เศรษฐมิติทางการเงินเบื้องต้น
Introduction to Financial Econometrics

เฉลิมพงษ์ คงเจริญ

คณะเศรษฐศาสตร์ มหาวิทยาลัยธรรมศาสตร์

http://econ.tu.ac.th/archan/Chaleampong/

ร่างแรก พฤศจิกายน 2555 ปรับปรุงครั้งล่าสุด 13 สิงหาคม 2562

สารบัญ

บทที่ 0

บทนำ: ความหมายของเศรษฐมิติทางการเงิน

เศรษฐมิติทางการเงิน (financial econometrics) คือการใช้วิธีการทางเศรษฐมิติในการตอบปัญหาทางการเงิน โดยที่เศรษฐมิติทางการเงินจะเป็นประโยชน์ในการทดสอบทฤษฎีทางการเงิน การกำหนดราคาของสินทรัพย์และผลตอบแทน การทดสอบความสัมพันธ์ระหว่างตัวแปรทางการเงินด้วยกัน ความสัมพันธ์ระหว่างตลาดการเงินและตัวแปรเศรษฐกิจมหภาค และการทำนายค่าในอนาคตของตัวแปรทางการเงิน เป็นต้น

ตัวอย่างคำถามทางการเงินที่สามารถนำเครื่องมือเศรษฐมิติไปใช้[1]

- ทดสอบความตลาดการเงินมีประสิทธิภาพหรือไม่ (weak-form efficient)
- ทดสอบว่าแบบจำลอง Capital Asset Pricing Model (CAPM) หรือ Arbitrage Pricing Theory (APT) เป็นแบบจำลองที่ดีกว่ากันในการอธิบายผลตอบแทนของสินทรัพย์เสี่ยง
- วัดและคาดการณ์ความผันผวน (volatility) ของผลตอบแทนในพันธบัตร
- อธิบายตัวกำหนด credit rating
- สร้างแบบจำลองอธิบายความสัมพันธ์ระยะยาวระหว่างราคาและอัตราแลกเปลี่ยน
- กำหนด optimal hedge ratio สำหรับ spot position
- ทดสอบ technical trading rules เพื่อหาวิธีที่จะทำกำไรได้มากที่สุด
- ทดสอบสมมุติฐานว่าการประกาศผลตอบแทนหรือเงินปันผลมีผลต่อราคาหุ้นหรือไม่
- การทดสอบว่าตลาด spot หรือ future ตอบสนองต่อข่าวต่างกันอย่างไร
- การทำนายความสหสัมพันธ์ระหว่างตลาดหลักทรัพย์ในสองประเทศ

[1] Brooks (2007, 2)

ความแตกต่างระหว่างเศรษฐมิติ (economic econometrics) กับเศรษฐมิติทางการเงิน

เครื่องมือพื้นฐานที่เราใช้ในวิชาเศรษฐมิติทางการเงินจะเหมือนกับในวิชาเศรษฐมิติโดยทั่วไป อย่างไรก็ตาม ข้อมูลนตลาดการเงินที่เรานำมาพิจารณาจะมีความแตกต่างในแง่ของ **ความถี่ ความถูกต้อง ฤดูกาล** กับข้อมูลที่เราพบในวิชาเศรษฐมิติทั่วไป ในวิชาเศรษฐศาสตร์เรามักพบปัญหาที่ข้อมูลไม่เพียงพอและสนใจปัญหาที่เกิดจากข้อมูลจำนวนน้อย

ในตลาดการเงินข้อมูลที่เราพบจะมีรูปแบบที่ค่อนข้างต่างกัน แต่โดยทั่วไปที่เราสนใจคือข้อมูลของราคา ณ ระดับที่มีการซื้อขาย ดังนั้นจะมีความถี่ที่สูงกว่าข้อมูลเศรษฐกิจมหภาค ราคาของสินทรัพย์ใดๆมักมีตั้งแต่ รายนาที ชั่วโมงหรือวัน ดังนั้นกลุ่มตัวอย่างในกรณีข้อมูลการเงินมักจะเป็นขนาดใหญ่

ประเภทของข้อมูล

ข้อมูลภาคตัดขวาง (cross sectional data) เป็นข้อมูลที่มีการจัดเก็บ ณ เวลาใดๆ เช่น ราคาหลักทรัพย์ของทั้งตลาด ณ ปีใดปีหนึ่ง, กำไรสุทธิของบริษัทในตลาดหลักทรัพย์ ณ ปีใดปีหนึ่ง ปัญหาที่สามารถใช้ข้อมูลภาคตัดขวางในการศึกษา ตัวอย่างเช่น ความสัมพันธ์ระหว่างขนาดของบริษัทกับผลตอบแทนในการลงทุนในหุ้นนั้นๆ, ความสัมพันธ์ระหว่างระดับของ GDP ของประเทศกับโอกาสในการผิดนัดชำระหนี้ของประเทศใดๆในปีใดปีหนึ่ง

ข้อมูลอนุกรมเวลา (time series) เป็นข้อมูลที่มีการจัดเก็บตามเวลา โดยจะมีความสัมพันธ์กับความถี่ในการจัดเก็บ อย่างไรก็ตามข้อมูลทางการเงินบางส่วนมีการบันทึกตามธุรกรรมที่เกิดเช่นการซื้อขายหุ้น แบบจำลองเดียวกันจำเป็นที่จะต้องมีความถี่ที่เหมือนกัน ตารางที่ **??** เป็นตัวอย่างของข้อมูลอนุกรมเวลา

ตารางที่ 1: ตัวอย่างข้อมูลอนุกรมเวลา

สัปดาห์	ปริมาณซื้อขาย	ราคาปิด
7/16/2012	6301500	191.63
7/23/2012	3759100	195.56
7/30/2012	2988800	197.68
8/6/2012	2475200	199.29
8/13/2012	2474700	201.22
8/20/2012	2828000	197.77
8/27/2012	2622700	194.85

ปัญหาที่สามารถใช้ข้อมูลอนุกรมเวลาในการศึกษา ตัวอย่างเช่น ดัชนีราคาหลักทรัพย์แปรผันกับตัวแปรทางเศรษฐกิจของประเทศอย่างไร, มูลค่าของหลักทรัพย์แปรผันตามการประกาศเงินปันผลอย่างไร, อัตราแลกเปลี่ยนมีผลต่อการขาดดุลการค้าอย่างไร

ข้อมูลแพนอล (panel) เป็นข้อมูลที่มีทั้งอนุกรมเวลาและภาคตัดขวางเช่น ราคารายวันของ

หุ้นใน SET50 ตลอดระยะเวลาสองปี

ในตำราเล่มนี้เราจะเน้นการเนื้อหาไปยังการศึกษาโดยใช้ข้อมูลอนุกรมเวลา**ข้อมูลอนุกรมเวลา**

เป้าหมายของตำรานี้

วัตถุประสงค์หลักของตำรานี้คือต้องการให้ผู้อ่านมีความรู้เบื้องต้นในการวิเคราะห์ข้อมูลอนุกรมเวลาทางการเงิน เช่น ความเบ้ ความโด่ง ตลอดจนอัตสหสัมพันธ์(autocorrelation)ของตัวแปรผลได้ตอบแทน นอกจากนี้ยังนำเสนอแบบเครื่องมือทางสถิติและแบบจำลองเศรษฐมิติที่เป็นประโยชน์ในการวิเคราะห์ข้อมูลอนุกรมเวลา และคาดหวังให้ผู้อ่านมีความรู้ความเข้าใจในแบบจำลองเศรษฐมิติทางการเงิน ตลอดจนสามารถนำเครื่องมือดังกล่าวไปประยุกต์ใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลาทางการเงิน

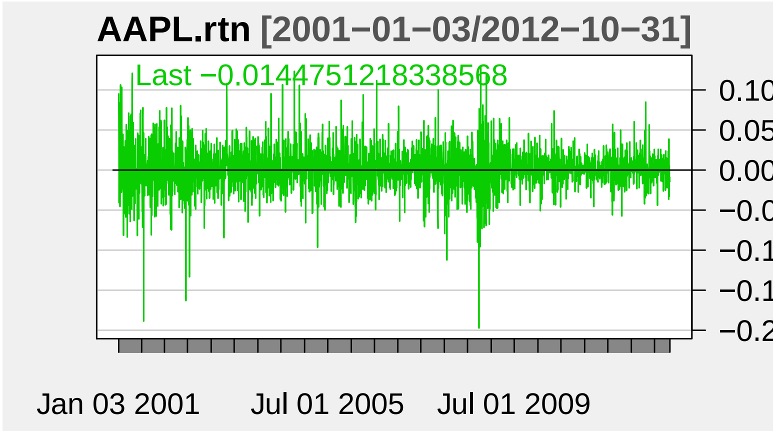
ตัวอย่างอนุกรมเวลาทางการเงิน

ตัวอย่างของข้อมูลอนุกรมเวลาถูกนำมาแสดงในหัวข้อนี้ รูปที่ **??** แสดงราคาปิดของหุ้น APPLE ระหว่างปี 2001 และ 2012 ซึ่งข้อมูลดังกล่าวนำมาคำนวณผลได้ตอบแทนรายวันของหุ้นและแสดงอยู่ในรูปที่ **??**

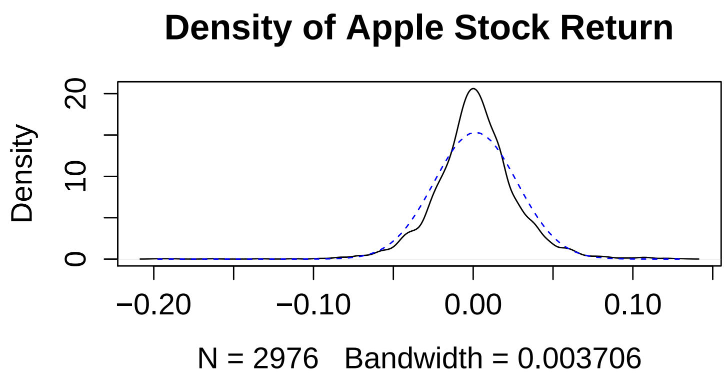
รูปที่ 1: ราคาปิดหุ้น APPLE co. รายวัน

จากรูปที่ **??** จะเห็นได้ว่าผลได้ตอบแทนมีค่าอยู่รอบๆศูนย์ โดยที่บางช่วงจะมีความผันผวนสูงกว่าบางช่วง (มีผลได้ตอบแทนอยู่ห่างจากศูนย์มากกว่าปกติ) รูปที่ **??** แสดงการแจกแจงผลได้ตอบแทนของหุ้น APPLE โดที่เส้นทึบแสดงการแจกแจงจากข้อมูลจริง ในขณะที่เส้นประแสดงการแจกแจงแบบปกติ

รูปที่ 2: ผลได้ตอบแทนรายวันของหุ้น APPLE co.

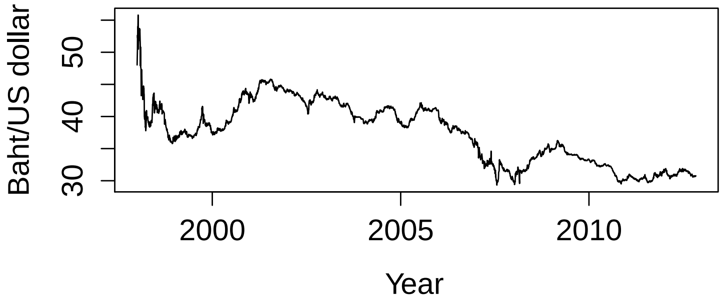


รูปที่ 3: การแจกแจงผลได้ตอบแทนรายวันของหุ้น APPLE co.

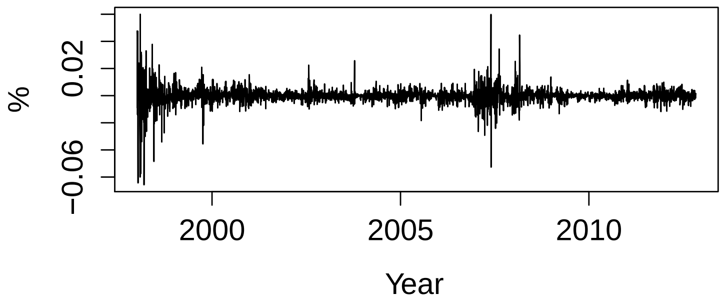


รูปที่ **??** แสดงอัตราแลกเปลี่ยนเงินบาทและดอลลาร์สหรัฐรายวันระหว่างปี 1997 ถึง 2011 เมื่อพิจารณาผลได้ตอบแทนในรูปที่ **??** จะเห็นได้ว่าช่วงหลังวิกฤติการณ์เงินตราปี 1997 และช่วงวิกฤติแฮมเบอร์เกอร์ ผลได้ตอบแทนมีความผันผวนมากกว่าช่วงอื่นๆ โดยที่แบบจำลองอธิบายความผันผวนจะถูกพิจารณาในบทที่ 3

รูปที่ 4: อัตราแลกเปลี่ยนบาทต่อดอลลาร์สหรัฐรายวัน

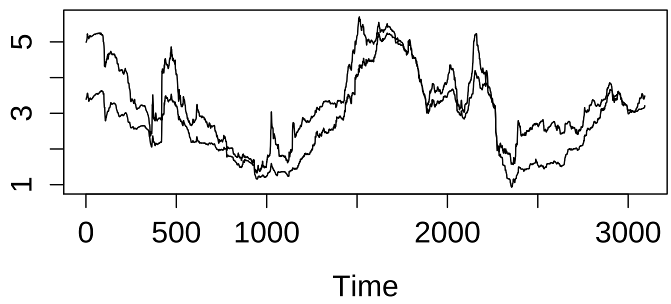


รูปที่ 5: อัตราแลกเปลี่ยนบาทต่อดอลลาร์สหรัฐรายวัน



รูปที่ **??** แสดงอัตราดอกเบี้ยพันธบัตรรัฐบาลอายุไถ่ถอน 1 ปีและ 3 ปี พบว่าข้อมูลทั้งสองมีความเคลื่อนไหวไปในทิศทางเดียวกัน ซึ่งความสัมพันธ์ระหว่างอนุกรมเวลามากกว่า 1 อนุกรมเวลาจะถูกพิจารณาในบทที่ 5 และ 6

รูปที่ 6: อัตราดอกเบี้ยพันธบัตรรัฐบาล



บทที่ 1

อนุกรมเวลาทางการเงินและคุณลักษณะ

1.1 การคำนวณผลตอบแทน (return)

การศึกษาปริมาณทางการเงินส่วนใหญ่เรามักจะสนใจผลได้ตอบแทน(return)ของทรัพย์สินมากกว่าราคา (price) เนื่องจากผลได้ตอบแทนของทรัพย์สินมีความสมบูรณ์ (complete) และปราศจากผลของหน่วยวัด (Campbell, Lo and Mackinley (1997)) คุณสมบัติดังกล่าวใช้ในการประเมินโอกาสในการลงทุน เช่นการระบุว่าสินทรัพย์ที่เราถืออยู่ราคาเพิ่มขึ้น 10 บาทมิได้แสดงให้เห็นว่าผลได้ตอบแทนนั้นดีหรือไม่ จำเป็นต้องระบุในรูปของผลได้ตอบแทนเป็นร้อยละ นอกจากนี้อนุกรม (series) ของผลได้ตอบแทนมีคุณสมบัติทางสถิติที่ดีกว่าอนุกรมของราคาในการดำเนินการทางสถิติซึ่งเราจะได้พิจารณาในเนื้อหาบทต่อไป

1.1.1 นิยามของผลได้ตอบแทนของทรัพย์สิน

สมมุติให้การซื้อสินทรัพย์ เช่นหุ้น พันธบัตร หรือกองทุนรวม ณ เวลา t_0 ด้วยราคา P_{t_0} บาท และขายสินทรัพย์ ณ เวลา t_1 ด้วยราคา P_{t_1} บาท การเปลี่ยนแปลงของราคาสินทรัพย์ในรูปร้อยละสามารถคำนวณได้โดย

$$R(t_0, t_1) = \frac{P_{t_1} - P_{t_0}}{P_{t_0}} \tag{1.1}$$

เราเรียกระยะเวลาระหว่าง t_0 และ t_1 ว่า*ระยะเวลาการถือสินทรัพย์ (holding period)* และสมการ (**??**) ว่า*ผลตอบแทนจากระยะเวลาการถือสินทรัพย์* โดยทั่วไปเราอาจจะกำหนดระยะเวลาการถือสินทรัพย์เป็นระยะเวลาใดๆก็ได้ เช่น ราย 15 นาที รายครึ่งเดือน อย่างไรในการศึกษาวิชานี้เราจะสมมุติให้ระยะเวลาการถือมีลักษณะเพิ่มขึ้นเช่นเดียวกับปฏิทิน เช่น รายนาที รายวัน รายเดือน หรือรายปี และในหัวข้อนี้เราจะสมมุติระยะเวลาการพิจารณาเป็นรายเดือน

กำหนดให้ P_t เป็นราคาของสินทรัพย์ ณ เวลาสิ้นเดือน t และไม่มีการจ่ายเงินปันผล และกำหนดให้ P_{t-1} เป็นราคาของสินทรัพย์ ณ เวลาสิ้นเดือน $t-1$ แล้ว **ผลได้ตอบแทนรวมอย่างง่าย**

หนึ่งเดือน (one-month simple gross return) สามารถคำนวณได้ดังนี้

$$1 + R_t = \frac{P_t}{P_{t-1}} \tag{1.2}$$

และ**ผลได้ตอบแทนสุทธิอย่างง่ายหนึ่งเดือน (one-month simple net return)** หรือ ผลได้ตอบแทนอย่างง่ายหนึ่งเดือน (one-month simple return)

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{1.3}$$

โดยที่ผลได้ตอบแทนอย่างง่ายหนึ่งเดือนอธิบายว่าหากเราลงทุนด้วยเงิน 1 บาทในเดือนที่ $t-1$ เราจะได้รับเงินคืนมาทั้งหมดเท่าไหร่ หรืออาจมองในแง่ของมูลค่าอนาคตของเงิน 1 บาท นอกจากนี้เราจะสังเกตเห็นได้ว่าราคาของสินทรัพย์ใดๆจะต้องเป็นค่าที่ไม่เป็นลบดังนั้น ค่าต่ำสุดของ R_t คือ -1 หรือขาดทุน 100%

Example 1

ตัวอย่างที่ 1.1 การคำนวณผลได้ตอบแทนอย่างง่าย

สมมุติว่าเราพิจารณาผลตอบแทนจากการลงทุนในหุ้น APPLE โดยที่เราซื้อหุ้นเมื่อเดือน $t-1$ ด้วยราคา $P_{t-1} = 190$ ดอลลาร์สหรัฐและขายไปในเดือนต่อไปด้วยราคา $P_t = 200$ ดอลลาร์สหรัฐและไม่มีการจ่ายเงินปันผลในระหว่างที่ถือหุ้น ดังนั้นผลได้ตอบแทนสุทธิเท่ากับ

$$R_t = \frac{200-190}{190} = \frac{200}{190} - 1 = 1.0526 - 1 = 0.0526$$

และผลได้ตอบแทนอย่างง่ายรวมหนึ่งเดือนจะเท่ากับ

$$1 + R_t = \frac{200}{190} = 1.0526$$

ผลได้ตอบแทนสุทธิหนึ่งเดือนจะเท่ากับ 5.25 % ต่อเดือน หรือการลงทุนในหุ้น APPLE 1 ดอลลาร์สหรัฐจะได้เงินคืนพร้อมเงินต้น 1.0525 ดอลลาร์สหรัฐ

ผลได้ตอบแทนหลายเดือน

ผลได้ตอบแทนจากการลงทุนสองเดือนสามารถคำนวณได้จากการเปลี่ยนแปลงของราคา ณ เดือน P_t และ P_{t-2} หรือผลได้ตอบแทนอย่างง่ายสองเดือน, $R_t(2)$, จะเท่ากับ

$$\begin{aligned} R_t(2) &= \frac{P_t - P_{t-2}}{P_{t-2}} = \frac{P_t}{P_{t-2}} - \frac{P_{t-2}}{P_{t-2}} \\ &= \frac{P_t}{P_{t-2}} - 1 \\ &= \frac{P_t}{P_{t-1}} \frac{P_{t-1}}{P_{t-2}} - 1 \\ &= (1 + R_t)(1 + R_{t-1}) - 1 \end{aligned} \tag{1.4}$$

และผลได้ตอบแทนรวมอย่างง่ายสองเดือนจะเท่ากับ

$$1 + R_t(2) = (1 + R_t)(1 + R_{t-1}) = 1 + R_t + R_{t-1} + R_t R_{t-1} \tag{1.5}$$

ซึ่งคือการรวมเรขาคณิต (geometric sum) ของผลได้ตอบแทนรวมอย่างง่ายหนึ่งเดือนของเดือน t และ $t - 1$ ดังนั้นจะเห็นได้ว่า $R_t(2)$ จะไม่เท่ากับผลรวมของ R_t และ R_{t-1}

ตัวอย่างที่ 1.2 การคำนวณผลได้ตอบแทนอย่างง่ายสองเดือน

พิจารณาต่อเนื่องจากตัวอย่าง 1.1 สมมุติให้เราซื้อหุ้น ณ เดือนที่ $t-2$ ด้วยราคา $P_{t-2} = 180$ ดอลลาร์และไม่มีการจ่ายเงินปันผล ผลได้ตอบแทนสุทธิสองเดือนจะเท่ากับ

$$R_t(2) = \frac{200 - 180}{180} = \frac{200}{180} - 1 = 0.1111$$

หรือ 11.11 % โดยที่ผลได้ตอบแทนหนึ่งเดือนของแต่ละเดือนจะเท่ากับ

$$\begin{aligned} 1 + R_{t-1} &= \frac{190 - 180}{180} = 1.0556 - 1 = 0.0556 \\ 1 + R_t &= \frac{200 - 190}{190} = 1.0526 - 1 = 0.0526 \end{aligned}$$

และผลได้ตอบแทนรวมสองเดือนจะเท่ากับ

$$1 + R_t(2) = 1.0556 \times 1.0526 = 1.1111$$

เราสามารถเขียนผลได้ตอบแทนในรูปทั่วไปได้ดังนี้ ผลได้ตอบแทนรวมอย่างง่าย k เดือน

(k-month simple gross return):

$$1 + R_t(k) = \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}}$$
$$= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}) \tag{1.6}$$

ซึ่งคือการรวมเรขาคณิตของผลได้ตอบแทนรวมอย่างง่ายรายเดือนของเดือนที่ t ถึงเดือนที่ $t-k+1$

ผลได้ตอบแทนของกลุ่มสินทรัพย์ลงทุน (Portfolio)

หากเราลงทุนด้วยเงินจำนวน V บาทในสินทรัพย์สองตัวคือ A และ B โดยสัดส่วนการลงทุนในสินทรัพย์ทั้งสองคือ x_A และ x_B แล้วมูลค่าการลงทุนในสินทรัพย์แต่ละชนิดจะเท่ากับ Vx_A และ Vx_B โดยเราสมมุติให้ $x_A + x_B = 1$ หากกำหนดให้ผลได้ตอบแทนอย่างง่ายหนึ่งเดือนของ A และ B คือ $R_{A,t}$ และ $R_{B,t}$ แล้ว มูลค่าของกลุ่มสินทรัพย์ลงทุนเมื่อสิ้นเดือนจะเท่ากับ

$$V \times [x_A(1 + R_{A,t}) + x_B(1 + R_{B,t})]$$

โดย $x_A(1 + R_{A,t}) + x_B(1 + R_{B,t})$ คือผลได้ตอบแทนรวมของกลุ่มสินทรัพย์ลงทุน และอัตราผลตอบแทนจากการลงทุนจะเท่ากับ $x_A R_{A,t} + x_B R_{B,t}$

การปรับกรณีเงินปันผล

ถ้าสินทรัพย์มีการจ่ายเงินปันผลเท่ากับ D_t ในช่วงเวลาระหว่างเดือน t และ $t-1$ แล้วการคำนวณผลได้ตอบแทนสุทธิทั้งหมด (total net return) สามารถคำนวณได้ด้วยสมการต่อไปนี้

$$R_t^{total} = \frac{P_t + D_t - P_{t-1}}{P_{t-1}} = \frac{P_t - P_{t-1}}{P_{t-1}} + \frac{D_t}{P_{t-1}} \tag{1.7}$$

โดยที่ส่วนแรกเป็นกำไรส่วนเกินทุน (capital gain) และส่วนที่สองคืออัตราผลตอบแทนเงินปันผล (dividend yield)

การแปลงผลได้ตอบแทนเป็นผลได้ตอบแทนรายปี

โดยทั่วไปเรามักจะรายงานผลได้ตอบแทนเป็นรายปีเพื่อใช้ในการตัดสินใจลงทุน ดังนั้นเราจำเป็นต้องมีการแปลงผลได้ตอบแทนที่มีความถี่อื่นๆเป็นรายปี สมมุติว่าเราต้องการหาผลได้ตอบแทนหนึ่งปี $(1 + R_A)$ จากข้อมูลผลได้ตอบแทนรายเดือน $(R_t, R_{t-1}, ..., R_{t-11})$ เราสามารถคำนวณได้โดย

$$1 + R_A = 1 + R_t(12) = \frac{P_t}{P_{t-12}} = \frac{P_t}{P_{t-1}} \cdot \frac{P_{t-1}}{P_{t-2}} \cdots \frac{P_{t-11}}{P_{t-12}}$$
$$= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-11}) \tag{1.8}$$

นอกจากนี้หากเราทราบผลได้ตอบแทนรายหนึ่งเดือน และต้องการประมาณการผลได้ตอบแทนรายปีภายใต้ข้อสมมุติว่าผลได้ตอบแทนคงที่ทุกเดือนเท่ากับ R เราจะได้ผลได้ตอบแทนรายหนึ่งปีเท่ากับ

$$1 + R_A = 1 + R_t(12) = (1 + R)^{12}$$

ซึ่งก็คือผลได้ตอบแทนทบต้น 12 เดือนนั่นเอง

1.1.2 ผลได้ตอบแทนทบต้นอย่างต่อเนื่อง

ในหัวข้อนี้เราจะพิจารณาผลของการคิดผลตอบแทนทบต้น(compounding) สมมุติว่าธนาคารจ่ายดอกเบี้ยเงินฝากในอัตรา 10 % ต่อไปี และมีเงินต้นเท่ากับ 100 บาท หากธนาคารจ่ายดอกเบี้ยหนึ่งครั้งเมื่อสิ้นปี มูลค่าสุทธิของเงินฝากดังกล่าวตอนปลายปีจะเท่ากับ $100(1 + 0.1) = 110$ บาท หากธนาคารแบ่งการจ่ายดอกเบี้ยออกเป็นสองครั้งครั้งละ 5 % ทุกครึ่งปี จะพบว่ามูลค่าสุทธิจะเท่ากับ $100(1 + 0.1/2)^2 = 110.25$ บาทหลังจากสิ้นปี จากการพิจารณารูปแบบข้างต้น หากธนาคารจ่ายดอกเบี้ย m ครั้ง มูลค่าสุทธิ ณ สิ้นปีที่หนึ่งจะมีค่า $100(1 + 0.1/m)^m$ และถ้าธนาคารจ่ายดอกเบี้ยทบต้นอย่างต่อเนื่อง (continuous compounding) มูลค่าสุทธิ ณ สิ้นปีที่หนึ่งจะเท่ากับ $100(\exp(0.1)) = 100.517$

ผลได้ตอบแทนหนึ่งเดือน

กำหนดให้ R_t เป็นผลได้ตอบแทนอย่างง่ายรายเดือนของการลงทุน เนื่องจากมูลค่าในอนาคตจะเท่ากับมูลค่าปัจจุบันคูณกับตัวคูณทบต้น ($P_t = P_{t-1}\exp(r_t)$) เราสามารถคำนวณผลได้ตอบแทนทบต้นต่อเนื่องหนึ่งเดือน (one-month continuously compounding return) หรือผลได้ตอบแทนในรูปลอการิทึม(one-month log return) ซึ่งใช้สัญลักษณ์แทน r_t ได้โดย

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(P_t) - \ln(P_{t-1}) = p_t - p_{t-1} \tag{1.9}$$

โดยที่ $p_t = \ln(P_t)$

ตัวอย่างที่ 1.3 การคำนวณผลได้ตอบแทนในรูปล็อก

ใช้ข้อมูลจากตัวอย่างที่ 1.1 ผลได้ตอบแทนทบต้นอย่างต่อเนื่องหนึ่งเดือนหรือผลได้ตอบแทนในรูปลอการิทึมหนึ่งเดือนจะเท่ากับ

$$r_t = \ln(1.0526) = 0.0513$$
$$r_t = \ln(200) - \ln(190) = 5.2983 - 5.2470 = 0.0513$$

เนื่องจาก $1 + R_t = \frac{P_t}{P_{t-1}}$ เราสามารถเขียนความสัมพันธ์ระหว่างผลได้ตอบแทนอย่างง่ายและผลได้ตอบแทนในรูปลอการิทึมได้เป็น

$$r_t = \ln(1 + R_t)$$

และเราสามารถคำนวณผลตอบแทนอย่างง่ายสุทธิได้จากผลได้ตอบแทนในรูปลอการิทึมโดย

$$R_t = \exp(r_t) - 1 \tag{1.10}$$

ตัวอย่างที่ 1.4 ความสัมพันธ์ระหว่างผลได้ตอบแทน

ใช้ข้อมูลจากตัวอย่างที่ผ่านมา ผลได้ตอบแทนในรูปลอการิทึมหนึ่งเดือนเท่ากับ 5.13 % ดังนั้น

$$R_t = \exp(0.0513) - 1 = 0.0526$$

หากผลได้ตอบแทนมีค่าต่ำ เช่น ผลได้ตอบแทนรายชั่วโมงหรือรายวัน ผลได้ตอบแทนในรูปลอการิทึมและผลได้ตอบแทนอย่างง่ายจะมีค่าใกล้เคียงกัน อย่างไรก็ตาม การคำนวณผลได้ตอบแทนทั้งสองวิธีมีคุณสมบัติที่แตกต่างกัน ดังนี้ หากราคาของสินทรัพย์ลดลงเหลือศูนย์ผลได้ตอบแทนอย่างง่ายจะเท่ากับ -1 ซึ่งเป็นขอบเขตต่ำสุดของผลได้ตอบแทนอย่างง่าย ในขณะที่ขอบเขตล่างของผลได้ตอบแทนในรูปลอการิทึมจะเท่ากับ $-\infty$ ($r_t = \ln(1 + R_t) = \ln(1 - 1) = \ln(0)$) ดังนั้นหากเราคำนวณผลได้ตอบแทนในรูปลอการิทึมมีค่าติดลบมากๆ จึงไม่ได้หมายความว่าเราสูญเสียเงินมากกว่าเงินต้น หากเราต้องการทราบอัตราผลขาดทุนเราจำเป็นต้องใช้สมการ (**??**) ในการคำนวณกลับไปเป็นผลได้ตอบแทนอย่างง่าย

เนื่องจากผลได้ตอบแทนในรูปลอการิทึมมีคุณสมบัติในการรวมกันที่ง่ายกว่าผลได้ตอบแทนอย่างง่าย ดังนั้นการวิเคราะห์เชิงสถิติหรือการสร้างแบบจำลองเรามักจะใช้ผลได้ตอบแทนในรูปลอการิทึม

ผลได้ตอบแทนหลายเดือน

หากเราพิจารณาผลได้ตอบแทนในรูปลอการิทึมสองเดือนจะเท่ากับ

$$r_t(2) = \ln(1 + R_t(2)) = \ln\left(\frac{P_t}{P_{t-2}}\right) = p_t - p_{t-2}$$

นอกจากนี้เราสามารถหาความสัมพันธ์ระหว่างผลได้ตอบแทนในรูปลอการิทึมสองเดือนกับผลได้ตอบแทนในรูปลอการิทึมหนึ่งเดือนได้โดย

$$\begin{aligned} r_t(2) &= \ln\left(\frac{P_t}{P_{t-2}}\right) \\ &= \ln\left(\frac{P_t}{P_{t-1}}\frac{P_{t-1}}{P_{t-2}}\right) \\ &= \ln\left(\frac{P_t}{P_{t-1}}\right) + \ln\left(\frac{P_{t-1}}{P_{t-2}}\right) \\ &= r_t + r_{t-1} \end{aligned}$$

ดังนั้นผลได้ตอบแทนในรูปลอการิทึมสองเดือนก็จะคือผลรวมของผลได้ตอบแทนในรูปลอการิทึมรายเดือนของสองเดือนมารวมกันนั่นเอง ซึ่งแตกต่างจากในกรณีของผลได้ตอบแทนอย่างง่ายที่ต้องใช้ผลรวมเรขาคณิต

ตัวอย่างที่ 1.5 การคำนวณผลได้ตอบแทนหลายเดือน

จากข้อมูลในตัวอย่างที่ 1.2 เราสามารถคำนวณผลได้ตอบแทนในรูปลอการิทึมสองเดือนได้สองวิธีคือ

$$r_t(2) = \ln(200) - \ln(180) = 5.2983 - 5.1930 = 0.1053$$

หรือรวมผลได้ตอบแทนในรูปลอการิทึมหนึ่งเดือนของแต่ละเดือนเข้าด้วยกัน โดย $r_t = \ln(200) - \ln(190) = 0.0513$ และ $r_{t-1} = \ln(190) - \ln(180) = 0.0540$ ดังนั้น

$$r_t(2) = 0.0513 + 0.0540 = 0.1053$$

ในกรณีทั่วไป ผลได้ตอบแทนในรูปลอการิทึม k เดือนสามารถคำนวณได้จาก

$$\begin{aligned} r_t(k) = \ln(1 + R_t(k)) = \ln\left(\frac{P_t}{P_{t-k}}\right) &= p_t - p_{t-k} \\ = \sum_{j=0}^{k-1} r_{t-j} \end{aligned}$$

ในขณะที่ผลได้ตอบแทนในรูปลอการิทึมของกลุ่มหลักทรัพย์จะเท่ากับ

$$r_{p,t} = \ln(1 + R_{p,t}) = \ln\left(1 + \sum_{i=1}^{n} x_i R_{i,t}\right) \neq \sum_{i=1}^{n} x_i r_{i,t} \tag{1.11}$$

โดยผลได้ตอบแทนในรูปลอการิทึมของกลุ่มหลักทรัพย์จะไม่เท่ากับค่าเฉลี่ยถ่วงน้ำหนักของผลได้ตอบแทนในรูปลอการิทึมของสินทรัพย์

การคำนวณผลได้ตอบแทนด้วยโปรแกรม R

เลือกโฟลเดอร์ที่จะใช้เก็บข้อมูล เราเรียกว่า work directory (ในกรณีนี้คือ D:/Teaching/EC435/R/) พร้อมกับคัดลอก (copy) ไฟล์ aapl-m.csv ซึ่งเป็นข้อมูลราคาของหุ้น APPLE รายเดือน ตั้งแต่ มกราคม 2010 ถึงตุลาคม 2012 จาก Moodle ไปยังโฟลเดอร์ดังกล่าว หลังจากนั้นเปิดโปรแกรม RStudio ขึ้นมาพร้อมนำเข้าข้อมูลตามคำสั่งต่อไปนี้

```
> setwd("D:/Teaching/EC435/R/")
> aapl<-read.csv("aapl-m.csv", header=T)
> head(aapl)
    Date  Price
1 4/1/10 190.37
2 1/2/10 202.82
3 1/3/10 232.93
4 1/4/10 258.79
5 3/5/10 254.62
6 1/6/10 249.32
```

คำนวณผลได้ตอบแทนอย่างง่ายสุทธิ โดยการกำหนดเวกเตอร์ P_t และ P_{t-1} แล้วคำนวณตามสมการ **??**

```
> n=nrow(aapl)
> aapl.ret=aapl$Price[2:n]/aapl$Price[2:n-1]-1
> head(aapl.ret)
[1]  0.06539896  0.14845676  0.11102048 -0.01611345 -0.02081533  0.02274186
```

คำนวณผลได้ตอบแทนในรูปลอการิทึมจากสมการ **??**

```
> aapl.lret=diff(log(aapl$Price))
> head(aapl.lret)
[1]  0.06334934  0.13841909  0.10527894 -0.01624468 -0.02103503  0.02248712
```

1.2 แบบจำลองสำหรับข้อมูลอนุกรมเวลา

เป้าหมายหลักของการวิเคราะห์ข้อมูลอนุกรมเวลาคือการสร้างแบบจำลองทางคณิตศาสตร์เพื่ออธิบายคุณลักษณะของข้อมูลตัวอย่าง เช่น ข้อมูลผลได้ตอบแทนของหุ้น เพื่อที่จะให้เราสามารถสร้างแบบจำลองที่จะอธิบายข้อมูลได้ เราจะสมมุติว่าอนุกรมเวลาสามารถนิยามเป็นตัวแปรสุ่ม (random variables) ที่มีการจัดลำดับตามเวลา เช่นลำดับของข้อมูล $y_1, y_2, y_3 ....$ โดยที่ y_1 เป็นค่าของอนุกรมที่คาบเวลาที่ 1, y_2 เป็นค่าของอนุกรมที่คาบเวลาที่ 2 ตามลำดับ โดยทั่วไป เราจะเรียกกลุ่มของตัวแปรสุ่มที่มีการจัดลำดับตามเวลา $Y_t = \{y_t, t = 0, \pm1, \pm2, ...\}$ ว่ากระบวนการสโตแคสติก (Stochastic process) ในวิชานี้เราจะเน้นการวิเคราะห์ข้อมูลอนุกรมเวลาด้วยแบบจำลองเชิงเส้นตรง (Linear model) หมายถึง Y_t มีความสัมพันธ์ขึ้นอยู่กับค่าในอดีตในลักษณะเป็นเส้นตรง

1.2.1 การแจกแจงของอนุกรมเวลา (Distribution of time series)

ตัวแปรสุ่ม Y_t คือฟังก์ชันที่ใช้อธิบายค่าของเหตุการณ์ที่จะเกิดขึ้นในอนาคต โดยที่เรารู้ค่าทั้งหมดที่จะเป็นไปได้แต่ไม่รู้ว่าจะเกิดอะไรขึ้นแน่นอน เช่น ราคาของหลักทรัพย์(P) เราไม่ทราบว่าราคาในเดือนหน้าจะเป็นเท่าไหร่แน่นอน แต่เราคิดว่าราคาจะต้องเป็นบวกและไม่สูงมากนั้น ดังนั้น P จึงเป็นตัวแปรสุ่มมีค่าระหว่าง $[0, M]$ โดยที่ M เป็นค่าที่ไม่สูงมากนัก คำถามอีกคำถามหนึ่งคือตัวแปรสุ่มจะมีการแจกแจงอย่างไร โดยที่หนึ่งในทางเลือกคือการแจกแจงแบบ log-normal

นอกจาก นี้ เรา อาจ จะ สนใจ ว่าการ ลงทุน ของ เรา ใน หนึ่ง เดือน ข้าง หน้า จะ มี ผล ได้ตอบแทน(R_t)เป็นอย่างไร ดังนั้น R_t ก็จะเป็นตัวแปรสุ่มเนื่องจากเราไม่ทราบว่าผลได้ตอบแทนในคาบข้างหน้าจะเป็นอย่างไร ทราบเพียงว่าค่าดังกล่าวเป็นได้ทั้งบวกและลบ โดยที่ลบ(ขาดทุน)ได้อย่างมากไม่เกิน 100% การแจกแจงที่ใช้ประมาณค่าสำหรับผลได้ตอบแทนอย่างง่ายรายเดือนที่ดีคือการแจกแจงแบบปกติ

บางครั้งนักลงทุนอาจจะสนใจว่าราคาจะเปลี่ยนแปลงอย่างไร โดยอาจจะสนใจกรณีที่ราคาปรับตัวขึ้นแทนด้วย 0 และกรณีที่ราคาคงที่หรือลดลงแทนด้วย 1 ในกรณีนี้ปริภูมิตัวอย่าง (sample space) ของตัวแปรสุ่มวิยุต (discrete random variable)

***นิยาม* 1.1.** *ตัวแปรสุ่มวิยุต Y คือตัวแปรสุ่มที่มีปริภูมิตัวอย่างเซตที่มีค่าจำกัด* $S_Y = \{y_1, y_2, ..., y_n\}$ *หรือมีค่าเป็นอนันต์ที่นับได้* $S_Y = \{y_1, y_2, ...\}$

ฟังก์ชันความหนาแน่นของความน่าจะเป็น (probability density function: pdf) ของตัวแปรสุ่มวิยุตสามารถเขียนแทนด้วย $p(y)$ จะเป็นฟังก์ชัน $p(y) = Pr(Y = y)$ ความน่าจะเป็นที่ตัวแปรสุ่ม Y เท่ากับค่า y โดย pdf จะต้องมีคุณสมบัติคือ (1) $p(y) \geq 0$ สำหรับทุกค่า $y \in S_y$ (2) $p(y) = 0$ สำหรับทุกค่า $y \notin S_y$ และ (3) $\sum_{y \in S_y} p(y) = 1$

ตัวอย่างที่ 1.6 การแจกแจงความน่าจะเป็น

สมมุติว่าผลได้ตอบแทนจากหุ้น A มีค่าที่เป็นไปได้และความน่าจะเป็นดังนี้ ตัวอย่างดัง

S_Y	$p(y) = Pr(Y = y)$
-0.30	0.05
0.0	0.20
0.10	0.50
0.20	0.20
0.50	0.05

กล่าวพิจารณาค่าน่าจะเป็นไปได้ 5 ค่าของผลได้ตอบแทนของหุ้น A เราสามารถทำนายค่าที่เป็นไปได้ของแต่ละเหตุการณ์ได้

ตัวอย่างของการแจกแจงแบบ discrete อื่นๆ ได้แก่ Bernoulli และ Binomial

นิยาม 1.2. *ตัวแปรสุ่มแบบต่อเนื่อง Y จะมีค่าเป็นเท่าใดก็ได้บนเส้นจำนวนจริง ดังนั้น* $S_Y = \{y : y \in \mathbb{R}\}$

ฟังก์ชันความหนาแน่นของความน่าจะเป็น pdf ของตัวแปรสุ่มต่อเนื่อง Y จะเป็นฟังก์ชันที่ไม่เป็นค่าลบ f นิยามบนเส้นจำนวนจริงโดยที่สำหรับช่วง A ใดๆ

$$Pr(Y \in A) = \int_A f(y)dy$$

ดังนั้น $Pr(Y \in A)$ คือพื้นที่ใต้กราฟความน่าจะเป็นตลอดช่วง A โดยที่ pdf $f(y)$ จะมีคุณสมบัติดังนี้ (1) $f(y) \geq 0$ และ (2) $\int_{-\infty}^{\infty} f(y)dy = 1$

Example 2 *ตัวอย่างเช่นกราฟรูประฆังรูปที่ **??** เป็น pdf ฟังก์ชัน โดยที่พื้นที่ใต้กราฟระหว่าง* $y = -2$ *ถึง* $y = 1$ *จะแสดงถึง* $Pr(-2 \leq Y < 1)$

ตัวอย่างของการแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องคือการแจกแจงปกติมาตรฐาน (standard normal distribution)

นิยาม 1.3. *ฟังก์ชันการแจกแจงสะสม (cumulative distribution function: cdf) ของตัวแปรสุ่ม Y (ไม่ว่าจะเป็นตัวแปรสุ่มวิยุตหรือต่อเนื่อง) จะมีสัญลักษณ์เป็น F_Y ซึ่งแสดงค่าความน่าจะเป็นที่ Y ใดๆจะมีค่าน้อยกว่าหรือเท่ากับ y*

$$F_Y(y) = P(Y \leq y), \qquad -\infty \leq y \leq \infty$$

ฟังก์ชันการแจกแจงสะสม (cdf) จะมีคุณสมบัติดังนี้

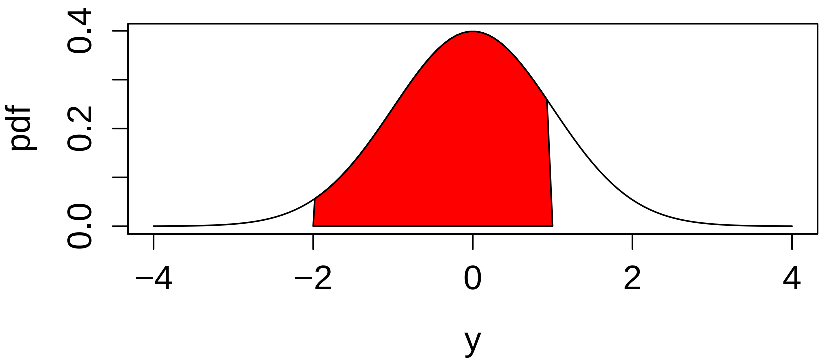
1. ถ้า $y_1 < y_2$ แล้ว $F_Y(y_1) \leq F_Y(y_2)$
2. $F_Y(-\infty) = 0$ และ $F_Y(\infty) = 1$
3. $Pr(Y > y) = 1 - F_Y(y)$
4. $Pr(y_1 < X \leq y_2) = F_Y(y_2) - F_Y(y_1)$
5. $\frac{\partial F_Y(y)}{\partial y} = f_Y(y)$ ถ้า Y เป็นตัวแปรสุ่มต่อเนื่องและ $F_Y(y)$ เป็นฟังก์ชันต่อเนื่องและหาอนุพันธ์ได้

ควอนไทล์ของการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม

หากเราพิจารณาตัวแปรสุ่ม Y ที่มีฟังก์ชันการแจกแจงสะสม $F_Y(y)$ หากค่า $0 \leq \alpha \leq 1$ แล้วค่าควอนไทล์ที่ $100\alpha\%$ ของการแจกแจง Y คือค่า q_α ที่ตรงกับเงื่อนไข

$$F_Y(q_\alpha) = Pr(Y \leq q_\alpha) = \alpha$$

รูปที่ 1.1: ฟังก์ชัน pdf ของการแจกแจงต่อเนื่อง



ตัวอย่างเช่น ควอนไทล์ที่ 5% ของ Y เขียนแทนด้วย $q_{0.05}$ จะเท่ากับค่าที่ทำให้ $F_Y(q_{0.05}) = Pr(Y \leq q_{0.05}) = 0.05$ และถ้า F_Y สามารถหาฟังก์ชันผกผันได้ ดังนั้น $q_\alpha = F_Y^{-1}(\alpha)$

ตัวอย่างที่ 1.7 การหาความน่าจะเป็นจากฟังก์ชันแจกแจงความถี่

กำหนดให้ $Y \sim N(0, 1)$ ค่าควอนไทล์ของการแจกแจงแบบปกติมาตรฐานจะเท่ากับ

$$q_\alpha = \Phi^{-1}(\alpha) \tag{1.12}$$

โดยที่ Φ^{-1} คือฟังก์ชันผกผันของฟังก์ชันการแจกแจงสะสม Φ ซึ่งค่าควอนไทล์จะเป็นค่าที่แสดงในตารางสถิติในหนังสือเกือบทุกเล่ม โดยค่าดังกล่าว

$$q_{0.005} = \Phi^{-1}(0.005) = -2.58, q_{0.01} = \Phi^{-1}(0.01) = -2.33,$$

$$q_{0.025} = \Phi^{-1}(0.025) = -1.96, q_{0.05} = \Phi^{-1}(0.05) = -1.645$$

โดยทั่วไปค่าควอนไทล์ของการแจกแจงปกติมาตรฐานมักจะแทนด้วย Z_α อย่างไรก็ตามหนังสือบางเล่มอาจจะเขียนสัญลักษณ์ดังกล่าวแทนค่าควอนไทล์นับจากด้านบน (upper quantile) โดยที่ค่าควอนไทล์นับจากด้านบน 5% จะเท่ากับค่าควอนไทล์ 95 % เป็นต้น

การคำนวณพื้นที่ใต้กราฟด้วย R

ใน โปรแกรม R เราสามารถคำนวณพื้นที่ใต้กราฟ pdf หรือ $Pr(Y < z)$ ด้วยคำสั่ง `pnorm(z)` โดยที่ $Pr(Y < -2)$ คำนวณได้โดย `pnorm(-2)` นอกจากนี้เราสามารถหาค่าควอนไทล์ที่ $100\alpha\%$ ได้ด้วยคำสั่ง `qnorm(`α`)` เช่น `qnorm(0.01)=-2.326`

1.2.2 คุณลักษณะเรื่องรูปร่างของการแจกแจงความน่าจะเป็น

ในการพิจารณาข้อมูลทางการเงิน เราสนใจเกี่ยวกับรูปร่างของการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่เราสนใจ เช่น เราต้องการที่จะทราบว่าค่ากลางของการแจกแจงอยู่ที่ตรงไหน หรือการแจกแจงมีการแผ่ขยายไปข้างๆอย่างไร นอกจากนี้เรายังสนใจว่าการแจกแจงมีความเป็นสมมาตรหรือไม่ มีรูปร่างที่เบ้ไปทางซ้ายหรือทางขวา หรือความน่าจะเป็นที่จะเกิดเหตุการณ์ที่ห่างจากค่ากลางมากๆ (extreme value) โดยเฉพาะค่าที่ติดลบมากๆ จากสิ่งที่เราสนใจที่ได้กล่าวมาแล้ว คุณลักษณะด้านรูปร่างของการแจกแจงความน่าจะเป็นใดๆจะวัดได้ด้วยคุณลักษณะ 4 ประการ

1. ค่าคาดหมาย (expected value) หรือค่าเฉลี่ยเป็นการวัดค่ากลางของการแจกแจง
2. ความแปรปรวนและค่าเบี่ยงเบนมาตรฐานวัดการแผ่จากค่ากลาง
3. ความเบ้ (skewness) วัดความสมมาตรรอบๆค่ากลาง
4. ค่าความโด่ง (kurtosis) วัดความอ้วนของหาง

นิยาม 1.4. ฟังก์ชันค่าคาดหมาย (mean function)*ของตัวแปรสุ่ม Y ใดๆ ใช้สัญลักษณ์ $E(Y)$ สามารถคำนวณได้ดังนี้ ในกรณีที่ Y เป็นตัวแปรสุ่มวิยุต ค่าคาดหมายจะเท่ากับ*

$$\mu_Y = E(Y) = \sum_{y \in S_Y} y \times P(Y = y) \tag{1.13}$$

หรือในกรณีตัวแปรสุ่มต่อเนื่อง ค่าคาดหมายจะเท่ากับ

$$\mu_Y = E(Y) = \int_{-\infty}^{\infty} y f(y) dy \tag{1.14}$$

โดยที่ E คือสัญลักษณ์แทนค่าคาดหมาย (Expected value)

ค่าคาดหมายของฟังก์ชันของตัวแปรสุ่ม

ลักษณะเฉพาะอื่นด้านรูปร่างของการแจกแจงของตัวแปรสุ่ม Y มีพื้นฐานจากค่าคาดหมายของฟังก์ชันของ Y เช่นสมมุติให้ $g(Y)$ เป็นฟังก์ชันของ Y ถ้า Y เป็นตัวแปรสุ่มแบบ discrete แล้ว

$$E[g(Y)] = \sum_{y \in S_Y} g(y) P(Y = y)$$

และถ้า Y เป็นตัวแปรสุ่มแบบต่อเนื่องด้วย pdf $f(y)$ แล้ว

$$E[g(Y)] = \int_{-\infty}^{\infty} g(y) f(y) dy$$

ค่าความแปรปรวนและค่าเบี่ยงเบนมาตรฐาน

ค่าแปรปรวนของตัวแปรสุ่ม Y เขียนแทนด้วย $Var(Y)$ หรือ σ_Y^2 วัดการแผ่ของการแจกแจงจากค่าเฉลี่ย โดยที่ค่าแปรปรวนสามารถนิยามได้โดย

$$\sigma_Y^2 = Var(Y) = E\left[(Y - \mu_Y)^2\right] \tag{1.15}$$

นอกจากนี้มีอีกสูตรหนึ่งที่เรามักใช้บ่อยในการคำนวณค่าความแปรปรวน

$$\sigma_Y^2 = Var(Y) = E\left[Y^2\right] - \mu_Y^2 \tag{1.16}$$

เนื่องจากหน่วยวัดของค่าความแปรปรวนมีหน่วยที่แตกต่างจากตัวแปรสุ่ม Y ดังนั้น เรามักจะพิจารณาค่าเบี่ยงเบนมาตรฐานเขียนแทนด้วย $sd(Y)$ หรือ σ_Y ซึ่งเท่ากับค่ารากที่สองของค่าความแปรปรวน ($\sqrt{\sigma_Y^2}$)

ค่าความเบ้

ค่าความเบ้ซึ่งสามารถเขียนแทนด้วย ($S(Y)$) วัดความสมมาตรรอบๆค่ากลาง สามารถทำได้โดย

$$S(Y) = \frac{E[(Y - \mu_Y)^3]}{\sigma_Y^3} \tag{1.17}$$

โดยที่หากค่าความเบ้เป็นศูนย์แสดงว่าการแจกแจงมีความสมมาตร ถ้าค่าความเบ้เป็นบวกแสดงว่าข้อมูลถูกดึงไปด้านขวาของการแจกแจง(หางไปทางขวา) แต่หากค่าความเบ้เป็นลบแสดงว่าข้อมูลถูกดึงไปทางซ้าย(หางไปทางซ้าย)ในแง่ของการลงทุนหากผลได้ตอบแทนถูกดึงไปด้านซ้าย แสดงว่ามีโอกาสที่เราจะได้ผลลัพธ์ที่สุดโต่งค่อนข้างมาก แต่ถ้าหางไปทางขวาโอกาสที่จะเกิดเหตุการณ์ร้ายที่รุนแรงค่อนข้างน้อย ตัวแปรสุ่มที่มีการแจกแจงแบบปกติจะมีค่าความเบ้เป็นศูนย์ เนื่องจากมีการแจกแจงที่เป็นสมมาตร

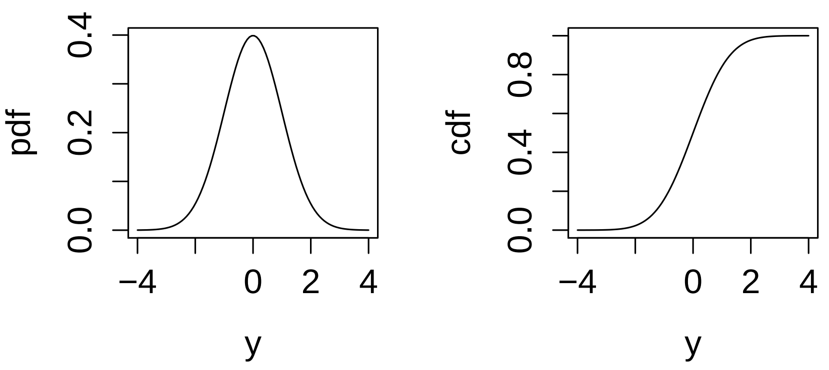
ค่าความโด่ง

ค่าความโด่งใช้วัดความหนาของหางของการแจกแจงซึ่งสามารถเขียนแทนได้ด้วย $K(Y)$ และสามารถคำนวณได้จาก

$$K(Y) = \frac{E[(Y - \mu_Y)^4]}{\sigma_Y^4} \tag{1.18}$$

เนื่องจากค่าความโด่งเป็นการหาความแตกต่างจากค่ากลางโดยการยกกำลัง 4 ดังนั้นค่าที่ห่างจากจุดศูนย์กลางมากๆจะทำให้ค่าถ่วงน้ำหนักสูงขึ้น และค่าความโด่งสูง ในทางตรงข้ามหากค่าความโด่งต่ำแสดงว่าข้อมูลกระจุกตัวอยู่ตรงกลางและมีโอกาสน้อยที่จะพบข้อมูลที่มีค่าสุดโต่ง ค่าความโด่งของตัวแปรสุ่มที่มีการแจกแจงปกติเท่ากับ 3 โดยเรามักใช้ค่าความโด่งดังกล่าวเป็นมาตรฐานความหนาของหาง หากการแจกแจงใดมีค่าความโด่งมากกว่า 3 แสดงว่าการแจกแจงนั้นมีหางที่

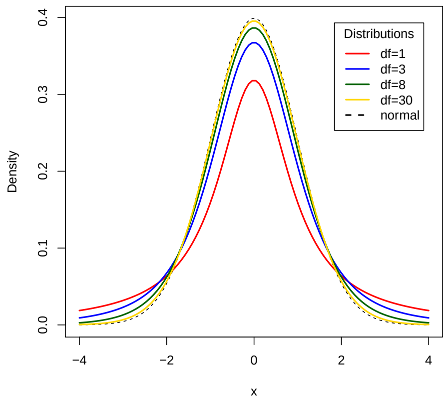
รูปที่ 1.2: ฟังก์ชันความหนาแน่นของความน่าจะเป็นและการแจกแจงสะสมของการแจกแจงปกติมาตรฐาน



อ้วนกว่า (thicker tail)การแจกแจงแบบปกติ แต่ถ้าการแจกแจงมีความความโด่งน้อยกว่า 3 แสดงว่าการแจกแจงมีหางที่ผอมกว่าการแจกแจงปกติ

บางครั้งเราแสดงค่าความโด่งในรูปความโด่งเปรียบเทียบกับการแจกแจงปกติ โดยค่าดังกล่าวเรียกว่าความโด่งส่วนเกิน (excess kurtosis) โดย excess $K(Y) = K(Y) - 3$ ถ้าค่าความโด่งส่วนเกินเท่ากับศูนย์แสดงว่าตัวแปรสุ่มมีค่าความโด่งเท่ากับการแจกแจงปกติ ถ้าค่าความโด่งส่วนเกินเป็นบวกแสดงว่าการแจกแจงนั้นมีหางที่อ้วนกว่าการแจกแจงแบบปกติ

รูปที่ 1.4: การแจกแจงแบบที



ค่าความโด่งของการแจกแจงปกติเป็นเกณฑ์มาตรฐานสำหรับความหนาของหางของการแจกแจงที่เป็นสมมาตร ตัวอย่างการแจกแจงอื่นที่มีลักษณะสมมาตรและมีหางที่อ้วนกว่าการแจกแจงแบบปกติคือ **การแจกแจงแบบที (Student's t)** โดยที่ถ้า Y มีการแจกแจกแบบทีด้วยองศาเสรี (degree of freedom) v จะมีค่าเฉลี่ยเท่ากับ 0, ค่าความแปรปรวนเท่ากับ $v/v-2$ โดยที่ $v > 2$, ค่าความเบ้เท่ากับ 0, และค่าความโด่งเท่ากับ $\frac{6}{(v-4)} - 4$ โดยที่ $v > 4$ เราจะเห็นได้ว่าองศาเสรี v เป็นตัวกำหนดการแผ่และความหนาของหางของการแจกแจง ถ้าค่าองศาเสรี v เข้าใกล้ 4 การแจกแจงจะมีหางที่อ้วนมาก แต่ถ้าค่าองศาเสรี v มีค่าเข้าใกล้อนันต์ ฟังก์ชันการแจกแจง(pdf)ของการแจกแจงทีจะมีลักษณะเข้าใกล้ฟังก์ชันการแจกแจง(pdf)ของการแจกแจงปกติ การแจกแจงความน่าจะเป็นแบบทีสำหรับค่าองศาเสรีที่ต่างกันสามารถดูได้จากรูปที่ **??**

ตัวประมาณค่า(estimator) ของค่าคุณลักษณะของการแจกแจง

ในการประยุกต์เราสามารถประมาณค่าเฉลี่ย ค่าความแปรปรวน ค่าความเบ้และค่าความโด่งได้ด้วยข้อมูลจากกลุ่มตัวอย่าง สมมุติให้เราสุ่มตัวอย่าง $y_1, ..., y_T$ ซึ่งมีขนาดเท่ากับ T ค่าเฉลี่ยของตัวอย่าง (sample mean) สามารถคำนวณได้โดย

$$\hat{\mu}_Y = \frac{1}{T}\sum_{t=1}^{T} y_t \tag{1.19}$$

ค่าความแปรปรวนของตัวอย่างสามารถคำนวณได้โดย

$$\hat{\sigma}_Y^2 = \frac{1}{T-1}\sum_{t=1}^{T}(y_t - \hat{\mu}_Y)^2 \tag{1.20}$$

ค่าความเบ้ของตัวอย่าง (sample skewness) สามารถคำนวณได้โดย

$$\hat{S}(Y) = \frac{1}{(T-1)\hat{\sigma}_Y^3}\sum_{t=1}^{T}(y_t - \hat{\mu}_Y)^3 \tag{1.21}$$

และค่าความโด่งของตัวอย่าง (sample kurtosis) สามารถคำนวณได้โดย

$$\hat{K}(Y) = \frac{1}{(T-1)\hat{\sigma}_Y^4}\sum_{t=1}^{T}(y_t - \hat{\mu}_Y)^4 \tag{1.22}$$

การทดสอบสมมุติฐาน

ภายใต้ข้อสมมุติที่ว่า Y_t มีการแจกแจงแบบปกติ $N(0, \sigma^2)$ ค่าเฉลี่ยของตัวอย่างจะมีการแจกแจงแบบปกติ $N(0, \frac{\sigma^2}{T})$ ดังนั้นในการทดสอบสมมุติฐานที่ว่าค่าเฉลี่ยของประชากรมีค่าเท่ากับศูนย์หรือไม่ เราสามารถทำได้โดยใช้ค่าสถิติสัดส่วน t (t-ratio)

$$t = \frac{\hat{\mu}_Y}{\hat{\sigma}_Y/\sqrt{T}}$$

โดยเราจะปฏิเสธสมมุติฐานหลักด้วยระดับนัยสำคัญ(significant level) α หรือความเชื่อมั่น $(1-\alpha/2)100\%$ ถ้า $|t| > Z_{1-\alpha/2}$ โดยที่ $Z_{1-\alpha/2}$ คือค่าควอนไทล์ที่ $100(1-\alpha/2)\%$ ของการแจกแจงปกติมาตรฐาน (standard normal)

ค่าความเบ้ของตัวอย่าง ($\hat{S}(Y)$) จะมีการแจกแจงเมื่อกลุ่มตัวอย่างใหญ่เข้าใกล้การแจกแจงปกติโดยมีค่าเฉลี่ยเท่ากับศูนย์และค่าความแปรปรวนเท่ากับ $6/T$ ส่วนค่าความโด่งส่วนเกินของตัวอย่าง ($\hat{K}(Y)-3$) จะมีการแจกแจงเมื่อกลุ่มตัวอย่างใหญ่เข้าใกล้การแจกแจงปกติโดยมีค่าเฉลี่ยเท่ากับศูนย์และค่าความแปรปรวนเท่ากับ $24/T$ ซึ่งคุณสมบัติทั้งสองนำมาใช้ในการทดสอบว่าตัว-

แปรสุ่ม Y_t มีการแจกแจงแบบปกติหรือไม่ โดยหากตัวแปรสุ่มมีการแจกแจงแบบปกติค่าความเบ้และค่าความโด่งส่วนเกินจะต้องเท่ากับศูนย์

กำหนดให้เรามีข้อมูลผลได้ตอบแทน $y_1, ..., y_T$ และต้องการทดสอบความเบ้ว่าเท่ากับศูนย์หรือไม่ เราจะตั้งสมมุติฐานหลักว่า $H_0 : S(Y) = 0$ และสมมุติฐานทางเลือกว่า $H_1 : S(Y) \neq 0$ โดยมีตัวสถิติสัดส่วน t (t-ratio)

$$t = \frac{\hat{S}(Y)}{\sqrt{6/T}}$$

โดยเราจะปฏิเสธสมมุติฐานหลักที่ระดับนัยสำคัญ α ถ้า $|t| > Z_{1-\alpha/2}$ โดยที่ $Z_{1-\alpha/2}$ คือค่าควอนไทล์ที่ $100(1-\alpha/2)\%$ ของการแจกแจงปกติมาตรฐาน หรือเราอาจจะใช้การคำนวณค่าพี (p-value)

เราสามารถทดสอบค่าความโด่งส่วนเกินได้ด้วยการตั้งสมมุติฐานว่า $H_0 : \hat{K}(Y) - 3 = 0$ และ $H_1 : \hat{K}(Y) - 3 \neq 0$ และมีตัวสถิติคือ

$$t = \frac{\hat{K}(Y) - 3}{\sqrt{24/T}}$$

โดยเราจะปฏิเสธสมมุติฐานหลักที่ระดับนัยสำคัญ α ถ้า $|t| > Z_{100-\alpha/2}$ โดยที่ $Z_{100-\alpha/2}$ คือค่าควอนไทล์ที่ $100(1 - \alpha/2)$ ของการแจกแจงปกติมาตรฐาน

Jarque and Bera (1987) ได้รวมการทดสอบของค่าสถิติทั้งสอบเข้าด้วยกันและเสนอตัวสถิติ

$$JB = \frac{\hat{S}^2(Y)}{6/T} + \frac{[\hat{K}(Y) - 3]^2}{24/T}$$

ซึ่งมีการแจกแจงเมื่อกลุ่มตัวอย่างใหญ่เข้าใกล้ไคกำลังสอง(χ^2)ที่มีองศาอิสระเท่ากับ 2 โดยเราจะปฏิเสธสมมุติฐานหลักที่ระดับนัยสำคัญ α ถ้า $JB > \chi^2_{(1-\alpha),df=2}$ โดยที่ $\chi^2_{(1-\alpha),df=2}$ คือค่าควอนไทล์ที่ $100(1 - \alpha)\%$ ของการแจกแจงไคกำลังสองที่มีองศาอิสระเท่ากับ 2

นอกจากนี้เรายังมีวิธีการที่เป็นที่นิยมอันหนึ่งคือการวาดแผนภาพ Normal quantile-quantile หรือเรียกย่อๆว่า QQ plot โดยเป็นการวาดจุด (scatterplot) ค่าที่เรียงจากต่ำ (quantile) ของอนุกรมเวลา y_t กับค่าที่เรียงจากต่ำของตัวแปรสุ่มที่มีการแจกแจงแบบปกติ ดังนั้นถ้าจุดดังกล่าวอยู่ใกล้กับเส้น 45 องศาแสดงว่ากระบวนการ y_t มีการแจกแจงแบบปกติ

1.2.3 ตัวอย่างการคำนวณค่าสถิติและการทดสอบสมมุติฐานจากผลได้ตอบแทนหุ้น PTT

ตัวอย่างนี้ใช้ข้อมูลผลได้ตอบแทนจากภาคผนวก A [ptt.Rdata] ซึ่งเราได้คำนวณค่าผลได้ตอบแทนในรูปล็อกของหุ้น PTT(`ptt$lret`) ซึ่งในหัวข้อนี้เราจะคำนวณค่าสถิติเบื้องต้นของ `ptt$lret` โดยใช้ package `fBasics` โดยใช้ฟังก์ชัน `basicStats`

```
> library(fBasics)
> basicStats(ptt$lret)
             X..ptt.lret
nobs         2089.000000
NAs             1.000000
Minimum        -0.185899
Maximum         0.149532
1. Quartile    -0.009569
3. Quartile     0.010989
Mean            0.000774
Median          0.000000
Sum             1.617043
SE Mean         0.000490
LCL Mean       -0.000187
UCL Mean        0.001736
Variance        0.000502
Stdev           0.022405
Skewness       -0.068862
Kurtosis        6.228529
```

จะเห็นได้ว่าค่าเฉลี่ยของผลได้ตอบแทนในรูปล็อกมีค่าเท่ากับ 0.00077 หรือ 0.077 % ต่อวัน ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.0224 ค่าความเบ้ของตัวอย่าง (sample skewness) เท่ากับ -0.0669 และค่าความโด่งส่วนเกินของตัวอย่าง (sample excess kurtosis) เท่ากับ 6.2285 [`Kurtosis` ที่รายงานใน R เป็นค่าความโด่งส่วนเกินของตัวอย่าง]

การทดสอบค่าเฉลี่ย

หากต้องการทราบว่าค่าเฉลี่ยของผลได้ตอบแทนมีค่าต่างจากศูนย์หรือไม่ก็สามารถทำได้โดยการคำนวณตัวสถิติ t โดยที่จำนวนตัวอย่างเท่ากับ `nobs-NA` (จำนวนตัวอย่าง-จำนวนข้อมูลที่ไม่มี) = 2089-1=2088 ดังนั้นค่าสถิติ t จะเท่ากับ

$$t = \frac{0.00077}{0.0224/\sqrt{2088}} = 1.579$$

เมื่อเปรียบเทียบกับค่า $Z_{1-0.05/2} = 1.96$ ค่าดังกล่าวสามารถคำนวณได้โดย `qnorm(0.975)` เราพบว่าค่า $|t| < 1.96$ ดังนั้นเราไม่สามารถปฏิเสธสมมุติฐานหลักว่า "ค่าเฉลี่ยของผลได้ตอบแทนมีค่าเท่ากับศูนย์" นอกจากนี้เราสามารถทดสอบค่าเฉลี่ยเท่ากับศูนย์ได้ด้วยฟังก์ชัน `t.test`

```
> t.test(ptt$lret)
  One Sample t-test
data:  ptt$lret
t = 1.5795, df = 2087, p-value = 0.1144
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.000187103  0.001735994
sample estimates:
   mean of x
0.0007744456
```

จะเห็นได้ว่าฟังก์ชันดังกล่าวคำนวณค่า $t = 1.5795$ และค่าพี (p-value) เท่ากับ 0.1144 ซึ่งเราไม่สามารถปฏิเสธสมมุติฐานหลักที่นัยสำคัญเท่ากับ 0.05

การทดสอบความเป็นสมมาตร

จากค่าความเบ้ของตัวอย่าง เราสามารถคำนวณค่าสถิติ t ที่ใช้ทดสอบ H_0 ว่าข้อมูลมีความสมมาตร ($S(Y) = 0$) ได้เท่ากับ

$$t = \frac{-0.0689}{\sqrt{6/2088}} = -1.2846$$

เมื่อเปรียบเทียบกับค่า $|t|$ กับ $Z_{1-0.05/2} = Z_{0.975} = 1.96$ เราจะได้ $|t| < Z_{0.975} = 1.96$ เราจึงไม่สามารถปฏิเสธ H_0 ที่ว่าผลได้ตอบแทนในรูปล็อกมีลักษณะสมมาตร ที่ระดับความเชื่อมั่น 95 %

การทดสอบความหนาของหาง

$$t = \frac{6.2285}{\sqrt{24/2088}} = 58.0959$$

เมื่อเปรียบเทียบกับค่า $|t|$ กับ $Z_{1-0.05/2} = Z_{0.975} = 1.96$ เราจะได้ $|t| > Z_{0.975} = 1.96$ เราจึงปฏิเสธสมมุติฐานหลัก H_0 ที่ว่าผลได้ตอบแทนในรูปล็อกมีความหนาของหางเท่ากับการแจกแจงแบบปกติ และจากค่าความโด่งของตัวอย่างเราสามารถอนุมานได้ว่าการแจกแจงมีลักษณะเป็นหางอ้วน (ความโด่งส่วนเกินมากกว่าศูนย์)

การทดสอบการแจกแจงปกติ

$$JB = \frac{-0.0689^2}{6/2088} + \frac{6.2285^2}{24/2088} = 3376$$

เมื่อเปรียบเทียบกับค่า $JB > 5.99 (= \chi^2_{0.95,df=2})$ ค่าดังกล่าวสามารถคำนวณโดยใช้ฟังก์ชัน `qchisq(1-0.05,2)` เราจึงปฏิเสธสมมุติฐานหลัก H_0 ที่ว่าผลได้ตอบแทนในรูปล็อกมีการแจกแจงเป็นแบบปกติ โดยการทดสอบข้างต้นสามารถใช้คำสั่ง `normalTest` ใน package `fBasics` โดยที่ระบุ argument คือ ตัวแปรที่ต้องการทดสอบ `lret` และวิธีการทดสอบ `method="jb"` สำหรับการทดสอบ Jarque-Bera ซึ่งคำสั่งและผลสามารถแสดงได้ดังนี้

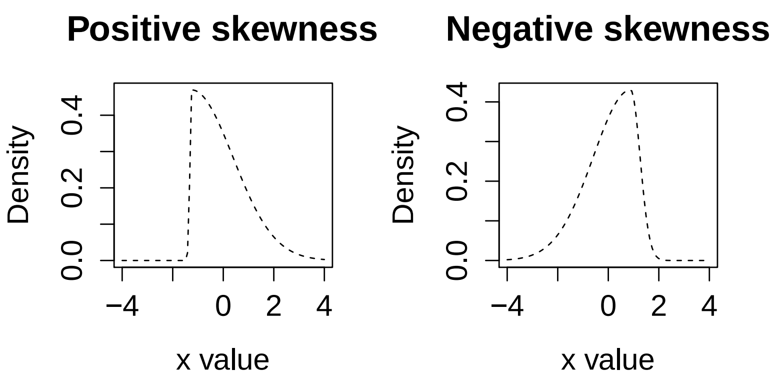
```
> normalTest(ptt$lret, method="jb")

Title:
 Jarque — Bera Normalality Test

Test Results:
  STATISTIC:
    X—squared: 3386.3742
  P VALUE:
    Asymptotic p Value: < 2.2e—16
```

การทดสอบดังกล่าวได้ค่าต่างจากการคำนวณด้วยมือเล็กน้อยเนื่องจากการปัดเศษ นอกจากนี้ในการทดสอบโปรแกรมจะคำนวณค่าพีซึ่งมีค่าน้อยมาก ($< 2.2 \times 10^{-16}$) ดังนั้นเราจึงปฏิเสธสมมุติฐานหลักด้วยนัยสำคัญใดๆมากกว่า 2.2×10^{-16} %

เราสามารถสร้างแผนภาพการแจกแจงข้อมูล (empirical density) โดยใช้คำสั่ง `density` ซึ่งแสดงในเส้น kernel density ในรูปที่ **??** นอกจากนี้ยังมีแผนภาพแสดงการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ 0.00077 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.0244

```
> plot(density(ptt$lret), main="Density␣of␣log␣return")
```

รูปที่ 1.3: ฟังก์ชันความถี่ความน่าจะเป็นที่ไม่สมมาตร



การแจกแจงของผลได้ตอบแทน

โดยทั่วไปแล้ว เรามักจะสมมุติให้พิจารณาผลได้ตอบแทนในรูปของล็อกและมักจะสมมุติให้มีการแจกแจงแบบปกติ นักศึกษาบางคนอาจจะสงสัยว่าทำไม่ไม่ใช้ผลได้ตอบแทนอย่างง่าย โดยเราสามารถตอบคำถามดังกล่าวได้โดยการสมมุติให้ ผลได้ตอบแทนอย่างง่ายมีการแจกแจงแบบปกติ $R_t \sim N(0.05, (0.05)^2)$ ซึ่งเราทราบว่าราคาสินทรัพย์ใดๆจะมีค่าติดลบไม่ได้ ดังนั้น R_t จะต้องมีค่ามากกว่า -1 ซึ่งหากพิจารณาจากข้อสมมุติการแจกแจงปกติจะเห็นได้ว่า $Pr(R_t < -1) = 0.018$ หรือมีโอกาส 1.8 % ที่ราคาหุ้นจะติดลบ! ซึ่งเป็นไปไม่ได้ ดัง

นั้นจึงมีความเหมาะสมมากกว่าที่จะสมมุติให้ผลได้ตอบแทนในรูปล็อกมีการแจกแจงแบบปกติ $\ln(1+R_t) = r_t \sim N(0.05,(0.5)^2)$ โดยในกรณีผลได้ตอบแทนในรูปล็อกสามารถจะมีค่าน้อยกว่า -1 ได้ เช่น หาก $r_t = -2$ จะสัมพันธ์กับ $R_t = \exp(-2) - 1 = -0.865$ ดังนั้น $Pr(r_t \leq -2) = Pr(R_t \leq -0.865) = 0.00002$ ดังนั้นในแบบจำลองในหัวข้อต่อๆไปเวลาเราพูดถึงผลได้ตอบแทนเรามักจะใช้ผลได้ตอบแทนในรูปล็อก (log return)

ฟังก์ชันเส้นตรงของตัวแปรสุ่ม

หากกำหนดให้ X เป็นตัวแปรสุ่มที่มี $E[X] = \mu_X$ และ $Var(X) = \sigma_X^2$ และ a และ b เป็นค่าคงที่ หากเรากำหนดตัวแปรสุ่มใหม่ Y เป็นฟังก์ชันเส้นตรงกับตัวแปรสุ่ม X โดยที่ $Y = a + bX$ แล้ว

- $\mu_Y = E(Y) = a + bE(X) = a + b\mu_X$
- $\sigma_Y^2 = Var(Y) = b^2 Var(X) = b^2\sigma_X^2$

ฟังก์ฟันค่าแปรปรวนร่วมในตนเอง (Autocovariance function)

ฟังก์ชันที่ใช้ในการสรุปความสัมพันธ์ขึ้นอยู่ต่อกันตามเวลา (temporal dependence) ในข้อมูลอนุกรมเวลาคือ ฟังก์ฟันค่าแปรปรวนร่วมในตนเองซึ่งสามารถคำนวณได้โดย

$$\begin{aligned}\gamma_{k,t} &= Cov(Y_t, Y_{t-k}) = E\left((y_t - E(Y_t))(y_{t-k} - E(Y_{t-k}))\right) \\ &= E(y_t y_{t-k}) - E(Y_t)E(Y_{t-k}) \quad \text{for} \quad k = 0, 1, 2, ... \end{aligned} \tag{1.23}$$

โดยสมการดังกล่าวสอดคล้องกับค่าแปรปรวนร่วม (covariance) ระหว่างตัวแปรสุ่มสองตัวใดๆ

นิยาม 1.5. (Strict stationary) *อนุกรมเวลา Y_t ใดๆจะถูกเรียกว่า strictly stationary ถ้าการแจกแจงร่วม (joint distribution) ของ $(y_{t_1}, ..., y_{t_k})$ เหมือนกันกับการแจกแจงร่วมของ $(y_{t_1+t}, ..., y_{t_k+t})$ สำหรับทุกค่าของ t*

นิยาม 1.6. (weakly stationary หรือ covariance stationary) *ข้อมูลอนุกรมเวลา Y_t ใดๆ จะเรียกว่าเป็น **Weakly stationary** ถ้ามีคุณสมบัติดังนี้*

1. $E(Y_t) = \mu$
2. $Var(Y_t) = \sigma^2 < \infty$
3. $\gamma_{k,t} = \gamma_k = Cov(Y_t, Y_{t-k})$ *for all k and t*

โดยสรุปแล้วอนุกรมจะมีค่าเฉลี่ยคงที่, ค่าแปรปรวนจำกัด (finite) และฟังก์ฟันค่าแปรปรวนร่วมในตนเองที่เป็นฟังก์ชันที่ขึ้นกับค่าล่าเท่านั้นและเป็นอิสระกับเวลา

นิยาม 1.7. Ergodic *อนุกรมเวลาใดๆ จะเป็นอนุกรมเวลาเออร์โกดิก (ergodic) ถ้าค่าโมเมนต์ของตัวอย่างมีค่าลู่เข้าในความน่าจะเป็น (converge in probability) ไปสู่ค่าโมเมนต์ของประชากร ตัวอย่างเช่น*

$$\bar{y} \xrightarrow{p} \mu$$
$$\widehat{\gamma}_j \xrightarrow{p} \gamma_j$$
$$\widehat{\rho}_j \xrightarrow{p} \rho_j$$

ฟังก์ชันสหสัมพันธ์ในตัวเอง (Autocorrelation function; ACF)

ในการวิเคราะห์สถิติทั่วไปเรามักจะสนใจสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) มากกว่าค่าแปรปรวนร่วม(covariance)เนื่องจากปราศจากผลของหน่วยของข้อมูล ดังนั้น **ฟังก์ชันสหสัมพันธ์ในตัวเองที่ช้ากว่ากัน k คาบเวลา** สำหรับอนุกรมที่เป็น weakly stationary จะคำนวณได้จาก

$$\rho_k = \frac{Cov(Y_t, Y_{t-k})}{[Var(Y_t)Var(Y_{t-k})]^{1/2}}, \qquad k = 0, 1, 2, ... \tag{1.24}$$

โดยที่ $\rho_0 = 1$ และ $|\rho_k| \leq 1$ สำหรับทุกค่า k. สำหรับข้อมูลที่เป็น weakly stationary $Var(Y_t) = Var(Y_{t-k}) = \gamma_0$ ดังนั้น

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \qquad k = 0, 1, 2, ... \tag{1.25}$$

เนื่องจากฟังก์ชันฟังก์ฟันค่าแปรปรวนร่วมในตนเองมีความเป็นสมมาตรดังนั้น $\gamma_k = \gamma_{-k}$ และ $\rho_k = \rho_{-k}$ นอกจากนี้กราฟที่แสดงค่าฟังก์ชันสหสัมพันธ์ในตัวเองในแกนตั้งและค่าล่า k ในแกนนอนเราจะเรียกว่า*โครีโลแกรม (Correlogram)* โดยสรุปแล้วการพิจารณาว่าอนุกรมเวลานิ่งไม่ก็จะพิจารณาจากค่าเฉลี่ย, ค่าความแปรปรวนและฟังก์ชันสหสัมพันธ์ในตัวเอง

นอกจากนี้เรายังมีคุณสมบัติที่น่าสนใจว่าฟังก์ชันของอนุกรมเวลานิ่งก็จะเป็นอนุกรมเวลานิ่งด้วย เช่นถ้า Y_t เป็นอนุกรมเวลานิ่งแล้ว $Z_t = g(Y_t)$ ก็จะเป็นอนุกรมเวลานิ่งด้วย

เราสามารถคำนวณค่าแปรปรวนร่วมในตนเองที่ช้ากว่ากัน k คาบเวลาของตัวอย่าง (lag k^{th} sample autocovariance) และสหสัมพันธ์ร่วมในตนเองที่ช้ากว่ากัน k คาบเวลาของตัวอย่าง (lag k^{th} sample autocorrelation) ได้จากสูตรดังต่อไปนี้

$$\hat{\gamma}_k = \frac{1}{T}\sum_{t=k+1}^{T}(y_t - \bar{Y})(y_{t-k} - \bar{Y}) \tag{1.26}$$

และ

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \tag{1.27}$$

โดยที่ $\bar{Y} = \frac{1}{T}\sum_t^T y_t$ คือค่าเฉลี่ยของตัวอย่าง โดยที่ฟังก์ชันสหสัมพันธ์ร่วมในตนเองของตัวอย่าง (sample ACF) จะวาด $\hat{\rho}_k$ กับ k

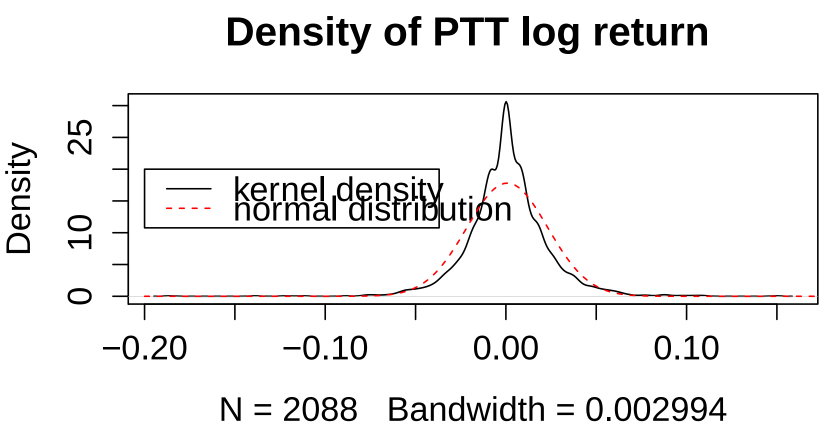
ตัวอย่างหนึ่งของอนุกรมเวลานิ่งคือ กระบวนการเกาซเซียนไวทนอซที่เป็นอิสระต่อกัน (independent Gaussian white noise process) โดยที่เกาซเซียนหมายถึงการแจกแจงเป็นปกตินั่นเอง ถ้ากระบวนการ y_t มีการแจกแจงแบบเกาซเซียนไวทนอซที่เป็นอิสระต่อกันโดยมีค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ σ^2 กระบวนการ y_t จะมีค่าเฉลี่ยเท่ากับ 0 นอกจากนี้ค่าความแปรปรวนร่วมในตัวเองและสหสัมพันธ์ในตัวเองก็จะเท่ากับ 0

เราสามารถสร้างข้อมูลที่เป็นเกาซเซียนไวทนอซที่มีจำนวนตัวอย่างเท่ากับ 500 ได้ด้วยคำสั่ง rnorm ใน R

```
> set.seed(123456)
> y=rnorm(500,0,1)
> library(TSA)
> par(mfrow=c(1,2))
> plot.ts(y,ylab="Y",main="Y")
> acf(y,lag.max=20,main="ACF")
```

จะได้แผนภาพดังนี้

รูปที่ 1.5: density ของผลได้ตอบแทนในรูปล็อกของ PTT



ในโปรแกรม R แผนภาพสำหรับฟังก์ชันสหสัมพันธ์ในตัวเองจะมีเส้นประแสดงความเชื่อมั่น 95% จาก 0 ให้เสมอ โดยที่เส้นประดังกล่าวมีพื้นฐานจากการที่ $y_t \sim iidN(0,\sigma^2)$ ดังนั้น

$$\hat{\rho}_k \overset{a}{\sim} N(0,\frac{1}{T}), \qquad \text{for} \qquad k>0$$

หมายความว่าการแจกแจงของ $\hat{\rho}_k$ มีประมาณใกล้เคียงกับการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และค่าแปรปรวนเท่ากับ $\frac{1}{T}$ และเมื่อใช้ผลจาก central limit theorem เราจะได้ว่า $\sqrt{T}\hat{\rho}_k \xrightarrow{d} N(0,1)$ และเส้นประที่แสดงค่าขอบเขตความเชื่อมั่น 95% ที่ค่าสหสัมพันธ์ในตัวเองเท่ากับศูนย์จะเท่ากับ $\pm\frac{1.96}{\sqrt{T}}$ ดังนั้นหากค่าสหสัมพันธ์ในตัวเองของตัวอย่าง ณ คาบล่าช้า k ใดๆอยู่เกินจากขอบเขตดังกล่าวแสดงว่าค่าสหสัมพันธ์ในตัวเองมีค่าแตกต่างจากศูนย์อย่างมีนัยสำคัญด้วยความเชื่อมัน 95%

ในกรณีที่ $y_t = \mu + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i}$ โดยที่ $\varepsilon_i \sim iid(0, \sigma^2)$ การแจกแจงจะเปลี่ยนไปเป็น

$$\hat{\rho}_k \overset{a}{\sim} N(0, \frac{(1 + 2\sum_{i=1}^{k-1} \rho_i^2)}{T}), \qquad \text{for} \qquad k > 0$$

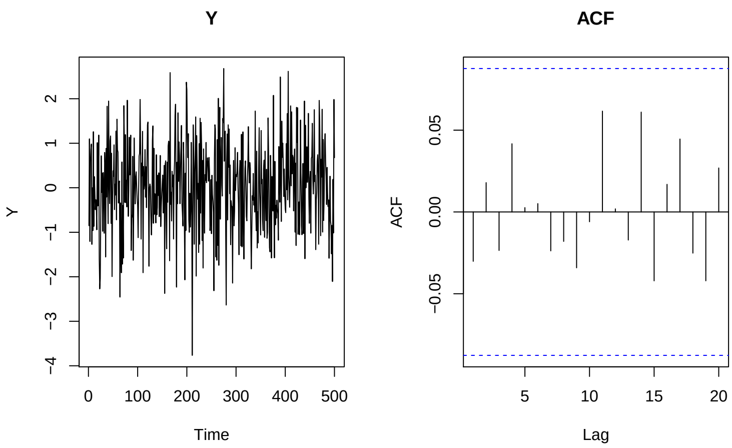
และ ทำให้ตัวสถิติที่ใช้ในการทดสอบ $H_0 : \rho_k = 0$ กับ $H_1 : \rho_k \neq 0$ เปลี่ยนไปเป็น $\frac{\hat{\rho}_k}{\sqrt{(1+2\sum_{i=1}^{k-1}\rho_i^2)/T}}$ แต่อย่างไรก็ตามโปรแกรมสำเร็จรูปส่วนมากยังใช้ $\sqrt{T}\hat{\rho}$ ในการทดสอบอยู่

ตัวอย่างการคำนวณฟังก์ชันสหสัมพันธ์ในตัวเอง (ACF) ของผลได้ตอบแทนในรูปล็อกของ PTT

เราพิจารณาผลได้ตอบแทนในรูปล็อกของหุ้น PTT ต่อจากตัวอย่างที่แล้ว โดยเราสามารถคำนวณค่า autocorrelation ได้ด้วยฟังก์ชัน `acf` โดยเรากำหนดให้ argument `lag.max` คือจำนวนคาบย้อนหลังสูงสุดที่เราพิจารณา โดยเราจะได้โครีโลแกรมของผลได้ตอบแทนในรูปล็อกด้านซ้ายของรูป **??** แต่เราจะสังเกตเห็นได้ว่าแผนภาพดังกล่าวจะรวม ฟังก์ชันสหสัมพันธ์ในตัวเองที่ค่าความล่าเท่ากับศูนย์ (ρ_0) ซึ่งเท่ากับ 1 ไว้เสมอ ซึ่งทำให้เราอ่านค่าอื่นๆได้ยาก ดังนั้นเราสามารถเรียกใช้ฟังก์ชัน `acf` หลังจากเรียก package `library(TSA)` ซึ่งผลปรากฏในรูปด้านขวา โดยที่ lag ที่ 1,6,13,14 มีความสูงเกินจากเส้นประ ($\pm 2/\sqrt{T}$) แสดงว่าค่าสหสัมพันธ์ในตัวเอง ณ ค่าล่าข้างต้นมีความแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่นัยสำคัญเท่ากับ 5 %

```
> acf(lret, lag.max=25)
> library(TSA)
> acf(lret, lag.max=25)
```

รูปที่ 1.6: ข้อมูลเกาซเซียนไวทนอซและฟังก์ชันสหสัมพันธ์ในตัวเอง



1.2.4 การทดสอบพอรทแมนโท (Portmanteau test) สำหรับสหสัมพันธ์ในตัวเอง

ในการวิเคราะห์ข้อมูลทางการเงิน เรามักจะเริ่มต้นด้วยการทดสอบว่าข้อมูลมีสหสัมพันธ์ในตัวเองในหลายๆคาบ (k) ว่าเท่ากับ 0 พร้อมๆกันหรือไม่ Box and Pierce (1970) ได้เสนอตัวสถิติพอรทแมนโทที่คำนวณดังสูตรต่อไปนี้

$$Q^*(m) = T\sum_{k=1}^{m}\hat{\rho}_k^2 \tag{1.28}$$

ในการทดสอบสมมุติฐานหลัก $H_0 : \rho_1 = ... = \rho_m = 0$ กับสมมุติฐานทางเลือก $H_1 : \rho_i \neq 0$ สำหรับค่าสหสัมพันธ์ในตัวเองบางคาบย้อนหลังใน $i \in 1, 2, ..., m$ โดยภายใต้ข้อสมมุติว่า Y_t เป็นลำดับที่แจกแจงเป็นอิสระและเหมือนกัน (identical independent distribution: iid) แล้ว $Q^*(m)$ จะมีการแจกแจงเชิงเส้นกำกับ (asymptotically) เป็นไคกำลังสอง (chi-square)ที่มีองศาเสรีเท่ากับ m หรือเขียนเป็นสัญลักษณ์ χ^2_m

Ljung and Box (1978) ได้ปรับปรุงตัวสถิติ $Q^*(m)$ โดยการเพิ่มพลัง(power)ในการทดสอบเมื่อมีตัวอย่างจำกัดโดยเสนอตัวสถิติ

$$Q(m) = T(T+2)\sum_{k=1}^{m}\frac{\hat{\rho}_k^2}{T-k} \tag{1.29}$$

โดยเราจะตัดสินใจปฏิเสธสมมุติฐานหลัก $H_0 : \rho_1 = ... = \rho_m = 0$ ถ้า $Q(m) > \chi^2_{1-\alpha,m}$ โดยที่ $\chi^2_{1-\alpha,m}$ แสดงถึงควอนไทล์ที่ $100(1-\alpha)$ ของการแจกแจงไคกำลังสองที่มีองศาเสรี m หรือโปรแกรมสำเร็จรูปส่วนมากก็รายงานค่าพี (p-value) เราก็จะปฏิเสธ H_0 ถ้าค่าพีน้อยกว่าระดับนัยสำคัญ α

ในทางปฏิบัติการเลือกค่า m จะส่งผลต่อความสามารถในการทดสอบ ดังนั้นงานศึกษาเชิงประจักษ์ส่วนใหญ่มักจะรายงานค่า Ljung-Box $Q(m)$ หลายๆค่าเช่น $m = 5, 10, 20$ หรืองานวิจัยบางงานพบว่าค่า $m = \ln(T)$ ให้ผลการทดสอบที่ดี

ตัวอย่างการทดสอบพอรทแมนโทของผลได้ตอบแทนของ PTT

ฟังก์ชันที่ใช้ในการทดสอบพอรทแมนโทคือ `Box.test` โดยเราต้องกำหนด argument คือข้อมูลที่ใช้ทดสอบ, จำนวนคาบที่รวมมาทดสอบ (m) และชนิดของการทดสอบ (`type="Ljung"`) สำหรับ Ljung and Box (1978)

```
> Box.test(lret, lag=5, type="Ljung")
  Box-Ljung test
data:  lret
X-squared = 16.2609, df = 5, p-value = 0.006137
```

จากการทดสอบจะเห็นได้ว่า $Q(5) = 16.26$ และค่าพีเท่ากับ 0.0061 ซึ่งเราสามารถปฏิเสธ H_0 ที่ว่าฟังก์ชันสหสัมพันธ์ในตัวเองจากคาบ 1 ถึง $m(= 5)$ มีค่าเท่าศูนย์ แสดงว่า Y_t มีความสัมพันธ์กับตัวเองในคาบใดคาบหนึ่งย้อนหลังไป 1 ถึง 5 คาบ

1.2.5 ตัวดำเนินการขยับไปข้างหลัง

นิยาม* 1.8.** *เรานิยาม* ***ตัวดำเนินการขยับไปข้างหลัง (backshift operator) *โดย*

$$LY_t = Y_{t-1}$$

และสามารถขยายค่ายกกำลังเป็น $L^2Y_t = L(LY_t) = LY_{t-1} = Y_{t-2}$ *ดังนั้น*

$$L^kY_t = Y_{t-k}$$

นอกจากจากนี้การหาผลต่าง (differencing) มีบทบาทสำคัญในการวิเคราะห์อนุกรมเวลา ดังนั้นเรามีเครื่องหมายสำหรับการหาผลต่าง โดยการหาผลต่างอันดับที่หนึ่ง (first difference) สามารถเขียนได้โดย

$$\Delta Y_t = Y_t - Y_{t-1} = (1 - L)Y_t \tag{1.30}$$

โดยที่เราสามารถขยายการหาผลต่างต่อไป เช่นการหาผลต่างอันดับที่สอง (second difference) เท่ากับ

$$\Delta^2 Y_t = (1 - L)^2 Y_t = (1 - 2L + L^2)Y_t$$

นิยาม* 1.9.** ***การหาผลต่างอันดับ d ***(differences of order d*** *สามารถคำนวณได้โดย*

$$\Delta^d = (1 - L)^d \tag{1.31}$$

บทที่ 2

แบบจำลองอนุกรมเวลาเชิงเส้นตรง

ข้อมูลอนุกรมเวลาที่ง่ายที่สุดคือกระบวนการไวท์นอยซ์ (white noise) โดยที่เรามักจะเขียนแทนกระบวนการไวท์นอยซ์ด้วย ตัวแปรสุ่ม (random variable) ε_t โดยส่วนใหญ่เราจะเรียกตัวแปรสุ่ม กระบวนการไวท์นอยซ์ว่า innovation, shock หรือ disturbance term ของข้อมูลอนุกรมเวลา

***นิยาม* 2.1** (กระบวนการไวท์นอยซ์). *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* $k \neq 0$

ทฤษฎีบทการแยกส่วนประกอบของโวลด์ (Wold's decomposition theorem) ฟูลเลอร์ (Fuller, 1996) ระบุว่าอนุกรมเวลานิ่ง y_t สามารถเขียนในรูปกระบวนการเส้นตรงหรือตัวแทนรูปแบบค่าเฉลี่ยเคลื่อนที่ที่มีอันดับเป็นอนันต์ (infinite moving average) และถ้าสามารถเขียน y_t ในรูป

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \tag{2.1}$$

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

ถ้า y_t เป็นข้อมูลที่เป็นอนุกรมนิ่ง (stationary) แล้ว y_t จะต้องมีค่าเฉลี่ยคงที่ $E(y_t) = \mu$ และ $Var(y_t) = \sigma_\varepsilon^2 \sum_{i=1}^{\infty} \psi_i^2$ จะต้องเป็นจำนวนที่จำกัด (finite) ซึ่งจะเกิดขึ้นถ้า $\sum_{i=1}^{\infty} \psi_i^2$ เข้าลู่เข้าหา

(converge) ค่าใดค่าหนึ่ง นอกจากนี้ค่าความแปรปรวนร่วม (covariance)

$$\gamma_l = Cov(y_t, y_{t-l}) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+l}$$

$$\rho_l = \frac{\sum_{j=0}^{\infty} \psi_j \psi_{j+l}}{\sum_{i=1}^{\infty} \psi_i^2}$$

ดังนั้นรูปแบบของสหสัมพันธ์ในตัวเองของอนุกรมเวลานิ่งใดๆจะถูกกำหนดด้วยตัวถ่วงน้ำหนักค่าเฉลี่ยเคลื่อนที่ (moving average weights; ψ_i) นอกจากนี้เราเรียกตัวถ่วงน้ำหนักค่าเฉลี่ยเคลื่อนที่อีกอย่างว่า**การตอบสนองแรงกระตุ้น (impulse responses)**

$$\frac{\partial y_{t+s}}{\partial \varepsilon_t} = \psi_s, \qquad s = 1, 2, 3, ... \tag{2.2}$$

ในกรณีที่อนุกรมเวลาเป็นอนุกรมนิ่งและเออะโกดิก $\lim_{s\to\infty} \psi_s = 0$ และการตอบสนองแรงกระตุ้นในระยะยาว (long-run cumulative impulse responses)มีค่าจำกัด $[\sum_{s=0}^{\infty} \psi_s < \infty]$ นอกจากนี้แผนภาพที่วาด ψ_s กับ s เรียกว่าฟังก์ชันการตอบสนองแรงกระตุ้น (impulse response function (IRF))

2.1 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

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t \tag{2.3}$$

โดยที่ y_t เป็นข้อมูลที่นิ่งและ ε_t เป็นไวท์นอซส์ที่มีค่าเฉลี่ยเท่ากับศูนย์และค่าความแปรปรวนคงที่เท่ากับ σ_ε^2 ในที่นี้ค่าเฉลี่ยของ y_t เท่ากับศูนย์

$$\underbrace{E(y_t)}_{=\mu} = \phi_1 \underbrace{E(y_{t-1})}_{=\mu} + \phi_2 \underbrace{E(y_{t-2})}_{=\mu} + ... + \phi_p \underbrace{E(y_{t-p})}_{=\mu} + \underbrace{E(\varepsilon_t)}_{=0}$$

$$\mu = 0/(1 - \phi_1 - \phi_2 - ... - \phi_p)$$

หากค่าเฉลี่ยของ y_t เท่ากับ μ เราสามารถแสดงสมการ $AR(p)$ ได้เป็น

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

หรือเขียนได้เป็น

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t \tag{2.5}$$

โดยที่ $\phi_0 = \mu(1 - \phi_1 - \phi_2 - ... - \phi_p)$ หรือ $E(y_t) = \mu = \frac{\phi_0}{(1-\phi_1-\phi_2-...-\phi_p)}$

นอกจากนี้เราสามารถเขียนสมการ (**??**) ได้ด้วยเครื่องมือ backshift operator เป็นสมการดังนี้

$$(1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p) y_t = \varepsilon_t \tag{2.6}$$

หรือ

$$\phi(L) y_t = \varepsilon_t \tag{2.7}$$

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

2.1.1 แบบจำลองออโตรีเกรสซีฟอันดับหนึ่ง ($AR(1)$)

เราจะพิจารณาคุณสมบัติของแบบจำลองออโตรีเกรสซีฟจากกระบวนการ $AR(1)$ ซึ่งมีค่าเฉลี่ยเท่ากับศูนย์สามารถแสดงได้โดย

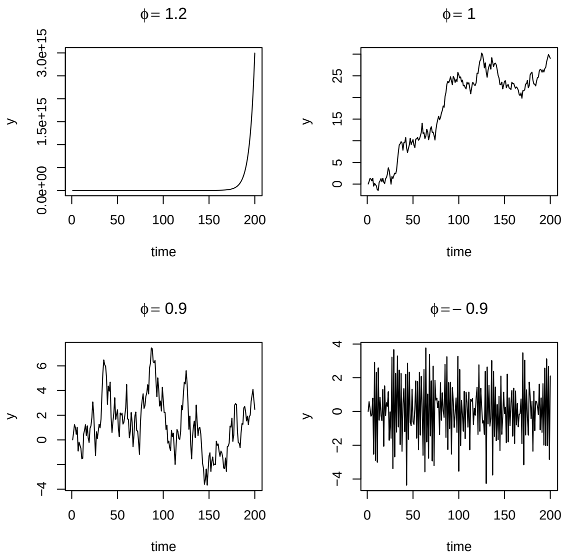
$$y_t = \phi y_{t-1} + \varepsilon_t \tag{2.8}$$

จากสมการข้างต้นหากเราเริ่มพิจารณากรณีที่ $y_0 = 0$ และให้ ε_t มีการแจกแจงแบบ $N(0,1)$ เราจะได้การเดินทางตามเวลา (time path) ของ y_t สำหรับค่าสัมประสิทธิ์(ϕ)ที่ต่างกันดังที่แสดงในรูป **??** จากรูป **??** มุมบนซ้ายซึ่งค่า $|\phi| > 1$ ในกรณีดังกล่าว y_t จะเพิ่มขึ้นอย่างต่อเนื่องในอัตราที่รุนแรง(explosive)โดยที่คาบที่ $t = 200$ ค่า $y_t > 10^{15}$ ในขณะที่มุมบนขวาซึ่งค่า $|\phi| = 1$ ค่า y_t จะขยับไปจุดหนึ่งๆและคงค้างอยู่บริเวณนั้นสักพักแล้วก็ไม่เดินทางกลับมาที่จุดเริ่มต้น จะเห็นได้ว่าค่าเฉลี่ยของแต่ละช่วงมีค่าที่แตกต่างกันและค่าความแปรปรวนมีแนวโน้มเพิ่มขึ้นเรื่อย ในกรณีนี้เราจะเรียกกระบวนการนี้ว่า "random walk" รูปข้างบนทั้งสองเป็นตัวอย่างของกระบวนการ $AR(1)$ ที่ไม่นิ่ง(non-stationary) ในขณะที่รูปล่างทั้งสองรูป $|\phi| < 1$ ค่า y_t จะขยับอยู่รอบๆค่าศูนย์แสดงถึงการเป็นกระบวนการนิ่ง(stationary)

จากสมการ **??** เราสามารถแทนค่า y_t ในอดีตไปเรื่อยๆแบบเวียนเกิด (recursive) k ครั้งดังนี้

$$\begin{aligned}
y_t &= \phi y_{t-1} + \varepsilon_t = \phi(\phi y_{y-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \phi^2 y_{t-2} + \phi\varepsilon_{t-1} + \varepsilon_t \\
&\vdots \\
&= \phi^k y_{t-k} + \sum_{j=0}^{k-1} \phi^j \varepsilon_{t-j}
\end{aligned}$$

รูปที่ 2.1: การจำลองกระบวนการ $AR(1)$ ที่ค่าสัมประสิทธิ์(ϕ)ที่ต่างกัน



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

$$y_t = \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + ... = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} \tag{2.9}$$

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

ค่าคาดหมายหรือค่าเฉลี่ยของ y_t

หากใส่ค่าคาดหมาย (take expectation) ทั้งสองข้างของสมการ **??** จะได้ข้อสรุปว่ากระบวน-การ $AR(1)$ ดังที่ได้แสดงในสมการ **??** มีค่าเฉลี่ยเท่ากับศูนย์

$$E(y_t) = E\left[\varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + ...\right]$$

เนื่องจาก $\varepsilon_t, \varepsilon_{t-1}, ...$ เป็นอิสระต่อกัน [ทวนความจำ: หาก X และ Z เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน $E(X + Z) = E(X) + E(Z)$] ดังนั้น

$$\begin{aligned} E(y_t) &= E(\varepsilon_t) + E(\phi\varepsilon_{t-1}) + E(\phi^2\varepsilon_{t-2}) + ... \\ &= \underbrace{E(\varepsilon_t)}_{=0} + \phi \underbrace{E(\varepsilon_{t-1})}_{=0} + \phi^2 \underbrace{E(\varepsilon_{t-2})}_{=0} + ... = 0 \end{aligned}$$

จะเห็นได้ว่าค่าเฉลี่ยของ y_t เท่ากับศูนย์

ค่าความแปรปรวนของ y_t

เนื่องจาก $E(y_t) = 0$ ค่าความแปรปรวน $Var(y_t) = E[y_t - E(y_t)]^2$ จะเท่ากับ $Var(y_t) = E(y_t^2)$ หากเราแทนค่า y_t จากสมการ **??** 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

$$
\begin{aligned}
E(y_t^2) &= E[(\varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + ...)(\varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + ...)] \\
&= E[\varepsilon_t\varepsilon_t + \varepsilon_t\phi\varepsilon_{t-1} + \varepsilon_t\phi^2\varepsilon_{t-2} + ... + \phi\varepsilon_{t-1}\varepsilon_t + \phi\varepsilon_{t-1}\phi\varepsilon_{t-1} + \phi\varepsilon_{t-1}\phi^2\varepsilon_{t-2} + ... \\
&\quad \phi^2\varepsilon_{t-2}\varepsilon_t + \phi^2\varepsilon_{t-2}\phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2}\phi^2\varepsilon_{t-2} + ...] \\
&= E[\varepsilon_t^2 + \phi\varepsilon_t\varepsilon_{t-1} + \phi^2\varepsilon_t\varepsilon_{t-2} + ... + \phi\varepsilon_{t-1}\varepsilon_t + \phi^2\varepsilon_{t-1}^2 + \phi^3\varepsilon_{t-1}\varepsilon_{t-2} + ... \\
&\quad + \phi^2\varepsilon_{t-2}\varepsilon_t + \phi^3\varepsilon_{t-2}\varepsilon_{t-1} + \phi^4\varepsilon_{t-2}^2 + ...] \\
&= \underbrace{E(\varepsilon_t^2)}_{=\sigma_\varepsilon^2} + \phi^2\underbrace{E(\varepsilon_{t-1}^2)}_{=\sigma_\varepsilon^2} + \phi^4\underbrace{E(\varepsilon_{t-2}^2)}_{=\sigma_\varepsilon^2} + ... + \phi\underbrace{E(\varepsilon_t\varepsilon_{t-1})}_{=0} + \phi^2\underbrace{E(\varepsilon_t\varepsilon_{t-2})}_{=0} + \phi\underbrace{E(\varepsilon_{t-1}\varepsilon_t)}_{=0} + ...
\end{aligned}
$$

$$
E(y_t^2) = (1 + \phi^2 + \phi^4 + ...)\sigma_\varepsilon^2 = \frac{\sigma_\varepsilon^2}{(1-\phi^2)} \equiv \gamma_0 \tag{2.10}
$$

จากบรรทัดที่ 6 จะเห็นได้ว่าค่าคาดหมายของพจน์ที่คูณกันที่มีตัวห้อยต่างกัน (cross terms) จะเท่ากับศูนย์ นอกจากนี้เราจะใช้สัญลักษณ์ γ_0 แทนค่าความแปรปรวน

ค่าความแปรปรวนร่วมในตัวเองและฟังก์ชันสหสัมพันธ์ร่วมในตัวเอง

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

$$
\begin{aligned}
\gamma_l = Cov(y_t, y_{t-l}) &= E[(y_t - E(y_t))(y_{t-l} - E(y_{t-l}))] = E(y_t y_{t-l}) \\
&= E[(\varepsilon_t + ... + \phi^l\varepsilon_{t-l} + \phi^{l+1}\varepsilon_{t-l-1} + \phi^{l+2}\varepsilon_{t-l-2} + ...)(\varepsilon_{t-l} + \phi\varepsilon_{t-l-1} + \phi^2\varepsilon_{t-l-2} + ...)] \\
&= E[\text{cross terms} + ... + \phi^l\varepsilon_{t-l}\varepsilon_{t-l} + \phi^l\varepsilon_{t-l}\phi\varepsilon_{t-l-1} + \phi^l\varepsilon_{t-l}\phi^2\varepsilon_{t-l-2} + ... \\
&\quad + \phi^{l+1}\varepsilon_{t-l-1}\varepsilon_{t-l} + \phi^{l+1}\varepsilon_{t-l-1}\phi\varepsilon_{t-l-1} + \phi^{l+1}\varepsilon_{t-l-1}\phi^2\varepsilon_{t-l-2} + ... \\
&\quad + \phi^{l+2}\varepsilon_{t-l-2}\varepsilon_{t-l} + \phi^{l+2}\varepsilon_{t-l-2}\phi\varepsilon_{t-l-1} + \phi^{l+2}\varepsilon_{t-l-2}\phi^2\varepsilon_{t-l-2} + ... + \text{cross terms}] \\
&= E\left[\phi^l\varepsilon_{t-l}\varepsilon_{t-l} + \phi^{l+1}\varepsilon_{t-l-1}\phi\varepsilon_{t-l-1} + \phi^{l+2}\varepsilon_{t-l-2}\phi^2\varepsilon_{t-l-2} + ... + \text{cross terms}\right] \\
&= \phi^l\underbrace{E(\varepsilon_{t-l}^2)}_{=\sigma_\varepsilon^2} + \phi^{l+2}\underbrace{E(\varepsilon_{t-l-1}^2)}_{=\sigma_\varepsilon^2} + \phi^{l+4}\underbrace{E(\varepsilon_{t-l-2}^2)}_{=\sigma_\varepsilon^2} + ... + \underbrace{E(\text{cross terms})}_{=0}
\end{aligned}
$$

$$
\gamma_l = \phi^l(1 + \phi^2 + \phi^4 + ...)\sigma_\varepsilon^2 = \frac{\phi^l\sigma_\varepsilon^2}{(1-\phi^2)} \tag{2.11}
$$

และฟังก์ชันสหสัมพันธ์ร่วมในตัวเอง (ACF) เท่ากับ

$$\rho_l = \frac{\gamma_l}{\gamma_0} = \frac{\frac{\phi^l \sigma_\varepsilon^2}{(1-\phi^2)}}{\frac{\sigma_\varepsilon^2}{(1-\phi^2)}} = \phi^l \tag{2.12}$$

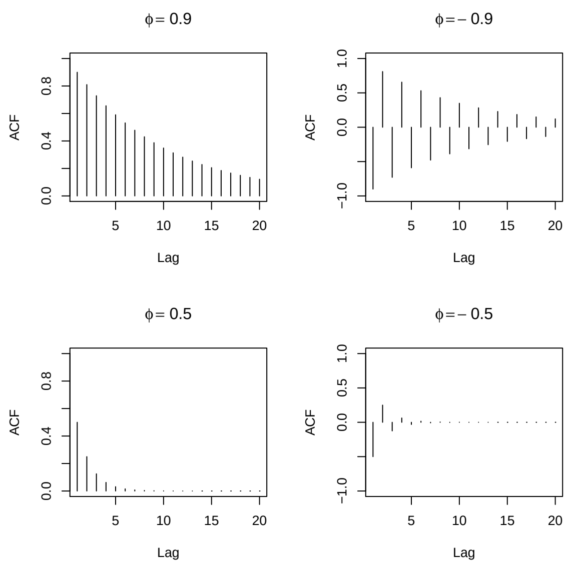
หากลองแทนค่า ϕ ด้วยค่าเท่ากับ $0.9, -0.9, 0.5, -0.5$ จะได้ฟังก์ชันสหสัมพันธ์ร่วมในตัวเองดังตารางต่อไปนี้

ตารางที่ 2.1: ค่าสหสัมพันธ์ร่วมในตัวเองของกระบวนการ $AR(1)$ ที่ค่าสัมประสิทธิ์(ϕ)ที่ต่างกัน

ϕ \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

จะเห็นได้ว่าค่าสหสัมพันธ์ร่วมในตัวเองของกระบวนการ $AR(1)$ ในรูปค่าสมบูรณ์มีค่าลดลงเรื่อยๆ โดยที่หากค่า ϕ เป็นลบ ACF จะมีลักษณะสวิงไปทางบวกและลบสลับกัน นอกจากนี้ค่า $|\phi|$ ที่เข้าใกล้หนึ่งจะมี ACF ที่ลดลงช้ากว่ากรณีที่ $|\phi|$ ที่เข้าใกล้ศูนย์ โดยสามารถดูได้จากรูป **??**

รูปที่ 2.2: ค่าสหสัมพันธ์ร่วมในตัวเองของกระบวนการ $AR(1)$ ที่ค่าสัมประสิทธิ์(ϕ)ที่ต่างกัน



การพิจารณาเงื่อนไขความเป็นอนุกรมนิ่งจากพหุนามออโตรีเกรสซีฟ

เราสามารถเขียนกระบวนการ $AR(1)$ ในรูปของพหุนามออโตรีเกรสซีฟได้เป็น

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

โดยที่เราสามารถเขียนสมการช่วย (auxiliary equation) สำหรับสมการ **??** ได้เป็น

$$(1-\phi m) = 0 \tag{2.14}$$

โดยที่รากของพหุนามออโตรีเกรสซีฟจะเท่ากับ $m = 1/\phi$ โดยที่เราทราบว่าเงื่อนไขที่กระบวนการ $AR(1)$ จะเป็นกระบวนการนิ่ง $|\phi|$ จะต้องมีค่าน้อยกว่าหนึ่ง ดังนั้น**ค่าสัมบูรณ์ของรากของพหุนามออโตรีเกรสซีฟจะต้องมีค่ามากกว่าหนึ่ง** $[|m| = |1/\phi| > 1]$

ใน ขณะ เดียวกัน เรา สามารถ เขียน สมการ ลักษณะ เฉพาะ(characteristic equation)สำหรับกระบวนการ $AR(1)$ จาก $z^{-1}(z-\phi)y_t = \varepsilon_t$ ได้เป็น

$$(z-\phi) = 0 \tag{2.15}$$

จะเห็นได้ว่ารากของสมการลักษณะเฉพาะจะเท่ากับ ϕ $[z = \phi]$ ดังนั้นเราจะได้เงื่อนไขกระบวนการนิ่งว่า **ค่าสัมบูรณ์ของรากของสมการลักษณะเฉพาะจะต้องมีค่าน้อยหนึ่ง** $[|z| = |\phi| < 1]$

ตัวอย่างที่ 2.1: กระบวณการนิ่ง

จงพิจารณาว่ากระบวนการต่อไปนี้เป็นกระบวนการนิ่งหรือไม่

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

เราสามารถเขียน AR(1) ในรูปทั่วไปเช่นสมการที่ **??** หรือ **??** โดยที่ในกรณีสมการ **??**

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

ถ้าเราต้องการหา $E(y_t)$ เราสามารถทำได้โดยการใส่ค่าคาดหมายทั้งสองข้างและใช้ข้อสมมุติที่ว่า

y_t เป็น stationary ($E(y_t) = E(y_{t-1})$) จะได้

$$\begin{aligned} E(y_t) - \mu &= \phi \underbrace{E(y_{t-1})}_{=E(y_t)} - \phi\mu + \underbrace{E(\varepsilon_t)}_{=0} \\ E(y_t) - \mu &= \phi E(y_t) - \phi\mu \\ E(y_t) = \frac{1-\phi}{1-\phi}\mu &= \mu \end{aligned}$$

ในกรณีสมการที่ **??**

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

ถ้าเราต้องการหา $E(y_t)$ เราสามารถทำได้โดยการใส่ค่าคาดหมายทั้งสองข้างและใช้ข้อสมมุติที่ว่า y_t เป็น stationary ($E(y_t) = E(y_{t-1})$) จะได้

$$\begin{aligned} E(y_t) &= \phi_0 + \phi_1 \underbrace{E(y_{t-1})}_{=E(y_t)} + \underbrace{E(\varepsilon_t)}_{=0} \\ (1-\phi_1)E(y_t) &= \phi_0 \\ E(y_t) &= \frac{\phi_0}{1-\phi_1} \end{aligned}$$

ดังนั้นเราจะได้ว่า $\mu = \frac{\phi_0}{1-\phi_1}$ เราสามารถขยายการวิเคราะห์ดังกล่าวไปยัง $AR(p)$

2.1.2 แบบจำลอง $AR(2)$

แบบจำลอง $AR(2)$ สามารถเขียนอธิบายได้ด้วยสมการ

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t \tag{2.16}$$

เงื่อนไขการเป็นกระบวนการนิ่ง

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

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

และพิจารณาสมการช่วย

$$\phi(m) = 1 - \phi_1 m - \phi_2 m^2 = 0 \tag{2.17}$$

โดยที่รากของสมการพหุนามออโตรีเกรสซีฟที่อยู่ในรูปสมการกำลังสอง (quadratic) จะมีด้วยกันสองจำนวน(m_1, m_2)เท่ากับ $\frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}$ ซึ่งรากดังกล่าวสามารถเป็นจำนวนเชิงซ้อน โดยที่

เงื่อนไขที่กระบวนการ $AR(2)$ จะเป็นกระบวนการนิ่งคือ รากของสมการพหุนามออโตรีเกรสซีฟจะต้องมากกว่าหนึ่ง (หรือมอดุลัส (modulus) มากกว่าหนึ่งในกรณีจำนวนเชิงซ้อน) ซึ่งจะเกิดขึ้นเมื่อ

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

ตัวอย่างที่ 2.2 เงื่อนไขอนุกรมเวลานิ่งกรณีออโตรีเกรสซีฟ

กำหนดให้ $y_t = 0.5y_{t-1} - 0.8y_{t-2} + \varepsilon_t$ เราสามารถพิจารณาสมการพหุนามออโตรีเกรสซีฟ

$$(1 - 0.5m + 0.8m^2) = 0$$

จะเห็นได้ว่า $\phi_1 = 0.5$ และ $\phi_2 = -0.8$ เป็นไปตามเงื่อนไข $\phi_1 + \phi_2 < 1, \quad \phi_2 - \phi_1 < 1, \quad |\phi_2| < 1$ ดังนั้นกระบวนการ y_t เป็นกระบวนการนิ่ง นอกจากนี้เราสามารถพิจารณารากของสมการพหุนามออโตรีเกรสซีฟ โดยเราจะเห็นได้ว่า

$$\begin{aligned} m_1, m_2 = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2} &= \frac{0.5 \pm \sqrt{0.5^2 + 4(-0.8)}}{-2(-0.8)} \\ &= \frac{0.5}{1.6} \pm \frac{\sqrt{-2.95}}{1.6} = \frac{0.5}{1.6} \pm \frac{\sqrt{2.95}\sqrt{-1}}{1.6} \\ &= 0.3125 \pm 1.073473i \end{aligned}$$

โดยที่รากดังกล่าวสามารถคำนวณด้วย R โดยใช้คำสั่ง `polyroot(c(1,-0.5,0.8))`

ค่ามอดุลัสของจำนวนเชิงซ้อน $m_1 = a + bi$ จะเท่ากับ $\sqrt{a^2 + b^2}$ ซึ่งในกรณีนี้เท่ากับ `1.118034` ซึ่งมีค่ามากกว่าหนึ่ง เราสามารถคำนวณมอดุลัสได้โดยใช้คำสั่ง `Mod(arg)` โดยที่ `arg` คือจำนวนเชิงซ้อน

คุณลักษณะของกระบวนการ $AR(2)$

เราสามารถพิจารณาค่าเฉลี่ยของกระบวนการในสมการ **??** และพบว่ามีค่าเฉลี่ยเท่ากับศูนย์ สำหรับฟังก์ชันความแปรปรวนร่วมเราสามารถพิจารณาได้โดยคูณสมการ **??** ทั้งสองข้างด้วย y_{t-k} แล้วใส่ค่าคาดหมาย (take expectation) เราจะได้สมการ

$$\begin{aligned} E(y_t y_{t-k}) &= \phi_1 E(y_{t-1} y_{t-k}) + \phi_2 E(y_{t-2} y_{t-k}) + \underbrace{E(\varepsilon_t y_{t-1})}_{=0} \\ \gamma_k &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} \end{aligned} \tag{2.18}$$

เนื่องจากหากเราเขียน y_t ในรูปของ infinite MA จะได้ $E(\varepsilon_t y_{t-1}) = E(\varepsilon_t\varepsilon_{t-1}+\varepsilon_t\psi_1\varepsilon_{t-2}+...) = 0$ และหากหารสมการ (**??**) ด้วย γ_0 จะได้

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} \tag{2.19}$$

เราเรียกสมการ **??** และ **??** ว่าสมการ Yule-Walker จากสมการที่ **??** 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 $k > 2$

2.1.3 แบบจำลอง $AR(p)$

แบบจำลอง $AR(p)$ ที่ได้มีปรับเอาค่าเฉลี่ยออกสามารถเขียนในรูป

$$y_t - \mu = \phi_1(y_{t-1}-\mu) + \phi_2(y_{t-2}-\mu) + ... + \phi_p(y_{t-p}-\mu) + \varepsilon_t \tag{2.20}$$

หรือเขียนในรูปตัวดำเนินการขยับไปข้างหลัง

$$\phi(L)(y_t - \mu) = \varepsilon_t \tag{2.21}$$

โดยที่ $\phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$ หรือในรูปสมการถดถอยในตัวเอง

$$\phi(L)y_t = c + \varepsilon_t \tag{2.22}$$

โดยที่เราสามารถแสดงเงื่อนไขที่ $AR(p)$ จะเป็นอนุกรมนิ่งและอะโกดิกได้ด้วยการพิจารณาราก ของสมการพหุนามออโตรีเกรสซีฟ

$$\phi(m) = 1 - \phi_1 m - \phi_2 m^2 - ... - \phi_p m^p = 0$$

หากค่าสัมบูรณ์ของรากทุกตัวของสมการพหุนามออโตรีเกรสซีฟมีค่ามากกว่าหนึ่ง (หรือมอดุลัสมากกว่าหนึ่งในกรณีจำนวนเชิงซ้อน) อนุกรมเวลา y_t จะเป็นอนุกรมเวลานิ่ง

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

หากเราสมมุติให้ค่าเฉลี่ย(μ)เท่าศูนย์และกระบวนการนิ่ง หากเราคูณสมการ **??** ด้วย y_{t-k} แล้วใส่ค่าคาดหมายทั้งสองข้างของสมการและหารด้วยค่าความแปรปรวนในตัวเอง (γ_0) เราจะได้

สมการแสดงความสัมพันธ์แบบเวียนเกิด

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + ... + \phi_p\rho_{k-p}, \tag{2.23}$$

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

$$\begin{aligned}
\rho_1 &= \phi_1 + \phi_2\rho_1 + \phi_3\rho_2 + ... + \phi_p\rho_{p-1} \\
\rho_2 &= \phi_1\rho_1 + \phi_2 + \phi_3 + ... + \phi_p\rho_{p-2} \\
&\vdots \\
\rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \phi_3\rho_{p-3} + ... + \phi_p
\end{aligned} \tag{2.24}$$

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 **??** ด้วย y_t และใส่ค่าคาดหมาย เราจะได้

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + ... + \phi_p\gamma_p + \sigma^2 \tag{2.25}$$

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

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

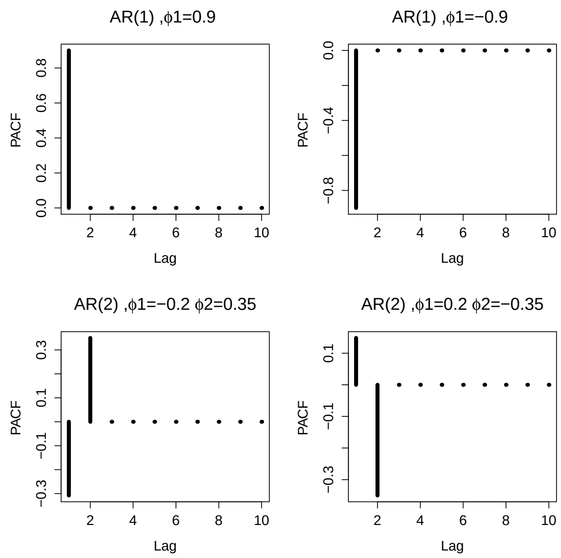
2.1.4 ฟังก์ชัน สห สัมพันธ์ ใน ตัว เอง บาง ส่วน (Partial Autocorrelation Function; PACF)

ฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วนเป็นเครื่องมือที่เป็นประโยชน์ในการระบุแบบจำลอง $AR(p)$ โดยที่ฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วนมีพื้นฐานจากการประมาณค่าแบบจำลอง AR เป็นลำดับดังนี้

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

โดยที่ $z_t = y_t - \mu$ คือข้อมูลที่ได้กำจัดค่าเฉลี่ยแล้ว (demeaned) เราเรียกค่าสัมประสิทธิ์ ϕ_{jj} สำหรับ $j = 1, 2, ..., p$ (ค่าสัมประสิทธิ์สุดท้ายของแต่ละสมการ) ว่า*ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน* ในกรณีที่เราพิจารณาแบบจำลอง $AR(1)$ ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนตัวแรก ϕ_{11} จะไม่เท่ากับศูนย์ ส่วนสัมประสิทธิ์ที่เหลือจะเท่ากับศูนย์ เช่นเดียวกันหากเราพิจารณาแบบจำลอง $AR(2)$ ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนตัวแรกและตัวที่สอง (ϕ_{11} และ ϕ_{22}) จะไม่เท่ากับศูนย์ ส่วนสัมประสิทธิ์ที่เหลือ (ϕ_{jj} สำหรับ $j > 2$) จะเท่ากับศูนย์ โดยสรุปแล้ว สำหรับแบบจำลอง $AR(p)$ ใดๆ ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน p ตัวแรกจะไม่เท่ากับศูนย์ และสัมประสิทธิ์ที่เหลือจะเท่ากับศูนย์ โดยจากตัวอย่างใดๆ เราสามารถประมาณค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนจากตัวอย่าง (sample PACF) ด้วยวิธีการกำลังสองน้อยที่สุดทีละสมการ และเก็บค่าประมาณสัมประสิทธิ์ $\hat{\phi}_{jj}$ เมื่อนำค่า $\hat{\phi}_{jj}$ มาวาดกราฟกับ j เราจะได้ฟังก์ชันสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนจากตัวอย่าง (sample PACF) ซึ่งสามารถคำนวณได้โดยใช้ฟังก์ชัน `pacf` ใน R

รูปที่ 2.3: ฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วนของข้อมูลจำลอง $AR(1)$ และ $AR(2)$



ตัวอย่างที่ 2.3: การสร้างแบบจำลองสำหรับความแตกต่างระหว่างอัตราดอกเบี้ยด้วย AR(p)

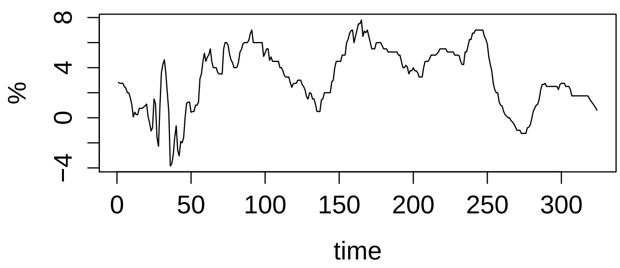
ในตัวอย่างนี้เราจะพิจารณาความแตกต่างระหว่างอัตราดอกเบี้ย MLR ของไทย[1] และอัตราดอกเบี้ยลูกค้าชั้นดีของสหรัฐอเมริกา[2] รายเดือนระหว่างเดือนมกราคม 1978 ถึงธันวาคม 2004 ซึ่งอยู่ในไฟล์ `mlr.csv` โดยที่คอลัมน์ที่หนึ่งเป็นเดือน คอลัมน์ที่สองเป็นอัตราดอกเบี้ยของไทย คอลัมน์

[1] ข้อมูลจากธนาคารแห่งประเทศไทย
[2] ข้อมูลจากธนาคารกลางแห่งสหรัฐอเมริกา

ที่สามเป็นอัตราดอกเบี้ยของสหรัฐอเมริกา และคอลัมน์สี่เป็นส่วนต่าง `diff_th_us` เราจะนำเข้าข้อมูล วาดกราฟ(รูปที่ **??**)

```
>
    int<-read.csv("https://raw.githubusercontent.com/chaleampong/EC435/master/mlr.csv",
    header = TRUE)
> head(int)
> ts.plot(int$diff_th_us, ylab ="%", xlab="time")
```

รูปที่ 2.4: ความแตกต่างระหว่างอัตราดอกเบี้ยของไทยและสหรัฐอเมริกา



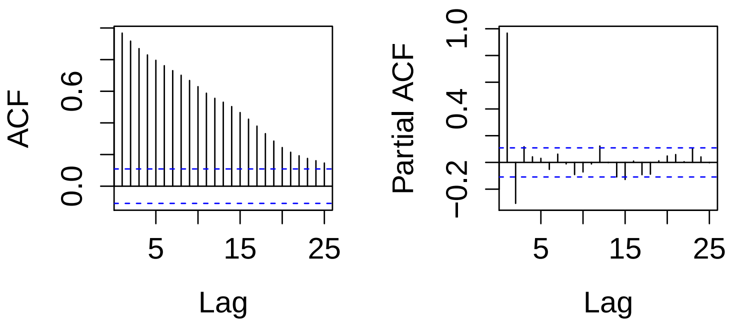
เราพิจารณา ACF และ PACF จะได้รูปที่ **??** เมื่อพิจารณาจากรูป ACF จะเห็นได้ว่าผลต่างของอัตราดอกเบี้ย (y_t) มีความสัมพันธ์กับตัวเองในอดีตอย่างมีนัยสำคัญ จากรูป ACF จะพอที่จะเดาได้ว่า y_t น่าจะมีลักษณะคล้ายกับกระบวนการ $AR(p)$ ซึ่งเราสามารถพิจารณาได้ด้วยกราฟ PACF ในด้านขวาจะเห็นได้ PACF ต่างจากศูนย์อย่างมีนัยสำคัญจนกระทั่งถึงค่าล้าที่ 3 ดังนั้นเราน่าจะใช้แบบจำลอง $AR(3)$ ในการประมาณค่าผลต่างของอัตราดอกเบี้ย

```
> acf(int$diff_th_us)
> pacf(int$diff_th_us)
```

2.1.5 การประมาณค่าและตรวจสอบแบบจำลอง

การประมาณค่าพารามิเตอร์ของแบบจำลอง $AR(p)$: สามารถดำเนินการได้โดยใช้วิธีการกำลังสองน้อยที่สุด โดยตัวแปรอิสระคือตัวแปร y_t ที่ค่าล่า $1, 2, ..., p$ โดยที่ตัวประมาณค่า $\hat{\sigma}^2 = \frac{\sum_{t=p+1}^{T} \hat{\varepsilon}_t^2}{T-2p-1}$

รูปที่ 2.5: ACF และ PACF ของความแตกต่างระหว่างอัตราดอกเบี้ยของไทยและสหรัฐอเมริกา



การประมาณค่าความควรจะเป็นสูงที่สุด (Maximum Likelihood 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) ดังนั้นหากเราเก็บข้อมูล $y_1, ..., y_T$ แล้วฟังก์ชันค่าความควรจะเป็นสามารถเขียนได้ดังนี้

$$L(\mu, \phi, \sigma^2) = f(y_1)f(y_2|y_1)f(y_3|y_2)\cdots f(y_T|y_{T-1}) \tag{2.26}$$

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

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

และสามารถเขียนฟังก์ชันค่าความควรจะเป็นได้เป็น

$$L(\mu, \phi, \sigma^2) = f(y_1)\Pi_{t=2}^{T} f_\varepsilon[(y_t - \mu) - \phi(y_{t-1} - \mu)] \tag{2.27}$$

โดยที่เราจำเป็นต้องหาการแจกแจงของ $f(y_1)$ เราทราบว่าเราสามารถเขียน y_1 ใดๆในรูปของ MA ที่มีอันดับเป็นอนันต์ $y_1 = \mu + \sum_{j=0}^{\infty} \phi^j \varepsilon_{1-j}$ ซึ่งจะเห็นได้ว่ามีการแจกแจงเป็นแบบปกติที่มีค่า

เฉลี่ย μ และค่าความแปรปรวนเท่ากับ $\sigma^2/(1-\phi^2)$ ดังนั้นฟังก์ชันค่าความควรจะเป็นจะเท่ากับ

$$L(\mu,\phi,\sigma^2) = (2\pi\sigma^2)^{-1/2}(1-\phi^2)^{-1/2}\exp\left(\frac{-(y_1-\mu)^2}{2\sigma^2/(1-\phi^2)} + (2\pi\sigma^2)^{-(T-1)/2}\right)\times$$
$$\exp\left(\frac{\sum_{t=2}^{T}[(y_t-\mu)-\phi(y_{t-1}-\mu)]}{2\sigma^2}\right) \tag{2.28}$$

การหาประมาณค่าโดยหาค่าพารามิเตอร์ที่ทำให้ log ของสมการ (**??**) สูงที่สุดเรียกว่า การประมาณค่าความควรจะเป็นสูงสุดแบบแม่นตรง (exact maximum likelihood estimation) โดยที่ในกรณีนี้เราไม่สามารถหาค่าสูงสุดได้ด้วยการวิเคราะห์หาเงื่อนไขจำเป็น เราต้องใช้กระบวนการด้วยเครื่องคอมพิวเตอร์ในการหาค่าสูงสุด

ในกรณีแบบจำลอง AR เราสามารถละเลยค่าเริ่มต้นซึ่งก่อนให้เกิดความไม่เป็นเส้นตรงได้ โดยภายใต้เงื่อนไขสมมุติให้ค่าเริ่มต้นเท่ากับศูนย์ เราจะได้ค่าความควรจะเป็นแบบมีเงื่อนไข (conditional likelihood)

$$L(\mu,\phi,\sigma^2) = \Pi_{t=2}^{T} f_\varepsilon[(y_t-\mu)-\phi(y_{t-1}-\mu)] \tag{2.29}$$
$$= (2\pi\sigma^2)^{-(T-1)/2}\exp\left(-\frac{\sum_{t=2}^{T}[(y_t-\mu)-\phi(y_{t-1}-\mu)]^2}{2\sigma^2}\right) \tag{2.30}$$

ตัวประมาณค่าที่ได้จากการแสวงหาค่าสูงสุดของสมการ (**??**) เรียกว่าการประมาณค่าความควรจะเป็นสูงสุดแบบมีเงื่อนไข (conditional MLE) หรือเราสามารถเขียนฟังก์ชันในสมการ (**??**) ใหม่เป็น

$$L(\mu,\phi,\sigma^2) = (2\pi\sigma^2)^{-(T-1)/2}\exp\left(-\frac{\sum_{t=2}^{T}\varepsilon_t^2}{2\sigma^2}\right) \tag{2.31}$$

หรือเราสามารถพิจารณาฟังก์ชัน log-likelihood

$$\ln L(\mu,\phi,\sigma^2) = \frac{-(T-1)}{2}\ln(2\pi) - \frac{(T-1)}{2}\ln(\sigma^2) - \frac{\sum_{t=2}^{T}\varepsilon_t^2}{2\sigma^2} \tag{2.32}$$

โดยที่ค่าพารามิเตอร์ที่ให้ค่าสูงสุดของ $\ln L$ จะให้ค่าต่ำสุดของ $S(\mu,\theta) = \frac{\sum_{t=2}^{T}\varepsilon_t^2}{2\sigma^2}$ ซึ่งเราเรียกว่า conditional sum of squares (CSS) ดังเราจะเรียกการประมาณค่าด้วยวิธี conditional MLE ว่า การประมาณค่าด้วย conditional sum of squares

ในกรณีที่เราประมาณค่าด้วยการหาค่าสูงสุดของสมการ **??** เราสามารถตัวประมาณค่าได้

ด้วยการหาเงื่อนไขอันดับหนึ่งสำหรับค่าสูงสุด (first order condition)

$$\frac{\partial \ln L}{\partial \phi} = -\frac{1}{2\sigma^2} 2 \sum_{t=2}^{T} \varepsilon_t(-\phi(y_{t-1} - \mu)) \tag{2.33}$$

$$\frac{\partial \ln L}{\partial \sigma^2} = \frac{-(T-1)}{2} \frac{1}{\sigma^2} + \frac{\sum_{t=2}^{T} \varepsilon_t^2}{2\sigma^2} \tag{2.34}$$

จากเงื่อนไขแรก

$$\sum_{t=2}^{T} \varepsilon_t(\hat{\phi}(y_{t-1} - \mu)) = 0$$

$$\sum_{t=2}^{T} [(y_t - \mu) - \hat{\phi}(y_{t-1} - \mu)](\hat{\phi}(y_{t-1} - \mu)) = 0$$

หลังจากที่จัดรูปใหม่จะได้

$$\hat{\phi} = \frac{\sum_{t=2}^{T} (y_t - \hat{\mu})(y_{t-1} - \hat{\mu})}{\sum_{t=2}^{T} (y_t - \hat{\mu})^2}$$

โดยที่ค่า $\hat{\mu} = \frac{\sum_{t=2}^{T} y_t}{(T-1)}$ และ จากเงื่อนไขที่สอง $\hat{\sigma}^2 = \frac{\sum_{t=2}^{T} \hat{\varepsilon}_t}{T-2}$ จะเห็นได้ว่าตัวประมาณค่าที่ได้จาก CSS เป็นตัวประมาณค่า OLS ด้วย

การประเมินความถูกต้องของแบบจำลอง (Model Checking)

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

เราอาจจะเริ่มจากการวาดแผนภาพ $\hat{\varepsilon}_t$ และฟังก์ชันสหสัมพันธ์ในตัวเองของ $\hat{\varepsilon}_t$ เพื่อดูว่าอนุกรมดังกล่าวยังมีความขึ้นอยู่ต่อกันหรือไม่ โดยที่ภายใต้สมมุติฐานหลักที่ว่า $\varepsilon_t \sim N(0,1)$ เราสามารถสรุปได้ว่าสหสัมพันธ์ในตัวเองจะมีการแจกแจงเป็น $N(0, 1/T)$ และค่าที่ใช้ในการทดสอบคือ $\pm 2/\sqrt{T}$

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

ตัวอย่าง

ต่อเนื่องจากตัวอย่างข้างต้น เราต้องการประมาณค่าแบบจำลอง $AR(3)$ เราสามารถประมาณค่าได้โดยใช้คำสั่ง `arima` โดยที่ arg ที่เราต้องใส่คือ (1) ชื่ออนุกรมเวลาซึ่งในที่นี้คือ `int$diff_th_us` (2) อันดับของอนุกรม `order=c(p,d,q)` โดยที่ p คืออันดับของออโตรีเกรสซีฟ ส่วน d และ q คืออันดับของ integration และ moving average ซึ่งเราจะพูดถึงในหัวข้อต่อไป ในหัวข้อนี้เราจะกำหนดให้ $d = 0, q = 0$ โดยที่เราจะเก็บแบบจำลองไว้ใน object ชื่อว่า $m1$

```
> 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 $\phi_1 = 1.3294, \phi_2 = -0.4986, \phi_3 = 0.1334, \mu(= intercept) = 3.0553$ และมีค่า standard errors อยู่ในบรรทัด `s.e.` ในโปรแกรม R ค่า `intercept` ที่ได้คือ μ มิใช่จุดตัดแกนในสมการที่ **??** ดังนั้นถ้าต้องการหากเราต้องการค่า $c = \mu(1-\phi_1-\phi_2-\phi_3) = 0.10937$ หรือเราสามารถเขียนผลการประมาณค่าได้โดย

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

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

ในการประมาณค่าข้างต้น R จะเลือกใช้การประมาณค่าด้วย CSS ก่อนเพื่อหาจุดเริ่มต้นสำหรับ MLE นอกจากนี้เราสามารถเลือกวิธีการประมาณค่าได้โดยการเพิ่ม argument `method=c("ML")` และ `method=c("CSS")` สำหรับการประมาณค่าด้วย MLE และ CSS ตามลำดับ โดยที่ผลการประมาณค่าจะมีค่าแตกต่างกันเล็กน้อยดังที่ได้แสดงไว้ข้างล่าง

```
> m2<-arima(int$diff_th_us, order=c(3,0,0),method=c("ML"))
> m2
Call:
arima(x = int$diff_th_us, order = c(3, 0, 0), method = "ML")
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$diff_th_us, order=c(3,0,0),method=c("CSS"))
> m3
Call:
arima(x = int$diff_th_us, order = c(3, 0, 0), method = "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
```

ถ้าในการประมาณค่าแบบจำลอง AR แล้วเราทดสอบค่าเฉลี่ยด้วย t-test แล้วไม่มีความแตกต่างจากศูนย์อย่างมีนัยสำคัญ เราอาจจะปรับรูปแบบสมการเป็นสมการที่ค่าเฉลี่ยเท่ากับศูนย์ และประมาณค่าโดยการเพิ่ม argument `include.mean=FALSE` ในคำสั่ง `arima`

นอกจากนี้เราสามารถหาค่ารากของสมการพหุนามออโตรีเกรสซีฟได้ตามคำสั่งข้างล่าง

```
> p1<-m1$coef
> p1
       ar1        ar2        ar3  intercept
 1.3294074 -0.4985723  0.1333653  3.0553267
> roots=polyroot(p1)
> roots
[1]  0.4008889+0.5949388i -0.8454280-0.0000000i  0.4008889-0.5949388i
> Mod(roots)
[1] 0.7174009 0.8454280 0.7174009
```

จะเห็นได้ว่าค่ารากของสมการพหุนามเป็นจำนวนเชิงเส้นที่มีมอดุลัสน้อยกว่าหนึ่งแสดงว่าอนุกรมที่เราประมาณค่าได้เป็นอนุกรมไม่นิ่ง (ซึ่งเราจะกลับมาพิจารณาอนุกรมนี้ในหัวข้อต่อไป)

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

```
> adqtest<-Box.test(m1$residuals, lag=12, type="Ljung")
> adqtest
  Box-Ljung test
data:  m1$residuals
X-squared = 14.009, df = 12, p-value = 0.3002
> #calculate p-value with df=12-3
> pv<-1-pchisq(adqtest$statistic, 9)
> pv
X-squared
0.1220215
```

เราทดสอบว่า residuals ของแบบจำลอง m1 เป็น white noise หรือไม่ โดยใช้คำสั่ง `Box.test` กับ `m1$residuals` โดยเก็บผลไว้ที่ `adqtest`

จากค่า LB statistics =14.009 เราสามารถนำไปเปรียบเทียบกับ critical chi-square (df=12-3, $\alpha = 0.05$)=16.92 เนื่องจาก LB statistics < critical chi-sq เราไม่สามารถปฏิเสธสมมุติฐานหลักว่า residuals เป็น white noise แสดงว่าแบบจำลอง AR(3) เพียงพอในการอธิบาย $diff_t h_u s$

นอกจากนี้เราสามารถคำนวณค่า p-value โดยใช้คำสั่ง `pchisq(value, df)` สำหรับคำนวณหา CDF ไปยังจุดที่ chi-square เท่ากับ 14.009 และสามารถคำนวณพื้นที่ปลายหางด้วย 1-CDF ก็จะเป็นค่า p-value

ในกรณีนี้ p-value = 0.12 > α (=0.05) เราไม่สามารถปฏิเสธสมมุติฐานหลักที่ว่า residuals เป็น white noise แสดงว่า**แบบจำลองของเราเพียงพอที่จะใช้อธิบายความแตกต่างของอัตราดอกเบี้ย**

2.1.6 การพยากรณ์ด้วยแบบจำลอง $AR(p)$

การ ใช้ ประโยชน์ จาก แบบ จำลอง อนุกรม เวลา ที่ สำคัญ ประการ หนึ่ง คือ การ พยากรณ์(forecasting) หากสมมุติให้เรากำลังอยู่ 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 คาบเมื่อจุดเริ่มต้นการพยากรณ์อยู่ที่ h (l-step ahead forecast)

ค่าพยากรณ์ไปข้างหน้า 1 คาบจากแบบจำลอง $AR(p)$

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

$$\begin{aligned}\hat{y}_h(1) &= E(y_{h+1}|F_h) \\ &= E(\phi_0 + \phi_1 y_h + ... + \phi_p y_{h+1-p}|F_h) + \underbrace{E(\varepsilon_{h+1}|F_h)}_{=0} \\ &= \phi_0 + \phi_1 y_h + ... + \phi_p y_{h+1-p} = \phi_0 + \sum_{i=1}^{p} \phi_i y_{h+1-i} \end{aligned} \tag{2.35}$$

และค่าผิดพลาดจากการพยากรณ์เท่ากับ

$$\begin{aligned}e_h(1) &= y_{t+1} - \hat{y}_h(1) \\ &= (\phi_0 + \phi_1 y_h + ... + \phi_p y_{h+1-p} + \varepsilon_{h+1}) - (\phi_0 + \phi_1 y_h + ... + \phi_p y_{h+1-p}) = \varepsilon_{h+1}\end{aligned}$$

และค่าแปรปรวนจากการพยากรณ์ไปข้างหน้าหนึ่งคาบเท่ากับ

$$Var(e_h(1)) = Var(\varepsilon_{h+1}) = \sigma^2$$

เราสามารถสร้างช่วงความเชื่อมั่น 95 % ของการพยากรณ์ไปข้างหน้า 1 คาบได้โดยช่วงดังกล่าวจะเท่ากับ

$$\hat{y}_h(1) \pm 1.96 sd(e_h(1))$$

โดยที่ $sd(e_h(1)) = \sqrt{Var(e_h(1)} = \sqrt{\sigma^2}$ หากแทนค่า σ^2 ด้วยตัวประมาณค่า($\widehat{\sigma}^2$)เราจะเรียกค่าดังกล่าวว่าค่าคาดเคลื่อนมาตรฐาน (standard error)

ค่าพยากรณ์ไปข้างหน้า 2 คาบจากแบบจำลอง $AR(p)$

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

$$\begin{aligned}\hat{y}_h(2) &= E(y_{h+2}|F_h) \\ &= E(\phi_0 + \phi_1 y_{h+1} + \phi_2 y_h + ... + \phi_p y_{h+2-p}|F_h) + \underbrace{E(\varepsilon_{h+2}|F_h)}_{=0} \\ &= \phi_0 + \phi_1 \underbrace{E(y_{h+1}|F_h)}_{=\hat{y}_h(1)} + \phi_2 y_h + ... + \phi_p y_{h+2-p} \\ &= \phi_0 + \phi_1 \hat{y}_h(1) + \phi_2 y_{h+2-2} + ... + \phi_p y_{h+2-p}\end{aligned} \tag{2.36}$$

และค่าผิดพลาดจากการพยากรณ์เท่ากับ

$$\begin{aligned} e_h(2) &= y_{h+2} - \hat{y}_h(2) \\ &= (\phi_0 + \phi_1 y_{h+1} + ... + \phi_p y_{h+2-p} + \varepsilon_{h+2}) - (\phi_0 + \phi_1 \hat{y}_h(1) + \phi_2 y_{h+2-2} + ... + \phi_p y_{h+2-p}) \\ &= \varepsilon_{h+2} + \phi_1(y_{h+1} - \hat{y}_h(1)) = \varepsilon_{h+2} + \phi_1 \varepsilon_{h+1} \end{aligned}$$

และค่าแปรปรวนจากการพยากรณ์ไปข้างหน้าหนึ่งคาบเท่ากับ

$$Var(e_h(2)) = Var(\varepsilon_{h+2} + \phi_1 \varepsilon_{h+1}) = (1 + \phi_1^2)\sigma^2$$

เราจะสังเกตได้ว่า $Var(e_h(2)) \geq Var(e_h(1))$ ดังนั้นช่วงความเชื่อมั่น 95% ของการพยากรณ์จะมีขนาดที่กว้างขึ้นเรื่อยๆ ซึ่งสะท้อนความแม่นยำของแบบจำลองอนุกรมเวลาเมื่อเราพยากรณ์ไปในอนาคตที่ไกลขึ้น

ตัวอย่างที่ 2.4 การพยากรณ์ การเปรียบเทียบผลการพยากรณ์กับข้อมูลจริง (in-sample evaluation)

ต่อเนื่องจากตัวอย่าง **??** เราจะพยากรณ์ไปข้างหน้าโดยใช้ตัวอย่างความแตกต่างของอัตราดอกเบี้ย จำนวนข้อมูลที่เรามีอยู่คือ 324 เดือน เพื่อแสดงให้เห็นความแตกต่างระหว่างค่าจริงและค่าพยากรณ์ เราจะประมาณค่าโดยใช้ข้อมูลเดือนมกราคม 1978 ถึงธันวาคม 2003 เป็นข้อมูลจำนวน 312 โดยเรียกว่าตัวแปร `intdiff.in` ดังคำสั่งในบรรทัดที่ 3 และประมาณค่าด้วยแบบจำลอง $AR(3)$ และเก็บค่าประมาณไว้ใน Object m_4

```
1 > length(int$diff_th_us)
2 [1] 324
3 > intdiff.in<-int$diff_th_us[1:312]
4 > m4=arima(intdiff.in,order=c(3,0,0))
```

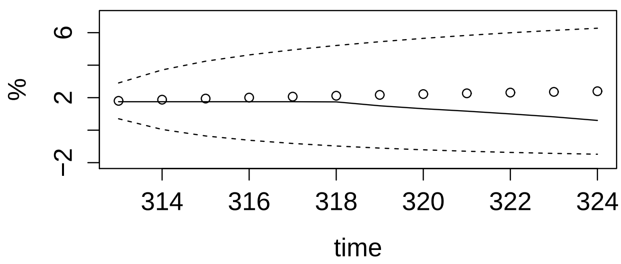
เราสามารถพยากรณ์ไปข้างหน้าโดยใช้คำสั่ง `predict` โดยคำสั่งดังกล่าวจะต้องระบุแบบจำลอง (m_4) และจำนวนคาบไปข้างหน้า (`n.ahead`) จะเก็บค่าดังกล่าวไว้ใน object `m4.pred` โดยใน Object ดังกล่าวจะประกอบด้วย

- `$pred` คือค่าพยากรณ์ไปข้างหน้า 1 ถึง 12 คาบซึ่งแสดงค่าในบรรทัดที่ 8-9 ตัวอย่างเช่น $\hat{y}_{312}(1) = 1.803$ และ $\hat{y}_{312}(12) = 2.397$
- `$se` คือค่าคาดเคลื่อนมาตรฐานของค่าพยากรณ์ไปข้างหน้า 1 ถึง 12 คาบซึ่งแสดงค่าในบรรทัดที่ 15-16 ตัวอย่างเช่น $se(e_{312}(1)) = 0.562$ และ $se(e_{312}(12)) = 1.977$

และสามารถวาดรูปเปรียบเทียบค่าจริง (จากเดือนที่ 313 ถึง 324) และค่าพยากรณ์ (`m4.pred$pred`) พร้อมช่วงความเชื่อมั่น 95 % [ดูคำสั่งในบรรทัด 17 ถึง 20] โดยกราฟดังกล่าวแสดงไว้ที่รูป **??**

```
> 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
$se
Time Series:
Start = 313
End = 324
Frequency = 1
 [1] 0.5622779 0.9342948 1.1737163 1.3399272 1.4696143 1.5777365 1.6700163
 [8] 1.7494732 1.8183586 1.8784737 1.9312318 1.9777417
> plot(seq(313,324),int$diff_th_us[313:324],type="l", ylim=c(-2,7), ylab="%",
    xlab="time")
> points(seq(313,324), m4.pred$pred)
> lines(seq(313,324), m4.pred$pred+1.96*m4.pred$se,lty=2)
> lines(seq(313,324), m4.pred$pred-1.96*m4.pred$se,lty=2)
```

รูปที่ 2.6: การพยากรณ์ความแตกต่างระหว่างอัตราดอกเบี้ยของไทยและสหรัฐอเมริกาจาก $AR(3)$



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

แบบจำลองนี้มีพื้นฐานมากจากแนวคิดที่ว่าอนุกรมเวลาเส้นตรงนิ่ง (stationary) สามารถเขียนให้อยู่ในรูปมูวิ่งเอเวอเรจอันดับเป็นอนันต์ (infinite moving average representation) และสามารถลดจำนวนความล่า(lag) ให้เหลือ q เรียกว่าแบบจำลองมูวิ่งเอเวอเรจอันดับ q และเขียนแทนด้วย $MA(q)$ และแสดงได้ด้วยสมการ

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

โดยที่ $\theta_0 = 1$

2.2.1 แบบจำลอง MA(1)

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

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} \tag{2.38}$$

เราเริ่มต้นด้วยการหาค่าคาดหมายของ y_t โดยการใส่ค่าคาดหมาย (take expectation) ทั้งสองข้างของสมการ (**??**) จะได้

$$E(y_t) = \mu + \underbrace{E(\varepsilon_t)}_{=0} + \theta_1 \underbrace{E(\varepsilon_{t-1})}_{=0} = \mu$$

ต่อไปเราจะพิจารณาค่าความแปรปรวนของ y_t (γ_0)

$$\begin{aligned}
\gamma_0 = Var(y_t) &= E(y_t - E(y_t))^2 \\
&= E(\varepsilon_t + \theta_1 \varepsilon_{t-1})^2 \\
&= E(\varepsilon_t^2 + \theta_1 \varepsilon_t \varepsilon_{t-1} + \theta_1^2 \varepsilon_{t-1}^2) \\
&= \underbrace{E(\varepsilon_t^2)}_{=\sigma_\varepsilon^2} + \theta_1 \underbrace{E(\varepsilon_t \varepsilon_{t-1})}_{=0} + \theta_1^2 \underbrace{E(\varepsilon_{t-1}^2)}_{=\sigma_\varepsilon^2} \\
\gamma_0 &= (1 + \theta_1^2)\sigma_\varepsilon^2
\end{aligned} \tag{2.39}$$

ต่อไปเราจะพิจารณาค่าความแปรปรวนร่วมในตัวเองย้อนไปหนึ่งคาบ (one-lag autocovariance)

$$\begin{aligned}
\gamma_1 &= Cov(y_t, y_{t-1}) = E\left[(y_t - E(y_t))(y_{t-1} - E(y_{t-1}))\right] \\
&= E\left[(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_{t-1} + \theta_1 \varepsilon_{t-2})\right] \\
&= E\left[\varepsilon_t \varepsilon_{t-1} + \theta_1 \varepsilon_{t-1}^2 + \theta_1 \varepsilon_t \varepsilon_{t-2} + \theta_1^2 \varepsilon_{t-1} \varepsilon_{t-2}\right] \\
&= \underbrace{E(\varepsilon_t \varepsilon_{t-1})}_{=0} + \theta_1 \underbrace{E(\varepsilon_{t-1}^2)}_{\sigma_\varepsilon^2} + \theta_1 \underbrace{E(\varepsilon_t \varepsilon_{t-2})}_{=0} + \theta_1^2 \underbrace{E(\varepsilon_{t-1} \varepsilon_{t-2})}_{=0} \\
\gamma_1 &= \theta_1 \sigma_\varepsilon^2
\end{aligned} \tag{2.40}$$

พิจารณาค่าความแปรปรวนร่วมในตัวเองย้อนไปที่ห่างกันสองช่วงเวลา

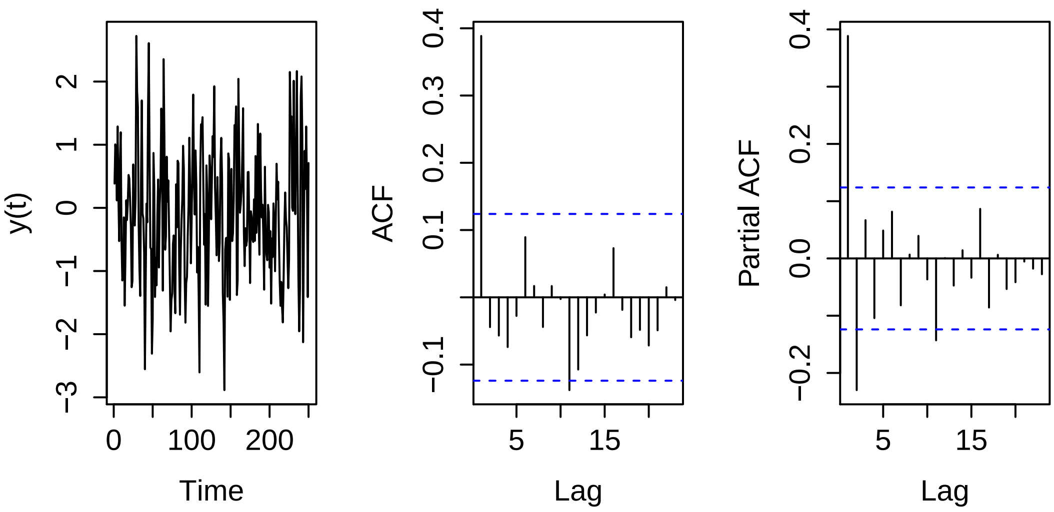
$$\begin{aligned}
\gamma_2 = Cov(y_t, y_{t-2}) &= E\left[(y_t - E(y_t))(y_{t-2} - E(y_{t-2}))\right] \\
&= E\left[(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_{t-2} + \theta_1 \varepsilon_{t-3})\right] \\
&= E\left[\varepsilon_t \varepsilon_{t-2} + \theta_1 \varepsilon_{t-1} \varepsilon_{t-2} + \theta_1 \varepsilon_t \varepsilon_{t-3} + \theta_1^2 \varepsilon_{t-1} \varepsilon_{t-3}\right] \\
\gamma_2 &= \underbrace{E(\varepsilon_t \varepsilon_{t-2})}_{=0} + \theta_1 \underbrace{E(\varepsilon_{t-1} \varepsilon_{t-2})}_{=0} + \theta_1 \underbrace{E(\varepsilon_t \varepsilon_{t-3})}_{=0} + \theta_1^2 \underbrace{E(\varepsilon_{t-1} \varepsilon_{t-3})}_{=0} = 0
\end{aligned} \tag{2.41}$$

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

$$\rho_k = \begin{cases} \frac{\theta_1}{1+\theta_1^2} & , k = 1 \\ 0 & , k > 1 \end{cases} \tag{2.42}$$

จะเห็นได้ว่า y_t มีสหสัมพันธ์กับ y_{t-1} แต่ไม่สัมพันธ์กับ $y_{t-2}, \ldots$ ซึ่งแตกต่างจาก $AR(1)$ รูปที่ **??** และ **??** แสดงอนุกรมเวลา y_t และ ACF กรณีที่ $\theta = 0.5$ และ $\theta = -0.5$ ตามลำดับ

รูปที่ 2.7: การจำลองข้อมูลกระบวนการ $MA(1)$ ที่มีค่าสัมประสิทธิ์ $\theta = 0.5$ และกลุ่มตัวอย่าง 500 หน่วย



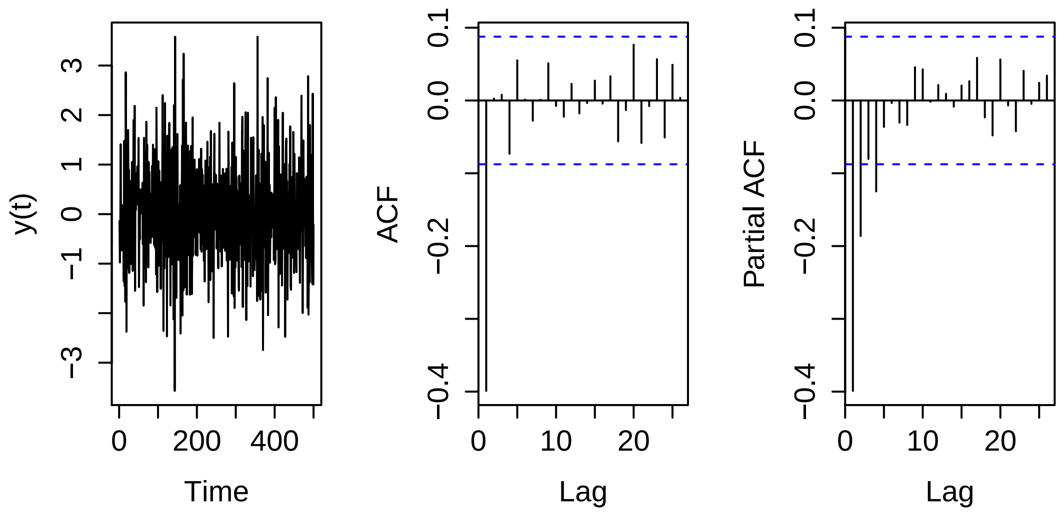
เราจะเห็นได้ว่าอนุกรมเวลา y_t ที่เป็นกระบวนการมูววิ่งเอเวเรจจะเป็นอนุกรมเวลานิ่ง (stationary) เสมอ เนื่องจากกระบวนการดังกล่าวเป็นการรวมเชิงเส้นตรง (linear combination) ของอนุกรมเวลาไวท์นอซเข้าด้วยกัน ซึ่งอนุกรมเวลาไวท์นอซจะมีค่าโมเมนต์ที่หนึ่งและสองที่ไม่ขึ้นกับเวลา

กระบวนการที่หาค่าผกผันได้ (invertible)

นอกจากนี้หากพิจารณาค่าสัมประสิทธิ์ของ $MA(1)$ จะเห็นได้ว่าค่าความแปรปรวนร่วมในตัวเองของกระบวนการ $MA(1)$ ที่มีค่า $\theta_1 = 5$ และ $\theta_1 = 1/5$ จะมีค่าเท่ากัน นอกจากนี้ฟังก์ชันสหสัมพันธ์ร่วมในตัวเองของกระบวนการ $MA(1)$ สองกระบวนการต่อไปนี้จะมีลักษณะที่เหมือนกัน

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

รูปที่ 2.8: การจำลองข้อมูลกระบวนการ $MA(1)$ ที่มีค่าสัมประสิทธิ์ $\theta = -0.5$ และกลุ่มตัวอย่าง 500 หน่วย



และ

$$y_t = v_t + 5v_{t-1}, \qquad v_t \sim iidN(0,1)$$

เมื่อเราสังเกตเห็นเพียงแต่อนุกรม x_t หรือ y_t โดยที่ไม่ทราบอนุกรมของช็อก w_t หรือ v_t ดังนั้นเราต้องเลือกว่าจะใช้แบบจำลองไหน ซึ่งเราจะเลือกแบบจำลองที่สามารถเขียนแบบจำลองในรูปของ infinite AR representation โดยที่อนุกรมที่มีลักษณะดังกล่าวเราจะเรียกว่ากระบวนการที่หาตัวผกผันได้(invertible process) ยกตัวอย่างเช่น $y_t = (1+\theta_1 L)\varepsilon_t$ เราสามารถเขียนแบบจำลองดังกล่าวในรูปต่อไปนี้

$$\begin{aligned}
(1+\theta_1 L)^{-1} y_t &= \varepsilon_t \\
(1 + (-\theta_1 L) + (-\theta_1 L)^2 + (-\theta_1 L)^3 + ...)y_t &= \varepsilon_t \\
y_t - \theta_1 y_{t-1} + \theta_1^2 y_{t-2} - \theta_1^3 y_{t-3} + ... &= \varepsilon_t \\
y_t &= \varepsilon_t + \theta_1 y_{t-1} - \theta_1^2 y_{t-2} + \theta_1^3 y_{t-3} + ...
\end{aligned} \tag{2.43}$$

ซึ่งสมการ **??** คือรูปแบบ AR ที่มีอันดับเป็นอนันต์ของ $y_t \sim MA(1)$ ซึ่งจะเป็นอนุกรมนิ่งหาก $|\theta_1| < 1$ ดังนั้น เงื่อนไขที่ $MA(1)$ จะเป็นกระบวนการที่หาตัวผกผันได้ คือ $|\theta_1| < 1$

เราสามารถเขียนเงื่อนไขที่สามารถเขียนกระบวนการที่หาตัวผกผันได้ด้วยการเขียนกระ-

บวนการ $MA(1)$ ในรูปสมการพหุนาม

$$y_t - \mu = (1 + \theta_1 L)\varepsilon_t$$

โดยที่สมการช่วยคือ $(1 + \theta_1 m) = 0$ ดังนั้นรากของสมการพหุนามมูววิงเอเวเรจคือ $m = \frac{-1}{\theta_1}$ เมื่อพิจารณารากของสมการพหุนาม $|m| = |\frac{-1}{\theta_1}| = \frac{1}{|\theta_1|_{<1}} > 1$ จะได้ว่าค่าสัมบูรณ์ของรากของสมการพหุนามมูววิงเอเวเรจจะต้องมีค่ามากกว่า 1

2.2.2 แบบจำลอง $MA(2)$

กำหนดให้ y_t เป็นกระบวนการ $MA(2)$ ซึ่งสามารถแสดงได้ด้วยสมการ

$$y_t - \mu = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \tag{2.44}$$

ซึ่งรูปแบบดังกล่าวจะมีค่าเฉลี่ยเท่ากับ μ

$$\begin{aligned} E(y_t) - \mu &= \underbrace{E(\varepsilon_1)}_{=0} + \theta_1 \underbrace{E(\varepsilon_{t-1})}_{=0} + \theta_2 \underbrace{E(\varepsilon_{t-2})}_{=0} \\ E(y_t) &= \mu \end{aligned}$$

ค่าความแปรปรวนของ y_t สามารถหาได้โดย

$$\begin{aligned} Var(y_t) = E(y_t - \mu)^2 &= E\left[(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2})(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2})\right] \\ &= E\left[\varepsilon_t^2 + \theta_1^2 \varepsilon_{t-1}^2 + \theta_2^2 \varepsilon_{t-2}^2 + \text{cross term}\right] \\ &= \underbrace{E(\varepsilon_t^2)}_{=\sigma_\varepsilon^2} + \theta_1^2 \underbrace{E(\varepsilon_{t-1}^2)}_{=\sigma_\varepsilon^2} + \theta_2^2 \underbrace{E(\varepsilon_{t-2}^2)}_{=\sigma_\varepsilon^2} + \underbrace{E(\text{cross term})}_{=0} \\ &= (1 + \theta_1^2 + \theta_2^2)\sigma_\varepsilon^2 \equiv \gamma_0 \end{aligned} \tag{2.45}$$

ต่อไปเราจะพิจารณาค่าความแปรปรวนในตัวเองย้อนไปหนึ่งคาบ

$$\begin{aligned} \gamma_1 = Cov(y_t, y_{t-1}) &= E\left[(y_t - \mu)(y_{t-1} - \mu)\right] \\ &= E\left[(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2})(\varepsilon_{t-1} + \theta_1 \varepsilon_{t-2} + \theta_2 \varepsilon_{t-3})\right] \\ &= E\left[\theta_1 \varepsilon_{t-1}^2 + \theta_2 \theta_1 \varepsilon_{t-2}^2 + \text{cross term}\right] \\ &= \theta_1 \underbrace{E(\varepsilon_{t-1}^2)}_{=\sigma_\varepsilon^2} + \theta_2 \theta_1 \underbrace{E(\varepsilon_{t-2}^2)}_{=\sigma_\varepsilon^2} + \underbrace{E(\text{cross term})}_{=0} \\ &= (\theta_1 + \theta_2 \theta_1)\sigma_\varepsilon^2 \end{aligned} \tag{2.46}$$

ค่าความแปรปรวนในตัวเองย้อนไปสองคาบ

$$\begin{aligned}
\gamma_2 = Cov(y_t, y_{t-2}) &= E\left[(y_t - \mu)(y_{t-2} - \mu)\right] \\
&= E\left[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2})(\varepsilon_{t-2} + \theta_1\varepsilon_{t-3} + \theta_2\varepsilon_{t-4})\right] \\
&= E\left[\theta_2\varepsilon_{t-2}^2 + \text{cross term}\right] \\
&= \theta_2 \underbrace{E(\varepsilon_{t-2}^2)}_{=\sigma_\varepsilon^2} + \underbrace{E(\text{cross term})}_{=0} \\
&= \theta_2\sigma_\varepsilon^2
\end{aligned} \tag{2.47}$$

และค่าความแปรปรวนในตัวเองย้อนไปสามคาบ

$$\begin{aligned}
\gamma_3 = Cov(y_t, y_{t-2}) &= E\left[(y_t - \mu)(y_{t-3} - \mu)\right] \\
&= E\left[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2})(\varepsilon_{t-3} + \theta_1\varepsilon_{t-4} + \theta_2\varepsilon_{t-5})\right] \\
&= E\left[\text{cross term}\right] = 0
\end{aligned} \tag{2.48}$$

เมื่อพิจารณาค่าสหสัมพันธ์ในตัวเองของอนุกรมเวลา $MA(2)$ จะได้ว่า

$$\begin{aligned}
\rho_0 &= \frac{\gamma_0}{\gamma_0} = 1 \\
\rho_1 &= \frac{\gamma_1}{\gamma_0} = \frac{(\theta_1 + \theta_2\theta_1)}{(1 + \theta_1^2 + \theta_2^2)} \\
\rho_2 &= \frac{\gamma_2}{\gamma_0} = \frac{\theta_2}{(1 + \theta_1^2 + \theta_2^2)} \\
\rho_3 &= \frac{\gamma_3}{\gamma_0} = 0 \\
\rho_j &= \frac{\gamma_j}{\gamma_0} = 0
\end{aligned}$$

กรณีที่ $j > 2$ จากสมการข้างบนจะเห็นได้ว่า ACF ของ $MA(2)$ จะมีค่ามากกว่าศูนย์ในกรณีที่ $j = 1$ และ $j = 2$ และจะลดลงเท่ากับศูนย์ในกรณีที่ $j > 2$

2.2.3 แบบจำลอง $MA(q)$

แบบจำลอง $MA(q)$ สามารถเขียนในรูปสมการดังนี้

$$y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + ... + \theta_q\varepsilon_{t-q}, \qquad \varepsilon_t \sim WN(0, \sigma^2) \tag{2.49}$$

แบบจำลอง $MA(q)$ นิ่งและ ergodic[3] ถ้าค่า $\theta_1, ..., \theta_q$ มีค่าจำกัด และแบบจำลองสามารถหาตัวผกผันได้ถ้าค่าราก (roots) ของพหุนาม MA ต่อไปนี้

$$\theta(m) = 1 + \theta_1 m + ... + \theta_q m^{t-q} = 0 \tag{2.50}$$

มีค่าสัมบูรณ์(หรือค่าโมดุลัส)ของรากของพหุนามมากกว่า 1 และเราสามารถสรุปโมเมนต์ของกระบวนการ $MA(q)$ ได้ดังนี้

$$\begin{aligned} E(y_t) &= \mu \\ \gamma_0 &= \sigma_\varepsilon^2(1 + \theta_1^2 + ... + \theta_q^2) \\ \gamma_j &= \begin{cases} (\theta_j + \theta_{j+1}\theta_1 + \theta_{t+2}\theta_2 + ... + \theta_q\theta_{q-j})\sigma_\varepsilon^2, & , \quad j = 1, 2, ..., q \\ 0 & , \quad j > q \end{cases} \end{aligned}$$

และค่าสหสัมพันธ์ในตัวเองจะเท่ากับ

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \frac{(\theta_j + \theta_{j+1}\theta_1 + \theta_{t+2}\theta_2 + ... + \theta_q\theta_{q-j})}{(1 + \theta_1^2 + ... + \theta_q^2)}$$

กรณีที่ $j = 1, 2, ..., q$ และ $\rho_j = 0$ กรณีที่ $j > q$ จะเห็นได้ว่า ACF ของ $MA(q)$ จะมีค่าไม่เท่ากับศูนย์จะกระทั่งช่วงล่า q และจะเท่ากับศูนย์หลังจากนั้น ในขณะที่ PACF จะมีค่อยลดลงไป

2.2.4 การประมาณค่าแบบจำลอง $MA(q)$

การประมาณค่าความควรจะเป็นสูงสุด (maximum likelihood estimation) เป็นวิธีการทั่วไปที่ใช้ในการประมาณค่าแบบจำลอง $MA(q)$

วิธีการประมาณค่าความควรจะเป็นแบบมีเงื่อนไข (Conditional maximum likelihood) สูงสุด

ในหัวข้อนี้จะประมาณค่าแบบจำลอง $MA(1)$ ซึ่งสามารถเขียนได้ดังนี้

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

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

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

[3]เนื่องจากแนวคิดเกี่ยวกับอนุกรมเวลานิ่งมิได้อธิบายความความสัมพันธ์ระหว่างข้อมูลในช่วงเวลาที่แตกต่างกัน อนุกรมเวลาจะมีคุณสมบัติเป็น ergodic ถ้าช่วงของข้อมูลสองช่วงเวลายิ่งมีความสัมพันธ์ระหว่างกันที่ลดลงเมื่อช่วงเวลาทั้งสองอยู่ห่างกันมากขึ้น

และเมื่อสามารถสังเกตค่า y_1 เราก็สามารถหาค่า $\varepsilon_1 = y_1 - \mu$ และจากข้อมูลในคาบที่ 1 จะได้ว่า

$$y_2|y_1 = \mu + \varepsilon_2 + \theta\varepsilon_1$$

และเราสามารถหาค่า $\varepsilon_2 = y_2 - \mu - \theta\varepsilon_1$ และได้ฟังก์ชันความน่าจะเป็น

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

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

$$\begin{aligned} L(\mu, \theta, \sigma_\varepsilon^2) &= \Pi_{t=2}^{T} f_{y_t|y_{t-1}} \\ &= (2\pi\sigma_\varepsilon^2)^{-(T-2)/2} \exp\left[\frac{\sum_{t=2}^{T} - \overbrace{(y_t - \mu - \theta\varepsilon_{t-1})^2}^{\varepsilon_t^2}}{2\sigma_\varepsilon^2}\right] \end{aligned} \tag{2.51}$$

สมการ **??** เราไม่สามารถหาตัวประมาณค่าได้ด้วยการวิเคราะห์อนุพันธ์เช่นในกรณี AR ในกรณีนี้ เราจำเป็นต้องใช้วิธีการทางตัวเลขด้วยคอมพิวเตอร์เพื่อหาค่า $\hat{\theta}$ ที่ทำให้ $L(\mu, \theta, \sigma_\varepsilon^2)$ มีค่าสูงที่สุด

วิธีการประมาณค่าความควรจะเป็นแบบแม่นตรง (Exact maximum likelihood) สูงสุด

โดยทั่วไปแล้วเราใช้วิธีการสองวิธีการในการคำนวณฟังก์ชันค่าความควรจะเป็นแบบแม่นตรง (Exact likelihood function) สำหรับ $MA(1)$ คือ Kalman filter และการใช้ triangular factorization ของ var-cov matrix ซึ่งผู้ที่สนใจสามารถหาอ่านได้จาก Hamilton (1993)

ตัวอย่างที่ 2.5: การสร้างแบบจำลองสำหรับผลได้ตอบแทนรายเดือนจาก SET โดย MA(q)

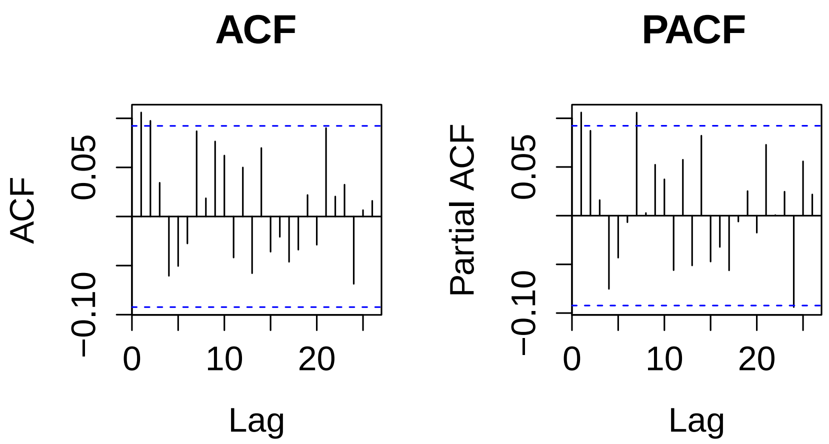
ในตัวอย่างนี้เราจะพิจารณาผลได้ตอบแทนรายเดือนของการลงทุนในตลาดหลักทรัพย์แห่งประเทศไทย โดยใช้ข้อมูลราคาปิดรายเดือนตั้งแต่ เมษายน 2518 ถึงเดือนพฤศจิกายน 2555 (ข้อมูลจาก `www.set.or.th/th/market/market_statistics.html`) ซึ่งอยู่ในไฟล์ `mset.csv` โดยเป็นข้อมูลคอลัมน์เดียว เราสามารถนำเข้าข้อมูลและคำนวณผลได้ตอบแทนได้ด้วยคำสั่งต่อไปนี้

```
> mset <-
    read.csv("https://raw.githubusercontent.com/chaleampong/EC435/master/mset.csv",
    header = FALSE)
> head(mset)
> ret<-diff(log(mset$V1))
> ts.plot(ret)
```

หลังจากนั้นเราจะพิจารณาแบบจำลองที่เหมาะสมโดยการพิจารณา ACF และ PACF

```
> acf(ret)
> pacf(ret)
```

รูปที่ 2.9: ACF และ PACF ของผลได้ตอบแทนรายเดือนจาก SET



จากรูป **??** จะเห็นได้ว่า ACF มีค่าลดลงจนมีค่าไม่แตกต่างจากศูนย์ที่ค่าล่าเท่ากับ 3 ดังนั้นแบบจำลองที่เหมาะสมน่าจะเป็น $MA(2)$ 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
```

จากผลการประมาณค่าเราได้ค่าสัมประสิทธิ์ดังนี้ $\theta_1 = 0.0886, \theta_2 = 0.1001$ และ $\mu = 0.0057$ และมีค่าความคาดเคลื่อนมาตรฐาน (s.e.) อยู่ในบรรทัด `s.e.`

เราสามารถพิจารณาความพอเพียงของแบบจำลอง `m1` ได้ด้วยการทดสอบ L-B สำหรับ residuals ดังคำสั่งต่อไปนี้


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

เราได้ค่า $Q(12)$ เท่ากับ 13.6075 ซึ่งสามารถหาค่าพีที่มีองศาอิสระเท่ากับ $(m-g) = 12-2$ โดยที่ g คือจำนวนอันดับของ MA จะได้ค่าพีเท่ากับ 0.192 แสดงว่าค่าส่วนเกิน (residuals) ไม่มีความสัมพันธ์ระหว่างกัน แสดงว่าแบบจำลองของเราเพียงพอที่ใช้อธิบายผลได้ตอบแทนรายเดือนของการลงทุนในตลาดหลักทรัพย์แห่งประเทศไทย

2.2.5 การพยากรณ์จากแบบจำลอง $MA(q)$

ในแบบจำลอง $MA(q)$ ความจำของข้อมูลค่อนข้างจำกัดซึ่งเราสามารถสังเกตเห็นได้จากค่าพยากรณ์ ตัวอย่างเช่นหากเราพิจารณาแบบจำลอง $MA(1)$: $y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1}$ โดยเราอยู่ ณ คาบเวลา h

การพยากรณ์ไปข้างหน้า 1 คาบ

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

$$\hat{y}_h(1) = E(y_{h+1}|F_h) = \mu + \underbrace{E(\varepsilon_{h+1}|F_h)}_{=0} + \theta_1\varepsilon_h = \mu + \theta_1\varepsilon_h \qquad (2.52)$$

และค่าผิดพลาดจากการพยากรณ์เท่ากับ

$$e_h(1) = y_{h+1} - \hat{y}_h(1) = \varepsilon_{h+1}$$

และค่าแปรปรวนจากการพยากรณ์ไปข้างหน้าหนึ่งคาบเท่ากับ

$$Var(e_h(1)) = Var(\varepsilon_{h+1}) = \sigma^2$$

ในการพยากรณ์จากแบบจำลอง $MA(1)$ เราต้องใช้ค่าช็อกซึ่งเราไม่สามารถสังเกตได้ ในทางปฏิบัติค่า ε_h จะสามารถคำนวณได้สองวิธีคือ

- สมมุติให้ $\varepsilon_0 = 0$ แล้วแทนค่าใน $\varepsilon_1 = y_1 - \mu$ และ $\varepsilon_t = y_t - \mu - \theta_1\varepsilon_{t-1}$ สำหรับ $2 \le t \le h$
- หรือใช้ค่า $\hat{\varepsilon}_h$ ที่เป็นค่าตกค้าง (residual) จากการประมาณค่า $MA(1)$

การพยากรณ์ไปข้างหน้า 2 คาบ

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

$$\hat{y}_h(2) = E(y_{h+2}|F_h) = \mu + \underbrace{E(\varepsilon_{h+2}|F_h)}_{=0} + \theta_1 \underbrace{\varepsilon_{h+1}}_{=0} = \mu \tag{2.53}$$

ซึ่งเท่ากับค่าเฉลี่ยของอนุกรม และค่าผิดพลาดจากการพยากรณ์เท่ากับ

$$e_h(2) = y_{h+2} - \hat{y}_h(2) = \varepsilon_{h+2} + \theta_1 \varepsilon_{h+1}$$

และค่าแปรปรวนจากการพยากรณ์ไปข้างหน้าสองคาบเท่ากับ

$$Var(e_h(2)) = Var(\varepsilon_{h+2} + \theta_1 \varepsilon_{h+1}) = (1 + \theta_1^2)\sigma^2$$

ซึ่งเราสามารถเขียนสรุปค่าพยากรณ์สำหรับแบบจำลอง $MA(1)$ ได้ดังนี้

$$\hat{y}_h(k) = \begin{cases} \mu + \theta_1 \varepsilon_k, & k = 1 \\ \mu, & k > 1 \end{cases}$$

คุณสมบัติดังกล่าวสามารถขยายผลไปยัง $MA(q)$ ใดๆว่าค่าพยากรณ์ไปข้างหน้า k คาบจะเท่ากับค่าเฉลี่ย หากคาบที่พยากรณ์ไปข้างหน้ามีค่าสูงกว่าอันดับของ MA $(k > q)$

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

เราสามารถสร้างแบบจำลองที่มีทั้งส่วนประกอบของออโตรีเกรสซีฟและมูววิ่งเอเวอเรจ ซึ่งหาก y_t เป็นกระบวนการที่เรียกว่าแบบจำลองออโตรีเกรสซีฟมูววิ่งเอเวอเรจอันดับ (p,q) หรือเรียกย่อๆว่า อารมา $ARMA(p,q)$ ถ้า y_t มีค่าเฉลี่ยไม่เท่ากับศูนย์ เราจะสามารถอธิบาย y_t ได้ด้วย

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

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

$$y_t = \alpha + \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q} \tag{2.55}$$

นิยาม 2.2. *เราสามารถเขียนสมการอารมาได้ในรูปพหุนามออโตรีเกรสซีฟ และมูววิ่งเอเวอเรจได้ดังนี้*

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

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

เราสามารถพิจารณาคุณสมบัติของแบบจำลองอารมาได้ด้วยการพิจารณา $ARMA(1,1)$

2.3.1 แบบจำลอง $ARMA(1,1)$

กำหนดให้ y_t เป็นกระบวนการ $ARMA(1,1)$ ซึ่งสามารถอธิบายได้ด้วยสมการข้างล่าง

$$y_t = \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1} \tag{2.56}$$

โดยที่ $\varepsilon_t \sim iidN(0, \sigma^2)$ เพื่อสร้างสมการ Yule-Walker เราใช้ความสัมพันธ์ต่อไปนี้

$$\begin{aligned} E(\varepsilon_t y_t) &= E\left[\varepsilon_t(\phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1})\right] \\ &= \phi \underbrace{E(\varepsilon_t y_{t-1})}_{=0} + \underbrace{E(\varepsilon_t^2)}_{=\sigma^2} + \theta \underbrace{E(\varepsilon_t \varepsilon_{t-1})}_{=0} = \sigma^2 \end{aligned}$$

และ

$$\begin{aligned} E(\varepsilon_{t-1} y_t) &= E\left[\varepsilon_{t-1}(\phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1})\right] \\ &= \phi \underbrace{E(\varepsilon_{t-1} y_{t-1})}_{=\sigma^2} + \underbrace{E(\varepsilon_{t-1} \varepsilon_t)}_{=0} + \theta \underbrace{E(\varepsilon_{t-1}^2)}_{=\sigma^2} \\ &= \phi \sigma^2 + \theta \sigma^2 = (\phi + \theta) \sigma^2 \end{aligned}$$

หากเราคูณสมการ **??** ทั้งสองข้างด้วย y_{t-k} กรณีที่ $k = 0, 1, ..., j, ...$ และใส่ค่าคาดหวัง (take expectation) จะได้

$$\begin{aligned}
\gamma_0 = E(y_t y_t) &= E((\phi y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1})y_t) \\
&= \phi \underbrace{E(y_t y_{t-1})}_{=\gamma_1} + \underbrace{E(y_t\varepsilon_t)}_{=\sigma^2} + \theta \underbrace{E(y_t\varepsilon_{t-1})}_{=(\phi+\theta)\sigma^2} \\
&= \phi\gamma_1 + [1 + \theta(\phi + \theta)]\,\sigma^2 \qquad (2.57)
\end{aligned}$$

$$\begin{aligned}
\gamma_1 = E(y_t y_{t-1}) &= E((\phi y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1})y_{t-1}) \\
&= \phi \underbrace{E(y_{t-1} y_{t-1})}_{=\gamma_0} + \underbrace{E(y_{t-1}\varepsilon_t)}_{=0} + \theta \underbrace{E(y_{t-1}\varepsilon_{t-1})}_{=\sigma^2} \\
&= \phi\gamma_0 + \theta\sigma^2 \qquad (2.58)
\end{aligned}$$

$$\begin{aligned}
\gamma_j = E(y_t y_{t-j}) &= E((\phi y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1})y_{t-j}) \\
&= \phi \underbrace{E(y_{t-1} y_{t-j})}_{=\gamma_{j-1}} + \underbrace{E(\varepsilon_t y_{t-j})}_{=0} + \theta \underbrace{E(\varepsilon_{t-1} y_{t-j})}_{=0} \qquad (2.59) \\
&= \phi\gamma_{j-1} \qquad (2.60)
\end{aligned}$$

กรณีที่ $j \geq 2$ จากสมการ **??** และ **??** เราจะได้

$$Var(y_t) = \gamma_0 = \frac{(1 + 2\phi\theta + \theta^2)}{1 - \phi^2}\sigma^2$$

และจากสมการ **??** เราสามารถสามารถหาค่าสหสัมพันธ์ในตัวเอง ณ ค่าล่า j ใดๆที่ $j \geq 2$ ได้

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

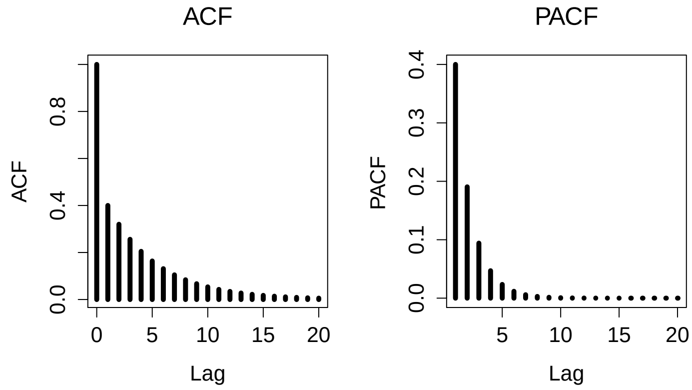
กรณีที่ $j \geq 2$ จะเห็นได้ว่า ACF มีค่าที่ลดลงเรื่อยๆแบบฟังก์ชันเลขชี้กำลัง(exponential) นอกจากนี้เราสามารถเขียนแบบจำลอง $ARMA(1,1)$ ในรูปของมูววิ่งเอเวอเรจที่มีอันดับเป็นอนันต์ (infinite MA)ได้เป็น

$$y_t = \varepsilon_t + (\phi - \theta)\sum_{j=1}^{\infty}\phi^{j-1}\varepsilon_{t-j} \qquad (2.61)$$

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

กระบวนการ $ARMA(p,q)$ จะนิ่งและ ergodic ถ้าค่าสัมบูรณ์ของค่ารากของพหุนามออโตรีเกรสซีฟ $\phi(m) = 0$ มีค่ามากกว่าหนึ่ง (หรือมอดุลัสมีค่ามากกว่าหนึ่งในกรณีจำนวนเชิงซ้อน)

รูปที่ 2.10: ACF และ PACF ของอนุกรมเวลา ARMA(1,1)



และสามารถหาค่าผกผันได้ถ้ารากของพหุนามมูววิ่งเอเวอเรจ $\theta(m) = 0$ มีค่ามากกว่าหนึ่ง (หรือมอดุลัสมีค่ามากกว่าหนึ่งในกรณีจำนวนเชิงซ้อน) และพหุนาม AR และ MA ไม่มีตัวประกอบร่วมกระบวนการ $ARMA(p, q)$ ที่นิ่งจะมีค่าเฉลี่ยเท่ากับ

$$\mu = \frac{c}{1 - \phi_1 - ... - \phi_p}$$

และค่าความแปรปรวนร่วมในตัวเอง ค่าสหสัมพันธ์ในตัวเองและตัวถ่วงน้ำหนักการตอบสนองแรงกระตุ้นสามารถเขียนในลักษณะเวียนเกิด (recursive) ได้ดังนี้

$$\gamma_j = \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} + ... + \phi_p\gamma_{j-p}$$
$$\rho_j = \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + ... + \phi_p\rho_{j-p}$$
$$\psi_j = \phi_1\psi_{j-1} + \phi_2\psi_{j-2} + ... + \phi_p\psi_{j-p}$$

โดยที่รูปทั่วไปของ ACF ของกระบวนการ $ARMA(p, q)$ ค่อนข้างยุ่งยาก โดยสรุปแล้วทั้ง ACF และ PACF จะค่อยๆลดลงเรื่อยแบบฟังก์ชันเลขชี้กำลัง (exponential) ดังตัวอย่างในรูปต่อไปนี้

ตัวอย่าง ความฟุ่มเฟือยของพารามิเตอร์ สมมุติว่าเรากำลังพิจารณากระบวนการไวท์นอซ $y_t = \varepsilon_t$ แล้วเราคูณทั้งสองและขยับไปข้างหลังหนึ่งคาบ $0.5y_{t-1} = 0.5\varepsilon_{t-1}$ หลังจากนั้น เรานำสมการดังกล่าวไปลบจากสมการดั้งเดิมจะได้

$$y_t - 0.5y_{t-1} = \varepsilon_t - 0.5\varepsilon_{t-1} \tag{2.62}$$

จะมีลักษณะเหมือนกับแบบจำลอง $ARMA(1, 1)$ แต่ y_t ก็ยังเป็นไวท์นอซ การใช้พารามิเตอร์อย่างฟุ่มเฟือยหรือใส่พารามิเตอร์มากเกินไป จะทำให้เราเข้าใจผิดว่ากระบวนการมีความสัมพันธ์ระ-หว่างกัน อย่างไรก็ตามปัญหานี้สามารถหลีกเลี่ยงได้ด้วยการพิจารณา polynomial และตัวคูณร่วม

ของ polynomial ตัวอย่างเช่นสมการ (??) สามารถเขียนได้เป็น

$$(1-0.5L)y_t = (1-0.5L)\varepsilon_t$$

ซึ่งสามารถเปลี่ยนกลับไปเป็นกระบวนการไวท์นอซด้วยการคูณทั้งสองข้างด้วย $(1-0.5L)^{-1}$

2.3.2 การประมาณค่าสมการแบบจำลอง ARMA

เราสามารถประมาณค่าแบบจำลอง $ARMA(p,q)$ ได้ด้วยวิธีการค่าความควรจะเป็นสูงที่สุด (MLE) ซึ่งเราสามารถใช้รูปแบบ state-space ในการสร้างฟังก์ชันล็อกไลลิฮูด อย่างไรก็ตาม ฟังก์ชันดังกล่าวที่แน่นอน (exact log-likelihood) นั้นมีรูปแบบที่ค่อนข้างซับซ้อน นอกจากนี้จุดเริ่มต้นของข้อมูลในกรณี exact likelihood จะใช้คุณสมบัติของการแจกแจงที่นิ่งเพื่อสร้างฟังก์ชันค่าควรจะเป็นของค่า y_t p ค่าแรก และ ε_t q ค่าแรก ในขณะที่ conditional likelihood จะสมมุติให้ y_t p ค่าแรก และ ε_t q ค่าแรกเท่ากับศูนย์ ตัวประมาณค่า Exact MLE จะเกิดจากการหาค่าสูงสุดจาก exact log-likelihood และ Conditional MLE จะเกิดจากการหาค่าสูงสุดจาก conditional log-likelihood โดยในกรณีที่จำนวนตัวอย่างมีจำนวนเข้าใกล้อนันต์จะมีค่าใกล้เคียงกัน แต่ค่าประมาณทั้งสองวิธีจะแตกต่างกันในกรณีตัวอย่างมีจำนวนน้อย

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

2.3.3 เกณฑ์การเลือกแบบจำลอง (Model Selection Criteria)

ก่อนที่เราจะประมาณค่าแบบจำลอง $ARMA(p,q)$ สำหรับอนุกรมเวลา y_t ใดๆ เราจะเป็นต้องระบุลำดับของ AR (p) และ MA (q) เสียก่อน โดยที่เราสามารถใช้การสังเกต ACF ของตัวอย่างในกรณีของ MA หรือ PACF ของตัวอย่าง ในกรณีของ AR อีกทางเลือกหนึ่งเราสามารถจะใช้เกณฑ์

ตารางที่ 2.2: สรุปลักษณะของ ACF และ PACF สำหรับแบบจำลอง ARMA

	$AR(p)$	$MA(q)$	$ARMA(p,q)$
ACF	ค่อยๆลดลง	ค่าเท่ากับศูนย์หลังจากช่วงล่าที่ q	ค่อยๆลดลง
PACF	เท่ากับศูนย์หลังจากช่วงล่าที่ p	ค่อยๆลดลง	ค่อยๆลดลง

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

$$MSC(p,q) = \ln(\tilde{\sigma}^2(p,q)) + c_T\varphi(p,q) \tag{2.63}$$

โดยที่ $\tilde{\sigma}^2(p,q)$ คือตัวประมาณค่าความควรจะเป็นสูงสุดของตัวแปรปรวนของช็อก c_T คือลำดับที่ขึ้นกับจำนวนตัวอย่าง และ $\varphi(p,q)$ คือฟังก์ชันเบี้ยปรับ (penalty function) กรณีที่เราเลือก p และ q มากๆ โดยที่แบบเกณฑ์ที่เป็นที่นิยมสามเกณฑ์ได้แก่ Akaike (AIC) Schwarz-Bayesian (BIC) และ Hannan-Quinn (HQIC) ซึ่งแสดงได้ด้วยสมการดังนี้

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

โดยที่ในกรณีที่กลุ่มตัวอย่างใหญ่ เกณฑ์ AIC จะเลือกอันดับที่มากกว่าที่ควรจะเป็น ในขณะที่เกณฑ์ BIC และ HQIC เลือกค่าที่ใกล้เคียงความเป็นจริง (consistent) อย่างไรก็ตามในตัวอย่างขนาดเล็กเกณฑ์ทั้งสามจะให้ค่าที่ไม่แตกต่างกันนัก

ตัวอย่างที่ 2.6: การสร้างแบบจำลองสำหรับผลได้ตอบแทนของการซื้อขายเงินตราต่างประเทศโดย ARMA(p,q)

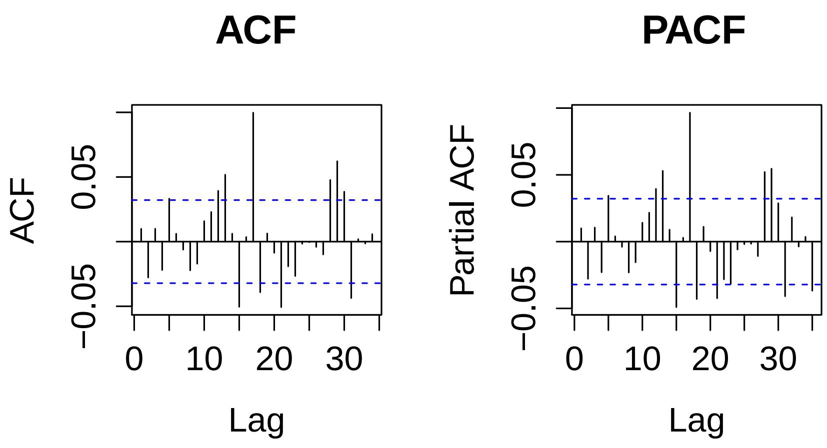
ตัวอย่างนี้เราพิจารณาผลได้ตอบแทนของการซื้อขายเงินตราต่างประเทศโดยคำนวณจากอัตราแลกเปลี่ยนบาทต่อดอลลาร์สหรัฐระหว่างปี 1998 ถึง 2012 ซึ่งอยู่ในไฟล์ `thbusd.csv` เราสามารถนำเข้าข้อมูลและคำนวณผลได้ตอบแทนได้ดังนี้

```
> exc<-
    read.csv("https://raw.githubusercontent.com/chaleampong/EC435/master/thbusd.csv",
    header = TRUE)
> head(exc)
  Jul.Day YYYY.MM.DD Wdy THB.USD
1 2450816   1/2/1998 Fri  48.023
2 2450819   1/5/1998 Mon  50.366
3 2450820   1/6/1998 Tue  52.665
4 2450821   1/7/1998 Wed  52.203
5 2450822   1/8/1998 Thu  53.957
6 2450823   1/9/1998 Fri  53.204
> excret <- diff(log(exc$THB.USD))
> ts.plot(excret, main = "Exchange rate return")
```

เมื่อพิจารณา ACF และ PACF จากรูป **??** จะพบว่าแบบจำลองที่เหมาะสมน่าจะเป็น $ARMA(p,q)$ เนื่องจากฟังก์ชันทั้งสองมิได้ลดลงอย่างมีรูปแบบ ดังนั้นเราจำเป็นต้องเลือกอันดับ p และ q ที่เหมาะสม โดยที่เราสามารถเปลี่ยนค่า p และ q ในคำสั่ง `arima` ไปเรื่อยๆ $(0,0),(1,0),(2,0),(0,1),(0,2),(1,1),(2,2),\ldots$ และบันทึกค่า AIC หรือ BIC ที่ได้แล้วเลือกแบบจำลองที่มีค่า AIC หรือ BIC น้อยที่สุด

อย่างไรก็ตาม `package "forecast"` มีคำสั่ง `auto.arima(series, arguments)` ที่จะทำการทดลองสร้างแบบจำลองตามค่า p และ q สูงสุดที่เรากำหนด แล้วเลือกแบบจำลองตาม

รูปที่ 2.11: ACF และ PACF ของผลได้ตอบแทนรายวันจากการซื้อขายเงินดอลลาร์สหรัฐ



Information criteria ที่เราใส่ arguments ต่อไปนี้ ค่า $d = 0$, $D = 0$, `max.p=6`, `max.q=6`, `max.order=[max.p+max.q]`, `ic=c("aic")` โดยที่เราสามารถเปลี่ยนเป็น bic, `stepwise=FALSE` แปลว่าให้ทดลองทุกค่าของ p และ q ที่น้อยกว่า max.p และ max.q และ `trace=TRUE` แปลว่าให้แสดงค่า IC จากทุกการทดลอง

```
> 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
```

แบบจำลองที่เราเลือกจากค่า AIC น้อยที่สุดคือแบบจำลอง $ARMA(4,1)$ ผู้อ่านอาจจะลองใช้เกณฑ์ BIC แล้วเปรียบเทียบว่าแบบจำลองที่แนะนำต่างกันหรือไม่

2.3.4 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

$$\frac{\theta(L)}{\phi(L)} = 1 + \psi_1 L + \psi_2 L + ... = \psi(L) \tag{2.64}$$

และ

$$\frac{\phi(L)}{\theta(L)} = 1 - \pi_1 L - \pi_2 L - ... = \pi(L) \tag{2.65}$$

2.3.5 การพยากรณ์ (forecasting)

จากรูปแบบของแบบจำลอง $ARMA$ ในหัวข้อที่ผ่านมา เราสามารถแสดงแบบจำลอง $ARMA(p,q)$ ได้สามรูปแบบคือ รูปแบบที่หนึ่งคือรูปแบบ $ARMA(p,q)$

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

รูปแบบที่สองคือการแสดงในรูปมูวิ่งเอเวอเรจอนันต์ (Infinite moving average representation)

$$y_t = \frac{\theta(L)}{\phi(L)}\varepsilon_t = \psi(L)\varepsilon_t$$

และรูปแบบที่สามคือการแสดงในรูปออโตรีเกรสซีฟอนันต์ (Infinite autoregressive representation)

$$\frac{\phi(L)}{\theta(L)}y_t = \pi(L)y_t = \varepsilon_t$$

โดยที่ $\psi(L) = 1 + \psi_1 L + \psi_2 L^2 + ...$ และ $\pi(L) = 1 - \pi_1 - \pi_2 - ...$ ดังนั้นเราสามารถพยากรณ์ตัวแปร y_{h+l} ด้วยสมการใดสมการหนึ่งในสามรูปแบบนี้ได้ ยกตัวอย่างเช่น กรณีใช้รูปแบบ $ARMA$ การพยากรณ์ไปข้างหน้าหนึ่งคาบ ณ ข้อมูลที่คาบ h, $\hat{y}_h(1)$ จะเท่ากับ

$$\begin{aligned} \hat{y}_h(1) &= E(\phi_1 y_h + ...\phi_p y_{h-p} + \varepsilon_{h+1} - \theta_1\varepsilon_h + ... + \theta_q\varepsilon_{h+1-q}|F_h) \\ &= \phi_1 y_h + ...\phi_p y_{h-p} - \theta_1\varepsilon_h + ... + \theta_q\varepsilon_{h+1-q} \end{aligned} \tag{2.66}$$

กรณีใช้รูปแบบออโตรีเกรสซีฟอนันต์ การพยากรณ์ไปข้างหน้าหนึ่งคาบ ณ ข้อมูลที่คาบ h, $\hat{y}_h(1)$ จะเท่ากับ

$$\begin{aligned}\hat{y}_h(1) &= E(\pi_1 y_h + \pi_2 y_{h-1} + ...|F_h) \\ &= \pi_1 y_h + \pi_2 y_{h-1} + ...\end{aligned} \tag{2.67}$$

และกรณีการแสดงในรูปมูวิ่งเอเวอเรจอนันต์ การพยากรณ์ไปข้างหน้าหนึ่งคาบ ณ ข้อมูลที่คาบ h, $\hat{y}_h(1)$ จะเท่ากับ

$$\begin{aligned}\hat{y}_h(1) &= E(\varepsilon_{h+1} + \psi_1 \varepsilon_h + \pi_2 \varepsilon_{h-1} + ...|F_h) \\ &= \psi_1 \varepsilon_h + \pi_2 \varepsilon_{h-1} + ...\end{aligned} \tag{2.68}$$

2.4 Unit Root Nonstationary

อนุกรมเวลาทางเศรษฐกิจและการเงินจำนวนมากมีพฤติกรรมที่มีแนวโน้ม(trend)หรือมีลักษณะไม่นิ่ง (non-stationary) ตัวอย่างเช่น ราคาสินทรัพย์ อัตราแลกเปลี่ยน หรือตัวแปรเศรษฐกิจมหภาคเช่น ผลิตภัณฑ์มวลรวมภายในประเทศ (GDP) ดังนั้นในการวิเคราะห์อนุกรมเวลาเราพิจารณาลักษณะของรูปแบบแนวโน้มของข้อมูลว่าเกิดจากความไม่นิ่งหรือเป็นแนวโน้ม เนื่องจากคุณลักษณะดังกล่าวจะส่งผลกระทบต่อการนำข้อมูลไปใช้ เช่นในการสร้างแบบจำลอง $ARMA$ เราจำเป็นต้องทำการแปลงข้อมูลในนิ่งเสียก่อน ในขณะที่ถ้าข้อมูลมีแนวโน้มเราจำเป็นต้องขจัดแนวโน้มเสียก่อน โดยที่กระบวนการที่มีลักษณะแนวโน้มอาจจะแบ่งได้เป็นสองกรณีคือกระบวนการแนวเดินแบบสุ่ม (random walk) แบบมีแนวโน้มด้วย (with drift) และกระบวนการที่นิ่งหลังจากกำจัดแนวโน้ม (trend stationary)

กรณีแรก คือกรณีที่ y_t เป็นกระบวนการที่นิ่งหลังจากดำเนินการผลต่าง (difference-stationary) หากกระบวนการ y_t เป็นกระบวนการแนวเดินแบบสุ่มแบบมีแนวโน้มด้วย จะสามารถเขียนสมการอธิบาย y_t ได้ดังนี้

$$y_t = \mu + y_{t-1} + \varepsilon_t \tag{2.69}$$

โดย $\varepsilon_t \sim WN(0, \sigma^2)$ และเราเรียก μ ว่าตัวสร้างแนวโน้ม (drift) หากเราสมมุติให้ $y_0 = 0$ และแทนค่าแบบ recursive เราจะสามารถเขียนสมการ (**??**) ได้เป็น

$$y_t = \mu t + \sum_{j=1}^{t} \varepsilon_j \tag{2.70}$$

จากสมการที่ (**??**) กระบวนการแนวเดินแบบสุ่มแบบมีแนวโน้มจะมีค่าคาดหมาย ($E(y_t)$) เท่ากับ μt และค่าความแปรปรวน ($Var(y_t)$) เท่ากับ $\sigma^2 t$ จะเห็นได้ว่าค่าทั้งสองขึ้นอยู่กับเวลา t ดังนั้น y_t จากสมการ (**??**) จะเป็นกระบวนการที่ไม่นิ่ง (non-stationary)

กรณีสอง คือกรณีที่ z_t เป็นกระบวนการที่นิ่งหลังจากกำจัดแนวโน้ม(trend-stationary) เรา

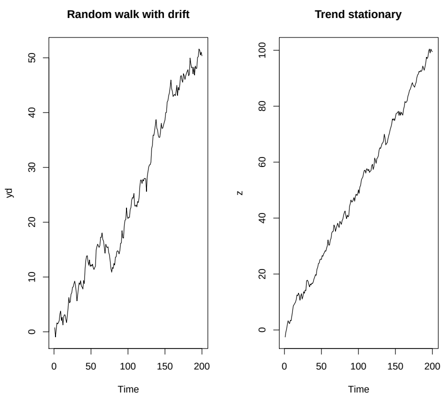
สามารถเขียนสมการอธิบาย z_t ได้ด้วย

$$z_t = \beta_0 + \beta_1 t + v_t \tag{2.71}$$

โดยที่ v_t เป็นอนุกรมนิ่งใดๆเช่น กระบวนการไวท์น๊อซหรือ $AR(1)$ ที่มีค่าสัมบูรณ์ของสัมประสิทธิ์น้อยกว่า 1 จากสมการที่ (**??**) เราสามารถพิสูจน์ได้ว่าค่าคาดหมายของ ($E(z_t)$) จะเท่ากับ $\beta_0 + \beta_1 t$ ซึ่งเป็นฟังก์ชันเส้นตรงกับเวลา ดังนั้น z_t จึงเป็นกระบวนการ non-stationary โดยที่ z_t จะแกว่งตัวอยู่รอบๆเส้นแนวโน้ม ในขณะที่ ค่าความแปรปรวน ($Var(z_t)$) จะเท่ากับ $Var(v_t)$ ซึ่งจำกัดและไม่แปรผันตามเวลา (เนื่องจากเราสมมุติให้ v_t เป็นอนุกรมนิ่ง)

ตัวอย่างของกราฟข้อมูลกระบวนการนิ่งหลังจากดำเนินการผลต่าง (difference-stationary) และกระบวนการนิ่งหลังจากกำจัดแนวโน้ม (trend-stationary) ถูกแสดงไว้ในรูปที่ **??** โดยมองดูเผินๆจะเห็นได้ว่าภาพทั้งสองมีลักษณะที่คล้ายคลึงกัน

รูปที่ 2.12: ข้อมูลจำลองของกระบวนการแนวเดินแบบสุ่มแบบมีแนวโน้มและกระบวนการที่นิ่งหลังจากขจัดแนวโน้ม



กระบวนการทั้งสองที่ได้ยกตัวอย่างมา เป็นกระบวนการที่ไม่นิ่ง (non-stationary) แต่มีคุณลักษณะที่แตกต่างกันดังนั้น วิธีการที่ใช้ในทำให้อนุกรมนิ่งจึงแตกต่างกัน โดยวิธีการที่ใช้ทำให้อนุกรมนิ่งมีอยู่สองวิธีคือ การดำเนินการผลต่าง (differencing) และการประมาณค่าถดถอยเพื่อกำจัดแนวโน้ม (detrended series)

วิธีการดำเนินการผลต่าง คือการนำข้อมูลที่คาบ t ลบออกด้วยข้อมูลที่คาบ $t-1$ หรือเขียนเป็นสัญลักษณ์ได้โดย $\Delta y_t = y_t - y_{t-1}$ ซึ่งเราเรียกกรณีนี้ว่าการดำเนินการผลต่างครั้งที่หนึ่ง (first difference) หากเรานำอนุกรมที่ได้จากการดำเนินการผลต่างครั้งที่หนึ่งไปดำเนินการผลต่างอีกครั้ง $\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$ เราจะเรียกว่า การดำเนินการผลต่างครั้งที่สอง (second difference)วิธีการดำเนินการผลต่างจะเหมาะกับอนุกรมเวลาที่เป็นกระบวนการนิ่งหลังจากดำเนินการผลต่าง โดยที่หากอนุกรมเวลา y_t เป็นอนุกรมที่นิ่งหลังจากการดำเนินการผลต่างครั้งที่หนึ่ง เราจะเรียกอนุกรมเวลา y_t ว่าอนุกรมเวลาที่มีอันดับอินทิเกรตเท่ากับหนึ่ง (Integrated of order 1 หรือ $I(1)$) หากอนุกรมเวลา y_t เป็นอนุกรมที่นิ่งหลังจากการดำเนินการผลต่างครั้งที่สอง เราจะเรียกอนุกรมเวลา y_t ว่า

อนุกรมเวลาที่มีอันดับอินทิเกรตเท่ากับสอง (Integrated of order 2 หรือ $I(2)$) ในขณะที่อนุกรมเวลานิ่งจะคือ $I(0)$

วิธีการที่สองเหมาะกับอนุกรมที่เป็นกระบวนการนิ่งหลังจากกำจัดแนวโน้ม โดยเราจะประมาณค่าสมการ **??** ด้วย OLS เมื่อได้ค่า $\widehat{\beta}_0$ และ $\widehat{\beta}_1$ เราสามารถสร้างอนุกรมเวลาที่ขจัดแนวโน้มออกไปได้โดยการคำนวณ $z_t - \widehat{\beta}_0 - \widehat{\beta}_1 t$ โดยเครื่องมือสำคัญที่ใช้ในการระบุว่าควรจะใช้วิธีการใดในการกำจัดแนวโน้มคือ การทดสอบยูนิทรูท (unit root test)

เพื่อ ให้ เรา เข้าใจ แนวคิด พื้น ฐาน ใน การ ทดสอบ ยู นิ ทรูท และ การ ทดสอบ ความ นิ่ง (stationary) เราพิจารณาอนุกรม y_t ซึ่งสามารถแยกได้เป็นส่วนของแนวโน้มและวัฏจักรโดยที่

$$
\begin{aligned}
y_t &= TD_t + z_t \\
TD_t &= k + \delta t \\
z_t &= \phi z_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim WN(0, \sigma^2)
\end{aligned}
$$

โดยที่ TD_t เป็นแนวโน้มเส้นตรงเชิงกำหนด (deterministic linear trend) และ z_t คือกระบวนการ $AR(1)$ ถ้า $|\phi| < 1$ แล้ว y_t จะเป็นอนุกรมนิ่ง (I(0)) รอบๆแนวโน้มเชิงกำหนด TD_t แต่ถ้า $\phi = 1$ แล้วอนุกรม z_t จะกลายเป็นแนวโน้มสโทแคสติก และ y_t เป็น I(1) ที่มี drift

การทดสอบยูนิทรูทในรูปของแบบถดถอยในตัวเองมีพื้นฐานจากการทดสอบสมมุติฐานหลักว่า $\phi = 1$ (อนุกรมนิ่งหลักจาก differencing) กับสมมุติฐานทางเลือก $\phi < 1$ (อนุกรมนิ่งหลังจากขจัดแนวโน้ม) โดยที่เราเรียกการทดสอบนี้ว่ายูนิทรูท เนื่องจากภายใต้สมมุติฐานหลักรากของพหุนาม $\phi(z) = (1 - \phi z) = 0$ มีค่าเท่ากับหนึ่ง

2.4.1 การทดสอบยูนิทรูทในรูปสมการถดถอยในตัวเอง

สมมุติว่าเราพิจารณาแบบจำลอง $AR(1)$ ซึ่งเขียนได้ดังนี้

$$
y_t = \phi y_{1-t} + \varepsilon_t, \qquad \varepsilon_t \sim WN(0, \sigma^2) \tag{2.72}
$$

โดยที่สมมุติฐานที่เราต้องการทดสอบคือ

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

เราสามารถประมาณค่าสมการ **??** ด้วย OLS และใช้ตัวสถิติที่เรียกว่าตัวสถิติ Dickey-Fuller t ซึ่งคำนวณได้โดย

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

โดยที่ $\hat{\phi}$ คือค่าประมาณจากวิธีการกำลังสองน้อยที่สุดและ $se(\hat{\phi})$ คือค่าผิดพลาดมาตรฐาน (standard error) โดยที่การทดสอบนี้เป็นการทดสอบหางเดียวทางซ้าย คือเราจะปฏิเสธสมมุติฐานหลักถ้าค่า-ตัวสถิติน้อยกว่าค่าวิกฤติ และสรุปว่าอนุกรมเวลาเป็นอนุกรมนิ่ง หากเราไม่สามารถปฏิเสธสมมุติ-ฐานหลักได้ แสดงว่าอนุกรมดังกล่าวเป็นยูนิทรูท นอกจากนี้เราสามารถใช้ตัวสถิติ normalized bias ซึ่งคำนวณได้ด้วย $T(\hat{\phi}-1)$ ในการทดสอบยูนิทรูทเช่นเดียวกัน

สมการ **??** ที่เราใช้ในการทดสอบยูนิทรูทยังสามารถเขียนในรูป

$$\Delta y_t = \pi y_{1-t} + \varepsilon_t, \tag{2.73}$$

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

ตัวสถิติที่ใช้ในการทดสอบยูนิทรูทจะมีการแจกแจงที่ไม่เป็นมาตรฐานและไม่เป็นการแจก-แจงแบบปกติเมื่อข้อมูลขนาดใหญ่เหมือนในกรณีทดสอบอื่นๆที่เราเคยพบมา Phillips (1987) ได้พิสูจน์ว่าหาก y_t เป็นยูนิทรูท การแจกแจงของตัวสถิติทั้งสองที่ใช้ทดสอบยูนิทรูทจะมีการแจกแจงดังนี้

$$t_{\phi=1} \xrightarrow{d} \frac{\int_0^1 W(r)dW(r)}{\left(\int_0^1 W(r)^2 dr\right)^{1/2}} \tag{2.74}$$

$$T(\hat{\phi}-1) \xrightarrow{d} \frac{\int_0^1 W(r)dW(r)}{\int_0^1 W(r)^2 dr} \tag{2.75}$$

โดยที่ $W(r)$ คือการเคลื่อนที่แบบบราวเนียน (Brownian motion) จากสมการข้างต้น ตัวประมาณค่าและตัวสถิติจะไม่มีการแจกแจงเข้าสู่การแจกแจงปกติ เราจะเรียกการแจกแจงของ $t_{\phi=1}$ ว่า การแจกแจง Dickey-Fuller t และการแจกแจงของ $T(\hat{\phi}-1)$ ว่าการแจกแจง Dickey-Fuller normalized bias

ตาราง สำหรับ การ แจกแจง Dickey-Fuller สามารถ ใช้ คำ สั่ง `adfTable` ใน `library(fUnitRoots)` โดยระบุรูปแบบของสมการที่ใช้ในการทดสอบ เช่น `trend=c("nc")` กรณีที่สมการทดสอบไม่มีพจน์ค่าคงที่อยู่ (no constant)เช่นในสมการ **??** (ในหัวข้อย่อยต่อไปจะพิ-จารณากรณีมีพจน์ค่าคงที่ 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

attr(,"class")
[1] "gridData"
attr(,"control")
    table     trend statistic
    "adf"      "nc"       "t"
```

โดยที่ x เป็นจำนวนตัวอย่างและ y เป็นความถี่สะสม และค่า critical value จะอยู่ในส่วน z ตัวอย่างเช่นกลุ่มตัวอย่าง 100 ตัวอย่างที่ค่า ADF เท่ากับ -1.95 จะมีความถี่สะสมเท่ากับ 0.05 หรือพื้นที่ใต้กราฟการแจกแจง Dickey-Fuller ทางซ้ายของ -1.95 จะเท่ากับ 5 %

ตารางที่ 2.3: ค่าวิกฤตของการแจกแจง Dickey-Fuller t statistics กรณีไม่มีค่าคงที่

จำนวนตัวอย่าง	ความถี่สะสม							
	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
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
∞	-2.58	-2.23	-1.95	-1.62	0.89	1.28	1.62	2.00

กรณีอนุกรมเวลามีค่าคงที่และแนวโน้ม

สิ่งสำคัญในการทดสอบยูนิทรูทคือการระบุว่าสมมุติฐานหลักและสมมุติฐานทางเลือกมีความเหมาะสมกับข้อมูลหรือ โดยเฉพาะแนวโน้ม โดยกรณีที่เราพบบ่อยมีสองกรณีคือกรณีมีค่าคงที่อย่างเดียว และกรณีมีค่าคงที่และแนวโน้ม

ในกรณีค่าคงที่อย่างเดียว สมการที่ใช้ในการประมาณค่าตัวสถิติทดสอบคือ

$$y_t = c + \phi y_{t-1} + \varepsilon_t \tag{2.76}$$

และสมมุติฐานที่เกี่ยวเนื่องคือ

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

ในกรณีที่มีค่าคงที่และแนวโน้ม สมการที่ใช้ในการประมาณค่า

$$y_t = c + \delta t + \phi y_{t-1} + \varepsilon_t \tag{2.77}$$

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

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

สำหรับสมการที่มีค่าคงที่และมีค่าคงที่และแนวโน้มตามลำดับ ดังนั้นการตัดสินใจว่าจะใช้สมการรูปแบบไหน สามารถทำได้โดยการพิจารณากราฟเส้นของ Δy_t มีค่าเฉลี่ยต่างจากศูนย์หรือมีเส้นแนวโน้มหรือไม่ โดยที่ตารางสำหรับการแจกแจง Dickey-Fuller ของแต่ละรูปแบบของสมการจะมีค่าที่แตกต่างกัน โดยที่กรณีมีค่าคงที่จะใช้ตัวเลือก `trend=c("c")` และกรณีมีค่าคงที่และเส้นแนวโน้มจะใช้ `trend=c("ct")` ในคำสั่ง

```
> adfTable(trend=c("c"), statistic=c("t"))
> adfTable(trend=c("ct"), statistic=c("t"))
```

แล้วจะได้ตารางดังนี้

การทดสอบยูนิทรูทดิกกี ฟูลเลอร์แบบแต่งเติม(Augmented Dickey-Fuller)

การทดสอบยูนิทรูทที่ได้กล่าวไว้ข้างต้นนั้นจะถูกต้องถ้าอนุกรมเวลาเป็น $AR(1)$ ที่มีช็อกเป็นไวทน้อซ Said and Dickey (1984) ได้เสริมการทดสอบยูนิทรูทด้วยพจน์ที่เพิ่มเข้าไปโดยมีพื้นฐานจากแบบจำลอง $ARMA(p,q)$ แล้วเรียกว่า การทดสอบยูนิทรูทดิกกี ฟูลเลอร์แบบแต่ง

ตารางที่ 2.4: ค่าวิกฤตของการแจกแจง Dickey-Fuller กรณีมีค่าคงที่

จำนวน	ความถี่สะสม							
ตัวอย่าง	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
25	-3.75	-3.33	-3.00	-2.63	-0.37	0.00	0.34	0.72
50	-3.58	-3.22	-2.93	-2.60	-0.40	-0.03	0.29	0.66
100	-3.51	-3.17	-2.89	-2.58	-0.42	-0.05	0.26	0.63
250	-3.46	-3.14	-2.88	-2.57	-0.42	-0.06	0.24	0.62
500	-3.44	-3.13	-2.87	-2.57	-0.43	-0.07	0.24	0.61
∞	-3.43	-3.12	-2.86	-2.57	-0.44	-0.07	0.23	0.60

ตารางที่ 2.5: ค่าวิกฤตของการแจกแจง Dickey-Fuller กรณีมีค่าคงที่และแนวโน้ม

จำนวน	ความถี่สะสม							
ตัวอย่าง	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
25	-4.38	-3.95	-3.60	-3.24	-1.14	-0.80	-0.50	-0.15
50	-4.15	-3.80	-3.50	-3.18	-1.19	-0.87	-0.58	-0.24
100	-4.04	-3.73	-3.45	-3.15	-1.22	-0.90	-0.62	-0.28
250	-3.99	-3.69	-3.43	-3.13	-1.23	-0.92	-0.64	-0.31
500	-3.98	-3.68	-3.42	-3.13	-1.24	-0.93	-0.65	-0.32
∞	-3.96	-3.66	-3.41	-3.12	-1.25	-0.94	-0.66	-0.33

เติม(Augmented Dickey-Fuller) ซึ่งตัวทดสอบสามารถประมาณค่าได้จากสมการ

$$y_t = \boldsymbol{\beta}^{'}(D_t) + \phi y_{t-1} + \sum_{j=1}^{p} \psi_j \Delta y_{t-j} + \varepsilon_t \tag{2.78}$$

หรือเราสามารถประมาณค่าสมการนี้

$$\Delta y_t = \beta^{'} D_t + \pi y_{t-1} + \sum_{j=1}^{p} \psi_j \Delta y_{t-j} + \varepsilon_t \tag{2.79}$$

โดยที่ $\pi = \phi - 1$ ซึ่งภายใต้สมมุติฐานหลัก Δy_t จะเป็น $I(0)$ และอนุมานได้ว่า $\pi = 0$ ดังนั้น ตัวสถิติ t สำหรับยูนิทรูทด้วยดิกกี ฟูลเลอร์แบบแต่งเติม จะเป็นค่าทีของสัมประสิทธิ์หน้า y_t

ตัวอย่างที่ 2.7 การทดสอบ Augmented Dickey Fuller Unit root

ตัวอย่าง นี้ เรา ต้องการ ทดสอบ ยู นิ ทรูท ของ SET index ราย วัน จาก ข้อมูล ใน ไฟล์ `setindex.txt` ซึ่งมีตัวอย่างเท่ากับ 3642 ตัวอย่าง โดยที่ใน โปรแกรมอาร์มีชุดคำสั่งหลายชุดคำสั่งที่สามารถทำการทดสอบ ADF ได้เช่น `urca`, `stats`, `tseries` แต่ในที่นี้เราจะใช้ชุดคำสั่ง

urca โดยเบื้องต้นเราจะพิจารณา ACF ของอนุกรมเวลาและอนุกรมเวลาที่ได้ปรับผลต่างอันดับหนึ่ง (first-differenced) จาก ACF ของอนุกรมเวลา setd$index เราพบว่าน่าจะมีปัญหายูนิทรูท เนื่องจากค่า ACF ลดลงช้ามาก

```
> setd <-
    read.csv("https://raw.githubusercontent.com/chaleampong/EC435/master/setdaily.csv",
    header = TRUE)
> head(setd)
  X      date  index
1 1  1/5/1998 366.18
2 2  1/6/1998 370.27
3 3  1/7/1998 370.31
4 4  1/8/1998 360.17
5 5  1/9/1998 349.67
6 6 1/12/1998 339.17
> ts.plot(setd$index)
> acf(setd$index)
```

เราสามารถทดสอบ unit root ด้วยคำสั่ง ur.df จาก package urca ซึ่งเป็นการทดสอบ Dickey Fuller Unit root โดยเรากำหนดตัวแปรที่ต้องการทดสอบ setd$index แล้วเลือกรูปแบบของสมการที่เราใช้ทดสอบซึ่งสามารถเลือกได้เป็นสมการที่ไม่มีค่าคงที่ (none) มีค่าคงที่ (drift) และมีเส้นแนวโน้ม (trend) ด้วยการระบุใน type เช่น `type=c("trend"`) แล้วเลือกว่าจะมี lag ของตัวแปรตามเท่ากับเท่าใด ซึ่งมีทางเลือกในการระบุจำนวนเลข เช่น lags=1 หรือใช้ model selection criteria เช่น selectlags = c("AIC")

ในกรณีแรกที่ระบุ lag=1 เราเก็บผลไว้ในชื่อ setd.df ค่าสถิติที่ได้คือ -1.863 ซึ่งมากกว่า critical value ที่ $\alpha = 0.05$ เท่ากับ -3.41 ดังนั้น เราไม่สามารถปฏิเสธสมมุติฐานที่ว่า setd เป็น unit root

```
> library(urca)
> setd.df<-ur.df(setd$index, type=c("trend"), lags=1 )
> summary(setd.df)
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################
Test regression trend
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
Residuals:
     Min       1Q   Median       3Q      Max
-108.380   -4.745   -0.105    4.837   74.125
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.4325591  0.4122648   1.049  0.29414
z.lag.1     -0.0021279  0.0011419  -1.864  0.06247 .
tt           0.0006223  0.0002840   2.191  0.02851 *
z.diff.lag   0.0478816  0.0165720   2.889  0.00388 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘’. 0.1 ‘’  1

Residual standard error: 9.355 on 3636 degrees of freedom
Multiple R-squared:  0.003526,  Adjusted R-squared:  0.002704
F-statistic: 4.289 on 3 and 3636 DF,  p-value: 0.004982

Value of test-statistic is: -1.8635 2.3752 2.4005

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.96 -3.41 -3.12
phi2  6.09  4.68  4.03
phi3  8.27  6.25  5.34
```

ในกรณีแรกที่ระบุ `selectlags = c("AIC")` เราเก็บผลไว้ในชื่อ `setd.df2` ค่าสถิติที่ได้คือ -1.864 ซึ่งมากกว่า critical value ที่ $\alpha = 0.05$ เท่ากับ -3.41 ดังนั้น เราไม่สามารถปฏิเสธสมมุติฐานที่ว่า setd เป็น unit root

```
> setd.df2<-ur.df(setd$index, type = c("trend"), selectlags = c("AIC"))
> summary(setd.df2)
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################
Test regression trend
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
Residuals:
     Min       1Q   Median       3Q      Max
-108.380   -4.745   -0.105    4.837   74.125
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.4325591  0.4122648   1.049  0.29414
z.lag.1     -0.0021279  0.0011419  -1.864  0.06247 .
tt           0.0006223  0.0002840   2.191  0.02851 *
z.diff.lag   0.0478816  0.0165720   2.889  0.00388 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 ''. 0.1 ''  1
Residual standard error: 9.355 on 3636 degrees of freedom
Multiple R-squared:  0.003526,  Adjusted R-squared:  0.002704
F-statistic: 4.289 on 3 and 3636 DF,  p-value: 0.004982

Value of test-statistic is: -1.8635 2.3752 2.4005

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.96 -3.41 -3.12
phi2  6.09  4.68  4.03
phi3  8.27  6.25  5.34
```

เราได้ค่า ADF t statistics เท่ากับ -0.68 เมื่อเปรียบเทียบกับค่าวิกฤตจากตาราง **??** ที่นัยสำคัญ 5 % และ $N = \infty$ มีค่าเท่ากับ -2.86 จะเห็นได้ว่า ADF t statistics มีค่ามากกว่าค่าวิกฤต ดังนั้นเราไม่สามารถปฏิเสธสมมุติฐานที่ว่า set index เป็นยูนิทรูท หรือเราสามารถพิจารณาค่าพีซึ่งมีค่าเท่ากับ 0.95 ซึ่งเราไม่สามารถปฏิเสธว่า set index เป็นยูนิทรูท

ตัวทดสอบยูนิทรูทของฟิลิปส์ เปอรรอง

Phillips and Perron (1998) ได้พัฒนาการทดสอบยูนิทซึ่งเป็นการทดสอบอีกวิธีหนึ่งที่เป็นที่นิยม โดยการทดสอบฟิลิปส์-เปอรรอง มีความแตกต่างจากการทดสอบยูนิทรูทดิกกี ฟูลเลอร์แบบแต่งเติมที่สำคัญค่อการดำเนินการที่เกี่ยวข้องกับปัญหา serial correlation และ heteroskedasticity ในช็อก โดยที่ในกรณีของการทดสอบ ADF จะใช้โครงสร้างที่มีรูปแบบ ARMA ในการประมาณตัวแปรช็อก ในขณะที่การทดสอบ PP เพิกเฉยต่อ serial correlation ในสมการทดสอบ โดยใช้สมการต่อไปนี้ในการประมาณค่า

$$\Delta y_t = \boldsymbol{\beta}' \boldsymbol{D}_t + \pi y_{t-1} + u_t \tag{2.80}$$

โดยที่ u_t เป็น $I(0)$ และอาจจะมีปัญหา heteroskedasticity การทดสอบ PP จะแก้ไขปัญหา serial correlation และ heteroskedasticity ในตัวแปรช็อก u_t โดยการปรับตัวสถิติทดสอบ $t_{\pi=0}$ และ $T\hat{\pi}$

โดยตัวสถิติที่ปรับจะอยู่ในรูปดังต่อไปนี้

$$Z_t = \left(\frac{\hat{\sigma}^2}{\hat{\lambda}^2}\right)^{1/2} \times t_{\pi=0} - \frac{1}{2}\left(\frac{\hat{\lambda}^2 - \hat{\sigma}^2}{\hat{\lambda}^2}\right) \times \left(\frac{T \times se(\hat{\pi})}{\hat{\sigma}^2}\right)$$

$$Z_\pi = T\hat{\pi} - \frac{1}{2}\frac{T^2 \times se(\hat{\pi})}{\hat{\sigma}^2}(\hat{\lambda}^2 - \hat{\sigma}^2)$$

โดยที่ $\hat{\sigma}^2$ และ $\hat{\lambda}^2$ เป็นตัวประมาณค่าที่ consistent ของพารามิเตอร์ค่าความแปรปรวน

$$\sigma^2 = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E[u_t^2]$$

$$\lambda^2 = \lim_{T\to\infty} \sum_{t=1}^{T} E[T^{-1} S_t^2]$$

โดยที่ $S_T = \sum_{t=1}^{T} u_t$

ภายใต้สมมุติฐานหลักที่ว่า $\pi = 0$ ตัวสถิติ Z_t และ Z_π ของ PP มีการแจกแจงที่ลู่เข้าหา ADF t-stat และ Normalized bias statistics เมื่อขนาดตัวอย่างใหญ่ ข้อได้เปรียบของ PP เมื่อเทียบกับ ADF คือ PP พิจารณาความน่าจะเป็นที่จะเกิด heteroskedasticity ของช็อก นอกจากนี้ผู้ใช้ไม่จำเป็นต้องกำหนดอันดับของตัวแปร y_t ในสมการทดสอบเช่นเดียวกับ ADF

ตัวอย่างที่ 2.8 การทดสอบ Phillips Perron Unit root

```
> setd.pp<-ur.pp(setd$index, type="Z-tau", model="trend")
> summary(setd.pp)

##################################
# Phillips-Perron Unit Root Test #
##################################

Test regression with intercept and trend


Call:
lm(formula = y ~ y.l1 + trend)

Residuals:
     Min       1Q   Median       3Q      Max
-108.674   -4.712   -0.062    4.765   68.930

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.4901935  0.7300869   2.041   0.0413 *
y.l1        0.9980139  0.0011418 874.061   <2e-16 ***
trend       0.0005988  0.0002841   2.108   0.0351 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘’. 0.1 ‘’  1

Residual standard error: 9.363 on 3638 degrees of freedom
Multiple R-squared:  0.9987,  Adjusted R-squared:  0.9987
F-statistic: 1.42e+06 on 2 and 3638 DF,  p-value: < 2.2e-16


Value of test-statistic, type: Z-tau  is: -1.8587

           aux. Z statistics
Z-tau-mu              1.7473
Z-tau-beta            2.1804

Critical values for Z statistics:
                     1pct      5pct     10pct
critical values -3.966098 -3.413711 -3.128565
```

ค่าสถิติ Phillips Perron มีค่าเท่ากับ -1.8587 มากกว่า critical value แสดงว่าเราไม่สามารถปฏิเสธสมมุติฐานว่า `setd$index` เป็น unit root

ตัวทดสอบความนิ่ง(Stationary tests)

ในการทดสอบยูนิทรูทด้วย ADF หรือ PP จะพิจารณาสมมุติฐานหลักว่าอนุกรมเวลา y_t เป็น $I(1)$ ในทางตรงข้าม การทดสอบอนุกรมนิ่งมีสมมติฐานหลักว่า y_t เป็น $I(0)$ โดยที่ตัวทดสอบที่เป็นที่นิยมคือการทดสอบ KPSS ซึ่งเสนอโดย Kwiatkowski, Phillips, Schmidt, and Shin (1992)

2.5 ARIMA (Integrated ARMA)

นิยาม 2.3. *เราจะเรียกกระบวนการ y_t ว่า $ARIMA(p,d,q)$ ถ้า $\Delta^d y_t = (1-L)^d y_t$ เป็นกระบวนการ $ARMA(p,q)$ หรือเราสามารถเขียนในรูปทั่วไปเป็น*

$$\phi(L)(1-L)^d y_t = \theta(L)\varepsilon_t \qquad (2.81)$$

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

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

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

ในที่นี้เราอาจจะนึกถึงแบบจำลอง $ARIMA(p,d,q)$ ว่าเป็นแบบจำลอง ARMA(p,q) ของอนุกรมเวลาที่ได้ดำเนินการผลต่างไปแล้ว d ครั้ง โดยที่ $z_t = \Delta^d$ และ $\phi(L)z_t = \theta(L)\varepsilon_t$ โดยทั่วไปแล้ว d จะมีค่าไม่เกิน 3

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

1. พิจารณา ACF ของข้อมูลว่ามีข้อบ่งชี้ว่ามีปัญหาข้อมูลไม่นิ่ง (nonstationary) หรือไม่
2. ทดสอบยูนิทรูทว่ามีอินทิเกรตที่อันดับเท่าใด หรือหาค่า d นั่นเอง
3. หลังจากนั้นเราจะพิจารณาอนุกรม $z_t = \Delta^d y_t$ ซึ่งเป็นอนุกรมเวลาที่นิ่งแล้วหลังจากได้ดำเนินการผลต่างไปแล้ว d ครั้ง ว่าแบบจำลองที่เหมาะสมในการอธิบาย z_t ควรจะมีอันดับของ p และ q เท่าใด โดยที่เราอาจจะพิจารณา ACF และ PACF ของ z_t หรือใช้ information criteria ก็ได้

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

ตัวอย่าง

ต่อจากตัวอย่างที่ **??** เราพิจารณาว่า set index เป็นอินทิเกรตที่อันดับเท่าไหร่ โดยการดำเนินการหาผลต่างอันดับหนึ่ง(first differenced) แล้วทดสอบยูนิทรูท พบว่าอนุกรมเวลาที่ดำเนินการหาผลต่างอันดับหนึ่งไม่เป็นยูนิทรูท ดังนั้น set index เป็น I(1) และในแบบจำลอง ARIMA เราจะกำหนดให้ $d = 1$ ซึ่งเราจะใส่ค่าดังกล่าวในคำสั่ง `auto.arima` เพื่อหา p และ q ที่เหมาะสมต่อไป

ตัวอย่างที่ 2.9 การสร้างแบบจำลอง ARIMA

ขั้น ตอน แรก ของ การ สร้าง แบบ จำลอง คือ กการ หา ลำดับ ขอ งอิ นทิ เก ร ชัน (order of integration) หรือค่า d

จากตัวอย่างที่ผ่านมาเราทราบว่า `setd$index` เป็น unit root ดังนั้น เราแปลงข้อมูลด้วย first difference ($\Delta setd_t = setd_t - setd_{t-1}$) แล้วทดสอบ Unit root

```
> setd.df.diff<-ur.df(diff(setd$index), type = c("trend"), selectlags = c("AIC"))
> summary(setd.df.diff)
<Omitted>
Value of test-statistic is: -39.9724 532.5983 798.8972

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.96 -3.41 -3.12
phi2  6.09  4.68  4.03
phi3  8.27  6.25  5.34
```

จากการทดสอบ unit root กับตัวแปรในรูป first difference ค่าสถิติเท่ากับ -39.97 ซึ่งน้อยกว่า critical value -3.41 ดังนั้นเราสามารถปฏิเสธสมมุติฐานหลักที่ว่า diff(setd) เป็น unit root แสดงว่า diff(setd) เป็น stationary และเราสามารถสรุปได้ว่า setd เป็น I(1) ซึ่งต้องทำ first difference ถึงจะเป็น stationary

แบบจำลองที่เหมาะสมกับ setd คือแบบจำลองที่ d=1 ดังนั้นเรากำหนดค่าดังกล่าวในคำสั่ง `auto.arima` และให้คำสั่งทดลองหา p, q ที่เหมาะสม ซึ่งในกรณีนี้ p

```
> library(forecast)
> setd_mod<-auto.arima(setd$index, d=1, max.p=6, max.q=6 , ic= c("bic"), stepwise
    = FALSE, trace = TRUE)
<Omitted>
 Best model: ARIMA(0,1,0)

> setd_mod
Series: setd$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

2.6 แบบจำลองตามฤดูกาล

อนุกรมเวลาทางการเงินบางอนุกรมเช่น ผลตอบแทนต่อหุ้น (earning per share) มีพฤติกรรมเป็นวงจรหรือเป็นช่วงเวลา (cyclical) โดยที่ในบางไตรมาสอาจจะมีผลตอบแทนต่อหุ้นที่สูง

กว่าไตรมาสอื่น ดังนั้นในการสร้างแบบจำลองเพื่ออธิบายอนุกรมดังกล่าวจำเป็นต้องผลของฤดูกาลด้วย

สมมุติให้อนุกรมเวลามีลักษณะตามฤดกาล s เราสามารถกำจัดผลขอฤดูกาลได้โดยการจัดการผลต่างฤดูกาล (seasonal differencing) หรือเขียนในรูปที่ใช้ ตัวดำเนินการความล่า-s (lag-s operator) ในพจน์สุดท้ายของสมการต่อไปนี้

$$\Delta_s y_t = y_t - y_{t-s} = y_t - L^s y_t = (1 - L^s) y_t \tag{2.82}$$

จากตัวดำเนินการดังกล่าวนำมาสู่แนวคิดในการสร้างแบบจำลอง seasonal $ARMA(P,Q)_s$ ซึ่งสามารถแสดงได้ด้วยสมการต่อไปนี้

$$\Phi_P(L^s) y_t = \Theta_Q(L^s) \varepsilon_t \tag{2.83}$$

โดยที่ $\Phi_P(L^s) = 1 - \Phi_1 L^s - \Phi_2 L^{2s} - ... - \Phi_P L^{Ps}$ และ $\Theta_Q(L^s) = 1 + \Theta_1 L^s - \Theta_2 L^{2s} - ... - \Theta_Q L^{qs}$ โดยที่คุณสมบัติคล้ายกับ $ARMA(p,q)$ คืออนุกรมเวลาจะนิ่งถ้ารากของ $\Phi_P(z^s)$ อยู่นอกวงกลมหนึ่งหน่วย และสามารถเขียนหาตัวผกผันได้ถ้ารากของ $\Theta_Q(z^s)$ อยู่นอกวงกลมหนึ่งหน่วย

ตัวอย่างเช่นแบบจำลองฤดูกาล $SARMA(1,1)_{12}$ โดยที่ข้อมูลเป็นรายเดือนค่าฤดูกาล $s = 12$ สามารถเขียนเป็นสมการได้ดังนี้

$$(1 - \Phi L^{12}) y_t = (1 + \Theta L^{12}) \varepsilon_t \tag{2.84}$$

หรือ $y_t = \Phi x_{12} + \varepsilon_t + \Theta \varepsilon_{t-12}$ โดยที่อนุกรมเวลานี้จะนิ่งถ้า $|\Phi| < 1$ และหาตัวผกผันได้ถ้า $|\Theta| < 1$

สำหรับแบบจำลอง $SMA(1)_{12}$, $y_t = \varepsilon_t + \Theta \varepsilon_{t-12}$ จะมีค่าความแปรปรวนในตัวเองและค่าสหสัมพันธ์ในตัวเองดังนี้

$$\begin{aligned}
\gamma_0 &= (1 + \Theta^2)\sigma^2 \\
\gamma_{12} &= \Theta \sigma^2 \\
\gamma_j &= 0, \quad j \neq 0, j \neq 12 \\
\rho_{12} &= \Theta / (1 + \Theta^2)
\end{aligned}$$

ส่วนแบบจำลอง $SAR(1)_{12}$, $y_t = \Phi y_{t-12} + \varepsilon_t$ จะมีค่าความแปรปรวนในตัวเองและค่า

สหสัมพันธ์ในตัวเองดังนี้

$$
\begin{aligned}
\gamma_0 &= \sigma^2/(1-\Phi^2) \\
\gamma_{12k} &= \sigma^2\Phi^k/(1-\Phi^2), \quad k = 1, 2, 3, ... \\
\gamma_j &= 0, \quad j \neq 0, j \neq 12 \\
\rho_{12k} &= \Phi^k, \quad k = 0, 1, 2, ...
\end{aligned}
$$

โดย ทั่วไป เรา จะ รวม ส่วน ที่ ไม่ เป็น ฤดูกาล และ เป็น ฤดูกาล เข้า ด้วย กัน ใน แบบ จำลอง multiplicative seasonal $ARMA(p,q) \times (P,Q)_s$

$$\Phi_P(L^s)\phi(L)y_t = \Theta_Q(L^s)\theta(L)\varepsilon_t \tag{2.85}$$

ตัวอย่าง $ARMA(0,1) \times (1,0)_{12}$

$$y_t = \Phi y_{t-12} + \varepsilon_t - \theta\varepsilon_{t-1} \tag{2.86}$$

แบบจำลอง multiplicative seasonal $ARIMA(p,d,q) \times (P,D,Q)_s$

$$\Phi_P(L^s)\phi(L)\Delta_s^D\Delta^d y_t = \Theta_Q(L^s)\theta(L)\varepsilon_t \tag{2.87}$$

โดยที่ $\Delta^d = (1-L)^d$ และ $\Delta_s^D = (1-L^s)^D$

ตัวอย่าง $ARIMA(0,1,1) \times (0,1,1)_{12}$

$$(1-L^{12})(1-L)y_t = (1+\Theta L^12)(1+\theta L)\varepsilon_t \tag{2.88}$$

ซึ่งสามารถขยายได้เป็น

$$(1-L-L^{12}+L^{13})y_t = (1+\theta L+\Theta L^{12}+\Theta\theta L^{13})\varepsilon_t$$

บทที่ 3

สมการถดถอยสำหรับข้อมูลอนุกรมเวลา

วิธี การ ประมาณ สมการ ถดถอย สำหรับ ข้อมูล อนุกรม เวลา ถูก นำ มา ใช้ แพร่ หลาย ใน การ วิเคราะห์ ข้อมูล ทางการ เงิน และ ใช้ ใน การ ประมาณ ค่า และ ทดสอบ แบบ จำลอง เช่น แบบ จำลอง ราคาสินทรัพย์และผลได้ตอบแทนของสินทรัพย์เช่น Capital Asset Pricing Model (CAPM) หรือ Arbitrage Price Model นอกจากนี้ยังใช้ศึกษาความสัมพันธ์ระหว่างผลได้ตอบแทนของทรัพย์สิน กับสัดส่วนทางการเงิน เพื่อใช้ในการทำนายการเปลี่ยนแปลงของผลได้ตอบแทน

อย่างไรก็ตาม การศึกษาความสัมพันธ์ทางการเงินที่ใช้ข้อมูลอนุกรมเวลาและวิธีการถดถอย จำเป็นที่จะต้องระมัดระวัง เนื่องจากคุณสมบัติของการประมาณค่าถดถอยบางประการอาจจะได้รับ ผลกระทบจากข้อมูลที่เป็นอนุกรมเวลา อันจะส่งผลต่อการอธิบายผลและทดสอบผล โดยทั่วไป ข้อมูลที่เราจะนำมาใช้ในการศึกษามักจะเป็นข้อมูลที่นิ่ง ($I(0)$) เช่น ผลได้ตอบแทนของทรัพย์สิน

3.1 แบบจำลองถดถอยเชิงเส้นตรงสำหรับข้อมูลอนุกรมเวลา

$$y_t = \beta_0 + \beta_1 x_{1t} + ... + \beta_k x_{kt} + \varepsilon_t = \boldsymbol{x}_t^{'}\boldsymbol{\beta} + \varepsilon_t, \quad t = 1, ..., T \tag{3.1}$$

โดยที่เราสามารถเขียนให้อยู่ในรูปเมตริกซ์ได้

ข้อสมมุติมาตรฐานสำหรับแบบจำลองถดถอยเชิงเส้นในกรณีข้อมูลอนุกรมเวลาคือ

- แบบจำลองในสมการ (**??**) เป็นแบบจำลองที่ระบุถูกต้อง
- $y_t, \boldsymbol{x}_t$ เป็นอนุกรมนิ่งและ ergodic
- ตัวแปรต้น $\boldsymbol{x}_t$ ถูกกำหนดไว้ก่อน (predetermined): $E(x_{is}\varepsilon_t) = 0$ สำหรับทุกค่าที่ $s \leq t$
- $E(\boldsymbol{x}_t\boldsymbol{x}_t^{'})$ เป็น full rank
- $\boldsymbol{x}_t\varepsilon_t$ เป็นกระบวนการที่ไม่มีสหสัมพันธ์กันในระยะจำกัด

3.1.1 การประมาณค่าวิธีกำลังสองน้อยที่สุด

วิธีการประมาณค่ากำลังสองน้อยที่สุดทั่วไป (Ordinary Least Square:OLS) เป็นการประมาณค่าโดยหาตัวประมาณค่าของพารามิเตอร์ที่ทำให้

$$SSR(\boldsymbol{\beta}) = \sum_{t=1}^{T}(y_t - \boldsymbol{x}_t'\boldsymbol{\beta})^2 = \sum_{t=1}^{T}\varepsilon_t^2 \tag{3.2}$$

และได้แบบจำลอง fitted

$$y_t = \boldsymbol{x}_t'\hat{\boldsymbol{\beta}} + \hat{\varepsilon}_t, \qquad t = 1, ..., T$$

โดยที่ $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ และ $\hat{\varepsilon}_t = y_t - \hat{y}_t = y_t - \boldsymbol{x}_t'\hat{\boldsymbol{\beta}}$ โดยมีค่าประมาณ error variance เท่ากับ $\hat{\sigma}^2 = \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}/(T-k-1)$ และภายใต้ข้อสมมุติข้างต้นตัวประมาณค่า OLS จะ consistent และ asymptotically normally distributed และตัวประมาณค่าของค่าความแปรปรวน

$$\widehat{Aver(\hat{\boldsymbol{\beta}})} = \hat{\sigma}^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

และค่า standard errors ของแต่ละตัวประมาณค่า β_j คือค่ารากที่สองของสมาชิกลำดับที่ j ในแนวทแยงมุม

การวิเคราะห์ค่าส่วนเกิน (residuals)

หลังจากการประมาณค่าแล้วค่าสถิติสองสามตัวจะถูกใช้ในการตรวจสอบตัวสถิติ โดยที่ตัวสถิติเหล่านี้จะทำหน้าที่ในการตรวจสอบว่าค่าส่วนเกินเป็นไปตามทฤษฎีหรือไม่ โดยตัวสถิติที่ใช้โดยปกติคือการตรวจสอบ serial correlation ด้วยการวิเคราะห์ตัวประมาณค่าสวนเกิน $\hat{\varepsilon}_t$ ที่เรียกว่าตัวสถิติ Durbin-Watson

$$DW = \frac{\sum_{t=2}^{T}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=2}^{T}\hat{\varepsilon}_t^2}$$

โดยที่ $DW \approx 2(1-\hat{\rho})$ โดยที่ ρ คือความสัมพันธ์ระหว่าง ε_t และ ε_{t-1} และค่า DW จะมีค่าอยู่ระหว่าง 0 และ 4 โดยที่ค่าใกล้สองแสดงว่าไม่มีสหสัมพันธ์ระหว่างกัน นอกจากนี้เรายังมีตัวสถิติอีกตัวคือ Ljung-Box Q

3.2 Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation

$$
\begin{aligned}
Var(\hat{\beta}_1|\boldsymbol{X}) &= Var(\beta_1 + \frac{\frac{1}{T}\sum_{t=1}^{T}(x_t-\bar{x})\varepsilon_t|X)}{\frac{1}{T}\sum_{t=1}^{T}(x_t-\bar{x})2} \\
&= \frac{Var(\frac{1}{T}\sum_{t=1}^{T}(x_t-\bar{x})\varepsilon_t|X))}{\left(\frac{1}{T}\sum_{t=1}^{T}(x_t-\bar{x})2\right)^2}
\end{aligned} \tag{3.3}
$$

โดยที่ตัวส่วนของสูตรข้างบนสามารถเขียนในรูปต่อไปนี้

$$
\begin{aligned}
Var(\frac{1}{T}\sum_{t=1}^{T}(x_t-\bar{x})\varepsilon_t|X)) &= \frac{1}{T^2}Var(\sum_{t=1}^{T}(x_t-\bar{x})\varepsilon_t|X) \\
&= \frac{1}{T^2}\sum_{t=1}^{T}Var\left((x_t-\bar{x})\varepsilon_t|X\right) + \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=t+1}^{T}Cov\left((x_t-\bar{x})\varepsilon_t,(x_s-\bar{x})\varepsilon_s|X\right)
\end{aligned} \tag{3.4}
$$

ในกรณีที่เราสมมุติให้ค่าผิดพลาดไม่มีปัญหา heteroskedasticity และ autocorrelation

$$
\underbrace{\frac{1}{T^2}\sum_{t=1}^{T}Var\left((x_t-\bar{x})\varepsilon_t|X\right)}_{=\sigma^2\sum_{t=1}^{T}(x_t-\bar{x})^2} + \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=t+1}^{T}\underbrace{Cov\left((x_t-\bar{x})\varepsilon_t,(x_s-\bar{x})\varepsilon_s|X\right)}_{=0}
$$

ดังนั้น $Var(\hat{\beta}_1) = \frac{\sigma^2}{(x_t-\bar{x})^2}$

3.2.1 White

3.2.2 Newey-West

ในกรณีที่เราต้องการแก้ไขความคลาดเคลื่อนมาตรฐานเราจำเป็นต้องประมาณค่าสมการ (**??**) อย่างไรก็ตามสมการดังกล่าวมีพจน์ที่คูณกันอยู่ค่อนข้างเยอะ Newey-West เสนอว่าเราไม่จำเป็นต้องใช้พจน์ทั้งหมดที่คูณกันอยู่เพียงประมาณค่าแค่ช่วงหนึ่งก็พอ โดยการเลือกค่า L ค่าหนึ่งซึ่งทำให้

$$
Var(\hat{\beta}_1) = \frac{1}{T}\sum_{l=1}^{L}\sum_{t=l+1}^{T}W_{l\varepsilon_t\varepsilon_{t-l}(\tilde{x}_t\tilde{x}_{t-l}-\tilde{x}_{t-l}\tilde{x}_t)}(3.5)
$$

โดยที่ $W_l = \frac{l}{L+1}$ และ $\tilde{x}_t = x_t - \bar{x}$ ปัญหาในทางปฏิบัติคือเราจะเลือก L เท่าใด โดยหลักแล้วถ้าข้อมูลยาวมากๆ L ก็น่าจะยาวด้วย Newey and West เสนอให้ใช้ $L = 4(T/100)^{2/9}$

บทที่ 4

แบบจำลอง $GARCH$

เป้าหมายที่สำคัญประการหนึ่งของการศึกษาเศรษฐมิติทางการเงินคือการศึกษาความผัน-ผวน (volatility) ของผลได้ตอบแทนของสินทรัพย์ โดยที่ความผันผวนจะหมายถึงความเบี่ยงเบนมาตรฐานที่มีเงื่อนไข (conditional standard deviation) ของผลได้ตอบแทนของสินทรัพย์

ความผันผวนที่คำนวณได้ถูกนำไปใช้ประยุกต์ในการคำนวณ Value-at-Risk(VaR) ของการจัดการความเสี่ยง, การจัดสรรการลงทุนภายใต้วิธีการ mean-variance และการเพิ่มประสิทธิภาพในการทำนาย

4.1 ความผันผวนของสินทรัพย์

4.1.1 คุณลักษณะของความผันผวน

ปัญหาเบื้องต้นของการศึกษาความผันผวนคือเราไม่สามารถสังเกตความผันผวนของสิน-ทรัพย์เช่นหลักทรัพย์ได้เหมือนกับการสังเกตเห็นราคาทรัพย์สินหรือการคำนวณผลตอบแทนได้จากราคาทรัพย์สิน ในอดีตนักวิชาการได้นำเสนอตัวแทนของความผันผวนเช่น หากกำหนดให้ P_t เป็นราคาหลักทรัพย์ และกำหนดให้ $y_t = \Delta \ln(P_t)$ เป็นผลได้ตอบแทน เราพบว่าอนุกรมของผลได้ตอบแทนมักจะไม่มีสหสัมพันธ์ระหว่างกันผลได้ตอบแทนในอดีต แต่ผลได้ตอบแทนกำลังสอง (y_t^2) , ค่าสัมบูรณ์ของผลได้ตอบแทน ($|y_t|$), หรือค่ายกกำลังของค่าสัมบูรณ์ของผลได้ตอบแทน ($|y_t|^\delta$) มีสหสัมพันธ์ค่อนข้างสูง ดังนั้นในอดีตเรามักใช้ค่าดังกล่าวเป็นตัวแทนของความผันผวน อย่างไรก็ตามในปัจจุบันมีการคำนวณความผันผวนด้วยวิธีการอื่นๆเช่น Realized volatility และ Implied volatility ด้วย Black-Scholes

4.1.2 การทดสอบ ARCH effect

หลังจากที่เราสร้างแบบจำลองเพื่ออธิบายผลได้ตอบแทน (y_t) แล้ว หากนิยาม $\varepsilon_t = y_t - \mu_t$ เป็น residuals ของแบบจำลองที่ใช้อธิบายค่าเฉลี่ย (mean equation) ของผลได้ตอบแทนซึ่งแทน

ด้วยสัญลักษณ์ μ_t ซึ่งแบบจำลองที่ใช้อาจจะเป็นแบบจำลอง $AR(p)$ หรือ $ARIMA(p,d,q)$ ในบทที่ผ่านมาเราจะทดสอบว่าแบบจำลองเพียงพอที่จะอธิบายค่าเฉลี่ยหรือไม่โดยพิจารณาอัตสหสัมพันธ์ของ residuals ซึ่งหากไม่หลงเหลือความสัมพันธ์แล้วแสดงว่าแบบจำลองเพียงพอ อย่างไรก็ตามจากการพิจารณาข้อมูลอนุกรมเวลาเราในอดีต เราพบว่า residuals อาจจะมีความผันผวนที่เปลี่ยนแปลงตามเวลา ดังนั้น เราต้องการทดสอบว่าข้อมูลของเรามีความแปรปรวนที่เปลี่ยนแปลงตามเวลา(conditional heteroskedasticity)หรือไม่ เราเรียกการทดสอบนี้ว่าการทดสอบ ARCH effect ซึ่งลักษณะของอนุกรมเวลาที่จะมีค่าความแปรปรวนเปลี่ยนแปลงตามเวลาก็ต่อเมื่อค่ากำลังสองของอนุกรมเวลานั้นๆมีความสัมพันธ์กัน จากแนวคิดดังกล่าวมีการทดสอบสองวิธีที่เป็นที่นิยมใช้ได้แก่ (1) การใช้การทดสอบ Ljung-Box เพื่อทดสอบอนุกรมเวลา ε_t^2 (McLeod and Li, 1983) และ (2) การทดสอบ Lagrange multiplier (Engle 1982)

วิธีแรกเป็นการทดสอบ Ljung-Box $Q(m)$ สำหรับค่ายกกำลังสองของ residuals (ε_t^2) โดยเราใช้ ε_t^2 เป็นตัวแทนค่าแปรปรวนที่เปลี่ยนแปลงตามเวลา ดังนั้นหากค่าความแปรปรวนขึ้นอยู่ต่อกันจะต้องมีสหสัมพันธ์ระหว่างกัน ดังนั้นเราจะตั้งสมมุติฐานหลักค่ายกกำลังสองของ residuals ไม่มีสหสัมพันธ์ระหว่างกันย้อนไปถึงคาบที่ m หากสมมุติฐานหลักเป็นจริงแสดงว่า residuals ไม่มีลักษณะของ ARCH (หรือไม่มี ARCH effect)

วิธีสองเป็นการทดสอบ Lagrange Multiplier (LM) โดยเราใช้ ε_t^2 เป็นตัวแทนของค่าความแปรปรวน ซึ่งหากเราต้องการทดสอบว่าค่าแปรปรวนขึ้นกับค่าแปรปรวนในอดีตหรือไม่ สามารถประมาณค่าสมการ

$$\varepsilon_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + ... + a_m\varepsilon_{t-m}^2 + e_t \tag{4.1}$$

และพิจารณาว่าค่าสัมประสิทธิ์มีนัยสำคัญเชิงสถิติต่างจากศูนย์หรือไม่ โดยสมมุติฐานหลักคือ $a_1 = a_2 = ... = a_m = 0$ หรือไม่มี ARCH effect โดยเราสามารถทำการทดสอบได้โดยการประมาณค่าสมการ **??** จะได้ค่า R^2 แล้วไปคำนวณค่าสถิติ $LM = TR^2$ โดยที่ T คือจำนวนตัวอย่าง และ $LM \sim \chi^2_{df=m}$ เราจะปฏิเสธสมมุติฐานหลักถ้า $LM > \chi^2_{df=m}(1-\alpha)$

ตัวอย่างที่ 4.1: การทดสอบ ARCH Effect ผลตอบแทนรายเดือนจากการลงทุนในตลาดหลักทรัพย์แห่งประเทศไทย

เราจะพิจารณาผลได้ตอบแทนรายเดือนในรูปล็อกของการลงทุนในตลาดหลักทรัพย์แห่งประเทศไทย โดยใช้ข้อมูลราคาปิดรายเดือนตั้งแต่ เมษายน 2518 ถึงเดือนพฤศจิกายน 2555 ซึ่งอยู่ในไฟล์ `mset.csv` เราสามารถสร้างอนุกรมเวลาผลได้ตอบแทนในรูปล็อกและหารูปแบบจำลองที่เหมาะสมได้ด้วยคำสั่ง `auto.arima` ได้แบบจำลอง $ARMA(2,2)$ ตามที่แสดงในคำสั่งข้างล่าง

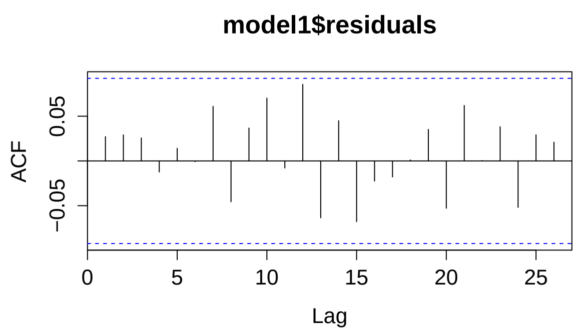

```
> mset <-
    read.csv("https://raw.githubusercontent.com/chaleampong/EC435/master/mset.csv",
    header = FALSE)
> head(mset)
> ret<-diff(log(mset$V1))
> library(forecast)
> auto.arima(ret, max.p=6, max.q=6, ic=c("aic"))
Series: ret
ARIMA(2,0,2) with zero mean

Coefficients:
         ar1      ar2      ma1     ma2
      1.1154  -0.9020  -1.0542  0.9141
s.e.  0.0511   0.0754   0.0646  0.0740

sigma^2 estimated as 0.006749:  log likelihood=489.01
AIC=-968.03   AICc=-967.89   BIC=-947.47
```

หลังจากนั้นเราประมาณค่าแบบจำลอง $ARMA(2,2)$ และให้ชื่อว่า m เราสามารถพิจารณาความพอเพียงของแบบจำลองได้ด้วย ACF ของ residuals หรือ Ljung-Box test ของ residuals

```
> m <- arima(ret, order=c(2,0,2))
> acf(m$residuals, lag.max=20)
> Box.test(m$residuals, lag=12, type="Ljung")

  Box-Ljung test

data:  m$residuals
X-squared = 10.191, df = 12, p-value = 0.5992
```

จะได้

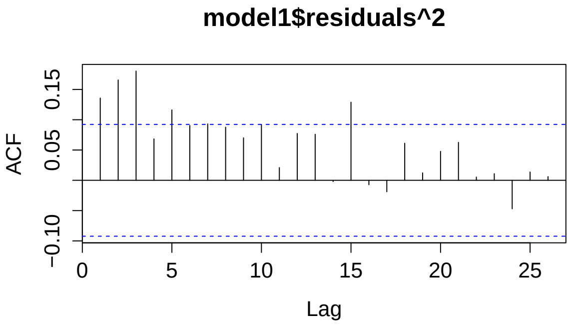
รูปที่ 4.1: ACF ของ residuals



จากรูป **??** และค่า Ljung-Box statistics =10 ซึ่งน้อยกว่า critical value จะเห็นได้ว่าแบบจำลองดังกล่าวเพียงพอในการอธิบาย log return และไม่มีสหสัมพันธ์ระหว่าง log return เหลืออยู่

แต่หากเราพิจารณา ACF ของค่ากำลังสองของ residuals หรือ Ljung-Box test ของค่ากำลังของ residuals กำลังสอง หรือทดสอบ LM ARCH effect จะเห็นได้ว่ายังมีผลของ ARCH เหลืออยู่

รูปที่ 4.2: ACF ของ residuals



ใน residuals โดยในการทดสอบ LM ARCH effect จะต้องใช้คำสั่ง `ArchTest(series, q)` จาก `library(FinTS)` โดยต้องระบุ series ที่ต้องการทดสอบและค่าล่า (q)

```
1  > acf(m$residuals^2, lag.max = 20)
2  > Box.test(m$residuals^2, lag=12, type="Ljung")
3    Box-Ljung test
4  data:  m$residuals^2
5  X-squared = 64.509, df = 12, p-value = 3.36e-09
6  > library(FinTS)
7  > ArchTest(ret)
8    ARCH LM-test; Null hypothesis: no ARCH effects
9  data:  ret
10 Chi-squared = 35.547, df = 12, p-value = 0.000383
```

กรณีการทดสอบ Ljung-Box ของค่ายกกำลังสองของ residuals (บรรทัด 2-7) ค่าพี < 0.0001 เราปฏิเสธสมมุติฐานที่ว่าค่ากำลังสองของ residuals ไม่มีสหสัมพันธ์กัน *กรณีการทดสอบ LM ARCH effect* (บรรทัด 9-14) ค่าพี < 0.002 เราปฏิเสธสมมุติฐานหลัก $a_1 = a_2 = ... = a_m = 0$ สรุปว่า เราพบว่า residuals ของแบบจำลอง $ARMA(2,2)$ มี ARCH effect

4.2 แบบจำลอง ARCH

เพื่อให้ผู้อ่านเข้าใจแรงจูงใจในการสร้างแบบจำลอง ARCH ผู้เขียนขอเน้นอีกครั้งว่าแบบจำลองอนุกรมเวลาที่เราสร้างขึ้นมาในบทที่ผ่านมา เราพยายามที่จะอธิบายค่าในอนาคตของตัวแปรสุ่มที่เราสนใจโดยที่ค่าเฉลี่ยที่เราพยากรณ์อาจจะอยู่ในรูปค่าเฉลี่ยที่มีเงื่อนไข (conditional mean) หรือค่าเฉลี่ยที่ไม่มีเงื่อนไข (unconditional mean) เช่นในกรณีแบบจำลอง $AR(1)$ ซึ่งสามารถแสดงได้ด้วย $y_t = \phi y_{t-1} + \varepsilon_t$ ซึ่งจากบทที่ผ่านมาเราทราบว่าค่าเฉลี่ย (ที่ไม่มีเงื่อนไข) $E(y_t) = 0$ และค่าความแปรปรวน $Var(y_t) = \sigma^2/(1-\phi^2)$ ซึ่งหากใช้ค่าเฉลี่ยดังกล่าวสำหรับการทำนายไปข้างหน้าจะได้ค่าคาดการณ์ของ y เท่ากับศูนย์ แต่วิธีการพยากรณ์ที่จะทำให้ค่า MSE ต่ำที่สุดควรจะใช้

ค่าเฉลี่ยที่มีเงื่อนไขจากแบบจำลอง $AR(1)$ ซึ่งเท่ากับ

$$y_h(1) = E(y_{h+1}|F_h) = \phi y_h$$

และมีค่าความแปรปรวนที่มีเงื่อนไขเท่ากับ

$$Var_h(y_{h+1}) = E_h(y_{h+1} - E_h(y_{h+1}))^2 = E_h(\phi y_h + \varepsilon_{h+1} - \phi y_h)^2 = E_h(\varepsilon_{h+1}^2) = \sigma_{h+1}^2$$

โดยที่ $E_h() = E(|F_h)$ และในบทที่ผ่านมาเราสมมุติให้ σ_{h+1}^2 มีค่าคงที่

Engle (1982) ได้วิเคราะห์โมเมนต์ที่สองที่มีเงื่อนไข โดยที่สร้างแบบจำลองที่ค่าความแปรปรวนสามารถเปลี่ยนแปลงได้ตามเวลา

4.2.1 รูปแบบของแบบจำลอง ARCH(q)

ค่าสหสัมพันธ์ในตัวเองของผลได้ตอบแทนกำลังสองหรือ conditional heteroskedasticity สามารถอธิบายได้ด้วยแบบจำลอง AR ของค่าส่วนเกิน (residuals) กำลังสอง ตัวอย่างเช่น กำหนดให้ y_t อนุกรมเวลานิ่ง เช่น ผลได้ตอบแทน แล้วกำหนดให้ y_t สามารถอธิบายได้ด้วยแบบจำลอง

$$y_t = \mu_t + \varepsilon_t \tag{4.2}$$

โดยที่ μ_t เป็นส่วนอธิบายค่าเฉลี่ย (mean equation) ซึ่งอาจจะเป็นค่าคงที่ c หรือ แบบจำลอง $ARMA(p,q)$ และ ε_t เป็น iid ที่มีค่าเฉลี่ยเท่ากับศูนย์ เพื่อที่จะให้แบบจำลองดังกล่าวมีความผันผวนเป็นกลุ่ม (volatility clustering) หรือ conditional heteroskedasticity เราจะสมมุติให้ความแปรปรวนที่ขึ้นอยู่กับข้อมูล ณ คาบที่ $t-1$ $Var_{t-1} = \sigma_t^2$ เขียนเป็นสมการต่อไปนี้

$$\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + ... + \alpha_q\varepsilon_{t-q}^2 \tag{4.3}$$

เนื่องจาก ε_t มีค่าเฉลี่ยเท่ากับศูนย์ $Var_{t-1}(\varepsilon_t) = E_{t-1}(\varepsilon_t^2) = \sigma_t^2$ ดังนั้นเราสามารถเขียนสมการ (**??**) ได้เป็น

$$\varepsilon_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + ... + \alpha_q\varepsilon_{t-q}^2 + u_t \tag{4.4}$$

โดยที่ $u_t = \varepsilon_t^2 - E_{t-1}(\varepsilon_t^2)$ ซึ่งเป็นกระบวนการไวท์นอซที่มีค่าเฉลี่ยเท่ากับศูนย์ สมการ (**??**),(**??**) และ (**??**) รวมกันเรียกว่า autoregressive conditional heteroskedasticity (ARCH) หรือ แบบจำลอง $ARCH(q)$

รูปแบบของ $ARCH(q)$ อีกแบบจำลองสามารถแสดงได้ดังนี้

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t \\ \varepsilon_t &= z_t\sigma_t \\ \sigma_t^2 &= \omega + \alpha_1\varepsilon_{t-1}^2 + ... + \alpha_q\varepsilon_{t-q}^2 \end{aligned} \tag{4.5}$$

โดยที่ z_t เป็นตัวแปรสุ่มเป็นอิสระและเหมือนกันที่มีค่าเฉลี่ยเท่ากับศูนย์และค่าความแปรปรวนเท่ากับหนึ่ง, $iid(0,1)$ ตัวอย่างเช่น การแจกแจงปกติมาตรฐาน

4.2.2 คุณสมบัติของ ARCH(1)

เพื่อให้ผู้อ่านเข้าใจแบบจำลอง $ARCH$ เราจะมาพิจารณาคุณสมบัติของแบบอนุกรมเวลา $ARCH(1)$ สมมุติว่าให้ y_t มีคุณลักษณะตามแบบจำลอง $ARCH(1)$ โดยที่

$$y_t = \varepsilon_t \tag{4.6}$$

$$\varepsilon_t = z_t\sigma_t \tag{4.7}$$

และ

$$\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 \tag{4.8}$$

หากเรายกกำลังสองทั้งสองข้างของ (**??**) แล้วสลับข้างสมการ (**??**) แล้วนำมาลบจะได้

$$\begin{aligned}\varepsilon_t^2 - \omega - \alpha_1\varepsilon_{t-1}^2 &= (z_t^2-1)\sigma_t^2\\ \varepsilon_t^2 &= \omega + \alpha_1\varepsilon_{t-1}^2 + (z_t^2-1)\sigma_t^2\end{aligned}$$

ซึ่งเป็นแบบจำลองของ ε_t^2 ในรูป AR โดยที่ $u_t = (z_t^2-1)\sigma_t^2$ และจากความสัมพันธ์ที่ $y_t^2 = \varepsilon_t^2$ เราสามารถสรุปได้ว่า $ARCH(1)$ สามารถเขียนในรูปกำลังสองของ y_t ได้

ค่าเฉลี่ยที่ไม่มีเงื่อนไข(unconditional mean)ของ ε_t เท่ากับศูนย์ ซึ่งการคำนวณสามารถแสดงได้ดังนี้

$$\begin{aligned}E(\varepsilon_t) = E(E(\varepsilon_t|F_{t-1})) &= E(E(z_t\sigma_t|F_{t-1}))\\ &= E(\sigma_t\underbrace{E(z_t|F_{t-1})}_{=0}) = 0\end{aligned}$$

โดยสมการแรกใช้ Law of Iterated Expectation

ค่าความแปรปรวน (unconditional variance) ของ ε_t สามารถคำนวณได้โดย

$$\begin{aligned}Var(\varepsilon_t) = E(\varepsilon_t^2) &= E(E(\varepsilon_t^2|F_{t-1}))\\ &= E(E(z_t^2\sigma_t^2|F_{t-1}))\\ &= E(\sigma_t^2\underbrace{E(z_t^2|F_{t-1})}_{=1})\\ &= E(\omega + \alpha_1\varepsilon_{t-1}^2) = \omega + \alpha_1 E(\varepsilon_{t-1}^2)\end{aligned}$$

หากต้องการให้ ε_t เป็นกระบวนการนิ่ง $Var(\varepsilon_t) = E(\varepsilon_t^2) = E(\varepsilon_{t-1}^2)$ ดังนั้น $Var(\varepsilon_t) = \frac{\omega}{1-\alpha_1}$

และเงื่อนไขที่จะทำให้ค่าความแปรปรวนเป็นบวกคือ $0 \leq \alpha_1 < 1$

ในทางปฏิบัติ เราต้องการทราบว่าค่าโมเมนต์ที่สูงกว่าโมเมนต์ที่สองเป็นอย่างไร ภายหลังจากการดำเนินการบางอย่างเราพบว่า โมเมนต์ที่สี่เท่ากับ

$$E(\varepsilon^4) = 3\frac{\omega(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)} \tag{4.9}$$

ดังนั้นเงื่อนไขที่ทำให้โมเมนต์ที่สี่มีค่าจำกัดคือ $0 \leq \alpha_1 < 1/3$ และค่าความโด่งเท่ากับ

$$\kappa = \frac{E(\varepsilon_t^4)}{[Var(\varepsilon_t)]^2} = 3\frac{\omega(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)} \times \frac{(1-\alpha_1)^2}{\omega^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2} \tag{4.10}$$

ซึ่งค่าความโด่งจะมากกว่าสามเสมอ แสดงว่าการแจกแจงของช็อกในแบบจำลอง ARCH จะมีหางที่อ้วนกว่าปกติ

4.2.3 การประมาณค่าแบบจำลอง $ARCH(1)$

การประมาณค่าพารามิเตอร์ในแบบจำลอง $ARCH(1)$ สามารถดำเนินการได้ด้วยวิธีการ conditional maximum likelihood หากสมมุติให้ z_t มีการแจกแจงแบบปกติมาตรฐาน จากสมการ (**??**) เราจะได้ $\varepsilon_t \sim N(0, \sigma_t^2)$ และฟังก์ชันล็อกไลลิฮูดจะเท่ากับ

$$\begin{aligned} \ln L(,\alpha_1;\boldsymbol{y}) &= \sum_{t=1}^{T} \ln f(\varepsilon_t|\varepsilon_{t-1}) \\ &= -\frac{1}{2}\sum_{t=1}^{T}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln(\sigma_t^2) - \sum_{t=1}^{T}\left(\frac{y_t^2}{2\sigma_t^2}\right) \end{aligned} \tag{4.11}$$

โดยที่ค่า ε_0 และ σ_0^2 เท่ากับค่าเฉลี่ย เราสามารถหาตัวประมาณค่าพารามิเตอร์ได้จากการหาค่าสูงสุดของสมการ (**??**) อย่างไรก็ตามเราสามารถกำหนดการแจกแจงรูปแบบอื่นที่มีค่าเฉลี่ยเท่ากับศูนย์และค่าความแปรปรวนเท่ากับหนึ่ง เช่น standardized student's t distribution

4.3 แบบจำลอง GARCH(p,q)

ถ้าเราทดสอบ ARCH effect ด้วยตัวทดสอบเช่น LM ARCH effect แล้วพบว่าอนุกรมเวลามีผลของ ARCH เราสามารถใช้แบบจำลอง $ARCH(q)$ เพื่อประมาณค่าความผันผวนที่เปลี่ยนแปลงตามเวลา (time varying volatility: σ_t^2) ซึ่งขึ้นอยู่กับข้อมูลในอดีต อย่างไรก็ตามในทางปฏิบัติ ค่าล่า q ที่เหมาะสมสำหรับแบบจำลอง $ARCH(q)$ ค่อนข้างที่จะยาว ส่งผลต่อจำนวนพารามิเตอร์ที่ต้องประมาณค่าและความเหมาะสม (goodness of fit) ของแบบจำลอง ดังนั้น Bollerslev (1986) ได้เสนอ

รูปแบบของสมการความแปรปรวนเป็น

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{4.12}$$

โดยที่มีเงื่อนไขค่าสัมประสิทธิ์ α_i $(i = 1, ..., q)$ และ β_j $(j = 1, ..., p)$ มีค่าเป็นบวกเพื่อให้ค่าแปรปรวนเป็นบวก โดยที่สมการที่ (**??**) และ (**??**) รวมกันเรียกว่าแบบจำลอง Generalized ARCH หรือ $GARCH(p, q)$ และหากค่า $p = 0$ แบบจำลองดังกล่าวจะลดรูปเหลือแค่ $ARCH(q)$

ในแบบจำลอง $GARCH(p, q)$ ค่าความแปรปรวนที่มีเงื่อนไข จะขึ้นอยู่กับค่ากำลังสองของค่าส่วนเกิน (squared residuals) ในช่วง q คาบก่อน และ ค่าความแปรปรวนที่มีเงื่อนไขในช่วง p คาบก่อน อย่างไรก็ดี โดยปกติแล้ว แบบจำลอง $GARCH(1, 1)$ ก็เพียงพอที่จะใช้อธิบายอนุกรมเวลาทางการเงิน

4.3.1 ARMA representation of GARCH

ในขณะที่แบบจำลอง $ARCH$ ที่สามารถเขียนให้อยู่ในรูปของแบบจำลอง AR ของค่า residuals ยกกำลังสอง แบบจำลอง $GARCH$ สามารถเขียนให้อยู่ในรูปของแบบจำลอง $ARMA$ ของค่า residuals ยกกำลังสอง ตัวอย่างเช่นในแบบจำลอง $GARCH(1, 1)$

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{4.13}$$

เนื่องจาก $E_{t-1}(\varepsilon_t^2) = \sigma_t^2$ ดังนั้นสมการ **??** สามารถเขียนใหม่ได้เป็น

$$\varepsilon_t^2 = \omega + (\alpha_1 + \beta_1)\varepsilon_{t-1}^2 + u_t - \beta_1 u_{t-1} \tag{4.14}$$

ซึ่งเป็นแบบจำลอง $ARMA(1, 1)$ โดยที่ $u_t = \varepsilon_t^2 - E_{t-1}(\varepsilon_t^2)$ เป็นตัวแปรช็อกไวท์น๊อซ

แบบจำลอง $GARCH$ ในรูปแบบของ $ARMA$ สามารถใช้แสดงคุณสมบัติของแบบจำลอง $GARCH$ ได้ง่ายตัวอย่างเช่น $GARCH(1, 1)$ จะเป็นอนุกรมนิ่งถ้า $\alpha_1 + \beta_1 < 1$ เนื่องจาก ε_t^2 อยู่ในรูป $AR(1)$ และ $\alpha_1 + \beta_1$ คือสัมประสิทธิ์ของ ε_{t-1}^2 ซึ่งจะต้องมีค่าน้อยกว่าหนึ่ง (อย่าลืมว่า α_1 และ β_1 จะต้องมีค่ามากกว่าศูนย์)

หากเราสมมุติให้แบบจำลองเป็นอนุกรมนิ่ง เราสามารถหาความแปรปรวนที่ไม่มีเงื่อนไข (unconditional variance)ของ ε_t จะเท่ากับ $Var(\varepsilon_t) = E(\varepsilon_t^2) = \omega/(1 - \alpha_1 - \beta_1)$ เนื่องจาก

$$E(\varepsilon_t^2) = \omega + (\alpha_1 + \beta_1) \underbrace{E(\varepsilon_{t-1}^2)}_{=E(\varepsilon_t^2)}$$

$$E(\varepsilon_t^2) = \omega + (\alpha_1 + \beta_1) E(\varepsilon_t^2)$$

ในกรณี $GARCH(p, q)$ ทั่วไปเราสามารถแสดงให้เห็นว่าเราสามารถเขียนแบบจำลอง

ดังกล่าวให้อยู่ในรูป $ARMA(\max(p,q),p)$ ของ residuals ยกกำลังสอง และแบบจำลองจะเป็นอนุกรมนิ่งถ้า $\sum_{i=1}^{q}\alpha_i + \sum_{j=1}^{p}\beta_j < 1$ และความแปรปรวนที่ไม่มีเงื่อนไข (unconditional variance)ของ ε_t จะเท่ากับ

$$Var(\varepsilon_t) = \frac{\omega}{1-\left(\sum_{i=1}^{q}\alpha_i + \sum_{j=1}^{p}\beta_j\right)}$$

4.3.2 คุณลักษณะเฉพาะของแบบจำลอง GARCH

ในทางปฏิบัตินักวิจัยได้พบคุณลักษณะหลายประการของความผันผวนของข้อมูลทางการเงิน โดยที่เราสามารถใช้รูปแบบ $ARMA$ ของ $GARCH$ ในการแสดงคุณลักษณะต่างๆ โดยคุณลักษณะที่สำคัญสามประการคือ ความผันผวนเป็นกลุ่ม (volatility clustering), หางอ้วน (fat tails) และความผันผวนย้อนกลับมาที่ดุลยภาพ (volatility mean reversion)

ความผันผวนเป็นกลุ่ม (volatility clustering)

ในแบบจำลอง $GARCH(1,1)$ ที่แสดงในสมการที่ **??** ค่าที่ประมาณได้จากข้อมูลรายสัปดาห์และรายวันของ β_1 จะมีค่าประมาณ 0.9 แสดงว่าหากค่า σ_{t-1}^2 มีขนาดใหญ่จะทำให้ σ_t^2 มีขนาดใหญ่ด้วยเช่นเดียวกัน แต่ถ้า σ_{t-1}^2 มีขนาดเล็กจะทำให้ σ_t^2 มีขนาดเล็ก ดังนั้นความผันผวนจะมีลักษณะที่เป็นกลุ่ม (clustering) คือความผันผวนขนาดใหญ่จะอยู่ใกล้ๆกัน

หางอ้วน(fat tails)

ข้อมูลทางการเงินที่มีความถี่สูงจะมีหางที่อ้วนมากกว่าการแจกแจงแบบปกติ คือมีความน่าจะเป็นที่จะเกิดการเปลี่ยนแปลงขนาดใหญ่(ที่ปลายหางของการแจกแจง)มากกว่ากรณีการแจกแจงแบบปกติ Bollerslev แสดงให้เห็นว่าภายใต้เงื่อนไขบางประการ เขาสามารถแสดงให้เห็นว่า $GARCH(1,1)$ จะให้ค่าเคอร์โทซิสที่มากกว่า 3 ซึ่งคือค่าเคอร์โทซิสของการแจกแจงแบบปกติ ดังนั้นแบบจำลอง $GARCH$ สามารถเลียนแบบพฤติกรรมหางอ้วนในข้อมูลจริง

ความผันผวนย้อนกลับมาที่ดุลยภาพ(volatility mean reversion)

แม้ว่าความผันผวนจะสูงขึ้นในบางช่วงแต่สุดท้ายแล้วจะย้อนกลับไปยังดุลยภาพระยะยาว โดยที่ค่าความแปรปรวนของ $GARCH(1,1)$ จะเท่ากับ $\omega/(1-\alpha_1-\beta_1)$ และถ้า $\alpha_1+\beta_1<1$ กระบวนการ $GARCH(1,1)$ จะเป็นกระบวนการนิ่ง และค่าความแปรปรวนจะกลับเข้าสู่ดุลยภาพในระยะยาว $\omega/(1-\alpha_1-\beta_1)$ หรือเรียกว่า "mean reverts" กลับไปสู่ดุลยภาพระยะยาว แต่ถ้า $\alpha_1+\beta_1 \geq 1$ กระบวนการ $GARCH(1,1)$ จะเป็นกระบวนการไม่นิ่ง (nonstationary)

4.4 การประมาณค่าแบบจำลอง GARCH

ในหัวข้อนี้เราจะพิจารณาการประมาณค่าแบบจำลอง $GARCH(p,q)$ โดยที่แบบจำลองสามารถเขียนในรูปแบบ

$$y_t = \mu_t + \varepsilon_t \tag{4.15}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{4.16}$$

สำหรับ $t = 1,2,...,T$ โดยที่ $\sigma_t^2 = Var_{t-1}(\varepsilon_t)$ และ μ_t แทนแบบจำลองค่าเฉลี่ยเช่น อาจจะกำหนดให้เท่ากับค่าคงที่ c หรือแบบจำลอง $ARMA(p,q)$ หากเราสมมุติให้ ε_t มีการแจกแจงแบบปกติภายใต้ข้อมูลในอดีตเราสามารถเขียน log-likelihood function ของแบบจำลอง $GARCH(p,q)$ ได้

$$\log L = -\frac{1}{2}\sum_{t=1}^{T} \ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T} \ln(\sigma_t^2) - \sum_{t=1}^{T} \left(\frac{y_t^2}{2\sigma_t^2}\right) \tag{4.17}$$

หากสมมุติให้ $\mu_t = c$ ค่าพารามิเตอร์ที่ต้องประมาณค่าคือ $c, \alpha_i(i = 1,...,q), \beta_j(j = 1,...,p)$ โดยที่หลังจากที่เราประมาณค่าพารามิเตอร์ดังกล่าวแล้ว เราจะได้ค่า σ_t สำหรับ $t = 1,...,T$

4.4.1 การตรวจสอบหลังจากการประมาณค่า GARCH

หลังจากที่เราประมาณค่าแบบจำลอง GARCH แล้วเราจำเป็นต้องตรวจสอบว่าแบบจำลองดังกล่าวเพียงพอต่อการอธิบายค่าเฉลี่ยที่มีเงื่อนไข (conditional mean) และค่าความแปรปรวนที่มีเงื่อนไข (conditional variance) หรือไม่ โดยมีพื้นฐานจากข้อสมมุติที่ว่า $_t = \frac{\varepsilon_t}{\sigma_t} \sim iidN(0,1)$ ดังนั้นเราจำเป็นต้องทดสอบ z_t ว่ามีคุณสมบัติเป็นอิสระต่อกัน และมีการแจกแจงแบบปกติหรือไม่ โดยที่ตัวแทนของ z_t คือ standardized residuals ($\hat{z}_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t}$ โดยเราจะทดสอบดังต่อไปนี้

ประการแรก เราจะทดสอบว่าแบบจำลองเพียงพอต่อการอธิบายค่าเฉลี่ยที่มีเงื่อนไขหรือไม่ โดยการทดสอบว่า standardized residuals เป็นอิสระต่อกันหรือไม่ โดยการทดสอบ Ljung-Box ซึ่งสมมุติฐานหลักคือ standardized residuals เป็นอิสระต่อกัน หากเราไม่สามารถปฏิเสธสมมุติฐานหลักได้แสดงว่าแบบจำลองนั้นเพียงพอต่อการอธิบายค่าเฉลี่ยที่มีเงื่อนไข

ประการที่สอง เราจะทดสอบว่าแบบจำลองเพียงพอต่อการอธิบายความแปรปรวนแบบที่มีเงื่อนไขหรือไม่ หากแบบจำลองไม่เพียงพอต่อการอธิบายความแปรปรวนแบบมีเงื่อนไข ค่ายกกำลังสองของ standardized residuals จะต้องมีความสัมพันธ์ระหว่างกัน ซึ่งเราสามารถทดสอบ Ljung-Box กับค่ายกกำลังสองของ standardized residuals หรืออาจจะทดสอบ ARCH effect ของ standardized residuals หากเราไม่สามารถปฏิเสธสมมุติฐานหลักแสดงว่าแบบจำลองเพียงพอต่อการอธิบายความแปรปรวนที่มีเงื่อนไข

ประการ ที่ สาม เรา สามารถ ทดสอบ ว่า ข้อ สมมุติ เกี่ยว กับ การ แจกแจง ของ standardized residuals ที่เราสมมุติว่ามีการแจกแจงแบบปกติมีความสมเหตุสมผลหรือไม่ โดยตั้งสมมุติหลักว่า standardized residuals มีการแจกแจงแบบปกติ หากเราปฏิเสธสมมุติฐานหลัก เราสามารถที่จะปรับปรุงแบบจำลองโดยการสมมุติการแจกแจงอื่นที่มีความคล่องตัวมากกว่าการแจกแจงแบบปกติ ซึ่งจะนำเสนอตัวอย่างในหัวข้อต่อไป

ตัวอย่างที่ 4.2 การสร้างแบบจำลอง GARCH

ต่อ จาก ตัวอย่าง 4.1 เรา จะ ประมาณ ค่า สมการ ค่า เฉลี่ย ด้วย $ARMA(2,2)$ และ สมการ ค่า ความ แปรปรวน ที่ มี เงื่อนไข ด้วย $GARCH(1,1)$ หรือ ใช้ สัญลักษณ์ ว่า $ARMA(2,2) - GARCH(1,1)$ โดยเราจะใช้คำสั่ง `ugarchfit` จาก package `rugarch` ในการประมาณค่าแบบจำลอง GARCH ซึ่งเป็นส่วนที่ประมาณค่า conditional variance เพิ่มขึ้นจากแบบจำลองที่อธิบาย conditional mean เช่น ARMA(2,2)

ใน การ สร้าง แบบ จำลอง เรา จะ ต้อง ระบุ รูป แบบ ของ แบบ จำลอง `ugarchspec` โดยที่ `ugarchspec` คือการระบุรูปแบบของสมการ univariate GARCH ซึ่งประกอบด้วยส่วนหลัก ๆ สามส่วนได้แก่

- `variance.model=list(model="sGARCH", garchOrder=c(1,1))` ส่วน นี้ ระบุ รูปแบบ ของ conditional variance เป็น standard GARCH (sGARCH) และ อันดับ c(1,1) คือ GARCH(1,1)
- `mean.model=list(armaOrder=c(2,2), include.mean=TRUE` ส่วนนี้ระบุรูปแบบของ conditional mean เป็น ARMA(2,2) และมีส่วนของค่าคงที่ (include.mean = TRUE)
- `distribution.model = "norm"` ส่วนนี้ระบุการแจกแจงเป็นแบบปกติ

จะได้ผลดังนี้

```
> library(rugarch)
> spec.garch11 <- ugarchspec(variance.model=list(model="sGARCH",
    garchOrder=c(1,1)),
+ mean.model=list(armaOrder=c(2,2), include.mean=TRUE), distribution.model="norm")
> fit.garch11<-ugarchfit(spec=spec.garch11, data = ret)
> show(fit.garch11)

*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model : sGARCH(1,1)
Mean Model  : ARFIMA(2,0,2)
Distribution  : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.005168    0.003323   1.5552 0.119893
ar1     0.929774    0.105519   8.8114 0.000000
ar2    -0.741403    0.131759  -5.6270 0.000000
ma1    -0.885132    0.079630 -11.1155 0.000000
ma2     0.813870    0.126564   6.4305 0.000000
omega   0.000194    0.000084   2.3165 0.020528
alpha1  0.221012    0.045507   4.8567 0.000001
beta1   0.777524    0.038916  19.9795 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.005168    0.004069   1.2701 0.204050
ar1     0.929774    0.134138   6.9315 0.000000
ar2    -0.741403    0.245777  -3.0166 0.002557
ma1    -0.885132    0.082263 -10.7597 0.000000
ma2     0.813870    0.253515   3.2103 0.001326
omega   0.000194    0.000120   1.6161 0.106077
alpha1  0.221012    0.063004   3.5079 0.000452
beta1   0.777524    0.051986  14.9564 0.000000

LogLikelihood : 542.9161
```

เราสามารถแสดงผลการประมาณค่าได้โดย

$$
\begin{aligned}
y_t &= 0.93y_{t-1} - 0.74y_{t-2} + \varepsilon_t - 0.89\varepsilon_{t-1} + 0.81\varepsilon_{t-2} \\
\sigma_t^2 &= 0.000194 + 0.22\varepsilon_{t-1}^2 + 0.78\sigma_{t-1}^2
\end{aligned}
$$

นอกจากนี้ ผลการประมาณค่ายังแสดงผล Information Criteria สำหรับการเปรียบเทียบรูปแบบแบบจำลอง

```
Information Criteria
------------------------------------

Akaike       -2.3721
Bayes        -2.2992
Shibata      -2.3727
Hannan-Quinn -2.3434
```

นอกจากนี้ คำสั่งดังกล่าวยังแสดงผลการทดสอบ Standardized Residuals ดังนี้

สำหรับกรณีของ Standardized Residuals หากพิจารณาที่ lag 1, 11,19 พบว่าค่า p-value น้อยกว่าระดับนัยสำคัญที่เราเลือก แสดงว่าแบบจำลองในส่วนของ conditional mean ไม่สามารถอธิบายตัวแปรที่เราสนใจได้ (เราจำเป็นต้องปรับแบบจำลองในส่วนของ ARMA(p,q))

ในขณะที่กรณีของ Standardized Squared Residuals หากพิจารณาที่ lag 1, 5, 9 พบว่าค่า p-value มากกว่าระดับนัยสำคัญที่เลือก เช่น 0.05 แสดงว่าแบบจำลองในส่วนของ conditional variance สามารถอธิบายตัวแปรได้

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                          statistic   p-value
Lag[1]                        3.351 6.717e-02
Lag[2*(p+q)+(p+q)-1][11]      9.220 1.777e-06
Lag[4*(p+q)+(p+q)-1][19]     13.350 8.841e-02
d.o.f=4
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                         statistic p-value
Lag[1]                      0.0957  0.7571
Lag[2*(p+q)+(p+q)-1][5]     1.8392  0.6569
Lag[4*(p+q)+(p+q)-1][9]     3.0164  0.7563
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]     1.876 0.500 2.000  0.1708
ARCH Lag[5]     3.360 1.440 1.667  0.2417
ARCH Lag[7]     3.586 2.315 1.543  0.4103
```

4.4.2 การแจกแจงที่ไม่ใช่การแจกแจงแบบปกติ

ในหัวข้อที่ผ่านมาเราสมมุติให้ ε_t^2 มีการแจกแจงแบบปกติมาตรฐานซึ่งข้อมูลทางการเงินบางข้อมูลอาจจะมีลักษณะหางอ้วน (fat tail) ดังนั้น ข้อสมมุติว่า ε_t^2 มีการแจกแจงอื่นที่มีหางอ้วนน่าจะเหมาะสมกว่า เช่นการแจกแจงแบบ student't และ generalized error distribution (GED)

การแจกแจงแบบ Student's t

ถ้าตัวแปรสุ่ม u_t มีการแจกแจงแบบ t ด้วยองศาเสรี (degree of freedom) เท่ากับ ν "นิว" และตัวแปรสเกล s_t จะมีฟังก์ชันความถี่ความน่าจะเป็น (pdf) เท่ากับ

$$f(u_t) = \frac{\Gamma[(\nu+1)/2]}{(\pi\nu)^{1/2}\Gamma(\nu/2)} \frac{s_t^{-1/2}}{[1+u_t^2/(s_t\nu)]^{(\nu+1)/2}}$$

โดยที่ $\Gamma()$ เป็นฟังก์ชันแกมมา และค่าความแปรปรวนของ u_t เท่ากับ

$$Var(u_t) = \frac{s_t\nu}{\nu-2}, \qquad \nu > 2$$

ถ้าค่าช็อก ε_t ในแบบจำลอง $GARCH$ มีการแจกแจงแบบ t ที่มีองศาอิสระเท่ากับ ν และมีค่า $Var_{t-1}(\varepsilon_t) = \sigma_t^2$ แล้วพารามิเตอร์ s_t จะถูกเลือกเพื่อให้ $s_t = \frac{\sigma_t^2(\nu-2)}{\nu}$

เงื่อนไขเพิ่มเติมในคำสั่ง `garchFit` คือ `cond.dist=c("std")` สำหรับการแจกแจงแบบ student's t

Generalized Error Distribution

Nelson (1991) ได้เสนอให้ใช้การแจกแจง generalized error distribution (GED) ในการผนวกคุณลักษณะหางอ้วนในแบบจำลอง ถ้า u_t มีการแจกแจงแบบ GED ที่มีค่าเฉลี่ยเท่ากับศูนย์และค่าความแปรปรวนเท่ากับหนึ่ง ฟังก์ชันความถี่ความน่าจะเป็น (pdf) เท่ากับ

$$f(u_t) = \frac{\nu \exp[-(1/2)|u_t/\lambda|^\nu}{\lambda 2^{(\nu+1)/\nu}\Gamma(1/\nu)}$$

โดยที่

$$\lambda = \left[\frac{2^{-2/\nu}\Gamma(1/\nu)}{\Gamma(3/\nu)}\right]^{1/2}$$

โดยที่ nu เป็นค่าพารามิเตอร์ที่เป็นบวกที่บอกความหนาของหาง ถ้า $\nu = 2$ pdf จะลดลงเหลือ pdf ของการแจกแจงปกติมาตรฐาน และถ้า $\nu < 2$ ความถี่ที่หางจะหนากว่าการแจกแจงแบบปกติ

เงื่อนไขเพิ่มเติมในคำสั่ง `garchFit` คือ `cond.dist=c("ged")` สำหรับ generalized error distribution

ตัวอย่าง 4.3: การสร้างแบบจำลอง GARCH ด้วยข้อสมมุติการแจกแจงแบบ t

ตัวอย่างนี้ใช้ข้อมูลต่อจากตัวอย่าง 4.2 เนื่องจากในตัวอย่างที่ผ่านมาเราพบว่า residuals ไม่ ได้ มี การ แจกแจง แบบ ปกติ ดัง นั้น เรา อาจ จะ ใช้ การ แจกแจง อื่น ที่ มี หาง อ้วน กว่า เช่น การ แจกแจง แบบ student's t และ ประมาณ ค่า แบบ จำลอง `fit.garch11.t` โดยที่ เปลี่ยน

distribution.model="std")

```
> spec.garch11.t <- ugarchspec(variance.model=list(model="sGARCH",
    garchOrder=c(1,1)),
+ mean.model=list(armaOrder=c(2,2), include.mean=TRUE), distribution.model="std")
> fit.garch11.t<-ugarchfit(spec=spec.garch11.t, data = ret)
> show(fit.garch11.t)

*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model : sGARCH(1,1)
Mean Model  : ARFIMA(2,0,2)
Distribution  : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.005531    0.003261   1.6962 0.089852
ar1     0.977050    0.175427   5.5695 0.000000
ar2    -0.602677    0.150885  -3.9943 0.000065
ma1    -0.880078    0.155925  -5.6443 0.000000
ma2     0.648087    0.154766   4.1875 0.000028
omega   0.000156    0.000104   1.4966 0.134492
alpha1  0.184988    0.056949   3.2483 0.001161
beta1   0.814012    0.051417  15.8317 0.000000
shape   4.918851    1.144499   4.2978 0.000017
<Omitted>
Information Criteria
------------------------------------

Akaike       -2.4373
Bayes        -2.3552
Shibata      -2.4381
Hannan-Quinn -2.4050
```

ในการประมาณค่าด้วยการแจกแจงแบบ student's t เราจะได้ค่า estimated shape คือตัวประมาณค่าองศาเสรี (degree of freedom) ในที่นี้คือ shape เท่ากับ 4.92

4.4.3 การเลือกแบบจำลองที่เหมาะสม (Model selection)

จากหัวข้อที่ผ่านมาเราจะเห็นได้ว่าเราสามารถประมาณค่าแบบจำลอง $GARCH$ ได้หลายรูปแบบ ดังนั้นการเลือกแบบจำลองที่เหมาะสมกับข้อมูลของเราเป็นขั้นตอนที่สำคัญ เนื่องจากเราสามารถเขียนแบบจำลองในรูปของ $ARMA$ สำหรับ ε_t^2 ได้ เราจึงสามารถใช้เกณฑ์ในการเลือกแบบจำลองที่เราได้เสนอไว้แล้วในบทที่แล้วเช่น Akaike information criteria (AIC) หรือ Baysian information criteria (BIC) ในการเลือกแบบจำลอง

ตัวอย่าง

จาก ตัวอย่างที่ 5.2และ 5.3 เราสามารถนำค่า AIC และ BIC มาเปรียบเทียบกันตามตาราง

ตารางที่ 4.1: การเปรียบเทียบแบบจำลอง $GARCH$

เกณฑ์	$ARMA(2,2) - GARCH(1,1)$ `fit.garch11`	$ARMA(2,2) - GARCH(1,1)$ t-dist `fit.garch11.t`
AIC	-2.372	-2.437
BIC	-2.299	-2.355

จะเห็นได้ว่าค่า AIC และ BIC ของ $GARCH$ ที่มีการแจกแจงแบบ student's t มีค่าน้อยกว่า AIC และ BIC ของ $GARCH$ ที่มีการแจกแจงแบบปกติ แสดงให้เห็นว่าแบบจำลองที่มีการแจกแจงแบบ student's t น่าจะอธิบายข้อมูลได้ดีกว่า

4.4.4 การขยายแบบจำลอง $GARCH$

แบบจำลอง $GARCH$ สามารถปรับปรุงเพื่อให้สามารถอธิบายความผันผวนของข้อมูลจริงได้ดีขึ้น

Integrated $GARCH$

จากตัวอย่างการประมาณค่า $GARCH(1,1)$ สำหรับผลตอบแทนของหลักทรัพย์มักจะให้ค่าประมาณ $\hat{\alpha}_1 + \hat{\beta}_1 \approx 1$ ตัวอย่างเช่นการประมาณค่าในตัวอย่างที่ 4.2 $\hat{\alpha}_1 + \hat{\beta}_1 = 0.22 + 0.777 = 0.997$

จากสมการที่ xxx เราพบว่าอนุกรมเวลาความผันผวนมีลักษณะใกล้เคียงกับ unit root หากเขาขยายแบบจำลอง GARCH ให้มีลักษณะ unit root เราเรียกแบบจำลองดังกล่าวว่า Integrated GARCH (IGARCH) โดยที่เราสามารถผนวกเงื่อนไขที่ว่าความผันผวนเป็น unit root ได้โดยการกำหนดให้ $\alpha_1 + \beta_1 = 1$ หรือแทนค่า $\beta_1 = 1 - \alpha_1$ ในสมการ (**??**) ส่งผลให้ conditional variance เท่ากับ

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + (1 - \alpha_1)\sigma_{t-1}^2 \tag{4.18}$$

จะเห็นได้ว่าจำนวนพารามิเตอร์ในแบบจำลองดังกล่าวเหลือเพียงแค่ 2 ตัวได้แก่ ω กับ α_1

การเพิ่มตัวแปรในแบบจำลอง GARCH

แนวทางหนึ่งในการขยายแบบจำลอง GARCH คือการเพิ่มตัวแปรในสมการค่าเฉลี่ยและสมการค่าความแปรปรวน เช่นในกรณีของแบบจำลอง $GARCH(1,1)$ สามารถขยายแบบจำลอง

ได้ดังสมการต่อไปนี้

$$
\begin{aligned}
y_t &= \mu_t + \gamma_1 x_t + \epsilon_t \\
\sigma_t^2 &= \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \lambda w_t
\end{aligned}
$$

โดยที่ x_t และ w_t คือตัวแปรอธิบายเพิ่มเติมสำหรับสมการค่าเฉลี่ยและสมการค่าความแปรปรวนตามลำดับ ตัวอย่างเช่นค่าความแปรปรวนอาจจะอธิบายได้ด้วยปริมาณการซื้อขาย การประกาศนโยบายของทางการ วันในสัปดาห์ และผลของวันหยุด

การเปลี่ยนแปลงที่ไม่สมมาตร

ในแบบจำลอง $GARCH$ เราพิจารณาตัวกำหนดความผันผวนด้วย ε_{t-1}^2 โดยที่ไม่ได้สนใจว่าช็อกที่เกิดขึ้นเป็นช็อกทางบวกหรือทางลบ ทั้งๆที่ช็อกทางลบ(ในขนาดที่เท่ากัน)น่าจะทำให้เกิดความผันผวนมากกว่าช็อกทางบวก ดังนั้นแนวทางการขยายแบบจำลอง $GARCH$ อีกทางหนึ่งคือการปรับแบบจำลองให้ผลของช็อกในทางบวกและทางลบมีค่าที่แตกต่างกัน โดยที่แบบจำลองในตลาดหลักทรัพย์เรามักจะพบว่าการปรับตัวลดลงของราคาหลักทรัพย์ส่งผลให้เกิดความผันผวนมากกว่าการปรับตัวขึ้นของราคา แบบจำลองที่ขยายขึ้นมานี้เราเรียกว่าเป็นแบบจำลองที่มีผลไม่สมมาตร

วิธีการแรกในการสร้างแบบจำลองที่มีผลไม่สมมาตรคือ threshold GARCH ($TGARCH$) ซึ่งสมการค่าความแปรปรวนสามารถอธิบายได้ดังนี้

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \phi d_{t-1} \epsilon_{t-1}^2 \tag{4.19}$$

โดยที่ d_{t-1} เป็นตัวแปรดัมมีที่นิยามได้ดังนี้

$$d_{t-1} = \begin{cases} 1: & \epsilon_{t-1} < 0 \\ 0: & \epsilon_{t-1} \geq 0 \end{cases}$$

เพื่อให้ค่าความแปรปรวนเป็นบวก เรากำหนดให้ $\phi > 0$ ดังนั้นหากเกิดข่าวดี ($\epsilon_{t-1} \geq 0$) จะส่งผลต่อความผันผวนเท่ากับ α_1 ในทางตรงข้ามหากเกิดข่าวร้าย ($\epsilon_{t-1} < 0$) จะส่งผลต่อความผันผวนเท่ากับ $\alpha_1 + \phi$ จะเห็นได้ว่าช็อกทางลบส่งผลต่อความผันผวนมากกว่าช็อกทางบวก

อีกวิธีในการสร้างแบบจำลองที่มีผลไม่สมมาตรคือแบบจำลอง Exponential GARCH ($EGARCH$) ซึ่งสามารถแสดงได้ด้วยสมการต่อไปนี้

$$\ln \sigma_t^2 = \omega + \alpha_1 \frac{\epsilon_{t-1}}{\sigma_{t-1}} + \phi \| \frac{\epsilon_{t-1}}{\sigma_{t-1}} \| + \beta_1 \ln \sigma_{t-1}^2 \tag{4.20}$$

การพิจารณาค่าความเสี่ยงในสมการค่าเฉลี่ย

ในทางการเงินความเสี่ยงที่สูงจะนำมาสู่ผลตอบแทนที่สูงด้วย หรือเราเรียกว่า "เบี้ยความเสี่ยง (risk premium)" Engle, Lilien, and Robins (1987) เสนอให้ขยาย $GARCH$ เพื่อให้พิจารณาเบี้ยความเสี่ยงเป็นปัจจัยหนึ่งในการกำหนดผลตอบแทนในสมการค่าเฉลี่ย หรือสามารถอธิบายได้ด้วยสมการ

$$y_t = c + \alpha g(\sigma_t) + \varepsilon_t$$

และเราเรียกแบบจำลองนี้ว่า GARCH-in-the-mean ($GARCH - M$)

4.5 การพยากรณ์ด้วยแบบจำลอง $GARCH$

เป้าหมายสำคัญของการสร้างแบบจำลองสำหรับ conditional volatility คือการสร้างการคาดการณ์ที่แม่นยำขึ้นสำหรับการทำนายค่าอนุกรมเวลาที่เราสนใจในอนาคตและการพยากรณ์ conditional variance

เนื่องจากในสมการค่าเฉลี่ยเราสมมุติให้อนุกรมเวลาถูกอธิบายด้วยแบบจำลอง $ARMA$ ดังนั้นค่าพยากรณ์ของอนุกรมเวลา y_t สามารถใช้การพยากรณ์ของ $ARMA$ ได้อย่างปกติ อย่างไรก็ตามการที่เราอนุญาตให้ความผันผวนสามารถเปลี่ยนแปลงได้ตามเวลา เช่นแบบจำลอง $GARCH$ จะทำให้เราได้ค่าผันผวนในอนาคตที่แม่นยำขึ้นโดยเฉพาะการพยากรณ์ในระยะสั้น ในที่นี้เราจะยกตัวอย่างแบบจำลอง $GARCH(1,1)$ ซึ่งสมการความผันผวนสามารถแสดงได้โดย

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

หากเราพิจารณาข้อมูลจนถึงคาบที่ h และการพยากรณ์ไปข้างหน้าหนึ่งคาบสามารถคำนวณได้โดยใช้ conditional expectation

$$\begin{aligned} E_h(\sigma_{h+1}^2) &= \omega + \alpha_1 E_h(\varepsilon_h^2) + \beta E_h(\sigma_h^2) \\ &= \omega + \alpha_1 \varepsilon_h^2 + \beta_1 \sigma_h^2 \end{aligned}$$

เนื่องจากเราทราบว่า ε_h^2 และ σ_h^2 จากการประมาณค่าแล้ว ต่อไปพิจารณาคาบที่ $h+2$

$$\begin{aligned} E_h(\sigma_{h+2}^2) &= \omega + \alpha_1 E_h(\varepsilon_{h+1}^2) + \beta E_h(\sigma_{h+1}^2) \\ &= \omega + (\alpha_1 + \beta_1) E_h(\sigma_{h+1}^2) \end{aligned}$$

เนื่องจาก $E_h(^2_{h+1}) = E_h(\sigma_{h+1}^2)$ จากวิธีการข้างต้นเราสามารถสร้างสมการสำหรับการพยากรณ์ ณ

คาบ k ใดๆได้เป็น

$$E_h(\sigma^2_{h+k}) = \omega \sum_{i=1}^{k-2} (\alpha_1 + \beta_1)^i + (\alpha_1 + \beta_1)^{k-1} E_h(\sigma^2_{h+1}) \tag{4.21}$$

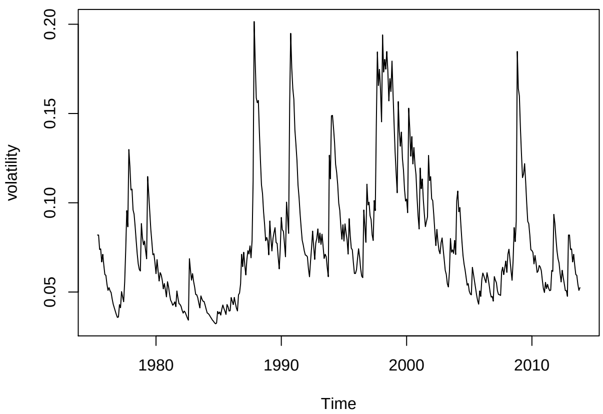
สำหรับ $k \geq 2$ และจะเห็นได้ว่าถ้า $k \to \infty$ และ $GARCH$ เป็นอนุกรมนิ่ง ($\alpha_1 + \beta_1 < 1$) การพยากรณ์ค่าความผันผวนจะเข้าใกล้ unconditional variance $\omega/(1 - \alpha_1 - \beta_1)$

ตัวอย่าง 4.4 การคำนวณ predicted conditional variance และ การพยากรณ์

จาก แบบ จำลอง GARCH เรา สามารถ คำนวณ ค่า predicted conditional variance หรือ standard deviation กับข้อมูลภายในกลุ่มตัวอย่าง (in-sample) เพื่อใช้เป็นตัวแทนของความผันผวน (volatility) สำหรับแบบจำลองทางการเงินต่าง ๆ ในตัวอย่างที่ 5.2 เราได้ประมาณค่าแบบจำลอง ARMA(2,2)-GARCH(1,1) ซึ่งเราสามารถคำนวณค่า predicted conditional standard deviation ได้จากคำสั่ง |sigma|

```
sig.garch11<-sigma(fit.garch11)
volatility<-ts(data=sig.garch11[,1], frequency=12, start=c(1975,5),
    end=c(2013,11))
plot.ts(volatility, type="l")
```

รูปที่ 4.3: ACF ของ residuals



จากรูป **??** จะเห็นได้ว่าช่วงที่ตลาดหุ้นไทยผันผวนค่อนข้างสูงและยาวนานคือหลังวิกฤติ 1997

นอกจากนี้ เราสามารถพยากรณ์ออกไปนอกช่วงข้อมูล (out-of sample) ซึ่งเราสามารถพยากรณ์ค่าเฉลี่ยและค่าความแปรปรวน เราสามารถใช้คำสั่ง `ugarchforecast` และระบุแบบจำลอง |

fit.garch11| และจำนวนคาบที่จะพยากรณ์ไปข้างหน้า `n.ahead=5` และได้ผลดังต่อไปนี้

```
> garch11.fcst=ugarchforecast(fit.garch11, n.ahead=5)
> garch11.fcst

*------------------------------------*
*       GARCH Model Forecast         *
*------------------------------------*
Model: sGARCH
Horizon: 5
Roll Steps: 0
Out of Sample: 0

0-roll forecast [T0=1971-03-28 07:00:00]:
        Series   Sigma
T+1  0.0008783 0.04438
T+2 -0.0026793 0.04648
T+3  0.0010525 0.04849
T+4  0.0071598 0.05041
T+5  0.0100715 0.05226
```

หรือสามารถเขียนสรุปได้เป็น

horizon	1	2	3	4	5
log return	0.0009	-0.0027	0.0011	0.0072	0.0101
volatility (σ_t)	0.0444	0.0465	0.0485	0.0504	0.0523

บทที่ 5

แบบจำลองอนุกรมเวลาเชิงพหุ (Multivariate time series model)

เนื่องจากตลาดการเงินของโลกมีความเชื่อมโยงกันมากขึ้น การเปลี่ยนแปลงที่เกิดขึ้นในตลาดการเงินหนึ่งสามารถส่งผ่านต่อไปยังตลาดอื่นได้ง่าย ดังนั้น บางครั้งเราจำเป็นต้องพิจารณาการเปลี่ยนแปลงของอนุกรมเวลาทางเงินไปพร้อมๆกัน โดยเราเรียกแบบจำลองที่พิจารณาอนุกรมเวลาหลายๆอนุกรมพร้อมกันว่าแบบจำลองอนุกรมเวลาเชิงพหุ(multivariate time series) โดยเราสามารถเขียนอนุกรมเวลาในรูปของเวกเตอร์ $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, ..., y_{nt})'$ โดยที่ y_{it} แทนอนุกรมเวลา i และ n คือจำนวนอนุกรมเวลาที่เราพิจารณาร่วมกัน[1] เช่น $y_{1t}, y_{2t}, ...y_{nt}$ แทนผลได้ตอบแทนในรูปล็อกของหุ้นในตลาดหลักทรัพย์แห่งประเทศไทย

5.1 ความนิ่งของอนุกรมเวลาและเมตริกซ์สหสัมพันธ์ระหว่างกัน

$\boldsymbol{Y}_t = (y_{1t}, y_{2t}, ..., y_{nt})'$ จะเป็นอนุกรมเวลานิ่ง (stationary) ถ้าโมเมนต์ที่หนึ่งและสองไม่ขึ้นอยู่กับเวลา (time-invariant) ซึ่งในที่นี้เราจะต้องพิจารณาเวกเตอร์ของค่าเฉลี่ย (mean vector) $\boldsymbol{\mu} = E(\boldsymbol{Y}_t)$ โดยที่ $\boldsymbol{\mu} = (\mu_1, ..., \mu_n)'$

เมตริกซ์ของความแปรปรวนร่วม (covariance matrix) ใช้สัญลักษณ์ Γ_0 สามารถคำนวณได้

[1]สัญลักษณ์ $'$ แทน transpose ของเมทริกซ์

โดย

$$
\begin{aligned}
\Gamma_0 &= E[(\boldsymbol{Y_t} - \boldsymbol{\mu})(\boldsymbol{Y_t} - \boldsymbol{\mu})^{'}] \\
&= \begin{pmatrix} E(y_{1t}-\mu_1)^2 & E(y_{1t}-\mu_1)(y_{2t}-\mu_2) & \dots & E(y_{1t}-\mu_1)(y_{nt}-\mu_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(y_{nt}-\mu_n)(y_{1t}-\mu_1) & E(y_{nt}-\mu_n)(y_{2t}-\mu_2) & \dots & E(y_{nt}-\mu_n)^2 \end{pmatrix} \\
&= \begin{pmatrix} Var(y_{1t}) & Cov(y_{1t}, y_{2t}) & \dots & Cov(y_{1t}, y_{nt}) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(y_{nt}, y_{1t}) & Cov(y_{nt}, y_{2t}) & \dots & Var(y_{nt}) \end{pmatrix}
\end{aligned}
$$

จะเห็นได้ว่าพจน์ในแนวแทยงมุมจะเป็นค่าความแปรปรวนของแต่ละอนุกรมเวลา ในขณะที่พจน์ที่ (i, j) จะเป็นสหสัมพันธ์ระหว่าง y_{it} และ y_{jt} ในคาบเวลาเดียวกัน

5.1.1 cross-correltaion matrics

กำหนดให้ $\boldsymbol{D}$ เป็นเมทริกซ์ $n \times n$ ที่ประกอบด้วยค่าเบี่ยงเบนมาตรฐาน (standard deviation) ของ y_{it} ในพจน์แทยงมุม โดยที่ $\Gamma_{ii}(0)$ เป็นสัญลักษณ์แทนค่าความแปรปรวนของ y_{it} ดังนั้นเราสามารถเขียนเมทริกซ์ $\boldsymbol{D}$ ได้ดังนี้

$$
\boldsymbol{D} = \begin{pmatrix} \sqrt{\Gamma_{11}(0)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sqrt{\Gamma_{nn}(0)} \end{pmatrix}
$$

เรา สามารถ นิยาม เม ทริกซ์ สห สัมพันธ์ ข้าม ตัวแปร (cross-correlation) ที่ คาบ เวลา เดียวกัน (concurrent) ได้ด้วย

$$
\boldsymbol{\rho}_0 = [\rho_{ij}(0)] = \boldsymbol{D}^{-1}\boldsymbol{\Gamma}_0\boldsymbol{D}^{-1}
$$

โดยที่ $\rho_{ij}(0) = \frac{\Gamma_{ij}(0)}{\sqrt{\Gamma_{ii}(0)\Gamma_{jj}(0)}} = \frac{Cov(y_{it}, y_{jt})}{std(y_{it})std(y_{jt})}$ ซึ่งคือสหสัมพันธ์ระหว่าง y_{it} และ y_{jt} โดยที่ในการวิเคราะห์อนุกรมเวลาเราจะเรียกสหสัมพันธ์ดังกล่าวว่าเป็นสหสัมพันธ์ในคาบเดียวกัน (concurrent หรือ contemporaneous)

นอกจากนี้ การวิเคราะห์อนุกรมเวลาเชิงพหุ เราสนใจความสัมพันธ์ในเชิงตัวแปรที่นำหรือตาม (lead-lag relationship) ระหว่างตัวแปรที่เราสนใจ ดังนั้นเราต้องการหาเมทริกซ์สหสัมพันธ์ข้ามตัวแปรที่ใช้วัดความสัมพันธ์ระหว่างช่วงเวลาเช่น เมทริกซ์ความแปรปรวนร่วมข้ามตัวแปรที่มี

ค่าล่าเท่ากับหนึ่งใช้สัญลักษณ์ Γ_1 ดังที่ได้แสดงในสมการต่อไปนี้

$$\begin{aligned}\Gamma_1 &= E[(\boldsymbol{Y_t}-\boldsymbol{\mu})(\boldsymbol{Y_{t-1}}-\boldsymbol{\mu})'] \\ &= \begin{pmatrix} E(y_{1t}-\mu_1)(y_{1,t-1}-\mu_1) & E(y_{1t}-\mu_1)(y_{2,t-1}-\mu_2) & \dots & E(y_{1t}-\mu_1)(y_{n,t-1}-\mu_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(y_{nt}-\mu_n)(y_{1,t-1}-\mu_1) & E(y_{nt}-\mu_n)(y_{2,t-1}-\mu_2) & \dots & E(y_{nt}-\mu_n)(y_{n,t-1}-\mu_n) \end{pmatrix} \\ &= \begin{pmatrix} Cov(y_{1t},y_{1,t-1}) & Cov(y_{1t},y_{2,t-1}) & \dots & Cov(y_{1t},y_{n,t-1}) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(y_{nt},y_{1,t-1}) & Cov(y_{nt},y_{2,t-1}) & \dots & Cov(y_{nt},y_{n,t-1}) \end{pmatrix}\end{aligned}$$

เราสามารถคำนวณเมทริกซ์สหสัมพันธ์ข้ามตัวแปร (cross-correlation) ที่มีค่าล่าเท่ากับหนึ่ง ได้ด้วย

$$\boldsymbol{\rho}_1 = [\rho_{ij}(1)] = \boldsymbol{D}^{-1}\boldsymbol{\Gamma}_1\boldsymbol{D}^{-1}$$

โดยที่ $\rho_{ij}(1) = \frac{\Gamma_{ij}(1)}{\sqrt{\Gamma_{ii}(0)\Gamma_{jj}(0)}} = \frac{Cov(y_{it},y_{j,t-1})}{std(y_{it})std(y_{jt})}$ ซึ่งคือสหสัมพันธ์ระหว่าง y_{it} และ $y_{j,t-1}$ หากค่า $\rho_{ij}(1) \neq 0$ แสดงว่าการเปลี่ยนแปลงของ y_{jt} ในคาบที่ผ่านมาส่งผลต่อการเปลี่ยนแปลงของ y_{it} ในคาบปัจจุบัน ดังนั้น y_{jt} เป็นตัวแปรนำ (lead) y_{it}

ความสัมพันธ์ดังกล่าวสามารถขยายไปยังกรณี $\Gamma_l, l = ..., -2, -1, 0, 1, 2, ...$ ถ้าหากค่า $\rho_{ij}(l) \neq 0$ กรณีที่ $l > 0$ แสดงว่า y_{jt} เป็นตัวแปรนำ (lead) y_{it} แต่ถ้า $\rho_{ij}(l) \neq 0$ กรณีที่ $l < 0$ แสดงว่า y_{it} เป็นตัวแปรนำ (lead) y_{jt}

5.2 แบบจำลองเวกเตอร์ออโตรีเกรสซีฟ (Vector Autoregressive Model)

กำหนดให้ $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, ..., y_{nt})'$ แทน $(n \times 1)$ เวกเตอร์ของตัวแปรอนุกรมเวลาจำนวน n ตัวแปร แล้วแบบจำลองเวกเตอร์ออโตรีเกรสซีฟที่มีค่าล่าเท่ากับ p ซึ่งใช้สัญลักษณ์ $VAR(p)$ สามารถเขียนในรูป

$$\boldsymbol{Y}_t = \boldsymbol{c} + \Phi_1\boldsymbol{Y}_{t-1} + ... + \Phi_p\boldsymbol{Y}_{t-p} + \boldsymbol{\varepsilon}_t, \qquad t = 1, 2, ..., T \tag{5.1}$$

โดยที่ Φ_i เป็นเมทริกซ์สัมประสิทธิ์ที่มีขนาด $n \times n$ และ $\boldsymbol{\varepsilon}_t$ เป็นเวคเตอร์ของกระบวนการไวทนอซที่มีค่าเฉลี่ยเท่ากับศูนย์ และมีเมตริกซ์ความแปรปรวนร่วมที่ไม่แปรผันตามเวลา $\boldsymbol{\Sigma}$ ตัวอย่างเช่น หากเราสนใจแบบจำลอง VAR ที่มีตัวแปรที่เราสนใจสองตัวและมีค่าล่าเท่ากับ 1 เราสามารถเขียนใน

รูปเมตริกซ์ได้เป็น

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \phi^1_{11} & \phi^1_{12} \\ \phi^1_{21} & \phi^1_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{5.2}$$

หรือเขียนเป็นรูปสมการของแต่ละตัวแปรที่เราสนใจ

$$y_{1t} = c_1 + \phi^1_{11} y_{1t-1} + \phi^1_{12} y_{2t-1} + \varepsilon_{1t}$$
$$y_{2t} = c_2 + \phi^1_{21} y_{1t-1} + \phi^1_{22} y_{2t-1} + \varepsilon_{2t}$$

โดยที่ $Cov(\varepsilon_{1t}, \varepsilon_{2s}) = \sigma_{12}$ ถ้า $s = t$ และ $Cov(\varepsilon_{1t}, \varepsilon_{2s}) = 0$ ถ้า $s \neq t$ (คือไม่มีสหสัมพันธ์ข้ามเวลาระหว่างช็อกของตัวแปรที่ต่างกัน) จากแบบจำลองข้างต้นจะเป็นได้ว่าตัวแปร y_{1t} ถูกอธิบายด้วย ค่าในอดีตของตัวเอง y_{1t-1}และตัวแปรอื่น y_{2t-1}

ค่าสัมประสิทธิ์แต่ละตัวจะเป็นค่าที่บอกความสัมพันธ์ระหว่างตัวแปรที่ต่างเวลากัน ตัวอย่างเช่น ϕ^1_{12} จะแสดงความสัมพันธ์ระหว่าง y_{1t} และ y_{2t-2} ซึ่งถ้า $\phi^1_{12} = 0$ แล้ว y_{1t} จะไม่ขึ้นกับค่า y_{2t-1} และขึ้นอยู่กับค่าในอดีตของ y_{1t} เท่านั้น เช่นเดียวกับค่า ϕ^1_{21} จะแสดงความสัมพันธ์ระหว่าง y_{2t} และ y_{1t-1} ซึ่งจะเห็นได้ว่าค่าสัมประสิทธิ์ในสมการ **??** อธิบายพลวัตของอนุกรมเวลาที่เราสนใจ (lead-lag) แต่มิได้อธิบายความสัมพันธ์ระหว่างอนุกรมเวลาในช่วงเวลาเดียวกัน (concurrent หรือ contemporaneous) ไว้อย่างชัดเจน เราสามารถดูความสัมพันธ์ดังกล่าวได้ในเมทริกซ์ความแปรปรวนของช็อก หรือพูดให้ชัดเจนคือการพิจารณาค่า $Cov(\varepsilon_{1t}, \varepsilon_{2t})$ นั่นเอง เราเรียกแบบจำลองเวกเตอร์ออโตรีเกรสซีฟในรูป **??** ว่า**สมการในรูปลดรูป (reduced form)**

เราสามารถเขียนความสัมพันธ์ระหว่างอนุกรมเวลาในช่วงเดียวกันได้อย่างชัดเจนโดยใช้การแปลงรูปสมการที่ **??** ได้ดังนี้ เนื่องจาก $\mathbf{\Sigma}$ เป็นเมตริกซ์ที่มีคุณลักษณะเป็น positive definite ดังนั้นเราสามารถสร้างเมทริกซ์ lower triangular

$$\boldsymbol{L} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & 1 \end{pmatrix}$$

และเมตริกซ์แทยงมุม $\boldsymbol{G}$ ที่ทำให้ $\mathbf{\Sigma} = \boldsymbol{LGL}'$ เราเรียกการแยกส่วนประกอบนี้ว่า Cholesky decomposition

กำหนดให้ $\boldsymbol{\eta}_t = \boldsymbol{L}^{-1}\boldsymbol{\varepsilon}_t$ เราจะได้ว่า

$$E(\boldsymbol{\eta}_t) = \boldsymbol{L}^{-1}E(\boldsymbol{\varepsilon}_t) = \mathbf{0}$$

$$Var(\boldsymbol{\eta}_t) = \boldsymbol{L}^{-1}Var(\boldsymbol{\varepsilon}_t)(\boldsymbol{L}^{-1})' = \boldsymbol{L}^{-1}\mathbf{\Sigma}(\boldsymbol{L}')^{-1} = \boldsymbol{L}^{-1}\boldsymbol{LGL}'(\boldsymbol{L}')^{-1} = \boldsymbol{G}$$

เนื่องจาก $\boldsymbol{G}$ เป็นเมตริกซ์แทยงมุม ดังนั้นส่วนประกอบของ $\boldsymbol{\eta}_t$ จะไม่มีความสัมพันธ์ระหว่างกัน

หากเราคูณข้างหน้าสมการ **??** ด้วย $\boldsymbol{L}^{-1}$ จะได้

$$\boldsymbol{L}^{-1}\boldsymbol{Y}_t = \boldsymbol{L}^{-1}\boldsymbol{c} + \boldsymbol{L}^{-1}\Phi_1\boldsymbol{Y}_{t-1} + ... + \boldsymbol{L}^{-1}\Phi_p\boldsymbol{Y}_{t-p} + \boldsymbol{L}^{-1}\boldsymbol{\varepsilon}_t, \qquad t = 1, 2, ..., T$$
$$\boldsymbol{L}^{-1}\boldsymbol{Y}_t = \boldsymbol{c}^* + \Phi_1^*\boldsymbol{Y}_{t-1} + ... + \Phi_p^*\boldsymbol{Y}_{t-p} + \boldsymbol{\eta}_t \tag{5.3}$$

โดยที่เมตริกซ์ $\boldsymbol{L}^{-1}$ จะแสดงความสัมพันธ์ระหว่างอนุกรมเวลาในคาบเดียวกัน และเราเรียกแบบจำลองเวกเตอร์ออโตรีเกรสซีฟในรูป **??** ว่า**สมการในรูปโครงสร้าง (structural equation)**

ตัวอย่าง 5.1:การแปลงสมการในรูปลดรูปไปเป็นรูปโครงสร้าง]

กำหนดให้แบบจำลองเวกเตอร์ออโตรีเกรสซีฟที่มีสองตัวแปรและมีอันดับเท่ากับหนึ่ง (หรือเรียกย่อๆว่า Bivariate VAR(1)) สามารถแสดงในรูปลดรูปได้ดังนี้

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{5.4}$$

โดยที่ $\Sigma = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ โดยที่เราสามารถใช้ Cholesky decomposition หาเมทริกซ์ $\boldsymbol{L}^{-1} = \begin{pmatrix} 1 & 0 \\ -0.5 & 1 \end{pmatrix}$ หากเราคูณสมการ **??** ด้วย $\boldsymbol{L}^{-1}$ จะได้

$$\begin{pmatrix} 1 & 0 \\ -0.5 & 1 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} 0.2 & 0.3 \\ -0.7 & 0.95 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} \tag{5.5}$$

โดยที่ $\boldsymbol{G} = Var(\eta_t) = \begin{pmatrix} 2 & 0 \\ 0 & 0.5 \end{pmatrix}$ จากสมการที่สองของ **??** เราสามารถเขียนความสัมพันธ์ได้เป็น

$$y_{2t} = 0.5y_{1t} - 0.7y_{1,t-1} + 0.95y_{2,t-1} + \eta_{2t}$$

ซึ่งแสดงความสัมพันธ์ระหว่าง y_{1t} และ y_{2t} ในช่วงเวลาเดียวกันไว้อย่างชัดเจน

เราจะเห็นได้ว่าเราสามารถแปลงแบบจำลองเวกเตอร์ออโตรีเกรสซีฟในรูปลดรูปไปเป็นรูปโครงสร้างได้ด้วย Cholesky decomposition อย่างไรก็ตามในการวิเคราะห์อนุกรมเวลาเรามักจะใช้สมการในรูปลดรูป เนื่องจากเป็นรูปที่ง่ายในการประมาณค่า และเรามักจะเน้นผลในการพยากรณ์ของแบบจำลองซึ่งเราไม่สามารถใช้แบบจำลองในรูปโครงสร้างพยากรณ์ข้อมูลได้เนื่องจากเราไม่ทราบข้อมูลของอนุกรมเวลาอื่นๆในคาบเดียวกัน

5.2.1 เงื่อนไขการเป็นอนุกรมเวลานิ่ง

แบบจำลอง $VAR(p)$ สามารถเขียนในรูปเครื่องหมายขยับไปข้างหลัง (lag operator) ได้ดังนี้

$$\mathbf{\Phi}(L)\boldsymbol{Y}_t = \boldsymbol{c} + \boldsymbol{\varepsilon}_t \tag{5.6}$$

โดยที่ $\mathbf{\Phi}(L) = \boldsymbol{I}_n - \mathbf{\Phi}_1 L - ... - \mathbf{\Phi}_p L^p$ แบบจำลอง $VAR(p)$ จะมีเสถียรภาพถ้ารากของ

$$\det(\boldsymbol{I}_n - \mathbf{\Phi}_1 z - ... - \mathbf{\Phi}_p z^p) = 0$$

มีค่ามากกว่าหนึ่ง (หรือโมดูลัสมากกว่าหนึ่งในกรณีจำนวนเชิงซ้อน) ถ้าสมมุติให้กระบวนการดังกล่าวมีค่าในอดีตที่เป็นอนันต์ แล้วกระบวนการ $VAR(p)$ ที่มีเสถียรภาพจะเป็นกระบวนการนิ่ง และ ergodic โดยมีค่าเฉลี่ย ค่าความแปรปรวนและค่าความแปรปรวนร่วมในตัวเองที่ไม่ขึ้นกับเวลา

ตัวอย่างที่ 5.2

สมมุติว่าเราพิจารณาแบบจำลอง $VAR(1)$ ของสองตัวแปรดังมีเมทริกซ์สัมประสิทธิ์ดังตัวอย่างที่ 5.1 เราสามารถเขียนแบบจำลอง VAR ในรูปของ polynomial matrix ได้ดังนี้

$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix} L\right) \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

โดยที่ polynomial matrix สามารถจัดรูปได้เป็น

$$\mathbf{\Phi}(L) = \begin{bmatrix} 1 - 0.2L & -0.3L \\ 0.6L & 1 - 1.1L \end{bmatrix}$$

ซึ่งดีเทอร์มิแนนต์เท่ากับ $|\mathbf{\Phi}(L)| = (1-0.2L)(1-1.1L) + (0.6)(0.3)L^2 = 1 - 1.3L + 0.4L^2$ เราสามารถหาคำตอบของสมการพหุนามได้เท่ากับ 2 และ 1.25 ซึ่งต่างมีค่าสัมบูรณ์มากกว่าหนึ่ง เราสามารถสรุปได้ว่าตัวแปรในแบบจำลอง VAR ที่มีเมทริกซ์สัมประสิทธิ์ดังกล่าวเป็นตัวแปรนิ่ง (stationary) ด้วยกันทั้งคู่

5.3 การประมาณค่าและการเลือกค่าล่า

5.3.1 การประมาณค่า

การเราพิจารณาสมการ **??** แต่ละสมการจะมีตัวแปรที่เหมือนกัน เราสามารถพิจารณาสมการสำหรับแต่ละตัวแปรออกเป็น

$$\boldsymbol{y}_i = \boldsymbol{Z}\boldsymbol{\phi}_i + \boldsymbol{e}_i, \quad i = 1, ..., n$$

โดยที่ $\boldsymbol{y}_i$ คือ เวกเตอร์ $(T \times 1)$ ของตัวแปรตามในสมการที่ i ในขณะที่ $\boldsymbol{Z}$ คือเมตริกซ์ $(T \times k)$ ที่แถว t แทนด้วย $\boldsymbol{Z}_t' = (1, \boldsymbol{Y}_{t-1}', ..., \boldsymbol{Y}_{t-p}')$, $k = np + 1$ และ $\boldsymbol{\phi}_i$ คือ $(k \times 1)$ เวกเตอร์ของพารามิเตอร์ และ $\boldsymbol{e}_i$ คือ error term ที่มีเวกเตอร์ค่าความแปรปรวนร่วม $\sigma_i^2 \boldsymbol{I}_T$ จะเห็นได้ว่าแต่ละสมการสามารถประมาณค่าได้ด้วย ordinary least squares (OLS) แยกแต่ละสมการ และจะได้ $\hat{\boldsymbol{\Phi}} = \left[\hat{\boldsymbol{\phi}}_1, ..., \hat{\boldsymbol{\phi}}_n\right]$ เป็นเมตริกซ์ $(k \times n)$ ของค่าสัมประสิทธิ์จากการประมาณค่าด้วย OLS หากกำหนดให้เครื่องหมาย vec แทนการนำเมตริกซ์มาเรียงต่อกัน เราจะได้

$$vec(\hat{\Phi}) = \begin{pmatrix} \hat{\boldsymbol{\phi}}_1 \\ \vdots \\ \hat{\boldsymbol{\phi}}_n \end{pmatrix}$$

ภายใต้ข้อสมมุติที่แบบจำลองเป็นแบบกระบวนการนิ่งและ ergodic เราจะได้ $vec(\hat{\Phi})$ มีคุณสมบัติ consistent และมีการแจกแจงแบบปกติเมื่อกลุ่มตัวอย่างใหญ่โดยมีเมตริกซ์ค่าความแปรปรวนเท่ากับ

$$\widehat{aver}(vec(\hat{\Phi})) = \hat{\boldsymbol{\Sigma}} \otimes (\boldsymbol{Z}'\boldsymbol{Z})^{-1}$$

โดยที่ $\hat{\Sigma} = \frac{1}{T-k}\sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t'$ และ $\hat{\varepsilon}_t = \boldsymbol{Y}_t - \hat{\boldsymbol{\Phi}}'\boldsymbol{Z}_t$ เป็น residuals จากการประมาณค่า

5.3.2 การเลือกจำนวนความล่าที่เหมาะสม

การหาความล่าที่เหมาะสมของแบบจำลองเวคเตอร์ออโตรีเกรสซีฟอาจจะทำได้โดยการใช้สูตรการเลือกความล่า โดยการประมาณค่าแบบจำลองเวคเตอร์ออโตรีเกรสซีฟที่มีความล่าเท่ากับ $0, 1, ..., p_{max}$ แล้วเลือกค่า p ที่ทำให้ค่าที่ใช้เลือกมีค่าต่ำที่สุด โดยจะมีสูตรดังนี้

$$IC(p) = \ln|\tilde{\boldsymbol{\Sigma}}(p)| + c_T(n, p) \tag{5.7}$$

โดยที่ $\tilde{\boldsymbol{\Sigma}}(p) = T^{-1}\sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t'$ คือ residual covariance matrix จาก $VAR(p)$, c_T คือลำดับที่ขึ้นกับจำนวนตัวอย่าง และ $c_T(n, p)$ คือฟังก์ชันเบี้ยปรับ (penalty function) กรณีที่เราเลือก n และ p มากๆ โดยที่แบบเกณฑ์ที่เป็นที่นิยมสามเกณฑ์ได้แก่ Akaike (AIC) Schwarz-Bayesian (BIC) และ

Hannan-Quinn (HQIC) ซึ่งแสดงได้ด้วยสมการดังนี้

$$AIC(p) = \ln|\tilde{\Sigma}(p)| + \frac{2}{T}(pn^2)$$
$$BIC(p) = \ln|\tilde{\Sigma}(p)| + \frac{\ln T}{T}(pn^2)$$
$$HQIC(p) = \ln|\tilde{\Sigma}(p)| + \frac{2\ln(\ln T)}{T}(pn^2)$$

ตัวอย่าง 5.3: แบบ จำลอง VAR สำหรับ ผล ตอบแทน ราย วัน ของ SET และ DOWJONES

ในตัวอย่างนี้เราจะพิจารณาความสัมพันธ์ระหว่างผลได้ตอบแทนรายวันจากการลงทุนในตลาดหลักทรัพย์แห่งประเทศไทย (`lret_set`) และตลาดหุ้นดาวโจนส์ (`lret_dowjones`) โดยใช้ข้อมูลระหว่างวันที่ 2 มกราคม 2540 ถึง 18 มกราคม 2556 (ข้อมูลจากธนาคารแห่งประเทศไทย โดยข้อมูลที่ตลาดใดตลาดหนึ่งปิดจะถูกลบออกจากข้อมูล และข้อมูลดัชนีราคาของตลาดหุ้นทั้งสองอยู่ในไฟล์ `set_dj.csv`)

เราสามารถนำเข้าข้อมูลและคำนวณผลได้ตอบแทนในรูปล็อกได้จากคำสั่งต่อไปนี้

```
> setdj<-read.csv(file="set_dj.csv", header=TRUE)
> head(setdj)
> set<-setdj$set
> lret_set<-diff(log(set))*100
> dowjones<-setdj$dowjones
> lret_dowjones<-diff(log(dowjones))*100
```

เราจะได้ข้อมูลสองชุดคือ `lret_set` และ `lret_dowjones` เราจะเขียนข้อมูลทั้งสองในรูปของเมตริกซ์ที่นำเวกเตอร์ทั้งสองมาต่อกันโดยใช้คำสั่ง `cbind`

```
> lret<-cbind(lret_set, lret_dowjones)
```

เราสามารถพิจารณาค่าสถิติเบื้องต้นได้จากคำสั่ง `basicStats` และ `acf` เหมือนกับที่เราเคยใช้ในหัวข้อก่อนได้

หาก เรา ต้องการ ประมาณ ค่า ด้วย แบบ จำลอง เวก เตอร์ ออ โต รี เก รส ซีฟ เรา จะ ใช้ `library(vars)` โดยที่เราอาจจะเริ่มจากการหาค่า p ที่เหมาะสมโดยการใช้คำสั่ง `VAR` พร้อมระบุเมตริกซ์ของข้อมูล `lret` จำนวนค่าล่าสูงสุด `lag.max` และ เกณฑ์ในการเลือกค่าล่า `ic=c("AIC")` โดยที่ `"SC"` จะแทน Schwarz-Bayesian (BIC) และ `"HQ"` Hannan-Quinn (HQIC)

```
> library(vars)
> msel=VAR(lret, lag.max=10, ic=c("AIC"))
> msel

VAR Estimation Results:
=======================

Estimated coefficients for equation lret_set:
=============================================
Call:
lret_set = lret_set.l1 + lret_dowjones.l1 + lret_set.l2 + lret_dowjones.l2 +
    lret_set.l3 + lret_dowjones.l3 + lret_set.l4 + lret_dowjones.l4 + lret_set.l5
    + lret_dowjones.l5 + lret_set.l6 + lret_dowjones.l6 + lret_set.l7 +
    lret_dowjones.l7 + lret_set.l8 + lret_dowjones.l8 + lret_set.l9 +
    lret_dowjones.l9 + const

     lret_set.l1 lret_dowjones.l1      lret_set.l2 lret_dowjones.l2
      0.05556382       0.22898590      -0.33541086       0.32270310
     lret_set.l3 lret_dowjones.l3      lret_set.l4 lret_dowjones.l4
      0.01244410       0.21790099      -0.13694929       0.17900004
     lret_set.l5 lret_dowjones.l5      lret_set.l6 lret_dowjones.l6
     -0.01107252       0.04790222      -0.06473588       0.12631799
     lret_set.l7 lret_dowjones.l7      lret_set.l8 lret_dowjones.l8
      0.01517100       0.05964304      -0.02897568       0.02785310
     lret_set.l9 lret_dowjones.l9            const
      0.03837222       0.08783703      -0.00332124


Estimated coefficients for equation lret_dowjones:
==================================================
<Omitted>
```

จากเกณฑ์ AIC จะเห็นได้ว่าค่าล่าที่เหมาะสมคือ $p = 9$ ถ้าลองใช้เกณฑ์ SIC จะได้ค่าล่าที่เหมาะสมเท่ากับ $p = 4$ ดังนั้นเราจะเลือกใช้ $p = 4$ ในการสร้างแบบจำลอง `model1`

```
> msel2=VAR(lret, lag.max=10, ic=c("SC"))
<omitted>

> model1=VAR(lret, p=4)
> summary(model1)

VAR Estimation Results:
=========================
Endogenous variables: lret_set, lret_dowjones
Deterministic variables: const
Sample size: 3803
Log Likelihood: -15650.859
Roots of the characteristic polynomial:
 0.62  0.62 0.5867 0.5867 0.3557 0.3557 0.2886 0.2886
Call:
VAR(y = lret, p = 4)


Estimation results for equation lret_set:
=========================================
lret_set = lret_set.l1 + lret_dowjones.l1 + lret_set.l2 + lret_dowjones.l2 +
    lret_set.l3 + lret_dowjones.l3 + lret_set.l4 + lret_dowjones.l4 + const

                  Estimate Std. Error t value Pr(>|t|)
lret_set.l1       0.059621   0.016288   3.660 0.000255 ***
lret_dowjones.l1  0.222804   0.032471   6.862 7.91e-12 ***
lret_set.l2      -0.320847   0.016256 -19.738  < 2e-16 ***
lret_dowjones.l2  0.313004   0.032696   9.573  < 2e-16 ***
lret_set.l3       0.023840   0.016122   1.479 0.139313
lret_dowjones.l3  0.205041   0.032958   6.221 5.47e-10 ***
lret_set.l4      -0.103471   0.016011  -6.462 1.16e-10 ***
lret_dowjones.l4  0.151876   0.032994   4.603 4.30e-06 ***
const             0.002529   0.043532   0.058 0.953684
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘’. 0.1 ‘’  1


Residual standard error: 2.683 on 3794 degrees of freedom
Multiple R-Squared: 0.1409, Adjusted R-squared: 0.1391
F-statistic: 77.78 on 8 and 3794 DF,  p-value: < 2.2e-16


Estimation results for equation lret_dowjones:
==============================================
<omitted>

Covariance matrix of residuals:
              lret_set lret_dowjones
lret_set         7.199        -0.556
lret_dowjones   -0.556         1.839

Correlation matrix of residuals:
              lret_set lret_dowjones
lret_set        1.0000       -0.1528
lret_dowjones  -0.1528        1.0000
```

จากผลการประมาณค่าเราสามารถเขียนสมการแสดงผลการประมาณค่าสำหรับสมการของ

log return ของ SET(set_t) ได้ดังนี้

$$\begin{aligned} set_t = 0.0025 + 0.0596set_{t-1} + 0.2228dj_{t-1} - 0.3208set_{t-2} + 0.313dj_{t-2} \\ 0.0238set_{t-3} + 0.205dj_{t-3} - 0.1035set_{t-4} + 0.1519dj_{t-4} \end{aligned}$$

5.4 การพยากรณ์จากแบบจำลองเวคเตอร์ออโตรีเกรสซีฟ

เป้าหมายหลักประการหนึ่งของการพัฒนาแบบจำลองเวคเตอร์ออโตรีเกรสซีฟ ($VAR(p)$) คือการพยากรณ์จากแบบจำลองดังกล่าวเช่นเดียวกับแบบจำลองออโตรีเกรสซีฟ ($AR(p)$) หากเราพิจารณาแบบจำลองเวคเตอร์ออโตรีเกรสซีฟอันดับหนึ่ง ($VAR(1)$) ซึ่งแสดงได้ด้วยสมการต่อไปนี้

$$\boldsymbol{Y}_t = \boldsymbol{\Phi}\boldsymbol{Y}_{t-1} + \boldsymbol{\varepsilon}_t$$

เราต้องการพยากรณ์ไปข้างหน้าหนึ่งคาบด้วยข้อมูลที่มี ณ เวลาที่ h หรือพยากรณ์ค่า

$$\boldsymbol{Y}_{h+1} = \boldsymbol{\Phi}\boldsymbol{Y}_h + \boldsymbol{\varepsilon}_{h+1} \tag{5.8}$$

เราหาค่าพยากรณ์โดยการคำนวณค่าคาดการณ์อย่างมีเงื่อนไข (conditional expectation) ด้วยข้อมูล h, $E_h(Y_{h+1})$, เราจะได้ค่าพยากรณ์เท่ากับ

$$\widehat{\boldsymbol{Y}}_h(1) = E_h(\boldsymbol{Y}_{h+1}) = \boldsymbol{\Phi}\boldsymbol{Y}_h + \underbrace{E_h(\boldsymbol{\varepsilon}_{h+1})}_{=\boldsymbol{0}} = \boldsymbol{\Phi}\boldsymbol{Y}_h \tag{5.9}$$

และค่าคาดเคลื่อนของการพยากรณ์ (forecast error) ของการพยากรณ์ไปข้างหน้าหนึ่งคาบ, $\boldsymbol{e}(1)$, จะเท่ากับ

$$\boldsymbol{e}(1) = \boldsymbol{Y}_{h+1} - E_h(\boldsymbol{Y}_{h+1}) = \boldsymbol{\varepsilon}_{h+1} \tag{5.10}$$

หากพิจารณากรณีตัวแปรสองตัวแปร $\boldsymbol{Y}_t = [Y_{1t} \quad Y_{2t}]'$ เราจะพบว่า

$$\begin{aligned} \widehat{Y}_{1,h}(1) &= \phi_{11}Y_{1,h} + \phi_{12}Y_{2,h} \\ \widehat{Y}_{2,h}(1) &= \phi_{21}Y_{1,h} + \phi_{22}Y_{2,h} \\ e_1(1) &= \varepsilon_{1,h+1} \\ e_2(1) &= \varepsilon_{2,h+1} \end{aligned}$$

ซึ่งแสดงให้เห็นว่าค่าพยากรณ์ของ Y_{1t} และ Y_{2t} ขึ้นอยู่กับตัวแปรทั้งสองในอดีต

สำหรับค่าพยากรณ์สองคาบข้างหน้า เมื่อเราพิจารณาค่าจริง ณ เวลาที่ $h+2$, Y_{h+2}, จะเท่ากับ

$$\boldsymbol{Y}_{h+2} = \boldsymbol{\Phi}\boldsymbol{Y}_{h+1} + \boldsymbol{\varepsilon}_{h+2} = \boldsymbol{\Phi}(\boldsymbol{\Phi}Y_h + \boldsymbol{\varepsilon}_{h+1}) + \boldsymbol{\varepsilon}_{h+2} = \boldsymbol{\Phi}^2\boldsymbol{Y}_h + \boldsymbol{\Phi}\boldsymbol{\varepsilon}_{h+1} + \boldsymbol{\varepsilon}_{h+2} \tag{5.11}$$

เราหาค่าพยากรณ์โดยการคำนวณค่าคาดหมายอย่างมีเงื่อนไขด้วยข้อมูล h เท่ากับ

$$\widehat{\boldsymbol{Y}}_h(2) = E_h(\boldsymbol{Y}_{h+2}) = \boldsymbol{\Phi}^2\boldsymbol{Y}_h + \boldsymbol{\Phi}\underbrace{E_h(\boldsymbol{\varepsilon}_{h+1})}_{=\mathbf{0}} + \underbrace{E_h(\boldsymbol{\varepsilon}_{h+2})}_{=\mathbf{0}} = \boldsymbol{\Phi}^2\boldsymbol{Y}_h \tag{5.12}$$

และค่าคาดเคลื่อนของการพยากรณ์ (forecast error) ของการพยากรณ์ไปข้างหน้าสองคาบ, $\boldsymbol{e}(2)$ จะเท่ากับ

$$\boldsymbol{e}(2) = \boldsymbol{Y}_{h+2} - E_h(\boldsymbol{Y}_{h+2}) = \boldsymbol{\Phi}\boldsymbol{\varepsilon}_{h+1} + \boldsymbol{\varepsilon}_{h+2} \tag{5.13}$$

เมื่อเราพิจารณากรณีของตัวแปรสองตัวแปร เราจะเห็นได้ว่าค่าคาดเคลื่อนของการพยากรณ์ไปข้างหน้าสองคาบของแต่ละตัวแปรจะขึ้นอยู่กับช๊อคของตัวแปรทุกตัวในแบบจำลองดังที่แสดงในสมการต่อไปนี้

$$e_1(2) = \phi_{11}\varepsilon_{1,h+1} + \phi_{12}\varepsilon_{2,h+1} + \varepsilon_{1,h+2} \tag{5.14}$$

$$e_2(2) = \phi_{21}\varepsilon_{1,h+1} + \phi_{22}\varepsilon_{2,h+1} + \varepsilon_{2,h+2} \tag{5.15}$$

หากเราดำเนินการด้วยวิธีการเดิมไปเรื่อยๆ จะได้ค่าพยากรณ์ไปข้างหน้า l คาบ $(\boldsymbol{Y}_{h+l})$ ด้วยข้อมูลที่มี ณ คาบที่ h เท่ากับ

$$\widehat{\boldsymbol{Y}}_h(l) = E_h(\boldsymbol{Y}_{h+l}) = \boldsymbol{\Phi}^l\boldsymbol{Y}_h \tag{5.16}$$

ในขณะที่ค่าคลาดเคลื่อนของการพยากรณ์ไปข้างหน้า l คาบจะเท่ากับ

$$\boldsymbol{e}(l) = \boldsymbol{\Phi}^{l-1}\boldsymbol{\varepsilon}_{h+1} + \boldsymbol{\Phi}^{l-2}\boldsymbol{\varepsilon}_{h+2} + \boldsymbol{\Phi}^{l-3}\boldsymbol{\varepsilon}_{h+3} + ... + \boldsymbol{\varepsilon}_{h+l} \tag{5.17}$$

จากสมการ (**??** เราจะเห็นได้ว่าแบบจำลอง VAR สามารถนำมาใช้ในการสร้างค่าพยากรณ์ตัวแปรต่างๆได้ง่าย และใช้เพียงแค่ตัวแปร ณ เวลาที่ h หรือก่อนหน้า h

นอกจากนี้หากเราต้องพยากรณ์แบบช่วง เราสามารถแสดงช่วงความเชื่อมั่น $(1-\alpha)$% ของการพยากรณ์ได้ด้วย

$$\widehat{Y_i}, h(l) \pm z_{\alpha/2}\sqrt{Var(e_i(l))}$$

โดยที่ z คือค่าวิกฤตจากการแจกแจงแบบปกติ และ $e_i(l)$ คือค่าคลาดเคลื่อนการพยากรณ์ของตัวแปร i

ในกรณีของแบบจำลองเวคเตอร์ออโตรีเกรสซีฟในรูปแบบโครงสร้าง (structural form VAR) เราสามารถสร้างค่าพยากรณ์ได้ด้วยวิธีการเดียวกัน ยกตัวอย่างเช่น แบบจำลองเวคเตอร์ออโตรีเกรสซีฟในรูปแบบโครงสร้างอันดับหนึ่งที่มีตัวแปรสองตัวแปร และอธิบายความสัมพันธ์ด้วยสมการดังต่อไปนี้

$$\begin{bmatrix} 1 & 0 \\ l_{21} & 1 \end{bmatrix} \begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \phi^*_{11} & \phi^*_{12} \\ \phi^*_{21} & \phi^*_{22} \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix} \tag{5.18}$$

ค่าพยากรณ์ Y_{1t} ไปข้างหน้าหนึ่งคาบเท่ากับ

$$E_h(Y_{1,h+1}) = \phi_{11}^* Y_{1,h} + \phi_{12}^* Y_{2,h}$$

และค่าคลาดเคลื่อนของการพยากรณ์ Y_{1t} ไปข้างหน้าหนึ่งคาบ, $e_1(1)$, เท่ากับ

$$e_1(1) = Y_{1,h+1} - E_h(Y_{1,h+1}) = \eta_{1,h+1}$$

ในขณะที่ค่าพยากรณ์ Y_{2t} ไปข้างหน้าหนึ่งคาบเท่ากับ

$$E_h(Y_{2,h+1}) = -l_{21} E_h(Y_{1,h+1}) + \phi_{21}^* Y_{1,h} + \phi_{22}^* Y_{2,h}$$

และค่าคลาดเคลื่อนของการพยากรณ์ Y_{2t} ไปข้างหน้าหนึ่งคาบ, $e_2(1)$,เท่ากับ

$$\begin{aligned} e_2(1) = Y_{2,h+1} - E_h(Y_{2,h+1}) &= -l_{21}(Y_{1,h+1} - E_h(Y_{1,h+1})) + \eta_{2,h+1} \\ &= -l_{21}\eta_{1,h+1} + \eta_{2,h+1} \end{aligned} \tag{5.19}$$

ตัวอย่างที่ 5.4: การพยากรณ์จากแบบจำลอง VAR

เราสามารถทำนายผลได้ตอบแทนไปข้างหน้าของทั้งสองตลาดที่เราสนใจได้ด้วยคำสั่ง `predict` โดยระบุแบบจำลองที่ใช้และคาบที่ทำนายไปข้างหน้า

```
> model1.predict=predict(model1, n.ahead=5)
> model1.predict
$lret_set
            fcst     lower    upper       CI
[1,]  0.23354446 -5.025286 5.492375 5.258830
[2,] -0.01504318 -5.311043 5.280957 5.296000
[3,]  0.08053962 -5.572718 5.733797 5.653258
[4,] -0.02118987 -5.682718 5.640339 5.661528
[5,] -0.04475454 -5.707153 5.617644 5.662399

$lret_dowjones
             fcst     lower    upper       CI
[1,] -0.046835523 -2.705029 2.611358 2.658193
[2,]  0.023569365 -2.639844 2.686983 2.663413
[3,]  0.005536266 -2.686514 2.697586 2.692050
[4,]  0.021958580 -2.670142 2.714059 2.692101
[5,]  0.031192947 -2.661288 2.723674 2.692481
```

เราจะได้ผลการพยากรณ์ดังแสดงได้ด้วยตารางดังต่อไปนี้

คาบไปข้างหน้า	1	2	3	4	5
log return ของ SET	0.233	-0.015	0.081	-0.021	-0.044
log return ของ Dow Jones	-0.467	0.024	0.006	0.022	0.031

5.5 การวิเคราะห์หลังจากการประมาณค่าแบบจำลองเวคเตอร์ออโตรีเกรสซีฟ

แบบจำลองเวคเตอร์ออโตรีเกรสซีฟโดยทั่วไปแล้วจะมีพารามิเตอร์จำนวนมากและอาจจะมีความยุ่งยากในการพิจารณาเนื่องจากความขึ้นอยู่ต่อกันอันจะส่งผลไปมาระหว่างตัวแปร ดังนั้นในการอธิบายความสัมพันธ์ระหว่างตัวแปรจะต้องมีการวิเคราะห์เพิ่มเติมไม่สามารถดูความสัมพันธ์ได้จากแบบจำลองเลยเช่นเดียวกับแบบจำลองอนุกรมเวลาอย่างง่าย

5.5.1 Granger Causality

โครงสร้างของแบบจำลอง VAR ช่วยให้ข้อมูลกับเราว่าตัวแปรหรือกลุ่มของตัวแปรมีความสามารถในการทำนายตัวแปรอื่นหรือไม่ โดยแนวคิดเกี่ยวกับความสามารถในการทำนายถูกนำเสนอโดย Granger (1969) ว่า ถ้าตัวแปร y_1 ช่วยในการทำนายตัวแปรอื่น y_2 เราจะเรียกว่า y_1 แกรงเจอรคอส (Granger-cause) y_2 แต่ถ้า y_1 ไม่ช่วยในการทำนายตัวแปร y_2 เราจะพูดว่า y_1 ไม่แกรงเจอรคอส (does not Granger-cause) y_2

นิยามอย่างเป็นทางการคือ หาก y_1 ไม่แกรงเจอรคอส (does not Granger-cause) y_2 แสดงว่าทุกค่าของ $s > 0$ ค่า MSE ของการพยากรณ์ $y_{2,t+s}$ ด้วยข้อมูลจาก $(y_{2,t}, y_{2,t-1}, ...)$ ไม่มีความแตกต่างจาก ค่า MSE ของการพยากรณ์ $y_{2,t+s}$ ด้วยข้อมูลจาก $(y_{2,t}, y_{2,t-1}, ...)$ และ $(y_{1,t}, y_{1,t-1}, ...)$ เราจะต้องระลึกไว้อยู่เสมอว่าแกรงเจอรคอสเป็นการอนุมานไปยังความสามารถในการพยากรณ์เท่านั้น ไม่ได้เป็นความเป็นเหตุเป็นผลจริงๆของตัวแปร

ตัวอย่างกรณีตัวแปรสองตัว

ในหัวข้อนี้เราพิจารณาแบบจำลอง $VAR(2)$ ที่มีตัวแปรสองตัว ซึ่งสามารถเขียนเป็นสมการได้ดังนี้

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \phi_{11}^1 & \phi_{12}^1 \\ \phi_{21}^1 & \phi_{22}^1 \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \phi_{11}^2 & \phi_{12}^2 \\ \phi_{21}^2 & \phi_{22}^2 \end{pmatrix} \begin{pmatrix} y_{1t-2} \\ y_{2t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{5.20}$$

ในสมการที่ (**??**) หาก y_{2t} ไม่แกรงเจอรคอส y_{1t} ค่าสัมประสิทธิ์ ϕ_{12}^1 และ ϕ_{12}^2 จะมีค่าเท่ากับศูนย์ ในทางตรงข้าม ถ้า y_{1t} ไม่แกรงเจอรคอส y_{2t} ค่าสัมประสิทธิ์ ϕ_{21}^1 และ ϕ_{21}^2 จะมีค่าเท่ากับศูนย์

ดังนั้นในการทดสอบว่า y_{2t} แกรงเจอร์คอส y_{1t} หรือไม่ เราจะตั้งสมมุติฐานหลักว่า H_0 : $\phi_{12}^1 = \phi_{12}^2 = 0$ และสมมุติฐานทางเลือกคือมีสัมประสิทธิ์ ϕ_{12}^1 หรือ ϕ_{12}^2 ตัวใดตัวหนึ่งมีค่าไม่เท่ากับศูนย์ ซึ่งการทดสอบดังกล่าวคือการทดสอบหลายเงื่อนไขหรือการทดสอบเอฟ(F-test) นั่นเอง

ตัวอย่าง 5.5 การทดสอบ Granger Causality จากแบบจำลอง VAR

ต่อเนื่องจากตัวอย่าง 5.3 เราต้องการทดสอบว่า `lret_dowjones` แกรงเจอร์คอส `lret_set` หรือไม่ (สมมุติฐานหลักคือ `lret_dowjones` ไม่แกรงเจอร์คอส `lret_set`) เราสามารถทดสอบโดยใช้คำสั่ง `causality` โดยที่เราจะต้องระบุแบบจำลองที่ใช้และตัวแปรที่เป็นสาเหตุ (cause) ซึ่งในที่นี้คือ `lret_dowjones`

```
1  > causality(model1, cause="lret_dowjones")
2  $Granger
3    Granger causality H0: lret_dowjones do not Granger-cause lret_set
4  data:  VAR object model1
5  F-Test = 43.5947, df1 = 4, df2 = 7588, p-value < 2.2e-16
6  
7  $Instant
8    H0: No instantaneous causality between: lret_dowjones and lret_set
9  data:  VAR object model1
10 Chi-squared = 86.7677, df = 1, p-value < 2.2e-16
```

จากบรรทัดที่ 2-5 เป็นผลการทดสอบแกรงเจอร์คอส โดยที่ค่าสถิติ F ที่ได้คือ 43 หรือมีค่าพีเท่ากับ $< 2.2 \times 10^{-16}$ ดังนั้นเราสามารถปฏิเสธสมมุติฐานหลักได้ด้วยระดับนัยสำคัญ 5 % (หรือค่าใดๆที่มากกว่า 2.2×10^{-16}) เราสรุปได้ว่า `lret_dowjones` แกรงเจอร์คอส `lret_set` หรือการทราบผลตอบแทนของตลาดหุ้นดาวโจนส์ช่วยในการพยากรณ์ผลตอบแทนในตลาดหุ้นไทย

นอกจาก นี้ บรรทัด ที่ 7-10 ได้ ทดสอบ ความ สัมพันธ์ ใน ระหว่าง ช่วง เวลา เดียวกัน (instantaneous causality)

5.5.2 ฟังก์ชันตอบสนองแรงกระตุ้น

การทดสอบสมมุติฐานสำหรับกลุ่มสัมประสิทธิ์(การทดสอบด้วยตัวสถิติ F)หรือการทดสอบแกรงเจอร์คอสช่วยระบุว่าตัวแปรใดๆมีผลกระทบอย่างมีนัยสำคัญต่อตัวแปรอื่นหรือไม่ อย่างไรก็ตาม การทดสอบดังกล่าวไม่ได้บอกทิศทางของความสัมพันธ์และระยะเวลาที่การเปลี่ยนแปลงของตัวแปรหนึ่งจะส่งผลต่ออีกตัวแปรว่าจะมีผลยาวนานแค่ไหน โดยที่เครื่องมือที่จะใช้ศึกษาความสัมพันธ์ดังกล่าวคือฟังก์ชันตอบสนองแรงกระตุ้น (impulse response function) โดยที่นิยามคือ การตอบสนองของตัวแปรตามอันเนื่องมาจากการเปลี่ยนแปลงของตัวแปรอีกตัวหนึ่ง โดยที่หากเกิดการเปลี่ยนแปลงในช็อกหนึ่งหน่วยจะส่งผลอย่างไรต่อแต่ละตัวแปรในคาบที่ต่อจากการเปลี่ยนแปลงนั้นๆ(เนื่องจากบางกรณีขนาดของตัวแปรที่เราสนใจอาจจะมีหน่วยที่ต่างกันมาก เราอาจจะใช้การเปลี่ยนแปลงหนึ่ง s.d.)

ตัวอย่างเช่น หากเราพิจารณาแบบจำลอง $VAR(1)$ ที่มีตัวแปรสองตัวแปร ซึ่งสามารถแสดงได้ด้วยสมการ

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{5.21}$$

สมมุติว่าเราพิจารณาว่าในปีที่ 0 เกิดช็อกในตัวแปร y_{1t} หนึ่งหน่วย ($\varepsilon_{10} = 1$) และช็อกในตัวแปร

y_{2t} เท่ากับศูนย์ และช็อกในปีอื่นๆเท่ากับศูนย์ สมมุติให้ y_{1t} และ y_{2t} เท่ากับศูนย์ในปีก่อนหน้าปี 0 จะได้ว่า

$$\begin{aligned}
\boldsymbol{Y}_0 &= \begin{pmatrix} \varepsilon_{10} \\ \varepsilon_{20} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
\boldsymbol{Y}_1 &= \begin{pmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{pmatrix} \boldsymbol{Y}_0 = \begin{pmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0 \end{pmatrix} \\
\boldsymbol{Y}_2 &= \begin{pmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{pmatrix} \boldsymbol{Y}_1 = \begin{pmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{pmatrix} \begin{pmatrix} 0.25 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.25 \\ 0 \end{pmatrix} \\
&\vdots
\end{aligned}$$

หรือในหากเราพิจารณา เกิดช็อกในตัวแปร y_{2t} หนึ่งหน่วย ($\varepsilon_{20} = 1$) และช็อกในตัวแปร y_{1t} เท่ากับศูนย์ และช็อกในปีอื่นๆเท่ากับศูนย์ สมมุติให้ y_{1t} และ y_{2t} เท่ากับศูนย์ในปีก่อนหน้าปี 0 จะได้ว่า

$$\begin{aligned}
\boldsymbol{Y}_0 &= \begin{pmatrix} \varepsilon_{10} \\ \varepsilon_{20} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
\boldsymbol{Y}_1 &= \begin{pmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{pmatrix} \boldsymbol{Y}_0 = \begin{pmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.3 \\ 0.2 \end{pmatrix} \\
&\vdots
\end{aligned}$$

ฟังก์ชันตอบสนองแรงกระตุ้นในกรณีทั่วไป

ในกรณี $VAR(p)$ ที่เป็นกระบวนการนิ่งใดๆ จะสามารถแสดงในรูป Wold ดังนี้

$$\boldsymbol{Y}_t = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t + \boldsymbol{\Psi}_1 \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\Psi}_2 \boldsymbol{\varepsilon}_{t-2} + \dots \tag{5.22}$$

โดยที่ $\boldsymbol{\Psi}_i$ คือ $n \times n$ เมทริกซ์ของสัมประสิทธิ์ moving average โดยที่ค่าสัมประสิทธิ์ดังกล่าวสามารถคำนวณได้จาก

$$(\boldsymbol{I} - \boldsymbol{\Phi}_1 L - \dots - \boldsymbol{\Phi}_p L^p)(\boldsymbol{I} + \boldsymbol{\Psi}_1 L + \boldsymbol{\Psi}_2 L^2 + \dots) = \boldsymbol{I}$$

จากสมการ (**??**) เราอาจต้องการที่จะอธิบาย ϕ_{ij}^s ว่าเป็นการเปลี่ยนแปลงของ

$$\frac{\partial y_{i,t+s}}{\partial \varepsilon_{j,t}} = \frac{\partial y_{i,t}}{\partial \varepsilon_{j,t-s}} = \psi_{i,j}^s$$

อย่างไรก็ตามคำอธิบายข้างต้นจะเป็นจริงหากเมตริกซ์ความแปรปรวนและความแปรปรวนร่วม $var(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}$ เป็นเมทริกซ์แทยงมุม หรือไม่มีความสัมพันธ์ $\boldsymbol{\varepsilon}_t$ ในคาบเดียวกัน

Sims (1980) เสนอให้ประมาณค่าด้วยแบบจำลอง triangular structural $VAR(p)$

$$
\begin{aligned}
y_{1t} &= c_1 + \boldsymbol{\gamma}_{11}^{'}\boldsymbol{Y}_{t-1} + ... + \boldsymbol{\gamma}_{1p}^{'}\boldsymbol{Y}_{t-p} + \eta_{1t} \\
y_{2t} &= c_2 + \beta_{21}y_{1t} + \boldsymbol{\gamma}_{21}^{'}\boldsymbol{Y}_{t-1} + ... + \boldsymbol{\gamma}_{2p}^{'}\boldsymbol{Y}_{t-p} + \eta_{2t} \\
y_{3t} &= c_3 + \beta_{31}y_{1t} + \beta_{32}y_{2t} + \boldsymbol{\gamma}_{31}^{'}\boldsymbol{Y}_{t-1} + ... + \boldsymbol{\gamma}_{3p}^{'}\boldsymbol{Y}_{t-p} + \eta_{3t} \\
&\vdots \\
y_{nt} &= c_n + \beta_{n1}y_{nt} + ... + \beta_{n,n-1}y_{n-1,t} + \boldsymbol{\gamma}_{n1}^{'}\boldsymbol{Y}_{t-1} + ... + \boldsymbol{\gamma}_{np}^{'}\boldsymbol{Y}_{t-p} + \eta_{nt}
\end{aligned} \tag{5.23}
$$

เราสามารถเขียนความสัมพันธ์ดังกล่าวในรูปของเมทริกซ์ดังนี้

$$\boldsymbol{B}\boldsymbol{Y}_t = \boldsymbol{c} + \boldsymbol{\Gamma}_1\boldsymbol{Y}_{t-1} + ... + \boldsymbol{\Gamma}_p\boldsymbol{Y}_{t-p} + \boldsymbol{\eta}_t \tag{5.24}$$

โดยที่

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -\beta_{21} & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\beta_{n1} & -\beta_{n2} & -\beta_{n3} & \dots & 1 \end{pmatrix}$$

โดยที่ Orthogonal error ($\boldsymbol{\eta}_t$) หรือ error ที่ไม่มีความสัมพันธ์ระหว่างกัน ซึ่งก็คือ structural errors

โครงสร้างแบบจำลองโครงสร้าง (structural model)อย่างในสมการ (**??**) จะกำหนดให้ความสัมพันธ์ของตัวแปรมีลำดับคือ y_1 เป็นตัวกำหด y_2 และ y_2 เป็นตัวกำหนด y_3 เป็นลำดับไปเรื่อยๆ

$$y_1 \longrightarrow y_2 \longrightarrow y_3 \longrightarrow ... \longrightarrow y_n$$

แสดงว่าค่าของตัวแปรด้านซ้ายจะกระทบตัวแปรด้านขวาในคาบเดียวกัน (contemporaneous) แต่ตัวแปรด้านขวาจะต้องใช้เวลาหนึ่งคาบถึงจะกระทบกลับมายังตัวแปรด้านซ้าย โดยที่ผลกระทบในคาบเดียวกันจะสะท้อนผ่านตัวแปร β_{ij} โดยที่ลำดับของการเรียงตัวแปรในแบบจำลอง $VAR(p)$ ตามสมการที่ (**??**)จะขึ้นอยู่กับทฤษฎีเป็นตัวกำหนด

หลังจากที่เราได้จัดเรียงตัวแปรแล้ว Wold representation ของ $\boldsymbol{Y}_t$ จากค่าช็อกเชิงตั้งฉาก (orthogonal errors) ได้เป็น

$$\boldsymbol{Y}_t = \boldsymbol{\mu} + \boldsymbol{\Theta}_0\boldsymbol{\eta}_t + \boldsymbol{\Theta}_1\boldsymbol{\eta}_{t-1} + \boldsymbol{\Theta}_2\boldsymbol{\eta}_{t-2} + ... \tag{5.25}$$

โดยที่ $\boldsymbol{\Theta}_0 = \boldsymbol{B}^{-1}$ และการตอบสนองแรงกระตุ้นจากช็อกเชิงตั้งฉาก η_{jt} จะเท่ากับ

$$\frac{\partial y_{i,t+s}}{\partial \eta_{j,t}} = \frac{\partial y_{i,t}}{\partial \eta_{j,t-s}} = \theta_{ij}^s, \qquad i,j = 1,...,n; s > 0 \tag{5.26}$$

โดยที่ θ_{ij}^s คือ ตำแหน่ง (i,j) ของเมทริกซ์ $\boldsymbol{\Theta}_s$ เราเรียกแผนภาพที่วาด θ_{ij}^s ในแกนตั้งและ s ในแกน

นอนว่า ฟังก์ชันตอบสนองแรงกระตุ้นเชิงตั้งฉาก (orthogonal impulse response function) ของ y_i จากช็อก η_j

ในทางปฏิบัติเราสามารถคำนวณ ฟังก์ชันตอบสนองแรงกระตุ้นเชิงตั้งฉากจากผลการประมาณค่า reduced VAR ในสมการ **??** ได้ด้วยการแยกส่วนประกอบ(decompose)ของ เมทริกซ์ค่าความแปรปรวนร่วมของ residuals $\boldsymbol{\Sigma}$ ด้วย Cholesky decomposition ดังนี้ $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{G}\boldsymbol{L}'$ และ $\boldsymbol{\eta}_t = \boldsymbol{L}^{-1}\boldsymbol{\varepsilon}_t$ จากความสัมพันธ์ดังกล่าวเราสามารถเขียน Wold representation ในสมการ **??** ใหม่ได้เป็น

$$\begin{aligned} \boldsymbol{Y}_t &= \boldsymbol{\mu} + \boldsymbol{L}\boldsymbol{L}^{-1}\boldsymbol{\varepsilon}_t + \boldsymbol{\Psi}_1\boldsymbol{L}\boldsymbol{L}^{-1}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\Psi}_2\boldsymbol{L}\boldsymbol{L}^{-1}\boldsymbol{Y}_{t-2} + \dots \\ &= \boldsymbol{\mu} + \boldsymbol{\Theta}_0\boldsymbol{\eta}_t + \boldsymbol{\Theta}_1\boldsymbol{\eta}_{t-1} + \boldsymbol{\Theta}_2\boldsymbol{\eta}_{t-2} + \dots \end{aligned} \tag{5.27}$$

โดยที่ $\boldsymbol{\Theta}_j = \boldsymbol{\Psi}_j\boldsymbol{L}$ โดยที่เมทริกซ์ $\boldsymbol{B}$ ในสมการ **??** คือ $\boldsymbol{L}^{-1}$ นั่นเอง

ตัวอย่างที่ 5.6: การสร้าง Impulse Response Function จากแบบจำลอง VAR

ต่อเนื่องจากตัวอย่าง โดยเราจะสร้าง Impulse response function ด้วยคำสั่ง `irf` โดยที่ default ในคำสั่งนี้คือการสร้าง orthogonal impulse response function (`ortho=TRUE`) โดยในคำสั่งนี้เราจะระบุแบบจำลองที่ใช้และคาบที่เราดูผลหลังจากที่เรา shock 1 s.d. โดยที่การเรียงตัวแปรใน `lret` ที่เรียง `lret_set` ก่อน `lret_dowjones` หมายความว่าเราสมมุติให้ structural shock ของ `lret_dowjones` ไม่มีผลทันที(contemporaneous)ต่อ `lret_set` แต่ structural shock ของ `lret_set` มีผลทันที(contemporaneous)ต่อ `lret_dowjones` [ข้อสมมุตินี้อาจจะไม่สมเหตุสมผลนัก ผู้อ่านอาจจะลองสร้างเมทริกซ์ใหม่ `lret2 <-cbind(lret_dowjones, lret_set)` แล้วประมาณค่าและหา IRF ใหม่]

```
1  > model1.irf=irf(model1, n.ahead=10)
2  > model1.irf
3  Impulse response coefficients
4  $lret_set
5              lret_set lret_dowjones
6   [1,]  2.6831259181 -0.2072364167
7   [2,]  0.1137971489  0.0030468719
8   [3,] -0.9182763350  0.1667746830
9  <omitted>
10 [11,]  0.0030760182 -0.0002311969
11 $lret_dowjones
12             lret_set lret_dowjones
13  [1,]  0.0000000000  1.3403194745
14  [2,]  0.2986284632 -0.0849824924
15  [3,]  0.4183950540 -0.1100409438
16 <omitted>
17 [11,]  0.0006173344 -0.0002757296
18
19
20 Lower Band, CI= 0.95
21 <omitted>
22
23 > plot(model1.irf)
```

บรรทัดที่ 4-10 เป็นผลจากการเปลี่ยนแปลงใน structural shock ของ `lret_set` หนึ่ง s.d. (=2.683) ในคาบที่ศูนย์ (ในที่นี้แทนด้วยบรรทัดที่ [,1]) ทำให้ `lret_dowjones` เปลี่ยนแปลงทันที -0.207 และส่งผลต่อ `lret_set` และ `lret_dowjones` ในคาบที่ 1 (หลังจาก shock) เท่ากับ 0.114 และ 0.003 ตามลำดับ (ในที่นี้แทนด้วยบรรทัดที่ [,2])

ส่วนบรรทัดที่ 11-16 เป็นผลจากการเปลี่ยนแปลงใน structural shock ของ `lret_dowjones` หนึ่ง s.d.(=1.34) ในคาบที่ศูนย์ (ในที่นี้แทนด้วยบรรทัดที่ [,1]) ไม่ทำให้ `lret_set` เปลี่ยนแปลงในคาบที่เกิดช็อก แต่ส่งผลในคาบต่อๆไป

นอกจากนี้คำสั่งดังกล่าวยังให้ช่วงความเชื่อมั่น 95% และเราสามารถสร้างกราฟ IRF ได้ด้วยคำสั่ง `plot(model1.irf)`

5.5.3 การแยกความแปรปรวนของค่าคลาดเคลื่อนการพยากรณ์

จากหัวข้อการพยากรณ์จากแบบจำลองเวคเตอร์ออโตรีเกรสซีฟ เราจะเห็นได้ว่าค่าพยากรณ์ของตัวแปรแต่ละตัวแปรนั้นขึ้นอยู่กับทั้งตัวแปรนั้นๆ และตัวแปรอื่นๆในแบบจำลอง ดังนั้น แนวทางอีกประการในการอธิบายผลจากแบบจำลองเวคเตอร์ออโตรีเกรสซีฟ คือการแยกส่วนประกอบของค่าคาดเคลื่อนของการพยากรณ์ (Forecast Error Variance Decomposition) ว่ามีที่มาจากตัวแปรแต่ละตัวมากน้อยแค่ไหน

ยกตัวอย่าง แบบจำลอง $VAR(1)$ ที่มีตัวแปรสองตัว เราพบว่าการพยากรณ์ Y_{1t} ไปข้างหน้าหนึ่งคาบมีค่าคาดเคลื่อนจากการพยากรณ์ (forecast error) ตามสมการที่ (**??**) เท่ากับ

$$e_1(1) = \varepsilon_{1,h+1}$$

ซึ่งค่าคาดเคลื่อนเกิดจากช๊อคในตัวแปร Y_{1t} เพียงอย่างเดียว และมีค่าความแปรปรวนของค่าความคาดเคลื่อนจากการพยากรณ์ไปข้างหน้าหนึ่งคาบ (one-period forecast error variance) เท่ากับ

$$Var(e_1(1)) = Var(\varepsilon_{1,h+1}) = \sigma_1^2$$

ในขณะที่การพยากรณ์ Y_{1t} ไปข้างหน้าสองคาบมีค่าคาดเคลื่อนจากการพยากรณ์ตามสมการที่ **??** เท่ากับ

$$e_1(2) = \phi_{11}\varepsilon_{1,h+1} + \phi_{12}\varepsilon_{2,h+1} + \varepsilon_{1,h+2}$$

ซึ่งค่าคาดเคลื่อนเกิดจากช๊อคของตัวแปร Y_{1t} และตัวแปร Y_{2t} เราสามารถคำนวณหาค่าความแปรปรวนของค่าความคลาดเคลื่อนจากการพยากรณ์ไปข้างหน้าสองคาบได้ดังสมการต่อไปนี้

$$\begin{aligned} Var(e_1(2)) &= Var(\phi_{11}\varepsilon_{1,h+1}) + Var(\phi_{12}\varepsilon_{2,h+1}) + Var(\varepsilon_{1,h+2}) \\ &= \phi_{11}^2\sigma_1^2 + \phi_{12}^2\sigma_2^2 + \sigma_1^2 && (5.28) \\ &= (1+\phi_{11}^2)\sigma_1^2 + \phi_{12}^2\sigma_2^2 && (5.29) \end{aligned}$$

และสามารถแยกส่วนประกอบของค่าความแปรปรวนของค่าความคลาดเคลื่อนจากการพยากรณ์ไปข้างหน้าสองคาบว่ามาจาก Y_{1t} เท่ากับ

$$\frac{(1+\phi_{11}^2)\sigma_1^2}{(1+\phi_{11}^2)\sigma_1^2 + \phi_{12}^2\sigma_2^2}$$

และมาจาก Y_{2t} เท่ากับ

$$\frac{\sigma_2^2}{(1+\phi_{11}^2)\sigma_1^2 + \phi_{12}^2\sigma_2^2}$$

เราสามารถอธิบายได้ว่าตัวแปรแต่ละตัวมีอิทธิพลต่อค่าความแปรปรวนการพยากรณ์ตัวแปรที่เราสนใจมากน้อยแค่ไหน

ตัวอย่างที่ 5.7 การหา Variance Decomposition

ต่อเนื่องจากตัวอย่าง 5.3 เราสามารถแยกความแปรปรวนของค่าผิดพลาดจากการพยากรณ์ได้ด้วยคำสั่ง `fevd`

```
1  > fevd(model1, n.ahead = 5)
2  $lret_set
3        lret_set lret_dowjones
4  [1,] 1.0000000    0.00000000
5  [2,] 0.9877859    0.01221413
6  [3,] 0.9682396    0.03176043
7  [4,] 0.9655329    0.03446710
8  [5,] 0.9654222    0.03457779
9
10 $lret_dowjones
11         lret_set lret_dowjones
12 [1,] 0.02334830     0.9766517
13 [2,] 0.02326189     0.9767381
14 [3,] 0.03751278     0.9624872
15 [4,] 0.03751192     0.9624881
16 [5,] 0.03750568     0.9624943
```

บรรทัด ที่ 4-8 แสดง การ แยก ความ แปรปรวน ของ ค่า คลาด เคลื่อน ของ การ พยากรณ์ `lret_set` ไป 1 ถึง 5 คาบ ยกตัวอย่างเช่น การพยากรณ์ไปข้างหน้า 5 คาบ ความแปรปรวนของค่าพยากรณ์เกิดจาก `lret_set` เอง 96.56 % และเกิดจาก `lret_downjones` 3.46 %

บทที่ 6

โคอินทิเกรชัน (Cointegration)

ข้อมูลที่เหมาะสมกับการวิเคราะห์ถดถอยเชิงเส้นสำหรับอนุกรมเวลาและแบบจำลองเวกเตอร์ออโตรีเกรสซีฟจะต้องเป็นข้อมูลที่นิ่ง ($I(0)$) เช่นอัตราผลได้ตอบแทนของสินทรัพย์หรืออัตราการเติบโตของตัวแปรเศรษฐกิจมหภาค อย่างไรก็ตาม ในการพิจารณาความสัมพันธ์ตามทฤษฎีเศรษฐศาสตร์บางครั้งเราจำเป็นต้องพิจารณาข้อมูลที่มิได้มีการปรับรูปแบบซึ่งมักจะเป็น $I(1)$

6.1 สมการความสัมพันธ์เทียม (spurious regression)

ในการวิเคราะห์ถดถอยสำหรับข้อมูลอนุกรมเวลา ตัวแปรที่เราใช้ในการวิเคราะห์จำเป็นจะต้องเป็น $I(0)$ เพื่อที่จะทำให้การทดสอบเชิงสถิติมีความเหมาะสม หากตัวแปรในแบบจำลองบางตัวหรือทั้งหมดเป็น $I(1)$ การทดสอบสถิติอาจจะไม่ถูกต้อง ตัวอย่างเช่นในกรณีที่ตัวแปรทั้งหมดเป็น $I(1)$ และไม่ได้โคอินทิเกรตกัน ตัวทดสอบที่ได้จากแบบจำลองจะไม่เหมาะสม เราจะพบความสัมพันธ์ระหว่างตัวแปรทั้งๆที่จริงๆแล้วตัวแปรทั้งหลายไม่มีความสัมพันธ์กัน หรือเราเรียกสมการความสัมพันธ์เทียม (spurious regression)

ตัวอย่างที่ 6.1

หากเราพิจารณาอนุกรมเวลาสองกระบวนการที่เป็น $I(1)$ และเป็นอิสระต่อกัน เราสามารถแสดงความสัมพันธ์ของแต่ละตัวแปรได้ดังนี้

$$y_{1t} = y_{1,t-1} + \varepsilon_{1t} \qquad \varepsilon_{1t} \sim WN(0,1)$$

$$y_{2t} = y_{2,t-1} + \varepsilon_{2t} \qquad \varepsilon_{2t} \sim WN(0,1)$$

โดยที่ $Cov(\varepsilon_{1t}, \varepsilon_{2t}) = 0$ ซึ่งแสดงว่าทั้งสองอนุกรมเวลาเป็นอิสระต่อกัน เราสามารถใช้โปรแกรม R สร้างตัวแปรทั้งสองด้วยคำสั่งต่อไปนี้

```
> set.seed(123456)
> e1<-rnorm(1000)
> e2<-rnorm(1000)
> y1<-cumsum(e1)
> y2<-cumsum(e2)
```

แผนภาพของอนุกรมเวลาทั้งสองสามารถแสดงได้ดังรูปที่ **??**

เมื่อนำอนุกรมเวลาทั้งสองมาพิจารณาความสัมพันธ์เชิงถดถอย $y_{1t} = \alpha + \beta y_{2t} + u_t$ จะพบว่า $\hat{\beta}$ มีค่าเท่ากับ 0.16 และแตกต่างจากศูนย์อย่างมีนัยสำคัญ ($t_{\hat{\beta}} = 5.6$)

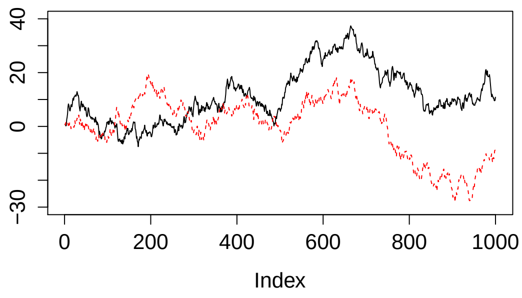
```
> summary(lm(y1~y2))
Call:
lm(formula = y1 ~ y2)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 11.43603    0.31329  36.503  < 2e-16 ***
y2           0.16237    0.02885   5.628 2.36e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '  1

Residual standard error: 9.905 on 998 degrees of freedom
Multiple R-squared: 0.03077,  Adjusted R-squared: 0.02979
F-statistic: 31.68 on 1 and 998 DF,  p-value: 2.363e-08
```

หากเราพิจารณาสมการ $\Delta y_{1t} = \beta \Delta y_{2t} + v_t$ จะพบว่า $\hat{\beta}$ มีค่าเท่ากับ 0.0034 และไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ

```
> summary(lm(diff(y1)~diff(y2)-1))

Call:
lm(formula = diff(y1) ~ diff(y2) - 1)

Coefficients:
         Estimate Std. Error t value Pr(>|t|)
diff(y2) 0.003455   0.030827   0.112    0.911

Residual standard error: 0.9921 on 998 degrees of freedom
Multiple R-squared: 1.258e-05,  Adjusted R-squared: -0.0009894
F-statistic: 0.01256 on 1 and 998 DF,  p-value: 0.9108
```

จากตัวอย่างข้างต้นเราจำเป็นต้องระมัดระวังในการพิจารณาความสัมพันธ์ระหว่างอนุกรมเวลาในกรณีที่ตัวแปรที่เราสนใจมีลักษณะเป็น $I(1)$ ซึ่งในกรณีทั่วไปเราจำเป็นต้องปรับให้ตัวแปรดังกล่าวเป็น $I(0)$ ก่อนที่จะพิจารณาความสัมพันธ์ด้วยสมการเช่น $\Delta y_t = \beta \Delta x_t + v_t$ โดยค่าสัมประสิทธิ์ β ก็จะอธิบายการเปลี่ยนแปลงของ Δx_t ต่อการเปลี่ยนแปลงของ Δy_t อย่างไรก็ตามในบางกรณีที่เราอาจจะสามารถศึกษาความสัมพันธ์ระหว่างตัวแปรที่เป็น $I(1)$ ก็ได้ซึ่งในกรณีดังกล่าวคือกรณีที่ตัวแปรเป็น cointegration

รูปที่ 6.1: การจำลองกระบวนการ $I(1)$ สองอนุกรมเวลา



6.2 นิยามของโคอินทิเกรชัน

กำหนดให้ $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, ..., y_{nt})'$ แทนเวกเตอร์ $(n \times 1)$ ของอนุกรมเวลาที่เป็น $I(1)$ แล้ว ตัวแปรใน $\boldsymbol{Y}_t$ จะโคอินทิเกรตกัน ถ้ามีเวกเตอร์ $\boldsymbol{\beta} = (\beta_1, ..., \beta_n)'$ ที่ทำให้

$$\boldsymbol{\beta}'\boldsymbol{Y}_t = \beta_1 y_{1t} + ... + \beta_n y_{nt} \sim I(0) \tag{6.1}$$

ซึ่งหมายความว่าผลรวมเชิงเส้น (linear combination) ของตัวแปรที่เป็น $I(1)$ มีลักษณะเป็น $I(0)$ โดยที่ผลรวมเชิงเส้นมักจะได้มาจากทฤษฎีทางเศรษฐศาสตร์ และเรียกสมการดังกล่าวว่าเป็นความสัมพันธ์ใน**ดุลยภาพระยะยาว (long-run equilibrium)** ซึ่งสะท้อนการที่อนุกรมเวลา $I(1)$ มีความสัมพันธ์ในดุลยภาพระยะยาวด้วยกันอยู่ดังนั้นตัวแปรดังกล่าวจะไม่แยกห่างจากกันที่ดุลยภาพ โดยจะมีแรงปรับให้ตัวแปรเข้าสู่ดุลยภาพ

เวกเตอร์ $\boldsymbol{\beta}$ ไม่ได้มีลักษณะเป็นรูปแบบอย่างเดียว (unique) เนื่องจากเวกเตอร์ที่เป็นจำนวนเท่าของ $\boldsymbol{c\beta}'$ ก็สามารถสอดคล้องกับเงื่อนไขในคำนิยามข้างต้นได้ ดังนั้น เรามักจะสนใจในรูปแบบที่ปรับเป็นบรรทัดฐาน (normalized)ซึ่งเขียนสมการโคอินทิเกรชันได้ดังนี้

$$\boldsymbol{\beta}'\boldsymbol{Y}_t = y_{1t} - \beta_2 y_{2t} - ... - \beta_n y_{nt} \sim I(0)$$

หรือเป็นสมการ

$$y_{1t} = \beta_2 y_{2t} + ... + \beta_n y_{nt} + u_t \tag{6.2}$$

โดยที่ $u_t \sim I(0)$ และเราเรียกพจน์ u_t ว่าค่าคลาดเคลื่อนจากดุลยภาพ (disequilibrium error) หรือ cointegrating residuals ซึ่งหากเราอยู่ในดุลยภาพระยะยาวค่าคลาดเคลื่อนจากดุลยภาพจะมีค่าเท่ากับศูนย์ เราจะได้ความสัมพันธ์ระยะยาวจะเท่ากับ $y_{1t} = \beta_2 y_{2t} + ... + \beta_n y_{nt}$

6.2.1 Cointegration และแบบจำลอง Error Correction Models

จากหัวข้อที่ผ่านมาเราพบว่าหากตัวแปรมีคุณลักษณะเป็น cointegration เราสามารถอธิบายได้ว่าในระยะยาวการเปลี่ยนแปลงของตัวแปรหนึ่งส่งผลอย่างไรต่อตัวแปรอื่นในดุลยภาพระยะยาว สิ่งที่เราสนใจต่อไปคือหากเกิดภาวะที่ในระยะสั้นตัวแปรขยับออกจากดุลยภาพในระยะยาวจะมีกลไกในการปรับตัวเข้าสู่ดุลยภาพระยะยาวอย่างไร ซึ่งสามารถอธิบายผ่านแบบจำลอง Error Correction Model (ECM)

หากเราพิจารณา cointegration ที่มีตัวแปรสองตัวแปรที่มีลักษณะ $I(1)$ คือ y_{1t} และ y_{2t} โดยมีความสัมพันธ์ cointegration เขียนได้เป็น $y_{1t} - \beta_2 y_{2t} \sim I(0)$ Engle and Granger (1987) แสดงให้เห็นว่ามีแบบจำลอง error correction model (ECM) ที่สามารถแสดงได้ในรูป

$$\begin{aligned}\Delta y_{1t} &= c_1 + \alpha_1(y_{1,t-1} - \beta_2 y_{2,t-1}) + \textstyle\sum_{j=1}^{p} \psi_{11}^{j} \Delta y_{1,t-j} + \sum_{j=1}^{p} \psi_{12}^{j} \Delta y_{2,t-j} + \varepsilon_{1t} \\ \Delta y_{2t} &= c_2 + \alpha_2(y_{1,t-1} - \beta_2 y_{2,t-1}) + \textstyle\sum_{j=1}^{p} \psi_{21}^{j} \Delta y_{1,t-j} + \sum_{j=1}^{p} \psi_{22}^{j} \Delta y_{2,t-j} + \varepsilon_{2t} \quad (6.3)\end{aligned}$$

สมการข้างต้นแสดงพฤติกรรมเชิงพลวัตรของ y_{1t} และ y_{2t} โดยแสดงการปรับตัวระยะสั้นเมื่อตัวแปรตัวใดตัวหนึ่งเคลื่อนตัวออกไปจากดุลยภาพ ซึ่งแบบจำลอง ECM ส่งผลให้เกิดประโยชน์ในการวิเคราะห์ข้อมูลการเงิน

ตัวอย่าง 6.2: ECM สำหรับราคาหุ้นและเงินปันผล

กำหนดให้ล็อกของราคาหุ้น (s_t) และล็อกของเงินปันผล (d_t) เป็น $I(1)$ ถ้าล็อกของสัดส่วนเงินปันผลต่อราคา ($\log(D_t/S_t)$) เป็น $I(0)$ ดังนั้น $d_t - s_t \sim I(0)$ หรือเขียนเป็นสมการความสัมพันธ์ระยะยาวได้เป็น

$$d_t = s_t + \mu + u_t \quad (6.4)$$

โดยที่ $u_t \sim I(0)$ และสามารถเขียน ECM ได้ดังนี้

$$\Delta s_t = c_s + \alpha_s(d_{t-1} - s_{t-1} - \mu) + \varepsilon_{st} \quad (6.5)$$

$$\Delta d_t = c_d + \alpha_d(d_{t-1} - s_{t-1} - \mu) + \varepsilon_{dt} \quad (6.6)$$

โดยที่สมการ **??** แสดงความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงของราคาหุ้นกับ disequilibrium error ($d_{t-1} - s_{t-1} - \mu$) ในขณะที่สมการ **??** แสดงความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงของเงินปันผลกับ disequilibrium error โดยมีการตอบสนองของราคาหุ้นและเงินปันผลขึ้นอยู่กับค่าสัมประสิทธิ์การปรับตัว (adjustment coefficient) α_s และ α_d

สมมุติให้ $\alpha_s = 0.5$ และ $\alpha_d = 0$ จะได้

$$\begin{aligned}\Delta s_t &= c_s + 0.5(d_{t-1} - s_{t-1} - \mu) + \varepsilon_{st} \\ \Delta d_t &= c_d + \varepsilon_{dt}\end{aligned}$$

ดังนั้นมีเพียงราคาหุ้นที่ตอบสนองต่อ disequilibrium error โดยที่ค่าเฉลี่ยของการเปลี่ยนแปลงของตัวแปรทั้งสองจะขึ้นอยู่กับข้อมูลในอดีต

$$
\begin{aligned}
E(\Delta s_t|F_{t-1}) &= c_s + 0.5(d_{t-1} - s_{t-1} - \mu) \\
E(\Delta d_t|F_{t-1}) &= c_d
\end{aligned}
$$

ซึ่งเราสามารถพิจารณาการเปลี่ยนแปลงได้ดังนี้

1. หาก $d_{t-1} - s_{t-1} - \mu = 0$ ดังนั้น $E(\Delta s_t|F_{t-1}) = c_s$ และ $E(\Delta d_t|F_{t-1}) = c_d$ หากไม่มีการเปลี่ยนแปลงออกจากดุลยภาพระยะยาวในคาบที่ผ่านมา อัตราการเปลี่ยนแปลงของทั้งสองตัวแปรจะเท่ากับค่าเฉลี่ยของการเปลี่ยนแปลง

2. หาก $d_{t-1} - s_{t-1} - \mu > 0$ ดังนั้น $E(\Delta s_t|F_{t-1}) = c_s + 0.5(d_{t-1} - s_{t-1} - \mu) > c_s$ เนื่องจากเงินปันผลสูงกว่าราคาหุ้นที่จะรักษาดุลยภาพระยะยาว ดังนั้น ECM พยากรณ์ว่า s_t จะเปลี่ยนแปลงที่อัตราสูงกว่าอัตราที่ดุลยภาพในระยะยาว ($> c_s$) เพื่อรักษาสัดส่วนราคาหุ้นต่อเงินปันผล โดยที่ความเร็วในการปรับตัว (speed of adjustment) แสดงได้ด้วยพารามิเตอร์ $\alpha_s = 0.5$

6.2.2 multiple cointegrating relationship

เวกเตอร์ของอนุกรมเวลา $\boldsymbol{Y}_t = (y_{1t}, ..., y_{nt})'$ ซึ่งเป็นอนุกรมเวลาที่ cointegrated กัน อาจจะมีความสัมพันธ์ระยะยาวได้ r รูปแบบ ซึ่ง $o < r < n$ เช่นกรณีที่ $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, y_{3t})'$ อาจจะมีความสัมพันธ์ในรูป $\beta_{11}y_{1t} + \beta_{12}y_{2t} + \beta_{13}y_{3t} \sim I(0)$ และ $\beta_{21}y_{1t} + \beta_{22}y_{2t} + \beta_{23}y_{3t} \sim I(0)$ และเราเรียกเวกเตอร์ $(\beta_{11} \quad \beta_{12} \quad \beta_{13})'$ และ $(\beta_{21} \quad \beta_{22} \quad \beta_{23})'$ ว่า cointegrating vector

6.3 การทดสอบ cointegration

หากเรากำหนดให้ $\boldsymbol{Y}_t$ เป็นเวกเตอร์ $n \times 1$ ของอนุกรมเวลา $I(1)$ ซึ่งมี r $(0 < r < n)$ cointegrating vectors ซึ่งแสดงเป็นสัญลักษณ์ได้ดังนี้

$$
\begin{pmatrix} \boldsymbol{\beta}_1'\boldsymbol{Y}_t \\ \vdots \\ \boldsymbol{\beta}_r'\boldsymbol{Y}_t \end{pmatrix} = \begin{pmatrix} u_{1t} \\ \vdots \\ u_{rt} \end{pmatrix} \sim I(0) \tag{6.7}
$$

การทดสอบ cointegration อาจจะแบ่งได้เป็นสองกรณีคือ

1. การทดสอบว่ามีความสัมพันธ์ cointegrating อย่างมากหนึ่งรูปแบบหรือไม่ (Engle and Granger(1986))

2. การทดสอบว่ามี cointegrating vector $0 \leq r < n$ หรือไม่ (Johansen (1988))

การทดสอบในประเด็นแรกได้ถูกนำเสนอโดย Engle and Granger (1986) ซึ่งเสนอการทดสอบสองขั้นตอน (two-step cointegration test)ซึ่งประกอบด้วย

1. สร้าง cointegrating residuals โดยที่ $u_t = \beta_1 y_{1t} + ... + \beta_n y_{nt}$
2. ทดสอบ unit root ของ u_t เพื่อแสดงว่า u_t เป็น $I(0)$ หรือไม่ โดยสมมุติฐานหลักของการทดสอบนี้คือไม่มี cointegration ซึ่งสมมุติฐานดังกล่าวสอดคล้องกับ u_t เป็น $I(1)$

การทดสอบ cointegration ตามวิธีการของ Engle and Granger สามารถแบ่งออกได้เป็นสองกรณี คือ *กรณีที่หนึ่ง* เราทราบความสัมพันธ์ของอนุกรมเวลาที่เราสนใจ เช่น จากทฤษฎีเราทราบความสัมพันธ์ระหว่างล็อกของราคาหุ้น (s_t) และล็อกของเงินปันผล (d_t) ว่าสามารถแสดงเป็นสมการ $s_t - d_t = u_t$ หรือ cointegrating vector สามารถแสดงได้โดย $\beta = (1 - 1)'$ ดังนั้นเราสามารถหา cointegrating residuals ได้ทันที *กรณีที่สอง* เราจำเป็นต้องประมาณค่าสำหรับ cointegrating vector และสร้าง cointegrating vector จาก $\hat{u}_t = \hat{\boldsymbol{\beta}}' \boldsymbol{Y}_t$

6.3.1 การทดสอบกรณีที่เราทราบค่า cointegrating vector

กำหนดให้ $\boldsymbol{Y}_t$ เป็นเวกเตอร์ $n \times 1$ ของ $I(1)$ และ cointegrated กันโดยที่มีความสัมพันธ์สามารถแสดงได้ด้วย cointegrating vector $\boldsymbol{\beta}$ ซึ่ง $u_t = \boldsymbol{\beta}'\boldsymbol{Y}_t$ เราต้องการทดสอบสมมุติฐานหลัก $H_0 : u_t = \boldsymbol{\beta}'\boldsymbol{Y}_t \sim I(1)$ หรือไม่มี cointegration กับสมมุติฐานรอง $H_1 : u_t = \boldsymbol{\beta}'\boldsymbol{Y}_t \sim I(0)$ หรือมี cointegration

เราสามารถทดสอบสมมุติฐานข้างต้นได้ด้วยการทดสอบยูนิทรูทเช่น ADF หรือ PP กับตัวแปร u_t ซึ่งเราจะสรุปว่า $\boldsymbol{Y}_t$ เป็น cointegrated ถ้าเราปฏิเสธสมมุติฐานหลัก

ตัวอย่างที่ 6.3: การพิจารณาความสัมพันธ์ระหว่างราคาปัจจุบันและราคาฟิวเจอร์

ในตัวอย่างนี้เราจะพิจารณาความสัมพันธ์ระหว่างราคาปัจจุบัน (spot price: S_t) และราคาฟิวเจอร์ (future price: F_t) ของอัตราแลกเปลี่ยนโดยที่ตัวแปรทั้งสองมีความสัมพันธ์ตามแบบจำลอง cost of carry model ซึ่งล็อกของราคาปัจจุบัน (s_t) จะเท่ากับล็อกของราคาฟิวเจอร์ (f_t) บวกกับค่าคงที่(ซึ่งสะท้อนต้นทุนในการถือสินทรัพย์)

$$f_t = s_t + c$$

อย่างไรก็ตาม ก่อนที่เราจะพิจารณาสมการ cointegration เราอาจจะพิจารณา order of integration ของตัวแปรทั้งสอง โดยการทดสอบ Augmented Dickey Fuller Unit root ด้วยคำสั่ง *ur.df* จาก package *urca*

กรณีของ lspot พบว่าค่าสถิติเท่ากับ -1.3151 มากกว่าค่า critical value (-2.88) แสดงว่า lspot เป็น unit root แล้วเมื่อพิจารณา diff(lspot) ค่าสถิติเท่ากับ -7.78 น้อยกว่า critical value แสดงว่า diff(lspot) เป็น stationary ดังนั้น lspot เป็น I(1)

เมื่อเราพิจารณา lfuture เราก็ได้ข้อสรุปเช่นเดียวกัน คือ lfuture เป็น I(1) ดังนั้น เราไม่สามารถประมาณค่า OLS กับตัวแปรทั้งสองได้ ยกเว้นในกรณี Cointegration

```
> set50_m <- read.csv("C:/teaching/ec435/r/set50_m.csv")
> lfutures<-log(set50_m$futures)
> lspot<-log(set50_m$spot)
> library(urca)
> summary(ur.df(lspot, type="drift", selectlags = "BIC"))
<Omitted>

Value of test-statistic is: -1.3151 1.15

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81

> summary(ur.df(diff(lspot), type="drift", selectlags = "BIC"))
<Omitted>
Value of test-statistic is: -7.7769 30.2417

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81

> summary(ur.df(lfutures, type="drift", selectlags = "BIC"))
<Omitted>
Value of test-statistic is: -1.3358 1.1551

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81

> summary(ur.df(diff(lfutures), type="drift", selectlags = "BIC"))
<Omitted>
Value of test-statistic is: -8.1164 32.9392

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81
```

ขั้นตอนถัดไป เราสามารถทดสอบ Cointegration จากความสัมพันธ์ระหว่างตัวแปรทั้งสองข้างต้น เราสามารถเขียนสมการความสัมพันธ์ระยะยาวได้เป็น

$$f_t = s_t + c + u_t$$

ดังนั้นในกรณีนี้เราสามารถเขียน $u_t = f_t - s_t$ หรือ $\boldsymbol{\beta} = (1 \quad -1)'$ แล้วทดสอบ Augmented Dickey Fuller Unit root กับ u ด้วยคำสั่งต่อไปนี้

```
> u<-lfutures-lspot
> summary(ur.df(u, type="drift", selectlags = "BIC"))

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
      Min        1Q    Median        3Q       Max
-0.053465 -0.005260  0.000870  0.006958  0.035017

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.003307   0.001250  -2.645  0.00924 **
z.lag.1     -0.899482   0.110577  -8.134 4.14e-13 ***
z.diff.lag   0.175245   0.089293   1.963  0.05199 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 ''. 0.1 ''  1

Residual standard error: 0.0132 on 121 degrees of freedom
Multiple R-squared:  0.4018,  Adjusted R-squared:  0.3919
F-statistic: 40.64 on 2 and 121 DF,  p-value: 3.151e-14


Value of test-statistic is: -8.1345 33.0866

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81
```

จากการทดสอบ Augmented Dickey Fuller unit root กับตัวแปร u_t พบว่าค่าสถิติเท่ากับ -8.13 ซึ่งน้อยกว่า ค่า critical value (-2.88) เราสามารถปปฏิเสธสมมุติฐานว่า u_t เป็น unit root และสรุปว่าตัวแปรทั้งสอง Cointegrated กัน

6.3.2 การทดสอบกรณีที่เราไม่ทราบค่า cointegrating vector

ในกรณีนี้เราจำเป็นต้องประมาณค่า cointegrating vector ขึ้นมาก ดังนั้นในขั้นตอนแรกของการทดสอบเราจำเป็นต้องประมาณค่า β โดยที่เราอาจจะเขียนในรูปบรรทัดฐาน (normalized) ซึ่งเขียนเป็นสมการได้

$$y_{1t} = c + \beta_2 y_{2t} + ... + \beta_n y_{nt} + u_t \tag{6.8}$$

การทดสอบ cointegration จะเป็นการทดสอบ unit root ของ $\hat{u}_t = y_{1t} - \hat{c} - \hat{\beta}_2 y_{2t} - ... - \hat{\beta}_n y_{nt}$ โดยที่ $\hat{c}, \hat{\beta}_2, ..., \hat{\beta}_n$ เป็นตัวประมาณค่า OLS ของสัมประสิทธิ์

Phillips and Ouliaris (1990) ได้แสดงให้เห็นว่าตัวทดสอบ ADF และ PP ของ cointegrating residual ในกรณีที่เราต้องประมาณค่าสัมประสิทธิ์ของ cointegrating vector ไม่ได้มีการแจกแจง

แบบ ADF distribution แบบปกติ เนื่องจากตัวประมาณค่า OLS ในขั้นตอนแรกเผชิญกับปัญหา spurious regression ดังนั้นในการทดสอบ cointegration ต้องใช้การแจกแจงจากตารางข้างล่าง

Phillips and Orliaris (1990) Distribution

	No constant			Constant			Constant and trend		
	Significance level			Significance level			Significance level		
n-1	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
1	-2.4505	-2.7619	-3.3865	-3.0657	-3.3654	-3.9618	-3.5184	-3.8	-4.3628
2	-2.9873	-3.2667	-3.8395	-3.4494	-3.7675	-4.3078	-3.8429	-4.1567	-4.6451
3	-3.4446	-3.7371	-4.3038	-3.8329	-4.1121	-4.7325	-4.195	-4.4895	-5.0433
4	-3.8068	-4.1261	-4.672	-4.1565	-4.4542	-5.0728	-4.4625	-4.7423	-5.3576
5	-4.1416	-4.3999	-4.9897	-4.4309	-4.7101	-5.2812	-4.7311	-5.0282	-5.5849

Note: n is the number of variables.

ตัวอย่าง 6.4: การทดสอบความสัมพันธ์ระหว่างราคาปัจจุบันและราคาฟิวเจอร์กรณีไม่ทราบค่าสัมประสิทธิ์

ต่อเนื่องจากตัวอย่างที่ผ่านมา ในกรณีนี้สมมุติว่าเราไม่ทราบค่าของ cointegrating vector ดังนั้นสมมุติให้ $\boldsymbol{\beta} = (1 \quad -\beta_2)'$ ดังนั้นเราสามารถพิจารณาสมการถดถอย $f_t = c + \beta_2 s_t + u_t$ ซึ่งหลังจากที่เราประมาณค่าด้วย OLS เราสามารถที่จะหาค่า $\hat{u}_t = f_t - \hat{c} - \hat{\beta}_2 s_t$ ตามขั้นตอนในคำสั่งข้างล่าง

```
> m1<—lm(lfutures~lspot)
> summary(m1)

Call:
lm(formula = lfutures ~ lspot)

Residuals:
      Min        1Q    Median        3Q       Max
—0.053040 —0.005474  0.001128  0.006023  0.037175

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) —0.044097   0.023113  —1.908   0.0587 .
lspot        1.006206   0.003535 284.610   <2e—16 ***
———
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 ''. 0.1 ''  1

Residual standard error: 0.01349 on 124 degrees of freedom
Multiple R—squared:  0.9985,  Adjusted R—squared:  0.9985
F—statistic: 8.1e+04 on 1 and 124 DF,  p—value: < 2.2e—16
```

ค่าสัมประสิทธิ์ β_2 เท่ากับ 1.006 ซึ่งใกล้เคียงกับ 1 มาก หลังจากนั้นเราก็จะทดสอบ cointegrating residuals โดยใช้การทดสอบ Augmented Dickey Fuller Unit root

```
> uhat<- m1$residuals
> summary(ur.df(uhat, type="drift", selectlags = "BIC"))

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
      Min        1Q    Median        3Q       Max
-0.052712 -0.005953  0.001036  0.006614  0.035313

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.0001084  0.0011708  -0.093   0.9264
z.lag.1     -0.9224120  0.1112721  -8.290 1.81e-13 ***
z.diff.lag   0.1846947  0.0890896   2.073   0.0403 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 ''. 0.1 ''  1

Residual standard error: 0.01304 on 121 degrees of freedom
Multiple R-squared:  0.4104,  Adjusted R-squared:  0.4007
F-statistic: 42.12 on 2 and 121 DF,  p-value: 1.311e-14


Value of test-statistic is: -8.2897 34.364

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81
```

ได้ค่าสถิติ ADF เท่ากับ -8.4232 นำไปเปรียบเทียบกับค่า critical value ในตาราง Phillips-Orliaris กรณีที่มีค่าคงที่ (เนื่องจากรูปแบบสมการโคอินทิเกรชันมีพจน์ค่าคงที่) และ n-1=1 (เนื่องจากในกรณีนี้เรามีตัวแปรในสมการ cointegration สองตัวแปร) ที่นัยสำคัญ 0.05 มีค่าเท่ากับ -3.3656 เนื่องจากค่าสถิติที่ได้มีค่าต่ำกว่า critical value ดังนั้นเราสามารถปฏิเสธ H_0 และสรุปได้ว่าตัวแปรทั้งสองเป็น cointegration กัน

6.4 การประมาณค่า ECM ด้วย OLS

สมมุติว่าเราสนใจตัวแปร $(y_{1t} y_{2t})'$ ซึ่งมีความสัมพันธ์ cointegration แสดงได้ด้วยสมการ $y_{1t} = \beta_2 y_{2t} + u_t$ (หรือ $y_{1t} - \beta_2 y_{2t} = u_t \sim I(0)$) และเรามีตัวประมาณค่า $\hat{\beta}_2$ ที่มีคุณลักษณะเป็น consistent และเราต้องประมาณค่า ECM ในรูปแบบ

$$\Delta y_{1t} = c_1 + \alpha_1(y_{1,t-1} - \beta_2 y_{2,t-1}) + \sum_{i=1}^{p} \psi_{11}^{j} \Delta y_{1,t-j} + \sum_{i=1}^{p} \psi_{12}^{j} \Delta y_{2,t-j} + \varepsilon_{1t} \qquad (6.9)$$

$$\Delta y_{2t} = c_2 + \alpha_2(y_{1,t-1} - \beta_2 y_{2,t-1}) + \sum_{i=1}^{p} \psi_{21}^{j} \Delta y_{1,t-j} + \sum_{i=1}^{p} \psi_{22}^{j} \Delta y_{2,t-j} + \varepsilon_{2t} \quad (6.10)$$

เราสามารถแทนค่า $\hat{\beta}_2$ ลงในสมการเหมือนเป็นพารามิเตอร์ที่เราทราบค่า ดังนั้นเราสามารถพิจารณา disequilibrium error $(y_{1,t-1} - \beta_2 y_{2,t-1})$ เหมือนเป็นตัวแปรที่ทราบค่า เนื่องจากตัวแปรทุกตัวในสมการ **??** และ **??** เป็น $I(0)$ เราสามารถประมาณค่าสมการดังกล่าวด้วย OLS

ตัวอย่างที่ 6.5 การประมาณค่า ECM สำหรับสมการราคาปัจจุบันและราคาฟิวเจอร์

เราต้องการประมาณค่าสมการ ECM สำหรับล็อกของราคาปัจจุบันและล็อกของราคาฟิวเจอร์ซึ่งเราจำเป็นต้องหาอันดับของค่าล่าของผลต่างที่เหมาะสม ซึ่งอาจจะใช้ AIC, BIC ในการเลือก แต่ในที่นี้เราสมมุติให้แบบจำลองของเป็นดังนี้

$$\Delta s_t = c_1 + \alpha_1 \hat{u}_{t-1} + \psi_{11} \Delta s_{t-1} + \psi_{12} \Delta f_{t-1} + \varepsilon_{ft} \quad (6.11)$$

$$\Delta f_t = c_2 + \alpha_2 \hat{u}_{t-1} + \psi_{21} \Delta s_{t-1} + \psi_{22} \Delta f_{t-1} + \varepsilon_{st} \quad (6.12)$$

เราเริ่มจากการสร้าง series dlspot และ dlfuture ซึ่งจะมีสมาชิกจากคาบที่ t=2,...,126 โดยที่เวกเตอร์ $dlspot$ และ $dlfuture$ จะมีสมาชิกทั้งหมด 125 ตัวโดยตัวแรกเป็นข้อมูล ณ คาบที่ t=2 นอกจากนี้เรายังปรับให้เวกเตอร์ $uhat$ เริ่มต้นที่ช่วงเวลาเดียวกันคือ t=2

```
> dlspot<-diff(lspot)
> dlfutures<-diff(lfutures)
> uhat<-uhat[2:126]
```

เนื่องจากการเรียงของข้อมูลความสัมพันธ์ของสมการที่ **??** และ **??** จะเริ่มที่ t=3,...,126 หรือสมาชิกตำแหน่งที่ 2:125 ในเวกเตอร์ $dlspot$ และ $dlfuture$ หรือคำสั่งที่ใช้ประมาณค่าคือ

```
> ecm1<-lm(dlfutures[2:125]~uhat[1:124]+dlfutures[1:124]+dlspot[1:124])
> summary(ecm1)

Call:
lm(formula = dlfutures[2:125] ~ uhat[1:124] + dlfutures[1:124] +
    dlspot[1:124])

Residuals:
     Min       1Q   Median       3Q      Max
-0.31503 -0.03288  0.01358  0.04233  0.14965

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       0.004115   0.006281   0.655   0.5136
uhat[1:124]      -1.152182   0.598961  -1.924   0.0568 .
dlfutures[1:124]  0.071715   0.478109   0.150   0.8810
dlspot[1:124]     0.137282   0.498649   0.275   0.7836
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘’. 0.1 ‘’  1

Residual standard error: 0.06973 on 120 degrees of freedom
Multiple R-squared:  0.06491, Adjusted R-squared:  0.04154
F-statistic: 2.777 on 3 and 120 DF,  p-value: 0.04424

> ecm2<-lm(dlspot[2:125]~uhat[1:124]+dlfutures[1:124]+dlspot[1:124])
> summary(ecm2)

Call:
lm(formula = dlspot[2:125] ~ uhat[1:124] + dlfutures[1:124] +
    dlspot[1:124])

Residuals:
     Min       1Q   Median       3Q      Max
-0.34142 -0.03319  0.01113  0.04038  0.13716

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       0.004314   0.005893   0.732    0.466
uhat[1:124]      -0.212277   0.561883  -0.378    0.706
dlfutures[1:124] -0.102953   0.448512  -0.230    0.819
dlspot[1:124]     0.288494   0.467780   0.617    0.539

Residual standard error: 0.06541 on 120 degrees of freedom
Multiple R-squared:  0.03337, Adjusted R-squared:  0.009204
F-statistic: 1.381 on 3 and 120 DF,  p-value: 0.252
```

เขียนเป็นผลการประมาณค่าได้ดังนี้

$$\Delta f_t = 0.004 + -1.152\hat{u}_{t-1} + 0.072\Delta f_{t-1} + 0.137\Delta s_{t-1} + \varepsilon_{st}$$
$$\Delta s_t = 0.004 - 0.212\hat{u}_{t-1} - 0.103\Delta f_{t-1} + 0.289\Delta s_{t-1} + \varepsilon_{ft}$$

หากพิจารณาจากสมการแรก ถ้า $f_{t-1} > \hat{\beta}s_{t-1} + \hat{c}$ จะส่งผลให้ $\hat{u}_{t-1}$ มากกว่าศูนย์ จากค่าสัมประสิทธิ์ $\hat{\alpha}_1 = -1.152$ แสดงว่าอัตราการเปลี่ยนแปลงของอัตราแลกเปลี่ยน futures จะลดลงมาเพื่อปรับเข้าสู่ดุลยภาพ นอกจากนี้สัมประสิทธิ์หน้า Δf_{s-1} สะท้อนว่าหากอัตราการเปลี่ยนแปลงของอัตราแลกเปลี่ยน future เปลี่ยนแปลงในคาบก่อนหน้าจะส่งผลให้อัตราการเปลี่ยนแปลงของอัตรา

แลกเปลี่ยน futures เปลี่ยนแปลงในทิศทางเดียวกัน

6.5 แบบ จำลอง เวคเตอร์ เอ เรอ คอ เร คชัน (Vector Error Correction Model: VECM)

Granger นำเสนอแบบจำลองโคอินทิเกรตที่อธิบายความสัมพันธ์ระยะยาว และแบบจำลองเอเรอคอเรคชันซึ่งนำเสนอการปรับตัวระยะสั้น Johansen (1991) ได้รวมความสัมพันธ์ทั้งระยะยาวและการปรับตัวระยะสั้นในรูปแบบของสมการของตัวแปรในรูปเวคเตอร์ โดยมีพื้นฐานจากแบบจำลองเวคเตอร์ออโตรีเกรสชัน ซึ่งหากพิจารณาตัวแปร $y_{1t}, ..., y_{nt}$ ในรูปเวกเตอร์ $(n \times 1)$ เขียนเป็นสัญลักษณ์ $\boldsymbol{Y}_t$ ในรูปสมการ $VAR(P)$ ดังต่อไปนี้

$$\boldsymbol{Y}_t = \boldsymbol{\Pi}_1 \boldsymbol{Y}_{t-1} + ... + \boldsymbol{\Pi}_p \boldsymbol{Y}_{t-p} + \boldsymbol{\epsilon}_t, t = 1, ..., T$$

ถ้า VAR(p) มีลักษณะเป็นมีคุณสมบัติไม่นิ่ง(nonstationary) แสดงว่าตัวแปรบางตัวหรือทั้งหมดใน $\boldsymbol{Y}_t$ มีลักษณะเป็น $I(1)$ ซึ่งเราสามารถพิจารณาต่อไปว่าตัวแปรเหล่านั้นโคอินทิเกรตกันหรือไม่

6.5.1 รูปแบบของสมการ

Johansen เสนอว่าเราสามารถทดสอบความสัมพันธ์ระยะยาว ได้จากการแปลงแบบจำลองเวคเตอร์ออโตรีเกรสชันของตัวแปร $\boldsymbol{Y}_t$ ให้อยู่ในรูปของแบบจำลองเวคเตอร์เอเรอคอเรคชัน ซึ่งเขียนเป็นสมการได้ดังนี้

$$\Delta \boldsymbol{Y}_t = \boldsymbol{\Pi} \boldsymbol{Y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{Y}_{t-1} + ... + \boldsymbol{\Gamma}_{p-1} \Delta \boldsymbol{Y}_{t-(p-1)} + \boldsymbol{\epsilon}_t$$

ซึ่งจะเห็นได้ว่าสมการนี้คือแบบจำลอง $VAR(p-1)$ ของ $\Delta \boldsymbol{Y}_t$ โดยที่เมตริกซ์ $\boldsymbol{\Pi}$ คือเมทริกซ์สัมประสิทธิ์ที่แสดงความสัมพันธ์ระยะยาว และ $\boldsymbol{\Gamma}_k$ คือเมทริกซ์สัมประสิทธิ์ที่อธิบายผลของการปรับตัวระยะสั้นในคาบที่ผ่านมา

ในแบบจำลองเวคเตอร์เอเรอคอเรคชันเราพบว่า $\Delta \boldsymbol{Y}_t$ และค่าล่าของ $\Delta \boldsymbol{Y}_t$ เป็น $I(0)$ จะเห็นได้ว่า เหลือเพียงพจน์ $\boldsymbol{\Pi} \boldsymbol{Y}_{t-1}$ ที่มีโอกาสเป็น $I(1)$ และหากเราพิจารณา $\Delta \boldsymbol{Y}_t$ เราพบว่า $\Delta \boldsymbol{Y}_t$ จะเป็น $I(0)$ ถ้า $\boldsymbol{\Pi} \boldsymbol{Y}_{t-1}$ เป็น $I(0)$ ดังนั้น $\boldsymbol{\Pi} \boldsymbol{Y}_{t-1}$ จะเป็นเมตริกซ์ที่ระบุความสัมพันธ์ระยะยาวหากตัวแปรโคอินทิเกรตกัน

กรณีที่ $\boldsymbol{Y}_t$ มีลักษณะเป็น $I(1)$ จะเกิดขึ้นเมื่อ $\boldsymbol{\Pi}$ จะเป็นเมทริกซ์เอกฐาน (singular matrix) ซึ่งกรณีดังกล่าวเกิดขึ้นเมื่อค่าลำดับของเมทริกซ์มีค่าน้อยกว่าจำนวนตัวแปร ($rank(\boldsymbol{\Pi}) = r < n$) ซึ่งค่าดังกล่าวเป็นไปได้สองกรณีคือ

1. $rank(\boldsymbol{\Pi}) = 0$ แสดงว่า $\boldsymbol{\Pi}$ และ $\boldsymbol{Y}_t$ เป็น $I(1)$ และตัวแปรมิได้โคอินทิเกรตกัน และแบบจำลอง $VECM(p-1)$ จะลดลงเหลือแค่ $VAR(p-1)$ ของตัวแปรในรูปผลต่างหนึ่งอันดับ

2. $0 < rank(\mathbf{\Pi}) = r < n$ กรณีนี้แสดงว่า $\boldsymbol{Y}_t$ เป็น $I(1)$ และตัวแปรโคอินทิเกรตกัน โดยที่มี r จำนวนความสัมพันธ์เท่ากับ r

จากคุณสมบัติข้างต้น การทดสอบโคอินทิเกรชันตามวิธีของ Johansen จะพิจารณา $\mathbf{\Pi}$

6.5.2 การทดสอบโคอินทิเกรชันตามวิธีของ Johansen

การทดสอบโคอินทิเกรชันจะพิจารณาจำนวนลำดับ (rank) ของเมทริกซ์ $\mathbf{\Pi}$ ซึ่งสัมพันธ์กับค่าลักษณะเฉพาะ (eigenvalues) โดยที่ค่าลักษณ์ะเฉพาะคือรากลักษณะเฉพาะ (characteristic roots) ที่มีค่าไม่เท่ากับศูนย์ กำหนดให้ λ_i แทนค่าลักษณะเฉพาะที่ i โดยที่ $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ และค่ารากมีคุณสมบัติ $|\lambda_i| < 1$ และ $\lambda_i \geq 0$

ถ้าตัวแปรใน $\boldsymbol{Y}_t$ ไม่ได้โคอินทิเกรตกัน ค่าลำดับของ $\mathbf{\Pi}$ จะไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ ดังนั้น $\lambda_i \approx 0$ สำหรับทุกค่าของ i อย่างไรก็ตาม ในการทดสอบเราจะพิจารณา $\ln(1-\lambda_i)$ ซึ่งหาก $\lambda_i = 0$ แล้ว $\ln(1-\lambda_i) = 0$

หากค่าอันดับของ $\mathbf{\Pi}$ เท่ากับ 1 แล้ว $\ln(1-\lambda_1)$ จะมีค่าเป็นลบ ในขณะที่ $\ln(1-\lambda_i) = 0$ สำหรับ $i > 1$

Johansen เสนอ ตัว สถิติ ที่ ใช้ ทดสอบ สอง ค่า ได้แก่ Johansen's Trace statistics และ Johansen's Maximum Eigenvalue statistics

Johansen's Trace Statistic

Johansen's Trace statistic ทดสอบสมมุติฐานต่อไปนี้

$$H_0(r_0) : r = r_0 \quad \text{vs} \quad H_1(r_0) : r > r_0$$

โดยที่ LR Trace statistics และคำนวณได้ดังนี้

$$LR_{trace}(r_0) = -T \sum_{i=r_0+1}^{n} \ln(1-\hat{\lambda}_i)$$

โดยที่ r_0 คือจำนวนความสัมพันธ์ในสมมุติฐานหลัก และ $\hat{\lambda}_i$ คือค่าประมาณของค่าลักษณะเฉพาะอันดับที่ i ซึ่งหาก $rank(\mathbf{\Pi}) = r_0$ แล้ว $\hat{\lambda}_{r_0+1}, ..., \hat{\lambda}_n$ จะมีค่าเข้าใกล้ศูนย์ และ $LR_{trace}(r_0)$ ควรจะมีค่าน้อย

Johansen's Maximum Eigenvalue Statistic

Johansen ยังได้เสนอการทดสอบ LR Maximum Eigenvalue statistic สำหรับสมมุติฐานต่อไปนี้

$$H_0(r_0) : r = r_0 \quad \text{vs} \quad H_1(r_0) : r = r_0 + 1$$

โดยเราเรียก LR statistic ว่า *maximum eigenvalue statistic* และคำนวณได้ดังนี้

$$LR_{max}(r_0) = -T\ln(1-\hat{\lambda}_{r_0+1})$$

Johansen and Juselius (1990) เสนอว่าค่าวิกฤตที่ใช้ทดสอบสำหรับตัวสถิติทั้งสองมีการแจกแจงขึ้นอยู่กับรูปแบบของสมการ ซึ่งหากค่าสถิติสูงกว่าค่าวิกฤตที่ Johansen เสนอ เราจะปฏิเสธสมมุติฐานหลักที่ว่าตัวแปรมีความสัมพันธ์ r_0 สมการ และยอมรับว่ามีความสัมพันธ์มากกว่า r_0 สำหรับการทดสอบ Trace statistics ในขณะที่หากเราใช้การทดสอบ Max Eigenvalue statistic เราจะยอมรับว่ามีความสัมพันธ์เท่ากับ r_0+1 ความสัมพันธ์

Johansen เสนอการทดสอบในลักษณะที่เป็นอันดับ (sequential testing procedure) โดยที่กระบวนการดังกล่าวจะระบุจำนวนของความสัมพันธ์ได้อย่างคงเส้นคงว่า โดยที่กระบวนการเป็นดังนี้

- ขั้นตอนแรกเราทดสอบ $H_0(r_0=0)$ กับ $H_1(r_0>0)$ ถ้าเราไม่สามารถปฏิเสธสมมุติฐานหลัก ตัวแปรไม่โคอินทิเกรตกัน

  ถ้าเราปฏิเสธ $H_0(r_0=0)$ เราสามารถสรุปได้ว่า มีจำนวนความสัมพันธ์อย่างน้อยหนึ่งความสัมพันธ์อย่างน้อยหนึ่งความสัมพันธ์และทดสอบในขั้นต่อไป

- ทดสอบ $H_0(r_0=1)$ กับ $H_1(r_0>1)$ ถ้าเราไม่สามารถปฏิเสธสมมุติฐานหลัก เราสรุปว่ามีมีจำนวนความสัมพันธ์เท่ากับหนึ่ง

  แต่ถ้า เราปฏิเสธสมมุติฐาน แสดงว่า มีมีจำนวนความสัมพันธ์อย่างน้อยสองสมการ

- กระบวนการที่เป็นอันดับดังกล่าวจะดำเนินไปเรื่อยๆจนกระทั่งสมมุติฐานหลักถูกปฏิเสธ

ในกรณีที่เราทดสอบแล้วพบว่ามีจำนวนความสัมพันธ์เท่ากับ r เราสามารถพิจารณาเวคเตอร์ $\mathbf{\Pi}$ ซึ่งสามารถแยกออกได้เป็นเมทริกซ์ $\boldsymbol{\alpha}$ และ $\boldsymbol{\beta}'$ ซึ่งมีมิติเท่ากับ $n\times r$ และ $r\times n$ ตามลำดับ

$$\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$$

ตัวอย่างเช่น หากเราพิจารณากรณีที่ตัวแปรเท่ากับ 4 หากเราพบความสัมพันธ์เท่ากับหนึ่ง

$$\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}' = \begin{bmatrix}\alpha_{11}\\ \alpha_{12}\\ \alpha_{13}\\ \alpha_{14}\end{bmatrix}\begin{bmatrix}\beta_{11} & \beta_{12} & \beta_{13} & \beta_{14}\end{bmatrix}$$

หากเราพบความสัมพันธ์เท่ากับสอง

$$\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}' = \begin{bmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \\ \alpha_{13} & \alpha_{23} \\ \alpha_{14} & \alpha_{24} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} \end{bmatrix}$$

ในกรณีของความสัมพันธ์เท่ากับหนึ่ง เราสามารถเขียนส่วนของ $\mathbf{\Pi}\boldsymbol{Y}_{t-1}$ ได้เป็น

$$\mathbf{\Pi}\boldsymbol{Y}_{t-1} = \begin{bmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{14} \end{bmatrix} (\beta_{11}y_{1,t-1} + \beta_{12}y_{2,t-1} + \beta_{13}y_{3,t-1} + \beta_{14}y_{4,t-1})$$

ซึ่งเราสามารถนำไปเขียนสมการเอเรอคอเรคชันสำหรับแต่ละตัวแปรได้

6.5.3 ขั้นตอนในการทดสอบโคอินทิเกรชันและสร้างแบบจำลองเวคเตอร์เอเรอคอเรคชัน

ขั้นตอนในการสร้างแบบจำลอง Cointegration ของ Johansen คือ

- ระบุ Order และประมาณค่า VAR(p) สำหรับ $\boldsymbol{Y}_t$
- สร้าง likelihood ratio tests สำหรับการทดสอบ rank of $\mathbf{\Pi}$ เพื่อระบุจำนวน cointegrating vectors
- หากจำเป็นเราอาจจะสมมุติ normalization และระบุเงื่อนไขของ cointegrating vectors
- จาก normalized cointegrating vectors ที่ได้ระบุไว้ เราสามารถประมาณค่า cointegrated VECM ด้วย MLE

Johansen (1995) ใช้ specification ของ VECM ในรูปแบบต่อไปนี้

$$\boldsymbol{Y}_t = \mathbf{\Phi}\boldsymbol{D}_t + \mathbf{\Pi}_1\boldsymbol{Y}_{t-1} + ... + \mathbf{\Pi}_p\boldsymbol{Y}_{t-p} + \boldsymbol{\epsilon}_t, t = 1, ..., T$$

โดยที่ $\mathbf{\Phi}\boldsymbol{D}_t = \boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$ เป็น deterministic terms โดยที่พฤติกรรมของ deterministic term ของ $\boldsymbol{Y}_t$ สามารถแยกออกได้เป็น 5 กรณี

1. Model $H_2(r)$: $\boldsymbol{\mu}_0$ (no constant)
2. Model $H_1^*(r)$: $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 = \boldsymbol{\alpha}\boldsymbol{\rho}_0$ (restricted constant)

3. Model $H_1(r)$: $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0$ (unrestricted constant)
4. Model $H^*(r) = \boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha\rho_1} t$ (restricted trend)
5. Model $H(r) = \boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu_1} t$ (unrestricted constant and trend)

ตัวอย่างที่ 6.6 การประมาณค่าสมการ VECM สมการราคาปัจจุบันและราคาฟิวเจอร์

เราจะพิจารณาข้อมูลในตัวอย่างที่ 6.5 ด้วยแบบจำลอง VECM ซึ่งประมาณค่าสมการ cointegration และ VECM ไปพร้อม ๆ กัน

ขั้นตอนแรกในการทดสอบ คือการหาอันดับที่เหมาะสมสำหรับ VECM(p-1) โดยจะเป็นอันดับของ VAR(p) ซึ่งในที่นี้เราจะใช้ package *vars* และคำสั่ง VAR ในการประมาณค่าข้อมูล fsprice ซึ่งคือเมทริกซ์ที่ประกอบด้วยข้อมูล lfutures กับ lspot และเลือกใช้ Model selection AIC ในการหาอันดับที่เหมาะสม โดยชุดคำสั่งแสดงในรูปต่อไปนี้

```
> fsprice<-cbind(lfutures,lspot)
> head(fsprice)
      lfutures    lspot
[1,] 6.287115 6.280134
[2,] 6.198682 6.199535
[3,] 6.151242 6.156004
[4,] 6.164367 6.179250
[5,] 6.167307 6.178836
[6,] 6.193384 6.174411
> library(vars)
> var.mod<-VAR(fsprice, lag.max = 4, ic = c("AIC"))
> var.mod

VAR Estimation Results:
=======================

Estimated coefficients for equation lfutures:
=============================================
Call:
lfutures = lfutures.l1 + lspot.l1 + lfutures.l2 + lspot.l2 + lfutures.l3 +
    lspot.l3 + lfutures.l4 + lspot.l4 + const

lfutures.l1    lspot.l1 lfutures.l2    lspot.l2 lfutures.l3    lspot.l3
    lfutures.l4    lspot.l4
 0.03901903  1.20084532  0.06312604 -0.48169298 -0.91804903  1.40139010
     0.37178420 -0.69722519
      const
 0.13406246


Estimated coefficients for equation lspot:
==========================================
Call:
lspot = lfutures.l1 + lspot.l1 + lfutures.l2 + lspot.l2 + lfutures.l3 + lspot.l3
    + lfutures.l4 + lspot.l4 + const

lfutures.l1    lspot.l1 lfutures.l2    lspot.l2 lfutures.l3    lspot.l3
    lfutures.l4    lspot.l4
 -0.1738544   1.3769259   0.1714170  -0.5291331  -0.6237252   1.0354564
     0.3430320  -0.6255168
      const
  0.1684990
```

จากผลการประมาณค่า เราจะเห็นได้ว่า lag ที่โปรแกรมเลือก คือ 4 ซึ่งแสดงด้วยสัญลักษณ์ .l4 แสดงว่าอันดับที่เหมาะสมสำหรับ VAR คือ VAR(4) ดังนั้น ในการประมาณค่าแบบจำลอง VECM เราจะเลือกใช้แบบจำลอง VECM(3)

ในการทดสอบ Johansen และ ประมาณค่า VECM เราจะใช้คำสั่ง *ca.jo* ใน package urca โดยที่เราต้องระบุข้อมูล ($fsprice$) วิธีในการทดสอบ (trace) รูปแบบของสมการ ($const$) และจำนวน lag ในแบบจำลอง VAR ซึ่งในที่นี้คือ `K=4 =4`` โดยเก็บผลไว้ในชื่อ `fsprice.rc` และเรียกดูผลด้วย `summary(fsprice.rc)`

```
> library(urca)
> fsprice.rc <—ca.jo(fsprice, type = c("trace"), ecdet = c("const"), K=4)
> summary(fsprice.rc)

######################
# Johansen—Procedure #
######################

Test type: trace statistic , without linear trend and constant in cointegration

Eigenvalues (lambda):
[1] 2.847322e—01 2.434360e—02 2.081668e—17

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  3.01  7.52  9.24 12.97
r = 0  | 43.89 17.85 19.96 24.60

Eigenvectors, normalised to first column:
(These are the cointegration relations)

            lfutures.l4  lspot.l4  constant
lfutures.l4  1.00000000  1.000000  1.000000
lspot.l4    —1.00451521 —1.422650 —1.256829
constant     0.03337561  2.822588  1.482432

Weights W:
(This is the loading matrix)

           lfutures.l4   lspot.l4      constant
lfutures.d  —1.5094647 0.06534493 —1.823573e—13
lspot.d     —0.3469295 0.06379890 —3.903367e—14
```

จะได้ค่าสถิติซึ่งมีการทดสอบแบบเป็นลำดับ

- $H_0 : r = 0$ vs $H_1 : r > 0$ ค่าสถิติ 43.89 มากกว่าค่า c.v.19.96 กรณี significance level = 0.05 เราสามารถปฏิเสธสมมุติฐานหลัก และสรุปว่า $r > 0$
- $H_0 : r = 1$ vs $H_1 : r > 1$ ค่าสถิติ 3.01 น้อยกว่าค่า c.v. 9.24 กรณี significance level = 0.05 เราไม่สามารถปฏิเสธสมมุติฐานหลัก และยอมรับว่า $r = 1$ หรือ $lfutures$ และ $lspot$ มี cointegration 1 ความสัมพันธ์

เราสามารถสร้างสมการเวกเตอร์เอเรอคอเรกชัน ได้ด้วยคำสั่ง $cajorls$ โดยระบุว่าใช้รูปแบบจาก $fsprice.rc$ และจำนวนความสัมพันธ์เท่ากับ 1 ($r = 1$)

```
> fsprice.vecm <-cajorls(fsprice.rc, r= 1)
> fsprice.vecm
$`rlm`

Call:
lm(formula = substitute(form1), data = data.mat)

Coefficients:
              lfutures.d  lspot.d
ect1          -1.50946    -0.34693
lfutures.dl1  -0.98859    -0.20081
lspot.dl1      1.24993     0.42485
lfutures.dl2  -0.92914    -0.03298
lspot.dl2      0.76844    -0.10408
lfutures.dl3  -1.86016    -0.66937
lspot.dl3      2.18017     0.94147

$beta
                   ect1
lfutures.l4  1.00000000
lspot.l4    -1.00451521
constant     0.03337561
```

จากคำสั่งดังกล่าว เราสามารถเขียนสมการ ECM ซึ่งอยู่ในส่วน `$`rlm`` และแต่ละสมการเรียงตามแต่ละคอลัมน์ได้ดังนี้

$$\Delta f_t = -1.51u_{t-1} - 0.99\Delta f_{t-1} + 1.25\Delta s_{t-1} - 0.93\Delta f_{t-2} + 0.77\Delta s_{t-2} - 1.86\Delta f_{t-3} + 2.18\Delta s_{t-3}$$

$$\Delta s_t = -0.35u_{t-1} - 0.20\Delta f_{t-1} + 0.42\Delta s_{t-1} - 0.03\Delta f_{t-2} - 0.1\Delta s_{t-2} - 0.67\Delta f_{t-3} + 0.94\Delta s_{t-3}$$

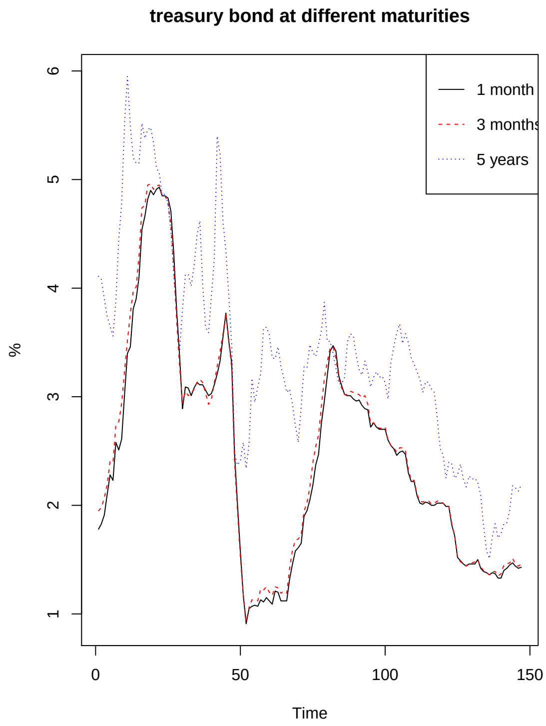
โดยที่สมการความสัมพันธ์ระยะยาวได้แก่

$$u_t = f_t - 1.005s_t + 0.033$$

ตัวอย่างที่ 6.7: การทดสอบและประมาณค่า VECM สมการอัตราผลตอบแทนพันธบัตรรัฐบาล

```
> tbond
   <-read.csv("https://raw.githubusercontent.com/chaleampong/EC435/master/tbond.csv",
   header = TRUE)
> head(tbond)
> m1<-tbond$m1
> m3<-tbond$m3
> y5<-tbond$y5
> ts.plot(cbind(m1,m3,y5), lty=c(1:3), ylab="%",col=c("black", "red", "blue"),
   main="treasury bond at different maturities")
> legend("topright", legend = c("1 month", "3 months", "5 years"), col=c("black",
   "red", "blue"), lty = 1:3, xjust = 1, yjust = 1)
```

รูปที่ 6.2: ผลตอบแทนพันธบัตร



ขั้นตอนแรกของการทดสอบคือการหาอันดับที่เหมาะสมของ VECM โดยการจัดตัว-แปรทั้งสามให้อยู่ในรูปเมทริกซ์ `rterm` ด้วยคำสั่ง `cbind` หลังจากนั้น ประมาณค่า VAR ด้วย package `vars` และคำสั่ง `VAR` โดยระบุข้อมูลที่ประมาณค่าคือ `rterm` จำนวนอันดับที่สูงที่สุด `lag.max=6`และเลือก model selection คือ AIC ด้วย `ic=c("AIC")`

```
> library(vars)
> rterm <- cbind(m1, m3, y5)
> var.mod <-VAR(rterm, lag.max=6, ic= c("AIC"))
> var.mod

VAR Estimation Results:
=======================

Estimated coefficients for equation m1:
=======================================
Call:
m1 = m1.l1 + m3.l1 + y5.l1 + m1.l2 + m3.l2 + y5.l2 + m1.l3 + m3.l3 + y5.l3 + const

      m1.l1        m3.l1        y5.l1        m1.l2        m3.l2        y5.l2
                 m1.l3
-0.148170687  1.468544693  0.017887883  0.319187528 -0.565154976  0.035796258
    -0.100551087
      m3.l3        y5.l3        const
 0.015396105 -0.054976949 -0.009814944


Estimated coefficients for equation m3:
=======================================
Call:
m3 = m1.l1 + m3.l1 + y5.l1 + m1.l2 + m3.l2 + y5.l2 + m1.l3 + m3.l3 + y5.l3 + const

     m1.l1       m3.l1       y5.l1       m1.l2       m3.l2       y5.l2
         m1.l3       m3.l3
-0.77262617  2.06919377  0.07347698  0.40365078 -0.63296798 -0.00663435
    -0.35137281  0.25663998
     y5.l3       const
-0.05718647  0.00346167


Estimated coefficients for equation y5:
=======================================
Call:
y5 = m1.l1 + m3.l1 + y5.l1 + m1.l2 + m3.l2 + y5.l2 + m1.l3 + m3.l3 + y5.l3 + const

     m1.l1       m3.l1       y5.l1       m1.l2       m3.l2       y5.l2
         m1.l3       m3.l3
-0.94700137  1.02677142  1.27187191 -0.41862278  0.01796799 -0.41027244
    0.34032262  0.01702885
     y5.l3       const
 0.06041757  0.12019345
```

หากพิจารณาแบบจำลอง VAR ที่เหมาะสมคือ VAR(3) ดังนั้น เราจะประมาณค่าแบบจำลอง VECM(2)

เราทดสอบ Johansen's test โดยใช้คำสั่ง `ca.jo` ใน package `urca` โดยระบุตัวสถิติที่ใช้คือ trace statistic ด้วย `type=c("trace")` และรูปแบบของ cointegration มีค่าคงที่ `ecdet=c("const")` และจำนวนอันดับของ VAR `k=3` โดยเก็บผลไว้ในชื่อ `rterm.rc` และเรียกดูผลด้วย `summary(rterm.rc)`

```
> rterm.rc <-ca.jo(rterm, type=c("trace"), ecdet=c("const"), K=3)
> summary(rterm.rc)

######################
# Johansen-Procedure #
######################

Test type: trace statistic , without linear trend and constant in cointegration

Eigenvalues (lambda):
[1] 0.15818825 0.10127853 0.02951317 0.00000000

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 2 |  4.31  7.52  9.24 12.97
r <= 1 | 19.69 17.85 19.96 24.60
r = 0  | 44.49 32.00 34.91 41.07

Eigenvectors, normalised to first column:
(These are the cointegration relations)

               m1.l3      m3.l3      y5.l3  constant
m1.l3     1.00000000  1.0000000  1.0000000  1.000000
m3.l3    -1.01768859 -1.8270107 -1.1379178 -1.443085
y5.l3     0.02286757  0.6652551  0.4414349  1.046099
constant  0.01082500 -0.1360052 -1.0450896 -4.791725

Weights W:
(This is the loading matrix)

          m1.l3        m3.l3        y5.l3     constant
m1.d -0.9588657  0.034340910 -0.005009448 2.724167e-16
m3.d -0.7547702  0.052367170 -0.017945128 2.186883e-16
y5.d -0.8972888 -0.004266062 -0.123746712 3.509375e-16
```

จากค่าสถิติซึ่งมีการทดสอบแบบเป็นลำดับ (ในกรณีนี้ใช้ significance level เท่ากับ 0.1)

- $H_0 : r = 0$ vs $H_1 : r > 0$ ค่าสถิติเท่ากับ 44.49 > Critical value (=32) เราสามารถปฏิเสธสมมุติฐานหลักที่ว่า $r = 0$ และยอมรับว่า $r > 0$
- $H_0 : r = 1$ vs $H_1 : r > 1$ ค่าสถิติเท่ากับ 19.69 > Critical value (=17.85) เราสามารถปฏิเสธสมมุติฐานหลักที่ว่า $r = 1$ และยอมรับว่า $r > 1$
- $H_0 : r = 2$ vs $H_1 : r > 2$ ค่าสถิติเท่ากับ 4.31 < Critical value (=7.52) เราไม่สามารถปฏิเสธสมมุติฐานหลักที่ว่า $r = 2$

สรุปว่าตัวแปรทั้ง 3 cointegrated กัน และมีความสัมพันธ์ 2 สมการ จากผลการทดสอบเราสามารถประมาณ VECM ด้วยคำสั่ง `cajorls` โดยระบุรูปแบบสมการเช่นเดียวกับ `rterm.rc` และจำนวนความสัมพันธ์ `` `r=2 `` โดยเก็บผลการประมาณค่าไว้ในชื่อ `rterm.vecm`


```
> rterm.vecm <-cajorls(rterm.rc, r=2)
> rterm.vecm
$`rlm`

Call:
lm(formula = substitute(form1), data = data.mat)

Coefficients:
        m1.d      m3.d      y5.d
ect1    -0.92452  -0.70240  -0.90155
ect2     0.91309   0.67245   0.92095
m1.dl1  -1.14885  -0.77507  -0.96384
m3.dl1   1.46761   1.06584   1.00361
y5.dl1   0.01948   0.07918   0.31122
m1.dl2  -0.82876  -0.36819  -1.36022
m3.dl2   0.90168   0.43010   1.00247
y5.dl2   0.05479   0.07082  -0.11098


$beta
                  ect1          ect2
m1.l3     1.000000e+00  2.220446e-16
m3.l3    -4.440892e-16  1.000000e+00
y5.l3    -7.849079e-01 -7.937354e-01
constant  1.954578e-01  1.814237e-01
```

ซึ่งเมื่อเรียกผลออกมา จะสามารถแบ่งออกเป็นสองส่วน

ส่วนแรกในบริเวณ `$beta` จะระบุความสัมพันธ์ระยะยาว หรือสมการ cointegration ตามคอลัมน์ โดยที่ที่และคอมันน์จะระบุด้วยชื่อ $ect1$ และ $ect2$ ซึ่งสามารถเขียนเป็นสมการได้ดังนี้

$ect1_{t-1} = 1m1_t - (-4.44x10^{-16})m3_t - 0.78y5_t + 0.195$ ซึ่งเราค่าสัมประสิทธิ์หน้า m3 มีค่าน้อยมาก เราสามารถตัดออก และเขียนสมการใหม่ได้เป็น $m1_t = 0.78y5_t - 0.195 + ect1_{t-1}$

$ect2_{t-1} = (2.22x10^{-16})m1_t + 1m3_t - 0.794y5 - t + 0.181$ ซึ่งเราค่าสัมประสิทธิ์หน้า m1 มีค่าน้อยมาก เราสามารถตัดออก และเขียนสมการใหม่ได้เป็น $1m3_t = 0.794y5 - t - 0.181 + ect2_{t-1}$

ส่วนสองในบริเวณ `$rlm` จะระบุการปรับตัวในระยะสั้น หรือ VECM โดยแต่ละคอลัมน์จะแทนแต่ละสมการ ได้แก่ m1.d ($\Delta m1_t$) m3.d ($\Delta m3_t$) และ y5.d ($\Delta y5_t$) ยกตัวอย่างเช่น สมการ m1.d สามารถเขียนได้ดังนี้

$\Delta m1_t = -0.92ect1_{t-1} + 0.91ect2_{t-1} - 1.14\Delta m1_{t-1} + 1.46\Delta m3_{t-1} + 0.02\Delta y5_{t-1} - 0.82\Delta m1_{t-2} + 0.90\Delta m3_{t-2} + 0.05\Delta y5_{t-2}$

ภาคผนวก A

การใช้โปรแกรม R เบื้องต้น

โปรแกรม R เป็นโปรแกรมภาษาคอมพิวเตอร์ขั้นสูงที่ถูกออกแบบมาสำหรับการคำนวณทางสถิติ และใช้กันอย่างแพร่หลายในแวดวงสถิติและการเงิน โปรแกรม R เป็นโปรแกรมที่ให้ผู้ใช้สามารถใช้ได้**ฟรี** และมีคู่มือมากมายที่อธิบายวิธีการใช้ สำหรับการสอนในวิชานี้เราจะใช้โปรแกรมต่อประสาน (interface) ชื่อว่า RStudio สำหรับการทำงานร่วมกับโปรแกรม R เนื่องจากมีหน้าต่างของคำสั่งที่ได้ใช้ตลอดจนผลการวาดกราฟในหน้าต่างย่อย นอกจากนี้เนื่องจากนักศึกษาจำนวนมากอาจจะคุ้นเคยการใช้โปรแกรมผ่าน Graphic User Interface ในวิชานี้ เราจะใช้ package `RCmdr` ช่วยในการนำเข้าข้อมูลและวิเคราะห์ทางสถิติเบื้องต้น

A.1 การติดตั้ง R

เพื่อที่จะติดตั้งโปรแกรม R เราสามารถดาวน์โหลดได้ที่ `http://www.cran.r-project.org` โดยที่เราสามารถเลือกระบบปฏิบัติการที่สอดคล้องกับเครื่องคอมพิวเตอร์ของเราได้ไม่ว่าเป็น Microsoft Windows, Linux หรือ Mac OS X ซึ่งในการติดตั้งนี้จะอธิบายในกรณีของระบบปฏิบัติการ Windows

- คลิก CRAN ตรงเมนูด้านซ้ายมือ
- เลือกเวปที่จะใช้ดาวน์โหลดซึ่งในไทยคือ `http://mirrors.psu.ac.th/pub/cran/`
- เลือก Operation ซึ่งในที่นี้ผมเลือก `R for Windows`
- เลือก `Base`
- คลิก `Download R x.x.x for Windows` และเลือก default ทุกคำถามที่ถาม

A.2 การติดตั้ง RStudio

หลังจากที่เราได้ติดตั้งโปรแกรม R แล้วเราจะมี Icon R อยู่บน Desktop เมื่อคลิกที่ไอคอนดังกล่าวจะมี Interface ในการทำงานขึ้นมา อย่างไรก็ตามในวิชานี้เราจะใช้ RStudio Interface ในการ

ทำงาน โดยเราสามารถดาวน์โหลด RStudio ได้ที่ `http://www.rstudio.com/ide/download/` ซึ่งสามารถเลือกระบบปฏิบัติการที่สอดคล้องกับเครื่องที่ใช้อยู่ได้

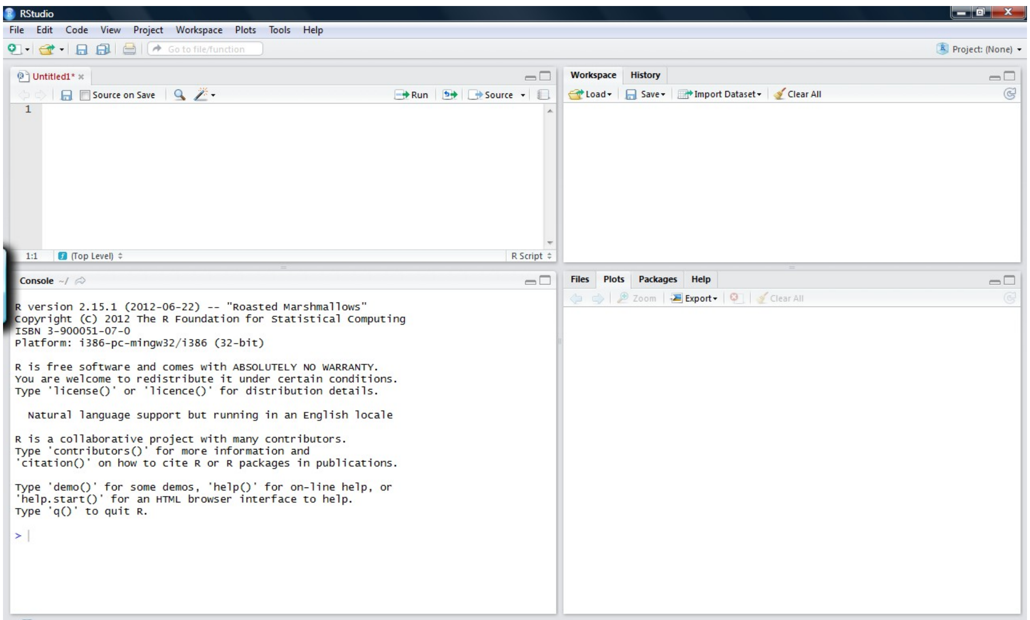
- คลิก `Download RStudio Desktop` เพื่อดาวน์โหลด
- คลิกไฟล์ที่อยู่ใต้หัวข้อ `Recommended For Your System`
- `Save File` และรันโปรแกรม

เนื่องจาก โปรแกรม R จำเป็น ต้องมี การ ติด ตั้ง โปรแกรม สำเร็จรูป ย่อย (package) ดัง นั้น เรา จะ ประมวลผล โปรแกรม R จาก Flashdrive หรือ External harddisk หรือ Drive D: ในห้องคอมพิว-เตอร์ ที่ เรา มี สิทธิ ที่ จะ แก้ไข เปลี่ยนแปลง ข้อมูล ได้ โดย การ คัด ลอก โฟลเดอร์ `R/R-x.x.x` และ `RStudio` จากโฟล์เดอร์ Program Files แล้ววางใน Flashdrive และเรียกโปรแกรมมาใช้โดยการคลิก `rstudio.exe` ในโฟลเดอร์ `RStudio/bin` โดยในครั้งแรกอาจจะมีการถามถึงโฟลเดอร์ที่มีโปร-แกรม R อยู่ใน Flashdrive

A.3 ผังของโปรแกรม RStudio

ผังของโปรแกรม RStudio จะประกอบด้วยหน้าต่างดังที่แสดงในรูป **??**

รูปที่ A.1: ผังโปรแกรม RStudio



หน้าต่างล่างซ้ายเรียกว่า **Console window** หรือ **Command window** ในหน้าต่างดังกล่าวจะมี ``>" prompt ซึ่งเราจะใส่คำสั่งหลังเครื่องหมายนั้นเพื่อให้โปรแกรม R ดำเนินการ

หน้าต่างบนซ้ายเรียกว่า **Editor window** หรือ **Script window** เป็นหน้าต่างที่ใช้เก็บชุดคำสั่งเพื่อที่จะเก็บและสามารถเรียกมาใช้ในภายหลังได้

หน้าต่างบนขวาเป็นหน้าต่าง **Workspace/ History window** เป็นหน้าต่างเราสามารถดูข้อ-มูลหรือค่าต่างๆที่เรากำลังเรียกมาอยู่ในหน่วยความจำ และคำสั่งเก่าๆที่ได้ดำเนินการไปแล้ว

หน้าต่างล่างขวาเป็นหน้าต่าง **Files/Plots/Packages/Help window** เป็นหน้าต่างที่เราสามารถเรียกข้อมูลขึ้นมา ดูกราฟ และติดตั้งหรือเรียกชุดคำสั่ง (packages)

A.4 ชุดคำสั่ง (package) หรือห้องสมุด (library)

โปรแกรม R สามารถวิเคราะห์ข้อมูลและสถิติได้อย่างหลากหลาย โดยชุดคำสั่งที่ใช้ร่วมกันจะถูกเก็บใน *packages* หรือ *libraries* โดยในการติดตั้งทั่วไปจะมีการติดตั้ง packages ที่ใช้บ่อยๆอยู่แล้ว โดยที่เราสามารถเรียกดู packages ที่ติดตั้งไว้แล้วในหน้าต่า packages หรือพิมพ์คำสั่ง `library()` หากกล่องหน้า packages มีเครื่องหมายถูกแสดงว่า packages นั้นได้ถูกโหลดมาไว้ในหน่วยความจำและสามารถใช้งานได้ทันที

ในขณะที่ packages ที่ยังไม่ได้ติดตั้งหรือยังไม่ได้โหลดเข้ามาในหน่วยความจสามารถดำเนินการได้โดยำ ใน RStudio ตัวอย่างเช่น หากเราต้องการติดตั้ง package ที่ชื่อว่า `RCmdr`

- การติดตั้ง package (install package): คลิก `install package` ใน package window แล้วพิมพ์ชื่อ package ที่ต้องการเช่น `Rcmdr` แล้วคลิก Install หรือพิมพ์ install.packages("Rcmdr") ในหน้าต่างคำสั่ง
- การใช้งาน package: เลือกกล่องเพื่อให้มีเครื่องหมายถูกหน้า package ที่เราต้องการเช่น `Rcmdr` หรือพิมพ์ `library("Rcmdr")` ในหน้าต่างคำสั่ง

A.5 การทำงานบน Working directory

Working directory เป็นโฟลเดอร์ที่เราจะทำงาน โดยโปรแกรม R จะหาข้อมูลตลอดจนโปรแกรมที่เราเขียนไว้จากโฟลเดอร์ดังกล่าว ตลอดจนข้อมูลและรูปภาพที่เราบันทึกไว้จะอยู่ในโฟลเดอร์นั้น ดังนั้นก่อนการเริ่มงานบน R เราควรกำหนดโฟลเดอร์ดังกล่าวด้วยคำสั่งใน command window `setwd("directoryname")` เช่น

```
> setwd("D:/Teaching/EC435/R")
```

หรือบน RStudio เมนูเลือก `Tools` → `Set Working Directory` → `Choose Directory`

A.6 การใช้งานเบื้องต้น

A.6.1 เครื่องคิดเลข

โปรแกรม RStudio เราสามารถคำนวณกระบวนการทางคณิตศาสตร์ได้ด้วยเครื่องหมายดังต่อไปนี้ โดยพิมพ์คำสั่งดังกล่าวในหน้าต่าง Console แล้ว Enter

```
> 2+3 # Addition
> 2-3 # Subtraction
> 2*3 # Multiplication
> 2/3 # Division
> 2^3 # 2 to the power of 3
> sqrt(3) # Square roots
> log(3) # Logarithms (to the base e)
```

A.6.2 พื้นที่ทำงาน (workspace)

เราสามารถเก็บผลการคำนวณไว้ในตัวแปรใดๆได้โดยใช้ "assignment operator" <- , หรือ = เช่น

```
> a <- 2 * 3
> 2 * a
[1] 12
```

A.6.3 สเกลาร์ เวกเตอร์ และเมทริกซ์

ข้อมูลตัวเลขใน R จะถูกบันทึกในรูปสเกลาร์ เวกเตอร์ หรือเมทริกซ์ เช่นในตัวอย่างที่ผ่านมา `a` คือสเกลาร์ที่มีค่าเท่ากับ 6 หรือเราสามารถสร้างเวกเตอร์ได้โดย ใช้ฟังก์ชัน `c` ดังนี้ เช่นเวกเตอร์ `b` ประกอบด้วยสมาชิกคือ 2,3,4

```
> b=c(2,3,4)
> b
[1] 2 3 4
```

A.6.4 ฟังก์ชัน

บางครั้งเราต้องการคำนวณค่าบางค่าที่ต้องใช้คำสั่งหลายบรรทัด บางครั้งเราสามารถใช้ฟังก์ชันที่มีอยู่แล้วในโปรแกรมหลักของ R หรือใน package ต่างๆ เช่น การหาค่าเฉลี่ยสามารถใช้ฟังก์ชัน `mean()` เช่นเราต้องการหาค่าเฉลี่ยของ 5,9,13,4 เราต้องรวมค่าทั้งหมดไว้ในเวกเตอร์และหาค่าเฉลี่ยดังนี้

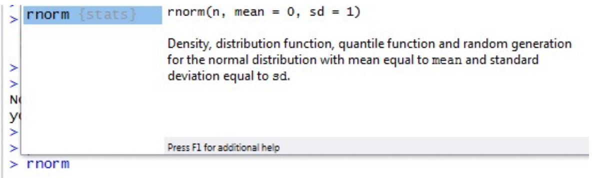
```
> x=c(5,9,13,4)
> mean(x)
[1] 7.75
```

โดยที่ค่าหรือตัวเลือกที่ต้องกำหนดในวงเล็บเรียกว่า *arguments* ตัวอย่างของฟังก์ชันอีกอันของฟังก์ชันคือฟังก์ชัน `rnorm` ที่ใช้ในการสร้างค่าทดลอง (generate) ค่าจากการแจกแจงแบบปกติ

```
> rnorm(10)
 [1]  0.60299 -0.75383  0.97095  1.06657  0.26545 -0.51994 -0.89247 -0.01616
 [9]  1.11192 -0.31573
```

RStudio มีส่วนประกอบที่ช่วยเหลือเกี่ยวกับ arguments ที่เราจะต้องใส่สำหรับแต่ละฟังก์ชันเราสามารถเขียนฟังก์ชันใน command window แล้วกดปุ่ม `Tab` จะมีตัวอย่างของ arguments ขึ้นมาให้เช่นรูป ??

รูปที่ A.2: Tab ช่วยเหลือใน RStudio



A.6.5 การวาดแผนภาพ

```
> x=rnorm(100)
> plot(x)
```

A.7 การช่วยเหลือและเอกสารประกอบฟังก์ชัน

เมื่อพิมพ์ฟังก์ชัน `help` จะมีเอกสารช่วยเหลือในหน้าต่าง help

```
> help(rnorm)
```

หรือเราสามารถใช้ฟังก์ชัน example เพื่อดูตัวอย่างคำสั่งและผลของฟังก์ชัน

```
> example(rnorm)
```

A.8 Scripts

นอกจากการพิมพ์คำสั่งใน command window แล้วเราสามารถประมวลผลโปรแกรมจาก script ซึ่งมีนามสกุล `.R` โดยการเลือกเปิดไฟล์จาก `File` → `Open File` แล้วเลือก Script ขึ้นมาแล้วคลิก `run` ในเมนู editor window

A.9 โครงสร้างข้อมูล

A.9.1 เวกเตอร์

เราสามารถสร้างเวกเตอร์ได้ด้วยฟังก์ชัน `c` เช่นคำสั่งในบรรทัดที่ 1 นอกจากนี้เราสามารถเรียกดูข้อมูลในตำแหน่งที่ i ด้วยคำสั่ง `vectorname[i]` เช่นบรรทัดที่ 4 นอกจากนี้เราสามารถสร้างเวกเตอร์ที่มีระยะเท่ากันด้วยฟังก์ชัน `seq` ดังเช่นในบรรทัดที่ 9

```
> vec1=c(2,4,6,8,10)
> vec1
[1]  2  4  6  8 10
> vec1[5]
[1] 10
> vec1[3]=12
> vec1
[1]  2  4 12  8 10
> vec2=seq(from=0,to=1,length=5)
> vec2
[1] 0.00 0.25 0.50 0.75 1.00
> vec1+vec2
[1]  2.00  4.25 12.50  8.75 11.00
```

A.9.2 เมตริกซ์

เราสามารถสร้างเมตริกซ์โดยการระบุข้อมูลและขนาดของเมตริกซ์ด้วย argument `nrow` หรือ `ncol` ซึ่งระบุจำนวนแถวหรือคอลัมน์ของเมตริกซ์

```
> m=matrix(data=c(1,2,3,4,5,6),ncol=3)
> m
     [,1] [,2] [,3]
[1,]    1    3    5
[2,]    2    4    6
```

arguments `data` ในฟังก์ชัน `matrix` กำหนดข้อมูลที่จะใส่ในเมตริกซ์โดยที่ `ncol` เป็นตัวกำหนดจำนวนคอลัมน์ของเมตริกซ์ โดยที่เราอาจเลือกกำหนด `nrow` ก็ได้

การเรียกข้อมูลแต่ละตำแหน่งสามารถทำได้โดยใช้ `[row,column]`

```
> m[1,2]
[1] 3
> m[1,]
[1] 1 3 5
> m[,2]
[1] 3 4
```

A.9.3 data frame

ข้อมูลอนุกรมเวลามักถูกจัดเก็บในรูปของ data frame โดยที่หัวคอลัมน์เป็นชื่อของคอลัมน์ ดังนั้นเราสามารถเรียกคอลัมน์นั้นๆมาใช้โดยไม่ต้องทราบตำแหน่งของข้อมูลเช่น

```
> t=data.frame(x=c(1,2,3),y=c(10,20,30),z=c(100,200,300))
> t
  x  y   z
1 1 10 100
2 2 20 200
3 3 30 300
> mean(t$y)
[1] 20
> mean(t[["z"]])
[1] 200
```

A.9.4 list

โครงสร้างของข้อมูลอีกแบบคือ list โดยที่ข้อมูลจะถูกเก็บเป็นคอลัมน์ซึ่งไม่ได้เรียงลำดับและไม่จำเป็นต้องมีความยาวเท่ากัน โดยที่ฟังก์ในการสร้าง list ก็คือ `list` นั่นเอง โดยที่ใน argument เราจะใส่ตัวแปรต่างๆลงไป โดยในที่ตัวแปร one เป็นตัวแปรสเกลาร์ที่มีค่าเท่ากับ 1 ตัวแปร two เป็นตัวแปรเวกเตอร์ และตัวแปร `five` เป็นเวกเตอร์ที่มีสมาชิกห้าตัว

```
> L=list(one=1, two=c(1,2), five=seq(1,4,length=5))
> L
$one
[1] 1
$two
[1] 1 2
$five
[1] 1.00 1.75 2.50 3.25 4.00
> L$five+10
[1] 11.00 11.75 12.50 13.25 14.00
```

A.10 Graphic

คำสั่งในการสร้างรูปภาพคือ `plot` โดยที่ argument ตัวแรกคือ ข้อมูลที่เราจะ plot ในที่นี้คือข้อมูลสุ่มที่สร้างขึ้นจากการแจกแจงแบบปกติจำนวน 100 ตัวอย่าง `rnorm(100)` สำหรับ argument ที่เหลือเป็นการกำหนดรูปและส่วนประกอบได้แก่ `type` คือประเภทของกราฟในที่นี้ ``l" แทนกราฟเส้น `col` คือสีของกราฟ `ylab` คือชื่อแกน Y `xlab` คือชื่อแกน X และ `main` คือชื่อกราฟ

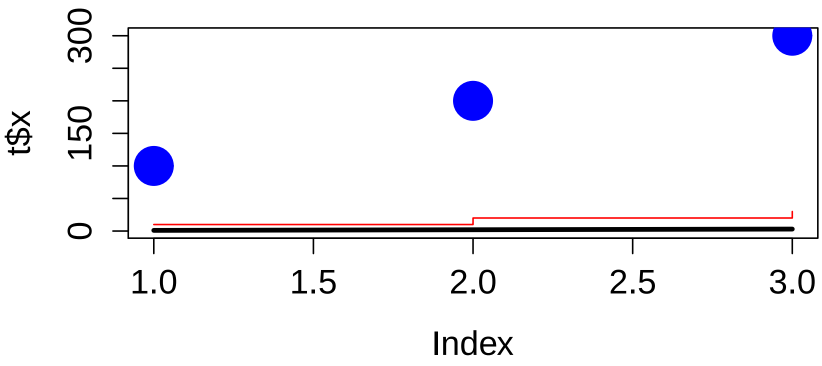
```
>
    plot(rnorm(100),type="l",col="blue",ylab="values",xlab="observations",main="Generated␣
    Data␣of␣N(0,1)")
```

หรือหากเราจะสร้างเส้นหลายเส้นสามารถทำได้โดยการเพิ่มเส้นทีละเส้น โดยในคำสั่งแรกจะต้องกำหนดระยะแกน Y ให้ครอบคลุมทั้งสามเส้นด้วยคำสั่ง `ylim`

```
> plot(t$x,type="l",ylim=range(t),lwd=3)
> lines(t$y,type="s",col="red")
> points(t$z,pch=20,cex=5,col="blue")
```

และจะได้รูปที่ **??**

รูปที่ A.3: ตัวอย่างการสร้างกราฟฟิก



A.11 การจัดเก็บข้อมูลและนำเข้าข้อมูล

A.11.1 การจัดเก็บข้อมูลในรูปไฟล์ text

เราสามารถนำเข้าข้อมูลโดยการสร้าง data frame และจัดเก็บในรูปไฟล์ text (นามสกุล .txt) โดยใช้คำสั่ง `write.table` ใน argument ประกอบด้วย ชื่อ data frame, ชื่อไฟล์ที่จัดเก็บ และชื่อของตัวแปร ซึ่งในที่นี้ได้กำหนดไว้แล้วใน data frame

```
> d=data.frame(a=c(3,4,5),b=c(12,13,15))
> d
  a  b
1 3 12
2 4 13
3 5 15
> write.table(d,file="ts0.txt",row.names=FALSE)
```

A.11.2 การเรียกข้อมูลในรูปไฟล์ text

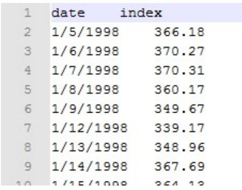
เราสามารถเรียกข้อมูลที่จัดเก็บในรูปไฟล์ text ด้วยฟังก์ชัน `write.table` โดย argument คือชื่อไฟล์และการระบุว่าไฟล์ดังกล่าวมีชื่อตัวแปรหรือไม่ซึ่งในที่นี้คือมีชื่อตัวแปร `header=TRUE`

```
> d2=read.table(file="ts0.txt",header=TRUE)
> d2
  a  b
1 3 12
2 4 13
3 5 15
```

ผู้อ่านสามารถทดลองนำเข้าข้อมูลไฟล์ text ชื่อว่า setindex.txt[1] เมื่อเปิดไฟล์ดังกล่าวด้วย

[1] download จาก Moodle ไปยัง workspace directory

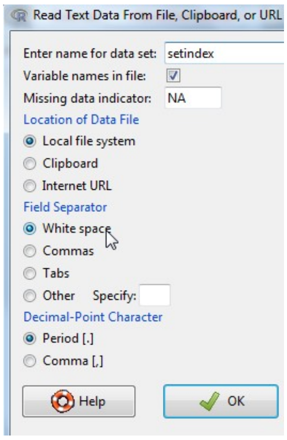
โปรแกรมเช่น Notepad จะเห็นได้ว่าข้อมูลมีการแบ่งด้วยช่องว่าง เราสามารถนำเข้าข้อมูลด้วย R



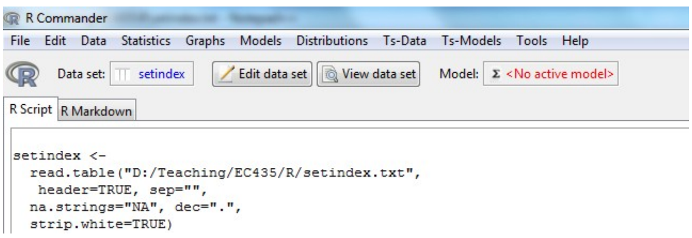
```
date      index
1/5/1998    366.18
1/6/1998    370.27
1/7/1998    370.31
1/8/1998    360.17
1/9/1998    349.67
1/12/1998   339.17
1/13/1998   348.96
1/14/1998   367.69
```

Commander โดยขั้นตอนต่อไปนี้

1. เลือก `Data/Import Data/From text file, clipboard, or URL` แล้วคลิก
2. จะได้กล่องให้ใส่เงื่อนไขต่างๆ ใส่ชื่อ Data set เป็น setindex และดูว่า มีการแบ่งข้อมูล(Field separator) เป็น white space ตามรูป



3. จะได้กล่องให้เราระบุตำแหน่งของไฟล์ เลือก Folder และ ไฟล์ที่เราต้องการ (ในที่นี้คือ setindex.txt) แล้ว Open
4. จะเห็นได้ว่าในพื้นที่ใต้เมนู จะมีชื่อ Dataset เป็น setindex ที่เรานำเข้ามา และเราสามารถเรียกดูข้อมูลได้โดยคลิก view data



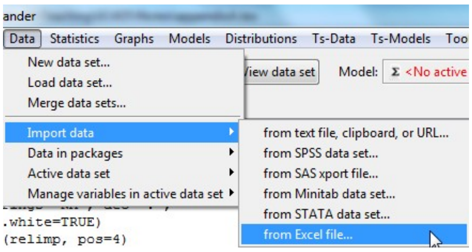
ไฟล์ text อีกประเภทหนึ่งที่เรามักจะเจอเวลาเราดาวน์โหลดข้อมูลคือ `.csv` โดยเราสามารถนำเข้าข้อมูลได้โดยใช้คำสั่ง `read.csv` เช่นในที่นี้ชื่อไฟล์ `pttstock.csv` เมื่อเปิดไฟล์ดังกล่าวด้วยโปรแกรมเช่น Notepad จะเห็นได้ว่าข้อมูลมีการแบ่งด้วย Comma เราสามารถนำเข้าข้อมูลด้วย R Commander โดยขั้นตอนต่อไปนี้

```
date,price
6/17/2003,65.5
6/18/2003,66
6/19/2003,67
6/20/2003,66.5
6/23/2003,68
6/24/2003,66.5
6/25/2003,66
6/26/2003,67.5
6/27/2003,66
6/30/2003,66.5
7/1/2003,66.5
```

1. เลือก `Data/Import Data/From text file, clipboard, or URL` แล้วคลิก
2. จะได้กล่องให้ใส่เงื่อนไขต่างๆ ใส่ชื่อ Data set เป็น `ptt` และดูว่า มีการแบ่งข้อมูล(Field separator) เป็น comma
3. จะได้กล่องให้เราระบุตำแหน่งของไฟล์ เลือก Folder และ ไฟล์ที่เราต้องการ (ในที่นี้คือ pttstock.csv) แล้ว Open
4. จะเห็นได้ว่าในพื้นที่ใต้เมนู จะมีชื่อ Dataset เป็น ptt ที่เรานำเข้ามา

A.11.3 การนำเข้าไฟล์ประเภทอื่นๆ

เราสามารถนำเข้าไฟล์ได้หลายประเภทเช่น SPSS, STATA และ Excel ได้ด้วยการเลือก Data/Import Data แล้วเลือกประเภทของไฟล์ที่ต้องการนำเข้าตามรูป



A.11.4 การบันทึกข้อมูล

เราสามารถบันทึกข้อมูลที่เรานำเข้าและสร้างตัวแปรใหม่ได้โดยการเลือก `Data/Active data set/Save active data set...` แล้วเลือก Folder ที่เราต้องการบันทึก และกำหนดชื่อ (ชนิดไฟล์ Rdata) แล้ว Save

ในการเปิดไฟล์ที่เราบันทึกครั้งต่อไปเราสามารถเลือก `Data/Load data set`

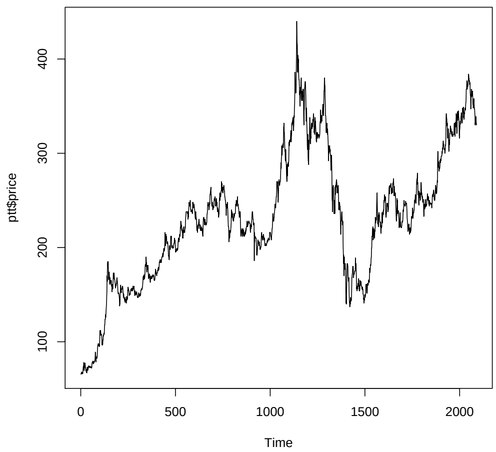
นอกจากนี้เราสามารถ Export ข้อมูลในรูปไฟล์ชนิดอื่นเช่น .txt หรือ .csv สำหรับนำเข้าในโปรแกรมอื่นได้ด้วยการเลือก `Data/Active data set/Export active data set...` เลือกวิธีการแบ่งคอลัมน์แล้วคลิก OK หลังจากนั้นกำหนดชื่อไฟล์และชนิดไฟล์ที่ต้องการ

A.12 การคำนวณผลได้ตอบแทนจากตัวอย่างหุ้น PTT

A.12.1 การวาดกราฟ

จากข้อมูล `ptt` เราสามารถวาดกราฟได้โดย เปลี่ยน active data set เป็น ptt โดยคลิกที่ Data set แล้วเลือก ptt

เราสามารถสร้างกราฟได้โดยการเลือก `Graphs` ในที่นี้ข้อมูลของเราเป็นอนุกรมเวลา เราจะเลือก `Graphs/epack- time series plots` จะได้รูปดังต่อไปนี้ เราสามารถบันทึกรูปไว้โดยเลือก



`File/Save as` แล้วเลือกประเภทไฟล์ที่ต้องการ หรือหากต้องการนำไปวางในโปรแกรมเอกสาร เช่น Word เราสามารถเลือก `File/Copy to the clipboard`

เนื่องจาก data frame ที่เรานำเข้ามายังไม่ได้ระบุตัวแปรเวลาสำหรับการเรียงข้อมูล เราจะใช้ package `zoo` ในการจัดการข้อมูลให้เรียงตามเวลาที่ระบุใน data frame ดังต่อไปนี้ โดยที่ตัวแปรใหม่จะถูกจัดเก็บใน data set `ptt` โดยการใส่ชื่อ `ptt.z` ต่อท้าย `ptt$`

```
> library(zoo)
> ptt$ptt.z<-zoo(x=ptt$price, order.by=as.Date(ptt$date,"%m/%d/%Y"))
> head(ptt$ptt.z)
2003-06-17 2003-06-18 2003-06-19 2003-06-20 2003-06-23 2003-06-24
      65.5       66.0       67.0       66.5       68.0       66.5
> plot(ptt$ptt.z,xlab="Year",ylab="Baht/share", main="PTT Stock Price")
```

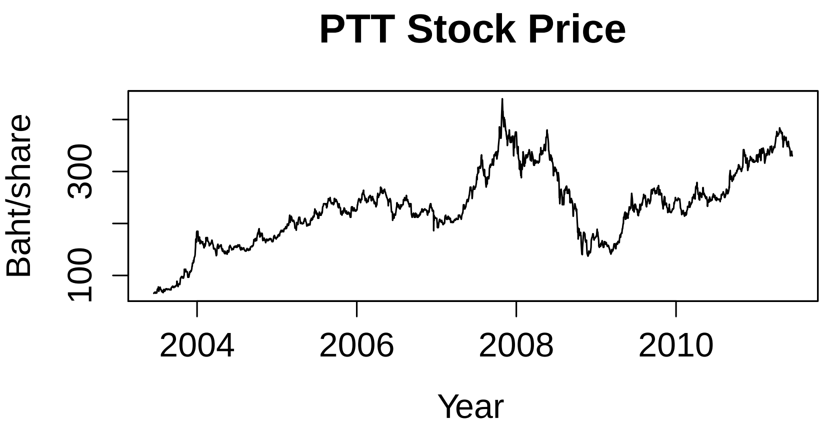
แล้วเราจะได้รูป **??**

A.12.2 การคำนวณผลได้ตอบแทน

การคำนวนผลได้ตอบแทนด้วยคำสั่งหลักใน R

เราสามารถคำนวณผลตอบแทนอย่างง่ายได้โดยนำราคาในตำแหน่งที่ i ลบด้วยตำแหน่งที่ i-1 แล้วหารด้วย i-1 โดยการเขียนคำสั่งในบรรทัดที่ 1 อย่างไรก็ตามผลลัพธ์ที่ได้จะมีจำนวนหายไป

รูปที่ A.4: ราคาหุ้น PTT



หนึ่งตำแหน่ง หากเราจะนำข้อมูลไปรวมใน data frame จะต้องมีการสมมุติให้ตำแหน่งแรกเป็นศูนย์ตามบรรทัดที่ 2

```
1 > n=length(ptt$price)
2 > sret=(ptt$price[2:n]−ptt$price[1:n−1])/(ptt$price[1:n−1])
```

หากเราต้องการรวมตัวแปรใน data frame เราต้องเพิ่มตัวอย่างแรกในตัวแปร `sret` เพื่อให้ความยาวเท่ากัน

```
> head(sret)
[1]  0.007634  0.015152 −0.007463  0.022556 −0.022059 −0.007519
> ptt$sret=c(NA,sret)
> head(ptt)
       date  price      sret
1 6/17/2003 65.5        NA
2 6/18/2003 66.0  0.007634
```

หรือสามารถคำนวณผลได้ตอบแทนในรูปของล็อกได้ด้วยคำสั่งในบรรทัดที่ 1 หรือบรรทัดที่สอง หรือแปลงค่าจากผลได้ตอบแทนอย่างง่าย

```
1  > lret=log(ptt$price[2:n])−log(ptt$price[1:n−1])
2  > lret2=diff(log(ptt$price))
3  > lret3=log(1+sret)
4  > head(lret)
5  [1]  0.007605  0.015038 −0.007491  0.022306 −0.022306 −0.007547
6  > ptt$lret=c(NA,lret)
7  > head(ptt)
8         date  price      sret      lret
9  1 6/17/2003 65.5        NA        NA
10 2 6/18/2003 66.0  0.007634  0.007605
11 3 6/19/2003 67.0  0.015152  0.015038
12 4 6/20/2003 66.5 −0.007463 −0.007491
13 5 6/23/2003 68.0  0.022556  0.022306
14 6 6/24/2003 66.5 −0.022059 −0.022306
```

เราสามารถบันทึก data frame ptt เป็นไฟล์ csv ด้วยคำสั่ง

```
> write.csv(ptt, file="pttreturn.csv", row.names=FALSE)
```

R package สำหรับคำนวณผลได้ตอบแทน

เรา สามารถ คำนวณ ผล ได้ ตอบแทน ได้ ด้วย package PerformanceAnalytics โดย-ที่ ตัวแปร นั้น ต้อง อยู่ ใน รูป zoo ซึ่ง ใน ที่ นี้ คือ ptt$pptz คำ สั่ง ที่ ใช้ ใน การ คำนวณ คือ Return.calculate(price, method) โดยที่ discrete คือ simple return

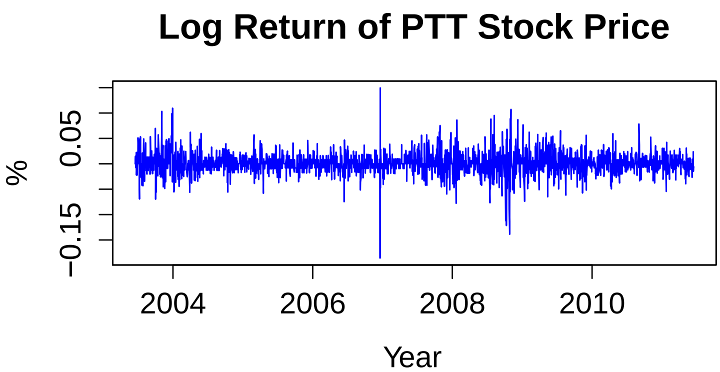
```
> ptt$sret2=Return.calculate(ptt$ptt.z, method = "discrete")
> ptt$lret2=Return.calculate(ptt$ptt.z, method = "log")
> head(ptt)
       date price ptt.z         sret         lret        sret2        lret2
1 6/17/2003  65.5  65.5           NA           NA           NA           NA
2 6/18/2003  66.0  66.0  0.007633588  0.007604599  0.007633588  0.007604599
3 6/19/2003  67.0  67.0  0.015151515  0.015037877  0.015151515  0.015037877
4 6/20/2003  66.5  66.5 -0.007462687 -0.007490672 -0.007462687 -0.007490672
5 6/23/2003  68.0  68.0  0.022556391  0.022305758  0.022556391  0.022305758
6 6/24/2003  66.5  66.5 -0.022058824 -0.022305758 -0.022058824 -0.022305758
```

เนื่องจาก ข้อมูล ใน แถว แรก เป็น NA เรา อาจ จะ ลบ ข้อมูล ใน แถว แรก โดย ที่ ในท R Commander menu เลือก **Data>Active data set> Remove row(s) from active data set** จะมีกล่องให้ระบุแถวที่ต้องการลบซึ่งในที่นี้คือ 1 หลังจากนั้น save ข้อมูลในไฟล์ชื่อ ptt.Rdata สำหรับการวิเคราะห์ทางสถิติต่อไป

แล้วสามารถวาดแผนภาพของผลได้ตอบแทนได้ในรูปล็อกได้ตามรูปที่ **??**

```
> plot(ptt$lret2,xlab="Year",ylab="%", main="Log Return of PTT Stock
    Price",col="blue")
```

รูปที่ A.5: ผลได้ตอบแทนจากราคาหุ้น PTT



บรรณานุกรม

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', Journal of econometrics 31(3), 307--327.

[0] [1] Brooks, C. (2008), Introductory econometrics for finance, Cambridge university press.

[2] Campbell, J. Y.; Lo, A. W.-C. MacKinlay, A. C. (1997), The econometrics of financial markets, princeton University press.

[3] Engle, R. F. Granger, C. W. (1987), 'Co-integration and error correction: representation, estimation, and testing', Econometrica: journal of the Econometric Society, 251--276.

[4] Engle, R. F.; Lilien, D. M. Robins, R. P. (1987), 'Estimating time varying risk premia in the term structure: the ARCH-M model', Econometrica: Journal of the Econometric Society, 391--407.

[5] Johansen, S. (1988), 'Statistical analysis of cointegration vectors', Journal of economic dynamics and control 12(2), 231--254.

[6] Kwiatkowski, D.; Phillips, P. C.; Schmidt, P. Shin, Y. (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?', Journal of econometrics 54(1), 159--178.

[7] Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: A new approach', Econometrica: Journal of the Econometric Society, 347--370.

[8] Phillips, P. C. Ouliaris, S. (1990), 'Asymptotic properties of residual based tests for cointegration', Econometrica: Journal of the Econometric Society, 165--193.

[9] Phillips, P. C. Perron, P. (1988), 'Testing for a unit root in time series regression', Biometrika 75(2), 335--346.

[10] Tsay, R. S. (2010), Analysis of financial time series, Vol. 543, Wiley-Interscience.

[11] Zivot, E. Wang, J. (2003), Modeling financial time series with S-PLUS, Vol. 191, Springer.