Central European Review of Economic Issues

EKONOMICKÁ REVUE

The impact of association measures within the portfolio dimensionality reduction problem

Sergio ORTOBELLI, University of Bergamo[i]
Tomáš TICHÝ, VŠB-TU Ostrava[ii]

Abstract

The dependency structure of random sources plays a crucial role in portfolio theory and in several pricing and risk management problems. In this paper, we discuss the possible usage of alternative association measures in portfolio problems. Among association measures, we highlight those that are consistent with the choices of risk-averse investors and we characterise semidefinite positive association measures. Additionally, we propose new portfolio selection problems that optimise the association between the portfolio and market benchmarks and follow a dimensionality reduction problem. Finally, by carrying out an empirical analysis, we show the impact of selected association measures within the portfolio problem. This analysis proves that the proper usage of both a risk measure and an association measure can increase the portfolio performance substantially.

Keywords

Concordance measure, dimensionality reduction, large-scale portfolio selection, reward measure.

1. Introduction

The dependency structure of random sources plays a crucial role in portfolio theory and in several pricing and risk management problems. In particular, the classic Pearson linear correlation measure is regularly used to measure and optimise the dispersion of portfolio returns and to reduce the dimensionality of large-scale portfolio problems. However, it is not clear why this measure of linear correlation is still so popular despite its drawbacks.

For example, it is well known that Pearson linear correlation works well only with elliptically distributed vectors (that admit a finite variance/covariance matrix). Unfortunately, the Gaussian distributional assumption of return series is usually rejected, see e.g. Mandelbrot (1963a,b) and Fama (1965), or Rachev and Mittnik (2000) and the references therein. Moreover, the empirical evidence (see, among others, Rachev et al. (2008) and Biglova et al. (2009)) suggests that the dependence model should account for the dependence of tail events (*huge losses go together*). Many other measures have been proposed in the literature to deal and summarise the dependence among random variables (see, among others, Scarsini (1984), Cherubini et al. (2004), Nelsen (2006) and the references therein). However, most of these measures cannot be used directly to order investors' choices

since they are not consistent with investors' preferences.

In this paper, we propose using alternative association measures in order to reduce the dimensionality of large-scale portfolio selection problems and applying them subsequently within the optimal portfolio selection problem. In order to show the performance of both steps we provide an empirical analysis on the US stock market.

We proceed as follows. In Section 2, we summarise some of the basic characteristics of concordance/association measures and characterise the most important examples. In Section 3, we discuss when and how we can use association measures for portfolio problems. In Section 4, we provide the results of an empirical comparison of selected portfolio strategies and the principal findings are summarised in Section 5.

2. Concordance and semidefinite positive association measures

One of the most essential tasks of financial decision-making is the measurement of the dependency among the realisations of particular random variables. Specifically, let us consider n risky assets with gross returns[1] $z = [z_1, z_2, \ldots, z_n]'$. As a consequence of the Sklar theorem (Sklar, 1959), the joint distribution function is given by:

$$F_{\mathbf{z}}(\mathbf{x}) = C\big(F_{z_1}(x_1), F_{z_2}(x_2), \ldots, F_{z_n}(x_n)\big), \qquad (1)$$

where $F_{z_i}(x_i) = \Pr(z_i \leq x_i)$ are the marginal distribution functions and $C : [0,1]^n \to [0,1]$ is the copula function. The copula function can therefore be defined by inverting (1) as follows:

$$C(\mathbf{u}) = F_{\mathbf{z}}\big(F_{z_1}^{-1}(u_1), F_{z_2}^{-1}(u_2), \ldots, F_{z_n}^{-1}(u_n)\big) \qquad (2)$$

Therefore, the dependency among particular variables is fully described by a suitable copula function C. It follows that any copula function can be regarded as the joint distribution function of marginal distribution functions, although in several financial contexts it is convenient to express the dependency between random variables using a single number (more generally, for n random variables we get an n-dimensional

matrix). Moreover, copula functions are useful when defining these dependency measures.

Generally, a concordance measure is used to measure the concordance/dependence association between random variables. In the following example (Nelsen, 2006), two random variables (X, Y) with independent replications, (X_1, Y_1) and (X_2, Y_2), are concordant if $X_1 < X_2$ $(X_1 > X_2)$ implies $Y_1 < Y_2$ $(Y_1 > Y_2)$. Similarly, the two variables are discordant if $X_1 < X_2$ $(X_1 > X_2)$ implies $Y_1 > Y_2$ $(Y_1 < Y_2)$. The concordance measures are easily definable by copula functions, since they rely only on the *joint* features, being not related to the marginal characteristics.

Formally, a concordance measure ρ defined on a space of random variables H is any function that satisfies the following seven properties:

1. $\rho : H \times H \to [-1,1]$;
2. for any random variable $X \in H$: $\rho(X, X) = 1$; $\rho(X, -X) = -1$;
3. $\rho(X, Y) = \rho(Y, X)$;
4. $\rho(-X, Y) = \rho(X, -Y) = -\rho(X, Y)$;
5. if X and Y are independent random variables, then $\rho(X, Y) = 0$;
6. if we consider two bivariate random vectors $\mathbf{X} = (X_1, X_2)$, $\mathbf{Y} = (Y_1, Y_2)$, with the same marginal distributions (F_1, F_2) such that $F_{\mathbf{X}}(\mathbf{x}) = \Pr(X_1 \leq x_1, X_2 \leq x_2) \leq F_{\mathbf{Y}}(\mathbf{x})$ for any $\mathbf{x} = (x_1, x_2) \in R^2$ (i.e. $\mathbf{X}$ dominates $\mathbf{Y}$ with respect to concordance ordering[2]) then $\rho(X_1, X_2) \leq \rho(Y_1, Y_2)$ (or $\rho_{C_1} \leq \rho_{C_2}$ where C_1, C_2 are the copulas associated with bivariate vectors $\mathbf{X}, \mathbf{Y}$);
7. given a sequence of continuous bivariate random vectors $\{(X_n, Y_n)\}_{n \geq 1}$ with copulas C_n that converge pointwise to the copula C, then ρ_{C_n} converge to ρ_C.

Observe that $\rho(X_1, X_2) = \rho(h_1(X_1), h_2(X_2))$ for any concordance measure ρ, for any two continuous

---

[1] Generally, we assume the discrete definition of gross returns between time t and time $t + 1$ of asset i as $z_{i,t+1} = \dfrac{S_{i,t+1} + d_{i,[t,t+1]}}{S_{i,t}}$, where $S_{i,t}$ is the price of the i-th asset at time t and $d_{i,[t,t+1]}$ is the total amount of cash dividends between t and $t + 1$.

[2] Similarly, we say that X dominates Y in the sense of concordance ordering if and only if the copulas C_1, C_2 associated to $\mathbf{X}, \mathbf{Y}$ are ordered i.e. $C_1 \leq C_2$. This definition is also equivalent to saying that $\mathrm{cov}(h_1(X_1), h_2(X_2)) \leq \mathrm{cov}(h_1(Y_1), h_2(Y_2))$ for any increasing function h_1, h_2 such that covariance exists.

random variables (X_1, X_2) and for any two strictly monotone (either both increasing or both decreasing) functions h_1, h_2. The Pearson correlation coefficient is not a concordance measure, since it does not satisfy Property 7 of the concordance measures. For further details on all properties of concordance measures and their proofs, see Cherubini et al. (2004) and Nelsen (2006).

The most popular measures of concordance are Kendall's tau, Spearmann's rho, Gini's gamma and Blomqvist's beta. We will state some examples below.

Example 1. *Kendall's tau*, τ_K (also called Kendall correlation) is defined as the probability of concordance reduced by the probability of discordance:

$$\tau_K(X,Y) = \Pr((X_1 - X_2)(Y_1 - Y_2) > 0) - \Pr((X_1 - X_2)(Y_1 - Y_2) < 0), \quad (3)$$

where (X_1, Y_1) and (X_2, Y_2) are independent replications of (X,Y). Therefore:

$$\tau_K(X,Y) = E(sign((X_1 - X_2)(Y_1 - Y_2))),$$

where $sign(x) = 1$ if $x > 0$, $sign(x) = 0$ if $x = 0$ and $sign(x) = -1$ if $x < 0$. Clearly, Kendall's tau can be defined in terms of the copula function:

$$\tau_K(C) = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1, \quad (4)$$

where C is the copula associated to the bivariate vector (X,Y).

Example 2. The second most popular measure of concordance, *Spearman's rho*, ρ_S, is given by:

$$\rho_S = 3\Pr((X_1 - X_2)(Y_1 - Y_3) > 0) - \Pr((X_1 - X_2)(Y_1 - Y_3) < 0) \quad (5)$$
$$= 3E(sign((X_1 - X_2)(Y_1 - Y_3))),$$

where (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) are independent replications of (X,Y).

This measure is the Pearson linear correlation coefficient applied to the marginal distribution functions of random variables, i.e.:

$$\rho_S = cor(F_X(X), F_Y(Y)) = \frac{cov(F_X(X), F_Y(Y))}{\sqrt{var(F_X(X)), var(F_Y(Y))}}. \quad (6)$$

It follows that it can be regarded as the correlation of copula functions:

$$\rho_S(X,Y) = 12\int_0^1\int_0^1 uv dC(u,v) - 3$$
$$= 12\int_0^1\int_0^1 C(u,v)du dv - 3, \quad (7)$$

where C is the copula associated to the bivariate vector (X,Y).

Example 3. Another measure used to quantify concordance among random variables is *Gini's gamma*,

γ_G. This can be defined in terms of copula functions as follows:

$$\gamma_G(C) = 4\left[\int_0^1 C(u, 1-u)du - \int_0^1 [u - C(u,u)]du\right], \quad (8)$$

where C is the copula associated to the bivariate vector (X,Y). Its sample estimation is given by ranks p_i and q_i of the random variables X and Y, respectively:

$$\gamma_G(X,Y) = \frac{1}{\lfloor n^2/2\rfloor}\left[\sum_{i=1}^n |p_i + q_i - n - 1| - \sum_{i=1}^n |p_i - q_i|\right]. \quad (9)$$

Example 4. *Blomqvist's beta*, β_B, is defined as follows:

$$\beta_B(X,Y) =$$
$$\Pr[(X - \tilde{x})(Y - \tilde{y}) > 0] - \Pr[(X - \tilde{x})(Y - \tilde{y}) < 0] \quad (10)$$
$$= E(sign((X - \tilde{x})(Y - \tilde{y}))),$$

where $\tilde{x}$ and $\tilde{y}$ are the medians of given continuous random variables X and Y, respectively. With certain simplifications, this measure may also be rewritten in terms of copula functions:

$$\beta_B(C) = 4C\left(\frac{1}{2}, \frac{1}{2}\right) - 1. \quad (11)$$

The proof that all these measures are really measures of concordance can be found, for example, in Nelsen (2006).

Example 5. A further example shows us a generalisation of the Pearson correlation:

$$O_{p,\Im_1}(X,Y) =$$

$$cor\left(\left(X^{<\frac{p}{2}>} - E\left(X^{<\frac{p}{2}>} \mid \Im_1\right)\right), \left(Y^{<\frac{p}{2}>} - E\left(Y^{<\frac{p}{2}>} \mid \Im_1\right)\right)\right)^{<\min(\frac{p}{2}, 2)>},$$

where $X^{<p>} = sign(X)|X|^p$, $\Im_1$ is a sub-sigma algebra of $\Im$ (i.e. $\Im_1 \subset \Im$) and X and Y are not $\Im_1$ measurable.

It is evident that the measure above is a logical extension of the Pearson correlation measure, since we obtain the Pearson correlation measure if $p = 2$ and $\Im_1 = \{\varnothing; \Omega\}$.

For example, we can use sigma algebra $\Im_1$ generated by a finite partition of Ω, that is, $\Im_1 = <\{A_i; i = 1, ..., n\}>$, where $A_i \in \Im$, $A_i \cap A_j = \varnothing$, $\forall i \neq j$ and $\cup_{i=1}^n A_i = \Omega$. In portfolio problems, we can think that:

$$A_1 = \{z_b \leq F_{z_b}^{-1}(\alpha_1)\}, \quad A_i = \{F_{z_b}^{-1}(\alpha_{i-1}) < z_b \leq F_{z_b}^{-1}(\alpha_i)\}$$

for $i = 2, ..., n-1$,

$$A_n = \{z_b > F_{z_b}^{-1}(\alpha_{n-1})\},$$

where $0 < \alpha_1 < ... < \alpha_{n-1} < 1$, and

$$F_{z_b}^{-1}(\beta) = \inf\{u \mid \Pr(z_b \leq u) \geq \beta\}$$

and z_b is a benchmark of the market.[3] Under these assumptions, the conditional expectation can be easily estimated, since it is given by the simple function:

$$E(X/\Im_1)(w) = \sum_{i=1}^{n} I_{[X(w)\in A_i]} \frac{1}{\Pr(A_i)} \int_{A_i} X d\Pr \quad \forall w \in \Omega,$$

where

$$I_{[X(w)\in A]} = \begin{cases} 1 & \text{if } X(w)\in A, \\ 0 & \text{otherwise.} \end{cases}$$

Given a sample of n i.i.d. copies (X_i, Y_i) of the bivariate vector (X,Y) and assuming a suitable sigma algebra $\Im_1$ as above we get:

- $\frac{1}{\#(X_i \in A)} \sum_{X_i \in A} X_i^{<p/2>}$ (where $\#(X_i \in A)$ is the number of observations of X_i belonging to A). This is a consistent estimator of $\frac{1}{\Pr(A)} \int_A X^{<p/2>} d\Pr$ and thus we can estimate $O_{p,\Im_1}(X,Y)$.

3. The portfolio dimensional problem

Papp et al. (2005) and Kondor et al. (2007) showed that the number of observations should increase proportionally with the number of assets in order to get a good approximation of the portfolio risk/reward measures. Therefore, it is necessary to find the right trade-off between a statistical approximation of the historical series depending only on a few parameters and the number of historical observations. In practice, portfolio managers reduce the dimensionality of the problem approximating the return series with a *k*-fund separation model (or other regression-type models) that depends on an adequate number (not too large) of parameters.

Thus, we can perform a principal component analysis (PCA) of the gross returns of the stocks used in order to identify the few factors (portfolios) with the highest return variability (see Biglova et al. (2009)). Therefore, we replace the original n correlated time series z_i with the n uncorrelated time series R_i

assuming that each z_i is a linear combination of the series R_i. This is always possible when we use a linear correlation measure ρ. Then, we implement a dimensionality reduction by choosing only those factors whose variability is significantly different from zero. We call portfolio factors f_i the s time series R_i with a significant dispersion measure, while the remaining $n-s$ series with very small dispersion measures are summarised by an error. Thus, each series z_i is a linear combination of the factors plus a small uncorrelated noise:

$$z_i = \sum_{j=1}^{s} a_{ij} f_j + \sum_{j=s+1}^{n} a_{ij} R_j = \sum_{j=1}^{s} a_{ij} f_j + \varepsilon_i. \quad (12)$$

We can apply the PCA either to the Pearson correlation matrix or to any other linear correlation measure, for example $Q = [\rho_{i,j}]$, where $\rho_{i,j} = O_{2,\Im_1}(z_i, z_j)$ for a suitable sigma algebra $\Im_1$. Having identified the s factors, $f_j = \sum_{i=1}^{n} x_i z_i$ ($j = 1,...,s$; such that $\sum_{k=1}^{n} x_k^2 = 1$), which account for most of the variability in gross returns, we further reduce the variability of the error by regressing the series on the factors f_j so that we get:

$$z_i = b_{i,0} + \sum_{j=1}^{s} b_{i,j} f_j + \varepsilon_i. \quad (13)$$

Once we have reduced the dimensionality of the problem, we can apply the portfolio selection optimisation problem:

$$\max_x \frac{v(x'\mathbf{z})}{q(x'\mathbf{z})} \left(\rho_1\left(x'\mathbf{z}, \max_i z_i\right) - \rho_2\left(x'\mathbf{z}, \min_i z_i\right) \right) \quad (14)$$

with $\sum_{i=1}^{n} x_i = 1$, $x_i \geq 0$ and $v(x'\mathbf{z})$ and $q(x'\mathbf{z})$ denote the risk and reward measure, respectively.

Recall from portfolio theory that the portfolio that provides the maximum reward per unit of risk is called the market portfolio. In particular, when reward and risk are both positive measures, the market portfolio is the solution for the optimisation problem (14) after a suitable choice of association measures.

In our case, the portfolio selection problem (14) is applied to the approximated portfolio of gross returns $x'\mathbf{z} \approx x'\hat{\mathbf{b}}_0 + \sum_{j=1}^{s} x'\hat{\mathbf{b}}_j f_j$, where $\hat{\mathbf{b}}_j = [\hat{b}_{1,j},...,\hat{b}_{n,j}]'$ is the vector of estimated coefficients $\hat{b}_{i,j}$ ($j = 0,1,...,s$). Such a procedure is computationally efficient and can be applied using any linear correlation measure.

---

[3] Generally, we can distinguish two different types of benchmarks: *artificial benchmarks* and *traded benchmarks*. Traded benchmarks are some indices traded on the market that represent some sectors and/or markets. For these benchmarks we can obtain historical observations. Artificial benchmarks are not traded on the market and they are artificially created by portfolio manager to represent the best/worst indicators of the assets used.

4. Empirical application using US data

In this section, we employ various uncertainty measures and association measures within the portfolio dimensionality reduction and large-scale portfolio selection, as defined in Sections 2 and 3. In this context, we consider 1,304 equities on the US stock market (450 equities from the Nasdaq and 854 equities from the NYSE). We use daily data for the period from January 1997 to December 2009 (3,258 observations in total). All data are obtained from DataStream.

Hence, our task is to reduce the dimensionality of the portfolio problem in order to compare portfolio strategies in a large-scale framework. Generally, we can consider two possible criteria for selecting *the best* principal components:

- Take the first principal components that together explain at least 50% of the dispersion; or

- Take only those principal components that explain not less than $100/N\%$ of the dispersion measure (Kaiser-type rule) (where $N = 1,304$ is the number of all assets in our empirical analysis).

However, in several large portfolio problems we should choose only few principal components in order to guarantee a sufficiently good approximation to the optimal portfolio problem (see, among others, Papp et al. (2005) and Kondor et al. (2007)). According to these studies and considering that we have 3258 historical observations, we can guarantee a good approximation of the portfolio problem if we take the first 35 principal components. For this reason, in the portfolio dimensionality reduction analysis we also consider the criterion:

- Take the first 35 principal components and show how much dispersion is explained.

As linear correlation measures, we consider (see Section 2):

a) *The Pearson correlation measure* $\rho_{i,j} = \mathrm{cor}\big(z_i, z_j\big)$ and *its conditional version* $\rho_{i,j} = \mathrm{cor}\big(z_i, z_j \mid x'\mathbf{z} \le F_{x'z}^{-1}(0.05)\big)$, where $x'\mathbf{z} = \frac{1}{N}\sum_{i=1}^{N} z_i$ is the equidiversified portfolio;

b) *The linear correlation measure* $\rho_{i,j} = O_{2,\mathfrak{I}_1}\big(z_i, z_j\big)$ and *its conditional version* $\rho_{i,j} = O_{2,\mathfrak{I}_1}\big(z_i, z_j \mid x'\mathbf{z} \le F_{x'z}^{-1}(0.1)\big)$, where $x'\mathbf{z} = \frac{1}{N}\sum_{i=1}^{N} z_i$ and $\mathfrak{I}_1 = <\{A_i; i=1,...,10\}>$ where

$$A_1 = \big\{\max_i z_i \le F_{\max_i z_i}^{-1}(0.1)\big\},$$
$$A_i = \big\{F_{\max_k z_k}^{-1}\big(0.1(i-1)\big) < \max_k z_k \le F_{\max_k z_k}^{-1}(0.1 i)\big\}$$

for $i = 2,...,9$; and $A_{10} = \big\{\max_k z_k > F_{\max_k z_k}^{-1}(0.9)\big\}$; and

c) *The linear correlation measure* $\rho_{i,j} = O_{2,\mathfrak{I}_2}\big(z_i, z_j\big)$ and *its conditional version* $\rho_{i,j} = O_{2,\mathfrak{I}_2}\big(z_i, z_j \mid x'\mathbf{z} \le F_{x'z}^{-1}(0.05)\big)$, where $x'\mathbf{z} = \frac{1}{N}\sum_{i=1}^{N} z_i$ and $\mathfrak{I}_2 = <\{A_i; i=1,...,40\}>$ where $A_1 = \big\{\max_i z_i \le F_{\max_i z_i}^{-1}(0.025)\big\}$, $A_i = \big\{F_{\max_k z_k}^{-1}\big(0.025(i-1)\big) < \max_k z_k \le F_{\max_k z_k}^{-1}(0.025 i)\big\}$ for $i = 2,...,39$; and

$$A_{40} = \big\{\max_k z_k > F_{\max_k z_k}^{-1}(0.975)\big\}.$$

Thus, we suggest using different linear correlation measures to perform a PCA that identifies the main portfolio factors whose dispersion is significantly different from zero. These factors are then used to approximate the portfolio returns in large-scale portfolio selection problems. Therefore, using more than 1,300 assets of the US stock market, we analyse the results obtained by the PCA applied to different linear correlation matrices. Then, we compare the performances of some large-scale portfolio selection strategies based on different return approximations.

In this context, we propose new portfolio optimisation models that account for logical investor behaviour in two ways:

a) Investors want to maximise the concordance and/or association with the upper stochastic bound of the market; and

b) Investors want to minimise the concordance and/or association with the lower stochastic bound of the market.

These results are provided in Table 1. We can observe very small differences between the $O_{2,\mathfrak{I}_1}$ and $O_{2,\mathfrak{I}_2}$ measures, including both conditional and unconditional versions. Moreover, we need more than 12% (22%)[4] of the principal components to explain most of the variability using the Kaiser rule applied to the conditional (unconditional) measure. Therefore, the number of principal components selected with the Kaiser rule is still too big to apply a portfolio selection of type (14) to the approximated portfolio of gross returns. In addition, fewer than 20 components are sufficient to explain more than 50% of the variability when we use conditional correlation measures, while using only 35 principal components for the unconditional (conditional) correlation measures can explain more than 40% (61%) of the variability.

---

[4] According to the results in Table 1, 12% is about 157/1304 and 22% is about 294/1304.

Table 1 Portfolio dimensionality reduction: Number of principal components used for the Kaiser rule and to explain the 50% of the variability with conditional and unconditional correlation measures. Similarly, percentages explained with the Kaiser rule and with the first 35 principal components using conditional and unconditional correlation measures.

Measure		cor	$O_{2,\Im_1}$	$O_{2,\Im_2}$
Unconditional	# PCs Kaiser rule	294	300	300
	% explained Kaiser r.	69.48%	69.55%	69.60%
	# PCs to explain 50%	97	102	102
	% explained 35 PCs	40.92%	40.15%	40.13%
Conditional	# PCs Kaiser rule	158	157	157
	% explained Kaiser r.	99.80%	99.73%	99.73%
	# PCs to explain 50%	19	18	18
	% explained 35 PCs	61.98%	63.03%	62.95%

Given these results, we want to compare the portfolio strategies considering two different approximations of gross returns: one based on 35 factors derived from the conditional and unconditional Pearson correlation measures and the second based on 35 factors derived from the conditional and unconditional $O_{2,\Im_1}$ correlation measures. Here we only include $O_{2,\Im_1}$, because we do not observe very big differences by using $O_{2,\Im_2}$. In both cases, we regress the gross return series on 35 factors f_j (i.e. $z_i = b_{i,0} + \sum_{j=1}^{35} b_{i,j} f_j + \varepsilon_i$) and we approximate the vector of gross returns $\hat{\mathbf{z}} \approx \hat{\mathbf{b}}_0 + \sum_{j=1}^{35} \hat{\mathbf{b}}_j f_j$ using ordinary least square estimates of parameters $\mathbf{b}_j$. In particular, in the first case the 35 factors f_j are obtained as follows:

- The first 19 principal components obtained using the conditional Pearson correlation measure (since they explain at least 50% of the conditional dispersion) and the first 16 principal components obtained using the unconditional Pearson correlation measure.

Similarly, in the second case we combine:

- The first 18 principal components obtained using the conditional $O_{2,\Im_1}$ correlation measure (that explain at least 50% of the conditional dispersion) and the first 17 principal components obtained using the unconditional $O_{2,\Im_1}$ correlation measure.

Using an approximation of gross returns $\hat{\mathbf{b}}_0 + \sum_{j=1}^{35} \hat{\mathbf{b}}_j f_j$ the randomness of the choices is uniquely determined by the 35 factors f_j. We propose

an ex-post portfolio comparison where the portfolio decision is taken using the approximated returns and the valuation of future wealth is made using real ex-post gross returns.

In particular, we recalibrate the portfolio every six months (125 working days) using the daily approximated observations over the past 10 years (i.e. 2,600 working days). Recall that the objective function has already been given in (14) with a reward measure $v(x'\hat{z})$ as follows:

$$v(x'z) = \mathrm{E}\left[x'\mathbf{z} \mid x'\mathbf{z} \geq F_{x'z}^{-1}(0.95) \right] \quad (15)$$

Moreover, the risk measure

$$q(x'\hat{\mathbf{z}}) = AVaR_{0.05}(x'\hat{\mathbf{z}} - \mathrm{E}(x'\hat{\mathbf{z}})) \quad (16)$$

is the average value at risk of the centred approximated portfolio, see Rockafellar et al. (2006).

We consider three possible strategies based on different association measures. The first strategy (SpeaAVaR) is based on the Spearman concordance measure (i.e. $\rho_1 = \rho_2 = \rho_S$); the second strategy (PearAVaR) is based on the Pearson correlation measure (i.e. $\rho_1 = \rho_2 = \mathrm{cor}$); and the third strategy (KendAVaR) uses as ρ the Kendall correlation measure (i.e. $\rho_1 = \rho_2 = \tau_K$).

Then, we value the effects in the portfolio selection of the two different dimensional reductions. In particular, we compare the ex-post wealth sample paths of the three strategies obtained considering either the dimensional reduction derived from the (conditional and unconditional) Pearson correlation measures or the reduction derived from the (conditional and unconditional) $O_{2,\Im_1}$ correlation measures. The results of this comparison are reported in Figure 1.

According to the particular results, it is apparent that in general it is better to reduce the dimensionality using $O_{2,\Im_1}$ (instead of using Pearson linear correlation). Thus, it makes sense to introduce other linear correlations into the PCA and its usage in the portfolio reduction of dimensionality.

Moreover, although Kendall provides the best ex-post wealth, it also presents the highest ex-post variability. In this sense, the Spearman strategy provides a better performance in terms of the ex-post Sharpe ratio. From the figure, we can also observe that there is a strong impact of both recent crises; in particular, the usage of Kendall τ in the portfolio strategy gives us the highest sensitivity.

5. Conclusion

In this paper, we discussed when and how association measures can be used in portfolio problems. First, we
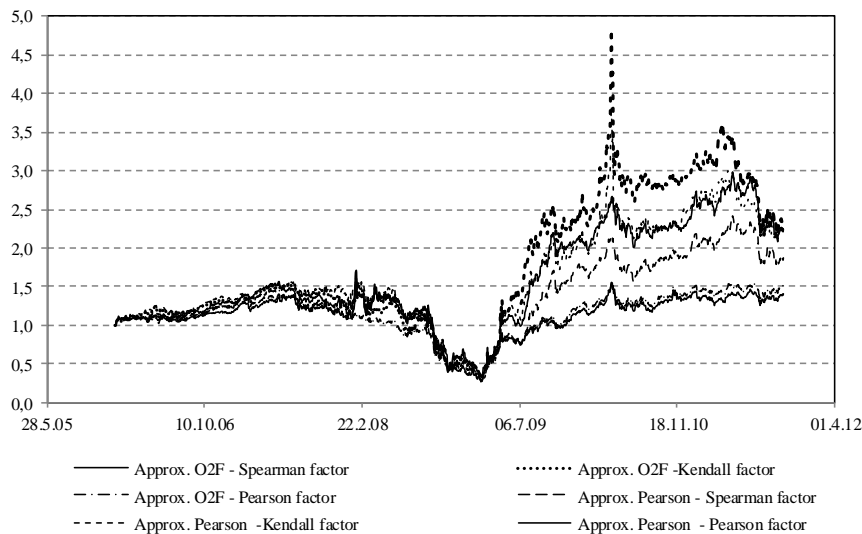
Figure 1 Ex-post final wealth after dimensionality reduction

Ex-post final wealth obtained using three portfolio strategies (PearAVaR, SpeaAVaR, KendAVaR) considering returns approximated either with the dimensional reduction derived from the Pearson correlation measure (namely, Approx. Pear) or with the reduction derived from the $O_{2,3_1}$ correlation measures (namely, Approx. O2F).

characterised the semidefinite positive association measures. Then, we discussed the use of association measures concerning portfolio performance optimisation and portfolio dimensionality reduction.

In particular, we formulated new portfolio selection problems using the association between portfolios and the stochastic bounds of the market; this was based on the approximated portfolio of returns. Finally, we carried out an empirical experiment showing the impact of particular measures of the dependency on (*i*) dimensionality reduction and (*ii*) portfolio performance. As the main results, we documented that the strategies using Kendall and Spearman present much better performances compared with similar strategies based on Pearson linear correlation.

Clearly this preliminary empirical analysis is the starting point for several new theoretical and empirical studies which are necessary in order to understand the class of linear correlation measures and their connection with the portfolio selection problem.

References

BIGLOVA, A., ORTOBELLI, S., RACHEV, S., FABOZZI, F. (2009). Modeling, estimation, and optimization of equity portfolios with heavy-tailed distributions. In: Satchel, S. (ed.). *Optimizing Optimization: The Next Generation of Optimization Applications and Theory*. Amsterdam: Academic Press, 117–141.

CHERUBINI, U., LUCIANO, E., VECCHIATO, W. (2004). *Copula Methods in Finance*. Chichester: Wiley.

FAMA, E. (1965). The behavior of stock market prices. *Journal of Business* 38: 34–105. http://dx.doi.org/10.1086/294743

KONDOR, I., PAFKA, S., NAGY, G. (2007). Noise sensitivity of portfolio selection under various risk measures. *Journal of Banking and Finance* 31: 1545–1573. http://dx.doi.org/10.1016/j.jbankfin.2006.12.003

MANDELBROT, B. (1963a). New methods in statistical economics. *Journal of Political Economy* 71: 421–440. http://dx.doi.org/10.1086/258792

MANDELBROT, B. (1963b). The variation of certain speculative prices. *Journal of Business* 26: 394–419. http://dx.doi.org/10.1086/294632

NELSEN, R.B. (2006). *An Introduction to Copulas*. New York: Springer.

PAPP, G., PAFKA, S., NOWAK, M.A., KONDOR, I. (2005). Random matrix filtering in portfolio optimization. *ACTA Physica Polonica B* 36: 2757–2765.

RACHEV, S., MITTNIK, S. (2000). *Stable Models in Finance*. Chichester: Wiley.

RACHEV, S., ORTOBELLI, S., STOYANOV, S., FABOZZI, F., BIGLOVA A. (2008). Desirable properties of an ideal risk measure in portfolio theory. *International Journal of Theoretical and Applied Finance* 11: 19–54. http://dx.doi.org/10.1142/S0219024908004713

ROCKAFELLAR, R.T., URYASEV, S., ZABARANKIN, M. (2006). Generalized deviations in risk analysis. *Finance and Stochastics* 10: 51–74. http://dx.doi.org/10.1007/s00780-005-0165-8

SCARSINI, M. (1984). On measures of concordance. *Stochastica* 8: 201–218.

SKLAR, A. (1959). Fonctions de réparitition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de Universitè de Paris* 8: 229–231.