Lecture 7: Duality (Section 5.1, 5.2.1 – 5.2.2)

Optimization Problem

$$\min_{x \in \mathbb{R}^n} f_0(x) \qquad \begin{array}{lll} f_i(x) \leq 0 & \forall i = 1, \dots, m & \text{Inequality} \\ h_i(x) = 0 & \forall i = 1, \dots, p & \text{Equality} \end{array}$$

Domain $\mathcal{D} = \bigcap_{i=0}^{m+1} \text{dom } f_i \cap \bigcap_{i=1}^{p} \text{dom } h_i$

Optimal value $p^* = \inf\{f_0(x) : x \text{ feasible}\}$

$$x \text{ feasible} \equiv \begin{array}{ll} f_i(x) \leq 0 & \forall i = 1, \dots, m \\ h_i(x) = 0 & \forall i = 1, \dots, p \end{array}$$

main idea of duality

Idea: "Approximate" the original constrained problem to an unconstrained problem.

objective

inequality constraint

new function is linear combi of all the functions seen before

Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$.

equality constraints.

original decision variable

vector of Lagrange multipliers associated to inequality constrains

$$L(x, \lambda, v) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x)$$

$\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^p$ $\quad v = (v_1, \dots, v_p) \in \mathbb{R}^p$

vector of Lagrange multiplier associated with equality constrains

$\text{dom } L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$

λ_i: Lagrange multiplier associated to i^{th} ineq. constraint $f_i(x) \leq 0$

ν_i: Lagrange multiplier associated to i^{th} eq. constraint $h_i(x) = 0$

$\underset{\in \mathbb{R}^m}{\lambda}, \underset{\in \mathbb{R}^p}{\nu}$: dual variables / Lagrange multiplier vectors.

Def: Lagrange dual function (concave) $g: \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ → lower bound to p^*

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

$$= \inf_{x \in \mathcal{D}} \left\{ f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right\}$$

Claim: g is a concave function. (does not matter what original function f are)

$\forall (\lambda, \nu), (\lambda', \nu')$

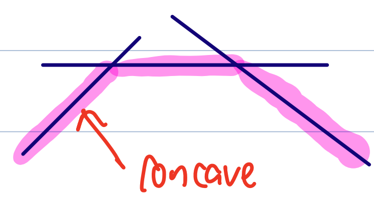
$$g(\theta\lambda + (1-\theta)\lambda', \theta\nu + (1-\theta)\nu') \geq \theta g(\lambda, \nu) + (1-\theta) g(\lambda', \nu')$$

Pf: $(\lambda, \nu) \mapsto f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$ is affine

for any $\{f_i\}_{i=0}^{m}$ & $\{h_i\}_{i=1}^{p}$.

g is the pointwise inf. of a bunch of affine functions.

$\Rightarrow g$ is concave.



The Lagrangian $L(x,\lambda,\nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \cancel{\sum_{i=1}^{p} \nu_i h_i(x)}$
is a linear approx of original opt. problem.
(without equality constrain)

p^*: $\min_{x \in \mathbb{R}^n} f_0(x)$ s.t. $f_i(x) \leq 0$, $\forall i = 1, \dots, p$
$\lambda \geq 0 \Rightarrow \lambda_i \geq 0$, $\forall i = 1, \dots, m$.

no feasible solution (not inside the domain of optimisation problem)
If the i^{th} constraint is violated, i.e., $f_i(x) > 0$, then
$p^* = \infty$

Lagragian Value of Lagrangian $p^* + \underbrace{\underset{\geq 0}{\lambda_i}\, \underset{\text{positive}}{f_i(x)}}_{\text{linear penalty}}$ is positively penalized.
= approximation to the original problem which was penalising any violation infinitely much

not exactly ∞ but larger than p^*

Lower bounds on optimal value

$\swarrow$ primal
$p^* = \inf\{f_0(x) : x \text{ feasible}\}$.

$$x \text{ feasible} \equiv \begin{array}{ll} f_i(x) \leq 0 & \forall i = 1, \dots, m \\ h_i(x) = 0 & \forall i = 1, \dots, p \end{array}$$

Claim: For any $\lambda \geq 0$, $\nu \in \mathbb{R}^p$,
lower bound to p^*
$$g(\lambda, \nu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \nu) \leq p^*.$$
(precursor to weak duality)

Pf: Suppose $\tilde{x} \in \mathbb{R}^n$ is a feasible point of original problem
$f_i(\tilde{x}) \leq 0 \quad \forall i = 1, \dots, m$
$h_i(\tilde{x}) = 0 \quad \forall i = 1, \dots, p$
inequality and equality constrains all satisfied.

Then consider the Lagrangian:

$$L(\tilde{x},\lambda,v) = f_0(\tilde{x}) + \underbrace{\sum_{i=1}^{m} \overset{\lambda_i \geq 0}{\lambda_i} f_i(\tilde{x})}_{\leq 0} + \underbrace{\sum_{i=1}^{p} v_i h_i(\tilde{x})}_{=0}$$

$$\leq f_0(\tilde{x}) \quad \text{for any feasible } \tilde{x}$$

$$g(\lambda,v) = \inf_{x\in D} L(x,\lambda,v) \qquad \text{(by def}^{\underline{n}} \text{ of dual f}^{\underline{n}})$$

$$\leq L(\tilde{x},\lambda,v) \leq f_0(\tilde{x}) \quad \text{for any feasible } \tilde{x}$$

$$g(\lambda,v) \leq f_0(\tilde{x})$$

Minimize the RHS w.r.t. every feasible $\tilde{x}$

$\forall \lambda \geq 0, v \in \mathbb{R}^p$

$$g(\lambda,v) \leq \inf\{f_0(\tilde{x}) : \tilde{x} \text{ feasible}\} = p^* \qquad ///$$

Weak duality : $\displaystyle\max_{\lambda\geq 0, v\in\mathbb{R}^p} g(\lambda,v) \leq p^*$

$$\|$$

$$d^* \leq p^*$$

<u>Example:</u>

Least Squares sol$^{\underline{n}}$ of linear equations. $\qquad$ 1° equality constraints

least norm problem $\qquad \displaystyle\min_{x\in\mathbb{R}^n} \|x\|^2$ s.t. $Ax=b$. $\qquad A\in\mathbb{R}^{p\times n}$

$$Ax-b=0$$

Lagrangian : $L(x, v) = x^T x + v^T(Ax - b)$

$\text{dom } L = \mathbb{R}^n \times \mathbb{R}^p$

$\nabla_x \; c^T x = c$

only one argument as there is no inequality constrains

$g(v) = \inf_{x \in D} \underbrace{x^T x + v^T(Ax - b)}_{L}$

$\nabla_x [x^T x + \underbrace{v^T(Ax}_{(A^T v)^T x} - b)] = 2x + A^T v = 0$

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

sub x back into $L(x, v)$ function

Lagrange dual f^n: $g(v) = \inf_{x \in D} L(x, v)$

$g(v) = (-\frac{1}{2}A^T v)^T(-\frac{1}{2}A^T v) + v^T(A(-\frac{1}{2}A^T v) - b)$

$= -\frac{1}{4} v^T A A^T v - v^T b$

Is this concave in v? Yes! $\because v \mapsto -\frac{1}{4} v^T \overset{\text{PSD.}}{AA^T} v$ is concave

$g(v)$: concave quadratic f^n.

Lower bd property

$\underbrace{-\frac{1}{4} v^T A A^T v - v^T b}_{g(v) \leq p^*} \leq p^* = \inf\{x^T x : Ax = b\}$

for all $v \in \mathbb{R}^p$

$=$ strong duality

$\sup_{v \in \mathbb{R}^p} \{-\frac{1}{4} v^T A A^T v - v^T b\} \leq p^* = \inf\{x^T x : Ax = b\}$

holds for all so it holds for the maximum

Example: Standard form LP

$-x \leq 0 \Leftrightarrow -x_i \leq 0 \quad \forall i = 1, \dots, m$

$$\min_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad Ax = b, \; x \geq 0$$

Lagrangian: $L(x, \lambda, \nu) = c^T x + \sum_{i=1}^{m} \lambda_i (-x_i) + \nu^T \underbrace{(Ax - b)}_{\sum_{i=1}^{p} \nu_i (a_i^T x - b_i)}$

$= c^T x - \lambda^T x + \nu^T (Ax - b)$

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

Lagrange dual function: $g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu)$

minimum

$$g(\lambda, \nu) = \inf_{x \in D} \{ -b^T \nu + (\underbrace{c + A^T \nu - \lambda}_{z})^T x \}$$

i) $c + A^T \nu - \lambda = 0 \Rightarrow g(\lambda, \nu) = -b^T \nu$ nothing to minimise, this is a constant

ii) $\underbrace{c + A^T \nu - \lambda}_{z \neq 0} \neq 0 \Rightarrow$ Take $x = -tz$, $t > 0$

$\Rightarrow (c + A^T \nu - \lambda)^T x = z^T(-tz) = -t \|z\|^2$ not 0

Take $t \to \infty$, $g(\lambda, \nu) = -\infty$.

In summary, $g(\lambda, \nu) = \begin{cases} -b^T \nu & c + A^T \nu - \lambda = 0 \\ -\infty & c + A^T \nu - \lambda \neq 0 \end{cases}$

note: dual of an LP is an LP

Connection of dual/Lagrangian to conjugate function

$f: \mathbb{R}^n \to \mathbb{R}$

$f^*: \mathbb{R}^n \to \mathbb{R}$ $\quad f^*(y) = \sup_{x \in \mathrm{dom} f} \{x^T y - f(x)\}$

f^* & $g(\lambda, \nu)$ are easily related.

$x_i = 0 \ \forall i = 1, \dots, n$

Consider the opt. problem : $\min_{x \in \mathbb{R}^n} f_0(x)$ s.t. $x = 0$.

Lagrangian : $L(x, \nu) = f_0(x) + \nu^T x$ $= \sum_{i=1}^{n} \nu_i x_i$

Lagrange dual function : $g(\nu) = \inf_{x \in D} \{ f_0(x) + \nu^T x \}$

$$
\begin{aligned}
g(\nu) &= - \sup_{x \in D} \{ - f_0(x) - \nu^T x \} \\
&= - \sup_{x \in D} \{ x^T(-\nu) - f_0(x) \} \\
&= - f_0^*(-\nu)
\end{aligned}
$$

Lagrange dual $\Leftrightarrow$ conjugate up to all these multiplier

More general result : $\min_{x \in \mathbb{R}^n} f_0(x)$ s.t. $Ax \le b$, $Cx = d$

$$g(\lambda, \nu) = -b^T\lambda - d^T\nu - f_0^*(-A^T\lambda - C^T\nu)$$

Examples: Equality constrained norm min

$$\min_{x \in \mathbb{R}} \underbrace{\|x\|_2}_{= f_0(x)} \quad \text{s.t.} \quad \underline{Ax = b}.$$

dual function domain

$f_0(x) = \|x\|_2 \Rightarrow f_0^*(y) = \begin{cases} 0 & \text{if } \|y\| \le 1 \\ +\infty & \text{else.} \end{cases}$

Given the form of the conjugate, the Lagrange dual f^n is

$$\boxed{g(\lambda, v) = -\cancel{b^T\lambda} - d^T v - f_0^*(-\cancel{A^T\lambda} - C^T v)}$$

no inequality constraint

$Cx = d.$

$$g(v) = -b^T v - f_0^*(-A^T v)$$

$$= \begin{cases} -b^T v & \|-A^T v\| \le 1 \quad (-f_0^*(-A^T v) = 0) \\ -\infty & \|-A^T v\| > 1 \end{cases}$$

$$g(v) = \begin{cases} -b^T v & \|A^T v\| \le 1 \\ -\infty & \|A^T v\| > 1 \end{cases}$$

<u>Entropy Maximization</u> $\sim \max H(x) \quad H(x) = -\sum x_i \log x_i$

$$\min_{x \in \mathbb{R}^n} f_0(x) = \sum_{i=1}^n x_i \log x_i$$

$\text{dom } f_0 = \mathbb{R}^n_{++}$ (otherwise cannot take log)

s.t. $Ax \le b, \quad \mathbf{1}^T x = 1.$

$C^T x = d$

$f_0^*(y) = \sum_{i=1}^n e^{y_i - 1} \quad \text{dom } f^* = \mathbb{R}^n$

$$\boxed{g(\lambda,\nu) = -b^T\lambda - d^T\nu - f_0^*(-A^T\lambda - C^T\nu)}$$

$$g(\lambda,\nu) = -b^T\lambda - \nu - f_0^*(-A^T\lambda - \nu)$$
$$= -b^T\lambda - \nu - \sum_{i=1}^{n} e^{-a_i^T\lambda - \nu - 1}$$

<u>Lagrange Dual Problem</u> - optimisation problem

We know that $g(\lambda,\nu) \leq p^* \;\; \forall \lambda \geq 0, \nu \in \mathbb{R}^p$.
We want the tightest lower bound.
Why don't we conside maximizing the LB over all $\lambda \geq 0$, $\nu \in \mathbb{R}^p$

g is concave

Consider: $\max_{\lambda \geq 0, \nu \in \mathbb{R}^p} g(\lambda,\nu)$
This called the <u>Lagrange Dual Problem</u>

Fact: The <u>Lagrange Dual Problem</u> is a convex opt. prob.

(λ,ν): <u>dual feasible</u> if $\lambda \geq 0$ ($\lambda_i \geq 0 \; \forall i=1,\dots,m$), $\nu \in \mathbb{R}^p$
all components non -ve

(λ^*,ν^*): <u>dual optimal</u> if it is dual feasible &

maximises the Lagrange dual function

$$g(\lambda^*, v^*) = \max_{\lambda \geq 0, v \in \mathbb{R}^p} g(\lambda, v)$$

Example: Lagrange dual of a standard LP

$$\min_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad \begin{array}{l} Ax = b \\ x \geq 0 \end{array}$$

Lagrange dual function.

$$g(\lambda, v) = \begin{cases} -b^T v & A^T v - \lambda + c = 0 \\ -\infty & \text{else.} \end{cases}$$

Lagrange dual problem:

$$\max_{\lambda \geq 0, v \in \mathbb{R}^p} g(\lambda, v)$$

$$\equiv \max_{\lambda \geq 0, v \in \mathbb{R}^p} -b^T v \quad \text{s.t.} \quad A^T v - \lambda + c = 0$$

$$\lambda = A^T v + c$$

constrain to be non-negative

$$\equiv \max_{v \in \mathbb{R}^p} -b^T v \quad \text{s.t} \quad A^T v + c \geq 0$$

$$\equiv \min_{v \in \mathbb{R}^p} b^T v \quad \text{s.t.} \quad A^T v + c \geq 0$$

This is an LP but in inequality form.

Weak Duality :

We know that $\forall \lambda \geq 0, v \in \mathbb{R}^p$, $g(\lambda, v) \leq p^*$

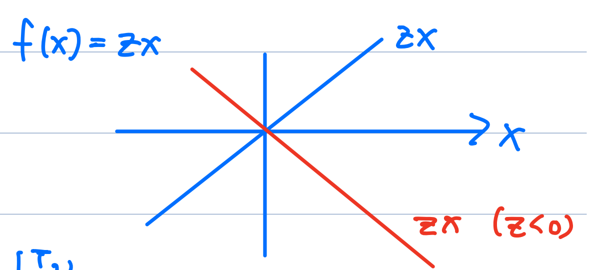
Define $d^* = \max_{\lambda \geq 0, \nu \in \mathbb{R}^p} g(\lambda, \nu)$: dual optimal value

Thm: $d^* \leq p^*$ [Weak Duality]

Weak Duality hold $\forall$ opt problems even if primal is not convex.

Def: Duality gap $p^* - d^* \geq 0$.

Sometimes, $p^* = d^*$ (i.e., zero duality gap), we say that strong duality holds.



$$-b^T\nu + \inf_x z^T x$$

If $z=0$, then the above is $= -b^T\nu$

If $z \neq 0$, $\inf_x z^T x = -\infty$

$x = -tz$, $z^T x = -t\|z\|^2 \to -\infty$ as $t \to +\infty$