Journal of Econometrics 163 (2011) 23–28

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

The general dynamic factor model: One-sided representation results

Mario Forni [a], Marco Lippi [b,*]

[a] Dipartimento di Economia Politica, Università di Modena e Reggio Emilia, CEPR and RECent, Italy

[b] Dipartimento di Economia, Università di Roma La Sapienza, and EIEF, Italy

ARTICLE INFO

Article history:
Available online 12 November 2010

JEL classification:
E0
C1

Keywords:
Dynamic factor models
Frequency domain approach
Estimation by one-sided filters

ABSTRACT

Recent dynamic factor models have been almost exclusively developed under the assumption that the common components span a finite-dimensional vector space. However, this finite-dimension assumption rules out very simple factor-loading patterns and is therefore severely restrictive. The general case has been studied, using a frequency domain approach, in Forni et al. (2000). That paper produces an estimator of the common components that is consistent but is based on filters that are two-sided and therefore unsuitable for prediction. The present paper, assuming a rational spectral density for the common components, obtains a one-sided estimator without the finite-dimension assumption.

1. Introduction

The dynamic factor model

$$x_{it} = \chi_{it} + \xi_{it} = b_{i1}(L)u_{1t} + b_{i2}(L)u_{2t} + \cdots + b_{iq}(L)u_{qt} + \xi_{it}, \tag{1.1}$$

where $i \in \mathbb{N}$, $t \in \mathbb{Z}$, has been studied in a vast literature starting with Stock and Watson (2002a,b), Forni et al. (2000) and Forni and Lippi (2001).

The components ξ_{it}, called *idiosyncratic*, are assumed to be orthogonal to the *common* components χ_{it} and cross-sectionally weakly correlated (see Section 2), so the comovement of the x's is mainly accounted for by the q *common shocks* u_{jt}. Usually, the assumptions also include that the Hilbert space spanned by the common components χ_{it}, *for a given* t and $i \in \mathbb{N}$, is finite dimensional. Under this assumption, the components χ_{it} and ξ_{it} can be consistently estimated, as n and T (the number of series and the number of observations for each series, respectively) tend to infinity, using principal components (standard or generalized) of the observable series x_{it} (see Stock and Watson, 2002a,b; Bai and Ng, 2002; Forni et al., 2005, 2009). Moreover, these estimators only involve present and past values of the variables x_{it}.

Dynamic principal components, based on the spectral density of the x's, have been used in Forni et al. (2000), where the above mentioned finite-dimension assumption is not required. However, dynamic principal components result in two-sided filters, involving present and past but also future values of the variables x_{it}, with the consequence that the estimates are unreliable at the end of the sample and therefore useless for prediction.

The present paper starts with the observation that the finite-dimension assumption is very strict, as it does not include a model as simple as

$$x_{it} = \frac{1}{1 - \alpha_i L} u_t + \xi_{it}, \tag{1.2}$$

with the coefficients α_i independently drawn, for example, from the uniform distribution between -0.9 and 0.9.

This seems sufficient motivation to go back to model (1.1) without the finite-dimension assumption. Combining the approach taken in Forni et al. (2000) with recent results obtained by Anderson, Deistler and coauthors (see Section 3), we show that under the assumption that the filters $b_{ij}(L)$ are rational, plus reasonable technical assumptions, model (1.1) can be rewritten as

$$H_n(L)\boldsymbol{x}_{nt} = R_n \boldsymbol{u}_t + H_n(L)\boldsymbol{\xi}_{nt}, \tag{1.3}$$

where $\boldsymbol{x}_{nt}$ and $\boldsymbol{\xi}_{nt}$ stack the first n series x_{it} and ξ_{it} respectively, $\boldsymbol{u}_t = (u_{1t}\ u_{2t}\ \cdots\ u_{qt})'$, $H_n(L)$ is a finite matrix polynomial. Moreover:

(i) $H_n(L)$, which is $n \times n$, and R_n, which is $n \times q$, can be obtained from the spectral density of $\boldsymbol{\chi}_{nt}$.
(ii) $H_n(L)\boldsymbol{\xi}_{nt}$ is idiosyncratic (this is not obvious; see Section 4).

Though the paper is limited to representation results, Eq. (1.3), combined with the estimate of the spectral density of $\boldsymbol{\chi}_{nt}$ proposed in Forni et al. (2000), can be seen as a basis for estimating the common components χ_{it}, without the finite-dimension assumption and using only contemporaneous and past values of the series x_{it}.

* Corresponding address: Dipartimento di Economia, Circonvallazione Tiburtina 4, 00185 Roma, Italy. Tel.: +39 0649917062; fax: +39 0649917060.
E-mail address: ml@lippi.ws (M. Lippi).

0304-4076/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2010.11.003

Section 2 reviews previous results on model (1.1). Section 3 introduces and discusses the main assumptions. Section 4 derives representation (1.3). Section 5 discusses estimation based on (1.3). Section 6 concludes.

2. Previous results

2.1. The general model

Let us rewrite model (1.1) in vector form:

$$\begin{aligned} \boldsymbol{x}_{nt} &= \boldsymbol{\chi}_{nt} + \boldsymbol{\xi}_{nt} \\ \boldsymbol{\chi}_{nt} &= B_n(L)\boldsymbol{u}_t \end{aligned} \tag{2.4}$$

with $b_{ij}(L)$ being the (i, j) entry of $B_n(L)$ for all $n \geq i$ (the matrices $B_n(L)$ are nested). We assume that:

A1. (Common components) $\boldsymbol{u}_t$ is an orthonormal q-dimensional white noise. The filters $b_{ij}(L)$ are square summable.
A2. (Idiosyncratic components) $\boldsymbol{\xi}_{nt}$ is weakly stationary.
A3. (Orthogonality of common and idiosyncratic components) $\boldsymbol{\xi}_{nt} \perp \boldsymbol{u}_s$, for all n, t, s.
A4. (Eigenvalues of the idiosyncratic components) Let $\Sigma_n^{\xi}(\theta)$ be the spectral density matrix of $\boldsymbol{\xi}_{nt}$ and $\lambda_{n1}^{\xi}(\theta)$ its first eigenvalue (in descending order). We assume that there exists a positive real number λ such that $\lambda_{n1}^{\xi}(\theta) \leq \lambda$ for all n.
A5. (Eigenvalues of the common components) Let $\Sigma_n^{\chi}(\theta)$ be the spectral density matrix of $\boldsymbol{\chi}_{nt}$ and $\lambda_{nq}^{\chi}(\theta)$ its qth eigenvalue. We assume that $\lambda_{nq}^{\chi}(\theta) \to \infty$, for all θ, for $n \to \infty$.

Forni and Lippi (2001) prove that (2.4) and Assumptions A1 through A5 impose little structure on the x's. They show that the following two assumptions: (1) $\boldsymbol{x}_{nt}$ is stationary for all n, (2) there exists an integer q such that, for $n \to \infty$, the qth eigenvalue of the spectral density matrix of $\boldsymbol{x}_{nt}$ diverges for all frequencies while the $(q + 1)$th is uniformly bounded, imply that the x's can be represented as in (2.4) with A1 through A5 holding.

Under Assumptions A1 through A5, the decomposition of the x's into common and idiosyncratic components is unique. To be precise, if

$$\begin{aligned} x_{it} = \chi_{it}' + \xi_{it}' &= b_{i1}'(L)u_{1t}' + b_{i2}'(L)u_{2t}' \\ &\quad + \cdots + b_{iq}'(L)u_{q't}' + \xi_{it}' \end{aligned} \tag{1.1$'$}$$

for all $i \in \mathbb{N}$ and $t \in \mathbb{Z}$, and Assumptions A1 through A5 are fulfilled for (1.1′), then

$$q' = q, \qquad \chi_{it}' = \chi_{it}, \qquad \xi_{it}' = \xi_{it}$$

for all $i \in \mathbb{N}$ and $t \in \mathbb{Z}$ (see Forni and Lippi, 2001).

Note that the asymptotic condition in Assumption A4 does not require mutual orthogonality of the idiosyncratic components, a standard identification condition in finite-n factor models. For example, a non-zero correlation of ξ_{it} with $\xi_{i+1,t}$ does not conflict with A4. As a consequence, the decomposition of the x's into common and idiosyncratic components is identified only under $x_{it} = \chi_{it} + \xi_{it} = \chi_{it}' + \xi_{it}'$, *for all* $i \in \mathbb{N}$ and $t \in \mathbb{Z}$. Note also that uniqueness does not extend to $B_n(L)$ or the common shocks $\boldsymbol{u}_t$. For, if $\mathcal{B}(L)$ is a $q \times q$ filter such that $\mathcal{B}(z)\mathcal{B}'(z^{-1}) = I_q$ for $|z| = 1$, then defining

$$\tilde{B}_n(L) = B_n(L)\mathcal{B}(L), \qquad \tilde{\boldsymbol{u}}_t = \mathcal{B}'(L^{-1})\boldsymbol{u}_t, \tag{2.5}$$

we have $\boldsymbol{\chi}_{nt} = \tilde{B}_n(L)\tilde{\boldsymbol{u}}_t$, which can replace the second equation in (2.4).

Now consider $\Sigma_n^x(\theta)$, its first q eigenvalues and corresponding eigenvectors:

$$\lambda_{n1}^x(\theta)\ \lambda_{n2}^x(\theta) \cdots \lambda_{nq}^x(\theta),$$
$$p_{n1}^x(\theta)\ p_{n2}^x(\theta) \cdots p_{nq}^x(\theta),$$

where $|p_{n1}^x(\theta)|^2 + |p_{n2}^x(\theta)|^2 + \cdots + |p_{nq}^x(\theta)|^2 = 1$ for all $\theta \in [-\pi\ \pi]$. Define $P_{nj}(L)$ as the inverse Fourier transform of

$$\frac{1}{\sqrt{\lambda_{nj}^x(\theta)}} p_{nj}^x(\theta).$$

The vector

$$\boldsymbol{u}_{t,n} = (P_{n1}(L)' P_{n2}(L)' \cdots P_{nq}(L)')' \boldsymbol{x}_{nt}$$

is a q-dimensional orthonormal white noise. Moreover, define

$$\chi_{it,n} = \text{Proj}(x_{it} | \overline{\text{span}}(\boldsymbol{u}_{s,n},\ s \in \mathbb{Z})).$$

Then as $n \to \infty$ we have $\chi_{it,n} \to \chi_{it}$ in quadratic mean.

The above matrices and vectors have sample counterparts:

$$\Sigma_{nT}^x(\theta), \qquad P_{nT,j}(L), \qquad \boldsymbol{u}_{t,nT}, \qquad \chi_{it,nT},$$

and the result is that

$$\chi_{it,nT} \to \chi_{it}$$

in probability as $n, T \to \infty$ (see Forni et al., 2000).

The following elementary example shows how the dynamic principal components work and their main drawback:

$$\chi_{it} = \begin{cases} u_{t-1} & \text{if } i \text{ is odd} \\ u_t & \text{if } i \text{ is even.} \end{cases} \tag{2.6}$$

Moreover, assume that $\Sigma_n^{\xi}(\theta) = \frac{1}{2\pi} I_n$ (the idiosyncratic components are orthogonal to one another and have unit variance). Then

$$\Sigma_n^x(\theta) = \frac{1}{2\pi} \begin{pmatrix} e^{-i\theta} \\ 1 \\ \vdots \\ e^{-i\theta} \\ 1 \end{pmatrix} \left(e^{i\theta}\ 1 \cdots e^{i\theta}\ 1\right) + \frac{1}{2\pi} I_n.$$

The first eigenvalue is $1 + n$, with eigenvector $\frac{1}{\sqrt{n}}\left(e^{i\theta}\ 1\ \cdots\ e^{i\theta}\ 1\right)$, so

$$P_{n1}(L) = \frac{1}{\sqrt{n(1+n)}} \left(L^{-1}\ 1 \cdots L^{-1}\ 1\right),$$

where L^{-1} is the forward shift operator: $L^{-1}x_{it} = x_{i,t+1}$. As a consequence this estimator can be used only for $t \leq T - 1$.

2.2. The restricted model

An important simplification is obtained with the following assumption, which is used in Stock and Watson (2002a,b), Bai and Ng (2002), Forni et al. (2005) and Forni et al. (2009). For a given t, we will denote by S_t^{χ} the Hilbert space $\overline{\text{span}}(\chi_{it},\ i \in \mathbb{N})$, i.e. the closure of the set of all linear combinations of the variables χ_{it}. Note that stationarity of the vectors $\boldsymbol{\chi}_{nt}$ implies that the dimension of S_t^{χ} is independent of t.

AF. The space S_t^{χ} is finite dimensional.

Under A1 through A5 plus AF, denote by r the dimension of S_t^{χ}. There exist:

(I) an r-dimensional stationary process $\boldsymbol{F}_t$, which has the representation

$$\boldsymbol{F}_t = N(L)\boldsymbol{u}_t, \tag{2.7}$$

$N(L)$ being a square-summable $r \times q$ filter;

(II) nested $n \times r$ matrices C_n, such that

$$\boldsymbol{\chi}_{nt} = B_n(L)\boldsymbol{u}_t = C_n\boldsymbol{F}_t \tag{2.8}$$

(for a proof of this fairly trivial statement, see Forni et al. (2009). The processes F_{jt} are called the *static factors*. Note that the static factors evolve according to a dynamic equation; see (2.7). "Static" only refers to the loading of $\boldsymbol{F}_t$ by the χ's; see (2.8).

Summing up, in general, the stochastic variables $\{\chi_{it},\ i \in \mathbb{N},\ t \in \mathbb{Z}\}$, span an infinite-dimensional Hilbert space $\mathcal{X}$, which is contained in the Hilbert space spanned by $\{u_{jt},\ j = 1, \ldots, q,\ t \in \mathbb{Z}\}$. Under AF the Hilbert space spanned by $\{\chi_{it},\ i \in \mathbb{Z}\}$, for any given t, is finite dimensional with stationary basis $\boldsymbol{F}_t$. Of course in that case $\mathcal{X}$ is also contained in the Hilbert space spanned by $\{F_{jt},\ j = 1, \ldots, r,\ t \in \mathbb{Z}\}$.

Let Γ_n^x be the covariance matrix of $\boldsymbol{x}_{nt}$. Under AF, estimation of the common components can be achieved using the first r eigenvalues and corresponding eigenvectors of Γ_n^x to obtain $\boldsymbol{F}_{t,n}$, then projecting x_{it} on $\boldsymbol{F}_{t,n}$. In this case only contemporaneous values of the x's are involved, so no two-sidedness problem arises.

3. Back to the general model

As we have observed in the Introduction, taking the simple case (1.2), rewritten here:

$$x_{it} = \frac{1}{1 - \alpha_i L} u_t + \xi_{it},$$

where α_i is drawn from the uniform distribution on the interval $[-.9\ .9]$, we see that S_t^χ is not finite dimensional; thus, so to speak, we have an infinite number of static factors.

Criteria for determining r, the number of static factors, when applied to models like (1.2), will produce wrong results, with the estimated r growing to infinity with n. Moreover, all criteria for determining q that are based on firstly estimating $\boldsymbol{F}_t$, then estimating a VAR for $\boldsymbol{F}_t$, are misspecified. To our knowledge, the only criterion for determining q, which does not depend on the assumption of a finite r and has therefore general applicability, is that of Hallin and Liška (2007).

3.1. Fundamental and zeroless representations

We believe that model (1.2) provides a strong motivation for not assuming AF. Instead, we assume here that:

A6. The spectral density of $\boldsymbol{\chi}_{nt}$ is rational.

Assumptions A6 and A5 imply that there exists $\bar{n} \geq q$ such that for $n \geq \bar{n}$, rank $(\Sigma_n^\chi(\theta)) = q$ for θ a.e. in $[-\pi\ \pi]$. As a consequence, for $n \geq \bar{n}$ the vector $\boldsymbol{\chi}_{nt}$ has a *fundamental* rational representation of rank q, i.e.

$$\boldsymbol{\chi}_{nt} = C_n(L)\boldsymbol{v}_t^{(n)}, \tag{3.9}$$

where: (1) the entries of $C_n(L)$, denoted by $c_{ij}(L)$, are rational functions

$$c_{ij}(L) = \frac{d_{ij}(L)}{e_{ij}(L)},$$

where d_{ij} and e_{ij} have no common roots and $e_{ij}(0) = 1$; (2) $\boldsymbol{v}_t^{(n)}$ is a q-dimensional orthonormal white noise; (3) $C_n(z)$ has no zeros for $|z| < 1$, a zero of $C_n(z)$ being defined as a complex number ζ such that the rank of $C_n(\zeta)$ is lower than the maximum rank of $C_n(z)$, and no poles for $|z| \leq 1$, the poles of $C_n(z)$ being defined as the poles of the polynomials $e_{ij}(z)$. This implies that $\boldsymbol{v}_t^{(n)}$ belongs to the space spanned by $\boldsymbol{\chi}_{n,t-k}$, for $k \geq 0$. As (3.9) implies that $\boldsymbol{\chi}_{nt}$ belongs to the space spanned by $\boldsymbol{v}_{t-k}^{(n)}$, for $k \geq 0$, the two spaces coincide.

Fundamental representations are unique up to an orthogonal matrix. To be precise,

$$\boldsymbol{\chi}_{nt} = \tilde{C}_n(L)\tilde{\boldsymbol{v}}_t^{(n)}$$

is fundamental if and only if there exists an orthogonal matrix K_n such that

$$\tilde{C}_n(L) = C_n(L)K_n$$
$$\tilde{\boldsymbol{v}}_t^{(n)} = K_n'\boldsymbol{v}_t^{(n)}.$$

To understand the relationship between representations (3.9) and (2.4), consider again example (2.6):

$$\chi_{it} = \begin{cases} u_{t-1} & \text{for } i \text{ odd} \\ u_t & \text{for } i \text{ even.} \end{cases}$$

In this case a fundamental white noise for $\boldsymbol{\chi}_{nt}$ is u_{t-1} for $n = 1$, u_t for $n > 1$. Note also that the $(1, 1)$ entry of $C_n(L)$ is 1 for $n = 1$, L for $n > 1$. The example shows that, firstly, reference to n in $\boldsymbol{v}_t^{(n)}$ is necessary and, secondly, that the matrices $C_n(L)$, unlike the matrices $B_n(L)$, are not necessarily nested.

In the following example, though $C_n(L) \neq B_n(L)$ for all n, the matrices $C_n(L)$ are nested. Let $q = 1$ and let representation (1.1) be

$$\chi_{it} = b(L)u_t, \quad b(L) = \frac{1 - \alpha^{-1}L}{1 - \alpha L},$$

with $|\alpha| < 1$. As the polynomial $1 - \alpha^{-1}L$ is not invertible, the white noise u_t does not belong to the space spanned by present and past values of the χ's. However, elementary calculations show that

$$\frac{1 - \alpha^{-1}L}{1 - \alpha L} u_t = -\alpha^{-1}\left[\frac{1 - \alpha L^{-1}}{1 - \alpha L}(Lu_t)\right] = -\alpha^{-1}w_t,$$

and that the spectral density of w_t is equal to unity at all frequencies. Thus w_t is a unit-variance white noise. Representation (3.9) is immediately obtained:

$$\chi_{it} = cw_t, \quad c = -\alpha^{-1}.$$

Thus the matrices $C_n(L)$ are nested and $\boldsymbol{v}_t^{(n)} = w_t$ is independent of n.

More generally, under Assumption A7′, to be introduced below, we can choose the fundamental representations (3.9) in such a way that $\boldsymbol{v}_t^{(n)}$ is independent of n and the matrices $C_n(L)$ are nested.

Now consider the set of all $n \times q$ matrices $D(L)$, with rational entries

$$d_{ij}(K) = \frac{f_{ij}(L)}{g_{ij}(L)},$$

with $g_{ij}(0) = 1$, such that

$$\text{degree}(f_{ij}) \leq p_1, \qquad \text{degree}(g_{ij}) \leq p_2.$$

The parameter space for $D(L)$ has dimension $nq(p_1 + p_2 + 1)$. If the matrix $D(L)$ is *tall*, i.e. if $n > q$, then, *for generic values of the parameters*, $D(L)$ is *zeroless*, i.e. the rank of $D(z)$ is q for all complex numbers z.

To see why this result holds, consider firstly the following example, in which $q = 1$:

$$\chi_{it} = (\alpha_i + \beta_i L)u_t, \tag{3.10}$$

for $i = 1, \ldots, n$, with $n > 1$. Obviously in this case $D(z)$ is zeroless unless $\alpha_i/\beta_i = \gamma$ for all i. In general, existence of a zero of $D(z)$ means that the determinants of all the $q \times q$ submatrices of $D(z)$ vanish for the same complex number. This implies algebraic restrictions on the coefficients of $D(L)$, as argued in Forni et al. (2009) and Zinner (2008). For a formal proof see Anderson and Deistler (2008a) and Deistler et al. (2010).

This motivates the following assumption, which will be enhanced in the next section:

A7. For $n \geq q + 1$, the matrix $C_n(z)$, corresponding to the fundamental representation $\boldsymbol{\chi}_{nt} = C_n(L)\boldsymbol{v}_t^{(n)}$, is zeroless.

3.2. Autoregressive representations for $n > q$

Tall, zeroless moving average rational matrices possess a finite inverse:
(F) Let $n > q$. Consider the rational representation $\boldsymbol{y}_t = D(L)\boldsymbol{z}_t$, where $\boldsymbol{y}_t$ is n-dimensional and $\boldsymbol{z}_t$ is an orthonormal q-dimensional white noise. If $D(L)$ is zeroless then $\boldsymbol{y}_t$ has a finite autoregressive representation

$$A(L)\boldsymbol{y}_t = D(0)\boldsymbol{z}_t.$$

For a formal proof see Anderson and Deistler (2008b) and Deistler et al. (2010). Example (3.10) for $n = 2$ provides an intuition:

$$\begin{aligned}\chi_{1t} &= \alpha_1 u_t + \beta_1 u_{t-1}\\ \chi_{2t} &= \alpha_2 u_t + \beta_2 u_{t-1}.\end{aligned} \tag{3.11}$$

We see that

$$u_t = \frac{1}{\alpha_1\beta_2 - \alpha_2\beta_1}(\beta_2\chi_{1t} - \beta_1\chi_{2t}),$$

and so

$$\begin{pmatrix} 1-\delta\beta_1\beta_2 L & \delta\beta_1^2 L \\ -\delta\beta_2^2 & 1+\delta\beta_1\beta_2 L\end{pmatrix}\begin{pmatrix}\chi_{1t}\\ \chi_{2t}\end{pmatrix} = \begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} u_t,$$

where $\delta = 1/(\alpha_1\beta_2 - \alpha_2\beta_1)$. Note that the autoregressive representation exists if and only if $\alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$, that is when $D(z)$ is zeroless. Moreover, χ_{1t-1} and χ_{2t-1} are linearly independent. Therefore the autoregressive representation of order 1 is unique.

But as soon as $n = 3$,

$$\begin{aligned}\chi_{1t} &= \alpha_1 u_t + \beta_1 u_{t-1}\\ \chi_{2t} &= \alpha_2 u_t + \beta_2 u_{t-1}\\ \chi_{3t} &= \alpha_3 u_t + \beta_3 u_{t-1},\end{aligned} \tag{3.12}$$

we see that infinitely many autoregressive representations of order 1 are possible. For, setting $\boldsymbol{\alpha} = (\alpha_1\ \alpha_2\ \alpha_3)'$, we have

$$u_t = \frac{1}{\boldsymbol{c}\boldsymbol{\alpha}}(c_1\chi_{1t} + c_2\chi_{2t} + c_3\chi_{3t}), \tag{3.13}$$

where $\boldsymbol{c} = (c_1\ c_2\ c_3)$ is any vector orthogonal to $(\beta_1\ \beta_2\ \beta_3)$ and such that $\boldsymbol{c}\boldsymbol{\alpha} \neq 0$. Using (3.13) to replace u_{t-1} in (3.12), we obtain an autoregressive representation of order one depending on $\boldsymbol{c}$.

Consider now $q+1$ integers $i_1, i_2, \dots, i_{q+1}$, with $1 \le i_k < i_{k+1} \le n$, and let

$$\boldsymbol{\chi}_{i_1,\dots,i_{q+1},t} = (\chi_{i_1 t}\ \chi_{i_2 t}\cdots\chi_{i_{q+1} t})' = C_{n;i_1,\dots,i_{q+1}}(L)\boldsymbol{v}_t^{(n)} \tag{3.14}$$

be obtained from (3.9) by selecting the rows $i_1, i_2, \dots, i_{q+1}$. The vector (3.14) is tall (it has dimension $q+1$ and rank q), so for generic values of the parameters the matrix $C_{n;i_1,\dots,i_{q+1}}(L)$ is zeroless. As a consequence, by Proposition (F), for generic values of the parameters the vector $\boldsymbol{\chi}_{i_1,\dots,i_{q+1},t}$ has a finite autoregressive representation. This motivates almost all of Assumption A7′ below, which enhances Assumption A7. The uniqueness in part (ii) is motivated by the discussion of examples (3.11) and (3.12).

A7′. For all n and all choices of $i_1, i_2, \dots, i_{q+1}$, we assume that (i) $C_{n;i_1,\dots,i_{q+1}}(z)$ is zeroless, and that (ii) $\boldsymbol{\chi}_{i_1,\dots,i_{q+1},t}$ has a unique minimum-lag autoregressive representation.

As the vector (3.14) is tall, being of dimension $(q+1)$ but of rank q, part (i) of A7′ can be motivated by the genericity argument. Part (ii) has a motivation in the discussion of examples (3.11) and (3.12).

A consequence of A7′(i) is that the space spanned by present and past values of $\{\chi_{it},\ i \in \mathbb{Z}\}$ is equal to that spanned by present and past values of any $q+1$ among the variables χ_{it}. For, present and past values of $\boldsymbol{\chi}_{i_1,\dots,i_{q+1},t}$ span the same space as is spanned by present and past values of $\boldsymbol{v}_t^{(n)}$, and therefore by present and past values of $\boldsymbol{\chi}_{nt}$, for any n.

Assumption A7′ rules out examples like (2.6), which fulfills A7. Note however that (2.6) is a special case of (3.10), in which Assumption A7′ is fulfilled for generic values of α_i and β_i.

Lastly, consider a fundamental representation for $\boldsymbol{\chi}_{q+1,t}$:

$$\boldsymbol{\chi}_{q+1,t} = F(L)\boldsymbol{v}_t.$$

By A7′, for $i > q+1$, χ_{it} belongs to the space spanned by present and past values of $\boldsymbol{\chi}_{q+1,t}$ and therefore of $\boldsymbol{v}_t$, so

$$\chi_{it} = f_{i1}(L)v_{1t} + f_{i2}(L)v_{2t} + \cdots + f_{iq}(L)v_{qt},$$

for all $i \in \mathbb{N}$. Thus under A7′ representation (3.9) can be written with a white noise $\boldsymbol{v}_t$, which is independent of n, and nested matrices $C_n(L)$:

$$\boldsymbol{\chi}_{nt} = C_n(L)\boldsymbol{v}_t. \tag{3.15}$$

3.3. Non-stationary variables; cointegration

Application of our dynamic factor model requires stationarity. If the data set contains non-stationary variables, as is the case with macroeconomic data sets, the data must be transformed either by removing a deterministic trend or by differencing (this is current practice in dynamic factor literature). Name as y_{it} the variables in the data set and x_{it} the corresponding transformed stationary variables. The question that we want to briefly discuss here is whether some of our assumptions may fail to hold for the transformed variables x_{it}. We find that strong cointegration relationships among the common components of the y's imply that Assumption A7′ does not hold for some choice of $i_1, i_2, \dots, i_{q+1}$.

Assume for simplicity that all the variables y_{it} in the data set are I(1) and that

$$y_{it} = \phi_{it} + \psi_{it},$$

where:

(i) ϕ_{it} is I(1) for all $i \in \mathbb{N}$.
(ii) The variables x_{it}, χ_{it} and ξ_{it}, defined as the first differences of y_{it}, ϕ_{it} and ψ_{it} respectively, evolve according to model (2.4) and fulfill Assumptions A1 through A6.

Consider now a q-dimensional vector obtained by selecting q variables among the ϕ's:

$$\boldsymbol{\phi}_{i_1,\dots,i_q,t} = (\phi_{i_1 t}\ \phi_{i_2 t}\cdots\phi_{i_q t})'.$$

The vector $\boldsymbol{\phi}_{i_1,\dots,i_q,t}$ has the representation

$$\boldsymbol{\chi}_{i_1,\dots,i_q,t} = B_{i_1,\dots,i_q}(L)\boldsymbol{u}_t,$$

which is obtained by selecting the rows $i_1, i_2, \dots, i_q$ in (2.4). The vector $\boldsymbol{\phi}_{i_1,\dots,i_q,t}$ is cointegrated if and only if $B_{i_1,\dots,i_q,t}(1)$ is singular.

Now consider a $(q+1)$-dimensional vector

$$\boldsymbol{\phi}_{i_1,\dots,i_{q+1},t} = (\phi_{i_1 t}\ \phi_{i_2 t}\cdots\phi_{i_{q+1} t})',$$

whose representation is

$$\boldsymbol{\chi}_{i_1,\dots,i_{q+1},t} = B_{i_1,\dots,i_{q+1}}(L)\boldsymbol{u}_t,$$

where the matrix $B_{i_1,\dots,i_{q+1}}(L)$ is $(q+1)\times q$. If we assume that all q-dimensional subvectors of $\boldsymbol{\phi}_{i_1,\dots,i_{q+1},t}$ are cointegrated, the matrix $B_{i_1,\dots,i_{q+1}}(z)$ has a zero at $z = 1$. Thus A7′ does not hold for $\boldsymbol{\chi}_{i_1,\dots,i_{q+1},t}$. In particular, $\boldsymbol{\chi}_{i_1,\dots,i_{q+1},t}$ has no finite autoregressive representation.

The problem has no obvious solution as the variables ϕ_{it} and χ_{it} are not observable. Direct estimation of non-stationary factors and common components has been obtained in Bai and Ng (2004), but only for the restricted model. Methods allowing estimation of the components ϕ_{it} and testing for their cointegration in the general model are not available. On the other hand, we do not really need as much as Assumption A7′. In the next section we show that what is needed to obtain a finite autoregressive representation for $\boldsymbol{\chi}_{nt}$ is the existence of a partition of $\boldsymbol{\chi}_{nt}$ into $(q+1)$-dimensional subvectors each fulfilling A7′. In empirical situations, careful grouping of the variables, based for example on their economic relationships, should help with avoiding "dangerous" $(q+1)$-dimensional vectors.

4. Transforming the dynamic model into a static model with *q* factors

We assume for convenience that $n = (q+1)m$ and partition $\boldsymbol{\chi}_{nt}$ as

$$\boldsymbol{\chi}_{nt} = (\boldsymbol{\chi}'_{[1]t}\ \boldsymbol{\chi}'_{[2]t} \cdots \boldsymbol{\chi}'_{[m]t})',$$

where $\boldsymbol{\chi}'_{[s]t} = (\chi_{(s-1)(q+1)+1,t}\ \chi_{(s-1)(q+1)+2,t} \cdots \chi_{s(q+1),t})'$.

We start with (3.15) and denote by

$$A_{[s]}(L)\boldsymbol{\chi}_{[s]t} = R_{[s]}\mathbf{v}_t \tag{4.16}$$

the minimum-lag autoregressive representation of the $(q+1)$-dimensional vector $\boldsymbol{\chi}_{[s]t}$ (see Assumption A7′). Combining Eq. (4.16), $\boldsymbol{\chi}_{nt}$ has the following autoregressive representation:

$$\begin{pmatrix} A_{[1]}(L) & 0 & \cdots & 0 \\ 0 & A_{[2]}(L) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & A_{[m]}(L) \end{pmatrix} \boldsymbol{\chi}_{nt} = R_n \mathbf{v}_t, \tag{4.17}$$

where $R_n = (R'_{[1]} \cdots R'_{[m]})'$. Of course other representations like (4.17) can be obtained by reordering the components of $\boldsymbol{\chi}_{nt}$. However, the component of $R_n\mathbf{v}_t$ which corresponds to a given component of $\boldsymbol{\chi}_{nt}$ is independent of which ordering has been chosen.

A8. We assume that the qth eigenvalue of $R_nR'_n$, call it ν_n, tends to infinity as $n \to \infty$.

Assumption A8 is not a consequence of A5. In example (3.10), A5 requires that $\sum |\alpha_i + \beta_i e^{-i\theta}|^2$ diverges for all θ, while A8 requires that $\sum \alpha_i^2$ diverges. Note that A8 is not affected if R_n is multiplied on the right by an orthogonal matrix.

We denote by G^* the complex conjugate of the matrix G.

A9. Let $A_{i_1,\ldots,i_{q+1}}(L)$ be the minimum-lag autoregressive matrix of $\boldsymbol{\chi}_{i_1,\ldots,i_{q+1},t}$. Denote by $\mu_{i_1,\ldots,i_{q+1}}(\theta)$ the maximum eigenvalue of

$$A_{i_1,\ldots,i_{q+1}}(e^{-i\theta})A_{i_1,\ldots,i_{q+1}}(e^{-i\theta})^*.$$

We assume that $\mu_{i_1,\ldots,i_{q+1}}(\theta) \le \mu$ for a positive real μ, for all choices of i_k, $k = 1, \ldots, q+1$, for all θ.

Assumption A9 is reasonable but not trivial. Take

$$A(L) = \begin{pmatrix} 1 & \alpha L \\ \beta L & 1 \end{pmatrix}.$$

The trace of $A(e^{-i\theta})A(e^{-i\theta})^*$ is $|1+\alpha e^{-i\theta}|^2 + |1+\beta e^{-i\theta}|^2$, which is not bounded under the stability condition $|\alpha\beta| < 1$.

Defining $H_n(L)$ as the autoregressive matrix in (4.17), we have

$$H_n(L)\mathbf{x}_{nt} = R_n\mathbf{v}_t + H_n(L)\boldsymbol{\xi}_{nt} \tag{4.18}$$

or, setting $\tilde{\mathbf{x}}_{nt} = H_n(L)\mathbf{x}_{nt}$, $\tilde{\boldsymbol{\chi}}_{nt} = R_n\mathbf{v}_t$ and $\tilde{\boldsymbol{\xi}}_{nt} = H_n(L)\boldsymbol{\xi}_{nt}$,

$$\tilde{\mathbf{x}}_{nt} = R_n\mathbf{v}_t + \tilde{\boldsymbol{\xi}}_{nt} = \tilde{\boldsymbol{\chi}}_{nt} + \tilde{\boldsymbol{\xi}}_{nt}. \tag{4.19}$$

Let us prove that this is a static factor model with q factors, i.e. that as $n \to \infty$ the first q eigenvalues of the covariance matrix of $\tilde{\boldsymbol{\chi}}_{nt}$ diverge and the first eigenvalue of the covariance matrix of $\tilde{\boldsymbol{\xi}}_{nt}$ is bounded. The first statement is a consequence of A8. Moreover, using A4 and A9, we have

$$\begin{aligned} a\Sigma_n^{\tilde{\xi}}(\theta)a^* &= aH_n(e^{-i\theta})\Sigma_n^{\xi}(\theta)H_n(e^{-i\theta})^*a^* \\ &\le \lambda_{n1}^{\xi}(\theta)aH_n(e^{-i\theta})H_n(e^{-i\theta})^*a^* \le \lambda\mu|a|^2. \end{aligned}$$

Thus the first eigenvalue of the spectral density $\Sigma_n^{\tilde{\xi}}(\theta)$, call it $\lambda_{n1}^{\tilde{\xi}}(\theta)$, is bounded by $\lambda\mu$. On the other hand, the first eigenvalue of the covariance matrix of $\tilde{\boldsymbol{\xi}}_{nt}$ is bounded by

$$\int_{-\pi}^{\pi} \lambda_{n1}^{\tilde{\xi}}(\theta)d\theta.$$

The result follows.

Other choices of the autoregressive representation of $\boldsymbol{\chi}_{nt}$ may turn out into representations $\check{\mathbf{x}}_{nt} = \check{\boldsymbol{\chi}}_{nt} + \check{\boldsymbol{\xi}}_{nt}$ with a non-idiosyncratic $\check{\boldsymbol{\xi}}_{nt}$. As an example, consider again model (3.10):

$$\boldsymbol{\chi}_{nt} = \boldsymbol{\alpha}_n u_t + \boldsymbol{\beta}_n u_{t-1}.$$

If $\mathbf{c} = (c_1\ c_2\ \cdots\ c_n)$ is orthogonal to $\boldsymbol{\beta}_n$, then an autoregressive representation is

$$\left[I - (\delta\boldsymbol{\beta}_n\mathbf{c}L)\right]\boldsymbol{\chi}_{nt} = \boldsymbol{\alpha}_n u_t,$$

where $\delta = (\mathbf{c}\boldsymbol{\alpha}_n)^{-1}$ and therefore

$$\left[I - (\delta\boldsymbol{\beta}_n\mathbf{c}L)\right]\mathbf{x}_{nt} = \boldsymbol{\alpha}_n u_t + \left[I - (\delta\boldsymbol{\beta}_n\mathbf{c}L)\right]\boldsymbol{\xi}_{nt} = \check{\boldsymbol{\chi}}_{nt} + \check{\boldsymbol{\xi}}_{nt}.$$

We have

$$\check{\xi}_{it} = \xi_{it} + \delta\beta_i[\mathbf{c}\boldsymbol{\xi}_{n,t-1}].$$

Thus the vector $\check{\boldsymbol{\xi}}_{nt}$ is not idiosyncratic.

5. Estimation; a sketch

In the previous section we have shown that Assumption A7′ implies the existence of representation (4.18). We now provide a procedure for constructing $H_n(L)$, R_n and $\mathbf{v}_t$ starting with the spectral density of the common components $\Sigma_n^{\chi}(\theta)$. As we assume that $\Sigma_n^{\chi}(\theta)$ is known, this is to be considered only as a sketch of an estimation procedure. In practical situations $\Sigma_n^{\chi}(\theta)$ is not known; we start with an estimate $\hat{\Sigma}_n^{\chi}(\theta)$ and compute the corresponding sample-dependent $\hat{H}_n(L)$, $\hat{R}_n$ and $\hat{\mathbf{v}}_t$. A proof of consistency of such estimates, for n and T tending to infinity, is beyond the scope of the present paper and left for future research. Let us only observe here that our assumptions, A1 through A9, must be enhanced with conditions ensuring consistency of a smoothed periodogram of $\mathbf{x}_{nt}$ (see e.g. Brockwell and Davis, 1991, pp. 445–7).

Firstly we determine $H_n(L)$ and R_n. We keep assuming that $n = (q+1)m$. Using the m diagonal $(q+1)\times(q+1)$ blocks of $\Sigma_n^{\chi}(\theta)$ we can obtain the matrices

$$G_{[j]}(L), \qquad \Gamma_{[j]}, \quad j = 1, 2, \ldots, m,$$

corresponding to the Wold representation

$$\boldsymbol{\chi}_{[j]t} = G_{[j]}(L)\mathbf{w}_{[j]t}. \tag{5.20}$$

Note that neither the $\boldsymbol{\chi}_{[j]t}$ nor the $\mathbf{w}_{[j]t}$ are observable. The matrix $G_{[j]}(L)$ is $(q+1)\times(q+1)$ and has rational entries. Moreover, $G_{[j]}(0) = I_{q+1}$. The matrix $\Gamma_{[j]}$ is the covariance matrix of the $(q+1)\times 1$ one-step-ahead prediction error vector $\mathbf{w}_{[j]t}$. The matrix

$\Gamma_{[j]}$ (like $\boldsymbol{w}_{[j]t}$) is of rank q. By Assumption A7′(ii), (5.20) can be rewritten as

$$A_{[j]}(L)\boldsymbol{\chi}_{[j]t} = \boldsymbol{w}_{[j]t},$$

where $A_{[j]}(L)$ is the unique minimum-lag left inverse of $G_{[j]}(L)$. The matrix $\Gamma_{[j]}$ can be factored as

$$\Gamma_{[j]} = \left[P_{[j]}\Lambda_{[j]}^{\frac{1}{2}}\right]\left[\Lambda_{[j]}^{\frac{1}{2}}P'_{[j]}\right],$$

the matrix $P_{[j]}$ being $(q+1)\times q$ with the normalized first q eigenvectors of $\Gamma_{[j]}$ on the columns, while $\Lambda_{[j]}$ is $q\times q$ with the (non-zero) corresponding eigenvectors on the diagonal. The columns of $P_{[j]}$ are mutually orthogonal. We define

$$\boldsymbol{v}_{[j]t} = \Lambda_{[j]}^{-\frac{1}{2}}P'_{[j]}\boldsymbol{w}_{j[]t} = \Lambda_{[j]}^{-\frac{1}{2}}P'_{[j]}A_{[j]}(L)\boldsymbol{\chi}_{[j]t}. \quad (5.21)$$

It is easily seen that $\boldsymbol{v}_{[j]t}$ is an orthonormal q-dimensional white noise. Moreover, projecting $\boldsymbol{w}_{[j]t}$ on $\boldsymbol{v}_{[j]t}$ we find $\boldsymbol{w}_{[j]t} = P_{[j]}\Lambda_{[j]}^{\frac{1}{2}}$ $\boldsymbol{v}_{[j]t}$. Defining $S_{[j]} = P_{[j]}\Lambda_{[j]}^{\frac{1}{2}}$, we obtain

$$A_{[j]}(L)\boldsymbol{\chi}_{[j]t} = S_{[j]}\boldsymbol{v}_{[j]t}.$$

The white noise vectors $\boldsymbol{v}_{[j]t}$ are different in general but, by Assumption A7′, span the same space. Therefore, for $j = 2, \dots, m$,

$$\boldsymbol{v}_{[j]t} = K_{[j]}\boldsymbol{v}_{[1]t},$$

where K_j is orthogonal. Using (5.21),

$$\begin{aligned} K_{[j]} &= \mathrm{E}\left(\boldsymbol{v}_{[j]t}\boldsymbol{v}'_{[1]t}\right) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\Lambda_j^{-\frac{1}{2}}P'_{[j]}A_{[j]}(\mathrm{e}^{-\mathrm{i}\theta})\Sigma^{\chi}_{[j1]}(\theta)A_{[1]}(\mathrm{e}^{-\mathrm{i}\theta})^{*}P_{[1]}\Lambda_{[1]}^{-\frac{1}{2}}\right]\mathrm{d}\theta, \end{aligned}$$

where $\Sigma^{\chi}_{[j1]}(\theta)$ is the $(q+1)\times(q+1)$ cross-spectrum of $\boldsymbol{\chi}_{[j]t}$ and $\boldsymbol{\chi}_{[1]t}$ (a submatrix of $\Sigma^{\chi}_{n}(\theta)$). In conclusion, setting $\boldsymbol{v}_t = \boldsymbol{v}_{[1]t}$, we have

$$\begin{pmatrix} A_{[1]}(L) & 0 & \cdots & 0 \\ 0 & A_{[2]}(L) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & A_{[m]}(L) \end{pmatrix}\begin{pmatrix}\boldsymbol{\chi}_{[1]t} \\ \boldsymbol{\chi}_{[2]t} \\ \vdots \\ \boldsymbol{\chi}_{[m]t}\end{pmatrix} = \begin{pmatrix} S_{[1]} \\ S_{[2]}K'_{[2]} \\ \vdots \\ S_{[m]}K'_{[m]}\end{pmatrix}\boldsymbol{v}_t, \quad (5.22)$$

and therefore

$$H_n(L)\boldsymbol{x}_{nt} = R_n\boldsymbol{v}_t + H_n(L)\boldsymbol{\xi}_{nt},$$

where $H_n(L)$ and R_n are defined in (5.22).

The next step determines $\boldsymbol{v}_t$. Note that the matrix R_n has mutually orthogonal columns. As a consequence, R'_nR_n has the eigenvalues of $R_nR'_n$ on the main diagonal (this is easily seen) and zero elsewhere. Setting $M_n = (R'_nR_n)^{-1}$,

$$\begin{aligned} M_nR'_nH_n(L)\boldsymbol{x}_{nt} &= M_nR'_nR_n\boldsymbol{v}_t + M_nR'_nH_n(L)\boldsymbol{\xi}_{nt} \\ &= \boldsymbol{v}_t + M_nR'_nH_n(L)\boldsymbol{\xi}_{nt}. \end{aligned}$$

Denoting by R_{ij} the entries of R_n, the sth row of $M_nR'_n$ is

$$\frac{1}{\sum_{k=1}^{n}R_{ks}^2}\left(R_{1s}\,R_{2s}\cdots R_{ns}\right).$$

Thus the sum of its squares is

$$\frac{1}{\sum_{k=1}^{n}R_{ks}^2},$$

i.e. the reciprocal of the sth eigenvalue of $R_nR'_n$. By Assumption A8, this reciprocal tends to zero as $n\to\infty$. Because $H_n(L)\boldsymbol{\xi}_{nt}$ is idiosyncratic, the term $M_nR'_nH_n(L)\boldsymbol{\xi}_{nt}$ tends to zero in mean square as $n\to\infty$ (see e.g. Forni and Lippi, 2001). Thus

$$M_nR'_nH_n(L)\boldsymbol{x}_{nt} \to \boldsymbol{v}_t$$

in mean square as $n\to\infty$. Lastly, $\boldsymbol{\chi}_{nt}$ results from inversion of $H_n(L)$.

6. Conclusions

Forni et al. (2000) estimate $\Sigma^{\chi}_{n}(\theta)$, the spectral density of the common components of model (1.1), by means of q dynamic principal components, and provide a factorization of $\Sigma^{\chi}_{n}(\theta)$. However, the estimator of the common components based on such factorization, though consistent, applies two-sided filters to the observable variables x_{it}.

In the present paper, under the assumption of rationality for $\Sigma^{\chi}_{n}(\theta)$ and other mild requirements, we obtain a factorization of $\Sigma^{\chi}_{n}(\theta)$ which only employs one-sided filters.

An important feature of our method is that the problem of factoring $\Sigma^{\chi}_{n}(\theta)$, which is of dimension n and rank q, is solved by separately factoring many spectral matrices of dimension $q+1$.

Acknowledgements

We would like to thank, for suggestions and criticism, Manfred Deistler, Marc Hallin, Hashem Pesaran and Paolo Zaffaroni.

References

Anderson, B.D.O., Deistler, M., 2008a. Properties of zero-free transfer function matrices. SICE Journal of Control, Measurement and System Integration 1, 1–9.
Anderson, B.D.O., Deistler, M., 2008b. Generalized linear dynamic factor models—a structure theory. In: 2008 IEEE Conference on Decision and Control.
Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. Econometrica 70, 191–221.
Bai, J., Ng, S., 2004. A PANIC attack on unit roots and cointegration. Econometrica 72, 1127–1177.
Brockwell, P.J., Davis, R.A., 1991. Time Series: Theory and Methods. Springer-Verlag, New York.
Deistler, M., Filler, A., Zinner, C., Chen, W., 2010. Generalized linear dynamic factor models: an approach via singular autoregressions. European Journal of Control 16 (3), 211–224.
Forni, M., Giannone, D., Lippi, M., Reichlin, L., 2009. Opening the black box: structural factor models with large cross-sections. Econometric Theory 25, 1319–1347.
Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2000. The generalized dynamic factor model: identification and estimation. The Review of Economics and Statistics 82, 540–554.
Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2005. The generalized factor model: one-sided estimation and forecasting. Journal of the American Statistical Association 100, 830–840.
Forni, M., Lippi, M., 2001. The generalized dynamic factor model: representation theory. Econometric Theory 17, 1113–1141.
Hallin, M., Liška, R., 2007. The generalized dynamic factor model: determining the number of factors. Journal of the American Statistical Association 102, 103–117.
Stock, J.H., Watson, M.W., 2002a. Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics 20, 147–162.
Stock, J.H., Watson, M.W., 2002b. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167–1179.
Zinner, C., 2008. Modeling of high-dimensional time series by generalized dynamic factor models. Ph.D. Dissertation. Technische Universität Wien.