The Kernel Trick for Nonlinear Factor Modeling

Varlam Kutateladze

11th ECB Conference on Forecasting Techniques,
June 2021

Motivation

"Linear" Factor Model

$$\boxed{\underset{N\times 1}{X_t} = \Lambda \underset{r\times 1}{F_t} + e_t}$$

Can use $\widehat{F}$ for forecasting (e.g. AR-DI)

Motivation

"Linear" Factor Model

$$\boxed{\underset{N\times 1}{X_t} = \Lambda \underset{r\times 1}{F_t} + e_t}$$

Can use $\widehat{F}$ for forecasting (e.g. AR-DI)

> *"Further forecast improvements will need to come from models with nonlinearities and/or time variation."*
>
> Stock and Watson 2012

> *"Nonlinear factor-augmented regression should be considered."*
>
> Cheng and Hansen 2015

Motivation

"Linear" Factor Model

$$\boxed{\underset{N\times 1}{X_t} = \Lambda \underset{r\times 1}{F_t} + e_t}$$

Can use $\widehat{F}$ for forecasting (e.g. AR-DI)

> *"Further forecast improvements will need to come from models with nonlinearities and/or time variation."*
>
> Stock and Watson 2012

> *"Nonlinear factor-augmented regression should be considered."*
>
> Cheng and Hansen 2015

✓ Nonlinear models often dominate their linear counterparts: Giovannetti 2013, Kim and Swanson 2014, Coulombe et al. 2019.

Motivation

"Linear" Factor Model

$$\boxed{\underset{N\times 1}{X_t} = \Lambda \underset{r\times 1}{F_t} + e_t}$$

Can use $\widehat{F}$ for forecasting (e.g. AR-DI)

> *"Further forecast improvements will need to come from models with nonlinearities and/or time variation."*
>
> Stock and Watson 2012

> *"Nonlinear factor-augmented regression should be considered."*
>
> Cheng and Hansen 2015

✓ Nonlinear models often dominate their linear counterparts: Giovannetti 2013, Kim and Swanson 2014, Coulombe et al. 2019.

This paper: Nonlinearization via the kernel method.

Kernel trick

Idea: Implicit nonlinearization of inputs $\varphi(\cdot) : \mathcal{X} \to \mathcal{F}$ $(\mathbb{R}^N \to \mathbb{R}^M)$.

How: Substitute $\langle \mathrm{x}_i, \mathrm{x}_j \rangle$ with $\langle \varphi(\mathrm{x}_i), \varphi(\mathrm{x}_j) \rangle = k(\mathrm{x}_i, \mathrm{x}_j)$,

e.g. $k(\mathrm{x_i}, \mathrm{x_j}) = e^{-\gamma \|\mathrm{x_i} - \mathrm{x_j}\|^2}$

"Nonlinear" Factor Model

$$\boxed{\underset{M \times 1}{\varphi(X_t)} = \Lambda \underset{r \times 1}{F_t} + e_t}$$

where $\varphi(\cdot)$ is very flexible and high-dimensional

Kernel factors $\Longrightarrow \widehat{F}_\varphi$

Interesting Results

Proposition 1 (*Simplified*)

Kernel factors and factors by Connor and Korajczyk 1993 when nonlinearized have identical column spaces.

Proposition 2 (*Simplified*)

Kernel factors can nest linear (PCA) factors.

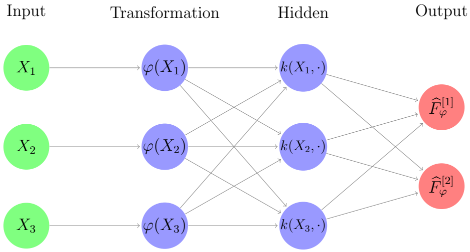


Figure 1: Neural network interpretation

Interesting Results

High-dimensional approximate static factor model

Theorem 1 (Very simplified)

Consistent estimation is possible for kernels with $M < \infty$

Theorem 2 (Very simplified)

Consistent estimation is possible for kernels with $M = \infty$

Interesting Results

Forecasting application:

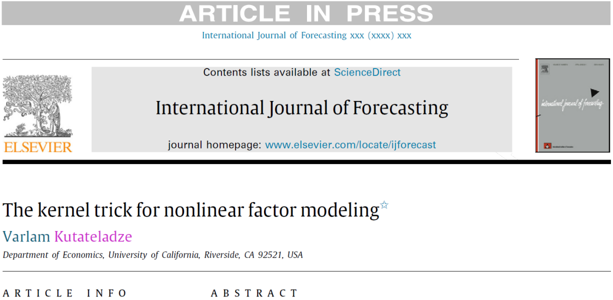
- McCracken and Ng 2016 dataset, 1959:01 to 2020:04
- 8 variables to forecast at $h = 1, 3, 6, 9, 12, 18, 24$
- Competing models:
  - AR-DI with PCA factors (Stock and Watson 2002)
  - AR-DI with SPCA factors (Bai and Ng 2008)
  - AR-DI with PC^2 factors (Bai and Ng 2008)
  - AR-DI with different kernel factors

Main result:

Kernel-based approach generally outperforms the competition, especially at mid to long horizons

Main Takeaways

- Constructing factor estimates nonlinearly can be beneficial forecasting
- Nesting of linear factor estimator
- Connection with neural networks
- Consistency
- Good empirical performance



ARTICLE IN PRESS

International Journal of Forecasting xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Forecasting

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

ELSEVIER

The kernel trick for nonlinear factor modeling☆

Varlam Kutateladze

Department of Economics, University of California, Riverside, CA 92521, USA

ARTICLE INFO

ABSTRACT