Financial Times Series

Lecture 10

Multivariate Time Series

- In many situations we want to model how returns of different assets evolve simultaneously

$$\{\boldsymbol{r}_t\}_{t\in\mathbb{N}} = \{(r_{1t}, \dots, r_{kt})'\}_{t\in\mathbb{N}}$$

- Typically we will have correlations between returns of different assets

Weak Stationarity

- We say that $\{\boldsymbol{r}_t\}_{t\in\mathbb{N}}$ is weakly stationary if its first and second moments are time invariant

- In particular the mean vector $\boldsymbol{\mu} = E(\boldsymbol{r}_t)$ and covariance matrix $\boldsymbol{\Gamma}_0 = E[(\boldsymbol{r}_t - \boldsymbol{\mu})(\boldsymbol{r}_t - \boldsymbol{\mu})']$ are time invariant

- The diagonal $\Gamma_{ii}(0)$ elements of $\boldsymbol{\Gamma}_0$ are the variances of$r_{1t}, \dots, r_{kt}$and the off-diagonal elements $\Gamma_{ij}(0)$ are the covariances $Cov\left(r_{it}, r_{jt}\right)$

Cross-Correlation Matrices

- Let $\boldsymbol{D}$ be the diagonal matrix containing the standard deviations $\sqrt{\Gamma_{11}(0)}, \dots, \sqrt{\Gamma_{kk}(0)}$

- We define the cross-correlation matrix of $\{\boldsymbol{r}_t\}_{t\in\mathbb{N}}$ as

$$\boldsymbol{\rho}_0 = \boldsymbol{D}^{-1}\boldsymbol{\Gamma}_0\boldsymbol{D}^{-1}$$

- The elements are $\rho_{ij}(0) = \frac{\Gamma_{ij}(0)}{\sqrt{\Gamma_{ii}(0)\Gamma_{jj}(0)}}$

- Such elements are called concurrent or contemporaneous since the are "at lag zero"

Cross-Correlation Matrices

- We may also define the lag-l cross-correlation matrix

$$\boldsymbol{\rho}_l = \boldsymbol{D}^{-1}\boldsymbol{\Gamma}_l\boldsymbol{D}^{-1}$$

where

$$\boldsymbol{\Gamma}_l = E[(\boldsymbol{r}_t - \boldsymbol{\mu})(\boldsymbol{r}_{t-l} - \boldsymbol{\mu})']$$

- The elements of $\boldsymbol{\rho}_l$ are $\rho_{ij}(l) = \dfrac{\Gamma_{ij}(l)}{\sqrt{\Gamma_{ii}(0)\Gamma_{jj}(0)}}$

Sample Cross-Correlation Matrices

- In practice we use the sample versions of the lag-l cross-correlation matrix

$$\hat{\boldsymbol{\rho}}_l = \widehat{\boldsymbol{D}}^{-\mathbf{1}} \hat{\boldsymbol{\Gamma}}_l \widehat{\boldsymbol{D}}^{-\mathbf{1}}$$

where

$$\hat{\boldsymbol{\Gamma}}_l = \frac{1}{T} \sum_{t=l+1}^{T} (\boldsymbol{r}_t - \bar{\boldsymbol{r}})(\boldsymbol{r}_{t-l} - \bar{\boldsymbol{r}})'$$

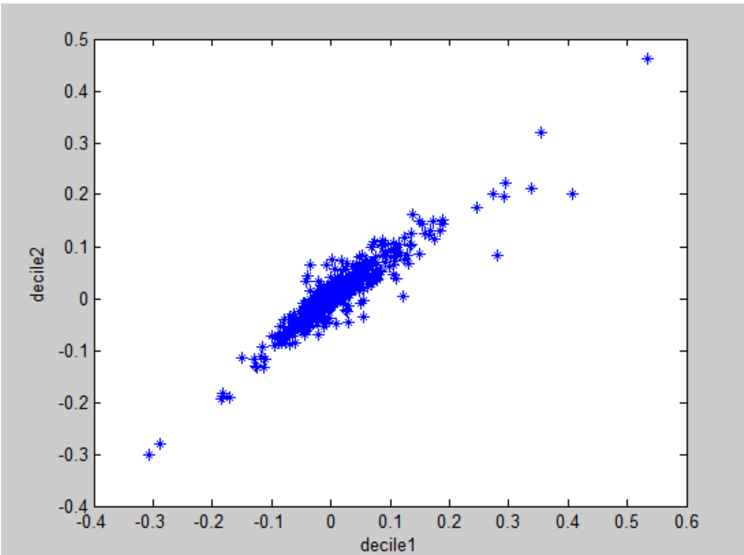
- Here $\bar{\boldsymbol{r}} = \frac{1}{T} \sum_{i=1}^{T} \boldsymbol{r}_t$ and $\widehat{\boldsymbol{D}}$ is the diagonal matrix with sample standard deviations

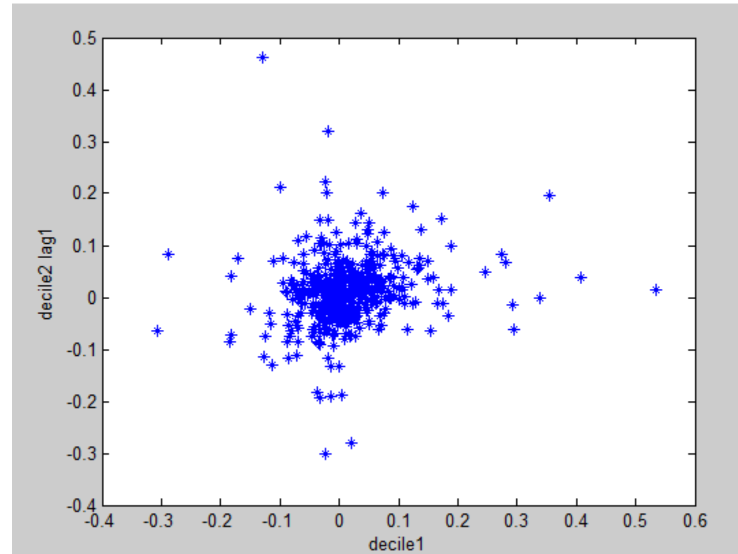
Concurrent CCM

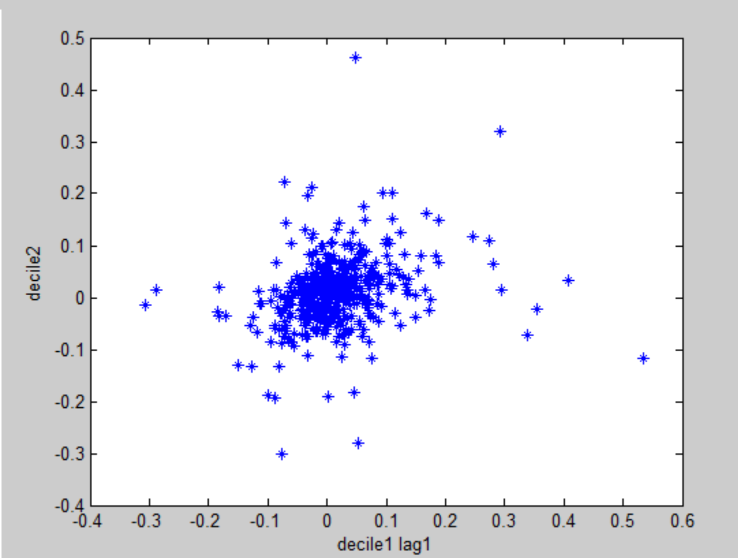
- The concurrent CCM for decile1,2,9,10 is given by

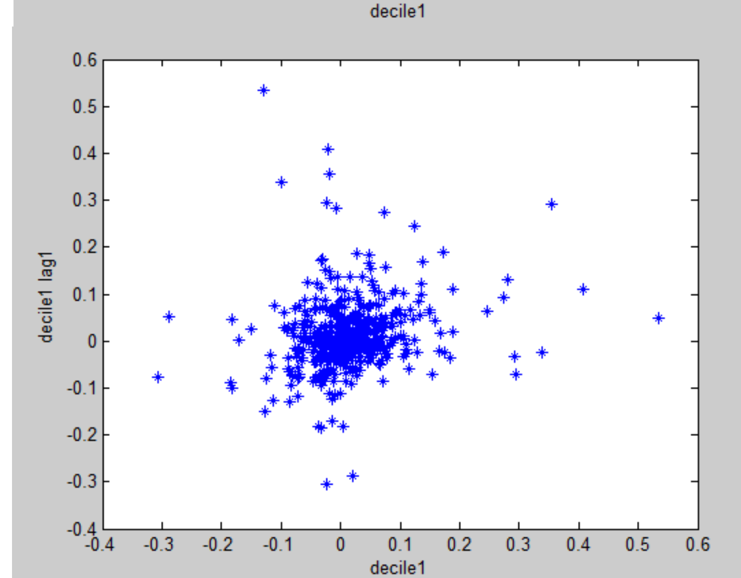
$$\widehat{\boldsymbol{\rho}}_0 = \begin{bmatrix} \mathbf{1.00} & \mathbf{0.93} & \mathbf{0.65} & \mathbf{0.51} \\ \mathbf{0.93} & \mathbf{1.00} & \mathbf{0.77} & \mathbf{0.63} \\ \mathbf{0.64} & \mathbf{0.77} & \mathbf{1.00} & \mathbf{0.92} \\ \mathbf{0.51} & \mathbf{0.63} & \mathbf{0.92} & \mathbf{1.00} \end{bmatrix}$$

Example decile1 vs.decile 2, lag 0, lag 1









Descriptives and cross-correlations

Series	Mean	Stdev	Skew	Kurt	Min	Max
Decile1	0.0133	0.0771	1.4084	11.401	-0.306	0.5348
Decile2	0.0105	0.0653	0.6184	10.697	-0.301	0.4620

Lag1		Lag2		Lag3		Lag4		Lag5	
0.18	0.16	-0.03	-0.02	-0.10	-0.09	-0.06	-0.05	-0.05	-0.06
0.25	0.20	-0.03	-0.03	-0.09	-0.06	-0.06	-0.04	-0.04	-0.05

Significant correlations									
Lag1		Lag2		Lag3		Lag4		Lag5	
+	+			-	-				
+	+			-					

Checking for significant cross-correlations

- Just as for the univariate case we may use 95% percent confidence bands

$$\left(-\frac{1.96}{\sqrt{T}}, \frac{1.96}{\sqrt{T}}\right)$$

Multivariate Portmanteau

- The null hypothesis $\boldsymbol{\rho}_{\mathbf{1}} = \cdots = \boldsymbol{\rho}_m = \mathbf{0}$ and its alternative $\boldsymbol{\rho}_i \neq \mathbf{0}$ for some $i \in \{1, \ldots, m\}$ may be tested using (the chi-square with $df = k^2 m$ distributed)

$$T^2 \sum_{l=1}^{m} \frac{1}{T-l} \operatorname{tr}\left(\hat{\boldsymbol{\Gamma}}_l{}' \hat{\boldsymbol{\Gamma}}_0{}^{-1} \hat{\boldsymbol{\Gamma}}_l \hat{\boldsymbol{\Gamma}}_0{}^{-1}\right)$$

where $\operatorname{tr}(A)$ is the trace of the matrix A, i.e. the sum of its diagonal elements (and k is the dimension of $\boldsymbol{r}_t$)

VAR (not VaR)

- Based on the decile1, decile2 significant lag 1 correlations it may be useful to model the two series (simultaneously) as an autoregressive series

- In the multivariate setting this is referred to as VAR (Vector AutoRegressive)

VAR(1)

- The simplest form is given by

$$\boldsymbol{r}_t = \boldsymbol{\varphi}_0 + \boldsymbol{\Phi}\boldsymbol{r}_{t-1} + \boldsymbol{a}_t$$

where $\boldsymbol{\varphi}_0$ is a k-dimensional vector, $\boldsymbol{\Phi}$ is a $k \times k$ matrix and $\{\boldsymbol{a}_t\}$ is a sequence of serially uncorrelated (k-dimensional) random vectors with mean zero and (positive definite) covariance matrix $\boldsymbol{\Sigma}$

2-dimensional VAR(1)

- As an example we write the components of the dimensional VAR(1)

$$r_{1t} = \varphi_{10} + \Phi_{11} r_{1t-1} + \Phi_{12} r_{2t-1} + a_{1t}$$
$$r_{2t} = \varphi_{20} + \Phi_{21} r_{1t-1} + \Phi_{22} r_{2t-1} + a_{2t}$$

- We may interpret Φ_{12} as the conditional effect of r_{2t-1} on r_{1t} given r_{1t-1} (v.v. for Φ_{21}).

- If $\Phi_{12} = 0$ then r_{1t} only depends on ”its own past”

2-dimensional VAR(1)

- If $\Phi_{12} = 0$ and $\Phi_{21} \neq 0$ then ”1 feeds 2” but not vice versa
- If $\Phi_{12} = \Phi_{21} = 0$ then ”1 and 2” are uncoupled
- If $\Phi_{12} \neq 0$ and $\Phi_{21} \neq 0$ there is a feedback relationship between ”1 and 2”

2-dimensional VAR(1)

- So the elements of the matrix $\mathbf{\Phi}$ determine the dynamic relation between ”1 and 2”

- The concurrent relation is given by the off-diagonal elements of $\mathbf{\Sigma}$

- Sometimes VAR models are written in reduced or structural forms, see p400 Tsay 3rd ed.

Stationarity and moments

- It can be shown that, given the existence of the WN covariance matrix $\mathbf{\Sigma}$**,** a necessary and sufficient condition for weak stationarity of the VAR(1) is that the eigenvalues of $\mathbf{\Phi}$ are less than one i modulus

- Sketched proof is found on p.402 Tsay 3rd ed.

Stationarity and moments

- If the VAR(1) as described above is weakly stationary we may write (since $\boldsymbol{E}(\boldsymbol{a}_t) = \boldsymbol{0}$)

$$\boldsymbol{E}(\boldsymbol{r}_t) = \boldsymbol{\varphi}_0 + \boldsymbol{\Phi}\boldsymbol{E}(\boldsymbol{r}_{t-1})$$

and get

$$\boldsymbol{\mu} = \boldsymbol{E}(\boldsymbol{r}_t) = (\boldsymbol{I} - \boldsymbol{\Phi})^{-\mathbf{1}}\boldsymbol{\varphi}_0$$

Stationarity and moments

- Using $\boldsymbol{\varphi_0} = (\boldsymbol{I} - \boldsymbol{\Phi})\boldsymbol{\mu}$ we may write

$$(\boldsymbol{r}_t - \boldsymbol{\mu}) = \boldsymbol{\Phi}(\boldsymbol{r}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{a}_t$$

and letting $\tilde{\boldsymbol{r}}_t = \boldsymbol{r}_t - \boldsymbol{\mu}$ we have

$$\begin{aligned} \tilde{\boldsymbol{r}}_t &= \boldsymbol{\Phi}\tilde{\boldsymbol{r}}_{t-1} + \boldsymbol{a}_t = \cdots \\ &= \boldsymbol{a}_t + \boldsymbol{\Phi}\boldsymbol{a}_{t-1} + \boldsymbol{\Phi}^2\boldsymbol{a}_{t-2} + \cdots \end{aligned}$$

Stationarity and moments

- Using the above expression gives (exercise)

$$Cov(\boldsymbol{r}_t, \boldsymbol{a}_t) = \boldsymbol{\Sigma}$$

and

$$Cov(\boldsymbol{r}_t, \boldsymbol{r}_t) = \boldsymbol{\Gamma}_0 = \boldsymbol{\Sigma} + \boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{\Phi}' + \boldsymbol{\Phi}^2\boldsymbol{\Sigma}(\boldsymbol{\Phi}')^2 + \cdots$$

- We also get (exercise)

$$\boldsymbol{\Gamma}_l = \boldsymbol{\Phi}^{\boldsymbol{l}}\boldsymbol{\Gamma}_0 \text{ and } \boldsymbol{\rho}_l = \left(\boldsymbol{D}^{-1/2}\boldsymbol{\Phi}\boldsymbol{D}^{1/2}\right)^{\boldsymbol{l}}\boldsymbol{\rho}_0$$

VAR(p)

- Generalization is as expected;

$$\boldsymbol{r}_t = \boldsymbol{\varphi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{r}_{t-1} + \cdots + \boldsymbol{\Phi}_p \boldsymbol{r}_{t-p} + \boldsymbol{a}_t$$

- Under weak stationarity (necessary and sufficient condition given further down)

$$\boldsymbol{\mu} = \boldsymbol{E}(\boldsymbol{r}_t) = \left(\boldsymbol{I} - \boldsymbol{\Phi}_1 - \ldots - \boldsymbol{\Phi}_p\right)^{-\mathbf{1}} \boldsymbol{\varphi}_{\mathbf{0}}$$

$$\boldsymbol{\Gamma}_l = \boldsymbol{\Phi}_1 \boldsymbol{\Gamma}_{l-1} + \cdots + \boldsymbol{\Phi}_p \boldsymbol{\Gamma}_{l-p}$$

$$\boldsymbol{\rho}_l = \boldsymbol{D}^{-\mathbf{1}/\mathbf{2}} \boldsymbol{\Phi}_1 \boldsymbol{D}^{\mathbf{1}/\mathbf{2}} \boldsymbol{\rho}_{l-1} + \cdots + \boldsymbol{D}^{-\mathbf{1}/\mathbf{2}} \boldsymbol{\Phi}_p \boldsymbol{D}^{\mathbf{1}/\mathbf{2}} \boldsymbol{\rho}_{l-p}$$

Stationarity

- Letting $\tilde{\boldsymbol{r}}_t = \boldsymbol{r}_t - \boldsymbol{\mu}$, $\boldsymbol{x}_t = \left(\tilde{\boldsymbol{r}}'_{t-p+1}, \ldots, \tilde{\boldsymbol{r}}'_t\right)'$ and $\boldsymbol{b}_t = (0, . . , 0, \boldsymbol{a}'_t)'$ we may write the VAR(p) as

$$\boldsymbol{x}_t = \boldsymbol{\Phi}^* \boldsymbol{x}_{t-1} + \boldsymbol{b}_t$$

where

$$\boldsymbol{\Phi}^* = \begin{bmatrix} \boldsymbol{0} & & \boldsymbol{I} & \cdots & \boldsymbol{0} & & \boldsymbol{0} \\ \boldsymbol{0} & & \boldsymbol{0} & & \boldsymbol{0} & & \boldsymbol{0} \\ & \vdots & & \ddots & & \vdots & \\ \boldsymbol{0} & & \boldsymbol{0} & & \boldsymbol{0} & & \boldsymbol{I} \\ \boldsymbol{\Phi}_p & & \boldsymbol{\Phi}_{p-1} & \cdots & \boldsymbol{\Phi}_2 & & \boldsymbol{\Phi}_1 \end{bmatrix}$$

- A necessary and sufficient condition for weak stationarity of the VAR(p) is that the eigenvalues of $\boldsymbol{\Phi}^*$ are less than one in modulus

Estimation

- To determine the order of our model we consider the models

$$\boldsymbol{r}_t = \boldsymbol{\varphi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{r}_{t-1} + \boldsymbol{a}_t$$
$$\boldsymbol{r}_t = \boldsymbol{\varphi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{r}_{t-1} + \boldsymbol{\Phi}_2 \boldsymbol{r}_{t-2} + \boldsymbol{a}_t$$
$$\vdots$$
$$\boldsymbol{r}_t = \boldsymbol{\varphi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{r}_{t-1} + \cdots + \boldsymbol{\Phi}_i \boldsymbol{r}_{t-i} + \boldsymbol{a}_t$$
$$\vdots$$

- Parameters may be estimated using OLS or ML

Estimation

- For the ”order i-model” we let $\widehat{\mathbf{\Phi}}_j{}^{(i)}$ be the OLS estimate of $\mathbf{\Phi}_j$

- The residuals are given by

$$\widehat{\boldsymbol{a}}_t{}^{(i)} = \boldsymbol{r}_t - \widehat{\boldsymbol{\varphi}}_0{}^{(i)} - \widehat{\mathbf{\Phi}}_1{}^{(i)} \boldsymbol{r}_{t-1} - \cdots - \widehat{\mathbf{\Phi}}_j{}^{(i)} \boldsymbol{r}_{t-i}$$

and

$$\widehat{\mathbf{\Sigma}}_i = \frac{1}{T-2i-1} \sum_{t=i+1}^{T} \widehat{\boldsymbol{a}}_t{}^{(i)} \left(\widehat{\boldsymbol{a}}_t{}^{(i)}\right)'$$

Estimation

- To specify the order we may (sequentially) test

$H_0: \mathbf{\Phi}_l = \mathbf{0}$ vs $H_a: \mathbf{\Phi}_l \neq \mathbf{0}$ for $l = 1,2, \ldots$ using the (asymptotically chi-squared with $df = k^2$ distributed) test function

$$M(l) = -\left(T - k - l - \frac{3}{2}\right) \ln\left(\frac{|\widehat{\mathbf{\Sigma}}_l|}{|\widehat{\mathbf{\Sigma}}_{l-1}|}\right)$$

Estimation

- Under assumption of Gaussian noise we may also use different information criteria

- The parameter estimates may be found using ML or OLS

- If using ML the estimate of the WN covariance matrix is

$$\widetilde{\boldsymbol{\Sigma}}_i = \frac{1}{T}\sum_{t=i+1}^{T} \widehat{\boldsymbol{a}}_t{}^{(i)}\left(\widehat{\boldsymbol{a}}_t{}^{(i)}\right)'$$

Estimation

- Under the assumption of Gaussian WN we have for the VAR(i) (smaller is better)

- $AIC(i) = \ln\left(\left|\widetilde{\mathbf{\Sigma}}_i\right|\right) + \frac{2k^2 i}{T}$
- $BIC(i) = \ln\left(\left|\widetilde{\mathbf{\Sigma}}_i\right|\right) + \frac{2k^2 i \ln(T)}{T}$

Estimation

- For the decile1,2,9,10 data we get

Order (l)	1	2	3	4	5
$M(l)$	$-$	30.16	46.00	22.94	32.50
$BIC(l)$	-3.37	-3.37	-3.40	-3.38	-3.39

- The critical values for the chi-square with $df = 16$ for $\alpha = 0.05$ and $\alpha = 0.01$ are 29.30 and 32.000 respectively

- So, we choose to use the VAR(3) model

- However it turns out that resdials do not pass the Portmanteau test.

Forecasting

- Forecasting is similar to the univariate case

- A one-step forecast is given by (treating the estimated model as the true model)

$$\boldsymbol{r}_h(1) = \boldsymbol{\varphi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{r}_{h-1} + \cdots + \boldsymbol{\Phi}_p \boldsymbol{r}_{h-p}$$

- A two-step forecast is given by (treating the estimated model as the true model)

$$\boldsymbol{r}_h(2) = \boldsymbol{\varphi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{r}_h(1) + \cdots + \boldsymbol{\Phi}_p \boldsymbol{r}_{h-p}$$

- And so forth…

VMA

- Of course there are also Vector MA models

- The VMA(q) is given by

$$\boldsymbol{r}_t = \boldsymbol{\theta}_0 + \boldsymbol{a}_t - \boldsymbol{\Theta}_1 \boldsymbol{a}_{t-1} - \cdots - \boldsymbol{\Theta}_q \boldsymbol{a}_{t-q}$$

where $\boldsymbol{\theta}_0$ is a k-dimensional vector, $\boldsymbol{\Theta}_i$ are a $k \times k$ matrices and $\{\boldsymbol{a}_t\}$ is a sequence of serially uncorrelated (k-dimensional) random vectors with mean zero and (positive definite) covariance matrix $\boldsymbol{\Sigma}$

- The VMA(q) is weakly stationary if $\boldsymbol{\Sigma}$ exists

Properties of VMA(q)

- $\boldsymbol{\mu} = \boldsymbol{E}(\boldsymbol{r}_t) = \boldsymbol{\theta}_0$
- $Cov(\boldsymbol{r}_t, \boldsymbol{a}_t) = \boldsymbol{\Sigma}$
- $\boldsymbol{\Gamma}_0 = \boldsymbol{\Sigma} + \boldsymbol{\Theta}_1\boldsymbol{\Sigma}\boldsymbol{\Theta}'_1 + \cdots + \boldsymbol{\Theta}_q\boldsymbol{\Sigma}\boldsymbol{\Theta}'_q$
- $\boldsymbol{\Gamma}_l = \boldsymbol{0}, l > q$
- $\boldsymbol{\Gamma}_l = \sum_{j=l}^{q} \boldsymbol{\Theta}_j\boldsymbol{\Sigma}\boldsymbol{\Theta}'_{j-l}\,, 1 \leq l \leq q, \boldsymbol{\Theta}_0 = -\boldsymbol{I}$

Estimation

- We may use conditional ML for

$$\boldsymbol{a}_1 = \boldsymbol{r}_1 - \boldsymbol{\theta}_0,$$
$$\boldsymbol{a}_2 = \boldsymbol{r}_1 - \boldsymbol{\theta}_0 + \boldsymbol{\Theta}_1 \boldsymbol{a}_1$$
$$\vdots$$
$$\boldsymbol{a}_t = \boldsymbol{r}_t - \boldsymbol{\theta}_0 + \boldsymbol{\Theta}_1 \boldsymbol{a}_{t-1} + \cdots + \boldsymbol{\Theta}_q \boldsymbol{a}_{t-q}$$

Estimation

- The the log likelihood function under assumption of Gaussianity is

$$l\left(\boldsymbol{\theta}_0, \boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_q, \boldsymbol{\Sigma} \,|\boldsymbol{r}_1, \ldots, \boldsymbol{r}_T\right)$$
$$= -\frac{1}{2}\sum_{t=1}^{T}\left(\ln(|\boldsymbol{\Sigma}|) + \boldsymbol{a}'_t\boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{a}_t\right)$$

VMA(1)

- We fit a four-dimensional VMA(1) to the decile1,2,9,10 data and get

$$\widehat{\boldsymbol{\theta}}_0 = [\mathbf{0.0133} \quad \mathbf{0.0105} \quad \mathbf{0.0099} \quad \mathbf{0.0083}]'$$

$$\widehat{\boldsymbol{\Theta}}_1 = \begin{bmatrix} \mathbf{0.13} & \mathbf{-0.39} & \mathbf{0.08} & \mathbf{0.09} \\ \mathbf{-0.02} & \mathbf{-0.06} & \mathbf{-0.09} & \mathbf{0.15} \\ \mathbf{-0.00} & \mathbf{0.01} & \mathbf{0.04} & \mathbf{0.04} \\ \mathbf{0.03} & \mathbf{-0.07} & \mathbf{0.16} & \mathbf{-0.02} \end{bmatrix}$$

Are the parameters statistically significant?

- To check this we (may) use profile likelihood

- The idea is to use that twice the difference between the log likelihood function for all parameters and the log likelihood function for all parameters but one follows a chi-squared distribution with one degree of freedom

- The 95% quantile for the chi-squared with one degree of freedom is 3.84 which means that we find a 95% CI for Θ_{11} say by finding the values of Θ_{11} for which the difference of the log likelihood function with all ML estimates plugged in and the log likelihood where we all the ML estimates except the one for Θ_{11} is 3.84/2=1.92

Parameter	95% LCL	95% UCL
Θ_{11}	0.1109	0.1540
Θ_{12}	-0.4210	-0.3708
Θ_{13}	0.0145	0.0934
Θ_{14}	0.0378	0.1211
Θ_{21}	-0.0403	-0.0104
Θ_{22}	-0.0824	-0.0481
Θ_{23}	-0.1164	-0.0720
Θ_{24}	0.1242	0.1781
Θ_{31}	-0.0138	0.0118
Θ_{32}	-0.0086	0.0228
Θ_{33}	0.0168	0.0569
Θ_{34}	0.0164	0.0614
Θ_{41}	0.0082	0.0363
Θ_{42}	-0.0920	-0.0597
Θ_{43}	0.1387	0.1809
Θ_{44}	-0.0450	0.0007

Noise Covariance and Correlation

- The estimated noise covariance matrix is

$$\widehat{\boldsymbol{\Sigma}} = \begin{bmatrix} \mathbf{0.0061} & \mathbf{0.0049} & \mathbf{0.0029} & \mathbf{0.0020} \\ \mathbf{0.0049} & \mathbf{0.0045} & \mathbf{0.0029} & \mathbf{0.0021} \\ \mathbf{0.0029} & \mathbf{0.0029} & \mathbf{0.0029} & \mathbf{0.0023} \\ \mathbf{0.0020} & \mathbf{0.0021} & \mathbf{0.0023} & \mathbf{0.0022} \end{bmatrix}$$

- If we transform $\widehat{\boldsymbol{\Sigma}}$ into a correlation matrix $\widehat{\boldsymbol{\rho}}$ we get

$$\widehat{\boldsymbol{\rho}} = \begin{bmatrix} \mathbf{1.00} & \mathbf{0.94} & \mathbf{0.68} & \mathbf{0.57} \\ \mathbf{0.94} & \mathbf{1.00} & \mathbf{0.80} & \mathbf{0.69} \\ \mathbf{0.68} & \mathbf{0.80} & \mathbf{1.00} & \mathbf{0.93} \\ \mathbf{0.57} & \mathbf{0.69} & \mathbf{0.93} & \mathbf{1.00} \end{bmatrix}$$

- So there are some strong correlations between the noise series

VARMA and co-integration

- There is of course also a generalization of univariate to multivariate or vector ARMA
- Also unit root non-stationarity applies for multivariate models
- More information found in Tsay.