Mathematical Brush Up for Machine Learning

- Saurav Poudel

Topics

- Linear Algebra
- Calculus

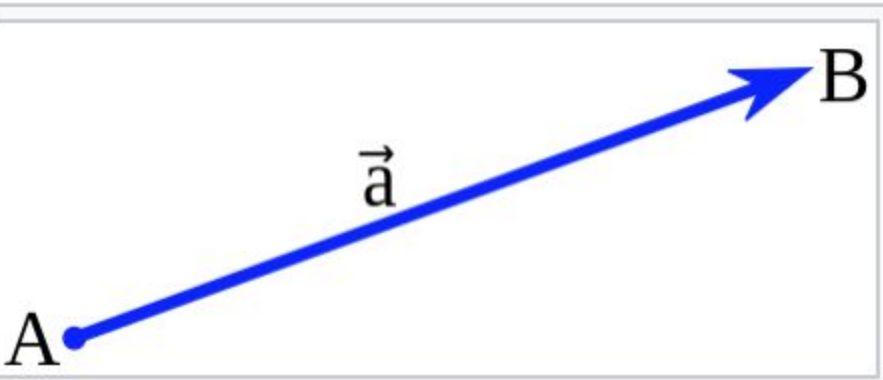
Linear Algebra

- Vector
- Vector Operations
- Dot Product / Similarity
- Matrix / Transformation
- Dot Product using Matrix Notation
- Matrix Operations

Calculus

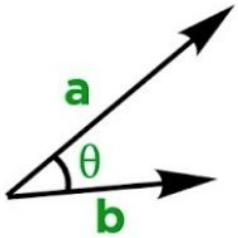
- Derivative
- Minima / Maxima
- Optimization
- Integration

Vector vs Scalar
(Examples, with real Data Case too)



A vector pointing from *A* to *B*

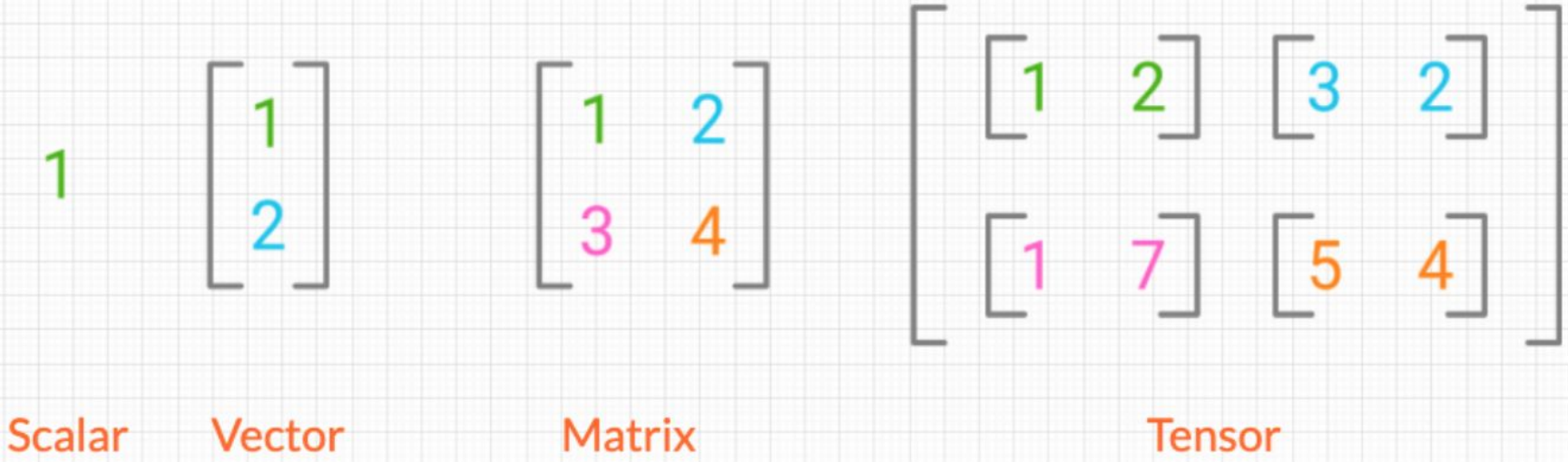
The Vector Dot Product

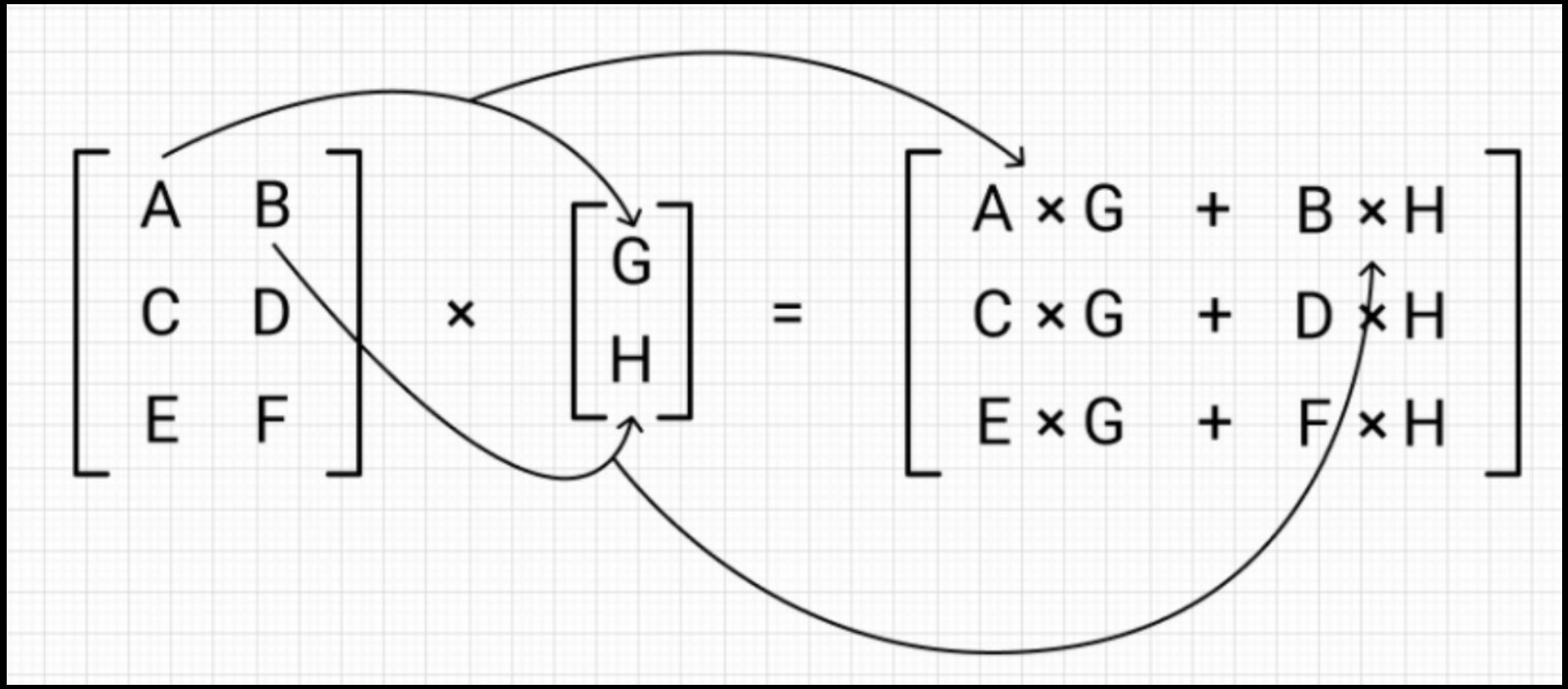


$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$$

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$$

Scalars, Vectors, Matrices & Tensors

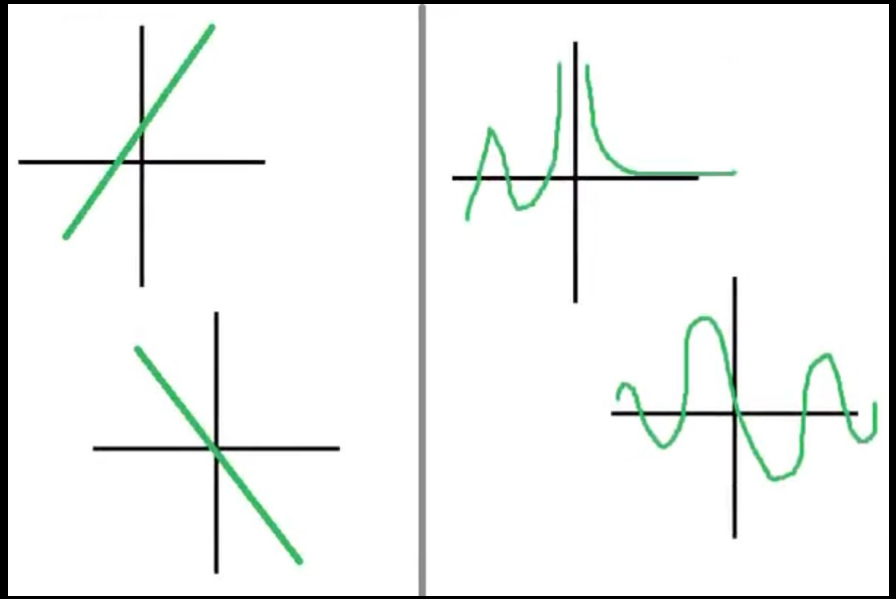



A B
C D
E F
×
G
H
=
A × G + B × H
C × G + D × H
E × G + F × H

Modeling / ML

- Modeling in terms of f(x) approx
- Finding f(x) meaning finding a vector
- Finding f(x) meaning finding a matrix
- Matrix as transformation
- Purpose to find that f(x) in terms of vector / matrix
- Whole Modeling and ML in summary

Concept of Linearity (and non linearity)



Linearity

- What do you mean by Linear?
- Linearity in terms of 3 terms.
- Role of Matrix.

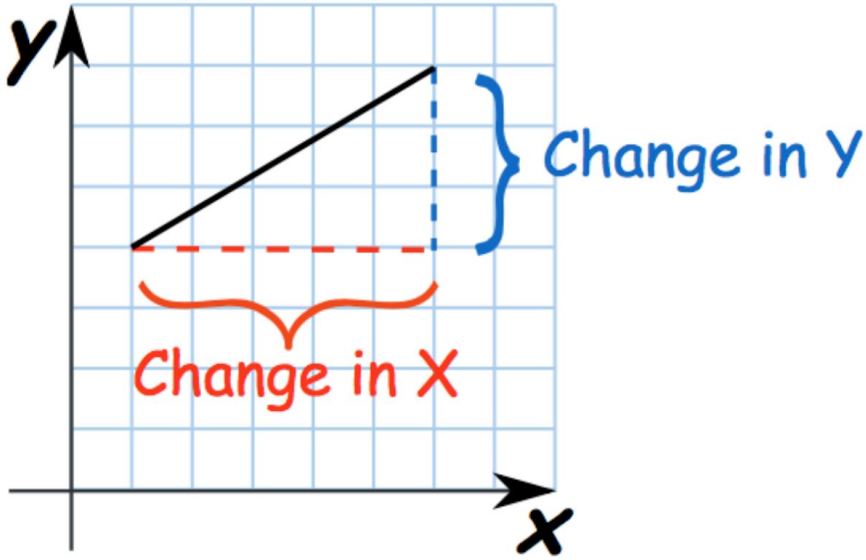
Non-Linearity

- What do you mean by Non-Linear?
- Examples of Non-Linear
- How to deal with Non-Linear

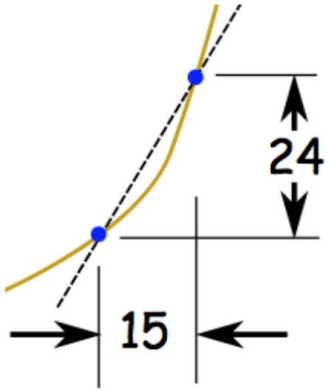
Calculus time!

It is all about slope!

$$\text{Slope} = \frac{\text{Change in Y}}{\text{Change in X}}$$



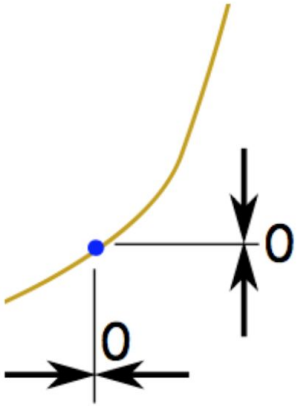
We can find an **average** slope between two points.



average slope = $\frac{24}{15}$

But how do we find the slope **at a point**?

There is nothing to measure!



$$\text{slope} = \frac{0}{0} = ???$$

$$\text{Slope} = \frac{\text{Change in Y}}{\text{Change in X}} = \frac{\Delta y}{\Delta x}$$

$$\frac{\Delta y}{\Delta x} = \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

$f(x) = x^2$

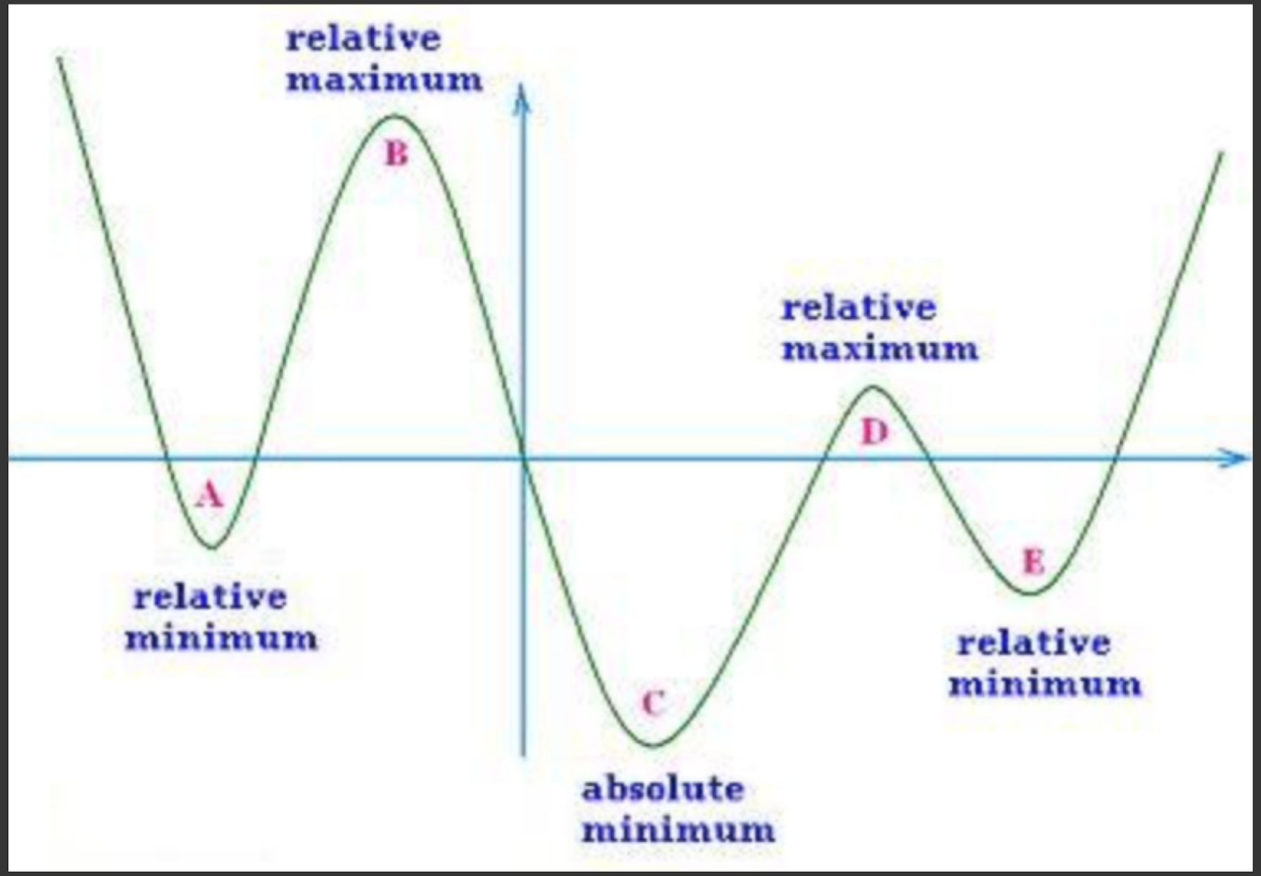
The slope formula is: $\dfrac{f(x+\Delta x) - f(x)}{\Delta x}$

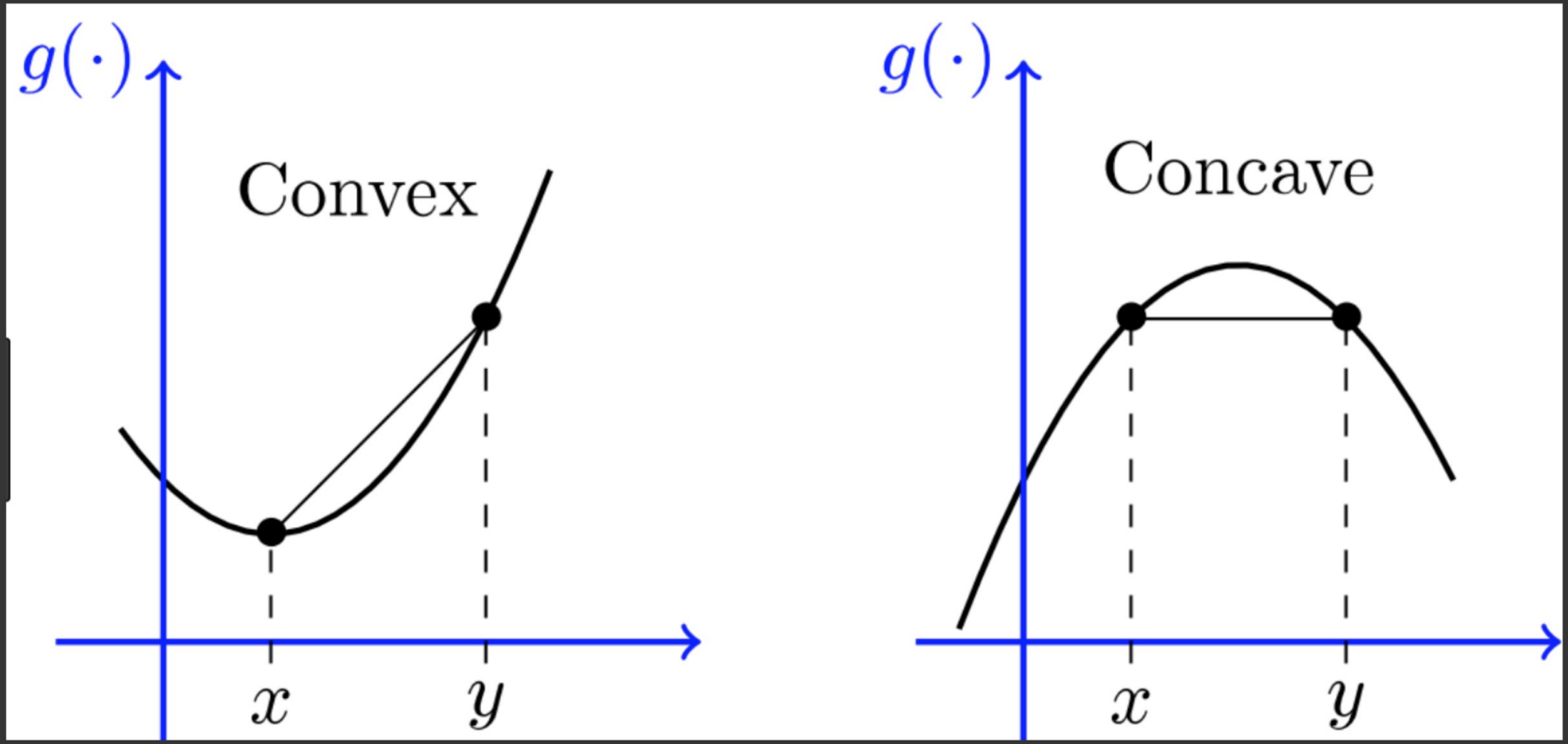
Put in **f(x+Δx)** and **f(x)**: $\dfrac{x^2 + 2x\,\Delta x + (\Delta x)^2 - x^2}{\Delta x}$

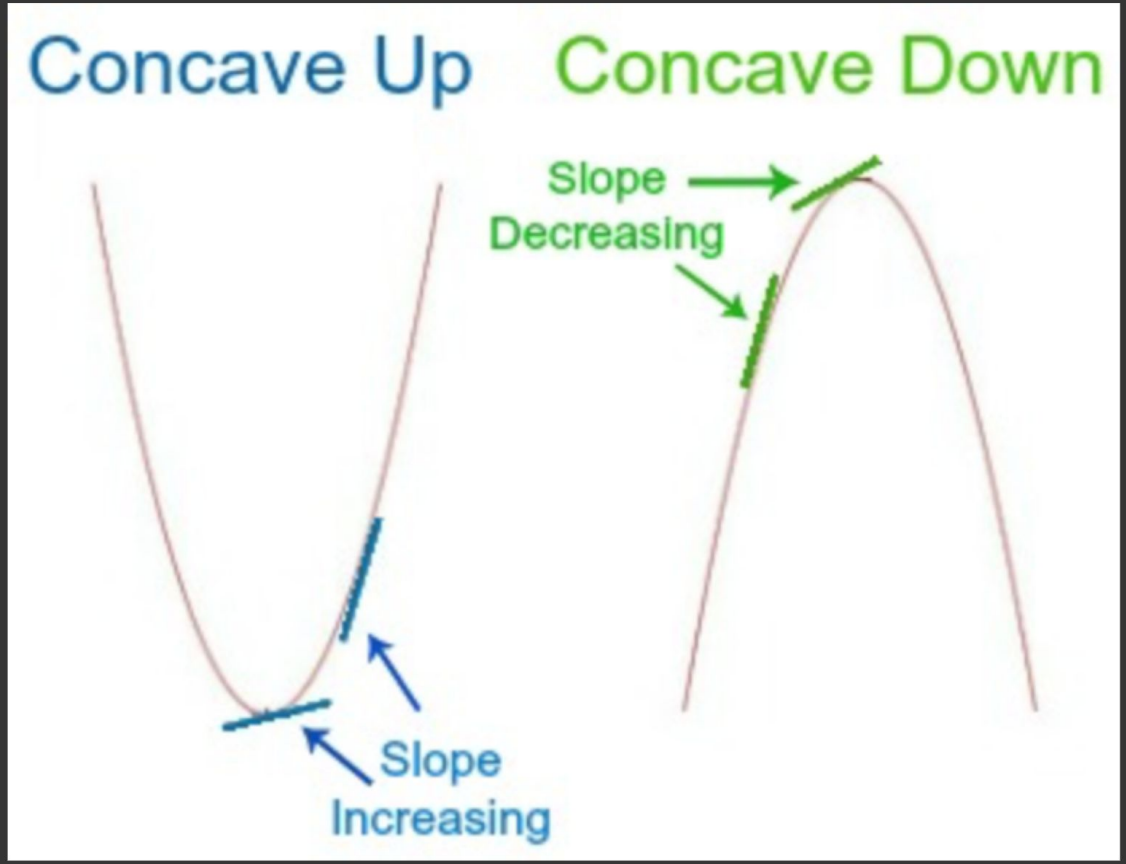
Simplify (x^2 and $-x^2$ cancel): $\dfrac{2x\,\Delta x + (\Delta x)^2}{\Delta x}$

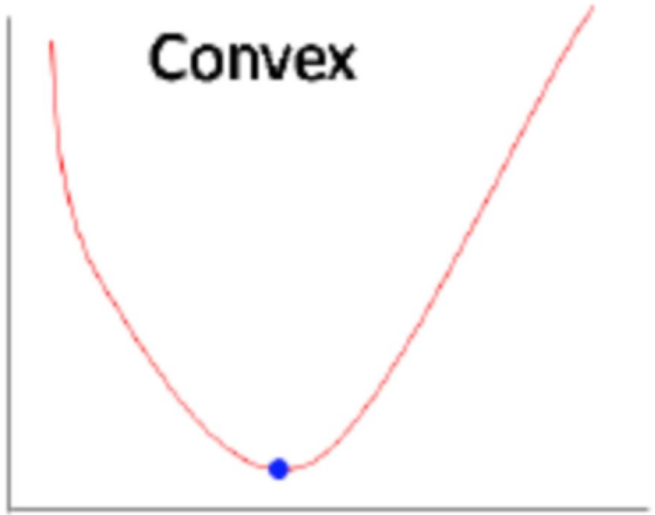
Simplify more (divide through by Δx): $= 2x + \Delta x$

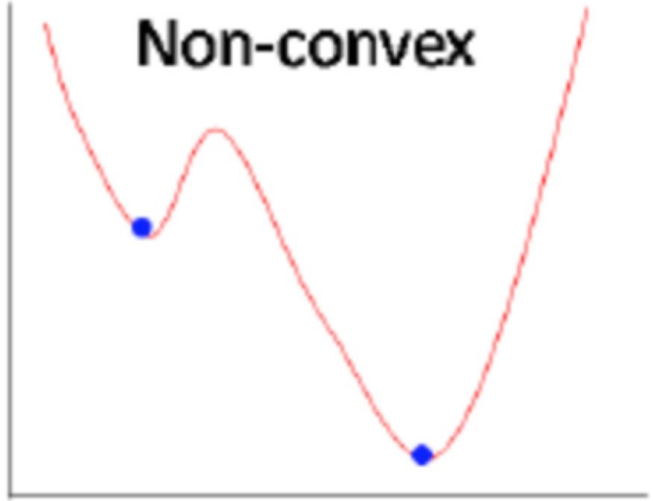
Then **as Δx heads towards 0** we get: $= 2x$


relative
maximum
B
relative
maximum
D
A
E
relative
minimum
relative
minimum
C
absolute
minimum


$g(\cdot)$
Convex
x
y
$g(\cdot)$
Concave
x
y


Concave Up
Concave Down
Slope
Decreasing
Slope
Increasing


Convex


Non-convex

End Note!

Thank You!