Estimating large-scale multivariate local level models with application to stochastic volatility

Stima di modelli local level di grandi dimensioni con applicazioni ai modelli di volatilità stocastica

Matteo Pelagatti and Giacomo Sbrana

Abstract We derive the closed-form solution to the Riccati equation for the steady-state Kalman filter of the multivariate local linear trend model. Based on this result we propose a fast EM algorithm that provides approximated maximum likelihood estimates of the model's parameters and apply it to large-scale stochastic volatility models.

Abstract *Deriviamo la soluzione dell'equazione di Riccati relativa al filtro di Kalman in steady-state per il local linear trend multivariato. Utilizziamo tale risultato per proporre un algoritmo EM, che calcola un'approssimazione della stima di massima verosimiglianza dei parameteri del modello. Applichiamo tale metodo di stima a modelli di volatilità stocastica di grandi dimensione.*

Key words: Algebraic Riccati equation, state-space models, Kalman filtering, smoothing, EM, multivariate stochastic volatility.

1 Introduction

Multivariate stochastic volatility (SV) models are useful for portfolio managers only if they can be applied to portfolios of tens or hundreds of assets. Indeed, in the GARCH literature the most successful multivariate models are those, such as the Constant Conditional Correlation (CCC) and the Dynamic Conditional Correlation (DCC) that, by splitting the estimation processes is a sequence of computationally feasible steps, make the application to large portfolios possible. However, the SV

Matteo Pelagatti
Università degli Studi di Milano-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano, Italy
e-mail: matteo.pelagatti@unimib.it

Giacomo Sbrana
NEOMA Business School, 1 Rue du Marechal Juin, 76130 Mont-Saint-Aignan, France e-mail: giacomo.sbrana@neoma-bs.fr

literature is till lacking of a computationally feasible procedure to estimate large scale models, possibly dividing the estimation process into sub-steps.

In this work we consider the multivariate extensions of the SV models of Harvey et al (1994) and Alizadeh et al (2002), which have a state-space representation as multivariate local linear trend models, provide the steady-state Kalman filter recursions by solving the related Riccati equation in closed form and propose an EM algorithm that approximates the (quasi) maximum likelihood estimate of SV model.

2 Main results

It is well known that, for time-invariant state-space models, the Kalman filter eventually converges to the steady-state solution such that the error covariance matrix satisfies the so called algebraic Riccati equation. Explicit solutions for such matrix equation are in general not available. A notable exception is represented by the univariate local level model (see Harvey, 1989) because the Kalman filter covariance matrix reduces to a scalar. In this paper we show that an analytical solution exists also for the multivariate local level model (also known as multivariate exponential smoothing or exponentially weighted moving average or EWMA). In what follows we use this convention: $\boldsymbol{M}$ is a matrix and $\boldsymbol{M}'$ is its transpose, $\boldsymbol{m}$ is a column vector such that $\boldsymbol{m}'$ is a row vector. A lower-case letter, such as x, represents a scalar. Finally, 0 is used indiscriminately for matrices, vectors and scalars.

Consider the time-invariant state-space representation of the multivariate exponential smoothing process:

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \qquad & \boldsymbol{\varepsilon}_t &\sim \mathrm{WN}(\boldsymbol{0}, \boldsymbol{\Sigma}_\varepsilon) \\ \boldsymbol{\alpha}_{t+1} &= \boldsymbol{\alpha}_t + \boldsymbol{\eta}_t, \qquad & \boldsymbol{\eta}_t &\sim \mathrm{WN}(\boldsymbol{0}, \boldsymbol{\Sigma}_\eta) \end{aligned} \tag{1}$$

where WN denotes a white noise sequence. In what follows it is assumed that $\boldsymbol{\Sigma}_\varepsilon$ and $\boldsymbol{\Sigma}_\eta$ are symmetric positive definite matrices and $\mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\eta}_t') = 0$. Note that all vectors and matrices in (1) have dimension d and $d \times d$ respectively. The Kalman filter recursions for this model can then be written as follows:

Innovation : $\boldsymbol{v}_t = \boldsymbol{y}_t - \boldsymbol{a}_t$,
Innovation variance : $\boldsymbol{F}_t = \boldsymbol{P}_t + \boldsymbol{\Sigma}_\varepsilon$,
Kalman gain : $\boldsymbol{K}_t = \boldsymbol{P}_t \boldsymbol{F}_t^{-1}$,
Prediction : $\boldsymbol{a}_{t+1} = \boldsymbol{a}_t + \boldsymbol{K}_t \boldsymbol{v}_t$,
Prediction error : $\boldsymbol{P}_t = \boldsymbol{P}_t - \boldsymbol{P}_t \boldsymbol{F}_t^{-1} \boldsymbol{P}_t + \boldsymbol{\Sigma}_\eta$.

In the steady-state the the covariance matrix $\boldsymbol{P}_t$ converges to the so called algebraic Riccati equation, that is,

$$\boldsymbol{P} = \boldsymbol{P} - \boldsymbol{P}(\boldsymbol{P} + \boldsymbol{\Sigma}_\varepsilon)^{-1} \boldsymbol{P} + \boldsymbol{\Sigma}_\eta, \tag{2}$$

where $\boldsymbol{P}$ is a symmetric positive definite matrix. The following proposition provides the analytical (matrix) solution of (2), that is the algebraic link between $\boldsymbol{P}$ and the pair $\boldsymbol{\Sigma}_\varepsilon$, $\boldsymbol{\Sigma}_\eta$.

Theorem 1. *Consider the system as in (1) where $\boldsymbol{\Sigma}_\eta$ is positive semi-definite while $\boldsymbol{\Sigma}_\varepsilon$ is strictly positive definite. Moreover, consider the following Cholesky decomposition $\boldsymbol{\Sigma}_\varepsilon = \boldsymbol{M}\boldsymbol{M}'$ and defining $\boldsymbol{Q} = \boldsymbol{M}^{-1}\boldsymbol{\Sigma}_\eta \boldsymbol{M}^{-1'} = \boldsymbol{\Psi}\boldsymbol{\Delta}\boldsymbol{\Psi}'$ such that for the last eigendecomposition $\boldsymbol{\Psi}$ is a matrix of eigenvectors (i.e. $\boldsymbol{\Psi}\boldsymbol{\Psi}' = \boldsymbol{I}$) and $\boldsymbol{\Delta} =$* diag$(\delta_1, \delta_2, \cdots, \delta_d)$ *is a diagonal matrix of eigenvalues. Then there exists a unique positive definite solution for $\boldsymbol{P}$. The solution is*

$$\boldsymbol{P} = \frac{1}{2}\boldsymbol{M}\boldsymbol{\Psi}\left[\boldsymbol{\Delta} + (\boldsymbol{\Delta}^2 + 4\boldsymbol{\Delta})^{\frac{1}{2}}\right]\boldsymbol{\Psi}'\boldsymbol{M}'. \tag{3}$$

Going back to the Kalman filter recursion, it is immediate to see that, in steady state, the only step to compute is the *prediction step*, because all the other quantities are time-invariant. Also the implementation of the smoothing process results greatly simplified. Indeed, the smoothing algorithm proposed by de Jong (1988, 1989) (see also Ansley and Kohn, 1985; Koopman, 1997) becomes $\boldsymbol{r}_n = \boldsymbol{0}$, $\boldsymbol{N}_n = \boldsymbol{0}$,

$$\boldsymbol{r}_{t-1} = \boldsymbol{F}^{-1}\boldsymbol{v}_t + \boldsymbol{L}'\boldsymbol{r}_t \tag{4}$$

$$\boldsymbol{N}_{t-1} = \boldsymbol{F}^{-1} + \boldsymbol{L}'\boldsymbol{N}_t\boldsymbol{L} \tag{5}$$

where $\boldsymbol{L} = \boldsymbol{I} - \boldsymbol{K}$ is also time-invariant.

The maximum likelihood estimation of a model in state space form using the EM algorithm is fully discussed by Shumway and Stoffer (2017) (see also Koopman, 1993; Durbin and Koopman, 2001). In practice, the EM algorithm for the multivariate local level can be implemented using the simple updating expressions (3.5), (3.6) and (3.7) of Section 3 in Koopman (1993). More specifically, for model (1) these expressions can be restated as follows:

$$\boldsymbol{\Sigma}_\varepsilon(\iota+1) = \boldsymbol{\Sigma}_\varepsilon(\iota) + \boldsymbol{\Sigma}_\varepsilon(\iota)\boldsymbol{\Theta}_e\boldsymbol{\Sigma}_\varepsilon(\iota) \tag{6}$$

$$\boldsymbol{\Sigma}_\eta(\iota+1) = \boldsymbol{\Sigma}_\eta(\iota) + \boldsymbol{\Sigma}_\eta(\iota)\boldsymbol{\Theta}_r\boldsymbol{\Sigma}_\eta(\iota) \tag{7}$$

where $\iota = 0, 1, \ldots$. Here $\boldsymbol{\Sigma}_\varepsilon(0)$, $\boldsymbol{\Sigma}_\eta(0)$ are the starting values. In addition,

$$\boldsymbol{\Theta}_r = \frac{1}{n}\sum_{t=1}^{n}(\boldsymbol{r}_t\boldsymbol{r}_t' - \boldsymbol{N}_t) \tag{8}$$

where $\boldsymbol{r}_t$ and $\boldsymbol{N}_t$ are constructed as in (4) and (5), and

$$\boldsymbol{\Theta}_e = \frac{1}{n}\sum_{t=1}^{n}(\boldsymbol{e}_t\boldsymbol{e}_t' - \boldsymbol{D}_t) \tag{9}$$

with

$$\boldsymbol{e}_t = \boldsymbol{F}^{-1}\boldsymbol{v}_t - \boldsymbol{K}'\boldsymbol{r}_t \tag{10}$$

and

$$\boldsymbol{D}_t = \boldsymbol{F}^{-1} + \boldsymbol{K}'\boldsymbol{N}_t\boldsymbol{K}. \tag{11}$$

Using the steady-state filter and smoother, we obtain a computationally feasible procedure to approximate the maximum likelihood estimation of the multivariate local level model.

Algorithm 1 (Approximate maximum likelihood estimation) *Fix arbitrary initial covariance matrices $\boldsymbol{\Sigma}_\varepsilon$ and $\boldsymbol{\Sigma}_\eta$ and iterate as follows.*

1. *Obtain $\boldsymbol{P}$ from equation (3) and compute the steady-state Kalman filter matrices and start the Kalman filter with $\boldsymbol{a}_1 = \boldsymbol{y}_1$ (see Harvey, 1989, p. 26).*
2. *Run $\boldsymbol{a}_t = \boldsymbol{K}\boldsymbol{y}_t + (\boldsymbol{I} - \boldsymbol{K})\boldsymbol{a}_{t-1}$ and for $t = 2, \dots, n$.*
3. *Run the steady-state smoothing formulae (4), (5), (10), (11), (8) and (9).*
4. *Run the EM step updating the parameters using (6) and (7).*
5. *Go to step 1. until the likelihood increment is negligible.*

Table 1 compares the execution time and the precision of our algorithm to exact MLE on simulated vector time series of dimension up to $d = 20$. For $d = 100$ our algorithm provides estimation in few minutes, while it was not possible to get results for MLE.

Table 1 Comparisons of Kalman filter based maximum likelihood estimates versus estimates obtained by Algorithm 1 (MAD = mean absolute difference, MAE = mean absolute error).

	Execution time (sec)			MAD		MAE MLE		MAE Alg. 1	
d	Classic	Alg. 1	Ratio	$\boldsymbol{\Sigma}_\varepsilon$	$\boldsymbol{\Sigma}_\eta$	$\boldsymbol{\Sigma}_\varepsilon$	$\boldsymbol{\Sigma}_\eta$	$\boldsymbol{\Sigma}_\varepsilon$	$\boldsymbol{\Sigma}_\eta$
3	1.3	0.1	10.6	0.002	0.030	0.006	0.006	0.007	0.031
5	5.7	0.5	11.8	0.002	0.022	0.006	0.007	0.006	0.023
10	60.3	1.6	37.0	0.001	0.012	0.006	0.007	0.006	0.015
20	1490.4	10.1	147.8	0.001	0.006	0.006	0.008	0.006	0.011

3 Application to multivariate stochastic volatility models

In this section, we demonstrate how the stochastic volatility (SV) models of Harvey et al (1994) (from now on HRS) and Alizadeh et al (2002) (from now on ABD) can be successfully estimated on large portfolios using our Algorithm 1. In particular, while Harvey et al (1994) discuss how to estimate small scale multivariate stochastic volatility model using state space methods, Alizadeh et al (2002) cover only the univariate model. Therefore, we first show how to extend the model by Alizadeh et al (2002) to the multivariate case and then apply our approximate maximum likelihood estimation to a large portfolio of stocks.

Let y_{it} be the, possibly mean-adjusted, return of stock $i \in \{1, 2, \dots, d\}$ at time $t \in \{1, 2, \dots, n\}$, then the multivariate stochastic volatility model Harvey et al (1994) consider is defined by

$$\begin{aligned} y_{it} &= \exp(h_{it}/2)\zeta_{it}, \\ h_{it+1} &= h_{it} + \eta_{it}, \end{aligned} \tag{12}$$

where $\exp(h_{it})$ plays the role of the time-varying variance, whose logarithm evolves according to a random walk. The random vectors $\boldsymbol{\zeta}_t$ and $\boldsymbol{\eta}_t$ obtained by stacking the random variables $\{\zeta_{1t},\ldots,\zeta_{dt}\}$ and $\{\eta_{1t},\ldots,\eta_{dt}\}$, respectively, into vectors are assumed to be normally distributed with zero means and covariance matrices $\boldsymbol{\Sigma}_\zeta$ and $\boldsymbol{\Sigma}_\eta$. The matrix $\boldsymbol{\Sigma}_\zeta$ is constrained to be a correlation matrix. Harvey et al (1994) model the logarithm of the squared return, so that equation (12) can be rewritten as

$$\begin{aligned} \log(y_{it}^2) &= h_{it} + \log(\zeta_{it}^2), \\ h_{it+1} &= h_{it} + \eta_{it}, \end{aligned} \tag{13}$$

This set of equations can be easily cast into the linear state space form as

$$\begin{aligned} \boldsymbol{w}_t &= \boldsymbol{h}_t + \boldsymbol{\varepsilon}, \\ \boldsymbol{h}_{t+1} &= \boldsymbol{h}_t + \boldsymbol{\eta}_{it}, \end{aligned} \tag{14}$$

where the i-th element of $\boldsymbol{w}_t$ is equal to $\log(y_{it}^2)+1.27$ and $\boldsymbol{\varepsilon}_t$ is a non-Gaussian random i.i.d. sequence whose ij-th element of the covariance matrix $\boldsymbol{\Sigma}_\varepsilon$ is given by equation B.9 in Harvey et al (1994).

Harvey et al (1994) propose to estimate the unknown covariance matrices $\boldsymbol{\Sigma}_\varepsilon$ and $\boldsymbol{\Sigma}_\eta$ by Gaussian quasi maximum likelihood, approximating the distribution of $\boldsymbol{\varepsilon}_t$ with a normal with the same mean and covariance matrix. The log-variances h_{it} can be estimated using Kalman filtering and state smoothing, which in this case provide just best linear estimates.

Alizadeh et al (2002) propose a stochastic volatility model based on the logarithm of stock price ranges (daily maximum minus daily minimum) instead of log-squared returns. Let $v_{it} = \log(P_{it}^{\max} - P_{it}^{\min})$ be the sequence of daily log-ranges, with $P_{it}^{\max}$ and $P_{it}^{\min}$ representing the maximum and minimum price of stock i for the day t. If each price process is well described by a Brownian motion, than the distribution of the log-range is approximately Gaussian (cf. Alizadeh et al, 2002, Table I and Figure 1). Assuming that the approximation remains valid also in a multivariate context, then we can write a multivariate stochastic volatility model based on log-ranges exactly as in equation (14), where now the generic element of $\boldsymbol{w}_t$ is $w_{it} = v_{it} - 0.43$, and the errors ε_{it} have all the same standard deviation $\sigma_\varepsilon = 0.29$ (cf Alizadeh et al, 2002, Table I).

We estimated both SV models for a portfolio composed by 94 daily stock returns belonging to the SP100 index ranging from 2007-01-04 to 2017-04-28 ($n = 2598$). As customary in many large scale multivariate GARCH models, we split the estimation process in two steps: first, we estimated the correlation matrix of the returns, from which we obtained the covariance matrix $\boldsymbol{\Sigma}_\varepsilon$, and then we applied our Algorithm 1 by running the EM update only for the matrix $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$, as in equation (7), and using the $\boldsymbol{\Sigma}_\varepsilon$ estimated in the first step.

If we concentrate on the correlation matrix of the disturbances that drive the h_{it} processes, we notice that all correlations are positive, but those of model HRS are generally larger than those of model ABD. The same message can be derived from the cumulated eigenvalues plot in the left panel of Figure 1: for model HRS the first three principal components cover more than 90% of the variance, while the same share of total variance is reached for model ABD with 20 components. If we use the scores of the first principal component of the estimated $\boldsymbol{h}_t$ in the two models and take the suitable transforms to derive volatility indicators, we get the volatility profiles depicted in the right panel of Figure 1. The profiles are similar, but they differ in some extreme event and in the level, which is higher for the HRS model.

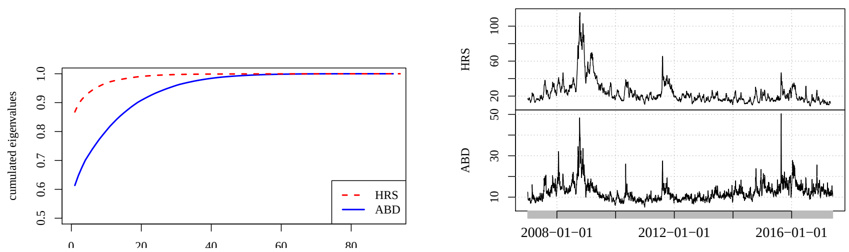

Fig. 1 Left) cumulated eigenvalues of the correlation matrix derived from $\boldsymbol{\Sigma}_\eta$ estimated in the two models. Right) Ensemble volatility indicators derived from the first principal component scores

References

Alizadeh S, Brandt MW, Diebold FX (2002) Range-based estimation of stochastic volatility models. The Journal of Finance 57(3):1047–1091

Ansley C, Kohn R (1985) Estimation, filtering and smoothing in state space models with incomplitely specified initial conditions. The Annals of Statistics 13:1286–1316

Durbin J, Koopman S (2001) Time Series Analysis by State Space Methods. Oxford University Press

Harvey A (1989) Forecasting Structural Time Series and the Kalman Filter. Cambridge University Press

Harvey A, Ruiz E, Shephard N (1994) Multivariate stochastic variance models. Review of Economic Studies 61(2):247–264

de Jong P (1988) A cross-validation filter for time series models. Biometrika 76:594–600

de Jong P (1989) Smoothing and interpolation with state-space models. Journal of the American Statistical Association 75:594–600

Koopman S (1993) Disturbance smoother for state space model. Biometrika 80(1):117–126

Koopman S (1997) Exact initial kalman filtering and smoothing for nonstationary time series models. Journal of the American Statistical Association 92(400):1630–1638

Shumway RH, Stoffer DS (2017) Time Series Analysis and Its Applications. With R Examples. Springer