"A Cross-Decomposition Algorithm for Two-Stage Stochastic Linear Programming with Recourse"
Abstract: We consider a paradigm of linear optimization in the face of uncertainty, in which (first-stage) decisions must be made before the uncertainty is resolved, and then recourse (second-stage decisions) is available to compensate. When a finite set of scenarios can be identified and their probability estimated, and the objective is to minimize the sum of the first-stage cost and the expected value of the second-stage cost, a (generally large) deterministic equivalent LP problem can be constructed. Benders' (primal) decomposition and Lagrangian (dual) decomposition each yields a family of smaller subproblems, one for each scenario, and a coordinating "master" problem. Cross-decomposition is a hybrid primal-dual iterative approach which eliminates the master problems and uses the primal and dual subproblems to provide both upper and lower bounds on the optimal expected cost at each iteration. A small example illustrates the computation.
EXAMPLE

- A farmer raises wheat, corn, and sugar beets on 500 acres of land. Before the planting season he wants to decide how much land to devote to each crop.
- At least 200 tons of wheat and 240 tons of corn are needed for cattle feed, which can be purchased from a wholesaler if not raised on the farm.
- Any grain in excess of the cattle feed requirement can be sold at $170 and $150 per ton of wheat and corn, respectively.
- The wholesaler sells the grain for 40% more (namely $238 and $210 per ton, respectively.)
- Up to 6000 tons of sugar beets can be sold for $36 per ton; any additional amounts can be sold for $10/ton.
<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Yield</td>
<td>2.5 T/Acre</td>
<td>3 T/Acre</td>
<td>20 T/Acre</td>
</tr>
<tr>
<td>Planting cost</td>
<td>$150/Acre</td>
<td>$230/Acre</td>
<td>$260/Acre</td>
</tr>
<tr>
<td>Selling price</td>
<td>$170/T</td>
<td>$150/T</td>
<td>$36/T first 6000T $10/T otherwise</td>
</tr>
<tr>
<td>Purchase price</td>
<td>$238/T</td>
<td>$210/T</td>
<td></td>
</tr>
<tr>
<td>Minimum Rqmt</td>
<td>200T</td>
<td>240T</td>
<td></td>
</tr>
</tbody>
</table>
DECISION VARIABLES

We distinguish between two types of decisions:

First stage (before growing season):

\[ x_1 = \text{acres of land planted in wheat} \]
\[ x_2 = \text{acres of land planted in corn} \]
\[ x_3 = \text{acres of land planted in beets} \]

Second stage (after harvest):

\[ w_1 = \text{tons of wheat sold} \]
\[ w_2 = \text{tons of corn sold} \]
\[ w_3 = \text{tons of beets sold at } \$36/\text{T} \]
\[ w_4 = \text{tons of beets sold at } \$10/\text{T} \]
\[ y_1 = \text{tons of wheat purchased} \]
\[ y_2 = \text{tons of corn purchased} \]
LINEAR PROGRAMMING MODEL

Minimize $150x_1 + 230x_2 + 260x_3 + 238y_1 - 170w_1 + 210y_2 - 150w_2 - 36w_3 - 10w_4$

subject to

\[ x_1 + x_2 + x_3 \leq 500 \]
\[ 2.5x_1 + y_1 - w_1 \geq 200 \]
\[ 3x_2 + y_2 - w_2 \geq 240 \]
\[ w_3 + w_4 \leq 20x_3 \]
\[ w_3 \leq 6000 \]

\[ x_i \geq 0, \ i=1,2,3; \ y_i \geq 0, \ i=1,2; \ w_i \geq 0, \ i=1,2,3,4 \]
OPTIMAL SOLUTION

Profit = $118,600

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant</td>
<td>120 Acres</td>
<td>80 Acres</td>
<td>300 Acres</td>
</tr>
<tr>
<td>Yield</td>
<td>300T</td>
<td>240T</td>
<td>6000T</td>
</tr>
<tr>
<td>Sales</td>
<td>100T</td>
<td>--</td>
<td>6000T</td>
</tr>
<tr>
<td>Purchase</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
In actuality, crop yields are uncertain, depending upon weather conditions during the growing season.

Three scenarios have been identified:

- "good" (20% higher than average)
- "fair" (average)
- "bad" (20% below average),

each equally likely:

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Wheat yield (tons/acre)</th>
<th>Corn yield (tons/acre)</th>
<th>Beet yield (tons/acre)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Good</td>
<td>3</td>
<td>3.6</td>
<td>24</td>
</tr>
<tr>
<td>2. Fair</td>
<td>2.5</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>3. Bad</td>
<td>2</td>
<td>2.4</td>
<td>16</td>
</tr>
</tbody>
</table>
Scenario #1: "Good" Yield: Optimal Profit = $167,667

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant</td>
<td>183.333 Acres</td>
<td>66.67 Acres</td>
<td>250 Acres</td>
</tr>
<tr>
<td>Yield</td>
<td>550T</td>
<td>240T</td>
<td>6000T</td>
</tr>
<tr>
<td>Sales</td>
<td>350T</td>
<td>--</td>
<td>6000T</td>
</tr>
<tr>
<td>Purchase</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Scenario #3: "Bad" Yield: Optimal Profit = $59,950

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant</td>
<td>100 Acres</td>
<td>25 Acres</td>
<td>375 Acres</td>
</tr>
<tr>
<td>Yield</td>
<td>200T</td>
<td>60T</td>
<td>6000T</td>
</tr>
<tr>
<td>Sales</td>
<td>--</td>
<td>--</td>
<td>6000T</td>
</tr>
<tr>
<td>Purchase</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

If a perfect forecast was available, then, the expected profit would be

\[
\frac{1}{3} \times 167,667 + \frac{1}{3} \times 118,600 + \frac{1}{3} \times 59,950 = 115,406
\]
The stochastic decision problem is to optimize the first-stage cost plus the *expected* second-stage costs:

Minimize \[ 150x_1 + 230x_2 + 260x_3 + \frac{1}{3} \sum_{k=1}^{3} Q_k(x) \]

subject to \[ x_1 + x_2 + x_3 \leq 500 \]
\[ x_j \geq 0, \ j=1,2,3 \]

where

\[ Q_k(x) = \text{second-stage costs in scenario } k, \text{ if first-stage decisions } x \text{ have been implemented} \]
\[ Q_1(x) = \text{Minimum} \ 170w_1 + 150w_2 + 36w_3 + 10w_4 - 238y_1 - 210y_2 \]
\[ \text{s.t.} \ y_1 - w_1 \geq 200 - 3x_1 \]
\[ y_2 - w_2 \geq 240 - 3.6x_2 \]
\[ w_3 + w_4 \leq 24x_3 \]
\[ y_1 \geq 0, \ y_2 \geq 0, \ w_1 \geq 0, \ w_2 \geq 0, \ 0 \leq w_3 \leq 6000, \ w_4 \geq 0 \]

\[ Q_2(x) = \text{Minimum} \ 170w_1 + 150w_2 + 36w_3 + 10w_4 - 238y_1 - 210y_2 \]
\[ \text{s.t.} \ y_1 - w_1 \geq 200 - 2.5x_1 \]
\[ y_2 - w_2 \geq 240 - 3x_2 \]
\[ w_3 + w_4 \leq 20x_3 \]
\[ y_1 \geq 0, \ y_2 \geq 0, \ w_1 \geq 0, \ w_2 \geq 0, \ 0 \leq w_3 \leq 6000, \ w_4 \geq 0 \]

\[ Q_3(x) = \text{Minimum} \ 170w_1 + 150w_2 + 36w_3 + 10w_4 - 238y_1 - 210y_2 \]
\[ \text{s.t.} \ y_1 - w_1 \geq 200 - 2x_1 \]
\[ y_2 - w_2 \geq 240 - 2.4x_2 \]
\[ w_3 + w_4 \leq 16x_3 \]
\[ y_1 \geq 0, \ y_2 \geq 0, \ w_1 \geq 0, \ w_2 \geq 0, \ 0 \leq w_3 \leq 6000, \ w_4 \geq 0 \]
TWO-STAGE LINEAR PROGRAMMING WITH RECOURSE

Minimize $z = cx + E \left[ \min q(\omega) y(\omega) \right]$ 

subject to 

$Ax = b$

$T(\omega)x + Wy(\omega) = h(\omega)$,

$x \geq 0, y(\omega) \geq 0$

where

$x$ = first-stage decision

and

$y(\omega)$ = second-stage decision after random event $\omega$ is observed

which must satisfy the second-stage constraints

$T(\omega)x + Wy(\omega) = h(\omega)$,

where $q(\omega), T(\omega) & h(\omega)$ are random variables
DETERMINISTIC EQUIVALENT PROBLEM

Assume a finite number of scenarios.

For each scenario $k$, define a set of second-stage variables, $y^k$, and arrays $T_k$, $q_k$, and $h_k$

The objective is to minimize the expected total costs of first and second stages

$$\text{Minimize } cx + \sum_{k=1}^{K} p_k Q_k(x)$$

subject to $x \in X$

where the cost of the second stage is

$$Q_k(x) = \text{Minimum } \left\{ q_k y : W y = h_k - T_k x, \ y \geq 0 \right\}$$
Consider the *deterministic LP* derived from the 2-stage stochastic LP:

\[
Z = \min cx + \sum_{k=1}^{K} p_k q_k y^k
\]

subject to

\[
T_k x + W y^k = h_k, k = 1, \ldots, K;
\]

\[
x \in X
\]

\[
y^k \geq 0, k = 1, \ldots, K
\]

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

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

Second stage decisions:

For each scenario k (k=1,2,3), define a set of decision variables:

\[ w_1^k = \text{tons of wheat sold} \]
\[ w_2^k = \text{tons of corn sold} \]
\[ w_3^k = \text{tons of beets sold at $36/T} \]
\[ w_4^k = \text{tons of beets sold at $10/T} \]
\[ y_1^k = \text{tons of wheat purchased} \]
\[ y_2^k = \text{tons of corn purchased} \]
**DETERMINISTIC EQUIVALENT LP:**

Minimize $150x_1 + 230x_2 + 260x_3 + \frac{1}{3}(238y_1^1 - 170w_1^1 + 210y_2^1 - 150w_2^1 - 36w_3^1 - 10w_4^1) + \frac{1}{3}(238y_1^2 - 170w_1^2 + 210y_2^2 - 150w_2^2 - 36w_3^2 - 10w_4^2) + \frac{1}{3}(238y_1^3 - 170w_1^3 + 210y_2^3 - 150w_2^3 - 36w_3^3 - 10w_4^3)$

subject to

$x_1 + x_2 + x_3 \leq 500$

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3x_1 + y_1^1 - w_1^1 \geq 200$</td>
<td>$2.5x_1 + y_1^2 - w_1^2 \geq 200$</td>
<td>$2x_1 + y_1^3 - w_1^3 \geq 200$</td>
</tr>
<tr>
<td>$3.6x_2 + y_2^1 - w_2^1 \geq 240$</td>
<td>$3x_2 + y_2^2 - w_2^2 \geq 240$</td>
<td>$2.4x_2 + y_2^3 - w_2^3 \geq 240$</td>
</tr>
<tr>
<td>$24x_3 - w_3^1 - w_4^1 \geq 0$</td>
<td>$20x_3 - w_3^2 - w_4^2 \geq 0$</td>
<td>$16x_3 - w_3^3 - w_4^3 \geq 0$</td>
</tr>
<tr>
<td>$w_3^1 \leq 6000$</td>
<td>$w_3^2 \leq 6000$</td>
<td>$w_3^3 \leq 6000$</td>
</tr>
</tbody>
</table>

$x_i \geq 0, i=1,2,3$;

$y_i^k \geq 0, i=1,2 \& k=1,2,3$;

$w_i^k \geq 0, i=1,2,3,4 \& k=1,2,3$

*Thus, all possible second-stage decisions are made simultaneously, in a single large LP.*
Optimal Solution: Expected profit = $108,390

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>First stage</td>
<td>Plant: 170 Acres</td>
<td>80 Acres</td>
<td>250 Acres</td>
</tr>
<tr>
<td>k=1 &quot;Good yield&quot;</td>
<td>Yield 510 T</td>
<td>288 T</td>
<td>6000 T</td>
</tr>
<tr>
<td></td>
<td>Sales 310 T</td>
<td>48 T</td>
<td>6000 T</td>
</tr>
<tr>
<td></td>
<td>Purchase --</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>k=2 &quot;Fair yield&quot;</td>
<td>Yield 425 T</td>
<td>240 T</td>
<td>5000 T</td>
</tr>
<tr>
<td></td>
<td>Sales 225 T</td>
<td>--</td>
<td>5000 T</td>
</tr>
<tr>
<td></td>
<td>Purchase --</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>k=3 &quot;Bad yield&quot;</td>
<td>Yield 340 T</td>
<td>192 T</td>
<td>4000 T</td>
</tr>
<tr>
<td></td>
<td>Sales 140 T</td>
<td>--</td>
<td>4000 T</td>
</tr>
<tr>
<td></td>
<td>Purchase --</td>
<td>48 T</td>
<td>--</td>
</tr>
</tbody>
</table>

♦ Using the original solution (where expected values of yields were assumed, i.e., planting 120 acres of wheat, 80 acres of corn, & 300 acres of beets) his expected profit would be $107,240 (which is $1,150 less than the optimal expected value).

♦ The Expected Value of Perfect Information is $115,406 - $108,390 = $7016
LAGRANGIAN DECOMPOSITION:

"SPLITTING" FIRST-STAGE VARIABLES

For each scenario $k$, define a first-stage decision $x^k$ which must equal the original first-stage decision (which we now denote by $x^0$). We can then write the equivalent LP:

$$Z = \min cx_0 + \sum_{k=1}^{K} p_k q_k y^k$$

subject to

$$x^0 \in X$$

In order to separate the LP by scenario, we need to "relax" the constraints

$$x^0 = x^k, k = 1, \ldots, K;$$
LAGRANGIAN RELAXATION

Given a family of Lagrangian multiplier vectors $\lambda_k$, $k=1, \ldots K$, we define the relaxation:

$$D(\lambda) = \min cx^0 + \sum_{k=1}^K p_k q_k y^k + \sum_{k=1}^K \lambda_k (x^k - x^0)$$

subject to

$$x^0 \in X$$

$$T_k x^k + W y^k = h_k, k = 1, \ldots K;$$

$$x^k \geq 0, k = 1, \ldots K; y^k \geq 0, k = 1, 2, \ldots K$$

That is,

$$D(\lambda) = \min \left( c - \sum_{k=1}^K \lambda_k \right) x^0 + \sum_{k=1}^K \left[ \lambda_k x^k + p_k q_k y^k \right]$$

subject to the above constraints.
This is motivated by the fact that the problem then separates into $K+1$ subproblems:

$$D(\lambda) = D_0(\lambda_1, \ldots, \lambda_K) + \sum_{k=1}^{K} D_k(\lambda_k)$$

where

$$D_0(\lambda) = \min \left( c - \sum_{k=1}^{K} \lambda_k \right) x^0$$

subject to $x^0 \in X$

and, for each $k = 1, \ldots, K$:

$$D_k(\lambda) = \min \lambda_k x^k + p_k q_k y^k$$

subject to $T_k x^k + W y^k = h_k$

$$x^k \geq 0, y^k \geq 0$$
### Dual Subproblem 0

**for 1st Stage**

\[
\min \left( c - \sum_{k=1}^{K} \lambda_k \right) x^0
\]

subject to \( x^0 \in X \)

### Dual Subproblem for Scenario k, k=1, …K

\[
\text{Min } \lambda_k x^k + p_k q_k y^k
\]

subject to

\[
T_k x^k + W y^k = h_k
\]

\[
x^k \geq 0, y^k \geq 0
\]

The value \( D(\lambda) = D_0(\lambda_1, \ldots \lambda_K) + \sum_{k=1}^{K} D_k(\lambda_k) \) provides a lower bound on the optimal cost \( Z \).

The *Lagrangian dual* problem is to select the multipliers which will produce the tightest such lower bound:

\[
\hat{D} = \max_{\lambda} D(\lambda)
\]

*Note:* In the linear case, \( \hat{D} = Z \) and there is no "duality gap".
Master Problem: Adjust multipliers $\lambda$

Lagrangian Subproblems: $D_k(\lambda), k=0,1,...,K$

Converged?

- Yes → STOP
- No → Discrepancies $x_0 - x_k$
BENDERS' DECOMPOSITION

Benders' partitioning (commonly known in stochastic programming as the "L-Shaped Method") achieves separability of the second stage decisions, but in a different manner. Given a first-stage decision $x^0$, solve for each scenario $k=1, \ldots, K$ the second-stage LP:

$$P_k\left(x^0\right) = \min q_k y^k$$

subject to $W y^k = h_k - T_k x^0, \ y^k \geq 0$

Then $P\left(x^0\right) = cx^0 + \sum_{k=1}^{k} p_k P_k\left(x^0\right)$ provides us with an upper bound on the optimal cost $Z$, i.e.,

$$D(\lambda) \leq Z \leq P\left(x^0\right)$$
Furthermore, solving each LP provides us with a vector $\lambda_k$ of dual variables corresponding to the constraints $x^0 = x^k$.

If $\pi_k$ is the dual solution of the LP

$$P_k\left(x^0\right) = \min q_k y^k$$

subject to $Wy^k = h_k - T_k x^0, y^k \geq 0$

then $\lambda_k = -T_k^T \pi_k$
An aside: Computing $\lambda_k$:

The dual of

$$\begin{align*}
\text{Min } q_k y^k \\
\text{subject to } & \\
T_k x^k + W y^k = h_k, \\
x^k = x^0, \\
x^k \geq 0
\end{align*}$$

is the LP

$$\begin{align*}
\text{Max } h_k \pi_k + x^0 \lambda_k \\
\text{subject to: } & \\
T_k^T \pi_k + I \lambda_k = 0 \\
W^T \pi_k \leq q_k
\end{align*}$$

If we eliminate $\lambda_k$ using the equality constraint, we obtain $\lambda_k = -T_k^T \pi_k$ and the dual LP

$$\begin{align*}
\text{Max } (h_k - T_k x^0) \pi_k \\
\text{subject to } & \\
W^T \pi_k \leq q_k
\end{align*}$$
The original problem now is seen to be equivalent to

$$\text{Min } cx^0 + \sum_{k=1}^{K} p_k P_k(x^0)$$

subject to $x^0 \in X$

By making use of dual information obtained after $M$ evaluations of $P_k(x^0)$, Benders' procedure forms an approximation (a convex piecewise-linear function) of $P_k(x^0)$:

$$P_k(x^0) \geq \max_{i=1, \ldots, M} \left\{ \alpha^i_k x^0 + \beta^i_k \right\}$$

so that the original problem reduces (with introduction of new variables $\theta_k$) to

$$\text{Min } cx^0 + \sum_{k=1}^{K} p_k \theta_k$$

subject to $x^0 \in X$

and

$$\theta_k \geq \alpha^i_k x^0 + \beta^i_k, \quad i=1, \ldots, M; \quad k=1, \ldots, K$$
That is, we have approximated $P_k(x^0)$ by the maximum of a finite number of linear functions, i.e., by a \textit{piecewise-linear convex function}:
Benders’ Master Problem:
Select first-stage decisions $x^0$

Benders’ Subproblems:
Solve $P_k(x^0)$, $k=1,2,...,K$

Converged?

Yes

STOP

No

Lagrangian multipliers $\lambda$
In either the Lagrangian relaxation approach or Benders' decomposition, the burden of the computation lies in the respective master problems: searching for the optimal $\lambda$ in the case of Lagrangian relaxation, & searching for the optimal $x^0$ in the case of Benders' decomposition.

The subproblems, being LPs separable by scenario, are easily solved in comparison.
CROSS-DECOMPOSITION

Cross-decomposition is a hybrid of Benders' decomposition and Lagrangian relaxation, in which the subproblem of each algorithm serves the purpose of the master problem of the other.

That is, Benders' subproblem receives the first-stage decisions $x^0$ from the Dual subproblem $D_0$ rather than from the Benders' master problem.

Likewise, the Dual subproblem $D_0$ receives the necessary multipliers $\lambda$ from the Benders' subproblem, rather than from the Dual master problem.
CROSS-DECOMPOSITION

Note that the algorithm can be "streamlined"-- only one of the dual subproblems $D_0(\lambda)$ needs to be solved at each iteration, except when the termination criterion

$$P(x^0) - D(\lambda) \leq \varepsilon$$

is to be tested.
Convergence is improved if the mean of all previously generated Lagrangian multipliers and first-stage decisions are sent to the Lagrangian and Benders' subproblems, respectively.
**EXAMPLE**

The cross-decomposition algorithm described above was implemented in the APL language (APL+WIN 3.0). First, the mean of all prior primal & dual solutions was used at each iteration. The result after 100 iterations was as follows:

<table>
<thead>
<tr>
<th>Total cost: ( -106456.94 ), found at iteration #72</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best lower bound: ( -110752.17 )</td>
</tr>
<tr>
<td>Gap= 4295.23, or 4.03%</td>
</tr>
<tr>
<td>Stage One Variables:</td>
</tr>
<tr>
<td>i</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>
The plot of upper & lower bounds at each iteration:

![Graph showing cost vs iterations with labels for UB and LB]
The mean values of first-stage variables used in the primal subproblems at each iteration.
As an alternative, exponential smoothing (with smoothing factor 10%) was used for both primal and dual solutions. After 100 iterations, the following was the best solution found:

Total cost: \$108210.7881, found at iteration #68

Best lower bound: \$111187.0364
Gap = 2976.24833, or 2.750417387%

Stage One Variables:

\[
\begin{array}{c|c}
 i & X[i] \\
 \hline
 - & ------ \\
 1 & 166.70 \\
 2 & 81.90 \\
 3 & 250.87 \\
 4 & 0.14 \\
\end{array}
\]

This solution is very nearly optimal. (Optimal solution is \$108390.)
Benders’ & Lagrangian subproblems at each iteration

Best upper & lower bounds at each iteration
RESEARCH ISSUES

1. Given that the number of scenarios is *extremely* large (or probability distributions are continuous and not discrete), how does one do "sampling" of scenarios in the cross-decomposition algorithm?

2. How can the cross-decomposition algorithm be extended to multi-(i.e., greater than 2) stages?

3. Given uncertainty in the parameters of the probability distributions describing future scenarios, perhaps it is not appropriate to continue iterations until the duality gap between upper & lower bounds is nearly zero-- can we determine an appropriate gap between upper & lower bound for a termination criterion for the cross-decomposition algorithm?

4. Case of integer first-stage decisions.
   ♦ The Lagrangian subproblems $D_k(\lambda)$ for scenarios $k=1,…K$ are now mixed-integer LP problems, which are substantially more difficult to solve.
   ♦ The computational savings obtained by solving only the Lagrangian subproblem $D_0(\lambda)$ and not the Lagrangian subproblems $D_k(\lambda)$ for scenarios $k=1,…K$ at every iteration become more significant!
   ♦ The Lagrangian subproblems $D_k(\lambda)$ for scenarios $k=1,…K$ may occasionally be solved, in order to test the duality gap as a termination criterion. How can information about the dual variables gathered from Benders' subproblems be accumulated in order to construct a Benders' master problem for each individual $D_k(\lambda)$?
REFERENCES

A comprehensive textbook:


Illustration of use of "variable splitting":


Development of the Cross-Decomposition algorithm:

