Chapter 1 of Calculus^{++}: Differential calculus with several variables

Gradients, Hessians and Jacobians for functions of two variables

by

Eric A Carlen
Professor of Mathematics
Georgia Tech

Spring 2006

Table of Contents

Section 1: Continuity for functions of a vector variable

1.1 Functions of several variables

We will be studying functions of several variables, say $x_1, x_2 \ldots, x_n$. It is often convenient to organize this list of input variables into a vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$. When n is two or three, we usually dispense with the subscripts and write $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ or $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

For example, consider the function f from $I\!R^2$ to $I\!R$ defined by

$$f(x, y) = x^2 + y^2 \ . \tag{1.1}$$

With $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, we can write this as

$$f(\mathbf{x}) = x^2 + y^2 \ . \tag{1.2}$$

As we shall see, sometimes it is very helpful to think of the input variables as united into a single vector variable $\mathbf{x}$, while other times it is helpful to think of them individually and separately.

We will also be considering functions from $I\!R^n$ to $I\!R^m$. These take vector variables as input, and return vector variables as output. For example, consider the function $\mathbf{F}$ from $I\!R^2$ to $I\!R^3$ given by

$$\mathbf{F}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x^2 + y^2 \\ xy \\ x^2 - y^2 \end{bmatrix} \ . \tag{1.3}$$

Introducing the functions

$$g(x, y) = xy \qquad \text{and} \qquad h(x, y) = x^2 - y^2 \ ,$$

and with $f(x, y)$ defined as in (1.1), we can rewrite as

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \\ h(\mathbf{x}) \end{bmatrix} \ .$$

Often, the questions we ask about $\mathbf{F}(\mathbf{x})$ can be answered by considering the functions f, g and h one at a time.

What kinds of questions will we be asking about such functions? Many of the questions have to do with *solving equations involving* $\mathbf{F}$. For example, consider the equation

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \ .$$

We can rewrite this as a *system of equations* using the functions f g and h introduced above:

$$\begin{aligned} f(x,y) &= 2 \\ g(x,y) &= 1 \\ h(x,y) &= 0 \ . \end{aligned}$$

More explicitly, this is

$$\begin{aligned} x^2 + y^2 &= 2 \\ xy &= 1 \\ x^2 - y^2 &= 0 \ . \end{aligned}$$

This is not a linear system of equations, so the methods of linear algebra cannot be applied directly. However, the point of the main algorithm in linear algebra – row reduction, also known as Gaussian elimination – is to eliminate variables. We can still do that here. Adding the first and third equation, we find $2x^2 = 2$, or $x^2 = 1$. The third equation now tells us $y^2 = x^2 = 1$. So $x = \pm 1$ and $y = \pm 1$. Going to the third equation, we we that if $x = 1$, the $y = 1$ also, and if $x = -1$, then $y = -1$ also. Hence this equation has exactly two solutions

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{x}_1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \ .$$

That is, for these vector $\mathbf{x}_1$ and $\mathbf{x}_2$,

$$\mathbf{F}(\mathbf{x}_1) = \mathbf{F}(\mathbf{x}_2) = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \ ,$$

and no other input vectors $\mathbf{x}$ yield the desired output.

Notice that this system of equations has exactly two solutions. This is something that never happens for a linear system, which either has no solution, a unique solution or infinitely many solutions. In general, it is not easy to solve systems of equations involving non linear functions. There are good methods for obtaining arbitrarily accurate approximate solutions though. One of the best is a multivariable version of Newton's method that we will learn to use here.

It may be disappointing to learn that when faced with an equation of the form

$$\mathbf{F}(\mathbf{x}) = \mathbf{b}$$

for some given $\mathbf{b}$ in $I\!R^m$, we will only be able to come up with approximate solutions. This is not so bad, since by running enough iterations of Newton's method, we can get as much accuracy as we want. This raises the question; *How much accuracy is enough?* This leads us right into the notion of *continuity* for functions of several variables.

1.2 Continuity in several variables

In plain words, this is what continuity at $\mathbf{x}_0$ means for a function $\mathbf{F}$ from $I\!R^n$ to $I\!R^m$:

• *If* $\mathbf{x} \approx \mathbf{x}_0$ *with some small enough margin of error, then* $\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}_0)$ *up to some specified small margin of error.*

Think about this. There are two margins of error referred to – one on the input, and one on the output. Let ϵ denote the specified margin of error on the output, Then we interpret

"$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}_0)$ *up to some specified small margin of error*"

as $|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0)| \le \epsilon$.

Now, no matter how small a non zero value* you pick for ϵ, say 10^{-3}, 10^{-9} or even something wildly impractical like 10^{-300}, there is supposed to be another margin of error, which, following tradition, we will call δ so that if $\mathbf{x} \approx \mathbf{x}_0$ up to an error of size δ, we get the desired level of accuracy on the output.

In general, the level of accuracy required at the input depends on how much accuracy has been required at the output, so generally, the smaller ϵ is, the smaller δ will have to be. That is, δ depends on ϵ, and it helps to signify this by writing $\delta(\epsilon)$. We can now express our notion of continuity in precise quantitative terms

Definition (continuity) A function $\mathbf{F}$ from $I\!R^n$ to $I\!R^m$ is *continuous* at $\mathbf{x}_0$ in case for every $\epsilon > 0$, there is a $\delta(\epsilon) > 0$ so that

$$|\mathbf{x} - \mathbf{x}_0| \le \delta(\epsilon) \qquad \Rightarrow \qquad |\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0)| \le \epsilon \ .$$

for all $\mathbf{x}$ in the domain of $\mathbf{F}$. The function $\mathbf{F}$ is *continuous* if it is continuous at each $\mathbf{x}_0$ in its domain.

Please make sure that you understand the relation between the intuitive definition at the beginning of this subsection, and the precise one that we just gave. Also, make sure you understand this:

• *If a function* $\mathbf{F}$ *from* $I\!R^n$ *to* $I\!R^m$ *is not continuous at a solution* $\mathbf{x}_0$ *of* $\mathbf{F}(\mathbf{x}) = \mathbf{b}$*, it is no use at all to find a vector* $\mathbf{x}_1$ *with even* $|\mathbf{x}_1 - \mathbf{x}_0| < 10^{-300}$ *since without continuity, there is no guarantee that* $\mathbf{F}(\mathbf{x}_1)$ *is at all close to* $\mathbf{F}(\mathbf{x}_0) = \mathbf{b}$.

---

* If you pick a value of zero, it isn't really a margin of error, is it?

Without continuity, only exact solutions are meaningful. But these will often involve irrational numbers that cannot be exactly represented on a computer. Therefore, whether a function is continuous or not is a serious practical matter.

How do we tell if a function is continuous? The advantage of the precise mathematical definition over the intuitive one is precisely that it is checkable. Here is a simple example in the case in which $m = 1$, so only the input is a vector:

Example 1 (Checking continuity) Consider the function f from $I\!R^2$ to $I\!R$ given by $f(x, y) = x$. Then

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| = |x - x_0| \leq \sqrt{|x - x_0|^2 + |y - y_0|^2} = |\mathbf{x} - \mathbf{x}_0| \ . \tag{1.4}$$

Hence we can take $\delta(\epsilon) = \epsilon$ since (1.4) says that

$$|\mathbf{x} - \mathbf{x}_0| \leq \epsilon \qquad \Rightarrow \qquad |f(\mathbf{x}) - f(\mathbf{x}_0)| \leq \epsilon \ .$$

The choice $\delta(\epsilon) = \epsilon$ works at every $\mathbf{x}_0$, so this function is continuous on all of $I\!R^2$.

The same analysis would have applied to $g(x, y) = y$, and even more simply to the constant function $h(x, y) = 1$. After this, it gets more complicated to use the definition, but there is no need to do this:

Theorem 1 (Building Continuous Functions) *Let f and g be continuous functions from some domain U in $I\!R^n$ to $I\!R$. Define the functions fg and $f + g$ by $fg(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ and $(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$. Then fg and $f + g$ are continuous on U. furthermore, if $g \neq 0$ anywhere in U, then f/g defined by $(f/g)(\mathbf{x}) = f(\mathbf{x})/g(\mathbf{x})$ is continuous in U. Finally, if h is a continuous function for $I\!R$ to R, then the composition $h \circ f$ is continuous on U.*

Proof: Consider the case of $f + g$. Fix any $\epsilon > 0$, and any $\mathbf{x}_0$ in U. Since f ang g are continuous there is a $\delta_f(\epsilon/2) > 0$ and a $\delta_g(\epsilon/2) > 0$ so that

$$|\mathbf{x} - \mathbf{x}_0| \leq \delta_f(\epsilon/2) \qquad \Rightarrow \qquad |f(\mathbf{x}) - f(\mathbf{x}_0)| \leq \epsilon/2$$

and

$$|\mathbf{x} - \mathbf{x}_0| \leq \delta_g(\epsilon/2) \qquad \Rightarrow \qquad |g(\mathbf{x}) - g(\mathbf{x}_0)| \leq \epsilon/2$$

Now define

$$\delta(\epsilon) = \max\{\delta_f(\epsilon/2) \ , \ \delta_g(\epsilon/2)\} \ .$$

Then, whenever $|\mathbf{x} - \mathbf{x}_0| \leq \delta(\epsilon)$,

$$|(f + g)(\mathbf{x}) - (f + g)(\mathbf{x}_0)| \leq |f(\mathbf{x}) - f(\mathbf{x}_0)| + |g(\mathbf{x}) - g(\mathbf{x}_0)| \leq \epsilon/2 + \epsilon/2 = \epsilon$$

so that

$$|\mathbf{x} - \mathbf{x}_0| \leq \delta(\epsilon) \qquad \Rightarrow \qquad |(f + g)(\mathbf{x}) - (f + g)(\mathbf{x}_0)| \leq \epsilon \ .$$

This proves the continuity of $f + g$. The other cases are similar, and the proofs work just like the proofs of the corresponding statements about functions of a single variable. They are therefore left as exercises. ■

To apply the theorem, we take a function apart, and try to recognize it as built out of continuous building–blocks. For example, consider $z(x,y) = \cos\left((1+x^2+y^2)^{-1}\right)$.

This is built out of the continuous building blocks

$$f(x,y) = x \qquad g(x,y) = y \qquad \text{and} \qquad h(x,y) = 1 \ .$$

Indeed,

$$(1+x^2+y^2)^{-1} = \cos\left(\frac{h}{h+ff+gg}\right)(\mathbf{x}) \ .$$

Repeated application of the theorem shows this function is continuous.

Finally, consider a function $\mathbf{F}$ from $I\!R^n$ to $I\!R^m$. We can write it in terms of m functions f_j from $I\!R^n$ to $I\!R$:

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix} \ .$$

Then it is easy to see that $\mathbf{F}$ is continuous if and only if each f_j is continuous. Hence a good understanding of the case in which only the input is a vector suffices for the general case. The proof is left as an exercise.

This last point is the reason we will focus on functions from $I\!R^n$ to $I\!R$ in the next few sections, in which we introduce the differential calculus for functions of several variables. So that we can draw pictures, we will furthermore focus first on $n = 2$.

1.3 Separate continuity is not continuity

Definition (Separate continuity) A function $f(x,y)$ on $I\!R^2$ is *separately continuous* in case for each x_0, the function $y \mapsto f(x_0,y)$ is a continuous function of y, and if for y_0, the function $x \mapsto f(x,y_0)$ is a continuous function of x

The definition is extended to arbitrarily many variables in the obvious way. It would be nice if all separately continuous functions were continuous. Then we could check for continuity using what we know about continuity of functions of a single variable – we could just check one variable at a time. Unfortunately, this is not the case.

Example 2 (Separately continuous but not continuous) Consider the function f from $I\!R^2$ to $I\!R$ given by

$$f(x,y) = \begin{cases} \dfrac{2xy}{x^4+y^4} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases} \ . \tag{1.5}$$

To see that this function is separately continuous, fix any x_0, and define

$$g(y) = f(x_0, x) \ .$$

If $x_0 = 0$, then $g(y) = 0$ for all y. This is certainly continuous. If $x_0 \neq 0$, then the denominator in $g(y)$, $x_0^4 + y^4$, is a strictly positive polynomial in y. Hence $g(y)$ is a ration function of y with a denominator that is strictly positive for all y, Such functions are continuous. This shows that no matter how x_0 is chosen, $g(y)$ is a continuous function of y.

The same argument applies when we fix any y_0 and consider $x \mapsto f(x, y_0)$. Indeed, the function is symmetric in x and y, so whatever we prove for y is also true for x. Hence this function is separately continuous.

However, it is *not continuous*. For each $t \geq 0$, let $\mathbf{x}(t) = \begin{bmatrix} t \\ t \end{bmatrix}$. Then

$$f(\mathbf{x}(t)) = \frac{2t^2}{t^4 + t^4} = \frac{1}{t^2} \ . \tag{1.6}$$

Therefore, since $f(\mathbf{x}(0)) = 0$,

$$|f(\mathbf{x}(t)) - f(\mathbf{x}(0))| = \frac{1}{t^2} \ .$$

On the other hand,

$$|\mathbf{x}(t) - \mathbf{x}(0)| = \sqrt{2}t \ .$$

As t approaches zero, so does $|\mathbf{x}(t) - \mathbf{x}(0)|$. If f we continuous, would then be the case that $|f(\mathbf{x}(t)) - f(\mathbf{x}(0))|$ would tends to zero as well. Instead, it does quite the opposite: it "blows up" as t tends to zero.

Separate continuity is easier to check than continuity – it can be done one variable at a time. Maybe this is actually a better generalization of continuity to several variables? This is not the case. Mathematical definitions are made because of what can be done with them. That is, they are made because they capture concepts that are useful in solving concrete problems.

Part of the problem solving value of the concept of continuity lies in its relevance to minimum–maximum problems. You know from the theory of function of a single variable that if g is any *continuous* function of x, then g *attains its maximum* on any closed interval $[a, b]$. That is, there is a point x_0 with $a \leq x_0 \leq b$ so that for every x with $a \leq x \leq b$,

$$g(x_0) \geq g(x) \ .$$

In this case, we say that x_0 is a *maximizer* of g on $[a, b]$. Finding maximizers is one of the important applications of the differential calculus.

In the next subsection, we show that continuity is the right hypothesis for proving a multi variable version of this important theorem, and that separate continuity is not enough. Separate continuity is easier to check, but alas, it is just not that useful.

1.4 Continuity and maximizers

We can learn more form the function in Example 2. Let's restrict its domain to the unit square

$$0 \leq x, y \leq 1 \ .$$

As you see from (1.6), the function f defined in (1.5) is *unbounded* on $0 \leq x, y \leq 1$. That is, given any number B, there is a point $\mathbf{x}$ in the closed unit square so that $f(\mathbf{x}) > B$. Hence we see that it is *not the case* that separately continuous functions attain their maxima in nice sets like the unit square. We have a counter example, and things go wrong in the worst possible way: *Not only is the bound not attained, the function is not even bounded.*

However, for continuous functions of several variables, there is an analog of the single variable theorem:

Theorem 2 (Continuity and Maximizers) *Let f be a continuous function defined on a closed bounded set C in $I\!R^n$. Then there is point $\mathbf{x}_0$ in C so that*

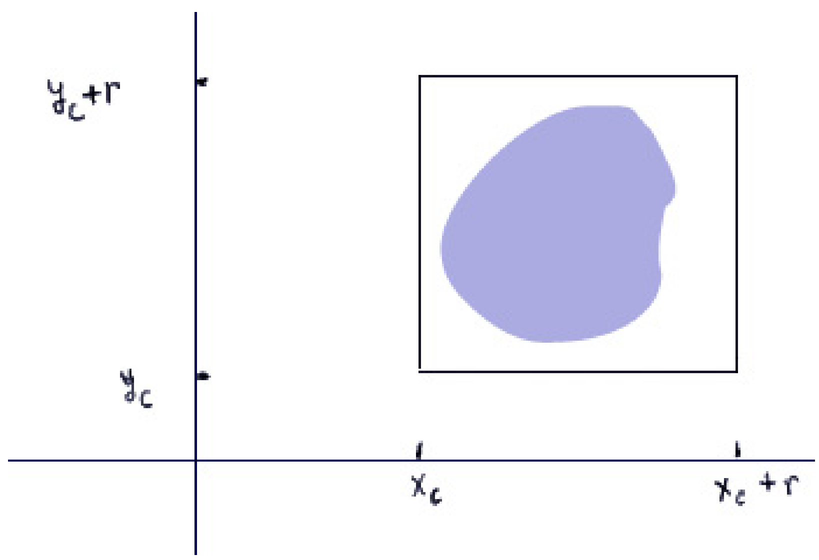
$$f(\mathbf{x}_0) \geq f(\mathbf{x}) \qquad \text{for all} \qquad \mathbf{x} \text{ in } C\ . \tag{1.7}$$

Proof: Let B be the least upper bound of f on C. That is, B is either infinity if f is unbounded on C, or else it is the smallest number that is greater than or equal to $f(\mathbf{x})$ for all $\mathbf{x}$ in C. We aim to show that B is finite, and that there is an $\mathbf{x}_0$ in C with $f(\mathbf{x}_0) = B$. This is the maximizer we seek.

We will do this for $n = 2$ so that we can draw pictures. Once you understand the idea for $n = 2$, you will see that it applies for all n.

Since C is a closed and bounded set, there are numbers x_c, y_c and r so that C is contained in the square

$$x_c \leq x \leq x_c + r \qquad \text{and} \qquad y_c \leq y \leq y_c + r\ .$$



The shaded region is the closed, bounded set C.

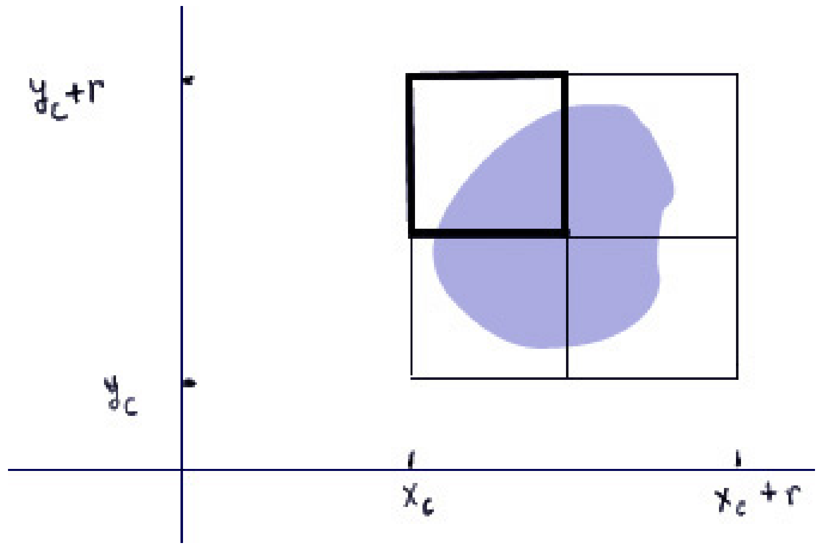
We now build a sequence $\{\mathbf{x}_n\}$ in $I\!R^2$ that will converge to the maximizer we seek. Here is how this goes:

By the definition of B, for each n, C contains points that satisfy either

$$f(\mathbf{x}) \geq B - 1/n \qquad \text{or, if } B = \infty, \qquad f(\mathbf{x}) \geq n\ . \tag{1.8}$$

Divide the square into four congruent squares. Since the four squares cover C, at least one of the four squares must be such that for infinitely many n, it contains points in C satisfying (1.8). Pick such a square, and pick $\mathbf{x}_1$ in it satisfying

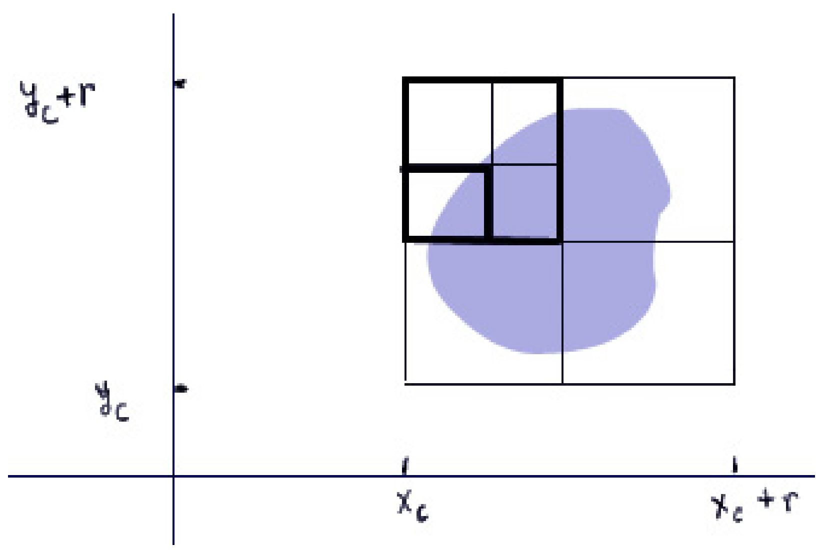
$$f(\mathbf{x}_1) \geq B - 1 \qquad \text{or, if } B = \infty, \qquad f(\mathbf{x}_1) \geq 1\ . \tag{1.9}$$



Here we chose the upper left square.

Next, subdivide the previously chosen square into four smaller squares as in the diagram below. Again, one of these must be such that for infinitely many n, it contains points in C satisfying (1.8). Pick such a square, and pick $\mathbf{x}_2$ in it satisfying

$$f(\mathbf{x}_2) \geq B - 1/2 \qquad \text{or, if } B = \infty, \qquad f(\mathbf{x}_2) \geq 2 \ . \tag{1.10}$$



Here we chose the lower left square in the previously selected square.

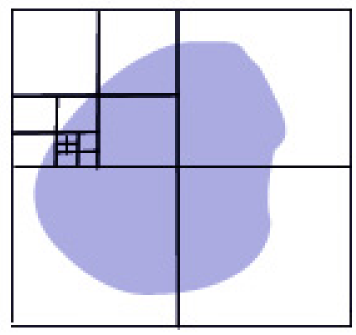
Iterating this procedure produces a sequence of points $\{\mathbf{x}_n\}$ such that

$$f(\mathbf{x}_n) \geq B - 1/n \qquad \text{or, if } B = \infty, \qquad f(\mathbf{x}_n) \geq n \ . \tag{1.11}$$

Moreover, the sequence $\{\mathbf{x}_n\}$ is *convergent*. That is, there is an $\mathbf{x}_0$ so that for every $\epsilon > 0$, there is an N so that

$$n \geq N \quad \Rightarrow \quad |\mathbf{x}_n - \mathbf{x}_0| \leq \epsilon \ .$$

This follows from the fact that the procedure described above produces a nested set of squares.



Since the side length is reduced by a factor of 2 with each subdivision, and since it starts at r, at the nth stage we have a square of side length $2^{-n}r$. As n goes to infinity, the squares shrink down to a single point, $\mathbf{x}_0$.*. Since $\mathbf{x}_n$ is contained in the nth square,

$$\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}_0 \ .$$

Now, since C is closed, $\mathbf{x}_0$ belongs to C, and since f is continuous,

$$\lim_{n\to\infty} f(\mathbf{x}_n) = f(\mathbf{x}_0) \ . \tag{1.12}$$

Next, observe if it were true that $b = \infty$, we would have $f(\mathbf{x}_n) \geq n$ for all n. This and (1.12) would imply that f is infinite at $\mathbf{x}_0$. But by hypothesis, f is a real valued function, so $f(\mathbf{x}_0)$ is a finite real number. Hence $B < \infty$, and it must be the case that

$$f(\mathbf{x}_n) \geq B - 1/n$$

for all n. Hence $\lim_{n\to\infty} f(\mathbf{x}_n) = B$. By (1.12), $f(\mathbf{x}_0) = B$, and since B is the least upper bound of f on C, it is in particular an upper bound, and we have (1.7). ■

Problems

1. Complete the proof of Theorem 1.

2. Consider the function defined by

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

Show that this function is continuous.

* Some of the exercises are designed to elaborate on this point, which is probably intuitively clear, if somewhat subtle

3. A function f from a domain U in $I\!R^n$ to $I\!R$ is called a *Lipschitz continuous* function in case there is some number M so that

$$|f(\mathbf{x}) - f(\mathbf{y})| \le M|\mathbf{x} - \mathbf{y}| \tag{1.13}$$

for all $\mathbf{x}$ and $\mathbf{y}$ in U. Show that a Lipschitz continuous function is continuous by finding a valid margin of error on the input; i.e., a valid $\delta(\epsilon)$.

4. Show that the function f defined in Problem 2 is not Lipschitz continuous. More specifically, show that for $\mathbf{y} = 0$, there is no finite M for which (1.13) holds in any neighborhood of 0. Together with the previous problem, this shows that Lipschitz continuity is a strictly less general concept than continuity. However, it can be very easy to check, as in the next problem.

5. Consider the function f defined by

$$f(x_1, x_2) = \sin(x_1)\cos(x_2) \ .$$

Note that

$$f(x_1, x_2) - f(y_1, y_2) = (\sin(x_1) - \sin(y_1))\cos(x_2) + \sin(y_1)(\cos(x_2) - \cos(y_2)) \tag{1.14}$$

Show that

$$|\sin(x_1) - \sin(y_1)| \le |x_1 - y_1| \qquad \text{and} \qquad |\cos(x_2) - \cos(y_2)| \le |x_2 - y_2| \ .$$

(This is a single variable problem, and the fundamental theorem of calculus can be applied). Combine this with the identity (1.14) to show that f satisfies (1.13) with $M = \sqrt{2}$.

6. Prove the assertion that a function $\mathbf{F}$ from $I\!R^n$ to $I\!R^m$ is continuous if and only if each of its component functions f_j is continuous.

7. Consider the function defined by

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

where $r > 0$. Show that this function is continuous for $0 < r < 2$, and discontinuous for $r \ge 2$. For which values of r is it bounded.

8 Show that every continuous function is separately continuous, so that separate continuity is a strictly more general, if less useful, concept.

Section 2: Derivatives of functions of a vector variable

2.1 Understanding functions of several variables – one variable at a time

One *very useful* method for analyzing a function $f(x, y)$ of two variables is to allow only one variable at a time to vary. That is, at every stage in the analysis, "freeze" one variable so that only the other one is allowed to vary. In effect, this produces a function of one variable, and we can apply what we know about calculus in one variable to study $f(x, y)$. In the section on continuity, we saw that this approach is not suited to every sort of question, but here we will see that is is very well suited to many others.

Consider a function $f(x, y)$, and "freeze" the value of y at $y = y_0$. We then obtain a single variable function $g(x)$ where

$$g(x) = f(x, y_0) \ . \tag{2.1}$$

In the case $f(x, y) = \sqrt{1 + x^2 + y^2}$ and $y_0 = 2$, we would have

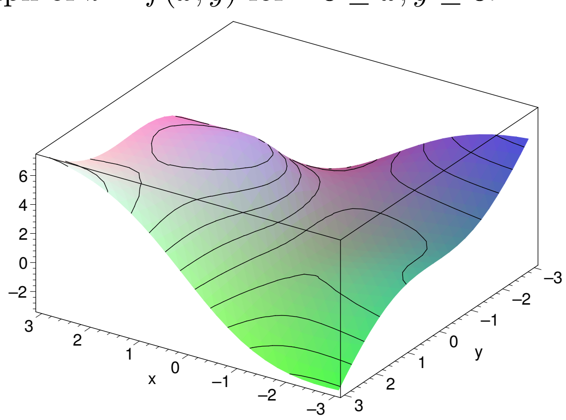
$$g(x) = \sqrt{5 + x^2} \ .$$

The graph of $z = g(x)$ is the "slice" of the graph $z = f(x, y)$ lying above the line $y = 2$ in the x, y plane. The slope of the tangent line to $z = g(x)$ at $x = x_0$ is $g'(x_0)$, This is the rate at which $g(x)$ increases as x is increased through x_0. But by the definition of $g(x)$, this same number tells us something about the behavior of f: It is the rate at which $f(x, 2)$ increases as x increases through x_0.

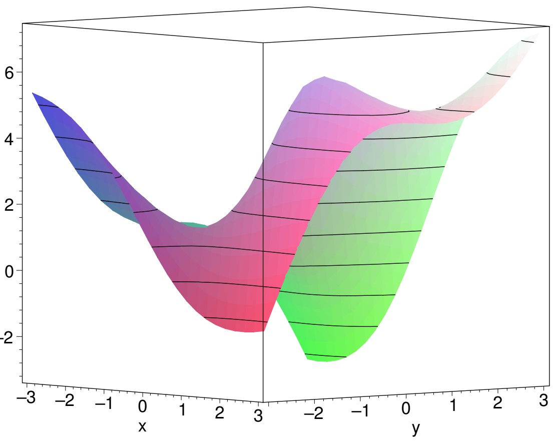
This "slicing" idea is fundamental, and deserves a closer look – with pictures. Consider the function $f(x, y)$ given by

$$f(x, y) = \frac{3(1 + x)^2 + xy^3 + y^2}{1 + x^2 + y^2} \ .$$

Here is a plot of the graph of $z = f(x, y)$ for $-3 \le x, y \le 3$:



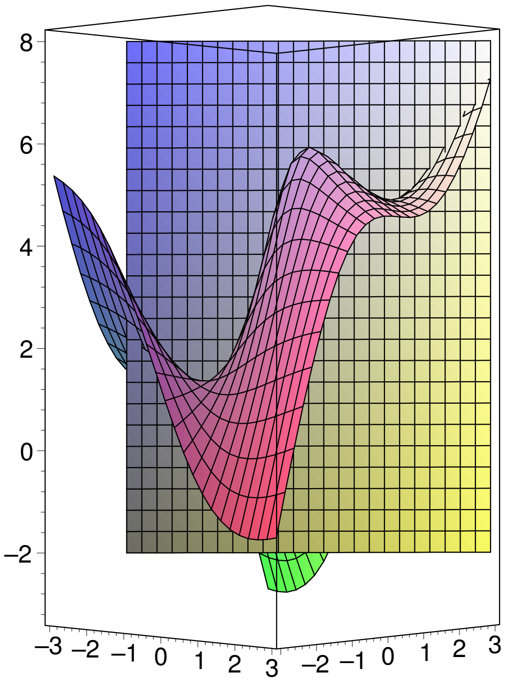
Here is another picture of the same graph, but from a different angle that give more of a side view:



In both graphs, the curves drawn on the surface show points that have the same z value, which we can think of as representing "altitude". Drawing them in helps us get a good visual understanding of the "landscape" in the graph.*

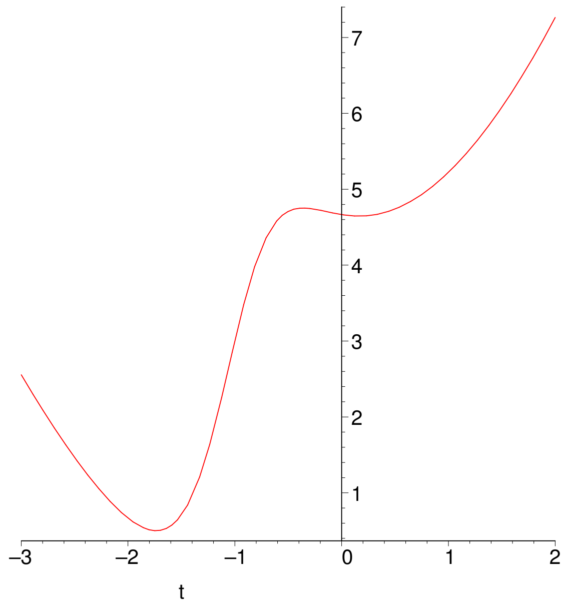
Now that we understand what the graph of $z = f(x, y)$ looks like, let's *slice it along a line*. Suppose for example that you are walking on a path in this landscape, and in the x, y plane, your path runs along the line $y = x$, passing through the point $(1, 1)$ at time $t = 0$.

In $I\!R^3$, the equation $y = x$, or $x - y = 0$ is the equation of a *vertical* plane. Here is a picture of this vertical plane slicing through the graph of $z = f(x, y)$:



The next graph shows the "altitude profile" as we walk along the graph; this curve is where the surface intersects our vertical plane:

* The formula for $f(x, y)$ was chosen to produce a graph that looks like a bit of a mountain landscape. "Contour curves", which are the curves of constant altitude on a topographic map, will be studied in the next section.



Compare the last two graphs, and make sure you see how the curve in the second one corresponds to the intersection of the plane and the surface in the first one.

Next, notice that in the second graph, the horizontal coordinate is t. Where did this new variable t come from? It came from parameterizing the the line $y = x$ around the base point $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The direction vector $\mathbf{v}$ is also $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and this leads to the parameterization

$$\mathbf{x}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} 1+t \\ 1+t \end{bmatrix} .$$

Then, plugging $x(t) = 1 + t$ and $y(t) = 1 + t$ into $f(x, y)$ gives us

$$g(t) = f(x(t), y(t)) = f(1+t, 1+t) = \frac{3(2+t)^2 + (1+t)^4 + (1+t)^2}{(1 + 2(1+t)^2)} .$$

This is the function that is plotted above, for $-3 \le t \le 2$.

The function $g(t)$ is a familiar garden variety function of a single variable, and we can use everything we know about single variable calculus to analyze it. For example, you could compute the derivative $g'(t)$. As you can see from the picture, you would find exactly 3 values of t in the range $-3 \le t \le 2$ at which $g'(t) = 0$. At all other points on the path you, are going either up hill or down hill on a non zero slope.

The notion of the slope along a slice, or a path, brings us to the notion of a *partial derivative.*

2.2 Partial derivatives and how to compute them

First, the definition:

Definition (partial derivatives) Given a function $f(x, y)$ defined in a neighborhood of (x_0, y_0), the partial derivative of f with respect to x at (x_0, y_0) is denoted by $\frac{\partial}{\partial x} f(x_0, y_0)$ and is defined by

$$\frac{\partial}{\partial x} f(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \tag{2.2}$$

provided that the limit exists. Likewise, the partial derivative of f with respect to y at (x_0, y_0) is denoted by $\frac{\partial}{\partial y} f(x_0, y_0)$ and is defined by

$$\frac{\partial}{\partial y} f(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} \tag{2.3}$$

Now, how to compute partial derivatives: This turns out to be easy! If $g(x)$ is related to $f(x, y)$ through (2.1), then

$$\frac{\partial}{\partial x} f(x_0, y_0) = \lim_{h \to 0} \frac{g(x_0 + h) - g(x_0)}{h} = g'(x_0)\ . \tag{2.4}$$

This is wonderful news! We will not need to make explicit use of the definition of partial derivatives very often to compute them. Rather, we can use (2.4) and everything we know about computing $g'(x_0)$ for functions of a single variable.

Example 1 (Differentiating in one variable) Let $f(x, y) = \sqrt{1 + x^2 + y^2}$ and $(x_0, y_0) = (1, 2)$. Then with $g(x)$ defined as in (2.1), $g(x) = \sqrt{5 + x^2}$. By a simple computation,

$$g'(x) = \frac{x}{\sqrt{5 + x^2}}$$

and in particular

$$\frac{\partial}{\partial x} f(1, 2) = \frac{1}{\sqrt{9}} = \frac{1}{3}\ .$$

In single variable calculus, the derivative function $g'(x)$ is the function assigning the "output " $g'(x_0)$ to the "input" x_0. In the same way, we let $\frac{\partial}{\partial y} f(x, y)$ denote the function of the two variables x and $y$that assigns the "output" $\frac{\partial}{\partial y} f(x_0, y_0)$ to the "input" (x_0, y_0). The same considerations apply to $\frac{\partial}{\partial y} f(x, y)$, *mutatis–mutandis.*

Example 2 (Computing partial derivatives) Let $f(x, y) = \sqrt{1 + x^2 + y^2}$. holding y fixed – as a parameter instead of a variable – we differentiate with respect to x as in the single variable calculus, and find

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

Likewise,

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

Because computing partial derivatives is just a matter of differentiating with respect to one chosen variable, everything we know about differentiating with respect to one variable can be applied – in particular the chain rule and the product rule.

Example 3 (Using the chain rule) The function $f(x, y) = \sqrt{1 + x^2 + y^2}$ that we considered in Example 2 can be written as a composition $f(x, y) = g(h(x, y))$ where

$$g(z) = \sqrt{z + 1} \qquad \text{and} \qquad h(x, y) = x^2 + y^2 \ .$$

Since

$$g'(z) = \frac{1}{2\sqrt{1 + z}} \qquad \text{and} \qquad \frac{\partial}{\partial x} h(x, y) = 2x \ ,$$

we have

$$\frac{\partial}{\partial x} f(x, y) = \frac{\partial}{\partial x} g(h(x, y)) = g'(h(x, y)) \frac{\partial}{\partial x} h(x, y) = \frac{1}{2\sqrt{1 + h(x, y)}} 2x = \frac{x}{\sqrt{x^2 + y^2}} \ ,$$

as before.

What we saw in Example 3 is a generally useful fact about partial derivatives: If g is a differentiable function of a single variable, and h is a function of two (or more) variables with $\partial h/\partial x$ defined, then

$$\frac{\partial}{\partial x} g(h(x, y)) = g'(h(x, y)) \frac{\partial}{\partial x} h(x, y) \ ,$$

and similarly with y and any other variables.

In short, as far as *computing* partial derivatives goes, there is nothing much new: Just pay attention to one variable at a time, and differentiate with respect to it as usual.

However, understanding exactly what the partial derivatives of f tell us about f is more subtle. For example, you know that whenever a function g of a single variable is differentiable, it is continuous. As we'll see next, a function of two variables can be discontinuous at a point *even if both partial derivatives exist everywhere.*

2.3 Partial derivatives, directional derivatives, and continuity

Consider the function f defined by

$$f(x, y) = \begin{cases} \dfrac{2xy}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases} \ . \qquad (2.5)$$

If we try to compute the partial derivatives of f at a point (x_0, y_0), the "two rule" definition of f causes no difficulty when $(x_0, y_0) \neq (0,0)$, but we have to take it into account and use the definitions (2.2) and (2.3) when $(x_0, y_0) = (0,0)$.

When $(x_0, y_0) = (0,0)$,

$$\frac{f(0+h,0) - f(0,0)}{h} = \frac{0-0}{h} = 0 \ ,$$

and so, by (2.2), $\dfrac{\partial}{\partial x} f(0,0) = 0$. In the same way, we find from (2.3) that $\dfrac{\partial}{\partial y} f(0,0) = 0$.

When $(x,y) \neq (0,0)$, we can calculate

$$\frac{\partial}{\partial x} f(x,y) = \frac{2y}{x^2+y^2} - \frac{4x^2y}{(x^2+y^2)^2} = 2y\frac{y^2-x^2}{(x^2+y^2)^2} \ ,$$

and

$$\frac{\partial}{\partial y} f(x,y) = \frac{2x}{x^2+y^2} - \frac{4y^2x}{(x^2+y^2)^2} = 2x\frac{x^2-y^2}{(x^2+y^2)^2} \ .$$

Both partial derivatives of the function $f(x,y)$ defined by (2.5) exist at every point (x,y) of the plane. We now come to a striking difference between calculus in one variable and more than one.

Roughly speaking, to say that f is continuous at (x_0, y_0) means that $f(x,y) \approx f(x_0, y_0)$. Now consider the function f defined by (2.5): For any $t \neq 0$,

$$f(t,t) = \frac{2t^2}{t^2+t^2} = 1 \ .$$

But by (2.5), $f(0,0) = 0$. Hence

$$\lim_{t \to 0} f(t,t) = 1 \neq 0 = f(0,0) \ .$$

This is not continuity!

The reason that this functions lack of continuity did not interfere when we computed the partial derivatives is that to compute these *we only make variations in the input along lines that are parallel to the axes.* The discontinuity of f only manifests itself as we approach the origin along some line that is not parallel to the axes.

• *To really understand the nature of functions, we are going to need to study how they vary along general slices, not just slices parallel to the axes*

Therefore, our strategy of reducing to functions of a single variable will become much more incisive if we do not limit ourselves to varying (x,y) in directions that are parallel to the coordinate axes. Let's consider all lines on an equal footing – parallel to the coordinate axes of not. This brings us to the notion of a *directional derivative*

Definition (directional derivatives) Given a function $f(\mathbf{x})$ defined in a neighborhood of some point $\mathbf{x}_0$ in $I\!R^2$, and also a non zero vector $\mathbf{v}$ in $I\!R^2$, *the directional derivative of f at $\mathbf{x}_0$ in the direction $\mathbf{v}$* is defined by

$$\lim_{h\to 0} \frac{f(\mathbf{x}_0 + h\mathbf{v}) - f(\mathbf{x}_0)}{h} \ , \tag{2.6}$$

provided this limit exists. If the limit does not exist, the directional derivative does not exist.

Given f, $\mathbf{x}_0$ and $\mathbf{v}$, note that if we define

$$g(t) = f(\mathbf{x}_0 + t\mathbf{v}) \ , \tag{2.7}$$

then the directional derivative of f at $\mathbf{x}_0$ in the direction $\mathbf{v}$ is just

$$g'(0) \ .$$

This means that directional derivatives, like partial derivatives, can be computed by single familiar variable methods.

Example 4 (Slicing a function along a line) For example, if $f(x) = \dfrac{xy^2}{1 + x^2 + y^2}$, $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, then

$$g(t) = f(1+t, 1+2t) = \frac{(1+t)(1+2t)^2}{1 + (1+t)^2 + (1+2t)^2} = \frac{1 + 5t + 8t^2 + 4t^3}{3 + 6t + 5t^2} \ .$$

The result is a familiar garden variety function of a single variable t. It is a laborious but straightforward task to now compute that $g'(0) = 1$. Please do the calculation; you will then appreciate the better way of computing directional derivatives that we shall soon explain!

There is an important observation to make at this point:

•*Partial derivatives are special cases of directional derivatives – the case in which the direction vector $\mathbf{v}$ is one of the standard basis vectors $\mathbf{e}_i$.*

Please compare the definitions, and make sure that you see this point. Next, we show that the "badly behaved" function considered at the beginning of this subsection has directional derivatives *only* for directions parallel to the coordinate axes.

Example 5 (Sometimes there are directional derivatives only in special directions) Let f be the function defined in (2.5), let $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and let $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ for some numbers a and b.

The question we now ask is: *for which values of a and b does there exists the directional derivative of f at $\mathbf{x}_0$ in direction $\mathbf{v}$?*

To answer this, let's compute

$$f(\mathbf{x}_0 + h\mathbf{v}) - f(\mathbf{x}_0) \ ,$$

divide by h, and try to take the limit $h \to 0$. We find that

$$f(\mathbf{x}_0) = 0 \qquad \text{and} \qquad f(bx_0 + h\mathbf{v}) = \frac{2ab}{a^2+b^2} \ .$$

(Since by definition, the direction vector $\mathbf{v} \neq 0$, we do not divide by zero on the right.) Therefore

$$\frac{f(\mathbf{x}_0 + h\mathbf{v}) - f(\mathbf{x}_0)}{h} = \frac{1}{h}\left(\frac{2ab}{a^2+b^2}\right) \ .$$

As $h \to 0$, this "blows up", unless either $a = 0$ or $b = 0$, in which case the the right hand side is zero for every $h \neq 0$, and so the limit does exist, and is zero. Therefore, for this "bad" function, the directional derivative exists if and only if the direction vector $\mathbf{v}$ is parallel to one of the coordinate axes.

2.4 Gradients and directional derivatives

If you worked through all the calculations in Example 4, you know that computing directional derivatives "straight from the definition" as we did there can be pretty laborious. The good new is that there is a *much better* way!

In fact, it turns out that if you can compute the partial derivatives of f, *and these partials derivatives are continuous*, then f will have partial derivatives in every direction, and moreover, these directional derivatives will be certain simple linear combinations of the partial derivatives.

The problem with the "bad" function f considered in Example 5 is precisely that its partial derivatives, while they do exist, fail to be continuous. In this subsection, we will give a very useful formula for directional derivatives as linear combinations of partial derivatives. We will then explain why the formula works whenever the partial derivatives are continuous.

To express the formula for directional derivatives in a clean and clear way, we first organize the partial derivatives into a vector:

Definition (Gradient) Let f be a function on the plane having both of its partial derivatives well defined at $\mathbf{x}_0$. Then the *gradient of f at* $\mathbf{x}_0$ is the vector $\nabla f(\mathbf{x}_0)$ given by

$$\nabla f(\mathbf{x}_0) = \begin{bmatrix} \dfrac{\partial}{\partial x} f(\mathbf{x}_0) \\ \dfrac{\partial}{\partial y} f(\mathbf{x}_0) \end{bmatrix} \ .$$

Since you know how to compute partial derivatives, you know how to compute gradient: It is just a matter of listing the partial derivatives once you have computed them. Here is an example:

Example 6 (Computing a gradient) With $f(x) = \dfrac{xy^2}{1+x^2+y^2}$, we compute that

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

and

$$\frac{\partial}{\partial y} f(\mathbf{x}) = \frac{2xy(1+x^2)}{(1+x^2+y^2)^2} \ .$$

Therefore,

$$\nabla f(\mathbf{x}) = \frac{1}{(1+x^2+y^2)^2} \begin{bmatrix} y^2(1+y^2-x^2) \\ 2xy(1+x^2) \end{bmatrix} \ .$$

We can now state the theorem on differentiating $f(\mathbf{x}_0 + t\mathbf{v})$:

Theorem 1 (Directional derivatives and gradients) *Let f be any function defined in an open set U of $\mathbb{R}^2$ Suppose that both partial derivatives of f are defined and continuous at every point of U. Then for any $\mathbf{x}_0$ in U, and any direction vector $\mathbf{v}$ in $\mathbb{R}^2$,*

$$\lim_{h\to 0} \frac{f(\mathbf{x}_0 + h\mathbf{v}) - f(\mathbf{x}_0)}{h} = \mathbf{v}\cdot\nabla f(\mathbf{x}_0) \ . \tag{2.8}$$

Example 7 (Computing a directional derivative using the gradient) With $f(x) = \dfrac{xy^2}{1+x^2+y^2}$, $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ as in Example 4. In that example, we computed (the hard way) that the corresponding directional derives is 1.

But now, from Example 6, we have that

$$\nabla f(\mathbf{x}) = \frac{1}{(1+x^2+y^2)^2} \begin{bmatrix} y^2(1+y^2-x^2) \\ 2xy(1+x^2) \end{bmatrix} \ ,$$

and hence, substituting $x = 1$ and $y = 1$, we have

$$\nabla f(\mathbf{x}_0) = \frac{1}{9}\begin{bmatrix} 1 \\ 4 \end{bmatrix} \ .$$

Therefore,

$$\mathbf{v}\cdot\nabla f(\mathbf{x}_0) = 1 \ .$$

Hence, we get the same value, namely 1, for the directional derivative.

The reason that we did not already introduce a special notation for the directional derivative of f at $\mathbf{x}_0$ in the direction $\mathbf{v}$ is that Theorem 1 provides one, namely $\mathbf{v}\cdot\nabla f(\mathbf{x}_0)$. We couldn't use it in the last subsection because we hadn't yet defined gradients, but now that we have, this will be our standard notation for directional derivatives, at least when we are dealing with "nice" functions whose partial derivatives are continuous.

Note that if $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$,

$$\mathbf{v}\cdot\nabla f(\mathbf{x}_0 = a\frac{\partial}{\partial x} f(\mathbf{x}_0) + b\frac{\partial}{\partial y} f(\mathbf{x}_0)) \ .$$

The right hand side is a linear combination of the partial derivatives of f, as we indicated earlier.

Theorem 1 provides an efficient means for computing directional derivatives, because if is easy to compute partial derivatives – even if there are many variables, just one is varying at a time. Also, one you have computed the gradient, you are done with that once and for all. You can take the dot product with lots of different direction vectors and compute lots of directional derivatives without doing any more serious work. In the approach used in Example 4, you would have to start from scratch each time you considered a new direction vector.

We will prove Theorem 1 in the final subsection of this section. Before coming to that, there are some important geometric issues to discuss.

2.5 The geometric meaning of the gradient

The gradient of a function is a vector. As such, it has a *magnitude*, and a *direction.* To understand the gradient in geometric terms, let's try to understand what the magnitude and direction are telling us.

The key to this is the formula

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos(\theta) \tag{2.9}$$

which says that the dot product of two vectors in $I\!R^n$ is the product of their magnitudes times the cosine of the angle between their directions.

Now pick any point $\mathbf{x}_0$ and any unit vector $\mathbf{u}$ in $I\!R^2$. Suppose f has continuous partial derivatives at $\mathbf{x}_0$, and consider the directional derivative of f at $\mathbf{x}_0$ in the direction $\mathbf{u}$. By Theorem 1, this is

$$\mathbf{u} \cdot \nabla f(\mathbf{x}_0)\ .$$

By (2.9) and the fact that $\mathbf{u}$ is a unit vector (i.e., a pure direction vector),

$$\mathbf{u} \cdot \nabla f(\mathbf{x}_0) = |\nabla f(\mathbf{x}_0)|\cos(\theta)$$

where θ is the angle between $\nabla f(\mathbf{x}_0)$ and $\mathbf{u}$. (This is defined as long a $\nabla f(\mathbf{x}_0) \neq 0$, in which case the right hand side is zero.)

Now, as $\mathbf{u}$ varies of the unit circle $\cos(\theta)$ varies between -1 and 1. That is,

$$-|\nabla f(\mathbf{x}_0)| \leq \mathbf{u} \cdot \nabla f(\mathbf{x}_0) \leq |\nabla f(\mathbf{x}_0)|$$

Recall that by Theorem 1, $\mathbf{u} \cdot \nabla f(\mathbf{x}_0)$ is the slope at $\mathbf{x}_0$ of the slice of the graph $z = f(\mathbf{x})$ that you get when slicing along $\mathbf{x}_0 + t\mathbf{u}$. Hence we can rephrase this as

$$-|\nabla f(\mathbf{x}_0)| \ \leq \ \text{[slope of a slice at } \mathbf{x}_0] \ \leq \ |\nabla f(\mathbf{x}_0)|$$

That is,

• *The magnitude of the gradient $|\nabla f(\mathbf{x}_0)|$ tells us the minimum and maximum values of the slopes of all slices of $z = f(\mathbf{x})$ through $\mathbf{x}_0$.*

The slope has the maximal value, $|\nabla f(\mathbf{x}_0)|$, exactly when $\theta = 0$; i.e., when $\mathbf{u}$ and $\nabla f(\mathbf{x}_0)$ point in the same direction. In other words:

- *The gradient of f at $\mathbf{x}_0$ points in the direction of steepest increase of f at $\mathbf{x}_0$*

For the same reasons, we get the steepest negative slope by taking $\mathbf{u}$ to point in the direction of $-\nabla f(\mathbf{x}_0)$.

Example 8 (Which way the water runs) Let $f(x) = \dfrac{xy^2}{1+x^2+y^2}$, $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ as in Example 6. Let $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. If $z = f(\mathbf{x})$ denotes the altitude at $\mathbf{x}$, and you stood at $\mathbf{x}_0$, and spilled a glass of water, which way would the water run?

For purposes of this question, let's say that the direction of the positive x axis is due East, and the direction of the positive y axis is due North.

But now, from Example 6, we have that

$$\nabla f(\mathbf{x}) = \frac{1}{(1+x^2+y^2)^2} \begin{bmatrix} y^2(1+y^2-x^2) \\ 2xy(1+x^2) \end{bmatrix} ,$$

and hence, substituting $x = 0$ and $y = 1$, we have

$$\nabla f(\mathbf{x}_0) = \frac{1}{4} \begin{bmatrix} 2 \\ 0 \end{bmatrix} .$$

Thus, the gradient points *due East*. This is the "straight uphill" direction. The water will run in the "straight downhill" direction, which is opposite. That is the water will run due West.

2.6 A chain rule for functions of a vector variable

In this section, we shall prove Theorem 1. In fact, we shall prove something even more general and even more useful: a chain rule for functions of a vector variable.

Let $\mathbf{x}(t)$ be a differentiable vector values function in $I\!R^2$. Let f be a function from $I\!R^2$ to $I\!R$. Consider the composite function $g(t)$ defined by

$$g(t) = f(\mathbf{x}(t)) .$$

Here we ask the question:

- *Under what conditions on f is g differentiable, and can we compute $g'(t)$ in terms of $\mathbf{x}'(t)$ and $\nabla f(\mathbf{x})$?*

Theorem 2 (The chain rule for functions from $I\!R^2$ to $I\!R$ *Let f be any function defined in an open set U of $I\!R^2$ Suppose that both partial derivatives of f are defined and continuous at every point of U. Let $\mathbf{x}(t)$ be a differentiable function from $I\!R$ to $I\!R^2$. Then, for all values of t so that $\mathbf{x}(t)$ lies in U,*

$$\lim_{h\to 0} \frac{f(\mathbf{x}(t+h)) - f(\mathbf{x}(t))}{h} = \mathbf{x}'(t) \cdot \nabla f(\mathbf{x}(t)) . \tag{2.10}$$

Note that Theorem 2 reduces to Theorem 1 if we consider the special case in which $\mathbf{x}(t)$ is just the line $\mathbf{x}_0 + t\mathbf{v}$, and we evaluate at the derivative at $t = 0$. The special case is used often enough that it deserves to be stated a a separate Theorem. Still, it is no harder to prove Theorem 2, and we shall also make frequent use of the more general form.

The key to the proof, in which we finally explain the importance of continuity for the partial derivatives if the *Mean Value Theorem* from single variable calculus:

The Mean Value Theorem says that if $g(s)$ has a *continuous* first derivative $g'(s)$, then for any numbers a and b, with $a < b$, there is a value of c in between; i.e., with $a < c < b$

$$\frac{g(b) - g(a)}{b - a} = g'(c) \ . \tag{2.11}$$

The principle expressed here is the one by which the police know that if you drove 100 miles in one hours, then at some point on your trip, you were driving at 100 miles per hour.

Proof of Theorem 2: Fix some t, and some $h > 0$. To simplify the notation, define the numbers x_0, y_0, x_1 and y_1 by

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \mathbf{x}(t) \qquad \text{and} \qquad \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \mathbf{x}(t+h) \ .$$

To enable the analysis of $f(\mathbf{x}(t+h)) - f(\mathbf{x}(t))$ by single variable methods, note that

$$\begin{aligned} f(\mathbf{x}(t+h)) - f(\mathbf{x}(t)) &= f(x_1, y_1) - f(x_0, y_0) \\ &= \left[f(x_1, y_1) - f(x_0, y_1)\right] + \left[f(x_0, y_1) - f(x_0, y_0)\right] \ . \end{aligned}$$

Notice that in going from the first line to the second, we have subtracted and added back in the quantity $f(x_0, y_1)$, and grouped the terms in brackets.

In the first group, only the x variable is varying, and in the second group, only the y variable is varying. Thus, we can use single variable methods on these groups.

To do this for the first group, define the function $g(s)$ by

$$g(s) = f(x_0 + s(x_1 - x_0), y_1) \ .$$

Notice that

$$g(1) - g(0) = f(x_1, y_1) - f(x_0, y_1) \ .$$

Then, if g is continuously differentiable, the Mean Value Theorem tells us that

$$f(x_1, y_1) - f(x_0, y_1) = \frac{g(1) - g(0)}{1 = 0} = g'(c)$$

for some c between 0 and 1.

But by the definition of $g(s)$, we can compute $g'(s)$ by taking a partial derivative of f, since as s varies, only the x component of the input to f is varied. Thus,

$$g'(s) = \frac{\partial f}{\partial x}(x_0 + s(x_1 - x_0), y_1)(x_1 - x_0) \ .$$

Therefore, for some c between 0 and 1,

$$\big[f(x_1, y_1) - f(x_0, y_1)\big] = \left[\frac{\partial f}{\partial x}(x_0 + c(x_1 - x_0), y_1)\right](x_1 - x_0) \ .$$

In the exact same way, we deduce that for some $\tilde{c}$ between 0 and 1,

$$\big[f(x_0, y_1) - f(x_0, y_0)\big] = \left[\frac{\partial f}{\partial y}(x_0, y_0 + \tilde{c}(y_1 - y_0))\right](y_1 - y_0) \ .$$

Therefore,

$$\begin{aligned}\frac{f(\mathbf{x}(t+h)) - f(\mathbf{x}(t))}{h} &= \left[\frac{\partial f}{\partial x}(x_0 + c(x_1 - x_0), y_1)\right]\frac{x_1 - x_0}{h} \\ &+ \left[\frac{\partial f}{\partial y}(x_0, y_0 + \tilde{c}(y_1 - y_0))\right]\frac{y_1 - y_0}{h} \ .\end{aligned}$$

Up to now, h has been fixed. But having derived this identity, it is now easy to analyze the limit $h \to 0$.

First, as $h \to 0$, $x_1 \to x_0$ and $y_1 \to y_0$. Therefore,

$$\lim_{h\to 0}\frac{\partial f}{\partial x}(x_0 + c(x_1 - x_0), y_1) = \frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial x}(\mathbf{x}(t)) \ ,$$

and

$$\lim_{h\to 0}\frac{\partial f}{\partial y}(x_0, y_0 + \tilde{c}(y_1 - y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = \frac{\partial f}{\partial y}(\mathbf{x}(t)) \ .$$

Also, since $\mathbf{x}(t)$ is differentiable

$$\lim_{h\to 0}\frac{x_1 - x_0}{h} = \lim_{h\to 0}\frac{x(t+h) - x(t)}{h} = x'(t)$$

and

$$\lim_{h\to 0}\frac{y_1 - y_0}{h} = \lim_{h\to 0}\frac{y(t+h) - y(t)}{h} = y'(t)$$

Since the limit of a product is the product of the limits,

$$\lim_{h\to 0} \frac{f(\mathbf{x}(t+h)) - f(\mathbf{x}(t))}{h} = \left[\frac{\partial f}{\partial x}(\mathbf{x}(t))\right] x'(t) + \left[\frac{\partial f}{\partial y}(\mathbf{x}(t))\right] y'(t)$$
$$= \nabla f(\mathbf{x}(t)) \cdot \mathbf{x}'(t) \ .$$

This is what we had to show. ■

Exercises

1. Compute the gradients of $f(x,y)$ and $h(x,y)$ where $f(x,y) = 2x^2 + xy + y^2$ and $h(x,y) = \sqrt{f(x,y)}$. Also compute $\nabla f(1,1)$ and $\nabla h(1,1)$.

2. Compute the gradients of $f(x,y)$ and $h(x,y)$ where $f(x,y) = 2x^2 + x\cos(y)$ and $h(x,y) = \sin(f(x,y))$. Also compute $\nabla f(0,0)$ and $\nabla h(0,0)$.

3. (a) Let f and g be two functions on $I\!R^2$ such that each has a well defined gradient. Show that

$$(\nabla fg)(x) = f(x)(\nabla g(x)) + g(x)(\nabla f(x)) \ .$$

(This is the *product rule* for the gradient.)
(b) Let $f(x,y) = x\cos(x^2y)$, $g(x,y) = \sqrt{x^2+y^2}$ and $h(x,y) = f(x,y)g(x,y)$. Compute $\nabla f(x,y)$, $\nabla g(x,y)$ and $\nabla h(x,y)$.

4. Let $f(x,y) = xy - x^2y^2$, $g(x,y) = e^{-(x^2+y^2)}$ and $h(x,y) = f(x,y)g(x,y)$. Compute $\nabla f(x,y)$, $\nabla g(x,y)$ and $\nabla h(x,y)$.

5. Show that if g is a differentiable function of one variable, and h is a function of two variables that has a gradient at (x_0, y_0), then so does $f(x,y) = g(h(x,y))$, and

$$\nabla f(x_0,y_0) = g'(h(x_0,y_0))\nabla h(x_0,y_0) \ .$$

6. (a) The distance of (x,y) from the origin $(0,0)$ is $\sqrt{x^2+y^2}$. A function f on $I\!R^2$ is called *radial* in case f depends on (x,y) only through $\sqrt{x^2+y^2}$, which means that there is a single variable function g so that

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

Show that if $g(z)$ is differentiable for all z, then f has a gradient at all points (x,y) except possibly $(0,0)$.
(b) Show that if $g'(0) = 0$, then both partial derivatives of f and the gradient of f are well defined also at $(0,0)$.

7. Let f be the "mountain landscape" function

$$f(x,y) = \frac{3(1+x)^2 + xy^3 + y^2}{1+x^2+y^2}$$

considered at the beginning of this section. If you stood in this landscape at the point with horizontal coordinates $(x,y) = (-1,2)$ and spilled a glass of water, in which direction (more or less) would it run: North, Northeast, East, Southeast, South, Southwest, West, or Northwest?

8. Let f be the "mountain landscape" function

$$f(x,y) = \frac{3(1+x)^2 + xy^3 + y^2}{1 + x^2 + y^2}$$

considered at the beginning of this section. If you stood in this landscape at the point with horizontal coordinates $(x, y) = (-1, -2)$ and spilled a glass of water, in which direction (more or less) would it run: North, Northeast, East, Southeast, South, Southwest, West, or Northwest?

9. $f(x, y) = 2x^2 + xy^3 + y^2$, $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Define $g(t) = f(\mathbf{x}_0 + t\mathbf{v})$. Compute an explicit formula for $g(t)$, and using this compute its derivative at $t = 1$. Also use Theorem 1 to do this computation using the gradient of f.

10. $f(x, y) = \cos(x^2 y + yx^2)$, $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Define $g(t) = f(\mathbf{x}_0 + t\mathbf{v})$. Compute an explicit formula for $g(t)$, and using this compute its derivative at $t = 1$. Also use Theorem 1 to do this computation using the gradient of f.

11. Let $f(x, y) = 2x^2 + xy^3 + y^2$, and let $\mathbf{x}(t) = \begin{bmatrix} t^2 + t \\ 1/(1+t^2) \end{bmatrix}$. Use Theorem 2 to compute $g'(t)$ where $g(t) = f(\mathbf{x}(t))$.

12. Let $f(x, y) = \cos(x^2 y + yx^2)$, and let $\mathbf{x}(t) = \begin{bmatrix} t^2 + t \\ 1/(1+t^2) \end{bmatrix}$. Use Theorem 2 to compute $g'(t)$ where $g(t) = f(\mathbf{x}(t))$.

13. Let $\mathbf{x}(t)$ be a differentiable function form $\mathbb{R}$ to $\mathbb{R}^2$. Let f be a functions from $\mathbb{R}^2$ to $\mathbb{R}$ with continuous partial derivatives. Suppose that for some value t_0, it is the case that

$$f(\mathbf{x}(t_0)) \geq f(\mathbf{x}(t)) \qquad \text{for all} \quad t\ .$$

Show that then it must be the case that $\nabla f(\mathbf{x}(t_0)$ and $\mathbf{x}'(t_0)$ are orthogonal.

Section 3: Level curves

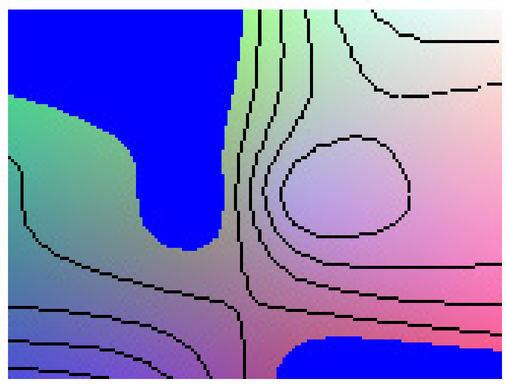
3.1: Horizontal slices and contour curves

In Sections 1 and 2, we have considered *vertical* slices of the graph of $z = f(x, y)$. We can gain a new perspective by considering *horizontal* slices.

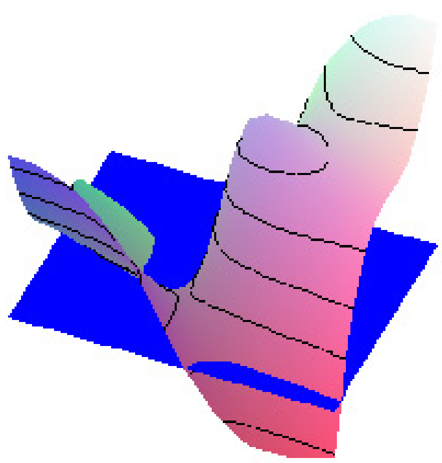
Consider once again the "mountain landscape" graphed in the first section:

$$f(x, y) = \frac{3(1 + x)^2 + xy^3 + y^2}{1 + x^2 + y^2} \tag{3.1}$$

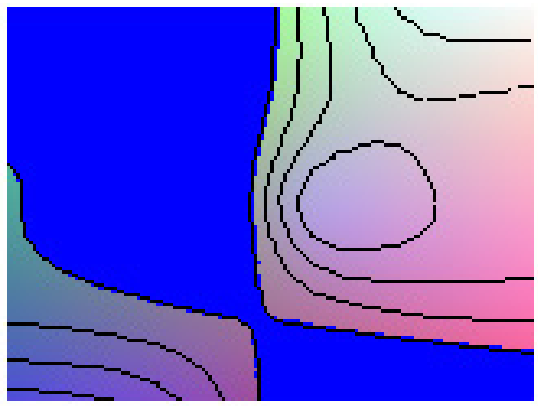
Suppose that a dam is built, and this landscape is flooded, up to an altitude 0.5 in the vertical distance units. This produces a lake that is shown below, in a top view; i.e., an aerial or satellite image:



The other lines of the land are the lines at other constant altitudes, specifically $x = 1.5$, $z = 2.5$, $z = 3.5$ and so on. These curves are called *contour curves.* Here is a sort of side view showing the lake as a horizontal "slice" through the graph $z = f(x, y)$ at height $z = 1.5$:



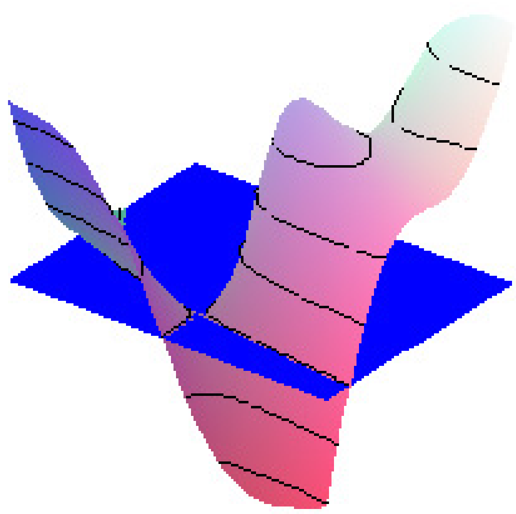
If the water level is raised further, say to the altitude $z = 1.5$, everything will be flooded up to the next contour curve:



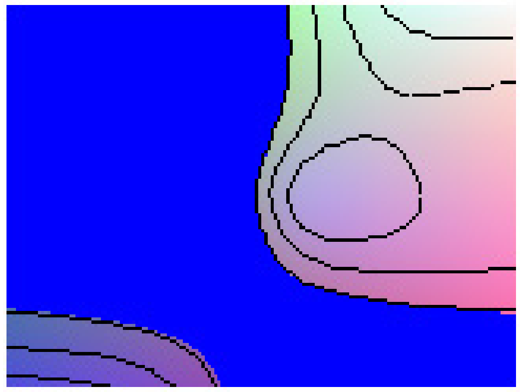
Comparing with the first picture, you clearly see that everything has been flooded up to the $z = 1.5$ contour curve. The isthmus joining the two tall hills is now submerged, and the two regions of the lake in the first graph have merged.

If you walked along the lake shore, your path would trace out the contour curve at $z = 1.5$ in the first picture.

Here is a side view showing the lake at this level. It shows it as a horizontal "slice" through the graph $z = f(x, y)$ at height $z = 1.5$:



If the water level is raised further, to the height $z = 2.5$, the shore line moves up to the next contour line. Now a walk along the shoreline would trace out the path along the $x = 2.5$ contour line in the first picture. Here is the top view showing the lake at this stage:



The contour curves, which are the results of horizontal slices of the graph of $z = f(x, y)$, tell us a lot about the function $f(x, y)$. This section is an introduction to what they tell us.

Definition (level set) Let $f(x, y)$ be a function on $I\!R^2$, and let c be any number. The set of points (x, y) satisfying

$$f(x, y) = c$$

is called the *level set* of f at c

If we think of $f(x, y)$ as representing the altitude at the point with coordinates (x, y), then the level set of f at height c is the set of all points at which the altitude is c. The level set at height c would be the "shore line" curve if the landscape were flooded up to an altitude c.

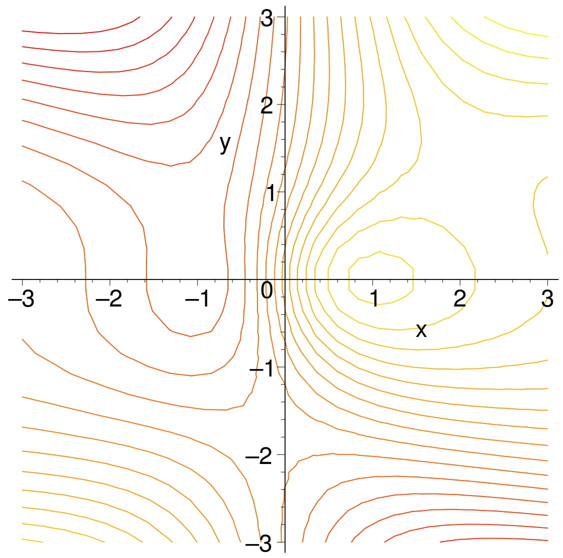
Now, here is a very important point, whose validity you can more or less see from the pictures we have displayed:

• *Under normal circumstances, the level set of f at c will be a curve in the plane, possible with several disconnected components.*

It is for this reason that we often refer to level sets as contour curves.

We can plot a number of the level sets on a common graph. A *contour plot* of a function $f(x, y)$ is graph in which level curves of f are plotted at several different "altitudes" $c_1, c_2, c_3, \ldots$. You have probably seen these on maps for hiking.

Here is a contour plot for the function "mountain landscape" function $f(x, y)$ in (3.1):



3.2: Implicit and explicit descriptions of planar curves

How could one go about actually drawing the contour curves starting from a formula like (3.1)? That is not so easy in general. You can see a hint of this in the convoluted form of the contour curves plotted here. The difficulty lies here:

• *The description of contour curves given by the defining equation $f(x, y) = c$ is, alas, just an implicit description. However, plotting a curve is a simple matter only when one has an explicit description.*

To really appreciate this point, one has to understand the distinction between an implicit and and explicit description of a curve. The unit circle is a great example with which to start.

Let $f(x, y) = x^2 + y^2$ and let $c = 1$. Then the level set of f at height c is the set of points (x, y) satisfying

$$x^2 + y^2 = 1\ . \tag{3.2}$$

This set, of course, is the unit circle. If we drew a contour plot of f showing the level curves at several altitudes "altitudes" $c_1, c_2, c_3, \ldots$, you would see, several concentric circles.

The equation (3.2) is the *implicit* equation for the unit circle. To get an *explicit description*, just solve the equation (3.2) for y as a function of x. The result is

$$y(x) = \pm\sqrt{1 - x^2}\ . \tag{3.3}$$

There are two "branches", corresponding to the two signs.

Given y as an explicit function of x, it is easy to plot the curve. For instance, taking the "upper branch" $y = \sqrt{1 - x^2}$ from (3.3), and plugging in a sequence of values for x, we get the table

$$
\begin{array}{rl}
x = 0 & \qquad y = 1 \\
x = 1/8 & \qquad y = \sqrt{63}/8 \\
x = 1/4 & \qquad y = \sqrt{15}/4 \\
x = 3/8 & \qquad y = \sqrt{55}/8 \\
x = 1/2 & \qquad y = \sqrt{3}/2 \\
x = 5/8 & \qquad y = \sqrt{39}/8
\end{array}
$$

and so forth. Connecting the dots, we get a good picture of part of the curve. (To get the rest, just extend the table so that the x values range all the way from -1 to 1, and do the same for the "lower branch" $y = -\sqrt{1-x^2}$).

As long as one variable is given as an *explicit* function of another, or several explicit functions in case there are several branches, one can produce the graph by making a table and connecting the dots. This is what is nice about "explicitly defined curves".

There is a more general kind of explicit description – *parametric*. As t varies between 0 and 2π,

$$\begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} \tag{3.4}$$

traces out the unit circle in the counterclockwise direction. This is another sort of explicit description since if you plug in any value of t, you get a point (x, y) on the unit circle, and as you vary t, you "sweep out" all such points. Again, there are no equations to solve, just computations to do.

If we take $t = x$, we can rewrite (3.3) as

$$x(t) = t \qquad \text{and} \qquad y = \pm\sqrt{1-t^2} \qquad \text{for} \quad -1 \le t \le 1\ , \tag{3.5}$$

so (3.3) is just a particular parametric representation of the circle in which we use x itself as the parameter.

Definition (Implicit and explicit descriptions of curves) An equation of the form $f(x, y) = c$ provides an *implicit* description of a curve. A parameterization $\mathbf{x}(t)$, possibly with $t = x$ and $y(x)$ given as an explicit function of x, provides an *explicit* description of a curve.

Once one has an explicit description, it is easy to generate a plot, just by plugging in values for the parameter, plotting the resulting points, and "connecting the dots". Passing from an implicit description to an explicit description involves *solving the equation* $f(x, y) = c$ to find an explicit description. Generally, that is easier said than done.

Example 1 (From implicit to explicit by means of algebra) Consider the function

$$f(x, y) = 2x^2 - 2xy + y^2\ .$$

The level curve at $c = 1$ for this function is given implicitly by the equation

$$2x^2 - 2xy + y^2 = 1 \ .$$

This can be rewritten as

$$y^2 - 2xy = 1 - 2x^2 \ .$$

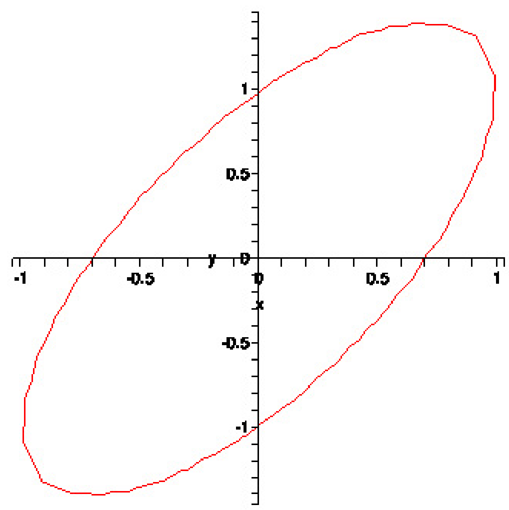
Completing the square in y, we have

$$(y - x)^2 = 1 - 2x^2 + x^2 = 1 - x^2 \ .$$

Therefore, we can solve for y as a function of x, finding

$$y(x) = x \pm \sqrt{1 - x^2} \ .$$

If we take x as the parameter, evidently y has a real value only for $-1 \leq x \leq 1$. It is now easy to plot the contour curve:



In this example, it was not so horrible passing from an implicit description to an explicit description; i.e., a parameterization, since the equation $f(x, y) = 1$ was quadratic in both x and y. We know how to deal with quadratic equations, so in this case, we were able to make the transition from implicit to explicit.

However, in general, this will not be possible to do. In general, we are going to need to extract information on the contour curves *directly* from the implicit description. Fortunately, what we have learned about gradients can help us to do this!

3.3 The direction of a contour curve as it passes through a point – tangent lines

Let $\mathbf{x}(t)$ be a parameterization of a contour curve. As this curve passes through the point $\mathbf{x}(t_0)$, supposing it does so in a reasonably nice way, one can ask about the *direction of motion* as the contour curve passes though the point $\mathbf{x}(t_0)$. We can answer this question using Theorem 2 of the previous section.

Suppose $\mathbf{x}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$ is *some* parameterization of the contour curve of f through c. Then, by definition,

$$g(t) = f(\mathbf{x}(t)) = c$$

for all t.

Since $g(t)$ is constant, $g'(t) = 0$. But by Theorem 2 of the previous section, $g'(t) = \mathbf{x}'(t) \cdot \nabla f(\mathbf{x}(t))$, and hence

$$\mathbf{x}'(t) \cdot \nabla f(\mathbf{x}(t)) = 0$$

for all t.

In other words, the velocity vector $\mathbf{x}'(t)$ is orthogonal to the gradient $\nabla f(\mathbf{x}(t))$ at $\mathbf{x}(t)$ for each t.

• *The direction of motion along the level curve of f passing through $\mathbf{x}_0$ is orthogonal to $\nabla f(\mathbf{x}_0)$.*

This observation brings us to the notion of the *tangent line of a differentiable curve.*

Definition (tangent line) Let $\mathbf{x}(t)$ be a differentiable curve in $I\!R^2$. The *tangent line* of this curve at t_0 is the line

$$\mathbf{x}(t_0) + s\mathbf{x}'(t_0) \ ,$$

where s is the parameter, $\mathbf{x}(t_0)$ is the base point, and $\mathbf{x}'(t_0)$ is the velocity vector.

Example 2 (Computing a tangent line) Consider the parameterized curve given by (3.4). We easily compute that

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} \ .$$

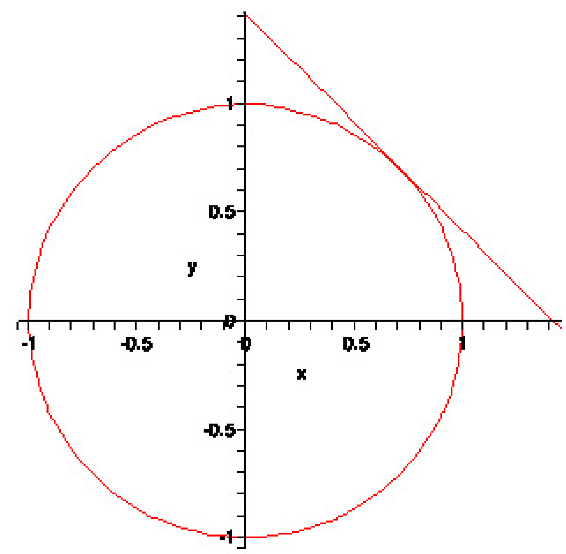
For $t_0 = \pi/4$, we have

$$\mathbf{x}(t_0) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \ ,$$

and

$$\mathbf{v}(t_0) = \mathbf{x}'(t_0) = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \ .$$

Here is a graph showing the curve, $\mathbf{x}(t_0)$ and $\mathbf{v}(t_0) = \mathbf{x}'(t_0)$.



The velocity vector depends not only on the curve, but also on the particular parameterization of it. For example,

$$\mathbf{x}(t) = \begin{bmatrix} \cos(-2t) \\ \sin(-2t) \end{bmatrix}$$

is another parameterization of the circle; it sweeps over the circle clockwise, and twice as fast as the one in (3.3). It passes through $\dfrac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ at $t_0 = 3\pi/8$, and you can compute that this time the velocity is $\mathbf{v}(t_0) = \sqrt{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Here is the picture:

Using two different parameterization, we have computed two different velocity vectors for motion along the circle at the point $\dfrac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. These two velocity vectors do have something important in common – they lie on the same line. The line itself is characteristic of the curve itself – it does not depend on the parameterization. It gives the *angle* at which the contour curve passes through the point in question.

Moving now from the particular case of the circle to the general situation, think of a point particle moving in the plane $\mathbb{R}^2$ with its position at time t being $\mathbf{x}(t)$. Suppose that the motion is smooth enough that $\mathbf{x}'(t) = \mathbf{v}(t)$ exists at each t. Fix any t_0, and consider the line parameterized by

$$\mathbf{x}(t_0) + s\mathbf{v}(t_0)\ .$$

This describes the straight line motion that passes through $\mathbf{x}(t_0)$ at $s = 0$ with the same velocity as the curve $\mathbf{x}(t)$. This is the *tangent line* to the curve at this point.

Example 3 (Computing another tangent line) Consider the parameterized curve given by $\mathbf{x}(t) = \begin{bmatrix} t^2 - t \\ t - t^3 \end{bmatrix}$. What is the tangent line to this curve at $t_0 = 2$? We readily compute that $\mathbf{x}(2) = \begin{bmatrix} 2 \\ -6 \end{bmatrix}$ and that $\mathbf{v}(2) = \begin{bmatrix} 3 \\ -11 \end{bmatrix}$. Hence the line in question is given by

$$\begin{bmatrix} 2 \\ -6 \end{bmatrix} + s \begin{bmatrix} 3 \\ -11 \end{bmatrix} \ .$$

As you see, computing tangent lines is a straightforward thing to do as long as you have an explicit parameterization of the curve in question. But what if you do not?

3.4: Computing tangent lines for implicitly defined curves

It is actually quite simple to compute the tangent line to an implicitly defined curve through a point *directly from the implicit description*, without first finding a parameterization. Here is how: Recall that the equation of a line in the plane has the form

$$\mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) = 0 \ ,$$

where $\mathbf{x}_0$ is any base point on the line, and $\mathbf{a}$ is any non zero vector orthogonal to the line. Now, as we have observed, the gradient vector $\nabla f(\mathbf{x}_0)$ is orthogonal to the contour curve, and hence to the tangent line, and thus we may take $\mathbf{a} = \nabla f(\mathbf{x}_0)$. This gives us the equation of the tangent line!

Example 4 (Computing a tangent line from an implicit description) Consider the implicit description of the unit circle given by the equation $x^2 + y^2 = 1$, and the point $\mathbf{x}_0 = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, which is on this circle.

We compute $\nabla f(\mathbf{x}_0) = \sqrt{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then with $\mathbf{a} = \nabla f(\mathbf{x}_0)$, $\mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) = 0$ is

$$\sqrt{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \left(\begin{bmatrix} x \\ y \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = (\sqrt{2}x - 1) + \sqrt{2}y - 1) = 0 \ ,$$

and so the equation of the tangent line is

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

In summary:

• *If there is a differentiable level curve of f, $\mathbf{x}(t)$, passing through $\mathbf{x}_0$ at $t = t_0$, $\nabla f(\mathbf{x}_0)$ is orthogonal to $\mathbf{x}'(t_0)$, and therefore to the tangent line to the level curve of f through $\mathbf{x}_0$.*

The point of this is that while you need to know an explicit parameterization to compute $\mathbf{x}'(t_0)$, you do not need this to compute $\nabla f(\mathbf{x}_0)$ – computing $\nabla f(\mathbf{x}_0)$ gives you a direction orthogonal to the tangent line, which is all we need to write down the equation of the line.

For a parametric description, since we already know the base point $\mathbf{x}_0$, we only need to know the *direction* of the line. As we have just seen, this is orthogonal to $\nabla f(\mathbf{x}_0)$. It is easy to see that if $\nabla f(\mathbf{x}_0) = \begin{bmatrix} p \\ q \end{bmatrix}$, then we can take $\mathbf{v} = \begin{bmatrix} -q \\ p \end{bmatrix}$. Indeed, no matter what p and q are,

$$\begin{bmatrix} p \\ q \end{bmatrix} \cdot \begin{bmatrix} -q \\ p \end{bmatrix} = -pq + qp = 0 \ .$$

This simple device is very useful, and so we make a definition:

Definition ($\mathbf{v}^\perp$) For any vector $\mathbf{v} = \begin{bmatrix} p \\ q \end{bmatrix}$ in $I\!R^2$, the vector $\mathbf{v}^\perp$ is defined by $\mathbf{v}^\perp = \begin{bmatrix} -q \\ p \end{bmatrix}$.

Example 5 (Tangent line to a level curve) Let $f(x,y) = (x^2 - y^2)^2 - 2xy$. We will now compute the tangent line to the level curve of f through $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. We first work out that

$$\nabla f(x,y) = \begin{bmatrix} 4x(x^2 - y^2) - 2y \\ -4y(x^2 - y^2) - 2x \end{bmatrix} .$$

Therefore,

$$\nabla f(1,2) = \begin{bmatrix} -16 \\ 22 \end{bmatrix} .$$

The equation for the tangent line is

$$0 = \begin{bmatrix} -16 \\ 22 \end{bmatrix} \cdot \begin{bmatrix} x-1 \\ y-2 \end{bmatrix} = -16(x-1) + 22(y-2) ,$$

or

$$-16x + 22y = 14 .$$

To find the parametric form, we compute

$$\mathbf{v} = \mathbf{x}' = (\nabla f(\mathbf{x}_0))^\perp = \begin{bmatrix} -16 \\ 22 \end{bmatrix}^\perp = -\begin{bmatrix} 22 \\ 16 \end{bmatrix} .$$

Hence the tangent line is

$$\mathbf{x}_0 + t\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - t\begin{bmatrix} 22 \\ 16 \end{bmatrix} .$$

3.5: Is the contour curve always a curve?

All of our conclusions so far in this discussion were predication on the assumption that there was a differentiable level curve of f through $\mathbf{x}_0$. Was this a reasonable assumption? more to the point, can we check the validity of this assumption?

It turns out that if both of the partial derivatives of f are continuous in a neighborhood of $\mathbf{x}_0$, and $\nabla f(\mathbf{x}_0) \neq 0$, then at least nearby $\mathbf{x}_0$, this level set is actually a curve that has a parameterization $\mathbf{x}(t)$, and moreover, $\mathbf{x}(t)$ is itself differentiable.

The theorem that guarantees this is called *Implicit Function Theorem*. We will prove it later on, but what it is saying can be pictured easily enough now. The point is that if the gradient is not zero, then there is a well defined “uphill” direction, given by the gradient, and perpendicular to this are the two “level” directions that are neither uphill nor downhill. The level curve through $\mathbf{x}_0$ is just the contour curve through $\mathbf{x}_0$, and it must proceed away from $\mathbf{x}_0$ in the level directions. If you were standing on the landscape described by $z = f(x,y)$, you would not need to solve anything to walk along the level curve: Just step away from $\mathbf{x}_0$, and walk along without changing your altitude. You would be walking along the the curve whose existence is asserted by the Implicit Function Theorem.

However, if the gradient were zero, in which direction would you walk? In this case there might be no preferred direction, and hence no path. The level set is not always a curve, even when it is not empty.

Example 6 (A level set that is a single point) Let $f(x, y) = x^2 + y^2$. Then the level set of f at height 0, which is the level set of f passing through $(0, 0)$ is just the single point $(0, 0)$. It is not a curve. At every other point (x, y), f takes on a strictly positive value. Indeed, for this reason, you see that the level set of f at height -1, or any other strictly negative value, is empty. Notice that $\nabla f(0, 0)$ is zero, as the Implicit Function Theorem says it must be.

Example 7 (A level set that is not a simple curve) Let $f(x, y) = xy$. Then the level set of f at height 0, which is the level set of f passing through $(0, 0)$, consists of both the x and y axes. Here there are *two* level directions, and the level set of f passing through $(0, 0)$ is not a curve, but a union of two curves. Again, notice that $\nabla f(0, 0)$ is zero, as the Implicit Function Theorem says it must be.

3.6: Points on a contour curve where the tangent line has a given slope

Consider the curve given implicitly by

$$x^4 + y^4 + 4xy = 0 \ .$$

This is quartic, so it is not so easy to express y as a function of x (though it can be done). A good way to graph this curve is to notice that it is the level set of $f(x, y) = x^4 + y^4 + 4xy$ at $c = 0$. We can use what we have learned to find al points (x, y) on the curve at which the tangent line through then has a given slope s. For example, when $s = 0$, the line is horizontal. When $s = \pm\infty$, it is vertical. When $s = 1$, the tangent line runs at an angle of $\pi/4$ with respect to the x–axis. To graph the contour curve, one can find all such points, and draw a small bit if a line segment through them at the corresponding slope. Connecting these up, one has a sketch of the curve.

Example 8 (Points on a level curve where the tangent has a given direction) Let $f(x, y) = x^4 + y^4 + 4xy$, and consider the level curve given in implicit form by

$$x^4 + y^4 + 4xy = 0 \ . \tag{3.6}$$

We will now find all points on this level curve at which the tangent is horizontal; i.e., parallel to the x axis.

We compute that

$$\nabla f(x, y) = 4 \begin{bmatrix} x^3 + y \\ y^3 + x \end{bmatrix} \ .$$

Notice that the gradient is zero exactly when

$$x^3 + y = 0 \qquad \text{and} \qquad y^3 + x = 0 \ .$$

The first equation says that $y = -x^3$. substituting this into the second equation, we eliminate y and get $x^9 + x = 0$. One solution is $x = 0$. If $x \neq 0$, we can divide by x and get the equation $x^8 + 1 = 0$, which clearly has no solution. Hence, for any solution $x = 0$. But then from $x^3 + y = 0$ we see that $y = 0$ too. Hence the only point at which the gradient is zero is $(0, 0)$.

At all other points, the level set specified implicitly by (3.6) is a differentiable curve – by the Implicit Function Theorem – though here it may be somewhat more complicated.

Now let's focus on points other than $(0,0)$, and see at which of the the tangent line is horizontal. The tangent to the level curve is horizontal exactly when the perpendicular direction is vertical. The perpendicular direction is the direction of the gradient, so the tangent line is horizontal only in case the first component of the gradient is zero; i.e.,

$$x^3 + y = 0 \ . \tag{3.7}$$

The equation (3.7), together with (3.6), gives us a nonlinear system of equations for the points we seek. To solve it, notice that while it is not at all easy to eliminate either variable in the original equation (3.6), it is obvious from (3.7) that

$$y = -x^3 \ . \tag{3.8}$$

Substituting this into (3.6), we get

$$x^{12} - 3x^4 = 0 \ .$$

This has exactly 3 solutions: $x = 0$, $x = 3^{1/8}$ and $x = 3^{-1/8}$. Going back to (3.8), we can now easily find the corresponding points. They are:

$$(0,0) \qquad (3^{1/8}, -3^{3/8}) \qquad \text{and} \qquad (-3^{1/8}, 3^{3/8}) \ .$$

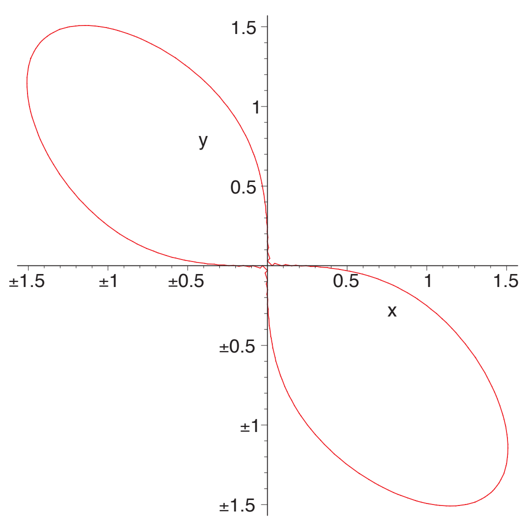
At the first of these, the gradient is zero. That is, $\nabla f(0,0)$ is zero, and so it has no direction at all. At the other two, we can be sure that the level set is a differentiable curve, and that its tangent vector is vertical.

Deciding what actually happens at $(0,0)$ is more subtle. For x and y very close to zero, $x^4 + y^4 + 4xy \approx 4xy$, and so for such x and y, the level set of f at height 0 looks pretty much like the level set of $4xy$ at height zero. This is easy to find: $4xy = 0$ if and only if $x = 0$ or $y = 0$. Hence this level set consists of *two* lines, namely the x and y axes, as in Example 7. At this point, there is a branch of the level set that is vertical, but also a branch that is horizontal, so we cannot properly say that the level curve has a horizontal tangent here.

However, using the same procedure that we used to find the points of vertical tangency, we find that the level curve of f at height 0 has a horizontal tangent exactly at

$$(3^{1/8}, -3^{3/8}) \qquad \text{and} \qquad (-3^{1/8}, 3^{3/8}) \ .$$

By doing this for a few more slopes, and connecting the points up, you could get a pretty good sketch of the curve. Here is such a sketch:



We close with an important observation that can be made in the last example. Notice that the level curve crosses itself at $(0,0)$. At any such point $\mathbf{x}_0$, it must be the case that

$\nabla f(\mathbf{x}_0) = 0$. The reason is that $\nabla f(\mathbf{x}_0)$ must be orthogonal to both of the level curves passing through $\mathbf{x}_0$, and the only vector that is orthogonal to two directions – or more – at once is the zero vector.

Exercises

1. (a) Find an explicit representation of the level set $f(x,y) = c$ for $f(x,y) = xy^2$ and $c = -1$, in the form $y(x)$, and sketch this curve.

(b) Notice that $f(-1,1) = -1$, so this point is on the curve from part **(a)**. Compute the tangent line to the curve at this point, and draw it in on your sketch from part **(a)**.

2. (a) Find an explicit representation of the level set $f(x,y) = c$ for $f(x,y) = y^2 - xy$ and $c = -1$, in the form $y(x)$, and sketch this curve.

(b) Notice that $f(2,1) = -1$, so this point is on the curve from part **(a)**. Compute the tangent line to the curve at this point, and draw it in on your sketch from part **(a)**.

3. Define $f(x,y)$ to be the imaginary part of $(x+iy)^n$. Describe the level set of f at height 0. How many branches does it have? Draw a sketch for $n = 3$.

4. Consider the curve given implicitly by $f(x,y) = 1$ where $f(x,y) = x^4 + y^4$. Find all points on this curve at which the tangent line is vertical, horizontal, has slope -1 or has slope $+1$. Draw these points and the tangent lines through them on a common graph. Using this "frame", draw a sketch of the curve.

Section 4: The tangent plane

Section 4.1: Finding the equation of the tangent plane

For a function $g(x)$ of a single variable x, the tangent line to the graph of g at x_0 is the line that "best fits" the graph $y = g(x)$ at x_0. It is the line that passes through the point $(x_0, g(x_0))$ with the same slope at the graph of $y = g(x)$ at x_0. Hence, it is given by the formula

$$y = g(x_0) + g'(x_0)(x - x_0) \ .$$

Since the graph of $y = g(x)$ and the tangent line to this graph at x_0 pass through the *exact same point* with the *exact same slope*, if you "zoom in" really close around a picture of both of them as they pass through $(x_0, g(x_0)$, you will not be able to distinguish between the two of them.

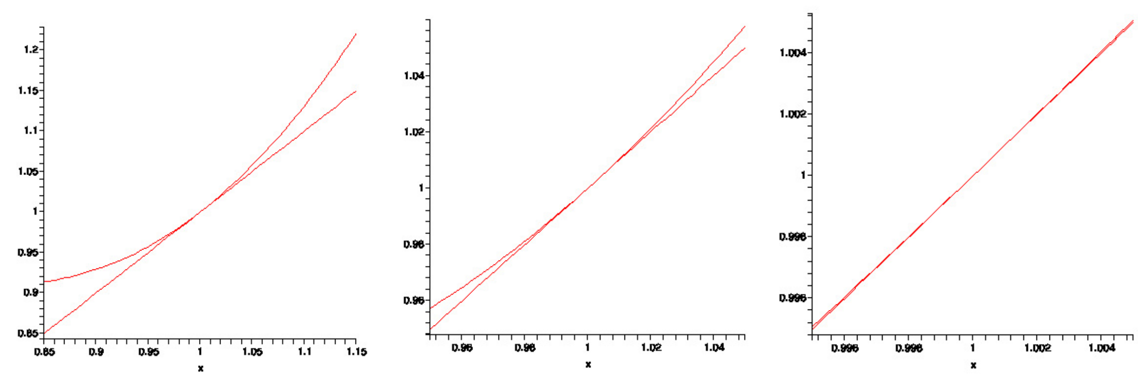
Example 1 (How well tangent lines fit) Consider $g(x) = x^3 - 2x + 2$, and $x_0 = 1$. The tangent line to the graph of $y = g(x)$ at $x_0 = 1$ is the graph of

$$y = g(1) + g'(1)(x - 1) \ .$$

Since $g(1) = 1$ and $g'(1) = 1$, the tangent line is the graph of

$$y = 1 + (x - 1) = x \ .$$

Here are three graphs of $y = g(x)$ that are "zoomed in" closer and closer about the point $(1, 1)$:

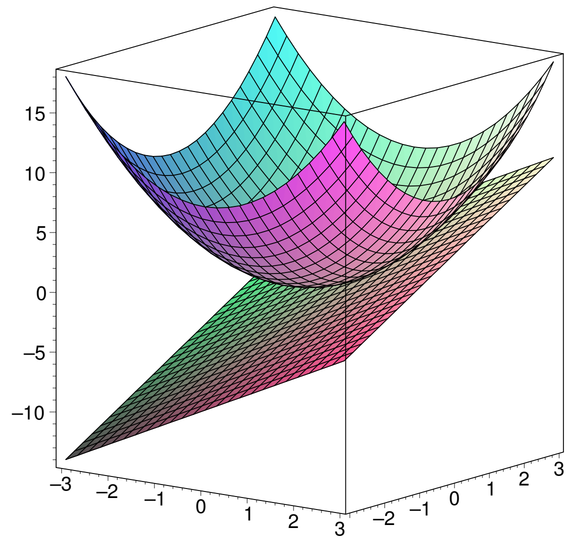


In the first graph, the x values range from 0.85 to 1.15. In the second, they range from 0.95 to 1.05. In the third, they range from 0.995 to 1.005. In the third graph, the curve and the tangent line are almost indistinguishable.

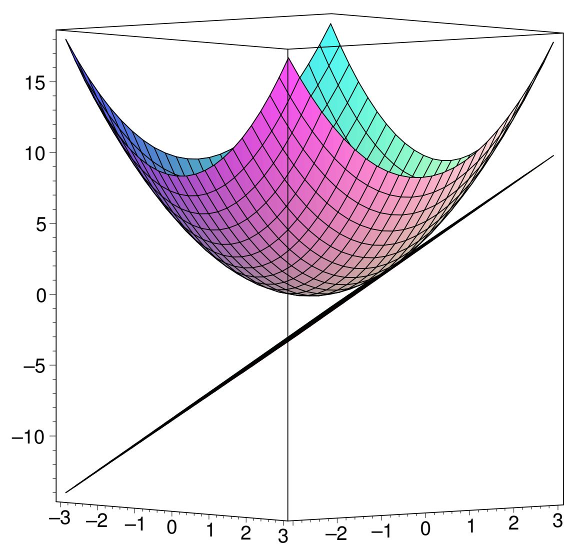
Clearly, there is just one line that fits so well: If the line had any other slope, the two graphs would not even look parallel, let alone the same, when we "zoom in". Also, they clearly must both pass through the point $(x_0, g(x_0))$ to have any sort of fit at all. Since the point and the slope determine the line, there is just one line that fits this well.

Now, lets move on to functions of two variables. The graph of $z = f(x, y)$ is a surface. The tangent plane to the graph of $z = f(x, y)$ at (x_0, y_0) is the plane that "best fits" this graph at (x_0, y_0) in the same way that the tangent line to the graph of $y = g(x)$ is the line that "best fits" this graph at x_0.

For example here is the graph of $z = x^2 + y^2$, together with the tangent plane to this graph at the point $(1, 1)$.



Here is another picture of the same thing from a different vantage point, giving a better view of the point of contact:



We can find an equation for this plane by looking along slices, and computing tangent lines. Here is how this goes. A non–vertical plane in $I\!R^3$ is the graph of the equation

$$z = Ax + By + C$$

for some A, B and C.

Definition (Linear function) A function $h(x, y)$ of the form

$$h(x, y) = Ax + By + C \tag{4.1}$$

is called a *linear function*

The terminology is a bit unfortunate since $h(0,0) = 0$ only in case $D = 0$, and *linear transformations* from $I\!R^2$ to $I\!R$ would always send the zero vector to zero. But this terminology is standard, and shouldn't cause trouble.

The point of the definition is that the graph of any linear function is a plane. How should A, B and D be chosen to ensure the "best possible fit", as in the pictures above?

• *To get the best fit, we require that*

$$h(x_0, y_0) = f(x_0, y_0) \tag{4.2}$$

and that every corresponding pair of slices of f and h have the same slopes at (x_0, y_0) in the slices parallel to the x and y axes.

As we now show, these three requirements determine the three coefficients A, B and C, and hence give us a formula for the tangent plane.

From (4.1) and (4.2) we see that

$$Ax_0 + By_0 + C = f(x_0.y_0) \ . \tag{4.3}$$

Next, in the slice parallel to the x axis, in which we fix $y = y_0$ and let x vary, the slopes in question are given by the x partial derivatives, so the requirement is that $\dfrac{\partial}{\partial x}h(x_0, y_0) = \dfrac{\partial}{\partial x}f(x_0, y_0)$. From (4.1), $\dfrac{\partial}{\partial x}h(x_0, y_0) = A$. Therefore,

$$A = \frac{\partial}{\partial x}f(x_0, y_0) \ . \tag{4.4}$$

The same reasoning with x and y varying gives

$$B = \frac{\partial}{\partial y}f(x_0, y_0) \ . \tag{4.5}$$

Combining (4.3), (4.4) and (4.5), we see that

$$h(x,y) = \left(\frac{\partial}{\partial x}f(x_0, y_0)\right)(x - x_0) + \left(\frac{\partial}{\partial y}f(x_0, y_0)\right)(y - y_0) + f(x_0, y_0) \ . \tag{4.6}$$

Introducing $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\mathbf{x}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ we can write this more compactly as

$$h(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \ . \tag{4.7}$$

Thus, (4.7) gives us a general formula for computing h in terms of $f(\mathbf{x}_0)$ and $\nabla f(\mathbf{x}_0)$.

The function h has an important relation to f: It is the *best linear approximation to f at $\mathbf{x}_0$* in that, first of all, it is a linear function, and second, its graph fits the graph of f at $\mathbf{x}_0$ better than any other linear function*

Definition (Best linear approximation and the tangent plane) If f has continuous first order partial derivatives in a neighborhood of a point $\mathbf{x}_0$, then the *best linear approximation to f at $\mathbf{x}_0$* is the function

$$h(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \ .$$

The *tangent plane of f at $\mathbf{x}_0$* is the graph of $\mathbf{z} = h(\mathbf{x})$. That is, it is the graph of

$$z = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \ .$$

The reason that we require continuity of the partial derivatives is that without this continuity, the tangent plane will not "fit" itself to the graph of $z = f(\mathbf{x})$ in a reasonable way. What we have done so far is to show that *if there is some plane that fits well*, like in the pictures we drew earlier, *it can only be* given by the graph of $z = h(\mathbf{x})$, where h is given by (4.7).

However, some nasty functions f have well defined partial derivatives, so that $h(\mathbf{x})$ is well defined, but its graph doesn't fit the graph of $z = f(\mathbf{x})$ very well at all. (The function defined in (2.5) is an example of such a function; it has no tangent plane at the origin.) Fortunately, such examples are rare and contrived. Moreover, in all of them, the partial derivatives are discontinuous at $\mathbf{x}_0$. As we shall see:

- *When the partial derivatives of f are continuous at $\mathbf{x}_0$, the graphs of $z = f(\mathbf{x})$ and $z = h(\mathbf{x})$ with h given by (4.7) do indeed match up, like in the pictures of tangent planes at the beginning of this section.*

We will come back to this soon. However, let's first focus on computing tangent planes for nice functions, and worry later about how well they actually fit the graph of $z = f(\mathbf{x})$.

Example 2 (Computing the equation of a tangent plane) Consider the function $f(x, y) = x^2 + y^2$ and $(x_0, y_0) = (1, 1)$, as in the graphs above. Then

$$f(1, 1) = 2 \qquad \text{and} \qquad \nabla f(1, 1) = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \ .$$

Hence from (4.7), the best linear approximation is

$$h(x, y) = 2 + \begin{bmatrix} 2 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} x - 1 \\ y - 1 \end{bmatrix} = 2x + 2y - 2 \ .$$

---

* As we have just seen, the graph of any other linear function would either have at least the wrong height at $\mathbf{x}_0$, or the wrong slope in some slice through $\mathbf{x}_0$. The graphs wouldn't even look parallel, let alone the same when viewed close up.

The tangent plane then has the equation

$$z = 2x + 2y - 2 \ .$$

The standard geometric form for the equation of a plane in $I\!R^3$ is $\mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) = 0$. To avoid confusion between two and three dimensional vectors at this crucial point, we will (temporarily) use capital letters to denote vectors in $I\!R^3$. That is we write $\mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $\mathbf{X}_0 = \begin{bmatrix} x_0 \\ y_0 \\ f(x_0, y_0) \end{bmatrix}$. Notice that $\mathbf{X}_0$ is the point on the graph of $z = f(x, y)$ above (x_0, y_0).

Now, from (4.6), we have that $z = h(x, y)$, the equation for the tangent plane, is

$$z = \left(\frac{\partial}{\partial x} f(x_0, y_0) \right) (x - x_0) + \left(\frac{\partial}{\partial y} f(x_0, y_0) \right) (y - y_0) + f(x_0, y_0)$$

or, equivalently, with $z_0 = f(x_0, y_0)$,

$$\left(\frac{\partial}{\partial x} f(x_0, y_0) \right) (x - x_0) + \left(\frac{\partial}{\partial y} f(x_0, y_0) \right) (y - y_0) + (-1)(z - z_0) = 0 \ . \tag{4.8}$$

Now introduce the vector

$$\mathbf{N} = \begin{bmatrix} \dfrac{\partial f}{\partial x}(\mathbf{x}_0) \\ \dfrac{\partial f}{\partial y}(\mathbf{x}_0) \\ -1 \end{bmatrix} \ . \tag{4.9}$$

Then (4.8) can be written compactly as

$$\mathbf{N} \cdot (\mathbf{X} - \mathbf{X}_0) = 0 \ . \tag{4.10}$$

In other words, the vector $\mathbf{N}$ defined in (4.9) is orthogonal to the tangent plane of f at $\mathbf{x}_0$. The formulas (4.10) (4.9) are useful for computing the equations of tangent planes.

Example 3 (Direct tangent plane computations) Let $f(x, y) = x^4 + y^4 + 4xy$, as in Example 8 of the previous section. From the computations of the partial derivatives there,

$$\mathbf{N} = \begin{bmatrix} 36 \\ 12 \\ -1 \end{bmatrix}$$

We can always take

$$\mathbf{X}_0 = \begin{bmatrix} x_0 \\ y_0 \\ f(x_0, y_0) \end{bmatrix}$$

since this is the point on the graph of $z = f(x, y)$ at which the tangent plane must touch. In this case we get

$$\mathbf{X}_0 = \begin{bmatrix} 2 \\ 1 \\ 25 \end{bmatrix}$$

Hence, (4.10) becomes

$$\begin{bmatrix} 36 \\ 12 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} x-2 \\ y-1 \\ z-25 \end{bmatrix} = 0 \ .$$

Computing the dot product, we find $36x + 12y - z = 59$, or, equivalently, $z = 36x + 12y - 59$, but for many purposes, such as computing the distance to the tangent plane, it is $\mathbf{N}$ and $\mathbf{X_0}$ that we want to know.

4.2 Critical points

The formula (4.9) has important application to *optimization problems*, which are problems in which we look for minimum and maximum values of f, and the inputs $\mathbf{x}$ that produce them. Indeed, you see that:

• *If $\nabla f(\mathbf{x}_0) \neq 0$, then $\mathbf{N}$ is not purely vertical, and so the tangent plane at $\mathbf{x}$ is not horizontal. Hence, there is an "uphill" direction and a "downhill" direction at $\mathbf{x}_0$.*

If it is possible to "move uphill" from $\mathbf{x}_0$, then $f(\mathbf{x}_0)$ cannot possibly be a maximum value of f. Likewise, it is possible to "move downhill" from $\mathbf{x}_0$, then $f(\mathbf{x}_0)$ cannot possibly be a minimum value of f.

• *If we are looking for either minimum values of f or maximum values of f in some open set U, and f has continuous partial derivatives everywhere in U, then it suffices to look among only at those points $\mathbf{x}$ at which $\nabla f(\mathbf{x}) = 0$.*

This leads us to the definition of a critical point:

Definition (Critical point) Suppose that f is defined and has a gradient in a neighborhood of some point $\mathbf{x}_0$. Then $\mathbf{x}_0$ is a *critical point* of f in case $\nabla f(\mathbf{x}_0) = 0$.

Example 4 (Computing critical points) Let $f(x, y) = x^4 + y^4 + 4xy$, as in Example 3. We have already computed the gradient of f, and we see that $\nabla f(x, y) = 0$ if and only if

$$\begin{aligned} x^3 + y &= 0 \\ y^3 + x &= 0 \ . \end{aligned}$$

This is an nonlinear system of equations. To solve it we have to find a way to eliminate variables. That is easy here: The first equation says $y = -x^3$, and using this to eliminate y from the second equation, we have $-x^9 + x = 0$. This has exactly three solutions: $x = 0$, $x = 1$ and $x = -1$. Since $y = -x^3$, the corresponding critical points are

$$(0, 0) \qquad (1, -1) \qquad \text{and} \qquad (-1, 1) \ .$$

The three critical points found in Example 4 are the *only* points at which f can possibly take on either a maximum value or a minimum value. We will see later that f is minimized at both $(1.-1)$ and $(-1,1)$, and that f has no maximum.

4.3 What differentiability means for a function on $\mathbb{R}^2$

Tangent planes are as central to the concept of differentiability in $\mathbb{R}^2$ as tangent lines are to the concept of differentiability in $\mathbb{R}$.

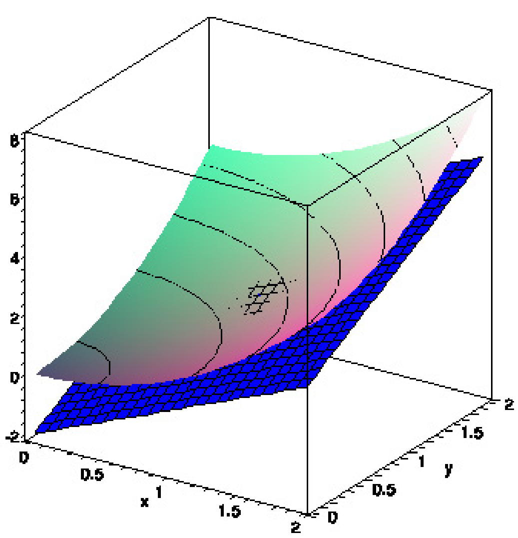
Roughly speaking, a function $f(x,y)$ is *differentiable* at (x_0,y_0) means that there is a linear function $h(x,y) = Ax + By + C$ such that

$$h(x,y) \approx f(x,y) \tag{4.11}$$

nearby (x_0,y_0), and the approximation in (4.11) is good enough that if you "zoom in" sufficiently closely on a joint graph of $z = f(x,y)$ and $z = h(x,y)$, you cannot see the difference.

For example, consider the function $f(x,y) = x^2 + y^2$ with $x_0 = 1$ and $y_0 = 1$. We have compute the best linear approximation h at $\mathbf{x}_0$ in Example 2 and found $h(x,y) = 2x+2y-2$. Here is a three dimensional graph of f and h together for the region

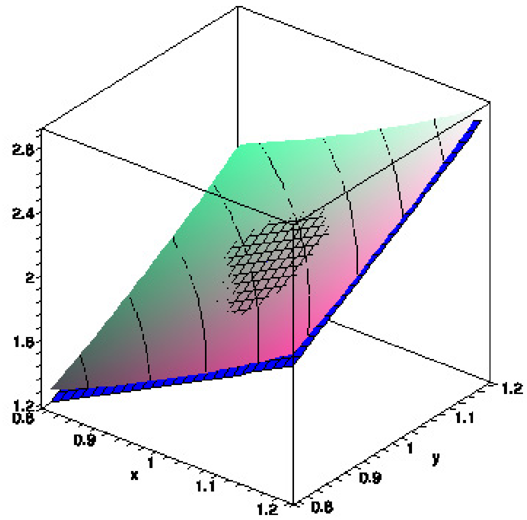
$$|x-1| \le 1 \qquad \text{and} \qquad |y-1| < 1 \ :$$



As you see, the graphs are almost indistinguishable for x and y in the region

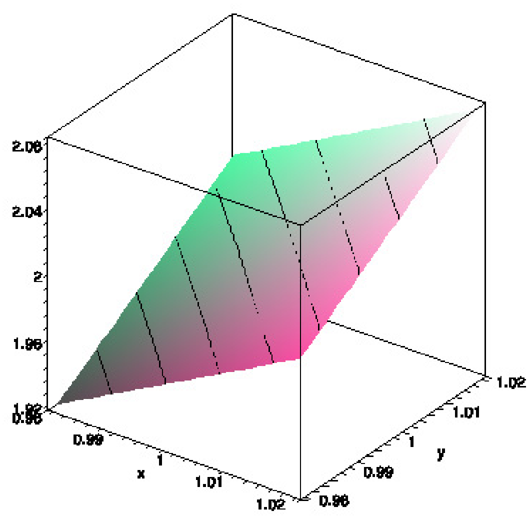
$$|x-1| \le 0.2 \qquad \text{and} \qquad |y-1| < 0.2 \ .$$

Let's "zoom in" on this region:



The vertical separation between the graphs is getting to be a pretty small percentage of the displayed distances. The graphs are almost indistinguishable. Let's zoom in by a factor of 10, and have a look for

$$|x - 1| \le 0.02 \qquad \text{and} \qquad |y - 1| < 0.02 \ .$$



Now, the graphs really are indistinguishable.

On the other hand, consider the function f defined by

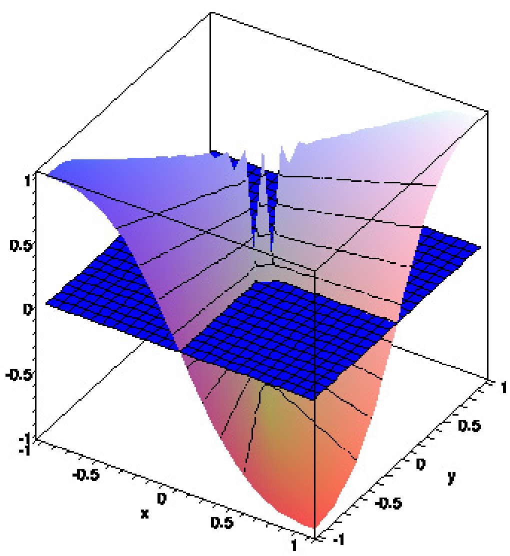
$$f(x,y) = \begin{cases} \dfrac{2xy}{x^4 + y^4} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases} \ . \tag{4.12}$$

This is a close relative of the function defined in (1.5), and almost as nasty. As you can check, both partial derivatives are defined everywhere and

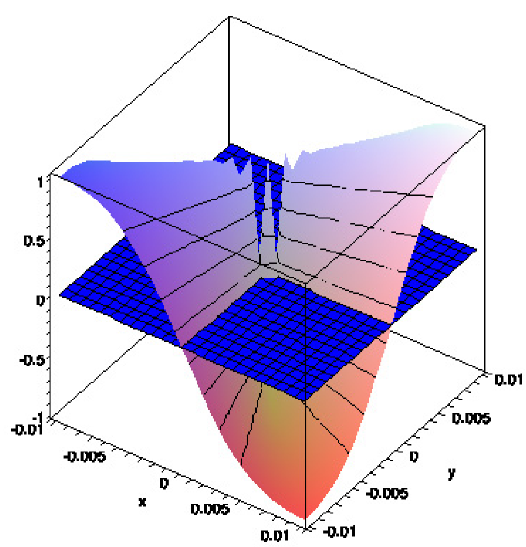
$$f(0,0) = 0 \qquad \frac{\partial f}{\partial x}(0,0) = 0 \qquad \text{and} \qquad \frac{\partial f}{\partial y}(0,0) = 0 \ .$$

Hence our formula for the best linear approximation gives us $h(x, y) = 0$. It may be the best linear approximation, but it is a pretty rotten approximation all the same. Here is a graph of f and h together for

$$|x - 1| \leq 0.1 \qquad \text{and} \qquad |y - 1| < 0.1 \ .$$



Anyone who calls that a "good fit" has pretty lax standards. Maybe things get better if we zoom in? Unfortunately, no. Here is what we get when we zoo in by a factor of 10:



The graph looks just the same, except that as the scales indicate, we have zoomed in on the region

$$|x - 1| \leq 0.1 \qquad \text{and} \qquad |y - 1| < 0.1 \ .$$

The same thing happens no matter how much you zoom in.

• *The function* $f(x, y) = x^2 + y^2$ **is differentiable** *at* $x_0 = 1$ *and* $y_0 = 1$ *because if you zoom in enough, its graph is indistinguishable from that of its tangent plane. The function*

f defined by (4.12) **is not differentiable** *at $x_0 = 0$ and $y_0 = 0$ because no matter how close you zoom, the graph never looks planar – there is no plane that really fits in the way a tangent plane should fit.*

Now that we understand in visual terms what differentiability is all about, let's phase it in mathematical terms.

In each of the graphs we drew above, we picked some number $\delta > 0$ which determined the *horizontal display size*. It determined the horizontal size simply because what we graphed was the part of the surfaces over

$$|x - x_0| \leq \delta \qquad \text{and} \qquad |y - y_0| < \delta \qquad (4.13)\ .$$

For x and y in this region, the vertical separation between the two graphs is

$$|f(x, y) - h(x, y)|\ .$$

For the graphs to be indistinguishable on this scale, we require that the vertical separation is a "small" fraction of δ at each x and y.

What passes for "small" depends on the resolution in our graph. For example, suppose we use a computer to draw the graphs and zoom in so that the *height and width** of the screen correspond to the distance 2δ. Then the two graphs will fill the screen horizontally. If the graph is 1000 pixels by 1000 pixels, and if for each x and y in the region (4.13), the difference

$$\frac{|f(x, y) - h(x, y)|}{2\delta} < \frac{1}{1000}\ ,$$

then the differences in the heights will be less than one pixel everywhere on the screen, and the two graphs will be indistinguishable.

So, indistinguishability has to do with the *ratio* of the vertical separation $|f(x, y) - h(x, y)|$ and the horizontal separation $|\mathbf{x} - \mathbf{x}_0|$. Therefore, define the function $r(\mathbf{x})$ by

$$r(\mathbf{x}) = \begin{cases} \dfrac{|h(\mathbf{x}) - f(\mathbf{x})|}{|\mathbf{x} - \mathbf{x}_0|} & \text{if } \mathbf{x} \neq \mathbf{x}_0 \\ 0 & \text{if } \mathbf{x} = \mathbf{x}_0 \end{cases}\ . \qquad (4.14))$$

The function $r(\mathbf{x})$ is continuous at $\mathbf{x}_0$ if and only if the ratio of the vertical to horizontal separations, $\dfrac{|h(\mathbf{x}) - f(\mathbf{x})|}{|\mathbf{x} - \mathbf{x}_0|}$, tends to zero as $\mathbf{x}$ approaches $\mathbf{x}_0$.

---

* Although we are now discussing a graph using the same scale for vertical and horizontal axes, or at least something close to that, we did not always use that in the graphs we drew above. You can check and see that sometimes we used different scales for the vertical direction just to fit things on the page. But if we had used "true proportions" in the last graph of the nasty function, the differences would have been even more dramatic – and the vertical oscillations would not have come close to fitting on the page.

Recalling the ϵ and δ definition of continuity, this amounts to requiring that if $\epsilon > 0$ is any given threshold of "indistinguishability" that we care to set, then there is then a $\delta > 0$ so that

$$|\mathbf{x} - \mathbf{x}_0| < \delta \quad \Rightarrow \quad |h(\mathbf{x}) - f(\mathbf{x})| \le \epsilon|\mathbf{x} - \mathbf{x}_0| \ . \tag{4.15}$$

In this case, if you "zoom in" on a joint graph of $z = h(x, y)$ and $z = f(x, y)$ until the region with $|\mathbf{x} - \mathbf{x}_0| < \delta$ fills your screen, the vertical separation of the two graphs will be a negligible fraction – ϵ at most – of the width.

Definition (Differentiability) Let f be a function defined on an open set U in $I\!R^2$. then, for any $\mathbf{x}_0$ in U, f is differentiable at $\mathbf{x}_0$ in case it there is a linear function

$$h(x, y) = Ax + By + D$$

so that

$$r(\mathbf{x}) = \begin{cases} \dfrac{|h(\mathbf{x}) - f(\mathbf{x})|}{|\mathbf{x} - \mathbf{x}_0|} & \text{if } \mathbf{x} \neq \mathbf{x}_0 \\ 0 & \text{if } \mathbf{x} = \mathbf{x}_0 \end{cases} \ . \qquad (4.16))$$

is continuous at $\mathbf{x}_0$.

Note that the continuity of r at $\mathbf{x}_0$ is equivalent to the requirement that

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{|h(\mathbf{x}) - f(\mathbf{x})|}{|\mathbf{x} - \mathbf{x}_0|} = 0 \ .$$

Now that we know what it means for a function to be differentiable, we can prove the following theorem which relates differentiability to the existence of continuous partial derivatives.

Theorem 1 (Differentiability and partial derivatives) *Let f be any function defined in an open set U of $I\!R^2$ Suppose that both partial derivatives of f are defined and continuous at some point $\mathbf{x}_0$ in U. Then f is differentiable at $\mathbf{x}_0$, and the unique linear function h for which*

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

is given by

$$h(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \ .$$

At this point you may be wondering: What is the derivative of a function of two variables? The derivative is the gradient. If you recall that in single variable calculus, the derivative gives you the slope of the tangent line to the graph, and recall that as we have just seen, the gradient determines the slope of the tangent plane for a function of two variables, you see that the terminology makes sense.

Proof of Theorem 1 The proof will be very similar to the proof of Theorem 2 from Section 2. Fix $\mathbf{x}$ and $\mathbf{x}_0$ and define numbers a and b by

$$x - x_0 = a \qquad \text{and} \qquad y - y_0 = b$$

Define single variable functions ϕ and ψ by

$$\phi(s) = h(x_0, y_0 + sb) - f(x_0, y_0 + sb)$$

and

$$\psi(t) = h(x_0 + ta, y_0 + b) - f(x_0 + ta, y_0 + b) \ .$$

As you can check,

$$\begin{aligned}
[\psi(1) - \psi(0)] + [\phi(1) - \phi(0)] &= \left[\Big(f(x,y) - h(x,y)\Big) - \Big(h(x_0,y) - f(x_0,y)\Big)\right] \\
&\quad + \left[\Big(f(x_0,y) - h(x_0,y)\Big) - \Big(f(x_0,y_0) - h(x_0,y_0)\Big)\right] \\
&= [h(\mathbf{x}) - f(\mathbf{x})] - [h(\mathbf{x}_0) - f(\mathbf{x}_0)] \ .
\end{aligned}$$

But since, by definition, $h(\mathbf{x}_0) = f(\mathbf{x}_0)$,

$$h(\mathbf{x}) - f(\mathbf{x}) = [\psi(1) - \psi(0)] + [\phi(1) - \phi(0)] \ . \tag{4.17}$$

We can now apply the Mean Value Theorem to the single variable functions $\psi(t)$ and $\phi(s)$ to find

$$[\psi(1) - \psi(0)] = \psi'(\tilde{t}) \qquad \text{and} \qquad [\phi(1) - \phi(0)] = \phi'(\tilde{s}) \tag{4.18}$$

for some $\tilde{s}$ and $\tilde{t}$ between 0 and 1.

Using the chain rule, we compute

$$\begin{aligned}
\psi'(t) &= \left[\frac{\partial h}{\partial x}(x_0 + t\tilde{a}, y_0 + b) - \frac{\partial f}{\partial x}(x_0 + ta, y_0 + b)\right] b \\
&= \left[A - \frac{\partial f}{\partial x}(x_0 + \tilde{t}a, y_0 + b)\right] a \ .
\end{aligned} \tag{4.19}$$

Likewise,

$$\begin{aligned}
\phi'(\tilde{s}) &= \left[\frac{\partial h}{\partial y}(x_0, y_0 + \tilde{s}b) - \frac{\partial f}{\partial y}(x_0, y_0 + \tilde{s}b)\right] b \\
&= \left[B - \frac{\partial f}{\partial y}(x_0, y_0 + \tilde{s}b)\right] b \ .
\end{aligned} \tag{4.20}$$

Combining (4.17), (4.18), (4.19) and (4.20),

$$h(\mathbf{x}) - f(\mathbf{x}) = \left[A - \frac{\partial f}{\partial x}(x_0 + \tilde{t}a, y_0 + b)\right] a + \left[B - \frac{\partial f}{\partial y}(x_0, y_0 + \tilde{s}b)\right] b \ .$$

The right hand side is the dot product of

$$\begin{bmatrix} A - \frac{\partial f}{\partial x}(x_0 + \tilde{t}a, y_0 + b) \\ B - \frac{\partial f}{\partial y}(x_0, y_0 + \tilde{s}b) \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} a \\ b \end{bmatrix} = \mathbf{x} - \mathbf{x}_0 \ .$$

Therefore, by the Schwarz inequality,

$$\frac{|h(\mathbf{x}) - f(\mathbf{x})|}{|\mathbf{x} - \mathbf{x}_0|} \le \sqrt{\left(A - \frac{\partial f}{\partial x}(x_0 + \tilde{t}a, y_0 + b)\right)^2 + \left(B - \frac{\partial f}{\partial y}(x_0, y_0 + \tilde{s}b)\right)^2} \ .$$

Now, as $\mathbf{x} \to \mathbf{x}_0$, $a \to 0$ and $b \to 0$. Therefore, *using the continuity of the partial derivatives at this point*,

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \sqrt{\left(A - \frac{\partial f}{\partial x}(x_0 + \tilde{t}a, y_0 + b)\right)^2 + \left(B - \frac{\partial f}{\partial y}(x_0, y_0 + \tilde{s}b)\right)^2} =$$

$$\sqrt{\left(A - \frac{\partial f}{\partial x}(x_0, y_0)\right)^2 + \left(B - \frac{\partial f}{\partial y}(x_0, y_0)\right)^2} \ .$$

Clearly, this is zero if and only if

$$A = \frac{\partial f}{\partial x}(x_0, y_0) \qquad \text{and} \qquad B = \frac{\partial f}{\partial y}(x_0, y_0) \ .$$

Making these choices for A and B we see that f is indeed differentiable at $\mathbf{x}_0$.

We see also that there is just one choice of A and B for which $\lim_{\mathbf{x} \to \mathbf{x}_0} r(\mathbf{x}) = 0$. We have already assumed that $h(\mathbf{x}_0) = f(\mathbf{x}_0)$, which determines C, but if we do not make this assumption, $\lim_{\mathbf{x} \to \mathbf{x}_0} r(\mathbf{x}) = 0$ is impossible. Hence the values of A, B and C for which $\lim_{\mathbf{x} \to \mathbf{x}_0} r(\mathbf{x}) = 0$ are uniquely determined. In other words, there is just one linear function h that fits f so well at $\mathbf{x}_0$. ■

Exercises

1. Let $f(x, y) = x\sin((x^2 + y)\pi) + xy$ and let $(x_0, y_0) = (1, 2)$ Compute the equation of the tangent plane to $z = f(x, y)$ at (x_0, y_0).

2. Let $f(x, y) = (2xy + x^2 - y^3)/(1 + x^4 + y^4)$ and let $(x_0, y_0) = (2, 0)$ Compute the equation of the tangent plane to $z = f(x, y)$ at (x_0, y_0).

3. Let $f(x, y) = x\sin((x^2 + y)\pi) + xy$ and let $(x_0, y_0) = (1, 2)$. Find vectors $\mathbf{N}$ and $\mathbf{X}_0$ in $I\!R^3$ that with $\mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, the equation for the tangent plane to $z = f(x, y)$ can be written as $\mathbf{N} \cdot (\mathbf{X} - \mathbf{X}_0) = 0$.

4. Let $f(x, y) = (2xy + x^2 - y^3)/(1 + x^4 + y^4)$ and let $(x_0, y_0) = (2, 0)$. Find vectors $\mathbf{N}$ and $\mathbf{X}_0$ in $I\!R^3$ that with $\mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, the equation for the tangent plane to $z = f(x, y)$ can be written as $\mathbf{N} \cdot (\mathbf{X} - \mathbf{X}_0) = 0$.

Section 5: Curvature of the graph of $z = f(x, y)$

5.1 Directional second derivatives and the Hessian of f

For a function $g(x)$ of a single variable x, the first derivative $g'(x)$ measures the *slope* of the graph of $y = g(x)$ at x, and the second derivative measures the *curvature* of this graph at x. Our goal now is to understand how the surface $z = f(x, y)$ "curves away" from its tangent plane at a point (x, y). You probably guess the strategy:

- *To understand how the surface $z = f(x, y)$ "curves away" from its tangent plane at (x, y), we will compute second derivatives along slices through (x, y).*

Here is how to do this: Consider a function f from $I\!R^2$ to $I\!R$. Let $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ be any to vectors in R^2. Consider the line in the plane paramterized by $\mathbf{x} + t\mathbf{v}$.

We have seen how to take directional (first) derivatives: The directional derivative of f at $\mathbf{x}$ in the direction $\mathbf{v}$ is just

$$\mathbf{v} \cdot \nabla f(\mathbf{x}) \ .$$

This is the slope of the slice of f passing through $\mathbf{x}$ in direction $\mathbf{v}$; i.e., the derivative of $g(t) = f(\mathbf{x} + t\mathbf{v})$ at $t = 0$.

Now, with $\mathbf{v}$ fixed, and $\mathbf{x}$ variable, the function that returns the value $\mathbf{v} \cdot \nabla f(\mathbf{x})$ for the input $\mathbf{x}$ is just another function of $\mathbf{x}$. Hence we can take its directional derivative in the direction $\mathbf{v}$. This is a second directional derivative, and it give the curvature of the slice of $z = f(\mathbf{x})$ passing through $\mathbf{x}$ with direction $\mathbf{v}$.

Example 1 (A second directional derivative) Consider the function $f(x, y) = x^2y - xy + x^3$, and the line $\mathbf{x}(t) = \mathbf{x}_0 + t\mathbf{v}$ where $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. We compute

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2xy - y + 3x^2 \\ x^2 - x \end{bmatrix}$$

so that

$$\mathbf{v} \cdot \nabla f(\mathbf{x}) = -(2xy - y + 3x^2) + 2(x^2 - x) = -2xy - x^2 + y - 2x \ .$$

The right hand side is just another function of x and y. Let's call it $s(x, y)$, so we have

$$s(x, y) = -2xy - x^2 + y - 2x \ .$$

The reason we call this function s is that it gives us the *slope* of the graph of $z = f(\mathbf{x})$ on a path that passes through $\mathbf{x}$ along the line $\mathbf{x} + t\mathbf{v}$. In particular, at $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $s(\mathbf{x}_0) = -4$, so the slope of the slice passing through $\mathbf{x}_0$ is downward – four steps down for each step forward, which is pretty steep.

As you walk along the path, the slope changes. Is is getting less steep or more steep? To see how the slope changes with t, take another directional derivative. We compute

$$\nabla s(\mathbf{x}) = \begin{bmatrix} -2y - 2x - 2 \\ -2x + 1 \end{bmatrix}$$

In particular, evaluating this at $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$,

$$\mathbf{v} \cdot \nabla s(\mathbf{x}_0) = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -6 \\ -1 \end{bmatrix} = 4 \ .$$

The slope is increasing, and therefore becoming less negative. That is, this slice is "curving upwards" at $\mathbf{x}_0$.

The computation done in particular in Example 1 can be done in general. Given any function $f(\mathbf{x})$ and any vector $\mathbf{v}$, we compute a "slope function" $s(\mathbf{x})$ defined by

$$s(\mathbf{x}) = \mathbf{v} \cdot \nabla f(\mathbf{x}) \ . \tag{5.1}$$

This represents the slope of the slice of the graph of $z = f(\mathbf{x})$ moving through $\mathbf{x}$ with the velocity vector $\mathbf{v}$. That is, with $g(t) = f(\mathbf{x}_0 + t\mathbf{v})$, $g'(t) = s(\mathbf{x}_0 + t\mathbf{v})$.

To see the rate of change of the slope, we differentiate again. By the directional derivative formula,

$$g''(t) = (g'(t))' = \mathbf{v} \cdot \nabla s(\mathbf{x}_0 + t\mathbf{v}) \ . \tag{5.2}$$

Let's do the computations explicitly with $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. First,

$$s(x,y) = \mathbf{v} \cdot \nabla f(x,y) = a\frac{\partial f}{\partial x}(x,y) + b\frac{\partial f}{\partial y}(x,y) \ , \tag{5.3}$$

and

$$\mathbf{v} \cdot \nabla f(x,y) = a\frac{\partial s}{\partial x}(x,y) + b\frac{\partial s}{\partial y}(x,y) \ . \tag{5.4}$$

The partial derivative of s can be computed using (5.3). This will involve *repeated partial derivatives*. There is nothing deep about this; once you've taken the x partial derivative of f, you've got a new function of x and y. You can now go on and take the y derivative of that. We now introduce some notation for these repeated partial derivatives. Define

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \\
\frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \\
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \\
\frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \ .
\end{aligned}$$

Example 2 (Second partial derivatives) Let $f(x,y) = x^2y - xy + x^3$. Then $\dfrac{\partial f}{\partial x} = 2xy - y + 3x^2$. Differentiating once more with respect to x and y respectively, we find

$$\frac{\partial^2 f}{\partial x^2} = 2y - 6x \qquad \text{and} \qquad \frac{\partial^2 f}{\partial y \partial x} = 2x - 1 \ .$$

We have also seen that $\frac{\partial f}{\partial y} = x^2 - x$. Differentiating once more with respect to x and y respectively, we find

$$\frac{\partial^2 f}{\partial x \partial y} = 2x - 1 \qquad \text{and} \qquad \frac{\partial^2 f}{\partial y^2} = 0$$

Notice that in Example 2, it turned out that

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} \ . \tag{5.5}$$

As we shall see later on, this is no accident: This equality always holds, and that will turn out to be a very useful fact.

However, our immediate goal is to deduce a *convenient formula* for second direction derivatives in terms of second partial derivatives. When we did the same thing for first directional derivatives and first partial derivatives, we organized the two first partial derivatives into a vector, the gradient. We will now organize the four second partial derivatives into a matrix.

To deduce the formula, plug (5.3) into (5.4) and obtain

$$\begin{aligned}
\mathbf{v} \cdot \nabla s(\mathbf{x}) &= \mathbf{v} \cdot \nabla \left(\mathbf{v} \cdot \nabla f \right) \\
&= \mathbf{v} \cdot \nabla \left(a \frac{\partial f}{\partial x} + b \frac{\partial f}{\partial y} \right) \\
&= a\mathbf{v} \cdot \nabla \left(\frac{\partial f}{\partial x} \right) + b\mathbf{v} \cdot \nabla \left(\frac{\partial f}{\partial y} \right) \\
&= a \left(a \frac{\partial}{\partial x} \frac{\partial f}{\partial x} + b \frac{\partial}{\partial y} \frac{\partial f}{\partial x} \right) + b \left(a \frac{\partial}{\partial x} \frac{\partial f}{\partial y} + b \frac{\partial}{\partial y} \frac{\partial f}{\partial y} \right) \\
&= a^2 \frac{\partial^2 f}{\partial x^2} + ab \frac{\partial^2 f}{\partial y \partial x} + ba \frac{\partial^2 f}{\partial x \partial y} + b^2 \frac{\partial^2 f}{\partial y^2} \ .
\end{aligned} \tag{5.6}$$

We can write this in a simpler form if we introduce the *Hessian matrix* H_f:

Definition Let f be a function that possesses second order partial derivatives at $\mathbf{x}$. The the *Hessian of f at x* is the 2×2 matrix defined by

$$H_f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x^2} & \frac{\partial^2 f(\mathbf{x})}{\partial y \partial x} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x \partial y} & \frac{\partial^2 f(\mathbf{x})}{\partial y^2} \end{bmatrix} \ .$$

Note that the entries are all functions of $\mathbf{x}$. However, to simplify the notation, we will simply write H_f instead of $H_f(\mathbf{x})$ wherever this is not ambiguous.

The point of the definition is that with $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$,

$$\mathbf{v} \cdot H_f \mathbf{v} = a^2 \frac{\partial^2 f}{\partial x^2} + ab \frac{\partial^2 f}{\partial y \partial x} + ba \frac{\partial^2 f}{\partial x \partial y} + b^2 \frac{\partial^2 f}{\partial y^2}$$

as you can easily check.

This, together with (5.2) and (5.6), gives us the following Theorem:

Theorem 1 *Let f be any function with continuous second partial derivatives in a neighborhood of $\mathbf{x}_0$. Then for any vector $\mathbf{v}$.*

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\mathbf{x}_0 + t\mathbf{u}) = \mathbf{v} \cdot H_f(\mathbf{x}_0 + t\mathbf{v})\mathbf{v} \ . \tag{5.7}$$

Example 3 (Computing second directional derivatives with the Hessian) Consider the function $f(x, y) = x^2 y - xy + x^3$, and the line $\mathbf{x}(t) = \mathbf{x}_0 + t\mathbf{v}$ where $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Let $g(t) = f(\mathbf{x}_0 + t\mathbf{v})$, and let's compute $g''(0)$.

We have already worked out all of the second order partial derivatives of f in Example 2, so we can write down H_f:

$$H_f = \begin{bmatrix} 2y + 6x & 2x - 1 \\ 2x - 1 & 0 \end{bmatrix} \ .$$

Evaluating this at the point $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we find

$$H_f(\mathbf{x}_0) = \begin{bmatrix} 8 & 1 \\ 1 & 0 \end{bmatrix} \ .$$

Therefore, by (5.7),

$$g''(0) = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 8 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -6 \\ -1 \end{bmatrix} = 4 \ .$$

This agrees with out calculation in Example 1

The Hessian encodes *all* information about *all* directional second derivatives into one compact form – a single 2×2 matrix. Therefore, by analyzing the Hessian, we can learn the answers to such questions of the type:

- *How does the graph of $z = f(\mathbf{x})$ curve away from the tangent plane at $\mathbf{x}_0$?*

Because the equality (5.5) of the two mixed partial derivatives is always true, the Hessian is always a symmetric matrix. The most important fact about symmetric matrices is that they always have real eigenvalues and there is always an orthonormal basis of eigenvectors for them. Finding these eigenvalues and eigenvectors will be the key to understanding the Hessian, and therefore the curvature of $z = f(x, y)$.

Example 4 (Symmetry of the Hessian matrix) Suppose that f is a second degree polynomial in x and y. The the general form of such a polynomial is

$$C + Ax + By + \frac{D}{2}x^2 + Exy + \frac{F}{2}y^2 \ .$$

No matter how the coefficients are chosen, we will have in this case that

$$\frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x_0, y_0) = E = \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x_0, y_0) \ .$$

Computing the rest, we find that the Hessian is the constant symmetric matrix

$$H_f = \begin{bmatrix} D & E \\ E & F \end{bmatrix} \ .$$

So at least it is true in this case that the Hessian of a quadratic polynomial is always symmetric.

5.2 Symmetry of the Hessian matrix

Theorem 2 (Equality of mixed partial derivatives) *Let $f(x, y)$ be a function of the two variables x and y such that all partial derivatives of f order 2 are continuous in a neighborhood of (x_0, y_0). Then*

$$\frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x_0, y_0) = \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x_0, y_0) \ .$$

In particular, $H_f(x_0, y_0)$ is a symmetric matrix.

It is worthwhile to go through the proof: Doing this will help to understand what $\dfrac{\partial^2}{\partial x \partial y}f(x, y)$ itself represents. That is easy for $\dfrac{\partial^2}{\partial x^2}f(x, y)$ and $\dfrac{\partial^2}{\partial y^2}f(x, y)$: According to Theorem 1, these are just the curvatures of the slices of $z = f(\mathbf{x})$ along slices in the x and y directions respectively.

Although the values of the mixed partial derivatives are used to compute a directional second derivative, and hence the curvature along a slice, in a general direction, they are not themselves curvatures in any particular direction. After working through the following proof, you will understand what they actually are.

Proof By hypothesis, all partial derivatives of order 2 or less of f are continuous in a neighborhood of (x_0, y_0). Pick any $\epsilon > 0$. Then choose $h > 0$ so small that for all x and y with

$$x_0 \le x \le x_0 + h \qquad \text{and} \qquad y_0 \le y \le y_0 + h \ ,$$

$$\left| \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x_0, y_0) - \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x, y) \right| < \epsilon \qquad \text{and} \qquad \left| \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x_0, y_0) - \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x, y) \right| < \epsilon \ . \tag{5.8}$$

Then, by the Fundamental Theorem of Calculus,

$$f(x_0 + h, y_0) - f(x_0, y_0) = \int_0^h \frac{\partial f}{\partial x}(x_0 + t, y_0)\mathrm{d}t$$

and

$$f(x_0+h, y_0+h) - f(x_0+h, y_0) = \int_0^h \frac{\partial f}{\partial x}(x_0+h, y_0+t)\mathrm{d}t \ .$$

Together we have

$$f(x_0+h, y_0+h) - f(x_0, y_0) = \int_0^h \frac{\partial f}{\partial x}(x_0+t, y_0)\mathrm{d}t + \int_0^h \frac{\partial f}{\partial y}(x_0+h, y_0+t)\mathrm{d}t \ .$$

In this computation, we are keeping track how the value of $f(x,y)$ changes as we move from (x_0, y_0) to (x_0+h, y_0+h) moving along the bottom and right sides of the square with side length h and lower left corner at (x_0, y_0).

We can also compute it by keeping track of how it changes as we move along the left and upper sides. This gives us

$$f(x_0+h, y_0+h) - f(x_0, y_0) = \int_0^h \frac{\partial f}{\partial y}(x_0, y_0+t)\mathrm{d}t + \int_0^h \frac{\partial f}{\partial x}(x_0+t, y_0+h)\mathrm{d}t \ .$$

The two different sets of integrals ad up to the same thing, so

$$\begin{gathered}\int_0^h \frac{\partial f}{\partial x}(x_0+t, y_0)\mathrm{d}t + \int_0^h \frac{\partial f}{\partial y}(x_0+h, y_0+t)\mathrm{d}t = \\ \int_0^h \frac{\partial f}{\partial y}(x_0, y_0+t)\mathrm{d}t + \int_0^h \frac{\partial f}{\partial x}(x_0+t, y_0+h)\mathrm{d}t \ .\end{gathered}$$

We can rearrange this by putting both integrals involving $\partial f/\partial y$ on the left, and both integrals involving $\partial f/\partial x$ on the right. This gives us

$$\begin{gathered}\int_0^h \left(\frac{\partial f}{\partial y}(x_0+h, y_0+t) - \frac{\partial f}{\partial y}(x_0, y_0+t)\right)\mathrm{d}t = \\ \int_0^h \left(\frac{\partial f}{\partial x}(x_0+t, y_0+h) - \frac{\partial f}{\partial x}(x_0+t, y_0)\right)\mathrm{d}t \ .\end{gathered}$$

Now fix any value of t with $0 \le t \le h$, and define a function $\phi(x)$ by $\phi(x) = \dfrac{\partial f}{\partial y}(x, y_0+t)$ for this value of t. Then

$$\left(\frac{\partial f}{\partial y}(x_0+h, y_0+t) - \frac{\partial f}{\partial y}(x_0, y_0+t)\right) = \phi(x_0+h) - \phi(x_0) \ .$$

By the Mean Value Theorem,

$$\phi(x_0+h) - \phi(x_0) = \phi'(c)h$$

for some value of c between x_0 and $x_0 + h$. Note that

$$\phi'(x) = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x, y_0 + t) \ ,$$

and so

$$\left(\frac{\partial f}{\partial y}(x_0 + h, y_0 + t) - \frac{\partial f}{\partial y}(x_0, y_0 + t)\right) = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(c(t), y_0 + t)h$$

where we have written $c(t)$ to indicate that the value of c depends on the value of t we have fixed.

We are finally ready to use (5.8): Since for each t, $(c(t), y_0 + t)$ lies in the square where (5.8) is true, we have

$$\frac{\partial}{\partial x}\frac{\partial f}{\partial y}(c(t), y_0 + t) = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x_0, y_0) \pm \epsilon \ .$$

Putting it all together,

$$\int_0^h \left(\frac{\partial f}{\partial y}(x_0 + h, y_0 + t) - \frac{\partial f}{\partial y}(x_0, y_0 + t)\right) \mathrm{d}t = h^2 \left(\frac{\partial}{\partial y}\frac{\partial f}{\partial y}(x_0, y_0) \pm \epsilon\right) \ .$$

The exact same sort of reasoning leads to

$$\int_0^h \left(\frac{\partial f}{\partial x}(x_0 + t, y_0 + h) - \frac{\partial f}{\partial x}(x_0 + t, y_0)\right) \mathrm{d}t = h^2 \left(\frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x_0, y_0) \pm \epsilon\right) \ .$$

This finally gives us

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial y}(x_0, y_0) = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x_0, y_0) \pm 2\epsilon$$

and since this is true for all $\epsilon > 0$ no matter how small, we have the equality. ■

5.3: Directions of minimal and maximal curvature

Depending on how you slice the graph of $z = f(\mathbf{x})$, you will see more or less curvature as the slice passes through $\mathbf{x}_0$: The surface may curve up in some directions, and down in others.

- *Which directions give the minimum and the maximum of the curvature?*

The answer involves the eigenvectors of the the Hessian matrix. Here is how they come in:

First, let $\mathbf{u}$ be any unit vector, and consider the directional second derivative

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\mathbf{x}_0 + t\mathbf{u})\bigg|_{t=0} = \mathbf{u} \cdot H_f(\mathbf{x}_0)\mathbf{u} \ .$$

This measures the curvature of the graph of $z = f(\mathbf{x})$ as you move through $\mathbf{x}_0$ in the direction $\mathbf{u}$. We can rephrase our question as

• *Which direction vector* $\mathbf{u}$ *gives the smallest value for* $\mathbf{u} \cdot H_f(\mathbf{x}_0)\mathbf{u}$*, and which gives the largest?*

Second, recall that for any 2×2 symmetric matrix A, there is an orthogonal pair of unit vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, and a pair of numbers μ_1 and μ_2 with and

$$A\mathbf{u}_1 = \mu_1\mathbf{u}_1 \qquad \text{and} \qquad A\mathbf{u}_2 = \mu_2\mathbf{u}_2 \ . \tag{5.9}$$

Now consider any unit vector $\mathbf{u}$ in $I\!R^2$. Since any pair of orthonormal vectors in $I\!R^2$ is a basis for $I\!R^2$, there are numbers α and β so that

$$\mathbf{u} = \alpha\mathbf{u}_1 + \beta\mathbf{u}_2 \ . \tag{5.10}$$

Since $\mathbf{u}$ is a unit vector, and since $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthonormal, $1 = \mathbf{u} \cdot \mathbf{u} = (\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) \cdot (\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) = \alpha^2 + \beta^2$. That is,

$$\alpha^2 + \beta^2 = 1 \ . \tag{5.11}$$

We can now use (5.9) and (5.10) to compute $\mathbf{u} \cdot A\mathbf{u}$:

$$\begin{aligned}
\mathbf{u} \cdot A\mathbf{u} &= (\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) \cdot A(\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) \\
&= (\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) \cdot (\alpha A\mathbf{u}_1 + \beta A\mathbf{u}_2) \\
&= (\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) \cdot (\alpha\mu_1\mathbf{u}_1 + \beta\mu_2\mathbf{u}_2) \\
&= \alpha^2\mu_1\mathbf{u}_1 \cdot \mathbf{u}_1 + \alpha\beta(\mu_1 + \mu_2)\mathbf{u}_1 \cdot \mathbf{u}_2 + \beta^2\mu_2\mathbf{u} \cdot \mathbf{u}_2) \\
&= \alpha^2\mu_1 + \beta^2\mu_2 \ .
\end{aligned} \tag{5.12}$$

The last equality uses the fact that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthonormal. From this and (5.11), we see that $\mathbf{u} \cdot A\mathbf{u}$ *is a weighted average of* $\mathbf{u}_1 \cdot A\mathbf{u}_1$ *and* $\mathbf{u}_2 \cdot A\mathbf{u}_2$.

Now let's assume that $\mu_1 \le \mu_2$, which we can do by switching the subscripts if need be. Then since any weighted average of numbers lies between the minimum and maximum,

$$\mu_1 \le \alpha^2\mu_1 + \beta^2\mu_2 \le \mu_2 \ .$$

That is, for all unit vectors $\mathbf{u}$,

$$\mu_1 \le \mathbf{u} \cdot A\mathbf{u} \le \mu_2 \ . \tag{5.13}$$

Moreover, you see from (5.12) that one makes $\mathbf{u} \cdot A\mathbf{u}$ as large as possible by choosing $\alpha = 0$ and $\beta = \pm 1$, which means choosing $\mathbf{u} = \pm\mathbf{u}_2$. Likewise, one makes $\mathbf{u} \cdot A\mathbf{u}$ as small as possible by choosing $\alpha = \pm 1$ and $\beta = 0$, which means choosing $\mathbf{u} = \pm\mathbf{u}_1$. When $\mu_1 < \mu_2$, all other choices give something in between.

Applying this to the case in which $A = H_f(\mathbf{x}_0)$, we have the following theorem:

Theorem 3 (Directions of minimum and maximum curvature) *Let f be any function whose second partial derivatives are continuous in a neighborhood of $\mathbf{x}_0$. Let μ_1 and μ_2 be the eigenvalues of $H_f(\mathbf{x}_0)$, and choose the subscripts so that $\mu_1 \le \mu_2$. Then for any unit vector $\mathbf{u}$.*

$$\mu_1 \ \le \ \frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\mathbf{x}_0 + t\mathbf{u})\bigg|_{t=0} \ \le \ \mu_2 \tag{5.14}$$

Moreover, let $\mathbf{u}_1$ and $\mathbf{u}_2$ be normalized eigenvectors corresponding to μ_1 and μ_2 respectively. Then, when $\mu_1 < \mu_2$, the maximal value in (5.14) is obtained with, and only with, $\mathbf{u} = \pm\mathbf{u}_2$, and the minimal value is obtained with, and only with, $\mathbf{u} = \pm\mathbf{u}_2$

We will often apply Theorem 3 at critical points. Recall that a critical point $\mathbf{x}_0$ is one at which the gradient is zero, and therefore the tangent plane is flat. If the surface $z = f(\mathbf{x})$ "curves upward" in all directions through $\mathbf{x}_0$, then $\mathbf{x}_0$ is a local minimum of f. On the other hand, if the surface $z = f(\mathbf{x})$ "curves downward" in all directions through $\mathbf{x}_0$, then $\mathbf{x}_0$ is a local maximum of f.

Now suppose the μ_1 and μ_2 are the two eigenvalues of $H_f(\mathbf{x}_0)$, and are labeled so that $\mu_1 \le \mu_2$. If $\mu_1 > 0$, then by Theorem 3,

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\mathbf{x}_0 + t\mathbf{u})\bigg|_{t=0} \ge \mu_1 > 0$$

and so the the surface curves upward in every direction (and $\mathbf{x}_0$ is a local minimum).

Likewise if $\mu_2 < 0$, then by Theorem 3,

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\mathbf{x}_0 + t\mathbf{u})\bigg|_{t=0} \le \mu_2 < 0$$

and so the the surface curves downward in every direction (and $\mathbf{x}_0$ is a local maximum).

However, there are other possibilities: If $\mu_1 < 0 < \mu_2$, the graph of $z = f(\mathbf{x})$ curves downward in some directions and upward in others at $\mathbf{x}_0$. We will describe this situation by saying that $\mathbf{x}_0$ is a *saddle point* of f:

Definition (Saddle point) Let f be a function with continuous second partial derivatives. Let $\mathbf{x}_0$ be a critical point of f. If $H_f(\mathbf{x}_0)$ has at least one strictly negative and one strictly positive eigenvalue at $\mathbf{x}_0$, then $\mathbf{x}_0$ is a *saddle point of f*.

Note that for $\mathbf{x}_0$ to be a saddle point, it must be a critical point; i.e., $\nabla f(\mathbf{x}_0) = 0$. Otherwise, it is not a saddle point no matter what the eigenvalues of the Hesisan are.

Example 5 (Computing minimum and maximum curvature) Consider the function $f(x, y) = 2x^2 - 2xy + y^2$, and $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. We easily compute that

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 4x - 2y \\ 2y - 2x \end{bmatrix} \qquad \text{and} \qquad H_f(\mathbf{x}_0) = \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} .$$

We see that the tangent plane is horizontal at $\mathbf{x}_0$, and nowhere else, since only for $x = 0$ and $y = 0$ is the gradient zero. If we compute the eigenvalues of $H_f(\mathbf{x}_0)$, we can learn how the graph of $z = f(x, y)$ curves away from this tangent plane.

The eigenvalues of a square matrix A are the roots of its characteristic polynomial $p(t)$, which is given by $\det(A - tI)$. The characteristic polynomial of the Hessian matrix $H_f(\mathbf{x}_0)$

$$p(t) = \det\left(\begin{bmatrix} 4 - t & -2 \\ -2 & 2 - t \end{bmatrix} \right) = (4 - t)(2 - t) - 4 = t^2 - 6t + 4 .$$

Completing the square, we easily find that the roots of this polynomial are

$$\mu_1 = 3 - \sqrt{5} \qquad \text{and} \qquad \mu_2 = 3 + \sqrt{5} .$$

Notice that both μ_1 and μ_2 are strictly positive. This means that no matter how you slice through the graph at $\mathbf{x}_0$, it is *curving upwards. Therefore,* $f(\mathbf{x}_0) \leq f(\mathbf{x})$ *for all* $\mathbf{x}$ *near* $\mathbf{x}_0$*, this means that* $\mathbf{x}_0$ *is a local minimum of* f*.*

Exercises

1. Let $f(x, y) = x^3 y - 2xy + y$, and let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

(a) Compute $H_f(\mathbf{x})$ and $H_f(\mathbf{x}_0)$.

(b) Compute the minimum and maximum curvatures of f at $\mathbf{x}_0$.

(c) Check that $\mathbf{x}_0$ is a critical point of f. Is it a local minimum, maximum, or saddle point?

(d) Compute $\|H_f(\mathbf{x}_0)\|$.

(e) Let U be the region given by $|x - x_0| < 1$ and $|y - y_0| < 1$. Find a value M so that

$$|f(\mathbf{x}) - (f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0))| \leq (M/2)|\mathbf{x} - \mathbf{x}_0|^2$$

for all $\mathbf{x}$ in U.

(f) Find all of the critical points of f, and determine whether each is a local minimum, maximum, or saddle point.

2 Let $f(x, y) = \sin(\pi(x^2 - xy))$, and let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(a) Compute $H_f(\mathbf{x})$ and $H_f(\mathbf{x}_0)$.

(b) Compute the minimum and maximum curvatures of f at $\mathbf{x}_0$.

(c) Check that $\mathbf{x}_0$ is a critical point of f. Is it a local minimum, maximum, or saddle point?

(d) Compute $\|H_f(\mathbf{x}_0)\|$.

(e) Let U be the region given by $|x - x_0| < 1$ and $|y - y_0| < 1$. Find a value M so that

$$|f(\mathbf{x}) - (f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0))| \leq (M/2)|\mathbf{x} - \mathbf{x}_0|^2$$

for all $\mathbf{x}$ in U.

3. Let $f(x, y) = (x^2 + y^2)^2 - 2xy$, and let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(a) Compute $H_f(\mathbf{x})$ and $H_f(\mathbf{x}_0)$.
(b) Compute the minimum and maximum curvatures of f at $\mathbf{x}_0$.
(c) Check that $\mathbf{x}_0$ is a critical point of f. Is it a local minimum, maximum, or saddle point?
(d) Compute $\|H_f(\mathbf{x}_0)\|$.
(e) Let U be the region given by $|x - x_0| < 1$ and $|y - y_0| < 1$. Find a value M so that

$$|f(\mathbf{x}) - (f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0))| \leq (M/2)|\mathbf{x} - \mathbf{x}_0|^2$$

for all $\mathbf{x}$ in U.

4. Let $f(x, y) = x^3 y - 2xy^3 + y^4$, and let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.
(a) Compute $H_f(\mathbf{x})$ and $H_f(\mathbf{x}_0)$.
(b) Compute the minimum and maximum curvatures of f at $\mathbf{x}_0$.
(c) Check that $\mathbf{x}_0$ is a critical point of f. Is it a local minimum, maximum, or saddle point?
(d) Compute $\|H_f(\mathbf{x}_0)\|$.
(e) Let U be the region given by $|x - x_0| < 1$ and $|y - y_0| < 1$. Find a value M so that

$$|f(\mathbf{x}) - (f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0))| \leq (M/2)|\mathbf{x} - \mathbf{x}_0|^2$$

for all $\mathbf{x}$ in U.

Section 6: Second derivartives

6.1: The Hessian and Taylor's Theorem with Remainder

Before plunging in with two variables, let us refresh our memories about Taylor approximation in one variable.

Let $g(t)$ be a twice continuously differentiable function of the single variable t. Then for any given t_0,

$$g(t) = g(t_0) + g'(t_0)(t - t_0) + \frac{1}{2}g''(c)(t - t_0)^2 \ , \tag{6.1}$$

where c is some number lying between t_0 and t. Here, we are using the Lagrange form of the remainder.

There are two natural approximations to make: First, if we simply throw out the remainder term, we get

$$g(t) \approx g(t_0) + g'(t_0)(t - t_0) \ , \tag{6.2}$$

which will be a good approximation if t is very close to t_0. This is the tangent line approximation; if you graph the linear function on the right, you get the tangent line to the graph of g at t_0.

While the tangent line approximation is very useful, it is pretty drastic to just throw out the whole remainder term. We have to do something about it, since Lagrange does not – and cannot – tell us the value of c. But since c lies between t_0 and t, c will be close to t_0 when t is close to t_0. Hence if the second derivatives of g are continuous, it makes sense to consider the approximation

$$g''(c) \approx g''(t_0) \ . \tag{6.3}$$

Using this approximation in the exact formula (6.1), we get

$$g(t) \approx g(t_0) + g'(t_0)(t - t_0) + \frac{1}{2}g''(t_0)(t - t_0)^2 \ . \tag{6.4}$$

This is the best quadratic approximation to g. The graph of the quadratic function on the right is the *tangent parabola* to the graph of g at t_0.

As you may expect, making the approximation (6.3) is much better than simply discarding the remainder term altogether, and so the tangent parabola gives a much better fit to the graph of g than does the tangent line. In particular, it accurately represents the curvature of the the graph at t_0, and this is missing altogether in the tangent line approximation.

Example 1 (Tangent lines and tangent parabolas in one variable) Consider the function $g(t) = \sin(t)$ and $t_0 = \pi/4$. Then

$$g(t_0) = g'(t_0) = \frac{1}{\sqrt{2}} \qquad \text{and} \qquad g''(t_0) = -\frac{1}{\sqrt{2}} \ .$$

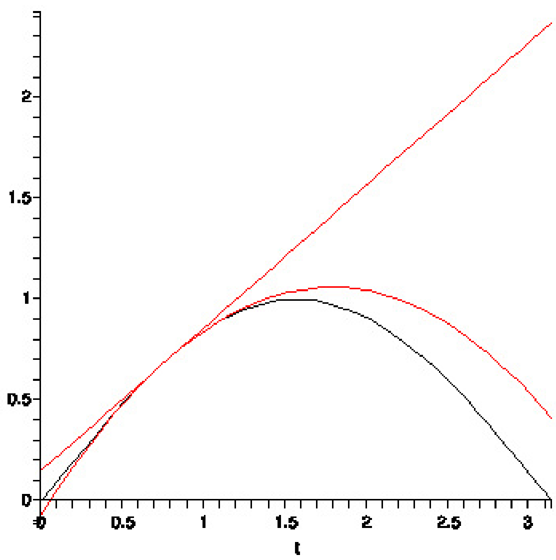
The tangent line approximation to $\sin(t)$ at $\pi/4$ is therefore

$$\sin(t) \approx \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}(t - \pi/4) \ ,$$

and the tangent parabola approximation to $\sin(t)$ at $\pi/4$ is therefore

$$\sin(t) \approx \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}(t - \pi/4) - \frac{1}{2\sqrt{2}}(t - \pi/4)^2 \ .$$

Here is a graph of $\sin(t)$ together with its tangent line and tangent parabola at $t = \pi/4$, in which you can see how the tangent parabola fits nicely over a much wider interval about $\pi/4$:



The region graphed is $0 \le t \le \pi$, and the curve that passes through 0 at both $t = 0$ and $t = \pi$ is the graph of $\sin(t)$. Notice that the graph of $\sin(t)$ and its tangent parabola are essentially indistinguishable over a reasonably wide interval around $t = \pi/4$.

Now we explain how this approximation scheme can be extended to two (or more) variables: Let f be any function whose second partial derivatives are continuous in an open and convex set U in $I\!R^2$.

Fix some $\mathbf{x}$ in U, and let $\mathbf{v} = \mathbf{x} - \mathbf{x}_0$. Define

$$g(t) = f(\mathbf{x}_0 + t\mathbf{v}) \ . \tag{6.5}$$

The convexity of U ensures that the line segment running from $\mathbf{x}_0$ to $\mathbf{x}$ is in the domain of definition of f, so that g is defined for $0 \le t \le 1$.

Now, by Taylor's Theorem with remainder, exactly as above,

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(c)t^2 \tag{6.6}$$

for some c between 0 and t.

Let us evaluate all of the terms in (6.6) for $t = 1$: First of all, $g(1) = f(\mathbf{x})$ and $g(0) = f(\mathbf{x}_0)$. Next, by Theorem 1 of Section 2,

$$g'(0) = \mathbf{v} \cdot \nabla f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \ ,$$

and then by Theorem 1 of Section 5,

$$g''(c) = \mathbf{v} \cdot [H_f(\mathbf{x}_0 + c(\mathbf{x} - \mathbf{x}_0))]\, \mathbf{v} = (\mathbf{x} - \mathbf{x}_0) \cdot [H_f(\mathbf{x}_0 + c(\mathbf{x} - \mathbf{x}_0))]\, (\mathbf{x} - \mathbf{x}_0) \ .$$

Plugging all of this back into (6.6), we have proved the following Theorem:

Theorem 1 (Taylor's Theorem With Remainder in More Variables) *Let f be a function with continuous second partial derivatives defined on an open convex set U in $I\!R^2$. Then for any $\mathbf{x}$ and $\mathbf{x}_0$ in U, the following identity holds:*

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot [H_f(\mathbf{x}_0 + c(\mathbf{x} - \mathbf{x}_0))]\, (\mathbf{x} - \mathbf{x}_0) \ , \qquad (6.7)$$

where c is some number with $0 \le c \le 1$.

Notice that we have gotten our two variable version of Taylor's Theorem with remainder directly from the one variable version by the slicing method. We could do the same thing for higher order Taylor approximations; the slicing method provides powerful leverage. But before moving on to higher orders, let us analyze the identity provided by Theorem 1.

There are two approximations we can make, starting from the identity (6.7). First, we can simply discard the remainder term altogether. This gives us

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \qquad (6.8)$$

which is nothing other than the familiar tangent plane approximation.

We can also appeal to the continuity of the second partial derivatives of f to make the approximation

$$H_f(\mathbf{x}_0 + c(\mathbf{x} - \mathbf{x}_0)) \approx H_f(\mathbf{x}_0) \ . \qquad (6.9)$$

Inserting this in (6.7), we get the approximation

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot [H_f(\mathbf{x}_0)]\, (\mathbf{x} - \mathbf{x}_0) \ , \qquad (6.10)$$

The function on the right in (6.10) is a *quadratic function of x and y*. That is, it is a polynomial in x and y in which no term has a total power higher than 2.

The general quadratic function $q(x, y)$ has the form

$$q(x, y) = A + Bx + Cy + D\frac{x^2}{2} + Exy + F\frac{y^2}{2} \ , \qquad (6.11)$$

where A through F are given coefficients.

Definition (The tangent quadratic surface) Let f be a function defined on an open set U in $I\!R^2$ whose second partial derivatives are continuous functions in U. Then for any $\mathbf{x}_0$ in U the function the graph of

$$z = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot [H_f(\mathbf{x}_0)]\, (\mathbf{x} - \mathbf{x}_0) \qquad (6.12)$$

is the *tangent quadratic surface* to f at $\mathbf{x}_0$

The tangent quadratic surface is the higher dimensional analog of the tangent parabola. Just as in Example 1, the tangent quadratic surface will fit much more closely than the tangent plane.

Example 2 (Tangent planes and tangent quadratic surfaces) Consider the function

$$f(x,y) = 3x^2y - 6x - y^3 \ .$$

We will now work out the tangent plane approximation to $z = f(\mathbf{x})$ and the quadratic surface approximation to $z = f(\mathbf{x})$ at the point $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

We compute

$$\nabla f(x,y) = \begin{bmatrix} 6xy - 6 \\ 3x^2 - 3y^2 \end{bmatrix} \qquad \text{and} \qquad H_f(\mathbf{x}) = 6 \begin{bmatrix} y & x \\ x & -y \end{bmatrix} \ .$$

Evaluating at $x = 1$ and $y = 1$,

$$f(\mathbf{x}_0) = -4 \qquad \nabla f(\mathbf{x}_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{and} \qquad H_f(\mathbf{x}_0) = 6 \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \ .$$

Notice that $\mathbf{x}_0$ is a critical point, and therefore the tangent plane approximation reduce to only

$$f(x,y) \approx -4 \ .$$

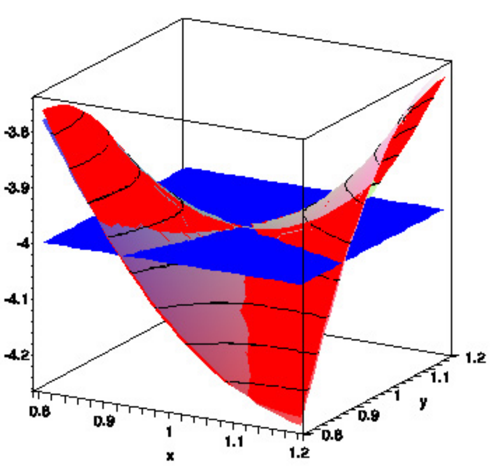
On the other hand, the tangent quadratic surface approximation works out to

$$\begin{aligned} f(x,y) &\approx -4 + 3 \begin{bmatrix} x-1 \\ y-1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x-1 \\ y-1 \end{bmatrix} \\ &= 3x^2 + 6xy - 3y^2 - 12x + 2 \ . \end{aligned}$$

As you see, the right hand side has the form (6.11). Now let us graph f together with its tangent plane and its tangent quadratic surface over a region around $\mathbf{x}_0$. We will first look in the region with

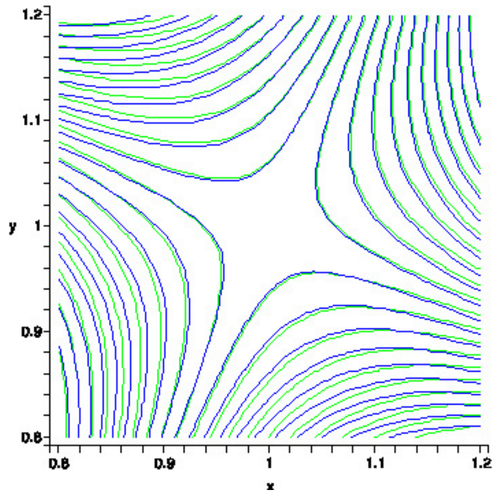
$$|x-1| \le 0.2 \qquad \text{and} \qquad |y-1| \le 0.2 \ .$$

Here is the result:



It *looks* like one level plane, and one curved surface. but that is only because the actual graph of $z = f(\mathbf{x})$ and its tangent quadratic surface fit so well in the region that is graphed. If you look in the upper left corner, you can see the difference between the two – barely.

To get another look at how well these surfaces match, let us look at superimposed contour plots for the graph of $z = f(\mathbf{x})$ and its tangent quadratic surface:

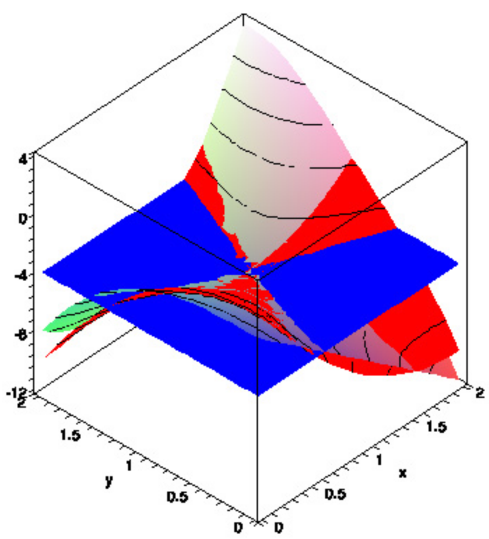


The two sets of contour curve begin to differ as one moves to the corners, but in the center, the contour plots look pretty much the same. (A contour plot of the tangent plane, which is level, would be devoid of content.)

Let us try and push our luck, and draw a graph of f, its tangent plane and its tangent quadratic surface over a wider region around $\mathbf{x}_0$. Let us do this for

$$|x - 1| \leq 1 \qquad \text{and} \qquad |y - 1| \leq 1 \ .$$

Here is the result:



Now you clearly see the graphs of $z = f(\mathbf{x})$ and the tangent quadratic surface pulling apart. But then, we are graphing a reasonably large region.

From the graphs we have just above, you might get the idea that the tangent quadratic surface approximation is good, while the tangent plane approximation is just plain rotten. However, at a critical point like $\mathbf{x}_0$, where a contour plot of the tangent plane is devoid of content, we looking at the tangent plane approximation in the worst possible light. At a non critical point, it is *much* more informative.

To see this, lets us look at the tangent plane and the tangent quadratic surface at a non critical point.

For this example, let us now take $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Evaluating that this point, we find

$$f(\mathbf{x}_0) = -8 \qquad \nabla f(\mathbf{x}_0) = \begin{bmatrix} 6 \\ -9 \end{bmatrix} \qquad \text{and} \qquad H_f(\mathbf{x}_0) = 6 \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} .$$

We now compute, as before, that the tangent plane approximation is

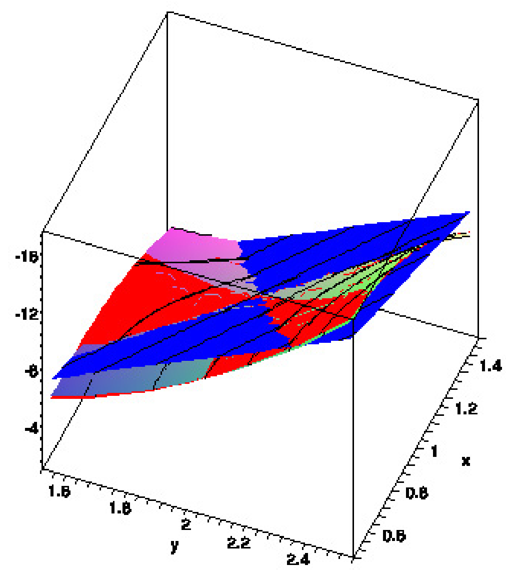
$$f(x, y) \approx 6x - 9y + 4 ,$$

and that the quadratic surface approximation is

$$f(x, y) \approx 6x^2 + 6xy - 6y^2 - 18x + 9y - 2 .$$

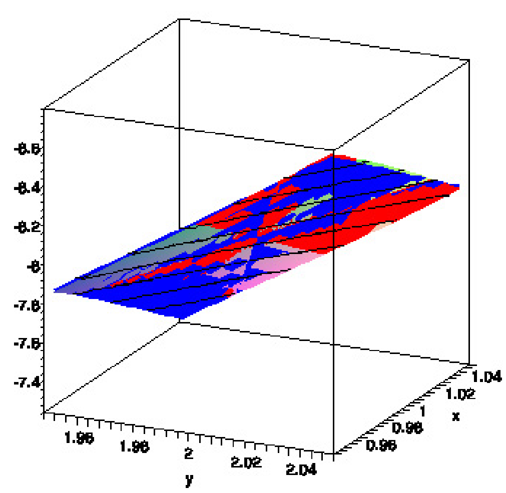
Here is a graph of $z = f(\mathbf{x})$ together with tangent plane and the tangent quadratic surface in the region

$$|x - 1| \le 0.5 \qquad \text{and} \qquad |y - 2| \le 0.5 .$$



All three surfaces are reasonably close together. Let us zoom in by a factor of 10, and draw the graphs for

$$|x - 1| \le 0.05 \qquad \text{and} \qquad |y - 2| \le 0.05 .$$



Now all three graphs are virtually indistinguishable.

6.2: Accuracy of the tangent approximations

What we have seen in the previous example raises the following important questions about the accuracy of the two tangent approximations:

• *In how large a region around $\mathbf{x}_0$ is the tangent plane approximation useful? In how large a region around $\mathbf{x}_0$ is the tangent quadratic surface approximation useful?*

For example, suppose we are drawing a graph in which vertical distances of less than 1/100 will amount to less than one pixel. Then what we want to know is: *How small does $|\mathbf{x}-\mathbf{x}_0|$ have to be to ensue that the remainder terms in our approximations are no more than* 1/100*?*

In both cases, the remainder terms involve the Hessian. We are going to throw out either

$$H_f(\mathbf{x}_0 + c(\mathbf{x}-\mathbf{x}_0))$$

or else

$$H_f(\mathbf{x}_0 + c(\mathbf{x}-\mathbf{x}_0)) - H_f\mathbf{x}_0) \ ,$$

and estimating the size what we are ignoring therefore involves estimating the "size" of the matrices. This raises the question:

• *How does one measure the size of a matrix?*

One measure of the size of a matrix A is in terms of a quantity called the *Hilbert–Schmidt norm** of A, and denoted $\|A\|_{\mathrm{HS}}$:

Definition (The Hilbert–Schmidt norm of a matrix) For an $m \times n$ matrix A, the *Hilbert–Schmidt norm* of A is the number $\|A\|_{\mathrm{HS}}$ which is defined by

$$\|A\|_{\mathrm{HS}} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}^2} \ .$$

Example 3 (Computing the Hilbert Schmidt norm of a matrix) Let $A = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$. To compute $\|A\|_{\mathrm{HS}}$, we first sum the squares of the entries, finding

$$4 + 9 + 9 + 16 = 38 \ .$$

We then take the square root, and have

$$\|A\|_{\mathrm{HS}} = \sqrt{38} \ .$$

* The Hilbert–Schmidt norm is sometimes also called the Frobenius norm. There are other measures of the size of matrices, also called norms. The most important of these is the *operator norm*, which is usually just called "the" matrix norm. We will encounter it later on.

It is pretty much as easy as computing the length of a vector. In fact, it is pretty much the same thing.

Here is why this is relevant: The terms that we "throw away" in our approximation are of the form $\mathbf{v} \cdot A\mathbf{v}$. How big can they be? The following Lemma gives a simple answer in terms of the Hilbert–Schmidt norm.

Lemma 1 (The Hilbert–Schmidt norm of a matrix and its "size")) *Let A be an $n \times m$ matrix. Then for any $\mathbf{v}$ in $I\!R^n$,*

$$|\mathbf{v} \cdot A\mathbf{v}| \le \|A\|_{\mathrm{HS}} |\mathbf{v}|^2 \ . \tag{6.13}$$

Proof: Write A as a list of its rows: $A = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_m \end{bmatrix}$. Then $A\mathbf{v} = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{v} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{v} \end{bmatrix}$. Now, by the Schwarz inequality, $|\mathbf{r}_i \cdot \mathbf{v}| \le |\mathbf{r}_i||\mathbf{v}|$, and so

$$|A\mathbf{v}| = \sqrt{A\mathbf{v} \cdot A\mathbf{v}} = \sqrt{\sum_{i=1}^{m} (\mathbf{r}_i \cdot \mathbf{v})^2} \le |\mathbf{v}| \sqrt{\sum_{i=1}^{m} (|\mathbf{r}_i|^2} \ .$$

But by the Schwarz inequality again, $|\mathbf{v} \cdot A\mathbf{v}| \le |\mathbf{v}||A\mathbf{v}|$. Combining results, we have

$$|\mathbf{v} \cdot A\mathbf{v}| \le \left(\sqrt{\sum_{i=1}^{m} |\mathbf{r}_i|^2} \right) |\mathbf{v}|^2 \ .$$

But since for each i, $|\mathbf{r}_i|^2 = \sum_{j=1}^{n} A_{i,j}^2$, this is the same as (6.13). $\square$

We can now combine Lemma 1 and Theorem 1 to prove a useful result on the accuracy of the tangent plane approximation:

Theorem 2 (Accuracy of the tangent plane approximation) *Let f be defined on an open convex subset U of $I\!R^2$. Suppose that all of the partial derivatives of f of order 2 exist and are continuous throughout U so that, in particular, the Hessian matrix H_f is defined throughout U, and so is $\|H_f(\mathbf{x})\|_{\mathrm{HS}}$. Suppose M is a number such that*

$$\|H_f(\mathbf{x})\|_{\mathrm{HS}} \le M \qquad \text{for all} \quad \mathbf{x} \quad \text{in } U \ .$$

Then for any $\mathbf{x}_0$ and $\mathbf{x}$ in U,

$$\left| f(\mathbf{x}) - [f(\mathbf{x}_0) + \cdot \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)] \right| \le \frac{M}{2} |\mathbf{x} - \mathbf{x}_0|^2$$

Proof: By Theorem 1,

$$\left|f(\mathbf{x}) - [f(\mathbf{x}_0) + \cdot\nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)]\right| \le \frac{1}{2}\left|(\mathbf{x}-\mathbf{x}_0)\cdot[H_f(\mathbf{x}_0 + c(\mathbf{x}-\mathbf{x}_0))](\mathbf{x}-\mathbf{x}_0)\right|$$

By the Lemma and the definition of M,

$$|(\mathbf{x}-\mathbf{x}_0)\cdot[H_f(\mathbf{x}_0 + c(\mathbf{x}-\mathbf{x}_0))](\mathbf{x}-\mathbf{x}_0)| \le M|\mathbf{x}-\mathbf{x}_0|^2 \ .$$

This proves the Theorem. $\square$

Example 4 (How fine a mesh should we use with a piecewise tangent plane approximation?) Suppose that we want to analyze the surface $z = f(x)$ over some region U.

A useful strategy for many purposes is "chop up" U into small square "tiles", and replace f in each square by its tangent plane approximation at the center of the square. Suppose each little square in the subdivision has a width 2δ.

• *How small should δ be to ensure that in each little square, the difference between f and its tangent plane approximation taking $\mathbf{x}_0$ to be the center of the square) is no more than $1/100$?*

Let us work this out for

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

We compute

$$\nabla f(x,y) = \begin{bmatrix} 4x - y - 3 \\ 2y - x + 1\end{bmatrix} \qquad \text{and} \qquad H_f(x,y) := \begin{bmatrix} 4 & -1 \\ -1 & 2\end{bmatrix} \ .$$

Notice that H_f is a constant matrix. This is because f itself was a quadratic function. We now compute

$$\|H_f\|_{\mathrm{HS}} = \sqrt{16+1+1+4} = \sqrt{22} \ .$$

Therefore, for *any* $\mathbf{x}_0$ in $I\!R^2$,

$$|f(\mathbf{x}) - [f(\mathbf{x}_0) + \cdot\nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)]| \le \frac{\sqrt{22}}{2}|\mathbf{x}-\mathbf{x}_0|^2 \ .$$

Now if $\mathbf{x}_0$ is the center of a square with side length 2δ, and $\mathbf{x}$ is any point in the square, then

$$|\mathbf{x}-\mathbf{x}_0| \le \sqrt{2}\delta \ ,$$

and it is this large only if $\mathbf{x}$ is in one of the corners of the square, which is the worst case. Therefore,

$$|f(\mathbf{x}) - [f(\mathbf{x}_0) + \cdot\nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)]| \le \sqrt{22}\delta^2$$

whenever $\mathbf{x}$ belongs to a square with side length 2δ centered at $\mathbf{x}_0$.

We need $\sqrt{22}\delta^2 \le 1/100$, or $\delta \le 0.046173....$ In particular, we get the desired accuracy if we take $\delta = 0.047$, so that $2\delta = 0.094$. This is more than $1/11$, so if we "chop up" our region into tiles of this size, we can replace f in each tile by its tangent plane approximation, the resulting error will be less than $1/100$ everywhere.

We now turn to the tangent quadratic surface approximation. Suppose that f has continuous second derivatives everywhere in an open convex set U in $I\!R^2$. Our goal now is to determine the size of the error that we might make in using the approximation (6.9); i.e.,

$$H_f(\mathbf{x}_0 + c(\mathbf{x} - \mathbf{x}_0)) \approx H_f(\mathbf{x}_0)$$

in the identity (6.7) from Theorem 1.

For this purpose, fix $\mathbf{x}_0$ in U, and define the function $m(\mathbf{y})$ in U by

$$m(\mathbf{y}) = \|H_f(\mathbf{y}) - H_f(\mathbf{x}_0)\|_{\mathrm{HS}} \ . \tag{6.14}$$

Because the second partial derivatives of f are continuous, and since m is built out of them by summing squares of differences, and then taking a square root, m is a continuous function of $\mathbf{y}$.

Now for all $r > 0$, define

$$M(r) = \max\{ \ m(\mathbf{y}) \ : \ |\mathbf{y} - \mathbf{x}_0| \le r \ \} \ . \tag{6.15}$$

Notice that the set of point $\mathbf{y}$ with $|\mathbf{y} - \mathbf{x}_0| \le r$ is just the closed disk of radius r centered on $\mathbf{x}_0$. Since continuous functions always have a maximum value on any closed bounded set, $M(r)$ is properly defined. Notice that $M(r)$ is a monotone increasing function of r, since as r is increased, we are taking the maximum over a larger disk in forming M.

Moreover, since m is a continuous function of $\mathbf{y}$, and since $m(\mathbf{x}_0) = 0$,

$$\lim_{\mathbf{y}\to\mathbf{x}_0} m(\mathbf{y}) = 0 \ .$$

It follows directly from this, and the definition of limiting values, that

$$\lim_{r\to 0} M(r) = 0 \ .$$

Now by Theorem 1, for any $\mathbf{x}$ in U,

$$\begin{aligned} &\left| f(\mathbf{x}) - \left[f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\cdot(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)\cdot[H_f(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0) \right] \right| \\ &\le \frac{1}{2}\left| (\mathbf{x} - \mathbf{x}_0)\cdot[H_f(\mathbf{y})) - H_f(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0) \right| \ , \end{aligned}$$

where $\mathbf{y} = \mathbf{x}_0 + c(\mathbf{x} - \mathbf{x}_0)$.

Notice that $|\mathbf{y} - \mathbf{x}_0| \le |\mathbf{x} - \mathbf{x}_0|$. Therefore, by Lemma 1 and the definition of M, the right hand side is no greater than $\frac{1}{2}|\mathbf{x} - \mathbf{x}_0|^2 M(|\mathbf{x} - \mathbf{x}_0|)$. This proves the following Theorem:

Theorem 3 (Accuracy of the tangent quadratic surface approximation) *Let f be defined on an open convex subset U of $I\!R^2$. Suppose that all of the partial derivatives of*

f of order 2 exist and are continuous throughout U. Let $M(r)$ be defined by (6.14) and (6.15), so that in particular $M(r)$ is continuous and $M(0) = 0$. Let

$$h(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot [H_f(\mathbf{x}_0)]\,(\mathbf{x} - \mathbf{x}_0)$$

be the tangent quadratic approximation to f at $\mathbf{x}_0$. Then for any $\mathbf{x}_0$ and $\mathbf{x}$ in U,

$$|f(\mathbf{x}) - h(\mathbf{x})| \leq \frac{1}{2} M(|\mathbf{x} - \mathbf{x}_0)|\mathbf{x} - \mathbf{x}_0|^2 \ . \tag{6.16}$$

6.3 Twice differentiable functions and tangent quadric surfaces

Recall that a function f on $I\!R^2$ is differentiable at $\mathbf{x}_0$ in case there is a linear function h so that

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{|f(\mathbf{x}) - h(\mathbf{x})|}{|\mathbf{x} - \mathbf{x}_0|} = 0 \ .$$

Definition (Twice differentiable) A function f on $I\!R^2$ is *twice differentiable at* $\mathbf{x}_0$ in case there is a quadratic function h so that

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{|f(\mathbf{x}) - h(\mathbf{x})|}{|\mathbf{x} - \mathbf{x}_0|^2} = 0 \ . \tag{6.17}$$

It is now easy to prove the following Theorem;

Theorem 4 (Twice differentiable functions) *Let f be defined on an open convex subset U of $I\!R^2$. Suppose that all of the partial derivatives of f of order 2 exist and are continuous throughout U. Then U is twice differentiable at each $\mathbf{x}_0$ in U, and moreover,*

$$h(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot [H_f(\mathbf{x}_0)]\,(\mathbf{x} - \mathbf{x}_0)$$

is the unique quadratic function for which (6.17) is true.

Proof Let h be given by the tangent quadratic surface approximation. By Theorem 3,

$$0 \leq \frac{|f(\mathbf{x}) - h(\mathbf{x})|}{|\mathbf{x} - \mathbf{x}_0|^2} \leq \frac{1}{2} M(|\mathbf{x} - \mathbf{x}_0) \ .$$

Since $\lim_{r \to 0} M(r) = 0$, (6.17) is true by comparison.

To see that (6.17) cannot hold true for any other quadratic function, suppose that h_1 and h_2 are two such quadratic functions. Then

$$h_1(\mathbf{x}) - h_2(\mathbf{x}) = (h_1(\mathbf{x}) - f(\mathbf{x})) - (h_2(\mathbf{x}) - f(\mathbf{x}))$$

so that

$$|h_1(\mathbf{x}) - h_2(\mathbf{x})| \le |f(\mathbf{x}) - h_1(\mathbf{x})| + |f(\mathbf{x}) - h_2(\mathbf{x})| \ .$$

From this it follows that

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{|h_1(\mathbf{x}) - h_2(\mathbf{x})|}{|\mathbf{x} - \mathbf{x}_0|^2} = 0 \ .$$

It is not hard to see that this is only possible if all of the coefficients of h_1 and h_2 are the same so that the numerator cancels is identically zero. Hints for doing this are provided in the problems. ■

The uniqueness statement in Theorem 4 justifies the following definition

---

Definition (Best quadratic approximation) Let f be function on $I\!R^2$ that is twice differentiable function near $\mathbf{x}_0$. Then the quadratic function function

$$f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\cdot(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)\cdot[H_f(\mathbf{x}_0)]\,(\mathbf{x} - \mathbf{x}_0)$$

is *the best quadratic approximation to f at $\mathbf{x}_0$.*

Example 5 (Computing a quadratic approximation) This is the same as computing a tangent quadratic surface approximation, which we have done before. But another example will not hurt.

Let $f(x,y) = x^4 + y^4 + 4xy$, and let $(x_0, y_0) = (2,1)$. We found in a previous example that

$$\begin{aligned} f(2,1) &= 25 \\ \nabla f(2,1) &= \begin{bmatrix} 36 \\ 12 \end{bmatrix} \\ H_f((2,1) &= 4\begin{bmatrix} 12 & 1 \\ 1 & 3 \end{bmatrix} \ . \end{aligned}$$

Hence the quadratic approximation $h(x,y)$ is

$$h(x,y) = 25 + 36(x-2) + 12(y-1) + 24(x-2)^2 + 4(x-2)(y-1) + 6(y-1)^2 \ .$$

We *could* multiply this all out and collect terms to find

$$h(x,y) = 51 - 74x - 8y + 24x^2 + 4xy + 6y^2 \ .$$

However, it is generally much more useful to leave it in terms of $x - x_0$ and $y - y_0$.

6.4 The Implicit Function Theorem

what we have learned so far makes it easy to prove an important result about contour curves that we have discussed earlier:

Theorem 5 (The Implicit Function Theorem) *Let f be defined on an open convex subset U of $I\!R^2$. Suppose that all of the partial derivatives of f of order 2 exist and are continuous throughout U. Then for each $\mathbf{x}_0$ in U for which*

$$\nabla f(\mathbf{x}_0) \ne 0 \ ,$$

there is an $r_0 > 0$ and a differentiable curve $\mathbf{x}(t)$ defined for $|t| < r_0$ such that $\mathbf{x}(0) = \mathbf{x}_0$,

$$\mathbf{x}'(t) \neq 0 \qquad \text{and} \qquad f(\mathbf{x}(t)) = f(\mathbf{x}_0) \qquad \text{for all} \quad |t| < r_0 \ ,$$

and last but not least, such that every solution of $f(\mathbf{x}) = f(\mathbf{x}_0)$ with $|\mathbf{x} - \mathbf{x}_0| < r$ lies on the curve $\mathbf{x}(t)$ for $|t| < r_0$.

The fact that $\mathbf{x}(0) = \mathbf{x}_0$ and $\mathbf{x}'(t) \neq 0$ for any t means that the parameterized curve passes through $\mathbf{x}_0$, and doesn't just "sit there" as it would if $\mathbf{x}'(t)$ were zero for all t. That is, it is a genuine curve. The fact that $f(\mathbf{x}(t)) = f(\mathbf{x}_0)$ means that f is constant on this curve; i.e., it is a contour curve. And finally, the last part of the theorem says that *all* solutions of $f(\mathbf{x}) = f(\mathbf{x}_0)$ in some open disk centered on $\mathbf{x}_0$ lie on our curve $\mathbf{x}(t)$.

Clearly the condition that $\nabla f(\mathbf{x}) \neq 0$ is crucial. For example, if $\mathbf{x}_0$ is a saddle point of f, the solution set of $f(\mathbf{x}) = f(\mathbf{x}_0)$ has two branches, and cannot be parameterized as a single segment of a curve. However, the theorem says that as long as $\nabla f(\mathbf{x}) \neq 0$, nothing like this can happen.

The idea of the proof is simple enough: If $\nabla f(\mathbf{x}_0) \neq 0$, then there is a well defined "uphill direction", and perpendicular to that is the level direction. Therefore, the tangent line

$$\mathbf{x}_0 + t\,(\nabla f(\mathbf{x}_0))^{\perp}$$

at least starts out moving in a level direction. We are going to show that one can "correct" this by adding a small displacement $y(t)\nabla f(\mathbf{x}_0)$ parallel (or anti–parallel) to the gradient so that

$$\mathbf{x}(t) = t\,(\nabla f(\mathbf{x}_0))^{\perp} + y(t)\nabla f(\mathbf{x}_0) \tag{6.18}$$

is the level curve that we desire. At least near by to $\mathbf{x}_0$, everything on the "uphill side " will be "too high", and everything on the "downhill side" will be "too low", and so there are no solutions of $f(\mathbf{x}(t)) = f(\mathbf{x}_0)$ that are close to $\mathbf{x}_0$, but not on our curve. Finally, we show that the curve has the stated differentiability properties.

The theorem can be proved, in a more technical way, assuming only that the first partial derivatives are continuous. In most application, the second derivatives are just as well behaves as the first, so we do not loose very much in generality by requiring continuity of the second partial derivative – and the proof will be greatly simplified.

Proof of Theorem 5: We break the proof up into steps.

Step 1: (Choosing a small disk centered on $\mathbf{x}_0$ in which the gradient points in more or less the same direction, and the Hessian is not too big). Suppose that $\nabla f(\mathbf{x}_0) \neq 0$. Since we will make frequent reference to $\nabla f(\mathbf{x}_0)$, define

$$\mathbf{n} = \nabla f(\mathbf{x}_0) \qquad \text{and} \qquad \mathbf{b} = \mathbf{n}^{\perp} \ .$$

By continuity, for $\mathbf{x}$ close to $\mathbf{x}_0$, $\nabla(\mathbf{x}) \approx \nabla f(\mathbf{x}_0)$, so that $\nabla f(\mathbf{x}) \cdot \mathbf{n} \approx |\mathbf{n}|^2$. Therefore, we can choose an $r > 0$ so that when $|\mathbf{x} - \mathbf{x}_0| < r$, $\mathbf{x}$ is in U and

$$|\mathbf{x} - \mathbf{x}_0| < r \Rightarrow f(\mathbf{x}) \cdot \mathbf{n} \geq |\mathbf{n}|^2/2 \ . \tag{6.19}$$

For any numbers s and t, consider the point

$$\mathbf{x} = \mathbf{x}_0 + s\mathbf{n} + t\mathbf{b} \ .$$

Then $\mathbf{x} - \mathbf{x}_0 = s\mathbf{n} + t\mathbf{b}$, and

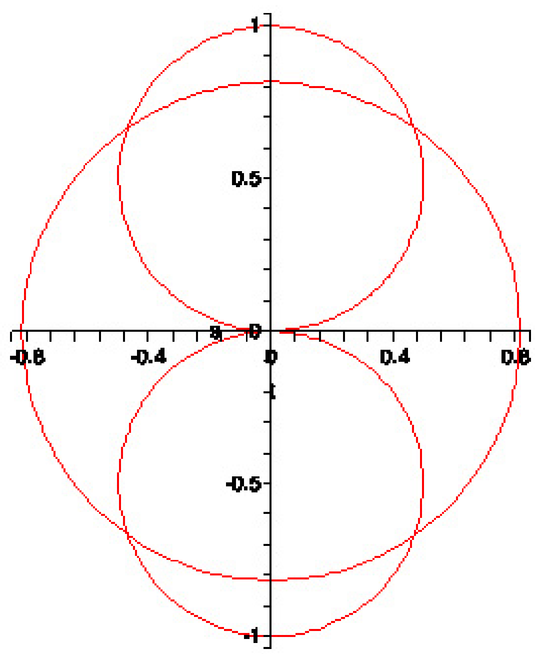
$$\nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) = \mathbf{n}\cdot(s\mathbf{n}+t\mathbf{b}) \ = s|\mathbf{n}|^2 \qquad \text{and} \qquad |\mathbf{x}-\mathbf{x}_0| = |\mathbf{n}|^2(s^2+t^2) \ . \tag{6.20}$$

Now let M be some number so that $M \ge \|H_f(\mathbf{x})\|_{\rm HS}$ in the disk of radius r about $\mathbf{x}_0$. Such an M exists since $\|H_f(\mathbf{x})\|_{\rm HS}$ is a continuous function in some closed disk about $\mathbf{x}_0$, and continuous functions always have a finite maximum on close bounded sets. Increasing M if need be, we can assume that $1/M < r$. We do so, as this will be convenient for the picture in the next step.

Step 2: (Determining the parts of the disk in which $f(\mathbf{x})$ is surely too big or too small) We will show that all solutions of $f(\mathbf{x}) = f(\mathbf{x}_0)$ for $|\mathbf{x}-\mathbf{x}_0| < r$ lie "squeezed" between two disks that are tangent to the line

$$\mathbf{x}_0 + t\mathbf{b} \ .$$

Here is a picture of what we are after:



The horizontal axis is the t axis, and the vertical axis is the s axis. The circle centered on the origin represents the set of s and t values for which $|\mathbf{x}-\mathbf{x}_0| < r$ with $\mathbf{x} = \mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}$. The t axis is the line $\mathbf{x}_0 + t\mathbf{b}$ which we expect to turn out to be the tangent line to our contour curve. We shall show in this step that there are two "forbidden circles" tangent to this line, as in the picture, such that $f(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}) > f(\mathbf{x}_0)$ for all (t,s) in the upper circle, and such that $f(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}) < f(\mathbf{x}_0)$. This will make it easy to show that there is a contour curve of f that gets "squeezed" between the circles. This squeezing will what forces it to have the right tangent line.

To validate the picture, we use Theorem 2 and (6.20) to conclude that

$$\left| f(\mathbf{x}) - [f(\mathbf{x}_0) + s|\mathbf{n}|^2] \right| \le \frac{M}{2}|\mathbf{n}|^2(s^2+t^2) \ .$$

In particular, $f(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}) \ge f(\mathbf{x}_0) + |\mathbf{n}|^2(s - (M/2)(s^2+t^2))$, so that

$$s - (M/2)(s^2+t^2) > 0 \Rightarrow f(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}) > f(\mathbf{x}_0) \ . \tag{6.21}$$

Completing the square, $s - (M/2)(s^2+t^2) > 0$ can be rewritten as

$$(s - 1/M)^2 + t^2 < 1/M^2 \ . \tag{6.22}$$

This the open disk bounded by the circle of radius $1/M$ centered on $(0, 1/M)$, which is tangent to the horizontal axis, as drawn. The inequality (6.21) is satisfied inside the circle, and so there are no points (t,s) inside this circle for which $f(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}) = f(\mathbf{x}_0)$.

Likewise,

$$f(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}) \le f(\mathbf{x}_0) + |\mathbf{n}|^2(s + (M/2)(s^2+t^2)) \ ,$$

and hence

$$s + (M/2)(s^2+t^2) < 0 \Rightarrow f(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}) < f(\mathbf{x}_0) \ .$$

Just as before, the inequality on the left can be rewritten as

$$(s + 1/M)^2 + t^2 < 1/M^2 \ . \tag{6.23}$$

This the open disk bounded by the circle of radius $1/M$ centered on $(0, 1/M)$, which is tangent to the horizontal axis, as drawn, and so there are no points (t,s) inside this circle for which $f(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}) = f(\mathbf{x}_0)$.

Step 3: (Finding the solutions of $f(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}) = f(\mathbf{x}_0)$ *using the Intermediate Value Theorem.)*

Fix any value of t with $|t| < 1/M$, and let $g(s) = f(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b})$. Clearly, g is a continuous function of s. Since $|t| < 1/M$, which is the radius of the "forbidden" circles, the point $(t, 1/M)$ lies inside the upper "forbidden circle", and so this point $g(1/M) > f(\mathbf{x}_0)$.

Likewise, as $(t, -1/M)$ lies in the lower "forbidden circle", and so $g(-1/M) < f(\mathbf{x}_0)$. Now, by the intermediate Value Theorem and the continuity of g, there is some value of s in between $-M$ and M for which $g(s) = f(\mathbf{x}_0)$.

In fact, for each t, there is *exactly one* such value of s. To see this, observe that $g(s)$ is strictly monotone increasing, so one it has passes through the level $f(\mathbf{x}_0)$, it cannot come back. To see that g is strictly monotone increasing, just check that its derivative is strictly positive:

$$g'(s) = \mathbf{n} \cdot \nabla(\mathbf{x}_0 + s\mathbf{n} + t\mathbf{b}) > |\mathbf{n}|^2/2$$

for all s and t under consideration, by (6.19). We now define $s(t)$ to the this unique s value, which of course can depend on t. We then also define

$$\mathbf{x}(t) = \mathbf{x}_0 + t\mathbf{b} + s(t)\mathbf{n} \ . \tag{6.24}$$

By the definition of $s(t)$, $f(\mathbf{x}(t)) = f(\mathbf{x}_0)$, and by the uniqueness that followed from the strict monotonicity of g, If $f(\mathbf{x}_0 + t\mathbf{b} + s\mathbf{n}) = f(\mathbf{x}_0)$ with $|t| < 1/M$ and $|s| < 1/M$, then it must be the case that $s = s(t)$. Therefore, *all solutions* of $f(\mathbf{x}) = f(\mathbf{x}_0)$ lies on the curve parameterized by (6.24) for $-1/M < t < 1/M$. We now see that the value of r_0 in the Theorem can be taken to be $1/M$, and we have our parameterization of the solution set of $f(\mathbf{x}) = f(\mathbf{x}_0)$ for $\mathbf{x} - \mathbf{x}_0|, r_0$.

Step 4: (The vector valued function $\mathbf{x}(t) = \mathbf{x}_0 + t\mathbf{b} + s(t)\mathbf{n}$ is a continuously differentiable curve) It remains to show that $\mathbf{x}(t)$ is differentiable, and that $\mathbf{x}'(t) \neq 0$. First consider $t = 0$. We claim that

$$\lim_{h\to 0} \frac{1}{h}(\mathbf{x}(h) - \mathbf{x}(0)) = \mathbf{b} \ . \tag{6.25}$$

Indeed,

$$\mathbf{x}(h) - \mathbf{x}(0) = s(h)\mathbf{n} + h\mathbf{b}$$

so that

$$\frac{1}{h}(\mathbf{x}(h) - \mathbf{x}(0)) = \mathbf{b} + \frac{s(h)}{h}\mathbf{n} \ .$$

Now, let $s_+(t)$ be the function whose graph is the lower arc of the upper "forbidden circle" and let $s_-(t)$ be the function whose graph is the upper arc of the lower "forbidden circle". Clearly

$$s_+(t) = \frac{1}{M} - \sqrt{\frac{1}{M^2} - t^2} \qquad \text{and} \qquad s_-(t) = -\frac{1}{M} + \sqrt{\frac{1}{M^2} - t^2} \ .$$

Moreover, $s_-(t) \le s(t) \le s_+(t)$ for all $|t| < 1/M$. Therefore, $\dfrac{s_-(h)}{h} \le \dfrac{s(h)}{h} \le \dfrac{s_+(h)}{h}$. By L'Hospital's rule,

$$\lim_{h\to 0} \frac{s_\pm(h)}{h} = (s_\pm)'(0) = 0 \ .$$

By the Squeeze Theorem, $\lim_{h\to 0}(s(h)/h) = 0$, and this proves (6.25). This means that $\mathbf{x}(t)$ is differentiable at $t = 0$ at least, and that $\mathbf{x}'(t) = \mathbf{b} \neq 0$.

Now it is easy to see that $\mathbf{x}(t)$ must be differentiable at each t_0 for which it is defined. Indeed, the same argument would show that the solution set of $f(\mathbf{x}) = f(\mathbf{x}(t_0))$ is a parameterized curve that is differentiable as it passes through $\mathbf{x}(t_0)$. But since $f(\mathbf{x}(t_0)) = f(\mathbf{x}_0)$, this is the same solution set that we have just parameterized above. From here one easily concludes that $\mathbf{x}(t)$ is differentiable at t_0, and hence at each t. ■

Exercises

1. Let $f(x, y) = x^3y - 2xy + y$, and let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

(a) Compute $H_f(\mathbf{x})$ and $H_f(\mathbf{x}_0)$.

(b) Compute the minimum and maximum curvatures of f at $\mathbf{x}_0$.

(c) Check that $\mathbf{x}_0$ is a critical point of f. Is it a local minimum, maximum, or saddle point?
(d) Compute $\|H_f(\mathbf{x}_0)\|$.
(e) Let U be the region given by $|x - x_0| < 1$ and $|y - y_0| < 1$. Find a value M so that

$$|f(\mathbf{x}) - (f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0))| \leq (M/2)|\mathbf{x} - \mathbf{x}_0|^2$$

for all $\mathbf{x}$ in U.

(f) Find all of the critical points of f, and determine whether each is a local minimum, maximum, or saddle point.

2 Let $f(x, y) = \sin(\pi(x^2 - xy))$, and let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
(a) Compute $H_f(\mathbf{x})$ and $H_f(\mathbf{x}_0)$.
(b) Compute the minimum and maximum curvatures of f at $\mathbf{x}_0$.
(c) Check that $\mathbf{x}_0$ is a critical point of f. Is it a local minimum, maximum, or saddle point?
(d) Compute $\|H_f(\mathbf{x}_0)\|$.
(e) Let U be the region given by $|x - x_0| < 1$ and $|y - y_0| < 1$. Find a value M so that

$$|f(\mathbf{x}) - (f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0))| \leq (M/2)|\mathbf{x} - \mathbf{x}_0|^2$$

for all $\mathbf{x}$ in U.

3. Let $f(x, y) = (x^2 + y^2)^2 - 2xy$, and let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
(a) Compute $H_f(\mathbf{x})$ and $H_f(\mathbf{x}_0)$.
(b) Compute the minimum and maximum curvatures of f at $\mathbf{x}_0$.
(c) Check that $\mathbf{x}_0$ is a critical point of f. Is it a local minimum, maximum, or saddle point?
(d) Compute $\|H_f(\mathbf{x}_0)\|$.
(e) Let U be the region given by $|x - x_0| < 1$ and $|y - y_0| < 1$. Find a value M so that

$$|f(\mathbf{x}) - (f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0))| \leq (M/2)|\mathbf{x} - \mathbf{x}_0|^2$$

for all $\mathbf{x}$ in U.

4. Let $f(x, y) = x^3y - 2xy^3 + y^4$, and let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.
(a) Compute $H_f(\mathbf{x})$ and $H_f(\mathbf{x}_0)$.
(b) Compute the minimum and maximum curvatures of f at $\mathbf{x}_0$.
(c) Check that $\mathbf{x}_0$ is a critical point of f. Is it a local minimum, maximum, or saddle point?
(d) Compute $\|H_f(\mathbf{x}_0)\|$.
(e) Let U be the region given by $|x - x_0| < 1$ and $|y - y_0| < 1$. Find a value M so that

$$|f(\mathbf{x}) - (f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0))| \leq (M/2)|\mathbf{x} - \mathbf{x}_0|^2$$

for all $\mathbf{x}$ in U.

Section 7: Quadratic functions and quadratic surfaces

7.1 Some basic examples

Quadratic surfaces are just the surfaces we get when graphing a quadratic function. These are the surfaces that arise when we make the tangent quadratic surface approximation.

It is easy to draw contour plots of quadratic quadratic functions, and thus we can easily visualize tangent quadratic surfaces. This is because if

$$h(x,y) = C + Ax + By + D\frac{x^2}{2} + Exy + F\frac{y^2}{2} \ , \tag{7.1}$$

and we want to graph the contour curve implicitly defined by

$$h(x,y) = c$$

for some given value of c, we can always solve for y as an explicit function of x, or x as an explicit function of y. We can do this, because with x fixed, say,

$$C + Ax + By + D\frac{x^2}{2} + Exy + F\frac{y^2}{2} = c$$

is just

$$F\frac{y^2}{2} + (B + Ex)y = c - \left(C + Ax + D\frac{x^2}{2}\right) \ .$$

This is a quadratic equation in y with coefficients depending on x. You can use the quadratic formula to find the roots, which in general give two branches of the contour curve.

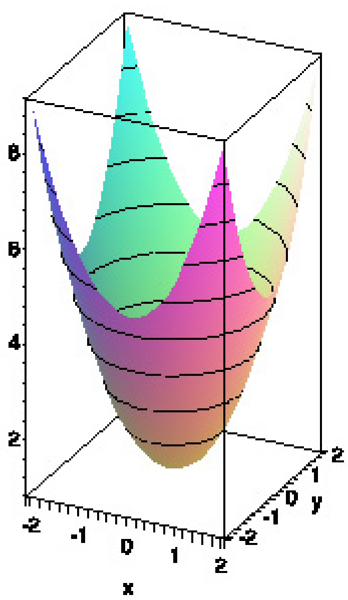
It turns out that the contour curves arising this way are all just conic sections: ellipses, hyperbolas and parabolas. And only one such type arises for any given function. Therefore, although quadratic surfaces are more complicated that tangent planes, they are still pretty simple, and there are just three types: elliptic, hyperbolic and parabolic, as we shall explain.

If one went on to cubic functions, one would encounter much greater algebraic difficulties – you could not longer use the quadratic formula – and there would be many more kinds of "tangent cubic surfaces" to consider. This is one reason why we generally stop at the tangent quadratic surface approximation: An approximation is useful only in so far as it allows one to replace something complicated by something simple. Quadratic surfaces are simple, and cubic surfaces are not.

Example 1 (A cup shaped quadratic surface) Consider the quadratic function

$$h(x,y) = 1 + x^2 + y^2 \ .$$

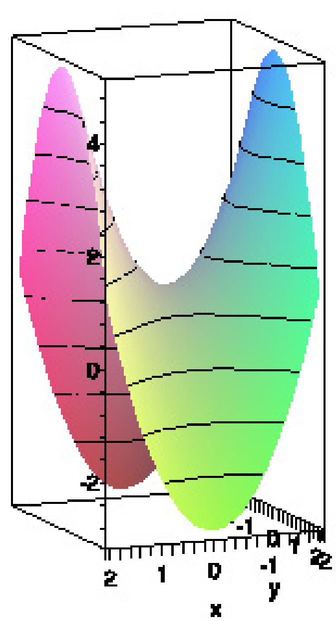
Clearly, for all x and y, $h(x,y) \geq h(0,0) = 1$, and there is equality if and only if $x = y = 0$. Hence the graph of this function must "curve up". Here is the graph:



Example 2 (A saddle shaped quadratic surface) Consider the quadratic function

$$h(x,y) = 1 + x^2 - y^2 \ .$$

This quadratic function differs from the one in Example 1 by the negative sign before y^2. If you move away from the origin along the y axis, you "curve downward", while if you move away from the origin along the x axis, you curve upward. This gives the corresponding quadric surface a "saddle shape". Here is a graph:



The quadratic functions in Examples 1 and 2 had critical points at the origin $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and hence their tangent planes would be flat there. Both would have the very simple linear approximation

$$h(x, y) \approx 1$$

at the origin. This is simple, but not very informative. The two surfaces are quite different, despite having the same linear approximation. One curves upwards in a cup shape, making the origin a minimum, while the other is saddle shaped.

7.2 Contour plots of quadratic functions

While quadratic surfaces are more complicated than planes, they are still pretty simple. First of all, they typically have exactly one critical point. Indeed, with $h(x, y)$ given by (7.1),

$$\nabla h(x, y) = \begin{bmatrix} A + Dx + Ey \\ B + Ex + Fy \end{bmatrix} .$$

That is,

$$\nabla h(\mathbf{x}) = \begin{bmatrix} A \\ B \end{bmatrix} + \begin{bmatrix} D & E \\ E & F \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} . \tag{7.2}$$

Note that the Hessian of h is constant, with

$$H_h(x, y) = \begin{bmatrix} D & E \\ E & F \end{bmatrix} \tag{7.3}$$

for all x and y. Notice also that

$$\begin{bmatrix} A \\ B \end{bmatrix} = \nabla h(0, 0) . \tag{7.4}$$

Definition (Degenerate and non degenerate quadratic functions) A quadratic function h is *non degenerate* in case its Hessian is invertible, which is the case exactly when $\det(H_h) \neq 0$. Otherwise, it is called *degenerate*.

Non degenerate quadratic functions always have exactly one critical point. From (7.2), (7.3) and (7.4), we see that (x, y) is a critical point if and only if

$$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix} &= -\begin{bmatrix} D & E \\ E & F \end{bmatrix}^{-1} \begin{bmatrix} A \\ B \end{bmatrix} \\ &= -[H_h]^{-1} \nabla h(0, 0) . \end{aligned} \tag{7.5}$$

Example 4 (Finding the critical point) Consider again the quadratic function from Example 3:

$$h(x, y) = 1 + x - 2y + x^2 + 4xy + y^2 .$$

In this case we have

$$\nabla h(0) = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \qquad \text{and} \qquad H_h = \begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix} \ .$$

Hence $[H_h]^{-1} = -\frac{1}{12}\begin{bmatrix} 2 & -4 \\ -4 & 2 \end{bmatrix}$, and the critical point $\mathbf{x}_0$ is given by

$$\mathbf{x}_0 = \frac{1}{6}\begin{bmatrix} -5 \\ 4 \end{bmatrix} \ .$$

Now that we have found the critical point, what do we do with it?

• *It is very simple to graph the contour curves of a non degenerate quadratic function if you use a new coordinate system whose origin is at the critical point* $\mathbf{x}_0$*, and whose axes point along the eigenvector of* H_f*.*

In fact, if μ_1 and μ_2 are the two eigenvalues of H_f, and if $\mathbf{u}_1$ and $\mathbf{u}_2$ are the corresponding eigenvectors, consider a new Cartesian coordinate system with coordinates $\tilde{x}$ and $\tilde{y}$ centered at $\mathbf{x}_0$, and with the $\tilde{x}$ axis running along the $\mathbf{u}_1$ direction, and the $\tilde{y}$ axis running along the $\mathbf{u}_2$ direction. Then the contour curves of h are just the contour curves of

$$\mu_1\tilde{x}^2 + \mu_2\tilde{y}^2 \ . \tag{7.6}$$

Now, if both μ_1 and μ_2 have the same sign, then

$$\mu_1\tilde{x}^2 + \mu_2\tilde{y}^2 = c$$

is the equation of an ellipse, provided c has the same sign as μ_1 and μ_2. (Otherwise, there are no solutions).

However, if μ_1 and μ_2 have opposite signs, then

$$\mu_1\tilde{x}^2 + \mu_2\tilde{y}^2 = c$$

is the equation of an hyperbola.

Before we go into the justification of these facts, let us go through some examples where we use them to graph contour curves.

Example 5 (Using the Hessian to graph a contour plot: hyperbolic case) Consider once more the quadratic function from Examples 3 and 4:

$$h(x,y) = 1 + x - 2y + x^2 + 4xy + y^2 \ .$$

As we have seen,

$$\nabla h(0,0) = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \qquad \text{and} \qquad H_f = \begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix} \ .$$

Since $\det(H_h) = -12 \neq 0$, h is non degenerate. Next, from (7.5), the unique critical point is

$$\mathbf{x}_0 = \frac{1}{6}\begin{bmatrix} -5 \\ 4 \end{bmatrix} \ .$$

As you can compute, the eigenvalues are $\mu_1 = 6$ and $\mu_2 = -2$. The egienvectors corresponding to μ_1 and μ_2 are $\mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Now let's use this information to draw a contour plot of h. We do this in two steps. According to what we have claimed above, there is a system of coordinates $\tilde{x}, \tilde{y}$ in which these contour curves are given by

$$6\tilde{x}^2 - 2\tilde{y}^2 = c\ . \tag{7.7}$$

This is the equation of a hyperbola in the $\tilde{x}, \tilde{y}$ plane. The asymptotes of the hyperbola are the two lines we get for the solution of (7.7) with $c = 0$; i.e., $6\tilde{x}^2 - 2\tilde{y}^2 = 0$.

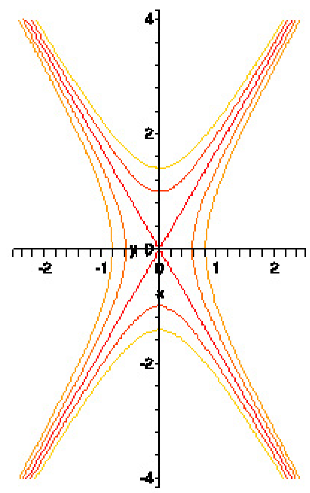
We can divide through by 2, and, since the left hand side is a difference of squares, we can factor it:

$$(\sqrt{3}\tilde{x} - \tilde{y})(\sqrt{3}\tilde{x} + \tilde{y}) = 0\ .$$

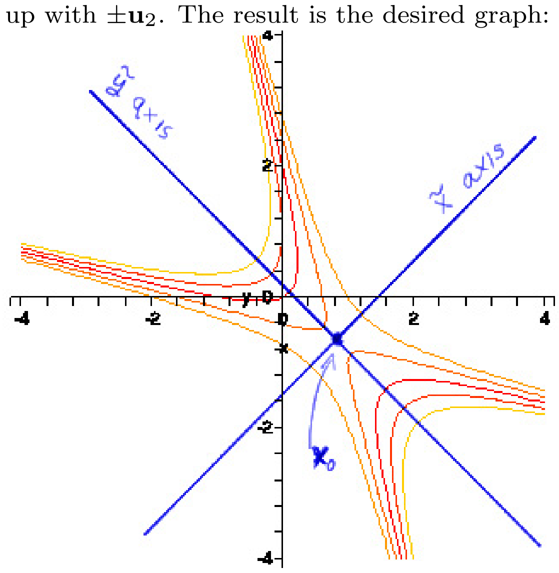
The asymptotes are the lines

$$\tilde{y} = \sqrt{3}\tilde{x} \qquad \text{and} \qquad \tilde{y} = -\sqrt{3}\tilde{x}\ .$$

To get your sketch in the $\tilde{x}, \tilde{y}$ plane, draw in these lines, and then draw in a few hyperbolas that are asymptotic to them. The result is:



To get the graph in the x, y plane, we just use the fact that our two coordinate systems are related by a translation and a rotation. Locate the critical point $\mathbf{x}_0$, and translate the graph so that the asymptotes cross at $\mathbf{x}_0$. Then rotate the graph until the $\tilde{x}$ axis is lines up with $\mathbf{u}_1$. Since $\mathbf{u}_2$ and $\mathbf{u}_1$ are orthogonal, the $\tilde{y}$ axes will also be lined up with $\pm\mathbf{u}_2$. The result is the desired graph:



In practical terms, the best procedure is probably to mark the critical points $\mathbf{x}_0$. Then, using the direction of $\mathbf{u}_1$, draw in the $\tilde{x}$ axis through $\mathbf{x}_0$. The $\tilde{y}$ axis is perpendicular to this, so you can now easily draw it in. Next, draw in the asymptotes relative to the $\tilde{x}, \tilde{y}$ axes, and finally sketch in a few hyperbolas that are asymptotic to them.

Example 6 (Using the Hessian to graph a contour plot: elliptic case) Consider the quadratic function

$$h(x, y) = 4x^2 - 3xy_{+}8y^2 + 8x - 3y + 2 \ .$$

Then,

$$\nabla h(0,0) = \begin{bmatrix} 8 \\ -3 \end{bmatrix} \qquad \text{and} \qquad H_f = \begin{bmatrix} 8 & -3 \\ -3 & 16 \end{bmatrix} \ .$$

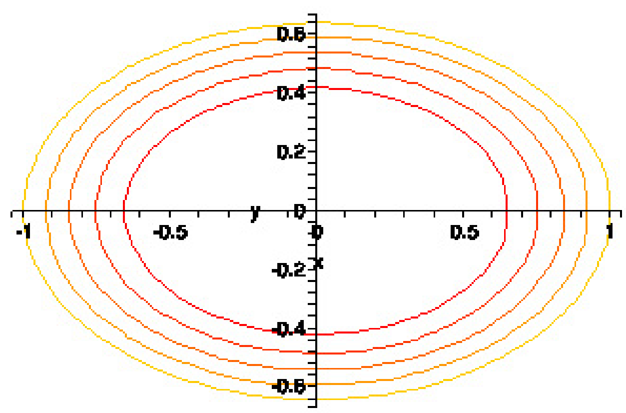
Since $\det(H_h) = 119 \neq 0$, h is non degenerate. From (7.5), the unique critical point is $\mathbf{x}_0 = -\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

As you can compute, the eigenvalues are $\mu_1 = 7$ and $\mu_2 = 17$. The egienvectors corresponding to μ_1 and μ_2 are $\mathbf{u}_1 = \dfrac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\mathbf{u}_2 = \dfrac{1}{\sqrt{10}} \begin{bmatrix} -1 \\ 3 \end{bmatrix}$.

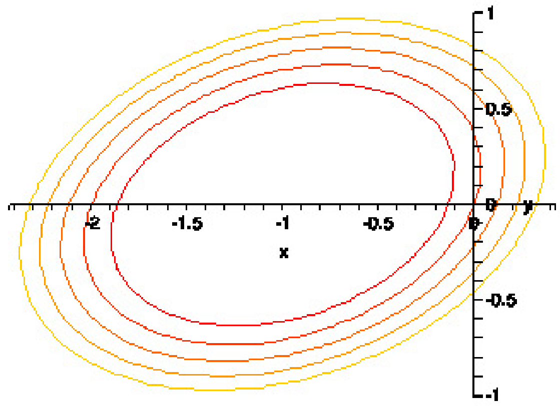
Now let's use this information to draw a contour plot of h. According to what we have claimed above, there is a system of coordinates $\tilde{x}, \tilde{y}$ in which these contour curves are given by

$$7\tilde{x}^2 + 17\tilde{y}^2 = c \ .$$

For $c > 0$, this is the equation of a centered ellipse in the $\tilde{x}, \tilde{y}$ plane. For $c = 1$, the major axis has the length $1/\sqrt{7}$, and runs along the $\tilde{x}$ axis, while the minor axis has the length $1/\sqrt{17}$, and runs along the $\tilde{y}$ axis. Here is the graph in the $\tilde{x}, \tilde{y}$ plane, for 5 values of c:



To get the graph in the x, y plane, we just shift the graph so that the center of the ellipse is at the critical point, and turn it so that the the $\tilde{x}$ axis is lines up with $\mathbf{u}_1$, and the $\tilde{y}$ axes being lined up with $\pm\mathbf{u}_2$. The result is the desired graph, again for 5 values of c:



Example 7 (The contour plot of a degenerate quadratic function) Consider the quadratic function

$$h(x, y) = x^2 - 2xy_+y^2 + x - y + 2 \ .$$

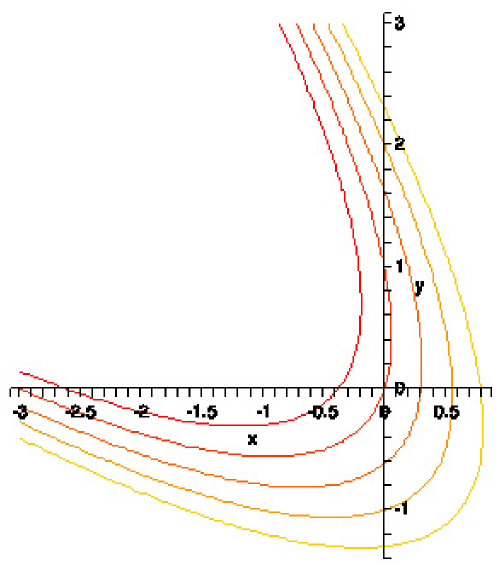
Then,

$$\nabla h(0, 0) = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \qquad \text{and} \qquad H_f = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \ .$$

Since $\det(H_h) = 0$, h is degenerate. The (7.5) has no solution, since clearly $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ cannot be written as a linear combination of the columns of H_h. Therefore, there is no critical point.

The graphing method that we have used on the non degenerate case breaks down here. In the next subsection, we will explain why it works in the non degenerate case, and from that analysis, we will see how to deal with the degenerate case.

The resulting contour plot consists of a parallel family of parabolas. Here it is:



Though we cannot use the simple method for the non degenerate case here, we will see that the tilt of the parabolas is still determined by the direction of the eigenvectors, and that the curvature of the parabolas is determined by the non zero eigenvalue of H_h.

You can see from Examples 5 and 6 that for any non degenerate quadratic function h, the contour curves of h are either all hyperbolas, or all ellipses, depending on whether the eigenvalues have different signs, or the same sign.

Definition (elliptic and hyperbolic quadratic surfaces) A non degenerate quadratic surface is *elliptic* if all of the eigenvalues of its Hessian matrix have the same sign, and it is *hyperbolic* if its Hessian has both positive and negative eigenvalues.

Moreover, the degenerate case is fairly rare: It only arises when $\det(H_h) = 0$. One can make an arbitrarily small change in the coefficients of h, with the result that $\det(H_h)$ becomes either positive or negative.

If you think back to the theory of conic sections, you can see why it is that in the degenerate case, the contour curves will be parabolas: Recall that one gets parabolas by slicing a cone vertically. If you tilt the plane a little bit, your slice now results in either an ellipse or a branch of a hyperbola depending on which way you tilt. Parabolas are the "borderline" between ellipses and hyperbolas, just as Hessians with zero determinant are the "borderline" between Hessians with negative determinant (eigenvalues of different sign) and positive determinant (eigenvalues of the same sign).

7.3 How to choose coordinates to simplify quadratic functions

In this subsection, we explain how to make the change of coordinates $(x, y) \to (\tilde{x}, \tilde{y})$ that simplifies the contour curve equation of h to something of the form

$$\mu_1 \tilde{x}^2 + \mu_2 \tilde{y}^2 = c$$

in the non degenerate case. This will completely justify the procedure that we have used in Examples 5 and 6. We shall also see how to deal with the degenerate case, and why the contour curves are always parabolas here.

Let h be q quadratic function of the form (7.1). The best quadratic approximation of h at any point $\mathbf{x}_0$ is given by

$$h(\mathbf{x}_0) + \nabla h(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot [H_h](\mathbf{x} - \mathbf{x}_0) \ . \tag{7.8}$$

But since h is quadratic itself, h is its own best quadratic approximation, so that (7.8) is just another way of writing h.

Now, as we have seen, in the non degenerate case, there is always exactly one critical point. Choosing $\mathbf{x}_0$ to be the critical point, we have $\nabla h(\mathbf{x}_0) = 0$, and so (7.8) simplifies, and we have

$$h(\mathbf{x}) = h(\mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot [H_h](\mathbf{x} - \mathbf{x}_0) \ . \tag{7.9}$$

Now, since the matrix H_h is symmetric, there is an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ of $I\!R^2$ consisting of eigenvectors of H_h, and the matrix $U = [\mathbf{u}_1, \mathbf{u}_2]$ diagonalizes H_h. That is,

$$H_h = U \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} U^t$$

where μ_1 and μ_2 are the eigenvalues of H_h corresponding to $\mathbf{u}_1$ and $\mathbf{u}_2$ respectively. (Notice that since the columns of U are orthonormal, U is an *orthogonal matrix*, and $U^t = U^{-1}$.)

Now, using this diagonalization formula, and the transpose formula for moving a matrix to the other side of a dot product,

$$\begin{aligned}(\mathbf{x}-\mathbf{x}_0)\cdot[H_h](\mathbf{x}-\mathbf{x}_0) &= (\mathbf{x}-\mathbf{x}_0)\cdot\left(U\begin{bmatrix}\mu_1 & 0\\ 0 & \mu_2\end{bmatrix}U^t(\mathbf{x}-\mathbf{x}_0)\right)\\ &= U^t(\mathbf{x}-\mathbf{x}_0)\cdot\left(\begin{bmatrix}\mu_1 & 0\\ 0 & \mu_2\end{bmatrix}U^t(\mathbf{x}-\mathbf{x}_0)\right)\ .\end{aligned} \tag{7.10}$$

We now introduce our new coordinates as follows: Let $\tilde{x}$ and $\tilde{y}$ be given by

$$\begin{bmatrix}\tilde{x}\\ \tilde{y}\end{bmatrix} = U^t\begin{bmatrix}x-x_0\\ y-y_0\end{bmatrix}\ . \tag{7.11}$$

Then the right hand side of (7.10) reduces to

$$\begin{bmatrix}\tilde{x}\\ \tilde{y}\end{bmatrix}\cdot\begin{bmatrix}\mu_1 & 0\\ 0 & \mu_2\end{bmatrix}\begin{bmatrix}\tilde{x}\\ \tilde{y}\end{bmatrix} = \mu_1\tilde{x}^2 + \mu_2\tilde{y}^2\ .$$

In the new coordinates, the formula for h is simply

$$h(\mathbf{x}_0) + \frac{1}{2}(\mu_1\tilde{x}^2 + \mu_2\tilde{y}^2)\ .$$

Notice that $h(\mathbf{x}_0)$ is just a number – the value of h at its critical point. Hence setting h to be constant amounts to setting $\mu_1\tilde{x}^2 + \mu_2\tilde{y}^2$ to be constant, so that, as claimed,

$$\mu_1\tilde{x}^2 + \mu_2\tilde{y}^2 = c$$

gives the equation for the contour curves of h, in the new coordinates. We have proved the following:

Theorem 1 (Nice coordinates for non degenerate quadratic functions) *Let h be any non-degenerate quadratic function on $\mathbb{R}^2$. Let $\mathbf{x}_0$ be its unique critical point, and let μ_1 and μ_2 be the eigenvalues of its Hessian. Then there is an orthogonal matrix U so that if $\tilde{\mathbf{x}} = \begin{bmatrix}\tilde{x}\\ \tilde{y}\end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix}x\\ y\end{bmatrix}$ are related by*

$$\tilde{\mathbf{x}} = U^t(\mathbf{x}-\mathbf{x}_0) \qquad \textit{or equivalently,} \quad \mathbf{x} = \mathbf{x}_0 + U\tilde{\mathbf{x}}\ , \tag{7.12}$$

and if $\tilde{h}(\tilde{\mathbf{x}})$ is defined by $\tilde{h}(\tilde{\mathbf{x}}) = h(\mathbf{x})$, then

$$\tilde{h}(\tilde{x},\tilde{y}) = h(\mathbf{x}_0) + \frac{1}{2}\mu_1\tilde{x}^2 + \mu_2\tilde{y}^2\ .$$

What about the degenerate case? In the degenerate case, at least one of the eigenvalues is zero. If both eigenvalues of a symmetric* 2×2 matrix are zero, then the matrix is the zero matrix.

If the Hessian of H is zero, then (7.1) reduces to

$$h(x,y) = Ax + By + C \ .$$

This is just a linear function, and the contour curves are straight lines – a degenerate sort of parabola.

Hence we may as well assume that exactly one eigenvalue is zero, say $\mu_1 \neq 0$, and $\mu_2 = 0$.

Let $\{\mathbf{u}_1, \mathbf{u}_2\}$ be the corresponding orthonormal basis of eigenvectors. Let $U = [\mathbf{u}_1, \mathbf{u}_2]$, and

$$D = \begin{bmatrix} \mu_1 & 0 \\ 0 & 0 \end{bmatrix} \ .$$

We will take $\mathbf{x}_0 = 0$ as our base point, since in the degenerate case, there may not be any critical point. Inserting $H_h = UDU^t$ into (7.8), with $\mathbf{x}_0 = 0$, we have from (7.1) and (7.2)

$$\begin{aligned} h(\mathbf{x}) &= h(0) + \nabla h(0) \cdot \mathbf{x} + \frac{1}{2}\mathbf{x} \cdot (U[H_h]U^t)\mathbf{x} \\ &= C + \begin{bmatrix} A \\ B \end{bmatrix} \cdot \mathbf{x} \ + \ \frac{1}{2}\mathbf{x} \cdot (U[H_h]U^t)\mathbf{x} \ . \end{aligned} \tag{7.13}$$

Introduce the new coordinates $\tilde{x}$ and $\tilde{y}$ through

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = U^t \begin{bmatrix} x \\ y \end{bmatrix} \ . \tag{7.14}$$

This is just what we did before, except that now we use $\mathbf{x}_0 = 0$. Also, define $\mathbf{p} = \begin{bmatrix} P \\ Q \end{bmatrix}$ by

$$\mathbf{p} = U^t \begin{bmatrix} A \\ B \end{bmatrix} \ .$$

---

* The symmetry is important here. The only eigenvalue of the matrix $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is zero, but it is not the zero matrix. The key is that symmetric matrices can always be diagonalized, and written in the form $U \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} U^t$. If 0 is the only eigenvalue of the matrix, the diagonal matrix in the middle is the zero matrix, and so the whole product is the zero matrix.

Then since $\mathbf{x} = \tilde{\mathbf{x}}$, (7.14) becomes

$$\begin{aligned} h(\mathbf{x}) &= C + \mathbf{p} \cdot \tilde{\mathbf{x}} + \frac{1}{2}\tilde{\mathbf{x}} \cdot \begin{bmatrix} \mu_1 & 0 \\ 0 & 0 \end{bmatrix} \tilde{\mathbf{x}} \\ &= C + P\tilde{x} + Q\tilde{y} + \frac{\mu_1}{2}\tilde{x}^2 \ . \end{aligned} \tag{7.15}$$

If $Q \neq 0$, then we can solve $h(\mathbf{x}) = c$ for $\tilde{y}$, finding

$$\tilde{y} = \left(\frac{c - C}{Q}\right) - \left(\frac{P}{Q}\right)\tilde{x} - \left(\frac{\mu_1}{2Q}\right)\tilde{x}^2 \ ,$$

and the graph of this equation is clearly a parabola. If $Q = 0$, then $\tilde{y}$ is free, and we get 0, 1 or 2 lines parallel to the $\tilde{x}$ axis, depending on the number of solutions of the quadratic equation $c = C + P\tilde{x} + (\mu_1/2)\tilde{x}^2$. Again, lines are degenerate cases of parabolas.

We have now explained why the level curves in the degenerate case are parabolas, but we shall not dwell on drawing the contour plots in this case: As we have explained, it is a "marginal" and "borderline" case.

7.4 Contour plots of more general functions

We can combine what we have learned so far in this section with what we have learned about "best quadratic approximation" to sketch informative contour plots of functions that are not quadratic.

You may doubt the utility of this, since good graphing programs are readily available. However, the point of this is not so much the resulting contour plots *per se*. Rather, it is that to produce the plot, we just do a few computations with gradients and Hessians, and then we know the "lay of the land" as far as the graph of $z\phi(\mathbf{x})$ is concerned.

Therefore, if you can draw a contour plot just using information coming from the computation of a few gradients and Hessians, you certainly know how to extract an understanding of the behavior of f from the gradients and Hessians. *In other words, you will then understand what the gradient and Hessian are saying about the behavior of the function f.* It turns out that this understanding will readily apply with more variables, where we cannot really draw nice graphs.

That said, let us begin with an example. Consider the function

$$f(x, y) = x^4 + y^4 + 4xy \tag{7.16}$$

that we discussed in Example 8 from Section 3. Note that

$$\nabla f(x, y) = 4 \begin{bmatrix} x^3 + y \\ y^3 + x \end{bmatrix} \ .$$

Hence (x, y) is a critical point if and only if

$$\begin{aligned} x^3 + y &= 0 \\ y^3 + x &= 0 \ . \end{aligned}$$

This is a non linear system of equations, but it is very easy to solve since we can very easily eliminate variables* The first equation says $y = -x^3$, and using this to eliminate y from the second, we get

$$x - x9 = 0 \ .$$

One solution is $x = 0$. If $x \neq 0$, we divide through by x and get $1 - x^8 = 0$, so $x = \pm 1$ are the other two solutions. Now that we know the three possible x values, the equation $y = -x^3$ gives us the y values, and we find the three critical points:

$$(0,0)qquad(1,-1) \quad \text{and} \quad (-1,1) \ .$$

Next, we compute the Hessian at each of the critical points: $H_f(x,y) = \begin{bmatrix} 12x^2 & 4 \\ 4 & 12y^2 \end{bmatrix}$, so that

$$H_f(0,0) = \begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix} \qquad H_f(1,-1) = \begin{bmatrix} 12 & 4 \\ 4 & 12 \end{bmatrix} \quad \text{and} \quad H_f(-1,1) = \begin{bmatrix} 12 & 4 \\ 4 & 12 \end{bmatrix} \ .$$

We now sketch a contour plot of the tangent quadratic at each of these points. Since the tangent quadratic is a good fit, this will be a pretty close contour plot of f near each of these points. Then we can "connect the curves" and get a contour plot of f in the vicinity of the critical points.

Here is how this goes: First, at $(0,0)$, the eigenvalues of the Hessian, $\begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix}$ are given by $\mu_1 = -4$ and $\mu_2 = 4$. To get an eigenvector $\mathbf{u}_1$ corresponding to μ_1, form

$$H_f(0,0) - 4I = \begin{bmatrix} -4 & 4 \\ 4 & -4 \end{bmatrix} \ .$$

The vector $\mathbf{u}_1$ must be orthogonal to the rows of this matrix. So "perping" the first row, we get $\begin{bmatrix} -4 \\ -4 \end{bmatrix}$. We can drop the factor of -4, and take $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. (We could normalize it with a factor of $1/sqrt2$, but for our present purposes, we only need the direction it provides.)

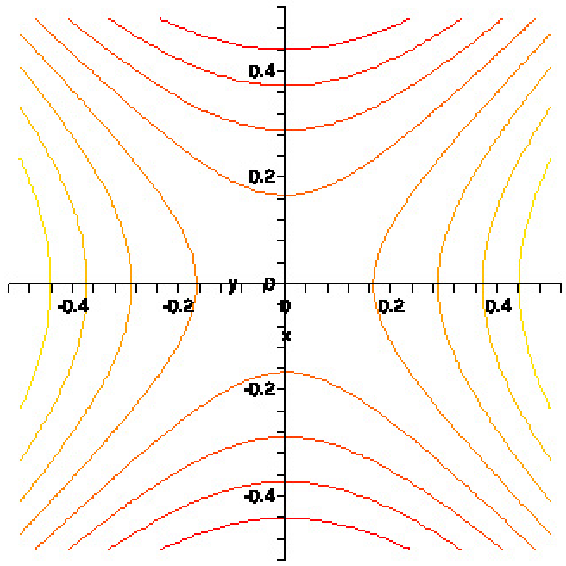
Then, since the Hessian is symmetric, $\mathbf{u}_2 = \mathbf{u}_1^\perp$ is an eigenvector with eigenvalue -4.

In the "nice coordinates" provided by Theorem 1, the equation for the contour curves of the tangent quadratic is

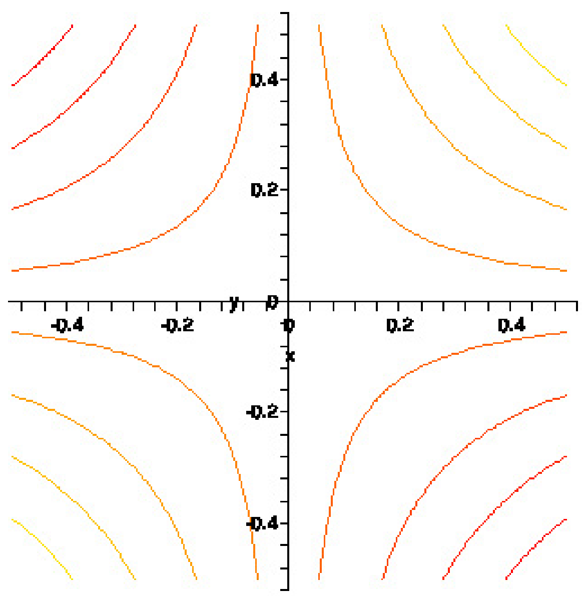
$$4\tilde{x}^2 - 4\tilde{y}^2 = c \ .$$

This is describes a family of hyperbolas. You can easily graph them, and find

* Remember, this is what solving equations algebraically is all about, and is why the main method for solving systems of linear equations is known as Gaussian elimination, as well as row reduction.



Tilting the hyperbolas so that the $\tilde{x}$ axis lines up with $\mathbf{u}_1$, we see:



(In the second graph, the axes are the x and y axes, while in the first, they are the $\tilde{x}$ and $\tilde{y}$ axes.)

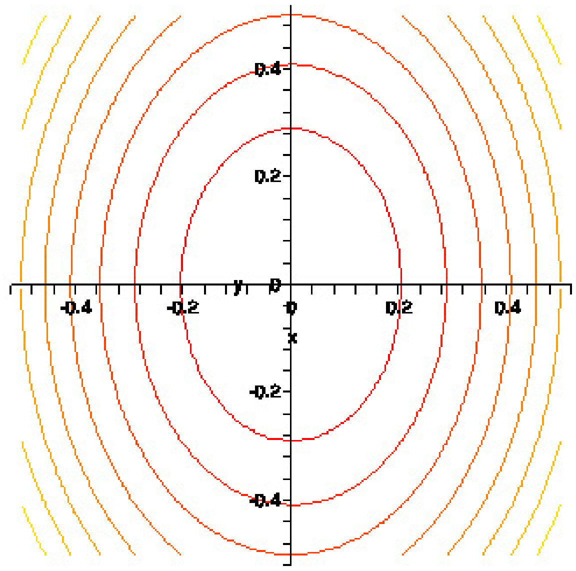
The second graph is a good representation of the contour plot of f itself in the vicinity of $(0,0)$ since the tangent quadratic approximation is very good there.

Next, consider the vicinity of the critical point $(1,-1)$. As you can easily compute, the Hessian here has the eigenvalues 16 and 8, and again the corresponding eigenvectors are $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{u}_2 = \mathbf{u}_1^\perp$.

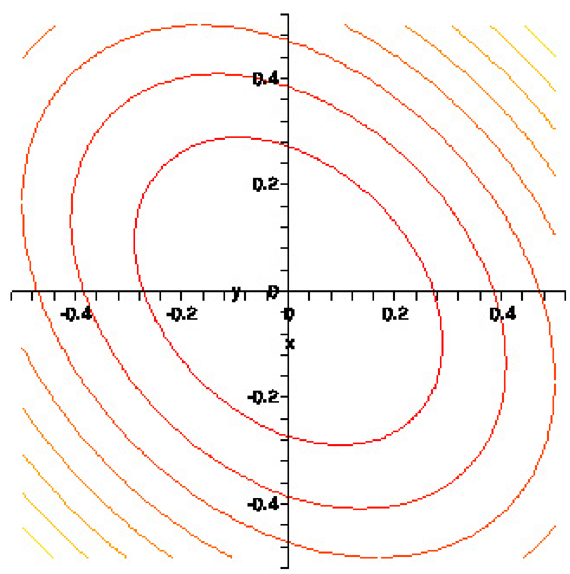
In the "nice coordinates" provided by Theorem 1, the equation for the contour curves of the tangent quadratic is

$$16\tilde{x}^2 + 8\tilde{y}^2 = c\ .$$

This describes a family of ellipses. You can easily graph them and find:

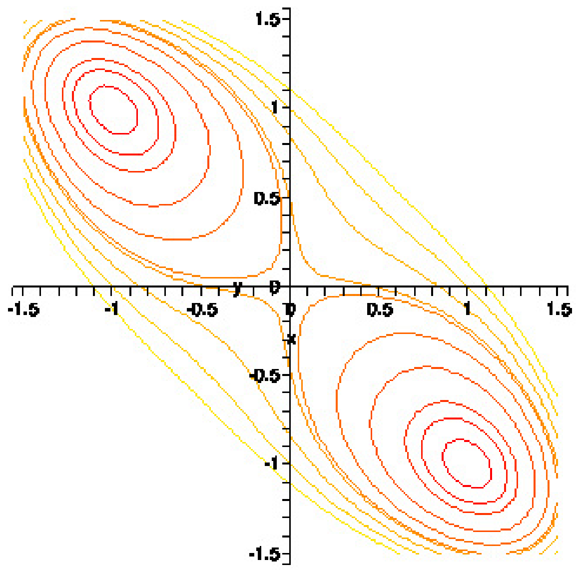


Tillting the hyperbolas so that the $\tilde{x}$ axis lines up with $\mathbf{u}_1$, we see:



(The axes in the graph are the lines $x = 1$ and $y = -1$, through the critical point, not the x and y axes.) Of course, the Hessian at the other critical point is the same, so the graphs there look exactly the same.

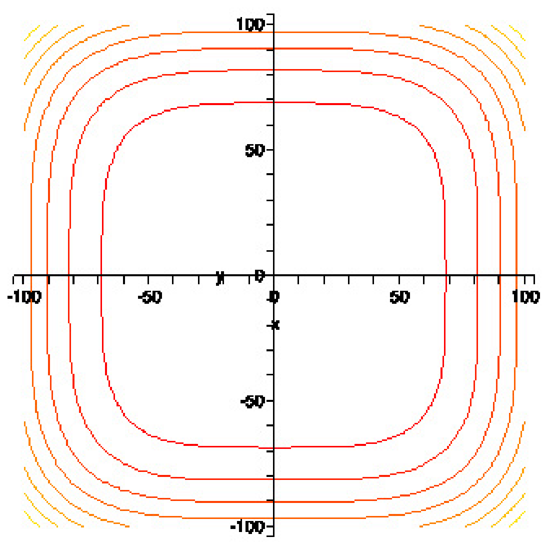
Now combine these graphs, drawing them in nearby each of the critical points, and connecting up the curves. The result is a pretty close contour plot of f in the vicinity of its critical points:



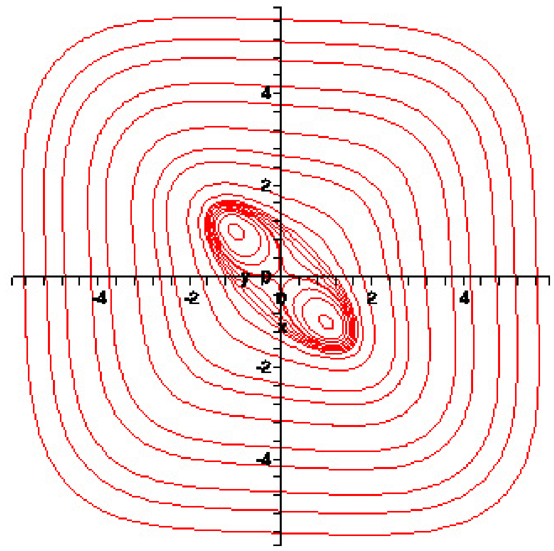
Next: What does the graph look like farther away form the critical points? Quite different, actually. If $|\mathbf{x}|$ is large, then $4xy$ is negligible compared to $x^4 + y^4$, and we can get a good sketch of the contour curves through such points by graphing

$$x^4 + y^4 = c\ .$$

This gives us "distorted circles" that "bulge out" toward the corners:



We now get a pretty good "global" contour plot by interpolating between this, and the plot we found for the vicinity of the critical points:



Notice that the contour curves of the *quartic* function f are not all simple rescalings or translations of one another, as in the quadratic case. They can be quite different, with some having two components and others just having one. This already happens with cubic functions, as you will see in the exercises. Quadratic functions really are quite simple, and this is what makes the tangent quadratic approximation so useful.

Exercises

6.1 Consider the quadratic function

$$h(x, y) = 2x^2 + 6xy + 10y^2 - 5x - 13y + 2 \ .$$

Determine whether h is elliptic, hyperbolic or degenerate. If it is non degenerate, find the critical point, and sketch a contour plot of h.

6.2 Consider the quadratic function

$$h(x, y) = 5x^2 - 4xy + 8y^2 + 2x + y - 3 \ .$$

Determine whether h is elliptic, hyperbolic or degenerate. If it is non degenerate, find the critical point, and sketch a contour plot of h.

6.3 Consider the quadratic function

$$h(x, y) = 6x^2 - 6xy - 2y^2 - 2x + 4y + 1 \ .$$

Determine whether h is elliptic, hyperbolic or degenerate. If it is non degenerate, find the critical point, and sketch a contour plot of h.

6.4 Consider the quadratic function

$$h(x, y) = x^2 - 2xy + y^2 - 2x + 4y + 1 \ .$$

Determine whether h is elliptic, hyperbolic or degenerate. If it is non degenerate, find the critical point, and sketch a contour plot of h.

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

a There are two critical points for f. Find them both.

b Compute the best quadratic approximation $h(x, y)$ to $f(x, y)$ at each critical point.

c Use the method explained in this section to draw a contour plot of f in the vicinity of its critical points. (Note: One of the Hessians has messy eigenvalues. Just work then out in decimal form, which is fine for purposes of drawing a sketch.)

6.7 Let $f(x, y) = x^3y + y^2x$.

a Find all of the critical points for f.

b Compute the best quadratic approximation $h(x, y)$ to $f(x, y)$ at each critical point.

c Use the method explained in this section to draw a contour plot of f in the vicinity of its critical points. (Note: When you find a Hessians with messy eigenvalues, just work then out in decimal form, which is fine for purposes of drawing a sketch.)

Section 8: The Jacobian of a transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$.

8.1 Jacobian matrices

Let $\mathbf{F}$ be a transformation, or in other words, a function, form $\mathbb{R}^2$ to $\mathbb{R}^2$. This is a function that takes vectors in $\mathbb{R}^2$ as input, and returns vectors in $\mathbb{R}^2$ as output. We are already familiar with many examples.

Example 1 (The gradient as a transformation) Let $f(x, y)$ be a function on $\mathbb{R}^2$ with values in $\mathbb{R}$. Then its gradient, $\nabla f(x, y)$ is a function on $\mathbb{R}^2$ with values in $\mathbb{R}^2$.

Example 2 (Linear transformations) Let A be a 2×2 matrix, and define $\mathbf{F}(\mathbf{x})$ by $\mathbf{F}(\mathbf{x}) = A\mathbf{x}$. This is a function on $\mathbb{R}^2$ with values in $\mathbb{R}^2$. More generally, we can add a constant vector to $A\mathbf{x}$. With a slight abuse of notation, we shall also call a transformation $\mathbf{F}$ of the from

$$\mathbf{F}(\mathbf{x}) = \mathbf{y}_0 + A\mathbf{x} \tag{8.1}$$

a linear transformation. Here, $\mathbf{y}_0$ is any fixed vector in $\mathbb{R}^2$.

Example 3 (General vector valued functions) Let $f(x, y)$ and $g(x, y)$ be any two functions on $\mathbb{R}^2$ with values in $\mathbb{R}$. Then we can define $F(\mathbf{x})$ by

$$\mathbf{F}(x, y) = \begin{bmatrix} f(x, y) \\ g(x, y) \end{bmatrix} .$$

This is a function on $\mathbb{R}^2$ with values in $\mathbb{R}^2$, and it should be clear that every such function can be written this way for some $\mathbf{F}$ and g.

In this section, we will see how to approximate a general non linear transformation $\mathbf{F}$, as in Example 3, by a linear transformation, as in (8.1), at least when the entries f and g are differentiable. As we shall see, this approximation scheme is very useful for solving non linear systems of equations.

Here is the idea:

- *To approximate* $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{bmatrix}$ *by a linear transformation of the type (8.1), use the tangent plane approximation for each component of* $\mathbf{F}$*; i.e., use the tangent plane approximation to replace* f *and* g *in* $F = \begin{bmatrix} f \\ g \end{bmatrix}$ *by their best linear approximations.*

Here is how this goes: Pick a base point $\mathbf{x}_0$. Then for $\mathbf{x}$ close to $\mathbf{x}_0$ we have both

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$

and

$$g(\mathbf{x}) \approx g(\mathbf{x}_0) + \nabla g(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) .$$

Hence with $F(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{bmatrix}$,

$$\begin{aligned} F(\mathbf{x}) &\approx \begin{bmatrix} f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \\ g(\mathbf{x}_0) + \nabla g(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \end{bmatrix} \\ &= \begin{bmatrix} f(\mathbf{x}_0) \\ g(\mathbf{x}_0) \end{bmatrix} + \begin{bmatrix} \nabla f(\mathbf{x}_0) \\ \nabla g(\mathbf{x}_0) \end{bmatrix} (\mathbf{x} - \mathbf{x}_0) \\ &= F(\mathbf{x}_0) + J_F(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \end{aligned} \tag{8.2}$$

where $J_F(\mathbf{x}_0)$ is the 2×2 matrix whose first row is $\nabla f(\mathbf{x}_0)$, and whose second row is $\nabla g(\mathbf{x}_0)$. This is a very useful matrix, and so it has a name: It is called the *Jacobian matrix* of the transformation $\mathbf{F}$ at $\mathbf{x}_0$.

Definition (Jacobian Matrix) Let $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{bmatrix}$ be a transformation from $I\!R^2$ to $I\!R^2$ such that both f and g are differentiable in some open set U in $I\!R^2$. Then for each $\mathbf{x}$ in U, the *Jacobian matrix* of $\mathbf{F}$ at $\mathbf{x}$, $J_{\mathbf{F}}(\mathbf{x})$, is given by

$$J_{\mathbf{F}}(\mathbf{x}) = \begin{bmatrix} \nabla f(\mathbf{x}) \\ \nabla g(\mathbf{x}) \end{bmatrix} .$$

Example 4 (Computing a Jacobian matrix) This is as easy as computing two gradients. Suppose $F(x,y)$ is given by

$$F(\mathbf{x}) = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = \begin{bmatrix} x^2 + 2yx \\ xy \end{bmatrix} .$$

Then $\nabla f(x,y) = \begin{bmatrix} 2x + 2y \\ 2x \end{bmatrix}$ and $\nabla g(x,y) = \begin{bmatrix} y \\ x \end{bmatrix}$. Hence

$$J_F(x,y) = \begin{bmatrix} 2x + 2y & 2x \\ y & x \end{bmatrix} .$$

Example 5 (Computing a linear approximation) Let $F(x,y)$ be given by

$$F(\mathbf{x}) = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = \begin{bmatrix} x^2 + 2yx \\ xy \end{bmatrix} ,$$

as in the last example. Let's compute the linear approximation to $F(x,y)$ at the point $(x_0, y_0) = (2,1)$. We have $F(2,1) = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$ and, from the previous example,

$$J_F(2,1) = \begin{bmatrix} 6 & 4 \\ 1 & 2 \end{bmatrix} .$$

Hence the linear approximation is

$$\begin{aligned} F(x,y) &\approx \begin{bmatrix} 8 \\ 2 \end{bmatrix} + \begin{bmatrix} 6 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix} \\ &= \begin{bmatrix} 24 \\ 6 \end{bmatrix} + \begin{bmatrix} 6x + 4y \\ x + 2y \end{bmatrix} \\ &= \begin{bmatrix} 24 + 6x + 4y \\ 6 + x + 2y \end{bmatrix} . \end{aligned}$$

Example 6 (The Jacobian matrix of a linear transformation) Suppose that $F(\mathbf{x}) = A\mathbf{x}$ where A is the 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then

$$F(x, y) = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} \ .$$

In this case, you easily compute that for all (x, y),

$$J_F(x, y) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = A \ .$$

This is important: The Jacobean matrix of the linear transformation $\mathbf{F}$ induced by a 2×2 matrix A is just A itself. Such a linear transformation is therefore its own linear approximation, as it should be.

Example 7 (The Jacobian of ∇f is H_f) Let f be a nice function defined on $I\!R^2$. Now ∇f itself is a transformation from $I\!R^2$ to $I\!R^2$. Hence, we can compute its Jacobian, which will be an 2×2 matrix. Consulting the definitions, it is not hard to see that the Jacobian of ∇f is the Hessian of f.

The observation made in the last example deserves emphasis:

- *If f be a function on $I\!R^n$ with values in $I\!R$. Then*

$$J_f = (\nabla f)^t \qquad \text{and} \qquad J_{\nabla f} = H_f \ .$$

8.2 Accuracy of the linear approximation

Because the approximation (8.2) is based on the tangent plane approximation, whose accuracy we have already examined, we can easily estimate its accuracy.

Indeed, let $h(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$ and $k(\mathbf{x}) = g(\mathbf{x}_0) + \nabla g(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$ so that

$$\mathbf{F}(\mathbf{x}_0) + J_F(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = \begin{bmatrix} h(\mathbf{x}) \\ k(\mathbf{x}) \end{bmatrix} \ .$$

Then

$$\begin{aligned} |\mathbf{F}(\mathbf{x}) - [\mathbf{F}(\mathbf{x}_0) + J_F(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)]| &= \left| \begin{bmatrix} f(\mathbf{x}) - h(\mathbf{x}) \\ g(\mathbf{x}) - k(\mathbf{x}) \end{bmatrix} \right| \\ &= \sqrt{(f(\mathbf{x}) - h(\mathbf{x}))^2 + (g(\mathbf{x}) - k(\mathbf{x}))^2} \ . \end{aligned}$$

Now suppose that M is a number such that

$$\|H_f(\mathbf{x})\|_{\mathrm{HS}}^2 + \|H_g(\mathbf{x})|_{\mathrm{HS}}^2 \le M^2$$

for all $\mathbf{x}$ in U. Then by Theorem 2 of Section 5,

$$\sqrt{(f(\mathbf{x}) - h(\mathbf{x}))^2 + (g(\mathbf{x}) - k(\mathbf{x}))^2} \le \frac{1}{2} M |\mathbf{x} - \mathbf{x}_0|^2 \ .$$

This proves the following theorem:

Theorem 1 (Accuracy of the linear approximation) *Let* $\mathbf{F} = \begin{bmatrix} f \\ g \end{bmatrix}$ *be defined on an open convex subset* U *of* $\mathbb{R}^2$*, with values in* $\mathbb{R}^2$*. Suppose that all of the partial derivatives of each* f *and* g *of order* 2 *exist and are continuous throughout* U*. Suppose there is a finite number* M *such that*

$$\|H_f(\mathbf{x})\|_{\mathrm{HS}}^2 + \|H_g(\mathbf{x})|_{\mathrm{HS}}^2 \le M^2$$

for all $\mathbf{x}$ *in* U*. Then*

$$|\mathbf{F}(\mathbf{x}) - [\mathbf{F}(\mathbf{x}_0) + J_F(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)]| \le \frac{1}{2}|\mathbf{x} - \mathbf{x}_0|^2 \ .$$

We remark that the hypothesis on the existence of a finite value of M is always satisfies U when the second order partial derivatives are continuous up to and including the boundary of U, since then $\|H_f(\mathbf{x})\|_{\mathrm{HS}}^2 + \|H_g(\mathbf{x})|_{\mathrm{HS}}^2$ is a continuous function of $\mathbf{x}$, and continuous functions always have a finite maximum on closed, bounded sets.

Once again, when $|\mathbf{x} - \mathbf{x}_0|$ is small, $|\mathbf{x} - \mathbf{x}_0|^2$ will be *very small*, so that

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}_0) + J_{\mathbf{F}}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

is a useful approximation.

8.3 Differentiability of vector valued functions

Recall that a function f from $\mathbb{R}^2$ to $\mathbb{R}$ is differentiable at $\mathbf{x}_0$ in case there is a vector $\mathbf{a}$ in $\mathbb{R}^n$ so that

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{|f(\mathbf{x}) - [f(\mathbf{x}_0) + \mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0)]|}{|\mathbf{x} - \mathbf{x}_0|} = 0 \ .$$

In this case, $\mathbf{a} = \nabla f(\mathbf{x}_0)$, and The function $h(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0)$ is the linear approximation to f at $\mathbf{x}_0$. We extend this directly to functions from $\mathbb{R}^2$ to $\mathbb{R}^2$:

Definition (Differentiable functions from $\mathbb{R}^2$ to $\mathbb{R}^2$) A function $\mathbf{F} = \begin{bmatrix} f \\ g \end{bmatrix}$ on $\mathbb{R}^2$ with values in $\mathbb{R}^2$ is *differentiable at* $\mathbf{x}_0$ in case there is an 2×2 matrix A so that

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{|\mathbf{F}(\mathbf{x}) - (\mathbf{F}(\mathbf{x}_0) + A(\mathbf{x} - \mathbf{x}_0))\,|}{|\mathbf{x} - \mathbf{x}_0|} = 0 \ . \qquad (8.3)$$

It is easy to see that there can be at most one matrix A for which (8.3) is true: suppose that both both

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{|\mathbf{F}(\mathbf{x}) - (\mathbf{F}(\mathbf{x}_0) + A(\mathbf{x} - \mathbf{x}_0))\,|}{|\mathbf{x} - \mathbf{x}_0|} = 0$$

and

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{|\mathbf{F}(\mathbf{x}) - (\mathbf{F}(\mathbf{x}_0) + B(\mathbf{x}-\mathbf{x}_0))\,|}{|\mathbf{x}-\mathbf{x}_0|} = 0\ .$$

Then since

$$(A-B)(\mathbf{x}-\mathbf{x}_0) = [\mathbf{F}(\mathbf{x}) - (\mathbf{F}(\mathbf{x}_0) + A(\mathbf{x}-\mathbf{x}_0))] - [\mathbf{F}(\mathbf{x}) - (\mathbf{F}(\mathbf{x}_0) + B(\mathbf{x}-\mathbf{x}_0))]\ ,$$

it follows that

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{|(A-B)(\mathbf{x}-\mathbf{x}_0)|}{|\mathbf{x}-\mathbf{x}_0|} = 0\ .$$

Now let $\mathbf{r}_i$ be the ith row of $A-B$, and define $\mathbf{x} = \mathbf{x}_0 + t\mathbf{r}_i$, Then the ith entry of $(A-B)(\mathbf{x}-\mathbf{x}_0)$ is $|\mathbf{r}_i|^2 t$, and $|\mathbf{x} - bx_0| = |\mathbf{r}_i|t$. Hence for this choice of $\mathbf{x}$,

$$\frac{|(A-B)(\mathbf{x}-\mathbf{x}_0)|}{|\mathbf{x}-\mathbf{x}_0|} \geq \frac{|\mathbf{r}_i|^2 t}{|\mathbf{r}_i t} = |\mathbf{r}_i|\ ,$$

and this can go to zero as t goes to zero (and hence $\mathbf{x} \to 0$) if and only if $\mathbf{r}_i = 0$. Since i is arbitrary, each row of $A-B$ must be zero, and so $A = B$.

Next, if f satisfies the condition of Theorem 1, and if $\mathbf{x}_0$ belongs to U, and if we take $A = J_{\mathbf{F}}(\mathbf{x}_0)$, then (8.3) is true, by Theorem 1. Thus we have:

Theorem 2 (Differentiability of vector valued functions) *Let* $\mathbf{F} = \begin{bmatrix} f \\ g \end{bmatrix}$ *be defined on an open convex subset* U *of* $I\!R^2$*, with values in* $I\!R^2$*. Suppose that all of the partial derivatives of each* f *and* g *of order* 2 *exist and are continuous throughout* U*. Then* $\mathbf{F}$ *is differentiable at each* $\mathbf{x}_0$ *in* U*, and the unique matrix* A *for which (8.3) is true is* $J_{\mathbf{F}}(\mathbf{x}_0)$.

The uniqueness assertion in Theorem 2 justifies the following definition:

Definition (Best linear approximation) Let $\mathbf{F}$ be function from $I\!R^n$ to $I\!R^m$ that is differentiable function at $\mathbf{x}_0$. Then the function $\mathbf{H}$ defined by

$$\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_0) + J_{\mathbf{F}}(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) \tag{8.4}$$

is called *the best linear approximation to* $\mathbf{F}$ *at* $\mathbf{x}_0$.

8.3 The chain rule for vector valued functions

Let $\mathbf{F}$ be a function on $I\!R^2$ with values in $I\!R^2$. Let $\mathbf{G}$ be a function on $I\!R^2$ with values in $I\!R^2$. Then since the range of $\mathbf{F}$ is contained in the domain of $\mathbf{G}$, we can compose these functions to define $\mathbf{H} = \mathbf{G}\circ\mathbf{F}$. That is,

$$\mathbf{H}(\mathbf{x}) = \mathbf{G}(\mathbf{F}(\mathbf{x}))\ .$$

We now claim that if both $\mathbf{G}$ and $\mathbf{F}$ are differentiable, then so is $\mathbf{H}$, and moreover, there is a simple relation between the Jacobian matrices of these three functions:

$$J_{\mathbf{G}\circ\mathbf{F}} = J_{\mathbf{G}} J_{\mathbf{F}} \ .$$

Theorem 3 (The chain rule for transformation from $I\!R^2$ to $I\!R^2$) *Let $\mathbf{F}$ be a function on $I\!R^n$ with values in $I\!R^m$. Let $\mathbf{G}$ be a function on $I\!R^m$ with values in $I\!R^p$. Suppose that $\mathbf{F}$ is differentiable at $\mathbf{x}_0$ and that $\mathbf{G}$ is differentiable at $\mathbf{y}_0 = \mathbf{F}(\mathbf{x}_0)$. Then $\mathbf{F}\circ\mathbf{G}$ is differentiable at $\mathbf{x}_0$ and*

$$J_{\mathbf{G}\circ\mathbf{F}}(\mathbf{x}_0) = J_{\mathbf{G}}(\mathbf{y}_0) J_{\mathbf{F}}(\mathbf{x}_0) \ . \tag{8.5}$$

Proof: We give the proof under slightly stronger assumptions; namely that the components of $\mathbf{F}$ and $\mathbf{G}$ are all twice continuously differentiable in a neighborhood of $\mathbf{x}_0$ and $\mathbf{y}_0$ respectively. This allows us to apply Theorem 1, and to explain the main idea without tangling it up in technicalities. Moreover, most differentiable functions you meet in applications are twice continuously differentiable too. Under these assumptions,

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_0) + J_{\mathbf{F}}(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) + \mathcal{O}(|\mathbf{x}-\mathbf{x}_0|^2) \tag{8.6}$$

and

$$\mathbf{G}(\mathbf{y}) = \mathbf{G}(\mathbf{y}_0) + J_{\mathbf{G}}(\mathbf{y}_0)(\mathbf{y}-\mathbf{y}_0) + \mathcal{O}(|\mathbf{y}-\mathbf{y}_0|^2) \ . \tag{8.7}$$

Now plugging $\mathbf{y} = \mathbf{F}(\mathbf{x})$ and $\mathbf{y}_0 = \mathbf{F}(\mathbf{x}_0)$ into (8.7), we get

$$\mathbf{G}(\mathbf{F}(\mathbf{x})) = \mathbf{G}(\mathbf{F}(\mathbf{x}_0)) + J_{\mathbf{G}}(\mathbf{y}_0)(\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0)) + \mathcal{O}(|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0)|)^2$$

Then using $\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0) = J_{\mathbf{F}}(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) + \mathcal{O}(|\mathbf{x}-\mathbf{x}_0|^2)$ in the first instance of $\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0)$, we get

$$\mathbf{G}(\mathbf{F}(\mathbf{x})) = \mathbf{G}(\mathbf{F}(\mathbf{x}_0)) + J_{\mathbf{G}}(\mathbf{y}_0)(J_{\mathbf{F}(\mathbf{x}_0)}(\mathbf{x}-\mathbf{x}_0)) + \mathcal{O}|\mathbf{x}-\mathbf{x}_0|^2 + \mathcal{O}(|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0)|)^2$$

Therefore, with $A = J_{\mathbf{G}}(\mathbf{y}_0) J_{\mathbf{F}}(\mathbf{x}_0)$,

$$\mathbf{G}(\mathbf{F}(\mathbf{x})) - [\mathbf{G}(\mathbf{F}(\mathbf{x}_0)) + A(\mathbf{x}-\mathbf{x}_0)]| = \mathcal{O}|\mathbf{x}-\mathbf{x}_0|^2 + \mathcal{O}(|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0)|)^2 \tag{8.8}$$

Clearly, $\displaystyle\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{|\mathbf{x}-\mathbf{x}_0|^2}{|\mathbf{x}-\mathbf{x}_0|} = 0$, and since $|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0)|^2 = \mathcal{O}(|\mathbf{x}-\mathbf{x}_0|^2)$,

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0)|^2}{|\mathbf{x}-\mathbf{x}_0|} = 0 \ .$$

This means that

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{|\mathbf{G}(\mathbf{F}(\mathbf{x})) - [\mathbf{G}(\mathbf{F}(\mathbf{x}_0)) + A(\mathbf{x}-\mathbf{x}_0)]\,|}{|\mathbf{x}-\mathbf{x}_0|} = 0 \ .$$

This shows that $\mathbf{G} \circ \mathbf{F}$ is differentiable, and that (8.5) holds. $\blacksquare$

8.5 Newton's method

We now explain one important application of Jacobian matrices: The two variable version of Newton's method for solving non linear equations.

Consider the following system of non linear equations:

$$\begin{aligned} x^2 + 2yx &= 4 \\ xy &= 1 \ . \end{aligned} \tag{8.9}$$

Any system of two equations in two variables can be written in the form

$$\begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \ . \end{aligned} \tag{8.10}$$

In this case we define $f(x, y) = x^2 + 2xy - 4$ and $g(x, y) = xy - 1$. All you have to do is to take whatever is on the right hand side of each equations, and subtract it off of both sides, leaving zero on the right. Just so we can standardize our methods, we shall always assume our equations have zero on the right hand side. If you run into one that doesn't, you know what to do as your first step: Cancel off the right hand sides.

Next, introducing $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{bmatrix}$, we can write this as a single vector equation

$$\mathbf{F}(\mathbf{x}) = 0 \ . \tag{8.11}$$

In the case of (8.9), we have

$$\mathbf{F}(x, y) = \begin{bmatrix} x^2 + 2yx - 4 \\ xy - 1 \end{bmatrix} \ . \tag{8.12}$$

In this case, we can solve (8.11) by algebra alone. For this purpose, the original formulation is most convenient. Using the second equation in (8.9) to eliminate y, the first equation becomes $x^2 = 2$. Hence $x = \pm\sqrt{2}$. The second equation says that $y = 1/x$ and so we have two solutions

$$(\sqrt{2}, 1/\sqrt{2}) \qquad \text{and} \qquad (-\sqrt{2}, -1/\sqrt{2}) \ .$$

In general, it may be quite hard to eliminate either variable, and algebra alone cannot deliver solutions.

There is a way forward: Newton's method is a very effective algorithm for solving such equations. This is a "successive approximations method". It takes a starting guess for the solution $\mathbf{x}_0$, and iteratively improves the guess. The iteration scheme produces an *infinite* *sequence* of approximate solutions $\{\mathbf{x}_n\}$. Under favorable circumstances, this sequence will converge *very rapidly* toward an exact solution. In fact, the number of correct digits x_n and y_n will more or less double double at each step. If you have one digit right at the outset, you may expect about a million correct digits after 20 iterations – more than you are ever likely to want to keep!

To explain the use of Newton's method, we have to cover three points:

(i) How one picks the starting guess $\mathbf{x}_0$.

(ii) How the iterative loop runs; i.e., the rule for determining $\mathbf{x}_{n+1}$ given $\mathbf{x}_n$.

(iii) How to break out of the iterative loop – we need a "stopping rule" that ensures us our desired level of accuracy has been achieved when we stop iterating.

We begin by explaining *(ii)*, the nature of the loop. Once we are familiar with it, we can better understand what we have to do to start it and stop it.

The basis of the method is the linear approximation formula for $\mathbf{F}$ at $\mathbf{x}_0$:

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}_0) + J_{\mathbf{F}}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\ . \tag{8.13}$$

Using this, we replace (8.11) with the approximate equation

$$\mathbf{F}(\mathbf{x}_0) + J_{\mathbf{F}}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = 0\ . \tag{8.14}$$

Don't let the notation obscure the simplicity of this: $\mathbf{F}(\mathbf{x}_0)$ is just a constant vector in $I\!R^2$ and $J_{\mathbf{F}}(\mathbf{x}_0)$ is just a constant 2×2 matrix. Using the shorter notation $\mathbf{F}(\mathbf{x}_0) = \mathbf{b}$ and $J_{\mathbf{F}}(\mathbf{x}_0) = A$, we can rewrite (8.14) as

$$A(\mathbf{x} - \mathbf{x}_0) = -\mathbf{b}\ .$$

We know what to do with this! We can solve this by row reduction. In fact, if A is invertible, we have $\mathbf{x} - \mathbf{x}_0 = A^{-1}\mathbf{b}$, or, what is the same thing,

$$\mathbf{x} = \mathbf{x}_0 - A^{-1}\mathbf{b}\ .$$

Writing this out in the full notation, we have a formula for the solution of (8.14)

$$\mathbf{x} = \mathbf{x}_0 - \left(J_{\mathbf{F}}(\mathbf{x}_0)\right)^{-1} \mathbf{F}(\mathbf{x}_0)\ . \tag{8.15}$$

We now define $\mathbf{x}_1$ to be this solution. To get $\mathbf{x}_2$ from $\mathbf{x}_1$, we do the same thing starting from $\mathbf{x}_1$. In general, we define $\mathbf{x}_{n+1}$ to be the solution of

$$\mathbf{F}(\mathbf{x}_n) + J_{\mathbf{F}}(\mathbf{x}_n)(\mathbf{x} - \mathbf{x}_n) = 0\ . \tag{8.16}$$

If $J_{\mathbf{F}}(\mathbf{x}_n)$ is invertible, this gives us

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \left(J_{\mathbf{F}}(\mathbf{x}_n)\right)^{-1} \mathbf{F}(\mathbf{x}_n)\ . \tag{8.17}$$

Now let's run through an example.

Example 8 (Using Newton's iteration) Consider the system of equations $\mathbf{F}(\mathbf{x}) = 0$ where $\mathbf{F}$ is given by (8.12). We will choose a starting point so that at least one of the equations in the system is satisfied, and the other is not *too* far off. This seems reasonable enough. Notice that with $x = y = 1$, $xy - 1 = 0$, while $x^2 - 2xy - 4 = -1$. Hence with $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ we have

$$\mathbf{F}(\mathbf{x}_0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix} .$$

Now let's write our system in the form $F(x, y) = 0$. We can do this with

$$F(\mathbf{x}) = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = \begin{bmatrix} x^2 + 2yx - 4 \\ xy - 1 \end{bmatrix} .$$

Computing the Jacobian, we find that

$$J_F(\mathbf{x}) = \begin{bmatrix} 2x + 2y & 2x \\ y & x \end{bmatrix} , \tag{8.18}$$

and hence

$$J_F(\mathbf{x}_0) = \begin{bmatrix} 4 & 2 \\ 1 & 1 \end{bmatrix} , \tag{8.19}$$

Hence (8.17) is

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 4 & 2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ 0 \end{bmatrix} .$$

Since

$$\begin{bmatrix} 4 & 2 \\ 1 & 1 \end{bmatrix}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -2 \\ -1 & 4 \end{bmatrix} ,$$

we find

$$\mathbf{x}_1 = [\,3/2, 1/2\,] .$$

Notice that $\mathbf{x}_1$ is indeed considerably closer to the exact solution $\left[\sqrt{2}, 1/\sqrt{2}\right]$ than $\mathbf{x}_0$. Moreover,

$$F(\mathbf{x}_1) = -\frac{1}{4} \begin{bmatrix} 1 \\ 1 \end{bmatrix} .$$

This is a better approximate solution; it is much closer to the actual solution. If you now iterate this further, you will find a sequence of approximate solutions converging to the exact solution $(\sqrt{2}, 1/\sqrt{2})$. You should compute $\mathbf{x}_2$ and $\mathbf{x}_3$ and observe the speed of convergence.

8.6 Choosing a starting point for Newton's method

With two variables, we can use what we know about generating plots of implicitly defined curves to locate good starting points. In fact, we can use such plots to determine the number of solutions. To do this, write $\mathbf{F}$ in the form

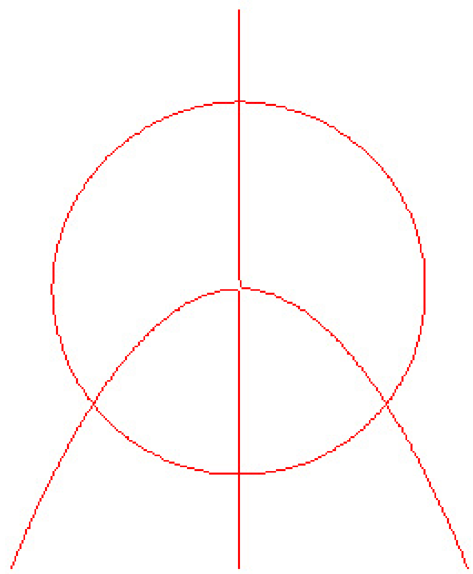
$$\mathbf{F}(x,y) = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} .$$

Each of the equations

$$f(x,y) = 0 \qquad \text{and} \qquad g(x,y) = 0$$

is an implicit definition of a curve. Points where the two curves intersect are points belonging to the solution set of both equations; i.e., to the solution set of $\mathbf{F}(\mathbf{x}) = 0$.

Example 9 (Using a graph to find a starting point for Newton's iteration) Consider the system $\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = 0$ where $f(x,y) = x^3 + xy$, and $g(x,y) = 1 - y^2 - x^2$. This is non linear, but simple enough that we can easily plot the curves. The equation $g(x,y) = 0$ is equivalent to $x^2 + y^2 = 1$, which is the equation for the unit circle. Since $f(x,y) = x(x^2 + y)$, $f(x,y) = 0$ is and only if $x = 0$, which is the equation of the y axis, or $y = -x^2$, which is the equation of a parabola. Here is a graph showing the intersection of the implicitly defined curves:



The axes have been left off since one branch of one curve is the y axis. Since one curve is the unit circle though, you can easily estimate the coordinates of the intersections anyway. As you see, there are exactly 4 solutions. Two of them are clearly the exact solutions $(0, \pm 1)$. The other two are where the parabola crosses the circle. Carefully measuring on the graph, you could determine (axes would now help) that $y \approx -0.618$ and $x \approx \pm 0.786$. This would give us tow good approximate solutions. applying Newton's method, we could improve them to compute as many digits as we desire of the exact solution.

If you have more than two variables, graphs become harder to use. An alternative to drawing the graph is to evaluate $\mathbf{F}(\mathbf{x})$ at all of the points in some grid, in some limited range of the variables. Use whichever grid points give $\mathbf{F}(\mathbf{x}) \approx= 0$ as your starting points.

8.7 Geometric interpretation of Newton's method

Newton's method is based on the tangent plane approximation, and so it has a geometric interpretation. This will help us to understand why it works when it does, and how we can reliably stop it.

Here is how this goes for the system

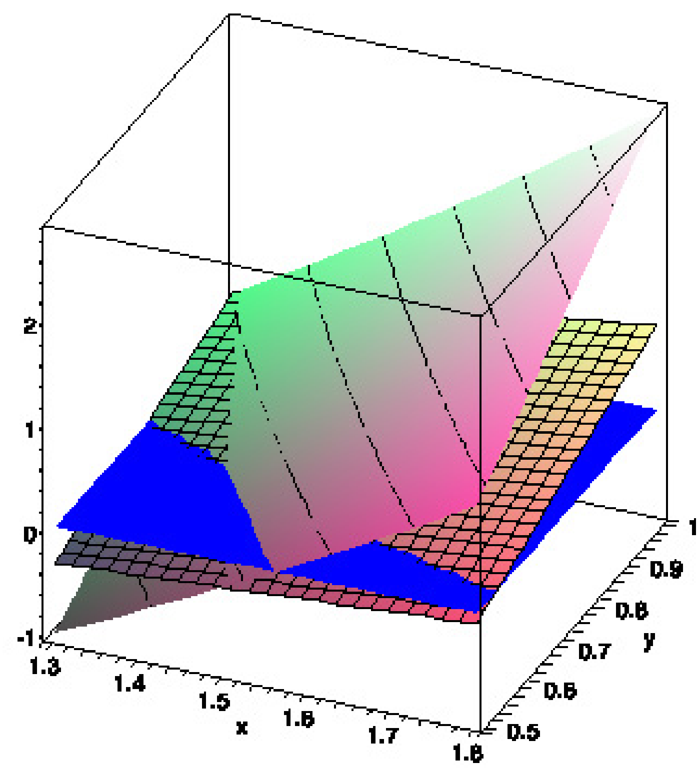
$$\begin{aligned} f(x,y) &= 0 \\ g(x,y) &= 0 \ . \end{aligned} \tag{8.20}$$

Replace this by the equivalent system

$$\begin{aligned} z &= f(x,y) \\ z &= g(x,y) \\ z &= 0\ . \end{aligned} \tag{8.21}$$

From an algebraic standpoint, we have taken a step backwards – we have gone from two equations in two variables to three equations in three variables. However, (8.21) has an interesting geometric meaning: The graph of $z = f(x,y)$ is a surface in $I\!R^3$, as is the graph of $z = g(x,y)$. The graph of $z = 0$ is just the x,y plane – a third surface. Hence the solution set of (8.21) is given by the intersection of 3 surfaces.

For example, here you a plot of the three surfaces in (8.21) when $f(x,y) = x^2 + 2xy - 4$ and $g(x,y) = xy - 1$, as in Example 8. Here, we have plotted $1.3 \le x \le 1.8$ and $0.5 \le y \le 1$, which includes one exact solution of the system (8.20) in this case. The plane $z = 0$ is the surface in solid color, $z = f(x,y)$ shows the contour lines, and $z = g(x,y)$ is the surface showing a grid. You see where all three surfaces intersect, and that is the where the solution lies.



You also see in this graph that the tangent plane approximation is pretty good in this region, so replacing the surfaces by their tangent planes will not wreak havoc on the graph. So here is what we do: Take any point (x_0, y_0) so that the three surface intersect *near* $(x_0, y_0, 0)$. Then replace the surfaces $z = f(x,y)$ and $z = g(x,y)$ by their tangent planes at (x_0, y_0), and compute the intersection of the tangent planes with the plane $z = 0$. This is a linear algebra problem, and hence is easily solved. Replacing $z = f(x,y)$ and $z = g(x,y)$

by the equations of their tangent planes at (x_0, y_0) amounts to the replacement

$$z = f(x,y) \qquad \to \qquad z = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)$$

and

$$z = g(x,y) \qquad \to \qquad z = g(\mathbf{x}_0) + \nabla g(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)$$

where $\mathbf{x}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$. This transforms (8.21) into

$$\begin{aligned} z &= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) \\ z &= g(\mathbf{x}_0) + \nabla g(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) \\ z &= 0\ . \end{aligned} \tag{8.22}$$

Now we can eliminate z, and pass to the simplified system

$$\begin{aligned} f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) &= 0 \\ g(\mathbf{x}_0) + \nabla g(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) &= 0\ . \end{aligned} \tag{8.23}$$

Since $J_{\mathbf{F}}(\mathbf{x}_0) = \begin{bmatrix} \nabla f(\mathbf{x}_0) \\ \nabla g(\mathbf{x}_0) \end{bmatrix}$, this is equivalent to (8.14) by the usual rules for matrix multiplication.

We see from this analysis that how close we come to an exact solution in one step of Newton's method depends on, among other things, how good the tangent plane approximation is at the current approximate solution. We know that tangent plane approximations are good when the norm of the Hessian is not too large. We can also see that there will be trouble if $J_{\mathbf{F}}$ is not invertible, or even if it ∇f and ∇g are nearly proportional, in which case $(J_{\mathbf{F}})^{-1}$ will have a large norm. There is a precise theorem, due the the 20th century Russian mathematician Kantorovich that can be paraphrased as saying that if $\mathbf{x}_0$ is not too far from an exact solution, $\|(J_{\mathbf{F}})^{-1}\|$ is not too large, and each component of $\mathbf{F}$ has a Hessian that is not too large, Newton's method works and converges very fast. The precise statement makes it clear what "not too large" means. We will neither sate it nor prove it here – it is quite intricate, and in practice one simply uses the method as described above, and stops the iteration when the answers stop changing in the digits that one is concerned with.

Exercises

1. Let $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{bmatrix}$ where $f(x,y) = x^3 + xy$, and $g(x,y) = 1 - y^2 - x^2$. let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Compute $J_{\mathbf{F}}(\mathbf{x})$ and $J_{\mathbf{F}}(\mathbf{x}_0)$.

2. Let $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{bmatrix}$ where $f(x,y) = \sqrt{x} + \sqrt{y} - 3$, and $g(x,y) = 4 - \sqrt{xy}$. let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Let $\mathbf{x}_0 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$. Compute $J_{\mathbf{F}}(\mathbf{x})$ and $J_{\mathbf{F}}(\mathbf{x}_0)$.

3. Let $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{bmatrix}$ where $f(x, y) = \sin(xy) - x$, and $g(x, y) = x^2y - 1$. let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Compute $J_{\mathbf{F}}(\mathbf{x})$ and $J_{\mathbf{F}}(\mathbf{x}_0)$.

4. Let $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{bmatrix}$ where $f(x, y) = x^3 + xy$, and $g(x, y) = 1 - 4y^2 - x^2$. Let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

a Compute $J_{\mathbf{F}}(\mathbf{x})$ and $J_{\mathbf{F}}(\mathbf{x}_0)$.

b Use $\mathbf{x}_0$ as a starting point for Newton's method, and compute the next approximate solution $\mathbf{x}_1$.

c Evaluate $\mathbf{F}(\mathbf{x}_1)$, and compare this with $\mathbf{F}(\mathbf{x}_0)$.

d Draw graphs of the curves implicitly defined by $f(x, y) = 0$ and $g(x, y) = 0$. How many solutions are there of this non linear system?

5. Let $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{bmatrix}$ where $f(x, y) = \sqrt{x} + \sqrt{y} - 3$, and $g(x, y) = x^2 + 4y^2 = 18$.

a Compute $\mathbf{F}(\mathbf{x}_0)$ for $\mathbf{x}_0 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$. does this look like a reasonable starting point? Compute $J_{\mathbf{F}}(\mathbf{x}_0)$. What happens if you try to use $\mathbf{x}_0$ as your starting point for Newton's method?

b Draw graphs of the curves implicitly defined by $f(x, y) = 0$ and $g(x, y) = 0$. How many solutions are there of this non linear system? Find starting points $\mathbf{x}_0$ near each of them with integer entries.

c Let $\mathbf{x}_0$ be the starting point that you found in part (b) that is closest to the x-axis. Compute the next approximate solution $\mathbf{x}_1$.

d Evaluate $\mathbf{F}(\mathbf{x}_1)$, and compare this with $\mathbf{F}(\mathbf{x}_0)$.

6. Let $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{bmatrix}$ where $f(x, y) = \sin(xy) - x$, and $g(x, y) = x^2y - 1$. Let $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

a Compute $J_{\mathbf{F}}(\mathbf{x})$ and $J_{\mathbf{F}}(\mathbf{x}_0)$.

b Use $\mathbf{x}_0$ as a starting point for Newton's method, and compute the next approximate solution $\mathbf{x}_1$.

c Evaluate $\mathbf{F}(\mathbf{x}_1)$, and compare this with $\mathbf{F}(\mathbf{x}_0)$.

d How many solutions of this system are there in the region $-2 \leq x \leq 2$ and $0 \leq y \leq 10$? Compute each of them to 10 decimal places of accuracy – using a computer, of course.

Section 9: Optimization problems in two variables

9.1 What is an optimization problem?

A *optimization problem* in two variables is one in which we are given a function $f(x,y)$, and a set D of *admissible points* in $I\!R^2$, and we are asked to find either the maximum or minimum value of $f(x,y)$ as (x,y) ranges over D.

It can be the case there is neither a maximum nor a minimum. Consider for example $f(x,y) = x+y$ and $D = I\!R^2$. Then

$$\lim_{t\to\infty} f(t,t) = \infty \qquad \text{and} \qquad \lim_{t\to\infty} f(-t,-t) = -\infty \ .$$

Also, for $f(x,y) = 1/(1+x^2+y^2)$ and $D = I\!R^2$, $f(x,y) \geq 0$ for all (x,y), and for any $m > 0$, $f(x,y) < m$ for all (x,y) with $x^2+y^2 > 1/\sqrt{m}$. Hence 0 is the greatest lower bound on the values of $f(x,y)$ as (x,y) ranges over D. Nonetheless, there is no point (x,y) in D such that $f(x,y) = 0$, so 0 is not a minimum value of f – it is not a value of f at all.

However, by Theorem 2 of Section 1, if D is a bounded and closed domain, and if f is a continuous function, then there is always point (x_1,y_1) in D with the property that

$$f(x_1,y_1) \geq f(x,y) \qquad \text{for all} \quad (x,y) \text{ in } D \ , \tag{9.1}$$

and there is always a point (x_0,y_0) in D with the property that

$$f(x_0,y_0) \leq f(x,y) \qquad \text{for all} \quad (x,y) \text{ in } D \ . \tag{9.2}$$

Definition (Maximizer and minimizer) Any point (x_1,y_1) satisfying (9.1) is called a *maximizer of f in D*, and any point (x_0,y_0) satisfying (9.1) is called a *minimizer of f in D*. The value of f at a maximizer is the *maximum value of f in D*, and the value of f at a minimizer is the *minimum value of f in D*.

To solve an optimization problem is to find all maximizers and minimizers, if any, and the corresponding maximum and minimum values. Our goal in this section is to explain a strategy for doing this. Another interesting example of what can go wrong even if D is closed and bounded, but f is merely separately continuous, and not continuous, is given right before the Theorem 2 in Section 1. But as long as D is closed and bounded, and f is continuous, minimizers and maximizers will exist, and our goal now is compute them.

9.2 A strategy for solving optimization problems

Recall that if $g(t)$ is a function of the single variable t, and we seek to maximize it on the closed bounded interval $[a,b]$, we proceed in three steps:

(1) We find all values of t in (a,b) at which $g'(t) = 0$. Hopefully there are only finitely many of these, say $\{t_1, t_2, \ldots, t_n\}$.

(2) Compute $g(t_1), g(t_2,)\ldots, g(t_n)$, together with $g(a)$ and $g(b)$. The largest number on this finite list is the maximum value, and the smallest is the minimum value. The

maximizers are exactly those numbers from among $\{t_1, t_2, \ldots, t_n\}$ together with a and b, at which f takes on the maximum values, and similarly for the minimizers.

The reason this works is that if t belongs to the open interval (a, b), and $g'(t) \neq 0$, it is possible to move either "uphill" or "downhill" while staying within $[a, b]$ by moving a bit to the right or the left, depending on whether the slope is positive or negative. *Hence no such point can be a maximizer or a minimizer.* This reasoning does not apply to exclude a or b, since at a, taking a step to the left is not allowed, and at b, taking a step to the right is not allowed. Thus, we have a short "list of suspects", namely the set of solutions of $g'(t) = 0$, together with a and b, and the maximizers and minimizers are there on this list.

In drawing up this "list of suspects", we are applying the Sherlock Holmes principle:

• *When you have eliminated the impossible, whatever else remains, however unlikely, is the truth.*

When all goes well, the elimination procedure reduces an *infinite* sets of suspects – all of the points in $[a, b]$ – to a *finite* list of suspects: $\{t_1, t_2, \ldots, t_n\}$ together with a and b. Finding the minimizers and maximizers among a finite list of points is easy – just compute the value of f at each point on the list, and see which ones give the largest and smallest values.

We now adapt this to two or more dimensions, focusing first on two. Suppose that D is a closed bounded domain. Let U be the interior of D and let B be the boundary. For example, if D is the closed unit disk

$$D = \{\mathbf{x} \ : \ |\mathbf{x} \leq 1\}$$

we have

$$U = \{\mathbf{x} \ : \ |\mathbf{x} < 1\} \qquad \text{and} \qquad B = \{\mathbf{x} \ : \ |\mathbf{x} = 1\} \ .$$

Notice that in this case, the boundary consists of infinitely many points.

• *In optimization problems, the big difference between one dimension and two or more dimensions is that in one dimension, the interval $[a, b]$ has only finitely boundary points – two– and there is no problem with throwing them into the list of suspects. But in two or more dimensions, the boundary will generally consist of infinitely many points, and if we throw them all onto the list of suspects, we make the list infinite, and therefore useless.*

We will therefore have to develop a "sieve" to filter the boundary B for suspects, as well as a "sieve" to filter the interior U for suspects. Here is the interior sieve:

If $\mathbf{x}$ belongs to U, the interior of D, and $\nabla f(\mathbf{x}) \neq 0$, then $\mathbf{x}$ is not a suspect. Recall that $\nabla f(\mathbf{x})$ points in the direction of steepest ascent. so if one moves a bit away from $\mathbf{x}$ in the direction $\nabla f(\mathbf{x})$, one moves to higher ground. Likewise, if one moves a bit away from $\mathbf{x}$ in the direction $-\nabla f(\mathbf{x})$, one moves to lower ground.

Since $\mathbf{x}$ is in the interior U of D, it is possible to move some positive distance from $\mathbf{x}$ in any direction and stay inside D. Hence, if $\mathbf{x}$ belongs to U, and $\nabla f(\mathbf{x}) \neq 0$, there are nearby points at which f takes on strictly higher and lower values. Such a point $\mathbf{x}$ cannot be a maximizer!

Apply the Sherlock Holmes principle: Eliminate all points $\mathbf{x}$ in U at which $\nabla f(\mathbf{x}) \neq 0$, and the remaining points are the only valid suspects in U. There are exactly the critical points in U – the points in U at which $\nabla f(\mathbf{x}) = 0$.

• *The suspect list from the interior U consists exactly of the critical points in U*

This is the "sieve" with which we filter the interior: We filter out the non critical points.

9.3 Lagrange's method for dealing with the boundary

Next, we need a "sieve" for the boundary. Here we need to make some assumptions. We will suppose first that the boundary B of D is the set of solutions of some equation

$$g(x, y) = 0 \ .$$

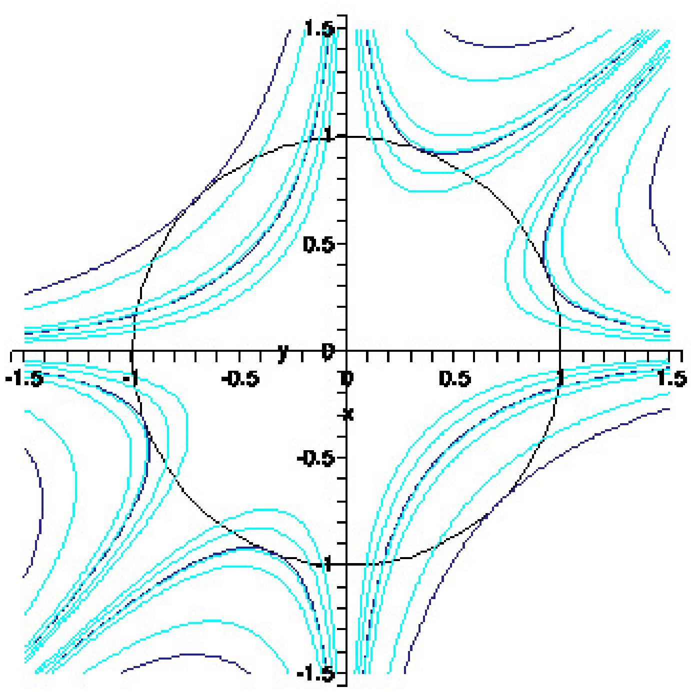
For example, when B is the unit circle, then we have

$$g(x, y) = x^2 + y^2 - 1 \ .$$

We can understand the geometrical idea behind Lagrange's method by looking at a contour plot of f with the boundary curve g superimposed. For example, let us take

$$f(x, y) = x^2y - y^2x \ .$$

Here is a contour plot showing the circle and some contour curves of g:



Most of the contour curves are drawn in lightly, but several others, drawn in more darkly are *tangent to the boundary at some point.*

These points of tangency are the ones to pay attention to. Here is why: Suppose that you are walking along the circle. If you path is "cutting across" contour curves of f at a point where ∇f is not zero, then you are going uphill or downhill at that point, and you cannot be at a minimum or maximum value on your path.

For example, suppose that you walk counterclockwise on the circle from $(1,0)$ to $(0,-1)$, and think of f as representing the altitude. Note that $f(1,0) = 0$ and $f(0,-1) = 0$, so that you are at sea level at the beginning and the end of the walk. However, in the middle of the trip, you pass through the point

$$\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) ,$$

and the value of f at this point is $-1/\sqrt{2}$.

If you look back at the contour plot, it is evident that you have been cutting across contour lines going downhill on the first part of the walk, and then cutting across them again going uphill on the second part of the walk.

Right in the middle, when you are neither going downhill nor uphill, you cannot be cutting across a contour curve, you must be walking parallel to one, so that the tangent line to your path is parallel to the tangent line of the contour curve. This is the point of minimal altitude along your walk.

In conclusion, we can eliminate any points on the boundary where $\nabla f \neq 0$ *and* where the tangent line to the boundary and the tangent line to the contour curve are not parallel, because if $\nabla f \neq 0$ at such a point, the path is cutting across contour curves, and headed uphill or downhill,

By the Sherlock Holmes principle, the only points to consider are points where the boundary curve is tangent to the contour curves, or where $\nabla f = 0$.

Now, we need to convert this tangency condition into an equation that we can solve. The tangent direction of the contour curve of f through a point is $(\nabla f)^\perp$ at that point. Also, since the boundary is a level curve of g, its tangent direction at any point on the boundary curve is given by $(\nabla g)^\perp$ at that point.

Therefore, the tangency means that these are parallel, or that $(\nabla f)^\perp$ is a multiple of $(\nabla g)^\perp$. But then we can undo the "perp"; the gradients must themselves be multiples:

$$\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}) \ . \tag{9.3}$$

We can rewrite this equation with λ eliminated as follows: The matrix $[\nabla f(\mathbf{x}), \nabla g(\mathbf{x})]$ has linearly independent columns, and therefore zero determinant, if and only if (9.3) is true. Hence (9.3) is equivalent to

$$\det\left([\nabla f(\mathbf{x}), \nabla g(\mathbf{x})]\right) = 0 \ . \tag{9.4}$$

Of course we are only interested in points x such that (9.4) is satisfied *and* $\mathbf{x}$ is on the boundary; i.e.,

$$g(\mathbf{x}) = 0 \ . \tag{9.5}$$

Together, (9.4) and (9.5) give us a system of two equations in two unknowns that we can solve to find all possible candidates for minima and maxima on the boundary. This method is often called *the method of Lagrange multipliers*, and presented with the equation (9.4) in the equivalent form (9.3). This is a vector equation, so it gives two numerical equations, which together with (9.5) give three equation for the three unknowns x, y and λ. It is often convenient to eliminate λ at the outset, which is what (9.4) does.

We have deduces the following Theorem, due to Lagrange,

Theorem 1 (Lagrange's Theorem) *Suppose that f and g are two functions on $I\!R^2$ with continuous first order partial derivatives. Let B denote the level curve of g given by $g(x,y) = 0$. Suppose that $\nabla g(x,y) \neq 0$ at any point along B.*

Then if $\mathbf{x}_0$ is a maximizer or minimizer of f on B,

$$\det\left([\nabla f(\mathbf{x}_0), \nabla g(\mathbf{x}_0)]\right) = 0 \ . \tag{9.6}$$

We will give a more formal proof at the end of this section. It is well worth going through it, but our first priority is to get to some examples. We now have a strategy for searching out maximizers and minimizers in a region D bounded by a level curve given by an equation of the form $g(x,y) = 0$.

(1) Find all critical points in U, the interior of D.

(2) Find all points on B, the boundary of D, at which (9.6) holds.

(3) The combined list of points found in *(1)* and *(2)* is a comprehensive list of suspected maximizers and minimizers. Hopefully it is a finite list. Now interrogate the suspects: Evaluate f at each of them, and see which produce the largest and smallest values. Case closed.

Example 1 (Finding minimizers and maximizers) Let $f(x,y) = x^4 + y^4 + 4xy$, which we have used in previous examples. Let D be the closed disk of radius 4 centered on the origin. We will now find the maximizers and minimizers of f in D.

We can write the equation for the boundary in the form $g(x,y) = 0$ by putting

$$g(x,y) = x^2 + y^2 - 16 \ .$$

Part (1): First, we look for the critical points. We have already examined this function in Section 7.4, and even drawn a contour plot of it. There, we found that f has exactly 3 critical points in all of $I\!R^2$, and all of them happen to be in the interior of D. They are

$$(0,0) \qquad (1,-1) \quad \text{and} \quad (-1,1) \ . \tag{9.7}$$

Part (2): Next, we look for solutions of (9.6) and $g(x,y) = 0$. Since

$$\nabla f(x,y) = 4\begin{bmatrix} x^3 + y \\ y^3 + x \end{bmatrix} \qquad \text{and} \qquad \nabla g(x,y) = 2\begin{bmatrix} x \\ y \end{bmatrix} \ ,$$

$$\det([\nabla f, \nabla g]) = 8\det\left(\begin{bmatrix} x^3 + y & x \\ y^3 + x & y \end{bmatrix}\right) = 8(x^3 y + y^2 - y^3 x - x^2) \ .$$

Hence (9.6) gives us the equation $x^3y + y^2 - y^3x - x^2 = 0$. Combining this with $g(x, y) = 0$ we have the system of equations

$$\begin{aligned} x^3y + y^2 - y^3x - x^2 &= 0 \\ g(x, y) = x^2 + y^2 - 16 &= 0 \ . \end{aligned}$$

The rest is algebra. The key to solving this system of equations is to notice that $x^3y + y^2 - y^3x - x^2$ can be factored:

$$x^3y + y^2 - y^3x - x^2 = (x^2 - y^2)(xy - 1) \ ,$$

so that the first equation can be written as $(x^2 - y^2)(xy - 1) = 0$. Now it is clear that either $x^2 - y^2 = 0$, or else $xy - 1 = 0$.

Suppose that $x^2 - y^2 = 0$. Then we can eliminate y from the second equation, obtaining $2x^2 = 16$, or $x = \pm 2\sqrt{2}$. If $y^2 = x^2$, then $y = \pm x$, so we get 4 solutions of the system this way:

$$(\pm 2\sqrt{2}, \pm 2\sqrt{2}) \ . \tag{9.8}$$

On the other hand, if $xy - 1 = 0$, $y = 1/x$, and eliminating y from the second equation gives us

$$x^2 + x^{-2} - 16 = 0 \tag{9.9}$$

Multiplying through by x^2, and writing $u = x^2$, we get $u^2 - 16u = -1$, so $u = 8 \pm \sqrt{63}$. Since $u = x^2$, there are four values of x that solve (9.9), namely $\pm\sqrt{8 \pm \sqrt{63}}$. The corresponding y values are given by $y = 1/x$. We obtain the final 4 solutions of (9.6):

$$(a, 1/a) \qquad \text{with} \qquad a = \pm\sqrt{8 \pm \sqrt{63}} \ . \tag{9.10}$$

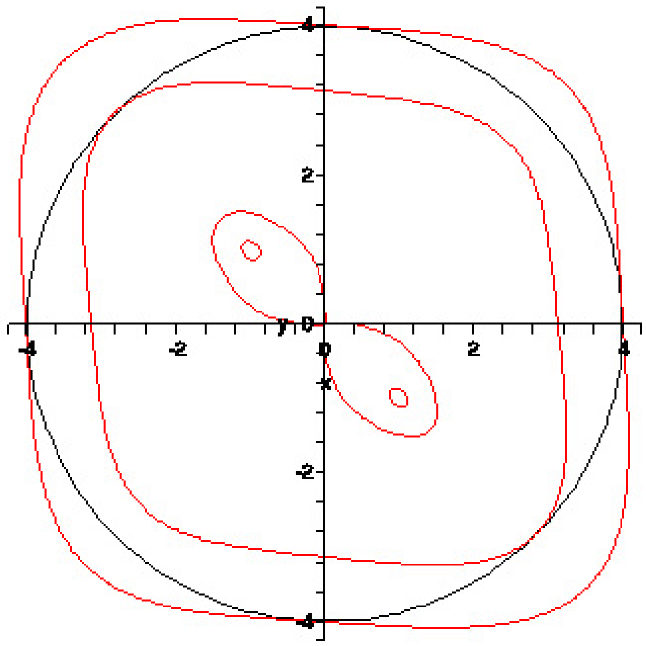
Part (3): We now round up and interrogate the suspects. There are 11 of them: Three from (9.7), four from (9.8), and four from (9.10).

Now for the interrogation phase:

$$\begin{aligned} f(0, 0) &= 0 \\ f(1, -1) &= -2 \\ f(-1, 1) &= -2 \\ f(2\sqrt{2}, 2\sqrt{2}) &= 96 \\ f(-2\sqrt{2}, 2\sqrt{2}) &= 96 \\ f(2\sqrt{2}, -2\sqrt{2}) &= 96 \\ f(-2\sqrt{2}, -2\sqrt{2}) &= 96 \\ f(\sqrt{8 + \sqrt{63}}, 1/\sqrt{8 + \sqrt{63}}) &= 258 \\ f(-\sqrt{8 + \sqrt{63}}, -1/\sqrt{8 + \sqrt{63}}) &= 258 \\ f(\sqrt{8 - \sqrt{63}}, 1/\sqrt{8 - \sqrt{63}}) &= 258 \\ f(-\sqrt{8 - \sqrt{63}}, -1/\sqrt{8 - \sqrt{63}}) &= 258 \end{aligned}$$

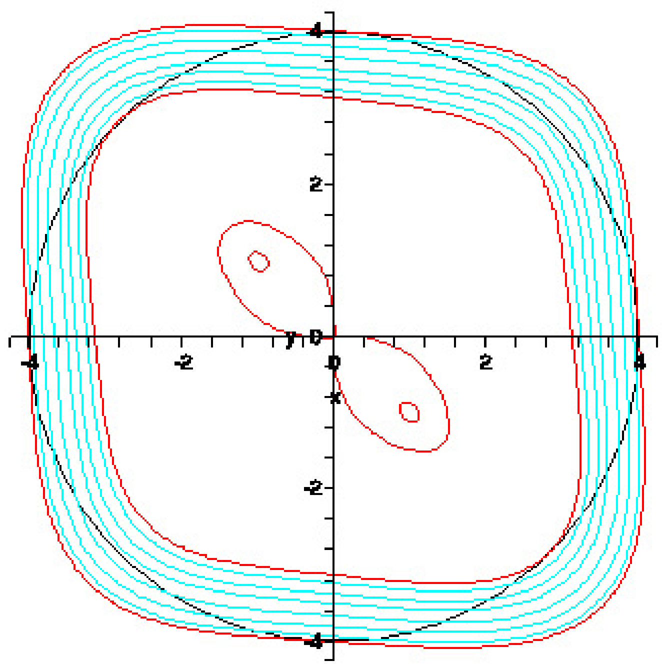
Evidently, the maximum value of f in D is 258, and the corresponding maximizers are the four points in (9.10). Also evidently, the minimum value of f in D is -2, and the corresponding maximizers are the two points $(-1, 1)$ and $(1, -1)$. Notice that the maximizers lie on the boundary, and the minimizers lie in the interior.

Here is a graph showing the boundary curve $g(x, y) = 0$, which is the circle of radius 4, and three contour curves of f, namely the contour curves at levels 258, 96, 0, and -1.9. These are the levels that showed up in our interrogation of the suspects, except that we have used the level -1.9 instead of -2 in the graph, since there are just two points, $(-1,1)$ and $(1,-1)$, for which $f(x,y) = -2$, and that would not graph well.



As you see, we get points of tangency in the graph exactly where our computations says we should.

Here is a graph showing some contour curves for levels between 96 and 258. As you can see, at all other levels, the circles cuts across the contour curves, which is what our computations say should be the case:



We finally close this example with one more observation: We can see from our computations that the minimum value of f on the boundary is 96.

In our next example, we use Theorem 1 to compute the distance between a point and a parabola. The idea is to write the equation for the parabola in the form $g(x,y) = 0$. For instance, if the parabola is given by $y = 2x^2$, we can take $g(x, y) = 2x^2 - y$.

Suppose for example that we want to fins the distance from $(1, 3)$ to this parabola. The square of the distance from any point (x, y) to $(1, 3)$ is

$$(x-1)^2 + (y-3)^2 \ .$$

To find the point on the parabola that is closest to $(1, 3)$, we use Theorem 1 to find the point (x_0, y_0) on the parabola that minimizes $(x-1)^2 + (y-3)^2$ – this is the point on the parabola that is closest to $(1, 3)$. Then, by definition, the distance from $(1, 3)$ to the parabola is the distance from $(1, 3)$ to this closest point.

Before proceeding to the calculations, note that the parabola is closed, but not bounded. Therefore, minima and maxima are not guaranteed to exist. Indeed, the parabola reaches upward for ever and ever, so there are points on the parabola that are arbitrarily far away from $(1, 3)$. That is, there is no furthest point.

But is it geometrically clear that there is a closest points.*

This particular problem can be solved using single variables methods. for each x, there is just one point on the parabola, namely $(x, 2x^2)$, the squared distance from this point to $(1, 3)$ is

$$(x-1)^2 + (2x^2-3)^2 = 4x^4 - 11x^2 - 2x + 10 \ .$$

Taking the derivative with respect to x, if x is a minimizer of this expression,

$$16x^3 - 22x - 2 = 0 \ .$$

This is a cubic equation, and so there are either one or three real roots. In fact, it has three real roots. Though you could solve for them exactly, the exact answers is very complicated, and involves cube roots of $(108 + 61\sqrt{7662})$. However, you can use Newton's method to compute the three roots numerically, finding

$$x_0 = -1.12419264... \qquad x_2 = -0.091465598... \qquad \text{and} \quad x_3 = 1.21565823...$$

Evaluating $4x^4 - 5x^2 - 2x + 10$ at each of these points, we see that x_3 is the minimizer.

We shall now look at this same problem from the point of view of Theorem 1:

Example 2 (Finding the distance to a parabola) Consider the parabola $y = 2x^2$, and the point $(1, 3)$. As explained above, to find the point on the parabola that is closest to $(1, 3)$, we minimize

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

on the curve $g(x, y) = 0$ where

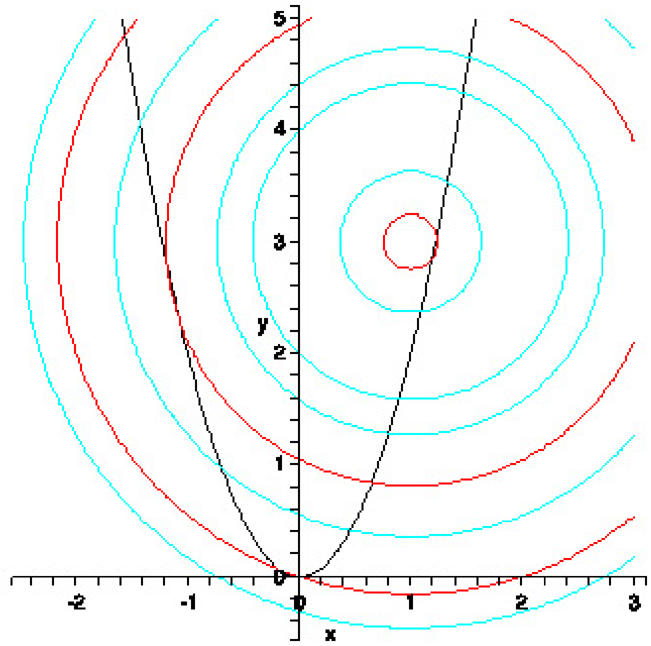
$$g(x, y) = 2x^2 - y \ .$$

Here is a graph showing the parabola, and some of the contour curves of f. As you can see, there are three places at which the parabola is tangent to the contour curves. Therefore, when we set up and

---

* Indeed, since $(1, 2)$ is on the parabola, and the distance of this point from $(1, 3)$ is 1, we only need to look for the minimum on the part of the parabola that lies in the closed unit disk centered on $(1, 3)$. This is closed and bounded, so a minimizer will exist.

solve the system consisting of (9.6) and $g(x, y) = 0$, we will find three solutions, as we found in our single variable approach.



To apply Theorem 1, we compute $\nabla f(x, y) = 2\begin{bmatrix} x-1 \\ y-3 \end{bmatrix}$ and $\nabla g(x, y) = \begin{bmatrix} 4x \\ -1 \end{bmatrix}$. Therefore, (9.6) reduces to

$$0 = \det\left(\begin{bmatrix} x-1 & 4x \\ y-3 & -1 \end{bmatrix}\right) = 1 + 11x - 4xy \ .$$

Now using $2x^2 - y = 0$, to eliminate y, we are left with

$$8x^3 - 11x + 1 = 0 \ ,$$

which is the same cubic equation that we encountered above. Solving it by Newton's method as above, we get the three suspect points:

$$\begin{aligned}
(x_1, y_1) &= (-1.12419264... \ , \ 2.527618184...) \\
(x_2, y_2) &= (-0.0914655981... \ , \ 0.01673191127...) \\
(x_3, y_3) &= (1.21565823... \ , \ 2.95564990...)
\end{aligned}$$

Evaluating f at each of these three points, we find

$$\begin{aligned}
f(x_1, y_1) &= 4.73533895... \\
f(x_2, y_2) &= 10.0911856... \\
f(x_3, y_3) &= 0.048475410...
\end{aligned}$$

and clearly (x_3, y_3) is the closest point. Its distance is the square root of the minimum value found above, namely 0.220171319..., since f is the squared distance.

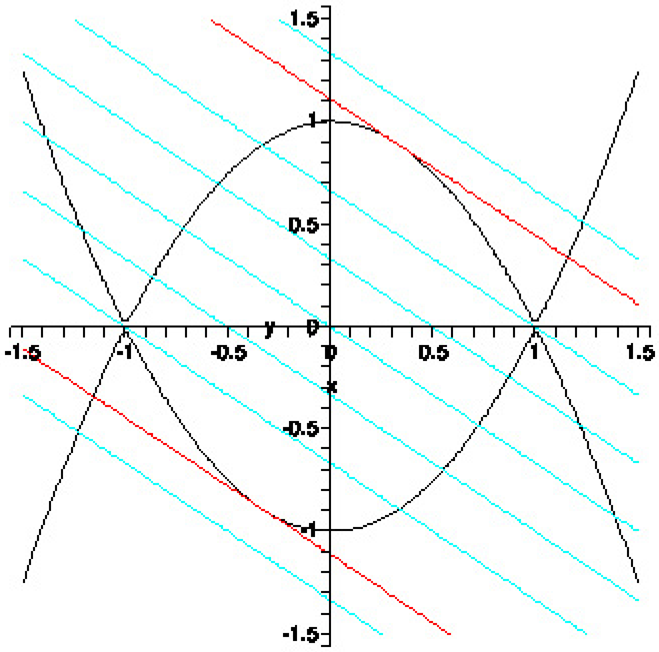
The reason we could use single variable methods to solve this problem was that it was easy to parameterize the parabola. But such explicit parameterizations can difficult to work with in general. The strength of Theorem 1 is that it allows us to work with implicit descriptions of the boundary. with more variables, this will be even more useful.

In our next example, we use Theorem 1 to solve an optimization problem in which the boundary is given in two pieces. This introduces one new feature into the method, as we shall see.

Example 3 (An optimization problem with a two piece boundary) Let D be the region lying on or above the parabola $y = x^2 - 1$, and on or below the parabola $y = 1 - x^2$. That is, D consists of the points (x, y) satisfying

$$x^2 - 1 \le y \le 1 - x^2 \ .$$

Let $f(x, y) = 2x + 3y$, and let us find all of the optimizers for f in D. Here is a graph showing the two parabolas that bound D, and some contour curves of f. (Since f is linear, these are just straight lines.)



Now, let us find the minima and maxima. First, $\nabla f(x, y) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, so there are no critical points.

Next, when we come to the boundary, we notice it is given in two pieces:

$$y = x^2 - 1 \qquad \text{for} \quad -1 \le x \le 1$$

and

$$y = 1 - x^2 \qquad \text{for} \quad -1 \le x \le 1 \ .$$

We now apply Theorem 1 to look for minima and maxima on these curves. The new twist is that we might have a maxima or a minima at the endpoints, no matter what the gradients are doing, just as we have to include the endpoints in an single variable optimization problem. But it is no problem to add these endpoints, $(-1, 0)$ and $(0, 1)$ to our list of suspects since there are only a finite number of them.

Now, for the upper arc, define

$$g(x, y) = 1 - x^2 - y \ .$$

Then $\nabla g(x, y) = \begin{bmatrix} -2x \\ -1 \end{bmatrix}$, so that $\det([\nabla f, \nabla g]) = 0$ reduces to

$$-6x + 2 = 0$$

so the point $(x, y) = (1/3, 8/9)$ goes on our list of suspects. (Notice that $x = 1/3$ does satisfy $-1 \le x \le 1$, or else we would not add it to the list of suspects, since then it would not have been on the boundary.)

For the lower arc, define

$$g(x,y) = x^2 - 1 - y \ .$$

Then $\nabla g(x,y) = \begin{bmatrix} 2x \\ -1 \end{bmatrix}$, so that $\det([\nabla f, \nabla g]) = 0$ reduces to

$$6x + 2 = 0$$

so the point $(x,y) = (-1/3, -8/9)$ goes on our list of suspects. (Again, $x = -1/3$ does satisfy $-1 \le x \le 1$, or else we would not add it to the list of suspects, since then it would not have been on the boundary.)

We now have our whole list of suspects:

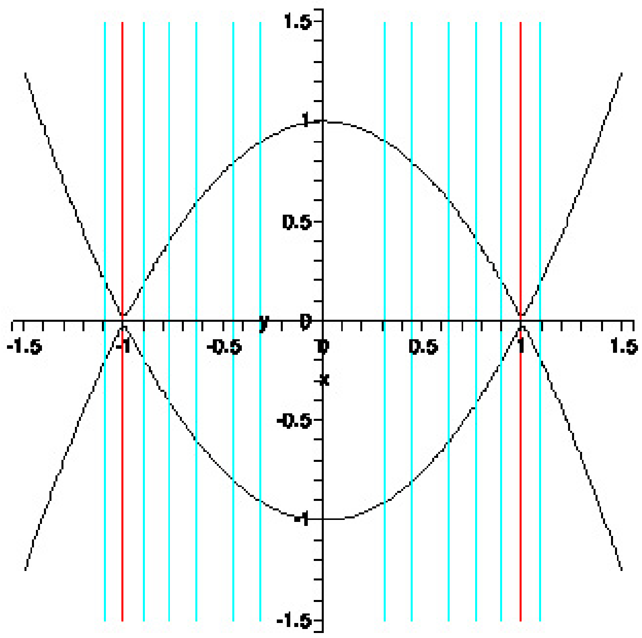
$$(-1,0) \qquad (1,0) \qquad (1/3, 8/9) \qquad (-1/3, -8/9) \ .$$

Proceeding to the interrogation,

$$\begin{gathered} f(-1,0) = -2 \\ f(1,0) = 2 \\ f(1/3,8/9) = 10/3 \\ f(-1/3,-8/9) = -10/3 \end{gathered}$$

You see that $(1/3, 8/9)$ is the maximizer, and $(-1/3, -8/9)$ is the minimizer. As you see, the contour curves are tangent to the boundary curves at these points.

The two points where the pieces of the boundary met turn out to be nothing special – in this problem. However, if you replace f by $f(x,y) = x^2$, and keep D the same, they both would have been maximizers. Here is a picture with some level curves of f, which once again are straight lines:



In this case, no point of the boundary is tangent to a level curve of f. There are infinitely many critical points in the region; any point (x,y) with $x = 0$ is a critical point. All the same, there is no problem with these, since $f(0,y) = 0$ for all y; all of the critical points are minima?

What are the maxima? Clearly $f(x,y)$ is largest in D when $x = \pm 1$, i.e., at $(-1,0)$ and $(1,0)$.

To close this subsection, we work out the minimizers and maximizers for the optimization problem that we discussed at the beginning of this section.

Example 4 (Minimizers and maximizers for $f(x,y) = x^2y - y^2x$ on the unit disk)

Let $f(x,y) = x^2y - y^2x$, and let D be given by $x^2 + y^2 \leq 1$. The boundary is then given by $g(x,y) = 0$ where $g(x,y) = x^2 + y^2 - 1$.

Part (1): First, we look for the critical points, and compute $\nabla f(x,y) = \begin{bmatrix} 2xy - y^2 \\ x^2 - 2xy \end{bmatrix}$. If (x,y) is a critical point, then

$$\begin{aligned} 2xy &= y^2 \\ 2xy &= x^2 \ . \end{aligned}$$

This tells us that $x^2 = y^2$. One solution is $x = y = 0$. If there is any other solution, neither x nor y is zero, and we can divide our equations through by a and y, getting

$$\begin{aligned} 2x &= y \\ 2y &= x \ . \end{aligned}$$

But then $x = 2y = 4x$, which is impossible for $x \neq 0$. Hence $(0,0)$ is the only critical point.

Part (2): To deal with the boundary, we compute

$$[\nabla f, \nabla g] = \begin{bmatrix} 2xy - y^2 & 2x \\ x^2 - 2xy & 2y \end{bmatrix} \ .$$

Taking the determinant, we deduce $4y^2x - 2y^3 - 2x^3 + 4x^2y = 0$. Together with $g(x,y) = 0$, this gives us the system of equations

$$\begin{aligned} 4y^2x - 2y^3 - 2x^3 + 4x^2y &= 0 \\ x^2 + y^2 - 1 &= 0 \ . \end{aligned}$$

This may look like a messy system system of non linear equations. Often in these problems, that is what one gets. Since we can always use Newton's method to solve such a system, that is not so bad.

However, in this case, one can solve the system using only algebra alone. To do this, notice that if (x,y) satisfies the first equation, and either $y = 0$, or $y = 0$, then both x and y are zero. But then (x,y) cannot satisfy the second equation. Therefore, if (x,y) is any solution, then neither $x = 0$ nor $y = 0$.

Armed with this information, we can divide the first equation through by y^3, and defining $z = x/y$, the first equation becomes

$$-z^3 + 2z^2 + 2z - 1 = 0 \ .$$

This is a cubic equation, but it is easy to see that $z = -1$ is a root. (In fact, $z = -1$ is the same as $y = -x$, and we have already observed that there should be a point of tangency on this line.) Knowing one root, it is easy to factor the cubic:

$$-z^3 + 2z^2 + 2z - 1 = (z+1)(-z^2 + 3z - 1) \ .$$

Factoring the quadratic, we find that the three roots are

$$z = -1 \qquad z = \frac{3 + \sqrt{5}}{2} \qquad \text{and} \qquad z = \frac{3 - \sqrt{5}}{2} \ .$$

Since $z = x/y$, so that $y = x/z$, we can use these values of z to eliminate y from the equation $x^2 + y^2 = 1$, getting

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

which mean that

$$x = \pm \frac{1}{\sqrt{1 + z^{-2}}}$$

for each of the three values of z found above, and for this same value of z, $y = x/z$. This gives us the six points of tangency that are shown in the first graph in this section.

For $z = -1$, we get $x = \pm 1/\sqrt{2}$, so this value of z gives us the two suspects

$$(x_1, y_1) = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right) \qquad \text{and} \qquad (x_2, y_2) = \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) .$$

For $z = (3 + \sqrt{5})/2$, we get the two suspects

$$(x_3, y_3) = \left(\frac{1+\sqrt{5}}{2\sqrt{3}}, \frac{1-\sqrt{5}}{2\sqrt{3}}\right) \qquad \text{and} \qquad (x_4, y_4) = \left(\frac{-1-\sqrt{5}}{2\sqrt{3}}, \frac{-1+\sqrt{5}}{2\sqrt{3}}\right) .$$

For $z = (3 - \sqrt{5})/2$, we get the two suspects

$$(x_5, y_5) = \left(\frac{-1+\sqrt{5}}{2\sqrt{3}}, \frac{-1-\sqrt{5}}{2\sqrt{3}}\right) \qquad \text{and} \qquad (x_6, y_6) = \left(\frac{1-\sqrt{5}}{2\sqrt{3}}, \frac{1+\sqrt{5}}{2\sqrt{3}}\right) .$$

Part (3): Now for the interrogation:

$$\begin{aligned}
f(0,0) &= 0\\
f(x_1, y_1) &= -\frac{1}{\sqrt{2}} = -0.707106781...\\
f(x_2, y_2) &= \frac{1}{\sqrt{2}} = 0.707106781...\\
f(x_3, y_3) &= -\frac{1}{3}\sqrt{\frac{5}{3}} = -0.430331483...\\
f(x_4, y_4) &= \frac{1}{3}\sqrt{\frac{5}{3}} = -0.430331483...\\
f(x_5, y_5) &= -\frac{1}{3}\sqrt{\frac{5}{3}} = -0.430331483...\\
f(x_6, y_6) &= \frac{1}{3}\sqrt{\frac{5}{3}} = -0.430331483...
\end{aligned}$$

As you can see (x_1, y_1) is the minimizer, and (x_2, y_2) is the maximizer.

9.4 Proof of Lagrange's Theorem

Proof of Theorem 1: let $\mathbf{x}(t)$ be a parameterization of a portion of the level curve of g through $\mathbf{x}_0$. As long as g has continuous partial derivatives in a neighborhood of $\mathbf{x}_0$, and $\nabla g(\mathbf{x}_0) \neq 0$, the Implicit Function Theorem guarantees the existence of such a parameterized curve with the property that

$$\mathbf{v} = \mathbf{x}'(0) \neq 0 .$$

Since $g(\mathbf{x}(t))$ is constant, by the definition of a level curve,

$$\frac{\mathrm{d}}{\mathrm{d}t} g(\mathbf{x}(t))\bigg|_{t=0} = \mathbf{v} \cdot \nabla g(\mathbf{x}_0) = 0 . \tag{9.11}$$

What is your instantaneous rate of change of altitude as you pass through $\mathbf{x}_0$? This is the directional derivative

$$\mathbf{v} \cdot \nabla f(\mathbf{x}_0) \ .$$

This quantity *must* be zero if $\mathbf{x}_0$ is to either minimize or maximize f along B. The reason is that if $\mathbf{v} \cdot \nabla f(\mathbf{x}_0) > 0$, you are going strictly uphill as you pass through $\mathbf{x}_0$, and so there is higher ground just ahead on your path, and lower ground just behind. Likewise, if $\mathbf{v} \cdot \nabla f(\mathbf{x}_0) < 0$, you are going strictly downhill as you pass through $\mathbf{x}_0$, and so there is lower ground just ahead on your path, and higher ground just behind.

Hence the only way that $\mathbf{x}_0$ can be a minimizer if

$$\mathbf{v} \cdot \nabla f(\mathbf{x}_0) = 0 \ . \tag{9.12}$$

To make this equation more useful, let's express it directly in terms of f and g. This is easy: Since $\mathbf{v} \neq 0$, and since we have from (9.11) and (9.12) that both $\nabla g(\mathbf{x}_0)$ and $\nabla f(\mathbf{x}_0)$ are orthogonal to $\mathbf{v}$, they must be parallel to each other. Hence the columns of $[\nabla f(\mathbf{x}_0), \nabla g(\mathbf{x}_0)]$ are proportional, and its rank is one, and so its determinant is zero.

On the other hand, if

$$\det\left([\nabla f(\mathbf{x}_0), \nabla g(\mathbf{x}_0)]\right) = 0 \ , \tag{9.13}$$

then the rank of the matrix cannot be two. Since $\nabla g(\mathbf{x}_0) \neq 0$, it is at least one, and hence it is exactly one. So this implies that $\nabla g(\mathbf{x}_0)$ and $\nabla f(\mathbf{x}_0)$ are parallel, and so (9.12) follows from (9.11) whenever (9.13) holds. ■

Exercises

1. Let D be the elliptical region bounded by the ellipse $x^2 + 4y^2 + 3y = 8$ so that D consists of all points (x, y) satisfying

$$x^2 + 4y^2 + 3y \leq 8 \ .$$

Let $f(x, y) = 2 - x - 2y$. Find the minimum and maximum values of f on D, and find all minimizers and maximizers.

2. Let D be the elliptical region bounded by the ellipse $x^2 + 4y^2 + 3y = 8$ so that D consists of all points (x, y) satisfying

$$x^2 + 4y^2 + 3y \leq 8 \ .$$

Let $f(x, y) = (x + y)(2 - x - 2y)$. Find the minimum and maximum values of f on D, and find all minimizers and maximizers.

3. Let D be the elliptical region bounded by the ellipse $x^2 + xy + y^2 = 1$ so that D consists of all points (x, y) satisfying

$$x^2 + xy + y^2 \leq 1 \ .$$

Let $f(x, y) = e^{xy-x}$. Find the minimum and maximum values of f on D, and find all minimizers and maximizers.

4. Find the point on the graph of $y = x^2$ that is closest to $(3, 1)$.

5. Find the maximum value of xy given that $x, y \geq 0$ and $x + y \leq 4$.

6. Find the points on the ellipse $x^2 + xy + y^2 = 1$ that are closest to $(3, 1)$, and furthest from $(3, 1)$.

7. Let D be the region consisting of all points (x, y) satisfying

$$x^2 \leq y \leq 3 + 2x \ .$$

Let $f(x, y) = x^2y - 3x$. Find the minimum and maximum values of f on D, and find all minimizers and maximizers.

8. Let D be the region consisting of all points (x, y) satisfying

$$x^2 - 4 \leq y \leq 4 + x^2 \ .$$

Let $f(x, y) = (x-1)^2+(y-1)^2$. Find the minimum and maximum values of f on D, and find all minimizers and maximizers.

9. Let f and g be functions on $I\!R^2$ with continuous partial derivatives, and suppose that the region D given by $\gamma(x, y) \leq 0$ is bounded. Let ϕ be a function on $I\!R$ with $\phi'(z) > 0$ for all z. Explain why if one applies the method of this section for finding minima and maxima of both f and $\phi(f)$ on D, one gets the same "list of suspects" in each case.