Hsing, Klüppelberg, Kuhn:

Modelling, Estimation and Visualization of Multivariate Dependence for Risk Management

Sonderforschungsbereich 386, Paper 375 (2004)

Online unter: http://epub.ub.uni-muenchen.de/

Projektpartner

Modelling, Estimation and Visualization of Multivariate Dependence for Risk Management

Tailen Hsing Claudia Klüppelberg* Gabriel Kuhn

C. Klüppelberg & G. Kuhn
Center of Mathematical Sciences
Munich University of Technology
D-85747 Garching, Germany
email: {cklu , gabriel}@ma.tum.de
phone: +49 89 289 17432
http://www.ma.tum.de/stat/

T. Hsing
Department of Statistics
Texas A&M University
College Station TX 77843, USA
email: thsing@stat.tamu.edu

*Author for correspondence.

Abstract

Dependence modelling and estimation is a key issue in the assessment of portfolio risk. When measuring extreme risk in terms of the Value-at-Risk, the multivariate normal model with linear correlation as its natural dependence measure is by no means an ideal model. We suggest a large class of models and a new dependence function which allows us to capture the complete extreme dependence structure of a portfolio. We also present a simple nonparametric estimation procedure. To show our new method at work we apply it to a financial data set of zero coupon swap rates and estimate the extreme dependence in the data.

JEL Classifications: C15, C52.

Keywords: Risk management, extreme risk assessment, multivariate models, dependence function.

1 Risk management for extreme risk

According to the Capital Adequacy Directive of the Basel Committee the risk capital of a bank must be sufficient to cover losses on the bank's trading portfolio over a 10-day holding period with a probability of 99%, a measure usually referred to as Value-at-Risk (VaR).

Estimation of a portfolio VaR faces two different problems: firstly, a time series structure of the portfolio value may affect the estimation of the VaR over a 10-day holding period, and, secondly, the underlying portfolio may consist of a large number of different instruments, whose multivariate dependence structure can introduce serious errors, when not modelled correctly.

Time series or serial dependence is not an issue of the present paper and we refer to Klüppelberg (2004) and McNeil and Frey (2000) for discussions on this important aspect in VaR modelling and estimation.

We are interested here in the influence of the multivariate dependence within the portfolio. We first recall that under the condition that the portfolio P/L follows a multivariate normal distribution and if there is no serial dependence, the $\mathrm{VaR}(\alpha, T)$ that corresponds to the α-quantile and holding period T (the Basel Committee requires $\alpha = 0.01$ and $T = 10$ days) is given by

$$\mathrm{VaR}(\alpha, T) = z_\alpha \sigma \sqrt{T}\,,$$

where z_α is the α-quantile of the standard normal distribution. The portfolio P/L standard deviation σ is calculated by the square root of its variance

$$\sigma^2 = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i \neq j} w_i w_j \sigma_i \sigma_j \rho_{ij}\,, \tag{1.1}$$

where the portfolio consists of n different instruments with nominal amount w_i invested into asset i. The standard deviation of asset i is given by σ_i and the pairwise correlation coefficients are ρ_{ij} $(i, j = 1, \ldots, n)$.

Definition 1.1 *For two random variables X and Y their* correlation *is defined as*

$$\rho(X, Y) = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}(X)\,\mathrm{var}(Y)}}\,,$$

where $\mathrm{cov}(X, Y) = E((X - EX)(Y - EY))$ *is the covariance of X and Y, and* $\mathrm{var}(X)$ *and* $\mathrm{var}(Y)$ *are the variances of X and Y, respectively.*

Correlation measures linear dependence: we have $|\rho(X, Y)| = 1$ if and only if $Y = aX + b$ with probability 1 for $a \in \mathbb{R} \setminus \{0\}$ and $b \in \mathbb{R}$. Furthermore, correlation is invariant

under strictly increasing *linear* transformations; i.e. for $\alpha, \gamma \in \mathbb{R} \setminus \{0\}$ and $\beta, \delta \in \mathbb{R}$

$$\rho(\alpha X + \beta, \gamma Y + \delta) = \text{sign}(\alpha\gamma)\, \rho(X, Y)\,.$$

Also for high-dimensional models correlation is easy to handle. For random (column) vectors $X, Y \in \mathbb{R}^n$ we denote by $\text{cov}(X, Y) = E((X - EX)(Y - EY)^T)$ the covariance matrix of X and Y. Then for $m \times n$ matrices A, B and vectors $a, b \in \mathbb{R}^m$ we calculate

$$\text{cov}(AX + a, BY + b) = A\,\text{cov}(X, Y)\, B^T\,,$$

where B^T denotes the transpose of the matrix B. From this it follows for $w \in \mathbb{R}^n$

$$\text{var}(w^T X) = w^T \text{cov}(X, X) w\,,$$

which is exactly formula (1.1) above. The popularity of correlation is also based on the fact that it is very easy to calculate and estimate. It is a natural dependence measure for elliptical distributions such as the multivariate normal or t distributions. Multivariate portfolios, however, are quite often not elliptically distributed, and there may be a more complex dependence structure than linear dependence. Indeed, data may be uncorrelated, i.e. with correlation 0, but still may be highly dependent. In the context of risk management, when measuring extreme risk as the Value-at-Risk does, measuring dependence by correlation may be grossly misleading. For a very readable article about different dependence measures, their limitations and pitfalls we refer to Embrechts, McNeil and Straumann (2002).

We turn to a measure for tail dependence, which relates large values of the components of a portfolio. In the bivariate context, consider random variables X and Y with marginal distribution functions G_X and G_Y and (generalized) inverses $G_X^{\leftarrow}$ and $G_Y^{\leftarrow}$. For

any distribution function G its *generalized inverse or quantile function* is defined as

$$G^{\leftarrow}(t) = \inf\{x \in \mathbb{R} \mid G(x) \geq t\}\,, \quad 0 < t < 1\,.$$

If G is strictly increasing, then $G^{\leftarrow}$ coincides with the usual inverse of G.

Definition 1.2 *The* upper tail dependence coefficient *of* (X, Y) *is defined by*

$$\rho_U = \lim_{u \uparrow 1} P(Y > G_X^{\leftarrow}(u) \mid X > G_Y^{\leftarrow}(u))\,, \tag{1.2}$$

provided the limit exists. If $\rho_U \in (0, 1]$*, then* X *and* Y *are called* asymptotically dependent in the upper tail*, if* $\rho_U = 0$*, they are called* asymptotically upper tail independent*.*

For some situations, this measure may be an appropriate extreme dependence measure; this is true, in particular, when the bivariate distribution is symmetric; see Example 2.8. However, ρ_U is not a very informative measure for asymmetric models since the extreme dependence around the 45 degree line does not reveal much about what happens elsewhere; see Example 2.9. As a remedy we suggest an extension of the upper tail dependence coefficient to a function of the angle, which measures extreme dependence in any direction in the first quadrant of $\mathbb{R}^2$. Its derivation is based on multivariate extreme value theory and we indicate this relationship in Section 2. We shall, however, refrain ourselves from a precise derivation and rather refer to Hsing, Klüppelberg and Kuhn (2003) for details. We also want to emphasize that one-dimensional extreme value theory has been applied successfully to risk management problems; see Embrechts (2000). We remark further that one-dimensional extreme value theory has meanwhile reached a consolidated state; we refer to Embrechts, Klüppelberg and Mikosch (1997) or Coles (2001) as standard references.

We shall illustrate our results by a direct application to a real data set. The complete data set we investigated consists of returns (daily differences) of *Annually Compounded*

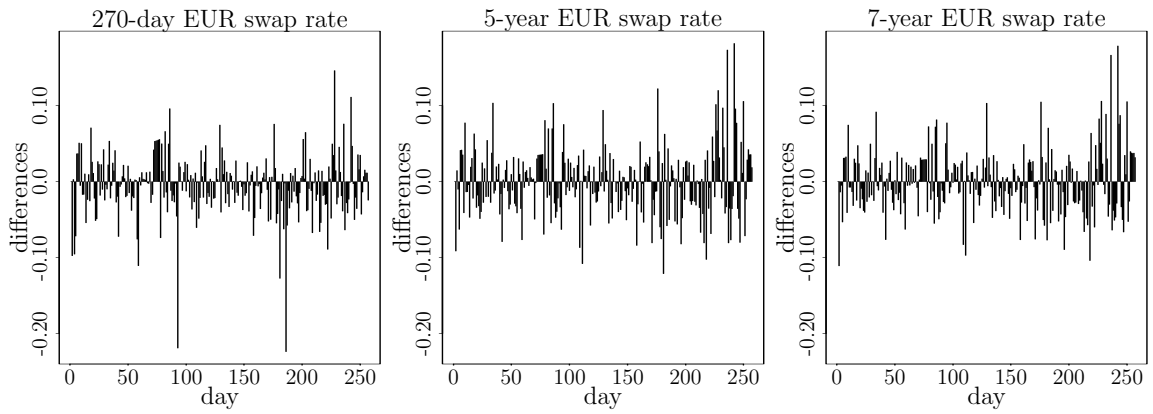

Figure 1.3 Example of data: time series of swap rates with maturities 270 days, 5 years and 7 years.

Zero Coupon Swap Rates with different maturities (between 7 days and 30 years) and different currencies (EUR, USD and GBP).

Each of the time series consists of 257 daily returns during the year 2001. In an exploratory data analysis we investigated first each single time series; typical examples are plotted in Figure 1.3. Plots of the autocorrelation functions of the single time series, their moduli and squares exhibited no significant serial dependence structure; hence we assume the data being iid. Moreover, the histograms and a tail analysis showed that the marginals are well modelled (at least in the tails) by a two-sided exponential distribution.

In this paper we illustrate our methods and models on swap rates in EUR currency only. The first column of Figure 1.4 shows scatter plots of different combinations of swap rates. We have estimated the mean, variance and correlation of the data in each scatter plot. The second column shows simulated normal data with the same estimated parameters. Recall that for risk assessment we are particularly interested in the left lower corner. None of the normal models seems to be able to capture the dependence structure in this area.

After introducing a new dependence function in Section 2 we shall present some ex-

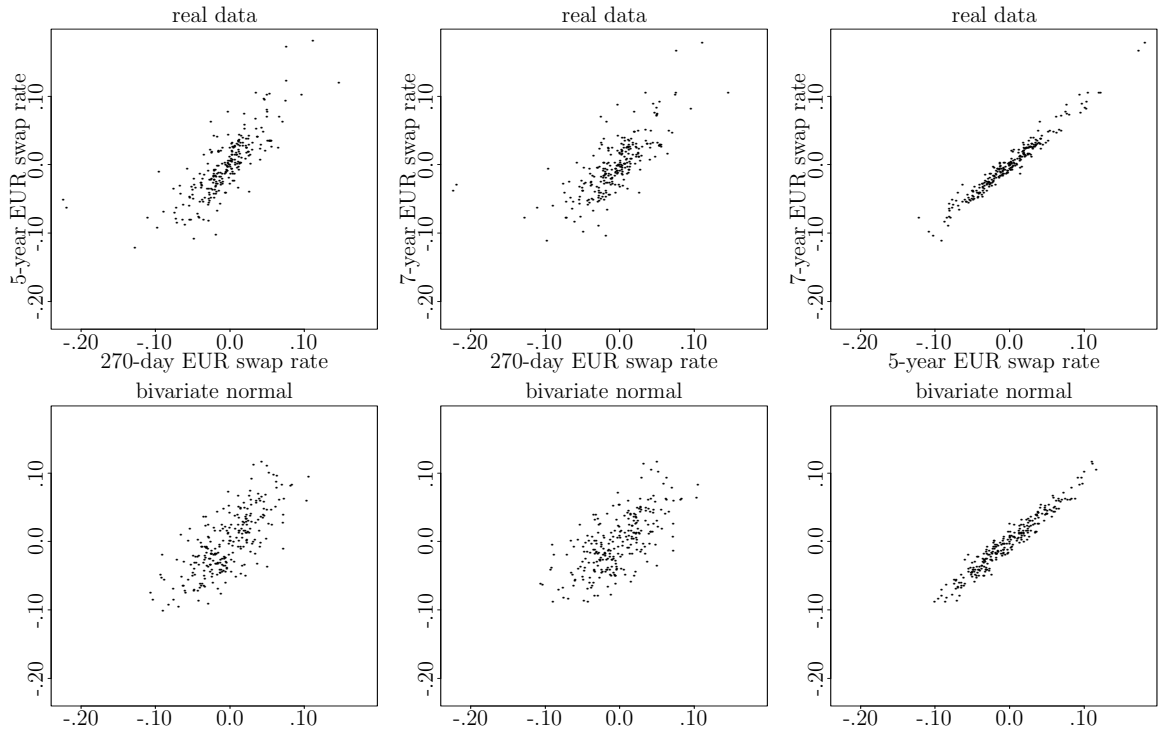

Figure 1.4 Upper row: Scatter plot of data of Figure 1.3, 270-day vs. 5-year (left plot), 270-day vs. 7-year (middle plot) and 5-year vs. 7-year EUR swap (right plot).

Lower row: Scatter plot of simulated 3-variate normal distribution with mean and variance of the data in the upper row.

amples including the bivariate normal, where we show that it exhibits no dependence in the far-out tails, no matter in which direction we go, thus extending the well-known result along the diagonal.

In Section 3 we introduce a simple nonparametric estimation procedure of the dependence function. We show its performance in various simulation examples and plots. Finally, we investigate some of our swap rate data in more detail and estimate the dependence function. We also show various plots to visualise our results.

Our findings in the whole data set with all maturities and all the different currencies

can be summarized as follows: for swap rates in the same currency a high dependence for similar maturities, and a low dependence between very different maturities. Between different currencies we observed only very little dependence except for similarly long maturities, where we detected some moderate dependence. For plots and more details on these effects we refer to Kuhn (2002).

2 Measuring extreme dependence

Although the upper tail dependence coefficient and its functional extension we are aiming at can be defined for random vectors of any dimension, we restrict ourselves in our presentation to the bivariate case. For a general treatment in any dimension we refer to Hsing et al. (2003).

Suppose $(X_i, Y_i)_{i=1,\dots,n}$ is a sequence of iid vectors and (X, Y) is a generic random vector with the same distribution function $G(x, y) = P(X \leq x, Y \leq y)$ for $(x, y) \in \mathbb{R}^2$ with continuous marginals. For $n \in \mathbb{N}$ define the vector of componentwise maxima

$$M_n = (\max_{i=1,\dots,n} X_i \,,\, \max_{i=1,\dots,n} Y_i) \,.$$

As a first goal we want to describe the behaviour of M_n for large n.

It is a standard approach in extreme value theory to first transform the marginals to some appropriate common distribution and then model the dependence structure separately. As copulas have become a fairly standard notion for modelling dependence we follow this approach and transform the marginal distributions G_X and G_Y to uniform (0,1). Then we have a bivariate uniform distribution, which is called a *copula* and is given for $0 < u, v < 1$ by

$$C_G(u, v) = P\left(G_X(X) \leq u \,,\, G_Y(Y) \leq v\right) = P\left(X \leq G_X^{\leftarrow}(u) \,,\, Y \leq G_Y^{\leftarrow}(v)\right) \,.$$

For more details on copulas and dependence structures in general we refer to Joe (1997); for applications of copulas in risk management see Embrechts, Lindskog and McNeil (2001). The transformation of the marginals to uniforms is illustrated in Figure 2.1.

Under weak regularity conditions on $G(x, y)$ we obtain

$$\lim_{n\to\infty} P\left(\max_{i=1,\dots,n} G_X(X_i) \le 1 + \frac{1}{n}\ln u\,,\ \max_{i=1,\dots,n} G_Y(Y_i) \le 1 + \frac{1}{n}\ln v\right)$$
$$= \exp(-\Lambda(-\ln u, -\ln v))) \;=\; C(u, v)\,.$$

Such a copula is called *extreme copula* and satisfies for all $t > 0$

$$C^t(u, v) = C(u^t, v^t)\,, \quad 0 < u, v < 1\,.$$

$C(u, v)$ has various integral representations. The *Pickands' representation* yields an extreme event intensity measure (we write $a \wedge b = \min(a, b)$ and $a \vee b = \max(a, b)$):

$$\begin{aligned} \Lambda(x, y) &= \lim_{n\to\infty} nP\left(G_X(X) > 1 - \frac{x}{n} \text{ or } G_Y(Y) > 1 - \frac{y}{n}\right) \\ &= \int_0^{\pi/2} \left(\frac{x}{1 \vee \cot\theta} \vee \frac{y}{1 \vee \tan\theta}\right) \Phi(d\theta)\,. \end{aligned} \tag{2.1}$$

Φ is a finite measure on $(0, \pi/2)$ satisfying

$$\int_0^{\pi/2} (1 \wedge \tan\theta)\, \Phi(d\theta) = \int_0^{\pi/2} (1 \wedge \cot\theta)\, \Phi(d\theta) = 1\,.$$

The definition of Λ as a limit of $n \times$ *success probability* is a version of the classical limit theorem of Poisson. For large n the measure Λ can be interpreted as the mean number of data in a strip near the upper and right boundary of the uniform distribution; see Figure 2.1. We also recall some properties of $\tan\theta = \frac{1}{\cot\theta} = \frac{\sin\theta}{\cos\theta}$: $\tan 0 = 0$, $\tan\theta$ is increasing in $\theta \in (0, \pi/2)$ and $\lim_{\theta\to\pi/2} \tan\theta = \infty$. Then $\cot\theta$ is its reflection on the 45 degree line, corresponding to $\theta = \pi/4$. Moreover, $\tan\pi/4 = \cot\pi/4 = 1$. Finally, arctan is the inverse function of tan.

The fact that $\Lambda(x, y) = x\Lambda(1, y/x)$ motivates the following definition.

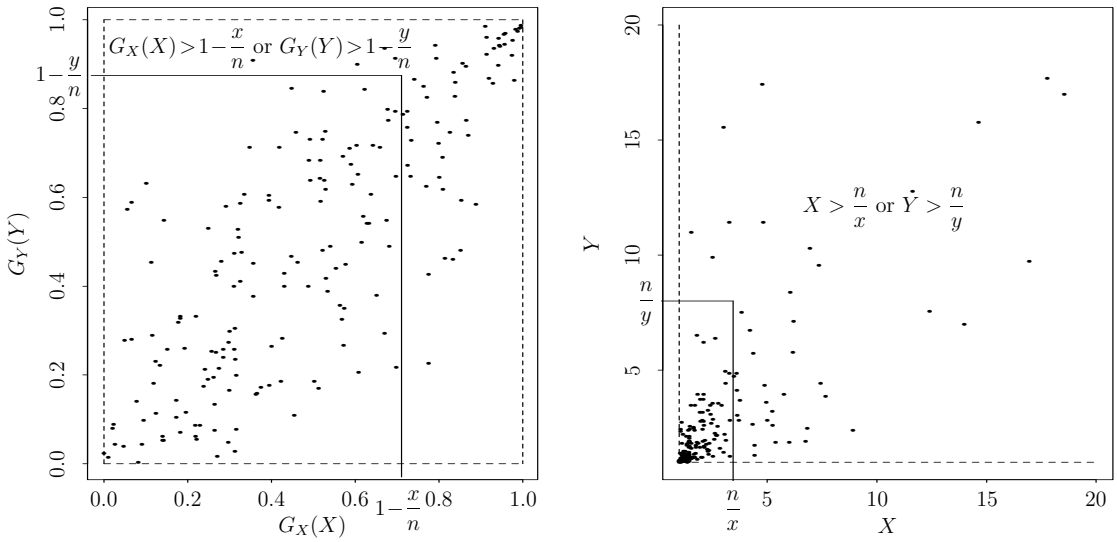

Figure 2.1 Left plot: Illustration of the intensity measure Λ as defined in equation (2.1) by simulated data for X and Y Fréchet distributed with distribution functions $G_X(x) = G_Y(x) = \exp(-1/x)$ for $x > 0$. Then Λ measures the strip near the upper and right boundary of the uniform distribution on the left. Right plot: In the original data this corresponds to the indicated region of large data points.

Definition 2.2 *For any random vector (X, Y) such that (2.1) holds we define the* dependence function *as*

$$\psi(\theta) = \Lambda(1, \cot\theta)\,, \quad 0 < \theta < \pi/2\,.$$

Note that $\psi(\cdot)$ is a function of the angle θ only and measures dependence in any direction of the positive quadrant of a bivariate distribution. The following result shows that $\psi(\cdot)$ allows us to approximate for large x_1 and y_1 the probability for X or Y to become large. We write $a(x) \sim b(x)$ as $x \to x_0$ for $\lim_{x\to x_0} a(x)/b(x) = 1$.

Proposition 2.3 *Let (X, Y) be a random vector. If $x_1, y_1 \to \infty$ such that $P(X > x_1)/P(Y > y_1) \to \tan\theta$, the following quotient converges for all $\theta \in (0, \pi/2)$,*

$$\frac{P(X > x_1 \text{ or } Y > y_1)}{P(X > x_1)}\,,$$

then the limit defines the dependence function $\psi(\theta)$.

Proof. From (2.1) we have for large x_1, y_1 and $x = n\overline{G}_X(x_1)$ and $y = n\overline{G}_Y(y_1)$ as $n \to \infty$ (note that x, x_1, y, y_1 depend on n):

$$\begin{aligned} P(X > x_1 \text{ or } Y > y_1) &\sim \frac{1}{n}\Lambda(n\overline{G}_X(x_1), n\overline{G}_Y(y_1)) \\ = \quad \overline{G}_X(x_1)\Lambda\left(1, \frac{\overline{G}_Y(y_1)}{\overline{G}_X(x_1)}\right) &= \overline{G}_X(x_1)\psi\left(\arctan\left(\frac{\overline{G}_X(x_1)}{\overline{G}_Y(y_1)}\right)\right). \end{aligned} \tag{2.2}$$

We set

$$\theta = \arctan\left(\frac{\overline{G}_X(x_1)}{\overline{G}_Y(y_1)}\right)$$

and obtain the result. □

Corollary 2.4 *(a) For X and Y independent we calculate*

$$\frac{P(X > x_1 \text{ or } Y > y_1)}{P(X > x_1)} \sim \frac{P(X > x_1) + P(Y > y_1)}{P(X > x_1)} \to 1 + \cot\theta =: \psi_0(\theta)$$

for $x_1, y_1 \to \infty$ such that $P(Y > y_1)/P(X > x_1) \to \cot\theta$.

(b) For X and Y completely dependent, i.e. $X = g(Y)$ with probability 1 for some increasing function g, we obtain

$$\frac{P(X > x_1 \text{ or } Y > y_1)}{P(X > x_1)} = \frac{P(X > x_1) \vee P(X > y_1)}{P(X > x_1)} \to 1 \vee \cot\theta =: \psi_1(\theta)$$

for $x_1, y_1 \to \infty$ such that $P(Y > y_1)/P(X > x_1) \to \cot\theta$.

(c) Furthermore,

$$\psi_1(\theta) \;\le\; \psi(\theta) \;\le\; \psi_0(\theta)\,, \quad 0 < \theta < \pi/2\,.$$

We normalize $\psi(\cdot)$ to the interval $[0, 1]$ as follows.

Definition 2.5 *The normalised function*

$$\rho(\theta) = \frac{\psi_0(\theta) - \psi(\theta)}{\psi_0(\theta) - \psi_1(\theta)}\,, \quad 0 < \theta < \pi/2\,,$$

we call tail dependence function.

Note that it describes the tail dependence of (X, Y) in any direction of the bivariate distribution on the positive quadrant of $\mathbb{R}^2$.

By this definition we have $\rho(\theta) \in [0, 1]$ for all $0 < \theta < \pi/2$, $\rho(\theta) \equiv 0$ in case of independence and $\rho(\theta) \equiv 1$ in case of complete dependence. Consequently, $\rho(\theta)$ being close to 0/1 corresponds to weak/strong extreme dependence.

The function $\rho(\cdot)$ is invariant under monotone transformation of the marginal distributions. We show this by calculating it as a function of the copula.

Proposition 2.6 *Let (X, Y) be a random vector with marginal distribution functions G_X and G_Y, which are continuous functions. Then $G_X(X) \stackrel{\mathrm{d}}{=} U$ and $G_Y(Y) \stackrel{\mathrm{d}}{=} V$ for uniform random variables U and V with the same dependence structure as (X, Y). Denote by $C(u, v) = P(U \leq u\,, V \leq v)$ the corresponding copula. We also relate the arguments by $G_X(x_1) = u$ and $G_Y(y_1) = v$. Then, provided that the limits exist,*

$$\rho(\theta) \;=\; \lim_{\substack{u,v\to 1 \\ (1-u)/(1-v)\to\tan\theta}} \frac{1-u-v+C(u,v)}{(1-u)\wedge(1-v)}\,, \quad 0<\theta<\pi/2\,.$$

Proof.

$$\begin{aligned}\psi(\theta) \;&=\; \lim_{\substack{x_1,y_1\to\infty \\ \overline{G}_X(x_1)/\overline{G}_Y(y_1)\to\tan\theta}} \frac{1-P(X\leq x_1\,,\,Y\leq y_1)}{P(X>x_1)} \\ &=\; \lim_{\substack{u,v\to 1 \\ (1-u)/(1-v)\to\tan\theta}} \frac{1-C(u,v)}{1-u}\,.\end{aligned}$$

□

Remark 2.7 Note also that the quantity $\rho(\pi/4)$ is nothing but the *(upper) tail dependence coefficient* ρ_U as defined in (1.2). Thus, the function ρ extends this notion from a single direction, the 45 degree line corresponding to $\theta = \pi/4$, to all directions in $(0, \pi/2)$.

This extension is illustrated by the following examples.

Example 2.8 [Gumbel copula]

Let (X, Y) be a bivariate random vector with dependence structure given by a *Gumbel* copula

$$C(u,v) = \exp\left\{-\left[(-\ln u)^\delta + (-\ln v)^\delta\right]^{1/\delta}\right\}, \quad \delta \in [1,\infty)\,. \tag{2.3}$$

The dependence arises from δ. To calculate $\psi(\theta)$ we use the relationship of ψ to its copula. We use also the fact that for $u, v \to 1$ we have

$$\frac{-\ln v}{-\ln u} \sim \frac{1-v}{1-u} \to \cot\theta\,.$$

Then by continuity of u^x in x we obtain for $u, v \to 1$ such that $(1-v)/(1-u) \to \cot\theta$

$$1 - C(u,v) \;=\; 1 - \exp\left(\ln u\left[1 + \left(\frac{-\ln v}{-\ln u}\right)^\delta\right]^{1/\delta}\right) \;\sim\; 1 - u^{(1+(\cot\theta)^\delta)^{1/\delta}}\,.$$

Using the l'Hospital rule and the fact that $u \to 1$, we obtain

$$\frac{1-C(u,v)}{1-u} \to \left(1 + (\cot\theta)^\delta\right)^{1/\delta}\,,$$

and hence

$$\rho(\theta) = \frac{1 + \cot\theta - \left(1 + (\cot\theta)^\delta\right)^{1/\delta}}{1 \wedge \cot\theta}\,, \quad 0 < \theta < \pi/2\,.$$

We also obtain the well-known upper tail dependence coefficient $\rho_U = \rho(\pi/4) = 2 - 2^{1/\delta}$.

□

Our next example is a typical model to capture risk in the extremes. We write again $a \wedge b = \min(a,b)$ and $a \vee b = \max(a,b)$.

Example 2.9 [Asymmetric Pareto model]

For $p_1, p_2 \in (0,1)$ set $\overline{p}_1 = 1 - p_1$ and $\overline{p}_2 = 1 - p_2$ and consider the model

$$X = p_1 Z_1 \vee \overline{p}_1 Z_2 \quad \text{and} \quad Y = p_2 Z_1 \vee \overline{p}_2 Z_3$$

with Z_1, Z_2, Z_3 iid Pareto(1) distributed; i.e., $P(Z_i > x) = x^{-1}$ for $x \geq 1$. Clearly, the dependence between X and Y arises from the common component Z_1. Hence the dependence is stronger for larger values of p_1, p_2. We calculate the function ρ, and observe first that by independence of the Z_i for $x \to \infty$,

$$\begin{aligned} P(X > x) &= 1 - P(p_1 Z_1 \vee \overline{p}_1 Z_2 \leq x) = 1 - P(p_1 Z_1 \leq x) P(\overline{p}_1 Z_2 \leq x) \\ &= 1 - \left(1 - \frac{p_1}{x}\right)\left(1 - \frac{\overline{p}_1}{x}\right) \sim \frac{1}{x}\left(p_1 + \overline{p}_1\right) = \frac{1}{x}\,. \end{aligned}$$

Consequently, we choose $y = x \tan\theta$, which satisfies the conditions of Proposition 2.3 and calculate similarly,

$$\begin{aligned} &P(X > x \text{ or } Y > x \tan\theta) = 1 - P(X \leq x\,,\, Y \leq x \tan\theta) \\ &= 1 - P\left(Z_1 \leq \frac{x}{p_1} \wedge \frac{x\tan\theta}{p_2}\right) P\left(Z_2 \leq \frac{x}{\overline{p}_1}\right) P\left(Z_3 \leq \frac{x\tan\theta}{\overline{p}_2}\right) \\ &\sim \frac{1}{x}(p_1 \vee p_2 \cot\theta + \overline{p}_1 + \overline{p}_2 \cot\theta)\,, \end{aligned}$$

which implies $\psi(\theta) = 1 + \cot\theta - p_1 \wedge p_2 \cot\theta$ for $0 < \theta < \pi/2$ and

$$\rho(\theta) = \frac{p_1 \wedge p_2 \cot\theta}{1 \wedge \cot\theta}\,, \quad 0 < \theta < \pi/2\,. \tag{2.4}$$

□

We conclude with the multivariate normal distribution. It is well-known that for correlation $\rho < 1$ the upper tail dependence coefficient is $\rho_U = 0$. We shall calculate the dependence function for this important model.

Example 2.10 [Bivariate normal distribution]

Let X and Y be $N(0,1)$ with distribution function Φ and correlation ρ between X and Y. Set $\Phi(X) = U$ and $\Phi(Y) = V$, and relate the arguments as required in (2.2) by $\overline{\Phi}(x_1) = 1 - u \sim 1 - x/n$ and $\overline{\Phi}(y_1) = 1 - v \sim 1 - y/n$. Then we set $v = u\cot\theta - \cot\theta + 1$

in order to meet the restrictions of Proposition 2.6; we even have here $1-v=(1-u)\cot\theta$.

Now recall that (provided the limits exist for all $\theta\in(0,\pi/2)$

$$\frac{1-u-v+C(u,v)}{(1-u)\vee(1-v)}\;\to\;\rho(\theta)\,.$$

Note first that $1-u>1-v$ if and only if $\cot\theta<1$, i.e. $\theta>\pi/4$; so assume that this is the case. Then

$$\begin{aligned}\frac{1-u-v+C(u,v)}{1-u} &= \frac{(1-u)\cot\theta-u+C(u,1-(1-u)\cot\theta)}{1-u}\\ &= \cot\theta+\frac{-u+C(u,1-(1-u)\cot\theta)}{1-u}\,.\end{aligned}$$

We want to use the l'Hospital rule and need the derivative of the copula:

$$\begin{aligned}\frac{dC(u,1-\cot\theta(1-u))}{du} &= \frac{\partial}{\partial s}C(s,t)\mid_{\substack{s=u\\ t=1-(1-u)\cot\theta}}+\cot\theta\,\frac{\partial}{\partial t}C(s,t)\mid_{\substack{s=u\\ t=1-(1-u)\cot\theta}}\\ &= P(U\le u\mid V=1-(1-u)\cot\theta)+\cot\theta P(V\le 1-(1-u)\cot\theta\mid U=u)\\ &= P(\Phi(X)\le u\mid\Phi(Y)=1-\cot\theta(1-u))\\ &\quad+\cot\theta\,P(\Phi(Y)\le 1-(1-u)\cot\theta\mid\Phi(X)=u)\\ &= P(X\le x_1\mid Y=y_1)+\cot\theta P(Y\le y_1\mid X=x_1)\,.\end{aligned}$$

Now recall that for the bivariate normal distribution also the conditional distribution $X|Y=y_1$ is $N(\rho y_1,1-\rho^2)$ and, analogously, $Y|X=x_1$ is $N(\rho x_1,1-\rho^2)$; in particular,

$$\begin{aligned}P(X\le x_1\mid Y=y_1) &= \Phi\left(\frac{x_1-\rho y_1}{\sqrt{1-\rho^2}}\right)\\ P(Y\le y_1\mid X=x_1) &= \Phi\left(\frac{y_1-\rho x_1}{\sqrt{1-\rho^2}}\right)\,.\end{aligned}$$

From this we conclude by the l'Hospital rule (provided the limit on the right hand side exists)

$$\begin{aligned}\lim_{u,v\to1}\frac{-u+C(u,1-\cot\theta(1-u))}{1-u} &= 1-\lim_{u,v\to1}\frac{dC(u,1-\cot\theta(1-u))}{du}\\ &= 1-\lim_{x_1,y_1\to\infty}\left(P(X\le x_1\mid Y=y_1)+\cot\theta\,P(Y\le y_1\mid X=x_1)\right)\,,\end{aligned}$$

where x_1 and y_1 are related to u and v as explained before. From this we obtain

$$\rho(\theta) \;=\; \lim_{x_1,y_1\to\infty}\left(\overline{\Phi}\left(\frac{x_1}{\sqrt{1-\rho^2}}\left(1-\rho\frac{y_1}{x_1}\right)\right)+\cot\theta\,\overline{\Phi}\left(\frac{y_1}{\sqrt{1-\rho^2}}\left(1-\rho\frac{x_1}{y_1}\right)\right)\right)\,,$$

where we still have to evalute the right hand side. Recall that by asymptotic inversion, for $1-u\to 0$, which implies $-\ln(1-u)\to\infty$,

$$\frac{x_1}{y_1} \;=\; \frac{\overline{\Phi}^{\leftarrow}(1-u)}{\overline{\Phi}^{\leftarrow}(1-v)} \;\sim\; \frac{\sqrt{-2\ln(1-u)}}{\sqrt{-2\ln(1-v)}} \;=\; \sqrt{\frac{-\ln(1-u)}{-\ln((1-u)\cot\theta)}} \;\to\; 1\,.$$

This means that

$$\rho(\theta) \;=\; \lim_{x_1,y_1\to\infty}\left(\overline{\Phi}\left(x_1\sqrt{\frac{1-\rho}{1+\rho}}\right)+\cot\theta\,\overline{\Phi}\left(y_1\sqrt{\frac{1-\rho}{1+\rho}}\right)\right) \;=\; 0\,.$$

As the normal copula is symmetric, we can reverse the roles of U and V and obtain the same result for $0<\theta<\pi/4$.

Note that the above calculation only covers the case, where $x_1/y_1\to 1$. All other directions $x_1/y_1\to c$ with $c\in(0,1)$ or $c\in(1,\infty)$ correspond to $\cot\theta\to 0$ and $\cot\theta\to\infty$, respectively, yielding also $\rho(0)=\rho(\pi/4)=0$. □

Remark 2.11 For the upper tail dependence index, which corresponds to $\theta=\pi/4$, hence $u=v$ and $x_1=y_1$ we find

$$\rho(\pi/4)=\lim_{x_1\to\infty}2\overline{\Phi}\left(x_1\sqrt{\frac{1-\rho}{1+\rho}}\right)=2\overline{\Phi}\left(x_1\sqrt{\frac{1-\rho}{1+\rho}}\right)=0\,,$$

for all $\rho<1$ as is well-known; see e.g. Embrechts, et al. (2001, 2002).

3 Extreme dependence estimation

To assess extreme dependence in a data set we want to estimate the dependence function $\rho(\cdot)$ on the positive quadrant. We use a nonparametric estimator as suggested in Hsing

et al. (2003) based on the empirical distribution function, which yields a simple nonparametric estimator of $\psi(\cdot)$ and hence of $\rho(\cdot)$, Recall that the empirical distribution function is given by

$$\widehat{G}_X(x) = \widehat{P}_n(X \le x) = \frac{1}{n}\sum_{j=1}^{n} I(X_i \le x)\,, \quad x \in \mathbb{R}\,,$$

is the standard estimate for the distribution function G_X of iid data ($I(\boldsymbol{A})$ denotes the indicator function of the set $\boldsymbol{A}$). The empirical distribution function can be rewritten in terms of the rank of a random variable in the ordered sample and we write

$$\widehat{G}_X(X_i) = \widehat{P}_n(X \le X_i) = \frac{1}{n}\mathrm{rank}(X_i)$$

for any sample variable X_i for $i = 1, \ldots, n$.

We still have to explain one important issue of our estimation procedure. Recall from (2.1), denoting by $\overline{G}_X(\cdot) = 1 - G_X(\cdot)$,

$$\begin{aligned}
\Lambda_n(x,y) &:= n\,P\left(G_X(X) > 1 - \frac{x}{n} \text{ or } G_Y(Y) > 1 - \frac{y}{n}\right) \\
&= n\,P\left(n\,\overline{G}_X(X) \le x \text{ or } n\,\overline{G}_Y(Y) \le y\right) \\
&= n\,P(n\,(\overline{G}_X(X), \overline{G}_Y(Y)) \in \boldsymbol{A}) \qquad (3.5) \\
&\to \Lambda(x,y)\,.
\end{aligned}$$

By a continuity argument we can replace $n \in \mathbb{N}$ by $t \in (0,\infty)$ and also replace in a first step the probability measure P by its empirical counterpart $\widehat{P}_n$. Then we obtain

$$\begin{aligned}
\widetilde{\Lambda}_{t,n}(x,y) &= t\widehat{P}_n\left(t\,(\overline{G}_X(X), \overline{G}_Y(Y)) \in \boldsymbol{A}\right) \\
&= \frac{t}{n}\sum_{i=1}^{n} I(t\,(\overline{G}_X(X), \overline{G}_Y(Y)) \in \boldsymbol{A})\,.
\end{aligned}$$

Now estimate the two distribution tails by their empirical counterparts:

$$\widehat{\overline{G}}_X(X_i) := \frac{1}{n}R_i^X := \frac{1}{n}\mathrm{rank}(-X_i) \quad \text{and} \quad \widehat{\overline{G}}_Y(Y_i) := \frac{1}{n}R_i^Y := \frac{1}{n}\mathrm{rank}(-Y_i)\,.$$

Then setting $\varepsilon = t/n$ we obtain

$$\widehat{\Lambda}_{\varepsilon,n}(\boldsymbol{A}) \;=\; \varepsilon \sum_{i=1}^{n} I(\varepsilon\,(R_i^X\,,\,R_i^Y) \in \boldsymbol{A})\,. \tag{3.6}$$

This yields in combination with Definition 2.2 an estimator for the function ρ.

$$\widehat{\rho}_{\varepsilon,n}(\theta) = \frac{1 + \cot\theta - \widehat{\Lambda}_{\varepsilon,n}(1, \cot\theta)}{1 \wedge \cot\theta}, \tag{3.7}$$

where $\widehat{\Lambda}_{\varepsilon,n}(1, \cot\theta)$ can be rewritten as

$$\varepsilon \sum_{i=1}^{n} I(R_i^X \le \varepsilon^{-1}\,,\, R_i^Y \le \varepsilon^{-1}),\,.$$

This estimator has good convergence properties: for appropriately small ε and $n \to \infty$ it converges in probability and almost surely; see Hsing et al. (2003) and references therein. Moreover, it has the advantage that it is only based on the ranks of our data. This estimator can be smoothed in the usual way, for instance, by averaging it over a window of size $2m + 1$ for $m \in \mathbb{N}$, we call this smoothed estimator $\widehat{\rho}_{\varepsilon,n}^{(m)}(\cdot)$.

The middle column of Figure 3.1 visualizes the extreme dependence structure by means of ranks. Plotted are the $(1/R_i^X\,,\,1/R_i^Y)$. Points on the axes correspond to independent extreme points; all points in the open quadrant exhibit some extreme dependence structure. Completely dependent points are to be found on the 45-degree line.

Example 3.2 [Gumbel copula: continuation of Example 2.8]

In Figure 3.1 we simulated the model (with exponential margins) for $n = 10\,000$ iid observations of (X, Y). We estimate the dependence function $\rho(\cdot)$ for this model. We stay away from the boundaries $\theta = 0$ and $\theta = \pi/2$, since in the numerator of (3.7) we have the difference of two quantities which both tend to ∞ as $\theta \to 0$, and for θ near $\pi/2$ there is a lack of data. The three sets of plots on the three rows correspond to the cases:

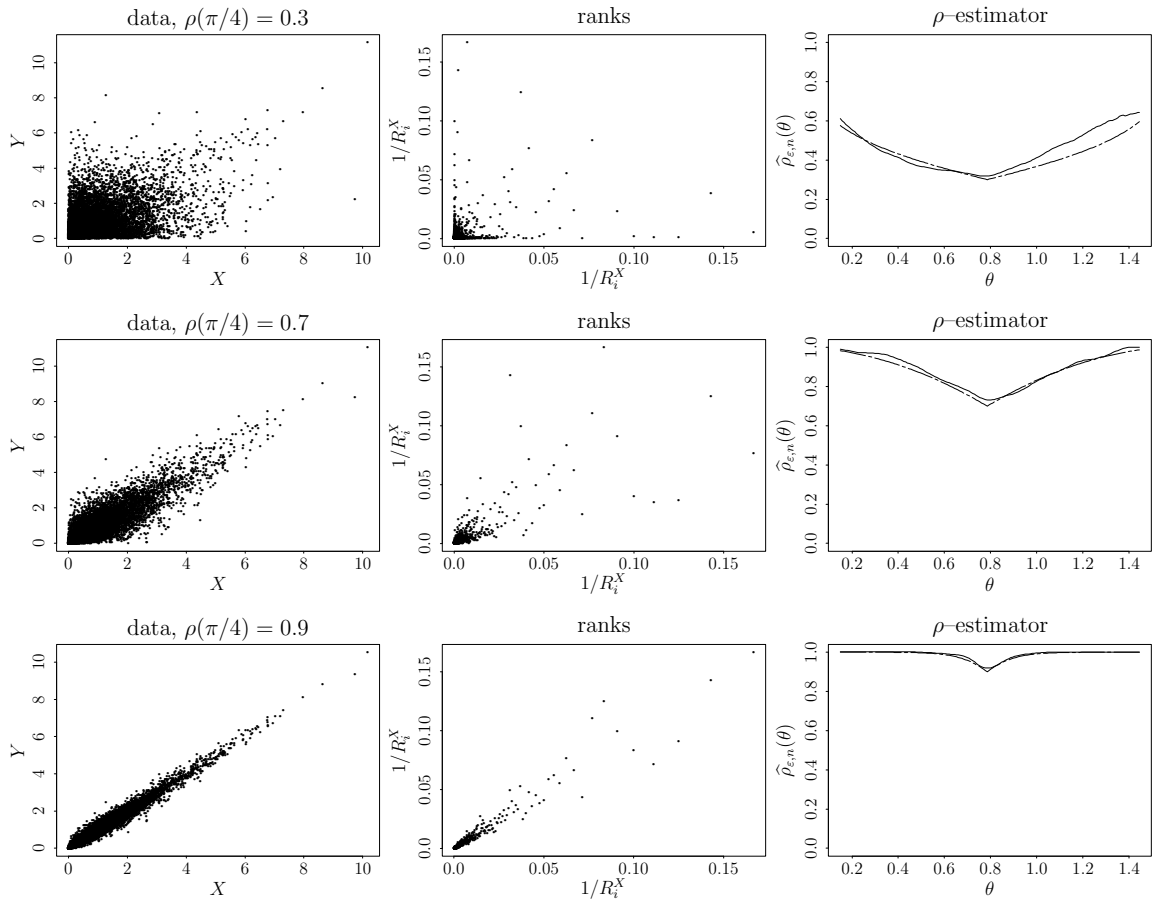

Figure 3.1 Left column: Plots of the data (X, Y), for $C_{\boldsymbol{X}}$ given in (2.3) with standard exponential margins and $\rho(\pi/4) = 0.3$ (upper row), $\rho(\pi/4) = 0.7$ (middle row), $\rho(\pi/4) = 0.9$ (lower row).

Middle column: Plots of the ranks $(1/R_i^X\,,\, 1/R_i^Y)$.

Right column: Smoothed versions $\widehat{\rho}_{\varepsilon,n}^{(5)}(\theta)$ (solid line) overlaid with true function $\rho(\theta)$.

$\rho(\pi/4) = 0.3$ (upper row), $\rho(\pi/4) = 0.7$ (middle row) and $\rho(\pi/4) = 0.9$ (lower row). On each row the left plots contain a simulated sample of size 10 000. The corresponding ranks $\left(1/R_i^X, 1/R_i^Y\right)$, $1 \le i \le n$, are shown in the middle plots. The right plots show the true functions $\rho(\theta)$ in (2.4) (dashed line) overlaid with the smoothed version of $\widehat{\rho}_{\varepsilon,n}(\theta)$ (solid line) based on the simulated sample. Note that $\rho(\pi/4)$ is the upper tail dependence coefficient, which is an appropriate and simple measure of extreme dependence for this symmetric model. The level of dependence is manifested by the data scattered around the diagonal. □

Example 3.3 [Asymmetric Pareto model: continuation of Example 2.9]

In Figure 3.4 we simulated this model for $n = 10\,000$ iid observations of (X, Y). The three sets of plots on the three rows correspond to the cases: $(p_1\,,\,p_2) = (0.7\,,\,0.3)$, $(p_1\,,\,p_2) = (0.5\,,\,0.5)$ and $(p_1\,,\,p_2) = (0.2\,,\,0.8)$. On each row the left plots contain a simulated sample of size 10 000. The corresponding ranks $\left(1/R_i^X, 1/R_i^Y\right)$, $1 \le i \le n$, are shown in the middle plots. The right plots show the true functions $\rho(\theta)$ in (2.4) (dashed line) overlaid with the smoothed version of $\widehat{\rho}_{\varepsilon,n}(\theta)$ (solid line) based on the simulated sample.

In the first row of plots, ρ is larger for small θ than for large θ; this is reflected by the left plot in which the violation of independence can be seen to be more severe below the diagonal. In the second row of plots, ρ is constant; which is reflected by having a portion of extreme points lined up on the diagonal in the left plot. The third row of plots is the converse situation to the first row, which is reflected by the pattern of extreme points above the diagonal.

This is an example of a situation where the tail dependence coefficient does not convey a good picture of extreme dependence, in that $\rho(\pi/4)$ is not sufficient to describe the full dependence structure of this model. □

Example 3.5 [Swap rate data]

Finally, we use the estimator $\widehat{\rho}$ for the *swap rate data* described in Section 1. First we show plots of $\widehat{\rho}_{\varepsilon,n}(\theta)$ for $\theta \in (0, \pi/2)$, as defined in (3.7) We stay away from the boundaries $\theta = 0$ and $\theta = \pi/2$, since in the numerator of (3.7) we have the difference of two quantities which both tend to ∞ as $\theta \to 0$, and for θ near $\pi/2$ there is a lack of data.

In Figure 3.6 the tail dependence function is estimated for various combinations of swap rates of different maturities with $\widehat{\rho}_{\varepsilon,n}(\theta_i)$ (zigzag-line) and the smoothed version

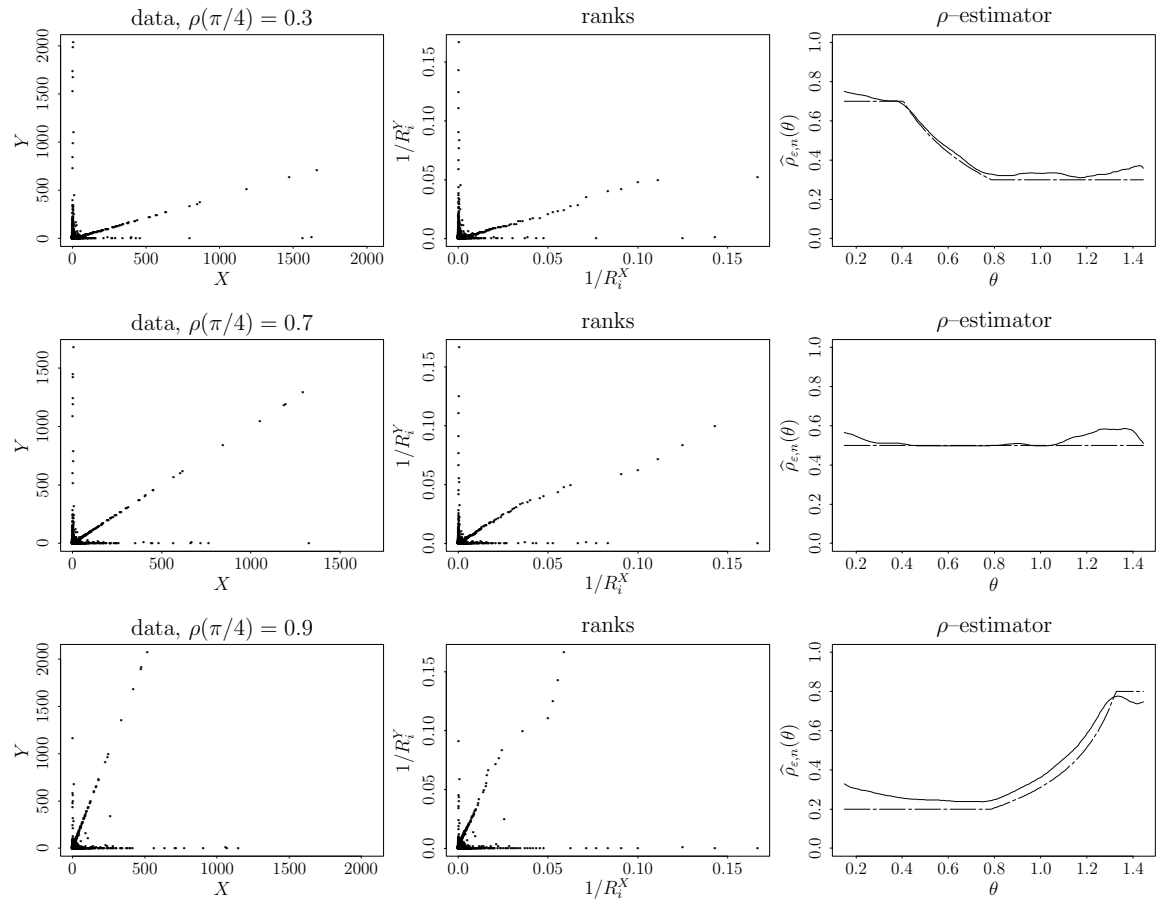

Figure 3.4 Left column: Plots of data (X, Y) from Example 3.3:

$p_1 = 0.7$, $p_2 = 0.3$ (upper row), $p_1 = 0.5$, $p_2 = 0.5$ (middle row) and $p_1 = 0.2$, $p_2 = 0.8$ (lower row).

Middle column: Plots of the ranks $(1/R_i^X, 1/R_i^X)$.

Right column: Smoothed version of $\widehat{\rho}_{\varepsilon,n}^{(5)}(\theta)$ (solid line) overlaid with true function $\rho(\theta)$.

$\widehat{\rho}_{\varepsilon,n}^{(m)}(\theta_i)$ (dashed line) for $\varepsilon = 0.06$, $m = 5$ and $\theta_i = \frac{i}{200}\frac{\pi}{2}$, $1 \le i \le 200$.

The upper row of plots exhibit some dependence for small θ and again for large θ, there is less dependence for θ around $\pi/4$. The middle row shows a similar dependence structure. This is not so surprising from the chosen maturities. The third row of plots, however, clearly shows high dependence for the whole range of θ. □

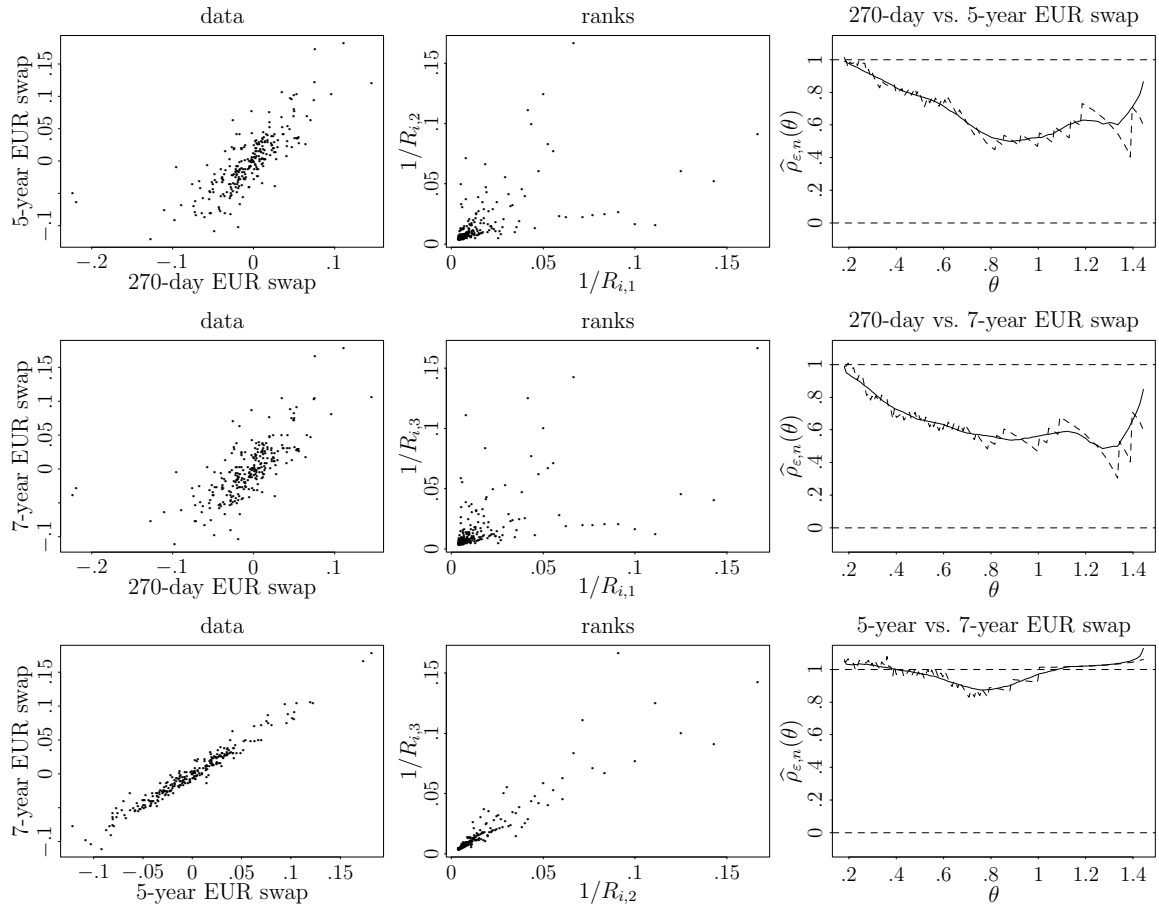

Figure 3.6 Left column: Scatter plot of data of Figure 1.3, 270-day vs. 5-year (first row), 270-day vs. 7-year (middle row) and 5-year vs. 7-year EUR swap (lower row).

Middle column: Plots of ranks $(1/R_{i,j}, 1/R_{i,k})$, $1 \le j < k \le 3$.

Right column: Smoothed version of $\widehat{\rho}^{(5)}_{\varepsilon,n}(\theta)$ (solid line) overlaid with true function $\rho(\theta)$ (zig-zag-line).

4 Conclusion

We have introduced a new tail dependence function which is tailor made to assess the extreme dependence structure in data. As it measures dependence in every direction it is also able to measure extreme dependence for data with an assymmetric dependence structure. We treat various examples, including the ubiquitous normal distribution as well as an asymmetric heavy-tailed Pareto model. Moreover, our approach is based on copulas and provides thus a general treatment for any dependence structure. In addition to a small simulation study we also show our new dependence function at work for real data and estimate extreme dependence for annually compounded zero coupon swap rates.

References

[1] Coles, S. G. (2001) *An Introduction to Statistical Modeling of Extreme Values.* Springer, London.

[2] Crouhy, M., Galai, D. and Mark, R. (2001) *Risk Management.* McGraw Hill, New York.

[3] Embrechts, P. (Ed.) (2000) *Extremes and Integrated Risk Management.* UBS Warburg and Risk Books.

[4] Embrechts, P., Klüppelberg, C. und Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[5] Embrechts, P., Lindskog, F. and McNeil, A. (2001) Modelling dependence with copulas and applications to risk management. In: Rachev, S. (Ed.) *Handbook of Heavy Tailed Distributions in Finance.* Elsevier, Chapter 8, pp. 329-384.

[6] Embrechts, P., McNeil, A. and Straumann, D. (2002) Correlation and dependence in risk managemant: properties and pitfalls. In: Dempster, M. and Moffatt, H.K. (Eds.) *Risk Management: Value at Risk and Beyond.* Cambridge University Press, Cambridge.

[7] Hsing, T., Klüppelberg, C. and Kuhn, G. (2003) Dependence estimation and visualization in multivariate extremes with applications to financial data. Submitted for Publication. Available at www.ma.tum.de/stat/.

[8] Joe, H. (1997) *Multivariate Models and Dependence Concepts.* Chapman & Hall, London.

[9] Klüppelberg, C. (2004) Risk management with extreme value theory. In: B. Finkenstädt and H. Rootzén (Eds.) *Extreme Values in Finance, Telecommunications and the Environment*, pp. 101-168. Chapman & Hall/CRC Press, Boca Raton.

[10] Kuhn, G. (2002) *Multivariate Value at Risk Schätzung von Zins Swap Sätzen.* Diploma Thesis. Munich University of Technology.
Available at www.ma.tum.de/stat/.

[11] McNeil, A. and Frey, R. (2000) Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance* **7**, 271-300.

[12] Smith, R. L. (2004) Statistics of extremes, with applications in environment, insurance and finance. In: B. Finkenstädt and H. Rootzén (Eds.) *Extreme Values in Finance, Telecommunications and the Environment*, pp. 1-78. Chapman & Hall/CRC Press, Boca Raton.