Solving Convex Optimization problems (or, optimality in convex optimization)

Consider

$$\text{Minimize} \quad f_0(\underline{x})$$
$$\text{st} \quad f_i(\underline{x}) \leq 0 \qquad i = 1, 2, \ldots m$$
$$h_i(\underline{x}) = 0 \qquad i = 1, 2 \ldots p.$$

Assume domain $\neq \phi$ & $\exists$ an optimum f^*

Any x satisfying constraints is called a feasible pr.

Define the Lagrangian $L : \mathbb{V} \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

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

$\downarrow \; \downarrow$

dual variables / Lagrange multipliers

Lagrange dual function

$$g(\underline{\lambda}, \underline{\nu}) = \inf_{\underline{x} \in \mathcal{D}} L(\underline{x}, \underline{\lambda}, \underline{\nu})$$

$\swarrow$

$\cap \, \mathrm{dom}(f_i) \cap \mathrm{dom}(h_i)$

① $g(\underline{\lambda}, \underline{\nu})$ is always concave $(\forall f_i, h_i)$

② $g(\underline{\lambda}, \underline{\nu}) \leq f^*$ $\quad \forall \underline{\lambda} \geq \underline{0} \quad \forall \underline{\nu} \in \mathbb{R}^p$

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

If $\underline{x}$ satisfies constraints & $\lambda_i \geq 0 \ \forall i$,

$$L(\underline{x}, \underline{\lambda}, \underline{\nu}) \leq f_0(\underline{x})$$

$$g(\underline{\lambda}, \underline{\nu}) = \inf_{\underline{x} \in \mathcal{D}} L(\underline{x}, \underline{\lambda}, \underline{\nu}) \leq \inf_{\underline{x} \in \text{Constraint set}} L(\underline{x}, \underline{\lambda}, \underline{\nu})$$

$$\leq f^*$$

Dual optimization :

$$\underset{\underline{\lambda} \geq 0,\ \underline{\nu} \in \mathbb{R}^p}{\text{Maximize}} \ g(\underline{\lambda}, \underline{\nu}) \rightarrow g^*$$

$g^* \le f^*$ $\longrightarrow$ Weak duality.

$g^* = f^*$ $\longrightarrow$ Strong duality

Example

$$\text{Minimize} \quad \underline{c}^T \underline{x} \qquad \underline{x} \in \mathbb{R}^n$$

$$\text{ST} \quad A\underline{x} = \underline{b} \qquad A : p \times n$$

$$\underline{x} \geq 0 \qquad m = n$$

$$\underbrace{}$$

$$\underline{x} \geq 0$$

$$(A\underline{x} - \underline{b}) = \underline{0}$$

$$L(\underline{x}, \underline{\lambda}, \underline{\nu}) = \underline{c}^T \underline{x} + \sum_{i=1}^{p} \nu_i (\underline{a}_i \underline{x} - b_i)$$

ith row of A

$$+ \sum_{i=1}^{n} \lambda_i (-x_i)$$

$$= \underline{c}^T \underline{x} + \underline{\nu}^T (A\underline{x} - \underline{b}) - \underline{\lambda}^T \underline{x}$$

$$= (c^T - \lambda^T + \nu^T A)\,\underline{x} - \underline{\nu}^T \underline{b}$$

$$L(\underline{x}, \underline{\lambda}, \underline{\nu}) = (A^T\nu + c - \lambda)^T \underline{x} - \underline{\nu}^T \underline{b}$$

$$g(\underline{\lambda}, \underline{\nu}) = \inf_{\underline{x} \in \mathbb{R}^n} L(\underline{x}, \underline{\lambda}, \underline{\nu})$$

$$= \begin{cases} -\underline{\nu}^T \underline{b} & \text{if} \quad A^T \underline{\nu} + \underline{c} = \underline{\lambda} \\ -\infty & \text{else} \end{cases}$$

Dual problem:

Maximize $-\nu^T b$

ST $A^T \underline{\nu} + \underline{c} = \underline{\lambda}$

$\lambda \geq 0$

$\equiv$ Maximize $-\underline{\nu}^T \underline{b}$

ST $A^T \underline{\nu} + \underline{c} \geq 0$

Example:

$\min \; \underline{c}^T \underline{x}$

$A\underline{x} = \underline{b}$

$\underline{x} \geq 0$

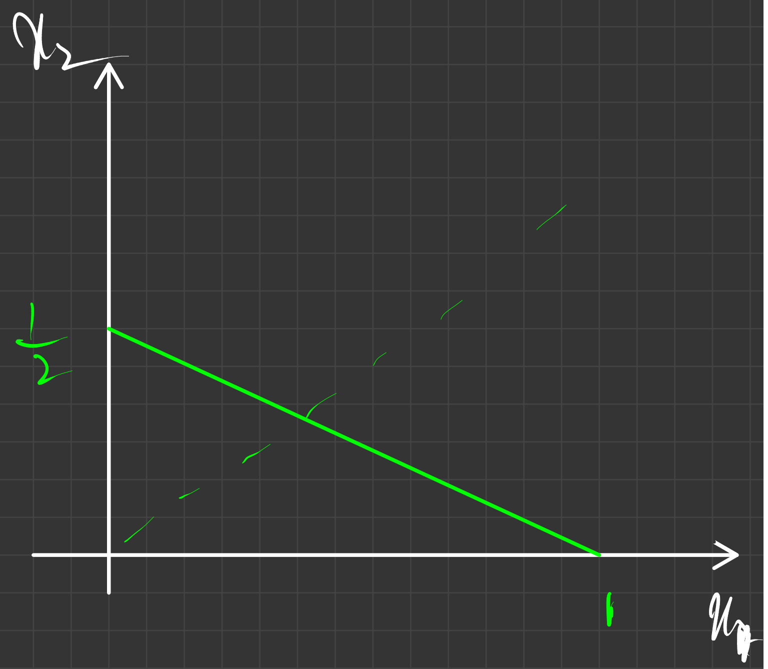
Minimize $x_1 + x_2 \rightarrow \frac{1}{2}$

ST

$x_1 + 2x_2 = 1$

$x_1 \geq 0$

$x_2 \geq 0$



$x_1 = 1 - 2x_2$

$\Rightarrow \quad f_0(\underline{x}) = 1 - 2x_2 + x_2$

$= 1 - x_2$

$f^* = \frac{1}{2}$

Dual: $\quad \text{Max} \; -\nu \quad \Rightarrow \quad g^* = \frac{1}{2}$

ST

$\begin{bmatrix}1\\2\end{bmatrix}\nu + \begin{bmatrix}1\\1\end{bmatrix} \geq 0 \Rightarrow \left.\begin{matrix}\nu \geq -1\\ 2\nu \geq -1\end{matrix}\right\} \Rightarrow \nu \geq -\frac{1}{2}$

KKT conditions (Karush, Kuhn, Tucker)

Solve for $\underline{x}^*, \underline{\lambda}^*, \underline{\nu}^*$ satisfying:

1. $f_i(x^*) \leq 0 \qquad i = 1, 2 \ldots m$

2. $h_i(x^*) = 0 \qquad i = 1, 2 \ldots p$

3. $\underline{\lambda}^* \succeq 0$

4. $\lambda_i^* f_i(x^*) = 0 \qquad i = 1, 2 \ldots m$

5. $$\nabla f_0(x^*) + \sum_{i=1}^{n} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{p} \nu_i^* \nabla h_j(x^*) = \underline{0}$$

In general:

(Slater's condition)

1. If prob satisfies strong duality, then KKT are necessary
2. If prob is convex (f_i's are convex & h_j are affine) KKT is sufficient

$$h_j(x) = x^2 - 2 = 0$$

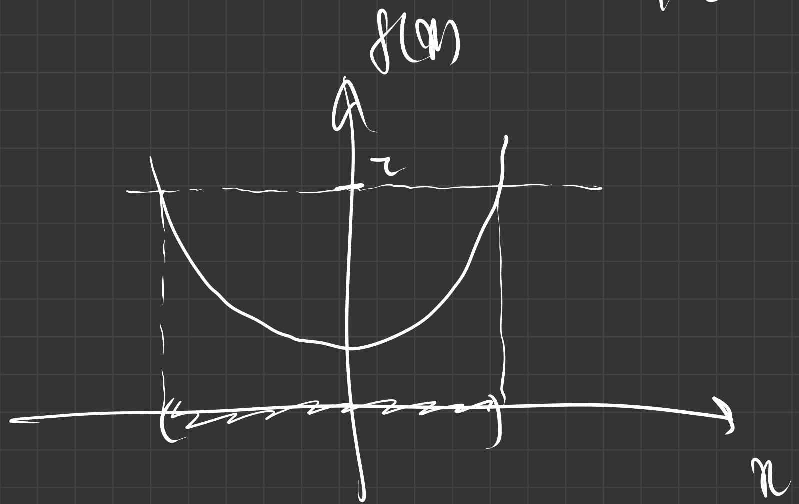
$$x^2 - 2 \le 0 \equiv x^2 \le 2$$

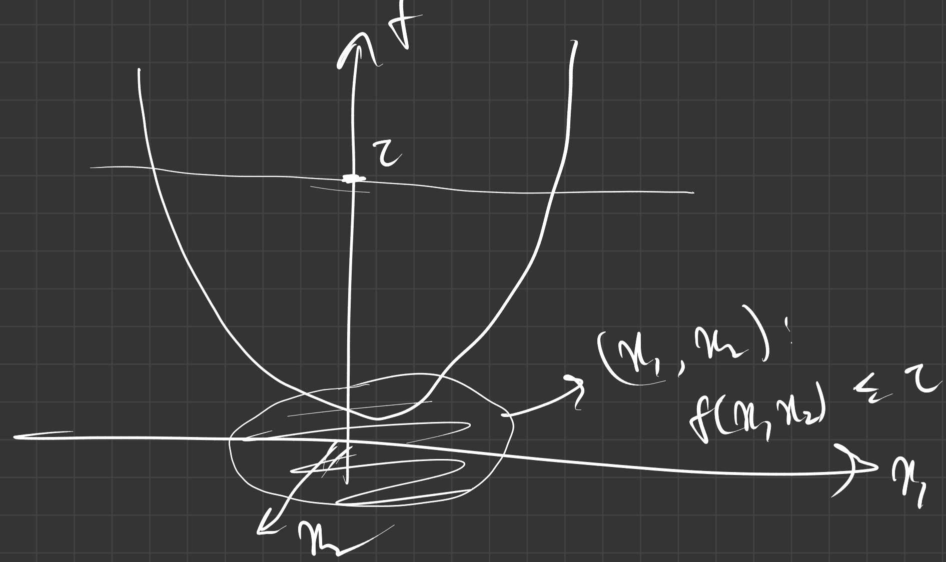
<u>Exercise</u>: If f is convex & $z \in \mathbb{R}$

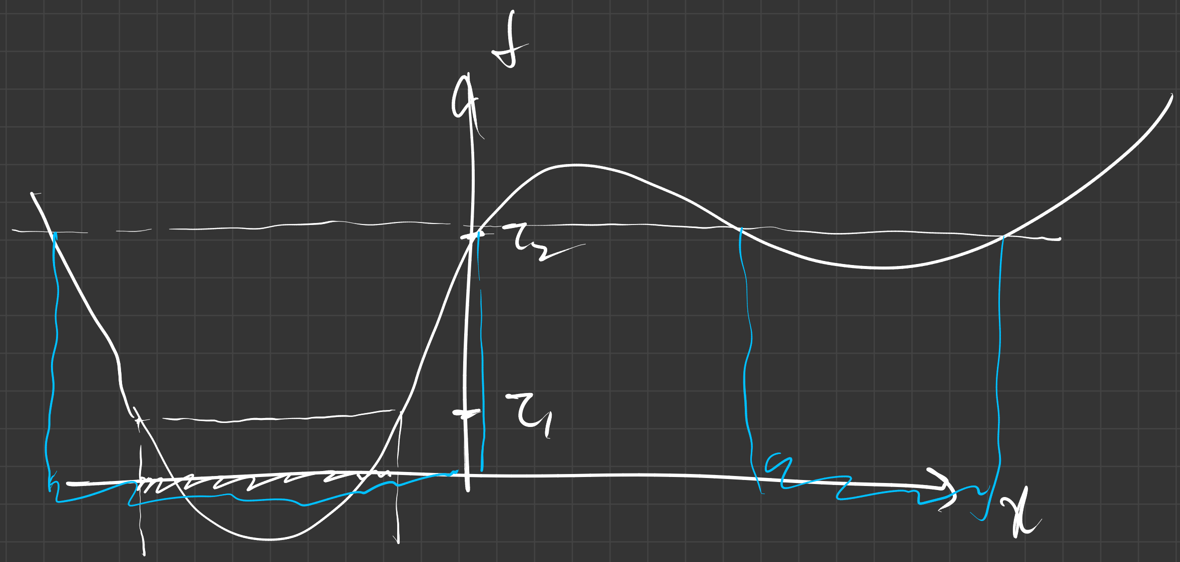
$\{ \underline{x} : f(\underline{x}) \le z \}$ is convex

If g is Concave,

$\{ \underline{x} : g(\underline{x}) \ge z \}$ is convex







Exercise: If f is both convex & concave, then it is affine

i.e, $f(x) = a^T x + b$ for some a, b

Hint: Use defn of convexity

Example 1 :

Minimize x^2 $\equiv$ Minimize x^2

ST $x \in [1, 2]$ ST

$-x+1 \leq 0 \rightarrow f_1(x) \leq 0$

$x-2 \leq 0 \rightarrow f_2(x) \leq 0$

① $f_i(x) \leq 0$

$-x+1 \leq 0$

$x-2 \leq 0$

② $\lambda \geq 0$ $\lambda_1 \geq 0$ $\lambda_2 \geq 0$

③ $\lambda_i f_i(x) = 0$ $\lambda_1(-x+1) = 0 \Rightarrow \lambda_1 = 0$ OR $x = 1$

$\lambda_2(x-2) = 0 \Rightarrow \lambda_2 = 0$ OR $x = 2$

④ $2x + \lambda_1(-1) + \lambda_2 = 0$

$2x = \lambda_1 - \lambda_2$

$\lambda_2 = 0$ & $x = 1$

$\lambda_1 = 2$

Example

$$\text{Minimize } \frac{1}{2}\underline{x}^T Q \underline{x} + b^T \underline{x} \qquad Q : n \times n \quad PSD$$

$$\text{ST} \quad A\underline{x} = \underline{0} \qquad A : p \times n$$

① $A\underline{x} = \underline{0}$

② $L(\underline{x}, \underline{\nu}) = f_0(x) + \sum_i \nu_i h_i(x) = f_0(x) + \underline{\nu}^T \underline{h}(\underline{x})$

$$= \underline{x}^T Q \underline{x} + \underline{b}^T \underline{x} + \underline{\nu}^T A \underline{x}$$

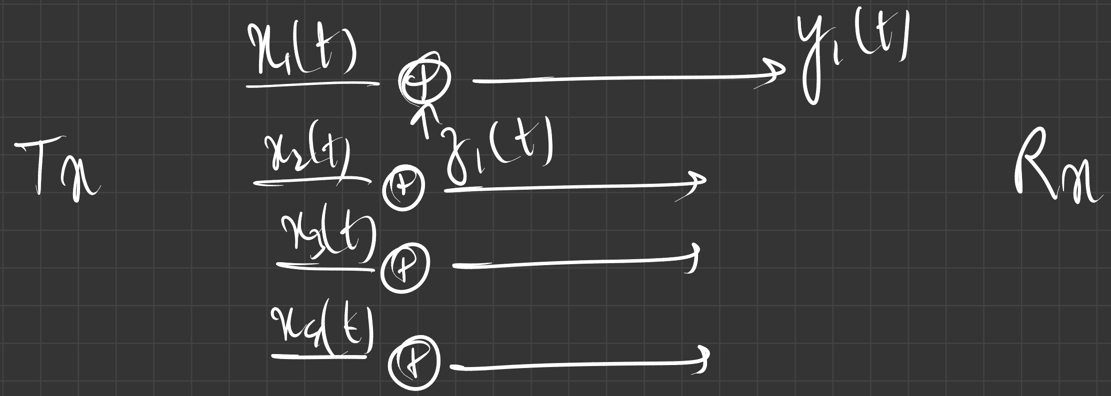
$$\nabla L(\underline{x}, \underline{\nu}) = \underline{0}$$

$$Q\underline{x} + \underline{b} + A^T \underline{\nu} = 0$$

$$A\underline{x} = 0$$

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{\nu} \end{bmatrix} = \begin{bmatrix} -\underline{b} \\ 0 \end{bmatrix}$$

Example 3



For any time instant t, $\forall$ i'th channel,

$$y_i(t) = x_i(t) + z_i(t)$$

$z_i(t)$: Gaussian, 0 mean, σ_i^2 var.

$$\frac{1}{T}\sum_{t=1}^{T} x_i^2(t) = P_i$$

Average power allocated to / used by ith channel

Want $\sum_{i=1}^{m} P \le P$ (Total power budget)

# of information bits that you can "reliably" send across the ith channel in T time slots

$$= T \frac{1}{2} \log_2\left(1 + \frac{P_i}{\sigma_i^2}\right)$$

$$\text{Rate} = \sum_{i=1}^{m} \frac{1}{2} \log\left(1 + \frac{P_i}{\sigma_i^2}\right)$$

$$\text{maximize}_{\underline{P}} \sum_{i=1}^{m} \log\left(1 + \frac{P_i}{\sigma_i^2}\right)$$

ST

$$\left.\begin{array}{l} \underline{P} \geq 0 \\ \sum_{i=1}^{m} P_i \leq P \end{array}\right\} \equiv \begin{array}{l} -P_i \leq 0 \quad \forall i \\ \sum_{i=1}^{m} P_i - P \leq 0 \end{array}$$

KKT conditions:

① $-p_i \le 0 \qquad p_i \ge 0 \quad \forall i \longrightarrow \lambda_1 \cdots \lambda_m$

$\sum_{i=1}^{m} p_i \le P \longrightarrow \lambda_{m+1}$

② $\underline{\lambda} \in \mathbb{R}^{m+1} \qquad \underline{\lambda} \succeq 0$

③ $\lambda_i(-p_i) = 0 \quad \forall i \in \{1, \cdots m\}$

$\lambda_i = 0 \quad \text{OR} \quad p_i = 0$

$\lambda_{m+1}\left(\sum_{i=1}^{m} p_i - P\right) = 0$

$\underline{\lambda_{m+1} = 0}$ ✗ $\quad \text{OR} \quad \boxed{\sum_{i=1}^{m} p_i = P}$ ✓

④ $L(\underline{p}, \underline{\lambda}) = -\sum_{i=1}^{m} \log\left(1 + \frac{p_i}{\sigma_i^2}\right) + \sum_{i=1}^{m} \lambda_i(-p_i) + \lambda_{m+1}\left(\sum_i p_i - P\right)$

$\nabla_{\underline{p}} L = 0 \Rightarrow$

$$\forall i, \quad -\frac{1}{1+\frac{P_i}{\sigma_i^2}} \frac{1}{\sigma_i^2} + (-\lambda_i) + \lambda_{m+1} = 0$$

$$-\frac{1}{P_i+\sigma_i^2} = \lambda_i - \lambda_{m+1} \quad \forall i$$

At least one $P_{i'} > 0 \Rightarrow \lambda_{i'} = 0$

$$-\frac{1}{P_{i'}+\sigma_{i'}^2} = -\lambda_{m+1}$$

$$P_{i'}+\sigma_{i'}^2 = \frac{1}{\lambda_{m+1}}$$

If $P_i = 0$

$$\frac{1}{\sigma_i^2} = \lambda_i - \lambda_{m+1}$$

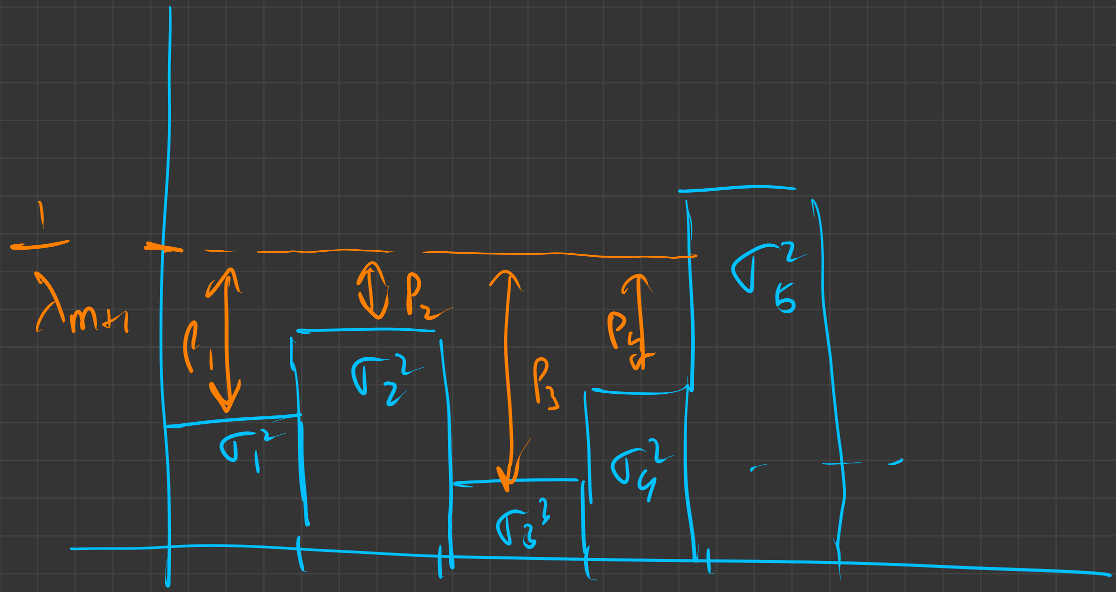
$$\lambda_i = \frac{1}{\sigma_i^2} + \lambda_{m+1}$$

If $P_i > 0$, $\quad P_i = \frac{1}{\lambda_{m+1}} - \sigma_i^2$

$$\sum_{i=1}^{m} P_i = P$$

Choose $\quad P_i = \max\left\{0, \frac{1}{\lambda_{m+1}} - \sigma_i^2\right\}$

$$\sum_{i=1}^{m} P_i = P$$



Waterfilling solution