

# A closed-form estimator for the multivariate GARCH(1,1) model

Giacomo Sbrana\* and Federico Poloni†

We provide a closed-form estimator based on the VARMA representation for the unrestricted multivariate GARCH(1,1). We show that all parameters can be derived using basic linear algebra tools. We show that the estimator is consistent and asymptotically normal distributed. Our results allow also to derive a closed form for the parameters in the context of temporal aggregation of multivariate GARCH(1,1) by solving the equations as in Hafner [2008].

**Keywords:** Multivariate GARCH(1,1), VARMA, Temporal Aggregation, Estimation.

## 1. Introduction

Estimating a multivariate GARCH(1,1) model is a challenging task. The most common tool for this purpose is the quasi maximum likelihood (QML) estimator which requires rather sophisticated optimization techniques. In this paper we present a simple and fast method of moments which makes the estimation of the multivariate GARCH(1,1) model more accessible. Our results represent the multivariate generalization of the analytical results already achieved by Kristensen and Linton [2006] for the scalar case.

Our estimator is consistent and, under additional assumptions on the moments, asymptotically normal distributed. Due to the difficulties in estimating multivariate GARCH(1,1) models our estimator may then be used to provide a consistent initial estimate when implementing numerical optimization techniques for the QML estimation. This is especially true when large-scale models are employed.

Several restricted models have been proposed by the previous literature in order to reduce the number of parameters, such as Diagonal VEC (Bollerslev et al. [1988]), BEKK-GARCH (Engle and Kroner [1995]), CCC-GARCH (Bollerslev [1990]). Interestingly,

---

\*Rouen Business School. 1, rue du Maréchal Juin, 76130 Mont-Saint-Aignan, France. E-mail [gsb@rouenbs.fr](mailto:gsb@rouenbs.fr)

†Dipartimento di Informatica, Università di Pisa. Largo Pontecorvo 3, 56127 Pisa, Italy. E-mail [fpoloni@di.unipi.it](mailto:fpoloni@di.unipi.it)

our results are valid in general. Therefore in the framework we stick to the unrestricted multivariate GARCH(1,1).

Finally, our results extend the results of Hafner [2008] in the context of temporal aggregation of multivariate GARCH(1,1). Indeed, our results allow to derive the parameters of the temporally aggregated GARCH for any aggregation frequency. In other words, given the parameters of the disaggregated process, those of the aggregate one are analytical functions of the disaggregate parameters. Alternatively, one can also use the moments of the disaggregated GARCH to produce an initial estimate of the parameters of temporally aggregated processes. The former estimator is again consistent and asymptotically normal when some moments conditions hold.

## 2. Framework

Consider the following unrestricted multivariate GARCH(1,1) model

$$y_t = H_t^{1/2} \epsilon_t, \quad t = 1, 2, \dots, n,$$

where  $y_t$  is a  $d$ -dimensional zero-mean, serially uncorrelated process. In addition, we have that  $\epsilon_t \in \mathbb{R}^{d \times 1}$  is an i.i.d. white noise vector with zero mean and variance  $I_d$ . Moreover, the conditional covariance matrix is given by

$$\text{vech}(H_t) = c + A \text{vech}(y_{t-1} y_{t-1}^T) + B \text{vech}(H_{t-1}), \quad t = 2, 3, \dots, n, \quad (1)$$

garchrec

where  $\text{vech}(M)$  represents the operator that stacks the elements of the lower triangular part of a symmetric matrix  $M$  to form a  $\bar{d} \times 1$  vector, with  $\bar{d} = \frac{d(d+1)}{2}$ .

In what follows we make the following assumptions:

1.  $H_t$  is positive definite almost surely for each  $t$ .
2. All eigenvalues of the matrix  $A + B$  have modulus smaller than one.
3. The process  $y_t$  is ergodic,  $\beta$ -mixing, and strictly stationary.
4. The fourth moments of  $y_t$  exist and are finite.

moments

Boussama [2006] provides sufficient conditions that ensure a strictly stationary, ergodic and  $\beta$ -mixing solution of the vector GARCH process (these can also be found in [Francq and Zakoïan, 2010, Theorem 11.5]).

The following stronger assumption is used only in some central limit results:

- 5 The eighth moments of  $y_t$  exist and are finite.

When the distribution of  $\epsilon_t$  is spherical, [Hafner, 2003, Theorem 3] has given an algebraic condition equivalent to Assumption 4 that is easy to test in practice. However, we do not need to assume sphericity here.

The VARMA(1,1) representation of the multivariate GARCH(1,1) is obtained by defining  $h_t \equiv \text{vech}(H_t)$ ,  $x_t \equiv \text{vech}(y_t y_t^T)$  and  $\xi_t \equiv x_t - h_t$ ; the recurrence relation (1) is equivalent to

$$x_t = c + \Phi x_{t-1} + \xi_t - B \xi_{t-1}.$$

By eliminating  $x_{t-1}$  recursively, we find that asymptotically the following formula holds

$$x_t = h + \sum_{i=0}^{\infty} \Theta_i \xi_{t-i},$$

where  $\Phi = A + B$ ;  $h = \text{vech}(H) = (I - \Phi)^{-1}c$ ;  $\Theta_0 = I_d$  and  $\Theta_i = (A + B)^{i-1}A$  for  $i \geq 1$ . The interest of this formulation lies in the fact that  $\xi_t$  is a martingale difference sequence. We define  $\Sigma \equiv E[\xi_t \xi_t^T]$ ; note that  $E[\xi_t \xi_{t+s}^T] = 0$  for  $s \geq 1$ .

### 3. Closed form Estimation

sec:closedForm

Under Assumptions 1–4, the autocovariances of  $x_t$  exist and are finite, and they are given by

$$M_k = E[(x_{t+k} - h)(x_t - h)^T] = \sum_{i=0}^{\infty} \Theta_{i+k} \Sigma \Theta_i^T.$$

From the VARMA(1,1) representation using the standard Yule-Walker results we have that

$$M_{k+1} = \Phi M_k, \quad \text{for all } k \geq 1,$$

thus  $\Phi$  can be obtained analytically as

$$\Phi = M_{k+1} M_k^{-1}, \quad \text{for all } k \geq 1. \quad (2) \quad \boxed{\text{Phi}}$$

These results are well known and can also be found for example in the book of Reinsel [1997] as well as in Hafner [2003] page 32 (for the univariate case see Kristensen and Linton [2006]). Therefore, we can estimate  $\hat{\Phi} = \hat{M}_2 \hat{M}_1^{-1}$ . Consider now the first-order moving average vector

$$j_t = x_t - \Phi x_{t-1} = c + \xi_t - B \xi_{t-1}.$$

The autocovariances of  $j_t$  are

$$\begin{aligned} \Gamma_0 &\equiv E[(j_t - c)(j_t - c)^T] = \Sigma + B \Sigma B^T = M_0 - M_1 \Phi^T - \Phi M_1^T + \Phi M_0 \Phi^T, \\ \Gamma_1 &\equiv E[(j_t - c)(j_{t-1} - c)^T] = -B \Sigma = M_1 - M_2 \Phi^T - \Phi M_0 + \Phi M_1 \Phi^T = M_1 - \Phi M_0. \end{aligned} \quad (3) \quad \boxed{\text{Gamma}}$$

We can combine the former two equations with simple manipulations to derive two separate equations for  $B$  and  $\Sigma$

$$\Gamma_1^T + \Gamma_0 B^T + \Gamma_1 (B^T)^2 = 0, \quad (4) \quad \boxed{\text{pme}}$$

$$\Gamma_0 = \Sigma + \Gamma_1 \Sigma^{-1} \Gamma_1^T. \quad (5) \quad \boxed{\text{nme}}$$

In the scalar case, (4) is a quadratic equation; the approach suggested by Kristensen and Linton [2006] consists essentially in deriving an estimator by solving this equation.

This method, however, need not be restricted to the univariate GARCH. In the multivariate case, basic linear algebra techniques can be used to derive a closed form in terms of eigenvalues and eigenvectors. We present them in the next section.

### 3.1. Closed formula for $B$

The following procedure can be used to obtain  $B$  as analytical function of  $\Gamma_0$  and  $\Gamma_1$ .

1. Form the  $2\bar{d} \times 2\bar{d}$  matrix

$$P = \begin{bmatrix} 0 & I \\ -\Gamma_1^{-1}\Gamma_1^T & -\Gamma_1^{-1}\Gamma_0 \end{bmatrix}. \quad (6) \quad \boxed{\text{defP}}$$

2. One can prove (Lemma 5 in the following) that the eigenvalues of  $P$  come in pairs  $(\lambda, 1/\lambda)$ . Therefore, unless there are eigenvalues that lie exactly on the unit circle, half of the  $2\bar{d}$  eigenvalues satisfy  $|\lambda| < 1$ , and we may reorder them so that  $|\lambda_i| < 1$  for  $i = 1, 2, \dots, \bar{d}$ . Moreover, consider the associated eigenvectors  $w_i$  and partition them as

$$w_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix}, \quad u_i, v_i \in \mathbb{R}^{\bar{d}}.$$

3. Now, a solution to the matrix equation (4) is given by

$$B = (U^T)^{-1}DU^T = \begin{bmatrix} u_1 & u_2 & \cdots & u_{\bar{d}} \end{bmatrix}^{-T} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_{\bar{d}} \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_{\bar{d}} \end{bmatrix}^T. \quad (7) \quad \boxed{\text{B}}$$

In Section 5 and Appendix A, we recall the theoretical results in linear algebra that ensure the functioning of this procedure.

### 3.2. Estimation using the closed formula

Using this formula, an estimation procedure can be derived as following:

1. Given the data, compute the observed average and the first three autocovariances of  $x_t$ :

$$\begin{aligned} \hat{h} &= \frac{1}{n} \sum_{t=1}^n x_t, & \hat{M}_0 &= \frac{1}{n} \sum_{t=1}^n [(x_t - \hat{h})(x_t - \hat{h})^T], \\ \hat{M}_1 &= \frac{1}{(n-1)} \sum_{t=1}^{n-1} [(x_{t+1} - \hat{h})(x_t - \hat{h})^T], & \hat{M}_2 &= \frac{1}{(n-2)} \sum_{t=1}^{n-2} [(x_{t+2} - \hat{h})(x_t - \hat{h})^T]. \end{aligned}$$

2. Evaluate  $\hat{\Phi} = \hat{M}_2 \hat{M}_1^{-1}$ ,  $\hat{I}_0 = \hat{M}_0 - \hat{M}_1 \hat{\Phi}^T - \hat{\Phi} \hat{M}_1^T + \hat{\Phi} \hat{M}_0 \hat{\Phi}^T$ , and  $\hat{I}_1 = \hat{M}_1 - \hat{\Phi} \hat{M}_0$ , as provided by (3).
3. Use the above procedure based on eigenvalue computation to get an estimated  $\hat{B}$ .
4. Finally, recover the other two parameters as  $\hat{A} = \hat{\Phi} - \hat{B}$  and  $\hat{c} = (I - \hat{A} - \hat{B})\hat{h}$ .

### 3.3. Asymptotic properties

In this section and the following, the symbols  $\xrightarrow{p}$  and  $\xrightarrow{\text{dist}}$  stand for convergence in probability and distribution, respectively.

The consistency of the QML estimator has been shown by Jeantreau [1998] while Comte and Lieberman [2003] provides the asymptotic normality of the QML estimator in the context of BEKK formulation. However, as noted in Bauwens et al. [2006]: *The asymptotic properties of ML and QML estimators in multivariate GARCH models are not yet firmly established, and are difficult to derive from low level assumptions [...] Asymptotic normality of the QMLE is not established generally. [...] Researchers who use MGARCH models have generally proceeded as if asymptotic normality holds in all cases.* Here we provide the asymptotic properties of our closed-form estimator  $\Lambda = (c, \text{vec}(A), \text{vec}(B))$ , which is function of the moments of  $x_t$  only, that is,  $\Lambda = F(h, M_0, M_1, M_2)$ .

Convergence

**Lemma 1.** *Let*

$$\mathbf{m} = \begin{bmatrix} h \\ \text{vec } M_0 \\ \text{vec } M_1 \\ \text{vec } M_2 \end{bmatrix}.$$

*Under Assumptions 1–4,  $\hat{\mathbf{m}} \xrightarrow{p} \mathbf{m}$ . In addition, if Assumption 5 holds, we have*

$$\sqrt{n}(\hat{\mathbf{m}} - \mathbf{m}) \xrightarrow{\text{dist}} N(0, \Psi), \quad (8)$$

*where  $\Psi = \text{Cov}[\mathbf{m}]$ .*

*Proof.* Under these assumptions, both  $h$  and  $M_k$  are finite, in addition  $\hat{h}$  and  $\hat{M}_k$  are consistent given the law of large numbers for stationary and ergodic processes. Finally, the joint distribution converges to a normal distribution thanks to the standard central limit theory for strongly mixing sequences.  $\square$

Note that  $\mathbf{m}$  is consistent even if Assumption 5 does not hold; in this case, however,  $\hat{\mathbf{m}}$  converges with a slower rate (see the discussion in Kristensen and Linton [2006]). Moreover, the limit joint distribution of the moments is not Gaussian.

If the noise distribution is spherical, then, as noted by Hafner [2003], the cross-covariance between  $h$  and  $\text{vec}(M_k)$  vanishes for each  $k$ , as it is an odd-power function of the noise  $\epsilon_t$ . In this case, the explicit expression for  $\Psi$  is

$$\Psi = \begin{bmatrix} \text{Cov}[h] & 0_{\bar{d} \times 3\bar{d}^2} \\ 0_{3\bar{d}^2 \times \bar{d}} & \text{Cov}[M_k, M_l] \end{bmatrix}$$

where:

$$\text{Cov}[h] = M_0 + \sum_{k=1}^{\infty} M_k + \sum_{k=1}^{\infty} M_k^T$$

and

$$\text{Cov}[M_k, M_l] = E \left[ \left( \text{vec} [(x_{t+k} - h)(x_t - h)^T] - \text{vec } M_k \right) \left( \text{vec} [(x_{t+l} - h)(x_t - h)^T] - \text{vec } M_l \right)^T \right]$$

for  $k, l = 0, 1, 2$ .

The following theorem represents a direct consequence of Lemma 1.

Normality

**Theorem 2.** *Under Assumptions 1-4, the estimator  $\hat{\Lambda} = (\hat{c}, \text{vec}(\hat{A}), \text{vec}(\hat{B})) = F(\mathbf{m})$  is consistent, i.e.  $\hat{\Lambda} \xrightarrow{p} \Lambda$ . In addition, if Assumption 5 holds,  $\sqrt{n}(\hat{\Lambda} - \Lambda) \rightarrow N(0, \Xi)$  with*

$$\Xi = \left( \frac{\partial F(\mathbf{m})}{\partial \mathbf{m}} \right) \Psi \left( \frac{\partial F(\mathbf{m})}{\partial \mathbf{m}} \right)^T. \quad (9)$$

usesPartialDer

*Proof.* The theorem follows from the continuous mapping theorem; in addition, the explicit expression for the covariance is obtained using the delta method. It remains to show that the partial derivative  $\frac{\partial F(\mathbf{m})}{\partial \mathbf{m}}$  exists, which is proven in the end of Section 5.  $\square$

Note that the asymptotic properties of the closed form estimator might be employed to prove these of the QML in general (as discussed by Bauwens et al. [2006]). However, we leave this for future research.

## 4. Temporal aggregation

An interesting consequence of the results above is the direct extension to the temporal aggregation. In fact, we can now derive (as well as estimate) the parameters of the temporally aggregated multivariate GARCH(1,1) as discussed in Hafner [2008].

Temporal aggregation of a GARCH can be conducted in two different forms, depending on whether we are interested in *stock* or *flow* variables. We are interested in deriving a GARCH representation for the process  $y^{(m)}$  aggregated over  $m$  periods, which is defined in the two cases as

$$\begin{aligned} y_{mt}^{(m)} &= y_{mt}, && \text{(stock variables)} \\ y_{mt}^{(m)} &= y_{mt} + y_{mt-1} \dots + y_{mt-m+1} + w_{mt}^{(m)}. && \text{(flow variables)} \end{aligned}$$

We denote  $x_{mt}^{(m)} = \text{vech}(y_{mt}^{(m)} y_{mt}^{(m)T})$ . While in the stock case  $x_{mt}^{(m)} = x_{mt}$ , the relation between the second moments is more involved for flow variables.

Hafner [2008] shows that the temporally aggregated process follows a weak VARMA(1,1)

$$x_{mt}^{(m)} = c^{(m)} + \Phi^{(m)} x_{m(t-1)}^{(m)} + \xi_{mt}^{(m)} - B^{(m)} \xi_{m(t-1)}^{(m)}. \quad (10)$$

Then, equations for  $B^{(m)}$  formally analogous to (3) are derived:

$$\begin{aligned}\Gamma_0^{(m)} &= \Sigma^{(m)} + B^{(m)} \Sigma^{(m)} (B^{(m)})^T, \\ \Gamma_1^{(m)} &= -B^{(m)} \Sigma^{(m)},\end{aligned}$$

with

$$\begin{aligned}\Gamma_0^{(m)} &= \sum_{i=0}^m J_i^s \Sigma J_i^{sT}, \\ \Gamma_v^{(m)} &= J_m^s \Sigma, \\ J_0^s &= I_{\bar{d}}, \\ J_i^s &= \Phi^{i-1} A, \\ J_m^s &= -\Phi^{m-1} B,\end{aligned}$$

for the case of stock variables, and

$$\begin{aligned}\Gamma_0^{(m)} &= \sum_{i=0}^{2m-1} J_i^f \Sigma J_i^{fT} + \Sigma_w^m + \Phi^m \Sigma_w^m (\Phi^T)^m, \\ \Gamma_1^{(m)} &= \sum_{i=0}^{m-1} J_{i+m}^f \Sigma J_i^{fT} - \Phi^m \Sigma_w^m, \\ J_0^f &= I_{\bar{d}}, \\ J_i^f &= I_{\bar{d}} + A + \Phi A + \dots + \Phi^{i-1} A, \quad i = 1, \dots, m-1 \\ J_i^f &= [I_{\bar{d}} + \Phi + \dots + \Phi^{m-2}] A - \Phi^{m-1} B, \quad i = m \\ J_i^f &= [\Phi^{i-m} + \Phi^{i-m+1} + \dots + \Phi^{m-2}] A - \Phi^{m-1} B, \quad i = m+1, \dots, 2m-2, \\ J_{2m-1}^f &= -\Phi^{m-1} B,\end{aligned}$$

for the case of flow variables (where an explicit expression for  $\Sigma_w^m$  can be found in Hafner [2008] eq.19). The author, however, rearranges them to eliminate  $\Sigma^{(m)}$  in a form that differs slightly from our (4), and for which deriving an explicit solution is more complicated. In his words, “*As of  $B^{(m)}$ , (29) is a system of nonlinear equations that cannot be solved explicitly.*”. With the tools provided in this paper, an explicit solution is now available. Equation (29) in Hafner [2008] can be replaced with (4) in this paper. Therefore using

$$P^{(m)} = \begin{bmatrix} 0 & I \\ -\Gamma_1^{(m)-1} \Gamma_1^{(m)T} & -\Gamma_1^{(m)-1} \Gamma_0^{(m)} \end{bmatrix}. \quad (11) \quad \boxed{\text{defPm}}$$

we can carry on the procedure described in Section 3. Again, the eigenvalues to choose are those inside the unit circle.

Hafner [2008] shows that  $\Gamma_0^{(m)}$  and  $\Gamma_1^{(m)}$  are analytical functions of  $A$ ,  $B$  and  $E(\xi_t \xi_t^T)$  and these are function of the moments of  $x_t$  only. It turns out that a closed form estimator of the temporally aggregated GARCH(1,1) can be derived as an analytical function  $A^{(m)} = (c^{(m)}, \text{vec } A^{(m)}, \text{vec } B^{(m)}) = G^{(m)}(h, \text{vec } M_0, \text{vec } M_1, \text{vec } M_2)$ . In particular, we can use for the estimation of the aggregated GARCH(1,1) the estimated moments of the high-frequency data, for which more information is available.

The estimator enjoys the same asymptotic properties.

**Theorem 3.** *Under Assumptions 1–4, the estimator  $\hat{\Lambda}^{(m)} = (\hat{c}^{(m)}, \text{vec } \hat{A}^{(m)}, \text{vec } \hat{B}^{(m)}) = G^{(m)}(\mathbf{m})$  is consistent, i.e.  $\hat{\Lambda}^{(m)} \xrightarrow{P} \Lambda^{(m)}$ . In addition, if Assumption 5 holds as well, we have:  $\sqrt{n}(\hat{\Lambda}^{(m)} - \Lambda) \rightarrow N(0, \Xi^{(m)})$ , with*

$$\Xi^{(m)} = \left( \frac{\partial G^{(m)}(\mathbf{m})}{\partial \mathbf{m}} \right) \Psi \left( \frac{\partial G^{(m)}(\mathbf{m})}{\partial \mathbf{m}} \right)^T. \quad (12)$$

Hafner and Rombouts [2007] discuss the estimation of temporally aggregated multivariate GARCH(1,1). Our estimator can be employed as a simple estimator when the number of observations is sufficiently high. Alternatively, it can be used as a consistent starting value for the QML estimation.

## 5. Palindromic matrix equations and eigenvalue problems

In this section, we present the linear algebra results that lead to the estimator of  $B$  (as well as  $B^{(m)}$ ). These matrix equations have been studied by many authors in linear algebra literature, see e.g. Gohberg et al. [1982], Engwerda et al. [1993], Meini [2002] and the references therein; we focus here on providing a non-technical exposition. Proofs of some of the results are presented in Appendix A.

The problem of computing one or more pairs  $(\lambda, u)$  satisfying

$$(\lambda^2 \Gamma_1 + \lambda \Gamma_0 + \Gamma_1^T)u = 0, \quad u \neq 0, \quad (13)$$

is known as *palindromic quadratic eigenvalue problem* Mackey et al. [2006]. The complex numbers  $\lambda$  are called *generalized eigenvalues* and the vectors  $u$  *generalized eigenvectors*. It is indeed a generalization of the standard eigenvalue problem, i.e., given a matrix  $A$  finding pairs  $(\lambda, u)$  satisfying  $(A - \lambda I)u = 0$ .

First, we show that all the solutions to (4) can be constructed from generalized eigenvalues and eigenvectors of (13).

**Lemma 4.** *Let  $(\lambda_1, u_1), (\lambda_2, u_2), \dots, (\lambda_{\bar{d}}, u_{\bar{d}})$  be  $\bar{d}$  different pairs of generalized eigenvalues and eigenvectors of the problem (13), such that the matrix  $U$  (as in (7)) is invertible. Then,  $B = (U^T)^{-1} D U^T$  is a solution of (4). All the solutions of (4) can be obtained in this way.*

Moreover, in our problem the possible values of  $\lambda$  can be “paired”.

pairinglemma

**Lemma 5.** Let  $\Gamma_0$  and  $\Gamma_1$  be real matrices, with  $\Gamma_0$  symmetric. If  $\lambda \neq 0$  is a generalized eigenvalue of (13), then  $1/\lambda$  is one as well.

Finally, the following result shows that we can reduce the palindromic eigenvalue problem to a standard eigenvalue problem.

eigenlemma

**Lemma 6.** Suppose that the matrix  $\Gamma_1$  is invertible.

1. Let  $(\lambda, u)$  be a solution of (13). Then,  $\lambda$  is an eigenvalue of  $P$ , with corresponding eigenvector

$$w = \begin{bmatrix} u \\ \lambda u \end{bmatrix}$$

2. Conversely, if  $\lambda$  is an eigenvalue of  $P$  with eigenvector  $w$ , then  $(\lambda, u)$  satisfy (13), where  $u$  is the vector formed by the first  $\bar{d}$  entries of  $w$ .

By combining all the above results, we can obtain a closed form for  $B$ . Note that different solutions are possible; namely, every choice of  $\bar{d}$  eigenvalues out of the  $2\bar{d}$  of  $P$  gives a different  $B$  satisfying (4); however, only the one with  $|\lambda_i| < 1$ ,  $i = 1, 2, \dots, \bar{d}$  results in a  $B$  with all its roots inside the unit circle.

*Remark 7.* The invertibility of  $\Gamma_1$  is not a crucial assumption. If  $\Gamma_1$  is singular, we can obtain in a similar way not an eigenvalue problem of the form  $Mv = \lambda v$ , but the slightly more general form  $Mv = \lambda Nv$ . This is known as *generalized eigenvalue problem*, and there are plenty of algorithms to find a closed-form solution to it. For instance, the Matlab command `eig(M,N)`. Similarly, if  $P$  is not diagonalizable, solutions  $B$  to the matrix equation (4) can be defined in terms of its Jordan canonical form.

The existence of the partial derivatives  $\frac{\partial F(h, \text{vec } M_0, \text{vec } M_1, \text{vec } M_2)}{\partial (h, \text{vec } M_0, \text{vec } M_1, \text{vec } M_2)}$  that are needed in Theorem 2 can be shown, again under the condition that  $P$  has no unimodular eigenvalues.

analytical

**Lemma 8.** Suppose that the matrix  $P$  has no eigenvalues on the unit circle. Then,  $B$  is an analytical function of the equation coefficients  $\Gamma_0, \Gamma_1$ .

Since  $\Gamma_0$  and  $\Gamma_1$  are in turn analytical functions of the  $M_i$ , the partial derivative exists. A sketch of proof of this result is in the appendix, together with a more explicit expression for the Jacobian.

## 6. Small sample issues

unimodular

The results provided in this paper should be employed with caution whenever the sample size  $n$  is not large enough. The closed-form estimator is based on the sample estimates of  $\hat{M}_0, \hat{M}_1, \hat{M}_2$ . However, it may be the case that the sample moments do not respect all the stated assumptions. More specifically, three different kind of issues can arise:

**Positivity**  $\hat{c}, \hat{A}$  and  $\hat{B}$  do not guarantee that  $H_t$  is positive definite.

**Stationarity** The roots of  $\hat{\Phi} = \hat{M}_{k+1} \hat{M}_k^{-1}$  lie on or outside the unit circle.

**Invertibility**  $\hat{B}$  has unimodular eigenvalues (i.e., on the unit circle).

When the GARCH parameters are estimated via maximum likelihood, the constraints of respecting these conditions are usually imposed when solving the optimization problem; see for instance Chrétien and Ortega [2012]. Since black-box optimization routines are used, additional constraints are easy to impose, but they make the resulting problem more complicated to solve. On the plus side, they guarantee that the resulting model has the desired properties, provided that the iterative optimization procedure does not fail. Instead, with an exact moment-based estimator, if one or more of these conditions fail, then the best way out is modifying the sample moments or the estimates *a posteriori* to make sure that they satisfy these constraints. We discuss briefly these problems that may arise when our closed form estimation is employed.

### 6.1. Positivity

Sufficient conditions for positivity are discussed by several authors (see Gouriéroux [1997], Chrétien and Ortega [2012], Francq and Zakoïan [2010]). However, as far as we know, the problem of finding necessary conditions has not been dealt with in literature. Indeed, even the simpler problem of finding all linear maps among symmetric matrix spaces that preserve positive semi-definiteness has no simple solution, see for instance [Bhatia, 2007, Chapters 2 and 3]. Our estimation procedure does not always produce an estimated GARCH satisfying the sufficient conditions cited above. We do not deal here with the problem of finding a weaker set of conditions that can be preserved.

### 6.2. Stationarity

In small samples the estimate of  $\hat{\Phi} = \hat{M}_{k+1}\hat{M}_k^{-1}$  can have eigenvalues on or outside the unit circle; moreover, the values  $\hat{\Phi}^{(k)}$  computed by choosing different values of  $k$  in the former expression will in general be different. The choice described above of taking  $\hat{\Phi}^{(1)}$  and ignoring all the other autocovariances ratios is the simplest way out of the latter problem. Kristensen and Linton [2006] discuss this problem in the scalar case, and suggest as another valid approach taking  $\frac{1}{3}(\hat{\Phi}^{(1)} + \hat{\Phi}^{(2)} + \hat{\Phi}^{(3)})$ , or in general any convex combination  $\sum w_i \hat{\Phi}^{(i)}$ .

To avoid problems with noninvertibility or outliers, in the multivariate case it is more advisable use instead a least-square solution  $\hat{\Phi}_*$  of the system

$$\hat{\Phi}_* \begin{bmatrix} w_1 \hat{M}_1 & w_2 \hat{M}_2 & \cdots & w_{n-1} \hat{M}_{n-1} \end{bmatrix} = \begin{bmatrix} w_1 \hat{M}_2 & w_2 \hat{M}_3 & \cdots & w_{n-1} \hat{M}_n \end{bmatrix},$$

again with suitably-chosen weights  $w_i$ . None of these solutions (and no choice of weights) clearly stands out. In particular, all of them may result in estimates  $\hat{\Phi}$  with eigenvalues equal or larger than 1. When this happens, a simple fix is projecting the estimate on the space of acceptable GARCH solutions by altering the eigenvalues that lie on or outside the unit circle.

### 6.3. Invertibility

Our assumptions on the solution  $B$  guarantee that it has all eigenvalues inside the unit disc, and thus that the matrix  $P$  has no unimodular eigenvalues. However, once again the sample autocovariances from a finite-time realization of a GARCH process may lead to sample estimates of  $\hat{\Gamma}_0$  and  $\hat{\Gamma}_1$  that do not necessarily guarantee that  $\hat{P}$  has no eigenvalues on the unit circle. This is especially true when the process is close to a non-invertible one. When  $\hat{B}$  has unimodular eigenvalues the invertibility condition does not hold. In addition the following theorem sheds light on the consequences of having eigenvalues lying on the unit disk.

**Theorem 9.** *The following results hold.*

1. *If  $P$  has no eigenvalues lying on the unit circle, then there exist unique solutions  $B$  and  $\Sigma$ , where  $\Sigma = \Sigma^T$  and  $|\lambda| < 1$  for each eigenvalue  $\lambda$  of  $B$ , and they can be computed with the above procedure.*
2. *If  $P$  has eigenvalues lying on the unit circle, then  $B$  and a positive definite  $\Sigma$  satisfying (5) exist only if some strong additional conditions are satisfied (in particular, all unimodular eigenvalues should have even multiplicity). In this case,  $B$  always has unimodular eigenvalues.*

A full proof is more technical than those for the other linear algebra results that we reported; we omit it and refer to Engwerda et al. [1993] for a complete presentation. However, the last assertion is clear in view of our derivation: since the eigenvalues of  $B$  are a subset of those of  $P$ ,  $B$  cannot have all its eigenvalues inside the unit circle if  $P$  has less than  $\bar{d}$  eigenvalues in that domain. If the existence conditions are not satisfied, we can still compute solutions  $B$  with  $\rho(B) = 1$  with the procedure of Section 3; there are multiple solutions, according to which unimodular eigenvalues we choose, but none of them will result in a symmetric  $\Sigma$ . Ad-hoc modifications of  $P$  can be made when unimodular eigenvalues are detected, but in general the accuracy of the computed solution is expected to decrease. Indeed, we show in Appendix B that the derivative of  $B$  with respect to the moments can become unbounded when it has eigenvalues equal to 1.

We are currently working on developing a general procedure for computing a small-norm modification of the  $\hat{M}_k$  that makes the estimated model invertible, rooted on results in linear algebra and eigenvalue perturbation theory.

We point out that the same problem arises in the scalar case treated by Kristensen and Linton [2006]: when (in our notation)  $\Gamma_0/\Gamma_1 > 2$ , the scalar quadratic equation (4) has two complex conjugate solutions with modulus 1, and the procedure breaks down.

## 7. Acknowledgments

We are grateful to Christian Francq, Mark M. Meerschaert, Hans-Peter Scheffler, Dalibor Volny and Jean-Michel Zakoïan for providing interesting and useful comments on an early

revision of this paper. All remaining errors in the paper are solely the responsibility of the authors.

F. Poloni thanks the Rouen Business School for support for a brief research visit, which was instrumental in completing the paper.

## A. Proofs

secproofs

### Proof of Lemma 4

*Proof.* Using the fact that the  $(\lambda_i, u_i)$  are generalized eigenvalues, one can check directly that each column of the matrix

$$\Gamma_1 U D^2 + \Gamma_0 U D + \Gamma_1^T U$$

is zero; therefore,

$$0 = (\Gamma_1 U D^2 + \Gamma_0 U D + \Gamma_1^T U) U^{-1} = \Gamma (U D U^{-1}) (U D U^{-1}) + \Gamma_0 (U D U^{-1}) + \Gamma_1^T,$$

as required. For the converse implication, let  $B^T = (U D U^{-1})$  be the spectral decomposition of a solution; we can reverse all the steps and obtain that each  $(\lambda_i, u_i)$  is a generalized eigenpair.  $\square$

### Proof of Lemma 5

*Proof.* Let  $\lambda$  satisfy (13) for some choice of  $u \neq 0$ . Since  $\Gamma_0, \Gamma_1$  are real, we can take the complex conjugate of every term and get

$$(\bar{\lambda}^2 \Gamma_1 + \bar{\lambda} \Gamma_0 + \Gamma_1^T) \bar{u} = 0,$$

where  $\bar{\lambda}$  and  $\bar{u}$  denote (componentwise) complex conjugation. In particular, this implies that

$$\det(\bar{\lambda}^2 \Gamma_1 + \bar{\lambda} \Gamma_0 + \Gamma_1^T) = 0.$$

Then the determinant of its conjugate transpose must be 0 as well, and thus

$$\det(\lambda^2 \Gamma_1^T + \lambda \Gamma_0 + \Gamma_1) = 0.$$

Multiply everything by  $\frac{1}{\lambda^2}$ , to obtain

$$0 = \det \frac{1}{\lambda^2} (\lambda^2 \Gamma_1^T + \lambda \Gamma_0 + \Gamma_1) = \det \left( \Gamma_1^T + \frac{1}{\lambda} \Gamma_0 + \left( \frac{1}{\lambda} \right)^2 \Gamma_1 \right).$$

Since this determinant is zero, the matrix is singular and there must be a vector  $\tilde{u}$  such that

$$\left( \Gamma_1^T + \frac{1}{\lambda} \Gamma_0 + \left( \frac{1}{\lambda} \right)^2 \Gamma_1 \right) \tilde{u} = 0.$$

But this equation shows that the pair  $(\frac{1}{\lambda}, \tilde{u})$  is also a generalized eigenpair of the polynomial eigenvalue problem.  $\square$

## Proof of Lemma 6

*Proof.* 1. Let  $\lambda, u$  be a solution to (13). We may verify explicitly that the equation

$$\begin{bmatrix} 0 & I \\ -\Gamma_1^{-1}\Gamma_1^T & -\Gamma_1^{-1}\Gamma_0 \end{bmatrix} \begin{bmatrix} u \\ \lambda u \end{bmatrix} = \lambda \begin{bmatrix} u \\ \lambda u \end{bmatrix}$$

holds.

2. Let  $w$  be an eigenvector of  $P$  with eigenvalue  $\lambda$ , and partition it as

$$w = \begin{bmatrix} u \\ v \end{bmatrix}.$$

From the eigenvalue condition  $Pw = \lambda w$  we obtain

$$\begin{bmatrix} 0 & I \\ -\Gamma_1^{-1}\Gamma_1^T & -\Gamma_1^{-1}\Gamma_0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix}.$$

Expanding the two blocks and eliminating  $v$  from the resulting equations, one gets (13). □

## Proof of Lemma 8

The proof follows from some classical results in matrix polynomials that can be found, for instance, in Gohberg et al. [1982]. We give the sketch of a self-contained proof here. We start from a classical result in complex analysis, the Cauchy integral formula

$$\frac{1}{2\pi i} \int_{|z|=1} (z - \lambda)^{-1} dz = \begin{cases} 1 & |\lambda| < 1, \\ 0 & |\lambda| > 1. \end{cases}$$

From this, a matrix version of the same integral follows for diagonal matrices

$$\frac{1}{2\pi i} \int_{|z|=1} (zI - D)^{-1} dz = \Pi,$$

where  $\Pi$  is the diagonal matrix such that  $\Pi_{ii}$  is zero if  $|D_{ii}| > 1$  and one if  $|D_{ii}| < 1$ . Now a change of bases in both sides of the equation yields for all diagonalizable  $A$  without unimodular eigenvalues

$$\frac{1}{2\pi i} \int_{|z|=1} (zI - A)^{-1} dz = \Pi_A, \tag{14} \quad \boxed{\text{intformula}}$$

with  $\Pi_A$  the projector on the invariant subspace of  $A$  associated to the eigenvalues inside the unit circle.

We may generalize further this formula to all  $A$  without unimodular eigenvalues, removing the diagonalizability of  $A$  from the requirements. Indeed, for a non-diagonalizable  $A$ , let us consider a sequence of matrices  $A_k$ , each of them diagonalizable, that converge

uniformly to  $A$ . Such a sequence exists because diagonalizable matrices are dense in  $\mathbb{C}^{n \times n}$ . Since the integrand function is holomorphic on the integration contour, limits and derivatives can be moved inside the integral. This shows that  $\Pi_A$  is an analytical function of  $A$ .

In particular, we apply the formula for  $A$  equal to the  $2\bar{d} \times 2\bar{d}$  matrix  $P$  in (6), for which  $\Pi_P$  has rank  $\bar{d}$  due to the eigenvalue pairing. Let us take any  $2\bar{d} \times \bar{d}$  matrix  $V$  such that  $\Pi_P V$  has full rank and spans the range  $\text{im } \Pi_P$ . Since  $\begin{bmatrix} I \\ B^T \end{bmatrix}$  is another basis for the same subspace, it follows that

$$B^T = \left( \begin{bmatrix} 0 & I_{\bar{d}} \end{bmatrix} \Pi_P V \right) \left( \begin{bmatrix} I_{\bar{d}} & 0 \end{bmatrix} \Pi_P V \right)^{-1}. \quad (15) \quad \boxed{\text{BTformula}}$$

Since invertibility is a condition that holds on an open domain, (15) holds locally with a constant  $V$  and provides an analytical expression for  $B$  in terms of  $\Pi_P$  and thus of  $P$ . The matrix  $P$  is in turn a function of  $\Gamma_0, \Gamma_1$ .

## B. Expressions for the derivatives

appDer

In this section, we give a computable form for the Jacobian of  $F : (h, M_0, M_1, M_2) \mapsto (c, A, B)$ , the function considered in Theorem 2. Rather than using vectorization to give an unwieldy matrix expression, we focus on describing its action as a linear map that takes a first-order perturbation of the moments (denoted by  $(\dot{h}, \dot{M}_0, \dot{M}_1, \dot{M}_2)$ ) to one of the parameters  $(\dot{c}, \dot{A}, \dot{B})$ . We shall use several times the expression for the derivative of the matrix inverse  $\frac{d}{dt}(M^{-1}) = -M^{-1}(\frac{d}{dt}M)M^{-1}$ .

The relation between  $(\dot{h}, \dot{M}_0, \dot{M}_1, \dot{M}_2)$  and  $\dot{\Gamma}_0, \dot{\Gamma}_1$  is easy to compute, by simply differentiating (3):

$$\begin{aligned} \dot{\Gamma}_1 &= \dot{M}_1 - \dot{\Phi}M_0 - \Phi\dot{M}_0, \\ \dot{\Gamma}_0 &= \dot{M}_0 - \dot{M}_1\Phi^T - M_1\dot{\Phi}^T - \dot{\Phi}M_1^T - \Phi\dot{M}_1^T + \dot{\Phi}M_0\Phi^T + \Phi\dot{M}_0\Phi^T + \Phi M_0\dot{\Phi}^T, \end{aligned} \quad (16) \quad \boxed{\text{der1}}$$

with  $\dot{\Phi} = \dot{M}_2M_1^{-1} - M_2M_1^{-1}\dot{M}_1M_1^{-1} = \dot{M}_2M_1^{-1} - \Phi\dot{M}_1M_1^{-1}$ , obtained by differentiating (2) for  $k = 1$ .

We now differentiate (5) to obtain

$$\dot{\Sigma} - B\dot{\Sigma}B^T = \dot{\Sigma} - \Gamma_1\Sigma^{-1}\dot{\Sigma}\Sigma^{-1}\Gamma_1^T = \dot{\Gamma}_0 - \dot{\Gamma}_1\Sigma^{-1}\Gamma_1^T - \Gamma_1\Sigma^{-1}\dot{\Gamma}_1^T = \dot{\Gamma}_0 + \dot{\Gamma}_1B^T + B\dot{\Gamma}_1^T. \quad (17) \quad \boxed{\text{der2}}$$

This is a discrete-time Lyapunov equation (see for instance Gajic and Qureshi [1995]) for  $\dot{\Sigma}$ , which can be solved in closed form by vectorization or numerically by procedures such as Matlab's `dlyap`. The equation is uniquely solvable since we are assuming that  $\rho(B) < 1$ .

Once we have  $\dot{\Sigma}$ , we differentiate  $B = -\Gamma_1\Sigma^{-1}$  to obtain

$$\dot{B} = -\dot{\Gamma}_1\Sigma^{-1} + \Gamma_1\Sigma^{-1}\dot{\Sigma}\Sigma^{-1} = -(\dot{\Gamma}_1 + B\dot{\Sigma})\Sigma^{-1}. \quad (18) \quad \boxed{\text{der3}}$$

The derivatives of the remaining two parameters are given by  $\dot{A} = \dot{\Phi} - \dot{B}$  and  $\dot{c} = (I - \dot{\Phi})h + (I - \Phi)\dot{h}$ .

Putting together (16), (17), (18), one can get to an expression for  $(\dot{c}, \dot{A}, \dot{B})$  as a function of  $(\dot{h}, \dot{M}_0, \dot{M}_1, \dot{M}_2)$ . It does not look like there are any significant simplifications in the resulting expressions. The main message to infer from this computation is that the norm of  $(I - B \otimes B)^{-1}$ , which appears when solving the discrete-time Lyapunov equation, has an impact on the magnitude of the derivatives; the closer  $B$  is to having unimodular eigenvalues, the more ill-conditioned the solution becomes.

## References

- BLR** Luc Bauwens, Sébastien Laurent, and Jeroen V. K. Rombouts. Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, 21(1):79–109, 2006.
- Bhatia** Rajendra Bhatia. *Positive definite matrices*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2007. ISBN 978-0-691-12918-1; 0-691-12918-5.
- B** Tim Bollerslev. Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach. *Review of Economics and Statistics*, 72:498–505, 1990.
- BEW** Tim Bollerslev, Robert F Engle, and Jeffrey M Wooldridge. A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, 96(1):116–31, February 1988. URL <http://ideas.repec.org/a/ucp/jpolec/v96y1988i1p116-31.html>.
- Bou06** Farid Boussama. Ergodicité des chaînes de Markov à valeurs dans une variété algébrique: application aux modèles GARCH multivariés. *C. R. Math. Acad. Sci. Paris*, 343(4):275–278, 2006. ISSN 1631-073X. doi: 10.1016/j.crma.2006.06.027. URL <http://dx.doi.org/10.1016/j.crma.2006.06.027>.
- CO** Stéphane Chrétien and Juan-Pablo Ortega. Multivariate GARCH estimation via a Bregman-proximal trust-region method. *Computational Statistics and Data Analysis*, 2012. To appear. Available as arXiv:1101.5475.
- CL** F. Comte and O. Lieberman. Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis*, 84(1):61 – 84, 2003. ISSN 0047-259X. doi: 10.1016/S0047-259X(02)00009-X. URL <http://www.sciencedirect.com/science/article/pii/S0047259X0200009X>.
- EK** Robert F. Engle and Kenneth F. Kroner. Multivariate simultaneous generalized ARCH. *Econometric Theory*, 11(1):122–150, 1995. ISSN 0266-4666. doi: 10.1017/S0266466600009063. URL <http://dx.doi.org/10.1017/S0266466600009063>.
- EngRR93** Jacob C. Engwerda, André C. M. Ran, and Arie L. Rijkeboer. Necessary and sufficient conditions for the existence of a positive definite solution of

the matrix equation  $X + A^*X^{-1}A = Q$ . *Linear Algebra Appl.*, 186:255–275, 1993. ISSN 0024-3795. doi: 10.1016/0024-3795(93)90295-Y. URL [http://dx.doi.org/10.1016/0024-3795\(93\)90295-Y](http://dx.doi.org/10.1016/0024-3795(93)90295-Y).

**FZ** C. Francq and J.M. Zakoïan. *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley, 2010. ISBN 9780470683910.

**Lyap** Zoran Gajic and Muhammad Tahir Javed Qureshi. *Lyapunov matrix equation in system stability and control*, volume 195 of *Mathematics in Science and Engineering*. Academic Press Inc., San Diego, CA, 1995. ISBN 0-12-273370-3.

**GohLR** I. Gohberg, P. Lancaster, and L. Rodman. *Matrix polynomials*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1982. ISBN 0-12-287160-X. Computer Science and Applied Mathematics.

**G** Christian Gouriéroux. *ARCH models and financial applications*. Springer Series in Statistics. Springer-Verlag, New York, 1997. ISBN 0-387-94876-7. doi: 10.1007/978-1-4612-1860-9. URL <http://dx.doi.org/10.1007/978-1-4612-1860-9>.

**H1** Christian M. Hafner. Fourth moment structure of multivariate GARCH models. *Journal of Financial Econometrics*, 1(1):26–54, 2003. doi: 10.1093/jjfnec/nbg001. URL <http://jfec.oxfordjournals.org/content/1/1/26.abstract>.

**H2** Christian M. Hafner. Temporal aggregation of multivariate GARCH processes. *J. Econometrics*, 142(1):467–483, 2008. ISSN 0304-4076. doi: 10.1016/j.jeconom.2007.08.001. URL <http://dx.doi.org/10.1016/j.jeconom.2007.08.001>.

**HR** Christian M. Hafner and Jeroen V. K. Rombouts. Estimation of temporally aggregated multivariate GARCH models. *J. Stat. Comput. Simul.*, 77(7-8): 629–650, 2007. ISSN 0094-9655. doi: 10.1080/10629360600616252. URL <http://dx.doi.org/10.1080/10629360600616252>.

**J** Thierry Jeantheau. Strong consistency of estimators for multivariate ARCH models. *Econometric Theory*, 14(1):70–86, 1998. ISSN 0266-4666. doi: 10.1017/S0266466698141038. URL <http://dx.doi.org/10.1017/S0266466698141038>.

**LK** Dennis Kristensen and Oliver Linton. A closed-form estimator for the GARCH (1,1) model. *Econometric Theory*, 22(2):323–337, 2006. ISSN 0266-4666. doi: 10.1017/S0266466606060142. URL <http://dx.doi.org/10.1017/S0266466606060142>.

**Vibes** D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Structured polynomial eigenvalue problems: good vibrations from good linearizations. *SIAM J. Matrix Anal. Appl.*, 28(4):1029–1051 (electronic), 2006. ISSN 0895-4798. doi: 10.1137/050628362. URL <http://dx.doi.org/10.1137/050628362>.

**Mei02** Beatrice Meini. Efficient computation of the extreme solutions of  $X + A^*X^{-1}A = Q$  and  $X - A^*X^{-1}A = Q$ . *Math. Comp.*, 71(239):1189–1204 (electronic), 2002. ISSN 0025-5718. doi: 10.1090/S0025-5718-01-01368-0. URL <http://dx.doi.org/10.1090/S0025-5718-01-01368-0>.

- R** Gregory C. Reinsel. *Elements of multivariate time series analysis*. Springer Series in Statistics. Springer-Verlag, New York, second edition, 1997. ISBN 0-387-94918-6.