

# 8 Aggregate risk

- 8.1 Coherent and convex risk measures
- 8.2 Law-invariant coherent risk measures
- 8.3 Risk measures for linear portfolios
- 8.4 Risk aggregation
- 8.5 Capital allocation

## 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  $\varrho : \mathcal{M} \rightarrow \mathbb{R}$ ;  $\varrho(L)$  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  $\varrho$  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  $\varrho$  the *acceptance set of  $\varrho$*  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  $\varrho$  be monotone and translation-invariant with associated  $A_\varrho$ . Then

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

$$L \in A_\varrho \text{ and } \tilde{L} \leq L \Rightarrow \tilde{L} \in A_\varrho. \quad (35)$$

2)  $\varrho$  can be reconstructed from  $A_\varrho$  via

$$\varrho(L) = \inf\{m \in \mathbb{R} : L - m \in A_\varrho\}. \quad (36)$$

*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 (35) and define

$$\varrho_A(L) = \inf\{m \in \mathbb{R} : L - m \in A\}. \quad (37)$$

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

*Proof.* These properties of  $\varrho_A$  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 (37) is given by

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

## Proposition 8.6

- 1) Let  $\varrho$  be monotone and translation-invariant. Then
  - 1.1)  $\varrho$  is convex if and only if  $A_\varrho$  is convex.
  - 1.2)  $\varrho$  is coherent if and only if  $A_\varrho$  is a convex cone.
- 2) More generally, consider a set of acceptable positions  $A$  and the associated risk measure  $\varrho_A$  (whose acceptance set may be larger than  $A$ ). If  $A$  is convex, so is  $\varrho_A$ ; if  $A$  is a convex cone, then  $\varrho_A$  is coherent.

## Example 8.7 (Risk measures based on loss functions)

Consider a strictly increasing and convex *loss function*  $\ell : \mathbb{R} \rightarrow \mathbb{R}$  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

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

and the associated risk measure by

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

- $\varrho_A$  is translation invariant and monotone by Proposition 8.4 since  $A$  satisfies (35).
- $\varrho_A$  is convex by Proposition 8.6; to see this consider acceptable positions  $L_1$  and  $L_2$  and observe that the convexity of  $\ell$  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*  $\gamma : \mathcal{Q} \rightarrow \mathbb{R}$  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*  $\varrho$  is defined by

$$\varrho(L) = \sup\{\mathbb{E}_{\mathbb{Q}}(L) - \gamma(\mathbb{Q}) : \mathbb{Q} \in \mathcal{Q}\}. \quad (38)$$

The corresponding acceptance set is given by

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

- $A_{\varrho}$  is convex, and thus so is  $\varrho$ .
- Every convex risk measure can be represented as (38); see Theorem 8.10.
- If  $\gamma(\cdot) \equiv 0$  on  $\mathcal{Q}$ ,  $\varrho$  is positive homogeneous and therefore coherent.
- The stress test risk measure of Example 8.8 is a special case of (38) 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  $\mathcal{M} := \{L : \Omega \rightarrow \mathbb{R}\}$ . Then:

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

$$\varrho(L) = \max\{\mathbb{E}_{\mathbb{Q}}(L) - \alpha_{\min}(\mathbb{Q}) : \mathbb{Q} \in \mathcal{S}^1(\Omega, \mathcal{F})\}, \quad (39)$$

where  $\mathcal{S}^1(\Omega, \mathcal{F})$  denotes the set of all probability measures on  $\Omega$ , 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  $\varrho$  is coherent, it has the representation

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

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)

Let  $(x)_+ = \max\{x, 0\}$ . For  $\alpha \in (0, 1)$ ,

$$\begin{aligned} 1) \quad \text{ES}_\alpha(L) &= \frac{\mathbb{E}((L - F_L^{\leftarrow}(\alpha))_+)}{1 - \alpha} + F_L^{\leftarrow}(\alpha); \\ 2) \quad \text{ES}_\alpha(L) &= \frac{\mathbb{E}(LI_{\{L > F_L^{\leftarrow}(\alpha)\}}) + F_L^{\leftarrow}(\alpha)(1 - \alpha - \bar{F}_L(F_L^{\leftarrow}(\alpha)))}{1 - \alpha}. \end{aligned}$$

### Corollary 8.12 (ES formulas under continuous $F_L$ )

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

$$\begin{aligned} 1) \quad \text{ES}_\alpha(L) &= \frac{\mathbb{E}(LI_{\{L > F_L^{\leftarrow}(\alpha)\}})}{1 - \alpha} \\ 2) \quad \text{ES}_\alpha(L) &= \mathbb{E}(L \mid L > F_L^{\leftarrow}(\alpha)) \text{ (i.e. } \textcolor{blue}{\textit{conditional VaR (CVaR)}}) \end{aligned}$$

With dual representations one can give a proof for  $\text{ES}_\alpha$  being subadditive; see the following result.

### Theorem 8.13

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

$$\text{ES}_\alpha(L) = \max\{\mathbb{E}^{\mathbb{Q}}(L) : \mathbb{Q} \in \mathcal{Q}_\alpha\}, \quad (40)$$

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)

- 1) A *convex distortion function*  $D$  is a convex, increasing and absolutely continuous function on  $[0, 1]$  satisfying  $D(0) = 0$  and  $D(1) = 1$ .
- 2) The *distortion risk measure* associated with a convex distortion function  $D$  is defined by

$$\varrho(L) = \int_0^1 q_u(L) dD(u). \quad (41)$$

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

$$\varrho(L) = \int_0^1 q_u(L) \phi(u) du. \quad (42)$$

A risk measure of this form is known as *spectral risk measure* and  $\phi$  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  $\alpha$ -quantile).

### Lemma 8.15

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

$$\varrho(L) = \int_{\mathbb{R}} x dD \circ F_L(x), \quad (43)$$

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

*Proof.*  $G(x) = D \circ F_L(x)$  has quantile function  $G^{\leftarrow} = F_L^{\leftarrow} \circ D^{\leftarrow}$ . Thus (43) can be written as

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

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

$$\int_{\mathbb{R}} x dD \circ F_L(x) = \mathbb{E}(F_L^{\leftarrow}(V)) = \int_0^1 F_L^{\leftarrow}(v) dD(v). \quad \square$$

$D$  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  $\varrho$  be a distortion risk measure associated with the convex distortion function  $D$ . Then, for a probability measure  $\mu$ ,

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

## Properties of distortion risk measures

### Definition 8.17 (Comonotone additivity)

A risk measure  $\varrho$  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  $L_1, \dots, L_d$ .

- Quantile functions (so value-at-risk) are comonotone additive. Comonotone additivity of distortion risk measures then follows from (41).



- Distortion risk measures are coherent. Monotonicity, translation invariance and positive homogeneity are obvious. Subadditivity follows from Proposition 8.16 and subadditivity of  $\text{ES}_\alpha$  (e.g., Theorem 8.13) by observing that

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

- In summary, we have verified that distortion risk measures are law invariant, coherent and comonotone additive.
- It may also be shown that, on a probability space without atoms ( $\mathbb{P}(\{\omega\}) = 0$  for all  $\omega$ ), a law-invariant, coherent, comonotone-additive risk measure must be of the form (41) for some convex distortion function  $D$ .

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

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

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

- ▶ Such a family of convex distortion functions is strictly decreasing in  $\alpha$  for fixed  $u$ .
- ▶  $D_0(u) = u$  (corresponding to the risk measure  $\varrho(L) = \mathbb{E}(L)$ ) and  $\lim_{\alpha \rightarrow 1} D(u) = 1_{\{u=1\}}$ .
- ▶ For  $\alpha_1 < \alpha_2$  and  $0 < u < 1$  we have  $D_{\alpha_1}(u) > D_{\alpha_2}(u)$ , so that  $D_{\alpha_2}$  distorts the original probability measure more than  $D_{\alpha_1}$  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  $\alpha$ -expectile  $e_\alpha(L)$  is given by the unique solution  $y$  of

$$\alpha \mathbb{E}((L - y)^+) = (1 - \alpha) \mathbb{E}((L - y)^-) \quad (44)$$

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

$$\min_{y \in \mathbb{R}} \mathbb{E}(S(y, L)) \quad (45)$$

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

tion that yields the expectile is

$$S_{\alpha}^e(y, L) = |1_{\{L \leq y\}} - \alpha|(L - y)^2. \quad (46)$$

In fact we can compute that  $\frac{d}{dy}\mathbb{E}(S_{\alpha}^e(y, L))$  equals

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

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

- One can show that the  $\alpha$ -quantile  $F_L^{\leftarrow}(\alpha)$  is also a minimizer of the form (45); consider the scoring function  $S_{\alpha}^q(y, L) = |1_{\{L \leq y\}} - \alpha||L - y|$ .

The following result shows uniqueness of the  $\alpha$ -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

$$\tilde{F}_L(y) = \frac{yF_L(y) - \mu(y)}{2(yF_L(y) - \mu(y)) + \mu - y} \quad (47)$$

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

### Example 8.20 (Bernoulli)

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

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

from which it follows that  $\tilde{F}_L(y) = \frac{y(1-p)}{y(1-2p)+p}$ ,  $0 \leq y \leq 1$  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  $\text{VaR}_\alpha(L) \in \{0, 1\}$ ,  $\alpha \in (0, 1]$ .

## 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

$$\mathcal{M} = \{L : L = m + \boldsymbol{\lambda}'\mathbf{X}, m \in \mathbb{R}, \boldsymbol{\lambda} \in \mathbb{R}^d\}, \quad (48)$$

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

- **Many standard approaches** to risk aggregation and capital allocation are based on the **assumption** that losses have a linear relationship to underlying risk factors.
- 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  $\varrho : \mathcal{M} \rightarrow \mathbb{R}$  be a positive-homogeneous risk measure. Define a **risk-measure function**  $r_\varrho(\boldsymbol{\lambda}) = \varrho(\boldsymbol{\lambda}'\mathbf{X})$  (function of portfolio weights).



- There is a one-to-one relationship between  $\varrho$  and  $r_\varrho$  given by

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

### **Lemma 8.22 (Properties of $r_\varrho$ )**

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

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

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

### Proposition 8.23 (Coherent risk measures for linear portfolios as stress tests)

$\varrho$  is a coherent risk measure on the set of linear portfolios  $\mathcal{M}$  in (48) if and only if for every  $L = m + \boldsymbol{\lambda}'\mathbf{X} \in \mathcal{M}$  we have the representation

$$\varrho(L) = m + r_{\varrho}(\boldsymbol{\lambda}) = \sup\{m + \boldsymbol{\lambda}'\mathbf{x} : \mathbf{x} \in S_{\varrho}\}. \quad (49)$$

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

## 8.3.2 Elliptically distributed risk factors

### Theorem 8.24 (Risk measurement for elliptical risk factors)

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

- 1) For any  $L = m + \boldsymbol{\lambda}'\mathbf{X} \in \mathcal{M}$ ,  $\varrho(L) = m + \boldsymbol{\lambda}'\boldsymbol{\mu} + \sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}}\varrho(Y_1)$  for  $Y_1 \sim S_1(\psi)$ .
- 2) If  $\varrho(Y_1) \geq 0$ ,  $\varrho$  is subadditive on  $\mathcal{M}$  (e.g.,  $\text{VaR}_\alpha$  for  $\alpha \geq 0.5$ ).
- 3) If  $\mathbb{E}\mathbf{X}$  exists then, for any  $L = m + \boldsymbol{\lambda}'\mathbf{X} \in \mathcal{M}$  and  $\rho_{ij} = \wp(\Sigma)_{ij}$ ,

$$\varrho(L - \mathbb{E}L) = \sqrt{\sum_{i=1}^d \sum_{j=1}^d \rho_{ij} \lambda_i \lambda_j \varrho(X_i - \mathbb{E}X_i) \varrho(X_j - \mathbb{E}X_j)}.$$

- 4) If  $\text{cov}(\mathbf{X})$  exists and  $\varrho(Y_1) > 0$  then, for every  $L \in \mathcal{M}$ ,  
 $\varrho(L) = \mathbb{E}(L) + k_\varrho \sqrt{\text{var}(L)}$  for some  $k_\varrho > 0$  depending on  $\varrho$ .
- 5) If  $\Sigma^{-1}$  ex.,  $\varrho(Y_1) > 0$  then  $S_\varrho = \{\mathbf{x} : (\mathbf{x} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) \leq \varrho(Y_1)^2\}$ .

*Proof.*

- 1) Let  $\mathbf{Y} \sim S_k(\psi)$ ,  $AA' = \Sigma$ .  $L = m + \boldsymbol{\lambda}'\mathbf{X} \stackrel{d}{=} m + \boldsymbol{\lambda}'\boldsymbol{\mu} + \boldsymbol{\lambda}'A\mathbf{Y}$ . By Theorem 6.15 3),  $L \stackrel{d}{=} m + \boldsymbol{\lambda}'\boldsymbol{\mu} + \|A'\boldsymbol{\lambda}\|Y_1$ . 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  $L_1 = m_1 + \boldsymbol{\lambda}'_1\mathbf{X}$  and  $L_2 = m_2 + \boldsymbol{\lambda}'_2\mathbf{X}$ . Subadditivity follows from 1) and  $\|A'(\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2)\| \leq \|A'\boldsymbol{\lambda}_1\| + \|A'\boldsymbol{\lambda}_2\|$  and  $\varrho(Y_1) \geq 0$ .
- 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  $\boldsymbol{\lambda} = \mathbf{e}_j$ ,  $\varrho(X_j - \mathbb{E}X_j) = \varrho(\mathbf{e}'_j\mathbf{X} - \mathbb{E}(\mathbf{e}'_j\mathbf{X})) = \sigma_j\varrho(Y_1)$ , from which the result follows.

- 4)  $\text{cov}(\mathbf{X}) = c\Sigma$  for some  $c > 0$ . Since  $\text{var}(L) = \text{var}(\boldsymbol{\lambda}'\mathbf{X}) = \boldsymbol{\lambda}'c\Sigma\boldsymbol{\lambda}$ , 3) implies that  $\varrho(L) = \mathbb{E}(L) + \sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}}\varrho(Y_1) = \mathbb{E}(L) + \sqrt{\text{var}(L)}\varrho(Y_1)/\sqrt{c}$ .

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

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

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

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

- 2) gives a **special case where VaR is subadditive** and thus coherent. In particular, if  $(L_1, \dots, L_d)$  is jointly elliptical,  $\text{VaR}_\alpha$  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  $\varrho$  also minimize the variance**. The portfolio minimizing  $\varrho$  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  $\mathbf{X}$  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

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

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

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

$$\text{EC} = \sqrt{\sum_{i=1}^d \sum_{j=1}^d \rho_{ij} \text{EC}_i \text{EC}_j}, \quad (50)$$

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  $L = L_1 + \dots + L_d$  where  $L_1, \dots, L_d$  are the losses arising from sub-units (e.g., business units, asset classes). Consider a translation-invariant  $\varrho$  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

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

and the aggregate capital should be

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

- We require an aggregation rule  $f$  such that  $f(\text{EC}_1, \dots, \text{EC}_d) = \text{EC}$ .



- If  $\varrho(L) = \mathbb{E}(L) + k \text{sd}(L)$ ,  $k > 0$ , and  $\mathbb{E}(L^2) < \infty$  then

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

where  $(\rho_{ij})_{i,j} = \text{corr}(\mathbf{L})$ , so correlation adjusted summation follows by noting that  $\text{sd}(L) = \varrho^{\text{mean}}(L)/k = \text{EC}/k$  (and  $\text{sd}(L_j) = \text{EC}_j/k$ ).

- If  $L_j = m_j + \boldsymbol{\lambda}'_j \mathbf{X}$  for  $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \psi)$  with existing  $\text{cov}(\mathbf{X})$ , 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  $\varrho$ .
- As the following result shows, the assumption on  $\text{cov}(\mathbf{X})$  can be dropped.

**Proposition 8.25 (Correlation adjusted sum. for linear portfolios)**

Let  $\mathbf{X} \sim E_k(\boldsymbol{\mu}, \Sigma, \psi)$  with  $\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu}$ . Let  $\mathcal{M} = \{L : L = m + \boldsymbol{\lambda}'\mathbf{X}, \boldsymbol{\lambda} \in \mathbb{R}^k, m \in \mathbb{R}\}$  and  $\varrho$  be a pos.-hom., translation- and law-invariant risk measure on  $\mathcal{M}$ . For  $L_1, \dots, L_d \in \mathcal{M}$ , let  $\text{EC}_j = \varrho^{\text{mean}}(L_j)$  and  $\text{EC} = \varrho^{\text{mean}}(L_1 + \dots + L_d)$ . Then  $\text{EC}, \text{EC}_1, \dots, \text{EC}_d$  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  $L_j = m_j + \boldsymbol{\lambda}_j'\mathbf{X}$ . By Theorem 8.24 1),  $\text{EC}_j = \varrho(L_j) - \mathbb{E}(L_j) = \sqrt{\boldsymbol{\lambda}_j'\Sigma\boldsymbol{\lambda}_j}\varrho(Y_1)$  where  $Y_1 \sim S_1(\psi)$  and that

$$\begin{aligned} \text{EC} &= \sqrt{(\boldsymbol{\lambda}_1 + \dots + \boldsymbol{\lambda}_d)' \Sigma (\boldsymbol{\lambda}_1 + \dots + \boldsymbol{\lambda}_d)} \varrho(Y_1) \\ &= \sqrt{\sum_{i=1}^d \sum_{j=1}^d \boldsymbol{\lambda}_i' \Sigma \boldsymbol{\lambda}_j \varrho(Y_1)^2} = \sqrt{\sum_{i=1}^d \sum_{j=1}^d \frac{\boldsymbol{\lambda}_i' \Sigma \boldsymbol{\lambda}_j}{\sqrt{(\boldsymbol{\lambda}_i' \Sigma \boldsymbol{\lambda}_i)(\boldsymbol{\lambda}_j' \Sigma \boldsymbol{\lambda}_j)}} \text{EC}_i \text{EC}_j}. \end{aligned}$$

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  $\lambda_i' \Sigma \lambda_j / \sqrt{(\lambda_i' \Sigma \lambda_i)(\lambda_j' \Sigma \lambda_j)}$  and thus

$$\text{EC} = \sqrt{\sum_{i=1}^d \sum_{j=1}^d \frac{\lambda_i' \Sigma \lambda_j}{\sqrt{(\lambda_i' \Sigma \lambda_i)(\lambda_j' \Sigma \lambda_j)}} \text{EC}_i \text{EC}_j} = \sqrt{\sum_{i=1}^d \sum_{j=1}^d \rho_{ij} \text{EC}_i \text{EC}_j}. \quad \square$$

- Correlation adjusted summation can thus be justified under the mean-adjusted VaR or ES if  $\mathbf{L}$  is elliptical.
- The formula requires the pairwise correlations  $\rho_{ij}$  between the losses  $L_1, \dots, L_d$ . It is difficult to obtain estimates of  $\rho_{ij}$  (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  $\varrho$  is coherent.

## 8.4.2 Aggregation based on stressing risk factors

- Correlation adjusted summation is used in the aggregation of capital contributions  $EC_1, \dots, EC_d$  computed by stressing individual risk factors. An example is the standard formula approach to Solvency II.
- Let  $x = X(\omega)$  be a scenario defined in terms of changes in risk factors and  $L(x)$  the corresponding loss. Assume  $L(x)$  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

$$EC_j = L(k_j e_j) - L(\mathbb{E}(X_j) e_j)$$

where  $k_j > \mathbb{E}(X_j)$  so that  $EC_j > 0$  (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  $\alpha$ .

- The following justifies correlation adjusted summation as a risk aggregation rule if  $k_j = \varrho(X_j)$  for elliptical  $X$  and  $L(X) = m + \lambda' X$ .

### Proposition 8.26 (Justification for correlation adjusted summation)

Let  $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \psi)$  with  $\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu}$ . Let  $\mathcal{M}$  be the space of linear portfolios (48) and  $\varrho$  be a pos. hom., translation- and law-invariant risk measure on  $\mathcal{M}$ . Then, for any  $L = L(\mathbf{X}) = m + \boldsymbol{\lambda}'\mathbf{X} \in \mathcal{M}$ ,

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

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

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

- 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  $\rho_{ij}$ 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  $\text{EC}_j = \varrho(L_j) - \mathbb{E}(L_j)$ ). In the *margins-plus-copula approach*, we could attempt to choose a suitable copula  $C$  for  $\mathbf{L} \sim F(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d))$ ; 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  $\mathbf{L}$  is typically sparse.

In the next section, we address the case where  $C$  is unknown and quantify the “best-to-worst” gaps/ranges in VaR and ES estimation.

- 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  $\mathbf{X} = (X_1, \dots, X_k)$  (more data

exists) and for the functionals  $g_j : \mathbb{R}^k \mapsto \mathbb{R}$  which give the losses  $L_j = g_j(\mathbf{X})$ ,  $j \in \{1, \dots, d\}$ , for the different portfolios/business units. Risk measures are then derived from the distribution of  $L = g_1(\mathbf{X}) + \dots + g_d(\mathbf{X})$ .



## 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  $\text{VaR}_\alpha$  and  $\text{ES}_\alpha$  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

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

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

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

## Summary of existing results

$d = 2$ : Fully solved analytically

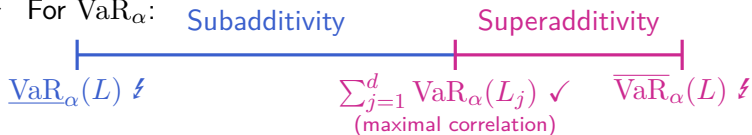
$d \geq 3$ : Here we distinguish:

- ▶ **Homogeneous case** ( $F_1 = \dots = F_d$ ):
  - $\underline{\text{ES}}_\alpha(L)$  solved analytically for decreasing densities (e.g. Pareto, Exponential)
  - $\underline{\text{VaR}}_\alpha(L)$ ,  $\overline{\text{VaR}}_\alpha(L)$  solved analytically for tail-decreasing densities (e.g. Pareto, Log-normal, Gamma)
- ▶ **Inhomogeneous case:**
  - Few **analytical** results: current research
  - **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)  $F_1, \dots, F_d$ . The copula  $C$  is **unknown**.
- We need to find the **dependence uncertainty gaps** ( $\underline{\text{VaR}}_\alpha(L), \overline{\text{VaR}}_\alpha(L)$ ) (or ( $\underline{\text{ES}}_\alpha(L), \overline{\text{ES}}_\alpha(L)$ )) for  $L = L_1 + \dots + L_d$ .
- The **dependence uncertainty spreads** can be visualized as follows.

- ▶ For  $\text{VaR}_\alpha$ :



- ▶ For  $\text{ES}_\alpha$ :



The Rearrangement Algorithm (RA) can find approximate solutions to the " $\nlessgtr$ " cases.

### Proposition 8.27 ( $\text{VaR}_\alpha$ in the homogeneous case)

Let  $F := F_1 = \dots = F_d$  with decreasing density on  $[b, \infty)$ . Then, for  $\alpha \in [F(b), 1)$  and  $X \sim F$ ,

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

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

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

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

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

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

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

**Proposition 8.28 ( $\underline{\text{ES}}_\alpha$  in the homogeneous case)**

Let  $F := F_1 = \dots = F_d$  with finite first moment and decreasing density on its support. Then, for  $\alpha \in [1 - dc, 1)$ ,  $\beta = (1 - \alpha)/d$ , and  $X \sim F$ ,

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

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

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

and  $\text{LES}_\alpha(X) = \frac{1}{\alpha} \int_0^\alpha \text{VaR}_u(X) du = -\text{ES}_{1-\alpha}(-X)$  (*lower ES*).

*Proof.* See Bernard et al. (2014). □

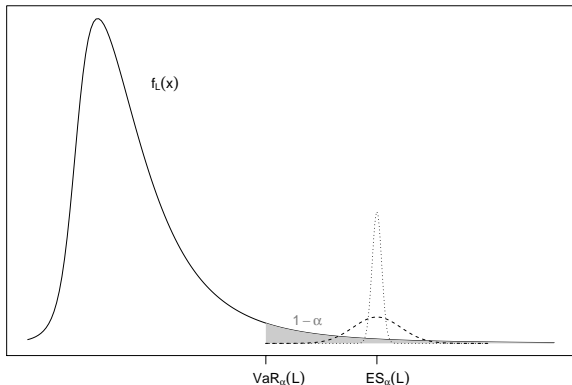
## The Rearrangement Algorithm (RA)

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

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

- 1) Fix  $\alpha \in (0, 1)$ ,  $F_1^\leftarrow, \dots, F_d^\leftarrow$ ,  $N \in \mathbb{N}$  (# of discr. points),  $\varepsilon \geq 0$  (tol.)
- 2) Compute the lower bound  $\underline{s}_N$ :
  - 2.1) Define the  $(N, d)$ -matrix  $\underline{X}^\alpha = \left( F_j^\leftarrow \left( \alpha + \frac{(1-\alpha)(i-1)}{N} \right) \right)_{i,j}$ .
  - 2.2) Permute randomly each column of  $\underline{X}^\alpha$  (to avoid  $\bar{s}_N - \underline{s}_N \nrightarrow 0$ )
  - 2.3) Set  $\underline{Y}^\alpha = \underline{X}^\alpha$ . For  $1 \leq j \leq d$ , rearrange the  $j$ th column of  $\underline{Y}^\alpha$  so that it becomes oppositely ordered to the sum of all others.
  - 2.4) While  $s(\underline{Y}^\alpha) - s(\underline{X}^\alpha) > \varepsilon$ , set  $\underline{X}^\alpha$  to  $\underline{Y}^\alpha$  and repeat Step 2.3).
  - 2.5) Set  $\underline{s}_N = s(\underline{Y}^\alpha)$ .
- 3) Similarly, compute  $\bar{s}_N = s(\bar{Y}^\alpha)$  based on  $\bar{X}^\alpha = \left( F_j^\leftarrow \left( \alpha + \frac{(1-\alpha)i}{N} \right) \right)_{i,j}$ .
- 4) Return  $(\underline{s}_N, \bar{s}_N)$  (*rearrangement range*; taken as bounds on  $\overline{\text{VaR}}_\alpha(L)$ )

**Intuition:** The RA is based on the idea of joint mixability. The dfs  $F_1, \dots, F_d$  are *jointly mixable* if there exists a  $c \in \mathbb{R}$  such that  $\mathbb{P}(L_1 + \dots + L_d = c) = 1$  where  $L_1 \sim F_1, \dots, L_d \sim F_d$ . This is a notion of negative dependence and can be illustrated as follows.



$\Rightarrow$  Minimizing the variance of  $L | L > F_L^{\leftarrow}(\alpha)$  increases  $\text{VaR}_\alpha(L)$ .

### Example 8.30 (How the RA works)

1) Where it works (to compute the optimum of the maximin problem):

$$\begin{array}{ccc}
 \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 5 & 4 \\ 4 & 7 & 8 \end{pmatrix} & \xRightarrow{\Sigma_{-1} = \begin{pmatrix} 2 \\ 5 \\ 9 \\ 15 \end{pmatrix}} & \begin{pmatrix} 4 & 1 & 1 \\ 3 & 3 & 2 \\ 2 & 5 & 4 \\ 1 & 7 & 8 \end{pmatrix} \xRightarrow{\text{stable sort}} \begin{pmatrix} 4 & 5 & 1 \\ 3 & 7 & 2 \\ 2 & 3 & 4 \\ 1 & 1 & 8 \end{pmatrix} \xRightarrow{\Sigma_{-3} = \begin{pmatrix} 9 \\ 10 \\ 5 \\ 2 \end{pmatrix}} \\
 \begin{pmatrix} 4 & 5 & 2 \\ 3 & 7 & 1 \\ 2 & 3 & 4 \\ 1 & 1 & 8 \end{pmatrix} & \xRightarrow{\Sigma_{-1} = \begin{pmatrix} 7 \\ 8 \\ 7 \\ 9 \end{pmatrix}} & \begin{pmatrix} 3 & 5 & 2 \\ 2 & 7 & 1 \\ 4 & 3 & 4 \\ 1 & 1 & 8 \end{pmatrix} \checkmark \xRightarrow{\Sigma = \begin{pmatrix} 10 \\ 10 \\ 11 \\ 10 \end{pmatrix}} \widehat{\text{VaR}}_{\alpha}(L) \approx 10
 \end{array}$$

2) Where it fails (to compute the optimum of the maximin problem):

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} \xRightarrow{\Sigma_{-1} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}} \begin{pmatrix} 3 & 1 & 1 \\ 2 & 2 & 2 \\ 1 & 3 & 3 \end{pmatrix} \checkmark \xRightarrow{\Sigma = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}} \widehat{\text{VaR}}_{\alpha}(L) \approx 5 < 6$$



### Example 8.31 (Par(2) margins)

Let  $L_j \sim \text{Par}(2)$  with  $\bar{F}_j(x) = (1+x)^{-2}$ ,  $j \in \{1, \dots, d\}$  (homogeneous case) and  $\alpha = 0.999$ . One obtains:

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

- The “+” and “ $\perp$ ” denote the comonotonic and independent case, resp.
- $\frac{\overline{\text{ES}}_\alpha(L)}{\overline{\text{VaR}}_\alpha(L)} \approx 1$  can be explained; see McNeil et al. (2015, Prop. 8.36).
- The dependence uncertainty spread  $\overline{\text{VaR}}_\alpha(L) - \underline{\text{VaR}}_\alpha(L) \geq \overline{\text{ES}}_\alpha(L) - \underline{\text{ES}}_\alpha(L)$  can be explained; see McNeil et al. (2015, Prop. 8.37).

## Remark 8.32

- The RA finds approximate solutions to *maximin* (for  $\overline{\text{VaR}}_\alpha(L)$ ) and *minimax* (for  $\underline{\text{VaR}}_\alpha(L)$ ) *problems* and is thus of wider interest (e.g., in Operations Research).
- For  $\underline{\text{ES}}_\alpha(L)$ , discretize the whole support of each margin, rearrange, and approximate  $\underline{\text{ES}}_\alpha(L)$  by the nonparametric  $\text{ES}_\alpha$  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  $\bar{s}_N$  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  $\overline{\text{VaR}}_\alpha$ ,  $\overline{\text{ES}}_\alpha$ )**

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  $\alpha \in (0, 1)$ ,  $\liminf_{d \rightarrow \infty} \frac{1}{d} \sum_{j=1}^d \text{ES}_\alpha(L_j) > 0$ .

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

**Proposition 8.34 (Dependence uncertainty spread of  $\text{VaR}_\alpha$  vs  $\text{ES}_\alpha$ )**

Let  $0 < \alpha_1 \leq \alpha_2 < 1$ , assume Proposition 8.33 i) to hold and that

$\liminf_{d \rightarrow \infty} \frac{1}{d} \sum_{j=1}^d \text{LES}_{\alpha_1}(X_j) > 0$  and  $\limsup_{d \rightarrow \infty} \frac{\sum_{j=1}^d \mathbb{E}(X_j)}{\sum_{j=1}^d \text{ES}_{\alpha_1}(X_j)} < 1$ . Then

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

## 8.5 Capital allocation

We now study how 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  $\varrho(L) = \varrho(L_1 + \dots + L_d)$ . Then allocate  $\varrho(L)$  to the investments according to a capital allocation principle such that

$$\varrho(L) = \sum_{j=1}^d AC_j,$$

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

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

- For a risk measure  $\varrho$ , 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  $r_{\varrho}$  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  $r_{\varrho}$  is positive homogeneous if  $\varrho$  is.

### Definition 8.35 (Euler capital allocation principle)

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

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

## Examples

### 1) Standard deviation and the covariance principle

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

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

This formula is known as *covariance principle*.

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

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

and hence

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

## 2) VaR and VaR contributions

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

$$\text{AC}_j^{\varrho} = \frac{\partial r_{\text{VaR}}^{\alpha}}{\partial \lambda_j}(\mathbf{1}) = \mathbb{E}(L_j \mid L = F_L^{\leftarrow}(\alpha)), \quad j \in \{1, \dots, d\}.$$

## 3) Expected shortfall and shortfall contributions

Now consider  $r_{\text{ES}}^{\alpha}(\boldsymbol{\lambda}) = \mathbb{E}(L \mid L \geq q_{\alpha}(L(\boldsymbol{\lambda})))$ . Then

$$r_{\text{ES}}^{\alpha}(\boldsymbol{\lambda}) = \frac{1}{1 - \alpha} \int_{\alpha}^1 r_{\text{VaR}}^u(\boldsymbol{\lambda}) \, du,$$

Assuming the differentiability of  $r_{\text{VaR}}^u(\lambda)$ , the Euler principle implies that

$$\frac{\partial r_{\text{ES}}^\alpha}{\partial \lambda_j}(\mathbf{1}) = \frac{1}{1-\alpha} \int_\alpha^1 \frac{\partial r_{\text{VaR}}^u}{\partial \lambda_j}(\mathbf{1}) du = \frac{1}{1-\alpha} \int_\alpha^1 \mathbb{E}(L_j \mid L = q_u(L)) du.$$

If  $F_L$  has a differentiable inverse,

$$\frac{\partial r_{\text{ES}}^\alpha}{\partial \lambda_j}(\mathbf{1}) = \frac{1}{1-\alpha} \int_{F_L^\leftarrow(\alpha)}^\infty \mathbb{E}(L_j \mid L = v) f_L(v) dv = \frac{\mathbb{E}(L_j; L \geq F_L^\leftarrow(\alpha))}{1-\alpha}.$$

Hence the Euler capital allocation takes the form

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

$\text{AC}_j^e$  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  $\mathbf{L} \sim E_d(\mathbf{0}, \Sigma, \psi)$ : Calculate the total risk capital and then use a simple partitioning formula (regardless of the pos.-hom. risk measure).



### Corollary 8.36 (Euler allocation under ellipticality)

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

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

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

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

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

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

from which the result follows. □

### 8.5.3 Economic properties of the Euler principle

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

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

denote the associated risk contributions under the Euler principle.

#### Compatibility with a RORAC approach

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

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

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

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

- 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  $0 < h \leq \varepsilon$

$$\text{RORAC}(L_j | L) > \text{RORAC}(L) \Rightarrow \text{RORAC}(L + hL_j) > \text{RORAC}(L).$$

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

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

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

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

## Diversification benefit

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

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

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

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

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

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

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

For  $\mathbf{y} = \boldsymbol{\lambda}$  and  $\mathbf{x} = \boldsymbol{\lambda} + \tilde{\boldsymbol{\lambda}}$ , 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  $\varrho$  is pos.-hom. and subadditive),  $\boldsymbol{\lambda} = \mathbf{e}_j$  and  $\tilde{\boldsymbol{\lambda}} = \mathbf{1} - \mathbf{e}_j$  to obtain

$$\varrho(L_j) = r_{\varrho}(\mathbf{e}_j) \geq \frac{\partial r_{\varrho}}{\partial \lambda_j}(\mathbf{1}) = \text{AC}_j^{\varrho} \quad \square$$

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