Lecture 14:

Recall Taylor's Thm (in 1 variable).

If a function $f: \mathbb{R} \to \mathbb{R}$ has derivatives of all orders (is C^∞, or smooth), at x_0, and has a power series expansion, then

p.159

$$f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(x_0)}{i!}(x-x_0)^i$$

$$= f(x_0) + \sum_{i=1}^{\infty} \frac{f^{(i)}(x_0)}{i!}(x-x_0)$$

$$= f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2}(x-x_0)^2 + \frac{f'''(x_0)}{3!}(x-x_0)^3 + \cdots$$

Notice that in the above description, we refer to the function value itself as the 0th derivative of f at x_0, and that $0! = 1! = 1$.

Notes ① Chopping off the series at the kth term creates a degree k polynomial in x: This is called the kth Taylor Polynomial of f at x_0, Notes ① cont'd.

And is considered the best kth-degree polynomial to approximate $f(x)$ "near" $x=x_0$. why? Because it will have all of the same derivatives at x_0 as f does, up to order k:

So call $T_k(f, x_0)$ the kth Taylor Poly of f at x_0

$$T_k(f, x_0) = f(x_0) + \sum_{i=1}^{k} \frac{f^{(i)}(x_0)}{i!}(x-x_0)^i$$

ex. Let $f(x) = \ln x$. Then we can calculate $T_5(\ln x, 1)$ as follows:

$$f(1) = \ln 1 = 0, \quad f'(1) = \frac{d}{dx}\Big|_{x=1}(\ln x) = \frac{1}{x}\Big|_{x=1} = 1$$

$$f''(1) = \frac{d^2}{dx^2}\Big|_{x=1}(\ln x) = \frac{d}{dx}\Big|_{x=1}\frac{1}{x} = -\frac{1}{x^2}\Big|_{x=1} = -1.$$

$$f'''(1) = \frac{2}{x^3}\Big|_{x=1} = 2, \quad f^{(iv)}(1) = \frac{-6}{x^4}\Big|_{x=1} = -6, \quad f^{(5)}(1) = \frac{24}{x^5}\Big|_{x=1} = 24$$

And so $T_5(\ln x, 1) = f(1) + f'(1)(x-1) + \frac{f''(1)}{2}(x-1)^2 + \cdots + \frac{f^{(5)}(1)}{120}(x-1)^5$

$$= 1 + (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4 + \frac{1}{5}(x-1)^5$$

Notes ② How good an approximation is the kth Taylor Polynomial $T_k(f, x_0)$ to f at x_0?

Taylor's Thm stipulated that if f had k derivatives at x_0, then there exists a function $R_k(f, x_0)$, called the kth Remainder Function, where

p. 159.

$$f(x) = T_k(f, x_0) + R_k(f, x_0)$$

where $\lim\limits_{x \to x_0} R_k(f, x_0) = 0$.

It was determined later that $R_k(f, x_0)$ has a form:

$$R_k(f, x_0) = \int_{x_0}^{x} \frac{(x-t)^k}{k!} f^{(k+1)}(t)\, dt$$

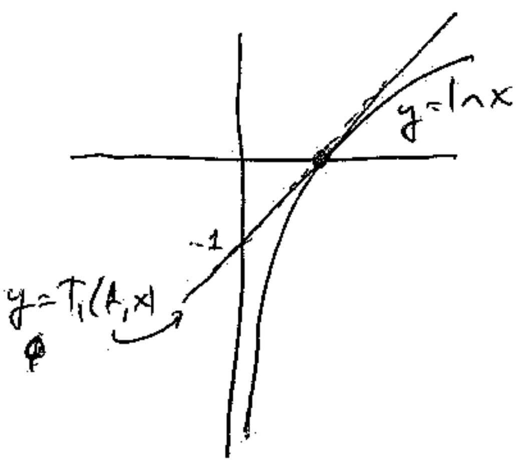
and that not only does $R_k(f, x_0) \to 0$ as $x \to x_0$, but it goes to 0 quickly:

equation 2, p. 159.

$$\lim_{x \to x_0} \frac{R_k(f, x_0)}{(x-x_0)^k} = 0.$$

Notes ③ The whole idea of a tangent line approximation to the graph of f at x_0 is simply the 1st Taylor Polynomial
$$T_1(f, x_0) = f(x_0) + f'(x_0)(x - x_0).$$

If you graph $y = f(x)$ near $x = x_0$ and $y = T_1(f, x_0)$, you will see that the latter is the tangent line equation



ex. For $f(x) = \ln x$,

$$T_1(\ln x, 1) = f(1) + f'(1)(x-1)$$
$$= 0 + 1(x-1)$$
$$y = T_1(\ln x, 1) = x - 1$$

One can also calculate $R_1(\ln x, 1)$ here:

$$R_1(\ln x, 1) = \int_1^x \frac{(x-t)^1}{1!} f''(t)\,dt = \int_1^x (x-t)\left(-\frac{1}{t^2}\right)dt = \int_1^x \left(-\frac{x}{t^2} + \frac{1}{t}\right)dt$$

$$= \left(\frac{x}{t} + \ln|t|\right)\Big|_1^x = \frac{x}{x} + \ln|x| - \frac{x}{1} - \ln|1| =$$
$$= 1 + \ln x - x \quad \text{on } (0, \infty).$$

Here $R_1(\ln x, 1)$ is a measure of how far off $T_1(\ln x, 1)$ as an approximation, and one can see that

$$\ln x = f(x) = T_1(\ln x, 1) + R_1(\ln x, 1)$$
$$= x - 1 + (1 - x + \ln x).$$

- Here, there are two things to say:

(A) Back in Taylor's time, the very existence of a function like $R_k(f, x_0)$ with its properties made for the use of $T_k(f, x_0)$ as an approximation to $f(x)$ at x_0 valid. And

(B) Now, we tend to, instead of calculating $R_k(f, x_0)$ as an integral, which is often hard, we bound $R_k(f, x_0)$ instead, saying that the error in approximation is not too big, and gets smaller when k gets large.

<u>exercise</u>: Compute $T_2(\ln x, 1)$ and $R_2(\ln x, 1)$.

Taylor's Thm holds for real-valued functions on n-variables also, as long as one understands the parts.

Taylor's Thm

Let $f: \mathbb{R}^n \to \mathbb{R}$ have derivatives up to order k at a pt $\vec{x}_0 \in \mathbb{R}^n$. Then there exists a function $R_k(f, \vec{x}_0): \mathbb{R}^n \to \mathbb{R}$, where

1st and 2nd order on pg 160

$$f(\vec{x}) = \underbrace{f(\vec{x}_0) + \sum_{i=1}^{k} \frac{D^{(i)} f(\vec{x}_0)}{i!} (\vec{x}-\vec{x}_0)^i}_{T_k(f, \vec{x}_0)} + R_k(f, \vec{x}_0)$$

and $\displaystyle \lim_{\vec{x} \to \vec{x}_0} \frac{|R_k(f, \vec{x}_0)|}{\|\vec{x}-\vec{x}_0\|^k} = 0$.

Notes ① Hence the kth Taylor Polynomial, now a polynomial of degree k in the n-input variables, is still the "best" degree k polynomial to approximate $f(\vec{x})$ near $\vec{x} = \vec{x}_0$

Notes ② To be able to calculate $T_k(f, \vec{x}_0)$, we need to understand better just what $D^{(i)}f(\vec{x}_0)$ actually is, for $k \geq 1$.

Here $D^{(i)}f(\vec{x}_0)$ is the ith derivative of f at $\vec{x}_0$. What is this?

For a function f which is C^1, we know

$$Df(\vec{x}_0) = D^{(1)}f(\vec{x}_0) = \left[\frac{\partial f}{\partial x_1}(\vec{x}_0) \ \cdots \ \frac{\partial f}{\partial x_n}(\vec{x}_0)\right]$$

is an $1 \times n$ matrix. It is a row of n, real-valued functions on $\mathbb{R}^n$.

One can think, then, that

$$Df(\vec{x}_0) : \underset{\substack{n\text{-input} \\ \text{variables}}}{\mathbb{R}^n} \longrightarrow \underset{\substack{n\text{-output functions given by} \\ \text{the partials.}}}{\mathbb{R}^n}$$

If, then, f is also C^2, then $D(Df) = D^{(2)}f(\vec{x}_0)$ will be an $n \times n$ matrix of 2^{nd} partials of f.

Notes ② cont'd.

So

$$D^{(2)}f(\vec{x}_0) = D(Df)(\vec{x}_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\vec{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\vec{x}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\vec{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\vec{x}_0) \end{bmatrix}$$

Q: So what kind of object would $D^{(3)}f(\vec{x}_0)$ look like?

A: It would actually be a 3-dimensional array comprised of all of the 3rd partial derivatives of f. Don't worry about this. We are not going to need this part!

Note: There is also an integral form for $R_k(f,\vec{x}_0)$ in the multivariable case. We will not derive this either, or use it.

We are in a position now, though, to be able to calculate the 2 basic approximations $T_1(f,\vec{x}_0)$ and $T_2(f,\vec{x}_0)$ for f at $\vec{x}_0$:

Calculation of $T_1(f,\vec{x}_0)$

Let $\vec{x}-\vec{x}_0=\vec{h}=\begin{bmatrix}h_1\\ \vdots \\ h_n\end{bmatrix}$ to reduce the number of variables

Then for $f:\mathbb{R}^n\to\mathbb{R}$ at least C^1 at $\vec{x}_0$, we have

$$f(\vec{x}) = f(\vec{x}_0) + Df(\vec{x}_0)(\vec{x}-\vec{x}_0) + R_1(f,\vec{x}_0)$$

[in coordinate form] $$= f(\vec{x}_0) + \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}(\vec{x}_0)\right)h_i + R_1(f,\vec{x}_0)$$

[in matrix form] $$= \underbrace{f(\vec{x}_0) + \begin{bmatrix}\frac{\partial f}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f}{\partial x_n}(\vec{x}_0)\end{bmatrix}\begin{bmatrix}h_1\\ \vdots \\ h_n\end{bmatrix}}_{\text{This is } T_1(f,\vec{x}_0)} + R_1(f,\vec{x}_0)$$

and when only calculating $T_1(f,\vec{x}_0)$, we can ignore $R_1(f,\vec{x}_0)$.

Special Note: For $f:\mathbb{R}^2\to\mathbb{R}$ differentiable at $\vec{x}_0$, the tangent plane to graph(f) in $\mathbb{R}^3$ is given by

$$z = \underbrace{f(\vec{x}_0) + \frac{\partial f}{\partial x}(\vec{x}_0)(x-x_0) + \frac{\partial f}{\partial y}(\vec{x}_0)(y-y_0)}_{\text{This is } T_1(f,\vec{x}_0),\text{ for } \vec{x}_0=\begin{bmatrix}x_0\\ y_0\end{bmatrix}}$$

Calculation of $T_2(f, \vec{x}_0)$

if f is at least C^2 at $\vec{x}_0$, then by Taylor Thm

$$f(\vec{x}) = \underbrace{f(\vec{x}_0) + Df(\vec{x}_0)(\vec{x}-\vec{x}_0) + \frac{D^{(2)}f(\vec{x}_0)}{2}(\vec{x}-\vec{x}_0)^2}_{} + R_2(f, \vec{x}_0)$$

This is $T_2(f, \vec{x}_0)$, provided we can make sense of the quadratic term

How to make sense of the term $\frac{D^{(2)}f(\vec{x})}{2}(\vec{x}-\vec{x}_0)^2$

- Here $D^{(2)}f(\vec{x}_0)$ is just the $n \times n$ matrix of 2nd partials of f.
- $(\vec{x}-\vec{x}_0)^2$ doesn't make sense as a vector squared.

But the entire term is an example of a ~~function~~ real-valued function on $\mathbb{R}^n$ called a quadratic form: The book calls this function the Hessian form; it looks like either

$$Hf(\vec{x}_0)(\vec{h}) = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0) h_i h_j \ , \text{ or}$$

$$= \frac{1}{2}\vec{h}^T D^{(2)}f(\vec{x}_0)\vec{h} \ .$$

See page 172.

p. 172

So that $T_2(f, \vec{x}_0)$ can be written

$$T_2(f, \vec{x}_0) = f(\vec{x}_0) + Df(\vec{x}_0)\vec{h} + \frac{1}{2}\vec{h}^T D^{(2)} f(\vec{x}_0)\vec{h}$$
$$= f(\vec{x}_0) + Df(\vec{x}_0)\vec{h} + Hf(\vec{x}_0)\vec{h}$$

or
$$T_2(f, \vec{x}_0) = f(\vec{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x}_0) h_i + \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0) h_i h_j$$

Note: There is nothing special about where the $\vec{h}^T$ comes from or why we pull the 2 $\vec{h}$'s apart like this: More, it is just a convenient form to put the calculation in. It is the general form for this type of calculation.

And the graph of $T_2(f, \vec{x}_0)$ is an example of what is called a quadric surface when $n=2$ ~~(quadratic~~ (quadric hypersurface, for $n>2$). Think about the best parabolic bowl to approximate the graph of $f(x,y)$ in $\mathbb{R}^3$ at $\vec{x}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$.

ex. Given $f(x,y) = e^{2x+3y}$, calculate $T_2(f,(0,0))$.

Solution: Calculate and fill in the constituent parts to $T_2(f,(0,0))$ in the previous page.

$$f(0,0) = e^{2(0)+3(0)} = 1 \quad , \quad \vec{h} = \begin{bmatrix} x-0 \\ y-0 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$$

$$Df(0,0) = \left[\frac{\partial f}{\partial x}(0,0), \frac{\partial f}{\partial y}(0,0)\right] = \left[2e^{2x+3y}, 3e^{2x+3y}\right]\Big|_{(0,0)} = [2,3]$$

$$D^{(2)}f(0,0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(0,0) & \frac{\partial^2 f}{\partial y \partial x}(0,0) \\ \frac{\partial^2 f}{\partial x \partial y}(0,0) & \frac{\partial^2 f}{\partial y^2}(0,0) \end{bmatrix} = \begin{bmatrix} 4e^{2x+3y} & 6e^{2x+3y} \\ 6e^{2x+3y} & 9e^{2x+3y} \end{bmatrix}\Bigg|_{(0,0)} = \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}.$$

and

$$T_2(f,(0,0)) = 1 + \underbrace{[2,3]}_{Df(0,0)}\underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\vec{h}} + \frac{1}{2}\underbrace{[x\ y]}_{\vec{h}^T}\underbrace{\begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}}_{D^{(2)}f(0,0)}\underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\vec{h}}$$

$$= 1 + 2x + 3y + \frac{1}{2}[x\ y]\begin{bmatrix} 4x+6y \\ 6x+9y \end{bmatrix}$$

$$= 1 + 2x + 3y + \frac{1}{2}(4x^2 + 6xy + 6xy + 9y^2)$$

$$= 1 + 2x + 3y + 2x^2 + 6xy + \frac{9}{2}y^2.$$