

# Chapter 3. Duality in convex optimization

Version: 30-10-2011

---

## 3.1 Lagrangean dual

Recall the program ( $J = \{1, \dots, m\}$ )

$$(CO) \quad \min_{x \in \mathcal{C}} \{ f(x) \mid g_j(x) \leq 0, j \in J \}$$

with convex functions  $f, g_j$ , a convex set  $\mathcal{C}$ , minimal value  $v(CO)$ , feasible set  $\mathcal{F}$ ; the Lagrangean function:

$$L(x, y) = f(x) + \sum_{j=1}^m y_j g_j(x), \quad x \in \mathcal{C}, y \geq 0$$

$L(x, y)$  is **convex** in  $x$  (for fixed  $y$ ) and **linear** in  $y$  (for fixed  $x$ ).

## Defining

$$\psi(\mathbf{y}) := \inf_{\mathbf{x} \in \mathcal{C}} \left\{ f(\mathbf{x}) + \sum_{j=1}^m y_j g_j(\mathbf{x}) \right\}$$

the Lagrangian-dual problem of (CO) can be written as:

$$(D) : \sup_{\mathbf{y} \geq 0} \psi(\mathbf{y})$$

The (*sup*)value of (D) is denoted by  $v(D)$ .

**L.3.2** *(Even without convexity assumptions on (CO))*

The function  $\psi(\mathbf{y})$  is a concave function and thus, (D), the Lagrangian dual of (CO) is a convex program.

**Note:** By definition  $\psi(\mathbf{y})$  is called concave if  $-\psi(\mathbf{y})$  is convex. So that  $\max_{\mathbf{y}} \psi \equiv \min_{\mathbf{y}} -\psi$  (a convex problem).

**Note:!!** The following weak and strong duality results are more general than the results of Th.3.4, Th.3.6 in KRT.

**Th.3.4'** [weak duality] (without convexity assumptions on  $(CO)$ )

$$\underbrace{\sup_{y \geq 0} \psi(y)}_{v(D)} \leq \underbrace{\inf_{x \in C} \{f(x) \mid g_j(x) \leq 0, j \in J\}}_{v(CO)}.$$

In particular:

- if equality holds for  $\bar{y} \geq 0$ :

$$\psi(\bar{y}) = \inf_{x \in C} \{f(x) \mid g_j(x) \leq 0, j \in J\}$$

then  $\bar{y}$  is a solution of  $(D)$ .

- If we have  $\psi(\bar{y}) = f(\bar{x})$  for  $\bar{y} \geq 0, \bar{x} \in \mathcal{F}$  then  $\bar{y}, \bar{x}$  are optimal solutions of  $(D), (CO)$ , respectively.

**Def.** The value  $v(CO) - v(D) \geq 0$  is called duality gap. (This gap may be positive).

**We show** in the next theorem that there is no duality gap if  $(CO)$  satisfies the Slater condition

**Th.2** [*strong duality*] **Suppose (CO) satisfies the Slater condition ( so  $v(\text{CO}) < \infty$ ). Then**

$$\sup_{y \geq 0} \psi(y) = \inf_{x \in C} \{ f(x) \mid g_j(x) \leq 0, j \in J \}$$

**If moreover  $v(\text{CO})$  is bounded from below (i.e., it is bounded), then the dual optimal value is attained.**

*Proof.* (For the case  $-\infty < v(\text{CO}) < \infty$ ) Let  $a = v(\text{CO})$  be the optimal value of (CO). From the Farkas Lemma 2.25 it follows that there exists a vector  $\bar{y} = (\bar{y}_1; \dots; \bar{y}_m) \geq 0$  such that

$$L(x, \bar{y}) = f(x) + \sum_{j=1}^m \bar{y}_j g_j(x) \geq a, \quad \forall x \in C, \quad \text{or} \quad \inf_{x \in C} L(x, \bar{y}) \geq a.$$

By the definition of  $\psi(\bar{y})$  this implies  $\psi(\bar{y}) \geq a$ . Using the weak duality theorem, it follows  $a \leq \psi(\bar{y}) \leq v(D) \leq v(\text{CO}) = a$  and thus  $a = \psi(\bar{y}) = v(\text{CO})$ . So,  $\bar{y}$  is a maximizer of (D).  $\diamond$

**Note** that the existence of a minimizer of (the primal) (CO) is not assured in Th. 2.

**Example 1.** (Slater holds so zero duality gap; solutions of (D) and (CO) exist.)

**Consider**

$$(CO) \min \left\{ f := x_1^2 + x_2^2 \mid x_1 + x_2 \geq 1, (x_1, x_2) \in \mathbb{R}^2 \right\}.$$

Here:  $L(x, y) = x_1^2 + x_2^2 + y(1 - x_1 - x_2)$ . Since  $L(x, y)$  is convex in  $x$ , (for fixed  $y$ )  $L(x, y)$  is minimal if and only if  $\nabla_x L(x, y) = 0$ , e.g. if  $2x_1 - y = 0$  and  $2x_2 - y = 0$ . This yields

$$x_1 = \frac{y}{2}, \quad x_2 = \frac{y}{2}.$$

Substitution gives  $\psi(y) = \frac{y^2}{2} + y(1 - y) = y - \frac{y^2}{2}$ .

So the dual problem is:  $\max \left\{ y - \frac{y^2}{2} \mid y \geq 0 \right\}$

with maximizer  $y = 1$ . So  $x_1 = x_2 = \frac{1}{2}$  is the minimizer of (CO) and  $\mathbf{v}(CO) = \mathbf{v}(D) = 1/2$ .

**Example 2.** *(No Slater but without duality gap)*

$$(\mathbf{CO}) \quad \min x \quad \text{s.t.} \quad x^2 \leq 0, \quad x \in \mathbb{R}$$

This **(CO)** problem is not Slater regular and has solution  $\bar{x} = 0$  with  $v(\mathbf{CO}) = 0$ . On the other hand, we have

$$\psi(\mathbf{y}) = \inf_{x \in \mathbb{R}} (x + yx^2) = \begin{cases} -\frac{1}{4y} & \text{for } y > 0 \\ -\infty & \text{for } y = 0. \end{cases}$$

We see that  $\psi(\mathbf{y}) < 0$  for all  $\mathbf{y} \geq 0$ . One has

$$\sup \{\psi(\mathbf{y}) \mid \mathbf{y} \geq 0\} = 0.$$

So the Lagrange-dual (D) has the same optimal value as the primal problem. In spite of the lack of Slater regularity there is no duality gap; however a solution of (D) doesn't exist.

## 3.2 The Wolfe-dual

We reconsider (CO) with  $C = \mathbb{R}^n$ ,  $f, g_j \in C^1$ , convex and its Lagrangean dual:

$$(D) \quad \sup_{y \geq 0} \{ \psi(y) := \inf_{x \in \mathbb{R}^n} L(x, y) \}$$

where  $L(x, y) = f(x) + \sum_{j=1}^m y_j g_j(x)$ .

**Recall:**  $\bar{x}$  is a minimizer of  $\psi(y)$  iff  $\nabla_x L(\bar{x}, y) = 0$ .

**In that case:**  $\inf_{x \in C} L(x, y) = L(\bar{x}, y)$ . So, (provided, a solution of  $\psi(y)$  exists) the Lagrangean dual takes the form:

$$(WD) \quad \sup_{x, y} \{ f(x) + \sum_{j=1}^m y_j g_j(x) \} \quad \text{s.t.} \quad y \geq 0 \quad \text{and} \\ \nabla_x L(x, y) \equiv \nabla f(x) + \sum_{j=1}^m y_j \nabla g_j(x) = 0 .$$

**Def.3.8** (WD) is called the Wolfe-Dual of (CO). (In general, (WD) is not a convex problem.)

**Th.3.9** (*weak duality*) If  $\hat{x}$  is feasible for (CO) and  $(\bar{x}, \bar{y})$  is a feasible point of (WD) then

$$L(\bar{x}, \bar{y}) \leq f(\hat{x}).$$

If  $L(\bar{x}, \bar{y}) = f(\hat{x})$  holds then  $(\bar{x}, \bar{y})$  is a solution of (WD) and  $\hat{x}$  is a minimizer of (CO)

From Corollary 2.33 (or Theorem 2.30) we obtain:

**Th.3.10** (*partial strong duality*) Let (CO) satisfy the Slater condition. If the feasible point  $\bar{x} \in \mathcal{F}$  is a minimizer of (CO) then there exists  $\bar{y} \geq 0$  such that  $(\bar{x}, \bar{y})$  is an optimal solution of (WD) and  $L(\bar{x}, \bar{y}) = f(\bar{x})$ .

**Rem.** Only under additional assumptions (e.g., second order conditions) also the converse of Th.3.10 holds.

**Rem.** Together with the results in Ch. 2, the Th.3.10 also says: If  $(\bar{x}, \bar{y})$  is a KKT point (or equivalently a saddlepoint of  $L(x, y)$ ) then  $(\bar{x}, \bar{y})$  is a maximizer of (WD).

**Example 3.11.** (Application of the Weak Duality for (WD))

Consider the convex optimization problem

$$(CO) \quad \min_{x \in \mathbb{R}^2} x_1 + e^{x_2} \quad \text{s.t.} \quad 3x_1 - 2e^{x_2} \geq 10, \quad x_2 \geq 0.$$

Note that  $\bar{x} = (4, 0)$  is feasible with objective value  $f(\bar{x}) = 5$ . To show that this is the minimizer, by Th. 3.9, it suffices to show that the value of the Wolfe-dual of (CO) is also 5:

$$(WD) \quad \sup_{x, y} \quad L = x_1 + e^{x_2} + y_1(10 - 3x_1 + 2e^{x_2}) - y_2x_2$$
$$\text{s.t.} \quad 1 - 3y_1 = 0$$
$$e^{x_2} + 2e^{x_2}y_1 - y_2 = 0, \quad x \in \mathbb{R}^2, y \geq 0.$$

which is a non-convex problem. The first constraint gives

$y_1 = \frac{1}{3}$ , and thus the second constraint becomes

$$\frac{5}{3}e^{x_2} - y_2 = 0.$$

Now we can eliminate  $y_1$  and  $y_2$  from  $L(x, y)$  to obtain The Lagrangean reduces to

$$L = \frac{5}{3}e^{x_2} - \frac{5}{3}x_2e^{x_2} + \frac{10}{3}.$$

This function has a maximum when

$$L'(x_2) = -\frac{5}{3}x_2e^{x_2} = 0,$$

which implies  $x_2 = 0$  and  $L = 5$ . Hence the optimal value of (WD) is 5 with solution  $(x, y) = (4, 0, \frac{1}{3}, \frac{5}{3})$ . Weak duality implies that  $\bar{x} = (4, 0)$  is minimizer of (CO).

**Remark:** A substitution  $z = e^{x_2} \geq 1$  would lead to a linear problem.

## General remarks on duality:

- The Wolfe-duality concept and duality via saddle points is restricted to the case where a solution  $\bar{x}$  of the primal program (CO) exists.
- The Lagrange duality approach is more general. The existence of a minimizer of (CO) is not required.
- Wolfe's dual is useful to find a suitable form of the dual of a convex program.

In 3.4 [KRT] are some examples with positive duality gaps.

*The following example in particular shows that Wolfe- and Lagrangean-duality are not (always) the same*

Ex. 3.12. (CO)  $\min_{x \in \mathbb{R}^2} e^{-x_2} \quad \text{s.t.} \quad \sqrt{x_1^2 + x_2^2} \leq x_1$

Here the feasible set is  $\mathcal{F} = \{(x_1, x_2) \mid x_2 = 0, x_1 \geq 0\}$  (all feasible points are minimizers) and

$$v(WD) = -\infty < v(D) = 0 < v(CO) = 1.$$

## 3.3 Examples of dual problems

The Wolfe dual can be used to find the Lagrangean dual.

### Duality for Linear Optimization (LP)

Let  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  and  $c, x \in \mathbb{R}^n$ . Consider the LP:

$$(P) \quad \max_y \{ b^T y \mid A^T y \leq c \}.$$

The corresponding (Wolfe-) dual is

$$(D) \quad \min_x \{ c^T x \mid Ax = b, x \geq 0 \}.$$

Ex. Show that (D) is also the Lagrangean dual of (P).

Rem. Since the Slater condition is (automatically) fulfilled, strong (Lagrangean) duality holds if (P) is feasible.

## Convex quadratic programming duality

Consider with positive semidefinite  $Q \in \mathbb{R}^{n \times n}$  the quadratic program

$$(P) \quad \min_x \left\{ c^T x + \frac{1}{2} x^T Q x \mid Ax \geq b, x \geq 0 \right\},$$

with linear constraints and convex object function. The Wolfe-dual is:

$$(D) \quad \max_{y \geq 0, x} \left\{ b^T y - \frac{1}{2} x^T Q x \mid A^T y - Qx \leq c \right\}.$$

Proof: Writing (P) as

$$\min_{x \in \mathbb{R}^n} \left\{ c^T x + \frac{1}{2} x^T Q x \mid b - Ax \leq 0, -x \leq 0 \right\}$$

the Wolfe-dual becomes:  $\max_{y \geq 0, s \geq 0, x} \{ c^T x + \frac{1}{2} x^T Q x + y^T (b - Ax) - s^T x \mid c + Qx - A^T y - s = 0 \}$ .

*Substituting  $c = -Qx + A^T y + s$  in the objective, we get*

$$\begin{aligned} & c^T x + \frac{1}{2} x^T Q x + y^T (b - Ax) - s^T x \\ = & (-Qx + A^T y + s)^T x + \frac{1}{2} x^T Q x + y^T (b - Ax) - s^T x \\ = & (-Qx)^T x + \frac{1}{2} x^T Q x + y^T b \\ = & b^T y - \frac{1}{2} x^T Q x. \end{aligned}$$

*Hence the problem simplifies to*

$$\max_{y \geq 0, s \geq 0, x} \left\{ b^T y - \frac{1}{2} x^T Q x \mid c + Qx - A^T y - s = 0 \right\}.$$

*By eliminating  $s$  this becomes*

$$\max_{y \geq 0, x} \left\{ b^T y - \frac{1}{2} x^T Q x \mid A^T y - Qx \leq c \right\}.$$

## 3.5 Semidefinite optimization

Semidefinite programs (SDP) provide an important generalization of Linear programs.

**Def.** Let  $\mathbb{S}^m$  denote the set of all symmetric matrices. For  $A, B \in \mathbb{S}^m$  we define the inner product  $A \bullet B$  by

$$A \bullet B = \sum_{i,j} a_{ij} b_{ij} \quad (= \text{trace } A \cdot B)$$

$\mathbb{S}_+^m$  denotes the set of all psd. matrices  $A$  ( $A \succeq 0$ ).

**Semidefinite program:** Let  $A_0, A_1, \dots, A_n \in S^m$ ,  $c \in \mathbb{R}^n$  and let  $x \in \mathbb{R}^n$  be the vector of unknowns.

The primal SDP is:

$$(P) \quad \min_x \left\{ c^T x \mid F(x) := -A_0 + \sum_{k=1}^n A_k x_k \succeq 0 \right\}$$

The dual problem is (as we shall see):

$$(D) \quad \max_Z \{ A_0 \bullet Z \mid A_k \bullet Z = c_k, k = 1, \dots, n, Z \succeq 0 \}$$

**Ex.** Show: Both programs, (P) and (D) are convex problems.

**Rem.** If the matrices  $A_0, A_1, \dots, A_n$  are diagonal matrices then (P) is simply a linear program with (linear) dual (D).

Weak duality: We need the following facts:

Exercise. (Let  $A \in \mathbb{S}^m$ ,  $x \in \mathbb{R}^m$ ) Then

- 1  $x^T A x = A \bullet x x^T$
- 2 There is an orthonormal basis of eigenvectors  $x_i \in \mathbb{R}^m$  of  $A$  and corresponding eigenvalues  $\lambda_i$ ,  $i = 1, \dots, m$  such that:

$$A = \sum_{i=1}^m \lambda_i x_i x_i^T .$$

If in addition  $A \succeq 0$  holds then  $A = B B^T$ , where  $B$  is the matrix with columns  $\sqrt{\lambda_i} x_i$ .

- 3  $S, Z \in \mathbb{S}_+^m \Rightarrow S \bullet Z \geq 0$  (and “=” holds iff  $S \cdot Z = 0$ ).

4

$$\inf_{S \in \mathbb{S}_+^m} S \bullet Z = \begin{cases} 0 & \text{if } Z \succeq 0, \\ -\infty & \text{otherwise.} \end{cases}$$

**Th.3.17** (*weak duality*) If  $x \in \mathbb{R}^n$  is primal feasible and  $Z \in \mathbb{R}^{m \times m}$  is dual feasible, then

$$c^T x - A_0 \bullet Z = F(x) \bullet Z \geq 0.$$

Equality holds iff  $F(x) \bullet Z = 0$  (and in this case  $x, Z$  are optimal for  $(P), (D)$ , respectively).

**Remark.** [*Without proof (see, e.g., [FKS, Ch.12])*]  
*Strong duality holds under the assumption that  $(P)$  is solvable and that there exists a strictly feasible  $x^*$ :*

$$F(x^*) \succ 0$$

## 3.6 Duality in cone-linear optimization

**Rem.** A cone-linear optimization problem is a natural generalization of a standard LP

**Recall:** The dual  $\mathcal{K}^*$  of a cone  $\mathcal{K} \subset \mathbb{R}^n$  is defined by:

$$\mathcal{K}^* = \{ \mathbf{y} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{y} \geq 0, \forall \mathbf{x} \in \mathcal{K} \}$$

### Examples of cones

- the nonnegative orthant:

$$\mathbb{R}_+^m = \{ \mathbf{x} \in \mathbb{R}^m \mid \mathbf{x} \geq \mathbf{0} \}.$$

- the Lorentz (or ice-cream) cone;

$$\mathbb{L}_+^m = \left\{ \mathbf{x} \in \mathbb{R}^m \mid \mathbf{x}_m \geq \sqrt{\sum_{i=1}^{m-1} \mathbf{x}_i^2} \right\}$$

- the cone of semidefinite matrices:

$$\mathbb{S}_+^m = \{ A \in \mathbb{R}^{m \times m} \mid A = A^T, \mathbf{x}^T A \mathbf{x} \geq 0, \forall \mathbf{x} \in \mathbb{R}^m \}$$

We define the primal cone-linear program as

$$(P) \quad \min \left\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} - \mathbf{b} \in \mathcal{C}_1, \mathbf{x} \in \mathcal{C}_2 \right\}$$

where  $\mathcal{C}_1 \subset \mathbb{R}^m, \mathcal{C}_2 \subset \mathbb{R}^n$  are closed convex cones.

The dual of this program is:

$$(D) \quad \max \left\{ \mathbf{b}^T \mathbf{y} \mid \mathbf{c} - \mathbf{A}^T \mathbf{y} \in \mathcal{C}_2^*, \mathbf{y} \in \mathcal{C}_1^* \right\}$$

**Th.3.18** (*weak duality*) If  $\mathbf{x} \in \mathbb{R}^n$  is feasible for (P) and  $\mathbf{y} \in \mathbb{R}^m$  is feasible for (D) then

$$\mathbf{c}^T \mathbf{x} - \mathbf{b}^T \mathbf{y} = \mathbf{x}^T (\mathbf{c} - \mathbf{A}^T \mathbf{y}) + \mathbf{y}^T (\mathbf{Ax} - \mathbf{b}) \geq 0.$$