Appendix 1

Pair-copula decomposition

We start defining the copula C. Let $X = (X_1, ..., X_k)$ be a multivariate random vector of k continuous variables and F is its joint cumulative distribution function (cdf), which contains information about the individual behaviour of these variables and their dependence structure. The complex non-linear dependence structure often existing between risk factors makes it necessary to search for approaches that provide a way to separate dependence structure from marginal behavior. The copulae provide an easy way to do this.

Sklar (1959), pioneer in copulas definition, through his famous Slark's theorem, demonstrates the existence of a copula C, through which it is possible to establish a functional relationship between a multivariate distribution function and its marginal distribution functions. Given F the joint cdf associated to the k-dimensional random vector X, with univariate marginal $F_1, \ldots, F_k$, then there exists a copula C such that:

$$F(x_1, \ldots, x_k) = C(F_1(x_1), F_2(x_2), \ldots, F_k(x_k)). \tag{14}$$

If we denote $u_j = F_j(x_j),\ j = 1, \ldots, k$, the values from a Uniform$[0,1]$, then copula C is defined as:

$$C(u_1, \ldots, u_k) = F(F_1^{-1}(u_1), \ldots, F_k^{-1}(u_k)), \tag{15}$$

where $F_j^{-1},\ j = 1, \ldots, k$, is the inverse of the distribution function associated with the marginal j. Moreover, if $F_j(x_j),\ j = 1, \ldots, k$, are continuous C is unique.

From expression (14) the joint probability density function (pdf) is obtained

$$f(x_1, \ldots, x_k) = c_{1,\ldots,k}(F_1(x_1), \ldots, F_k(x_k)) \cdot f_1(x_1) \ldots f_k(x_k), \tag{16}$$

where $c_{1,\ldots,k}$ is a k-dimensional copula density.

On the other hand, the joint pdf can be expressed, for example, using the following decomposition:

$$\begin{aligned} f(x_1, \ldots, x_k) &= f_k(x_k) \cdot f_{k-1|k}(x_{k-1}|x_k) \cdot f_{k-2|k-1,k}(x_{k-2}|x_{k-1}, x_k) \ldots f_{1|2,\ldots,k}(x_1|x_2, \ldots, x_k) \\ &= f_k(x_k) \cdot \prod_{j=1}^{k-1} f_{j|j+1,\ldots,k}(x_j|x_{j+1}, \ldots, x_k), \end{aligned} \tag{17}$$

where $f_{A|B}(A|B)$ is the conditional density function of A given B. We know that the conditional density function can be expressed as:

$$f_{j|j+1,\ldots,k}(x_j|x_{j+1}, \ldots, x_k) = \frac{f_{j,j+1|j+2,\ldots,k}(x_j, x_{j+1}|x_{j+2}, \ldots, x_k)}{f_{j+1|j+2,\ldots,k}(x_{j+1}|x_{j+2}, \ldots, x_k)}. \tag{18}$$

Every decomposition similar to the one expressed in (17) supposes a different structure of dependency that defines the interrelations between variables.

Using Sklar's theorem and calculating the derivative for obtaining the probability density function, expression (18) in terms of copula is equal to:

$$\begin{aligned} & f_{j|j+1,\dots,k}(x_j|x_{j+1},\dots,x_k) = \\ = \; & c_{j,j+1|j+2,\dots,k}(F_{j|j+2,\dots,k}(x_j|x_{j+2},\dots,x_k),F_{j+1|j+2,\dots,k}(x_{j+1}|x_{j+2},\dots,x_k)) \\ \cdot \; & f_{j|j+2,\dots,k}(x_j|x_{j+2},\dots,x_k). \end{aligned} \tag{19}$$

Replacing recursively in (19) we obtain that:

$$\begin{aligned} & f_{j|j+1,\dots,k}(x_j|x_{j+1},\dots,x_k) = \\ = \; & c_{j,j+1|j+2,\dots,k}(F_{j|j+2,\dots,k}(x_j|x_{j+2},\dots,x_k),F_{j+1|j+2,\dots,k}(x_{j+1}|x_{j+2},\dots,x_k)) \\ \cdot \; & c_{j,j+2|j+3,\dots,k}(F_{j|j+3,\dots,k}(x_j|x_{j+3},\dots,x_k),F_{j+2|j+3,\dots,k}(x_{j+2}|x_{j+3},\dots,x_k)) \\ \cdot \; & \dots \\ \cdot \; & c_{j,k}(F_j(x_j),F_k(x_k)) \cdot f_j(x_j), \end{aligned} \tag{20}$$

simplifying notation as $c_{A,B}(F_A(x_A),F_B(x_B))) = c_{A,B}$ and $f_A(x_A) = f_A$, with $A \neq B$, then:

$$\begin{aligned} & f_{j|j+1,\dots,k}(x_j|x_{j+1},\dots,x_k) = \\ = \; & f_j \cdot \prod_{i=1}^{k-j} c_{j,j+i|j+i+1,\dots,k}, \; \forall j = 1,\dots,k-1. \end{aligned} \tag{21}$$

Substituting result (21) in (17) we obtain a possible factorization of the join distribution expressed in term of bivariate copulae. The joint pdf is:

$$f(x_1,\dots,x_k) = \prod_{j=1}^{k-1}\prod_{i=1}^{k-j} c_{j,j+i|j+i+1,\dots,k} \prod_{l=1}^{k} f_l \tag{22}$$

and the joint cdf is:

$$F(x_1,\dots,x_k) = \prod_{j=1}^{k-1}\prod_{i=1}^{k-j} C_{j,j+i|j+i+1,\dots,k}. \tag{23}$$

In general, expression (23) represents a multivariate distribution in function of $k(k-1)/2$ dependence parameters.

D-Vine

A D-vine is a sequence of m trees that we denote as $T_1,\ldots,T_m$. Every tree has n nodes and the associations between the nodes of each tree are called edges. In this case, for each tree with n nodes we have $n-1$ edges. For each edge a pair-copula arises.

The D-vine is a special case of R-vine in which it is necessary to choose the order of the variables (nodes) in the upper tree, T_1. An example of trees associated with a D-vine that represent a multivariate distribution with $k=6$ dimensions and its decomposed pdf is shown in (24).

$$\begin{aligned} f(x_1,x_2,x_3,x_4,x_5,x_6) \;=\;& \underbrace{f_1\cdot f_2\cdot f_3\cdot f_4\cdot f_5\cdot f_6}_{\textit{nodes in } T_1}\cdot\underbrace{c_{12}\cdot c_{23}\cdot c_{34}\cdot c_{45}\cdot c_{56}}_{\textit{edges in } T_1-\textit{nodes in } T_2} \\ \cdot\;& \underbrace{c_{13|2}\cdot c_{24|3}\cdot c_{35|4}\cdot c_{46|5}}_{\textit{edges in } T_2-\textit{nodes in } T_3}\cdot\underbrace{c_{14|23}\cdot c_{25|34}\cdot c_{36|45}}_{\textit{edges in } T_3-\textit{nodes in } T_4} \\ \cdot\;& \underbrace{c_{15|234}\cdot c_{26|345}}_{\textit{edges in } T_4-\textit{node in } T_5}\cdot\underbrace{c_{16|2345}}_{\textit{node in } T_5}\cdot \end{aligned} \tag{24}$$

The resulting 6-dimensional joint cdf is:

$$\begin{aligned} F(x_1,x_2,x_3,x_4,x_5,x_6) = \\ C_{12}(F_1(x_1),F_2(x_2))\cdot C_{23}(F_2(x_2),F_3(x_3))\cdot C_{34}(F_3(x_3),F_4(x_4)) \\ \cdot C_{45}(F_4(x_4),F_5(x_5))\cdot C_{56}(F_5(x_5),F_6(x_6)) \\ \cdot C_{13|2}(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2))\cdot C_{24|3}(F_{2|3}(x_2|x_3),F_{4|3}(x_4|x_3)) \\ \cdot C_{35|4}(F_{3|4}(x_3|x_4),F_{5|4}(x_5|x_4))\cdot C_{46|5}(F_{4|5}(x_4|x_5),F_{6|7}(x_6|x_5)) \\ \cdot C_{14|23}(F_{1|23}(x_1|x_2,x_3),F_{4|23}(x_4|x_2,x_3)) \\ \cdot C_{25|34}(F_{2|34}(x_2|x_3,x_4),F_{5|34}(x_5|x_3,x_4)) \\ \cdot C_{36|45}(F_{3|45}(x_3|x_4,x_5),F_{6|45}(x_6|x_4,x_5)) \\ \cdot C_{15|234}(F_{1|234}(x_1|x_2,x_3,x_4),F_{5|234}(x_5|x_2,x_3,x_4)) \\ \cdot C_{26|345}(F_{2|345}(x_2|x_3,x_4,x_5),F_{6|345}(x_6|x_3,x_4,x_5)) \\ \cdot C_{16|2345}(F_{1|2345}(x_1|x_2,x_3,x_4,x_5),F_{6|2345}(x_6|x_2,x_3,x_4,x_5)). \end{aligned} \tag{25}$$

Each pair-copula $C_{A|B}$ requires at least one parameter $\theta_{A|B}$ to be estimated. At the end, we will have as many estimated parameters as edges plus one exist in the D-Vine, in our case $6(6-1)/2=15$ parameters of pair-copula.

On the other hand, we know that conditional cdfs can be calculate as:

$$F_{j|B} = \frac{\partial C_{j,i|B}(F_{j|B}(x_j|B),F_{i|B}(x_i|B))}{\partial F_{i|B}(x_i|B)}, \tag{26}$$

where B is the group of conditional variables different to i and j. Furthermore, we know that for two variables:

$$F_{j|i} = \frac{\partial C_{j,i}(F_j(x_j), F_i(x_i))}{\partial F_i(x_i)}. \tag{27}$$

So, if we can approximate the univariate cdfs we can obtain conditional cdfs recursively starting with (27) and, therefore, we can express copula and its density in functions of these partial derivatives.

Appendix 2

Table 10: Companies in the share portfolios.

	Stock 1 (R1)	**Stock 2 (R2)**	**Stock 3 (R3)**	**Stock 4 (R4)**	**Stock 5 (R5)**	**Stock 6 (R6)**
P1	Petrobras	Google	Coca-cola	General Motors	DirecTV	Moody's Corporation
P2	Petrobras	Google	Coca-cola	Procter & Gamble	Telefónica Brasil	Telecom Italia
P3	Merck	Novartis	Pfizer	Deutsche Bank	ING	Santander
P4	Merck	Novartis	Pfizer	Sanofi	GlaxoSmithKline	AstraZeneca

Table 11: Fitted ARMA(P,Q)-GARCH(p,q) for each return time series and scale descriptive statistics of each filtered return time series.

Share	**ARMA**	**GARCH**	**STD**	**IQR**	**RANGE**
Petrobras	(0,0)	(0,0)	0.024	0.027	0.256
Google	(0,0)	(0,0)	0.016	0.015	0.217
Coca-cola	(1,1)	(1,1)	0.010	0.011	0.096
General Motors	(0,0)	(1,1)	0.021	0.024	0.206
DirecTV	(0,0)	(1,1)	0.014	0.015	0.142
Moody's Corporation	(0,0)	(1,1)	0.021	0.019	0.224
Procter & Gamble	(0,0)	(1,1)	0.009	0.010	0.100
Telefónica Brasil	(0,0)	(0,0)	0.018	0.020	0.167
Telecom Italia	(0,0)	(1,1)	0.027	0.031	0.215
Merck	(0,0)	(1,1)	0.012	0.012	0.114
Novartis	(1,0)	(1,1)	0.011	0.012	0.107
Pfizer	(1,1)	(1,1)	0.012	0.012	0.101
Sanofi	(0,0)	(1,1)	0.016	0.018	0.140
GlaxoSmithKline	(0,0)	(1,1)	0.011	0.013	0.101
AstraZeneca	(0,0)	(1,1)	0.012	0.013	0.135
Deutsche Bank	(1,1)	(1,1)	0.029	0.030	0.288
ING	(1,1)	(1,1)	0.018	0.020	0.167
Santander	(0,0)	(1,1)	0.025	0.027	0.219

Table 12: Parameters of models used in simulation study.

		Model 1	Model 2	Model 3	Model 4
θ_{12}	ρ	0.402	0.408	0.469	0.474
	d.f.	7.8	8.9	11.2	11.8
θ_{23}	ρ	0.431	0.433	0.514	0.515
	d.f.	12.2	13.1	7.1	6.9
θ_{34}	ρ	0.413	0.408	0.562	0.559
	d.f.	9.2	8.6	6.7	5.8
θ_{45}	ρ		0.424		0.581
	d.f.		7.0		5.8
θ_{56}	ρ		0.491		0.668
	d.f.		11.1		5.3
$\theta_{13\|2}$	ρ	0.200	0.209	0.514	0.513
	d.f.	61.1	57.7	4.2	4.2
$\theta_{24\|3}$	ρ	0.367	0.366	0.522	0.523
	d.f.	21.0	22.2	9.7	9.8
$\theta_{35\|4}$	ρ		0.362		0.234
	d.f.		10.4		8.9
$\theta_{46\|5}$	ρ		0.451		0.441
	d.f.		18.1		8.3
$\theta_{14\|23}$	ρ	0.255	0.257	0.151	0.151
	d.f.	24.9	22.7	19.0	19.1
$\theta_{25\|34}$	ρ		0.170		0.271
	d.f.		96.9		11.1
$\theta_{36\|45}$	ρ		0.259		0.181
	d.f.		37.4		62.6
$\theta_{15\|234}$	ρ		0.198		0.077
	d.f.		12.9		83.687
$\theta_{26\|345}$	ρ		0.185		0.191
	d.f.		60.6		92.7
$\theta_{16\|2345}$	ρ		0.166		0.148
	d.f.		12.5		80.1

Table 13: Theoretical values for VaR and CVaR obtained from theoretical models in simulation study.

		Model 1	Model 2	Model 3	Model 4
VaR	99%	3.764	5.582	3.841	5.720
	99.50%	3.844	5.524	3.905	5.827
CVaR	99%	0.090	0.158	0.069	0.116
	99.50%	0.061	0.300	0.042	0.074