Lecture Notes on Factor Models

John C. Chao

Econ 721 Lecture Notes

December 1, 2022

Factor Models

- **Model, Notations, and Assumptions:** Consider the model

$$\underset{N\times 1}{X_{\cdot t}} = \underset{N\times m}{\Lambda^0} \underset{m\times 1}{F_t^0} + \underset{N\times 1}{e_{\cdot t}} \text{ for } t = 1, ..., T$$

where

Λ^0 - factor loadings,

F_t^0 - common factors (latent, unobserved),

$\Lambda^0 F_t^0$ - common component,

$e_{\cdot t}$ - idiosyncratic (or error) component

A Taxonomy of Factor Models

- **Exact Factor Model:** $\{e_{it}\}$ exhibits no cross-sectional dependence and no temporal dependence, so that

$$Cov\left(e_{it}, e_{js}\right) = 0 \text{ for all } i \neq j \text{ and/or } t \neq s$$

- **Approximate Factor Model:** $\{e_{it}\}$ can exhibit some cross-sectional dependence, so that

$$Cov\left(e_{it}, e_{jt}\right) \neq 0 \text{ for at least some } i \neq j$$

- **Generalized Factor Model:** $\{e_{it}\}$ can exhibit both cross-sectional and temporal dependence.

(Classical) Exact Factor Models

- **Model:** Consider the model

$$\underset{p\times 1}{X_t} = \underset{p\times m}{\Lambda^0} \underset{m\times 1}{F_t^0} + \underset{p\times 1}{e_t}, \quad t = 1, \ldots, n$$

- **Assumptions:**

1. $E[e_t] = \underset{p\times 1}{0}$,
2. $E[F_t^0 e_t'] = \underset{m\times p}{0}$ (orthogonality between factors and errors),
3. $E[e_t e_t'] = \underset{p\times p}{\Psi} = diag(\psi_{11}, .., \psi_{pp}) > 0$ (no correlation between idiosyncratic components),
4. $E[F_t^0] = \underset{m\times 1}{0}$,
5. $E[F_t^0 F_t^{0\prime}] = \underset{m\times m}{\Phi^0} > 0$.
6. m and p are both fixed.

(Classical) Exact Factor Models

- **Remark:** A key feature of the classical factor model is that $E\left[e_t e_t'\right]$ is diagonal so that the components of e_t are uncorrelated. Hence, the only correlation between the components of X_t comes from the common factors.

Identification

- Note that in the absence of further restrictions, there is an indeterminancy in the specification given above. In particular, for any nonsingular matrix C, we can write

$$\begin{aligned} X_t &= \Lambda^0 F_t^0 + e_t \\ &= \Lambda^0 C C^{-1} F_t^0 + e_t \\ &= \Lambda^* F_t^* + e_t, \end{aligned}$$

where $\Lambda^* = \Lambda^0 C$ and $F_t^* = C^{-1} F_t^0$. It follows that the structures $\left(\Lambda^0, \Phi^0, \Psi\right)$ and $(\Lambda^*, \Phi^*, \Psi) = \left(\Lambda^0 C, C^{-1} \Phi^0 C'^{-1}, \Psi\right)$ are **observationally equivalent** in the sense that they give the same value of the likelihood function, so that it will not be possible to differentiate between them based on data.

Orthogonal Factor Case

- Some of the indeterminancy can be removed by assuming that

$$E\left[F_t^0 F_t^{0\prime}\right] = \Phi^0 = I_m.$$

- **Remark:** In this case, the factors are said to be **orthogonal**. On the other hand, if Φ^0 is some more general symmetric, positive definite matrix that is not diagonal; then the factors are said to be **oblique**.

Orthogonal Factor Case

- In the case of orthogonal factors, we have

$$\begin{aligned}\Sigma &= E\left[X_{\cdot t}X_{\cdot t}'\right] \\ &= E\left[\left(\Lambda^0 F_t^0 + e_t\right)\left(F_t^{0\prime}\Lambda^{0\prime} + e_t'\right)\right] \\ &= \Lambda^0 E\left[F_t^0 F_t^{0\prime}\right]\Lambda^{0\prime} + \Psi \\ &= \Lambda^0\Lambda^{0\prime} + \Psi \;\left(\text{since } E\left[F_t^0 F_t^{0\prime}\right] = I_m\right)\end{aligned}$$

- Note that there is still an indeterminancy in this case since, for any orthogonal matrix C, i.e., $C'C = CC' = I_m$; we have

$$\Sigma = \Lambda^0\Lambda^{0\prime} + \Psi = \Lambda^0 CC'\Lambda^{0\prime} + \Psi = \Lambda^*\Lambda^{*\prime} + \Psi$$

where $\Lambda^* = \Lambda^0 C$, so that with orthogonal factors, this indeterminancy is only caused by an orthogonal transformation, as opposed to any nonsingular transformation.

Additional Identifying Assumptions

- Additional identifying restrictions have been introduced in the literature to try to fully identify the classical factor model. An often used assumption is to take

$$\underset{m \times m}{\Gamma} = \Lambda^{0\prime} \Psi^{-1} \Lambda^{0} = \begin{pmatrix} \gamma_{11} & 0 & \cdots & 0 \\ 0 & \gamma_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \gamma_{mm} \end{pmatrix}$$

If the diagonal elements of Γ are ordered and all different, i.e.,

$$\gamma_{11} > \gamma_{22} > \cdots > \gamma_{mm};$$

then Λ^{0} is uniquely determined up to a sign.

Additional Identifying Assumptions

- In this case, suppose, for example, we let

$$\underset{m\times m}{S} = diag\,(1, .., 1. - 1., 1..., 1)$$

i.e., a diagonal matrix with ones in all the diagonal positions except for the i^{th} position which has -1 instead. Then,

$$\begin{aligned}\underset{m\times m}{\Gamma} &= \Lambda^{0\prime}\Psi^{-1}\Lambda^{0} = S\Lambda^{0\prime}\Psi^{-1}\Lambda^{0}S, \\ \Sigma &= \Lambda^{0}\Lambda^{0\prime} + \Psi = \Lambda^{0}S^{2}\Lambda^{0\prime} + \Psi\end{aligned}$$

- This indeterminancy in the sign can be removed by restricting the elements in the first row of Λ^0 to be positive.

Maximum Likelihood Estimation of Exact Factor Models

- Suppose we make a multivariate Gaussian assumption

$$\begin{pmatrix} F_t \\ X_t \end{pmatrix} \equiv i.i.d.N\left(0, \Sigma\right),$$

where we can partition Σ conformably with $\left(F_t', X_t'\right)'$ as

$$\Sigma = \begin{pmatrix} I & \Lambda' \\ \Lambda & \Lambda\Lambda' + \Psi \end{pmatrix}.$$

By well-known property of the multivariate normal, we have

$$F_t | X_t \sim N\left(\Lambda' \left[\Lambda\Lambda' + \Psi\right]^{-1} X_t, I_m - \Lambda' \left[\Lambda\Lambda' + \Psi\right]^{-1} \Lambda\right)$$

Direct Approach to ML Estimation of Exact Factor Models

- Now, suppose we take a direct approach to maximum likelihood estimation of the parameters of the exact factor model. Then, note that since $\{F_t\}$ is not observed, we would need maximize the log-likelihood function

$$\begin{aligned}\ell(\theta; X) &= \ln p(X|\theta) \\ &= \text{const} - \frac{n}{2}\ln\left|\Lambda\Lambda' + \Psi\right| - \sum_{t=1}^{n} X_t'\left(\Lambda\Lambda' + \Psi\right)^{-1} X_t\end{aligned}$$

where

$$X = \begin{pmatrix} X_1' \\ X_2' \\ \vdots \\ X_n' \end{pmatrix},\ \theta = \begin{pmatrix} vec(\Lambda) \\ diag(\Psi) \end{pmatrix},\ diag(\Psi) = \begin{pmatrix} \psi_{11} \\ \psi_{22} \\ \vdots \\ \psi_{pp} \end{pmatrix}.$$

Note, however, this log-likehood is likely to be very difficult to maximize.

Alternative Approach to ML Estimation of Exact Factor Models Based on the EM Algorithm

- .Recall that the EM algorithm proceeds as follows:

(i.) Write out the complete log-likelihood ℓ_c. Note that in this case the missing data is $F = (F_1, .., F_n)'$. so let

$$W = (X, F)$$

(ii.) **E-Step:** Calculate

$$Q\left(\theta', \widehat{\theta}^{(k-1)}\right) = E\left[\ell_c\left(\theta', W\right)|X, \widehat{\theta}^{(k-1)}\right].$$

Alternative Approach to ML Estimation of Exact Factor Models Based on the EM Algorithm

(ii.) **E-Step (con't):** In this case, we can show that

$$
\begin{aligned}
& Q\left(\theta', \widehat{\theta}^{(k-1)}\right) \\
= \quad & \text{const} - \frac{n}{2} \ln |\Psi| - \frac{1}{2} \sum_{t=1}^{n} E\left[F_t' F_t | X_t, \widehat{\theta}^{(k-1)}\right] \\
& - \frac{n}{2} tr \left\{ \left[\frac{1}{n} \sum_{t=1}^{n} \left(X_t X_t' - \Lambda E\left[F_t | X_t, \widehat{\theta}^{(k-1)}\right] X_t' \right.\right.\right. \\
& \qquad \left.\left.\left. - X_t E\left[F_t' | X_t, \widehat{\theta}^{(k-1)}\right] \Lambda' + \Lambda E\left[F_t F_t' | X_t, \widehat{\theta}^{(k-1)}\right] \Lambda' \right) \right] \Psi^{-1} \right\}
\end{aligned}
$$

where

$$
E\left[F_t | X_t, \widehat{\theta}^{(k-1)}\right] = \widehat{\Lambda}^{(k-1)\prime} \left[\widehat{\Lambda}^{(k-1)} \widehat{\Lambda}^{(k-1)\prime} + \widehat{\Psi}^{(k-1)}\right]^{-1} X_t.
$$

Alternative Approach to ML Estimation of Exact Factor Models Based on the EM Algorithm

(ii.) **E-Step (con't):** In addition,

$$
\begin{aligned}
& E\left[F_t F_t' | X_t, \widehat{\theta}^{(k-1)}\right] \\
= \; & I_m - \widehat{\Lambda}^{(k-1)\prime}\left[\widehat{\Lambda}^{(k-1)}\widehat{\Lambda}^{(k-1)\prime} + \widehat{\Psi}^{(k-1)}\right]^{-1}\widehat{\Lambda}^{(k-1)} \\
& + \widehat{\Lambda}^{(k-1)\prime}\left[\widehat{\Lambda}^{(k-1)}\widehat{\Lambda}^{(k-1)\prime} + \widehat{\Psi}^{(k-1)}\right]^{-1} X_t X_t' \\
& \qquad \times \left[\widehat{\Lambda}^{(k-1)}\widehat{\Lambda}^{(k-1)\prime} + \widehat{\Psi}^{(k-1)}\right]^{-1}\widehat{\Lambda}^{(k-1)}.
\end{aligned}
$$

Alternative Approach to ML Estimation of Exact Factor Models Based on the EM Algorithm

(iii) **M-Step:** Maximize $Q\left(\theta', \widehat{\theta}^{(k-1)}\right)$ as a function of the dummy argument θ', i.e., determine

$$\widehat{\theta}^{(k)} = \arg\max_{\theta'} Q\left(\theta', \widehat{\theta}^{(k-1)}\right).$$

Remark: Note, in particular, that maximizing $Q\left(\theta', \widehat{\theta}^{(k-1)}\right)$ with respect to Λ here is essentially the same as doing maximum likelihood for a multivariate linear regression under Gaussian errors. Hence, we obtain

$$\widehat{\Lambda}^{(k)} = \left(\sum_{t=1}^{n} X_t E\left[F_t' | X_t, \widehat{\theta}^{(k-1)}\right]\right)\left(\sum_{t=1}^{n} E\left[F_t F_t' | X_t, \widehat{\theta}^{(k-1)}\right]\right).$$

(iv) Iterate between the E-step and the M-step until convergence.

Generalized Factor Model

- .Bai and Ng (2002) and Bai (2003) studied the following generalized factor model

$$X_{it} = \underset{1\times m}{\lambda_i^{0\prime}} \underset{m\times 1}{F_t^0} + e_{it} = c_{it} + e_{it}, \text{ where } i = 1, ..., N;\ t = 1, ..., T$$

- Stacking the observations, we can obtain the representation

$$\underset{T\times N}{X} = \underset{T\times m}{F^0} \underset{m\times N}{\Lambda^{0\prime}} + \underset{T\times N}{e}$$

where $\underset{N\times m}{\Lambda^0} = \begin{pmatrix} \lambda_1^0 & \lambda_2^0 & \cdots & \lambda_N^0 \end{pmatrix}'$,

$$X = \begin{pmatrix} X_{N,1}' \\ X_{N,2}' \\ \vdots \\ X_{N,T}' \end{pmatrix}, \underset{N\times 1}{X_{N,t}} = \begin{pmatrix} X_{1,t} \\ X_{2,t} \\ \vdots \\ X_{N,t} \end{pmatrix}, \text{ and } \underset{T\times m}{F^0} = \begin{pmatrix} F_1^{0\prime} \\ F_2^{0\prime} \\ \vdots \\ F_T^{0\prime} \end{pmatrix}.$$

- **Remark:** As we will discuss in more details below, in addition to allowing for more general assumptions on the error term $\{e_{it}\}$, Bai and Ng (2002) and Bai (2003) also consider the case where N is large.

Assumptions on the Generalized Factor Model (from Bai and Ng, 2002, and Bai, 2003)

1. **Factors:** There exists a positive constant M such that

$$E\left\|F_t^0\right\|^4 \leq M < \infty$$

and

$$\frac{1}{T}\sum_{t=1}^{T} F_t^0 F_t^{0\prime} \xrightarrow{p} \Sigma_F > 0.$$

2. **Factor Loadings:** There exists a positive constant $\overline{\lambda}$ such that

$$\left\|\lambda_i^0\right\| \leq \overline{\lambda} < \infty \; \forall i$$

and

$$\left\|\frac{\Lambda^{0\prime}\Lambda^0}{N} - \Sigma_\Lambda\right\| \to 0 \text{ for some } \Sigma_\Lambda > 0.$$

Assumptions on the Generalized Factor Model (from Bai and Ng, 2002, and Bai, 2003)

3. **Time and Cross-Sectional Dependence and Heterogeneity:** There exists a positive constant M such that $\forall N, T$ the following conditions hold.

(a)
$$E\left[e_{it}\right] = 0,\ E\left[e_{it}^{8}\right] \leq M < \infty$$

(b) Let
$$\gamma_N\left(s,t\right) = \frac{1}{N}\sum_{i=1}^{N} E\left[e_{is}e_{it}\right]$$

and
$$\max_{1\leq t\leq T}\sum_{s=1}^{T}\left|\gamma_N\left(s,t\right)\right| \leq M < \infty$$

(**Note:** This latter condition puts restriction on the amount of temporal dependence.)

Assumptions on the Generalized Factor Model (from Bai and Ng, 2002, and Bai, 2003)

3. **Time and Cross-Sectional Dependence and Heterogeneity (con't):** There exists a positive constant M such that $\forall N, T$ the following conditions hold.

(c) Let

$$\tau_{ij,t} = E\left[e_{it}e_{jt}\right]$$

and assume that

$$\left|\tau_{ij,t}\right| \leq \left|\tau_{ij}\right| \ \forall t$$

for some τ_{ij} and

$$\max_{1 \leq i \leq N} \sum_{j=1}^{N} \left|\tau_{ij}\right| \leq M < \infty.$$

(**Note:** This condition puts restriction on the amount of cross-sectional dependence.)

Some Remarks

1. Note that the factor models studied in Bai and Ng (2002) and Bai (2003) are large dimensional since $X_{\cdot t}$ is $N \times 1$ for $t = 1, .., T$, and both N and T are allowed to approach infinity.
2. Consider the special case where there is no temporal dependence and heterogeneity. Let

$$E\left[e_{\cdot t} e_{\cdot t}'\right] = \Gamma_N^e$$

and note that in this case

$$\begin{aligned} \lambda_{\max}\left(\Gamma_N^e\right) &= \sqrt{\lambda_{\max}\left(\Gamma_N^{e\prime}\Gamma_N^e\right)} \\ &\leq \sqrt{\left\|\Gamma_N^e\right\|_1 \left\|\Gamma_N^e\right\|_\infty} \\ &= \sqrt{\left(\max_{1\leq j\leq N} \sum_{i=1}^N \left|\tau_{ij}\right|\right)\left(\max_{1\leq i\leq N} \sum_{j=1}^N \left|\tau_{ij}\right|\right)} \\ &\leq M \end{aligned}$$

so that $\lambda_{\max}\left(\Gamma_N^e\right) = O(1)$.

Some Remarks

3. Suppose that $\left\{F_t^0\right\} \equiv i.i.d.\left(0, I_m\right)$; then,

$$\Lambda^0 E\left[F_t^0 F_t^{0\prime}\right] \Lambda^{0\prime} = \Lambda^0 \Lambda^{0\prime}$$

so that the assumption that

$$\left\| \frac{\Lambda^{0\prime}\Lambda^0}{N} - \Sigma_\Lambda \right\| \to 0 \text{ for some } \Sigma_\Lambda > 0$$

implies that

$$\lambda_{\min}\left(\Lambda^{0\prime}\Lambda^0\right) \sim N.$$

More Notations

- We can stack the observations to obtain

$$\underset{T\times N}{X} = \underset{T\times m}{F^0}\underset{m\times N}{\Lambda^{0\prime}} + \underset{T\times N}{e}$$

where

$$\underset{T\times N}{X} = \begin{pmatrix} X'_{\cdot 1} \\ X'_{\cdot 2} \\ \vdots \\ X'_{\cdot T} \end{pmatrix}, \; X_{\cdot t} = \begin{pmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{Nt} \end{pmatrix}, \; \underset{T\times m}{F^0} = \begin{pmatrix} F_1^{0\prime} \\ F_2^{0\prime} \\ \vdots \\ F_T^{0\prime} \end{pmatrix},$$

$$\underset{N\times m}{\Lambda^0} = \begin{pmatrix} \lambda_1^{0\prime} \\ \lambda_2^{0\prime} \\ \vdots \\ \lambda_T^{0\prime} \end{pmatrix}.$$

- For estimation, we want to choose $\widehat{\Lambda}, \widehat{F}_1, ..., \widehat{F}_T$ so as to minimize

$$
\begin{aligned}
& Q_{NT}\left(\Lambda, F_1, ..., F_T\right) \\
= \; & \frac{1}{NT}\sum_{t=1}^{T}\left(X_{\cdot t} - \Lambda F_t\right)'\left(X_{\cdot t} - \Lambda F_t\right) \\
= \; & \frac{1}{NT}\sum_{t=1}^{T} tr\left\{\left(X_{\cdot t} - \Lambda F_t\right)\left(X_{\cdot t} - \Lambda F_t\right)'\right\} \\
= \; & tr\left\{\frac{1}{NT}\sum_{t=1}^{T}\left(X_{\cdot t}X_{\cdot t}' - \Lambda F_t X_{\cdot t}' - X_{\cdot t}F_t'\Lambda' + \Lambda F_t F_t'\Lambda'\right)\right\} \\
= \; & tr\left\{\frac{X'X}{NT} - \frac{\Lambda F'X}{NT} - \frac{X'F\Lambda'}{NT} + \frac{\Lambda F'F\Lambda'}{NT}\right\}
\end{aligned}
$$

subject to the constraint $F'F/T = I_m$.

Estimation

- Next, note that, proceeding as if F is observed, we obtain

$$\widetilde{\Lambda}' = \left(F'F\right)^{-1} F'X.$$

- Concentrating the above objective function by evaluating it at $\Lambda = \widetilde{\Lambda}$, we obtain

$$\begin{aligned} & \widetilde{Q}_{NT}\left(F\right) \\ = \quad & tr\left\{ \frac{X'X}{NT} - \frac{\widetilde{\Lambda}F'X}{NT} - \frac{X'F\widetilde{\Lambda}'}{NT} + \frac{\widetilde{\Lambda}F'F\widetilde{\Lambda}'}{NT} \right\} \\ = \quad & tr\left\{ \frac{X'X}{NT} - \frac{X'F\left(F'F\right)^{-1}F'X}{NT} \right\} \\ = \quad & tr\left\{ \frac{X'M_F X}{NT} \right\} \end{aligned}$$

where $M_F = I_T - F\left(F'F\right)^{-1} F'$.

- Now, observe that minimizing $\widetilde{Q}_{NT}(F)$ w.r.t. F is the same as maximizing

$$\widehat{Q}_{NT}(F) = tr\left\{\frac{X'P_F X}{NT}\right\} \quad \left(\text{where } P_F = F\left(F'F\right)^{-1}F'\right)$$
$$= tr\left\{\left(F'F\right)^{-1/2}F'\frac{XX'}{NT}F\left(F'F\right)^{-1/2}\right\}$$

with respect to F. To solve this maximization problem, we consider the spectral decomposition

$$\frac{XX'}{NT} = CD_l C',$$

where

$$\underset{T\times T}{C} \in \mathcal{O}(T), \text{ i.e., } C'C = CC' = I_T$$

and $D_l = diag(l_1, l_2, ..., l_T)$.

Estimation

- Without loss of generality, we assume the ordering

$$l_1 > l_2 > \cdots > l_T$$

  noting that for continuously distributed X, the eigenvalues would differ with probability one.

- Next, we partition

$$\underset{T\times T}{C} = \left[\begin{array}{cc} \underset{T\times m}{C_1} & \underset{T\times(T-m)}{C_2} \end{array} \right]$$

  so that the columns of C_1 are the eigenvectors corresponding to the m largest eigenvalues.

Estimation

- We choose
$$\widehat{F} = \sqrt{T} C_1$$
and note that
$$\frac{\widehat{F}'\widehat{F}}{T} = \frac{\sqrt{T} C_1' C_1 \sqrt{T}}{T} = I_m,$$
so that this choice satisfies our normalization on the factors.

Estimation

- We now take

$$
\begin{aligned}
\underset{N\times m}{\widehat{\Lambda}} &= X'\widehat{F}\left(\widehat{F}'\widehat{F}\right)^{-1} \\
&= \frac{X'\widehat{F}}{T}\left(\frac{\widehat{F}'\widehat{F}}{T}\right)^{-1} \\
&= \frac{X'\widehat{F}}{T} \quad \left(\text{given that } \frac{\widehat{F}'\widehat{F}}{T} = I_m\right) \\
&= \frac{X'C_1\sqrt{T}}{T} \\
&= \frac{X'C_1}{\sqrt{T}}.
\end{aligned}
$$

Estimation

- To relate this estimator to principal component analysis, note that

$$e_{1,m}'\widehat{\Lambda}' = \frac{e_{1,m}'C_1'X}{\sqrt{T}} = \frac{c_{11}'X}{\sqrt{T}},$$

where c_{11} is the first column of C_1, i.e., the eigenvector associated with the largest eigenvalue λ_1 of $XX'/(NT)$.

Asymptotic Results

- Under Assumptions 1-3 and some additional conditions, there exists a nonsingular $m \times m$ matrix H such that

(i)

$$\sqrt{N}\left(\widehat{F}_t - H' F_t^0\right) \xrightarrow{d} N\left(0, V^{-1} Q \Gamma_t Q' V^{-1}\right) \text{ for each } t$$

if $\sqrt{N}/T \to 0$ as $N, T \to \infty$. Here,

$$\begin{aligned} Q &= p \lim_{N,T\to\infty} \frac{\widehat{F}' F^0}{T}, \\ \Gamma_t &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i^0 \lambda_j^{0\prime} E\left[e_{it} e_{jt}\right], \\ V &= diag\left(v_1, ..., v_m\right) \end{aligned}$$

where $v_1 > \cdots > v_m > 0$ are the eigenvalues of $\Sigma_\Lambda^{1/2} \Sigma_F \Sigma_\Lambda^{1/2}$.

(ii)

$$\sqrt{T}\left(\widehat{\lambda}_i - H^{-1}\lambda_i^0\right) \xrightarrow{d} N\left(0, \left(Q'\right)^{-1}\Phi_i Q^{-1}\right) \text{ for each } i$$

if $\sqrt{T}/N \to 0$ as $N, T \to \infty$. Here,

$$\Phi_i = \lim_{T\to\infty} \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T} E\left[F_s^0 F_t^{0\prime} e_{is} e_{it}\right].$$

- An explicit form can be found for H as

$$H = \left(\frac{\Lambda^{0\prime}\Lambda^0}{N}\right)\frac{F^{0\prime}\widehat{F}}{T}\widetilde{V}_{NT}^{-1}$$

where $\widetilde{V}_{NT}$ is an $m \times m$ matrix containing the m largest eigenvalues of

$$\frac{XX'}{NT} \;\; (T \times T)$$