Stochastic Programming Models

Risk Measures

Sirong Luo

Faculty of Statistics and Data Science
Shanghai University of Finance and Economics

Table of Contents

Risk Measures

- In the classical Markowitz model, variance (equivalently standard deviation) is used as a measure of risk. This measure of risk is relatively easy to compute, and, as we have seen in Chapter 6, leads to a quadratic programming model when we are interested in finding efficient portfolios.
  As we illustrate below, variance has some shortcomings as a measure of risk. This has motivated the introduction of other risk measures.

  Dispersion Measures
  Let r denote the (random) return of an asset. The variance

$$\sigma^2 = \mathrm{var}(r) = \mathbb{E}\left((r-\mu)^2\right)$$

  is a measure of dispersion of the distribution of r. Another dispersion measure is the mean absolute deviation (MAD) favored by Konno and Yamazaki (1991):

$$\mathbb{E}(|r-\mu|)$$

Risk Measures

- For the special case of normally distributed returns, the mean absolute deviation and the standard deviation are equivalent. Indeed, the following property is a straightforward exercise in probability:

  Proposition 11.1 If $r \sim N\left(\mu, \sigma^2\right)$ then $\mathbb{E}(|r-\mu|) = \sqrt{2/\pi}\sigma$.

  A major difference between mean absolute deviation and standard deviation is their sensitivity to outliers. The mean absolute deviation is more robust to outliers. When the distribution of joint returns is represented via a set of scenarios, the computation of efficient portfolios for the mean absolute deviation can be formulated as a linear program. This offers an alternative with potential advantages as we will show next. Suppose the investment universe has n assets with (random) returns $r_1, r_2, \ldots, r_n$. Let $\mu_j = \mathbb{E}\left(r_j\right), j = 1, \ldots, n$.

Risk Measures

- Recall the portfolio optimization problem that finds the minimum-variance portfolio among a set of portfolios $\mathcal{X}$:

$$\begin{aligned} &\min_{\mathbf{x}} \operatorname{var}\left(\mathbf{r}^\top \mathbf{x}\right) = \mathbb{E}\left(\left[(\mathbf{r}-\boldsymbol{\mu})^\top \mathbf{x}\right]^2\right) \\ &s.t. \mathbf{x} \in \mathcal{X} \end{aligned}$$

  Consider now the model obtained by using instead the mean absolute deviation as a measure of risk:

$$\begin{aligned} &\min_{\mathbf{x}} \mathbb{E}\left(\left|(\mathbf{r}-\boldsymbol{\mu})^\top \mathbf{x}\right|\right) \\ &\text{s.t. } \mathbf{x} \in \mathcal{X} \end{aligned} \tag{11.1}$$

  Not only does the computation of efficient portfolios based on formulation (11.1) involve solving a linear program as opposed to a quadratic program, but also the linear program solves the problem directly over the set of scenarios thereby circumventing the estimation of the covariance matrix.

- **Mean Absolute Deviation via Scenario Optimization**

Assume the possible scenarios for the vector of returns $\mathbf{r} = \begin{bmatrix} r_1 & \cdots & r_n \end{bmatrix}^\top$ are

$$\mathbf{r}^k = \begin{bmatrix} r_1^k & \cdots & r_n^k \end{bmatrix}^\top, \quad k = 1, \ldots, S,$$

and scenario k occurs with probability p_k, $k = 1, \ldots, S$. Then we can write the above mean absolute deviation model (11.1) as

$$\begin{aligned} \min_{\mathbf{x},\mathbf{w}} & \sum_{k=1}^{S} p_k w_k \\ \text{s.t. } & w_k = \left| \left(\mathbf{r}^k - \boldsymbol{\mu}\right)^\top \mathbf{x} \right| \qquad \text{for } k = 1, \ldots, S \\ & \mathbf{x} \in \mathcal{X} \end{aligned}$$

Risk Measures

- We now turn this formulation into a linear program as follows:

$$\begin{aligned}
\min_{\mathbf{x},\mathbf{w}} \ & \sum_{k=1}^{S} p_k w_k \\
\text{s.t.} \ & w_k \geq \left(\mathbf{r}^k - \boldsymbol{\mu}\right)^\top \mathbf{x} && \text{for } k = 1, \dots, S \\
& w_k \geq -\left(\mathbf{r}^k - \boldsymbol{\mu}\right)^\top \mathbf{x} && \text{for } k = 1, \dots, S \\
& \mathbf{x} \in \mathcal{X}.
\end{aligned}$$

Note that, because $p_k > 0$ for $k = 1, \dots, S$ and the objective is minimized, w_k in an optimal solution satisfies at equality the constraint with the larger right-hand side, that is, $w_k = \left|\left(\mathbf{r}^k - \boldsymbol{\mu}\right)^{\mathrm{T}} \mathbf{x}\right|$.

- **Downside Risk Measures**

  Dispersion measures, such as variance and mean absolute deviation, measure the degree of uncertainty in the random return. These measures treat both positive and negative deviations from the mean as equally risky. In particular, these types of measures are blind to skewed distributions.
  We will next discuss two popular downside risk measures:*value at risk* and *conditional value at risk*. Value at risk (VaR) was first introduced by a team at J.P. Morgan and made available through RiskMetrics. VaR is used by many financial institutions to track and report the market risk exposure of their trading portfolios.

Risk Measures

- VaR is a measure of the worst possible loss that a portfolio may sustain with a pre-specified likelihood. For that reason, VaR is generally measured in dollar terms, instead of percentage units. The formal definition is as follows. Assume that Y is a (random) loss function, and $\alpha \in (0,1)$ is a confidence level (typically $99\%, 95\%$, or 90%). The α value at risk of Y is the $(1-\alpha)$ quantile of Y; that is, the value γ such that

$$\mathbb{P}(Y \geq \gamma) = 1 - \alpha$$

We shall denote this value by $\mathrm{VaR}_\alpha(Y)$.

- The value at risk has the following interpretation. Given a loss function Y and a confidence level $\alpha \in (0,1)$, the loss Y will exceed γ with probability $(1-\alpha)$. In the special case when the loss function is normally distributed, it is easy to compute VaR via well-known quantiles of the normal distribution.

Example 11.2 If $Y \sim N\left(\mu, \sigma^2\right)$ then

$$\mathrm{VaR}_{0.95}(Y) = \mu + 1.645\sigma, \mathrm{VaR}_{0.99}(Y) = \mu + 2.33\sigma$$

When Y has a discrete distribution, VaR can be computed by sorting the values of Y as detailed in the following example.

- Example 11.3 Assume there are S possible scenarios for the loss Y :

$$\mathbb{P}(Y = y_k) = p_k, k = 1, \ldots, S,$$

where

$$y_1 \leq y_2 \leq \cdots \leq y_S$$

Then

$$\mathrm{VaR}_\alpha(Y) = y_K,$$

where K is the smallest index such that

$$\sum_{i=K}^{S} p_i \geq 1 - \alpha$$

Risk Measures

In spite of its wide popularity, VaR is known to have the following two major shortcomings (see the exercises at the end of this chapter):

- VaR is not "subadditive": The VaR of two positions combined may be greater than the sum of the VaR of each, meaning that diversification can actually increase VaR.
- VaR does not distinguish loss size beyond the VaR threshold.

Risk Measures

These deficiencies of VaR led Artzner et al. (1999) to propose the following formal set of properties that a reasonable risk measure $\rho(Y)$ of a loss function Y should satisfy:

- Monotonicity: If $Y \geq 0$ then $\rho(Y) \geq 0$.
- Subadditivity: $\rho(Y + Z) \leq \rho(Y) + \rho(Z)$.
- Positive homogeneity: For $c > 0, \rho(cY) = c\rho(Y)$.
- Translational invariance: For any $c \in \mathbb{R}, \rho(Y + c) = \rho(Y) + c$.

A risk measure is *coherent* if it satisfies the above four properties. Neither standard deviation nor VaR are coherent. However, there is a modification of VaR that is coherent, namely the *conditional value at risk* introduced by Rockafellar and Uryasev (2000). Conditional value at risk (CVaR) is also known as expected tail loss.

Risk Measures

- CVaR can be motivated as follows. Since $\text{VaR}_\alpha(Y)$ is the most we can lose with probability α, it is equivalent to saying that with probability $(1-\alpha)$ the loss Y will be at least $\text{VaR}_\alpha(Y)$. CVaR is the answer to the following question: What should we expect the value of that loss to be? More precisely, CVaR is defined as follows. Given a loss function Y and confidence level $\alpha \in (0,1)$, the conditional value at risk is the expected loss Y, conditional on this loss being at least $\text{VaR}_\alpha(Y)$:

$$\mathbb{E}\left(Y \mid Y \geq \text{VaR}_\alpha(Y)\right)$$

We shall denote this expected value as $\mathrm{CVaR}_\alpha(Y)$. Again, in the special case when the loss function is normally distributed, it is easy to compute CVaR by using properties of the quantiles and expected tails of the normal distribution.

- **Example 11.4** If $Y \sim N\left(\mu, \sigma^2\right)$ then

$$\mathrm{CVaR}_{0.95}(Y) = \mu + 2.06\sigma, \mathrm{CVaR}_{0.99}(Y) = \mu + 2.67\sigma.$$

When Y has a discrete distribution, CVaR can be computed by sorting the values of Y.

- Example 11.5 Assume Y takes values $y_k, k = 1, \ldots, S$, in S possible scenarios:

$$\mathbb{P}(Y = y_k) = p_k, k = 1, \ldots, S,$$

where

$$y_1 \leq y_2 \leq \cdots \leq y_S$$

Then

$$\mathrm{CVaR}_\alpha(Y) = \frac{1}{1-\alpha} \sum_{i=K}^{S} p_i y_i$$

where K is the smallest index such that

$$\sum_{i=K}^{S} p_i = 1 - \alpha.$$

Note that here we may have to split the probability p_K.

A Key Property of CVaR

- We next present a key property of CVaR that makes it possible to solve portfolio optimization problems with CVaR via convex optimization. Proposition 11.6 Assume Y is a loss function. Then for $\alpha \in (0, 1)$

$$\mathrm{CVaR}_\alpha(Y) = \min_\gamma \left(\gamma + \frac{1}{1-\alpha} \mathbb{E}[\max(Y - \gamma, 0)] \right).$$

  Furthermore, the optimal solution (i.e., the minimizer) $\bar{\gamma}$ of this problem is $\mathrm{VaR}_\alpha(Y)$.

A Key Property of CVaR

As consequence of Proposition 11.6, it follows that CVaR is subadditive. Indeed, CVaR is a coherent risk measure (see exercises at the end of the chapter for details). Another consequence of Proposition 11.6 is that CVaR can be computed as the following linear two-stage stochastic program:

$$\text{CVaR}_\alpha(Y) = \min_\gamma [\gamma + \mathbb{E}(Q(\gamma, Y))],$$

where

$$\begin{aligned} Q(\gamma, Y) := \min_z \; & \frac{1}{1-\alpha} \cdot z \\ \text{s.t. } \; & z \geq Y - \gamma \\ & z \geq 0. \end{aligned}$$

More concisely

$$\begin{aligned} \mathrm{CVaR}_\alpha(Y) = \min_{\gamma, z} \; & \gamma + \frac{1}{1-\alpha} \cdot \mathbb{E}(z) \\ \text{s.t. } \; & z \geq Y - \gamma \\ & z \geq 0 \end{aligned}$$

A Key Property of CVaR

- In this formulation the first-stage and second-stage decision variables are γ and z respectively. Notice that z is adapted to the random outcome Y. In the particular case when Y is discrete and takes values y_k, for $k = 1, \ldots, S$, in S possible scenarios (not necessarily sorted):

$$\mathbb{P}(Y = y_k) = p_k, k = 1, \ldots, S,$$

A Key Property of CVaR

- we obtain the following linear programming formulation for $\mathrm{CVaR}_\alpha(Y)$.

Variables:

$$\gamma, z_1, \ldots, z_S$$

Linear programming formulation of CVaR:

$$\begin{array}{lll} \min_{\gamma,\mathbf{z}} & \gamma + \frac{1}{1-\alpha}\sum_{k=1}^{S} p_k z_k & \\ \text{s.t.} & z_k \geq y_k - \gamma, & \text{for } k = 1, \ldots, S \\ & z_k \geq 0, & \text{for } k = 1, \ldots, S \end{array}$$

An advantage of this formulation is that it allows us to minimize the CVaR of a portfolio via linear programming as we next explain.

Portfolio Optimization with CVaR

- The discussion in this section is based on Andersson et al. (2001). This study uses CVaR for measuring and controlling the credit risk of a portfolio of bonds. The loss function of interest is the loss due to credit risk; that is, the loss that the portfolio may suffer due to default or credit migration in its positions. This type of loss function is characterized by having a large likelihood of no loss and a small likelihood of a substantial loss. The loss distribution is heavily skewed. In this case, standard mean-variance analysis to characterize market risk is inadequate. VaR and CVaR are more appropriate criteria for minimizing portfolio credit risk.

Portfolio Optimization with CVaR

- **Distribution of Future Values for One Single Bond**
  Consider a risky bond and a fixed time horizon, e.g., one year. The future value of the bond depends on the forward curve that applies to its coupon payments. The forward curve in turn depends on the current rating of the bond. The benchmark future value of the bond is the future value of the bond if there is no change on its credit rating. However, in the event of credit migration, the future value of the bond may differ from the benchmark value. In particular, if the credit rating deteriorates, the coupon payments will be subject to higher discount values and the future value of the bond will be lower than its benchmark value.

- For a concrete illustration, suppose the one-year forward interest curves for the S& P credit ratings are as follows:

Category	Year 1	Year 2	Year 3	Year 4
AAA	0.036	0.0417	0.0473	0.0512
AA	0.0365	0.0422	0.0478	0.0517
A	0.0372	0.0432	0.0493	0.0532
BBB	0.041	0.0467	0.0525	0.0563
BB	0.0555	0.0602	0.0678	0.0727
B	0.0605	0.0702	0.0803	0.0852
CCC	0.1505	0.1502	0.1403	0.1352

Portfolio Optimization with CVaR

- Suppose the probabilities of credit rating migration for A, BBB, and B in one year are as follows:

	Rating at year end							
Initial rating	AAA	AA	A	BBB	BB	B	CCC	Default
A	0.09%	2.27%	91.05%	5.52%	0.74%	0.26%	0.01%	0.06%
BBB	0.02%	0.33%	5.95%	86.93%	5.30%	1.17%	0.12%	0.18%
B	0.00%	0.11%	0.24%	0.43%	6.48%	83.47%	4.07%	5.20%

Portfolio Optimization with CVaR

- Assuming a 50% recovery rate in default, the possible future values of a fiveyear, $6\%\mathrm{BBB}$ bond with face value 100 are as follows:

Year-end rating	Future value	Probability
AAA	109.352908	0.0002
AA	109.1723709	0.0033
A	108.6429921	0.0595
BBB	107.5309439	0.8693
BB	102.0063855	0.053
B	98.08591318	0.0117
CCC	83.6257912	0.0012
Default	50	0.0018

For example, for BBB rated bonds, the future value 107.5309439 was obtained as follows:

$$107.5309439 = 6 \cdot \left(1 + \frac{1}{1.041} + \frac{1}{1.0467^2} + \frac{1}{1.0525^3} + \frac{1}{1.0563^4}\right) + 100 \cdot \frac{1}{1.0563^4}$$

Portfolio Optimization with CVaR

- **Credit Risk Optimization for a Portfolio of Bonds**

  Now suppose we construct a portfolio of risky bonds. Assume there are n risky bonds and let x_j be the percentage of portfolio invested in bond j. Then the loss function of our portfolio is

$$Y(\mathbf{x}) := (\mathbf{b} - \omega)^\top \mathbf{x} = \sum_{j=1}^{n} (b_j - \omega_j)\, x_j,$$

  where each b_j is the future bond value of bond j with no credit migration, and ω_j is the (random) possible future bond value of bond j with credit migration. Suppose we want to select the portfolio in the constraint set $\mathcal{X}$ with minimum CVaR_α. In other words, we want to solve

$$\min_{\mathbf{x}} \mathrm{CVaR}_\alpha(Y(\mathbf{x}))$$
$$\mathbf{x} \in \mathcal{X}.$$

Portfolio Optimization with CVaR

- Suppose that the possible scenarios for the vector of future bond values $\boldsymbol{\omega} = \begin{bmatrix} \omega_1 & \cdots & \omega_n \end{bmatrix}^\top$ are

$$\boldsymbol{\omega}^k = \begin{bmatrix} \omega_1^k & \cdots & \omega_n^k \end{bmatrix}^\top, \quad k = 1, \ldots, S.$$

Then by Proposition 11.6, this problem has the following formulation:

$$\begin{aligned} \min_{\gamma, \mathbf{x}, \mathbf{z}} & \gamma + \frac{1}{1-\alpha} \sum_{k=1}^{S} p_k z_k \\ \text{s.t.} \ & z_k \geq \left(\mathbf{b} - \omega^k\right)^\top \mathbf{x} - \gamma, k = 1, \ldots, S \\ & z_k \geq 0, k = 1, \ldots, S \\ & \mathbf{x} \in \mathcal{X} \\ & \gamma \text{ free.} \end{aligned} \tag{11.2}$$

If the constraint set $\mathcal{X}$ is defined by linear constraints, then (11.2) is a linear program.

- **Scenario Generation in the Credit-Risk Example**

  When there is a single bond, the probability distribution of the possible future values of the bond depends on the probability of credit migration and the bond value in each of these scenarios. For instance, for the S&P ratings, the scenarios correspond to the ratings $\mathrm{AAA}, \mathrm{AA}, \mathrm{A}, \mathrm{BBB}, \mathrm{BB}, \mathrm{B}, \mathrm{CCC}$, and default. The likelihood of each of these scenarios is given by the migration matrix, which estimates the probability of migrating from one rating to the others over a specified time period.
  The discrete distribution readily yields the set of possible scenarios for the bond. Scenarios can also be generated via *normal sampling* as Figure 11.1 suggests (assuming we are working with a BB bond).

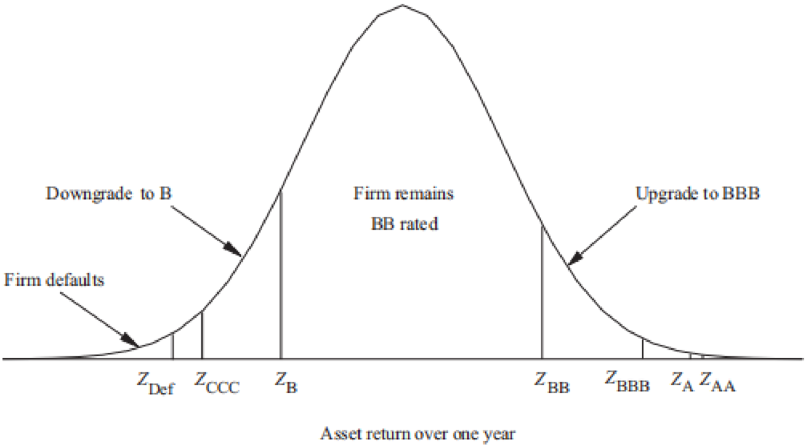
Downgrade to B
Firm remains
BB rated
Upgrade to BBB
Firm defaults
Z_{Def}
Z_{CCC}
Z_B
Z_{BB}
Z_{BBB}
Z_A
Z_{AA}
Asset return over one year

Figure11.1

Portfolio Optimization with CVaR

More precisely, normal sampling goes as follows:

- compute Z-scores associated with the probabilities of each of the scenarios,
- draw samples from a standard normal distribution,
- use the Z-scores to determine the sampled scenario.

Some interesting challenges arise in the scenario generation when we need to work with multiple bonds. Under the simple assumption that the credit migrations are statistically independent, we can generate scenarios via discrete sampling or independent normal sampling. Notice that, although discrete, the joint probability distribution for a set of ten or more bonds is extremely large. Hence it is generally impractical to exhaustively generate the entire set of scenarios.

Portfolio Optimization with CVaR

In the case when credit migrations are correlated, the scenario generation problem becomes more interesting. In this case a possible solution is to use correlated normal sampling. That is, draw samples from a correlated joint multivariate random variable. Then map each of the components in the random sample to a possible credit rating of the bonds. Some statistical packages, like the statistics toolbox in MATLAB, readily provide routines to sample from correlated multivariate normal variables. However, it is easy to generate correlated normal sampling from independent normal sampling. More precisely, to sample from a general n-dimensional normal distribution $N(\boldsymbol{\mu}, \mathbf{V})$, proceed as follows:

- Let $\mathbf{L}\mathbf{L}^\top = \mathbf{V}$ be the Cholesky factorization of the covariance matrix $\mathbf{V}$.
- Sample n standard independent normals $x_i \sim N(0, 1)$. Put $\mathbf{y} = \boldsymbol{\mu} + \mathbf{L}\mathbf{x}$.
- The resulting variable $\mathbf{y}$ has the desired distribution $\mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{V})$.

Solution of a Real-World Bond Example

Andersson et al. (2001) considered a portfolio of 197 bonds from 29 different countries with a market value of \$8.8 billion and duration of approximately five years. Their goal was to rebalance the portfolio in order to minimize credit risk. The one-year portfolio credit loss was generated using a Monte Carlo simulation: 20,000 scenarios of joint credit states of obligators and related losses. The distribution of portfolio losses had a long fat tail, as expected. The authors rebalanced the portfolio by minimizing CVaR using formulation (11.2). For $\alpha =$ 99%, the original bond portfolio had an expected portfolio return of 7.26%. The expected loss was 95 million dollars with a standard deviation of 232 million. The VaR was 1.03 billion dollars and the CVaR was 1.32 billion. After optimizing the portfolio (with expected return of 7.26%), the expected loss was only 5000 dollars, with a standard deviation of 152 million. The VaR was reduced to 210 million and the CVaR to 263 million dollars. So all around, the characteristics of the portfolio were much improved. Positions were reduced in bonds from Brazil, Russia, and Venezuela, whereas positions were increased in bonds from Thailand, Malaysia, and Chile. Positions in bonds from Colombia, Poland, and Mexico remained high and each accounted for about 5% of the optimized portfolio.

Notes

- As early as the 1970s and 1980s, some major financial institutions developed internal systems for risk management. The best known of these systems was RiskMetrics developed in the late 1980s at J.P. Morgan when chairman Dennis Weatherstone requested his staff provide a " $4:15\mathrm{pm}$ " daily one-page report measuring and explaining the risks and potential losses over the next 24 hours across the bank's entire portfolio. The RiskMetrics system featured and popularized the use of value at risk as a risk measure. The interest in a rigorous treatment of risk measures led a set of prominent scholars to develop a formal theory of *coherent measures of risk* in a landmark paper (Artzner et al., 1999). Conditional value at risk is one of the most popular coherent measures of risk.