

CLASS NOTES

JOHN B. ETNYRE

1. THE TOTAL DERIVATIVE

Recall, from calculus I, that if $f : \mathbb{R} \rightarrow \mathbb{R}$ is a function then

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}.$$

We can rewrite this as

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - f'(a)h}{h} = 0.$$

Written this way we could then say that f is **differentiable at a** if there is a number $\lambda \in \mathbb{R}$ such that

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - \lambda h}{h} = 0.$$

Then if such a number λ exists we define $f'(a) = \lambda$. This might seem a bit of a round about way to define the derivative, but it generalizes nicely to functions of more than one variable. The way to interpret this definition is that $f'(a)$ is the “best linear approximation to $f(x)$ at a ” (or more precisely, the best linear approximation to $f(x+a) - f(a)$ at 0).

Now if $\bar{f} : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a function and $\bar{a} \in D$ then generalizing the idea above we want the derivative of \bar{f} at \bar{a} to be a linear map that approximates \bar{f} at \bar{a} . Recall, a linear map from \mathbb{R}^n to \mathbb{R}^m can be given by an $m \times n$ matrix T .

Indeed we say $\bar{f} : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ is **differentiable at \bar{a} in D** if there is an $n \times m$ matrix T , which we think of as a linear map $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, such that

$$\lim_{\bar{h} \rightarrow \bar{0}} \frac{\|\bar{f}(\bar{a} + \bar{h}) - \bar{f}(\bar{a}) - T\bar{h}\|}{\|\bar{h}\|} = 0.$$

If T exists it is called the **total derivative of \bar{f} at a** and we write

$$D\bar{f}(\bar{a}) = T.$$

Theorem 1.1. *If $\bar{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a differentiable function at \bar{a} then \bar{f} is continuous at \bar{a} .*

Example 1.2. Suppose $\bar{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is given by

$$\bar{f}(\bar{x}) = B\bar{x} + \bar{b},$$

where B is an $m \times n$ -matrix and $\bar{b} \in \mathbb{R}^m$. For example if $B = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 1 & 2 \end{bmatrix}$ and $\bar{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ then

$$f(x, y, z) = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2x + 3y + 4z + 1 \\ x + y + 2z + 2 \end{bmatrix}.$$

We claim that $D\bar{f}(\bar{a}) = B$ for all \bar{a} . To check this note

$$\begin{aligned} \lim_{\bar{h} \rightarrow \bar{0}} \frac{\|\bar{f}(\bar{a} + \bar{h}) - \bar{f}(\bar{a}) - B\bar{h}\|}{\|\bar{h}\|} &= \\ \lim_{\bar{h} \rightarrow \bar{0}} \frac{\|(B(\bar{a} + \bar{h}) - \bar{b}) - (B\bar{a} + \bar{b}) - B\bar{h}\|}{\|\bar{h}\|} &= \\ \lim_{\bar{h} \rightarrow \bar{0}} \frac{\|\bar{0}\|}{\|\bar{h}\|} &= 0. \end{aligned}$$

So indeed, $D\bar{f}(\bar{a}) = B$.

Example 1.3. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function and \bar{u} be a unit vector. Suppose that f is differentiable at \bar{a} . Then

$$\lim_{h \rightarrow 0} \frac{\|f(\bar{a} + h\bar{u}) - f(\bar{a}) - Df(\bar{a})(h\bar{u})\|}{\|h\bar{u}\|} = 0, \quad \text{by definition.}$$

so

$$\lim_{h \rightarrow 0} \left(\frac{f(\bar{a} + h\bar{u}) - f(\bar{a})}{h} - Df(\bar{a})(\bar{u}) \right) = 0.$$

Thus

$$[Df(\bar{a})]\bar{u} = \lim_{h \rightarrow 0} \frac{f(\bar{a} + h\bar{u}) - f(\bar{a})}{h} = D_{\bar{u}}f(\bar{a}) = [\nabla f(\bar{a})] \cdot \bar{u}.$$

Where $D_{\bar{u}}f(\bar{a})$ is the directional derivative of f in the direction of \bar{u} . Since it is easy to see check that if \bar{v} and \bar{w} are two vectors then $\bar{v} \cdot \bar{w} = \bar{v}^t \bar{w}$, where \bar{v}^t means the transpose of \bar{v} , we have

$$[Df(\bar{a})]\bar{u} = [\nabla f(\bar{a})]^t \bar{u}.$$

We formalize this in a theorem.

Theorem 1.4. *If $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ has continuous first partial derivatives in a neighborhood of some point \bar{a} in D , then*

$$Df(\bar{a}) = [\nabla f(\bar{a})]^t.$$

Example 1.5. Find the total derivative of $f(x, y) = \sin xy + x^2y$ at the point $(1, \pi)$.

$$\begin{aligned}
 Df(1, \pi) &= \left[\frac{\partial f}{\partial x}(1, \pi) \quad \frac{\partial f}{\partial y}(1, \pi) \right] \\
 &= \left[(y \cos xy + 2xy) \quad (x \cos xy + x^2) \right] \Big|_{(1, \pi)} \\
 &= [\pi \quad 0]
 \end{aligned}$$

In general we have

Theorem 1.6. Let $\bar{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be given by

$$\bar{f}(\bar{x}) = (f_1(\bar{x}), \dots, f_m(\bar{x})),$$

where $f_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \dots, m$. If $\frac{\partial f_i}{\partial x_j}$ is continuous at \bar{a} for all i and j , then \bar{f} is differentiable at \bar{a} and the total derivative is

$$D\bar{f}(\bar{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\bar{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\bar{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\bar{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\bar{a}) \end{bmatrix}.$$

That is, $D\bar{f}(\bar{a})$ is the $m \times n$ -matrix with the i^{th} row being the transpose of the gradient of f_i .

Example 1.7. Let $\bar{f}(x, y) = (e^{xy} + x^2y, x^2 + y^2, \frac{x}{y})$. Find the total derivative of \bar{f} at $(1, -1)$.

Before we start this problem let's note that f is a function from \mathbb{R}^2 to \mathbb{R}^3 so we expect to get a 3×2 -matrix as our answer. We can think of \bar{f} as

$$\bar{f} = (f_1(x, y), f_2(x, y), f_3(x, y)),$$

where $f_1(x, y) = e^{xy} + x^2y, f_2(x, y) = x^2 + y^2$, and $f_3(x, y) = \frac{x}{y}$. Then the matrix in the theorem is

$$\begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} \end{bmatrix} = \begin{bmatrix} (ye^{xy} + 2xy) & (xe^{xy} + x^2) \\ 2x & 2y \\ \frac{1}{y} & \frac{-x}{y^2} \end{bmatrix}$$

Since all the partial derivatives in this matrix are continuous at $(1, -1)$ we can just evaluate the terms at $(1, -1)$ to compute the total derivative. So

$$D\bar{f}(1, -1) = \begin{bmatrix} (-e^{-2} - 2) & (e^{-1} + 1) \\ 2 & -2 \\ -1 & -1 \end{bmatrix}.$$

Example 1.8. Let $\bar{f}(x, y) = (x \cos y, x \sin y)$. Find the total derivative. We know the partials of the functions $x \cos y$ and $x \sin y$ are continuous so

$$D\bar{f} = \begin{bmatrix} \cos y & -x \sin y \\ \sin y & x \cos y \end{bmatrix}$$

2. THE CHAIN RULE

Now suppose you have two functions $\bar{f} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ and $\bar{g} : \mathbb{R}^p \rightarrow \mathbb{R}^m$. If you compose these functions you get $\bar{g} \circ \bar{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$. We know that $D\bar{f}(\bar{a})$ is the best linear approximation to \bar{f} at \bar{a} and that $D\bar{g}(\bar{f}(\bar{a}))$ is the best linear approximation to \bar{g} at $\bar{f}(\bar{a})$. So what is the best linearly approximation to $\bar{g} \circ \bar{f}$ at \bar{a} ? Well, you would expect it just to be the composition of the other two linear approximations and compositions of linear maps corresponding to matrices is achieved by matrix multiplications. This leads to the multi-variable chain rule.

Theorem 2.1 (Mutli-Variable Chain Rule). *Let $\bar{f} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ and $\bar{g} : \mathbb{R}^p \rightarrow \mathbb{R}^m$ be two functions. If \bar{f} is differentiable at \bar{a} and \bar{g} is differentiable at $\bar{f}(\bar{a})$ then $\bar{g} \circ \bar{f}$ is differentiable at \bar{a} and*

$$D(\bar{g} \circ \bar{f})(\bar{a}) = [D\bar{g}(\bar{f}(\bar{a}))] [D\bar{f}(\bar{a})].$$

Example 2.2. Let $\bar{f}(x, y) = (x^2 + y^2, xy)$ and $\bar{g}(x, y) = (4xy, x - y, 3x^2 + 2y^2)$. Find the total derivative of $\bar{g} \circ \bar{f}$ at $(-3, 1)$. Use the chain rule to compute this.

$$\begin{aligned} D(\bar{g} \circ \bar{f})(-3, 1) &= [D\bar{g}(\bar{f}(-3, 1))] [D\bar{f}(-3, 1)] = [D\bar{g}(10, -3)] [D\bar{f}(-3, 1)] \\ &= \begin{bmatrix} 4y & 4x \\ 1 & -1 \\ 6x & 4y \end{bmatrix}_{(10, -3)} \begin{bmatrix} 2x & 2y \\ y & x \end{bmatrix}_{(-3, 1)} \\ &= \begin{bmatrix} -12 & 40 \\ 1 & -1 \\ 60 & -12 \end{bmatrix} \begin{bmatrix} -6 & 2 \\ 1 & -3 \end{bmatrix} \\ &= \begin{bmatrix} 112 & -144 \\ -7 & 5 \\ -372 & 156 \end{bmatrix} \end{aligned}$$

Example 2.3. Here we will check that a special case of the chain rule used when studying the gradient follows from the general chain rule. To this end let $\bar{f} : \mathbb{R} \rightarrow \mathbb{R}^n$ and $g : \mathbb{R}^n \rightarrow \mathbb{R}$ and write $\bar{f}(t) = (f_1(t), \dots, f_n(t))$. Then we have

$$\begin{aligned} \frac{d}{dt}(g \circ \bar{f})(t) &= D(g \circ \bar{f})(t) = [Dg(\bar{f}(t))] [D\bar{f}(t)] \\ &= \begin{bmatrix} \frac{\partial g}{\partial x_1}(\bar{f}(t)) & \dots & \frac{\partial g}{\partial x_n}(\bar{f}(t)) \end{bmatrix} \begin{bmatrix} \frac{\partial f_1}{\partial t}(t) \\ \vdots \\ \frac{\partial f_n}{\partial t}(t) \end{bmatrix} \\ &= \left(\frac{\partial g}{\partial x_1}(\bar{f}(t)) \right) \frac{\partial f_1}{\partial t}(t) + \dots + \left(\frac{\partial g}{\partial x_n}(\bar{f}(t)) \right) \frac{\partial f_n}{\partial t}(t). \end{aligned}$$

Example 2.4. Here we show how to use the multi-variable chain rule to compute derivatives in different coordinate systems.

Recall cartesian coordinates (x, y) are related to polar coordinates (r, θ) by

$$x = r \cos \theta, \quad y = r \sin \theta$$

and

$$r = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1} \frac{x}{y}.$$

Now suppose we are given a function expressed in xy -coordinates, like $g(x, y) = x^2y^3$. We can compute the total derivative of g

$$Dg(x, y) = [2xy^3 \quad 3x^2y^2].$$

But this is the total derivative with respect to the coordinates x and y . What about the derivative with respect to the coordinates r and θ ? We could plug $x = r \cos \theta$ and $y = r \sin \theta$ into g to express g in terms of r and θ . Then we could compute the total derivative of g with respect to r and θ . It turns out that it is just as easy (and in other cases easier!) to use the multi-variable chain rule. To this end consider

$$f(r, \theta) = (r \cos \theta, r \sin \theta).$$

Then g thought of as a function of r and θ is just

$$g(r, \theta) = g \circ f(r, \theta).$$

Now the chain rule gives us

$$\begin{aligned} D(g \circ f)(r, \theta) &= [Dg(f(r, \theta))] [Df(r, \theta)] \\ &= [2xy^3 \quad 3x^2y^2]_{(r \cos \theta, r \sin \theta)} \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}_{(r, \theta)} \\ &= [2(r \cos \theta)(r \sin \theta)^3 \quad 3(r \cos \theta)^2(r \sin \theta)^2] \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix} \\ &= [(2r^4 \cos^2 \theta \sin^3 \theta + 3r^4 \cos^2 \theta \sin^3 \theta) \quad (-2r^5 \cos \theta \sin^4 \theta + 3r^5 \cos^3 \theta \sin^2 \theta)] \\ &= [(5r^4 \cos^2 \theta \sin^3 \theta) \quad (r^5 \cos \theta \sin^2 \theta(3 \cos^2 \theta - 2 \sin^2 \theta))]. \end{aligned}$$

This is the total derivative with respect to r and θ . Notice, now that we have this worked out we know what $\frac{\partial g}{\partial r}$ and $\frac{\partial g}{\partial \theta}$ are:

$$\frac{\partial g}{\partial r} = 5r^4 \cos^2 \theta \sin^3 \theta, \quad \frac{\partial g}{\partial \theta} = r^5 \cos \theta \sin^2 \theta(3 \cos^2 \theta - 2 \sin^2 \theta).$$

Again this problem could have been worked out by just writing g in terms of r and θ and computing the derivative, but you cannot always do this. For example, in the next problem we don't know what our function is, just that it satisfies an equation in xy -coordinates and we want to see what the corresponding equation is in $r\theta$ -coordinate.

Example 2.5. A function $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ is said to satisfy Laplace's equation if

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

This equation is used in heat conduction, electrodynamics and many other areas. It is frequently useful to have Laplace's equation written in polar coordinates. To do this we need to figure out how to write the partial derivatives with respect to x and y in terms of partial derivatives with respect to r and θ . So again consider the change of coordinates

$$f(r, \theta) = (r \cos \theta, r \sin \theta).$$

Now as above

$$Du(r, \theta) = D(u \circ f)(r, \theta) = [Du]_{f(r, \theta)} [Df(r, \theta)].$$

So

$$\begin{bmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \end{bmatrix}_{f(r, \theta)} \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}$$

Denote the 2×2 -matrix on the right by A and notice that

$$A^{-1} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\frac{\sin \theta}{r} & \frac{\cos \theta}{r} \end{bmatrix}.$$

Thus

$$\begin{aligned} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \end{bmatrix}_{f(r, \theta)} &= \begin{bmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial \theta} \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\frac{\sin \theta}{r} & \frac{\cos \theta}{r} \end{bmatrix} \\ &= \left[\cos \theta \frac{\partial u}{\partial r} - \frac{\sin \theta}{r} \frac{\partial u}{\partial \theta} \quad \sin \theta \frac{\partial u}{\partial r} + \frac{\cos \theta}{r} \frac{\partial u}{\partial \theta} \right]. \end{aligned}$$

So

$$\begin{aligned} \frac{\partial u}{\partial x} &= \cos \theta \frac{\partial u}{\partial r} - \frac{\sin \theta}{r} \frac{\partial u}{\partial \theta}, \quad \text{and} \\ \frac{\partial u}{\partial y} &= \sin \theta \frac{\partial u}{\partial r} + \frac{\cos \theta}{r} \frac{\partial u}{\partial \theta}. \end{aligned}$$

So we have the first derivatives of u in terms of x and y written in terms of r and θ . Of course these same formulas hold for a function v as well. So if $v = \frac{\partial u}{\partial x}$ we have

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial x} \right) = \frac{\partial}{\partial x} v \\ &= \cos \theta \frac{\partial v}{\partial r} - \frac{\sin \theta}{r} \frac{\partial v}{\partial \theta} = \cos \theta \frac{\partial}{\partial r} v - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} v \\ &= \cos \theta \frac{\partial}{\partial r} \frac{\partial u}{\partial x} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \frac{\partial u}{\partial x} \\ &= \cos \theta \frac{\partial}{\partial r} \left(\cos \theta \frac{\partial u}{\partial r} - \frac{\sin \theta}{r} \frac{\partial u}{\partial \theta} \right) - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \left(\cos \theta \frac{\partial u}{\partial r} - \frac{\sin \theta}{r} \frac{\partial u}{\partial \theta} \right) \\ &= \cos^2 \theta \frac{\partial^2 u}{\partial r^2} - 2 \frac{\cos \theta \sin \theta}{r^2} \frac{\partial^2 u}{\partial r \partial \theta} + \frac{\sin^2 \theta}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\sin^2 \theta}{r} \frac{\partial u}{\partial r} + 2 \frac{\sin \theta \cos \theta}{r^2} \frac{\partial u}{\partial \theta}. \end{aligned}$$

Similarly

$$\frac{\partial^2 u}{\partial y^2} = \sin^2 \theta \frac{\partial^2 u}{\partial r^2} + 2 \frac{\cos \theta \sin \theta}{r^2} \frac{\partial^2 u}{\partial r \partial \theta} + \frac{\cos^2 \theta}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\cos^2 \theta}{r} \frac{\partial u}{\partial r} - 2 \frac{\sin \theta \cos \theta}{r^2} \frac{\partial u}{\partial \theta}.$$

Thus

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{1}{r} \frac{\partial u}{\partial r}.$$

And finally, Laplace's equation becomes

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{1}{r} \frac{\partial u}{\partial r} = 0.$$

3. MULTI-VARIABLE TAYLOR EXPANSIONS

Recall, from calculus I, that the single variable Taylor polynomial is the following: given a function $f : \mathbb{R} \rightarrow \mathbb{R}$ that is $(k + 1)$ -times differentiable near a then the k^{th} order Taylor polynomial of f about a is

$$P_{f,a}^k(x) = f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2 + \cdots + \frac{1}{k!} f^{(k)}(a)(x - a)^k$$

which we can write more succinctly

$$P_{f,a}^k(x) = \sum_{i=0}^k \frac{1}{i!} \left(\frac{d}{dx} \right)^i f(a)(x - a)^i.$$

This polynomial satisfies

$$|f(x) - P_{f,a}^k(x)| \leq K|x - a|^{k+1}$$

for some constant K and x near a . In other words, $P_{f,a}^k(x)$ is the k^{th} order polynomial that best approximates f near a . The Taylor series is just

$$\lim_{k \rightarrow \infty} P_{f,a}^k(x).$$

We will see that a formula almost identical to the boxed one above will also give the multi-variable Taylor polynomial. The difficult part is interpreting what this should mean in a multi-variable context. If we are in dimension n then a **multi-index** I is an ordered sequence of n non-negative integers

$$I = (i_1, \dots, i_n).$$

If $\bar{x} = (x_1, \dots, x_n)$ is a point in \mathbb{R}^n then define

$$\bar{x}^I = x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}.$$

Example 3.1. For example if $\bar{x} = (x, y, z)$ then

$$\bar{x}^{(1,0,0)} = x^1 y^0 z^0 = x,$$

$$\bar{x}^{(2,0,3)} = x^2 y^0 z^3 = x^2 z^3,$$

$$\bar{x}^{(1,1,1)} = xyz.$$

The **degree** of a multi-index $I = (i_1, \dots, i_n)$ is

$$|I| = i_1 + \cdots + i_n.$$

Example 3.2. We can write any multi-variable polynomial in multi-index notation. For example

$$xy^2 + xz + 3y + 7$$

can be written

$$\bar{x}^{(1,2,0)} + \bar{x}^{(1,0,1)} + 3\bar{x}^{(0,1,0)} + 7\bar{x}^{(0,0,0)},$$

where, of course, $\bar{x} = (x, y, z)$.

Example 3.3. A general second order polynomial in two variables can be written

$$(a_{(2,0)}x^2 + a_{(1,1)}xy + a_{(0,2)}y^2) + (a_{(1,0)}x + a_{(0,1)}y) + a_{(0,0)},$$

which we can write

$$\sum_{|I|=2} a_I \bar{x}^I + \sum_{|I|=1} a_I \bar{x}^I + \sum_{|I|=0} a_I \bar{x}^I,$$

where, of course, $\bar{x} = (x, y)$. Note all the multi-indices in two dimensions with degree 2 are $(2, 0)$, $(1, 1)$, $(0, 2)$. So the first sum in this formula is over these three multi-indices and give the first three terms in the polynomial. Similarly the multi-indices with degree 1 are $(1, 0)$, $(0, 1)$ and these two terms give us the next two terms in the polynomial. Finally the constant term is given from them multi-index $(0, 0)$ of degree 0. Now we can write this last expression for the polynomial even more succinctly as

$$\sum_{i=0}^2 \sum_{|I|=i} a_I \bar{x}^I.$$

Note this same formula gives an expression for a second order polynomial in any number of variables. The only difference is the multi-indices which equal a given degree will be different. For example in 3 dimensions, the multi-indices of degree 2 are $(2, 0, 0)$, $(0, 2, 0)$, $(0, 0, 2)$, $(1, 1, 0)$, $(1, 0, 1)$, $(0, 1, 1)$.

In general, a k^{th} order polynomial in the variable $\bar{x} = (x_1, \dots, x_n)$ is just

$$\sum_{i=0}^k \sum_{|I|=i} a_I \bar{x}^I.$$

We are almost ready for the multi-variable Taylor polynomial, we just need a few more definitions. If $I = (i_1, \dots, i_n)$ then

$$I! = (i_1!)(i_2!) \cdots (i_n!).$$

We abbreviate $\frac{\partial}{\partial x_i}$ by D_i , in other words

$$D_i f = \frac{\partial f}{\partial x_i}.$$

Finally,

$$D^I f = D_1^{i_1} D_2^{i_2} \cdots D_n^{i_n} f = \left(\frac{\partial}{\partial x_1} \right)^{i_1} \left(\frac{\partial}{\partial x_2} \right)^{i_2} \cdots \left(\frac{\partial}{\partial x_n} \right)^{i_n} f.$$

Example 3.4. For example if $I = (2, 1, 1)$ then

$$\begin{aligned} D^I(x^5y^2z^3) &= \left(\frac{\partial}{\partial x}\right)^2 \left(\frac{\partial}{\partial y}\right)^1 \left(\frac{\partial}{\partial z}\right)^1 (x^5y^2z^3) \\ &= 5 \cdot 4x^3 \cdot 2y \cdot 3z^2 = 120x^3yz^2. \end{aligned}$$

We can finally state:

Theorem 3.5 (Multi-variable Taylor Polynomials). *Let $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be a function with continuous mixed partial derivatives of order less than or equal to $k + 1$ near \bar{a} . Then*

$$P_{f,\bar{a}}^k(\bar{x}) = \sum_{i=0}^k \sum_{|i|=i} \frac{1}{i!} D^i f(\bar{a})(\bar{x} - \bar{a})^i$$

is the best k^{th} order polynomial approximation of f near \bar{a} in the sense that

$$|f(\bar{x}) - P_{f,\bar{a}}^k(\bar{x})| \leq K|\bar{x} - \bar{a}|^{k+1},$$

for some constant K and all \bar{x} near \bar{a} .

Notice that this theorem is almost identical to the one for single variable Taylor polynomials. The main difference is interpreting everything in terms of multi-indices.

Example 3.6. Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a function of two variables. Write out the 2^{nd} order Taylor polynomial at $\bar{a} = (a, b)$.

$$\begin{aligned}
P_{f,(a,b)}^2(x,y) &= \sum_{i=0}^2 \sum_{|i|=i} \frac{1}{i!} D^i f(\bar{a})(\bar{x} - \bar{a})^i \\
&= \sum_{|I|=0} \frac{1}{I!} D^I f(\bar{a})(\bar{x} - \bar{a})^I + \sum_{|I|=1} \frac{1}{I!} D^I f(\bar{a})(\bar{x} - \bar{a})^I + \sum_{|I|=2} \frac{1}{I!} D^I f(\bar{a})(\bar{x} - \bar{a})^I \\
&= \frac{1}{0!0!} D^{(0,0)} f(a,b)((x,y) - (a,b))^{(0,0)} \\
&\quad + \left(\frac{1}{1!0!} D^{(1,0)} f(a,b)((x,y) - (a,b))^{(1,0)} + \frac{1}{0!1!} D^{(0,1)} f(a,b)((x,y) - (a,b))^{(0,1)} \right) \\
&\quad + \left(\frac{1}{2!0!} D^{(2,0)} f(a,b)((x,y) - (a,b))^{(2,0)} + \frac{1}{1!1!} D^{(1,1)} f(a,b)((x,y) - (a,b))^{(1,1)} \right. \\
&\quad \left. + \frac{1}{0!2!} D^{(0,2)} f(a,b)((x,y) - (a,b))^{(0,2)} \right) \\
&= f(\bar{a})(x-a)^0(y-b)^0 \\
&\quad + \left(\frac{\partial f}{\partial x}(a,b)(x-a)^1(y-b)^0 + \frac{\partial f}{\partial y}(a,b)(x-a)^0(y-b)^1 \right) \\
&\quad + \left(\frac{1}{2} \frac{\partial^2 f}{\partial x^2}(a,b)(x-a)^2(y-b)^0 + \frac{\partial^2 f}{\partial x \partial y}(a,b)(x-a)^1(y-b)^1 \right. \\
&\quad \left. + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(a,b)(x-a)^0(y-b)^2 \right) \\
&= f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) \\
&\quad + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(a,b)(x-a)^2 + \frac{\partial^2 f}{\partial x \partial y}(a,b)(x-a)(y-b) + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(a,b)(y-b)^2.
\end{aligned}$$

OK that took some work. Let's look at a specific function.

Example 3.7. Compute the 2nd order Taylor polynomial of

$$f(x,y) = \sin(x + y^2)$$

at $(0,0)$. We know from that last example that we need to compute all derivatives up to order 2 and evaluate them at $(0,0)$. We have

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \cos(x + y^2), \\
\frac{\partial f}{\partial y} &= 2y \cos(x + y^2), \\
\frac{\partial^2 f}{\partial x^2} &= -\sin(x + y^2), \\
\frac{\partial^2 f}{\partial x \partial y} &= -2y \sin(x + y^2),
\end{aligned}$$

and

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

When we evaluate at $(0, 0)$ we get

$$f(0, 0) = 0, \quad \frac{\partial f}{\partial x}(0, 0) = 1, \quad \frac{\partial f}{\partial y}(0, 0) = 0,$$

and

$$\frac{\partial^2 f}{\partial x^2}(0, 0) = 0, \quad \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 0, \quad \frac{\partial^2 f}{\partial y^2}(0, 0) = 2.$$

Plugging this into the formula from the last exercise we get

$$\begin{aligned} P_{f,(0,0)}^2(x, y) &= 0 + 1(x - 0) + 0(y - 0) + 0(x - 0)^2 + 0(x - 0)(y - 0) + \frac{1}{2}2(y - 0)^2 \\ &= x + y^2. \end{aligned}$$

This is a fairly tedious way to compute Taylor polynomials. Let's find some shortcuts. First, recall the standard Taylor polynomials from single variable calculus. More precisely we have the Taylor series:

$$e^x = 1 + x + \frac{1}{2}x^2 + \dots = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!},$$

$$\cos x = 1 - \frac{1}{2}x^2 + \frac{1}{4!}x^4 + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!},$$

and

$$\ln(1 + x) = x - \frac{1}{2}x^2 + \dots = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}.$$

Now a theorem that helps us put Taylor polynomials together is the following.

Theorem 3.8. *We have the following*

- (1) *If f is a polynomial then $P_{f,\bar{0}}^k$ is all the terms in f of degree less than or equal to k .*
- (2) *$P_{f \pm g, \bar{a}}^k = P_{f, \bar{a}}^k \pm P_{g, \bar{a}}^k$,*
- (3) *$P_{fg, \bar{a}}^k$ is the terms of degree less than or equal to k in $P_{f, \bar{a}}^k P_{g, \bar{a}}^k$,*
- (4) *$P_{f \circ g, \bar{a}}^k$ is the terms of degree less than or equal to k in $P_{f, g(\bar{a})}^k \circ P_{g, \bar{a}}^k$*

Example 3.9. Compute the order 4 Taylor polynomial of $f(x, y) = \sin(x + y^2)$ at $(0, 0)$.

Note we can think of $f(x, y)$ as $g \circ h(x, y)$ where $g(x) = \sin x$ and $h(x, y) = x + y^2$. From item (1) in the theorem we know the Taylor series of a polynomial is just that polynomial so

$$P_{h,(0,0)}^4 = x + y^2.$$

Also $h(0, 0) = 0$ so from item (4) in the theorem we are interested in the Taylor polynomial of g at $0 = h(0, 0)$. From above we know the order 4 Taylor polynomial of $g(x)$ at 0 is

$$P_{g,0}^4(x) = x - \frac{x^3}{6}.$$

Let's compute $P_{g,0}^4 \circ P_{h,(0,0)}^4$.

$$\begin{aligned} P_{g,0}^4 \circ P_{h,(0,0)}^4 &= P_{h,(0,0)}^4 - \frac{1}{6} \left(P_{h,(0,0)}^4 \right)^3 \\ &= (x + y^2) + \frac{1}{6} (x + y^2)^3 \\ &= (x + y^2) + \frac{1}{6} (x^3 + 3x^2y^2 + 3xy^4 + y^6) \end{aligned}$$

We want all the terms of degree less than or equal to 4. So we have

$$P_{f,(0,0)}^4(x, y) = x + y^2 - \frac{1}{6}x^3 - \frac{1}{2}x^2y^2.$$

Example 3.10. Compute the order 2 Taylor polynomial of $f(x, y) = \cos(\pi e^x \cos y)$ at $(0, 0)$.

We will think of $f(x, y)$ as the composition $g \circ h(x, y)$ where $g(x) = \cos x$ and $h(x, y) = \pi e^x \cos y$. Now $P_{h,(0,0)}^2(x, y)$ is the terms of order less than or equal to 2 in $P_{e^x,0}^2 P_{\cos y,0}^2$. We have

$$\begin{aligned} P_{e^x,0}^2 P_{\cos y,0}^2 &= \pi \left(1 + x + \frac{x^2}{2} \right) \left(1 - \frac{y^2}{2} \right) \\ &= \pi \left[1 - \frac{y^2}{2} + x - \frac{xy^2}{2} + \frac{1}{2} \left(x^2 - \frac{x^2y^2}{2} \right) \right] \end{aligned}$$

Thus

$$P_{h,(0,0)}^2(x, y) = \pi \left(1 + x - \frac{y^2}{2} + \frac{x^2}{2} \right).$$

Now $h(0, 0) = \pi$ so we need to find the expansion of $g(x)$ at π . We know

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!},$$

so

$$\cos(x - \pi) = \sum_{n=0}^{\infty} (-1)^n \frac{(x - \pi)^{2n}}{(2n)!},$$

but

$$\cos(x - \pi) = \cos x \cos \pi - \sin x \sin \pi = -\cos x.$$

Thus

$$\cos x = - \sum_{n=0}^{\infty} (-1)^n \frac{(x - \pi)^{2n}}{(2n)!},$$

and

$$P_{g,\pi}^2 = -1 + \frac{1}{2}(x - \pi)^2.$$

We now need to consider $P_{g,\pi}^2 \circ P_{h,(0,0)}^2$:

$$\begin{aligned} P_{g,\pi}^2 \circ P_{h,(0,0)}^2 &= -1 + \frac{1}{2} \left(\pi \left(1 + x - \frac{y^2}{2} + \frac{x^2}{2} \right) - \pi \right)^2 \\ &= -1 + \frac{\pi^2}{2} \left(x^2 - y^2 x + x^3 + \frac{y^4}{4} - \frac{x^2 y^2}{2} + \frac{x^4}{4} \right). \end{aligned}$$

Taking the terms of order less than or equal to 2 we have

$$P_{f,(0,0)}^2(x, y) = -1 + \frac{\pi^2}{2} x^2.$$