15

Fitting Finite Order VAR Models to Infinite Order Processes

15.1 Background

In the previous chapters, we have derived properties of models, estimators, forecasts, and test statistics under the assumption of a true model. We have also argued that such an assumption is virtually never fulfilled in practice. In other words, in practice, all we can hope for is a model that provides a useful approximation to the actual data generation process of a given multiple time series. In this chapter, we will, to some extent, take into account this state of affairs and assume that an approximating rather than a true model is fitted. Specifically, we assume that the true data generation process is an infinite order VAR process and, for a given sample size T, a finite order VAR(p) is fitted to the data.

In practice, it is likely that a higher order VAR model is considered if the sample size or time series length is larger. In other words, the order p increases with the sample size T. If an order selection criterion is used in choosing the VAR order, the maximum order to be considered is likely to depend on T. This again implies that the actual order chosen depends on the sample size because it will depend on the maximum order. In summary, the actual order selected may be regarded as a function of the sample size T. In order to derive statistical properties of estimators and forecasts, we will make this assumption in the following. More precisely, we will assume that the VAR order goes to infinity with the sample size. Under that assumption, an asymptotic theory has been developed that will be discussed in this chapter.

In Section 15.2, the assumptions for the underlying true process and for the order of the process fitted to the data are specified in detail and asymptotic estimation results are provided for stable processes. In Section 15.3, the consequences for forecasting are discussed and impulse response analysis is considered in Section 15.4. Our standard investment/income/consumption example is used to contrast the present approach to that considered in Chapter 3, where a true finite order process is assumed. Finally, in Section 15.5, extensions to cointegrated processes are discussed.

H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*

15.2 Multivariate Least Squares Estimation

Suppose the generation process of a given multiple time series is a stationary, stable, K-dimensional, infinite order VAR process,

$$y_t = \sum_{i=1}^{\infty} \Pi_i y_{t-i} + u_t, \tag{15.2.1}$$

with absolutely summable Π_i, that is,

$$\sum_{i=1}^{\infty} \|\Pi_i\| < \infty \tag{15.2.2}$$

(see Appendix C.3) and canonical MA representation

$$y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i}, \quad \Phi_0 = I_K, \tag{15.2.3}$$

satisfying

$$\det\left(\sum_{i=0}^{\infty} \Phi_i z^i\right) \neq 0 \quad \text{for } |z| \leq 1 \quad \text{and} \quad \sum_{i=1}^{\infty} i^{1/2} \|\Phi_i\| < \infty. \tag{15.2.4}$$

The zero mean assumption implied by these conditions is not essential and is imposed for convenience only. Stable, invertible VARMA processes satisfy the foregoing conditions. The assumptions allow for more general processes, however. Of course, the generation process may also be a stable, finite order VAR(p) in which case $\Pi_i = 0$ for $i > p$.

We have argued in the previous section that in practice the true structure will usually be unknown and the investigator may consider fitting a finite order VAR process with the VAR order depending on the length T of the available time series. For this situation, Lewis & Reinsel (1985) have shown consistency and asymptotic normality of the multivariate LS estimators. For univariate processes, similar results were discussed earlier by Berk (1974) and Bhansali (1978).

To state these results formally, we use the following notation:

$$\Pi(n) := [\Pi_1, \ldots, \Pi_n],$$

$$\boldsymbol{\pi}(n) := \text{vec } \Pi(n).$$

Fitting a VAR(n) process, the i-th estimated coefficient matrix is denoted by $\widehat{\Pi}_i(n)$,

$$\widehat{\Pi}(n) := [\widehat{\Pi}_1(n), \ldots, \widehat{\Pi}_n(n)],$$

and

$$\widehat{\boldsymbol{\pi}}(n) := \text{vec } \widehat{\Pi}(n).$$

Now we can state a result of Lewis & Reinsel (1985).

Proposition 15.1 (*Properties of the LS Estimator of an Approximating VAR Model*)
Let the multiple time series $y_1, \dots, y_T$ be generated by a potentially infinite order VAR process satisfying (15.2.1)–(15.2.4) with standard white noise u_t. Suppose finite order VAR(n_T) processes are fitted by multivariate LS and assume that the order n_T depends upon the sample size T such that

$$n_T \to \infty, \quad n_T^3/T \to 0, \quad \text{and} \quad \sqrt{T} \sum_{i=n_T+1}^{\infty} \|\Pi_i\| \to 0 \quad \text{as} \quad T \to \infty. \tag{15.2.5}$$

Furthermore, let c_1, c_2 be positive constants and $\mathbf{f}(n)$ a sequence of $(K^2 n \times 1)$ vectors such that

$$0 < c_1 \le \mathbf{f}(n)'\mathbf{f}(n) \le c_2 < \infty \quad \text{for } n = 1, 2, \dots.$$

Then

$$\frac{\sqrt{T - n_T}\,\mathbf{f}(n_T)'[\widehat{\boldsymbol{\pi}}(n_T) - \boldsymbol{\pi}(n_T)]}{[\mathbf{f}(n_T)'(\Gamma_{n_T}^{-1} \otimes \Sigma_u)\mathbf{f}(n_T)]^{1/2}} \xrightarrow{d} \mathcal{N}(0, 1), \tag{15.2.6}$$

where

$$\Gamma_n := E\left(\begin{bmatrix} y_t \\ \vdots \\ y_{t-n+1} \end{bmatrix} [y_t', \dots, y_{t-n+1}']\right). \tag{15.2.7}$$

■

Remark 1 The assumption (15.2.5) means that, although the VAR order has to go to infinity with the sample size, it has to do so at a much slower rate because $n_T^3/T \to 0$. The requirement

$$\sqrt{T} \sum_{i=n_T+1}^{\infty} \|\Pi_i\| \to 0 \tag{15.2.8}$$

is always satisfied if y_t is actually a finite order VAR process and $n_T \to \infty$. For infinite order VAR processes, this condition implies a lower bound for the rate at which n_T goes to infinity. To see this, consider the *univariate* MA(1) process

$$y_t = u_t - m u_{t-1},$$

where $0 < |m| < 1$ to ensure invertibility. Its AR representation is

$$y_t = -\sum_{i=1}^{\infty} m^i y_{t-i} + u_t$$

and condition (15.2.8) becomes

$$\begin{aligned}\sqrt{T}\sum_{i=n+1}^{\infty}|m^i| &= \sqrt{T}|m|^{n+1}\sum_{i=0}^{\infty}|m|^i \\ &= \sqrt{T}\frac{|m|^{n+1}}{1-|m|}\underset{T\to\infty}{\longrightarrow} 0. \qquad (15.2.9)\end{aligned}$$

Here the subscript T has been dropped from n_T for notational simplicity. In this example, $n_T = T^{1/\epsilon}$ with $\epsilon > 3$ is a possible choice for the sequence n_T that satisfies both (15.2.9) and $n_T^3/T \to 0$. On the other hand, $n_T = \ln\ln T$ is not a permissible choice because in this case

$$\sqrt{T}|m|^{n_T+1}$$

does not approach zero as $T \to \infty$. This result is easily established by considering the logarithm of (15.2.9),

$$\tfrac{1}{2}\ln T + (n_T+1)\ln|m| - \ln(1-|m|),$$

which goes to infinity for $n_T = \ln\ln T$.

In summary, (15.2.8) is a lower bound and $n_T^3/T \to 0$ establishes an upper bound for the rate at which n_T has to go to infinity with the sample size T. ■

Remark 2 Proposition 15.1 implies that for fixed m,

$$\sqrt{T-n_T}\,\mathrm{vec}([\widehat{\Pi}_1(n_T),\dots,\widehat{\Pi}_m(n_T)] - [\Pi_1,\dots,\Pi_m])$$

has an asymptotic multivariate normal distribution with mean zero and covariance matrix $V \otimes \Sigma_u$, where V is obtained as follows: Let V_n be the upper left-hand $(Km \times Km)$ block of the inverse of Γ_n, for $n \geq m$. Then $V = \lim_{n\to\infty} V_n$. Loosely speaking, V is the upper left-hand $(Km \times Km)$ block of the inverse of the infinite order matrix

$$E\left(\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \end{bmatrix}\left[y_t', y_{t-1}', \dots\right]\right).$$

Thus, the result can be used for inference on a finite number of parameters. It is also possible, however, to use the result from Proposition 15.1 to construct tests for hypotheses involving an infinite number of restrictions. Such hypotheses can arise in studying Granger-causality in infinite order VAR processes. This case was considered explicitly by Lütkepohl & Poskitt (1996). ■

Remark 3 If the data generation process has nonzero mean originally, the sample mean $\overline{y}$ may be subtracted initially from the data. It is asymptotically independent of the $\widehat{\Pi}_i(n_T)$ and has an asymptotic normal distribution,

$$\sqrt{T}(\overline{y} - \mu) \xrightarrow{d} \mathcal{N}(0, \Sigma_{\overline{y}}),$$

where

$$\Sigma_{\overline{y}} = \left(\sum_{i=0}^{\infty} \Phi_i\right) \Sigma_u \left(\sum_{i=0}^{\infty} \Phi_i\right)'.$$

■

A corresponding result from Lütkepohl & Poskitt (1991) for the white noise covariance matrix is stated next.

Proposition 15.2 (*Asymptotic Properties of the White Noise Covariance Matrix Estimator*)
Let

$$\widehat{u}_t(n) := y_t - \sum_{i=1}^{n} \widehat{\Pi}_i(n) y_{t-i}, \quad t = 1, \dots, T,$$

be the multivariate LS residuals from a VAR(n) model fitted to a multiple time series of length T, let

$$\widetilde{\Sigma}_u(n) := \frac{1}{T} \sum_{t=1}^{T} \widehat{u}_t(n) \widehat{u}_t(n)'$$

be the corresponding estimator of the white noise covariance matrix and let $U := [u_1, \dots, u_T]$ so that

$$\frac{1}{T} UU' = \frac{1}{T} \sum_{t=1}^{T} u_t u_t'$$

is an estimator of Σ_u based on the true white noise process u_t. Then, under the conditions of Proposition 15.1,

$$\operatorname{plim} \sqrt{T}(\widetilde{\Sigma}_u(n_T) - T^{-1}UU') = 0.$$

■

We know from Chapter 3, Propositions 3.2 and 3.4, that, for a Gaussian process, $T^{-1}UU'$ has an asymptotic normal distribution,

$$\sqrt{T}\, \mathrm{vech}(T^{-1}UU' - \Sigma_u) \xrightarrow{d} \mathcal{N}(0, 2\mathbf{D}_K^+ (\Sigma_u \otimes \Sigma_u) \mathbf{D}_K^{+\prime}), \qquad (15.2.10)$$

where, as usual, $\mathbf{D}_K^+ = (\mathbf{D}_K' \mathbf{D}_K)^{-1} \mathbf{D}_K'$ is the Moore-Penrose inverse of the $(K^2 \times \frac{1}{2}K(K+1))$ duplication matrix $\mathbf{D}_K$. Using Proposition C.2(2) of Appendix C.1, Proposition 15.2 implies that

$$\sqrt{T}\,\mathrm{vech}(\widetilde{\Sigma}_u(n_T) - \Sigma_u)$$

has precisely the same asymptotic distribution as the one in (15.2.10). Obviously, this distribution does not depend on the VAR structure of y_t or the VAR coefficients. In addition, the estimator $\widetilde{\Sigma}_u(n_T)$ is asymptotically independent of $\widehat{\boldsymbol{\pi}}(n_T)$. In the following, the consequences of these results for prediction and impulse response analysis will be discussed.

15.3 Forecasting

15.3.1 Theoretical Results

Suppose the VAR(n_T) model estimated in the previous section is used for forecasting. In that case, the usual h-step forecast at origin T, $\widetilde{y}_T(h)$, can be computed recursively for $h = 1, 2, \dots$, using

$$\widetilde{y}_T(h) = \sum_{i=1}^{n_T} \widehat{\Pi}_i(n_T)\widetilde{y}_T(h-i), \tag{15.3.1}$$

where $\widetilde{y}_T(j) := y_{T+j}$ for $j \leq 0$ (see Section 3.5). We use the notation

$$\widetilde{\mathbf{y}}_T(h) := \begin{bmatrix} \widetilde{y}_T(1) \\ \vdots \\ \widetilde{y}_T(h) \end{bmatrix}, \quad \mathbf{y}_T(h) := \begin{bmatrix} y_T(1) \\ \vdots \\ y_T(h) \end{bmatrix}, \quad \mathbf{y}_{T,h} := \begin{bmatrix} y_{T+1} \\ \vdots \\ y_{T+h} \end{bmatrix}$$

and

$$\boldsymbol{\Sigma}_{\mathbf{y}}(h) := E\left\{[\mathbf{y}_{T,h} - \mathbf{y}_T(h)][\mathbf{y}_{T,h} - \mathbf{y}_T(h)]'\right\},$$

where $y_T(j)$, $j = 1, \dots, h$, is the optimal j-step forecast at origin T based on the infinite past, that is,

$$y_T(j) = \sum_{i=1}^{\infty} \Pi_i y_T(j-i)$$

with $y_T(i) := y_{T+i}$ for $i \leq 0$ (see Section 11.5). The following result is also essentially due to Lewis & Reinsel (1985) (see also Lütkepohl (1987, Section 3.3, Proposition 3.2)).

Proposition 15.3 (*Asymptotic Distributions of Estimated Forecasts*)
Under the conditions of Proposition 15.1, if y_t is a Gaussian process and if independent processes with identical stochastic structures are used for estimation and forecasting, respectively, then

$$\sqrt{\frac{T}{n_T}}[\widetilde{\mathbf{y}}_T(h) - \mathbf{y}_T(h)] \xrightarrow{d} \mathcal{N}(0, K\boldsymbol{\Sigma}_{\mathbf{y}}(h))$$

for $h = 1, 2 \dots$. ■

Remark 1 The proposition implies that for large samples the forecast vector $\widetilde{\mathbf{y}}_T(h)$ has approximate MSE matrix

$$\mathbf{\Sigma}_{\widetilde{\mathbf{y}}}(h) = \left(1 + \frac{Kn_T}{T}\right) \mathbf{\Sigma}_{\mathbf{y}}(h). \tag{15.3.2}$$

This result can be seen by noting that

$$\begin{aligned} & E\{[\mathbf{y}_{T,h} - \widetilde{\mathbf{y}}_T(h)][\mathbf{y}_{T,h} - \widetilde{\mathbf{y}}_T(h)]'\} \\ & \quad = \; E\{[\mathbf{y}_{T,h} - \mathbf{y}_T(h)][\mathbf{y}_{T,h} - \mathbf{y}_T(h)]'\} \\ & \qquad + E\{[\mathbf{y}_T(h) - \widetilde{\mathbf{y}}_T(h)][\mathbf{y}_T(h) - \widetilde{\mathbf{y}}_T(h)]'\} \end{aligned}$$

and approximating the last term via the asymptotic result of Proposition 15.3. ■

Remark 2 An approximation for the MSE matrix of an h-step forecast $\widetilde{y}_T(h)$ follows directly from (15.3.2),

$$\Sigma_{\widetilde{y}}(h) = \left(1 + \frac{Kn_T}{T}\right) \Sigma_y(h), \quad h = 1, 2, \ldots. \tag{15.3.3}$$

In Section 3.5.1, we have obtained an approximate MSE matrix

$$\Sigma_{\widehat{y}}(h) = \Sigma_y(h) + \frac{1}{T}\Omega(h) \tag{15.3.4}$$

for an h-step forecast based on an estimated VAR process with known finite order. If in Chapter 3 the process mean is known to be zero and is not estimated, it can be shown that $\Omega(h)$ approaches zero as $h \to \infty$. In other words, the MSE part due to estimation variability goes to zero as the forecast horizon increases. The same does not hold in the present case. In fact, the $\Sigma_{\widetilde{y}}(h)$'s are monotonically nondecreasing for growing h, that is,

$$\Sigma_{\widetilde{y}}(h) \geq \Sigma_{\widetilde{y}}(i), \qquad \text{for } h \geq i.$$

The explanation for this result is that, under the present assumptions, increasingly many parameters are estimated with growing sample size. For a zero mean VAR process with known finite order, the optimal forecast approaches the process mean of zero when the forecast horizon gets large and, thus, the estimated VAR parameters do not contribute to the forecast uncertainty for long-run forecasts. The same is not true under the present conditions, where the VAR order goes to infinity. ■

Remark 3 We have also seen in Section 3.5.2 that $\Omega(1) = (Kp + 1)\Sigma_u$ for a K-dimensional VAR(p) process with estimated intercept term. It is easy to see that, if the process mean is known to be zero and the mean term is not estimated, $\Omega(1) = Kp\Sigma_u$. Hence, in that case,

$$\Sigma_{\widehat{y}}(1) = \Sigma_y(1) + \frac{Kp}{T}\Sigma_u = \; \Sigma_u + \frac{Kp}{T}\Sigma_u = \; \Sigma_{\widetilde{y}}(1),$$

if $n_T = p$. In other words, for 1-step ahead forecasts, the two MSE approximations are identical if the same VAR orders are used in both approaches. It is easy to see that the same does not hold in general for predictions more than 1 step ahead (see Problem 15.2). ■

Remark 4 Because forecasts can be obtained from finite order approximations to infinite order VAR processes, we may also base the prediction tests for structural change considered in Sections 4.6.2 and 13.5.3 on such approximations. Of course, in that case the MSE approximation implied by Proposition 15.3 should be used in setting up the test statistics. For instance, a test statistic based on h-step forecasts would be

$$\widetilde{\tau}_h = (y_{T+h} - \widetilde{y}_T(h))' \widetilde{\Sigma}_{\widetilde{y}}(h)^{-1}(y_{T+h} - \widetilde{y}_T(h)),$$

where $\widetilde{\Sigma}_{\widetilde{y}}(h)$ is an estimator of $\Sigma_{\widetilde{y}}(h)$. ■

Remark 5 If y_t is a process with nonzero mean vector μ, then the sample mean may be subtracted from the original data and the previous analysis may be performed with the mean-adjusted data. If the sample mean is added to the forecasts, an extra term should be added to the approximate MSE matrix. A term similar to that resulting from an estimated mean term in a finite order VAR setting with known order may be added (see Problem 3.9, Chapter 3). ■

15.3.2 An Example

To illustrate the effects of approximating a potentially infinite order VAR process by a finite order model, we use again the West German investment/income/consumption data from File E1. The variables y_1, y_2, and y_3 are defined as in Chapter 3, Section 3.2.3, and we use the same sample period 1960–1978 and a VAR order $n_T = 2$. That is, we assume that the VAR order depends on the sample size in such a way that $n_T = 2$ for $T = 73$. Note that the condition (15.2.5) for the VAR order is an asymptotic condition that leaves open the actual choice in finite samples. Therefore, we choose the VAR order that was suggested by the AIC criterion in Chapter 4 and, thus, we use the same VAR order as in Chapter 3. As a consequence, the point forecasts obtained under our present assumptions are the same one gets from a mean-adjusted model under the conditions of Chapter 3. The interval forecasts obtained under the different sets of assumptions are different for $h > 1$, however, because the approximate MSE matrices are different. We have estimated $\Sigma_{\widetilde{y}}(h)$ by

$$\widetilde{\Sigma}_{\widetilde{y}}(h) = \left(1 + \frac{3n_T}{T}\right) \sum_{i=0}^{h-1} \widehat{\Phi}_i \widehat{\Sigma}_u \widehat{\Phi}_i' + \frac{1}{T}\widehat{G}_y(h), \tag{15.3.5}$$

where the $\widehat{\Phi}_i$'s and $\widehat{\Sigma}_u$ are obtained from the VAR(2) estimates, as in Section 3.5.3, and $\widehat{G}_y(h)/T$ is a term that takes account of the fact that the mean term is estimated in addition to the VAR coefficients. It is the same term that is used if a VAR(2) process with true order $p = 2$ is assumed and the model is estimated in mean-adjusted form (see Problem 3.9).

Table 15.1. Interval forecasts from a VAR(2) model for the investment/income/consumption example series based on different asymptotic theories

			95% interval forecasts	
variable	forecast horizon	point forecast	based on known order assumption	based on infinite order assumption
investment	1	−.010	[−.105, .085]	[−.105. .085]
	2	.012	[−.087, .110]	[−.088, .112]
	3	.022	[−.075, .119]	[−.078, .122]
	4	.013	[−.084. .111]	[−.088, .114]
income	1	.020	[−.004, .044]	[−.004, .044]
	2	.020	[−.004, .045]	[−.005, .045]
	3	.017	[−.007, .042]	[−.008, .042]
	4	.021	[−.004, .045]	[−.005, .047]
consumption	1	.022	[.002, .041]	[.002, .041]
	2	.015	[−.005, .035]	[−.005, .035]
	3	.020	[−.002, .042]	[−.002, .042]
	4	.019	[−.003, .041]	[−.003, .041]

We have used the approximate forecast MSEs from (15.3.5) to set up forecast intervals under Gaussian assumptions and give them in Table 15.1. For comparison purposes we also give forecast intervals obtained from a VAR(2) process in mean-adjusted form based on the asymptotic theory of Chapter 3, assuming that the true order is $p = 2$. As we know from Remark 3 in Section 15.3.1, the 1-step forecast MSEs are the same under the two competing assumptions. For larger forecast horizons, most of the intervals based on the infinite order assumption become slightly wider than those based on the known finite order assumption, as expected on the basis of Remark 2 in Section 15.3.2. For our sample size, the differences are quite small, though.

Which of the two sets of forecast intervals should we use in practice? This question is difficult to answer. Assuming a known finite VAR order is, of course, more restrictive and less realistic than the assumption of an unknown and possibly infinite order. The additional uncertainty introduced by the latter assumption is reflected in the wider forecast intervals. It may be worth noting, however, that such a result is not necessarily obtained in all practical situations. In other words, there may be time series and generation processes for which the infinite order assumption actually leads to smaller forecast intervals than the assumption of a known finite VAR order (see Problem 15.2).

Under both sets of assumptions, the MSE approximations are derived from asymptotic theory and little is known about the small sample quality of these approximations. Both approaches are based on a set of assumptions that may not hold in practice. Notably the stationarity and normality assumptions may be doubtful in many practical situations. Given all these reservations, there is still one argument in favor of the present approach, assuming a potentially infinite VAR order. For $h > 1$, the MSE approximation in (15.3.3) is generally simpler to compute than the one obtained in Chapter 3.

15.4 Impulse Response Analysis and Forecast Error Variance Decompositions

15.4.1 Asymptotic Theory

For a researcher who does not know the true structure of the data generating process, it is possible to base an impulse response analysis or forecast error variance decomposition on an approximating finite order VAR process. Given the results of Section 15.2, we can now study the consequences of such an approach. As in Sections 2.3.2 and 2.3.3, the quantities of interest here are the forecast error impulse responses,

$$\Phi_i = \sum_{j=1}^{i} \Phi_{i-j} \Pi_j, \quad i = 1, 2, \dots, \quad \Phi_0 = I_K,$$

the accumulated forecast error impulse responses,

$$\Psi_m = \sum_{i=0}^{m} \Phi_i, \quad m = 0, 1, \dots,$$

the responses to orthogonalized impulses,

$$\Theta_i = \Phi_i P, \quad i = 0, 1, \dots,$$

where P is the lower triangular matrix obtained by a Choleski decomposition of Σ_u, the accumulated orthogonalized impulse responses,

$$\Xi_m = \sum_{i=0}^{m} \Theta_i, \quad m = 0, 1, \dots,$$

and the forecast error variance components,

$$\omega_{jk,h} = \sum_{i=0}^{h-1} (e_j' \Theta_i e_k)^2 / \mathrm{MSE}_j(h), \quad h = 1, 2, \dots,$$

where e_k is the k-th column of I_K and

$$\mathrm{MSE}_j(h) = \sum_{i=0}^{h-1} e_j' \Phi_i \Sigma_u \Phi_i' e_j$$

is the j-th diagonal element of the MSE matrix, $\Sigma_y(h)$, of an h-step forecast.

Estimators of these quantities are obtained from the $\widehat{\Pi}_i(n_T)$ and $\widetilde{\Sigma}_u(n_T)$ in the obvious way. For instance, estimators for the Φ_i's are obtained recursively as

$$\widetilde{\Phi}_i(n_T) = \sum_{j=1}^{i} \widetilde{\Phi}_{i-j}(n_T)\widehat{\Pi}_j(n_T), \quad i = 1, 2, \ldots,$$

with $\widetilde{\Phi}_0(n_T) = I_K$, and

$$\widetilde{\Theta}_i(n_T) = \widetilde{\Phi}_i(n_T)\widetilde{P}(n_T), \quad i = 0, 1, \ldots,$$

are estimators of the Θ_i's. Here $\widetilde{P}(n_T)$ is the unique lower triangular matrix with positive main diagonal for which

$$\widetilde{P}(n_T)\widetilde{P}(n_T)' = \widetilde{\Sigma}_u(n_T).$$

The asymptotic distributions of all the estimators are given in the next proposition. Proofs, based on Propositions 15.1 and 15.2, are given by Lütkepohl (1988a) and Lütkepohl & Poskitt (1991).

Proposition 15.4 (*Asymptotic Distributions of Impulse Responses*)
Under the conditions of Proposition 15.2, the impulse responses and forecast error variance components have the following asymptotic normal distributions:

$$\sqrt{T}\ \mathrm{vec}(\widetilde{\Phi}_i(n_T) - \Phi_i) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_u^{-1} \otimes \sum_{j=0}^{i-1} \Phi_j \Sigma_u \Phi_j'\right), \quad i = 1, 2, \ldots; \tag{15.4.1}$$

$$\sqrt{T}\ \mathrm{vec}(\widetilde{\Psi}_m(n_T) - \Psi_m) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_u^{-1} \otimes \sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{j=0}^{l-1} \Phi_j \Sigma_u \Phi_{k-l+j}'\right), \tag{15.4.2}$$

$m = 1, 2, \ldots$, with $\Phi_j := 0$ for $j < 0$;

$$\sqrt{T}\ \mathrm{vec}(\widetilde{\Theta}_i(n_T) - \Theta_i) \xrightarrow{d} \mathcal{N}(0, \Omega_\theta(i)), \quad i = 0, 1, \ldots, \tag{15.4.3}$$

where

$$\Omega_\theta(i) = \left(I_K \otimes \sum_{j=0}^{i-1} \Phi_j \Sigma_u \Phi_j' \right) + (I_K \otimes \Phi_i) H \Sigma_{\tilde{\boldsymbol{\sigma}}} H' (I_K \otimes \Phi_i'),$$

$H = \mathbf{L}_K' [\mathbf{L}_K (I_{K^2} + \mathbf{K}_{KK})(P \otimes I_K) \mathbf{L}_K']^{-1}$,

$\mathbf{L}_K$ is the $(\frac{1}{2}K(K+1) \times K^2)$ elimination matrix,

$\mathbf{K}_{KK}$ is the $(K^2 \times K^2)$ commutation matrix,

and $\Sigma_{\tilde{\boldsymbol{\sigma}}}$ is the asymptotic covariance matrix of $\sqrt{T}\ \mathrm{vech}(T^{-1}\sum_{t=1}^T u_t u_t' - \Sigma_u)$;

$$\sqrt{T}\ \mathrm{vec}(\widetilde{\Xi}_m(n_T) - \Xi_m) \xrightarrow{d} \mathcal{N}(0, \Omega_\xi(m)), \quad m = 1, 2, \dots, \tag{15.4.4}$$

where

$$\Omega_\xi(m) = \sum_{k=0}^{m} \sum_{l=0}^{m} \left[I_K \otimes \sum_{j=0}^{l-1} \Phi_j \Sigma_u \Phi_{k-l+j}' + (I_K \otimes \Phi_l) H \Sigma_{\tilde{\boldsymbol{\sigma}}} H' (I_K \otimes \Phi_k') \right]$$

with $\Phi_j := 0$ for $j < 0$;

$$\sqrt{T}(\widetilde{\omega}_{jk,h}(n_T) - \omega_{jk,h}) \xrightarrow{d} \mathcal{N}(0, \sigma_{jk,h}^2), \quad h = 1, 2, \dots, \quad j, k = 1, \dots, K, \tag{15.4.5}$$

where

$$\sigma_{jk,h}^2 = \sum_{l=0}^{h-1} \sum_{m=0}^{h-1} g_{jk,h}(l) \left[I_K \otimes \sum_{i=0}^{m-1} \Phi_i \Sigma_u \Phi_{l-m+i}' \right.$$
$$\left. + \ (I_K \otimes \Phi_m) H \Sigma_{\tilde{\boldsymbol{\sigma}}} H' (I_K \otimes \Phi_l') \right] g_{jk,h}(m)'$$

with

$$\begin{aligned} g_{jk,h}(m) \ &= \ 2\Big[(e_k' \otimes e_j')(e_j' \Theta_m e_k) \mathrm{MSE}_j(h) \\ &\qquad - \ (e_j' \Theta_m \otimes e_j') \sum_{i=0}^{h-1} (e_j' \Theta_i e_k)^2 \ \Big] \Big/ \ \mathrm{MSE}_j(h)^2. \end{aligned}$$

■

Remark 1 In the proposition, it is ignored that $\sigma_{jk,h}^2$ may be zero, in which case the asymptotic normal distribution is degenerate. In particular, $\sigma_{jk,h}^2 = 0$ if $\omega_{jk,h} = 0$. This result is easily seen by noting that $\omega_{jk,h}$ is zero if and only if $\theta_{jk,0} = \cdots = \theta_{jk,h-1} = 0$, where $\theta_{jk,m}$ is the jk-th element of Θ_m. Thus, the asymptotic distribution in (15.4.5) is not immediately useful for testing

$$H_0 : \omega_{jk,h} = 0 \quad \text{against} \quad H_1 : \omega_{jk,h} \neq 0, \tag{15.4.6}$$

which is a set of hypotheses of particular interest in practice. The significance of $\omega_{jk,h}$ may be checked, however, by testing $\theta_{jk,0} = \cdots = \theta_{jk,h-1} = 0$. Using a minor generalization of (15.4.3), this hypothesis can be tested (see Lütkepohl & Poskitt (1991)). ■

Remark 2 In sharp contrast to the case where the VAR order is assumed to be known and finite (see Proposition 3.6), the asymptotic variances of all impulse responses are nonzero in the present case. Another difference between the finite and infinite order VAR cases is that in the former the asymptotic standard errors of the Φ_i and Θ_i go to zero as i increases, while the covariance matrix in (15.4.1) is a nondecreasing function of i and the covariance matrix in (15.4.3) is bounded away from zero, for $i > 0$. ■

Remark 3 For $i = 1$, the asymptotic covariance matrix of $\widetilde{\Phi}_1(n_T)$ in (15.4.1) is $\Sigma_u^{-1} \otimes \Sigma_u$. It can be shown that the same asymptotic covariance matrix is obtained for $\widehat{\Phi}_1$ from Proposition 3.6, if a VAR(n) process is fitted with n greater than the true order p (see Lütkepohl (1988a)). A similar result is obtained for $\widetilde{\Theta}_i(n_T)$ and $\widehat{\Theta}_i$ for $i = 0, 1$ (see Problem 15.4). ■

Remark 4 The results in Proposition 15.4 can also be used to construct tests for zero impulse responses. This case was considered by Lütkepohl (1996b). ■

Remark 5 Although forecast error and orthogonalized impulse responses are considered only in Proposition 15.4, similar results can also be obtained for the structural impulse responses discussed in Chapter 9. ■

15.4.2 An Example

To illustrate the consequences of the finite and infinite VAR order assumptions, we use again the VAR(2) model for the investment/income/consumption data. Of course, the same estimated impulse responses are obtained as in Section 3.7. (The intercept form of the model is used now.) The standard errors are different, however. In Figures 15.1 and 15.2, consumption responses to income impulses are depicted and the two-standard error bounds obtained from both sets of assumptions are shown. In both figures, the two-standard error bounds based on Proposition 3.6 decline almost to zero for longer lags while the two-standard error bounds from Proposition 15.4 are seen to grow with the time lag. This behavior reflects the additional estimation uncertainty that results from assuming that the VAR order goes to infinity with the sample size. Thereby more and more parameters are estimated as the sample size gets large.

In Table 15.2, forecast error variance decompositions of the system are shown. Again most standard errors based on the infinite VAR assumption are slightly larger than those from Chapter 3, which are also given in the table. Although it is tempting to use the estimated standard errors in checking the

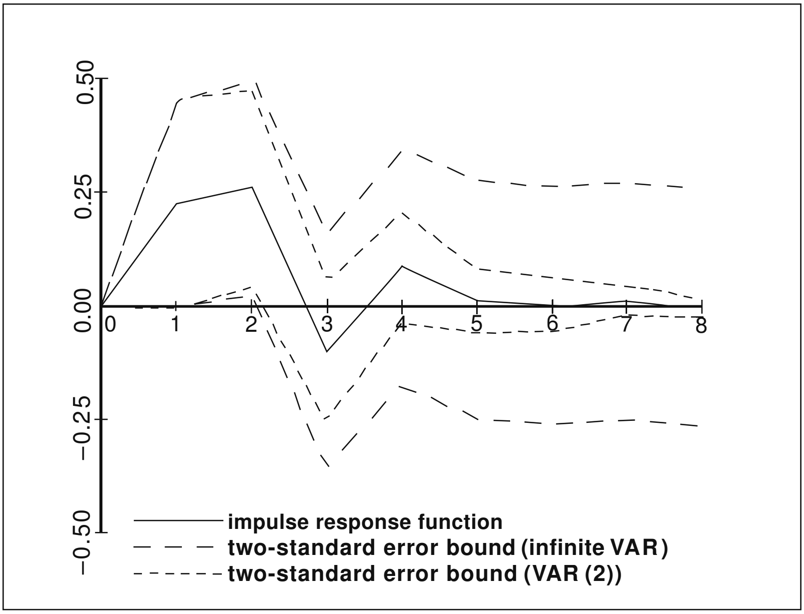

Fig. 15.1. Estimated responses of consumption to a forecast error impulse in income with two-standard error bounds based on finite and infinite VAR order assumptions.

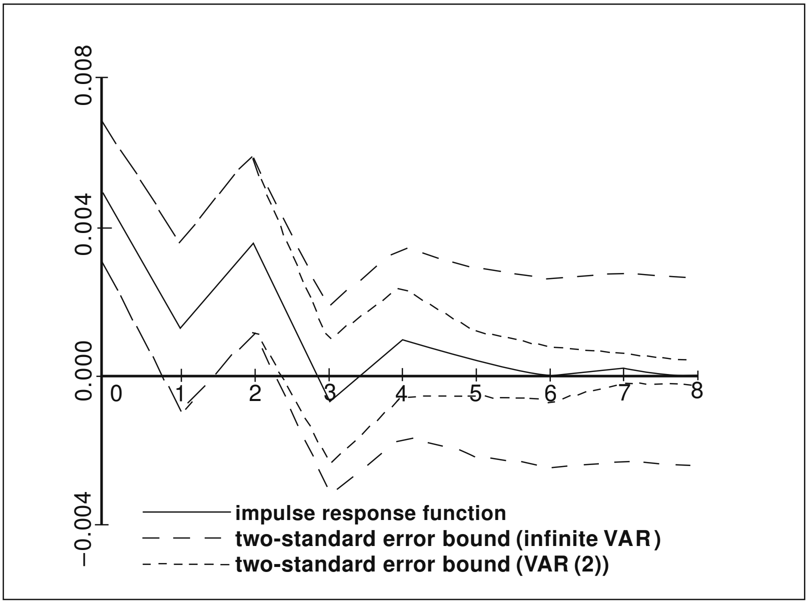

Fig. 15.2. Estimated responses of consumption to an orthogonalized impulse in income with two-standard error bounds based on finite and infinite VAR order assumptions.

Table 15.2. Forecast error variance decompositions of the investment/income/consumption system with standard errors from two different asymptotic theories

	forecast	proportions of forecast error variance, h periods ahead, accounted for by innovations in[a]		
forecast error in	horizon h	investment $\omega_{j1,h}$	income $\omega_{j2,h}$	consumption $\omega_{j3,h}$
investment	1	1.000(.000)[.000]	.000(.000)[.000]	.000(.000)[.000]
($j = 1$)	2	.960(.042)[.044]	.018(.030)[.030]	.023(.031)[.033]
	3	.946(.042)[.045]	.028(.033)[.033]	.026(.029)[.032]
	4	.941(.045)[.048]	.029(.031)[.032]	.030(.032)[.036]
	8	.938(.048)[.050]	.031(.032)[.034]	.032(.035)[.039]
income	1	.018(.031)[.031]	.983(.031)[.031]	.000(.000)[.000]
($j = 2$)	2	.060(.054)[.053]	.908(.063)[.064]	.032(.037)[.040]
	3	.070(.057)[.058]	.896(.066)[.068]	.035(.039)[.041]
	4	.068(.056)[.057]	.892(.067)[.069]	.039(.041)[.045]
	8	.069(.057)[.058]	.891(.068)[.070]	.040(.041)[.045]
consumption	1	.080(.061)[.061]	.273(.086)[.086]	.647(.090)[.090]
($j = 3$)	2	.077(.059)[.059]	.274(.082)[.082]	.649(.088)[.088]
	3	.130(.080)[.080]	.334(.089)[.091]	.537(.091)[.092]
	4	.129(.079)[.079]	.335(.088)[.090]	.536(.089)[.091]
	8	.129(.080)[.081]	.340(.089)[.092]	.532(.091)[.093]

[a]Estimated standard error based on a finite known VAR order assumption in parentheses and estimated standard error based on an infinite VAR order assumption in brackets.

significance of individual forecast error variance components, we know from Remark 1 of Section 15.4.1 that they are not useful for that purpose because the asymptotic standard errors from Proposition 15.4 corresponding to zero forecast error variance components are zero.

15.5 Cointegrated Infinite Order VARs

In Chapter 6, it was discussed that assuming a fixed finite starting date is advantageous if integrated variables are considered. Therefore, some modifications will be necessary in defining infinite order processes for integrated variables. Details of the model setup will be given in Section 15.5.1. The properties of estimators of the parameters of such models are considered in Section 15.5.2, and testing for the cointegrating rank will be discussed in Section 15.5.3.

15.5.1 The Model Setup

The general framework presented in the following is that of Saikkonen (1992) and Saikkonen & Lütkepohl (1996). Given a K-dimensional system of time series variables y_t with cointegrating rank r, we assume that the variables are arranged such that for $t = 1, 2, \ldots,$

$$\begin{aligned} y_t^{(1)} &= -\beta'_{(K-r)} y_t^{(2)} + z_t^{(1)}, \\ \Delta y_t^{(2)} &= z_t^{(2)}, \end{aligned} \tag{15.5.1}$$

where $y_t^{(1)}$ and $y_t^{(2)}$ are $(r \times 1)$ and $((K-r) \times 1)$, respectively, such that

$$y_t = \begin{bmatrix} y_t^{(1)} \\ y_t^{(2)} \end{bmatrix},$$

as in the triangular representation discussed in Section 6.3 (see (6.3.10)). Hence, $\beta_{(K-r)}$ is $((K-r) \times r)$ such that

$$\begin{bmatrix} I_r \\ \beta_{(K-r)} \end{bmatrix}$$

is the cointegration matrix and

$$z_t = \begin{bmatrix} z_t^{(1)} \\ z_t^{(2)} \end{bmatrix}$$

is a strictly stationary process with $E(z_t) = 0$ and positive definite covariance matrix $\Sigma_z = E(z_t z_t')$. As a further technical condition which is needed in some proofs, we also assume that z_t has a continuous spectral density matrix which is positive definite at zero frequency. For a discussion of spectral density matrices of vector processes see, e.g., Fuller (1976, Section 4.4). The initial vector y_0 is assumed to be such that the process Δy_t is stationary.

In matrix form, the process y_t may be written as

$$\begin{bmatrix} I_r & \beta'_{(K-r)} \\ 0 & I_{K-r} \end{bmatrix} y_t = \begin{bmatrix} 0 & 0 \\ 0 & I_{K-r} \end{bmatrix} y_{t-1} + z_t. \tag{15.5.2}$$

Multiplying by the inverse of the left-hand matrix,

$$\begin{bmatrix} I_r & -\beta'_{(K-r)} \\ 0 & I_{K-r} \end{bmatrix},$$

subtracting y_{t-1} on both sides of the equation and rearranging terms gives

$$\Delta y_t = -\begin{bmatrix} I_r & \beta'_{(K-r)} \\ 0 & 0 \end{bmatrix} y_{t-1} + v_t = -\begin{bmatrix} \beta' \\ 0 \end{bmatrix} y_{t-1} + v_t, \tag{15.5.3}$$

where $\beta' = [I_r : \beta'_{(K-r)}]$ and

$$v_t = \begin{bmatrix} I_r & -\beta'_{(K-r)} \\ 0 & I_{K-r} \end{bmatrix} z_t$$

is a stationary process. It is assumed to have an infinite order VAR representation,

$$v_t = \sum_{j=1}^{\infty} G_j v_{t-j} + u_t, \quad t \in \mathbb{Z}, \tag{15.5.4}$$

where u_t is again standard white noise. Notice that, due to the stationarity of v_t, there is no problem in defining it for all integers t. Moreover, because the process v_t is stationary, it also has an MA representation for which we could make similar assumptions as in (15.2.4). We do not need that representation here, however, and therefore we formulate the required assumptions directly for the VAR coefficients. In particular, the G_j's are assumed to satisfy

$$\det\left(I_K - \sum_{j=1}^{\infty} G_j z^j\right) \neq 0 \text{ for } |z| \leq 1 \quad \text{and} \quad \sum_{j=1}^{\infty} j\|G_j\| < \infty. \tag{15.5.5}$$

This condition imposes weak restrictions on the autocorrelation structure of the process v_t and is, for example, satisfied for VARMA processes. From the previous assumptions, it follows that if the infinite order VAR is approximated by a finite order process, the approximation error gets sufficiently small for our purposes, if the order of the approximating process is chosen as in (15.2.5) with $\sqrt{T}\sum_{i=n_T+1}^{\infty} \|G_j\| \to 0$ as $T \to \infty$.

Defining

$$G_j^* := -(G_{j+1} + \cdots + G_n), \quad j = 0, 1, \ldots, n-1,$$

and

$$G_{n-1}^*(L) := \sum_{j=0}^{n-1} G_j^* L^j,$$

it follows that

$$G_n(L) := I_K - \sum_{j=1}^{n} G_j L^j = G_n(1) - G_{n-1}^*(L)(1-L) \tag{15.5.6}$$

(see Problem 15.6). Multiplying (15.5.3) by $G_n(L)$ and rearranging terms gives the VECM representation

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{j=1}^{n} \mathbf{\Gamma}_j \Delta y_{t-j} + e_t, \quad t = n+1, n+2, \ldots, \tag{15.5.7}$$

where

$$\alpha := -\left(I_K - \sum_{j=1}^{n} G_j\right)\begin{bmatrix} I_r \\ 0 \end{bmatrix},$$

$$e_t := u_t + \sum_{j=n+1}^{\infty} G_j v_{t-j}$$

and

$$\begin{aligned}\sum_{j=1}^{n} \mathbf{\Gamma}_j L^j &= \sum_{j=1}^{n} G_j L^j + G^*_{n-1}(L)\begin{bmatrix} \beta' \\ 0 \end{bmatrix} L \\ &= \sum_{j=1}^{n}\left(G_j + G^*_{j-1}\begin{bmatrix} \beta' \\ 0 \end{bmatrix}\right) L^j.\end{aligned}$$

Hence,

$$\mathbf{\Gamma}_j = G_j - (G_j + \cdots + G_n)\begin{bmatrix} \beta' \\ 0 \end{bmatrix}, \quad j = 1, \ldots, n$$

(see Problem 15.7). Although this fact is not specifically indicated, the coefficient matrices α and $\mathbf{\Gamma}_j$, $j = 1, \ldots, n$, depend on n. In particular, $\mathbf{\Gamma}_n = [0 : \mathbf{\Gamma}_{n2}]$, where $\mathbf{\Gamma}_{n2}$ is $(K \times (K - r))$. It can be shown that assumption (15.5.5) implies that the $\mathbf{\Gamma}_i$'s are absolutely summable, that is, $\lim_{n\to\infty} \sum_{j=1}^{n} \|\mathbf{\Gamma}_j\|$ exists, and the process y_t is well-defined (Phillips & Solo (1992, 2.1 Lemma)).

Rearranging terms, the VECM (15.5.7) can also be rewritten in levels VAR form as

$$y_t = \sum_{j=1}^{n+1} \Pi_j y_{t-j} + e_t, \quad t = n+1, n+2, \ldots, \tag{15.5.8}$$

where

$$\begin{aligned}\Pi_1 &= I_K + \alpha\beta' + \mathbf{\Gamma}_1 = I_K + G_1 - \begin{bmatrix} \beta' \\ 0 \end{bmatrix}, \\ \Pi_j &= \mathbf{\Gamma}_j - \mathbf{\Gamma}_{j-1} = G_j - [0 : G_{j-1,1}\beta'_{(K-r)} - G_{j-1,2}], \quad j = 2, \ldots, n, \\ \Pi_{n+1} &= -\mathbf{\Gamma}_n.\end{aligned}$$

Here $G_{j-1,1}$ and $G_{j-1,2}$ are submatrices of G_{j-1} consisting of the first r and last $K - r$ columns, respectively. Thus, although the $\mathbf{\Gamma}_j$ depend on n, the same is not true for the Π_j, except for Π_{n+1}. In the following subsection, the asymptotic properties of the LS estimators of the VECM and the levels VAR representations will be considered.

15.5.2 Estimation

Suppose a levels VAR($n_T + 1$) model of the form (15.5.8) is estimated by multivariate LS based on a sample of size T. Notice that the order of the model now depends explicitly on the sample size T, as in the case where stationary processes were approximated by finite order VARs, discussed earlier in this chapter. We assume again that the VAR order goes to infinity with the sample size although at a smaller rate than T. The following proposition, which is similar to Theorem 2 of Saikkonen & Lütkepohl (1996), gives the details. In stating the proposition, the LS estimators are denoted by $\widehat{\Pi}_j$, $\Pi(n) := [\Pi_1, \dots, \Pi_n]$, and $\widehat{\Pi}(n) := [\widehat{\Pi}_1, \dots, \widehat{\Pi}_n]$, as before. Now we can present the result.

Proposition 15.5 (*Asymptotic Distribution of the LS Estimator of the VAR Coefficients*)
Suppose that finite order VAR($n_T + 1$) processes are fitted by multivariate LS to a multiple time series generated by the process specified in Section 15.5.1 and assume that the order n_T depends on the sample size T such that

$$n_T \to \infty, \quad n_T^3/T \to 0, \quad \text{and} \quad \sqrt{T} \sum_{i=n_T+1}^{\infty} \|G_i\| \to 0 \quad \text{as} \quad T \to \infty. \tag{15.5.9}$$

Furthermore, let c_1, c_2 be fixed constants and $\mathbf{f}(n)$ a sequence of nonzero $((Kr + K^2 n) \times 1)$ vectors such that

$$0 < c_1 \le \mathbf{f}(n)'\mathbf{f}(n) \le c_2 < \infty \quad \text{for } n = 1, 2, \dots .$$

Then

$$\frac{\sqrt{T - n_T}\, \mathbf{f}(n_T)'[\widehat{\boldsymbol{\pi}}(n_T) - \boldsymbol{\pi}(n_T)]}{[\mathbf{f}(n_T)'(H_{n_T}' \Gamma_{n_T,VECM}^{-1} H_{n_T} \otimes \Sigma_u)\mathbf{f}(n_T)]^{1/2}} \xrightarrow{d} \mathcal{N}(0, 1), \tag{15.5.10}$$

where $\boldsymbol{\pi}(n) := \operatorname{vec} \Pi(n)$ and $\widehat{\boldsymbol{\pi}}(n) := \operatorname{vec} \widehat{\Pi}(n)$, as in Section 15.2, H_{n_T} is a $((r + Kn_T) \times Kn_T)$ matrix defined such that

$$[\Pi_1 : \cdots : \Pi_{n_T}] = [\boldsymbol{\alpha} : \boldsymbol{\Gamma}_1 : \cdots : \boldsymbol{\Gamma}_{n_T}]H_{n_T} + [I_K : 0 : \cdots : 0]$$

and

$$\Gamma_{n_T,VECM} := E\left(\begin{bmatrix} u_{t-1}^{(1)} \\ \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-n_T} \end{bmatrix} [u_{t-1}^{(1)\prime}, \Delta y_{t-1}', \dots, \Delta y_{t-n_T}']\right). \tag{15.5.11}$$

Here $u_{t-1}^{(1)}$ denotes the vector of the first r components of u_{t-1}. ■

This proposition can be proven analogously to Theorem 2 in Saikkonen & Lütkepohl (1996). Clearly, the proposition is similar to Proposition 15.1. Note, however, that in the present proposition, only the first n_T coefficient matrices are considered, although a VAR(n_T+1) process is fitted to the data. Dropping the last lag in deriving the asymptotic distribution of the estimators ensures that standard asymptotic properties are obtained. This devise was also used in Section 7.6.3 in deriving a Wald test for Granger-causality in a finite order cointegrated VAR context.

Consider now the VECM with n_T lagged differences,

$$\Delta y_t = \mathbf{\Pi} y_{t-1} + \sum_{j=1}^{n_T} \mathbf{\Gamma}_j \Delta y_{t-j} + e_t. \tag{15.5.12}$$

Suppose that the model is also estimated by multivariate LS based on a sample of size T. The estimators are denoted by $\widehat{\mathbf{\Pi}}$ and $\widehat{\mathbf{\Gamma}}_j$ and the residuals are signified as $\widehat{u}_t(n_T)$. Using this notation,

$$\widetilde{\Sigma}_u = \frac{1}{T - n_T - 1} \sum_{t=n_T+2}^{T} \widehat{u}_t(n_T)\widehat{u}_t(n_T)' \tag{15.5.13}$$

is an estimator of the white noise covariance matrix Σ_u. The loading matrix α may be estimated as $\widehat{\alpha} = \widehat{\mathbf{\Pi}}_1$, where the latter matrix consists of the first r columns of $\widehat{\mathbf{\Pi}}$, as in the EGLS procedure presented in Section 7.2.2. As in that procedure, the matrix $\beta'_{(K-r)}$ may be estimated as

$$\widehat{\widehat{\beta}}'_{(K-r)} = (\widehat{\alpha}' \widetilde{\Sigma}_u^{-1} \widehat{\alpha})^{-1} \widehat{\alpha}' \widetilde{\Sigma}_u^{-1} \widehat{\mathbf{\Pi}}_2, \tag{15.5.14}$$

where $\widehat{\mathbf{\Pi}}_2$ consists of the last $K-r$ columns of $\widehat{\mathbf{\Pi}}$ (see Remark 4 of Section 7.2.2). The next proposition summarizes the asymptotic properties of the estimators. Proofs can be found in Saikkonen (1992) and Saikkonen & Lütkepohl (1996).

Proposition 15.6 (*Asymptotic Distribution of VECM Estimators*)
Under the conditions of Proposition 15.5,

$$T(\widehat{\widehat{\beta}}'_{(K-r)} - \beta'_{(K-r)}) \xrightarrow{d} \left(\int_0^1 \mathbf{W}^{\#}_{K-r} d\mathbf{W}^{\#\prime}_r \right)' \left(\int_0^1 \mathbf{W}^{\#}_{K-r} \mathbf{W}^{\#\prime}_{K-r} ds \right)^{-1}, \tag{15.5.15}$$

where $\mathbf{W}^{\#}_{K-r}$ and $\mathbf{W}^{\#}_r$ are independent $(K-r)$- and r-dimensional Wiener processes, respectively, as in Proposition 7.2. Furthermore,

$$\frac{\sqrt{T - n_T}\, \mathbf{f}(n_T)'[\widehat{\gamma}(n_T) - \gamma(n_T)]}{[\mathbf{f}(n_T)'(\Gamma^{-1}_{n_T,VECM} \otimes \Sigma_u)\mathbf{f}(n_T)]^{1/2}} \xrightarrow{d} \mathcal{N}(0,1), \tag{15.5.16}$$

where $\gamma(n) := \mathrm{vec}[\alpha : \mathbf{\Gamma}_1 : \cdots : \mathbf{\Gamma}_n]$ and $\widehat{\gamma}(n) := \mathrm{vec}[\widehat{\alpha} : \widehat{\mathbf{\Gamma}}_1 : \cdots : \widehat{\mathbf{\Gamma}}_n]$. ■

Notice that the asymptotic distribution of the cointegration parameters in (15.5.15) is the same as that of the corresponding ML estimator for Gaussian finite order VAR processes, as discussed in Section 7.2 (see in particular Proposition 7.2 and Remark 3 for Proposition 7.4). The loading and short-run parameters have asymptotic properties similar to those of finite order processes as well. Their asymptotic properties are the same that would be obtained if the true $\boldsymbol{\beta}$ matrix were known and used in the estimation procedure.

Moreover, Saikkonen & Lütkepohl (1996) showed that the white noise covariance matrix estimator $\widetilde{\Sigma}_u$ also has similar asymptotic properties as in the finite order case. Furthermore, they stated slightly more general versions of Propositions 15.5 and 15.6 and discussed how the results can be used in testing hypotheses of parameter restrictions. In particular, they considered the case of testing for Granger-causality. They also discussed adding an intercept term to the model. Saikkonen & Lütkepohl (2000a) presented extensions which can be used in deriving, for example, asymptotic properties of impulse responses in the present framework.

In practice, the cointegrating rank is usually unknown and has to be determined from the given multiple time series. How to do so in the present framework of an infinite order process is discussed next.

15.5.3 Testing for the Cointegrating Rank

In Section 8.2, we have discussed testing the cointegrating rank of a finite order VAR process by considering pairs of hypotheses of the form

$$H_0 : \text{rk}(\mathbf{\Pi}) = r_0 \quad \text{against} \quad H_1 : r_0 < \text{rk}(\mathbf{\Pi}) \le r_1. \tag{15.5.17}$$

In particular, the cases $r_1 = r_0 + 1$ and $r_1 = K$ were discussed and suitable likelihood ratio tests were introduced. Suppose now that the test statistics are computed in precisely the same way as in Section 8.2.1, based on the VECM (15.5.12) with n_T lagged differences of y_t. In other words, we compute the statistic as if (15.5.12) were a Gaussian process with lag order n_T. To emphasize the dependence on the lag order, we denote the test statistic corresponding to the pair of hypotheses in (15.5.17) by $\lambda_{LR}^{(n_T)}(r_0, r_1)$. Lütkepohl & Saikkonen (1999b) proved the following result.

Proposition 15.7 (*Asymptotic Distributions of Tests for the Cointegrating Rank*)
Suppose y_t is generated by an infinite order process as described in Section 15.5.1. Moreover, suppose that

$$n_T \to \infty \quad \text{and} \quad n_T^3/T \to 0 \quad \text{as} \quad T \to \infty. \tag{15.5.18}$$

Then $\lambda_{LR}^{(n_T)}(r_0, r_0 + 1)$ and $\lambda_{LR}^{(n_T)}(r_0, K)$ have the same limiting null distributions as for a Gaussian finite order process given in Proposition 8.2. ∎

Notice that in this proposition we just have an upper bound for the rate at which the lag order n_T has to go to infinity. No lower bound for the rate of divergence is needed. In fact, Lütkepohl & Saikkonen (1999b) considered also processes with nonzero mean term and, in addition, they treated the case where the lag order is chosen by some model selection procedure instead of a deterministic rule derived from (15.5.18). In summary, these results show that as far as asymptotic theory is concerned, the cointegrating rank of an $I(1)$ process may be chosen on the basis of a finite order approximation rather than a correctly specified model. This result is not only important because, in practice, models are usually just approximations to the true DGP and, hence, allowing explicitly for some approximation error is more realistic, it is also important because we have proposed this approach for choosing the cointegrating rank of a VARMA process in Chapter 14, Section 14.4.2.

15.6 Exercises

Problem 15.1
For the invertible MA(1) process $y_t = u_t + Mu_{t-1}$ and $n = 1, 2$, determine the matrix Γ_n defined in (15.2.7).

Problem 15.2
Suppose the true data generation mechanism is a univariate AR(1) process, $y_t = \alpha y_{t-1} + u_t$. Assume that a univariate AR(1) is indeed fitted to the data and compare the resulting approximate forecast MSEs $\Sigma_{\widehat{y}}(h)$ (given in Section 3.5) and $\Sigma_{\widetilde{y}}(h)$ (given in Section 15.3.1) for $h = 1, 2, \ldots$. (Hint: See Lütkepohl (1987, pp. 76, 77).)

Problem 15.3
Suppose the true data generation process is an invertible MA(1), as in Problem 15.1. Write down explicit expressions for the asymptotic covariance matrices of $\widetilde{\Phi}_i(n_T)$, $\widetilde{\Theta}_i(n_T)$, $i = 1, 2$, and of $\widetilde{\Psi}_m(n_T)$, $\widetilde{\Xi}_m(n_T)$, $m = 1, 2$.

Problem 15.4
Let $\widetilde{\Theta}_i(n_T)$ and $\widehat{\Theta}_i$ be estimators of the orthogonalized impulse responses Θ_i obtained under the conditions of Propositions 15.4 and 3.6, respectively. If the true data generation mechanism is a finite order VAR(p) process and the actual process fitted to the data has order $n_T > p$, show that the asymptotic covariance matrices in (15.4.3) and (3.7.8) are identical for $i = 0, 1$.

Problem 15.5
Consider the investment/income/consumption system of Section 15.3.2 and fit a VAR(4) process to the data.

(a) Determine 95% interval forecasts for all three variables and forecast horizons $h = 1, 2, 3$ under the assumption of a known true VAR order of $p = 4$ and under the assumption of an infinite order true generation process.

(b) Determine Φ_i and Θ_i impulse responses and their asymptotic standard errors for $i = 1, 2, 3, 4$ under both the assumption of a finite and an infinite true VAR order. Compare the estimated standard errors obtained under the two alternative scenarios for all variables.

Problem 15.6
Show that the relation in (15.5.6) holds.

Problem 15.7
Derive the model representation (15.5.7). (Hint: See Saikkonen (1992, Section 2).)