Thus,

$$\begin{aligned}\left(T^{-1}\sum_{t=1}^{T} y_{t-1}y'_{t-1}\right)^{-1} &= Q'\begin{bmatrix} (\Gamma_z^{(1)})^{-1}+o_p(1) & o_p(1) \\ o_p(1) & o_p(1)\end{bmatrix} Q \\ &= \beta(\Gamma_z^{(1)})^{-1}\beta' + o_p(1),\end{aligned}$$

which proves Result 2. ■

Thus, the limiting distribution of $\sqrt{T}\mathrm{vec}(\widehat{\mathbf{\Pi}}-\mathbf{\Pi})$ is singular because $\Gamma_z^{(1)}$ is an $(r\times r)$ matrix. Still, we can use the usual estimator of the covariance matrix based on the regressor matrix. Thus, t-ratios can be set up in the standard way and may have their usual asymptotic standard normal distributions, if a consistent estimator of Σ_u is used. In Result 8, we will see that the usual residual covariance matrix is in fact a consistent estimator for Σ_u, as in the stationary case. On the other hand, it is not difficult to see that the covariance matrix in the limiting distribution (7.1.7) has rank rK. Therefore, setting up a Wald test for more general restrictions may be problematic. As explained in Appendix C.7, a nonsingular weighting matrix is needed for the Wald test to have its usual limiting χ^2-distribution under the null hypothesis. Thus, if we want to test, for example,

$$H_0 : \mathbf{\Pi} = 0 \quad \text{versus} \quad H_1 : \mathbf{\Pi} \neq 0,$$

the corresponding Wald statistic is

$$\lambda_W = T\mathrm{vec}(\widehat{\mathbf{\Pi}})'\left(\left(T^{-1}\sum_{t=1}^{T} y_{t-1}y'_{t-1}\right)\otimes \widehat{\Sigma}_u^{-1}\right)\mathrm{vec}(\widehat{\mathbf{\Pi}}).$$

Under H_0, the arguments in the proof of Result 2 can be used to show that $(T^{-1}\sum_{t=1}^{T} y_{t-1}y'_{t-1})^{-1}$ converges to zero in probability and, hence, the weighting matrix in the Wald statistic diverges. Thus, λ_W will not have an asymptotic $\chi^2(K^2)$-distribution. Therefore, caution is necessary in setting up F-tests, for example. In the nonstationary case, they may not have an asymptotic justification. We will provide more discussion of this problem in Section 7.6 in the context of testing for Granger-causality.

It is interesting to note that the asymptotic distribution in (7.1.7) is the same one that is obtained if the cointegration matrix β is known and only α is estimated by LS. To see this result, we consider the LS estimator

$$\widehat{\alpha} = \left(\sum_{t=1}^{T} \Delta y_t y'_{t-1}\beta\right)\left(\sum_{t=1}^{T} \beta' y_{t-1}y'_{t-1}\beta\right)^{-1}. \tag{7.1.8}$$

This estimator has the following properties.