

# การเรียนรู้เหตุและผลแบบเกรนเจอร์สำหรับข้อมูลอนุกรมเวลาด้วยวิธีทดสอบแบบวาลด์ บนแบบจำลองที่มีเสถียรภาพ

## Exploring Granger Causality for Time series via Wald Tests on Estimated Models with Guaranteed Stability

นันทนัช รักษาศรี และ จิตโกมุท สังศิริ

ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

astronomer.nut@gmail.com, jitkomut.s@chula.ac.th

### บทคัดย่อ

บทความนี้จะอธิบายความสัมพันธ์ระหว่างตัวแปรของข้อมูลอนุกรมเวลาผ่านเงื่อนไขความสัมพันธ์เชิงเหตุและผลแบบเกรนเจอร์ (GC) บนแบบจำลองถดถอยตัวเอง และหารูปแบบความสัมพันธ์นั้นผ่านการทดสอบวาลด์ (Wald) บนพารามิเตอร์ของแบบจำลอง ผลลัพธ์ของบทความนี้คือ รูปแบบปัญหาการประมาณแบบจำลองด้วยวิธีกำลังสองต่ำสุด ที่รวมเงื่อนไข GC และเสถียรภาพของแบบจำลอง โดยเงื่อนไขเพียงพอสถียรภาพ อยู่ในรูปนอร์มแบบอนันต์ของเมทริกซ์พลวัตของแบบจำลองมีค่าน้อยกว่าหนึ่ง ปัญหาที่ได้จะอยู่ในรูปคอนเวกซ์แบบกำลังสอง และผลเฉลยของปัญหาคือแบบจำลองที่ гаранต์ว่ามีเสถียรภาพ อย่างไรก็ตาม แบบจำลองจะมีค่าความคลาดเคลื่อนที่มากขึ้นจากความจำเป็นเนื่องจากเงื่อนไขเสถียรภาพเป็นแค่เงื่อนไขเพียงพเท่านั้น

**คำสำคัญ:** เหตุและผลแบบเกรนเจอร์, แบบจำลองถดถอยตัวเอง, การทดสอบวาลด์

### Abstract

This paper aims to explain relationships between time series by using the Granger causality (GC) concept through autoregressive (AR) models and examine such relationship via the model parameters using the Wald test. The resulting formulation is a least-squares estimation with Granger causality and stability constraints. Our main contribution is to provide a formulation that guarantees the system stability or the stationarity of estimated AR process. A stability sufficient condition is added as the infinity norm constraint on the dynamic matrix. Estimated models are guaranteed to be stable but the model fitting error could be conservatively increased due to the selected stability condition.

**Keywords:** Granger causality, Autoregressive model, Wald test

### 1 บทนำ

ปัญหาหนึ่งในการอธิบายพลวัตของอนุกรมเวลา คือการอธิบายความสัมพันธ์ของตัวแปรต่างๆ ว่ามีผลมาจากตัวแปรอื่นใดบ้าง แบบจำลองถดถอยตัวเองหลายตัวแปร (Vector Autoregressive (VAR)) อันดับ  $p$  เป็นแบบจำลองหนึ่งที่นิยมถูกนำมาใช้อธิบายพลวัตของอนุกรมเวลาซึ่งบรรยายได้ด้วยสมการ

$$y(t) = A_1 y(t-1) + \dots + A_p y(t-p) + v(t) \quad (1)$$

เมื่อ  $y(t) = (y_1(t), \dots, y_n(t)) \in \mathbf{R}^n$  คือสัญญาณขาออก  $v(t)$  คือ สัญญาณรบกวนที่มีเมทริกซ์ความแปรปรวนร่วมเป็น  $\Sigma$  และ  $A_k \in \mathbf{R}^{n \times n}$  สำหรับ  $k = 1, \dots, p$  คือเมทริกซ์สัมประสิทธิ์ของแบบจำลอง การอธิบายความสัมพันธ์ของตัวแปรในข้อมูลอนุกรมเวลาหรือโครงสร้างเชิงสาเหตุสามารถอธิบายด้วยเงื่อนไขความสัมพันธ์ระหว่างเหตุและผล (Granger causality หรือ GC) โดยจาก [1] เงื่อนไข GC แบบจำลอง AR นั้นสามารถเขียนได้ในรูปสมการเชิงเส้นในพารามิเตอร์ของแบบจำลอง กล่าวคือ  $y_j$  ไม่ใช่สาเหตุแบบ Granger ต่อ  $y_i$  ก็ต่อเมื่อสำหรับทุกค่า  $k = 1, 2, \dots, p$

$$(A_k)_{ij} = 0 \quad (2)$$

เมื่อ  $(A_k)_{ij}$  คือสมาชิกตัวที่  $(i, j)$  ของเมทริกซ์  $A_k$  ซึ่งคือพารามิเตอร์ของแบบจำลอง AR

การประมาณค่าพารามิเตอร์ของแบบจำลองจะเลือกใช้วิธีการประมาณแบบความควรจะเป็นสูงสุด (maximum likelihood) หรือ ML ถ้าค่าประมาณของ  $A_k$  ในตำแหน่ง  $(i, j)$  ใดๆ มีค่าเท่ากับศูนย์ จะสามารถบอกได้ว่าข้อมูลอนุกรมเวลาในตำแหน่ง  $j$  กับ  $i$  นั้นไม่สัมพันธ์กัน แต่โดยปกติแล้วเมทริกซ์  $A_k$  ที่ประมาณได้อาจจะมีบางสมาชิกที่ไม่เท่ากับศูนย์ หรือมีค่าน้อยระดับหนึ่ง จึงต้องตั้งสมมติฐานขึ้นมาเพื่อทดสอบสมการ (2) ว่าเป็นจริงหรือไม่ กล่าวคือการทดสอบหารูปแบบโครงสร้างของค่าที่เป็นศูนย์ในพารามิเตอร์ของแบบจำลอง AR จะนำหลักการของการทดสอบทางสถิติแบบ Wald [2] มาใช้ โดยมีหลักการคือ ถ้าสมมติฐานว่าค่าที่แท้จริงของพารามิเตอร์มีค่าเท่ากับศูนย์เป็นจริง แล้วค่าพารามิเตอร์ที่ประมาณได้ควรมีค่าเข้าใกล้ศูนย์อย่างมีนัยยะสำคัญ จากนั้นจะสามารถประมาณค่าพารามิเตอร์ใหม่ได้ภายใต้ข้อกำหนดที่ว่าค่าของสมาชิกบางตัวของพารามิเตอร์มีค่าเป็นศูนย์ ซึ่งอยู่ในรูปแบบปัญหาการหาค่าที่เหมาะสมที่สุดภายใต้ข้อกำหนดค่าที่เป็นศูนย์ (optimization problem with zero constraints)

กระบวนการสุ่ม (1) จะเป็นกระบวนการคงที่ (stationary process) ก็ต่อเมื่อระบบเชิงเส้นที่สอดคล้องกันมีเสถียรภาพ แต่ค่าพารามิเตอร์ของแบบจำลองที่ประมาณได้จากวิธีหนึ่งๆ นั้นอาจส่งผลให้ระบบเชิงเส้นที่ประมาณได้ไม่มีเสถียรภาพ ดังนั้นจึงต้องมีการรวมเงื่อนไขเสถียรภาพเข้าไปในการประมาณพารามิเตอร์ เพื่อให้กระบวนการ (1) จากค่าพารามิเตอร์ที่ประมาณได้เป็นกระบวนการที่คงที่ ซึ่งทำให้น่าไปใช้เพื่อวัตถุประสงค์อื่น เช่นการทำนาย ได้ต่อไป

บทความนี้มีจุดประสงค์ที่จะศึกษารูปแบบปัญหาการประมาณแบบจำลอง AR ที่รวม 2 คุณลักษณะที่สำคัญคือ รูปแบบโครงสร้าง

ของศูนย์ในพารามิเตอร์ที่จะแสดงถึงความสัมพันธ์แบบ GC และแบบจำลองที่ได้จะมีเสถียรภาพ โดยในบทความนี้จะอธิบายการประมาณแบบจำลองด้วยวิธี ML การทดสอบ Wald และการรวมเงื่อนไขความสัมพันธ์ GC และเสถียรภาพ พร้อมทั้งแสดงผลการจำลองเชิงเลขด้วยวิธีดังกล่าว

## 2 การประมาณแบบจำลอง AR

เราสามารถเขียนแบบจำลอง (1) ในรูปแบบสมการเชิงเส้นในพารามิเตอร์ได้ดังนี้

$$y(t) = AH(t) + v(t) \quad (3)$$

เมื่อ  $y(t) = (y_1(t), \dots, y_n(t))$  เป็นสัญญาณขาออกของแบบจำลอง,  $v(t)$  เป็นสัญญาณรบกวนแบบเกาส์  $v(t) \sim \mathcal{N}(0, \Sigma)$ ,  $H(t) = [y(t-1)^T \ y(t-2)^T \ \dots \ y(t-p)^T]^T$  และค่าพารามิเตอร์ของแบบจำลองที่ต้องการประมาณคือ

$$A = [A_1 \ A_2 \ \dots \ A_p]$$

ในงานนี้เราใช้วิธีความควรจะเป็นสูงสุด (ML) มาใช้ประมาณค่าพารามิเตอร์ จากข้อมูล  $y(1), y(2), \dots, y(N)$  และ  $N$  เป็นจำนวนข้อมูลอนุกรมเวลาที่เก็บได้ รูปแบบของปัญหาจะเป็น

$$\underset{A, \Sigma}{\text{maximize}} \quad \frac{N-p}{2} \log \det \Sigma^{-1} - \frac{1}{2} \|L(Y - AH)\|_F^2$$

โดยที่  $L$  เป็นฟังก์ชันของ  $\Sigma^{-1}$  ด้วยความสัมพันธ์  $L^T L = \Sigma^{-1}$  และ  $Y$  กับ  $H$  เป็นพารามิเตอร์ของปัญหาที่ประกอบไปด้วยค่าของ  $y(t)$  ในเวลาต่างๆ (สามารถดูรายละเอียดได้ใน [3]) ผลเฉลยของ  $A, \Sigma$  ในรูปแบบปิด (closed-form) คือ

$$\hat{A} = YH^T(HH^T)^{-1}$$

$$\hat{\Sigma} = \frac{1}{N-p} \sum_{t=p+1}^N (y(t) - \hat{A}H(t))(y(t) - \hat{A}H(t))^T$$

โดยจะสามารถคำนวณค่า  $\hat{\Sigma}$  ได้เมื่อรู้ค่า  $\hat{A}$  แล้ว เป็นที่ทราบกันดีว่า รูปแบบของปัญหา ML นั้นสมมูลกันกับปัญหาค่ากำลังสองน้อยสุด (Least-Squares) หรือ LS

$$\underset{A}{\text{minimize}} \quad \|Y - AH\|_F^2 \quad (4)$$

ซึ่งเป็นจริงภายใต้เงื่อนไขที่ว่า  $v(t)$  ต้องเป็นสัญญาณรบกวนเกาส์ ในหัวข้อต่อไปเราจะพิจารณาปัญหา LS นี้เพื่อรวมเงื่อนไข GC และเงื่อนไขเสถียรภาพบนแบบจำลอง

## 3 การทดสอบwald สำหรับ Granger causality

การหาเรียนรู้ความสัมพันธ์แบบ GC คือการทดสอบสมมติฐานที่ว่า  $(\hat{A}_k)_{ij} = 0$  หรือไม่ เพื่อหารูปแบบโครงสร้างของศูนย์ในพารามิเตอร์ของแบบจำลองที่ถูกประมาณได้ ในการทดสอบนี้ เราสามารถทดสอบสำหรับแต่ละคู่  $(i, j)$  ของเมทริกซ์ AR โดยให้ฟังก์ชันข้อกำหนดคือ

$$r(\hat{\theta}) = ((\hat{A}_1)_{ij}, (\hat{A}_2)_{ij}, \dots, (\hat{A}_p)_{ij}) \quad (5)$$

เมื่อ  $\hat{\theta}$  คือค่าประมาณของพารามิเตอร์  $\theta$  และ  $\theta = \text{vec}(A)$  (การทำให้  $A$  เป็นเวกเตอร์) ในการทดสอบนี้ เรามีสสมมติฐานว่าง  $H_0$  คือ

$$H_0 : r(\hat{\theta}) = 0$$

จากนั้น ค่าทางสถิติของwald ( $W_{ij}$ ) สำหรับการทดสอบนี้สามารถจัดรูปได้เป็น [3]

$$W_{ij} = r(\hat{\theta})^T [\widehat{\text{Avar}}(\hat{\theta})_{ij}]^{-1} r(\hat{\theta})$$

โดยกำหนดให้  $\widehat{\text{Avar}}(\hat{\theta})_{ij}$  คือเมทริกซ์ย่อยขนาด  $p \times p$  ในแนวทแยงมุม (main diagonal block) ของ  $\widehat{\text{Avar}}(\hat{\theta})$  ที่สอดคล้องกับพารามิเตอร์  $r(\hat{\theta})$  (รายละเอียดดูได้ใน [3]) โดยค่าทางสถิติของwald ( $W_{ij}$ ) จะมีการแจกแจงเข้าสู่การแจกแจงแบบไคสแควร์ (Chi-square distribution) ที่มีองศาอิสระเท่ากับ  $p$

ในการทดสอบwaldนั้น กำหนดให้  $\chi_{\alpha, m}^2$  คือค่าวิกฤต (critical value) ที่ได้จากค่าระดับนัยสำคัญ  $\alpha$  จากนิยาม  $\alpha = \text{Prob}(W_{ij} > \chi_{\alpha, m}^2)$  โดยเราจะปฏิเสธ  $H_0$  ถ้า  $W_{ij} > \chi_{\alpha, m}^2$

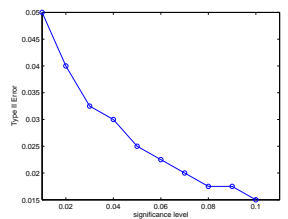
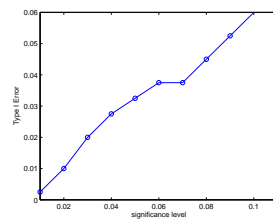
จากการทดสอบข้างต้นสำหรับที่ละตำแหน่ง  $(i, j)$  บนเมทริกซ์ AR ดังนั้น เราสามารถวนซ้ำการทดสอบดังกล่าว สำหรับทุกๆ  $j \neq i, i, j = 1, 2, \dots, n$  ดังนั้นเราสามารถกำหนดให้  $W = [W_{ij}]$  เป็นเมทริกซ์ทางสถิติของwald และเปรียบเทียบค่า  $W_{ij}$  กับค่าวิกฤตที่สอดคล้องได้ โดยแสดงผลพร้อมออกมาในรูปแบบของเมทริกซ์ที่มีค่าเป็น 0 หรือ 1 (binary matrix) ซึ่งมีความหมายคือ หากตำแหน่ง  $(i, j)$  ใดมีค่าเป็น 1 แสดงว่า  $y_j$  มีผลแบบ Granger ต่อตัวแปร  $y_i$  และถ้ามีค่าเป็น 0 เมื่อ  $y_j$  ไม่มีผลแบบ Granger ต่อตัวแปร  $y_i$

เพื่อทดสอบการเรียนรู้ GC บนข้อมูลที่จำลองขึ้น เราได้สร้างแบบจำลอง AR อันดับ  $n = 20$  และ  $p = 4$  โดยมีจำนวนศูนย์ในแบบจำลองจริงประมาณ 50% ของจำนวนพารามิเตอร์ทั้งหมด และจำลองข้อมูล  $y(t)$  มา 1,000 จุดทางเวลา เมื่อคำนวณตัวประมาณแบบ LS และทดสอบwald แล้ว ได้ผลลัพธ์ดังในรูป 2 ซึ่งจะเห็นว่าค่า  $\alpha$  จากนิยามนั้นสัมพันธ์โดยตรงกับ

$$\text{Type I Error} = \text{Prob}((\hat{A}_k)_{ij} \neq 0 \mid (A_k)_{ij} = 0)$$

นั่นคือ เมื่อ  $\alpha$  มีค่ามากขึ้น จะมีค่าผิดพลาดที่แสดงโดยวงกลมมากขึ้น ดังผลการจำลองในรูป 1 และ 2 นอกจากนี้ค่าความผิดพลาดแบบที่สอง (Type II error) คือค่าความน่าจะเป็นที่ค่าพารามิเตอร์ที่ประมาณได้เท่ากับศูนย์เมื่อกำหนดให้ค่าจริงของพารามิเตอร์ไม่เท่ากับศูนย์ ซึ่งจะมีค่าลดลงเมื่อ  $\alpha$  เพิ่มขึ้น

$$\text{Type II Error} = \text{Prob}((\hat{A}_k)_{ij} = 0 \mid (A_k)_{ij} \neq 0)$$



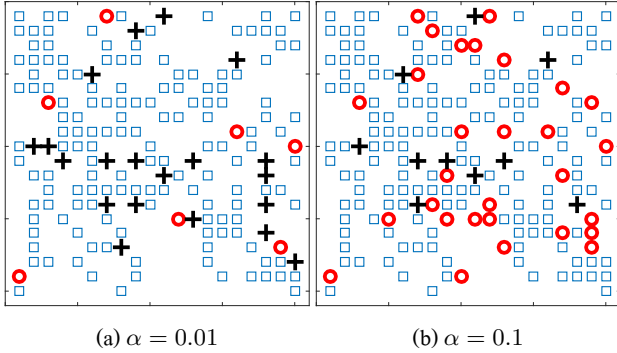
รูปที่ 1: รูปแสดงค่าความผิดพลาดทั้งสองแบบ เมื่อเปลี่ยนแปลงค่าระดับนัยสำคัญ ( $\alpha$ )

เมื่อเรียนรู้ GC จากข้อมูลแล้วนั้น รูปแบบศูนย์ของเมทริกซ์ AR ที่สอดคล้องกับ GC ดังกล่าวสามารถอธิบายผ่านเซตคู่อันดับ  $I$  (index set) ได้ ตัวอย่างเช่น  $y(t) = (y_1(t), y_2(t), y_3(t))$  และ GC ที่เรียนรู้ได้คือ  $y_1$  ไม่ส่งผลกับ  $y_2$  และ  $y_3$  ไม่ส่งผลกับ  $y_1$  จะได้ว่า  $I = \{(2, 1), (1, 3)\}$  และเราสามารถประมาณหาแบบจำลอง AR

ที่รวมเงื่อนไข GC ที่เรขียนไว้ได้นั้น จากการแก้ปัญหาค่าประมาณแบบ LS แบบมีเงื่อนไข ดังปัญหา

$$\begin{aligned} & \underset{A}{\text{minimize}} \quad \|Y - AH\|_F^2 \\ & \text{subject to} \quad (A_k)_{ij} = 0, \quad k = 1, \dots, p, \quad (i, j) \in I \end{aligned} \quad (6)$$

ความหมายของผลเฉลยจากปัญหา (6) คือเราจะได้แบบจำลอง AR ที่แสดงถึงความสัมพันธ์ของตัวแปรแบบ GC ที่อธิบายข้อมูลได้ดีที่สุด โดยที่ปัญหา (6) นั้นมีคำตอบในรูปแบบปิด สามารถคำนวณได้อย่างมีประสิทธิภาพเช่นกัน [4]



รูปที่ 2: รูปแบบโครงสร้างของศูนย์ของ AR พารามิเตอร์ที่ประมาณได้เมื่อเทียบกับค่าจริง เมื่อ □ แทนจุดข้อมูลที่ไม่เท่ากับศูนย์ทั้งคู่ของค่าจริงและค่าประมาณ ○ แทนจุดที่ค่าจริงเท่ากับศูนย์แต่ค่าประมาณไม่เป็นศูนย์ + แทนจุดที่ค่าจริงไม่เท่ากับศูนย์แต่ค่าประมาณเป็นศูนย์และช่องว่างคือจุดข้อมูลที่เป็นศูนย์ทั้งคู่ของค่าจริงและค่าประมาณ

### การรวมเงื่อนไขเสถียรภาพในการประมาณ

การประมาณแบบจำลองด้วยวิธี LS ไม่ว่าจะในปัญหา (4) หรือ (6) นั้นไม่มีสิ่งใดที่การันตีได้ว่า แบบจำลองที่ประมาณได้จะมีเสถียรภาพ เพื่อให้แบบจำลองที่ได้นั้นสอดคล้องกับคุณสมบัติทางกายภาพ เราจึงควรพิจารณาการรวมเงื่อนไขเสถียรภาพเข้าไปในการประมาณพารามิเตอร์ พิจารณา สมการเชิงเส้นเวลาวิฤตในรูปแบบทั่วไป

$$x(t+1) = Ax(t) + Bu(t) \quad (7)$$

เมื่อ  $x$  คือตัวแปรสถานะและ  $A$  คือเมทริกซ์พลวัต แบบจำลอง (1) เมื่อจัดให้อยู่ในรูปแบบ (7) จะมีเมทริกซ์พลวัตจะเป็น

$$A = \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \end{bmatrix}_{np \times np} \quad (8)$$

สำหรับเงื่อนไขจำเป็นและเพียงพอสำหรับเสถียรภาพ ระบบเชิงเส้นเวลาวิฤตจะมีเสถียรภาพก็ต่อเมื่อค่าเฉพาะ (eigenvalue) ทุกตัวของ  $A$  มีขนาดน้อยกว่าหนึ่ง นั่นคือ  $|\lambda_i(A)| < 1$  สำหรับ  $i = 1, 2, \dots, np$

สำหรับการประมาณพารามิเตอร์  $A$  ที่มีเสถียรภาพนั้น คือการหาค่า  $A_1, \dots, A_p$  ที่ทำให้ค่าเฉพาะของ  $A$  อยู่ในวงกลมหนึ่งหน่วย เงื่อนไขดังกล่าวอยู่ในรูปสมการไม่เชิงเส้นของพารามิเตอร์  $A$  และยากต่อการนำไปการคำนวณจริง

ในงานวิจัยที่ผ่านมา มีการทดลองรวมเงื่อนไขเสถียรภาพในการประมาณแบบจำลองด้วยวิธีต่างๆ เช่น การใช้ Jury's test บนระบบเชิงเส้นเวลาวิฤตแบบสัญญาณขาเข้าและสัญญาณขาออกเดียว (SISO) ใน [5] โดยจะหาค่าประมาณพารามิเตอร์ของระบบเชิงเส้นที่สอดคล้องกับตารางของ Jury เพื่อรับประกันว่าระบบเชิงเส้นจะมีเสถียรภาพแบบ BIBO แต่ในงานนี้ระบบที่เราสนใจเป็นระบบหลายสัญญาณขาเข้าและหลายสัญญาณขาออก (MIMO) การนำทฤษฎีบทเลียปูนอฟมาประยุกต์ใช้กับแบบจำลอง Autoregressive Moving Average (ARMA) แบบเวกเตอร์ จาก [6, 7] โดยจะประมาณค่าพารามิเตอร์ของแบบจำลอง และการแก้สมการเลียปูนอฟ สลับกันไป แต่อย่างไรก็ตามแนวทางนี้ไม่สามารถต่อยอดได้กับปัญหาในงานนี้ เนื่องจากเงื่อนไข GC ที่เราสนใจเพิ่มเติมใน [8] พิจารณาระบบ AR ที่มี  $p = 1$  และใช้เงื่อนไขขอบเขตบน 2-นอร์มของเมทริกซ์พลวัต นอกจากนี้มีการใช้ทฤษฎีของ Gershgorin circle บนแบบจำลอง ARMAX ใน [9] โดยเขียนแบบจำลองให้อยู่ในรูปของสมการสถานะ และหาค่าประมาณพารามิเตอร์ของแบบจำลองที่สอดคล้องกับการบังคับให้ Gershgorin circle นั้นอยู่ภายในวงกลมหนึ่งหน่วย เราจะแสดงให้เห็นว่าแนวทางที่เสนอขึ้นในบทความนี้เป็นกรณีทั่วไปของ [8] และจะนำไปสู่ผลลัพธ์ที่คล้ายกันกับ [9]

กำหนดให้ spectral radius ของ  $A$  คือ  $\rho(A)$  เราจะพบว่าเสถียรภาพของ  $A$  นั้นสมมูลกับเงื่อนไข  $\rho(A) < 1$  นอกจากนี้ พิจารณา induced norm ใดๆ ของ  $A$  ที่นิยามจาก  $\|A\| = \max_{x \neq 0} \|Ax\|/\|x\|$  เราสามารถแสดงให้เห็นได้ว่า  $\rho(A) \leq \|A\|$  เสมอ ด้วยเหตุนี้ เงื่อนไข  $\|A\| < 1$  จึงเป็นเงื่อนไขเพียงพอ สำหรับเสถียรภาพของ  $A$  เราได้พิจารณาใน [3] เงื่อนไข  $\|A\|_2 < 1$  หรือ  $\|A\|_1 < 1$  และพบว่าทั้งสองนอร์มนั้น เมื่อนำไปใช้กับ  $A$  ที่มีโครงสร้างดัง (8) แล้วนั้น จะนำไปสู่ผลเฉลยที่ไม่มีความหมาย

พิจารณาเงื่อนไข  $\|A\|_2 \leq 1$  จาก  $A$  ที่มีโครงสร้างดัง (8) กำหนดให้  $\bar{A} = [A_1 \ A_2 \ \dots \ A_{p-1}]$  และเนื่องจากเงื่อนไข  $\|A\|_2 \leq 1$  สมมูลกับ  $A^T A \preceq I$  จะได้ว่า

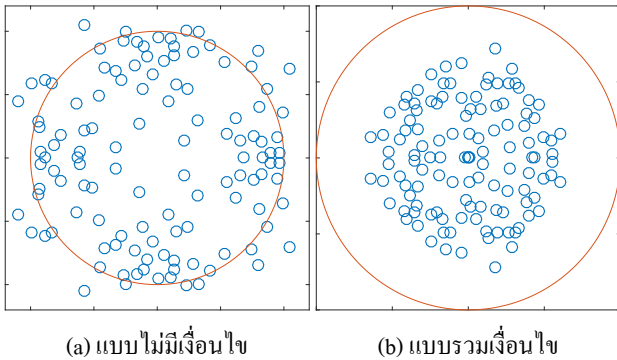
$$I - A^T A = \begin{bmatrix} -\bar{A}^T \bar{A} & -\bar{A}^T A_p \\ -A_p^T \bar{A} & I - A_p^T A_p \end{bmatrix} \succeq 0$$

จะเป็นจริงเมื่อ  $-\bar{A}^T \bar{A} \succeq 0$  ซึ่งเกิดขึ้นได้กรณีเดียวเมื่อ  $\bar{A} = 0$  หรือก็คือ  $A_1, \dots, A_{p-1}$  ต้องมีค่าเป็นศูนย์ทั้งหมด ซึ่งจะกลายเป็นผลตอบที่ไม่มีความหมาย เนื่องจากพารามิเตอร์ของแบบจำลอง  $A_1, \dots, A_p$  การมีรูปแบบโครงสร้างของศูนย์แบบเดียวกัน และจะได้ว่า  $y(t)$  ขึ้นกับ  $A_p$  อย่างเดียว ซึ่งไม่จำเป็น และในทำนองคล้ายกัน จากนิยาม  $\|A\|_1$  คือค่ามากที่สุดของผลรวมค่าสัมบูรณ์ของ  $A_{ij}$  ในคอลัมน์แต่ละคอลัมน์ และจะพบว่าแต่ละคอลัมน์ของ  $A$  ดัง (8) จะมี 1 อยู่เสมอ ยกเว้นคอลัมน์บล็อกที่มี  $A_p$  ดังนั้น ถ้าเงื่อนไข  $\|A\|_1 \leq 1$  เป็นจริง เราจะได้ว่า  $A_1, \dots, A_{p-1}$  ต้องมีค่าเป็นศูนย์ทั้งหมด ซึ่งจะกลายเป็นผลตอบที่ไม่มีความหมายเช่นกัน

ในบทความนี้เราเลือกใช้  $\|A\|_\infty < 1$  เป็นเงื่อนไขเพียงพอสำหรับเสถียรภาพ เนื่องจากเป็นเงื่อนไขแบบคอนเวกซ์ และเหมาะสมกับ  $A$  ใน (8) ที่จะนำไปสู่ผลเฉลยที่  $A_k$  ที่มีความหมาย กล่าวโดยสรุป เรานำเสนอรูปแบบปัญหาการประมาณแบบจำลอง AR ที่มีเสถียรภาพและสอดคล้องกับเงื่อนไข GC ดังนี้

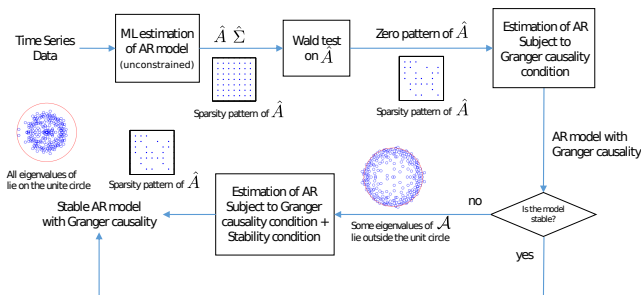
$$\begin{aligned} & \underset{A, \hat{A}}{\text{minimize}} && \|Y - AH\|_F^2 \\ & \text{subject to} && \|A\|_\infty \leq 1, \\ & && (A_k)_{ij} = 0, \quad (i, j) \in I \end{aligned} \quad (9)$$

เมื่อ  $I \subset \{1, 2, \dots, n\} \times \{1, 2, \dots, n\}$  เป็นเซตของคู่อันดับที่เป็นตัวกำหนดโครงสร้างของศูนย์ใน  $A_k$  และรูปแบบโครงสร้างของศูนย์ใน  $A_k$  นั้นหาได้จากการทดสอบwald นอกจากนั้น  $\hat{A}$  สัมพันธ์กับ  $A_k$  ดัง (8) ปัญหาที่เสนอขึ้นเป็นปัญหาคอนเวกซ์ จึงสามารถแก้หาค่าเฉลยเชิงเลขได้จากตัวแก้ปัญหามากมายวิธี ในงานนี้เราใช้ CVX [10] ในการเปรียบเทียบการทดลองประมาณแบบจำลองแบบ LS ที่ไม่มีเงื่อนไขใดๆ และการประมาณที่รวมเงื่อนไข GC และเสถียรภาพ โดยผลลัพธ์แสดงในรูป 3



รูปที่ 3: รูปแสดงตำแหน่งค่าจะของเมทริกซ์  $A$  (สีน้ำเงิน) บนระนาบเชิงซ้อนเมื่อเทียบกับวงกลมหนึ่งหน่วย (สีแดง) ที่คำนวณจากการประมาณแบบกำลังสองต่ำสุดในสองกรณี

จะเห็นว่าตำแหน่งของค่าจะของ  $A$  เมื่อรวมเงื่อนไขเสถียรภาพเข้าไปจะอยู่ภายในวงกลม 1 หน่วยทั้งหมด ซึ่งแสดงถึงการมีเสถียรภาพของระบบเชิงเส้นที่ประมาณได้ โดยจะเห็นว่าตำแหน่งค่าจะจะถูกบีบให้แคบลงมาก ซึ่งค่าฟังก์ชันจุดประสงค์ของ (9) จะเพิ่มมากขึ้นเมื่อมีการกำหนดเงื่อนไข GC และเสถียรภาพเข้าไป นั่นคือการบ่งบอกว่า การถ่วงน้ำหนักเสถียรภาพและโครงสร้าง GC นั้น ต้องแลกกับการทำให้แบบจำลองประมาณได้จะอธิบายข้อมูลที่วัดได้ด้วยความแม่นยำที่น้อยลง



รูปที่ 4: Diagram of stable AR model estimation with Granger causality.

## 4 บทสรุป

บทความนี้หาเงื่อนไขความสัมพันธ์เชิงเหตุและผลของพารามิเตอร์ของแบบจำลอง AR ผ่านรูปแบบโครงสร้างของศูนย์ มีจุดประสงค์เพื่ออธิบายความสัมพันธ์ระหว่างตัวแปรต่างๆ ของอนุกรมเวลาผ่านค่าพารามิเตอร์ของแบบจำลอง ซึ่งคำนึงถึงเสถียรภาพของแบบจำลองด้วย ในรูป 4 ได้สรุปขั้นตอนการประมาณแบบจำลอง ที่เริ่มจากการหาค่าประมาณเมทริกซ์ AR ด้วยวิธี ML (หรือ LS) ที่โดยปกติแล้ว มักจะมีค่าไม่เป็นศูนย์เกือบทั้งหมด เราจึงใช้การทดสอบแบบwaldเพื่อทดสอบสมมติฐานของ GC และผลลัพธ์คือรูปแบบของศูนย์ในเมทริกซ์ AR (รูปแบบของ GC) จากนั้น รูปแบบความสัมพันธ์นี้นำไปเป็นเงื่อนไขในการประมาณแบบจำลอง AR อีกครั้งหนึ่ง แบบจำลองที่ได้จะนำมาตรวจสอบความมีเสถียรภาพ ซึ่งถ้าหากไม่สอดคล้องกับเงื่อนไขดังกล่าวแล้ว เราสามารถประมาณหาแบบจำลองที่รวมเงื่อนไขเสถียรภาพ ที่อยู่ในรูปของนอร์มอันต์ของเมทริกซ์พลวัตของระบบ ผลลัพธ์สุดท้าย เราจะได้แบบจำลองที่มีเสถียรภาพและสอดคล้องกับเงื่อนไขความสัมพันธ์ระหว่างเหตุและผล และสามารถนำแบบจำลองไปใช้ในประโยชน์อื่นได้ต่อไป

## เอกสารอ้างอิง

- [1] J. Hamilton, *Time series analysis*, vol. 2. Princeton university press Princeton, 1994.
- [2] W. H. Greene, "Econometric analysis," *Upper Saddle River, NJ: Prentice Hall*, vol. 24, pp. 25–26, 2008.
- [3] N. Raksasri and J. Songsiri, "Exploring granger causality for time series via statistical tests on estimated models," tech. rep., Chulalongkorn University, 2016. Available in Thai at <http://jitkornut.lecturer.eng.chula.ac.th/group.html#current-students>.
- [4] A. Pongrattarakul, P. Lerdkultanon, and J. Songsiri, "Sparse system identification for discovering brain connectivity from fMRI time series," in *Proceedings of SICE Annual Conference*, pp. 949–954, 2013.
- [5] V. Cerone, D. Piga, and D. Regruto, "Bounding the parameters of linear systems with stability constraints," in *American Control Conference (ACC), 2010*, pp. 2152–2157, IEEE, 2010.
- [6] S. L. Lacy and D. S. Bernstein, "Subspace identification with guaranteed stability using constrained optimization," *Automatic Control, IEEE Transactions on*, vol. 48, no. 7, pp. 1259–1263, 2003.
- [7] J. Mari, P. Stoica, and T. McKelvey, "Vector ARMA estimation: A reliable subspace approach," *Signal Processing, IEEE Transactions on*, vol. 48, no. 7, pp. 2092–2104, 2000.
- [8] L. Buesing, J. H. Macke, and M. Sahani, "Learning stable, regularised latent models of neural population dynamics," *Network: Computation in Neural Systems*, vol. 23, no. 1-2, pp. 24–47, 2012.
- [9] K. Turksoy, E. S. Bayrak, L. Quinn, E. Littlejohn, and A. Cinar, "Guaranteed stability of recursive multi-input-single-output time series models," in *American Control Conference (ACC), 2013*, pp. 77–82, IEEE, 2013.
- [10] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1." <http://cvxr.com/cvx>, Mar. 2014.