Duality

$p^* = \text{minimize} \quad f_0(x)$

$\text{s.t} \quad f_i(\vec{x}) \leq 0 \quad \forall i \quad 1 \leq i \leq m$

$\qquad h_i(\vec{x}) = 0 \quad \forall i \quad 1 \leq i \leq p$

} primal

Lagrangian

$$L(\vec{x}, \vec{\lambda}, \vec{\nu}) = f_0(\vec{x}) + \sum_{i=1}^{m} \lambda_i f_i(\vec{x}) + \sum_{i=1}^{p} \nu_i h_i(\vec{x}) \qquad \lambda_i \geq 0$$

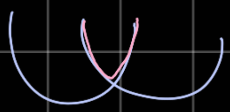
↳ affine fcn of $\vec{\lambda}$ & $\vec{\nu}$

=> convex fcn of $\vec{\lambda}$ & $\vec{\nu}$

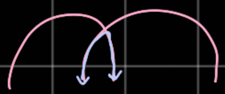
$$\underset{\vec{x}}{\text{minimize}} \quad L(\vec{x}, \vec{\lambda}, \vec{\nu}) := g$$
$$\text{s.t} \quad \vec{\lambda} \geq 0$$

↳ let's examine g

Recall



↳ pointwise max of convex fcns is convex



↳ pointwise min of concave fcns is concave

Properties

1. g is a fcn of only $\vec{\lambda}, \vec{\nu}$
2. $L(\vec{x}, \vec{\lambda}, \vec{\nu})$ is an affine fcn of $\vec{\lambda}, \vec{\nu}$
3. what kind of fcn is g? what can we say about g?

   ↳ g is a pointwise min of concave (affine) fcns

   recall affine both conv & conc.

   => g is a concave fcn of $\vec{\lambda}, \vec{\nu}$

$$g(\mu, \lambda)$$

$$g := \min \begin{cases} L(\vec{x}_1, \vec{\lambda}, \vec{\nu}) = k_1(\vec{\lambda}, \vec{\nu}) \\ L(\vec{x}_2, \vec{\lambda}, \vec{\nu}) = k_2(\vec{\lambda}, \vec{\nu}) \\ \vdots \\ L(\vec{x}_n, \vec{\lambda}, \vec{\nu}) = k_n(\vec{\lambda}, \vec{\nu}) \end{cases}$$

4. $g(\vec{\lambda}, \vec{\nu})$ is a lower bound on the primal optimal p^*

<u>Proof</u>: $\forall \vec{\lambda} \geq 0, \vec{\nu}$ $\quad g(\vec{\lambda}, \vec{\nu}) \leq p^*$

Consider some $\tilde{\vec{x}}$ that's feasible for the primal, ie that

$$f_i(\tilde{\vec{x}}) \leq 0 \qquad h_i(\tilde{\vec{x}}) = 0$$

$$L(\tilde{\vec{x}}, \vec{\lambda}, \vec{\nu}) = f_0(\tilde{\vec{x}}) + \underbrace{\sum \lambda_i f_i(\tilde{\vec{x}})}_{\leq 0} + \underbrace{\sum \nu_i h_i(\tilde{\vec{x}})}_{=0}$$

$$\leq f_0(\tilde{\vec{x}})$$

↳ recall g is the smallest of all of these ⇒ g a lower bound

$$g(\vec{\lambda}, \vec{\nu}) = \min_{\tilde{\vec{x}}} L(\tilde{\vec{x}}, \vec{\lambda}, \vec{\nu}) \Rightarrow g(\vec{\lambda}, \vec{\nu}) \leq p^*$$

Alternate form of Lagrangian:

$$M(\vec{x}) = f_0(\vec{x}) + \sum \mathbb{1}_{\leq 0}\{f_i(\vec{x})\} + \sum \mathbb{1}_0(h_i(\vec{x}))$$

indicator fcn

doing some "hard thresholding"

$$\mathbb{1}_{\leq 0}\{f_i(\vec{x})\} = \begin{cases} 0 & \text{if } f_i(\vec{x}) \leq 0 \\ \infty & \text{if } f_i(\vec{x}) > 0 \end{cases}$$

$$\mathbb{1}_0\{h_i(\vec{x})\} = \begin{cases} 0 & \text{if } h_i(\vec{x}) = 0 \\ \infty & \text{if } h_i(x) \neq 0 \end{cases}$$

<u>Let's compare</u>:

$$\sum \mathbb{1}_0\{f_i(x)\} \longleftrightarrow \sum \lambda_i f_i(\vec{x})$$

if $f_i(\vec{x}) \leq 0$ $\quad 0 \quad \geq \quad \leq 0$ → indicator provides linear lower bound

if $f_i(\vec{x}) > 0$ $\quad \infty \quad >$

Example:

$\min \vec{x}^T \vec{x} = p^*$ $\quad A :=$ wide matrix $\in \mathbb{R}^{m \times n}$ $\quad m < n$

s.t. $A\vec{x} = \vec{b}$

$A\vec{x} - \vec{b} = \vec{0}$

no inequality constraints

$$L(\vec{x}, \vec{\nu}) = f_0(\vec{x}) + \sum \lambda_i f_i(\vec{x}) + \sum \nu_i h_i(\vec{x})$$

$$= \vec{x}^T \vec{x} + \sum \nu_i (A\vec{x} - \vec{b})$$

$$= \vec{x}^T \vec{x} + \vec{\nu}^T (A\vec{x} - \vec{b})$$

quadratic term

linear term

convex!

↳ gradient to find min for g

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

$$\nabla_x L = 2\vec{x} + A^T \vec{x}$$

set eq to 0

$$0 = 2\vec{x} + A^T \vec{\nu}$$

$$\vec{x} = -\frac{1}{2} A^T \vec{\nu}$$

$$g(\vec{\nu}) = L\left(-\frac{1}{2} A^T \vec{\nu}, \vec{\nu}\right) = \left(\frac{1}{4} \nu^T A A^T \nu - \nu^T \left(\frac{A A^T \vec{\nu}}{2} + \vec{b} \right) \right.$$

$$= -\frac{1}{4} \vec{\nu}^T A A^T \vec{\nu} - \vec{\nu}^T \vec{b} \quad \Big\} \text{ concave!}$$

$\forall \vec{\nu}$ $g(\vec{\nu})$ is a lower bound on p^*

↳ what is the largest lower bound?

↳ want to maximize $g(\vec{\nu})$

$$\max_{\vec{\nu}} g(\vec{\nu})$$

Note: Dual problem : maximization of concave fcn with linear constraints

$$\max_{\substack{\lambda \geq 0 \\ \vec{\nu}}} g(\vec{\lambda}, \vec{\nu}) = d^*$$

↳ this is a $\boxed{\text{convex program}}$

↳ Note:

1. # of variables = # of constraints of primal
2. Always a convex fcn, even if primal is not

$$d^* \leq p^* \quad \Big\} \text{ weak duality}$$

Continuing w/example:

$$\max_{\vec{\nu}} g(\vec{\nu})$$

$$\nabla_{\vec{\nu}} g(\vec{\nu}) = -\frac{1}{4}(2 A A^T \vec{\nu}) - \vec{b}$$

$$\vec{\nu}^* = -2(A A^T)^{-1} \vec{b}$$

$$\vec{x}^* = -\frac{1}{2}(A^T \vec{\nu}) = -\frac{1}{2} A^T (-2 A A^T)^{-1} \vec{b}$$

$$= A^T (A A^T)^{-1} \vec{b} \quad \Big\} \text{ min norm solution!}$$

⇓

$$g(\vec{\nu}) \leq p^* \quad \forall \vec{\nu}$$

strong duality (when $p^* = g(\vec{\nu})$)

Partitioning Problem:

$$\min \quad x^T W \vec{x} \qquad W \in S^n$$

W → weight matrix (happiness of each person to be in the +1 or -1 group)

s.t. $x_i^2 = 1$

$i = 1, \dots, n$

→ $x_i = \pm 1$ ← non-convex domain

⇓

NOT CONVEX

$$L(\vec{x}, \vec{\nu}) = \vec{x}^T W \vec{x} + \sum_{i=1}^{n} \nu_i (x_i^2 - 1)$$

$$= \vec{x}^T W \vec{x} + \vec{x}^T \operatorname{diag}(\vec{\nu}) \vec{x} - \sum_{i=1}^{n} \nu_i$$

$$= \vec{x}^T (W + \operatorname{diag}(\vec{\nu})) \vec{x} - \sum_{i=1}^{n} \nu_i$$

$$\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \nu_1 & & & \\ & \nu_2 & & \\ & & \ddots & \\ & & & \nu_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$g(\vec{\nu}) = \min_{\vec{x}} \; L(\vec{x}, \vec{\nu})$$

$$= \begin{cases} -\sum_{i=1}^{n} \nu_i & \text{if } W + \operatorname{diag}(\vec{\nu}) \succeq 0 \text{ (PSD)} \\ -\infty & \text{o/w (if it's negative semidefinite)} \end{cases}$$

dual:

$$\max \quad -\sum_{i=1}^{n} \nu_i$$

$$W + \operatorname{diag}(\vec{\nu}) \succeq 0 \quad \Big\} \text{ SDP (semi-definite problem)}$$

choose $\vec{\nu} = \lambda_{min}(W)$ ← largest $\vec{\nu}$ that we can choose that preserves PSD

$$p^* \geq n \lambda_{min}(W)$$