Lecture 10
Duality (part 2)

- duality in algorithms
- sensitivity analysis via duality
- duality for general LPs
- examples
- mechanics interpretation
- circuits interpretation
- two-person zero-sum games

Duality in algorithms

many algorithms produce at iteration k

- a primal feasible $x^{(k)}$
- and a dual feasible $z^{(k)}$

with $c^T x^{(k)} + b^T z^{(k)} \to 0$ as $k \to \infty$

hence at iteration k we **know** $p^\star \in \left[-b^T z^{(k)}, c^T x^{(k)}\right]$

- useful for stopping criteria
- algorithms that use dual solution are often more efficient

Nonheuristic stopping criteria

- (absolute error) $c^T x^{(k)} - p^\star$ is less than ϵ if

$$c^T x^{(k)}) + b^T z^{(k)} < \epsilon$$

- (relative error) $(c^T x^{(k)} - p^\star)/|p^\star|$ is less than ϵ if

$$-b^T z^{(k)} > 0 \quad \& \quad \frac{c^T x^{(k)} + b^T z^{(k)}}{-b^T z^{(k)}} \leq \epsilon$$

or

$$c^T x^{(k)} < 0 \quad \& \quad \frac{c^T x^{(k)}) + b^T z^{(k)}}{-c^T x^{(k)}} \leq \epsilon$$

- target value ℓ is achievable ($p^\star \leq \ell$) if

$$c^T x^{(k)}) \leq \ell$$

- target value ℓ is unachievable ($p^\star > \ell$) if

$$-b^T z^{(k)} > \ell$$

Sensitivity analysis via duality

perturbed problem:

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & Ax \le b + \epsilon d\end{array}$$

$A \in \mathbf{R}^{m\times n}$; $d \in \mathbf{R}^m$ given; optimal value $p^\star(\epsilon)$

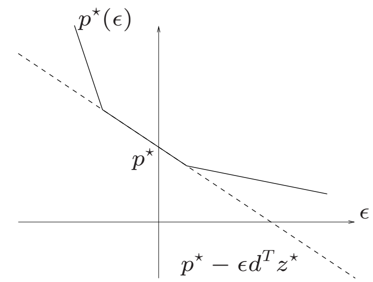
global sensitivity result: if $z^\star$ is (any) dual optimal solution for the unperturbed problem, then for all ϵ

$$p^\star(\epsilon) \ge p^\star - \epsilon d^T z^\star$$

proof. $z^\star$ is dual feasible for all ϵ; by weak duality,

$$p^\star(\epsilon) \ge -(b+\epsilon d)^T z^\star = p^\star - \epsilon d^T z^\star$$

interpretation



- $d^T z^\star > 0$: $\epsilon < 0$ increases $p^\star$
- $d^T z^\star > 0$ and large: $\epsilon < 0$ greatly increases $p^\star$
- $d^T z^\star > 0$ and small: $\epsilon > 0$ does not decrease $p^\star$ too much
- $d^T z^\star < 0$: $\epsilon > 0$ increases $p^\star$
- $d^T z^\star < 0$ and large: $\epsilon > 0$ greatly increases $p^\star$
- $d^T z^\star < 0$ and small: $\epsilon > 0$ does not decrease $p^\star$ too much

Local sensitivity analysis

assumption: there is a nondegenerate optimal vertex $x^\star$, *i.e.*,

- $x^\star$ is an optimal *vertex*: $\textbf{rank}\,\bar{A} = n$, where

$$\bar{A} = \left[\begin{array}{cccc} a_{i_1} & a_{i_2} & \cdots & a_{i_K}\end{array}\right]^T$$

  and $I = \{i_1, \ldots, i_K\}$ is the set of active constraints at $x^\star$
- $x^\star$ is nondegenerate: $\bar{A} \in \mathbf{R}^{n\times n}$

w.l.o.g. we assume $I = \{1, 2, \ldots, n\}$

consequence: dual optimal $z^\star$ is unique

proof: by complementary slackness, $z_i^\star = 0$ for $i > n$

by dual feasibility,

$$\sum_{i=1,\ldots,n} a_i z_i^\star = \bar{A}^T \left[\begin{array}{c} z_1^\star \\ \vdots \\ z_n^\star\end{array}\right] = -c \quad\Longrightarrow\quad z^\star = \left[\begin{array}{c} -\bar{A}^{-T}c \\ 0\end{array}\right]$$

optimal solution of the perturbed problem (for small ϵ):

$$x^\star(\epsilon) = x^\star + \epsilon \bar{A}^{-1}\bar{d} \quad (\text{with } \bar{d} = (d_1, \ldots, d_n))$$

- $x^\star(\epsilon)$ is feasible for small ϵ:

$$a_i^T x^\star(\epsilon) = b_i + \epsilon d_i, \quad i = 1, \ldots, n, \qquad a_i^T x^\star(\epsilon) < b_i + \epsilon d_i, \quad i = n+1, \ldots, m$$

- $z^\star$ is dual feasible and satisfies complementary slackness:

$$(b + \epsilon d - Ax^\star(\epsilon))^T z^\star = 0$$

optimal value of perturbed problem (for small ϵ):

$$p^\star(\epsilon) = c^T x^\star(\epsilon) = p^\star + \epsilon c^T \bar{A}^{-1}\bar{d} = p^\star - \epsilon d^T z^\star$$

- $z_i^\star$ is sensitivity of cost w.r.t. righthand side of ith constraint
- $z_i^\star$ is called *marginal cost* or *shadow price* associated with ith constraint

Dual of a general LP

method 1: express as LP in inequality form and take its dual

example: standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \begin{bmatrix} -I \\ A \\ -A \end{bmatrix} x \leq \begin{bmatrix} 0 \\ b \\ -b \end{bmatrix} \end{array}$$

dual:

$$\begin{array}{ll} \text{maximize} & -b^T(v-w) \\ \text{subject to} & -u + A^T(v-w) + c = 0 \\ & u \geq 0, \quad v \geq 0, \quad w \geq 0 \end{array}$$

method 2: apply Lagrange duality (this lecture)

Lagrangian

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array}$$

Lagrangian $L : \mathbf{R}^{n+m} \to \mathbf{R}$

$$L(x, \lambda) = c^T x + \sum_{i=1}^m \lambda_i (a_i^T x - b_i)$$

- λ_i are called *Lagrange multipliers*
- objective is *augmented* with weighted sum of constraint functions

lower bound property: if $Ax \leq b$ and $\lambda \geq 0$, then

$$c^T x \geq L(x, \lambda) \geq \min_{\tilde{x}} L(\tilde{x}, \lambda)$$

hence, $p^\star \geq \min_{\tilde{x}} L(\tilde{x}, \lambda)$ for $\lambda \geq 0$

Lagrange dual problem

Lagrange dual function $g : \mathbf{R}^m \to \mathbf{R} \cup \{-\infty\}$

$$\begin{aligned} g(\lambda) = \min_x L(x, \lambda) &= \min_x \left(-b^T\lambda + (A^T\lambda + c)^T x\right) \\ &= \begin{cases} -b^T\lambda & \text{if } A^T\lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

(Lagrange) dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda) \\ \text{subject to} & \lambda \geq 0 \end{array}$$

yields the dual LP:

$$\begin{array}{ll} \text{maximize} & -b^T\lambda \\ \text{subject to} & A^T\lambda + c = 0, \quad \lambda \geq 0 \end{array}$$

finds best lower bound $g(\lambda)$

Lagrangian of a general LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \\ & g_i^T x = h_i, \quad i = 1, \ldots, p \end{array}$$

define **Lagrangian** $L : \mathbf{R}^{n+m+p} \to \mathbf{R}$ as

$$L(x, \lambda, \nu) = c^T x + \sum_{i=1}^m \lambda_i (a_i^T x - b_i) + \sum_{i=1}^p \nu_i (g_i^T x - h_i)$$

lower bound property: if x is feasible and $\lambda \geq 0$,

$$c^T x \geq L(x, \lambda, \nu) \geq \min_{\tilde{x}} L(\tilde{x}, \lambda, \nu)$$

hence, $p^\star \geq \min_x L(x, \lambda, \nu)$ if $\lambda \geq 0$

multipliers associated with equality constraints can have either sign

Lagrange dual function:

$$
\begin{aligned}
g(\lambda, \nu) = \min_x L(x, \lambda, \nu) &= \min_x \left(c^T x + \lambda^T (Ax - b) + \nu^T (Gx - h)\right) \\
&= \min_x \left(-b^T\lambda - h^T\nu + x^T(c + A^T\lambda + G^T\nu)\right) \\
&= \begin{cases} -b^T\lambda - h^T\nu & \text{if } A^T\lambda + G^T\nu + c = 0 \\ -\infty & \text{otherwise} \end{cases}
\end{aligned}
$$

Lagrange dual problem:

$$
\begin{array}{ll}
\text{maximize} & -b^T\lambda - h^T\nu \\
\text{subject to} & A^T\lambda + G^T\nu + c = 0 \\
& \lambda \geq 0
\end{array}
$$

variables λ, ν; optimal value $d^\star$

- an LP (in general form)
- weak duality $p^\star \geq d^\star$
- strong duality holds: $p^\star = d^\star$ (except when both problems are infeasible)

Example: standard form LP

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax = b, \quad x \geq 0
\end{array}
$$

Lagrangian: $L(x, \nu, \lambda) = c^T x + \nu^T(Ax - b) - \lambda^T x$

dual function

$$
g(\lambda, \nu) = \min_x L(x, \nu, \lambda) = \begin{cases} -b^T\nu & \text{if } A^T\nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}
$$

dual problem

$$
\begin{array}{ll}
\text{maximize} & -b^T\nu \\
\text{subject to} & A^T\nu - \lambda + c = 0, \quad \lambda \geq 0
\end{array}
$$

equivalent to dual on page 10–9

$$
\begin{array}{ll}
\text{maximize} & b^T z \\
\text{subject to} & A^T z \leq c
\end{array}
$$

Price or tax interpretation

- x: describes how an enterprise operates; $c^T x$: cost of operating at x
- $a_i^T x \leq b_i$: limits on resources, regulatory limits

optimal operating point is solution of

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m
\end{array}
$$

optimal cost: $p^\star$

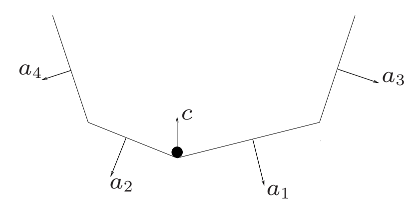
scenario 2: constraint violations can be bought or sold at unit cost $\lambda_i \geq 0$

$$
\text{minimize} \quad c^T x + \textstyle\sum_{i=1}^m \lambda_i (a_i^T x - b_i)
$$

optimal cost: $g(\lambda)$

interpretation of strong duality: there exist prices $\lambda_i^\star$ s.t. $g(\lambda^\star) = p^\star$, $i.e.$, there is no advantage in selling/buying constraint violations

Mechanics interpretation



- mass subject to gravity, can move freely between walls described by $a_i^T x = b_i$
- equilibrium position minimizes potential energy, $i.e.$, solves

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject} & a_i^T x \leq b_i, \quad i = 1, \ldots, m
\end{array}
$$

optimality conditions:

1. $a_i^T x^\star \leq b_i$, $i = 1.\dots, m$

2. $\sum_{i=1}^m z_i^\star a_i + c = 0$, $z^\star \geq 0$

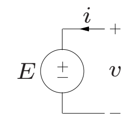
3. $z_i^\star(b_i - a_i^T x^\star) = 0$, $i = 1, \dots, m$

interpretation: $-z_i a_i$ is contact force with wall i; nonzero only if the ball touches the ith wall

Circuits interpretation

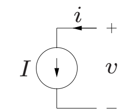
circuit components:

ideal voltage source



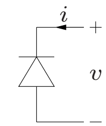
$v = E$

ideal current source



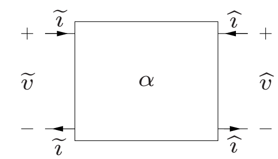
$i = I$

ideal diode



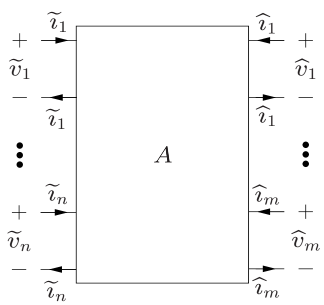
$v \geq 0, \quad i \leq 0$
$vi = 0$

ideal transformer



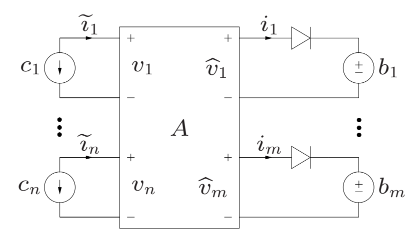
$$\begin{aligned} \widehat{v} &= \alpha \widetilde{v} \\ \widetilde{\imath} &= -\alpha \widehat{\imath} \end{aligned}$$

ideal multiterminal transformer ($A \in \mathbf{R}^{m \times n}$)



$$\widehat{v} = A\widetilde{v}$$
$$\widetilde{\imath} = -A^T \widehat{\imath}$$

example



circuit equations:

$$\widehat{v} = Av \leq b, \qquad i \geq 0, \qquad \widetilde{\imath} + c = A^T i + c = 0$$

$$i_k(b_k - a_k^T v) = 0, \quad k = 1, \dots, m$$

$i.e.$, optimality conditions for LP

minimize	$c^T v$	maximize	$-b^T i$
subject to	$Av \leq b$	subject to	$A^T i + c = 0$
			$i \geq 0$

interpretation: two 'potential functions'

- *content* (a function of the voltages)
- *co-content* (a function of the currents)

contribution of each component (notation of page 10–18)

- content of current source is Iv
  co-content is 0 if $i = I$, $-\infty$ otherwise
- content of voltage source is 0 if $v = E$, ∞ otherwise
  co-content is $-Ei$
- content of diode is 0 if $v \geq 0$, $+\infty$ otherwise
  co-content is 0 if $i \leq 0$ and $-\infty$ otherwise
- content of transformer is 0 if $\widehat{v} = A\widetilde{v}$, ∞ otherwise
  co-content is 0 if $\widetilde{i} = -A^T\widehat{\imath}$, $-\infty$ otherwise

primal problem: voltages minimize total content

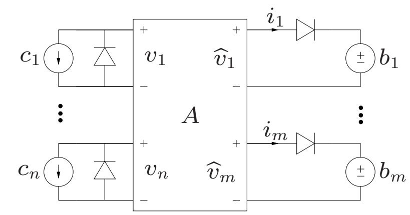
dual problem: currents maximize total co-content

example

primal problem

$$\begin{array}{ll} \text{minimize} & c^T v \\ \text{subject to} & Av \leq b \\ & v \geq 0 \end{array}$$

circuit equivalent:



dual problem:

$$\begin{array}{ll} \text{maximize} & -b^T i \\ \text{subject to} & A^T i + c \geq 0 \\ & i \geq 0 \end{array}$$

Two-person zero-sum games (matrix games)

described by a **payoff matrix**

$$A \in \mathbf{R}^{m \times n}$$

- player 1 chooses a number in $\{1, \ldots, m\}$ (corresponding to m possible actions or strategies)
- player 2 chooses a number in $\{1, \ldots, n\}$
- players make their choice simultaneously and independently
- if P1's choice is i and P2's choice is j, then P1 pays a_{ij} to P2 (negative a_{ij} means P2 pays $-a_{ij}$ to P1)

Mixed (randomized) strategies

players make random choices according to some probability distribution

- player 1 chooses randomly according to distribution $x \in \mathbf{R}^m$:
$$\mathbf{1}^T x = 1, \qquad x \geq 0$$
  (x_i is probability of choosing i)
- player 2 chooses randomly (and independently from 1) according to distribution $y \in \mathbf{R}^n$:
$$\mathbf{1}^T y = 1, \qquad y \geq 0$$
  (y_j is probability of choosing j)

expected payoff from player 1 to 2, if they use mixed stragies x and y:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j a_{ij} = x^T A y$$

Optimal mixed strategies

optimal strategy for player 1:

$$\begin{array}{ll} \text{minimize}_x & \max_{\mathbf{1}^T y=1,\, y\geq 0} x^T A y \\ \text{subject to} & \mathbf{1}^T x = 1, \quad x \geq 0 \end{array}$$

note:

$$\max_{\mathbf{1}^T y=1,\, y\geq 0} x^T A y = \max_{j=1,\ldots,n} (A^T x)_j$$

optimal strategy $x^\star$ can be computed by solving an LP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & A^T x \leq t\mathbf{1} \\ & \mathbf{1}^T x = 1, \quad x \geq 0 \end{array} \qquad (1)$$

(variables x, t)

optimal strategy for player 2:

$$\begin{array}{ll} \text{maximize}_y & \min_{\mathbf{1}^T x=1,\, x\geq 0} x^T A y \\ \text{subject to} & \mathbf{1}^T y = 1, \quad y \geq 0 \end{array}$$

note:

$$\min_{\mathbf{1}^T x=1,\, x\geq 0} x^T A y = \min_{i=1,\ldots,m} (Ay)_i$$

optimal strategy $y^\star$ can be computed by solving an LP

$$\begin{array}{ll} \text{maximize} & w \\ \text{subject to} & Ay \geq w\mathbf{1} \\ & \mathbf{1}^T y = 1, \quad y \geq 0 \end{array} \qquad (2)$$

(variables y, w)

The minimax theorem

for all mixed strategies x, y,

$$x^{\star T} A y \leq x^{\star T} A y^\star \leq x^T A y^\star$$

proof: the LPs (1) and (2) are duals, so they have the same optimal value

example

$$A = \begin{bmatrix} 4 & 2 & 0 & -3 \\ -2 & -4 & -3 & 3 \\ -2 & -3 & 4 & 1 \end{bmatrix}$$

optimal strategies

$$x^\star = (0.37, 0.33, 0.3), \qquad y^\star = (0.4, 0, 0.13, 0.47)$$

expected payoff: $x^{\star T} A y^\star = 0.2$