Lecture: Duality

1. Duality for LP in canonical form.
   - Duality, complementarity, and interpretations.
2. Duality for general LP problems.
3. Duality for LP in standard form

Reading instructions for the slides:

- The geometrical interpretation of the complementarity conditions will be considered in more detail later in the course for non-linear problems. It is probably easier to understand then, so do not spend to much time on it here.
- The economical interpretation of primal and dual problems can be skimmed through

Duality for LP in canonic form

Primal	Dual
$(P) \quad \begin{array}{ll} \text{minimize} & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array}$	$(D) \quad \begin{array}{ll} \text{maximize} & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{s.t.} & \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c} \\ & \mathbf{y} \geq 0 \end{array}$

Primal feasible region

$\mathcal{F}_P = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{A}\mathbf{x} \geq \mathbf{b};\ \mathbf{x} \geq 0\}$

$\hat{\mathbf{x}} \in \mathcal{F}_P$ is the optimal solution if

$\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} \leq \mathbf{c}^\mathsf{T}\mathbf{x},\ \forall \mathbf{x} \in \mathcal{F}_P$

Dual feasible region

$\mathcal{F}_D = \{\mathbf{y} \in \mathbf{R}^m : \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c};\ \mathbf{y} \geq 0\}$

$\hat{\mathbf{y}} \in \mathcal{F}_D$ is the optimal solution if

$\mathbf{b}^\mathsf{T}\hat{\mathbf{y}} \geq \mathbf{b}^\mathsf{T}\mathbf{y},\ \forall \mathbf{y} \in \mathcal{F}_D$

How are the primal and dual related?

Consider the following pair of primal and dual optimization problems

Primal

$$
(P) \quad \begin{aligned} \text{minimize} \quad & x_1 + x_2 \\ \text{s.t.} \quad & \tfrac{1}{40}x_1 + \tfrac{1}{50}x_2 \geq 200 \\ & \tfrac{1}{60}x_1 + \tfrac{1}{50}x_2 \geq 400 \\ & x_1, x_2 \geq 0 \end{aligned}
$$

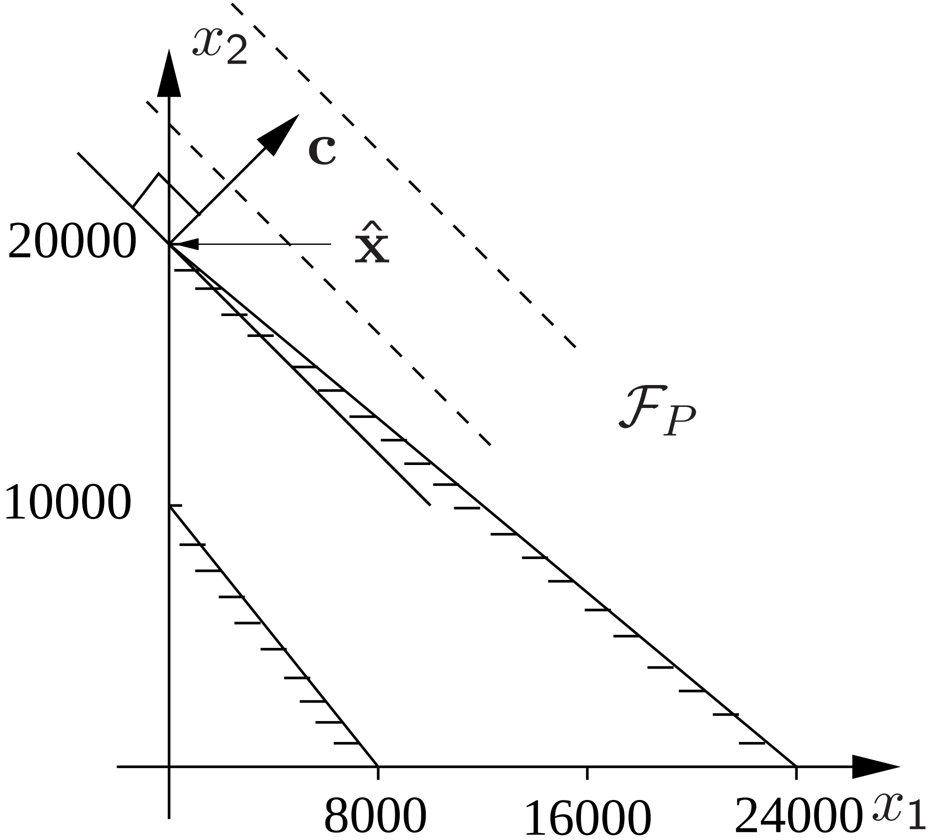
Dual

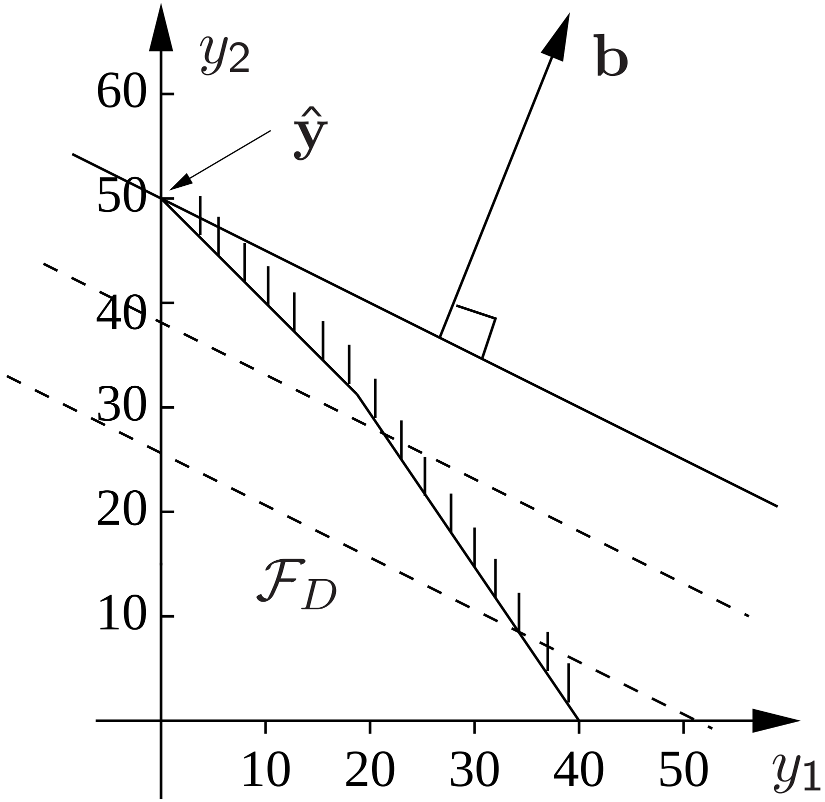
$$
(D) \quad \begin{aligned} \text{maximize} \quad & 200y_1 + 400y_2 \\ \text{s.t.} \quad & \tfrac{1}{40}y_1 + \tfrac{1}{60}y_2 \leq 1 \\ & \tfrac{1}{50}y_1 + \tfrac{1}{50}y_2 \leq 1 \\ & y_1, y_2 \geq 0 \end{aligned}
$$

What is the relation between

- the optimal values?
- the objective function values (for an arbitrary feasible point) ?
- the optimal solutions and the constraints at the optimum?

Illustration of the feasible region, iso-cost lines, and the optimal solution





- The Optimal values are the same $\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = \mathbf{b}^\mathsf{T}\hat{\mathbf{y}} = 20000$
- For each $\mathbf{x} \in \mathcal{F}_P$ and $\mathbf{y} \in \mathcal{F}_D$ it holds that $\mathbf{c}^\mathsf{T}\mathbf{x} \geq \mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = \mathbf{b}^\mathsf{T}\hat{\mathbf{y}} \geq \mathbf{b}^\mathsf{T}\mathbf{y}$

Let

$$\hat{\mathbf{s}} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{b} = \begin{bmatrix} 200 \\ 0 \end{bmatrix} \qquad (S)$$

$$\hat{\mathbf{r}} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \mathbf{c} - \mathbf{A}^\mathsf{T}\hat{\mathbf{y}} = \begin{bmatrix} \frac{1}{6} \\ 0 \end{bmatrix} \qquad (R)$$

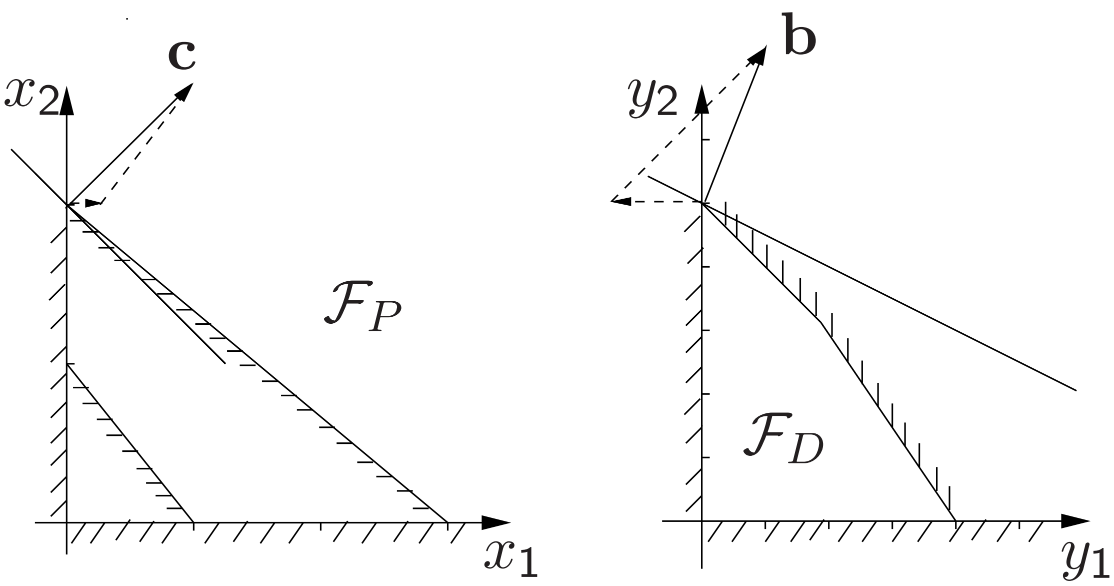
The optimal solution satisfies

1. $\hat{\mathbf{s}} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{b} \geq \mathbf{0}$, $\hat{\mathbf{x}} \geq \mathbf{0}$ (primal feasibility)
2. $\hat{\mathbf{r}} = \mathbf{c} - \mathbf{A}^\mathsf{T}\hat{\mathbf{y}} \geq \mathbf{0}$, $\hat{\mathbf{y}} \geq \mathbf{0}$ (dual feasibility)
3. $x_j r_j = 0$, $j = 1, 2$ and $y_i s_i = 0$, $i = 1, 2$ (complementarity)

The complementarity conditions can be interpreted geometrically. From (R) and (S) we get $\mathbf{c} = \hat{\mathbf{r}} + \mathbf{A}^T\hat{\mathbf{y}}$ and $\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{s}}$, *i.e.*,

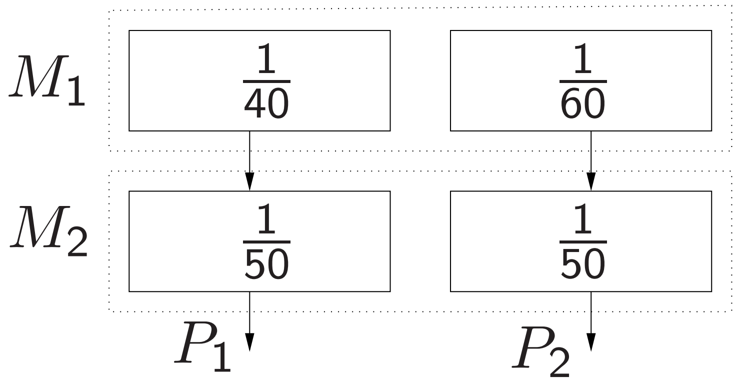
$$(R)\quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{6} + \begin{bmatrix} \frac{1}{60} \\ \frac{1}{50} \end{bmatrix} 50 \qquad (S)\quad \begin{bmatrix} 200 \\ 400 \end{bmatrix} = \begin{bmatrix} \frac{1}{50} \\ \frac{1}{50} \end{bmatrix} 20000 - \begin{bmatrix} 1 \\ 0 \end{bmatrix} 200$$

it follows that the gradients of the objective functions ($\mathbf{c}$ and $\mathbf{b}$ respectively) can be given as linear combinations of the active constraints. This is illustrated in the figure. (Note: incorrect scale)



Economical interpretation

- We consider the production of two products P_1 and P_2
- Machines M_1 and M_2 are used for the production of P_1 and P_2
- The Production capacities for the machines are [day/unit]:



- The price for the products [SEK/unit]:

product	price
P_1	$b_1 = 200$
P_2	$b_2 = 400$

- In the dual, production volume is determined for profitmaximization.

Assume that you have the option of renting out the machines M_1 and M_2. What is the lowest possible price x_k SEK/day for renting out M_k, $k = 1, 2$.

The lowest price is determined from the following optimization problem

$$\begin{aligned} \text{minimize} \quad & x_1 + x_2 \\ \text{s.t.} \quad & \tfrac{1}{40}x_1 + \tfrac{1}{50}x_2 \geq 200 \\ & \tfrac{1}{60}x_1 + \tfrac{1}{50}x_2 \geq 400 \\ & x_1, x_2 \geq 0 \end{aligned}$$

where the constraints ensure that you earn at least as much by renting out the machines as the profit you would have made producing products P_1 and P_2.

- The primal and dual can be seen as alternatives to each others.
  - If your business idea is to rent out M_1 and M_2, then the dual provides information about the lowest production volume necessary for instead using the machines for producing P_1 and P_2.
  - if your business idea is to produce P_1 and P_2, then the primal provides information of the lowest renting price that makes it profitable to rent out the machines instead. (break-even price)
- What is called primal and dual is actually arbitrary, since it can be shown that the dual of the dual is the primal.
  (at least for linear programs)

Theoretical results

The following result from the book confirms the relations in the example:

Proposition 6.1 (Weak duality) For every $\mathbf{x} \in \mathcal{F}_P$ and every $\mathbf{y} \in \mathcal{F}_D$ it holds that $\mathbf{c}^\mathsf{T}\mathbf{x} \geq \mathbf{b}^\mathsf{T}\mathbf{y}$.

A direct implication from this is that, if $\hat{\mathbf{x}} \in \mathcal{F}_P$ and $\hat{\mathbf{y}} \in \mathcal{F}_D$ and $\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = \mathbf{b}^\mathsf{T}\hat{\mathbf{y}}$, then $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are optimal to (P) and (D), respectively.

Theorem 6.3 *(The duality Theorem)*
There are four alternatives for the optimization problems (P) *and* (D)

(a) $\mathcal{F}_P \neq \emptyset$, $\mathcal{F}_d \neq \emptyset$. *Then there is (at least) one optimal solution* $\hat{\mathbf{x}}$ *to* (P) *and (at least) one optimal solution to (D).*
They satisfy $\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = \mathbf{b}^\mathsf{T}\hat{\mathbf{y}}$.

(b) $\mathcal{F}_P \neq \emptyset$ *but* $\mathcal{F}_D = \emptyset$. *Then the primal is unbounded (from below).*

(c) $\mathcal{F}_D \neq \emptyset$ *but* $\mathcal{F}_P = \emptyset$. *Then the dual is unbounded (from above).*

(d) $\mathcal{F}_P = \emptyset$ *and* $\mathcal{F}_D = \emptyset$.

- The proof of (a) is non-trivial and is usually based on Farkas' Lemma.
- Case (b) can be interpreted as a consequence of the lack of a lower limit for the primal due to the lack of solutions to the dual.
  Similar interpretation holds for (c).

Notation

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{a}}_1 & \dots & \hat{\mathbf{a}}_n \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{a}}_1^\mathsf{T} \\ \vdots \\ \bar{\mathbf{a}}_m^\mathsf{T} \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} s_1 \\ \vdots \\ s_m \end{bmatrix},$$

$$\mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$$

Theorem 6.5 *(The Complementarity Theorem)*
$\mathbf{x} \in \mathbf{R}^n$ *is optimal for* (P) *and* $\mathbf{y} \in \mathbf{R}^m$ *is optimal for* (D) *if and only if*

$$x_j \geq 0, \; r_j \geq 0, \; x_j r_j = 0, \quad j = 1, \ldots, n,$$
$$y_i \geq 0, \; s_i \geq 0, \; y_i s_i = 0, \quad i = 1, \ldots, m,$$

where $\mathbf{s} = \mathbf{A}\mathbf{x} - \mathbf{b}$ *och* $\mathbf{r} = \mathbf{c} - \mathbf{A}^\mathsf{T}\mathbf{y}$.

Terminology:

- Non-negativity constraints for the primal: $x_j \geq 0$, $j = 1, \ldots, n$.
- General constraints for the primal: $s_i = \bar{\mathbf{a}}_i^\mathsf{T}\mathbf{x} - b_i \geq 0$, $i = 1, \ldots, m$.
- Non-negativity constraints for the dual: $y_i \geq 0$, $i = 1, \ldots, m$.
- General constraints for the dual: $\mathbf{r}_j = \mathbf{c}_j - \hat{\mathbf{a}}_j^\mathsf{T}\mathbf{y} \geq 0$, $j = 1, \ldots, n$.

A consequence of the complementarity Theorem is that

- if $r_j = c_j - \hat{\mathbf{a}}_j^\mathsf{T}\mathbf{y} > 0$ then $x_j = 0$,
  i.e., the j:th non-negativity constraint of the primal is active.

- if $y_i > 0$ then $s_i = \bar{\mathbf{a}}_i^\mathsf{T}\mathbf{x} - b_i = 0$,
  i.e., the i:th general constraint of the primal is active.

- if $s_i = \bar{\mathbf{a}}_i^\mathsf{T}\mathbf{x} - b_i > 0$ then $y_i = 0$,
  i.e., the i:th non-negativity constraint of the dual is active.

- if $x_j > 0$ then $r_j = c_j - \hat{\mathbf{a}}_j^\mathsf{T}\mathbf{y} = 0$,
  i.e., the j:th general constraint of the dual is active.

Geometric interpretation of the complementarity constraints

Let $\mathbf{e}_j = \begin{bmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \end{bmatrix}^\mathsf{T} \in \mathbf{R}^n$, where the 1 is in position j. Then

$$\mathbf{c} = \mathbf{r} + \mathbf{A}^\mathsf{T}\mathbf{y} = \sum_{j:r_j>0} \mathbf{e}_j r_j + \sum_{i:y_i>0} \bar{\mathbf{a}}_i y_i$$

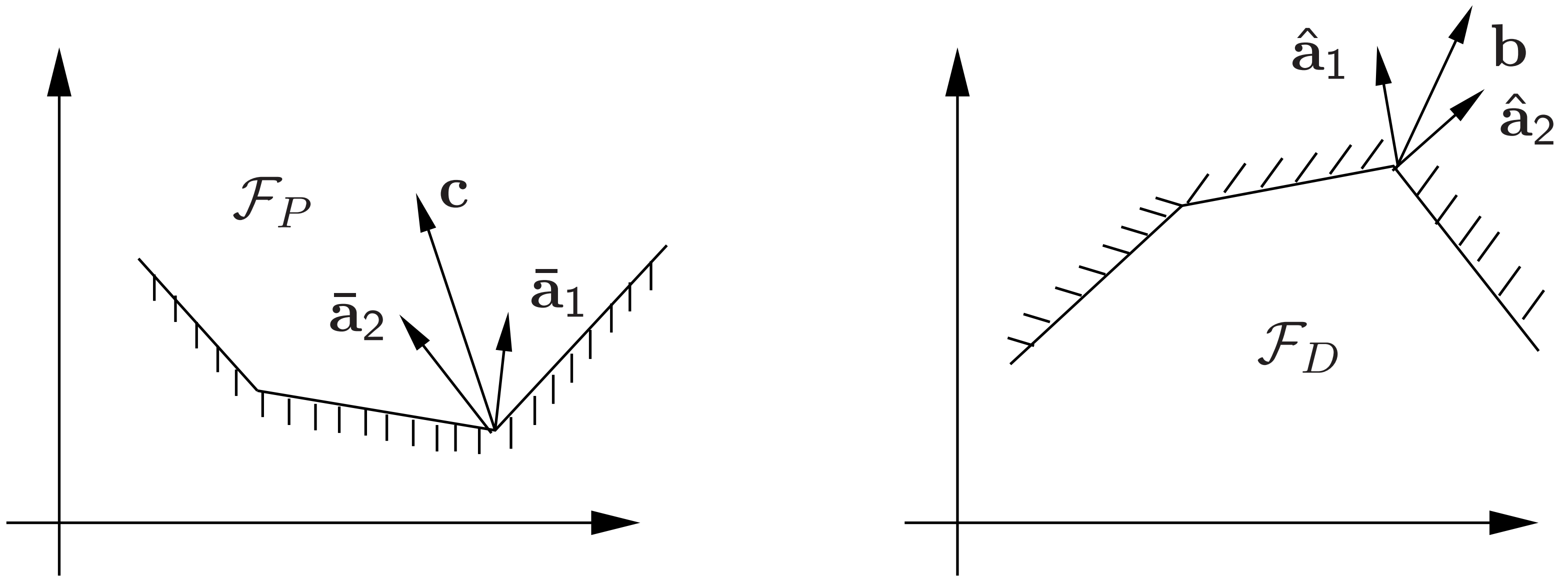
i.e., the gradient of the primal objective function is a linear combination (with coefficients > 0) of the normal vectors of the active constraints.

Let $\mathbf{e}_i = \begin{bmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \end{bmatrix}^\mathsf{T} \in \mathbf{R}^m$, where the 1 is in position i. Then

$$\mathbf{b} = \mathbf{A}\mathbf{x} - \mathbf{s} = \sum_{j:x_j>0} \hat{\mathbf{a}}_j x_j + \sum_{i:s_i>0} -\mathbf{e}_i s_i$$

i.e., the gradient of the dual objective function is a linear combination of the normal vectors of the active constraints.

Graphic illustration of the geometrical interpretation



- $\mathbf{c}$ can be written as a linear combination of $\bar{\mathbf{a}}_1$ and $\bar{\mathbf{a}}_2$
- $\mathbf{b}$ can be written as a linear combination of $\hat{\mathbf{a}}_1$ and $\hat{\mathbf{a}}_2$

Duality for general LP-problems

Primal

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}_1^T \mathbf{x}_1 + \mathbf{c}_2^T \mathbf{x}_2 \\ \text{s.t.} \quad & \mathbf{A}_{11}\mathbf{x}_1 + \mathbf{A}_{12}\mathbf{x}_2 \geq \mathbf{b}_1 \\ & \mathbf{A}_{21}\mathbf{x}_1 + \mathbf{A}_{22}\mathbf{x}_2 = \mathbf{b}_2 \\ & \mathbf{x}_1 \geq 0, \mathbf{x}_2 \text{ free} \end{aligned}$$

Dual

$$\begin{aligned} \text{maximize} \quad & \mathbf{b}_1^T \mathbf{y}_1 + \mathbf{b}_2^T \mathbf{y}_2 \\ \text{s.t.} \quad & \mathbf{A}_{11}^T\mathbf{y}_1 + \mathbf{A}_{21}^T\mathbf{y}_2 \leq \mathbf{c}_1 \\ & \mathbf{A}_{12}^T\mathbf{y}_1 + \mathbf{A}_{22}^T\mathbf{y}_2 = \mathbf{c}_2 \\ & \mathbf{y}_1 \geq 0, \ \mathbf{y}_2 \text{ free} \end{aligned}$$

Two methods for deriving the dual

- Reformulate the problem in canonical form, for which the dual is known.
- By using Lagrange-relaxation (more about this later in the course)

Duality for LP in standard form

Primal | Dual

$$(P)\quad \begin{aligned} \text{minimize}\quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.}\quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \qquad\qquad (D)\quad \begin{aligned} \text{maximize}\quad & \mathbf{b}^\mathsf{T}\mathbf{x} \\ \text{s.t.}\quad & \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c} \end{aligned}$$

$$\mathcal{F}_P = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b};\ \mathbf{x} \geq 0\} \qquad \mathcal{F}_D = \{\mathbf{y} \in \mathbf{R}^m : \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}\}$$

Weak duality: $\mathbf{x} \in \mathcal{F}_P$ and $\mathbf{y} \in \mathcal{F}_D$

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

Complementarity: $\mathbf{x} \in \mathcal{F}_P$ and $\mathbf{y} \in \mathcal{F}_D$ is optimal for (P) and (D) respectively, if and only if $(\mathbf{c} - \mathbf{A}^\mathsf{T}\mathbf{y})^\mathsf{T}\mathbf{x} = 0$, i.e. if $(c_i - \hat{\mathbf{a}}_i^\mathsf{T}\mathbf{y})x_i = 0$

A consequence of complementarity

Assume that we with the simplex method determined an optimal solution of the primal (P). This means that we have basic- and non-basic variables with indecies $\beta = (\beta_1, \ldots, \beta_m)$ and $\nu = (\nu_1, \ldots, \nu_l)$ so that

$$\begin{cases} \mathbf{x}_\beta = \mathbf{A}_\beta^{-1}\mathbf{b} \\ \mathbf{x}_\nu = \mathbf{0} \\ \mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta \\ \mathbf{r}_\nu = \mathbf{c}_\nu - \mathbf{A}_\nu^\mathsf{T}\mathbf{y} \geq 0 \end{cases} \Rightarrow \begin{cases} \mathbf{A}_\beta^\mathsf{T}\mathbf{y} = \mathbf{c}_\beta \\ \mathbf{A}_\nu^\mathsf{T}\mathbf{y} \leq \mathbf{c}_\nu \\ \mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{c}_\beta^\mathsf{T}\mathbf{x}_\beta = \mathbf{b}^\mathsf{T}\mathbf{y} \end{cases} \Rightarrow \begin{cases} \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c} \\ \mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{b}^\mathsf{T}\mathbf{y} \end{cases}$$

The last implication means that $\mathbf{y}$ is optimal for (D).

Conversely, assume that we have found an optimal solution $\mathbf{y}$ to the dual (D). We can construct an optimal solution $\mathbf{x}$ to the primal by using the complementarity constraints:

$$(c_i - \hat{\mathbf{a}}_i^{\mathsf{T}} \mathbf{y}) \cdot x_i = 0, \qquad i = 1, \ldots, n$$

If $c_i - \hat{\mathbf{a}}_i^{\mathsf{T}} \mathbf{y} > 0$ it holds that $x_i = 0$, i.e. $i \in \nu$ (non-basic variable). If there are $n - m$ such non-active dual constraints, then we have found a set of basic variable indices β for which the optimal x is given by

$$\mathbf{x}_\beta = \mathbf{A}_\beta^{-1} \mathbf{b} \text{ and } \mathbf{x}_\nu = 0$$

If it is easy to solve the dual, this is a good alternative. This is usually the case if $n >> m$.

Example

$$\begin{aligned} \text{minimize} \quad & x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 \\ \text{s.t.} \quad & 5x_1 + 4x_2 + 3x_3 + 2x_2 + x_1 = 1 \\ & x_k \geq 0, k = 1, \ldots, 5 \end{aligned}$$

The dual is

$$\begin{aligned} \text{maximize} \quad & y \\ \text{s.t.} \quad & 5y \leq 1,\ 4y \leq 2,\ 3y \leq 3,\ 2y \leq 4,\ y \leq 5 \end{aligned}$$

which has the optimal solution $\hat{y} = \frac{1}{5}$. From the complementarity constraint it is clear that $\hat{x}_2 = \hat{x}_3 = \hat{x}_4 = \hat{x}_5 = 0$ and for the primal constraint to be satisfied $\hat{x}_1 = \frac{1}{5}$. This is the optimal solution to the primal.

Reading instructions

- Chapter 6 (6.7) in the book.