# EECS16B DIS11B remote

learning objectives

1. Writing linearizations of a function at an operating point for functions of one variable
2. Extending linearizations to functions of multiple variables

---

Linearization : Coming up with a linear model for a nonlinear system that is "true" for states close to some state (operating point)

Q: Why do we want to look at a nonlinear system as if it were linear?

A: Use all of our tools we have developed up until now!

Control system

We've seen

$$\left.\begin{aligned} \vec{x}[i+1] &= A\vec{x}[i] + B\vec{u}[i] \\ \frac{d\vec{x}}{dt}(t) &= A\vec{x}(t) + B\vec{u}(t) \end{aligned}\right\} \begin{aligned} &\text{discrete time (linear system)} \\ &\text{(continuous time linear system)} \end{aligned}$$

Most systems in real life are nonlinear

$$\rightarrow \vec{x}[i+1] = f(\vec{x}[i], \vec{u}[i])$$

Ex:

$$x_1[i+1] = x_1^2[i] + x_2^2[i] \leftarrow \text{nonlinear function of } \vec{x}[i] = \begin{bmatrix} x_1[i] \\ x_2[i] \end{bmatrix}$$

$$x_2[i+1] = x_2[i]$$

EECS 16B Fall 2021 — Designing Information Devices and Systems II — Discussion Worksheet

# Discussion 11B

The following notes are useful for this discussion: Note 19

## 1. Linear Approximation

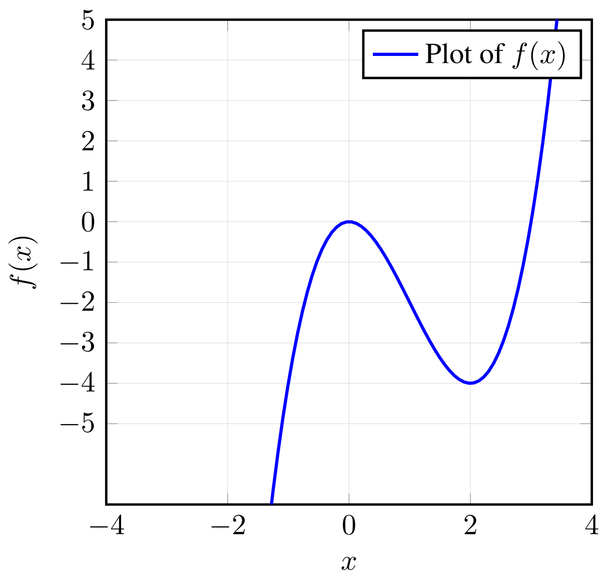
A common way to approximate a nonlinear function is to perform linearization near a point. In the case of a one-dimensional function $f(x)$, the linear approximation of $f(x)$ at a point $x_\star$ is given by

$$f(x) \approx f(x_\star) + f'(x_\star) \cdot (x - x_\star), \tag{1}$$

where $f'(x_\star) := \frac{\mathrm{d}f}{\mathrm{d}x}(x_\star)$ is the derivative of $f(x)$ at $x = x_\star$.

Keep in mind that wherever we see $x_\star$, this denotes a *constant value* or operating point.

(a) Suppose we have the single-variable function $f(x) = \boxed{x^3 - 3x^2}$. We can plot the function $f(x)$ as follows:




To linearize a function of one variable
Compute $f(x_\star)$
$f'(x_\star)$

i. **Write the linear approximation of the function around an arbitrary point $x_\star$.**

$f(x_\star) = x_\star^3 - 3x_\star^2$

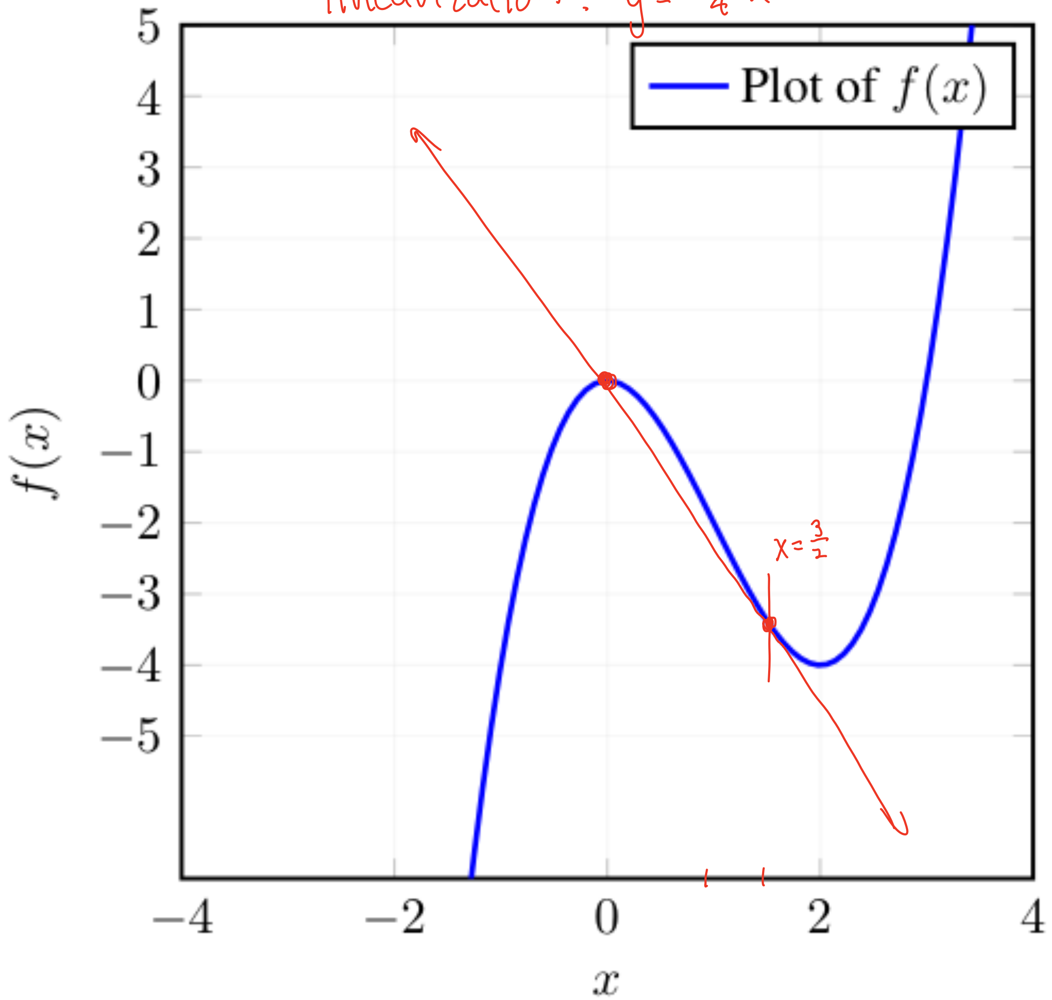
$f'(x) = \frac{d}{dx}(\underbrace{x^3 - 3x^2}_{f(x)}) = 3x^2 - 6x \qquad f'(x_\star) = 3x_\star^2 - 6x_\star$

$f(x) \approx \boxed{\underbrace{(x_\star^3 - 3x_\star^2)}_{f(x_\star)} + (3x_\star^2 - 6x_\star)(x - x_\star)}$

↳ the linearization of $f(x)$ about $x_\star$

linearization : $y = -\frac{9}{4}x$




ii. Using the expression above, **linearize the function around the point $x = 1.5$. Draw the linearization into the plot of part i).**

$$f(x) \approx (x_*^3 - 3x_*^2) + (3x_*^2 - 6x_*)(x - x_*)$$

$$x_* = 1.5 = \frac{3}{2}$$

$$f(x) \approx \left(\left(\tfrac{3}{2}\right)^3 - 3\left(\tfrac{3}{2}\right)^2\right) + \left(3\left(\tfrac{3}{2}\right)^2 - 6\left(\tfrac{3}{2}\right)\right)\left(x - \tfrac{3}{2}\right) \qquad f(x_*) = -\frac{27}{8}$$

$$= \frac{27}{8} - \frac{27}{4} + \left(\frac{27}{4} - \frac{18}{2}\right)\left(x - \frac{3}{2}\right)$$

$$= -\frac{27}{8} + \left(\frac{27-36}{4}\right)\left(x - \frac{3}{2}\right)$$

$$= -\frac{27}{8} + \left(-\frac{9}{4}\right)\left(x - \frac{3}{2}\right)$$

$$= -\frac{27}{8} - \frac{9}{4}x + \frac{27}{8}$$

$$= \boxed{-\frac{9}{4}x}$$

ii. Using the expression above, **linearize the function around the point $x = 1.5$. Draw the linearization into the plot of part i).**

Done above.

Now that we have this specific point's linearization, we understand how the function behaves around the point. Let's use this linearization to evaluate the function's approximation at $x = 1.7$ (based on our approximation at $x = 1.5$, we want to see how a $\delta = +0.2$ shift in the $x$ value changes the corresponding $f(x)$ value). How does this approximation compare to the exact value of the function at $x = 1.7$?

$$f(1.7) \approx -3.375 + (-2.25) \cdot (1.7 - 1.5) \tag{2}$$
$$\approx -3.375 - 0.45 \tag{3}$$
$$\approx \underline{-3.825} \tag{4}$$

close because 1.7 is close to 1.5 (operating point)

linearization point

Comparing to the exact value $f(1.7) = 1.7^3 - 3 \cdot 1.7^2 = \underline{-3.757}$, we find that the difference is $0.068$. Not too bad! What if we repeat with $\delta = 1$? To do so, we must use the approximation around $x = 1.5$ to compute $x = 2.5$, and compare to the exact value $f(2.5)$. How does our new approximation compare to the exact result?

$$f(2.5) \approx -3.375 + (-2.25) \cdot (2.5 - 1.5) \tag{5}$$
$$\approx -3.375 - 2.25 \tag{6}$$
$$\approx \underline{-5.625} \tag{7}$$

very different

Comparing to the exact value $f(2.5) = 2.5^3 - 3 \cdot 2.5^2 = \underline{-3.125}$, we find that the difference is much larger; the error jumped to 2.5! This is an error multiplication of $\frac{2.5}{0.068} \approx 37$, even though our $\delta$ only multiplied by 5. What happened?

Looking at the actual function, we see that the function has a significant curvature between our "anchor point" of $x_\star = 1.5$ and $x = 2.5$. Our linear model is unable to capture this curvature, and so we estimated $f(2.5)$ as if the function kept decreasing, as it did around $x = 1.5$ (where the slope was $-2.25$).

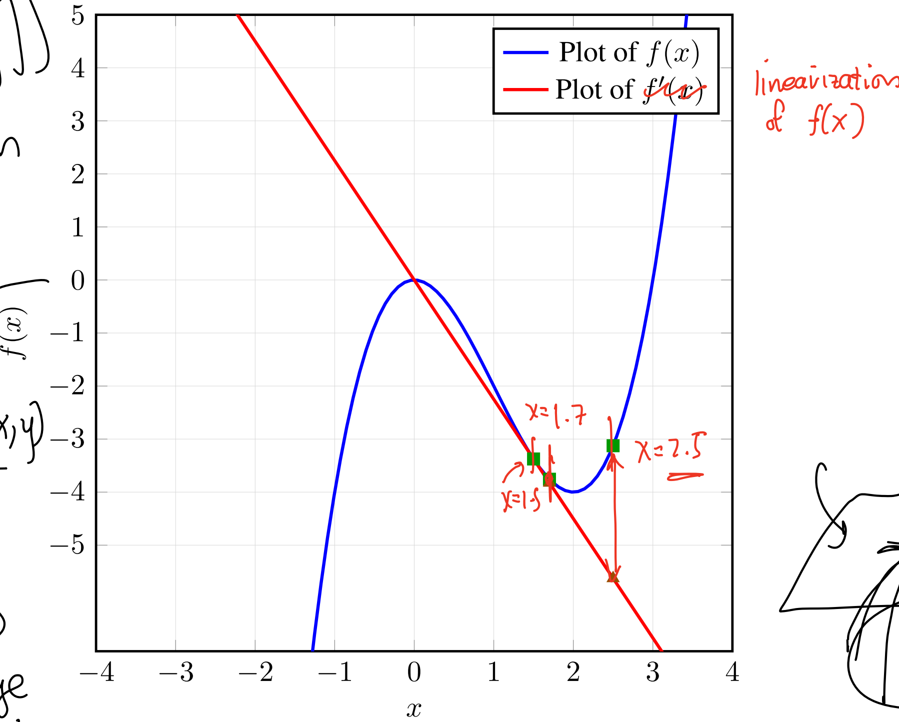
$$\begin{bmatrix}a_1 & a_2\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} = f\left(\begin{bmatrix}x\\ y\end{bmatrix}\right) = a_1x + a_2y$$

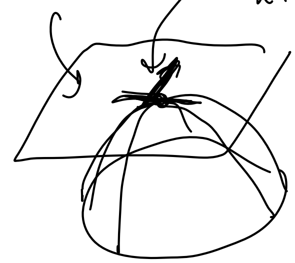
linear function of $\begin{bmatrix}x\\ y\end{bmatrix}$ (affine)

$f(x,y)$

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

hold $y$ constant in $f(x,y)$, then look @ the rate of change in the $x$ coordinate




Now, we can extend this to higher dimensional functions. In the case of a two-dimensional function $f(x, y)$, the linear approximation of $f(x, y)$ at a point $(x_\star, y_\star)$ is given by

small changes in y

$$f(x,y) \approx f(x_\star, y_\star) + \underbrace{\frac{\partial f}{\partial x}(x_\star, y_\star) \cdot (x - x_\star)}_{\text{account for small changes in x}} + \frac{\partial f}{\partial y}(x_\star, y_\star) \cdot (y - y_\star). \qquad (8)$$

where $\frac{\partial f}{\partial x}(x_\star, y_\star)$ is the partial derivative of $f(x, y)$ with respect to $x$ at the point $(x_\star, y_\star)$, and similarly for $\frac{\partial f}{\partial y}(x_\star, y_\star)$

(b) Now, let's see how we can find partial derivatives. When we are given a function $f(x, y)$, we calculate the partial derivative of $f$ with respect to $x$ by fixing $y$ and taking the derivative with respect to $x$. **Given the function $f(x, y) = x^2y$, find the partial derivatives $\frac{\partial f(x,y)}{\partial x}$ and $\frac{\partial f(x,y)}{\partial y}$.**

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x^2y) = y\frac{\partial}{\partial x}(x^2) = 2xy, \quad \frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(x^2y) = x^2\frac{\partial}{\partial y}y = x^2(1) = x^2$$

operating point

(c) **Write out the linear approximation of $f$ near $(x_\star, y_\star)$.**

$$f(x,y) \approx f(x_\star, y_\star) + 2x_\star y_\star(x - x_\star) + x_\star^2(y - y_\star)$$

(d) We want to see if the approximation arising from linearization of this function is reasonable for a point close to our point of evaluation. **First, approximate $f(x, y)$ at the point $(2.01, 3.01)$ using $(x_\star, y_\star) = (2, 3)$. Next, compare the result to $f(2.01, 3.01)$.**

$f(x,y) = x^2 y \qquad f(2,3) = 4 \cdot 3 = 12$

$f(x,y) \approx 12 + 2(2)(3)(x-2) + (2)^2(y-3)$

$= \cancel{12} + 12x - 24 + 4y - \cancel{12}$

$\frac{1}{100} = h \qquad = 12x + 4y - 24$

compare

$f(2.01, 3.01) = (2.01)^2(3.01)$

$= (2+h)^2(3+h)$

$= (4 + 4h + h^2)(3+h)$

$= 12 + 12h + 3h^2 + 4h + 4h^2 + h^3$

$= \underline{12 + 16h} + \underbrace{7h^2 + h^3}$

$12(2.01) + 4(3.01) - 24$

$12(2+h) + 4(3+h) - 24$

$\cancel{24} + 12h + 12 + 4h - \cancel{24}$

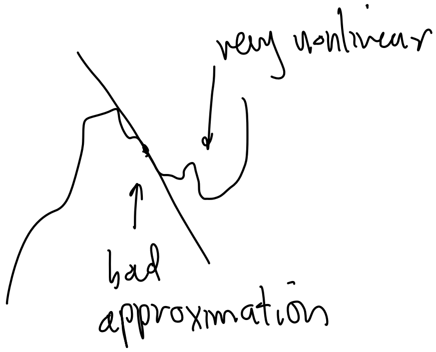
$\underline{16h + 12}$

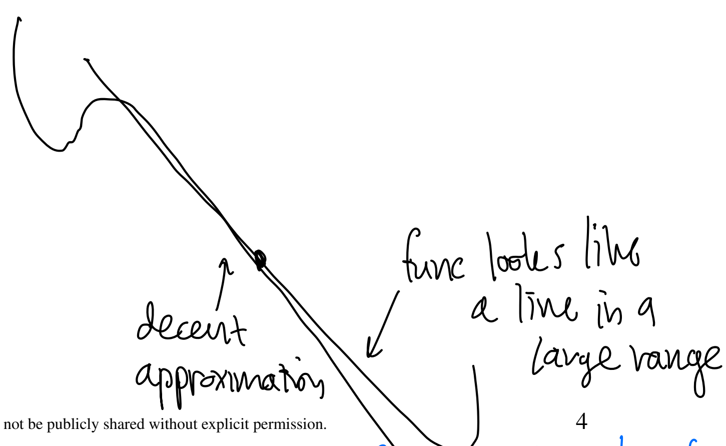
If $h$ small, how does $h^2$ & $h^3$ compare to $h$?

$h = \frac{1}{100} \qquad h^2 = \frac{1}{100^2} = \frac{1}{10^4} \qquad h^3 = \frac{1}{100^3} = \frac{1}{10^6}$

$\left(h = \frac{1}{100}\right) \quad 12.16|0701$

$12.16$






My point here is just "do linearizations only work for small values away from the operating point? A: No depends on what small means"

Note: inner product operates on two column vectors

(e) We will now define the notion of a derivative as a function, and take a look at one possible representation of that function.

**Given the representation of the derivative as a row-vector, describe a function that can take this representation, along with some column vector defining a change in direction, to return a scalar value (which is the change in the real-valued output).** *Don't worry if this seems abstract for now, the next subpart will clarify.*

row ← $\vec{d}^{T}$ Derivative → row vector

column ← $\vec{v}$ direction → column vector

How to combine these two into a single number? (Hope: that this # describes a rate of change along the direction)

$\boxed{\text{matrix product}} \to \boxed{\vec{d}^{T}\vec{v}} \to$ scalar

this is the answer

(f) Suppose we have now a scalar-valued function $f(\vec{x},\vec{y})$, which takes in vector-valued arguments $\vec{x} \in \mathbb{R}^n$, $\vec{y} \in \mathbb{R}^k$ and outputs a scalar $\in \mathbb{R}$. That is, $f(\vec{x},\vec{y})$ is $\mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$. For this new model involving a ~~vector~~-valued function, how can we adapt our previous linearization method?

scalar

One way to linearize the function $f$ is to do it for every single element in $\vec{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^{\top}$ and $\vec{y} = \begin{bmatrix} y_1 & y_2 & \dots & y_k \end{bmatrix}^{\top}$. Then, when we are looking at $x_i$ or $y_j$, we fix everything else as constant. This would give us the linear approximation

$f(\vec{x},\vec{y})$ is a scalar (output) but has vector inputs

$$f(\vec{x},\vec{y}) \approx f(\vec{x}_\star,\vec{y}_\star) + \sum_{i=1}^{n} \frac{\partial f(\vec{x},\vec{y})}{\partial x_i}(x_i - x_{i,\star}) + \sum_{j=1}^{k} \frac{\partial f(\vec{x},\vec{y})}{\partial y_j}(y_j - y_{j,\star}). \tag{9}$$

In order to simplify this equation, we can define the rows $D_{\vec{x}}$ and $D_{\vec{y}}$ as

Derivative of f with respect to $\vec{x}$

$$D_{\vec{x}} f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}, \tag{10}$$

Derivative of f with respect to $\vec{y}$

$$D_{\vec{y}} f = \begin{bmatrix} \boxed{\frac{\partial f}{\partial y_1}} & \cdots & \boxed{\frac{\partial f}{\partial y_k}} \end{bmatrix}. \tag{11}$$

partial derivative of f with respect to $y_k$

Then, eq. (9) can be rewritten as

$$f(\vec{x},\vec{y}) \approx f(\vec{x}_\star,\vec{y}_\star) + (D_{\vec{x}} f)\Big|_{(\vec{x}_\star,\vec{y}_\star)} \cdot (\vec{x} - \vec{x}_\star) + (D_{\vec{y}} f)\Big|_{(\vec{x}_\star,\vec{y}_\star)} \cdot (\vec{y} - \vec{y}_\star). \tag{12}$$

or $\vec{y}$ and $\vec{x}$ both in $\mathbb{R}^n$

**Assume that** $n = k$ **and we define the function** $f(\vec{x},\vec{y}) = \vec{x}^{\top}\vec{y} = \sum_{i=1}^{k} x_i y_i$**. Find** $D_{\vec{x}} f$ **and** $D_{\vec{y}} f$**.**

**[Practice] Next, suppose** $g(\vec{x},\vec{y}) = x_1 x_2^2 y_1 + x_1 y_2^3 + x_2 x_1 y_2 y_1 + \frac{x_1^2}{x_2^3 y_1}$**. Find** $D_{\vec{x}} g$ **and** $D_{\vec{y}} g$

*Hint: it can help to look at eq. (8), and match the terms in eq. (9) to that formulation.*

$$\vec{x} - \vec{x}_\star = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} x_{1\star} \\ x_{2\star} \\ \vdots \\ x_{n\star} \end{bmatrix}$$

column in $\mathbb{R}^n$

$$f(\vec{x},\vec{y}) = \vec{x}^{T}\vec{y} = \sum_{i=1}^{n} x_i y_i \qquad \frac{\partial f}{\partial x_j} = y_j \qquad \frac{\partial f}{\partial y_j} = x_j$$

$$D_{\vec{x}} f = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = \vec{y}^{T}$$

$$D_{\vec{y}} f = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} = \vec{x}^{T}$$

Derivatives are row vectors

we contrast this to columns of partial derivatives

(g) Following the above part, **find the linear approximation of $f(\vec{x}, \vec{y})$ near $\vec{x}_\star = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{y}_\star = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.**

**Recall that** $f(\vec{x}, \vec{y}) = \vec{x}^\top \vec{y} = \sum_{i=1}^{k} x_i y_i$. $= x_1 y_1 + x_2 y_2$

$D_{\vec{x}} f = [y_1 \ y_2] = \vec{y}^T \qquad D_{\vec{x}} f(\vec{x}_\star, \vec{y}_\star) = [-1 \ 2]$

$D_{\vec{y}} f = [x_1 \ x_2] = \vec{x}^T \qquad D_{\vec{y}} f(\vec{x}_\star, \vec{y}_\star) = [1 \ 2]$

$f(\vec{x}, \vec{y}) \approx f(\vec{x}_\star, \vec{y}_\star) + [-1 \ 2](\vec{x} - \vec{x}_\star) + [1 \ 2](\vec{y} - \vec{y}_\star)$

$\underline{\vec{x}_\star^T \vec{y}_\star = 1(-1)+2(2)} = 3$

$= 3 + (-1)(x_1 - 1) + 2(x_2 - 2) + (y_1 + 1) + 2(y_2 - 2)$

function that has a scalar value

(linearization of $f(\vec{x}, \vec{y})$ at $(\vec{x}_\star, \vec{y}_\star)$)

These linearizations are important for us because we can do many easy computations using linear functions.

Something that might help clarify for $f(\vec{x}, \vec{y}) = \vec{x}^T \vec{y}$

$f(\vec{x}, \vec{y}) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$

$\frac{\partial}{\partial x_2} f(\vec{x}, \vec{y}) = \cancel{\frac{\partial}{\partial x_2}(x_1 y_1)}^{0} + \frac{\partial}{\partial x_2}(x_2 y_2) + \cancel{\cdots}^{0} + \cancel{\frac{\partial}{\partial x_2}(x_n y_n)}^{0}$

no $x_2$! no $x_2$!

$= 0 + y_2 + \underbrace{0 + \cdots + 0}_{n-2 \text{ zeroes from } x_3 y_3, \ldots, x_n y_n}$

$= y_2$

**Contributors:**

- Neelesh Ramachandran.
- Kuan-Yun Lee.

Practice: $g(\vec{x}, \vec{y}) = x_1 x_2^2 y_1 + x_1 y_2^3 + x_2 x_1 y_2 y_1 + \frac{x_1^2}{x_2^3 y_1}$

$$D_{\vec{x}} g = \begin{bmatrix} \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} \end{bmatrix}$$

$$D_{\vec{y}} g = \begin{bmatrix} \frac{\partial g}{\partial y_1} & \frac{\partial g}{\partial y_2} \end{bmatrix}$$

$$\frac{\partial g}{\partial x_1} = \frac{\partial}{\partial x_1}\left( x_1 x_2^2 y_1 + x_1 y_2^3 + x_2 x_1 y_2 y_1 + \frac{x_1^2}{x_2^3 y_1} \right)$$

$$= x_2^2 y_1 + y_2^3 + x_2 y_2 y_1 + \frac{1}{x_2^3 y_1} \frac{\partial}{\partial x_1}(x_1^2)$$

$$= x_2^2 y_1 + y_2^3 + x_2 y_2 y_1 + \frac{1}{x_2^3 y_1}(2x_1)$$

$$\frac{\partial g}{\partial x_2} = \frac{\partial}{\partial x_2}\left( x_1 x_2^2 y_1 + x_1 y_2^3 + x_2 x_1 y_2 y_1 + \frac{x_1^2}{x_2^3 y_1} \right)$$

$$= x_1 y_1 \frac{\partial}{\partial x_2}(x_2^2) + 0 + x_1 y_2 y_1 + \frac{x_1^2}{y_1} \frac{\partial}{\partial x_2}\left(\frac{1}{x_2^3}\right)$$

$$= 2x_1 x_2 y_1 + x_1 y_2 y_1 + \frac{x_1^2}{y_1}(-3x_2^{-4})$$

$$\frac{\partial g}{\partial y_1} = \frac{\partial}{\partial y_1}\left( x_1 x_2^2 y_1 + x_1 y_2^3 + x_2 x_1 y_2 y_1 + \frac{x_1^2}{x_2^3 y_1} \right)$$

$$= x_1 x_2^2 + 0 + x_2 x_1 y_2 + \frac{x_1^2}{x_2^3} \frac{\partial}{\partial y_1}\left(\frac{1}{y_1}\right)$$

$$= x_1 x_2^2 + x_2 x_1 y_2 + \frac{x_1^3}{x_2^3}(-y_1^{-2})$$

$$\frac{\partial g}{\partial y_2} = \frac{\partial}{\partial y_2}\left( x_1 x_2^2 y_1 + x_1 y_2^3 + x_2 x_1 y_2 y_1 + \frac{x_1^2}{x_2^3 y_1} \right)$$

$$= 0 + x_1 \frac{\partial}{\partial y_2}(y_2^3) + x_2 x_1 y_1 + 0$$

$$= 3x_1 y_2^2 + x_2 x_1 y_1$$

Stuff all four up there!
into $D_{\vec{x}} g$, $D_{\vec{y}} g$