Taylor’s Formula

Consider the function $f(x,y)$. Recall that we can approximate $f(x,y)$ with a linear function in $x$ and $y$:

$$f(x,y) \approx f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$$

Notice that again this is just a linear polynomial in two-variables that does a good job of approximating $f$ near the point $(x,y) = (a,b)$. It’s also exactly the equation of the tangent plane to the surface $f$ at the point $(a,b)$.

**Example 1**

Find the linear approximation to $f(x,y) = xe^y$ at the point $(0,0)$.

We need evaluate the function and its first partial derivatives at the point $(0,0)$. We have

$$f = xe^y \quad f(0,0) = 0$$
$$f_x = e^y \quad f_x(0,0) = 1$$
$$f_y = xe^y \quad f_y(0,0) = 0$$

Then the linear approximation is

$$f(x,y) \approx 0 + (1 \cdot (x-0) + 0 \cdot (y-0)) = x = L(x,y)$$

**Example 2**

Use $L(x,y)$ to approximate $f(x,y) = xe^y$ at the point $(0.05,0.05)$ and find the error in the approximation.

$$L(0.05,0.05) = 0.05 \quad f(0.05,0.05) = 0.05e^{0.05} = 0.052564$$

$$|E(0.05,0.05)| = |L(0.05.0.05) - f(0.05,0.05)| = 2.5 \times 10^{-3}$$

OK, that’s pretty good. But what if we need to do better? The linearization is the best approximation by a linear polynomial of $f(x,y)$ near the point $(0,0)$. It’s natural to ask if we can get a better approximation if we use a quadratic polynomial.
It turns out that we can. The quadratic approximation of \( f(x, y) \) near the general point \((a, b)\) is given by

\[
f(x, y) \approx f(a, b) + f_x(a, b) (x - a) + f_y(a, b) (y - b) + \frac{1}{2} \left[f_{xx}(a, b) (x - a)^2 + 2f_{xy}(a, b) (x - a) (y - b) + f_{yy}(a, b) (y - b)^2\right]
\]

Notice that the first three terms in the approximation are just the linearization of \( f(x, y) \) about the point \((a, b)\). The additional terms are quadratic in \( x \) and \( y \) and involve the second partial derivatives of \( f \) evaluated at the point \((a, b)\).

**Example 3**

Find a quadratic approximation to \( f(x, y) = xe^y \) at the point \((0, 0)\).

We already computed the value of the function and it’s first partial derivatives at the point \((0, 0)\) when computing the linearization in the previous example. Now we need the second partials.

\[
\begin{align*}
  f_{xx} &= 0 & f_{xx}(0, 0) &= 0 \\
  f_{xy} &= e^y & f_{xy}(0, 0) &= 1 \\
  f_{yy} &= xe^y & f_{yy}(0, 0) &= 0 
\end{align*}
\]

Then the quadratic approximation is

\[
f(x, y) \approx L(x, y) + \frac{1}{2} \left(0 \cdot (x - 0)^2 + 2 \cdot 1 \cdot (x - 0) (y - 0) + 0 \cdot (y - 0)^2\right) \\
\approx x + xy = Q(x, y)
\]

**Example 4**

Use \(Q(x, y)\) to approximate \( f(x, y) = xe^y \) at the point \((0.05, 0.05)\) and find the error in the approximation.

\[
Q(0.05, 0.05) = 0.05 + (0.05)^2 = 0.0525 \quad f(0.05, 0.05) = 0.05e^{0.05} = 0.052564
\]

\[
|E(0.05, 0.05)| = |Q(0.05, 0.05) - f(0.05, 0.05)| = 6.4 \times 10^{-5}
\]

Notice that, not surprisingly, the quadratic approximation has a smaller error than the linear approximation.
OK, so we’ve found a linear approximation and quadratic approximation to \( f(x, y) \) near the point \((a, b)\). It turns out that we can come up with a polynomial approximation of any degree we like that approximates \( f(x, y) \) near the point \((a, b)\). This result is called **Taylor’s Theorem** for functions of two variables. Of course, we’ve seen this before.

Recall that in Calc II we used Taylor’s Formula to approximate a function \( f(x) \) near a point \( x = a \) by a sequence of polynomials.

**Theorem:** If \( f \) has \( n + 1 \) continuous partial derivatives in an open interval \( I \) around \( x = a \), then

\[
f(x) = f(a) + f'(a) (x - a) + \frac{f''(a)}{2} (x - a)^2 + \frac{f'''(a)}{3!} (x - a)^3 + \ldots + \frac{f^{(n)}(a)}{n!} (x - a)^n + \frac{f^{(n+1)}(c)}{(n+1)!} (x - a)^{n+1}
\]

for some \( c \in I \).

Notice that if we take \( n = 1 \)

\[
f(x) = f(a) + f'(a) (x - a) + \frac{f''(c)}{2} (x - a)^2
\]

then the first two terms are exactly the equation of the tangent line to \( f \) at the point \( x = a \), which in turn is exactly the linearization of \( f \) about the point \( x = a \).

The remainder term is just the next term in the Taylor Series. Notice that the second derivative in the remainder term is evaluated at some point \( x = c \) instead of \( x = a \). It turns out that for some value \( c \) between \( x \) and \( a \) this expression is exact. The hitch is that we don’t know exactly what \( c \) is. The remainder term is useful because it can be used to get an upper bound on the error incurred by using the linear approximation to approximate values of \( f \) near \( x = a \).

Of course, the power of Taylor’s Formula is that we can use it to obtain higher-order polynomial approximations to \( f \) near \( x = a \) for any degree polynomial that we like. If we want to approximate \( f \) using a quadratic polynomial then we use

\[
f(x) = f(a) + f'(a) (x - a) + \frac{f''(a)}{2} (x - a)^2 + \frac{f'''(c)}{3!} (x - a)^3
\]
Taylor’s Theorem for Functions of Two Variables

OK, so how do we do this for functions of two variables? It turns out it’s pretty straightforward and very similar to Taylor’s Theorem for functions of one variable. But to do this we need to introduce some new notation. First, let \( \Delta x = (x - a) \) and \( \Delta y = (y - b) \). Then we define a special operator as follows

\[
\left( \frac{\Delta x}{\partial x} + \frac{\Delta y}{\partial y} \right) f \bigg|_{(a, b)} = \Delta x f_x(a, b) + \Delta y f_y(a, b) = (x - a) f_x(a, b) + (y - b) f_y(a, b)
\]

Notice that the operator is a rule for applying this particular sum of partial derivatives to the function \( f \) and then evaluating them at the point \( (a, b) \). Notice also that this is exactly the linear part of the linearization \( L(x, y) \).

To get the quadratic term for the quadratic approximation we do this twice. Note that when taking derivatives, we treat \( \Delta x \) and \( \Delta y \) as constants.

\[
\left( \frac{\Delta x}{\partial x} + \frac{\Delta y}{\partial y} \right)^2 f \bigg|_{(a, b)} = \left( \frac{\Delta x}{\partial x} + \frac{\Delta y}{\partial y} \right) \left( \frac{\Delta x}{\partial x} + \frac{\Delta y}{\partial y} \right) f \bigg|_{(a, b)}
\]

\[
= \left( \frac{\Delta x}{\partial x} + \frac{\Delta y}{\partial y} \right) \left( \Delta x f_x + \Delta y f_y \right) \bigg|_{(a, b)}
\]

\[
= \Delta x^2 f_{xx} + 2 \Delta x \Delta y f_{xy} + \Delta y^2 f_{yy} \bigg|_{(a, b)}
\]

\[
= f_{xx}(a, b)(x - a)^2 + 2 f_{xy}(a, b)(x - a)(y - b) + f_{yy}(a, b)(y - b)^2
\]

So the quadratic approximation to \( f(x, y) \) at the point \( (a, b) \) can be written as

\[
f(x, y) \approx Q(x, y) = f(a, b) + \left( \frac{\Delta x}{\partial x} + \frac{\Delta y}{\partial y} \right) f \bigg|_{(a, b)} + \frac{1}{2} \left( \frac{\Delta x}{\partial x} + \frac{\Delta y}{\partial y} \right)^2 f \bigg|_{(a, b)}
\]

OK, so how do we generalize to an \( n \)-degree polynomial approximation of \( f(x, y) \), and what about that remainder term? It turns out that it exactly follows the same pattern of Taylor’s Theorem for functions of one variable, but the regular derivatives are replaced by powers of the operator described above. We have the following
Taylor’s Theorem. Suppose \( f(x, y) \) has \( n + 1 \) continuous partial derivatives in an open region \( R \) near \( (x, y) = (a, b) \), then for \( \Delta x = (x - a) \) and \( \Delta y = (y - b) \) we have

\[
f(x, y) = f(a, b) + \left( \Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right) \bigg|_{(a,b)} + \frac{1}{2} \left( \Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^2 \bigg|_{(a,b)} + \frac{1}{3!} \left( \Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^3 \bigg|_{(a,b)} + \cdots + \frac{1}{n!} \left( \Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^n \bigg|_{(a,b)} + \frac{1}{(n + 1)!} \left( \Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^{n+1} \bigg|_{(c_1,c_2)}
\]

where here the remainder term is evaluated at some (unknown) point \( (c_1, c_2) \) on the line connecting \( (a, b) \) and \( (x, y) \).

Taylor’s Theorem is powerful for a couple of reasons. The first is that it allows us a methodical way of coming up with polynomial approximations to \( f(x, y) \) near a point \( (a, b) \) for any degree polynomial that we like. The second is that it allows us to use the remainder term to get an upper bound on the error incurred when using the approximation.
Example 5

Use Taylor’s Formula to find a cubic approximation to \( f(x, y) = xe^y \) at the point \((0, 0)\).

If we want to do the cubic approximation then we need to evaluate the cubic term in the series. We have

\[
\left( \Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^3 f \bigg|_{(a,b)} = \Delta x^3 f_{xxx} + 3\Delta x^2 \Delta y f_{xxy} + 3\Delta x \Delta y^2 f_{xyy} + \Delta y^3 f_{yyy}
\]

It turns out that you can easily get the coefficients of the expansion from Pascal’s Triangle

\[
\begin{array}{c}
1 \\
1 1 \\
1 2 1 \\
1 3 3 1 \\
1 4 6 4 1
\end{array}
\]

To get the cubic terms in the example we need to evaluate some third-order partial derivatives

\[
\begin{align*}
f_{xxx} &= 0 & f_{xxx} (0, 0) &= 0 \\
f_{xxy} &= 0 & f_{xxy} (0, 0) &= 0 \\
f_{xyy} &= e^y & f_{xyy} (0, 0) &= 1 \\
f_{yyy} &= xe^y & f_{yyy} (0, 0) &= 0
\end{align*}
\]

Then

\[
f(x, y) \approx x + xy + \frac{1}{3!} \left( 0 \cdot (x - 0)^3 + 3 \cdot 0 \cdot (x - 0)^2 (y - 0) + 3 \cdot 1 \cdot (x - 0) (y - 0)^2 + 0 \cdot (y - 0)^3 \right)
\]

\[
= x + xy + \frac{xy^2}{2}
\]

If we use the cubic approximation to approximate the function at the point \((0.05, 0.05)\) we find that the exact error incurred is around \(1 \times 10^{-6}\), which is again better than the linear and quadratic approximations.
Error in the Taylor Approximation

The remainder term in Taylor’s Theorem gives us a way to find an upper bound on the error incurred by approximating a function \( f(x,y) \) using a Taylor polynomial. The remainder term is always taken to be the next term in the series beyond those used in the approximation. The only difference is that the remainder term is evaluated at some unknown point \((c_1, c_2)\) instead of \((a, b)\). For instance, if we want to bound the error in the linear approximation, the remainder term is the quadratic term in the polynomial.

Recall (one more time) that the linear approximation of \( f(x,y) \) at \((a, b)\) is given by

\[
L(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)
\]

Then from the theorem we see that

\[
E(x,y) = f(x,y) - L(x,y) = \frac{1}{2} [ f_{xx}(c_1, c_2)(x-a)^2 + 2f_{xy}(c_1, c_2)(x-a)(y-b) + f_{yy}(c_1, c_2)(y-b)^2 ]
\]

Then to get a bound on the worst-case scenario error we put absolute values around everything in the remainder term. This guarantees that we don’t get any helpful cancellation in the remainder from some terms being positive and some being negative.

\[
|E(x,y)| \leq \frac{1}{2} [ |f_{xx}| |x-a|^2 + 2 |f_{xy}| |x-a| |y-b| + |f_{yy}| |y-b|^2 ]
\]

If \( M \) is an upper bound on each of the second partial derivatives in the region of interest such that \( |f_{xx}|, |f_{xy}|, |f_{yy}| \leq M \) then we have

\[
|E(x,y)| \leq \frac{M}{2} [ |x-a|^2 + 2 |x-a| |y-b| + |y-b|^2 ] = \frac{M}{2} (|x-a| + |y-b|)^2
\]
Example 6

Consider again the function \( f(x, y) = xe^y \) near the point \((0, 0)\). Find a bound on the error if we use the linearization to approximate \( f \) for any \( x \) and \( y \) satisfying \(|x| \leq 0.1\) and \(|y| \leq 0.1\).

Note here that we want to find an upper bound when using the approximation to approximate \( f \) at any point in the region of interest. From the previous example we know that \( L(x, y) = x \).

To use the error formula we derived previously we need to find an upper bound on the second partial derivatives in the region \(|x| \leq 0.1\) and \(|y| \leq 0.1\). The second partials were

\[
\begin{align*}
  f_{xx} &= 0 \\
  f_{xy} &= e^y \\
  f_{yy} &= xe^y
\end{align*}
\]

We want to find the worst-case scenario for the error when \(|x| \leq 0.1\) and \(|y| \leq 0.1\). So we need to choose points that make the second derivatives as large as possible in the given region shown below.

To find \( M \) we need to figure out the largest values that any of the second partials can take on in the desired region. We have

\[
\begin{align*}
  |f_{xx}| &= 0 \\
  |f_{xy}| &= |e^y| \leq e^{0.1} \\
  |f_{yy}| &= |xe^y| \leq 0.1e^{0.1}
\end{align*}
\]

The biggest value that the three partials take on in the given region is \( M = e^{0.1} \), so we have

\[
|E(x, y)| \leq \frac{e^{0.1}}{2} (|x| + |y|)^2 \leq \frac{e^{0.1}}{2} (0.1 + 0.1)^2 \approx 2.2 \times 10^{-2}
\]

Note that this error bound is larger than the actual error incurred when we approximated \( f \) at the point \((0.05, 0.05)\). This makes sense because this error bound is valid for any point in the region of interest.
**Example 7**

**Example.** Use Taylor’s Theorem to find the linear approximation to $f(x, y) = y \cos x$ at the point $(\pi, 0)$ and use it to approximate $f$ at the point $(3.1, 0.15)$. Find a bound on the error if the linear approximation is used to approximate $f$ for $x$ in $[\pi - 0.1, \pi + 0.1]$ and $y$ in $[-0.2, 0.2]$.

For the linearization we need to evaluate $f$ and it’s first partial derivatives at $(\pi, 0)$.

\[
\begin{align*}
  f &= y \cos x \\
  f(x, 0) &= 0 \\
  f_x &= -y \sin x \\
  f_x(\pi, 0) &= 0 \\
  f_y &= \cos x \\
  f_y(\pi, 0) &= -1
\end{align*}
\]

Then the linearization of $f$ at $(\pi, 0)$ is given by

\[
L(x, y) = -y
\]

Evaluating both $f$ and $L$ at $(3.1, 0.15)$ we find

\[
f(3.1, 0.15) = -0.1494 \quad L(3.1, 0.15) = -0.15 \quad \text{and} \quad |E(x, y)| = 6.29 \times 10^{-4}
\]

To find an upper bound on the error we need to bound the second-partial derivatives of $f$ in the desired region. We have

\[
\begin{align*}
  f_{xx} &= -y \cos x \\
  f_{xy} &= -\sin x \\
  f_{yy} &= 0
\end{align*}
\]

In the region that we care about, we have the following bounds on the second-partial derivatives:

\[
\begin{align*}
  |f_{xx}| &= |-y \cos x| \leq 0.2 \\
  |f_{xy}| &= |-\sin x| \leq |-\sin(\pi + 0.1)| \approx 0.1 \\
  |f_{xx}| &= 0.0
\end{align*}
\]

So we pick $M = 0.2$. The error bound for general $(x, y)$ is then

\[
|E(x, y)| \leq \frac{0.2}{2} (|x - \pi| + |y|)^2
\]

Then, plugging in $x = \pi + 0.1$ and $y = 0.2$ we have the following bound on the error the linear approximation

\[
|E(x, y)| \leq \frac{0.2}{2} (0.1 + 0.2)^2 = 9 \times 10^{-3}
\]

which is greater than the exact error for the approximation at the point $(3.1, 0.15)$. 
Example 8

Consider again the function \( f(x, y) = xe^y \) near the point \((0, 0)\). Find a bound on the error if we use the quadratic approximation of \( f \) for any \( x \) and \( y \) satisfying \(|x| \leq 0.1 \) and \(|y| \leq 0.1\).

To find a bound on the quadratic approximation to \( f \) we use the cubic remainder term in the Taylor Polynomial. For a general \( f \) we have

\[
E(x, y) = f(x, y) - Q(x, y) = \frac{1}{3!} \left[ f_{xxx}(c_1, c_2) (x - a)^3 + 3 f_{xxy}(c_1, c_2) (x - a)^2 (y - b) + 
3 f_{xyy}(c_1, c_2) (x - a) (y - b)^2 + f_{yyy}(c_1, c_2) (y - b)^3 \right]
\]

Then putting absolute values around everything in the remainder term we get the following upper bound.

\[
|E(x, y)| \leq \frac{1}{3!} \left[ |f_{xxx}| |x - a|^3 + 3 |f_{xxy}| |x - a|^2 |y - b| + |f_{xyy}| |x - a| |y - b|^2 + |f_{yyy}| |y - b|^3 \right]
\]

If \( M \) is an upper bound on each of the second partial derivatives in the region of interest such that \(|f_{xxx}|, |f_{xxy}|, |f_{xyy}|, |f_{yyy}| \leq M \) then we have

\[
|E(x, y)| \leq \frac{M}{3!} \left[ |x - a|^3 + 3 |x - a|^2 |y - b| + 3 |x - a| |y - b|^2 + |y - b|^3 \right] \quad \Rightarrow \\
|E(x, y)| \leq \frac{M}{3!} (|x - a| + |y - b|)^3
\]

To determine the upper bound on the error in the example we need to bound each of the third partial derivatives in the region of interest. We have

\[
|f_{xxx}| = 0 \\
|f_{xxy}| = 0 \\
|f_{xyy}| = |e^y| \leq e^{0.1} \\
|f_{yyy}| = |xe^y| \leq 0.1 e^{0.1}
\]

Again we see that the largest value that the third partial derivatives take on on the region is \( M = e^{0.1} \). Then, plugging this into the error formula we have

\[
|E(x, y)| \leq \frac{M}{3!} (|x - 0| + |y - 0|)^3 \leq \frac{e^{0.1}}{3!} (0.1 + 0.1)^3 = 1.47 \times 10^{-3}
\]

This is the worst-case scenario error that can be incurred by using the quadratic approximation to approximate \( f \) in the region of interest.