CME 307 / MS&E 311: Optimization

LP modeling and solution techniques

Professor Udell

Management Science and Engineering
Stanford

September 22, 2024

Course survey

You're interested in

- duality
- modeling real-world problems
- hyperparameter and blackbox optimization
- fairness and ethics in optimization
- ...

Outline

Linear programming: standard form

standard form linear program (LP)

$$\begin{array}{lll} \text{minimize} & c^T x & \\ \text{subject to} & Ax = b : & \quad \text{dual } y \\ & x \geq 0 & \end{array}$$

optimal value $p^\star$, solution $x^\star$ (if it exists)

- any x with $Ax = b$ and $x \geq 0$ is called a **feasible point**
- if problem is infeasible, we say $p^\star = \infty$
- $p^\star$ can be finite or $-\infty$

Linear programming: standard form

standard form linear program (LP)

$$\begin{array}{lll} \text{minimize} & c^T x & \\ \text{subject to} & Ax = b : & \quad \text{dual } y \\ & x \geq 0 & \end{array}$$

optimal value $p^\star$, solution $x^\star$ (if it exists)

- any x with $Ax = b$ and $x \geq 0$ is called a **feasible point**
- if problem is infeasible, we say $p^\star = \infty$
- $p^\star$ can be finite or $-\infty$

Q: if $p^\star = -\infty$, does a solution exist? is it unique?
what about $p^\star = \infty$?

Linear programming: standard form

standard form linear program (LP)

$$\begin{array}{lll} \text{minimize} & c^T x & \\ \text{subject to} & Ax = b : & \quad \text{dual } y \\ & x \geq 0 & \end{array}$$

optimal value $p^\star$, solution $x^\star$ (if it exists)

- any x with $Ax = b$ and $x \geq 0$ is called a **feasible point**
- if problem is infeasible, we say $p^\star = \infty$
- $p^\star$ can be finite or $-\infty$

Q: if $p^\star = -\infty$, does a solution exist? is it unique?
what about $p^\star = \infty$?
henceforth assume $A \in \mathbf{R}^{m \times n}$ has full row rank m
Q: why? how to check?

Linear programming: standard form

standard form linear program (LP)

$$\begin{array}{lll} \text{minimize} & c^T x & \\ \text{subject to} & Ax = b : & \quad \text{dual } y \\ & x \geq 0 & \end{array}$$

optimal value $p^\star$, solution $x^\star$ (if it exists)

- any x with $Ax = b$ and $x \geq 0$ is called a **feasible point**
- if problem is infeasible, we say $p^\star = \infty$
- $p^\star$ can be finite or $-\infty$

Q: if $p^\star = -\infty$, does a solution exist? is it unique?
what about $p^\star = \infty$?
henceforth assume $A \in \mathbf{R}^{m \times n}$ has full row rank m
Q: why? how to check?
A: otherwise infeasible or redundant rows; use gaussian elimination to check and remove

LP example: diet problem

- x_j servings of food j, $j = 1, \ldots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \ldots, m$

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

LP example: diet problem

- x_j servings of food j, $j = 1, \ldots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \ldots, m$

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

extensions:

- foods come from recipes? $x = By$

LP example: diet problem

- x_j servings of food j, $j = 1, \ldots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \ldots, m$

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

extensions:

- foods come from recipes?
- ensure diversity in diet?

LP example: diet problem

- x_j servings of food j, $j = 1, \ldots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \ldots, m$

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

extensions:

- foods come from recipes?
- ensure diversity in diet? $y \leq u$

LP example: diet problem

- x_j servings of food j, $j = 1, \dots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \dots, m$

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

extensions:

- foods come from recipes?
- ensure diversity in diet? $y \leq u$
- ranges of nutrients?

LP example: diet problem

- x_j servings of food j, $j = 1, \dots, n$
- c_j cost per serving
- a_{ij} amount of nutrient i in food j
- b_i required amount of nutrient i, $i = 1, \dots, m$

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

extensions:

- foods come from recipes?
- ensure diversity in diet? $y \leq u$
- ranges of nutrients? $Ax + s = b, \; l \leq s \leq u$

Geometry of LP

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

the **feasible set** is the set of points x that satisfy all constraints.

Geometry of LP

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

the **feasible set** is the set of points x that satisfy all constraints.

interpretation: conic hull

- define the **cone** generated by $A = [a_1, \ldots a_n]$:

$$\{Ax \mid x \geq 0\} = \{\sum_{i=1}^{n} a_i x_i \mid x \geq 0\} = (a_1, \ldots, a_n)$$

Geometry of LP

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

the **feasible set** is the set of points x that satisfy all constraints.

interpretation: conic hull

- define the **cone** generated by $A = [a_1, \dots a_n]$:

$$\{Ax \mid x \geq 0\} = \{\sum_{i=1}^{n} a_i x_i \mid x \geq 0\} = (a_1, \dots, a_n)$$

- LP is feasible if $b \in (a_1, \dots, a_n)$

Geometry of LP

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

the **feasible set** is the set of points x that satisfy all constraints.

interpretation: conic hull

- define the **cone** generated by $A = [a_1, \ldots a_n]$:

$$\{Ax \mid x \geq 0\} = \{\sum_{i=1}^{n} a_i x_i \mid x \geq 0\} = (a_1, \ldots, a_n)$$

- LP is feasible if $b \in (a_1, \ldots, a_n)$

interpretation: intersection of hyperplane and halfspaces

- $Ax = b$ defines a **hyperplane**

Geometry of LP

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

the **feasible set** is the set of points x that satisfy all constraints.

interpretation: conic hull

- define the **cone** generated by $A = [a_1, \ldots a_n]$:

$$\{Ax \mid x \geq 0\} = \{\sum_{i=1}^{n} a_i x_i \mid x \geq 0\} = (a_1, \ldots, a_n)$$

- LP is feasible if $b \in (a_1, \ldots, a_n)$

interpretation: intersection of hyperplane and halfspaces

- $Ax = b$ defines a **hyperplane** (dimension? normal vectors?)
- define a **halfspace** $\{x \mid a^T x \geq b\}$

Geometry of LP

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

the **feasible set** is the set of points x that satisfy all constraints.

interpretation: conic hull

- define the **cone** generated by $A = [a_1, \ldots a_n]$:

$$\{Ax \mid x \geq 0\} = \{\sum_{i=1}^{n} a_i x_i \mid x \geq 0\} = (a_1, \ldots, a_n)$$

- LP is feasible if $b \in (a_1, \ldots, a_n)$

interpretation: intersection of hyperplane and halfspaces

- $Ax = b$ defines a **hyperplane** (dimension? normal vectors?)
- define a **halfspace** $\{x \mid a^T x \geq b\}$
- the **positive orthant** $x \geq 0$ is an intersection of halfspaces

Geometry of LP

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

the **feasible set** is the set of points x that satisfy all constraints.

interpretation: conic hull

- define the **cone** generated by $A = [a_1, \dots a_n]$:

$$\{Ax \mid x \geq 0\} = \{\sum_{i=1}^{n} a_i x_i \mid x \geq 0\} = (a_1, \dots, a_n)$$

- LP is feasible if $b \in (a_1, \dots, a_n)$

interpretation: intersection of hyperplane and halfspaces

- $Ax = b$ defines a **hyperplane** (dimension? normal vectors?)
- define a **halfspace** $\{x \mid a^T x \geq b\}$
- the **positive orthant** $x \geq 0$ is an intersection of halfspaces

Geometry of LP: convexity

- define **convex combination** of x, $y \in \mathbf{R}^n$: $\theta x + (1-\theta)y$ for $\theta \in [0,1]$

Geometry of LP: convexity

- define **convex combination** of x, $y \in \mathbf{R}^n$: $\theta x + (1-\theta)y$ for $\theta \in [0,1]$
- define **convex set**: C is convex if for any $x, y \in C$,

$$\theta x + (1-\theta)y \in C, \qquad \theta \in [0,1]$$

- define the **convex hull** of a set S:

$$\mathbf{conv}(S) = \left\{ \sum_{i=1}^{k} \theta_i x_i \mid x_i \in S,\ \theta_i \geq 0,\ \sum_{i=1}^{k} \theta_i = 1 \right\}$$

- define **polytope**: the convex hull of a finite set: $\mathbf{conv}(\{x_1, \ldots, x_k\})$

some useful convex sets:

Geometry of LP: convexity

- define **convex combination** of x, $y \in \mathbf{R}^n$: $\theta x + (1-\theta)y$ for $\theta \in [0,1]$
- define **convex set**: C is convex if for any $x, y \in C$,

$$\theta x + (1-\theta)y \in C, \qquad \theta \in [0,1]$$

- define the **convex hull** of a set S:

$$\mathbf{conv}(S) = \left\{ \sum_{i=1}^{k} \theta_i x_i \mid x_i \in S,\ \theta_i \geq 0,\ \sum_{i=1}^{k} \theta_i = 1 \right\}$$

- define **polytope**: the convex hull of a finite set: $\mathbf{conv}(\{x_1, \ldots, x_k\})$

some useful convex sets:

- a hyperplane is convex

Geometry of LP: convexity

- define **convex combination** of x, $y \in \mathbf{R}^n$: $\theta x + (1-\theta)y$ for $\theta \in [0,1]$
- define **convex set**: C is convex if for any $x, y \in C$,

$$\theta x + (1-\theta)y \in C, \qquad \theta \in [0,1]$$

- define the **convex hull** of a set S:

$$\mathbf{conv}(S) = \left\{ \sum_{i=1}^{k} \theta_i x_i \mid x_i \in S,\ \theta_i \geq 0,\ \sum_{i=1}^{k} \theta_i = 1 \right\}$$

- define **polytope**: the convex hull of a finite set: $\mathbf{conv}(\{x_1, \ldots, x_k\})$

some useful convex sets:

- a hyperplane is convex
- a halfspace is convex

Geometry of LP: convexity

- define **convex combination** of x, $y \in \mathbf{R}^n$: $\theta x + (1-\theta)y$ for $\theta \in [0,1]$
- define **convex set**: C is convex if for any $x, y \in C$,

$$\theta x + (1-\theta)y \in C, \qquad \theta \in [0,1]$$

- define the **convex hull** of a set S:

$$\mathbf{conv}(S) = \left\{ \sum_{i=1}^{k} \theta_i x_i \mid x_i \in S,\ \theta_i \geq 0,\ \sum_{i=1}^{k} \theta_i = 1 \right\}$$

- define **polytope**: the convex hull of a finite set: $\mathbf{conv}(\{x_1, \ldots, x_k\})$

some useful convex sets:

- a hyperplane is convex
- a halfspace is convex
- the intersection of convex sets is convex

Geometry of LP: convexity

- define **convex combination** of x, $y \in \mathbf{R}^n$: $\theta x + (1-\theta)y$ for $\theta \in [0,1]$
- define **convex set**: C is convex if for any $x, y \in C$,

$$\theta x + (1-\theta)y \in C, \qquad \theta \in [0,1]$$

- define the **convex hull** of a set S:

$$\mathbf{conv}(S) = \left\{ \sum_{i=1}^{k} \theta_i x_i \mid x_i \in S,\ \theta_i \geq 0,\ \sum_{i=1}^{k} \theta_i = 1 \right\}$$

- define **polytope**: the convex hull of a finite set: $\mathbf{conv}(\{x_1, \ldots, x_k\})$

some useful convex sets:

- a hyperplane is convex
- a halfspace is convex
- the intersection of convex sets is convex
- the feasible set $\{x : Ax = b,\ x \geq 0\}$ is convex

Outline

LP inequality form

another useful form for LP is **inequality form**

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

LP inequality form

another useful form for LP is **inequality form**

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

interpretation: halfspaces

- $a_i^T x \leq b_i$ defines a **halfspace**
- $Ax \leq b$ defines a **polyhedron**: intersection of halfspaces
- LP is feasible if polyhedron $\{x \mid Ax \leq b\}$ is nonempty

LP example: production planning

- x_i units of product i
- c_i cost per unit
- a_{ij} amount of resource j used by product i
- b_j amount of resource j available
- d_i demand for product i

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax \leq b \\
& 0 \leq x \leq d
\end{array}
$$

LP example: production planning

- x_i units of product i
- c_i cost per unit
- a_{ij} amount of resource j used by product i
- b_j amount of resource j available
- d_i demand for product i

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax \leq b \\
& 0 \leq x \leq d
\end{array}
$$

extensions:

- fixed cost for producing product i at all?

LP example: production planning

- x_i units of product i
- c_i cost per unit
- a_{ij} amount of resource j used by product i
- b_j amount of resource j available
- d_i demand for product i

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \\ & 0 \leq x \leq d \end{array}$$

extensions:

- fixed cost for producing product i at all?
  $c^T x + f^T z$, $z_i \in \{0, 1\}$, $x_i \leq M z_i$ for M large

LP inequality form to standard form

standard form to inequality form

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

LP inequality form to standard form

standard form to inequality form

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

LP inequality form to standard form

standard form to inequality form

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

inequality form to standard form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array} \quad \rightarrow$$

LP inequality form to standard form

standard form to inequality form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \qquad \rightarrow \qquad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \\ & Ax \geq b \\ -x \leq 0 & \end{array}$$

inequality form to standard form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array} \qquad \rightarrow \qquad \begin{array}{ll} \text{minimize} & c^T (x_+ - x_-) \\ \text{subject to} & A(x_+ - x_-) + s = b \\ & s, x_+, x_- \geq 0 \end{array}$$

Outline

Extreme points

define **extreme point**: $x \in \mathbf{R}^n$ is extreme in $C \subset \mathbf{R}^n$ if it cannot be written as a convex combination of other points in C: for $\theta \in [0, 1]$,

$$x \in C \quad \text{and} \quad x = \theta y + (1 - \theta) z \quad \Longrightarrow \quad x = y = z$$

Extreme points

define **extreme point**: $x \in \mathbf{R}^n$ is extreme in $C \subset \mathbf{R}^n$ if it cannot be written as a convex combination of other points in C: for $\theta \in [0, 1]$,

$$x \in C \quad \text{and} \quad x = \theta y + (1 - \theta)z \qquad \Longrightarrow \qquad x = y = z$$

fact: if $x^\star$ is the unique optimal solution, then $x^\star$ is an extreme point of the feasible set

Extreme points

define **extreme point**: $x \in \mathbf{R}^n$ is extreme in $C \subset \mathbf{R}^n$ if it cannot be written as a convex combination of other points in C: for $\theta \in [0,1]$,

$$x \in C \quad \text{and} \quad x = \theta y + (1-\theta)z \quad \Longrightarrow \quad x = y = z$$

fact: if $x^\star$ is the unique optimal solution, then $x^\star$ is an extreme point of the feasible set proof: suppose by way of contradiction that $x^\star$ is not extreme in the feasible set $F = \{x \mid Ax = b, x \geq 0\}$:

$$\begin{aligned} x^\star &= \theta y + (1-\theta)z \qquad \text{for} y, z \in C, \theta \in (0,1) \\ p^\star c^T x^\star &= \theta c^T y + (1-\theta) c^T z \end{aligned}$$

Extreme points

define **extreme point**: $x \in \mathbf{R}^n$ is extreme in $C \subset \mathbf{R}^n$ if it cannot be written as a convex combination of other points in C: for $\theta \in [0, 1]$,

$$x \in C \quad \text{and} \quad x = \theta y + (1-\theta)z \quad \Longrightarrow \quad x = y = z$$

fact: if $x^\star$ is the unique optimal solution, then $x^\star$ is an extreme point of the feasible set proof: suppose by way of contradiction that $x^\star$ is not extreme in the feasible set $F = \{x \mid Ax = b, x \geq 0\}$:

$$\begin{aligned} x^\star &= \theta y + (1-\theta)z \qquad \text{for} y, z \in C, \theta \in (0,1) \\ p^\star c^T x^\star &= \theta c^T y + (1-\theta) c^T z \end{aligned}$$

by optimality $p^\star \leq c^T y$ and $p^\star \leq c^T z$, so $p^\star = c^T y = c^T z = c^T x^\star$ and all three are optimal.

Extreme points

define **extreme point**: $x \in \mathbf{R}^n$ is extreme in $C \subset \mathbf{R}^n$ if it cannot be written as a convex combination of other points in C: for $\theta \in [0, 1]$,

$$x \in C \quad \text{and} \quad x = \theta y + (1-\theta) z \quad \implies \quad x = y = z$$

fact: if $x^\star$ is the unique optimal solution, then $x^\star$ is an extreme point of the feasible set proof: suppose by way of contradiction that $x^\star$ is not extreme in the feasible set $F = \{x \mid Ax = b, x \geq 0\}$:

$$\begin{aligned} x^\star &= \theta y + (1-\theta) z \qquad \text{for} y, z \in C, \theta \in (0,1) \\ p^\star c^T x^\star &= \theta c^T y + (1-\theta) c^T z \end{aligned}$$

by optimality $p^\star \leq c^T y$ and $p^\star \leq c^T z$, so $p^\star = c^T y = c^T z = c^T x^\star$ and all three are optimal.

Q: Example of a problem with a non-extreme solution? Does there always exist an extreme solution?

Vertices

define **vertex**: $x \in \mathbf{R}^n$ is a vertex of set $C \subset \mathbf{R}^n$ if for some vector $c \in \mathbf{R}^n$,

$$c^T x < c^T y \quad \forall y \in C \setminus \{x\}$$

Vertices

define **vertex**: $x \in \mathbf{R}^n$ is a vertex of set $C \subset \mathbf{R}^n$ if for some vector $c \in \mathbf{R}^n$,

$$c^T x < c^T y \quad \forall y \in C \setminus \{x\}$$

fact: x is a vertex of $C \implies x$ is an extreme point of C

Vertices

define **vertex**: $x \in \mathbf{R}^n$ is a vertex of set $C \subset \mathbf{R}^n$ if for some vector $c \in \mathbf{R}^n$,

$$c^T x < c^T y \quad \forall y \in C \setminus \{x\}$$

fact: x is a vertex of $C \implies x$ is an extreme point of C proof: similar to previous proof

Solution of LP is a vertex

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

fact: if a solution exists and the feasible set has a vertex, then some vertex of the feasible set is optimal

Solution of LP is a vertex

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

fact: if a solution exists and the feasible set has a vertex, then some vertex of the feasible set is optimal

cases: solution $x^\star$ is unique / not unique

Solution of LP is a vertex

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

fact: if a solution exists and the feasible set has a vertex, then some vertex of the feasible set is optimal

cases: solution $x^\star$ is unique / not unique

- unique: so $c^T x < c^T y$ for all $y \in F \setminus \{x\}$

Solution of LP is a vertex

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

fact: if a solution exists and the feasible set has a vertex, then some vertex of the feasible set is optimal

cases: solution $x^\star$ is unique / not unique

- unique: so $c^T x < c^T y$ for all $y \in F \setminus \{x\}$
- not unique: optimal set $X^\star = \{x : c^T x = c^T x^\star, x \in F\}$ is a polyhedron. It is not empty (a solution exists) and its complement is not empty (optimal value is bounded). So, it has at least one vertex. That vertex is also a vertex of F.

Basic feasible solution

define: $x \in \mathbf{R}^n$ is a **basic feasible solution** (BFS) if there is a set $S \subset \{1, \ldots, n\}$ of m columns so that A_S is invertible and

$$x_S = A_S^{-1} b, \qquad x_{\bar{S}} = 0, \qquad x \geq 0.$$

- $A_S \in \mathbf{R}^{m \times m}$, submatrix of A with columns in S, is invertible
- BFS $\iff$ extreme point
- two BFS with S, S' are neighbors if they share $m = 1$ columns: $|S \cap S'| = m - 1$

Basic feasible solution

define: $x \in \mathbf{R}^n$ is a **basic feasible solution** (BFS) if there is a set $S \subset \{1, \ldots, n\}$ of m columns so that A_S is invertible and

$$x_S = A_S^{-1} b, \qquad x_{\bar{S}} = 0, \qquad x \geq 0.$$

- $A_S \in \mathbf{R}^{m \times m}$, submatrix of A with columns in S, is invertible
- BFS $\iff$ extreme point
- two BFS with S, S' are neighbors if they share $m = 1$ columns: $|S \cap S'| = m - 1$

define: **active set** is set of nonzero variables in x

Basic feasible solution

define: $x \in \mathbf{R}^n$ is a **basic feasible solution** (BFS) if there is a set $S \subset \{1, \ldots, n\}$ of m columns so that A_S is invertible and

$$x_S = A_S^{-1} b, \qquad x_{\bar{S}} = 0, \qquad x \geq 0.$$

- $A_S \in \mathbf{R}^{m \times m}$, submatrix of A with columns in S, is invertible
- BFS $\iff$ extreme point
- two BFS with S, S' are neighbors if they share $m = 1$ columns: $|S \cap S'| = m - 1$

define: **active set** is set of nonzero variables in x

Q: how to find a BFS?

Basic feasible solution

define: $x \in \mathbf{R}^n$ is a **basic feasible solution** (BFS) if there is a set $S \subset \{1, \ldots, n\}$ of m columns so that A_S is invertible and

$$x_S = A_S^{-1} b, \qquad x_{\bar{S}} = 0, \qquad x \geq 0.$$

- $A_S \in \mathbf{R}^{m \times m}$, submatrix of A with columns in S, is invertible
- BFS $\iff$ extreme point
- two BFS with S, S' are neighbors if they share $m = 1$ columns: $|S \cap S'| = m - 1$

define: **active set** is set of nonzero variables in x

Q: how to find a BFS?
A: start at a feasible point; move in a **feasible direction** until you hit another constraint; continue until you reach a BFS

Extreme point $\iff$ vertex $\iff$ BFS

for any nonempty polyhedron $P = \{Ax \leq b\}$ in $\mathbf{R}^n$, the following are equivalent:

- x is an extreme point of P
- x is a vertex of P
- x is a BFS of P

Outline

Solving LPs

algorithms:

- enumerate all vertices and check
- fourier-motzkin elimination
- simplex method
- ellipsoid method
- interior point methods
- first-order methods
- ...

Solving LPs

algorithms:

- enumerate all vertices and check
- fourier-motzkin elimination
- simplex method
- ellipsoid method
- interior point methods
- first-order methods
- ...

remarks:

- enumeration and elimination are simple but not practical
- simplex was the first practical algorithm; still used today
- ellipsoid method is the first polynomial-time algorithm; not practical
- interior point methods are polynomial-time and practical
- first-order methods are practical and scale to large problems

Enumerate vertices of LP

can generate all extreme points of LP: for each $S \subseteq \{1, \ldots, n\}$ with $|S| = m$,

- $A_S \in \mathbf{R}^{m \times m}$, submatrix of A with columns in S, is invertible
- solve $A_S x_S = b$ for x_S and set $x_{\bar{S}} = 0$
- if $x_S \geq 0$, then x is a BFS
- evaluate objective $c^T x$

the best BFS is optimal!

Enumerate vertices of LP

can generate all extreme points of LP: for each $S \subseteq \{1, \dots, n\}$ with $|S| = m$,

- $A_S \in \mathbf{R}^{m \times m}$, submatrix of A with columns in S, is invertible
- solve $A_S x_S = b$ for x_S and set $x_{\bar{S}} = 0$
- if $x_S \geq 0$, then x is a BFS
- evaluate objective $c^T x$

the best BFS is optimal!

problem: how many BFSs are there?

Enumerate vertices of LP

can generate all extreme points of LP: for each $S \subseteq \{1, \ldots, n\}$ with $|S| = m$,

- $A_S \in \mathbf{R}^{m \times m}$, submatrix of A with columns in S, is invertible
- solve $A_S x_S = b$ for x_S and set $x_{\bar{S}} = 0$
- if $x_S \geq 0$, then x is a BFS
- evaluate objective $c^T x$

the best BFS is optimal!

problem: how many BFSs are there?
n choose m is $\binom{n}{m} = \frac{n!}{m!(n-m)!}$ ("exponentially many")

Simplex algorithm

basic idea: local search on the vertices of the feasible set

- start at BFS x and evaluate objective $c^T x$
- move to a neighboring BFS x' with better objective $c^T x'$
- repeat until no improvement possible

Simplex algorithm

basic idea: local search on the vertices of the feasible set

- start at BFS x and evaluate objective $c^T x$
- move to a neighboring BFS x' with better objective $c^T x'$
- repeat until no improvement possible

discuss in groups:

- how to find an initial BFS?
- how to find a neighboring BFS with better objective?
- how to prove optimality?

Finding an initial BFS

solve an auxiliary problem for which a BFS is known:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{m} z_i \\ \text{subject to} & Ax + Dz = b \\ & x, z \geq 0 \end{array}$$

where $D \in \mathbf{R}^{m \times m}$ is a diagonal matrix with $D_{ii} = \mathbf{sign}(b_i)$ for $i = 1, \ldots, m$.

- $x = 0$, $z = |b|$ is a BFS of this problem
- $(x, z) = (x, 0)$ is a BFS of this problem $\iff$ x is a BFS of the original problem

Find a better neighboring BFS

start with BFS x with active set S and turn on variable $j \notin S$

$$x^+ \leftarrow x + \theta d, \qquad \theta > 0$$

where $d_j = 1$ and $d_i = 0$ for $i \notin S \cup \{j\}$. need to solve for d_S.

Find a better neighboring BFS

start with BFS x with active set S and turn on variable $j \notin S$

$$x^+ \leftarrow x + \theta d, \qquad \theta > 0$$

where $d_j = 1$ and $d_i = 0$ for $i \notin S \cup \{j\}$. need to solve for d_S.

- need to stay feasible wrt equality constraints, so

$$Ax = b, \quad A(x + \theta d) = b, \implies Ad = 0$$

Find a better neighboring BFS

start with BFS x with active set S and turn on variable $j \notin S$

$$x^+ \leftarrow x + \theta d, \qquad \theta > 0$$

where $d_j = 1$ and $d_i = 0$ for $i \notin S \cup \{j\}$. need to solve for d_S.

- need to stay feasible wrt equality constraints, so

$$Ax = b, \quad A(x + \theta d) = b, \implies Ad = 0$$

- construct the j**th basic direction**

$$Ad = A_S d_S + A_j = 0 \implies d_S = -A_S^{-1} A_j$$

Find a better neighboring BFS

start with BFS x with active set S and turn on variable $j \notin S$

$$x^+ \leftarrow x + \theta d, \qquad \theta > 0$$

where $d_j = 1$ and $d_i = 0$ for $i \notin S \cup \{j\}$. need to solve for d_S.

- need to stay feasible wrt equality constraints, so

$$Ax = b, \quad A(x + \theta d) = b, \implies Ad = 0$$

- construct the j**th basic direction**

$$Ad = A_S d_S + A_j = 0 \implies d_S = -A_S^{-1} A_j$$

- if $x_S > 0$ is **non-degenerate**, then $\exists \theta > 0$ st $x^+ \geq 0$

Find a better neighboring BFS

start with BFS x with active set S and turn on variable $j \notin S$

$$x^+ \leftarrow x + \theta d, \qquad \theta > 0$$

where $d_j = 1$ and $d_i = 0$ for $i \notin S \cup \{j\}$. need to solve for d_S.

- need to stay feasible wrt equality constraints, so

$$Ax = b, \quad A(x + \theta d) = b, \implies Ad = 0$$

- construct the j**th basic direction**

$$Ad = A_S d_S + A_j = 0 \implies d_S = -A_S^{-1} A_j$$

- if $x_S > 0$ is **non-degenerate**, then $\exists \theta > 0$ st $x^+ \geq 0$
- how does objective change?

$$c^T x^+ = c^T x + \theta c^T d = c^T x + \theta c_j - \theta c_S^T A_S^{-1} A_j$$

Reduced cost

define **reduced cost** $\bar{c}_j = c_j - c_S^T A_S^{-1} A_j$, $j \notin S$

Reduced cost

define **reduced cost** $\bar{c}_j = c_j - c_S^T A_S^{-1} A_j$, $j \notin S$

fact:

- if $\bar{c} \geq 0$, x is optimal
- if x is optimal and nondegenerate ($x_S > 0$), then $\bar{c} \geq 0$

Outline

Why duality?

- certify optimality
  - turn $\forall$ into $\exists$
  - use dual lower bound to derive stopping conditions
- new algorithms based on the dual
  - solve dual, then recover primal solution

Duality notation

- inner product

$$y^T x = \langle y, x \rangle = y \cdot x = \sum_{i=1}^{n} y_i x_i$$

- conjugate

$$\langle y, Ax \rangle = \langle A^T y, x \rangle$$

Warmup: Farkas lemma

Theorem (Farkas lemma)

Given $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$, exactly one of the following is true:

- *there exists $x \in \mathbf{R}^n$ so that $Ax = b$ and $x \geq 0$*
- *there exists $y \in \mathbf{R}^m$ so that $A^T y \geq 0$ and $\langle b, y \rangle < 0$*

$\Longrightarrow$ can efficiently certify infeasibility of a linear program

Warmup: Farkas lemma

Theorem (Farkas lemma)

Given $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$, exactly one of the following is true:

- *there exists $x \in \mathbf{R}^n$ so that $Ax = b$ and $x \geq 0$*
- *there exists $y \in \mathbf{R}^m$ so that $A^T y \geq 0$ and $\langle b, y \rangle < 0$*

$\implies$ can efficiently certify infeasibility of a linear program

proof: suppose we have $x \in \mathbf{R}^n$ so that $Ax = b$ and $x \geq 0$.
then for any $y \in \mathbf{R}^m$,

$$\begin{aligned} 0 &= \langle y, b - Ax \rangle = \langle y, b \rangle - \langle A^T y, x \rangle \\ \langle y, b \rangle &= \langle A^T y, x \rangle \end{aligned}$$

so if $A^T y \geq 0$, then use $x \geq 0$ to conclude $\langle y, b \rangle \geq 0$.

Warmup: Farkas lemma

Theorem (Farkas lemma)

Given $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$, exactly one of the following is true:

- *there exists $x \in \mathbf{R}^n$ so that $Ax = b$ and $x \geq 0$*
- *there exists $y \in \mathbf{R}^m$ so that $A^T y \geq 0$ and $\langle b, y \rangle < 0$*

$\implies$ can efficiently certify infeasibility of a linear program

proof: suppose we have $x \in \mathbf{R}^n$ so that $Ax = b$ and $x \geq 0$.
then for any $y \in \mathbf{R}^m$,

$$\begin{aligned} 0 &= \langle y, b - Ax \rangle = \langle y, b \rangle - \langle A^T y, x \rangle \\ \langle y, b \rangle &= \langle A^T y, x \rangle \end{aligned}$$

so if $A^T y \geq 0$, then use $x \geq 0$ to conclude $\langle y, b \rangle \geq 0$.

(opposite direction is similar)

Lagrange duality

primal problem with solution $x^\star \in \mathbf{R}^n$, optimal value $p^\star$:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b : \qquad \text{dual } y \\ & x \geq 0 \end{array} \qquad (\mathcal{P})$$

if x is feasible, then $Ax = b$, so $\langle y, Ax - b\rangle = 0$ for $y \in \mathbf{R}^m$.

Lagrange duality

primal problem with solution $x^\star \in \mathbf{R}^n$, optimal value $p^\star$:

$$\begin{array}{lll} \text{minimize} & c^T x & \\ \text{subject to} & Ax = b : & \quad \text{dual } y \\ & x \geq 0 & \end{array} \tag{$\mathcal{P}$}$$

if x is feasible, then $Ax = b$, so $\langle y, Ax - b \rangle = 0$ for $y \in \mathbf{R}^m$.
define the **Lagrangian**

$$\mathcal{L}(x, y) \quad := \quad c^T x - \langle y, Ax - b \rangle$$

Lagrange duality

primal problem with solution $x^\star \in \mathbf{R}^n$, optimal value $p^\star$:

$$\begin{array}{lll} \text{minimize} & c^T x & \\ \text{subject to} & Ax = b : & \quad \text{dual } y \\ & x \geq 0 & \end{array} \qquad (\mathcal{P})$$

if x is feasible, then $Ax = b$, so $\langle y, Ax - b\rangle = 0$ for $y \in \mathbf{R}^m$.
define the **Lagrangian**

$$\begin{aligned} \mathcal{L}(x,y) &:= c^T x - \langle y, Ax - b\rangle \\ p^\star &= \inf_{x:Ax=b,\ x\geq 0} \mathcal{L}(x,y) \geq \inf_{x\geq 0} \mathcal{L}(x,y) \end{aligned}$$

Lagrange duality

primal problem with solution $x^\star \in \mathbf{R}^n$, optimal value $p^\star$:

$$\begin{array}{lll} \text{minimize} & c^T x & \\ \text{subject to} & Ax = b: & \quad \text{dual } y \\ & x \geq 0 & \end{array} \qquad (\mathcal{P})$$

if x is feasible, then $Ax = b$, so $\langle y, Ax - b \rangle = 0$ for $y \in \mathbf{R}^m$.
define the **Lagrangian**

$$\begin{aligned} \mathcal{L}(x, y) &:= c^T x - \langle y, Ax - b \rangle \\ p^\star &= \inf_{x: Ax = b,\ x \geq 0} \mathcal{L}(x, y) \geq \inf_{x \geq 0} \mathcal{L}(x, y) \\ &= \inf_{x \geq 0} c^T x + \langle y, b - Ax \rangle \\ &= \langle y, b \rangle + \inf_{x \geq 0} \left(c^T x - \langle A^T y, x \rangle \right) \\ &= \langle y, b \rangle + \inf_{x \geq 0} \left(\langle c - A^T y, x \rangle \right) \end{aligned}$$

Lagrange duality, ctd

we have a lower bound on $p^\star$ for any y, and a useful one whenever $c + A^T y = 0$.
maximize bound:

$$p^\star \geq \begin{array}{ll} \text{maximize} & \langle y, b \rangle \\ \text{subject to} & A^T y \leq c \\ \text{variable} & y \in \mathbf{R}^m \end{array}$$

define the **dual function**

$$g(y) = \begin{cases} \langle y, b \rangle & A^T y \leq c \\ -\infty & \textit{otherwise} \end{cases}$$

Lagrange duality

weak duality asserts that $p^\star \geq g(y)$ for all $y \in \mathbf{R}^m$.

$$\begin{aligned} p^\star &\geq g(y) \quad \forall y \in \mathbf{R}^m \\ &\geq \underbrace{\sup_y g(y)}_{\mathcal{D}} =: d^\star \end{aligned}$$

$p^\star \geq d^\star$ dual optimal value

Strong duality

Definition (Duality gap)

The **duality gap** for a primal-dual pair (x, y) is $c^T x - b^T y \geq 0$

by weak duality, duality gap is always nonnegative

Strong duality

Definition (Duality gap)

The **duality gap** for a primal-dual pair (x, y) is $c^T x - b^T y \geq 0$

by weak duality, duality gap is always nonnegative

Definition (Strong duality)

A primal-dual pair $(x^\star, y^\star)$ satisfies **strong duality** if

$$p^\star = d^\star \iff c^T x - b^T y = 0$$

Strong duality

Definition (Duality gap)

The **duality gap** for a primal-dual pair (x, y) is $c^T x - b^T y \geq 0$

by weak duality, duality gap is always nonnegative

Definition (Strong duality)

A primal-dual pair $(x^\star, y^\star)$ satisfies **strong duality** if

$$p^\star = d^\star \iff c^T x - b^T y = 0$$

strong duality holds

- for feasible LPs
- (later) for convex problems under **constraint qualification** aka **Slater's condition**. feasible region has an **interior point** x so that all inequality constraints hold strictly

strong duality fails if either primal or dual problem is infeasible or unbounded

Strong duality for LPs

primal and dual LP in standard form:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \qquad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

claim: if primal LP has a bounded feasible solution $x^\star$, then strong duality holds
i.e., dual LP has a bounded feasible solution $y^\star$ and $p^\star = d^\star$

Logic of strong duality proof

$x \in \mathbf{R}^n$ is optimal for the primal LP with optimal value $p^\star$

$\Downarrow$ (see next slide)

the following linear system has no solution

$$\begin{bmatrix} A & -b \\ c^T & -p^\star \end{bmatrix} \begin{bmatrix} x \\ \tau \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \qquad \text{or} \qquad \begin{bmatrix} A^T & c \\ -b^T & -p^\star \end{bmatrix}$$

$\Downarrow$ (Farkas lemma)

$$\begin{bmatrix} A^T & c \\ -b^T & -p^\star \end{bmatrix} \begin{bmatrix} -y \\ \sigma \end{bmatrix} \geq 0, \ \ \sigma > 0$$

$\Downarrow$

y/σ is dual feasible with optimal value as least as good as $p^\star$

Proof of strong duality for LPs

consider the following system with variables $x' \in \mathbf{R}^n$, $\tau \in \mathbf{R}$

$$Ax' - b\tau = 0, \quad c^T x' = p^\star \tau - 1, \quad (x', \tau) \geq 0$$

Proof of strong duality for LPs

consider the following system with variables $x' \in \mathbf{R}^n$, $\tau \in \mathbf{R}$

$$Ax' - b\tau = 0, \quad c^T x' = p^\star \tau - 1, \quad (x', \tau) \geq 0$$

claim: this system has no solution. pf by contradiction:

- if $\tau > 0$, then x'/τ is feasible for LP and $c^T x'/\tau < p^\star$
- if $\tau = 0$, then $x^\star + x'$ is feasible for LP and $c^T(x^\star + x') < p^\star$

Proof of strong duality for LPs

consider the following system with variables $x' \in \mathbf{R}^n$, $\tau \in \mathbf{R}$

$$Ax' - b\tau = 0, \quad c^T x' = p^\star \tau - 1, \quad (x', \tau) \geq 0$$

claim: this system has no solution. pf by contradiction:

- if $\tau > 0$, then x'/τ is feasible for LP and $c^T x'/\tau < p^\star$
- if $\tau = 0$, then $x^\star + x'$ is feasible for LP and $c^T(x^\star + x') < p^\star$

so use Farkas' lemma:

$$\bar{A}\bar{x} = \bar{b},\ \bar{x} \geq 0 \qquad \text{or} \qquad \bar{A}^T \bar{y} \geq 0, \quad \bar{b}^T \bar{y} < 0$$

Proof of strong duality for LPs

consider the following system with variables $x' \in \mathbf{R}^n$, $\tau \in \mathbf{R}$

$$Ax' - b\tau = 0, \quad c^T x' = p^\star \tau - 1, \quad (x', \tau) \geq 0$$

claim: this system has no solution. pf by contradiction:

- if $\tau > 0$, then x'/τ is feasible for LP and $c^T x'/\tau < p^\star$
- if $\tau = 0$, then $x^\star + x'$ is feasible for LP and $c^T(x^\star + x') < p^\star$

so use Farkas' lemma:

$$\bar{A}\bar{x} = \bar{b}, \ \bar{x} \geq 0 \qquad \text{or} \qquad \bar{A}^T \bar{y} \geq 0, \quad \bar{b}^T \bar{y} < 0$$

$$\begin{bmatrix} A & -b \\ c^T & -p^\star \end{bmatrix} \begin{bmatrix} x \\ \tau \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \qquad \text{or} \qquad \begin{bmatrix} A^T & c \\ -b^T & -p^\star \end{bmatrix} \begin{bmatrix} y \\ \sigma \end{bmatrix} \geq 0, \ \sigma > 0$$

Proof of strong duality for LPs

consider the following system with variables $x' \in \mathbf{R}^n$, $\tau \in \mathbf{R}$

$$Ax' - b\tau = 0, \quad c^T x' = p^\star \tau - 1, \quad (x', \tau) \geq 0$$

claim: this system has no solution. pf by contradiction:

- if $\tau > 0$, then x'/τ is feasible for LP and $c^T x'/\tau < p^\star$
- if $\tau = 0$, then $x^\star + x'$ is feasible for LP and $c^T(x^\star + x') < p^\star$

so use Farkas' lemma:

$$\begin{array}{ccc} \bar{A}\bar{x} = \bar{b}, \ \bar{x} \geq 0 & \text{or} & \bar{A}^T \bar{y} \geq 0, \quad \bar{b}^T \bar{y} < 0 \\ \begin{bmatrix} A & -b \\ c^T & -p^\star \end{bmatrix} \begin{bmatrix} x \\ \tau \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} & \text{or} & \begin{bmatrix} A^T & c \\ -b^T & -p^\star \end{bmatrix} \begin{bmatrix} y \\ \sigma \end{bmatrix} \geq 0, \ \sigma > 0 \end{array}$$

second system is feasible $\implies$ y/σ is dual feasible and optimal

Outline

Duality as stopping condition

want to optimize until **primal suboptimality** $p^\star - c^T x \geq 0$ or **dual suboptimality** $d^\star - b^T y \geq 0$ are small enough. how?

Duality as stopping condition

want to optimize until **primal suboptimality** $p^\star - c^T x \geq 0$ or **dual suboptimality** $d^\star - b^T y \geq 0$ are small enough. how?

duality gap $c^T x - b^T y \geq 0$ bounds both!

for x feasible, y dual feasible,

$$c^T x \geq c^T x^\star \geq b^T y^\star \geq b^T y$$

Duality as stopping condition

want to optimize until **primal suboptimality** $p^\star - c^T x \geq 0$ or **dual suboptimality** $d^\star - b^T y \geq 0$ are small enough. how?

duality gap $c^T x - b^T y \geq 0$ bounds both!

for x feasible, y dual feasible,

$$c^T x \geq c^T x^\star \geq b^T y^\star \geq b^T y$$

in practice: improve primal and dual iterates in parallel until duality gap is small enough

How to use duality to estimate sensitivity?

primal and dual LP in standard form:

$$p^\star = \begin{array}{ll} \text{min} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \qquad\qquad d^\star = \begin{array}{ll} \text{max} & b^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

optimal primal and dual solution $x^\star$, $y^\star$

perturbed problem: primal and dual LP in standard form:

$$\tilde{p}^\star = \begin{array}{ll} \text{min} & c^T x \\ \text{subject to} & Ax = b + \epsilon \\ & x \geq 0 \end{array} \qquad\qquad \tilde{d}^\star = \begin{array}{ll} \text{max} & (b+\epsilon)^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

How to use duality to estimate sensitivity?

primal and dual LP in standard form:

$$p^\star = \begin{array}{ll} \min & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \qquad d^\star = \begin{array}{ll} \max & b^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

optimal primal and dual solution $x^\star$, $y^\star$

perturbed problem: primal and dual LP in standard form:

$$\tilde{p}^\star = \begin{array}{ll} \min & c^T x \\ \text{subject to} & Ax = b + \epsilon \\ & x \geq 0 \end{array} \qquad \tilde{d}^\star = \begin{array}{ll} \max & (b + \epsilon)^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

$y^\star$ is feasible for perturbed problem, so

$$\tilde{p}^\star = \tilde{d}^\star \geq (b + \epsilon)^T y^\star = d^\star + \epsilon^T y^\star$$

Outline

Column / constraint generation

primal and dual LP, $A \in \mathbf{R}^{m \times n}$, $n \gg m$:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \qquad \leftrightarrow^{\text{dual}} \qquad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

Column / constraint generation

primal and dual LP, $A \in \mathbf{R}^{m \times n}$, $n \gg m$:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \quad \overset{\text{dual}}{\leftrightarrow} \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

approximate by using $S \subset \{1, \ldots, n\}$: fewer variables (primal) or constraints (dual)

$$\begin{array}{ll} \text{minimize} & c_s^T x_S \\ \text{subject to} & A_S x_S = b \\ & x_S \geq 0 \end{array} \quad \overset{\text{dual}}{\leftrightarrow} \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A_S^T y \leq c_S \end{array}$$

Column / constraint generation

primal and dual LP, $A \in \mathbf{R}^{m \times n}$, $n \gg m$:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \quad \overset{\text{dual}}{\leftrightarrow} \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

approximate by using $S \subset \{1, \ldots, n\}$: fewer variables (primal) or constraints (dual)

$$\begin{array}{ll} \text{minimize} & c_s^T x_S \\ \text{subject to} & A_S x_S = b \\ & x_S \geq 0 \end{array} \quad \overset{\text{dual}}{\leftrightarrow} \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A_S^T y \leq c_S \end{array}$$

if x_S is optimal for $\mathcal{P}_S$ and reduced cost $\bar{c} \geq 0$, then x_S is optimal for $\mathcal{P}$

if y is optimal for $\mathcal{D}_S$ and feasible for $\mathcal{D}$, then y is optimal for $\mathcal{D}$

Column / constraint generation

primal and dual LP, $A \in \mathbf{R}^{m \times n}$, $n \gg m$:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \quad \overset{\text{dual}}{\leftrightarrow} \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

approximate by using $S \subset \{1, \ldots, n\}$: fewer variables (primal) or constraints (dual)

$$\begin{array}{ll} \text{minimize} & c_s^T x_S \\ \text{subject to} & A_S x_S = b \\ & x_S \geq 0 \end{array} \quad \overset{\text{dual}}{\leftrightarrow} \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A_S^T y \leq c_S \end{array}$$

if x_S is optimal for $\mathcal{P}_S$ and reduced cost $\bar{c} \geq 0$, then x_S is optimal for $\mathcal{P}$

if y is optimal for $\mathcal{D}_S$ and feasible for $\mathcal{D}$, then y is optimal for $\mathcal{D}$

otherwise?

Column / constraint generation

primal and dual LP, $A \in \mathbf{R}^{m \times n}$, $n \gg m$:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \quad \stackrel{\text{dual}}{\leftrightarrow} \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

approximate by using $S \subset \{1, \ldots, n\}$: fewer variables (primal) or constraints (dual)

$$\begin{array}{ll} \text{minimize} & c_s^T x_S \\ \text{subject to} & A_S x_S = b \\ & x_S \geq 0 \end{array} \quad \stackrel{\text{dual}}{\leftrightarrow} \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A_S^T y \leq c_S \end{array}$$

if x_S is optimal for $\mathcal{P}_S$ and reduced cost $\bar{c} \geq 0$, then x_S is optimal for $\mathcal{P}$

if y is optimal for $\mathcal{D}_S$ and feasible for $\mathcal{D}$, then y is optimal for $\mathcal{D}$

otherwise? find i with $\bar{c}_i = c_i - c_S^T A_S^{-1} a_i < 0$ (primal) or $a_i^T y > c_i$ (dual) and add to S

Column / constraint generation

primal and dual LP, $A \in \mathbf{R}^{m \times n}$, $n \gg m$:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \quad \leftrightarrow^{\text{dual}} \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y \leq c \end{array}$$

approximate by using $S \subset \{1, \dots, n\}$: fewer variables (primal) or constraints (dual)

$$\begin{array}{ll} \text{minimize} & c_S^T x_S \\ \text{subject to} & A_S x_S = b \\ & x_S \geq 0 \end{array} \quad \leftrightarrow^{\text{dual}} \quad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A_S^T y \leq c_S \end{array}$$

if x_S is optimal for $\mathcal{P}_S$ and reduced cost $\bar{c} \geq 0$, then x_S is optimal for $\mathcal{P}$

if y is optimal for $\mathcal{D}_S$ and feasible for $\mathcal{D}$, then y is optimal for $\mathcal{D}$

otherwise? find i with $\bar{c}_i = c_i - c_S^T A_S^{-1} a_i < 0$ (primal) or $a_i^T y > c_i$ (dual) and add to S

- if dual constraints are all binding, $A_S^T y = c_S$, so these conditions are the same!
- active set of non-zero primal variables dual to active set of constraints that hold

Presolve

Often many constraints are redundant or can be simplified. example:

$$
\begin{array}{ll}
\text{minimize} & x_3 \\
\text{subject to} & x_1 = 1 \\
& x_2 = x_3 - x_1 \\
& x_3 - x_2 \geq 0 \\
& x \geq 0
\end{array}
$$

a good presolve can often reduce problem from 1000s of variables and constraints down to 10s!

reference: Achterberg, Tobias, et al. "Presolve reductions in mixed integer programming." INFORMS Journal on Computing 32.2 (2020): 473-506.

Outline

MILP solution vs LP solution

mixed-integer linear program (MILP):

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax + Bz = b \\ & x \geq 0, z \geq 0 \in \mathbb{Z} \end{array} \quad \rightarrow^{\text{relax}} \quad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax + Bz = b \\ & x, z \geq 0 \end{array}$$

MILP solution vs LP solution

mixed-integer linear program (MILP):

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax + Bz = b \\ & x \geq 0, z \geq 0 \in \mathbb{Z} \end{array} \quad \rightarrow^{\text{relax}} \quad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax + Bz = b \\ & x, z \geq 0 \end{array}$$

example:

$$\begin{array}{ll} \text{maximize} & x \\ \text{subject to} & x \leq z \\ & x \leq 1 - z \\ & x \geq 0, z \in \{0, 1\} \end{array}$$

MILP solution vs LP solution

mixed-integer linear program (MILP):

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax + Bz = b \\ & x \geq 0, z \geq 0 \in \mathbb{Z} \end{array} \quad \to^{\text{relax}} \quad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax + Bz = b \\ & x, z \geq 0 \end{array}$$

example:

$$\begin{array}{ll} \text{maximize} & x \\ \text{subject to} & x \leq z \\ & x \leq 1 - z \\ & x \geq 0, z \in \{0, 1\} \end{array}$$

draw picture: where is solution of MILP? of LP relaxation?

Branch and bound

given MILP with integer variable z in rectangle $R = (l, u)$, $l \leq z \leq u$, optimal value $p^\star(R)$, solution $z^\star(R)$

- solve LP relaxation to produce lower bound $\mathrm{LB}(R) \leq p^\star(R)$
- round z to nearest feasible integer z' to produce upper bound $\mathrm{UB}(R) \geq p^\star(R)$

Branch and bound

given MILP with integer variable z in rectangle $R = (l, u)$, $l \leq z \leq u$, optimal value $p^\star(R)$, solution $z^\star(R)$

- solve LP relaxation to produce lower bound $\mathrm{LB}(R) \leq p^\star(R)$
- round z to nearest feasible integer z' to produce upper bound $\mathrm{UB}(R) \geq p^\star(R)$

if $\mathrm{LB}(R) = \mathrm{UB}(R)$, then $p^\star(R) = \mathrm{LB}(R) = \mathrm{UB}(R)$ and we are done.

Branch and bound

given MILP with integer variable z in rectangle $R = (l, u)$, $l \leq z \leq u$, optimal value $p^\star(R)$, solution $z^\star(R)$

- solve LP relaxation to produce lower bound $\mathrm{LB}(R) \leq p^\star(R)$
- round z to nearest feasible integer z' to produce upper bound $\mathrm{UB}(R) \geq p^\star(R)$

if $\mathrm{LB}(R) = \mathrm{UB}(R)$, then $p^\star(R) = \mathrm{LB}(R) = \mathrm{UB}(R)$ and we are done.

otherwise, branch

- split R into two subrectangles $R_1 = (l_1, u_1)$, $R_2 = (l_2, u_2)$ so that $\mathbb{Z} \cap R = (\mathbb{Z} \cap R_1) \cup (\mathbb{Z} \cap R_2)$
- compute bounds $\mathrm{LB}(R_1)$, $\mathrm{UB}(R_1)$, $\mathrm{LB}(R_2)$, $\mathrm{UB}(R_2)$
- $R \subset R_1 \cup R_2$ so $\mathrm{LB}(R) \leq \min(\mathrm{LB}(R_1), \mathrm{LB}(R_2))$
- keep best solution so far $\mathrm{UB} \leftarrow \min(\mathrm{UB}, \mathrm{UB}(R_1), \mathrm{UB}(R_2))$
- prune: eliminate rectangle from consideration if $\mathrm{LB}(R) > \mathrm{UB}$

draw picture in 2D