CME 307 / MS&E 311: Optimization

LP modeling and solution techniques

Professor Udell

Management Science and Engineering
Stanford

January 29, 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)

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

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

Linear algebra review

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

- span of A:

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A:

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- how to compute basis for span and nullspace of A?

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- how to compute basis for span and nullspace of A?
  can use QR factorization or SVD

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: **nullspace**$(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- how to compute basis for span and nullspace of A?
  can use QR factorization or SVD
- how to solve $Ax = b$?

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- how to compute basis for span and nullspace of A?
  can use QR factorization or SVD
- how to solve $Ax = b$? factor-solve with QR or SVD; form normal equations $A^T Ax = A^T b$ and use CG; other Krylov methods like LSQR (positive definite), MINRES (indefinite), GMRES (general)

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- how to compute basis for span and nullspace of A?
  can use QR factorization or SVD
- how to solve $Ax = b$? factor-solve with QR or SVD; form normal equations $A^T Ax = A^T b$ and use CG; other Krylov methods like LSQR (positive definite), MINRES (indefinite), GMRES (general)
- solution to $Ax = b$ is unique if

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- how to compute basis for span and nullspace of A?
  can use QR factorization or SVD
- how to solve $Ax = b$? factor-solve with QR or SVD; form normal equations $A^T Ax = A^T b$ and use CG; other Krylov methods like LSQR (positive definite), MINRES (indefinite), GMRES (general)
- solution to $Ax = b$ is unique if $m = n$ and A is full rank

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\textbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- how to compute basis for span and nullspace of A?
  can use QR factorization or SVD
- how to solve $Ax = b$? factor-solve with QR or SVD; form normal equations $A^T Ax = A^T b$ and use CG; other Krylov methods like LSQR (positive definite), MINRES (indefinite), GMRES (general)
- solution to $Ax = b$ is unique if $m = n$ and A is full rank
- if $m < n$ and A is full rank
  - solution set is a hyperplane of dimension

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- how to compute basis for span and nullspace of A?
  can use QR factorization or SVD
- how to solve $Ax = b$? factor-solve with QR or SVD; form normal equations $A^T Ax = A^T b$ and use CG; other Krylov methods like LSQR (positive definite), MINRES (indefinite), GMRES (general)
- solution to $Ax = b$ is unique if $m = n$ and A is full rank
- if $m < n$ and A is full rank
  - solution set is a hyperplane of dimension $n - m$
  - null space of A, $\mathbf{nullspace}(A)$, is a hyperplane of dimension

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- how to compute basis for span and nullspace of A?
  can use QR factorization or SVD
- how to solve $Ax = b$? factor-solve with QR or SVD; form normal equations $A^T Ax = A^T b$ and use CG; other Krylov methods like LSQR (positive definite), MINRES (indefinite), GMRES (general)
- solution to $Ax = b$ is unique if $m = n$ and A is full rank
- if $m < n$ and A is full rank
  - solution set is a hyperplane of dimension $n - m$
  - null space of A, $\mathbf{nullspace}(A)$, is a hyperplane of dimension $n - m$
  - solution set is $\{x : Ax = b\} = \{x_0 + Vz\}$ where columns of $V \in \mathbf{R}^{n \times n-m}$ span $\mathbf{nullspace}(A)$

Linear algebra review

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

- span of A: $\mathrm{span}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$
- nullspace of A: $\mathbf{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\} \subseteq \mathbf{R}^n$
- how to compute basis for span and nullspace of A?
  can use QR factorization or SVD
- how to solve $Ax = b$? factor-solve with QR or SVD; form normal equations $A^T Ax = A^T b$ and use CG; other Krylov methods like LSQR (positive definite), MINRES (indefinite), GMRES (general)
- solution to $Ax = b$ is unique if $m = n$ and A is full rank
- if $m < n$ and A is full rank
  - solution set is a hyperplane of dimension $n - m$
  - null space of A, $\mathbf{nullspace}(A)$, is a hyperplane of dimension $n - m$
  - solution set is $\{x : Ax = b\} = \{x_0 + Vz\}$ where columns of $V \in \mathbf{R}^{n \times n-m}$ span $\mathbf{nullspace}(A)$

if these are confusing: review linear algebra and prove them all!

LP example: diet problem

- x_i servings of food i
- c_i cost per serving
- a_{ij} amount of nutrient j in food i
- b_j required amount of nutrient j

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

LP example: diet problem

- x_i servings of food i
- c_i cost per serving
- a_{ij} amount of nutrient j in food i
- b_j required amount of nutrient j

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

extensions:

- foods come from recipes? $x = By$

LP example: diet problem

- x_i servings of food i
- c_i cost per serving
- a_{ij} amount of nutrient j in food i
- b_j required amount of nutrient j

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

extensions:

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

LP example: diet problem

- x_i servings of food i
- c_i cost per serving
- a_{ij} amount of nutrient j in food i
- b_j required amount of nutrient j

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

extensions:

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

LP example: diet problem

- x_i servings of food i
- c_i cost per serving
- a_{ij} amount of nutrient j in food i
- b_j required amount of nutrient j

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

extensions:

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

LP example: diet problem

- x_i servings of food i
- c_i cost per serving
- a_{ij} amount of nutrient j in food i
- b_j required amount of nutrient j

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

extensions:

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

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

- interpretation: add up columns of A so they match b
- $Ax = b$ defines a **hyperplane**
- $x_i \geq 0$ is a **halfspace**
- $x \geq 0$ is the **positive orthant**

Geometry of LP: convexity

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

- define the **feasible set** $\{x : Ax = b,\ x \geq 0\}$

Geometry of LP: convexity

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

- define the **feasible set** $\{x : Ax = b,\ x \geq 0\}$
- define **convex set**: C is convex if for any $x, y \in C$,

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

Geometry of LP: convexity

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

- define the **feasible set** $\{x : Ax = b,\ x \geq 0\}$
- define **convex set**: C is convex if for any $x, y \in C$,

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

- fact: the feasible set is convex

Geometry of LP: convexity

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

- define the **feasible set** $\{x : Ax = b,\ x \geq 0\}$
- define **convex set**: C is convex if for any $x, y \in C$,

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

- fact: the feasible set is convex
- define **extreme point**: x is extreme in C if it cannot be written as a linear combination of other points in C:

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

Geometry of LP: convexity

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

- define the **feasible set** $\{x : Ax = b,\ x \geq 0\}$
- define **convex set**: C is convex if for any $x, y \in C$,

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

- fact: the feasible set is convex
- define **extreme point**: x is extreme in C if it cannot be written as a linear combination of other points in C:

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

- fact: if a solution exists, then some extreme point of the feasible set is optimal

Geometry of LP: polytopes

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

- define **polytope** P: convex hull of its extreme points $v_1, \ldots, v_k \in \mathbf{R}^n$:

$$P = \{x \in \mathbf{R}^n \mid x = \sum_{i=1}^{k} \theta_i v_i,\ \theta_i \geq 0,\ \sum_{i=1}^{k} \theta_i = 1\}$$

- if feasible set is bounded, it is a polytope
- prove: if a solution exists, then some extreme point of the feasible set is optimal

Outline

Let’s do some modeling!

- OptiMUS: https://optimus-solver.vercel.app/
- power systems: https://jump.dev/JuMP.jl/stable/tutorials/applications/power_systems/
- multicast routing: https://colab.research.google.com/drive/1iOn1T1Muh51KaA7mf7UIQOdhSFZhZyry?usp=sharing

Let's do some modeling!

- OptiMUS: https://optimus-solver.vercel.app/
- power systems: https://jump.dev/JuMP.jl/stable/tutorials/applications/power_systems/
- multicast routing: https://colab.research.google.com/drive/1iOn1T1Muh51KaA7mf7UIQOdhSFZhZyry?usp=sharing

practical solvers for MILP:

- Gurobi and COPT (cardinal optimizer) are the state-of-the-art commercial solvers
- GLPK is a free solver that is not as fast
- JuliaOpt/JuMP is a modeling language in Julia that calls solvers like Gurobi and is specialized for MILP applications
- CVX* (including CVXPY in python) are modeling languages that call solvers like Gurobi with good support for convex problems
- OptiMUS is a LLM-based modeling tool for MILP

Modeling challenges

model the following as standard form LPs:

1. **inequality constraints.** $Ax \leq b$
2. **free variable.** $x \in \mathbf{R}$
3. **absolute value.** constraint $|x| \leq 10$
4. **piecewise linear.** objective $\max(x_1, x_2)$
5. **assignment.** *e.g.*, every class is assigned exactly one classroom
6. **logic.** *e.g.*, class enrollment $\leq$ capacity of assigned room
7. **flow.** *e.g.*, the least cost way to ship an item from s to t

Modeling challenges

model the following as standard form LPs:

1. **inequality constraints.** $Ax \leq b$
2. **free variable.** $x \in \mathbf{R}$
3. **absolute value.** constraint $|x| \leq 10$
4. **piecewise linear.** objective $\max(x_1, x_2)$
5. **assignment.** *e.g.*, every class is assigned exactly one classroom
6. **logic.** *e.g.*, class enrollment $\leq$ capacity of assigned room
7. **flow.** *e.g.*, the least cost way to ship an item from s to t

(see chapter 1 of Bertsimas and Tsitsiklis for more details on 1–6. see https://colab.research.google.com/drive/1iOn1T1Muh51KaA7mf7UIQOdhSFZhZyry?usp=sharing for a detailed treatment of a flow problem.)

Use slack variables to represent inequality constraints

to represent the following problem in standard form,

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax \leq b \\ &x \geq 0 \end{aligned}$$

Use slack variables to represent inequality constraints

to represent the following problem in standard form,

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

introduce slack variable $s \in \mathbf{R}^m$: $Ax + s = b$, $s \geq 0 \iff Ax \leq b$

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

Split variable into parts to represent free variables

to represent the following problem in standard form,

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

Split variable into parts to represent free variables

to represent the following problem in standard form,

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

introduce positive variables x_+, x_- so $x = x_+ - x_-$:

$$\begin{array}{ll} \text{minimize} & c^T x_+ - c^T x_- \\ \text{subject to} & Ax_+ - Ax_- = b \\ & x_+, x_- \geq 0 \end{array}$$

Use epigraph variables to handle absolute value

to represent the following problem in standard form,

$$
\begin{array}{ll}
\text{minimize} & \|x\|_1 = \sum_i = 1^n |x_i| \\
\text{subject to} & Ax = b \\
x \geq 0 &
\end{array}
$$

Use epigraph variables to handle absolute value

to represent the following problem in standard form,

$$\begin{array}{ll} \text{minimize} & \|x\|_1 = \sum_i = 1^n |x_i| \\ \text{subject to} & Ax = b \\ x \geq 0 & \end{array}$$

introduce epigraph variable $t \in \mathbf{R}^n$ so $|x_i| \leq t_i$:

$$\begin{array}{ll} \text{minimize} & 1^T t = \sum_{i=1}^n t_i \geq \|x\|_1 \\ \text{subject to} & Ax = b \\ & -t \leq x \leq t \\ & x, t \geq 0 \end{array}$$

Use epigraph variables to handle absolute value

to represent the following problem in standard form,

$$
\begin{array}{ll}
\text{minimize} & \|x\|_1 = \sum_i = 1^n |x_i| \\
\text{subject to} & Ax = b \\
x \geq 0 &
\end{array}
$$

introduce epigraph variable $t \in \mathbf{R}^n$ so $|x_i| \leq t_i$:

$$
\begin{array}{ll}
\text{minimize} & 1^T t = \sum_{i=1}^n t_i \geq \|x\|_1 \\
\text{subject to} & Ax = b \\
& -t \leq x \leq t \\
& x, t \geq 0
\end{array}
$$

Q: Why does this work? For what kinds of functions can we use this trick?

Use binary variables to handle assignment

every class is assigned exactly one classroom:
define variable $X_{ij} \in \{0,1\}$ for each class $i = 1, \ldots, n$ and room $j = 1, \ldots, m$

$$X_{ij} = \begin{cases} 1 & \text{class } i \text{ is assigned to room } j \\ 0 & \text{otherwise} \end{cases}$$

Use binary variables to handle assignment

every class is assigned exactly one classroom:
define variable $X_{ij} \in \{0, 1\}$ for each class $i = 1, \ldots, n$ and room $j = 1, \ldots, m$

$$X_{ij} = \begin{cases} 1 & \text{class } i \text{ is assigned to room } j \\ 0 & \text{otherwise} \end{cases}$$

now solve the problem

$$\begin{array}{lll} \text{minimize} & \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij} & \\ \text{subject to} & \sum_{i=1}^{n} X_{ij} = 1, \ \forall j & \text{(every class assigned one room)} \\ & \sum_{j=1}^{m} X_{ij} = 1, \ \forall i & \text{(no more than one class per room)} \\ & X_{ij} \in \{0, 1\} & \text{(binary variables)} \end{array}$$

where C_{ij} is the cost of assigning class i to room j.

Use binary variables to handle logic

model class enrollment $n_i \leq$ capacity c_j of assigned room:
define variable $X_{ij} \in \{0, 1\}$ for each class $i = 1, \ldots, n$ and room $j = 1, \ldots, m$

$$X_{ij} = \begin{cases} 1 & \text{class } i \text{ is assigned to room } j \\ 0 & \text{otherwise} \end{cases}$$

Use binary variables to handle logic

model class enrollment $n_i \leq$ capacity c_j of assigned room:
define variable $X_{ij} \in \{0,1\}$ for each class $i = 1, \dots, n$ and room $j = 1, \dots, m$

$$X_{ij} = \begin{cases} 1 & \text{class } i \text{ is assigned to room } j \\ 0 & \text{otherwise} \end{cases}$$

solve the problem

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij} \\ \text{subject to} & \sum_{i=1}^{n} X_{ij} = 1, \ \forall j \quad \text{(every class assigned one room)} \\ & \sum_{j=1}^{m} X_{ij} = 1, \ \forall i \text{(no more than one class per room)} \\ & \sum_{i=1}^{n} p_i X_{ij} \leq c_j, \ \forall j \qquad \text{(capacity constraint)} \\ & X_{ij} \in \{0,1\} \qquad \text{(binary variables)} \end{array}$$

where C_{ij} is the cost of assigning class i to room j.

Use binary variables to handle logic

model class enrollment $n_i \leq$ capacity c_j of assigned room:
define variable $X_{ij} \in \{0,1\}$ for each class $i = 1,\dots,n$ and room $j = 1,\dots,m$

$$X_{ij} = \begin{cases} 1 & \text{class } i \text{ is assigned to room } j \\ 0 & \text{otherwise} \end{cases}$$

solve the problem

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}X_{ij} \\ \text{subject to} & \sum_{i=1}^{n} X_{ij} = 1,\ \forall j \quad \text{(every class assigned one room)} \\ & \sum_{j=1}^{m} X_{ij} = 1,\ \forall i \text{(no more than one class per room)} \\ & \sum_{i=1}^{n} p_i X_{ij} \leq c_j,\ \forall j \quad \text{(capacity constraint)} \\ & X_{ij} \in \{0,1\} \quad \text{(binary variables)} \end{array}$$

where C_{ij} is the cost of assigning class i to room j.

what if we want p to be a variable, too?

. . . or use a big-M relaxation!

model class enrollment $n_i \leq$ capacity c_j of assigned room:
define variable $X_{ij} \in \{0, 1\}$ for each class $i = 1, \ldots, n$ and room $j = 1, \ldots, m$

$$X_{ij} = \begin{cases} 1 & \text{class } i \text{ is assigned to room } j \\ 0 & \text{otherwise} \end{cases}$$

suppose M is a very large number.

. . . or use a big-M relaxation!

model class enrollment $n_i \leq$ capacity c_j of assigned room:
define variable $X_{ij} \in \{0, 1\}$ for each class $i = 1, \ldots, n$ and room $j = 1, \ldots, m$

$$X_{ij} = \begin{cases} 1 & \text{class } i \text{ is assigned to room } j \\ 0 & \text{otherwise} \end{cases}$$

suppose M is a very large number. solve the problem

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij} \\ \text{subject to} & \sum_{i=1}^{n} X_{ij} = 1, \ \forall j \quad \text{(every class assigned one room)} \\ & \sum_{j=1}^{m} X_{ij} = 1, \ \forall i \text{(no more than one class per room)} \\ & p_i \leq c_j + (1 - X_{ij}) M, \ \forall i, j \qquad \text{(capacity constraint)} \\ & X_{ij} \in \{0, 1\} \qquad \text{(binary variables)} \end{array}$$

where C_{ij} is the cost of assigning class i to room j.

Outline

LP inequality form

another common form for LP is **inequality form**

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

how to transform to standard form?

- inequality constraints $Ax \leq b$?

LP inequality form

another common form for LP is **inequality form**

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

how to transform to standard form?

- inequality constraints $Ax \leq b$? slack variables $s \geq 0$
- free variable $x \in \mathbf{R}^n$?

LP inequality form

another common form for LP is **inequality form**

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

how to transform to standard form?

- inequality constraints $Ax \leq b$? slack variables $s \geq 0$
- free variable $x \in \mathbf{R}^n$? split into positive and negative parts

we will see later that these forms are also related by **duality**

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

Geometry of LP: inequality form

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

- $Ax \leq b$ defines a **polyhedron**

Geometry of LP: inequality form

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

- $Ax \leq b$ defines a **polyhedron**
- $\implies$ feasible set $P = \{x : Ax \leq b\}$ is a polyhedron

Geometry of LP: inequality form

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

- $Ax \leq b$ defines a **polyhedron**
- $\implies$ feasible set $P = \{x : Ax \leq b\}$ is a polyhedron
- x is a **vertex** of polyhedron P if there is some v so that

$$v^T x < v^T y, \qquad \forall y \in P \setminus \{x\}$$

Geometry of LP: inequality form

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

- $Ax \leq b$ defines a **polyhedron**
- $\implies$ feasible set $P = \{x : Ax \leq b\}$ is a polyhedron
- x is a **vertex** of polyhedron P if there is some v so that

$$v^T x < v^T y, \qquad \forall y \in P \setminus \{x\}$$

fact: vertex $\iff$ extreme point

Solution of LP is extreme point

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

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

Solution of LP is extreme point

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

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

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

Solution of LP is extreme point

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

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

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

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

Solution of LP is extreme point

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

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

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

- unique: so $c^T x < c^T y$ for all $y \in P \setminus \{x\}$
- not unique: $\{X^\star : c^T x = c^T x^\star, x \in P\}$ 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 P.

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

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

Discuss: how to solve LPs?

write down a method to solve LPs; discuss in groups.

- idea
- math
- pseudocode

complete https://forms.gle/JbP2fLd6cRVbNUoW9 when you're ready (and before Friday noon)
(link also available from course schedule)

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 $S \subseteq \{1, \ldots, n\}$ 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?

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 $\langle b, y \rangle < 0$*

$\implies$ 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 $\langle b, y \rangle < 0$*

$\Longrightarrow$ 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 $p^\star$:

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

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

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

Lagrange duality

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

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

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

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:

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

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:

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

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

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:

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

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

use second system to show 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,

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

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:

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

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

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

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

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

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

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

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

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

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

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

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):

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

example:

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

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 $\mathrm{UB}(R) \geq p^(R)$

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 $\mathrm{UB}(R) \geq p^(R)$

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 $\mathrm{UB}(R) \geq p^(R)$

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