8 Aggregate risk

8.1 Coherent and convex risk measures

- Consider a linear space $\mathcal{M} \subseteq \mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P})$ (a.s. finite rvs).
- Each $L \in \mathcal{M}$ (incl. constants) represents a loss over a fixed time horizon.
- A *risk measure* is a mapping $\varrho : \mathcal{M} \to \mathbb{R}$; $\varrho(L)$ gives the total amount of capital needed to back a position with loss L.
- $C \subseteq \mathcal{M}$ is *convex* if $(1-\gamma)x + \gamma y \in C$ for all $x, y \in C$, $0 < \gamma < 1$. C is a *convex cone* if, additionally, $\lambda x \in C$ when $x \in C$, $\lambda > 0$.
- Axioms for ϱ we consider are:

  Monotonicity: $L_1 \leq L_2 \Rightarrow \varrho(L_1) \leq \varrho(L_2)$.

  Translation invariance: $\varrho(L + m) = \varrho(L) + m$ for all $m \in \mathbb{R}$.

  Subadditivity: $\varrho(L_1 + L_2) \leq \varrho(L_1) + \varrho(L_2)$ for all $L_1, L_2 \in \mathcal{M}$.

  Positive homogeneity: $\varrho(\lambda L) = \lambda \varrho(L)$ for all $\lambda \geq 0$.

  Convexity: $\varrho(\gamma L_1 + (1-\gamma)L_2) \leq \gamma\varrho(L_1) + (1-\gamma)\varrho(L_2)$ for all $0 \leq \gamma \leq 1$, $L_1, L_2 \in \mathcal{M}$.

Definition 8.1 (Convex, coherent risk measures)

- A risk measure which satisfies monotonicity, translation invariance and convexity is called *convex*.
- A risk measure which satisfies monotonicity, translation invariance, subadditivity and positive homogeneity is called *coherent*.

A coherent risk measure is convex; the converse is not true, see below. On the other hand, for a positive-homogeneous risk measure, convexity and coherence are equivalent.

8.1.1 Risk measures and acceptance sets

Definition 8.2 (Acceptance set)

For a monotone and translation-invariant risk measure ϱ the *acceptance set of ϱ* is $A_\varrho = \{L \in \mathcal{M} : \varrho(L) \leq 0\}$ (so it contains the positions that are acceptable without any backing capital).

Proposition 8.3

Let ϱ be monotone and translation-invariant with associated A_ϱ. Then

1) $A_\varrho \neq \emptyset$ and A_ϱ satisfies

$$L \in A_\varrho \text{ and } \tilde{L} \leq L \Rightarrow \tilde{L} \in A_\varrho. \quad (34)$$

2) ϱ can be reconstructed from A_ϱ via

$$\varrho(L) = \inf\{m \in \mathbb{R} : L - m \in A_\varrho\}. \quad (35)$$

Proof. 1) is clear. For 2), note that $\inf\{m : L - m \in A_\varrho\} = \inf\{m : \varrho(L - m) \leq 0\} = \inf\{m : \varrho(L) - m \leq 0\}$ and this is equal to $\varrho(L)$. $\square$

Proposition 8.4

Suppose that A satisfies (34) and define

$$\varrho_A(L) = \inf\{m \in \mathbb{R} : L - m \in A\}. \tag{36}$$

Suppose $\varrho_A(L)$ is finite for all $L \in \mathcal{M}$. Then ϱ_A is monotone and translation-invariant on $\mathcal{M}$ and A_{ϱ_A} satisfies $A_{\varrho_A} \supseteq A$.

Proof. These properties of ϱ_A are easily checked. □

Example 8.5 (Value-at-risk)

For $\alpha \in (0,1)$, suppose we call $L \in \mathcal{M}$ *acceptable* if $\mathbb{P}(L > 0) \leq 1 - \alpha$. Then (36) is given by

$$\begin{aligned}\varrho_\alpha(L) &= \inf\{m \in \mathbb{R} : \mathbb{P}(L - m > 0) \leq 1 - \alpha\} \\ &= \inf\{m \in \mathbb{R} : \mathbb{P}(L \leq m) \geq \alpha\} = \mathrm{VaR}_\alpha(L).\end{aligned}$$

Proposition 8.6

1) Let ϱ be monotone and translation-invariant. Then

   1.1) ϱ is convex if and only if A_ϱ is convex.

   1.2) ϱ is coherent if and only if A_ϱ is a convex cone.

2) More generally, consider a set of acceptable positions A and the associated risk measure ϱ_A (whose acceptance set may be larger than A). If A is convex, so is ϱ_A; if A is a convex cone, then ϱ_A is coherent.

Example 8.7 (Risk measures based on loss functions)

Consider a strictly increasing and convex *loss function* $\ell : \mathbb{R} \to \mathbb{R}$ and some $c \in \mathbb{R}$. Assume that $\mathbb{E}(\ell(L))$ is finite for all $L \in \mathcal{M}$. Define an acceptance set by

$$A = \{L \in \mathcal{M} : \mathbb{E}\big(\ell(L)\big) \le \ell(c)\},$$

and the associated risk measure by

$$\varrho_A = \inf\{m \in \mathbb{R} : \mathbb{E}\big(\ell(L-m)\big) \le \ell(c)\}.$$

- ϱ_A is translation invariant and monotone by Proposition 8.4 since A satisfies (34).
- ϱ_A is convex by Proposition 8.6; to see this consider acceptable positions L_1 and L_2 and observe that the convexity of ℓ implies

$$\begin{aligned}\mathbb{E}(\ell(\gamma L_1 + (1-\gamma)L_2)) &\leq \mathbb{E}(\gamma\ell(L_1) + (1-\gamma)\ell(L_2)) \\ &\leq \gamma\ell(c) + (1-\gamma)\ell(c) = \ell(c),\end{aligned}$$

  where we have used that $\mathbb{E}(\ell(L_i)) \leq \ell(c)$ for acceptable positions. Hence $\gamma L_1 + (1-\gamma)L_2 \in A$, so A is convex.
- Example: $\ell(x) = \exp(\alpha x)$ for some $\alpha > 0$. Then

$$\begin{aligned}\varrho_{\alpha,c}(L) :&= \inf\{m : \mathbb{E}(e^{\alpha(L-m)}) \leq e^{\alpha c}\} = \inf\{m : \mathbb{E}(e^{\alpha L}) \leq e^{\alpha c + \alpha m}\} \\ &= \frac{1}{\alpha}\log(\mathbb{E}(e^{\alpha L})) - c.\end{aligned}$$

  Note that $\varrho_{\alpha,c}(0) = -c$, so $\varrho_{\alpha,c}$ cannot be coherent. For $c = 0$ and

$\lambda > 1$, the *entropic risk measure* $\varrho_{\alpha,0}$ satisfies

$$\varrho_{\alpha,0}(\lambda L) = \frac{1}{\alpha}\ln\{\mathbb{E}(e^{\alpha\lambda L})\} \geq \frac{1}{\alpha}\ln\{\mathbb{E}(e^{\alpha L})^{\lambda}\} = \lambda\varrho_{\alpha,0}(L),$$

where the inequality is strict if L is non-degenerate. This shows that $\varrho_{\alpha,0}$ is convex but not coherent. If L are insurance claims, $\varrho_{\alpha,0}$ is known as *exponential premium principle*.

Example 8.8 (Stress test or worst case risk measure)

Given *stress scenarios* $S \subseteq \Omega$, a *stress test risk measure* can be defined by

$$\varrho(L) = \sup\{L(\omega) : \omega \in S\},$$

that is, the worst loss on S. The associated acceptance set is

$$A_{\varrho} = \{L : L(\omega) \leq 0 \text{ for all } \omega \in S\}.$$

The choice of S is often guided by the underlying probability measure $\mathbb{P}$.

Example 8.9 (Generalized scenario risk measures)

Consider a set $\mathcal{Q}$ of probability measures on $(\Omega, \mathcal{F})$ and a *penalty function* $\gamma : \mathcal{Q} \to \mathbb{R}$ such that $\inf\{\gamma(\mathbb{Q}) : \mathbb{Q} \in \mathcal{Q}\} > -\infty$. Suppose $\sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}|L| < \infty$ for all $L \in \mathcal{M}$. The *generalized scenario risk measures* ϱ is defined by

$$\varrho(L) = \sup\{\mathbb{E}_{\mathbb{Q}}(L) - \gamma(\mathbb{Q}) : \mathbb{Q} \in \mathcal{Q}\}. \tag{37}$$

The corresponding acceptance set is given by

$$A_\varrho = \{L \in \mathcal{M} \colon \sup\{\mathbb{E}_{\mathbb{Q}}(L) - \gamma(\mathbb{Q}) : \mathbb{Q} \in \mathcal{Q}\} \le 0\}.$$

- A_ϱ is convex, and thus so is ϱ.
- Every convex risk measure can be represented as (37); see Theorem 8.10.
- If $\gamma(\cdot) \equiv 0$ on $\mathcal{Q}$, ϱ is positive homogeneous and therefore coherent.
- The stress test risk measure of Example 8.8 is a special case of (37) in which $\gamma \equiv 0$ and $\mathcal{Q}$ is the set of all Dirac measures $\delta_\omega(\cdot)$, $\omega \in S$, that is, $\delta_\omega(B) = I_B(\omega)$ for arbitrary measurable sets $B \subseteq \Omega$.

8.1.2 Dual representation of convex measures of risk

Theorem 8.10 (Dual representation for risk measures)
Suppose $|\Omega| = n < \infty$. Let $\mathcal{F} = \mathcal{P}(\Omega)$ (power set) and $\mathcal{M} := \{L : \Omega \to \mathbb{R}\}$. Then:

1) Every convex risk measure ϱ on $\mathcal{M}$ can be written in the form

$$\varrho(L) = \max\{\mathbb{E}_{\mathbb{Q}}(L) - \alpha_{\min}(\mathbb{Q}) : \mathbb{Q} \in \mathcal{S}^1(\Omega, \mathcal{F})\}, \tag{38}$$

where $\mathcal{S}^1(\Omega, \mathcal{F})$ denotes the set of all probability measures on Ω, and where the penalty function $\alpha_{\min}$ is given by $\alpha_{\min}(\mathbb{Q}) = \sup\{\mathbb{E}_{\mathbb{Q}}(L) : L \in A_\varrho\}$.

2) If ϱ is coherent, it has the representation

$$\varrho(L) = \max\{\mathbb{E}_{\mathbb{Q}}(L) \colon \mathbb{Q} \in \mathcal{Q}\}$$

for some set $\mathcal{Q} = \mathcal{Q}(\varrho) \subseteq \mathcal{S}^1(\Omega, \mathcal{F})$.

One can show that $\alpha_{\min}(\mathbb{Q}) = \sup_{L \in \mathcal{M}}\{\mathbb{E}_{\mathbb{Q}}(L) - \varrho(L)\}$.

8.1.3 Examples of dual representations

Proposition 8.11 (ES formulas)

For $\alpha \in (0,1)$,

1) $\mathrm{ES}_\alpha(L) = \dfrac{\mathbb{E}((L - F_L^{\leftarrow}(\alpha))_+)}{1-\alpha} + F_L^{\leftarrow}(\alpha)$;
2) $\mathrm{ES}_\alpha(L) = \dfrac{\mathbb{E}(LI_{\{L > F_L^{\leftarrow}(\alpha)\}}) + F_L^{\leftarrow}(\alpha)(1 - \alpha - \bar{F}_L(F_L^{\leftarrow}(\alpha)))}{1-\alpha}$.

Corollary 8.12 (ES formulas under continuous F_L)

Let F_L be continuous at $F_L^{\leftarrow}(\alpha)$. Then

1) $\mathrm{ES}_\alpha(L) = \frac{\mathbb{E}(LI_{\{L > F_L^{\leftarrow}(\alpha)\}})}{1-\alpha}$
2) $\mathrm{ES}_\alpha(L) = \mathbb{E}(L \,|\, L > F_L^{\leftarrow}(\alpha))$ (i.e. *conditional VaR (CVaR)*)

With dual representations one can give a proof for ES_α being subadditive; see the following result.

Theorem 8.13

For $\alpha \in [0,1)$, ES_α is coherent on $\mathcal{M} = \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. The dual representation is given by

$$\mathrm{ES}_\alpha(L) = \max\{\mathbb{E}^{\mathbb{Q}}(L) : \mathbb{Q} \in \mathcal{Q}_\alpha\}, \tag{39}$$

where $\mathcal{Q}_\alpha$ is the set of all probability measures on $(\Omega, \mathcal{F})$ that are absolutely continuous with respect to $\mathbb{P}$ and for which the measure-theoretic density $d\mathbb{Q}/d\mathbb{P}$ is bounded by $1/(1-\alpha)$.

8.2 Law-invariant coherent risk measures

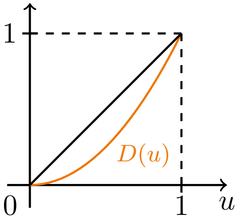
8.2.1 Distortion risk measures

Distortion risk measures are important coherent risk measures. We summarize important representations and investigate their properties.

Representations of distortion risk measures

Definition 8.14 (Distortion risk measure)

A *convex distortion function* D is a convex, increasing and absolutely continuous function on $[0,1]$ satisfying $D(0) = 0$ and $D(1) = 1$.



The *distortion risk measure* associated with D is defined by

$$\varrho(L) = \int_0^1 F_L^{\leftarrow}(u)\,\mathrm{d}D(u). \tag{40}$$

Note:

- A distortion risk measure is law-invariant (average of the L-quantiles).
- $D(u) = \int_0^u \phi(s)\,\mathrm{d}s$ for an increasing, positive function ϕ (the right-sided derivative of D), hence

$$\varrho(L) = \int_0^1 F_L^{\leftarrow}(u)\phi(u)\,\mathrm{d}u. \tag{41}$$

  A risk measure of this form is known as *spectral risk measure* and ϕ as *spectrum*.
- For $D_\alpha(u) = (1-\alpha)^{-1}(u-\alpha)^+$ one obtains expected shortfall. The spectrum is $\phi(u) = (1-\alpha)^{-1} I_{\{u \geq \alpha\}}$ (equal weight is placed on all quantiles beyond the α-quantile).

Lemma 8.15

The distortion risk measure ϱ associated with a convex distortion function D can be written in the form

$$\varrho(L) = \int_{\mathbb{R}} x \,\mathrm{d}D \circ F_L(x), \tag{42}$$

where $D \circ F_L(x) = D\big(F_L(x)\big)$.

Proof. $G(x) = D \circ F_L(x)$ has quantile function $G^{\leftarrow} = F_L^{\leftarrow} \circ D^{\leftarrow}$. Thus (42) can be written as

$$\int_{\mathbb{R}} x \,\mathrm{d}G(x) \underset{u=G(x)}{=} \int_0^1 G^{\leftarrow}(u)\,\mathrm{d}u = \int_0^1 F_L^{\leftarrow} \circ D^{\leftarrow}(u)\,\mathrm{d}u = \mathbb{E}(F_L^{\leftarrow} \circ D^{\leftarrow}(U)),$$

where $U \sim \mathrm{U}(0,1)$. Now introduce $V = D^{\leftarrow}(U) \sim D$ and note that

$$\int_{\mathbb{R}} x \,\mathrm{d}D \circ F_L(x) = \mathbb{E}(F_L^{\leftarrow}(V)) = \int_0^1 F_L^{\leftarrow}(v)\,\mathrm{d}D(v). \qquad \square$$

D distorts F_L. Since D is convex, $D(u) \leq u$, so $G = D \circ F_L$ puts more mass on high values of L than F_L.

Distortion risk measure can be represented as a weighted average of expected shortfall; see the appendix for a proof.

Proposition 8.16 (Distortion risk measures as weighted ES)
Let ϱ be a distortion risk measure associated with the convex distortion function D. Then, for a probability measure μ,

$$\varrho(L) = \int_0^1 \mathrm{ES}_\alpha(L)\,\mathrm{d}\mu(\alpha).$$

Properties of distortion risk measures

Definition 8.17 (Comonotone additivity)
A risk measure ϱ on a space of random variables $\mathcal{M}$ is said to be *comonotone additive* if $\varrho(L_1 + \cdots + L_d) = \varrho(L_1) + \cdots + \varrho(L_d)$ for comonotone $L_1, \ldots, L_d$.

- Quantile functions (so value-at-risk) are comonotone additive. Comonotone additivity of distortion risk measures then follows from (40).

- Distortion risk measures are coherent. Monotonicity, translation invariance and positive homogeneity are obvious. Subadditivity follows from Proposition 8.16 and subadditivity of ES_α (e.g., Theorem 8.13) by observing that

$$\begin{aligned}\varrho(L_1+L_2) &= \int_0^1 \mathrm{ES}_\alpha(L_1+L_2)\,\mathrm{d}\mu(\alpha)\\ &\le \int_0^1 \mathrm{ES}_\alpha(L_1)\,\mathrm{d}\mu(\alpha)+\int_0^1 \mathrm{ES}_\alpha(L_2)\,\mathrm{d}\mu(\alpha)\\ &= \varrho(L_1)+\varrho(L_2).\end{aligned}$$

- In summary, we have verified that distortion risk measures are law invariant, coherent and comonotone additive.
- It may also be shown that, on an atomless probability space (where there exists a continuous random variable), a law-invariant, coherent, comonotone-additive risk measure must be of the form (40) for some convex distortion function D.

- Parametric families of distortion risk measures can be based on convex distortion functions of the form

$$D_\alpha(u) = \Psi(\Psi^{-1}(u) + \ln(1-\alpha)), \quad 0 \le \alpha < 1,$$

where Ψ is a continuous df on $\mathbb{R}$; for $\Psi(u) = 1 - \exp(-u)$, $u \ge 0$, one obtains the distortion function for ES.

  - Such a family of convex distortion functions is strictly decreasing in α for fixed u.
  - $D_0(u) = u$ (corresponding to the risk measure $\varrho(L) = \mathbb{E}(L)$) and $\lim_{\alpha\to 1} D(u) = 1_{\{u=1\}}$.
  - For $\alpha_1 < \alpha_2$ and $0 < u < 1$ we have $D_{\alpha_1}(u) > D_{\alpha_2}(u)$, so that D_{α_2} distorts the original probability measure more than D_{α_1} and places more weight on outcomes in the tail.

8.2.2 The expectile risk measure

Definition 8.18 (Expectiles)
Let $L \in \mathcal{M} := L^1(\Omega, \mathcal{F}, \mathbb{P})$, so $\mathbb{E}|L| < \infty$. Then, for $\alpha \in (0,1)$, the *α-expectile $e_\alpha(L)$* is given by the unique solution y of

$$\alpha \mathbb{E}((L-y)^+) = (1-\alpha)\mathbb{E}((L-y)^-) \tag{43}$$

where $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$.

- Since $x^+ - x^- = x$, $e_{0.5}(L) = \mathbb{E}(L)$ as $\mathbb{E}(L-y)^- = \mathbb{E}(L-y)^+$ iff $\mathbb{E}((L-y)^+ - (L-y)^-) = 0$ iff $\mathbb{E}(L-y) = 0$.
- $\mathbb{E}(L^2) < \infty$, $e_\alpha(L)$ is the minimizer of

$$\min_{y \in \mathbb{R}} \mathbb{E}\left(S(y, L)\right) \tag{44}$$

for *scoring function $S(y, L)$*. This could be relevant for the out-of-sample testing of expectile-estimates (so-called *backtesting*). The scoring function that yields the expectile is

$$S^e_\alpha(y, L) = |1_{\{L\le y\}} - \alpha|(L-y)^2. \tag{45}$$

In fact we can compute that $\frac{d}{dy}\mathbb{E}\left(S^e_\alpha(y, L)\right)$ equals

$$\begin{aligned}
&\frac{d}{dy}\int_{-\infty}^{\infty} |1_{\{y\ge x\}} - \alpha|(y-x)^2\,\mathrm{d}F_L(x) \\
&= \frac{d}{dy}\int_{-\infty}^{y} (1-\alpha)(y-x)^2\,\mathrm{d}F_L(x) + \frac{d}{dy}\int_{y}^{\infty} \alpha(y-x)^2\,\mathrm{d}F_L(x) \\
&= 2(1-\alpha)\int_{-\infty}^{y} (y-x)\,\mathrm{d}F_L(x) + 2\alpha\int_{y}^{\infty} (y-x)\,\mathrm{d}F_L(x) \\
&= 2(1-\alpha)\mathbb{E}((L-y)^-) - 2\alpha\mathbb{E}((L-y)^+)
\end{aligned}$$

and setting this equal to zero yields the definition of an expectile.

- One can show that the α-quantile $F_L^{\leftarrow}(\alpha)$ is also a minimizer of the form (44); consider the scoring function $S^q_\alpha(y, L) = |1_{\{L\le y\}} - \alpha||L-y|$.

The following result shows uniqueness of the α-expectile and provides a helpful formula for computing expectiles of certain distributions; see the appendix for a proof.

Proposition 8.19

Let $\alpha \in (0,1)$ and L a rv such that $\mu := \mathbb{E}(L) < \infty$. Then $e_\alpha(L)$ may be written as $e_\alpha(L) = \tilde{F}_L^{-1}(\alpha)$ where

$$\tilde{F}_L(y) = \frac{yF_L(y) - \mu(y)}{2(yF_L(y) - \mu(y)) + \mu - y} \tag{46}$$

is a continuous df that is strictly increasing on its support and $\mu(y) := \int_{-\infty}^{y} x \,\mathrm{d}F_L(x)$.

Example 8.20 (Bernoulli)

Let $L \sim \text{Be}(p)$ be a Bernoulli-distributed loss. Then

$$F_L(y) = \begin{cases} 0, & y < 0 \\ 1-p, & 0 \le y < 1, \\ 1, & y \ge 1 \end{cases} \quad \mu(y) = \begin{cases} 0, & y < 1 \\ p, & y \ge 1 \end{cases}$$

from which it follows that $\tilde{F}_L(y) = \frac{y(1-p)}{y(1-2p)+p}$, $0 \le y \le 1$ and

$$e_\alpha(L) = \frac{\alpha p}{(1-\alpha) + p(2\alpha - 1)}.$$

Note that this can take any value in zero and one, whereas $\mathrm{VaR}_\alpha(L) \in \{0, 1\}$, $\alpha \in (0, 1]$.

Properties of expectiles

Proposition 8.21 (Coherence of expectile risk measures)
$\varrho = e_\alpha$ is a coherent risk measure on $\mathcal{M} = L^1(\Omega, \mathcal{F}, \mathbb{P})$ for $\alpha \geq 0.5$.

- See the appendix for a proof.
- Expectiles are not comonotone additive and thus are not distortion risk measures.
- If L_1 and L_2 are comonotonic and of the same type (so that $L_2 = kL_1 + m$ for some $m \in \mathbb{R}$ and $k > 0$) then we do have comonotone additivity (by translation invariance and positive homogeneity), but for comonotonic variables that are not of the same type one can find examples where $e_\alpha(L_1 + L_2) < e_\alpha(L_1) + e_\alpha(L_2)$ for $\alpha > 0.5$.

8.3 Risk measures for linear portfolios

We now consider linear portfolios in

$$\mathcal{M} = \{L : L = m + \boldsymbol{\lambda}'\boldsymbol{X},\ m \in \mathbb{R}, \boldsymbol{\lambda} \in \mathbb{R}^d\}, \tag{47}$$

for a fixed d-dimensional random vector $\boldsymbol{X}$.

- Many standard approaches to risk aggregation and capital allocation are based on the assumption that losses have a linear relationship to underlying risk factor changes.
- It is common to use linear approximations for losses due to market risks over short time horizons.

8.3.1 Coherent risk measures as stress tests

- Let $\varrho : \mathcal{M} \to \mathbb{R}$ be a positive-homogeneous risk measure. Define a *risk-measure function* $r_\varrho(\boldsymbol{\lambda}) = \varrho(\boldsymbol{\lambda}'\boldsymbol{X})$ (function of portfolio weights).

- If ϱ is translation-invariant, there is a one-to-one relationship between ϱ and r_ϱ given by

$$\varrho(m + \boldsymbol{\lambda}'\boldsymbol{X}) = m + r_\varrho(\boldsymbol{\lambda}).$$

Lemma 8.22 (Properties of r_ϱ)

Consider a translation-invariant risk measure $\varrho : \mathcal{M} \to \mathbb{R}$ with associated risk-measure function r_ϱ. Then

1) ϱ is a positive-homogeneous risk measure if and only if r_ϱ is a positive-homogeneous function, that is $r_\varrho(t\boldsymbol{\lambda}) = t r_\varrho(\boldsymbol{\lambda})$ for all $t > 0$, $\boldsymbol{\lambda} \in \mathbb{R}^d$.
2) Suppose that ϱ is positive-homogeneous. Then ϱ is subadditive if and only if r_ϱ is convex.

The main result of this section is that coherent risk measures for linear portfolios are stress tests as in Example 8.8 where the scenario set is

$$S_\varrho := \{\boldsymbol{x} \in \mathbb{R}^d \colon \boldsymbol{u}'\boldsymbol{x} \le r_\varrho(\boldsymbol{u}) \text{ for all } \boldsymbol{u} \in \mathbb{R}^d\}.$$

Proposition 8.23 (Coherent risk measures for linear portfolios as stress tests)
ϱ is a coherent risk measure on the set of linear portfolios $\mathcal{M}$ in (47) if and only if for every $L = m + \boldsymbol{\lambda}'\boldsymbol{X} \in \mathcal{M}$ we have the representation

$$\varrho(L) = m + r_\varrho(\boldsymbol{\lambda}) = \sup\{m + \boldsymbol{\lambda}'\boldsymbol{x} : \boldsymbol{x} \in S_\varrho\}. \tag{48}$$

- S_ϱ is an intersection of the half-spaces $H_u = \{\boldsymbol{x} \in \mathbb{R}^d \colon \boldsymbol{u}'\boldsymbol{x} \le r_\varrho(\boldsymbol{u})\}$, so that S_ϱ is a closed convex set. The precise form of S_ϱ depends on the df of $\boldsymbol{X}$ and on ϱ.
- If $\varrho = \mathrm{VaR}_\alpha$, S_ϱ has an interpretation as a *depth set*. Suppose that $\boldsymbol{u}'\boldsymbol{X}$ is continuously distributed for all $\boldsymbol{u} \in \mathbb{R}^d \setminus \{\boldsymbol{0}\}$. Then for $H_u = \{\boldsymbol{x} \in \mathbb{R}^d \colon \boldsymbol{u}'\boldsymbol{x} \le \mathrm{VaR}_\alpha(\boldsymbol{u}'\boldsymbol{X})\}$, $\mathbb{P}(\boldsymbol{u}'\boldsymbol{X} \in H_u) = \alpha$ so that S_{VaR_α} is the intersection of all half-spaces with probability α.

8.3.2 Elliptically distributed risk factors

Theorem 8.24 (Risk measurement for elliptical risk factors)

Let $\boldsymbol{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \psi)$ and ϱ be any positive-homogeneous, translation-invariant and law-invariant risk measure on $\mathcal{M}$. Then:

1) For any $L = m + \boldsymbol{\lambda}'\boldsymbol{X} \in \mathcal{M}$, $\varrho(L) = m + \boldsymbol{\lambda}'\boldsymbol{\mu} + \sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}}\varrho(Y_1)$ for $Y_1 \sim S_1(\psi)$.
2) If $\varrho(Y_1) \geq 0$, then ϱ is subadditive on $\mathcal{M}$ (e.g., VaR_α for $\alpha \geq 0.5$).
3) If $\mathbb{E}\boldsymbol{X}$ exists then, $\forall\, L = m + \boldsymbol{\lambda}'\boldsymbol{X} \in \mathcal{M}$ and $\rho_{ij} = \wp(\Sigma)_{ij} = P_{ij}$,
$$\varrho(L - \mathbb{E}L) = \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{d} \rho_{ij}\lambda_i\lambda_j\varrho(X_i - \mathbb{E}X_i)\varrho(X_j - \mathbb{E}X_j)}.$$
4) If $\mathrm{cov}(\boldsymbol{X})$ exists and $\varrho(Y_1) > 0$ then, for every $L \in \mathcal{M}$, $\varrho(L) = \mathbb{E}(L) + k_\varrho\sqrt{\mathrm{var}(L)}$ for some $k_\varrho > 0$ depending on ϱ.
5) If Σ^{-1} ex., $\varrho(Y_1) > 0$ then $S_\varrho = \{\boldsymbol{x} : (\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) \leq \varrho(Y_1)^2\}$.

Proof.

1) Let $\boldsymbol{Y} \sim S_k(\psi)$, $AA' = \Sigma$. $L = m + \boldsymbol{\lambda}'\boldsymbol{X} \stackrel{\text{d}}{=} m + \boldsymbol{\lambda}'\boldsymbol{\mu} + \boldsymbol{\lambda}' A\boldsymbol{Y}$. By Theorem 6.15 3), $L \stackrel{\text{d}}{=} m + \boldsymbol{\lambda}'\boldsymbol{\mu} + \|A'\boldsymbol{\lambda}\| Y_1$. Thus $\varrho(L) = m + \boldsymbol{\lambda}'\boldsymbol{\mu} + \|A'\boldsymbol{\lambda}\|\varrho(Y_1) = m + \boldsymbol{\lambda}'\boldsymbol{\mu} + \sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}}\varrho(Y_1)$.
2) Set $L_1 = m_1 + \boldsymbol{\lambda}_1'\boldsymbol{X}$ and $L_2 = m_2 + \boldsymbol{\lambda}_2'\boldsymbol{X}$. Subadditivity follows from 1) and $\|A'(\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2)\| \le \|A'\boldsymbol{\lambda}_1\| + \|A'\boldsymbol{\lambda}_2\|$ and $\varrho(Y_1) \ge 0$.
3) $\varrho(L - \mathbb{E}L) = \varrho(L) - \mathbb{E}(L) = \varrho(L) - (m + \boldsymbol{\lambda}'\boldsymbol{\mu}) = \sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}}\varrho(Y_1)$, so

$$\varrho(L - \mathbb{E}L) = \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{d} \rho_{ij}\lambda_i\lambda_j\sigma_i\sigma_j\varrho(Y_1)},$$

where $\sigma_j = \sqrt{\Sigma_{jj}}$ for $j \in \{1, \dots, d\}$. For $\boldsymbol{\lambda} = \boldsymbol{e}_j$, $\varrho(X_j - \mathbb{E}X_j) = \varrho(\boldsymbol{e}_j'\boldsymbol{X} - \mathbb{E}(\boldsymbol{e}_j'\boldsymbol{X})) = \sigma_j\varrho(Y_1)$, from which the result follows.
4) $\operatorname{cov}(\boldsymbol{X}) = c\Sigma$ for some $c > 0$. Since $\operatorname{var}(L) = \operatorname{var}(\boldsymbol{\lambda}'\boldsymbol{X}) = \boldsymbol{\lambda}'c\Sigma\boldsymbol{\lambda}$, 3) implies that $\varrho(L) = \mathbb{E}(L) + \sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}}\varrho(Y_1) = \mathbb{E}(L) + \sqrt{\operatorname{var}(L)}\varrho(Y_1)/\sqrt{c}$.

5) 2) implies that $r_\varrho(\boldsymbol{\lambda}) = \|A'\boldsymbol{\lambda}\|\varrho(Y_1) + \boldsymbol{\lambda}'\boldsymbol{\mu}$ so that S_ϱ is

$$\begin{aligned} S_\varrho &= \left\{\boldsymbol{x} \in \mathbb{R}^d \ : \ \boldsymbol{u}'\boldsymbol{x} \le \boldsymbol{u}'\boldsymbol{\mu} + \|A'\boldsymbol{u}\| \ \varrho(Y_1) \ \forall\, \boldsymbol{u} \in \mathbb{R}^d\right\} \\ &= \left\{\boldsymbol{x} \in \mathbb{R}^d \ : \ \boldsymbol{u}'AA^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \le \|A'\boldsymbol{u}\| \ \varrho(Y_1) \ \forall\, \boldsymbol{u} \in \mathbb{R}^d\right\} \\ &= \left\{\boldsymbol{x} \in \mathbb{R}^d \ : \ \boldsymbol{v}'\frac{A^{-1}(\boldsymbol{x} - \boldsymbol{\mu})}{\varrho(Y_1)} \le \|\boldsymbol{v}\| \ \forall\, \boldsymbol{v} \in \mathbb{R}^d\right\}, \end{aligned}$$

where the last line follows because $\mathbb{R}^d = \{A'\boldsymbol{u} : \boldsymbol{u} \in \mathbb{R}^d\}$. Since $\{\boldsymbol{y} \in \mathbb{R}^d : \boldsymbol{y}'\boldsymbol{y} \le 1\}$ can be written as $\{\boldsymbol{y} \in \mathbb{R}^d : \boldsymbol{v}'\boldsymbol{y} \le \|\boldsymbol{v}\| \ \forall\, \boldsymbol{v} \in \mathbb{R}^d\}$, we conclude that, for $\boldsymbol{x} \in S_\varrho$, the vectors $\boldsymbol{y} = A^{-1}(\boldsymbol{x} - \boldsymbol{\mu})/\varrho(Y_1)$ describe the unit ball and therefore

$$S_\varrho = \{\boldsymbol{x} \in \mathbb{R}^d : (\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \le \varrho(Y_1)^2\}. \qquad \square$$

- 2) gives a special case where VaR is subadditive and thus coherent. In particular, if $(L_1, \ldots, L_d)$ is jointly elliptical, VaR_α is subadditive for $\alpha \geq 0.5$.
- 3) provides a useful interpretation of risk measures on $\mathcal{M}$ in terms of the aggregation of stress tests.
- 4) is relevant to portfolio optimization. If we consider losses $L \in \mathcal{M}$ for which $\mathbb{E}(L)$ is fixed, the weights that minimize ϱ also minimize the variance. The portfolio minimizing ϱ is thus the same as the Markowitz variance-minimizing portfolio.
- 5) shows that the scenario sets in the stress test representation of coherent risk measures are ellipsoids when $\boldsymbol{X}$ is elliptical. Different law-invariant coherent risk measures simply lead to ellipsoids of differing radius $\varrho(Y_1)$. Scenario sets of ellipsoidal form are often used in practice and this result provides a justification for this practice in the case of linear portfolios of elliptical risk factors.

8.4 Risk aggregation

- A *risk aggregation rule* is a mapping

$$f(\mathrm{EC}_1, \dots, \mathrm{EC}_d) = \mathrm{EC}$$

which maps the individual capital amounts $\mathrm{EC}_1, \dots, \mathrm{EC}_d$ to the aggregate capital EC (economic capital). Examples are:

  - *Simple summation* $\mathrm{EC} = \mathrm{EC}_1 + \cdots + \mathrm{EC}_d$ (a special case of and upper bound for correlation adjusted summation)
  - *Correlation adjusted summation*

$$\mathrm{EC} = \sqrt{\sum_{i=1}^{d} \sum_{j=1}^{d} \rho_{ij} \, \mathrm{EC}_i \, \mathrm{EC}_j}, \qquad (49)$$

where $\rho_{ij} \in [0, 1]$ are parameters (referred to as *correlations*).

- Applying such rules *without* considering a multivariate model or risk measures is known as *rules-based aggregation*, otherwise, *principles-based aggregation*; we focus on the latter.
- In what follows we show that correlation adjusted summation is justified as a risk aggregation rule under various setups.

8.4.1 Aggregation based on loss distributions

- Suppose that the overall loss is $L = L_1 + \cdots + L_d$ where $L_1, \ldots, L_d$ are the losses arising from sub-units (e.g., business units, asset classes). Consider a translation-invariant ϱ and define

$$\varrho^{\text{mean}}(\cdot) = \varrho(\cdot - \mathbb{E}(\cdot)) = \varrho(\cdot) - \mathbb{E}(\cdot),$$

  that is, the capital required to cover unexpected losses.
- The capital requirements for the sub-units are

$$\mathrm{EC}_j = \varrho^{\text{mean}}(L_j), \quad j \in \{1, \ldots, d\},$$

and the aggregate capital should be

$$\mathrm{EC} = \varrho^{\text{mean}}(L).$$

- We require an aggregation rule f such that $f(\mathrm{EC}_1, \dots, \mathrm{EC}_d) = \mathrm{EC}$.
- If $\varrho(L) = \mathbb{E}(L) + k\,\mathrm{sd}(L)$, $k > 0$, and $\mathbb{E}(L^2) < \infty$ then

$$\mathrm{sd}(L) = \sqrt{\mathrm{var}(\mathbf{1}'\boldsymbol{L})} = \sqrt{\mathbf{1}'\,\mathrm{cov}(\boldsymbol{L})\mathbf{1}} = \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{d} \rho_{ij}\,\mathrm{sd}(L_i)\,\mathrm{sd}(L_j)},$$

  where $(\rho_{ij})_{i,j} = \mathrm{corr}(\boldsymbol{L})$, so correlation adjusted summation follows by noting that $\mathrm{sd}(L) = \varrho^{\text{mean}}(L)/k = \mathrm{EC}/k$ (and $\mathrm{sd}(L_j) = \mathrm{EC}_j/k$).
- If $L_j = m_j + \boldsymbol{\lambda}_j'\boldsymbol{X}$ for $\boldsymbol{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \psi)$ with existing $\mathrm{cov}(\boldsymbol{X})$, then this formula and Theorem 8.24 4) imply that correlation adjusted summation is justified for any positive-homogeneous, translation-invariant and law-invariant risk measure ϱ.
- As the following result shows, the assumption on $\mathrm{cov}(\boldsymbol{X})$ can be dropped.

Proposition 8.25 (Correlation adjusted sum. for linear portfolios) Let $\boldsymbol{X} \sim E_k(\boldsymbol{\mu}, \Sigma, \psi)$ with $\mathbb{E}(\boldsymbol{X}) = \boldsymbol{\mu}$. Let $\mathcal{M} = \{L : L = m + \boldsymbol{\lambda}'\boldsymbol{X},\ \boldsymbol{\lambda} \in \mathbb{R}^k,\ m \in \mathbb{R}\}$ and ϱ be a pos.-hom., translation- and law-invariant risk measure on $\mathcal{M}$. For $L_1, \dots, L_d \in \mathcal{M}$, let $\mathrm{EC}_j = \varrho^{\text{mean}}(L_j)$ and $\mathrm{EC} = \varrho^{\text{mean}}(L_1 + \dots + L_d)$. Then $\mathrm{EC}, \mathrm{EC}_1, \dots, \mathrm{EC}_d$ satisfy the correlation adjusted summation for $P = \wp(\tilde{\Sigma}) = (\rho_{ij})_{ij}$ and $\tilde{\Sigma}$ is the scale matrix of the (elliptical) $(L_1, \dots, L_d)$.

Proof. Let $L_j = m_j + \boldsymbol{\lambda}_j'\boldsymbol{X}$. By Theorem 8.24 1), $\mathrm{EC}_j = \varrho(L_j) - \mathbb{E}(L_j) = \sqrt{\boldsymbol{\lambda}_j'\Sigma\boldsymbol{\lambda}_j}\,\varrho(Y_1)$ where $Y_1 \sim S_1(\psi)$ and that

$$\mathrm{EC} = \sqrt{(\boldsymbol{\lambda}_1 + \dots + \boldsymbol{\lambda}_d)'\,\Sigma\,(\boldsymbol{\lambda}_1 + \dots + \boldsymbol{\lambda}_d)}\,\varrho(Y_1)$$

$$= \sqrt{\sum_{i=1}^d \sum_{j=1}^d \boldsymbol{\lambda}_i'\Sigma\boldsymbol{\lambda}_j\,\varrho(Y_1)^2} = \sqrt{\sum_{i=1}^d \sum_{j=1}^d \frac{\boldsymbol{\lambda}_i'\Sigma\boldsymbol{\lambda}_j}{\sqrt{(\boldsymbol{\lambda}_i'\Sigma\boldsymbol{\lambda}_i)(\boldsymbol{\lambda}_j'\Sigma\boldsymbol{\lambda}_j)}}\mathrm{EC}_i\,\mathrm{EC}_j}.$$

The scale matrix $\tilde{\Sigma}$ of $(L_1, \dots, L_d)$ is $\tilde{\Sigma} = \Lambda\Sigma\Lambda'$ where $\Lambda = (\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_d)'$. The corresponding $P = (\rho_{ij})_{ij}$ has elements $\boldsymbol{\lambda}_i'\Sigma\boldsymbol{\lambda}_j / \sqrt{(\boldsymbol{\lambda}_i'\Sigma\boldsymbol{\lambda}_i)(\boldsymbol{\lambda}_j'\Sigma\boldsymbol{\lambda}_j)}$ and thus

$$\mathrm{EC} = \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\boldsymbol{\lambda}_i'\Sigma\boldsymbol{\lambda}_j}{\sqrt{(\boldsymbol{\lambda}_i'\Sigma\boldsymbol{\lambda}_i)(\boldsymbol{\lambda}_j'\Sigma\boldsymbol{\lambda}_j)}} \mathrm{EC}_i \, \mathrm{EC}_j} = \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{d} \rho_{ij} \, \mathrm{EC}_i \, \mathrm{EC}_j}. \qquad \square$$

- Correlation adjusted summation can thus be justified under the mean-adjusted VaR or ES if $\boldsymbol{L}$ is elliptical.
- The correlations ρ_{ij} between $L_1, \dots, L_d$ are typically difficult to estimate (data is rather available for risk factors than losses). If they are chosen by *expert judgement*, there are compatibility requirements. If $(L_1, \dots, L_d)$ is non-elliptical, the limited range of attainable correlations for each pair (L_i, L_j) is also a relevant constraint; see Chapter 7.
- No obvious way to incorporate tail dependence between $L_1, \dots, L_d$.
- Simple summation only offers a conservative upper bound if ϱ is coherent.

8.4.2 Aggregation based on stressing risk factors

- Correlation adjusted summation is used in the aggregation of capital contributions $\mathrm{EC}_1, \dots, \mathrm{EC}_d$ computed by stressing individual risk factors (example: Standard formula approach to Solvency II).
- Let $\boldsymbol{x} = \boldsymbol{X}(\omega)$ be a scenario defined in terms of changes in risk factors and $L(\boldsymbol{x})$ the corresponding loss. Assume $L(\boldsymbol{x})$ is known and componentwise increasing.
- The d risk factors are stressed individually by amounts $k_1, \dots, k_d$. Capital contributions for each risk factor are computed by

$$\mathrm{EC}_j = L(k_j \boldsymbol{e}_j) - L(\mathbb{E}(X_j)\boldsymbol{e}_j)$$

  where $k_j > \mathbb{E}(X_j)$ so that $\mathrm{EC}_j > 0$ (interpreted as the loss incurred by stressing risk factor j by k_j relative to the impact of stressing it by its expected change); an example is $k_j = q_\alpha(X_j)$ for large α.
- The following justifies correlation adjusted summation as a risk aggregation rule if $k_j = \varrho(X_j)$ for elliptical $\boldsymbol{X}$ and $L(\boldsymbol{X}) = m + \boldsymbol{\lambda}'\boldsymbol{X}$.

Proposition 8.26 (Justification for correlation adjusted summation)
Let $\boldsymbol{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \psi)$ with $\mathbb{E}(\boldsymbol{X}) = \boldsymbol{\mu}$. Let $\mathcal{M}$ be the space of linear portfolios (47) and ϱ be a pos. hom., translation- and law-invariant risk measure on $\mathcal{M}$. Then, for any $L = L(\boldsymbol{X}) = m + \boldsymbol{\lambda}'\boldsymbol{X} \in \mathcal{M}$,

$$\mathrm{EC} = \varrho(L - \mathbb{E}(L)) = \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{d} \rho_{ij}\, \mathrm{EC}_i\, \mathrm{EC}_j},$$

where $\mathrm{EC}_j = L(\varrho(X_j)\boldsymbol{e}_j) - L(\mathbb{E}(X_j)\boldsymbol{e}_j)$ and $\rho_{ij} = \wp(\Sigma)_{i,j}$.

Proof. Note that $\mathrm{EC}_j = m + \lambda_j \varrho(X_j) - (m + \lambda_j \mathbb{E}X_j) = \lambda_j \varrho(X_j - \mathbb{E}X_j)$ and plug this into Theorem 8.24 3) to see that the claim holds. □

- Thus under linearity of the losses in jointly elliptical risk-factor changes, we can aggregate the effects of single-risk-factor stresses to an aggregate capital; this applies to VaR, ES or distortion risk measures. This idea underscores correlation adjusted summation in Solvency II.

- For market risk factors (returns on prices), the data may be available to estimate the ρ_{ij}s. For other risk factors (e.g. mortality and policy lapse rates in Solvency II), they are set by expert judgement (see issues mentioned earlier).

8.4.3 Modular versus fully integrated aggregation approaches

- The approaches of Sections 8.4.1 and 8.4.2 are *modular approaches*. In Sections 8.4.1 the *modules* (or *silos*) are business units or asset classes; in Section 8.4.2 they were individual risk factors; the former approach is more natural because losses are additive (and it is possible to remove risks from the enterprise by selling parts of the business).
- The aggregation approaches involved correlations and the correlation adjusted summation; however, correlations give only a partial description of dependence. It is natural to consider using copulas in aggregation.

- Consider simple summation and suppose we know/have estimated the marginal distributions $F_1, \dots, F_d$ for each of the modules (necessary for computing $\mathrm{EC}_j = \varrho(L_j) - \mathbb{E}(L_j)$). In the *margins-plus-copula approach*, we could attempt to choose a suitable copula C for $\boldsymbol{L} \sim F(\boldsymbol{x}) = C(F_1(x_1), \dots, F_d(x_d))$; see the converse of Sklar's Theorem. Computing the aggregate capital is then typically done by simulation and estimating the risk measures empirically.
- Problems: (Mis)specification of the copula C (*dependence uncertainty*); Data from $\boldsymbol{L}$ is typically sparse.
- It is generally easier to follow a *fully integrated approach* by building a margins-plus-copula model or more dynamic models (*economic scenario generators*) for the risk-factor changes $\boldsymbol{X} = (X_1, \dots, X_k)$ (more data exists) and for $g_j : \mathbb{R}^k \mapsto \mathbb{R}$ which give the losses $L_j = g_j(\boldsymbol{X})$, $j \in \{1, \dots, d\}$, for the different portfolios/business units. Risk measures are then derived from the distribution of $L = g_1(\boldsymbol{X}) + \dots + g_d(\boldsymbol{X})$.

8.4.4 Risk aggregation and Fréchet problems

- Consider the margins-plus-copula approach where $L_j \sim F_j$, $j \in \{1, \dots, d\}$, are treated as known (estimated or postulated) and C is unknown.
- Consider $L = L_1 + \dots + L_d$. Due to the unknown C (*dependence uncertainty*), risk measures can no longer be computed explicitly.
- Our goal is to find bounds on VaR_α and ES_α under all possible C. Let

$$\mathcal{S}_d := \mathcal{S}_d(F_1, \dots, F_d) := \left\{ L = \sum_{j=1}^d L_j : L_j \sim F_j,\ j = 1, \dots, d \right\}$$

  and consider

$$\overline{\varrho}(L) := \overline{\varrho}(\mathcal{S}_d) := \sup\{\varrho(L) : L \in \mathcal{S}_d(F_1, \dots, F_d)\} \quad \text{(worst } \varrho\text{)}$$

$$\underline{\varrho}(L) := \underline{\varrho}(\mathcal{S}_d) := \inf\{\varrho(L) : L \in \mathcal{S}_d(F_1, \dots, F_d)\} \quad \text{(best } \varrho\text{)}$$

- If $\varrho = \mathrm{ES}_\alpha$, $\overline{\mathrm{ES}}_\alpha(L) = \sum_{j=1}^d \mathrm{ES}_\alpha(L_j)$ (subadditivity, com. additivity). $\underline{\mathrm{ES}}_\alpha$, $\underline{\mathrm{VaR}}_\alpha$, $\overline{\mathrm{VaR}}_\alpha$ depend on whether the portfolio is *homogeneous* (that is, $F_1 = \dots = F_d$); we focus on $\overline{\mathrm{VaR}}_\alpha$.

Summary of existing results

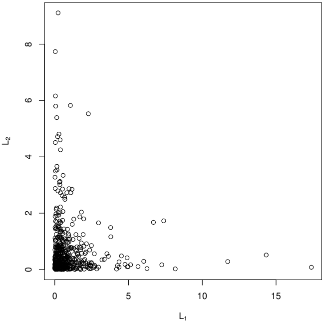
$d = 2$: Fully solved analytically

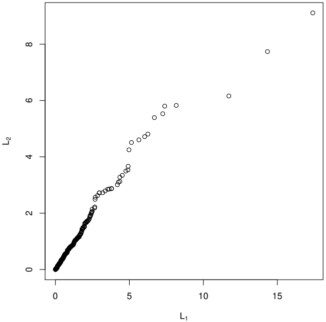
$d \geq 3$: Here we distinguish:

- Homogeneous case ($F_1 = \cdots = F_d$):
  - $\underline{\mathrm{ES}}_\alpha(L)$ solved analytically for decreasing densities (e.g. Pareto, Exponential)
  - $\underline{\mathrm{VaR}}_\alpha(L)$, $\overline{\mathrm{VaR}}_\alpha(L)$ solved analytically for tail-decreasing densities (e.g. Pareto, Log-normal, Gamma)
- Inhomogeneous case:
  - Some analytical results available
  - Numerical methods: (Adaptive/Block) Rearrangement Algorithm

The general problem

- We have one-period risks $L_1 \sim F_1, \dots, L_d \sim F_d$ with given (estimated or postulated) $F_1, \dots, F_d$ and unknown copula C and want to compute $\overline{\mathrm{VaR}_\alpha}(L)$ for $L = L_1 + \cdots + L_d$.
- Iman and Conover (1982) idea for $\mathrm{Par}(2)$, $\mathrm{Par}(2.5)$ sample of size 500:





$\Rightarrow$ Reordering columns changes the dependence of (L_1, L_2) and F_L.

Proposition 8.27 (VaR_α in the homogeneous case)
Let $F := F_1 = \cdots = F_d$ with decreasing density on $[b, \infty)$. Then, for $\alpha \in [F(b), 1)$ and $X \sim F$,

$$\overline{\mathrm{VaR}}_\alpha(\mathcal{S}_d) = d\mathbb{E}(X \,|\, X \in [F^{-1}(\alpha + (d-1)c,\ F^{-1}(1-c)]),$$

where c is the smallest number in $[0, (1-\alpha)/d]$ such that

$$\int_{\alpha+(d-1)c}^{1-c} F^{-1}(t)\,\mathrm{d}t \geq \frac{1-\alpha-dc}{d}((d-1)F^{-1}(\alpha+(d-1)c) + F^{-1}(1-c)).$$

If the density f of F is decreasing on its support, then for $\alpha \in (0,1)$,

$$\underline{\mathrm{VaR}}_\alpha(\mathcal{S}_d) = \max\{(d-1)F^{-1}(0) + F^{-1}(\alpha),\ d\mathbb{E}(X \,|\, X \leq F^{-1}(\alpha))\}.$$

Proof. See Wang et al. (2013) and Bernard et al. (2014). □

- The underlying numerics are non-trivial; see Hofert et al. (2015) and `qrmtools::VaR_bounds_hom()`.

Proposition 8.28 ($\underline{\mathrm{ES}}_\alpha$ in the homogeneous case)
Let $F := F_1 = \cdots = F_d$ with finite first moment and decreasing density on its support. Then, for $\alpha \in [1-dc, 1)$, $\beta = (1-\alpha)/d$, and $X \sim F$,

$$\underline{\mathrm{ES}}_\alpha(\mathcal{S}_d) = \frac{1}{\beta}\int_0^\beta ((d-1)F^{-1}((d-1)t) + F^{-1}(1-t))\,\mathrm{d}t$$
$$= (d-1)^2\,\mathrm{LES}_{(d-1)\beta}(X) + \mathrm{ES}_{1-\beta}(X),$$

where c is the smallest number in $[0, 1/d]$ such that

$$\int_{(d-1)c}^{1-c} F^{-1}(t)\,\mathrm{d}t \geq \frac{1-dc}{d}((d-1)F^{-1}((d-1)c) + F^{-1}(1-c))$$

and $\mathrm{LES}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_u(X)\,\mathrm{d}u = -\mathrm{ES}_{1-\alpha}(-X)$ (*lower ES*).

Proof. See Bernard et al. (2014). $\square$

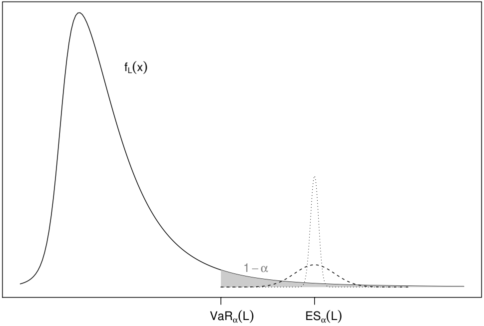
The Rearrangement Algorithm (RA)

- Two columns $\boldsymbol{a}, \boldsymbol{b}$ are *oppositely ordered* if $(a_i - a_j)(b_i - b_j) \le 0 \;\forall\, i, j$.
- *Minimum row-sum operator* $s(X) = \min_{1 \le i \le N} \sum_{1 \le j \le d} x_{ij}$

Algorithm 8.29 (RA for computing $\overline{\mathrm{VaR}}_\alpha(L)$)

1) Fix $\alpha \in (0,1)$, $F_1^{\leftarrow}, \dots, F_d^{\leftarrow}$, $N \in \mathbb{N}$ (# of discr. points), $\varepsilon \ge 0$ (tol.)
2) Compute the lower bound $\underline{s}_N$:
   2.1) Define the (N,d)-matrix $\underline{X}^\alpha = \Bigl(F_j^{\leftarrow}\bigl(\alpha + \frac{(1-\alpha)(i-1)}{N}\bigr)\Bigr)_{i,j}$.
   2.2) Randomly permute each column of $\underline{X}^\alpha$ (to avoid $\overline{s}_N - \underline{s}_N \not\to 0$)
   2.3) Iterate over all columns of $\underline{X}^\alpha$ and oppositely order each to the sum of all others $\Rightarrow$ Matrix $\underline{Y}^\alpha$
   2.4) Repeat Step 2.3) until $s(\underline{Y}^\alpha) - s(\underline{X}^\alpha) \le \varepsilon$, then set $\underline{s}_N = s(\underline{Y}^\alpha)$.
3) Similarly, compute $\overline{s}_N = s(\overline{Y}^\alpha)$ based on $\overline{X}^\alpha = \Bigl(F_j^{\leftarrow}\bigl(\alpha + \frac{(1-\alpha)i}{N}\bigr)\Bigr)_{i,j}$.
4) Return $(\underline{s}_N,\ \overline{s}_N)$ (*rearrangement range*; taken as bounds on $\overline{\mathrm{VaR}}_\alpha(L)$)

- The RA aims at maximizing the minimal row sums (solving a *maximin problem*; minimax problem for $\underline{\mathrm{VaR}}_\alpha$).
- **Intuition:** A *completely mixable matrix* (equal row sums), would minimize the variance of $L\,|\,L > F_L^-(\alpha)$ and thus concentrate more of the $1-\alpha$ mass of F_L around the constant $\mathbb{E}[L\,|\,L > \mathrm{VaR}_\alpha(L)] \underset{\text{cont.}}{=} \mathrm{ES}_\alpha(L) \geq \mathrm{VaR}_\alpha(L)$, so $\mathrm{VaR}_\alpha(L)$ increases (F_L jumps to 1 in $\mathrm{VaR}_\alpha(L)$ so $\mathrm{VaR}_\alpha(L)$ is largest).



Example 8.30 (How the RA works)

1) Where it works (to compute the maximal minimal row sum):

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 5 & 4 \\ 4 & 7 & 8 \end{pmatrix} \underset{\sum_{-1}=\begin{pmatrix} 2\\5\\9\\15 \end{pmatrix}}{\Longrightarrow} \begin{pmatrix} 4 & 1 & 1 \\ 3 & 3 & 2 \\ 2 & 5 & 4 \\ 1 & 7 & 8 \end{pmatrix} \underset{\sum_{-2}=\begin{pmatrix} 5\\5\\6\\9 \end{pmatrix}}{\overset{\text{here: stable}}{\Longrightarrow}} \begin{pmatrix} 4 & 5 & 1 \\ 3 & 7 & 2 \\ 2 & 3 & 4 \\ 1 & 1 & 8 \end{pmatrix} \underset{\sum_{-3}=\begin{pmatrix} 9\\10\\5\\2 \end{pmatrix}}{\Longrightarrow}$$

$$\begin{pmatrix} 4 & 5 & 2 \\ 3 & 7 & 1 \\ 2 & 3 & 4 \\ 1 & 1 & 8 \end{pmatrix} \underset{\sum_{-1}=\begin{pmatrix} 7\\8\\7\\9 \end{pmatrix}}{\overset{\text{here: not}}{\Longrightarrow}} \begin{pmatrix} 3 & 5 & 2 \\ 2 & 7 & 1 \\ 4 & 3 & 4 \\ 1 & 1 & 8 \end{pmatrix} \quad \checkmark \underset{\sum=\begin{pmatrix} 10\\10\\11\\10 \end{pmatrix}}{\Longrightarrow} \widehat{\mathrm{VaR}}_\alpha(L^+) \approx 10$$

2) The RA can also fail:

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} \underset{\sum_{-1}=\begin{pmatrix} 2\\4\\6 \end{pmatrix}}{\Longrightarrow} \begin{pmatrix} 3 & 1 & 1 \\ 2 & 2 & 2 \\ 1 & 3 & 3 \end{pmatrix} \quad \checkmark \underset{\sum=\begin{pmatrix} 5\\6\\7 \end{pmatrix}}{\Longrightarrow} \widehat{\mathrm{VaR}}_\alpha(L^+) \approx 5 < 6$$

Example 8.31 ($\mathrm{Par}(\theta)$ margins)

Let $L_j \sim \mathrm{Par}(\theta)$ with $\bar{F}_j(x) = (1+x)^{-\theta}$, $j \in \{1,\dots,d\}$ (homogeneous case) and $\alpha = 0.999$. One obtains:

	$d=8$		$d=56$	
	$\theta=2$	$\theta=0.8$	$\theta=2$	$\theta=0.8$
$\overline{\mathrm{VaR}}_\alpha(L)$	465	300 182	3454	4 683 172
$\mathrm{VaR}^+_\alpha(L) = d\,\mathrm{VaR}_\alpha(L_1)$	245	44 979	1715	314 855
$\mathrm{VaR}^\perp_\alpha(L)$	96	75 877	293	862 855
$\underline{\mathrm{VaR}}_\alpha(L)$	31	5622	53	5622
$\overline{\mathrm{ES}}_\alpha(L) = d\,\mathrm{ES}_\alpha(L_1)$	498	–	3486	–
$\mathrm{ES}^\perp_\alpha(L)$	184	–	518	–
$\underline{\mathrm{ES}}_\alpha(L)$	178	–	472	–

- The "$+$" and "$\perp$" denote the comonotonic and independent case, resp.
- $\dfrac{\overline{\mathrm{ES}}_\alpha(L)}{\overline{\mathrm{VaR}}_\alpha(L)} \underset{d\uparrow\infty}{\approx} 1$ can be explained; see McNeil et al. (2015, Prop. 8.36).
- The dependence uncertainty spread $\overline{\mathrm{VaR}}_\alpha(L) - \underline{\mathrm{VaR}}_\alpha(L) \ge \overline{\mathrm{ES}}_\alpha(L) - \underline{\mathrm{ES}}_\alpha(L)$ can be explained; see McNeil et al. (2015, Prop. 8.37).

Remark 8.32

- The RA finds approximate solutions to *maximin* (for $\overline{\mathrm{VaR}}_\alpha(L)$) and *minimax* (for $\underline{\mathrm{VaR}}_\alpha(L)$) *problems* and is thus of wider interest (e.g., in Operations Research).
- For $\underline{\mathrm{ES}}_\alpha(L)$, discretize the whole support of each margin, rearrange, and approximate $\underline{\mathrm{ES}}_\alpha(L)$ by the nonparametric ES_α estimate of the row sums.
- The *Adaptive Rearrangement Algorithm (ARA)*
  - uses relative (instead of absolute) individual tolerances;
  - uses a relative joint tolerance to guarantee that $\underline{s}_N$ and $\overline{s}_N$ are close;
  - chooses N adaptively to reach the joint tolerance; and
  - determines convergence after each rearranged column.
- The *Block Rearrangement Algorithm* rearranges blocks of columns.

Proposition 8.33 (Asymptotic equivalence of $\overline{\mathrm{VaR}}_\alpha$, $\overline{\mathrm{ES}}_\alpha$)

Suppose that $L_j \sim F_j$, $j \geq 1$ and that

i) for some $k > 1$, $\mathbb{E}(|L_j - \mathbb{E}(L_j)|^k)$ is uniformly bounded, and

ii) for some $\alpha \in (0,1)$, $\liminf_{d\to\infty} \frac{1}{d}\sum_{j=1}^d \mathrm{ES}_\alpha(L_j) > 0$.

Then, as $d \to \infty$, $\dfrac{\overline{\mathrm{ES}}_\alpha(\mathcal{S}_d)}{\overline{\mathrm{VaR}}_\alpha(\mathcal{S}_d)} = 1 + O(d^{\frac{1}{k}-1})$.

Proposition 8.34 (Dependence uncertainty spread of VaR_α vs ES_α)

Let $0 < \alpha_1 \leq \alpha_2 < 1$, assume Proposition 8.33 i) to hold and that $\liminf_{d\to\infty} \dfrac{1}{d}\sum_{j=1}^d \mathrm{LES}_{\alpha_1}(X_j) > 0$ and $\limsup_{d\to\infty} \dfrac{\sum_{j=1}^d \mathbb{E}(X_j)}{\sum_{j=1}^d \mathrm{ES}_{\alpha_1}(X_j)} < 1$. Then

$$\liminf_{d\to\infty} \frac{\overline{\mathrm{VaR}}_{\alpha_2}(\mathcal{S}_d) - \underline{\mathrm{VaR}}_{\alpha_2}(\mathcal{S}_d)}{\overline{\mathrm{ES}}_{\alpha_1}(\mathcal{S}_d) - \underline{\mathrm{ES}}_{\alpha_1}(\mathcal{S}_d)} \geq 1$$

Example 8.35 (Superadditivity of VaR under special dependence)

Let $\alpha \in (0,1)$, $L_1 \sim \mathrm{U}(0,1)$ and define $L_2 \overset{\text{a.s.}}{=} \begin{cases} L_1, & \text{if } L_1 < \alpha, \\ 1+\alpha-L_1, & \text{if } L_1 \geq \alpha. \end{cases}$

One can show that $L_2 \sim \mathrm{U}(0,1)$. Also, $L_1 + L_2 = \begin{cases} 2L_1, & \text{if } L_1 < \alpha, \\ 1+\alpha, & \text{if } L_1 \geq \alpha, \end{cases}$

from which one can show that

$$F_{L_1+L_2}(x) = \begin{cases} 0, & \text{if } x < 0, \\ x/2, & \text{if } x \in [0, 2\alpha), \\ \alpha, & \text{if } x \in [2\alpha, 1+\alpha), \\ 1, & \text{if } x \geq 1+\alpha. \end{cases}$$

For all $\varepsilon \in \left(0, \frac{1-\alpha}{2}\right)$, we thus obtain that

$$\mathrm{VaR}_{\alpha+\varepsilon}(L_1+L_2) = 1+\alpha \underset{\varepsilon \in \left(0, \frac{1-\alpha}{2}\right)}{>} 2(\alpha+\varepsilon) = \mathrm{VaR}_{\alpha+\varepsilon}(L_1) + \mathrm{VaR}_{\alpha+\varepsilon}(L_2).$$

8.5 Capital allocation

How can the overall capital requirement may be disaggregated into additive contributions/units/investments? Motivation: How can we measure the risk-adjusted performance of different investments?

8.5.1 The allocation problem

- The performance of investments is usually measured using a *RORAC (return on risk-adjusted capital) approach* by considering

$$\frac{\text{expected profit of investment } j}{\text{risk capital for investment } j}.$$

- The risk capital of investment j with loss L_j can be computed as follows: Compute $\varrho(L) = \varrho(L_1 + \cdots + L_d)$. Then allocate $\varrho(L)$ to the investments according to a *capital allocation principle* such that

$$\varrho(L) = \sum_{j=1}^{d} \mathrm{AC}_j,$$

where the *risk contribution* AC_j is the capital allocated to investment j.

The formal set-up

- Consider an open set $\mathbf{1} \in \Lambda \subseteq \mathbb{R}^d \setminus \{\mathbf{0}\}$ of portfolio weights and define

$$L(\boldsymbol{\lambda}) = \boldsymbol{\lambda}'\boldsymbol{L} = \sum_{j=1}^{d} \lambda_j L_j, \quad \boldsymbol{\lambda} \in \Lambda.$$

- For a risk measure ϱ, define the *associated risk-measure function*

$$r_\varrho(\boldsymbol{\lambda}) = \varrho(L(\boldsymbol{\lambda})),$$

  so that $r_\varrho(\mathbf{1}) = \varrho(L)$.

8.5.2 The Euler principle and examples

- If r_ϱ is positive homogeneous and differentiable at $\boldsymbol{\lambda} \in \Lambda$, Euler's rule (see the appendix) implies that

$$r_\varrho(\boldsymbol{\lambda}) = \sum_{i=1}^{d} \lambda_i \frac{\partial r_\varrho}{\partial \lambda_i}(\boldsymbol{\lambda}) \quad \text{so} \quad \varrho(L) = r_\varrho(\mathbf{1}) = \sum_{j=1}^{d} \frac{\partial r_\varrho}{\partial \lambda_j}(\mathbf{1}).$$

  Note that r_ϱ is positive homogeneous if ϱ is.

Definition 8.36 (Euler capital allocation principle)

If r_ϱ is a pos.-hom. risk-measure function and differentiable at $\boldsymbol{\lambda} = \mathbf{1}$, then the *Euler capital allocation principle* has risk contributions

$$\mathrm{AC}_j = \mathrm{AC}_j^\varrho := \frac{\partial r_\varrho}{\partial \lambda_j}(\mathbf{1}), \quad j \in \{1, \dots, d\}.$$

Examples

1) Standard deviation and the covariance principle

- Consider $r_{\mathsf{SD}}(\boldsymbol{\lambda}) = \sqrt{\mathrm{var}(L(\boldsymbol{\lambda}))} = \sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}}$ where Σ is the covariance matrix of $(L_1, \dots, L_d)$. Therefore

$$\mathrm{AC}_j^\varrho = \frac{\partial r_{\mathsf{SD}}}{\partial \lambda_j}(\mathbf{1}) = \frac{(\Sigma\mathbf{1})_j}{r_{\mathsf{SD}}(\mathbf{1})} = \frac{\sum_{k=1}^d \mathrm{cov}(L_j, L_k)}{r_{\mathsf{SD}}(\mathbf{1})} = \frac{\mathrm{cov}(L_j, L)}{\sqrt{\mathrm{var}(L)}}.$$

  This formula is known as *covariance principle*.

- If we consider the more general $\varrho(L) = \mathbb{E}(L) + \kappa\,\mathrm{SD}(L)$ for some $\kappa > 0$ we get

$$r_\varrho(\boldsymbol{\lambda}) = \boldsymbol{\lambda}'\mathbb{E}(\boldsymbol{L}) + \kappa r_{\mathsf{SD}}(\boldsymbol{\lambda})$$

and hence

$$\mathrm{AC}_j^\varrho = \mathbb{E}(L_j) + \kappa\frac{\mathrm{cov}(L_j, L)}{\sqrt{\mathrm{var}(L)}}.$$

2) VaR and VaR contributions

Suppose that $r^\alpha_{\mathsf{VaR}}(\boldsymbol{\lambda}) = q_\alpha(L(\boldsymbol{\lambda}))$. In this case it can be shown (non-trivial) that, subject to technical conditions,

$$\mathrm{AC}_j^\varrho = \frac{\partial r^\alpha_{\mathsf{VaR}}}{\partial\lambda_j}(\mathbf{1}) = \mathbb{E}(L_j \,|\, L = F_L^{\leftarrow}(\alpha)), \quad j \in \{1,\dots,d\}.$$

3) Expected shortfall and shortfall contributions

Now consider $r^\alpha_{\mathrm{ES}}(\boldsymbol{\lambda}) = \mathbb{E}(L \,|\, L \geq q_\alpha(L(\boldsymbol{\lambda})))$. Then

$$r^\alpha_{\mathsf{ES}}(\boldsymbol{\lambda}) = \frac{1}{1-\alpha}\int_\alpha^1 r^u_{\mathsf{VaR}}(\boldsymbol{\lambda})\,\mathrm{d}u,$$

Assuming the differentiability of $r^u_{\mathsf{VaR}}(\boldsymbol{\lambda})$, the Euler principle implies that

$$\frac{\partial r^\alpha_{\mathsf{ES}}}{\partial \lambda_j}(\mathbf{1}) = \frac{1}{1-\alpha}\int_\alpha^1 \frac{\partial r^u_{\mathsf{VaR}}}{\partial \lambda_j}(\mathbf{1})\,\mathrm{d}u = \frac{1}{1-\alpha}\int_\alpha^1 \mathbb{E}(L_j \mid L = F_L^{\leftarrow}(u))\,\mathrm{d}u.$$

If F_L has a differentiable inverse,

$$\frac{\partial r^\alpha_{\mathsf{ES}}}{\partial \lambda_j}(\mathbf{1}) = \frac{1}{1-\alpha}\int_{F_L^{\leftarrow}(\alpha)}^\infty \mathbb{E}(L_j \mid L = v) f_L(v)\,\mathrm{d}v = \frac{\mathbb{E}(L_j; L \geq F_L^{\leftarrow}(\alpha))}{1-\alpha}.$$

Hence the Euler capital allocation takes the form

$$\mathrm{AC}_j^\varrho = \mathbb{E}(L_j \mid L \geq \mathrm{VaR}_\alpha(L)), \quad L := L(\mathbf{1});$$

AC_j^ϱ is known as the *expected shortfall contribution* of investment j. This is a popular allocation principle in practice.

4) Euler allocation for elliptical loss distributions

The following result shows that allocation is very simple in the case of $\boldsymbol{L} \sim E_d(\mathbf{0}, \Sigma, \psi)$: Calculate the total risk capital and then use a simple partitioning formula (regardless of the pos.-hom. risk measure).

Corollary 8.37 (Euler allocation under ellipticality)

Assume that r_ϱ is the risk-measure function of a positive-homogeneous and law invariant ϱ. Let $\boldsymbol{L} \sim E_d(\mathbf{0}, \Sigma, \psi)$. Then, under an Euler allocation,

$$\frac{\mathrm{AC}_j^\varrho}{\mathrm{AC}_k^\varrho} = \frac{\sum_{l=1}^d \Sigma_{jl}}{\sum_{l=1}^d \Sigma_{kl}}, \quad j, k \in \{1, \dots, d\}.$$

Proof. The proof of Theorem 8.24 implies that, by positive homogeneity,

$$r_\varrho(\boldsymbol{\lambda}) = \varrho(L(\boldsymbol{\lambda})) = \varrho\biggl(\sum_{j=1}^d \lambda_j L_j\biggr) = \sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}}\,\varrho(Y_1),$$

where Y_1 is the first component of $\boldsymbol{Y} \sim S_d(\psi)$. For the Euler allocation we get

$$\mathrm{AC}_j^\varrho = \frac{\partial r_\varrho}{\partial \lambda_j}(\mathbf{1}) = \frac{\sum_{k=1}^d \Sigma_{jk}}{\sqrt{\mathbf{1}'\Sigma\mathbf{1}}}\,\varrho(Y_1)$$

from which the result follows. $\square$

8.5.3 Economic properties of the Euler principle

- We show that the Euler principle has good economic properties.
- Assume that r_ϱ is continuously differentiable in $\mathbb{R}^d \setminus \{\mathbf{0}\}$ and by

$$\mathrm{AC}_j^\varrho = \frac{\partial r_\varrho}{\partial \lambda_j}(\mathbf{1}), \quad j \in \{1, \dots, d\},$$

denote the associated risk contributions under the Euler principle.

Compatibility with a RORAC approach

- The *RORAC (return on risk adjusted capital)* is defined as

$$\mathrm{RORAC}(L) := \frac{\mathbb{E}(-L)}{\varrho(L)}$$

and the *portfolio-related RORAC* of investment j is defined as

$$\mathrm{RORAC}(L_j \,|\, L) := \frac{\mathbb{E}(-L_j)}{\mathrm{AC}_j^\varrho}.$$

- The Euler principle is compatible with a RORAC approach: If investment j performs better than the overall portfolio L in the RORAC metric, then the latter is increased if one increases the weight of unit j. Hence the Euler principle gives correct signals for investment decisions.
- In mathematical terms, RORAC compatibility means that there is some $\varepsilon > 0$ such that for all $0 < h \le \varepsilon$

$$\mathrm{RORAC}(L_j \,|\, L) > \mathrm{RORAC}(L) \;\Rightarrow\; \mathrm{RORAC}(L + hL_j) > \mathrm{RORAC}(L).$$

Proof. $\frac{d}{dh}\mathrm{RORAC}(L + hL_j)|_{h=0}$

$$= \frac{d}{dh}\frac{\mathbb{E}(-(L+hL_j))}{r_\varrho(\mathbf{1}+h\boldsymbol{e}_j)}\Big|_{h=0} = \frac{1}{r_\varrho(\mathbf{1})^2}\Big(\mathbb{E}(-L_j)r_\varrho(\mathbf{1}) - \mathbb{E}(-L)\frac{\partial r_\varrho(\mathbf{1})}{\partial \lambda_j}\Big),$$

$$= \frac{1}{r_\varrho(\mathbf{1})^2}\big(\mathbb{E}(-L_j)\varrho(L) - \mathbb{E}(-L)\mathrm{AC}_j^\varrho\big) > 0$$

if $\frac{\mathbb{E}(-L_j)}{\mathrm{AC}_j^\varrho} = \mathrm{RORAC}(L_j \,|\, L) > \mathrm{RORAC}(L) = \frac{\mathbb{E}(-L)}{\varrho(L)}$. □

Diversification benefit

- For a subadditive ϱ, $\sum_{j=1}^{d} \varrho(L_j) - \varrho(L) > 0$ is known as *diversification benefit*.
- It is reasonable to require that each business unit profits from the diversification benefit in the sense that

$$\mathrm{AC}_j^{\varrho} \le \varrho(L_j), \quad j \in \{1, \dots, d\}.$$

- We now show that the Euler principle does indeed satisfy this inequality.

Proof. Let $f : \mathbb{R}^d \to \mathbb{R}$ be convex, pos.-hom. and continuously differentiable in $\mathbb{R}^d \setminus \{\mathbf{0}\}$. By convexity,

$$f(\boldsymbol{y}) \ge f(\boldsymbol{x}) + \sum_{j=1}^{d} (y_j - x_j) \frac{\partial f}{\partial x_j}(\boldsymbol{x}), \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d,\ \boldsymbol{x} \ne \mathbf{0}.$$

By Euler's rule, $f(\boldsymbol{x}) = \sum_{j=1}^{d} x_j \frac{\partial f}{\partial x_j}(\boldsymbol{x})$ and hence

$$f(\boldsymbol{y}) \geq \sum_{j=1}^{d} y_j \, \frac{\partial f}{\partial x_j}(\boldsymbol{x}).$$

For $\boldsymbol{y} = \boldsymbol{\lambda}$ and $\boldsymbol{x} = \boldsymbol{\lambda} + \tilde{\boldsymbol{\lambda}}$, we obtain

$$f(\boldsymbol{\lambda}) \geq \sum_{j=1}^{d} \lambda_j \frac{\partial f}{\partial \lambda_j}(\boldsymbol{\lambda} + \tilde{\boldsymbol{\lambda}}) \quad \text{for all } \boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}} : \boldsymbol{\lambda} \neq -\tilde{\boldsymbol{\lambda}}.$$

Apply this inequality with $f = r_\varrho$ (which is convex as ϱ is pos.-hom. and subadditive), $\boldsymbol{\lambda} = \boldsymbol{e}_j$ and $\tilde{\boldsymbol{\lambda}} = \mathbf{1} - \boldsymbol{e}_j$ to obtain

$$\varrho(L_j) = r_\varrho(\boldsymbol{e}_j) \geq \frac{\partial r_\varrho}{\partial \lambda_j}(\mathbf{1}) = \mathrm{AC}_j^\varrho \, . \qquad \square$$

- From a practical point of view, expected shortfall and expected shortfall contributions are typically a reasonable choice in many applications.