Lecture 1

Linear Programming: Formulation and Solution

Qiong Wang

Department of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana-Champaign

August 21-23, 2023

Example

You plan to invest in three companies. Expected returns from investment in companies 1, 2, and 3 are $r_1 = 0.05$, $r_2 = 0.12$, and $r_3 = 0.09$ respectively. For strategic reasons, you do not want to invest more than 40% your budget in company 1, 50% in company 2, and at least 30% in company 3. Determine the percentage of your budget to be invested in each company to maximize your expected return?

- *variables:* x_1, x_2, x_3 be the fractions invested in three companies.
- *objective function:*

$$\max_{x_1, x_2, x_3} \{0.05x_1 + 0.12x_2 + 0.09x_3\}$$

- *constraints:*

$$x_1 + x_2 + x_3 \leq 1,$$

$$x_1 \leq 0.4, \;\; x_2 \leq 0.5, \;\; x_3 \geq 0.3,$$

$$x_1 \geq 0, \;\; x_2 \geq 0, \;\; x_3 \geq 0.$$

More general formulation

You plan to invest in n different assets with the expected returns r_i ($i = 1, \cdots, n$). You are required to invest at least a fraction of s_i and at most S_i of your budget in asset i ($i = 1, \cdots, n$). Determine an investment solution to maximize the total expected returns.

x_i: fraction of the budget in excess of s_i invested in asset i (the actual fraction is $x_i + s_i$), $i = 1, \cdots, n$.

let

$$\Delta s_i = S_i - s_i, \quad i = 1, \cdots, n, \quad \bar{s} = \sum_{i=1}^{n} s_i, \quad \bar{R} = \sum_{i=1}^{n} r_i s_i.$$

$$\begin{aligned} & \max_{x_1, \ldots, x_n} \{r_1 x_1 + ... + r_n x_n\} + \bar{R} \\ \text{subject to:} \quad & x_1 \le \Delta s_1, \; x_2 \le \Delta s_2, ..., x_n \le \Delta s_n, \\ & x_1 + ... + x_n \le 1 - \bar{s}, \\ & x_1 \ge 0, \; x_2 \ge 0, ..., x_n \ge 0. \end{aligned}$$

Definition of Linear Programming Problem

Linear Programming (LP) is an optimization model in which

- *all* decision variables take real values,
- the objective is a linear function of decision variables,
- *all* constraints are linear (in)equalities of decision variables.

$$\max_{x_1,\ldots,x_n} \{r_1 x_1 + ... + r_n x_n\} + \bar{R}$$

$$\begin{aligned} \text{subject to:} \quad & x_1 \le \Delta s_1,\ x_2 \le \Delta s_2, ..., x_n \le \Delta s_n, \\ & x_1 + ... + x_n \le 1 - \bar{s}, \\ & x_1 \ge 0,\ x_2 \ge 0, ..., x_n \ge 0. \end{aligned}$$

n decision variables: $x_1, .., x_n$.

$2n + 2$ parameters: $\bar{R}$, $\bar{s}$, $(r_1, ..., r_n)$ and $(\Delta s_1, ..., \Delta s_n)$.

$(n + 1)$ "regular" constraints and n nonnegativity constraints.

Standard Form

LP in standard form:

1. Maximize an objective.
2. Left Hand Side (LHS) of each constraint, which is a linear combination of decision variables, is less than or equal to its Right Hand Side (RHS), which is a constant.
3. All decision variables are positive.

LP formulation with n decision variables and $m = n + k$ constraints

$$\max_{x_1,\dots,x_n} \{c_1x_1 + ... + c_nx_n\}$$

$$\begin{aligned}\text{subject to:} \quad & a_{11}x_1 + ... + a_{1n}x_n \leq b_1, \\ & .... \\ & a_{k1}x_1 + ... + a_{kn}x_n \leq b_k, \\ & x_1 \geq 0, ..., x_n \geq 0.\end{aligned}$$

Transform an LP into Its Standard Form

$$\max_{x_1,\dots,x_n} \{c_1x_1 + ... + c_nx_n\}$$

subject to: $a_{j1}x_1 + ... + a_{jn}x_n \leq b_j, \quad j = 1, \cdots, k$

$$x_1 \geq 0, ..., x_n \geq 0.$$

how to fit following situations into the standard LP form?

- the objective is $\min_{x_1,\dots,x_n} \{c_1x_1 + ... + c_nx_n\}$?

$$\max_{x_1,\dots,x_n} \{-c_1x_1 - .... - c_nx_n\}$$

- constraint $a_{j1}x_1 + ... + a_{jn}x_n \geq b_j$?

$$-a_{j1}x_1 - ... - a_{jn}x_n \leq -b_j,$$

- constraint $a_{j1}x_1 + ... + a_{jn}x_n = b_j$?

$$a_{j1}x_1 + ... + a_{jn}x_n \leq b_j \quad \text{and} \quad a_{j1}x_1 + ... + a_{jn}x_n \geq b_j.$$

i.e, $a_{j1}x_1 + ... + a_{jn}x_n \leq b_j$ and $-a_{j1}x_1 - ... - a_{jn}x_n \leq -b_j$.

- a decision variable x can be negative?

$$x = y_1 - y_2, \; y_1 \geq 0, y_2 \geq 0.$$

Another Standard Formulation

Dual of the maximization problem (to be discussed later in the course):

1. instead of
$$\max_{x_1,\ldots,x_n} \{c_1x_1 + ... + c_nx_n\}$$
write the objective function as
$$\min_{y_1,\ldots,y_n} \{h_1y_1 + ... + h_ny_n\}$$

2. instead of
$$a_{j1}x_1 + ... + a_{jn}x_n \leq b_j$$
make the LHS of constraints greater than or equal to the RHS:
$$a_{11}y_1 + ... + a_{n1}y_n \geq d_1,$$

3. all decision variables are positive.
$$y_1 \geq 0, ..., y_n \geq 0.$$

Linear Programming (LP) in Standard Form

$$
\begin{aligned}
\max_{x_1,\ldots,x_n} \quad & \{c_1x_1 + ... + c_nx_n\} \\
\text{subject to} \quad & a_{11}x_1 + ... + a_{1n}x_n \leq b_1, \\
& \ldots\ldots\ldots \\
& a_{k1}x_1 + ... + a_{kn}x_n \leq b_k, \\
& x_1 \geq 0, ..., x_n \geq 0.
\end{aligned}
$$

- all decision variables take non-negative real values;
- the objective is to maximize a linear function of decision variables;
- all constraints are linear inequalities of decision variables;

  variables are on the LHS and the constant is on the RHS

$$\text{LHS} \leq \text{RHS}.$$

Why Solving LP Needs a New Approach

Basic calculus, we know how to solve a (simple) optimization problem like

$$\max_{x_1, x_2} \left\{ g(x_1, x_2) = 6x_1 - x_1^2 + x_2 - 2x_2^2 \right\},$$

by the use of the first-order condition

$$\frac{\partial g}{\partial x_1} = 6 - 2x_1 = 0 \quad \text{and} \quad \frac{\partial g}{\partial x_2} = 1 - 4x_2 = 0,$$

the optimal solution $x_1^* = 3,\ x_2^* = 1/4$ and $g(x_1^*, x_2^*) = 9\frac{1}{8}$.

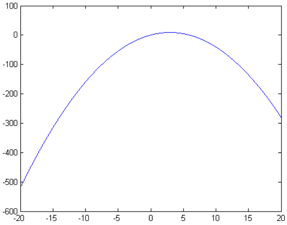
Try this approach on a linear objective function, e.g.,

$$g(x_1, x_2) = 6x_1 + x_2?$$

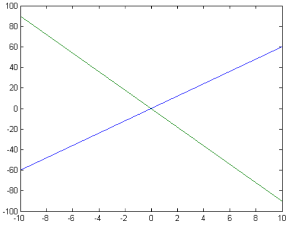
what happened, and why?

Linear versus Nonlinear Optimization

nonlinear function:
$g(x) = 6x - x^2$



linear function:
$g(x) = 6x$ or $g(x) = -9x$

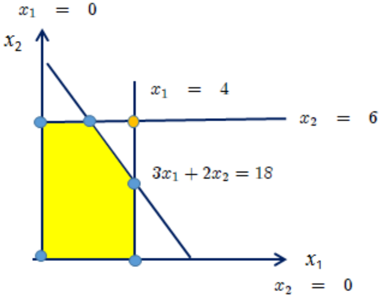


implications of the above comparisons:

- a meaningful LP problem must have constraints.
- optimal solution at the boundary of some constraints.

"Push to the Boundary" in the 2-Dimensional Case

$$\begin{aligned} \max \quad & \{2x_1 + x_2\} \\ \text{subject to} \quad & x_1 \leq 4, \\ & x_2 \leq 6, \\ & 3x_1 + 2x_2 \leq 18, \\ & x_1 \geq 0, \;\; x_2 \geq 0. \end{aligned}$$

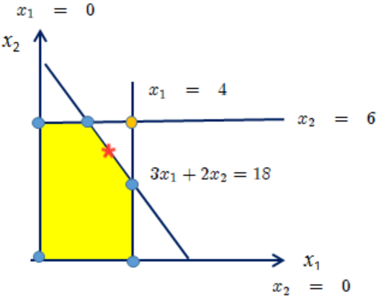


"Corner Feasible Point" (CFP) solutions

- $(x_1, x_2) = (0, 0) \quad \longrightarrow \quad 2x_1 + x_2 = 0 \quad (x_1 = 0, x_2 = 0).$
- $(x_1, x_2) = (0, 6) \quad \longrightarrow \quad 2x_1 + x_2 = 6 \quad (x_1 = 0, x_2 = 6).$
- $(x_1, x_2) = (2, 6) \quad \longrightarrow \quad 2x_1 + x_2 = 10 \quad (3x_1 + 2x_2 = 18, x_2 = 6).$
- $(x_1, x_2) = (4, 0) \quad \longrightarrow \quad 2x_1 + x_2 = 8 \quad (x_1 = 4, x_2 = 0).$
- $(x_1, x_2) = (4, 3) \quad \longrightarrow \quad 2x_1 + x_2 = 11 \quad (x_1 = 4, 3x_1 + 2x_2 = 18).$

What about Other Points at the "Boundary"

$$\begin{aligned} \max \quad & \{2x_1 + x_2\} \\ \text{subject to} \quad & x_1 \le 4, \\ & x_2 \le 6, \\ & 3x_1 + 2x_2 \le 18, \\ & x_1 \ge 0,\ x_2 \ge 0. \end{aligned}$$



Let point * in the figure be $\mathbf{x} = (x_1, x_2)$.

$\mathbf{x} = \alpha * (2,6) + (1-\alpha) * (4,3)$ i.e. $x_1 = 2\alpha + 4(1-\alpha)$, $x_2 = 6\alpha + 3(1-\alpha)$

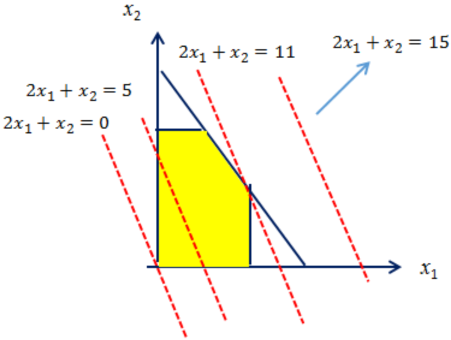
and

$2x_1 + x_2 = \alpha(2*2+6) + (1-\alpha)*(2*4+3) \le \max\{2*2+6, 2*4+3\}$.

so the objective value at * cannot be better than both CFP solutions.

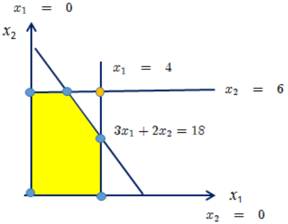
A Fundamental Fact about LP

If the optimal objective value of an LP exists and is finite, then this value can always be obtained at one of the "corner points" (solutions to *n* simultaneous equations).



Systems of Equations and Feasible Solutions

$$\begin{aligned} \max \quad & \{2x_1 + x_2\} \\ \text{subject to} \quad & x_1 \leq 4, \\ & x_2 \leq 6, \\ & 3x_1 + 2x_2 \leq 18, \\ & x_1 \geq 0,\ x_2 \geq 0. \end{aligned}$$



observation: to solve for $n = 2$ variables, we need a pair of equations.

question: with $3 + 2$ *(regular+non-negativity) constraints, how many pairs of equations can we possibly have?* 10.

question: why we only get five pairs of equations?

- some pairs of equations have no solution: e.g.,

$$x_1 = 0 \text{ and } x_1 = 4.$$

- some solutions violate other constraints: e.g.,

$$x_1 = 4 \text{ and } x_2 = 6 \text{ or } x_1 = 0 \text{ and } 3x_1 + 2x_2 = 18.$$

Possible Procedure

select n out of $n+m$ (m regular, n non-negativity) constraints to form a system of equations

- if the system has no solution, next!
- if the solution does not satisfy another constraint, next!
- if the solution is feasible for all constraints, evaluate the corresponding objective value.
- compare all objective values to pick the best solution.

Question: how many systems of equations we need to evaluate?

- $n = 2$ and $m + n = 10$?

$$\frac{10 \times 9}{2} = 45$$

- $n = 20$ and $m + n = 30$?

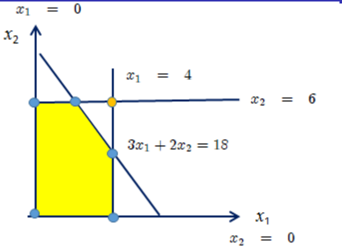
$$\frac{30 \times 29 \times .... \times 11}{20!} = \text{huge (about 30 million)}$$

- $n = 10k$ and $m + n = 30k$?

don't even think about it!

Searching Optimal Solutions

$$\begin{aligned} \max \quad & \{2x_1 + x_2\} \\ \text{subject to} \quad & x_1 \le 4, \\ & x_2 \le 6, \\ & 3x_1 + 2x_2 \le 18, \\ & x_1 \ge 0, \;\; x_2 \ge 0. \end{aligned}$$



$x_1 = 0, x_2 = 0, (2x_1 + x_2 = 0)$:

$$\text{neighbors:} \quad \begin{aligned} x_1 = 4, x_2 = 0 &\;\rightarrow\; 2x_1 + x_2 = 8 \\ x_1 = 0, x_2 = 6 &\;\rightarrow\; 2x_1 + x_2 = 6. \end{aligned}$$

$x_1 = 4, x_2 = 0, (2x_1 + x_2 = 8)$:

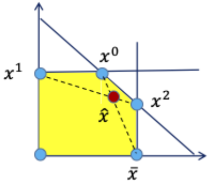
new neighbor: $x_1 = 4, 3x_1+2x_2 = 18 \rightarrow (x_1, x_2) = (4, 3) \rightarrow 2x_1+x_2 = 11.$

$x_1 = 4, x_2 = 3, (2x_1 + x_2 = 11)$:

new neighbor: $3x_1+2x_2 = 18, x_2 = 6 \rightarrow (x_1, x_2) = (2, 6) \rightarrow 2x_1+x_2 = 10.$

Illustration: Optimality of the Solution

write the objective function in vector form $c^T x = c_1 x_1 + \cdots c_n x_n$



x^0 is a corner solution and x^1 and x^2 are its neighbors, and

$$c^T x^0 \geq c^T x^1, \quad c^T x^0 \geq c^T x^2.$$

- if x^0 is not optimal, then there exists $\bar{x}$ such that $c^T \bar{x} > c^T x^0$.
- lines $x^1 - x^2$, and $x^0 - \bar{x}$ must intersect at some point $\hat{x}$, so

$$\hat{x} = \alpha x^1 + (1-\alpha)x^2 = \beta x^0 + (1-\beta)\bar{x}, \quad 0 < \alpha, \beta < 1.$$

- a contradiction:

$$c^T \hat{x} = \alpha c^T x^1 + (1-\alpha) c^T x^2 \leq c^T x^0 < \beta c^T x^0 + (1-\beta) c^T \bar{x} = c^T \hat{x}.$$

General Idea of SIMPLEX Method

1. **Fact 1:** The objective function of an LP is optimized by solutions to one of n simultaneous equations.
2. **Fact 2:** If a solution is not optimal, then it is always possible to get a better solution by swapping in/out one equation.

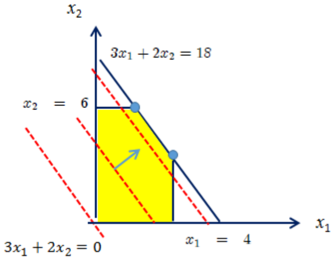
SIMPLEX procedure:

1. start from a feasible solution to a system of n equations, compare its objective value with that of its "adjacent solutions".
2. if the current solution is better than all its adjacent solutions, stop! the current solution is optimal.
3. otherwise, move to an adjacent solution with higher objective value.
4. repeat steps 2 and 3 until the optimal solution is found.

https://www.youtube.com/watch?v=x8rYyQjEMMs

Special Case (Non-unique Optimal Solutions)

$$
\begin{array}{rl}
\max & \{3x_1 + 2x_2\} \\
\text{subject to} & x_1 \le 4, \\
& x_2 \le 6, \\
& 3x_1 + 2x_2 \le 18, \\
& x_1 \ge 0, \ \ x_2 \ge 0.
\end{array}
$$

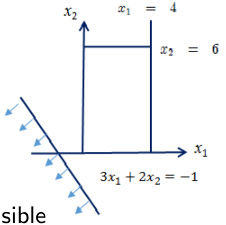


- every points on a segment of $3x_1 + 2x_2 = 18$ is an optimal solution.
- both $(2, 6)$ and $(4, 3)$ are optimal.

so we can still maximize the LP objective function by going through solutions of n simultaneous equations.

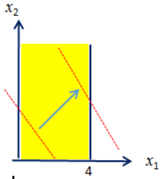
Exceptions (Don't solve, fix the model!)

$$\begin{aligned} \max \quad & \{2x_1 + x_2\} \\ \text{subject to} \quad & x_1 \leq 4,\ \ x_2 \leq 6, \\ & 3x_1 + 2x_2 \leq -1, \\ & x_1 \geq 0,\ \ x_2 \geq 0. \end{aligned}$$



infeasible

$$\begin{aligned} \max \quad & \{2x_1 + x_2\} \\ \text{subject to} \quad & x_1 \leq 4, \\ & x_1 \geq 0,\ \ x_2 \geq 0. \end{aligned}$$



unbounded