A novel approach for forecasting high-dimensional conditional covariance matrices using general dynamic factor models

Carlos Trucíos
Department of Statistics, Universidade Estadual de Campinas (UNICAMP)

65th ISI World Statistics Congress
The Hague, 2025.

FAPESP

UNICAMP

Table of contents

Introduction

Introduction

- The conditional covariance matrix (Σ_t) is an input parameter for several economic and financial applications.

Introduction

- The conditional covariance matrix (Σ_t) is an input parameter for several economic and financial applications.
- Σ_t evolves over time.

Introduction

- The conditional covariance matrix (Σ_t) is an input parameter for several economic and financial applications.
- Σ_t evolves over time.
- In the multivariate setting, MGARCH models are widely used in the literature.

Introduction

- The conditional covariance matrix (Σ_t) is an input parameter for several economic and financial applications.
- Σ_t evolves over time.
- In the multivariate setting, MGARCH models are widely used in the literature.
- Although MGARCH models are useful for predicting conditional covariance matrices in small to moderate dimensions, they are severely affected by the "curse of dimensionality."

Introduction

- The conditional covariance matrix (Σ_t) is an input parameter for several economic and financial applications.
- Σ_t evolves over time.
- In the multivariate setting, MGARCH models are widely used in the literature.
- Although MGARCH models are useful for predicting conditional covariance matrices in small to moderate dimensions, they are severely affected by the "curse of dimensionality."
- To address these issues, alternative procedures for estimating the conditional covariance matrix in high-dimensional settings have been proposed, such as:

Introduction

- The conditional covariance matrix (Σ_t) is an input parameter for several economic and financial applications.
- Σ_t evolves over time.
- In the multivariate setting, MGARCH models are widely used in the literature.
- Although MGARCH models are useful for predicting conditional covariance matrices in small to moderate dimensions, they are severely affected by the "curse of dimensionality."
- To address these issues, alternative procedures for estimating the conditional covariance matrix in high-dimensional settings have been proposed, such as:
    - Composite likelihood

Introduction

- The conditional covariance matrix (Σ_t) is an input parameter for several economic and financial applications.
- Σ_t evolves over time.
- In the multivariate setting, MGARCH models are widely used in the literature.
- Although MGARCH models are useful for predicting conditional covariance matrices in small to moderate dimensions, they are severely affected by the "curse of dimensionality."
- To address these issues, alternative procedures for estimating the conditional covariance matrix in high-dimensional settings have been proposed, such as:
    - Composite likelihood
    - Shrinkage

Introduction

- The conditional covariance matrix (Σ_t) is an input parameter for several economic and financial applications.
- Σ_t evolves over time.
- In the multivariate setting, MGARCH models are widely used in the literature.
- Although MGARCH models are useful for predicting conditional covariance matrices in small to moderate dimensions, they are severely affected by the "curse of dimensionality."
- To address these issues, alternative procedures for estimating the conditional covariance matrix in high-dimensional settings have been proposed, such as:
  - Composite likelihood
  - Shrinkage
  - Non-parametric approaches

Introduction

- The conditional covariance matrix (Σ_t) is an input parameter for several economic and financial applications.
- Σ_t evolves over time.
- In the multivariate setting, MGARCH models are widely used in the literature.
- Although MGARCH models are useful for predicting conditional covariance matrices in small to moderate dimensions, they are severely affected by the "curse of dimensionality."
- To address these issues, alternative procedures for estimating the conditional covariance matrix in high-dimensional settings have been proposed, such as:
  - Composite likelihood
  - Shrinkage
  - Non-parametric approaches
  - Dimension reduction techniques

Introduction

Dimension reduction techniques to forecast the conditional covariance matrix:

- Principal components analysis (PCA),
- Independent component analysis (ICA),
- Conditionally uncorrelated components (CUC),
- Dynamic orthogonal components (DOC),
- Principal volatility components (PVC), etc.

However, most dimension reduction techniques are based on a static approach which is not optimal in a time series context (Hallin et al., 2018).

The general dynamic factor model

GDFM

Flexivel, general and based on representation results.

GDFM

Flexivel, general and based on representation results.

- Sucessfully applied in several fields.
- Proposed in the early 2000s.
- N and $T \to \infty$

General dynamic factor model

Let $\mathbf{X}_t = (X_{1t}\ X_{2t} \ldots\ X_{nt})'$, $t = 1, ...$ denotea double-indexed stationary stochastic process with zero mean and finite second order moments. The GDFM is based on the decomposition

$$X_{it} = \chi_{it} + \xi_{it} \quad (1)$$

$$\chi_{it} = b_{i1}(L)u_{1t} + \ldots + b_{iq}(L)u_{qt}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \quad (2)$$

where L stands for a lag operator and the unobservable $u_{i,t}$, $\chi_{i,t}$ and $\xi_{i,t}$, stand for the common shocks, common components and idiosyncratic components, respectively.

General dynamic factor model

Under the assumption that the space spanned by the common components is finite-dimensional, the decomposition (1) takes the static form

$$X_{it} = \underbrace{\lambda_{i1} F_{1t} + \ldots + \lambda_{ir} F_{rt}}_{\chi_{it}} + \xi_{it}, \quad r \geq q \tag{3}$$

However, this assumption rules out simple factor-loading patterns (Forni and Lippi, 2011; Forni et al., 2015, 2017) such as

$$X_{i,t} = \overbrace{\underbrace{a_i(1-\alpha_i L)^{-1} u_t}_{a_i(u_t+\alpha_i u_{t-1}+\alpha_i^2 u_{t-2}+\alpha_i^3 u_{t-3}+\ldots)}}^{\chi_{it}} + \xi_{it}. \tag{4}$$

General dynamic factor model

- Forni et al. (2000) proposed a procedure that does not assume that the space spanned by the common components is finite-dimensional. However, is based on a **two-sided filter**, which is not satisfactory for forecasting.

General dynamic factor model

- Forni et al. (2000) proposed a procedure that does not assume that the space spanned by the common components is finite-dimensional. However, is based on a **two-sided filter**, which is not satisfactory for forecasting.
- Forni et al. (2005) proposed a procedure which allows for one-sided filter estimation. However, assume **finite-dimensional** factor space.

General dynamic factor model

- Forni et al. (2000) proposed a procedure that does not assume that the space spanned by the common components is finite-dimensional. However, is based on a **two-sided filter**, which is not satisfactory for forecasting.
- Forni et al. (2005) proposed a procedure which allows for one-sided filter estimation. However, assume **finite-dimensional** factor space.
- Forni et al. (2015, 2017) proposed a procedure which allows **one-sided** filter estimation and **infinite-dimensional** factor space.

General dynamic factor model

- Forni et al. (2000) proposed a procedure that does not assume that the space spanned by the common components is finite-dimensional. However, is based on a **two-sided filter**, which is not satisfactory for forecasting.
- Forni et al. (2005) proposed a procedure which allows for one-sided filter estimation. However, assume **finite-dimensional** factor space.
- Forni et al. (2015, 2017) proposed a procedure which allows **one-sided** filter estimation and **infinite-dimensional** factor space.
- Trucíos et al. (2023) and Hallin and Trucíos (2023) extend the procedure to allow for the estimation of the **conditional covariance matrix**.

Conditional covariance matrix

Forni et al. (2015, 2017) show that, $\chi_t = (\chi_{1t}\ \chi_{2t} \dots \chi_{nt})'$ admits a block-structure autoregressive representation

$$\mathbf{A}(L)\chi_t = \mathbf{R}u_t. \tag{5}$$

where

$$\mathbf{A}(L) = \begin{bmatrix} \mathbf{A}^1(L) & 0 & \dots & 0 \\ 0 & \mathbf{A}^2(L) & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \mathbf{A}^m(L) \end{bmatrix} \text{ and } \mathbf{R} = \begin{bmatrix} \mathbf{R}^1 \\ \mathbf{R}^2 \\ \vdots \\ \mathbf{R}^m \end{bmatrix}$$

with $A^k(L)$ being a $(q+1) \times (q+1)$ polynomial matrix with finite degree and R^k a $(q+1) \times q$ matrix, $n = m(q+1)$.

Conditional covariance matrix

Under assumptions in Barigozzi and Hallin (2020) and additionally assuming that u_t and ξ_t are conditionally uncorrelated for any lead and lag, Trucíos et al. (2023) show that the conditional variance-covariance matrix of $\mathbf{X}_t$ is given by

$$V(\mathbf{X}_t|\mathcal{F}_{t-1}) = \mathbf{R}\; V(\mathbf{u}_t|\mathcal{F}_{t-1})\mathbf{R}' + V(\xi_t|\mathcal{F}_{t-1}). \qquad (6)$$

Conditional covariance matrix

Under assumptions in Barigozzi and Hallin (2020) and additionally assuming that u_t and ξ_t are conditionally uncorrelated for any lead and lag, Trucíos et al. (2023) show that the conditional variance-covariance matrix of $\mathbf{X}_t$ is given by

$$V(\mathbf{X}_t|\mathcal{F}_{t-1}) = \mathbf{R}\, V(\mathbf{u}_t|\mathcal{F}_{t-1})\mathbf{R}' + V(\xi_t|\mathcal{F}_{t-1}). \tag{6}$$

- $u_t \sim$ MGARCH.

Conditional covariance matrix

Under assumptions in Barigozzi and Hallin (2020) and additionally assuming that u_t and ξ_t are conditionally uncorrelated for any lead and lag, Trucíos et al. (2023) show that the conditional variance-covariance matrix of $\mathbf{X}_t$ is given by

$$V(\mathbf{X}_t|\mathcal{F}_{t-1}) = \mathbf{R}\, V(\mathbf{u}_t|\mathcal{F}_{t-1})\mathbf{R}' + V(\xi_t|\mathcal{F}_{t-1}). \tag{6}$$

- $u_t \sim$ MGARCH.
- The conditional covariance matrix of the idiosyncratic factors can be modelled as a full or diagonal matrix, where each conditional variance is modelled independently by a GARCH-type model.

Estimation

- **Step 1.** Determine the number q of common shocks via an information criterion, for instance, using Hallin and Liška (2007).

Estimation

- **Step 1.** Determine the number q of common shocks via an information criterion, for instance, using Hallin and Liška (2007).
- **Step 2.** Randomly reorder the n observed series.

Estimation

- **Step 1.** Determine the number q of common shocks via an information criterion, for instance, using Hallin and Liška (2007).
- **Step 2.** Randomly reorder the n observed series.
- **Step 3.** Estimate the spectral density matrix of $\mathbf{X}$ by

$$\widehat{\boldsymbol{\Sigma}}^X_{nT}(\theta) = \frac{1}{2\pi} \sum_{k=-M_T}^{M_T} e^{-ik\theta} \underbrace{K\left(\frac{k}{B_T}\right)}_{1-\frac{|k|}{[|\sqrt{T}|]+1}} \widehat{\boldsymbol{\Gamma}}^X_k \quad \theta \in [0, 2\pi] \tag{7}$$

where $K(\cdot)$ is a kernel function, M_T a truncation parameter, B_T a bandwidth, and $\widehat{\boldsymbol{\Gamma}}^X_k$ the sample lag-k cross-covariance matrix.

Estimation

- **Step 4.** Estimate the spectral density matrix of the **common components** by

$$\widehat{\mathbf{\Sigma}}^{\chi}_{nT}(\theta) := \widehat{\mathbf{P}}^{X}_{nT}(\theta)\widehat{\mathbf{\Lambda}}^{X}_{nT}(\theta)\widehat{\mathbf{P}}^{X*}_{nT}(\theta) \quad \theta \in [0, 2\pi]$$

where $\widehat{\mathbf{\Lambda}}^{X}_{nT}(\theta)$ is a $q \times q$ diagonal matrix with diagonal elements the q largest eigenvalues of $\widehat{\mathbf{\Sigma}}^{X}_{nT}(\theta)$ and $\widehat{\mathbf{P}}^{X}_{nT}(\theta)$ (with complex conjugate $\widehat{\mathbf{P}}^{X*}_{nT}$) is the $n \times q$ matrix with the associated eigenvectors.

Estimation

- **Step 5.** By inverse Fourier transform of $\widehat{\mathbf{\Sigma}}^{\chi}_{n^* T}(\theta)$, estimate the autocovariance matrices $\widehat{\mathbf{\Gamma}}^{\chi}_{k}$ of the m sub-vectors

$$\boldsymbol{\chi}^k_t = (\chi_{(k-1)(q+1)+1,t} \cdots \chi_{k(q+1),t})', \quad k = 1, ..., m$$

of dimension $(q+1)$. Based on these, compute, after order identification, the Yule-Walker estimators $\hat{\mathbf{A}}^k(L)$ of the m VAR filters $\mathbf{A}^k(L)$ and stack them into a block-diagonal matrix $\hat{\mathbf{A}}_{n^*}(L)$. Compute $\hat{\mathbf{A}}_{n^*}(L)\mathbf{X}_{n^* t} = \widehat{\mathbf{Y}}_{n^* t} = \widehat{\mathbf{R}}_{n^*}\widehat{\mathbf{u}}_t + \widehat{\varepsilon}_t$

Estimation

- **Step 5.** By inverse Fourier transform of $\widehat{\boldsymbol{\Sigma}}^{\chi}_{n^* T}(\theta)$, estimate the autocovariance matrices $\widehat{\boldsymbol{\Gamma}}^{\chi}_{k}$ of the m sub-vectors

$$\boldsymbol{\chi}^k_t = (\chi_{(k-1)(q+1)+1,t} \cdots \chi_{k(q+1),t})', \quad k = 1, \ldots, m$$

  of dimension $(q+1)$. Based on these, compute, after order identification, the Yule-Walker estimators $\hat{\mathbf{A}}^k(L)$ of the m VAR filters $\mathbf{A}^k(L)$ and stack them into a block-diagonal matrix $\hat{\mathbf{A}}_{n^*}(L)$. Compute $\hat{\mathbf{A}}_{n^*}(L)\mathbf{X}_{n^* t} = \widehat{\mathbf{Y}}_{n^* t} = \widehat{\mathbf{R}}_{n^*}\widehat{\mathbf{u}}_t + \widehat{\varepsilon}_t$
- **Step 6.** Based on the first q **standard principal components** of $\widehat{\mathbf{Y}}_{n^* t}$, obtain estimates $\widehat{\mathbf{R}_{n^*}\mathbf{u}_t}$ of $\mathbf{R}_{n^*}\mathbf{u}_t$ and, via a Cholesky identification constraint the estimates $\widehat{\mathbf{R}}_{n^*}$ and $\widehat{\mathbf{u}}_t$ of $\mathbf{R}_{n^*}$ and $\mathbf{u}_t$; then, an estimate of the impulse-response function is $\widehat{\mathbf{B}}_{n^*}(L) := [\hat{\mathbf{A}}_{n^*}(L)]^{-1}\widehat{\mathbf{R}}_{n^*}$.

- **Step 7.** Repeat Steps 2 through 7 M times: the final estimates (denoted as $\widehat{\mathbf{R}}_n$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_n$) are obtained by averaging the estimates $\widehat{\mathbf{R}}_{n^*}$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_{n^*}$ associated with each iteration. Let $\widehat{\chi}_{nt} := \widehat{\mathbf{B}}_n(L)\widehat{\mathbf{u}_t}$ and $\hat{\boldsymbol{\xi}}_{nt} := \mathbf{X}_{nt} - \widehat{\chi}_{nt}$.

Estimation

- **Step 7.** Repeat Steps 2 through 7 M times: the final estimates (denoted as $\widehat{\mathbf{R}}_n$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_n$) are obtained by averaging the estimates $\widehat{\mathbf{R}}_{n^*}$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_{n^*}$ associated with each iteration. Let $\widehat{\chi}_{nt} := \widehat{\mathbf{B}}_n(L)\widehat{\mathbf{u}}_t$ and $\hat{\xi}_{nt} := \mathbf{X}_{nt} - \widehat{\chi}_{nt}$.
- **Step 8.** The estimator of $\mathrm{V}(\mathbf{X}_t|\mathcal{F}_{t-1})$ is given by

$$\widehat{\mathrm{V}}(\mathbf{X}_t|\mathcal{F}_{t-1}) := \widehat{\mathbf{R}}\widehat{\mathrm{V}}(\hat{\mathbf{u}}_t|\mathcal{F}_{t-1})\widehat{\mathbf{R}}' + \widehat{\mathrm{V}}(\hat{\xi}_t|\mathcal{F}_{t-1}). \tag{8}$$

- **Step 7.** Repeat Steps 2 through 7 M times: the final estimates (denoted as $\widehat{\mathbf{R}}_n$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_n$) are obtained by averaging the estimates $\widehat{\mathbf{R}}_{n^*}$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_{n^*}$ associated with each iteration. Let $\widehat{\chi}_{nt} := \widehat{\mathbf{B}}_n(L)\widehat{\mathbf{u}_t}$ and $\hat{\boldsymbol{\xi}}_{nt} := \mathbf{X}_{nt} - \widehat{\chi}_{nt}$.
- **Step 8.** The estimator of $\mathrm{V}(\mathbf{X}_t|\mathcal{F}_{t-1})$ is given by

$$\widehat{\mathrm{V}}(\mathbf{X}_t|\mathcal{F}_{t-1}) := \widehat{\mathbf{R}}\widehat{\mathrm{V}}(\hat{\mathbf{u}}_t|\mathcal{F}_{t-1})\widehat{\mathbf{R}}' + \widehat{\mathrm{V}}(\hat{\xi}_t|\mathcal{F}_{t-1}). \quad (8)$$

Estimation

- **Step 7.** Repeat Steps 2 through 7 M times: the final estimates (denoted as $\widehat{\mathbf{R}}_n$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_n$) are obtained by averaging the estimates $\widehat{\mathbf{R}}_{n^*}$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_{n^*}$ associated with each iteration. Let $\widehat{\chi}_{nt} := \widehat{\mathbf{B}}_n(L)\widehat{\mathbf{u}_t}$ and $\hat{\xi}_{nt} := \mathbf{X}_{nt} - \widehat{\chi}_{nt}$.
- **Step 8.** The estimator of $\mathrm{V}(\mathbf{X}_t|\mathcal{F}_{t-1})$ is given by

$$\widehat{\mathrm{V}}(\mathbf{X}_t|\mathcal{F}_{t-1}) := \widehat{\mathbf{R}}\widehat{\mathrm{V}}(\hat{\mathbf{u}}_t|\mathcal{F}_{t-1})\widehat{\mathbf{R}}' + \widehat{\mathrm{V}}(\hat{\xi}_t|\mathcal{F}_{t-1}). \tag{8}$$

New idea:

Can we do better? What happends if we replace the PCA in Step 6 with a dimension-reduction technique designed specifically to extract volatility components?

Principal Volatility Components

Principal Volatility Components (PVC)

Similar to PCA.

Principal Volatility Components (PVC)

Similar to PCA.

- Hu, Y. P., & Tsay, R. S. **(2014)**. Principal volatility component analysis. Journal of Business & Economic Statistics, 32(2), 153-164.
- Li, W., Gao, J., Li, K., & Yao, Q. **(2016)**. Modeling multivariate volatilities via latent common factors. Journal of Business & Economic Statistics, 34(4), 564-573.

Principal Volatility Components (PVC)

PVC is based on a similar idea that PCA.

- PCA decomposes a N-dimensional vector into N contemporaneous uncorrelated components according with the amount of variability explained by the components.
- PCA uses the sample covariance matrix.
- Hu and Tsay (2014) and Li et al. (2016) proposed PVC: A generalization of PCA that takes into account the dynamic dependence between the volatility processes.
- In PVC the covariance matrix is replaced by a matrix that summarizes the dynamic dependence of volatilities.
- With PVC we identify common volatility components and also components with no conditional heteroscedasticity.

Principal Volatility Components (PVC)

Let the *Cumulative Generalized Kurtosis Matrix* defined by

$$\Gamma_\infty = \sum_{\ell=1}^{\infty} \sum_{i=1}^{k} \sum_{j=1}^{m} E\left[(y_t' y_t - E(y_t' y_t))\,(x_{ij,t-\ell} - E(x_{ij,t}))\right], \tag{9}$$

where $x_{ij,t-k}$ is a function of the cross product $y_{i,t-k}$ and $y_{j,t-k}$

Principal Volatility Components (PVC)

Let the *Cumulative Generalized Kurtosis Matrix* defined by

$$\Gamma_\infty = \sum_{\ell=1}^{\infty} \sum_{i=1}^{k} \sum_{j=1}^{m} E\left[\left(y_t' y_t - E(y_t' y_t) \right) \left(x_{ij,t-\ell} - E(x_{ij,t}) \right) \right], \tag{9}$$

where $x_{ij,t-k}$ is a function of the cross product $y_{i,t-k}$ and $y_{j,t-k}$

Additionally,

$$\Gamma_\infty M = M\Lambda, \qquad \text{where}$$

- $M = [m_1, ..., m_k]$ is the matrix of normalized eigenvectors and
- Λ is the diagonal matrix of ordered eigenvalues,

Principal Volatility Components (PVC)

Hu and Tsay (2014) proves that, under mild conditions, there exist r linear combination of y_t that have ARCH effects and $k - r$ linear combination of y_t that have no ARCH effects if and only if $rank\left(\Gamma_\infty\right) = r$.

Principal Volatility Components (PVC)

Hu and Tsay (2014) proves that, under mild conditions, there exist r linear combination of y_t that have ARCH effects and $k - r$ linear combination of y_t that have no ARCH effects if and only if $rank\left(\Gamma_{\infty}\right) = r$.

The v-th PVC is defined as

$$z_{vt} = m_v' y_t$$

Principal Volatility Components (PVC)

Hu and Tsay (2014) proves that, under mild conditions, there exist r linear combination of y_t that have ARCH effects and $k - r$ linear combination of y_t that have no ARCH effects if and only if $rank\left(\Gamma_\infty\right) = r$.

The v-th PVC is defined as

$$z_{vt} = m_v' y_t$$

- z_{vt}, z_{ut} are (contemporaneously) uncorrelated if $\lambda_v^2 \neq \lambda_u^2$.
- z_{vt} may still be correlated with lagged values of z_{ut}.

Principal Volatility Components (PVC)

Li et al. (2016) proposed an alternative PVC characterized only by the second moment. In this approach, the matrix Γ_∞ is replaced by

$$G = \sum_{k=1}^{g} \sum_{t=1}^{T} \omega(y_t) E^2 \left[(y_t y_t' - \Sigma) \, I(\|y_{t-k}\| \leq \|y_t\|) \right], \tag{10}$$

where $\omega(\cdot)$ is a weight function and $\| \cdot \|$ is the L_1 norm.

Principal Volatility Components (PVC)

Li et al. (2016) proposed an alternative PVC characterized only by the second moment. In this approach, the matrix Γ_∞ is replaced by

$$G = \sum_{k=1}^{g} \sum_{t=1}^{T} \omega(y_t) E^2 \left[(y_t y_t' - \Sigma) \, I(\|y_{t-k}\| \leq \|y_t\|) \right], \tag{10}$$

where $\omega(\cdot)$ is a weight function and $\| \cdot \|$ is the L_1 norm.

The alternative procedure has the same properties of the proposal of Hu and Tsay (2014) but only requires finite second-order moments.

Principal Volatility Components

In practice, the matrix G is estimated in a natural way by

$$\hat{G} = \sum_{k=1}^{g} \sum_{\tau=1}^{T} \omega(y_\tau) \left[\frac{1}{T-k} \sum_{t=k+1}^{T} \left[\left(y_t y_t' - \hat{\Sigma} \right) I(\|y_{t-k}\| \leq \|y_\tau\|) \right] \right]^2 .$$

Principal Volatility Components

In practice, the matrix G is estimated in a natural way by

$$\hat{G} = \sum_{k=1}^{g} \sum_{\tau=1}^{T} \omega(y_\tau) \left[\frac{1}{T-k} \sum_{t=k+1}^{T} \left[\left(y_t y_t' - \hat{\Sigma} \right) I(\|y_{t-k}\| \leq \|y_\tau\|) \right] \right]^2 .$$

Thus, we use PVC instead PCA in Step 6 and the estimator of $\mathrm{V}(\mathbf{X}_t|\mathcal{F}_{t-1})$ is given by

$$\widehat{\mathrm{V}}(\mathbf{X}_t|\mathcal{F}_{t-1}) := \underbrace{\widehat{\mathbf{R}}}_{GPVC} \underbrace{\widehat{\mathrm{V}}(\hat{\mathbf{u}}_t|\mathcal{F}_{t-1})}_{MGARCH} \underbrace{\widehat{\mathbf{R}}'}_{GPVC} + \underbrace{\widehat{\mathrm{V}}(\hat{\xi}_t|\mathcal{F}_{t-1})}_{\text{Has constant conditional variance}} \tag{11}$$

Empirical Application

Empirical Applications

- **Minimum Variance Portfolios**

- **Minimum Variance Portfolios**
- Universe: 534 stocks traded on the NYSE.

- **Minimum Variance Portfolios**
- Universe: 534 stocks traded on the NYSE.
- Sample period: January 2, 2010, through December 31, 2024 ($T = 3774$ trading days).

Empirical Applications

- **Minimum Variance Portfolios**
- Universe: 534 stocks traded on the NYSE.
- Sample period: January 2, 2010, through December 31, 2024 ($T = 3774$ trading days).
- **Rolling Window Scheme:** Estimation window of 1,250 days.

Empirical Applications

- **Minimum Variance Portfolios**
- Universe: 534 stocks traded on the NYSE.
- Sample period: January 2, 2010, through December 31, 2024 ($T = 3774$ trading days).
- **Rolling Window Scheme:** Estimation window of 1,250 days.
- Out-of-sample period: 2,524 days.

Empirical Applications

- **Minimum Variance Portfolios**
- Universe: 534 stocks traded on the NYSE.
- Sample period: January 2, 2010, through December 31, 2024 ($T = 3774$ trading days).
- **Rolling Window Scheme:** Estimation window of 1,250 days.
- Out-of-sample period: 2,524 days.
- Daily portfolio rebalancing.

Empirical Application

At time $t = 1250, \ldots, 3774$ the one-step ahead conditional covariance matrix is estimated and used to obtain the optimal portfolio allocation weights, that is, minimise

$$\boldsymbol{\omega}'\widehat{\Sigma}_{T+1|T}\boldsymbol{\omega},$$

subject to $\omega_i \geq 0$ and $\sum_{i=1}^{n} \omega_i = 1$.

Empirical Application

At time $t = 1250, \ldots, 3774$ the one-step ahead conditional covariance matrix is estimated and used to obtain the optimal portfolio allocation weights, that is, minimise

$$\boldsymbol{\omega}' \widehat{\Sigma}_{T+1|T} \boldsymbol{\omega},$$

subject to $\omega_i \geq 0$ and $\sum_{i=1}^{n} \omega_i = 1$.

Then, the resulting (out-of-sample) portfolio return is given by

$$R_{p,T+1} := \sum_{i=1}^{n} \widehat{\omega}_{i;T+1|T} \times r_{i,T+1}$$

Empirical Application

- **Annualized average portfolio (AV)**: is given by $\sqrt{12} \times \bar{R}_p$, where $\bar{R}_p$ is the sample mean of the realized out-of-sample portfolio returns. The larger the AV, the better the portfolio performance.

Empirical Application

- **Annualized average portfolio (AV)**: is given by $\sqrt{12} \times \bar{R}_p$, where $\bar{R}_p$ is the sample mean of the realized out-of-sample portfolio returns. The larger the AV, the better the portfolio performance.
- **Annualized standard deviation (SD)**: is given by $\sqrt{12} \times \hat{\sigma}_p$, where $\hat{\sigma}_p$ is the sample standard deviation of the realized out-of-sample portfolio returns. The smaller the SD, the less risky the portfolio and, consequently, the better the portfolio performance.

Empirical Application

- **Annualized average portfolio (AV)**: is given by $\sqrt{12} \times \bar{R}_p$, where $\bar{R}_p$ is the sample mean of the realized out-of-sample portfolio returns. The larger the AV, the better the portfolio performance.
- **Annualized standard deviation (SD)**: is given by $\sqrt{12} \times \hat{\sigma}_p$, where $\hat{\sigma}_p$ is the sample standard deviation of the realized out-of-sample portfolio returns. The smaller the SD, the less risky the portfolio and, consequently, the better the portfolio performance.
- **Annualized Sharpe ratio (SR)**: is given by $\sqrt{12} \times SR$, where SR is the Sharpe ratio, a risk-adjusted performance measure defined by $SR = \bar{R}_p - \bar{R}_f \big/ \hat{\sigma}_{p-f}$, where $\bar{R}_f$ is the average risk-free rate. The higher the annualized SR, the better the portfolio performance.

Empirical Application

- **Annualized average portfolio (AV)**: is given by $\sqrt{12} \times \bar{R}_p$, where $\bar{R}_p$ is the sample mean of the realized out-of-sample portfolio returns. The larger the AV, the better the portfolio performance.
- **Annualized standard deviation (SD)**: is given by $\sqrt{12} \times \hat{\sigma}_p$, where $\hat{\sigma}_p$ is the sample standard deviation of the realized out-of-sample portfolio returns. The smaller the SD, the less risky the portfolio and, consequently, the better the portfolio performance.
- **Annualized Sharpe ratio (SR)**: is given by $\sqrt{12} \times SR$, where SR is the Sharpe ratio, a risk-adjusted performance measure defined by $SR = \bar{R}_p - \bar{R}_f \big/ \hat{\sigma}_{p-f}$, where $\bar{R}_f$ is the average risk-free rate. The higher the annualized SR, the better the portfolio performance.
- **Annualized adjusted Sharpe ratio (ASR)**: is given by $\sqrt{12} \times ASR$, where ASR is the adjusted Sharpe ratio which is defined by $ASR = SR\Big[1 + (\frac{\mu_3}{6}) SR - (\frac{\mu_4 - 3}{24}) SR^2\Big]$, where μ_3 and μ_4 stand for the skewness and kurtosis of the out-of-sample portfolio returns. The higher the ASR, the better the portfolio performance.

Empirical Application

- **Annualized average portfolio (AV)**: is given by $\sqrt{12} \times \bar{R}_p$, where $\bar{R}_p$ is the sample mean of the realized out-of-sample portfolio returns. The larger the AV, the better the portfolio performance.
- **Annualized standard deviation (SD)**: is given by $\sqrt{12} \times \hat{\sigma}_p$, where $\hat{\sigma}_p$ is the sample standard deviation of the realized out-of-sample portfolio returns. The smaller the SD, the less risky the portfolio and, consequently, the better the portfolio performance.
- **Annualized Sharpe ratio (SR)**: is given by $\sqrt{12} \times SR$, where SR is the Sharpe ratio, a risk-adjusted performance measure defined by $SR = \bar{R}_p - \bar{R}_f \big/ \hat{\sigma}_{p-f}$, where $\bar{R}_f$ is the average risk-free rate. The higher the annualized SR, the better the portfolio performance.
- **Annualized adjusted Sharpe ratio (ASR)**: is given by $\sqrt{12} \times ASR$, where ASR is the adjusted Sharpe ratio which is defined by $ASR = SR\left[1 + (\frac{\mu_3}{6})SR - (\dfrac{\mu_4 - 3}{24})SR^2\right]$, where μ_3 and μ_4 stand for the skewness and kurtosis of the out-of-sample portfolio returns. The higher the ASR, the better the portfolio performance.
- **Annualized Sortino ratio (SO)**: is given by $\sqrt{12} \times SO$, where SO is the Sortino ratio which is defined by $SO = \bar{R}_p \big/ \sqrt{\text{semi-variance}}$. The higher the SO, the better the portfolio performance.

Empirical Application

Since the GDFM-CHF proposed by Trucíos et al. (2023) has proven to be quite competitive, we compare our new proposal only with GDFM-CHF (which applies PCA to the static representation)

Empirical Application

Since the GDFM-CHF proposed by Trucíos et al. (2023) has proven to be quite competitive, we compare our new proposal only with GDFM-CHF (which applies PCA to the static representation)

Table 1: Out-of-sample performance measures of the minimum variance portfolio with short-selling constraints: AV, SD, SR, ASR, and SO, stand for the average, standard deviation, Sharpe ratio, Adjusted Sharpe ratio, and Sortino ratio, respectively.

	AV	SD	SR	ASR	SO
GDFM-CHF	0.2720	2.6871	0.1012	0.1006	0.1513
GDFM-GPVC	**0.3391**	**2.6718**	**0.1269**	**0.1258**	**0.1895**

Conclusion

- Building on previous results based on the GDFM (Trucíos et al., 2023) and PVC (**?**), we propose a new procedure to forecast the conditional covariance matrix in large portfolios.

Conclusion

- Building on previous results based on the GDFM (Trucíos et al., 2023) and PVC (**?**), we propose a new procedure to forecast the conditional covariance matrix in large portfolios.
- The new procedure does not require estimating the conditional covariance matrix of the idiosyncratic components using a conditional heteroscedastic method, as they have constant conditional variance.

Conclusion

- Building on previous results based on the GDFM (Trucíos et al., 2023) and PVC (**?**), we propose a new procedure to forecast the conditional covariance matrix in large portfolios.
- The new procedure does not require estimating the conditional covariance matrix of the idiosyncratic components using a conditional heteroscedastic method, as they have constant conditional variance.
- The new procedure delivers encouraging results, being competitive with the GDFM-CHF.

Conclusion

- Building on previous results based on the GDFM (Trucíos et al., 2023) and PVC (**?**), we propose a new procedure to forecast the conditional covariance matrix in large portfolios.
- The new procedure does not require estimating the conditional covariance matrix of the idiosyncratic components using a conditional heteroscedastic method, as they have constant conditional variance.
- The new procedure delivers encouraging results, being competitive with the GDFM-CHF.
- Further empirical applications and theoretical results are in progress.

References

Barigozzi, M. and Hallin, M. (2020). Generalized dynamic factor models and volatilities: Consistency, rates, and prediction intervals. *Journal of Econometrics*.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *Review of Economics and Statistics*, 82(4):540–554.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005). The generalized dynamic factor model: one-sided estimation and forecasting. *Journal of the American Statistical Association*, 100(471):830–840.

Forni, M., Hallin, M., Lippi, M., and Zaffaroni, P. (2015). Dynamic factor models with infinite-dimensional factor spaces: One-sided representations. *Journal of Econometrics*, 185(2):359–371.

Forni, M., Hallin, M., Lippi, M., and Zaffaroni, P. (2017). Dynamic factor models with infinite-dimensional factor space: asymptotic analysis. *Journal of Econometrics*, 199(1):74–92.

Forni, M. and Lippi, M. (2011). The general dynamic factor model: One-sided representation results. *Journal of Econometrics*, 163(1):23–28.

Hallin, M., Hörmann, S., and Lippi, M. (2018). Optimal dimension reduction for high-dimensional and functional time series. *Statistical Inference for Stochastic Processes*, 21(2):385–398.

Hallin, M. and Liška, R. (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association*, 102(478):603–617.

Hallin, M. and Trucíos, C. (2023). Forecasting value-at-risk and expected shortfall in large portfolios: A general dynamic factor model approach. *Econometrics and Statistics*, 27:1–15.

References iii

Hu, Y.-P. and Tsay, R. S. (2014). Principal volatility component analysis. *Journal of Business & Economic Statistics*, 32(2):153–164.

Li, W., Gao, J., Li, K., and Yao, Q. (2016). Modeling multivariate volatilities via latent common factors. *Journal of Business & Economic Statistics*, 34(4):564–573.

Trucíos, C., Mazzeu, J. H. G., Hallin, M., Hotta, L. K., Valls Pereira, P. L., and Zevallos, M. (2023). Forecasting conditional covariance matrices in high-dimensional time series: a general dynamic factor approach. *Journal of Business and Economic Statistics*.