Symposium on Big Data in Finance, Retail and Commerce: Statistical and Computational Challenges
Lisbon, 2-3 November 2017, Portugal

Forecasting conditional covariance matrix via principal volatility components in the presence of additive outliers.

Carlos Trucíos (Speaker)
São Paulo School of Economics - FGV, *ctrucios@gmail.com*

Luiz K. Hotta
Department of Statistics - University of Campinas, *hotta@ime.unicamp.br*

Pedro Valls
São Paulo School of Economics - FGV, *pedro.valls@fgv.br*

Abstract
In this work, we analyse a recently procedure called principal volatility components. This procedure overcome several difficulties in modelling and forecasting the conditional covariance matrix in large dimensions. We show that outliers have a devastating effect on the construction of the principal volatility components and on the forecast of the conditional covariance matrix. We propose a robust procedure and analyse its finite sample properties by means of Monte Carlo experiments and present an empirical application. The robust procedure outperforms the classical method in contaminated series and has a similar performance in uncontaminated ones.

Keywords: Conditional covariance matrix; Constant volatility; Curse of dimensionality; Jumps; Principal components.

Introduction
Based on the idea that comovements in the market can be driven by a few components, factor models appear in the economic and financial literature as a way to achieve dimension reduction and tackling the curse of dimensionality. In this spirit, an innovative approach based on the classical principal component analysis (PCA), so-called principal volatilty components (PVC), has been recently proposed by [1] and [2]. This methodology produces two types of components. The first type corresponds to components with conditional covariance matrix evolving over time while the other type corresponds to components with constant conditional covariance matrix.

Several works show that additive outliers (AO) affect dramatically the forecast of (co)volatilities and consequently financial applications ([3], [4], [5]). Furthermore, there are evidence showing that PCA is very sensitive to the presence of AO ([6], [7]). Thus, procedures based on similar methodology are expected to be sensitive to outliers too.
We first analyse by means of Monte Carlo experiments the performance of the PVC in the presence of AO showing that outliers have a devastating effect on this procedure, even when moderate outliers are present. Then, we propose a robust procedure with good finite sample properties.

Principal volatility components
For $t=1,...,T$, let $y_t=(y_{1,t},...,y_{N,t})$ be a N-dimensional random vector weakly stationary with $E\left(y_t|\mathcal{F}_{t-1}\right)=0$ and T the sample size. Consider $\hat{\Gamma}_r M=\Lambda M$, where Λ is a decreasing ordered diagonal matrix of eigenvalues, M the associated normalized eigenvectors and $\hat{\Gamma}_r$ is defined as

$$\hat{\Gamma}_r=\sum_{k=1}^{g}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(1-\frac{k}{T}\right)^2\left[\frac{1}{T}\sum_{t=k+1}^{T}\left[\left(y_ty_t'-\hat{\Sigma}\right)\left(x_{ij,t-k}-\bar{x}_{ij}\right)\right]\right]^2, \quad (1)$$

with $\hat{\Sigma}$ being the sample covariance matrix, $\bar{x}_{ij}$ the sample mean of $x_{ij,t}$, where $x_{ij,t}$ is a function of $y_{i,t}y_{j,t}$ and g is a lag order (usually 5 or 10).
Recently, [2] proposes an alternative PVC, called Generalized PVC which requires only finite second order moment. This procedure replace (1) by

$$\hat{G}=\sum_{k=1}^{g}\sum_{\tau=1}^{T}\omega(y_\tau)\left[\frac{1}{T-k}\sum_{t=k+1}^{T}\left[\left(y_ty_t'-\hat{\Sigma}\right)I(\|y_{t-k}\|\leq\|y_\tau\|)\right]\right]^2, \quad (2)$$

where $\omega(\cdot)$ is a weight function and $\|\cdot\|$ is the L_1 norm

Robust principal volatility components (RPVC)
We will show in the next section that PVC and GPVC are not robust to AO. Thus, in order to obtain a procedure less sensitive to AO, we robustify the estimator given in (2). The robust procedure is based on a robust estimator of the unconditional covariance matrix and a weighted estimator of $E\left[(y_ty_t'-\Sigma)I(\|y_{t-k}\|\leq\|y_t\|)\right]$. In our robust proposal, we replace (2) by

$$\hat{G}^R=\sum_{k=1}^{g}\sum_{\tau=1}^{T}\omega(y_\tau)\left[\sum_{t=k+1}^{T}\delta^*(d_t^2)\left\{(y_ty_t'-\hat{\Sigma}^R)I(\|y_{t-k}\|\leq\|y_\tau\|)\right\}\right]^2, \quad (3)$$

where $d_t^2=(y_t-\hat{\mu}^R)'\hat{\Sigma}_t^{-1}(y_t-\hat{\mu}^R)$ with $\hat{\Sigma}_t=0.01\rho(y_{t-1}'y_{t-1})+0.99\hat{\Sigma}_{t-1}$, $\hat{\Sigma}_1=\hat{\Sigma}^R$, $\hat{\mu}^R$ and $\hat{\Sigma}^R$ are robust estimates of μ and Σ and $\delta^*(\cdot)=\delta(\cdot)/||\delta(\cdot)||$. Finally, $\rho(\cdot)$ and $\delta(\cdot)$ are given by

$$\rho(x) = \begin{cases} x, & \text{if } x \le c, \\ \hat{\Sigma}^R, & \text{if } x > c, \end{cases} \qquad \delta(x) = \begin{cases} 1, & \text{if } x \le c, \\ \frac{1}{x}, & \text{if } x > c. \end{cases}$$

Monte Carlo experiments

We analyse the effects of AO on the one-step-ahead prediction of the conditional covariance by Monte Carlo experiments with 1000 replicates. Following [1] and [2], we simulate a factor model driven by just one common factor which follows a GARCH(1,1) with parameters $\omega = 1$, $\alpha = 0.07$ and $\beta = 0.83$. The idiosyncratic factors were simulated following a multivariate standard Normal distribution. Fig. 1 reports the MSE of the predicted conditional covariance matrix for uncontaminated and contaminated series by isolated ($t = 500$ and 999) and consecutive ($t = 500$ and 501, 998 and 999) outliers. The results show a devastating impact of outliers on the forecasting of the conditional covariance matrix when the PVC and GPVC procedures are used but not in the robust procedure. Similar results are observed when other performance measures are used.

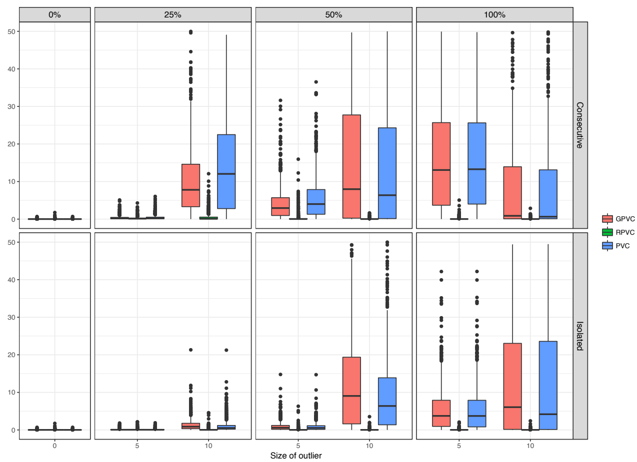

Figure 1: Boxplot of MSE for uncontaminated (0%) and contaminated series with 25%, 50% and 100% of series contaminated at time t. $T = 1000$ and outliers of size $\omega = 0$, 5 and 10 standard deviations of the univariate uncontaminated process.

We also applied the RPVC to 73 stock returns of the Nasdaq 100 index and construct the minimum variance portfolio. RPVC shows a better performance in terms of annualized performance measures compared with the obtained using the non-robust version.

Conclusions

We analyse the problem of modelling and forecast the conditional covariance matrix via principal volatility components in the presence of AO and show that a few AO are sufficient to affect drastically the volatility components and the estimation of the conditional covariance matrix.

A new procedure called robust principal volatility components based on a robust estimator of the unconditional covariance matrix and on a weighted estimator is proposed. The superiority of our procedure has been shown in both empirical and simulated data.

Acknowledgments

The authors acknowledge financial support from São Paulo Research Foundation (FAPESP) grants 2016/18599-4, 2013/00506-1 and 2013/22930-0. The third author is also grateful with the National Council for Scientific and Technological Development (CNPq) grant 309158/2016-8. Support from Laboratory EPIFISMA is also grateful.

References

[1] Hu, Y. P., & Tsay, R. S. (2014). Principal volatility component analysis. *Journal of Business & Economic Statistics*, 32(2), 153-164.

[2] Li, W., Gao, J., Li, K., & Yao, Q. (2016). Modeling multivariate volatilities via latent common factors. *Journal of Business & Economic Statistics*, 34(4), 564-573.

[3] Vaz de Melo Mendes, B., & Pereira Câmara Leal, R. (2005). Robust multivariate modeling in finance. *International Journal of Managerial Finance*, 1(2), 95-106.

[4] Trucíos, C., & Hotta, L. K. (2016). Bootstrap prediction in univariate volatility models with leverage effect. *Mathematics and Computers in Simulation*, 120, 91-103.

[5] Trucíos, C., Hotta, L. K., & Ruiz, E. (2017). Robust bootstrap densities for dynamic conditional correlations: Implications for portfolio selection and value-at-risk. *Working paper available at SSRN: https://ssrn.com/abstract=2969908*

[6] Croux, C., & Haesbroeck, G. (2000). Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies. *Biometrika*, 87(3), 603-618.

[7] Hubert, M., Rousseeuw, P. J., & Vanden Branden, K. (2005). ROBPCA: a new approach to robust principal component analysis. *Technometrics*, 47(1), 64-79.