18. Linear optimization

• linear program
• examples
• geometrical interpretation
• extreme points
• simplex method
Linear program

minimize \[ c_1 x_1 + c_2 x_2 + \cdots + c_n x_n \]
subject to \[ \begin{align*}
a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n & \leq b_1 \\
a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n & \leq b_2 \\
\vdots \\
a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n & \leq b_m 
\end{align*} \]

- \( n \) optimization (decision) variables \( x_1, \ldots, x_n \)
- \( m \) linear inequality constraints

matrix notation

minimize \( c^T x \)
subject to \( Ax \leq b \)

the inequality between vectors \( Ax \leq b \) is interpreted elementwise
Production planning

- a company makes three products, in quantities $x_1, x_2, x_3$ per month
- profit per unit is 1.0 (for product 1), 1.4 (product 2), 1.6 (product 3)
- products use different amounts of resources (labor, material, ...)

| Fraction of total available resource needed per unit of each product |
|-----------------|-----------------|-----------------|
| resource 1      | 1/1000          | 1/800           | 1/500           |
| resource 2      | 1/1200          | 1/700           | 1/600           |

optimal production plan

maximize $x_1 + 1.4x_2 + 1.6x_3$
subject to $(1/1000)x_1 + (1/800)x_2 + (1/500)x_3 \leq 1$
             $(1/1200)x_1 + (1/700)x_2 + (1/600)x_3 \leq 1$
             $x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0$

solution: $x_1 = 462, \quad x_2 = 431, \quad x_3 = 0$
Network flow optimization

maximize \( t \)
subject to
flow conservation at nodes
capacity constraints on the arcs

capacity constraints

\[
\begin{align*}
0 & \leq x_1 \leq 3 \\
0 & \leq x_2 \leq 2 \\
0 & \leq x_3 \leq 1 \\
0 & \leq x_4 \leq 2 \\
0 & \leq x_5 \leq 1 \\
0 & \leq x_6 \leq 3 \\
0 & \leq x_7 \leq 3 \\
0 & \leq x_8 \leq 1 \\
0 & \leq x_9 \leq 1
\end{align*}
\]
linear programming formulation

maximize \( t \)
subject to \( t = x_1 + x_2, \ x_1 + x_3 = x_4, \) et cetera
\( 0 \leq x_1 \leq 3, \ 0 \leq x_2 \leq 2, \) et cetera

\( t = x_1 + x_2 \) is equivalent to inequalities \( t \leq x_1 + x_2, \ t \geq x_1 + x_2, \ldots \)

solution
Data fitting

fit a straight line to $m$ points $(u_i, v_i)$ by minimizing sum of absolute errors

$$\text{minimize } \sum_{i=1}^{m} |\alpha + \beta u_i - v_i|$$

dashed: least-squares solution; solid: minimizes sum of absolute errors
linear programming formulation

minimize \[ \sum_{i=1}^{m} y_i \]
subject to \[-y_i \leq \alpha + \beta u_i - v_i \leq y_i, \quad i = 1, \ldots, m\]

- variables \( \alpha, \beta, y_1, \ldots, y_m \)
- inequalities are equivalent to
  \[ y_i \geq |\alpha + \beta u_i - v_i| \]
- the optimal \( y_i \) satisfies
  \[ y_i = |\alpha + \beta u_i - y_i| \]
Terminology

minimize \( c^T x \)
subject to \( Ax \leq b \)

problem data: \( n \)-vector \( c \), \( m \times n \)-matrix \( A \), \( m \)-vector \( b \)

- \( x \) is **feasible** if \( Ax \leq b \)
- **feasible set** is set of all feasible points
- \( x^* \) is **optimal** if it is feasible and \( c^T x^* \leq c^T x \) for all feasible \( x \)
- **optimal value**: \( p^* = c^T x^* \)
- **unbounded problem**: \( c^T x \) is unbounded below on feasible set
- **infeasible problem**: feasible set is empty
Example

minimize $-x_1 - x_2$

subject to

$2x_1 + x_2 \leq 3$

$x_1 + 4x_2 \leq 5$

$x_1 \geq 0, \quad x_2 \geq 0$

feasible set is shaded
solution

minimize \(-x_1 - x_2\)
subject to
\[\begin{align*}
2x_1 + x_2 &\leq 3 \\
x_1 + 4x_2 &\leq 5 \\
x_1 &\geq 0, \quad x_2 &\geq 0
\end{align*}\]

optimal solution is \(x^* = (1, 1)\), optimal value is \(p^* = -2\)

Linear optimization
Hyperplanes and halfspaces

hyperplane

solution set of one linear equation with nonzero coefficient vector \( a \)

\[
a^T x = b
\]

halfspace

solution set of one linear inequality with nonzero coefficient vector \( a \)

\[
a^T x \leq b
\]
Geometrical interpretation

\[ G = \{ x \mid a^T x = b \} \]

\[ H = \{ x \mid a^T x \leq b \} \]

- \( u = (b/\|a\|^2) a \) satisfies \( a^T u = b \)
- \( x \) is in \( G \) if \( a^T (x - u) = 0 \), i.e., \( x - u \) is orthogonal to \( a \)
- \( x \) is in \( H \) if \( a^T (x - u) \leq 0 \), i.e., angle \( \angle(x - u, a) \geq \pi/2 \)
Example

\[
\begin{align*}
\mathbf{a}^T \mathbf{x} &= 0 \\
\mathbf{a}^T \mathbf{x} &= 10 \\
\mathbf{a}^T \mathbf{x} &= -5 \\
\mathbf{a}^T \mathbf{x} &= 5
\end{align*}
\]

\[
\mathbf{a} = (2, 1)
\]

\[
\mathbf{a}^T \mathbf{x} \leq 3
\]
definition: the solution set of a finite number of linear inequalities

\[ a_1^T x \leq b_1, \quad a_2^T x \leq b_2, \quad \ldots, \quad a_m^T x \leq b_m \]

matrix notation: \( Ax \leq b \) where

\[
A = \begin{bmatrix}
  a_1^T \\
  a_2^T \\
  \vdots \\
  a_m^T 
\end{bmatrix}, \quad b = \begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m 
\end{bmatrix}
\]

geometrical interpretation: intersection of \( m \) halfspaces
example \((n = 2)\)

\[
x_1 + x_2 \geq 1, \quad -2x_1 + x_2 \leq 2, \quad x_1 \geq 0, \quad x_2 \geq 0
\]
example \((n = 3)\)

\[0 \leq x_1 \leq 2, \quad 0 \leq x_2 \leq 2, \quad 0 \leq x_3 \leq 2, \quad x_1 + x_2 + x_3 \leq 5\]
Extreme points

let $\hat{x} \in \mathcal{P}$, where $\mathcal{P}$ is a polyhedron defined by inequalities $a_k^T x \leq b_k$

- if $a_k^T \hat{x} = b_k$, we say the $k$th inequality is active at $\hat{x}$
- if $a_k^T \hat{x} < b_k$, we say the $k$th inequality is inactive at $\hat{x}$

$\hat{x}$ is called an extreme point of $\mathcal{P}$ if

$$A_{I(\hat{x})} = \begin{bmatrix} a_{k_1}^T \\ a_{k_2}^T \\ \vdots \\ a_{k_p}^T \end{bmatrix} \text{ has a zero nullspace}$$

where $I(\hat{x}) = \{k_1, \ldots, k_p\}$ are the indices of the active constraints at $\hat{x}$

an extreme point $\hat{x}$ is nondegenerate if $p = n$, degenerate if $p > n$
example

\[-x_1 - x_2 \leq -1, \quad -2x_1 + x_2 \leq 1, \quad -x_1 \leq 0, \quad -x_2 \leq 0\]

- \(\mathbf{\hat{x}} = (1, 0)\) is an extreme point
  
  \[A_{I(\mathbf{\hat{x}})} = \begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix}\]

- \(\mathbf{\hat{x}} = (0, 1)\) is an extreme point
  
  \[A_{I(\mathbf{\hat{x}})} = \begin{bmatrix} -1 & -1 \\ -2 & 1 \\ -1 & 0 \end{bmatrix}\]

- \(\mathbf{\hat{x}} = (1/2, 1/2)\) is not an extreme point
  
  \[A_{I(\mathbf{\hat{x}})} = \begin{bmatrix} -1 & -1 \end{bmatrix}\]

Linear optimization
Simplex algorithm

minimize \( x_1 + x_2 - x_3 \)
subject to \( 0 \leq x_1 \leq 2, \quad 0 \leq x_2 \leq 2, \quad 0 \leq x_3 \leq 2 \)
\( x_1 + x_2 + x_3 \leq 5 \)

move from one extreme point to another extreme point with lower cost
One iteration of the simplex method

suppose $\hat{x}$ is a nondegenerate extreme point

renumber constraints so that active constraints are 1, $\ldots$, $n$

**active constraints** at $\hat{x}$:

$$a_1^T \hat{x} = b_1, \ldots, a_n^T \hat{x} = b_n$$

**inactive constraints** at $\hat{x}$:

$$a_{n+1}^T \hat{x} < b_{n+1}, \ldots, a_m^T \hat{x} < b_m$$

**matrix of active constraints**: define $I = I(\hat{x}) = \{1, \ldots, n\}$

$$A_I = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix} \quad \text{(an $n \times n$ nonsingular matrix)}$$
step 1 (test for optimality) solve

\[ A^T_I z = c \]

\( i.e., \) find \( z \) that satisfies

\[ \sum_{k=1}^{n} z_k a_k = c \]

if \( z_k \leq 0 \) for all \( k \), then \( \hat{x} \) is optimal and we return \( \hat{x} \)

proof. consider any feasible \( x \):

\[ c^T x = \sum_{k=1}^{n} z_k a_k^T x \geq \sum_{k=1}^{n} z_k b_k = \sum_{k=1}^{n} z_k a_k^T \hat{x} = c^T \hat{x} \]

(the inequality follows from \( z_k \leq 0, \ a_k^T x \leq b_k \))
step 2 (compute step) choose a $j$ with $z_j > 0$ and solve

$$A_Iv = -e_j$$

i.e., find $v$ that satisfies

$$a_j^Tv = -1, \quad a_k^Tv = 0 \quad \text{for } k = 1, \ldots, n \text{ and } k \neq j$$

for small $t > 0$, $\hat{x} + tv$ is feasible and has a smaller cost than $\hat{x}$

proof:

$$a_j^T(\hat{x} + tv) = b_j - t$$
$$a_k^T(\hat{x} + tv) = b_k \quad \text{for } k = 1, \ldots, n \text{ and } k \neq j$$

and for sufficiently small positive $t$

$$a_k^T(\hat{x} + tv) \leq b_k \quad \text{for } k = n + 1, \ldots, m$$

finally, $c^T(\hat{x} + tv) = c^T\hat{x} + tz_j < c^T\hat{x}$
step 3 (compute step size) maximum $t$ such that $\hat{x} + tv$ is feasible, i.e.,

$$a_k^T \hat{x} + ta_k^Tv \leq b_k, \quad k = 1, \ldots, m$$

the maximum step size is

$$t_{\text{max}} = \min_{k: a_k^Tv > 0} \frac{b_k - a_k^T \hat{x}}{a_k^Tv}$$

(if $a_k^Tv \leq 0$ for all $k$, then the problem is unbounded below)

step 4 (update $\hat{x}$)

$$\hat{x} := \hat{x} + t_{\text{max}}v$$
Example

minimize \( x_1 + x_2 - x_3 \)
subject to \( 0 \leq x_1 \leq 2, \quad 0 \leq x_2 \leq 2, \quad 0 \leq x_3 \leq 2 \)
\( x_1 + x_2 + x_3 \leq 5 \)

at \( \hat{x} = (2, 2, 0) \)

\[ A_I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \]

1. \( z = A_I^{-T}c = (1, 1, 1) \)
2. for \( j = 3, \; v = A_I^{-1}(0, 0, -1) = (0, 0, 1) \)
3. \( \hat{x} + tv \) is feasible for \( t \leq 1 \)
Example

minimize \[ x_1 + x_2 - x_3 \]
subject to \[ 0 \leq x_1 \leq 2, \quad 0 \leq x_2 \leq 2, \quad 0 \leq x_3 \leq 2 \]
\[ x_1 + x_2 + x_3 \leq 5 \]

at \( \hat{x} = (2, 2, 1) \)

\[ A_I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \]

1. \( z = A_I^{-T}c = (2, 2, -1) \)

2. for \( j = 2 \), \( v = A_I^{-1}(0, -1, 0) = (0, -1, 1) \)

3. \( \hat{x} + tv \) is feasible for \( t \leq 1 \)
Example

minimize \quad x_1 + x_2 - x_3
subject to \quad 0 \leq x_1 \leq 2, \quad 0 \leq x_2 \leq 2, \quad 0 \leq x_3 \leq 2
\quad x_1 + x_2 + x_3 \leq 5

at \: \hat{x} = (2, 1, 2)

\begin{align*}
A_I &= \begin{bmatrix}
1 & 0 & 0 \\
0 & 0 & 1 \\
1 & 1 & 1
\end{bmatrix}
\end{align*}

1. \: z = A_I^{-T}c = (0, -2, 1)

2. \: for \: j = 3, \: v = A_I^{-1}(-1, 0, 0) = (0, -1, 0)

3. \: \hat{x} + tv \: is \: feasible \: for \: t \leq 1
Example

minimize \( x_1 + x_2 - x_3 \)
subject to \( 0 \leq x_1 \leq 2, \quad 0 \leq x_2 \leq 2, \quad 0 \leq x_3 \leq 2 \)
\( x_1 + x_2 + x_3 \leq 5 \)

at \( \hat{x} = (2, 0, 2) \)

\[
A_I = \begin{bmatrix}
1 & 0 & 0 \\
0 & -1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\]

1. \( z = A_I^{-T}c = (1, -1, -1) \)

2. for \( j = 1 \), \( v = A_I^{-1}(-1, 0, 0) = (-1, 0, 0) \)

3. \( \hat{x} + tv \) is feasible for \( t \leq 2 \)
Example

minimize \[ x_1 + x_2 - x_3 \]
subject to \[ 0 \leq x_1 \leq 2, \quad 0 \leq x_2 \leq 2, \quad 0 \leq x_3 \leq 2 \]
\[ x_1 + x_2 + x_3 \leq 5 \]

at \( \hat{x} = (0, 0, 2) \)

\[ A_I = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \]

\[ z = A_I^{-T} c = (-1, -1, -1) \]

therefore, \( \hat{x} \) is optimal
Practical aspects

- at each iteration, we solve two sets of linear equations

\[ A^T_I z = c, \quad A_I v = -e_j \]

using an LU factorization or sparse LU factorization of \( A_I \)

- implementation requires a ‘phase 1’ to find first extreme point

- simple modifications handle problems with degenerate extreme points

- very large LPs (several 100,000 variables and constraints) are routinely solved in practice