UNIVERSITÄT
BIELEFELD
Fakultät für
Wirtschaftswissenschaften

Using Subspace Algorithms for the Estimation of Linear State Space Models in the Context of Approximate Dynamic Factor Models

Dietmar Bauer

1. Motivation

- Approximate dynamic factor models (aDFMs) are used for joint modelling of many time series, wherein the cross sectional dimension N is of the same size as the time series length T.
- In typical applications both range in the low hundreds.
- Classical joint models (like VAR) for such number of series necessarily contain many parameters.
- Factor models reduce the dimensionality:

$$y_{it} = \chi_{it} + \xi_{it} = \underbrace{\lambda_i' F_t}_{\text{factor component}} + \underbrace{\xi_{it}}_{\text{idiosyncratic component}},$$

  for $i = 1, .., N, t = 1, ..., T$..
- Factors and idiosyncratic component are assumed uncorrelated.
- In matrix notation: $Y_t^N = \Lambda_N F_t + \Xi_t^N \in \mathbb{R}^N$.
- I use asymptotic identification of factors and idiosyncratics as in Chamerblain and Rothschild (1983): $\lambda_{max}(\mathbb{E}\,\Xi_t^N(\Xi_t^N)') \leq M_\Xi$ and $\Lambda_N'\Lambda_N/N \overset{N\to\infty}{\to} M_\circ, \mathbb{E}\,F_tF_t' > 0$.

2. Assumptions

Assumption

- The processes $\Delta(F_t)_{t\in\mathbb{Z}}, (\Xi_t^N)_{t\in\mathbb{Z}}$ are wide sense stationary with mean zero for all N and possess spectral densities.
- For $\Delta F_t, \Xi_t^N, y_t^N$ we have uniform (in lag and in N) LIL rates $Q_T := \sqrt{\frac{(\log\log T)}{T}}$ for the estimation of covariances up to lag $H_T = O((\log T)^a)$ for some $a > 1$.
- Factor loadings λ_i are deterministic, $\frac{1}{N}\Lambda_N'\Lambda_N \to M_\circ > 0, \sup_N \max_i \|\lambda_i\| \leq M_\lambda$.
- $\tilde{I}_r'\Lambda_N = I_r$ for a selector matrix (other normalisations are possible.).
- $(F_t)_{t\in\mathbb{Z}}, F_t \in \mathbb{R}^r$, is a rational I(0) or I(1) process, i.e. it has a minimal state space representation as $(q \leq r)$

$$F_t = \underbrace{\begin{pmatrix} C_1 & C_\bullet \end{pmatrix}}_{C} x_t + Du_t, x_{t+1} = \underbrace{\begin{pmatrix} I_c & 0 \\ 0 & A_\bullet \end{pmatrix}}_{A} x_t + \underbrace{\begin{pmatrix} B_1 \\ B_\bullet \end{pmatrix}}_{B} u_t, \quad \begin{matrix} u_t \in \mathbb{R}^q, \\ x_t \in \mathbb{R}^n. \end{matrix}$$

- $\lambda_{|max|}(A_\bullet) < 1$. $[I_q, 0]D$ is positive lower triangular; $\exists D^\dagger : D^\dagger D = I_q$ such that $A - BD^\dagger C$ is stable with $\lambda_{|max|}(A - BD^\dagger C) < \rho_\circ < 1$.
- $(u_t)_{t\in\mathbb{Z}}$ is a stationary, ergodic martingale difference sequence and has expectation zero, variance matrix $\Omega = I_q$ and finite fourth moments.

Canonical Variate Analysis (CVA)

$$F_t = Cx_t + Du_t, \qquad x_{t+1} = Ax_t + Bu_t.$$

$\Rightarrow$ estimation could be done using OLS, if the state x_t and u_t were known!

1. The state x_t can be approximated by past observations:

$$x_t(p) = \sum_{j=1}^{p} \mathcal{K}_j(p) F_{t-j} = \mathcal{K}_p F_t^- \overset{p\to\infty}{\to} x_t$$

2. Predictions of $F_{t+h}, h \geq 0$, based on $F_{t-1}, F_{t-2} \ldots$:

$$F_t^+ = \mathcal{O}x_t + \mathcal{U}U_t^+ = \boxed{\mathcal{O}_f \mathcal{K}_p} F_t^- + \tilde{U}_t. \qquad (*)$$

Main idea:

- Apply CVA with PCA estimated common factors $\hat{F}_t$.
- Estimate the state as $\hat{x}_t = \widehat{\mathcal{K}}_p \hat{F}_t^-$ where $\widehat{\mathcal{K}}$ is obtained from RRR in $(*)$.
- Different weights W_f^+ in the RRR step should be used (see paper)!

Theorem (Consistency)

- Let the data be generated according to the before mentioned Assumptions.
- Assume that $T/N^2 \to 0, T \to \infty$, such that $O(Q_T + 1/N) = O(Q_T)$.
- Let $p(T) \leq H_T, f \geq n, p = p(T) \geq -\frac{e \log T}{\log \rho_0}, e > 1$, for $\rho_\circ > 0$ and $p \geq p_0$ for $\rho_\circ = 0$.
- Further let $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ denote the CVA estimators with appropriate (see paper) weighting matrix W_f^+ in the RRR step based on the PCA estimates $\hat{F}_t$ for given order n converted to an overlapping form.

Then

$$\max\{\|\hat{A} - A\|, \|\hat{B} - B\|, \|\hat{C} - C\|, \|\hat{D} - D\|\} = O(Q_T).$$

References:

- Bauer, D. (2025): Using Subspace Algorithms for the Estimation of Linear State Space Models in the Context of Approximate Dynamic Factor Models, working paper see: version on Github
- Chamberlain, G., Rothschild, M. (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. Econometrica 51 (5), 1281–1304.