Summary notes on financial econometrics

Radu Briciu

March 5, 2025

Contents

Preface

My name is Radu Briciu and I write this book to attempt the impossible task of keeping up to date with protoeconometrics. I cordially invite you to join my descent into madness.

https://www.lambda.art

Econometric Interpretation:

Let $\{y_{1,t}\}_{t=1}^{T}$ and $\{y_{2,t}\}_{t=1}^{T}$ be two time series, each decomposable as

$y_{i,t} = \mu_t + \epsilon_{i,t}, \quad i = 1, 2,$

where μ_t is a common trend component and $\epsilon_{i,t}$ are idiosyncratic shocks.

In our prototypist framework, each series is encoded as a normalized state:

$$|y_i\rangle = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_{i,t} \, |t\rangle, \quad i = 1, 2.$$

The similarity between the series is measured via the inner product,

$$\langle y_1 | y_2 \rangle = \frac{1}{T} \sum_{t=1}^{T} y_{1,t} y_{2,t},$$

which (after normalization) corresponds to the sample covariance or correlation.

Chapter 1

Estimation Methods

1.1 OLS vs Maximum Likelihood

1.1.1 Ordinary Least Squares (OLS)

OLS minimises the sum of squared residuals:

$$\hat{\beta} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y} \tag{1.1}$$

where:

- $\mathbf{X}$ is the matrix of predictors.
- $\mathbf{y}$ is the vector of observations.
- $\hat{\beta}$ is the vector of estimated coefficients.

1.1.2 Maximum Likelihood Estimation (MLE)

MLE maximises the likelihood function:

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} f(y_i; \theta) \tag{1.2}$$

The log-likelihood is given by:

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \log f(y_i; \theta) \tag{1.3}$$

The MLE estimate $\hat{\theta}$ solves:

$$\max_{\theta} \frac{\partial \mathcal{L}(\theta)}{\partial \theta} = 0 \tag{1.4}$$

An alternative to maximisation of the exact log-likelihood function with numerical procedures is to consider as deterministic and maximise the likelihood conditioned on the first observation

Chapter 2

Elementary Constructions

2.1 Two-Step Engle and Granger Approach

1. Estimate a linear relationship:

$$y_t = \alpha + \beta x_t + \epsilon_t \tag{2.1}$$

2. Test residuals $\hat{\epsilon}_t$ for stationarity using the ADF test.

2.2 ARMA models

2.2.1 Autoregressive (AR) model

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t \tag{2.2}$$

where heteroscedasticity is assumed.

2.2.2 Moving Average (MA) model

$$y_t = \mu_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t \tag{2.3}$$

where $\epsilon_t \sim$ **G.W.N**.

2.2.3 ARMA(p, q) model

$$y_t = \delta + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t \tag{2.4}$$

$$y_t = \delta + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \tag{2.5}$$

2.3 VAR model

We can conduct a vector autoregression by ordering the variables in this manner:

$$y_t = \alpha + \theta_1 y_1 + \cdots + \theta_p y_p + \epsilon_t \tag{2.6}$$

where variable y_t is essentially an envelopment of p observed quantitative variables. This allows us to then construct a VAR model as such:

$$Y_t = \mu + \Gamma_1 y_{t-1} + \sum_{i=2}^{q} \Gamma_i Y_{t-i} + \omega \tag{2.7}$$

where the vector Y_t is estimated by its last available discretely observed value, and $q - 1$ equidistant lags of equal duration in time t. The exogenous components μ and ϵ_t are the long run mean of our process and the measurement error, respectively.

Chapter 3

Modeling Volatility

3.1 Introduction

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models and Multivariate GARCH (MGARCH) models are widely used in financial econometrics to capture time-varying volatility. This chapter provides a detailed comparison of the two model classes.

3.2 Dimensionality

- **GARCH Models**: These are *univariate* models designed to capture time-varying volatility in a single time series.
- **MGARCH Models**: These extend GARCH to the *multivariate* case, modeling the time-varying covariance structure across multiple assets or time series.

3.3 Purpose

- **GARCH**: Primarily used to model *conditional variance* for a single financial asset or time series.
- **MGARCH**: Used to model the *conditional covariance matrix* between multiple assets, which is crucial for *portfolio risk management and asset allocation.*

3.4 Conditional Variance vs. Conditional Covariance

- **GARCH** models the conditional variance of a single asset return.
- **MGARCH** models both the conditional variances and conditional covariances of multiple asset returns.

3.5 Mathematical Formulation

3.5.1 GARCH(1,1)

A univariate GARCH(1,1) model is typically written as:

$$h_t = \omega + \alpha\epsilon_{t-1}^2 + \beta h_{t-1}, \tag{3.1}$$

where h_t is the conditional variance.

3.5.2 MGARCH

MGARCH models introduce a *matrix formulation* for the covariance structure:

$$H_t = C + A\epsilon_{t-1}\epsilon_{t-1}'A' + BH_{t-1}B', \tag{3.2}$$

where H_t is the conditional covariance matrix.

3.6 Model Variants

Some common MGARCH models include:

- **Constant Conditional Correlation (CCC) GARCH**: Assumes constant correlations but time-varying variances.
- **Dynamic Conditional Correlation (DCC) GARCH**: Allows for dynamic correlations over time.
- **BEKK GARCH**: Ensures positive-definiteness of the covariance matrix but has high computational costs.

3.7 Computational Complexity

- **GARCH**: Computationally simpler as it only deals with a *single variance equation*.
- **MGARCH**: More complex because it must estimate *multiple covariance equations*, making it computationally demanding.

3.8 Application Areas

- **GARCH**: Used in *option pricing, risk management, and volatility forecasting* for individual assets.
- **MGARCH**: Essential for *portfolio management, systemic risk analysis, and asset pricing*, where interactions between multiple assets are critical.

3.9 Conclusion

While GARCH models are useful for univariate volatility modeling, MGARCH models provide a framework for capturing interdependencies across multiple assets. The choice between the two depends on the specific application and computational considerations.

Chapter 4

Multivariate Additions

4.1 Introduction

Volatility modeling plays a central role in modern financial econometrics. While univariate GARCH models capture time-varying conditional variance in a single asset return series, many practical applications—such as portfolio risk management and asset pricing—require understanding the joint dynamics of several assets. Multivariate GARCH (MGARCH) models extend the univariate framework by modeling the evolution of the entire conditional covariance matrix. This chapter briefly reviews several key formulations of MGARCH models and outlines their estimation and applications.

4.2 MGARCH Model Formulations

In general, let $\mathbf{r}_t$ denote an N-dimensional vector of asset returns with zero conditional mean and conditional covariance matrix $\mathbf{H}_t$. A basic multivariate GARCH model may be written as:

$$\mathbf{r}_t = \mathbf{H}_t^{1/2}\mathbf{z}_t, \tag{4.1}$$

where $\mathbf{z}_t$ is an N-vector of i.i.d. random variables (typically standard normal).

Several variants of MGARCH have been proposed; below we briefly describe three popular classes.

4.2.1 The VECH Model

The VECH model directly parameterizes the vectorized lower triangular part of the conditional covariance matrix. Let

$$\operatorname{vech}(\mathbf{H}_t) = \mathbf{c} + \sum_{i=1}^{q} \mathbf{A}_i \operatorname{vech}(\mathbf{r}_{t-i}\mathbf{r}_{t-i}') + \sum_{j=1}^{p} \mathbf{B}_j \operatorname{vech}(\mathbf{H}_{t-j}), \tag{4.2}$$

where $\operatorname{vech}(\cdot)$ denotes the half-vectorization operator. Although this formulation is conceptually straightforward, the number of parameters increases rapidly with the dimension N.

4.2.2 The BEKK Model

A common (restricted) BEKK specification is

$$\mathbf{H}_t = \mathbf{C}'\mathbf{C} + \sum_{i=1}^{q} \mathbf{A}_i' \mathbf{r}_{t-i}\mathbf{r}_{t-i}'\mathbf{A}_i + \sum_{j=1}^{p} \mathbf{B}_j' \mathbf{H}_{t-j}\mathbf{B}_j, \tag{4.3}$$

where $\mathbf{C}$ is a lower triangular matrix. This structure guarantees that $\mathbf{H}_t$ is positive definite.

4.2.3 The Dynamic Conditional Correlation Model

An alternative approach is to decouple the dynamics of individual volatilities from the evolution of correlations. In the DCC model, one first estimates univariate GARCH models for each return series to obtain conditional standard deviations, and then models the time-varying correlation matrix separately:

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t, \tag{4.4}$$

where $\mathbf{D}_t$ is a diagonal matrix of conditional standard deviations and $\mathbf{R}_t$ is the dynamic conditional correlation matrix.

4.3 Estimation and Inference

Estimation of MGARCH models is typically achieved via maximum likelihood methods. In practice, one must impose parameter restrictions to ensure stationarity and positive definiteness of $\mathbf{H}_t$. For example, in the BEKK formulation, it is common to require that all eigenvalues of the matrices

$$\sum_{i=1}^{q} \mathbf{A}_i'\mathbf{A}_i + \sum_{j=1}^{p} \mathbf{B}_j'\mathbf{B}_j$$

lie within the unit circle.

In addition to full maximum likelihood estimation, quasi-maximum likelihood and two-step estimation procedures (such as those used in the DCC model) are widely applied in empirical research.

4.4 Applications

MGARCH models are employed extensively in financial applications. They provide a framework for:

- **Portfolio optimization:** Capturing the dynamics of asset covariances is essential for optimal asset allocation.
- **Risk management:** Accurate forecasts of the conditional covariance matrix are critical for computing value-at-risk (VaR) measures.
- **Asset pricing:** MGARCH models can be used to study the time-varying risk premia associated with different sources of systematic risk.

4.5 Conclusion

Multivariate GARCH models represent an important class of econometric models for analyzing the dynamic behavior of asset return volatilities and covariances. Despite challenges related to the high dimensionality of the parameter space, models such as the BEKK and DCC have made practical applications feasible. Ongoing research continues to explore more parsimonious structures and improved estimation techniques to better capture the complex dynamics observed in financial markets.

Chapter 5

Empirical exercises

5.1 Realised Volatility (RV)

Realised volatility is computed as:

$$RV_t = \sum_{i=1}^{n} r_{t,i}^2 \tag{5.1}$$

where $r_{t,i}$ are intraday returns.

Chapter 6

Innovations

6.1 Wave methods

Impulse response functions (IRFs) are a fundamental tool in time series analysis and econometrics for tracing the effects of a one-time shock on a dynamic system. Traditionally developed in the context of autoregressive (AR) or vector autoregressive (VAR) models, IRFs describe the evolution of the system following an impulse (or shock).

In parallel, the concept of *network wave theory* has emerged to capture how shocks propagate not only in time but also across nodes in a network. In behavioral econometrics, agents (such as consumers, firms, or policy makers) are interconnected through social or economic ties, and shocks—such as changes in sentiment, policy announcements, or external news—can diffuse through these networks much like waves.

This chapter synthesizes these ideas. In Section 6.2 we review the theory and computation of IRFs in simple dynamic models. Section 6.3 introduces network wave theory and extends the standard IRF framework to incorporate spatial (or network) dimensions. Section 6.4 discusses applications in behavioral econometrics, and Section **??** concludes.

6.2 Impulse Response Functions: Theory and Estimation

Consider first a univariate AR(1) process:

$$x_t = \phi x_{t-1} + u_t, \quad |\phi| < 1, \tag{6.1}$$

where u_t is a white noise disturbance. Since the process is stationary, it can be rewritten in its infinite moving average representation:

$$x_t = \sum_{j=0}^{\infty} \phi^j u_{t-j} = u_t + \phi u_{t-1} + \phi^2 u_{t-2} + \cdots . \tag{6.2}$$

The impulse response function (IRF) in this case is given by:

$$\frac{\partial x_{t+j}}{\partial u_t} = \phi^j, \quad j = 0, 1, 2, \ldots . \tag{6.3}$$

Now, consider a vector autoregressive (VAR) system:

$$x_t = A_1 x_{t-1} + u_t, \quad x_t \in \mathrm{R}^n, \tag{6.4}$$

where the innovations u_t can be decomposed as:

$$u_t = B\varepsilon_t, \tag{6.5}$$

with B being an $n \times n$ full-rank matrix and ε_t a vector of orthogonal structural shocks. The moving average representation is then:

$$x_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j}, \quad \text{with } C_j = A_1^j B. \tag{6.6}$$

Thus, the IRF is the sequence $\{C_j\}_{j\geq 0}$, where the (i,k) element of C_j represents the response of variable i at time $t+j$ to a shock in structural disturbance $\varepsilon_{t,k}$.

6.3 Network Wave Theory

Traditional IRFs capture the temporal evolution following a shock, but in many behavioral settings agents are interconnected in a network, and shocks may diffuse spatially. In network wave theory, the dynamic process is augmented with an explicit network structure.

Consider a simple network autoregressive model:

$$x_t = \phi x_{t-1} + \theta W x_{t-1} + \varepsilon_t, \tag{6.7}$$

where:

- x_t is an $n \times 1$ vector of outcomes across n agents,
- W is an $n \times n$ matrix representing the connectivity or influence weights among agents,
- θ is a scalar capturing the strength of the network (spatial) dependence,
- ε_t is an $n \times 1$ vector of shocks.

The moving average representation of Equation (6.7) is:

$$x_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}, \tag{6.8}$$

where the matrices Ψ_j capture both temporal and network propagation. In particular, the (i, k) entry of Ψ_j gives the response of agent i at time $t + j$ to a shock in agent k at time t.

A central insight is that the effect of a shock decays with both time and the network distance. If $d(i, k)$ denotes the (shortest path) distance between agents i and k in the network, then typically the response decreases as $d(i, k)$ increases.

6.4 Applications in Behavioral Econometrics

Behavioral econometrics focuses on modeling the decisions and interactions of economic agents. Often, these agents are embedded in networks where peer effects and spillovers are significant. For example:

- **Consumer Behavior:** A viral marketing campaign might shock a subgroup of consumers, and the effect can propagate through social networks.
- **Financial Contagion:** Behavioral biases can lead to herding, where a shock in one market spills over into others.
- **Policy Interventions:** A regulatory change may have both direct impacts on certain agents and indirect effects on their peers.

In such settings, incorporating network wave dynamics allows researchers to trace both the temporal and spatial (network) diffusion of shocks. For instance, one can estimate a network VAR and compute the IRFs as functions of time and network distance:

$$\Psi_j = (\phi I + \theta W)^j B. \tag{6.9}$$

The entry $(\Psi_j)_{ik}$ shows how a one-unit shock to agent k affects agent i after j periods.

6.5 Empirical Illustration

Suppose we have data on consumers' spending x_{it} observed over time, and a network matrix W that represents their social interactions. A simple network VAR could be specified as:

$$x_t = \phi x_{t-1} + \theta W x_{t-1} + \varepsilon_t, \tag{6.10}$$

where x_t aggregates the spending of all consumers at time t. After estimating parameters ϕ and θ (using appropriate econometric techniques), the IRF is given by:

$$\Psi_j = (\phi I + \theta W)^j B. \tag{6.11}$$

This IRF describes how a shock to one consumer propagates through the network over time and across agents. The decay with time and network distance helps quantify both direct and indirect effects, providing valuable insights into behavioral contagion and spillover effects.

6.6 Conclusion

In this chapter we have merged the classical concept of impulse response functions with the novel framework of network wave theory. Traditional IRFs capture the evolution of a system over time following a shock, while network wave theory extends this to incorporate the diffusion of shocks across interconnected agents. This combined approach enhances behavioral econometric models by enabling a detailed analysis of both temporal and cross-sectional responses to shocks. Future work may explore nonlinearities and heterogeneous network structures to further enrich this framework.

Chapter 7

Ultimatum

In case any of the content was not clear, kindly contact g@r.ch

Bibliography

[1] Bauer, E., Hasselmann, K., Young, I. R., & Hasselmann, S. (1996). Assimilation of wave data into the wave model WAM using an impulse response function method. *Journal of Geophysical Research: Atmospheres*, **101**, 3801–3816. doi:10.1029/95jc03306.

[2] Aït-Sahalia, Y. & Jacod, J. (2014). *High-frequency Financial Econometrics*. Princeton University Press.

[3] Campbell, J. Y., Lo, A. W., & Mackinlay, A. C. (2012). *The Econometrics of Financial Markets*. Princeton University Press.

[4] Hautsch, N. (2012). *Econometrics of Financial High-Frequency Data*. Springer Berlin Heidelberg.

[5] Taylor, S. J. (2013). *Modelling Financial Time Series*. World Scientific.