DUALITY

The main motivation for a dual problem is to find bounds on the optimal objective value of a problem by solving a relevant model to the original.

Consider a problem with equality constraints:

$$(P)\quad \left(\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & h_j(x) = 0,\ j=1,\dots,p \end{array}\right)$$

Now, if we write the problem as:

$$\left(\begin{array}{ll} \text{minimize} & f_0(x) + \sum_{j=1}^{p} \mu_j h_j(x) \\ \text{subject to} & h_j(x) = 0,\ j=1,\dots,p \end{array}\right) \quad (P')$$

where $\mu_1, \dots, \mu_j \in \mathbb{R}$ are given, then clearly both models are equivalent.

If we relax the constraints and solve the unconstraint problem given by

$$(P_\mu) \qquad \underset{x \in \mathbb{R}^n}{\text{minimize}} \left(f_0(x) + \sum_{j=1}^{p} \mu_j h_j(x) \right),$$

we find a lower bound on the value of the original problem. This is the case for any choice of $\mu_1, \dots, \mu_j \in \mathbb{R}$.

If we denote the optimal objective function value of problem (P_μ) by $q(\mu)$, whereas the optimal objective function value of (P) by p^*, we have $\quad q(\mu) \leq p^* \quad$ for all $\mu \in \mathbb{R}^p$.

To find the "best" lower bound, one would maximize $q(\mu)$ over $\mu \in \mathbb{R}^p$.

Now, let's consider a problem with both inequality & equality constraints:

$$\begin{aligned}(P) \quad & \text{minimize} \quad f_0(x) \\ & \text{subject to} \quad f_i(x) \le 0, \quad i=1,\dots,m \\ & \qquad\qquad\quad\; h_j(x) = 0, \quad j=1,\dots,p\end{aligned}$$

Let p^* be the optimal objective function value of (P). Note that if the problem is unbounded, then $p^* = -\infty$. Moreover, if it's not feasible, then we say that $p^* = +\infty$.

The Lagrangian for problem (P) is a function $L: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ given by

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

where $\lambda_1, \dots, \lambda_m \ge 0$ and $\mu_1, \dots, \mu_p \in \mathbb{R}$ are Lagrange multipliers (or dual variables) associated with "$f_i(x) \le 0$" and "$h_j(x) = 0$" constraints, respectively.

We define the (Lagrange) dual function $g: \mathbb{R}^m_+ \times \mathbb{R}^p \to \mathbb{R}$ as

$$g(\lambda, \mu) := \inf_x L(x, \lambda, \mu),$$

where the infimum is taken over the domain of the original problem, that is, the intersection of the domains of $f_0, f_1, \dots, f_m, h_1, \dots, h_p$.

Observations:

1.) If we consider the Lagrangian as a function of λ, μ (for some fixed x), we see that $L(x,\lambda,\mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{p} \mu_j h_j(x)$ is an affine function of these variables. The same is true for $-L(x,\lambda,\mu)$.

Note that

$$g(\lambda,\mu) = \inf_x L(x,\lambda,\mu) = \underbrace{-\underbrace{\sup_x \underbrace{-L(x,\lambda,\mu)}_{\substack{\text{affine} \\ \Downarrow \\ \text{convex}}}}_{\substack{\text{pointwise supremum} \\ \text{of convex functions} \\ \text{is also convex}}}}_{\substack{\text{negative of a convex} \\ \text{function is concave!}}}$$

is a concave function of λ, μ.

2.) Let $\bar{x}$ be feasible for (P) and $\lambda_i \geq 0$. Then, we have

$$L(\bar{x},\lambda,\mu) = f_0(\bar{x}) + \underbrace{\underbrace{\sum_{i=1}^{m} \underbrace{\lambda_i}_{\geq 0} \underbrace{f_i(\bar{x})}_{\leq 0}}_{\leq 0}}_{\leq 0} + \underbrace{\sum \mu_j \underbrace{h_j(\bar{x})}_{=0}}_{=0} \leq f_0(\bar{x}).$$

This implies:

$$\begin{aligned} g(\lambda,\mu) &= \inf_x \left\{ f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{p} \mu_j \cdot h_j(x) \right\} \\ &\leq f_0(\bar{x}) + \sum_{i=1}^{m} \lambda_i f_i(\bar{x}) + \sum_{j=1}^{p} \mu_j h_j(\bar{x}) \\ &\leq f_0(\bar{x}) \end{aligned}$$

holds true for any feasible $\bar{x}$. Then, $g(\lambda,\mu) \leq p^*$ holds true.
(The inequality is trivial if (P) is infeasible as $p^* = +\infty$ in that case.)

In other words for any $\lambda \in \mathbb{R}^m_+$, $\mu \in \mathbb{R}^p$, the value of the dual function gives a lower bound for the optimal objective value p^* of problem (P).

Then, the question is to find the best lower bound that can be obtained by the function, which leads to the Lagrange dual problem:

$$\left(\begin{array}{lll} (D) & \text{maximize} & g(\lambda,\mu) \\ & \text{subject to} & \lambda_i \geq 0, \ i=1,\dots,m \\ & & \mu_j \in \mathbb{R}, \ j=1,\dots,p \end{array} \right)$$

We denote the optimal objective function value of (D) by d^*.

From the observation above, we have seen that $d^* \leq p^*$ holds for any problem (P) and its Lagrange dual (D). This is called the "weak duality" theorem.

Moreover, from the first observation we have $g(\lambda,\mu)$ a concave function.

Problem (D) is equivalent to $\left(\begin{array}{ll} \text{minimize} & -g(\lambda,\mu) \\ \text{subject to} & \lambda \geq 0 \\ & \mu \in \mathbb{R}^p \end{array} \right)$ which clearly is a convex problem.

So, we show that the Lagrange dual problem of any problem (P) (which is not necessarily convex) is always a convex optimization problem.

Terminology and notation:

The value of the primal problem (P) can be written as:

$$p^* = \inf \{ f_0(x) \mid f_i(x) \leq 0,\ h_j(x) = 0 \text{ for } i=1,\dots,m,\ j=1,\dots,p \}.$$

In particular, p^* can take values $\pm\infty$.

If the problem is unbounded, then $p^* = -\infty$.

If the problem is infeasible, then $p^* = \inf \phi = +\infty$ (convention)

Note that the dual objective function $g(\lambda,\mu) := \inf_x L(x,\lambda,\mu)$ is also an extended valued function; that is, for some λ,μ, the value can be $-\infty$. We say that the domain of g is $\operatorname{dom} g = \{ (\lambda,\mu) : g(\lambda,\mu) \in \mathbb{R} \}$.

Now, the value of the dual problem (D) can be written as:

$$d^* = \sup \{ g(\lambda,\mu) \mid \lambda \geq 0 \}.$$

In particular, d^* can take values $\pm\infty$.

If the problem is unbounded, then $d^* = +\infty$.

If the problem is infeasible, then $d^* = \sup \phi = -\infty$.

Note that the problem would be infeasible if $g(\lambda,\mu) = -\infty$ for all $\lambda \geq 0$.

Weak duality is: (*) $(d^* \leq p^*)$ and it holds in every possible case.

- If (P) is unbounded, i.e., $p^* = -\infty$, then $d^* \overset{(*)}{=} -\infty$ holds. That is, the dual problem (D) is infeasible.

Lagrangian: $L(x,\mu) = x_1^2 - 3x_2^2 + \mu(x_1 - x_2^3)$

$$= x_1^2 + \mu x_1 - 3x_2^2 - \mu x_2^3 .$$

$$g(\mu) = \inf_{x \in \mathbb{R}^2} \{ x_1^2 + \mu x_1 - 3x_2^2 - \mu x_2^3 \}$$

$$= \inf_{x_1 \in \mathbb{R}} \{ x_1^2 + \mu x_1 \} + \underbrace{\inf_{x_2 \in \mathbb{R}} \{ -3x_2^2 - \mu x_2^3 \}}_{-\infty \text{ for any choice of } \mu. \text{ (Why?)}}$$

So, $g(\mu) = -\infty$ for all $\mu \in \mathbb{R}$. (Dual problem is infeasible!)

$\Rightarrow d^* = -\infty.$

Clearly weak duality holds: $d^* \le p^*$. But dual problem gives the trivial lower bound.

Strong duality: Under some assumptions, the primal and dual problems may give the same value, that is, $p^* = d^*$ may hold. This is called strong duality.

e.g. For linear programming problems, we have strong duality.

Next, we will state that the strong duality holds also for convex problems BUT only under certain conditions. In particular, the constraints have to satisfy further conditions, these are called constraint qualifications.

One such condition is the Slater's condition that we have seen before.

- Recall that Slater's condition is satisfied if there exists $x \in \mathbb{R}^n$ (or the domain of the problem) such that $f_i(x) < 0$, $i = 1, \dots, m$ and $h_j(x) = 0$, $j = 1, \dots, p$.

- If (D) is unbounded, i.e., $d^* = +\infty$, then $+\infty = d^* \leq p^*$ implies $p^* = +\infty$. That is, the primal problem (P) is infeasible.

- Note that if the primal problem is infeasible, i.e., $p^* = +\infty$, then the dual problem may be 'anything'! In any case $d^* \leq p^* = +\infty$ would hold. Hence, it's possible that both problems are infeasible at the same time for instance. Or, $d^* \in \mathbb{R}$ may also happen in general.

Example: $\begin{pmatrix} \text{minimize} & x_1^2 - 3x_2^2 \\ \text{subject to} & x_1 = x_2^3 \end{pmatrix}$ (P)

Let's find p^* for this problem. Note that it's equivalent to the following problem:

$$\text{minimize } (x_2^6 - 3x_2^2).$$

This is unconstrained with objective function $f(x_2) = x_2^6 - 3x_2^2$. Since $\lim_{|x_2| \to \infty} f(x_2) = +\infty$, it's coercive for minimization. A minimum exists and it has to be among stationary points: $f'(x_2) = 6x_2^5 - 6x_2 = 6x_2(x_2^4 - 1) = 0$ is satisfied in 3 cases: $x_2 = 0$, $x_2 = 1$, $x_2 = -1$

Note that $f(1) = f(-1) = -2$ and $f(0) = 0$. Then $x_2^* = 1$ or $x_2^* = -1$ are optimal. (Optimal solutions for the original problem: $(1,1)$ and $(-1,-1)$.)

The value of the primal problem is $p^* = -2$.

Now, let's write the dual problem & solve it:

Theorem (Strong duality): If (P) is a convex optimization problem and the Slater's condition is satisfied, then strong duality holds. Moreover, if $p^* = d^* \in \mathbb{R}$ (value of the problems are finite), then, there exists a solution to the dual problem.

- Recall that in general $p^* \in \mathbb{R}$ does not mean that there exists a solution to (D). Similarly $d^* \in \mathbb{R}$ doesn't " " " " " to the primal problem.

e.g. $\inf_{x \geq 0} \frac{1}{x} = 0$ but $\frac{1}{x} \neq 0$ for any $x \in \mathbb{R}_+$.

Example: Consider the problem: $\begin{pmatrix} \text{minimize} & x_1^2 - x_2 \\ \text{subject to} & x_2^2 \leq 0 \end{pmatrix}$ (P)

This is a convex problem. But the Slater's condition doesn't hold since for no x we have $x_2^2 < 0$.

Let's check the value of the problem: Clearly, the feasible region is $\{x \in \mathbb{R}^2 \mid x_2 = 0\}$. Hence the problem is equivalent to

$$\begin{pmatrix} \underset{x_1 \in \mathbb{R}}{\text{minimize}} & x_1^2 \end{pmatrix}.$$

$\hat{x}_1 = 0$ is the optimal solution. Hence, for (P), $\hat{x} = (0,0)$ is the optimal solution and $p^* = 0$.

Let's write the dual problem.

Lagrangian: $L(x, \lambda) = x_1^2 - x_2 + \lambda(x_2^2)$

Dual function: $g(\lambda) = \inf_{x \in \mathbb{R}^2} L(x, \lambda)$.

$$\Rightarrow g(\lambda) = \inf_{x_1, x_2} \left(x_1^2 + \lambda x_2^2 - x_2 \right)$$

$$= \underbrace{\inf_{x_1} x_1^2}_{=0} + \underbrace{\inf_{x_2} \left(\lambda x_2^2 - x_2 \right)}_{\text{Let's check the value for } \lambda \geq 0.}$$

If $\lambda > 0$: $(\lambda x_2^2 - x_2)$ is quadratic convex. Let's check the stationary points:

$$2\lambda x_2 - 1 = 0 \Leftrightarrow x_2 = \frac{1}{2\lambda}$$ is the optimal solution for the inf.

$$\inf_{x_2} (\lambda x_2^2 - x_2) = \lambda \left(\frac{1}{2\lambda} \right)^2 - \frac{1}{2\lambda} = \frac{1}{4\lambda} - \frac{1}{2\lambda} = \frac{-1}{4\lambda}.$$

If $\lambda = 0$: $\inf_{x_2} (-x_2) = -\infty$.

So, the dual function is found as $$g(\lambda) = \begin{cases} -\infty, & \lambda = 0 \\ -\frac{1}{4\lambda} & \lambda > 0 \end{cases}$$

<u>The Dual problem:</u>

$$\begin{pmatrix} \text{maximize} & \frac{-1}{4\lambda} \\ \text{s.t.} & \lambda > 0 \end{pmatrix} \quad (D)$$

Note that $\frac{-1}{4\lambda} < 0$ for all $\lambda > 0$. Moreover, $\frac{-1}{4\lambda} \to 0$ as $\lambda \to \infty$.

Hence, the value of the dual problem is $d^* = 0$. Strong duality holds!

However, there's no solution to the dual problem! This may happen as the Slater's condition does not hold (hence the theorem does not apply this case).

Example: Consider the problem: minimize e^{-x_2}

subject to $\sqrt{x_1^2+x_2^2} - x_1 \le 0$

First, note that the problem is convex: $f_0(x) = e^{-x_2}$ is convex.

$f_1(x) = \|x\|_2 - x_1$ is convex.

Let's check the feasible region: $(x_1 \geqslant 0)$

Note that $\sqrt{x_1^2+x_2^2} \le x_1 \Rightarrow x_1^2 + x_2^2 \le x_1^2$ (taking square is increasing on the nonnegative real line)

$\Rightarrow x_2^2 \le 0$

$\Rightarrow x_2 = 0.$

On the other hand, $\sqrt{x_1^2} \le x_1$ holds true for all $x_1 \geqslant 0$. Then, the feasible region is $S = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geqslant 0,\ x_2 = 0\}$.

There's no $x \in S$ such that $f_1(x) < 0$ holds. Slater's condition is not satisfied.

Let's check p^* & d^*:

- Note that $p^* = e^{-0} = 1$ (For any point from S, the value is 1) (Any feasible solution is optimal)

- Let's write the dual problem:

Lagrangian: $L(x, \lambda) = e^{-x_2} + \lambda\left(\sqrt{x_1^2+x_2^2} - x_1\right)$, $\lambda \geqslant 0$.

Dual function: $g(\lambda) = \inf_x L(x, \lambda)$.

$$g(\lambda) = \inf_{x} \left(e^{-x_2} + \lambda \sqrt{x_1^2 + x_2^2} - \lambda x_1 \right)$$

First of all, note that $\sqrt{x_1^2 + x_2^2} - x_1 \geq \sqrt{x_1^2} - x_1 \geq x_1 - x_1 \geq 0$ for all $x_1, x_2 \in \mathbb{R}$.

$$\Rightarrow \lambda \left(\sqrt{x_1^2 + x_2^2} - x_1 \right) \geq 0 \quad \text{for all } \lambda \geq 0.$$

Also, $e^{-x_2} \geq 0$ for all $x_2 \in \mathbb{R}$.

Hence, $g(\lambda) \geq 0$ holds for all $\lambda \geq 0$.

On the other hand, one can show that the infimum is zero.

Indeed, if one takes $x_2 = -\log \varepsilon$, $x_1 = \dfrac{x_2^2 - \varepsilon^2}{2\varepsilon}$ for some $\varepsilon > 0$, then

we have

$$L(x, \lambda) = e^{-x_2} + \lambda \left(\sqrt{\frac{(x_2^2 - \varepsilon^2)^2}{4\varepsilon^2} + x_2^2} - x_1 \right) = \varepsilon + \lambda \left[\frac{x_2^2 + \varepsilon^2}{2\varepsilon^2} - \frac{x_2^2 - \varepsilon^2}{2\varepsilon^2} \right] = \varepsilon(\lambda + 1).$$

$$\longrightarrow 0 \text{ as } \varepsilon \downarrow 0.$$

This means that $g(\lambda) = 0$ for all $\lambda \geq 0$; hence $d^* = 0$.

The problem is convex but Slater's condition does not hold. This time the strong duality does not hold as we have

$$0 = d^* < p^* = 1.$$

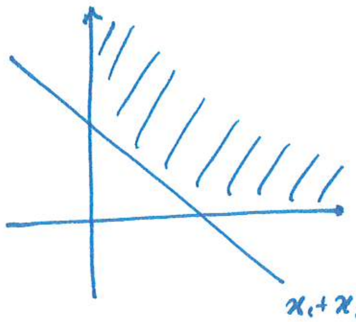
- <u>Summary</u>: Theorem says that if the problem is convex & Slater condition holds, then strong duality holds and a solution to the dual problem exists.

Note that there are also examples which are not convex but the strong duality holds.

Example: Consider:

$$\begin{aligned} \text{minimize} \quad & x_1^3 + x_2^3 \\ \text{subject to} \quad & x_1 + x_2 \geq 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

This is not a convex program. It's not difficult to verify that $\hat{x} = \left(\frac{1}{2}, \frac{1}{2}\right)$ is the optimal solution. The value of the problem is then $p^* = \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^3 = \frac{1}{4}$.



We will see that we can derive different dual problems by considering different Lagrangian functions and the domains of the problem:

① Lagrangian:
$$\begin{aligned} L(x, \lambda) &= x_1^3 + x_2^3 + \lambda_1(1 - x_1 - x_2) + \lambda_2(-x_1) + \lambda_3(-x_2) \\ &= x_1^3 - \lambda_1 x_1 - \lambda_2 x_1 + x_2^3 - \lambda_1 x_2 - \lambda_3 x_2 + \lambda_1 \\ &= x_1^3 - (\lambda_1 + \lambda_2) x_1 + x_2^3 - (\lambda_1 + \lambda_3) x_2 + \lambda_1 \end{aligned}$$

Dual function:
$$\begin{aligned} g(\lambda) &= \inf_{x \in \mathbb{R}^2} L(x, \lambda) \\ &= \underbrace{\inf_{x_1 \in \mathbb{R}} \{ x_1^3 - (\lambda_1 + \lambda_2) x_1 \}}_{\substack{-\infty \\ (\text{for any } \lambda_1, \lambda_2)}} + \underbrace{\inf_{x_2 \in \mathbb{R}} \{ x_2^3 - (\lambda_1 + \lambda_3) x_2 \}}_{\substack{-\infty \\ (\text{for any } \lambda_1, \lambda_3)}} + \lambda_1 \end{aligned}$$

Then, $g(\lambda) = -\infty$ for any $\lambda \in \mathbb{R}^3_+$. Hence, $d^* = -\infty$.

Weak duality holds as expected, but it's the trivial lower bound.

(2) Now, assume that the domain of the problem is $x_1, x_2 \geq 0$.

Lagrangian: $L(x, \lambda) = x_1^3 + x_2^3 + \lambda(1 - x_1 - x_2)$ (there's a single constraint)

Dual function: $g(\lambda) = \inf_{\substack{x_1 \geq 0 \\ x_2 \geq 0}} L(x, \lambda)$

$$= \inf_{\substack{x_1 \geq 0 \\ x_2 \geq 0}} \left\{ (x_1^3 - \lambda x_1) + (x_2^3 - \lambda x_2) \right\} + \lambda$$

$$= \underbrace{\inf_{x_1 \geq 0} (x_1^3 - \lambda x_1)} + \underbrace{\inf_{x_2 \geq 0} (x_2^3 - \lambda x_2)} + \lambda$$

$x_1^3 - \lambda x_1$ is coercive for minimization for $x_1 \geq 0$.

(Stationary points) $3x_1^2 - \lambda = 0 \Rightarrow x_1^2 = \frac{\lambda}{3}$

(only stat. point is) $x_1 = \sqrt{\frac{\lambda}{3}} > 0$ is optimal.

(Value of infimum) $\frac{\lambda}{3}\sqrt{\frac{\lambda}{3}} - \lambda\sqrt{\frac{\lambda}{3}} = \frac{-2\lambda\sqrt{\lambda}}{3\sqrt{3}}$

$x_2^3 - \lambda x_2$ is coercive for minimization for $x_2 \geq 0$.

$3x_2^2 - \lambda = 0$

$\Rightarrow x_2 = \sqrt{\frac{\lambda}{3}} > 0$. is optimal

same here!

$$g(\lambda) = -\frac{4}{3\sqrt{3}} \lambda\sqrt{\lambda} + \lambda$$

Dual Problem: maximize $g(\lambda)$

subject to $\lambda \geq 0$.

This is equivalent to: $(-)\underset{\lambda \geq 0}{\text{minimize}} \left(\frac{4}{3\sqrt{3}} \lambda^{3/2} - \lambda \right)$

We know that this is a convex program. Any KKT point is optimal.

Lagrangian: $L(\lambda, y) = \frac{4}{3\sqrt{3}} \lambda^{3/2} - \lambda + y(-\lambda)$

$$\nabla_\lambda L(\lambda, y) = \frac{3}{2} \frac{4}{3\sqrt{3}} \lambda^{1/2} - 1 - y = 0 \Rightarrow \lambda^{1/2} = \frac{\sqrt{3}}{2}(y+1).$$

- Compl. slackness: $\lambda . y = 0$
  - $\lambda = 0 \Rightarrow y = -1$ (not dual feasible)
  - $y = 0 \Rightarrow \lambda^{1/2} = \frac{\sqrt{3}}{2} \Rightarrow \lambda = \frac{3}{4}$ (KKT point, hence optimal.)

The value of the dual problem:

$$d^* = g\left(\frac{3}{4}\right) = \frac{-4}{3\sqrt{3}} \cdot \frac{3}{4} \cdot \frac{\sqrt{3}}{2} + \frac{3}{4} = \frac{1}{4}.$$

Note that the strong duality holds! Moreover, there's a solution to the dual problem, too!

Indeed, this was expected! Why?

When we consider $f_0(x) = x_1^3 + x_2^3$ over $\mathbb{R}^2_+$, it's a convex function.

$$\nabla f_0(x) = \begin{bmatrix} 3x_1^2 \\ 3x_2^2 \end{bmatrix}, \quad \nabla^2 f_0(x) = \begin{bmatrix} 6x_1 & 0 \\ 0 & 6x_2 \end{bmatrix} \succeq 0 \quad \text{for } x_1, x_2 \geq 0.$$

Moreover, the Slater's condition holds, as there exists $x_1, x_2 \geq 0$ s.t. $x_1 + x_2 \geq 1$. (e.g. $x = (1,1)$).

Hence, the theorem applies!

Example: Dual of a standard LP

Consider the program:
$$\left(\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \le b \\ & x \geq 0 \end{array}\right)$$

Let's take the dual. First, realize that it's equivalent to

$$\begin{array}{lll} \text{minimize} & -c^T x & \\ \text{subject to} & Ax \le b & (\lambda) \\ & x \geq 0 & (\text{let's consider it as the domain}) \end{array}$$

Lagrangian:
$$L(x,\lambda) = -c^T x + \underbrace{\lambda^T (Ax - b)}_{\sum_i \lambda_i (a_i x - b_i)}$$

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

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

Dual function:
$$g(\lambda) = \inf_{x \geq 0} \{ \underbrace{(A^T\lambda - c)^T x}_{} \} - \lambda^T b$$

If there's a negative component, one can let corresponding $x_i \to +\infty$ and the product would tend to $-\infty$.

Otherwise, the inner product is always nonnegative and the infimum is attained at $x = 0$

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

The dual problem:
$$\begin{aligned} &\text{maximize} && g(\lambda) \\ &\text{subject to} && \lambda \geq 0. \end{aligned}$$

This can be written as:
$$\begin{aligned} &\text{maximize} && -\lambda^T b \\ &\text{subject to} && A^T\lambda \geq c \\ & && \lambda \geq 0 \end{aligned}$$

Equivalently,
$$\left(\begin{aligned} &\text{minimize} && b^T\lambda \\ &\text{subject to} && A^T\lambda \geq c \\ & && \lambda \geq 0 \end{aligned} \right) \quad (D) \qquad \text{as expected!}$$

Recall that for LP's taking the dual has same rules:

Primal		Dual
maximize	—	minimize
$\leq$	—	non-negative variable
$=$	—	free variable
non-neg. var	—	$\geq$
free var	—	$=$

Exercise:

$$(P) \quad \begin{aligned} &\text{maximize} && c^T x + d^T y \\ &\text{subject to} && A_1 x + B_1 y \leq b_1 \quad (u) \\ & && A_2 x + B_2 y = b_2 \quad (v) \\ & && x \geq 0 \\ & && y: \text{free} \end{aligned}$$

$$\xrightarrow{\text{Show!}}$$

$$(D) \quad \begin{aligned} &\text{minimize} && b_1^T u + b_2^T v \\ &\text{subject to} && A_1^T u + A_2^T v \geq c \\ & && B_1^T u + B_2^T v = d \\ & && u \geq 0 \\ & && v: \text{free} \end{aligned}$$

Example: (strictly convex quadratic programming)

Let $Q \in \mathbb{R}^{n\times n}$ be positive definite, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Consider

$$(P)\begin{pmatrix} \text{minimize} & x^T Q x + c^T x \\ \text{subject to} & Ax \leq b \end{pmatrix}$$

- This is a convex problem as $Q \succ 0$, and constraints are linear. Moreover, if it's feasible, then the strong duality holds. Indeed, we would look for $Ax < b$ to satisfy Slater's condition. However, if the inequality constraints are linear, then one can relax this condition.

Let's derive the Lagrange dual.

Lagrangian:

$$L(x, \lambda) = x^T Q x + c^T x + \lambda^T (Ax - b)$$

where $\lambda \in \mathbb{R}^m$, $\lambda \geqslant 0$.

Dual function: $$g(\lambda) = \inf_{x \in \mathbb{R}^n} \{ x^T Q x + c^T x + \lambda^T A x \} - \lambda^T b$$

$$= \inf_{x \in \mathbb{R}^n} \{ \underbrace{x^T Q x + (c + A^T \lambda)^T x}_{\text{Convex function of } x \text{ as } Q \text{ is p.d.}} \} - \lambda^T b$$

Any stationary point is optimal

$$\nabla_x \left(x^T Q x + (c + A^T\lambda)^T x \right) = 2Qx + c + A^T \lambda = 0 \iff \hat{x} = -\frac{1}{2} Q^{-1} (c + A^T \lambda)$$

This is possible as $Q \succ 0$, hence Q^{-1} exists!

Then, the infimum is attained at $\hat{x}$. Let's compute the value:

$$g(\lambda) = -\lambda^T b + \hat{x}^T Q \hat{x} + (c + A^T\lambda)^T \hat{x}$$

$$= -\lambda^T b + \frac{1}{4}(c + A^T\lambda)^T Q^{-1} \underbrace{Q Q^{-1}}_{I} (c + A^T\lambda) + (c + A^T\lambda)^T \left[-\frac{1}{2} Q^{-1} (c + A^T\lambda) \right]$$

$$= -\lambda^T b - \frac{1}{4} (c + A^T\lambda)^T Q^{-1} (c + A^T\lambda)$$

$$= -\lambda^T b - \frac{1}{4} \left[c^T Q^{-1} c + \lambda^T A Q^{-1} c + c^T Q^{-1} A^T \lambda + \lambda^T A Q^{-1} A^T \lambda \right]$$

$$= -b^T \lambda - \frac{1}{4} \left[\lambda^T (A Q^{-1} A) \lambda + 2 c^T Q^{-1} A^T \lambda + c^T Q^{-1} c \right]$$

$$= -\frac{1}{4} \lambda^T (A Q^{-1} A) \lambda - \left(b + \frac{1}{2} A Q^{-1} c\right)^T \lambda - \frac{1}{2} c^T Q^{-1} c$$ – quadratic function of λ!

Dual problem: $$\begin{pmatrix} \text{maximize } g(\lambda) \\ \text{subject to } \lambda \geq 0 \end{pmatrix} \quad (D)$$ again a convex quadratic optimization problem.

Example: (Convex quadratic programming)

Consider the same problem. Now assume that Q is posive semidefinite but not necessarily positive definite. Then, Q^{-1} may not exist, hence we can not use the previous approach to derive a dual model.

Instead, we use the fact that for any $Q \succcurlyeq 0$, there exists a matrix $D \in \mathbb{R}^{n\times n}$ such that $Q = D^T D$. Now, let's write the equivalent problem using D:

$$\begin{aligned} &\text{minimize} && x^T D^T D x + c^T x \\ &\text{subject to} && Ax \leq b \end{aligned}$$

Moreover, we observe that defining $z = Dx$, we can write it further as

Decision variables are $x \in \mathbb{R}^n$, $z \in \mathbb{R}^n$

$$\left(\begin{aligned} &\text{minimize} && z^T z + c^T x \\ &\text{subject to} && Ax \leq b \\ & && Dx = z \end{aligned}\right) \quad (P)$$

Lagrangian: $L(x,z,\lambda,\mu) = z^T z + c^T x + \lambda^T (Ax - b) + \mu^T (Dx - z)$

$$= z^T z - \mu^T z + (c + A^T\lambda + D^T\mu)^T x - \lambda^T b.$$

$$g(\lambda,\mu) = \inf_{x,z} L(x,z,\lambda,\mu)$$

$$= \underbrace{\inf_z \{z^T z - \mu^T z\}}_{\substack{\text{quadratic unconstrained} \\ \text{optimization (convex)}}} + \inf_x \underbrace{(c + A^T\lambda + D^T\mu)^T x}_{\substack{\text{linear function} \\ \text{of } x}} - \lambda^T b$$

$\Downarrow$

since this is convex any stationary point is optimal

$\Downarrow$

infimum is $-\infty$. unless $c + A^T\lambda + D^T\mu = 0$.

$$\inf_x (c + A^T\lambda + D^T\mu)^T x = \begin{cases} 0 & \text{if } c + A^T\lambda + D^T\mu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$\nabla_z (z^T z - \mu^T z) = 2z - \mu = 0 \Rightarrow \hat{z} = \frac{\mu}{2}.$$

Then, $\inf_z \{ z^T z - \mu^T z \} = \frac{1}{4}\mu^T\mu - \frac{1}{2}\mu^T\mu = -\frac{1}{4}\mu^T\mu.$

The dual function can be written as:

$$g(\lambda,\mu) = \begin{cases} -\frac{1}{4}\mu^T\mu - \lambda^T b & \text{if } c + A^T\lambda + D^T\mu = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem is:

$$\begin{aligned} (D) \quad & \text{maximize} && -\frac{1}{4}\mu^T\mu - \lambda^T b \\ & \text{subject to} && c + A^T\lambda + D^T\mu = 0 \\ & && \lambda \geq 0 \\ & && \mu : \text{free} \end{aligned}$$