Convex Optimization

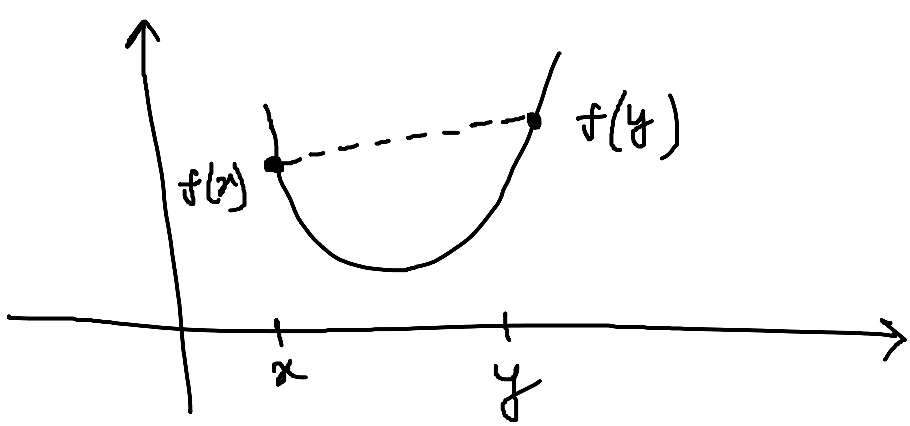
What are Convex Sets?

A subset S of a vector space is called convex, if $\forall\, x, y \in S$ and $\forall\, t \in [0,1]$

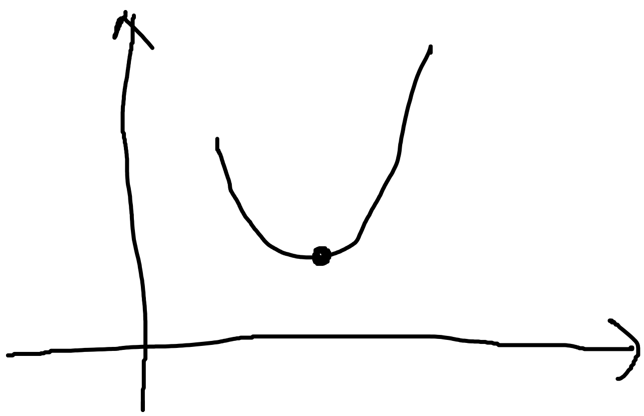
$$tx + (1-t)y \in S$$

Convex functions?

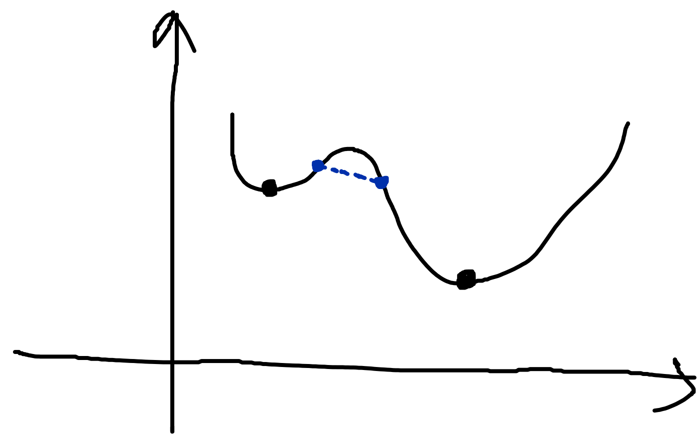
$$f(tx+(1-t)y) \leq t \cdot f(x) + (1-t)\, f(y)$$



Convex functions have only one local minimum, which is equal to the global minimum.



Convex



Non - Convex

Unconstrained Convex optimization -

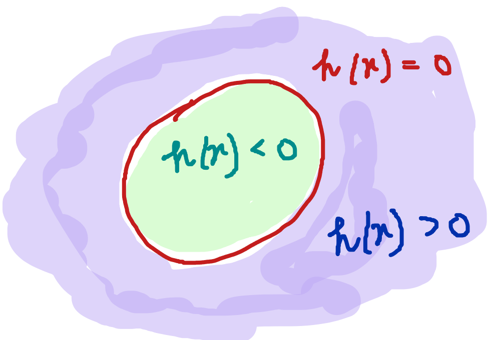
Minimize the convex function $f_0(x)$.

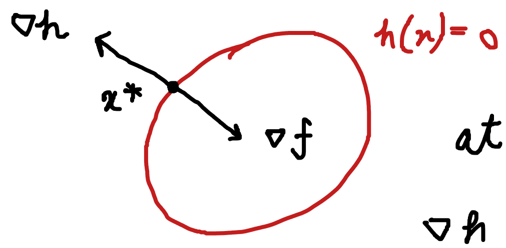
How will you solve this?

Equality Constraints -

Minimize $f(x)$ subject to $h(x) = 0$

$f(x)$ and $h(x)$ are convex functions.





at optimal point x^*, ∇h and ∇f would be parallel.

$$\nabla f + v \cdot \nabla h = 0$$

Lagrange Multiplier.

$$\text{Lagrangian} \Rightarrow L(x, v) = f(x) + v\, h(x)$$

We can find the optimal x^* by finding the critical points of the lagrangian.

$$\boxed{\nabla_{x,v} L(x, v) = 0}$$

General Constrained Optimization

minimize the convex function : $f_0(x)$ (objective function)

Subject to inequality constraints : $f_i(x) \leq 0$; $i \in \{1, 2, \cdots, m\}$

equality constraints : $h_j(x) = 0$; $j \in \{1, 2, \cdots, n\}$

① Lagrangian $\boxed{L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{n} \nu_j h_j(x)}$

$\lambda_i \geq 0$

② The constrained optimal value $f(x^*)$ is denoted by p^* (primal)

③ The dual function is defined as :-

$$\boxed{g(\lambda, \nu) = \min_{x} L(x, \lambda, \nu)}$$

g is a concave function.

④ The dual problem is defined as -

$$\boxed{\max_{\lambda, \nu} g(\lambda, \nu)}$$

where $\lambda_i \geq 0$

$\nu \in R$

⑤ Solution to dual problem is d^* (dual) attained at $g(\lambda^*, \nu^*)$

⑥ Weak duality $\boxed{d^* \leq p^*}$

Strong Duality & KKT conditions :

When $p^* = d^*$, we say strong duality holds

If strong duality holds, then $[x^*, \lambda^*, \nu^*]$ are saddle points of the Lagrangian

$\boxed{\text{KKT conditions are equivalent to strong duality.}}$

$$\mathcal{L}(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i \cdot f_i(x) + \sum_{j=1}^{n} \nu_j \cdot h_j(x)$$

1. $f_i(x^*) \leq 0 \quad \forall i \in \{1 \dots m\}$ (primal feasibility 1)
2. $h_j(x^*) = 0 \quad \forall j \in \{1 \dots n\}$ (primal feasibility 2)
3. $\lambda^* \geq 0$ (dual feasibility)
4. $\lambda_i^* \cdot f_i(x^*) = 0 \quad \forall i \in \{1 \dots m\}$ (Complementary Slackness)
5. $\nabla_x \mathcal{L}(x, \lambda, \nu) \big|_{x^*, \lambda^*, \nu^*} = 0$ (Stationary)

then $p^* = d^*$, and they are the optimal values.

How to use KKT conditions to solve problems?

Hint: Complementary Slackness.

Case 1 : $\lambda = 0$

- Term eliminated.
- Find saddle points of Lagrangian.
- The point which satisfies KKT is optimal.

Case 2 : $f_i(x) = 0$

- Solution at boundary.
- Replace inequality constraint with equality constraint.
- Find optimal point like before.

Need to try all 2^N cases.

Example 1

$$\begin{aligned} \text{minimize} \quad & -x-3y \\ \text{subject to} \quad & x+y=6 \\ & -x+y \leq 4 \end{aligned}$$

Langrangian function can be written as –

$$L(x,y,\lambda,\nu) = -x-3y + \lambda(-x+y-4) + \nu(x+y-6)$$

where $\lambda \geq 0$

First we will solve the primal problem :

From Complimentary Slackness, $\lambda \cdot f(x,y) = 0$

① $\boxed{\lambda = 0}$

$$L(x,y,\lambda,\nu) = -x-3y + \nu(x+y-6)$$

$$\left.\begin{aligned} \frac{\partial L}{\partial x} &= -1+\nu = 0 \\ \frac{\partial L}{\partial y} &= -3+\nu = 0 \end{aligned}\right\} \quad \nu = 1, 3$$

$$\left.\frac{\partial L}{\partial \nu} = x+y-6 = 0\right\} \quad (x,y) = (1,5), (-3,9), \ldots\ldots$$

satisfies KKT

does not satisfy KKT

$p^* = -16, \; x^* = 1, \; y^* = 5$

② $\boxed{f(x,y) = 0}$

$$L(x,y,\lambda,\nu) = -x-3y + \lambda(-x+y-4) + \nu(x+y-6)$$

$$\left.\begin{aligned} \frac{\partial L}{\partial \lambda} &= -x+y-4 = 0 \\ \frac{\partial L}{\partial \nu} &= x+y-6 = 0 \end{aligned}\right\} \quad x = 1,\ y = 5$$

$$\left.\begin{aligned} \frac{\partial L}{\partial x} &= -1-\lambda+\nu = 0 \\ \frac{\partial L}{\partial y} &= -3+\lambda+\nu = 0 \end{aligned}\right\} \quad \nu = 2,\ \lambda = 1$$

from both cases $\Rightarrow$ $\min f(x,y) = -16$

$x = 1,\ y = 5$

Now let's solve the dual problem:

The dual problem is

$$\max_{\lambda \geq 0, \nu} \left(\min_{x,y} L(x,y,\lambda,\nu) \right)$$

First let's solve the inner minimization $\rightarrow$

$$L(x,y,\lambda,\nu) = -x - 3y + \lambda(-x + y - 4) + \nu(x + y - 6)$$

$$\frac{\partial L}{\partial x} = -1 - \lambda + \nu = 0$$

$$\frac{\partial L}{\partial y} = -3 + \lambda + \nu = 0$$

$$\min_{x,y} L(x,y,\lambda,\nu) = x(-1-\lambda+\nu) + y(-3+\lambda+\nu) - 4\lambda - 6\nu$$
$$= -4\lambda - 6\nu$$

The dual problem becomes =

$$\boxed{\begin{array}{ll} \text{maximise} & g = -4\lambda - 6\nu \\ \text{subject to} & \begin{pmatrix} -1 - \lambda + \nu = 0 \\ -3 + \lambda + \nu = 0 \\ \lambda \geq 0 \end{pmatrix} \end{array}}$$

$\rightarrow$ required by Lagrangian

We can easily solve the linear constraints $\rightarrow$

$$-4 + 2\nu = 0$$
$$\nu = 2 \quad , \quad \lambda = 1$$

$$\therefore \; (\lambda^*, \nu^*) = (1, 2)$$

Maximal value of $g = -4 - 12$
$$= \boxed{-16}$$

$d^* = -16$ $\quad (p^* = d^*)$ $\quad$ Strong duality holds

Example 2

$$\text{minimize} \quad e^{-x}$$
$$\text{subject to} \quad \frac{x^2}{y} \le 0$$
$$\text{and domain } D = \{(x,y) \mid y > 0\} \quad x, y \in \mathbb{R}^2$$

Solution to primal problem :

$$L(x,y,\lambda) = e^{-x} + \lambda\left(\frac{x^2}{y}\right)$$

given $y > 0$, $\frac{x^2}{y} \le 0$ would be satisfied only for $x = 0$

$$\therefore \quad p^* = e^{-0} = 1 \;, \quad x^* = 0$$

Solution to dual problem :

$$\max_{\lambda}\left(\min_{x,y} L(x,y,\lambda)\right)$$

$$g(\lambda) = \min_{x,y} L(x,y,\lambda) = \begin{cases} 0 & \lambda \ge 0 \\ -\infty & \lambda < 0 \end{cases}$$

dual problem : maximize 0 subject to $\lambda \ge 0$ $\Rightarrow$ $d^* = 0 \;\; \forall \lambda \ge 0$

$\boxed{p^* \neq d^*}$ Strong duality does not hold

KKT?

Example 3

minimize $x^2 + y^2 + z^2$

subject to

$$x + y = 3$$
$$x - y = 3$$

Example 3

minimize $x^2 + y^2 + z^2$

subject to $x + y = 3$

$x - y = 3$

$$L(x, y, v_1, v_2) = x^2 + y^2 + z^2 + v_1(x + y - 3) + v_2(x - y - 3)$$

$\nabla_x L \Rightarrow 2x + v_1 + v_2 = 0$ $\boxed{v_1 = v_2 = -3}$

$\nabla_y L \Rightarrow 2y + v_1 - v_2 = 0$

$\nabla_z L \Rightarrow 2z = 0$ $\boxed{z = 0}$

$$\left.\begin{array}{l} \nabla_{v_1} L \Rightarrow x + y - 3 = 0 \\ \nabla_{v_2} L \Rightarrow x - y - 3 = 0 \end{array}\right\} \quad \boxed{x = 3} \quad \boxed{y = 0}$$

Example 4

$$\text{minimize} \quad f_0(x,y) = x^2 + y^2$$

$$\text{subject to} \quad x + y - 1 \leq 0$$

$$x - y + 2 \leq 0$$

Example 4

$$\text{minimize } f_0(x,y) = x^2 + y^2$$

subject to $\quad x + y - 1 \le 0$

$\quad x - y + 2 \le 0$

$$L(\,) = x^2 + y^2 + \lambda_1 (x + \overset{f_1}{y} - 1) + \lambda_2 (x - \overset{f_2}{y} + 2)$$

① $\lambda_1 = 0, \quad \lambda_2 = 0$

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

$\nabla_x L \Rightarrow \boxed{x = 0}$

$\nabla_y L \Rightarrow \boxed{y = 0}$

✗ $x - y + 2 \le 0$ does not hold

② $\lambda_1 = 0, \quad \lambda_2 \ne 0$

$\therefore f_2 = 0$

$L(x, y, \lambda_2) = x^2 + y^2 + \lambda_2 (x - y + 2)$

$\nabla_x L \Rightarrow 2x + \lambda_2 = 0$

$\nabla_y L \Rightarrow 2y - \lambda_2 = 0$

$\nabla_{\lambda_2} L \Rightarrow x - y + 2 = 0$

$\boxed{x = -1}, \ \boxed{y = 1}, \ \boxed{\lambda_2 = 2}$

③ $\lambda_1 \neq 0, \lambda_2 = 0$

$$\mathcal{L} = x^2 + y^2 + \lambda_1 (x+y-1)$$

$$\nabla_x \mathcal{L} \Rightarrow 2x + \lambda_1 = 0$$

$$\nabla_y \mathcal{L} \Rightarrow 2y + \lambda_1 = 0$$

$$\nabla_{\lambda_1} \mathcal{L} \Rightarrow x + y - 1 = 0$$

$\boxed{x = y = \frac{1}{2}}$ $\boxed{\lambda_1 = -1}$

$\lambda \geq 0$

does not hold

④ $\lambda_1 \neq 0, \lambda_2 \neq 0$

$$\mathcal{L} = x^2 + y^2 + \lambda_1 (x+y-1) + \lambda_2 (x-y+2)$$

$$\nabla_x \mathcal{L} \Rightarrow 2x + \lambda_1 + \lambda_2 = 0$$

$$\nabla_y \mathcal{L} \Rightarrow 2y + \lambda_1 - \lambda_2 = 0$$

$$\nabla_{\lambda_1} \mathcal{L} \Rightarrow x + y - 1 = 0$$

$$\nabla_{\lambda_2} \mathcal{L} \Rightarrow x - y + 2 = 0$$

$\boxed{x = -1/2}$ $\boxed{y = 3/2}$

$\boxed{\lambda_1 = -1}$ $\boxed{\lambda_2 = 2}$

$\lambda \geq 0$ does not hold.