Linear Programming

**Significance.**
- Quintessential tool for optimal allocation of scarce resources, among a number of competing activities.
- Powerful model generalizes many classic problems:
  - shortest path, max flow, multicommodity flow, MST, matching, 2-person zero sum games
- Ranked among most important scientific advances of 20th century.
  - accounts for a major proportion of all scientific computation
- Helps find "good" solutions to NP-hard optimization problems.
  - optimal solutions (branch-and-cut)
  - provably good solutions (randomized rounding)

---

Brewery Problem

Small brewery produces ale and beer.
- Production limited by scarce resources: corn, hops, barley malt.
- Recipes for ale and beer require different proportions of resources.

<table>
<thead>
<tr>
<th>Beverage</th>
<th>Corn (pounds)</th>
<th>Hops (ounces)</th>
<th>Malt (pounds)</th>
<th>Profit ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ale</td>
<td>5</td>
<td>4</td>
<td>35</td>
<td>13</td>
</tr>
<tr>
<td>Beer</td>
<td>15</td>
<td>4</td>
<td>20</td>
<td>23</td>
</tr>
</tbody>
</table>

How can brewer maximize profits?
- Devote all resources to beer: 32 barrels of beer ⇒ $736.
- Devote all resources to ale: 34 barrels of ale ⇒ $442.
- 7½ barrels of ale, 29½ barrels of beer ⇒ $776.
- 12 barrels of ale, 28 barrels of beer ⇒ $800.

**Mathematical Formulation**

\[
\begin{align*}
\text{maximize} & \quad 13A + 23B \\
\text{subject to} & \quad 5A + 15B \leq 480 \\
& \quad 4A + 4B \leq 160 \\
& \quad 35A + 20B \leq 1190 \\
& \quad A, B \geq 0
\end{align*}
\]
Brewery Problem: Feasible Region

- Hops: $4A + 4B \leq 160$
- Malt: $35A + 20B \leq 1190$
- Corn: $5A + 15B \leq 480$

Brewery Problem: Objective Function

- Objective Function:
  
  $\text{Maximize } 13A + 23B$
  
  Subject to:

  - $13A + 23B = 800$
  - $13A + 23B = 1600$
  - $13A + 23B = 442$

Brewery Problem: Geometry

- Brewery problem observation: Regardless of coefficients of linear objective function, there exists an optimal solution that is an extreme point.

Linear Programming

LP "standard" form.
- Input data: rational numbers $c_j, b_i, a_{ij}$.
- Maximize linear objective function.
- Subject to linear inequalities.

\[
\begin{align*}
\text{(P)} & \quad \text{max } \sum_{j=1}^{n} c_j x_j \\
\text{s.t.} & \quad \sum_{j=1}^{n} a_{ij} x_j \leq b_i \quad 1 \leq i \leq m \\
& \quad x_j \geq 0 \quad 1 \leq j \leq n
\end{align*}
\]
LP: Geometry

Geometry.
- Forms an n-dimensional polyhedron.
- Convex: if y and z are feasible solutions, then so is \( \frac{1}{2}y + \frac{1}{2}z \).
- Extreme point: feasible solution \( x \) that can't be written as \( \frac{1}{2}y + \frac{1}{2}z \) for any two distinct feasible solutions \( y \) and \( z \).

LP: Algorithms

Simplex. (Dantzig 1947)
- Developed shortly after WWII in response to logistical problems: used for 1948 Berlin airlift.
- Practical solution method that moves from one extreme point to a neighboring extreme point.
- Finite (exponential) complexity, but no polynomial implementation known.

LP: Polynomial Algorithms

Ellipsoid. (Khachian 1979, 1980)
- Solvable in polynomial time: \( O(n^4 L) \) bit operations.
  - \( n \) = # variables
  - \( L \) = # bits in input
- Theoretical tour de force.
- Not remotely practical.

Karmarkar's algorithm. (Karmarkar 1984)
- \( O(n^{3.5} L) \).
- Polynomial and reasonably efficient implementations possible.

Interior point algorithms.
- \( O(n^3 L) \).
- Competitive with simplex!
  - will likely dominate on large problems soon
- Extends to even more general problems.
Primal problem.

\[
\begin{align*}
\text{max} & \quad 4x_1 + x_2 + 5x_3 + 3x_4 \\
\text{s.t.} & \quad x_1 - x_2 - x_3 + 3x_4 \leq 1 \\
& \quad 5x_1 + x_2 + 3x_3 + 8x_4 \leq 55 \\
& \quad -x_1 + 2x_2 + 3x_3 - 5x_4 \leq 3 \\
& \quad x_1, x_2, x_3, x_4 \geq 0
\end{align*}
\]

Find a lower bound on optimal value.

- \((x_1, x_2, x_3, x_4) = (0, 0, 1, 0)\) \Rightarrow z^* \geq 5.
- \((x_1, x_2, x_3, x_4) = (2, 1, 1, 1/3)\) \Rightarrow z^* \geq 15.
- \((x_1, x_2, x_3, x_4) = (3, 0, 2, 0)\) \Rightarrow z^* \geq 22.
- \((x_1, x_2, x_3, x_4) = (0, 14, 0, 5)\) \Rightarrow z^* \geq 29.

Find an upper bound on optimal value.

- Multiply 2\text{nd} inequality by 2: \(10x_1 + 2x_2 + 6x_3 + 16x_4 \leq 110\).
  \Rightarrow z^* = 4x_1 + x_2 + 5x_3 + 3x_4 \leq 10x_1 + 2x_2 + 6x_3 + 16x_4 \leq 110.
- Adding 2\text{nd} and 3\text{rd} inequalities: \(4x_1 + 3x_2 + 6x_3 + 3x_4 \leq 58\).
  \Rightarrow z^* = 4x_1 + x_2 + 5x_3 + 3x_4 \leq 4x_1 + 3x_2 + 6x_3 + 3x_4 \leq 58.

General idea: add linear combination \((y_1, y_2, y_3)\) of the constraints.

\[
\begin{align*}
(y_1 + 5y_2 - y_3) x_1 + (-y_1 + y_2 + 2y_3) x_2 + \\
(-y_1 + 3y_2 + 3y_3) x_3 + (3y_1 + 8y_2 - 5y_3) x_4 & \leq y_1 + 55y_2 + 3y_3
\end{align*}
\]

Dual problem.

\[
\begin{align*}
\text{min} & \quad y_1 + 55y_2 + 3y_3 \\
\text{s.t.} & \quad y_1 + 5y_2 - y_3 \geq 4 \\
& \quad -y_1 + y_2 + 2y_3 \geq 1 \\
& \quad -y_1 + 3y_2 + 3y_3 \geq 5 \\
& \quad 3y_1 + 8y_2 - 5y_3 \geq 3 \\
& \quad y_1, y_2, y_3 \geq 0
\end{align*}
\]
**LP Duality**

Primal and dual linear programs: given rational numbers $a_{ij}, b_i, c_j$, find values $x_i, y_j$ that optimize (P) and (D).

(P) \[ \max \sum_{j=1}^{n} c_j x_j \]  
\[ \text{s.t.} \sum_{j=1}^{n} a_{ij} x_j \leq b_i \quad 1 \leq i \leq m \]  
\[ x_j \geq 0 \quad 1 \leq j \leq n \]

(D) \[ \min \sum_{i=1}^{m} b_i y_i \]  
\[ \text{s.t.} \sum_{i=1}^{m} a_{ij} y_i \geq c_j \quad 1 \leq j \leq n \]  
\[ y_i \geq 0 \quad 1 \leq i \leq m \]

Duality Theorem (Gale-Kuhn-Tucker 1951, Dantzig-von Neumann 1947). If (P) and (D) are nonempty then $\max = \min$.

- Dual solution provides certificate of optimality $\Rightarrow$ decision version $\in$ NP $\cap$ co-NP.
- Special case: max-flow min-cut theorem.
- Sensitivity analysis.

**LP Duality: Economic Interpretation**

Brewer’s problem: find optimal mix of beer and ale to maximize profits.

$\begin{align*}
(P) & \quad \max 13A + 23B \\
\text{s.t.} & \quad 5A + 15B \leq 480 \\
& \quad 4A + 4B \leq 160 \\
& \quad 35A + 20B \leq 1190 \\
& \quad A, B \geq 0
\end{align*}$

Entrepreneur’s problem: Buy individual resources from brewer at minimum cost.

- C, H, M = unit price for corn, hops, malt.
- Brewer won’t agree to sell resources if $5C + 4H + 35M < 13$.

$\begin{align*}
(D) & \quad \min 480C + 160H + 1190M \\
\text{s.t.} & \quad 5C + 4H + 35M \geq 13 \\
& \quad 15C + 4H + 20M \geq 23 \\
& \quad C, H, M \geq 0
\end{align*}$

**Standard Form**

Standard form. $\begin{align*}
(P) & \quad \max \sum_{j=1}^{n} c_j x_j \\
\text{s.t.} & \quad \sum_{j=1}^{n} a_{ij} x_j \leq b_i \quad 1 \leq i \leq m \\
& \quad x_j \geq 0 \quad 1 \leq j \leq n
\end{align*}$

Easy to handle variants.

- $x + 2y - 3z \geq 17 \Rightarrow -x - 2y + 3z \leq -17$.
- $x + 2y - 3z = 17 \Rightarrow x + 2y - 3z \leq 17, -x - 2y + 3z \leq -17$.
- $\min x + 2y - 3z \Rightarrow \max -x - 2y + 3z$.
- $x$ unrestricted $\Rightarrow x = y - z, y \geq 0, z \geq 0$.
LP Application: Weighted Bipartite Matching

Assignment problem. Given a complete bipartite network $K_{n,n}$ and edge weights $c_{ij}$, find a perfect matching of minimum weight.

$$\min \sum_{1 \leq i \leq n} \sum_{1 \leq j \leq n} c_{ij} x_{ij}$$

s.t.

$$\sum_{1 \leq j \leq n} x_{ij} = 1 \quad 1 \leq i \leq n$$

$$\sum_{1 \leq i \leq n} x_{ij} = 1 \quad 1 \leq j \leq n$$

$$x_{ij} \geq 0 \quad 1 \leq i,j \leq n$$

Birkhoff-von Neumann theorem (1946, 1953). All extreme points of the above polytope are {0-1}-valued.

Corollary. Can solve assignment problem using LP techniques since LP algorithms return optimal solution that is an extreme point.

Remark. Polynomial combinatorial algorithms also exist.

Proof (by contradiction). Suppose $x$ is a fractional feasible solution.

- Consider $A = \{(i, j) : 0 < x_{ij}\}$.
- Claim: there exists a perfect matching in $(V, A)$.
  - Fractional flow gives fractional perfect matching

Define $\varepsilon = \min \{ x_{ij} : x_{ij} > 0 \}$,

$$x^1 = (1 - \varepsilon) x + \varepsilon y,$$

$$x^2 = (1 + \varepsilon) x - \varepsilon y.$$
Multicommodity flow. Given a network $G = (V, E)$ with edge capacities $u(e) \geq 0$, edge costs $c(e) \geq 0$, set of commodities $K$, and supply / demand $d^k(v)$ for commodity $k$ at node $v$, find a minimum cost flow that satisfies all of the demand.

$$\min \sum_{k \in K} \sum_{e \in E} c^k(e) x^k(e)$$
$$\text{s.t. } \sum_{e \in \text{in to } v} x^k(e) - \sum_{e \in \text{out of } v} x^k(e) = d^k(v) \quad v \in V, k \in K$$
$$\sum_{k \in K} \sum_{e \in E} x^k(e) \leq u(e) \quad e \in E$$
$$x^k(e) \geq 0 \quad e \in E, k \in K$$

Applications.
- Transportation networks.
- Communication networks (Akamai).
- Solving $Ax = b$ with Gaussian elimination, preserving sparsity.
- VLSI design.

Weighted vertex cover. Given an undirected graph $G = (V, E)$ with vertex weights $w_v \geq 0$, find a minimum weight subset of nodes $S$ such that every edge is incident to at least one vertex in $S$.

- NP-hard even if all weights are 1.

Integer programming formulation.

$$(ILP) \min \sum_{v \in V} w_v x_v$$
$$\text{s.t. } x_v + x_w \geq 1 \quad (v, w) \in E$$
$$x_v \in \{0, 1\} \quad v \in V$$

If $x^*$ is optimal solution to (ILP), then $S = \{v \in V : x_v^* = 1\}$ is a min weight vertex cover.

Linear programming relaxation.

$$(LP) \min \sum_{v \in V} w_v x_v$$
$$\text{s.t. } x_v + x_w \geq 1 \quad (v, w) \in E$$
$$x_v \geq 0 \quad v \in V$$

- Note: optimal value of (LP) is $\leq$ optimal value of (ILP), and may be strictly less.
  - clique on $n$ nodes requires $n-1$ nodes in vertex cover
  - LP solution $x^* = 1/2$ has value $n / 2$

- Provides lower bound for approximation algorithm.

- How can solving LP help us find good vertex cover?

  - Round fractional values.
Weighted Vertex Cover

**Theorem.** If $x^*$ is optimal solution to (LP), then $S = \{v \in V : x^*_v \geq \frac{1}{2}\}$ is a vertex cover within a factor of 2 of the best possible.

- Provides 2-approximation algorithm.
- Solve LP, and then round.

$S$ is a vertex cover.
- Consider an edge $(v,w) \in E$.
- Since $x^*_v + x^*_w \geq 1$, either $x^*_v \geq \frac{1}{2}$ or $x^*_w \geq \frac{1}{2} \Rightarrow (v,w)$ covered.

$S$ has small cost.
- Let $S^*$ be optimal vertex cover.

\[
\sum_{v \in S^*} w_v \geq \sum_{v \in V} w_v x^*_v \\
\geq \sum_{v \in S} w_v x^*_v \\
\geq \frac{1}{2} \sum_{v \in S} w_v
\]

\[x^*_v \geq \frac{1}{2}\]

**Good news.**
- 2-approximation algorithm is basis for most practical heuristics.
  - can solve LP with min cut $\Rightarrow$ faster
  - primal-dual schema $\Rightarrow$ linear time 2-approximation
- PTAS for planar graphs.
- Solvable on bipartite graphs using network flow.

**Bad news.**
- NP-hard even on 3-regular planar graphs with unit weights.
- If $P \neq NP$, then no $\rho$-approximation for $\rho < 4/3$, even with unit weights. (Dinur-Safra, 2001)

Maximum Satisfiability

MAX-SAT: Given clauses $C_1, C_2, \ldots, C_m$ in CNF over Boolean variables $x_1, x_2, \ldots, x_n$, and integer weights $w_j \geq 0$ for each clause, find a truth assignment for the $x_i$ that maximizes the total weight of clauses satisfied.

- NP-hard even if all weights are 1.
- Ex.

<table>
<thead>
<tr>
<th>$C_i$</th>
<th>$w_i$</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$C_2$</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$C_3$</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$C_4$</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$C_5$</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Weight = 14

**Randomized polynomial time (RP).** Polynomial time extended to allow calls to `random()` call in unit time.

- Polynomial algorithm $A$ with one-sided error:
  - if $x$ is YES instance: $Pr[A(x) = YES] \geq \frac{1}{2}$
  - if $x$ is NO instance: $Pr[A(x) = YES] = 0$
- Fundamental open question: does $P = RP$?

Johnson's Algorithm: Flip a coin, and set each variable true with probability $\frac{1}{2}$, independently for each variable.

**Theorem:** The "dumb" algorithm is a 2-approximation for MAX-SAT.
Maximum Satisfiability: Johnson’s Algorithm

Proof: Consider random variable \( Y_j = \begin{cases} 1 & \text{if clause } C_j \text{ is satisfied} \\ 0 & \text{otherwise.} \end{cases} \)

Let \( W = \sum_{j=1}^{m} w_j Y_j. \)

Let \( \text{OPT} = \text{weight of the optimal assignment.} \)

Let \( \ell_j \) be the number of distinct literals in clause \( C_j. \)

\[
E[W] = \sum_{j=1}^{m} w_j E[Y_j] \\
= \sum_{j=1}^{m} w_j \Pr[\text{clause } C_j \text{ is satisfied}] \\
= \sum_{j=1}^{m} w_j (1 - \left(\frac{1}{2}\right)^{\ell_j}) \\
\geq \frac{1}{2} \sum_{j=1}^{m} w_j \\
\geq \frac{1}{2} \text{OPT.}
\]

weights are \( \geq 0 \)

Maximum Satisfiability: Randomized Rounding

Idea 1. Used biased coin flips, not 50-50.

\[
z_j = \begin{cases} 1 & \text{if clause } C_j \text{ is satisfied} \\ 0 & \text{otherwise.} \end{cases} \quad y_j = \begin{cases} 1 & x_i \text{ is true} \\ 0 & \text{otherwise.} \end{cases}
\]

\( P_j = \text{indices of variables that occur un-negated in clause } C_j \)
\( N_j = \text{indices of variables that occur negated in clause } C_j \)

\[
\text{(LP)} \quad \max \sum_{j=1}^{m} w_j z_j \\
\text{s.t. } \sum_{i \in P_j} y_i + \sum_{i \in N_j} (1 - y_i) \geq z_j \\
0 \leq z_j \leq 1
\]

Theorem (Goemans-Williamson, 1994). The algorithm is an \( \frac{1}{e} \approx 0.582 \)-approximation algorithm for MAX-SAT.

Maximum Satisfiability: Randomized Rounding

Fact 1. For any nonnegative \( a_1, \ldots, a_k \), the geometric mean is \( \leq \) the arithmetic mean.

\[
k \sqrt[k]{a_1 a_2 \cdots a_k} \leq \frac{1}{k} (a_1 + a_2 + \cdots + a_k)
\]

Theorem (Goemans-Williamson, 1994). The algorithm is an \( \frac{1}{e} \approx 0.582 \)-approximation algorithm for MAX-SAT.

Proof. Consider an arbitrary clause \( C_j \).

\[
\Pr[\text{clause } C_j \text{ is not satisfied}] = \prod_{i \in P_j} (1 - y_i) \prod_{i \in N_j} y_i \\
\leq \left[ \frac{1}{\ell_j} \left( \sum_{i \in P_j} (1 - y_i) + \sum_{i \in N_j} y_i \right) \right]^{\ell_j} \\
= \left[ 1 - \frac{1}{\ell_j} \left( \sum_{i \in P_j} y_i + \sum_{i \in N_j} (1 - y_i) \right) \right]^{\ell_j} \\
\leq \left( 1 - \frac{z_j}{\ell_j} \right)^{\ell_j}
\]

geometric-arithmetic mean

LP constraint
Maximum Satisfiability: Randomized Rounding

Proof (continued).

\[ \Pr[\text{clause } C_j \text{ is satisfied}] \geq 1 - \left( 1 - \frac{z^*}{x} \right)^k \]

1 - \left( 1 - \frac{z^*}{x} \right)^k \text{ is concave}

(1-1/x)^k converges to \( e^{-1} \)

\[ f(z^*) = 1 - \left( 1 - \frac{z^*}{x} \right)^k \]

\[ f(z^*) = 1 - \left( 1 - \frac{z^*}{x} \right)^k \]

Optimize

\[ \min \left\{ \sum w_j \Pr[\text{clause } C_j \text{ is satisfied}] \right\} \]

\[ \geq (1 - \frac{1}{e}) \cdot \text{OPT} \]

Corollary. If all clauses have length at most \( k \)

\[ E[W] \geq \left[ 1 - (1 - \frac{1}{k})^k \right] \cdot \text{OPT}. \]

Maximum Satisfiability: Best of Two

Observation. Two approximation algorithms are complementary.

- Johnson’s algorithm works best when clauses are long.
- LP rounding algorithm works best when clauses are short.

How can we exploit this?

- Run both algorithms and output better of two.
- Re-analyze to get 4/3-approximation algorithm.
- Better performance than either algorithm individually!

Best-of-Two (\( C_1, C_2, \ldots, C_m \))

\[ (x^1, W^1) \leftarrow \text{Johnson} (C_1, \ldots, C_m) \]

\[ (x^2, W^2) \leftarrow \text{LPround} (C_1, \ldots, C_m) \]

IF \( W^1 > W^2 \)

RETURN \( x^1 \)
ELSE

RETURN \( x^2 \)

Theorem (Goemans-Williamson, 1994). The Best-of-Two algorithm is a 4/3-approximation algorithm for MAX-SAT.

Proof.

\[ E \left[ \max(W^1, W^2) \right] \geq E \left[ \frac{3}{4} W^1 + \frac{1}{4} W^2 \right] \]

\[ = \frac{1}{2} \sum w_j \Pr[\text{clause } C_j \text{ is satisfied by Alg } 1] + \frac{1}{2} \sum w_j \Pr[\text{clause } C_j \text{ is satisfied by Alg } 2] \]

\[ \geq \frac{1}{2} \sum w_j \left[ (1 - \left( \frac{1}{x} \right)^k) + \left( 1 - \left( 1 - \frac{1}{x} \right)^k \right) z^* \right] \]

\[ \geq \frac{1}{4} \sum w_j \left( \frac{3}{4} z^* \right) \]

\[ \geq \frac{3}{4} \text{OPT}_{LP} \]

\[ \geq \frac{3}{4} \text{OPT}. \]
Maximum Satisfiability: Best of Two

Lemma. For any integer \( \ell \geq 1 \), \( (1 - (\frac{1}{\ell})^\ell) + \left[ 1 - (1 - \frac{1}{\ell})^\ell \right] z_j^* \geq \frac{1}{2} z_j^* \).

Proof.
- Case 1 (\( \ell = 1 \)): \( \frac{1}{2} + 1 z_j^* \geq \frac{3}{2} z_j^* \).
- Case 2 (\( \ell = 2 \)): \( \frac{3}{4} + \frac{3}{4} z_j^* \geq \frac{3}{2} z_j^* \).
- Case 3 (\( \ell \geq 3 \)): \( (1 - (\frac{1}{\ell})^\ell) + \left[ 1 - (1 - \frac{1}{\ell})^\ell \right] z_j^* \geq \frac{3}{2} z_j^* \).

Maximum Satisfiability: State of the Art

Observation. Can’t get better than 4/3-approximation using our LP.
- If all weights = 1, \( \text{OPT}_{LP} = 4 \) but \( \text{OPT} = 3 \).

Lower bound.
- Unless \( P = NP \), can’t do better than \( 8 / 7 \approx 1.142 \).

Semi-definite programming.
- 1.275-approximation algorithm.
- 1.2-approximation algorithm if certain conjecture is true.

Open research problem.
- 4 / 3 - approximation algorithm without solving LP or SDP.

Appendix: Proof of LP Duality Theorem

LP Duality Theorem. For \( A \in \mathbb{R}^{m \times n} \), \( b \in \mathbb{R}^m \), \( c \in \mathbb{R}^n \), if (P) and (D) are nonempty then \( \max = \min \).

\[ \begin{align*}
(P) \quad & \max c^T x \\
& \text{s.t. } A x \leq b \\
& \quad \quad x \geq 0
\end{align*} \]

\[ \begin{align*}
(D) \quad & \min y^T b \\
& \text{s.t. } y^T A \geq c \\
& \quad \quad y \geq 0
\end{align*} \]

Proof (easy).
- Suppose \( x \in \mathbb{R}^m \) is feasible for (P) and \( y \in \mathbb{R}^n \) is feasible for (D).
  - \( x \geq 0 \), \( y^T A \geq c \) \( \Rightarrow \) \( y^T A x \geq c^T x \).
  - \( y \geq 0 \), \( A x \leq b \) \( \Rightarrow \) \( y^T A x \leq y^T b \).
  - combining two inequalities: \( c^T x \leq y^T A x \leq y^T b \).
**Weierstrass’ Theorem.** Let $X$ be a compact set, and let $f(x)$ be a continuous function on $X$. Then $\min \{ f(x) : x \in X \}$ exists.

**Lemma 1.** Let $X \subseteq \mathbb{R}^m$ be a nonempty closed convex set, and let $y \in X$. Then there exists $x^* \in X$ with minimum distance from $y$. Moreover, for all $x \in X$ we have $(y - x^*)^T (x - x^*) \leq 0$.

**Proof.** (existence)
- Define $f(x) = ||y - x||$.
- Want to apply Weierstrass:
  - $f$ is continuous
  - $X$ closed, but maybe not bounded
- $X' = \{ x \in X : ||y - x|| \leq ||y - x'|| \}$
  - is closed and bounded.
- $\min \{ f(x) : x \in X \} = \min \{ f(x) : x \in X' \}$

**Proof.** (moreover)
- $x^*$ min distance
- $||y - x^*|| \leq ||y - x||$ for all $x \in X$.
- By convexity: if $x \in X$, then $x^* + \varepsilon (x - x^*) \in X$ for all $0 \leq \varepsilon \leq 1$.
- $||y - x^*||^2 \leq ||y - x^* - \varepsilon (x - x^*)||^2$
  - $= ||y - x'||^2 + \varepsilon^2 ||(x - x^*)||^2 - 2 \varepsilon (y - x*)^T (x - x^*)$
- Thus, $(y - x^*)^T (x - x^*) \leq \frac{1}{2} \varepsilon ||(x - x^*)||^2$.
- Letting $\varepsilon \to 0^+$, we obtain the desired result.

**Separating Hyperplane Theorem**

**Separating Hyperplane Theorem.** Let $X \subseteq \mathbb{R}^m$ be a nonempty closed convex set, and let $y \in X$. Then there exists a hyperplane $H = \{ x \in \mathbb{R}^m : a^T x = \alpha \}$ where $a \in \mathbb{R}^m, \alpha \in \mathbb{R}$ that separates $y$ from $X$.
- $a^T x \leq \alpha$ for all $x \in X$.
- $a^T y > \alpha$.

**Proof.**
- Let $x^*$ be closest point in $X$ to $y$.
  - $L1 \Rightarrow (y - x^*)^T (x - x^*) \leq 0$ for all $x \in X$
- Choose $a = y - x^*$ and $\alpha = a^T x^*$.
  - $a^T y = a^T (a + x^*) = ||a||^2 + \alpha > \alpha$
  - If $x \in X$, then $a^T (x - x^*) \leq 0$
  - $\Rightarrow a^T x \leq a^T x^* = \alpha$

**Fundamental Theorem of Linear Inequalities**

**Farkas’ Theorem** (Farkas 1894, Minkowski 1896). For $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ exactly one of the following two systems holds.

(I) $\exists x \in \mathbb{R}^n$
- s.t. $Ax = b$
- $x \geq 0$

(II) $\exists y \in \mathbb{R}^m$
- s.t. $y^T A \leq 0$
- $y^T b > 0$

**Proof (not both).** Suppose $x$ satisfies (I) and $y$ satisfies (II).
- Then $0 < y^T b = y^T Ax \leq 0$, a contradiction.

**Proof (at least one).** Suppose (I) infeasible. We will show (II) feasible.
- Consider $S = \{ Ax : x \geq 0 \}$ so that $S$ closed, convex, $b \in S$.
  - there exists hyperplane $y \in \mathbb{R}^m, \alpha \in \mathbb{R}$ separating $b$ from $S$:
    - $y^T b > \alpha, y^T s \leq \alpha$ for all $s \in S$.
    - $0 \in S \Rightarrow \alpha \geq 0 \Rightarrow y^T b > 0$
  - $y^T A x \leq \alpha$ for all $x \geq 0 \Rightarrow y^T A \leq 0$ since $x$ can be arbitrarily large.
Corollary. For \( A \in \mathbb{R}^{m \times n} \), \( b \in \mathbb{R}^m \) exactly one of the following two systems holds.

\[
\begin{align*}
(I) & \quad \exists x \in \mathbb{R}^n \\
& \quad \text{s.t. } Ax \leq b \\
& \quad x \geq 0
\end{align*}
\]

\[
\begin{align*}
(II) & \quad \exists y \in \mathbb{R}^m \\
& \quad \text{s.t. } y^T A \geq 0 \\
& \quad y^T b < 0 \\
& \quad y \geq 0
\end{align*}
\]

Proof.
- Define \( A' \in \mathbb{R}^{m \times (m+n)} = [A \mid I] \), \( x' = [x \mid s] \), where \( s \in \mathbb{R}^m \).
- Farkas’ Theorem to \( A', b' \): exactly one of (I’) and (II’) is feasible.

\[
\begin{align*}
(I') & \quad \exists x \in \mathbb{R}^n, s \in \mathbb{R}^m \\
& \quad \text{s.t. } Ax + Is = b \\
& \quad x, s \geq 0
\end{align*}
\]

\[
\begin{align*}
(II') & \quad \exists y \in \mathbb{R}^m \\
& \quad \text{s.t. } y^T A \leq 0 \\
& \quad I y \leq 0 \\
& \quad y^T b > 0
\end{align*}
\]

- (I’) equivalent to (I), (II’) equivalent to (II).

LP Strong Duality

LP Duality Theorem. For \( A \in \mathbb{R}^{m \times n} \), \( b \in \mathbb{R}^m \), \( c \in \mathbb{R}^n \), if (P) and (D) are nonempty then \( \max = \min \).

Proof (\( \max \leq \min \)). Weak LP duality.
Proof (\( \min \leq \max \)). Suppose \( \max < \alpha \). We show \( \min < \alpha \).

\[
\begin{align*}
\text{(I)} & \quad \exists x \in \mathbb{R}^n \\
& \quad \text{s.t. } Ax \leq b \\
& \quad x \geq 0
\end{align*}
\]

\[
\begin{align*}
\text{(II)} & \quad \exists y \in \mathbb{R}^m, z \in \mathbb{R} \\
& \quad \text{s.t. } y^T A - cz \geq 0 \\
& \quad y^T b - \alpha z < 0 \\
& \quad y, z \geq 0
\end{align*}
\]

- By definition of \( \alpha \), (I) infeasible \( \Rightarrow \) (II) feasible by Farkas Corollary.

Case 1: \( z = 0 \).
- Then, \( \{ y \in \mathbb{R}^m : y^T A \geq 0, y^T b < 0, y \geq 0 \} \) is feasible.
- Farkas Corollary \( \Rightarrow \) \( \{ x \in \mathbb{R}^n : Ax \leq b, x \geq 0 \} \) is infeasible.
- Contradiction since by assumption (P) is nonempty.

Case 2: \( z > 0 \).
- Scale \( y, z \) so that \( y \) satisfies (II) and \( z = 1 \).
- Resulting \( y \) feasible to (D) and \( y^T b < \alpha \).