Optimization Models

EECS 127 / EECS 227AT

Laurent El Ghaoui

EECS department
UC Berkeley

Fall 2018

LECTURE 10

Linear Programs

> *Luckily the particular geometry used in my thesis was the one associated with the columns of the matrix instead of its rows.*
>
> George Dantzig

Outline

1. Linear programs
   - Half-spaces
   - Linear inequalities and polyhedra
   - The LP model
   - Polyhedral functions

2. Examples
   - ℓ_∞ and ℓ_1 regression problems
   - Drug production problem
   - Cash-flow management
   - Network flows

Half-spaces

For given $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$, the set $\mathcal{H}$ of points $x \in \mathbb{R}^n$ satisfying the linear inequality

$$a^\top x \leq b$$

is a closed half-space. Its boundary is the hyperplane defined by the *equality* $a^\top x = b$.

- vector a is normal to the boundary of the half-space and points outwards.
- scalar b tells us where along a the boundary of the half-space sits.

- When $b = 0$, the set $\mathcal{H}$ is the set of points forming an obtuse angle with a.
- When $b \neq 0$, choose any x_0 on the boundary of $\mathcal{H}$, for example $x_0 = (b/a^T a)a$; then $\mathcal{H}$ is the set of points such that $x - x_0$ forming an obtuse angle with a.

Geometry

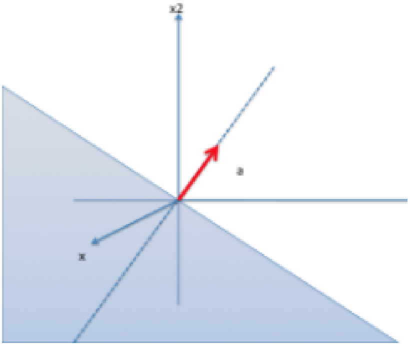


Figure: $b = 0$

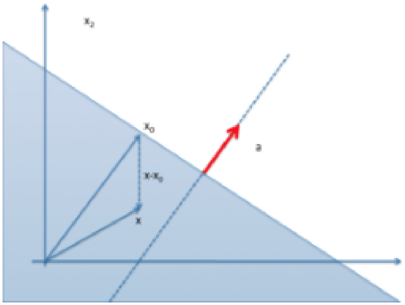


Figure: $b \neq 0$

Linear inequalities and polyhedra

For given $a_1, \dots, a_m \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$, the collection of m linear inequalities

$$a_i^\top x \leq b_i, \quad i = 1, \dots, m,$$

defines a region in $\mathbb{R}^m$ which is the intersection of m half-spaces and it is named a *polyhedron*.

Depending on the actual inequalities, these region can be unbounded, or bounded; in this latter case it is called a *polytope*.

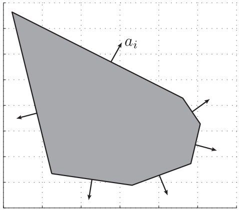
Linear inequalities and polyhedra

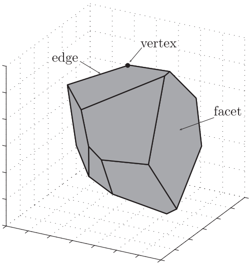
- It is often convenient to group several linear inequalities using matrix notation: we define

$$A = \begin{bmatrix} a_1^\top \\ a_2^\top \\ \vdots \\ a_m^\top \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

and then write inequalities in the equivalent matrix form (note: the inequality is taken component-wise)

$$Ax \leq b.$$





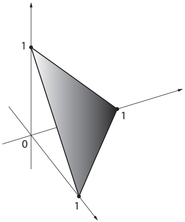
Example: the probability simplex

- The probability simplex is the polytope defined as

$$P = \{x \in \mathbb{R}^n : \, x \geq 0, \, \sum_{i=1}^{n} x_i = 1\}.$$

- The name suggests the fact that any x in the probability simplex has a natural interpretation of a discrete probability distribution, i.e. the x_i's are nonnegative and they sum up to one.
- The probability simplex in $\mathbb{R}^n$ has n vertices, which correspond to the standard orthonormal basis vectors for $\mathbb{R}^n$, that is

$$P = \mathrm{co}\{e^{(1)}, \ldots, e^{(n)}\}.$$



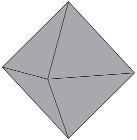
Example: the ℓ_1-norm ball

- The ℓ_1-norm ball is the set $\{x \in \mathbb{R}^n : \|x\|_1 \leq 1\}$, that is the set where $\sum_{i=1}^n |x_i| \leq 1$.
- This set is indeed a polytope, since the previous inequality is equivalent to a collection of 2^n linear inequalities. To see this fact, consider sign variables $s_i \in \{-1, 1\}$, $i = 1, \dots, n$. Then,

$$\sum_{i=1}^n |x_i| \quad = \quad \max_{s_i \in \{-1,1\}} \sum_{i=1}^n s_i x_i.$$

- Therefore $\|x\|_1 \leq 1$ if and only if $\max_{s_i \in \{-1,1\}} \sum_{i=1}^n s_i x_i \leq 1$, which is in turn equivalent to requiring that

$$\sum_{i=1}^n s_i x_i \leq 1, \quad \text{for all } s_i \in \{-1, 1\}, \; i = 1, \dots, m.$$



The LP model

- A linear optimization problem (or, linear program, LP) is one of standard form where every function $f_0, f_1, \ldots, f_m$ is affine. Thus, the feasible set of an LP is a polyhedron.
- Linear optimization problems admits several standard forms. E.g.,

$$\begin{aligned} p^* = \min_x \quad & c^\top x + d \\ \text{s.t.:} \quad & A_{\text{eq}} x = b_{\text{eq}} \\ & Ax \leq b, \end{aligned}$$

  where the inequalities are understood componentwise; we shall denote this form as the *inequality form* of the LP.
- The constant term d in the objective function is, of course, immaterial: it offsets the value of the objective but it has no influence on the minimizer.

The LP model

Geometric interpretation

- The set of points that satisfy the constraints of an LP (i.e., the feasible set) is a polyhedron (or a polytope, when it is bounded):
  $\mathcal{X} = \{x \in \mathbb{R}^n : A_{\mathrm{eq}}x = b_{\mathrm{eq}},\ Ax \leq b\}$.
- Let $x_f \in \mathcal{X}$ be a feasible point. To such point is associated the objective level $c^\top x_f$ (from now on, we assume without loss of generality that $d = 0$).
- A point $x_f \in \mathcal{X}$ is an optimal point, hence a solution of our LP, if and only if there is no other point $x \in \mathcal{X}$ with lower objective, that is:

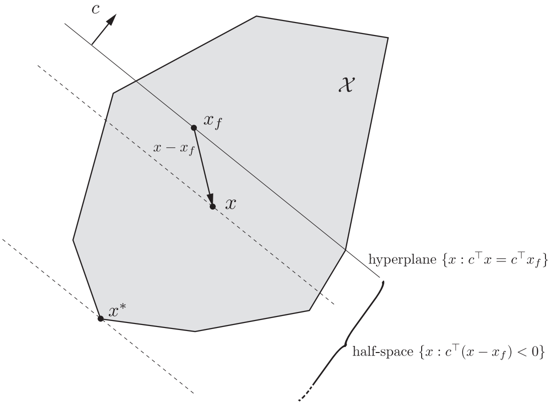
$$x_f \in \mathcal{X} \text{ is optimal for LP } \Leftrightarrow c^\top x \geq c^\top x_f,\ \forall y \in \mathcal{X}.$$

- Vice versa, the objective can be improved if one can find $x \in \mathcal{X}$ such that $c^\top(x - x_f) < 0$.

The LP model

Geometric interpretation

- Geometrically, the latter condition means that $\exists x$ in the intersection of the feasible set $\mathcal{X}$ and the open half-space $\{x : c^\top(x - x_f) < 0\}$, i.e., that we can move away from x_f in a direction that forms a negative inner product with direction c (descent direction), while maintaining feasibility. At an optimal point x^* there is no feasible descent direction,



Example: a toy LP

$$\min_{x \in \mathbb{R}^2} 3x_1 + 1.5x_2 \text{ subject to: } -1 \leq x_1 \leq 2,\ 0 \leq x_2 \leq 3$$

- The problem is an LP, and it can be put in standard inequality form:

$$\min_x : \min_{x \in \mathbb{R}^2} 3x_1 + 1.5x_2 \text{ subject to: } -x_1 \leq 1,\ x_1 \leq 2,\ -x_2 \leq 0,\ x_2 \leq 3$$

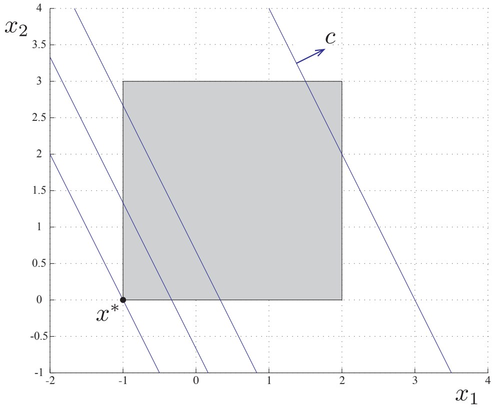
  or, using matrix notation, $\min_x c^\top x$ subject to $Ax \leq b$, with

$$c^\top = [3\ 1.5], \quad A = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 3 \end{bmatrix}.$$

- The level curves (curves of constant value) of the objective function are straight lines orthogonal to the objective vector, $c^\top = [3\ 1.5]$.

Example: a toy LP

- The problem amounts to find the smallest value of p such that $p = c^\top x$ for some feasible x.
- The optimal point is $x^* = [-1\ 0]^\top$, and the optimal objective value is $p^* = -3$.



Polyhedral functions

- We say that a function $f : \mathbb{R}^n \to \mathbb{R}$ is polyhedral if its epigraph is a polyhedron, that is if $\mathrm{epi}\, f = \left\{(x,t) \in \mathbb{R}^{n+1} : f(x) \leq t\right\}$ can be represented as

$$\mathrm{epi}\, f = \left\{(x,t) \in \mathbb{R}^{n+1} : C \begin{bmatrix} x \\ t \end{bmatrix} \leq d \right\},$$

  for some matrix $C \in \mathbb{R}^{m,n+1}$ and vector $d \in \mathbb{R}^m$.

- Polyhedral functions include in particular functions that can be expressed as a maximum of a finite number of affine functions:

$$f(x) = \max_{i=1,\ldots,m} a_i^\top x + b_i,$$

  where $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $i = 1, \ldots, m$.

- The epigraph of f

$$\mathrm{epi}\, f = \left\{(x,t) \in \mathbb{R}^{n+1} : \max_{i=1,\ldots,m} a_i^\top x + b_i \leq t \right\}$$

  can indeed be expressed as the polyhedron

$$\mathrm{epi}\, f = \left\{(x,t) \in \mathbb{R}^{n+1} : a_i^\top x + b_i \leq t,\ i = 1, \ldots, m \right\}.$$

Polyhedral functions

- The ℓ_∞-norm function $f(x) = \|x\|_\infty$, $x \in \mathbb{R}^n$, is polyhedral since it can be written as the maximum of $2n$ affine functions:

$$f(x) = \max_{i=1,\ldots,n} \max(x_i, -x_i).$$

- Polyhedral functions also include functions that can be expressed as

$$f(x) = \sum_{j=1}^{q} \max_{i=1,\ldots,m} a_{ij}^\top x + b_{ij},$$

- Condition $(x, t) \in \mathrm{epi}\ f$ is equivalent to $\exists u \in \mathbb{R}^q$ such that

$$\sum_{j=1}^{q} u_j \leq t, \quad a_{ij}^\top x + b_{ij} \leq u_j,\ i = 1, \ldots, m;\ j = 1, \ldots, q$$

  hence, $\mathrm{epi}\ f$ is the projection (on the space of (x, t)-variables) of a polyhedron, which is itself a polyhedron.

- The ℓ_1-norm function $f(x) = \|x\|_1$, $x \in \mathbb{R}^n$, is polyhedral since it can be written as the sum of maxima of affine functions: $f(x) = \sum_{i=1,\ldots,n} \max(x_i, -x_i)$.

Minimization of polyhedral functions

- The problem of minimizing a polyhedral function, under linear equality or inequality constraints, can be cast as an LP.
- If f is polyhedral, then

$$\min_{x} f(x) \quad \text{s.t.: } Ax \leq b,$$

  is cast as

$$\min_{x,t} t \quad \text{s.t.: } Ax \leq b, \ (x,t) \in \mathrm{epi}\ f.$$

- Since $\mathrm{epi}\ f$ is a polyhedron, it can be expressed as in (15), hence the problem above is an LP.
- Notice however that explicit representation of the LP in a standard form may require introduction of additional slack variables, which are needed for representation of the epigraph.

ℓ_∞ regression problems

$$\min_x \|Ax - b\|_\infty, \quad A \in \mathbb{R}^{m,n}, b \in \mathbb{R}^m.$$

- The problem may be first rewritten in epigraphic form, adding a slack scalar variable t

$$\min_{x,t} t \quad \text{s.t.: } \|Ax - b\|_\infty \leq t$$

- Then we observe that

$$\|Ax - b\|_\infty \leq t \Leftrightarrow \max_{i=1,\ldots,m} |a_i^\top x - b_i| \leq t \Leftrightarrow |a_i^\top x - b_i| \leq t, \; i = 1, \ldots, m.$$

- Hence, problem is equivalent to the following LP in variables $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$:

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{s.t.:} \quad & a_i^\top x - b_i \leq t, \quad i = 1, \ldots, m \\ & a_i^\top x - b_i \geq -t, \quad i = 1, \ldots, m. \end{aligned}$$

ℓ_1 regression problems

$$\min_x \|Ax - b\|_1, \quad A \in \mathbb{R}^{m,n}, b \in \mathbb{R}^m.$$

- Equivalent to a problem with a vector u of additional slack variables $u \in \mathbb{R}^m$:

$$\min_{x,u} \sum_{i=1}^{m} u_i \quad \text{s.t.: } |a_i^\top x - b_i| \le u_i, \ i = 1, \ldots, m,$$

- This is in turn easily cast as a standard LP as follows

$$\begin{array}{lll} \min\limits_{x,u} & \mathbf{1}^\top u & \\ \text{s.t.:} & a_i^\top x - b_i \le u_i, & i = 1, \ldots, m \\ & a_i^\top x - b_i \ge -u_i, & i = 1, \ldots, m. \end{array}$$

Drug production problem

A company produces two kinds of drugs, DrugI and DrugII, containing a specific active agent A, which is extracted from raw materials purchased on the market.

There are two kinds of raw materials, RawI and RawII, which can be used as sources of the active agent. The related production, cost and resource data are given next. The goal is to find the production plan which maximizes the profit of the company.

Problem data

	DrugI	DrugII
Selling price, $ per 1000 packs	5,500	6,100
Content of agent A, g per 1000 packs	0.500	0.600
Manpower required, hours per 1000 packs	90.0	100.0
Equipment required, hours per 1000 packs	40.0	50.0

	Purchasing price ($ per kg)	Content of agent A (g per kg)
RawI	100.00	0.01
RawII	199.90	0.02

Budget ($)	Manpower (hours)	Equipment (hours)	Capacity of raw materials' storage (kg)
100,000	2,000	800	1,000

LP model: variables, objective, constraints

- Variables: denote by $x_{rmDrugI}$, $x_{rmDrugII}$ the amounts (in 1000 of packs) of Drug I and II produced, while x_{rmRawI}, $x_{rmRawII}$ denote the amounts (in kg) of raw materials to be purchased.
- Objective function: $f_0(x) = f_{\text{costs}}(x) - f_{\text{income}}(x)$, where

$$f_{\text{costs}}(x) = 100x_{\text{RawI}} + 199.90x_{\text{RawII}} + 700x_{\text{DrugI}} + 800x_{\text{DrugII}}$$

  represents the purchasing and operational costs, and

$$f_{\text{income}}(x) = 5,500x_{\text{DrugI}} + 6100x_{\text{DrugII}}$$

  contain the unit market prices as coefficients.
- Balance of active agent: says that the amount of raw material must be enough to produce the drugs

$$0.01x_{\text{RawI}} + 0.02x_{\text{RawII}} - 0.50x_{\text{DrugI}} - 0.60x_{\text{DrugII}} \geq 0.$$

LP model: more constraints & full model

- Storage: $x_{\text{RawI}} + x_{\text{RawII}} \leq 1000$.
- Manpower: $90.0x_{\text{DrugI}} + 100.0x_{\text{DrugII}} \leq 2000$
- Equipment: $40.0x_{\text{DrugI}} + 50.0x_{\text{DrugII}} \leq 800$.
- Budget: $100.0_{\text{RawI}} + 199.90x_{\text{RawII}} + 700x_{\text{DrugI}} + 800x_{\text{DrugII}} \leq 100,000$.
- Sign constraints: all variables are non-negative.

Putting this together we get the LP:

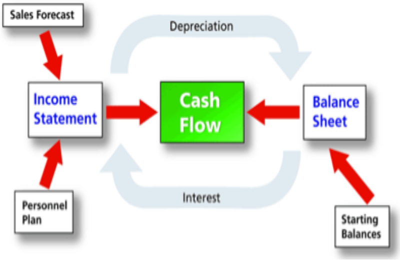
$$\min c^\top x \ : \ Ax \leq b, \ \ x \geq 0,$$

where $x = (x_{\text{RawI}}, x_{\text{RawII}}, x_{\text{DrugI}}, x_{\text{DrugII}})$, and

$$A = \begin{pmatrix} -0.01 & -0.02 & 0.500 & 0.600 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 90.0 & 100.0 \\ 0 & 0 & 40.0 & 50.0 \\ 100.0 & 199.9 & 700 & 800 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1000 \\ 2000 \\ 800 \\ 100000 \end{pmatrix}, \quad c = \begin{pmatrix} 100 \\ 199.9 \\ -5500 \\ -6100 \end{pmatrix}.$$

Examples

Cash-flow management



A company needs to choose between three financial instruments to cover its liabilities over a six-months period into the future:

- A line of credit of maximum amount $\$100k$, with interest rate 1% per month;
- In any of the first 3 months it can issue 90-day commercial paper (a type of unsecured debt) bearing a total interest of 2% for the 3-month period;
- Excess funds (cash) can be invested at 0.3% per month.

Examples

Cash-flow management: variables and decision problem

Variables:

- Balance on the credit line x_i for month $i = 1, 2, 3, 4, 5$.
- Amount y_i of commercial paper issued $(i = 1, 2, 3)$.
- Excess funds z_i for month $i = 1, 2, 3, 4, 5$.
- z_6, the company's wealth at the end of the 6-month period.

With these variables we have to meet certain cash-flow requirements each month.

Decision problem:

$$\text{maximize } z_6 \text{ subject to } \begin{cases} \text{Bounds on variables,} \\ \text{Cash-flow balance equations.} \end{cases}$$

Examples

Cash-flow management: optimization model

$$
\begin{array}{lrcl}
\max\limits_{x,y,z} & z_6 & & \\
\text{s.t.} & x_1 + y_1 - z_1 & = & 150, \\
& x_2 + y_2 - 1.01x_1 + 1.003z_1 - z_2 & = & 100, \\
& x_3 + y_3 - 1.01x_2 + 1.003z_2 - z_3 & = & -200, \\
& x_4 - 1.02y_1 - 1.01x_3 + 1.003z_3 - z_4 & = & 200, \\
& x_5 - 1.02y_2 - 1.01x_4 + 1.003z_4 - z_5 & = & -50, \\
& -1.02y_3 - 1.01x_5 + 1.003z_5 - z_6 & = & -300, \\
& 0 \leq x \leq 100, \quad y \geq 0, \quad z \geq 0 & &
\end{array}
$$

The right-hand side contains the liabilities that must be met.

Challenges:

- Future liabilities are uncertain.
- Some instruments may have varying (thus, uncertain) interest rates.

Example

Network flows[1]

We consider a network (directed graph) having m nodes connected by n directed arcs (ordered pairs (i, j)). We assume there is at most one arc from any node i to any node j, and no self-loops.

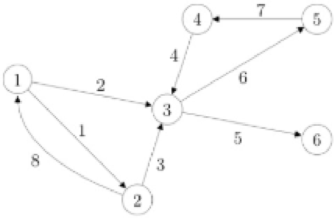
We define the *arc-node incidence matrix* $A \in \mathbb{R}^{m,n}$ to be the matrix with coefficients

$$A_{ij} = \begin{cases} 1 & \text{if arc } j \text{ starts at node } i, \\ -1 & \text{if it ends at node } i, \\ 0 & \text{otherwise.} \end{cases}$$

Note that the column sums of A are zero: $\mathbf{1}^\top A = 0$.

[1] Thanks to L. Vandenberghe for his slides, at `http://www.seas.ucla.edu/~vandenbe/ee236a/lectures/networks.pdf`

Incidence matrix example



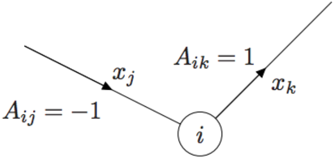
The figure shows the graph associated with the arc-node incidence matrix

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & -1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}.$$

Flows and balance equations

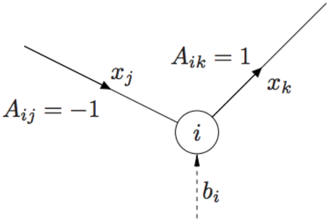
A *flow* (of traffic, information, charge) is represented by a vector $xin\mathbb{R}^n$, and the *total flow leaving node i* is then the vector with components

$$(Ax)_i = \sum_{j=1}^{n} A_{ij} x_j, \quad i = 1, \ldots, m.$$



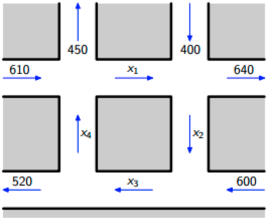
Supply vector

Define a *supply* vector $b \in \mathbb{R}^m$, with possibly positive or negative components, representing supply and demand. We assume that $\mathbf{1}^T b = 0$, so that the total supply equals the total demand.



The condition $Ax = b$ expresses the *balance equations* of the network.

Example: traffic equations[2]



$x_1 = ?, \quad x_2 = ?, \quad x_3 = ?, \quad x_4 = ?$

The traffic balance equations are $Ax = b$, with

$$A = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} -160 \\ 240 \\ -600 \\ 520 \end{pmatrix}.$$

[2]See `http://livebooklabs.com/keeppies/c5a5868ce26b8125/ac283f2cc94c29f3`

Minimum cost network flow problem

We also define a cost vector $c \in \mathbb{R}^n$, containing the unit cost of traversing each edge, so that the total cost of a flow x is

$$c^\top x.$$

The minimum cost network flow problem has the LP form

$$\min_x : c^T x \ : \ Ax = b, \ \ l \leq x \leq u,$$

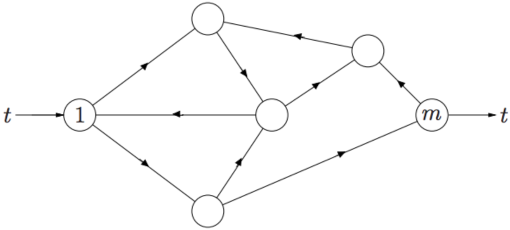
where $l \leq u$ are two vectors that provide upper and lower bounds on the flow.

Maximum flow problem[3]

In the maximum flow problem, we seek to maximize the flow between node 1 (the source) and node m (the sink):

$$\min_{x,t} : t \;\; : \;\; Ax = te, \;\; l \leq x \leq u,$$

with $e^T = (1, 0, \ldots, 0, -1)$.



Both problems can be solved at scale with dedicated algorithms.

[3] See also: http://www.mathcs.emory.edu/~cheung/Courses/323/Syllabus/NetFlow/max-flow-lp.html.

Example

Swimming coach's problem[4]

The coach of a swim team needs to assign swimmers to a 200-yard medley relay team to compete in a tournament. The problem is that his best swimmers are good in more than one stroke, so its not clear which swimmer to assign to which stroke. Here are the best times for each swimmer:

Stroke	Carl	Chris	David	Tony	Ken
Backstroke	37.7	32.9	33.8	37.0	35.4
Breaststroke	43.4	33.1	42.2	34.7	41.8
Butterfly	33.3	28.5	38.9	30.4	33.6
Freestyle	29.2	26.4	29.6	28.5	31.1

Table: 4×5 matrix of best times for every stroke and swimmer.

[4] From B. Van Roy, K. Mason

Assignment problem

Variable is $X \in \mathbb{R}^{4\times 5}$

$$\min_{X\geq 0} \ \mathrm{trace}\, M^\top X \ : \quad \begin{array}{l} X\mathbf{1} \leq \mathbf{1} \text{ (at most one swimmer per stroke)} \\ X^\top \mathbf{1} = \mathbf{1} \text{ (one stroke per swimmer)} \end{array}$$

Solving the LP above, we actually get an integer solution:

$$X = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

(It can be proven that integral solutions always exist for such LPs; we don't have to consider fractional swimmers ...)