Critical Points, Implicit Function Theorem

Friday, February 26, 2021 8:16 AM

Last Lecture

Single-variable

say f is def'd on an open domain $D \subseteq \mathbb{R}$, $x_0 \in D$

Two definitions of derivatives:

① $f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h}$

diff'able at x_0 iff this limit exists

② We say L is the derivative of f at x_0 if

$$0 = \lim_{h \to 0} \frac{|f(x_0+h) - f(x_0) - hL|}{|h|}$$

We say f is diff'able at x_0 iff such an $L \in \mathbb{R}$

Note: if such an L exists, it's unique.

$f'(x_0) := L$ if L exists

Generalization to multiple dimensions

say $f: D \to \mathbb{R}^p$ with $D \subseteq \mathbb{R}^n$ an open domain and $\vec{x}_0 \in D$

We say f is diff'able at $\vec{x}_0$ if there's a linear map $L: \mathbb{R}^n \to \mathbb{R}^p$ s.t.

$$\lim_{\vec{h} \to 0} \frac{\| f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - L(\vec{h}) \|}{\|\vec{h}\|} = 0$$

Recall: $\vec{h}, \vec{x}_0 \in \mathbb{R}^n$

$f(\underbrace{-}_{\text{anything}}), L(\vec{h}) \in \mathbb{R}^p$

Q/ How do we compute L?

Q/ eg. given

$$f(x,y) = \left(3\cos(xy) - ye^x,\ xy^2 + \frac{y}{x^2+1}\right)$$

$$\mathbb{R}^2 \to \mathbb{R}^2$$

what 2×2 matrix L?

A/ Partial derivatives!

Recall: $\lim_{h \to 0}$ means $\vec{h}$ can approach 0 from any direction and the limit should be the same.

Let's assume f is diff'able and compute the lim from a particular direction and $L = f'(\vec{x}_0)$

Say $x_1, \dots, x_n$ are coords on $\mathbb{R}^n$

Consider: $\vec{h} = (h, 0, 0, \dots, 0) \in \mathbb{R}^n$ with $h \in \mathbb{R}$

ie, $\vec{h}$ approaches 0 along x_1-axis

So $\|\vec{h}\| = h$

write: $\vec{x}_0 = (s_1, s_2, \dots, s_n)$ and

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - L(\vec{h}) = f(s_1+h, s_2, \dots, s_n) - f(s_1, s_2, \dots, s_n) - L(h\vec{e}_1)$$

because $\vec{h} = h\vec{e}_1$, $\vec{e}_1 = (1, 0, \dots, 0) \in \mathbb{R}^n$

$$= f(s_1+h, s_2, \dots, s_n) - f(s_1, \dots, s_n) - hL(\vec{e}_1)$$

because L is linear

Recall saying that $L = f'(\vec{x}_0)$ is, by definition, say that:

$$0 = \lim_{\vec{h} \to 0} \frac{\| f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - L(\vec{h}) \|}{\|\vec{h}\|}$$

$$= \lim_{\vec{h} \to 0} \frac{\| f(s_1 + h, s_2, \dots, s_n) - f(s_1, s_2, \dots, s_n) - hL(\vec{e}_1) \|}{h}$$

use $\|a\vec{v}\| = a\|\vec{v}\|$

looks like partial derivative

$$= \lim_{\vec{h} \to 0} \left\| \frac{f(s_1 + h, s_2, \dots, s_n) - f(s_1, s_2, \dots, s_n)}{h} - L(\vec{e}_1) \right\|$$

first column of L matrix

so, this limit is 0

In other words, as $h \to 0$, $\frac{f(s_1 + h, s_2, \dots, s_n)}{h}$ approaches $L(\vec{e}_1)$

$$\Rightarrow L(\vec{e}_1) = \lim_{h \to 0} \frac{f(s_1 + h, s_2, \dots, s_n) - f(s_1, \dots, s_n)}{h}$$

$$= \frac{\partial f}{\partial x_1}$$

Q/ We talked about partial derivatives of fcns w/ multiple inputs but one output. What does $\frac{\partial f}{\partial x_1}$ mean if $f : D \to \mathbb{R}^p$ and $p > 1$?

A/ Apply $\frac{\partial}{\partial x_1}$ to each of the p components

Explicitly: If

$f(x_1, \dots, x_n) = (u_1(x_1, \dots, x_n), u_2(x_1, \dots, x_n), \dots, u_p(x_1, \dots, x_n))$

eg. a fcn $f: \mathbb{R}^3 \to \mathbb{R}^2$ is the same as pair of fcns u_1 and u_2, each from $\mathbb{R}^3$ to $\mathbb{R}^1$

then

$$\frac{\partial f}{\partial x_1} = \begin{bmatrix} \partial u_1 / \partial x_1 \\ \partial u_2 / \partial x_1 \\ \vdots \\ \partial u_p / \partial x_1 \end{bmatrix}$$

Conclusion

If f is diff'able at $\vec{x}_0$ and $L = f'(\vec{x}_0)$ then

$$L(\vec{e}_1) = 1^{st} \text{ column of } L$$

$$= \frac{\partial f}{\partial x_1}$$

In general, let $1 \le j \le n$, $L(e_j) = j^{th}$ column of L

$$= \frac{\partial f}{\partial x_j}(\vec{x}_0)$$

∂x_j
partial derivative at $\vec{x}_0$

So

$$L = \begin{bmatrix} \partial u_1/\partial x_1 & \partial u_1/\partial x_2 & \cdots & \partial u_1/\partial x_n \\ \partial u_2/\partial x_1 & \partial u_2/\partial x_2 & \cdots & \partial u_2/\partial x_n \\ \vdots & \vdots & & \vdots \\ \partial u_p/\partial x_1 & \partial u_p/\partial x_2 & & \partial u_p/\partial x_n \end{bmatrix}$$

"total derivative" $\quad F'(\vec{x}_0)$

Caveat: this says that if L exists, then L is the linear map $\mathbb{R}^n$ to $\mathbb{R}^p$ given by the matrix of partials

Q/ How to show F is diff'able?

A/ see Apostol on BCourses for proof

Thm
If $F: D \to \mathbb{R}^p$ and $D \overset{\text{open}}{\subseteq} \mathbb{R}^n$ and if $\forall\ 1 \le j \le n$,

$\frac{\partial F}{\partial x_j}(\vec{x})$ exists and is continuous for all $\vec{x} \in D$, then F is diff'able at all $\vec{x} \in D$

and then $F'(\vec{x})$ is represented by a matrix whose j^{th} column is $\frac{\partial F}{\partial x_j}(\vec{x})$

Notice: the i^{th} row of $F'(\vec{x})$ is $\nabla u_i(\vec{x})$ viewed as row vector where $F = (u_1, u_2, \ldots, u_p)$

FACT: ∇u_i is a vector of length n

so, the j^{th} column corresponds to x_j and the i^{th} row corresponds to u_i

Example Corollary

$F(x,y) = \left(3\cos(xy) - ye^x,\ xy^2 + \frac{y}{x^2+1}\right)$ is diff'able at all $\vec{x} = (x,y) \in \mathbb{R}^2$

PROOF

$$\frac{\partial F}{\partial x} = \begin{bmatrix} -3y\sin(xy) - ye^x \\ y^2 - \frac{2xy}{(x^2+1)^2} \end{bmatrix}$$

$$\frac{\partial F}{\partial y} = \begin{bmatrix} -3x\sin(xy) - e^x \\ 2xy + \frac{1}{x^2+1} \end{bmatrix}$$

Notice both $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$ are defined and continuous for all $(x,y) \in \mathbb{R}^2$ so, by the Thm, F is diff'able on all of $\mathbb{R}^2$. ∎ QED

Note: if $f : D \to \mathbb{R}^1$ is diff'able at $\vec{x}_0$, then $D_{\vec{u}} f(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot \vec{u}$

This follows from the Chain Rule:

Suppose $f : D \to \mathbb{R}^p$, $D \subseteq \mathbb{R}^n$, $g : E \to \mathbb{R}^q$, $E \subseteq \mathbb{R}^p$

suppose $\vec{x}_0 \in D$, $f(\vec{x}_0) \in D$

f is diff'able at $\vec{x}_0$ and g is diff'able at $f(\vec{x}_0)$

Then

$$\underbrace{(g \circ f)'(\vec{x}_0)}_{\substack{\text{linear map} \\ \text{from} \\ \mathbb{R}^n \text{ to } \mathbb{R}^q}} = \underbrace{g'(f(\vec{x}_0))}_{\substack{\text{linear map} \\ \text{from} \\ \mathbb{R}^p \text{ to } \mathbb{R}^q}} \circ \underbrace{f'(\vec{x}_0)}_{\substack{\text{linear} \\ \text{map from} \\ \mathbb{R}^n \text{ to } \mathbb{R}^p}}$$

Concretely — this says you can compute partials of $g \circ f$ in terms of partials of g and f using matrix mult.