

รายงานฉบับสมบูรณ์โครงการวิศวกรรมไฟฟ้า วิชา 2102499 ปีการศึกษา 2558

ชื่อโครงการ การเรียนรู้ความสัมพันธ์ของตัวแปรจากข้อมูล EEG

Project Title Learning dependence structure in EEG time series

ชื่อนิสิต นายปวีศร ใจทหาร หมายเลขประจำตัว 5531044621

กลุ่มวิจัย กลุ่มวิจัยการควบคุมขั้นสูงและการหาค่าเหมาะที่สุด

อาจารย์ที่ปรึกษา ผศ.ดร. จิตโกมุท ส่งศิริ ห้องปฏิบัติการวิจัย ระบบควบคุม

ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อ

โครงการนี้ต้องการที่จะอธิบายความสัมพันธ์ระหว่างตัวแปรต่างๆของข้อมูล EEG ผ่านเงื่อนไขความสัมพันธ์ระหว่างเหตุและผลแบบจำลองถดถอยตัวเอง ในการทดลองจะใช้เกณฑ์ข้อสนเทศของเบส์สำหรับเลือกอันดับที่เหมาะสมของแบบจำลองถดถอยตัวเอง และใช้การทดสอบ Wald ในการหาโครงสร้างความสัมพันธ์ระหว่างเหตุและผล จุดประสงค์ของโครงการคือหาโครงสร้างความสัมพันธ์ระหว่างเหตุและผลของสัญญาณคลื่นไฟฟ้าทางสมองที่ปรากฏในคนไข้ทุกคน โดยมีผลลัพธ์ที่ได้คือโครงสร้างความสัมพันธ์ระหว่างเหตุและผลของสัญญาณคลื่นไฟฟ้าทางสมองที่ปรากฏในผู้ป่วยสถานะชักเท่านั้น

คำสำคัญ: ข้อมูลอนุกรมเวลา, สัญญาณคลื่นไฟฟ้าทางสมอง, แบบจำลองถดถอยตัวเอง, เกณฑ์ข้อสนเทศของเบส์, การทดสอบ Wald, ความสัมพันธ์ระหว่างเหตุและผล

Abstract

This project describes the relationship of variables in EEG using Granger Causality on Autoregressive Model. Bayesian Information Criterion is used for choosing the appropriate order of Autoregressive Model. Granger Causality is defined by Wald test. The aim of this project is to define Granger Causality of EEG in patients. The result obtained from this project is the Granger Causality of EEG in seizure patients only.

Keywords: Time series, EEG, Autoregressive model, Bayesian Information Criterion, Wald test, Granger causality

สารบัญ

1	บทนำ	3
2	ทฤษฎีพื้นฐานที่ใช้ในโครงการ	4
2.1	การประมาณและการเลือกอันดับของแบบจำลองสำหรับสัญญาณ EEG	4
2.2	Granger Causality	6
3	ระเบียบวิธีที่ใช้ในการทำโครงการ	6
3.1	Conditional Granger Causality	6
3.2	การทดสอบ Granger Causality โดยวิธีการทางสถิติ	10
3.2.1	การทดสอบค่าศูนย์ของ AR coefficients ด้วยวิธี Wald	10
3.2.2	การทดสอบค่าศูนย์ของ Conditional Granger Causality โดยวิธีการทางสถิติ	11
3.2.3	การเปรียบเทียบค่า Granger Causality จากวิธีทางสถิติแบบวาลด์กับ Conditional Granger Causality.	11
3.2.4	ผลการหา Granger Causality โดยวิธีการทางสถิติแบบ Wald	13
3.3	การใช้ Granger Causality แยกแยะสถานะคนไข้	14
4	ผลการทดลองบนข้อมูล EEG	15
4.1	รายละเอียดของข้อมูล EEG	15
4.2	การประมาณแบบจำลอง AR ของข้อมูล EEG	16
4.3	โครงสร้างความสัมพันธ์ของสมองบนตำแหน่งต่างๆ ของผู้ป่วยในสองสถานะ	17
5	บทสรุป	21
6	กิตติกรรมประกาศ	21
7	ภาคผนวก	23
7.1	รายละเอียดของข้อมูล EEG	23
7.2	MATLAB codes	25

1 บทนำ

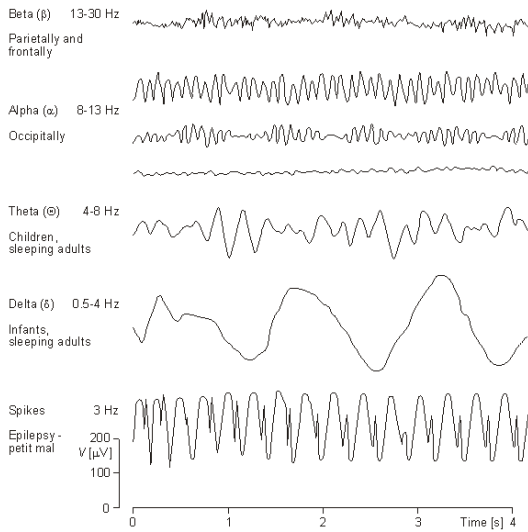
ในปัจจุบันการวิเคราะห์โรคทางสมองได้พัฒนาไปมาก หลังจากพัฒนาเครื่องวัดสัญญาณสมอง electroencephalogram หรือ EEG ซึ่งทำให้เห็นสัญญาณความต่างศักย์ของจุดต่างๆบนศีรษะ สัญญาณ EEG ที่วัดได้มาจากหลักการที่ว่า ระบบประสาทภายในร่างกายใช้กระแสไฟฟ้า ซึ่งเกิดจากการเคลื่อนที่ของโซเดียมไอออน (Na^+) และโพแทสเซียมไอออน (K^+) ผ่านเยื่อหุ้มเซลล์ประสาท ในการติดต่อสื่อสารกัน กระแสไฟฟ้านี้จะเหนี่ยวนำให้เกิดสนามไฟฟ้าขึ้น ก่อให้เกิดความต่างศักย์ระหว่างจุดต่างๆบนร่างกาย ความต่างศักย์ที่วัดค่าได้จากศีรษะเทียบกับตำแหน่งอ้างอิงที่สนใจ (นิยมใช้บริเวณที่มีประสาทร้อยเป็นจุดอ้างอิง เช่น ตีงหู) ก็คือสัญญาณ EEG [11] สามารถแบ่งสัญญาณ EEG ตามช่วงความถี่เป็น 5 ประเภท ได้แก่

- Delta (0.5-4 Hz) เป็นสัญญาณพื้นฐานของสมอง เกิดขึ้นเมื่อตอนหลับลึก สมองทำงานตามความจำเป็นเท่านั้น
- Theta (4-7.5 Hz) เป็นสัญญาณที่จะปรากฏขึ้นเมื่อมีสมาธิมากๆ ผ่อนคลายระดับลึก
- Alpha (8-13 Hz) เกิดเมื่อเราพักผ่อนและมีความสงบ แต่อยู่ในสภาวะที่รู้สึกตัว พบบ่อยในเด็กที่มีความสุข และผู้ใหญ่ที่ขบถนั่งสมาธิ โดยทั่วไปจะปรากฏเป็นสัญญาณไซน์
- Beta (14-26 Hz) เป็นสัญญาณสมองที่สื่อสารได้เร็วที่สุด เกิดเมื่อสมองมีการขบคิดอย่างตั้งใจ จดจ่อกับสิ่งภายนอก คลื่นนี้พบได้โดยทั่วไปในผู้ใหญ่ นอกจากนี้ที่ความถี่สูง มักเกิดตอนที่มีเหตุการณ์ที่ทำให้ตกใจ
- Spike มีความถี่ประมาณ 3 Hz มักเกิดเมื่อสมองปล่อยกระแสไฟฟ้าผิดปกติ ซึ่งเป็นอาการของผู้ป่วยโรคลมชัก

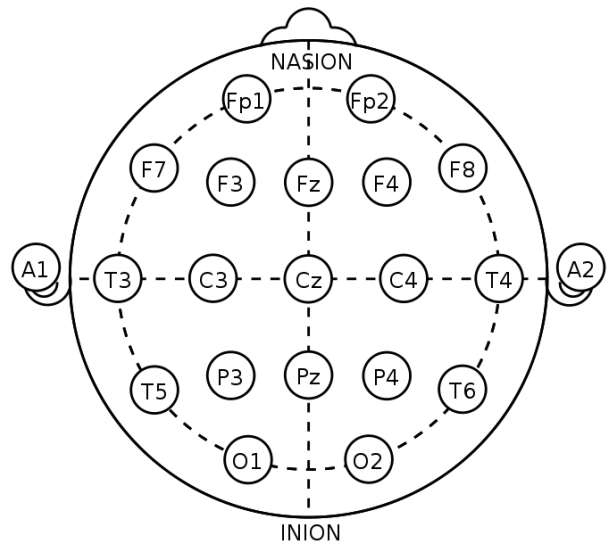
เครื่องวัด EEG มีส่วนประกอบหลักได้แก่ ขั้วไฟฟ้า (electrodes) ซึ่งประกอบไปด้วยขั้วบวก และขั้วลบ ทำหน้าที่ร่วมกันคือรับสัญญาณจากสมอง, เครื่องขยายสัญญาณ (amplifiers), ตัวกรอง (filter) ทำหน้าที่กรองสัญญาณส่วนเกินออก เช่น สัญญาณจากกล้ามเนื้อ ECG (Electrocardiography), ตัวแปลงสัญญาณแอนะล็อกเป็นดิจิทัล (ADC) [11] การวัดสัญญาณ EEG เริ่มจากวางขั้วไฟฟ้าไว้บนตำแหน่งต่างๆดังรูปที่ 2 โดยขั้วที่ตีงหู (A1 และ A2) คือขั้วที่ใช้เป็นจุดอ้างอิงในการอ่านค่าความต่างศักย์จากขั้วอื่นๆ [11]

จากข้อมูล EEG ที่วัดมาจากหลายตำแหน่งบนศีรษะ เราอยากทราบว่าข้อมูลแต่ละตำแหน่งเหล่านั้นมีความสัมพันธ์กันและสามารถอธิบายได้ด้วยหลักการทางคณิตศาสตร์หรือไม่ ซึ่งหลักการที่เราสนใจคือเงื่อนไขความสัมพันธ์ระหว่างเหตุและผล (Granger causality หรือ GC) [13] ที่มีหลักการว่าหากตัวแปร x มีผลต่อตัวแปร z แล้วการที่รู้ค่าตัวแปร x ก่อน จะช่วยในการทำนายค่าของตัวแปร z ในส่วนของโครงงานนี้เราสนใจพิจารณาโครงสร้างความสัมพันธ์ของสมองคนไข้ทั้งสองสภาวะ โดยมีสมมติฐานว่าแต่ละสภาวะจะมีโครงสร้างเฉพาะตัว ดังนั้นหลังจากหาแบบจำลองสำหรับสัญญาณ EEG และวิเคราะห์ความเหมาะสม

สมของแบบจำลองแล้ว เราจะวิเคราะห์โครงสร้างความสัมพันธ์ GC ของสัญญาณ EEG จากตำแหน่งต่างๆบนศีรษะของคนไข้แต่ละสภาวะว่าสอดคล้องกับสมมติฐานเพียงใด



รูป 1: ตัวอย่างสัญญาณ EEG ที่ความถี่ต่างๆ S. Sanei and J. Chambers, EEG Signal Processing. Wiley, 2008, Page 12



รูป 2: ตำแหน่งการวางขั้วไฟฟ้า แบบ 21 ขั้ว (channel) S. Sanei and J. Chambers, EEG Signal Processing. Wiley, 2008, Page 16

2 ทฤษฎีพื้นฐานที่ใช้ในโครงการ

2.1 การประมาณและการเลือกอันดับของแบบจำลองสำหรับสัญญาณ EEG

เมื่อพิจารณาสัญญาณ EEG ที่วัดมาได้ว่าเป็นสัญญาณต่อเนื่องทางเวลา ซึ่งค่าของสัญญาณในเวลาปัจจุบันขึ้นกับค่าของสัญญาณในอดีต พบว่าสามารถใช้แบบจำลองที่มีคุณสมบัติดังกล่าวทำนายสัญญาณ EEG ได้ดี จากการศึกษาการใช้แบบจำลองที่เหมาะสมในการอธิบายพลวัตของสัญญาณ EEG นั้น พบว่าแบบจำลองที่นิยมใช้ได้แก่ แบบจำลองถดถอยด้วยตัวเองค่าเฉลี่ยเคลื่อนที่ (autoregressive-moving-average หรือ ARMA) และแบบจำลองถดถอยด้วยตัวเอง (autoregressive หรือ AR) [11]

แบบจำลอง ARMA ที่มี n ตัวแปร อันดับของถดถอยด้วยตัวเอง p อันดับของค่าเฉลี่ยเคลื่อนที่ q สามารถแสดงได้ดังนี้

$$y(t) = \sum_{k=1}^p A_k y(t-k) + \sum_{j=1}^q B_j u(t-j) \quad (1)$$

โดยที่ $y(t) \in \mathbb{R}^n$, $A_k \in \mathbb{R}^{n \times n}$, $B_j \in \mathbb{R}^{n \times n}$, $k = 1, 2, \dots, p$, $j = 1, 2, \dots, q$ และ $u(t)$ คือสัญญาณรบกวน ณ เวลา t

แบบจำลอง AR ที่มี n ตัวแปร อันดับ p สามารถแสดงได้ดังนี้

$$y(t) = c + \sum_{k=1}^p A_k y(t-k) + u(t) \quad (2)$$

โดยที่ $y(t) \in \mathbb{R}^n$ คือเวกเตอร์ของสัญญาณที่เวลา t , $c \in \mathbb{R}^n$, $A_k \in \mathbb{R}^{n \times n}$, $k = 1, 2, \dots, p$ และ $u(t)$ คือสัญญาณรบกวน จาก (1) และ (2) จะเห็นว่าแบบจำลอง AR เป็นกรณีพิเศษของแบบจำลอง ARMA โดยแตกต่างกันที่สัญญาณของ ARMA ในปัจจุบันขึ้นกับสัญญาณรบกวนย้อนเข้าไปในอดีต q ค่าด้วย ส่วนสัญญาณของ AR ในปัจจุบันไม่ขึ้นกับสัญญาณย้อนเข้าไป

การทำให้ผลตอบจากแบบจำลองมีค่าความผิดพลาดน้อย และเรียบง่ายเป็นสิ่งที่เราต้องการ จากข้อมูลสัญญาณสมอง [1] เป็นข้อมูลอนุกรมเวลา กล่าวคือ ไม่มีสัญญาณเข้าแบบระบุได้ ดังนั้นในโครงการนี้เราได้เลือกแบบจำลอง AR สำหรับประมาณสัญญาณ EEG

วิธีการประมาณแบบจำลอง AR การประมาณแบบจำลอง ในการหาแบบจำลองถดถอยด้วยตัวเองของสัญญาณ EEG สำหรับแบบจำลอง AR ที่มีอันดับ p จาก (2) เขียนได้ว่า

$$y(t) = c + A_1 y(t-1) + A_2 y(t-2) + \dots + A_p y(t-p) + u(t) \quad (3)$$

โดยที่ $y(t) \in \mathbb{R}^n$ คือเวกเตอร์สัญญาณ EEG ที่เวลา t , $A_1, A_2, \dots, A_p \in \mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$ คือพารามิเตอร์ของแบบจำลอง, และ $u(t)$ คือสัญญาณรบกวนที่เวลา t จาก (3) จัดรูปใหม่ได้ว่า

$$y(t) = AH(t) + u(t) \quad (4)$$

โดยที่ $A = [c \ A_1 \ A_2 \ \dots \ A_p] \in \mathbb{R}^{n \times (np+1)}$ และ

$$H(t) = [1 \ y(t-1) \ y(t-2) \ \dots \ y(t-p)]^T \in \mathbb{R}^{np+1}$$

ตามที่ [8] กล่าวไว้ จาก (4) เราสามารถประมาณพารามิเตอร์ของแบบจำลองด้วยการประมาณค่าสูงสุดของค่าควรจะเป็น (maximum likelihood) โดยปัญหาดังกล่าวจะสมมูลกับปัญหากำลังสองต่ำสุด (least square) minimize $\|L[Y - AH]\|_F^2$ และมีผลเฉลยเป็น

$$\hat{A} = (YH^T)(HH^T)^{-1} \quad (5)$$

โดยที่

$$Y = [y(p+1) \ y(p+2) \ \dots \ y(N)] \in \mathbb{R}^{n \times (N-p)}$$

$$H = \begin{bmatrix} 1 & 1 & \dots & 1 \\ y(p) & y(p+1) & \dots & y(N-1) \\ y(p-1) & y(p) & \dots & y(N-2) \\ \vdots & \vdots & & \vdots \\ y(1) & y(2) & \dots & y(N-p) \end{bmatrix} \in \mathbb{R}^{(np+1) \times (N-p)}$$

สำหรับพารามิเตอร์ $\hat{\Sigma}$ มีผลเฉลยดังนี้

$$\hat{\Sigma} = \frac{1}{N-p} \sum_{t=p+1}^N (Y - \hat{A}H)(Y - \hat{A}H)^T \quad (6)$$

ดังนั้นผลการประมาณแบบจำลอง AR สามารถเขียนได้เป็น

$$\hat{Y} = \hat{A}H \quad (7)$$

โดยที่ $\hat{Y} = [\hat{y}(p+1) \ \hat{y}(p+2) \ \dots \ \hat{y}(N)] \in \mathbb{R}^{n \times (N-p)}$ คือ ผลตอบของแบบจำลอง AR ที่ประมาณได้จากสัญญาณ EEG เพื่อให้สะดวกกับการทดลองกับข้อมูลจริง เราจึงสร้างโปรแกรม MATLAB สำหรับหาแบบจำลอง AR ขึ้น จะได้โปรแกรกดัง ARestLS.m ในหัวข้อที่ 7.2

การเลือกอันดับของแบบจำลอง หลังจากได้แบบจำลองมาแล้ว ปัญหาถัดมาคือการเลือกค่าอันดับ p ของแบบจำลอง AR ที่จะทำให้แบบจำลองที่ได้มีค่าความคลาดเคลื่อนต่ำ และสามารถอธิบายลักษณะโครงสร้างเชิงสาเหตุได้อย่างสมเหตุสมผล โดยจะเห็นว่า การลดค่าอันดับของแบบจำลอง AR (p) จะทำให้แบบจำลองมีความซับซ้อนน้อยลง แต่อาจทำให้ค่าความผิดพลาดของแบบจำลองมีค่ามากขึ้น ในทางกลับกัน หากเพิ่มค่า p จะทำให้ค่าความผิดพลาดของแบบจำลองมีค่าลดลง แต่จะทำให้แบบจำลองมีความซับซ้อนมากขึ้น ดังนั้นเพื่อหาค่า p ที่เหมาะสม เราจะใช้เกณฑ์การเลือกแบบจำลอง Bayesian Information Criterion (BIC) ซึ่งค่าคะแนน BIC มีนิยามคือ

$$\text{BIC score} = -2\mathcal{L} + d \log N \quad (8)$$

โดย $\mathcal{L}(A, \Sigma) = \frac{N-p}{2} \log \det \Sigma^{-1} - \frac{1}{2} \|L(Y - AH)\|_F^2$, $d = n(np + 1)$ และ N คือจำนวนข้อมูลตัวอย่าง (อ้างอิงการคำนวณค่า \mathcal{L} , d และ N จาก [8]) โดยเราจะเลือกอันดับที่ BIC มีค่าต่ำสุดเป็นอันดับของแบบจำลอง AR สำหรับสัญญาณ EEG

2.2 Granger Causality

สำหรับแบบจำลอง AR ที่มีสัญญาณ n ตำแหน่ง อันดับ p ที่สามารถแสดงได้ดัง (2) นั้น ใน [13] ได้แสดงให้เห็นว่าเงื่อนไข GC ของแบบจำลอง AR นั้น กล่าวได้ว่า ถ้า y_j ไม่ใช่สาเหตุแบบ Granger ต่อ y_i แล้ว

$$(A_k)_{ij} = 0 \quad (9)$$

ทุกค่า $k = 1, 2, \dots, p$ โดยที่ $(A_k)_{ij}$ คือ สมาชิกตัวที่ (i, j) ของเมทริกซ์ A_k ดังนั้นหลังจากได้แบบจำลอง AR อันดับ p สำหรับสัญญาณ EEG แล้ว เราสามารถวิเคราะห์โครงสร้างความสัมพันธ์ GC ของข้อมูล EEG ได้จากตำแหน่งค่าศูนย์ในเมทริกซ์ A_k

3 ระเบียบวิธีที่ใช้ในการทำโครงงาน

3.1 Conditional Granger Causality

นักประสาทวิทยาได้ทำการศึกษาความสัมพันธ์ระหว่างบริเวณต่างๆ และพลวัตขณะเกิดปฏิสัมพันธ์ภายในสมองจากสัญญาณ EEG [14] ที่วัดมาหลาย channel โดยนิยมใช้แบบจำลอง AR [10, 9] ในการประมาณสัญญาณ ซึ่ง Granger Causality เป็นเทคนิคที่ดีในการตรวจสอบความสัมพันธ์เชิงสาเหตุของสัญญาณ

ทางเวลา เทคนิคที่นิยมใช้หาค่า Granger Causality ในวงการประสาทวิทยามี 2 วิธี ได้แก่ Pairwise Granger causality และ Condition Granger causality [3]

Pairwise Granger Causality. Pairwise GC เป็นเทคนิคที่นิยมใช้ในการตรวจสอบความสัมพันธ์ระหว่างสัญญาณ 2 กลุ่ม ข้อดีของวิธีนี้คือ สามารถหาค่าตามผลเฉลยได้ง่าย และให้คำตอบที่ครบถ้วนชัดเจน สำหรับคู่สัญญาณสเกลาร์ $x(t) \in \mathbb{R}$ และ $y(t) \in \mathbb{R}$ ที่ใช้แบบจำลอง AR ในการประมาณสัญญาณจะอธิบายสัญญาณได้ดังนี้

$$x(t) = \sum_{k=1}^p a_{1k}x(t-k) + \epsilon_1(t), \quad \text{var}(\epsilon_1(t)) = \Sigma_1 \quad (10)$$

$$y(t) = \sum_{k=1}^p c_{1k}y(t-k) + \nu_1(t), \quad \text{var}(\nu_1(t)) = \Gamma_1$$

ถ้าสัญญาณสองกลุ่มนี้มีความสัมพันธ์เชิงสาเหตุต่อกันสามารถเขียนได้ว่า

$$x(t) = \sum_{k=1}^p a_{2k}x(t-k) + \sum_{k=1}^p b_{2k}y(t-k) + \epsilon_2(t) \quad (11)$$

$$y(t) = \sum_{k=1}^p c_{2k}x(t-k) + \sum_{k=1}^p d_{2k}y(t-k) + \nu_2(t) \quad (12)$$

จาก (11), (12) กำหนดเมทริกซ์ความแปรปรวนร่วมดังนี้

$$\Sigma = \begin{bmatrix} \Sigma_2 & \Upsilon_2 \\ \Upsilon_2 & \Gamma_2 \end{bmatrix}$$

เมื่อให้ $\Sigma_2 = \text{var}(\epsilon_2(t))$, $\Gamma_2 = \text{var}(\nu_2(t))$, $\Upsilon_2 = \text{cov}(\epsilon_2(t), \nu_2(t))$ ถ้า $x(t)$ และ $y(t)$ เป็นอิสระต่อกันแล้ว b_2 และ c_2 จะเป็นศูนย์, $\Upsilon_2 = 0$

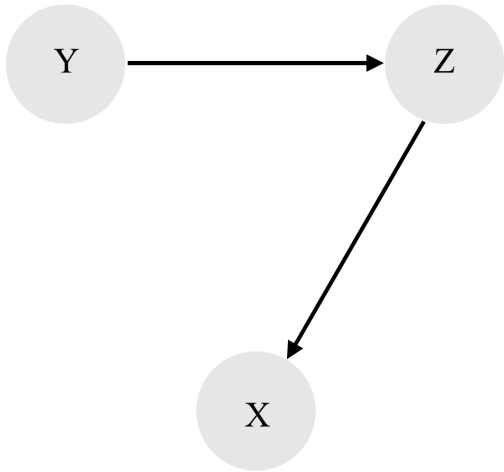
ตามที่ Granger [5] กล่าวเอาไว้ เราจะหาปริมาณที่ $y(t)$ มีผลกระทบเชิงสาเหตุต่อ $x(t)$ ได้จาก

$$F_{y \rightarrow x} = \log \frac{\Sigma_1}{\Sigma_2} \quad (13)$$

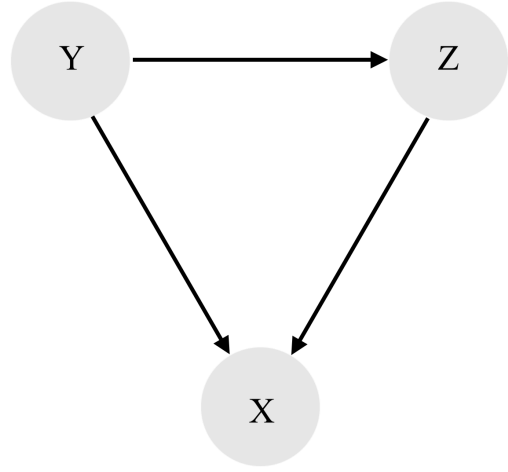
โดยที่ $F_{y \rightarrow x} = 0$ เมื่อ $y(t)$ ไม่มีผลต่อ $x(t)$ เลย และ $F_{y \rightarrow x} > 0$ เมื่อ $y(t)$ มีผลกระทบเชิงสาเหตุต่อ $x(t)$ ทั้งนี้จาก (10) และ (11) จะเห็นว่าการรวม $y(t)$ เข้าไปในแบบจำลอง (11) ส่งผลให้ค่าของ Σ_1 ลดลงเป็น Σ_2 ได้ว่า $\Sigma_2 < \Sigma_1$ แต่ถ้า $b_{2k} = 0$ จะส่งผลให้ (11) ลดรูปลงเหมือน (10) ทำให้ $\Sigma_2 = \Sigma_1$ ดังนั้นจะได้ $\Sigma_2 \leq \Sigma_1$ ทำให้ $F_{y \rightarrow x}$ มีค่ามากกว่า 0 เสมอ

อนึ่ง Pairwise GC เป็นเทคนิคที่มีข้อเสียคือหากมีข้อมูลตั้งแต่ 3 ชุดขึ้นไป Pairwise GC จะมีข้อจำกัดในการอธิบายความสัมพันธ์ของข้อมูลอยู่ จากรูปที่ 3 และ 4 เป็นรูปที่แสดงความสัมพันธ์เชิงเหตุผล

ของสัญญาณ $y(t) \in \mathbb{R}$ และ $z(t) \in \mathbb{R}$ ที่มีผลต่อ $x(t) \in \mathbb{R}$ สำหรับเทคนิค Pairwise GC เมื่อหาค่าความสัมพันธ์เชิงสาเหตุที่ $y(t)$ ส่งผลต่อ $x(t)$ แม้ว่า $y(t)$ จะไม่มีผลทางตรงต่อ $x(t)$ แต่มีผลทางอ้อมผ่าน $z(t)$ แบบรูปที่ 3 ก็จะทำให้ค่าที่บ่งบอกว่า $y(t)$ มีผลทางตรงต่อ $x(t)$ แบบรูปที่ 4 ซึ่งไม่ถูกต้อง



รูป 3: ผลของ $y(t)$ ที่มีต่อ $x(t)$ ทางอ้อมผ่าน $z(t)$



รูป 4: ผลของ $y(t)$ ที่มีต่อ $x(t)$ ทั้งทางตรงและทางอ้อม

ดังนั้นหากข้อมูลมีมากกว่า 3 กลุ่มขึ้นไป Condition GC จึงเป็นเทคนิคที่นิยมใช้ในการหาค่า GC โดยสามารถแยกความแตกต่างของความสัมพันธ์เชิงเหตุผลแบบในรูปที่ 3 และ 4 ได้

Conditional Granger Causality. ในการหาค่า Condition GC สำหรับกลุ่มข้อมูล $x(t)$, $y(t)$ และ $z(t)$ เมื่อต้องการหาว่า $y(t)$ มีผลต่อ $x(t)$ หรือไม่ ต้องพิจารณาเพิ่มด้วยว่า $y(t)$ มีผลต่อ $x(t)$ ทางตรง หรือมีผลทางอ้อมผ่าน $z(t)$ จาก (11) หาก $x(t)$ และ $z(t)$ มีผลต่อกันแล้วจะได้ว่า

$$x(t) = \sum_{k=1}^p a_{3k}x(t-k) + \sum_{k=1}^p b_{3k}z(t-k) + \epsilon_3(t)$$

$$z(t) = \sum_{k=1}^p c_{3k}x(t-k) + \sum_{k=1}^p d_{3k}z(t-k) + \gamma_3(t)$$

ได้เมทริกซ์ความแปรปรวนร่วมดังนี้

$$\Sigma_3 = \begin{bmatrix} \Sigma_3 & \Upsilon_3 \\ \Upsilon_3 & \Gamma_3 \end{bmatrix}$$

เนื่องจากสัญญาณสามกลุ่มนี้มีความสัมพันธ์เชิงสาเหตุต่อกันสามารถเขียนได้ว่า

$$\begin{aligned}x(t) &= \sum_{k=1}^p a_{4k}x(t-k) + \sum_{k=1}^p b_{4k}y(t-k) + \sum_{k=1}^p c_{4k}z(t-k) + \epsilon_4(t) \\y(t) &= \sum_{k=1}^p d_{4k}x(t-k) + \sum_{k=1}^p e_{4k}y(t-k) + \sum_{k=1}^p g_{4k}z(t-k) + \nu_4(t) \\z(t) &= \sum_{k=1}^p u_{4k}x(t-k) + \sum_{k=1}^p v_{2k}y(t-k) + \sum_{k=1}^p w_{4k}z(t-k) + \gamma_4(t)\end{aligned}$$

ได้เมทริกซ์ความแปรปรวนร่วมดังนี้

$$\Sigma_4 = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} & \Sigma_{xz} \\ \Sigma_{yx} & \Sigma_{yy} & \Sigma_{yz} \\ \Sigma_{zx} & \Sigma_{zy} & \Sigma_{zz} \end{bmatrix}$$

จากผลข้างต้นเราจะหาหาปริมาณที่ $y(t)$ มีผลกระทบต่อเชิงสาเหตุต่อ $x(t)$ โดยกำจัดตัวแปร $z(t)$ ได้จาก

$$F_{y \rightarrow x|z} = \log \frac{\Sigma_3}{\Sigma_{xx}}$$

เมื่อ Σ_3 คือความแปรปรวนร่วมที่มาจาก residual error ของแบบจำลองที่ไม่รวมผลของ $y(t)$ โดยที่ $F_{y \rightarrow x|z} = 0$ เมื่อความสัมพันธ์เป็นแบบรูปที่ 3 กล่าวคือ $\Sigma_{xx} = \Sigma_3$ ซึ่งหมายความว่ากรรวม $y(t)$ เข้าไปในแบบจำลองไม่ส่งผลต่อความแปรปรวนร่วม ได้ว่า $y(t)$ ไม่ส่งผลทางตรงต่อ $x(t)$ แต่หากความสัมพันธ์เป็นแบบรูปที่ 4 จะได้ $F_{y \rightarrow x|z} > 0$ กล่าวคือ $\Sigma_{xx} < \Sigma_3$ ซึ่งหมายความว่ากรรวม $y(t)$ เข้าไปในแบบจำลองส่งผลให้ความแปรปรวนร่วมน้อยลง ได้ว่า $y(t)$ ส่งผลทางตรงต่อ $x(t)$ ด้วย

Conditional Granger Causality for Multivariate Variables. จากตัวอย่างการหาค่า Condition GC จากกลุ่มข้อมูล 3 ตัวแปรข้างต้น สามารถขยายออกไปเป็นการหา Condition GC ของกลุ่มข้อมูล n ตัวแปรได้ดังนี้ ให้ $X(t) = [x_1(t) \ x_2(t) \ \dots \ x_n(t)]^T$ เราสนใจหาว่า x_j มีผลต่อ x_i แบบ Condition GC หรือไม่ และจะนิยามตัวแปร $Z(t)$ คือตัวแปร $X(t)$ ที่สลับ x_i และ x_j มาเป็นสมาชิก 2 แฉวบนสุด

เพื่อให้่ายต่อการอธิบายเราจะยกตัวอย่างกลุ่มข้อมูล 6 ตัวแปร คือ $X(t) = [x_1(t) \ x_2(t) \ \dots \ x_6(t)]^T$ ซึ่งเราสนใจหาว่า x_5 มีผลต่อ x_3 หรือไม่ เราจะให้ $Z(t) = [x_3(t) \ x_5(t) \ x_1(t) \ x_2(t) \ x_4(t) \ x_6(t)]^T$ คือเวกเตอร์ $X(t)$ ที่สลับ $x_3(t)$ มาไว้แฉวบนสุด และ $x_5(t)$ มาไว้แฉวที่สอง ส่วนแฉวที่ 3 จนถึง n คือข้อมูลสัญญาณอื่นๆที่เหลือ นอกจากนี้เรายังกำหนดให้ $\tilde{Z}(t)$ คือ เวกเตอร์ $Z(t)$ ที่ตัดแฉวที่ตรงกับ x_j ออก ในกรณีนี้คือแฉวที่ 2 จะได้ $\tilde{Z}(t) = [x_3(t) \ x_1(t) \ x_2(t) \ x_4(t) \ x_6(t)]^T$

ดังนั้นจากนิยามตัวแปรดังข้างต้นจะได้ว่าแบบจำลองที่รวมพลวัตของทุกตัวแปรคือ

$$Z(t) = \sum_{k=1}^p A_k Z(t-k) + \epsilon(t)$$

โดยที่ $Z(t) \in \mathbb{R}^n$ จะมี $\Sigma \in \mathbb{R}^{n \times n}$ เป็นเมทริกซ์ความแปรปรวนร่วมของ $\epsilon(t)$ ส่วนแบบจำลองที่ตัดผลของ $x_j(t)$ ออกไปจะได้

$$\tilde{Z}(t) = \sum_{k=1}^p B_k \tilde{Z}(t-k) + \gamma(t)$$

สังเกตว่า $\tilde{Z}(t) \in \mathbb{R}^{n-1}$ จะมี $\tilde{\Sigma} \in \mathbb{R}^{(n-1) \times (n-1)}$ เป็นเมทริกซ์ความแปรปรวนร่วมของ $\gamma(t)$ ดังนั้นจากผลข้างต้นเราจะหาปริมาณ Condition GC ที่ $x_j(t)$ มีผลต่อ $x_i(t)$ จะได้ว่า

$$F_{x_j \rightarrow x_i | \text{others}} = \log \frac{\tilde{\Sigma}_{11}}{\Sigma_{11}} \quad (14)$$

โดยที่ $F_{x_j \rightarrow x_i | \text{others}}$ หมายถึงผลของ $x_j(t)$ ต่อ $x_i(t)$ ที่คำนึงถึง $x_k(t), k \neq \{i, j\}$

3.2 การทดสอบ Granger Causality โดยวิธีการทางสถิติ

หลังจากได้แบบจำลองสำหรับสัญญาณ EEG แล้ว ขั้นตอนถัดไปคือการหาโครงสร้างความสัมพันธ์ของสัญญาณโดยพิจารณาค่าศูนย์ของ GC ซึ่งในการทดลองของเราจะหาค่าศูนย์ของ GC โดยใช้วิธีการดังต่อไปนี้

3.2.1 การทดสอบค่าศูนย์ของ AR coefficients ด้วยวิธี Wald

จากการทำการทดลองหาค่า A_k จาก (5) จะพบว่าค่าในเมทริกซ์ A_k ที่ประมาณได้นั้นมีค่าหลากหลาย แต่จากที่กล่าวไปตั้ง (9) เราจะสนใจเฉพาะค่าศูนย์ใน A_k เท่านั้น ซึ่งจากการพิจารณาค่าศูนย์ในเมทริกซ์ A_k พบว่าค่าที่ประมาณได้อาจจะไม่ได้มีค่าเป็นศูนย์จริงๆ เพียงแต่มีค่าใกล้เคียงกับศูนย์

ดังนั้นปัญหาถัดมาคือ การตัดสินใจว่าตำแหน่งใดใน A_k บ้างที่มีค่าเป็น 0 สำหรับการทดลองนี้ เราได้ตั้งสมมติฐานขึ้นมาทดสอบว่าค่าตำแหน่งใดๆใน A_k มีค่าเป็นศูนย์หรือไม่ โดยการทดสอบสมมติฐานนั้นจะนำหลักการของการทดสอบ Wald มาใช้โดยจาก [6] การทดสอบ Wald มีหลักการคือถ้าสมมติฐานว่าค่าที่แท้จริงของพารามิเตอร์มีค่าเท่ากับศูนย์เป็นจริง แล้วค่าพารามิเตอร์ที่ประมาณได้ควรมีค่าเข้าใกล้ศูนย์อย่างมีนัยยะสำคัญ

ในการตัดสินใจว่าตำแหน่งนั้นมีค่าเป็น 0 หรือไม่ ทำได้โดยโดยการเปรียบเทียบค่าทางสถิติของ Wald (W) ซึ่งมีการแจกแจงเข้าสู่การแจกแจงแบบไคสแควร์ (Chi-square distribution) กับ ค่า c (critical value) ของแต่ละตำแหน่งใน A_k เมื่อ α คือระดับนัยสำคัญ (significance level) ได้ว่า $c = G^{-1}(1-\alpha)$ โดยหาก

$$W \leq c \quad (15)$$

เราจะให้ค่าในตำแหน่งนั้นเป็น 0 (ดูรายละเอียดการหาค่า W และ c ได้จาก [8])

3.2.2 การทดสอบค่าศูนย์ของ Conditional Granger Causality โดยวิธีการทางสถิติ

สำหรับแบบจำลอง AR ที่มีสัญญาณ n สัญญาณ ที่สามารถแสดงได้ดัง (2) จาก (14) ได้แสดงให้เห็นว่า ปริมาณความสัมพันธ์ Condition GC ของแต่ละคู่สัญญาณในแบบจำลอง AR นั้น สามารถเขียนให้อยู่ ในรูปเมทริกซ์ของ Condition GC ได้ดังนี้

$$F = [F_{ij}] \quad (16)$$

โดยที่ $F \in \mathbb{R}^{n \times n}$ และ $F_{ij} = F_{x_j \rightarrow x_i | \text{others}}$ จากนิยาม หากการรู้ค่า y_j ไม่มีผลต่อการประมาณ y_i แล้ว จะได้ว่า

$$F_{ij} = 0 \quad (17)$$

เงื่อนไขข้างต้นจึงเป็นอีกวิธีที่เราสามารถวิเคราะห์โครงสร้างความสัมพันธ์ GC ของข้อมูล EEG ได้ โดย วิเคราะห์จากตำแหน่งค่าศูนย์ในเมทริกซ์ F

จากนิยามค่าในเมทริกซ์ F จะมีค่ามากกว่าหรือเท่ากับศูนย์เสมอ ในการทดลองหาค่า F จาก (14) พบว่าผลการประมาณค่าในเมทริกซ์ F อาจจะได้ไม่มีค่าเป็นศูนย์จริงๆ เพียงแต่มีค่าใกล้เคียงกับศูนย์

ดังนั้นปัญหาถัดมาคือ การตัดสินใจว่าตำแหน่งใดใน F บ้างที่มีค่าเป็น 0 สำหรับการทดลองนี้ เราได้ตั้ง สมมติฐานขึ้นมาทดสอบว่าค่าตำแหน่งใดๆใน F มีค่าเป็นศูนย์หรือไม่ โดยการทดสอบสมมติฐานนั้นจะ นำหลักการของการทดสอบ Chi-square มาใช้ โดยจาก [2] การทดสอบ Chi-square มีหลักการคือถ้า สมมติฐานว่าค่าที่แท้จริงของพารามิเตอร์มีค่าเท่ากับศูนย์เป็นจริงแล้ว ค่าพารามิเตอร์ที่ประมาณได้ควร มีค่าเข้าใกล้ศูนย์อย่างมีนัยยะสำคัญ (ดูรายละเอียดการตัดสินใจให้ค่า F_{ij} เป็นศูนย์เพิ่มเติมได้ที่ [4])

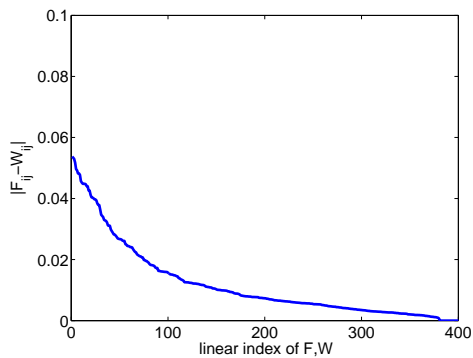
3.2.3 การเปรียบเทียบค่า Granger Causality จากวิธีทางสถิติแบบวาลด์กับ Conditional Granger Causality.

ในการทดลองหาค่า GC เบื้องต้นเราสนใจใช้ วิธีทางสถิติแบบวาลด์ และ Conditional GC ซึ่งทั้งสองวิธี ให้ค่า GC ต่างกัน ดังนั้นเราจึงอยากพิจารณาความต่างระหว่างค่าจากทั้งสองวิธีนี้ จึงสร้างข้อมูลจำลอง ขึ้นมาให้มี $n = 20$ channel, มีจำนวนข้อมูล N และทดสอบหาความสัมพันธ์ระหว่างข้อมูล เพื่อให้ ง่ายต่อการเปรียบเทียบเนื่องจาก F (16) และ W (15) มีค่าต่างกันมาก เราจึงปรับขนาดค่าใน F และ W ให้มีค่าอยู่ในมาตราส่วนเดียวกัน กำหนดให้มีค่าน้อยสุดเป็น 0 มากสุดเป็น 1 โดยมีสูตรการคำนวณ ปรับขนาดมาตราส่วน F และ W ดังนี้

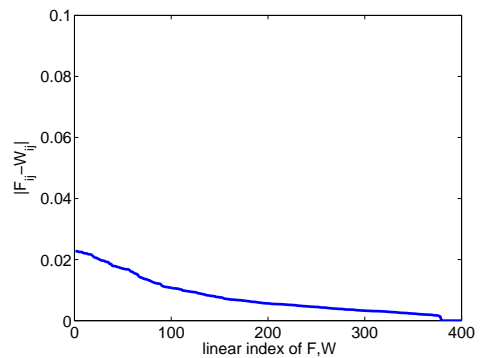
$$F_{\text{scale}} = \frac{F_{ij} - F_{\min}}{F_{\max} - F_{\min}}, \quad W_{\text{scale}} = \frac{W_{ij} - W_{\min}}{W_{\max} - W_{\min}}$$

จากการทดลองหลังจากปรับขนาดมาตราส่วนแล้วเปรียบเทียบค่า F และ W พบว่ายังมีจำนวนข้อมูล มาก ค่า GC จากทั้งสองวิธีจะยังมีค่าใกล้เคียงกันมากขึ้น เมื่อทำการทดลองหลายๆครั้งให้ผลประมาณ ได้ดังรูปที่ 5

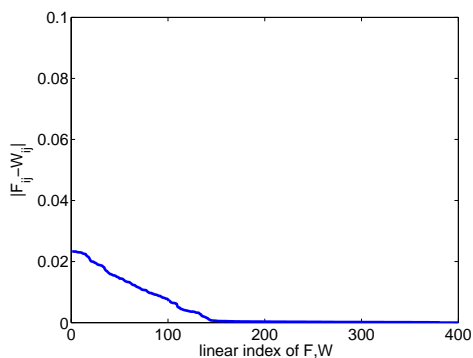
จากการทดลองจะเห็นว่าค่า F และ W ที่ N มากกว่า 1,000 มีขนาดต่างกันสูงสุดประมาณ 2% เท่านั้น เราจึงคาดการณ์ว่าไม่ว่าเราจะหาค่า GC ด้วยวิธีการใดก็จะได้ผลที่ใกล้เคียงกัน ดังนั้นในการทดลองนี้เพื่อความเรียบง่ายเราจะเลือกใช้วิธีการเดียวเท่านั้นในการหาค่า GC ซึ่งจากการพิจารณากระบวนการหาโครงสร้างความสัมพันธ์ของแบบจำลอง AR พบว่าการหาค่า GC ด้วยวิธีการทางสถิติแบบวาลด์ (2 ขั้นตอน) นั้นกระชับกว่าวิธีแบบ Condition GC (3 ขั้นตอน) ดังแสดงในรูปที่ 6 ดังนั้นในการทดลองหาโครงสร้างความสัมพันธ์ถัดจากนี้ เราจะใช้วิธีการทางสถิติแบบวาลด์เป็นวิธีดำเนินการ



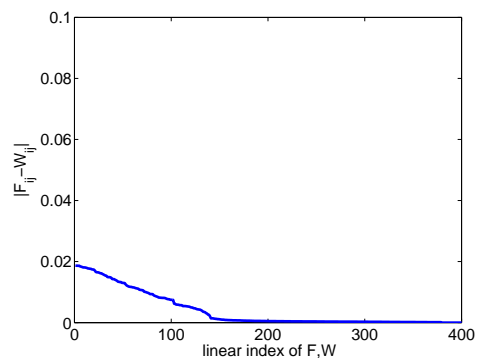
(a) $|F - W|$ เมื่อ $N = 500$



(b) $|F - W|$ เมื่อ $N = 1,000$

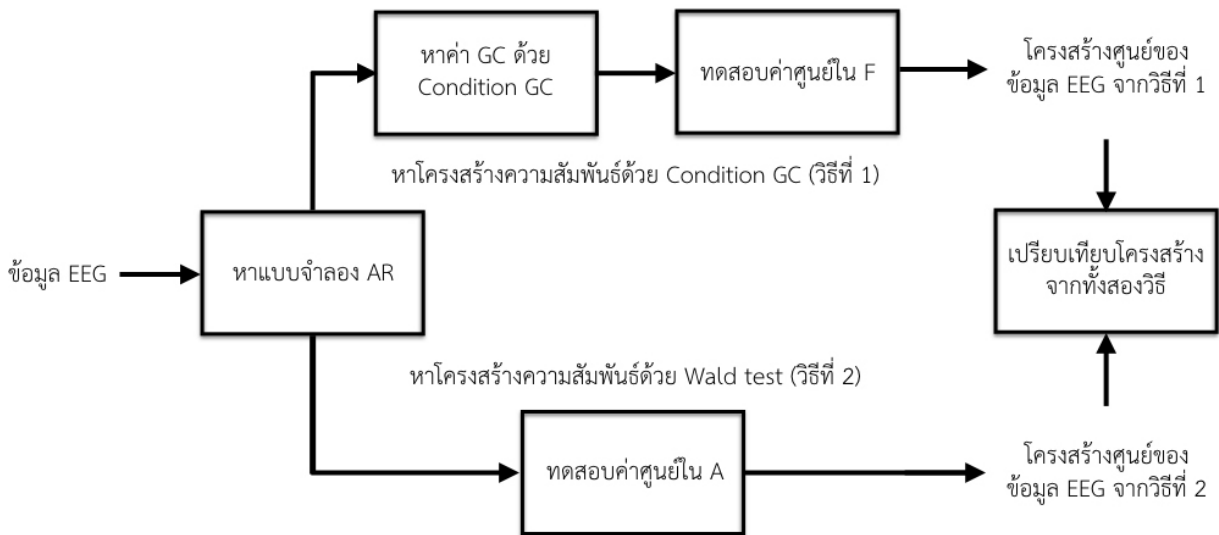


(c) $|F - W|$ เมื่อ $N = 5,000$



(d) $|F - W|$ เมื่อ $N = 10,000$

รูป 5: ขนาดค่า $|F - W|$ ซึ่งเป็นผลต่างของค่า GC จากวิธีการทดสอบของ Wald กับ Condition GC หลังปรับค่าภายในให้มีมาตราส่วนเท่ากัน และเรียงลำดับตัวชี้สมาชิกเชิงเส้นใหม่ตามขนาดจากมากไปน้อย

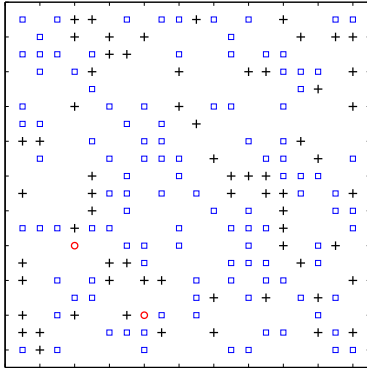


รูป 6: กระบวนการหาโครงสร้างความสัมพันธ์โดยวิธีการทางสถิติแบบวาลด์ และ Condition GC

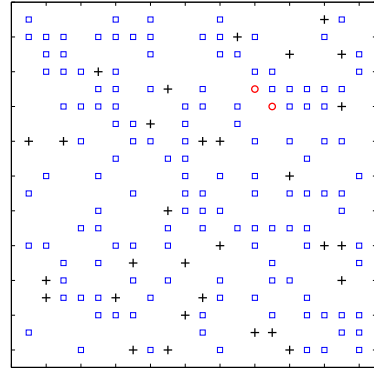
3.2.4 ผลการหา Granger Causality โดยวิธีการทางสถิติแบบ Wald

จาก (9) เราสามารถรู้โครงสร้างความสัมพันธ์ของข้อมูลได้จากค่าศูนย์ของพารามิเตอร์ใน A_k จากบทที่ 3.2.1 เราจะตัดสินใจว่าตำแหน่งใดใน A_k บ้างที่มีค่าเป็น 0 โดยใช้วิธีการทางสถิติแบบวาลด์ ในบทนี้เราได้ทำการทดลองเพื่อตรวจสอบความแม่นยำของวิธีการ โดยสร้างข้อมูลจำลองจากแบบจำลอง AR ที่เราทราบค่าตำแหน่งศูนย์ในเมทริกซ์ A_k ให้มีตำแหน่งศูนย์กระจายอยู่ใน A_k เป็นจำนวน 50% ของพารามิเตอร์ทั้งหมด แล้วเปรียบเทียบตำแหน่งศูนย์จริงกับตำแหน่งศูนย์ที่ประมาณได้

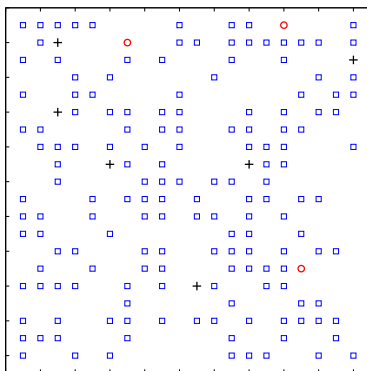
ทั้งนี้พบว่าความผิดพลาดในการประมาณตำแหน่งศูนย์มี 2 ประเภท ประกอบด้วย ความผิดพลาดประเภทที่ 1 คือ ที่ตำแหน่งจริงไม่เป็นศูนย์แต่ประมาณให้เป็นศูนย์ และความผิดพลาดประเภทที่ 2 คือ ที่ตำแหน่งจริงเป็นศูนย์แต่ประมาณได้ไม่เป็นศูนย์ จากการทดลองผลบ่งชี้ว่ายิ่งเรามีจำนวนข้อมูลมาก ความผิดพลาดในการทำนายตำแหน่งศูนย์ยิ่งน้อยลง เมื่อทำการทดลองหลายๆครั้งให้ผลประมาณได้ดังรูปที่ 7



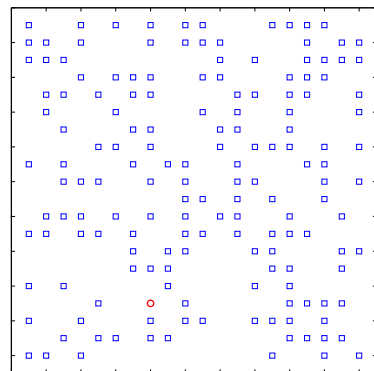
(a) เมื่อ $N = 500$ มีความผิดพลาดรวม 16.75%



(b) เมื่อ $N = 1,000$ มีความผิดพลาดรวม 8.00%



(c) เมื่อ $N = 5,000$ มีความผิดพลาดรวม 2.25%



(d) เมื่อ $N = 10,000$ มีความผิดพลาดรวม 0.25%

รูป 7: แสดงตำแหน่งศูนย์จริงใน A_k เทียบกับตำแหน่งศูนย์ที่ประมาณได้ โดย \square แทนตำแหน่งที่ไม่เป็นศูนย์ \circ แทนตำแหน่งที่ไม่เป็นศูนย์แต่ทำนายว่าเป็นศูนย์ $+$ แทนตำแหน่งที่เป็นศูนย์แต่ทำนายว่าไม่เป็นศูนย์

3.3 การใช้ Granger Causality แยกแยะสถานะคนไข้

โครงสร้างความสัมพันธ์ของสมองผู้ป่วยที่เราหาได้ สามารถนำไปประยุกต์ใช้ต่อยอดให้เกิดประโยชน์เพิ่มเติมได้มากมาย หนึ่งใน การประยุกต์ใช้ที่เราสนใจคือการใช้โครงสร้าง GC ที่เราหาได้มาแยกแยะสถานะคนไข้

จาก [7] คือหนึ่งในตัวอย่างการนำโครงสร้าง GC มาเป็นเวกเตอร์คุณลักษณะ (feature vector) สำหรับแยกแยะสถานะคนไข้ โดยเป็นการทดลองเพื่อแยกแยะสถานะรู้สึกตัวและไม่รู้สึกตัวของคนไข้ที่โดนยาสลบ เมื่อทำการแบ่งพื้นที่บนศีรษะเป็น 5 ส่วน ได้แก่ พื้นที่ส่วนกลาง (Z) พื้นที่ส่วนหน้าซ้าย (LF) พื้นที่ส่วนหน้าขวา (RF) พื้นที่ส่วนหลังขวา (RP) และ พื้นที่ส่วนหลังซ้าย (LP) จาก [7] พิจารณาความสัมพันธ์ระหว่างพื้นที่บนศีรษะโดยใช้ pairwise GC (13) เมื่อพิจารณาค่าความสัมพันธ์ระหว่างพื้นที่ต่างๆ บนศีรษะ ได้คู่ความสัมพันธ์ทั้งหมดดังนี้

$$F_{LF \leftrightarrow RF}, F_{LF \leftrightarrow LP}, F_{LF \leftrightarrow RP}, F_{LF \leftrightarrow Z}, F_{RF \leftrightarrow LP},$$

$$F_{RF \leftrightarrow RP}, F_{RF \leftrightarrow Z}, F_{LP \leftrightarrow RP}, F_{LP \leftrightarrow Z} \text{ และ } F_{RP \leftrightarrow Z}$$

จาก [11] หนึ่งในเทคนิคที่ให้ผลลัพธ์ที่ดีที่สุดและนิยมใช้แยกแยะข้อมูล EEG คือ เทคนิค Support Vector Machine (SVM) ซึ่ง [7] ก็ใช้เทคนิคนี้ในการแยกแยะข้อมูลสองสถานะเช่นเดียวกัน โดยมีเวกเตอร์คุณลักษณะดังนี้

$$X = [F_{LF \rightarrow LP} \quad F_{RF \rightarrow LP} \quad F_{LF \rightarrow RP} \quad F_{RF \rightarrow RP}]$$

จากการที่ใช้เวกเตอร์คุณลักษณะนี้แยกแยะข้อมูลในสภาวะรู้สึกตัวและไม่รู้สึกตัวของคนไข้ที่โดนยาสลบ สามารถแยกได้แม่นยำมากถึง 98%

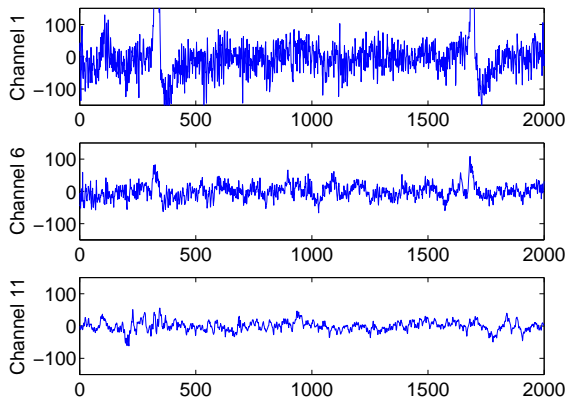
อนึ่งสำหรับการใช้โครงสร้าง GC แยกแยะสภาวะคนไข้ เราเพียงต้องการศึกษาเพื่อให้เห็นตัวอย่างการนำโครงสร้างความสัมพันธ์ GC ไปใช้ประโยชน์เท่านั้น ไม่ได้ทดลองแยกข้อมูลแต่อย่างใดเนื่องจากอยู่นอกเหนือวัตถุประสงค์ของโครงการ

4 ผลการทดลองบนข้อมูล EEG

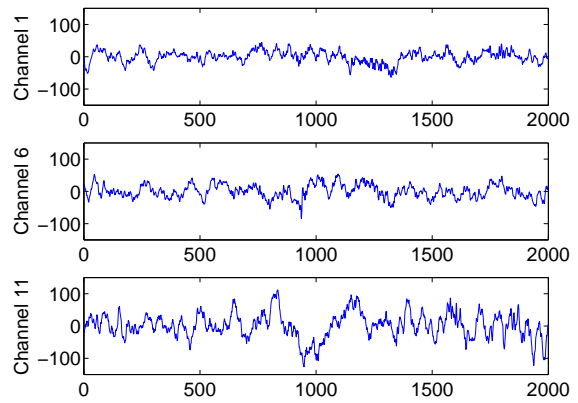
4.1 รายละเอียดของข้อมูล EEG

ในการทดลองนี้เรานำชุดข้อมูล EEG มาจาก [12] ซึ่งเป็นข้อมูลที่วัดมาจากผู้ป่วยโรคลมชัก 10 คน แบ่งเป็น ชาย 3 คน หญิง 7 คน ข้อมูล EEG 1 ไฟล์ประกอบด้วยข้อมูลสัญญาณ 23 ช่อง สุ่มด้วยความถี่ 256 Hz เป็นเวลา 1 ชั่วโมง (921,600 จุด) โดยทาง [12] ได้ระบุช่วงเวลาในสภาวะปกติ และสภาวะที่เกิดการชักของผู้ป่วยไว้ชัดเจนเราจึงทำการตัดสัญญาณเพื่อแยกข้อมูลเป็น 2 ส่วน มีตัวอย่างดังรูปที่ 18

จากรูปที่ 18 หากสังเกตด้วยสายตาก็เห็นว่า สัญญาณของผู้ป่วยในสภาวะปกติ (เป็นสัญญาณ Beta มีความถี่ 13-30 Hz [11]) มีความถี่มากกว่าสัญญาณของผู้ป่วยในสภาวะชัก (เป็นสัญญาณ Spikes มีความถี่ประมาณ 3 Hz [11]) ดังนั้นเราจึงมีสมมติฐานว่าแบบจำลองและโครงสร้างศูนย์ของสัญญาณทั้งสองชนิดจะต่างกัน ซึ่งเราจะทดสอบสมมติฐานเหล่านี้ในหัวข้อถัดไป



(a) สัญญาณของผู้ป่วยคนที่ 1 ในสภาวะปกติ



(b) สัญญาณของผู้ป่วยคนที่ 1 ในสภาวะชัก

รูป 8: ตัวอย่างสัญญาณ EEG ของผู้ป่วยโรคลมชักคนที่ 1 ที่สองสภาวะ

เมื่อทำการทดลองในโปรแกรม MATLAB ข้อมูล 1 ไฟล์ประกอบด้วยข้อมูลสัญญาณ 23 แฉว 921,600 หลัก โดยแต่ละแฉวหมายถึงแต่ละช่องสัญญาณ (channel) และแต่ละหลักคือกรอบเวลาที่สุ่มวัดสัญญาณ จากการทดลองพบว่าข้อมูลแฉวที่ 1,3,15,23 ไม่เป็นอิสระต่อกัน ทำให้หาเมทริกซ์ผกผันตาม (5) ไม่ได้ และจากการทดลองในหัวข้อที่ 3.2.4 สำหรับสัญญาณ 20 แฉว หากมีข้อมูลมากกว่า 10,000 หลัก จะทำให้เปอร์เซ็นต์ความผิดพลาดของการทำนายตำแหน่งศูนย์น้อยกว่า 1% ดังนั้นจากสัญญาณที่เรามีประกอบด้วยข้อมูล 23 แฉว 921,600 หลัก ซึ่งเยอะมากพอที่เราจะเลือกพิจารณาเป็นบางส่วนได้ เพื่อให้การทดลองมีความง่าย เร็ว และผิดพลาดน้อยที่สุด เราจะขอละการพิจารณาข้อมูลแฉวที่ 1,3,15,23 และพิจารณาสัญญาณแต่ละชุดให้มีกรอบเวลาไม่เกิน 40 วินาที ดังนั้นในการทดลองของเราสัญญาณ EEG หนึ่งชุดจะมีข้อมูล 19 แฉว 10,240 หลัก (ดูตัวอย่างสัญญาณ EEG ของผู้ป่วยคนหนึ่งในสภาวะปกติและสภาวะชักหลังตัดแต่งข้อมูลแล้วได้ในภาคผนวก)

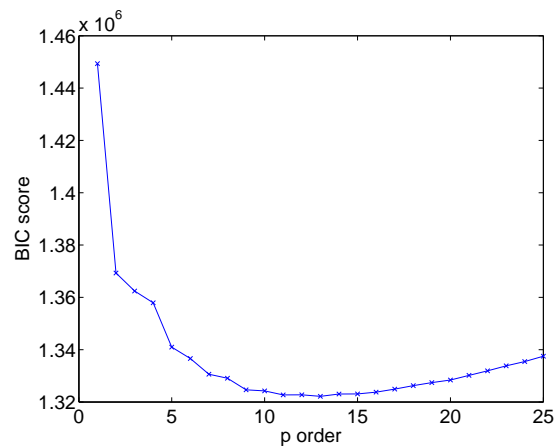
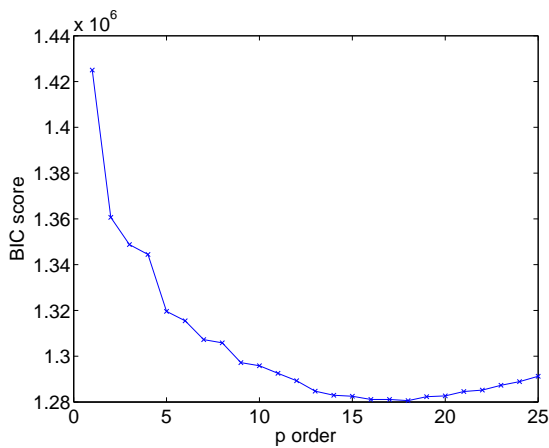
4.2 การประมาณแบบจำลอง AR ของข้อมูล EEG

ในขั้นต้นเราได้ทดลองศึกษาลักษณะผลตอบของแบบจำลอง AR สำหรับสัญญาณ EEG ของผู้ป่วยโรคลมชักในสภาวะปกติจาก 23 ช่องสัญญาณที่วัดมาจากตำแหน่งต่างๆบนศีรษะ หลังจากได้แบบจำลองมาแล้ว หากพิจารณาความคลาดเคลื่อนที่อันดับ p ต่างๆ จะเห็นว่าค่าความผิดพลาดของแบบจำลองแปรผกผันกับค่า p โดยการที่ค่าความผิดพลาดน้อยนั้นดี แต่หาก p มีค่ามากแล้วทำให้แบบจำลองมีความซับซ้อนมาก ดังนั้นจึงทำให้เกิดปัญหาข้อถัดไปขึ้น ซึ่งคือการเลือกอันดับที่เหมาะสมที่ทำให้ค่าความผิดพลาดไม่มากเกินไปและแบบจำลองมีความซับซ้อนน้อย

จากบทที่ 2.1 เราจะเลือกอันดับของแบบจำลองจากคะแนนของแบบจำลอง BIC (8) ที่ให้กับแต่ละค่าอันดับของแบบจำลอง AR โดยใช้ `BICARmodel.m` ในหัวข้อที่ 7.2 ซึ่งเราจะเลือกอันดับที่คะแนน BIC ต่ำสุดเป็นอันดับที่เหมาะสมสำหรับแบบจำลอง อนึ่งจากการพิจารณาข้อมูล EEG ในบทที่ 4.1 เรามี

สมมติฐานว่าแบบจำลองของสัญญาณที่สภาวะปกติและสภาวะชักในคนไข้แต่ละคนจะต่างกัน ดังนั้นจากข้อมูล EEG ที่เรามีทั้งหมดสำหรับคนไข้แต่ละคนเราจะแยกพิจารณาหาอันดับที่เหมาะสมสำหรับแบบจำลองของข้อมูลในสภาวะปกติและอันดับสำหรับแบบจำลองของข้อมูลในสภาวะชัก

จากการทดลองกับข้อมูลของผู้ป่วยในสภาวะปกติพบว่าอันดับของแบบจำลอง AR มีค่าตั้งแต่ 4-16 ส่วนการทดลองกับข้อมูลของผู้ป่วยในสภาวะชักพบว่าอันดับของแบบจำลอง AR มีค่าตั้งแต่ 6-13 ซึ่งเราเลือกใช้อันดับตามความเหมาะสมกับข้อมูลแต่ละชุด ตัวอย่างคะแนน BIC ของสัญญาณ EEG ของผู้ป่วยคนหนึ่งในสองสภาวะเทียบกับอันดับแสดงดังรูปที่ 9 และ 10

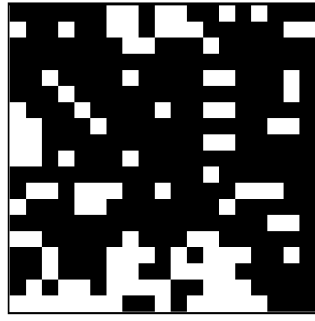


รูป 9: BIC score ของแบบจำลองในสภาวะปกติ

รูป 10: BIC score ของแบบจำลองในสภาวะชัก

4.3 โครงสร้างความสัมพันธ์ของสมองบนตำแหน่งต่างๆ ของผู้ป่วยในสองสภาวะ

หลังจากทำให้ข้อมูลแต่ละชุดมี 19 channel และกรอบเวลาไม่เกิน 40 วินาทีดังที่กล่าวไปในหัวข้อที่ 4.1 และได้แบบจำลองที่เหมาะสมแล้ว ขั้นตอนถัดไปคือการหาโครงสร้างความสัมพันธ์ของข้อมูล EEG จากแบบจำลอง AR ด้วยวิธีทางสถิติแบบวาลด์ ในการทดลองของเราข้อมูล 1 ชุดมี 19 แถว 10,240 หลัก เมื่อทำการทดลองหาเมทริกซ์ W ด้วย `funcWaldTest.m` ในหัวข้อที่ 7.2 ได้ผลลัพธ์ออกมาเป็น binary matrix ที่ค่าภายในเป็น 0 และ 1 ซึ่งคือค่า Wald statistic ที่ผ่านการตรวจสอบกับค่า critical value ดัง (15) โดยจะมีขนาด $n \times n$ เมื่อ n คือจำนวน channel ของสัญญาณ EEG ที่เราพิจารณา เพื่อให้ง่ายต่อการพิจารณาเมื่อนำมาวาดให้เป็นรูปโทเนตาด้วยโปรแกรม MATLAB จะได้ผลประมาณได้ดังรูปที่ 11 โดยตำแหน่งที่มีความสัมพันธ์คือ W_{ij} มีค่าไม่เป็นศูนย์ จะแทนด้วยสีดำ และตำแหน่งที่ไม่มีความสัมพันธ์คือ W_{ij} มีค่าเป็นศูนย์จะแทนด้วยสีขาว โดยจากข้อมูลของเราที่สนใจสัญญาณ 19 channel จะได้ $W \in \mathbb{R}^{19 \times 19}$

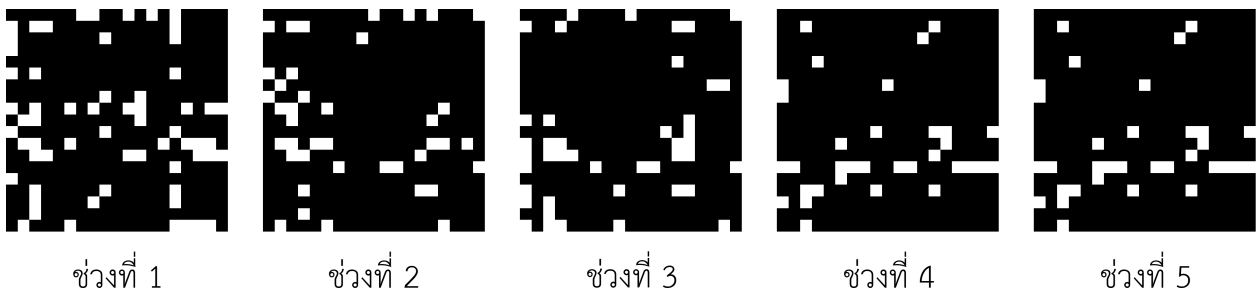


รูป 11: ตัวอย่างโครงสร้างความสัมพันธ์แบบ Granger Causality ของข้อมูล EEG 1 ชุด

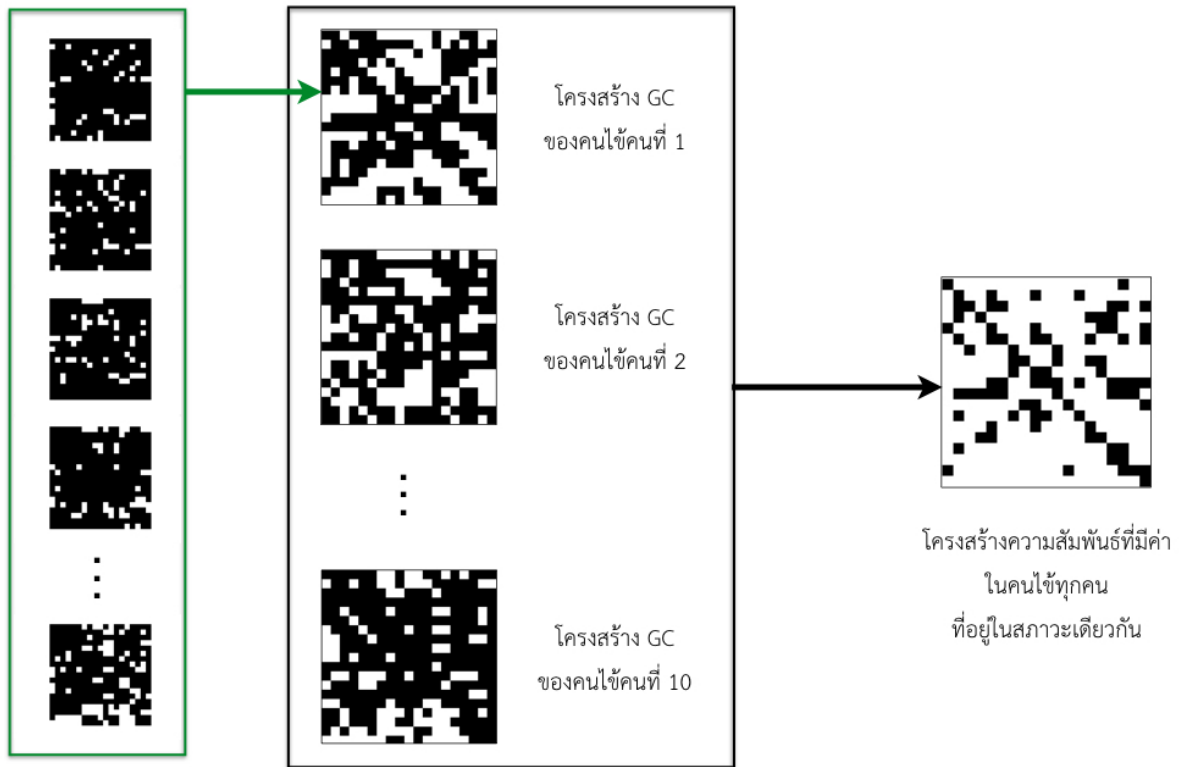
จากหัวข้อที่ 4.1 เราได้ตัดแบ่งข้อมูลคนไข้แต่ละคนออกเป็นหลายส่วน ทำให้ผลการหาโครงสร้าง GC ของสัญญาณ EEG ของคนไข้แต่ละคนในแต่ละช่วงเวลานั้นแตกต่างกัน ดังแสดงในรูปที่ 12 แต่ในการทดลองของเรา เราสนใจหาโครงสร้าง GC ของสัญญาณ EEG ในสภาวะปกติและสภาวะชักที่ปรากฏอยู่ในคนไข้ทุกคนทุกช่วงเวลา ดังนั้นเราจะพิจารณาหาโครงสร้าง GC ของคนไข้แต่ละคนในสภาวะปกติและสภาวะชัก แล้วจึงนำโครงสร้างที่ได้จากคนไข้แต่ละคนมาหาโครงสร้างที่มีความสัมพันธ์ร่วมกันต่อไป ดังแสดงในรูปที่ 13

หลังจากทำการทดลองดังรูปที่ 13 จากข้อมูลของเราที่ประกอบด้วยข้อมูลของคนไข้สภาวะชักและสภาวะปกติ สุดท้ายแล้วเราจะได้โครงสร้างความสัมพันธ์สองแบบซึ่งแทนโครงสร้างที่ปรากฏในคนไข้สภาวะปกติทุกคนในทุกช่วงเวลาทำการทดลองดังรูป 15a และโครงสร้างที่ปรากฏในคนไข้สภาวะชักทุกคนในทุกช่วงเวลาทำการทดลองดังรูป 15b

เพื่อให้เข้าใจได้ง่ายเราจะอ้างอิงตำแหน่งบนศีรษะบริเวณที่เป็นแหล่งกำเนิดสัญญาณ EEG แทนสัญญาณ EEG แต่ละ Channel จากรูปที่ 14 เป็นโครงสร้างความสัมพันธ์ binary matrix ที่แสดงถึงความสัมพันธ์ของสัญญาณ EEG แต่ละ Channel (ดูรายละเอียดตำแหน่งที่วัดสัญญาณทั้ง 23 Channel ได้หัวข้อที่ 7.1) เมื่อนำมาพิจารณาให้บริเวณที่เป็นแหล่งกำเนิดสัญญาณ EEG แทนสัญญาณ EEG แต่ละ Channel จะสามารถแสดงความสัมพันธ์ของตำแหน่งต่างๆบนสมองได้ดังรูปที่ 15



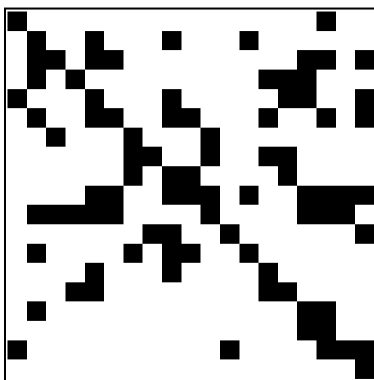
รูป 12: ตัวอย่างโครงสร้างศูนย์ของคนไข้คนหนึ่งในแต่ละช่วงเวลา



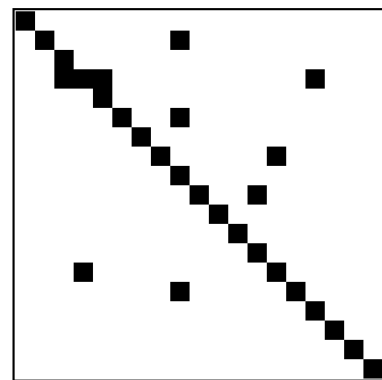
โครงสร้างศูนย์ของคนใช้คนที่ 1
ในแต่ละช่วงเวลา

โครงสร้างความสัมพันธ์ที่มีค่า
ทุกช่วงเวลาของคนใช้แต่ละคนในสถานะเดียวกัน

รูป 13: แสดงวิธีหาโครงสร้างความสัมพันธ์ของสัญญาณ EEG ที่มีในคนใช้ทุกคนในทุกช่วงเวลา โดย → แทนกระบวนการนำโครงสร้างในกลุ่มเดียวกันมาหาความสัมพันธ์ที่มีร่วมกัน

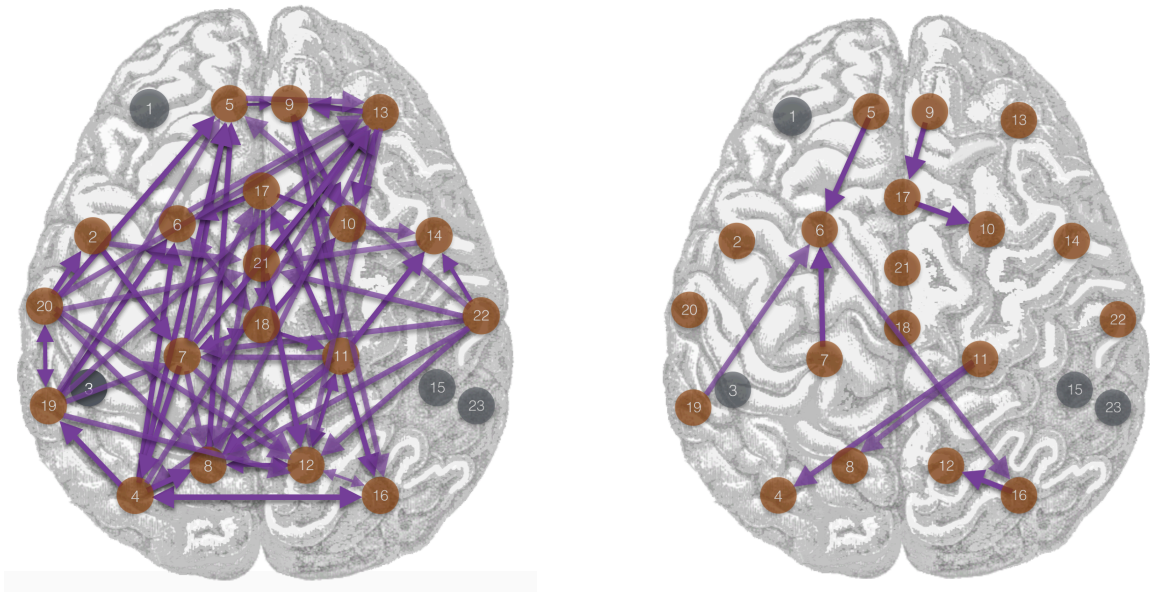


(a) โครงสร้างความสัมพันธ์ของคนใช้สถานะปกติ



(b) โครงสร้างความสัมพันธ์ของคนใช้สถานะชัก

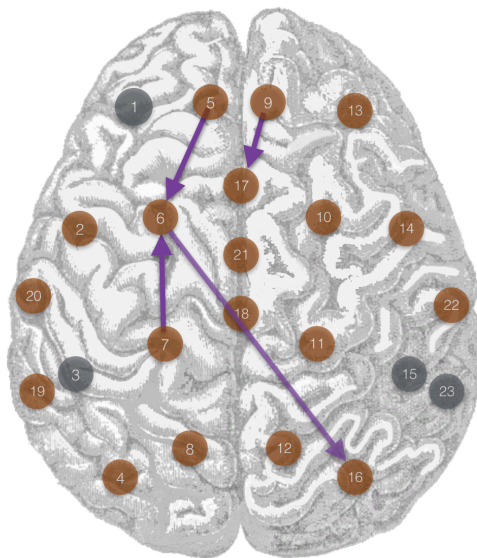
รูป 14: โครงสร้างความสัมพันธ์ของสัญญาณ EEG ของผู้ป่วยโรคลมชักในสถานะปกติและสถานะชักที่ปรากฏในคนใช้ทุกคนในทุกช่วงเวลา



(a) โครงสร้างความสัมพันธ์ของตำแหน่งต่างๆบนสมองของคนใช้สภาวะปกติ (b) โครงสร้างความสัมพันธ์ของตำแหน่งต่างๆบนสมองของคนใช้สภาวะชัก

รูป 15: โครงสร้างความสัมพันธ์ของตำแหน่งต่างๆบนสมองที่ให้กำเนิดสัญญาณ EEG ของผู้ป่วยโรคลมชักในสภาวะปกติและสภาวะชัก โดยตำแหน่งสีเทาคือช่องสัญญาณ 1,3,15,23 ที่เราละเว้นไม่พิจารณา

จากการทดลองพบว่าสัญญาณ EEG ที่วัดได้จากสมองของผู้ป่วยโรคลมชัก ไม่ว่าจะเวลาใดหรือใครก็ตาม ในสภาวะปกติจะมีโครงสร้างความสัมพันธ์ดังรูปที่ 15a เสมอ และในสภาวะชักจะมีโครงสร้างความสัมพันธ์ดังรูปที่ 15b ซึ่งจะเห็นว่าในสภาวะชักโครงสร้างความสัมพันธ์ของสมองจะมีความซับซ้อนน้อยกว่าโครงสร้างความสัมพันธ์ในสภาวะปกติมาก และมีบางความสัมพันธ์ที่มีเฉพาะในสภาวะชักเท่านั้นแต่ไม่มีในสภาวะปกติ ดังแสดงในรูปที่ 16



รูป 16: โครงสร้างความสัมพันธ์ของตำแหน่งต่างๆบนสมองที่ปรากฏในสถานะชักเท่านั้น

5 บทสรุป

โครงการนี้จัดทำขึ้นโดยมีจุดประสงค์เพื่อหาโครงสร้างความสัมพันธ์ระหว่างเหตุและผลของข้อมูล EEG ที่เราสนใจ เนื่องจากสัญญาณ EEG เป็นสัญญาณต่อเนื่องทางเวลาเราจึงมีสมมติฐานว่าค่าของสัญญาณ ปัจจุบันขึ้นอยู่กับอดีต ดังนั้นเราจึงใช้แบบจำลอง AR ในการหาแบบจำลองของสัญญาณ EEG ซึ่งพารามิเตอร์ศูนย์ที่ได้จากการหาแบบจำลอง AR นั้นสามารถสะท้อนถึงความสัมพันธ์ระหว่างเหตุและผลของสัญญาณ EEG ได้ ดังนั้นเราจึงสนใจหารูปแบบโครงสร้างของศูนย์ของสัญญาณ EEG ซึ่งสามารถหาได้จากการทดสอบสมมติฐานแบบ Wald สุดท้ายแล้วเราจะได้รูปแบบโครงสร้างศูนย์ที่สามารถอธิบายความสัมพันธ์ระหว่างเหตุและผลของข้อมูล EEG ที่เราสนใจได้

6 กิตติกรรมประกาศ

โครงการวิศวกรรมไฟฟ้า เรื่องการเรียนรู้ความสัมพันธ์ของตัวแปรจากข้อมูล EEG จะไม่สามารถสำเร็จได้เลย หากขาดความช่วยเหลือและการสนับสนุนจาก ผู้ช่วยศาสตราจารย์ ดร. จิตโกมุท สงศิริ อาจารย์ที่ปรึกษาในวิชาโครงการวิศวกรรมไฟฟ้า โดยการให้คำปรึกษา และคำแนะนำที่เป็นประโยชน์ในการทำโครงการนี้ ตลอดจนให้ความรู้ในการแก้ปัญหา ปรับปรุง และแก้ไขจนรายงานเล่มนี้เสร็จสมบูรณ์ ผู้จัดทำจึงขอขอบพระคุณเป็นอย่างสูง

References

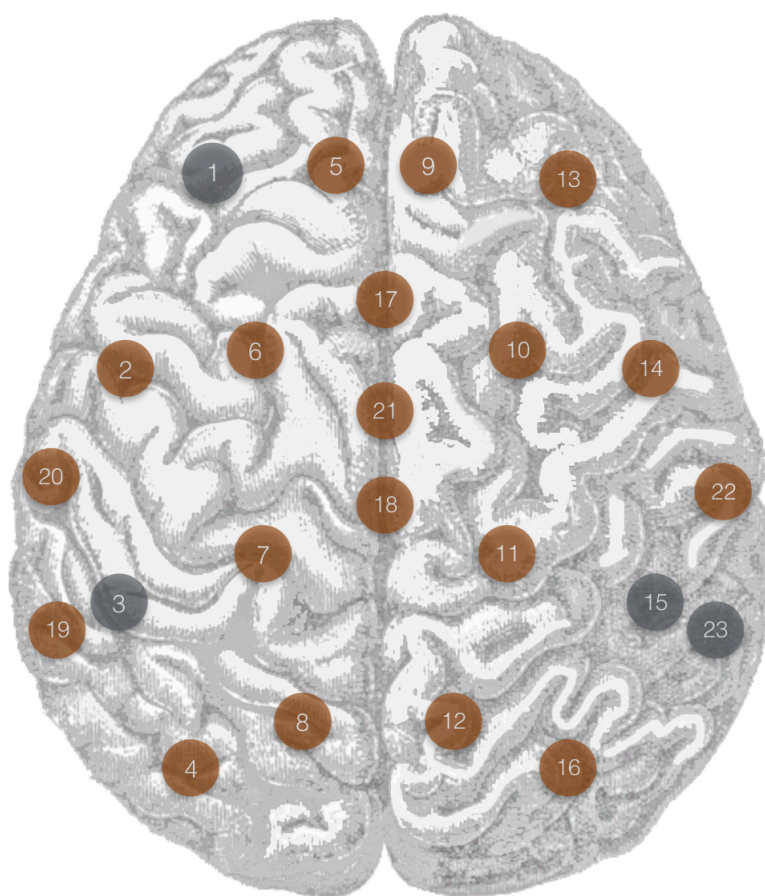
- [1] Ralph G Andrzejak, Kaspar Schindler, and Christian Rummel. Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients. *Physical Review E*, 86(4):046206, 2012.
- [2] Steven L Bressler and Anil K Seth. Wiener-Granger causality: a well established methodology. *Neuroimage*, 58(2):323–329, 2011.
- [3] Mingzhou Ding, Yonghong Chen, and Steven L Bressler. 17 Granger causality: basic theory and application to neuroscience. *Handbook of time series analysis: recent theoretical developments and applications*, 2006.
- [4] John Geweke. Measurement of linear dependence and feedback between multiple time series. *Journal of the American statistical association*, 77(378):304–313, 1982.
- [5] Clive WJ Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438, 1969.
- [6] W.H. Greene. *Econometric Analysis*. Pearson/Prentice Hall, 2008.

- [7] Nicoletta Nicolaou, Saverios Hourris, Pandelitsa Alexandrou, and Julius Georgiou. EEG-based automatic classification of ‘awake’ versus ‘anesthetized’ state in general anesthesia using Granger causality. *PLoS One*, 7(3):e33869, 2012.
- [8] Nuntanut Raksasri. Exploring Granger causality for time series via statistical tests on estimated models. In *Senior Project Report*. Electrical Engineering Department, Faculty of Engineering, Chulalongkorn University, 2015.
- [9] Alard Roebroek, Elia Formisano, and Rainer Goebel. Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage*, 25(1):230–242, 2005.
- [10] James B Rowe. Connectivity analysis is essential to understand neurological disorders. 2010.
- [11] S. Sanei and J.A. Chambers. *EEG Signal Processing*. Wiley, 2008.
- [12] Ali H Shoeb and John V Guttag. Application of machine learning to epileptic seizure detection. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 975–982, 2010.
- [13] Jitkomut Songsiri. Sparse autoregressive model estimation for learning Granger causality in time series. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 3198–3202. IEEE, 2013.
- [14] Le Zhang, Guangjin Zhong, Yukun Wu, Mark G Vangel, Beini Jiang, and Jian Kong. Using Granger-Geweke causality model to evaluate the effective connectivity of primary motor cortex (m1), supplementary motor area (SMA) and cerebellum. *Journal of biomedical science and engineering*, 3:848, 2010.

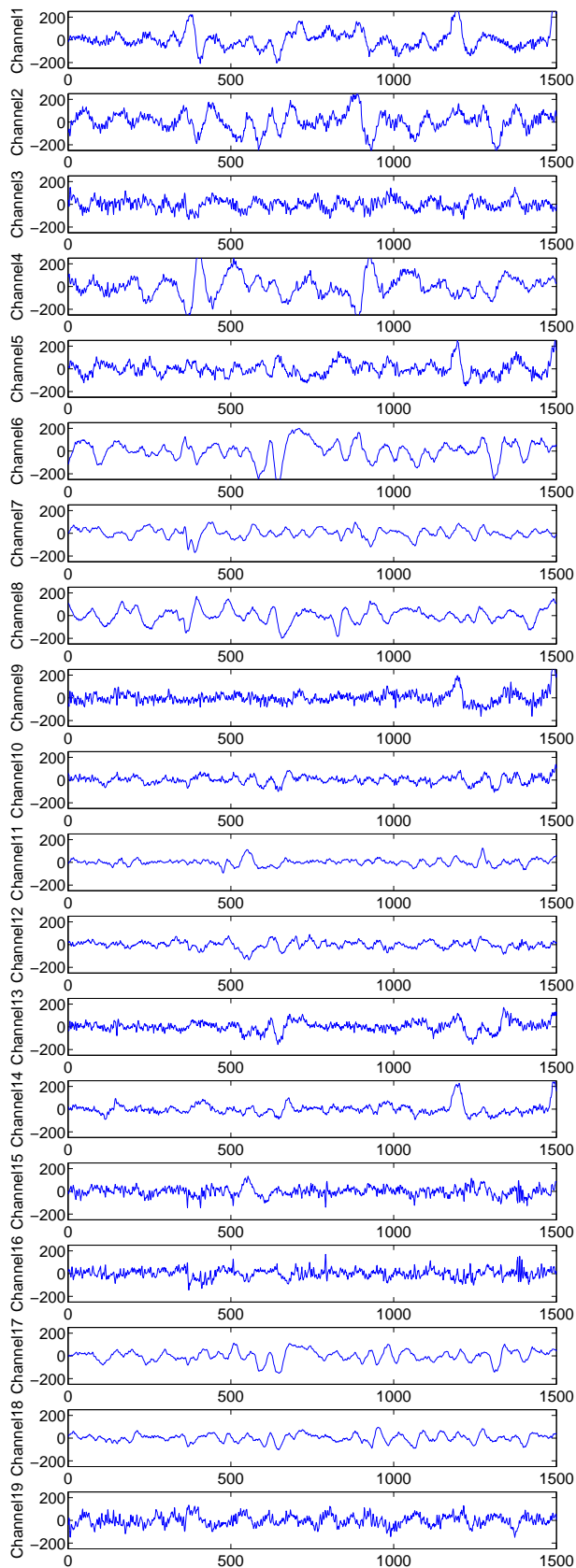
7 ภาคผนวก

7.1 รายละเอียดของข้อมูล EEG

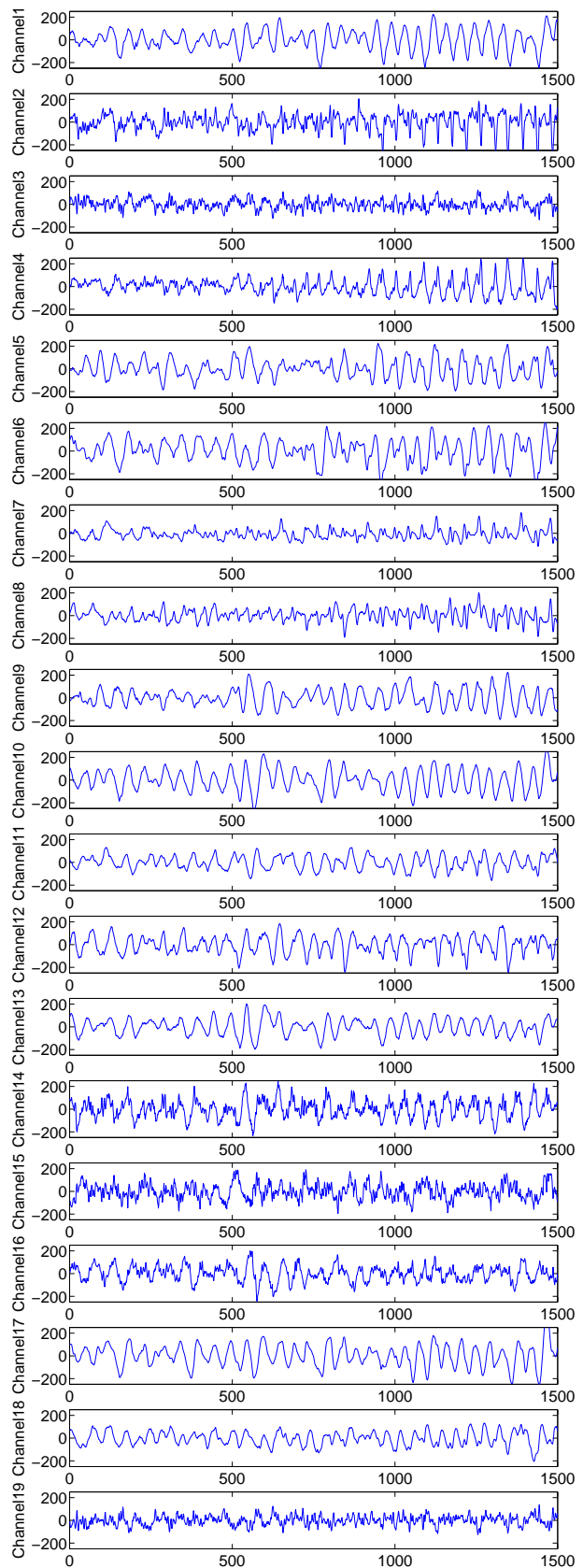
การทดลองทั้งหมดในโครงการนี้ทดลองกับข้อมูล EEG ที่วัดมาจากตำแหน่งบนศีรษะดังรูปที่ 17 ในขณะที่รูปที่ 18 เป็นตัวอย่างของสัญญาณ EEG ของคนใช้คนที่ 4 เฉพาะ channel ที่เราสนใจทั้งหมด 19 channel จะเห็นว่าสัญญาณที่สภาวะปกติส่วนมากมีลักษณะเป็นสัญญาณ Beta มีความถี่ 13-30 Hz [11] ส่วนสัญญาณที่สภาวะชักส่วนมากมีลักษณะเป็นสัญญาณ Spikes มีความถี่ประมาณ 3 Hz [11]



รูป 17: บริเวณที่เป็นแหล่งกำเนิดสัญญาณ EEG ทั้ง 23 Channel โดยตำแหน่งสีเทาคือ Channel 1,3,15,23 ที่เราละเว้นไม่พิจารณาเนื่องจากข้อมูลไม่เป็นอิสระต่อกัน



(a) สัญญาณ EEG ของคนไข้ที่สภาวะปกติ



(b) สัญญาณ EEG ของคนไข้ที่สภาวะชัก

รูป 18: สัญญาณ EEG 19 Channel ของคนไข้คนที่ 4

7.2 MATLAB codes

ใช้ `ARestLS.m` ในการประมาณแบบจำลอง AR สำหรับสัญญาณ EEG ในการทดลองในหัวข้อที่ 2.1

```
1 % for AR model :  $y(t) = c + A1 * y(t-1) + A2 * y(t-2) + \dots + Ap$   
2 %  $* y(t-p)$   
3 %  $y(t)$  is  $n \times 1$  vector  
4 %  
5 % Input:  
6 % - data is the datasets of EEG signal , data = [y(1)  
7 % y(2) ... y(N)].  
8 %  $y_x(t)$  is the value of data at time equals to t.  
9 % example:  
10 % | y1(1) y1(2) .... Y1(N) |  
11 % | y2(1) y2(2) .... y2(N) |  
12 % data = | . . . . | ; n is the  
13 % dimension of time series ,  
14 % | . . . . | N is the  
15 % number of time points in the data  
16 % | yn(1) yn(2) .... yn(N) |  
17 % - p is the lag-order of AR model.  
18 %  
19 % Output:  
20 % - c is estimated parameter of AR model  
21 % c is  $n \times 1$  vector  
22 %  
23 % - Ahat is estimated parameter of AR model.  
24 % Ahat = (A1, A2, ..., Ap)  
25 %  
26 % - Yhat is estimated of data .  
27 %  
28 % - deltaY is  $data(p+1:N) - Yhat(p+1:N)$ .  
29 %  
30 % - SIGMAhat is covariance matrix .  
31 %% write this as a function  
32 function [c, Ahat, Yhat, deltaY, SIGMAhat] = ARestLS(  
33 data, p)  
34  
35 [n, N] = size(data); %get n, N from data  
36 Y = data(:, p+1:N);  
37  
38 % create H matrix  
39 H = ones(1, N-p); %first row  
40 for i = 1:p  
41 %second row, third row and the orther  
42 H = [ H ; data(:, p+1-i:N-i) ];  
43 end  
44 %% check if  $H * H'$  is NOT invertible , return an error  
45 message  
46 if rcond(H * H') < 1e-12  
47 error('(H * transpose(H)) matrix is close to  
48 singular or badly scaled, you can try to change p  
,
```

```

49 %%
50 Aest = Y*H'/(H*H'); % Aest = [c A1 A2 ... Ap]
51 Yhatm = Aest*H;
52
53 c = Aest(:,1); % output
54 Ahat = reshape(Aest(:,2:n*p+1),n,n,p); % output
55 Yhat = [data(:,1:p) Yhatm]; % output
56 deltaY = Y-Yhatm; %Y-Yest when p+1-N % output
57 SIGMAhat = (1/(N-p))*(deltaY)*(deltaY)'; % output
58
59 end

```

Listing 1: ARestLS.m

ใช้ BICARmodel.m ในการเลือกอันดับของแบบจำลอง AR สำหรับสัญญาณ EEG ในการทดลองในหัวข้อที่ 2.1

```

1 % - deltaY is the difference matrix of data and
  %   estimated data .
2 %   example: deltaY = Y-Yest when p+1-N; Y is data ,
  %   Yest is estimated data ,
3 % - Sigma is the covariance matrix of AR model .
4 % - p is the lag-order of AR model , that you want to
  %   know score .
5
6 % Output :
7 % - BIC score
8
9 %% write this as a function
10 function [BIC] = BIC_ARmodel(deltaY , Sigma , p)
11     %% get parameter
12     [n, Nminusp] = size(deltaY); % get n ans (N-p)
13     N = Nminusp+p;
14
15     %% find right part of log-likelihood function
16     ivSigma = eye(n)/Sigma ;
17     rLogli = 0;
18     for t=1:Nminusp
19         rLogli = rLogli +(deltaY(:,t)')*ivSigma *deltaY
20         (:,t);
21     end
22
23     %% find BIC score at order p
24     L = ((N-p)/2)*log(det(ivSigma )) - (1/2)*rLogli ;
25     d = n*(n*p+1);
26     BIC = -2*L + d*log(N);
27 end

```

Listing 2: BICARmodel.m

ใช้ funcWaldTest.m ในการทดสอบค่าศูนย์ของพารามิเตอร์ของแบบจำลอง AR สำหรับสัญญาณ EEG ในการทดลองในหัวข้อที่ 3.2.1

```

1 % n is the dimension of the model output
2 % p is the lag-order of the model
3 % Avec is estimated A matrix into a vector form .
  % Avec has size n^2*p x 1.
4 % Avar is a consistent estimate of the asymptotic
  % covariance matrix of Aest

```

```

5 % alpha is a significance level of the test
6
7 % OUTPUT :
8 % show zero structure of estimated model parameter
9
10 %%
11 function [S,W]= funcWaldTest(n,p,Avec ,Avar ,alpha)
12
13 %--- Wald test ---;
14 W= zeros(n,n);
15 c= chi2inv(1-alpha ,p); % a critical value
16
17 Cvec(1:n) = c+1;
18 Cdiag = diag(Cvec);
19
20 W = W+Cdiag;
21 r = reshape(Avec ,p,n,n); % r = (B11 ,B12 ,... ,B1n)
    ;...;( Bn1 ,Bn2 ,... ,Bnn)
22
23 nsquar = 1;
24 for i=1:n
25     for j=1:n
26         rij = r(:,j,i);
27         Avarij = Avar(1+(nsquar-1)*p : nsquar*p , 1+(
    nsquar-1)*p : nsquar*p);
28
29         W(i,j) = rij'*(Avarij\rij); % Wald statistic
    value
30         nsquar = nsquar+1;
31     end
32 end
33
34 S=W>c;
35 % diagonal entries of S should be assigned to 1
36 for i=1:n
37     S(i,i)=1;
38 end
39 spy(S); title('Zero structure of the model parameter
    ');
40 end

```

Listing 3: funcWaldTest.m