is,

$$\nabla f(a,b) = \lambda\nabla g(a,b)$$

for some number λ. That's the Lagrange multiplier rule of Theorem 20.

© Joel Feldman. 2014. All rights reserved.