universität
innsbruck

Time Series Analysis

Multivariate Time Series

Multivariate Time Series

Contents

Contents

Multivariate Time Series

Motivation

Motivation

- Many economic relationships are multivariate.
- As a rule: *multivariate TSA* $\approx$ *univariate TSA* + *matrix algebra*.
- This chapter is confined to linear models for multivariate time series. Multivariate nonlinear models (e.g., ARCH-type) also exist.
- There exist multivariate versions of ARMA models, called VARMA models. Special cases are VAR and VMA models.
- In time series econometrics, vector autoregressions (VARs) constitute a large industry, motivated by empirical macroeconomics. Historically important paper: Sims (1980), Macroeconomics and Reality, *Econometrica*.
- Minority opinion of Harvey (1997), Trends, Cycles and Autoregressions, *Economic Journal*: *To many econometricians, VAR stands for 'very awful regression'.*

Motivation

- Multivariate (linear) time series analysis also important in electrical engineering or systems engineering. Often called *linear dynamical systems* there. Terminology differs from econometrics.
- Algebra for multivariate time series often makes use of special matrix operations, such as Kronecker products or the ‘vec’ and ‘vech’ operators. Avoided here.
- Software much more limited than for univariate time series. VARs are routinely available in econometric software, VARMA models are not so widely available.

Motivation

Issues: What is similar to univariate case, what is new?

- Basic building block was white noise process. What does change in multivariate setting?
- What are stationarity conditions in the multivariate case?
- What are model diagnostics in the multivariate case?
- What are new issues in the multivariate case?
  - Effects of shocks in one series on other series.
  - One series may help to forecast other series.

Multivariate Time Series

Basic Concepts

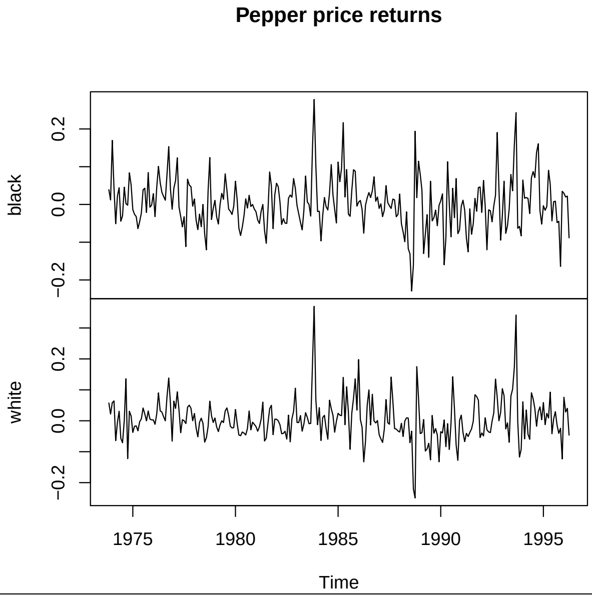
Basic concepts

Data: `PepperPrice`. Average monthly European spot prices for black and white pepper (USD/ton) from 1973(10) to 1996(4).

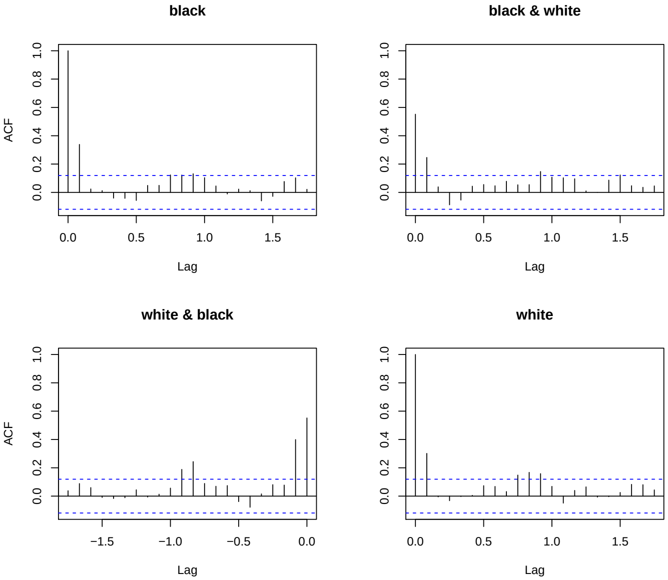
Illustration:

- Bivariate series of log-returns.
- Autocorrelation function of black and white pepper, respectively.
- Crosscorrelations of black and white pepper and vice versa.

Basic concepts



Basic concepts



Basic concepts

Notation: Consider vector-valued process $\{Y_t\}$, with typical element

$$Y_t = \begin{pmatrix} Y_{1t} \\ \vdots \\ Y_{kt} \end{pmatrix}.$$

As in the univariate case: Consider first two moments. For elements i, j, these are

$$\begin{aligned} \mathrm{E}(Y_{it}) &= \mu_{it}, \\ \gamma_{ij}(t, t+h) &= \mathrm{E}[(Y_{it} - \mu_{it})(Y_{j,t+h} - \mu_{j,t+h})]. \end{aligned}$$

Basic concepts

Matrix notation: Collect all first two moments in vectors and matrices.

$$
\begin{aligned}
\mathrm{E}(Y_t) &= \mu_t = \begin{pmatrix} \mu_{1t} \\ \vdots \\ \mu_{kt} \end{pmatrix}, \\
\Gamma(t, t+h) &= \begin{pmatrix} \gamma_{11}(t, t+h) & \cdots & \gamma_{1k}(t, t+h) \\ \vdots & \ddots & \vdots \\ \gamma_{k1}(t, t+h) & \cdots & \gamma_{kk}(t, t+h) \end{pmatrix}.
\end{aligned}
$$

Warning: Since $\gamma_{ij}(t, t+h) \neq \gamma_{ji}(t, t+h)$, the matrix $\Gamma(t, t+h)$ is in general not symmetric.

Basic concepts

Stationarity: Weak stationarity if first two moments are time-invariant.

$$\mathrm{E}(Y_t) \;=\; \mu, \quad \Gamma(t, t+h) \;=\; \Gamma(h) \;=\; [\gamma_{ij}(h)]_{i,j=1,\ldots,k}$$

Properties: For a weakly stationary process,

- $\Gamma(h) \;=\; \Gamma(-h)^\top$
- $|\gamma_{ij}(h)| \;\leq\; \sqrt{\gamma_{ii}(0)\gamma_{jj}(0)}$

Autocorrelation matrices: Define

$$R(h) \;=\; D^{-1/2}\Gamma(h)D^{-1/2}, \quad D \;=\; \mathrm{diag}(\gamma_{11}(0), \ldots, \gamma_{kk}(0))$$

Hence have matrix-valued ACF. By construction, $R(h) = R(-h)^\top$.
Elements $\varrho_{ij}(h)$ are called *cross correlation functions* (CCFs).

Warning: CCF not easy to interpret if components are not white noise! Hence, potentially *prewhiten* univariate series.

Basic concepts

Definition: A process $\{e_t\}$ is a (multivariate) weak white noise, if $\mathrm{E}(e_t) = \mu$ for all t, $\mathrm{Cov}(e_t, e_{t-h}) = 0$ for all $h \neq 0$ and $\mathrm{Cov}(e_t) = \Sigma_e$.

Autocovariance function:

$$\Gamma_e(h) = \begin{cases} \Sigma_e, & h = 0, \\ 0, & h \neq 0. \end{cases}$$

Note: If Σ_e is not diagonal, then there is *contemporaneous correlation* between components of e_t – a new phenomenon.

In the following: $\{e_t\}$ denotes multivariate white noise with mean $\mu = 0$ and covariance matrix Σ_e.

Basic concepts

Linear process: $\{Y_t\}$ with representation

$$Y_t \;=\; \sum_{j=-\infty}^{\infty} \Psi_j e_{t-j} \;=\; \Psi(B) e_t$$

with $\sum_{j=-\infty}^{\infty} |\psi_j(i,\ell)| < \infty$, for all $i, \ell = 1, 2, \ldots, k$.

This implies that autocovariances are

$$\Gamma(h) \;=\; \sum_{j=-\infty}^{\infty} \Psi_j \Sigma_e \Psi_{j-h}^{\top}.$$

Alternative representations:

- MA(∞) has $\Psi_j = 0$ for $j < 0$: $Y_t \;=\; \sum_{j=0}^{\infty} \Psi_j e_{t-j}$.
- AR(∞) is such that: $Y_t - \sum_{j=1}^{\infty} \Phi_j Y_{t-j} \;=\; e_t$.

Basic concepts

Estimation: Estimate theoretical μ and autocovariances $\Gamma(h)$ by empirical counterparts.

$$
\begin{aligned}
\hat{\mu} &= \bar{Y}_t = \frac{1}{T}\sum_{t=1}^{T} Y_t \\
\hat{\Gamma}(h) &= \frac{1}{T}\sum_{t=h+1}^{T} (Y_t - \bar{Y})(Y_{t-h} - \bar{Y})^\top \quad (0 \le h \le T-1).
\end{aligned}
$$

By symmetry, use $\hat{\Gamma}(h) = \hat{\Gamma}(-h)^\top$ for $-T+1 \le h < 0$.

Similarly:

$$
\begin{aligned}
\hat{R}(h) &= \hat{D}^{-1/2}\hat{\Gamma}(h)\hat{D}^{-1/2}, \\
\hat{D} &= \mathrm{diag}(\hat{\gamma}_{11}(0), \dots, \hat{\gamma}_{kk}(0)).
\end{aligned}
$$

Basic concepts

Test: Global test for white noise (compare Box-Pierce/Ljung-Box) employs

$$Q_k(m) \;=\; T^2 \sum_{h=1}^{m} \frac{1}{T-h}\, \mathrm{tr}\left[\hat{\Gamma}(h)\hat{\Gamma}(0)^{-1}\hat{\Gamma}(h)\hat{\Gamma}(0)^{-1}\right]$$

which has an asymptotic $\chi^2(k^2 m)$ distribution under the null hypothesis of white noise.

Note: Some authors express $Q_k(m)$ in terms of the Kronecker product $\otimes$.

Multivariate Time Series

Vector ARMA Models

Vector ARMA models

Definition: $\{Y_t\}$ is a vector ARMA (VARMA) process if stationary with

$$\begin{aligned} Y_t - \Phi_1 Y_{t-1} - \cdots - \Phi_p Y_{t-p} &= e_t - \Theta_1 e_{t-1} - \cdots - \Theta_q e_{t-q}, \\ \Phi(B) Y_t &= \Theta(B) e_t,. \end{aligned}$$

Additionally, a nonzero mean μ can be included. The characteristic polynomials are matrix-valued, e.g., $\Phi(B) = I_k - \Phi_1 B - \cdots - \Phi_p B^p$.

Vector ARMA models

Example: VAR(1) with $\mathrm{E}(Y_t) = 0$ is

$$Y_t \;=\; \Phi_1 Y_{t-1} + e_t,$$

Repeated substitution gives MA(∞) representation

$$Y_t \;=\; \sum_{j=0}^{\infty} \Phi_1^j Y_{t-j} \;=\; \sum_{j=0}^{\infty} \Psi_j e_{t-j}$$

This is valid if all eigenvalues λ_j of Φ_1 satisfy $|\lambda_j| < 1$.

Vector ARMA models

Stationarity: A VARMA process $\{Y_t\}$ is *stationary* (and causal) if there exist matrices Ψ_j such that

$$Y_t \;=\; \sum_{j=0}^{\infty} \Psi_j e_{t-j}.$$

Condition: (Sometimes called *stability*.)

$$\det(\Phi(z)) \neq 0 \quad \text{for all } z \in \mathbb{C} \text{ with } |z| \leq 1$$

Coefficients Ψ_j can be found recursively as in the univariate case:

$$\Psi_j = -\Theta_j + \sum_{k=1}^{\infty} \Phi_k \Psi_{j-k}, \quad j = 0, 1, \ldots$$

with $\Theta_0 = I_k$, $\Theta_j = 0$ for $j > q$ and all $j < 0$, and $\Phi_j = 0$ for $j > p$.

Vector ARMA models

Invertibility: A VARMA process $\{Y_t\}$ is *invertible* if there exist matrices Π_j such that

$$e_t \;=\; \sum_{j=0}^{\infty} \Pi_j Y_{t-j}.$$

Condition:

$$\det(\Theta(z)) \neq 0 \quad \text{for all } z \in \mathbb{C} \text{ with } |z| \leq 1$$

Coefficients Π_j can be found recursively as in the univariate case:

$$\Pi_j \;=\; -\Phi_j + \sum_{k=1}^{\infty} \Theta_k \Pi_{j-k}, \quad j = 0, 1, \ldots$$

with $\Phi_0 = -I_k$, $\Phi_j = 0$ for $j > p$, $\Theta_j = 0$ for $j > q$, and $\Pi_j = 0$ for $j < 0$.

Vector ARMA models

Identifiability: Consider VAR(1) with

$$\Phi_1 = \begin{pmatrix} 0 & \phi_{12} \\ 0 & 0 \end{pmatrix}$$

Its MA(∞) representation is

$$Y_t = \sum_{j=0}^{\infty} \Phi_1^j e_{t-j} = e_t + \Phi_1 e_{t-1}$$

Hence this VAR(1) is also a VMA(1) process!

Vector ARMA models

Remarks:

- Note that this object is stationary and invertible, yet it has both a finite-order MA and a finite-order AR representation.
- Lesson: Identification more complex in the multivariate case.
- Practical solution, especially in economics: Employ only VARs.
- Advantage: Estimation much easier for pure VARs, as long as there are no restrictions on parameters.

Vector ARMA models

Marginal models: Only for VAR models. Suppose $\{Y_t\}$ is a VAR process. What can be said about $\{Y_{it}\}$?

Recall: Inverse A^{-1} of a matrix A is given by

$$A^{-1} = \frac{1}{\det(A)} A^{\#}$$

with $A^{\#}$ the adjoint matrix (transpose of complex conjugates).
Thus: For a stationary VAR use VMA representation

$$\begin{aligned} \Phi(B)Y_t &= e_t \\ Y_t &= \Phi(B)^{-1}e_t \\ &= \frac{1}{\det(\Phi(B))}\Phi(B)^{\#}e_t. \end{aligned}$$

This is a matrix of rational functions and hence the marginal models of VARs are ARMA. Further justification of ARMAs in the univariate case.

Multivariate Time Series

Vector Autoregressions

Vector autoregressions

Definition: VAR(p) is given by

$$Y_t \;=\; \mu + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + e_t.$$

Remarks:

- Note that this has $p \cdot k^2$ autoregression parameters (excluding μ and Σ_e) – "curse of dimensionality". Hence VARs are only useful for moderate k.
- The *intercept* is denoted μ here (rather than using a vector of θ_0 as in the univariate case).
- Above form is called *reduced form* of the VAR – only lags on right-hand side, no contemporaneous relations.
- Also called *levels form* of the VAR – if there are cointegrating relations, there exists a further form that is neither levels nor differences, the *error correction form*.

Vector autoregressions

Higher-order VAR models: Rewrite

$$Y_t \;=\; \mu + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + e_t,$$

as a VAR(1) in the form

$$\tilde{Y}_t = \tilde{\mu} + \tilde{\Phi}\tilde{Y}_{t-1} + \tilde{e}_t$$

$$\begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix} = \begin{pmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\ I_k & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_k & 0 \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Y_{t-2} \\ \vdots \\ Y_{t-p} \end{pmatrix} + \begin{pmatrix} e_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Jargon: This is called the *companion form* of a higher-order VAR. $\tilde{\Phi}$ is called the *companion matrix*.

Vector autoregressions

Remark: This ‘trick’ has nothing to do with statistics or econometrics, it works for dynamical systems in general, e.g., also for differential equations.

Advantage: VAR(p) derivations can be reduced to VAR(1) derivations.

Example: Stationarity condition for VAR(p)

$$\det(\Phi(z)) \neq 0 \quad \text{for all } z \in \mathbb{C} \text{ with } |z| \leq 1$$

is satisfied if all eigenvalues λ_j of the companion matrix $\tilde{\Phi}$ are $|\lambda_j| < 1$.

Vector autoregressions

Estimation: Write

$$\begin{aligned} Y_t &= \mu + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + e_t \\ &= [\mu, \Phi_1, \Phi_2, \ldots, \Phi_p]\, Z_{t-1} + e_t, \end{aligned}$$

with $Z_{t-1} = (1_k, Y_{t-1}^\top, Y_{t-2}^\top, \ldots, Y_{t-p}^\top)^\top$.

Vector autoregressions

Remarks:

- *Multivariate* linear regression model.
- Conditional least squares using p starting values $Y_{-p+1}, \ldots, Y_0$

$$[\hat{\mu}, \hat{\Phi}_1, \hat{\Phi}_2, \ldots, \hat{\Phi}_p] = \left(\sum_{t=1}^{T} Y_t Z_{t-1}^\top \right) \left(\sum_{t=1}^{T} Z_t Z_{t-1}^\top \right)^{-1} .$$

- Identical to conditional ML under normality.
- Can be estimated separately for each equation $j = 1, \ldots, k$.
- Essentially, a VAR is a time series version of the seemingly unrelated regression (SUR) model.

Vector autoregressions

Properties: For j-th equation of a stationary VAR(p)

$$Y_j \;=\; Z\pi_j + e_j$$

where

- Y_j collects all T observations on equation j.
- Z is a $T \times \ell$ matrix with $\ell = kp + 1$ with the t-th row $Z_{t-1}^\top$ as above.
- π_i are the ℓ coefficients pertaining to equation j.

Central limit theorem: Writing $\beta = \text{vec}(\Pi) = \text{vec}(\pi_1, \pi_2, \ldots, \pi_p)$, it can be shown that, under some technical assumptions,

$$\sqrt{T}(\hat{\beta} - \beta) \quad \xrightarrow{d} \quad \mathcal{N}_{pk^2+k}(0, \Sigma_e \otimes \Gamma_y^{-1})$$

with $\frac{1}{T} Z^\top Z \xrightarrow{p} \Gamma_y$.

Vector autoregressions

Message: Estimates are approximately normally distributed, so *t*/*F* tests are available almost "as usual".

Remark: In the nonstationary case, the distribution depends on the presence of cointegrating relationships. If there are some, then *t*/*F* tests may or may not follow standard distributions, depending on lag length etc.

Model selection: As in the univariate case, use information criteria. The number of parameters is pk^2 for the autoregression coefficients and k for the constants μ. Additionally, the parameters for Σ_e may be counted (but their number does not depend on specification of the mean equation).

Vector autoregressions

In R:

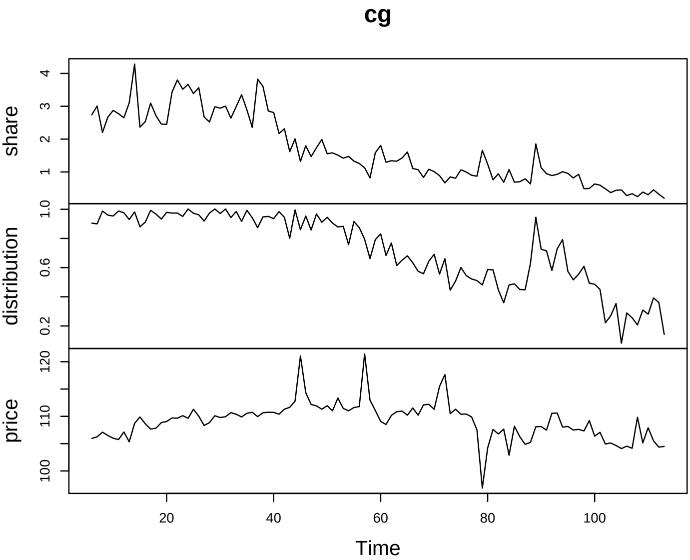
- `ar()` can fit (stationary) VAR(p) models by conditional OLS. Order p can be selected automatically via AIC.
- Package *vars* offers VAR(p) models as well as more elaborate models. `VAR()` estimates VAR(p) models and `VARselect()` supports various information criteria for selection of p. See Pfaff (2008, *Journal of Statistical Software* 27, 4).
- Package *dse* implements dynamic systems estimation, including VAR(p) models in function `estVARXls()` as well as more general VARIMA models.

Vector autoregressions

Example: Marketing data. Trivariate weekly time series of distribution, market share and price of a fast-moving consumer good.

```
R> data("ConsumerGood", package = "AER")
R> cg <- ConsumerGood[, c(2, 1, 3)]
```

Vector autoregressions



Vector autoregressions

Selection of order $p = 1, \dots, 5$:

```
R> library("vars")
R> VARselect(cg, 5)
```

```
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     3      1      1      3

$criteria
                1         2         3         4         5
AIC(n) -4.743504 -4.752737 -4.821923 -4.733922 -4.628702
HQ(n)  -4.619175 -4.535162 -4.511101 -4.329853 -4.131387
SC(n)  -4.436545 -4.215559 -4.054526 -3.736306 -3.400867
FPE(n)  0.008709  0.008633  0.008066  0.008828  0.009842
```

Estimate optimal BIC model: VAR(1).

```
R> cg_var1 <- VAR(cg, p = 1)
```

Vector autoregressions

```
R> coef(cg_var1)
$share
                 Estimate Std. Error t value  Pr(>|t|)
share.l1          0.76177    0.08043  9.4706 1.113e-15
distribution.l1   0.80547    0.37403  2.1535 3.361e-02
price.l1         -0.00985    0.01490 -0.6611 5.100e-01
const             0.86746    1.54392  0.5619 5.754e-01

$distribution
                  Estimate Std. Error t value  Pr(>|t|)
share.l1          0.046645   0.018158  2.5689 1.164e-02
distribution.l1   0.755778   0.084434  8.9511 1.580e-14
price.l1          0.003002   0.003364  0.8925 3.742e-01
const            -0.235803   0.348531 -0.6766 5.002e-01

$price
                 Estimate Std. Error t value  Pr(>|t|)
share.l1          -0.5718    0.45314  -1.262 2.099e-01
distribution.l1    5.3788    2.10713   2.553 1.216e-02
price.l1           0.5428    0.08394   6.466 3.417e-09
const             46.9599    8.69792   5.399 4.320e-07
```

Vector autoregressions

Eigenvalues of the companion matrix:

```
R> roots(cg_var1)
[1] 0.9551 0.6973 0.4080
```

Ljung-Box type test for residual autocorrelation:

```
R> serial.test(cg_var1, lags.pt = 10, type = "PT.adjusted")
        Portmanteau Test (adjusted)

data:  Residuals of VAR object cg_var1
Chi-squared = 98, df = 81, p-value = 0.09
```

Further diagnostic tests:

```
R> arch.test(cg_var1)
R> normality.test(cg_var1)
```

Vector autoregressions

Forecasting: As in the univariate case, VAR forecasts satisfy:

$$\hat{Y}_t(\ell) \;=\; \mu \;+\; \Phi_1 \hat{Y}_t(\ell-1) \;+\; \Phi_2 \hat{Y}_t(\ell-2) \;+\; \ldots \;+\; \Phi_p \hat{Y}_t(\ell-p)$$

with $\hat{Y}_t(\ell) = Y_t$ for $\ell \leq 0$. In practice, Φ_j have to be estimated.

The ℓ-step-ahead forecast errors follow MA processes:

$$Y_{t+\ell} - \hat{Y}_t(\ell) \;=\; \sum_{j=0}^{\ell-1} \Psi_j e_{t+\ell-j}$$

Example: For a VAR(1) with zero mean, $Y_t = \Phi Y_{t-1} + e_t$,

$$\hat{Y}_t(\ell) \;=\; \Phi \hat{Y}_t(\ell-1) \;=\; \ldots \;=\; \Phi^{\ell-1} \hat{Y}_t(1) \;=\; \Phi^{\ell} Y_t$$

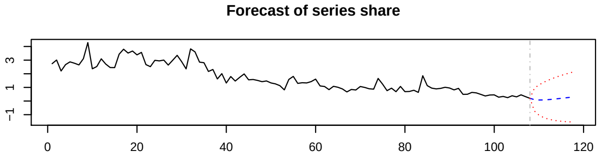
Therefore, the ℓ-step-ahead forecast error is

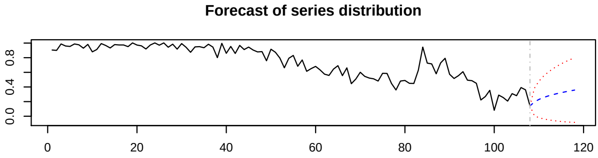
$$Y_{t+\ell} - \hat{Y}_t(\ell) \;=\; \sum_{j=0}^{\ell-1} \Phi^j e_{t+\ell-j}$$

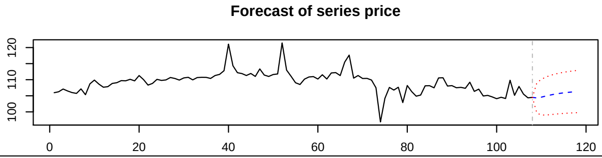
Vector autoregressions

```
R> predict(cg_var1, n.ahead = 3)
$share
        fcst   lower  upper     CI
[1,] 0.10405 -0.7337 0.9418 0.8377
[2,] 0.07569 -1.0295 1.1809 1.1052
[3,] 0.07965 -1.2022 1.3615 1.2819

$distribution
       fcst     lower  upper     CI
[1,] 0.1945  0.005379 0.3836 0.1891
[2,] 0.2292 -0.019684 0.4782 0.2489
[3,] 0.2549 -0.035384 0.5452 0.2903

$price
      fcst lower upper    CI
[1,] 104.3 99.61 109.1 4.720
[2,] 104.6 99.03 110.1 5.546
[3,] 104.9 99.01 110.8 5.907
R> plot(predict(cg_var1))
R> fanchart(predict(cg_var1))
```

Vector autoregressions

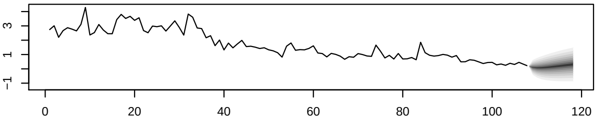




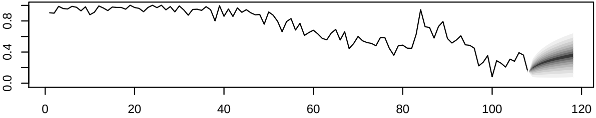


Vector autoregressions

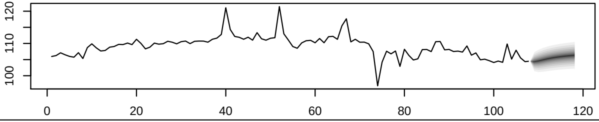
Fanchart for variable share



Fanchart for variable distribution



Fanchart for variable price



Multivariate Time Series

Interpreting VARs

Interpreting VARs

Interpretation: Three approaches for VAR results.

1. Granger ‘causality’ analysis (Granger, 1969, *Econometrica*).
2. Impulse response functions (IRFs).
3. Forecast error variance decomposition (FEVD).

Interpreting VARs

Granger 'causality' analysis: Partition $Y_t = (Y_t^{(1)\top}, Y_t^{(2)\top})^\top$, with dimensions $k_1 + k_2 = k$.

$$\begin{pmatrix} Y_t^{(1)} \\ Y_t^{(2)} \end{pmatrix} = \begin{pmatrix} \mu^{(1)} \\ \mu^{(2)} \end{pmatrix} + \sum_{j=1}^{p} \begin{pmatrix} \Phi_j^{(11)} & \Phi_j^{(12)} \\ \Phi_j^{(21)} & \Phi_j^{(22)} \end{pmatrix} \begin{pmatrix} Y_{t-j}^{(1)} \\ Y_{t-j}^{(2)} \end{pmatrix} + \begin{pmatrix} e_t^{(1)} \\ e_t^{(2)} \end{pmatrix}$$

Consider the information sets

$$\begin{aligned} \mathcal{I}_t &= \{Y_t, Y_{t-1}, Y_{t-2}, \dots\} \\ \mathcal{I}_t^{(1)} &= \{Y_t^{(1)}, Y_{t-1}^{(1)}, Y_{t-2}^{(1)}, \dots\} \end{aligned}$$

Definition: $Y_t^{(2)}$ does not Granger-cause $Y_t^{(1)}$, if

$$\mathsf{E}(Y_t \mid \mathcal{I}_{t-1}) = \mathsf{E}(Y_t \mid \mathcal{I}_{t-1}^{(1)})$$

Vector autoregressions

Thus: Given the information in lagged $Y_t^{(1)}$, lagged $Y_t^{(2)}$ does not help to forecast Y_t.

In a (linear) VAR, this means

$$\Phi_1^{(12)} = \Phi_2^{(12)} = \ldots = \Phi_p^{(12)} = 0$$

These are linear restrictions and can be assessed using an F/Wald test.

Remarks:

- Unfortunate terminology: Forecastability, not causality.
- If $Y_t^{(2)}$ does not Granger-cause $Y_t^{(1)}$, then a VAR for Y_t does not require lagged $Y_t^{(2)}$.
- However, $Y_t^{(2)}$ could still be useful to explain e.g., the conditional variance.

Vector autoregressions

Example: Granger causality of distribution and price for market share?

```
R> causality(cg_var1, cause = c("distribution", "price"))$Granger
        Granger causality H0: distribution price do not
        Granger-cause share

data:  VAR object cg_var1
F-Test = 2.5, df1 = 2, df2 = 310, p-value = 0.08
```

Interpreting VARs

Impulse response functions:

- How are shocks transmitted through the system?
- What is the effect of a unit (= standardized) shock in series j on itself and on other series?

Interpreting VARs

MA(∞) representation is

$$Y_t \;=\; \Phi(B)^{-1} e_t \;=\; \Psi(B) e_t \;=\; \sum_{j=0}^{\infty} \Psi_j e_t$$

Marginal (expected) response of $Y_{i,t+h}$ to a unit impulse in series j is

$$\frac{\partial \mathrm{E}(Y_{i,t+h} \mid \mathcal{I}_t)}{\partial e_{j,t}} \;=\; \Psi_{h,ij}$$

Mapping $h \mapsto \Psi_{h,ij}$ is called the *impulse response function* of j on i.

$\Psi_{h,ij}$ measures dynamic effect of shock in j on i after h periods. Hence also called *h-period dynamic multiplier*. In total, k^2 IRFs for k series.

Interpreting VARs

Problem: If Σ_e is not diagonal, the shocks are correlated.

'Solution': For a covariance matrix Σ_e (symmetric and positive definite), the Cholesky decomposition is $\Sigma_e = PP^\top$, P lower triangular. Now use orthogonalized shock $v_t = P^{-1}e_t$. This gives

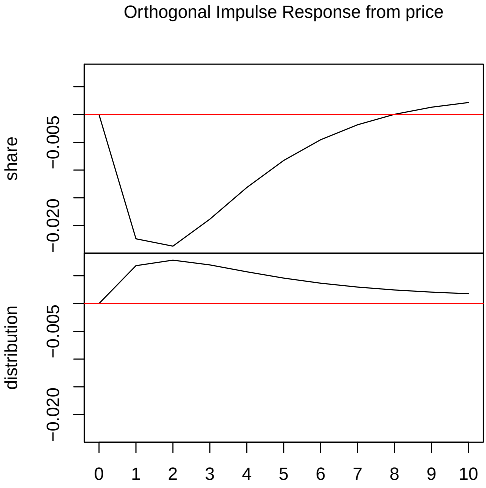
$$Y_t = \sum_{j=0}^{\infty} \Psi_j e_t = \sum_{j=0}^{\infty} (\Psi_j P)\left(P^{-1} e_t\right) = \sum_{j=0}^{\infty} \Psi_j^O v_t$$

So $\Psi_j^O = \Psi_j P$. Note that $\Psi_0^O = \Psi_0 P$ is lower triangular.

Call $h \mapsto \Psi_{h,ij}^O$ an *orthogonalized IRF*.

New problem: Cholesky decomposition is unique (up to sign of P) only for a given ordering of the variables. But different orderings lead to different IRFs. Hence shape of IRFs depends on 'identifying assumptions' $\rightarrow$ structural VAR analysis.

Interpreting VARs



Interpreting VARs

Forecast error variance decomposition: Which part of the forecast error variance is caused by which variable?

Using orthogonalized shocks v_t with $\Sigma_v = I_k$,

$$Y_{t+h} - \hat{Y}_t(h) \;=\; \sum_{j=0}^{h-1} \Psi_j^O v_{t+h-j}$$

forecast error variance for forecast horizon h is

$$\Sigma(h) \;=\; \mathrm{Var}(Y_{t+h} - \hat{Y}_t(h)) \;=\; \sum_{j=0}^{h-1} \Psi_j^O \Psi_j^{O\top}$$

Hence forecast error variance of i-th component is

$$\sigma_i^2(h) \;=\; [\Sigma(h)]_{ii} \;=\; \left[\sum_{j=0}^{h-1} \Psi_j^O \Psi_j^{O\top}\right]_{ii} \;=\; \sum_{j=0}^{h-1} \left[\Psi_j^O \Psi_j^{O\top}\right]_{ii}$$

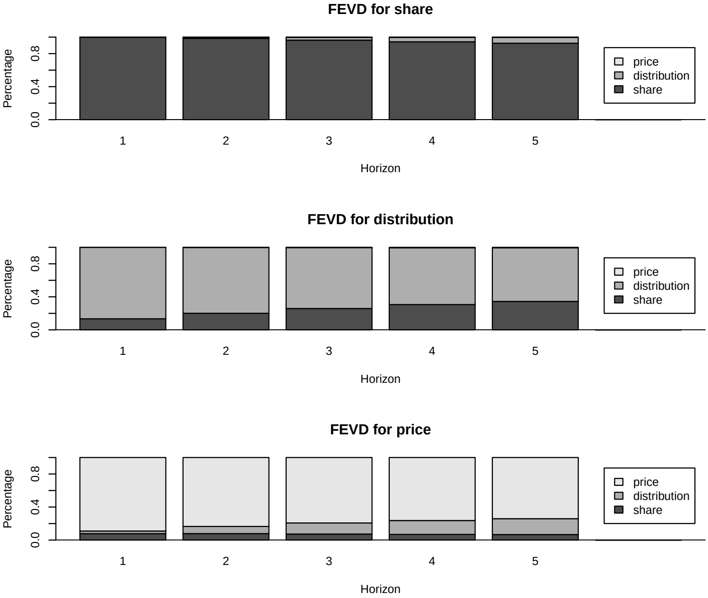
Interpreting VARs

In practice: Report relative contribution of *j*-th shock.

$$(\Psi^{2\,O}_{0,ij} + \cdots + \Psi^{2\,O}_{h-1,ij})/\sigma_i^2(h)$$

```
R> fevd(cg_var1, n.ahead = 2)
$share
      share distribution    price
[1,] 1.0000      0.00000 0.000000
[2,] 0.9838      0.01458 0.001575

$distribution
      share distribution    price
[1,] 0.1344       0.8656 0.000000
[2,] 0.2012       0.7959 0.002884

$price
       share distribution  price
[1,] 0.07799      0.03171 0.8903
[2,] 0.07841      0.08692 0.8347
```

Interpreting VARs



Multivariate Time Series

Structural VARs

Structural VARs

Note: VAR only allows *lagged* regressors and no contemporaneous terms. Economic theory might suggest contemporaneous terms.

Distinguish: *Reduced form* (VAR)

$$Y_t \;=\; \mu + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + e_t,$$

and *structural form* (structural VAR, SVAR)

$$\begin{aligned} AY_t \;&=\; A\mu \;+\; A\Phi_1 Y_{t-1} \;+\; A\Phi_2 Y_{t-2} \;+\; \cdots + A\Phi_p Y_{t-p} \;+\; Ae_t \\ &=\; \mu^* \;+\; \Phi_1^* Y_{t-1} \;+\; \Phi_2^* Y_{t-2} \;+\; \cdots + \Phi_p^* Y_{t-p} \;+\; v_t. \end{aligned}$$

Hence $\Sigma_v = A\,\Sigma_e\,A^\top$.

Note: Multiplication with any nonsingular B also yields structural form. Hence parameters are not identified without further restrictions. Sometimes $B = A^{-1}$ is of special interest.

Structural VARs

Approach: Structural VAR analysis imposes restrictions on A, B and/or Σ_v.

Jargon: Depending on form of restrictions, there are so-called A, B, and AB models.

Classical example: Blanchard and Quah (1989, *American Economic Review*) consider bivariate SVAR for output growth and unemployment rate.