Elements of Mathematical Analysis and Conic Duality

Yinyu Ye

Department of Management Science and Engineering

Stanford University

Stanford, CA 94305, U.S.A.

http://www.stanford.edu/˜yyye

Appendix B, Chapters 3.1-2, 6.1-6.4

Carathéodory's theorem

The following theorem states that a polyhedral cone can be generated by a set of basic directional vectors.

Theorem 1 *Given matrix* $A \in R^{m \times n}$, *let convex polyhedral cone* $C = \{A\mathbf{x} : \mathbf{x} \geq \mathbf{0}\}$. *For any* $\mathbf{b} \in C$,

$$\mathbf{b} = \sum_{i=1}^{d} \mathbf{a}_{j_i} x_{j_i},\ x_{j_i} \geq 0, \forall i$$

for some linearly independent vectors $\mathbf{a}_{j_1}$,...,$\mathbf{a}_{j_d}$ *chosen from* $\mathbf{a}_1$,...,$\mathbf{a}_n$.

There is a construct proof of the theorem (page 26 of the text).

Basic and Basic Feasible Solution I

Now consider the feasible set $\{\mathbf{x} : \; A\mathbf{x} = \mathbf{b}, \; \mathbf{x} \geq \mathbf{0}\}$ for given data $A \in R^{m \times n}$ and $\mathbf{b} \in R^m$. Select m linearly independent columns, denoted by the variable index set B, from A. Solve $A_B \mathbf{x}_B = \mathbf{b}$ for the m-dimension vector $\mathbf{x}_B$, and set the remaining variables, $\mathbf{x}_N$, to zero. Then, we obtain a solution $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$, that is called a basic solution to with respect to the basis A_B. If a basic solution $\mathbf{x}_B \geq \mathbf{0}$, then $\mathbf{x}$ is called a basic feasible solution, or BFS.

An equivalent statement of Carathéodory's theorem is:

Theorem 2 *If there is a feasible solution* $\mathbf{x}$ *to* $\{\mathbf{x} : \; A\mathbf{x} = \mathbf{b}, \; \mathbf{x} \geq \mathbf{0}\}$, *then there is a basic feasible solution to the system (page 26 of the text), and it is an extreme or corner point of the feasible set and vice versa.*

Corollary 1 *The set* $\{\mathbf{x} : \; A\mathbf{x} = \mathbf{b}, \; \mathbf{x} \geq \mathbf{0}\}$ *is a polyhedral set.*

Hyper-Planes

The most important type of convex set is hyperplane, also called *linear variety or affine set*: if for any two points are in H then their affine combination is also in H.

Hyperplanes dominate the entire theory of optimization. Let $\mathbf{a}$ be a nonzero n-dimensional (slope) vector, and let b be a real (intercept) number. The set

$$H = \{\mathbf{x} \in \mathcal{R}^n : \mathbf{a} \bullet \mathbf{x} = b\}$$

is a hyperplane in $\mathcal{R}^n$. Relating to hyperplane, upper and lower closed half spaces are given by

$$H_+ = \{\mathbf{x} : \mathbf{a} \bullet \mathbf{x} \geq b\}$$

$$H_- = \{\mathbf{x} : \mathbf{a} \bullet \mathbf{x} \leq b\}.$$

Separating and supporting hyperplane theorem

The most important theorem about the convex set is the following separating hyperplane theorem (page 510 of the text).

Theorem 3 *(Separating hyperplane theorem) Let C be a closed convex set in $\mathcal{R}^m$ and let $\mathbf{b}$ be a point exterior to C. Then there is a vector $\mathbf{y} \in \mathcal{R}^m$ such that*

$$\mathbf{b} \bullet \mathbf{y} > \sup_{\mathbf{x} \in C} \mathbf{x} \bullet \mathbf{y}.$$

Theorem 4 *(Supporting hyperplane theorem) Let C be a closed convex set and let $\mathbf{b}$ be a point on the boundary of C. Then there is a vector $\mathbf{y} \in \mathcal{R}^m$ such that*

$$\mathbf{b} \bullet \mathbf{y} = \sup_{\mathbf{x} \in C} \mathbf{x} \bullet \mathbf{y}.$$

Let C be a unit circle centered at point $(1; 1)$. That is, $C = \{x \in \mathcal{R}^2 : \ (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1\}$. If $\mathbf{b} = (2; 0)$, $\mathbf{y} = (1; -1)$ is a separating hyperplane vector. If $\mathbf{b} = (0; -1)$, $\mathbf{y} = (0; -1)$ is a separating hyperplane vector. It is worth noting that these separating hyperplanes are not unique.

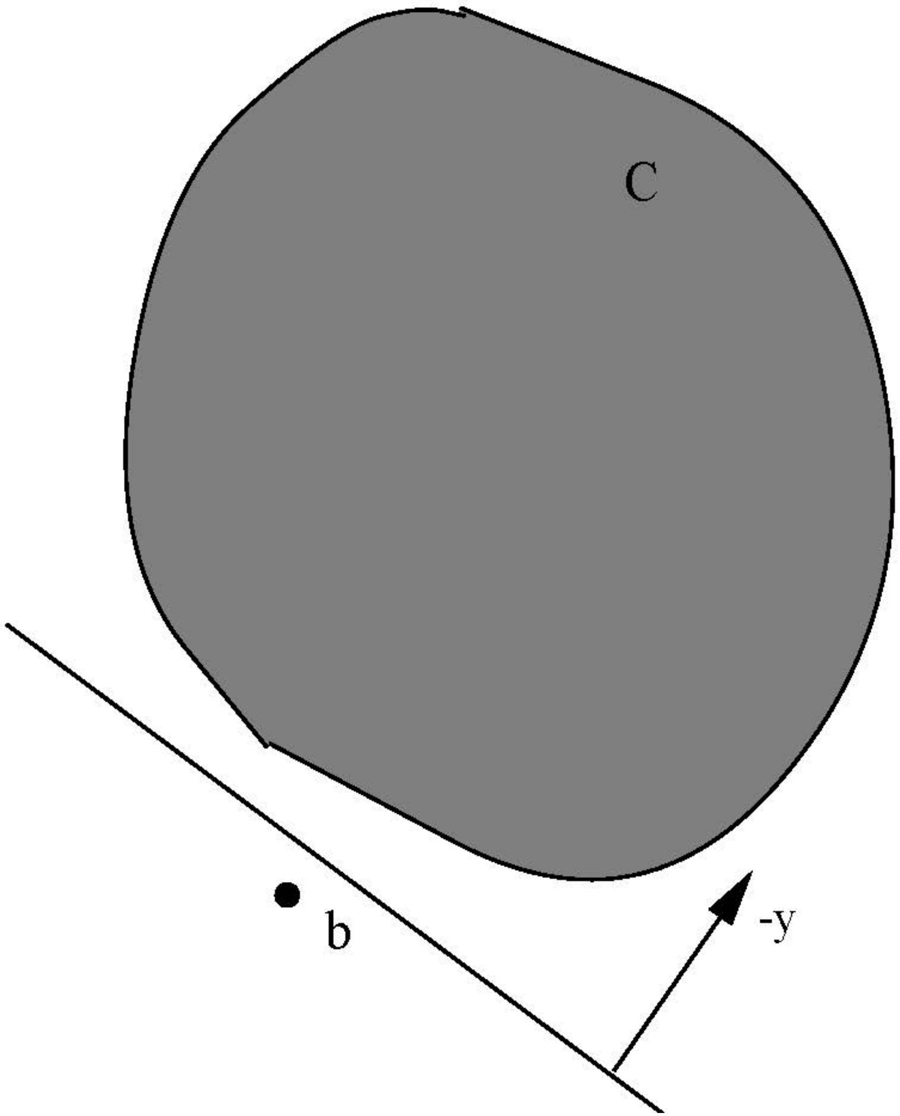


Figure 1: Illustration of the separating hyperplane theorem; an exterior point $\mathbf{b}$ is separated by a hyperplane from a convex set C.

Farkas' Lemma

The following results are Farkas' lemma and its variants.

Theorem 5 *Let* $A \in \mathcal{R}^{m \times n}$ *and* $\mathbf{b} \in \mathcal{R}^m$. *Then, the system* $\{\mathbf{x} : \; A\mathbf{x} = \mathbf{b}, \; \mathbf{x} \geq \mathbf{0}\}$ *has a feasible solution* $\mathbf{x}$ *if and only if that its alternative system* $-A^T\mathbf{y} \geq \mathbf{0}$ *and* $\mathbf{b}^T\mathbf{y} > 0$ *has no feasible solution* $\mathbf{y}$*.*

Geometrically, Farkas' lemma means that if a vector $\mathbf{b} \in \mathcal{R}^m$ does not belong to the convex cone generated by $\mathbf{a}_{.1}, ..., \mathbf{a}_{.n}$, then there is a hyperplane separating $\mathbf{b}$ from cone$(\mathbf{a}_{.1}, ..., \mathbf{a}_{.n})$.

Example Let $A = (1, 1)$ and $b = -1$. Then, there is $y = -1$ such that $-A^Ty \geq 0$ and $by > 0$..

Proof

Let $\{\mathbf{x}: \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$ have a feasible solution, say $\bar{\mathbf{x}}$. Then, $\{\mathbf{y}: \ A^T\mathbf{y} \leq \mathbf{0}, \ \mathbf{b}^T\mathbf{y} > 0\}$ is infeasible, since otherwise,

$$0 < \mathbf{b}^T\mathbf{y} = (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^T(A^T\mathbf{y}) \leq 0$$

from $\mathbf{x} \geq \mathbf{0}$ and $A^T\mathbf{y} \leq \mathbf{0}$.

Now let $\{\mathbf{x}: \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$ have no feasible solution, or $\mathbf{b} \notin C := \{A\mathbf{x}: \ \mathbf{x} \geq \mathbf{0}\}$. We now prove that its alternative system has a solution. We first prove

Lemma 1 $C = \{A\mathbf{x}: \ \mathbf{x} \geq \mathbf{0}\}$ *is a closed convex set.*

That is, any convergent sequence $\mathbf{b}^k \in C, \ k = 1.2....$ has its limit point $\bar{\mathbf{b}}$ also in C. Let $\mathbf{b}^k = A\mathbf{x}^k, \ \mathbf{x}^k \geq \mathbf{0}$. Then by Carathéodory's theorem, we must have $\mathbf{b}^k = A_{B^k}\mathbf{x}_{B^k}, \ \mathbf{x}_{B^k} \geq \mathbf{0}$ where A_{B^k} is a basis of A. Therefore, $\mathbf{x}_{B^k}$, together with zero values for the nonbasic variables, is **bounded** for all k, so that it has sub-sequence, say indexed by $l = 1, ...$, where $\mathbf{x}^l = \mathbf{x}_{B^l}$ has a limit point $\bar{\mathbf{x}}$ and $\bar{\mathbf{x}} \geq \mathbf{0}$. Consider this very sub-sequence $\mathbf{b}^l = A\mathbf{x}^l$ we must also have $\mathbf{b}^l \rightarrow \bar{\mathbf{b}}$. Then from

$$\|\bar{\mathbf{b}} - A\bar{\mathbf{x}}\| = \|\bar{\mathbf{b}} - \mathbf{b}^l + A\mathbf{x}^l - A\bar{\mathbf{x}}\| \leq \|\bar{\mathbf{b}} - \mathbf{b}^l\| + \|A\mathbf{x}^l - A\bar{\mathbf{x}}\| \leq \|\bar{\mathbf{b}} - \mathbf{b}^l\| + \|A\|\|\mathbf{x}^l - \bar{\mathbf{x}}\|$$

we must have $\bar{\mathbf{b}} = A\bar{\mathbf{x}}$, that is, $\bar{\mathbf{b}} \in C$; since otherwise the right-hand side of the above inequality is strictly greater than zero which is a contradiction.

Now since C is a closed convex set, by the separating hyperplane theorem, there is $\mathbf{y}$ such that

$$\mathbf{y} \bullet \mathbf{b} > \sup_{\mathbf{c} \in C} \mathbf{y} \bullet \mathbf{c}$$

or

$$\mathbf{y} \bullet \mathbf{b} > \sup_{\mathbf{x} \geq \mathbf{0}} \mathbf{y} \bullet (A\mathbf{x}) = \sup_{\mathbf{x} \geq \mathbf{0}} A^T\mathbf{y} \bullet \mathbf{x}. \tag{1}$$

From $\mathbf{0} \in C$ we have $\mathbf{y} \bullet \mathbf{b} > 0$.

Furthermore, $A^T\mathbf{y} \leq \mathbf{0}$. Since otherwise, say $(A^T\mathbf{y})_1 > 0$, one can have a vector $\bar{\mathbf{x}} \geq \mathbf{0}$ such that $\bar{x}_1 = \alpha > 0, \bar{x}_2 = ... = \bar{x}_n = 0$, from which

$$\sup_{\mathbf{x} \geq \mathbf{0}} A^T\mathbf{y} \bullet \mathbf{x} \geq A^T\mathbf{y} \bullet \bar{\mathbf{x}} = (A^T\mathbf{y})_1 \cdot \alpha$$

and it tends to ∞ as $\alpha \to \infty$. This is a contradiction because $\sup_{\mathbf{x} \geq \mathbf{0}} A^T\mathbf{y} \bullet \mathbf{x}$ is bounded from above by (1).

Farkas' Lemma Variant

Theorem 6 *Let $A \in \mathcal{R}^{m\times n}$ and $\mathbf{c} \in \mathcal{R}^n$. Then, the system $\{\mathbf{y} : \ \mathbf{c} - A^T\mathbf{y} \geq \mathbf{0}\}$ has a solution $\mathbf{y}$ if and only if that $A\mathbf{x} = \mathbf{0}$, $\mathbf{x} \geq \mathbf{0}$, and $\mathbf{c}^T\mathbf{x} < 0$ has no feasible solution $\mathbf{x}$.*

Example Let $A = (1; -1)$ and $\mathbf{c} = (1; -2)$. Then, there is $\mathbf{x} = (1; 1) \geq \mathbf{0}$ such that $A\mathbf{x} = 0$ and $\mathbf{c}^T\mathbf{x} < 0$.

Alternative System Pair I

$$A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}.$$

$$-A^T\mathbf{y} \geq \mathbf{0}, \quad \mathbf{b}^T\mathbf{y} = 1(> 0)$$

A vector $\mathbf{y}$, with $A^T\mathbf{y} \leq \mathbf{0}$ and $\mathbf{b}^T\mathbf{y} = 1$, is called an infeasibility certificate for the system $\{\mathbf{x} : \; A\mathbf{x} = \mathbf{b}, \; \mathbf{x} \geq \mathbf{0}\}$.

Alternative System Pair II

$$A\mathbf{x} = \mathbf{0}, \ \mathbf{x} \geq \mathbf{0}, \ \mathbf{c}^T\mathbf{x} = -1 (< 0).$$

$$\mathbf{c} - A^T\mathbf{y} \geq \mathbf{0}$$

A vector $\mathbf{x}$, with $A\mathbf{x} = \mathbf{0}, \ \mathbf{x} \geq \mathbf{0}$ and $\mathbf{c}^T\mathbf{x} = -1$, is called an infeasibility certificate for the system $\{\mathbf{y} : \ \mathbf{c} - A^T\mathbf{y} \geq \mathbf{0}\}$.

Farkas' Lemma for General Closed Convex Cones?

Consider the pair:

$$\{\mathbf{x} : \; A\mathbf{x} = \mathbf{b}, \; \mathbf{x} \in K\}$$

and

$$\{\mathbf{y} : \; -A^T\mathbf{y} \in K^*, \quad \mathbf{b}^T\mathbf{y} > 0\}.$$

Or in operator form: given data vector or matrix $\mathbf{a}_i$, $i = 1, ..., m$, and $\mathbf{b} \in \mathcal{R}^m$, an "alternative" system pair would be

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in K,$$

and

$$-\mathcal{A}^T\mathbf{y} \in K^*, \quad \mathbf{b}^T\mathbf{y} = 1(> 0)$$

where

$$\mathcal{A}\mathbf{x} = (\mathbf{a}_1 \bullet \mathbf{x}; ...; \mathbf{a}_m \bullet \mathbf{x}) \in \mathcal{R}^m \text{ and } \mathcal{A}^T\mathbf{y} = \sum_i^m y_i\mathbf{a}_i.$$

They hold for a general closed convex cone K?

An SDP Cone Example when "Alternative System" Failed

$$K = \mathcal{S}^2_+.$$

$$\mathbf{a}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

The Problem: $C := \{\mathcal{A}\mathbf{x} : \ \mathbf{x} \in K\}$ is not closed even when K is a closed convex cone.

When Farkas' Lemma Holds for General Cones?

Let K be a closed and convex cone in the rest of the course.

If there is $\mathbf{y}$ such that $-\mathcal{A}^T\mathbf{y} \in \operatorname{int} K^*$, then $C := \{\mathcal{A}\mathbf{x} : \ \mathbf{x} \in K\}$ is a closed convex cone. Consequently,

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in K,$$

and

$$-\mathcal{A}^T\mathbf{y} \in K^*, \quad \mathbf{b}^T\mathbf{y} = 1 (> 0)$$

are an alternative system pair.

And if there is $\mathbf{x}$ such that $\mathcal{A}^T\mathbf{x} = \mathbf{0}, \ \mathbf{x} \in \operatorname{int} K$, then

$$\mathcal{A}\mathbf{x} = \mathbf{0}, \quad \mathbf{x} \in K, \quad \mathbf{c} \bullet \mathbf{x} = -1 (< 0)$$

and

$$\mathbf{c} - \mathcal{A}^T\mathbf{y} \in K^*$$

are an alternative system pair.

Recall Conic LP

$$
\begin{array}{lll}
(CLP) & \text{minimize} & \mathbf{c} \bullet \mathbf{x} \\
& \text{subject to} & \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, ..., m, \ \mathbf{x} \in K, \\
& & (\ \mathcal{A}^T \mathbf{x} = \mathbf{b}\ ),
\end{array}
$$

where K is a closed and pointed convex cone.

Linear Programming (LP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = \mathcal{R}^n_+$

Second-Order Cone Programming (SOCP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = SOC = \{\mathbf{x} :\ x_1 \geq \|\mathbf{x}_{-1}\|_2\}$.

Semidefinite Programming (SDP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{S}^n$ and $K = \mathcal{S}^n_+$

p-Order Cone Programming (POCP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = POC = \{\mathbf{x} :\ x_1 \geq \|\mathbf{x}_{-1}\|_p\}$.

Here, $\mathbf{x}_{-1}$ is the vector $(x_2; ...; x_n) \in R^{n-1}$.

Cone K can be also a product of different cones, that is, $\mathbf{x} = (\mathbf{x}_1; \mathbf{x}_2; ...)$ where $\mathbf{x}_1 \in K_1$, $\mathbf{x}_2 \in K_2$,... and so on with linear constraints:

$$\mathcal{A}_1 \mathbf{x}_1 + \mathcal{A}_2 \mathbf{x}_2 + ... = \mathbf{b}.$$

LP, SOCP, and SDP Examples Again

$$
\begin{array}{lll}
(LP) & \text{minimize} & 2x_1 + x_2 + x_3 \\
 & \text{subject to} & x_1 + x_2 + x_3 = 1, \\
 & & (x_1; x_2; x_3) \geq \mathbf{0}.
\end{array}
$$

$$
\begin{array}{lll}
(SOCP) & \text{minimize} & 2x_1 + x_2 + x_3 \\
 & \text{subject to} & x_1 + x_2 + x_3 = 1, \\
 & & x_1 - \sqrt{x_2^2 + x_3^2} \geq 0.
\end{array}
$$

$$
\begin{array}{lll}
(SDP) & \text{minimize} & 2x_1 + x_2 + x_3 \\
 & \text{subject to} & x_1 + x_2 + x_3 = 1, \\
 & & \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0}.
\end{array}
$$

(SDP) can be rewriten as

$$
\begin{array}{ll}
\text{minimize} & \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix} \bullet \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \\
\text{subject to} & \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix} \bullet \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} = 1, \\
& \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0},
\end{array}
$$

that is

$$
\mathbf{c} = \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{a}_1 = \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix}.
$$

Dual of Conic LP

The dual problem to

$$(CLP) \quad \text{minimize} \quad \mathbf{c} \bullet \mathbf{x}$$
$$\text{subject to} \quad \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, ..., m, \ \mathbf{x} \in K.$$

is

$$(CLD) \quad \text{maximize} \quad \mathbf{b}^T \mathbf{y}$$
$$\text{subject to} \quad \sum_i^m y_i \mathbf{a}_i + \mathbf{s} = \mathbf{c}, \ \mathbf{s} \in K^*,$$

where $y \in \mathcal{R}^m$, $\mathbf{s}$ is called the dual slack vector/matrix, and K^* is the dual cone of K. The former is called the primal problem, and the latter is called dual problem.

Theorem 7 *The dual of the dual is the primal.*

The alternative system of a conic feasible set can be viewed as its dual with a positive objective value.

LP, SOCP, and SDP Examples

$$\begin{array}{ll} \text{min} & (2;\ 1;\ 1)^T\mathbf{x} \\ \text{s. t.} & \mathbf{e}^T\mathbf{x} = 1, \\ & \mathbf{x} \geq \mathbf{0}. \end{array} \qquad \begin{array}{ll} \text{max} & y \\ \text{s.t.} & \mathbf{e}\cdot y + \mathbf{s} = (2;\ 1;\ 1), \\ & \mathbf{s} \geq \mathbf{0}. \end{array}$$

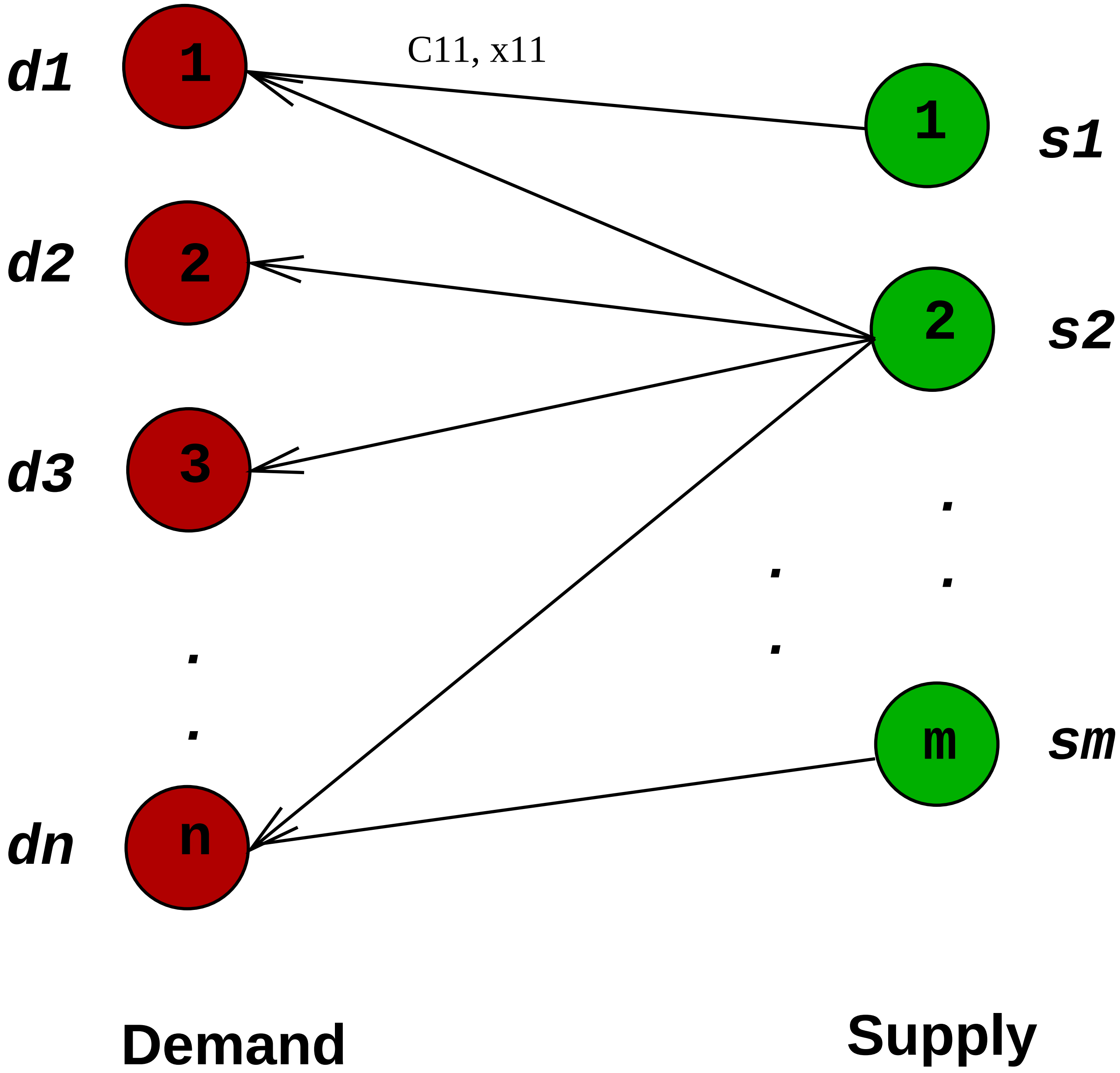
$$\begin{array}{ll} \text{min} & (2;\ 1;\ 1)^T\mathbf{x} \\ \text{s.t.} & \mathbf{e}^T\mathbf{x} = 1, \\ & x_1 - \|\mathbf{x}_{-1}\| \geq 0. \end{array} \qquad \begin{array}{ll} \text{max} & y \\ \text{s.t.} & \mathbf{e}\cdot y + \mathbf{s} = (2;\ 1;\ 1), \\ & s_1 - \|\mathbf{s}_{-1}\| \geq 0. \end{array}$$

$$
\begin{array}{ll}
\text{min} & \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \\
\text{s.t.} & \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} = 1, \\
& \mathbf{x} = \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0},
\end{array}
$$

$$
\begin{array}{ll}
\text{max} & y \\
\text{s.t.} & \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix} y + \mathbf{s} = \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix}, \\
& \mathbf{s} = \begin{pmatrix} s_1 & s_2 \\ s_2 & s_3 \end{pmatrix} \succeq \mathbf{0}.
\end{array}
$$

Recall Transportation Problem

$$
\begin{array}{lll}
\min & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} & \\
\text{s.t.} & \sum_{j=1}^{n} x_{ij} & = s_i,\ \forall i = 1, ..., m \\
 & \sum_{i=1}^{m} x_{ij} & = d_j,\ \forall j = 1, ..., n \\
 & x_{ij} & \geq 0,\ \forall i, j.
\end{array}
$$


C11, x11
d1
d2
d3
dn
1
2
3
n
1
2
m
s1
s2
sm
Demand
Supply

Transportation Dual: Economic Interpretation

$$\begin{array}{lll} \max & \sum_{i=1}^m s_i u_i + \sum_{j=1}^n d_j v_j & \\ \text{s.t.} & u_i + v_j & \le c_{ij}, \ \forall i, j. \end{array}$$

u_i: supply site unit price

v_i: demand site unit price

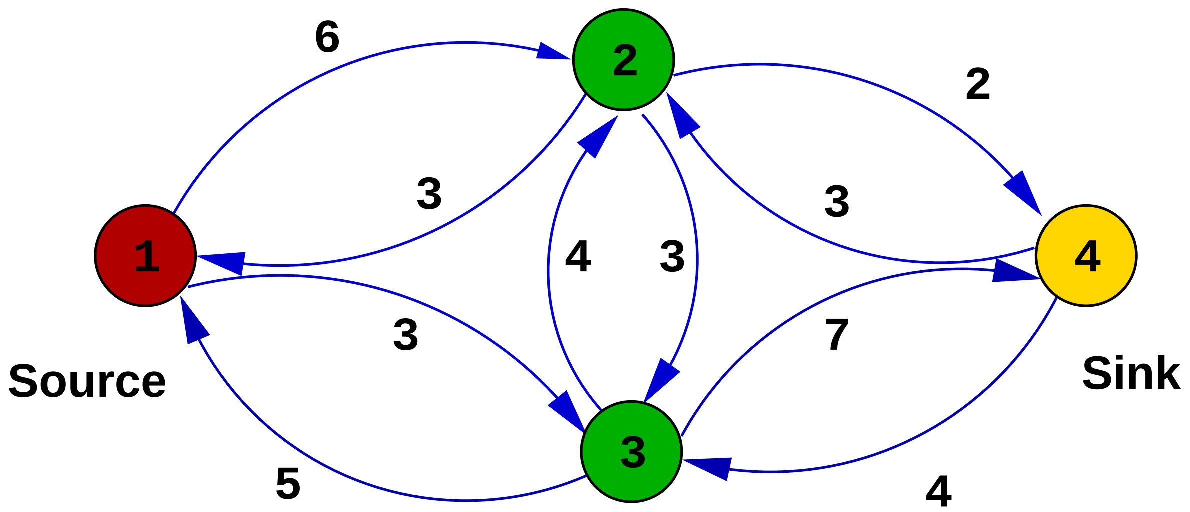
$u_i + v_j \le c_{ij}$: competitiveness

Max-Flow and Min-Cut

Given a directed graph with nodes $1, ..., m$ and edges $\mathcal{A}$, where node 1 is called source and node m is called the sink, and each edge (i, j) has a flow rate capacity k_{ij}. The Max-Flow problem is to find the largest possible flow rate from source to sink.

Let x_{ij} be the flow rate from node i to node j. Then the problem can be formulated as

$$
\begin{array}{ll}
\text{maximize} & x_{m1} \\
\text{subject to} & \sum_{j:(j,1)\in\mathcal{A}} x_{j1} - \sum_{j:(1,j)\in\mathcal{A}} x_{1j} + x_{m1} = 0, \\
& \sum_{j:(j,i)\in\mathcal{A}} x_{ji} - \sum_{j:(i,j)\in\mathcal{A}} x_{ij} = 0, \forall i = 2, ..., m-1, \\
& \sum_{j:(j,m)\in\mathcal{A}} x_{jm} - \sum_{j:(m,j)\in\mathcal{A}} x_{mj} - x_{m1} = 0, \\
& 0 \le x_{ij} \le k_{ij}, \ \forall (i,j) \in \mathcal{A}.
\end{array}
$$


6
2
3
3
1
4
3
4
3
7
Source
Sink
3
5
4

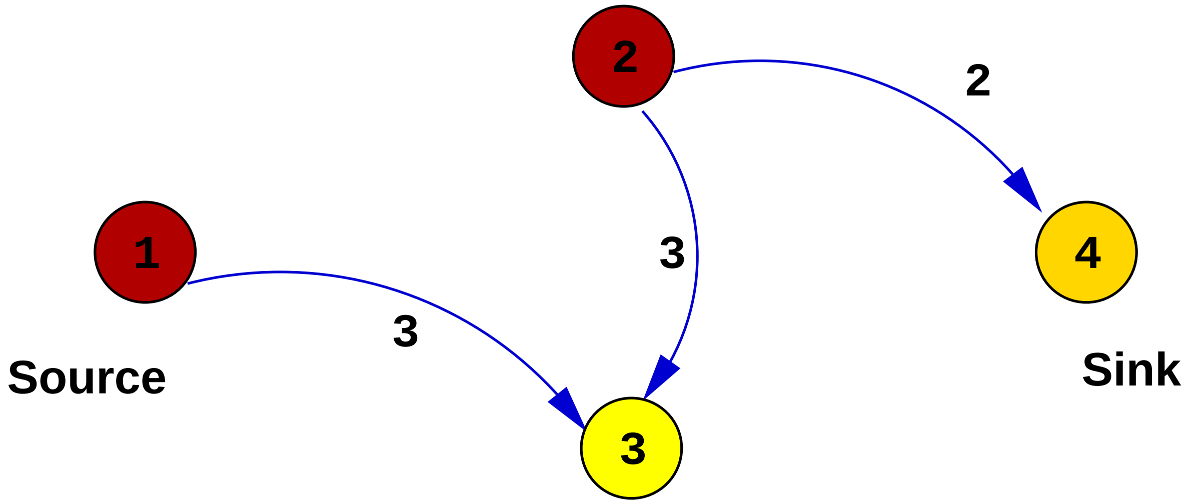
The dual of the Max-Flow problem

$$
\begin{array}{ll}
\text{minimize} & \sum_{(i,j)\in\mathcal{A}} k_{ij} z_{ij} \\
\text{subject to} & -y_i + y_j + z_{ij} \geq 0, \ \forall (i,j) \in \mathcal{A}, \\
 & y_1 - y_m = 1, \\
 & z_{ij} \geq 0, \ \forall (i,j) \in A.
\end{array}
$$

y_i: node potential value. At an optimal solution has property $y_1 = 1, \ y_m = 0$ and for all other i:

$$
y_i = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \notin S \end{cases}
$$

This problem is called the Min-Cut problem.


2
2
1
3
3
4
3
Source
Sink

The Dual of the MDP/Reinforcement Learning LP

Recall the cost-to-go value of the reinforcement learning LP problem:

$$
\begin{array}{lrcl}
\text{maximize}_{\mathbf{y}} & \sum_{i=1}^{m} y_i & & \\
\text{subject to} & y_1 - \gamma \mathbf{p}_j^T \mathbf{y} & \leq & c_j,\ j \in \mathcal{A}_1 \\
 & & \ldots & \\
 & y_i - \gamma \mathbf{p}_j^T \mathbf{y} & \leq & c_j,\ j \in \mathcal{A}_i \\
 & & \ldots & \\
 & y_m - \gamma \mathbf{p}_j^T \mathbf{y} & \leq & c_j,\ j \in \mathcal{A}_m.
\end{array}
$$

$$
\begin{array}{lccccc}
\text{minimize}_{\mathbf{x}} & \sum_{j \in \mathcal{A}_1} c_j x_j + & \ldots & + \sum_{j \in \mathcal{A}_m} c_j x_j & & \\
\text{subject to} & \sum_{j \in \mathcal{A}_1} (\mathbf{e}_1 - \gamma \mathbf{p}_j) x_j + & \ldots & + \sum_{j \in \mathcal{A}_m} (\mathbf{e}_m - \gamma \mathbf{p}_j) x_j & = & \mathbf{e}, \\
 & \ldots & x_j & \ldots & \geq & 0,\ \forall j,
\end{array}
$$

where $\mathbf{e}_i$ is the unit vector with 1 at the ith position and 0 everywhere else.

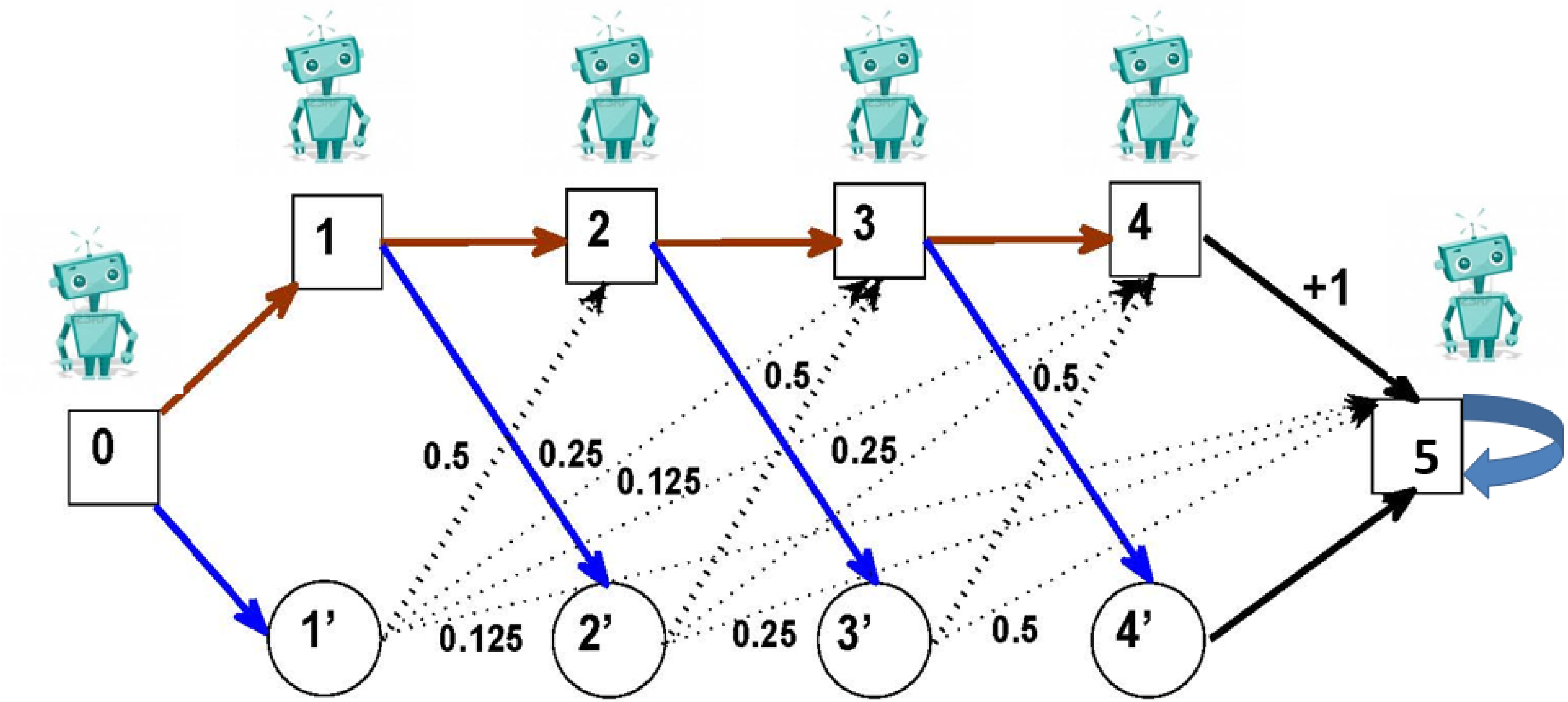
Interpretation of the Dual of the MDP/RL-LP

Variable x_j, $j \in \mathcal{A}_i$, is the state-action frequency or called flux, or the expected present value of the number of times that an individual is in state i and takes state-action j.

Thus, solving the problem entails choosing a state-action frequencies/fluxes that minimizes the expected present value of total costs for the infinite horizon, where the RHS is $(1; 1; 1; 1; 1; 1)$:

x:	(0_1)	(0_2)	(1_1)	(1_2)	(2_1)	(2_2)	(3_1)	(3_2)	(4_1)	(5_1)	**b**
c:	0	0	0	0	0	0	0	0	1	0	
(0)	1	1	0	0	0	0	0	0	0	0	1
(1)	$-\gamma$	0	1	1	0	0	0	0	0	0	1
(2)	0	$-\gamma/2$	$-\gamma$	0	1	1	0	0	0	0	1
(3)	0	$-\gamma/4$	0	$-\gamma/2$	$-\gamma$	0	1	1	0	0	1
(4)	0	$-\gamma/8$	0	$-\gamma/4$	0	$-\gamma/2$	$-\gamma$	0	1	0	1
(5)	0	$-\gamma/8$	0	$-\gamma/4$	0	$-\gamma/2$	0	$-\gamma$	$-\gamma$	$1-\gamma$	1

where state 5 is the absorbing state that has a infinite loops to itself.



The optimal dual solution is

$$x^*_{01} = 1,\ x^*_{11} = 1 + \gamma,\ x^*_{21} = 1 + \gamma + \gamma^2,\ x^*_{32} = 1 + \gamma + \gamma^2 + \gamma^3,\ x^*_{41} = 1,$$
$$x^*_{51} = \frac{1+2\gamma+\gamma^2+\gamma^3+\gamma^4}{1-\gamma}.$$

Two-Person Zero-Sum Game

Let P be the payoff matrix of a two-person, ”column” and ”row”, zero-sum game.

$$P = \begin{pmatrix} +3 & -1 & -4 \\ -3 & +1 & +4 \end{pmatrix}$$

Players usually use randomized strategies in such a game. A randomized strategy is a vector of probabilities, each associated with a particular decision.

Nash Equilibrium

In a Nash Equilibrium, if your (column) strategy is a pure strategy (one where you always play a single action), the expected payout for the (dominating) action that you are playing should be greater than or equal to the expected payout for any other action. If you are playing a randomized strategy, the expected payout for each action included in your strategy should be the same (if one were lower, you won't want to ever choose that action) and these payouts should be greater than or equal to the actions that aren't part of your strategy.

LP formulation of Nash Equilibrium

"Column" strategy:

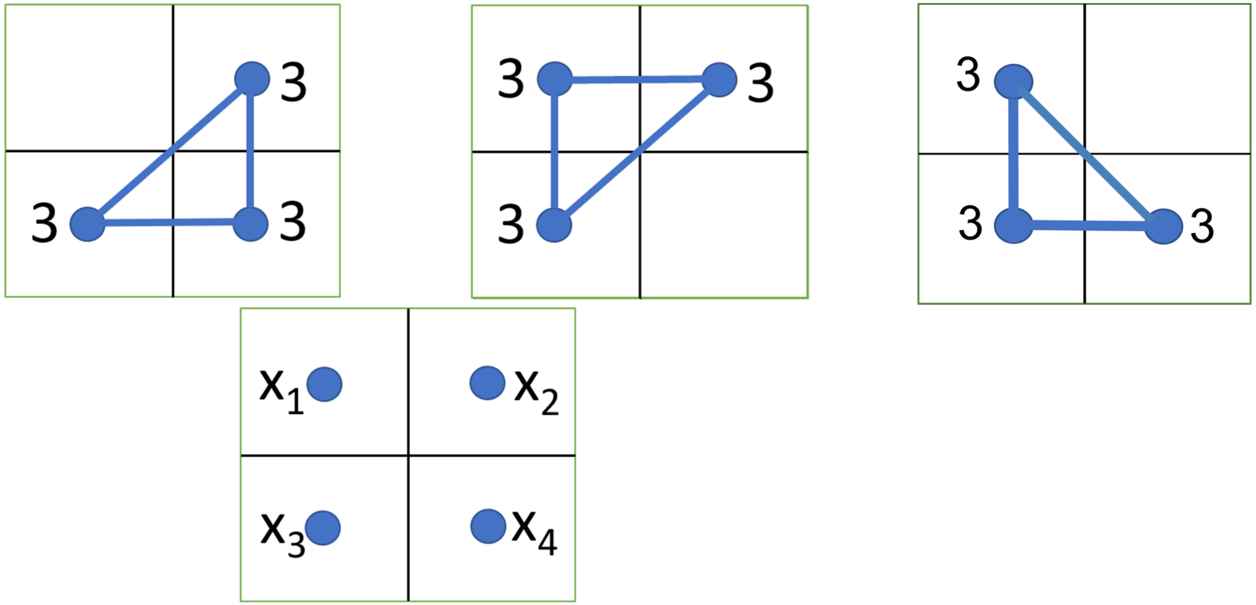
$$\begin{array}{ll} \max & v \\ s.t. & v\mathbf{e} \leq P\mathbf{x} \\ & \mathbf{e}^T\mathbf{x} = 1 \\ & \mathbf{x} \geq \mathbf{0}. \end{array}$$

"Row" strategy:

$$\begin{array}{ll} \min & u \\ s.t. & u\mathbf{e} \geq P^T\mathbf{y} \\ & \mathbf{e}^T\mathbf{y} = 1 \\ & \mathbf{y} \geq \mathbf{0}. \end{array}$$

They are dual to each other.

Recall Wassestein Barycenter Problem



Find distribution of $x_i, i = 1, 2, 3, 4$ to minimize

$$\begin{array}{ll} \min & WD_l(\mathbf{x}) + WD_m(\mathbf{x}) + WD_r(\mathbf{x}) \\ \text{s.t.} & x_1 + x_2 + x_3 + x_4 = 9, \qquad x_i \geq 0, \ i = 1, 2, 3, 4. \end{array}$$

The objective is a nonlinear function, but its gradient vector $\nabla WD_l(\mathbf{x})$, $\nabla WD_m(\mathbf{x})$ and $\nabla WD_l(\mathbf{x})$ are dual optimal solutions the three transportation problems, which can be solved in parallel or distributed fashion.