18-660: Numerical Methods for Engineering Design and Optimization

Xin Li
Department of ECE
Carnegie Mellon University
Pittsburgh, PA 15213

Electrical & Computer ENGINEERING

Overview

- Unconstrained Optimization
  - Golden section search
  - Downhill simplex method

Unconstrained Optimization

- Linear regression with regularization

$$A\alpha = B \quad \Rightarrow \quad \min_{\alpha} \; \|A\alpha - B\|_2^2 + \lambda \cdot \|\alpha\|_1$$

- Unconstrained optimization: minimizing a cost function without any constraint
  - Golden section search
  - Downhill simplex method

  → Non-derivative method (this lecture)

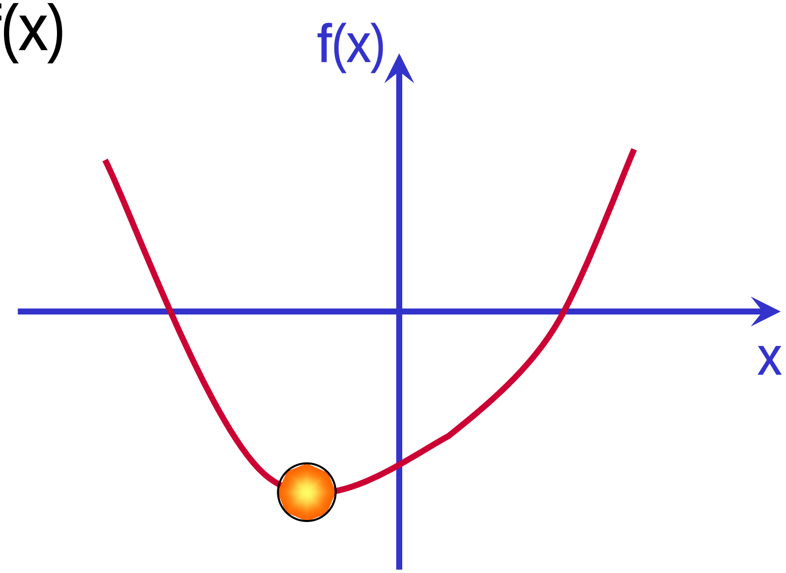
  - Gradient method
  - Newton method

  → Rely on derivatives

Golden Section Search

- Problem definition: find the optimum x that minimizes a one-dimensional cost function f(x)

$$\min_{x} \quad f(x)$$

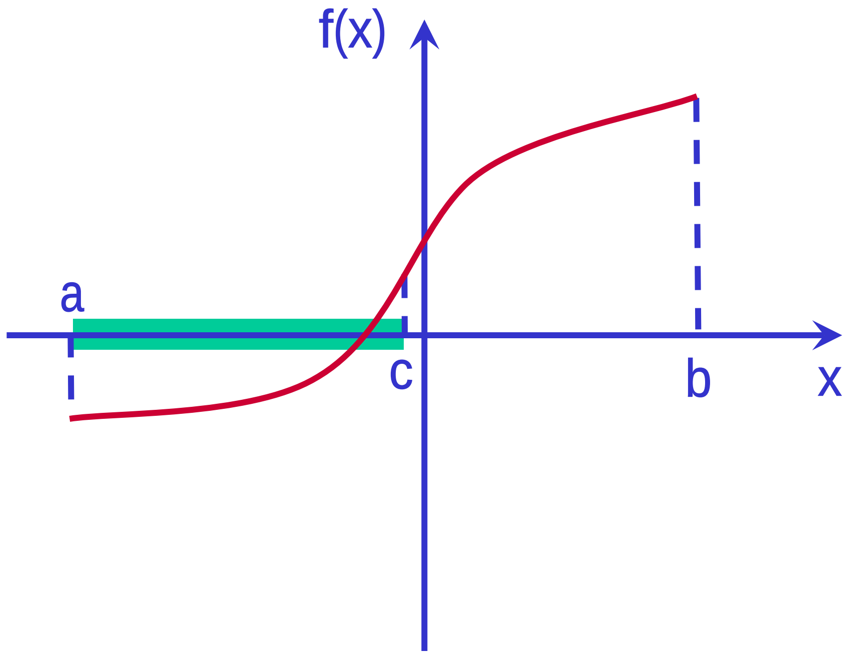


- Key idea
  - Assume that the optimum x is within an interval [a, b]
  - Iteratively shrink [a, b] to find the solution x

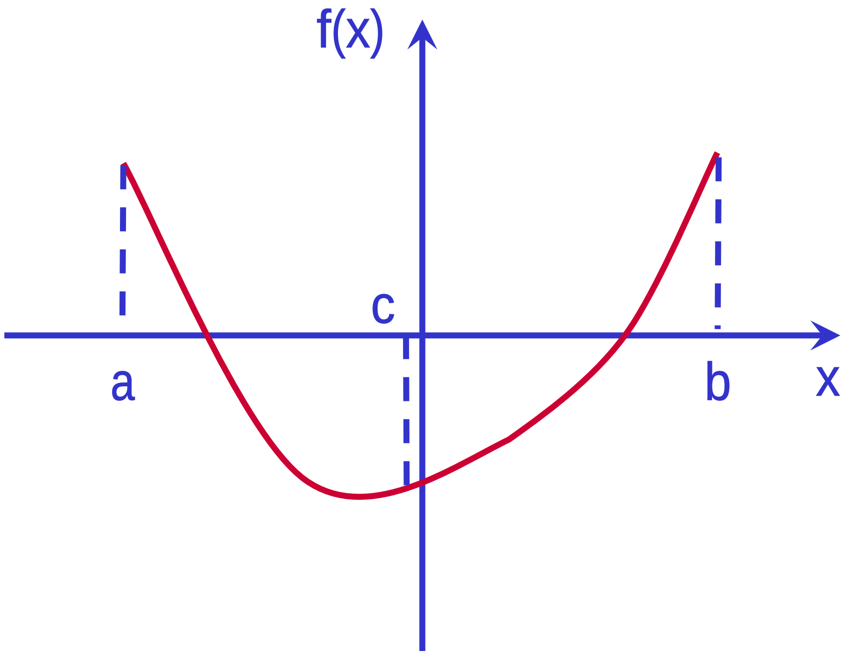
Similar (but not identical) to binary search

Golden Section Search

- It is not trivial to determine the interval that contains optimal solution x



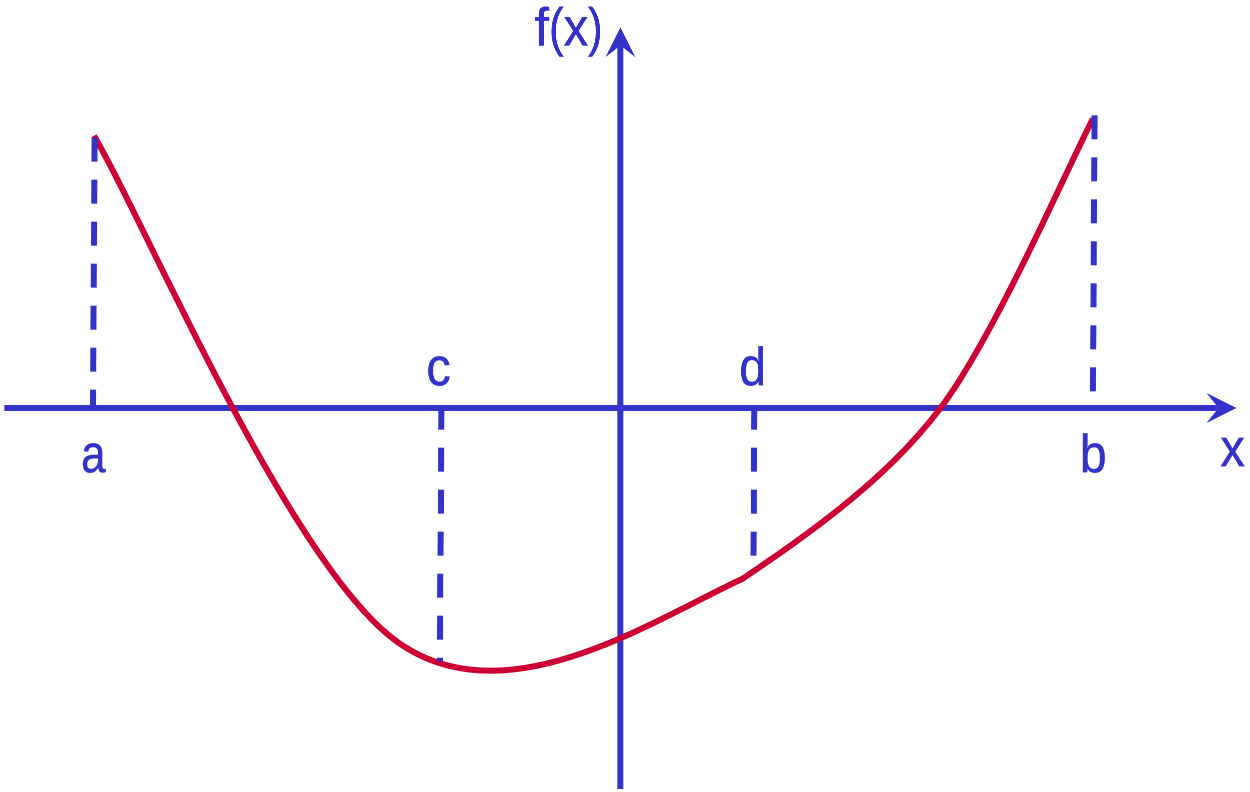
Binary search: f(x) = 0



Golden section search: minimize f(x)

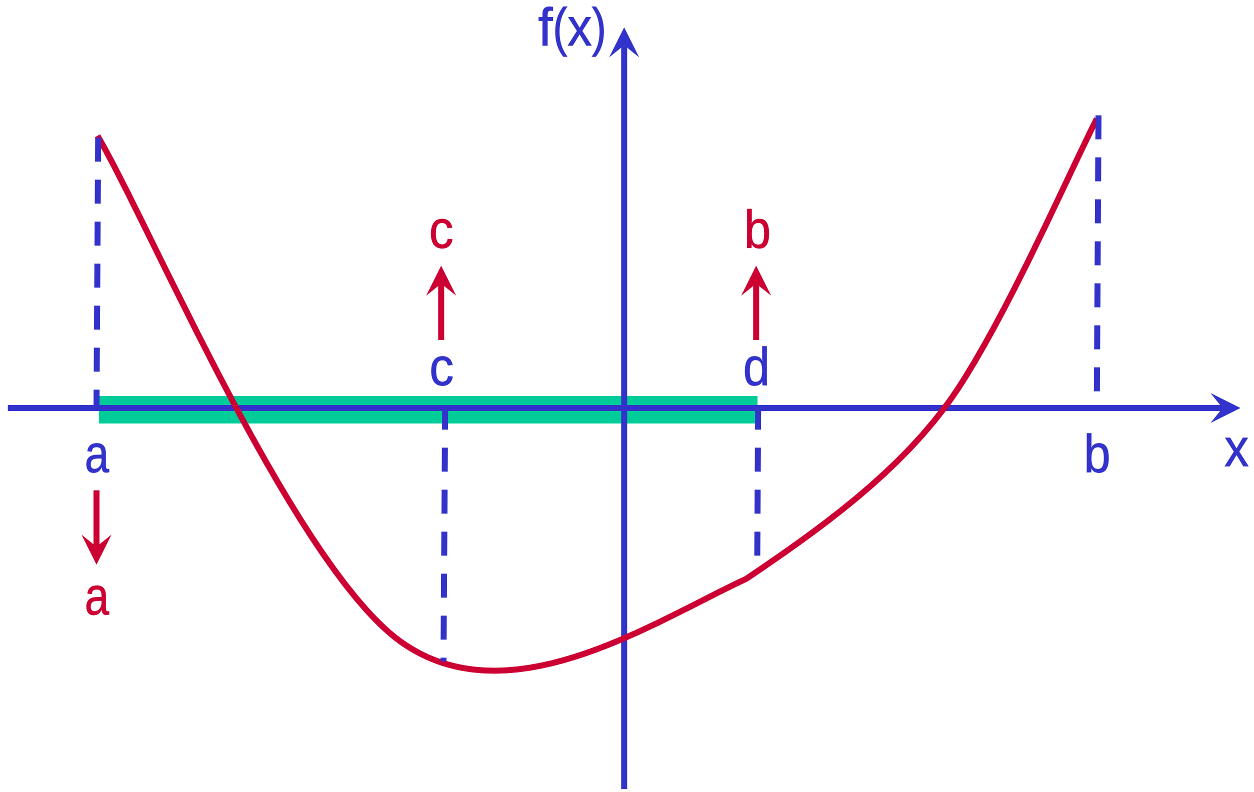
Golden Section Search

- Require a triplet (a, b, c) to decide the appropriate interval
  - Start from (a, b, c): f(c) < f(a) and f(c) < f(b)
  - Add one extra point d



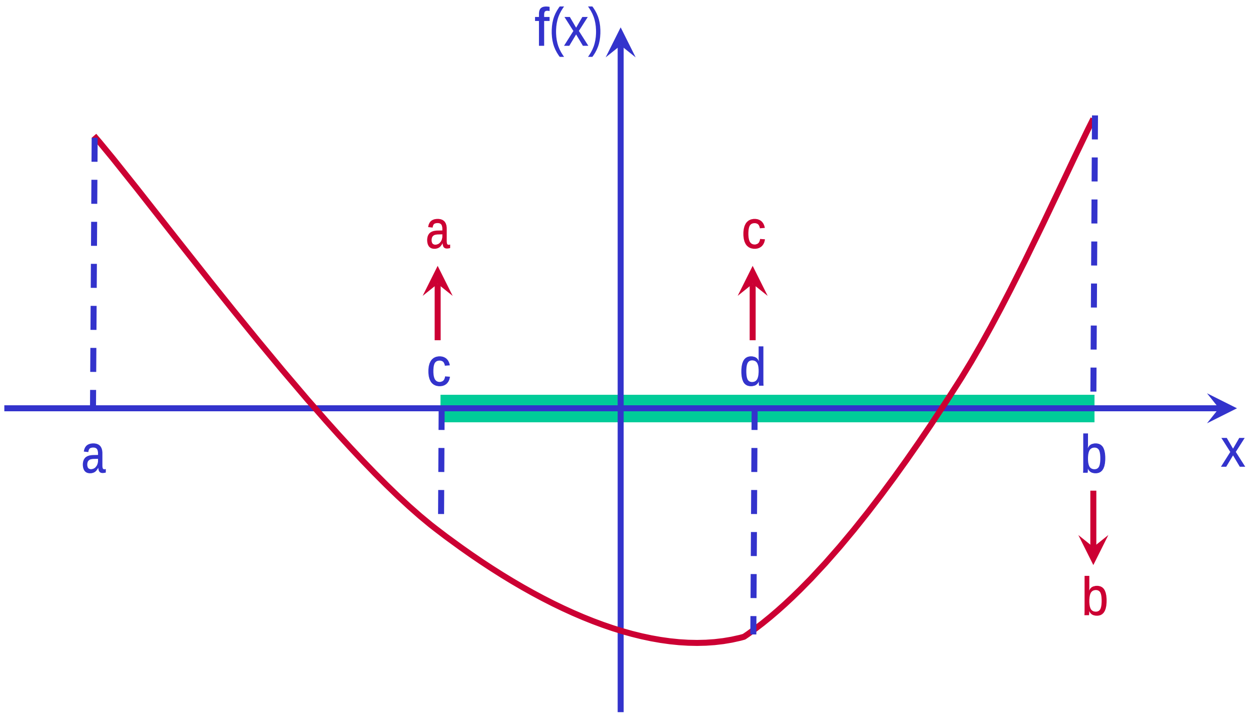
Golden Section Search

- If f(c) has the smallest value, optimal solution is within [a, d]



Golden Section Search

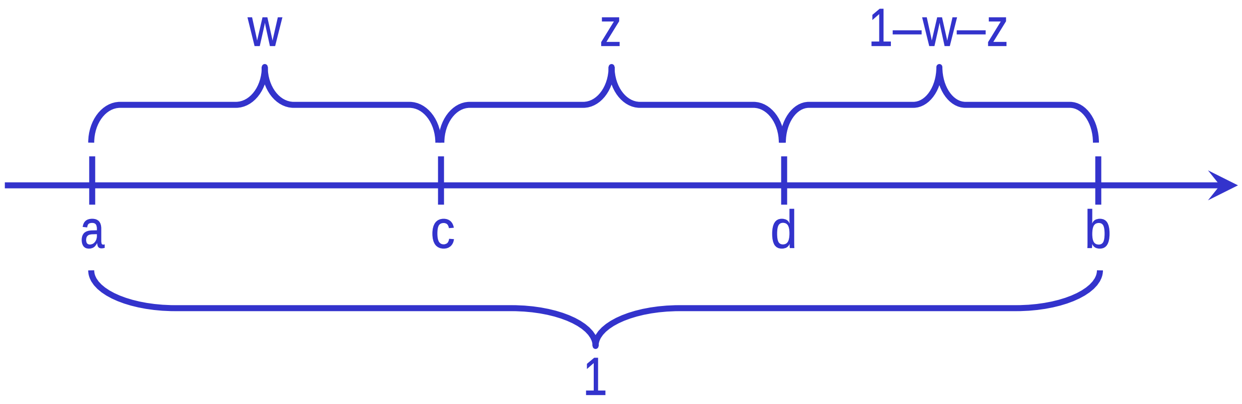
- If f(d) has the smallest value, optimal solution is within [c, b]



Continue iteration until (local) convergence is reached

Golden Section Search

- Open question: how to decide the location of c and d?
- Criterion: each iteration identically shrinks the interval
  - We will explain its meaning soon



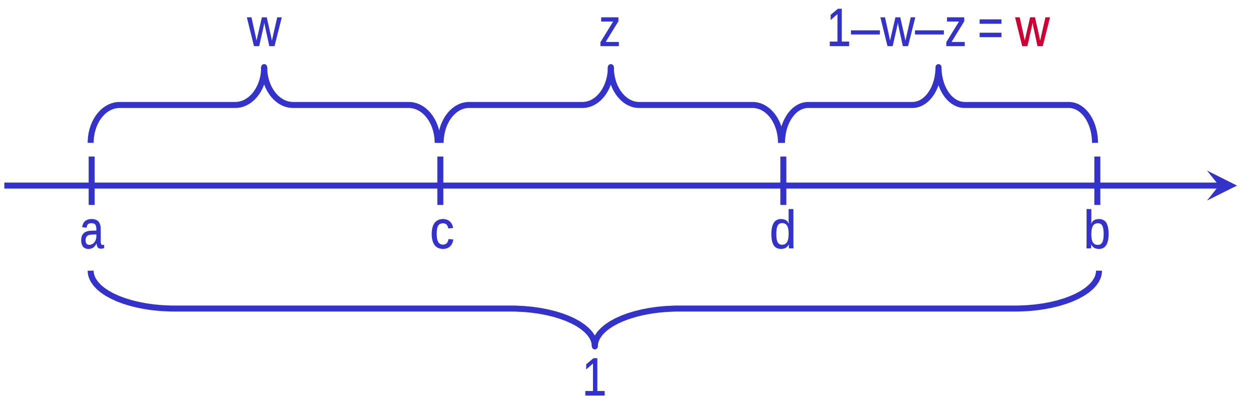
Golden Section Search

- Criterion: each iteration identically shrinks the interval
  - Starting from [a, b], the next interval can be [a, d] or [c, b]
  - The length of these two intervals should be identical

$$|a-d| = |c-b|$$

$$w + z = z + (1 - w - z)$$

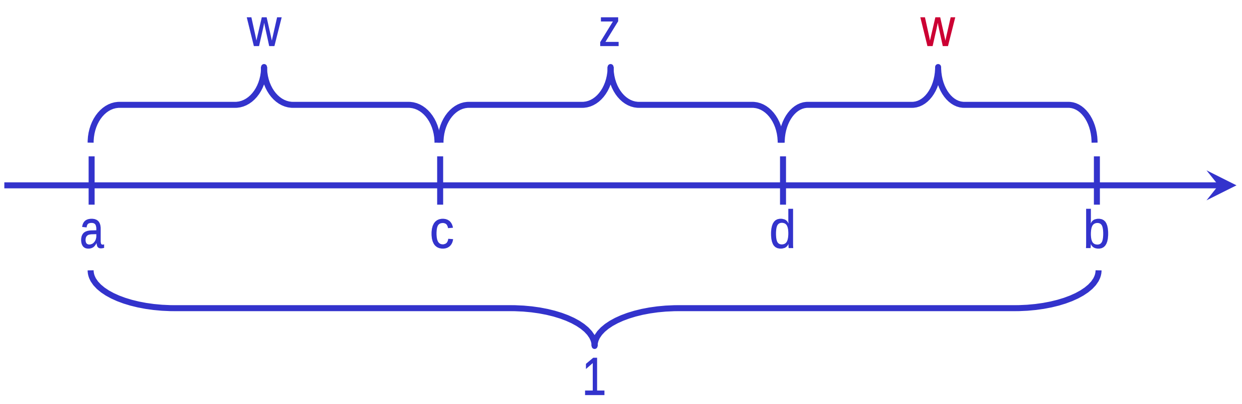
$$w = 1 - w - z$$



Golden Section Search

- Criterion: each iteration identically shrinks the interval
  - Starting from [a, b], the next interval can be [a, d] or [c, b]
  - The ratio of these two intervals is identical

$$\frac{|c-d|}{|a-c|} = \frac{|c-d|}{|d-b|} = \frac{z}{w}$$

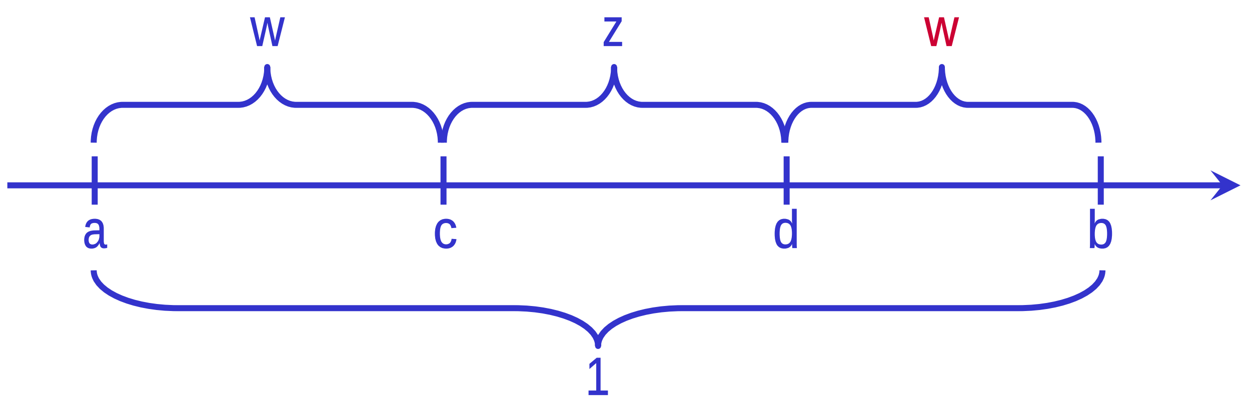


Golden Section Search

- Criterion: each iteration identically shrinks the interval
  - Starting from [a, b], the next interval can be [a, d] or [c, b]
  - The ratio of these three intervals should be identical

$$\frac{|c-d|}{|a-c|} = \frac{|c-d|}{|d-b|} = \frac{z}{w} \qquad \frac{|a-c|}{|c-b|} = \frac{w}{z+w}$$

$$\frac{w}{z+w} = \frac{z}{w}$$

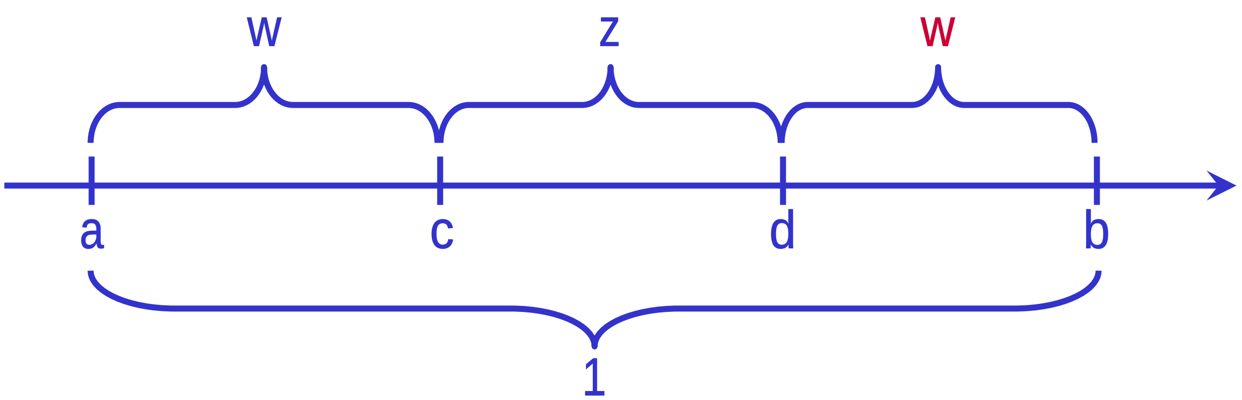


Golden Section Search

- Criterion: each iteration identically shrinks the interval

$$\frac{w}{z+w} = \frac{z}{w} \qquad \Rightarrow \qquad w = 0.382$$

$$2w + z = 1$$



A Simple Example

$$f(x) = x^2 \quad where \quad x \in [-1,1]$$

- Iteration #1

$$a = -1 \qquad f(-1) = 1$$

$$b = 1 \qquad f(1) = 1$$

$$c = a + |a - b| \cdot w = -0.236 \qquad f(-0.236) = 0.056$$

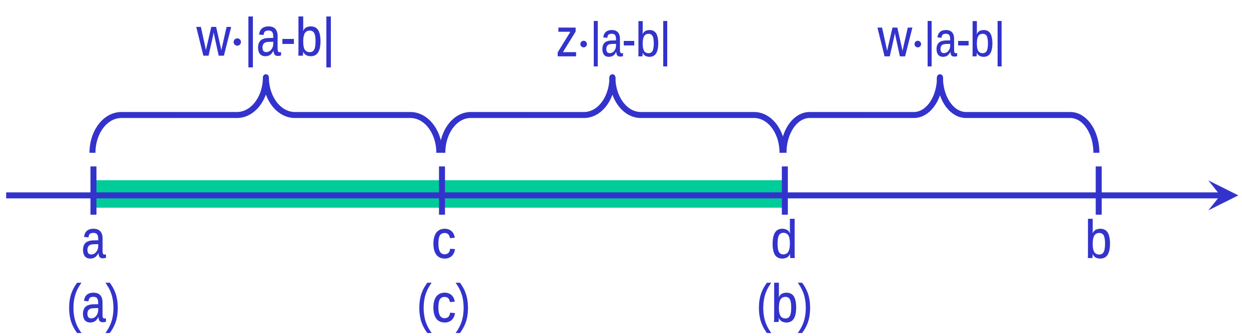
$$d = b - |a - b| \cdot w = 0.236 \qquad f(0.236) = 0.056$$

$\Rightarrow$

$$a = -1$$

$$b = 0.236$$

$$c = -0.236$$



A Simple Example

$$f(x) = x^2 \quad where \quad x \in [-1,1]$$

- Iteration #2

$$a = -1 \qquad f(-1) = 1$$

$$b = 0.236 \qquad f(0.236) = 0.056$$

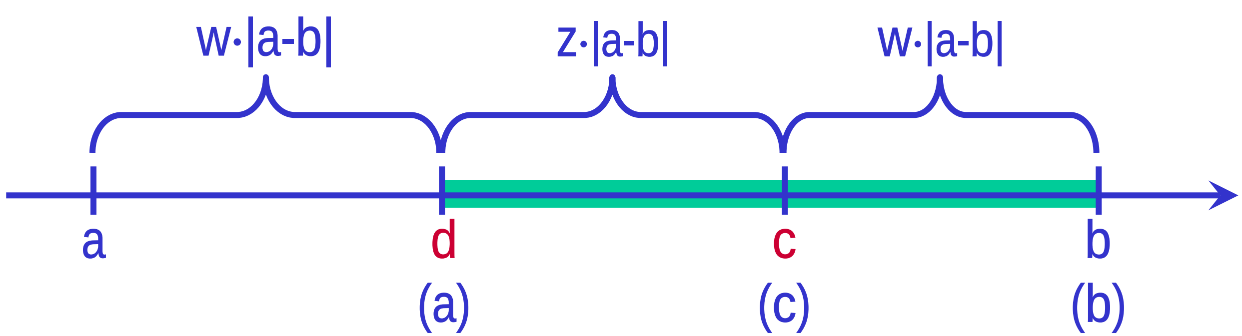
$$c = -0.236 \qquad f(-0.236) = 0.056$$

$$d = a + |a-b| \cdot w = -0.528 \qquad f(-0.528) = 0.279$$

$$a = -0.528$$

$$b = 0.236$$

$$c = -0.236$$



A Simple Example

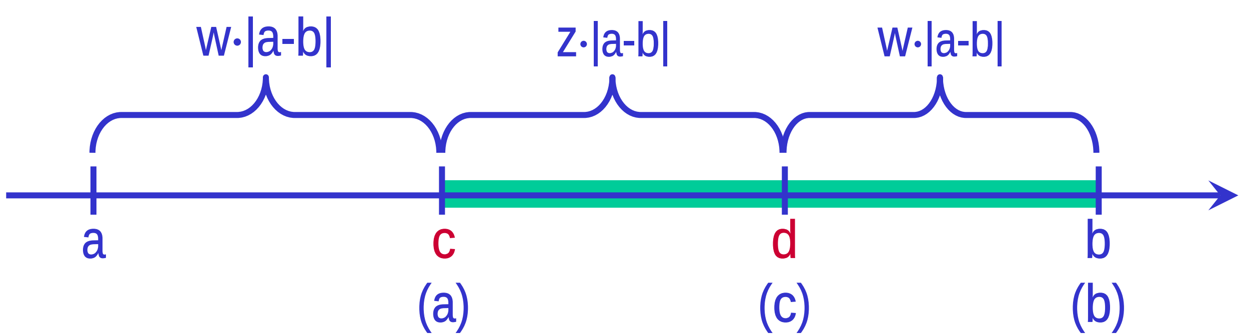
$$f(x) = x^2 \quad where \quad x \in [-1,1]$$

- Iteration #3

$$a = -0.528 \qquad f(-0.528) = 0.279$$
$$b = 0.236 \qquad f(0.236) = 0.056$$
$$c = -0.236 \qquad f(-0.236) = 0.056$$
$$d = b - |a-b| \cdot w = -0.056 \qquad f(-0.056) = 0.003$$

$\Rightarrow$

$$a = -0.236$$
$$b = 0.236$$
$$c = -0.056$$



Downhill Simplex Method

- Golden section search is easy to implement
  - However, it is typically used for one-dimensional problem only

- Multi-dimensional optimization can also be solved by non-derivative method

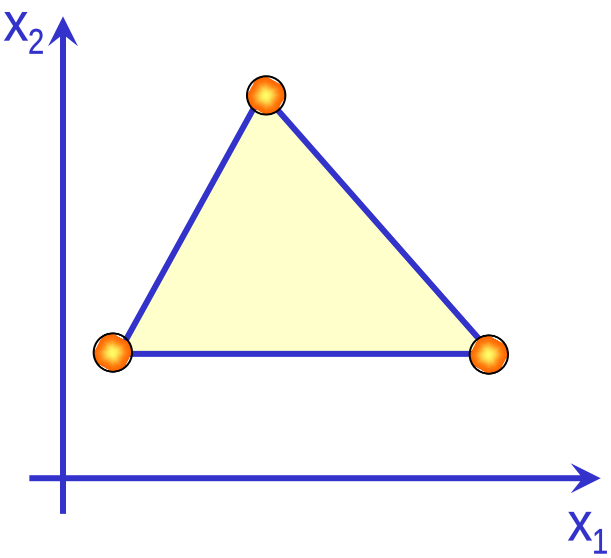
$$\min_X \quad f(X) \quad where \quad X \in R^N$$

  - E.g., downhill simplex method

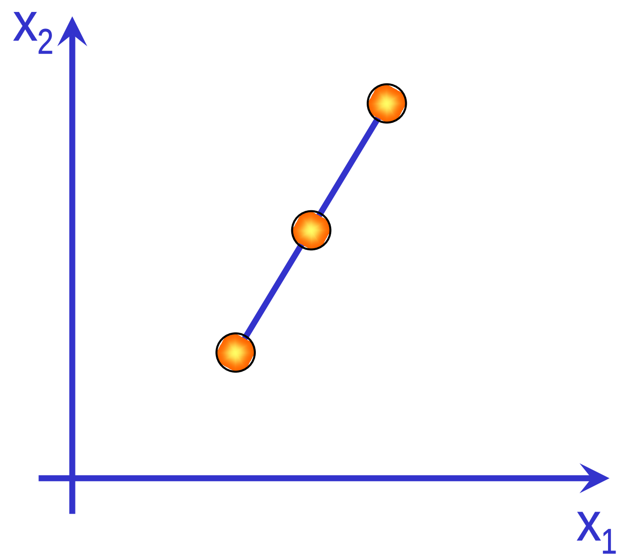
- Key idea
  - Find the optimum X by a sequence of reflection, expansion and contraction of a simplex

Downhill Simplex Method

- In N-dimensional space, a simplex is defined by N+1 non-degenerate points



Three non-degenerate points define a simplex in 2-D space



Three degenerate points does NOT define a simplex in 2-D space

Downhill Simplex Method

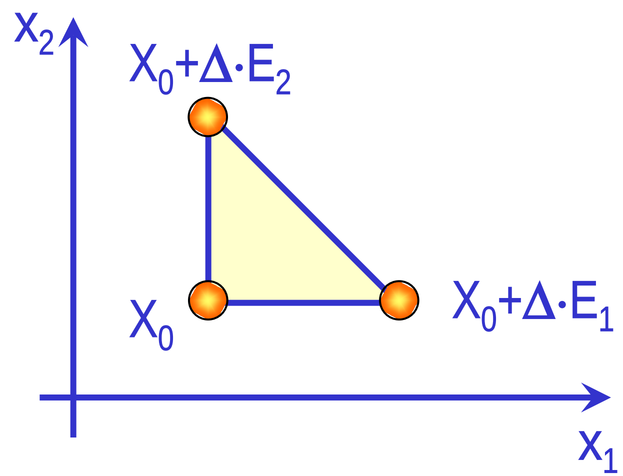
- Construct initial simplex
  - Start from an initial guess $X_0 \in R^N$
  - Apply perturbation in N orthogonal directions

$$X_i = X_0 + \underline{\Delta} \cdot \underline{E_i}$$

Small perturbation — Unit vector

$$E_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

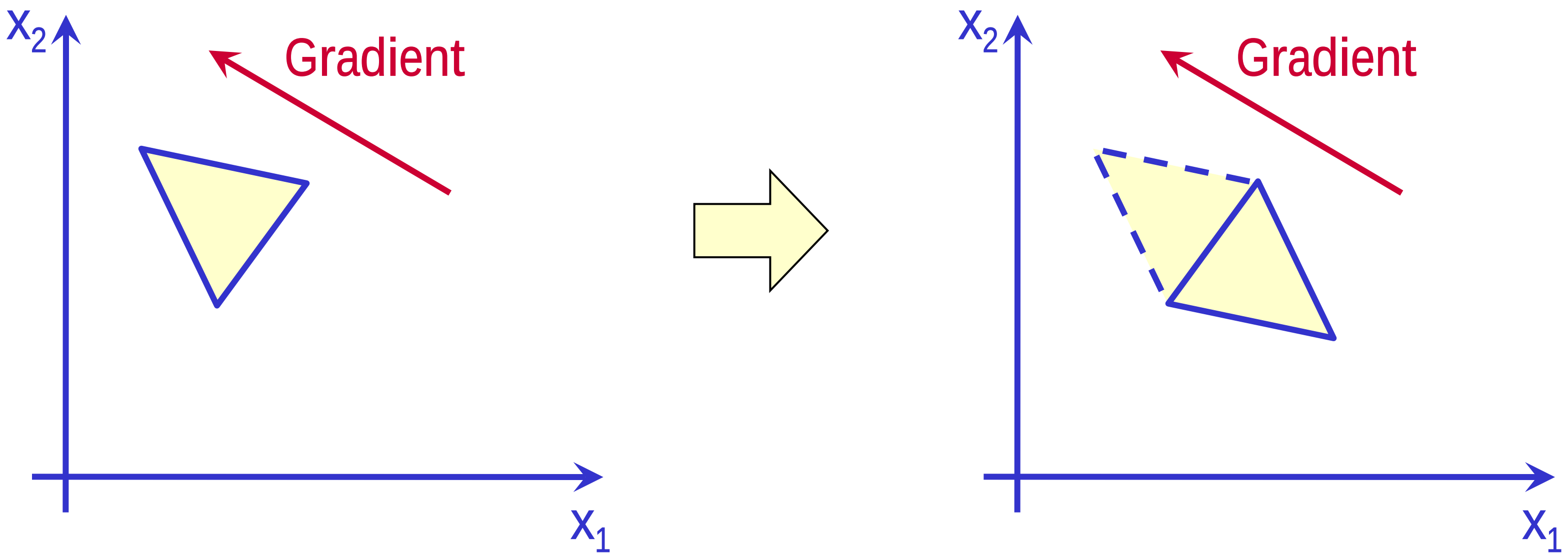
1 → i-th element



  - $\{X_0, X_1, ..., X_N\}$ define a simplex

Downhill Simplex Method

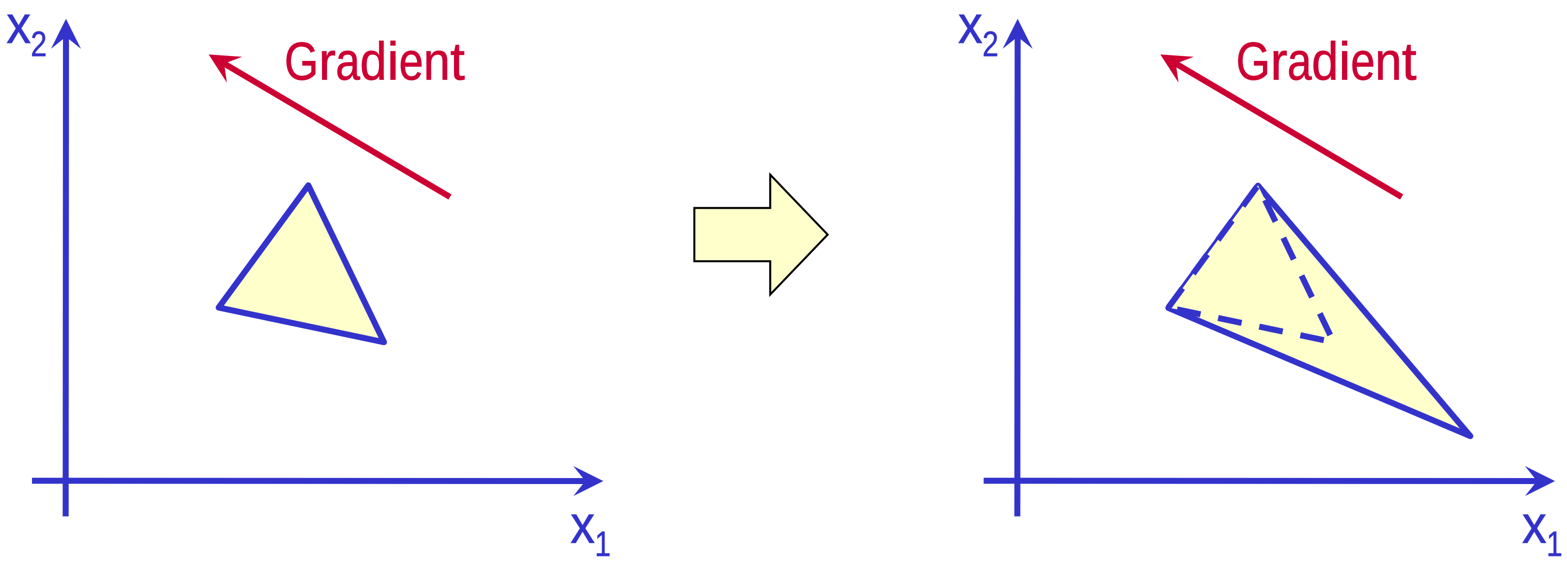
- Apply reflection, expansion and contraction



Reflect away from large cost function value

Downhill Simplex Method

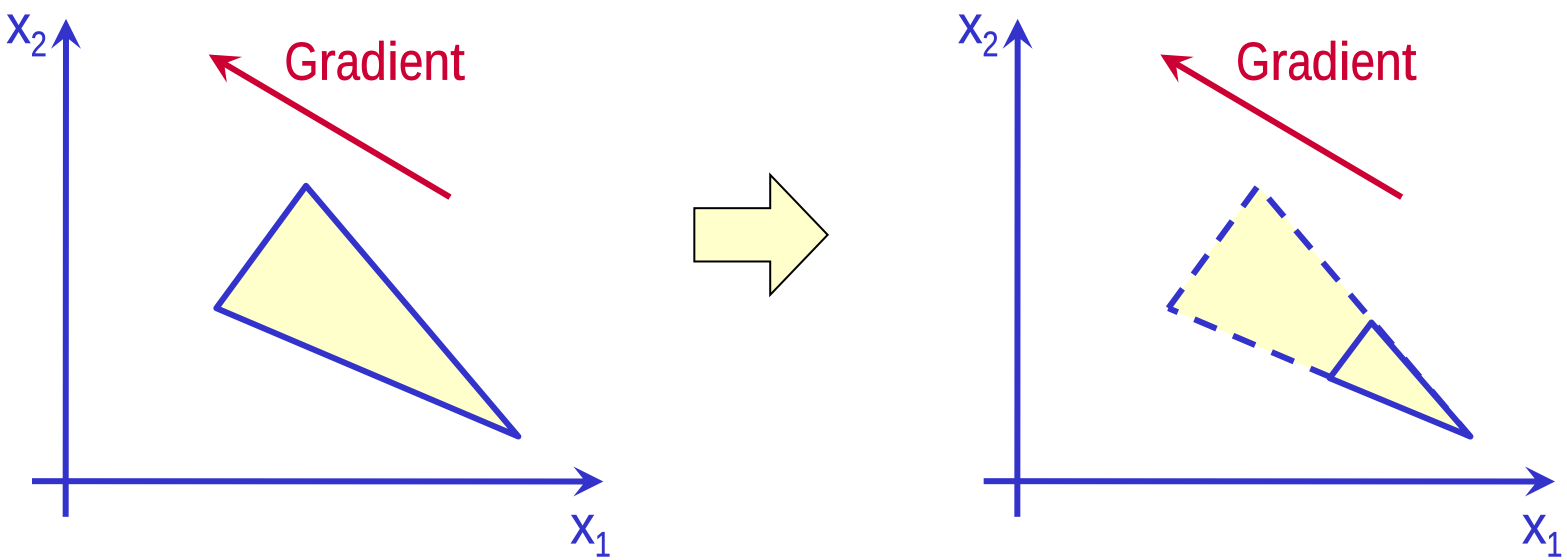
- Apply reflection, expansion and contraction



Expand towards small cost function value

Downhill Simplex Method

- Apply reflection, expansion and contraction



Contract towards small cost function value

Downhill Simplex Method

- Converge to a local minimum, if a sequence of reflection, expansion and contraction are appropriately applied
  - Slow convergence (especially in high-dimensional cases)
  - I.e., a large number of iteration steps

- If cost function is smooth, its derivative information can be used to search optimal solution
  - More details in future lectures

- However, downhill simplex method is still required if derivatives are not available or too difficult to calculate
  - Cost function is not smooth
  - Cost function contains sharp changes

Summary

- Unconstrained Optimization
  - Golden section search
  - Downhill simplex method