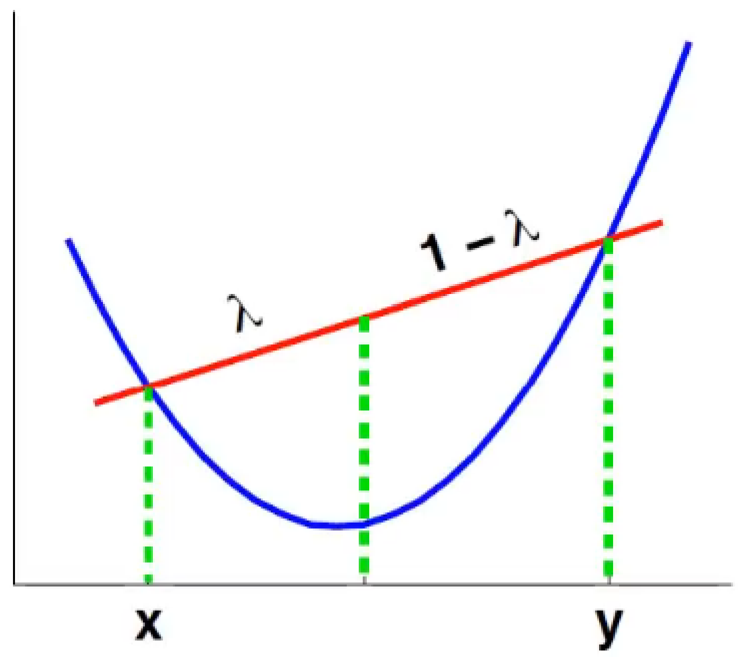
Convex Optimization

First let's define what's **Convex** function means, They take a graph shape :



- Where any two points line stays inside the function graph are called convex functions

Following this :

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y)$$

The Standard Form of the objective function what we will try to optimize will take this form :

$$\min f_0(x)$$
$$\text{s.t } f_i(x) < 0, i = 1, \ldots, m$$
$$h_i(x) = 0, i = 1, \ldots, p$$

- A point satisfy the constraints is called **Feasible set**
- The **Optimal value** p^* of the problem is defined as

$$p^* = \inf \{f_0(x) | x \text{ satisfies all constrains}\}$$

- x^* is an **optimal point** or a solution to the problem if x^* is feasible and $f(x^*) = p^*$

Lagrangian

For an optimization problem :

$$\min f_0(x)$$
$$\text{s.t } f_i(x) \leq 0$$

The **Lagrangian** for this optimization problem is :

$$\mathcal{L}(x, \lambda) = f_0(x) + \sum^{m} \lambda_i f_i(x)$$

- the second term is a **Linear Combination** of the constraint functions with weight λ_i
- They are called **dual variables**
- Simply Lagrangian removes the constrains and insert them into the function

So the Original optimization problem in **Primal Form** :

$$p^* = \min_x \max_{\lambda \geq 0} \mathcal{L}(x, \lambda)$$

- The X try to minimize the $\mathcal{L}(x, \lambda)$
- While the λ try to penalize if we don't meet the constraints

We get the **Lagrangian dual problem** by sapping the $\min$ and $\max$, So it makes the job of X minimizing the **convex** function easier :

$$d^* = \max_{\lambda \geq 0} \min_x \mathcal{L}(x, \lambda)$$

- Strong duality $\rightarrow p^* = d^*$
- Weak duality $\rightarrow p^* \geq d^*$

Lagrange dual function or **dual function** :

$$g(\lambda) = \min_x \mathcal{L}(x, \lambda) = \min_x \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \right)$$

The **Weak duality** can also be expressed as :

$$p^* = \max_{\lambda \geq 0} g(\lambda) = d^*$$

- λ is called **Dual Feasible** if :
  - $\lambda \geq_0$
  - $g(\lambda) > -\infty$
- Maximize $g(\lambda)$ subject to $\lambda \geq 0$
- The **Lagrange dual** makes the primal problem simple by removing the constraints and inserting them into the **Objective function**