Dynamic Factor Model for Matrix Time Series

Ruofan Yu
Rutgers University

Outline

1. Introduction
    1. Matrix time series
    2. Review of matrix time series models
2. Dynamic matrix factor models
3. Estimation and prediction of the model
4. Real examples

Joint work with Yuefeng Han, Rong Chen and Han Xiao

Introduction: Matrix Time Series

Matrix Time Series

A sequence of observations $\mathbf{X_t}$ in matrix form observed over time

where

$$\boldsymbol{X}_t = \begin{pmatrix} x_{t11} & \cdots & x_{t1n} \\ \vdots & \ddots & \vdots \\ x_{tm1} & \cdots & x_{tmn} \end{pmatrix}$$

Examples:

- A set of economic indicators of a group of countries observed over time;
- A set of financial characteristics of a group of companies observed over time;
- Import-Export among a group of countries observed over time;
- A sequence of time-indexed images.

Review of Matrix Time Series Models

Matrix Autoregressive Model (Chen, Xiao and Yang, 2021+, JoE)

$$\boldsymbol{X}_t = \boldsymbol{A}\boldsymbol{X}_{t-1}\boldsymbol{B}' + \boldsymbol{E}_t.$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are $m \times m$ and $n \times n$ coefficient matrices. $\boldsymbol{E}_t$ is a $m \times n$ series of (matrix) white noise.

Matrix Factor Model (Wang, Liu, and Chen (2019, JoE))

$$\boldsymbol{X}_t = \boldsymbol{U}_1\boldsymbol{F}_t\boldsymbol{U}_2' + \boldsymbol{E}_t$$

where $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ are $d_1 \times k_1$ and $d_2 \times k_2$ orthonormal loading matrices. $\boldsymbol{F}_t$ $(k_1 \times k_2)$ unobserved factor matrix. $(k_1 << d_1$ and $k_2 << d_2)$. $\boldsymbol{E}_t$ is a $d_1 \times d_2$ series of (matrix) white noise.

Dynamic Matrix Factor Model

$$
\begin{aligned}
\boldsymbol{X}_t &= \lambda \boldsymbol{U}_1 \boldsymbol{F}_t \boldsymbol{U}_2^\top + \boldsymbol{E}_t \\
\boldsymbol{F}_t &= \boldsymbol{A}_1 \boldsymbol{F}_{t-1} \boldsymbol{A}_2^\top + \boldsymbol{Z}_t
\end{aligned}
$$

where $\boldsymbol{E}_t$ and $\boldsymbol{Z}_t$ are independent matrix white noise processes. $||\boldsymbol{U}_1||_F = ||\boldsymbol{U}_2||_F = 1$ and orthogonal.

Why?

- Adding prediction capability to matrix factor model;
- Reducing dimension for matrix AR model :
  - Matrix AR involves $2d^2$ parameters, here only $2dr + 2r^2$
- More detailed dynamic structure in low dimension;
- Can be extended to allow more flexible dynamic structure, as $\boldsymbol{F}_t$ is of much lower dimension:
  - $vec(\boldsymbol{F}_t) \sim VAR$
  - Each f_{ijt} follows univariate AR model, and independent to each other

Prediction

Let $\mathcal{F}_t$ be the σ-field generated by $\boldsymbol{X}_t, \ldots, \boldsymbol{X}_1$. Under the model:

$$E(\boldsymbol{X}_{t+1} \mid \mathcal{F}_t) = \lambda \boldsymbol{U}_1 \boldsymbol{F}^*_{t+1} \boldsymbol{U}_2^\top.$$

where $\boldsymbol{F}^*_{t+1} = E[\boldsymbol{F}_{t+1} \mid \mathcal{F}_t] = \boldsymbol{A}_1 \boldsymbol{F}_t \boldsymbol{A}_2^\top$.

Hence the least squares predictor is

$$\begin{aligned}
\hat{\boldsymbol{X}}_{t+1} &= \lambda \boldsymbol{U}_1 \hat{\boldsymbol{F}}^*_{t+1} \boldsymbol{U}_2^\top \\
&= \lambda \boldsymbol{U}_1 \boldsymbol{A}_1 \hat{\boldsymbol{F}}_t \boldsymbol{A}_2^\top \boldsymbol{U}_2^\top \\
&= \boldsymbol{U}_1 \boldsymbol{A}_1 \boldsymbol{U}_1^\top \boldsymbol{X}_t \boldsymbol{U}_2 \boldsymbol{A}_2^\top \boldsymbol{U}_2^\top.
\end{aligned}$$

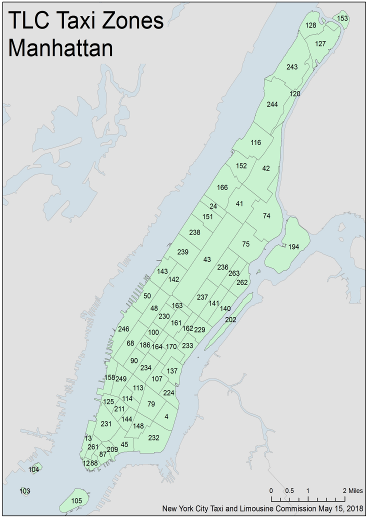
- It is in a Matrix AR(1) form – though the model is not a MAR(1).
- The coefficient matrices $\boldsymbol{U}_1 \boldsymbol{A}_1 \boldsymbol{U}_1^\top$ and $\boldsymbol{U}_2 \boldsymbol{A}_2^\top \boldsymbol{U}_2^\top$ are of low rank.

Estimation

A two-step estimation procedure through a combination of (an extended version of) PCA and (lower dimensional) LS:

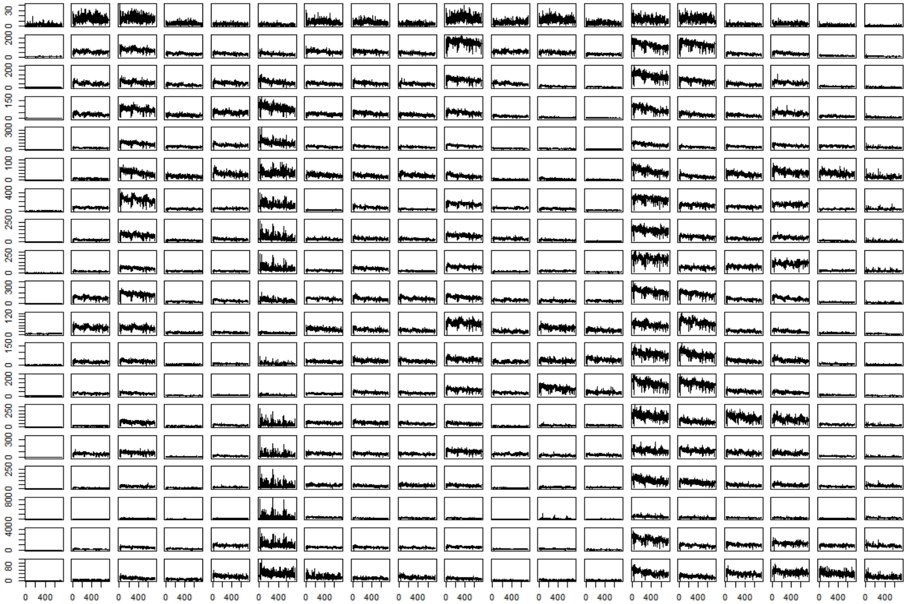
1. Estimate the matrix factor model to obtain $\hat{\boldsymbol{U}}_1, \hat{\boldsymbol{U}}_2, \hat{\boldsymbol{F}}_t$,
2. Estimate a matrix AR model (or VAR, or individual AR), treating $\hat{\boldsymbol{F}}_t$ as true

- Two stage estimation is commonly used to estimate the dynamic factor model in vector case.
- Joint likelihood estimation is possible, through a state-space model approach, with a modified Kalman filter.
    - Poor performance as the number of parameters (the loading matrices and the MAR coefficient matrices and the noise variances).



- The NYC taxi data on Manhattan is tensor data of dimensions $69 \times 69 \times 24 \times T$, which is the taxi counts from one region to another in a certain hour in a day
- Use a subset of the data: Sum of morning rush hours (7-10am) in 19 regions in middle town for years 2015 to 2017. The dimension is $19 \times 19 \times 750$.

Real example : NYC taxi data



Real example : NYC taxi data

Table: Use morning rush hours data to do rolling forecasts

Dynamic Factor		Vector factor		Autoregressive Models	
Method	RMSE	Method	RMSE	Method	RMSE
F.MAR1	**12.68**			X.MAR1	12.71
F.VAR1	12.75	VF.VAR1	12.71	X.VAR1	19.05
F.iAR1	12.81	VF.iAR1	12.74	X.iAR1	12.86
F.rw	14.69	VF.rw	14.9	X.rw	16.12
F.mean	13.34	VF.mean	13.33	X.mean(ES)	13.35

Note: ES means we're using exponential smoothing to do prediction. rw is the random walk prediction, where we use the data at last time point to predict next.

THANK YOU!