JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Algorithms for Data Science

Optimization: Mathematical Foundations

Convex Sets and Convex Hull

Most optimization algorithms are easier and more efficient on convex problems since convex sets ensure that any local optimum is also a global optimum.

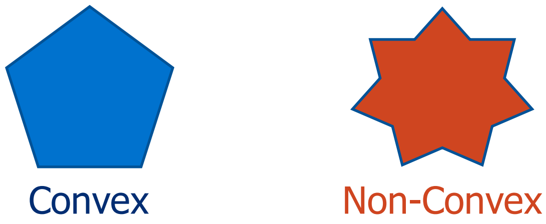
Convex Sets

- A set is convex if, for any two points x_1, x_2 in the set, the line segment joining them lies entirely within the set.
- Mathematically:

$$\lambda x_1 + (1-\lambda)x_2 \in S, \forall \lambda \in [0,1]$$

Convex Hull

- The smallest convex set that contains all the points in a given set.
- Used to simplify problems by approximating non-convex sets with convex boundaries.



Convex Functions

- A function $f(x)$ is convex if its domain is a convex set and:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2), \forall \lambda \in [0,1]$$

- This means that the line segment between two points on the function lies above or on the function curve.

- Properties:

1. First Derivative Test:
   - If $f(x)$ is differentiable, it is convex if:

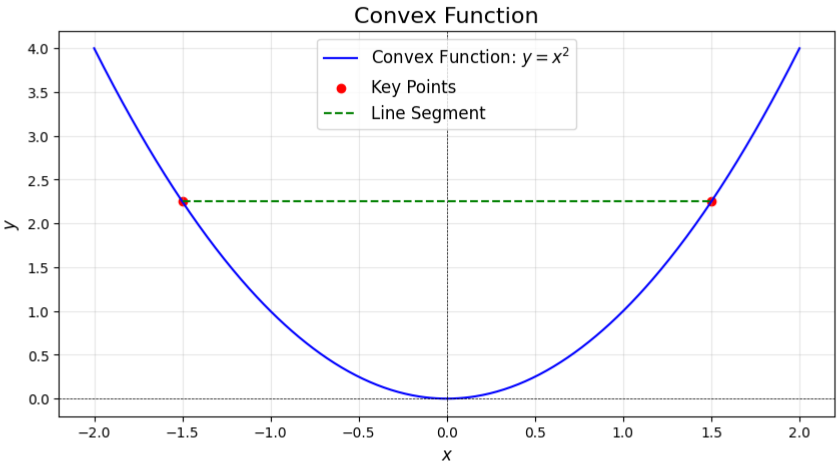
$$f(y) \geq f(x) + \nabla f(x)^T(y-x)$$

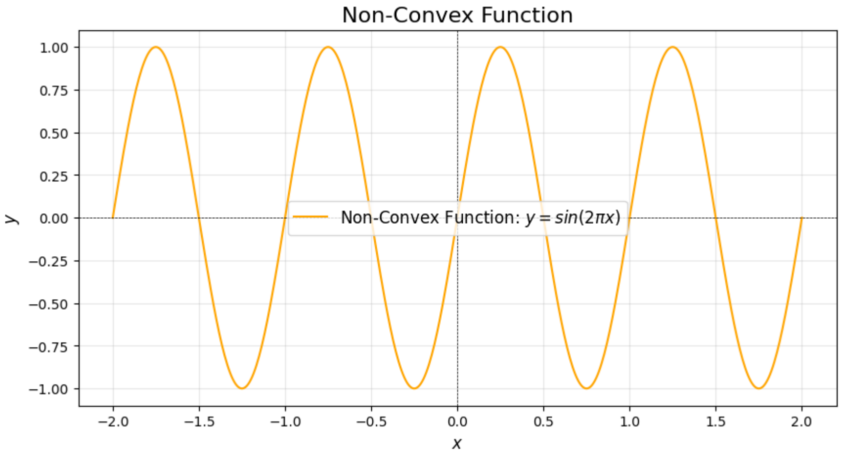
2. Second Derivative Test:
   - If $f(x)$ is twice differentiable, it is convex if the Hessian H(x) is positive semidefinite:

$$H(x) \geq 0$$

Convex Functions Examples

- Convex: $f(x) = x^2, f(x) = e^x$
- Non-Convex: $f(x) = -x^2, f(x) = \sin(x)$





JOHNS HOPKINS
WHITING SCHOOL of ENGINEERING

Linear Algebra Review

1. **Matrix-Vector Multiplication:**
   - Foundation for expressing systems of linear equations.
   - E.g. $Ax=b$, where A is a matrix, x is a vector of variables, and b is the result.

2. **Dot Product:**
   - Measures similarity between two vectors.
   - Formula:

$$a \cdot b = \sum_{i=1}^{n} a_i b_i$$

3. **Norms:**
   Measure the length or size of a vector.
   Formula:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

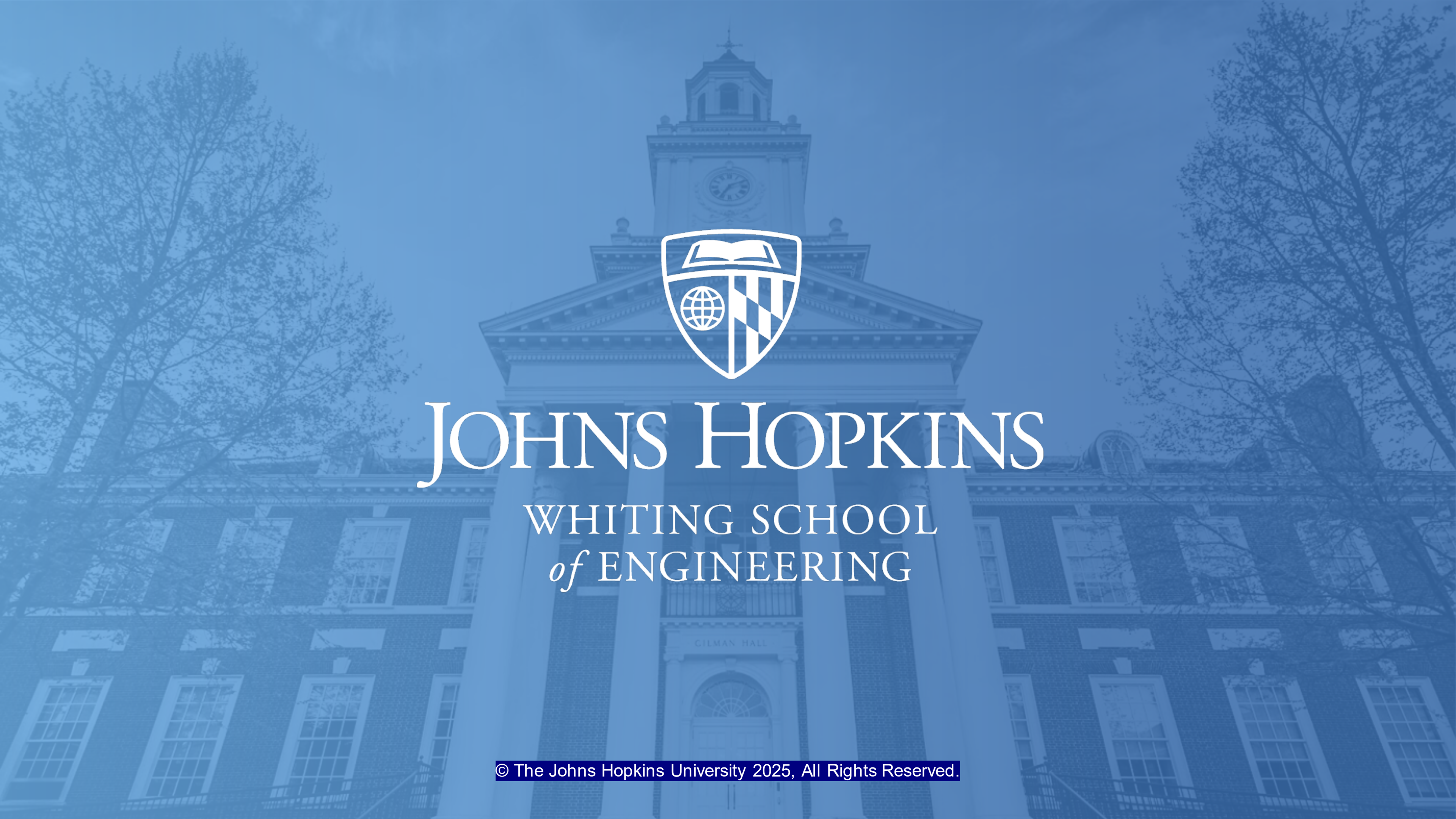
Gradient and Hessian

- **Gradient:** The vector of partial derivatives, representing the direction of steepest ascent.

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

- **Hessian:** The matrix of second-order partial derivatives, representing the curvature of the function.

$$H(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$


JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING