Consider the **Two-stage stochastic LP with recourse**:

\[
\text{Minimize } c^T x + \sum_{k=1}^{K} p_k Q_k(x) \\
\text{subject to } x \in X
\]

where, for example, the feasible set of first-stage decisions is defined by

\[X = \{ x \in R^n : Ax = b, x \geq 0 \}\]

Here \(k\) indexes the finitely-many possible realizations of a random vector \(\xi\), with \(p_k\) the probability of realization \(k\).

The first-stage variables \(x\) are to be selected before \(\xi\) is observed.

Then the set of **second-stage decision variables** \(y_k\) are to be selected, after \(x\) has been selected and the \(k\)th realization of \(\xi\) is observed.

The cost of the second stage when scenario \(k\) occurs is

\[Q_k(x) = \min \{ q^T y : W y = b_k - T x, y \geq 0 \}\]

That is, \(y\) is a recourse which must be chosen so as to satisfy some linear constraints in the least costly way.

Note that, in general,
- the coefficient matrices \(T\) and \(W\),
- the right-hand-side vector \(b\), and
- the second-stage cost vector \(q\)
are all random.

The **deterministic equivalent LP** is a large-scale problem which simultaneously selects
- the first-stage variables \(x\) and
- the second-stage variables \(y_k\) for every realization \(k\)

\[
P: \text{Find } Z = \min c^T x + \sum_{k=1}^{K} p_k Q_k(x) \\
\text{subject to } T x + W y_k = b_k, k = 1, \ldots, K; \\
x \in X; \\
y_k \geq 0, k = 1, \ldots, K
\]

This can be an extremely large LP, with \(K n_2\) variables and \(K m_2\) constraints.

The equivalent "L-Shaped Method" was later introduced independently by van Slyke & Wets for solving stochastic LP with Recourse (**SLPwR**) problems:


*SIAM Journal of Applied Mathematics* 17: 638-663.

The central idea is to **partition** the variables into 2 sets

- **MIP**: integer and continuous variables
- **SLPwR**: 1st stage & 2nd stage variables

and to **project** the problem onto the first set of variables.

- **MIP**: \(\min c^T x + V(x)\)
  where \(V(x) = \) optimal value of continuous LP after integer variables \(x\) have been fixed.

- **SLPwR**: \(\min c^T x + Q(x)\)
  where \(Q(x) = \) minimal expected cost of second stage when first-stage variables \(x\) have been fixed.
Given a first-stage decision $x_i$, define a function $Q_i(x_i)$ equal to the optimum of the second stage for each scenario $k=1, \ldots, K$:

$$Q_i(x_i) = \min q_i y_i,$$

subject to

$$\sum_{k=1}^{K} T_k x_i + W_j y_i = \theta_i,$$

$$y_i \geq 0.$$

Then $P(x_i) = c x_i + \sum_{k=1}^{K} Q_i(x_i)$ provides us with an upper bound on the optimal value $Z$.

If, as before, we introduce the variables $x_i$ for each scenario $k$, together with the nonanticipativity constraints, we obtain the second-stage problem for scenario $k$,

Minimize $q_i y_i$

subject to

$$\sum_{k=1}^{K} T_k x_i + W_j y_i = \theta_i,$$

$$x_i = 0,$$

$$y_i \geq 0$$

whose linear programming dual is the linear program

Maximize $\lambda \theta_i + \pi \pi_i$

subject to

$$\pi_i T_k + \lambda \lambda_i = 0,$$

$$\pi_i W \leq \pi_i.$$%  

It is not necessary to introduce the variables $x_i$, but it is done in anticipation of later defining a cross-decomposition algorithm, which is a hybrid of Benders’ decomposition and Lagrangian relaxation.

Denote by $\Pi_i = \{ \pi \geq W_i \pi_i \leq q_i \}$ the polyhedral feasible region of the second-stage problem for scenario $k$. Denote by $\pi_i$ the $i$th extreme point of $\Pi_i$, $i = 1, 2, \ldots, K$.

By enumerating the large (but finite) number of extreme points of $\Pi_i$, we can write

$$Q_i(x_i) = \max_{\pi_i \in \Pi_i} \pi_i (\theta_i - T_i x_i),$$

where $\lambda_i = -\pi_i \theta_i$ and $\alpha_i = \pi_i \theta_i$.

(Note that this demonstrates that $Q_i(x_i)$ is a piecewise-linear convex function.)

However, if a subset of the dual extreme points of $\Pi_i$ are available, e.g., $\pi_i^*, i=1, \ldots, M$, where $M \ll K$, then we obtain an underestimate of $Q_i(x_i)$, which we denote by

$$Q_i(x_i) \leq \max_{\pi_i \in \Pi_i^*} \pi_i (\theta_i - T_i x_i).$$

Thus, by making use of dual information obtained after $M$ evaluations of $Q_i(x_i)$, we obtain a Partial Master Problem,

$$\Phi_m = \min c x_i + \sum_{k=1}^{K} p_k \theta_i,$$

subject to $x_i \in X$, and

$$\theta_i \geq \lambda_i x_i + \alpha_i, i=1, \ldots, M, k=1, \ldots, K.$$%  

which provides a lower bound on the solution of $Z$.

Benders’ (Complete) Master Problem then uses this representation of $Q_i(x_i)$ to provide an alternate method for evaluating $Z$, namely

$$Z = \min c x_i + \sum_{k=1}^{K} p_k \theta_i,$$

subject to $x_i \in X$, and

$$\theta_i \geq \lambda_i x_i + \alpha_i, i=1, \ldots, M, k=1, \ldots, K.$$%  

While it is possible in principle to solve the problem using Benders’ Complete Master Problem, in practice the magnitude of the number of dual extreme points makes it prohibitively expensive.

Benders’ algorithm solves the current Partial Master Problem, obtaining

• $x_i$ (a ‘trial solution’) and
• an underestimate $\sum_{k=1}^{K} p_k Q_i(x_i)$ of the associated expected second-stage cost.

The actual expected second-stage cost, i.e., $\sum_{k=1}^{K} p_k Q_i(x_i)$, is then evaluated by solving the second-stage problem for each scenario. Additional constraints are added to the Partial Master Problem to complete the iteration.

At each iteration of Benders’ algorithm, then,

• the subproblem solution

$$P(x_i) = c x_i + \sum_{k=1}^{K} p_k Q_i(x_i)$$

provides an upper bound for $Z$, and

• the Partial Master Solution

$$\Phi_m = P(x_i) = c x_i + \sum_{k=1}^{K} p_k Q_i(x_i)$$

provides a lower bound for $Z$.  

We can eliminate $\lambda_i$ (the dual variables for the constraint $x_i = x_i$) by using the equality constraint to obtain

$$\lambda_i = -\pi_i \theta_i$$

and

$$Q_i(x_i) = \max (\theta_i - T_i x_i) \pi_i$$

subject to

$$\pi_i W \leq \pi_i.$$%  

The original problem now reduces to

$$Z = \min c x_i + \sum_{k=1}^{K} p_k Q_i(x_i).$$

Benders’ (Complete) Master Problem then uses this representation of $Q_i(x_i)$ to provide an alternate method for evaluating $Z$, namely

$$Z = \min c x_i + \sum_{k=1}^{K} p_k \theta_i,$$

subject to $x_i \in X$, and

$$\theta_i \geq \lambda_i x_i + \alpha_i, i=1, \ldots, M, k=1, \ldots, K.$$%  

While it is possible in principle to solve the problem using Benders’ Complete Master Problem, in practice the magnitude of the number of dual extreme points makes it prohibitively expensive.
Benders’ Algorithm— ‘Uni-cut’ Version

In the uni-cut version, at each iteration i the K constraints

\[
\theta \geq \sum_{k=1}^{K} \left[ l_k x_k + \alpha_k \right], \quad i=1, \ldots, I
\]

are aggregated before adding them to the Partial Master

Problem:

\[
Z = \min c x + \theta
\]

subject to \( x \in X \), and

\[
\theta \geq \sum_{k=1}^{K} \left[ l_k x_k + \alpha_k \right], \quad i=1, \ldots, I
\]

Generally, more iterations are required, but there are fewer cuts (& less computation) in each Partial Master Problem.

Benders’ algorithm is as follows:

\textbf{Step 0.} Select an arbitrary \( x_0 \in X \). Initialize the upper bound \( Z = +\infty \) and lower bound \( Z = -\infty \).

Note: This allows the user to make use of knowledge about his/her problem by using an initial “guess” at the solution.

Another alternative is to solve the Expected Value LP problem to obtain the initial \( x_0 \):

\[
\text{Minimize } c x
\]

subject to \( A x = b, \quad \sum_{k=1}^{K} p_k x_k = \sum_{k=1}^{K} p_k h_k \)

\( x \geq 0 \)

\textbf{Step 1a.} Solve the primal subproblems to evaluate \( P(x_k) \) and the optimal dual variables \( \pi_k \), \( k=1, \ldots, K \) and compute \( P(x_0) \).

\textbf{1b.} For each scenario, generate an optimality cut.

\textbf{1c.} Uni-cut version: Aggregate the K optimality cuts and add to Benders’ master problem.

\textbf{Multi-cut version:} Add each of the K optimality cuts to Benders’ master problem.

\textbf{1d.} Update the upper bound, \( Z = \min \{ Z, P(x_k) \} \).

\textbf{1e.} If \( Z - Z \leq \varepsilon \), STOP; else continue to Step 2.

\textbf{Step 2a.} Solve the Partial Master Problem to obtain

• an optimal \( x_0 \), and

• an underestimate \( P(x_0) = c x_0 + \sum_{k=1}^{K} p_k P(x_k) \) of the expected cost \( P(x_0) \).

\textbf{2b.} Update the lower bound, \( Z = \max \{ Z, P(x_0) \} \).

\textbf{2c.} If \( Z - Z \leq \varepsilon \), STOP; else return to Step 1a.

At each iteration, the number of constraints (and therefore the size of the basis) of the Partial Master Problem increases, adding to the computational burden.

Furthermore, because constraints have been added, the solution of each partial master problem is generally infeasible in the partial master problem which follows.

For these reasons, it is preferable to solve the dual of the partial master problem, which is formed by appending a column to the dual of the previous partial master problem, so that the solution of the dual of the previous Partial Master Problem may serve as an initial basic feasible solution for the Partial Master Problem which follows.

If \( X = \{ x : A x = b, x \geq 0 \} \), the linear programming dual of Benders’ Partial Master problem is

\[
\Phi_u = \max b u + \sum_{i=1}^{M} \sum_{j=1}^{N_i} \left( u_i v_{ij} \right)
\]

subject to \( A^T u - \sum_{i=1}^{M} \sum_{j=1}^{N_i} v_{ij} = c \)

\[
\sum_{j=1}^{N_i} v_{ij} = p_i, \quad k=1, \ldots, K
\]

\[
v_{ij} \geq 0, \quad i=1, \ldots, M; k=1, \ldots, K
\]

(The dual variable \( u \) is

unrestricted in sign if \( X \) is defined by \( A x = b \), but

nonnegative if \( A x \geq b \), and

nonpositive if \( A x \leq b \).)

It can be shown that, in fact, this dual of Benders’ Master Problem is identical to the Master Problem of Dantzig-Wolfe decomposition applied to the original large-scale deterministic equivalent LP!