

Learning objectives. By the end of class, you should be able to

1. define strong duality, primal attainment, and dual attainment, and explain when a certificate of global optimality exists;
2. write a linear program in standard form, derive its dual, and certify a solution by complementary slackness;
3. state Slater's condition, and check it for second-order cone and semidefinite constraints;
4. use the S-lemma to show that the trust-region subproblem has zero duality gap.

1 Recap: Weak duality

Every constrained problem has a primal and a dual,

$$p^* = \inf_{x \in \mathbb{R}^n} f_0(x) \text{ s.t. } f_i(x) \leq 0 \text{ for all } i,$$

$$d^* = \sup_{y \in \mathbb{R}_+^m} g(y), \quad g(y) = \inf_{x \in \mathbb{R}^n} \underbrace{\left\{ f_0(x) + \sum_{i=1}^m y_i f_i(x) \right\}}_{\mathcal{L}(x,y)}.$$

Weak duality. Every feasible $\hat{x}$ and every $\hat{y} \in \mathbb{R}_+^m$ satisfy

$$g(\hat{y}) \leq d^* \leq p^* \leq f_0(\hat{x}).$$

A zero duality gap $f_0(\hat{x}) = g(\hat{y})$ certifies that $\hat{x}$ is globally optimal.

Example 1. Certify that $\hat{x} = 1$ is globally optimal for

$$\min_{x \in \mathbb{R}} x^2 \text{ s.t. } x \geq 1.$$

Example 2. Consider the primal problem

$$p^* = \min_{x \in \mathbb{R}^2} x_1 \text{ s.t. } \left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\| \leq x_2.$$

Its dual problem is

$$d^* = \sup_{y \geq 0} 0 \text{ s.t. } \left\| \begin{bmatrix} 1 \\ -y \end{bmatrix} \right\| \leq y.$$

Compute p^* and d^* .

The problem is convex. The norm is convex, $-x_2$ is affine, and their sum is convex. We derive this dual later in the lecture.

2 Strong duality

Consider again the primal and the dual,

$$p^* = \inf_{x \in \mathbb{R}^n} f_0(x) \text{ s.t. } f_i(x) \leq 0 \text{ for all } i,$$
$$d^* = \sup_{y \in \mathbb{R}_+^m} \inf_{x \in \mathbb{R}^n} \left\{ f_0(x) + \sum_{i=1}^m y_i f_i(x) \right\}.$$

Weak duality says $d^* \leq p^*$, without knowing anything else about the two problems. In this lecture, we care about three properties.

- **Strong duality.** $p^* = d^*$.
- **Primal attainment.** Some feasible x_* has $f_0(x_*) = p^*$.
- **Dual attainment.** Some $y_* \in \mathbb{R}_+^m$ has $g(y_*) = d^*$.

Question 1. When does every globally optimal x_* have a certificate?

Example 2 shows that strong duality can fail, even for a convex problem. Two classic conditions guarantee it.

- The objective and the constraints are linear.
- The objective and the constraints are convex, and Slater's condition holds.

3 Linear programs

Theorem 3. LP strong duality. Let $f_0, f_1, \dots, f_m$ be affine. If the primal or the dual is feasible, then $p^* = d^*$. If, in addition, p^* is finite, then both are attained.

Theorem 3 follows from Farkas' lemma, which is its feasibility version. So it is not surprising that Theorem 3 is only a slight extension.

Question 2. Let $f_0(x) = c^T x$ and $f_i(x) = a_i^T x - b_i$, so that the primal is

$$\min_{x \in \mathbb{R}^n} c^T x \text{ s.t. } Ax \leq b.$$

Derive the dual problem.

Theorem 4. LP strong duality, restated. For the primal and dual linear programs

$$(P) \quad p^* = \inf_{x \in \mathbb{R}^n} c^T x \text{ s.t. } Ax \leq b,$$

$$(D) \quad d^* = \sup_{y \in \mathbb{R}^m} -b^T y \text{ s.t. } A^T y + c = 0, \quad y \geq 0,$$

if (P) or (D) is feasible, then $p^* = d^*$. If, in addition, p^* is finite, then both are attained.

Example 5. Find p^* and d^* for

$$\min_{x \in \mathbb{R}^2} x_2 \text{ s.t. } x_1 \leq 1, \quad -x_1 \leq -2.$$

Both problems are infeasible. This is the only way an LP can have a nonzero gap, and it is a corner case.

Question 3. Let x be feasible for (P) and y be feasible for (D), neither of them necessarily optimal,

$$(P) \quad \min_{x \in \mathbb{R}^n} c^T x \text{ s.t. } Ax \leq b,$$

$$(D) \quad \max_{y \in \mathbb{R}^m} -b^T y \text{ s.t. } A^T y + c = 0, \ y \geq 0.$$

What is the duality gap?

For feasible x and y , the duality gap is a sum of nonnegative terms,

$$c^T x + b^T y = \sum_{i=1}^m y_i (b_i - a_i^T x).$$

So the gap is zero exactly when every term is zero.

Complementary slackness. If

$$y_i (b_i - a_i^T x) = 0 \quad \text{for every } i,$$

then the duality gap is zero, and x and y are both optimal. A constraint with positive weight must be tight. A constraint that is slack must get zero weight.

Strict complementarity. The pair is strictly complementary if exactly one of y_i and $b_i - a_i^T x$ is zero for every i ,

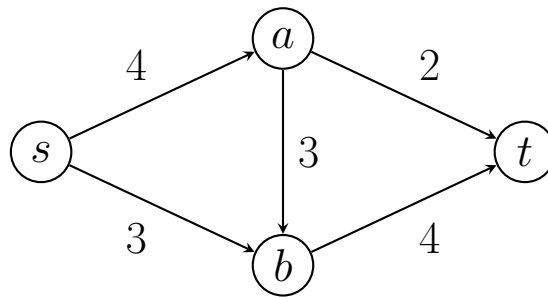
$$y_i + (b_i - a_i^T x) > 0 \quad \text{for every } i.$$

Theorem 6. Goldman–Tucker. If p^* is finite, then (P) and (D) have optimal x_* and y_* that are strictly complementary.

Linear programs are special in this respect. For other problems, even convex ones, strict complementarity can fail, and it is an assumption.

4 Example: Routing traffic

A data center s sends traffic to a city t . Each link has a capacity in Gb/s. Traffic on a route uses every link on that route.



Let x_1, x_2, x_3 be the traffic on the routes $s \rightarrow a \rightarrow t$, $s \rightarrow b \rightarrow t$, and $s \rightarrow a \rightarrow b \rightarrow t$. Maximizing the total traffic is the LP

$$\begin{array}{ll}
 \min_{x \in \mathbb{R}^3} & -x_1 - x_2 - x_3 \\
 \text{s.t.} & x_1 + x_3 \leq 4 \quad (s \rightarrow a), \quad x_2 \leq 3 \quad (s \rightarrow b), \\
 & x_1 \leq 2 \quad (a \rightarrow t), \quad x_2 + x_3 \leq 4 \quad (b \rightarrow t), \\
 & x_3 \leq 3 \quad (a \rightarrow b), \quad x \geq 0.
 \end{array}$$

Give each constraint of the routing LP a dual variable, shown in brackets:

$$\begin{array}{ll} \min_{x \in \mathbb{R}^3} & -x_1 - x_2 - x_3 \\ \text{s.t.} & x_1 + x_3 \leq 4 \quad [y_{sa}], \quad x_2 \leq 3 \quad [y_{sb}], \\ & x_1 \leq 2 \quad [y_{at}], \quad x_2 + x_3 \leq 4 \quad [y_{bt}], \\ & x_3 \leq 3 \quad [y_{ab}], \quad -x \leq 0 \quad [w]. \end{array}$$

Example 7. Write the routing LP in the standard form $\min c^T x$ s.t. $Ax \leq b$.

In standard form, the routing LP has

$$c = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}, \quad A = \begin{bmatrix} N \\ -I \end{bmatrix}, \quad b = \begin{bmatrix} u \\ 0 \end{bmatrix},$$
$$N = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad u = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 4 \\ 3 \end{bmatrix}.$$

Question 4. Derive the dual $\max -b^T(y, w)$ s.t. $A^T(y, w) + c = 0$, $(y, w) \geq 0$.

The dual of the routing LP is

$$\begin{aligned} \max_{y, w \geq 0} \quad & -4y_{sa} - 3y_{sb} - 2y_{at} - 4y_{bt} - 3y_{ab} \\ \text{s.t.} \quad & y_{sa} + y_{at} - w_1 = 1, \\ & y_{sb} + y_{bt} - w_2 = 1, \\ & y_{sa} + y_{bt} + y_{ab} - w_3 = 1. \end{aligned}$$

Example 8. The routing $\hat{x} = (2, 2, 2)$ delivers 6 Gb/s. Use complementary slackness to find a dual point with the same value.

The certificate puts weight 1 on the two links into t . Every route uses one of them, and their capacities add up to $2 + 4 = 6$ Gb/s.

5 Slater's condition

Theorem 9. Slater's condition. Let $f_0, f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex. For the primal and dual problems

$$p^* = \inf_{x \in \mathbb{R}^n} f_0(x) \text{ s.t. } Ax \leq b, \quad f_i(x) \leq 0 \text{ for all } i,$$

$$d^* = \sup_{y \geq 0, z \geq 0} \inf_{x \in \mathbb{R}^n} \left\{ f_0(x) + \sum_{i=1}^m y_i f_i(x) + z^T (Ax - b) \right\},$$

suppose some $\bar{x}$ satisfies

$$A\bar{x} \leq b, \quad f_i(\bar{x}) < 0 \text{ for all } i.$$

Then $p^* = d^*$. If p^* is finite, the dual is attained.

We state this without proof (Boyd and Vandenberghe, Sec. 5.3.2).

- The linear inequalities are non-strict. An equality $Gx = h$ enters as $Gx \leq h$ and $-Gx \leq -h$, which can never hold strictly.
- Slater's condition is sufficient for strong duality, but not necessary.

A **second-order cone program** (SOCP) minimizes a linear function subject to constraints

$$\|A_i x + b_i\| \leq c_i^T x + d_i.$$

Each constraint function is a norm plus an affine function, so it is convex.

Example 10. Derive the dual of the SOCP

$$p^* = \min_{x \in \mathbb{R}^n} c^T x \text{ s.t. } \|x\| \leq a^T x + b.$$

The SOCP of Example 10 and its dual are

$$p^* = \min_{x \in \mathbb{R}^n} c^T x \text{ s.t. } \|x\| \leq a^T x + b,$$
$$d^* = \sup_{y \geq 0} -by \text{ s.t. } \|c - ya\| \leq y.$$

Question 5. When does Slater's condition hold?

So strong duality holds for most SOCPs, but not for all.

Question 6. Recall Example 2,

$$p^* = \min_{x \in \mathbb{R}^2} x_1 \text{ s.t. } \left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\| \leq x_2.$$

Does Slater's condition hold? What is its dual?

Example 11. For which b does

$$\min_{x \in \mathbb{R}} x \text{ s.t. } M(x) = \begin{bmatrix} 1 & x \\ x & b \end{bmatrix} \succeq 0$$

satisfy Slater's condition?

A symmetric matrix is positive semidefinite exactly when its smallest eigenvalue is nonnegative.

For a matrix constraint $M(x) \succeq 0$, Slater's condition asks for a feasible point with $M(\bar{x}) \succ 0$, that is, a positive definite feasible point.

6 The trust-region subproblem

Trust-region Newton chooses each step by minimizing a quadratic model over a ball. With gradient c , Hessian H , and radius $\Delta > 0$, the **trust-region subproblem** is

$$p^* = \min_{v \in \mathbb{R}^n} c^T v + \frac{1}{2} v^T H v \text{ s.t. } \frac{1}{2} \left(\|v\|^2 - \Delta^2 \right) \leq 0.$$

It is nonconvex whenever H has a negative eigenvalue.

Question 7. Derive the dual.

The trust-region subproblem and its dual are

$$p^* = \min_{v \in \mathbb{R}^n} c^T v + \frac{1}{2} v^T H v \text{ s.t. } \frac{1}{2} \left(\|v\|^2 - \Delta^2 \right) \leq 0,$$
$$d^* = \sup_{y \geq 0, H+yI \succ 0} -\frac{1}{2} c^T (H + yI)^{-1} c - \frac{1}{2} y \Delta^2.$$

Question 8. When does Slater's condition hold?

For indefinite H , Slater's theorem says nothing. The next result covers every H .

Theorem 12. S-lemma. Let $q_0, q_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ be quadratic functions, not necessarily convex, and suppose $q_1(\bar{v}) < 0$ for some $\bar{v}$. Then

$$q_0(v) \geq 0 \text{ whenever } q_1(v) \leq 0$$

if and only if some $y \geq 0$ satisfies

$$q_0(v) + y q_1(v) \geq 0 \quad \text{for every } v.$$

We state this without proof (Boyd and Vandenberghe, App. B).

Question 9. Show that the trust-region subproblem has $p^* = d^*$ for every symmetric H .

7 Summary

Strong duality means $p^* = d^*$. With dual attainment, every globally optimal point has a certificate with zero gap.

Linear programs. The dual of $\min c^T x$ s.t. $Ax \leq b$ is $\max -b^T y$ s.t. $A^T y + c = 0$, $y \geq 0$. If either is feasible, then $p^* = d^*$. The gap is $\sum_i y_i(b_i - a_i^T x)$, so complementary slackness gives zero gap. LPs also have strictly complementary solutions.

Slater's condition. For convex $f_0, \dots, f_m$, a point with $A\bar{x} \leq b$ and $f_i(\bar{x}) < 0$ for all i gives $p^* = d^*$, and dual attainment when p^* is finite. Convexity alone is not enough.

Cone constraints. For an SOCP, Slater's condition asks for a point that satisfies every cone constraint strictly. For a matrix constraint $M(x) \succeq 0$, it asks for a point with $M(\bar{x}) \succ 0$.

Trust-region subproblem. Slater's condition covers only $H \succeq 0$. The S-lemma gives $p^* = d^*$ for every H .