More general chain rule.

Let $\vec{x} = \{x_1, \ldots, x_n\} / R$

Let $f(\vec{x}) = (f_1(\vec{x}), \ldots, f_m(\vec{x})) \in R[\vec{x}]^m$

Define the <u>total derivative</u> (Jacobian)

$$Df = \underbrace{\left.\begin{pmatrix} D_{x_1} f_1 & D_{x_2} f_1 & \cdots & D_{x_n} f_1 \\ \vdots & \vdots & & \vdots \\ D_{x_1} f_m & D_{x_2} f_m & \cdots & D_{x_n} f_m \end{pmatrix}\right\} m}_{n}$$

If $m = 1$, $Df = \nabla f$ is the gradient (row) vector.

Now let $\Phi(\vec{x}) = (\Phi_1(\vec{x}), \ldots, \Phi_n(\vec{x}))$ in $R[\vec{x}]^n$, and define formal composition

$$(f \circ \Phi)(\vec{x}) = ((f_1 \circ \Phi)(\vec{x}), \ldots, (f_m \circ \Phi)(\vec{x}))$$

and

$$(f_i \circ \Phi) := f_i(\Phi_1(\vec{x}), \dots, \Phi_n(\vec{x})).$$

[Think: $f : R^n \to R^m$

$\Phi : R^n \to R^n$ "change of coords"

$f \circ \Phi : R^n \to R^m$

Over infinite domain this is correct.]

Chain Rule:

$$\underset{m \times n}{D(f \circ \Phi)} = \underset{m \times n}{(Df \circ \Phi)} \, \underset{n \times n}{D\Phi}$$

Proof: Same as before, because we can work row by row. ///

Today: Taylor Exansion.

For every $f(\vec{x}) \in R[x]$ & $\vec{a} \in R^n$, there is a unique Taylor expansion of $f(\vec{x})$ "near $\vec{x} = \vec{a}$":

$$f(\vec{x}) = \sum c_I (\vec{x} - \vec{a})^I$$

$$= \sum c_I (x_1 - a_1)^{i_1} (x_2 - a_2)^{i_2} \cdots (x_n - a_n)^{i_n},$$

sum over all index vectors $I \in \mathbb{N}^n$.

We want to compute the coefficients. The answer is expressed in terms of differential operators.

Definition: For an $I = (i_1, \ldots, i_n) \in \mathbb{N}^n$ we define the operator

$$D^I_{\vec{x}} : R[\vec{x}] \to R[\vec{x}]$$

as

$$D^I_{\vec{x}} = D^{i_1}_{x_1} D^{i_2}_{x_2} \cdots D^{i_n}_{x_n}$$

$$= \text{“} \frac{\partial^{i_1 + i_2 + \cdots + i_n}}{\partial x_1^{i_1} \partial x_2^{i_2} \cdots \partial x_n^{i_n}} \text{”},$$

which is well-defined because

mixed partials commute.

Theorem (Taylor Expansion):
For any $f(\vec{x}) \in \mathbb{R}[x_1, \dots, x_n]$ and for any point $\vec{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$, I claim that

$$f(\vec{x}) = \sum_{I \in \mathbb{N}^n} \frac{(D_{\vec{x}}^{I} f)(\vec{a})}{I!} (\vec{x} - \vec{a})^I,$$

where $(\vec{x} - \vec{a})^I := (x_1 - a_1)^{i_1} \cdots (x_n - a_n)^{i_n}$

$I! := i_1! \, i_2! \cdots i_n!$

We observe that the sum is finite because $D_{\vec{x}}^{I} f = 0(\vec{x})$ whenever $i_k > \deg(f)$ for some k.

<u>Proof</u>: Given $\vec{a} \in \mathbb{R}^n$, $f(\vec{x}) \in \mathbb{R}[\vec{x}]$ we consider $g(\vec{x}) := f(\vec{x} + \vec{a}) \in \mathbb{R}[\vec{x}]$.
From the single variable chain rule,

we have $(D_{x_i} g)(\vec{x}) = (D_{x_i} f)(\vec{x}+\vec{a})$

because $D_{x_i}(x_j + a_j) = \begin{cases} 0 & i \neq j, \\ 1 & i = j. \end{cases}$

This implies that

$$(D^I_{\vec{x}} g)(\vec{x}) = (D^I_{\vec{x}} f)(\vec{x}+\vec{a})$$

$$(D^I_{\vec{x}} g)(\vec{0}) = (D^I_{\vec{x}} f)(\vec{a}).$$

Note that $g(\vec{x}) \in R[\vec{x}]$ can be expressed as $g(\vec{x}) = \sum c_I \vec{x}^I$,

as can any polynomial. If we can show that

$$c_I = \frac{(D^I_{\vec{x}} g)(\vec{0})}{I!} = \frac{(D^I_{\vec{x}} f)(\vec{a})}{I!},$$

then we will have

$$f(\vec{x}) = g(\vec{x}-\vec{a}) = \sum \frac{(D^I_{\vec{x}} f)(\vec{a})}{I!} (\vec{x}-\vec{a})^I$$

as desired.

To prove the red formula, we define a partial order on $\mathbb{N}^n$ by

$$\text{“} I \leq J \text{”} \iff i_k \leq j_k \quad \forall k$$

$$\text{“} I < J \text{”} \iff I \leq J \ \& \ I \neq J$$

$$\forall k,\ i_k \leq j_k \ \& \ \exists k,\ i_k \neq j_k$$

Then I claim that $D_{\vec{x}}^{I}$ acts on monomials as follows:

$$D_{\vec{x}}^{I}(\vec{x}^{J}) = \begin{cases} (J!_I)\,\vec{x}^{J-I} & I < J \\ I! & I = J \\ 0 & \text{else} \end{cases}$$

where we define

$$(r)_s = r(r-1)(r-2)\cdots(r-s+1)$$

$$J!_I = (j_1)_{i_1}(j_2)_{i_2}\cdots(j_n)_{i_n}.$$

Indeed, if $I \nleq J$ then we have $i_k > j_k$ for some k. It follows that $D_{x_k}^{i_k}\vec{x}^{J} = 0$ and hence $D_{\vec{x}}^{I}\vec{x}^{J} = 0$.

On the other hand, if $I \leq J$ then we have $i_k \leq j_k$ and hence

$$D_{x_k}^{i_k} \vec{x}^J = j_k(j_k-1)\cdots(j_k-i_k+1)\frac{\vec{x}^J}{x_k^{i_k}}$$

$$= (j_k)_{i_k}\, \vec{x}^J / x_k^{i_k}$$

for any k. It follows that

$$D_{\vec{x}}^I \vec{x}^J = (j_1)_{i_1}(j_2)_{i_2}\cdots(j_n)_{i_n}\frac{\vec{x}^J}{x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}}$$

$$= (J!_I)\, \vec{x}^J / \vec{x}^I$$

$$= (J!_I)\, \vec{x}^{J-I}.$$

In the special case $I = J$ this becomes

$$D_{\vec{x}}^I \vec{x}^I = I!\, \vec{x}^{\vec{0}} = I! \qquad ///$$

By applying $D_{\vec{x}}^I$ to

$$g(\vec{x}) = \sum c_J \vec{x}^J$$

we obtain

$$D^{I}_{\vec{x}}\, g = c_I\, I! + \sum_{J>I} c_J (J!_I)\, \vec{x}^{\,J-I}$$

Evaluate at $\vec{0}$ to get

$$(D^{I}_{\vec{x}}\, g)(\vec{0}) = c_I\, I!$$

$$c_I = \frac{(D^{I}_{\vec{x}}\, g)(\vec{0})}{I!}$$

$\square$

[Shreeram Abhyankar : There are three kinds of algebra:

High School, College, University.

None is harder than the others, they only differ in abstraction.]

The first few terms of the Taylor expansion have special meaning.

$$f(\vec{x}) = f(\vec{a}) + (\nabla f)_{\vec{a}}(\vec{x}-\vec{a})$$

$$+ \frac{1}{2}(\vec{x}-\vec{a})^T (Hf)_{\vec{a}} (\vec{x}-\vec{a})$$

$$+ \text{higher terms},$$

where $(\vec{x}-\vec{a})$ is a column vector,

$$(\nabla f)_{\vec{a}} = \left(D_{x_1} f(\vec{a}), \ldots, D_{x_n} f(\vec{a}) \right)$$

is the gradient row vector, and

Hf is the Hessian matrix of 2nd derivatives:

$$Hf = \begin{pmatrix} D_{x_1} D_{x_1} f & \cdots & D_{x_1} D_{x_n} f \\ \vdots & & \vdots \\ D_{x_n} D_{x_1} f & \cdots & D_{x_n} D_{x_n} f \end{pmatrix}$$