1 Quantities

- $J \in \mathbb{Z}^+$ = number of scales decomposition
- $T = 2^J$ = number of time periods
- $N \in \mathbb{Z}^+$ = number of cross-section elements
- $K(\leq N)$ = number of common factors

2 Multivariate Locally Stationary Wavelet process (Park et al. (2014))

The vector $(N \times 1)$ of stochastic processes $\boldsymbol{X_{t;T}}$ follows the given decomposition :

$$\boldsymbol{X_{t;T}} = \sum_{j=-J}^{-1} \sum_{k=0}^{T} \boldsymbol{W}_j\left(k/T\right) \boldsymbol{\xi}_{j,k} \psi_{j,k}(t) \tag{2.1}$$

where

- $\boldsymbol{W_j(z)}$ is a lower-triangular $(N \times N)$ matrix.
  For each (m,n)-element,

$$W_j^{(m,n)}(z) \text{ is a Lipschitz continuous function on } z \in (0,1) \tag{2.2}$$

$$\sum_{j=-\infty}^{-1} \left| W_j^{(m,n)}(z) \right|^2 < \infty, \qquad \forall z \in (0,1) \qquad \text{(finite energy)} \tag{2.3}$$

$$\sum_{j=-\infty}^{-1} 2^{-j} L_j^{(m,n)} < \infty \qquad \text{(uniformly bounded Lipschitz constants } L_j\text{)} \tag{2.4}$$

- $\boldsymbol{\xi}_{j,k}$ is the vector $(N \times 1)$ of random orthonormal increments.

$$\mathrm{E}\left[\xi_{j,k}^{(u)}\right] = 0, \qquad \forall j,k,u \tag{2.5}$$

$$\mathrm{Cov}\left[\xi_{j,k}^{(u)}, \xi_{j',k'}^{(u')}\right] = \delta_{j,j'}\delta_{k,k'}\delta_{u,u'}, \qquad \forall j,j',k,k',u,u' \tag{2.6}$$

- $\psi_{j,k}(t) = \psi_{j,k-t}$ is a scalar representing a non-decimated wavelet.

We can define the *Cross-Evoluationary Wavelet Spectrum* $(N \times N)$ matrix : $\boldsymbol{S}_j(z) = \boldsymbol{W}_j(z)\boldsymbol{W}_j(z)'$. This gives us the ability to express the *local autocovariance* : $c^{(u,u')}(z,\tau) = \sum_{j=-\infty}^{-1} S_j^{(u,u')}(z)\boldsymbol{\Psi}_j(\tau)$ where $\boldsymbol{\Psi}_j(\tau) = \sum_k \psi_{j,k}(0)\psi_{j,k}(\tau)$, the *autocorrelation wavelet*. The latter also define the *inner product matrix of discrete autocorrelation wavelets* : $A_{jl} = \sum_\tau \boldsymbol{\Psi}_j(\tau)\boldsymbol{\Psi}_l(\tau)$, $A = \{A_{jl}\}_{j,l\in\mathbb{N}}$ and its inverse : $\bar{A} = A^{-1}$.

Each (m,n)-element of the Cross-Evolutionary Wavelet Spectrum can be expressed as

$$S_j^{(m,n)}\left(k/T\right) = \sum_{u=1}^{N} W_j^{(m,u)}\left(k/T\right) W_j^{(u,n)}\left(k/T\right), \quad \forall j,k$$

From this definition it is not difficult to extend the CEWS to take into account the dependence structure between different scales and throught time :

$$S_{j,j'}^{m,n}\left(k/T, k'/T\right) = \sum_{u=1}^{N} W_j^{(m,u)}\left(k/T\right) W_{j'}^{(u,n)}\left(k'/T\right), \quad \forall j,j',k,k' \tag{2.7}$$

We make the following assumption regarding the latter object,

$$S_{j,j'}^{(m,n)}\left(k/T, k/T\right) = \begin{cases} S_j^{(m,n)}\left(k/T\right) & \text{if } j = j' \\ 0 & \text{otherwise} \end{cases} \tag{2.8}$$

This assumption (possible improvement : condition similar to Chamberlain ?) imposes no dependence between different scales of decomposition. Notice that we don't restrict the serial dependence.

2.1 Estimation of MvLSW

- $\boldsymbol{d}_{j,k} = \sum_{t=0}^{T-1} \boldsymbol{X}_t \psi_{j,k}(t)$ (empirical wavelet coefficients)
- $\boldsymbol{I}_{j,k} = \boldsymbol{d}_{j,k}\boldsymbol{d}'_{j,k}$ (raw wavelet periodogram)
  - $\mathrm{E}\left[\boldsymbol{I}_{j,k}\right] = \sum_{l=-J}^{-1} A_{jl}\boldsymbol{S}_l\left(k/T\right) + O(T^{-1})$ (biaised estimator)
- $\bar{\boldsymbol{I}}_{j,k} = \sum_{l=-J}^{-1} \bar{A}_{j,l}\boldsymbol{I}_{l,k}$ (corrected periodogram) $\Longrightarrow$ (unbiased estimator)
- $\widetilde{\boldsymbol{I}}_{j,k} = \frac{1}{2M+1}\sum_{m=-M}^{M} \boldsymbol{I}_{j,k+m}$ (smooth periodogram) $\Longrightarrow$ (consistent estimator)
- $\widehat{\boldsymbol{S}}_j\left(k/T\right) = \sum_{l=-J}^{-1} \bar{A}_{jl}\widetilde{\boldsymbol{I}}_{l,k} = \frac{1}{2M+1}\sum_{m=-M}^{M} \bar{\boldsymbol{I}}_{j,k+m} = \frac{1}{2M+1}\sum_{m=-M}^{M}\sum_{l=-J}^{-1} \bar{A}_{j,l}\boldsymbol{I}_{l,k+m}$ (final estimator of CEWS)

2.2 Notes

- The dependence structure is entirely in $\boldsymbol{W}_j(z)$, not in $\boldsymbol{\xi}_{j,k}$.
- The lower-triangular form of $\boldsymbol{W}_j(z)$ allows us to use the Cholesky decomposition on $\boldsymbol{S}_j(z)$.

3 Factor Model

- The factor structure is imposed on the following :

$$\boldsymbol{W}_j\left(k/T\right)\boldsymbol{\xi}_{j,k} = \boldsymbol{\Lambda}_{j,k}\boldsymbol{F}_k + \boldsymbol{\epsilon}_{j,k} \tag{3.1}$$

  , not only on $\boldsymbol{\xi}_{j,k}$ since they are assumed orthonormal.

- Assumptions :
  1. $\boldsymbol{F}_k \sim (\boldsymbol{0}, \boldsymbol{\Sigma}_F)$, where $\boldsymbol{\Sigma}_F$ is a diagonal positive definite $(K \times K)$ matrix.
  2. $\boldsymbol{F}_k \perp \boldsymbol{\epsilon}_{j,k'}, \forall j, k, k'$
  3. $\boldsymbol{\epsilon}_{j,k} \sim (\boldsymbol{0}, \boldsymbol{\Sigma}_\epsilon)$, where $\boldsymbol{\Sigma}_\epsilon$ has bounded eigenvalues. Note : make $\boldsymbol{\Sigma}_\epsilon$ dependent on time ? what about serial dependence ?
  4. $\boldsymbol{\Lambda}'_{j,k}\boldsymbol{\Lambda}_{l,m} = \boldsymbol{0}, \forall j \neq l, \forall k \neq m$.

- We can then represent the CEWS with the factor structure :

$$\begin{aligned}
\mathrm{Var}\left[\boldsymbol{W}_j\left(k/T\right)\boldsymbol{\xi}_{j,k}\right] &= \mathrm{Var}\left[\boldsymbol{\Lambda}_{j,k}\boldsymbol{F}_k\right] + \mathrm{Var}\left[\boldsymbol{\epsilon}_{j,k}\right] \\
\boldsymbol{W}_j\left(k/T\right)\mathrm{Var}\left[\boldsymbol{\xi}_{j,k}\right]\boldsymbol{W}_j\left(k/T\right)' &= \boldsymbol{\Lambda}_{j,k}\mathrm{Var}\left[\boldsymbol{F}_k\right]\boldsymbol{\Lambda}'_{j,k} + \mathrm{Var}\left[\boldsymbol{\epsilon}_{j,k}\right] \\
\boldsymbol{W}_j\left(k/T\right)\boldsymbol{W}_j\left(k/T\right)' &= \boldsymbol{\Lambda}_{j,k}\boldsymbol{\Sigma}_F\boldsymbol{\Lambda}'_{j,k} + \boldsymbol{\Sigma}_\epsilon \qquad \text{from (2.6)} \\
\boldsymbol{S}_j\left(k/T\right) &= \boldsymbol{\Lambda}_{j,k}\boldsymbol{\Sigma}_F\boldsymbol{\Lambda}'_{j,k} + \boldsymbol{\Sigma}_\epsilon
\end{aligned}$$

3.1 Estimation

The estimation of the loadings and common factors is carried out by a non-linear least square procedure in the wavelet domain.

$$\begin{aligned}
\min_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}_{\forall j,k}, \{\bar{\boldsymbol{F}}_k\}_{\forall j,k}} \quad & (NT)^{-1}\sum_t \left[\boldsymbol{X}_{t;T} - \sum_{j=-J}^{-1}\sum_{k=0}^{T}\left(\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{F}}_k\right)\psi_{j,k}(t)\right]'\left[\boldsymbol{X}_{t;T} - \sum_{j=-J}^{-1}\sum_{k=0}^{T}\left(\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{F}}_k\right)\psi_{j,k}(t)\right] \\
\text{s.t.} \quad & \frac{\bar{\boldsymbol{\Lambda}}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}}{N} = \boldsymbol{I}_K
\end{aligned} \tag{3.2}$$

After distributing the objective function becomes,

$$(NT)^{-1}\sum_t \left[\boldsymbol{X}'_{t;T}\boldsymbol{X}_{t;T} - \boldsymbol{X}'_{t;T}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{F}}_k\psi_{j,k}(t) - \sum_{j=-J}^{-1}\sum_{k=0}^{T}\psi_{j,k}(t)\bar{\boldsymbol{F}}'_k\bar{\boldsymbol{\Lambda}}'_{j,k}\boldsymbol{X}_{t;T} + \sum_{j=-J}^{-1}\sum_{k=0}^{T}\sum_{l=-J}^{-1}\sum_{m=0}^{T}\psi_{j,k}(t)\psi_{l,m}(t)\bar{\boldsymbol{F}}'_k\bar{\boldsymbol{\Lambda}}'_{j,k}\bar{\boldsymbol{\Lambda}}_{l,m}\bar{\boldsymbol{F}}_m\right]$$

$$(NT)^{-1}\Big[\sum_t \boldsymbol{X'_{t;T}}\boldsymbol{X_{t;T}} - \sum_{j=-J}^{-1}\sum_{k=0}^{T}\sum_t \boldsymbol{X'_{t;T}}\psi_{j,k}(t)\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{F}}_k - \sum_{j=-J}^{-1}\sum_{k=0}^{T}\bar{\boldsymbol{F}}'_k\bar{\boldsymbol{\Lambda}}'_{j,k}\sum_t \boldsymbol{X_{t;T}}\psi_{j,k}(t) + \sum_t\sum_{j=-J}^{-1}\sum_{k=0}^{T}\sum_{l=-J}^{-1}\sum_{m=0}^{T}\psi_{j,k}(t)\psi_{l,m}(t)\bar{\boldsymbol{F}}'_k\bar{\boldsymbol{\Lambda}}'_{j,k}\bar{\boldsymbol{\Lambda}}_{l,m}\bar{\boldsymbol{F}}_m\Big]$$

By definition of the empirical wavelet coefficents,

$$(NT)^{-1}\Big[\sum_t \boldsymbol{X'_{t;T}}\boldsymbol{X_{t;T}} - \sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{F}}_k - \sum_{j=-J}^{-1}\sum_{k=0}^{T}\bar{\boldsymbol{F}}'_k\bar{\boldsymbol{\Lambda}}'_{j,k}\boldsymbol{d}_{j,k} + \sum_t\sum_{j=-J}^{-1}\sum_{k=0}^{T}\sum_{l=-J}^{-1}\sum_{m=0}^{T}\psi_{j,k}(t)\psi_{l,m}(t)\bar{\boldsymbol{F}}'_k\bar{\boldsymbol{\Lambda}}'_{j,k}\bar{\boldsymbol{\Lambda}}_{l,m}\bar{\boldsymbol{F}}_m\Big]$$

By assumption on the loadings (possible improvement) and the fact that wavelets are normalized $\sum_t(\psi_{j,k}(t))^2 = 1, \forall j,k$,

$$(NT)^{-1}\Big[\sum_t \boldsymbol{X'_{t;T}}\boldsymbol{X_{t;T}} - \sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{F}}_k - \sum_{j=-J}^{-1}\sum_{k=0}^{T}\bar{\boldsymbol{F}}'_k\bar{\boldsymbol{\Lambda}}'_{j,k}\boldsymbol{d}_{j,k} + \sum_{j=-J}^{-1}\sum_{k=0}^{T}\bar{\boldsymbol{F}}'_k\bar{\boldsymbol{\Lambda}}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{F}}_k\Big]$$

The First Order Conditions with respect to the factors are given by :

$$\bar{\boldsymbol{F}}'_k\sum_{j=-J}^{-1}\bar{\boldsymbol{\Lambda}}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k} - \sum_{j=-J}^{-1}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k} = 0 \quad ,\forall k$$

$$\bar{\boldsymbol{F}}'_k\sum_{j=-J}^{-1}N = \sum_{j=-J}^{-1}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k} \quad ,\forall k \qquad \text{from (3.2)}$$

$$\bar{\boldsymbol{F}}'_k = (JN)^{-1}\sum_{j=-J}^{-1}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k} \tag{3.3}$$

Replace (3.3) in the original minimization problem,

$$\begin{aligned}\min_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}\forall j,k}\quad (NT)^{-1}\Big[\sum_t \boldsymbol{X'_{t;T}}\boldsymbol{X_{t;T}} &- \sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\left((JN)^{-1}\sum_{l=-J}^{-1}\bar{\boldsymbol{\Lambda}}'_{l,k}\boldsymbol{d}_{l,k}\right)\\ &- \sum_{j=-J}^{-1}\sum_{k=0}^{T}\left((JN)^{-1}\sum_{l=-J}^{-1}\boldsymbol{d}'_{l,k}\bar{\boldsymbol{\Lambda}}_{l,k}\right)\bar{\boldsymbol{\Lambda}}'_{j,k}\boldsymbol{d}_{j,k}\\ &+ \sum_{j=-J}^{-1}\sum_{k=0}^{T}\left((JN)^{-1}\sum_{l=-J}^{-1}\boldsymbol{d}'_{l,k}\bar{\boldsymbol{\Lambda}}_{l,k}\right)\bar{\boldsymbol{\Lambda}}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\left((JN)^{-1}\sum_{n=-J}^{-1}\bar{\boldsymbol{\Lambda}}'_{n,k}\boldsymbol{d}_{n,k}\right)\Big]\end{aligned}$$

$$\begin{aligned}\min_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}\forall j,k}\quad (NT)^{-1}\Big[\sum_t \boldsymbol{X'_{t;T}}\boldsymbol{X_{t;T}} &- (JN)^{-1}\sum_{l=-J}^{-1}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{\Lambda}}'_{l,k}\boldsymbol{d}_{l,k}\\ &- (JN)^{-1}\sum_{l=-J}^{-1}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{l,k}\bar{\boldsymbol{\Lambda}}_{l,k}\bar{\boldsymbol{\Lambda}}'_{j,k}\boldsymbol{d}_{j,k}\\ &+ (JN)^{-2}\sum_{l=-J}^{-1}\sum_{n=-J}^{-1}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{l,k}\bar{\boldsymbol{\Lambda}}_{l,k}\bar{\boldsymbol{\Lambda}}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{\Lambda}}'_{n,k}\boldsymbol{d}_{n,k}\Big]\end{aligned}$$

$$\begin{aligned}\min_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}\forall j,k}\quad (NT)^{-1}\Big[\sum_t \boldsymbol{X'_{t;T}}\boldsymbol{X_{t;T}} &- (JN)^{-1}\sum_{l=-J}^{-1}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{\Lambda}}'_{l,k}\boldsymbol{d}_{l,k}\\ &- (JN)^{-1}\sum_{l=-J}^{-1}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{l,k}\bar{\boldsymbol{\Lambda}}_{l,k}\bar{\boldsymbol{\Lambda}}'_{j,k}\boldsymbol{d}_{j,k}\\ &+ (JN)^{-2}JN\sum_{l=-J}^{-1}\sum_{n=-J}^{-1}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{l,k}\bar{\boldsymbol{\Lambda}}_{l,k}\bar{\boldsymbol{\Lambda}}'_{n,k}\boldsymbol{d}_{n,k}\Big] \qquad \text{from (3.2)}\end{aligned}$$

$$\min_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}\forall j,k}\quad (NT)^{-1}\Big[\sum_t \boldsymbol{X'_{t;T}}\boldsymbol{X_{t;T}} - (JN)^{-1}\sum_{l=-J}^{-1}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{\Lambda}}'_{l,k}\boldsymbol{d}_{l,k}\Big]$$

Minimizing the latter expression is equivalent to maximizing,

$$\max_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}\forall j,k}\quad (JT)^{-1}N^{-2}\sum_{l=-J}^{-1}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{\Lambda}}'_{l,k}\boldsymbol{d}_{l,k}$$

Each term in the triple sum is a scalar, i.e. (1×1) matrix. Therefore we can freely take its trace,

$$\max_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}_{\forall j,k}} \quad (JT)^{-1}N^{-2}\sum_{l=-J}^{-1}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\mathrm{tr}\left\{\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\bar{\boldsymbol{\Lambda}}'_{l,k}\boldsymbol{d}_{l,k}\right\}$$

From the cyclic property of the trace,

$$\max_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}_{\forall j,k}} \quad (JT)^{-1}N^{-2}\sum_{l=-J}^{-1}\sum_{j=-J}^{-1}\sum_{k=0}^{T}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{l,k}\boldsymbol{d}_{l,k}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\right\}$$

$$\max_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}_{\forall j,k}} \quad (JT)^{-1}N^{-2}\left[\sum_{k=0}^{T}\sum_{j=-J}^{-1}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{j,k}\boldsymbol{d}_{j,k}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\right\}+2\sum_{k=0}^{T}\sum_{j=-J}^{-1}\sum_{l=-J}^{j-1}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{l,k}\boldsymbol{d}_{l,k}\boldsymbol{d}'_{j,k}\bar{\boldsymbol{\Lambda}}_{j,k}\right\}\right]$$

We recognize the raw wavelet periodogram $\boldsymbol{d}_{l,k}\boldsymbol{d}'_{j,k}$. We replace the latter with the unbiased and consistent estimator of the CEWS, i.e. $\widehat{\boldsymbol{S}}_j\left(k/T\right)$.
Therefore the first term becomes,

$$\sum_{k=0}^{T}\sum_{j=-J}^{-1}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{j,k}\widehat{\boldsymbol{S}}_j\left(k/T\right)\bar{\boldsymbol{\Lambda}}_{j,k}\right\}$$

The second term needs a similar treatement. We replace $\boldsymbol{d}_{l,k}\boldsymbol{d}'_{j,k}$ with the unbiaised and consistent estimator of (2.7) :

$$\hat{\boldsymbol{S}}_{j,j'}(k/T,k/T)=\frac{1}{2M+1}\sum_{m=-M}^{M}\sum_{r=-J}^{-1}\sum_{l=-J}^{-1}\bar{A}_{j,l}^{(j',r)}\boldsymbol{d}_{j,k+m}\boldsymbol{d}'_{j',k+m} \tag{3.4}$$

where $\bar{A}_{j,l}^{(j',l')}=\sum_{\tau}\Psi_{j,j'}(\tau)\Psi_{l,l'}(\tau)$ and $\Psi_{j,j'}(\tau)=\sum_{t}\psi_{j,0}(t)\psi_{j',\tau}(t)$, the inner product operator of the cross-correlation wavelet functions and the *Cross-Correlation Wavelet Function*, respectively. In his thesis Koch showed that this CCWF inherit the same properties as the autocorrelation function. One of the latter is that the functions in that family are linearly indepedent of each other. Consequently, the inner product operator of that family is invertible.

The second term on the maximization problem thus reads,

$$2\sum_{k=0}^{T}\sum_{j=-J}^{-1}\sum_{l=-J}^{j-1}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{l,k}\hat{\boldsymbol{S}}_{j,l}(k/T,k/T)\bar{\boldsymbol{\Lambda}}_{j,k}\right\}$$

Consequently, the whole optimization problem is changed into

$$\max_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}_{\forall j,k}} \quad (JT)^{-1}N^{-2}\left[\sum_{k=0}^{T}\sum_{j=-J}^{-1}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{j,k}\widehat{\boldsymbol{S}}_j\left(k/T\right)\bar{\boldsymbol{\Lambda}}_{j,k}\right\}+2\sum_{k=0}^{T}\sum_{j=-J}^{-1}\sum_{l=-J}^{j-1}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{l,k}\hat{\boldsymbol{S}}_{j,l}(k/T,k/T)\bar{\boldsymbol{\Lambda}}_{j,k}\right\}\right]$$

which is asymptotically equivalent to (see Park et al. (2014)),

$$\max_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}_{\forall j,k}} \quad (JT)^{-1}N^{-2}\left[\sum_{k=0}^{T}\sum_{j=-J}^{-1}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{j,k}\boldsymbol{S}_j\left(k/T\right)\bar{\boldsymbol{\Lambda}}_{j,k}\right\}+2\sum_{k=0}^{T}\sum_{j=-J}^{-1}\sum_{l=-J}^{j-1}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{l,k}\boldsymbol{S}_{j,l}(k/T,k/T)\bar{\boldsymbol{\Lambda}}_{j,k}\right\}\right]$$

The last term is zero by assumption (2.8),

$$\max_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}_{\forall j,k}} \quad (JT)^{-1}N^{-2}\sum_{k=0}^{T}\sum_{j=-J}^{-1}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{j,k}\boldsymbol{S}_j\left(k/T\right)\bar{\boldsymbol{\Lambda}}_{j,k}\right\}$$

Finally, we get back the feasible problem,

$$\max_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}_{\forall j,k}} \quad (JT)^{-1}N^{-2}\sum_{k=0}^{T}\sum_{j=-J}^{-1}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{j,k}\widehat{\boldsymbol{S}}_j\left(k/T\right)\bar{\boldsymbol{\Lambda}}_{j,k}\right\}+O(M^{-1}) \tag{3.5}$$

where $M(T)\to\infty$ when $T\to\infty$.

This final problem can be decompose into sub-problems and the latter can be solved independently. In other words, the problem (3.5) is maximized when each term in the double sum is also maximized. (Proof)

$$\max_{\{\bar{\boldsymbol{\Lambda}}_{j,k}\}_{\forall j,k}} \quad (JT)^{-1}\sum_{k=0}^{T}\sum_{j=-J}^{-1}N^{-2}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{j,k}\widehat{\boldsymbol{S}}_j\left(k/T\right)\bar{\boldsymbol{\Lambda}}_{j,k}\right\} \quad = \quad (JT)^{-1}\sum_{k=0}^{T}\sum_{j=-J}^{-1}\max_{\bar{\boldsymbol{\Lambda}}_{j,k}}N^{-2}\mathrm{tr}\left\{\bar{\boldsymbol{\Lambda}}'_{j,k}\widehat{\boldsymbol{S}}_j\left(k/T\right)\bar{\boldsymbol{\Lambda}}_{j,k}\right\}$$

This solution of those final optimization problems are $\sqrt{N}\tilde{\mathbf{\Lambda}}_{j,k}$, where $\tilde{\mathbf{\Lambda}}_{j,k}$ is the $(N \times K)$ matrix whose columns are the first K orthonormal eigenvectors of $\widehat{\boldsymbol{S}}_j\,(k/T)$.

We obtain the wanted least squares estimators of the loadings and factors as

$$\widehat{\mathbf{\Lambda}}_{j,k} = \sqrt{N}\tilde{\mathbf{\Lambda}}_{j,k} \tag{3.6}$$

$$\widehat{\boldsymbol{F}}_k = (JN)^{-1} \sum_{j=-J}^{-1} \widehat{\mathbf{\Lambda}}'_{j,k}\boldsymbol{d}_{j,k} \tag{3.7}$$

4 Goals

- $\|\hat{\mathbf{\Lambda}}_{j,k} - \mathbf{\Lambda}_{j,k}\boldsymbol{R}_{j,k}\| = O(1)$
- $\|\hat{\boldsymbol{F}}_k - \boldsymbol{R}_{j,k}^{-1}\boldsymbol{F}_k\| = O(1)$