CSCI 1951-G – Optimization Methods in Finance
Part 11:
Stochastic Optimization

April 13, 2018
Outline

1) Uncertainty in optimization

2) The troubles of an European farmer

3) Two-stage problems with recourse

4) Benders decomposition

5) Multistage problems
This material is covered in Chapter 16 of the *Optimization Methods in Finance* textbook.

Additional reading:
Birge and Louveaux, *Introduction to Stochastic Programming*, 2nd Ed. 1.1–3, 2.4–6, 3.1, 3.4, 4, 5.1. Some of the figures and examples in these slides are from this book.
Important!

*No* class next Friday *4/20*

*Yes* class Friday *4/27*  
(last class)
Outline

1) Uncertainty in optimization

2) The troubles of an European farmer

3) Two-stage problems with recourse

4) Benders decomposition

5) Multistage problems
Motivation

\[
\begin{align*}
\text{min } & \quad c^T x \\
\text{s.t. } & \quad Ax = b \\
& \quad x \geq 0
\end{align*}
\]

Assumption: we know the *exact* value of the *constraint parameters* $A$ and $b$.

But life is full of *uncertainties*…

What if the parameters were *random variables*? (with *known distribution*)

*Stochastic programming*: optimization with of uncertainties
Modeling Uncertainty

\[ \Omega = \{\omega_1, \ldots, \omega_S\} \] \text{: sample space with } S \text{ disjoint events.}

\[ p_i = \Pr(\omega_i) \] \text{: probability of event } \omega_i, \sum_{i=1}^{S} p_i = 1.

\[ \omega \] \text{: random variable (or vector of r.v.'s) representing events in } \Omega;

\[ A = A(\omega), \ b = b(\omega). \]
If we could wait to make decisions until we know $\omega$, we could solve

$$\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad A(\omega)x = b(\omega) \\
& \quad x \geq 0
\end{align*}$$

Often, we must make some decisions before knowing $\omega$.

What’s our goal in making decisions when we face uncertainties?

How do we achieve this goal?
Outline

1) Uncertainty in optimization

2) The troubles of a European farmer

3) Two-stage problems with recourse

4) Benders decomposition

5) Multistage problems
Example: The uncertain life of a European farmer

Claire, a European farmer, has 500 acres of land.

She grows wheat, corn, and sugar beets. She also has some cattle.

Claire’s winter dilemma: how much land to devote to each crop?
Example: The uncertain life of a European farmer

Claire’s winter dilemma: *how much* of her 500 acres to devote to each crop?

Claire knows the *mean yield of each crop per acre*.

Planting crops incurs in *costs* (per acre);

Claire can *sell* up to $Z$ tons of beets at a favourable price, and the rest at a discounted price.

Claire *needs* enough corn and wheat to feed her cattle. She can *grow* them, *buy* them, and even *sell* them.

Claire wants to minimize the *losses* *(difference* betw. planting/buying expenses and selling earnings)*
Data and variables

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yield (T/acre)</td>
<td>2.5</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Planting cost ($/acre)</td>
<td>150</td>
<td>230</td>
<td>260</td>
</tr>
<tr>
<td>Selling price ($/T)</td>
<td>170</td>
<td>150</td>
<td></td>
</tr>
<tr>
<td>Purchase price ($/T)</td>
<td>238</td>
<td>210</td>
<td></td>
</tr>
<tr>
<td>Minimum requirement (T)</td>
<td>200</td>
<td>240</td>
<td></td>
</tr>
</tbody>
</table>

Total available land: 500 acres

- $a_w, a_c, a_b$: acres of land devoted wheat, corn, and beets
- $s_w, s_c, s_{b+}, s_{b-}$: sold T of wheat, corn, beets (at favorable price), beets (at unfavorable price);
- $p_w, p_c$: purchased T of wheat, corn;

Formulation:

$$\text{min } 150a_w + 230a_c + 260a_b + 238p_w - 170s_w + 210p_c$$
$$- 150s_c - 36s_{b+} - 10s_{b-}$$

s.t. $a_w + a_c + a_b \leq 500$
$$2.5a_w + p_w - s_w \geq 200, 3a_c + p_c - s_c \geq 240$$
$$s_{b+} + s_{b-} \leq 20a_b, s_{b+} \leq 6000$$
$$a_w, a_c, a_b, p_w, p_c, s_w, s_c, s_{b+}, s_{b-} \geq 0$$
Formulation

\[
\begin{align*}
\text{min} & \quad 150a_w + 230a_c + 260a_b + 238p_w - 170s_w + 210p_c \\
& \quad - 150s_c - 36s_{b+} - 10s_{b-} \\
\text{s.t.} & \quad a_w + a_c + a_b \leq 500 \\
& \quad 2.5a_w + p_w - s_w \geq 200, \ 3a_c + p_c - s_c \geq 240 \\
& \quad s_{b+} + s_{b-} \leq 20a_b, \ s_{b+} \leq 6000 \\
& \quad a_w, a_c, a_b, p_w, p_c, s_w, s_c, s_{b+}, s_{b-} \geq 0
\end{align*}
\]

Claire solves this LP with her favourite solver:

<table>
<thead>
<tr>
<th>Culture</th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface (acres)</td>
<td>120</td>
<td>80</td>
<td>300</td>
</tr>
<tr>
<td>Yield (T)</td>
<td>300</td>
<td>240</td>
<td>6000</td>
</tr>
<tr>
<td>Sales (T)</td>
<td>100</td>
<td>–</td>
<td>6000</td>
</tr>
<tr>
<td>Purchase (T)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Overall profit: $118,600
Crop yields have been very *different from year to year*, due to weather conditions, seed quality, …

Simplifying assumptions:
- years are *good, fair, or terrible* for all crops:
  - good year: yield is 20% above mean;
  - fair year: yield is exactly mean;
  - bad year: yield is 20% below mean;
- prices *stay the same* independently of what year it is;

Claire asks herself: *what would happen in a good/terrible year?*
1.1 A Farming Example and the News Vendor Problem

Table 3 Optimal solution based on above average yields (+ 20%). If the three scenarios have an equal probability of 1/3, the farmer's problem reads as follows:

Claire *adjust the yields* in the LP formulation and solves two LPs, one for good years, and one for terrible years;

<table>
<thead>
<tr>
<th>Culture</th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface (acres)</td>
<td>183.33</td>
<td>66.67</td>
<td>250</td>
</tr>
<tr>
<td>Yield (T)</td>
<td>550</td>
<td>240</td>
<td>6000</td>
</tr>
<tr>
<td>Sales (T)</td>
<td>350</td>
<td>–</td>
<td>6000</td>
</tr>
<tr>
<td>Purchase (T)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Overall profit: $167,667</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure: Good years**

<table>
<thead>
<tr>
<th>Culture</th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface (acres)</td>
<td>100</td>
<td>25</td>
<td>375</td>
</tr>
<tr>
<td>Yield (T)</td>
<td>200</td>
<td>60</td>
<td>6000</td>
</tr>
<tr>
<td>Sales (T)</td>
<td>–</td>
<td>–</td>
<td>6000</td>
</tr>
<tr>
<td>Purchase (T)</td>
<td>–</td>
<td>180</td>
<td>–</td>
</tr>
<tr>
<td>Overall profit: $59,950</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure: Terrible years**

The optimal solution is very *sensitive to changes in yields*, and so is the profit. (Profit for fair years: $118,600)
Claire cannot make a set of perfect decisions that would be best in all circumstances;

She would like to understand the benefits and losses of each decision in each situation;

• The decisions $a_w, a_c, a_b$ on land assignments must be done now. They are first stage or anticipative variables.

• Sales $s_i$ and purchases $p_i$ decisions depend on yields, must be made at later stage. They are adaptive or second-stage variables.

Let’s give the sales and purchases variables an additional subscript index, denoting the scenario $r$ 

$(r = g$: good year, $r = f$: fair year, $r = t$: terrible year):

• $s_is: i = w, c, b_+, b_-, r = g, f, t$

• $p_is: i = 1, 2, r = g, f, t$

E.g.: $s_{b+f}$: T of beets sold at the favorable price in a fair year.
Probabilistic weather

Claire wants to maximize her long-run profit, i.e., her expected profit. We need a probability distribution:

- good, fair, and terrible years are equiprobable (each has prob. 1/3)

Claire’s LP then becomes:

\[
\begin{align*}
\text{min } & 150x_1 + 230x_2 + 260x_3 \\
& - \frac{1}{3} (170w_{11} - 238y_{11} + 150w_{21} - 210y_{21} + 36w_{31} + 10w_{41}) \\
& - \frac{1}{3} (170w_{12} - 238y_{12} + 150w_{22} - 210y_{22} + 36w_{32} + 10w_{42}) \\
& - \frac{1}{3} (170w_{13} - 238y_{13} + 150w_{23} - 210y_{23} + 36w_{33} + 10w_{43})
\end{align*}
\]

s.t. \( x_1 + x_2 + x_3 \leq 500 \), \( 3x_1 + y_{11} - w_{11} \geq 200 \),
\[3.6x_2 + y_{21} - w_{21} \geq 240\], \( w_{31} + w_{41} \leq 24x_3 \), \( w_{31} \leq 6000 \),
\[2.5x_1 + y_{12} - w_{12} \geq 200\], \( 3x_2 + y_{22} - w_{22} \geq 240\),
\( w_{32} + w_{42} \leq 20x_3 \), \( w_{32} \leq 6000 \), \( 2x_1 + y_{13} - w_{13} \geq 200\),
\[2.4x_2 + y_{23} - w_{23} \geq 240\], \( w_{33} + w_{43} \leq 16x_3 \),
\( w_{33} \leq 6000 \), \( x, y, w \geq 0 \).
The stochastic LP

\[
\begin{align*}
\text{min } & 150x_1 + 230x_2 + 260x_3 \\
& - \frac{1}{3}(170w_{11} - 238y_{11} + 150w_{21} - 210y_{21} + 36w_{31} + 10w_{41}) \\
& - \frac{1}{3}(170w_{12} - 238y_{12} + 150w_{22} - 210y_{22} + 36w_{32} + 10w_{42}) \\
& - \frac{1}{3}(170w_{13} - 238y_{13} + 150w_{23} - 210y_{23} + 36w_{33} + 10w_{43}) \\
\text{s.t. } & x_1 + x_2 + x_3 \leq 500 , \ 3x_1 + y_{11} - w_{11} \geq 200 , \\
& 3.6x_2 + y_{21} - w_{21} \geq 240 , \ w_{31} + w_{41} \leq 24x_3 , \ w_{31} \leq 6000 , \\
& 2.5x_1 + y_{12} - w_{12} \geq 200 , \ 3x_2 + y_{22} - w_{22} \geq 240 , \\
& w_{32} + w_{42} \leq 20x_3 , \ w_{32} \leq 6000 , \ 2x_1 + y_{13} - w_{13} \geq 200 , \\
& 2.4x_2 + y_{23} - w_{23} \geq 240 , \ w_{33} + w_{43} \leq 16x_3 , \\
& w_{33} \leq 6000 , \ x, y, w \geq 0 .
\end{align*}
\]

This LP is \textit{much larger} than the original one:

For each scenario, there are:

Copies of the \textit{second stage} decision variables;

A copy of the \textit{set of constraints} involving the 2nd stage variables;
**Solution**

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area (acres)</td>
<td>170</td>
<td>80</td>
<td>250</td>
</tr>
<tr>
<td><strong>s = 1 Above</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yield (T)</td>
<td>510</td>
<td>288</td>
<td>6000</td>
</tr>
<tr>
<td>Sales (T)</td>
<td>310</td>
<td>48</td>
<td>6000</td>
</tr>
<tr>
<td>Purchase (T)</td>
<td>–</td>
<td>–</td>
<td>(favor. price)</td>
</tr>
<tr>
<td><strong>s = 2 Average</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yield (T)</td>
<td>425</td>
<td>240</td>
<td>5000</td>
</tr>
<tr>
<td>Sales (T)</td>
<td>225</td>
<td>–</td>
<td>5000</td>
</tr>
<tr>
<td>Purchase (T)</td>
<td>–</td>
<td>–</td>
<td>(favor. price)</td>
</tr>
<tr>
<td><strong>s = 3 Below</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yield (T)</td>
<td>340</td>
<td>192</td>
<td>4000</td>
</tr>
<tr>
<td>Sales (T)</td>
<td>140</td>
<td>–</td>
<td>4000</td>
</tr>
<tr>
<td>Purchase (T)</td>
<td>–</td>
<td>48</td>
<td>(favor. price)</td>
</tr>
</tbody>
</table>

Overall profit: $108,390

**Observations:**
- Claire never sells beets at the unfavorable price;
- Sometimes she produces fewer beets than the quota;
- She always sells wheat, never having to buy it;
- She may sell or buy corn, depending on the year;
Some of Claire’s decisions (e.g., underproducing beets, having to buy corn) would never take place if she had perfect information. They happen because decisions have to be balanced (hedged) against the various scenarios.

Assume that the weather is cyclical (bad, fair, good) and that Claire knows it.

Then she knows, for every year, the optimal decisions to make.
The value of perfect information

With *complete knowledge* of the weather cycle, Claire’s average long-term profit would be the *average of the three profits*:

\[
\frac{59,950 + 118,600 + 167,667}{3} = 115,456.
\]

*Expected Value of Perfect Information* (EVPI):

the *profit lost* due to the presence of uncertainty:

\[
115,456 - 108,390 = 7,016
\]
The value of the stochastic solution

What would Claire’s profits in the long run be if she planted following the expected yields? (i.e., using the solution in the first original LP)

For each scenario, compute:
  yields, buy/sell quantities, and thus profits.

The average profit is $107,240.

Value of the stochastic solution (VSS):
  the additional gain achieved by solving the stochastic model

  \[ 108,390 - 107,240 = 1,150 \]
EVPI and VSS

EVPI measures the value of *knowing the future with certainty*;

VSS measures the value of *knowing and using distributions* on future outcomes.

In practice, EVPI is difficult to measure, so the emphasis is on VSS.
Outline

1) Uncertainty in optimization

2) The troubles of an European farmer

3) Two-stage problems with recourse

4) Benders decomposition

5) Multistage problems
Two-stage programs with fixed recourse

Claire’s problem is a \textit{two-stage stochastic LP with fixed recourse}:

\[
\min z = c^T x + \mathbb{E}[\min q(\omega)^T y(\omega)] \\
\text{s.t. } Ax = b \\
B(\omega)x + Wy(\omega) = h(\omega) \\
x \geq 0, \quad y(\omega) \geq 0
\]

- \(\omega\): random event \(\omega \in \Omega\), which happens with probability \(p(\omega)\);
- \(x\): first-stage decisions, to be fixed \textit{before} the realization of \(\omega\) is known;
- \textit{After} the realization of \(\omega\), we know the second-stage problem data \(q(\omega), h(\omega), \text{ and } B(\omega)\) (i.e., each component of these vector is a r.v.)
- \(y(\omega)\) are the second stage decisions, which depend on \(\omega\) in the sense that they depend on the random constraints and costs;
Recourse

\[
\begin{align*}
\min z &= c^T x + \mathbb{E}[\min q(\omega)^T y(\omega)] \\
\text{s.t. } Ax &= b \\
B(\omega)x + Wy(\omega) &= h(\omega) \\
x &\geq 0, \quad y(\omega) \geq 0
\end{align*}
\]

“Recourse” allows to correct the first-stage decisions when additional information (i.e., \(\omega\)) is revealed.

In fixed-recourse models, \(W\) is fixed, but in general \(W(\omega)\) is revealed.

Example:

- We make an initial set of investments;
- As time passes we can then adjust our allocations to take into account changes in values.
- The losses /profits from the allocations depend on the random changes in values.
Second-stage problem

Second-stage or recourse problem:

\[ f(x, \omega) = \min q(\omega)^T y(\omega) \]
\[ Wy(\omega) = h(\omega) - B(\omega)x \]
\[ y(\omega) \geq 0 \]

Let \( f(x) = \mathbb{E}[f(x, \omega)] \) be the second stage value function.  
Then we can write:

\[ \min c^T x + f(x) \]
\[ Ax = b \]
\[ x \geq 0 \]
Scenarios

Assume $\Omega = \{\omega_1, \ldots, \omega_S\}$ (scenarios) and let $p = (p_1, \ldots, p_S)$ be the probability distribution on $\Omega$; Then

$$
\mathbb{E}[\min q(\omega)^T y(\omega)] = \sum_{k=1}^{S} p_k \min y(\omega_k) q^T y(\omega_k)
$$

Rewrite the stochastic program as:

$$
\min z = c^T x + \sum_{k=1}^{S} p_k \min q_k^T y_k
\text{subject to:}
Ax = b
B_k x + W_k y_k = h_k \text{ for } k = 1, \ldots, S
x \geq 0
y_k \geq 0 \text{ for } k = 1, \ldots, S
$$

There is a different 2nd stage decision vector $y_k$ for each scenario $k$
The optimization is over $x$ and all the $y_k$. 
The deterministic equivalent problem

\[
\min_{x,y_1,\ldots,y_k} c^T x + p_1 q_1^T y_1 + \cdots + p_S q_S^T y_S \\
Ax = b \\
B_1 x + W_1 y_i = h_1 \\
\vdots \\
B_S x + W_S y_S = h_S \\
x, y_1, \ldots, y_s \geq 0
\]

**Large** LP:

*S copies* of the 2nd stage decision variables and of the constraints.

The problem has a **very nice structure**:

We should be able to **exploit** it to **solve the problem efficiently**.
1) Uncertainty in optimization

2) The troubles of an European farmer

3) Two-stage problems with recourse

4) Benders decomposition

5) Multistage problems
Let’s consider a two-stage problem:

\[
\begin{align*}
\max_{x, y_1, \ldots, y_k} & \quad c^T x + p_1 q_1^T y_1 + \cdots + p_S q_S^T y_S \\
Ax & = b \\
B_1 x + W_1 y_i & = h_1 \\
& \vdots \\
B_S x + W_S y_S & = h_S \\
x, y_1, \ldots, y_S & \geq 0
\end{align*}
\]
Bender’s decomposition

Benders decomposition, (aka the L-Shaped method) solves a number of *smaller LPs*, leveraging the structure.

- Start by solving a “*master*” LP involving only $x$ and $Ax = b$;
- Solve a series of *small, independent, linear* “recourse problems”, each involving a *different vector of 2nd stage variables* $y_k$ (one per scenario).
- The recourse problems depend on the optimal solution $x^*$ of the master LP.
- The *recourse LPs optimal solutions* are used to *generate inequalities that are added to the master LP* (cuts!)
- Solve the *new master problem*, obtain a new $x^*$, and *iterate*. 
We can rewrite the two-stage problem as:

$$\max_x c^T x + P_1(x) + \cdots + P_S(x)$$

$$Ax = b$$

$$x \geq 0$$

where, for $k = 1, \ldots, S$, the recourse linear problem $P_k(x)$ is:

$$P_k(x) = \max_{y_k} p_k q_k^T y_k$$

$$W_k y_k = h_k - B_k x$$

$$y_k \geq 0$$

We solve the recourse LPs $P_k(x), k = 1, \ldots, S$ for a sequence of vectors $x^i, i = 0, \ldots$. 
The recourse linear problem

We solve the recourse LPs \( P_k(x), k = 1, \ldots, S \) for a sequence of vectors \( x^i, i = 0, \ldots \).

\( x^0 \) is obtained by solving the first master LP

\[
\max_x c^T x \\
Ax = b \\
x \geq 0
\]

Observations:
- \( x^0 \) may not be optimal for the original problem
- \( x^0 \) may make some of the recourse LPs infeasible.
Assume to have obtained $x^i$ from the master LP. The dual of a recourse LP $P_k(x)$, given $x^i$, is:

$$P_k(x^i) = \min_{u_k} u_k^T (h_k - B_k x^i)$$

$$W_k^T u_k \geq p_k q_k$$

Assume the primal is feasible with optimal solution $y^i_k$ and $u^i_k$ is the corresponding optimal dual. Then

$$P_k(x^i) = (u^i_k)^T (h_k - B_k x^i)$$

From LP duality we have that, for an optimal $x$,

$$P_k(x) \leq (u^i_k)^T (h_k - B_k x)$$

We can then add the following optimality cut to the current master linear program:

$$P_k(x) \leq (u^i_k)^T (B_k x^i - B_k x) + P_k(x^i)$$
If the primal recourse is *infeasible*, then the dual is *unbounded*

We want in first-stage decisions $x$ that leads to *feasible* second stage decisions $y_k$;

Let $u^i_k$ be a direction where the dual is unbounded, i.e.:

\[
(u^i_k)^T(h_k - B_k x^i) \leq 0 \quad \text{and} \quad W_k^T u^i_k \geq p_k q_k
\]

We can add the following *feasibility cut* to the current master program:

\[
(u^i_k)^T(h_k - B_k x) \geq 0
\]
After solving the recourse problems for each $k$ we have a lower bound to the optimal value of the stochastic program:

$$LB_i = c^T x^i + P_1(x^i) + \cdots + P_S(x^i)$$

where $P_k(x^i) = -\infty$ if the corresponding problem is infeasible
After adding all the optimality and feasibility cuts found so far (for $j = 0, \ldots, i$) to the master program, we obtain a new linear program:

$$\begin{align*}
\max \ x, z_1, \ldots, z_S \ c^T x + \sum_{k=1}^{S} z_k \\
Ax &= b \\
z_k &\leq (w_k^j)^T (B_k x^j - B_k x) + P_k(x^j) \text{ for some pairs (}j, k\text{)} \\
0 &\leq (w_k^j)^T (h_k - B_k x) \text{ for the remaining pairs} \\
x &\geq 0
\end{align*}$$

By solving this problem we obtain:

- new first stage decision variables $x^{i+1}$; and
- a new upper bound $UB_i$ to the optimal value of the original stochastic problem.

Bender decomposition stops when $LB_i$ and $UB_i$ are closer than a desired threshold.
Algorithm

- Start with the “simplest” master program to obtain $x_0$.
- Set $UB = +\infty$ and $LB = -\infty$.
- While $UB - LB >$threshold:
  - For $k = 1, \ldots, S$
    - Try to solve recourse problem $P_k$.
    - If feasible, add optimality cut
    - Otherwise, add feasibility cut
  - Compute new $LB$ using $x^i$
  - Solve the new master program and obtain the new $UB$
Outline

1) Uncertainty in optimization

2) The troubles of an European farmer

3) Two-stage problems with recourse

4) Benders decomposition

5) Multistage problems
Multi-stage programs with recourse

- The recourse decisions can be made at $n \geq 2$ points in time, called *stages*;
- The random event $\omega$ is a vector $(o_1, \ldots, o_{n-1})$ that gets revealed progressively over time;
- First stage decisions are taken before any component of $\omega$ is revealed;
- Then $o_1$ is revealed and the second stage decisions are taken
- Then $o_2$ and so on, alternating between revealing a new component and taking the current stage decisions.
Scenarios tree

- Assume $\Omega = \{\omega_1, \ldots, \omega_S\}$
- Some scenarios may be identical in their first components;
- They “become” differentiated in later stages;
- We can represent this situation as a scenario tree
Properties of the scenario tree

- Nodes are labeled from 1 to $N$, with 1 being the root.
- Each node $i$ in stage $k \geq 2$ has a single mother $m(i)$.
- The paths from the root to the leaves represent scenarios.
- The scenarios that pass through node $i$ in stage $k$ have identical components $o_1$ to $o_{k-1}$.
Formulation of a multi-stage stochastic problem

- For each node $i$ of the tree, there is a recourse decision vector $x_i$;
- For each node $i$, let $r_i$ be the sum of the $p_k$ for the scenarios $\omega_k$ that go through $i$;

Multi-stage stochastic program with recourse:

$$\min_{x_1, \ldots, x_N} \sum_{i=1}^{N} r_i c_i^T x_i$$

$Ax_1 = b$

$B_i x_{a(i)} + W_i x_i = h_i$ for $i = 2, \ldots, N$

$x_i \geq 0$
Example

- \( r_4 = p_1, r_5 = p_2, r_6 = p_3, r_7 = p_4 \)
- \( r_2 = p_1 + p_2 + p_3, r_3 = p_4, r_2 + r_3 = 1 = r_1. \)

\[
\min c^T x_1 + r_2 q_2^T x_2 + \cdots + r_7 q_7^T x_7 \\
A x_1 = b \\
B_{2}x_1 + W_2 x_2 = h_2, B_3 x_1 + W_3 x_3 = h_3 \\
B_{4}x_2 + W_4 x_4 = h_4, B_5 x_2 + W_5 x_5 = h_5, B_6 x_2 + W_6 x_6 = h_6 \\
B_7 x_3 + W_7 x_7 = h_7 \\
x_i \geq 0
\]
Observations

\[
\begin{align*}
\min & \quad c^T x_1 + r_2 q_2^T x_2 + \cdots + r_7 q_7^T x_7 \\
& \quad Ax_1 = b \\
& \quad B_2 x_1 + W_2 x_2 = h_2, B_3 x_1 + W_3 x_3 = h_3 \\
& \quad B_4 x_2 + W_4 x_4 = h_4, B_5 x_2 + W_5 x_5 = h_5, B_6 x_2 + W_6 x_6 = h_6 \\
& \quad B_7 x_3 + W_7 x_7 = h_7 \\
& \quad x_i \geq 0
\end{align*}
\]

- The size of the LP increases rapidly with the number of stages (e.g., 10 stages and binary tree means 1024 scenarios, 2047 decision vectors, 2048 constraints);
- The problem still has the “nice” structure of 2-stage stochastic problems;
Example from personal finance

- We have $55k to invest today in bonds or stocks;
- 15 years from now, we would like to have at least $G$ = $80k to pay for a 2-year master program;
- If we don’t have $80k after 15 years, we can borrow for a cost of 4% the amount we need;
- We can change the investments every 5 years, so we have 3 investment periods.
- We assume that over the 3 decision periods, 8 scenarios are possible and all equally likely ($p_i = 0.125$):
  - over each 5-year period, either stocks give a return of 1.25 and bonds of 1.14, or stocks give a return of 1.06 and bonds a return of 1.12
- We would like to maximize the expected amount of money we have left at the end of the 9 years (taking into account the eventual costs of borrowing and our tuition expenses)
Introduction and Examples

$s = 1, 2, 5, 6,$ and for $t = 3$, $s = 1, 3, 5, 7$. In the other cases, $\xi(1, t, s) = 1.06$, $\xi(2, t, s) = 1.12$.

Fig. 3 Tree of scenarios for three periods. The eight scenarios are represented by the tree in Figure 3. The scenario tree divides into branches corresponding to different realizations of the random returns. Because Scenarios 1 to 4, for example, have the same return for $t = 1$, they all follow the same first branch. Scenarios 1 and 2 then have the same second branch and finally divide completely in the last period. To show this more explicitly, we may refer to each scenario by the history of returns indexed by $s$ for periods $t = 1, 2, 3$ as indicated on the tree in Figure 3. In this way, Scenario 1 may also represent $(s_1, s_2, s_3) = (1, 1, 1)$.

With the tree representation, we need only have a decision vector for each node of the tree. The decisions at $t = 1$ are just $x(1, 1)$ and $x(2, 1)$ for the amounts invested in stocks (1) and bonds (2) at the outset. For $t = 2$, we would have $x(i, 2, s_1)$ where $i = 1, 2$ for the type of investment and $s_1 = 1, 2$ for the first-period return outcome. Similarly, the decisions at $t = 3$ are $x(i, 3, s_1, s_2)$.

With these decision variables defined, we can formulate a mathematical program to maximize expected utility. Because the concave utility function in Figure 1 is piecewise linear, we just need to define deficit or shortage and excess or surplus variables, $w(i_1, i_2, i_3)$ and $y(i_1, i_2, i_3)$, and we can main an linear model. The objective is simply a probability- and penalty-weighted sum of these terms, which, in general, becomes:

$$\text{Objective} = \sum \text{Probability} \cdot \sum \text{Penalty} \cdot (w(i_1, i_2, i_3) + y(i_1, i_2, i_3))$$
Formulation

\[
\begin{align*}
\max z &= \sum_{s_1=1}^{2} \sum_{s_2=1}^{2} \sum_{s_3=1}^{2} 0.125(y(s_1, s_2, s_3) - 4w(s_1, s_2, s_3)) \\
\text{s. t.} & \quad x(1, 1) + x(2, 1) = 55, \\
& \quad -1.25x(1, 1) - 1.14x(2, 1) + x(1, 2, 1) + x(2, 2, 1) = 0, \\
& \quad -1.06x(1, 1) - 1.12x(2, 1) + x(1, 2, 2) + x(2, 2, 2) = 0, \\
& \quad -1.25x(1, 2, 1) - 1.14x(2, 2, 1) + x(1, 3, 1, 1) + x(2, 3, 1, 1) = 0, \\
& \quad -1.06x(1, 2, 1) - 1.12x(2, 2, 1) + x(1, 3, 1, 2) + x(2, 3, 1, 2) = 0, \\
& \quad -1.25x(1, 2, 2) - 1.14x(2, 2, 2) + x(1, 3, 2, 1) + x(2, 3, 2, 1) = 0, \\
& \quad -1.06x(1, 2, 2) - 1.12x(2, 2, 2) + x(1, 3, 2, 2) + x(2, 3, 2, 2) = 0, \\
& \quad 1.25x(1, 3, 1, 1) + 1.14x(2, 3, 1, 1) - y(1, 1, 1) + w(1, 1, 1) = 80, \\
& \quad 1.06x(1, 3, 1, 1) + 1.12x(2, 3, 1, 1) - y(1, 1, 2) + w(1, 1, 2) = 80, \\
& \quad 1.25x(1, 3, 1, 2) + 1.14x(2, 3, 1, 2) - y(1, 2, 1) + w(1, 2, 1) = 80, \\
& \quad 1.06x(1, 3, 1, 2) + 1.12x(2, 3, 1, 2) - y(1, 2, 2) + w(1, 2, 2) = 80, \\
& \quad 1.25x(1, 3, 2, 1) + 1.14x(2, 3, 2, 1) - y(2, 1, 1) + w(2, 1, 1) = 80, \\
& \quad 1.06x(1, 3, 2, 1) + 1.12x(2, 3, 2, 1) - y(2, 1, 2) + w(2, 1, 2) = 80, \\
& \quad 1.25x(1, 3, 2, 2) + 1.14x(2, 3, 2, 2) - y(2, 2, 1) + w(2, 2, 1) = 80, \\
& \quad 1.06x(1, 3, 2, 2) + 1.12x(2, 3, 2, 2) - y(2, 2, 2) + w(2, 2, 2) = 80, \\
x(i, t, s_1, \ldots, s_{t-1}) \geq 0, \quad y(s_1, s_2, s_3) \geq 0, \quad w(s_1, s_2, s_3) \geq 0, \\
\text{for all } i, t, s_1, s_2, s_3.
\end{align*}
\]
Solving the problem in (2.1) yields a non-optimal expected utility value of \(-1.514\). We call this value, \(RP\), the expected recourse problem solution value. The optimal solution (in thousands of dollars) appears in Table 6.

Table 6: Optimal solution with three-period stochastic program.

<table>
<thead>
<tr>
<th>Period, Scenario</th>
<th>Stock</th>
<th>Bonds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,1-8</td>
<td>41.5</td>
<td>13.5</td>
</tr>
<tr>
<td>2,1-4</td>
<td>65.1</td>
<td>2.17</td>
</tr>
<tr>
<td>2,5-8</td>
<td>36.7</td>
<td>22.4</td>
</tr>
<tr>
<td>3,1-2</td>
<td>83.8</td>
<td>0.00</td>
</tr>
<tr>
<td>3,3-4</td>
<td>0.00</td>
<td>71.4</td>
</tr>
<tr>
<td>3,5-6</td>
<td>0.00</td>
<td>71.4</td>
</tr>
<tr>
<td>3,7-8</td>
<td>64.0</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Above G</th>
<th>Below G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24.8</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>8.87</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>1.43</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>1.43</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>8</td>
<td>0.00</td>
<td>12.2</td>
</tr>
</tbody>
</table>

- The initial solution is heavy in stocks;
- After the first period, we become either even more unbalanced towards stocks, or try to rebalance;
- The final investments are either all in stocks or all in bonds;
- Despite having to borrow money only once, because of the cost associated to it, the expected utility is negative: -$1,514.
Value of the stochastic solution

• What if we used a deterministic model replacing the random returns with their expectation?

• The expected return of stock is 1.155 in each period, while bonds return only 1.113

• The optimal investment plan would place all funds in stocks in each period

• The resulting utility is -$3,788.

•

$$VSS = -1,514 - (-3,788) = 2,274$$
Benders decomposition can be used for multi-stage problems:

- the stages are partitioned into:
  - a first set corresponding to the master problem; and
  - a second set corresponding to the recourse problems
- When the first set variables are fixed, then we have separate linear problems for each stage in the remaining set
- Solving these LPs gives additional feasibility cuts and optimality cuts

Benders composition is very easily parallelizable, thanks to the independence of the recourse problems
Outline

1) Uncertainty in optimization

2) The troubles of an European farmer

3) Two-stage problems with recourse

4) Benders decomposition

5) Multistage problems