SECTION 3.2 TAYLOR'S THEOREM

Homework: 1, 2, 3, 4, 5, 6 A, B, C

REVIEW OF TAYLOR SERIES AND TAYLOR POLYNOMIALS IN ONE VARIABLE

Suppose $f : \mathbf{R} \to \mathbf{R}$ and $a \in \mathbf{R}$. Suppose f and its derivatives of all orders are continuous. The *Taylor series* of f at a is given by

$$P(x,a) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

The *Taylor polynomial of order k* of f at a is given by

$$P_k(x,a) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

EXAMPLE 1: If $f(x) = e^x$ and $a = 0$, then for all n we have $f^{(n)}(x) = e^x$, and so $f^{(n)}(0) = 1$ and thus our n^{th} coefficient is $\frac{1}{n!}$, so we have

$$P(x,0) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots$$

and

$$P_k(x,0) = \sum_{n=0}^{k} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots + \frac{x^k}{k!}.$$

There are two basic issues to consider: For a fixed value of x what happens to $P_k(x,a)$ as $k \to \infty$? For a fixed value of k what happens to $P_k(x,a)$ as $x \to a$?

What one would hope is that as $k \to \infty$ one would have that $P_k(x,a) \to f(x)$. In this case we would say that $P(x,a)$ converges to $f(x)$, so that it would be legitimate to write $f(x) = P(x,a) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$. For $f(x) = e^x$ it can be shown that this does indeed happen. However, this need not always be the case.

EXAMPLE 2: Let $f(x) = \ln(x)$ and $a = 1$. Then $f'(x) = x^{-1}$, $f''(x) = -x^{-2}$, $f^{(3)}(x) = 2x^{-3}$, $f^{(4)}(x) = -3 \cdot 2x^{-4}$, etc. In general, for $n \geq 1$ we have $f^{(n)}(x) = (-1)^{n+1}(n-1)!\,x^{-n}$, and so $\frac{f^{(n)}(1)}{n!} = \frac{(-1)^{n+1}(n-1)!}{n!} = \frac{(-1)^{n+1}}{n}$. For $n = 0$ we have $f^{(0)}(1) = f(1) = \ln(1) = 0$. Thus we have

$$P(x,1) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(x-1)^n}{n} = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots.$$

It can be shown that if $0 < x \leq 2$, then this series will converge to $\ln(x)$. But look what happens for $x = 3$: We have that the n^{th} term of the series is $\frac{(-1)^{n+1}2^n}{n}$. By l'Hopital's Rule $\frac{2^n}{n} \to \infty$ as $n \to \infty$, so the series cannot converge to anything, much less to $\ln(3)$. More generally, the series doesn't converge at all when $x > 2$ even though $\ln(x)$ is defined for these values.

Suppose the Taylor series does converge. Does it converge to $f(x)$?

EXAMPLE 3: Let $f(x) = e^{-1/x^2}$ for $x \neq 0$ and set $f(0) = 0$. It can be shown that $f^{(n)}(0) = 0$ for all n, and hence $P(x, 0) = 0$. So the Taylor series converges to 0 for all x, but $f(x)$ is 0 only at $x = 0$. So the answer to our question is no.

Let's give a name to the difference between $f(x)$ and $P_k(x, a)$; we set $R_k(x, a) = f(x) - P_k(x, a)$. Thus $P(x, a)$ converging to $f(x)$ is equivalent to $R_k(x, a) \to 0$ as $k \to \infty$. There are several formulas for $R_k(x, a)$. One of them is the *Lagrange Remainder Formula* which says that there is a number t between x and a such that

$$R_k(x, a) = \frac{f^{(k+1)}(t)}{(k+1)!}(x-a)^{k+1}.$$

Note that the formula doesn't say how to find t, so it may appear pretty useless. However one can sometimes use it to estimate $P_k(x, a)$ well enough to show that $P(x, a)$ converges to $f(x)$. Consider $f(x) = e^x$ and $a = 0$ from Example 1. We have $R_k(x, 0) = \frac{e^t}{(k+1)!}x^{k+1}$. If $x \leq 0$, then $x \leq t \leq 0$, and so $e^x \leq e^t \leq e^0 = 1$. If $x \geq 0$, then $0 \leq t \leq x$ and so $1 = e^0 \leq e^t \leq e^x$. In either case this part of the formula won't go off to ∞ as $k \to \infty$. It can be shown that $\frac{x^{k+1}}{(k+1)!} \to 0$ as $k \to \infty$, so we have that $R_k(x, 0) \to 0$ as $k \to \infty$, and thus we can legitimately say that $e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots$ for all x.

We now consider the second question. For a fixed value of k what happens as $x \to a$? Here the results are much more encouraging. Look at $R_k(x, a)$. Note that as $x \to a$ the mysterious number t between x and a is forced to go to a as well, so we have that

$$R_k(x, a) = \frac{f^{(k+1)}(t)}{(k+1)!}(x-a)^{k+1} \to \frac{f^{(k+1)}(a)}{(k+1)!} \cdot 0 = 0.$$

Note that we can actually get a stronger result here; not only does $R_k(x, a)$ go to zero, but it goes to zero faster than $(x-a)^k$ does.

$$\frac{R_k(x, a)}{(x-a)^k} = \frac{f^{(k+1)}(t)}{(k+1)!}(x-a) \to \frac{f^{(k+1)}(a)}{(k+1)!} \cdot 0 = 0.$$

This is called *Taylor's Theorem.*

Theorem 1 (Taylor's Theorem) *Let* $P_k(x,a) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x-a)^n$.
Then $f(x) = P_k(x,a) + R_k(x,a)$, *where* $\frac{R_k(x,a)}{(x-a)^k} \to 0$ *as* $x \to a$.

For $k = 1$ note that $P_1(x,a) = f(a) + f'(a)(x-a)$, which is just the function giving the tangent line to the graph of f at $(a, f(a))$.

$$\lim_{x\to a} \frac{R_1(x,a)}{(x-a)} = \lim_{x\to a} \frac{f(x) - f(a) - f'(a)(x-a)}{x-a} = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} - f'(a) = f'(a) - f'(a) = 0,$$

so in this case Taylor's Theorem is just the fact that $f'(a)$ exists.

Consider Example 1 again. For $k = 1$ we have $P_1(x,0) = 1 + x$. Then

$$\lim_{x\to 0} \frac{R_1(x,0)}{(x-0)} = \lim_{x\to 0} \frac{e^x - 1 - x}{x} = \lim_{x\to 0} \frac{e^x - 1}{1} = \frac{e^0 - 1}{1} = \frac{0}{1} = 0$$

by l'Hopital's Rule. For $k = 2$ we have $P_2(x,0) = 1 + x + \frac{x}{2}$. Then

$$\lim_{x\to 0} \frac{R_2(x,0)}{(x-0)^2} = \lim_{x\to 0} \frac{e^x - 1 - x - x/2}{x^2} = \lim_{x\to 0} \frac{e^x - 1 - x}{2x} = 0,$$

where we have applied l'Hopital's Rule twice.

TAYLOR'S THEOREM FOR SEVERAL VARIABLES

Now consider a function $f : \mathbf{R}^n \to \mathbf{R}$ and a point $\mathbf{a} \in \mathbf{R}^n$. We assume that f and its partial derivatives of all orders are continuous. The general formulas for Taylor series and Taylor polynomials require a lot of notation, so we will concentrate on the cases $k = 1$ and $k = 2$. In order to further simplify the notation we set $\mathbf{h} = \mathbf{x} - \mathbf{a}$. Note that $\mathbf{x} \to \mathbf{a}$ is equivalent to $\mathbf{h} \to \mathbf{0}$.

Consider $k = 1$. Recall that in Chapter 2 we defined a linear approximation $g(\mathbf{x})$ to $f(\mathbf{x})$. This will be $P_1(\mathbf{x}, \mathbf{a})$.

$$P_1(\mathbf{x},\mathbf{a}) = g(\mathbf{x}) = f(\mathbf{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})(x_i - a_i) = f(\mathbf{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})\, h_i.$$

Note that by using the gradient $\nabla f(\mathbf{a}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{a})\right)$ we can rewrite this as

$$P_1(\mathbf{x},\mathbf{a}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{h}.$$

For the case $n = 2$ we can write $x = x_1$, $y = x_2$, $a = a_1$, $b = a_2$, $h = h_1$, and $k = h_2$. Then this is just the function whose graph is the tangent plane to the graph of f at $(a, b, f(a,b))$.

$$P_1(\mathbf{x},\mathbf{a}) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) = f(a,b) + \frac{\partial f}{\partial x}(a,b)h + \frac{\partial f}{\partial y}(a,b)k.$$

As before we set $R_1(\mathbf{x},\mathbf{a}) = f(\mathbf{x}) - g(\mathbf{x}) = f(\mathbf{x}) - P_1(\mathbf{x},\mathbf{a})$. Recall that we defined $f(\mathbf{x})$ to be differentiable at $\mathbf{a}$ if $\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - g(\mathbf{x})}{||\mathbf{x}-\mathbf{a}||} = 0$. With our new notation this translates into saying $\lim_{\mathbf{x}\to\mathbf{a}} \frac{R_1(\mathbf{x},\mathbf{a})}{||\mathbf{x}-\mathbf{a}||} = 0$. This is then the case $k = 1$ of Taylor's Theorem.

(The book, after introducing the notation $R_1(\mathbf{x},\mathbf{a})$, rewrites it on page 183 as $R_1(\mathbf{h},\mathbf{a})$! They do apologize for doing this, but even so, this is way too confusing, so I'm going to stick with $R_1(\mathbf{x},\mathbf{a})$. Just watch out for conflicting notation in the book.)

Note the following comparison with the one variable case. $f(a)$ is replaced by $f(\mathbf{a})$, and $f'(a)(x-a)$ is replaced by the sum of all the expressions of the form $\frac{\partial f}{\partial x_i}(\mathbf{a})(x_i - a_i)$. This gives us an idea of what the order two contribution should be for the case $k = 2$. Replace $\frac{1}{2}f^{(2)}(a)(x-a)^2$ by $\frac{1}{2}$ times the sum of all the expressions of the form $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})(x_i - a_i)(y_j - a_j)$. This gives the Taylor polynomial $P_2(\mathbf{x},\mathbf{a})$.

$$P_2(\mathbf{x},\mathbf{a}) = f(\mathbf{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})(x_i - a_i) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})(x_i - a_i)(y_j - a_j) =$$

$$f(\mathbf{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})h_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})h_i h_j.$$

There is also a shorthand version of this. Regard $\mathbf{x}$, $\mathbf{a}$, and $\mathbf{h}$ as column vectors and $\nabla f(\mathbf{a})$ as a row vector. Then the dot product $\nabla f(\mathbf{a}) \cdot \mathbf{h}$ can be regarded as the result of matrix multiplication $\nabla f(\mathbf{a})\mathbf{h}$. (A $1 \times n$ matrix times an $n \times 1$ matrix is a 1×1 matrix, i.e. just a number.) Now let $H(\mathbf{a})$ be the $n \times n$ matrix whose entry in the i^{th} row and j^{th} column is $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})$. This is called the *Hessian matrix* of f at $\mathbf{a}$. Let $\mathbf{h}^T$ be the transpose of $\mathbf{h}$, so it is just a row vector, or $1 \times n$ matrix. Then $\mathbf{h}^T H(\mathbf{a})\mathbf{h}$ is, as you can check, a 1×1 matrix, so is just a number, and its value is $\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})h_i h_j$. So we can write

$$P_2(\mathbf{x},\mathbf{a}) = f(\mathbf{a}) + \nabla f(\mathbf{a})\mathbf{h} + \frac{1}{2}\mathbf{h}^T H(\mathbf{a})\mathbf{h}.$$

The last term in this expression (including the factor of $\frac{1}{2}$) is sometimes called the *Hessian quadratic form of f at* $\mathbf{a}$.

Consider the case $n = 2$. Then, using the same notation as before, this becomes

$$P_2(\mathbf{x},\mathbf{a}) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) +$$

$$\frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(a,b)(x-a)^2 + \frac{\partial^2 f}{\partial x \partial y}(a,b)(x-a)(y-b) + \frac{\partial^2 f}{\partial y \partial x}(a,b)(y-b)(x-a) + \frac{\partial^2 f}{\partial y^2}(a,b)(y-b)^2\right] =$$

$$f(a,b)+\frac{\partial f}{\partial x}(a,b)h+\frac{\partial f}{\partial y}(a,b)k+\frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(a,b)h^2+\frac{\partial^2 f}{\partial x\partial y}(a,b)hk+\frac{\partial^2 f}{\partial y\partial x}(a,b)kh+\frac{\partial^2 f}{\partial y^2}(a,b)k^2\right]=$$

$$f(a,b)+\frac{\partial f}{\partial x}(a,b)h+\frac{\partial f}{\partial y}(a,b)k+\frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(a,b)h^2+2\frac{\partial^2 f}{\partial x\partial y}(a,b)hk+\frac{\partial^2 f}{\partial y^2}(a,b)k^2\right],$$

where in the last step we used the equality of mixed partials.

Warning: The most common mistakes on this material are (1) leaving out the factor of $\frac{1}{2}$ in front of the left bracket and (2) leaving out the factor of 2 inside the brackets.

EXAMPLE 4: Find the order two Taylor polynomial for $f(x,y)=\sqrt{x+2y+1}$ at $(0,0)$.

We first collect the computations we will need.

$f(x,y)=(x+2y+1)^{1/2}$, $f(0,0)=1$

$f_x(x,y)=\frac{1}{2}(x+2y+1)^{-1/2}$, $f_x(0,0)=\frac{1}{2}$

$f_y(x,y)=(x+2y+1)^{-1/2}$, $f_y(0,0)=1$

$f_{xx}(x,y)=-\frac{1}{4}(x+2y+1)^{-3/2}$, $f_{xx}(0,0)=-\frac{1}{4}$

$f_{xy}(x,y)=-\frac{1}{2}(x+2y+1)^{-3/2}$, $f_{xy}(0,0)=-\frac{1}{2}$

$f_{yy}(x,y)=-(x+2y+1)^{-3/2}$, $f_{yy}(0,0)=-1$

Plugging these values into our formula we get

$$P_2(\mathbf{x},\mathbf{0})=1+\frac{1}{2}h+k+\frac{1}{2}\left[-\frac{1}{4}h^2+2(-\frac{1}{2})hk-k^2\right]=1+\frac{1}{2}h+k-\frac{1}{8}h^2-\frac{1}{2}hk-\frac{1}{2}k^2$$

The book uses h_1 and h_2 instead of h and k and tacks on R_2 to get $f(\mathbf{x})$ rather than $P_2(\mathbf{x},\mathbf{0})$, so in this style the answer would be written

$$1+\frac{1}{2}h_1+h_2-\frac{1}{8}h_1^2-\frac{1}{2}h_1h_2-\frac{1}{2}h_2^2+R_2$$

EXAMPLE 5: Find the order two Taylor polynomial for $f(x,y)=\frac{1}{xy}$ at $(1,2)$.

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

$$f_x(x,y) = -\frac{1}{x^2y},\ f_x(1,2) = -\frac{1}{2}$$

$$f_y(x,y) = -\frac{1}{xy^2},\ f_y(1,2) = -\frac{1}{4}$$

$$f_{xx}(x,y) = \frac{2}{x^3y},\ f_{xx}(1,2) = 1$$

$$f_{xy}(x,y) = \frac{1}{x^2y^2},\ f_{xy}(1,2) = \frac{1}{4}$$

$$f_{yy}(x,y) = \frac{2}{xy^3},\ f_{yy}(1,2) = \frac{1}{4}$$

Plugging these values into our formula we get

$$P_2(\mathbf{x},\mathbf{a}) = \frac{1}{2}-\frac{1}{2}h_1-\frac{1}{4}h_2+\frac{1}{2}\left[h_1^2+2(\frac{1}{4}h_1h_2)+\frac{1}{4}h_2^2\right] = \frac{1}{2}-\frac{1}{2}h_1-\frac{1}{4}h_2+\frac{1}{2}h_1^2+\frac{1}{4}h_1h_2+\frac{1}{8}h_2^2$$

There are similar (but of course longer) formulas for higher values of k. They all have in common the general version of Taylor's Theorem.

Theorem 2 (Taylor's Theorem) *Let $P_k(\mathbf{x},\mathbf{a})$ be the Taylor polynomial of order k. Then $f(\mathbf{x}) = P_k(\mathbf{x},\mathbf{a}) + R_k(\mathbf{x},\mathbf{a})$, where $\frac{R_k(\mathbf{x},\mathbf{a})}{||\mathbf{x}-\mathbf{a}||^k} \to 0$ as $\mathbf{x} \to \mathbf{a}$.*