# Ph.D. Defense

Giulio Bottegal
Advisor: Prof. Giorgio Picci

Department of Information Engineering - University of Padova

February 28th, 2013

# Summary of the thesis

**Modeling, estimation and identification of stochastic systems with latent variables**

Contents

1. Generalized factor analysis models
2. Zero properties of tall multirate linear systems
3. Identifiability of errors-in-variables models
4. Nonparametric kernel-based spectrum estimation

# Summary of the thesis

**Modeling, estimation and identification of stochastic systems with latent variables**

Contents

# Generalized factor analysis models

# Factor analysis models

## History

- Introduced by psychologists (Spearman, 1904)
- Successively applied in econometrics (Ledermann, 1937)
- Extended to a dynamic context (Geweke 1977)
- Generalized to infinite cross-sectional dimension (Chamberlain and Rotschild 1982, Forni and Lippi 2001)

# Factor analysis models

## History

- Introduced by psychologists (Spearman, 1904)
- Successively applied in econometrics (Ledermann, 1937)
- Extended to a dynamic context (Geweke 1977)
- Generalized to infinite cross-sectional dimension (Chamberlain and Rotschild 1982, Forni and Lippi 2001)

## Aim in econometry

Describe the common core of a set of observations

# Factor analysis models

## History

- Introduced by psychologists (Spearman, 1904)
- Successively applied in econometrics (Ledermann, 1937)
- Extended to a dynamic context (Geweke 1977)
- Generalized to infinite cross-sectional dimension (Chamberlain and Rotschild 1982, Forni and Lippi 2001)

## Aim in econometry

Describe the common core of a set of observations

**Could these models be used also for other purposes?**

# Modeling of flocks

## Global perspective

We observe a group of agents with similar behaviors

**Observations = Common behavior + Local interactions**

# Modeling of flocks

## Global perspective

We observe a group of agents with similar behaviors

**Observations = Common behavior + Local interactions**

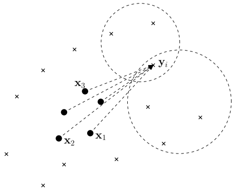
## Examples (from nature)



# Detection of emitters

## Scenario

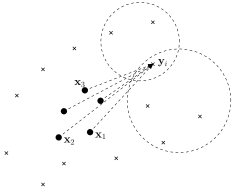
- $q$ impulse emitters $\mathbf{x}_i$
- large ($N \uparrow$) amount of receveirs $\mathbf{y}_i$
- local sources of noise $\tilde{\mathbf{y}}$



# Detection of emitters

## Scenario

- $q$ impulse emitters $\mathbf{x}_i$
- large ($N \uparrow$) amount of receveirs $\mathbf{y}_i$
- local sources of noise $\tilde{\mathbf{y}}$

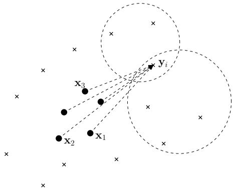



***Goal*: detect $q$ and estimate $\mathbf{x}$**

# Detection of emitters

## Scenario

- $q$ impulse emitters $\mathbf{x}_i$
- large ($N \uparrow$) amount of receveirs $\mathbf{y}_i$
- local sources of noise $\tilde{\mathbf{y}}$




***Goal*: detect $q$ and estimate $\mathbf{x}$**

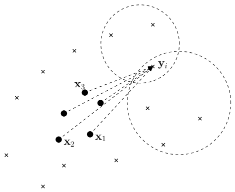
## Receiver equation

$$y_i(t) = f_{i1}x_1(t) + \ldots + f_{iq}x_q(t) + \tilde{y}_i(t) \qquad (f_{ij} \sim \textit{distance from emitter } j)$$

# Detection of emitters

## Scenario

- $q$ impulse emitters $\mathbf{x}_i$
- large ($N \uparrow$) amount of receveirs $\mathbf{y}_i$
- local sources of noise $\tilde{\mathbf{y}}$




***Goal*: detect $q$ and estimate $\mathbf{x}$**

## Receiver equation

$$y_i(t) = f_{i1}x_1(t) + \ldots + f_{iq}x_q(t) + \tilde{y}_i(t) \qquad (f_{ij} \sim \textit{distance from emitter } j)$$

$$\Downarrow$$

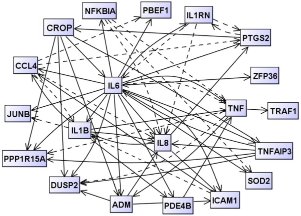
$$\mathbf{y} = F\mathbf{x} + \tilde{\mathbf{y}}$$

$$= \text{Common behavior} + \text{Local interactions}$$

# Gene regulatory network

## Scenario

- network of $N$ genes: $\mathbf{y}_i$ expression level of $i-$th gene
- $q$ transcription factors regulate the activity of the genes ($q << N$)
- Genes mutually influence the activity of their neighbors



# Gene regulatory network

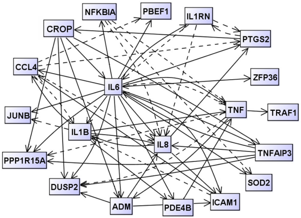
## Scenario

- network of $N$ genes: $\mathbf{y}_i$ expression level of $i-$th gene
- $q$ transcription factors regulate the activity of the genes ($q << N$)
- Genes mutually influence the activity of their neighbors

## Model for the network

$$\mathbf{y}_i = f_{i1}\mathbf{x}_1 + \ldots + f_{iq}\mathbf{x}_q + \tilde{\mathbf{y}}_i$$

*$x_i$ trans. factor, $f_{ij}$ strenght of influence*



# Gene regulatory network

## Scenario

- network of $N$ genes: $\mathbf{y}_i$ expression level of $i-$th gene
- $q$ transcription factors regulate the activity of the genes ($q << N$)
- Genes mutually influence the activity of their neighbors

## Model for the network

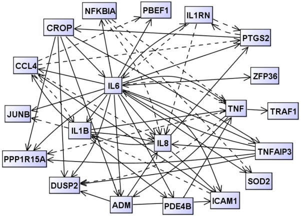
$$\mathbf{y}_i = f_{i1}\mathbf{x}_1 + \ldots + f_{iq}\mathbf{x}_q + \tilde{\mathbf{y}}_i$$

*$x_i$ trans. factor, $f_{ij}$ strenght of influence*

$$\Downarrow$$

$$\mathbf{y} = F\mathbf{x} + \tilde{\mathbf{y}}$$

$=$ Common behavior $+$ Local interactions



# Characterization

## Our scope

Formalize the concept

Observations = Common (and simple) behavior + Local interactions

# Characterization

## Our scope

Formalize the concept

Observations = Common (and simple) behavior + Local interactions

## Stochastic framework of generalized factor analysis

$\mathbf{y} := \{\mathbf{y}_k,\ k \in \mathbb{N}\}$ infinite collection of observed zero-mean r.v.'s

# Characterization

## Our scope

Formalize the concept

Observations = Common (and simple) behavior + Local interactions

## Stochastic framework of generalized factor analysis

$\mathbf{y} := \{\mathbf{y}_k,\ k \in \mathbb{N}\}$ infinite collection of observed zero-mean r.v.'s

**Goal: find a unique decomposition**

$$\mathbf{y}_k = \hat{\mathbf{y}}_k + \tilde{\mathbf{y}}_k \qquad k = 1, 2, \ldots$$

# Characterization

## Our scope

Formalize the concept

Observations = Common (and simple) behavior + Local interactions

## Stochastic framework of generalized factor analysis

$\mathbf{y} := \{\mathbf{y}_k,\ k \in \mathbb{N}\}$ infinite collection of observed zero-mean r.v.'s

**Goal: find a unique decomposition**

$$\mathbf{y}_k = \hat{\mathbf{y}}_k + \tilde{\mathbf{y}}_k \qquad k = 1, 2, \ldots$$

## Question

Which features shall $\hat{\mathbf{y}}_k$ and $\tilde{\mathbf{y}}_k$ have?

# Generalized Factor model

**Vector notation:** $\mathbf{y} = \hat{\mathbf{y}} + \tilde{\mathbf{y}}$

## Definitions

- $\hat{\mathbf{y}} := F\mathbf{x}$ form a *q*-aggregate sequence $\quad (F \in \mathbb{R}^{\infty \times q})$
- $\tilde{\mathbf{y}}$ form idiosyncratic noise

# Generalized Factor model

**Vector notation:** $\mathbf{y} = \hat{\mathbf{y}} + \tilde{\mathbf{y}}$

## Definitions

- $\hat{\mathbf{y}} := F\mathbf{x}$ form a $q$-aggregate sequence $\quad (F \in \mathbb{R}^{\infty \times q})$
- $\tilde{\mathbf{y}}$ form idiosyncratic noise

## Features

- $\mathbf{x} :=$ $q$-dimensional orthonormal random vector (latent factors)
  1. $q$ fixed $\longrightarrow$ "simple" behavior

# Generalized Factor model

**Vector notation:** $\mathbf{y} = \hat{\mathbf{y}} + \tilde{\mathbf{y}}$

## Definitions

- $\hat{\mathbf{y}} := F\mathbf{x}$ form a *q*-aggregate sequence $(F \in \mathbb{R}^{\infty \times q})$
- $\tilde{\mathbf{y}}$ form idiosyncratic noise

## Features

- $\mathbf{x} :=$ $q$-dimensional orthonormal random vector (latent factors)
  1. $q$ fixed $\longrightarrow$ "simple" behavior
- sequence $\tilde{\mathbf{y}}$:
  1. orthogonal to $\hat{\mathbf{y}}$;
  2. has "weak" cross–correlation $\longrightarrow$ $\mathbb{E}[\tilde{\mathbf{y}}_i \tilde{\mathbf{y}}_j] \to 0$ when $|i-j| \uparrow$.

# Covariance matrices decription

## Notation

- $\Sigma$ := infinite covariance matrix of $\mathbf{y}$
- $\Sigma_n$ := covariance matrix of first $n$ components

# Covariance matrices decription

## Notation

- $\Sigma :=$ infinite covariance matrix of $\mathbf{y}$
- $\Sigma_n :=$ covariance matrix of first $n$ components

From previous assumptions

$$\Sigma := \hat{\Sigma} + \tilde{\Sigma}$$

## Features

- $\hat{\Sigma} :=$ covariance matrix of $\hat{\mathbf{y}} \longrightarrow \mathrm{rank}\hat{\Sigma} = q$
- $\tilde{\Sigma} :=$ covariance matrix of $\tilde{\mathbf{y}} \longrightarrow$ "weak" cross-correlations

# Identifiability

$$\mathbf{y} = \hat{\mathbf{y}} + \tilde{\mathbf{y}}$$
$$= \textbf{q-aggregate} + \textbf{idiosyncratic}$$

**How to guarantee uniqueness of the decomposition?**

# Identifiability

$$\mathbf{y} = \hat{\mathbf{y}} + \tilde{\mathbf{y}}$$
$$= \textbf{q-aggregate} + \textbf{idiosyncratic}$$

**How to guarantee uniqueness of the decomposition?**

## Goal

We need to characterize:

1. Idiosyncratic sequences
2. q-aggregate sequences

# Idiosyncratic sequences

## Preliminar concept - Averaging sequences

$\{\mathfrak{a}_n\}_{n\in\mathbb{N}}$ := sequence of elements of $\ell^2$.

$\{\mathfrak{a}_n\}_{n\in\mathbb{N}}$ is an averaging sequence (AS) if $\lim_{n\to\infty} \|\mathfrak{a}_n\|_2 = 0$.

# Idiosyncratic sequences

## Preliminar concept - Averaging sequences

$\{\mathfrak{a}_n\}_{n\in\mathbb{N}} :=$ sequence of elements of $\ell^2$.

$\{\mathfrak{a}_n\}_{n\in\mathbb{N}}$ is an averaging sequence (AS) if $\lim_{n\to\infty} \|\mathfrak{a}_n\|_2 = 0$.

## Example

$\mathfrak{a}_n := \frac{1}{n}[\underbrace{1 \ldots 1}_{n} \; 0 \ldots]^\top$ is an averaging sequence ($\|\mathfrak{a}_n\|^2 = \frac{1}{n}$).

# Idiosyncratic sequences

## Preliminar concept - Averaging sequences

$\{\mathfrak{a}_n\}_{n\in\mathbb{N}} :=$ sequence of elements of $\ell^2$.

$\{\mathfrak{a}_n\}_{n\in\mathbb{N}}$ is an averaging sequence (AS) if $\lim_{n\to\infty} \|\mathfrak{a}_n\|_2 = 0$.

## Example

$\mathfrak{a}_n := \frac{1}{n}[\underbrace{1\ \ldots\ 1}_{n}\ 0\ \ldots]^\top$ is an averaging sequence ($\|\mathfrak{a}_n\|^2 = \frac{1}{n}$).

## Definition

$\tilde{\mathbf{y}}$ is an idiosyncratic sequence if $\lim_{n\to\infty} \|\mathfrak{a}_n^\top \tilde{\mathbf{y}}\| = 0$ for any AS $\mathfrak{a}_n$.

# Meaning of idiosyncratic

## Example 1

$\tilde{\mathbf{y}}$ = white noise (with uniformly bounded variance) $\Rightarrow$ $\tilde{\mathbf{y}}$ idiosyncratic

$$\lim_{n \to \infty} \mathfrak{a}_n^\top \mathrm{diag}\{\sigma_1^2, \sigma_2^2, \ldots, \} \mathfrak{a}_n = 0 \quad \forall \, \mathfrak{a}_n \text{ AS}$$

# Meaning of idiosyncratic

## Example 1

$\tilde{\mathbf{y}}$ = white noise (with uniformly bounded variance) $\Rightarrow$ $\tilde{\mathbf{y}}$ idiosyncratic

$$\lim_{n\to\infty} \mathfrak{a}_n^\top \mathrm{diag}\{\sigma_1^2, \sigma_2^2, \ldots,\}\mathfrak{a}_n = 0 \quad \forall\, \mathfrak{a}_n \text{ AS}$$

## Example 2

$\mathbf{y}$ = sequence with a single latent factor

$$\mathbf{y} = \mathbb{1}\mathbf{x} + \tilde{\mathbf{y}}, \qquad \tilde{\mathbf{y}} \text{ white noise}$$

$$\mathbf{z} := \lim_N \frac{1}{N} \sum_{k=1}^N \mathbf{y}_k \quad \Longrightarrow \quad \mathbf{z} = \mathbf{x}$$

1. We have recovered the latent factor
2. Idiosyncratic noise vanishes by averaging the observations

# A Strong characterization

## Eigenvalues of the sequence

- $\tilde{\lambda}_n :=$ largest (in magnitude) eigenvalue of $\tilde{\Sigma}_n$
- $\tilde{\lambda} := \lim_{n\to\infty} \tilde{\lambda}_n$ (well-defined)

# A Strong characterization

## Eigenvalues of the sequence

- $\tilde{\lambda}_n :=$ largest (in magnitude) eigenvalue of $\tilde{\Sigma}_n$
- $\tilde{\lambda} := \lim_{n\to\infty} \tilde{\lambda}_n$ (well-defined)

## Theorem (Chamberlain, Rotschild, Forni, Lippi)

The following conditions are equivalent:

1. $\tilde{\mathbf{y}}$ is idiosyncratic
2. $\lambda < \infty$
3. $\tilde{\Sigma}$ is a bounded linear operator in $\ell^2$ [B.P.]

# Interpretation

## Corollary

$\tilde{\mathbf{y}}$ is idiosyncratic $\Longrightarrow$ The rows (columns) of $\tilde{\Sigma}$ are square integrable.

## Consequence

If $\tilde{\mathbf{y}}$ is idiosyncratic, then $\mathbb{E}[\tilde{\mathbf{y}}_i \tilde{\mathbf{y}}_j] \to 0$ as $|i - j| \to \infty$.

# Interpretation

**Corollary**

$\tilde{\mathbf{y}}$ is idiosyncratic $\Longrightarrow$ The rows (columns) of $\tilde{\Sigma}$ are square integrable.

**Consequence**

If $\tilde{\mathbf{y}}$ is idiosyncratic, then $\mathbb{E}[\tilde{\mathbf{y}}_i \tilde{\mathbf{y}}_j] \to 0$ as $|i - j| \to \infty$.

**Idiosyncratic $\Longleftrightarrow$ Local interactions**

# Aggregate sequences

## Previuos concept

$\hat{\mathbf{y}}$ is aggregate $\Longrightarrow$ $\hat{\Sigma}$ has "low" rank

# Aggregate sequences

## Previuos concept

$\hat{\mathbf{y}}$ is aggregate $\Longrightarrow$ $\hat{\Sigma}$ has "low" rank

## Question

Does the "low" rank concept guarantee uniqueness of the decomposition q-aggregate + idiosyncratic ?

# Aggregate sequences

## Previuos concept

$\hat{\mathbf{y}}$ is aggregate $\Longrightarrow$ $\hat{\Sigma}$ has "low" rank

## Question

Does the "low" rank concept guarantee uniqueness of the decomposition q-aggregate + idiosyncratic ?

## Example

$$\hat{\mathbf{y}}_k := \alpha^k \mathbf{x},\ |\alpha| < 1$$

# Aggregate sequences

## Previuos concept

$\hat{\mathbf{y}}$ is aggregate $\Longrightarrow$ $\hat{\Sigma}$ has "low" rank

## Question

Does the "low" rank concept guarantee uniqueness of the decomposition q-aggregate + idiosyncratic ?

## Example

$$\hat{\mathbf{y}}_k := \alpha^k \mathbf{x}, \ |\alpha| < 1$$

- $\mathrm{rank}\hat{\Sigma} = 1 \quad \Longrightarrow \quad$ $\hat{\mathbf{y}}$ 1-aggregate?

# Aggregate sequences

## Previuos concept

$\hat{\mathbf{y}}$ is aggregate $\Longrightarrow$ $\hat{\Sigma}$ has "low" rank

## Question

Does the "low" rank concept guarantee uniqueness of the decomposition q-aggregate + idiosyncratic ?

## Example

$$\hat{\mathbf{y}}_k := \alpha^k \mathbf{x},\ |\alpha| < 1$$

- $\mathrm{rank}\hat{\Sigma} = 1 \quad \Longrightarrow \quad \hat{\mathbf{y}}$ 1-aggregate?
- $\hat{\lambda} = \frac{\alpha^2}{1-\alpha^2} \quad \Longrightarrow \quad \hat{\mathbf{y}}$ idiosyncratic!

# Aggregate sequences

## Previuos concept

$\hat{\mathbf{y}}$ is aggregate $\Longrightarrow$ $\hat{\Sigma}$ has "low" rank

## Question

Does the "low" rank concept guarantee uniqueness of the decomposition q-aggregate + idiosyncratic ?

## Example

$$\hat{\mathbf{y}}_k := \alpha^k \mathbf{x},\ |\alpha| < 1$$

- $\mathrm{rank}\hat{\Sigma} = 1 \quad \Longrightarrow \quad \hat{\mathbf{y}}$ 1-aggregate?
- $\hat{\lambda} = \frac{\alpha^2}{1-\alpha^2} \quad \Longrightarrow \quad \hat{\mathbf{y}}$ idiosyncratic!

**Characterization of $q$-aggregate sequences is needed**

# $q$-PD sequences

## Definition

Hilbert space $H := \operatorname{span}\{\mathbf{y}_k,\ k \in \mathbb{N}\}$

# $q$-PD sequences

## Definition

Hilbert space $H := \mathrm{span}\,\{\mathbf{y}_k,\ k \in \mathbb{N}\}$

## Definition [B.P.]

$\mathbf{y}$ := purely deterministic of rank $q$ ($q$-PD) if $\dim H = q$.

# $q$-PD sequences

## Definition

Hilbert space $H := \operatorname{span}\{\mathbf{y}_k,\ k \in \mathbb{N}\}$

## Definition [B.P.]

$\mathbf{y}$ := purely deterministic of rank $q$ ($q$-PD) if $\dim H = q$.

$\mathbf{y} = (q\text{-PD}) \quad \Longrightarrow \quad$ there exists $F = [\, f_1(\cdot),\ f_2(\cdot),\ \dots f_q(\cdot)\,]$ s.t.

$$\mathbf{y}_k = f^\top(k)\mathbf{x} = \sum_{i=1}^{q} f_i(k)\,\mathbf{x}_i\,, \qquad k \in \mathbb{N}$$

# $q$-PD sequences

## Definition

Hilbert space $H := \operatorname{span}\{\mathbf{y}_k,\ k \in \mathbb{N}\}$

## Definition [B.P.]

$\mathbf{y}$ := purely deterministic of rank $q$ ($q$-PD) if $\dim H = q$.

$\mathbf{y} = (q\text{-PD}) \quad \Longrightarrow \quad$ there exists $F = [\, f_1(\cdot),\ f_2(\cdot),\ \ldots f_q(\cdot)\,]$ s.t.

$$\mathbf{y}_k = f^\top(k)\mathbf{x} = \sum_{i=1}^{q} f_i(k)\,\mathbf{x}_i\,, \qquad k \in \mathbb{N}$$

**A ($q$-PD) sequence can be idiosyncratic ($\mathbf{y}_k := \alpha^k \mathbf{x}$)**

# Aggregate sequences

**How to guarantee uniqueness of decomposition?**

# Aggregate sequences

**How to guarantee uniqueness of decomposition?**

Definition [B.P.]

$q-$PD sequence := $q$-aggregate $\Longrightarrow$ all $q$ nonzero eigenvalues infinite.

# Aggregate sequences

**How to guarantee uniqueness of decomposition?**

## Definition [B.P.]

$q-$PD sequence := $q$-aggregate $\Longrightarrow$ all $q$ nonzero eigenvalues infinite.

## Theorem [B.P.]

$q-$PD sequence $(\mathbf{y}_k = \sum_{i=1}^{q} f_i(k)\mathbf{x}_i)$ = $q$-aggregate iff

$$\lim_{n\to\infty} \| f_i^n(\cdot) - \Pi[f_i^n(\cdot) \,|\, \mathcal{F}_i^n] \|_2 = +\infty\,, \quad (1)$$

where

$$\mathcal{F}_i^n = \mathrm{span}\,\{f_j^n(\cdot),\, j = 1,\, \ldots,\, q,\, j \neq i\} \quad (2)$$

# Meaning of the theorem

## Example

- $\mathbf{y} = 2-$PD sequence

$$\mathbf{y}_k := \mathbf{x}_1 + \left(1 - (0.5)^k\right) \mathbf{x}_2$$

# Meaning of the theorem

## Example

- $\mathbf{y} = 2{-}$PD sequence

$$\mathbf{y}_k := \mathbf{x}_1 + \left(1 - (0.5)^k\right)\mathbf{x}_2$$

- $f_1(k) = 1 \quad \forall k \quad , \quad f_2(k) = 1 - (0.5)^k$

# Meaning of the theorem

## Example

- $\mathbf{y} = 2-$PD sequence

$$\mathbf{y}_k := \mathbf{x}_1 + \left(1 - (0.5)^k\right) \mathbf{x}_2$$

- $f_1(k) = 1 \quad \forall k \quad , \quad f_2(k) = 1 - (0.5)^k$
- non zero eigenvalues of $\Sigma_n$ = eigenvalues of Gramian matrix of $f_i^n$'s

$$F^{n\top} F^n = \begin{bmatrix} \|f_1^n\|_2^2 & \langle f_1^n, f_2^n \rangle_2 \\ \langle f_1^n, f_2^n \rangle_2 & \|f_2^n\|_2^2 \end{bmatrix}$$

# Meaning of the theorem

## Example

- $\mathbf{y} = 2{-}\mathrm{PD}$ sequence

$$\mathbf{y}_k := \mathbf{x}_1 + \left(1 - (0.5)^k\right) \mathbf{x}_2$$

- $f_1(k) = 1 \quad \forall k \quad , \quad f_2(k) = 1 - (0.5)^k$
- non zero eigenvalues of $\Sigma_n$ = eigenvalues of Gramian matrix of $f_i^n$'s

$$F^{n\top} F^n = \begin{bmatrix} \|f_1^n\|_2^2 & \langle f_1^n, f_2^n \rangle_2 \\ \langle f_1^n, f_2^n \rangle_2 & \|f_2^n\|_2^2 \end{bmatrix}$$

- $n \to \infty \quad \Longrightarrow \quad$ second eigenvalue $= \frac{5}{3}$.

# $q$-factor sequences

## Definition - Generalized factor model

$\mathbf{y}$ := $q-$factor sequence ($q-$FS) if it admits a representation

$$\mathbf{y} = F\mathbf{x} + \tilde{\mathbf{y}} \qquad , \quad F \in \mathbb{R}^{\infty \times q}$$

# $q$-factor sequences

## Definition - Generalized factor model

$\mathbf{y} :=$ $q-$factor sequence ($q-$FS) if it admits a representation

$$\mathbf{y} = F\mathbf{x} + \tilde{\mathbf{y}} \qquad , \quad F \in \mathbb{R}^{\infty \times q}$$

1. $\hat{\mathbf{y}} := F\mathbf{x} =$ $q$-aggregate sequence

# $q$-factor sequences

## Definition - Generalized factor model

$\mathbf{y}$ := $q-$factor sequence ($q-$FS) if it admits a representation

$$\mathbf{y} = F\mathbf{x} + \tilde{\mathbf{y}} \qquad , \qquad F \in \mathbb{R}^{\infty \times q}$$

1. $\hat{\mathbf{y}} := F\mathbf{x} =$ $q$-aggregate sequence
2. $\tilde{\mathbf{y}} =$ idiosyncratic sequence (orth. to $\mathbf{x}$)

# $q$-factor sequences

## Definition - Generalized factor model

$\mathbf{y}$ := $q-$factor sequence ($q-$FS) if it admits a representation

$$\mathbf{y} = F\mathbf{x} + \tilde{\mathbf{y}} \qquad , \qquad F \in \mathbb{R}^{\infty \times q}$$

1. $\hat{\mathbf{y}} := F\mathbf{x} =$ $q$-aggregate sequence
2. $\tilde{\mathbf{y}} =$ idiosyncratic sequence (orth. to $\mathbf{x}$)

## Theorem (Forni, Lippi)

$\mathbf{y} = q-$FS iff $\lambda_q^y = \infty$ but $\lambda_{q+1}^y$ is bounded $\quad$ *($\lambda_i = i-th$ eigenvalue)*

# $q$-factor sequences

## Definition - Generalized factor model

$\mathbf{y}$ := $q-$factor sequence ($q-$FS) if it admits a representation

$$\mathbf{y} = F\mathbf{x} + \tilde{\mathbf{y}} \qquad , \quad F \in \mathbb{R}^{\infty \times q}$$

1. $\hat{\mathbf{y}} := F\mathbf{x} =$ $q$-aggregate sequence
2. $\tilde{\mathbf{y}} =$ idiosyncratic sequence (orth. to $\mathbf{x}$)

## Theorem (Forni, Lippi)

$\mathbf{y} = q-$FS iff $\lambda_q^y = \infty$ but $\lambda_{q+1}^y$ is bounded $\quad$ *($\lambda_i = i-th$ eigenvalue)*

**Identification: PCA on the $q$ unbounded eigenvalues of $\Sigma$**

# Stationary factor sequences

**Sequences with unbounded variance are ill-posed**

# Stationary factor sequences

**Sequences with unbounded variance are ill-posed**

## Assumption

$\mathbf{y}$ is stationary $\Longrightarrow \quad \mathbb{E}[\mathbf{y}_t \mathbf{y}_s] = r(t-s)$ *($\Rightarrow$ variance uniformely bounded)*

# Stationary factor sequences

**Sequences with unbounded variance are ill-posed**

## Assumption

$\mathbf{y}$ is stationary $\Longrightarrow \quad \mathbb{E}[\mathbf{y}_t \mathbf{y}_s] = r(t-s)$ *($\Rightarrow$ variance uniformely bounded)*

## Our purpose

Compare the $q$-FS decomposition with the Wold decomposition of stationary processes

# The Wold decomposition

## Definitions

$\mathbf{y}$ stationary sequence

# The Wold decomposition

## Definitions

**y** stationary sequence

- remote future: $H_\infty = \bigcap_{t \geq 0} H_t \qquad H_t := \mathrm{span}\left\{\mathbf{y}_k,\ k \geq t\right\}$

# The Wold decomposition

## Definitions

**y** stationary sequence

- remote future: $H_\infty = \bigcap_{t\geq 0} H_t \qquad H_t := \mathrm{span}\,\{\mathbf{y}_k,\ k \geq t\}$
- space of innovations: $\tilde{H} = \bigoplus_{t\geq 0} E_t \qquad E_t := H_t \ominus H_{t+1}$

# The Wold decomposition

## Definitions

$\mathbf{y}$ stationary sequence

- remote future: $H_\infty = \bigcap_{t\geq 0} H_t \qquad H_t := \mathrm{span}\,\{\mathbf{y}_k,\ k \geq t\}$
- space of innovations: $\tilde{H} = \bigoplus_{t\geq 0} E_t \qquad E_t := H_t \ominus H_{t+1}$

## Wold decomposition

Unique decomposition

$$\mathbf{y} = \hat{\mathbf{y}} + \tilde{\mathbf{y}}\,, \qquad \hat{\mathbf{y}}_k \in H_\infty \quad \tilde{\mathbf{y}}_k \in \tilde{H}$$

$$= \text{PD component} \ + \ \text{PND component}$$

# Seeking a correspondence

## Question

$$\begin{aligned} \mathbf{y} &= \text{PD component} \quad + \quad \text{PND component} \\ & \qquad\quad \Updownarrow \qquad\qquad\qquad\qquad \Updownarrow \\ &= \qquad\qquad\qquad\quad + \end{aligned}$$

# Seeking a correspondence

**Question**

$$\begin{aligned} \mathbf{y} &= \text{PD component} &+& \quad \text{PND component} \\ & \qquad\quad \Updownarrow & & \qquad\quad \Updownarrow \\ &= \quad q\text{-aggregate} &+& \quad \text{idiosyncratic} \end{aligned}$$

When does this hold?

# Seeking a correspondence

## Question

$$\begin{array}{ccccc} \mathbf{y} & = & \text{PD component} & + & \text{PND component} \\ & & \Updownarrow & & \Updownarrow \\ & = & q\text{-aggregate} & + & \text{idiosyncratic} \end{array}$$

When does this hold?

## Lemma [B.P.]

$\mathbf{y}$ stationary $+\ S_y(\omega) \in L^\infty([-\pi, \pi]) \implies \mathbf{y}$ idiosyncratic.

# Seeking a correspondence

## Question

$$\begin{array}{ccccc} \mathbf{y} & = & \text{PD component} & + & \text{PND component} \\ & & \Updownarrow & & \Updownarrow \\ & = & q\text{-aggregate} & + & \text{idiosyncratic} \end{array}$$

When does this hold?

## Lemma [B.P.]

$\mathbf{y}$ stationary $+ \; S_y(\omega) \in L^\infty([-\pi, \pi]) \quad \Longrightarrow \quad \mathbf{y}$ idiosyncratic.

**PND processes with bounded spectrum are idiosyncratic**

# Seeking a correspondence

## Question

$$\begin{aligned} \mathbf{y} &= \text{PD component} &+& \quad \text{PND component} \\ & \qquad\Updownarrow & & \qquad\Updownarrow \\ &= \quad q\text{-aggregate} &+& \quad \text{idiosyncratic} \end{aligned}$$

When does this hold?

## Lemma [B.P.]

$\mathbf{y}$ stationary $+\ S_y(\omega) \in L^\infty([-\pi, \pi]) \implies \mathbf{y}$ idiosyncratic.

**PND processes with bounded spectrum are idiosyncratic**

**What about PD processes?**

# Aggregation subspace and remote future

**Definition - Aggregation subspace $\mathcal{G}$**

$z \in \mathcal{G} \quad \Rightarrow \quad z = \lim_{n\to\infty} \mathfrak{a}_n^\top \mathbf{y}$ for a certain AS.

Consequences: $\mathbf{y}$ idiosyncratic $\Rightarrow \mathcal{G} = 0, \quad \mathbf{y} :=$ q-FS $\Rightarrow \dim \mathcal{G} = q$

# Aggregation subspace and remote future

## Definition - Aggregation subspace $\mathcal{G}$

$z \in \mathcal{G} \quad \Rightarrow \quad z = \lim_{n\to\infty} \mathfrak{a}_n^\top \mathbf{y}$ for a certain AS.

Consequences: $\mathbf{y}$ idiosyncratic $\Rightarrow \mathcal{G} = 0, \quad \mathbf{y} :=$ q-FS $\Rightarrow \dim \mathcal{G} = q$

## Lemma [B.P.]

$\mathbf{y}$ stationary $+ S_y(\omega) \in L^\infty([-\pi, \pi]) \Longrightarrow \mathcal{G} \subseteq H_\infty$ $(\mathbf{x}_i \in H_\infty)$.

# Aggregation subspace and remote future

## Definition - Aggregation subspace $\mathcal{G}$

$z \in \mathcal{G} \quad \Rightarrow \quad z = \lim_{n\to\infty} \mathfrak{a}_n^\top \mathbf{y}$ for a certain AS.

Consequences: $\mathbf{y}$ idiosyncratic $\Rightarrow \mathcal{G} = 0, \quad \mathbf{y} :=$ q-FS $\Rightarrow \dim \mathcal{G} = q$

## Lemma [B.P.]

$\mathbf{y}$ stationary $+ S_y(\omega) \in L^\infty([-\pi, \pi]) \implies \mathcal{G} \subseteq H_\infty$ $(\mathbf{x}_i \in H_\infty)$.

$H_\infty \subseteq \mathcal{G}$ **generally not true**

# Aggregation subspace and remote future

**Definition - Aggregation subspace $\mathcal{G}$**

$z \in \mathcal{G} \quad \Rightarrow \quad z = \lim_{n\to\infty} \mathfrak{a}_n^\top \mathbf{y}$ for a certain AS.

Consequences: $\mathbf{y}$ idiosyncratic $\Rightarrow \mathcal{G} = 0, \quad \mathbf{y} :=$ q-FS $\Rightarrow \dim \mathcal{G} = q$

**Lemma [B.P.]**

$\mathbf{y}$ stationary $+ S_y(\omega) \in L^\infty([-\pi, \pi]) \Longrightarrow \mathcal{G} \subseteq H_\infty$ $(\mathbf{x}_i \in H_\infty)$.

$H_\infty \subseteq \mathcal{G}$ **generally not true**

**Lemma [B.P.]**

$\mathbf{y}$ stationary $+ S_y(\omega) \in L^\infty([-\pi, \pi]) + \dim H_\infty < \infty \Rightarrow H_\infty \equiv \mathcal{G}$.

# $q$-FS and Wold decomposition

Using the previous conditions

**Theorem [B.P.]**

$\mathbf{y}$ stationary $+\ S_y(\omega) \in L^\infty([-\pi, \pi]) + \dim H_\infty < \infty$

# $q$-FS and Wold decomposition

Using the previous conditions

**Theorem [B.P.]**

$\mathbf{y}$ stationary $+\ S_y(\omega) \in L^\infty([-\pi, \pi]) + \dim H_\infty < \infty$

$$\Downarrow$$

$\mathbf{y}$ $q$−factor sequence with

$$\begin{array}{cccc} \mathbf{y} & = & \text{PD component} & + & \text{PND component} \\ & & \Updownarrow & & \Updownarrow \\ & = & q\text{-aggregate} & + & \text{idiosyncratic} \end{array}$$

# Flocking interpretion

$$\begin{aligned} \mathbf{y} &= \textbf{common behavior} + \textbf{local interactions} \\ & \qquad\quad \Updownarrow \qquad\qquad\qquad\qquad \Updownarrow \\ &= q\textbf{-aggregate} + \textbf{idiosyncratic} \end{aligned}$$

# Flocking interpretion

$$\begin{array}{ccccc} \mathbf{y} & = & \textbf{common behavior} & + & \textbf{local interactions} \\ & & \Updownarrow & & \Updownarrow \\ & = & q\textbf{-aggregate} & + & \textbf{idiosyncratic} \end{array}$$

## How can we model dynamic flocks?

# Modeling a flock

## Framework

- $\mathbf{y}(k,t) :=$ evolution in time of $k-$th element of a flock

# Modeling a flock

## Framework

- $\mathbf{y}(k,t) :=$ evolution in time of $k-$th element of a flock
- multiplicative structure assumption: $\mathbf{y}(k,t) = \mathbf{v}(k)\mathbf{u}(t)$

# Modeling a flock

## Framework

- $\mathbf{y}(k,t) :=$ evolution in time of $k-$th element of a flock
- multiplicative structure assumption: $\mathbf{y}(k,t) = \mathbf{v}(k)\mathbf{u}(t)$
- $\mathbf{v}(k)$ spatial behavior
- $\mathbf{u}(t)$ evolution in time

# Modeling a flock

## Framework

- $\mathbf{y}(k,t) :=$ evolution in time of $k-$th element of a flock
- multiplicative structure assumption: $\mathbf{y}(k,t) = \mathbf{v}(k)\mathbf{u}(t)$
- $\mathbf{v}(k)$ spatial behavior
- $\mathbf{u}(t)$ evolution in time

## Assumptions

- $\mathbf{v}(k) = \sum_{i=1}^{q} f_i(k)\mathbf{z}_i + \tilde{\mathbf{v}}(k)$ stationary *($\mathbf{v}$ $q-$factor sequence)*
- $\mathbb{E}_{\mathbf{v}}\{\mathbf{v}(k_1)\mathbf{v}(k_2) \mid \mathbf{u}(t_1)\mathbf{u}(t_2)\} = \mathbb{E}_{\mathbf{v}}\{\mathbf{v}(k_1)\mathbf{v}(k_2)\}$

# Estimation of the flocking behavior

## Procedure

1. Take a snapshot of the flock $\mathbf{y}(k, t_0)$

# Estimation of the flocking behavior

## Procedure

1. Take a snapshot of the flock $\mathbf{y}(k, t_0)$
2. Compute the sample covariance

$$\hat{\sigma}_N(h, t_0) := \frac{1}{N}\sum_{k=1}^{N} y(k+h, t_0)y(k, t_0) = \frac{1}{N}\sum_{k=1}^{N} v(k+h)v(k)\, u(t_0)^2\,,$$

# Estimation of the flocking behavior

## Procedure

1. Take a snapshot of the flock $\mathbf{y}(k, t_0)$
2. Compute the sample covariance

$$\hat{\sigma}_N(h, t_0) := \frac{1}{N} \sum_{k=1}^{N} y(k+h, t_0) y(k, t_0) = \frac{1}{N} \sum_{k=1}^{N} v(k+h) v(k)\, u(t_0)^2 \,,$$

3. Compute the sample Toeplitz matrix from sample covariance

# Estimation of the flocking behavior

## Procedure

1. Take a snapshot of the flock $\mathbf{y}(k, t_0)$
2. Compute the sample covariance

$$\hat{\sigma}_N(h, t_0) := \frac{1}{N}\sum_{k=1}^{N} y(k+h, t_0)y(k, t_0) = \frac{1}{N}\sum_{k=1}^{N} v(k+h)v(k)\, u(t_0)^2\,,$$

3. Compute the sample Toeplitz matrix from sample covariance
4. Do PCA on the infinite eigenvalue $\rightarrow$ recover the factor model for $\mathbf{v}$

# Estimation of the flocking behavior

## Procedure

1. Take a snapshot of the flock $\mathbf{y}(k, t_0)$
2. Compute the sample covariance

$$\hat{\sigma}_N(h, t_0) := \frac{1}{N}\sum_{k=1}^{N} y(k+h, t_0)y(k, t_0) = \frac{1}{N}\sum_{k=1}^{N} v(k+h)v(k)\, u(t_0)^2\,,$$

3. Compute the sample Toeplitz matrix from sample covariance
4. Do PCA on the infinite eigenvalue $\rightarrow$ recover the factor model for $\mathbf{v}$
5. Use the interpretation common behavior + local interactions

# Conclusions and future works

## Conclusions

- Rigorous formulation of Generalized factor analysis models
- Connection with stationary sequences and Wold decomposition
- Interpretation in terms of flocking modeling

# Conclusions and future works

## Conclusions

- Rigorous formulation of Generalized factor analysis models
- Connection with stationary sequences and Wold decomposition
- Interpretation in terms of flocking modeling

## Future works

- Connection with nonstationary Wold decomposition
- Modeling of nonstationary flocks

# Publications


1. G. Bottegal, G. Picci and S. Pinzoni. *On the identifiability of errors-in-variables models with white measurement errors.* Automatica, 47(3):545-551, Mar. 2011.
2. G. Bottegal and G. Picci. *A note on generalized factor analysis models.* IEEE CDC-ECC, 2011.
3. G. Bottegal and G. Pillonetto. *Regularized spectrum estimation in spaces induced by stable spline kernels.* IEEE ACC, 2012.
4. B.D.O. Anderson, M. Zamani and G. Bottegal. *On the zero properties of tall linear systems with single-rate and multirate outputs.* IFAC MTNS, 2012.
5. G. Picci, G. Bottegal. *Generalized Factor Analysis Models.* Control Theory: Mathematical Perspectives on Complex Networked Systems, 2012.
6. G. Bottegal and G. Pillonetto. *Regularized spectrum estimation using stable spline kernels.* Automatica (submitted).
7. G. Bottegal and G. Picci. *Flocking and generalized factor analysis.* IEEE ECC, 2013.
8. M. Zamani, G. Bottegal and B.D.O. Anderson. *On the zero freeness of tall multirate linear systems.* Automatica (submitted).
9. G. Bottegal and G. Picci. *Flocking and generalized factor analysis.* IEEE Transactions on Automatic Control (submitted).

# Ph.D. Defense

Giulio Bottegal
Advisor: Prof. Giorgio Picci

Department of Information Engineering - University of Padova

February 28th, 2013

bottegal@dei.unipd.it
automatica.dei.unipd.it/people/bottegal.html