Bounds for nested law invariant coherent risk measures

Linwei Xin* **Alexander Shapiro†**

Abstract. With every law invariant coherent risk measure is associated its conditional analogue. In this paper we discuss lower and upper bounds for the corresponding nested (composite) formulations of law invariant coherent risk measures. In particular, we consider the Average Value-at-Risk and comonotonic risk measures.

Key Words: Risk measures, law invariant, coherent, Average Value-at-Risk, comonotonic

*School of Industrial & Systems Engineering, Georgia Institute of Technology, 765 Ferst Drive, Atlanta, GA 30332, e-mail: lwxin@gatech.edu.

†School of Industrial & Systems Engineering, Georgia Institute of Technology, 765 Ferst Drive, Atlanta, GA 30332, e-mail: ashapiro@isye.gatech.edu. This research was partly supported by the NSF award DMS-0914785 and ONR award N000140811104.

1 Introduction

Let $\Xi \subset \mathbb{R}^d$ be a closed set equipped with its Borel sigma algebra $\mathcal{B}$ and a probability measure P. We can view an element $\xi \in \Xi$ as an elementary event of the probability space $(\Xi, \mathcal{B}, P)$ or as a random vector whose probability distribution P is supported on the set Ξ. Unless stated otherwise all probabilistic statements and notation will be made with respect to the reference probability measure P. For $p \in [1, \infty)$ consider the space $\mathcal{Z} := L_p(\Xi, \mathcal{B}, P)$ of measurable functions $Z : \Xi \to \mathbb{R}$ such that $\int_\Xi |Z|^p dP$ is finite[1]. We can view an element $Z \in \mathcal{Z}$ as a random variable having finite p-th order moment. To a real valued function $\rho : \mathcal{Z} \to \mathbb{R}$ we refer as[2] a *risk measure*. An important example of risk measures is the Average Value-at-Risk measure

$$\mathsf{AV@R}_\alpha(Z) := \inf_{t \in \mathbb{R}} \left\{ t + \alpha^{-1} \mathbb{E}[Z - t]_+ \right\}, \tag{1.1}$$

where $\alpha \in (0, 1]$ and $\mathcal{Z} := L_1(\Xi, \mathcal{B}, P)$. This risk measure is also known under the names Expected Shortfall, Expected Tail Loss and Conditional Value-at-Risk. It could be noted that $\mathsf{AV@R}_1(\cdot) = \mathbb{E}(\cdot)$ and $\mathsf{AV@R}_0(\cdot) := \lim_{\alpha \downarrow 0} \mathsf{AV@R}_\alpha(Z) = \operatorname{ess\,sup}(Z)$.

The $\rho = \mathsf{AV@R}_\alpha$ risk measure is an example of coherent risk measures in the sense of Artzner et al [1], and is law invariant in the sense that it is a function of the probability distribution of Z, i.e., if two random variables $Z, Z' \in \mathcal{Z}$ have the same probability distribution, then $\rho(Z) = \rho(Z')$. With every law invariant risk measure ρ can be associated its conditional analogue. That is, let Y be a random vector. Then conditional risk measure, denoted $\rho(Z|Y)$ or $\rho_{|Y}(Z)$, is defined as the risk measure ρ of the conditional distribution of Z given Y. For example, the conditional analogue of $\mathsf{AV@R}_\alpha$ can be written as

$$\mathsf{AV@R}_{\alpha|Y}(Z) := \inf_{t \in \mathbb{R}} \left\{ t + \alpha^{-1} \mathbb{E}_{|Y}[Z - t]_+ \right\}, \tag{1.2}$$

where $\mathbb{E}_{|Y}$ denotes the corresponding conditional expectation.

We consider the following construction. Let vector $\xi \in \Xi$ be partitioned as $\xi = (\xi_1, ..., \xi_T)$ with $\xi_t \in \mathbb{R}^{d_t}$ and $d_1 + ... + d_T = d$. Note that we can view each ξ_t as a random vector with the corresponding marginal distribution derived from the probability distribution of the random vector ξ. For $\mathcal{Z} := L_p(\Xi, \mathcal{B}, P)$ consider a law invariant coherent risk measure $\rho : \mathcal{Z} \to \mathbb{R}$ and the corresponding composite (nested) risk measure

$$\bar{\rho}^T(Z) := \rho \left(\rho_{|\xi_2} \left(\cdots \rho_{|\xi_2, ..., \xi_T}(Z) \right) \cdots \right), \tag{1.3}$$

where $Z = Z(\xi_1, ..., \xi_T) \in \mathcal{Z}$. It is also possible to define such nested risk measures in an abstract form (cf., [4],[7]). An advantage of the present construction is that it seems to be more intuitive and well suited for dealing with multistage stochastic programming problems where $\xi_1, ..., \xi_T$ is viewed as the corresponding data process.

The risk measure $\rho(\cdot)$ is not the same as its nested counterpart $\bar{\rho}^T(\cdot)$. In this paper we discuss various bounds and inequalities between law invariant coherent risk measures and their nested counterparts.

---

[1] More precisely $L_p(\Xi, \mathcal{B}, P)$ is formed by classes of functions which may differ from each other on sets of P-measure zero.

[2] For historical reasons we use here the terminology of "risk measures". It should not be confused with the concept of probability measure.

2 Dual representations

Every (real valued) coherent risk measure $\rho : \mathcal{Z} \to \mathbb{R}$ is continuous in the norm topology of $\mathcal{Z} = L_p(\Xi, \mathcal{B}, P)$ (cf., [6]), and has the following dual representation

$$\rho(Z) = \sup_{\zeta \in \mathfrak{A}} \mathbb{E}[Z\zeta]. \tag{2.1}$$

Here $\mathfrak{A}$ is a convex weakly* compact subset of the dual space $\mathcal{Z}^* = L_q(\Xi, \mathcal{B}, P)$ formed by probability density functions, i.e., for any $\zeta \in \mathfrak{A}$ it holds that $\zeta(\cdot)$ is nonnegative valued and $\mathbb{E}[\zeta] = 1$. Note that since $\mathfrak{A}$ is weakly* compact, the maximum in the right hand side of (2.1) is attained. For $\rho := \mathsf{AV@R}_\alpha$ the corresponding set $\mathfrak{A}$ is

$$\mathfrak{A} = \left\{\zeta \in \mathcal{Z}^* : 0 \preceq \zeta \preceq \alpha^{-1},\ \mathbb{E}[\zeta] = 1\right\}, \tag{2.2}$$

where $\mathcal{Z}^* = L_\infty(\Xi, \mathcal{B}, P)$ and the notation $\zeta \succeq 0$ means that ζ is nonnegative w.p.1.

The representation (2.1) can be also written in the form $\rho(Z) = \sup_{Q \in \mathfrak{Q}} \mathbb{E}_Q[Z]$, where $\mathfrak{Q} := \{Q : dQ = \zeta dP,\ \zeta \in \mathfrak{A}\}$, i.e., $\mathfrak{Q}$ is the set of probability measures Q absolutely continuous with respect to P and such that $dQ/dP \in \mathfrak{A}$. We also consider conditional risk measures of the form

$$\varrho_{|Y}(Z) := \sup_{Q \in \mathfrak{Q}} \mathbb{E}_{Q|Y}[Z], \tag{2.3}$$

where $\xi = (X, Y)$ is a partition of the random vector ξ and $\mathbb{E}_{Q|Y}$ denotes the conditional on Y expectation with respect to the probability distribution Q of ξ.

Note that $\varrho_{|Y}$ is not the same as the conditional risk measure $\rho_{|Y}$. In order to see the difference let us note that $\varrho_{|Y}$ can be written as

$$\varrho_{|Y}(Z) = \sup_{\zeta \in \mathfrak{A}} \mathbb{E}_{|Y}[Z(X, Y)\zeta(X, Y)]. \tag{2.4}$$

On the other hand the dual representation of the conditional risk measure $\rho_{|Y}$ can be written as follows. Given $Y = y$ consider function $Z_y(\cdot) := Z(\cdot, y)$. We can view Z_y as a random variable whose (conditional) distribution is supported on the set $\Xi_y := \{x : (x, y) \in \Xi\}$. For the considered law invariant coherent risk measure ρ we have the corresponding dual representation (2.1) of $\rho(Z_y)$ with the associated dual set $\mathfrak{A}_y$ of density functions $\zeta_1 : \Xi_y \to \mathbb{R}$. That is

$$\rho_{|Y}(Z) = \sup_{\zeta_1 \in \mathfrak{A}_Y} \mathbb{E}_{|Y}[Z(X, Y)\zeta_1(X)]. \tag{2.5}$$

Example 1 Suppose that random vectors X and Y are independent, i.e., $P = P_1 \times P_2$ with P_1 and P_2 being respective probability distributions of X and Y supported on Ξ_1 and Ξ_2, respectively, and $\Xi = \Xi_1 \times \Xi_2$. Then the set $\mathfrak{A}_y = \mathfrak{A}_1$ in (2.5) does not depend on $y \in \Xi_2$. Compared with (2.4) we have that the maximum in (2.4) is taken with respect to a larger set. It follows that $\rho_{|Y}(Z) \leq \varrho_{|Y}(Z)$. Let, further, $\rho := \mathsf{AV@R}_\alpha$ with $\alpha \in (0, 1]$. Then

$$\varrho_{|Y}(Z) = \sup\left\{\mathbb{E}_{|Y}[Z(X, Y)\zeta(X, Y)] : 0 \preceq \zeta(X, Y) \preceq \alpha^{-1},\ \mathbb{E}[\zeta(X, Y)] = 1\right\}. \tag{2.6}$$

If the measure P_2 is nonatomic, then we can choose $\zeta = \zeta(X, Y)$ in the right hand side of (2.6) such that $\mathbb{E}[\zeta] = 1$, and $\zeta(X, y) = 1$ for a given $y \in \Xi_2$ and all $X \in \Xi_1$. For $Y = y$ and $Z \succeq 0$ the maximum in (2.6) is attained at that ζ and hence $\varrho_{|Y}(Z) = \alpha^{-1}\mathbb{E}_{|Y}[Z]$. Of course, it follows directly from (2.2) that for $Z \succeq 0$ and $\alpha \in (0, 1]$,

$$\mathsf{AV@R}_{\alpha|Y}(Z) \preceq \alpha^{-1}\mathbb{E}_{|Y}[Z]. \tag{2.7}$$

We also consider nested risk measures of the form

$$\bar{\varrho}^T(Z) := \rho\left(\varrho_{|\xi_2}\left(\cdots \varrho_{|\xi_2,\ldots,\xi_T}(Z)\right)\cdots\right). \tag{2.8}$$

It is not difficult to show (cf., [9]) that for any $Z = Z(\xi_1, ..., \xi_T) \in \mathcal{Z}$ the following inequality holds

$$\rho(Z) \le \bar{\varrho}^T(Z). \tag{2.9}$$

Suppose now that the random vectors $\xi_1, ..., \xi_T$ are independent, i.e., $P = P_1 \times \cdots \times P_T$ and $\Xi = \Xi_1 \times \cdots \Xi_T$. As it was argued in Example 1, in that case the inequality $\bar{\rho}^T(Z) \le \bar{\varrho}^T(Z)$ holds for all $Z \in \mathcal{Z}$. Therefore, inequality (2.9) does not imply the inequality $\rho(Z) \le \bar{\rho}^T(Z)$. And indeed, for example, for $\rho := \mathsf{AV@R}_\alpha$ the inequality $\rho(Z) \le \bar{\rho}^T(Z)$ does not necessarily holds (e.g., [5, p.157]).

3 Upper bounds for the nested Average Value-at-Risk measure

In this section we discuss upper bounds for the nested Average Value-at-Risk measure. We assume in this section that $\mathcal{Z} := L_1(\Xi, \mathcal{B}, P)$. Let us start with the following inequality (cf., [5, Proposition 3.36(i)]).

Proposition 3.1 *Let $Z \in \mathcal{Z}$ be nonnegative valued, i.e., $Z \succeq 0$, and $\alpha \in (0, 1]$. Then*

$$\mathsf{AV@R}_\alpha\Big[\mathsf{AV@R}_{\alpha|\xi_2}\left[\cdots \mathsf{AV@R}_{\alpha|\xi_2,\ldots,\xi_T}[Z]\right]\,\cdots\Big] \le \alpha^{-(T-1)}\mathsf{AV@R}_\alpha\,[Z]\,. \tag{3.1}$$

Proof. By (2.7) we have

$$\mathsf{AV@R}_{\alpha|\xi_2,\ldots,\xi_T}[Z] \le \alpha^{-1}\mathbb{E}_{|\xi_2,\ldots,\xi_T}[Z].$$

We also have (e.g., [8, Corollary 6.30])

$$\mathsf{AV@R}_{\alpha|\xi_2,\ldots,\xi_{T-1}}\left[\mathbb{E}_{|\xi_2,\ldots,\xi_T}[Z]\right] \le \mathsf{AV@R}_{\alpha|\xi_2,\ldots,\xi_T}[Z].$$

The proof can be completed now by induction. ▮

It is tempting to conclude that (3.1) is true for all $Z \in \mathcal{Z}$ Unfortunately as the following example shows it does not hold for all $Z \in \mathcal{Z}$. This should be not surprising. Note that, for $\alpha \in (0, 1)$, the coefficient $\kappa = \alpha^{-(T-1)}$ is bigger than one and that for any coherent risk measure $\rho(Z+c) = \rho(Z)+c$ for any $c \in \mathbb{R}$. Therefore by replacing Z with $Z - c$ for sufficiently large c, the inequality (3.1) can be made wrong.

Example 2 Let $Z_1 \sim N(0, 1)$ and $Z_2 \sim N(0, 1)$ be two independent random variables having standard normal distribution, and let $Z = Z_1 + Z_2$. We have then that $Z \sim N(0, 2)$, and Z has the same distribution as $\sqrt{2}\, Z_1$. Thus

$$\mathsf{AV@R}_\alpha(Z) = \sqrt{2}\,\mathsf{AV@R}_\alpha(Z_1). \tag{3.2}$$

We also have that

$$\mathsf{AV@R}_\alpha(\mathsf{AV@R}_{\alpha|Z_1}(Z)) = \mathsf{AV@R}_\alpha(Z_1) + \mathsf{AV@R}_\alpha(Z_2) = 2\mathsf{AV@R}_\alpha(Z_1). \tag{3.3}$$

Now

$$\mathsf{AV@R}_\alpha(Z_1) = \frac{1}{\alpha}\int_{z_\alpha}^{+\infty} t f(t)dt, \tag{3.4}$$

where $z_\alpha = \Phi^{-1}(1-\alpha)$ and $f(\cdot)$ is the pdf of Z_1. For $\alpha \in (0, \frac{1}{2}]$ the following formula holds

$$\mathsf{AV@R}_\alpha(Z_1) = \frac{1}{\alpha\sqrt{2\pi}} \exp(-z_\alpha^2/2).$$

Anyway for any $\alpha \in (0,1)$ the integral in the right hand side of (3.4) is positive and hence $\mathsf{AV@R}_\alpha(Z_1) > 0$. Therefore the inequality

$$\mathsf{AV@R}_\alpha(\mathsf{AV@R}_{\alpha|Z_1}(Z)) \le \alpha^{-1}\mathsf{AV@R}_\alpha(Z) \tag{3.5}$$

holds iff $\alpha \le 1/\sqrt{2}$. It follows that for $\alpha \in (1/\sqrt{2}, 1)$ the inequality (3.5) does not hold.

For small $\alpha > 0$ the factor $\kappa = \alpha^{-(T-1)}$ grows quickly with increase of the number of stages T (the question of computing a smallest coefficient κ in the bound of the form (3.1) is discussed in the recent paper [2]). The following bound seems to be more accurate (see section 5).

Proposition 3.2 *For $Z \in \mathcal{Z}$ and $\alpha \in [0,1]$ it holds that*

$$\mathsf{AV@R}_\alpha\Big[\mathsf{AV@R}_{\alpha|\xi_2}\big[\cdots \mathsf{AV@R}_{\alpha|\xi_2,\dots,\xi_T}[Z]\big]\ \cdots\Big] \le \mathsf{AV@R}_{\alpha^T}\,[Z]\,. \tag{3.6}$$

Proof. Consider $\alpha, \beta \in (0,1]$ and partition $\xi = (X, Y)$. As it was pointed out, the maximum in the right hand side of (2.5) is attained, and hence

$$\mathsf{AV@R}_{\alpha|Y}(Z) = \mathbb{E}_{|Y}[Z(X,Y)\zeta_Y(X)] \tag{3.7}$$

for some $\zeta_Y \in \mathfrak{A}_Y$. Thus

$$\begin{aligned}
\mathsf{AV@R}_\beta(\mathsf{AV@R}_{\alpha|Y}(Z)) &= \sup\Big\{\mathbb{E}[\zeta_1(Y)\mathsf{AV@R}_{\alpha|Y}(Z)] : 0 \preceq \zeta_1 \preceq \beta^{-1},\ \mathbb{E}[\zeta_1] = 1\Big\} \\
&= \sup\Big\{\mathbb{E}\big[\zeta_1(Y)\mathbb{E}_{|Y}[Z(X,Y)\zeta_Y(X)]\big] : 0 \preceq \zeta_1 \preceq \beta^{-1},\ \mathbb{E}[\zeta_1] = 1\Big\} \\
&= \sup\Big\{\mathbb{E}\big[\mathbb{E}_{|Y}[Z(X,Y)\zeta_1(Y)\zeta_Y(X)]\big] : 0 \preceq \zeta_1 \preceq \beta^{-1},\ \mathbb{E}[\zeta_1] = 1\Big\} \\
&= \sup\Big\{\mathbb{E}\big[Z(X,Y)\zeta_1(Y)\zeta_Y(X)\big] : 0 \preceq \zeta_1 \preceq \beta^{-1},\ \mathbb{E}[\zeta_1] = 1\Big\}.
\end{aligned} \tag{3.8}$$

We also have that

$$\mathbb{E}[\zeta_1(Y)\zeta_Y(X)] = \mathbb{E}\left[\zeta_1(Y)\mathbb{E}_{|Y}[\zeta_Y(X)]\right] = 1$$

and $0 \preceq \zeta_1(Y)\zeta_Y(X) \preceq \beta^{-1}\alpha^{-1}$. It follows that the last maximum in (3.8) is less than or equal to

$$\sup\Big\{\mathbb{E}[Z(X,Y)\zeta(X,Y)] : 0 \preceq \zeta \preceq \alpha^{-1}\beta^{-1},\ \mathbb{E}[\zeta] = 1\Big\} = \mathsf{AV@R}_{\alpha\beta}(Z).$$

This proves the following inequality

$$\mathsf{AV@R}_\beta(\mathsf{AV@R}_{\alpha|Y}(Z)) \le \mathsf{AV@R}_{\alpha\beta}(Z). \tag{3.9}$$

Of course, for $T = 2$ the inequality (3.6) is a particular case of (3.9) with $\alpha = \beta$. For $\alpha \in (0,1]$ the proof can be completed now by induction. For $\alpha = 0$ the left and right hand sides of (3.6) are equal to each other. ■

4 Upper bounds for comonotonic risk measures

Recall that two random variables Z_1, Z_2 are said to be comonotonic if $(Z_1, Z_2) \stackrel{\mathcal{D}}{\sim} (F_1^{-1}(U), F_2^{-1}(U))$, where $F_1(\cdot)$ and $F_2(\cdot)$ are cumulative distribution functions of Z_1 and Z_2, respectively, U is a random variable uniformly distributed on the interval $[0,1]$ and the notation $\stackrel{\mathcal{D}}{\sim}$ means that two random vectors have the same probability distribution. It is said that a risk measure $\rho : \mathcal{Z} \to \mathbb{R}$ is comonotonic if for any two comonotonic random variables $Z_1, Z_2 \in \mathcal{Z}$ it follows that $\rho(Z_1 + Z_2) = \rho(Z_1) + \rho(Z_2)$. We have the following result due to Kusuoka [3].

Proposition 4.1 *Suppose that the probability space $(\Xi, \mathcal{B}, P)$ is nonatomic and let $\mathcal{Z} := L_p(\Xi, \mathcal{B}, P)$, $p \in [1, +\infty]$. Then a risk measure $\rho : \mathcal{Z} \to \mathbb{R}$ is law invariant, coherent and comonotonic iff there exists a probability measure μ on the interval $[0,1]$ such that*

$$\rho(Z) = \int_0^1 \mathsf{AV@R}_\alpha(Z) d\mu(\alpha), \text{ for all } Z \in \mathcal{Z}. \tag{4.1}$$

Recall that $\mathsf{AV@R}_0(Z) = \operatorname{ess\,sup}(Z)$. Therefore if measure μ, in representation (4.1), has a positive mass at zero, then in order for $\rho(\cdot)$ to be real valued we have to use space $\mathcal{Z} := L_\infty(\Xi, \mathcal{B}, P)$. For $p \in [1, +\infty)$ the measure μ is distributed on $(0, 1]$.

Proposition 4.2 *Let $\mathcal{Z} := L_p(\Xi, \mathcal{B}, P)$, $p \in [1, +\infty)$, and $\rho : \mathcal{Z} \to \mathbb{R}$ be a risk measure of the form (4.1). Then for $Z = Z(\xi_1, ..., \xi_T) \in \mathcal{Z}$ the following inequality holds*

$$\rho\left(\rho_{|\xi_2}\left(\cdots \rho_{|\xi_2,\dots,\xi_T}(Z)\right)\cdots\right) \le \int_0^1 \cdots \int_0^1 \mathsf{AV@R}_{\alpha_1\alpha_2\cdots\alpha_T}(Z) d\mu(\alpha_1) d\mu(\alpha_2) ... d\mu(\alpha_T). \tag{4.2}$$

Proof. Suppose for the moment that $T = 2$. By (4.1) we have

$$\rho(\rho_{|\xi_2}(Z)) = \int_0^1 \mathsf{AV@R}_\beta\left(\rho_{|\xi_2}(Z)\right) d\mu(\beta)$$

and

$$\rho_{|\xi_2}(Z) = \int_0^1 \mathsf{AV@R}_{\alpha|\xi_2}(Z) d\mu(\alpha).$$

Moreover, by convexity and positive homogeneity of $\mathsf{AV@R}_\beta$ we have the following inequality

$$\mathsf{AV@R}_\beta\left(\int_0^1 \mathsf{AV@R}_{\alpha|\xi_2}(Z) d\mu(\alpha)\right) \le \int_0^1 \mathsf{AV@R}_\beta\left(\mathsf{AV@R}_{\alpha|\xi_2}(Z)\right) d\mu(\alpha).$$

Using (3.9) we obtain

$$\begin{array}{rcl} \rho(\rho_{|\xi_2}(Z)) & \le & \int_0^1 \int_0^1 \mathsf{AV@R}_\beta\left(\mathsf{AV@R}_{\alpha|\xi_2}(Z)\right) d\mu(\alpha) d\mu(\beta) \\ & \le & \int_0^1 \int_0^1 \mathsf{AV@R}_{\alpha\beta}(Z) d\mu(\alpha) d\mu(\beta). \end{array}$$

This proves (4.2) for $T = 2$, it can be extended to $T > 2$ in a straightforward way. ∎

For example, let $\rho(Z) := \sum_{i=1}^m c_i \mathsf{AV@R}_{\alpha_i}(Z)$ for some $\alpha_i \in (0, 1]$ and $c_i > 0$, $i = 1, ..., m$, such that $\sum_{i=1}^m c_i = 1$. This risk measure has representation (4.1) with discrete measure $\mu = \sum_{i=1}^m c_i \Delta_{\alpha_i}$, where Δ_α denotes measure of mass one at point α. In that case the inequality (4.2) takes the form

$$\rho\left(\rho_{|\xi_2}\left(\cdots \rho_{|\xi_2,\dots,\xi_T}(Z)\right)\cdots\right) \le \int_0^1 \mathsf{AV@R}_\alpha(Z) d\eta(\alpha), \tag{4.3}$$

where $\eta = \left(\sum_{i=1}^m c_i \Delta_{\alpha_i}\right)^T$, i.e., $\left(\sum_{i=1}^m c_i \Delta_{\alpha_i}\right)^2 = \sum_{i=1}^m \sum_{j=1}^m c_i c_j \Delta_{\alpha_i \alpha_j}$, etc.

5 Discussion and examples

In order to see how tight is the bound (3.6) let us consider the following examples. Let $\mathcal{Z} := L_p(\Omega, \mathcal{F}, P)$, $p \in [1, +\infty)$, and $\rho : \mathcal{Z} \to \mathbb{R}$ be a law invariant coherent risk measure. Consider a sequence $Z_1, ..., Z_T \in \mathcal{Z}$ of identically distributed random variables, i.e., $Z_t \overset{\mathcal{D}}{\sim} Z$, $t = 1, ..., T$, and thus $\rho(Z_t) = \rho(Z)$, $t = 1, ..., T$. Let $\bar{Z}_T := T^{-1} \sum_{t=1}^T Z_t$. By convexity of ρ we have that

$$\rho(\bar{Z}_T) \le T^{-1} \sum_{t=1}^{T} \rho(Z_t) = \rho(Z). \tag{5.1}$$

Suppose that the random variables $Z_1, ..., Z_T$ are independent, i.e., this is an iid sequence. Then

$$\rho_{|Z_1,...,Z_{T-1}}(Z_1 + ... + Z_T) = Z_1 + ... + Z_{T-1} + \rho(Z_T),$$

and so on, and hence in that case

$$\rho\left(\rho_{|Z_1}\left(\cdots \rho_{|Z_1,...,Z_{T-1}}(\bar{Z}_T)\right)\cdots\right) = \rho(Z). \tag{5.2}$$

Let, moreover, $p = 2$, i.e., Z has finite second order moment. Then $\mathrm{Var}(\bar{Z}_T) = \mathrm{Var}(Z)/T$ tends to 0 as $T \to \infty$, and thus $\bar{Z}_T$ converges in L_2-norm to $\mathbb{E}[Z]$. By continuity of ρ it follows that $\rho(\bar{Z}_T)$ tends to $\mathbb{E}[Z]$. The equality $\rho(Z) = \mathbb{E}[Z]$ holds for all $Z \in \mathcal{Z}$ only if the set $\mathfrak{A}$, in the dual representation (2.1), consists of unique element $\zeta(\cdot) \equiv 1$, i.e., only if $\rho(\cdot) = \mathbb{E}(\cdot)$. On the other hand, if the random variables $Z_1, ..., Z_T$ are perfectly correlated, i.e., $Z_1 = ... = Z_T$, then $\rho(\bar{Z}_T) = \rho(Z)$ and the equality (5.2) holds as well.

Now let $\rho := \mathsf{AV@R}_\alpha$ and $Z_1, ..., Z_T$ be an iid sequence of uniformly distributed on [0,1] random variables. Then (see the Appendix) $\mathsf{AV@R}_\alpha(Z_1) = 1 - \alpha/2$ and

$$\mathsf{AV@R}_{\alpha^T}(\bar{Z}_T) \le 1 - \frac{(T!)^{1/T}\alpha}{T+1}. \tag{5.3}$$

By the inequality

$$\left(\frac{T+1}{e}\right)^T < T! < e\left(\frac{T+1}{e}\right)^{T+1},$$

where $e \approx 2.72$ is Euler's number, we have that the right hand side of (5.3) tends to $1 - \alpha/e$ as $T \to \infty$. It follows that

$$\limsup_{T\to\infty} \left\{ \mathsf{AV@R}_{\alpha^T}\left[\bar{Z}_T\right] - \mathsf{AV@R}_\alpha\left[\mathsf{AV@R}_{\alpha|Z_1}\left[\cdots \mathsf{AV@R}_{\alpha|Z_1,...,Z_{T-1}}\left[\bar{Z}_T\right]\right]\cdots\right]\right\} \le \frac{\alpha}{2} - \frac{\alpha}{e}. \tag{5.4}$$

As another example suppose that random vector $(Z_1, ..., Z_T)$ has a multivariate normal distribution with $Z_t \sim N(0,1)$, $t = 1, ..., T$, and equal correlations $\mathrm{Cov}(Z_{t_1}, Z_{t_2}) = r$ for $1 \le t_1 < t_2 \le T$. Note that the correlation coefficient r can vary in the range $r \in [-(T-1)^{-1}, 1]$. In that case $\bar{Z}_T$ has normal distribution with zero mean and standard deviation $\sqrt{\mathrm{Var}(\bar{Z}_T)} = A_T(r)$, where

$$A_T(r) := \sqrt{T^{-1} + (1 - T^{-1})r}.$$

Thus $\bar{Z}_T \overset{\mathcal{D}}{\sim} A_T(r)Z$, where $Z \sim N(0,1)$. It follows that

$$\rho(\bar{Z}_T) = A_T(r)\rho(Z). \tag{5.5}$$

It is also possible to show that (see the Appendix)

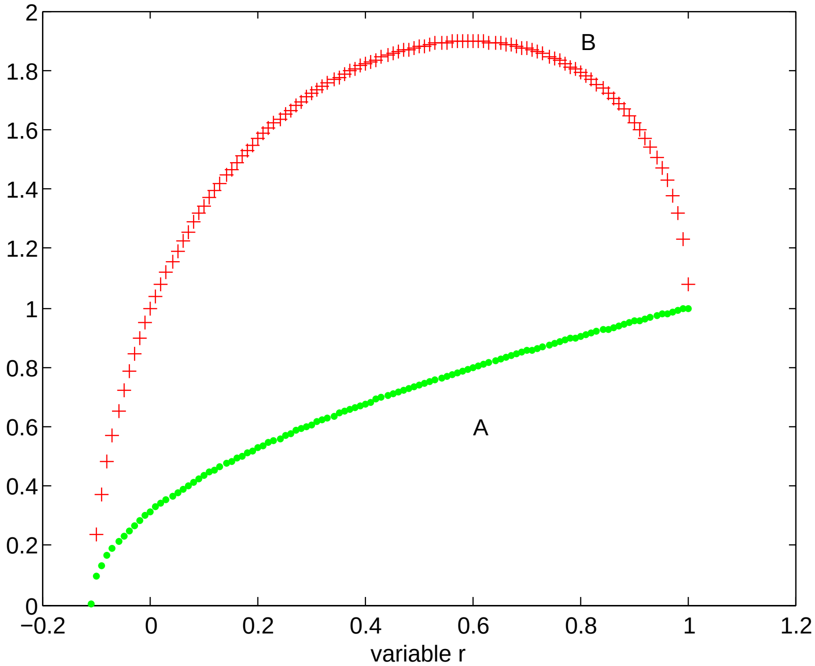
$$\rho\left(\rho_{|Z_1}\left(\cdots \rho_{|Z_1,\dots,Z_{T-1}}(\bar{Z}_T)\right)\cdots\right) = B_T(r)\rho(Z), \tag{5.6}$$

where

$$B_T(r) := \left[T^{-1} + (1-T^{-1})r\right]\left(1 + \sqrt{1-r}\sum_{k=1}^{T-1}\sqrt{\frac{1}{(1+kr)\left[1+(k-1)r\right]}}\right).$$

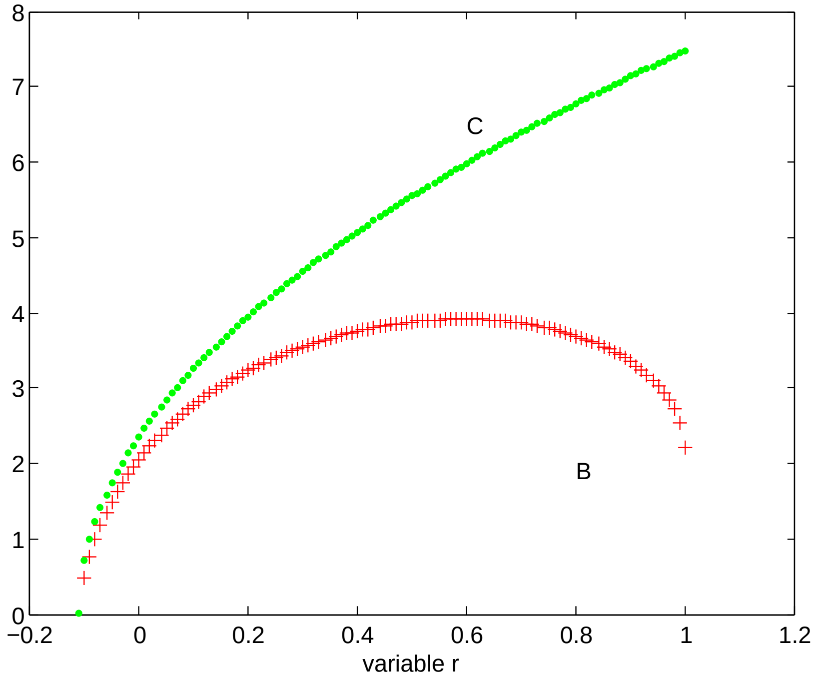
In Figure 1 values of $A_T(r)$ and $B_T(r)$ are plotted as a function of r for $T = 10$. Note that here $\rho(Z) \geq \mathbb{E}[Z] = 0$, and hence indeed the right hand side of (5.6) is bigger than the right hand side of (5.5).

Figure 1: Values of $A := A_T(r)$ and $B := B_T(r)$ $(T = 10)$



For $\rho :=$ AV@R$_\alpha$ we also have the upper bound (3.6). Values of AV@R$_{\alpha^T}(\bar{Z}_T)$ and the corresponding nested values, given by the left hand side of (3.6) for $Z = \bar{Z}_T$, as a function of r for $T = 10$, are plotted in Figure 2. It can be seen that for $r = 0$ (the iid case) the upper bound (3.6) is quite accurate here.

Figure 2: Values of $B := \mathsf{AV@R}_\alpha\left(\mathsf{AV@R}_{\alpha|Z_1}\left(\cdots\mathsf{AV@R}_{\alpha|Z_1,...,Z_{T-1}}(\bar Z_T)\right)\cdots\right)$ and $C := \mathsf{AV@R}_{\alpha^T}(\bar Z_T)$ $(T = 10,\ \alpha{=}0.05)$



6 Appendix

Let $\rho := \mathsf{AV@R}_\alpha$ and $Z_1, ..., Z_T$ be an iid sequence of uniformly distributed on [0,1] random variables. Then $\mathbb{E}[Z_1] = 1/2$ and[3]

$$\mathsf{AV@R}_\alpha(Z_1) = \alpha^{-1}\int_0^\alpha \mathsf{V@R}_\tau(Z_1)d\tau = \alpha^{-1}\int_0^\alpha (1-\tau)d\tau = 1 - \alpha/2.$$

Let $S := Z_1 + ... + Z_T$ and $\beta \in (0,1)$. In order to compute $\mathsf{AV@R}_\beta(S)$ we proceed as follows. We have that $\mathsf{Pr}(S \le x)$ is equal to 1 minus the volume of the set

$$\{z : z_1 + ... + z_T \ge x,\ 0 \le z_t \le 1,\ t = 1, ..., T\}.$$

Of course, this set is included in the set $\{z : z_1 + ... + z_T \ge x,\ z_t \le 1,\ t = 1, ..., T\}$, and for $x \ge T-1$ these two sets do coincide. Volume of the last set is $\frac{(T-x)^T}{T!}$. It follows that

$$\mathsf{Pr}(S \le x) \ge 1 - \frac{(T-x)^T}{T!},$$

and hence

$$\mathsf{V@R}_\tau(S) \le T - (T!)^{1/T}\tau^{1/T}.$$

Consequently

$$\mathsf{AV@R}_{\alpha^T}(\bar Z_T) \le T^{-1}\alpha^{-T}\int_0^{\alpha^T}\left[T - (T!)^{1/T}\tau^{1/T}\right]d\tau = 1 - \frac{(T!)^{1/T}\alpha}{T+1}.$$

[3] Recall that $\mathsf{V@R}_\tau(Z)$ of random variable Z is its left side $(1-\tau)$-quantile.

Next suppose that random vector $(Z_1, ..., Z_T)$ has a multivariate normal distribution with $\mathbb{E}[Z_t] = 0$, $\mathrm{Var}(Z_t) = 1$, $t = 1, ..., T$, and $\mathrm{Cov}(Z_{t_1}, Z_{t_2}) = r$ for $1 \le t_1 < t_2 \le T$. The conditional distribution of Z_T given $(Z_1, ..., Z_{T-1})$ is normal with mean $\Sigma_{21}\Sigma_{11}^{-1}(Z_1, ..., Z_{T-1})^{\mathsf{T}}$ and variance $\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$, where $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ is the covariance matrix of $(Z_1, ..., Z_{T-1}, Z_T)$, provided this matrix is nonsingular. It follows that

$$\rho(Z_T|Z_1, ..., Z_{T-1}) = \Sigma_{21}\Sigma_{11}^{-1}(Z_1, ..., Z_{T-1})^{\mathsf{T}} + \sqrt{\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}}\, \rho(Z). \tag{6.1}$$

Note that $\Sigma = (1-r)I_T + r\mathbf{1}_T\mathbf{1}_T^{\mathsf{T}}$, where I_T denotes $T \times T$ identity matrix and $\mathbf{1}_T$ denotes $T \times 1$ vector of ones, and

$$\Sigma^{-1} = (1-r)^{-1}\left[I_T - \frac{r}{1 + r(T-1)}\mathbf{1}_T\mathbf{1}_T^{\mathsf{T}}\right].$$

Similar formula holds for Σ_{11}^{-1}, with T replaced by $T - 1$. It follows that

$$\Sigma_{21}\Sigma_{11}^{-1}(Z_1, ..., Z_{T-1})^{\mathsf{T}} = \tfrac{r}{1+(T-2)r}(Z_1 + Z_2 + ... + Z_{T-1}),$$

$$\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} = \tfrac{(1-r)[1+(T-1)r]}{1+(T-2)r}.$$

Substituting this into (6.1) we obtain

$$\rho(Z_T|Z_1, ..., Z_{T-1}) = \frac{r(Z_1 + Z_2 + ... + Z_{T-1})}{1 + (T-2)r} + \sqrt{\frac{(1-r)[1 + (T-1)r]}{1 + (T-2)r}}\rho(Z).$$

Hence

$$\begin{aligned}
&\rho_{|Z_1,...,Z_{T-2}}\left[\rho_{|Z_1,...,Z_{T-1}}\left[Z_1 + Z_2 + ... + Z_T\right]\right] \\
=&\rho_{|Z_1,...,Z_{T-2}}\left[Z_1 + Z_2 + ... + Z_{T-1} + \rho_{|Z_1,...,Z_{T-1}}[Z_T]\right] \\
=&\rho_{|Z_1,...,Z_{T-2}}\left[Z_1 + Z_2 + ... + Z_{T-1} + \frac{r(Z_1 + Z_2 + ... + Z_{T-1})}{1 + (T-2)r}\right] + \sqrt{\frac{(1-r)\left[1 + (T-1)r\right]}{1 + (T-2)r}}\rho(Z) \\
=&\frac{1 + (T-1)r}{1 + (T-2)r}\rho_{|Z_1,...,Z_{T-2}}\left[Z_1 + Z_2 + ... + Z_{T-1}\right] + \sqrt{\frac{(1-r)\left[1 + (T-1)r\right]}{1 + (T-2)r}}\rho(Z).
\end{aligned}$$

Continuing by induction we obtain

$$\begin{aligned}
&\rho\left(\rho_{|Z_1}\left(\cdots \rho_{|Z_1,...,Z_{T-1}}(Z_1 + ... + Z_T)\right)\cdots\right) = \\
&[1 + (T-1)r]\left(1 + \sqrt{1-r}\sum_{k=1}^{T-1}\sqrt{\frac{1}{(1+kr)\left[1 + (k-1)r\right]}}\right)\rho(Z),
\end{aligned} \tag{6.2}$$

and hence (5.6) follows.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, Coherent measures of risk, *Mathematical Finance*, 9 (1999), 203–228.

[2] P. Huang, D.A. Iancu, M.Petrik and D. Subramanian, The Price of Dynamic Inconsistency for Distortion Risk Measures, `arXiv:1106.6102v1`.

[3] S. Kusuoka, On law-invariant coherent risk measures, in *Advances in Mathematical Economics*, Vol. 3, editors Kusuoka S. and Maruyama T., pp. 83-95, Springer, Tokyo, 2001.

[4] F. Riedel, Dynamic coherent risk measures. *Stochastic Processes and Their Applications*, 112 (2004), 185–200.

[5] G.Ch. Pflug and W. Römisch, *Modeling, Measuring and Managing Risk*, World Scientific Publishing Co., London, 2007.

[6] A. Ruszczyński and A. Shapiro, Optimization of convex risk functions. *Mathematics of Operations Research*, 31 (2006), 433–452.

[7] A. Ruszczyński and A. Shapiro, Conditional risk mappings. *Mathematics of Operations Research*, 31 (2006), 544–561.

[8] A. Shapiro, D. Dentcheva and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia, 2009.

[9] A. Shapiro, A, Minimax and risk averse multistage stochastic programming, *European Journal of Operational Research*, to appear.