parameters. The bootstrap error terms are obtained by resampling the row vectors $\tilde{\boldsymbol{U}}_t$, where $\tilde{\boldsymbol{U}}_t$ is equal to $(n/(n-1-gp))^{1/2}$ times the row vector $\hat{\boldsymbol{U}}_t$ of OLS residuals, and actual pre-sample values of $\boldsymbol{Y}_t$ are used to start the recursive process of generating the bootstrap data. Limited simulation evidence suggests that this procedure yields much more accurate $P$ values for tests based on (13.91) than using the $\chi^2(g^2)$ distribution.

If we wish to construct confidence intervals for, or test hypotheses about, individual parameters in a VAR, we can use the OLS standard errors, which are asymptotically valid. Similarly, if we wish to test hypotheses concerning two or more parameters in a single equation, we can compute Wald tests in the usual way based on the OLS covariance matrix for that equation. However, if we wish to test hypotheses concerning coefficients in two or more equations, we need the covariance matrix of the parameter estimates for the entire system.

We saw in Chapter 12 that the estimated covariance matrix for the feasible GLS estimates of a multivariate regression model is given by expression (12.19), and the one for the ML estimates is given by expression (12.38). These two covariance matrices differ only because they use different estimates of $\boldsymbol{\Sigma}$. As in Section 12.2, we let $\boldsymbol{X}_\bullet \equiv \mathbf{I}_g \otimes \boldsymbol{X}$, which is a $gn \times gk$ matrix. Then, if all the parameters are stacked into a column vector of dimension $gk$, both covariance matrices have the form

$$\big(\boldsymbol{X}_\bullet^\top(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_n)\boldsymbol{X}_\bullet\big)^{-1}.$$

Using the rules for manipulating Kronecker products given in equations (12.08), we see that

$$\big(\boldsymbol{X}_\bullet^\top(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_n)\boldsymbol{X}_\bullet\big)^{-1} = \big((\mathbf{I}_g \otimes \boldsymbol{X}^\top)(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_n)(\mathbf{I}_g \otimes \boldsymbol{X})\big)^{-1} = \hat{\boldsymbol{\Sigma}} \otimes (\boldsymbol{X}^\top\boldsymbol{X})^{-1}.$$

Thus the covariance matrix for all the coefficients of a VAR is easily computed from $\hat{\boldsymbol{\Sigma}}$, which is given in (13.90), and the inverse of the $\boldsymbol{X}^\top\boldsymbol{X}$ matrix.

The idea of using vector autoregressions instead of structural models to model macroeconomic dynamics is often attributed to Sims (1980). Our treatment has been very brief. For a more detailed introductory treatment, with many references, see Lütkepohl (2001). For a review of macroeconomic applications of VARs, see Stock and Watson (2001).

### Granger Causality

One common use of vector autoregressions is to test the hypothesis that one or more of the variables in a VAR do not "Granger cause" the others. The concept of **Granger causality** was developed by Granger (1969). Other, closely related, definitions of causality have been suggested, notably by Sims (1972). Suppose we divide the variables in a VAR into two groups, $\boldsymbol{Y}_{t1}$ and $\boldsymbol{Y}_{t2}$, which are row vectors of dimensions $g_1$ and $g_2$, respectively. Then we may say that $\boldsymbol{Y}_{t2}$ does not Granger cause $\boldsymbol{Y}_{t1}$ if the distribution of $\boldsymbol{Y}_{t1}$, conditional on past