Version: 10-10-2012

3.1 Lagrangean dual

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

$$(CO) \qquad \min_{x \in \mathcal{C}} \left\{ f(x) \mid g_j(x) \le 0, \;\; j \in J \right\}$$

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

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

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

Defining

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

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

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

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

L.3.2 (*Even without convexity assumptions on (CO)*)
The function $\psi(y)$ is a concave function and thus, (D), the Lagrangean dual of (CO) is a convex program.

Note: *By definition $\psi(y)$ is called concave if $-\psi(y)$ is convex. So that $\max_y \psi \equiv \min_y -\psi$ (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))*

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

In particular:

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

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

  then $\overline{y}$ is a solution of (D).
- **If we have $\psi(\overline{y}) = f(\overline{x})$ for $\overline{y} \geq 0$, $\overline{x} \in \mathcal{F}$ then $\overline{y}$, $\overline{x}$ 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* $v(CO) < \infty$*)*. **Then**

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

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

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

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

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

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

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

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

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*

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

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* $\boldsymbol{v(CO) = v(D) = 1/2}$.

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

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

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

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

We see that $\psi(y) < 0$ *for all* $y \ge 0$*. One has*

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

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:

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

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

Recall: $\overline{x}$ is a minimizer of $\psi(y)$ iff $\nabla_x L(\overline{x}, y) = 0$.
In that case: $\inf_{x \in \mathcal{C}} L(x, y) = L(\overline{x}, y)$. So, *(provided, a solution of* $\psi(y)$ *exists)* the Lagrangean dual takes the form:

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

$$\nabla_x L(x, y) \equiv \nabla f(x) + \sum_{j=1}^m y_j \nabla g_j(x) = 0 \, .$$

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 $(\overline{x}, \overline{y})$ is a feasible point of (WD) then

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

If $L(\overline{x}, \overline{y}) = f(\hat{x})$ holds then $(\overline{x}, \overline{y})$ 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 $\overline{x} \in \mathcal{F}$ is a minimizer of (CO) then there exists $\overline{y} \geq 0$ such that $(\overline{x}, \overline{y})$ is an optimal solution of (WD) and $L(\overline{x}, \overline{y}) = f(\overline{x})$.

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 $(\overline{x}, \overline{y})$ is a KKT point** *(or equivalently a saddlepoint of $L(x, y)$)* **then $(\overline{x}, \overline{y})$ is a maximizer of (WD).**

Example 3.11. *(Application of the Weak Duality for (WD))*
Consider the convex optimization problem

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

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

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

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

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

This function has a maximum when

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

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 $\overline{x} = (4, 0)$ 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 $\overline{x}$ 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.**

[KRT,Sec.3.4] gives some examples with positive duality gap. *The following example in particular shows that Wolfe- and Lagrangean-duality are not (always) the same*

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

Here the feasible set is $\mathcal{F} = \{(x_1, x_2) \mid x_2 = 0, x_1 \ge 0\}$ (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:

$$(P) \qquad \max_y \left\{ b^T y \mid A^T y \le c \right\}.$$

The corresponding (Wolfe-) dual is

$$(D) \qquad \min_x \{ c^T x \mid Ax = b \,,\; x \ge 0 \}.$$

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

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

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

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

Proof: *Writing (P) as*

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

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*

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

Hence the problem simplifies to

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

By eliminating s this becomes

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

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 $A_0, A_1, \ldots, A_n \in \mathbb{S}^m$, $c \in \mathbb{R}^n$ 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*):

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

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

Rem. *If the matrices* $A_0, A_1, \ldots, A_n$ *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 λ_i, $i = 1, .., m$ such that:
$$A = \sum_{i=1}^{m} \lambda_i x_i x_i^T \ .$$
If in addition $A \succeq 0$ holds then $A = BB^T$, 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 *(week 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 $F(x) \cdot Z = 0$ (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^*:

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