Aim

- What should a 1-hour PhD lecture on LP achieve?
- Audience members have
  - Different backgrounds in LP
  - Different reasons to know about LP
- Answer
  - "Mature" review of LP theory—don’t worry if you don’t follow!
  - Identify an alternative to the tableau simplex method
  - Establish a result which
    - Is nice in itself
    - Leads into "Structure and matrix sparsity": Wednesday 13:30–15:30

Overview

- What is LP?
- General LP problems
  - The feasible region
  - Basic solutions
  - Fundamental results
- The simplex method
  - Algorithm
  - Implementation
  - Linear algebra

What is linear programming (LP)?

- The most important model used in optimal decision-making
- Grew out of US Army Air Force logistics problems in WW2
- First practical LP problem was formulated by George Dantzig in 1946
- Dantzig invented the principal solution technique—the simplex algorithm—in 1947
  - In the “Top 10 algorithms of the 20th century”
  - “The algorithm that runs the world”
- LP and the simplex method have revolutionised organised human decision-making
**General LP problems**

**General LP problems: Introduction**

- General LP problem is
  \[
  \text{maximize } f = c^T x \text{ subject to } Ax \leq b, \quad x \geq 0
  \]

  Problem has \( n \) variables and \( m \) constraints

- Feasible region is a **convex polyhedron** in \( \mathbb{R}^n \)

- Vertices are at intersection of \( n \) constraint/bound planes

  Aid vertex characterization and algebraic analysis by introducing **slack variables**

  - For each inequality introduce a slack variable \( x_{n+i} \)
  - Transforms the inequality into an equation and bound on \( x_{n+i} \)
  - Inherent LP problem is unchanged, but is in **standard form**

  \[
  \text{maximize } f = \varepsilon^T x \text{ subject to } \overline{A}x = b, \quad x \geq 0
  \]

  where \( \varepsilon = \begin{bmatrix} c \\ 0 \end{bmatrix} \) and \( \overline{A} = [A \ \ I] \) has rank \( m \)

**General LP problems: The feasible region, vertices and solutions**

For an LP in standard form, \( \overline{A}x = b \) where \( \overline{A} = [A \ \ I] \in \mathbb{R}^{m \times (n+m)} \)

- **Characterization of the feasible region**
  - The feasible region \( K \) is the intersection of
    - The hyperplane of solutions of \( \overline{A}x = b \)
    - The orthant \( x \geq 0 \)
  - \( K \) is a **convex set**

- **Definition of a feasible vertex**
  - A feasible vertex is a point \( x \in K \) which does not lie strictly within any line segment joining two points in \( K \)

**Result:** If an LP has an optimal solution then there is an optimal solution at a vertex

**General LP problems: Basic solutions**

- The point \( x \in \mathbb{R}^{n+m} \) is a **basic solution** of an LP problem in standard form if there is a **partition** of \( \{1,2,\ldots,n+m\} \) into
  - A set \( N \) of \( n \) indices of **nonbasic variables** with value zero at \( x \)
  - A set \( B \) of \( m \) indices of **basic variables** whose values are then uniquely defined by the \( m \) equations

- Corresponding to \( N \) and \( B \) the following are defined
  - Corresponding to the \( n \) indices in \( N \)
    - The matrix \( N \in \mathbb{R}^{m \times n} \) is the \( n \) columns of \( \overline{A} \)
    - The vector \( x_N \) of **nonbasic variables** is the \( n \) components of \( x \)
    - The vector \( c_N \) of **nonbasic costs** is the \( n \) components of \( c \)
  - Corresponding to the \( m \) indices in \( B \)
    - The **basis matrix** \( B \in \mathbb{R}^{m \times m} \) is the \( m \) columns of \( \overline{A} \) and is nonsingular
    - The vector \( x_B \) of **basic variables** is the \( m \) components of \( x \)
    - The vector \( c_B \) of **basic costs** is the \( m \) components of \( c \)
The partitioned LP in standard form is

\[
\text{maximize } f = c_X^T x_B + c_N^T x_N \\
\text{subject to } B x_B + N x_N = b \\
x_B \geq 0, x_N \geq 0
\]

- Same LP as in standard form
- Corresponds to reordering the components of \( x \) according to sets \( B \) and \( N \)
- Equations are \( B x_B + N x_N = b \) so \( x_B = B^{-1} b - B^{-1} N x_N \)
- Substituting for \( x_B \) in the objective function gives \( f = \tilde{f} + \tilde{c}_N^T x_N \), where
  \[
  \tilde{f} = c_X \hat{b} \text{ is the objective value when } x_N = 0 \\
  \tilde{c}_N = c_N - N^T B^{-T} c_B \text{ is the vector of reduced costs}
  \]
- If \( \hat{b} \geq 0 \) then \( x \geq 0 \) is referred to as a basic feasible solution

Result: \( x \) is a vertex of \( K \) iff \( x \) is a basic feasible solution

Consequence: Solution methods need only consider basic feasible solutions

Result: A point \( x \in K \) is an optimal solution of an LP problem if it is a basic feasible solution with non-positive reduced costs \( \tilde{c}_N \leq 0 \)

Consequence:
- Condition \( \tilde{c}_N \leq 0 \) allows the optimality of a basic feasible solution to be checked
- If \( \tilde{c}_N \leq 0 \) the simplex algorithm identifies a basic feasible solution with better objective value
This is the key to solving LP problems

Simplex method: Choosing an improving direction

- Observe: If \( \tilde{c}_N \leq 0 \) there exists \( q \) such that \( \tilde{c}_q > 0 \)
- Let \( x_q' \) be the \( q \)th nonbasic variable
- If \( x \) is partitioned as \( [x_B \ x_N] \) then consider \( x + \alpha d \) for \( d \) partitioned as \( [d_B \ d_N] \)
- Only nonbasic variable \( x_q' \) is increased from zero if \( d_N = e_q \) \([e_q \text{ is column } q \text{ of } I]\)
- For feasibility \( \overline{A} d = 0 \) iff
  \[
  B d_B + N e_q = 0 \iff d_B = -B^{-1} N e_q = -\tilde{a}_q
  \]
  where \( B \tilde{a}_q = a_q \) and \( a_q \) is column \( q \) of \( N \)
- If \( x + \alpha d \) is feasible for \( \alpha > 0 \) then the objective increases strictly by \( \alpha \tilde{c}_q \)
Simplex method: Identifying the step length in the improving direction

- On $x + \alpha d$ for $\alpha \geq 0$, components of $x_N$ remain feasible since $d_N = e_q$
- Any limit on the feasibility of $x + \alpha d$ is given by the values of the basic variables
  $$x_B = \hat{b} - \alpha \hat{a}_q$$
- If $\hat{a}_q$ has positive components
  - For $\alpha$ sufficiently large, at least one component of $x_B$ will be zeroed
  - The smallest of these values of $\alpha$ is the greatest step $\alpha$ which can be made in the direction $d$ whilst maintaining feasibility
- If $\hat{a}_q \leq 0$ no basic variable is zeroed so the LP is **unbounded**

Simplex method: Identifying the new basic feasible solution

- At $x + \pi d$ let $x_{p'}$ be the zeroed basic variable
- Interchanging $p'$ and $q'$ between $B$ and $N$ yields a partition at $x + \pi d$ with
  - $x_p \geq 0$
  - $x_n = 0$
- $x + \pi d$ is a new basic feasible solution since its basis matrix is nonsingular
  - If $\pi > 0$ then
    - $x + \pi d$ is a distinct, “new” vertex
    - Objective at $x + \pi d$ is strictly greater (by $\pi \hat{c}_q$) than the objective at $x$
    - If $\pi > 0$ always holds then termination is provable
  - If $\pi = 0$ then $x$ is **degenerate** and the simplex algorithm can **cycle/stall**

Simplex method: Algorithm description

<table>
<thead>
<tr>
<th>Description of the simplex algorithm</th>
<th>Description of the simplex algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Given a basic feasible solution $x$</strong></td>
<td><strong>Given a basic feasible solution $x$</strong></td>
</tr>
</tbody>
</table>
| - If the reduced costs are non-positive then **stop**  
  **The solution is optimal** | - If the reduced costs are non-positive then **stop**  
  **The solution is optimal** |
| - Determine the nonbasic variable $x_{q'}$ with most positive reduced cost  
  Determine the feasible direction $d$ when $x_{q'}$ is increased from zero  
  If no basic variable is zeroed on $x + \alpha d$ then **stop**  
  **The LP is unbounded** | - Determine the nonbasic variable $x_{q'}$ with most positive reduced cost  
  Determine the feasible direction $d$ when $x_{q'}$ is increased from zero  
  If no basic variable is zeroed on $x + \alpha d$ then **stop**  
  **The LP is unbounded** |
| - Determine the first basic variable $x_{p'}$ to be zeroed on $x + \alpha d$  
  Make $x_{p'}$ nonbasic and $x_{q'}$ basic  
  Go to 1 | - Determine the first basic variable $x_{p'}$ to be zeroed on $x + \alpha d$  
  Make $x_{p'}$ nonbasic and $x_{q'}$ basic  
  Go to 1 |

Unbounded step if $q' = 3$  
Bounded step if $q' = 1$
Simplex method: Algorithm definition

**Definition of the simplex algorithm**

Given a basic feasible solution $x$ with $B$ and $N$

1. If $\hat{c}_N \leq 0$ then **stop: the solution is optimal**
2. Determine the index $q' \in N$ of the variable $x_{q'}$ with most positive reduced cost $\hat{c}_q$
3. Let $\hat{a}_q = B^{-1}a_q$, where $a_q$ is column $q$ of $N$
4. If $\hat{a}_q \leq 0$ then **stop: the LP is unbounded**
5. Determine the index $p' \in B$ of the variable $x_{p'}$ corresponding to $p = \arg\min_{i=1, \ldots, m} \hat{b}_i \hat{a}_{iq} > 0$
6. Exchange indices $p'$ and $q'$ between $B$ and $N$ to yield a new basic feasible solution
7. Go to 1

**Simplex method: Basis matrix update**

**Updating $B$**

- Each simplex iteration exchanges the $p$th entry of $B$ with the $q$th entry of $N$
  - Column $p$ of $B$ is replaced by the vector $a_q$ [column $q$ of $N$]
  - The updated basis matrix $B'$ is given by
    \[
    B' = B + a_q e_p^T - a_p e_q^T = B [I + (\hat{a}_q - e_p)e_p^T]
    \]
    \[
    = BE \quad \text{where} \quad E = I + (\hat{a}_q - e_p)e_p^T
    \]

This result has fundamental theoretical and practical consequences

---

**Simplex method: $B'$ is nonsingular**

- $B'$ is nonsingular iff $E$ is nonsingular
  [Since $B' = BE$, where $E = I + (\hat{a}_q - e_p)e_p^T$ and $B$ is nonsingular]

- Nonsingularity of $E$ may be established as follows
  - For a matrix $A = I + uv^T$ the (Sherman-Morrison) formula for $A^{-1}$ is
    \[
    A^{-1} = I - \frac{1}{1 + v^Tu}uv^T \quad \text{when} \quad v^Tu \neq -1
    \]
  - Hence, using $u = \hat{a}_q - e_p$ and $v = e_p$
    \[
    E^{-1} = I - \frac{1}{\hat{a}_{pq}} (\hat{a}_q - e_p)e_p^T
    \]
  - Since the pivot value in the simplex iteration is $\hat{a}_{pq} \neq 0$, it follows that $E$ is nonsingular

**The simplex method: Implementation**
Simplex method: Implementation

The data for a simplex iteration are

- **The reduced costs** \( \tilde{c}_N = c_N - N^T B^{-T} c_B \)
- **The pivotal column** \( \tilde{a}_q = B^{-1} a_q \)
- **The reduced RHS** \( \tilde{b} = B^{-1} b \)

How to obtain these vectors determines the efficiency of the simplex implementation

- **Standard simplex method (SSM)**: maintains \( B^{-1} N, \tilde{b} \) and \( \tilde{c}_N \) in a rectangular tableau
  - Requires \( O(mn) \) storage and \( O(mn) \) computation per iteration
  - Inefficient and prohibitively expensive for large problems
- **Revised simplex method (RSM)**: computes \( \tilde{a}_q = B^{-1} a_q \) as required, forms \( \tilde{c}_N \) and updates \( \tilde{b} = B^{-1} b \)
  - Requires up to \( O(m^2) \) storage and \( O(m^2) \) computation per iteration but vastly less for sparse LP problems
  - Efficient for large (sparse) problems

For the reduced costs \( \tilde{c}_N = c_N - N^T B^{-T} c_B \) observe that

\[
\tilde{c}_N = c_N - N^T \pi
\]

where \( \pi = B^{-T} c_B \)

so \( \tilde{c}_N \) may be formed as

- Solve \( B^T \pi = c_B \)
- Form \( z = N^T \pi \)
- Then \( \tilde{c}_N = c_N - z \)

For the pivotal column \( \tilde{a}_q = B^{-1} a_q \)

- Solve \( B \tilde{a}_q = a_q \)

For the reduced RHS \( \tilde{b} = B^{-1} b \)

- Exploit \( x_B = \tilde{b} - \pi \tilde{a}_q \) to update \( \tilde{b} \)

Simplex method: Implementation (RSM)

LU decomposition

- GE applied to \( B \) yields the decomposition \( B = LU \) at cost \( O(m^3) \)
  - \( L \) is the lower triangular matrix of elimination multipliers
  - Diagonal entries are all one
  - \( U \) is the upper triangular matrix after GE
- Linear systems with lower (upper) triangular coefficient matrix can be solved by forward (backward) substitution
- Given \( B = LU \), solve \( B \tilde{a}_q = a_q \) as
  \[
  L y = a_q \quad \text{then} \quad U \tilde{a}_q = y
  \]
- Since \( B^T = U^T L^T \), solve \( B^T \pi = c_B \) as
  \[
  U^T y = c_B \quad \text{then} \quad L^T \pi = y
  \]
- Cost of each forward and backward substitution is \( O(m^2) \)
Updating the invertible representation of $B$

- Recall
  
  \[ B' = BE \]
  
  where
  
  \[ E = \left[ I + (\hat{a}_q - e_p)e_p^T \right] \]
  
  and
  
  \[ E^{-1} = I - \frac{1}{\hat{a}_{pq}}(\hat{a}_q - e_p)e_p^T \]

- $B'x = b$ may be solved as
  
  \[ BEx = b \iff By = b; \quad x = E^{-1}y \]

- Similarly $B^T x = b$ may be solved as
  
  \[ E^T B^T x = b \iff y = E^{-T}b; \quad B^T x = y \]

- Since $E$ has only one non-trivial column, it may be stored using a single vector and the index $p$ and operations with $E^{-1}$ cost $O(m)$

- This product form (PF) update technique can be extended to perform multiple updates
  
  - Eventually the cost of working with all the matrices of the form $E$ will dominate
  
  - Then preferable to recompute the LU decomposition at cost $O(m^3)$

- \[ \text{Solving all systems at cost } O(m^2) \text{ each} \]

- Start with $B = \{n+1, n+2, \ldots, n+m\} \text{ so } B = I$ has trivial LU decomposition $L = U = I$

- Use the Fletcher-Matthews technique to update the LU decomposition itself
  
  - Allows the LU decomposition of $B'$ to be obtained by updating the LU decomposition of $B$ at a cost $O(m^2)$
  
  - Numerically stable (unlike PF update)
  
  - No recalculation of LU decomposition of $B$ is required

The simplex method: Why use it?

- Interior point methods (IPM) are the “modern” alternative to the simplex method

- For single LP problems IPM are frequently faster

- For some classes of single LP problems the simplex method is faster

- When solving sequences of related LP problems the simplex method is preferable
  
  - Branch-and-bound for discrete optimization
  
  - Sequential linear programming for nonlinear optimization

- Why?
  
  - Simplex method yields a basic feasible solution
  
  - Simplex method can be re-started easily from an optimal solution of one LP to solve a related LP quickly

- 67 years old and going strong!
Linear Programming 1: Summary

- Identified that solution methods need only consider basic feasible solutions of LPs
- Described the simplex algorithm
- Identified that efficient implementation depends on techniques for solving related systems of equations
- Remains to consider how to exploit LP problem **structure and matrix sparsity**.

Wednesday 13:30–15:30