MATH4425 (T1A) – Tutorial 12

Kazovskaia Anastasiia

May, 9

Important information

- T1A: **Thursday 19:00 - 19:50** (Rm 1033, LSK Bldg)
- Office hours: **Wednesday 14:00 - 14:50** (Math support center, 3rd floor, Lift 3)
- Any questions to be addressed to **akazovskaia@connect.ust.hk**

Please vote in the survey

Multivariate Time Series Models

Moving Average and Autoregressive Representation of Vector Processes

MA(∞) model

$$Z_t = \sum_{s=1}^{\infty} \Psi_s a_{t-s} + a_t,$$

where $\Psi_s := (\Psi_{s_{i,j}})_{k\times k}$ and $\sum_{s=0}^{\infty} |\Psi_{s_{i,j}}|^2 < \infty$.

AR(∞) model

$$Z_t = \sum_{s=1}^{\infty} \Pi_s Z_{t-s} + a_t,$$

where $\Pi_s := (\Pi_{s_{i,j}})_{k\times k}$ and $\sum_{s=0}^{\infty} |\Pi_{s_{i,j}}|^2 < \infty$.

Vector AR model

Vector AR(1) model

Model

$$(I - \Phi_1 B)Z_t = a_t \Leftrightarrow Z_t = \Phi_1 Z_{t-1} + a_t$$

Example: When $k = 2$,

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} \phi_{1,1} & \phi_{1,2} \\ \phi_{2,1} & \phi_{2,2} \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{pmatrix} + \begin{pmatrix} a_{1,t} \\ a_{2,t} \end{pmatrix},$$

which is equivalent to

$$\begin{cases} Z_{1,t} = \phi_{1,1} Z_{1,t-1} + \phi_{1,2} Z_{2,t-1} + a_{1,t} \\ Z_{2,t} = \phi_{2,1} Z_{1,t-1} + \phi_{2,2} Z_{2,t-1} + a_{2,t} \end{cases}$$

Stationarity condition

All **roots** of $|I - \Phi_1 z| = 0$ lie **outside** the unit circle, which is equivalent to the requirement that all **eigenvalues** of Φ_1 lie **inside** the unit circle.

In this case,

$$(I - \Phi_1 B)^{-1} = I + \Phi_1 B + \Phi_1^2 B^2 + \dots$$

and

$$Z_t = a_t + \Phi_1 a_{t-1} + \Phi_1^2 a_{t-2} + \dots$$

Covariance matrix function

Let's first notice

$$\mathbb{E} Z_t = \mathbb{E} \sum_{i=0}^{\infty} \Phi_1^i a_{t-i} = 0$$

Then,

$$\Gamma(0) = \mathbb{E} Z_t Z_t^T = \mathbb{E} \left(\sum_{i=0}^{\infty} \Phi_1^i a_{t-i} \right) \left(\sum_{j=0}^{\infty} \Phi_1^j a_{t-j} \right)^T =$$

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \Phi_1^i \mathbb{E} \left(a_{t-i} a_{t-j}^T \right) \Phi_1^{j^T} = \sum_{i=0}^{\infty} \Phi_1^i \Sigma \Phi_1^{i^T} = \Sigma + \Gamma(1) \Phi_1^T$$

$$\Gamma(1) = \mathbb{E} Z_t Z_{t-1}^T = \mathbb{E} \left(\sum_{i=0}^{\infty} \Phi_1^i a_{t-i} \right) \left(\sum_{j=0}^{\infty} \Phi_1^j a_{t-j-1} \right)^T = \mathbb{E} \left(\sum_{i=0}^{\infty} \Phi_1^i a_{t-i} \right) \left(\sum_{j=1}^{\infty} \Phi_1^{j-1} a_{t-j} \right)^T =$$

$$\sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \Phi_1^i \mathbb{E} \left(a_{t-i} a_{t-j}^T \right) \Phi_1^{j-1^T} = \sum_{i=1}^{\infty} \Phi_1^i \Sigma \Phi_1^{i-1^T}$$

$$\Gamma(l) = \mathbb{E}Z_t Z_{t-l}^T = \mathbb{E}(\Phi_1 Z_{t-1} + a_t) Z_{t-l}^T = \Phi_1 \mathbb{E} Z_{t-1} Z_{t-l}^T + \mathbb{E} a_t Z_{t-l}^T =$$

$$\Phi_1 \Gamma(l-1) + \mathbb{E} a_t \left(\sum_{i=0}^{\infty} \Phi_1^i a_{t-l-i} \right)^T = \Phi_1 \Gamma(l-1) + \sum_{i=0}^{\infty} \mathbb{E}\left(a_t a_{t-l-i}^T\right) \Phi_1^{i^T} = \Phi_1 \Gamma(l-1) \quad \forall l > 0$$

$$\Gamma(l) = \begin{cases} \Gamma(1)\Phi_1^T + \Sigma, & \text{if } l = 0 \\ \Phi_1 \Gamma(l-1) = \Phi_1^l \Gamma(0), & \text{if } l \geq 1 \end{cases}$$

Vector AR(p) model

Model

$$(I - \Phi_1 B - \Phi_2 B^2 - \cdots \Phi_p B^p) Z_t = a_t \Leftrightarrow Z_t = \Phi_1 Z_{t-1} + \Phi_2 Z_{t-2} + \cdots + \Phi_p Z_{t-p} + a_t$$

Stationarity condition
All **roots** of $|I - \Phi_1 z - \Phi_2 z^2 - \cdots \Phi_p z^p| = 0$ lie **outside** the unit circle.

In this case,

$$(I - \Phi_1 B - \Phi_2 B^2 - \cdots - \Phi_p B^p)^{-1} = I + \Psi_1 B + \Psi_2 B^2 + \cdots \Leftarrow$$

$$(I - \Phi_1 B - \Phi_2 B^2 - \cdots - \Phi_p B^p)(I + \Psi_1 B + \Psi_2 B^2 + \dots) = I$$

and

$$Z_t = a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \dots$$

Covariance matrix function
The covariance matrix function can be found via direct calculation, though it's tedious.

Fitting VAR(p) model

Assume that $Z_1, Z_2, \dots, Z_n$ are from the k-dimensional VAR(p) model with drift

$$Z_t = \Phi_{0_0} + \Phi_{1_0} Z_{t-1} + \Phi_{2_0} Z_{t-2} + \cdots + \Phi_{p_0} Z_{t-p} + a_t,$$

where $a_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \Sigma_0)$.

Let's rewrite it:

$$Z_t = \Pi_0 X_t + a_t,$$

where

$$\Pi_0 := \left[\Phi_{0_0} \quad \Phi_{1_0} \quad \Phi_{2_0} \quad \dots \quad \Phi_{p_0}\right],$$

which is a $k \times (kp+1)$-matrix, and

$$X_t := \begin{bmatrix} 1 \\ Z_{t-1} \\ Z_{t-2} \\ \vdots \\ Z_{t-p} \end{bmatrix},$$

which is a $(kp+1)$-vector.

Let's denote $\theta := [\Phi_0 \;\; \Phi_1 \;\; \Phi_2 \;\; \ldots \;\; \Phi_p \;\; \Sigma]$. **The conditional likelihood function** (given $Z_0^*, Z_{-1}^*, \ldots, Z_{1-p}^*$) is

$$\prod_{t=1}^{n} f(Z_t \mid Z_{t-1}, Z_{t-2}, \ldots, Z_0^*, Z_{-1}^*, \ldots, Z_{1-p}^*; \theta) =$$

$$\prod_{t=1}^{n} (2\pi)^{-\frac{k}{2}} |\Sigma|^{-\frac{1}{2}} \times \exp\left(-\frac{1}{2}(Z_t - \Pi X_t)^T \Sigma^{-1} (Z_t - \Pi X_t)\right)$$

Then, **the conditional log-likelihood function** is

$$L(\theta) := -\frac{nk}{2}\ln(2\pi) - \frac{n}{2}\ln|\Sigma| - \frac{1}{2}\sum_{t=1}^{n}(Z_t - \Pi X_t)^T \Sigma^{-1}(Z_t - \Pi X_t)$$

Taking the derivative w.r.t. Π and Σ, we can get the CMLE estimates. So,

$$\hat{\Pi} := \left(\sum_{t=1}^{n} Z_t X_t^T\right)\left(\sum_{t=1}^{n} X_t X_t^T\right)^{-1}$$

$$\hat{\Sigma} := \frac{1}{n}\sum_{t=1}^{n} \hat{a}_t \hat{a}_t^T = (\hat{\sigma}_{i,j}),$$

where $\hat{a}_t^T := Z_t - \hat{\Pi} X_t$, $\hat{\sigma}_{i,j} := \frac{1}{n}\sum_{t=1}^{n} \hat{a}_{i,t}\hat{a}_{j,t}$.

Denote

$$S_k(\theta) := \sum_{t=1}^{n}(Z_t - \Pi X_t)^T(Z_t - \Pi X_t) = \sum_{t=1}^{n} \|a_t(\Pi)\|^2$$

The *minimizer* of $S_k(\theta)$ is called **OLS of $\boldsymbol{\Pi}$**, which is the same as the CMLE.

Hypothesis testing

Testing zero parameters

We can test the hypothesis of the form

$$H_0 : \omega = 0,$$

where ω is a ν-dimensional vector consisting of some elements from θ.

Let $\hat{\omega}$ be the MLE of ω. Then

$$\hat{\omega}^T \hat{\Omega}^{-1} \hat{\omega} \sim \chi^2_\nu,$$

where $\hat{\Omega}$ is an estimator of the covariance matrix of $\hat{\omega}$. This is called **Wald test**.

Sequential Likelihood Ratio Tests

To adjust the lag order p, we can consider the hypothesis of the form

$$H_0 : \Phi_p = 0$$

against the alternative $H_a : \Phi_p \neq 0$ using **Likelihood Ratio Test**. Meaning that we are testing VAR($p-1$) model against VAR(p) model.

The test statistic is

$$M(l) := -(n - k - l - 1.5) \ln \left(\frac{|\hat{\Sigma}_l|}{|\hat{\Sigma}_{l-1}|} \right),$$

where $\hat{\Sigma}_i$ is based on the residuals of CMLE of VAR(i) model.

Model selection

In order to pick the final model, we need to **choose the lag *p***. It can be done by comparing fitted models with different $p = 1, 2, \ldots, p_{\max}$ using some model selection criteria.

The most common information criteria are the Akaike (AIC), Bayesian (BIC), and Hannan-Quinn (HQ):

$$AIC(p) := \ln |\hat{\Sigma}(p)| + \frac{2}{n} pk^2,$$

$$BIC(p) := \ln |\hat{\Sigma}(p)| + \frac{\ln(n)}{n} pk^2,$$

$$HQ(p) := \ln |\hat{\Sigma}(p)| + \frac{2 \ln \ln(n)}{n} pk^2,$$

where $\hat{\Sigma}(p) := \frac{1}{n} \sum_{t=1}^{n} \hat{a}_t \hat{a}_t^T$ is the residual covariance matrix without a degrees of freedom correction from VAR(p) model.

Model checking

The main assumption of the model is that $\boldsymbol{a}_t$ **are white noise**. Thus, we should check if the CCM of a_t are zero:

$$H_0 : \rho(1) = \rho(2) = \cdots = \rho(m) = 0$$

Test statistic:

$$Q_k(m) := n^2 \sum_{l=1}^{m} \frac{1}{n-l} tr(\hat{C}_l^T \hat{C}_0^{-1} \hat{C}_l \hat{C}_0^{-1}) \sim \chi^2((m-p)k^2),$$

where $\hat{C}_l := \frac{1}{n-l} \sum_{t=l+1}^{n} \hat{a}_t \hat{a}_{t-l}^T$

Forecasting

The **forecast** $\hat{Z}_n(l)$ of Z_{n+l} is calculated by

$$\hat{Z}_n(l) = \mathbb{E}(Z_{n+l} \mid Z_n, Z_{n-1}, \dots) = \Phi_0 + \Phi_1 \hat{Z}_n(l-1) + \dots + \Phi_p \hat{Z}_n(l-p),$$

where $\hat{Z}_n(j) = Z_{n+j}$ if $j \leq 0$.

Forecasting error can be calculated as follows

$$\hat{e}_n(l) = Z_{n+l} - \hat{Z}_n(l) = \sum_{j=0}^{l-1} \Psi_j a_{n+l-j},$$

where Ψ_j can be calculated recursively:

$$\Psi_j = \sum_{k=1}^{p} \Phi_k \Psi_{j-k} \quad \forall j = 1, 2, \dots, l-1$$

with $\Psi_0 = I_k$, $\Psi_j = 0$ for $j < 0$.

Note: It can be proved by induction.

Properties:

$$\mathbb{E}(Z_{n+l} - \hat{Z}_n(l)) = 0$$

$$\Sigma(l) = \mathrm{COV}(\hat{e}_n(l)) = \sum_{s=0}^{l-1} \Psi_s \Sigma \Psi_s^T$$

Forecast intervals:
Asymptotic CI for the individual elements $\hat{Z}_n(l)$ are

$$\left[\hat{Z}_{i,n}(l) - \mathcal{N}_{\frac{\alpha}{2}} \hat{\sigma}_i(l) \ , \ \hat{Z}_{i,n}(l) + \mathcal{N}_{\frac{\alpha}{2}} \hat{\sigma}_i(l) \right],$$

where $\mathcal{N}_{\frac{\alpha}{2}}$ is the $\frac{\alpha}{2}$-quantile of the standard normal distribution, i.e. $P(\mathcal{N}(0,1) > \mathcal{N}_{\frac{\alpha}{2}}) = \alpha/2$, $\hat{\sigma}_i(l)$ is the square root of the diagonal element of $\hat{\Sigma}(l)$ (**the standard errors of prediction**).

Previously, we ignored **the estimator effect**:

$$\tilde{Z}_n(l) = \hat{\Phi}_0 + \hat{\Phi}_1 \tilde{Z}_n(l-1) + \dots + \hat{\Phi}_p Z_n(l-p)$$

«Real» l-step forecasting error is

$$\tilde{e}_n(l) = Z_{n+l} - \tilde{Z}_n(l) = Z_{n+l} - \hat{Z}_n(l) + \hat{Z}_n(l) - \tilde{Z}_n(l) =$$

$$\hat{e}_n(l) + \hat{Z}_n(l) - \tilde{Z}_n(l)$$

It can be shown that

$$\sqrt{n-p}\tilde{e}_n(l) \sim \mathcal{N}(0, \Omega_l)$$

The square root of the diagonal element of Ω_l is called **the root mean squared errors of prediction**.

Vector MA model

Vector MA(1) model

Model

$$Z_t = (I - \Theta_1 B)a_t \Leftrightarrow Z_t = a_t - \Theta_1 a_{t-1}$$

Example: When $k = 2$,

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} a_{1,t} \\ a_{2,t} \end{pmatrix} - \begin{pmatrix} \theta_{1,1} & \theta_{1,2} \\ \theta_{2,1} & \theta_{2,2} \end{pmatrix} \begin{pmatrix} a_{1,t-1} \\ a_{2,t-1} \end{pmatrix},$$

which is equivalent to

$$\begin{cases} Z_{1,t} = a_{1,t} - \theta_{1,1}a_{1,t-1} - \theta_{1,2}a_{2,t-1} \\ Z_{2,t} = a_{2,t} - \theta_{2,1}a_{1,t-1} - \theta_{2,2}a_{2,t-1} \end{cases}$$

Invertibility condition
All **roots** of $|I - \Theta_1 z| = 0$ lie **outside** the unit circle, which is equivalent to the requirement that all **eigenvalues** of Θ_1 lie **inside** the unit circle.

In this case,

$$(I - \Theta_1 B)^{-1} = I + \Theta_1 B + \Theta_1^2 B^2 + \dots$$

and

$$a_t = Z_t + \Theta_1 Z_{t-1} + \Theta_1^2 Z_{t-2} + \dots$$

Covariance matrix function
Let's first notice

$$\mathbb{E}Z_t = \mathbb{E}(a_t - \Theta_1 a_{t-1}) = 0$$

Then,

$$\Gamma(0) = \mathbb{E}Z_t Z_t^T = \mathbb{E}\left(a_t - \Theta_1 a_{t-1}\right)\left(a_t - \Theta_1 a_{t-1}\right)^T =$$

$$\mathbb{E}a_t a_t^T - \mathbb{E}a_t a_{t-1}^T \Theta_1^T - \Theta_1 \mathbb{E}a_{t-1}a_t^T + \Theta_1 \mathbb{E}a_{t-1}a_{t-1}^T \Theta_1^T = \Sigma + \Theta_1 \Sigma \Theta_1^T$$

$$\Gamma(1) = \mathbb{E}Z_t Z_{t-1}^T = \mathbb{E}\left(a_t - \Theta_1 a_{t-1}\right)\left(a_{t-1} - \Theta_1 a_{t-2}\right)^T =$$

$$\mathbb{E}a_t a_{t-1}^T - \mathbb{E}a_t a_{t-2}^T \Theta_1^T - \Theta_1 \mathbb{E}a_{t-1}a_{t-1}^T + \Theta_1 \mathbb{E}a_{t-1}a_{t-2}^T \Theta_1^T = -\Theta_1 \Sigma$$

$$\Gamma(l) = \mathbb{E}Z_t Z_{t-l}^T = \mathbb{E}(a_t - \Theta_1 a_{t-1})Z_{t-l}^T = \mathbb{E}a_t Z_{t-l}^t - \Theta_1 \mathbb{E}a_{t-1}Z_{t-l}^T = 0 \quad \forall l \geq 2$$

$$\Gamma(l) = \begin{cases} \Sigma + \Theta_1 \Sigma \Theta_1^T, & \text{if } l = 0 \\ -\Theta_1 \Sigma, & \text{if } l = 1 \\ -\Sigma^T \Theta_1^T = -\Sigma \Theta_1^T, & \text{if } l = -1 \\ 0, & \text{if } |l| \geq 2 \end{cases}$$

Vector MA(q) model

Model

$$Z_t = (I - \Theta_1 B - \Theta_2 B^2 - \cdots \Theta_q B^q) a_t \Leftrightarrow Z_t = a_t - \Theta_1 a_{t-1} - \cdots - \Theta_q a_{t-q}$$

Invertibility condition
All **roots** of $|I - \Theta_1 z - \Theta_2 z^2 - \cdots \Theta_q z^q| = 0$ lie **outside** the unit circle.

In this case,

$$(I - \Theta_1 B - \Theta_2 B^2 - \cdots - \Theta_q B^q)^{-1} = I - \Pi_1 B - \Pi_2 B^2 + \cdots \Leftarrow$$

$$(I - \Theta_1 B - \Theta_2 B^2 - \cdots - \Theta_q B^q)(I - \Pi_1 B - \Pi_2 B^2 - \dots) = I$$

and

$$a_t = Z_t - \Pi_1 Z_{t-1} - \Pi_2 Z_{t-2} + \dots$$

or

$$Z_t = a_t + \Pi_1 Z_{t-1} + \Pi_2 Z_{t-2} + \dots$$

Covariance matrix function
The covariance matrix function can be found via direct calculation, though it's tedious.

Hypothesis testing

To pick the lag order q, we can consider the hypothesis of the form

$$H_0 : \rho(q) = \rho(q+1) = \cdots = \rho(m) = 0$$

The test statistic is

$$Q_k(q, m) := n^2 \sum_{l=q}^{m} \frac{1}{n-l} tr(\hat{\Gamma}^T(l) \hat{\Gamma}^{-1}(0) \hat{\Gamma}(l) \hat{\Gamma}^{-1}(0)) \sim \chi^2((m - q + 1)k^2),$$

Fitting VMA(q) model

Assume that $Z_1, Z_2, \dots, Z_n$ are from the k-dimensional VMA(q) model with drift

$$Z_t = \mu_0 + a_t - \Theta_{1_0} a_{t-1} - \cdots - \Theta_{q_0} a_{t-q},$$

where $a_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \Sigma_0)$.

Let's denote $\theta := [\mu \quad \Theta_1 \quad \Theta_2 \quad \dots \quad \Theta_q \quad \Sigma]$. Then, we're looking for **the conditional LSE-minimizer** — the minimizer $\hat{\theta}$ of

$$S_n(\theta) = \sum_{t=1}^{n} \|a_t(\theta)\|^2,$$

where $a_t(\theta) := Z_t - \mu + \Theta_1 a_{t-1}(\theta) + \Theta_2 a_{t-2}(\theta) + \cdots + \Theta_q a_{t-q}(\theta)$, introducing initial values $Z_t = a_t = 0$ for $t \leq 0$.

Model checking

The main assumption of the model is that $\boldsymbol{a_t}$ **are white noise**. Thus, we should check if the CCM of a_t are zero:

$$H_0 : \rho(1) = \rho(2) = \cdots = \rho(m) = 0$$

Test statistic for Ljung-Box $Q_k(m) \sim \chi^2(mg)$, where g is the number of estimated parameters.

Forecasting

One-step forecast for VMA(1) model:

$$\hat{Z}_n(1) = \mathbb{E}(Z_{n+1} \mid Z_n, Z_{n-1}, \dots) = \mu - \Theta_1 a_t$$

The associated forecast error and its covariance matrix are

$$e_n(1) = Z_{n+1} - \hat{Z}_n(1) = a_{t+1}$$

$$\mathrm{COV}(e_n(1)) = \Sigma$$

Vector ARMA model

Model

$$\Phi_p(B) Z_t = \Theta_q(B) a_t,$$

where $\Phi_p(z) := I - \Phi_1 z - \Phi_2 z^2 - \cdots - \Phi_p z^p$ and $\Theta_q(z) := I - \Theta_1 z - \Theta_2 z^2 - \cdots - \Theta_q z^q$.

Stationarity condition

All **roots** of $|\Phi_p(z)| = 0$ lie **outside** the unit circle.

Invertibility condition

All **roots** of $|\Theta_q(z)| = 0$ lie **outside** the unit circle.

Covariance matrix function

The covariance matrix function can be found via direct calculation, though it's tedious.

When $p = q = 1$,

$$\Gamma(l) = \begin{cases} \Phi_1\Gamma(-1) + \Sigma - \Theta_1\Sigma(\Phi_1 - \Theta_1)^T, & \text{if } l = 0 \\ \Phi_1\Gamma(0) - \Theta_1\Sigma, & \text{if } l = 1 \\ \Phi_1\Gamma(l-1), & \text{if } l \geq 2 \end{cases}$$

Identifiability condition

The only common left divisors of $\Phi_p(B)$ and $\Theta_q(B)$ are unimodular ones.

Fitting VARMA(p, q) model

Assume that $Z_1, Z_2, \ldots, Z_n$ are from the k-dimensional VARMA(p, q) model with drift

$$Z_t = \Phi_{0_0} + \Phi_{1_0} Z_{t-1} + \cdots + \Phi_{p_0} Z_{t-p} + a_t - \Theta_{1_0} a_{t-1} - \cdots - \Theta_{q_0} a_{t-q},$$

where $a_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \Sigma_0)$.

Let's denote $\theta := [\Phi_0 \quad \Phi_1 \quad \Phi_2 \quad \ldots \quad \Phi_p \quad \Theta_1 \quad \Theta_2 \quad \ldots \quad \Theta_q \quad \Sigma]$. Then, we're looking for **the conditional LSE-minimizer** — the minimizer $\hat{\theta}$ of

$$S_n(\theta) = \sum_{t=1}^{n} \|a_t(\theta)\|^2,$$

where $a_t(\theta) := Z_t - \Phi_0 - \Phi_1 Z_{t-1} - \Phi_2 Z_{t-2} - \cdots - \Phi_p Z_{t-p} + \Theta_1 a_{t-1}(\theta) + \Theta_2 a_{t-2}(\theta) + \cdots + \Theta_q a_{t-q}(\theta)$, introducing initial values $Z_t = a_t = 0$ for $t \leq 0$.

Identification

A Summary Two-Way Table via Multivariate Q-statistic $Q^{(m)}_{(j+1):l}$ in R.

Non-stationary Vector ARMA(p, q) models

There are multiple options to introduce non-stationarity:

1) $\Phi_p(B)(1-B)^d Z_t = \Theta_q(B) a_t$

2) $\Phi_p(B) D(B) Z_t = \Theta_q(B) a_t$, where

$$D(B) = \begin{bmatrix} (1-B)^{d_1} & \mathbb{O} & \ldots & \mathbb{O} \\ \mathbb{O} & (1-B)^{d_2} & \ldots & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{O} & \mathbb{O} & \ldots & (1-B)^{d_m} \end{bmatrix}$$

Usually, $d = 1$ and $W_t = (1-B) Z_t$, where $Z_t = \ln(P_t)$. W_t is called **log-return**.