RUTCOR
RESEARCH
REPORT

Decomposing CVaR minimization in two-stage stochastic models

Csaba I. Fábián[a]

RRR 32-2005, October 2005

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/∼rrr

[a]Department of Operations Research, Loránd Eötvös University of Budapest. P.O.Box 120, Budapest, H-1518 Hungary. E-mail: fabian@cs.elte.hu.

Rutcor Research Report
RRR 32-2005, October 2005

Decomposing CVaR minimization in two-stage stochastic models

Csaba I. Fábián

Abstract. Based on the polyhedral representation of Künzi-Bay and Mayer (2005), we propose a decomposition framework for the minimization of CVaR in two-stage stochastic models.
We show that the decomposed problems can be effectively solved by a special inexact-cut method.

Introduction

Value-at-risk (VaR) is a widely accepted risk measure that is not easily managed in optimizition problems. An alternative risk measure, namely conditional value-at-risk (CVaR), has been proposed by Rockafellar and Uryasev (2000). They derived a representation of CVaR as the optimum of a special minimization problem. This representation makes CVaR tractable in optimization problems. An overview of VaR- and CVaR-optimization methods can be found in Prékopa (2003).

In Section 1 of this paper we cite the representation result of Rockafellar and Uryasev. We also obtain a straightforward new proof of the Rockafellar-Uryasev representation formula, in the special case of discrete distributions.

Recently Künzi-Bay and Mayer proposed a new method for the minimization of CVaR in one-stage stochastic problems. They consider these problems as special two-stage problems, and present a special decomposition method for the solution of such two-stage problems. Their method is based on a polyhedral representation of their special master problem.

Künzi-Bay and Mayer implemented their method, and present extensive test results demonstrating the superiority of their solver over general-purpose solvers, in the solution of CVaR-minimization problems.

In Section 2 of this paper we outline the polyhedral representation proposed by Künzi-Bay and Mayer. Morevore, we propose another polyhedral representation that is similar to, but simpler than theirs. We trust in the usability of this simpler representation for the solution of one-stage CVaR-minimization problems.

In Section 3 we present a decomposition framework for the minimization of CVaR in two-stage models. This framework is based on the polyhedral representation of Künzi-Bay and Mayer.

The two-stage prototype model we examine prescribes the minimization of the CVaR of the end-of-horizon yield. This type of objective function had been used earlier by Topaloglou (2004). He developed elaborate financial models, even in multi-stage frameworks. He solved one-stage and two-stage problems with commercially available, general-purpose modeling systems and optimization solvers. From an algorithmic point of view, his two-stage models fit the prototype model we examine. Hence the decomposition framework we propose can also be applied to his financial problems.

We also show that the two-stage prototype problem can be effectively solved by the Level Decomposition Method.

In Section 4 we propose an improvement on the prototype model.

1 Conditional value-at-risk. Formulas for discrete distributions

Let us consider a one-period financial investment.

w denotes the total wealth at the end of the examined period. This is a random variable.

$w^{\mathcal{B}}$ denotes a benchmark for end-of-period wealth (i.e., the wealth that we intend to accumulate by the end of the examined period). We assume it is a parameter that has been set by the decision maker.

Then the loss relative to the benchmark can be expressed as: $w^{\mathcal{B}} - w$. Given a probability α, a heuristic definition of the risk measures is the following.

α-*Value-at-Risk* (VaR) answers the question: what is the maximum loss with the confidence level $\alpha * 100\%$?

α-*Conditional Value-at-Risk* (CVaR) is the (conditional) mean value of the worst $(1-\alpha) * 100\%$ losses.

Rockafellar and Uryasev (2000) proved that α-VaR and α-CVaR can be computed through the solution of the following problem:

$$\min_{z \in \mathbb{R}} \; z + \frac{1}{1-\alpha} \mathrm{E}\left([w^{\mathcal{B}} - w - z]_+ \right). \tag{1}$$

Namely, VaR is the optimal value of z; and CVaR is the optimal objective value.

In this paper we assume that w has a discrete distribution. Let the realizations be $w^{(1)}, \dots, w^{(N)}$ with probabilities $p_1, \dots, p_N$, respectively. Problem (1) takes the form

$$\min_{z \in \mathbb{R}} \; z + \frac{1}{1-\alpha} \sum_{j=1}^{N} p_j \left[w^{\mathcal{B}} - w^{(j)} - z \right]_+ . \tag{2}$$

Rockafellar and Uryasev (2002) showed (in their Proposition 8) that the above problem can be solved by just sorting the values $w^{(j)}$. We are going to reach the same conclusion through a different approach.

Rockafellar and Uryasev (2000) proposed transforming (2) into a linear programming problem by introducing new variables y_j $(j = 1, \dots, N)$:

$$\begin{gathered} \min \; z + \tfrac{1}{1-\alpha} \sum_{j=1}^{N} p_j y_j \\ \text{such that} \\ z + y_j \geq w^{\mathcal{B}} - w^{(j)} \quad (j = 1, \dots, N), \\ y_j \geq 0 \quad (j = 1, \dots, N). \end{gathered} \tag{3}$$

The dual of (3) can be written as

$$\max\ w^{\mathcal{B}} - \tfrac{1}{1-\alpha}\sum_{j=1}^{N} \pi_j\, w^{(j)}$$

$$\text{such that}$$

$$0 \le \pi_j \le p_j \quad (j = 1, \ldots, N), \tag{4}$$

$$\sum_{j=1}^{N} \pi_j = 1 - \alpha.$$

(In the objective function, we used $\frac{1}{1-\alpha}\sum_{j=1}^{N} \pi_j w^{\mathcal{B}} = w^{\mathcal{B}}$ that is a consequence of the constraint $\sum_{j=1}^{N} \pi_j = 1 - \alpha$.)

Since problem (4) has a single constraint, it can be solved without using linear programming algorithms. We only need sorting the objective coefficients $w^{(j)}$; assume that $w^{(1)} \le \cdots \le w^{(j)} \le \cdots \le w^{(N)}$ holds. Let $J \in \{1, \ldots, N-1\}$ be such that

$$P_J := \sum_{j=1}^{J-1} p_j\ < 1 - \alpha \quad \text{and} \quad \sum_{j=1}^{J} p_j \ge 1 - \alpha.$$

The optimal dual solution is

$$\hat{\pi}_j = \begin{cases} p_j & \text{for} \quad 1 \le j < J, \\ 1 - P_J & \text{for} \quad j = J, \\ 0 & \text{for} \quad J < j \le N. \end{cases} \tag{5}$$

Remark 1 *The dual problem (4) clearly expresses the (conditional) mean value of the worst* $(1-\alpha) * 100\%$ *losses. Indeed,*

the dual variable $0 \le \pi_j \le p_j$ *can be interpreted as the weight of the* j*th scenario in an event* $\mathcal{E}$*,*

the constraint $\sum_{j=1}^{N} \pi_j = 1 - \alpha$ *determines the probability of the event* $\mathcal{E}$*,*

the term $\frac{1}{1-\alpha}\sum_{j=1}^{N} \pi_j\, w^{(j)}$ *in the objective function can be interpreted as the conditional expectation of the end-period wealth given that* $\mathcal{E}$ *occurs. By maximizing the gap between the benchmark* $w^{\mathcal{B}}$ *and the above conditional expectation, we find the worst* $(1-\alpha)*100\%$ *cases.*

Hence the above discussion is a straightforward new proof of the validity of the Rockafellar-Uryasev representation formula (2) in case of discrete distributions.

2 Minimizing CVaR in a one-stage model

The end-of-period wealth w will be the yield of a *portfolio* selected by the decision maker. Assume there are n assets.

$\boldsymbol{x} = (x_1, \ldots, x_i, \ldots, x_n)^T$ represents the amounts of money invested in the different assets at the beginning of the examined period. (I.e., $\boldsymbol{x}$ is a portfolio.)

Let $X \subset \mathbb{R}^n$ denote the set of the feasible portfolios. Assume that X is a convex bounded polyhedron.

$\boldsymbol{r} = (r_1, \ldots, r_i, \ldots, r_n)^T$ denotes the returns for different assets. This is a random vector.

There are N realizations,

$$\boldsymbol{r}^{(1)}, \ldots, \boldsymbol{r}^{(j)}, \ldots, \boldsymbol{r}^{(N)}, \quad \text{where} \quad \boldsymbol{r}^{(j)} = \left(r_1^{(j)}, \ldots, r_i^{(j)}, \ldots, r_n^{(j)}\right)^T.$$

The probabilities of the different realizations are $p_1, \ldots, p_N$.

The total wealth at the end of the examined period is $w = \boldsymbol{r}^T\boldsymbol{x}$. This is a random variable with realizations $w^{(j)} = \boldsymbol{r}^{(j)\,T}\boldsymbol{x}$

Using (2) we can compute CVaR as a function of $\boldsymbol{x}$:

$$\mathcal{C}(\boldsymbol{x}) := \quad \min_{z \in \mathbb{R}} \; z + \frac{1}{1-\alpha} \sum_{j=1}^{N} p_j \left[w^{\mathcal{B}} - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} - z \right]_+ . \tag{6}$$

$\mathcal{C}(\boldsymbol{x})$ is a polyhedral convex function.

Suppose we want to find a balance between the end-of-period expected wealth $\mathrm{E}(\boldsymbol{r})^T\boldsymbol{x}$ and the risk as measured by $\mathcal{C}(\boldsymbol{x})$. A customary formulation of the problem is:

$$\begin{gathered} \min \; \boldsymbol{c}^T\boldsymbol{x} + \mathcal{C}(\boldsymbol{x}) \\ \text{such that} \\ \boldsymbol{x} \in X, \end{gathered} \tag{7}$$

where $\boldsymbol{c} := -\lambda \mathrm{E}(\boldsymbol{r})$ with a parameter $\lambda > 0$ that is interpreted as *risk tolerance* of the decision maker. We assume a known fixed λ.

Künzi-Bay and Mayer (2005) observe that problems of minimizing CVaR fit the prototype of the two-stage stochastic programming problem. We sketch their approach for the special problem (7). The master problem is:

$$\min\ \boldsymbol{c}^T\boldsymbol{x} + z + \tfrac{1}{1-\alpha}\sum_{j=1}^{N} p_j C_j(\boldsymbol{x}, z)$$

$$\text{such that}$$

$$z \in \mathbb{R}, \quad \boldsymbol{x} \in X.$$

The function value $C_j(\boldsymbol{x}, z)$ is defined by the recourse problem:

$$C_j(\boldsymbol{x}, z) := \min\ y \quad \text{such that} \quad y \geq w^{\mathcal{B}} - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} - z, \quad y \geq 0.$$

The dual of the above recourse problem has a single variable and takes the form: $\max\limits_{0\leq u\leq 1}\left(w^{\mathcal{B}} - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} - z\right)$ The optimal solution is either $u = 1$ or $u = 0$ depending on the sign of the objective coefficient $w^{\mathcal{B}} - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} - z$. On the basis of these observations, Künzi-Bay and Mayer propose an interesting equivalent formulation of the CVaR-minimization problem (7):

$$\begin{gathered}\min\ \boldsymbol{c}^T\boldsymbol{x} + z + \tfrac{1}{1-\alpha}v \\ \text{such that} \\ z, v \in \mathbb{R}, \quad \boldsymbol{x} \in X, \\ \sum_{j\in\mathcal{J}} p_j\left(w^{\mathcal{B}} - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} - z\right) \leq v \qquad (\,\mathcal{J} \subset \{1,\ldots,N\}\,)\,.\end{gathered} \tag{8}$$

A subset $\mathcal{J} \subset \{1,\ldots,N\}$ represents the possibility that we have 1 as optimal dual solution for the recourse problems belonging to $j \in \mathcal{J}$; and 0 for those belonging to $j \notin \mathcal{J}$. The cut belonging to the empty set $\mathcal{J} = \emptyset$ just prescribes the non-negativity of v. (Künzi-Bay and Mayer observe that the formulation (8) is the CVaR-analogue of a polyhedral representation result of Klein Haneveld and Van der Vlerk (2002).)

Künzi-Bay and Mayer adapt the L-shaped method to their special two-stage problem. They use cuts of the type that appear in the formulation (8).

An alternative method for the solution of the CVaR-minimization problem (7) is obtained by computing the CVaR function using the dual formulation (4):

$$\begin{aligned} \mathcal{C}(\boldsymbol{x}) = \quad & \max\ w^{\mathcal{B}} - \tfrac{1}{1-\alpha} \sum_{j=1}^{N} \pi_j\, \boldsymbol{r}^{(j)\,T} \boldsymbol{x} \\ & \text{such that} \\ & 0 \le \pi_j \le p_j \quad (j = 1, \ldots, N), \\ & \sum_{j=1}^{N} \pi_j = 1 - \alpha. \end{aligned} \tag{9}$$

Let $(\hat{\pi}_1, \ldots, \hat{\pi}_N)$ denote a feasible solution of problem (9). Then the linear function

$$l(\boldsymbol{x}) := w^{\mathcal{B}} - \frac{1}{1-\alpha} \sum_{j=1}^{N} \hat{\pi}_j\, \boldsymbol{r}^{(j)\,T} \boldsymbol{x} \tag{10}$$

obviously satisfies $l(\boldsymbol{x}) \le \mathcal{C}(\boldsymbol{x}) \quad (\boldsymbol{x} \in \mathbb{R}^n)$.

Given $\hat{\boldsymbol{x}} \in \mathbb{R}^n$, let $(\hat{\pi}_1, \ldots, \hat{\pi}_N)$ denote now an optimal solution of problem (9: $\boldsymbol{x} = \hat{\boldsymbol{x}}$). Then the linear function $l(\boldsymbol{x})$ constructed according to (10) will be a support function of $\mathcal{C}(\boldsymbol{x})$ at $\hat{\boldsymbol{x}}$. As shown in Section 1, problem (9: $\boldsymbol{x} = \hat{\boldsymbol{x}}$) can be solved by just sorting the potential yields $\boldsymbol{r}^{(j)\,T} \hat{\boldsymbol{x}} \quad (1 \le j \le N)$. Easy computation of the support functions suggests a direct cutting-plane or bundle-type approach for the solution of the CVaR-minimization problem (7).

In order to show similarity of the above alternative approach to the approach proposed by Künzi-Bay and Mayer, let us assume for the remaining part of this section that we have $p_1 = \ldots = p_N = \frac{1}{N}$. Assume moreover that $(1-\alpha)$ is an integer multiple of $\frac{1}{N}$. Hence $\frac{1-\alpha}{1/N} = N(1-\alpha)$ is an integer. The corresponding special form of (9) will be denoted by (9: $\boldsymbol{p} = \frac{\mathbf{1}}{N}$).

Proposition 2 *The following polyhedral representation is equivalent to problem (9: $\boldsymbol{p} = \frac{\mathbf{1}}{N}$).*

$$\begin{aligned} & \min\ \nu \\ & \text{such that} \quad \nu \in \mathbb{R}, \\ & w^{\mathcal{B}} - \tfrac{1}{N(1-\alpha)} \sum_{j \in \mathcal{J}} \boldsymbol{r}^{(j)\,T} \boldsymbol{x} \ \le \nu \qquad (\ \mathcal{J} \subset \{1, \ldots, N\},\ |\mathcal{J}| = N(1-\alpha)\ ). \end{aligned} \tag{11}$$

Proof. Given a set $\hat{\mathcal{J}} \subset \{1, \ldots, N\}$, $|\hat{\mathcal{J}}| = N(1-\alpha)$ that represents a cut in (11), let

$$\hat{\pi}_j := \begin{cases} 1/N & \text{if } j \in \hat{\mathcal{J}}, \\ 0 & \text{otherwise,} \end{cases} \qquad (1 \le j \le N).$$

This is a feasible solution of (9: $\boldsymbol{p} = \frac{\mathbf{1}}{N}$), hence $l(\boldsymbol{x}) \leq \mathcal{C}(\boldsymbol{x}) \quad (\boldsymbol{x} \in \mathbb{R}^n)$ holds with the linear function $l(\boldsymbol{x})$ constructed after (10).

Conversely, given $\hat{\boldsymbol{x}} \in \mathbb{R}^n$, let us construct a linear support function $l(\boldsymbol{x})$ of $\mathcal{C}(\boldsymbol{x})$ at $\hat{\boldsymbol{x}}$. As we saw above, it can be done according to (10) where $(\hat{\pi}_1, \ldots, \hat{\pi}_N)$ is an optimal solution of the problem (9: $\boldsymbol{p} = \frac{\mathbf{1}}{N}$, $\boldsymbol{x} = \hat{\boldsymbol{x}}$).

We saw in Section 1 that an optimal solution $(\hat{\pi}_1, \ldots, \hat{\pi}_N)$ can be found in the special form:

$$\text{either} \quad \hat{\pi}_j = 1/N \quad \text{or} \quad \hat{\pi}_j = 0 \qquad (1 \leq j \leq N).$$

The number of the positive components is $N(1-\alpha)$. With this optimal solution, let

$$\hat{\mathcal{J}} := \{\ j \ \mid 1 \leq j \leq N,\ \hat{\pi}_j = 1/N\ \}.$$

The corresponding cut is valid in (11). □

Corollary 3 *The following polyhedral representation is equivalent to problem (7):*

$$\begin{gathered} \min\ \boldsymbol{c}^T\boldsymbol{x} + \nu \\ \text{such that} \\ \nu \in \mathbb{R}, \quad \boldsymbol{x} \in X, \\ \tfrac{1}{N(1-\alpha)} \sum_{j \in \mathcal{J}} \left(w^{\mathcal{B}} - \boldsymbol{r}^{(j)T}\boldsymbol{x} \right) \leq \nu \qquad (\ \mathcal{J} \subset \{1,\ldots,N\},\ |\mathcal{J}| = N(1-\alpha)\ ). \end{gathered} \tag{12}$$

(In the last constraint, we used $w^{\mathcal{B}} = \frac{1}{N(1-\alpha)} \sum_{j \in \mathcal{J}} w^{\mathcal{B}}$ that is a consequence of $|\mathcal{J}| = N(1-\alpha)$.)

Let us consider now the polyhedral representation (8) proposed by Künzi-Bay and Mayer. In order to adapt it to the supposed case of approximate distribution, let us substitute $p_j = \frac{1}{N} \quad (j = 1, \ldots, N)$. Moreover, let us introduce a new variable $\nu := z + \frac{1}{1-\alpha} v$ instead of v. We get:

$$\begin{gathered} \min\ \boldsymbol{c}^T\boldsymbol{x} + \nu \\ \text{such that} \\ z, \nu \in \mathbb{R}, \quad \boldsymbol{x} \in X, \\ \tfrac{1}{N(1-\alpha)} \sum_{j \in \mathcal{J}} \left(w^{\mathcal{B}} - \boldsymbol{r}^{(j)T}\boldsymbol{x} \right) - \left(\tfrac{|\mathcal{J}|}{N(1-\alpha)} - 1 \right) z \leq \nu \qquad (\ \mathcal{J} \subset \{1,\ldots,N\}\ ). \end{gathered} \tag{13}$$

There is a strong similarity between the representations (13) and (12). In (13) the cuts are generated in the space $\mathbb{R}^{n+2} \ni (\boldsymbol{x}, z, \nu)$. In (12) the cuts are generated in the space

$\mathbb{R}^{n+1} \ni (\boldsymbol{x}, \nu)$. Equivalence of the two representations mean that the feasible region of (12) is a projection of the feasible region of (13).

Künzi-Bay and Mayer implemented their special decomposition method. They solved several CVaR-minimization test problems with their experimental solver called CVaRMin. They also solved the test problems with general-purpose LP solvers. These solvers were used to solve the LP-equivalents of the relevant two-stage stochastic programming problems. (The LP-equivalent form contains the constraints of the first-stage problem, and $2N$ additional constraints. These are *individual* constraints for each scenario, as opposed to the *aggregate* cuts that appear in the representation (8).) Additionally, they also solved the test problems in two-stage recourse forms, by employing an implementation of the Regularized Decomposition Method of Ruszczyński (1986).

In their problems the discrete distributions were samples taken from normal and log-normal distributions. Sample sizes varied between 500 and 20′000. (Some of the problems were portfolio optimization problems, others were randomly generated. In many cases they solved batteries of problems, built using different samples taken from the same continuous distributions.) Their experimental results show the clear superiority of the solver CVaRMin. For the largest test problems, CVaRMin was by at least one order of magnitude faster then either of the other solvers involved. The problems were solved with 8 to 10 accurate digits in the optimum. The number of the cuts that CVaRMin generated in course of the solution, varied between 24 and 106.

No special cutting-plane or bundle-type method has yet been implemented on the basis of the polyhedral representation (12) and of the formulas (9 - 10). But we are convinced of the effectiveness of such a method, on account of the simplicity of the representation (12).

For the decomposition of two-stage models, however, we need the more detailed representation (8) proposed by Künzi-Bay and Mayer:

2.1 A parametric version of the polyhedral representation (8)

Let us use a parameter ζ instead of the decision variable z. Moreover, let us introduce a new parameter ω to represent initial wealth. (Up to this point we assumed a fixed initial wealth that was implicitly represented in the feasible set X.)

$$\begin{aligned} \mathcal{D}(\omega,\zeta) := \quad & \min\ \boldsymbol{c}^T\boldsymbol{x} + \tfrac{1}{1-\alpha}v \\ & \text{such that} \\ & \boldsymbol{x} \in X, \quad v \in \mathbb{R}, \\ & \mathbf{1}^T\boldsymbol{x} = \omega, \\ & \sum_{j\in\mathcal{J}} p_j \left(w^{\mathcal{B}} - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} - \zeta \right) \le v \qquad (\ \mathcal{J} \subset \{1,\ldots,N\}\ ). \end{aligned} \tag{14}$$

The cutting plane method proposed by Künzi-Bay and Mayer can be easily adapted to the above problem:

0. Initialize.
Set the stopping tolerance ϵ.
The initial master problem contains the constraints $\boldsymbol{x} \in X,\ \mathbf{1}^T\boldsymbol{x} = \omega$;
and the cut belonging to the set $\mathcal{J}_0 := \{1, \dots, N\}$.
Set the iteration counter $\iota := 1$.

1. Solve master problem.
Let $(\boldsymbol{x}^\star, v^\star)$ be an optimal solution of the current cutting-plane model problem. Let

$$\mathcal{J}_\iota := \left\{ j \;\middle|\; w^{\mathcal{B}} - \boldsymbol{r}^{(j)\,T}\boldsymbol{x}^\star - \zeta > 0 \right\}.$$

2. Check for optimality
If

$$\sum_{j\in\mathcal{J}_\iota} p_j \left(w^{\mathcal{B}} - \boldsymbol{r}^{(j)\,T}\boldsymbol{x}^\star - \zeta \right) - v^\star \le (1-\alpha)\epsilon,$$

then $\boldsymbol{x}^\star$ is an ϵ-optimal solution; stop.

3. Append cut
Append the cut belonging to the set $\mathcal{J}_\iota$.
Increment ι, and repeat from *step 1.*

Alternatively, the problem can be solved by a bundle-type method.

Remark 4 *In the original method, Künzi-Bay and Mayer used a relative stopping tolerance. For the decomposition method we are going to describe in Section 3.2, we need an absolute stopping tolerance.*

However, our inexact-cut approach will never require solutions of irrationally high accuracy.

Suppose that after κ iterations, the cutting-plane model problem looks like this:

$$\begin{aligned} \mathcal{D}_\kappa(\omega,\zeta) := \quad & \min\ \boldsymbol{c}^T\boldsymbol{x} + \tfrac{1}{1-\alpha} v \\ & \text{such that} \\ & \boldsymbol{x} \in X, \quad v \in \mathbb{R}, \\ & \mathbf{1}^T\boldsymbol{x} = \omega, \\ & \textstyle\sum_{j\in\mathcal{J}_\iota} p_j \left(w^{\mathcal{B}} - \boldsymbol{r}^{(j)\,T}\boldsymbol{x} - \zeta \right) \le v \qquad (\iota = 0, \dots, \kappa). \end{aligned} \tag{15}$$

Given the parameter values $\hat{\omega}$ and $\hat{\zeta}$, let us solve the linear programming problem (15: $\omega = \hat{\omega}, \zeta = \hat{\zeta}$). A linear support function $L_\kappa(\omega, \zeta)$ of $\mathcal{D}_\kappa(\omega, \zeta)$ at $(\hat{\omega}, \hat{\zeta})$ can be constructed using the optimal dual variables. The function $\mathcal{D}_\kappa(\omega, \zeta)$ is obviously a polyhedral convex lower approximation of the convex function $\mathcal{D}(\omega, \zeta)$. Hence we have $L_\kappa(\omega, \zeta) \leq \mathcal{D}_\kappa(\omega, \zeta) \leq \mathcal{D}(\omega, \zeta)$.

Given $\hat{\omega}$ and $\hat{\zeta}$, suppose the solution actually terminated with the cutting-plane model problem (15: $\omega = \hat{\omega}, \zeta = \hat{\zeta}$). It means that $\mathcal{D}(\hat{\omega}, \hat{\zeta}) - \mathcal{D}_\kappa(\hat{\omega}, \hat{\zeta}) \leq \epsilon$ holds. Then the linear support function $L_\kappa(\omega, \zeta)$ constructed from the optimal dual variables of the final approximate problem, is also an ϵ-support function of $\mathcal{D}(\omega, \zeta)$ at $(\hat{\omega}, \hat{\zeta})$.

To summarize the above discussion: Given $\hat{\omega}, \hat{\zeta}$, and a tolerance ϵ, an ϵ-support function of $\mathcal{D}(\omega, \zeta)$ at $(\hat{\omega}, \hat{\zeta})$ can be constructed through the approximate solution of the problem (14: $\omega = \hat{\omega}, \zeta = \hat{\zeta}$).

3 A two-stage prototype model

Suppose that we wish to plan investments for two time periods. The investment can be restructured at the beginning of each period. We are going to use the following notation:

w_0 represents initial capital. This is a given parameter.

$\boldsymbol{x}_1 = (x_{11}, \ldots, x_{i1}, \ldots, x_{n1})^T$ represents the portfolio selected at the beginning of the first time period. This is a decision vector. Feasible portfolios are represented by $\boldsymbol{x}_1 \in X, \quad \mathbf{1}^T \boldsymbol{x}_1 = w_0$.

$\boldsymbol{r}_1 = (r_{11}, \ldots, r_{i1}, \ldots, r_{n1})^T$ denotes return occurring in course of the first time period. This is a random vector having a known distribution.

$w_1 := \boldsymbol{r}_1{}^T \boldsymbol{x}_1$ denotes the wealth at the end of the first time period.

$\boldsymbol{x}_2 = (x_{12}, \ldots, x_{i2}, \ldots, x_{n2})^T$ represents the portfolio selected at the beginning of the second time period. This is a decision vector. Feasible portfolios are represented by $\boldsymbol{x}_2 \in X, \quad \mathbf{1}^T \boldsymbol{x}_2 = w_1$.

$\boldsymbol{r}_2 = (r_{12}, \ldots, r_{i2}, \ldots, r_{n2})^T$ denotes return occurring in course of the second time period. This is a random vector. Assume that the joint distribution of the random vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ is known.

$w_2 := \boldsymbol{r}_2{}^T \boldsymbol{x}_2$ denotes the wealth at the end of the second time period.

Suppose that the set $X \subset \mathbb{R}^n$ representing feasible portfolios, is a convex bounded polyhedron.

Description of the random parameters. Assume there are N realizations of the random vector $\boldsymbol{r}_1$, namely

$$\boldsymbol{r}_1^{(1)}, \ldots, \boldsymbol{r}_1^{(j)}, \ldots, \boldsymbol{r}_1^{(N)} \quad \text{occurring with probabilities} \quad p_1^{(1)}, \ldots, p_1^{(j)}, \ldots, p_1^{(N)}, \quad \text{respectively.}$$

Assume, moreover, that the random vector $\boldsymbol{r}_2$ given that $\boldsymbol{r}_1 = \boldsymbol{r}_1^{(j)}$; has a known discrete distribution for each $1 \leq j \leq N$. Assume there are M_j realizations of this random vector, namely

$$\boldsymbol{r}_2^{(j1)}, \ldots, \boldsymbol{r}_2^{(jk)}, \ldots, \boldsymbol{r}_2^{(jM_j)} \quad \text{occurring with probabilities} \quad p_2^{(j1)}, \ldots, p_2^{(jk)}, \ldots, p_2^{(jM_j)}, \quad \text{respectively.}$$

Scenarios can identified by the elements of the set

$$\mathcal{S} := \{ \, (j,k) \mid 1 \leq j \leq N, \ 1 \leq k \leq M_j \, \},$$

where (j,k) means that the first- and second-period returns $\boldsymbol{r}_1^{(j)}$ and $\boldsymbol{r}_2^{(jk)}$ realize. The probability of this event will be denoted by

$$p_{(jk)} := p_1^{(j)} p_2^{(jk)} \qquad ((j,k) \in \mathcal{S}).$$

Realizations of the random wealth at the end of the first period will be denoted by

$$w_1^{(j)} := \boldsymbol{r}_1^{(j)^T} \boldsymbol{x}_1 \qquad (j = 1, \ldots, N).$$

Realizations of the random wealth at the end of the second period will be denoted by

$$w_2^{(jk)} := \boldsymbol{r}_2^{(jk)^T} \boldsymbol{x}_2 \qquad ((j,k) \in \mathcal{S}).$$

Problem formulation. Suppose we wish to find a trade-off between the expectation $\mathrm{E}(w_2)$ of the final wealth, and the conditional value-at-risk $\mathrm{CVaR}(w_2)$. This aim is expressed by the objective

$$\min \ -\lambda \mathrm{E}(w_2) + \mathrm{CVaR}(w_2), \tag{16}$$

where the risk tolerance parameter λ has been determined by the decision maker. Suppose that the benchmark $w^{\mathcal{B}}$ for the final wealth has also been determined by the decision maker.

The objective of minimizing end-of-horizon CVaR had been used earlier by Topaloglou (2004). He developed elaborate financial models, even in multi-stage frameworks. He solved one-stage and two-stage problems with commercially available, general-purpose modeling systems and optimization solvers. From an algorithmic point of view, his two-stage models fit the above described prototype. Hence the special decomposition framework and method we are going to describe can also be applied to his financial problems. (Topaloglou worked with scenario sets whose cardinality was in the magnitude of 10′000. We trust that our special decomposition method will enable the solution of problems with larger scenario sets.)

Using the polyhedral representation (8) of Künzi-Bay and Mayer, the prototype problem can be formulated as follows:

$$
\begin{aligned}
\min \quad & -\lambda \sum_{(j,k)\in\mathcal{S}} p_{(jk)} w_2^{(jk)} \quad + z \; + \tfrac{1}{1-\alpha} v \\
& \text{such that} \\
& z, v \in \mathrm{I\!R}, \quad \boldsymbol{x}_1 \in X, \\
& w_1^{(j)} \in \mathrm{I\!R}, \quad \boldsymbol{x}_2^{(j)} \in X, \qquad (j = 1, \ldots, N), \\
& w_2^{(jk)} \in \mathrm{I\!R} \qquad ((j,k) \in \mathcal{S}), \\
& \mathbf{1}^T \boldsymbol{x}_1 = w_0, \\
& \mathbf{1}^T \boldsymbol{x}_2^{(j)} \; = \; w_1^{(j)} \; = \; \boldsymbol{r}_1^{(j)T} \boldsymbol{x}_1 \qquad (j = 1, \ldots, N), \\
& w_2^{(jk)} \; = \; \boldsymbol{r}_2^{(jk)T} \boldsymbol{x}_2^{(j)} \qquad ((j,k) \in \mathcal{S}), \\
& \sum_{(j,k)\in\mathcal{L}} p_{(jk)} \left(w^{\mathcal{B}} - w_2^{(jk)} - z \right) \; \leq v \qquad (\, \mathcal{L} \subset \mathcal{S} \,) \, .
\end{aligned}
\tag{17}
$$

This is linear programming problem that may grow to a gigantic size depending on the cardinality of the scenario set $\mathcal{S}$.

3.1 Formulation of the problem in two-stage form

The first-stage problem will be:

$$\min \quad z + \sum_{j=1}^{N} p_1^{(j)}\, \mathcal{D}^{(j)}\left(w_1^{(j)},\, z\right)$$

$$\text{such that}$$

$$\begin{aligned} & z \in \mathrm{I\!R}, \quad \boldsymbol{x}_1 \in X, \\ & w_1^{(j)} \in \mathrm{I\!R} \quad (j = 1, \ldots, N), \\ & \mathbf{1}^T \boldsymbol{x}_1 = w_0, \\ & w_1^{(j)} = \boldsymbol{r}_1^{(j)^T} \boldsymbol{x}_1 \quad (j = 1, \ldots, N), \end{aligned} \tag{18}$$

where the functions $\mathcal{D}^{(j)} : \mathrm{I\!R}^2 \to \mathrm{I\!R} \quad (j = 1, \ldots, N)$ are defined by the second-stage problem:

$$\begin{aligned} \mathcal{D}^{(j)}(\omega, \zeta) := \quad & \min \quad -\lambda \sum_{k=1}^{M_j} p_2^{(jk)} w_2^{(k)} + \tfrac{1}{1-\alpha} v \\ & \text{such that} \\ & \boldsymbol{x}_2 \in X, \quad v \in \mathrm{I\!R}, \\ & w_2^{(k)} \in \mathrm{I\!R} \qquad (k = 1, \ldots, M_j), \\ & \mathbf{1}^T \boldsymbol{x}_2 = \omega, \\ & w_2^{(k)} = \boldsymbol{r}_2^{(jk)^T} \boldsymbol{x}_2 \qquad (k = 1, \ldots, M_j), \\ & \sum_{k \in \mathcal{K}} p_2^{(jk)} \left(w^{\mathcal{B}} - w_2^{(k)} - \zeta\right) \le v \qquad (\, \mathcal{K} \subset \{1, \ldots, M_j\}\,). \end{aligned} \tag{19}$$

Theorem 5 *The two-stage problem (18 - 19) is equivalent to the polyhedral representation problem (17).*

Proof. Let us construct the equivalent linear programming form (ELPF) of the two-stage problem (18 - 19) in the usual manner. It means defining a separate set of second-stage variables for each $j = 1, \ldots, N$. Let

$$\boldsymbol{x}_2^{(j)} \in X, \quad v^{(j)} \in \mathbb{R},$$

$$w_2^{(jk)} \in \mathbb{R} \qquad (k = 1, \ldots, M_j)$$

denote the set of the second-stage variables corresponding to j. There is a straightforward matching between the variables of the ELPF problem and those of the polyhedral representation problem (17). There is but one non-trivial case: variable v of the polyhedral representation problem will correspond to the weighted sum $\sum_{j=1}^{N} p_1^{(j)} v^{(j)}$ of the ELPF-variables $v^{(j)}$.

The objective function of the ELPF problem is

$$z + \sum_{j=1}^{N} p_1^{(j)} \left\{ -\lambda \sum_{k=1}^{M_j} p_2^{(jk)} w_2^{(jk)} + \frac{1}{1-\alpha} v^{(j)} \right\} = -\lambda \sum_{j=1}^{N} \sum_{k=1}^{M_j} p_1^{(j)} p_2^{(jk)} w_2^{(jk)} + z + \frac{1}{1-\alpha} \sum_{j=1}^{N} p_1^{(j)} v^{(j)},$$

and this is obviously equivalent to the objective function of the polyhedral representation problem (17).

The cuts in the ELPF problem are

$$\sum_{k \in \mathcal{K}^{(j)}} p_2^{(jk)} \left(w^{\mathcal{B}} - w_2^{(jk)} - z \right) \le v^{(j)} \qquad \left(\mathcal{K}^{(j)} \subset \{1, \ldots, M_j\}, \quad j = 1, \ldots, N \right).$$

Let us select a subset $\mathcal{K}^{(j)} \subset \{1, \ldots, M_j\}$ for each $j = 1, \ldots, N$, and aggregate the corresponding cuts with the weights $p_1^{(j)}$. (In case of an empty set $\mathcal{K}^{(j)} = \emptyset$, the corresponding cut is $0 \le v^{(j)}$.) We get

$$\sum_{j=1}^{N} p_1^{(j)} \sum_{k \in \mathcal{K}^{(j)}} p_2^{(jk)} \left(w^{\mathcal{B}} - w_2^{(jk)} - z \right) \le \sum_{j=1}^{N} p_1^{(j)} v^{(j)}. \tag{20}$$

Let

$$\mathcal{L} := \bigcup_{j=1}^{N} \left\{ (j,k) \in \mathcal{S} \;\middle|\; k \in \mathcal{K}^{(j)} \right\}.$$

With this $\mathcal{L}$, the aggregate ELPF-cut (20) takes the form

$$\sum_{(j,k) \in \mathcal{L}} p_{(jk)} \left(w^{\mathcal{B}} - w_2^{(jk)} - z \right) \le v,$$

that is a cut in the polyhedral representation problem (17).

In order to show that a single ELPF-cut corresponding to the set $\hat{\mathcal{K}}^{(\hat{\jmath})}$ is also valid in the polyhedral representation problem; let us select $\mathcal{K}^{(j)} := \emptyset$ for $j \ne \hat{\jmath}$ in the aggregation.

Conversely, given a cut in the polyhedral representation problem (17), i.e., a subset $\mathcal{L} \subset \mathcal{S}$, we can construct the sets

$$\mathcal{K}^{(j)} := \{ \, k \mid (j,k) \in \mathcal{L} \, \} \qquad (j = 1, \ldots, N).$$

In the manner of (20), let us aggregate the ELPF-cuts corresponding to the above sets. This results just the cut corresponding to $\mathcal{L}$ in the polyhedral representation problem (17). □

3.2 A solution method for the two-stage problem

Following the idea of Künzi-Bay and Mayer, the above two-stage problem could be solved as a special three-stage stochastic programming problem. We propose an alternative method.

The second-stage problem (19) has the form of the parametric polyhedral representation problem (14). (Indeed, let us substitute in (14)

$$\boldsymbol{r} := \left(\boldsymbol{r}_2 \middle| \boldsymbol{r}_1 = \boldsymbol{r}_1^{(j)} \right), \quad \boldsymbol{c} := -\lambda \mathrm{E} \left(\boldsymbol{r}_2 \middle| \boldsymbol{r}_1 = \boldsymbol{r}_1^{(j)} \right) = -\lambda \sum_{k=1}^{M_j} p_2^{(jk)} \boldsymbol{r}_2^{(jk)}$$

to obtain (19).)

Having fixed the values $\hat{z}$, $\hat{w}_1^{(j)}$ $(j = 1, \ldots, N)$ of the first-stage variables, and given a tolerance ϵ, we can construct ϵ-support functions $L^{(j)}(\omega, \zeta)$ of $\mathcal{D}^{(j)}(\omega, \zeta)$ at the point $\left(\omega = \hat{w}_1^{(j)}, \, \zeta = \hat{z}\right)$, by the method described in Section 2.1. Then the linear function $\zeta + \sum_{j=1}^{N} p_1^{(j)} L^{(j)}(\omega, \zeta)$ will be an ϵ-support function of the objective function of the first-stage problem (18), at the point $\left(\omega = \hat{w}_1^{(j)}, \, \zeta = \hat{z}\right)$.

Using the above described procedure as an *oracle* that provides information of the objective function, the first-stage problem (18) can be effectively solved by the Inexact Level Method. This is an inexact version of the Level Method of Lemaréchal, Nemirovskii, and Nesterov (1995). This inexact version was proposed by Fábián (2000).

The Inexact Level Method is a bundle-type method that successively builds a cutting-plane model of the objective function. On oracle is used to provide objective function data, in the form of ϵ-support functions. The accuracy tolerance ϵ of the cuts is gradually decreased as the optimum is approached. (At each step, we have an estimate of how closely the optimum has been approached. The successive cut is generated with an accuracy tolerance derived from that estimate.)

The above described procedure is a form of the Level Decomposition Method specialized for the two-stage stochastic problem (18 - 19). Level Decomposition was proposed by Fábián (2001) for the approximate solution of two-stage stochastic programming problems. (In the general framework, recourse function evaluations of appropriate accuracy are provided by an approximation scheme closely integrated with the inexact convex optimizer. The framework enables effective application of either bounding methods for the approximation of the distribution, or interior-point methods for the approximate solution of the recourse problems.)

4 An improvement on the two-stage prototype model

Jobst and Zenios (2001) describe an experiment: they solved one-stage portfolio selection problems with time periods of one year. The objective was the minimization of a risk measure. As risk measure, they used MAD in one problem, and CVaR in another problem. They then simulated the returns of the optimal portfolios, at the time points 3, 6, 9, and 12 months after the beginning of the time period.

With the MAD-optimal portfolio, they found that catastrophic losses are probable after the first 3 months.

But with the CVaR-optimal portfolio, worst losses were limited to around 10 % of the total wealth throughout. However, an interesting phenomenon occurred: the experimental distribution of the 9-month losses had a tail much heavier than the experimental distribution of the end-of-period losses had.

From the above phenomenon we conclude that the two-stage prototype model needs a correction: Beside restricting final risk, we also need restricting the risk at the end of the first period. We may penalize end-of-first-period risk $\mathrm{CVaR}(w_1)$ by including a new term in the objective function (16). The augmented objective will be:

$$\min\ \gamma\, \mathrm{CVaR}(w_1) - \lambda \mathrm{E}(w_2) + \mathrm{CVaR}(w_2), \tag{21}$$

where the risk trade-off parameter $\gamma > 0$ needs to be determined by the decision maker.

To include the new term into the prototype problem (17), we can use either the polyhedral representation (8) of Künzi-Bay and Mayer, or the simpler representation (12). The polyhedral cuts are then inherited by the first-stage problem (18).

This problem can also be effectively solved by the Level Decomposition Method.

If the decision maker can not determine the value of the parameters γ and λ, then we can help him by building an approximation of the efficient frontier. In the present case it is a 3-dimensional convex surface. Points from this surface can be found by minimizing the objective function (21) with different settings of the parameters γ and λ.

Künzi-Bay and Mayer proposed a warm-start procedure for the approximation of the efficient frontier in case of the one-stage problem (7). This procedure can be adapted to the 3-dimensional case.

Taking into account the running times of the special solver of Künzi-Bay and Mayer in the case of one-stage problems, we trust that two-stage problems of usable sizes can also be solved in reasonable time.

References

[1] FÁBIÁN, C.I. (2000). Bundle-type methods for inexact data. *Central European Journal of Operations Research* **8** (special issue, T. Csendes and T. Rapcsák, eds.), 35-55.

[2] Fábián, C.I. (2001). Adapting an approximate Level Method to the two-stage stochastic programming problem. *Stochastic Programming E-Print Series* 12-2001.

[3] Jobst, N.J. and S. Zenios (2001). The tail that wags the dog: Integrating credit risk in asset portfolios. *Working Paper* 01-05. HERMES, Center of Excellence on Computational Finance and Economics, University of Cyprus.

[4] Künzi-Bay, A. and J. Mayer (2005). Computational aspects of minimizing conditional value-at-risk. *Working Paper* No. 211. FINRISK, Swiss National Centre of Competence in Research. Financial Valuation and Risk Management.

[5] Klein Haneveld, W.K. and M.H. Van der Vlerk (2002). Integrated chance constraints: Reduced forms and an algorithm. *SOM Research Report* 02A33, University of Groningen.

[6] Lemaréchal, C., A. Nemirovskii, and Yu. Nesterov. (1995). New variants of bundle methods. *Mathematical Programming* **69**, 111-147.

[7] Prékopa, A. (2003). Probabilistic programming. In: *Stochastic Programming, Handbooks in Operations Research and Management Science* **10** (A. Ruszczyński and A. Shapiro, eds.), Elsevier, Amsterdam, 267-351.

[8] Rockafellar, R.T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of Risk* **2**, 21-41.

[9] Rockafellar, R.T. and S. Uryasev (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance* **26**, 1443-1471.

[10] Ruszczyński, A. (1986). A regularized decomposition method for minimizing the sum of polyhedral functions. *Mathematical Programming* **35**, 309-333.

[11] Topaloglou, N. (2004). A stochastic programming framework for international portfolio management. *Ph.D. Dissertation*. School of Economics and Management, University of Cyprus.