Optimization Methods in Finance

Part 1
Stochastic Programming

Introduction

Example:

Suppose we have a company producing (splittable) goods G for which we need resources R

- At time 0 (now): Decide the number $x_i \geqslant 0$ of units of resource $i \in R$ that we want to order (for a price of c_i)
- At time 1: The resources arrive. We then decide how many units y_j of good $j \in G$ shall be produced if we get a price w_j for good j, there is a demand of $\leqslant b_j$ and we need a_{ij} units of resource i to produce j.

$$\begin{aligned} \max\ -c^T x + w^T y & & & \\ \sum_{j \in G} a_{ij} y_j & \leqslant & x_i & \quad \forall i \in R \\ y_j & \leqslant & b_j & \quad \forall j \in G \\ x_i, y_j & \geqslant & 0 & \quad \forall i \in R, j \in G \end{aligned}$$

Introduction (2)

Problem: Parameters concerning future are uncertain

- demand b_j and price w_j depend on the development of the market
- coefficients a_{ij} may change over time

Suppose we know:

- finite set $\Omega = \{\omega_1, \ldots, \omega_S\}$ of possible scenarios and the probability $p(\omega_k)$ of scenario ω_k
- parameters $b_j(\omega_k), p_j(\omega_k), a_{ij}(\omega_k)$ are random variables

Goal:

- Determine x_i (what to order at time 0), then observe which scenario ω_k became true, then choose $y_j(\omega_k)$
- Maximize the expected total profit

Introduction (3)

Two-stage stochastic program with recourse:

$$
\begin{aligned}
\max_{x} \; -c^T x + E[\max_{y(\omega)} w^T y(\omega)] & & & \\
\sum_{j \in G} a_{ij}(\omega) y_j(\omega) & \leqslant & x_i & \quad \forall i \in R \\
y_j(\omega) & \leqslant & b_j(\omega) & \quad \forall j \in G \\
x_i, y_j(\omega) & \geqslant & 0 & \quad \forall i \in R,\, j \in G
\end{aligned}
$$

Two kinds of decision variables:

- x_i: anticipative (here-and-now decisions)
- $y_j(\omega_k)$: adaptive (wait-and-see decisions)

Introduction (4)

Abbreviate $y^k := y(\omega_k)$, $b_j^k := b_j(\omega_k)$, $w_j^k := w_j(\omega_k)$, $p_k := p(\omega_k)$.
Then

$$E[w(\omega)^T y(\omega)] = \sum_{k=1}^{S} p_k {w^k}^T y^k$$

Deterministic equivalent:

$$\begin{aligned}
\max_{x,y} -c^T x + \sum_{k=1}^{S} p_k (w^k)^T y^k & & & \\
\sum_{j \in G} a_{ij}^k y_j^k &\leqslant& x_i & \quad \forall i \in R \, \forall k = 1, \ldots, S \\
y_j^k &\leqslant& b_j^k & \quad \forall j \in G \, \forall k = 1, \ldots, S \\
x_i, y_j^k &\geqslant& 0 & \quad \forall i \in R, \, j \in G, \, k = 1, \ldots, S
\end{aligned}$$

More general

What is Stochastic Programming?

- It is a way to deal with uncertainty in the parameters.
- Goal: Transformation to a so-called deterministic equivalent
- two kind of decision variables: anticipative and/or adaptive
- i.e. here-and-now versus wait-and-see
- multi-stage with recourse: anticipative and adaptive variables

Two-stage problems with recourse

$$\begin{array}{llll} \max_x & a^T x & + & E[\max_{y(w)} c(w)^T y(w)] & \\ & Ax & & & = b \\ & B(w)x & + & C(w)y(w) & = d(w) \\ & x \geqslant 0, & & y(w) \geqslant 0. & \end{array}$$

$$\begin{array}{llll} \max_x & a^T x & + & f(x) \\ & Ax = b & & \\ & x \geqslant 0. & & \end{array}$$

with $f(x) = E[f(x,w)]$ and

$$\begin{array}{llll} f(x,w) = & \max_{y(w)} & c(w)^T y(w) & \\ & & C(w)y(w) & = d(w) - B(w)x \\ & & y(w) \geqslant 0. & \end{array}$$

Two-stage problems with recourse and finite state space

- Let $\Omega = \{\omega_1, \ldots, \omega_S\}$ with probabilities $p_1, \ldots, p_S$

Stochastic program

$$\begin{array}{llcll} \max_x & a^T x & + & E[\max_{y(w)} c(w)^T y(w)] & \\ & Ax & & & = b \\ & B(w)x & + & C(w)y(w) & = d(w) \\ & x \geqslant 0, & & y(w) \geqslant 0 & \end{array}$$

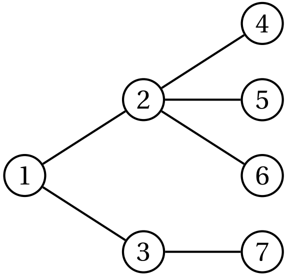
Deterministic equivalent

$$\begin{array}{llcll} \max_{x,y_k} & a^T x & + & \sum_{k=1}^{S} p_k \max_{y_k} c_k^T y_k & \\ & Ax & & & = b \\ & B_k x & + & C_k y_k & = d_k \quad \text{for } k = 1, \ldots, S \\ & x \geqslant 0 & & y_k \geqslant 0 & \quad\quad\ \ \text{for } k = 1, \ldots, S \end{array}$$

- $y_1, \ldots, y_k$ are independent.

Multi-stage

Scenario tree



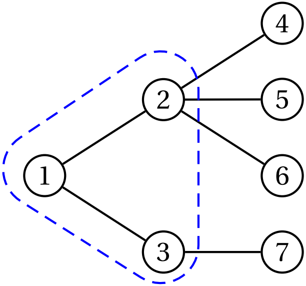
- $\{1\}$ root node
- $\{4, 5, 6, 7\}$ terminal nodes
- Four scenarios
- Three stages
- $a(i)$ is the father of i
- scenario tree could be huge

$$
\begin{array}{lcclll}
\max_{x_1,\dots,x_N} & c_1^T x_1 & + & \sum_{i=2}^N q_i c_i^T x_i & & \\
 & A x_1 & & & = b & \\
 & B_i x_{a(i)} & + & C_i x_i & = d_i & \text{for } i = 2,\dots,N \\
 & & & x_i \geqslant 0 & &
\end{array}
$$

- q_i is probability to reach node i

From Multi-stage to Two-stage

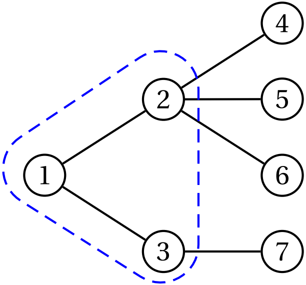
Scenario tree



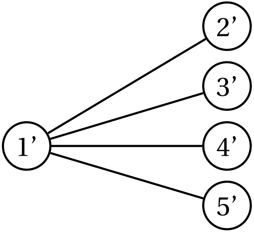
$$\begin{pmatrix} A & & & & & & \\ B_2 & C_2 & & & & & \\ B_3 & & C_3 & & & & \\ & B_4 & & C_4 & & & \\ & B_5 & & & C_5 & & \\ & B_6 & & & & C_6 & \\ & & B_7 & & & & C_7 \end{pmatrix}$$

$$\begin{array}{lccccl} \max_{x_1,\ldots,x_N} & c_1^T x_1 & + & \sum_{i=2}^N q_i c_i^T x_i & & \\ & Ax_1 & & & = b & \\ & B_i x_{a(i)} & + & C_i x_i & = d_i & \text{for } i = 2,\ldots,N \\ & & & x_i \geqslant 0. & & \end{array}$$

From Multi-stage to Two-stage



$$\begin{pmatrix} A & & & & & & \\ B_2 & C_2 & & & & & \\ B_3 & & C_3 & & & & \\ & B_4 & & C_4 & & & \\ & B_5 & & & C_5 & & \\ & B_6 & & & & C_6 & \\ & & B_7 & & & & C_7 \end{pmatrix}$$



$$\begin{pmatrix} A' & & & & \\ B'_2 & C'_2 & & & \\ B'_3 & & C'_3 & & \\ B'_4 & & & C'_4 & \\ B'_5 & & & & C'_5 \end{pmatrix}$$

Benders decomposition

We exploit the structure of the two-stage problem

$$\begin{pmatrix} A & & & \\ B_1 & C_1 & & \\ \vdots & & \ddots & \\ B_S & & & C_S \end{pmatrix}$$

Recall

$$\begin{array}{llll} \max_x & a^T x & + & \sum_{k=1}^{S} P_k(x) \\ & Ax = b & & \\ & x \geqslant 0. & & \end{array}$$

with

$$\begin{array}{llll} P_k(x) = & \max_{y_k} & p_k c_k^T y_k & \\ & & C_k y_k & = d_k - B_k x \\ & & y_k \geqslant 0. & \end{array}$$

Benders decomposition

Master LP

$$\max_{x,z_1,\ldots,z_S} a^T x + \sum_{k=1}^{S} z_k$$
$$Ax = b$$
$$x \geqslant \mathbf{0}$$

- z_k is auxiliary variable that gives an upper bound on $P_k(x)$

Algorithm:

(1) FOR $i = 0, \ldots$ DO

  (2) Solve master LP $\rightarrow (x^i, z^i)$

  (3) FOR $k = 1, \ldots, S$ DO

    (4) IF $P_k(x^i) = -\infty$ THEN add feasibility cut to master LP

    (5) ELSE add optimality cut to master LP

  (6) IF (x^i, z^i) still feasible to master LP THEN RETURN x^i

The recourse subproblem

Duality

$$
\begin{array}{rclrcl}
P_k(x) & = & \max_{y_k} & p_k c_k^T y_k & & \\
 & & & C_k y_k & = & d_k - B_k x \\
 & & & y_k & \geqslant & \mathbf{0} \\
 & & & & & \\
 & = & \min_{u_k} & u_k^T (d_k - B_k x) & & (D_k(x)) \\
 & & & C_k^T u_k & \geqslant & p_k c_k
\end{array}
$$

Feasibility Cuts

Case $P_k(x^i) = -\infty$:

- $D_k(x^i)$ unbounded
- Let u_k^i be a direction in which $D_k(x^i)$ is unbounded, i.e.

$$(u_k^i)^T(d_k - B_k x^i) < 0 \qquad \text{and} \qquad C_k^T u_k^i \geqslant 0$$

- Add feasibility cut to master LP

$$(u_k^i)^T(d_k - B_k x) \geqslant 0$$

- Cut forbids x^i
- Cut does not forbid any $(x, y_1, \ldots, y_S)$ which is feasible (to original problem)

$$(u_k^i)^T \underbrace{B_k x}_{=d_k - C_k y_k} = (u_k^i)^T(d_k - C_k y_k) = (u_k^i)^T d_k - \underbrace{(u_k^i)^T C_k}_{\geqslant \mathbf{0}} \underbrace{y_k}_{\geqslant \mathbf{0}} \leqslant (u_k^i)^T d_k$$

Optimality Cuts

Case $P_k(x^i)$ has optimum solution:

- Dual $D_k(x^i)$ has optimum solution u_k^i.
- $P_k(x^i) = (u_k^i)^T(d_k - B_k x^i)$
- u_k^i is feasible (not necessarily optimal) for $D_k(x)$ for any x

$$P_k(x) \leqslant (u_k^i)^T(d_k - B_k x)$$

- Adding up both equations/inequalities gives

$$\begin{aligned} P_k(x) &\leqslant (u_k^i)^T(d_k - B_k x) - (u_k^i)^T(d_k - B_k x^i) + P_k(x^i) \\ &= (u_k^i)^T(B_k x^i - B_k x) + P_k(x^i) \end{aligned}$$

- Add optimality cut $z_k \leqslant (u_k^i)^T(B_k x^i - B_k x) + P_k(x^i)$ to master LP

Optimality Cuts (2)

Lemma

Suppose $z_k^i > P_k(x^i)$. Then (x^i, z^i) is eliminated by the optimality cut.

Proof.

Plugging x^i and z_k^i into $z_k \leqslant (u_k^i)^T(B_k x^i - B_k x) + P_k(x^i)$ yields

$$z_k^i \leqslant \underbrace{(u_k^i)^T(B_k x^i - B_k x^i)}_{=0} + P_k(x^i) = P_k(x^i)$$

which is a contradiction. □

Conclusion

The master LP in iteration i:

$$
\begin{aligned}
\max_{x, z_1, \ldots, z_S} a^T x + \sum_{k=1}^{S} z_k & & \\
Ax &= b & \\
z_k &\leqslant (u_k^j)^T (B_k x^j - B_k x) + P_k(x^j) & \text{for some } (j,k) \\
(u_k^j)^T B_k x &\leqslant (u_k^j)^T d_k & \text{for some } (j,k) \\
x &\geqslant \mathbf{0} &
\end{aligned}
$$

Conclusion

One can prove that the algorithm finds an optimum solution in finite time.

Scenario Generation

- If the state space is too large or even infinite,
- we have to approximate by few samples, e.g.
- by random sampling,
- such that the statistical properties of the sample are as close as possible to the ones of the original distribution (in particular the moments)
- Caution: The approximation might introduce modeling errors,
- e.g. create arbitrage opportunities.