Financial Optimization
ISE 347/447

Lecture 23

Dr. Ted Ralphs
Reading for This Lecture

- C&T Chapter 17
Value at Risk (VaR)

• *Value at risk* is a risk measure developed at J.P. Morgan.
• Today it is in wide-spread use across the finance industry.
• $\text{VaR}_\alpha$ is defined as the lowest level for which the probability of experiencing a loss above this level is smaller than $1 - \alpha$.
• In other words, the loss will not exceed $\text{VaR}_\alpha$ with probability $\alpha$.
• Let us now define this notion more formally.
Definitions

• Consider an investment decision represented by the vector $x \in \mathbb{R}^n$.

• Let the loss over an investment period under the outcome $\omega \in \Omega$ be $L(x, \omega)$.

• Thus, for fixed $x$, the loss function $L(x, \cdot)$ is a random variable that takes positive values when a loss is incurred, and negative ones when a gain occurs.

• For any fixed value of $x$ let

$$
\Psi(x, \gamma) := P \left[ L(x, \cdot) \leq \gamma \right] = F_{L(x, \cdot)}(\gamma)
$$

be the cumulative distribution function of the loss function $L(x, \cdot)$ associated with holding the investment $x$. 
Definitions (cont.)

For any $\alpha \in [0, 1]$ (typically, $\alpha = .95$ is chosen), the value at risk on the confidence level $\alpha$ is defined by

$$\text{VaR}_\alpha(x) := \min_{\gamma \in \mathbb{R}} \gamma$$

$$\text{s.t.} \quad \Psi(x, \gamma) \geq \alpha.$$ 

Since $\Psi$ is typically a nonlinear function, computing the value at risk is a nonlinear programming problem.
Example 1

- A set of risky assets $S^1, \ldots, S^n$ have multivariate normal returns $R \sim N(\mu, Q)$ over the investment period $[0, 1]$.

- Suppose we want to find the portfolio $x^*$ that minimizes the value at risk on the confidence level $\alpha$.

- If the total value of the invested capital is $w$, then the loss incurred by the portfolio $x$ over the investment period is $-wR^Tx$.

- Therefore, we have to solve a bilevel optimization problem (see next slide).
Portfolio Optimization with VaR

(VM1) \[ x^* = \arg \min_{x \in \mathbb{R}^n} \text{VaR}_\alpha(x) \]

s.t. \[ Ax \geq a, \quad Bx = b, \]

where the objective function

\[
\text{VaR}_\alpha(x) = \min_{\gamma \in \mathbb{R}} \gamma \\
\text{s.t.} \quad \int_{\{r: -wr^\top x \leq \gamma\}} \exp \left\{ -\frac{1}{2}(r - \mu)^\top Q^{-1}(r - \mu) \right\} \frac{dr}{\sqrt{(2\pi)^n \det(Q)}} \geq \alpha
\]

is itself the optimal solution to an optimization problem.
Example 2

• The return vector $R$ of a set of risky assets $S^1, \ldots, S^n$ takes the values $r^1, \ldots, r^k \in \mathbb{R}^n$ with probability $1/k$ each.

• Find the vector $x^*$ of relative wealth allocation weights that minimizes the value at risk on the confidence level $\alpha$.

\[
(VM2) \quad x^* = \arg \min_{x \in \mathbb{R}^n} \text{VaR}_\alpha(x) \\
\text{s.t.} \quad Ax \geq a, \ Bx = b,
\]

with

\[
\text{VaR}_\alpha(x) = \min_{\gamma \in \mathbb{R}} \gamma \\
\text{s.t.} \quad \sum_{\{i : -wx^\top r^i \leq \gamma\}} \frac{1}{k} \geq \alpha.
\]
**Drawbacks**

$\text{VaR}_\alpha$-minimization has a number of serious drawbacks:

- The objective function $\text{VaR}_\alpha$ in Example 2 is nonlinear and nonsmooth with many local minimizers, rendering the problem (VM2) computationally difficult.

- Both (VM1) and (VM2) are bilevel optimization problems, which are generally computationally harder to solve than single-level problems.

- $\text{VaR}_\alpha$ is not subadditive: A good risk measure $f(x)$ should satisfy the inequality
  \[ f(x_1 + x_2) \leq f(x_1) + f(x_2), \]  
  that is, the risk of holding both investments $x_1$ and $x_2$ should be no larger than the sum of the risks of holding each individually. This guarantees that diversification reduces risk. However, when $\text{VaR}_\alpha$ is used as a risk-measure, diversification can actually increase the risk!

- $\text{VaR}_\alpha$ pays no attention to the magnitude of losses when the rare extremal event of experiencing a loss above the level $\text{VaR}_\alpha$ occurs.
Conditional Value at Risk (CVaR)

- To overcome this drawback, the notion of *conditional value at risk* (CVaR) has been developed.

- This is the same as *mean expected loss*, *mean shortfall*, *expected shortfall risk* and *tail-VaR*.

- As before, let $\alpha \in [0, 1]$ be a given confidence level.

- Then we define

\[
\text{CVaR}_\alpha(x) := \mathbb{E} [L(x, \omega) \mid L(x, \omega) \geq \text{VaR}_\alpha(x)] \\
= \frac{1}{1 - \alpha} \int_{\{\omega : L(x, \omega) \geq \text{VaR}_\alpha(x)\}} L(x, \omega) \, P[d\omega].
\]
**Example 3**

- A given investment generates losses of $L(j) = j - 80$ $(j = 1, \ldots, 100)$ each with probability 1%.

- We have

$$\text{VaR}_{0.95} = \min_{j=1,\ldots,100} L(j)$$

s.t. $\sum_{i=1}^{j} \frac{1}{100} \geq 0.95$.

- The constraint is satisfied for $j = 95, \ldots, 100$. Therefore,

$$\text{VaR}_\alpha = \min_{j=95,\ldots,100} (j - 80) = 15.$$

- The expected shortfall risk is

$$\text{CVaR}_{0.95} = \frac{1}{0.05} \sum_{j=95}^{100} \frac{j - 80}{100} = 17.5.$$
Comparing VaR and CVaR

• Note that

\[ \text{CVaR}_\alpha(x) \geq \frac{1}{1 - \alpha} \int_{\{\omega: L(x,\omega) \geq \text{VaR}_\alpha(x)\}} \text{VaR}_\alpha(x) P[d\omega] \]

\[ = \frac{\text{VaR}_\alpha(x)}{1 - \alpha} P[L(x, \omega) \geq \text{VaR}_\alpha(x)] \geq \text{VaR}_\alpha(x), \]

so minimizing CVaR\(_\alpha\) also makes VaR\(_\alpha\) small, but the opposite may not be true.

• CVaR\(_\alpha(x)\) can now be used as a risk measure in investment decision problems that take the form

\[(\text{CVM}) \quad x^* = \arg \min_{x \in \mathbb{R}^n} \text{CVaR}_\alpha(x) \]

\[\text{s.t.} \quad x \in \mathcal{F},\]

where \(\mathcal{F}\) is some set of feasible investments defined by a set of constraints.
Example 4

In Example 1, if we had proposed to find an investment that minimizes $\text{CVaR}_\alpha$, we would have had to solve

$$(\text{CVM1}) \quad x^* = \arg \min_{x \in \mathbb{R}^n} \text{CVaR}_\alpha(x)$$

subject to
$$Ax \geq a, \quad Bx = b,$$

where

$$\text{CVaR}_\alpha(x) = \frac{-w}{1 - \alpha} \int_{\{r: -wr^\top x \geq \text{VaR}_\alpha(x)\}} \frac{r^\top x \cdot \exp \left\{ -\frac{1}{2}(r - \mu)^\top Q^{-1}(r - \mu) \right\}}{\sqrt{(2\pi)^n \det(Q)}} \, dr$$

and

$$\text{VaR}_\alpha(x) = \min_{\gamma \in \mathbb{R}} \gamma$$

subject to
$$\int_{\{r: -wr^\top x \leq \gamma\}} \frac{\exp \left\{ -\frac{1}{2}(r - \mu)^\top Q^{-1}(r - \mu) \right\}}{\sqrt{(2\pi)^n \det(Q)}} \, dr \geq \alpha.$$
Example 5

In Example 2, if we had proposed to find an investment that minimizes $\text{CVaR}_\alpha$, we would have had to solve

$$(\text{CVM2}) \quad x^* = \operatorname{arg\ min}_{x \in \mathbb{R}^n} \text{CVaR}_\alpha(x)$$

subject to $Ax \geq a, Bx = b$,

where

$$\text{CVaR}_\alpha(x) = \frac{1}{|I|} \sum_{i \in I} -wx^\top r^i,$$

$$I = \{i : -wx^\top r^i \geq \text{VaR}_\alpha(x)\},$$

and

$$\text{VaR}_\alpha(x) = \min_{\gamma \in \mathbb{R}} \gamma$$

subject to

$$\sum_{\{i : -wx^\top r^i \leq \gamma\}} \frac{1}{k} \geq \alpha.$$
Computing CVaR

• These examples illustrate that computing $\text{CVaR}_\alpha(x)$ generally requires the computation of $\text{VaR}_\alpha(x)$.

• This suggests that the $\text{CVaR}_\alpha$-minimization problem

$$(\text{CVM}) \quad x^* = \arg \min_{x \in \mathbb{R}^n} \text{CVaR}_\alpha(x)$$

s.t. $x \in \mathcal{F}$

might be even harder than the $\text{VaR}_\alpha$-minimization problem

$$(\text{VM}) \quad x^* = \arg \min_{x \in \mathbb{R}^n} \text{VaR}_\alpha(x)$$

s.t. $x \in \mathcal{F}$.

• It thus comes as a surprise that under quite reasonable modeling assumptions, the opposite is true.
Let $\beta(x) = P[L(x, \omega) \geq \text{VaR}_\alpha(x)]$ and consider the auxiliary function

$$F_\alpha(x, \gamma) := \gamma + \int_\Omega \frac{(L(x, \omega) - \gamma)_+}{\beta(x)} P[d\omega].$$

**Theorem 1.**

i) For any fixed $x$, the function $\gamma \mapsto F_\alpha(x, \gamma)$ is convex.

ii) $\text{VaR}_\alpha(x)$ is a minimizer of the problem $\min_\gamma F_\alpha(x, \gamma)$.

iii) $F_\alpha(x, \text{VaR}_\alpha(x)) = \text{CVaR}_\alpha(x)$. 
**Proof:** i) Since \((L(x, \omega) - \gamma)_+\) is a convex function in \(\gamma\), it is true that for any \(\gamma_1, \gamma_2\) and \(\tau \in [0, 1]\),

\[
F_\alpha(x, \tau \gamma_1 + (1 - \tau) \gamma_2) \\
\leq \tau \gamma_1 + (1 - \tau) \gamma_2 \\
+ \int_\Omega \left( \tau \frac{(L(x, \omega) - \gamma_1)_+}{\beta(x)} + (1 - \tau) \frac{(L(x, \omega) - \gamma_2)_+}{\beta(x)} \right) P[d\omega] \\
= \tau F_\alpha(x, \gamma_1) + (1 - \tau) F_\alpha(x, \gamma_2).
\]

This shows that \(F_\alpha(x, \gamma)\) is convex in \(\gamma\).
ii) Since the problem of minimizing $F_\alpha(x, \gamma)$ with respect to $\gamma$ is convex, the KKT conditions are sufficient for optimality, i.e., we only need to check that the $F_\alpha(x, \gamma)$ is stationary at $\gamma = \text{VaR}_\alpha(x)$.

For any set $S \subset \Omega$ let $\chi_S$ be the associated indicator function

$$
\chi_S(\omega) = \begin{cases} 
1 & \text{if } \omega \in S, \\
0 & \text{otherwise}.
\end{cases}
$$

With this notation we have

$$
\frac{\partial}{\partial \gamma} F_\alpha(x, \text{VaR}_\alpha(x)) = 1 - \int_\Omega \frac{\chi\{\omega: L(x, \omega) \geq \text{VaR}_\alpha(x)\}(\omega)}{\beta(x)} P[d\omega] = 1 - \frac{P[L(x, \omega) \geq \text{VaR}_\alpha(x)]}{\beta(x)} = 0.
$$
iii) We have

\[ F_\alpha(x, \text{VaR}_\alpha(x)) = \text{VaR}_\alpha(x) + \int_\Omega \frac{(L(x, \omega) - \text{VaR}_\alpha(x))}{\beta(x)} P[d\omega] \]

\[ = \text{VaR}_\alpha(x) + \int_{\{\omega : L(x, \omega) \geq \text{VaR}_\alpha(x)\}} \frac{L(x, \omega)}{\beta(x)} P[d\omega] \]
\[ - \text{VaR}_\alpha(x) \frac{P[L(x, \omega) \geq \text{VaR}_\alpha(x)]}{\beta(x)} \]

\[ = \text{VaR}_\alpha(x) + \text{CVaR}_\alpha(x) - \text{VaR}_\alpha(x). \]
Minimizing $\text{CVaR}_\alpha$

- Theorem 1 now implies that the $\text{CVaR}_\alpha$-minimization problem

$$(\text{MCV}) \quad \min_{x \in \mathbb{R}^n} \text{CVaR}_\alpha(x)$$

s.t. $x \in \mathcal{F}$

can be reformulated as the single-level optimization problem

$$(\text{MCV}') \quad \min_{(x, \gamma) \in \mathbb{R}^{n+1}} F_\alpha(x, \gamma)$$

s.t. $x \in \mathcal{F}$.

- In applications, it is often the case that $F_\alpha$ is convex in $x$ as well, and $\mathcal{F}$ is a convex set.

- In this case $(\text{MCV}')$ is a convex minimization problem and can generally be well solved.
Example 6

- Problem (CVM1) from Example 4 is equivalent to

\[
\begin{align*}
\text{(CVM1')} & \quad \min_{x, \gamma} \gamma + \frac{1}{1 - \alpha} \int_{\mathbb{R}^n} (-wr^\top x - \gamma) + \exp \left\{ -\frac{1}{2} (r - \mu)^\top Q^{-1} (r - \mu) \right\} \frac{d}{\sqrt{(2\pi)^n \det(Q)}} \\
\text{s.t. } & \quad Ax \geq a, \quad Bx = b.
\end{align*}
\]

- Since \((-wr^\top x - \gamma)_+\) is convex in \(x\), the objective function of (CVM1') is a positive combination of convex functions and hence also convex in \(x\).

- By Theorem 1 the objective function is also convex in \(\gamma\).
Example 7

• Problem (MCV2) from Example 5 is equivalent to

\[
(MCV2') \quad \min_{x, \gamma} \gamma + \frac{1}{\beta(x)} \sum_{i=1}^{k} \frac{(-wx^\top r^i - \gamma)_+}{k}
\]

s.t. \( Ax \geq a, \ Bx = b. \)

• Since \( \beta(x) \approx 1 - \alpha \), Problem (MCV2) can be approximated by the convex problem

\[
(MCV2') \quad \min_{x, \gamma} \gamma + \frac{1}{1 - \alpha} \sum_{i=1}^{k} \frac{(-wx^\top r^i - \gamma)_+}{k}
\]

s.t. \( Ax \geq a, \ Bx = b. \)
Example 7 (cont.)

- Finally, Problem (MCV2') is equivalent to the following LP,

\[(\text{LMCV2'}) \quad \min_{x,z,\gamma} \gamma + \frac{1}{(1 - \alpha)k} \sum_{i=1}^{k} z_i\]

\[\text{s.t.} \quad z_i \geq -wx^\top r^i - \gamma, \quad (i = 1, \ldots, k)\]

\[Ax \geq a, \quad Bx = b,\]

\[z \geq 0,\]

- Note that we replaced a piecewise linear convex objective function by a linear objective by introducing extra variables and extra linear constraints.

- This is the same thing we did in the L-shaped method.
General Techniques

- Example 7 can be generalized to approximate any CVaR$_\alpha$-minimization problem via LP or QP:

- For this purpose we replace the probability measure $P$ on $\Omega$ by a finite set of equiprobable scenarios $\omega_1, \ldots, \omega_S$.

- These scenarios are typically obtained by statistical sampling.

- Next, we approximate $F_\alpha$ by

$$F_\alpha(x, \gamma) = \gamma + \frac{1}{(1 - \alpha)S} \sum_{s=1}^{S} (L(x, \omega_s) - \gamma)_+,$$

so that the problem (MCV) can be approximated.
Approximating

The approximation is then

\[(AMCV) \quad \min_{x, \gamma} \gamma + \frac{1}{(1 - \alpha)S} \sum_{s=1}^{S} (L(x, \omega_s) - \gamma)_+ \]

s.t. \(x \in \mathcal{F} \).

Introducing artificial variables to get rid of the break points of the objective function, we replace (AMCV) by the equivalent problem

\[(LAMCV) \quad \min_{x, z, \gamma} \gamma + \frac{1}{(1 - \alpha)S} \sum_{i=1}^{S} z_s \]

s.t. \(z_s \geq 0, \quad (s = 1, \ldots, S)\)

\(z_s \geq L(x, \omega_s) - \gamma, \quad (s = 1, \ldots, S)\)

\(Ax \geq b.\)
Remarks

- If $L(x, \omega)$ is linear in $x$, then (LAMCV) is an LP.

- More generally, $L(x, \omega)$ is typically convex in $x$, in which case (LAMCV) is well solved via standard NLP software.

- In applications in which $L(x, \omega)$ is not convex in $x$, (LAMCV) is often further approximated by replacing $L(x, \omega)$ by an approximation that is convex in $x$.

- Typically, NLP software will do this automatically.
Further Applications

• In risk management, one is often interested in controlling the expected loss at several confidence levels.

• The following model is typical,

\[
(RM) \quad \max_x \mu^\top x \quad \text{s.t.} \quad \text{CVaR}_{\alpha_j}(x) \leq u_{\alpha_j}, \quad (j = 1, \ldots, k) \\
A x \geq a, \quad B x = b.
\]

• To control the risk of the investment \( x \), we thus require that the conditional value at risk must not exceed thresholds \( u_{\alpha_j} \) on the confidence levels \( \alpha_1, \ldots, \alpha_k \).
Adapting the LAMCV

The reformulation of the finite scenario case can easily be adapted to such problems, which now become

\[
\begin{align*}
\text{(ARM)} \quad & \max_{x, \gamma, z} \mu^\top x \\
\text{s.t.} \quad & \gamma + \frac{1}{(1 - \alpha_j)S} \sum_{s=1}^{S} z_s \leq U_{\alpha_j}, \quad (j = 1, \ldots, k) \\
\quad & z_s \geq 0, \quad (s = 1, \ldots, S) \\
\quad & z_s \geq L(x, \omega_s) - \gamma, \quad (s = 1, \ldots, S), \\
\quad & Ax \geq a, \quad Bx = b.
\end{align*}
\]