For the actual minimization, it is convenient to work with a transformation of $\boldsymbol{\eta}$. Let $\boldsymbol{S}$ denote any $m \times m$ matrix with the property that $\boldsymbol{S}^\top \boldsymbol{S} = (\boldsymbol{Y}^*_{-p})^\top \boldsymbol{Y}^*_{-p}$, and define the $m \times r$ matrix $\boldsymbol{\zeta}$ as $\boldsymbol{S\eta}$. The ratio (20.41) becomes

$$\frac{\left|\boldsymbol{\zeta}^\top (\boldsymbol{S}^{-1})^\top (\boldsymbol{Y}^*_{-p})^\top \boldsymbol{M}^* \boldsymbol{Y}^*_{-p} \boldsymbol{S}^{-1} \boldsymbol{\zeta}\right|}{\left|\boldsymbol{\zeta}^\top \boldsymbol{\zeta}\right|}. \tag{20.42}$$

Since all that matters is the subspace spanned by the r columns of $\boldsymbol{\zeta}$, we may without loss of generality choose $\boldsymbol{\zeta}$ such that $\boldsymbol{\zeta}^\top \boldsymbol{\zeta} = \mathbf{I}_r$. Let us define the $m \times m$ positive definite matrix $\boldsymbol{A}$ to be the matrix that appears in the numerator of (20.42). Then we have to minimize $|\boldsymbol{\zeta}^\top \boldsymbol{A\zeta}|$ with respect to $\boldsymbol{\zeta}$ subject to the constraint that $\boldsymbol{\zeta}^\top \boldsymbol{\zeta} = \mathbf{I}$.

In order to perform this minimization, it turns out to be enough to consider the eigenvalue-eigenvector problem associated with $\boldsymbol{A}$. If we solve this problem, we will obtain an orthogonal matrix $\boldsymbol{Z}$, the columns of which are orthonormalized eigenvectors of $\boldsymbol{A}$, and a diagonal matrix $\boldsymbol{\Lambda}$, the diagonal elements of which are the eigenvalues of $\boldsymbol{A}$, which must evidently lie between zero and unity. Then $\boldsymbol{AZ} = \boldsymbol{Z\Lambda}$. If the columns of $\boldsymbol{Z}$ and $\boldsymbol{\Lambda}$ are arranged in increasing order of the eigenvalues $\lambda_1, \ldots, \lambda_m$, we may choose the ML estimate $\hat{\boldsymbol{\zeta}}$ to be the first r columns of $\boldsymbol{Z}$. Geometrically, the columns of $\hat{\boldsymbol{\zeta}}$ span the space spanned by the eigenvectors of $\boldsymbol{A}$ that correspond to the r smallest eigenvalues. The fact that $\boldsymbol{Z}$ is orthogonal means that $\hat{\boldsymbol{\zeta}}$ satisfies the constraint, and the choice of the *smallest* eigenvalues serves to minimize the determinant $|\boldsymbol{\zeta}^\top \boldsymbol{A\zeta}|$.

The ML estimate of the space of cointegrating vectors $\mathcal{S}(\boldsymbol{\eta})$ can now be recovered from $\hat{\boldsymbol{\zeta}}$ by the formula $\hat{\boldsymbol{\eta}} = \boldsymbol{S}^{-1}\hat{\boldsymbol{\zeta}}$. The matrix $\hat{\boldsymbol{\alpha}}$ needed in order to obtain ML estimates of the parameters contained in the matrix $\boldsymbol{\Pi}$ can then be obtained as the OLS estimates from the multivariate regression of $\Delta \boldsymbol{Y}^*$ on $\boldsymbol{Y}^*_{-p}\hat{\boldsymbol{\eta}}$. Subsequently, estimates of the matrices $\boldsymbol{\Gamma}_i$, $i = 1, \ldots, p-1$, can also be obtained by OLS.

Often, we are not especially interested in the parameters of the VAR (20.35). The focus of our interest is more likely to be testing the hypothesis of noncointegration against an alternative of cointegration of some chosen order. Should the null hypothesis that $r = 0$ be rejected, we may then wish to test the hypothesis that $r = 1$ against the alternative that $r = 2$, and so forth. The eigenvalues λ_i, $i = 1, \ldots, m$, provide a very convenient way to do this, in terms of a likelihood ratio test. It is clear that if we select some value of r, the minimized determinant $|\boldsymbol{\zeta}^\top \boldsymbol{A\zeta}|$ is just the product of the r smallest eigenvalues, $\lambda_1 \cdots \lambda_r$. The minimum of (20.40) is this product multiplied by $|(\Delta \boldsymbol{Y}^*)^\top \Delta \boldsymbol{Y}^*|$. If $r = 0$, then the minimum of (20.40) is simply this last determinant. Likelihood ratios for different values of r are therefore just products of some of the eigenvalues, raised to the power $n/2$; recall (9.65). If we take logs and multiply by 2 in order to obtain an LR statistic, we obtain $-n$ times the sum of the logs of the appropriate eigenvalues.