The L-Shaped Method

Operations Research

Anthony Papavasiliou

Contents

Table of Contents

Extensive Form 2-Stage Stochastic Linear Program

$$
\begin{aligned}
&(EF): \min c^T x + \sum_{k=1}^{K} p_k q_k^T y_k \\
&\text{s.t. } Ax = b \\
&T_k x + W y_k = h_k, k = 1, \ldots, K \\
&x \geq 0, y_k \geq 0, k = 1, \ldots, K
\end{aligned}
$$

- K realizations of random vector ξ, with probabilities $p_k, k = 1, \ldots, K$
- Randomness: q_k, h_k, T_k
- y_k: second-stage decision given realization k

Block Structure

Multi-stage decision making process

$$\begin{array}{cccc}
\boxed{A} & & & \\
\boxed{T_1} & \boxed{W} & & \\
\boxed{T_2} & & \boxed{W} & \\
\vdots & & & \\
\boxed{T_K} & \cdots & & \boxed{W}
\end{array}$$

Idea: ignore constraints of future stages

Master Problem

We know that

$$V(x) = \{\min \sum_{k=1}^{K} p_k q_k^T y_k | W y_k = h_k - T_k x, y_k \geq 0\}$$

is a *piecewise linear* function of x

Define **master problem** as

$$(M) : \min z = c^T x + \theta \quad (1)$$

$$\text{s.t. } Ax = b$$

$$D_l x \geq d_l, l = 1, \ldots, r \quad (2)$$

$$E_l x + \theta \geq e_l, l = 1, \ldots, s \quad (3)$$

$$x \geq 0, \theta \in \mathbb{R}$$

- Feasibility cuts: equation 2
- Optimality cuts: equation 3

Optimality Cuts

Consider a trial first-stage decision x^v

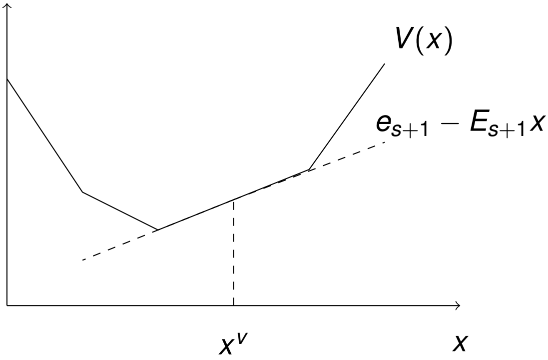
Let π_k^v be simplex multipliers of second-stage problem:

$$\min w = q_k^T y$$
$$\text{s.t. } Wy = h_k - T_k x^v$$
$$y \geq 0$$

$e_{s+1} - E_{s+1} x$ supports $V(x)$ at x^v, where

$$E_{s+1} = \sum_{k=1}^{K} p_k \cdot (\pi_k^v)^T T_k \quad (4)$$

$$e_{s+1} = \sum_{k=1}^{K} p_k \cdot (\pi_k^v)^T h_k \quad (5)$$


$V(x)$
$e_{s+1} - E_{s+1}x$
x^v
x

The L-Shaped Algorithm

- Step 0. Set $r = s = v = 0$
- Step 1. Solve master problem (M). Let (x^v, θ^v) be an optimal solution.
  If $s = 0$ (no optimality cuts), remove θ from (M) and set $\theta^0 = -\infty$
- Step 2. If $x \notin K_2$, add feasibility cut (equation 2) and return to Step 1. Otherwise, go to Step 3.
  $K_2 = \{x | \exists y : Wy = h_k - T_k x, y \geq 0, k = 1, \ldots, K\}$
- Step 3. Compute E_{s+1}, e_{s+1}. Let $w^v = e_{s+1} - E_{s+1} x^v$. If $\theta^v \geq w^v$, stop with x^v an optimal solution. Otherwise, set $s = s + 1$, add optimality cut to equation 3 and return to Step 1.

Table of Contents

Example 1

$$
\begin{aligned}
& z = \min 100x_1 + 150x_2 + \mathbb{E}_\xi(q_1 y_1 + q_2 y_2) \\
& \text{s.t. } x_1 + x_2 \leq 120 \\
& 6y_1 + 10y_2 \leq 60x_1 \\
& 8y_1 + 5y_2 \leq 80x_2 \\
& y_1 \leq d_1, y_2 \leq d_2 \\
& x_1 \geq 40, x_2 \geq 20, y_1, y_2 \geq 0
\end{aligned}
$$

$$
\xi = (d_1, d_2, q_1, q_2) = \begin{cases} (500, 100, -24, -28), & p_1 = 0.4 \\ (300, 300, -28, -32), & p_2 = 0.6 \end{cases}
$$

Iteration 1

- *Step 1.*
  $\min\{100x_1 + 150x_2 | x_1 + x_2 \leq 120, x_1 \geq 40, x_2 \geq 20\}$
- $x^1 = (40, 20)^T$, $\theta^1 = -\infty$
- *Step 3.* For $\xi = \xi_1$ solve

$$\min\{-24y_1 - 28y_2 | 6y_1 + 10y_2 \leq 2400, 8y_1 + 5y_2 \leq 1600$$

$$0 \leq y_1 \leq 500, 0 \leq y_2 \leq 100\}$$

$$w_1 = -6100, y^T = (137.5, 100), \pi_1^T = (0, -3, 0, -13)$$

For $\xi = \xi_2$ solve

$$\min\{-28y_1 - 32y_2 | 6y_1 + 10y_2 \leq 2400, 8y_1 + 5y_2 \leq 1600$$

$$0 \leq y_1 \leq 300, 0 \leq y_2 \leq 300\}$$

$$w_2 = -8384, y^T = (80, 192), \pi_2^T = (-2.32, -1.76, 0, 0)$$

Iteration 1: Optimality Cut

$h_1 = (0, 0, 500, 100)^T$, $h_2 = (0, 0, 300, 300)^T$
$T_{\cdot,1} = (-60, 0, 0, 0)^T$, $T_{\cdot,2} = (0, -80, 0, 0)^T$

- $e_1 = 0.4 \cdot \pi_1^T \cdot h_1 + 0.6 \cdot \pi_2^T \cdot h_2 = 0.4 \cdot (-1300) + 0.6 \cdot (0) = -520$
- $E_1 = 0.4 \cdot \pi_1^T T + 0.6 \cdot \pi_2^T T =$
  $0.4(0, 240) + 0.6(139.2, 140.8) = (83.52, 180.48)$
- $w^1 = -520 - (83.52, 180.48) \cdot x^1 = -7470.4$
- $w^1 = -7470.4 > \theta^1 = -\infty$, therefore add the cut
  $83.52x_1 + 180.48x_2 + \theta \geq -520$

Iteration 2

- *Step 1.* Solve master

$$\min\{100x_1 + 150x_2 + \theta | x_1 + x_2 \leq 120, x_1 \geq 40, x_2 \geq 20,$$
$$83.52x_1 + 180.48x_2 + \theta \geq -520\}$$
$$z = -2299.2, x^2 = (40, 80)^T, \theta^2 = -18299.2$$

- *Step 3.* Add the cut $211.2x_1 + \theta \geq -1584$

Iteration 3

- *Step 1.* Solve master.

$$z = -1039.375, x^3 = (66.828, 53.172)^T, \theta^3 = -15697.994$$

- *Step 3.* Add the cut $115.2x_1 + 96x_2 + \theta \geq -2104$

Iteration 4

- *Step 1.* Solve master.

$$z = -889.5, x^4 = (40, 33.75)^T, \theta^4 = -9952$$

- *Step 3.* There are multiple solutions for $\xi = \xi_2$. Select one, add the cut $133.44x_1 + 130.56x_2 + \theta \geq 0$

Iteration 5

- *Step 1.* Solve master

$$\min\{100x_1 + 150x_2 + \theta | x_1 + x_2 \leq 120, x_1 \geq 40, x_2 \geq 20,$$
$$83.52x_1 + 180.48x_2 + \theta \geq -520, 211.2x_1 + \theta \geq -1584$$
$$115.2x_1 + 96x_2 + \theta \geq -2104, 133.44x_1 + 130.56x_2 + \theta \geq 0\}$$
$$z = -855.833, x^5 = (46.667, 36.25)^T, \theta^5 = -10960$$

- *Step 3.* $w_5 = -520 - (83.52, 180.48) \cdot x^5 = -10960 = \theta^5$, stop. $x = (46.667, 36.25)^T$ is the optimal solution.

Example 2

$$z = \min \mathbb{E}_{\xi}(y_1 + y_2)$$

$$\text{s.t. } 0 \le x \le 10$$

$$y_1 - y_2 = \xi - x$$

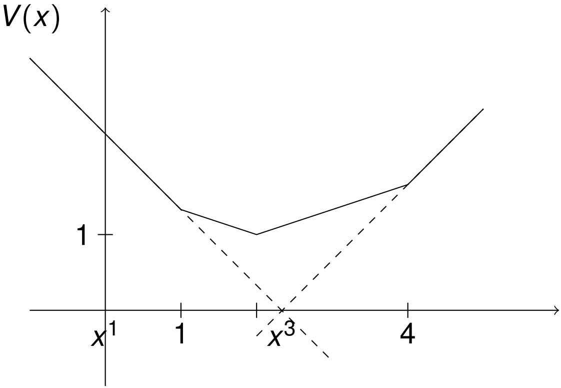
$$y_1, y_2 \ge 0$$

$$\xi = \begin{cases} 1 & p_1 = 1/3 \\ 2 & p_2 = 1/3 \\ 4 & p_3 = 1/3 \end{cases}$$

$K_2 = \mathbb{R}$

L-Shaped Method in Example 2

- Iteration 1, Step 1: $x^1 = 0$
- Iteration 1, Step 3: x^1 not optimal, add cut: $\theta \geq 7/3 - x$
- Iteration 2, Step 1: $x^2 = 10$
- Iteration 2, Step 3: x^2 not optimal, add cut: $\theta \geq x - 7/3$
- Iteration 3, Step 1: $x^3 = 7/3$
- Iteration 3, Step 3: x^3 not optimal, add cut: $\theta \geq (x + 1)/3$
- Iteration 4, Step 1: $x^4 = 1.5$
- Iteration 4, Step 3: x^4 not optimal, add cut: $\theta \geq (5 - x)/3$
- Iteration 3, Step 1: $x^5 = 2$
- Iteration 3, Step 3: x^5 is optimal



- $V(x^1) = 7/3$ and $V(x) = 7/3 - x$ 'around' x^1
- $(7 - x)/3$ is the optimality cut at x^1

Geometric Interpretation of Optimality Cuts

- $Q(x, \xi) = \min_y \{q(\omega)^T y | W(\omega) y = h(\omega) - T(\omega)x, y \geq 0\}$ yields dual optimal multipliers $\pi^T = q_B(\omega)^T \cdot B(\omega)^{-1}$
- From linear programming duality

$$Q(x, \xi) = q_B(\omega)^T \cdot B(\omega)^{-1}(h(\omega) - T(\omega)x)$$

- $B(\omega)$ is optimal basis as long as $y_B = B(\omega)^{-1}(h(\omega) - T(\omega)x) \geq 0$, $y_N = 0$ and $q_B(\omega)^T B(\omega)^{-1} \cdot W \leq q(\omega)^T$
- Taking expectation,

$$V(x) = \mathbb{E}_\xi \{q_B(\omega)^T \cdot B(\omega)^{-1}(h(\omega) - T(\omega)x)\}$$

  in the set $\cap_{\xi \in \Xi} \{x | B(\omega)^{-1}(h(\omega) - T(\omega)x) \geq 0\}$
- This is exactly the expression for optimality cuts at x^v

Table of Contents

Feasibility Cuts

Consider the following problem:

$$(F): \min w' = e^T v^+ + e^T v^-$$
$$\text{s.t. } Wy + Iv^+ - Iv^- = h_k - T_k x^v$$
$$y \geq 0, v^+ \geq 0, v^- \geq 0$$

with dual multipliers σ^v. Define

$$\begin{aligned} D_{r+1} &= (\sigma^v)^T T_k \\ d_{r+1} &= (\sigma^v)^T h_k \end{aligned}$$

Step 2 of L-shaped method: For $k = 1, \ldots, K$ solve (F).

- If $w' = 0$ for all k, go to Step 3.
- Else, add $D_{r+1} x \geq d_{r+1}$, set $r = r + 1$ and go to Step 1.

Example

$$\min 3x_1 + 2x_2 - \mathbb{E}_\xi(15y_1 + 12y_2)$$
$$\text{s.t. } 3y_1 + 2y_2 \leq x_1, 2y_1 + 5y_2 \leq x_2$$
$$0.8\xi_1 \leq y_1 \leq \xi_1, 0.8\xi_2 \leq y_2 \leq \xi_2$$
$$x, y \geq 0$$

$$\xi = \begin{cases} (4,4), p_1 = 0.25 \\ (4,8), p_2 = 0.25 \\ (6,4), p_3 = 0.25 \\ (6,8), p_4 = 0.25 \end{cases}$$

Generating a Feasibility Cut

For $x^1 = (0,0)^T$, $\xi = (6,8)^T$, solve

$$\begin{aligned}
&\min_{v^+,v^-,y} v_1^+ + v_1^- + v_2^+ + v_2^- + v_3^+ + v_3^- + \\
&v_4^+ + v_4^- + v_5^+ + v_5^- + v_6^+ + v_6^- \\
&\text{s.t. } v_1^+ - v_1^- + 3y_1 + 2y_2 \le 0,\ v_2^+ - v_2^- + 2y_1 + 5y_2 \le 0 \\
&v_3^+ - v_3^- + y_1 \ge 4.8,\ v_4^+ - v_4^- + y_2 \ge 6.4 \\
&v_5^+ - v_5^- + y_1 \le 6,\ v_6^+ - v_6^- + y_2 \le 8
\end{aligned}$$

We get $w' = 11.2$, $\sigma^1 = (-3/11, -1/11, 1, 1, 0, 0)$
$h = (0, 0, 4.8, 6.4, 6, 8)^T$, $T_{\cdot,1} = (-1, 0, 0, 0, 0, 0)^T$,
$T_{\cdot,2} = (0, -1, 0, 0, 0, 0)^T$
$D_1 = (-3/11, -1/11, 1, 1, 0, 0) \cdot T = (3/11, 1/11)$,
$d_1 = (-3/11, -1/11, 1, 1, 0, 0) \cdot h = 11.2$
$3/11 x_1 + 1/11 x_2 \ge 11.2$

Induced Constraints

Going by the book:

- Iteration 2 master problem: $x^2 = (41.067, 0)^T$
- Iteration 2 feasibility cut: $x_2 \geq 22.4$
- Iteration 3 master problem: $x^3 = (33.6, 22.4)^T$
- Iteration 3 feasibility cut: $x_2 \geq 41.6$
- Iteration 4 master problem: $x^4 = (27.2, 41.6)^T$ is feasible

Induced Constraints:

- Observe that for $\xi = (6, 8)^T$, $y_1 \geq 4.8$, $y_2 \geq 6.4$
- This implies $x_1 \geq 27.2$, $x_2 \geq 41.6$, which should be added directly to the master

Personal experience: feasibility cuts are impractical

Complete Recourse and Relative Complete Recourse

$\text{pos } W = \{z | z = Wy, y \geq 0\}$

$K_1 = \{x | Ax = b\}$, $K_2 = \{x | \exists y : Wy = h_k - T_k x, k = 1, \ldots, K\}$

Feasibility cuts are not necessary when we have:

- **Complete recourse**: $\text{pos } W = \mathbb{R}^{m_2}$
- **Relative complete recourse**: $K_2 \supseteq K_1$

Table of Contents

Proof of Convergence

Key observation: when we solve the master problem (M) subject to $x \in K_1 \cap K_2$ either

- $\theta^v \geq V(x^v)$ or
- $\theta^v < V(x^v)$

The first case is our termination criterion and implies optimality.

The second case can occur finitely many times.

$K_1 = \{x | Ax = b, x \geq 0\}$,

$K_2 = \{x | \exists y : Wy = h_k - T_k x, y \geq 0, k = 1, \ldots, K\}$

$\theta^v \geq V(x^v)$ Is Equivalent to the Termination Criterion

$$Q(x, \xi(\omega)) = \min_y \{q(\omega)^T y | Wy = h(\omega) - T(\omega)x, y \geq 0\}$$

- From linear programming duality, $Q(x^v, \xi_k) = (\pi_k^v)^T (h_k - T_k x^v)$
- Taking expectation, $V(x) = \mathbb{E}_\xi Q(x^v, \xi) = \sum_{k=1}^K p_k \cdot (\pi_k^v)^T (h_k - T_k x^v)$
- Our termination criterion is $\theta^v \geq w^v$, where

$$w^v = e_{s+1} - E_{s+1} x^v = (\sum_{k=1}^{K} p_k \cdot (\pi_k^v)^T h_k) - (\sum_{k=1}^{K} p_k \cdot (\pi_k^v)^T T_k) x^v$$

$\theta^v \geq V(x^v)$ Implies Optimality

- Fact: $Q(x, \xi)$ is convex in x, and π_k^v is a subgradient at x^v:

$$Q(x, \xi_k) \geq (\pi_k^v)^T h_k - (\pi_k^v)^T T_k x$$

- Taking expectation: $V(x) \geq \sum_{k=1}^K p_k \cdot (\pi_k^v)^T (h_k - T_k)x$
- Full problem can be written as

$$\begin{aligned} &\min c^T x + \theta \\ &\text{s.t. } V(x) \leq \theta \\ &x \in K_1 \cap K_2 \end{aligned}$$

  Therefore master (M) is a relaxation and gives lower bound
- If there were $\bar{x} \in K_1 \cap K_2$ with $c^T \bar{x} + V(\bar{x}) < c^T x^v + V(x^v)$ then x^v would not be optimal for (M) since (M) would be at most $c^T \bar{x} + V(\bar{x})$

$\theta^v < V(x^v)$ Occurs Finitely Many Times

- We have seen that each optimality cut is equivalent to $\theta \geq V(x^l) + \partial V(x^l)(x - x^l), l = 1, \ldots, s$
- Every time $\theta^v < V(x^v)$ occurs, the set of new multipliers $(\pi_k^v), k = 1, \ldots, K$ must be different from those generated previously
- Optimal multipliers π_k^v correspond to optimal bases of $\min_y \{q_k^T y | Wy = h_k - T_k x, y \geq 0\}$
- There are finitely many bases, therefore finitely many combinations of optimal multipliers $(\pi_k^v), k = 1, \ldots, K$

Finite Termination of Feasibility Cuts

Key observations:

1. Only a finite number of feasibility cuts can be generated
2. The feasibility cuts of equation 2 are **valid**

Valid feasibility cuts:

- remove current candidate x^{ν} from consideration
- do not remove any other candidates $x \in K_2$ from consideration

Finite Number of Possible Cuts

Recall feasibility cuts: $D_l x \geq d_l, l = 1, \ldots, r+1$, where

$$\begin{aligned} D_{r+1} &= (\sigma^v)^T T_k \\ d_{r+1} &= (\sigma^v)^T h_k \end{aligned}$$

and σ^v the dual multipliers of

$$(F) : \min w' = e^T v^+ + e^T v^-$$
$$\text{s.t. } Wy + Iv^+ - Iv^- = h_k - T_k x^v$$
$$y \geq 0, v^+ \geq 0, v^- \geq 0$$

- σ^v corresponds to a basis B of W: $\sigma^v = c_B^T B^{-1}$
- W has a finite number of bases

Valid Feasibility Cuts: x^v Is Cut Off

If $w' > 0$ then by linear programming duality $(\sigma^v)^T(h_k - T_k x^v) > 0$, therefore $D_{r+1}x \geq d_{r+1}$ removes x^v from consideration

Valid Feasibility Cuts: No Other $x \in K_2$ Is Cut Off

- Consider the following subset of $\mathbb{R}^{m_2}$:
  $\{z | z = h_k - T_k x, x \in K_2\}$
- This set is contained in $\mathrm{pos}\ W = \{z | z = Wy, y \geq 0\}$ by definition of K_2
- The half-space defined by $(\sigma^v)^T z \leq 0$ contains $\mathrm{pos}\ W$
  - To see this, note that the reduced cost of y must be ≥ 0:

    $$c_N^T - c_B^T B^{-1} N = -(\sigma^v)^T W \geq 0$$

    where $c_N^T = 0$, $(\sigma^v)^T = c_B^T B^{-1}$, $N = W$
  - Any element of $\mathrm{pos}\ W$ can be expressed as $z = Wy$, $y \geq 0$, hence from the previous inequality $(\sigma^v)^T z \leq 0$
- Therefore, $(\sigma^v)^T (h - Tx) \leq 0$ will be a valid feasibility cut

Feasibility Cuts

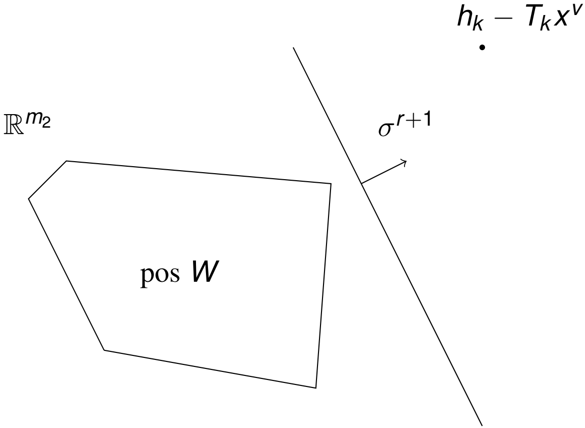


Table of Contents

Example: Capacity Expansion Planning

$$\min_{x,y\geq 0} \sum_{i=1}^{n} (I_i \cdot x_i + \mathbb{E}_\xi \sum_{j=1}^{m} C_i \cdot T_j \cdot y_{ij}(\omega))$$

$$\text{s.t.} \sum_{i=1}^{n} y_{ij}(\omega) = D_j(\omega), j = 1, \ldots, m$$

$$\sum_{j=1}^{m} y_{ij}(\omega) \leq x_i, i = 1, \ldots n$$

- I_i, C_i: fixed/variable cost of technology i
- $D_j(\omega)$, T_j: height/width of load block j
- $y_{ij}(\omega)$: capacity of i allocated to j
- x_i: capacity of i

Note: D_j is not dependent on ω

Problem Data

Two possible realizations of load duration curve:

- Reference scenario: 10%
- 10x wind scenario: 90%

	Duration (hours)	Level (MW) Reference scenario	Level (MW) 10x wind scenario
Base load	8760	0-7086	0-3919
Medium load	7000	7086-9004	3919-7329
Peak load	1500	9004-11169	7329-10315

Slave Problem

$$
\begin{aligned}
(S_\omega): \quad & \min_{y \geq 0} \sum_{i=1}^{n} \sum_{j=1}^{m} C_i \cdot T_j \cdot y_{ij} \\
(\lambda_j(\omega)): \quad & \sum_{i=1}^{n} y_{ij} = D_j(\omega), j = 1, \ldots, m \\
(\rho_i(\omega)): \quad & \sum_{j=1}^{m} y_{ij} \leq \bar{x}_i, i = 1, \ldots n
\end{aligned}
$$

where $\bar{x}$ has been fixed from the master problem

Sequence of Investment Decisions

Iteration	Coal (MW)	Gas (MW)	Nuclear (MW)	Oil (MW)
1	0	0	0	0
2	0	0	0	8736
3	0	0	0	15999.6
4	0	14675.5	0	0
5	10673.8	0	0	0
6	10673.8	0	0	13331.8
7	0	163.8	7174.5	3830.8
8	0	3300.6	7868.4	0
9	0	5143.4	7303.9	1679.4
10	3123.9	1948.1	4953.7	1143.3
11	1680	4322.4	6625	0
12	8747.6	1652.8	0	768.6
13	5701.9	464.9	4233.6	768.6
14	4935.9	1405	3994.7	0
15	6552.6	386.3	3173.7	882.9
16	5085	1311	3919	854

Observations

- Investment candidate in each iteration necessarily different from *all* past iterations
- 'Greedy' behavior: low capital cost in early iterations