Multivariate Time Series Analysis in R

Ruey S. Tsay
Booth School of Business
University of Chicago

May 2013, R/Finance Conference

Objective

Analysis of multivariate time-series data using R:

- To obtain parsimonious models for estimation
- To extract "useful" information when the dimension is high
- To make use of prior information or substantive theory
- To consider also multivariate volatility modeling and applications

Outline (3 main parts)

- Multivariate time series analysis ("MTS" package)
  1. VAR, VMA, VARMA, Seasonal VARMA, VARMAX, Factor models, Multivariate volatility models, etc.
  2. Simple demonstration
- Factor models (dimension reduction)
  1. Constrained factor models
  2. A motivating example
  3. Partially constrained factor models
  4. Application in risk management
- Principal volatility component analysis
  1. Generalized kurtosis matrix
  2. Simple illustration

Multivariate time series analysis

- Difficulties
  1. Too many parameters when the dimension is high
  2. Identifiability problems

- Solutions
  1. Stay away: focus on Vector autoregressive models (VAR)
  2. Structural specification: Kronecker index and Scalar components
     Tools used: Canonical correlation analysis, likelihood ratio test
  3. Factor models (with or without constraints)
     Tools used: PCA, LASSO, K means, model-based classification

A 2-dimensional VARMA(2,2) model

$$\begin{aligned}
\mathbf{z}_t &= \begin{bmatrix} 0.6 & -0.5 \\ -0.5 & 0.6 \end{bmatrix} \mathbf{z}_{t-1} + \begin{bmatrix} -0.1 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \mathbf{z}_{t-2} \\
&= \mathbf{a}_t - \begin{bmatrix} -1.05 & 0.15 \\ 0.15 & -1.05 \end{bmatrix} \mathbf{a}_{t-1} - \begin{bmatrix} -0.26 & 0.06 \\ 0.06 & -0.26 \end{bmatrix} \mathbf{a}_{t-2},
\end{aligned}$$

where $\text{Cov}(\mathbf{a}_t) = \boldsymbol{\Sigma}_a > 0$.

The model can be simplified! $w_t = z_{1t} + z_{2t}$ satisfies $(1 - 0.5B)w_t = (1 + 0.5B)b_t$.

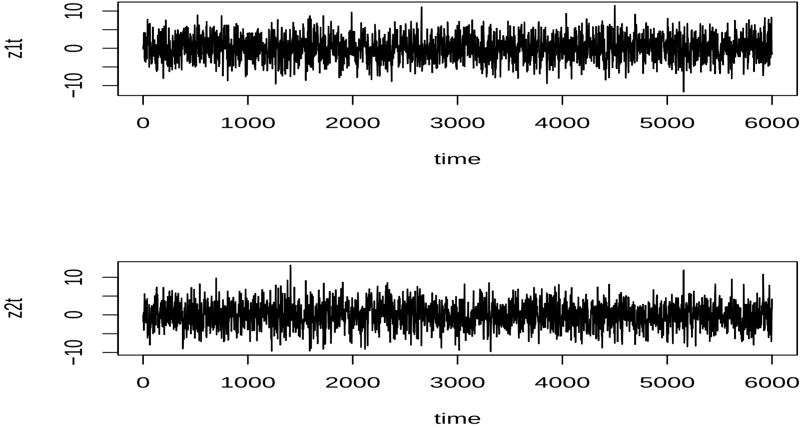


Figure: Simulation: a bivariate series with 6000 observations

Model specification

- VAR models
  1. AIC and HQ: VAR(12)
  2. BIC: VAR(10)
  3. Tiao-Box chi-square: VAR(13)
- VARMA models
  1. ECCM (Tiao-Tsay): VARMA(2,2)
  2. Kronecker index: $\{2, 1\}$, implying a VARMA(2,2) model, but with a VARMA(1,1) component.

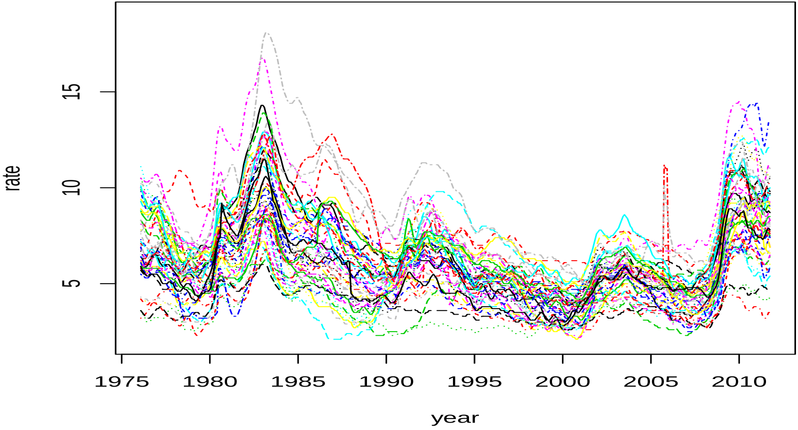


Figure: Time plots of the monthly unemployment rates of the 50 States in the U.S. from January 1976 to September 2011. The data are seasonally adjusted.

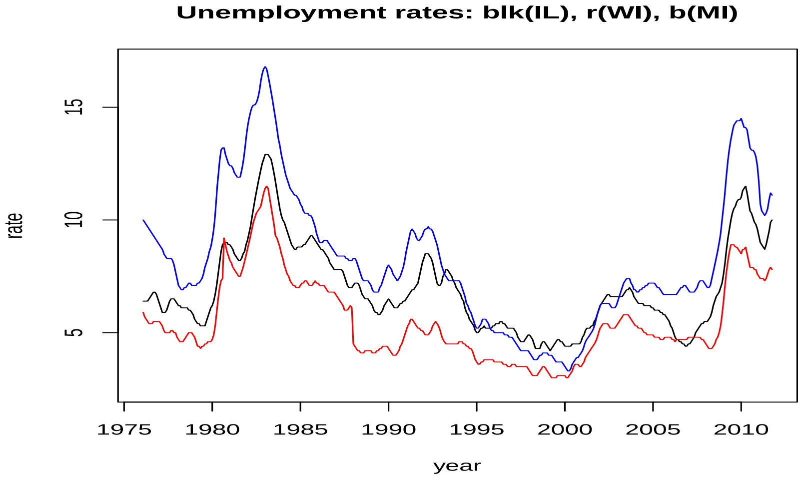


Figure: Time plots of the monthly unemployment rates of IL, WI, and MI from January 1976 to September 2011. The data are seasonally adjusted.

Summary of modeling

- VAR approach: AIC selects a VAR(7) model, with 35 parameters

- VARMA approach: a VARMA(2,2) model, with 29 parameters

Both models fit the data reasonably well. The two models are similar.

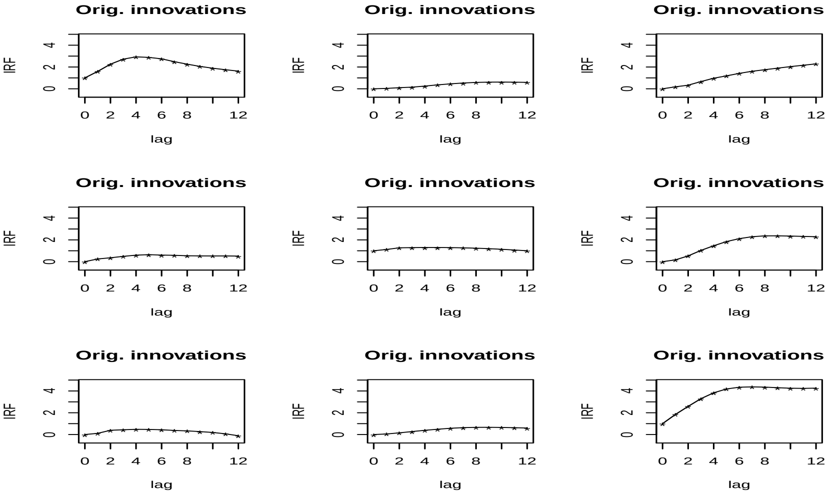


Figure: Impulse response functions of A VAR(7) model for monthly unemployment rates of IL, WI, and MI from January 1976 to September 2011.

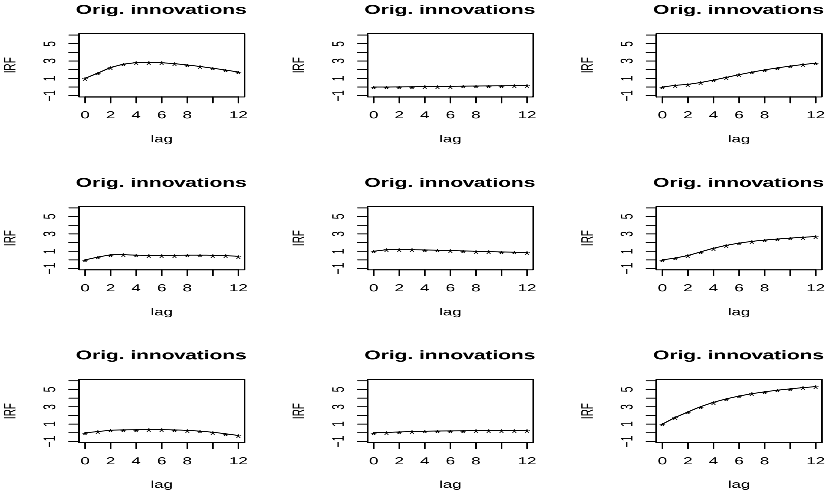


Figure: Impulse response functions of a VARMA(2,2) model for monthly unemployment rates of IL, WI, and MI from January 1976 to September 2011.

Factor models

For high-dimensional series: Dimension reduction and ease in interpretation

Approximate factor models

$$\begin{aligned} \mathbf{x}_t &= \mathbf{L}\mathbf{f}_t + \epsilon_t \\ y_{t+h} &= \beta'\mathbf{f}_t + \gamma'\mathbf{w}_t + v_{t+h} \end{aligned}$$

where $\mathbf{x}_t$ is an N-dimensional random vector, $\mathbf{L}$ is an $N \times r$ loading matrix, $\mathbf{f}_t$ is the r-dimensional common factors, $\mathbf{w}_t$ is a pre-determined vector that may contain lagged values of y_t, $h > 0$ is the forecast horizon, ϵ_t and v_t are the noise terms, respectively.

Usual assumptions:

- All variables have zero means.
- $E(\mathbf{f}_t\mathbf{f}_t') = \mathbf{I}_r$.
- $E(\epsilon_t\epsilon_t') = \mathbf{\Psi}$ (positive definite)
- $E(\mathbf{f}_t\epsilon_t') = \mathbf{0}$, $E(\mathbf{f}_t v_{t+h}) = \mathbf{0}$, & $E(\mathbf{w}_t v_{t+h}) = \mathbf{0}$.
- Rank$(\mathbf{L}) = r$ and $\frac{1}{N}\mathbf{L}'\mathbf{L}$ positive definite as $N \to \infty$.
- Additional conditions needed if $\mathbf{\Psi}$ is not diagonal, i.e. bounded eigenvalues.

Discussions

Some difficulties often encountered when N is large:

- Hard to understand or interpret the estimated common factors.
- Does a large N produce more accurate forecasts? (Not necessarily)
- y_t plays no role in factor estimation.
- Does not make use of any prior information or theory or past experience.

Our goal is to overcome some of these weaknesses.

Constrained factor model

$\mathbf{H}$ is an $N \times m$ matrix of **known** constraints. The model becomes

$$\mathbf{x}_t = \mathbf{H}\boldsymbol{\omega}\mathbf{f}_t + \epsilon_t$$

where $\boldsymbol{\omega}$ is an $m \times r$ matrix, Rank($\mathbf{H}$) $= m$ and Rank($\boldsymbol{\omega}$) $= r$. Typically, $r \leq m << N$. [Simply put, $\mathbf{L} = \mathbf{H}\boldsymbol{\omega}$.]

Examples:

- For stock returns, columns of $\mathbf{H}$ may indicate the industrial sectors of the stock.
- For interest rates, columns $\mathbf{H}$ may indicate *level, slope* and *curvature* of the yield curve.

Motivating example

Monthly excess returns of 10 stocks: (less 3-month T bill)

(a) Pharmaceutial: Abbott Labs, Eli Lilly, Merck, and Pfizer
(b) Auto: General Motors and Ford
(c) Oil: BP, Chevron, Royal Dutch, and Exxon-Mobil

Sample period: January 1990 to December 2003 for 168 observations.

Example continued: traditional factors

Results of traditional PCA using correlations:

- Eig. Values: 3.890, 1.971, 1.498, 0.586, 0.498, ..., 0.242
- first 3 vectors:

abt	0.280	-0.355	0.1196
lly	0.244	-0.463	0.0110
mrk	0.296	-0.432	0.0462
pfe	0.337	-0.337	0.1115
gm	0.249	0.007	-0.6311
f	0.180	0.070	-0.7030
bp	0.351	0.326	0.1977
cvx	0.376	0.346	0.1318
rd	0.411	0.244	0.1366
xom	0.364	0.261	0.0574

Example continued.

Make use of the knowledge of three industries:

$$\mathbf{H}' = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Perform a constrained analysis: (least-squares estimates)
Eigen Values:

- Constrained space: 3.813, 1.917, 1.362
- Residual space: 0.660, 0.575, 0.517, ..., 0.256.

Example continued: Loading matrix

stock	Unconstrained			Constrained		
		$\mathbf{L}$			$\mathbf{H}\omega$	
abt	0.551	-0.497	0.141	0.568	-0.556	0.074
lly	0.480	-0.649	0.013	0.568	-0.556	0.074
mrk	0.583	-0.605	0.054	0.568	-0.556	0.074
pfe	0.663	-0.471	0.131	0.568	-0.556	0.074
gm	0.490	0.009	-0.744	0.423	0.071	-0.783
f	0.353	0.098	-0.829	0.423	0.071	-0.783
bp	0.690	0.457	0.233	0.736	0.409	0.168
cvx	0.739	0.485	0.155	0.736	0.409	0.168
rd	0.809	0.342	0.161	0.736	0.409	0.168
xom	0.715	0.365	0.068	0.736	0.409	0.168

Example continued.

Discussions:

- Constrained model is more parsimonious (10×3 vs. 3×3)
- Sector variations explain the variability in the excess returns (equal loading for stocks in the same industry)
- The spaces spanned by the common factors are essentially the same with/without constraints
  Canonical correlations between the two sets of common factors are
  $$0.9997, 0.9990, 0.9952.$$
- Both maximum likelihood and least squares estimations available
- Test is available for checking the constraints. Tsai and Tsay (2010, JASA)

Partially constrained factor models

In practice, it is likely that only partial constraints are available.

$$\begin{aligned} \mathbf{x}_t &= \mathbf{H}\omega\mathbf{f}_t + \mathbf{L}\mathbf{g}_t + \epsilon_t, \\ y_{t+h} &= \beta_1'\mathbf{f}_t + \beta_2'\mathbf{g}_t + v_{t+h}, \quad t = 1, \ldots, T, \end{aligned}$$

where $\mathbf{L}$ is an $N \times p$ unconstrained loading matrix of rank p and $\mathbf{g}_t$ is a p-dimensional unconstrained common factors.
Additional assumptions:
$E(\mathbf{g}_t) = \mathbf{0}$, $E(\mathbf{g}_t\mathbf{g}_t') = \mathbf{I}_p$, $E(\mathbf{f}_t\mathbf{g}_t') = 0$ and $\mathbf{H}'\mathbf{L} = \mathbf{0}$.
$E(\mathbf{g}_t v_{t+h}) = \mathbf{0}$

Principal Volatility Components (PVC)

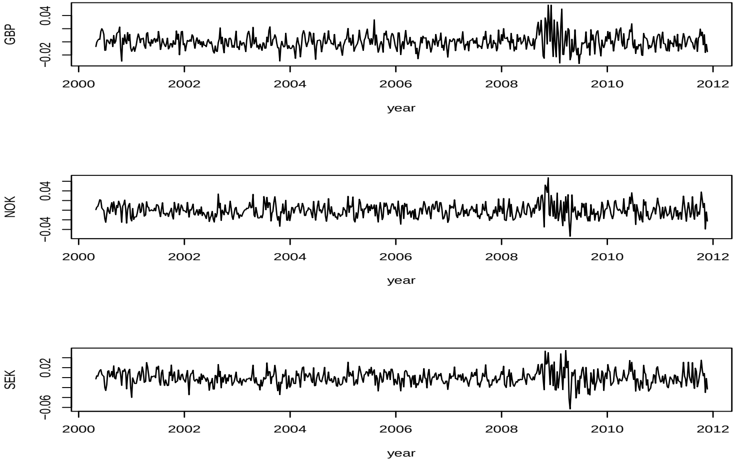
- Common volatility components
- On going research, joint with Y. Hu
- Different from applying PCA to asset returns

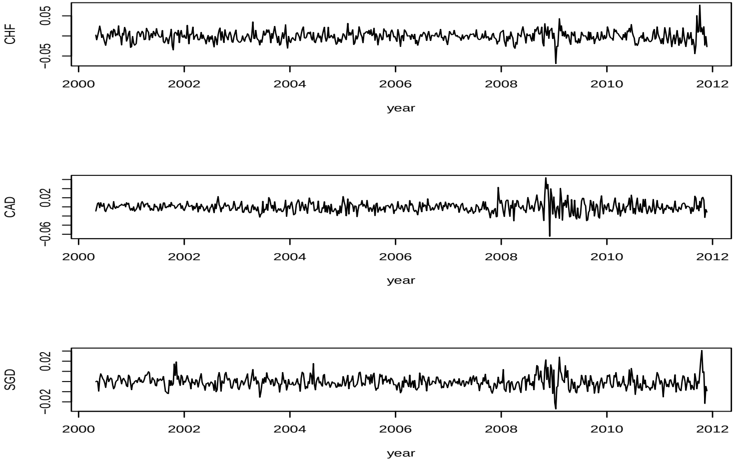
Outline of PVC

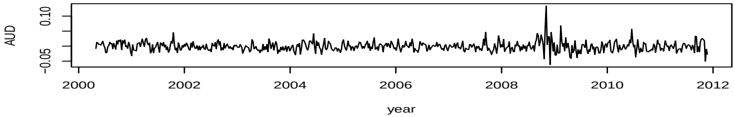
- Motivation: Are there no-ARCH portfolios among financial assets? Are there common volatility components?
- Definition of ARCH Dimension and Transformation
- Generalized Cross Kurtosis Matrices
- The Principal Volatility Components
- Estimation and Testing (skipped)
- Data Analysis

Motivation: data of seven exchange rates

- We consider exchange rates of seven currencies against US dollars: They are British Pound, Norwegian Kroner, Swedish Kroner, Swiss Francs, Canadian Dollar, Singapore Dollar, and Australian Dollar.
- We employed weekly log returns of the exchange rates from March 29, 2000 to October 26, 2011. Each series has 605 observations.


GBP
0.04
−0.02
2000
2002
2004
2006
2008
2010
2012
year
NOK
0.04
−0.04
2000
2002
2004
2006
2008
2010
2012
year
SEK
0.02
−0.06
2000
2002
2004
2006
2008
2010
2012
year


CHF
0.05
−0.05
CAD
0.02
−0.06
SGD
0.02
−0.02
2000
2002
2004
2006
2008
2010
2012
year


AUD
0.10
-0.05
2000
2002
2004
2006
2008
2010
2012
year

Motivation (continued)

- We are interested in the volatility movements of the exchange rates. A VAR(5) model is adopted to remove the dynamic linear dependence in the data. We employ the residual series in the following analysis.
- Based on the Ljung-Box test and Engle's LM test of the squared series, each residual series has significant ARCH effects.
- The ARCH effect implies that the conditional variance is time-varying.

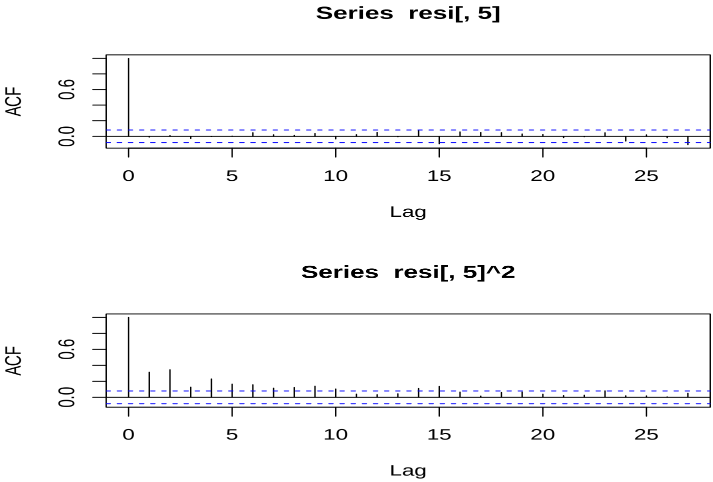


Figure: Autocorrelations of residuals and squared residuals of Canadian Dollars

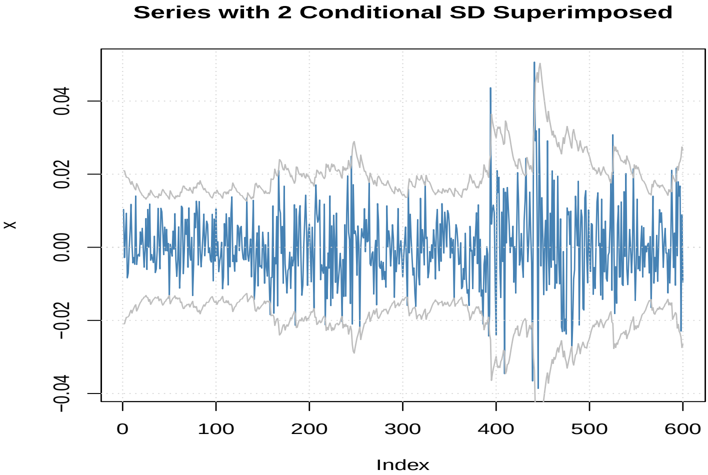


Figure: Volatility plot of log returns of Canadian Dollars

Summary statistics

Stat	BP	NK	SK	SF
s.e.	0.0101	0.0131	0.0131	0.0120
Skewness	0.4681	0.5578	0.4581	0.3369
Kurtosis	0.7419	0.5970	0.8454	1.3008
Q(10,ϵ_t)	5.68(.8)	2.80(1.)	7.84(.6)	9.58(.5)
Q(10,ϵ_t^2)	74.73	90.87	42.95	56.09

Stat	CD	SD	AD
s.e.	0.0105	0.0053	0.0144
Skewness	0.4765	0.5866	0.9295
Kurtosis	1.2973	1.1980	2.2019
Q(10,ϵ_t)	4.46(.9)	5.86(.8)	11.42(.3)
Q(10,ϵ_t^2)	235.3	41.46	93.04

Motivation (continued)

- Is there common volatility? (Global integration? Interdependence?)
  Although each series displays ARCH effects, is it possible to find some linear combinations of these seven variables that mitigate the ARCH effect? (to reduce risk)
- If yes, how to construct no-ARCH effect portfolios from these seven exchange rates?

ARCH Dimension

For k assets, how to quantify volatility factors?
Answer: ARCH dimension.

Suppose $\mathbf{y}_t = (y_{i,t}, \cdots, y_{k,t})'$ is a k-dimensional series with conditional mean zero, and a time-varying conditional covariance

$$\mathbf{\Sigma}_t = E(\mathbf{y}_t \mathbf{y}_t' | \Omega_{t-1}),$$

where Ω_{t-1} is the σ-field generated by $\{\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \cdots\}$.

ARCH Dimension and Transformation

- Seek simplifying structure in $\mathbf{\Sigma}_t$.
- Suppose a $k \times k$ matrix $\mathbf{M} = (\mathbf{M}_1', \mathbf{M}_2')'$ exists such that the transformed series $\mathbf{My}_t = [(\mathbf{M}_1\mathbf{y}_t)', (\mathbf{M}_2\mathbf{y}_t)']'$ satisfies

$$E[\mathbf{My}_t(\mathbf{My}_t)'|\Omega_{t-1}] = \mathbf{M\Sigma}_t\mathbf{M}' = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_{2t} \\ \mathbf{C}_{2t}' & \mathbf{\Delta}_t \end{bmatrix},$$

where $\mathbf{\Delta}_t$ is an $r \times r$ time-varying conditional covariance of $\mathbf{M}_2\mathbf{y}_t$, where $0 \leq r \leq k$.

ARCH Dimension and Transformation

- The **ARCH dimension** of $\mathbf{y}_t$ is r if a $(k - r) \times k$ matrix $\mathbf{M}_1$ exists such that $\text{Cov}(\mathbf{M}_1 \mathbf{y}_t | \Omega_{t-1}) = \mathbf{M}_1 \mathbf{\Sigma}_t \mathbf{M}_1'$ is a constant matrix.
- The $\mathbf{M}_1$ is referred to as a **no-ARCH transformation**.
- Engle and Kozicki (1993) and Engle and Susmel (1993) discussed a pairwise method to identify the common feature in volatility.

Generalized covariance matrix

Statistics of interest:
Volatility models are essentially concerned with the linear dynamic dependence of the matrix process $\{\mathbf{y}_t\mathbf{y}_t'|t = 1, \ldots, T\}$.
A multivariate ARCH(1) model:

$$E(\mathbf{y}_t\mathbf{y}_t'|\Omega_{t-1}) = \mathbf{A}_0\mathbf{A}_0' + \mathbf{A}_1(\mathbf{y}_{t-1}\mathbf{y}_{t-1}')\mathbf{A}_1'.$$

That is, elements of $\mathbf{y}_t\mathbf{y}_t'$ relate to $y_{i,t-1}y_{j,t-1}$.

Generalized covariance matrix (cont.)

- Define the **Generalized Covariance**

$$\text{Cov}(\mathbf{y}_t\mathbf{y}'_t, x_{t-1}) = [\text{Cov}(y_{it}y_{jt}, x_{t-1})]$$

  where $x_{t-1} \in \Omega_{t-1}$ is a univariate random variable.
- The generalized covariance satisfies that

$$\text{Cov}\{\mathbf{M}\mathbf{y}_t(\mathbf{M}\mathbf{y}_t)', x_{t-1}\} = \mathbf{M}\text{Cov}(\mathbf{y}_t\mathbf{y}'_t, x_{t-1})\mathbf{M}'.$$

The idea has been used before, e.g. Li (1992, pHd method)

Generalized lag-ℓ cross-kurtosis matrix

$$\boldsymbol{\gamma}_\ell = \sum_{i=1}^{k}\sum_{j=i}^{k} \text{Cov}^2(\mathbf{y}_t\mathbf{y}_t', y_{i,t-\ell}y_{j,t-\ell}) \equiv \sum_{i=1}^{k}\sum_{j=i}^{k} \boldsymbol{\gamma}_{\ell,ij}^2$$

where

$$\boldsymbol{\gamma}_{\ell,ij} = \text{Cov}(\mathbf{y}_t\mathbf{y}_t', y_{i,t-\ell}y_{j,t-\ell}),$$

and the square is used to ensure non-negative definite of $\boldsymbol{\gamma}_\ell$. Both $\boldsymbol{\gamma}_{\ell,ij}$ and $\boldsymbol{\gamma}_\ell$ are $k \times k$ symmetric matrix.

Implication of a zero eigenvalue

Consider γ_ℓ with ℓ fixed. Let $\mathbf{u}$ be a k-dimensional vector associated with a zero eigenvalue if it exists. That is, $\gamma_\ell \mathbf{u} = \mathbf{0}$. Then, $\gamma^2_{\ell,ij}\mathbf{u} = \mathbf{0}$, implying

$$\gamma_{\ell,ij}\mathbf{u} = \mathbf{0},$$

for all i and j. This implies that $\mathbf{y}_t\mathbf{y}'_t$ is not correlated with $y_{i,t-\ell}y_{j,t-\ell}$ for all i, j.

Data Analysis: FX

- Recall the data of seven exchange rates ($k = 7$). Each return series contains 605 observations. A VAR(5) model is adopted to remove the serial correlation to have residual series $\mathbf{y}_t$.
- Each residual series displays ARCH effects. We ask "whether there exists no-ARCH portfolios?"

Data Analysis (continue)

- We adopt $\widehat{\mathbf{\Gamma}}_m$ for $m = 10$ to estimate the transformation $\widehat{\mathbf{M}}$, where $\widehat{\mathbf{\Gamma}}_m = \sum_{h=1}^{m} \sum_{i=1}^{k} \sum_{j=i}^{k} (1 - h/n)^2 \widehat{\text{cov}}^2(\mathbf{y}_t \mathbf{y}_t', y_{i,t-h} y_{j,t-h})$.

- The estimates of principal volatility component are $\widehat{\mathbf{M}}\mathbf{y}_t$.

PVC	7th	6th	5th	4th	3rd	2nd	1st
Values	0.045	0.078	0.109	0.120	0.165	0.191	0.409
Vectors	-0.232	0.366	0.165	0.656	-0.002	0.214	-0.197
	-0.187	-0.754	0.177	0.159	-0.102	-0.029	-0.309
	-0.216	0.510	-0.329	-0.313	-0.331	-0.351	-0.235
	-0.219	-0.072	-0.236	-0.016	0.143	0.198	0.631
	0.569	-0.054	-0.399	0.294	0.698	-0.060	-0.038
	0.663	0.173	0.762	0.365	-0.078	-0.850	0.641
	-0.233	0.012	0.191	-0.476	0.605	0.254	0.028

s	$p_n = 5$	$p_n = 10$	$p_n = 15$	$p_n = 20$
	(a) $m = 10$			
1	-0.72(0.77)	-1.00(0.84)	-0.64(0.74)	-0.39(0.65)
2	3.51(0.00)	4.51(0.00)	5.17(0.00)	6.76(0.00)
3	11.30(0.00)	13.22(0.00)	13.45(0.00)	14.41(0.00)
4	26.61(0.00)	29.07(0.00)	28.92(0.00)	29.80(0.00)
5	31.33(0.00)	34.90(0.00)	33.76(0.00)	32.88(0.00)
6	44.20(0.00)	48.41(0.00)	48.24(0.00)	47.55(0.00)
7	57.09(0.00)	62.20(0.00)	63.84(0.00)	64.78(0.00)
	(c) $m = 20$			
1	-0.76(0.78)	-1.03(0.85)	-0.78(0.78)	-0.60(0.73)
2	4.95(0.00)	4.92(0.00)	4.29(0.00)	3.76(0.00)
3	17.47(0.00)	18.16(0.00)	17.14(0.00)	16.16(0.00)

Data Analysis: Seven exchange rates

- The results of ARCH dimension test using Generalized Ling-Li statistic indicate that there is only one no-ARCH portfolio.
- Further analysis of the principal volatility components confirms the findings.

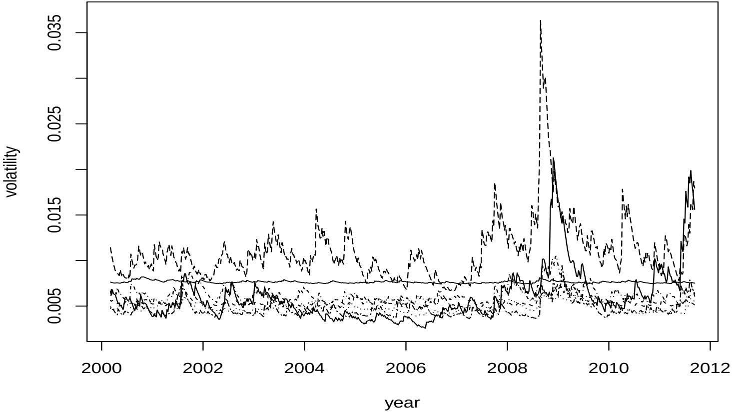


Figure: Volatility series of PVCs

Data Analysis: Seven exchange rates

- We have tried different values of the parameters to estimate $\widehat{\mathbf{\Gamma}}_m$ and $GT_{p_n,s}$. Results are quite robust to the parameters.

- In conclusion, there is a linear combination with no-ARCH effect among the seven exchange rates.

No ARCH poerfolio

$$e_{7t} \approx 0.2(\text{GBP}+\text{NOK}+\text{SEK}+\text{CHF}+\text{AUD}) - 0.6(\text{CAD}+\text{SGD}).$$

Conclusion

- General multivariate models can be used in R to analyze multiple processes
- Constraints can be used effectively in factor models to simplify the interpretation and estimation of the common factors
- Principal volatility components are proposed and they can be used to simplify multivariate volatility modeling.
- Many problems remain open! For instance, copula models, factor models for volatility, etc.