

# **VAR 101 – A Toolbox for Multivariate Time Series Modelling**

**03/06/2025**

# Why Multivariate Models?

- Real-world systems: many variables evolve together.
- Example: in finance, order flow and returns are intertwined.
- Need to model interactions and feedback loops across time.

# What Is a VAR Model?

- A Vector AutoRegression model of order p:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \epsilon_t$$

- $Y_t$ : a vector of k time series
- $A_i$ : a matrix of k x k coefficients
- $\epsilon_t$ : White noise

# When to Use a VAR?

- When all variables are:
  - Endogenous
  - Stationary
- Examples:
  - Macroeconomic indicators
  - Asset returns and volatilities
  - Trade, quote, and order flow variables

# VAR Workflow Overview

- Data preparation
- Stationarity checks (ADF test, differencing)
- Lag selection (AIC, BIC, HQIC)
- Model estimation
- Diagnostics
- Forecasting or impulse response analysis

# Stationarity & Preprocessing

- Time series must be stationary
- Check with:
  - Augmented Dickey-Fuller (ADF) test
  - KPSS test
- Transform if needed: differencing, log transform, detrending

# Choosing the Lag Order

- Use Information Criteria:
  - AIC: Akaike Information Criterion
  - BIC: Bayesian Information Criterion
  - HQIC: Hannan-Quinn
- Trade-off: fit vs complexity

# Model Diagnostics

- Check:
  - Residuals  $\rightarrow$  white noise?
  - Stability  $\rightarrow$  eigenvalues inside unit circle?
  - Autocorrelation  $\rightarrow$  Ljung-Box test
- Plot: residuals, ACF/PACF, impulse responses