Lecture 12

niceguy

October 4, 2023

Note: I skipped a lecture on Monday, so there is no lec11.

1 Recap

Recall the derivative of $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is

$$Df(x) = \begin{pmatrix} D_1 f & D_2 f & \dots & D_n f \end{pmatrix} = \frac{\partial f_j}{\partial x_i}$$

where $Df(x)$ uniquely satisfies

$$f(x+h) - f(x) = Df(x) \cdot h + o(h)$$

$f \in \mathcal{C}^r$ means the rth order partial derivatives or anything less exist and are continuous.
The chain rule states that $D(g \circ f)(x) = Dg(f(x)) \times Df(x)$.

2 Applications of Chain Rule

We have proven this before, but this may provide more insight. Note that defining $g(t) = x + tv$, we have

$$\frac{d}{dt} f(x+tv)\Big|_{t=0} = D(f \circ g) = Df(g(0)) \times Dg(0) = Df(x) \frac{d}{dt} g(t)\Big|_{t=0} = Df(x) \cdot \vec{v}$$

Fun fact: just like in single variable calculus, a local maximum or minimum (for a differentiable function) has to have a differential of 0 at its point. However, the opposite is not true. For $f(x,y) = x^2 - y^2$, its derivative at $(0,0)$ vanishes, but it looks like a chip, hence it is not a maximum nor minimum (this holds for any direction apart from $x = y, x = -y$).

If we apply the chain rule to an invertible function, then

$$\begin{aligned} D(g \circ f)(x) &= Dx \\ D(g)(f(x)) \times Df(x) &= I \end{aligned}$$

This implies $Df(x)$ is invertible, and

$$[Df(x)]^{-1} = Dg(f(x))$$

Rearranging,

$$Df^{-1}(y) = \left[Df\left(f^{-1}(y)\right)\right]^{-1}$$