GENERALIZED LINEAR DYNAMIC FACTOR MODELS - A STRUCTURE THEORY

Manfred Deistler[1]

[1]Research Group Econometrics and System Theory
TU Wien
deistler@tuwien.ac.at

joint work with:
B.D.O. Anderson (Research School of Information Sciences and Engineering, ANU, Canberra)
A. Filler (Vienna University of Technology and University of Vienna)
Ch. Zinner (BAWAG P.S.K., Market Risk, Vienna)

Linz, October 2008

Workshop on Inverse and Partial Information Problems

Contents

Outline

Modeling of high dimensional time series

Multivariate time series

$$y_t \in \mathbb{R}^N, t = 1, ..., T$$

Sample size T, cross-sectional dimension N

- "Traditional" approach, e.g. "unstructured" AR modeling: "Curse of dimensionality": Dimension of parameter space N^2p (p: maximal lag). Number of data points NT.

Alternatives:

- Factor models
  - Comovement allows for dimension reduction in the cross-sectional dimension.

Modeling of high dimensional time series

- Time series factor models: Complexity reduction in time and cross-section. Under certain assumptions the dimension of the parameter space is linear in N.

Note, there is no symmetry in the time and the cross-sectional dimension: Stationarity in time, “similarity” or “comovement” of time series; permutation invariant.

- Cointegration
- Panel-time series models
- Structural models; e.g. ARX models which are sparse due to “physical“ a priori knowledge.
- “Graphical“ time series models, where the inverse of the spectral density is sparse.

Applications

History:

Psychometrics: Intelligence factors (Burt 1909, Thurstone 1934)

Great range of applications: Signal processing, marketing econometrics, finance econometrics, ...

Recent applications for generalized factor models:

- Forecasting for macrovariables
- Structure and insigths for macroeconomics
- Cross-country studies
- Finance

Outline

GDFM's - The Model Class

Generalized linear dynamic factor models (GDFM's)

- Generalization of:
  - Linear dynamic factor models with (strictly) idiosyncratic noise (Geweke, in: Aigner and Goldberger (eds), Latent Variables in Socioeconomic Models 1977, Scherrer and Deistler, SIAM, J. Optim and Control 1998)
  - Generalized static linear factor models (Chamberlain, Chamberlain and Rothschild, Econometrica 1983)
- Main features:
  - Dynamics (here in a stationary context)
  - Uncorrelatedness of noise components in cross-section is replaced by weak dependence (risk diversification is possible)
  - Similarity: Information gain by adding additional time series

GDFM's - The Model Class

Main references for GDFM's

- Forni, Hallin, Lippi, Reichlin, RES, 2000
- Forni and Lippi, Econometric Theory, 2001
- Forni, Hallin, Lippi, Reichlin, JASA, 2005
- Stock and Watson, JASA, 2002
- Stock and Watson, Journal of Business and Economic Statistics, 2002

GDFM's - The Model Class

$$y_t^N = \hat{y}_t^N + u_t^N$$

y_t^N ... observations
$\hat{y}_t^N$ ... latent variables, strongly dependent in the cross-sectional dimension
u_t^N ... (wide sense) idiosyncratic noise, weakly dependent

Assumptions:

- $(\hat{y}_t^N)$, (u_t^N) wide sense stationary with absolutely summable covariances
- $\mathbb{E}\hat{y}_t^N u_s^{N\prime} = 0$
- $\mathbb{E}\hat{y}_t^N = \mathbb{E}u_t^N = 0$

GDFM's - The Model Class

Spectral densities:

$$f_y^N(\lambda) = f_{\hat{y}}^N(\lambda) + f_u^N(\lambda)$$

Asymptotic analysis: $T \to \infty$, $N \to \infty$ Sequence of GDFM's; Nested i.e. elements of $\hat{y}_t^N$ and u_t^N do not depend on N

Assumptions

Additional assumptions to separate the latent variables from the noise for $N \to \infty$:

A1 $f_{\hat{y}}^N$ is a rational spectral density with constant rank $q < N$, and of McMillan degree $2n < N$; q and n do not depend on N

A2 Weak dependence of (u_t^N): The largest eigenvalue of f_u^N is uniformly bounded for all frequencies λ and all N

A3 Strong dependence of $(\hat{y}_t^N)$: The first q eigenvalues of $f_{\hat{y}}^N$ diverge to infinity for all frequencies λ, as $N \to \infty$

Identifiability

GDFM's are identifiable only for $N \to \infty$:

- The elements of $f_{\hat{y}}^N$ and $\hat{y}_t^N$ are uniquely determined from y_t^N for $N \to \infty$
- Asymptotic equivalence to dynamic PCA

Dynamic PCA (Brillinger 1981)

We want to approximate $f_y(\lambda)$ by a spectral density $f_{\hat{y}}(\lambda)$ of rank q for all λ s.t. $\mathbb{E}u_t'u_t$ is minimal

$$f_y(\lambda) = O_1(e^{-i\lambda})\Omega_1(\lambda)O_1(e^{-i\lambda})^* + O_2(e^{-i\lambda})\Omega_2(\lambda)O_2(e^{-i\lambda})^*$$

Model:

$$y_t = \underbrace{O_1(z)O_1^*(z)y_t}_{\hat{y}_t} + \underbrace{O_2(z)O_2^*(z)y_t}_{u_t}$$

Note

- Here dimension reduction is in cross-section only; even for rational f_y, $f_{\hat{y}}$ may be non-rational
- $O_1(z)O_1^*(z)$ may be non-causal
- Estimation commences from a nonparametric estimator of f_y.

Characterization of GDFM's

y_t^N follows a generalized dynamic factor model if and only if

- the first q eigenvalues of f_y^N diverge to infinity for all frequencies λ, as $N \to \infty$
- the $q+1-th$ eigenvalue of f_y^N is uniformly bounded for all frequencies λ and all N

Aims of our Analysis

- Structural insight
- Obtain a state space or an ARMA model and in particular an AR model for $(\hat{y}_t)$ from the second moments of (y_t)
  - estimation of the integer valued parameters such as q (dimension of dynamic factors), r (dimension of static factors) and n (state dimension)
  - estimation of the real valued parameters such as (F, G, H)
- forecasting of y_t based on forecasts of $\hat{y}_t$ and eventually of u_t

Major Steps

- Factorization of $f_{\hat{y}}$
- Realization of a “tall“ spectral factor by a state space or an ARMA model
- Emphasise the zeroless case
- Averaging out of (weakly) idiosyncratic noise

Outline

Outline

$$f_{\hat{y}}(\lambda) = \underbrace{w(e^{-i\lambda})}_{N\times q} w(e^{-i\lambda})^* \tag{1}$$

Theorem

(i) Every rational spectral density $f_{\hat{y}}$ of constant rank q can be factorized as in (1) *where*

$$w(z) = \sum_{j=0}^{\infty} w_j z^j;\ w_j \in \mathbb{R}^{N\times q}$$

is rational, analytic in $|z| \leq 1$ and has rank q for all $|z| \leq 1$.

(ii) For given $f_{\hat{y}}$, w is unique up to postmultiplication by constant orthogonal matrices.

Wold decomposition

(Smith Mc Millan form)

$$w = u\ell v$$

u, v *... unimodular, polynomial*
ℓ *... diagonal, diagonal elements display poles and zeros of w*

w corresponds to the Wold decomposition:
There exists (ε_t), white noise, with $\mathbb{E}\varepsilon_t\varepsilon_t' = 2\pi I$, s.t.:

$$\hat{y}_t = w(z)\varepsilon_t = \sum_{j=0}^{\infty} w_j \varepsilon_{t-j}$$

$$\varepsilon_t = w^-(z)\hat{y}_t$$

causal left inverse $w^- = v^{-1}(\ell'\ell)^{-1}\ell' u^{-1}$

Outline

Realization

Realization: Find a system for a transfer function

- State space system - (F, G, H)
- ARMA system

  $$a(z)\hat{y}_t = b(z)\varepsilon_t$$
  $$w(z) = a(z)^{-1}b(z)$$

- Right MFD (Lippi)

  $$w(z) = c(z)d(z)^{-1}$$

ARMA Realizations

$$\underbrace{a(z)}_{N\times N}\,\hat{y}_t = \underbrace{b(z)}_{N\times q}\,\varepsilon_t$$

(a, b) left coprime

Stability: $det\ a(z) \neq 0, |z| \leq 1$

Miniphase condition: $b(z)$ has full rank, $|z| \leq 1$

State space Realizations

$$x_{t+1} = Fx_t + G\varepsilon_{t+1}$$
$$\hat{y}_t = Hx_t$$

$F \in \mathbb{R}^{n\times n}, G \in \mathbb{R}^{n\times q}, H \in \mathbb{R}^{N\times n}$

x_t ... state

$w(z) = H(I - Fz)^{-1}G$

(F, G, H) is minimal (i.e. controllable and observable)

Stability: $|\lambda_{max}(F)| < 1$

Miniphase condition:

$M(z) = \begin{pmatrix} I - Fz & -G \\ H & 0 \end{pmatrix}$ has rank $n + q$, $|z| \leq 1$

(F, G, H) State space systems

The state is unique up to basis changes:

$$\bar{F} = TFT^{-1}, \bar{G} = TG, \bar{H} = HT^{-1}, \det T \neq 0$$

Note that here x_t is a static factor but not necessarely a minimal one. x_t is a minimal static factor if and only if $rk(H) = n$ holds.

Realization: Kalman-Akaike procedure

$$\underbrace{\begin{pmatrix} \hat{y}_t \\ \hat{y}_{t+1|t} \\ \hat{y}_{t+2|t} \\ \vdots \end{pmatrix}}_{\hat{Y}_t} = \underbrace{\begin{pmatrix} HG & HFG & HF^2G & \dots \\ HFG & HF^2G & HF^3G & \dots \\ \vdots & & & \\ \vdots & & & \end{pmatrix}}_{\mathcal{H} \text{ Hankelmatrix of the transfer function}} \underbrace{\begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \vdots \end{pmatrix}}_{E_t^-}$$

$\hat{y}_{t+r|t}$ ... best linear least squares predictor of $\hat{y}_{t+r}$ from the infinite past $\hat{y}_t, \hat{y}_{t-1}, \dots$

$$X_t = \underbrace{\underbrace{S}_{n \times \infty}}_{rk(S)=n} \begin{pmatrix} \hat{y}_t \\ \hat{y}_{t+1|t} \\ \hat{y}_{t+2|t} \\ \vdots \end{pmatrix} = S\,\mathcal{H}\,E_t^-$$

$$= \underbrace{\underbrace{S \begin{pmatrix} HFG & HF^2G & \dots \\ HF^2G & HF^3G & \dots \\ \vdots & & \end{pmatrix}}_{\mathsf{F\,S}\,\mathcal{H}} E_{t-1}^-}_{\mathsf{F}\,X_{t-1}} + \underbrace{S \begin{pmatrix} HG \\ HFG \\ \vdots \end{pmatrix}}_{\mathsf{G}} \varepsilon_t$$

with

$$H\,S\,\mathcal{H} = \begin{pmatrix} HG & HFG & \dots \end{pmatrix}$$

S ... Selector matrix

Special choice for S: Select the first basis rows of $\mathcal{H}$: Echelon form, selection described by Kronecker indices

Outline

Latent Variables and Minimal Static Factors

$$\hat{y}_t = HTT^{-1}x_t = (H_1, 0)T^{-1}x_t = H_1 z_t$$

z_t ... r dimensional minimal static factor, $rk\ H_1 = r$

In general $n \geq r \geq q$;

q is the dimension of minmal dynamic factors

Note: Minimal static factors are unique up to premultiplication by constant nonsingular matrices

Static factors are obtained from $\Sigma_{\hat{y}} = E\hat{y}_t\hat{y}_t'$:

$$\Sigma_{\hat{y}} = MM', M = H_1R, M \in \mathbb{R}^{N\times r}, \operatorname{rk} M = r$$

as:

$$z_t = (M'M)^{-1}M'\hat{y}_t.$$

Static factors z_t and latent variables $\hat{y}_t$ are related by a linear static relation and thus have the same dynamics

$$z_t = (M'M)^{-1}M'w(z)\varepsilon_t = k(z)\varepsilon_t$$

z_t has smaller dimension, thus we prefer to model (z_t)
Echelon case:

$$S = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix}, S_1 \in \mathbb{R}^{r\times\infty}, S_2 \in \mathbb{R}^{(n-r)\times\infty}$$

$$x_t = S\hat{Y}_t, \quad z_t = S_1\hat{Y}_t$$

Clearly (z_t) has a rational spectral density

State space model for (z_t): (F, G, C) where

$$C = (M'M)^{-1}M'H$$

Identification of an ARMA model for (z_t), (Zinner, PhD-thesis TU Wien)

Outline

Zeroless Transfer Functions and (Singular) AR Systems

A transfer function $w(z)$ is called `zeroless` if all numerator polynomials in the diagonal of the matrix ℓ in the Smith-Mc-Millan form are equal to 1

Note: $k(z)$ is zeroless if and only if $w(z)$ is zeroless

Theorem (Anderson and Deistler, SICE J. Control 2008)

Consider a rational transfer function $w(z)$ with minimal state space realization (F, G, H) with state dimension n. If $N > q$ holds, then the transfer functions are zeroless for generic values of (F, G, H).

Theorem (Anderson and Deistler, CDC, 2008)

The following statements are equivalent:

(i) *The stable miniphase spectral factors* k *of the spectral density* f_z *of* (z_t) *are zeroless*

(ii) *There exists a polynomial left inverse* k^- *for* k

(iii) (z_t) *is a stable AR-process, i.e.*

$$z_t = e_1 z_{t-1} + \cdots + e_p z_{t-p} + \nu_t$$

where $\det \underbrace{(I - e_1 z - \cdots - e_p z^p)}_{e(z)} \neq 0,\ |z| \leq 1$

and $\operatorname{rk} \Sigma_\nu = q$, $\Sigma_\nu = E\nu_t \nu_t'$.

Let

$$\Gamma_m = \begin{pmatrix} \gamma_0 & \cdots & \cdots & \gamma'_{m-1} \\ \vdots & \gamma_0 & & \vdots \\ \vdots & & \ddots & \vdots \\ \gamma_{m-1} & \cdots & \cdots & \gamma_0 \end{pmatrix} \text{ where } \gamma_j = E z_{t+j} z'_t$$

If Σ_ν is nonsingular, then Γ_m is nonsingular for all m.

If Σ_ν is singular, then Γ_{p+1} will be singular and Γ_p may be singular.

Yule Walker Equations:

$$\begin{aligned} (e_1, \ldots, e_p)\Gamma_p &= (\gamma_1, \ldots, \gamma_p) \\ \Sigma_\nu &= \gamma_0 - (e_1, \ldots, e_p)(\gamma'_1, \ldots, \gamma'_p)' \end{aligned}$$

Zeroless Transfer Functions and (Singular) AR Systems

Solution of the Yule Walker equations may not be unique: Description of the class of all obervationally equivalent AR systems for given p

Theorem (Anderson and Deistler, CDC, 2008)

(i) *Every singular AR system with* $\mathrm{rk}\,\Sigma_\nu = q$ *can be written as*

$$e(z)z_t = f\varepsilon_t, f \in \mathbb{R}^{r\times q}$$

where (ε_t) *is white noise with* $E\varepsilon_t\varepsilon_t' = I_q$ *and where* $e(z)$ *and* f *are relatively left prime.*

Theorem (Anderson and Deistler, CDC, 2008)

(ii) *Let* $(e(z), f)$ *be relatively left prime, then the class of all observationally equivalent* $(\bar{e}(z), \bar{f})$ *satisfying the degree restrictions* $\delta(\bar{e}(z)) \leq p$, $\delta(\bar{f}) = 0$ *is given by*

$$\left(\bar{e}(z), \bar{f}\right) = u(z)(e(z), f)$$

where the polynomial matrix $u(z)$ *satisfies*

$$\begin{aligned} \det u(z) \neq 0, |z| \leq 1 \\ u(0) = I \\ \delta(u(z)e(z)) \leq p \\ \delta(u(z)f) = 0 \end{aligned}$$

Theorem (Anderson and Deistler, CDC, 2008)

(iii) $e(z)$ *is unique if and only if* $\mathrm{rk}(e_p, f) = r$ *holds.*

Outline

Let $\hat{\gamma}_j^T = 1/T \sum_{t=1}^T \hat{y}_{t+j}\hat{y}_t'$ and let

$$(\hat{e}_1, \ldots, \hat{e}_p)\hat{\Gamma}_p = (\hat{\gamma}_1, \ldots, \hat{\gamma}_p), \hat{\Gamma}_p = \begin{pmatrix} \hat{\gamma}_0 & \cdots & \cdots & \hat{\gamma}_{p-1}' \\ \vdots & \hat{\gamma}_0 & & \vdots \\ \vdots & & \ddots & \vdots \\ \hat{\gamma}_{p-1} & \cdots & \cdots & \hat{\gamma}_0 \end{pmatrix}$$

be the corresponding Yule Walker equations.

Typically $\hat{\Gamma}_p$ will be nonsingular, even if Γ_p is singular, however, truncation is appropriate in such a case

$$(\hat{e}_1, ...\hat{e}_p) = (\hat{\gamma}_1^T, ..., \hat{\gamma}_p^T)O_T\Lambda_T^{-1}O_T'$$

where

$$\hat{\Gamma}_p^T = O_T\Lambda_T O_T',\ O_T \in \mathbb{R}^{pr\times s},\ \Lambda_T \in \mathbb{R}^{s\times s}, s = \operatorname{rk}\Gamma_p$$

The Yule Walker Equations

Theorem (Anderson, Deistler, Filler and Zinner, ECC, 2009)

(i) *If* $\operatorname{rk}\Gamma_p = pr$ *holds, then the YW estimators correspond to a stable autoregression*

(ii) *For* $\operatorname{rk}\Gamma_p = s < pr$*, the truncation procedure above yields a stable autoregression*

Outline

Estimation of Integers

Work in progress:

Estimation of r, q, p and s for the AR case, or r, n, q and the Kronecker indices for the state space case.

Outline

Removing the (weakly) idiosyncratic noise from the observations

Note: $\hat{y}_t$ is not directly observed, how can we get rid of u_t?

$N \to \infty$

There are several approaches.

Here we only describe the (static) PCA based procedure

(i) Commence from $\hat{\Gamma}_y^N = T^{-1} \sum y_t^N y_t^{N\prime}$

Now, under suitable assumptions, for $T, N \to \infty$ the first r eigenvalues of $\hat{\Gamma}_y^N$ tend to infinity, and the other eigenvalues of $\hat{\Gamma}_y^N$ converge to finite values.

Removing the (weakly) idiosyncratic noise from the observations

This is used to estimate r and z_t :

$$\hat{\Gamma}_y^N = O_T \underbrace{\Lambda_T}_{\text{first } r \text{ eigenvalues}} O_T' + remainder$$

$$\hat{z}_t = O_T' y_t$$

$\hat{r}$ and $\hat{z}_t$ are consistent for $T, N \to \infty$

(ii) Use $\hat{z}_t$ to estimate the AR order p and the autoregressive coefficients

Outline

Conclusions

Aims:

- obtain „structural“ insights
- direct estimation procedure based on estimation of the second moments of the observations (computational simplicity)
- treatment of case more general case (compared to existing literature)
- forecasting

Zeroless transfer functions and spectra make things easier.

Further open questions:

- Properties of the estimators for the autoregression and of the estimators for (F, G, H) beyond consistency
- Properties of the Estimation of the integer-valued parameters such as r, q, p, s

Thank You