EC435
บทที่ 2 แบบจำลองอนุกรมเวลาเชิงเส้นตรง

เฉลิมพงษ์ คงเจริญ ©2556-61

คณะเศรษฐศาสตร์ มหาวิทยาลัยธรรมศาสตร์

September 3, 2019

กระบวนการไวท์นอยซ์ (white noise)

ข้อมูลอนุกรมเวลาที่ง่ายที่สุดคือกระบวนการไวท์นอยซ์ (white noise) โดยที่เรามักจะเขียนแทนด้วย random variable ε_t

นิยาม 2.1

ตัวแปร ε_t จะเรียกว่ากระบวนการไวท์นอยซ์ถ้ามีค่าเฉลี่ยเท่ากับศูนย์,มีค่าความแปรปรวนคงที่เท่ากับ σ^2 และไม่มีสหสัมพันธ์ข้ามเวลา (serially uncorrelated) หรือเขียนเป็นเงื่อนไขได้ดังนี้

1. $E(\varepsilon_t) = 0$
2. $Var(\varepsilon_t) = \sigma_\varepsilon^2$
3. $\gamma_k = Cov(\varepsilon_t, \varepsilon_{t-k}) = 0$ for $k \neq 0$

กระบวนการเส้นตรง (Linear process)

Wold's decomposition theorem ระบุว่าอนุกรมเวลานิ่ง y_t สามารถเขียนในรูปกระบวนการเส้นตรงหรือตัวแทนรูปแบบค่าเฉลี่ยเคลื่อนที่ที่มีอันดับเป็นอนันต์ (infinite moving average) และถ้าสามารถเขียน y_t ในรูป

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \tag{1}$$

โดยที่ μ คือค่าเฉลี่ย, $\psi_0 = 1$, และ ε_t คือ (iid) โดยที่เราสามารถพิจารณา ε_t ในฐานะเป็นข้อมูลใหม่ที่เข้ามาในช่วงเวลา t หรือเป็น innovation หรือ shock ณ เวลา t และ $\sum_{i=1}^{\infty} \psi_i^2 < \infty$

แบบจำลองออโตรีเกรสซีฟ

แนวคิดพื้นฐานของแบบจำลองออโตรีเกรสซีฟ (autoregressive) คือข้อมูลในเวลาปัจจุบัน (y_t) สามารถอธิบายได้ด้วยตัวแปรนั้นในอดีตย้อนหลังไป 1 ถึง p ช่วงเวลา ($y_{t-1}, ..., y_{t-p}$)

แบบจำลองออโตรีเกรสซีฟที่มีอันดับ (order) เท่ากับ p (เขียนแทนด้วย $AR(p)$) สามารถเขียนได้เป็น

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t, \qquad (2.3)$$

โดยที่ y_t เป็นข้อมูลที่นิ่งและ ε_t เป็นไวท์นอยซ์ที่มีค่าเฉลี่ยเท่ากับศูนย์และค่าความแปรปรวนคงที่เท่ากับ σ_ε^2

แบบจำลองออโตรีเกรสซีฟ:ค่าเฉลี่ย

ค่าเฉลี่ยของ y_t เท่ากับ
$E(y_t) =$

หากเราเขียน $AR(p)$ ในรูป

$$(y_t - \mu) = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + ... + \phi_p(y_{t-p} - \mu) + \varepsilon_t, \quad (2.4)$$

หรือ $y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t, (2.5)$
$E(y_t) =$

การเขียน AR(p) ในรูป backshift operator

นอกจากนี้เราสามารถเขียนสมการ (2.3) ได้ด้วยเครื่องมือ backshift operator เป็นสมการดังนี้

$$(1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p) y_t = \varepsilon_t, \qquad (2.5)$$

$$\phi(L) y_t = \varepsilon_t$$

โดยที่เราเรียก $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p$ ว่าพหุนามออโตรีเกรสซีฟ (autoregressive polynomial)

AR(1) model

เราจะพิจารณาคุณสมบัติของแบบจำลอง $AR(1)$ ซึ่งมีค่าเฉลี่ยเท่ากับศูนย์ สามารถแสดงได้โดย

$$y_t = \phi y_{t-1} + \varepsilon_t, \qquad (2.8)$$

จากสมการข้างต้นหากเราเริ่มพิจารณากรณีที่ $y_0 = 0$ และให้ ε_t มีการแจกแจงแบบ $N(0, 1)$ เราจะได้การเดินทางตามเวลา (time path) ของ y_t สำหรับค่าสัมประสิทธิ์(ϕ)ที่ต่างกันดังที่แสดงในรูปต่อไปนี้

AR(1) model

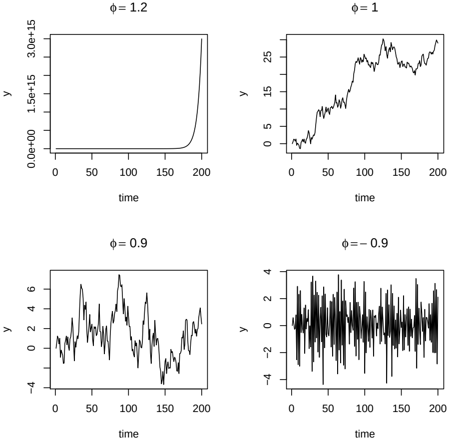


Figure: การจำลองกระบวนการ $AR(1)$ ที่ค่าสัมประสิทธิ์(ϕ)ที่ต่างกัน

จากสมการ (2.8) เราสามารถแทนค่า y_t ในอดีตไปเรื่อยๆแบบเวียนเกิด (recursive) k ครั้งดังนี้

ด้วยวิธีดังกล่าวข้างต้นและค่า $|\phi| < 1$ แล้ว $\lim_{k\to\infty} \phi^k y_{t-k} = 0$ จะทำให้เราสามารถเขียนแบบจำลอง $AR(1)$ ในรูปต่อไปนี้

$$y_t = \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \ldots = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}, \qquad (2.9)$$

เราเรียกรูปดังกล่าวว่าการแสดงแบบจำลอง $AR(1)$ ด้วย infinite moving average representation

ค่าความแปรปรวนของ y_t

เนื่องจาก $E(y_t) = 0$ ค่าความแปรปรวน $Var(y_t) = E\left[y_t - E(y_t)\right]^2$ จะเท่ากับ $Var(y_t) = E(y_t^2)$ หากเราแทนค่า y_t จากสมการ (2.9) ลงในสูตรดังกล่าว และใช้คุณสมบัติของ ε_t ที่ว่า $Var(\varepsilon_t) = E(\varepsilon_t^2) = \sigma_t^2$ และ $E(\varepsilon_j, \varepsilon_k) = 0$ สำหรับ $k \neq j$ เราจะได้

$$E(y_t^2) =$$

ค่าความแปรปรวนร่วม

หากนำค่า y_t และ y_{t-l} ที่เขียนในรูปสมการ (2.9) และค่าเฉลี่ยเท่ากับศูนย์ $E(y_t) = 0$ แทนค่าในสูตรค่าความแปรปรวนร่วมได้ฟังก์ชันค่าความแปรปรวนร่วมจะเท่ากับ

$$\gamma_l =$$

ฟังก์ชันสหสัมพันธ์ร่วมในตัวเอง (ACF)

ฟังก์ชันสหสัมพันธ์ร่วมในตัวเอง (ACF) เท่ากับ $\rho_l = \frac{\gamma_l}{\gamma_0} = \frac{\frac{\phi^l \sigma_\varepsilon^2}{(1-\phi^2)}}{\frac{\sigma_\varepsilon^2}{(1-\phi^2)}} = \phi^l$

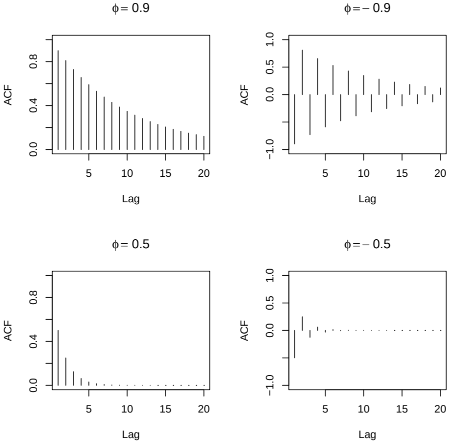
หากลองแทนค่า ϕ ด้วยค่าเท่ากับ $0.9, -0.9, 0.5, -0.5$ จะได้ฟังก์ชันสหสัมพันธ์ร่วมในตัวเองดังตารางต่อไปนี้

Table: Autocorrelation ของ $AR(1)$ ที่ค่าสัมประสิทธิ์(ϕ)ที่ต่างกัน

$\phi \backslash l$	1	2	3	4	5	6	7	8	9	10
0.9	0.900	0.810	0.729	0.656	0.590	0.531	0.478	0.430	0.387	0.349
-0.9	-0.900	0.810	-0.729	0.656	-0.590	0.531	-0.478	0.430	-0.387	0.349
0.5	0.500	0.250	0.125	0.063	0.031	0.016	0.008	0.004	0.002	0.001
-0.5	-0.500	0.250	-0.125	0.063	-0.031	0.016	-0.008	0.004	-0.002	0.001

ฟังก์ชันสหสัมพันธ์ร่วมในตัวเอง (ACF)

Figure: ค่าสหสัมพันธ์ร่วมในตัวเองของกระบวนการ $AR(1)$ ที่ค่าสัมประสิทธิ์(ϕ)ที่ต่างกัน



เงื่อนไขความเป็นอนุกรมนิ่ง

เราสามารถเขียนกระบวนการ $AR(1)$ ในรูปของพหุนามออโตรีเกรสซีฟได้เป็น

$$
\begin{aligned}
y_t - \phi y_{t-1} &= \varepsilon_t \\
(1 - \phi L) y_t &= \varepsilon_t \qquad (2.13)
\end{aligned}
$$

โดยที่เราสามารถเขียน auxiliary equation ได้เป็น

$$(1 - \phi m) = 0$$

Example 1

จงพิจารณาว่ากระบวนการต่อไปนี้เป็นกระบวนการนิ่งหรือไม่

- $y_t - 0.9y_{t-1} = \varepsilon_t$
- $y_t - y_{t-1} = \varepsilon_t$
- $y_t + 1.2y_{t-1} = \varepsilon_t$

แบบจำลอง AR(1) ที่มีค่าเฉลี่ย

เราสามารถเขียน AR(1) ในรูปทั่วไปเช่น $y_t - \mu = \phi(y_{t-1} - \mu) + \varepsilon_t$ ถ้าเราต้องการหา $E(y_t)$

กรณีสมการ $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$ ถ้าเราต้องการหา $E(y_t)$

แบบจำลอง $AR(2)$

แบบจำลอง $AR(2)$ สามารถเขียนอธิบายได้ด้วยสมการ

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

เราสามารถพิจารณาความเป็นกระบวนการนิ่งได้โดยการหารากของพหุนามออโตรีเกรสซีฟซึ่ง $(1 - \phi_1 L - \phi_2 L^2)y_t = \varepsilon_t$ สามารถเขียนได้เป็น

$$\phi(m) = 1 - \phi_1 m - \phi_2 m^2$$

$\phi(m) = 1 - \phi_1 m - \phi_2 m^2 = 0$ โดยที่รากของสมการพหุนามออโตรีเกรสซีฟที่อยู่ในรูปสมการกำลังสอง (quadratic) จะมีด้วยกันสองจำนวน(m_1, m_2)เท่ากับ $\frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}$ โดยที่เงื่อนไขที่คือ รากของสมการพหุนามออโตรีเกรสซีฟจะต้องมากกว่าหนึ่ง (หรือมอดุลัส (modulus) มากกว่าหนึ่งในกรณีจำนวนเชิงซ้อน) ซึ่งจะเกิดขึ้นเมื่อ

$$\phi_1 + \phi_2 < 1, \qquad \phi_2 - \phi_1 < 1, \qquad |\phi_2| < 1$$

คุณลักษณะของกระบวนการ $AR(2)$

ฟังก์ชันความแปรปรวนร่วมเราสามารถพิจารณาได้โดยคูณสมการ (2.16) ทั้งสองข้างด้วย y_{t-k} แล้ว take expectation เราจะได้

$E(y_t y_{t-k}) =$

และหากหารด้วย γ_0 จะได้

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}$$

เราเรียกสมการทั้งสองว่าสมการ Yule-Walker

คุณลักษณะของกระบวนการ $AR(2)$

หากเราพิจารณากรณีที่ $k = 1$, $\rho_1 = \rho_{-1}$ และ $\rho_0 = 1$ เราจะได้

$$\rho_1 = \frac{\phi_1}{1-\phi_2}$$

และกรณีที่ $k = 2$ เราจะได้

$$\rho_2 = \frac{\phi_2(1-\phi_2)+\phi_1^2}{1-\phi_2}$$

และเราสามารถแทนค่าแบบเวียนเกิด(recursive)เพื่อให้ได้ ρ_k กรณี $k > 2$

ให้ไปอ่านเนื้อหาสำหรับแบบจำลอง AR(p)

แบบจำลอง $AR(p)$

กรณีที่ $AR(p)$ เป็นอนุกรมนิ่ง ค่าคงที่ในรูปแบบถดถอยในตัวเอง(c)จะเท่ากับ $\mu(1-\phi_1-...-\phi_p)$ ในทางกลับกัน $\mu = c/(1-\phi_1-...-\phi_p)$

หากเราสมมุติให้ค่าเฉลี่ย(μ)เท่าศูนย์และกระบวนการนิ่ง เราจะได้

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + ... + \phi_p\rho_{k-p},$$

สำหรับทุกค่า $k \geq 1$ หากเราพิจารณากรณีที่ $k = 1, 2, ..., p$ และใช้ความสัมพันธ์ที่ $\rho_0 = 1$ และ $\rho_j = \rho_{-j}$ เราจะได้สมการ Yule-Walker

$$\begin{aligned}
\rho_1 &= \phi_1 + \phi_2\rho_1 + \phi_3\rho_2 + ... + \phi_p\rho_{p-1} \\
\rho_2 &= \phi_1\rho_1 + \phi_2 + \phi_3 + ... + \phi_p\rho_{p-2} \\
&\vdots \\
\rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \phi_3\rho_{p-3} + ... + \phi_p
\end{aligned} \tag{2}$$

แบบจำลอง $AR(p)$

ซึ่งหากเราทราบว่า $\phi_1, \phi_2, ..., \phi_p$ เราสามารถหาค่า $\rho_1, \rho_2, ..., \rho_3$ ได้ นอกจากนี้

$$E(\varepsilon_t y_t) = E[\varepsilon_t(\phi_1 y_{t-1} + ... + \phi_p y_{t-p} + \varepsilon_t)] = E(\varepsilon_t^2) = \sigma^2$$

และหากเราคูณสมการ (2.20) ด้วย y_t และใส่ค่าคาดหมาย เราจะได้

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + ... + \phi_p\gamma_p + \sigma^2 \tag{3}$$

และหากใช้ความสัมพันธ์ว่า $\rho_k = \gamma_k/\gamma_0$ จะได้

$$\gamma_0 = \frac{\sigma^2}{1 - \phi_1\rho_1 - \phi_2\rho_2 - ... - \phi_p\rho_p}$$

Partial Autocorrelation Function

PACF ปีนเครื่องมือที่เป็นประโยชน์ในการระบุแบบจำลอง $AR(p)$

$$\begin{aligned}
z_t &= \phi_{11} z_{t-1} + \varepsilon_{1t} \\
z_t &= \phi_{21} z_{t-1} + \phi_{22} z_{t-2} + \varepsilon_{2t} \\
&\vdots \\
z_t &= \phi_{p1} z_{t-1} + \phi_{p2} z_{t-2} + \ldots + \phi_{pp} z_{t-p} + \varepsilon_{pt}
\end{aligned}$$

โดยที่ $z_t = y_t - \mu$ คือข้อมูลที่ได้กำจัดค่าเฉลี่ยแล้ว (demeaned) เราเรียกค่าสัมประสิทธิ์ ϕ_{jj} สำหรับ $j = 1, 2, ..., p$ (ค่าสัมประสิทธิ์สุดท้ายของแต่ละสมการ) ว่า*ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน*

PACF

ในกรณีที่เราพิจารณาแบบจำลอง $AR(1)$ ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนตัวแรก ϕ_{11} จะไม่เท่ากับศูนย์ ส่วนสัมประสิทธิ์ที่เหลือจะเท่ากับศูนย์

หากเราพิจารณาแบบจำลอง $AR(2)$ ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนตัวแรกและตัวที่สอง (ϕ_{11} และ ϕ_{22}) จะไม่เท่ากับศูนย์ ส่วนสัมประสิทธิ์ที่เหลือ (ϕ_{jj} สำหรับ $j > 2$) จะเท่ากับศูนย์

โดยสรุปแล้ว สำหรับแบบจำลอง $AR(p)$ ใดๆ ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน p ตัวแรกจะไม่เท่ากับศูนย์ และสัมประสิทธิ์ที่เหลือจะเท่ากับศูนย์

สามารถคำนวณได้โดยใช้ฟังก์ชัน pacf ใน R

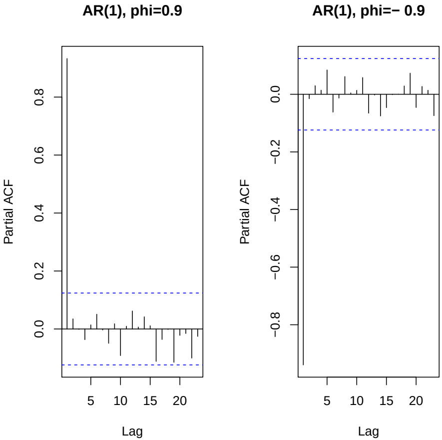
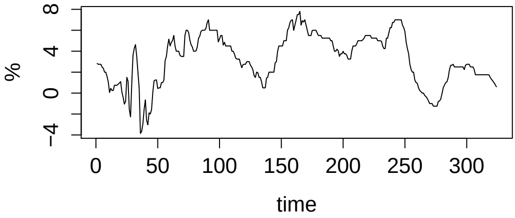


Figure: ฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วนของข้อมูลจำลอง $AR(1)$

ตัวอย่าง 2.3

```
> int<-read.csv(file="mlr.csv", header=T)
> head(int)
> plot(int$diff_th_us,type="l", ylab="%", xlab="time")
```

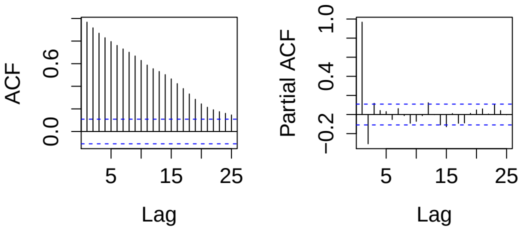
Figure: ความแตกต่างระหว่างอัตราดอกเบี้ยของไทยและสหรัฐอเมริกา



ตัวอย่างที่ 2.3: PACF

```
> par(mfrow=c(1,2))
> library(TSA)
> Acf(int$diff_th_us)
> pacf(int$diff_th_us)
```

Figure: ACF และ PACF ของความแตกต่างระหว่างอัตราดอกเบี้ยของไทยและสหรัฐอเมริกา



MLE estimation

หากเราสนใจที่จะประมาณค่าแบบจำลอง $AR(1)$ ซึ่งอยู่ในรูปดังต่อไปนี้

$$y_t = \mu + \phi(y_{t-1} - \mu) + \varepsilon_t$$

โดยที่ $\varepsilon_t \sim iidN(0, \sigma^2)$ เนื่องจากข้อมูลอนุกรมเวลาเป็นข้อมูลที่ขึ้นอยู่ต่อกัน (dependent) ดังนั้น หากเราเก็บข้อมูล $y_1, ..., y_T$ แล้วฟังก์ชันค่าความควรจะเป็นสามารถเขียนได้ดังนี้

$$L(\mu, \phi, \sigma^2) = f(y_1)f(y_2|y_1)f(y_3|y_2)\cdots f(y_T|y_{T-1}) \tag{4}$$

เนื่องจาก $y_t|y_{t-1} \sim N(\mu + \phi(y_{t-1} - \mu), \sigma^2)$ และฟังก์ชันการแจกแจงจะเป็นแบบปกติที่เหมือนกับช็อก

$$f(y_t|y_{t-1}) = f_\varepsilon[(y_t - \mu) - \phi(y_{t-1} - \mu)]$$

MLE estimation

$$L(\mu,\phi,\sigma^2) = \exp\left(\frac{\sum_{t=2}^{T}[(y_t-\mu)-\phi(y_{t-1}-\mu)]}{2\sigma^2}\right) \tag{2.30}$$

การหาประมาณค่าโดยหาค่าพารามิเตอร์ที่ทำให้ log ของสมการ (2.30) สูงที่สุด เรียกว่า การประมาณค่าความควรจะเป็นสูงสุดแบบแม่นตรง (exact maximum likelihood estimation)

กรณีนี้เราไม่สามารถหาค่าสูงสุดได้ด้วยการวิเคราะห์หาเงื่อนไขจำเป็น เราต้องใช้กระบวนการด้วยเครื่องคอมพิวเตอร์ในการหาค่าสูงสุด

Model Checking

หลังจากที่เราได้ตัวประมาณค่าสำหรับพารามิเตอร์แล้ว ขั้นต่อไปเราก็จะหาค่า fitted ของตัวแปรที่เราศึกษา ($\hat{y}_t$)

แล้วก็จะตรวจสอบแบบจำลองโดยการวิเคราะห์ค่าส่วนเกิน (residuals) โดยในที่นี้ค่าส่วนเกินสามารถคำนวณได้จาก $\hat{\varepsilon}_t = y_t - \hat{y}_t$ โดยเราจะประเมินค่าส่วนเกินดังต่อไปนี้

เราอาจจะเริ่มจากการวาดแผนภาพ ACF งของ $\hat{\varepsilon}_t$ เพื่อดูว่าอนุกรมดังกล่าวยังมีความขึ้นอยู่ต่อกันหรือไม่

นอกจากนี้เราสามารถทดสอบโดยใช้ค่าสถิติ $Q(m) = T(T+2)\sum_{k=1}^{m} \frac{\hat{\rho}_{\varepsilon}{}_k^2}{T-k}$ โดยที่ $Q(m) \sim \chi^2_{m-g}$ โดยที่ g คือจำนวนของอันดับในแบบจำลองซึ่งในแบบจำลอง $AR(p)$ ค่า $g = p$ และ m คือจำนวนค่าล่าของความสัมพันธ์ที่พิจารณา

ตัวอย่าง 2.3 ต่อ

```
> m1<-arima(int$diff_th_us, order=c(3,0,0))
> m1
Call:
arima(x = int$diff_th_us, order = c(3, 0, 0))
Coefficients:
         ar1      ar2     ar3  intercept
      1.3294  -0.4986  0.1334     3.0553
s.e.  0.0549   0.0878  0.0549     0.8102
sigma^2 estimated as 0.3054:  log likelihood = -269.11,  aic = 548.21
```

โดยที่ค่าสัมประสิทธิ์ $\phi_1 = 1.3294, \phi_2 = -0.4986, \phi_3 = 0.1334, \mu(= \textit{intercept}) = 3.0553$ และมีค่า standard errors อยู่ในบรรทัด `s.e.`

ตัวอย่าง 2.3 ต่อ

ในโปรแกรม R ค่า `intercept` ที่ได้คือ μ มิใช่จุดตัดแกนในสมการที่ **??** ดังนั้นถ้าต้องการหากเราต้องการค่า $c = \mu(1 - \phi_1 - \phi_2 - \phi_3) = 0.10937$
เราสามารถเขียนผลการประมาณค่าได้โดย

$$(1 - \underset{(0.0549)}{1.3294}\,L + \underset{(0.0878)}{0.4986}\,L^2 - \underset{(0.0549)}{0.1334}\,L^3)(y_t - \underset{(0.8102)}{3.0553}) = \varepsilon_t$$

โดยที่ค่าในวงเล็บคือ standard errors และ $\widehat{\sigma}^2 = 0.3054$

ตัวอย่าง 2.3 ต่อ

```
> m2<-arima(int$diff_th_us, order=c(3,0,0),method=c("ML"))
> m2
Coefficients:
         ar1      ar2     ar3  intercept
      1.3295  -0.4987  0.1334      3.056
s.e.  0.0549   0.0878  0.0549      0.810

sigma^2 estimated as 0.3054:  log likelihood = -269.11,  aic = 548.21
> m3<-arima(int$diff_th_us, order=c(3,0,0),method=c("CSS"))

Coefficients:
         ar1      ar2     ar3  intercept
      1.3334  -0.5016  0.1346     3.0793
s.e.  0.0551   0.0880  0.0552     0.9195
sigma^2 estimated as 0.3082:  part log likelihood = -269.05
```

ตัวอย่าง 2.3 ต่อ

หลังจากที่เราได้ตัวประมาณค่าของพารามิเตอร์ทั้งหมดแล้วเราจะพิจารณาว่าค่า residuals นั้นยังมีความสัมพันธ์ระหว่างช่วงเวลาหลงเหลืออยู่หรือไม่ หรือแบบจำลองเพียงพอหรือไม่ โดยการทดสอบ L-B test สำหรับ `m2$residuals` ด้วยคำสั่งข้างล่าง

```
> Box.test(m2$residuals,lag=12,type="Ljung")
^^IBox-Ljung test
data:  m2$residuals
X-squared = 14.0085, df = 12, p-value = 0.3002
> pv=1-pchisq(14.0085,9)
> pv
[1] 0.1220232
```

การพยากรณ์ด้วยแบบจำลอง $AR(p)$

สมมุติให้เรากำลังอยู่ ณ ช่วงเวลาที่ h แต่ต้องการที่จะพยากรณ์ไปข้างหน้า l คาบเวลา หรือสนใจค่าของ y_{h+l} โดยที่ $l \geq 1$ เราเรียก h ว่าจุดเริ่มต้นการพยากรณ์ (forecast origin) และ l ว่าขอบเขตการพยากรณ์ (forecast horizon)

กำหนดให้ $\hat{y}_h(l)$ เป็นค่าพยากรณ์ของ y_{h+l} โดยการพยากรณ์ดังกล่าวจะทำให้ฟังก์ชันสูญเสียที่เกิดจากการค่าผิดพลาดกำลังสอง (squared errors) น้อยที่สุด

$$E\left[(y_{h+l} - \hat{y}_h(l))^2 | F_h\right] \leq \min_g E\left[(y_{h+l} - g)^2 | F_h\right]$$

โดยที่ F_h เป็นข้อมูลที่มีทั้งหมดในคาบที่ h และเราเรียก $\hat{y}_h(l)$ ว่าค่าพยากรณ์ของ y_t ไปข้างหน้า l คาบเมื่อเริ่มต้นการพยากรณ์อยู่ที่ h

ค่าพยากรณ์ไปข้างหน้า 1 คาบจาก $AR(p)$

เราพบว่าค่า $y_{h+1} = \phi_0 + \phi_1 y_{h+1-1} + \ldots + \phi_p y_{h+1-p} + \varepsilon_{h+1}$ ตัวพยากรณ์ที่จะทำให้ค่าฟังก์ชันสูญเสียน้อยที่สุดคือค่าคาดการณ์แบบมีเงื่อนไข (conditional expectation)

$\hat{y}_h(1) =$

และค่าผิดพลาดจากการพยากรณ์เท่ากับ $e_h(1) =$

และค่าแปรปรวนจากการพยากรณ์ไปข้างหน้าหนึ่งคาบเท่ากับ

$Var(e_h(1)) =$

เราสามารถสร้างช่วงความเชื่อมั่น 95 % ของการพยากรณ์ไปข้างหน้า 1 คาบได้โดยช่วงดังกล่าวจะเท่ากับ $\hat{y}_h(1) \pm 1.96 sd(e_h(1))$

ค่าพยากรณ์ไปข้างหน้า 2 คาบจาก $AR(p)$

เราพบว่าค่า $y_{h+2} = \phi_0 + \phi_1 y_{h+2-1} + \ldots + \phi_p y_{h+2-p} + \varepsilon_{h+2}$ ตัวพยากรณ์ที่จะทำให้ค่าฟังก์ชันสูญเสียน้อยที่สุดคือค่าคาดการณ์แบบมีเงื่อนไข (conditional expectation)

$\hat{y}_h(2) =$

และค่าผิดพลาดจากการพยากรณ์เท่ากับ

$e_h(2) =$

และค่าแปรปรวนจากการพยากรณ์ไปข้างหน้าหนึ่งคาบเท่ากับ

$Var(e_h(2)) =$

เราจะสังเกตได้ว่า $Var(e_h(2)) \geq Var(e_h(1))$

ตัวอย่าง 2.3 ผลการพยากรณ์

```
> m4.pred=predict(m4,n.ahead=12)
> m4.pred
$pred
Time Series:
Start = 313
End = 324
Frequency = 1
 [1] 1.803828 1.875260 1.943271 2.005150 2.062937 2.117888 2.170297 2.220198
 [9] 2.267658 2.312788 2.355708 2.396527
$se
Time Series:
Start = 313
End = 324
Frequency = 1
 [1] 0.5622779 0.9342948 1.1737163 1.3399272 1.4696143 1.5777365 1.6700163
 [8] 1.7494732 1.8183586 1.8784737 1.9312318 1.9777417
```

แบบจำลองมูวิ่งเอเวอเรจ (Moving Average; $MA(q)$)

แบบจำลองนี้มีพื้นฐานมากจากแนวคิดที่ว่าอนุกรมเวลาเส้นตรงนิ่ง (stationary) สามารถเขียนให้อยู่ในรูป infinite MA

แบบจำลองมูวิ่งเอเวอเรจอันดับ q และเขียนแทนด้วย $MA(q)$ และแสดงได้ด้วยสมการ

$$\begin{aligned} y_t &= \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \\ &= \sum_{j=0}^{q} \theta_j \varepsilon_{t-j} = \theta(L)\varepsilon_t \qquad (2.37) \end{aligned}$$

โดยที่ $\theta_0 = 1$

แบบจำลอง MA(1)

สมมุติให้ y_t เป็นอนุกรม $MA(1)$ โดยที่ ε_t มีคุณสมบัติเป็นไวท์นอซ

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} \qquad (2.38)$$

ค่าคาดหมายของ y_t, $E(y_t)$, เท่ากับ

$E(y_t) =$

ค่าความแปรปรวนของ y_t (γ_0)

$\gamma_0 = Var(y_t) =$

แบบจำลอง MA(1)

one-lag autocovariance

$\gamma_1 = Cov(y_t, y_{t-1}) =$

two-lag autocovariance

$\gamma_2 = Cov(y_t, y_{t-2}) =$

แบบจำลอง MA(1)

จากสมการข้างต้นจะเห็นได้ว่า $Cov(y_t, y_{t-k}) = 0$ สำหรับ $k \geq 2$ ในกรณีที่ y_t เป็น $MA(1)$

k-lag autocorrelation เท่ากับ

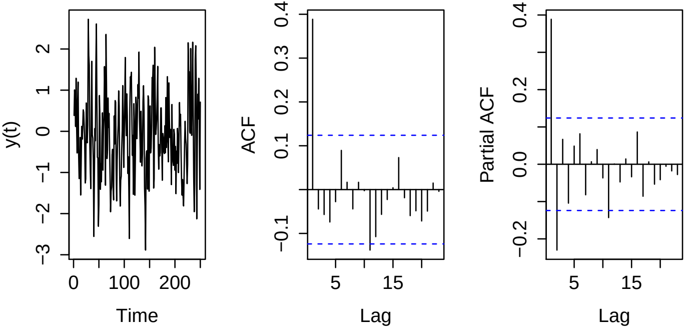
$$\rho_k = \begin{cases} \frac{\theta_1}{1+\theta_1^2} & , k = 1 \\ 0 & , k > 1 \end{cases}$$

จะเห็นได้ว่า y_t มีสหสัมพันธ์กับ y_{t-1} แต่ไม่สัมพันธ์กับ $y_{t-2}, \ldots$ ซึ่งแตกต่างจาก $AR(1)$

เราจะเห็นได้ว่าอนุกรมเวลา y_t ที่เป็นกระบวนการมูวิ่งเอเวเรจจะเป็นอนุกรมเวลานิ่ง (stationary) เสมอ

ตัวอย่างข้อมูลที่เป็น MA(1)

Figure: การจำลองข้อมูลกระบวนการ $MA(1)$ ที่มีค่าสัมประสิทธิ์ $\theta = 0.5$



Invertible process

ฟังก์ชันสหสัมพันธ์ร่วมในตัวเองของกระบวนการ $MA(1)$ สองกระบวนการต่อไปนี้จะมีลักษณะที่เหมือนกัน

$$x_t = w_t + \frac{1}{5}w_t, \qquad w_t \sim iidN(0, 25)$$

และ

$$y_t = v_t + 5v_t, \qquad v_t \sim iidN(0, 1)$$

เราต้องเลือกว่าจะใช้แบบจำลองไหน ซึ่งเราจะเลือกแบบจำลองที่สามารถเขียนแบบจำลองในรูปของ infinite AR representation โดยที่อนุกรมที่มีลักษณะดังกล่าวเราจะเรียกว่ากระบวนการที่หาตัวผกผันได้(invertible process)

เงื่อนไขที่ $MA(1)$ จะเป็นกระบวนการที่หาตัวผกผันได้ คือ $|\theta_1| < 1$

แบบจำลอง $MA(2)$

กำหนดให้ y_t เป็นกระบวนการ $MA(2)$ ซึ่งสามารถแสดงได้ด้วยสมการ

$$y_t - \mu = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}$$

เมื่อพิจารณาค่าสหสัมพันธ์ในตัวเองของ $MA(2)$ จะได้

$$\rho_0 = \frac{\gamma_0}{\gamma_0} = 1$$

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(\theta_1 + \theta_2\theta_1)}{(1 + \theta_1^2 + \theta_2^2)}$$

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \frac{\theta_2}{(1 + \theta_1^2 + \theta_2^2)}$$

$$\rho_3 = \frac{\gamma_3}{\gamma_0} = 0$$

$$\rho_j = \frac{\gamma_j}{\gamma_0} = 0$$

แบบจำลอง $MA(q)$

แบบจำลอง $MA(q)$ สามารถเขียนในรูปสมการดังนี้

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q}, \qquad \varepsilon_t \sim WN(0, \sigma^2)$$

แบบจำลอง $MA(q)$ นิ่งและ ergodic ถ้าค่า $\theta_1, ..., \theta_q$ มีค่าจำกัด และสามารถหาตัวผกผันได้ถ้าค่าราก (roots) ของพหุนาม MA $\theta(m) = 1 + \theta_1 m + ... + \theta_q m^{t-q} = 0$ มีค่าสัมบูรณ์(หรือค่าโมดุลัส)ของรากของพหุนามมากกว่า 1

แบบจำลอง $MA(q)$

โมเมนต์ของกระบวนการ $MA(q)$ ได้ดังนี้

$E(y_t) = \gamma_0 = \gamma_j =$

และค่าสหสัมพันธ์ในตัวเองจะเท่ากับ

$\rho_j = \frac{\gamma_j}{\gamma_0} =$

ในขณะที่ PACF จะมีค่อยลดลงไป

การประมาณค่าแบบจำลอง $MA(q)$

การประมาณค่า MLE เป็นวิธีการทั่วไปที่ใช้ในการประมาณค่าแบบจำลอง $MA(q)$

ประมาณค่าแบบจำลอง $MA(1)$ ซึ่งสามารถเขียนได้ดังนี้

$$y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$$

โดยที่ $\varepsilon_t \sim iidN(0, \sigma_\varepsilon^2)$ ในกรณีที่ conditional MLE เราจะสมมุติให้ $\varepsilon_0 = 0$ ดังนั้น $y_1 = \mu + \varepsilon_1$

$$y_1|\varepsilon_0 = 0 \sim N(\mu, \sigma_\varepsilon^2)$$

จากข้อมูลในคาบที่ 1 จะได้ว่า $y_2|y_1 = \mu + \varepsilon_2 + \theta\varepsilon_1$ และได้ฟังก์ชันความน่าจะเป็น

$$f_{y_2|y_1} = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left[\frac{-(y_2 - \mu - \theta\varepsilon_1)^2}{2\sigma_\varepsilon^2}\right]$$

การประมาณค่าแบบจำลอง $MA(q)$

หากใช้กระบวนการเดียวกันไปเรื่อย เราจะสามารถหาค่า $\varepsilon_t = y_t - \mu - \theta\varepsilon_{t-1}$ และสร้าง conditional likelihood function ได้

$$L(\mu, \theta, \sigma_\varepsilon^2) = \Pi_{t=2}^{T} f_{y_t|y_{t-1}}$$

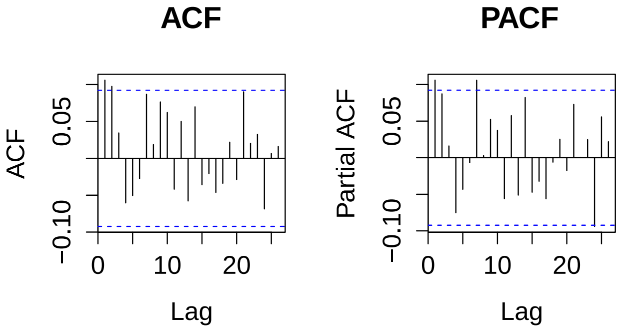
$$= (2\pi\sigma_\varepsilon^2)^{-(T-2)/2} \exp\left[\frac{\sum_{t=2}^{T} - \overbrace{(y_t - \mu - \theta\varepsilon_{t-1})^2}^{\varepsilon_t^2}}{2\sigma_\varepsilon^2}\right] \tag{2.50}$$

สมการ (2.50) เราไม่สามารถหาตัวประมาณค่าได้ด้วยการวิเคราะห์อนุพันธ์เช่นในกรณี AR ในกรณีนี้เราจำเป็นต้องใช้วิธีการทางตัวเลขด้วยคอมพิวเตอร์เพื่อหาค่า $\hat{\theta}$ ที่ทำให้ $L(\mu, \theta, \sigma_\varepsilon^2)$ มีค่าสูงที่สุด

ตัวอย่างที่ 2.6

ในตัวอย่างนี้เราจะพิจารณาผลได้ตอบแทนรายเดือนของการลงทุนในตลาดหลักทรัพย์แห่งประเทศไทย โดยใช้ข้อมูลราคาปิดรายเดือนตั้งแต่ เมษายน 2518 ถึงเดือนพฤศจิกายน 2555

Figure: ACF และ PACF ของผลได้ตอบแทนรายเดือนจาก SET



ตัวอย่างที่ 2.6

แบบจำลองที่เหมาะสมน่าจะเป็น $MA(2)$ ซึ่งประมาณค่าได้ด้วยคำสั่ง `arima` โดยกำหนดอันดับเป็น `c(0,0,2)`

```
> m1<-arima(ret, order=c(0,0,2))
> m1
Call:
arima(x = ret, order = c(0, 0, 2))

Coefficients:
         ma1     ma2  intercept
      0.0886  0.1001     0.0057
s.e.  0.0469  0.0487     0.0046

sigma^2 estimated as 0.006795:  log likelihood = 485.64,  aic = -963.28
```

ตัวอย่างที่ 2.6

เราสามารถพิจารณาความพอเพียงของแบบจำลอง `m1` ได้ด้วยการทดสอบ L-B สำหรับ residuals ดังคำสั่งต่อไปนี้

```
> Box.test(m1$residuals, lag=12, type="Ljung")
^^IBox-Ljung test
data:  m1$residuals
X-squared = 13.6075, df = 12, p-value = 0.3265
> pv=1-pchisq(13.6075,12-2)
> pv
[1] 0.1916591
```

การพยากรณ์ไปข้างหน้า 1 คาบ

พิจารณาแบบจำลอง $MA(1)$: $y_{h+1} = \mu + \varepsilon_{h+1} + \theta_1 \varepsilon_h$ ตัวพยากรณ์ที่จะทำให้ค่าฟังก์ชันสูญเสียน้อยที่สุดคือค่าคาดการณ์แบบมีเงื่อนไข (conditional expectation)

$\hat{y}_h(1)$

และค่าผิดพลาดจากการพยากรณ์เท่ากับ $e_h(1) =$

และค่าแปรปรวนจากการพยากรณ์ไปข้างหน้าหนึ่งคาบเท่ากับ $Var(e_h(1)) =$

ในทางปฏิบัติค่า ε_h จะสามารถคำนวณได้สองวิธีคือ

- สมมุติให้ $\varepsilon_0 = 0$
- หรือใช้ค่า $\hat{\varepsilon}_h$ ที่เป็นค่า residual จากการประมาณค่า $MA(1)$

การพยากรณ์ไปข้างหน้า 2 คาบ

พิจารณา $y_{h+2} = \mu + \varepsilon_{h+2} + \theta_1 \varepsilon_{h+1}$ ตัวพยากรณ์ที่จะทำให้ค่าฟังก์ชันสูญเสียน้อยที่สุดคือค่าคาดการณ์แบบมีเงื่อนไข

$\hat{y}_h(2) =$

และค่าผิดพลาดจากการพยากรณ์เท่ากับ $e_h(2) =$

และค่าแปรปรวนจากการพยากรณ์ไปข้างหน้าสองคาบเท่ากับ

$Var(e_h(2)) =$

Autoregressive Moving Average; $ARMA(p, q)$)

แบบจำลองที่มีทั้งส่วนประกอบของ AR และ MA ซึ่งหาก y_t เป็นกระบวนการที่เรียกว่าแบบจำลองออโตรีเกรสซีฟมูววิ่งเอเวอเรจอันดับ (p, q) หรือเรียกย่อๆว่า อารมา $ARMA(p, q)$ ถ้า y_t มีค่าเฉลี่ยไม่เท่ากับศูนย์ เราจะสามารถอธิบาย y_t ได้ด้วย

$$y_t = \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q} \qquad (2.53)$$

หากเรากำหนดให้ $\alpha = \mu(1 - \phi_1 - ... - \phi_p)$ และสามารถเขียนแบบจำลองสำหรับ y_t ได้เป็น

$$y_t = \alpha + \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q} \qquad (2.54)$$

Autoregressive Moving Average; $ARMA(p, q)$

เราสามารถเขียนสมการอารมาได้ในรูป AR และ MA polynomials ได้ดังนี้

$$\phi(L)y_t = \theta(L)\varepsilon_t$$

โดยที่ $\phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$ และ $\theta(L) = 1 + \theta_1 L + ... + \theta_q L^q$
พิจารณาคุณสมบัติของแบบจำลองอารมาได้ด้วยการพิจารณา $ARMA(1,1)$

$$y_t = \phi y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1} \qquad (2.55)$$

โดยที่ $\varepsilon_t \sim iidN(0, \sigma^2)$

คุณสมบัติ ARMA(1,1)

เราใช้ความสัมพันธ์ต่อไปนี้

$E(\varepsilon_t y_t) =$

$E(\varepsilon_{t-1} y_t) =$

$E(y_t) =$

คุณสมบัติ ARMA(1,1)

คุณสมบัติของ $ARMA(1, 1)$ จะได้

$\gamma_0 =$

$\gamma_1 =$

$\gamma_j =$

คุณสมบัติ ARMA(1,1)

กรณีที่ $j \geq 2$ จากสมการที่ (2.56) และ (2.57) เราจะได้
$Var(y_t) = \gamma_0 =$

ACF ณ ค่าล่า j ใดๆที่ $j \geq 2$ ได้

$$\rho_j = \phi \rho_{j-1}$$

กรณีที่ $j \geq 2$ จะเห็นได้ว่า ACF มีค่าที่ลดลงเรื่อยๆด้วยอัตราแบบ exponential)

คุณสมบัติ ARMA(1,1)

แบบจำลอง $ARMA(1,1)$ ในรูปของมูววิ่งเอเวอเรจที่มีอันดับเป็นอนันต์ (infinite MA)ได้เป็น

$$y_t = \varepsilon_t + (\phi - \theta)\sum_{j=1}^{\infty} \phi^{j-1}\varepsilon_{t-j}$$

หรือ $\psi_j = (\phi - \theta)\phi^{j-1}$ จะเห็นได้ว่าเงื่อนไขที่จะทำให้กระบวนการนี้เป็นกระบวนการนิ่งคือ $|\phi| < 1$

สรุปคุณสมบัติของ $ARMA(p, q)$

- กระบวนการ $ARMA(p, q)$ จะ stationary และ ergodic ถ้าค่าสัมบูรณ์ของค่ารากของพหุนามออโตรีเกรสซีฟ $\phi(m) = 0$ มีค่ามากกว่าหนึ่ง
- และสามารถหาค่าผกผันได้ถ้ารากของพหุนามมูวิ่งเอเวอเรจ $\theta(m) = 0$ มีค่ามากกว่าหนึ่ง
- พหุนาม AR และ MA ไม่มีตัวประกอบร่วม

รูปทั่วไปของ ACF ของกระบวนการ $ARMA(p, q)$ ค่อนข้างยุ่งยาก โดยสรุปแล้วทั้ง ACF และ PACF จะค่อยๆลดลงเรื่อยแบบเลขชี้กำลัง (exponential)

การประมาณค่าสมการแบบจำลอง ARMA

เราสามารถประมาณค่าแบบจำลอง $ARMA(p,q)$ ได้ด้วยวิธีการค่าความควรจะเป็นสูงที่สุด (MLE)

- ฟังก์ชันที่แน่นอน (exact log-likelihood) นั้นมีรูปแบบที่ค่อนข้างซับซ้อน นอกจากนี้จุดเริ่มต้นของข้อมูลในกรณี exact likelihood จะใช้คุณสมบัติของการแจกแจงที่นิ่งเพื่อสร้างฟังก์ชันค่าควรจะเป็นของค่า y_t p ค่าแรก และ ε_t q ค่าแรก
- conditional likelihood จะสมมุติให้ y_t p ค่าแรก และ ε_t q ค่าแรกเท่ากับศูนย์

กรณีที่จำนวนตัวอย่างมีจำนวนเข้าใกล้อนันต์ค่าประมาณจากสองวิธีจะมีค่าใกล้เคียงกัน แต่จะแตกต่างกันในกรณีตัวอย่างมีจำนวนน้อย

การประมาณค่าสมการแบบจำลอง ARMA

ทดสอบความเพียงพอของแบบจำลองได้เช่นเดียวกับในกรณีของแบบจำลอง AR และ MA โดยที่ตัวสถิติ $Q(m) \sim \chi^2_{m-p-q}$ และเราจะปฏิเสธสมมุติฐานหลักถ้า $Q(m)$ มีค่ามากกว่าควอนไทล์ที่ $(1-\alpha)$ ของ χ^2_{m-p-q}

ก่อนที่เราจะประมาณค่าแบบจำลอง $ARMA(p,q)$ สำหรับอนุกรมเวลา y_t ใดๆ เราจะเป็นต้องระบุลำดับของ AR (p) และ MA (q) เสียก่อน

เกณฑ์การเลือกแบบจำลอง (Model Selection Criteria)

เราสามารถใช้การสังเกต ACF ของตัวอย่าง ในกรณีของ MA หรือ PACF ของตัวอย่าง ในกรณีของ AR

Table: สรุปลักษณะของ ACF และ PACF สำหรับแบบจำลอง ARMA

	$AR(p)$	$MA(q)$	$ARMA(p, q)$
ACF	ค่อยๆลดลง	ค่าเท่ากับศูนย์ หลังจากช่วงล่าที่ q	ค่อยๆลดลง
PACF	เท่ากับศูนย์ หลังจากช่วงล่าที่ p	ค่อยๆลดลง	ค่อยๆลดลง

เกณฑ์การเลือกแบบจำลอง (Model Selection Criteria)

อีกทางเลือกหนึ่งเราสามารถจะใช้เกณฑ์การเลือกแบบจำลอง
แนวคิดคือการประมาณค่าแบบจำลอง $ARMA(p, q)$ สำหรับค่าอันดับ p และ q ต่างๆที่น้อยกว่าค่าหนึ่งที่เรากำหนดไว้ p_{max} และ q_{max} และเลือกค่า p และ q ที่ทำให้ค่าเกณฑ์การเลือกต่ำที่สุด โดยที่เกณฑ์การเลือกจะอยู่ในรูป

$$MSC(p, q) = \ln(\tilde{\sigma}^2(p, q)) + c_T \varphi(p, q)$$

เกณฑ์การเลือกแบบจำลอง (Model Selection Criteria)

โดยที่แบบเกณฑ์ที่เป็นที่นิยมสามเกณฑ์ได้แก่ Akaike (AIC) Schwarz-Bayesian (BIC) และ Hannan-Quinn (HQIC)

$$
\begin{aligned}
AIC(p,q) &= \ln(\tilde{\sigma}^2(p,q)) + \frac{2}{T}(p+q) \\
BIC(p,q) &= \ln(\tilde{\sigma}^2(p,q)) + \frac{\ln T}{T}(p+q) \\
HQIC(p,q) &= \ln(\tilde{\sigma}^2(p,q)) + \frac{2\ln(\ln T)}{T}(p+q)
\end{aligned}
$$

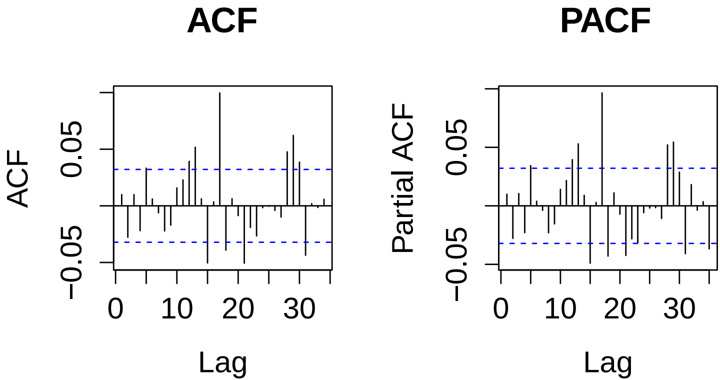
กรณี T ใหญ่ เกณฑ์ AIC จะเลือกอันดับที่มากกว่าที่ควรจะเป็น ในขณะที่เกณฑ์ BIC และ HQIC เลือกค่าที่ consistent

อย่างไรก็ตามใน T ขนาดเล็กเกณฑ์ทั้งสามจะให้ค่าที่ไม่แตกต่าง

ตัวอย่างที่ 2.7

log return on forex B/$ ปี 1998 ถึง 2012

Figure: ACF และ PACF ของผลได้ตอบแทนรายวันจากการซื้อขายเงินดอลลาร์สหรัฐ



ตัวอย่างที่ 2.7

อย่างไรก็ตาม package "forecast" มีคำสั่ง auto.arima(series, arguments) เลือก optimal p & q

```
> library(forecast)
> auto.arima(ret,d=0,D=0,max.p=6,max.q=6,ic=c("aic"),stepwise=FALSE,trace=TRUE)

 ARIMA(0,0,0) with zero mean     : -27687.05
 ARIMA(0,0,0) with non-zero mean : -27686.59
[omitted]
 ARIMA(5,0,0) with zero mean     : -27850.55
 ARIMA(5,0,0) with non-zero mean : -27851.11

Series: ret
ARIMA(4,0,1) with non-zero mean

Coefficients:
          ar1      ar2      ar3      ar4
      -0.5876  -0.0238  -0.0066  -0.0227
s.e.   0.2641   0.0196   0.0208   0.0199
         ma1  intercept
      0.5981     -1e-04
s.e.  0.2636      1e-04

sigma^2 estimated as 3.432e-05:  log likelihood=13849.44
AIC=-27684.88   AICc=-27684.85   BIC=-27641.33
```

การเขียนกระบวนการอารมา(ARMA)ในสามรูปแบบ

จากแบบจำลองที่เราอธิบาย y_t ด้วยแบบจำลอง *ARMA* ในรูปแบบพหุนาม

$$\phi(L)y_t = \theta(L)\varepsilon_t$$

โดยที่ $\phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i$ และ $\theta(L) = 1 + \sum_{i=1}^{q} \theta_i L^i$ เราสามารถแสดง y_t โดยใช้

$$\frac{\theta(L)}{\phi(L)} = 1 + \psi_1 L + \psi_2 L + ... = \psi(L) \qquad (2.63)$$

และ

$$\frac{\phi(L)}{\theta(L)} = 1 - \pi_1 L - \pi_2 L - ... = \pi(L) \qquad (2.64)$$

การพยากรณ์ (forecasting)

เราสามารถแสดงแบบจำลอง $ARMA(p, q)$ ได้สามรูปแบบ ซึ่งเราสามารถใช้รูปแบบใดรูปแบบหนึ่งในการสร้างการพยากรณ์

กรณีใช้รูปแบบ $ARMA$

$$\hat{y}_h(1) =$$

การพยากรณ์ (forecasting)

กรณีใช้รูปแบบออโตรีเกรสซีฟอนันต์

$\hat{y}_h(1)$

การพยากรณ์ (forecasting)

กรณีรูปแบบมูวิ่งเอเวอเรจอนันต์

$\hat{y}_h(1)$

Nonstationary series

อนุกรมเวลาทางเศรษฐกิจจำนวนมากมีพฤติกรรมที่มีแนวโน้ม(trend)หรือมีลักษณะไม่นิ่ง (non-stationary) ตัวอย่างเช่น ราคาสินทรัพย์ อัตราแลกเปลี่ยน GDP

รูปแบบแนวโน้มของข้อมูลว่าเกิดจากความไม่นิ่งหรือเป็นแนวโน้ม

การสร้างแบบจำลอง *ARMA* เราจำเป็นต้องทำการแปลงข้อมูลในนิ่งเสียก่อน ในขณะที่ถ้าข้อมูลมีแนวโน้มเราจำเป็นต้องขจัดแนวโน้มเสียก่อน

โดยที่กระบวนการที่มีลักษณะแนวโน้มอาจจะแบ่งได้เป็น (1) random walk with drift (2) trend stationary

Difference-stationary

กำหนดให้ y_t เป็น random walk with drift

$$y_t = \mu + y_{t-1} + \varepsilon_t \qquad (2.68)$$

โดย $\varepsilon_t \sim WN(0, \sigma^2)$ และเราเรียก μ ว่าตัวสร้างแนวโน้ม (drift) หากเราสมมุติให้ $y_0 = 0$ จะได้

กระบวนการแนวเดินแบบสุ่มแบบมีแนวโน้มจะมีค่าคาดหมาย ($E(y_t)$) เท่ากับ

ค่าความแปรปรวน ($Var(y_t)$) เท่ากับ

Trend- stationary

กำหนดให้ z_t เป็น trend-stationary

$$z_t = \beta_0 + \beta_1 t + v_t \qquad (2.70)$$

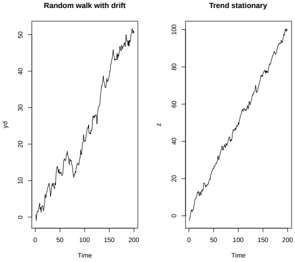
โดยที่ v_t เป็นอนุกรมนิ่งใดๆเช่น AR(1)

ค่าคาดหมายของ $(E(z_t))$ จะเท่ากับ

ค่าความแปรปรวน $(Var(z_t))$ จะเท่ากับ

Difference- and Trend- stationary

Figure: ข้อมูลจำลองของกระบวนการแนวเดินแบบสุ่มแบบมีแนวโน้มและกระบวนการที่นิ่งหลังจากขจัดแนวโน้ม



กระบวนการทั้งสองที่ได้ยกตัวอย่างมา เป็นกระบวนการที่ไม่นิ่ง (non-stationary) แต่มีคุณลักษณะที่แตกต่างกันดังนั้น วิธีการที่ใช้ในทำให้อนุกรมนิ่งจึงแตกต่างกัน

Difference-stationary

วิธีการการทำ first difference $\Delta y_t = y_t - y_{t-1}$
$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$ เราจะเรียกว่า second difference

วิธีการดำเนินการผลต่างจะเหมาะกับอนุกรมเวลาที่เป็นกระบวนการนิ่งหลังจากดำเนินการผลต่าง โดยที่หากอนุกรมเวลา y_t เป็น stationary หลังจากการทำ first difference เราจะเรียกอนุกรมเวลา y_t Integrated of order 1 หรือ $I(1)$

ในขณะที่อนุกรมเวลานิ่งจะคือ $I(0)$

วิธีการที่สองเหมาะกับอนุกรมที่เป็นกระบวนการนิ่งหลังจากกำจัดแนวโน้ม

การทดสอบยูนิทรูทในรูป AR(1)

เครื่องมือสำคัญที่ใช้ในการระบุว่าควรจะใช้วิธีการใดในการกำจัดแนวโน้มคือการทดสอบยูนิทรูท (unit root test)

สมมุติว่าเราพิจารณาแบบจำลอง $AR(1)$ ซึ่งเขียนได้ดังนี้

$$y_t = \phi y_{1-t} + \varepsilon_t, \qquad \varepsilon_t \sim WN(0, \sigma^2) \qquad (2.71)$$

การทดสอบยูนิทรูทมีพื้นฐานจากการทดสอบสมมุติฐานหลักว่าหาก $\phi = 1$ แล้ว y_t จะเป็น unit root และมีลักษณะเป็น random walk

ดังนั้นสมมุติฐานที่เราต้องการทดสอบคือ

$$\begin{aligned} H_0 &: \phi = 1 & \qquad (y_t \sim I(1)) \\ H_1 &: \phi < 1 & \qquad (y_t \sim I(0)) \end{aligned}$$

การทดสอบยูนิทรูทในรูป AR(1)

เราสามารถประมาณค่าสมการ (2.71) ด้วย OLS และใช้ตัวสถิติที่เรียกว่าตัวสถิติ Dickey-Fuller t ซึ่งคำนวณได้โดย

$$t_{\phi=1} = \frac{\hat{\phi} - 1}{se(\hat{\phi})}$$

การทดสอบนี้เป็นการทดสอบหางเดียวทางซ้าย คือเราจะปฏิเสธสมมุติฐานหลักถ้าค่าตัวสถิติน้อยกว่าค่าวิกฤติ และสรุปว่าอนุกรมเวลาเป็นอนุกรมนิ่ง หากเราไม่สามารถปฏิเสธสมมุติฐานหลักได้ แสดงว่าอนุกรมดังกล่าวเป็นยูนิทรูท

การทดสอบยูนิทรูทในรูป AR(1)

สมการ (2.71) ที่เราใช้ในการทดสอบยูนิทรูทยังสามารถเขียนในรูป

$$\Delta y_t = \pi y_{1-t} + \varepsilon_t, \qquad (2.72)$$

โดยที่ $\pi = (\phi - 1)$ ซึ่งในที่นี้การทดสอบยูนิทรูทจะเป็นการทดสอบ $H_0 : \pi = 0$ และ $H_1 : \pi < 0$

ตัวสถิติที่ใช้ในการทดสอบยูนิทรูทจะมีการแจกแจงที่ไม่เป็นมาตรฐานและไม่เป็นการแจกแจงแบบปกติเมื่อข้อมูลขนาดใหญ่เหมือนในกรณีทดสอบอื่นๆที่เราเคยพบมา Phillips (1987) ได้พิสูจน์ว่าหาก y_t เป็นยูนิทรูท $t_{\phi=1}$ และ $T(\hat{\phi} - 1)$ จะมีการแจกแจงแบบ Dickey-Fuller t และ Dickey-Fuller normalized bias

การแจกแจง Dickey-Fuller

ตารางสำหรับการแจกแจง Dickey-Fuller สามารถใช้คำสั่ง `adfTable` ใน `library(fUnitRoots)` โดยสมการทดสอบมีได้ 3 รูปแบบ

`trend=c("nc")` - no constant เช่นในสมการ (2.72) (ในหัวข้อย่อยต่อไปจะพิจารณากรณีมีพจน์ค่าคงที่ และจุดตัด) และแสดงการแจงแจงของ t โดยระบุ `statistics=c("t")` ส่วนการแจกแจง Dickey-Fuller normalized bias จะต้องระบุว่า `statistics=c("n")`

```
> library(fUnitRoots)
> adfTable(trend=c("nc"), statistic=c("t"))
$x
[1]   25   50  100  250  500  Inf

$y
[1] 0.010 0.025 0.050 0.100 0.900 0.950 0.975 0.990

$z
    0.010 0.025 0.050 0.100 0.900 0.950 0.975 0.990
 25 -2.66 -2.26 -1.95 -1.60  0.92  1.33  1.70  2.16
 50 -2.62 -2.25 -1.95 -1.61  0.91  1.31  1.66  2.08
100 -2.60 -2.24 -1.95 -1.61  0.90  1.29  1.64  2.03
250 -2.58 -2.23 -1.95 -1.62  0.89  1.29  1.63  2.01
500 -2.58 -2.23 -1.95 -1.62  0.89  1.28  1.62  2.00
Inf -2.58 -2.23 -1.95 -1.62  0.89  1.28  1.62  2.00
```

การทดสอบ unit root

สิ่งสำคัญในการทดสอบยูนิทรูทคือการระบุว่าสมมุติฐานหลักและสมมุติฐานทางเลือกมีความเหมาะสมกับข้อมูลหรือ โดยเฉพาะแนวโน้ม โดยกรณีที่เราพบบ่อยมีสองกรณีคือกรณีมีค่าคงที่อย่างเดียว และกรณีมีค่าคงที่และแนวโน้ม

ในกรณีค่าคงที่อย่างเดียว สมการที่ใช้ในการประมาณค่าตัวสถิติทดสอบคือ

$$y_t = c + \phi y_{t-1} + \varepsilon_t \tag{5}$$

และสมมุติฐานที่เกี่ยวเนื่องคือ

$$\begin{aligned} H_0 &: \phi = 1 & (y_t \sim I(1)\text{without drift}) \\ H_1 &: \phi < 1 & (y_t \sim I(0)\text{with non-zero mean}) \end{aligned}$$

การทดสอบ unit root

ในกรณีที่มีค่าคงที่และแนวโน้ม สมการที่ใช้ในการประมาณค่า

$$y_t = c + \delta t + \phi y_{t-1} + \varepsilon_t \tag{6}$$

และสมมุติฐานที่เกี่ยวเนื่องกันคือ

$$\begin{aligned} H_0 :& \phi = 1 & (y_t \sim I(1) \text{with drift}) \\ H_1 :& \phi < 1 & (y_t \sim I(0) \text{with deterministic time trend}) \end{aligned}$$

โดยที่รูปแบบของสมการที่ใช้ในการทดสอบมักจะเป็นในรูป

$$\Delta y_t = c + \pi y_{t-1} + \varepsilon_t$$

และ

$$\Delta y_t = c + \delta t + \pi y_{t-1} + \varepsilon_t$$

Augmented DF test

การทดสอบยูนิทรูทด้วยพจน์ที่เพิ่มเข้าไปโดยมีพื้นฐานจากแบบจำลอง $ARMA(p, q)$ แล้วเรียกว่า Augmented Dickey-Fuller ซึ่งตัวทดสอบสามารถประมาณค่าได้จากสมการ

$$y_t = \boldsymbol{\beta}^{'}(D_t) + \phi y_{t-1} + \sum_{j=1}^{p} \psi_j \Delta y_{t-j} + \varepsilon_t$$

หรือเราสามารถประมาณค่าสมการนี้

$$\Delta y_t = \beta^{'} D_t + \pi y_{t-1} + \sum_{j=1}^{p} \psi_j \Delta y_{t-j} + \varepsilon_t$$

โดยที่ $\pi = \phi - 1$ ซึ่งภายใต้สมมุติฐานหลัก Δy_t จะเป็น $I(0)$ และอนุมานได้ว่า $\pi = 0$

ตัวอย่าง 2.8

ทดสอบยูนิทรูทของ SET index รายวัน จากข้อมูลในไฟล์ `setindex.txt`

แต่ในที่นี้เราจะใช้ชุดคำสั่ง `fUnitRoots` โดยเบื้องต้นเราจะพิจารณา ACF ของอนุกรมเวลาและอนุกรมเวลาที่ได้ปรับผลต่างอันดับหนึ่ง (first-differenced) จาก ACF ของอนุกรมเวลา `set$index` เราพบว่าน่าจะมีปัญหายูนิทรูทเนื่องจากค่า ACF ลดลงช้ามาก

```
> set=read.table(file="setindex.txt", header=TRUE)
> library(fUnitRoots)
> acf(set$index)
> acf(diff(set$index))
> plot(diff(set$index),type="l")
```

เราจะทดสอบยูนิทรูทโดยใช้ Augmented Dickey-Fuller test โดยคำสั่งที่ใช้ในที่นี้ค่า `adfTest` โดยเราต้องระบุค่าล่าของ y_t ที่เราจะรวมในสมการ โดยในที่นี้จะเลือก `lags=10` และรูปแบบของสมการ `type` (ไม่มีค่าคงที่ (`nc`), มีค่าคงที่ (`c`) และมีค่าคงที่และเส้นแนวโน้ม (`ct`)) จากกราฟ Δy_t เราพบว่ามีค่าจุดตัดต่างจากศูนย์แต่ไม่มีแนวโน้มดังนั้นเราจะเลือก `type=c("c")`

```
> adfTest(set$index, lags=10, type=c("c"))

Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 10
  STATISTIC:
    Dickey-Fuller: -0.0682
  P VALUE:
    0.9501
```

ARIMA

Definition 2

เราจะเรียกกระบวนการ y_t ว่า $ARIMA(p,d,q)$ ถ้า $\Delta^d y_t = (1-L)^d y_t$ เป็นกระบวนการ $ARMA(p,d,q)$ หรือเราสามารถเขียนในรูปทั่วไปเป็น

$$\phi(L)(1-L)^d y_t = \theta(L)\varepsilon_t \tag{7}$$

และถ้าหาก $E(\Delta^d y_t) = \mu$ เราสามารถเขียนแบบจำลองได้เป็น

$$\phi(L)(1-L)^d y_t = \delta + \theta(L)\varepsilon_t$$

โดยที่ $\delta = \mu(1-\phi_1 - \ldots - \phi_p)$

ARIMA

ในทางปฏิบัติ เราสามารถแบ่งแยกขั้นตอนในการสร้างแบบจำลอง $ARIMA(p,d,q)$ ได้ดังนี้

1. พิจารณา ACF ของข้อมูลว่ามีข้อบ่งชี้ว่ามีปัญหาข้อมูลไม่นิ่ง (nonstationary) หรือไม่
2. ทดสอบยูนิทรูทว่ามีอินทิเกรตที่อันดับเท่าใด หรือหาค่า d นั่นเอง
3. หลังจากนั้นเราจะพิจารณาอนุกรม $z_t = \Delta^d y_t$ ว่าแบบจำลองที่เหมาะสมในการอธิบาย z_t ควรจะมีอันดับของ p และ q เท่าใด

หลังจากที่เราได้ค่า d, p และ q แล้วเราก็จะได้แบบจำลอง $ARIMA(p,d,q)$

ตัวอย่างที่ 2.9

ต่อจากตัวอย่างที่ 2.8 เราพิจารณาว่า set index

```
> adfTest(diff(set$index), lags=10, type=c("c"))
Title:
 Augmented Dickey-Fuller Test
Test Results:
  PARAMETER:
    Lag Order: 10
  STATISTIC:
    Dickey-Fuller: -16.7697
  P VALUE:
    0.01
> library(forecast)
> auto.arima(set$index
      ,d=1,D=0,max.p=12,max.q=12,ic=c("bic"),stepwise=FALSE,trace=FALSE)
Series: set$index
ARIMA(0,1,0)
sigma^2 estimated as 87.77:  log likelihood=-13312.56
AIC=26627.12   AICc=26627.12   BIC=26633.32
```

ซึ่งจากการใช้ BIC พบว่าอันดับที่เหมาะสมคือ $p = 0$ และ $q = 0$ ดังนั้นแบบจำลองที่เหมาะสมที่ใช้ประมาณค่า set index รายวันคือ `ARIMA(0,1,0)` หรือแบบจำลอง random walk นั่นเอง