Partial derivatives
Lecture 5b – 2021-06-09

MAT A35 – Summer 2021 – UTSC

Prof. Yun William Yu

What is a derivative?

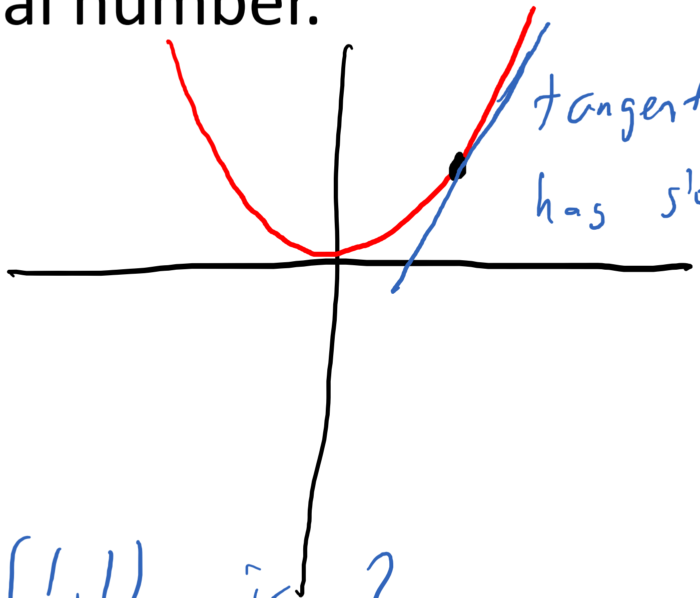
- A derivative measures the rate of change of a function as the variable it depends on changes.
- Given a function $f: \mathbb{R} \to \mathbb{R}$ written as $f(x)$, $\frac{df}{dx} = f'$ measures how quickly f changes when x changes.
- Note $f': \mathbb{R} \to \mathbb{R}$ since $f'(x)$ is a real number.

Ex.

$$f(x) = x^2$$

$$f'(x) = 2x$$

$$f'(1) = 2$$



tangent line has slope = $f'(x) = 2x$

tangent slope at point $(1,1)$ is 2

Partial derivatives of multivar. functions

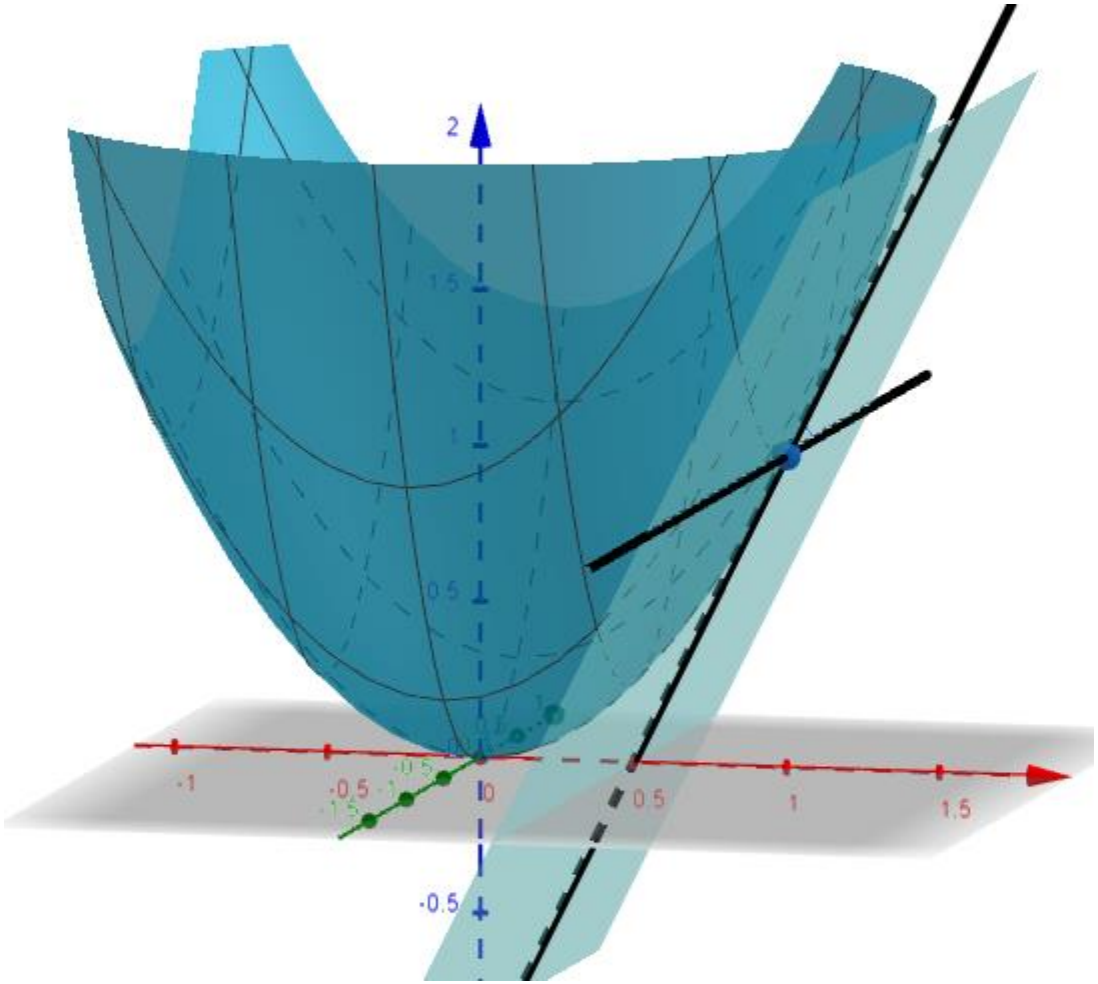
- We can measure the rate of change of the function with respect to each variable independently, assuming the other variable doesn't change.
- Given a function $f: \mathbb{R}^2 \to \mathbb{R}$ written as $f(x, y)$, the *partial derivative* $\frac{\partial f}{\partial x}$ measures how quickly f changes when x changes but y is fixed constant.
- Similarly, the partial derivative $\frac{\partial f}{\partial y}$ measures how quickly f changes when y changes but x is a fixed constant.
- Note $\frac{\partial f}{\partial x}: \mathbb{R}^2 \to \mathbb{R}$ takes as input a pair (x, y) and outputs a number

Pronunciation note: $\frac{\partial f}{\partial x}$ can be read several ways:

- del eff by del ecks
- del eff over del ecks
- del eff del ecks
- partial of eff with respect to ecks
- Sometimes even "dee eff dee ecks" if unambiguous

$$f(x, y) = x^2 + y^2$$

has slope $\frac{\partial f}{\partial x}$



Tangent slope at the point (1, 0, 1) depends on which direction we are going.

In the x-axis direction, slope is 2.

In the y-axis direction, slope is 0

https://www.geogebra.org/3d/j8ntyjzw

Formal definition of partial derivatives

- Recall: for $z = f(x)$, where $f\colon \mathbb{R} \to \mathbb{R}$, a 1-variable function
  - $\dfrac{dz}{dx} = \dfrac{df}{dx} = \lim\limits_{h \to 0} \dfrac{f(x+h) - f(x)}{h}$ $= f'$
- Let $z = f(x, y)$, where $f\colon \mathbb{R}^2 \to \mathbb{R}$, a 2-variable function.
  - $\dfrac{\partial z}{\partial x} = \dfrac{\partial f}{\partial x} = \lim\limits_{h \to 0} \dfrac{f(x+h, y) - f(x, y)}{h}$ $= f_x$
  - $\dfrac{\partial z}{\partial y} = \dfrac{\partial f}{\partial x} = \lim\limits_{h \to 0} \dfrac{f(x, y+h) - f(x, y)}{h}$ $= f_y$
- This generalizes in the natural way to n-variable functions, where you just treat all the other variables as constant.

Computing partial derivatives

- For the partial derivative with respect to a variable, treat all the other variables as constants and apply the normal derivative rules.

$f(x,y) = x^2 + y^2$

$f(1,0) = 1$

$\frac{\partial f}{\partial x}(\underline{x,y}) = \frac{\partial}{\partial x}[x^2 + y^2] = 2x$

constant

$\frac{\partial f}{\partial x}(1,0) = 2$

$\frac{\partial f}{\partial y}(x,y) = \frac{\partial}{\partial y}[x^2 + y^2] = 2y$

$\frac{\partial f}{\partial y}(1,0) = 0$

Example: $f(x, y) = x^2 + 2xy^2 + y^3$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}(x, y) = \frac{\partial}{\partial x}[x^2] + \frac{\partial}{\partial x}[2xy^2] + \frac{\partial}{\partial x}[y^3]$$

$$= 2x + 2y^2 + 0 = 2x + 2y^2$$

(treat y as a constant)

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial y}(x, y) = \frac{\partial}{\partial y}[x^2] + \frac{\partial}{\partial y}[2xy^2] + \frac{\partial}{\partial y}[y^3]$$

$$= 0 + 4xy + 3y^2 = 4xy + 3y^2$$

(treat x as a constant)

Try it out

- $f(x,y) = 3x^2y + xy^2$
  - Compute $\frac{\partial f}{\partial x}$ $= 6xy + y^2$
  - Compute f_y $= 3x^2 + 2xy$
- $w = g(x,y,z) = 5y^2 + 2yz$
  - Compute $\frac{\partial g}{\partial x}(x,y)$ $= 0$
  - Compute g_y $= 10y + 2z$
  - Compute $\frac{\partial w}{\partial z}$ $= \frac{\partial g}{\partial z} = 2y$
- Evaluating at a point
  - Compute $f_y(1,2)$ $= 3 + 4 = 7$
  - Compute $\frac{\partial w}{\partial z}(0,1,2)$ $= 2$

    $\alpha(x,y,z) = 2y = 2 \cdot 1 = 2$

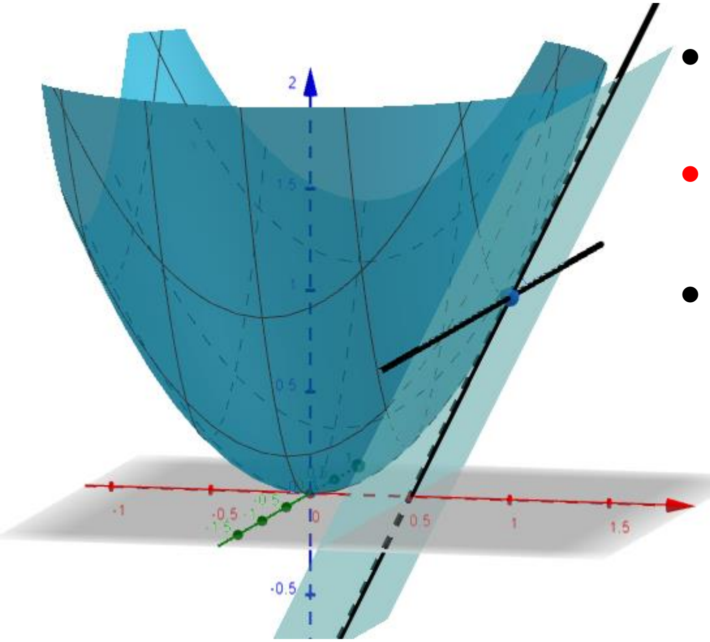
A: 0
B: $6xy + y^2$
C: $3x^2 + 2xy$
D: $3x^2 + 6xy + y^2 + 2xy$
E: None of the above

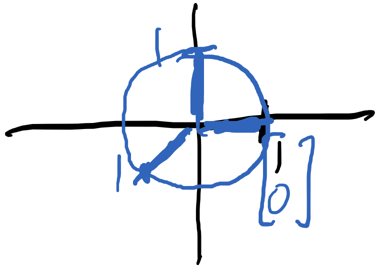
A: 0
B: $2y$
C: $10y + 2z$
D: $5y^2 + 2yz$
E: None of the above

A: 0
B: 2
C: 5
D: 7
E: None of the above

What about other directions?

$\begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\partial f}{\partial x}$





- We had an entire tangent plane.
- $\frac{\partial f}{\partial x}$ says how fast f grows in the x-direction.
- $\frac{\partial f}{\partial y}$ says how fast f grows in the y-direction.
- <u>Advanced (not on quiz 3):</u>
- Given a direction vector $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ where $u_1^2 + u_2^2 = 1$, we can compute how quickly f grows in the u-direction by computing the matrix product

$$\begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \frac{\partial f}{\partial x} \cdot u_1 + \frac{\partial f}{\partial y} \cdot u_2$$

where $\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}$ is the gradient of f.

Jacobian matrix

- Consider a function $h: \mathbb{R}^2 \to \mathbb{R}^2$ that takes a point in the plane to another point in the plane.
- We can write $h\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}$, where $f, g: \mathbb{R}^2 \to \mathbb{R}$.
- Then the Jacobian matrix of h (or of the pair of functions f and g) is given by:

$$J(x,y) = \begin{bmatrix} \nabla f \\ \nabla g \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix}$$

- The Jacobian matrix is the higher-dimensional analogue of a derivative, and tells you how the output of the function (a vector) changes as you go in a particular direction.

Example

$$h(x,y) = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = \begin{bmatrix} 2x - 3y^2 \\ 3xy^3 \end{bmatrix}$$

Jacobian $\quad J = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix} = \begin{bmatrix} 2 & -6y \\ 3y^3 & 9xy^2 \end{bmatrix}$

- Gradient ≈ "total" derivative of $f\colon \mathbb{R}^2 \to \mathbb{R}$ because it combines together all the partial derivatives.
- Jacobian ≈ "total" derivative of $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ because it combines together all the partial derivatives.

Higher-order partial derivatives

- Given $f(x, y)$ a function of two variables, $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}(x, y)$ is also a function of two variables.
- Define:
  - $\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\frac{\partial f}{\partial x} = f_{xx}$, which is taking the partial derivative by x twice
  - $\frac{\partial^2 f}{\partial y dx} = \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = f_{xy}$, which is taking partial-x, then partial-y
  - $\frac{\partial^2 f}{\partial x dy} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y} = f_{yx}$, which is taking partial-y, then partial-x
  - $\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\frac{\partial f}{\partial y} = f_{yy}$, which is taking the partial derivative by y twice

Example: $f(x,y) = x^3y^2 + y\sin x + xe^y$

$$\frac{\partial f}{\partial x} = 3x^2y^2 + y\cos x + e^y \qquad \frac{\partial f}{\partial y} = 2x^3y + \sin x + xe^y$$

$$\frac{\partial^2 f}{\partial x^2} = 6xy^2 - y\sin x \qquad \frac{\partial^2 f}{\partial x\,\partial y} = 6x^2y + \cos x + e^y$$

$$\frac{\partial^2 f}{\partial y\,\partial x} = 6x^2y + \cos x + e^y \quad = \qquad \frac{\partial^2 f}{\partial y^2} = 6x^2 + e^y$$

- Note: "usually" it is true that $\frac{\partial^2 f}{\partial y dx} = f_{xy} = f_{yx} = \frac{\partial^2 f}{\partial x dy}$.

Try it out

- $f(x, y) = x^2y^2 + 4xy$
- $\frac{\partial f}{\partial x} = 2xy^2 + 4y$
- $\frac{\partial f}{\partial y} = 2x^2y + 4x$
- $\frac{\partial^2 f}{\partial x^2} = 2y^2$
- $\frac{\partial^2 f}{\partial y dx} = 4xy + 4$
- $\frac{\partial^2 f}{\partial x dy} = 4xy + 4$
- $\frac{\partial^2 f}{\partial y^2} = 2x^2$

$\frac{\partial}{\partial x}$

A: $2x^2$
B: $2x^2y + 4x$
C: $2y^2$
D: $2xy^2 + 4y$
E: $4xy + 4$

Hessian matrix

- Hessian matrix corresponds to second derivative

$f(x,y)$, $f: \mathbb{R}^2 \to \mathbb{R}$

$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}$ 1st "total" derivative

Can think of ∇f as a function $\nabla f : \mathbb{R}^2 \to \mathbb{R}^2$

$$\text{Jacobian}(\nabla f) = \text{Jacobian}\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

$$= \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \text{Hessian}$$

Hessian matrix ≈ 2nd total derivative

- Say we have $f: \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y)$.
- 1st total derivative $\approx \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}$
- We can think of $\nabla f: \mathbb{R}^2 \to \mathbb{R}^2$ by transposing $\nabla f^T = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$
- Then we can the total derivative of ∇f by using the Jacobian, and we'll call that new matrix the Hessian of f.
- $\text{Hessian}(f) = \text{Jacobian}(\nabla \text{f}) = \text{Jacobian}\left(\frac{\partial f}{\partial x}, \quad \frac{\partial f}{\partial y}\right)$

$$= \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y dx} \\ \frac{\partial^2 f}{\partial x dy} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

- The Hessian includes all the 2nd partial derivatives of f.