Again rather than define derivative as the limit of difference quotients

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

Instead as a good linear approximation to F.

$$F: S \to \mathbb{R}^m \quad \text{where} \quad S \subseteq \mathbb{R}^n \text{ and } S \text{ open.}$$

Then F differentiable at $a \in S$ means there exists $L \in \mathbb{R}^{m \times n}$ that is a linear function $L: \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\frac{\left| F(x+h) - \overbrace{\left(F(x) + Lh\right)}^{\text{linear approx}} \right|}{|h|} \to 0 \quad \text{as } h \to 0.$$

Notation

If F is differentiable at a we denote L by $DF(a)$

To make such notation, need to know L is unique...

Suppose there were $L' \in \mathbb{R}^{m \times n}$ such that $L \neq L'$. Then there is $u \in \mathbb{R}^n$ such that $Lu \neq L'u$.

Let $h = tu$. Then

$$Lh - L'h = \underbrace{Lh + F(x) - F(x+h) - F(x) + F(x+h)}_{\text{add and subtract}} - L'h$$

$$0 \leq \frac{|Lh - L'h|}{|h|} \leq \frac{|Lh + F(x) - F(x+h)|}{|h|} + \frac{|-F(x) + F(x+h) - L'h|}{|h|}$$

$$\to 0 + 0 = 0 \quad \text{as } h \to 0$$

$h = tu$.

$$\frac{|Lh - L'h|}{|h|} = \frac{|Ltu - L'tu|}{|tu|} \overset{\text{linearity}}{=} \frac{|t||Lu - L'u|}{|t||u|} \neq 0$$

if $L \neq L'$

which contradicts it getting squeezed between 0 and 0. Thus L is unique.

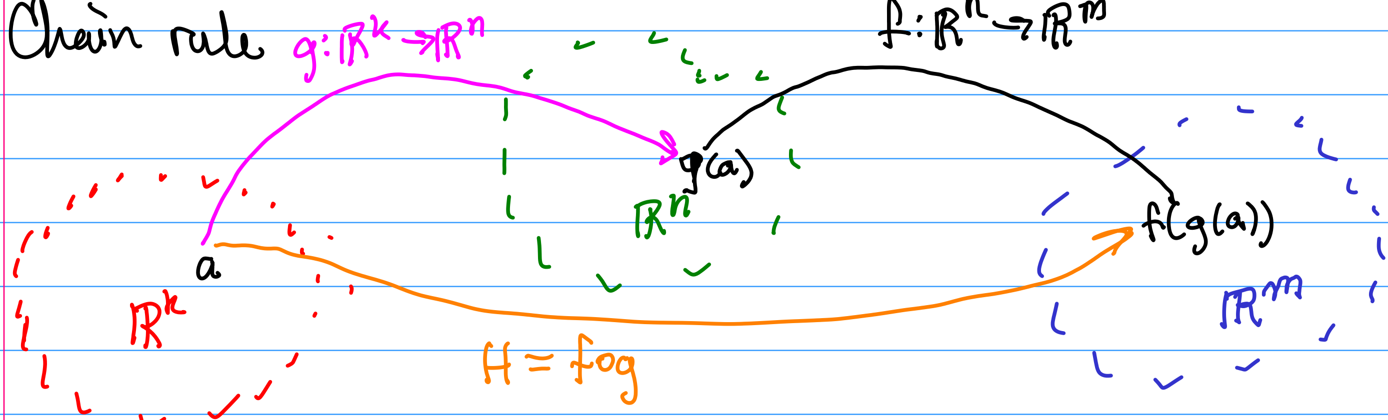
$$\mathbf{f}(\mathbf{x}) = \left(f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x})\right).$$

Remark that if f is differentiable, then each f_i is differentiable. So ∇f_i exists and equal the vector of partial derivatives of f_i. Thus

$$Df = L = \begin{bmatrix} \nabla f_1 \\ \nabla f_2 \\ \vdots \\ \nabla f_n \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

If the partials exist and are continuous then F is differentiable.

Chain rule



assume g is differentiable at a
f is differentiable at $g(a)$

Then $H = f \circ g$ is differentiable at a and

$$DH(a) = Df(g(a))\, Dg(a)$$

2.86 Theorem (Chain Rule III). *Suppose* $\mathbf{g} : \mathbb{R}^k \to \mathbb{R}^n$ *is differentiable at* $\mathbf{a} \in \mathbb{R}^k$ *and* $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ *is differentiable at* $\mathbf{g}(\mathbf{a}) \in \mathbb{R}^n$. *Then* $\mathbf{H} = \mathbf{f} \circ \mathbf{g} : \mathbb{R}^k \to \mathbb{R}^m$ *is differentiable at* $\mathbf{a}$, *and*

$$D\mathbf{H}(\mathbf{a}) = D\mathbf{f}(\mathbf{g}(\mathbf{a}))D\mathbf{g}(\mathbf{a}),$$

where the expression on the right is the product of the matrices $D\mathbf{f}(\mathbf{g}(\mathbf{a}))$ *and* $D\mathbf{g}(\mathbf{a})$.

Chapter 3

THE IMPLICIT FUNCTION THEOREM AND ITS APPLICATIONS

Let $F : \mathbb{R}^{n+1} \to \mathbb{R}$

$F(x, y)$ where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$.

Suppose $F(a, b) = 0$ for $(a, b) \in \mathbb{R}^n \times \mathbb{R}$.

We want to find a function $f : B \to \mathbb{R}$ such that

$$F(x, f(x)) = 0 \quad \text{for all } x \in B$$

Equivalently, I want a set $\mathcal{U} \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\left(F(x,y) = 0 \quad \text{if and only if} \quad y = f(x) \right) \quad \text{for } (x,y) \in \mathcal{U}.$$

Intuition: Linear case ... Let $F : \mathbb{R}^{n+1} \to \mathbb{R}$

$$F(x,y) = \overbrace{\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n + \beta y}^{L(x,y)} + c = 0$$

Scalars

Solve for y in terms of x ...

$$y = \frac{-c - \alpha_1 x_1 - \alpha_2 x_2 - \cdots - \alpha_n x_n}{\beta}$$

need $\beta \neq 0$ for this to work...

3.1 Theorem (The Implicit Function Theorem for a Single Equation). *Let $F(\mathbf{x}, y)$ be a function of class C^1 on some neighborhood of a point $(\mathbf{a}, b) \in \mathbb{R}^{n+1}$. Suppose that $F(\mathbf{a}, b) = 0$ and $\partial_y F(\mathbf{a}, b) \neq 0$. Then there exist positive numbers r_0, r_1 such that the following conclusions are valid.*

f exists
a. For each $\mathbf{x}$ in the ball $|\mathbf{x} - \mathbf{a}| < r_0$ there is a unique y such that $|y - b| < r_1$ and $F(\mathbf{x}, y) = 0$. We denote this y by $f(\mathbf{x})$; in particular, $f(\mathbf{a}) = b$.

f is C^1
b. The function f thus defined for $|\mathbf{x} - \mathbf{a}| < r_0$ is of class C^1, and its partial derivatives are given by

$$\partial_j f(\mathbf{x}) = -\frac{\partial_j F(\mathbf{x}, f(\mathbf{x}))}{\partial_y F(\mathbf{x}, f(\mathbf{x}))}. \tag{3.2}$$

(a) Suppose $\partial_y F(\mathbf{a}, b) \neq 0$. Assume $\partial_y F(a,b) > 0$ if not then work with $-F$ instead.

Since $F \in C^1$ then $\partial_y F$ is continuous. Therefore there is a neighborhood where $\partial_y F(x,y) > 0$ near (a,b).

For example, there is $r_1 > 0$ such that

$$(x,y) \in B(r_1, a) \times (b - r_1, b + r_1)$$

implies $\partial_y F(x,y) > 0$. Thus $y \to F(x,y)$ is a strictly increasing function of y holding x fixed.

$$F(a,b) = 0 \qquad \text{so} \qquad F(a, b+r_1) > 0$$
$$\text{and} \qquad F(a, b-r_1) < 0.$$

Since F is cont and $F(a, b+r_1) > 0$, $F(a, b-r_1) < 0$. then there is $r_0 > 0$ such that $x \in B(r_0, a)$ implies

$$F(x, b+r_1) > 0, \quad F(x, b-r_1) < 0$$

Thus for each $x \in B(r_0, a)$ there is a $y \in (b-r_1, b+r_1)$ such that $F(x,y) = 0$. y is unique because the function $y \to F(x,y)$ is strictly increasing. So denote y by $f(x)$.

Thus for each $x \in B(r_0, a)$ we have $F(x, f(x)) = 0$.

Therefore f exists. $\boxed{\checkmark}$

(b) Claim f is continuous at a.

For every $\varepsilon > 0$ there is $\delta > 0$ such that $|x - a| < \delta$ implies $|f(x) - f(a)| < \varepsilon$.

Just take $r_1 = \varepsilon$ above and r_0 is then the δ.