Forecasting intraday financial time series with sieve bootstrapping and dynamic updating

Han Lin Shang
Department of Actuarial Studies and Business Analytics
Macquarie University

Kaiying Ji
Discipline of Accounting, Governance and Regulation
The University of Sydney

September 13, 2023
Presentation at the DataX

Intraday financial time series

1. CAPM considers returns using *low-frequency* spot prices, where price changes are ignored

[1] T. Andersen, T. Su, V. Todorov and Z. Zhang (2023+), Intraday periodic volatility curves, *Journal of the American Statistical Association*, in press

[2] D. Donoho and J. Tanner (2009), Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing, *Philosophical Transactions of the Royal Society A*, 367, 4273-4293

Intraday financial time series

1. CAPM considers returns using *low-frequency* spot prices, where price changes are ignored
2. Intraday high-frequency[1] financial data take form of curves that can be sequentially observed over time

---

[1] T. Andersen, T. Su, V. Todorov and Z. Zhang (2023+), Intraday periodic volatility curves, *Journal of the American Statistical Association*, in press

[2] D. Donoho and J. Tanner (2009), Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing, *Philosophical Transactions of the Royal Society A*, 367, 4273-4293

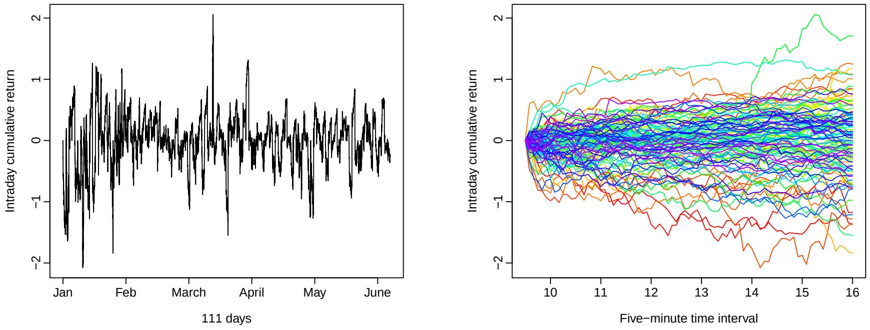
Intraday financial time series

1. CAPM considers returns using *low-frequency* spot prices, where price changes are ignored
2. Intraday high-frequency[1] financial data take form of curves that can be sequentially observed over time
3. High-frequency data give rise to (dense) functional time series -> 'bless of dimensionality' [2]

[1] T. Andersen, T. Su, V. Todorov and Z. Zhang (2023+), Intraday periodic volatility curves, *Journal of the American Statistical Association*, in press

[2] D. Donoho and J. Tanner (2009), Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing, *Philosophical Transactions of the Royal Society A*, 367, 4273-4293

Examples of functional time series (FTS)



- A time series of functions is generated from a stochastic process $\mathcal{X}_t(u)$ where $u \in \mathcal{I} \subset R$, $t \in \mathcal{Z}$
- Modeling temporal dependence within & among functions

Advantages of functional time series

1. Study temporal correlation of an intraday functional object & learn about how correlation progress over days

[3] G. Hooker and S (2022) Selecting the derivative of a functional covariate in scalar-on-function regression, Statistics and Computing, 32(3), 35

Advantages of functional time series

1. Study temporal correlation of an intraday functional object & learn about how correlation progress over days
2. Handle missing values via interpolation or smoothing

[3] G. Hooker and S (2022) Selecting the derivative of a functional covariate in scalar-on-function regression, Statistics and Computing, 32(3), 35

Advantages of functional time series

1. Study temporal correlation of an intraday functional object & learn about how correlation progress over days
2. Handle missing values via interpolation or smoothing
   - Interpolation is fine for dense functional data

[3] G. Hooker and S (2022) Selecting the derivative of a functional covariate in scalar-on-function regression, Statistics and Computing, 32(3), 35

Advantages of functional time series

1. Study temporal correlation of an intraday functional object & learn about how correlation progress over days
2. Handle missing values via interpolation or smoothing
   - Interpolation is fine for dense functional data
   - Smoothing is needed for sparse functional data

[3] G. Hooker and S (2022) Selecting the derivative of a functional covariate in scalar-on-function regression, Statistics and Computing, 32(3), 35

Advantages of functional time series

1. Study temporal correlation of an intraday functional object & learn about how correlation progress over days
2. Handle missing values via interpolation or smoothing
   - Interpolation is fine for dense functional data
   - Smoothing is needed for sparse functional data
3. Study not only level but also derivatives[3] of functions –> dynamic modeling

[3] G. Hooker and S (2022) Selecting the derivative of a functional covariate in scalar-on-function regression, Statistics and Computing, 32(3), 35

Road map

1. Introduce a functional time-series forecasting method for one-day-ahead prediction

Road map

1. Introduce a functional time-series forecasting method for one-day-ahead prediction
2. When partially observed data in most recent day becomes available, incorporate them to improve forecast accuracy

Road map

1. Introduce a functional time-series forecasting method for one-day-ahead prediction
2. When partially observed data in most recent day becomes available, incorporate them to improve forecast accuracy
3. Apply a sieve bootstrap method for uncertainty quantification

Data

1. S&P/ASX All Ordinaries (XAO), 500 largest companies in Australian equities market

[4] Hansen and Lunde (2006) Realized variance and market microstructure noise, JBES, 24(2), 127-161

Data

1. S&P/ASX All Ordinaries (XAO), 500 largest companies in Australian equities market
2. According to UBS, Australian market is highly concentrated, with financial, resources, technology companies

[4] Hansen and Lunde (2006) Realized variance and market microstructure noise, JBES, 24(2), 127-161

Data

1. S&P/ASX All Ordinaries (XAO), 500 largest companies in Australian equities market
2. According to UBS, Australian market is highly concentrated, with financial, resources, technology companies
3. As first major financial market to open each day, it's a world leader is raising capital & active managers with chances to beat market

[4] Hansen and Lunde (2006) Realized variance and market microstructure noise, JBES, 24(2), 127-161

Data

1. S&P/ASX All Ordinaries (XAO), 500 largest companies in Australian equities market
2. According to UBS, Australian market is highly concentrated, with financial, resources, technology companies
3. As first major financial market to open each day, it's a world leader is raising capital & active managers with chances to beat market
4. 5-minute[4] intraday close prices of XAO from January 4 to December 23, 2021 from Refinitiv

[4] Hansen and Lunde (2006) Realized variance and market microstructure noise, JBES, 24(2), 127-161

Cumulative intraday return (CIDR)

1. Let $P_t(u_i), t \in \mathbb{Z}_+, i = 2, \dots, \tau$, $\tau = 75$ be 5-minute close price of XAO at intraday time u_i between 10:00 & 16:10 Sydney time on day t

Cumulative intraday return (CIDR)

1. Let $P_t(u_i), t \in \mathbb{Z}_+, i = 2, \dots, \tau,\ \tau = 75$ be 5-minute close price of XAO at intraday time u_i between 10:00 & 16:10 Sydney time on day t
2. Any overnight trading will be reflected at beginning close price next day

Cumulative intraday return (CIDR)

1. Let $P_t(u_i), t \in \mathbb{Z}_+, i = 2, \dots, \tau,\ \tau = 75$ be 5-minute close price of XAO at intraday time u_i between 10:00 & 16:10 Sydney time on day t
2. Any overnight trading will be reflected at beginning close price next day
3. Apply a functional KPSS test of Horváth et al. (2014), series is trend stationary with p-value of 0.737

Cumulative intraday return (CIDR)

1. Let $P_t(u_i), t \in \mathbb{Z}_+, i = 2, \dots, \tau$, $\tau = 75$ be 5-minute close price of XAO at intraday time u_i between 10:00 & 16:10 Sydney time on day t
2. Any overnight trading will be reflected at beginning close price next day
3. Apply a functional KPSS test of Horváth et al. (2014), series is trend stationary with p-value of 0.737
4. For a stationary series, compute CIDR

$$\mathcal{X}_t(u_i) = 100 \times [\ln P_t(u_i) - \ln P_t(u_1)]$$

Cumulative intraday return (CIDR)

1. Let $P_t(u_i), t \in \mathbb{Z}_+, i = 2, \dots, \tau$, $\tau = 75$ be 5-minute close price of XAO at intraday time u_i between 10:00 & 16:10 Sydney time on day t
2. Any overnight trading will be reflected at beginning close price next day
3. Apply a functional KPSS test of Horváth et al. (2014), series is trend stationary with p-value of 0.737
4. For a stationary series, compute CIDR

$$\mathcal{X}_t(u_i) = 100 \times [\ln P_t(u_i) - \ln P_t(u_1)]$$

5. Via inverse transformation,

$$P_t(u_i) = \exp^{\frac{\mathcal{X}_t(u_i)}{100}} \times P_t(u_1)$$

Functional principal component regression

1. For a time series of functions $[\mathcal{X}_1(u), \dots, \mathcal{X}_n(u)]$, mean function

$$\overline{\mathcal{X}}(u) = \frac{1}{n}\sum_{t=1}^{n} \mathcal{X}_t(u)$$

Functional principal component regression

1. For a time series of functions $[\mathcal{X}_1(u), \dots, \mathcal{X}_n(u)]$, mean function

$$\overline{\mathcal{X}}(u) = \frac{1}{n}\sum_{t=1}^{n}\mathcal{X}_t(u)$$

2. Covariance function is

$$\text{cov}[\mathcal{X}(u), \mathcal{X}(v)] = \text{E}\{[\mathcal{X}(u) - \overline{\mathcal{X}}(u)][\mathcal{X}(v) - \overline{\mathcal{X}}(v)]\}$$

Mercer's lemma

Covariance function can be approximated by orthonormal eigenfunctions

$$\text{cov}[\mathcal{X}(u), \mathcal{X}(v)] = \sum_{k=1}^{\infty} \widehat{\lambda}_k \widehat{\phi}_k(u) \widehat{\phi}_k(v)$$

- $\widehat{\phi}_k(u)$: k^{th} orthonormal functional principal components
- $\widehat{\lambda}_k$: k^{th} eigenvalue

Karhunen-Loève expansion

1. Any functional realization $\mathcal{X}_t(u)$ can be expressed

$$\begin{aligned}\mathcal{X}_t(u) &= \overline{\mathcal{X}}(u) + \sum_{k=1}^{\infty} \underbrace{\widehat{\beta}_{t,k}}_{\langle \mathcal{X}_t(u) - \overline{\mathcal{X}}(u), \widehat{\phi}_k(u)\rangle} \widehat{\phi}_k(u) \\ &= \overline{\mathcal{X}}(u) + \sum_{k=1}^{K} \widehat{\beta}_{t,k}\widehat{\phi}_k(u) + e_t(u)\end{aligned}$$

Karhunen-Loève expansion

1. Any functional realization $\mathcal{X}_t(u)$ can be expressed

$$\mathcal{X}_t(u) = \overline{\mathcal{X}}(u) + \sum_{k=1}^{\infty} \underbrace{\widehat{\beta}_{t,k}}_{\langle \mathcal{X}_t(u) - \overline{\mathcal{X}}(u), \widehat{\phi}_k(u) \rangle} \widehat{\phi}_k(u)$$

$$= \overline{\mathcal{X}}(u) + \sum_{k=1}^{K} \widehat{\beta}_{t,k} \widehat{\phi}_k(u) + e_t(u)$$

- K: retained number of principal components

Karhunen-Loève expansion

1. Any functional realization $\mathcal{X}_t(u)$ can be expressed

$$\mathcal{X}_t(u) = \overline{\mathcal{X}}(u) + \sum_{k=1}^{\infty} \underbrace{\widehat{\beta}_{t,k}}_{\langle \mathcal{X}_t(u) - \overline{\mathcal{X}}(u), \widehat{\phi}_k(u) \rangle} \widehat{\phi}_k(u)$$

$$= \overline{\mathcal{X}}(u) + \sum_{k=1}^{K} \widehat{\beta}_{t,k} \widehat{\phi}_k(u) + e_t(u)$$

- K: retained number of principal components
- $e_t(u)$: error term

Eigenvalue ratio criterion

2. K is selected

$$K = \underset{1\leq k\leq k_{\max}}{\arg\min}\left\{\frac{\widehat{\lambda}_{k+1}}{\widehat{\lambda}_k}\times \mathbb{1}\left(\frac{\widehat{\lambda}_k}{\widehat{\lambda}_1}\geq \upsilon\right) + \mathbb{1}\left(\frac{\widehat{\lambda}_k}{\widehat{\lambda}_1} < \upsilon\right)\right\},$$

Eigenvalue ratio criterion

2. K is selected

$$K = \underset{1\leq k\leq k_{\max}}{\arg\min}\left\{\frac{\widehat{\lambda}_{k+1}}{\widehat{\lambda}_k}\times \mathbb{1}(\frac{\widehat{\lambda}_k}{\widehat{\lambda}_1}\geq \upsilon) + \mathbb{1}(\frac{\widehat{\lambda}_k}{\widehat{\lambda}_1} < \upsilon)\right\},$$

- $\upsilon = 1/\ln[\max(\widehat{\lambda}_1, n)]$ is a pre-specified positive number

Eigenvalue ratio criterion

2. K is selected

$$K = \underset{1\leq k\leq k_{\max}}{\arg\min}\left\{\frac{\widehat{\lambda}_{k+1}}{\widehat{\lambda}_k}\times \mathbb{1}\left(\frac{\widehat{\lambda}_k}{\widehat{\lambda}_1}\geq \upsilon\right) + \mathbb{1}\left(\frac{\widehat{\lambda}_k}{\widehat{\lambda}_1} < \upsilon\right)\right\},$$

- $\upsilon = 1/\ln[\max(\widehat{\lambda}_1, n)]$ is a pre-specified positive number
- $k_{\max} = \#\{k|\widehat{\lambda}_k \geq \sum_{k=1}^{n}\widehat{\lambda}_k/n\}$

Eigenvalue ratio criterion

2. K is selected

$$K = \mathop{\arg\min}_{1\leq k\leq k_{\max}} \left\{ \frac{\widehat{\lambda}_{k+1}}{\widehat{\lambda}_k} \times \mathbb{1}(\frac{\widehat{\lambda}_k}{\widehat{\lambda}_1} \geq \upsilon) + \mathbb{1}(\frac{\widehat{\lambda}_k}{\widehat{\lambda}_1} < \upsilon) \right\},$$

- $\upsilon = 1/\ln[\max(\widehat{\lambda}_1, n)]$ is a pre-specified positive number
- $k_{\max} = \#\{k|\widehat{\lambda}_k \geq \sum_{k=1}^{n} \widehat{\lambda}_k/n\}$
- $\mathbb{1}\{\cdot\}$: binary indicator function.

VAR(p) model

3. Let $\widehat{\boldsymbol{\beta}} = (\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2, \dots, \widehat{\boldsymbol{\beta}}_K)$

VAR(p) model

3. Let $\widehat{\boldsymbol{\beta}} = (\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2, \dots, \widehat{\boldsymbol{\beta}}_K)$
4. VAR(p) model

$$\widehat{\boldsymbol{\beta}}_t = \sum_{\xi=1}^{p} \widehat{\boldsymbol{A}}_{\xi,p} \widehat{\boldsymbol{\beta}}_{t-\xi} + \widehat{\boldsymbol{\epsilon}}_t, t = p+1, \dots, n$$

VAR(p) model

3 Let $\widehat{\boldsymbol{\beta}} = (\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2, \dots, \widehat{\boldsymbol{\beta}}_K)$

4 VAR(p) model

$$\widehat{\boldsymbol{\beta}}_t = \sum_{\xi=1}^{p} \widehat{\boldsymbol{A}}_{\xi,p}\widehat{\boldsymbol{\beta}}_{t-\xi} + \widehat{\boldsymbol{\epsilon}}_t, t = p+1, \dots, n$$

- $\widehat{\boldsymbol{A}}_{\xi,p}$: $(K \times K)$ coefficient matrix of forward score series

VAR(p) model

3. Let $\widehat{\boldsymbol{\beta}} = (\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2, \dots, \widehat{\boldsymbol{\beta}}_K)$
4. VAR(p) model

$$\widehat{\boldsymbol{\beta}}_t = \sum_{\xi=1}^{p} \widehat{\boldsymbol{A}}_{\xi,p} \widehat{\boldsymbol{\beta}}_{t-\xi} + \widehat{\boldsymbol{\epsilon}}_t, t = p+1, \dots, n$$

- $\widehat{\boldsymbol{A}}_{\xi,p}$: $(K \times K)$ coefficient matrix of forward score series
- $(\widehat{\boldsymbol{\epsilon}}_{p+1}, \dots, \widehat{\boldsymbol{\epsilon}}_n)$: residuals after fitting VAR(p) model to K-dimensional multivariate time series of scores

Order selection

1. Order p of VAR model can be chosen from AIC_c by minimizing

$$\text{AIC}_\text{c}(p) = n \ln |\widehat{\boldsymbol{\Sigma}}_{\widehat{\epsilon},p}| + \frac{n(nK + pK^2)}{n - K(p+1) - 1},$$

over a set of $p = \{1, 2, \ldots, 10\}$

Order selection

1. Order p of VAR model can be chosen from AIC_c by minimizing

$$\text{AIC}_\text{c}(p) = n \ln |\widehat{\mathbf{\Sigma}}_{\widehat{\epsilon},p}| + \frac{n(nK + pK^2)}{n - K(p+1) - 1},$$

over a set of $p = \{1, 2, \dots, 10\}$

2. After fitting VAR(p), compute residuals $\widehat{\mathbf{\Sigma}}_{\widehat{\epsilon},p} = \frac{1}{n-p} \sum_{t=p+1}^{n} \widehat{\boldsymbol{\epsilon}}_t \widehat{\boldsymbol{\epsilon}}_t^\top$

One-step-ahead point forecast

1. Conditional on

One-step-ahead point forecast

1. Conditional on
   - observed time series of functions $\boldsymbol{\mathcal{X}}(u)$

One-step-ahead point forecast

1. Conditional on
   - observed time series of functions $\boldsymbol{\mathcal{X}}(u)$
   - estimated mean function $\overline{\mathcal{X}}(u)$

One-step-ahead point forecast

1. Conditional on
   - observed time series of functions $\boldsymbol{\mathcal{X}}(u)$
   - estimated mean function $\overline{\mathcal{X}}(u)$
   - estimated functional principal components $\boldsymbol{\Phi}(u) = [\widehat{\phi}_1(u), \dots, \widehat{\phi}_K(u)]$

One-step-ahead point forecast

1. Conditional on
   - observed time series of functions $\boldsymbol{\mathcal{X}}(u)$
   - estimated mean function $\overline{\mathcal{X}}(u)$
   - estimated functional principal components $\boldsymbol{\Phi}(u) = [\widehat{\phi}_1(u), \dots, \widehat{\phi}_K(u)]$
2. One-step-ahead forecast is

$$\begin{aligned}\widehat{\mathcal{X}}_{n+1|n}(u) &= \mathsf{E}[\mathcal{X}_{n+1}(u)|\boldsymbol{\mathcal{X}}(u), \overline{\mathcal{X}}(u), \boldsymbol{\Phi}(u)] \\ &= \overline{\mathcal{X}}(u) + \sum_{k=1}^{K} \widehat{\beta}_{n+1|n,k}\widehat{\phi}_k(u)\end{aligned}$$

where $\widehat{\beta}_{n+1|n,k}$: one-step-ahead prediction from VAR(p)

One-step-ahead point forecast

1. Conditional on
   - observed time series of functions $\boldsymbol{\mathcal{X}}(u)$
   - estimated mean function $\overline{\mathcal{X}}(u)$
   - estimated functional principal components $\mathbf{\Phi}(u) = [\widehat{\phi}_1(u), \dots, \widehat{\phi}_K(u)]$
2. One-step-ahead forecast is

$$\begin{aligned}\widehat{\mathcal{X}}_{n+1|n}(u) &= \mathsf{E}[\mathcal{X}_{n+1}(u)|\boldsymbol{\mathcal{X}}(u), \overline{\mathcal{X}}(u), \mathbf{\Phi}(u)] \\ &= \overline{\mathcal{X}}(u) + \sum_{k=1}^{K} \widehat{\beta}_{n+1|n,k}\widehat{\phi}_k(u)\end{aligned}$$

   where $\widehat{\beta}_{n+1|n,k}$: one-step-ahead prediction from VAR(p)
3. If $K = 1$, VAR(p) reduces to AR(p)

Sieve bootstrap

1. zero-mean random element $\mathcal{X}_t$ is generated as

$$\mathcal{X}_t = f(\mathcal{X}_{t-1}, \mathcal{X}_{t-2}, \dots) + \varepsilon_t$$

Sieve bootstrap

1. zero-mean random element $\mathcal{X}_t$ is generated as

$$\mathcal{X}_t = f(\mathcal{X}_{t-1}, \mathcal{X}_{t-2}, \dots) + \varepsilon_t$$

- $f : \mathcal{H}^\infty \rightarrow \mathcal{H}$

Sieve bootstrap

1. zero-mean random element $\mathcal{X}_t$ is generated as

$$\mathcal{X}_t = f(\mathcal{X}_{t-1}, \mathcal{X}_{t-2}, \dots) + \varepsilon_t$$

- $f : \mathcal{H}^\infty \rightarrow \mathcal{H}$
- $\{\varepsilon_t\}$: zero-mean i.i.d. innovation process with $\mathrm{E}\|\varepsilon_t\|^2 < \infty$

Sieve bootstrap

1. zero-mean random element $\mathcal{X}_t$ is generated as

$$\mathcal{X}_t = f(\mathcal{X}_{t-1}, \mathcal{X}_{t-2}, \dots) + \varepsilon_t$$

   - $f : \mathcal{H}^\infty \to \mathcal{H}$
   - $\{\varepsilon_t\}$: zero-mean i.i.d. innovation process with $\mathrm{E}\|\varepsilon_t\|^2 < \infty$

2. Based on last ℓ observed functions, $\mathcal{X}_{n,\ell} = (\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n-\ell+1})$ for $\ell < n$, a predictor

$$\widehat{\mathcal{X}}_{n+1} = \widehat{g}(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n-\ell+1})$$

where $\widehat{g} : \mathcal{H}^\ell \to \mathcal{H}$ estimated operator

Prediction error

3. Prediction error $\mathcal{E}_{n+1} = \mathcal{X}_{n+1} - \widehat{\mathcal{X}}_{n+1}$ given $\mathcal{X}_{n,\ell}$

$$\begin{aligned}\mathcal{E}_{n+1} &= \mathcal{X}_{n+1} - \widehat{\mathcal{X}}_{n+1} \\ &= \vartheta_{n+1} + [f(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots) - g(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n+1-\ell}] + \\ &\quad [g(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n+1-\ell}) - \widehat{g}(\widehat{\mathcal{X}}_n, \widehat{\mathcal{X}}_{n-1}, \dots, \widehat{\mathcal{X}}_{n+1-\ell})] \\ &= \mathcal{E}_{I,n+1} + \mathcal{E}_{M,n+1} + \mathcal{E}_{E,n+1}\end{aligned}$$

Prediction error

3 Prediction error $\mathcal{E}_{n+1} = \mathcal{X}_{n+1} - \widehat{\mathcal{X}}_{n+1}$ given $\mathcal{X}_{n,\ell}$

$$\begin{aligned}\mathcal{E}_{n+1} &= \mathcal{X}_{n+1} - \widehat{\mathcal{X}}_{n+1} \\ &= \vartheta_{n+1} + [f(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots) - g(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n+1-\ell}] + \\ &\quad [g(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n+1-\ell}) - \widehat{g}(\widehat{\mathcal{X}}_n, \widehat{\mathcal{X}}_{n-1}, \dots, \widehat{\mathcal{X}}_{n+1-\ell})] \\ &= \mathcal{E}_{I,n+1} + \mathcal{E}_{M,n+1} + \mathcal{E}_{E,n+1}\end{aligned}$$

- $\mathcal{E}_{I,n+1}$: error attributable to i.i.d. innovation

Prediction error

3 Prediction error $\mathcal{E}_{n+1} = \mathcal{X}_{n+1} - \widehat{\mathcal{X}}_{n+1}$ given $\mathcal{X}_{n,\ell}$

$$\begin{aligned}\mathcal{E}_{n+1} &= \mathcal{X}_{n+1} - \widehat{\mathcal{X}}_{n+1} \\ &= \vartheta_{n+1} + [f(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots) - g(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n+1-\ell}] + \\ &\quad [g(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n+1-\ell}) - \widehat{g}(\widehat{\mathcal{X}}_n, \widehat{\mathcal{X}}_{n-1}, \dots, \widehat{\mathcal{X}}_{n+1-\ell})] \\ &= \mathcal{E}_{I,n+1} + \mathcal{E}_{M,n+1} + \mathcal{E}_{E,n+1}\end{aligned}$$

- $\mathcal{E}_{I,n+1}$: error attributable to i.i.d. innovation
- $\mathcal{E}_{M,n+1}$: model misspecification error

Prediction error

3 Prediction error $\mathcal{E}_{n+1} = \mathcal{X}_{n+1} - \widehat{\mathcal{X}}_{n+1}$ given $\mathcal{X}_{n,\ell}$

$$\begin{aligned}\mathcal{E}_{n+1} &= \mathcal{X}_{n+1} - \widehat{\mathcal{X}}_{n+1} \\ &= \vartheta_{n+1} + [f(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots) - g(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n+1-\ell}] + \\ &\quad [g(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n+1-\ell}) - \widehat{g}(\widehat{\mathcal{X}}_n, \widehat{\mathcal{X}}_{n-1}, \dots, \widehat{\mathcal{X}}_{n+1-\ell})] \\ &= \mathcal{E}_{I,n+1} + \mathcal{E}_{M,n+1} + \mathcal{E}_{E,n+1}\end{aligned}$$

- $\mathcal{E}_{I,n+1}$: error attributable to i.i.d. innovation
- $\mathcal{E}_{M,n+1}$: model misspecification error
- $\mathcal{E}_{E,n+1}$: error attributable to estimation of unknown operator g & random elements $(\mathcal{X}_n, \dots, \mathcal{X}_{n+1-\ell})$ used for one-step-ahead prediction

Prediction error

3 Prediction error $\mathcal{E}_{n+1} = \mathcal{X}_{n+1} - \widehat{\mathcal{X}}_{n+1}$ given $\mathcal{X}_{n,\ell}$

$$\begin{aligned}\mathcal{E}_{n+1} &= \mathcal{X}_{n+1} - \widehat{\mathcal{X}}_{n+1} \\ &= \vartheta_{n+1} + [f(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots) - g(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n+1-\ell}] + \\ &\quad [g(\mathcal{X}_n, \mathcal{X}_{n-1}, \dots, \mathcal{X}_{n+1-\ell}) - \widehat{g}(\widehat{\mathcal{X}}_n, \widehat{\mathcal{X}}_{n-1}, \dots, \widehat{\mathcal{X}}_{n+1-\ell})] \\ &= \mathcal{E}_{I,n+1} + \mathcal{E}_{M,n+1} + \mathcal{E}_{E,n+1}\end{aligned}$$

- $\mathcal{E}_{I,n+1}$: error attributable to i.i.d. innovation
- $\mathcal{E}_{M,n+1}$: model misspecification error
- $\mathcal{E}_{E,n+1}$: error attributable to estimation of unknown operator g & random elements $(\mathcal{X}_n, \dots, \mathcal{X}_{n+1-\ell})$ used for one-step-ahead prediction

4 Ultimate goal: Prediction band $[\widehat{\mathcal{X}}_{n+1}(u) - L_n(u), \widehat{\mathcal{X}}_{n+1}(u) + U_n(u)]$

$$\lim_{n\rightarrow\infty} \text{Pr}(\widehat{\mathcal{X}}_{n+1}(u) - L_n(u) \leq \mathcal{X}_{n+1}(u) \leq \widehat{\mathcal{X}}_{n+1}(u) + U_n(u), \forall u \in \mathcal{I} | \mathcal{X}_{n,\ell}) = 1 - \alpha$$

VAR(p) forward series

1. Sieve bootstrap uses VAR(p) to generate forward score forecasts

$$\boldsymbol{\beta}^*_{n+1} = \sum_{\xi=1}^{p} \widehat{\boldsymbol{A}}_{\xi,p}\boldsymbol{\beta}^*_{n+1-\xi} + \boldsymbol{\epsilon}^*_{n+1}$$

where $\boldsymbol{\beta}^*_{n+1-\xi} = \widehat{\boldsymbol{\beta}}_{n+1-\xi}$ for $n+1-\xi \leq n$

VAR(p) forward series

1. Sieve bootstrap uses VAR(p) to generate forward score forecasts

$$\boldsymbol{\beta}^*_{n+1} = \sum_{\xi=1}^{p} \widehat{\boldsymbol{A}}_{\xi,p}\boldsymbol{\beta}^*_{n+1-\xi} + \boldsymbol{\epsilon}^*_{n+1}$$

where $\boldsymbol{\beta}^*_{n+1-\xi} = \widehat{\boldsymbol{\beta}}_{n+1-\xi}$ for $n+1-\xi \leq n$

2. $\boldsymbol{\epsilon}^*_{n+1}$: i.i.d. resampled from centered residuals $\{\widehat{\epsilon}_t - \overline{\epsilon}, t = p+1,\dots,n\}$

VAR(p) forward series

1. Sieve bootstrap uses VAR(p) to generate forward score forecasts

$$\boldsymbol{\beta}^*_{n+1} = \sum_{\xi=1}^{p} \widehat{\boldsymbol{A}}_{\xi,p}\boldsymbol{\beta}^*_{n+1-\xi} + \boldsymbol{\epsilon}^*_{n+1}$$

   where $\boldsymbol{\beta}^*_{n+1-\xi} = \widehat{\boldsymbol{\beta}}_{n+1-\xi}$ for $n+1-\xi \leq n$
2. $\boldsymbol{\epsilon}^*_{n+1}$: i.i.d. resampled from centered residuals $\{\widehat{\epsilon}_t - \overline{\epsilon}, t = p+1, \dots, n\}$
3. Compute

$$\mathcal{X}^*_{n+1}(u) = \overline{\mathcal{X}}(u) + \sum_{k=1}^{K} \beta^*_{n+1,k}\widehat{\phi}_k(u) + e^*_{n+1}(u)$$

VAR(p) forward series

1. Sieve bootstrap uses VAR(p) to generate forward score forecasts

$$\boldsymbol{\beta}^*_{n+1} = \sum_{\xi=1}^{p} \widehat{\boldsymbol{A}}_{\xi,p}\boldsymbol{\beta}^*_{n+1-\xi} + \boldsymbol{\epsilon}^*_{n+1}$$

   where $\boldsymbol{\beta}^*_{n+1-\xi} = \widehat{\boldsymbol{\beta}}_{n+1-\xi}$ for $n+1-\xi \leq n$
2. $\boldsymbol{\epsilon}^*_{n+1}$: i.i.d. resampled from centered residuals $\{\widehat{\epsilon}_t - \overline{\epsilon}, t = p+1, \dots, n\}$
3. Compute

$$\mathcal{X}^*_{n+1}(u) = \overline{\mathcal{X}}(u) + \sum_{k=1}^{K} \beta^*_{n+1,k}\widehat{\phi}_k(u) + e^*_{n+1}(u)$$

   - $e^*_{n+1}(u)$: iid resampled from $\{e_t(u) - \overline{e}(u)\}$

VAR(p) forward series

1. Sieve bootstrap uses VAR(p) to generate forward score forecasts

$$\boldsymbol{\beta}^*_{n+1} = \sum_{\xi=1}^{p} \widehat{\boldsymbol{A}}_{\xi,p}\boldsymbol{\beta}^*_{n+1-\xi} + \boldsymbol{\epsilon}^*_{n+1}$$

   where $\boldsymbol{\beta}^*_{n+1-\xi} = \widehat{\boldsymbol{\beta}}_{n+1-\xi}$ for $n+1-\xi \leq n$

2. $\boldsymbol{\epsilon}^*_{n+1}$: i.i.d. resampled from centered residuals $\{\widehat{\epsilon}_t - \overline{\epsilon}, t = p+1, \dots, n\}$
3. Compute

$$\mathcal{X}^*_{n+1}(u) = \overline{\mathcal{X}}(u) + \sum_{k=1}^{K} \beta^*_{n+1,k}\widehat{\phi}_k(u) + e^*_{n+1}(u)$$

- $e^*_{n+1}(u)$: iid resampled from $\{e_t(u) - \overline{e}(u)\}$
- $e_t(u) = \mathcal{X}_t(u) - \overline{\mathcal{X}}(u) - \sum_{k=1}^{K} \widehat{\beta}_{t,k}\widehat{\phi}_k(u)$

VAR(p) backward series

1. Because of stationarity, VAR(p) can go backward in time to generate bootstrap samples of scores

$$\widehat{\boldsymbol{\beta}}_t = \sum_{\xi=1}^{p} \widehat{\boldsymbol{B}}_{\xi,p} \widehat{\boldsymbol{\beta}}_{t+\xi} + \boldsymbol{\eta}_t$$

VAR(p) backward series

1. Because of stationarity, VAR(p) can go backward in time to generate bootstrap samples of scores

$$\widehat{\boldsymbol{\beta}}_t = \sum_{\xi=1}^{p} \widehat{\boldsymbol{B}}_{\xi,p} \widehat{\boldsymbol{\beta}}_{t+\xi} + \boldsymbol{\eta}_t$$

- $\widehat{\boldsymbol{B}}_{\xi,p}$: $(K \times K)$ coefficient matrix for backward scores

VAR(p) backward series

1. Because of stationarity, VAR(p) can go backward in time to generate bootstrap samples of scores

$$\widehat{\boldsymbol{\beta}}_t = \sum_{\xi=1}^{p} \widehat{\boldsymbol{B}}_{\xi,p} \widehat{\boldsymbol{\beta}}_{t+\xi} + \boldsymbol{\eta}_t$$

- $\widehat{\boldsymbol{B}}_{\xi,p}$: $(K \times K)$ coefficient matrix for backward scores
- $\boldsymbol{\eta}_t$: VAR error term

VAR(p) backward series

1. Because of stationarity, VAR(p) can go backward in time to generate bootstrap samples of scores

$$\widehat{\boldsymbol{\beta}}_t = \sum_{\xi=1}^{p} \widehat{\boldsymbol{B}}_{\xi,p} \widehat{\boldsymbol{\beta}}_{t+\xi} + \boldsymbol{\eta}_t$$

- $\widehat{\boldsymbol{B}}_{\xi,p}$: $(K \times K)$ coefficient matrix for backward scores
- $\boldsymbol{\eta}_t$: VAR error term

2. Bootstrap samples η_t^*

$$\eta_t^* = \boldsymbol{B}_p(L^{-1})\boldsymbol{A}_p^{-1}(L)\epsilon_t^*$$

VAR(p) backward series

1. Because of stationarity, VAR(p) can go backward in time to generate bootstrap samples of scores

$$\widehat{\boldsymbol{\beta}}_t = \sum_{\xi=1}^{p} \widehat{\boldsymbol{B}}_{\xi,p} \widehat{\boldsymbol{\beta}}_{t+\xi} + \boldsymbol{\eta}_t$$

   - $\widehat{\boldsymbol{B}}_{\xi,p}$: $(K \times K)$ coefficient matrix for backward scores
   - $\boldsymbol{\eta}_t$: VAR error term

2. Bootstrap samples η_t^*

$$\eta_t^* = \boldsymbol{B}_p(L^{-1})\boldsymbol{A}_p^{-1}(L)\epsilon_t^*$$

   - $\boldsymbol{A}_p(z) = \boldsymbol{I}_K - \sum_{\xi=1}^{p} \boldsymbol{A}_{\xi,p} z^{\xi}$

VAR(p) backward series

1. Because of stationarity, VAR(p) can go backward in time to generate bootstrap samples of scores

$$\widehat{\boldsymbol{\beta}}_t = \sum_{\xi=1}^{p} \widehat{\boldsymbol{B}}_{\xi,p}\widehat{\boldsymbol{\beta}}_{t+\xi} + \boldsymbol{\eta}_t$$

   - $\widehat{\boldsymbol{B}}_{\xi,p}$: $(K \times K)$ coefficient matrix for backward scores
   - $\boldsymbol{\eta}_t$: VAR error term

2. Bootstrap samples η_t^*

$$\eta_t^* = \boldsymbol{B}_p(L^{-1})\boldsymbol{A}_p^{-1}(L)\epsilon_t^*$$

   - $\boldsymbol{A}_p(z) = \boldsymbol{I}_K - \sum_{\xi=1}^{p} \boldsymbol{A}_{\xi,p} z^{\xi}$
   - $\boldsymbol{B}_p(z) = \boldsymbol{I}_K - \sum_{\xi=1}^{p} \boldsymbol{B}_{\xi,p} z^{\xi}$

VAR(p) backward series

1. Because of stationarity, VAR(p) can go backward in time to generate bootstrap samples of scores

$$\widehat{\boldsymbol{\beta}}_t = \sum_{\xi=1}^{p} \widehat{\boldsymbol{B}}_{\xi,p} \widehat{\boldsymbol{\beta}}_{t+\xi} + \boldsymbol{\eta}_t$$

- $\widehat{\boldsymbol{B}}_{\xi,p}$: $(K \times K)$ coefficient matrix for backward scores
- $\boldsymbol{\eta}_t$: VAR error term

2. Bootstrap samples η_t^*

$$\eta_t^* = \boldsymbol{B}_p(L^{-1})\boldsymbol{A}_p^{-1}(L)\epsilon_t^*$$

- $\boldsymbol{A}_p(z) = \boldsymbol{I}_K - \sum_{\xi=1}^{p} \boldsymbol{A}_{\xi,p} z^{\xi}$
- $\boldsymbol{B}_p(z) = \boldsymbol{I}_K - \sum_{\xi=1}^{p} \boldsymbol{B}_{\xi,p} z^{\xi}$
- $\boldsymbol{I}_K$: $(K \times K)$ diagonal matrix

VAR(p) bootstrap scores

1. Bootstrap samples for backward series

$$\boldsymbol{\beta}_t^* = \sum_{\xi=1}^{p} \widehat{\boldsymbol{B}}_{\xi,p} \boldsymbol{\beta}_{t+\xi}^* + \boldsymbol{\eta}_t^*$$

VAR(p) bootstrap scores

1. Bootstrap samples for backward series

$$\boldsymbol{\beta}_t^* = \sum_{\xi=1}^{p} \widehat{\boldsymbol{B}}_{\xi,p}\boldsymbol{\beta}_{t+\xi}^* + \boldsymbol{\eta}_t^*$$

2. Bootstrap functional time series

$$\mathcal{X}_t^*(u) = \overline{\mathcal{X}}(u) + \sum_{k=1}^{K} \beta_{t,k}^* \widehat{\phi}_k(u) + e_t^*(u)$$

where $e_t^*(u)$: i.i.d. resampled from $\{e_t(u) - \overline{e}(u)\}$

Anyone has a favor method

FAR(1)

$$\widehat{\mathcal{X}}_{n+1} = \overline{\mathcal{X}}(u) + \gamma[\mathcal{X}_n(u) - \overline{\mathcal{X}}(u)]$$

where γ: bounded linear operator, measuring first-order autocorrelation

$$\widehat{\gamma} = \frac{\widehat{\Gamma}(1)}{\widehat{\Gamma}(0)}$$

$$\widehat{\Gamma}(0) = \frac{1}{n}\sum_{t=1}^{n}[\mathcal{X}_t(u) - \overline{\mathcal{X}}(u)] \otimes [\mathcal{X}_t(u) - \overline{\mathcal{X}}(u)]$$

$$\widehat{\Gamma}(1) = \frac{1}{n}\sum_{t=1}^{n-1}[\mathcal{X}_t(u) - \overline{\mathcal{X}}(u)] \otimes [\mathcal{X}_{t+1}(u) - \overline{\mathcal{X}}(u)]$$

Model calibration error

1. Distribution of prediction error $\mathcal{E}^*_{n+1}(u) = \mathcal{X}^*_{n+1}(u) - \widehat{\mathcal{X}}^*_{n+1}(u)$: proxy for distribution of $\mathcal{E}_{n+1}(u) = \mathcal{X}_{n+1}(u) - \widehat{\mathcal{X}}_{n+1}(u)$ given $[\mathcal{X}_{n-\ell+1}(u), \dots, \mathcal{X}_n(u)]$

Model calibration error

1. Distribution of prediction error $\mathcal{E}^*_{n+1}(u) = \mathcal{X}^*_{n+1}(u) - \widehat{\mathcal{X}}^*_{n+1}(u)$: proxy for distribution of $\mathcal{E}_{n+1}(u) = \mathcal{X}_{n+1}(u) - \widehat{\mathcal{X}}_{n+1}(u)$ given $[\mathcal{X}_{n-\ell+1}(u), \dots, \mathcal{X}_n(u)]$
2. $\widehat{\mathcal{X}}_{n+1}(u)$: one-step-ahead point forecast from the same FAR(1), applied to original functional time series

Model calibration error

1. Distribution of prediction error $\mathcal{E}^*_{n+1}(u) = \mathcal{X}^*_{n+1}(u) - \widehat{\mathcal{X}}^*_{n+1}(u)$: proxy for distribution of $\mathcal{E}_{n+1}(u) = \mathcal{X}_{n+1}(u) - \widehat{\mathcal{X}}_{n+1}(u)$ given $[\mathcal{X}_{n-\ell+1}(u), \dots, \mathcal{X}_n(u)]$
2. $\widehat{\mathcal{X}}_{n+1}(u)$: one-step-ahead point forecast from the same FAR(1), applied to original functional time series
3. From $\mathcal{E}^*_{n+1}(u)$, compute sd, $\sigma^*_{n+1}(u)$

Model calibration error

1. Distribution of prediction error $\mathcal{E}^*_{n+1}(u) = \mathcal{X}^*_{n+1}(u) - \widehat{\mathcal{X}}^*_{n+1}(u)$: proxy for distribution of $\mathcal{E}_{n+1}(u) = \mathcal{X}_{n+1}(u) - \widehat{\mathcal{X}}_{n+1}(u)$ given $[\mathcal{X}_{n-\ell+1}(u), \dots, \mathcal{X}_n(u)]$
2. $\widehat{\mathcal{X}}_{n+1}(u)$: one-step-ahead point forecast from the same FAR(1), applied to original functional time series
3. From $\mathcal{E}^*_{n+1}(u)$, compute sd, $\sigma^*_{n+1}(u)$
4. Normalized statistic

$$V^*_{n+1}(u) = \frac{\mathcal{X}^*_{n+1}(u) - \widehat{\mathcal{X}}^*_{n+1}(u)}{\sigma^*_{n+1}(u)}$$

Model calibration error

1. Distribution of prediction error $\mathcal{E}^*_{n+1}(u) = \mathcal{X}^*_{n+1}(u) - \widehat{\mathcal{X}}^*_{n+1}(u)$: proxy for distribution of $\mathcal{E}_{n+1}(u) = \mathcal{X}_{n+1}(u) - \widehat{\mathcal{X}}_{n+1}(u)$ given $[\mathcal{X}_{n-\ell+1}(u), \dots, \mathcal{X}_n(u)]$
2. $\widehat{\mathcal{X}}_{n+1}(u)$: one-step-ahead point forecast from the same FAR(1), applied to original functional time series
3. From $\mathcal{E}^*_{n+1}(u)$, compute sd, $\sigma^*_{n+1}(u)$
4. Normalized statistic

$$V^*_{n+1}(u) = \frac{\mathcal{X}^*_{n+1}(u) - \widehat{\mathcal{X}}^*_{n+1}(u)}{\sigma^*_{n+1}(u)}$$

5. $V^*_{n+1}(u)$: proxy for distribution of

$$V_{n+1}(u) = \frac{\mathcal{X}_{n+1}(u) - \widehat{\mathcal{X}}_{n+1}(u)}{\sigma_{n+1}(u)}$$

Prediction band

1. Let $M^* = \sup_{u\in\mathcal{I}}|V^*_{n+1}(u)|$, denote $Q^*_{1-\alpha}$ be $(1-\alpha)$ quantile of distribution of M^*

Prediction band

1. Let $M^* = \sup_{u \in \mathcal{I}} |V^*_{n+1}(u)|$, denote $Q^*_{1-\alpha}$ be $(1-\alpha)$ quantile of distribution of M^*
2. $(1-\alpha)$ uniform prediction band for $\mathcal{X}_{n+1}(u)$ is

$$\left[\widehat{\mathcal{X}}_{n+1}(u) - Q^*_{1-\alpha}\sigma^*_{n+1}(u), \widehat{\mathcal{X}}_{n+1}(u) + Q^*_{1-\alpha}\sigma^*_{n+1}(u)\right]$$

Dynamic updating

1. When a functional time series is formed as segments of a univariate time series, most recent curve is observed sequentially

Dynamic updating

1. When a functional time series is formed as segments of a univariate time series, most recent curve is observed sequentially
2. Let first m periods of $\mathcal{X}_{n+1}(u)$ be:
$\mathcal{X}_{n+1}(u_e) = [\mathcal{X}_{n+1}(u_2), \dots, \mathcal{X}_{n+1}(u_m)]^\top$

Dynamic updating

1. When a functional time series is formed as segments of a univariate time series, most recent curve is observed sequentially
2. Let first m periods of $\mathcal{X}_{n+1}(u)$ be:
$\mathcal{X}_{n+1}(u_e) = [\mathcal{X}_{n+1}(u_2), \dots, \mathcal{X}_{n+1}(u_m)]^\top$
3. Update forecasts in remainder of day $n+1$, $\mathcal{X}_{n+1}(u_l)$, $u_l \in (u_m, u_\tau]$

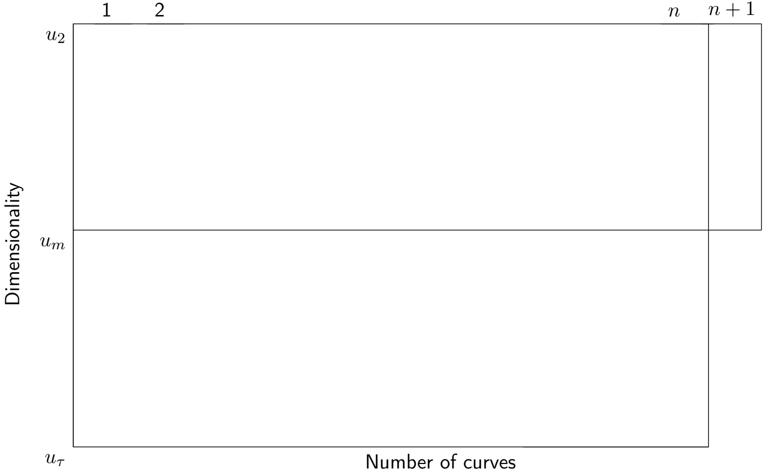


Figure: *Conceptual diagram of dynamic updating.*

Penalized least squares (PLS) method

1. Let $\mathcal{X}^c_{n+1}(u_e) = \mathcal{X}_{n+1}(u_e) - \overline{\mathcal{X}}(u_e)$

Penalized least squares (PLS) method

1. Let $\mathcal{X}^c_{n+1}(u_e) = \mathcal{X}_{n+1}(u_e) - \overline{\mathcal{X}}(u_e)$
2. Shrink regression coefficient estimates towards $\widehat{\beta}^{\text{TS}}_{n+1}$

Penalized least squares (PLS) method

1. Let $\mathcal{X}^c_{n+1}(u_e) = \mathcal{X}_{n+1}(u_e) - \overline{\mathcal{X}}(u_e)$
2. Shrink regression coefficient estimates towards $\widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1}$
3. PLS regression coefficient estimates minimize a penalized residual sum of squares

$$\underset{\boldsymbol{\beta}_{n+1}}{\arg\min}\Big\{[\mathcal{X}^c_{n+1}(u_e) - \boldsymbol{\mathcal{F}}_e\boldsymbol{\beta}_{n+1}]^\top[\mathcal{X}^c_{n+1}(u_e) - \boldsymbol{\mathcal{F}}_e\boldsymbol{\beta}_{n+1}] + \lambda(\boldsymbol{\beta}_{n+1} - \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n})^\top(\boldsymbol{\beta}_{n+1} - \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n})\Big\}$$

Penalized least squares (PLS) method

1. Let $\mathcal{X}^c_{n+1}(u_e) = \mathcal{X}_{n+1}(u_e) - \overline{\mathcal{X}}(u_e)$
2. Shrink regression coefficient estimates towards $\widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1}$
3. PLS regression coefficient estimates minimize a penalized residual sum of squares

$$\underset{\boldsymbol{\beta}_{n+1}}{\arg\min}\Big\{[\mathcal{X}^c_{n+1}(u_e) - \boldsymbol{\mathcal{F}}_e\boldsymbol{\beta}_{n+1}]^\top[\mathcal{X}^c_{n+1}(u_e) - \boldsymbol{\mathcal{F}}_e\boldsymbol{\beta}_{n+1}]+ \lambda(\boldsymbol{\beta}_{n+1} - \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n})^\top(\boldsymbol{\beta}_{n+1} - \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n})\Big\}$$

- $\lambda \in (0, \infty)$: shrinkage parameter

Penalized least squares (PLS) method

1. Let $\mathcal{X}^c_{n+1}(u_e) = \mathcal{X}_{n+1}(u_e) - \overline{\mathcal{X}}(u_e)$
2. Shrink regression coefficient estimates towards $\widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1}$
3. PLS regression coefficient estimates minimize a penalized residual sum of squares

$$\underset{\boldsymbol{\beta}_{n+1}}{\arg\min}\Big\{[\mathcal{X}^c_{n+1}(u_e) - \boldsymbol{\mathcal{F}}_e\boldsymbol{\beta}_{n+1}]^\top[\mathcal{X}^c_{n+1}(u_e) - \boldsymbol{\mathcal{F}}_e\boldsymbol{\beta}_{n+1}] + \lambda(\boldsymbol{\beta}_{n+1} - \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n})^\top(\boldsymbol{\beta}_{n+1} - \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n})\Big\}$$

- $\lambda \in (0, \infty)$: shrinkage parameter
- $\boldsymbol{\mathcal{F}}_e$: $(m \times K)$ matrix, whose $(i,k)^{\text{th}}$ entry is $\widehat{\phi}_k(u_i)$ for $2 \leq i \leq m$

PLS regression coefficient

1. By taking first derivative with respect to $\boldsymbol{\beta}_{n+1}$

$$\widehat{\boldsymbol{\beta}}_{n+1}^{\text{PLS}} = (\boldsymbol{\mathcal{F}}_e^\top \boldsymbol{\mathcal{F}}_e + \lambda \boldsymbol{I}_K)^{-1} \left[\boldsymbol{\mathcal{F}}_e^\top \mathcal{X}_{n+1}^c(u_e) + \lambda \widehat{\boldsymbol{\beta}}_{n+1|n}^{\text{TS}}\right]$$

PLS regression coefficient

1. By taking first derivative with respect to $\boldsymbol{\beta}_{n+1}$

$$\widehat{\boldsymbol{\beta}}^{\text{PLS}}_{n+1} = (\boldsymbol{\mathcal{F}}_e^\top \boldsymbol{\mathcal{F}}_e + \lambda \boldsymbol{I}_K)^{-1}\left[\boldsymbol{\mathcal{F}}_e^\top \mathcal{X}^c_{n+1}(u_e) + \lambda \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n}\right]$$

2. When shrinkage parameter

$$\widehat{\boldsymbol{\beta}}^{\text{PLS}}_{n+1} = \begin{cases} \widehat{\boldsymbol{\beta}}^{\text{OLS}}_{n+1} & \text{if } \lambda \to 0; \\ \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n} & \text{if } \lambda \to \infty; \\ (\widehat{\boldsymbol{\beta}}^{\text{OLS}}_{n+1}, \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n}) & \text{if } 0 < \lambda < \infty. \end{cases}$$

PLS regression coefficient

1. By taking first derivative with respect to $\boldsymbol{\beta}_{n+1}$

$$\widehat{\boldsymbol{\beta}}^{\text{PLS}}_{n+1} = (\boldsymbol{\mathcal{F}}_e^\top \boldsymbol{\mathcal{F}}_e + \lambda \boldsymbol{I}_K)^{-1} \left[\boldsymbol{\mathcal{F}}_e^\top \mathcal{X}^c_{n+1}(u_e) + \lambda \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n}\right]$$

2. When shrinkage parameter

$$\widehat{\boldsymbol{\beta}}^{\text{PLS}}_{n+1} = \begin{cases} \widehat{\boldsymbol{\beta}}^{\text{OLS}}_{n+1} & \text{if } \lambda \rightarrow 0; \\ \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n} & \text{if } \lambda \rightarrow \infty; \\ (\widehat{\boldsymbol{\beta}}^{\text{OLS}}_{n+1}, \widehat{\boldsymbol{\beta}}^{\text{TS}}_{n+1|n}) & \text{if } 0 < \lambda < \infty. \end{cases}$$

3. With optimal λ, PLS forecasts of $\mathcal{X}_{n+1}(u_l)$

$$\widehat{\mathcal{X}}^{\text{PLS}}_{n+1}(u_l) = \overline{\mathcal{X}}(u_l) + \sum_{k=1}^{K} \widehat{\beta}^{\text{PLS}}_{n+1,k} \widehat{\phi}_k(u_l)$$

Selection of λ

Split data into a training set, a validation set, a testing set

$1:150$	$151:200$	$201:250$
Training	Validation	Testing

Updating interval forecasts

1. Bootstrap B samples of TS forecast regression coefficient,

$$\widehat{\boldsymbol{\beta}}^{*,\text{TS}}_{n+1|n} = (\widehat{\beta}^{*,\text{TS}}_{n+1|n,1}, \dots, \widehat{\beta}^{*,\text{TS}}_{n+1|n,K})^\top$$

Updating interval forecasts

1. Bootstrap B samples of TS forecast regression coefficient,

$$\widehat{\boldsymbol{\beta}}^{*,\text{TS}}_{n+1|n} = (\widehat{\beta}^{*,\text{TS}}_{n+1|n,1}, \dots, \widehat{\beta}^{*,\text{TS}}_{n+1|n,K})^\top$$

2. For each $b = 1, \dots, B = 400$

$$\widehat{\mathcal{X}}^{*,\text{PLS}}_{n+1} = \overline{\mathcal{X}}(u_l) + \sum_{k=1}^{K} \widehat{\beta}^{*,\text{PLS}}_{n+1,k} \widehat{\phi}_k(u_l) + e^*_{n+1}(u_l)$$

where $e^*_{n+1}(u_l)$: bootstrapped residuals for updating period

Updating interval forecasts

1. Bootstrap B samples of TS forecast regression coefficient,

$$\widehat{\boldsymbol{\beta}}^{*,\text{TS}}_{n+1|n} = (\widehat{\beta}^{*,\text{TS}}_{n+1|n,1}, \dots, \widehat{\beta}^{*,\text{TS}}_{n+1|n,K})^{\top}$$

2. For each $b = 1, \dots, B = 400$

$$\widehat{\mathcal{X}}^{*,\text{PLS}}_{n+1} = \overline{\mathcal{X}}(u_l) + \sum_{k=1}^{K} \widehat{\beta}^{*,\text{PLS}}_{n+1,k} \widehat{\phi}_k(u_l) + e^*_{n+1}(u_l)$$

where $e^*_{n+1}(u_l)$: bootstrapped residuals for updating period

3. $(1-\alpha)$ PIs for updated forecasts are: $\alpha/2$ & $(1-\alpha/2)$ quantiles of

$$\left\{\widehat{\mathcal{X}}^{1,\text{PLS}}_{n+1}(u_l), \dots, \widehat{\mathcal{X}}^{B,\text{PLS}}_{n+1}(u_l)\right\}$$

Function-on-function linear regression

1. Regression

$$\mathcal{X}_{n+1}^{l}(u) = \overline{\mathcal{X}}^{l}(u) + \int [\mathcal{X}_{n+1}^{e}(v) - \overline{\mathcal{X}}^{e}(v)]\beta(u,v)dv + \xi_{n+1}^{l}(u)$$

Function-on-function linear regression

1. Regression

$$\mathcal{X}_{n+1}^{l}(u) = \overline{\mathcal{X}}^{l}(u) + \int [\mathcal{X}_{n+1}^{e}(v) - \overline{\mathcal{X}}^{e}(v)]\beta(u,v)dv + \xi_{n+1}^{l}(u)$$

- $v \in [u_2, u_m]$ & $u \in (u_m, u_\tau]$: function support ranges for observed & updating periods

Function-on-function linear regression

1. Regression

$$\mathcal{X}_{n+1}^{l}(u) = \overline{\mathcal{X}}^{l}(u) + \int [\mathcal{X}_{n+1}^{e}(v) - \overline{\mathcal{X}}^{e}(v)]\beta(u,v)dv + \xi_{n+1}^{l}(u)$$

- $v \in [u_2, u_m]$ & $u \in (u_m, u_\tau]$: function support ranges for observed & updating periods
- $\overline{\mathcal{X}}^{e}(v)$ & $\overline{\mathcal{X}}^{l}(u)$: mean functions

Function-on-function linear regression

1. Regression

$$\mathcal{X}_{n+1}^{l}(u) = \overline{\mathcal{X}}^{l}(u) + \int [\mathcal{X}_{n+1}^{e}(v) - \overline{\mathcal{X}}^{e}(v)]\beta(u,v)dv + \xi_{n+1}^{l}(u)$$

- $v \in [u_2, u_m]$ & $u \in (u_m, u_\tau]$: function support ranges for observed & updating periods
- $\overline{\mathcal{X}}^{e}(v)$ & $\overline{\mathcal{X}}^{l}(u)$: mean functions
- $\mathcal{X}_{n+1}^{e}(v)$ & $\mathcal{X}_{n+1}^{l}(u)$: functional predictor & response

Function-on-function linear regression

1. Regression

$$\mathcal{X}_{n+1}^{l}(u) = \overline{\mathcal{X}}^{l}(u) + \int [\mathcal{X}_{n+1}^{e}(v) - \overline{\mathcal{X}}^{e}(v)]\beta(u,v)dv + \xi_{n+1}^{l}(u)$$

- $v \in [u_2, u_m]$ & $u \in (u_m, u_\tau]$: function support ranges for observed & updating periods
- $\overline{\mathcal{X}}^{e}(v)$ & $\overline{\mathcal{X}}^{l}(u)$: mean functions
- $\mathcal{X}_{n+1}^{e}(v)$ & $\mathcal{X}_{n+1}^{l}(u)$: functional predictor & response
- $\beta(u,v)$: bivariate regression coefficient function

FPCA

$$\mathcal{X}_t^e(v) = \overline{\mathcal{X}}^e(v) + \sum_{r=1}^{\infty} \widehat{\theta}_{t,r}\widehat{\phi}_r^e(v) = \overline{\mathcal{X}}^e(v) + \sum_{r=1}^{R} \widehat{\theta}_{t,r}\widehat{\phi}_r^e(v) + \kappa_t^e(v)$$

$$\mathcal{X}_t^l(u) = \overline{\mathcal{X}}^l(u) + \sum_{s=1}^{\infty} \widehat{\vartheta}_{t,s}\widehat{\phi}_s^l(u) = \overline{\mathcal{X}}^l(u) + \sum_{s=1}^{S} \widehat{\vartheta}_{t,s}\widehat{\phi}_s^l(u) + \delta_t^l(u)$$

Ordinary least squares (OLS)

1. To estimate $\beta(u,v)$, let $\widehat{\boldsymbol{\theta}} = [\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2, \dots, \widehat{\boldsymbol{\theta}}_R]$, $\widehat{\boldsymbol{\vartheta}} = [\widehat{\boldsymbol{\vartheta}}_1, \widehat{\boldsymbol{\vartheta}}_2, \dots, \widehat{\boldsymbol{\vartheta}}_S]$

Ordinary least squares (OLS)

1. To estimate $\beta(u,v)$, let $\widehat{\boldsymbol{\theta}} = [\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2, \dots, \widehat{\boldsymbol{\theta}}_R]$, $\widehat{\boldsymbol{\vartheta}} = [\widehat{\boldsymbol{\vartheta}}_1, \widehat{\boldsymbol{\vartheta}}_2, \dots, \widehat{\boldsymbol{\vartheta}}_S]$
2. Via OLS, linear relationship between $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\vartheta}}$

$$\widehat{\boldsymbol{\vartheta}} = \widehat{\boldsymbol{\theta}} \times \boldsymbol{\rho}$$

$$\widehat{\boldsymbol{\rho}} = (\widehat{\boldsymbol{\theta}}^\top \widehat{\boldsymbol{\theta}})^{-1} \widehat{\boldsymbol{\theta}}^\top \widehat{\boldsymbol{\vartheta}}$$

Ordinary least squares (OLS)

1. To estimate $\beta(u,v)$, let $\widehat{\boldsymbol{\theta}} = [\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2, \dots, \widehat{\boldsymbol{\theta}}_R]$, $\widehat{\boldsymbol{\vartheta}} = [\widehat{\boldsymbol{\vartheta}}_1, \widehat{\boldsymbol{\vartheta}}_2, \dots, \widehat{\boldsymbol{\vartheta}}_S]$
2. Via OLS, linear relationship between $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\vartheta}}$

$$\widehat{\boldsymbol{\vartheta}} = \widehat{\boldsymbol{\theta}} \times \boldsymbol{\rho}$$
$$\widehat{\boldsymbol{\rho}} = (\widehat{\boldsymbol{\theta}}^\top \widehat{\boldsymbol{\theta}})^{-1} \widehat{\boldsymbol{\theta}}^\top \widehat{\boldsymbol{\vartheta}}$$

3. One-step-ahead forecast of $\mathcal{X}_{n+1}^l(u)$

$$\widehat{\mathcal{X}}_{n+1}^l(u) = \overline{\mathcal{X}}^l(u) + \sum_{s=1}^{S} \widehat{\vartheta}_{n+1,s} \widehat{\phi}_s^l(u)$$
$$\approx \overline{\mathcal{X}}^l(u) + \widehat{\boldsymbol{\theta}}_{n+1} \times \widehat{\boldsymbol{\rho}} \times \widehat{\boldsymbol{\phi}}^l(u)$$

Updating interval forecasts

1. One-step-ahead interval forecast of $\mathcal{X}^l_{n+1}(u)$ is

$$\widehat{\mathcal{X}}^{l,*}_{n+1}(u) = \overline{\mathcal{X}}^l(u) + \int [\mathcal{X}^e_{n+1}(v) - \overline{\mathcal{X}}^e(v)]\widehat{\beta}^*(u,v)dv + e^{l,*}_{n+1}(u)$$

Updating interval forecasts

1. One-step-ahead interval forecast of $\mathcal{X}^l_{n+1}(u)$ is

$$\widehat{\mathcal{X}}^{l,*}_{n+1}(u) = \overline{\mathcal{X}}^l(u) + \int [\mathcal{X}^e_{n+1}(v) - \overline{\mathcal{X}}^e(v)]\widehat{\beta}^*(u,v)dv + e^{l,*}_{n+1}(u)$$

   - $\widehat{\beta}^*(u,v)$: bootstrap regression coefficient estimates

Updating interval forecasts

1. One-step-ahead interval forecast of $\mathcal{X}^l_{n+1}(u)$ is

$$\widehat{\mathcal{X}}^{l,*}_{n+1}(u) = \overline{\mathcal{X}}^l(u) + \int [\mathcal{X}^e_{n+1}(v) - \overline{\mathcal{X}}^e(v)]\widehat{\beta}^*(u,v)dv + e^{l,*}_{n+1}(u)$$

- $\widehat{\beta}^*(u,v)$: bootstrap regression coefficient estimates
- $e^{l,*}_{n+1}(u)$: bootstrap residuals for updating period

Updating interval forecasts

1. One-step-ahead interval forecast of $\mathcal{X}^{l}_{n+1}(u)$ is

$$\widehat{\mathcal{X}}^{l,*}_{n+1}(u) = \overline{\mathcal{X}}^{l}(u) + \int [\mathcal{X}^{e}_{n+1}(v) - \overline{\mathcal{X}}^{e}(v)]\widehat{\beta}^{*}(u,v)dv + e^{l,*}_{n+1}(u)$$

   - $\widehat{\beta}^{*}(u,v)$: bootstrap regression coefficient estimates
   - $e^{l,*}_{n+1}(u)$: bootstrap residuals for updating period

2. Via sieve bootstrap, obtain bootstrap curves, and then estimate $\widehat{\beta}^{*}(u,v)$

Updating interval forecasts

1. One-step-ahead interval forecast of $\mathcal{X}_{n+1}^{l}(u)$ is

$$\widehat{\mathcal{X}}_{n+1}^{l,*}(u) = \overline{\mathcal{X}}^{l}(u) + \int [\mathcal{X}_{n+1}^{e}(v) - \overline{\mathcal{X}}^{e}(v)]\widehat{\beta}^{*}(u,v)dv + e_{n+1}^{l,*}(u)$$

   - $\widehat{\beta}^{*}(u,v)$: bootstrap regression coefficient estimates
   - $e_{n+1}^{l,*}(u)$: bootstrap residuals for updating period

2. Via sieve bootstrap, obtain bootstrap curves, and then estimate $\widehat{\beta}^{*}(u,v)$
3. $(1-\alpha)$ PI for updated forecasts are $\alpha/2$ & $(1-\alpha/2)$ quantiles of $\{\widehat{\mathcal{X}}_{n+1}^{l,1}(u), \dots, \widehat{\mathcal{X}}_{n+1}^{l,B}(u)\}$

Expanding-window scheme

1. Initial training samples are curves from Days 1 to 200, compute one-day-ahead forecast

Expanding-window scheme

1. Initial training samples are curves from Days 1 to 200, compute one-day-ahead forecast
2. Increase training samples from Days 1 to 201, compute one-day-ahead forecast

Expanding-window scheme

1. Initial training samples are curves from Days 1 to 200, compute one-day-ahead forecast
2. Increase training samples from Days 1 to 201, compute one-day-ahead forecast
3. Iterate this procedure until training samples cover entire 250 days

Mean squared forecast error

1. MSFE measures closeness of forecasts compared with actual values of variable being forecast

$$\text{MSFE}(u_i) = \frac{1}{n_{\text{test}}}\sum_{\iota=1}^{n_{\text{test}}}\left[\mathcal{X}_\iota(u_i) - \widehat{\mathcal{X}}_\iota(u_i)\right]^2$$

$$\text{MSFE} = \frac{1}{\tau - 1}\sum_{i=2}^{\tau}\text{MSFE}(u_i)$$

Mean squared forecast error

1. MSFE measures closeness of forecasts compared with actual values of variable being forecast

$$\text{MSFE}(u_i) = \frac{1}{n_{\text{test}}} \sum_{\iota=1}^{n_{\text{test}}} \left[\mathcal{X}_\iota(u_i) - \widehat{\mathcal{X}}_\iota(u_i)\right]^2$$

$$\text{MSFE} = \frac{1}{\tau - 1} \sum_{i=2}^{\tau} \text{MSFE}(u_i)$$

- $\mathcal{X}_\iota$: holdout samples for i^{th} intraday period on ι day

Mean squared forecast error

1. MSFE measures closeness of forecasts compared with actual values of variable being forecast

$$\text{MSFE}(u_i) = \frac{1}{n_{\text{test}}} \sum_{\iota=1}^{n_{\text{test}}} \left[\mathcal{X}_{\iota}(u_i) - \widehat{\mathcal{X}}_{\iota}(u_i)\right]^2$$

$$\text{MSFE} = \frac{1}{\tau - 1} \sum_{i=2}^{\tau} \text{MSFE}(u_i)$$

- $\mathcal{X}_{\iota}$: holdout samples for i^{th} intraday period on ι day
- $n_{\text{test}} = 50$: number of curves in forecasting period

Empirical coverage probability (ECP)

$$\text{ECP}_{\text{pointwise}} = 1 - \frac{1}{n_{\text{test}} \times (\tau - 1)} \sum_{\iota=1}^{n_{\text{test}}} \sum_{i=2}^{\tau} \left[\mathbb{1}\{\mathcal{X}_\iota(u_i) < \widehat{\mathcal{X}}_\iota^{\text{lb}}(u_i)\} + \mathbb{1}\{\mathcal{X}_\iota(u_i) > \widehat{\mathcal{X}}_\iota^{\text{ub}}(u_i)\} \right]$$

$$\text{ECP}_{\text{uniform}} = 1 - \frac{1}{n_{\text{test}}} \sum_{\iota=1}^{n_{\text{test}}} \left[\mathbb{1}\{\mathcal{X}_\iota(u) < \widehat{\mathcal{X}}_\iota^{\text{lb}}(u)\} + \mathbb{1}\{\mathcal{X}_\iota(u) > \widehat{\mathcal{X}}_\iota^{\text{ub}}(u)\} \right].$$

Uniform prediction bands are wider than pointwise PIs

Interval score[5]

$$
\begin{aligned}
S_\alpha\left[\widehat{\mathcal{X}}_\iota^{\text{lb}}(u_i), \widehat{\mathcal{X}}_\iota^{\text{ub}}(u_i), \mathcal{X}_\iota(u_i)\right] = &\left[\widehat{\mathcal{X}}_\iota^{\text{ub}}(u_i) - \widehat{\mathcal{X}}_\iota^{\text{lb}}(u_i)\right] \\
&+ \frac{2}{\alpha}\left[\widehat{\mathcal{X}}_\iota^{\text{lb}}(u_i) - \mathcal{X}_\iota(u_i)\right] \mathbb{1}\left\{\mathcal{X}_\iota(u_i) < \widehat{\mathcal{X}}_\iota^{\text{lb}}(u_i)\right\} \\
&+ \frac{2}{\alpha}\left[\mathcal{X}_\iota(u_i) - \widehat{\mathcal{X}}_\iota^{\text{ub}}(u_i)\right] \mathbb{1}\left\{\mathcal{X}_\iota(u_i) > \widehat{\mathcal{X}}_\iota^{\text{ub}}(u_i)\right\}.
\end{aligned}
$$

[5] T. Gneiting and A. E. Raftery (2007) Strictly proper scoring rules, prediction, and estimation, *Journal of the American Statistical Association*, 102(477), 359-378

Mean interval score

Averaged over different days in forecasting period, mean interval score

$$\overline{S}_\alpha(u_i) = \frac{1}{n_{\text{test}}}\sum_{\iota=1}^{n_{\text{test}}} S_\alpha\left[\widehat{\mathcal{X}}_\iota^{\text{lb}}(u_i), \widehat{\mathcal{X}}_\iota^{\text{ub}}(u_i), \mathcal{X}_\iota(u_i)\right]$$

$$\overline{S}_\alpha = \frac{1}{\tau-1}\sum_{i=2}^{\tau}\overline{S}_\alpha(u_i).$$

An illustration

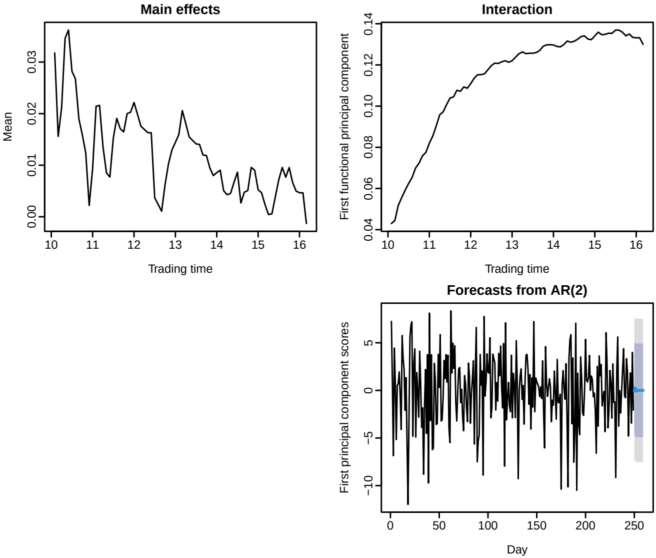
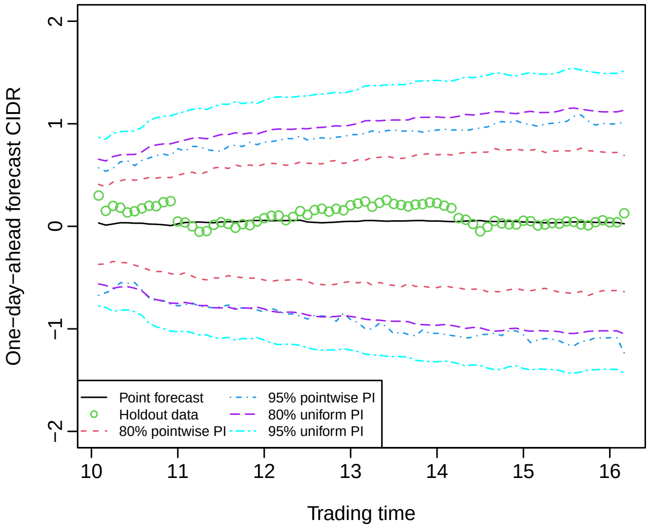


Figure: *From January 4 to December 22, 2021, forecast CIDR for December 23*

Forecast for December 23



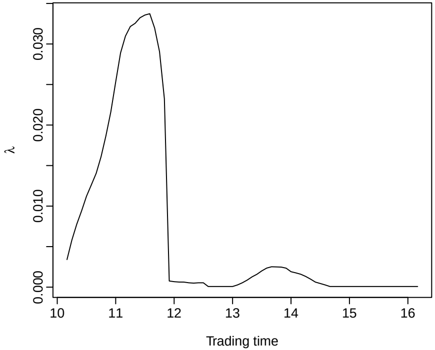
Point forecast accuracy

Averaging over 50 days in forecasting period & 73 different intraday updating periods

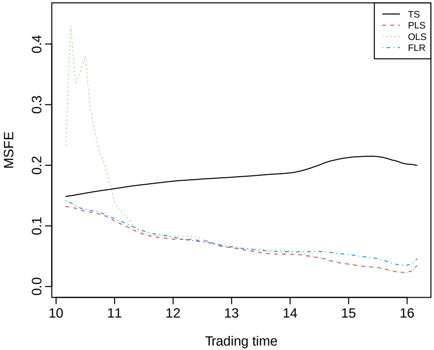
		$\text{ECP}_{\text{pointwise},1-\alpha}$		$\text{ECP}_{\text{uniform},1-\alpha}$		$\overline{S}_\alpha$	
Method	MSFE	$\alpha = 0.2$	$\alpha = 0.05$	$\alpha = 0.2$	$\alpha = 0.05$	$\alpha = 0.2$	$\alpha = 0.05$
TS	0.1474	0.89	0.98	0.88	0.96	1.43	2.08

Updating point forecasts

PLS method has best point forecast accuracy



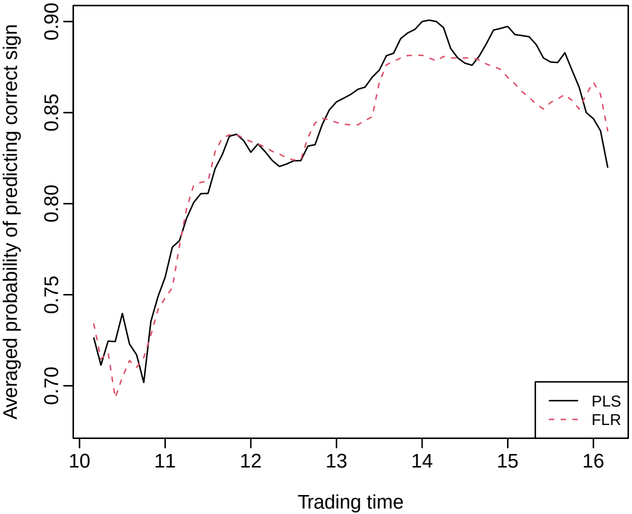
(a) *Selected optimal λ values*



(b) *Out-of-sample MSFE*

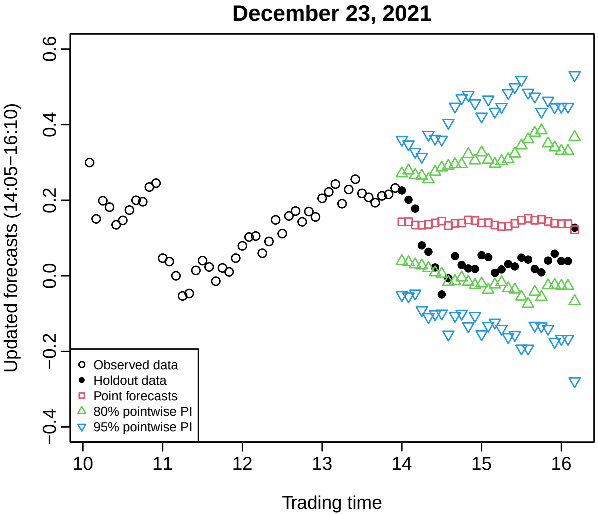
Predicting signs

CIDRs are hard to predict, but signs of future values are easier

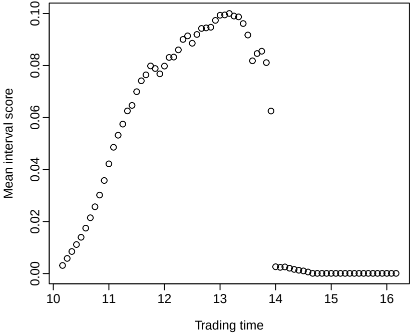


Updating prediction intervals

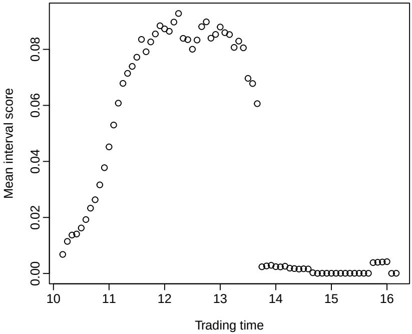
As we observe new data from beginning to 14:00, apply PLS with optimal λ to update point & interval forecast



Estimated optimal λ

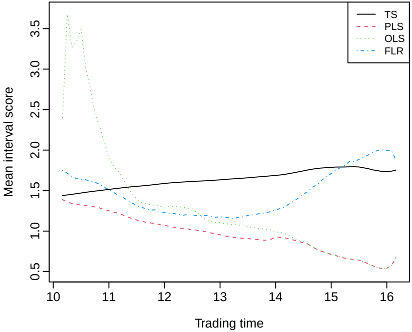


(c) *80% nominal coverage*

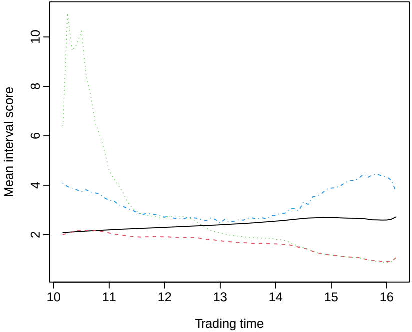


(d) *95% nominal coverage*

Mean interval score



(e) *80% nominal coverage*



(f) *95% nominal coverage*

Updating forecast accuracy

		$\text{ECP}_{\text{pointwise},1-\alpha}$		$\overline{S}_\alpha$	
Method	MSFE	$\alpha = 0.2$	$\alpha = 0.05$	$\alpha = 0.2$	$\alpha = 0.05$
TS	0.1838	0.8817	0.9688	1.6429	2.4277
PLS	0.0672	0.7275	0.8851	0.9564	1.6440
FLR	0.0735	0.5001	0.7266	1.4661	3.2729

Future research

1. Consider other sampling frequencies

Future research

1. Consider other sampling frequencies
2. Outliers can affect estimation of covariance, one can use robust FPCA

Future research

1. Consider other sampling frequencies
2. Outliers can affect estimation of covariance, one can use robust FPCA
3. With validation samples, PLS parameters can be adaptively chosen without recomputing

Thank you

Paper: https://onlinelibrary.wiley.com/doi/full/10.1002/for.3000

RG: https://www.researchgate.net/profile/Han-Lin-Shang