15. Duality

- Upper and lower bounds
- General duality
- Constraint qualifications
- Counterexample
- Complementary slackness
- Examples
- Sensitivity analysis

Laurent Lessard (www.laurentlessard.com)

Upper bounds

Optimization problem (not necessarily convex!):

$$\begin{aligned} \underset{x \in D}{\text{minimize}} \quad & f_0(x) \\ \text{subject to:} \quad & f_i(x) \leq 0 \quad \text{for } i = 1, \dots, m \\ & h_j(x) = 0 \quad \text{for } j = 1, \dots, r \end{aligned}$$

- D is the domain of all functions involved.
- Suppose the optimal value is $p^\star$.
- **Upper bounds:** if $x \in D$ satisfies $f_i(x) \leq 0$ and $h_j(x) = 0$ for all i and j, then: $p^\star \leq f_0(x)$.
- Any feasible x yields an upper bound for $p^\star$.

Lower bounds

Optimization problem (not necessarily convex!):

$$\begin{aligned} \underset{x \in D}{\text{minimize}} \quad & f_0(x) \\ \text{subject to:} \quad & f_i(x) \le 0 \quad \text{for } i = 1, \dots, m \\ & h_j(x) = 0 \quad \text{for } j = 1, \dots, r \end{aligned}$$

- As with LPs, use the constraints to find lower bounds
- For any $\lambda_i \ge 0$ and $\nu_j \in \mathbb{R}$, if $x \in D$ is feasible, then

$$f_0(x) \ge f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{r} \nu_j h_j(x)$$

Lower bounds

$$f_0(x) \geq \underbrace{f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{r} \nu_j h_j(x)}_{\text{Lagrangian } L(x,\lambda,\nu)}$$

This is a lower bound on f_0, but we want a lower bound on $p^\star$.
Minimize right side over $x \in D$ and left side over feasible x.

$$p^\star \geq \left\{ \inf_{x \in D} \; L(x,\lambda,\nu) \right\} = g(\lambda,\nu)$$

This inequality holds whenever $\lambda \geq 0$.

Lower bounds

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

Whenever $\lambda \geq 0$, we have:

$$g(\lambda, \nu) := \left\{ \inf_{x \in D} \; L(x, \lambda, \nu) \right\} \leq p^\star$$

Useful fact: $g(\lambda, \nu)$ is a **concave** function. This is true even if the original optimization problem is not convex! (because g is a pointwise minimum of affine functions)

General duality

Primal problem (P)

$$\begin{aligned} \underset{x \in D}{\text{minimize}} \quad & f_0(x) \\ \text{subject to:} \quad & f_i(x) \le 0 \quad \forall i \\ & h_j(x) = 0 \quad \forall j \end{aligned}$$

Dual problem (D)

$$\begin{aligned} \underset{\lambda, \nu}{\text{maximize}} \quad & g(\lambda, \nu) \\ \text{subject to:} \quad & \lambda \ge 0 \end{aligned}$$

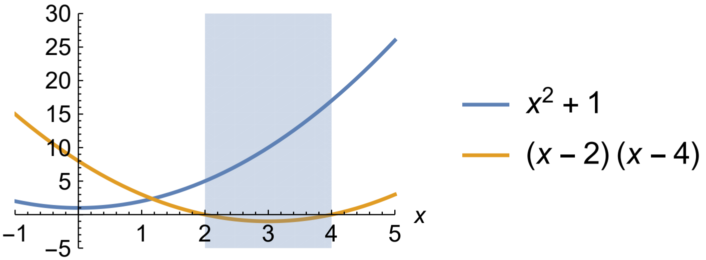
If x and λ are feasible points of (P) and (D) respectively:

$$g(\lambda, \nu) \le d^\star \le p^\star \le f_0(x)$$

This is called the **Lagrange dual**. Bad news: strong duality ($p^\star = d^\star$) does **not** always hold!

Example (Srikant)

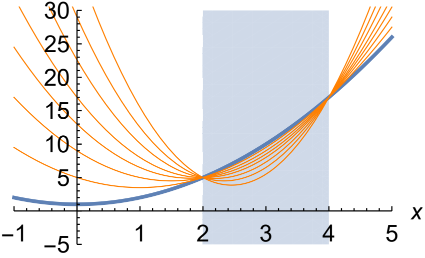
$$\begin{array}{ll} \underset{x}{\text{minimize}} & x^2 + 1 \\ \text{subject to:} & (x-2)(x-4) \leq 0 \end{array}$$



- optimum occurs at $x = 2$, has value $p^\star = 5$

Example (Srikant)

Lagrangian: $L(x, \lambda) = x^2 + 1 + \lambda(x-2)(x-4)$



- Plot for different values of $\lambda \geq 0$
- $g(\lambda) = \inf_x L(x, \lambda)$ should be a lower bound on $p^\star = 5$ for all $\lambda \geq 0$.

Example (Srikant)

Lagrangian: $L(x, \lambda) = x^2 + 1 + \lambda(x-2)(x-4)$

- Minimize the Lagrangian:

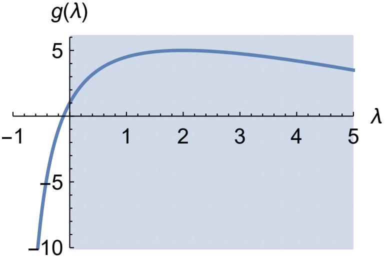
$$\begin{aligned} g(\lambda) &= \inf_x \; L(x, \lambda) \\ &= \inf_x \; (\lambda+1)x^2 - 6\lambda x + (8\lambda + 1) \end{aligned}$$

If $\lambda \leq -1$, it is unbounded. If $\lambda > -1$, the minimum occurs when $2(\lambda+1)x - 6\lambda = 0$, so $\hat{x} = \frac{3\lambda}{\lambda+1}$.

$$g(\lambda) = \begin{cases} -9\lambda^2/(1+\lambda) + 1 + 8\lambda & \lambda > -1 \\ -\infty & \lambda \leq -1 \end{cases}$$

Example (Srikant)

$$\begin{aligned} \underset{\lambda}{\text{maximize}} \quad & -9\lambda^2/(1+\lambda) + 1 + 8\lambda \\ \text{subject to:} \quad & \lambda \geq 0 \end{aligned}$$



- optimum occurs at $\lambda = 2$, has value $d^\star = 5$
- same optimal value as primal problem! (strong duality)

Constraint qualifications

- weak duality ($d^\star \leq p^\star$) always holds. Even when the optimization problem is not convex.
- strong duality ($d^\star = p^\star$) often holds for convex problems (but not always).

A **constraint qualification** is a condition that guarantees strong duality. An example we've already seen:

- If the optimization problem is an LP, strong duality holds

Slater's constraint qualification

$$\begin{aligned} \underset{x\in D}{\text{minimize}} \quad & f_0(x) \\ \text{subject to:} \quad & f_i(x) \le 0 \quad \text{for } i = 1,\dots,m \\ & h_j(x) = 0 \quad \text{for } j = 1,\dots,r \end{aligned}$$

Slater's constraint qualification:

> If the optimization problem is convex and strictly feasible, then strong duality holds.

- convexity requires: D and f_i are convex and h_j are affine.
- strict feasibility means there exists some $\tilde{x}$ in the interior of D such that $f_i(\tilde{x}) < 0$ for $i = 1,\dots,m$.

Slater's constraint qualification

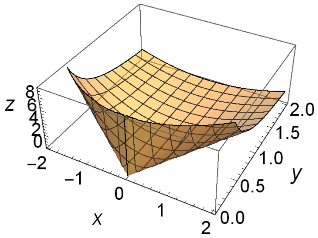
If the optimization problem is convex and strictly feasible, then strong duality holds.

- Good news: Slater's constraint qualification is rather weak. i.e. it is usually satisfied by convex problems.
- Can be relaxed so that strict feasibility is not required for the linear constraints.

Counterexample (Boyd)

$$\begin{aligned} \underset{x\in\mathbb{R},\, y>0}{\text{minimize}} \quad & e^{-x} \\ \text{subject to:} \quad & x^2/y \le 0 \end{aligned}$$

- The function x^2/y is convex for $y > 0$ (see plot)
- The objective e^{-x} is convex
- Feasible set: $\{(0, y) \mid y > 0\}$
- Solution is trivial ($p^\star = 1$)



Counterexample (Boyd)

$$\underset{x\in\mathbb{R},\, y>0}{\text{minimize}} \quad e^{-x}$$
$$\text{subject to:} \quad x^2/y \leq 0$$

- **Lagrangian:** $L(x, y, \lambda) = e^{-x} + \lambda x^2/y$
- **Dual function:** $g(\lambda) = \inf_{x,y>0}\left(e^{-x} + \lambda x^2/y\right) = 0.$
- The dual problem is:

$$\underset{\lambda\geq 0}{\text{maximize}} \quad 0$$

  So we have $d^\star = 0 < 1 = p^\star$.
- Slater's constraint qualification is **not** satisfied!

About Slater's constraint qualification

Slater's condition is only **sufficient**.
(Slater) $\Longrightarrow$ (strong duality)

- There exist problems where Slater's condition fails, yet strong duality holds.
- There exist nonconvex problems with strong duality.

Complementary slackness

Assume strong duality holds. If $x^\star$ is primal optimal and $(\lambda^\star, \nu^\star)$ is dual optimal, then we have:

$$g(\lambda^\star, \nu^\star) = d^\star = p^\star = f_0(x^\star)$$

$$\begin{aligned} f_0(x^\star) = g(\lambda^\star, \nu^\star) &= \inf_{x \in D} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i^\star f_i(x) + \sum_{j=1}^{r} \nu_j^\star h_j(x) \right) \\ &\le f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star f_i(x^\star) + \sum_{j=1}^{r} \nu_j^\star h_j(x^\star) \\ &\le f_0(x^\star) \end{aligned}$$

The last inequality holds because $x^\star$ is primal feasible. We conclude that the inequalities must all be equalities.

Complementary slackness

- We concluded that:

$$f_0(x^\star) = f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star f_i(x^\star) + \sum_{j=1}^{r} \nu_j^\star h_j(x^\star)$$

  But $f_i(x^\star) \le 0$ and $h_j(x^\star) = 0$. Therefore:

$$\lambda_i^\star f_i(x^\star) = 0 \quad \text{for } i = 1, \dots, m$$

- This property is called **complementary slackness**. We've seen it before for linear programs.

$$\lambda_i^\star > 0 \implies f_i(x^\star) = 0 \qquad \text{and} \qquad f_i(x^\star) < 0 \implies \lambda_i^\star = 0$$

Dual of an LP

$$\begin{aligned} \underset{x\geq 0}{\text{minimize}} \quad & c^\mathsf{T} x \\ \text{subject to:} \quad & Ax \geq b \end{aligned}$$

- Lagrangian: $L(x,\lambda) = c^\mathsf{T} x + \lambda^\mathsf{T}(b - Ax)$
- Dual function: $g(\lambda) = \min_{x\geq 0} \, (c - A^\mathsf{T}\lambda)^\mathsf{T} x + \lambda^\mathsf{T} b$

$$g(\lambda) = \begin{cases} \lambda^\mathsf{T} b & \text{if } A^\mathsf{T}\lambda \leq c \\ -\infty & \text{otherwise} \end{cases}$$

Dual of an LP

$$\begin{aligned} \underset{x \geq 0}{\text{minimize}} \quad & c^\mathsf{T} x \\ \text{subject to:} \quad & Ax \geq b \end{aligned}$$

- Dual is:

$$\begin{aligned} \underset{\lambda \geq 0}{\text{maximize}} \quad & \lambda^\mathsf{T} b \\ \text{subject to:} \quad & A^\mathsf{T} \lambda \leq c \end{aligned}$$

- This is the same result that we found when we were studying duality for linear programs.

Dual of an LP

What if we treat $x \geq 0$ as a constraint instead? ($D = \mathbb{R}^n$).

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^\mathsf{T} x \\ \text{subject to:} \quad & Ax \geq b \\ & x \geq 0 \end{aligned}$$

- Lagrangian: $L(x, \lambda, \mu) = c^\mathsf{T} x + \lambda^\mathsf{T}(b - Ax) - \mu^\mathsf{T} x$
- Dual function: $g(\lambda, \mu) = \min_x \, (c - A^\mathsf{T}\lambda - \mu)^\mathsf{T} x + \lambda^\mathsf{T} b$

$$g(\lambda) = \begin{cases} \lambda^\mathsf{T} b & \text{if } A^\mathsf{T}\lambda + \mu = c \\ -\infty & \text{otherwise} \end{cases}$$

Dual of an LP

What if we treat $x \geq 0$ as a constraint instead? ($D = \mathbb{R}^n$).

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^{\mathsf{T}} x \\ \text{subject to:} \quad & Ax \geq b \\ & x \geq 0 \end{aligned}$$

- Dual is:

$$\begin{aligned} \underset{\lambda \geq 0,\, \mu \geq 0}{\text{maximize}} \quad & \lambda^{\mathsf{T}} b \\ \text{subject to:} \quad & A^{\mathsf{T}} \lambda + \mu = c \end{aligned}$$

- Solution is the same, μ acts as the slack variable.

Dual of a convex QP

Suppose $Q \succ 0$. Let's find the dual of the QP:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \tfrac{1}{2}x^\mathsf{T} Q x \\ \text{subject to:} \quad & Ax \ge b \end{aligned}$$

- Lagrangian: $L(x,\lambda) = \frac{1}{2}x^\mathsf{T} Q x + \lambda^\mathsf{T}(b - Ax)$
- Dual function: $g(\lambda) = \min_x \left(\frac{1}{2}x^\mathsf{T} Q x + \lambda^\mathsf{T}(b - Ax)\right)$
  Minimum occurs at: $\hat{x} = Q^{-1}A^\mathsf{T}\lambda$

$$g(\lambda) = -\tfrac{1}{2}\lambda^\mathsf{T} A Q^{-1} A^\mathsf{T} \lambda + \lambda^\mathsf{T} b$$

Dual of a convex QP

Suppose $Q \succ 0$. Let's find the dual of the QP:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \tfrac{1}{2} x^\mathsf{T} Q x \\ \text{subject to:} \quad & Ax \geq b \end{aligned}$$

- Dual is also a QP:

$$\begin{aligned} \underset{\lambda}{\text{maximize}} \quad & -\tfrac{1}{2} \lambda^\mathsf{T} A Q^{-1} A^\mathsf{T} \lambda + \lambda^\mathsf{T} b \\ \text{subject to:} \quad & \lambda \geq 0 \end{aligned}$$

- It's still easy to solve (maximizing a concave function)

Sensitivity analysis

$$\begin{aligned} \min_{x \in D} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le u_i \quad \forall i \\ & h_j(x) = v_j \quad \forall j \end{aligned}$$

$$\begin{aligned} \max_{\lambda,\nu} \quad & g(\lambda,\nu) - \lambda^\mathsf{T} u - \nu^\mathsf{T} v \\ \text{s.t.} \quad & \lambda \ge 0 \end{aligned}$$

- As with LPs, dual variables quantify the sensitivity of the optimal cost to changes in each of the constraints.
- A change in u_i causes a bigger change in $p^\star$ if $\lambda_i^\star$ is larger.
- A change in v_j causes a bigger change in $p^\star$ if $\nu_j^\star$ is larger.
- If $p^\star(u,v)$ is differentiable, then:

$$\lambda_i^\star = -\frac{\partial p^\star(0,0)}{\partial u_i} \qquad \text{and} \qquad \nu_j^\star = -\frac{\partial p^\star(0,0)}{\partial v_j}$$