VaR and CVaR

Přemysl Bejda

premyslbejda@gmail.com

2014
Contents

1 Motivation
2 General concept of CVaR
3 Optimization
4 Comparison
Contents

1. Motivation
2. General concept of CVaR
3. Optimization
4. Comparison
## Contents

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Motivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>General concept of CVaR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Optimization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Comparison</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Contents

1. Motivation
2. General concept of CVaR
3. Optimization
4. Comparison
Picture of our risk measures

The curve represents a hypothetical Profit-and-Loss probability density function. It has mean one and standard deviation one, but fatter tails than a Normal distribution. The 5% VaR point is 1.82 standard deviations below the mean, versus 1.64 for a Normal distribution.

Blue area to the right of the line represents 95% of the total area under the curve.

Red area to the left of the line represents 5% of the total area under the curve.

Line at -0.82 means 5% Value-at-Risk is 0.82.

Figure: Artificial example of VaR.
General overview

- VaR is known measure among investors.
- VaR is also discussed in BASEL II.
- CVaR takes into account the distribution of the tail.
- CVaR has many advantages, but its interpretation is worse.
- Previous picture is just for illustration, in the next chapter we will focus on loss instead of profit.
Outline

1. Motivation
2. General concept of CVaR
3. Optimization
4. Comparison
Model and assumptions

- Define possible loss function as the function $z = f(x, y)$.
- $x \in X$ represents a decision vector, $X \subset \mathbb{R}^n$ express decision constraints.
- $y \in Y \subset \mathbb{R}^m$ represents the future values (random variable) of some variables like interest rates, weather data...
- $y$ and $x$ are independent (I cannot influence the interest rate by some policy).
- The distribution function for the loss $z$ is defined as

$$\psi(x, \zeta) = P(f(x, y) \leq \zeta),$$

where we have to make some technical assumptions (see [Rockafellar and Urysanov (2001)]).
Definition 1 (VaR).

The $\alpha$-VaR of the loss associated with a decision $x$ is the value

$$
\zeta_\alpha(x) = \min_\zeta (\Psi(x, \zeta) \geq \alpha).
$$

- Since $\Psi(x, \zeta)$ is nondecreasing and right continuous the minimum has to be attained.
- In case of $\Psi(x, \zeta)$ continuous and increasing we find a solution of $\Psi(x, \zeta) = \alpha$.
- Define $\alpha$-VaR$^+$ as $\zeta_\alpha^+(x) = \inf_\zeta \Psi(x, \zeta) > \alpha$.
- When $\zeta_\alpha$ and $\zeta_\alpha^+$ differ it means that $\Psi(x, \zeta)$ is constant on $(\zeta_\alpha, \zeta_\alpha^+)$ and small change in $\alpha$ influence the result a lot.
- In this sense the VaR is not robust enough.
Definition 2 (Atom).

When the difference

\[ \psi(x, \zeta) - \psi(x, \zeta^-) = P(f(x, y) = \zeta) > 0 \]

(the minus sign denotes convergence from the left to the \( \zeta \) is positive, so that \( \psi(x, \cdot) \) has a jump at \( \zeta \), we say that probability atom is presented at \( \zeta \).
Motivation

General concept of CVaR

Optimization

Comparison

CVaR

Definition 3 (CVaR).

The $\alpha$-CVaR of the loss associated with a decision $x$ is the value

$$\phi_\alpha(x) = \int_{-\infty}^{\infty} f(x, y) d\psi_\alpha(x, f(x, y)),$$

where the distribution in question is the one defined by

$$\psi_\alpha(x, \zeta) = \begin{cases} 0 & \text{for } \zeta < \zeta_\alpha(x), \\ \frac{\psi(x, \zeta) - \alpha}{1 - \alpha} & \text{for } \zeta \geq \zeta_\alpha(x). \end{cases}$$

What to do when in the point $\zeta_\alpha(x)$ appears an atom?

The definition is good enough to deal with this problem.

What should really be meant by CVaR in this case?
CVA$\text{R}^+$ and CVA$\text{R}^-$

Definition 4 (CVA$\text{R}^+$ and CVA$\text{R}^-$).

The $\alpha$-CVA$\text{R}^+$ of the loss associated with a decision $\mathbf{x}$ is the value

$$
\phi^+_{\alpha}(\mathbf{x}) = \mathbb{E}(f(\mathbf{x}, \mathbf{y}) : f(\mathbf{x}, \mathbf{y}) > \zeta_{\alpha}(\mathbf{x})),$$

whereas $\alpha$-CVA$\text{R}^-$ of the loss associated with a decision $\mathbf{x}$ is the value

$$
\phi^-_{\alpha}(\mathbf{x}) = \mathbb{E}(f(\mathbf{x}, \mathbf{y}) : f(\mathbf{x}, \mathbf{y}) \geq \zeta_{\alpha}(\mathbf{x})).$$

- $\phi_{\alpha}^-(\mathbf{x})$ is well defined because $\mathbb{P}(f(\mathbf{x}, \mathbf{y}) \geq \zeta_{\alpha}) \geq 1 - \alpha > 0$.
- On the other hand it can happen that $\mathbb{P}(f(\mathbf{x}, \mathbf{y}) > \zeta_{\alpha}) = 0$, when in $\zeta_{\alpha}$ is the appropriate atom.
Basic CVaR relations

Proposition 5.

*If there is no probability atom at* \( \zeta_\alpha(x) \) *one simply has*

\[
\phi^-_\alpha(x) = \phi_\alpha(x) = \phi^+_\alpha(x),
\]

*If a probability atom does exist at* \( \zeta_\alpha(x) \) *and*

\[
\Psi(x, \zeta_\alpha(x)^-) < \alpha < \Psi(x, \zeta_\alpha(x)) < 1 \text{ we get}
\]

\[
\phi^-_\alpha(x) < \phi_\alpha(x) < \phi^+_\alpha(x).
\]
Atom in $\zeta_\alpha$

Figure: The picture supports previous proposition.
Proposition 6 (CVaR as a weighted average).

Let
\[ \lambda_\alpha(x) = \frac{\Psi(x, \zeta_\alpha(x)) - \alpha}{1 - \alpha} \in [0, 1]. \]

If \( \Psi(x, \zeta_\alpha(x)) < 1 \), so there is chance of a loss greater than \( \zeta_\alpha(x) \), then
\[ \phi_\alpha(x) = \lambda_\alpha(x) \zeta_\alpha(x) + (1 - \lambda_\alpha(x)) \phi_\alpha^+(x), \]

whereas if \( \Psi(x, \zeta_\alpha(x)) = 1 \), so \( \zeta_\alpha(x) \) is the highest loss that can occur (and thus \( \lambda_\alpha(x) = 1 \) but \( \phi_\alpha^+ \) is not defined) then
\[ \phi_\alpha(x) = \zeta_\alpha(x). \]
Basic CVaR relations

• The proof of previous proposition comes from the definition of VaR and CVaR.
• From the proposition we see that probability atoms can be split such that we compute the CVaR on the proper part which has probability $\alpha$.
• Surprising is that CVaR, which is coherent (see later) can be gained as linear combination of two non coherent risk measures.

**Corollary 7.**

$\alpha$-CVaR dominates $\alpha$-VaR: $\phi_\alpha(x) \geq \zeta_\alpha(x)$. 
Motivation  General concept of CVaR  Optimization  Comparison

CVaR for scenario models

- In this case we suppose that probability measure is discrete, it means it is concentrated only in atoms and there is finitely many of these atoms.

Proposition 8 (CVaR for scenario models).

Suppose the probability measure $P$ is concentrated in finitely many points $y_k \in Y$, so that for each $x \in X$ the distribution of loss $z = f(x, y)$ is likewise concentrated in finitely many points and $\Psi(x, \cdot)$ is a step function with jumps at those points. Fixing $x$, let those corresponding loss points be ordered as $z_1 < z_2 < \cdots < z_N$ and $P(f(x, y) = z_k) = p_k > 0$. Let $k_\alpha$ is the unique index such that

$$
\sum_{k=1}^{k_\alpha} p_k \geq \alpha > \sum_{k=1}^{k_\alpha-1} p_k.
$$
Proposition 8 (CVaR for scenario models).

The $\alpha$-VaR of the loss is given then by $\zeta_\alpha(x) = z_{k_\alpha}$, whereas the $\alpha$-CVaR is given by

$$
\phi_\alpha(x) = \frac{\left(\sum_{k=1}^{k_\alpha} p_k - \alpha\right) z_{k_\alpha} + \sum_{k=k_\alpha+1}^{N} p_k z_k}{1 - \alpha}.
$$

Further

$$
\lambda_\alpha(x) = \frac{\sum_{k=1}^{k_\alpha} p_k - \alpha}{1 - \alpha} \leq \frac{p_{k_\alpha}}{p_{k_\alpha} + \cdots + p_N}.
$$

- It is enough to realize that $\Psi(x, \zeta_\alpha(x)) = \sum_{k=1}^{k_\alpha} p_k$,
  $\Psi(x, \zeta_\alpha(x)^-) = \sum_{k=1}^{k_\alpha-1} p_k$ and
  $\Psi(x, \zeta_\alpha(x)) - \Psi(x, \zeta_\alpha(x)^-) = p_{k_\alpha}$. Then we use the previous proposition.
- The last inequality follows from the $\alpha > \sum_{k=1}^{k_\alpha-1} p_k$. 

Goal

- We will show that $\alpha$-VaR and $\alpha$-CVaR of the loss $z = f(x, y)$ with given strategy $x$ can be computed simultaneously by solving one dimensional optimization problem of convex nature.

$$F_{\alpha}(x, \zeta) = \zeta + \frac{\mathbb{E}(\lfloor f(x, y) - \zeta \rfloor^+)}{1 - \alpha},$$

where $[t]^+ = \max\{0, t\}$.

- The following proposition is not valid for CVaR$^+$ or CVaR$^-$. 
Proposition 9 (Fundamental minimization formula).

As a function of $\zeta \in \mathbb{R}$, $F_\alpha(x, \zeta)$ is finite and convex (hence continuous), with

$$\phi_\alpha(x) = \min_{\zeta} F_\alpha(x, \zeta),$$

$$\zeta_\alpha(x) = \text{lower endpoint of } \arg\min_{\zeta} F_\alpha(x, \zeta),$$

$$\zeta^+_\alpha(x) = \text{upper endpoint of } \arg\min_{\zeta} F_\alpha(x, \zeta),$$

where the argmin refers to the set of $\zeta$ for which the minimum is attained. In this case is nonempty, closed, bounded interval (point). In particular, one always has

$$\zeta_\alpha(x) \in \arg\min_{\zeta} F_\alpha(x, \zeta), \quad \phi_\alpha(x) = F_\alpha(x, \zeta_\alpha(x)).$$
Definition 10 (Sublinear function).

The function $h(x)$ is sublinear if $h(x + y) \leq h(x) + h(y)$ and $h(\lambda x) = \lambda h(x)$ for $\lambda > 0$. The second property is called positive homogeneity.

- From the second property we know $\lim_{t \to 0} h(t) = 0$.
- We know further that $h(t) \leq h(0) + h(t)$. This yields $h(0) \geq 0$.
- Finally $0 \leq h(0) \leq h(t) + h(-t)$, but $h(t)$ and $h(-t)$ can be arbitrarily small, which gives $h(0) = 0$.
- Sublinearity is equivalent to the combination of convexity with positive homogeneity [Rockafellar (1997)].
Proposition 11 (Convexity of CVaR).

If \( f(x, y) \) is convex with respect to \( x \), then \( \phi_\alpha(x) \) is also convex to \( x \). In this case \( F_\alpha(x, \zeta) \) is jointly convex in \( (x, \zeta) \).

According to previous note if \( f(x, y) \) is sublinear with respect to \( x \), then \( \phi_\alpha(x) \) is also sublinear to \( x \). In this case \( F_\alpha(x, \zeta) \) is jointly sublinear in \( (x, \zeta) \).

- The proof come from the fact that \( F_\alpha(x, \zeta) \) is convex when \( f(x, y) \) is convex in \( x \).
- For convexity of \( \phi_\alpha \) we need the fact that when a convex function of two vector variables is minimized with respect to one of them, the result is a convex function of the other see [Rockafellar (1997)].
Definition 12 (Coherence).

We consider some risk measure $\rho$ as a functional on a linear space of random variables, where $z = f(x, y)$ is our random variable then the first requirement for coherence is that $\rho$ has to be sublinear

$$\rho(z + z') \leq \rho(z) + \rho(z') \quad \rho(\lambda z) = \lambda \rho(z) \text{ for } \lambda \geq 0,$$

$$\rho(z) = c \text{ when } z \equiv c (\text{constant})$$

and

$$\rho(z) \leq \rho(z') \text{ when } z \leq z'.$$

- We have shown, that we can employ also strict inequality in definition of sublinearity.
- The inequality $z \leq z'$ refers to first order stochastic dominance (i.e. $\Pr(z \leq a) \leq \Pr(z' \leq a)$ for all possible $a$).
Proposition 13 (Coherence of CVaR).

*When* $f(x, y)$ *is linear with respect to* $x$ *then* $\alpha$–CVaR *is a coherent risk measure*. $\phi_\alpha(x)$ *is not only sublinear with respect to* $x$, *but further it satisfies*

$$\phi_\alpha(x) = c \text{ when } f(x, y) = c$$

*(thus accurately reflecting a lack of risk)*, *and it obeys the monotonicity rule that*

$$\phi_\alpha(x) \leq \phi_\alpha(x') \text{ when } f(x, y) \leq f(x', y).$$

- $f(x, y)$ *is linear when* $f(x, y) = x_1 f_1(y) + \cdots + x_n f_n(y)$.
- We do not need linearity in the proposition. The result is more general.
- The proof is consequence of fundamental minimization formula.
Optimization shortcut

- In problems of optimization under uncertainty, risk measure can enter into the objective or the constraints or both.
- In this context CVaR has a big advantage in its convexity.

Theorem 14 (Optimization shortcut).

Minimizing $\phi_\alpha(x)$ with respect to $x \in X$ is equivalent to minimizing $F_\alpha(x, \zeta)$ over all $(x, \zeta) \in X \times \mathbb{R}$, in the sense that

$$\min_{x \in X} \phi_\alpha(x) = \min_{(x, \zeta) \in X \times \mathbb{R}} F_\alpha(x, \zeta),$$

where moreover

$$(x^*, \zeta^*) \in \arg\min_{(x, \zeta) \in X \times \mathbb{R}} F_\alpha(x, \zeta) \iff$$

$$x^* \in \arg\min_{x \in X} \phi_\alpha(x), \quad \zeta^* \in \arg\min_{\zeta \in \mathbb{R}} F_\alpha(x^*, \zeta).$$
Optimization shortcut

- The proof is based on an idea that minimization with respect to \((x, \zeta)\) can be carried out by minimizing with respect to \(\zeta\) for each \(x\) then minimizing the rest with respect to \(x\).

**Corollary 15 (VaR and CVaR calculation).**

If \((x^*, \zeta^*)\) minimizes \(F_\alpha\) over \(X \times \mathbb{R}\), then not only does \(x^*\) minimize \(\phi_\alpha\) over \(X\), but also

\[
\phi_\alpha(x^*) = F_\alpha(x^*, \zeta^*), \quad \zeta_\alpha(x^*) \leq \zeta^* \leq \zeta_\alpha^+(x^*),
\]

where actually \(\zeta_\alpha(x^*) = \zeta^*\) if \(\arg\min_\zeta F_\alpha(x^*, \zeta)\) reduces to a single point.
Optimization shortcut

- If $\text{argmin}_\zeta F_\alpha(x^*, \zeta)$ does not consist of just a single point, it is possible to have $\zeta_\alpha(x^*) < \zeta^*$. In this case the joint minimization in producing $(x^*, \zeta)$ yields the CVaR associated with $x^*$, it does not immediately yield the VaR associated with $x^*$.

- The fact that minimization of CVaR (if the problem is convex) does not have to proceed numerically through repeated calculations of $\phi_\alpha(x)$ for various decisions $x$ is powerful attraction to working with CVaR.

- VaR can be ill behaved and does not offer such a shortcut.
Outline

1. Motivation
2. General concept of CVaR
3. Optimization
4. Comparison
Pros of VaR

- Very easy idea behind, easy to understand and interpret. How much you may lose with certain confidence level.
- Two distributions can be ranked by comparing their VaRs for the same confidence level.
- VaR focuses on a specific part of the distribution specified by the confidence level.
- Stability of estimation procedure. It is not affected by very high tail losses.
- It is often estimated parametrically (historical or Monte Carlo simulation or by using approximations based on second-order Taylor expansion).
Cons of VaR

- VaR does not account for properties of the distribution beyond the confidence level.
- This can lead to taking undesirable risk.
- Risk control using VaR may lead to undesirable results for skewed distributions.
- VaR is a nonconvex and discontinuous function for discrete distributions w.r.t. portfolio positions when returns have discrete distributions.
Pros of CVaR

- Its interpretation is still straightforward. It measures the losses that hurts the most.
- Defining CVaR \((X)\) for all confidence levels \(\alpha\) in \((0, 1)\) completely specifies the distribution of \(X\).
- CVaR is a coherent risk measure.
- CVaR \((X)\) is continuous with respect to \(\alpha\).
- CVaR \((w_1X_1 + \cdots + w_nX_n)\) is a convex function with respect to \((w_1, \ldots, w_n)\).
Cons of CVaR

- CVaR is more sensitive than VaR to estimation errors.
- For instance, historical scenarios often do not provide enough information about tails; hence, we should assume a certain model for the tail to be calibrated on historical data.
- In financial setting, equally weighted portfolios may outperform CVaR-optimal portfolios when historical data have mean reverting characteristics.
- Slightly worse interpretation, not so used in practise.
Motivation

General concept of CVaR

Optimization

Comparison

**VaR or CVaR**

- VaR and CVaR measure different parts of the distribution. Depending on what is needed, one may be preferred over the other.
- A trader may prefer VaR to CVaR, because he may like high uncontrolled risks; VaR is not as restrictive as CVaR with the same confidence level.
- Nothing dramatic happens to a trader in case of high losses.
- A company owner will probably prefer CVaR; he has to cover large losses if they occur; hence, he “really” needs to control tail events.
- VaR may be better for optimizing portfolios when good models for tails are not available.
**VaR or CVaR**

- CVaR may not perform well out of sample when portfolio optimization is run with poorly constructed set of scenarios.

- Historical data may not give right predictions of future tail events because of mean-reverting characteristics of assets. High returns typically are followed by low returns; hence, CVaR based on history may be quite misleading in risk estimation.

- If a good model of tail is available, then CVaR can be accurately estimated and CVaR should be used. CVaR has superior mathematical properties and can be easily handled in optimization and statistics.

- Appropriate confidence levels for VaR and CVaR must be chosen, avoiding comparison of VaR and CVaR for the same level of $\alpha$ because they refer to different parts of the distribution.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Title</th>
<th>Publication Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Shapiro, D. Dentcheva, A. Ruszczyński</td>
<td>Lectures on stochastic programming modeling and theory</td>
<td>Mathematical modeling society, 2009.</td>
</tr>
</tbody>
</table>
Time for your questions
Thank you for your attention