The Wishart Autoregressive Process of Multivariate Stochastic Volatility

Gourieroux, C. [1], Jasiak, J. [2], and R., Sufana[3]

January 2004, revised September 2005

[1]CREST, CEPREMAP and University of Toronto.
[2]York University, Toronto.
[3]University of Toronto.
The authors thank H. Massam, E. Renault, C. Robert and N. Shephard for helpful comments.
The authors gratefully acknowledge financial support of NSERC Canada.

The Wishart Autoregressive Process of Multivariate Stochastic Volatility

Abstract

The Wishart Autoregressive (WAR) process is a multivariate process of stochastic positive definite matrices. The WAR is proposed in this paper as a dynamic model for stochastic volatility matrices. It yields simple nonlinear forecasts at any horizon and has factor representation, which separates white noise directions from those that contain all information about the past. For illustration, the WAR is applied to a sequence of intraday realized volatility-covolatility matrices.

Keywords: Stochastic Volatility, Car Process, Factor Analysis, Reduced Rank, Realized Volatility.

JEL number: G13, C51.

1 Introduction

Portfolio management of multiple risky assets demands tractable multivariate models of expected returns, volatilities, and covolatilities. While there exists a variety of stochastic volatility models for one risky asset, relatively few papers propose specifications of stochastic volatility in a multiasset framework. Among these, a majority is interested in a limited number of assets, such as 2, 3 or 4, except for the recent papers on conditional correlation GARCH model by Engle and Sheppard (2001), Fiorentini, Sentana and Shephard (2003), and the bayesian model by Chip, Nardari and Shephard (2002). Concerning empirical research, the existing multivariate models are typically fitted to series of exchange rates [4], interest rates[5], stock prices,[6] and volatility dependence between stock markets[7]. The reason that theoretical contributions in multivariate volatility analysis are so scarce, is the difficulty in finding a dynamic specification of a stochastic volatility matrix, which would satisfy all the following requirements:

i) the symmetry and positivity properties of each variance-covariance matrix in the process satisfied at each point in time.

ii)a reasonably low number of parameters, without compromising the flexibility of the model (to alleviate the curse of dimensionality).

iii) the availability of forecasts at any horizon in a closed form.

iv) the possibility of varifying in a straightforward manner the time series properties of the volatility process, such as stationarity and Markov property.

v) the invariance of the model with respect to time aggregation and portfolio allocation.

vi) the existence of a direct analogue in continuous time.

vii) the compatibility with the theoretical models of the term structure of interest rates and derivative pricing.

In the literature, we distinguish two types of multivariate models for the dynamic volatility-covolatility matrix:

$$Y_t = V_t\left(r_{t+1}\right),$$

where r_{t+1} is a n-dimensional vector of returns, (Y_t) is a (n,n) symmetric positive definite matrix, and V_t denotes the variance-covariance matrix conditional

[4]Bollerslev (1987), Diebold and Nerlove (1989), Bollerslev (1990), Baillie, Bollerslev (1990), Pelletier (2003).

[5]Engle, Ng and Rothschild (1990).

[6]Schwert, Seguin (1990).

[7]King and Whadwani (1990), King, Sentana, and Wadhwani (1994), Lin, Engle, and Ito (1994), Ledoit, Santa-Clara, Wolf (2001).

on the information available at date t. These are the multivariate ARCH models and the stochastic volatility models described below.

i) The multivariate ARCH model relies on an autoregressive specification of the volatility matrix, written as a linear combination of lagged volatilities and lagged squared returns. The basic multivariate ARCH model is the multivariate ARCH(1) model, in which the elements of the volatility matrix Y_t are linear affine functions of the elements of the matrix of squared returns : $vech\,(Y_t) = A\,vech\left(r_{t-1}r'_{t-1}\right) + b$, where $vech\,(Y)$ denotes the vector obtained by stacking the $\frac{n(n+1)}{2}$ different elements of Y. Regarding the requirements listed above, the multivariate ARCH(1) suffers from the curse of dimensionality [see Bollerslev, Engle, and Wooldridge (1988)], as the full unrestricted model involves $\left[\frac{n(n+1)}{2}\right]^2 + \frac{n(n+1)}{2}$ parameters. To resolve this difficulty, the multivariate ARCH literature came up with the following extensions: The diagonal-vech specification is based on the assumption that matrix A is diagonal and each series in the multivariate vector has a GARCH-like specification[8] [Bollerslev, Engle, and Wooldridge (1988) and e.g. Brandt and Diebold (2002) for an application]; the constant conditional correlation restriction was imposed in Bollerslev (1987) to make the estimation of a large model feasible and to ensure positive definiteness of the covariance matrix; this approach has been extended by Pelletier (2003), who considered a regime switching model with constant correlation in each regime. Recently, Tse, Tsui (2002), Engle (2002) introduced models with time varying correlations. They proposed a nonlinear GARCH type representation, which guarantees that correlations vary between -1 and 1.

An alternative stream of research focused on the spectral decomposition of the volatility matrix, assumed to be of some specific form [Baba, Engle, Kraft, and Kroner (1987)]. Recently, Alexander (2000) has advocated the use of factor ARCH models, initially proposed by Engle, Ng, and Rothschild (1990), which, in turn, were criticized by Engle and Sheppard (2001) for poor fit in empirical research.

The existing literature hasn't been fully successful in eliminating the multiple drawbacks of multivariate ARCH specifications. The symmetry and positivity constraints can only be satisfied under a set of complicated parameter restrictions, which are hard to interpret. Also, the models are not invariant with respect to a change of time unit[9], or with respect to change in portfolio allocation [see for example the Dynamic Conditional Correlation model by Engle, Sheppard (2001)].

ii) Stochastic volatility models in discrete time have been initially introduced by Taylor (1986), and extended to multivariate processes by Harvey, Ruiz, and Shephard (1994) [see, Chib, Nardari, Shephard (2002), Fiorentini, Sentana, Shephard (2002), also Ghysels, Harvey, and Renault (1996) for a survey on the so-called stochastic variance models]. Typically, in this literature the

[8] This approach has been recently extended by Engle and Sheppard (2001) to a model with time-varying correlation compatible with univariate GARCH.

[9] See Drost and Nijman (1993), Drost and Werker (1996), Meddahi and Renault (2004) for a discussion of time aggregation of ARCH and volatility models.

volatility matrix is written as:

$$Y_t = A \begin{pmatrix} \exp h_{1t} & & 0 \\ & \ddots & \\ 0 & & \exp h_{nt} \end{pmatrix} A',$$

where A is a (n,n) matrix and h_{it}, $i = 1, \ldots, n$, are independent volatility factor processes. The factor processes can be chosen so that $(h_{1t}, \ldots, h_{nt})$ is a Gaussian VAR process [see Harvey, Ruiz, and Shephard (1994)]. This specification ensures that stochastic matrices (Y_t) are symmetric positive definite and follow a Markov process. The stochastic variance model is easy to estimate from return data with zero expected value by using the Kalman filter, but quite hard to implement if (non-zero) volatility-in-mean is considered [10]. Moreover, other important drawbacks concern 1) the number n of latent factors, which is strictly less than the number of distinct elements of Y_t, and 2) the diagonal form of volatility matrix, which assumes stochastic weights, but constant factor loadings (given in the columns of A)[11] [12].

The existing multivariate models seem too restrictive to accommodate the complexity of data. Therefore, some researchers agree that new solutions need to be found [see Engle (2002a)]. The aim of the present paper is to introduce a multivariate dynamic specification, which is compatible with financial theory, satisfies the constraints on volatility matrices, has a flexible form, is easy to predict, invariant with respect to temporal aggregation and portfolio allocation, and easy to implement. Our approach is based on a dynamic extension of the Wishart distribution. It is known that a sample variance-covariance matrix computed from i.i.d. multivariate Gaussian observations [see Wishart (1928a,b) for the initial papers, and Anderson (1984), Muirhead (1978, 1982), Stuart and Ord (1994), Bilodeau and Brenner (1999) for surveys] follows the Wishart distribution. The extension consists in introducing serial dependence by considering multivariate serially correlated Gaussian processes, which are contemporaneously independent of one another.

The Wishart Autoregressive (WAR) process is defined in Section 2 from the specification of the conditional Laplace transform. The section explains how the WAR process of order 1 is constructed from underlying Gaussian VAR processes, when the degree of freedom parameter is an integer, and presents the first and second order conditional moments. The definition is next extended to WAR processes of autoregressive order higher than 1. Examples of WAR processes are discussed in Section 3, and some continuous time analogues are presented in Section 5. The WAR processes arise as special cases of compound autoregressive (Car) processes introduced in Darolles, Gourieroux, and Jasiak (2005). For this reason, nonlinear predictions at any horizon are easy to perform. The predictive

[10] See Kim, Shephard, Chib (1998) for an application to exchange rates.

[11] The specification resembles Bollerslev's constant correlation GARCH process, since the correlation is zero after a transformation A^{-1} of basic assets.

[12] Constant factor loadings are also assumed in the standard factor ARCH model [Diebold, Nerlove(1989), Engle, Ng, Rotschild (1990), Alexander (2000)].

distribution at horizon h and temporal aggregation are discussed in Section 4. Section 6 is examines models with reduced rank and their factor interpretations. This factor representation separates white noise directions from directions that contain all information about the past. The WAR-in-mean models are presented in Section 7. Next, the WAR-in-mean model is used as a representation for the dynamics of an efficient portfolio in the mean-variance framework, and structural interpretations are discussed. Statistical inference is covered in Section 8. First, we consider observable volatility matrices and discuss the identification. Next, we explain how to derive simple, consistent nonlinear least squares estimators. The nonlinear least squares estimates can be used as initial values in likelihood maximization, which is given as an alternative estimation method. In Section 9, the Wishart process is fitted to a series of intraday realized volatility matrices. In this application, the number and types of latent factors are examined. Section 10 concludes the paper. The proofs are gathered in Appendices.

2 The Wishart Autoregressive Process

This section defines the Wishart Autoregressive process and describes its dynamic properties. The Wishart process is a process (Y_t) formed by stochastic symmetric positive definite matrices of dimension $(n \times n)$. The dynamic of the Wishart process is specified by its conditional Laplace transform, which defines the conditional expectations of any exponential transform of elements of matrix Y_{t+1}. It is defined as follows:

$$\Psi(\Gamma) = E_t[\exp Tr(\Gamma Y_{t+1})],$$

where E_t denotes the expectation conditional on current and lagged elements of Y, Γ is a deterministic and symmetric matrix of real numbers, and Tr denotes the trace operator. In particular, for two symmetric matrices Γ and Y, we have:

$$Tr\,(\Gamma Y) = \sum_{i=1}^{n} (\Gamma Y)_{ii} = \sum_{i=1}^{n}\sum_{l=1}^{n} \gamma_{il} Y_{li} = \sum_{i=1}^{n}\sum_{l=1}^{n} \gamma_{il} Y_{il}.$$

Section 2.1 defines the Wishart Autoregressive process of order 1. It is a matrix process with Wishart conditional distribution [Wishart (1928) a,b], and a noncentrality parameter, which is an affine function of lagged values of Y_t. Section 2.2 provides an interpretation of the Wishart process of order 1 as the outer product of Gaussian VAR(1) processes for integer valued degree of freedom parameter. Using this approach, we interpret the structural parameters of the model, and give insights on the expressions of conditional moments and forecasts. Next, we derive the conditional first and second order moments of WAR(1), and show that this model is invariant with respect to portfolio allocation (Section 2.3). Finally, we discuss the extension of the WAR process of order one to a WAR process of any finite order p.

2.1 The Wishart Autoregressive process of order 1

Definition 1: The Wishart Autoregressive process of order 1, denoted WAR(1) is a matrix Markov process (Y_t) with the following conditional Laplace transform:

$$\begin{aligned}\Psi_t(\Gamma) &= E\left[\exp Tr\left(\Gamma Y_{t+1}\right)\right] \\ &= \frac{\exp Tr\left[M'\Gamma\left(Id-2\Sigma\Gamma\right)^{-1}MY_t\right]}{\left[\det\left(Id-2\Sigma\Gamma\right)\right]^{K/2}},\end{aligned}$$

The transition density of WAR(1) depends on the following parameters: K is the scalar degree of freedom, strictly larger than $n-1$, M is the $n \times n$ matrix of autoregressive parameters, and Σ is a $n \times n$ symmetric, positive definite matrix. The Laplace transform is defined for a matrix Γ such that [13] $\|2\Sigma\Gamma\| < 1$.

The transition density of this process is noncentered Wishart [see Muirhead (1982) p.442]:

$$\begin{aligned}f(Y_{t+1}|Y_t) &= \frac{1}{2^{Kn/2}}\frac{1}{\Gamma_n(K/2)}(det\Sigma)^{-K/2}(detY_{t+1})^{(K-n-1)/2} \\ &\quad \exp\{(-\frac{1}{2}Tr[\Sigma^{-1}(Y_{t+1}+MY_tM')]\}{}_0F_1(K/2;(1/4)MY_tM'Y_{t+1})\end{aligned}$$

where $\Gamma_n(K/2) = \int_{A>>0}\exp\{Tr(-A)\}(detA)^{(K-n-1)/2}dA$ is the multidimensional gamma function, ${}_0F_1$ is the hypergeometric function of matrix argument, and the density is defined on positive definite matrices. The hypergeometric function has a series expansion:

$${}_0F_1(K/2;(1/4)MY_tM'Y_{t+1}) = \sum_{p=0}^{\infty}\sum_l \frac{C_l((1/4)MY_tM'Y_{t+1})}{(K/2)_l p!},$$

where $\sum_l$ denotes summation over all partitions $l = (p_1,...,p_m), p_1 \geq ... \geq p_m \geq 0$ of p into integers, $(K/2)_l$ is the generalized hypergeometric coefficient $(K/2)_l = \prod_{i=1}^m (K/2-(i-1)/2)_{p_i}$ with $(a)_{p_i} = a(a+1)...(a+p_i-1)$, and $C_l((1/4)MY_tM'Y_{t+1})$ is the zonal polynomial associated with partition l. The zonal polynomials have no closed form expressions, but can be easily computed recursively (see Muirhead (1982), Chapter 7.2, and James (1968)).

The WAR(1) model alleviates the curse of dimensionality encountered in multivariate volatility models, where the number of reduced form parameters is of order $\left[\frac{n(n+1)}{2}\right]^2$. The WAR(1) process involves a much smaller number of parameters equal to $1+\frac{n(n+1)}{2}+n^2$, which corresponds to the order for the reduced-form parameters of n-dimensional VAR(1) process. The number of parameters can be reduced further by imposing some restrictions on matrices M, or Σ (see Section 6).

[13]The norm of a symmetric matrix is equal to its maximal eigenvalue.

2.2 WAR(1) with Integer Degree of Freedom K.

Let us consider the process Y_t defined by

$$Y_t = \sum_{k=1}^{K} x_{kt} x'_{kt}, \tag{1}$$

where the processes $x_{kt}, k = 1, ..., K$ are independent Gaussian VAR(1) processes of dimension n with the same autoregressive parameter matrix M and innovation variance Σ:

$$x_{kt} = M x_{k,t-1} + \epsilon_{k,t}, \quad \epsilon_{k,t} \sim N(0, \Sigma). \tag{2}$$

The Proposition below is proved in Appendix 1.

Proposition 1: When the processes (x_{kt}) , $k = 1, \ldots, K$, are independent with the same autoregressive parameter M and innovation variance Σ:
i) The process $Y_t = \sum_{k=1}^{K} x_{kt} x'_{kt}$ is a Markov process.
ii) Its (conditional) Laplace transform is given by:

$$\begin{aligned}
\Psi_t(\Gamma) &= E\left[\exp Tr\left(\Gamma Y_{t+1}\right)|x_t\right] \\
&= E\left[\exp\left(\sum_{k=1}^{K} x'_{k,t+1}\Gamma x_{k,t+1}\right)|x_t\right] \\
&= E\left[\exp\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{ij}Y_{ij,t+1}\right)|Y_t\right] \\
&= \frac{\exp Tr\left[\left(M'\Gamma\left(Id-2\Sigma\Gamma\right)^{-1}M\right)Y_t\right]}{\left[\det\left(Id-2\Sigma\Gamma\right)\right]^{K/2}}.
\end{aligned}$$

The conditional Laplace transform depends on $x_{k,t}, k = 1, ..., K$ by Y_t only, which is the Markov property of matrix process (Y_t).

Proposition 1 justifies the interpretation of parameters M and Σ as a latent autoregressive parameter, and a latent innovation variance, respectively. However, the interpretation of (Y_t) from latent Gaussian VAR(1) processes is valid for integer valued K only. In general, any economic or financial interpretation of the latent processes (x_{kt}) is not necessary, except for applications such as the quadratic term structure of interest rates[14]. In this paper, the latent processes are introduced mainly to provide an intuitive understanding of parameters and results.

It is interesting to note that, when $n = 1$, the Wishart distribution becomes a chi-square distribution, which explains the interpretation of Y_t as a sum of squared Gaussian variables. In particular, extending that Wishart distribution

[14] See Ahn, Dittmar and Gallant (2002) for the estimation of a basic quadratic term structure model, and Cheng and Scaillet (2002), Gourieroux and Sufana (2003) for the discussion and extension of such models.

to noninteger degrees of freedom, is analogous to extending the corresponding chi-square to a gamma distribution.

Let us now comment on the condition $K > n-1$. We shall see that it ensures the almost sure invertibility of Y_t. When $K=1$, the matrix $Y_t = x_{1t}x'_{1t}$ has rank equal to 1. In this case Y_t, it is not invertible, and does not have any continuous distribution on the set of symmetric positive definite matrices. When K is an integer greater or equal to n, Y_t is a sum of a sufficient number of independent matrices of rank 1 and Y_t is invertible.

2.3 Conditional moments

The conditional Laplace transform contains all information on the conditional distribution [15]. However, other summary statistics, such as the first and second order conditional moments, can also be considered, even though these are less informative. While the expression of the conditional expectation of a stochastic matrix is easy to define, its conditional variance-covariance matrix is cumbersome. Remember that the volatility matrix of a stochastic volatility matrix[16] is of dimension $\frac{n(n+1)}{4}\left[\frac{n(n+1)}{2}+1\right]$, which is very large. In order to provide some insights on the structure of that matrix, without complicated matrix notation, we calculate the conditional variance between two inner products $\gamma' Y_{t+1}\alpha$, $\delta' Y_{t+1}\beta$ based on Y_{t+1}. Given the formulas established for any real vectors α, β, γ, δ, we can compute all covariances of interest. For instance, the conditional covariance $cov_t\left(Y_{ij,t+1}, Y_{kl,t+1}\right)$ corresponds to $\alpha = e_j$, $\gamma = e_i$, $\beta = e_l$, $\delta = e_k$, where e_i is the i^{th} canonical vector with zero components except the i^{th} component, which is equal to 1.

The first and second order conditional moments of the WAR(1) process are derived in Appendix 2.

Proposition 2: We have:
i) $E_t\left(Y_{t+1}\right) = MY_tM' + K\Sigma$.
ii) For any set of four n-dimensional vectors α, β, γ, δ we get:

$$\begin{aligned} & cov_t\left(\gamma' Y_{t+1}\alpha, \delta' Y_{t+1}\beta\right) \\ = \; & \gamma' MY_tM'\delta\alpha'\Sigma\beta + \gamma' MY_tM'\beta\alpha'\Sigma\delta + \alpha' MY_tM'\delta\gamma'\Sigma\beta \\ & +\alpha' MY_tM'\beta\gamma'\Sigma\delta + K\left[\gamma'\Sigma\beta\alpha'\Sigma\delta + \alpha'\Sigma\beta\gamma'\Sigma\delta\right]. \end{aligned}$$

The first and second order conditional moments are affine functions of the lagged values of the volatility process, which is a direct consequence of the exponential affine expression of the conditional Laplace transform [see Darolles, Gourieroux, and Jasiak (2005)]. In particular, the WAR(1) process is a weak linear AR(1) process [see e.g. Grunwald, Hyndman, Tedesco, and Tweedie (2000) for a survey of linear AR(1) processes]. More precisely, we get :

[15] This is due to the positivity of process (Y_t) [see Feller (1971)].

[16] The volatility of the volatility is important for financial applications. Indeed, it is related to the volatility of derivatives written on underlying returns. For this reason, a market for derivatives on the market index volatility has opened in Chicago.

$$Y_{t+1} = MY_tM' + K\Sigma + \eta_{t+1}, \tag{3}$$

where η_{t+1} is a matrix of stochastic errors with conditional mean zero. Equivalently, we get:

$$vech(Y_{t+1}) = A(M)vech(Y_t) + vech(K\Sigma) + vech(\eta_{t+1}), \tag{4}$$

where $vech(Y)$ denotes the vector obtained by stacking the lower triangular elements of Y, and $A(M)$ is a matrix function of M. The error term η is a weak white noise, since it features conditional heteroscedasticity and, even after conditional standardization, is not identically distributed.

2.4 Invariance to linear invertible transformation

Let us consider a WAR(1) process Y_t of dimension n with parameters K, M, Σ, and a (n, n) invertible matrix A; the process: $Y_t\,(A) = A'Y_tA$ is another process of stochastic symmetric positive definite matrices. Moreover, for integer K, we have:

$$Y_t\,(A) = A' \sum_{k=1}^{K} x_{kt}x'_{kt}A = \sum_{k=1}^{K} A'x_{kt}x'_{kt}A = \sum_{k=1}^{K} z_{kt}z'_{kt},$$

where $z_{kt} = A'x_{kt}$ are also Gaussian autoregressive processes such that: $z_{k,t+1} = A'M\,(A')^{-1}\,z_{k,t} + A'\varepsilon_{k,t+1}$. This explains the property below for which the proof for noninteger K follows directly from the conditional Laplace transform.

Proposition 3: If (Y_t) is a WAR(1) process $W_n\,(K, M, \Sigma)$ and A is a (n, n) invertible matrix, then $Y_t\,(A) = A'Y_tA$ is also a WAR(1) process $W_n(K,\, A'M\,(A')^{-1},\, A'\Sigma A)$.

From a financial point of view,, Proposition 3 establishes the invariance of the family of Wishart processes with respect to portfolio allocation. Indeed, let us consider n basic assets with returns r_t and volatility Y_t, and n portfolios of various quantities of those assets. The quantities of each asset (positive or negative) in a given portfolio allocation form a column of matrix A. The returns on the portfolios are:

$$r_{t+1}\,(A) = A'r_{t+1},$$

whereas the portfolios' volatilities are $V_tr_{t+1}\,(A) = A'Y_tA$. Thus, if asset return volatility follows a Wishart process, the portfolios' volatility follows a Wishart process as well [17]. This invariance property is not satisfied by some constrained multivariate ARCH models such as the so-called diagonal model, the model with constant correlation and the Dynamic Conditional Correlation model.

Proposition 3 implies that any Wishart autoregressive process can be rewritten as a "standardized" WAR, with latent innovation variance equal to an identity matrix of dimension n.

[17] Similarly, the Wishart specification for a volatility matrix of log-exchange rates is invariant with respect to the currency unit.

Corollary 1: Any WAR(1) process $W_n(K, M, \Sigma)$ can be written as: $Y_t = \Sigma^{1/2} Y_t^* \Sigma^{1/2}$, where Y_t^* is a "standardized" WAR(1) process $W_n(K,\, \Sigma^{-1/2} M \Sigma^{1/2},\, Id)$.

Other linear invertible transformations can also be considered. For instance, let us assume that the autoregressive matrix M is diagonalizable[18]. M can be written as: $M = Q \Lambda Q^{-1}$, where Q is the matrix of eigenvectors and Λ the diagonal matrix of eigenvalues of M. The transformed process $Y_t^* = Q^{-1} Y_t \left(Q^{-1}\right)'$ is a WAR(1) process $W_n(K,\, \Lambda,\, Q^{-1}\Sigma\left(Q^{-1}\right)')$, with a diagonal autoregressive matrix. Thus, all interactions between latent variables are captured by the innovation variance.

For studies concerning porfolio allocations, we define the portfolio volatilities $\alpha' Y_t \alpha$, where α is a given vector of portfolio allocations. The second order dynamic properties of such portfolio volatilities follow from Proposition 2 (see Appendix 3).

Corollary 2: Let α, β, γ, δ be n-dimensional vectors. We obtain:

i) $V_t(\gamma' Y_{t+1} \alpha) = \gamma' M Y_t M' \gamma\ \alpha' \Sigma \alpha + 2\gamma' M Y_t M' \alpha\ \alpha' \Sigma \gamma + \alpha' M Y_t M' \alpha\ \gamma' \Sigma \gamma$
$+K\left[(\gamma'\Sigma\alpha)^2 + \alpha'\Sigma\alpha\gamma'\Sigma\gamma\right]$;

ii) $V_t(\alpha' Y_{t+1} \alpha) = 4\alpha' M Y_t M' \alpha\ \alpha' \Sigma \alpha + 2K(\alpha'\Sigma\alpha)^2$;

iii) $cov_t(\alpha' Y_{t+1} \alpha, \beta' Y_{t+1} \beta) = 4\alpha' M Y_t M' \beta\ \alpha' \Sigma \beta + 2K(\alpha'\Sigma\beta)^2$;

iv) $cov_t(\alpha' Y_{t+1} \alpha, \alpha' Y_{t+1} \beta) = 2\alpha' M Y_t M' \alpha\ \alpha' \Sigma \beta + 2\alpha' M Y_t M' \beta\ \alpha' \Sigma \alpha$
$+2K\alpha'\Sigma\beta\ \alpha'\Sigma\alpha$.

We see that :

i) the degree of freedom parameter determines the magnitude of overdispersion;

ii) the correlations between portfolio volatilities can be of any sign due to the first term in iii). Thus, it is easy to accommodate asymmetric reactions of volatilities and covolatilities [Ang, Chen (2002)].

2.5 WAR(p) processes

Due to nonlinear dynamics and the number n of components in Y_t, a WAR(1) process can accommodate a large spectrum of patterns of persistence in volatilities and covolatilities, including possibly long memory effects. Nevertheless, there may be cases when WAR processes with higher autoregressive order p (called WAR(p)) need to be considered. The Wishart processes are easily extended to include more autoregressive lags. Since the formula of the conditional Laplace transform in Definition 1 is valid for any conditioning matrix $M Y_t M'$, this matrix can be replaced by any symmetric positive semi-definite function of Y_t, $Y_{t-1}, \ldots, Y_{t-p+1}$.

[18]This assumption has been made for instance by Ahn, Dittmar and Gallant (2002) in the context of quadratic term structure models.

Definition 2: A Wishart autoregressive process of order p, denoted WAR(p), is a matrix process with conditional Laplace transform:

$$\begin{aligned}\Psi_t(\Gamma) &= E_t\left[\exp Tr\left(\Gamma Y_{t+1}\right)\right] \\ &= \frac{\exp Tr\left[\Gamma\left(Id-2\Sigma\Gamma\right)^{-1}\sum_{j=1}^{p} M_j Y_{t-j+1} M_j'\right]}{\left[\det\left(Id-2\Sigma\Gamma\right)\right]^{K/2}},\end{aligned}$$

where the matrices M_j of dimension (n,n) represent the sequence of latent "matrix autoregressive coefficients".The process is denoted $W_n\left(K; M_1,\ldots,M_p,\Sigma\right)$.

When the autoregressive order is larger than 1, the interpretation of the Wishart process as the sum of squares of autoregressive Gaussian processes is no longer valid, even for integer K. For instance, let us consider a Gaussian VAR(2) process: $x_{t+1} = M_1 x_t + M_2 x_{t-1} + \varepsilon_{t+1}$, $\varepsilon_{t+1} \sim IIN\left(0,\Sigma\right)$. The conditional Laplace transform of $Y_{t+1} = \left(x_{t+1}x_{t+1}'\right)$ given $\underline{x_t} = \left(x_t, x_{t-1},\ldots\right)$ becomes:

$$\begin{aligned}&\Psi_t(\Gamma)\\ =\ &\frac{\exp\left[\left(M_1x_t+M_2x_{t-1}\right)'\Gamma\left(Id-2\Sigma\Gamma\right)^{-1}\left(M_1x_t+M_2x_{t-1}\right)\right]}{\left[\det\left(Id-2\Sigma\Gamma\right)\right]^{1/2}}\\ =\ &\frac{\exp Tr\left[\Gamma\left(Id-2\Sigma\Gamma\right)^{-1}\left(M_1x_t+M_2x_{t-1}\right)\left(M_1x_t+M_2x_{t-1}\right)'\right]}{\left[\det\left(Id-2\Sigma\Gamma\right)\right]^{1/2}}\\ =\ &\frac{\exp Tr\left[\Gamma\left(Id-2\Sigma\Gamma\right)^{-1}\left(M_1Y_tM_1'+M_2Y_{t-1}M_2'+M_1x_tx_{t-1}'M_2'+M_2x_{t-1}x_t'M_1'\right)\right]}{\left[\det\left(Id-2\Sigma\Gamma\right)\right]^{1/2}}.\end{aligned}$$

We see that this is not the conditional Laplace transform of a Wishart process because of the presence of cross products $x_t x_{t-1}'$ [19].

The expressions of first and second order conditional moments of a WAR(p) process are similar to the expressions given in Proposition 2 and Corollary 2. We get, for instance:

$$\begin{aligned}E_t(Y_{t+1}) &= \sum_{j=1}^{p} M_j Y_{t+1-j} M_j' + K\Sigma,\\ V_t(\alpha' Y_{t+1}\alpha) &= 4\alpha'\left(\sum_{j=1}^{p} M_j Y_{t+1-j} M_j'\right)\alpha\alpha'\Sigma\alpha + 2K(\alpha'\Sigma\alpha)^2.\end{aligned}$$

In particular, a WAR(p) process admits a weak linear autoregressive representation of order p:

[19] Section 4.2 shows how such cross terms can be handled in a Wishart framework.

$$vech(Y_{t+1}) = \sum_{j=1}^{p} A_j(M_1, ..., M_p) vech(Y_{t+1-j}) + vech(K\Sigma) + vech(\eta_{t+1}), \quad \text{say,} \tag{5}$$

where $A_j(M_1, ..., M_p)$ is a matrix function of $M_1, ..., M_p$.

3 Examples

In this section, we give various examples of Wishart processes and describe special cases, which are known in the literature, such as the Wishart White Noise, the one-dimensional Wishart process, known as the Autoregressive Gamma (ARG) Process, and the Wishart unit root process.

3.1 The Wishart White Noise

When $M = 0$, the series (Y_t) is a sequence of independent matrices with identical centered Wishart distributions with parameters K and Σ. The first and second order moments are given by: $E(Y_t) = K\Sigma$, $cov(\gamma' Y_t \alpha, \delta' Y_t \beta) = K[\gamma'\Sigma\beta\alpha'\Sigma\delta + \alpha'\Sigma\beta\gamma'\Sigma\delta]$. In particular, $cov(\alpha' Y_t \alpha, \beta' Y_t \beta) = 2K(\alpha'\Sigma\beta)^2$. The two stochastic quadratic forms $\alpha' Y_t \alpha$ and $\beta' Y_t \beta$ are uncorrelated, if and only if, vectors α and β are orthogonal for the inner product associated with Σ. Such results are useful in the analysis of Wishart processes, since, as shown in the next section, the marginal distribution of a stationary Wishart process is a centered Wishart.

3.2 The limiting deterministic case

Let us consider the WAR(1) process with parameters K, $\Sigma_K = K^{-1}\Sigma_1$, $M_K = M_1$, where Σ_1, M_1 are constant matrices, and the limit of the WAR(1) process when the degree of freedom K tends to infinity. By definition, we have (for integer K):

$$Y_t = \sum_{k=1}^{K} x_{kt} x_{kt}',$$

where $x_{k,t} = M_K x_{k,t-1} + \varepsilon_{k,t}$, $\varepsilon_{k,t} \sim N(0, \Sigma_K)$. Equivalently, we can write:

$$Y_t = \frac{1}{K} \sum_{k=1}^{K} \widetilde{x}_{kt} \widetilde{x}_{kt}',$$

where $\widetilde{x}_{k,t} = \sqrt{K} x_{kt} = M_1 \widetilde{x}_{k,t-1} + \widetilde{\varepsilon}_{k,t}$, $\widetilde{\varepsilon}_{k,t} \sim N(0, \Sigma_1)$. Since the variables $\widetilde{x}_{k,t}$, $k = 1, \ldots, K$, are independent identically distributed[20], it follows that, for

[20] if the processes are stationary. Otherwise, the result is still valid, if we assume identical initial values for the different (x_{kt}) processes.

large K:

$$Y_t \sim E\left(\widetilde{x}_{kt}\widetilde{x}_{kt}'\right),$$

by the law of large numbers.

For instance, if the autoregressive coefficient M_1 admits eigenvalues with a modulus strictly less than 1, if $x_{k,o} = 0, \forall k$, Y_t tends to $\Sigma(\infty)$, where $\Sigma(\infty)$ is the unconditional variance of $\widetilde{x}_{kt}$. Thus, the WAR(1) process includes as a limiting case the constant process, formed by a sequence of constant matrices.

3.3 The univariate WAR process

In the univariate framework ($n = 1$), the conditional Laplace transform becomes:

$$\begin{aligned}\Psi_t(\gamma) &= E\left[\exp\left(\gamma Y_{t+1}\right)|Y_t\right] \\ &= \left(1-2\gamma\sigma^2\right)^{-K/2}\exp\left(\frac{\gamma m^2}{1-2\gamma\sigma^2}Y_t\right).\end{aligned} \tag{6}$$

This is the conditional Laplace transform of an autoregressive gamma process [see e.g. Gourieroux and Jasiak (2005), Darolles, Gourieroux, and Jasiak (2005)], up to a scale factor. The transition distribution is a path dependent noncentered gamma distribution up to a change of scale.

3.4 Unit root

A WAR(1) process for $M = Id, \Sigma = Id$ exists already in the literature [Bru (1989), Bru (1991), O'Connel (2003)]. If K is an integer, the latent processes $(x_{kt}), k = 1, \dots, K$, are independent Gaussian random walks, and the $W_n(K, Id, Id)$ process arises as time discretized counterpart of the continuous time process defined by:

$$dY_t = KId_n dt + Y_t^{1/2} d\tilde{W}_t' Y_t^{1/2}, \tag{7}$$

where $Y_t^{1/2}$ is the symmetric positive root of Y_t and $\tilde{W}_t$ is a (n, n) matrix, whose components are independent Brownian motions. This matrix process is a multivariate extension of the Bessel process used in Finance for time deformation [Geman, Yor (1999)], and therefore shares the properties of the Bessel process[21]. Several theoretical results have been derived in this special case [Bru (1991), Donati-Martin et alii (2003)] including the closed form expression of the transition density of the process and the joint distribution of the process of eigenvalues of matrix Y_t. For $n = 1$, this process is equivalent to an autoregressive gamma process with a unit root. This process is known to feature long memory (see Gourieroux, Jasiak (2005)).

[21] see Karlin, Taylor (1981) p175-176 for the definition of the Bessel process, and Revuz, Yor (1998), Chapter XI for its properties.

3.5 The bivariate WAR process

The bivariate WAR(1) process has three components and depends on eight parameters. Therefore, it accommodates a variety of dynamic patterns. In this section, we show the simulated paths of

i) Y_{11t}, Y_{22t}, interpreted as volatilities,

ii) correlation $Y_{12t}/\left(Y_{11t}Y_{22t}\right)^{1/2}$, and

iii) eigenvalues $\lambda_{1t} > \lambda_{2t}$ of the stochastic volatility matrix.

The spectral decomposition of the volatility matrix is important for financial applications. The largest eigenvalue λ_{1t} is equal to the maximum of portfolio volatilities $\alpha' Y_t \alpha$, computed for the portfolio allocations standardized by $\alpha'\alpha = 1$. It provides a measure of the highest risk, while the associated eigenvector defines the most risky portfolio allocation. Similarly, the smallest eigenvalue λ_{2t} is equal to the minimum of portfolio volatilities computed for portfolio allocations standardized as before. When the smallest eigenvalue is close to zero, the associated eigenvector is a basis for arbitragist strategies.

For illustration, let us consider three experiments involving a bivariate WAR(1) process with $T = 100$ observations, $K = 2$ latent processes and latent innovation variance $\Sigma = Id$. The autoregressive coefficients are as follows:

$$
\begin{aligned}
M &= \begin{pmatrix} 0.9 & 0 \\ 1 & 0 \end{pmatrix} \text{for experiment 1, } M = \begin{pmatrix} 0.3 & -0.3 \\ -0.3 & 0.3 \end{pmatrix} \text{for experiment 2,} \\
M &= \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix} \text{ for experiment 3.}
\end{aligned}
$$

The first experiment examines a recursive system for x_t, which involves a component (x_{1t}) with a root close to 1. The second experiment concerns an autoregressive matrix of rank 1, where all elements of volatility matrix are driven by a single dynamic factor [see Section 6]. Finally, in the third experiment, the two latent processes are independent with identical dynamics.

[Insert Figure 1: volatilities, example 1]

[Insert Figure 2: correlation, example 1]

[Insert Figure 3: canonical volatilities, example 1]

$\vdots$

[Insert Figure 9: canonical volatilities, example 3]

As expected, the bivariate WAR(1) model is able to reproduce volatility clustering phenomena, that is path dependent subperiods of large (resp. low) variances Y_{11t}, Y_{22t}, or path dependent subperiods of large (resp. low) λ_{1t}, λ_{2t}. We note that the clustering pattern is not necessarily identical for all portfolio volatilities.

In particular, we can observe simultaneously a cluster of high λ_{1t}, and a cluster of low λ_{2t}. In such a situation, the market has to manage the following two very different types of risks: 1) the common volatility risk associated with the first eigenvector, and 2) the risk due to the leverage effect of arbitragist strategies associated with the second eigenvector. This situation can occur, when some portfolio volatilities are negatively correlated.

Let us first discuss the volatility patterns. In experiment 1, we observe directly the lag of one time unit between the peaks and throughs, which is a consequence of the recursive form of matrix M [see Figure 1]. In experiment 2, the series are driven by the same factor, but the sensitivity coefficients with respect to the factors are different. Moreover, the conditional heteroscedasticity of volatility series renders the detection of the common factor difficult.

The correlations are quite specific in the case $n = K = 2$. Indeed, the case $K = 2$ is close to the degenerate case $K = 1$. If $K = 1$, the matrix Y_t is stochastic with rank 1 and the correlation alternates, taking randomly values +1 and -1. When $K = 2$, the matrix Y_t has rank 2 with probability 1, but the probability of correlation with absolute value close to one is significant. This feature is directly observed in Figures 2, 5, 8, in which we see highly fluctuating correlation. This effect generally diminishes when K increases, as shown in Section 3.2.

Finally, in Figures displaying the eigenvalues, we find dates at which λ_{1t} is rather large while λ_{2t} is close to zero. At such times, we face the aforementioned two types of risk, which are the common volatility risk and the arbitragist risk.

4 Predictions from WAR processes

The WAR processes belong to the family of compound autoregressive (Car) processes [see Darolles, Gourieroux, and Jasiak (2005)], which have simple prediction formulas due to the exponential affine representation of the conditional Laplace transform. In This section also discusses temporal aggregation of WAR processes. It will be shown that it resembles in many aspects the temporal aggregation of Gaussian VAR processes.

4.1 Prediction formulas and stationarity condition

Nonlinear forecasting of the matrix WAR(1) process Y at horizon h consists in computing the conditional distribution of Y_{t+h} given Y_t. The prediction formulas are based on the conditional Laplace transform at horizon h, which can be easily computed by recursions [see Darolles, Gourieroux, and Jasiak (2005)]. To keep our exposition simple, let us consider an integer-valued degree of freedom K. By definition, we have:

$$Y_{t+h} = \sum_{k=1}^{K} x_{k,t+h} x'_{k,t+h},$$

where $x_{k,t+h} = M^h x_{k,t} + \varepsilon_{k,t,h}$, $V(\varepsilon_{k,t,h}) = \Sigma + M\Sigma M' + \ldots + M^{h-1}\Sigma \left(M^{h-1}\right)' = \Sigma(h)$, say. This implies the following proposition.

Proposition 4: The transition distribution at horizon h of the WAR(1) process is the (conditional) Wishart $W_n\left(K, M^h, \Sigma(h)\right)$.

In particular, the WAR(1) process admits linear prediction formulas at any horizon. We have:

$$E\left[Y_{t+h}|Y_t\right] = M^h Y_t \left(M^h\right)' + K\Sigma\left(h\right).$$

The WAR(1) process is asymptotically strictly stationary if matrix M admits eigenvalues with a modulus strictly less than 1. The stationary (marginal) distribution of WAR(1) is centered Wishart $W\left[K, 0, \Sigma\left(\infty\right)\right]$, where $\Sigma\left(\infty\right)$ is the solution of the equation:

$$\Sigma\left(\infty\right) = M\Sigma\left(\infty\right)M' + \Sigma.$$

The prediction formulas are easily extended to WAR(p) process, which also is a compound autoregressive process (Car).

4.2 Temporal aggregation

Sections 2 and 3 examined a volatility matrix Y_t at horizon 1 based on the information set including the lagged values of Y_t and returns r_t. It is well-known that standard volatility models are not invariant with respect to time aggregation [see e.g. Drost and Nijman (1993), Drost and Werker (1996), Meddahi and Renault (2004)]. Let us consider a WAR(1) specification and study the volatilities and returns defined at a horizon of 2 time units, say. Let us first interpret the time aggregated volatility process:

$$\widetilde{Y}_{\tau+1} = Y_{2\tau} + Y_{2\tau+1}, \tau = 0, 1, 2, \ldots.$$

To do this, let us define the geometric return at horizon 2:

$$\widetilde{r}_{\tau+1} = r_{2\tau+1} + r_{2\tau+2},$$

and assume a zero expected return. We also assume that Y_t follows a WAR(1) stochastic volatility model on the time grid with time unit equal to one. When the information set at date $\tau = 2t$ includes the lagged values of the aggregate volatility and returns, we get:

$$\begin{aligned}
& V\left[\widetilde{r}_{\tau+1}|\underline{\widetilde{r}_{\tau}}, \underline{\widetilde{Y}_{\tau}}\right] \\
= \; & V\left[r_{2\tau+1} + r_{2\tau+2}|\underline{\widetilde{r}_{\tau}}, \underline{\widetilde{Y}_{\tau}}\right] \\
= \; & V\left[E\left(r_{2\tau+1} + r_{2\tau+2}|\underline{r_{2\tau}}, \underline{Y_{2\tau}}\right)|\underline{\widetilde{r}_{\tau}}, \underline{\widetilde{Y}_{\tau}}\right] \\
& + E\left[V\left(r_{2\tau+1} + r_{2\tau+2}|\underline{r_{2\tau}}, \underline{Y_{2\tau}}\right)|\underline{\widetilde{r}_{\tau}}, \underline{\widetilde{Y}_{\tau}}\right] \\
= \; & E\left[V\left(r_{2\tau+1} + r_{2\tau+2}|\underline{r_{2\tau}}, \underline{Y_{2\tau}}\right)|\underline{\widetilde{r}_{\tau}}, \underline{\widetilde{Y}_{\tau}}\right] \\
= \; & E\left[Y_{2\tau} + E\left(Y_{2\tau+1}|\underline{Y_{2\tau}}\right)|\underline{\widetilde{r}_{\tau}}, \underline{\widetilde{Y}_{\tau}}\right] \\
= \; & E\left[Y_{2\tau} + Y_{2\tau+1}|\underline{\widetilde{Y}_{\tau}}\right] \\
= \; & E\left[\widetilde{Y}_{\tau+1}|\widetilde{Y}_{\tau}\right].
\end{aligned}$$

Thus, the aggregate process $\widetilde{Y}_{\tau+1} = Y_{2\tau} + Y_{2\tau+1}$ is the process to be considered in the computation of volatility at horizon 2, equal to $E\left[\widetilde{Y}_{\tau+1}|\widetilde{Y}_{\tau}\right]$.

Let us now consider the expression of aggregate volatility in terms of the latent processes x (assuming an integer degree of freedom). We get:

$$Y_{2\tau} + Y_{2\tau+1} = \sum_{k=1}^{K} \left(x_{k,2\tau} x'_{k,2\tau} + x_{k,2\tau+1} x'_{k,2\tau+1} \right),$$

which is not a WAR(1) process, due to the presence of lags. However, the aggregate volatility $\tilde{Y}_{\tau+1}$ can be obtained from the (n^2, n^2) matrix:

$$Z_\tau = \sum_{k=1}^{K} \begin{pmatrix} x_{k,2\tau} \\ x_{k,2\tau+1} \end{pmatrix} \begin{pmatrix} x'_{k,2\tau} & x'_{k,2\tau+1} \end{pmatrix},$$

by summing the two diagonal blocks. The stacked process $\left(x'_{k,2\tau}, x'_{k,2\tau+1}\right)'$ is a Gaussian VAR(1) process:

$$\begin{pmatrix} x_{k,2\tau} \\ x_{k,2\tau+1} \end{pmatrix} = \begin{pmatrix} 0 & M \\ 0 & M^2 \end{pmatrix} \begin{pmatrix} x_{k,2\tau-2} \\ x_{k,2\tau-1} \end{pmatrix} + \begin{pmatrix} Id & 0 \\ M & Id \end{pmatrix} \begin{pmatrix} \varepsilon_{k,2\tau} \\ \varepsilon_{k,2\tau+1} \end{pmatrix}.$$

Since the process (Z_τ) is the sum of squares of the stacked Gaussian VAR, it follows that:

Proposition 5: The stochastic process (Z_τ) is a Wishart process of dimension $2n$: $W_{2n}\left(K, \begin{pmatrix} 0 & M \\ 0 & M^2 \end{pmatrix}, \begin{pmatrix} \Sigma & \Sigma M' \\ M\Sigma & \Sigma + M\Sigma M' \end{pmatrix}\right)$.

Thus, the process of aggregate volatilities is the sum of block diagonal elements of a Wishart process obtained by stacking the consecutive realizations of latent processes x_t. The approach, based on stacking, reveals that the cross-products $x_{k,2\tau} x'_{k,2\tau+1}$ have an affect on the distribution of block diagonal elements. We conclude that the volatility process at horizon 2, that is $E\left[\widetilde{Y}_{\tau+1}|\widetilde{Y}_{\tau}\right]$, is not a Wishart process of order 1, but a different process that can be computed from an augmented Wishart process of order 1.

5 Continuous time analogue

When the autoregressive coefficient M can be written as $M = \exp(A)$, where A is a matrix, the Wishart autoregressive process of order 1 is a time-discretized diffusion process. Moreover, if K is an integer the diffusion process is obtained by summing the squares of K independent multivariate Ornstein-Uhlenbeck processes.

Let us consider $K = 1$ and the multivariate Ornstein-Uhlenbeck process defined by:

$$dx_t = Ax_t dt + \Omega dw_t, \tag{8}$$

where (w_t) is a n-dimensional standard Brownian motion, and A and Ω are (n,n) matrices. It is well known that the time-discretized Ornstein-Uhlenbeck process is a Gaussian autoregressive process of order 1, where $M = \exp(A)$ and $\Sigma = \int_0^1 \exp(sA)\,\Omega\Omega'\,[\exp(sA)]'\,ds$.

The exponential function in the expression of autoregressive coefficient matrix implies restrictions on the dynamics of the associated discrete-time Gaussian AR(1) process. More precisely, the autoregressive matrix M cannot admit negative or zero eigenvalues. Thus, a number of Gaussian VAR(1) processes in discrete time, that are usually encountered in applications, cannot be considered as time discretizations of multivariate Ornstein-Uhlenbeck processes. Also, a number of Wishart processes of order one are not time discretized continuous-time processes. For M of dimension $(2,2)$ we have, for example,

i) the white noise Wishart process for $M = 0$;

ii) the periodic model with period 2 for $M = -Id$;

iii) the model with recursive dependence $M = \begin{pmatrix} 0 & 0 \\ -0.5 & 0 \end{pmatrix}$, where the latent process $x_{1t} = \varepsilon_{1t}$, $x_{2t} = \varepsilon_{2t} - 0.5\ x_{1,t-1} = \varepsilon_{2t} - 0.5\ \varepsilon_{1,t-1}$, is a moving average.

For any other $M = \exp(A)$, the WAR(1) process is a time-discretized diffusion process $Y_t = x_t x_t'$, where (x_t) is the Ornstein-Uhlenbeck process (equation 10).

Let us show the stochastic differential system satisfied by the continuous time matrix process (Y_t). It is proved in Appendix 4 that this matrix process satisfies:

$$\begin{aligned} dY_t &= (\Omega\Omega' + AY_t + Y_tA')\,dt + x_t\,(\Omega dw_t)' + \Omega dw_t x_t' \qquad (9) \\ &= (\Omega\Omega' + AY_t + Y_tA')\,dt + \sum_{l=1}^{n} (x_t\Omega_l' + \Omega_l x_t')\,dw_{lt}, \end{aligned}$$

where Ω_l, $l = 1, \ldots, n$, are the columns of matrix Ω. It is easy to check that the volatility matrix of $d\,(vecY_t)$ depends on Y_t only. Indeed, let us introduce: $vecY_t = \left(Y_t^{1\prime}, \ldots, Y_t^{n\prime}\right)'$, where Y_t^j, $j = 1, \ldots, n$, is the j^{th} column of Y_t. The Brownian component of dY_t^j is $\sum_{l=1}^{n} (x_t\omega_{jl} + \Omega_l x_{jt})\,dw_{lt}$. Thus, we infer:

$$\begin{aligned} & cov_t\left(dY_t^i, dY_t^j\right) \\ = \;& cov_t\left[\sum_{l=1}^{n} (x_t\omega_{il} + \Omega_l x_{it})\,dw_{lt}, \sum_{l=1}^{n} (x_t\omega_{jl} + \Omega_l x_{jt})\,dw_{lt}\right] \\ = \;& \left[\sum_{l=1}^{n} (x_t\omega_{il} + \Omega_l x_{it})\,(x_t\omega_{jl} + \Omega_l x_{jt})'\right] dt. \end{aligned}$$

This conditional covariance is a function of Σ and Y_t only:

$$cov_t\left(dY_t^i, dY_t^j\right) = \left(\sigma_{ij}Y_t + Y_t^j\left(\Sigma^i\right)' + \Sigma^j\left(Y_t^i\right)' + Y_{ij,t}\Sigma\right)dt, \tag{10}$$

where $\Sigma = \Omega\Omega'$. In particular, we find for any n-dimensional vectors α, β, γ, δ:

i) $cov_t\left(dY_t\alpha, dY_t\beta\right) = \left(\alpha'\Sigma\beta Y_t + Y_t\beta\alpha'\Sigma + \Sigma\beta\alpha' Y_t + \alpha' Y_t\beta\Sigma\right)dt,$

ii) $cov_t\left(\gamma' dY_t\alpha, \delta' dY_t\beta\right) = [(\alpha'\Sigma\beta)(\gamma' Y_t\delta) + (\gamma' Y_t\beta)(\alpha'\Sigma\delta)$
$+ (\gamma'\Sigma\beta)(\alpha' Y_t\delta) + (\alpha' Y_t\beta)(\gamma'\Sigma\delta)]dt,$

iii) $V_t\left(\gamma' dY_t\alpha\right) = [(\alpha'\Sigma\alpha)(\gamma' Y_t\gamma) + 2(\alpha' Y_t\gamma)(\alpha'\Sigma\gamma) + (\alpha' Y_t\alpha)(\gamma'\Sigma\gamma)]dt,$

iv) $V_t\left(\alpha' dY_t\alpha\right) = 4(\alpha'\Sigma\alpha)(\alpha' Y_t\alpha)\,dt,$

v) $cov_t\left(\alpha' dY_t\alpha, \beta' dY_t\beta\right) = 4(\alpha'\Sigma\beta)(\alpha' Y_t\beta)\,dt.$

These covariance formulas are the local counterparts of the discrete time formulas derived in Corollary 2. Indeed, for a small time increment dt, the formulas of Corollary 2 hold with $M = Id + o(dt)$ and Σ replaced by Σdt. In continuous time, only the terms of order dt in the volatility expression are relevant.

By construction, we know that the solution Y_t of the differential system (9-10) is symmetric positive semidefinite. The positivity condition follows from the the "drift" and "volatility" expressions. Indeed, let us consider a vector (portfolio allocation) α such that $\alpha' Y_t\alpha = 0$. The drift of $\alpha' dY_t\alpha$ is $\alpha'\Omega\Omega'\alpha dt \geq 0$, whereas its volatility is $V_t\left(\alpha' dY_t\alpha\right) = 0$. Thus, there is a reflection effect, which ensures that $\alpha' Y_t\alpha$ remains nonnegative. This argument is valid for any α.

The Wishart continuous-time process is easily extended to handle any degree of freedom K strictly greater than 0, integer or noninteger valued. To do that, we keep the volatility function unchanged, and change the drift to $K\Omega\Omega' + AY_t + Y_tA'$, and increase the number of independent Brownian motions up to dimension $\frac{n(n+1)}{2}$. When K is not an integer, the interpretation in terms of sums of squares of Ornstein-Uhlenbeck processes is no longer valid, but the symmetry and positivity of the solutions are ensured by the reflection argument given above.

The differential stochastic system satisfied by the Wishart process can be written as:

$$dvechY_t = \mu_t dt + \Lambda_t^{1/2} dW_t, \tag{11}$$

where (W_t) is an $n(n+1)/2$ dimensional Brownian motion, $\mu_t = vech(K\Omega\Omega' + AY_t + Y_tA')$ and $\Lambda_t \approx (1/dt)V_t(dvechY_t)$ has a complicated expression. An alternative representation of the continuous time process can be derived by analogy to the equation of unit root Wishart processes (see Section 3.4). It is easy to see that a continuous time Wishart process satisfies a system of the type :

$$dY_t = (\tilde{\Omega}\tilde{\Omega}' + \tilde{A}Y_t + Y_t\tilde{A}')dt + Y_t^{1/2}d\tilde{W}_tQ + Q'd\tilde{W}_t'Y_t^{1/2}, \tag{12}$$

where $\tilde{W}_t$ is a (n,n) stochastic matrix, whose components are independent Brownian motions and $\tilde{\Omega}, A, Q$ are (n,n) matrices . This representation is useful in computations, but can be misleading in the sense that the number of scalar Brownian motions is strictly larger than the number of linearly independent components of Y_t. Therefore, information generated by the $n(n+1)/2$ components of Y is strictly contained in the information set generated by the n^2 Brownian motions.

Example 1. The square of a univariate Ornstein-Uhlenbeck process $y_t = x_t^2$, where:

$$dx_t = ax_t dt + \omega dW_t,$$

satisfies the stochastic differential equation:

$$dy_t = \left(2ay_t + \omega^2\right) dt + 2\omega\sqrt{y_t}dW_t.$$

For another value K of the degree of freedom, we get:

$$dy_t = \left(2ay_t + K\omega^2\right) dt + 2\omega\sqrt{y_t}dW_t.$$

This is the Cox-Ingersoll-Ross (CIR) process [Cox, Ingersoll, Ross (1985)]. This result is not surprising since the CIR process is the continuous time analogue of the autoregressive gamma process. In particular, the square of an Ornstein-Uhlenbeck process is a special case of CIR process with a restriction on the mean reverting, volatility and equilibrium parameters [see Heston (1993)].

6 Reduced-rank (factor) models

In multivariate time series models, the number of parameters can be reduced by finding factor representations with a small number of factors. The factor representations can be defined a priori as in factor ARCH models, or else can be based on a coherent general-to-specific methodology as in multivariate linear autoregressive models. In this section, we develop a general-to-specific approach, which is based on the analysis of the rank, null-space and range of the autoregressive matrix. By considering a matrix M with reduced rank, we are able to define portfolio allocations with the following properties: 1) serially independent portfolio volatilities (white noise directions), 2) portfolio volatilities, which summarize relevant information (factor directions).

For ease of exposition, we first consider an autoregressive matrix of rank one, and next extend the results to matrices of any rank.

6.1 Matrix M of rank 1

Let us first consider a WAR(1) process with autoregressive matrix M of rank 1. This matrix can always be written as: $M = \beta\alpha'$, where β and α are two nonzero vectors of dimension n.

i) The conditional Laplace transform of the process Y_t is equal to:

$$\begin{aligned}\Psi_t(\Gamma) &= \frac{\exp Tr\left[\alpha\beta'\Gamma\left(Id - 2\Sigma\Gamma\right)^{-1}\beta\alpha' Y_t\right]}{\left[\det\left(Id - 2\Sigma\Gamma\right)\right]^{K/2}} \\ &= \frac{\exp Tr\left[\left(\beta'\Gamma\left(Id - 2\Sigma\Gamma\right)^{-1}\beta\right)\alpha' Y_t\alpha\right]}{\left[\det\left(Id - 2\Sigma\Gamma\right)\right]^{K/2}},\end{aligned}$$

since we can commute under the trace operator. Thus, the conditional Laplace transform depends on Y_t by the term $\alpha' Y_t \alpha$ only.

Proposition 6: When $M = \beta\alpha'$, the conditional Laplace transform depends on Y_t by the quadratic form (portfolio volatility) $\alpha' Y_t \alpha$ only.

Moreover, the dynamics of $\alpha' Y_t \alpha$ is easily characterized. Indeed, we have:

$$\begin{aligned}& E_t \exp\left(u\alpha' Y_{t+1}\alpha\right) \\ = \ & \Psi_t\left(u\alpha\alpha'\right) \\ = \ & \frac{\exp\left[\left(u\beta'\alpha\alpha'\left(Id - 2u\Sigma\alpha\alpha'\right)^{-1}\beta\right)\alpha' Y_t\alpha\right]}{\left[\det\left(Id - 2u\Sigma\alpha\alpha'\right)\right]^{K/2}}.\end{aligned}$$

This conditional Laplace transform represents a WAR(1) process of dimension 1, which has a noncentered chi-square transition distribution [see Appendix 5].

Proposition 7: When $M = \beta\alpha'$, the univariate process $(\alpha' Y_t \alpha)$ is a WAR(1) process $W_1(K,\, \alpha'\beta,\, \alpha'\Sigma\alpha)$.

Thus, we get a nonlinear one-factor model, with the dynamic factor $F_t = \alpha' Y_t \alpha$. More precisely, the factor process (F_t) admits autonomous dynamics, and, once the factor value is known, the conditional distribution of Y_{t+1} given Y_t is known and equal to the conditional distribution of Y_{t+1} given F_t. Note that in the standard CAPM model the asset return volatility matrix depends on the past by market portfolio volatility only, which implies that the matrix M is of rank one.

ii) It is also interesting to point out that there exist functions of the volatility matrix that destroy serial dependence. Let us consider a deterministic matrix C' with dimension (p, n). We focus on matrix process $(C' Y_t C)$, and consider integer K for ease of exposition. We get:

$$C' Y_{t+1} C = C' \sum_{k=1}^{K} x_{k,t+1} x'_{k,t+1} C$$

$$= C' \sum_{k=1}^{K} (\beta\alpha' x_{k,t} + \varepsilon_{k,t+1}) (\beta\alpha' x_{k,t} + \varepsilon_{k,t+1})' C.$$

This expression doesn't depend on the lagged values $(x_{k,t})$ if the columns of C are orthogonal to vector β. Moreover, $C'Y_{t+1}C = C' \sum_{k=1}^{K} \varepsilon_{k,t+1}\varepsilon'_{k,t+1}C$ will follow a WAR(1) process $W_p(K,\ 0,\ C'\Sigma C)$ of dimension p.

Proposition 8: Let us consider a matrix C of dimension $(n, n-1)$, whose columns span the vector space orthogonal to vector β. Then, the sequence of matrices $(C'Y_tC)$ is an i.i.d. sequence of Wishart variables $W_{n-1}(K,\ 0,\ C'\Sigma C)$ of dimension $n-1$.

Therefore, in the framework of a matrix M of rank one, we can define transformations of the stochastic volatility matrix, which either contain all sufficient information, or reveal the absence of serial dependence. Two cases can be distinguished:

1) If α is not orthogonal to β: $\alpha'\beta \neq 0$, we can compute portfolio volatilities with respect to a new basis of the vector space. More precisely, we can consider the transformed volatility matrix:

$$Y_{t+1}(A) = \begin{bmatrix} C'Y_{t+1}C & C'Y_{t+1}\alpha \\ \alpha'Y_{t+1}C & \alpha'Y_{t+1}\alpha \end{bmatrix},$$

corresponding to $A = (C, \alpha)$, where C is orthogonal to β. The first diagonal block is a white noise, while the second diagonal block captures all past information. The diagonal blocks are mutually independent.

2) If α and β are orthogonal: $\alpha'\beta = 0$, we can compute the volatilities with respect to a basis including the direction without serial dependence plus the β direction. In this case: $A = (C, \alpha, \beta)$, where C is a $(n, n-2)$ matrix with columns orthogonal to β and linearly independent of α. We get:

$$Y_{t+1}(A) = \begin{bmatrix} C'Y_{t+1}C & C'Y_{t+1}\alpha & C'Y_{t+1}\beta \\ \alpha'Y_{t+1}C & \alpha'Y_{t+1}\alpha & \alpha'Y_{t+1}\beta \\ \beta'Y_{t+1}C & \beta'Y_{t+1}\alpha & \beta'Y_{t+1}\beta \end{bmatrix}.$$

The portfolio volatility $\alpha'Y_{t+1}\alpha$ is a white noise process, which captures all relevant information.

6.2 Transformations of WAR(1) processes

We will now consider the general framework of matrix M of any rank and of an integer or noninteger valued degree of freedom. Let us consider a transformation $a'Y_{t+1}a$ of the volatility matrix, where a is a (n, p) matrix of full column rank. The conditional Laplace transform of this process is:

$$\widetilde{\Psi}_t(\gamma) = E\left[\exp Tr\left(\gamma a'Y_{t+1}a\right)|Y_t\right],$$

where γ is a symmetric (p,p) matrix. It can be written in terms of the basic Laplace transform :

$$\begin{aligned}\widetilde{\Psi}_t(\gamma) &= E\left[\exp Tr\left(\gamma a' Y_{t+1} a\right)|Y_t\right] \\ &= \Psi_t\left(a\gamma a'\right),\end{aligned}$$

since we can commute under the trace operator. We get:

$$\begin{aligned}\widetilde{\Psi}_t(\gamma) &= \frac{\exp Tr\left[M' a\gamma a' \left(Id - 2\Sigma a\gamma a'\right)^{-1} MY_t\right]}{\left[\det\left(Id - 2\Sigma a\gamma a'\right)\right]^{K/2}} \\ &= \frac{\exp Tr\left[\gamma a' \left(Id - 2\Sigma a\gamma a'\right)^{-1} MY_t M' a\right]}{\left[\det\left(Id - 2\Sigma a\gamma a'\right)\right]^{K/2}}.\end{aligned}$$

Thus, $(a'Y_t a)$ is a Markov process, if and only if, $MY_tM'a$ is function of $a'Y_ta$ (for any value Y_t), or equivalently if there exists a matrix Q such that $M'a = aQ'$. Moreover, it is easy to show that in this case $(a'Y_ta)$ still defines a Wishart process.

Proposition 9: Let us assume that (Y_t) is a Wishart process of order 1 $W_n(K, M, \Sigma)$ and consider a matrix a of dimension (n,p) and full column rank.

i) The transformed process $(a'Y_ta)$ is a Markov process if, and only if, there exists a (p,p) matrix Q such that $a'M = Qa'$.

ii) Under this condition, the process $(a'Y_ta)$ is also a Wishart process $W_p(K, Q, a'\Sigma a)$ of dimension p.

Condition i) of Proposition 9 is easy to interpret, when K is an integer and the Wishart process is written in terms of the latent processes x:

$$a'Y_ta = \sum_{k=1}^{K} a' x_{kt} x_{kt}' a = \sum_{k=1}^{K} z_{kt} z_{kt}',$$

where $z_{kt} = a'x_{kt} = a'Mx_{k,t-1} + a'\varepsilon_t$. The process (z_{kt}) is Gaussian autoregressive iff $a'Mx_{k,t-1}$ is a linear function of $z_{k,t-1}$, that is, iff there exists Q such that: $a'Mx_{k,t-1} = Qa'x_{k,t-1} = Qz_{k,t-1}$. Then, the parameters of the transformed Wishart process are the parameters of the new Gaussian autoregressive process (z_t).

6.3 Wishart processes with reduced rank

The results given above allow us to find the analogues of outcomes from Section 6.1 for a WAR(1) with an autoregressive matrix of any rank. Let us assume that the rank of this matrix is $l < n$. Then, the autoregressive matrix can be written as:

$$M = \beta\alpha', \tag{13}$$

where α and β are matrices with dimension (n, l) and full column rank.

The following two transformed processes have direct interpretations:

i) $(\alpha' Y_t \alpha)$ is a process which conveys all information, called the nonlinear dynamic factor process.

ii) $(C' Y_t C)$, where C is a matrix "orthogonal" to β, that is satisfying $C'\beta = 0$, is a white noise process.

Moreover, both transformed processes satisfy condition i) of Proposition 9 since:

i) $\alpha' M = \alpha' \beta \alpha' = Q\alpha'$, with $Q = \alpha'\beta$;

ii) $C'M = C'\beta\alpha' = 0 = 0\alpha'$, with $Q = 0$.

Proposition 9 implies the following properties.

Proposition 10: Let us assume $M = \beta\alpha'$, where α and β are (n, l) matrices with full column rank l.

i) The conditional distribution of Y_{t+1} depends on the past values Y_t by $\alpha' Y_t \alpha$ only.

ii) $(\alpha' Y_t \alpha)$ is a Wishart process $W_l\left(K, \alpha'\beta, \alpha'\Sigma\alpha\right)$ of dimension l.

iii) If C is a $(n, n-l)$ matrix such that $C'\beta = 0$, then $(C' Y_t C)$ is an i.i.d. Wishart process $W_{n-l}\left(K, 0, C'\Sigma C\right)$ of dimension $n - l$.

7 Stochastic volatility-in-mean

By analogy to the ARCH-in-mean process, we can formulate an expected return model with WAR-in-mean stochastic volatility [see Engle, Lilien, and Robbins (1987)]. The definition of the WAR-in-mean process is given in Section 7.1 and its predictive properties are described in Section 7.2.

7.1 Definition of the WAR-in-mean process

Let us consider the returns on n risky assets. The returns form a n-dimensional process (r_t). We assume that the distribution of r_{t+1} conditional on the lagged returns $\underline{r_t}$ and lagged volatilities $\underline{Y_t}$ is Gaussian with conditional variance Y_t and a conditional mean that is an affine function of Y_t[22].

Definition 3: The return process (r_t) is a WAR-in-mean process if the conditional distribution of r_{t+1} given $\underline{r_t}$, $\underline{Y_t}$ is Gaussian with a WAR(1) conditional variance-covariance matrix Y_t, and conditional mean $m_t = (m_{i,t})$ with components: $m_{i,t} = b_i + Tr\left(D_i Y_t\right)$, $i = 1, \ldots, n$, where b_i are scalars and D_i are (n, n) symmetric matrices of "risk premia".

[22] The assumption of normality concerns the distribution conditional on lagged returns and lagged volatilities. It is compatible with fat tails observed in the distribution conditional on lagged returns only.

For instance, for two asset returns the WAR-in-mean model becomes:

$$\left\{ \begin{array}{l} r_{1,t+1} = b_1 + d_{1,11}Y_{11,t} + 2d_{1,12}Y_{12,t} + d_{1,22}Y_{22,t} + \varepsilon_{1,t+1} \\ r_{2,t+1} = b_2 + d_{2,11}Y_{11,t} + 2d_{2,12}Y_{12,t} + d_{2,22}Y_{22,t} + \varepsilon_{2,t+1}, \end{array} \right.$$

where $V_t\left[\left(\varepsilon'_{1,t+1}, \varepsilon'_{2,t+1}\right)'\right] = Y_t$. The model allows for dependence of the expected return on volatilities and covolatilities.

The WAR-in-mean specification is useful for practical implementations, since the predictive distributions of returns are easy to compute by means of Laplace transforms. This is due to the expression of the conditional Laplace transform of the return r_{t+1} given $\underline{r_t}$, $\underline{Y_t}$, which is an exponential affine function of Y_t. Indeed, we have:

$$\begin{aligned} & E\left[\exp\left(z'r_{t+1}\right)|\underline{r_t}, \underline{Y_t}\right] \\ = \ & \exp\left[z'm_t + \frac{1}{2}z'Y_t z\right] \\ = \ & \exp\left[\sum_{i=1}^{n} z_i\left[b_i + Tr\left(D_i Y_t\right)\right] + \frac{1}{2}z'Y_t z\right] \\ = \ & \exp\left[z'b + Tr\left[\left(\sum_{i=1}^{n} z_i D_i + \frac{1}{2}zz'\right)Y_t\right]\right], \end{aligned}$$

Similar computations can easily be performed for more complicated specifications in which the conditional mean contains combinations of lagged returns or higher autoregressive orders.

Finally, note that, as mentioned in Section 5, under some parameter restrictions, some WAR processes can be seen as time discretized continuous time processes. The same remark holds for the WAR-in-mean process. When it admits a continuous time representation, the differential system for asset prices $S_{i,t}$ is:

$$d\log S_{i,t} = \left[b_i + Tr\left(D_i Y_t\right)\right]dt + Y_t^{1/2}dW_t^S,$$

where (Y_t) satisfies stochastic differential system (14) with a different multivariate Brownian motion. The tractability is due to the affine specification of the joint process $\left(vec\left(\log S_{i,t}\right), vec\left(Y_t\right)\right)$ that admits affine drift and volatility coefficients. This continuous-time specification can be considered as a multivariate extension[23] of the model:

$$\left\{ \begin{array}{rl} dS_t = & \left(\alpha + \beta\sigma_t^2\right)S_t dt + \sigma_t dW_t^S, \\ d\sigma_t^2 = & \left(\gamma_0 + \delta_0\sigma_t^2\right)dt + \sqrt{\gamma_1 + \delta_1\sigma_t^2}\,dW_t^\sigma, \end{array} \right.$$

introduced by Heston (1993).

[23]See Gourieroux, Sufana (2004)b for a use of this extended version to derive closed-form expressions for derivative prices in a multi-asset framework. This is another new frontier for ARCH models to be crossed, according to Engle (2002)b.

7.2 Mean-variance efficient portfolios

For a net return r_{it} defined as the difference between the return on asset i and the risk-free return, the Markowitz mean-variance efficient portfolio has an allocation proportional to:

$$a_t^* = (Y_t)^{-1} m_t.$$

Let us assume a WAR-in-mean process of net returns. When the volatility of net returns is equal to zero, the risky returns are equal to the risk-free return. Thus, we can assume $b_i = 0$, $\forall i$. Moreover, it is easy to see that the "risk premium" $Tr(D_i Y_t)$ is positive if the matrix D_i is positive definite. In this particular case, the risk premium is an increasing function of volatility Y_t[24], [25] . Thus, for a WAR-in-mean model, we get:

$$a_t^* = (Y_t)^{-1} vec\left[Tr(D_i Y_t)\right].$$

The positivity constraint on matrix D has a simple structural interpretation. The risk premium for asset i is equal to $Tr(D_i Y_t)$. Typically, it is a linear combination of volatilities and covolatilities such as: $d_{1,11}Y_{11,t} + 2d_{1,12}Y_{12,t} + d_{1,22}Y_{22,t}$ for $i = 1$, $n = 2$. The risk premium involves two components: Y_t measures the underlying joint risk, whereas $D = \begin{pmatrix} d_{1,11} & d_{1,12} \\ d_{1,12} & d_{1,22} \end{pmatrix}$ is a matrix of risk aversion coefficients describing the risk perceived by the market. As usual in a multiasset framework, the risk aversion is represented by a symmetric positive definite matrix. The combination of both effects determines the level of risk premium and explains the positivity of the risk premium, since $Tr(DY) \geq 0$, if $D \gg 0$ and $Y \gg 0$.

8 Statistical inference

Two types of statistical inference can be considered according to the type of available observations:

i) When a time-series of volatility matrices is available, a WAR model can be estimated directly from $Y_1, \ldots, Y_T$.

[24]Indeed, a positive definite matrix D can be written as $D = \sum_{k=1}^n d_k d_k'$. We get: $Tr(DY_t) = Tr\left(\sum_{k=1}^n d_k d_k' Y_t\right) = \sum_{k=1}^n Tr\left(d_k d_k' Y_t\right) = \sum_{k=1}^n Tr\left(d_k' Y_t d_k\right) = \sum_{k=1}^n d_k' Y_t d_k \geq 0$, since Y_t is a volatility.

Moreover, if two values of the volatility Y_t and Y_t^* are such that: $Y_t \gg Y_t^* \iff Y_t - Y_t^* \gg 0$, we deduce that: $Tr\left[D_i\left(Y_t - Y_t^*\right)\right] = Tr(D_i Y_t) - Tr\left(D_i Y_t^*\right) \geq 0$, which is the monotonicity property of the risk premium.

[25]However, Abel (1988), Backus, Gregory (1993) and Gennotte, Marsh (1993) offer models where a negative relation between expected return and variance is compatible with equilibrium. This is mainly due to the partial interpretation of this relationship, which does not necessarily account for all state variables. It would be natural to examine this financial puzzle in a multiasset framework to see how the matrix D and its positivity conditions depend on the number of assets.

ii) When asset returns are observed while the stochastic volatility is unobserved, a WAR-in-mean model can be estimated and latent volatilities approximated by a nonlinear filter.

In this section, we focus on the first type of statistical inference, which has at least two interesting applications.

i) From high frequency data, it is possible to compute daily volatility matrices of returns at a given frequency (for example sampled at 5 minute intervals) to obtain a series of intraday volatility matrices[26]. Due to different order matching procedures at market opening and closure (auction), and within the day (continuous trading), the dynamics of the intraday volatility matrices can be different from the dynamics of volatilities of daily returns computed from closing prices [see Gourieroux and Jasiak (2002), Chapter 14, for a description of electronic financial markets].

ii) Another application concerns the dynamics of derivative prices. In a multiasset framework, the Black-Scholes formula can be used to compute implied volatility matrices from derivative prices written on a set of n assets. The WAR specifications can be applied to series of implied volatilities and covolatilities [see e.g. Stapleton, Subrahmanyam (1984) for contingent claims whose payoffs are written on two or more assets].

In the sequel, we first discuss identification of the parameters of interest. Next, we introduce a first-order method of moments, which provides consistent estimators and is easy to implement. This method can be seen as the first step before numerical implementation of maximum likelihood based on the expression of the transition density given in Section 2.2. Finally, we discuss estimation of the WAR-in-mean model.

8.1 Identification

The identifiable [resp. first-order identifiable] parameters are obtained by considering the expressions of the conditional Laplace transform [resp. the conditional first-order moment]. The following identification results are proved in Appendix 6.

Proposition 11: Let us assume $K > (n-1)$.

i) K and Σ are identifiable while the autoregressive coefficient M is identifiable up to its sign.

ii) Σ is first-order identifiable[27] up to a scale factor and M is first-order identifiable up to its sign. The degree of freedom K is not first-order identifiable, but is second-order identifiable[28].

[26]Called realized volatility in the literature [see e.g. Andersen, Bollerslev, and Diebold (2002) for a survey].

[27]That is, identifiable from the first-order conditional moment.

[28]That is, identifiable from the first and second order conditional moments.

At order one the number of identifiable structural parameters is $n^2 + \frac{n(n+1)}{2}$ (for M and $\Sigma^* = K\Sigma$). The number of reduced form parameters in the prediction formula $E(Y_{t+1}|Y_t)$ is $\left[\frac{n(n+1)}{2}\right]^2 + \frac{n(n+1)}{2}$ (which are the number of slope plus intercept coefficients, respectively, in the seemingly unrelated regression of $vech(Y_t)$ on $vech(Y_{t-1})$ plus constant). The degree of (first-order) over-identification $\left[\frac{n(n+1)}{2}\right]^2 - n^2 = \frac{n^2(n-1)(n+3)}{4}$, is equal to zero for $n = 1$ and increases quickly with the number of assets.

Table 1. Degree of first-order over-identification.

Number of assets	1	2	3	4	5
Degree of over-identification	0	5	27	84	200

Thus, more accurate estimators are likely obtained when the cross sectional dimension n increases. This is due to the presence of second order cross moments among the moment restrictions.

Finally, statistical inference concerning the rank of M, its null-space and range can be performed (consistently) using conditional moments of order one, since these do not depend on the sign of matrix M.

8.2 First-order method of moments

The first-order conditional moments can be used to calibrate the parameters M and Σ, up to the sign and scale factor, respectively. The first-order method of moments is equivalent to nonlinear least squares. The ordinary nonlinear least squares estimators are defined as:

$$\left(\widehat{M}, \widehat{\Sigma}^*\right) = Arg \min_{M,\Sigma^*} S^2(M, \Sigma^*),$$

where:

$$\begin{aligned} S^2(M,\Sigma^*) &= \sum_{t=2}^{T} \sum_{i<j} \left(Y_{ij,t} - \sum_{k=1}^{n} \sum_{l=1}^{n} Y_{kl,t-1} m_{ik} m_{lj} - \sigma^*_{ij} \right)^2 \\ &= \sum_{t=2}^{T} \left\| vech(Y_t) - vech(MY_{t-1}M' + \Sigma^*) \right\|^2, \end{aligned}$$

and $\Sigma^* = K\Sigma$. This method can be applied by using any software that accounts for conditional heteroscedasticity. It can be improved by applying quasi-generalized nonlinear least squares, since the expression of $V_t[vech(Y_{t+1})]$ becomes known, once the degree of freedom K is estimated [see Corollary 2].

Once the parameters M and Σ^* are estimated, different tests can be performed on matrix M.

i) First we can check the rank of M, that is test for a reduced rank model. For instance, if the rank is equal to l the matrix M can be written as $M = \beta\alpha'$, where α and β have dimension (n, l) and are full column rank. Then an asymptotic least squares estimator of M under the hypothesis $RkM = l$ is defined by [see Gourieroux, Monfort, Renault (1995)] :

$$\hat{M}_l = \hat{\beta}_l' \hat{\alpha}_l,$$

where :

$$(\hat{\alpha}_l, \hat{\beta}_l) = \arg\min_{\alpha,\beta} [vec\, \hat{M} - vec\, (\beta\alpha')]' \hat{Var}(vec\hat{M})^{-1} [vec\hat{M} - vec(\beta\alpha')],$$

and the minimization is performed under the identifying restrictions $\alpha'\alpha = Id$. This optimization is similar to singular value decomposition of a well-chosen symmetric matrix computed from $\hat{M}$ and its asymptotic-covariance matrix.

ii) Second, we can test for embeddability, that is for the possibility to write $M = \exp A$. This test can be based on the spectral decomposition of $\hat{M}$.

8.3 Estimation of the degree of freedom

Finally, the degree of freedom K and the latent covariance matrix can be identified from the second-order moments. Indeed, the marginal distribution of the process (Y_t) is a centered Wishart distribution (see Section 4.1), such that :

$$\begin{aligned} V(\alpha' Y_t \alpha) &= 2K[\alpha'\Sigma(\infty)\alpha]^2 \\ &= \tfrac{2}{K^{-1}}[\alpha'\Sigma^*(\infty)\alpha]^2, \end{aligned}$$

where : $\Sigma^*(\infty) = M\Sigma^*(\infty)M' + \Sigma^*$.

Consistent estimators of the degree of freedom can be derived in the following way.

step 1 : Compute $\hat{\Sigma}^*(\infty)$ as a solution of :

$$\hat{\Sigma}^*(\infty) = \hat{M}\hat{\Sigma}^*(\infty)\hat{M}' + \hat{\Sigma}^*.$$

step 2 : Choose a portfolio allocation α, say, and compute its sample volatility

$$\hat{V}(\alpha' Y_t \alpha) = \frac{1}{T}\sum_{t=1}^{T}\left[\alpha' Y_t \alpha - \frac{1}{T}\sum_{t=1}^{T}\alpha' Y_t \alpha\right]^2.$$

step 3 : A consistent estimator of K is :

$$\hat{K}(\alpha) = 2[\alpha'\hat{\Sigma}^*(\infty)\alpha]^2 / \hat{V}(\alpha' Y_t \alpha).$$

step 4 : A consistent estimator of Σ is $\hat{\Sigma}(\alpha) = \hat{\Sigma}^* / \hat{K}(\alpha)$.

In practice, it can be useful to compare the estimators computed from different portfolio allocations to construct a specification test of the WAR process.

The two-step estimation method described above is simple to implement. It suggests associated specification tests and is suitable for a general-to-specific approach. However, it has a shortcoming of inefficiency.

Full efficiency can be reached in a second step by applying the maximum likelihood, that is by maximizing

$$
\begin{aligned}
L_T \quad = \quad & \sum_{t=1}^{T} \Big\{ -\frac{Kn}{2}\log 2 - \log \Gamma_n(K/2) - \frac{K}{2}\log\det\Sigma \\
& -\frac{K}{2}\log\det(Id - \frac{1}{2}\Sigma^{-1}MY_tM'\Sigma^{-1}Y_{t+1}) + \frac{K-n-1}{2}\log\det Y_{t+1} \\
& -\frac{1}{2}Tr[\Sigma^{-1}(Y_{t+1} + MY_tM')] + \log {}_0F_1(K/2;(1/4)MY_tM'Y_{t+1}) \Big\}.
\end{aligned}
$$

Similarly, some standard methods can be applied to the WAR-in-mean model, which is a special case of a nonlinear factor model. Such methods are Monte-Carlo Markov Chain and optimal filtering via particle filters [see Pitt, Shephard (1999), and Chib (2001) for an extensive review].

9 Dynamics of intraday volatility

9.1 The data

In this section, we consider a series of intraday historical volatility-covolatility matrices. They correspond to three stocks : ABX (Barrick Gold), BCE (Bell Canada Enterprise), NTL (Northern Telecom) traded on the Toronto Stock Exchange (TSX). Since the TSX is an electronic market with continuous trading throughout the day, high frequency data on quotes and trades are available. For each stock the (trade) returns are computed at 5 minute intervals, and used to compute the historical volatility-covolatility matrices at 5 minutes for each day [29]. This leads to 72 observations per day available to compute each matrix, since the market during the sampling period was opened between 9:30 a.m. and 4:30 p.m., and the first and last 30 minutes were deleted to remove the opening and closure effects. For estimation, we retained a sample covering one month of trading in October 1998, which consists of data on 21 working days intraday volatility matrices. Although a longer series could have been considered, this exercise allows us to check if the WAR model can be applied by rolling, as it is done by financial practitioners. It would also show if the WAR provides reasonable fit even when estimated from a sample of one month length. It is important to note that the number of observed variables is much greater than 21. Indeed, the observations concern a symmetric matrix (3,3) with 6 different elements. In particular, for a WAR model with lag one, we get : 120 = (21-1) x 6 observations, which is sufficient to estimate 16 parameters in M, Σ, K. Thus,

[29] All returns are multiplied by 10^3 for standardization.

the cross-sectional dimension permits to improve the accuracy of estimators (see the discussion of overidentification in Section 8.1).

The evolution of intraday volatility matrices is shown in Figures 10-12.

[Insert Figure 10 : Stock Return Volatilities].

The returns volatilities are displayed in Figure 10, in which some common market effects can be observed. For instance all volatilities increase simultaneously on day 10.

The evolution of return correlations is displayed in Figure 11. Here, other factor effects can be detected. For instance, on day 8, all correlations decrease quickly. The correlations take mostly values between 0.2 and 0.6 during the whole month.

[Insert Figure 11 : Stock Return Correlations]

Finally, the eigenvalues of the volatility matrices are displayed in Figure 12. On day 3, we observe a decrease of the smallest eigenvalue while the two other ones increase [see the discussion of the Monte-Carlo study of Section 3.5]. Such an effect contradicts the standard one-factor market model.

[Insert Figure 12 : Eigenvalues]

9.2 Unconstrained estimation

The WAR (1) model [30] is estimated by the first-order method of moments from the same data set. The unconstrained estimators of M and Σ^* are provided in Tables 2 and 3. The estimation time on an 1997 IBM Unix server was less than 1 minute.

The latent autoregressive coefficient matrix is highly significant, which leads to the rejection of the time deformed models with deterministic drift recently introduced in the literature for derivation of properties of (one-dimensional) observed realized volatilities [see e.g. Madan, Seneta (1990), Andersen, Bollerslev, Diebold, Labys (2001) for time deformed Brownian motion of the underlying return process, or Barndorff-Nielsen, Shephard (2003) for the extension to time deformed Levy processes].

[30] As already mentioned, the advantage of the WAR(1) process is that it naturally represents a process of symmetric positive definite matrices. An analogue domain restriction has not been taken into account by Andersen et alii (2003). In their paper, exchange rates data are studied and assumed to follow a normal model for (y_{1t}, y_{2t}, y_{3t}), where y_{1t}(resp y_{2t}, y_{3t}) is the logarithmic volatility for DM/\$ [resp. Y/\$, Y/DM]. Since the log-exchange rates satisfy a deterministic relationship, we see that $y_{1t} = \exp \sigma_{11t}, y_{2t} = \exp \sigma_{22t}, y_{3t} = \exp(\sigma_{11t} + \sigma_{22t} - 2\sigma_{12t})$, where σ_{12t} is the covolatility between the two first log-exchange rates. There is a one to one relationship between y_{1t}, y_{2t}, y_{3t} and $\sigma_{11t}, \sigma_{22t}, \sigma_{12t}$. The standard Cauchy-Schwartz inequality $\sigma_{12t}^2 \leq \sigma_{11t}\sigma_{22t}$ implies a complicated nonlinear constraint on the three log-volatilities. It is not taken into account in the multivariate Gaussian model (see Andersen et alii (2003), page 599).

The eigenvalues of the estimated matrix $\hat{M}$ are given in Table 4. They are all real, nonnegative and strictly less than one. This indicates that the process can be considered as a time discretized version of a continuous time process [31], and satisfies the stationarity conditions.

Table 2 : Estimated Latent Autoregressive Matrix M
(t-ratios in parentheses)

0.806 (4.09)	0.066 (0.63)	−0.474 (2.85)
0.377 (1.79)	0.300 (2.42)	0.168 (0.88)
1.017 (1.60)	0.120 (0.48)	−0.532 (1.42)

Table 3 : Estimated Latent Covariance Matrix Σ^*
(t-ratios in parentheses)

2.524 (1.28)	1.737 (1.68)	−1.361 (0.34)
	6.266 (4.48)	0.732 (0.55)
		7.040 (0.86)

Table 4 : Eigenvalues of $\hat{M}$

0.323	0.207	0.042

Table 5 provides the eigenvalues of $\hat{M}\hat{M}'$. We can see that the smallest eigenvalue is much smaller than all other ones. Thus, a two factor model can likely be considered a suitable representation.

Table 5 : Eigenvalues of $\hat{M}\hat{M}'$

2.291	0.179	$1.973e-0.5$

The degree of freedom has been estimated from the marginal second order moment corresponding to the equiweighted portfolio allocation $\alpha = (1,1,1)$. It is equal to : $\hat{K}(\alpha) = 4.25$, with a confidence interval of $[3.82, 5.54]$. The degree of freedom is strictly larger than 3, which ensures a nondegenerate Wishart process.

[31] This can be useful in further financial applications, like derivative pricing in continuous time [see e.g. Gourieroux, Sufana (2003), (2004), Gourieroux, Monfort, Sufana (2004)].

Moreover, other estimations of K based on different portfolio allocations have been considered [see Table 6]. They provide estimates within the confidence interval reported above, which favours the Wishart specification.

Table 6 : Estimated Degree of Freedom K

porfolio allocation	$(1,1,0)$	$(0,1,1)$	$(1,0,1)$
$\hat{K}(\alpha)$	3.82	4.89	4.66

9.3 Estimated reduced rank model

A two factor Wishart model has been reestimated from the same data set. The constrained estimators are provided in Tables 7 and 8.

Table 7 : Constrained Latent Autoregressive Matrix
(t-ratios in parentheses)

0.808 (3.14)	0.063 (0.38)	−0.472 (3.02)
0.377 (1.78)	0.299 (2.52)	0.167 (0.91)
1.014 (1.74)	0.121 (0.57)	−0.524 (1.51)

Table 8 : Constrained Latent Covariance Matrix
(t-ratios in parentheses)

2.519 (1.21)	1.739 (1.66)	−1.359 (0.34)
	6.266 (4.48)	0.730 (0.55)
		6.075 (0.89)

In the model of rank 2, the β-space is generated by the first two columns of the $\hat{M}$ matrix given in Table 7, whereas the α-space is generated by the rows of $\hat{M}$ and is orthogonal to the vector $(0.697, -1.439, 1)$.

Since the components of the first two columns of $\hat{M}$ are positive, the C vector orthogonal to these columns has some positive and negative elements. In some sense the "white noise" direction corresponds to a particular "arbitrage" portfolio.

10 Concluding remarks

The Wishart Autoregressive process provides an interesting alternative to standard multivariate GARCH and stochastic variance models. The WAR specification is quite flexible in the sense that it allows for the presence of autoregressive lags higher than one and provides a staightforward factor representation. The nonlinear prediction formulas have closed-form expressions at all horizons and are quite easy to compute as well. It is well-known that the CIR diffusion process can be interpreted as the limit of well-chosen ARCH processes [Nelson (1990)]. Likely, the continuous time WAR process could also be shown as the limit of well-chosen multivariate ARCH models. However, the discrete time WAR seems more convenient in many applications.

The WAR process can be used to model the dynamics of volatility matrices in financial applications, including derivative pricing and hedging. The WAR underlies the quadratic term structure model [Gourieroux, Sufana (2003)] and yields closed-form expressions of derivative prices in multivariate stochastic volatility models in which it arises as the multivariate extension of Heston's model [Gourieroux, Sufana (2004)]. The WAR also provides a coherent specification for the dynamics of stock prices, exchange rates and interest rates [Gourieroux, Monfort, Sufana (2004b)].

APPENDICES

Appendix 1 : Proof of Proposition 1

i) Let us first establish a preliminary lemma.

Lemma: For any symmetric semi-definite matrix Ω with dimension (n,n) and any vector $\mu \in R^n$, we get:

$$\int_{R^n} \exp\left(-x'\Omega x + \mu' x\right) dx = \frac{\pi^{n/2}}{(\det \Omega)^{1/2}} \exp\left(\frac{1}{4}\mu'\Omega^{-1}\mu\right).$$

Proof. Indeed, the integral on the left hand side is equal to:

$$\begin{aligned}
&\int_{R^n} \exp\left[-\left(x - \frac{1}{2}\Omega^{-1}\mu\right)' \Omega \left(x - \frac{1}{2}\Omega^{-1}\mu\right)\right] \exp\left(\frac{1}{4}\mu'\Omega^{-1}\mu\right) dx \\
= &\ \frac{\pi^{n/2}}{(\det \Omega)^{1/2}} \exp\left(\frac{1}{4}\mu'\Omega^{-1}\mu\right),
\end{aligned}$$

since the Gaussian multivariate distribution with mean $\frac{1}{2}\Omega^{-1}\mu$ and covariance matrix $2\Omega^{-1}$ admits unit mass. ∎

ii) Let us now prove Proposition 1, for $K = 1$. Let us consider the stochastic process (Y_t) defined by $Y_t = x_t x_t'$, $x_{t+1} = M x_t + \Sigma^{1/2}\xi_{t+1}$ and $\xi_{t+1} \sim IIN(0, Id)$. The conditional Laplace transform of the process (Y_t) is:

$$\begin{aligned}
&\Psi_t(\Gamma) \\
= &\ E\left[\exp\left(x_{t+1}'\Gamma x_{t+1}\right)|x_t\right] \\
= &\ E\left[\exp\left(\left(M x_t + \Sigma^{1/2}\xi_{t+1}\right)' \Gamma \left(M x_t + \Sigma^{1/2}\xi_{t+1}\right)\right)|x_t\right] \\
= &\ \exp\left(x_t' M'\Gamma M x_t\right) E\left[\exp\left(2x_t' M'\Gamma\Sigma^{1/2}\xi_{t+1} + \xi_{t+1}'\Sigma^{1/2}\Gamma\Sigma^{1/2}\xi_{t+1}\right)|x_t\right].
\end{aligned}$$

By using the pdf of standard normal,

$$f(\xi_{t+1}) = \frac{1}{2^{n/2}\pi^{n/2}} \exp -\frac{1}{2}\xi_{t+1}'\xi_{t+1},$$

and the Lemma, we get:

$$\begin{aligned}
&\Psi_t(\Gamma) \\
= &\ \frac{\exp\left(x_t' M'\Gamma M x_t\right)}{2^{n/2}\left[\det\left(\frac{1}{2}Id - \Sigma^{1/2}\Gamma\Sigma^{1/2}\right)\right]^{1/2}} \\
&\ \exp\left[\frac{1}{4}\left(2x_t' M'\Gamma\Sigma^{1/2}\right)\left(\frac{1}{2}Id - \Sigma_t^{1/2}\Gamma\Sigma_t^{1/2}\right)^{-1}\left(2\Sigma^{1/2}\Gamma M x_t\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\exp\left(x_t' M' \Gamma M x_t + 2 x_t' M' \Gamma \left(\Sigma^{-1} - 2\Gamma\right)^{-1} \Gamma M x_t\right)}{\left[\det\left(Id - 2\Sigma^{1/2}\Gamma\Sigma^{1/2}\right)\right]^{1/2}} \\
&= \frac{\exp\left[x_t' M' \Gamma \left(Id - 2\Sigma\Gamma\right)^{-1} M x_t\right]}{\left[\det\left(Id - 2\Sigma^{1/2}\Gamma\Sigma^{1/2}\right)\right]^{1/2}} \\
&= \frac{\exp Tr\left[M' \Gamma \left(Id - 2\Sigma\Gamma\right)^{-1} M Y_t\right]}{\left[\det\left(Id - 2\Sigma\Gamma\right)\right]^{1/2}}.
\end{aligned}
$$

This formula is valid whenever $Id - 2\Sigma\Gamma$ is a positive definite matrix.

iii) In the general case of integer K, the process can be written as: $Y_t = \sum_{k=1}^{K} Y_{kt}$, where the matrix processes $Y_{kt} = x_{kt} x_{kt}'$ are independent with the Laplace transform given above in ii). We find that:

$$
\begin{aligned}
\Psi_t(\Gamma) &= \prod_{k=1}^{K} \frac{\exp Tr\left[M' \Gamma \left(Id - 2\Sigma\Gamma\right)^{-1} M Y_{kt}\right]}{\left[\det\left(Id - 2\Sigma\Gamma\right)\right]^{1/2}} \\
&= \frac{\exp Tr\left[M' \Gamma \left(Id - 2\Sigma\Gamma\right)^{-1} M Y_t\right]}{\left[\det\left(Id - 2\Sigma\Gamma\right)\right]^{K/2}}.
\end{aligned}
$$

Appendix 2 : Conditional moments of the WAR(1) process

We provide the proofs for integer K. When K is not an integer, the results are derived from an expansion of the conditional log-Laplace transform.

Appendix A.2.1 : Conditional mean

We have:

$$
\begin{aligned}
E\left(Y_{t+1} | Y_t\right) &= E\left(\sum_{k=1}^{K} x_{k,t+1} x_{k,t+1}' | x_t\right) \\
&= \sum_{k=1}^{K} E\left(x_{k,t+1} x_{k,t+1}' | x_t\right) \\
&= \sum_{k=1}^{K} E\left(x_{k,t+1} | x_t\right) E\left(x_{k,t+1}' | x_t\right) + \sum_{k=1}^{K} V\left(x_{k,t+1} | x_t\right),
\end{aligned}
$$

where the last equality follows from the definition of the variance-covariance matrix. Thus, we obtain:

$$
\begin{aligned}
E\left(Y_{t+1} | Y_t\right) &= M \sum_{k=1}^{K} x_{k,t} x_{k,t}' M' + \sum_{k=1}^{K} (\Sigma) \\
&= M Y_t M' + K\Sigma.
\end{aligned}
$$

Appendix A.2.2 : Conditional variance

Let us consider $K = 1$. We get:

$$\begin{aligned}
& cov_t\left(\gamma' Y_{t+1}\alpha, \delta' Y_{t+1}\beta\right) && (14)\\
= \ & cov_t\left(\gamma'\left(Mx_t+\varepsilon_{t+1}\right)\alpha'\left(Mx_t+\varepsilon_{t+1}\right), \delta'\left(Mx_t+\varepsilon_{t+1}\right)\beta'\left(Mx_t+\varepsilon_{t+1}\right)\right)\\
= \ & cov_t(\gamma' Mx_t\alpha'\varepsilon_{t+1}+\gamma'\varepsilon_{t+1}\alpha' Mx_t+\gamma'\varepsilon_{t+1}\alpha'\varepsilon_{t+1},\\
& \delta' Mx_t\beta'\varepsilon_{t+1}+\delta'\varepsilon_{t+1}\beta' Mx_t+\delta'\varepsilon_{t+1}\beta'\varepsilon_{t+1})\\
= \ & E_t\left[\left(\gamma' Mx_t\alpha'\varepsilon_{t+1}+\gamma'\varepsilon_{t+1}\alpha' Mx_t\right)\left(\delta' Mx_t\beta'\varepsilon_{t+1}+\delta'\varepsilon_{t+1}\beta' Mx_t\right)\right]\\
& +cov_t\left(\gamma'\varepsilon_{t+1}\alpha'\varepsilon_{t+1}, \delta'\varepsilon_{t+1}\beta'\varepsilon_{t+1}\right),
\end{aligned}$$

where the other terms are zero, since they cannot be written as quadratic functions of x_t. Using the fact that $E_t\left(\varepsilon_{t+1}\varepsilon_{t+1}'\right) = \Sigma$, the first term in the above expression can be written as:

$$\begin{aligned}
& E_t\left[\left(\gamma' Mx_t\alpha'\varepsilon_{t+1}+\gamma'\varepsilon_{t+1}\alpha' Mx_t\right)\left(\delta' Mx_t\beta'\varepsilon_{t+1}+\delta'\varepsilon_{t+1}\beta' Mx_t\right)\right] && (15)\\
= \ & E_t\left[\left(\gamma' Mx_t\alpha'\varepsilon_{t+1}+\alpha' Mx_t\gamma'\varepsilon_{t+1}\right)\left(\varepsilon_{t+1}'\beta x_t' M'\delta+\varepsilon_{t+1}'\delta x_t' M'\beta\right)\right]\\
= \ & \gamma' Mx_t\alpha'\Sigma\beta x_t' M'\delta+\gamma' Mx_t\alpha'\Sigma\delta x_t' M'\beta+\alpha' Mx_t\gamma'\Sigma\beta x_t' M'\delta\\
& +\alpha' Mx_t\gamma'\Sigma\delta x_t' M'\beta\\
= \ & \gamma' MY_t M'\delta\alpha'\Sigma\beta+\gamma' MY_t M'\beta\alpha'\Sigma\delta+\alpha' MY_t M'\delta\gamma'\Sigma\beta\\
& +\alpha' MY_t M'\beta\gamma'\Sigma\delta.
\end{aligned}$$

Let $\varepsilon_{t+1} = \Sigma^{1/2}\xi_{t+1}$, where $\xi_{t+1} \sim IIN\left(0, Id\right)$. The second term in expression (16) becomes:

$$\begin{aligned}
& cov_t\left(\gamma'\varepsilon_{t+1}\alpha'\varepsilon_{t+1}, \delta'\varepsilon_{t+1}\beta'\varepsilon_{t+1}\right)\\
= \ & \gamma' cov_t\left(\varepsilon_{t+1}\varepsilon_{t+1}'\alpha, \varepsilon_{t+1}\varepsilon_{t+1}'\beta\right)\delta\\
= \ & \gamma' E_t\left(\varepsilon_{t+1}\varepsilon_{t+1}'\alpha\beta'\varepsilon_{t+1}\varepsilon_{t+1}'\right)\delta-\gamma' E_t\left(\varepsilon_{t+1}\varepsilon_{t+1}'\alpha\right)E_t\left(\beta'\varepsilon_{t+1}\varepsilon_{t+1}'\right)\delta\\
= \ & \gamma'\Sigma^{1/2}E_t\left[\xi_{t+1}\xi_{t+1}'\left(\Sigma^{1/2}\alpha\beta'\Sigma^{1/2}\right)\xi_{t+1}\xi_{t+1}'\right]\Sigma^{1/2}\delta-\gamma'\Sigma\alpha\beta'\Sigma\delta\\
= \ & \gamma'\Sigma^{1/2}E_t\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\xi_{i,t+1}\xi_{t+1}\left(b_{ij}\right)\xi_{j,t+1}\xi_{t+1}'\right]\Sigma^{1/2}\delta-\gamma'\Sigma\alpha\beta'\Sigma\delta\\
= \ & \gamma'\Sigma^{1/2}\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}E_t\left(\xi_{i,t+1}\xi_{j,t+1}\xi_{t+1}\xi_{t+1}'\right)\Sigma^{1/2}\delta-\gamma'\Sigma\alpha\beta'\Sigma\delta,
\end{aligned}$$

where $B = \Sigma^{1/2}\alpha\beta'\Sigma^{1/2}$. Let e_i be the canonical vector with zero components except the i^{th} component which is equal to 1, and δ_{ij} be the Kronecker symbol: $\delta_{ij} = 1$ if $i = j$, and 0, otherwise. Since $E_t\left(\xi_{i,t+1}\xi_{j,t+1}\xi_{t+1}\xi_{t+1}'\right) = \delta_{ij}Id + e_ie_j' + e_je_i'$ [see e.g. Bilodeau and Brenner (1999), page 75], we have:

$$cov_t\left(\gamma'\varepsilon_{t+1}\alpha'\varepsilon_{t+1}, \delta'\varepsilon_{t+1}\beta'\varepsilon_{t+1}\right) \qquad (16)$$

$$
\begin{aligned}
&= \gamma'\Sigma^{1/2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\left(\delta_{ij}Id + e_i e_j' + e_j e_i'\right)\Sigma^{1/2}\delta - \gamma'\Sigma\alpha\beta'\Sigma\delta \\
&= \gamma'\Sigma^{1/2}\left[Tr\left(B\right)Id + B + B'\right]\Sigma^{1/2}\delta - \gamma'\Sigma\alpha\beta'\Sigma\delta \\
&= \gamma'\Sigma^{1/2}Tr\left(B\right)\Sigma^{1/2}\delta + \gamma'\Sigma^{1/2}B\Sigma^{1/2}\delta + \gamma'\Sigma^{1/2}B'\Sigma^{1/2}\delta - \gamma'\Sigma\alpha\beta'\Sigma\delta \\
&= \gamma'\Sigma^{1/2}\left(\beta'\Sigma\alpha\right)\Sigma^{1/2}\delta + \gamma'\Sigma\alpha\beta'\Sigma\delta + \gamma'\Sigma\beta\alpha'\Sigma\delta - \gamma'\Sigma\alpha\beta'\Sigma\delta \\
&= \gamma'\Sigma\delta\beta'\Sigma\alpha + \gamma'\Sigma\beta\alpha'\Sigma\delta.
\end{aligned}
$$

Combining the results in (17) and (18) we obtain:

$$
\begin{aligned}
& cov_t\left(\gamma' Y_{t+1}\alpha, \delta' Y_{t+1}\beta\right) \\
= \; & \gamma' M Y_t M'\delta\alpha'\Sigma\beta + \gamma' M Y_t M'\beta\alpha'\Sigma\delta + \alpha' M Y_t M'\delta\gamma'\Sigma\beta \\
& +\alpha' M Y_t M'\beta\gamma'\Sigma\delta + \left[\gamma'\Sigma\beta\alpha'\Sigma\delta + \alpha'\Sigma\beta\gamma'\Sigma\delta\right].
\end{aligned}
$$

A similar proof can be constructed for an arbitrary positive integer K.

Appendix 3: Proof of Corollary 2

Let α, β, γ, δ be n-dimensional real vectors. i) Setting $\delta = \gamma$ and $\beta = \alpha$ in Proposition 2, we get:

$$
\begin{aligned}
& V_t\left(\gamma' Y_{t+1}\alpha\right) \\
= \; & cov_t\left(\gamma' Y_{t+1}\alpha, \gamma' Y_{t+1}\alpha\right) \\
= \; & \gamma' M Y_t M'\gamma\alpha'\Sigma\alpha + 2\gamma' M Y_t M'\alpha\alpha'\Sigma\gamma + \alpha' M Y_t M'\alpha\gamma'\Sigma\gamma.
\end{aligned}
$$

ii) The result above for $\gamma = \alpha$ implies:

$$
V_t\left(\alpha' Y_{t+1}\alpha\right) = 4\alpha' M Y_t M'\alpha\alpha'\Sigma\alpha + 2K\left(\alpha'\Sigma\alpha\right)^2.
$$

iii) Using again Proposition 2 with $\gamma = \alpha$ and $\delta = \beta$, we obtain:

$$
cov_t\left(\alpha' Y_{t+1}\alpha, \beta' Y_{t+1}\beta\right) = 4\alpha' M Y_t M'\beta\alpha'\Sigma\beta + 2K\left(\alpha'\Sigma\beta\right)^2.
$$

iv) Finally, Proposition 2 with $\gamma = \alpha$ and $\delta = \alpha$ implies:

$$
\begin{aligned}
cov_t\left(\alpha' Y_{t+1}\alpha, \alpha' Y_{t+1}\beta\right) &= 2\alpha' M Y_t M'\alpha\alpha'\Sigma\beta + 2\alpha' M Y_t M'\beta\alpha'\Sigma\alpha \\
&\quad +2K\alpha'\Sigma\beta\alpha'\Sigma\alpha.
\end{aligned}
$$

Appendix 4: Continuous-time analogue

We have:

$$
\begin{aligned}
dY_t &= Y_{t+dt} - Y_t \\
&= x_{t+dt}x_{t+dt}' - x_t x_t' \\
&= \left(x_t + Ax_t dt + \Omega dw_t\right)\left(x_t + Ax_t dt + \Omega dw_t\right)' - x_t x_t' \\
&= x_t x_t' A' dt + x_t dw_t'\Omega' + Ax_t x_t' dt + Ax_t x_t' A'\left(dt\right)^2 \\
&\quad +Ax_t\left(dw_t\right)'\Omega' dt + \Omega dw_t x_t' + \Omega dw_t x_t' A' dt + \Omega dw_t\left(dw_t\right)'\Omega'.
\end{aligned}
$$

The terms that cannot be neglected in the expression above are:

$$
\begin{aligned}
& dY_t \# x_t x_t' A' dt + x_t dw_t' \Omega' + A x_t x_t' dt + \Omega dw_t x_t' + \Omega E\left[dw_t (dw_t)'\right] \Omega' \\
& \# \left(Y_t A' + A Y_t + \Omega\Omega'\right) dt + x_t (\Omega dw_t)' + (\Omega dw_t) x_t'.
\end{aligned}
$$

Appendix 5 : Proof of Proposition 6

Let P denote an orthogonal matrix such that $P\Sigma^{1/2}\alpha = e_1\sqrt{\alpha'\Sigma\alpha}$, where e_1 denotes the canonical vector with zero components except the first component which is equal to 1. The conditional Laplace transform of $\alpha' Y_t \alpha$ is:

$$
\begin{aligned}
& \Psi_t (u\alpha\alpha') \\
= \; & \frac{\exp\left[\left(u\beta'\alpha\alpha' \left(Id - 2u\Sigma\alpha\alpha'\right)^{-1}\beta\right)\alpha' Y_t \alpha\right]}{\left[\det\left(Id - 2u\Sigma\alpha\alpha'\right)\right]^{K/2}} \\
= \; & \frac{\exp\left[\left(u\beta'\alpha\alpha'\Sigma^{1/2}\left(Id - 2u\Sigma^{1/2}\alpha\alpha'\Sigma^{1/2}\right)^{-1}\Sigma^{-1/2}\beta\right)\alpha' Y_t \alpha\right]}{\left[\det(\Sigma^{1/2}\det\left(Id - 2u\Sigma^{1/2}\alpha\alpha'\Sigma^{1/2}\right)\det(\Sigma^{-1/2})\right]^{K/2}} \\
= \; & \frac{\exp\left[\left(u\beta'\alpha\alpha'\Sigma^{1/2}\left(P^{-1}P - 2uP^{-1}\left(P\Sigma^{1/2}\alpha\right)\left(\alpha'\Sigma^{1/2}P^{-1}\right)P\right)^{-1}\Sigma^{-1/2}\beta\right)\alpha' Y_t \alpha\right]}{\left[\det\left(P^{-1}P - 2uP^{-1}\left(P\Sigma^{1/2}\alpha\right)\left(\alpha'\Sigma^{1/2}P^{-1}\right)P\right)\right]^{K/2}} \\
= \; & \frac{\exp\left[\left(u\beta'\alpha\alpha'\Sigma^{1/2}P'\left(Id - 2u\alpha'\Sigma\alpha e_1 e_1'\right)^{-1}P\Sigma^{-1/2}\beta\right)\alpha' Y_t \alpha\right]}{\left[\det\left(P^{-1}\left(Id - 2u\alpha'\Sigma\alpha e_1 e_1'\right)P\right)\right]^{K/2}} \\
= \; & \frac{\exp\left[\left(u\beta'\alpha\sqrt{\alpha'\Sigma\alpha}e_1'\left(Id - 2u\alpha'\Sigma\alpha e_1 e_1'\right)^{-1}P\Sigma^{-1/2}\beta\right)\alpha' Y_t \alpha\right]}{\left[\det\left(Id - 2u\alpha'\Sigma\alpha e_1 e_1'\right)\right]^{K/2}} \\
= \; & \frac{\exp\left[\left(u\beta'\alpha\left(1 - 2u\alpha'\Sigma\alpha\right)^{-1}\alpha'\Sigma^{-1/2}P'P\beta\right)\alpha' Y_t \alpha\right]}{\left(1 - 2u\alpha'\Sigma\alpha\right)^{K/2}} \\
= \; & \frac{\exp\left[\left(u\beta'\alpha\sqrt{\alpha'\Sigma\alpha}e_1'(1 - 2u\alpha'\Sigma\alpha)^{-1}P\Sigma^{-1/2}\beta\right)\alpha' Y_t \alpha\right]}{(1 - 2u\alpha'\Sigma\alpha)^{K/2}} \\
= \; & \left(1 - 2u\alpha'\Sigma\alpha\right)^{-K/2}\exp\left[\left(\frac{u\left(\alpha'\beta\right)^2}{1 - 2u\alpha'\Sigma\alpha}\right)\alpha' Y_t \alpha\right],
\end{aligned}
$$

which is the conditional Laplace transform of a WAR(1) process of dimension 1 (see Section 3.3) with $m = \alpha'\beta$ and $\sigma^2 = \alpha'\Sigma\alpha$.

Appendix 6 : Proof of Proposition 11

We have just to check the second part of the Proposition. From Proposition

2, we deduce that:

$$E\left[Y_{ij,t+1}|Y_t\right] = \sum_{k=1}^{n}\sum_{l=1}^{n} Y_{kl,t} m_{ik} m_{lj} + K\sigma_{ij}.$$

Since $K > (n-1)$, the admissible values of $Y_{kl,t}$ are not functionally dependent. Thus the product $m_{ik}m_{lj}$, $\forall i$, k, l, j, and the quantities $K\sigma_{ij}$, $\forall i$, j, are first-order identifiable. The result follows.

References

[1] Abel, A. (1988):"Stock Prices under Time Varying Dividend Risk: An Exact Solution in an Infinite-horizon General Equilibrium Model", Journal of Monetary Economics, 22, 375-393.

[2] Ahn, D., Dittmar, R., and A., Gallant (2002): "Quadratic Term Structure Models: Theory and Evidence," Review of Financial Studies, 15, 243-288.

[3] Alexander, C. (2000): "A Primer on the Orthogonal GARCH Model", University of Reading DP.

[4] Andersen, T. (1994): "Stochastic Autoregressive Volatility: A Framework for Volatility Modeling", Mathematical Finance, 4, 75-102.

[5] Andersen, T., and T., Bollerslev (19987) : "Answering the Skeptics : Yes, Standard Volatility Models do Provide Accurate Forecasts", International Economic Review, 39, 885-905.

[6] Andersen, T., Bollerslev, T., and F., Diebold (2002): "Parametric and Non-parametric Volatility Measurement", in Handbook of Financial Econometrics, Hansen, L., and Y., Aït-Sahalia (editors), Amsterdam: North Holland, forthcoming.

[7] Andersen, T., Bollerslev, T., Diebold, F., and P., Labys (2001) : "The Distribution of Exchange Rate Volatility", Journal of the American Statistical Association, 96, 42-55.

[8] Anderson, T. (1984): "An Introduction to Multivariate Statistical Analysis", 2^{nd} edition, Wiley, New-York.

[9] Ang, A., and J. Chen (2002): "Asymmetric Correlations of Equity Portfolios", Journal of Financial Economics, 63, 443-494.

[10] Baba, Y., Engle, R., Kraft, D., and K., Kroner (1987): "Multivariate Simultaneous Generalized ARCH Models", DP, University of California, San Diego.

[11] Backus, D., and A., Gregory (1993) : "Theoretical Relations Between Risk Premiums and Conditional Variance", Journal of Business and Economic Statistics, 11, 177-185.

[12] Baillie, R., and T., Bollerslev (1990) : "A Multivariate Generalized ARCH Approach to Modelling Risk Premia in Forward Foreign Exchange Rate Markets", Journal of International Money and Finance, 9, 309-324.

[13] Barndorff-Nielsen, O., and N., Shephard (2003) : "Impact of Jumps on Returns and Realized Variances : Econometric Analysis of Time-Deformed Levy Processes", DP Nuffield College, Oxford.

[14] Bilodeau, M., and D., Brenner (1999): "Theory of Multivariate Statistics", Springer-Verlag.

[15] Bollerslev, T. (1987): "A Multivariate GARCH Model with Constant Conditional Correlations for a Set of Exchange Rates", Northwestern University, DP.

[16] Bollerslev, T., (1990): "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model", Review of Economics and Statistics, 72, 498-505.

[17] Bollerslev, T., Engle, R., and J., Wooldridge (1988): "A Capital Asset Pricing Model with Time-Varying Covariances", Journal of Political Economy, 96, 116-131.

[18] Brandt, M., and F., Diebold (2002): "A No Arbitrage Approach to Range Based Estimation of Return Covariances and Correlations", University of Pennsylvania, DP.

[19] Bru, M. (1989) :"Diffusions of Perturbed Principal Component Analysis", Journal of Multivariate Analysis, 29, 127-136.

[20] Bru, M. (1991) : "Wishart Processes", J. Theo. Probab., 4, 725-751.

[21] Cheng, P., and O., Scaillet (2002): "Linear Quadratic Jump Diffusion Modelling with Application to Stochastic Volatility", HEC Geneva DP.

[22] Chib, S. (2001): "Markov Chain Monte-Carlo Methods: Computation and Inference", in J.J. Heckman and E. Leamer (eds), Handbook of Econometrics, Vol 5, 3569-3649, North Holland.

[23] Chib, S., Nardari, F., and N. Shephard (2002): "Analysis of High Dimensional Multivariate Stochastic Volatility Models", DP Oxford University.

[24] Cox, J., Ingersoll, J., and S., Ross(1985): "A Theory of the Term Structure of Interest Rates," Econometrica, 53, 385-407.

[25] Darolles, S., Gourieroux, C., and J., Jasiak (2005): "Structural Laplace Transform and Compound Autoregressive Models", Journal of Time Series Analysis, forthcoming.

[26] Diebold, F., and M., Nerlove (1989): "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model", Journal of Applied Econometrics, 4, 1-22.

[27] Ding, Z., and R., Engle (2001): "Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing", DP 01-029, Stern Business School, New York University.

[28] Donati-Martin, C., Doumerc, Y., Matsumoto, H., and M., Yor (2003) :"Some Properties of the Wishart Processes and a Matrix Extension of the Harman-Watson Laws", preprint Univ. Paris VI.

[29] Drost, F., and T., Nijman (1993): "Temporal Aggregation of GARCH Processes", Econometrica, 61, 909-928.

[30] Drost, F., and B., Werker (1996): "Closing the GARCH Gap: Continuous-Time GARCH Modeling", Journal of Econometrics, 74, 31 - 57.

[31] Engle, R. (2002)a: "Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models", Journal of Business and Economic Statistics, 20, 339-350.

[32] Engle, R. (2002)b: "New Frontiers for ARCH Models", Journal of Applied Econometrics, 17, 425-446.

[33] Engle, R., and K., Kroner (1995): "Multivariate Simultaneous Generalized ARCH", Econometric Theory, 11, 122-150.

[34] Engle, R., Lilien, D., and R., Robbins (1987): "Estimating Time-Varying Risk Premia in the Term Structure: The ARCH-M Model", Econometrica, 55, 391-407.

[35] Engle, R., Ng, V., and M., Rothschild (1990): "Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills", Journal of Econometrics, 45, 213-237.

[36] Engle, R., and K., Sheppard (2001): "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH", NBER working paper 8554.

[37] Feller, W. (1971): "An Introduction to Probability Theory and Its Applications", 2nd ed., Vol 2, Wiley.

[38] Fiorentini, G., Sentana, E., and N., Shephard (2003): "Likelihood Based Estimation of Latent Generalized ARCH Structures", DP Oxford University.

[39] Gantmacher, F. (1974): "Theory of Matrices", Chelsea, New York.

[40] Geman, H., and M., Yor (1993) : "Bessel Processes, Asian Options and Perpetuities", Math. Finance, 3., 349-375.

[41] Gennotte, G., and T., Marsh (1993) : "Variations in Economic Uncertainty and Risk Premiums on Capital Assets", European Economic Review, 37, 1021-1041.

[42] Ghysels, E., Harvey, A., and E., Renault (1996): "Stochastic Volatility", in Handbook of Statistics, Vol. 14, Statistical Methods in Finance, Maddala G., and C., Rao (editors), North Holland.

[43] Gourieroux, C., and J., Jasiak (2002): "Financial Econometrics", Princeton University Press.

[44] Gourieroux, C., and J., Jasiak (2005): "Autoregressive Gamma Processes", Journal of Forecasting, 24, forthcoming.

[45] Gourieroux, C., Monfort, A., and E., Renault (1995) : "Inference in Factor Models", in Advances in Econometrics and Quantitative Economics, 311-353, G. Maddala, P.C.B. Phillips and T., Srinivasan, editors, Blackwell, Oxford.

[46] Gourieroux, C., Monfort, A. and R., Sufana (2004b): "International Money and Stock Market Contingent Claims", DP Univ. of Toronto.

[47] Gourieroux, C., and R., Sufana (2003): "Wishart Quadratic Term Structure Models", CREF 03-10, HEC Montreal.

[48] Gourieroux, C., and R., Sufana (2004): "Derivative Pricing with Multivariate Stochastic Volatility", CREST DP.

[49] Grunwald, G., Hyndman, R., Tedesco, L., and R., Tweedie (2000): "Non-Gaussian Conditional Linear AR(1) Models", Australian New Zealand Journal of Statistics, 42, 479-495.

[50] Hamao, Y., Masulis, R., and V., Ng (1990) : "Correlations in Price Changes and Volatility Across International Stock Markets", Review of Financial Studies, 3, 281-307.

[51] Harvey, A., Ruiz, E., and E., Sentana (1994): "Unobserved Component Time Series Models with ARCH Disturbances", Journal of Econometrics, 52, 129-158.

[52] Harvey, A., Ruiz, E., and N., Shephard (1994): "Multivariate Stochastic Variance Models", Review of Economic Studies, 6, 247-264.

[53] Heston, S. (1993): "A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", Review of Financial Studies, 6, 327-343.

[54] James, A. (1968): "Calculation of Zonal Polynomial Coefficients by Use of Laplace-Beltrami Operator", The Annals of Mathematical Statistics, 39, 1711-1718.

[55] Karlin, S., and H., Taylor (1981) :" A Second Course in Stochastic Processes", Academic Press.

[56] Kim, S., Shephard, N., and S., Chib (1998) : "Stochastic Volatility : Likelihood Inference and Comparison with ARCH Models", Review of Economic Studies, 65, 361-393.

[57] King, M., Sentana, E., and S., Wadhwani (1994): "Volatility Links Between National Stock Markets", Econometrica, 62, 901-934.

[58] King, M., and S., Whadwani (1990): "Transmission of Volatility Between Stock Markets", Review of Financial Studies, 3, 5-33.

[59] Ledoit, O., Santa-Clara, P., and M., Wolf (2001) : "Flexible Multivariate GARCH Modelling with an Application to International Stock Markets", DP 578, University Pompeu Fabra.

[60] Lin, W., Engle, R., and T., Ito (1994): "Do Bulls and Bears Move Across Borders ?", International Transmission of Stock Returns and Volatility as the World Turns", Review of Financial Studies, 7, 507-538.

[61] Madan, D., and E., Seneta (1990) : "The VG Model for Share Market Returns", Journal of Business, 63, 511-524.

[62] Meddahi, N., and E., Renault (2004): "Temporal Aggregation of Volatility Models", Journal of Econometrics, 119, 355-379.

[63] Muirhead, R. (1978): "Latent Roots and Matrix Variates: A Review of Some Asymptotic Results", Ann. Stat., 6,5.

[64] Muirhead, R. (1982): "Aspects of Multivariate Statistical Theory", Wiley.

[65] Nelson, D. (1990): "ARCH Models as Diffusion Approximations", Journal of Econometrics, 45, 7-38.

[66] Ng, Y., Engle R., and M. Rotschild (1992): "A Multidynamic Factor Model for Stock Returns", Journal of Econometrics, 52, 245-266.

[67] O'Connel, N. (2003) : "Random Matrices, Non Colliding Particle System and Queues", Seminaire de Probabilités, XXXVI, Lect. Notes in Math., 1801, 165-182, Springer Verlag.

[68] Pelletier, D. (2003) : "Regime Switching for Dynamic Correlations", DP Montreal Univ.

[69] Pitt, M., and N., Shephard (1999):"Filtering via Simulation: Auxiliary Particle Filter", JASA, 94, 590-599.

[70] Revuz, D., and M., Yor (1998) : " Continuous Martingales and Brownian Motion", Springer Verlag.

[71] Schwert, G., and P., Seguin (1990) : "Heteroskedasticity in Stock Returns", The Journal of Finance, 44, 1115-1153.

[72] Stuart, A., and K., Ord (1994): "Kendall's Advanced Theory of Statistics, Volume 1, Distribution Theory", Sixth Edition, Edward Arnold.

[73] Taylor, S. (1986): "Modeling Financial Time Series", Wiley.

[74] Tse, Y. (2000): "A Test for Constant Correlations in a Multivariate GARCH Model", Journal of Econometrics, 98, 107-127.

[75] Tse, Y., and A., Tsui (2002): "A Multivariate GARCH Model with Time Varying Correlations", Journal of Business and Economic Statistics, 20, 351-362.

[76] Wishart, J. (1928a,b): "The Generalized Product Moment Distribution in Samples from a Multinomial Population", Biometrika, 20, 32 and 424.

Fig.1 Volatilities, Example 1

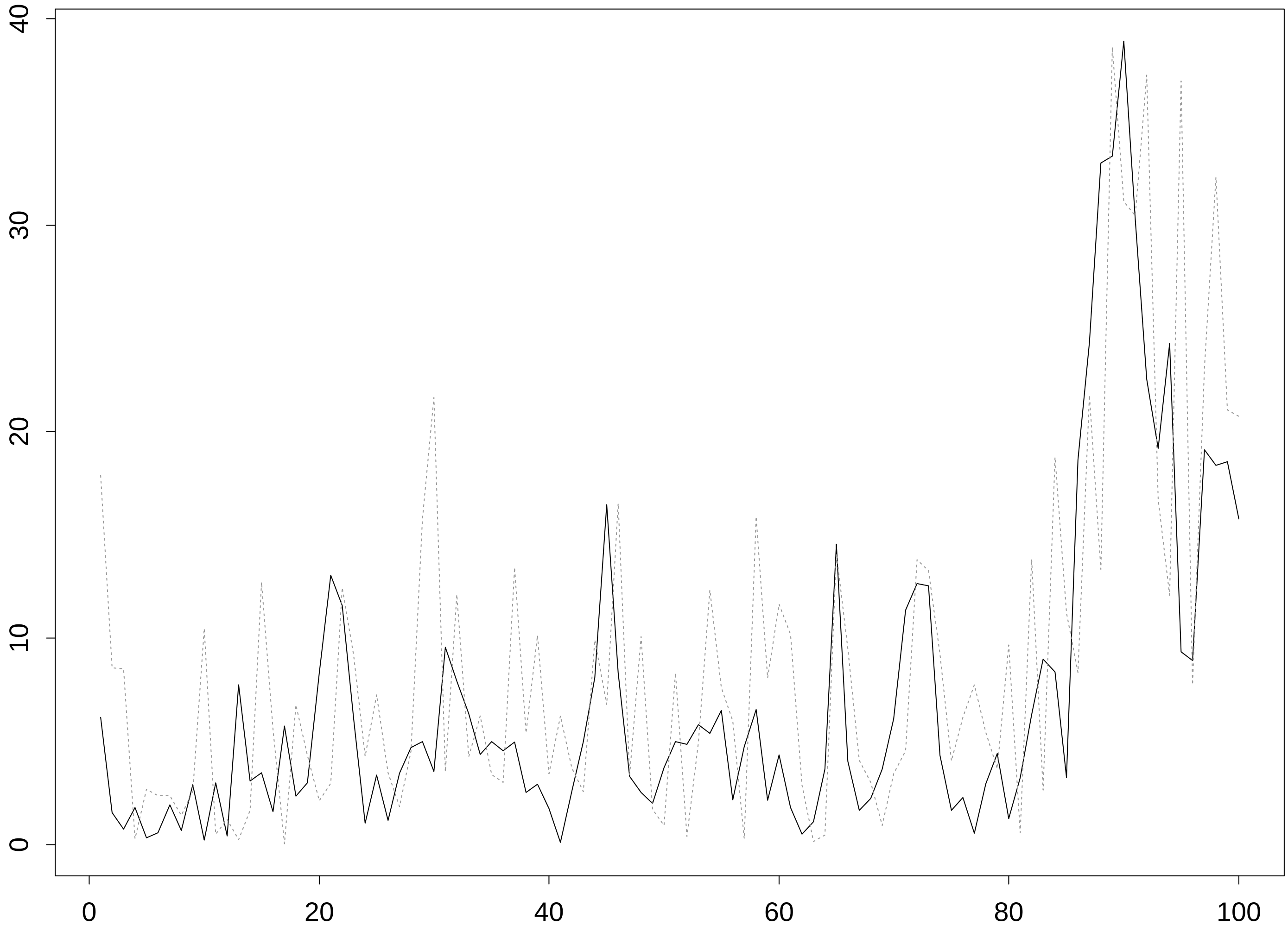

Fig.2 Correlation, Example 1

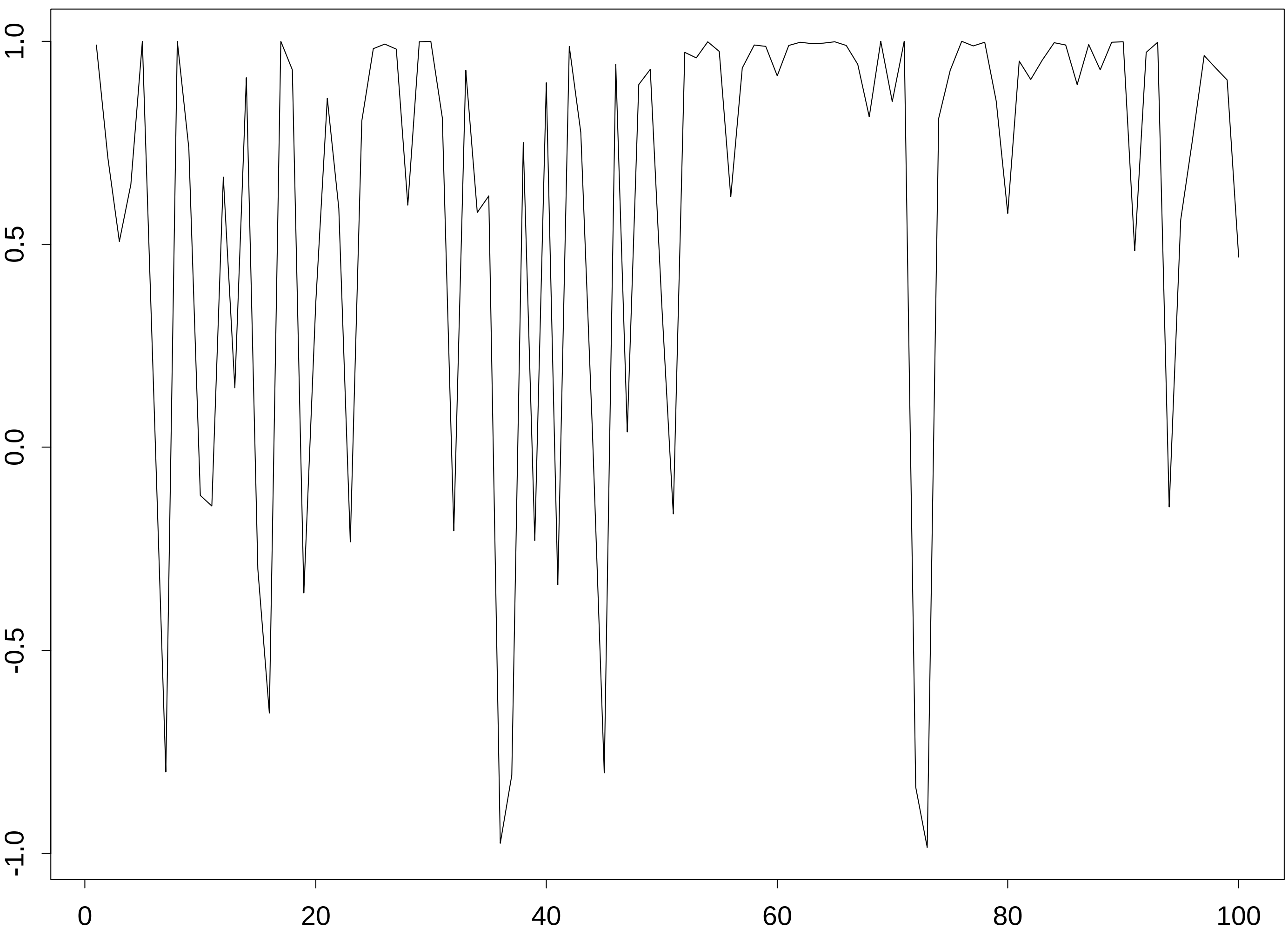

Fig.3 Canonical Volatilities, Example 1

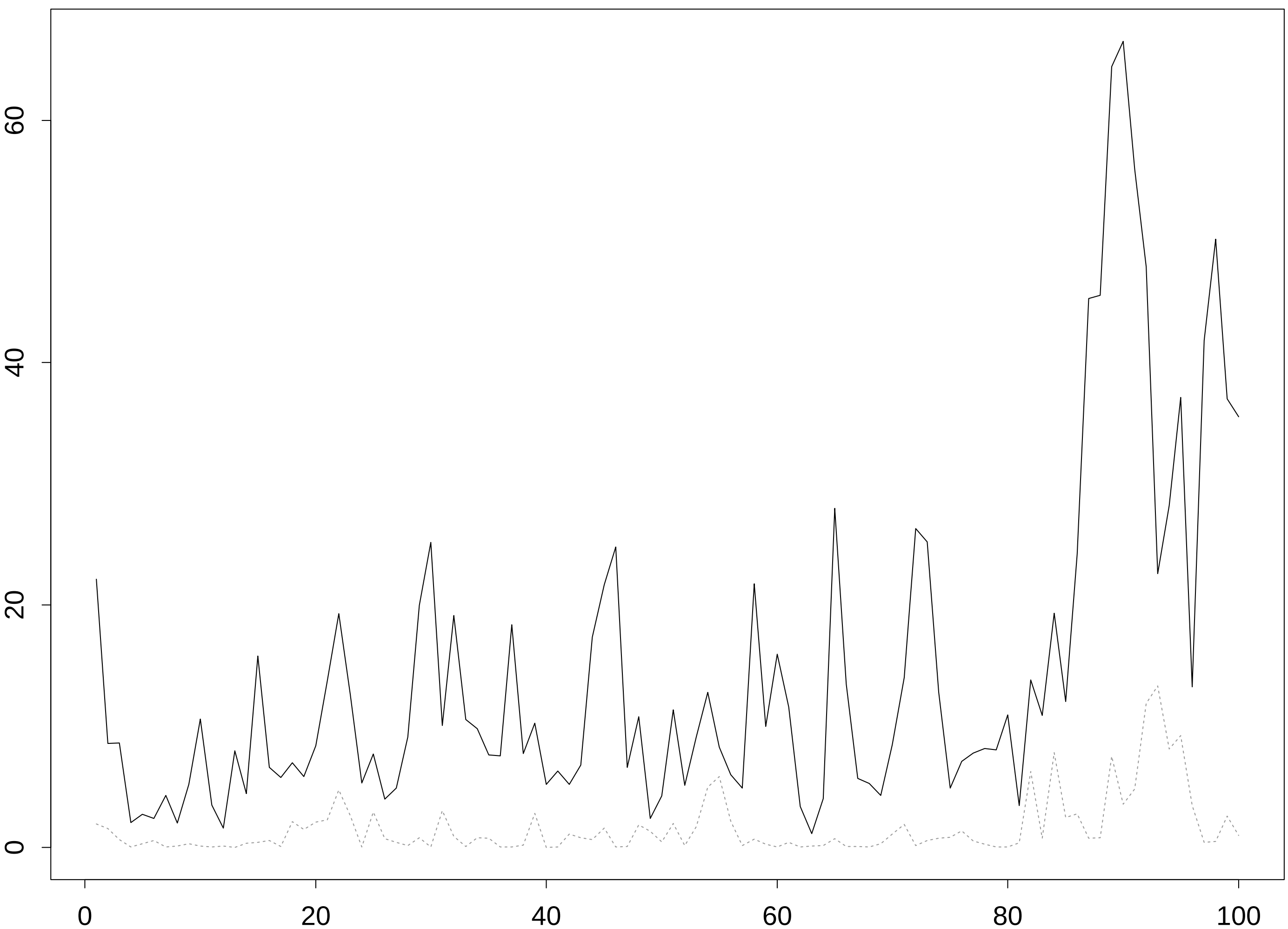

Fig.4 Volatilities, Example 2

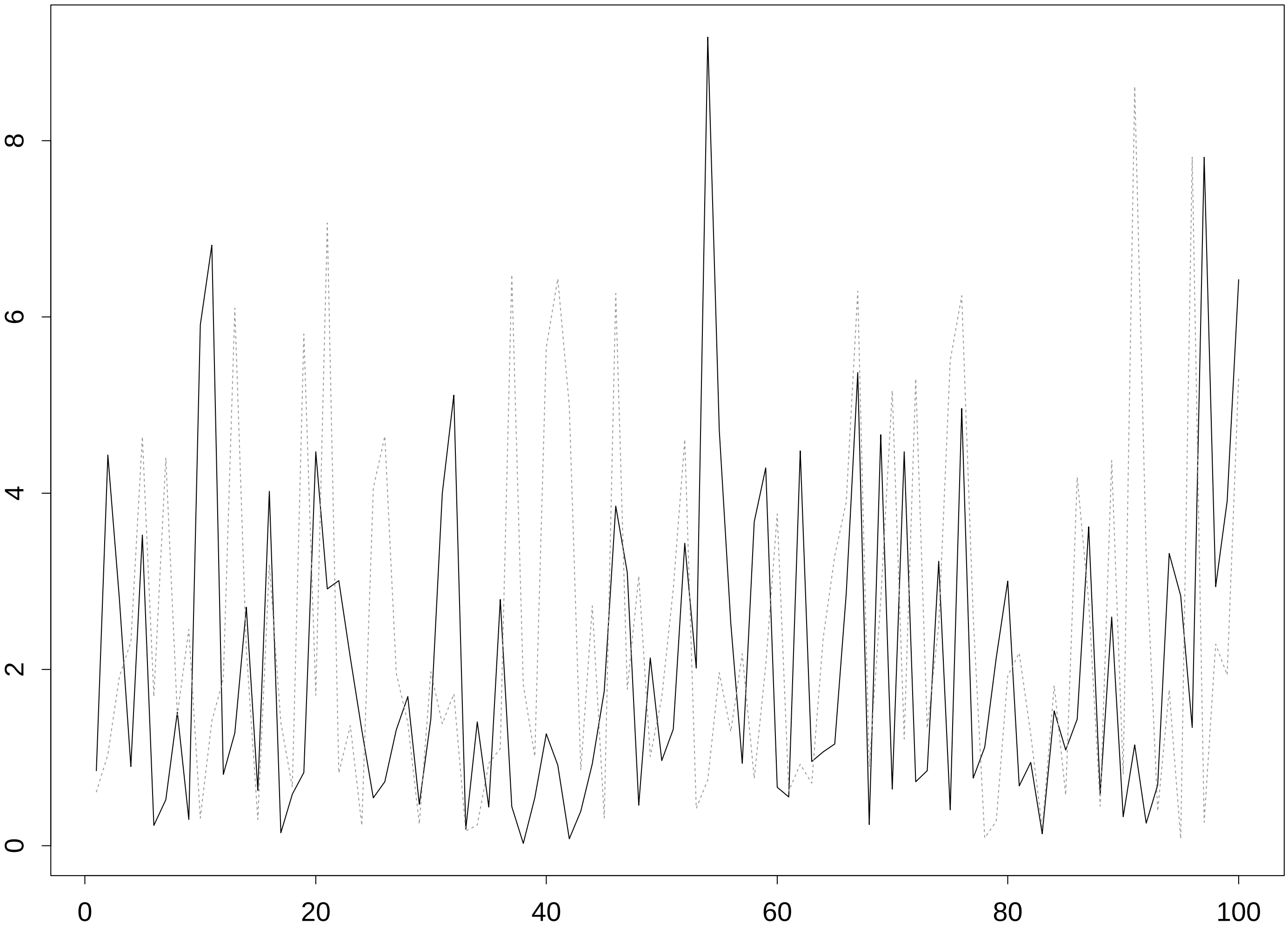

Fig.5 Correlation, Example 2

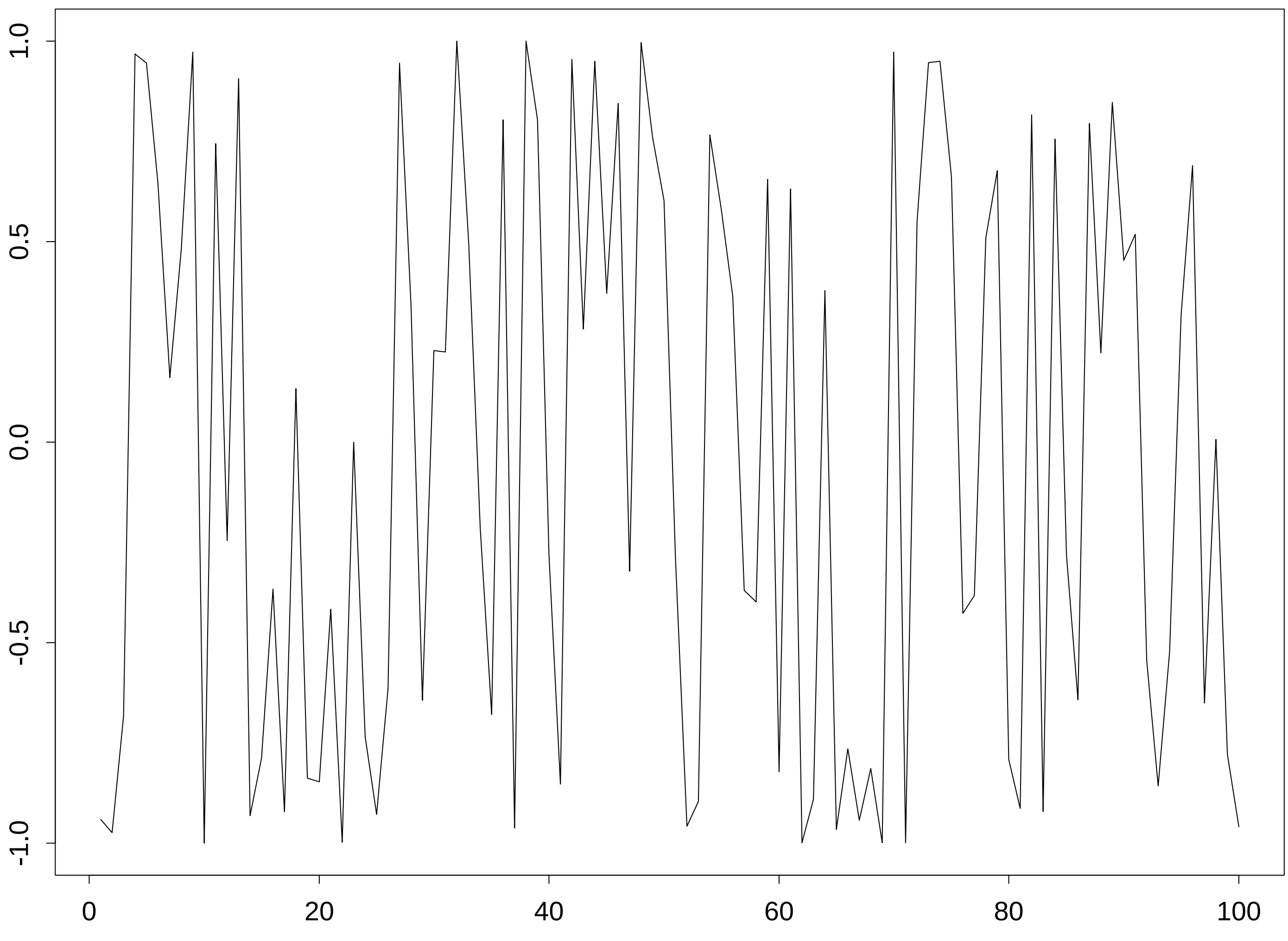

Fig.6 Canonical Volatilities, Example 2

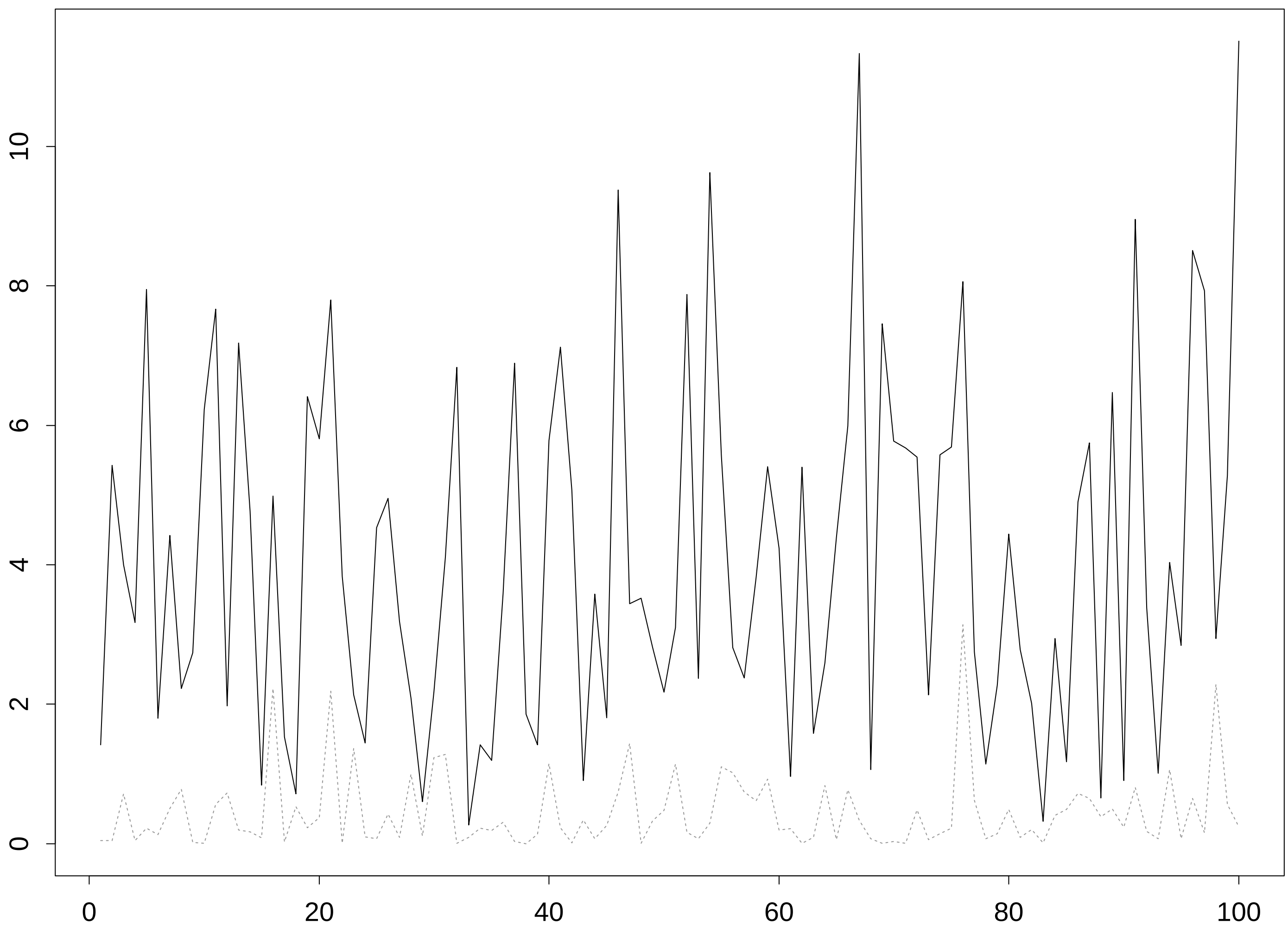

Fig.7 Volatilities, Example 3

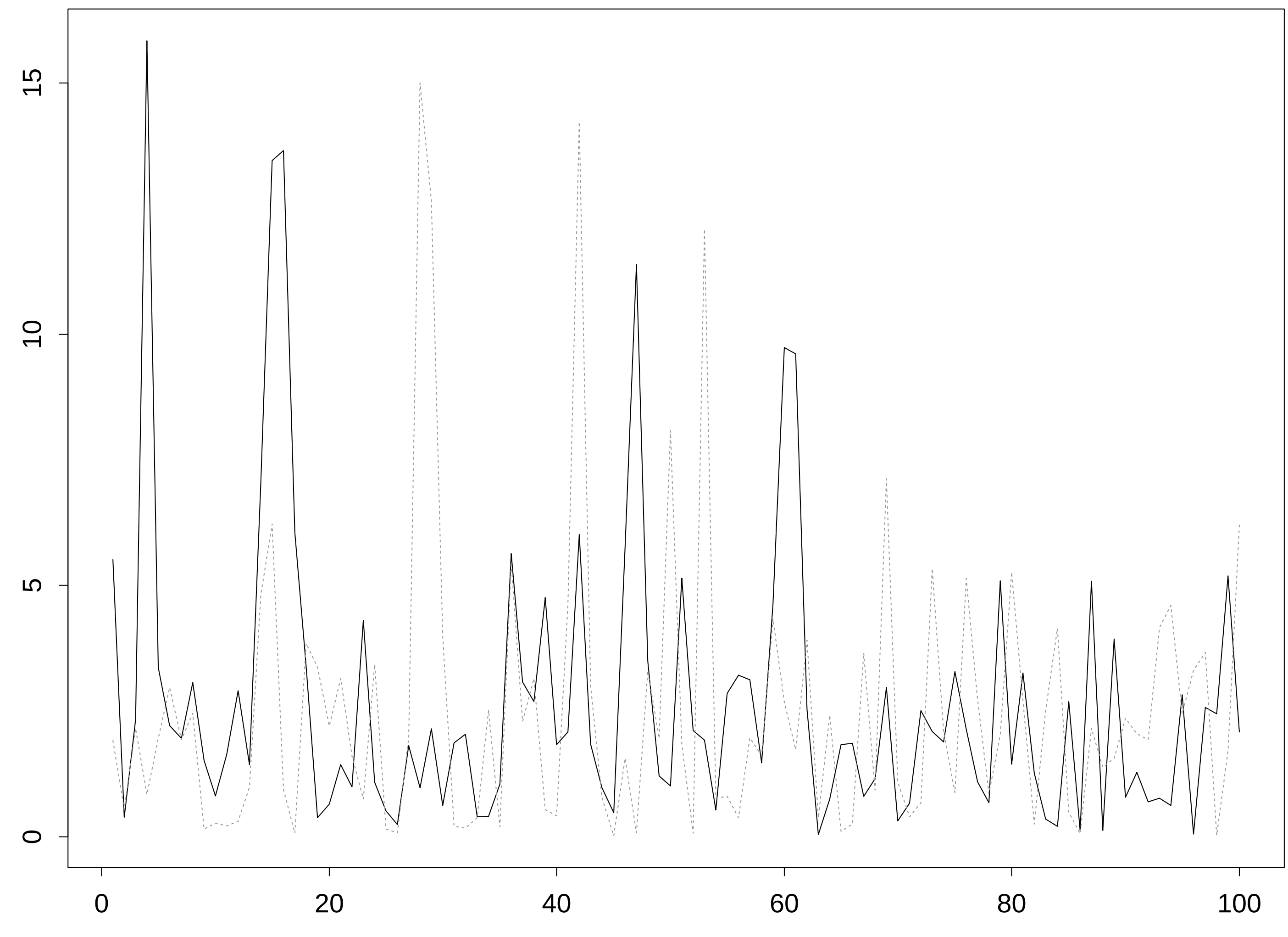

Fig.8 Correlation, Example 3

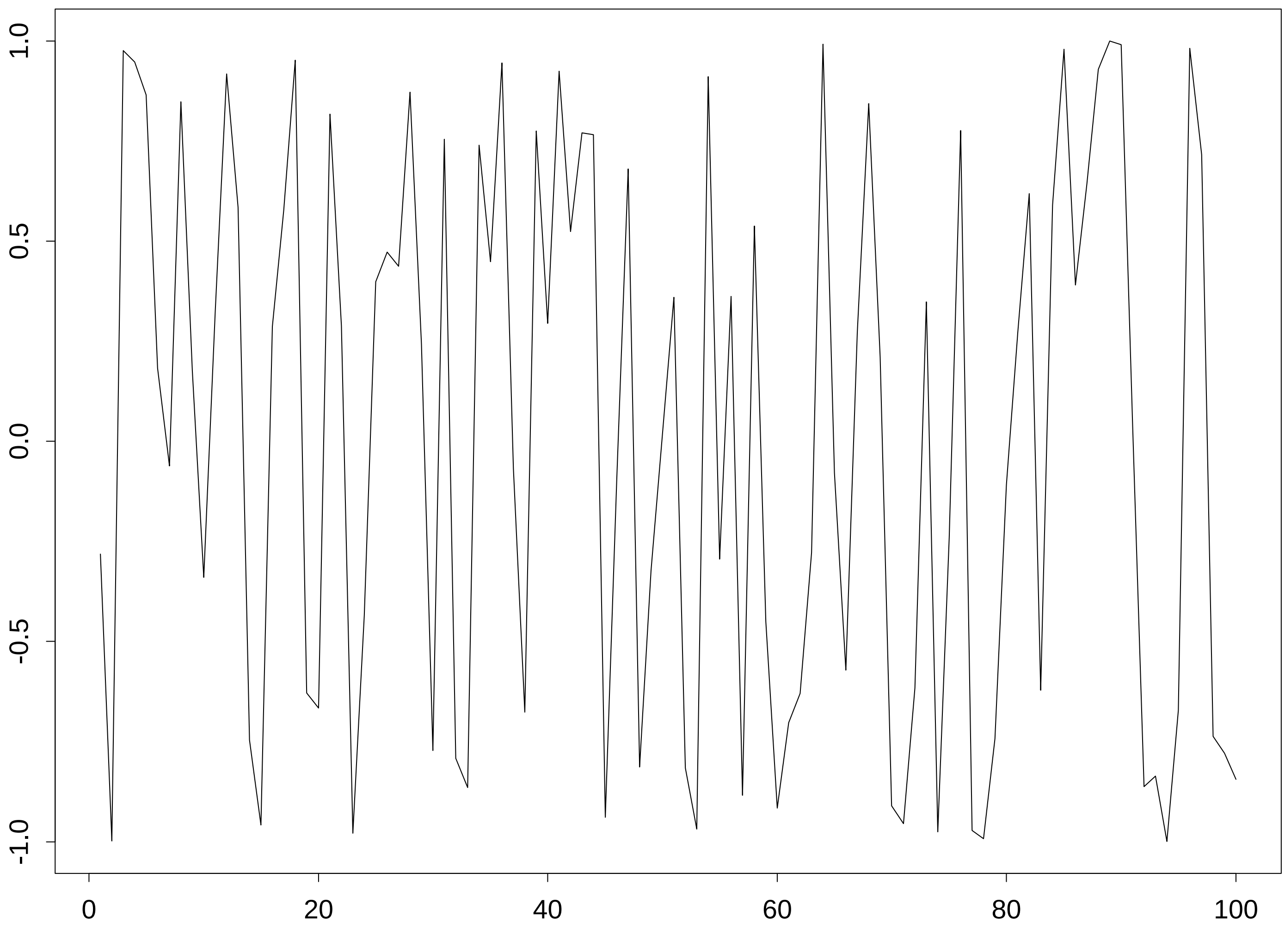

Fig.9 Canonical Volatilities, Example 3

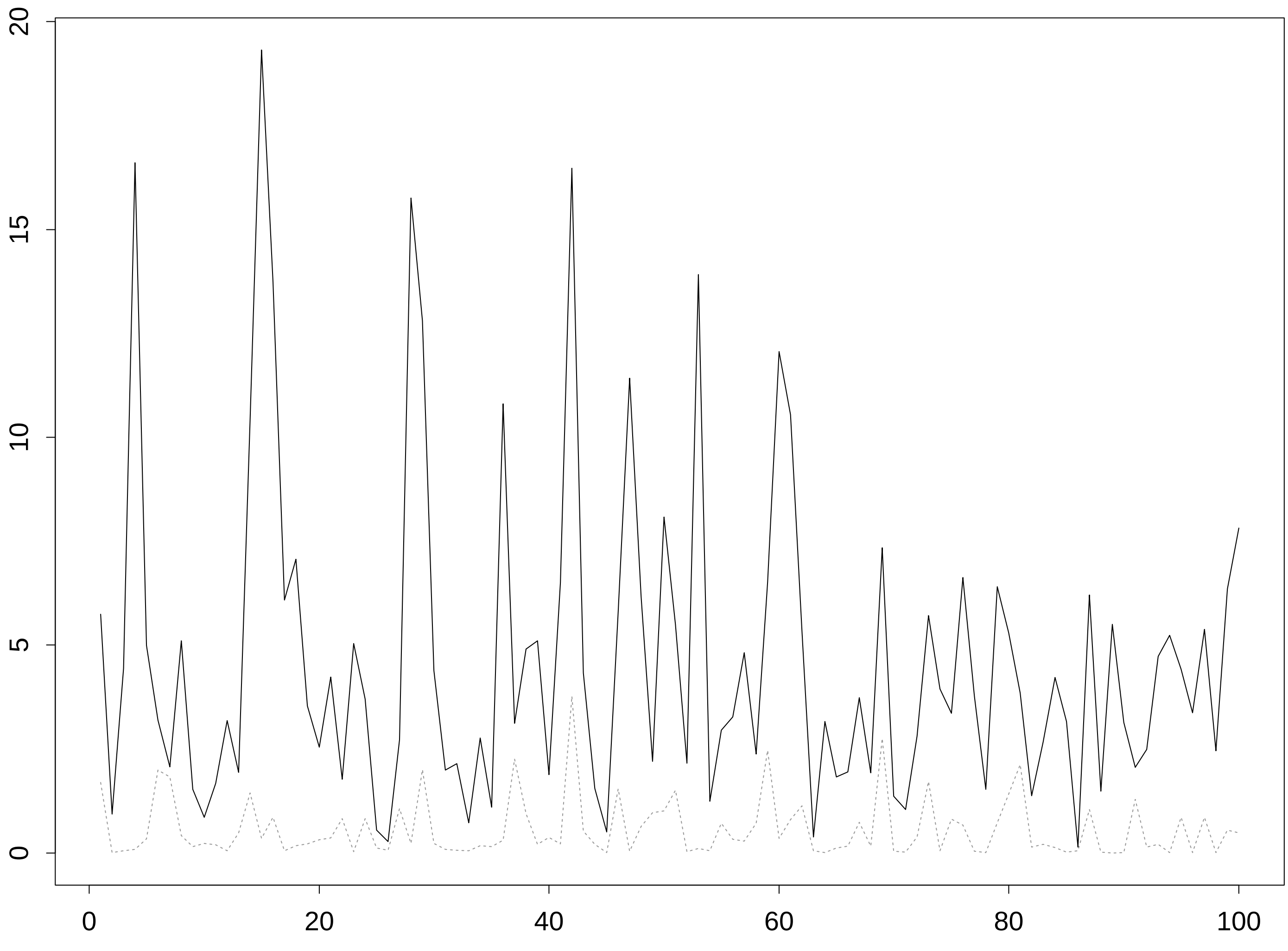

Figure 10 : Volatilities

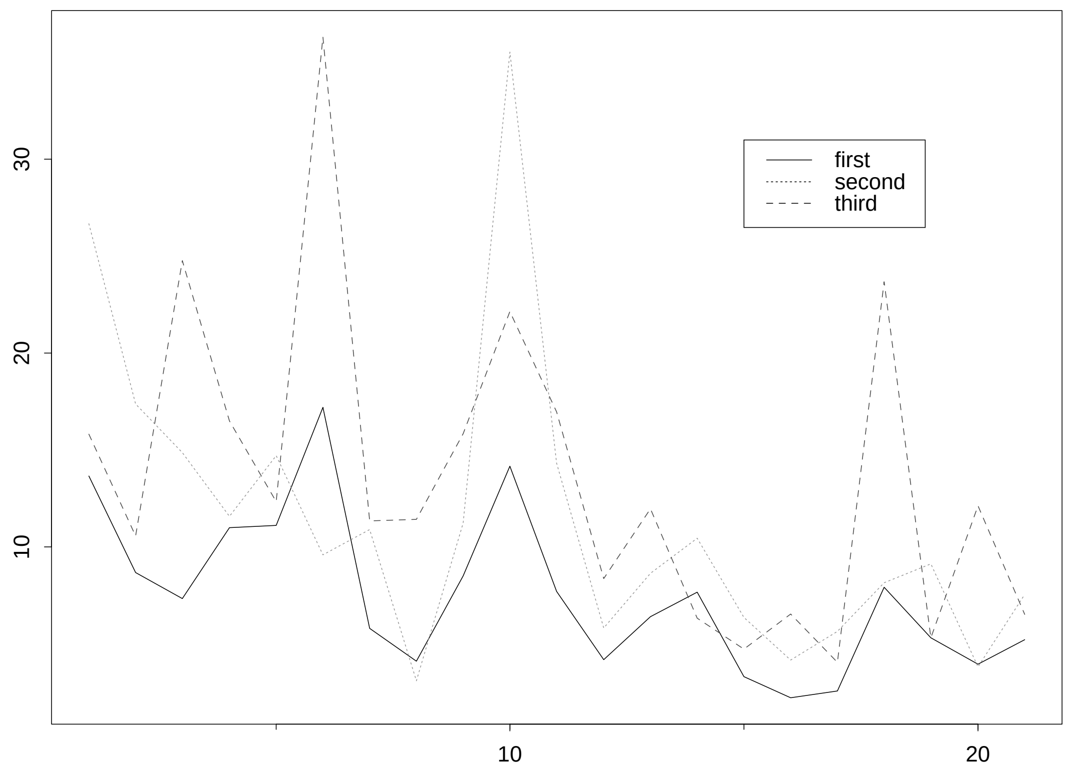

Figure 11 : Correlations

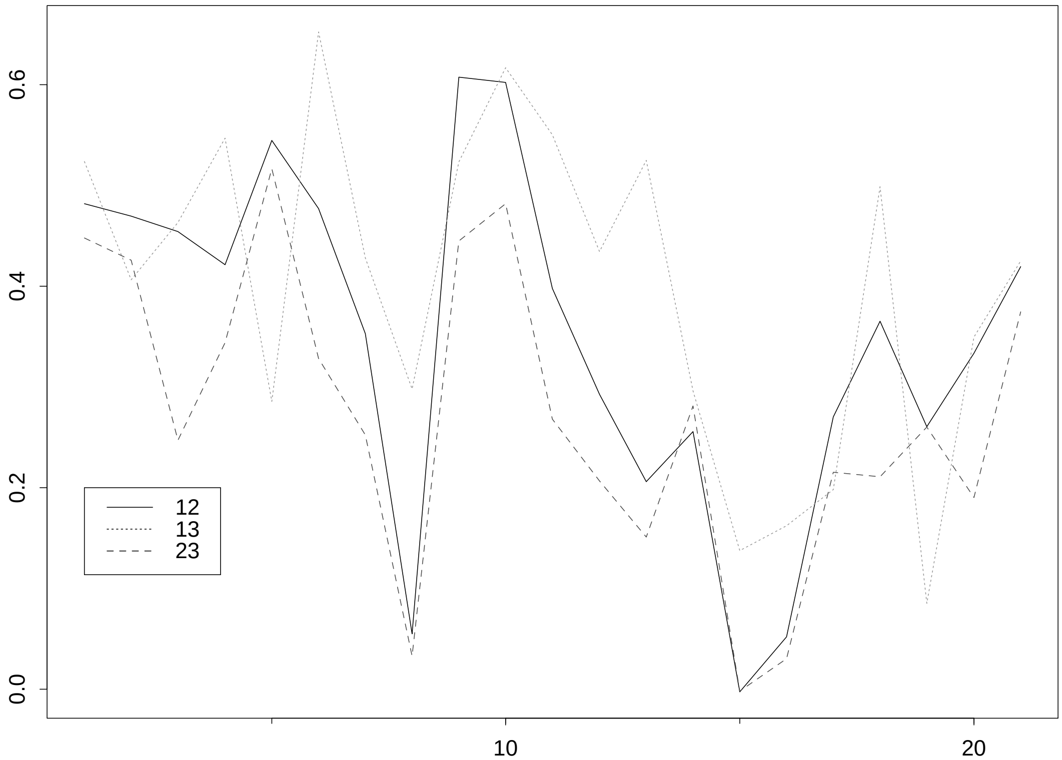

Figure 12 : Eigenvalues