Föreläsning: LP-Dualitet

1. Dualitet för LP på kanonisk form.
   - Dualitet, komplementaritet, och tolkningar.
2. Dualitet för allmäna LP problem
3. Dualitet för LP på standardform

Läsanvisning för OH-bilderna:

- Den geometriska tolkning av komplementaritetsvillkoren läses kursivt. Liknande tolkningar återkommer senare i kursen där det kommer att vara enklare att förstå.
- Den ekonomiska tolkningen av primal- och dualproblemet läses kursivt.

Dualitet för LP på kanonisk form

Primal	Dual
$$(P)\quad \begin{array}{rl} \text{minimera} & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{då} & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array}$$	$$(D)\quad \begin{array}{rl} \text{maximera} & \mathbf{b}^{\mathsf{T}}\mathbf{y} \\ \text{då} & \mathbf{A}^{\mathsf{T}}\mathbf{y} \leq \mathbf{c} \\ & \mathbf{y} \geq 0 \end{array}$$

Primalt tillåtet område

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

$\hat{\mathbf{x}} \in \mathcal{F}_P$ är optimal lösning om

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

Dualt tillåtet område

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

$\hat{\mathbf{y}} \in \mathcal{F}_D$ är optimal lösning om

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

Hur är primalen och dualen relaterade?

Betrakta följande primal och motsvarande dual

Primal

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

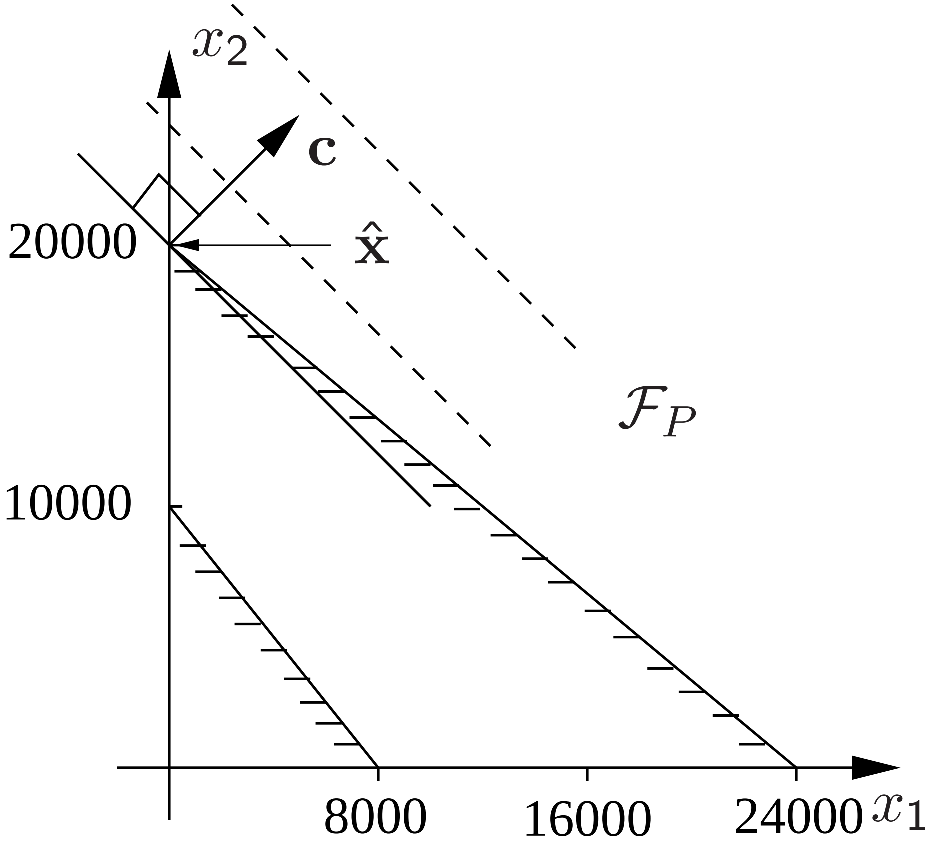
Dual

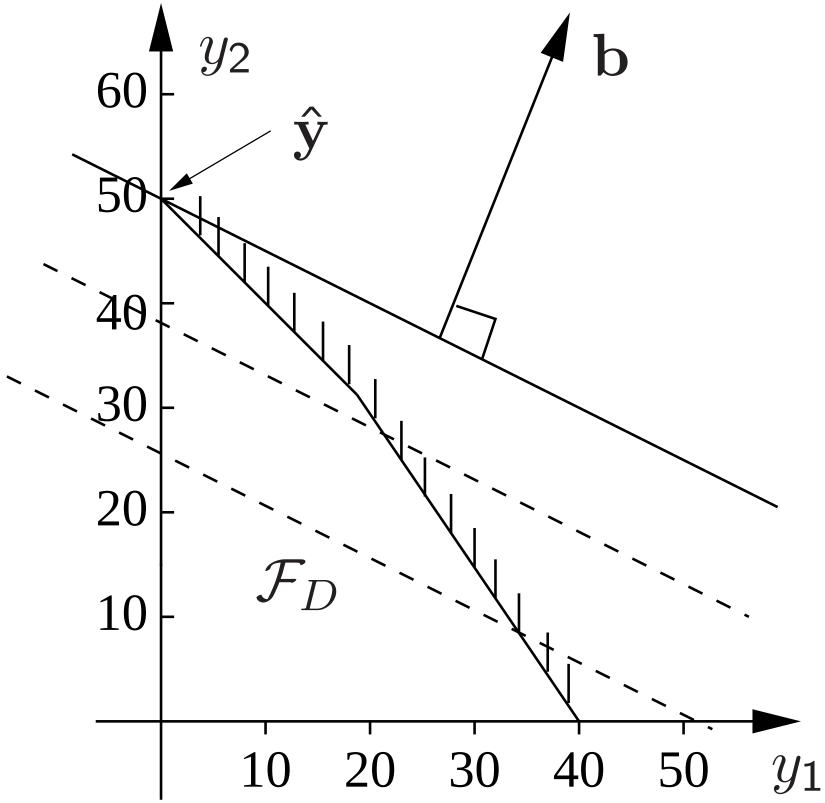
$$
(D)\quad \begin{aligned}
\text{maximera}\quad & 200y_1 + 400y_2 \\
\text{då}\quad & \tfrac{1}{40}y_1 + \tfrac{1}{60}y_2 \leq 1 \\
& \tfrac{1}{50}y_1 + \tfrac{1}{50}y_2 \leq 1 \\
& y_1, y_2 \geq 0
\end{aligned}
$$

Hur förhåller sig

- optimalvärdena?
- målfunktionernas värde? (för godtyckliga tillåtna punkter)
- optimallösningen och bivillkoren vid optimalitet?

Illustration av tillåtna området, isokostlinjer, samt optimal punkt





- Optimalvärdena är lika $\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = \mathbf{b}^\mathsf{T}\hat{\mathbf{y}} = 20000$

- För varje $\mathbf{x} \in \mathcal{F}_P$ och $\mathbf{y} \in \mathcal{F}_D$ gäller $\mathbf{c}^\mathsf{T}\mathbf{x} \geq \mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = \mathbf{b}^\mathsf{T}\hat{\mathbf{y}} \geq \mathbf{b}^\mathsf{T}\mathbf{y}$

Inför

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

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

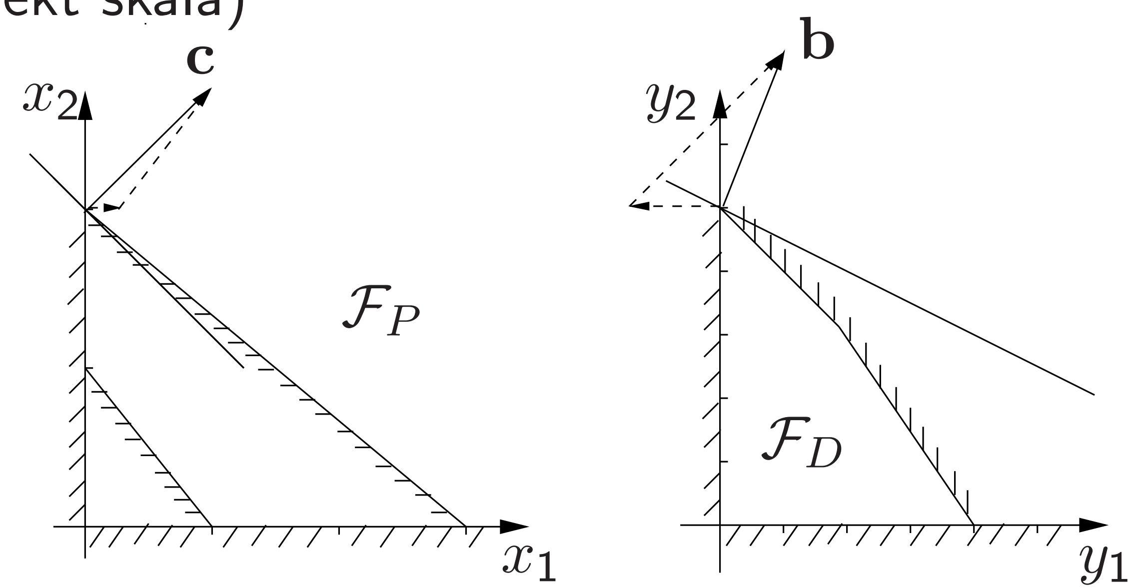
Optimallösningen uppfyller

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

Komplementaritetsvillkoren kan tolkas geometriskt. Från (R) och (S) ser vi att $\mathbf{c} = \hat{\mathbf{r}} + \mathbf{A}^T\hat{\mathbf{y}}$ och $\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{s}}$, dvs

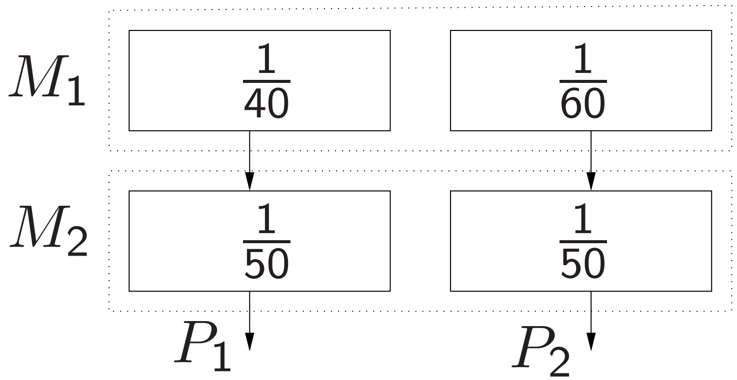
$$(R)\quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{6} + \begin{bmatrix} \frac{1}{60} \\ \frac{1}{50} \end{bmatrix} 50 \qquad (S)\quad \begin{bmatrix} 200 \\ 400 \end{bmatrix} = \begin{bmatrix} \frac{1}{50} \\ \frac{1}{50} \end{bmatrix} 20000 - \begin{bmatrix}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

följer att målfunktionernas gradienter ($\mathbf{c}$ respektive $\mathbf{b}$) kan skrivas som linjärkombinationer av de bindande bivillkoren. Detta illustreras i figuren (dock inte i korrekt skala)



Ekonomisk tolkning

- Vi betraktar produktion av två produkter P_1 och P_2
- Maskinerna M_1 och M_2 används vid produktion av såväl P_1 som P_2
- Produktionskapaciteten för maskinerna [dygn/enhet]:



- Priset för produkterna [kKr/enhet]:

produkt	pris
P_1	$b_1 = 200$
P_2	$b_2 = 400$

- I dualen bestäms produktionsvolymen så att profiten maximeras.

Antag att du har alternativet att hyra ut maskinerna M_1 och M_2. Vilket är det minsta tänkbara priset x_k kKr/dygn för att hyra ut M_k, $k = 1, 2$.

Minsta priset erhålls ur följande optimeringsproblem

$$
\begin{array}{rl}
\text{minimera} & x_1 + x_2 \\
\text{då} & \frac{1}{40}x_1 + \frac{1}{50}x_2 \geq 200 \\
& \frac{1}{60}x_1 + \frac{1}{50}x_2 \geq 400 \\
& x_1, x_2 \geq 0
\end{array}
$$

där bivillkoren kräver att du tjänar minst lika mycket på att hyra ut maskinerna jämfört med din inkomst från produktion av P_1 och P_2.

- Primalen och dualen kan ses som alternativ till varandra.
  - Om du har som affärsidé att hyra ut M_1 och M_2 så ger dualen information om minsta produktionsvolym som krävs för att det skall vara lönt att istället använda maskinerna för produktion av P_1 och P_2.
  - Om din affärsidé är produktion av P_1 och P_2 så ger primalen information om minsta dygnspris för att det skall löna sig att istället hyra ut maskinerna.
- Vad som kallas primal respektive dual är egentligen godtyckligt eftersom man kan visa att dualen av dualen är primalen.

Teoretiska resultat

Följande resultat i OK (Svanbergs gröna häfte) bekräftar våra iakttagelser från exemplet.

Proposition 6.1 (Svag dualitet) För varje $\mathbf{x} \in \mathcal{F}_P$ och varje $\mathbf{y} \in \mathcal{F}_D$ gäller att $\mathbf{c}^\mathsf{T}\mathbf{x} \geq \mathbf{b}^\mathsf{T}\mathbf{y}$.

En direkt följd av detta är att om $\hat{\mathbf{x}} \in \mathcal{F}_P$ och $\hat{\mathbf{y}} \in \mathcal{F}_D$ och $\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = \mathbf{b}^\mathsf{T}\hat{\mathbf{y}}$ så är $\hat{\mathbf{x}}$ och $\hat{\mathbf{y}}$ optimal till (P) respektive (D).

Theorem 6.3 (Stark dualitet) Vi har följande fyra alternativ för optimeringsproblemen (P) och (D)

(a) $\mathcal{F}_P \neq \emptyset$, $\mathcal{F}_d \neq \emptyset$. Då finns (minst) en optimal lösning $\hat{\mathbf{x}}$ till (P) och (minst) en optimal lösning $\hat{\mathbf{y}}$ till (D). Dessa uppfyller $\mathbf{c}^\mathsf{T}\hat{\mathbf{x}} = \mathbf{b}^\mathsf{T}\hat{\mathbf{y}}$.

(b) $\mathcal{F}_P \neq \emptyset$ men $\mathcal{F}_D = \emptyset$. Då är primalen nedåt obegränsad.

(c) $\mathcal{F}_D \neq \emptyset$ men $\mathcal{F}_P = \emptyset$. Då är dualen uppåt obegränsad.

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

- Beviset av (a) är icke-trivialt och använder sig typiskt av Farkas lemma.
- Fall (b) inses intuitivt eftersom det saknas nedre gräns för primalen om dualen inte har någon tillåten lösning. Motsvarande tolkning gäller för (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 (Komplementaritetssatsen) $\mathbf{x} \in \mathbf{R}^n$ är optimal till (P) och $\mathbf{y} \in \mathbf{R}^m$ är optimal till (D) om och endast om (omm)

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

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

Terminologi:

- Enkla bivillkor för primalen: $x_j \geq 0$, $j = 1, \dots, n$.
- Allmäna bivillkor för primalen: $s_i = \bar{\mathbf{a}}_i^\mathsf{T}\mathbf{x} - b_i \geq 0$, $i = 1, \dots, m$.
- Enkla bivillkor för dualen: $y_i \geq 0$, $i = 1, \dots, m$.
- Allmäna bivillkor för dualen: $\mathbf{r}_j = \mathbf{c}_j - \hat{\mathbf{a}}_j^\mathsf{T}\mathbf{y} \geq 0$, $j = 1, \dots, n$.

En konsekvens av komplementaritetssatsen är att

- Om $r_j = c_j - \hat{\mathbf{a}}_j^\mathsf{T}\mathbf{y} > 0$ så är $x_j = 0$,
  d.v.s primalens $j:e$ enkla bivillkor är bindande.

- Om $y_i > 0$ så är $s_i = \bar{\mathbf{a}}_i^\mathsf{T}\mathbf{x} - b_i = 0$,
  d.v.s. primalens $i:e$ allmäna bivillkor är bindande.

- Om $s_i = \bar{\mathbf{a}}_i^\mathsf{T}\mathbf{x} - b_i > 0$ så är $y_i = 0$,
  d.v.s. dualens $i:e$ enkla bivillkor är bindande.

- Om $x_j > 0$ så är $r_j = c_j - \hat{\mathbf{a}}_j^\mathsf{T}\mathbf{y} = 0$,
  d.v.s. dualens $j:e$ allmäna bivillkor är bindande.

Geometrisk tolkning av komplementaritetsvillkoren

Låt $\mathbf{e}_j = \begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{bmatrix}^{\mathsf{T}} \in \mathbf{R}^n$ med 1:an i position j.
Då kan vi skriva

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

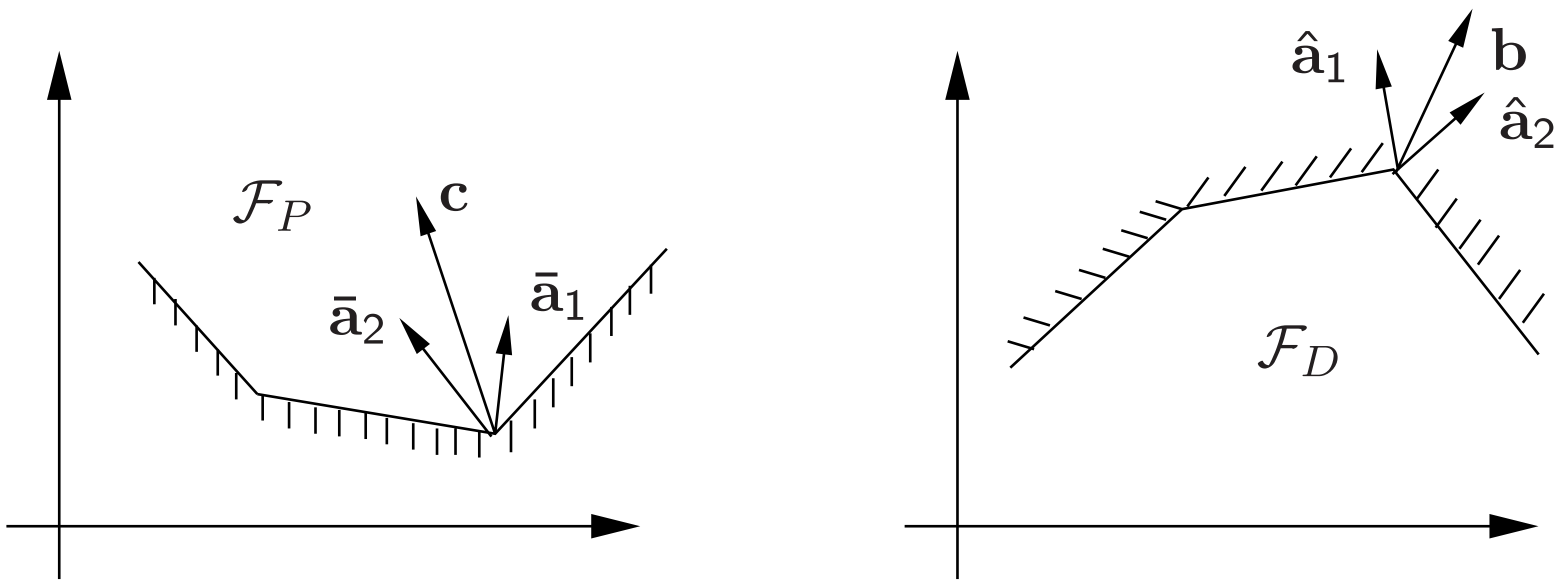
d.v.s. primala målfunktionens gradient är en linjärkombination (med positiva koefficienter) av de aktiva bivillkorens normalvektorer.

Låt $\mathbf{e}_i = \begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{bmatrix}^{\mathsf{T}} \in \mathbf{R}^m$ med 1:an i position i.
Då kan vi skriva

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

d.v.s. den duala målfunktionens gradient är en linjärkombination (med positiva koefficienter) av de aktiva bivillkorens normalvektorer.

Grafisk illustration av den geometriska tolkningen.



- $\mathbf{c}$ kan skrivas som en positiv linjärkombination av $\bar{\mathbf{a}}_1$ och $\bar{\mathbf{a}}_2$
- $\mathbf{b}$ kan skrivas som en positiv linjärkombination av $\hat{\mathbf{a}}_1$ och $\hat{\mathbf{a}}_2$

Dualitet för allmäna LP-problem

	Primal		Dual
minimera	$\mathbf{c}_1^T\mathbf{x}_1 + \mathbf{c}_2^T\mathbf{x}_2$	maximera	$\mathbf{b}_1^T\mathbf{y}_1 + \mathbf{b}_2^T\mathbf{y}_2$
då	$\mathbf{A}_{11}\mathbf{x}_1 + \mathbf{A}_{12}\mathbf{x}_2 \geq \mathbf{b}_1$	då	$\mathbf{A}_{11}^T\mathbf{y}_1 + \mathbf{A}_{21}^T\mathbf{y}_2 \leq \mathbf{c}_1$
	$\mathbf{A}_{21}\mathbf{x}_1 + \mathbf{A}_{22}\mathbf{x}_2 = \mathbf{b}_2$		$\mathbf{A}_{12}^T\mathbf{y}_1 + \mathbf{A}_{22}^T\mathbf{y}_2 = \mathbf{c}_2$
	$\mathbf{x}_1 \geq 0, \mathbf{x}_2$ fri		$\mathbf{y}_1 \geq 0,\ \mathbf{y}_2$ fri

Två metoder för härledning av dualen

- Överför på kanonisk form för vilken dualen är känd.
- Genom att använda Lagrangerelaxering

Dualitet för LP på standardform

	Primal	Dual
	$(P) \quad \begin{array}{rl} \text{minimera} & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{då} & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array}$	$(D) \quad \begin{array}{rl} \text{maximera} & \mathbf{b}^\mathsf{T}\mathbf{x} \\ \text{då} & \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c} \end{array}$
	$\mathcal{F}_P = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b};\ \mathbf{x} \geq 0\}$	$\mathcal{F}_D = \{\mathbf{y} \in \mathbf{R}^m : \mathbf{A}^\mathsf{T}\mathbf{y} \leq \mathbf{c}\}$

Svag dualitet: $\mathbf{x} \in \mathcal{F}_P$ och $\mathbf{y} \in \mathcal{F}_D$

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

Komplementaritet: $\mathbf{x} \in \mathcal{F}_P$ och $\mathbf{y} \in \mathcal{F}_D$ är optimal till (P) respektive (D) om och endast om $(\mathbf{c} - \mathbf{A}^\mathsf{T}\mathbf{y})^\mathsf{T}\mathbf{x} = 0$, d.v.s om $(c_i - \hat{\mathbf{a}}_i^\mathsf{T}\mathbf{y}) \cdot x_i = 0$

Konsekvens av komplementaritet

Antag att vi med hjälp av simplexmetoden hittat en optimal lösning till primalen (P). Detta betyder att vi har indexvektorer $\beta = (\beta_1, \ldots, \beta_m)$ och $\nu = (\nu_1, \ldots, \nu_l)$ så att

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

Den sista implikationen betyder att $\mathbf{y}$ är optimal till (D).

Antag omvänt att om vi har hittat en optimal lösning $\mathbf{y}$ till dualen (D). Vi kan konstruera en optimal lösning $\mathbf{x}$ till primalen genom att uttnyttja komplementaritetsvillkoret

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

Om $c_i - \hat{\mathbf{a}}_i^\mathsf{T}\mathbf{y} > 0$ så gäller att $x_i = 0$, d.v.s. $i \in \nu$ (icke-basvariabel). Om det finns $n - m$ sådana icke-bindande duala bivillkor så har vi funnit en basindexvektor β varvid optimala x ges av

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

Om dualen är billig att lösa så är detta ett bra alternativ, d.v.s. om $n >> m$.

Exempel

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

Dualen blir

$$\begin{aligned} \text{maximera} \quad & y \\ \text{då} \quad & 5y \leq 1, \ 4y \leq 2, \ 3y \leq 3, \ 2y \leq 4, \ y \leq 5 \end{aligned}$$

vilken har optimala lösningen $\hat{y} = \frac{1}{5}$. Från komplementaritetsvillkoret ser man att $\hat{x}_2 = \hat{x}_3 = \hat{x}_4 = \hat{x}_5 = 0$ och för att primala bivillkoret skall vara uppfyllt behöver vi då $\hat{x}_1 = \frac{1}{5}$. Detta är optimallösningen till primalen.

Läsanvisningar

I Optimeringskompendiet kapitel 6.

Gamla Materialet:

- Gröna häftet sidan 23 till 28.