$$= \left[(\widehat{\alpha}'\widehat{\Sigma}_u^{-1}\widehat{\alpha})^{-1}\widehat{\alpha}'\widehat{\Sigma}_u^{-1} - (\alpha'\Sigma_u^{-1}\alpha)^{-1}\alpha'\Sigma_u^{-1}\right]$$
$$\times \left(T^{-1}\sum_{t=1}^{T} u_t y_{t-1}^{(2)\prime}\right)\left(T^{-2}\sum_{t=1}^{T} y_{t-1}^{(2)} y_{t-1}^{(2)\prime}\right)^{-1}$$
$$+(\widehat{\alpha}'\widehat{\Sigma}_u^{-1}\widehat{\alpha})^{-1}\widehat{\alpha}'\widehat{\Sigma}_u^{-1}$$
$$\times \left(T^{-1}\sum_{t=1}^{T} (u_t^* - u_t) y_{t-1}^{(2)\prime}\right)\left(T^{-2}\sum_{t=1}^{T} y_{t-1}^{(2)} y_{t-1}^{(2)\prime}\right)^{-1}.$$

The term in brackets is $o_p(1)$ because $\widehat{\alpha}$ and $\widehat{\Sigma}_u$ are consistent estimators by assumption. Moreover, $T^{-1}\sum_{t=1}^{T}(u_t^* - u_t)y_{t-1}^{(2)\prime} = o_p(1)$ (see Problem 7.1). Thus, the desired result follows because all other terms converge as established previously. ■

If the process is assumed to be Gaussian, ML estimation may be used alternatively. In case α and Σ_u are known, the ML estimator is identical to the GLS estimator for $\beta'_{(K-r)}$ and, hence, $\widehat{\beta}'_{(K-r)}$ is also the ML estimator. If α and Σ_u are unknown, ML estimation under the constraint $\text{rk}(\mathbf{\Pi}) = r$ may be used. The log-likelihood function is

$$\ln l = -\frac{KT}{2}\ln 2\pi - \frac{T}{2}\ln|\Sigma_u| - \frac{1}{2}\sum_{t=1}^{T}(\Delta y_t - \mathbf{\Pi} y_{t-1})'\Sigma_u^{-1}(\Delta y_t - \mathbf{\Pi} y_{t-1}). \tag{7.1.20}$$

From Chapter 3, we know that maximizing this function is equivalent to minimizing the determinant

$$\left|T^{-1}\sum_{t=1}^{T}(\Delta y_t - \mathbf{\Pi} y_{t-1})(\Delta y_t - \mathbf{\Pi} y_{t-1})'\right|.$$

To impose the rank restriction $\text{rk}(\mathbf{\Pi}) = r$, we write $\mathbf{\Pi} = \alpha\beta'$, where α and β are $(K \times r)$ matrices with rank r. For the moment we do not impose any normalization restrictions and consider minimization of the determinant

$$\left|T^{-1}\sum_{t=1}^{T}(\Delta y_t - \alpha\beta' y_{t-1})(\Delta y_t - \alpha\beta' y_{t-1})'\right|$$

with respect to α and β. This minimization problem is solved in Proposition A.7 in Appendix A.14 and the solution is obtained by considering the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_K$ and the associated orthonormal eigenvectors $v_1, \ldots, v_K$ of the matrix

$$\left(\sum_{t=1}^{T} y_{t-1}y_{t-1}'\right)^{-1/2}\left(\sum_{t=1}^{T} y_{t-1}\Delta y_t'\right)\left(\sum_{t=1}^{T}\Delta y_t \Delta y_t'\right)^{-1}$$