*VAR

Vector autoregression.

Def'n VAR(1) process of K endogenous variables is defined as

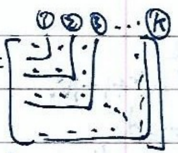
$$\underset{\sim}{y}_t = A\underset{\sim}{y}_{t-1} + \underset{\sim}{u}_t.$$

$$\underset{\sim}{y}_t = (y_{1t}, y_{2t}, \cdots, y_{Kt}),$$

$A_{K\times K}$ is a coefficient matrix,

$\underset{\sim}{u}_K$ is a K-dim WN process with time-invariant positive ~~coefficient~~ definite variance matrix $E(\underset{\sim}{u}_t \underset{\sim}{u}_t') = \Sigma_u$.

$[I-AB)\underset{\sim}{y}_t=\underset{\sim}{u}_t.$

A=


If det ⓢ > 0
$\forall s = 1, \cdots, k$,
then A is
⊕. d. (ift).

⊖. d.
⇔ det ⓢ < 0
s odd
and det ⓢ > 0
s even.

Note If $\underset{\sim}{y}_t \sim VAR(1)$, then

$$\underset{\sim}{y}_t = A\underset{\sim}{y}_{t-1} + \underset{\sim}{u}_t$$

$$= \cdots \qquad (VMA(\infty))$$

$$= \underset{\sim}{u}_t + A\underset{\sim}{u}_{t-1} + A^2\underset{\sim}{u}_{t-2} + A^3\underset{\sim}{u}_{t-3} + \cdots.$$

That is, for $\underset{\sim}{y}_t \sim VAR(1)$ to be stationary, $\lim_{j\to\infty} A^j = 0$. Mathematically, we require all K eigenvalues of A be less than one in abs. value.

$$VAR(1) \Rightarrow (I - A\underset{\sim}{B})\underset{\sim}{y}_t = \underset{\sim}{u}_t \sim N_k(\vec{0}, \Sigma_u)$$

$$VAR(p) \Rightarrow (I - A_1\underset{\sim}{B} - A_2\underset{\sim}{B}^2 - \cdots - A_p\underset{\sim}{B}^p)\underset{\sim}{y}_t = \underset{\sim}{u}_t + \vec{A}_0.$$

$$\Rightarrow \underset{\sim}{y}_t = A_1\underset{\sim}{y}_{t-1} + A_2\underset{\sim}{y}_{t-2} + \cdots + A_p\underset{\sim}{y}_{t-p} + \underset{\sim}{u}_t + \vec{A}_0,$$

Let $\dot{\underset{\sim}{y}}_t = \underset{\sim}{y}_t - \overset{E(y_t).}{\underset{\downarrow}{\vec{A}_0}}$. Then $\dot{\underset{\sim}{y}}_t \sim VAR(p)$

$\Leftrightarrow (I - A_1 \underset{\sim}{B} - A_2 \underset{\sim}{B}^2 - \cdots - A_p \underset{\sim}{B}^p) \dot{\underset{\sim}{y}}_t = \underset{\sim}{u}_t \overset{iid}{\sim} N_k(\vec{0}, \Sigma_u)$.

We can add const., trend, seasonal dummy variables, etc.

e.g. $\underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_t}_{\underset{\sim}{y}_t} = \underbrace{\begin{bmatrix} 5 \\ 10 \end{bmatrix}}_{\vec{A}_0} + \underbrace{\begin{bmatrix} .5 & .2 \\ -.2 & -.5 \end{bmatrix}}_{A_1} \underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-1}}_{\underset{\sim}{y}_{t-1}} + \underbrace{\begin{bmatrix} -.3 & -.7 \\ -.1 & .3 \end{bmatrix}}_{A_2} \underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-2}}_{\underset{\sim}{y}_{t-2}} + \underbrace{\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_t}_{\underset{\sim}{u}_t}$.

Write $\underset{\sim}{y}_t \sim VAR(p) \Leftrightarrow A_p(B) \underset{\sim}{y}_t = \underset{\sim}{u}_t$.

$\underset{\sim}{y}_t$ is stationary iff all the roots of
$\det(A_p(B)) = 0$ are greater than 1 in abs. value.

$A_p(B) = I_k - A_1 B - A_2 B^2 - \cdots - A_p B^p$.

- Companion form of VAR(p) process.

$\underset{\sim}{\xi}_t = A \underset{\sim}{\xi}_{t-1} + \underset{\sim}{V}_t$.

$\underset{\sim}{\xi}_t = \begin{bmatrix} \underset{\sim}{y}_t \\ \vdots \\ \underset{\sim}{y}_{t-(p-1)} \end{bmatrix}_{kp \times 1}$, $A = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}_{kp \times kp}$

~~$V_t = \begin{bmatrix} \underset{\sim}{y}_t \\ \vdots \\ \underset{\sim}{y}_{t-(p-1)} \end{bmatrix}$~~ $\underset{\sim}{V}_t = \begin{bmatrix} \underset{\sim}{u}_t \\ \underset{\sim}{0} \\ \vdots \\ \underset{\sim}{0} \end{bmatrix}_{kp \times 1}$.

The companion form of VAR(p) is VAR(1),
i.e. $\underset{\sim}{y}_t \sim VAR(p) \Rightarrow \underset{\sim}{\xi}_t \sim VAR(1)$,
where $\underset{\sim}{\xi}_t = \underset{\sim}{A} \underset{\sim}{\xi}_{t-1} + \underset{\sim}{V}_t$,

$\underset{\sim}{\xi}_t = \begin{bmatrix} \underset{\sim}{y}_t \\ \vdots \\ \underset{\sim}{y}_{t-(p-1)} \end{bmatrix}_{kp \times 1}$, $\underset{\sim}{A} = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}_{kp \times kp}$, $V_t = \begin{bmatrix} \underset{\sim}{u}_t \\ \underset{\sim}{0} \\ \vdots \\ \underset{\sim}{0} \end{bmatrix}_{kp \times 1}$

So from $\underset{\sim}{\xi}_t = \underset{\sim}{A}\,\underset{\sim}{\xi}_{t-1} + V_t$, if all the abs. value of λ's of A are less than 1, then $\underset{\sim}{y}_t$ is stationary!

e.g. $$\underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_t}_{\underset{\sim}{y}_t} = \underbrace{\begin{bmatrix} .5 & .2 \\ -.2 & -.5 \end{bmatrix}}_{A_1} \underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-1}}_{\underset{\sim}{y}_{t-1}} + \underbrace{\begin{bmatrix} -.3 & -.7 \\ -.1 & .3 \end{bmatrix}}_{A_2} \underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-2}}_{\underset{\sim}{y}_{t-2}} + \underset{\sim}{u}_t$$

$\underset{p}{\text{VAR}(2)}$ $k=2$.

$$A = \begin{bmatrix} A_1 & A_2 \\ I & 0 \end{bmatrix} = \begin{bmatrix} .5 & .2 & -.3 & -.7 \\ -.2 & -.5 & -.1 & .3 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

```
> r <- eigen(A)
> r$values
> abs(r$values)
[1] 0.81~    .59~   .572~   .572~ ,
```

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} = \begin{bmatrix} .5 & .2 & -.3 & -.7 \\ -.2 & -.5 & -.1 & .3 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ 0 \\ 0 \end{bmatrix}.$$

- Order selection
  - → sequential LRT : VAR(p) vs VAR(p-1).
  - → Information criterion (BIC, etc.)

- Analysis (analogous to Box-Jenkin's Approach)

Specification + estimation of reduced form VAR model

↓ ↑ model rejected

Model checking

↓ model accepted

Forecasting / Causality analysis / Structural specification test.

Structural specification test. → IRA (impulse response analysis) / Forecast error variance decomposition

Granger causality

- Idea: $y_{2,t}$ does not cause $y_{1,t}$ if the distribution of $y_{1,t} \mid$ (past values of both $y_{1,t}$ and $y_{2,t}$) = dist. of $y_{1,t} \mid$ (past values of only $y_{1,t}$).

Testing the entire distribution of $y_{1,t}$ is very difficult. Instead, we see that if conditional mean of $y_{1,t}$ depends on past values of $y_{2,t}$.

Consider VAR(p):

$$\underset{\sim}{y_t} = \vec{\alpha} + \sum_{j=1}^{p} A_j \underset{\sim}{y_{t-j}} + \underset{\sim}{a_t}.$$

Let $A_j = \left[\begin{array}{c|c} \phi_{11}^{(j)} & \boxed{\phi_{12}^{(j)}} \\ \hline \phi_{21}^{(j)} & \phi_{22}^{(j)} \end{array}\right]_{2\times 2}$. $\begin{bmatrix} y_{1,t-j} \\ y_{2,t-j} \end{bmatrix} = \underset{\sim}{y_{t-j}}$

If $y_{2,t}$ does not $\boxed{\text{Granger-cause}}$ $y_{1,t}$, then all of the $\phi_{12}^{(j)}$'s, $j=1,\cdots,p$, must be 0.

If $y_{1,t}$ does not Granger-cause $y_{2,t}$, then all of the $\phi_{21}^{(j)}$'s must be 0.

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \sum_{j=1}^{p} \begin{bmatrix} \phi_{11}^{(j)} y_{1,t-j} + \phi_{12}^{(j)} y_{2,t-j} \\ \phi_{21}^{(j)} y_{1,t-j} + \phi_{22}^{(j)} y_{2,t-j} \end{bmatrix} + \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix}$$

$$\Rightarrow y_{1,t} = \alpha_1 + \sum_{j=1}^{p} \phi_{11}^{(j)} y_{1,t-j} + \sum_{j=1}^{p} \phi_{12}^{(j)} y_{2,t-j} + a_{1,t}, \text{ and}$$

$$y_{2,t} = \alpha_2 + \sum_{j=1}^{p} \phi_{21}^{(j)} y_{1,t-j} + \sum_{j=1}^{p} \phi_{22}^{(j)} y_{2,t-j} + a_{2,t}.$$

$x_t \sim ARMA(p,q)$,

x_t stationary iff $\Phi_p(B)$'s roots all outside ⊙ iff $x_t = \Psi_\infty(B) a_t$ and $\sum_{j\ge0}^{\infty} |\psi_j| < \infty$.

x_t invertible iff $\Theta_q(B)$'s roots " iff $\Pi_\infty(B) x_t = a_t$ and $\sum_{j\ge0}^{\infty} |\pi_j| < \infty$.

· LRT

The first eqn:

$$y_{1,t} = \alpha_1 + \sum_{j=1}^{p} \phi_{11}^{(j)} y_{1,t-j} + \sum_{j=1}^{p} \phi_{12}^{(j)} y_{2,t-j} + a_{1,t}$$

From here, $H_0 : \phi_{12}^{(j)} = 0 \ \forall j$ vs. H_1: at least one $\neq 0$.

$$y_{1,t} = \alpha_1 + \sum_{j=1}^{p} \phi_{11}^{(j)} y_{1,t-j} + e_{1,t} \quad (H_0 \text{ true})$$

$$y_{1,t} = \alpha_1 + \sum_{j=1}^{p} \phi_{11}^{(j)} y_{1,t-j} + \sum_{j=1}^{p} \phi_{12}^{(j)} y_{2,t-j} + \varepsilon_{1,t} \quad (H_1 \text{ true}).$$

Obtain ML or OLS from both.

$$n(\log(\det \tilde{\Sigma}) - \log(\det \hat{\Sigma})) \sim \chi^2_p ,$$

where $\tilde{\Sigma}$ and $\hat{\Sigma}$ denote the residual variance ~~matrix~~ matrices.

(i.e. $E(\underset{\sim}{e}_t \underset{\sim}{e}_t')$ and $E(\underset{\sim}{\varepsilon}_t \underset{\sim}{\varepsilon}_t')$).

· Portmanteau test for Granger causality

$x_t \sim ARMA(p,q)$

$\Phi_p(B) x_t = \Theta_q(B) a_t$

AR(p) is invertible.

MA(q) is stationary.

AR(p) is stationary iff roots of $\Phi_p(B)$ lie outside of the unit circle.

MA(q) is invertible iff $\Theta_q(B)$'s roots are lying outside ⊙.

Let $\{x_t\}$ and $\{y_t\}$ be causal (stationary) and invertible univariate ARMA:

$$\Phi_X(B) \dot{x}_t = \Theta_X(B) u_t, \quad u_t \overset{iid}{\sim} N(0, \sigma_u^2)$$

$$\Phi_Y(B) \dot{y}_t = \Theta_Y(B) v_t, \quad v_t \overset{iid}{\sim} N(0, \sigma_v^2).$$

$$\rho_{uv}(k) = \frac{E(u_t v_{t+k})}{\sqrt{E(u_t^2) E(v_t^2)}} \left(= \frac{\overset{= \gamma_{uv}(k)}{\gamma_{uv}(t,t+k)}}{\sqrt{\underbrace{\gamma_u(t,t)}_{= \gamma_u(0)} \underbrace{\gamma_v(t,t)}_{= \gamma_v(t+k,t+k) = \gamma_v(0)}}} \right)$$

$$\boxed{\begin{aligned} &* H_0 : X \text{ does not cause } Y. \\ &: Q_L = n^2 \sum_{k=0}^{L} (n-k)^{-1} \gamma_{uv}^2(k) \sim \chi^2_{L+1}. \end{aligned}}$$

$$\gamma_{uv}(k) = \frac{\rho_{uv}(k)}{\rho_{uv}(0)}.$$

(Modern portfolio theory: $R_p = w_1 r_1 + \cdots + w_p r_p$. Minimize $Var(R_p)$ s.t. $E(R_p) = \mu_p$. $\min Var(R_p) = \sum_i \sum_j w_i w_j \sigma_{ij}$, $\sum w_i = 1$, $w_i \geq 0$ $\forall i = 1, \cdots, p$.)

* Modelling Private Assets

Dimson's model

Def'n Appraisal return: $y_t = \sum_{i=0}^{p} w_i r_{t-i}$, $\sum w_i = 1$, $w_i \geq 0$.

Def'n Linear factor model: $r_t - r_f = \alpha + \sum_j^k \beta_j f_{j,t} + u_t$.

$$\Rightarrow y_t = r_t + \alpha + \sum_{i=0}^{p} \sum_j^k w_i \beta_j f_{j,t-i} + \xi_t$$

(ξ_t is serially correlated).

r = contemporaneous and lagged economic returns
w = weight on the corresponding economic return.

- What's wrong with appraisal returns?
Appraisal returns are found to be smoothed and serially correlated ("stale-pricing bias"). This leads to underestimations of volatilities and inaccurate statistical inference on factor exposures.

- Two-step approaches: GLM and ~~Pedersen~~ PPH.
→ Getmansky, Lo, and Makarov (GLM): constrained MA(m).

$$y_t = \sum_{i=0}^{m} \theta_i a_{t-i}, \quad \theta_0 = 1; \ \theta_i \geq 0, \ (i = 1, \cdots, m).$$

$$\Rightarrow y_t = \frac{\theta_0}{\sum \theta_i} \sum \theta_i a_t + \frac{\theta_1}{\sum \theta_i} \sum \theta_i a_{t-1} + \cdots + \frac{\theta_m}{\sum \theta_i} \sum \theta_i a_{t-m}.$$

That is, $w_i = \theta_i / \sum_{j=0}^{m} \theta_j$, $r_t = a_t \sum_{j=0}^{m} \theta_j$.
$i = 0, \cdots, m$.

~~→ Pedersen, Page, and He (PPH):~~

~~From $y_t = r_t + \alpha + \sum_{i=0}^{p} \sum_j^k w_i \beta_j f_{j,t-i} + \xi_t$~~

→ Pedersen, Page, and He (PPH):

So $y_t = \sum_{i=0}^{m} w_i r_{t-i}$, $r_t - r_f = \alpha + \sum_j^k \beta_j f_{j,t} + u_t$.

$\xi_t = \sum_i w_i u_{t-i}$ Thus $y_t = r_f + \alpha + \sum_i^m \sum_j^k w_i \beta_j f_{j,t-i} + \xi_t$.

Ignoring [illegible] r_f, we have:

~~$y_t = \alpha + \sum_{j=1}^{m} \beta_j$~~ $y_t = \alpha + \sum_j^k \beta_j \underbrace{\sum_i^m w_i f_{j,t-i}}_{x_{j,t}} + \xi_t$.

① Estimate w_i based on GLM.

② Calculate "smoothed" factor as

$$x_{j,t} = \sum_i^m \hat{w}_i f_{j,t-i}, \ i = 1, \cdots, m, \ j = 1, \cdots, k.$$

③ Estimate factor loadings using the following equation:

$$y_t = \alpha + \sum_j^k \beta_j x_{j,t} + \xi_t,$$

where error terms are serially correlated.

Note: · Both methods retrive $\{w_i\}$ using only appraisal returns history.

· Both methods have the problem of Errors In Variable (EIV).

We see that PPH regression residuals are serially correlated. More advanced approaches are needed for estimating regression w/ serially correlated errors.

· ARMAX/TFNM

pg. 19 of ppt $y_t = \alpha + \beta_m X_{mkt} + \beta_b X_{bond} + \sum_{i=0}^{5} \theta_i u_{t-i}$.

- Appraisal formula w/ linear factor model

$$y_t = \alpha + \Sigma_{i=0}^{m} \Sigma_{j=1}^{k} \beta_{ij} f_{j,t-i} + \xi_t$$

$$\beta_j = \sum_{i=0}^{m} \beta_{ij} \quad , j = 1, \cdots, k.$$

$$\beta_{ij} = w_i \beta_j \Rightarrow w_i = \frac{\beta_{ij}}{\Sigma_i \beta_{ij}} = \frac{\beta_{ij}}{\beta_j}.$$

→ Challenges

① An ad-hoc one-step regression usually requires $\overset{a}{\wedge}$ large # of params.

② The # of statistically significant lagged factors tends to be different across factors, i.e. different m for different factors.

③ $sgn(\beta_{1j}) = \cdots = sgn(\beta_{mj}) \; \forall j = 1, \cdots, k.$

④ All factors are smoothed by the same appraisal weights, i.e. $\frac{\hat{\beta}_{i1}}{\hat{\beta}_1} = \frac{\hat{\beta}_{i2}}{\hat{\beta}_2} = \cdots = \frac{\hat{\beta}_{ik}}{\hat{\beta}_k}, \; i = 1, \cdots, m.$

⑤ Empirical studies show that the estimates of factor exposures for private equity based on appraisal returns tend to be smaller than practitioners' expectation or those estimated using the cash flow approach.

- PROPOSED METHOD IN ppt.

$$y_t = \alpha + \Sigma_{i=0}^{m} \Sigma_{j=1}^{k} w_i \beta_j f_{j,t-i} + \Sigma_{i=0}^{m} w_i u_{t-i}$$

"appraisal w/ k factors and m appraisal lags"

$\beta_{ij} f_{j,t-i}$

$\boxed{\beta_j}$

$= \frac{\beta_{ij}}{\beta_j} f_{j,t-i} \beta_j$

$= w_i \beta_j f_{j,t-i}$

1. constrained linear reg model w/ MA errors
2. multiple-input distributed lag model
3. The above model can be formulated as a state-space model and estimated using the Kalman Filter.

$$\frac{\beta_{ij}}{\sum_i \beta_{ij}} = \frac{\beta_{ij}}{\beta_j} = w_i$$

So proposed model is (again):

$$y_t = \alpha + \sum_{i=0}^{m} w_i \sum_{j=1}^{k} \beta_j f_{j,t-i} + \sum_{i=0}^{m} w_i u_{t-i}$$

"m=2". Maybe used in private equity!

e.g. single appraisal factor, 2 lags.
(appraisal w/ 1 factor)

$$\Rightarrow m=2,\ k=1.$$

$$\Rightarrow y_t = \alpha + \sum_{i=0}^{2} w_i \sum_{j=1}^{1} \beta_j f_{j,t-i} + \sum_{i=0}^{2} w_i u_{t-i}$$

$$= \alpha + \sum_{i=0}^{2} w_i \beta f_{t-i} + \sum_{i=0}^{2} w_i u_{t-i}.$$

$$= \alpha + \beta \textstyle\sum_{i=0}^{2} w_i f_{t-i} + e_t,$$

$$e_t = w_0 u_t + w_1 u_{t-1} + w_2 u_{t-2},$$

$$u_t \sim N(0,Q).$$

(f_t = factor return)

So what exactly is a state-space model and state eqn? To be cont.

State eqn:

$$\begin{bmatrix} u_t \\ u_{t-1} \\ u_{t-2} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u_{t-1} \\ u_{t-2} \\ u_{t-3} \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \\ 0 \end{bmatrix}.$$

$$y_t = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} \cdot \begin{bmatrix} u_t \\ u_{t-1} \\ u_{t-2} \end{bmatrix} + \begin{bmatrix} \alpha \\ \beta w_0 \\ \beta w_1 \\ \beta w_2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ f_t \\ f_{t-1} \\ f_{t-2} \end{bmatrix} + \varphi_t,$$

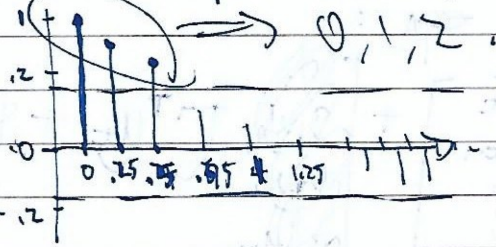
$$u_t \sim N(0,Q),\quad \varphi_t \sim N(0,R),\ R=0.$$

- Parametric weight function.
  For alternative assets w/ long appraisal lags (e.g. real estate), the estimation of m appraisal weights may not be feasible. The explicitly estimated appraisal weights tend to be sensitive to extreme obs..

In view of the mixed data sampling (MIDAS) reg. literature, we could use a parametric weight function such as normalized exponential Almon lag polynomial function to avoid the above issues.

- Empirical studies (e.g.)

→ We can identify the # of appraisal lags (m) using the sample ACF plot.



$\Longrightarrow 0, 1, 2.$ $\underline{m=2.}$

$$j = \begin{cases} 1 & \text{if mkt} \\ 2 & \text{if smb} \\ 3 & \text{if hml} \end{cases}$$

Unconstrained regression: $y_t = \alpha + \sum_{i=0}^{2}\sum_{j=1}^{3} w_i \beta_j f_{j,t-i} + e_t$

$$= \alpha + \sum_{i=0}^{2}\sum_{j=1}^{3} \beta_{ji} f_{j,t-i} + e_t$$

$$= \alpha + \sum_{i=0}^{2}\left[\beta_{m,i} f_{m,t-i} + \beta_{smb,i} f_{smb,t-i} + \beta_{hml,i} f_{hml,t-i}\right] + e_t.$$

$$\beta_m = \sum_{i=0}^{2} \beta_{m,i}, \quad \beta_{j,i} = w_i \beta_j \Rightarrow \frac{\beta_{ji}}{\beta_j} = w_i \ \forall j.$$

(For a fixed j, $\mathrm{sgn}(\beta_{ji})$ are all the same for each $j = 0, \dots, m$;
for a fixed i, $\frac{\hat{\beta}_{ji}}{\hat{\beta}_j} = \hat{w}_i$ are all the same (in value).).

$w_i = \frac{\beta_{ji}}{\beta_j}$, $w_i \beta_j = \beta_{ji}$.

$\beta_j = \sum_i \beta_{ji}$.

↓

$$= \alpha + \sum_{i=0}^{m} w_i \sum_{j=1}^{k} \beta_j f_{j,t-i} + \sum_{i=0}^{m} w_i \xi_{t-i}$$

$$y_t = \alpha + \sum_{i=0}^{m} \sum_{j=1}^{k} \beta_{ji} f_{j,t-i} + \sum_{i=0}^{m} w_i \xi_{t-i}$$

Problems:

① Require a large # of param.

② Difference m for difference factors

③ $\mathrm{sgn}(\beta_{ji})$ same $\forall i = 1, \dots, m$ (when j fixed)

④ $\frac{\beta_{ji}}{\beta_j}$ same $\forall j = 1, \dots, k$ (when i fixed)

⑤ Underestimate of factor exposures than practitioner's expectation and cash flow approach.

Appendix: appraisal w/ k factors and m lags

1) State eqn.:

$$\begin{bmatrix} \xi_t \\ \xi_{t-1} \\ \vdots \\ \xi_{t-m} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 \cdots & 0 & 0 \\ & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & 0 & \\ 0 & 0 & 0 \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \xi_{t-1} \\ \xi_{t-2} \\ \vdots \\ \xi_{t-(m+1)} \end{bmatrix} + \begin{bmatrix} \xi_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

2) Observation eqn.:

$$y_t = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_m \end{bmatrix} \cdot \begin{bmatrix} \xi_t \\ \xi_{t-1} \\ \vdots \\ \xi_{t-m} \end{bmatrix}_{(m+1)\times 1} + \cancel{\begin{bmatrix} \alpha \\ \beta_1 w_0 \\ \beta_1 w_1 \\ \vdots \\ \beta_1 w_m \\ \vdots \end{bmatrix}} \underset{\sim}{\Gamma}^T \underset{\sim}{u_t} + \varphi_t.$$

$\xi_t \sim N(0, Q)$,

$\varphi_t \sim N(0, R)$,

$R = 0$

$$\underset{\sim}{\Gamma} = \begin{bmatrix} \alpha \\ \beta_1 w_0 \\ \beta_1 w_1 \\ \vdots \\ \beta_1 w_m \\ \beta_2 w_0 \\ \beta_2 w_1 \\ \vdots \\ \beta_2 w_m \\ \beta_3 w_0 \\ \vdots \\ \beta_k w_0 \\ \vdots \\ \beta_k w_m \end{bmatrix}_{(k(m+1)+1)\times 1}, \quad \underset{\sim}{u_t} = \begin{bmatrix} 1 \\ f_{1,t} \\ f_{1,t-1} \\ \vdots \\ f_{1,t-m} \\ f_{2,t} \\ f_{2,t-1} \\ \vdots \\ f_{2,t-m} \\ f_{3,t} \\ \vdots \\ f_{k,t} \\ \vdots \\ f_{k,t-m} \end{bmatrix}_{(k(m+1)+1)\times 1}$$

* Cointegration

Recall $\{Z_t\}$ is $I(d)$, or an integrated process of order d if $(1-B)^d Z_t$ is stationary and invertible.
If $\{Z_t\}$ is stationary, then $Z_t \sim I(0)$.

Motivation: we used Box-Jenkin's approach ~~to~~ ~~to~~ and differencing to solve the issue of nonstationarity.
Cointegration is another technique to model nonstationary (multivariate) time series.

Intuition:
① "Balance" the (linear) reg. equation.
② Check if time series share the same "source" of the $I(1)$'ness, or time series move together in the long run.

Recall Box-Jenkin's Approach, 3 stages.

Start {
· TS realization
· understand a problem
· collect + plot data

↓

① Identify a prelim time series model
→ perform differencing + transformations to transform data into stationary
→ identify prelim ARMA(p,q) models using $\widehat{ACF}$ and $\widehat{PACF}$.

→ if the fitted model fails diagnostic test
→ identify another model.

② Estimate the model param.
→ MoM, MLE, Kalman Filter, etc.

③ Diagnose model adequacy
→ Examine if the res. of the fitted model are approx. uncorrelated
↳ Stop → if passes
→ use model for analysis

– Cointegration

<u>Def'n</u> Consider a multivariate time series Z_t.
If $z_{i,t}$ $\forall i$ are $I(1)$ processes but a non-trivial linear combination $\underset{\sim}{\beta}' \underset{\sim}{Z}_t$ is $I(0)$, then $\underset{\sim}{Z}_t$ is said to be <u>cointegrated of order one.</u>

<u>Def'n</u> Such $\underset{\sim}{\beta}$ is called a <u>cointegrating vector.</u>

<u>Def'n</u> If $z_{i,t}$ are $I(d)$ nonstationary and $\underset{\sim}{\beta}' \underset{\sim}{Z}_t$ is $I(h)$ with $h < d$, then we say $\underset{\sim}{Z}_t$ is <u>cointegrated</u>.

In practice, the case of $h=0, d=1$ is of major interest.
Thus, cointegration often means that a linear combination of individually-unit-root-nonstationary time series becomes a stationary and invertible series.

Properties of $I(0)$ and $I(1)$

① $X_t \sim I(0) \Rightarrow a + bX_t \sim I(0)$.
$X_t \sim I(1) \Rightarrow a + bX_t \sim I(1)$.

② $X_t, Y_t \sim I(0) \Rightarrow aX_t + bY_t \sim I(0)$.

③ $X_t \sim I(0)$ and $Y_t \sim I(1)$
$\Rightarrow aX_t + bY_t \sim I(1)$.

④ $X_t, Y_t \sim I(1) \Rightarrow aX_t + bY_t \sim I(1)$ in general.

· Common trends
Cointegration relationships : cointegrated variables sharing common stochastic trends.

e.g. Let $W_t \sim ARIMA(p,1,q)$, $\mathcal{X}_t \sim ARMA(p,q)$, $\mathcal{Y}_t \sim ARMA(p,q)$.
$X_t = \alpha W_t + \mathcal{X}_t$, $Y_t = W_t + \mathcal{Y}_t$.

Consider $z_t = X_t - \alpha Y_t$.

Then $$\begin{aligned} Z_t &= X_t - \alpha Y_t \\ &= \alpha W_t + \chi_t - \alpha(W_t + \psi_t) \\ &= \cancel{\alpha W_t} + \chi_t \cancel{- \alpha W_t} - \alpha \psi_t \\ &= \chi_t - \alpha \psi_t \sim I(0) \text{ by property.} \end{aligned}$$

That is, $\underset{\sim}{Z}_t = \begin{bmatrix} X_t \\ Y_t \end{bmatrix}$ and $\beta = \begin{bmatrix} 1 \\ -\alpha \end{bmatrix}$ ~~gives~~ give

$\beta' \underset{\sim}{Z}_t = X_t - \alpha Y_t \sim I(0)$, a stationary process.
So $\beta = \begin{bmatrix} 1 \\ -\alpha \end{bmatrix}$ is a cointegrating vector.

* If two $I(1)$ processes have a common $I(1)$ trend, and $I(0)$ idiosyncratic components, then we say they are cointegrated.

Def'n X_t and Y_t are cointegrated one another if both are $I(d)$ processes and have a common $I(d)$ trend, and $I(0)$ idiosyncratic components respectively.

e.g. 1 $X_t = W_t + V_t$, $Y_t = W_t + U_t$, $Z_t = W_t + S_t$,
$U_t, V_t, S_t \sim I(0)$ and $W_t \sim I(1)$.

Yes, 0 is allowed.

Then $(1)X_t + (-1)Y_t + (0)Z_t$
$= V_t + \cancel{W_t} \cancel{- W_t} - U_t = V_t - U_t \sim I(0)$.

Also, $(0)X_t + (1)Y_t + (-1)Z_t$
$= W_t + U_t \cancel{+} - W_t - S_t = U_t - S_t \sim I(0)$.

e.g. 2 ~~$X_t = W_t + U_t$, $Y_t =$~~
$X_t = W_t + R_t + V_t$, $Y_t = W_t + U_t$, $Z_t = R_t + S_t$.
$W_t, R_t \sim I(1)$, $U, V, S \sim I(0)$.

Then $(1)X_t + (-1)Y_t + (-1)Z_t$
$= \cancel{W_t} + \cancel{R_t} + V_t - \cancel{W_t} - U_t \cancel{+} - \cancel{R_t} - S_t$
$= V_t - U_t - S_t \sim I(0)$.

e.g. 1 $\underset{\sim}{I_t} = \begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} W_t + V_t \\ W_t + U_t \\ W_t + S_t \end{bmatrix}$

$\beta = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. $\beta \cdot I_t = aW_t + aV_t + bW_t + cW_t + cS_t$
$= (a+b+c)W_t + \sim\sim$.

So $\text{span}\{\beta \mid a+b+c\}$ is a space of all β's (cointegrating vectors of I_t).

e.g. $\underset{\sim}{I_t} = \begin{bmatrix} W_t + R_t + V_t \\ W_t + U_t \\ R_t + S_t \end{bmatrix}$, $\beta = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$

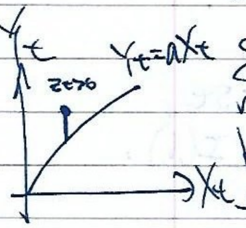
$\Rightarrow \underline{aW_t} + \underline{aR_t} + aV_t + \underline{bW_t} + bU_t + \underline{cR_t} + cS_t$
$= (a+b)W_t + (a+c)R_t + \sim\sim$.

So any $\beta \in \text{span}\{(a,b,c) \mid \underline{a+b=0, a+c=0}\}$ is a cointegrating vector of $\underset{\sim}{I_t}$.

$a+b=0$
$a+c=0$
$2a+b+c=0$
$b+c=-2a$

— ECM (error correction model)

Let $Z_t = Y_t - aX_t$ denote the deviation from the long-run equilibrium. If the system is going to return to long-run equilibrium, the short-run movements of the variables (at some of them) must be respond to the magnitude of disequilibrium. Hence, the path of a cointegrated system is influenced by the extend of deviation from the long-run equilibrium.



e.g.

$$\begin{bmatrix} \nabla r_{s,t} \\ \nabla r_{L,t} \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} \alpha_s & -\alpha_s\beta \\ -\alpha_L & \alpha_L\beta \end{bmatrix} \begin{bmatrix} r_{L,t-1} \\ r_{s,t-1} \end{bmatrix}$$

$\alpha_s, \alpha_L > 0$
$\varepsilon_{i,t} \sim WN(0, \sigma_i^2)$
$(i = s, L)$

$$+ \sum_i \begin{bmatrix} a_{11}^{(i)} & a_{12}^{(i)} \\ a_{21}^{(i)} & a_{22}^{(i)} \end{bmatrix} \begin{bmatrix} \nabla r_{s,t-i} \\ \nabla r_{L,t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{s,t} \\ \varepsilon_{L,t} \end{bmatrix}$$

$$\begin{bmatrix} -\alpha_s\beta & \alpha_s \\ \alpha_L\beta & -\alpha_L \end{bmatrix}\begin{bmatrix} r_{s,t-1} \\ r_{L,t-1} \end{bmatrix}$$

$$\nabla \underset{\sim}{r_t} = \vec{\alpha} + \overset{\Pi}{\cancel{}}\underset{\sim}{r_{t-1}} + \sum_i A_i \nabla \underset{\sim}{r_{t-i}} + \underset{\sim}{\varepsilon_t} .$$

$$\nabla \underset{\sim}{r_t} = \vec{\alpha} + \Pi \underset{\sim}{r_{t-1}} + \sum_i A_i \nabla \underset{\sim}{r_{t-i}} + \underset{\sim}{\varepsilon_t} , \quad \Pi = \begin{bmatrix} -\alpha_s\beta & \alpha_s \\ \alpha_L\beta & -\alpha_L \end{bmatrix} \ (*)$$

(similar to ADF eqn. : $\nabla x_t = \tau^T DR_t + \pi x_{t-1} + \sum_{j=1}^{k} \overset{k=p-1}{\cancel{}_j} \nabla x_{t-j} + a_t$)

$\tau = (\alpha_1, \alpha_2, \dots)$

$DR_t = (t, t^2, \dots)$

Granger Representation Theorem

THM If X_t and Y_t are cointegrated (i.e. $\exists a, b$ s.t. $aX_t + bY_t \sim I(0)$), then there exists an error correction model (ECM) representation.

Note Cointegration is a necessary condition for ECM and vise versa.

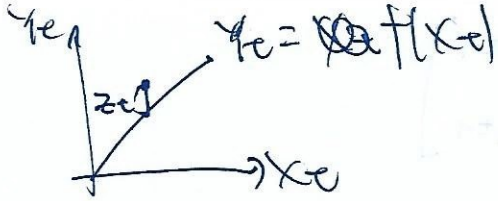
Vector autoregressions on differenced $I(1)$ processes will be a misspecification if the ~~components~~ component series are cointegrated. Engle and Granger (1987) showed that an equilibrium specification is missing from a VAR representation. However, when lagged disequilibrium terms are included as explanatory variables, the model becomes well specified. Such a model is called an ECM because the model is structured so that short-run deviation from the long-run equilibrium will be corrected. (just like in (*)).

- The procedure of E and G ('87)

① Test whether X_t and Y_t are $I(1)$ using URT.

② If both are $I(1)$, regress one series against the other using least squares.

③ Run an URT on regression residuals. If residuals are stationary, these two series are cointegrated.



(The reg line indicates the long-run equilibrium relationship btwn two variables. The disequilibrium term is simply the regression residuals)

④ Finally, consider the following ECM;

$$\nabla\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} + \begin{bmatrix} \hat{\alpha}\rho_1 & \hat{\alpha}\rho_1 \\ -\hat{\alpha}\rho_2 & \rho_2 \end{bmatrix}\begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix}$$

$$+ \begin{bmatrix} \beta_{x,1} \\ \beta_{x,2} \\ \vdots \\ \ \end{bmatrix} \cdot \nabla\begin{bmatrix} X_{t-1} \\ X_{t-2} \\ \ \\ \ \end{bmatrix} + \begin{bmatrix} \beta_{Y,1} \\ \beta_{Y,2} \\ \vdots \\ \vdots \end{bmatrix} \cdot \nabla\begin{bmatrix} Y_{t-1} \\ Y_{t-2} \\ \vdots \\ \vdots \end{bmatrix} + \begin{bmatrix} \varepsilon_{X,t} \\ \varepsilon_{Y,t} \end{bmatrix}.$$

$$\sum_{i=1}^{\infty} \begin{bmatrix} \beta_{x,i} \\ \beta_{Y,i} \end{bmatrix} \cdot \nabla\begin{bmatrix} X_{t-i} \\ Y_{t-i} \end{bmatrix}$$

$$\Rightarrow \nabla\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} + \begin{bmatrix} -\hat{\alpha}\rho_1 & \rho_1 \\ -\hat{\alpha}\rho_2 & \rho_2 \end{bmatrix}\begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \sum_{i=1}^{\infty}\begin{bmatrix} \beta_{x,i} \\ \beta_{Y,i} \end{bmatrix} \cdot \nabla\begin{bmatrix} X_{t-i} \\ Y_{t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{X,t} \\ \varepsilon_{Y,t} \end{bmatrix}$$

$$\nabla\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} + \begin{bmatrix} -\hat{\alpha}\rho_1 & \rho_1 \\ -\hat{\alpha}\rho_2 & \rho_2 \end{bmatrix}\begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \sum_{i=1}^{\infty}\begin{bmatrix} \beta_{x,i} & \beta_{Y,i} \\ \gamma_{x,i} & \gamma_{Y,i} \end{bmatrix}\begin{bmatrix} \nabla X_{t-i} \\ \nabla Y_{t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{X,t} \\ \varepsilon_{Y,t} \end{bmatrix}$$

→ Why using E-G Method

① It is very straightforward to implement and to interpret the Engle-Granger procedure.

② From the risk management POV, the EG criterion that minimizes variance is usually more important than the Johansen criterion that maximizes stationarity.

③ Sometimes there is a natural choice of dependent variables in the cointegrating regressions, for example, in equity index tracking.

→ Comments

① Assumption implicitly imposed in this approach : the E-G procedure is only applicable to systems with more than two variables in a very special circumstances.

② Another way to test cointegration :
The Johansen procedure ('88) seeks the linear combination which is most stationary whereas the E-G tests seek the l.c. having minvar. The Johansen tests are a multivariate generalization of the URTs.

③ The presence of change points will affect the effectiveness of cointegration analysis.

* Pairs trading based on cointegration

Idea If two ~~assets~~ asset prices are coinegrated, then the value of a wisely built portfolio (spread) btwn these two assets is stationary/mean-reverting.

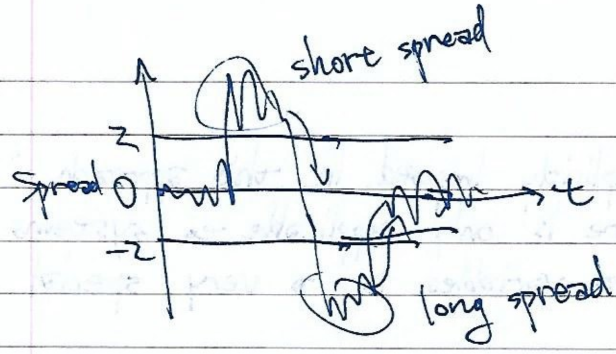
"Buy low and sell high (above its mean)."

Pairs trading is executed when spread diverges too much from its mean.

Warning : cointegration is the long-run relationship so the constructed spread may diverge substantially from these relationship in the short run.

Random Walk: $I(1)$
$X_t - X_{t-1} = a_t$
$(1-B)X_t = a_t$
$X_t \sim I(1)$.

- Pairs trading in stocks
  ① Construct a portfolio consisting of two stocks : S_1 and S_2.
  ② Define spread $:= S_1 - bS_2$.
  ③ Avoid guessing trends and explore the mkt inefficiency in a statistical sense. Difference methods for constructing spreads available.



A good spread is mean-reverting.
i.e. ① Constant mean
② Bounded, or finite second moment
⟺ Weak stationarity!

i.e. A good spread is weakly stationary.

- Procedure in Pairs Trading
  ① Find pairs of stocks
  ② Determine the trading signal.
  ③ Determine how to size your position as well as portfolio construction.
  ④ Size your position (risk / capital mgmt).
  ⑤ Calculate transaction cost.
  ⑥ Backtest if your strategy works
  ⑦ The better implementations and less assumptions used by your model, the more successful your trading strategy.

※ Multivariate Time Series

→ VAR(p) ⟹ VMA(∞).

→ Portmanteau tests

$$Q_h = T\sum_{j=1}^{h} tr(\hat{C}_j^T \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}) \sim \chi^2_{k^2(h-n)}, \quad \hat{C}_i = \frac{1}{T}\sum_{t=i+1}^{T} \underset{\sim}{\hat{u}_t} \underset{\sim}{\hat{u}_{t-i}}$$

$$Q_h^* = T^2 \sum_{j=1}^{h} \underset{\uparrow \text{ for small } h}{\frac{1}{T-j}} tr(\hat{C}_j^T \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1})$$

They test the absense of up to order h serially correlated disturbances in a stationary VAR(n).

Companion: $\begin{bmatrix} \underset{\sim}{z_t} \\ \underset{\sim}{z_{t-1}} \\ \vdots \\ \underset{\sim}{z_{t-(n-1)}} \end{bmatrix} = \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{n-1} & \Phi_n \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} \underset{\sim}{z_{t-1}} \\ \underset{\sim}{z_{t-2}} \\ \vdots \\ \underset{\sim}{z_{t-n}} \end{bmatrix} + \begin{bmatrix} \underset{\sim}{u_t} \\ \vec{0} \\ \vdots \\ \vec{0} \end{bmatrix}$

→ Granger causality

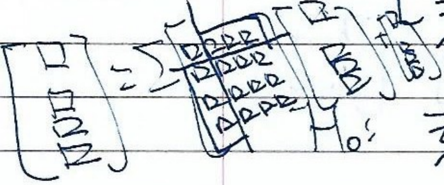
$(I - A_1 B - A_2 B^2 - \cdots - A_p B^p) \underset{\sim}{z_t} = \underset{\sim}{u_t}$.

$\underset{\sim}{z_t} \sim VAR(n)$ $\quad A_p(B) \underset{\sim}{z_t} = \underset{\sim}{u_t}$

$\Rightarrow \underset{\sim}{z_t} = \sum_{i=1}^{n} \Phi_i \underset{\sim}{z_{t-i}} + \underset{\sim}{u_t}$

Split $\underset{\sim}{z_t}$ into two subvectors $\underset{\sim}{z_{1t}}$ and $\underset{\sim}{z_{2t}}$ with each dim = K_1 and K_2, $K_1 + K_2 = K$.

So $\begin{bmatrix} \underset{\sim}{z_{1,t}} \\ \underset{\sim}{z_{2,t}} \end{bmatrix} = \sum_{i=1}^{n} \begin{bmatrix} \Phi_{11}^{(i)} & \boxed{\Phi_{12}^{(i)}} \\ \boxed{\Phi_{21}^{(i)}} & \Phi_{22}^{(i)} \end{bmatrix} \begin{bmatrix} \underset{\sim}{z_{1,t-i}} \\ \underset{\sim}{z_{2,t-i}} \end{bmatrix} + \begin{bmatrix} \underset{\sim}{u_{1,t}} \\ \underset{\sim}{u_{2,t}} \end{bmatrix}$



H_0: $\underset{\sim}{z_{1,t}}{}_{k_1 \times 1}$ doesn't G-cause $\underset{\sim}{z_{2,t}}{}_{k_2 \times 1}$

$\Phi_{21}^{(i)} = 0 \ \forall i$.

H_1: $\Phi_{21}^{(i)} \neq 0 \ \exists i$.

→ TFNM

$CO2_t = \alpha + \sum_i v_i gas_{t-i} + e_t$, e_t serially correlated.

CCR plot ⇒ $i = 3, \cdots, 7$.

→ TFNM model adequacy

→ Residual ACF and CCF plots to check if there are signs indicating the violation of assumptions.

→ Portmanteau test to check residuals of the fitted model.

```
> m40=Box.test(m1$res, type="Ljung", lag=40, fitdf=6)$p.value
[1] 1.521927e-11
```

⇒ model is fitted adequately.

— Cointegration

<u>Def'n</u> $\underset{\sim}{Z}_t = (Z_{1,t}, \cdots, Z_{k,t})^T$ is said to be <u>cointegrated of order d and b</u>, denoted $\underset{\sim}{Z}_t \sim CI(d,b)$, if

(1) All components $Z_{i,t}$ are $I(d)$

(2) $\exists \beta = (\beta_1, \cdots, \beta_k)^T$ s.t. $\beta \cdot \underset{\sim}{Z}_t \sim I(d-b)$, $b>0$.

β is called the <u>cointegrating vector</u>.

e.g.
$$\begin{bmatrix} \nabla Z_{1,t} \\ \nabla Z_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_1 & -\alpha_1 \\ \alpha_2 & -\alpha_2 \end{bmatrix} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix} + \begin{bmatrix} U_{1,t} \\ U_{2,t} \end{bmatrix}$$

$$\nabla \underset{\sim}{Z}_t = A \underset{\sim}{Z}_{t-1} + \underset{\sim}{U}_t$$

$$\underset{\sim}{Z}_t - \underset{\sim}{Z}_{t-1} = A \underset{\sim}{Z}_{t-1} + \underset{\sim}{U}_t$$

$$\underset{\sim}{Z}_t = (A+I)\underset{\sim}{Z}_{t-1} + \underset{\sim}{U}_t = (A+I)B\underset{\sim}{Z}_t + \underset{\sim}{U}_t$$

$$\cancel{= (AB+I)\underset{\sim}{Z}_t + \underset{\sim}{U}_t}$$

$$\cancel{(I - AB - I)\underset{\sim}{Z}_t = \underset{\sim}{U}_t}$$

$$(I - (A+I)B)\underset{\sim}{Z}_t = \underset{\sim}{U}_t .$$

$$C_1 = A+I$$

$$\Rightarrow (I - C_1 B)\underset{\sim}{Z}_t = \underset{\sim}{U}_t . \quad Z_t \sim VAR(1).$$

$\cancel{\beta = (}$

$$\beta = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\beta' \underset{\sim}{Z}_t = \beta' \left[(A+I)\underset{\sim}{Z}_{t-1} + \underset{\sim}{U}_t \right]$$

$$= \beta'(A+I)\underset{\sim}{Z}_{t-1} + \beta' \underset{\sim}{U}_t$$

$$= \underbrace{\begin{bmatrix} 1 & -1 \end{bmatrix}}_{1\times 2} \underbrace{\begin{bmatrix} \alpha_1+1 & -\alpha_1 \\ \alpha_2 & -\alpha_2+1 \end{bmatrix}}_{2\times 2} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix} + U_{1,t} - U_{2,t} .$$

$\alpha_1+1-\alpha_2$

$-\alpha_1+\alpha_2-1$

$\alpha_1-1+\alpha_2$

$$= \underbrace{\begin{bmatrix} 1+\alpha_1-\alpha_2 & -1-\alpha_1+\alpha_2 \end{bmatrix}}_{1\times 2} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix} + U_{1,t} - U_{2,t}$$

$$= (1+\alpha_1-\alpha_2)Z_{1,t-1} - (1+\alpha_1-\alpha_2)Z_{2,t-1} + U_{1,t} - U_{2,t} .$$

So $\beta' \overset{\overset{y_t}{\parallel}}{\underset{\sim}{Z_t}} = (1+\alpha_1-\alpha_2)(Z_{1,t-1} - Z_{2,t-1}) + u_{1,t} - u_{2,t}$.

$$\Rightarrow y_t = (1+\alpha_1-\alpha_2) y_{t-1} + e_t, \quad e_t = u_{1,t} - u_{2,t} \sim WN(0, \sigma_a^2)$$

$$= \underbrace{(1+\alpha_1-\alpha_2)}_{C} B y_t + e_t$$

$$\Rightarrow (1 - CB) y_t = e_t. \text{ So } y_t \sim AR(1).$$

For y_t to be stationary, we require all the roots of $1 - CB = 0$ to lie outside the unit circle. That is, $|B| = \left|\frac{1}{C}\right| > 1 \Rightarrow \underline{|C| < 1}$. //

→ Some e.g.

1. Money demand

expectations:

$$\hookrightarrow \underset{\underset{\text{(in log)}}{\uparrow}}{m_t} = \beta_0 + \beta_1 \overset{\oplus}{\underset{\underset{\text{price level in log}}{\uparrow}}{p_t}} + \beta_2 \overset{\oplus}{\underset{\underset{\text{real income in log}}{\uparrow}}{y_t}} + \beta_3 \overset{\ominus}{\underset{\underset{\text{interest rate}}{\uparrow}}{r_t}} + \underset{\underset{\text{stationary disturbance in log}}{\nwarrow}}{\varepsilon_t}$$

2. PPP

$$\text{Long-run PPP} = \underset{\underset{\text{price of foreign exchange in log}}{\uparrow}}{e_t} + \underset{\underset{\text{foreign price level in log}}{\uparrow}}{p_{ft}} - \underset{\underset{\text{domestic price level in log}}{\nwarrow}}{p_t}$$

3. Forward rate

$$E_t(\underset{\underset{\text{spot price}}{\nwarrow}}{S_{t+1}}) = \underset{\underset{\text{forward price in log.}}{\nwarrow}}{f_{t,t+1}}$$

$$S_{t+1} - E_t(S_{t+1}) = \varepsilon_t,$$

$$E_t(\varepsilon_{t+1}) = 0. \qquad \{S_t\} \sim I(1).$$

$$\{f_t\} \sim I(1).$$

$$\Rightarrow S_{t+1} = f_{t,t+1} + \underset{\underset{\text{forecast error}}{\nwarrow}}{\varepsilon_t}.$$

4. Pairs trading

Spread: the value of the portfolio containing aforesaid two stocks

Pairs trading: selling a higher priced stock and buying a lower priced stock simultaneously with the hope that the mispricing will correct itself in the future.

Key assumptions:

1) the law of one price
2) if the prices differ, then it is likely that one of the stocks is overpriced and the other underpriced.

- Statistical model for cointegration
  1) regression formulation
  2) autoregressive formulation ⟵ FOCUS
  3) unobserved components formulation

2) AR formulation

<u>Idea</u> Consider: $\nabla \underset{\sim}{Z_t} = \alpha\beta' \underset{\sim}{Z_{t-1}} + \underset{\sim}{U_t}$.

$\underset{\sim}{U_t} \overset{iid}{\sim}$ Errors,

$\alpha_{k\times r}$, $\beta_{k\times r}$ (so $\alpha\beta'$ is $k\times k$) $\beta'_{r\times k}$

The formulation above allows modelling

① the long-run relations $\beta' \underset{\sim}{Z}$, and

② the adjustment, or feedback coeff α towards the attractor set $\{\underset{\sim}{Z} \mid \beta' \underset{\sim}{Z}\}$ defined by the long-run relations.

Implication of the above modelling: there are possibly $r \in [0, k]$ cointegrating vectors.

So models for different cointegration ranks are nested and the smallest, for $\alpha=\beta=0$, corresponds to k random walks. The rank can be tested using LRTs.

<u>Exercise</u> We need $|\lambda|$'s of $I_r+\beta'\alpha$ ~~to~~ less than 1 for $\beta'\underset{\sim}{Z}_t$ to be stationary. Why?

$$\nabla \underset{\sim}{Z}_t = \alpha\beta' \underset{\sim}{Z}_{t-1} + \underset{\sim}{U}_t$$

$$\Rightarrow \underset{\sim}{Z}_t - \underset{\sim}{Z}_{t-1} = \alpha\beta' \underset{\sim}{Z}_{t-1} + U_t$$

$$\cancel{\Rightarrow \nabla \underset{\sim}{Z}_t = \nabla Z + \alpha\beta' B \underset{\sim}{Z}_t + \underset{\sim}{U}_t}.$$

$$\Rightarrow \beta' \underset{\sim}{Z}_t - \beta' \underset{\sim}{Z}_{t-1} = \beta'\alpha\beta' \underset{\sim}{Z}_{t-1} + \beta' \underset{\sim}{U}_t$$

$$\Rightarrow \beta' \underset{\sim}{Z}_t = [\beta'\alpha + I_r]\beta' \underset{\sim}{Z}_{t-1} + \beta' \underset{\sim}{U}_t .$$

So we require all abs. values of λ's of $\beta'\alpha+I_r$ to be less than 1. ✓

<u>Statistical model</u>

Consider $\underset{\sim}{Z}_t \sim VAR(n)$: $\underset{\sim}{Z}_t = \sum_{i=1}^{n} \Phi_i \underset{\sim}{Z}_{t-i} + \underset{\sim}{U}_t$.

$$\Rightarrow \Phi(B) = I_k - \Phi_1 B - \Phi_2 B^2 - \cdots - \Phi_n B^n .$$

Then $\Phi(1) = I_k - \Phi_1 - \Phi_2 - \cdots - \Phi_n$.

Define $\Gamma_i = -[I_k - \Phi_1 - \cdots - \Phi_i]$, $i=1,\cdots,n-1$,

and $\Pi = \Gamma_n = -[I_k - \Phi_1 - \cdots - \Phi_n] = -\Phi(1)$.

$$\text{So } \nabla \underset{\sim}{Z}_t = \sum_{i=1}^{n-1} \Gamma_i \underbrace{\nabla Z_{t-i}}_{\text{first-order differenced}} + \Pi \underset{\sim}{Z}_{t-n} + \underset{\sim}{U}_t .$$

This is a traditional VAR(n-1) except "$\Pi \underset{\sim}{Z}_{t-n}$". "$\Pi$" contains info about long-run relationships btwn the variables in the data vector.

$\underset{\sim}{Z}_t = \sum_{i=1}^{n} \Phi_i \underset{\sim}{Z}_{t-i} + \underset{\sim}{U}_t$

$$\Gamma_i = -I_k + \sum_{j=1}^{i} \Phi_j, \quad \Gamma_n = \Pi = -\Phi(1).$$

$$\nabla \underset{\sim}{Z}_t = \sum_{i=1}^{n-1} \Gamma_i \nabla \underset{\sim}{Z}_{t-i} + \Pi \underset{\sim}{Z}_{t-n} + U_t \quad (\ast)$$

"Π" contains lin. rel. btwn variables in the data vector.

Three possible cases:

① rank$(\Pi) = k$

$\Rightarrow$ $\Phi(B)$ contains no unit root or $\Phi(B) \neq 0$

$\Rightarrow$ Z_t is stationary!

② rank$(\Pi) = 0$

$\Rightarrow$ Π is the null matrix ($\Pi = 0_{k\times k}$)

$\Rightarrow$ $(\ast)$ corresponds to a classic first-order differenced vector time series model.

③ rank$(\Pi) = r \in (0, k)$

$\Rightarrow$ $\exists\, \alpha_{k\times r}$ and $\beta_{k\times r}$ s.t. $\Pi = \alpha\beta'$ and $W_t := \beta' \underset{\sim}{Z}_t \sim I(0)$.

W_t is referred to as cointegrating series, and α denotes the impact of the cointegrating series on $\nabla \underset{\sim}{Z}_t$.

- LRTs for cointegration

$\underset{\sim}{Z}_t \sim N_k(\vec{0}, \Sigma_z)$,

$\underset{\sim}{Z}_t' = (\underset{\sim}{X}_t', \underset{\sim}{Y}_t')$, $\dim(X) = p$, $\dim(y) = q$.

WLOG, let $p \geq q$.

$$\Sigma_z = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

Equivalent:
$X_i = \Pi y_i + U_i$,
$\Pi = 0$ vs $\neq$.

H_0: $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ are uncorrelated.

$\Rightarrow \Sigma_{xy} = 0$.

vs. H_1: $\Sigma_{xy} \neq 0$.

Under H_0, $\hat{\Sigma}_0 = \begin{bmatrix} \hat{\Sigma}_{xx} & 0 \\ 0 & \hat{\Sigma}_{yy} \end{bmatrix}$.

$$\hat{\Sigma}_{xx} = \frac{1}{T}\sum_{t=1}^{T} x_t x_t', \quad \hat{\Sigma}_{yy} = \frac{1}{T}\sum_{t=1}^{T} y_t y_t'.$$

MLE under H_0: $l_0 \propto |\hat{\Sigma}_0|^{-\frac{T}{2}} = (|\hat{\Sigma}_{xx}||\hat{\Sigma}_{yy}|)^{-\frac{T}{2}}$.

~~MLE under H_1: $\hat{\Sigma}_1 =$~~

Under H_1, $\hat{\Sigma}_1 = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} X_t \\ Y_t \end{bmatrix}[X_t' \; Y_t'] = \begin{bmatrix} \hat{\Sigma}_{xx} & \hat{\Sigma}_{xy} \\ \hat{\Sigma}_{yx} & \hat{\Sigma}_{yy} \end{bmatrix}$.

MLE under H_1: $l_1 \propto |\hat{\Sigma}_1|^{-\frac{T}{2}} = (\hat{\Sigma}_{yy} - \hat{\Sigma}_{yx}\hat{\Sigma}_{xx}^{-1}\hat{\Sigma}_{xy})^{-\frac{T}{2}}$.

So $\mathbb{L} = \frac{l_0}{l_1} = \left(\frac{|\hat{\Sigma}_1|}{|\hat{\Sigma}_0|}\right)^{\frac{T}{2}} = \left(|I - \hat{\Sigma}_{yy}^{-1}\hat{\Sigma}_{yx}\hat{\Sigma}_{xx}^{-1}\hat{\Sigma}_{xy}|\right)^{\frac{T}{2}}$.

Reject H_0 if $\mathbb{L}$ small.

Equivalently, if $(\lambda_i)_{i=1}^{q}$ are eigenvalues of $\hat{\Sigma}_{yy}^{-1}\hat{\Sigma}_{yx}\hat{\Sigma}_{xx}^{-1}\hat{\Sigma}_{xy}$,

then $(1-\lambda_i)_{i=1}^{q}$ are eigenvalues of $I - \hat{\Sigma}_{yy}^{-1}\hat{\Sigma}_{yx}\hat{\Sigma}_{xx}^{-1}\hat{\Sigma}_{xy}$.

Define $LR = -\frac{T}{2}\ln(|I - \hat{\Sigma}_{yy}^{-1}\hat{\Sigma}_{yx}\hat{\Sigma}_{xx}^{-1}\hat{\Sigma}_{xy}|)$.

$$= -\frac{T}{2}\ln\left|\prod_{i=1}^{q}(1-\lambda_i)\right)$$

$$= -\frac{T}{2}\sum_{i=1}^{q}\ln(1-\lambda_i).$$

~~Reje~~ Reject H_0 if LR large!

- Cointegration tests of VAR models

We can estimate Π using ~~one regression~~ two regressions:

- $\nabla \underset{\sim}{z}_t = \sum_{i=1}^{n-1} \Psi_i \nabla \underset{\sim}{z}_{t-i} + \underset{\sim}{u}_t$
- $\underset{\sim}{z}_{t-1} = \sum_{i=1}^{n-1} \Psi_i^{*} \nabla \underset{\sim}{z}_{t-i} + \underset{\sim}{v}_t$.

since Π is related to the covariance matrix btwn z_{t-1} and ~~z_t~~ ∇z_t,

$$\nabla \underset{\sim}{Z}_t = \sum_{i=1}^{n-1} \Psi_i \nabla \underset{\sim}{Z}_{t-i} + \underset{\sim}{u}_t$$

$$\underset{\sim}{Z}_{t-1} = \sum_{i=1}^{n-1} \Psi_i^* \nabla \underset{\sim}{Z}_{t-i} + \underset{\sim}{v}_t$$

Let $\hat{\underset{\sim}{u}}_t$ and $\hat{\underset{\sim}{v}}_t$ denote the least-square residuals of ~~the~~ above regressions.

We have $\hat{\underset{\sim}{u}}_t = \Pi \hat{\underset{\sim}{v}}_t + E_t$, E_t = error terms.

Let $H(0) \subset H(1) \subset \cdots \subset H(k)$ be the nested model s.t. under $H(r)$ there are r cointegrating vectors in $\underset{\sim}{Z}_t$ (rank$(\Pi) = r$).

Define
$$\begin{cases} \hat{\Sigma}_{00} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t \hat{u}_t' \\ \hat{\Sigma}_{10} = \frac{1}{T}\sum_{t=1}^{T} \hat{v}_t \hat{u}_t' \\ \hat{\Sigma}_{01} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t \hat{v}_t' \\ \hat{\Sigma}_{11} = \frac{1}{T}\sum_{t=1}^{T} \hat{v}_t \hat{v}_t' \end{cases}$$

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$ be the ordered eigenvalues of the sample matrix $\hat{\Sigma}_{11}^{-1}\hat{\Sigma}_{10}\hat{\Sigma}_{00}^{-1}\hat{\Sigma}_{01}$, and $\underset{\sim}{g}_i$ be the eigenvector of λ_i.

☆ 1) Trace test

$H_0: r = r_0$ vs. $H_1: r > r_0$.

r_0 is an integer btwn 0 and $k-1$.

Johansen's trace statistic:

$$\lambda_{trace}(r_0) = -\tilde{T} \cdot \sum_{i=r_0+1}^{k} \ln(1-\hat{\lambda}_i)$$

$\hat{\lambda}_i$: estimated eigenvalues from estimated Π

$\tilde{T}$: # of usable observations.

Under H_0, $\lambda_{r_0+1}, \lambda_{r_0+2}, \cdots, \lambda_k = 0$,
so the test statistic is small.

Under H_1, some λ_i $(i = r_0+1, \cdots, k)$ are nonzero,
so ~~the~~ test stat is large.

[Presence of unit root $\Rightarrow \lambda_{trace}(r_0) \not\sim \chi^2$]

☆ 2) Maximum eigenvalue test.

$H_0: r = r_0$ vs. $H_1: r = r_0+1$.

$\lambda_{max}(r_0) = -\tilde{T} \cdot \ln(1 - \hat{\lambda}_{r_0+1})$.

e.g. Danish money demand

pg. 25

Table 1. Trace test ($\lambda_{trace}(r_0) = -\tilde{T} \cdot \sum_{i=r_0+1}^{k} \ln(1-\hat{\lambda}_i)$) $H_0: r=r_0$ vs. $H_1: r>r_0$

Table 2. Max eigenvalue test ($\lambda_{max}(r_0) = -\tilde{T} \cdot \ln(1-\hat{\lambda}_{r_0+1})$) $H_0: r=r_0$ vs. $H_1: r=r_0+1$

	Eigen.max	10pct	5pct	1pct
$r \leq 3$	2.35	7.52	9.24	12.97
$r \leq 2$	6.34	13.75	15.67	20.20
$r \leq 1$	10.36	19.77	22.00	26.81
$r = 0$	30.09	25.56	28.14	33.24

→ Starting from r=1, Eigen.max < 10, 5, 1 pct's.
Choose r=1.

$\underset{\sim}{\nabla z_t} = \alpha\beta' \underset{\sim}{z}_{t-1} + u_t$

$\underset{\sim}{z}_t \sim VAR(n)$

$\underset{\sim}{z}_t = \sum_{i=1}^{n} \Phi_i z_{t-i} + \underset{\sim}{u}_t$

$\Gamma_i = -I_k + \sum_{j=1}^{i} \Phi_j$

~~Γ_n I_k~~

$\Gamma_n = \Pi = -\Phi(1)$

$\nabla \underset{\sim}{z}_t = \sum_{i=1}^{n-1} \Gamma_i \nabla \underset{\sim}{z}_{t-i} + \Pi \underset{\sim}{z}_{t-n} + u_t$

$rank(\Pi) = r \in (0,k)$

$\Rightarrow \exists$ β

s.t. $\alpha\beta' = \Pi$

and $\beta' \underset{\sim}{z}_t \sim I(0)$.

Table 3.

Estimates of eigenvalues(λ) and cointegrating vectors(β)

Table 4.

Estimates of speed adjusement (α)

3.

	LRM.l2	LRY.l2	IBO.l2	IDE.l2	constant
Eigenvalues	.43	.18	.11	.04	.00
LRM.l2	1	1	1	1	1
LRY.l2	-1.03	-1.37	-3.23	-1.88	-0.63
IBO.l2	5.21	.24	.54	24.40	1.70
IDE.l2	-4.22	6.84	-5.65	-14.30	-1.90
constant	-6.06	-4.27	7.90	-2.26	-8.03

$\Rightarrow \underline{m_2 = 1.03y - 5.21 i_b + 4.22 i_d + 6.06}$.

4.

	LRM.l2	LRY.l2	IBO.l2	IDE.l2	Constant
LRM.d	-0.213	-0.005	0.035	0.002	0
LRY.d	0.115	0.02	0.05	0.001	0
IBO.d	0.023	-0.011	0.003	-0.002	0
IDE.d	0.029	-0.03	-0.003	0.000	0

$$\hat{\alpha} = \begin{bmatrix} -0.213 \\ 0.115 \\ 0.023 \\ 0.029 \end{bmatrix}$$