Mutschler, Willi

Conference Paper

Note on Higher-Order Statistics for the Pruned-State-Space of nonlinear DSGE models

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Macroeconomic Modelling, No. A24-V2

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Mutschler, Willi (2015) : Note on Higher-Order Statistics for the Pruned-State-Space of nonlinear DSGE models, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Macroeconomic Modelling, No. A24-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:
https://hdl.handle.net/10419/113138

Note on Higher-Order Statistics for the Pruned-State-Space of nonlinear DSGE models

Willi Mutschler*

Center for Quantitative Economics, Westfälische Wilhelms-Universität Münster

Abstract

This note shows how to derive unconditional moments, cumulants and polyspectra of order higher than two for the Pruned-State-Space of nonlinear DSGE models. Useful Matrix tools and computational aspects are also discussed.

Keywords: higher-order moments, cumulants, polyspectra, nonlinear DSGE, pruning
JEL: C10, C51, E1

1. Introduction

Since a Gaussian process is completely characterized by its first two moments, most linear DSGE models focus on Gaussian innovations for simplicity. If, however, we relax linearity or use non-Gaussian innovations, it is natural to analyze whether we are able to exploit information from higher-order moments for the identification and estimation of parameters. Researchers in mathematics, statistics and signal processing have developed tools, called higher-order statistics (HOS), to solve detection, estimation and identification problems when the noise source is non-Gaussian or we are faced with nonlinearities; however, applications in the macroeconometric literature are rather sparse.[1] The basic tools of HOS are cumulants, which are defined as the coefficients in the Taylor expansion

*Corresponding author. Center for Quantitative Economics, Westfälische Wilhelms-Universität Münster, Am Stadtgraben 9, 48143 Münster, Tel.: +49-251-83-22914, Fax: +49-251-83-22012, Email: willi.mutschler@wiwi.uni-muenster.de.

[1] For introductory literature and tutorials see the textbooks by Brillinger (2001), Nikias & Petropulu (1993), Priestley (1983) and the references therein. Most theoretical and applied econometric literature is either concerned with tests on normality (e.g. Bao (2013); Rusticelli et al. (2008)) or method of moments estimation (e.g. Dagenais & Dagenais (1997); Erickson & Whited (2002)).

of the log moment generating function in the time-domain; and polyspectra, which are defined as Fourier transformations of the cumulants in the frequency-domain. In this note, we derive closed-form expressions for unconditional third- and fourth-order moments, cumulants and corresponding polyspectra for nonlinear or non-Gaussian DSGE models. We limit ourselves to fourth-order statistics, since third-order cumulants and the bispectrum capture nonlinearities (or non-Gaussianity) for a skewed process, whereas the fourth-order cumulants and the trispectrum can be used in the case of a non-Gaussian symmetric probability distribution. Regarding the approximation of the nonlinear DSGE models we focus on the pruning scheme proposed by Kim et al. (2008) and operationalized by Andreasen et al. (2014), since the Pruned-State-Space (PSS) is a linear, stationary and ergodic state-space system.[2] In the PSS, however, Gaussian innovations do not imply Gaussian likelihood, leaving scope for higher-order statistics to capture information from nonlinearities and non-Gaussianity.

The ideas and procedures derived are useful both from a theoretical and applied point of view. Theoretically, this paper adds to the literature on nonlinear DSGE models by establishing procedures to compute analytically unconditional moments, cumulants and polyspectra for higher-order approximations. An applied researcher may thus use information from higher-order statistics in a moment matching estimation approach or a particle likelihood-type estimation. Our Matlab-code is model-independent and can be found on the homepage of the author.

2. Pruned-State-Space system

Let E_t be the expectation operator conditional on information available at time t, then $E_t f\left(x_{t+1}, y_{t+1}, x_t, y_t|\theta\right) = 0$ is called the general DSGE model with deep parameters θ, states x_t, controls y_t, stochastic innovations u_t, and perturbation parameter σ, which can be cast into a nonlinear first-order system of expectational difference equations f. The solution of such rational expectation models is characterized by so-called *policy-functions*,

[2]Pruning may seem an ad-hoc procedure, however, it can also be theoretically founded as a Taylor expansion in the perturbation parameter (Johnston et al., 2014; Lombardo & Uhlig, 2014) or on an infinite moving average representation (Lan & Meyer-Gohde, 2013).

g and h, that solve (at least approximately) the system of equations f.

$$x_{t+1} = h(x_t, u_{t+1}, \sigma|\theta), \tag{1}$$

$$y_{t+1} = g(x_t, u_{t+1}, \sigma|\theta) \tag{2}$$

For the sake of notation, we assume that all control variables are observable. The vector of innovations u_t is temporally iid with $E(u_t) = 0$ and finite covariance matrix $E(u_t u_t') =: \Sigma = \sigma^2 \eta\eta'$; thus, σ is set to be dependent on the standard deviation of one of the shocks, while scaling all other variances and cross-correlations through η accordingly. Further, u_t has finite higher-order moments depending on the order of approximation.[3] Apart from the existence of moments we do not need to impose any distributional assumptions. In our Matlab we implement both the Gaussian as well as the Student-t distribution, as long as the moments exist.

The approximations of (1) and (1) are a straightforward application of Taylor series expansions in the state variables around the non-stochastic steady-state. However, simulation studies show that due to artificial fixed points, higher-order approximations may generate explosive time-paths even though the linear approximation is stable. Thus, the model may neither be stationary nor imply an ergodic probability distribution, both assumptions are essential for calibration, identification and estimation purposes. Thus, Kim et al. (2008) propose the pruning scheme, in which one leaves out terms in the solution that have higher-order effects than the approximation order. For instance, given a second-order approximation, we decompose the state vector into first-order ($\hat{x}_t^f$) and second-order ($\hat{x}_t^s$) effects ($\hat{x}_t = \hat{x}_t^f + \hat{x}_t^s$), where $\hat{x}_t = x_t - \bar{x}$ and $\hat{y}_t = y_t - \bar{y}$ denote deviations from steady-state, and set up the law of motions for these variables preserving

[3] For a second-order approximation we require at least finite eighth moments. A weaker assumption than iid is to consider u_t being an eighth-order white noise process, which implies y_t being stationary of order four, see Subba Rao & Gabr (1984) for a definition of n-th order stationarity.

only effects up to second-order (see Andreasen et al. (2014) for details):

$$\hat{x}^f_{t+1} = h_x \hat{x}^f_t + h_u u_{t+1} \tag{3}$$

$$\begin{aligned}\hat{x}^s_{t+1} = {} & h_x \hat{x}^s_t + \frac{1}{2} H_{xx} \left(\hat{x}^f_t \otimes \hat{x}^f_t \right) + \frac{1}{2} H_{uu} \left(u_{t+1} \otimes u_{t+1} \right) \\ & + \frac{1}{2} H_{xu} \left(\hat{x}^f_t \otimes u_{t+1} \right) + \frac{1}{2} H_{ux} \left(u_{t+1} \otimes \hat{x}^f_t \right) + \frac{1}{2} h_{\sigma\sigma} \sigma^2\end{aligned} \tag{4}$$

$$\begin{aligned}\hat{y}_{t+1} = {} & g_x (\hat{x}^f_t + \hat{x}^s_t) + g_u u_{t+1} + \frac{1}{2} G_{xx} \left(\hat{x}^f_t \otimes \hat{x}^f_t \right) + \frac{1}{2} G_{uu} \left(u_{t+1} \otimes u_{t+1} \right) \\ & + \frac{1}{2} G_{xu} \left(\hat{x}^f_t \otimes u_{t+1} \right) + \frac{1}{2} G_{ux} \left(u_{t+1} \otimes \hat{x}^f_t \right) + \frac{1}{2} g_{\sigma\sigma} \sigma^2\end{aligned} \tag{5}$$

with H_{xx} being an $n_x \times n_x^2$ matrix containing all second-order terms for the i-th state variable in the i-th row, whereas G_{xx} is an $n_y \times n_x^2$ matrix containing all second-order terms for the i-th control variable in the i-th row. H_{xu}, H_{ux}, G_{xu} and G_{ux} are accordingly shaped for the cross-terms of states and shocks, and H_{uu} and G_{uu} contain the second-order terms for the product of shocks. Thus, terms containing third-order and fourth-order effects are left out, since they are higher than the approximation order. Also there are no second-order effects in u_{t+1}. Further g_x and g_u are the gradients of g with respect to x_t and u_{t+1} respectively. The same notation applies to h_x and h_u. All matrices are evaluated at the non-stochastic steady-state.

Extending the state vector to $z_t := [(\hat{x}^f_t)',\ (\hat{x}^s_t)',\ (\hat{x}^f_t \otimes \hat{x}^f_t)']'$, then equations (3), (4) and (5) can be rewritten as a linear system of equations called the Pruned-State-Space:[4]

$$z_{t+1} = c + A z_t + B \xi_{t+1} \tag{6}$$

$$\hat{y}_{t+1} = d + C z_t + D \xi_{t+1} \tag{7}$$

[4]This approach also works for higher-order approximations. That is, appending the state vector accordingly, we are always able to establish a system linear in the extended state vector. See Andreasen et al. (2014) for the corresponding matrices of the third-order PSS. It has finite fourth moments, if u_t has finite twelfth moments.

where

$$\xi_{t+1} := \begin{bmatrix} u_{t+1} \\ u_{t+1} \otimes u_{t+1} - vec(\Sigma) \\ u_{t+1} \otimes x_t^f \\ x_t^f \otimes u_{t+1} \end{bmatrix}, \qquad c := \begin{bmatrix} 0 \\ \frac{1}{2} h_{\sigma\sigma} \sigma^2 + \frac{1}{2} H_{uu} vec(\Sigma) \\ (h_u \otimes h_u) vec(\Sigma) \end{bmatrix},$$

$$d := \left[\frac{1}{2} g_{\sigma\sigma} \sigma^2 + \frac{1}{2} G_{uu} vec(\Sigma) \right],$$

$$A := \begin{bmatrix} h_x & 0 & 0 \\ 0 & h_x & \frac{1}{2} H_{xx} \\ 0 & 0 & h_x \otimes h_x \end{bmatrix}, \qquad B := \begin{bmatrix} h_u & 0 & 0 & 0 \\ 0 & \frac{1}{2} H_{uu} & \frac{1}{2} H_{ux} & \frac{1}{2} H_{xu} \\ 0 & h_u \otimes h_u & h_u \otimes h_x & h_x \otimes h_u \end{bmatrix},$$

$$C := \left[g_x \quad g_x \quad \frac{1}{2} G_{xx} \right], \qquad D := \left[g_u \quad \frac{1}{2} G_{uu} \quad \frac{1}{2} G_{ux} \quad \frac{1}{2} G_{xu} \right].$$

Thus, conceptually we work in a time-invariant linear state-space system. It can be shown that if the first-order approximation is stable, i.e. all Eigenvalues of h_x have modulus less than one, then the Pruned-State-Space is also stable. In other words, all higher-order terms are unique and all Eigenvalues of A have modulus less than one. Further, if u_t has finite fourth moments, then the Pruned-State-Space system has finite second moments, see Andreasen et al. (2014) for closed-form expressions. We will show below that if u_t has finite eighth moments, then the Pruned-State-Space system has finite fourth moments. Note, that apart from the existence of moments and temporal independence we do not need to impose any distributional assumptions on u_t. Even in the (common) case of u_t being normally distributed, ξ_t is clearly non-Gaussian, therefore leaving scope for higher-order moments to contain additional information.

3. Unconditional moments, cumulants and polyspectra up to fourth-order

The mean of the extended state vector is equal to

$$\mu_z := E(z_t) = (I_{n_z} - A)^{-1} c. \tag{8}$$

with $n_z = 2n_x + n_x^2$. Since there is a linear relationship between y_t and z_{t-1} in (7), we get

$$\mu_y := E(y_t) = \bar{y} + C\mu_z + d. \tag{9}$$

For the derivation of moments, cumulants and spectra we will work with zero-mean processes to simplify notation and expressions. Therefore, we denote $\widetilde{z}_t := z_t - \mu_z$ and $\widetilde{y}_t := y_t - \mu_y$.

Formally, the kth (k=2,3,4)-order cumulants of the kth-order stationary, mean-zero vector process $\widetilde{z}_t$ $(t_1, t_2, t_3 \geq 0)$ are given by the n_z^k vectors $\mathcal{C}_{k,z}$ as

$$\begin{aligned}
\mathcal{C}_{2,z}(t_1) &:= E[\widetilde{z}_0 \otimes \widetilde{z}_{t_1}], \\
\mathcal{C}_{3,z}(t_1, t_2) &:= E[\widetilde{z}_0 \otimes \widetilde{z}_{t_1} \otimes \widetilde{z}_{t_2}], \\
\mathcal{C}_{4,z}(t_1, t_2, t_3) &:= E[\widetilde{z}_0 \otimes \widetilde{z}_{t_1} \otimes \widetilde{z}_{t_2} \otimes \widetilde{z}_{t_3}] - \mathcal{C}_{2,z}(t_1) \otimes \mathcal{C}_{2,z}(t_2 - t_3) \\
&\quad - P'_{n_z} \left(\mathcal{C}_{2,z}(t_2) \otimes \mathcal{C}_{2,z}(t_3 - t_1) \right) - P_{n_z} \left(\mathcal{C}_{2,z}(t_3) \otimes \mathcal{C}_{2,z}(t_1 - t_2) \right),
\end{aligned}$$

where $P_{n_z} = I_{n_z} \otimes U_{n_z^2 \times n_z}$ and $U_{n_z^2 \times n_z}$ is a $(n_z^3 \times n_z^3)$ permutation matrix with unity entries in elements $[(i-1)n_z + j, (j-1)n_z^2], i = 1, \ldots, n_z^2$ and $j = 1, \ldots, n_z$, and zeros else. Here we adopt the compact notation of Swami & Mendel (1990) and store all product-moments of a mean-zero vector-valued process in a vector using Kronecker-products. For example the second moments of $\widetilde{z}_t$ can either be stored in a $n_z \times n_z$ matrix $E(\widetilde{z}_t \cdot \widetilde{z}_t') := \Sigma_z$ or in the $n_z^2 \times 1$ vector $E(\widetilde{z}_t \otimes \widetilde{z}_t) = vec(\Sigma_z)$; this idea naturally carries over to higher orders. There is an intimate relationship between moments and cumulants: if two probability distributions have identical moments, they will have identical cumulants as well. In particular, the second cumulant is equal to the (auto-) covariance matrix and the third cumulant to the (auto-) coskewness matrix. The fourth-order cumulant, however, is the fourth-order product moment ((auto-) cokurtosis matrix) less the second-order moments. In general, for cumulants higher than three, we need knowledge of lower-order moments or cumulants. Nevertheless, using cumulants is for several reasons preferable: For instance, all cumulants of order three and above of a Gaussian process are zero, whereas the same holds true only for odd product-moments. Further, the cumulant of two statistically independent random processes equals the sum of the cumulants of the individual processes (which is not true for higher-order moments). And lastly, cumulants of a white-noise sequence, such as ξ_t, are Kronecker-delta functions, therefore, their polyspectra are flat (Mendel, 1991).[5]

[5]For a mathematical discussion of using cumulants instead of moments in terms of ergodicity and proper functions, see Brillinger (1965).

Assuming that $\mathcal{C}_{k,z}(t_1,\ldots,t_{k-1})$ is absolutely summable the kth-order polyspectrum $\mathcal{S}_{k,z}$ is defined as the (k-1)-dimensional Fourier transform of the kth-order cumulant

$$\mathcal{S}_{k,z}(\omega_1,\ldots,\omega_{k-1}) := \sum_{t_1=-\infty}^{\infty} \cdots \sum_{t_{k-1}=-\infty}^{\infty} \mathcal{C}_{k,z}(t_1,\ldots,t_{k-1}) \cdot exp\{-i\sum_{j=1}^{k-1}\omega_j t_j\},$$

with $\omega_j \in [-\pi;\pi]$ and imaginary i, see Swami et al. (1994) for further details. The second-, third- and fourth-order spectra are called power spectrum, bispectrum and trispectrum, respectively. The power spectrum corresponds to the well-studied spectral density which is a decomposition of the autocorrelation structure of the underlying process (Wiener-Khinchin theorem). The bispectrum can be viewed as a decomposition of the third moments (auto- and cross-skewness) over frequency and is useful for considering systems with asymmetric nonlinearities. In studying symmetric nonlinearities the trispectrum is a more powerful tool, as it represents a decomposition of (auto- and cross-) kurtosis over frequency. Further, both the bi- and trispectrum will be equal to zero for a Gaussian process, thus, departures from Gaussianity will be reflected in these higher-order spectra.

Standard results from VAR(1) systems and insights from HOS can be used regarding the computation of unconditional cumulants and polyspectra in the PSS. First, it is trivial to show that ξ_t is zero-mean white-noise with finite moments, since it is a function of $\hat{x}_t^f, u_{t+1}$ and $u_{t+1} \otimes u_{t+1}$. The kth-order cumulants of ξ_t are

$$\mathcal{C}_{k,\xi}(t_1,\ldots,t_{k-1}) = \begin{cases} \Gamma_{k,\xi} & \text{if } t_1 = \cdots = t_{k-1} = 0, \\ 0 & \text{otherwise,} \end{cases}$$

and corresponding polyspectra $\mathcal{S}_{k,\xi}(\omega_1,\ldots,\omega_{k-1}) = \Gamma_{k,\xi}$ are flat. Regarding the computation of $\Gamma_{k,\xi}$ see also Appendix A. There we show that even if the underlying shock process u_t is Gaussian, ξ_t is not normally distributed, since its higher-order cumulants are not equal to zero.

Letting $[\otimes_{j=1}^{k} X(j)] = X(1) \otimes X(2) \otimes \cdots \otimes X(k)$ for objects $X(j)$, Swami & Mendel (1990) show that given a zero-mean stationary, time-invariant linear state-space system, the cumulants of the state vector $\widetilde{z}_t$

$$\mathcal{C}_{k,z}(t_1,\ldots,t_{k-1}) = [\otimes_{j=0}^{k-1} A^{t_j}] \cdot \mathcal{C}_{k,z}(0,\ldots,0)$$

are given in terms of their zero-lag cumulants

$$\mathcal{C}_{k,z}(0,\ldots,0) = (I_{n_z^k} - [\otimes_{j=1}^k A])^{-1} \cdot [\otimes_{j=1}^k B] \cdot \Gamma_{k,\xi}.$$

which can be efficiently computed using iterative algorithms for generalized Sylvester equations, see Appendix B. Further there is a lot of symmetry (by using appropriate permutation matrices): in particular, all second-order cumulants can be computed from $t_1 > 0$, all third-order cumulants from $t_1 \geq t_2 \geq 0$ and all fourth-order cumulants from $t_1 \geq t_2 \geq t_3 \geq 0$. Since there is a linear relationship between y_t and z_{t-1} in (7), we get closed-form expressions for the kth-order cumulants of our observables (after substracting the mean). That is, for $t_j > 0$

$$\mathcal{C}_{k,y}(0,\ldots,0) = [\otimes_{j=1}^k C]\mathcal{C}_{k,z}(0,\ldots,0) + [\otimes_{j=1}^k D]\Gamma_{k,\xi} \tag{10}$$

$$\mathcal{C}_{k,y}(t_1,\ldots,t_{k-1}) = [\otimes_{j=1}^k C]\mathcal{C}_{k,z}(t_1,\ldots,t_{k-1}) \tag{11}$$

For the polyspectra consider the vector-moving-average representation (VMA) of $\widetilde{z}_t$

$$\widetilde{z}_t = z_t - \mu_z = \sum_{j=0}^{\infty} A^j B \xi_{t-j}.$$

Using equation (7) and lag-operator L, we get the VMA for our controls

$$\widetilde{y}_t = y_t - \bar{y} - C\mu_z - d = \sum_{j=0}^{\infty} CA^j B\xi_{t-j-1} + D\xi_t = H_\xi(L^{-1})\xi_t$$

with transfer function $H_\xi(\mathfrak{z}) = D + C\left(\mathfrak{z} I_{n_z} - A\right)^{-1} B$ for $\mathfrak{z} \in \mathbb{C}$. Setting $\mathfrak{z}_j = e^{-i\omega_j}$, with imaginary i and $\omega_j \in [-\pi;\pi]$, we obtain the Fourier transformations of the cumulants of $\widetilde{y}_t$, i.e. the power spectrum $\mathcal{S}_{2,y}$, bispectrum $\mathcal{S}_{3,y}$ and trispectrum $\mathcal{S}_{4,y}$:

$$\mathcal{S}_{2,y}(\mathfrak{z}_1) = \left[H(\mathfrak{z}_1^{-1}) \otimes H(\mathfrak{z}_1)\right] \Gamma_{2,\xi} \tag{12}$$

$$\mathcal{S}_{3,y}(\mathfrak{z}_1,\mathfrak{z}_2) = \left[H(\mathfrak{z}_1^{-1} \cdot \mathfrak{z}_2^{-1}) \otimes H(\mathfrak{z}_1) \otimes H(\mathfrak{z}_2)\right] \Gamma_{3,\xi} \tag{13}$$

$$\mathcal{S}_{4,y}(\mathfrak{z}_1,\mathfrak{z}_2,\mathfrak{z}_3) = \left[H(\mathfrak{z}_1^{-1} \cdot \mathfrak{z}_2^{-1} \cdot \mathfrak{z}_3^{-1}) \otimes H(\mathfrak{z}_1) \otimes H(\mathfrak{z}_2) \otimes H(\mathfrak{z}_3)\right] \Gamma_{4,\xi}. \tag{14}$$

4. Conclusion

Whenever we are confronted with nonlinearities or non-Gaussian stochastic innovations, it is natural to analyze whether we are able to exploit information from higher-order moments for the identification and estimation of parameters. In this note, we

derive expressions for unconditional moments, cumulants and polyspectra of the Pruned-State-Space representation of a nonlinear DSGE model. In particular, we see that an approximation to higher orders yields nonlinearities and non-Gaussian innovations. Since higher-order cumulants and polyspectra measure the departure from Gaussianity, these expressions can be used to gain additional information from higher-order statistics. For instance, Mutschler (2015) shows how to incorporate these expressions into formal identifiability criteria. Even though our exposition is based on the second-order, an extension to higher-orders is straightforward, since the Pruned-State-Space always results in a system which is linear in an extended state vector. The expressions can also be used for a linear DSGE model with non-Gaussian innovations.

5. References

Andreasen, M. M., Fernández-Villaverde, J., & Rubio-Ramírez, J. F. (2014). *The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications*. Working Paper Aarhus University.

Bao, Y. (2013). On sample skewness and kurtosis. *Econometric Reviews*, *32*, 415–448. doi:10.1080/07474938.2012.690665.

Brillinger, D. R. (1965). An introduction to polyspectra. *The Annals of Mathematical Statistics*, *36*, 1351–1374. doi:10.1214/aoms/1177699896.

Brillinger, D. R. (2001). *Time Series: Data Analysis and Theory*. Classics in Applied Mathematics (Book 36). San Francisco: Society for Industrial and Applied Mathematics.

Dagenais, M. G., & Dagenais, D. L. (1997). Higher moment estimators for linear regression models with errors in the variables. *Journal of Econometrics*, *76*, 193–221. doi:10.1016/0304-4076(95)01789-5.

Erickson, T., & Whited, T. M. (2002). Two-step GMM estimation of the errors-in-variables model using high-order moments. *Econometric Theory*, *3*, 776–799. doi:10.1017/S0266466602183101.

Johnston, M. K., King, R. G., & Lie, D. (2014). *Straightforward approximate stochastic equilibria for nonlinear Rational Expectations models*. Research Paper 59 Crawford School of Public Policy.

Kan, R. (2008). From moments of sum to moments of product. *Journal of Multivariate Analysis*, *99*, 542 – 554. doi:10.1016/j.jmva.2007.01.013.

Kim, J., Kim, S., Schaumburg, E., & Sims, C. A. (2008). Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models. *Journal of Economic Dynamics and Control*, *32*, 3397–3414. doi:10.1016/j.jedc.2008.02.003.

Lan, H., & Meyer-Gohde, A. (2013). *Pruning in Perturbation DSGE Models - Guidance from Nonlinear Moving Average Approximations*. Discussion Papers SFB 649 Humbold University Berlin.

Lombardo, G., & Uhlig, H. (2014). *A theory of pruning*. ECB Working Paper Series Number 1696 European Central Bank.

Magnus, J. R., & Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. (2nd ed.). Wiley.

Meijer, E. (2005). Matrix algebra for higher order moments. *Linear Algebra and its Applications*, *410*, 112 – 134. doi:10.1016/j.laa.2005.02.040.

Mendel, J. M. (1991). Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, *79*, 278–305. doi:10.1109/5.75086.

Nikias, C., & Petropulu, A. (1993). *Higher-order Spectra Analysis: A Nonlinear Signal Processing Framework*. Prentice Hall signal processing series. PTR Prentice Hall.

Priestley, M. (1983). *Spectral Analysis and Time Series, Two-Volume Set: Volumes I and II*. Elsevier Science.

Rusticelli, E., Ashley, R. A., Dagum, E. B., & Patterson, D. M. (2008). A new bispectral test for nonlinear serial dependence. *Econometric Reviews*, *28*, 279–293. doi:10.1080/07474930802388090.

Subba Rao, T., & Gabr, M. (1984). *An introduction to bispectral analysis and bilinear time series models*. Number 24 in Lecture notes in statistics. Springer.

Swami, A., Giannakis, G., & Shamsunder, S. (1994). Multichannel ARMA processes. *IEEE Transactions on Signal Processing*, *42*, 898–913. doi:10.1109/78.285653.

Swami, A., & Mendel, J. (1990). Time and lag recursive computation of cumulants from a state-space model. *IEEE Transactions on Automatic Control*, *35*, 4–17. doi:10.1109/9.45138.

Appendix A. Product moments of innovations

Given a second-order approximation, the innovations are defined as the $n_\xi \times 1$ vector

$$\xi_{t+1} = \begin{pmatrix} u'_{t+1} & (u_{t+1} \otimes u_{t+1} - vec(\Sigma))' & (u_{t+1} \otimes x_t^f)' & (x_t^f \otimes u_{t+1})' \end{pmatrix}'$$

with $n_\xi = n_u + n_u^2 + 2n_x n_u$ elements. We are interested in product moments $M_{2,\xi} := E(\xi_t \otimes \xi_t)$, $M_{3,\xi} := E(\xi_t \otimes \xi_t \otimes \xi_t)$ and $M_{4,\xi} := E(\xi_t \otimes \xi_t \otimes \xi_t \otimes \xi_t)$ with n_ξ^2, n_ξ^3 and n_ξ^4 elements, respectively. In order to compute these objects efficiently, we first reduce the dimension of ξ_t, since it has some duplicate elements. That is, we compute product-moments for the $n_{\tilde{\xi}} = n_u + n_u(n_u+1)/2 + n_u n_x$ vector

$$\tilde{\xi}_{t+1} := \begin{pmatrix} u'_{t+1} & (DP_{n_u}^+(u_{t+1} \otimes u_{t+1} - vec(\Sigma)))' & (u_{t+1} \otimes x_t^f)' \end{pmatrix}'$$

since

$$\xi_t = \begin{pmatrix} I & 0 & 0 \\ 0 & DP_{n_u} & 0 \\ 0 & 0 & I \\ 0 & 0 & K_{n_x,n_u} \end{pmatrix} \tilde{\xi}_t := F_\xi \cdot \tilde{\xi}_t$$

with $DP^{+}_{n_u}$ being the Moore-Penrose-Inverse of the duplication matrix DP_{n_u} and K_{n_x,n_u} the commutation matrix such that $K_{n_x,n_u}(u_{t+1} \otimes x^f_t) = (x^f_t \otimes u_{t+1})$. Then we have

$$M_{k,\xi} := [\otimes^k_{j=1} F_\xi] \cdot M_{k,\tilde{\xi}}$$

denoting the k-th (k=2,3,4)-order product moment of $\tilde{\xi}_t$. Since $[\otimes^k_{j=1} F_\xi]$ does not change with θ, we can focus on $M_{k,\tilde{\xi}}$. $M_{k,\tilde{\xi}}$, however, contains also many duplicate elements. Denote with $\widetilde{M}_{k,\tilde{\xi}}$ the unique elements of $M_{k,\tilde{\xi}}$, we have the following relationships:

$$M_{2,\tilde{\xi}} = DP_{n_{\tilde{\xi}}} \cdot \widetilde{M}_{2,\tilde{\xi}}, \quad M_{3,\tilde{\xi}} = TP_{n_{\tilde{\xi}}} \cdot \widetilde{M}_{3,\tilde{\xi}}, \quad M_{4,\tilde{\xi}} = QP_{n_{\tilde{\xi}}} \cdot \widetilde{M}_{4,\tilde{\xi}},$$

with the duplication matrix $DP_{n_{\tilde{\xi}}}$ defined by Magnus & Neudecker (1999), and the triplication matrix $TP_{n_{\tilde{\xi}}}$ and quadruplication matrix $QP_{n_{\tilde{\xi}}}$ similarly defined by Meijer (2005).[6] Note that these matrices are independent of θ and their Moore-Penrose-Inverse always exists, e.g. $(QP'_{n_{\tilde{\xi}}} QP_{n_{\tilde{\xi}}})^{-1} QP'_{n_{\tilde{\xi}}} \cdot M_{4,\tilde{\xi}} = \widetilde{M}_{4,\tilde{\xi}}$. Further, $DP_{n_{\tilde{\xi}}}$, $TP_{n_{\tilde{\xi}}}$ and $QP_{n_{\tilde{\xi}}}$ are constructed such that there is a unique ordering in $\widetilde{M}_{k,\tilde{\xi}}$, see Meijer (2005) for an example and more details.

To compute the product-moments of $\tilde{\xi}_t$ symbolically we therefore use the following procedure in Matlab given the number of shocks n_u, the number of state variables n_x and the order of product moments k=2,3,4.

1. Define $u_{t+1} = (u_{t+1,1}, \dots u_{t+1,n_u})'$, $x^f_t = (x^f_{t,1}, \dots x^f_{t,n_x})'$ and $\Sigma_u = [sig_{ij}]_{nu \times nu}$ symbolically with $i, j = 1, \dots n_u$. Set up

$$\tilde{\xi}_t = (u'_t, DP^{+}_{n_u}(u_{t+1} \otimes u_{t+1} - vec(\Sigma))', (u_{t+1} \otimes x^f_t)')'.$$

2. Get all integer permutations of $[i_1, i_2, \dots i_{n_{\tilde{\xi}}}]$ that sum up to k, with $i_j = 1, \dots, k$ and $j = 1, \dots, n_{\tilde{\xi}}$. Sort them in the ordering of Meijer (2005).
3. For each permutation $[i_1, i_2, \dots i_{n_{\tilde{\xi}}}]$ evaluate symbolically

$$E\left[(\tilde{\xi}_{1,t})^{i_1} \cdot (\tilde{\xi}_{2,t})^{i_2} \cdot \dots (\tilde{\xi}_{n_{\tilde{\xi}},t})^{i_{n_{\tilde{\xi}}}}\right]$$

   and store it in the vector $\widetilde{M}_{k,\xi}$.

[6] Actually $\widetilde{M}_{k,\tilde{\xi}}$ has some further duplicate terms for $n_u, n_x > 1$ due to higher-order cross terms of u_{t+1} and x^f_t, which we can further reduce using indices from the `unique` function of Matlab.

4. Optionally: Use Matlab's `unique` function to further reduce the dimension of $\widetilde{M}_{k,\xi}$.

The expressions we get in step 3 contain terms of the form

$$const. \cdot E[(u_{1,t+1})^{i_{u_1}} \cdot (u_{2,t+1})^{i_{u_2}} \cdot \ldots \cdot (u_{n_u,t+1})^{i_{u_{n_u}}}] \cdot E[(x_{1,t}^f)^{i_{x_1}} \cdot (x_{2,t}^f)^{i_{x_2}} \cdot \ldots \cdot (x_{n_x,t}^f)^{i_x^{n_x}}],$$

that is joint product moments of the elements of u_{t+1} and x_t^f (keeping in mind that x_t^f and u_{t+1} are independent due to the temporal independence of u_t). For instance, for $n_u = n_x = 1$ the third-order product moment of $\tilde{\xi}_t$ is equal to

$$\tilde{M}_{3,\xi} = vec\left(E\left[\begin{bmatrix} u^3 & u^4 - \sigma_u^2 u^2 \\ u^3 x & \sigma_u^4 u - 2\sigma_u^2 u^3 + u^5 \\ xu^4 - \sigma_u^2 x u^2 & u^3 x^2 \\ -\sigma_u^6 + 3\sigma_u^4 u^2 - 3\sigma_u^2 u^4 + u^6 & x\sigma_u^4 u - 2x\sigma_u^2 u^3 + xu^5 \\ u^4 x^2 - \sigma_u^2 u^2 x^2 & u^3 x^3 \end{bmatrix}' \right)$$

where we dropped sub- and superscripts and $E(u^2) = \sigma_u^2$. Given a function that evaluates the moment structure of x_t^f and u_{t+1} either analytically or numerically, we are able to calculate these terms individually and save them into script files. Note, that these computations need only to be done once for a model, after that we simply evaluate the script files numerically given model parameters θ. Our code can evaluate product moments from the Gaussian as well as Student-t distribution.

In the case that u_t is normally distributed, x_t^f is also Gaussian with covariance matrix Σ_x. Therefore,

$$\begin{pmatrix} u_{t+1} \\ x_t^f \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma_x \end{pmatrix} \right)$$

is multivariate normal. All joint product moments are therefore functions of the variances and covariances in Σ and Σ_x and can be computed analytically. To this end, we use the very efficient method and Matlab function of Kan (2008) to derive these joint product moments symbolically. For our example with $n_u = n_x = 1$ and Gaussian u_t, we get the

unique entries

$$\widetilde{M}_{2,\xi} = \left[\sigma_u^2,\ 0,\ 0,\ 2\sigma_u^4,\ 0,\ \sigma_u^2\sigma_x^2\right]'$$
$$\widetilde{M}_{3,\xi} = \left[0,\ 2\sigma_u^4,\ 0,\ 0,\ 0,\ 0,\ 8\sigma_u^6,\ 0,\ 2\sigma_u^4\sigma_x^2,\ 0\right]'$$
$$\widetilde{M}_{4,\xi} = \left[3\sigma_u^4,\ 0,\ 0,\ 10\sigma_u^6,\ 0,\ 3\sigma_u^4\sigma_x^2,\ 0,\ 0,\ 0,\ 0,\ 60\sigma_u^8,\ 0,\ 10\sigma_u^6\sigma_x^2,\ 0,\ 9\sigma_u^4\sigma_x^4\right]'$$

where $E(x_t^{f2}) = \sigma_x^2$. The cumulants can then be computed as outlined in section 3. Since the third-order cumulant of a Gaussian process must be zero, we now see, that ξ_t is clearly non-Gaussian, since its third-order cumulant is different from zero, even if the underlying distribution for u_t is Gaussian.

In the case that u_t is Student-t distributed with v degrees of freedom, we rewrite u_t in terms of a Inverse-Gamma distributed variable $W = v^{-1/2} \sim IGAM(v/2, v/2)$, and a normally distributed variable $\varepsilon_t \sim N(0, \Sigma)$, $u_t = v^{-1/2}\varepsilon_t$. Since W and ε_t are independent, we have $E(u_t u_t') = E(W)E(\varepsilon_t\varepsilon_t') = \frac{v}{v-2}\Sigma$. Whereas all odd product moments of u_t are zero, the even product moments ($n = \sum_{j=1}^{n_u} i_{u_j}$ is an even number) are given by

$$E[(u_{1,t})^{i_{u_1}} \cdot (u_{2,t})^{i_{u_2}} \cdot \ldots \cdot (u_{n_u,t})^{i_{u_{n_u}}}] = E[W^{\frac{n}{2}}] \cdot E[(\varepsilon_{1,t})^{i_{u_1}} \cdot (\varepsilon_{2,t})^{i_{u_2}} \cdot \ldots \cdot (\varepsilon_{n_u,t})^{i_{u_{n_u}}}].$$

The first term is equal to $E[W^n] = \frac{v/2}{(v/2-1)\ldots(v/2-n)}$ and since ε_t is multivariate normal, we can use Kan (2008)'s procedure and Matlab function for the second product. Similar arguments apply to the product moments of x_t^f, for instance the variance is given by

$$vec(\Sigma_x) = E[x_t^f \otimes x_t^f] = \underbrace{E[W]}_{\frac{v}{v-2}} \cdot (I_{n_x^2} - h_x \otimes h_x)^{-1}(h_u \otimes h_u) \cdot \underbrace{E[\varepsilon_t \otimes \varepsilon_t]}_{vec(\Sigma)}.$$

Thus, odd product moments are also zero, whereas even product moments can also be computed symbolically by Kan (2008)'s procedure and Matlab function, however, adjusted for $E[W^{n/2}]$.

Appendix B. Using generalized Sylvester equations for cumulants

The zero-lag cumulants (k=2,3,4)

$$\mathcal{C}_{k,z} = (I_{n_z^k} - [\otimes_{j=1}^k A])^{-1} \cdot [\otimes_{j=1}^k B] \cdot \Gamma_{k,\xi}$$

require the inversion of the big matrix $\left(I_{n_z^k} - [\otimes_{j=1}^{k} A]\right)$. Since $\mathcal{C}_{k,z}$ and $\Gamma_{k,\xi}$ are vectors, we can use properties of the Kronecker-product and rewrite the equations to

$$
\begin{aligned}
[\underset{n_z \times n_z}{\ \mathcal{C}_{2,z}\ }] &= A[\underset{n_z \times n_z}{\ \mathcal{C}_{2,z}\ }]A' + B[\underset{n_\xi \times n_\xi}{\ \Gamma_{2,\xi}\ }]B', \\
[\underset{n_z^2 \times n_z}{\ \mathcal{C}_{3,z}\ }] &= (A \otimes A)[\underset{n_z^2 \times n_z}{\ \mathcal{C}_{3,z}\ }]A' + (B \otimes B)[\underset{n_\xi^2 \times n_\xi}{\ \Gamma_{3,\xi}\ }]B', \\
[\underset{n_z^2 \times n_z^2}{\ \mathcal{C}_{4,z}\ }] &= (A \otimes A)[\underset{n_z^2 \times n_z^2}{\ \mathcal{C}_{4,z}\ }](A \otimes A)' + (B \otimes B)[\underset{n_\xi^2 \times n_\xi^2}{\ \Gamma_{4,\xi}\ }](B \otimes B)',
\end{aligned}
$$

where $[\underset{n \times m}{\ }]$ reshapes a $n \cdot m$ vector into a $n \times m$ matrix. In other words, we reduce the inversion problem to a generalized Sylvester equation, which can be efficiently solved.