the scalar case discussed in §25.1, just like the CLT itself. Consider the m.d. case in which $\boldsymbol{\Sigma}_n = \sum_{t=1}^n E(\boldsymbol{X}_t\boldsymbol{X}_t')$, and assume this matrix has rank p asymptotically in the sense defined above. Under the assumptions of **25.2**,

$$\frac{\sum_{t=1}^n(\boldsymbol{\alpha}'\boldsymbol{X}_t)^2}{\boldsymbol{\alpha}'\boldsymbol{\Sigma}_n\boldsymbol{\alpha}} = \frac{\boldsymbol{\alpha}'\left(\sum_{t=1}^n\boldsymbol{X}_t\boldsymbol{X}_t'\right)\boldsymbol{\alpha}}{\boldsymbol{\alpha}'\boldsymbol{\Sigma}_n\boldsymbol{\alpha}} \xrightarrow{pr} 1 \tag{25.29}$$

for any $\boldsymbol{\alpha}$ with $\boldsymbol{\alpha}'\boldsymbol{\alpha} = 1$, where the ratio is always well defined on taking n large enough, by assumption. This suggests that the positive semidefinite matrix $\hat{\boldsymbol{\Sigma}}_n = \sum_{t=1}^n\boldsymbol{X}_t\boldsymbol{X}_t'$ is the natural estimator for $\boldsymbol{\Sigma}_n$.

To be more precise: (25.29) says that $P(|\boldsymbol{\alpha}'\hat{\boldsymbol{\Sigma}}_n\boldsymbol{\alpha}/\boldsymbol{\alpha}'\boldsymbol{\Sigma}_n\boldsymbol{\alpha} - 1|)$ can be made as small as desired for *arbitrary* $\boldsymbol{\alpha} \neq \mathbf{0}$ by taking n large enough, since the normalization to unit length cancels in the ratio. This is therefore true in the particular case $\boldsymbol{\alpha}^* = \boldsymbol{L}_n^{-\prime}\boldsymbol{\alpha}$, and since $\boldsymbol{\alpha}^{*\prime}\boldsymbol{\Sigma}_n\boldsymbol{\alpha}^* = 1$, we are able to conclude that $\boldsymbol{\alpha}'\boldsymbol{L}_n^{-}\hat{\boldsymbol{\Sigma}}_n\boldsymbol{L}_n^{-\prime}\boldsymbol{\alpha} \xrightarrow{pr} 1$. We can further deduce from this fact that $\boldsymbol{L}_n^{-}\hat{\boldsymbol{\Sigma}}_n\boldsymbol{L}_n^{-\prime} \xrightarrow{pr} \boldsymbol{I}_p$. To show this, note that if a matrix $\boldsymbol{B}$ $(p \times p)$ is nonsingular, and $g(\boldsymbol{\alpha}) = \boldsymbol{\alpha}'\boldsymbol{B}\boldsymbol{\alpha}/\boldsymbol{\alpha}'\boldsymbol{\alpha} = 1$ for every $\boldsymbol{\alpha} \neq \mathbf{0}$, g has the gradient vector $g'(\boldsymbol{\alpha}) = 2\boldsymbol{B}\boldsymbol{\alpha}/\boldsymbol{\alpha}'\boldsymbol{\alpha} - 2\boldsymbol{\alpha}/\boldsymbol{\alpha}'\boldsymbol{\alpha}$, for any $\boldsymbol{\alpha}$, and the system of equations $\boldsymbol{B}\boldsymbol{\alpha}/\boldsymbol{\alpha}'\boldsymbol{\alpha} - \boldsymbol{\alpha}/\boldsymbol{\alpha}'\boldsymbol{\alpha} = \mathbf{0}$ has the unique solution $\boldsymbol{B} = \boldsymbol{I}_p$. If $\hat{\boldsymbol{L}}_n$ is the factorization of $\hat{\boldsymbol{\Sigma}}_n$, since $\boldsymbol{L}_n$ is asymptotically of rank p it follows by **18.10**(ii) that $\hat{\boldsymbol{L}}_n\boldsymbol{L}_n^{-} \xrightarrow{pr} \boldsymbol{I}_p$, and we arrive at the desired conclusion, for comparison with (25.27):

$$\hat{\boldsymbol{L}}_n^{-}\sum_{t=1}^n\boldsymbol{X}_t \xrightarrow{D} N(\mathbf{0}, \boldsymbol{I}_p). \tag{25.30}$$

The extension to general dependence is a matter of estimating $\boldsymbol{\Sigma}_n$ by a generalization of the consistent methods discussed in §25.1, either $\hat{\boldsymbol{\Sigma}}_{n1} = \sum_{i=1}^{r_n}\mathbf{Z}_{ni}^*\mathbf{Z}_{ni}^{*\prime}$ where $\mathbf{Z}_{ni}^* = \sum_{t=(i-1)b_n+1}^{ib_n}\boldsymbol{X}_t$, or, letting weights $\{w_{nk}\}$ represent the Bartlett kernel,

$$\hat{\boldsymbol{\Sigma}}_{n3} = \sum_{t=1}^n\boldsymbol{X}_t\boldsymbol{X}_t' + \sum_{k=1}^{b_n}w_{nk}\sum_{t=k+1}^n(\boldsymbol{X}_t\boldsymbol{X}_{t-k}' + \boldsymbol{X}_{t-k}\boldsymbol{X}_t'). \tag{25.31}$$

The latter matrix is assuredly positive definite, since $\boldsymbol{\alpha}'\hat{\boldsymbol{\Sigma}}_{n3}\boldsymbol{\alpha} \geq 0$ with arbitrary $\boldsymbol{\alpha}$ by application of (25.12) with $X_t = \boldsymbol{X}_t\boldsymbol{\alpha}$.

25.4 Error Estimation

There are a number of celebrated theorems in the classical probability literature on the rates at which the deviations of distributions from their weak limits, and also stochastic sequences from their almost sure limits, decline with n. Most are for the independent case, and the extensions to general dependent sequences are not much researched to date. These results will not be treated in detail, but it is useful to know of their existence. For details and proofs the reader is referred to texts such as Chung (1974) and Loève (1977).

If $\{F_n\}$ is a sequence of c.d.f.s and $F_n \Rightarrow \Phi$ (the Gaussian c.d.f.), the Berry-Esséen theorem sets limits on the largest deviation of F_n from Φ. The setting for