Paper Club

Feb 18, 2022

Can Wu

Factor Modeling for High-Dimensional Time Series

I. Introduction

AR(1)

Time series: $Y_t = \Phi Y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim N(0, \Sigma_\varepsilon)$

$Y_t \in \mathbb{R}^p$, $\Phi \in \mathbb{R}^{p \times p}$, $\varepsilon_t \in \mathbb{R}^p$

Factor Model: $Y_t = \Lambda F_t + \varepsilon_t$, ← idiosyncratic term

Σ_t, diagonal

$F_t \in \mathbb{R}^r$, r small, $\Lambda \in \mathbb{R}^{p \times r}$, $\varepsilon_t \in \mathbb{R}^p$, $Var(\varepsilon_t) = \Sigma_t$.

$\varepsilon_t \sim$ white noise, $\mathbb{E}(\varepsilon_t) = 0$, $cov(\varepsilon_t, \varepsilon_s) = 0$, $\forall t \neq s$.

static model

$$Y_t = \Lambda F_t + \varepsilon_t$$

$$Y' = \Lambda F + \varepsilon, \quad \in \mathbb{R}^{p\times n}$$

$$Y' = (Y_1, Y_2, \cdots, Y_n) = \begin{pmatrix} y_1' \\ \vdots \\ y_p' \end{pmatrix}$$

y_i : component of Y.

$$F = (F_1, F_2, \cdots, F_n) \in \mathbb{R}^{r\times n}$$

$$\varepsilon = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) \in \mathbb{R}^{p\times n}$$

dynamic model

$$Y_t = \Lambda_0 F_t + \Lambda_1 F_{t-1} + \cdots + \Lambda_s F_{t-s} + \varepsilon_t, \quad AR(s)$$

$$= \Lambda \mathcal{F}_t + \varepsilon_t, \quad \text{where } \underset{\in \mathbb{R}^{p\times r(s+1)}}{\Lambda = (\Lambda_0, \cdots, \Lambda_s)}, \quad \underset{\in \mathbb{R}^{r(s+1)}}{\mathcal{F}_t = (F_t', F_{t-1}', \cdots, F_{t-s}')'}$$

II. Identifiability

$$Y' = (\Lambda H)(H^{-1}F) + \varepsilon, \quad H \in \mathbb{R}^{r\times r}$$

Need $r\times r$ constraints.

$$\left\{ \begin{array}{l} \text{(i)} \ \Lambda'\Lambda = I_r, \qquad \frac{r(r+1)}{2} \\ \text{(ii)} \ \lambda_{ij} = 0 \ \text{for} \ 1 \le i < j \le n \qquad \frac{r(r-1)}{2} \\ \text{(iii)} \ \lambda_{ii} > 0, \qquad 1 \le i \le n. \end{array} \right.$$

$(\Lambda'\Lambda)_{ii} = 1$

$(\Lambda'\Lambda)_{ij} = 0, \ i<j. \quad \frac{r(r-1)}{2}$

(Λ_1, F_1)

(Λ_0, F_0)

$Y' = \Lambda_0 F_0 + \varepsilon.$

$(\boxed{\Lambda_0 H}, \boxed{H^{-1}F_0})$

(i) $H'\Lambda_0'\Lambda_0 H = I_r$

r

(ii) $(\Lambda_0 H)_{ij} = 0$

(iii) $(\Lambda_0 H)_{ii} > 0$

$$\Lambda'\Lambda = \begin{pmatrix} \lambda_1' \\ \lambda_2' \\ \vdots \\ \lambda_r' \end{pmatrix} \underbrace{(\lambda_1 \ \lambda_2 \cdots \lambda_r)}_{\lambda_i \in \mathbb{R}^{p\times 1}} = \begin{pmatrix} \lambda_1'\lambda_1 & \lambda_1'\lambda_2 & \cdots & \lambda_1'\lambda_r \\ \lambda_2'\lambda_1 & \lambda_2'\lambda_2 & \cdots & \lambda_2'\lambda_r \\ \vdots & \vdots & \ddots & \\ \lambda_r'\lambda_1 & \cdots & \cdots & \lambda_r'\lambda_r \end{pmatrix}$$

$$\underbrace{\{(\Lambda H, H^{-1}F)\}}_{A} \xRightarrow{r^2 \text{ constraints}} \{\underbrace{(\Lambda_0, F_0)}\}$$

implicitly determine H

Identifiability

$\hat{\Lambda} H \to \Lambda^*$

rotation matrix

III. Estimation

1. Least-squares, PCA

$$S(\Lambda, F) = \sum_{t=1}^{n} \| Y_t - \Lambda F_t \|^2$$

$$= \sum_{t=1}^{n} \varepsilon_t' \varepsilon_t$$

$$= \operatorname{tr}(\varepsilon' \varepsilon), \quad \varepsilon = (\varepsilon_1 \ \varepsilon_2 \cdots \varepsilon_n) \in \mathbb{R}^{p \times n}$$

$$= \operatorname{tr}\left((Y' - \Lambda F)' (Y' - \Lambda F) \right)$$

Let $(\hat{\Lambda}, \hat{F})$ be the minimizer. $\left.\frac{\partial S}{\partial F}\right|_{\substack{\Lambda = \hat{\Lambda} \\ F = \hat{F}}} = 0,$

$$-2 \hat{\Lambda}' Y' + 2 \hat{\Lambda}' \hat{\Lambda} \hat{F} = 0$$

Using identifiability condition, $\hat{\Lambda}' \hat{\Lambda} = I_r$, $\underset{\hat{}}{\hat{F}} = \underset{\hat{}}{\hat{\Lambda}}' Y'$.

$$S(\hat{\Lambda}, \hat{F}) = \operatorname{tr}(Y Y') - \operatorname{tr}(\hat{\Lambda}' Y' Y \hat{\Lambda}).$$

(i)

$\Rightarrow$ $\hat{\Lambda}$ maximizes $\operatorname{tr}(\hat{\Lambda}' Y' Y \hat{\Lambda})$. $\Rightarrow$ $\hat{\Lambda} = (V_1 \cdots V_r)$,

V_i eigenvectors for i-th largest eigenvalues of $Y'Y$.

2 Factor loading space estimation

First estimate column space of Λ,

$$\Sigma_y(k) = \mathrm{cov}(Y_{t+k}, Y_t) = \Lambda\, \mathrm{cov}(F_{t+k}, F_t)\Lambda' = \Lambda \boxed{\Sigma_f(k)\Lambda'}$$

$\Rightarrow \Sigma_y(k)$ lies in column space of Λ. $\quad \Sigma_y(k) = \Lambda G$.

$$M := \sum_{k=1}^{k_0} \Sigma_y(k)\Sigma_y'(k) = \Lambda\left(\sum_{k=1}^{k_0} \Sigma_f(k)\Sigma_f'(k)\right)\Lambda'$$

$= (m_1 \cdots m_r)$

$M = \Lambda G$

$= (\lambda_1 \cdots \lambda_r)\begin{pmatrix} g_{11} & & \\ & \ddots & \\ & & g_{rr} \end{pmatrix}$

$m_1 = g_{11}\lambda_1 + \cdots + g_{r1}\lambda_r$

$\Rightarrow$ column space of M = column space of Λ.

B_M = basis of $\mathrm{col}(M)$.

B_Λ = basis of $\mathrm{col}(\Lambda)$.

$B_M = H B_\Lambda$. H invertible

$\{m\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\}$ $\{n\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\}$

$\subseteq$, $\supseteq$ $\Rightarrow \sum_{k=1}^{k_0} \Sigma_f(k)\Sigma_f'(k)$ p.s.d.

$\hat{\Lambda} = (V_1 \cdots V_r)$. V_i eigenvector for the i-th largest eigenvalues of $\hat{M}$.

3. Likelihood-based estimation

$$\Phi(B) F_t = \Theta(B) Z_t, \quad \leftarrow \text{Vector ARMA}(\tilde{p},\tilde{q}) \text{ model}$$

$\Phi(B) = I - \Phi_1 B - \cdots - \Phi_{\tilde{p}} B^{\tilde{p}}$, B: backshift operator

$\Theta(B) = I + \Theta_1 B + \cdots + \Theta_{\tilde{q}} B^{\tilde{q}}$, $BY_t = Y_{t-1}$

$Z_t \sim N(0, \Sigma_Z)$, $Z_t \in \mathbb{R}^r$,

$\tilde{p} < \tilde{s}$

3.1. Kalman filtering $\quad Y' = \Lambda F + \mathcal{E}$

Also assume $\boxed{\mathcal{E}_t = \Psi_1 \mathcal{E}_{t-1} + \cdots + \Psi_{\tilde{s}} \mathcal{E}_{t-\tilde{s}} + e_t}$, $e_t \overset{iid.}{\sim} N(0, \Sigma_e)$ white noise

$\boxed{\Psi(B)\mathcal{E}_t = e_t}$, $\Psi(B) = I - \Psi_1 B - \cdots - \Psi_{\tilde{s}} B^{\tilde{s}}$.

$$\alpha_t = \begin{pmatrix} F_t \\ F_{t-1} \\ \vdots \\ F_{t-\tilde{p}+1} \end{pmatrix} = \begin{pmatrix} \Phi_1 & \Phi_2 & & \Phi_{\tilde{p}} \\ I_r & & & \\ & \ddots & & \\ & & I_r & \end{pmatrix} \begin{pmatrix} F_{t-1} \\ F_{t-2} \\ \vdots \\ F_{t-\tilde{p}} \end{pmatrix} + \begin{pmatrix} I_r & \Theta_1 & \cdots & \Theta_{\tilde{q}} \\ & O & & \end{pmatrix} \begin{pmatrix} Z_t \\ \vdots \\ Z_{t-\tilde{q}} \end{pmatrix}$$

$$= H\alpha_{t-1} + R\eta_t,$$

$$\boxed{\alpha_t = H\alpha_{t-1} + R\eta_t,}$$

$$Y_t = \Lambda F_t + \varepsilon_t, \qquad \Psi(B)$$

$$\Psi(B) Y_t = \Lambda \Psi(B) F_t + \Psi(B)\varepsilon_t \qquad \alpha_t = \begin{pmatrix} F_t \\ \vdots \\ F_{t-\tilde{p}+1} \end{pmatrix} \quad \tilde{p}$$

$$= Z\alpha_t + e_t$$

WLOG $\boxed{\tilde{p} \geq \tilde{s}}$

$$Z = (\Lambda \quad -\Psi_1\Lambda \quad \cdots \quad -\Psi_{\tilde{s}}\Lambda \quad 0) = (\bar{\Lambda} \quad 0) \quad A$$

$$\Lambda \Psi(B) F_t \neq A(F_t - \Psi_1 F_{t-1} - \cdots - \Psi_{\tilde{s}} F_{t-\tilde{s}}) \quad \tilde{s}$$

$F_t \quad F_{t-1} \quad \Delta F_{t-2} \cdots$

$A F_{t-s+1}$

$\boxed{\text{state space model}}$

$\Rightarrow \hat{F}$