1 Introduction

We will focus on stochastic two stage linear programs with recourse. The general formulation can be written

$$\min_x c^T x + E[Q(x, \xi)]$$
subject to $Ax = b$
$x \geq 0$ \hfill (1)

where the first stage decisions are $x \in \mathbb{R}^n$, the constraint matrix $A$ is $m \times n$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $\xi$ is the uncertainty, and $Q(x, \xi)$ is the cost of the recourse decision when the first stage decision is $x$ and the uncertainty is $\xi$. Thus, $Q(x, \xi)$ is the second stage cost. We take the expectation of the second stage cost over all scenarios $\xi$.

The second stage cost

$$Q(x, \xi) = \min_y q^T y$$
subject to $Wy = h(\xi) - T(\xi)x$
$y \geq 0$ \hfill (2)

where $y \in \mathbb{R}^p$, $W$ is a fixed $l \times p$ matrix, the right hand side $h(\xi) \in \mathbb{R}^l$ depends on the uncertainty $\xi$, and the $l \times n$ technology matrix $T(\xi)$ also depends on $\xi$. Note that the second stage optimization is over $y$, with $x$ taken as a parameter.

Sometimes the expectation does not capture the risk sufficiently. There are several alternatives.

2 Robust optimization

Replace expectation by worst possible outcome:

$$\min_{x,v} c^T x + v$$
subject to $Ax = b$
$v \geq Q(x, \xi)$ $\forall \xi$
$x \geq 0$ \hfill (3)

This doesn’t require knowledge of the density function for $\xi$. This is conservative.
A slightly less conservative choice is to only consider $\xi$ within some set:
\begin{align}
\min_x \quad & c^T x + v \\
\text{subject to} \quad & Ax = b \\
& v \geq Q(x, \xi) \quad \forall \xi \in \Xi \\
& x \geq 0
\end{align}
(4)

Typical choices for $\Xi$ are ellipsoids or boxes.

For example: we want to solve an LP, but the data is uncertain. If we assume each coefficient lies within some range then this is like taking a box around the data. Another popular model is to assume the data in each row of the constraint matrix sits within some ellipsoid: the resulting model can be solved as a second order cone program [3, 2].

For example, consider the robust LP:
\begin{align}
\min_x \quad & c^T x \\
\text{subject to} \quad & a^T x \geq b \quad \text{with} \quad a = \bar{a} + w, \forall w \text{ satisfying } w^T M^{-1} w \leq 1 \\
& x \geq 0
\end{align}

For any given $\bar{x}$, we have a subproblem to determine if the constraint holds:
\[ \min_w \{ \bar{x}^T w : w^T M^{-1} w \leq 1 \} . \]
The solution is $\bar{w} = \frac{-1}{\sqrt{\bar{x}^T M \bar{x}}} M \bar{x}$, so the constraint requires
\[ \bar{a}^T x - \sqrt{\bar{x}^T M \bar{x}} \geq b. \]
This is a second order cone constraint.

3 Chance constrained formulations

Require that $Q(x, \xi)$ be below some threshold $\beta$ with probability $1 - \alpha$:
\begin{align}
\min_x \quad & c^T x \\
\text{subject to} \quad & Ax = b \\
& \Pr(Q(x, \xi) \geq \beta) \leq \alpha \\
& x \geq 0
\end{align}
(5)

A special case is minimizing the Value-at-Risk (VaR):
\begin{align}
\min_{x, \eta} \quad & \eta \\
\text{subject to} \quad & Ax = b \\
& \Pr(Q(x, \xi) \geq \eta) \leq \alpha \\
& x \geq 0
\end{align}
(6)

With $\alpha = 0.1$, this is choosing $x$ so that 90th percentile outcome is as good as possible.

When the uncertainty $\xi$ consists of a finite number of scenarios, these problems can be modeled as equivalent integer programs [4].
4 Conditional Value-at-Risk (CVaR)

Instead of taking the expectation of all possible outcomes, take the expectation of the worst $\alpha$ outcomes. Perhaps, take the expectation of the worst 10% of outcomes.

$$\begin{align*}
\min_x & \quad c^T x + E_{\xi} [Q(x, \xi) \mid Q(x, \xi) \text{ is in worst } \alpha \% \text{ of outcomes}] \\
\text{subject to} & \quad Ax = b \\
& \quad x \geq 0
\end{align*}$$

(7)

With a finite set of scenarios, this can be modeled as a linear program:

$$\begin{align*}
\min_{x,v,\eta} & \quad c^T x + \eta + \frac{1}{\alpha} \sum_s p_s v_s \\
\text{subject to} & \quad Ax = b \\
& \quad Q(x, \xi_s) - \eta \leq v_s \quad \forall s \\
& \quad v_s \geq 0 \quad \forall s \\
& \quad x \geq 0
\end{align*}$$

(8)

This LP is only slightly more complicated than the original stochastic program \[\text{[1]}\]. The values $v_s$ give the excess of $Q(x, \xi_s)$ over $\eta$. The value of $\eta$ can fall in a range: anywhere between the best $(1 - \alpha)$ outcomes and the worst $\alpha$ outcomes. See \[\text{[5]}\].

Can combine the expectation with CVaR:

$$\begin{align*}
\min_{x,v,\eta} & \quad c^T x + E_{\xi} [Q(x, \xi)] + \mu \left( \eta + \frac{1}{\alpha} \sum_s p_s v_s \right) \\
\text{subject to} & \quad Ax = b \\
& \quad Q(x, \xi_s) - \eta \leq v_s \quad \forall s \\
& \quad v_s \geq 0 \quad \forall s \\
& \quad x \geq 0
\end{align*}$$

(9)

Here, we’ve weighted the CVaR by $\mu$, so we can trade it off with the expected return.
5 Coherent risk measures

Can combine minimizing the expected value with minimizing some measure of risk.

There is a theory of coherent risk measures [1, 6] which possess various nice properties:

- convexity
- monotonicity
- translation equivalence
- positive homogeneity

**Variance not a coherent risk measure.** It fails monotonicity, which requires:

if the outcome is always worse then the risk measure should be higher

For example, assume we have two strategies each with two possible outcomes:

- **Strategy A:**
  - in scenario $\xi^1$, value of outcome is 10; in scenario $\xi^2$, value of outcome is 20.

- **Strategy B:**
  - in scenario $\xi^1$, value of outcome is 25; in scenario $\xi^2$, value of outcome is 25.

So Strategy A always outperforms Strategy B, so we should choose strategy A. The variance of the outcomes with Strategy A is positive, whereas it is zero with Strategy B. So if we were seeking to minimize variance, we would choose Strategy B.

**Semivariance:** minimize the variance of the outcomes above some threshold. This is a coherent risk measure.

**CVaR** is coherent.

**VaR** is not coherent: it fails convexity.

Stating the other properties very informally:

- translation equivalence: if the value of every outcome is increased by the same amount, then the ranking of different strategies should not change.

- positive homogeneity: if the value of every outcome is scaled by the same positive amount, then the ranking of different strategies should not change.

- convexity: if $x^1$ and $x^2$ are two possible first stage decisions with risk measures $\rho(x^1)$ and $\rho(x^2)$ then the risk measure of their weighted average $\lambda x^1 + (1-\lambda)x^2$ with $0 < \lambda < 1$ satisfies

$$\rho(\lambda x^1 + (1-\lambda)x^2) \leq \lambda \rho(x^1) + (1-\lambda)\rho(x^2)$$

References


