MAT257 Notes

Jad Elkhaleq Ghalayini

October 5 2018

Partial Derivatives

Theorem 1. *If $U \subset \mathbb{R}^m$, $f : U \to \mathbb{R}^n$ is differentiable at a and $f = (f_1, ..., f_n)$ then each partial derivative $\frac{\partial f_i}{\partial x_j}(a)$ exists, and*

$$f'(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & ... & \frac{\partial f_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & ... & \frac{\partial f_n}{\partial x_m}(a) \end{pmatrix}$$

with

$$\frac{\partial f}{\partial x_j}(a) = \frac{\partial}{\partial x} f(a_1, ..., a_{j-1}, x, a_{j+1}, ..., a_m)\bigg|_{x=a_j}$$

Proof.

- Case $n = 1$: Let
$$h(x) = (a_1, ..., a_{j-1}, x, a_{j+1}, ..., a_m) : \mathbb{R} \to \mathbb{R}^m$$
Then
$$\frac{\partial f}{\partial x_j}(a) = D(f \circ h)(a_j) = Df(h(a_j)) \circ Dh(a_j) = f'(a) \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$
where the matrix is $1 \times m$ and 1 in the j^{th} place, i.e. the j^{th} entry in $f'(a)$.

- For n in general,
$$f = (f_1, ..., f_n), f'(a) = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_n'(a) \end{pmatrix}$$
Hence, we have a row for each f_i, and we can apply the case for $n = 1$ to each row.

□

We're going to do some calculations right away with this, since it's very important that this is all straight in your head. One thing we want to look at is the converse of this theorem, which turns out not to be true unless we make additional hypotheses.

But let's fist do some computations, and we'll see that this notion of partial derivatives lets us rethink what the chain rule says, which is very important when you want to use the chain rule.

The Chain Rule

Suppose we have

$$F(x) = f(g_1(x), ..., g_n(x)), x = (x_1, ..., x_m), g = (g_1, ..., g_n)$$

with g differentiable at a and f differentiable at $g(a)$. We want to compute $\frac{\partial F}{\partial x_i}(a)$. Let's say that $f = f(y_1, ..., y_n)$. Then what's the formula?

$$\frac{\partial F}{\partial x_i}(a) = \sum_{j=1}^{n} \frac{\partial f}{\partial y_j}(g(a)) \frac{\partial g_j}{\partial x_i}(a)$$

Why is this? Well, what does the chain rule say?

$$DF(a) = Df(g(a))Dg(a) = \begin{pmatrix} \frac{\partial f}{\partial y_1}(g(a)) & \ldots & \frac{\partial f}{\partial y_n}(g(a)) \end{pmatrix} \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \ldots & \frac{\partial g_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1}(a) & \ldots & \frac{\partial g_n}{\partial x_m}(a) \end{pmatrix}$$

So the result of this is a row matrix, and by taking $\frac{\partial F}{\partial x_i}$ is just indexing into it. You should really know this, you should even memorize it. When you come out of calculus, this should be as natural as addition when you come out of elementary school.

These things, sometimes people write them in a kind of symbolic, simpler way, which is useful. Here we have

$$y_j = g_j(x_1, ..., x_m)$$

and then we're composing to get

$$z = f(y_1, ..., y_n) = f(g(x))$$

So if you want to take the derivative with respect to x_i, you could write that as

$$\frac{\partial z}{\partial x_i}$$

But this is like short form, since z is not a function of x, it's a function of y. So when we write $\frac{\partial z}{\partial x_i}$, we're thinking of it as already composite. And the chain rule thinks of this as

$$\frac{\partial z}{\partial x_i} = \sum_{j=1}^{n} \frac{\partial z}{\partial y_j} \frac{\partial y_j}{\partial x_i}$$

with evaluation, this becomes

$$\frac{\partial z}{\partial x_i}(a) = \sum_{j=1}^{n} \frac{\partial z}{\partial y_j}\bigg|_{y=g(a)} \frac{\partial y_j}{\partial x_i}\bigg|_{x=a}$$

Now how does this notation sometimes show up in a confusing way? Some examples!

1. Suppose $F(t) = f(x(t), y(t), z(t), t)$. Let's call $F(t) = w$. Let's take the derivative with respect to t:

   $$\frac{dw}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} + ?$$

   What do we write in the fourth place? We want the derivative of f with respect to the fourth variable, so we just write $\frac{\partial w}{\partial t}$. So we have

   $$\frac{dw}{dt} \neq \frac{\partial w}{\partial t}$$

   See how this is confusing?

   This is *shorthand*, so there can be ambiguity. That ambiguity is somewhat resolved by the different ordinary versus partial derivative notation, but it's still confusing. That's why you've got to unserstand the full, genuine meaning behind the shorthand notation.

2. Let $w = f(x(s, t), y(s, t), s, t)$. Now it gets even worse...