Advanced Time Series Econometrics

Week 3: High-dimensional Volatility Models

Yi He

January 17, 2023

Plan for Today

1. Introduction
2. The Zoo of Volatility Models
3. Estimating + Evaluating High-Dimensional Volatility Model

Introduction

Conditional Variance Matrix

Consider the random vectors (e.g., daily stock returns)

$$y_t = (y_{1,t}, \ldots, y_{N,t})' \in \mathbb{R}^N, \quad t = 1, \ldots, T$$

such that

$$\mathbb{E}[y_t \mid \mathcal{F}_{t-1}] = \mathbf{0}_N, \qquad \mathrm{Var}[y_t \mid \mathcal{F}_{t-1}] = H_t \in \mathbb{R}^{N \times N}$$

where

- the sigma-algebra $\mathcal{F}_{t-1}$ contains information up to time $t-1$
- H_t is the conditional (co)variance matrix measurable by $\mathcal{F}_{t-1}$
- Under stationarity: $\Sigma = \mathbb{E}H_t = \mathbb{E}[y_t y_t']$ is the unconditional variance matrix

Harvey (1989, JFE):

> ... *The evidence indicates that the conditional covariances do change through time...*

- Conditional variance matrix H_t of vector of asset returns y_t is relevant for:
    - Portfolio selection: mean-variance optimisation (diversification) requires reliable estimate of variance matrix over the holding period.
    - Hedging: trying to eliminate exposure to certain risk factors requires conditional betas, which are functions of H_t.
    - Risk management: Value at Risk of portfolio return $w'y_t$ depends on variance $w'H_tw$. If portfolio weights are time-varying, this requires H_t.
- Today we are interested in the high-dimensional regime with $N \asymp T$, although most -if not all- techniques still work for small N.
- High-dimensional GARCH models often suffer from curse of dimensionality; some proposals provide a partial solution.

Catch Up: Univariate GARCH

- Define the marginal variances

$$h_{i,t} \quad = \quad \mathrm{Var}_{t-1}(y_{i,t}), \quad i = 1, \ldots, N$$

where $\mathrm{Var}_{t-1}(\cdot)$ is conditional on $\mathcal{F}_{t-1}$.

- Standard GARCH(1,1) specification for each series:

$$\begin{aligned} y_{i,t} &= \mu_{i,t} + \varepsilon_{i,t} = 0 + h_{i,t}^{1/2} z_{i,t}, \\ h_{i,t} &= (1 - \alpha_i - \beta_i)\sigma_i^2 + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1}, \\ z_{i,t} &\sim IID(0,1). \end{aligned}$$

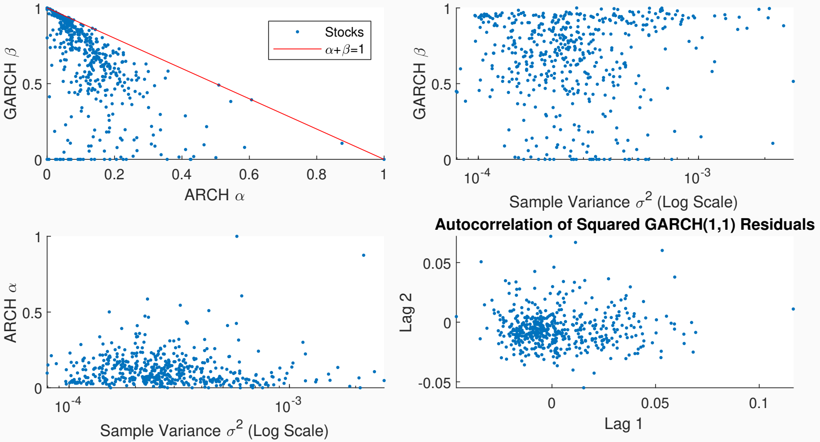
- Then decompose that

$$H_t = D_t Q_t D_t, \qquad D_t = \mathrm{diag}(h_{1,t}^{1/2}, \ldots, h_{N,t}^{1/2}),$$

with

$$Q_t = \mathrm{Var}_{t-1}(z_t), \quad z_t = (z_{1,t}, \ldots, z_{N,t})'.$$

Example: 500 US Stocks from 2015–2019



$N = 500$ US Stocks with: (1) largest market value at the beginning of 2015 (2) a complete daily return history over 2015–2019 ($T = 1258$)

High-Dimensional Volatility Models

- Our focus today is to model and estimate Q_t
- The marginal GARCH parameters σ_i^2, α_i, β_i can be fitted series-by-series.
- A trivial property we want to guarantee: H_t and Q_t are positive (semi)-definite
- For practitioners, one may want specification **flexible** enough to replicate properties such as
  - persistence in volatility and covariation;
  - time variation in correlation;
  - volatility spill-over effects;
  - asymmetries, leverage effects.
  - model is closed under linear transformations (taking portfolios)...

The Zoo of Volatility Models

DVEC Model: Bollerslev et al. (1988)

Analogous to univariate GARCH model, the (co)variances may depend on their own past, and on past squares $\varepsilon_{i,t-1}^2$ and cross-products $\varepsilon_{i,t-1}\varepsilon_{j,t-1}$:

$$H_t = (I - A - B) \circ \Sigma + A \circ \left(\varepsilon_{t-1}\varepsilon_{t-1}'\right) + B \circ H_{t-1}$$

where '$\circ$' denotes the entrywise product (Hadamard product), and A, B are symmetric matrices.

- Involves parameters at the *order* $N^2 + 2N^2$.
- Marginal GARCH models: $A_{ii} = \alpha_i$, $B_{ii} = \beta_i$, $\Sigma_{ii} = \sigma_i^2$
- Q_t is the correlation matrix corresponding to H_t
- If taking A and B as scalars (implying that $\alpha_{ii} = \alpha$, $\beta_{ii} = \beta$):

  $$H_t = (1 - \alpha - \beta)\Sigma + \alpha\varepsilon_{t-1}\varepsilon_{t-1}' + \beta H_{t-1}$$

  $\#$ parameters at the order $N^2 + 2$
- RiskMetrics ($\alpha + \beta = 1$): $H_t = (1 - \beta)\varepsilon_{t-1}\varepsilon_{t-1}' + \beta H_{t-1}$

BEKK Model: Engle and Kroner (1995)

BEKK(1,1,K) Model:

$$H_t = C^{*\prime}C^* + \sum_{k=1}^{K} A_k^{*\prime}\varepsilon_{t-1}\varepsilon_{t-1}^{\prime}A_k^* + \sum_{k=1}^{K} B_k^{*\prime}H_{t-1}B_k^*,$$

$$C^{*\prime}C^* = \Sigma - \sum_{k=1}^{K} A_k^{*\prime}\Sigma A_k^* - \sum_{k=1}^{K} B_k^{*\prime}\Sigma B_k^*$$

where A_k^* and B_k^* are non-singular $N \times N$ matrices, and C^* is $N \times N$ upper triangular matrix.

- Guarantee positive definiteness of H_t;
- Setting $A_k^* = \mathrm{diag}(a_k)$ and $B_k^* = \mathrm{diag}(b_k)$ diagonal gives the DVEC model with

$$A = \sum_{k=1}^{K} a_k a_k^{\prime}, \quad B = \sum_{k=1}^{K} b_k b_k^{\prime}.$$

  # parameters at the order $N^2 + KN$

Factor GARCH Model: Engle, Ng and Rothschild, 1990

$$y_t = \sum_{k=1}^{K} \lambda_k f_{kt} + e_t = \Lambda f_t + e_t,$$

where $\{f_{k,t} = w_k' y_t\}_{k=1}^{K}$ are independent GARCH processes, independent of $e_t \sim \text{IID}(0, \Omega^*)$.

- $\text{Var}_{t-1}(f_{k,t}) = h_{k,t} = (1 - \alpha_k - \beta_k)\sigma_k^2 + \alpha_k f_{k,t-1}^2 + \beta_k h_{k,t-1}$
- The conditional variance

$$H_t = \Sigma + \sum_{k=1}^{K} \lambda_k \lambda_k' (\alpha_k (f_{k,t-1}^2 - \sigma_k^2) + \beta_k (w_k' H_{t-1} w_k - \sigma_k^2))$$

- # parameters at the order $N^2 + KN$
- Special case of BEKK: Exercise 3

GO-GARCH Model

Consider the spectral decomposition of unconditional covariance matrix

$$\Sigma = PLP'$$

GO-GARCH (Alexander, 2002; van der Weide, 2002; Vrontis, Dellaportas and Pollitis, 2003):

$$y_t = \Lambda f_t, \qquad H_t = \Lambda V_t \Lambda'$$

where $\Lambda = PL^{1/2}U'$ is a non-singular matrix, U is some orthogonal matrix, and f_t is an N-dimensional vector of independent GARCH processes, with diagonal variance matrix V_t, and with unit unconditional variance.

- Closed under non-singular transformations;
- $\#$ parameters at the order $N^2 + N^2 + N$
- If $U = I$ then $N^2 + N$

CCC Model: Bollerslev (1990)

The constant conditional correlation (CCC) model

$$H_t = D_t Q D_t, \qquad D_t = \mathrm{diag}(h_{11,t}^{1/2}, \ldots, h_{NN,t}^{1/2}),$$

where $Q = \mathbb{E}[z_t z_t']$ is the covariance/correlation matrix of 'devolatilized' residuals z_t.

- Positive definite;
- # parameters of order $N^2 + N$;
- Not closed under linear transformations.

DCC Model: Engle (2002)

The dynamic conditional correlation (DCC) model :

$$H_t = D_t R_t D_t, \qquad D_t = \mathrm{diag}(h_{11,t}^{1/2}, \ldots, h_{NN,t}^{1/2}),$$

where R_t is a conditional correlation matrix, given by

$$\begin{aligned} R_t =& \mathrm{diag}(Q_t)^{-1/2} Q_t \,\mathrm{diag}(Q_t)^{-1/2} \\ Q_t =& (1-\alpha-\beta)\Sigma + \alpha z_{t-1} z_{t-1}' + \beta Q_{t-1} \end{aligned}$$

Here $\Sigma = \mathbb{E} Q_t$

- Positive definite;
- More flexible than CCC with 2 more parameters (α, β); not closed under linear transformations;
- Correction proposed by Aielli (2013, JBES): makes little difference in practice.

GAS Model: Creal, Koopmans and Lucas (2011)

The **generalized autoregressive score (GAS)** model builds on the DCC model, with

$$f_t = \begin{pmatrix} \mathrm{diag}(D_t^2) \\ \mathrm{vech}(Q_t) \end{pmatrix}$$

satisfying

$$f_t = \omega + A s_{t-1} + B f_{t-1}, \quad s_t = S_t \times \nabla_{f_t} \log p(y_t | f_t; \theta)$$

By taking p to be the multivariate $t(\nu)$ density, this gives an automatic model specification that downweighs extreme observations. In particular,

$$Q_t = (1-\beta)\Sigma + \alpha \widetilde{S}_{t-1} \nabla_{Q_{t-1}} \log p(y_{t-1} | f_{t-1}; \theta) + \beta Q_{t-1}$$

Asymmetry

- Equity returns display leverage effects: large negative shocks have stronger effect on volatility than large positive shocks.
- One can use asymmetric *univariate* GARCH models

$$h_{i,t} = \omega_i + (\alpha_i + \gamma_i \mathbb{1}[\varepsilon_{i,t-1} < 0])\varepsilon_{i,t-1}^2 + \beta_i h_t$$

  on the factors $\varepsilon_{i,t} = f_{i,t}$ or returns $\varepsilon_{i,t} = y_{i,t}$
- Can be included in a DCC model by

$$Q_t = (1 - \alpha - \beta)\Sigma + \alpha z_{t-1} z_{t-1}' + \beta Q_{t-1} + \gamma(v_{t-1} v_{t-1}' - \Sigma_-),$$

  where $v_t = \max(0, -z_t) = -\mathbb{1}[z_t < 0] \circ z_t$ and $\Sigma_- = \mathbb{E}[v_t v_t']$.

Estimating + Evaluating High-Dimensional Volatility Model

Estimation of multivariate GARCH models

- The log-(quasi-)likelihood for the model

$$y_t|\mathcal{F}_{t-1} \overset{\text{quasi}}{\sim} N(0, H_t(\theta))$$

  is given by $\ell(\theta) = \sum_{t=1}^{T} \ell_t(\theta)$, with

$$\ell_t(\theta) = -\frac{N}{2}\log 2\pi - \frac{1}{2}\log|H_t(\theta)| - \frac{1}{2}y_t' H_t(\theta)^{-1} y_t.$$

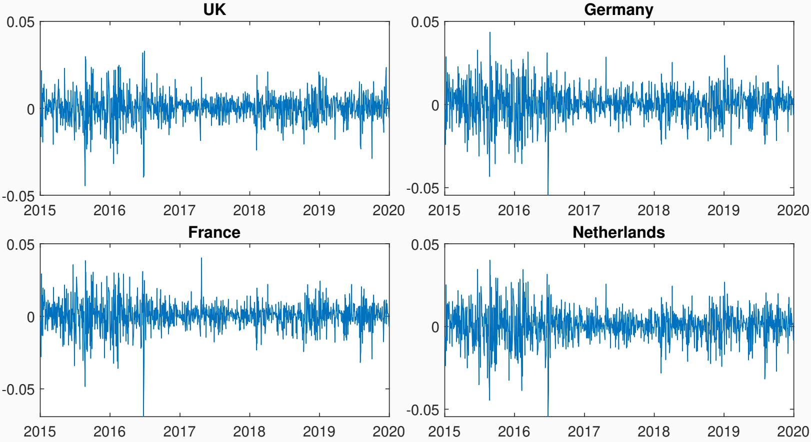
- For GAS model one may use the likelihood $t(\nu)$ and include ν as a parameter too
- Depending on N and the specification of $H_t(\theta)$, the dimension of the parameter space can be very high
- For (very) small $N/T < 100$, ML works fine. For specific models, multi-step approximations to ML have been proposed.
- For large $N \asymp T$, ML suffers from overfitting

Estimating Low-Dimensional GARCH Models

For small $N/T < 100$, say:

- For **GO-GARCH**, this is adapted as follows, using $y_t = \Lambda f_t$, $\Lambda = PL^{1/2}U$:
  1. Estimate P and L as eigenvectors and eigenvalues of sample variance matrix of y_t. Hence $\widehat{s}_t = \widehat{L}^{-1/2}\widehat{P}' y_t$;
  2. Maximize the log-likelihood for the model $f_t = U'\widehat{s}_t$, over orthogonal matrices U (if unknown) and univariate GARCH parameters for f_{it};
  3. U may also be estimated from the (cross-) autocorrelations in $\widehat{s}_t\widehat{s}_t'$: Boswijk and van der Weide (2011)
- For **DCC**, a three-step estimation procedure is:
  1. Estimate GARCH models for y_{it}, and compute $\widehat{z}_{it} := \widehat{h}_{i,t}^{-1/2} y_t$;
  2. Estimate $\Sigma = \mathbb{E}Q_t$ by $S = T^{-1}\sum_{t=1}^{T} \widehat{z}_t\widehat{z}_t'$.
  3. Maximize $-\frac{1}{2}\sum_{t=1}^{T}\left(\log|R_t| + \widehat{z}_t' R_t^{-1}\widehat{z}_t\right)$ over correlation dynamic parameters.

Example: European Stock Market Indices, 2015–2019



Daily Return On European Stock Market Indices

Example: Estimated DCC Model

- Univariate GARCH(1,1) models:

	UK	Germany	France	Netherlands
$\widehat{\sigma}_i$	0.0078	0.0097	0.0098	0.0096
$\widehat{\alpha}_i$	0.0237	0.0186	0.0050	0.0093
$\widehat{\gamma}_i$	0.2041	0.1935	0.1284	0.2332
$\widehat{\beta}_i$	0.8061	0.8376	0.9057	0.8509

- Sample correlation matrix of $\widehat{z}_t$

$$S = \begin{bmatrix} 1.0000 & 0.7366 & 0.7423 & 0.7859 \\ 0.7366 & 1.0000 & 0.8690 & 0.8860 \\ 0.7423 & 0.8690 & 1.0000 & 0.9063 \\ 0.7859 & 0.8860 & 0.9063 & 1.0000 \end{bmatrix}$$

- Correlation dynamics: $\widehat{\alpha} = 0.0683$, $\widehat{\beta} = 0.8764$.

Estimating Large-Dimensional Covariance Matrix

- Most - if not all - models involve estimating the unconditional covariance matrix in some sense:

$$\Sigma = \mathbb{E}H_t \quad \text{or} \quad \Sigma = \mathbb{E}Q_t$$

- In general, the sample covariance matrix

$$S = \frac{1}{T}\sum_{t=1}^{T} y_t y_t' = \frac{1}{T}Y'Y, \;\; Y = [y_1, \ldots, y_T]' \quad \text{or}$$

$$S = \frac{1}{T}\sum_{t=1}^{T} \widehat{z}_t \widehat{z}_t^T = \frac{1}{n}\widehat{Z}'\widehat{Z}, \;\; Z = [\widehat{z}_1, \ldots, \widehat{z}_T]'$$

  is inconsistent towards Σ in high dimensions with $N \asymp T$; even if one replace $\widehat{z}_t$ with z_t.
- In particular, $\|S - \Sigma\|_F^2 = \|\text{vec}(S) - \text{vec}(\Sigma)\|^2 \overset{\mathbb{P}}{\nrightarrow} 0$.

Rotation Equivariant Estimator

- Consider the spectral decomposition of sample covariance matrix S given by

$$S = U \Lambda U', \quad U = [u_1, \ldots, u_N],\ \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_N).$$

- The sample covariance matrix $S = S(Y') = \frac{1}{T} Y'Y$ is rotation equivariant: for all orthogonal matrix W

$$S(WY') = WS(Y)W'$$

- The class of rotation equivariant estimator is of the form

$$\widehat{\Sigma} = U \widehat{\Lambda} U', \quad \widehat{\Lambda} = \mathrm{diag}(d_1, \ldots, d_N),$$

where d_i are invariant with respect to data rotations.

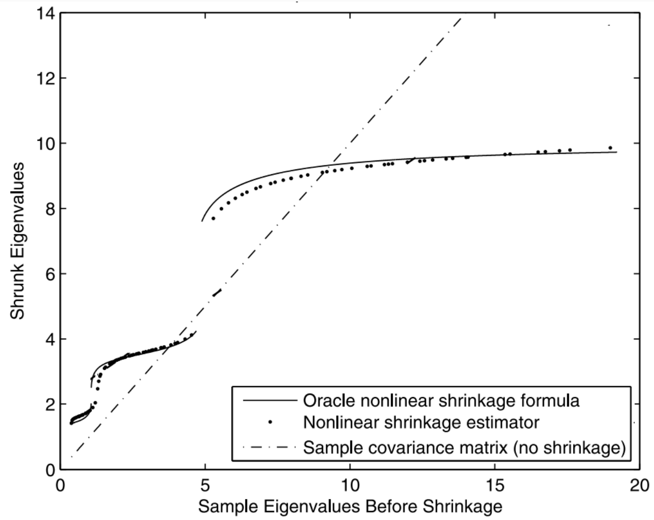
Optimal Rotation Equivariant Estimator

The sum of squared errors over all entries

$$\begin{aligned}\left\|\widehat{\Sigma}-\Sigma\right\|_F^2 &= \left\|\widehat{\Lambda}-U'\Sigma U\right\|_F^2 \\ &= \sum_{i=1}^{N}(d_i - u_i'\Sigma u_i)^2 + \underbrace{\sum_{i\neq j}(u_i'\Sigma u_j)^2}_{\overset{\mathbb{P}}{\not\to} 0}\end{aligned}$$

is minimized with $d_i = u_i'\Sigma u_i =: d_i^*$.

- Ledoit and Péché (2011): $d_i^* \approx \delta(\lambda_i)$ for some shrinkage function δ for $Y' = \Sigma^{1/2}Z'$ for Z with i.i.d. entries
- Ledoit and Wolf (2012): (Consistent) Estimation of δ



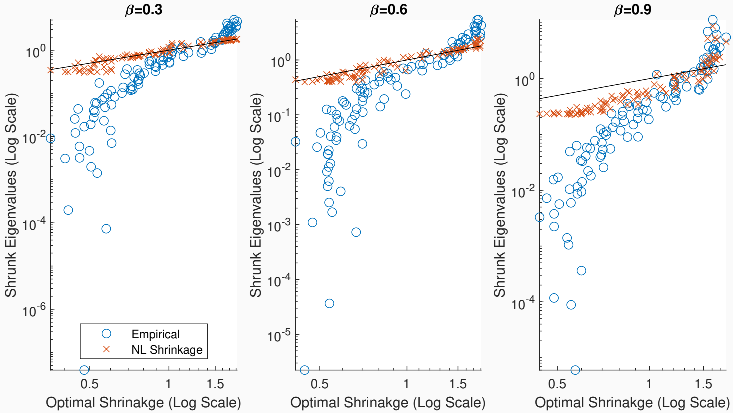
A typical simulation result for IID data with $N = 100$ and $T = 300$ in Ledoit and Wolf (2012): 20% of population eigenvalues of Σ are equal to 1, 40% are equal to 3 and 40% are equal to 10.

Engle, Ledoit, and Wolf (2019)

- For each asset, fit a univariate GARCH(1,1) model and use the fitted model to devolatilize the return series.
- Take LW2012's nonlinear shrinkage estimator 'off the shelf' and apply them to devolatilized returns $\widehat{z}_t$ to get $\widehat{\Sigma}$ (inconsistent but 'optimal' in terms of squared errors)
- Plug in $\widehat{\Sigma}$ and then fit the remaining 2 DCC model parameters $\widehat{\alpha}, \widehat{\beta}$ using the composite likelihood method in Pakel et al.(2021).

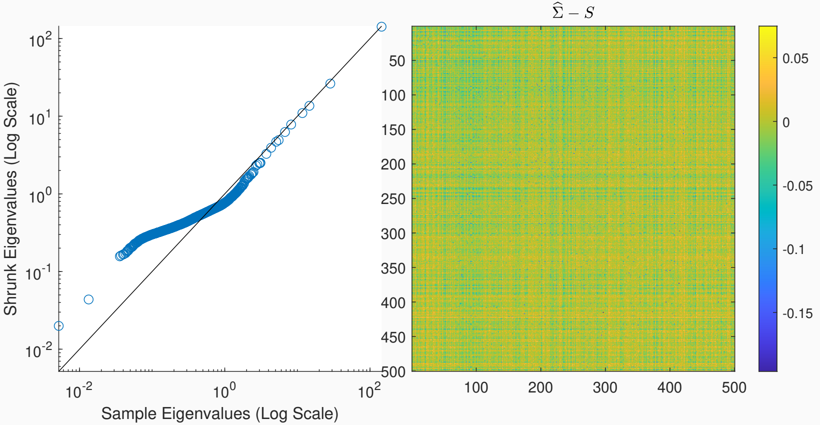
Simulation: Estimating Σ in DCC model

- Univariate GARCH: $\sigma_i^2 = 1$, $\alpha_i = 0.05$, $\beta_i = 0.9$
- Correlation dynamics: $\alpha = 0.05$, $\beta \in \{0.3, 0.6, 0.9\}$



Improvement $1 - \frac{\|\widehat{\Sigma}-\Sigma\|_F^2}{\|S-\Sigma\|_F^2} = 87.0\%, 87.3\%, 47.6\%$ from left to the right

Example: 500 US Stocks from 2015–2019



Correlation dynamics

- with shrinkage: $\widehat{\alpha} = 0.0417$ and $\widehat{\beta} = 0.8136$
  `https://youtu.be/SeDwJfykgqA`
- without shrinkage: $\widehat{\alpha} = 0.0413$ and $\widehat{\beta} = 0.8083$

Evaluating High-Dimensional Volatility Model

- Diagnostic testing are possible (only) in low dimensions: see Bauwens et al.(2006)
- In high dimensions, performance in risk management may be evaluated by testing implied Value at Risk for portfolio $w' y_t$. (Joint test of specification of H_t and distributional assumption.)
- ... or testing standardized residuals $\mathrm{SR}_t = \frac{(w_t' y_t)^2}{w_t' H_t w_t}$ via the regression

$$\mathrm{SR}_t = \mu + \rho \mathrm{SR}_{t-1} + \nu_t$$

- ... or checking the excess variance of the implied minimum-variance portfolio
- ... or tracking error of a particular portfolio

Exercises

1. Show that the matrix H_t defined by the BEKK model is positive definite.
2. Show that the BEKK model is a special case of the general VEC model specified by Equation (4) in Bollerslev et al.(1988), using the property $\text{vec}(ABC) = (C' \otimes A)\text{vec}(B)$ for conformable matrices A, B and C, and the duplication matrix D_N, defined by $\text{vec}(A) = D_N \text{vech}(A)$ for a symmetric $N \times N$ matrix A.
3. Show that the factor model is a special case of the BEKK model.