

Analysis of Multiple Time Series

Kevin Sheppard

<http://www.kevinsheppard.com>

Oxford MFE

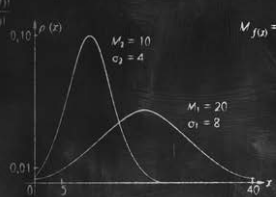
This version: February 18, 2019

February 21, 2019



$$D_x = \int_{-\infty}^{+\infty} (x - M_x)^2 \phi(x) dx$$

$$M_x = \int_{-\infty}^{+\infty} x \cdot \phi(x) dx$$

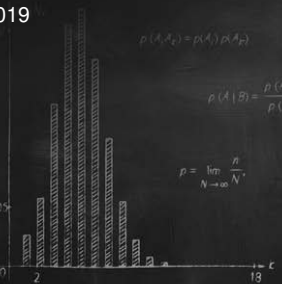


$$M_{f(x)} = \int_{-\infty}^{+\infty} f(x) \phi(x) dx$$

$$S = v\sigma^2 + \frac{\sigma^4}{2}$$

$$F = G \frac{m_1 m_2}{\rho^2}$$

$$f(v) = 4\pi \left(\frac{m_0}{2\pi kT} \right)^{3/2} v^2 e^{-\frac{mv^2}{2kT}}$$



$$\sigma_n = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

$$D_x = \sigma^2 = M_{x^2} - (M_x)^2$$

$$p_T(\lambda) = \frac{\lambda^c}{c!} e^{-\lambda}$$

$$p(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} \phi(x) dx$$

$$M_x = \sum_{i=1}^c p_i x_i$$

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

$$p = \lim_{N \rightarrow \infty} \frac{n}{N}$$

$$D_x = \sum_{i=1}^c p_i (x_i - M_x)^2$$

$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

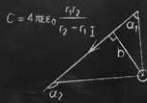
$$\phi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

$$\phi(v) = 4 \sqrt{\frac{k^3}{\pi}} v^2 e^{-kv^2}$$

$$\phi(\ln x) d(\ln x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} d(\ln x) = \frac{1}{\sqrt{2\pi\sigma x}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx$$

$$\langle v \rangle = \frac{\langle v \rangle t}{n \sqrt{2\pi} d^2}$$

$$C = \frac{\epsilon \epsilon_0 S}{d}$$



$$B = \frac{\mu_0 I}{2\pi b} (\cos \alpha_1 - \cos \alpha_2)$$

$$A^2 = A_1^2 + A_2^2 + 2A_1 A_2 \cos(\Phi_2 - \Phi_1)$$

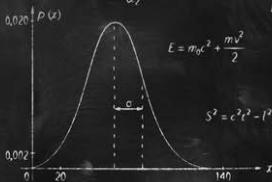
$$h\nu = A + \frac{mv^2}{2}$$

$$E = m_0 c^2 + \frac{mv^2}{2}$$

$$m = m_0 \sqrt{1 - \frac{v^2}{c^2}}$$

$$S^2 = c^2 t^2 - l^2 = \ln v$$

$$r_n = \frac{4\pi\epsilon_0 n^2 a^2}{m Z e^2}$$



This week's material

- Vector Autoregressions
- Basic examples
- Properties
 - ▶ Stationarity
- Revisiting univariate ARMA processes
- Forecasting
 - ▶ Granger Causality
 - ▶ Impulse Response functions
- Cointegration
 - ▶ Examining long-run relationships
 - ▶ Determining whether a VAR is cointegrated
 - ▶ Error Correction Models
 - ▶ Testing for Cointegration
 - Engle-Granger

Lots of revisiting univariate time series.

$$n! \approx \left(\frac{n}{e}\right)^n \cdot \sqrt{2\pi n}$$

$$P_n = \frac{n!}{(n-n)!} = \frac{n!}{0!}$$

$$A_n^k = \frac{n!}{(n-k)!}$$

$$A_n^k = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1)$$

$$\lambda_n^k = \underbrace{n \cdot n \cdot \dots \cdot n}_k = n^k$$

$$\tilde{C}_n^m = P_{m, n-m-1} = \frac{(n+m-1)!}{m!(n-1)!}$$

$$\tilde{P}_{n_1, n_2, \dots, n_k} = \frac{(n_1 + n_2 + \dots + n_k)!}{n_1! n_2! \dots n_k!}$$

$$C_n^k = \frac{n!}{k!(n-k)!}$$

$$(a+b)^n = C_n^0 a^n b^0 + C_n^1 a^{n-1} b^1 + C_n^2 a^{n-2} b^2 + \dots + C_n^{n-1} a^1 b^{n-1} + C_n^n a^0 b^n = \sum_{k=0}^n C_n^k a^{n-k} b^k$$

$$p(B) = p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + p(B|A_3)p(A_3) + \dots + p(B|A_n)p(A_n)$$

$$p(x) = \frac{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}$$



$$D_x = \int_{-\infty}^{+\infty} (x - M_x)^2 \phi(x) dx$$

$$M_x = \int_{-\infty}^{+\infty} x \cdot \phi(x) dx$$

$$M_{f(x)} = \int_{-\infty}^{+\infty} f(x) \phi(x) dx$$

$$S = \sqrt{\sigma^2} = \frac{d^2}{2}$$

$$E = G \frac{m_1 m_2}{Q^2}$$

$$f(x) = A \exp\left(-\frac{m_1 x^2}{2 \sigma^2}\right) \exp\left(-\frac{m_2 x^2}{2 \sigma^2}\right)$$



Vector Autoregressions



$$D_x = \sigma^2 = M_x^2 - (M_x)^2$$

$$p(x) = \frac{1}{\sigma} e^{-x^2/\sigma^2}$$

$$p(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} \phi(x) dx$$

$$M_x = \sum_{i=1}^n p_i x_i$$

$$D_x = \sum_{i=1}^n p_i (x_i - M_x)^2$$

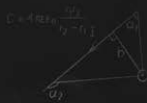
$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x-\mu}{\sigma}} \frac{1}{\sigma^2}$$

$$\phi(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

$$\phi(x) = A \sqrt{\frac{k^3}{\pi}} \sqrt{e^{-kx^2}}$$

$$C = \frac{2V}{\pi \sqrt{2\pi} d^2}$$

$$C = \frac{\pi r^2 S}{d}$$

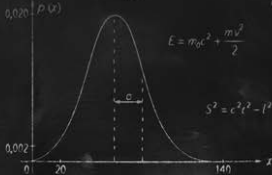


$$d = \frac{u_1}{\sqrt{2}} (\cos \varphi_1 - \cos \varphi_2)$$

$$d^2 = d_1^2 + d_2^2 + 2 d_1 d_2 \cos(\varphi_2 - \varphi_1)$$

$$nv = A + \frac{mv^2}{2}$$

$$0.020 \rho(x)$$



$$E = m_0 c^2 + \frac{mv^2}{2}$$

$$m = m_0 \sqrt{1 - \beta^2}$$

$$S^2 = c^2 t^2 - l^2 = in v$$

$$r_n = \frac{4\pi \epsilon_0 n^2 a^2}{m Z e^2}$$

Why VAR analysis?

- Stationary VARs
 - ▶ Determine whether variables feedback into one another
 - ▶ Improve forecasts
 - ▶ Model the effect of a shock in one series on another
 - ▶ Differentiate between short-run and long-run dynamics
- Cointegration
 - ▶ Link random walks
 - ▶ Uncover long run relationships
 - ▶ Can improve medium to long term forecasting [a lot](#)

$$f'_x = \frac{df}{dx}$$



$$Q_x = \int_{-\infty}^{\infty} (x - M_x) f(x) dx$$

- Pth order autoregression, AR(P):

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_P y_{t-p} + \epsilon_t$$

- Pth order vector autoregression, VAR(P):

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \dots + \Phi_P \mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t$$

where \mathbf{y}_t and $\boldsymbol{\epsilon}_t$ are k by 1 vectors

- Bivariate VAR(1):

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- Compactly expresses two linked models:

$$y_{1,t} = \phi_{01} + \phi_{11} y_{1,t-1} + \phi_{12} y_{2,t-1} + \epsilon_{1,t}$$

$$y_{2,t} = \phi_{02} + \phi_{21} y_{1,t-1} + \phi_{22} y_{2,t-1} + \epsilon_{2,t}$$

Stationarity Revisited

- Stationarity is a statistically meaningful form of regularity. A stochastic process $\{y_t\}$ is covariance stationary if

$$E[y_t] = \mu \quad \forall t$$

$$V[y_t] = \sigma^2 \quad \sigma^2 < \infty \forall t$$

$$E[(y_t - \mu)(y_{t-s} - \mu)] = \gamma_s \quad \forall t, s$$

- AR(1) stationarity: $y_t = \phi y_{t-1} + \epsilon_t$
 - ▶ $|\phi| < 1$
 - ▶ ϵ_t is white noise
- AR(P) stationarity: $y_t = \phi_1 y_{t-1} + \dots + \phi_P y_{t-P} + \epsilon_t$
 - ▶ Roots of $(z^P - \phi_1 z^{P-1} - \phi_2 z^{P-2} - \dots - \phi_{P-1} z - \phi_P)$ less than 1
 - ▶ ϵ_t is white noise
- No dependence on t



$$p_x = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

■ AR(1)

$$\begin{aligned}y_t &= \phi_0 + \phi_1 y_{t-1} + \epsilon_t \\&= \phi_0 + \phi_1(\phi_0 + \phi_1 y_{t-2} + \epsilon_{t-1}) + \epsilon_t \\&= \phi_0 + \phi_1 \phi_0 + \phi_1^2 y_{t-2} + \phi_1 \epsilon_{t-1} + \epsilon_t \\&= \phi_0 + \phi_1 \phi_0 + \phi_1^2(\phi_0 + \phi_1 y_{t-3} + \epsilon_{t-2}) + \phi_1 \epsilon_{t-1} + \epsilon_t \\&= \phi_0 \sum_{i=0}^{\infty} \phi_1^i + \sum_{i=0}^{\infty} \phi_1^i \epsilon_{t-i} \\&= (1 - \phi_1)^{-1} \phi_0 + \sum_{i=0}^{\infty} \phi_1^i \epsilon_{t-i}\end{aligned}$$

■ VAR(1)

$$\begin{aligned} \mathbf{y}_t &= \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \epsilon_t \\ &= \Phi_0 + \Phi_1 (\Phi_0 + \Phi_1 \mathbf{y}_{t-2} + \epsilon_{t-1}) + \epsilon_t \\ &= \Phi_0 + \Phi_1 \Phi_0 + \Phi_1^2 \mathbf{y}_{t-2} + \Phi_1 \epsilon_{t-1} + \epsilon_t \\ &= \Phi_0 + \Phi_1 \Phi_0 + \Phi_1^2 (\Phi_0 + \Phi_1 \mathbf{y}_{t-3} + \epsilon_{t-2}) + \Phi_1 \epsilon_{t-1} + \epsilon_t \\ &= \sum_{i=0}^{\infty} \Phi_1^i \Phi_0 + \sum_{i=0}^{\infty} \Phi_1^i \epsilon_{t-i} \\ &= (\mathbf{I}_k - \Phi_1)^{-1} \Phi_0 + \sum_{i=0}^{\infty} \Phi_1^i \epsilon_{t-i} \end{aligned}$$

Properties of a VAR(1) and AR(1)

$$\sigma_{\epsilon}^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

$$\text{AR}(1) : y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t$$

$$\text{VAR}(1) : \mathbf{y}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \epsilon_t$$

	AR(1)	VAR(1)
Mean	$\phi_0 / (1 - \phi_1)$	$(\mathbf{I}_k - \mathbf{\Phi}_1)^{-1} \mathbf{\Phi}_0$
Variance	$\sigma^2 / (1 - \phi_1^2)$	$(\mathbf{I} - \mathbf{\Phi}_1 \otimes \mathbf{\Phi}_1)^{-1} \text{vec}(\mathbf{\Sigma})$
s^{th} Autocovariance	$\gamma_s = \phi_1^s \text{V}[y_t]$	$\mathbf{\Gamma}_s = \mathbf{\Phi}_1^s \text{V}[\mathbf{y}_t]$
$-s^{\text{th}}$ Autocovariance	$\gamma_{-s} = \phi_1^s \text{V}[y_t]$	$\mathbf{\Gamma}_{-s} = \text{V}[\mathbf{y}_t] \mathbf{\Phi}_1^{s'}$

Autocovariances of vector processes are not symmetric, but $\mathbf{\Gamma}_s = \mathbf{\Gamma}'_{-s}$

■ Stationarity

- ▶ AR(1): $|\phi_1| < 1$
- ▶ VAR(1): $|\lambda_i| < 1$ where λ_i are the eigenvalues of $\mathbf{\Phi}_1$

Stock and Bond VAR

- VWM from CRSP
- TERM constructed from 10-year bond *return minus 1-year return* from FRED
- February 1962 until December 2018 (683 months)

$$\begin{bmatrix} VW M_t \\ TERM_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11,1} & \phi_{12,1} \\ \phi_{21,1} & \phi_{22,1} \end{bmatrix} \begin{bmatrix} VW M_{t-1} \\ TERM_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- Market model:

$$VW M_t = \phi_{01} + \phi_{11,1} VW M_{t-1} + \phi_{12,1} 10Y R_{t-1} + \epsilon_{1,t}$$

- Long bond model

$$TERM_t = \phi_{01} + \phi_{21,1} VW M_{t-1} + \phi_{22,1} TERM_{t-1} + \epsilon_{2,t}$$

- Estimates

$$\begin{bmatrix} VW M_t \\ TERM_t \end{bmatrix} = \begin{bmatrix} 0.801 \\ (0.000) \\ 0.232 \\ (0.041) \end{bmatrix} + \begin{bmatrix} 0.059 & 0.166 \\ (0.122) & (0.004) \\ -0.104 & 0.116 \\ (0.000) & (0.002) \end{bmatrix} \begin{bmatrix} VW M_{t-1} \\ TERM_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

■ Estimates from VAR

$$\begin{aligned}
 VWM_t &= 0.816 + 0.060 VWM_{t-1} + 0.168 TERM_{t-1} \\
 &\quad (0.000) \quad (0.117) \quad (0.003) \\
 TERM_t &= 0.228 - 0.104 VWM_{t-1} + 0.115 TERM_{t-1} \\
 &\quad (0.045) \quad (0.000) \quad (0.002)
 \end{aligned}$$

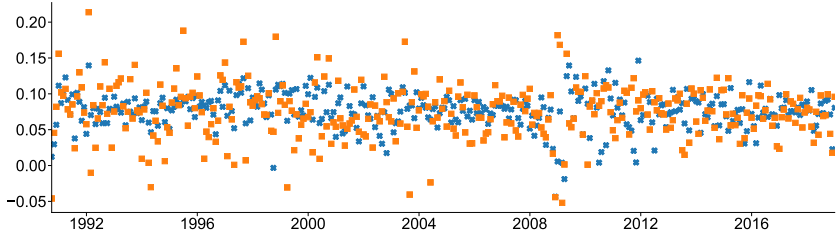
■ Estimates from AR

$$\begin{aligned}
 VWM_t &= 0.830 + 0.073 VWM_{t-1} \\
 &\quad (0.000) \quad (0.057) \\
 TERM_t &= 0.137 + 0.098 TERM_{t-1} \\
 &\quad (0.224) \quad (0.011)
 \end{aligned}$$

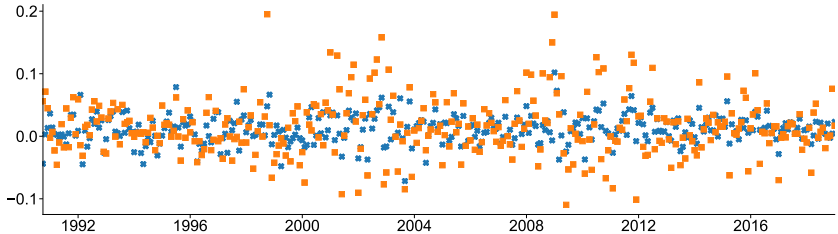
Comparing AR and VAR forecasts

$$\hat{y}_t = \mu + \sigma \epsilon_t$$

1-month-ahead forecasts of the VWM returns



1-month-ahead forecasts of 10-year bond returns



- Standard tool in monetary policy analysis

- ▶ Unemployment rate (differenced)
- ▶ Federal Funds rate
- ▶ Inflation rate (differenced)

$$\begin{bmatrix} \Delta \text{UNEMP}_t \\ \text{FF}_t \\ \Delta \text{INF}_t \end{bmatrix} = \Phi_0 + \Phi_1 \begin{bmatrix} \Delta \text{UNEMP}_{t-1} \\ \text{FF}_{t-1} \\ \Delta \text{INF}_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \epsilon_{3,t} \end{bmatrix}.$$

	$\Delta \ln \text{UNEMP}_{t-1}$	FF_{t-1}	ΔINF_{t-1}
$\Delta \ln \text{UNEMP}_t$	0.624 (0.000)	0.015 (0.001)	0.016 (0.267)
FF_t	-0.816 (0.000)	0.979 (0.000)	-0.045 (0.317)
ΔINF_t	-0.501 (0.010)	-0.009 (0.626)	-0.401 (0.000)

Interpreting Estimates



- Variable scale affects cross-parameter estimates
 - ▶ Not an issue in ARMA analysis
- Standardizing data can improve interpretation when scales differ

	$\Delta \ln \text{UNEMP}_{t-1}$	FF_{t-1}	ΔINF_{t-1}
$\Delta \ln \text{UNEMP}_t$	0.624 (0.000)	0.015 (0.001)	0.016 (0.267)
FF_t	-0.816 (0.000)	0.979 (0.000)	-0.045 (0.317)
ΔINF_t	-0.501 (0.010)	-0.009 (0.626)	-0.401 (0.000)

- Other important measures – statistical significance, persistence, model selection – are unaffected by standardization

VAR(P) is really a VAR(1)

- Companion form:

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_P \mathbf{y}_{t-P} + \epsilon_t$$

- Reform into a single VAR(1) where

$$\boldsymbol{\mu} = E[\mathbf{y}_t] = (\mathbf{I} - \Phi_1 - \dots - \Phi_P)^{-1} \Phi_0$$

$$\mathbf{z}_t = \Upsilon \mathbf{z}_{t-1} + \boldsymbol{\xi}_t$$

$$\mathbf{z}_t = \begin{bmatrix} \mathbf{y}_t - \boldsymbol{\mu} \\ \mathbf{y}_{t-1} - \boldsymbol{\mu} \\ \vdots \\ \mathbf{y}_{t-P+1} - \boldsymbol{\mu} \end{bmatrix}, \quad \Upsilon = \begin{bmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \dots & \Phi_{P-1} & \Phi_P \\ \mathbf{I}_k & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_k & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_k & \mathbf{0} \end{bmatrix}$$

- ▶ All results can be directly applied to the companion form.
- ▶ Can also be used to transform AR(P) into VAR(1)

$$n! \approx \left(\frac{n}{e}\right)^n \cdot \sqrt{2\pi n}$$

$$P_n = \frac{n!}{(n-n)!} = \frac{n!}{0!}$$

$$A_n^k = \frac{n!}{(n-k)!}$$

$$A_n^k = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1)$$

$$\lambda_n^k = \underbrace{n \cdot n \cdot \dots \cdot n}_k = n^k$$

$$\tilde{C}_n^m = P_{m, n-m-1} = \frac{(n+m-1)!}{m!(n-1)!}$$

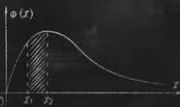
$$\tilde{P}_{n_1, n_2, \dots, n_k} = \frac{(n_1 + n_2 + \dots + n_k)!}{n_1! n_2! \dots n_k!}$$

$$C_n^c = \frac{n!}{c!(n-c)!}$$

$$(a+b)^n = C_n^0 a^n b^0 + C_n^1 a^{n-1} b^1 + C_n^2 a^{n-2} b^2 + \dots + C_n^{n-1} a^1 b^{n-1} + C_n^n a^0 b^n = \sum_{k=0}^n C_n^k a^{n-k} b^k$$

$$p(B) = p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + p(B|A_3)p(A_3) + \dots + p(B|A_n)p(A_n)$$

$$p(x) = \frac{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}$$



$$D_x = \int_{-\infty}^{+\infty} (x - M_x)^2 \phi(x) dx$$

$$M_x = \int_{-\infty}^{+\infty} x \cdot \phi(x) dx$$

$$M_{f(x)} = \int_{-\infty}^{+\infty} f(x) \phi(x) dx$$

$$S = \sqrt{\sigma^2} = \frac{d^2}{2}$$

$$E = G \frac{m_1 m_2}{Q^2}$$

$$f(x) = 4\pi \left(\frac{m_1}{15x^2}\right)^{1/2} \cdot 2 \cdot \frac{d^2}{2}$$



Forecasting



$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - M_x)^2}{n}}$$

$$D_x = \sigma^2 = M_{x^2} - (M_x)^2$$

$$P_c(t) = \frac{1}{t!} e^{-t}$$

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} \phi(x) dx$$

$$M_x = \sum_{i=1}^c p_i x_i$$

$$D_x = \sum_{i=1}^c p_i (x_i - M_x)^2$$

$$\phi(x) = \frac{1}{\sqrt{2\pi c}} e^{-\frac{(x-M_x)^2}{2c}}$$

$$\phi(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

$$\phi(v) = 4 \sqrt{\frac{k^3}{\pi}} \sqrt{v} e^{-kv^2}$$

$$C = \frac{2V}{\pi \sqrt{2} \pi d^2}$$

$$C = 4 \cos \alpha \frac{2V}{\pi \sqrt{2} \pi d^2}$$

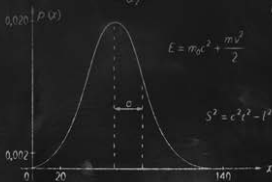


$$d = \frac{m_1}{2\pi} (\cos \alpha_1 - \cos \alpha_2)$$

$$d^2 = d_1^2 + d_2^2 + 2 d_1 d_2 \cos(\varphi_2 - \varphi_1)$$

$$C = \frac{\pi c S}{d}$$

$$nV = A + \frac{mV^2}{2}$$



$$E = m_0 c^2 + \frac{mV^2}{2}$$

$$m = m_0 \sqrt{1 - \beta^2}$$

$$S^2 = c^2 t^2 - l^2 = mV$$

$$r_n = \frac{4\pi \epsilon_0 n^2 n^2}{m^2 c^2}$$

Revisiting Univariate Forecasting

$$\sigma_y^2 = \int_{-\infty}^{\infty} (y - \mu_y)^2 g(y) dy$$

- Consider standard AR(1)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t$$

- Optimal 1-step ahead forecast:

$$\begin{aligned} E_t[y_{t+1}] &= E_t[\phi_0] + E_t[\phi_1 y_t] + E_t[\epsilon_{t+1}] \\ &= \phi_0 + \phi_1 y_t + 0 \end{aligned}$$

- Optimal 2-step ahead forecast:

$$\begin{aligned} E_t[y_{t+2}] &= E_t[\phi_0] + E_t[\phi_1 y_{t+1}] + E_t[\epsilon_{t+2}] \\ &= \phi_0 + \phi_1 E_t[y_{t+1}] + 0 \\ &= \phi_0 + \phi_1 (\phi_0 + \phi_1 y_t) \\ &= \phi_0 + \phi_1 \phi_0 + \phi_1^2 y_t \end{aligned}$$

- Optimal h -step ahead forecast:

$$E_t[y_{t+h}] = \sum_{i=0}^{h-1} \phi_1^i \phi_0 + \phi_1^h y_t$$

Forecasting with VARs



- Identical to univariate case

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \epsilon_t$$

- Optimal 1-step ahead forecast:

$$\begin{aligned} E_t[\mathbf{y}_{t+1}] &= E_t[\Phi_0] + E_t[\Phi_1 \mathbf{y}_t] + E_t[\epsilon_{t+1}] \\ &= \Phi_0 + \Phi_1 \mathbf{y}_t + \mathbf{0} \end{aligned}$$

- Optimal h-step ahead forecast:

$$\begin{aligned} E_t[\mathbf{y}_{t+h}] &= \Phi_0 + \Phi_1 \Phi_0 + \dots + \Phi_1^{h-1} \Phi_0 + \Phi_1^h \mathbf{y}_t \\ &= \sum_{i=0}^{h-1} \Phi_1^i \Phi_0 + \Phi_1^h \mathbf{y}_t \end{aligned}$$

- Higher order forecast can be recursively computed

$$E_t[\mathbf{y}_{t+h}] = \Phi_0 + \Phi_1 E_t[\mathbf{y}_{t+h-1}] + \dots + \Phi_P E_t[\mathbf{y}_{t+h-P}]$$

What makes a good forecast?

- Forecast residuals

$$\hat{e}_{t+h|t} = y_{t+h} - \hat{y}_{t+h|t}$$

- Residuals are *not* white noise
- Can contain an MA($h - 1$) component
 - ▶ Forecast error for $y_{t+1} - \hat{y}_{t+1|t-h+1}$ was not known at time t .
- Plot your residuals
- Residual ACF
- Mincer-Zarnowitz regressions
- Three period procedure
 - ▶ Training sample: Used to build model
 - ▶ Validation sample: Used to refine model
 - ▶ Evaluation sample: Ultimate test, ideally 1 shot

Multi-step Forecasting



- Two methods
- Iterative method
 - ▶ Build model for 1-step ahead forecasts

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \epsilon_t$$

- ▶ Iterate forecast out to period h

$$\hat{\mathbf{y}}_{t+h|t} = \sum_{i=0}^{h-1} \Phi_1^i \Phi_0 + \Phi_1^h \mathbf{y}_t$$

- ▶ Makes efficient use of information
 - ▶ Imposes a lot of structure on the problem
- Direct Method
 - ▶ Build model for h -step ahead forecasts

$$\mathbf{y}_t = \Phi_0 + \Phi_h \mathbf{y}_{t-h} + \epsilon_t$$

- ▶ Directly forecast using a pseudo 1-step ahead method

$$\hat{\mathbf{y}}_{t+h|t} = \Phi_0 + \Phi_h \mathbf{y}_t$$

- ▶ Robust to some nonlinearities

Multi-step Forecast Evaluation

- Multistep forecast evaluation is identical to one-step ahead forecast evaluation with one caveat
- h -step ahead forecast errors may be correlated with any forecast error not known at time t

$$\hat{\epsilon}_{t+1|t-h+1}, \hat{\epsilon}_{t+2|t-h+2}, \dots, \hat{\epsilon}_{t+h-1|t-1}$$

- Leads to a $MA(h-1)$ structure in the forecast errors
- Solutions:
 - ▶ Use regular GMZ regression with a Newey-West covariance estimator

$$y_{t+h} - \hat{y}_{t+h|t} = \beta_1 + \beta_2 \hat{y}_{t+h|t} + \gamma \mathbf{x}_t + \eta_t$$

$$H_0 : \beta_1 = \beta_2 = \gamma = 0, H_1 : \beta_1 \neq 0 \cup \beta_2 \neq 0 \cup \gamma_j \neq 0 \exists j$$

- ▶ Explicitly model the $MA(h-1)$ and use a standard covariance estimator

$$y_{t+h} - \hat{y}_{t+h|t} = \beta_1 + \beta_2 \hat{y}_{t+h|t} + \gamma \mathbf{x}_t + \eta_t + \sum_{i=1}^{h-1} \theta_i \eta_{t-i}$$

Note: Null is the same; does not impose a restriction on θ

Example: Monetary Policy VAR

- Forecasts produced iteratively for 1 to 8 quarters ahead
- Random walk (FF) or constant mean benchmark
- AR and VAR select lag length using BIC
- Restricted force reversion to in-sample mean using 2-step estimator
 1. Estimate sample mean, and subtract to produce $\tilde{y}_t = y_t - \hat{\mu}$
 2. Estimate VAR *without* a constant

$$\tilde{y}_t = \Phi_1 \tilde{y}_{t-1} + \dots + \Phi_P \tilde{y}_{t-P} + \epsilon_t$$

3. Forecast and then add the in-sample mean

$$E_t [\tilde{y}_{t+h}] + \hat{\mu}$$

- Evaluation based on relative MSE

$$\text{Rel. MSE} = \frac{\text{MSE}}{\text{MSE}_{bm}}, \quad \text{MSE} = 1/T-h-R \sum_{t=R}^{T-h} (y_{t+h} - \hat{y}_{t+h|t})^2$$

Example: Monetary Policy VAR

Horizon	Series	VAR		AR	
		Restricted	Unrestricted	Restricted	Unrestricted
1	Unemployment	0.522	0.520	0.507	0.507
	Fed. Funds Rate	0.887	0.903	0.923	0.933
	Inflation	0.869	0.868	0.839	0.840
2	Unemployment	0.716	0.710	0.717	0.718
	Fed. Funds Rate	0.923	0.943	<i>1.112</i>	<i>1.130</i>
	Inflation	<i>1.082</i>	<i>1.081</i>	<i>1.031</i>	<i>1.030</i>
4	Unemployment	0.872	0.861	0.937	0.940
	Fed. Funds Rate	0.952	0.976	<i>1.082</i>	<i>1.109</i>
	Inflation	<i>1.000</i>	0.999	0.998	0.998
8	Unemployment	0.820	0.806	0.973	0.979
	Fed. Funds Rate	0.974	<i>1.007</i>	<i>1.062</i>	<i>1.110</i>
	Inflation	<i>1.001</i>	1.000	0.998	0.997

$$n! \approx \left(\frac{n}{e}\right)^n \cdot \sqrt{2\pi n}$$

$$P_n = \frac{n!}{(n-n)!} = \frac{n!}{0!}$$

$$A_n^k = \frac{n!}{(n-k)!}$$

$$A_n^k = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1)$$

$$\lambda_n^k = \underbrace{n \cdot n \cdot \dots \cdot n}_k = n^k$$

$$\tilde{C}_n^m = P_{m, n-m-1} = \frac{(n+m-1)!}{m!(n-1)!}$$

$$\tilde{P}_{n_1, n_2, \dots, n_k} = \frac{(n_1 + n_2 + \dots + n_k)!}{n_1! n_2! \dots n_k!}$$

$$C_n^c = \frac{n!}{c!(n-c)!}$$

$$(a+b)^n = C_n^0 a^n b^0 + C_n^1 a^{n-1} b^1 + C_n^2 a^{n-2} b^2 + \dots + C_n^{n-1} a^1 b^{n-1} + C_n^n a^0 b^n = \sum_{k=0}^n C_n^k a^{n-k} b^k$$

$$p(B) = p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + p(B|A_3)p(A_3) + \dots + p(B|A_n)p(A_n)$$

$$p(x) = \frac{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}$$



$$D_x = \int_{-\infty}^{+\infty} (x - M_x)^2 \phi(x) dx$$

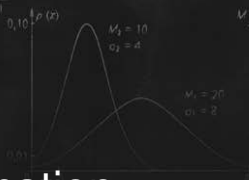
$$M_x = \int_{-\infty}^{+\infty} x \cdot \phi(x) dx$$

$$M_{f(x)} = \int_{-\infty}^{+\infty} f(x) \phi(x) dx$$

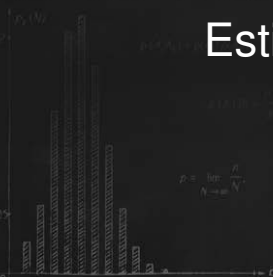
$$S = \sqrt{\sigma^2} = \frac{d^2}{2}$$

$$E = G \frac{m_1 m_2}{Q^2}$$

$$f(x) = A \exp\left(-\frac{m_1 x}{15 \pi^2}\right) \exp\left(-\frac{m_2 x}{15 \pi^2}\right)$$



Estimation



$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

$$D_x = \sigma^2 = M_x^2 - (M_x)^2$$

$$P_c(k) = \frac{1}{k!} e^{-k}$$

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} \phi(x) dx$$

$$M_x = \sum_{i=1}^c p_i x_i$$

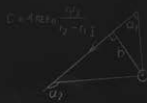
$$D_x = \sum_{i=1}^c p_i (x_i - M_x)^2$$

$$\phi(x) = \frac{1}{\sqrt{2\pi c}} e^{-\frac{(x-\bar{x})^2}{2c}}$$

$$\phi(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

$$\phi(v) = A \sqrt{\frac{k^3}{\pi}} \sqrt{v} e^{-kv^2}$$

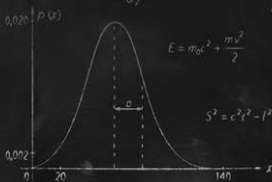
$$C = \frac{2V}{\pi \sqrt{2} \pi d^2}$$



$$a = \frac{a_1}{\sin \phi} = \frac{a_2}{\cos \phi}$$

$$a^2 = a_1^2 + a_2^2 + 2 a_1 a_2 \cos(\phi_2 - \phi_1)$$

$$nv = A + \frac{mv^2}{2}$$



$$E = m_0 c^2 + \frac{m v^2}{2}$$

$$m = m_0 \sqrt{1 - \beta^2}$$

$$S^2 = c^2 - l^2 = in v$$

$$\tau_n = \frac{4 \pi \epsilon_0 n^2 n^2}{m^2 c^2}$$



- Univariate Identification: Box-Jenkins
 - ▶ Use ACF and PACF to determine AR and MA lag order
 - ▶ Examine residuals
 - ▶ Parsimony principle
- The autocorrelation of a scalar process is defined

$$\rho_s = \frac{\gamma_s}{\gamma_0}$$

where γ_s is s^{th} the autocovariance

- ▶ Regression coefficient:

$$y_t = \mu + \rho_s y_{t-s} + \epsilon_t$$

- Partial autocorrelation ψ_s
 - ▶ Regression interpretation of s^{th} partial autocorrelation:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_{s-1} y_{t-s+1} + \psi_s y_{t-s} + \epsilon_t$$

- ▶ ψ is the s^{th} partial autocorrelation



$$\rho_{xy} = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}$$

- Multivariate equivalents
 - ▶ ACF and PACF have same regression definitions
 - ▶ Cross-correlation function

$$\rho_{xy,s} = \frac{\text{E}[(x_t - \mu_x)(y_{t-s} - \mu_y)]}{\sqrt{\text{V}[x_t]\text{V}[y_t]}}$$

$$\rho_{yx,s} = \frac{\text{E}[(y_t - \mu_y)(x_{t-s} - \mu_x)]}{\sqrt{\text{V}[x_t]\text{V}[y_t]}}$$

- ▶ Generally different
- ▶ Cross-partial-correlation function $\psi_{xy,s}$

$$\begin{aligned} x_t = & \phi_0 + \phi_{x1}x_{t-1} + \dots + \phi_{xs-1}x_{t-(s-1)} \\ & + \phi_{y1}y_{t-1} + \dots + \phi_{ys-1}y_{t-(s-1)} + \varphi_{xy,s}y_{t-s} + \epsilon_{x,t} \end{aligned}$$

– Can help identify VAR order

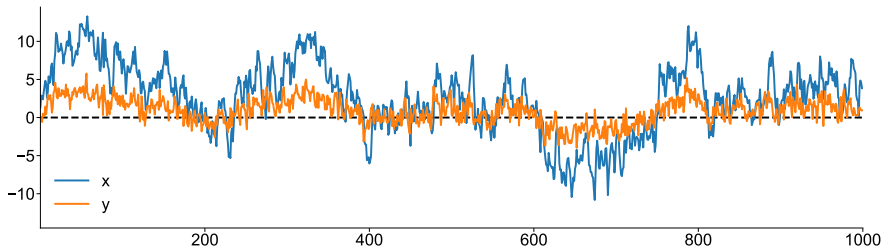
- Deeper issue: too many and too complicated
- Simple solution: Model selection

Interpreting CCFs and PCCFs

- y has HAR dynamics, spills over to x

$$\begin{aligned} \begin{bmatrix} x_t \\ y_t \end{bmatrix} &= \begin{bmatrix} 0.5 & 0.9 \\ .0 & 0.47 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \sum_{i=2}^5 \begin{bmatrix} 0 & 0 \\ 0 & 0.06 \end{bmatrix} \begin{bmatrix} x_{t-i} \\ y_{t-i} \end{bmatrix} \\ &+ \sum_{j=6}^{22} \begin{bmatrix} 0 & 0 \\ 0 & 0.01 \end{bmatrix} \begin{bmatrix} x_{t-j} \\ y_{t-j} \end{bmatrix} + \begin{bmatrix} \epsilon_{x,t} \\ \epsilon_{y,t} \end{bmatrix} \end{aligned}$$

- Simulated data



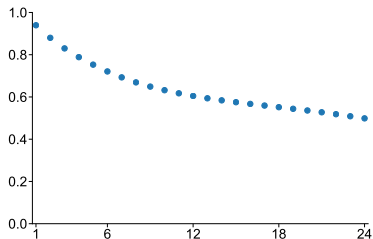
ACFs and CCFs

$$f_x = \frac{m!}{(m-x)!}$$

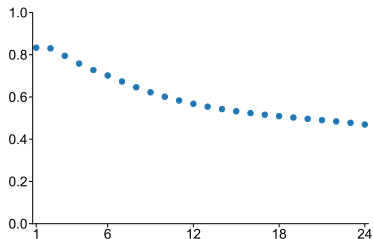


$$f_x = \int_{-\infty}^{\infty} (x - M_x)^2 f(x) dx$$

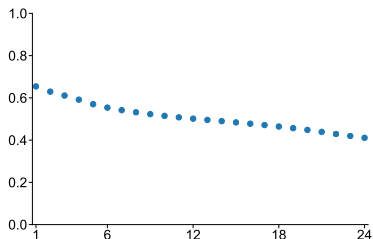
ACF (x on lagged x)



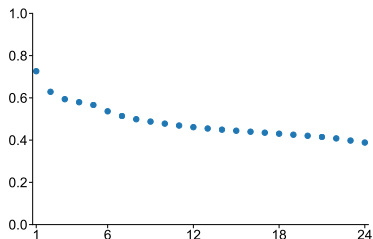
CCF (x on lagged y)



CCF (y on lagged x)



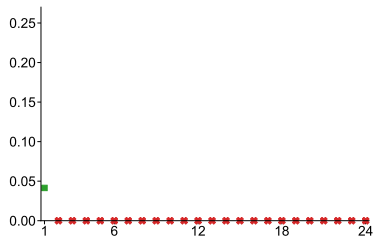
ACF (y on lagged y)



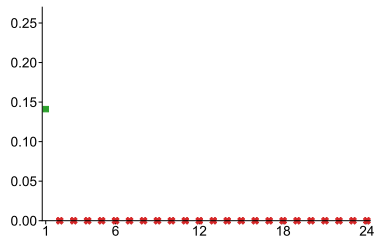
PACFs and Partial CCFs

$$\sigma_x^2 = \int_{-\pi}^{\pi} |u - M_x / \sigma_x|^2 dx$$

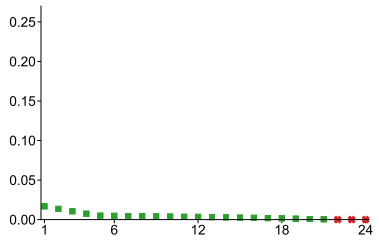
PACF (x on lagged x)



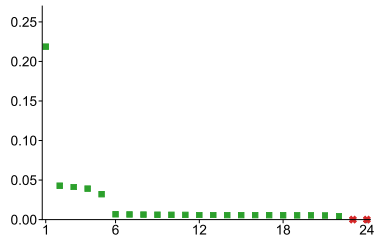
PCCF (x on lagged y)



PCCF (y on lagged x)



PACF (y on lagged y)



- Step 1: Pick maximum lag length
 - ▶ Information criteria

$$\text{AIC:} \quad \ln |\Sigma(P)| + k^2 P \frac{2}{T}$$

$$\text{Hannan-Quinn IC (HQIC):} \quad \ln |\Sigma(P)| + k^2 P \frac{\ln \ln T}{T}$$

$$\text{SIC:} \quad \ln |\Sigma(P)| + k^2 P \frac{\ln T}{T}$$

- $\Sigma(P)$ is the covariance of the residuals using P lags
- $|\cdot|$ is the determinant
- ▶ Hypothesis testing based
 - General to Specific
 - Specific to General
- ▶ Likelihood Ratio

$$(T - P_2 k^2) (\ln |\Sigma(P_1)| - \ln |\Sigma(P_2)|) \stackrel{A}{\sim} \chi_{(P_2 - P_1)k^2}^2$$

Lag Length Selection in Monetary Policy VAR

- Maximum lag: 12 (1 year)

Lag Length	AIC	HQIC	BIC	LR	P-val
0	4.014	3.762	3.605	925	0.000
1	0.279	0.079	0.000 ^{▼▲}	39.6	0.000
2	0.190	0.042	0.041	40.9	0.000
3	0.096	0.000 [▼]	0.076	29.0	0.001
4	0.050 [▼]	0.007	0.160	7.34	0.602 [▼]
5	0.094	0.103	0.333	29.5	0.001
6	0.047	0.108	0.415	13.2	0.155
7	0.067	0.180	0.564	32.4	0.000
8	0.007	0.172 [▲]	0.634	19.8	0.019
9	0.000 [▲]	0.217	0.756	7.68	0.566 [▲]
10	0.042	0.312	0.928	13.5	0.141
11	0.061	0.382	1.076	13.5	0.141
12	0.079	0.453	1.224	—	—

$$n! \approx \left(\frac{n}{e}\right)^n \cdot \sqrt{2\pi n}$$

$$P_n = \frac{n!}{(n-n)!} = \frac{n!}{0!}$$

$$A_n^k = \frac{n!}{(n-k)!}$$

$$A_n^k = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1)$$

$$\lambda_n^k = \underbrace{n \cdot n \cdot \dots \cdot n}_k = n^k$$

$$\tilde{C}_n^m = P_{m, n-m-1} = \frac{(n+m-1)!}{m!(n-1)!}$$

$$\tilde{C}_n^m = \frac{(n_1 + n_2 + \dots + n_p)!}{n_1! n_2! \dots n_p!} \quad C_n^c = \frac{n!}{c!(n-c)!}$$

$$(a+b)^n = C_n^0 a^n b^0 + C_n^1 a^{n-1} b^1 + C_n^2 a^{n-2} b^2 + \dots + C_n^{n-1} a^1 b^{n-1} + C_n^n a^0 b^n = \sum_{k=0}^n C_n^k a^{n-k} b^k$$

$$p(B) = p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + p(B|A_3)p(A_3) + \dots + p(B|A_n)p(A_n)$$

$$p(x) = \frac{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}$$



$$D_x = \int_{-\infty}^{+\infty} (x - M_x)^2 \phi(x) dx$$

$$M_x = \int_{-\infty}^{+\infty} x \cdot \phi(x) dx$$

$$M_{f(x)} = \int_{-\infty}^{+\infty} f(x) \phi(x) dx$$

$$S = \sqrt{\sigma^2} = \frac{d^2}{2}$$

$$E = G \frac{m_1 m_2}{Q^2}$$

$$f(x) = A \exp\left(-\frac{m v}{15 x^2}\right) \exp\left(-\frac{x^2}{2\sigma^2}\right)$$



Granger Causality



$$D_x = \sigma^2 = M_x^2 - (M_x^2)$$

$$p_c(x) = \frac{1}{\sigma} e^{-x/\sigma}$$

$$p(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} \phi(x) dx$$

$$M_x = \sum_{i=1}^c p_i x_i$$

$$D_x = \sum_{i=1}^c p_i (x_i - M_x)^2$$

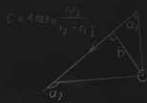
$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{|x-\mu|^2}{2\sigma^2}}$$

$$\phi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

$$\phi(v) = A \sqrt{\frac{k^3}{\pi}} \sqrt{v} e^{-kv^2}$$

$$C = \frac{2\gamma}{\pi \sqrt{2\pi} d^2}$$

$$C = \frac{\pi \epsilon S}{d}$$



$$d = \frac{h \lambda}{2 \sin \theta} (\cos \varphi_1 - \cos \varphi_2)$$

$$d^2 = d_1^2 + d_2^2 + 2 A_1 A_2 \cos(\varphi_2 - \varphi_1)$$

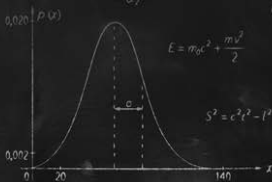
$$n v = A + \frac{m v^2}{2}$$

$$E = m_0 c^2 + \frac{m v^2}{2}$$

$$m = m_0 \sqrt{1 - \beta^2}$$

$$S^2 = c^2 t^2 - l^2 = i n v$$

$$r_n = \frac{4 \pi \epsilon_0 n^2 a^2}{m^2 c^2}$$



Granger Causality



$$D_t = \int_{-\infty}^{\infty} (x - M_t) f(x) dx$$

- First fundamentally new concept
- Examines whether lags of one variable are helpful in predicting another

Definition (Granger Causality)

A scalar random variable $\{x_t\}$ is said to **not** Granger cause $\{y_t\}$ if $E[y_t | x_{t-1}, y_{t-1}, x_{t-2}, y_{t-2}, \dots] = E[y_t | y_{t-1}, y_{t-2}, \dots]$. That is, $\{x_t\}$ does not Granger cause if the forecast of y_t is the same whether conditioned on past values of x_t or not.

- Translates directly into a restriction in a VAR
- Unrestricted

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- Restricted so that x_t does not GC y_t

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ 0 & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

$$x_t = \phi_{01} + \phi_{11}x_{t-1} + \phi_{12}y_{t-1} + \epsilon_{1,t}$$

$$y_t = \phi_{02} + \phi_{22}y_{t-1} + \epsilon_{2,t} \Leftarrow \text{No } x_t!$$

- In P lag model

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_P \mathbf{y}_{t-P} + \epsilon_t$$

the null hypothesis is

$$H_0 : \phi_{ij,1} = \phi_{ij,2} = \dots = \phi_{ij,P} = 0$$

- Alternative is

$$H_0 : \phi_{ij,1} \neq 0 \text{ or } \phi_{ij,2} \neq 0 \text{ or } \dots \text{ or } \phi_{ij,P} \neq 0$$

- Likelihood Ratio test

$$(T - Pk^2) (\ln |\Sigma_r| - \ln |\Sigma_u|) \overset{A}{\sim} \chi_P^2$$

- Σ_u is the covariance of the errors from unrestricted model
- Σ_r is the covariance of the errors from restricted model
- $T - Pk^2$ is number of observations minus number of free parameters in unrestricted model
 - ▶ Why χ_P^2 ?

- Standard tool in monetary policy analysis
 - ▶ Unemployment rate (differenced)
 - Federal Funds rate
 - Inflation rate (differenced)

$$\begin{bmatrix} \Delta \text{UNEMP}_t \\ \text{FF}_t \\ \Delta \text{INF}_t \end{bmatrix} = \Phi_0 + \Phi_1 \begin{bmatrix} \Delta \text{UNEMP}_{t-1} \\ \text{FF}_{t-1} \\ \Delta \text{INF}_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \epsilon_{3,t} \end{bmatrix}.$$

Granger Causality in Campbell's VAR

$\phi_{ij} = \sum_{k=1}^3 \phi_{ij,k} \alpha_{ij,k}$

- Using model with lags 3 lags (HQIC)
- $H_0 : \phi_{ij,1} = \phi_{ij,2} = \phi_{ij,3} = 0$
- $H_1 : \phi_{ij,1} \neq 0$ or $\phi_{ij,2} \neq 0$ or $\phi_{ij,3} \neq 0$
- i represent series being affected by lags of series j

Exclusion	Fed. Funds Rate		Inflation		Unemployment	
	P-val	Stat	P-val	Stat	P-val	Stat
Fed. Funds Rate	–	–	0.001	13.068	0.014	8.560
Inflation	0.001	14.756	–	–	0.375	1.963
Unemployment	0.000	19.586	0.775	0.509	–	–
All	0.000	33.139	0.000	18.630	0.005	10.472

$$n! \approx \left(\frac{n}{e}\right)^n \cdot \sqrt{2\pi n}$$

$$P_n = \frac{n!}{(n-n)!} = \frac{n!}{0!}$$

$$A_n^k = \frac{n!}{(n-k)!}$$

$$A_n^k = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1)$$

$$\lambda_n^k = \underbrace{n \cdot n \cdot \dots \cdot n}_k = n^k$$

$$\tilde{C}_n^m = P_{m, n-m-1} = \frac{(n+m-1)!}{m!(n-1)!}$$

$$\tilde{P}_{n_1, n_2, \dots, n_k} = \frac{(n_1 + n_2 + \dots + n_k)!}{n_1! n_2! \dots n_k!}$$

$$C_n^k = \frac{n!}{k!(n-k)!}$$

$$(a+b)^n = C_n^0 a^n b^0 + C_n^1 a^{n-1} b^1 + C_n^2 a^{n-2} b^2 + \dots + C_n^{n-1} a^1 b^{n-1} + C_n^n a^0 b^n = \sum_{k=0}^n C_n^k a^{n-k} b^k$$

$$p(B) = p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + p(B|A_3)p(A_3) + \dots + p(B|A_n)p(A_n)$$

$$p(x) = \frac{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}$$



$$D_x = \int_{-\infty}^{+\infty} (x - M_x)^2 \phi(x) dx$$

$$M_x = \int_{-\infty}^{+\infty} x \cdot \phi(x) dx$$

$$M_{f(x)} = \int_{-\infty}^{+\infty} f(x) \phi(x) dx$$

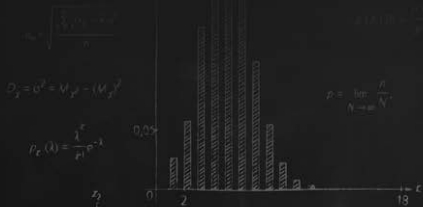
$$S = v_0 t + \frac{at^2}{2}$$

$$E = G \frac{m_1 m_2}{Q^2}$$

$$f(v) = A \exp\left(-\frac{mv}{kT}\right) \exp\left(-\frac{mv^2}{2kT}\right)$$



Impulse Response Functions



$$D_x = \sigma^2 = M_x^2 - (M_x)^2$$

$$p_x(x) = \frac{1}{\sigma} e^{-x/\sigma}$$

$$p(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} \phi(x) dx$$

$$M_x = \sum_{i=1}^n p_i x_i$$

$$D_x = \sum_{i=1}^n p_i (x_i - M_x)^2$$

$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\phi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

$$\phi(v) = A \sqrt{\frac{k^3}{\pi}} \sqrt{v} e^{-kv^2}$$

$$v = \frac{2V}{n \sqrt{2\pi} d^2}$$

$$C = \frac{v \pi n S}{d}$$



$$A = \frac{A_1}{\sin \alpha} (\cos \alpha_1 - \cos \alpha_2)$$

$$A^2 = A_1^2 + A_2^2 + 2A_1 A_2 \cos(\alpha_2 - \alpha_1)$$

$$nv = A + \frac{mv^2}{2}$$

$$0,020 \rho(v)$$



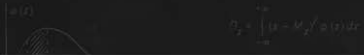
$$E = m_0 c^2 + \frac{mv^2}{2}$$

$$m = m_0 \sqrt{1 - \beta^2}$$

$$S^2 = c^2 t^2 - l^2 = in v$$

$$r_n = \frac{4\pi \epsilon_0 n^2 a^2}{m^2 Z e^2}$$

Impulse Response Functions



- Second fundamentally new concept
- Complicated dynamics of a VAR make direct interpretation of coefficients difficult
- Solution is to examine impulse responses
- The impulse response function of y_i with respect to a shock in ϵ_j , for any j and i , is defined as the change in y_{it+s} , $s \geq 0$ for a unit shock in ϵ_{jt}
 - ▶ Hard to decipher
- As long as \mathbf{y}_t is covariance stationarity it must have a VMA representation,

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\epsilon}_t + \boldsymbol{\Xi}_1 \boldsymbol{\epsilon}_{t-1} + \boldsymbol{\Xi}_2 \boldsymbol{\epsilon}_{t-2} + \dots$$

- $\boldsymbol{\Xi}_j$ are the impulse responses!
- Why?
 - ▶ Directly measure the effect in period j of any shock



$$f(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (u - M_y / \sigma_y^2) e^{i\omega u} du$$

- Any stationary AR(P)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_P y_{t-P} + \epsilon_t$$

can be represented as an MA(∞)

$$y_t = \phi_0 / (1 - \phi_1 - \phi_2 - \dots - \phi_P) + \epsilon_t + \sum_{i=1}^{\infty} \theta_i \epsilon_{t-i}$$

- AR(1)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t$$

becomes

$$y_t = \phi_0 / (1 - \phi_1) + \epsilon_t + \sum_{i=1}^{\infty} \phi_1^i \epsilon_{t-i}$$

- Stationary VAR(P) have the same relationship to VMA(∞)

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_P \mathbf{y}_{t-P} + \epsilon_t$$

$$\mathbf{y}_t = \boldsymbol{\mu} + \epsilon_t + \Xi_1 \epsilon_{t-1} + \Xi_2 \epsilon_{t-2} + \dots$$

- Easy in VAR(1)

$$\mathbf{y}_t = (\mathbf{I}_K - \Phi_1)^{-1} \Phi_0 + \epsilon_t + \Phi_1 \epsilon_{t-1} + \Phi_1^2 \epsilon_{t-2} + \dots$$

- $\Xi_j = \Phi_1^j$
- In the general VAR(P),

$$\Xi_j = \Phi_1 \Xi_{j-1} + \Phi_2 \Xi_{j-2} + \dots + \Phi_P \Xi_{j-P}$$

where $\Xi_0 = \mathbf{I}_k$ and $\Xi_m = \mathbf{0}$ for $m < 0$.

- In a VAR(2),

$$\mathbf{y}_t = \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \epsilon_t$$

$$- \Xi_0 = \mathbf{I}_k, \Xi_1 = \Phi_1, \Xi_2 = \Phi_1^2 + \Phi_2, \text{ and } \Xi_3 = \Phi_1^3 + \Phi_1 \Phi_2 + \Phi_2 \Phi_1.$$

- Confidence intervals are also somewhat painful
 - Explained in notes

Considerations for Shocks

- Simple bivariate VAR(1)

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- How you *shock* matters
- Depends on correlation between $\epsilon_{1,t}$ and $\epsilon_{2,t}$
- 3 methods
 - ▶ Ignore correlation and just shock $\epsilon_{j,t}$ with a 1 standard deviation shock
 - ▶ Use Cholesky to factor Σ and use $\Sigma^{1/2} \mathbf{e}_j$ where \mathbf{e}_j is a vector of zeros with 1 in the j^{th} position

$$\Sigma = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix} \quad \Sigma_C^{1/2} = \begin{bmatrix} 1 & 0 \\ .5 & .866 \end{bmatrix}$$

- Variable order matters
- ▶ “Generalized” impulse response that uses a projection method

Example of the different shocks

- Define the error covariance

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_x \sigma_y \rho \\ \sigma_x \sigma_y \rho & \sigma_y^2 \end{bmatrix}$$

- Standardized

$$\begin{bmatrix} 0 \\ \sigma_x \end{bmatrix}$$

- Cholesky

$$\Sigma_C^{1/2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sigma_x & 0 \\ \sigma_y \rho & \sigma_y \sqrt{1 - \rho^2} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \sigma_y \sqrt{1 - \rho^2} \end{bmatrix}$$

$$\begin{bmatrix} \sigma_x & 0 \\ \sigma_y \rho & \sigma_y \sqrt{1 - \rho^2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma_x \\ \sigma_y \rho \end{bmatrix}$$

Impulse Responses

- Federal Funds ordered first
- Response to Federal Funds Shock
- Cholesky factorization

