S.Co. 2009

S.Co.2009

SIXTH CONFERENCE

COMPLEX DATA MODELING AND COMPUTATIONALLY INTENSIVE STATISTICAL METHODS FOR ESTIMATION AND PREDICTION

POLITECNICO DI MILANO, SEPTEMBER 14-16,2009

PROCEEDINGS

Edited by:

A.M. Paganoni, L.M. Sangalli, P. Secchi, S. Vantini

politecnica

MAGGIOLI EDITORE

ISBN 978-88387-4385-1

Maggioli Editore è un marchio di Maggioli S.p.A.
Azienda con sistema qualità certificato ISO 9001:2000

47822 Santarcangelo di Romagna (RN) • Via del Carpino, 8
Tel. 0541/628111 • Fax 0541/622020

www.maggioli.it/servizioclienti
e-mail: servizio.clienti@maggioli.it

Finito di stampare nel mese di agosto 2009
da DigitalPrint Service s.r.l., via Torricelli, 9 – Segrate (Milano)

S.Co. 2009

SIXTH CONFERENCE

COMPLEX DATA MODELING AND COMPUTATIONALLY INTENSIVE STATISTICAL METHODS FOR ESTIMATION AND PREDICTION

POLITECNICO DI MILANO, SEPTEMBER 14-16, 2009

PROCEEDINGS

Edited by:
A.M. Paganoni, L.M. Sangalli, P. Secchi, S. Vantini

MAGGIOLI EDITORE

ESTIMATING VALUE-AT-RISK WITH PRODUCT PARTITION MODELS

Danilo Delpini[1,2], Giacomo Bormetti[2,3], Maria Elena De Giuli[4] and Claudia Tarantola[4]

[1] Dipartimento di Fisica Nucleare e Teorica
Università di Pavia
Via A. Bassi 6, 27100 Pavia, Italy
(e-mail: danilo.delpini@pv.infn.it)
[2] Istituto Nazionale di Fisica Nucleare
Sezione di Pavia
Via A. Bassi 6, 27100 Pavia, Italy
[3] Centro Studi Rischio e Sicurezza
Istituto Universitario di Studi Superiori
Viale Lungo Ticino Sforza 56, 27100 Pavia, Italy
[4] Dipartimento di Economia Politica e Metodi Quantitativi
Università di Pavia
Via San Felice 5, 27100 Pavia, Italy

ABSTRACT. We consider Bayesian estimation of Value-at-Risk (VaR) using parametric Product Partition Models (PPM). VaR is a standard tool to measure and control the market risk of an asset or a portfolio, and it is also required for regulatory purposes. We use PPM to provide robustly Bayesian estimators of VaR, remaining in a Normal setting, even in presence of outlying points. We consider two different scenarios: a product partition structure on the vector of means and a product partition structure on the vector of variances. In both frameworks we obtain a closed-form expression for VaR. The results are illustrated with an application to a set of shares from the Italian stock market. The methodology and the obtained results are described in details in Bormetti *et al.* (2009).

1 BACKGROUND AND PRELIMINARIES

Following the increase in financial uncertainty, there has been intensive research from financial institutions, regulators and academics to develop models for market risk evaluation. A common and easily understood measure of risk is VaR.

VaR is referred to the probability of extreme losses due to adverse market movements. In particular, for a given significance level α (typically 1% or 5%), VaR is defined as the maximum potential loss over a fixed time horizon of individual assets or portfolios of assets as well. For normally independent and identically distributed (i.i.d.) returns (with mean μ and variance σ^2), a closed-form expression for VaR normalised to the spot price is given by

$$\frac{\Lambda}{W_0} = -\mu + \sigma\sqrt{2}\,\mathrm{erfc}^{-1}(2\alpha),$$

where Λ is VaR, W_0 is the spot price and erfc^{-1} is the inverse of the complementary error function. In the following, with VaR we shall refer to the normalized VaR, Λ/W_0, if not otherwise specified. If this quantity is expressed in percentage term we name it percentage VaR, indicated with VaR(%).

For low liquidity markets and short time horizons, the normal i.i.d. assumption fails to be effective and has to be relaxed. Possible solutions are to resort to heavy tailed distributions or to abandon the hypothesis of identically distributed returns. We follow the latter approach and we use a Bayesian methodology based on parametric PPMs. We assume all the returns being normally distributed with a partition structure on the parameters of interest. We assign a prior distribution on the space of all possible partitions and we identify clusters of returns sharing the same mean and variance values. Returns belonging to different clusters are characterised by different values either of the mean or the variance. The hypothesis of identical distribution holds within but non between clusters.

2 OUR MODEL

Let $\mathbf{y}=(y_1,\ldots,y_t,\ldots y_T)$ denote the vector of returns of a generic asset at different time points t. The returns are independent, and jointly distributed with probability density function f parameterised by the vector $(\boldsymbol{\theta},\psi)$. The elements of $\boldsymbol{\theta}$ depend on the time point t, $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_T)$, whereas ψ is a parameter that is common to all observations. In the following we consider the model

$$\mathbf{y}|(\boldsymbol{\theta},\psi)\sim(\mathbf{y}|\boldsymbol{\theta},\psi),\quad \text{with}\quad y_t \overset{ind}{\sim} (y_t|\theta_t,\psi)\quad t=1,\ldots,T. \tag{1}$$

Given the model in (1), let $S_0=\{t:t=1,\ldots,T\}$ be the set of time periods. A partition of the set S_0, $\rho=\{S_1,\ldots,S_d,\ldots,S_{|\rho|}\}$ with cardinality $|\rho|$, is defined by the property that $S_d\cap S_{d'}=\emptyset$ for $d\neq d'$ and $\cup_d S_d=S_0$. The generic element of ρ is $S_d=\{t:\theta_t=\theta^*_d\}$, where $\boldsymbol{\theta}^*=\left(\theta^*_1,\ldots,\theta^*_{|\rho|}\right)$ is the vector of the unique values of $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_T)$. All θ_t whose subscripts t belong to the same set $S_d\in\rho$ are (*stochastically*) equal, in this sense they are regarded as a single *cluster*.

We assign to each partition ρ the following prior distribution

$$P\left(\rho=\{S_1,\ldots,S_{|\rho|}\}\right)=K\prod_{d=1}^{|\rho|}C(S_d)=K\prod_{d=1}^{|\rho|}c\times(|S_d|-1)!\,, \tag{2}$$

where $C(S_d)$ is a cohesion function, c is a positive parameter, $|S_d|$ denotes the cardinality of the set S_d and K is the normalising constant. Equation (2) is referred to as the *product distribution* for partitions. A moderate value of c, e.g. $c=1$, favours the formation of partitions with a reduced number of large subsets.

We present and compare two different PPMs; in the first one we impose a partition structure on the vector of means, and in the second one we consider partitions on the vector of variances. In the following the PPM applied to the vector of means will be shortly referred to as the $\boldsymbol{\mu}$-PPM approach, while $\boldsymbol{\sigma}^2$-PPM will refer to the PPM for the vector of variances. In $\boldsymbol{\mu}$-PPM the vector $\boldsymbol{\theta}$ is the vector of means while in $\boldsymbol{\sigma}^2$-PPM it corresponds to the vector of variances. In the former model ψ is the variance and in the latter it corresponds to the mean.

For comparative reason, we also consider the σ^2-CP model proposed by Loschi *et al.* (2003) where PPM are used to identify change points in the volatility time series.

We consider the following hierarchical structure

$$y_t|(\rho,(\theta_1^*,\ldots,\theta_{|\rho|}^*),\sigma^2) \overset{ind.}{\sim} N(y_t|(\theta_t,\psi)) ,$$
$$\theta_1^*,\ldots,\theta_{|\rho|}^*|(\rho,\psi) \overset{i.i.d.}{\sim} f(\cdot|\psi) ,$$
$$\rho \sim \text{product distribution, with } C(S_d) = c \times (|S_d|-1)! ,$$
$$\psi \sim g(\psi) ,$$

where f and g denote generic density functions and the product distribution is defined in equation (2).

To obtain a sample from the posterior distribution of the parameter of interested we apply specific Gibbs sampling algorithm, see Bormetti *et al.* (2009) for the details.

3 Bayesian estimation of VaR

We now present how the posterior distribution of VaR and consequently its Bayesian estimate can be obtained by using the output of the Gibbs algorithms. First we focus our attention on the PPM on the vector of means. Let indicate with $\boldsymbol{\mu}^*_{(\ell)} = \left(\mu^*_{1(\ell)},\ldots,\mu^*_{|\rho|(\ell)}\right)$ and $\sigma^2_{(\ell)}$ respectively the vector of means and the variance sampled at the ℓ-th iteration of the algorithm. At each iteration we obtain a peculiar clustering structure. All returns share the same value of $\sigma^2_{(\ell)}$, but each cluster is characterized by a different value $\mu^*_{d(\ell)}$. In order to provide a single VaR estimate for each iteration of the chain we propose to average the different entries of $\boldsymbol{\mu}^*_{(\ell)}$ and we consider the following equation

$$\frac{\Lambda_{(\ell)}}{W_0} = -\sum_{d=1}^{|\rho|} \frac{|S_{d(\ell)}|}{T}\mu^*_{d(\ell)} + \sigma_{(\ell)}\sqrt{2}\,\text{erfc}^{-1}(2\alpha) . \qquad (3)$$

If we impose a clustering structure over the vector of variances, VaR can be computed in an analogous way but the arithmetic average is performed over different values of $\sigma^*_{d(\ell)}$, that is

$$\frac{\Lambda_{(\ell)}}{W_0} = -\mu_{(\ell)} + \sum_{d=1}^{|\rho|} \frac{|S_{d(\ell)}|}{T}\,\sigma^*_{d(\ell)}\sqrt{2}\,\text{erfc}^{-1}(2\alpha) . \qquad (4)$$

In this case all returns share the same value of $\mu_{(\ell)}$ but each cluster is characterised by a different value of $\sigma^*_{d(\ell)}$. The resulting VaR estimate is obtained as the ergodic mean of the quantities $\Lambda_{(\ell)}$ in (3) or (4) for $\boldsymbol{\mu}$-PPM or σ^2-PPM respectively, by means of

$$\frac{\Lambda}{W_0} = \frac{1}{L}\sum_{\ell=1}^{L}\frac{\Lambda_{(\ell)}}{W_0}.$$

VaR under the σ^2-CP model is computed in a similar way.

4 Real data analysis

We test our models over the MIB30 index and its three components with the highest excess of kurtosis, where standard approaches based on Normal distributions usually fail. In particular we apply our analysis to the Italian assets Lottomatica (LTO.MI), Mediobanca (MB.MI) and Snam Rete Gas (SRG.MI). We consider time series of daily returns from April 2004 to March 2008. All time series are made of 1000 daily returns freely downloadable from the site *http://it.finance.yahoo.com*.

In the examples below we run 10000 MC sweeps with 10% burnin. The parameters are fixed as follows: $c = 1$, $m = 0$ (for short time horizon, typically from one day until one week, the value of the mean is usually neglected,), $\tau_0^2 = 10^3$, $\lambda_0 = 0.0101$ and $\nu_0 = 2.01$. This choice for the Inverted Gamma parameters reflects what is known from the past experience about the volatility behaviour for equity assets. For the $\boldsymbol{\sigma}^2$-CP model that we use as yardstick model we set the priors parameters following Loschi *et al.* (2003). In particular we consider the conjugate Normal-Inverted-Gamma model, with the probability p that a change occurs at any instant in the sequence equal to 0.1.

In table 1 we report Bayesian estimates of percentage VaR for $\alpha = 1\%$ and $\alpha = 5\%$ and the 68% posterior credible interval.

VaR(%)	α=5%			α=1%		
	$\boldsymbol{\mu}$-PPM	$\boldsymbol{\sigma}^2$-PPM	$\boldsymbol{\sigma}^2$-CP	$\boldsymbol{\mu}$-PPM	$\boldsymbol{\sigma}^2$-PPM	$\boldsymbol{\sigma}^2$-CP
MIB30.MI	$1.45^{+0.05}_{-0.05}$	$1.74^{+0.11}_{-0.12}$	$1.76^{+0.01}_{-0.01}$	$2.07^{+0.06}_{-0.06}$	$2.48^{+0.16}_{-0.17}$	$2.49^{+0.01}_{-0.01}$
LTO.MI	$2.08^{+0.07}_{-0.07}$	$2.78^{+0.15}_{-0.16}$	$2.66^{+0.02}_{-0.02}$	$2.95^{+0.09}_{-0.09}$	$3.94^{+0.21}_{-0.21}$	$3.78^{+0.03}_{-0.03}$
MB.MI	$1.91^{+0.07}_{-0.08}$	$2.40^{+0.12}_{-0.12}$	$2.36^{+0.01}_{-0.01}$	$2.72^{+0.11}_{-0.11}$	$3.40^{+0.17}_{-0.17}$	$3.35^{+0.02}_{-0.02}$
SRG.MI	$1.58^{+0.05}_{-0.05}$	$1.97^{+0.12}_{-0.13}$	$2.01^{+0.01}_{-0.01}$	$2.26^{+0.06}_{-0.06}$	$2.81^{+0.17}_{-0.17}$	$2.87^{+0.02}_{-0.02}$

Table 1. Daily estimated VaR (%) values at 5% and 1% significance level with 68% credible intervals.

The estimates of VaR obtained with $\boldsymbol{\sigma}^2$-PPM and $\boldsymbol{\sigma}^2$-CP are in good agreement even if the two approaches are quite different in spirit. The former approach is a natural extension of the $\boldsymbol{\mu}$-PPM to the vector of variances while the latter one is specific for change point identification. The PPM on the vector of means in general underestimates VaR with respect to the values given by the PPMs applied to the variances. This fact can be empirically justified by noticing that for daily time horizons the contribution to VaR due to the volatility σ is of order ten greater than that due to the mean μ. Figure 1 depicts posteriors distributions for VaR estimates at level $\alpha = 1\%$. In the first row we present the results based on the $\boldsymbol{\mu}$-PPM approach, while the second corresponds to $\boldsymbol{\sigma}^2$-PPM. The posterior distribution of VaR presents a higher variability under the $\boldsymbol{\sigma}^2$-PPM approach than under the $\boldsymbol{\mu}$-PPM VaR one. The posterior expectation of the number of clusters is low for both the $\boldsymbol{\mu}$-PPM and $\boldsymbol{\sigma}^2$-PPM approaches and, moreover, the partitions are characterised by a very large cluster and few small ones. The results are presented in table 2.

The arithmetic average in equations (3) and (4) is therefore dominated by the values of $\mu^*_{d(\ell)}$ and $\sigma^{2*}_{d\,(\ell)}$ that correspond to the largest cluster, while outlying clusters introduce corrections to VaR.

	Number of Clusters		Largest Cluster Weight	
	μ-PPM	σ^2-PPM	μ-PPM	σ^2-PPM
MIB30.MI	3.11	3.39	0.986	0.990
LTO.MI	5.02	4.52	0.963	0.944
MB.MI	4.11	3.72	0.968	0.970
SRG.MI	3.44	3.59	0.984	0.978

Table 2. Posterior mean of the number of clusters and relative weight of the largest cluster for μ-PPM and σ^2-PPM.

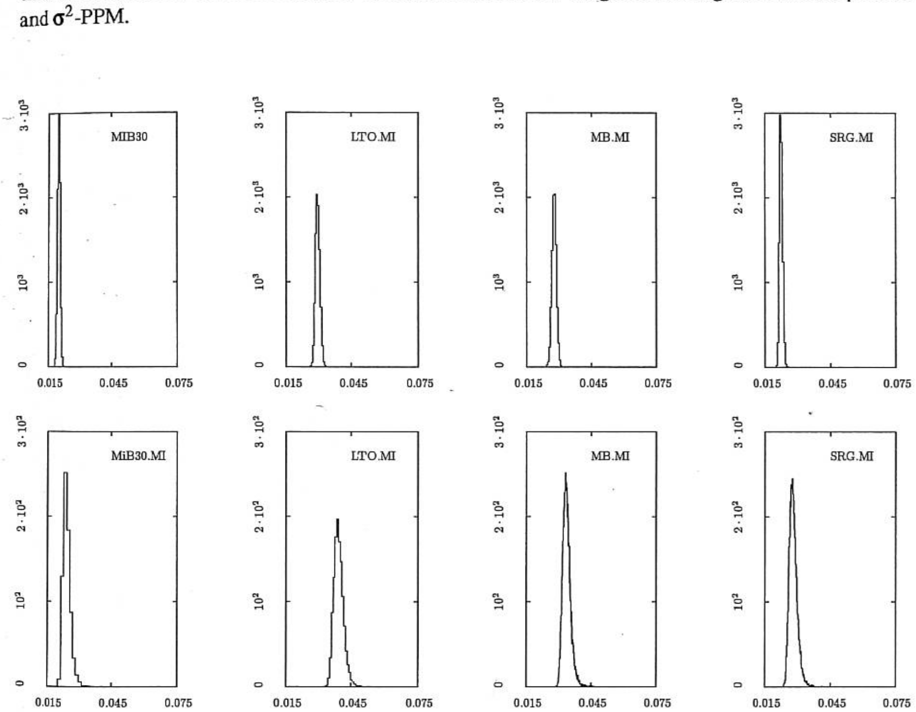

Figure 1. VaR posterior distribution for $\alpha = 1\%$, for μ-PPM (first row) and σ^2-PPM (second row).

We now compare our results with those obtained with standard parametric approaches based on ML estimators for the mean and variance. In particular we consider the results obtained with a Normal model and with the generalised Student-*t* (GST) distribution, see Bormetti *et al.* (2007). In the GST we set the tail index $\nu > 2$, in order to keep the variance finite, see last column of table 3. In the following we consider the GST as the benchmark for our analysis since it presents a good agreement with historical simulations. Numerical details are reported in tables 1 and 3. At $\alpha = 1\%$ the results obtained with σ^2-PPM and σ^2-CP are the ones in best agreement with the GST distribution, while Normal and μ-PPM underestimate VaR. The situation is different if we consider $\alpha = 5\%$. In this case μ-PPM is the only one in agreement with the GST distribution, while σ^2-PPM and σ^2-CP overestimate VaR.

VaR(%)	α=5%		α=1%		
	Normal	Student-*t*	Normal	Student-*t*	ν
MIB30.MI	$1.38^{+0.05}_{-0.07}$	$1.27^{+0.04}_{-0.06}$	$1.95^{+0.07}_{-0.09}$	$2.22^{+0.09}_{-0.10}$	$4.16^{+0.43}_{-0.48}$
LTO.MI	$2.50^{+0.14}_{-0.15}$	$2.15^{+0.08}_{-0.09}$	$3.55^{+0.19}_{-0.20}$	$4.05^{+0.19}_{-0.22}$	$3.26^{+0.28}_{-0.30}$
MB.MI	$2.07^{+0.06}_{-0.09}$	$1.89^{+0.05}_{-0.07}$	$2.95^{+0.09}_{-0.11}$	$3.37^{+0.12}_{-0.14}$	$3.93^{+0.35}_{-0.38}$
SRG.MI	$1.62^{+0.05}_{-0.08}$	$1.48^{+0.04}_{-0.07}$	$2.32^{+0.08}_{-0.10}$	$2.65^{+0.11}_{-0.12}$	$3.97^{+0.44}_{-0.44}$

Table 3. Daily ML estimated VaR(%) values at 5% and 1% significance level with 68% bootstrap intervals. In the last column we report central value and 68% bootstrap interval for the tail index ν of the GST.

5 CONCLUDING REMARKS

In this paper we have presented a novel Bayesian methodology for VaR computation based on parametric PPMs. The main advantages of our approach are that it allows us to remain in the Normal setting, to identify anomalous observations and to obtain a closed-form expression for the VaR measure. This expression generalizes the standard parametric formula that is used in the literature under the normality assumption. By means of PPMs we induce a clustering structure over the vector of means ($\boldsymbol{\mu}$-PPM) and we find the best agreement with ML approaches for significance level of order 5%. For lower values of α we obtained the best result by applying the PPMs to the vector of variances (σ^2-PPM).

REFERENCES

ANTONIAK, C.E., (1974): Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist. 2*, 1152–1174.

BARRY, D., HARTIGAN, J.A., (1992): Product partition models for change point problems. *Ann. Statist. 20*, 260–279.

BORMETTI, G., CISANA, E., MONTAGNA, G., NICROSINI, O., (2007): A non-Gaussian approach to risk measures. *Physica A 376*, 532–542.

BORMETTI, G., DE GIULI, M.E., DELPINI, D., TARANTOLA, C. (2009): Bayesian Value-at-Risk with Product Partition Models. *Preprint* available at http://lanl.arxiv.org/abs/0809.0241.

HARTIGAN, J.A., (1990): Partition models. *Commun. Statist. Theory Methods 19*, 2745–2756.

LOSCHI, R.H., CRUZ, F.R.B., IGLESIAS, P.L., ARELLANO-VALLE, R.B., (2003): A Gibbs sampling scheme to the product partition model: An application to change-point problems. *Comp. Oper. Res. 30*, 463–482.

QUINTANA, F.A., IGLESIAS, P.L., (2003): Bayesian clustering and product partition models. *J. Roy. Statist. Soc. B 65*, 557–574.