

การเปรียบเทียบเทคนิคการพยากรณ์เชิงสถิติ เมื่อข้อมูลมีการเคลื่อนไหวแบบไม่เชิงเส้น

A Comparison of Nonlinear Time Series Statistical Forecasting Technics

*กนิษฐา ยิ้มนาค
Kanittha Yimnak

บทคัดย่อ

การวิจัยครั้งนี้ได้เปรียบเทียบประสิทธิภาพของเทคนิคการพยากรณ์เชิงสถิติ ได้แก่ เทคนิคการวิเคราะห์อนุกรมเวลาแบบบ็อกซ์-เจนกินส์และโครงข่ายเรเดียลเบสฟังก์ชัน เมื่ออนุกรมเวลามีรูปแบบการเคลื่อนไหวแบบไม่เชิงเส้นใน 4 รูปแบบที่มักเกิดขึ้นบ่อย ได้แก่ รูปแบบการเคลื่อนไหวแบบ

ควอดเรติก คิวบิก เอ็กซ์โปเนนเชียลและรูปแบบการเคลื่อนไหวที่ไม่มีแบบแผน สำหรับข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลที่เป็นตัวแทนการเคลื่อนไหวทั้ง 4 แบบ ได้แก่ ข้อมูลปริมาณการจำหน่ายคอมพิวเตอร์ในประเทศไทย ข้อมูลปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศไทย ข้อมูลปริมาณการนำเข้าเชื้อเพลิงธรรมชาติและสารหล่อลื่น และ ข้อมูลปริมาณการส่งออกสุกรมีชีวิตของประเทศไทย ผลการวิจัยพบว่า โครงข่ายเรเดียลเบสฟังก์ชันมีประสิทธิภาพในการพยากรณ์ดีกว่าเทคนิคการวิเคราะห์อนุกรมเวลาแบบบ็อกซ์-เจนกินส์ในกรณีที่ข้อมูลมีรูปแบบการเคลื่อนไหวแบบไม่เชิงเส้นทั้ง 4 รูปแบบ

Abstract

Objective of this research was to comparison efficiency of statistical forecasting technics between Box-Jenkins method and Radial Basis Function method for four model nonlinear time series viz model : Quadratic , Cubic , Exponential and Unpattern .The substitute data for four nonlinear time series models for this study used are the sales volume of compressor and motorcycle in Thailand , the import volume of mineral fuel and lubricant, and the export volume of pigs in Thailand.

The results of this study show that Radial Basis Function model has a suitable model and more accuracy than Box-Jenkins Method in case of model nonlinear time series

Keywords Box-jenkins Technics, Radian Basis Function

คำสำคัญ เทคนิคการวิเคราะห์อนุกรมเวลาแบบบ็อกซ์-เจนกินส์,
โครงข่ายเรเดียลเบสฟังก์ชัน

การพยากรณ์อนุกรมเวลาโดยทั่วไปจะตั้งอยู่บนสมมติฐานที่ว่าอนุกรมเวลามีลักษณะเป็นเชิงเส้นและรูปแบบการเคลื่อนไหวของข้อมูลในอนาคตมีรูปแบบเดียวกับข้อมูลในอดีต ซึ่งอธิบายได้ด้วยตัวแบบ AR(Autoregressive model) หรือตัวแบบถดถอย(Regressive model) ในการพยากรณ์ แต่ในความเป็นจริงข้อมูลส่วนมากมักมีลักษณะที่ไม่เป็นเชิงเส้น หากใช้ตัวแบบดังกล่าวอาจทำให้เกิดความคลาดเคลื่อนในการพยากรณ์

สำหรับเทคนิคการวิเคราะห์อนุกรมเวลาที่เหมาะสมกับการพยากรณ์ข้อมูลที่มีการเคลื่อนไหวแบบไม่เชิงเส้นนั้นมีหลายวิธี ทั้งเทคนิคการพยากรณ์แบบดั้งเดิม (Traditional Statistical Method) และเทคนิคการพยากรณ์อีกรูปแบบหนึ่งซึ่งเป็นที่นิยม ได้แก่ วิธีโครงข่ายประสาทเทียม (Artificial Neural Networks) ซึ่งเป็นเทคนิคการพยากรณ์แบบใหม่ (Modern Statistical Method) สำหรับเทคนิคการพยากรณ์แบบดั้งเดิมที่นิยมใช้ในการพยากรณ์ ได้แก่ เทคนิคการวิเคราะห์อนุกรมเวลาแบบบ็อกซ์และเจนกินส์ ซึ่งเทคนิคดังกล่าวสามารถใช้กับข้อมูลที่มีการเคลื่อนไหวทุกประเภทและเป็นวิธีการที่มีความแม่นยำ (accuracy) สูง แต่ทั้งนี้วิธีการพยากรณ์แบบบ็อกซ์และเจนกินส์มีข้อจำกัดในเรื่องของข้อตกลงเบื้องต้น (assumption) หลายประการ ในเรื่องของตัวแบบที่จะนำมาวิเคราะห์อนุกรมเวลาต้องเป็นตัวแบบที่สเตชันนารี (stationary time series) กล่าวคือ ต้องมีคุณสมบัติทางสถิติที่ค่าเฉลี่ยและค่าความแปรปรวน และฟังก์ชันความน่าจะเป็นของค่าสังเกต ณ เวลาต่าง ๆ กัน คงที่ไม่เปลี่ยนแปลงตามเวลาที่เปลี่ยนไป หากอนุกรมเวลาไม่สเตชันนารี ก็จะต้องมีการแปลงอนุกรมเวลาดังกล่าวให้

เป็นสเตชันนารี โดยการหาผลต่างการใช้ลอการิทึม ฯลฯ ทำให้บางครั้งวิธีการพยากรณ์ดังกล่าว อาจวิเคราะห์ได้ไม่ตรงกับสภาพความเป็นจริงของข้อมูล

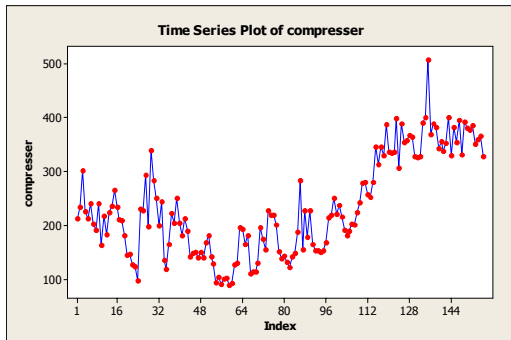
ส่วนวิธีโครงข่ายประสาทเทียม (Artificial Neural Networks) ซึ่งเป็นเทคนิคการพยากรณ์แบบใหม่ (Modern Statistical Method) มีลักษณะพิเศษ คือ เป็นวิธีที่ไม่มีข้อตกลงเบื้องต้น และสามารถวิเคราะห์ข้อมูลโดยไม่ขึ้นกับรูปแบบการแจกแจงข้อมูล นอกจากนี้เทคนิคดังกล่าวยังให้ความถูกต้องในการพยากรณ์ที่สูงกว่าวิธีการพยากรณ์แบบอื่นๆ เมื่อข้อมูลมีรูปแบบไม่เชิงเส้น (Makridakis, 1998) ปัจจุบันมีนักวิจัยอยู่จำนวนมากไม่น้อย ที่พัฒนาตัวแบบจากโครงข่ายประสาทเทียม ซึ่งเทคนิค ดังกล่าวประสบความสำเร็จค่อนข้างสูงในการนำมาใช้ในการพยากรณ์ข้อมูลที่มีลักษณะที่ไม่เป็นเชิงเส้น โดยเฉพาะโครงข่ายเรเดียลเบสฟังก์ชัน (Radial Basis Function) ซึ่งเป็นโครงข่ายประสาทเทียมแบบหนึ่ง ที่นิยมใช้กันมากในงานด้านการประมาณค่า

เนื่องจากเทคนิคการวิเคราะห์อนุกรมเวลาทั้ง 2 วิธี ได้แก่ เทคนิคการวิเคราะห์อนุกรมเวลาแบบบ็อกซ์และเจนกินส์ ซึ่งจัดเป็นเทคนิค การพยากรณ์แบบดั้งเดิม และ วิธีโครงข่ายเรเดียลเบสฟังก์ชัน ซึ่งเป็นเทคนิคการพยากรณ์แบบใหม่ ต่างก็เป็นเทคนิคที่ให้ความถูกต้องในการพยากรณ์ค่อนข้างสูงและเหมาะสมกับข้อมูลดังกล่าว จึงเป็นที่น่าสนใจที่จะทำการเปรียบเทียบวิธีการพยากรณ์ทั้ง 2 วิธี เพื่อศึกษาว่าวิธีใดที่จะให้ประสิทธิภาพในการพยากรณ์สูงกว่ากัน โดยการเปรียบเทียบประสิทธิภาพของการพยากรณ์นั้นจะตรวจสอบจากการวัดค่าความถูกต้องในการพยากรณ์ (Accuracy) โดยค่าสถิติที่ใช้วัดความถูกต้อง

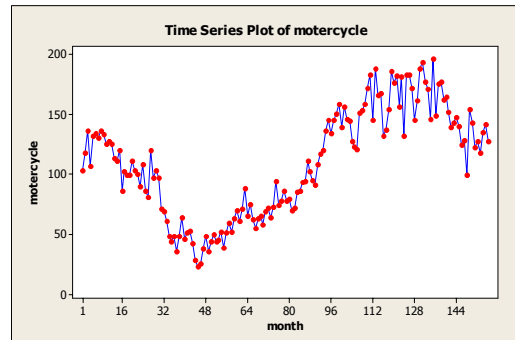
ได้แก่ รากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (RMSE)

สำหรับข้อมูลที่จะใช้ในการวิจัยครั้งนี้จะเป็นข้อมูลทางด้านเศรษฐกิจและการเกษตร ที่ทำการทดสอบแล้วว่ามีความเหมาะสมในการเคลื่อนไหวไม่เชิงเส้นในลักษณะต่าง ๆ ที่เกิดขึ้นบ่อยในประเทศไทย 4 รูปแบบ ได้แก่ รูปแบบการเคลื่อนไหวแบบควอดเรติก (Quadratic) หรือพหุนามกำลังสอง รูปแบบการเคลื่อนไหวแบบคิวบิก (Cubic) หรือพหุนามกำลังสาม รูปแบบการเคลื่อนไหวแบบเอ็กซ์โปเนนเชียล (Exponential) ซึ่งเป็นอินเวอร์สของลอการิทึม และรูปแบบการเคลื่อนไหวที่ไม่มีแบบแผน โดยในการวิจัยครั้งนี้ได้ทำการทดสอบข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวใกล้เคียงกับรูปแบบดังกล่าว ได้ข้อมูลอนุกรมเวลาที่จะนำมาการศึกษา ดังนี้ 1) ข้อมูลปริมาณการจำหน่ายคอมพิวเตอร์ในประเทศไทยตั้งแต่เดือนมกราคม 2538 ถึงเดือนธันวาคม 2550 รวม 156 เดือน

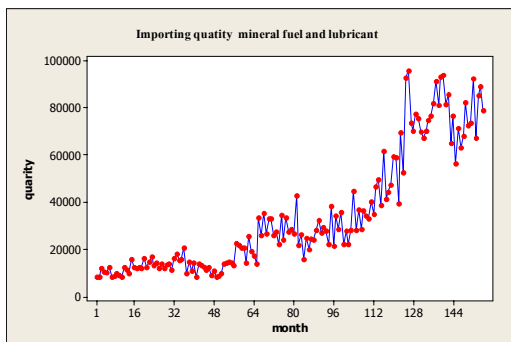
ซึ่งมีลักษณะการเคลื่อนไหวแบบควอดเรติก 2) ข้อมูลปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศไทยตั้งแต่เดือน มกราคม 2538 ถึงเดือนธันวาคม 2550 รวม 156 เดือน ซึ่งมีลักษณะการเคลื่อนไหวแบบคิวบิก 3) ข้อมูลปริมาณการนำเข้าเชื้อเพลิงธรรมชาติและสารหล่อลื่นตั้งแต่เดือนมกราคม 2538 ถึงเดือนธันวาคม 2550 รวม 156 เดือน ซึ่งมีการเคลื่อนไหวใกล้เคียงกับรูปแบบเอ็กซ์โปเนนเชียล และ 4) ข้อมูลปริมาณการส่งออกสุกรมชีวิตของประเทศไทย ตั้งแต่เดือนมกราคม 2542 ถึงเดือน ธันวาคม 2550 รวม 108 เดือน ซึ่งมีลักษณะการเคลื่อนไหวแบบไม่มีแบบแผน กล่าวคือ เมื่อทำการทดสอบสมการการเคลื่อนไหวในแบบต่าง ๆ แล้วปรากฏว่าไม่เหมาะสมกับรูปแบบการเคลื่อนไหวแบบใดเลย โดยรูปแบบการเคลื่อนไหวทั้ง 4 รูปแบบแสดงในรูปด้านล่างนี้



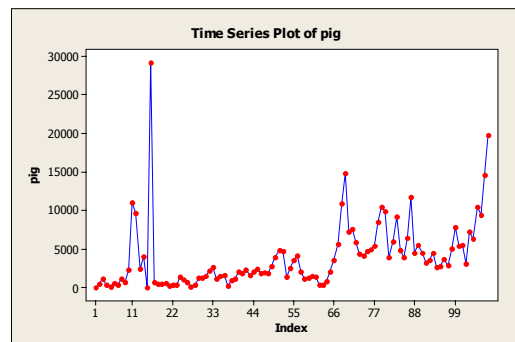
รูปที่ 1 กราฟปริมาณการจำหน่ายคอมเพรสเซอร์ในประเทศไทย



รูปที่ 2 กราฟปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศไทย



รูปที่ 3 กราฟปริมาณการนำเข้าเชื้อเพลิงธรรมชาติและสารหล่อลื่นในประเทศไทย



รูปที่ 4 กราฟปริมาณการส่งออกสุกรมีชีวิตในประเทศไทย

เทคนิคการวิเคราะห์อนุกรมเวลาแบบบ็อกซ์และเจนกินส์ (Box and Jenkins Technic)

เป็นวิธีวิเคราะห์ที่อาศัยขบวนการสโตคาสติก (Stochastic Process) โดยถือว่าข้อมูลที่เกิดขึ้นเป็นไปตามเวลาที่เปลี่ยนแปลง มีลักษณะการเกิดที่เป็นไปตามกฎความน่าจะเป็น กำหนดให้ Y_t เป็นค่าสังเกตของอนุกรมเวลา ณ เวลา t ซึ่ง Y_t มีฟังก์ชันความหนาแน่นน่าจะเป็น (Probability Density Function) เป็นแบบปกติมีค่าเฉลี่ยคือ μ_y และค่าความแปรปรวน σ_y^2 โดยที่ค่าความคลาดเคลื่อนของข้อมูล (ϵ_t) ณ เวลา t มีการแจกแจงแบบ

ปกติมีค่าเฉลี่ยเป็น 0 และมีค่าความแปรปรวน σ_{ϵ}^2 เท่ากัน

การวิเคราะห์อนุกรมเวลาเพื่อการพยากรณ์โดยวิธีบ็อกซ์และเจนกินส์นั้น ลักษณะของอนุกรมเวลาต้องเป็นอนุกรมเวลาที่เป็นสเตชันนารี (Stationary Time Series) กรณีที่อนุกรมเวลาไม่มีคุณสมบัติสเตชันนารีจะต้องแปลงอนุกรมเวลาให้มีคุณสมบัติดังกล่าวเสียก่อน ซึ่งอาจทำได้หลายวิธีเช่น การหาผลต่างของค่าสังเกตที่อยู่ติดกัน การหาลอการิทึมของค่าสังเกตในอนุกรมเวลาในกรณีที่ความแปรปรวนของอนุกรมเวลาไม่คงที่

อนุกรมเวลาที่จะนำมาวิเคราะห์โดยวิธีวิธี บ็อกซ์และเจนกินส์ เพื่อหาสมการพยากรณ์แยก ออกได้เป็น 2 กลุ่ม ได้แก่

1. อนุกรมเวลาเสถียร (Stationary Time Series) เป็นอนุกรมเวลาที่ค่าสังเกตมี คุณสมบัติทางสถิติคือ ค่าเฉลี่ย ค่าความแปรปรวน และฟังก์ชันความหนาแน่นน่าจะเป็น (Probability Density Function) ของค่าสังเกต ณ เวลาต่างๆ คงที่ กล่าวคือไม่เปลี่ยนแปลงไปตามเวลาที่เปลี่ยน ไป เช่น อนุกรมเวลา $\{Y_t\}$ ที่เป็นเสถียร จะมี ความสัมพันธ์เชิงสถิติระหว่าง $Y_t, Y_{t+1}, Y_{t+2}, \dots, Y_{t+T-1}$ ที่จุดเริ่มต้น t ไม่ต่างจากความสัมพันธ์เชิง สถิติระหว่าง $Y_{t+i}, Y_{t+i+1}, Y_{t+i+2}, \dots, Y_{t+i+T-1}$ ที่จุด เริ่มต้น $t+i$ สำหรับ $i = 1, 2, \dots, n$

2. อนุกรมเวลาไม่เป็นเสถียร (Nonstationary Time Series) เป็นอนุกรม เวลาที่ค่าสังเกตไม่มีคุณสมบัติเป็นเสถียร

การพิจารณาว่าอนุกรมเวลาเป็นเสถียรหรือไม่ นั้น จะพิจารณาจากกราฟของอนุกรมเวลา หรือพิจารณาจากคอเรลโรแกรมที่แสดงฟังก์ชันสห สัมพันธ์ในตนเอง (Autocorrelation Function : ACF) หรือ ฟังก์ชันสหสัมพันธ์บางส่วน (Partial Autocorrelation Function : PACF) ซึ่งมีความหมายดังนี้

1. ฟังก์ชันสหสัมพันธ์ในตัวเองที่ lag k (r_k) เป็นฟังก์ชันที่ใช้วัดสหสัมพันธ์ระหว่างค่าสังเกตใน อนุกรมเวลาที่ห่างกัน k ช่วงเวลา จะใช้ r_k แทน ฟังก์ชันสหสัมพันธ์ในตัวเองของตัวอย่างที่ lag k r_k เป็นตัวประมาณของ ρ_k ซึ่งเป็นฟังก์ชัน สหสัมพันธ์ในตัวเองของประชากรที่ lag k r_k มีคุณสมบัติดังนี้ r_k มีค่าอยู่ระหว่าง -1 ถึง 1 และ ฟังก์ชันสหสัมพันธ์ในตัวเองระหว่าง Y_t และ Y_{t+k} จะไม่ต่างจากฟังก์ชันสหสัมพันธ์ในตัวเองระหว่าง Y_t และ Y_{t-k} นั่นคือ $r_k = r_{-k}$ จะคำนวณค่า r_k ได้จาก

$$r_k = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}; k = 1, 2, \dots \quad (1)$$

เมื่อ $\bar{Y} = \sum_{t=1}^n \frac{Y_t}{n}$ เป็นค่าเฉลี่ยของอนุกรมเวลา

n เป็นขนาดของอนุกรมเวลา

ค่าของฟังก์ชันสหสัมพันธ์ในตัวเอง r_k มีค่า เข้าใกล้ $+1$ แสดงว่าค่าสังเกตที่อยู่ห่างกัน k ช่วง เวลา มีสหสัมพันธ์กันมาก และสหสัมพันธ์มีทิศทางตามกัน กล่าวคือ ถ้าค่า Y_t มีค่าสูง ค่า Y_{t+k} จะสูงด้วย หรือถ้าค่า Y_t มีค่าต่ำ ค่า Y_{t+k} จะต่ำ ด้วย เมื่อ r_k เข้าใกล้ -1 แสดงว่าค่าสังเกตที่อยู่

ห่างกัน k ช่วงเวลานั้น มีสหสัมพันธ์กันมากแต่มี ทิศทางตรงข้ามกัน กล่าวคือ ถ้าค่า Y_t มีค่าสูง ค่า Y_{t+k} จะมีค่าต่ำหรือถ้าค่า Y_t มีค่าต่ำ ค่า Y_{t+k} จะมี ค่าสูง เมื่อ r_k มีค่าเข้าใกล้ 0 แสดงว่า ค่าสังเกตที่ ห่างกัน k ช่วงเวลา มีสหสัมพันธ์กันน้อยมาก

อนุกรมเวลาสเตชันนารี จะมีคอเรโลแกรม ซึ่งเป็นกราฟที่เกิดจากการพล็อตระหว่างค่าของ k และ r_k ถ้าค่า r_k มีลักษณะลดลง (die down) อย่างรวดเร็ว เมื่อ k มีค่าเพิ่มขึ้นหรือค่าเข้าใกล้ 0 ที่ระยะห่างช่วงใดช่วงหนึ่งจากอนุกรมเวลา $\{Y_t\}$ ค่า r_k ที่คำนวณได้อาจจะใกล้ 0 มากจนกล่าวได้ว่ามาจากอนุกรมเวลาที่มี $\rho_k = 0$ การกำหนดว่า r_k มาจากอนุกรมเวลาที่มี ρ_k มีค่าเป็น 0 หรือไม่นั้น จะใช้การทดสอบสมมติฐาน โดยตั้งสมมติฐานหลักและสมมติฐานรองดังนี้

$$\begin{aligned} H_0 &: \rho_k = 0 \\ H_1 &: \rho_k \neq 0 \end{aligned}$$

$$r_{kk} = \begin{cases} r_1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} \end{cases} \quad \text{สำหรับ } k = 2, 3, \dots \quad (2)$$

$$\text{เมื่อ } r_{k,j} = r_{k-1,j} - r_{kk} r_{k-1,k-j}$$

$$\text{สำหรับ } j = 1, 2, \dots, k-1$$

อนุกรมเวลาสเตชันนารีจะมีคอเรโลแกรมที่เป็นกราฟที่เกิดจากการพล็อตของ k และ r_{kk} ค่า r_{kk} จะมีลักษณะลดลงอย่างรวดเร็ว เมื่อค่าสังเกตของอนุกรมเวลามีระยะห่างกันเพิ่มขึ้น หรือ r_{kk} มีค่าเข้าใกล้ 0 ที่ระยะห่าง k มีค่าเพิ่มขึ้น ค่า r_{kk} ที่คำนวณได้อาจจะเข้าใกล้ 0 มากจนกล่าวได้ว่ามาจากอนุกรมเวลาที่มี $\rho_{kk} = 0$ การพิจารณาว่าการกำหนดว่า r_{kk} มาจากอนุกรมเวลาที่มี ρ_{kk} มีค่าเป็น 0 หรือไม่นั้น จะใช้การทดสอบสมมติฐาน โดยตั้งสมมติฐานหลักและสมมติฐานรองดังนี้

$$\begin{aligned} H_0 &: \rho_{kk} = 0 \\ H_1 &: \rho_{kk} \neq 0 \end{aligned}$$

การทดสอบสมมติฐานจะต้องพิจารณาว่า r_{kk} อยู่ในช่วง ± 2 เท่าของค่าเบี่ยงเบนมาตรฐานของ r_{kk} หรือไม่ ถ้าอยู่ในช่วงดังกล่าวจะยอมรับ H_0 แสดงว่า $\rho_{kk} = 0$ ซึ่ง 2 เป็นค่าประมาณของ $Z_{.025}$ และค่าเบี่ยงเบนมาตรฐานของ r_{kk} ประมาณด้วยขั้นตอนการพยากรณ์อนุกรมเวลาแบบบ็อกซ์และเจนกินส์ แบ่งเป็น 4 ขั้นตอน มีรายละเอียดดังนี้

ขั้นที่ 1 กำหนดตัวแบบ (Identification)
ก่อนการกำหนดตัวแบบให้กับอนุกรมเวลา $\{Y_t\}$ จะต้องพิจารณาว่าอนุกรมเวลานั้นมีคุณสมบัติเป็นสเตชันนารีหรือไม่ หากพบว่าอนุกรมเวลาไม่เป็นสเตชันนารีต้องแปลงอนุกรมเวลาใหม่ให้เป็นสเตชันนารีก่อน แล้วจึงหาตัวแบบที่เหมาะสมให้กับอนุกรมเวลาใหม่

ตัวแบบอนุกรมเวลาที่ใช้ในการพยากรณ์แบบบ็อกซ์และเจนกินส์เมื่ออนุกรมเวลาที่มีคุณสมบัติเป็นสเตชันนารี มีดังนี้

1. ตัวแบบออโตรีเกรสซีฟอันดับที่ p (Autoregressive Model of Order p) : $AR(p)$ ซึ่ง p คือ อันดับของออโตรีเกรสซีฟ มีตัวแบบทั่วไปดังนี้

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (3)$$

โดยที่ Y_t คือ ค่าสังเกตอนุกรมเวลา ณ เวลา t
 δ คือ ค่าคงที่
 $\phi_1, \phi_2, \dots, \phi_p$ คือ ค่าพารามิเตอร์ของออโตรีเกรสซีฟ (Autoregressive Parameters)
 ε_t คือ ค่าความคลาดเคลื่อน ณ เวลา t มีการแจกแจงแบบปกติ ซึ่งมีค่าเฉลี่ยเท่ากับ 0 ค่าความแปรปรวนเท่ากับ σ^2

ตัวอย่างของตัวแบบออโตรีเกรสซีฟอันดับที่ p

1. ตัวแบบ $AR(1)$

$$Y_t = \delta + \phi_1 Y_{t-1} + \varepsilon_t$$

มี $|\phi_1| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลาที่มีคุณสมบัติสเตชันนารี

2. ตัวแบบ $AR(2)$

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

เมื่อ $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$ และ $|\phi_1| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลาที่มีคุณสมบัติสเตชันนารี

2. ตัวแบบค่าเฉลี่ยเคลื่อนที่อันดับที่ q (Moving Average Model of Order q) : MA(q)
ซึ่ง q คืออันดับที่ของตัวแบบค่าเฉลี่ยเคลื่อนที่ มีตัวแบบทั่วไปดังนี้

$$Y_t = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (4)$$

โดยที่ Y_t คือ ค่าสังเกตของอนุกรมเวลา ณ เวลา t
 δ คือ ค่าคงที่

$\theta_1, \theta_2, \dots, \theta_q$ คือ พารามิเตอร์ของค่าเฉลี่ยเคลื่อนที่ (Moving Average Parameters)

ε_t คือ ค่าความคลาดเคลื่อน ณ เวลา t มีการแจกแจงแบบปกติ มีค่าเฉลี่ยเท่ากับ 0 ค่าความแปรปรวนเท่ากับ

ตัวอย่างของตัวแบบค่าเฉลี่ยเคลื่อนที่อันดับที่ q

1. ตัวแบบ MA(1)

$$Y_t = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

เมื่อ $|\theta_1| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลามีคุณสมบัติเป็นอินเวอร์ติเบิล (Invertible) ซึ่งเป็นคุณสมบัติที่ทำให้เขียน ε_t ในเทอมของ $Y_t, Y_{t+1}, Y_{t+2}, \dots$ ได้ และประมาณค่าความคลาดเคลื่อนของ ε_t ได้

2. ตัวแบบ MA(2)

$$Y_t = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

เมื่อ $\theta_1 + \theta_2 < 1$, $\theta_2 - \theta_1 < 1$ และ $|\theta_2| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลามีคุณสมบัติเป็นอินเวอร์ติเบิล (Invertible)

3. ตัวแบบผสมค่าเฉลี่ยเคลื่อนที่และออโตรีเกรสซีฟอันดับที่ p และ q (Mixed Autoregressive and Moving Average Model of Order p and q) : ARMA (p, q) มีตัวแบบทั่วไปดังนี้

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (5)$$

ตัวอย่างของตัวแบบผสมค่าเฉลี่ยเคลื่อนที่และออโตรีเกรสซีฟอันดับที่ p และ q

1. ARMA (p,q) ที่มี $p = 1$, $q = 1$ หรือ ARMA (1,1)

$$Y_t = \delta + \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

เมื่อ $|\phi_1| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลามีคุณสมบัติสแตชันนารี

และ $|\theta_1| < 1$ เป็นเงื่อนไขที่ทำให้อนุกรมเวลามีคุณสมบัติเป็นอินเวอร์ติเบิล (Invertible)

กรณีที่อนุกรมเวลาไม่เป็นสแตชันนารี การหาตัวแบบที่เหมาะสมกับอนุกรมเวลานั้นจะต้องแปลงอนุกรมเวลาให้เป็นสแตชันนารีก่อน ซึ่งทำได้หลายวิธี วิธีที่นิยมใช้กันมากได้แก่ การหาผลต่าง (Differencing) ของอนุกรมเวลาเดิม ผลต่างของอนุกรมเวลานั้นจะทำให้ครั้งขึ้นอยู่กับว่าอนุกรมเวลาผลต่างนั้นมีคุณสมบัติสแตชันนารีแล้วหรือยัง ถ้าไม่ก็ต้องทำต่อไปจนเป็นสแตชันนารีแล้วจึงนำอนุกรมเวลาที่มีคุณสมบัติสแตชันนารี ไปหาตัวแบบที่เหมาะสมต่อไป

จากอนุกรมเวลา $\{Y_t\}$ หาผลต่างอนุกรมเวลา $\{Z_t\}$ จะใช้การกระทำ ∇ เพื่ออธิบายการหาผลต่างของอนุกรมเวลาเดิมดังนี้

$$\begin{aligned} \text{ผลต่างครั้งที่ 1} \quad Z_t &= \nabla Y_t \\ &= Y_t - Y_{t-1} \end{aligned}$$

$$\begin{aligned} \text{ผลต่างครั้งที่ 1} \quad Z_t &= \nabla^2 Y_t \\ &= Y_t - 2Y_{t-1} + Y_{t-2} \end{aligned}$$

ตัวแบบอนุกรมเวลาที่ต้องการหาผลต่างของอนุกรมเวลาเพื่อให้มีคุณสมบัติสแตชันนารี โดยกำหนดอันดับของการหาผลต่าง ($d > 1$) เป็นตัวแบบอินทิเกรทเฉลี่ยเคลื่อนที่และออโตรีเกรสซีฟ (Autoregressive Integrated Moving Average) : ARIMA (p,d,q) โดย d เป็นอันดับที่ผลต่างมีตัวแบบทั่วไปดังนี้

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

เมื่อ Z_t คือ ค่าสังเกต ณ เวลา t จากอนุกรมเวลา
 δ คือ ค่าคงที่

$\phi_1, \phi_2, \dots, \phi_p$ คือ พารามิเตอร์ออโตรีเกรสซีฟ (Autoregressive Parameters)

$\theta_1, \theta_2, \dots, \theta_q$ คือ พารามิเตอร์ของค่าเฉลี่ยเคลื่อนที่ (Moving Average Parameters)

ε_t คือ ค่าความคลาดเคลื่อน ณ เวลา t มีการแจกแจงแบบปกติมี ค่าเฉลี่ยเท่ากับ 0 ค่าความแปรปรวนเท่ากับ σ^2

ตัวอย่างของตัวแบบอินทีเกรทเฉลี่ยเคลื่อนที่และออโตรีเกรทชีฟ

1. ARIMA (0,1,1) หรือ IMA (1,1) มีตัวแบบ

$$Z_t = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad \text{เมื่อ } |\theta_1| < 1$$

$$Y_t = \delta + Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

2. ARIMA (1,1,0) หรือ ARI (1,1) มีตัวแบบ

$$Z_t = \delta - \phi_1 Z_{t-1} + \varepsilon_t \quad \text{เมื่อ } |\phi_1| < 1$$

$$Y_t = \delta + (1+\phi_1)Y_{t-1} - \phi_1 Y_{t-2} + \varepsilon_t$$

3. ARIMA (1,1,1) มีตัวแบบ

$$Z_t = \delta - \phi_1 Z_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad \text{เมื่อ } |\theta_1| < 1 \text{ และ } |\phi_1| < 1$$

$$Y_t = \delta + (1+\phi_1)Y_{t-1} - \phi_1 Y_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

4. ARIMA (0,1,0) มีตัวแบบ

$$Z_t = \delta + \varepsilon_t$$

$$Y_t = \delta + Y_{t-1} + \varepsilon_t$$

ถ้านุกรมเวลาที่มีทั้งแนวโน้มและความผันแปรเนื่องจากฤดูกาลเข้ามาเกี่ยวข้อง จะกำหนดตัวแบบ SARIMA (P,D,Q)_L (Seasonal Integrated Autoregressive and Moving Average) ที่ Oder (P,D,Q) โดย P เป็นอันดับของ SAR (Seasonal Autoregressive) Q เป็นอันดับของ SMA (Seasonal Moving Average) และ D เป็นจำนวนครั้งที่หาผลต่างฤดูกาลเพื่อทำให้นุกรมเวลา $\{Y_t\}$ ที่ไม่เป็นสเตชันนารีเนื่องจากฤดูกาลเป็นนุกรมเวลาชุดใหม่ $\{Z_t\}$ ที่มีคุณสมบัติเป็นสเตชันนารี โดย

$$Z_t = \nabla_L^D Y_t$$

เช่น สำหรับอนุกรมเวลารายเดือน $L = 12$

$$D = 1, \quad Z_t = \nabla_{12}^1 Y_t \\ = Y_t - Y_{t-12}$$

$$D = 2, \quad Z_t = \nabla_{12}^2 Y_t \\ = \nabla_{12} (Y_t - Y_{t-12}) \\ = Y_t - 2Y_{t-12} + Y_{t-24}$$

ตัวอย่างของตัวแบบอนุกรมเวลา SARIMA (P,D,Q)_L เช่น

อนุกรมเวลารายเดือนที่มีตัวแบบเป็น SARIMA (0,1,1)₁₂ จะเขียนตัวแบบเป็น

$$Z_t = \varepsilon_t - \Theta_1 \varepsilon_{t-12} \quad \text{เมื่อ } |\Theta_1| < 1$$

$$\text{หรือ } Y_t = Y_{t-12} + \varepsilon_t - \Theta_1 \varepsilon_{t-12}$$

โดยที่ Y_t คือ ค่าสังเกต ณ เวลา t และ Y_{t-12} คือค่าสังเกตที่อยู่ก่อนเวลา t 12 ช่วงเวลา

Θ_1 คือ พารามิเตอร์ของค่าเฉลี่ยเคลื่อนที่ฤดูกาล (Seasonal Moving Average Parameters)

ขั้นที่ 2 การประมาณค่าพารามิเตอร์ (Parameters Estimation)

การประมาณค่าพารามิเตอร์ในตัวแบบจะใช้การวิเคราะห์ตัวเลข (Numerical Analysis) ซึ่งจะต้องประมาณค่าพารามิเตอร์ในตัวแบบด้วยวิธีกำลังสองน้อยสุด (Least Squares Method) โดยกำหนดค่าประมาณเบื้องต้น (Initial Estimate) ของพารามิเตอร์เพื่อหาค่าประมาณสุดท้าย (Final Estimate) โดยการซ้ำ (Iteration) ซึ่งค่าประมาณที่ทำให้ผลรวมกำลังสองของความคลาดเคลื่อน ($\sum e_i^2$) มีค่าต่ำสุดจะเป็นค่าประมาณที่ใช้ในสมการพยากรณ์

ขั้นที่ 3 การตรวจสอบความเหมาะสมของ ตัวแบบ (Diagnostic Checking)

เมื่อกำหนดตัวแบบและประมาณค่าพารามิเตอร์แล้ว ต้องมีการตรวจสอบว่าตัวแบบที่เลือกไว้มีความเหมาะสมกับอนุกรมเวลาหรือไม่ การตรวจสอบความเหมาะสมของตัวแบบมีหลายวิธี แต่ละวิธีอาจให้ผลสรุปไม่เหมือนกัน ดังนั้นจึงควรใช้การตรวจสอบความเหมาะสมหลายวิธีพร้อมกัน

ในการตรวจสอบความเหมาะสมของตัวแบบ วิธีที่นิยมใช้กันมากได้แก่

1. การทดสอบของบ็อกซ์และจุง (Box-Ljung) เพื่อตรวจสอบความคลาดเคลื่อนของการพยากรณ์ (e_t) ว่าเป็นอิสระกันหรือไม่ ด้วยการทดสอบสมมติฐาน

$$H_0 : \rho_1(e_t) = \dots = \rho_m(e_t) = 0$$

$$H_1 : \rho_k(e_t) \text{ บางค่าไม่เท่ากับ } 0 \text{ สำหรับ } k = 1, 2, \dots, m$$

โดยใช้ตัวทดสอบสถิติ

$$Q = n(n+2) \sum_{i=1}^m \frac{r_k^2(e_t)}{n-k}$$

เมื่อ n คือ จำนวนข้อมูลที่จะใช้ในการวิเคราะห์อนุกรมเวลา

M คือ lag สูงสุดที่ต้องการทดสอบ

$r_k(e_t)$ คือ ค่าฟังก์ชันสหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อนที่อยู่ห่างกัน k ช่วงเวลา ตัวทดสอบสถิติ Q มีการแจกแจงแบบไคสแควร์ ที่มีองศาความเป็นอิสระเท่ากับ $m-n_p$ ซึ่ง

n_p คือ จำนวนพารามิเตอร์ในตัวแบบ

ตัวทดสอบสถิติ Q มีการแจกแจงแบบไคสแควร์ ที่มีองศาความเป็นอิสระเท่ากับ $m-n_p$

ซึ่ง n_p คือ จำนวนพารามิเตอร์ในตัวแบบ

จะปฏิเสธ H_0 เมื่อ แสดงว่าตัวแบบ ARMA(p,q) ที่กำหนดยังไม่เหมาะสม

2. พิจารณาค่าความคลาดเคลื่อนของการพยากรณ์ (e_t) ว่ามีสหสัมพันธ์กันหรือไม่ด้วยการทดสอบสมมติฐาน

$$H_0 : \rho_k(e_t) = 0$$

$$H_1 : \rho_k(e_t) \neq 0$$

จะปฏิเสธ H_0 ถ้า $r_k(e_t)$ ตกอยู่นอกช่วง 2 เท่าของส่วนเบี่ยงเบนมาตรฐานของ สรุปว่าตัวแบบ ARMA(p,q) ที่กำหนดยังไม่เหมาะสม

ขั้นที่ 4 การพยากรณ์ (Forecasting)

วิธีบ็อกซ์และเจนกินส์จะให้สมการพยากรณ์ที่จะใช้หาค่าพยากรณ์ล่วงหน้าได้หลายช่วงเวลา สมการพยากรณ์จะสร้างจากตัวแบบที่ได้ทดสอบแล้วว่าเหมาะสมกับการพยากรณ์ โดยส่วนใหญ่ค่าพยากรณ์ที่ได้จากสมการพยากรณ์ หากเป็นการพยากรณ์ระยะยาว ค่าพยากรณ์นั้นจะให้สาระจากข้อมูลจริงน้อยลง และความแม่นยำที่ได้จากการพยากรณ์ก็จะมีค่าลดลงกว่าการพยากรณ์ระยะสั้น

โครงข่ายเรเดียลเบซิสฟังก์ชัน (Radial Basis Function)

โครงข่ายเรเดียลเบซิสฟังก์ชัน เป็นโครงข่ายประสาทเทียมแบบหนึ่งที่ยอมรับใช้กันมากในงานด้านการประมาณค่าหรือการพยากรณ์ข้อมูล โครงข่ายเรเดียลเบซิสฟังก์ชันเป็นแบบฟีดฟอร์เวิร์ดเน็ตเวิร์คที่มีจำนวน 3 ชั้นได้แก่ ชั้นอินพุต ชั้นซ่อน และชั้นเอาต์พุต ซึ่งชั้นซ่อนจะมีเพียง 1 ชั้นเท่านั้น เอาต์พุตของโครงข่ายเรเดียลเบซิสฟังก์ชันสามารถเขียนเป็นสมการทางคณิตศาสตร์ ได้ดังนี้

$$f_r(x) = w_0 + \sum_{i=1}^{n_r} w_i \phi(\|x - c_i\|)$$

เมื่อ $f_r(\cdot)$ คือ เอาต์พุตของโครงข่าย

x คือ อินพุตเวกเตอร์ และ $x = [x_1 \ x_2 \ \dots \ x_p]^T$

p คือ ไคเมนชัน (Dimension) ของอินพุตเวกเตอร์

$\phi(\cdot)$ คือ เบซิสฟังก์ชัน (Basis Function) จาก R^+ ไปยัง R

$\|\cdot\|$ คือ Euclidean Norm

n_r คือ จำนวนเซ็นเตอร์

$c_i \in R^n$, $1 \leq i \leq n_r$ คือ เซ็นเตอร์ของโครงข่าย

w_i , $1 \leq i \leq n_r$ คือ ค่าน้ำหนักที่เชื่อมต่อระหว่างนิวรอนที่ i ในชั้นซ่อน กับนิวรอนของเอาต์พุต

เบซิสฟังก์ชัน $\phi(\cdot)$ มีให้เลือกได้หลายรูปแบบ ขึ้นอยู่กับความเหมาะสมของการนำไปใช้งาน เช่น

1. The Gaussian function

$$\varphi(x) = \exp\left(\frac{v^2}{\beta^2}\right)$$

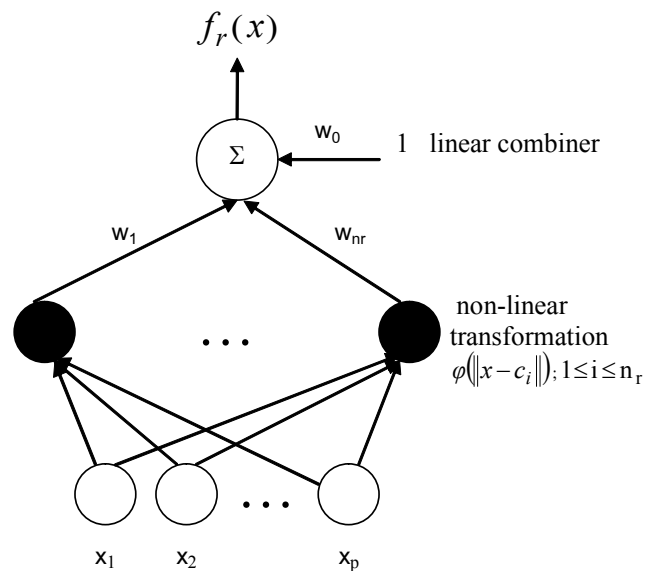
2. The Thin - Plate - Spline Function

$$\varphi(x) = v^2 \log v$$

3. The Multiquadratic Function

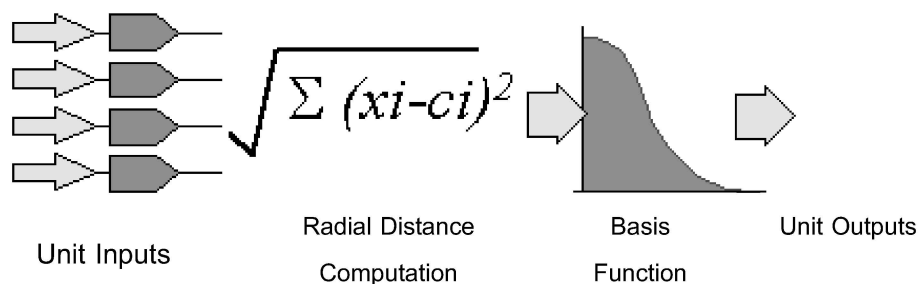
$$\varphi(x) = \exp\left(\frac{v^2}{\beta^2}\right)^{\frac{1}{2}}$$

โครงข่ายเรเดียลเบสฟังก์ชันสามารถเขียนในลักษณะของโครงสร้างทางสถาปัตยกรรมดังรูปที่ 1 ซึ่งประกอบไปด้วย ชั้นอินพุต ชั้นซ่อน และชั้นเอาต์พุต โดยมีชั้นซ่อนจะมีเพียง 1 ชั้นเท่านั้น



รูปที่ 5 ลักษณะโครงสร้างของโครงข่ายเรเดียลเบสฟังก์ชัน

ข้อมูลเอาต์พุตของชั้นซ่อน จะได้มาจากเบสฟังก์ชันของระยะห่าง(Radial Distance) ระหว่างข้อมูลอินพุตและเซ็นเตอร์ แสดงดังรูปที่ 2



รูปที่ 6 ลักษณะการทำงานในชั้นซ่อนของโครงข่ายเรเดียลเบสซิสฟังก์ชัน

อัลกอริทึมในการเรียนรู้แบบ LMS (Least Mean Squares)

อัลกอริทึมในการเรียนรู้ของโครงข่าย RBF มีด้วยกันหลาย ๆ อัลกอริทึม ในที่นี้จะกล่าวถึงเฉพาะอัลกอริทึมที่ใช้ในงานวิจัยฉบับนี้ คือ อัลกอริทึมแบบ LMS

เมื่อกำหนดให้ $w(n) = [w_0 \ w_1 \dots w_{N-1}]^T$ คือ เวกเตอร์น้ำหนัก (Weight)

$\bar{x}(n) = [x(n) \ x(n-1) \dots x(n-N+1)]^T$ คือ ข้อมูลนำเข้า (Input)

$y(n) = w^T \bar{x}(n)$ คือ ผลลัพธ์ (Output)

$e(n) = d(n) - y(n)$ คือ ความคลาดเคลื่อน

การคำนวณหาค่าน้ำหนักจะคำนวณได้จากสมการ

$$w(n+1) = w(n) + \frac{1}{2} \mu [-\nabla E(e^2(n))] \quad (8)$$

เมื่อ μ คือ ค่าพารามิเตอร์ของการปรับช่วงก้าว (step-size) โดยค่า μ จะเป็นค่าที่กำหนดขึ้น ซึ่งค่า μ ที่เหมาะสมจะมีค่าประมาณ $\mu \leq \frac{1}{3\text{trace}(R)}$

เพราะฉะนั้น $e^2(n) = [d(n) - w^T \bar{x}(n)]^2$

$$\text{โดยที่} \quad \nabla_w (E \{e^2(n)\}) = -2r + 2R w(n)$$

เมื่อ R คือ เมทริกซ์ความแปรปรวน ความแปรปรวนร่วม

$$R(n) = \bar{x}(n) \bar{x}^T(n)$$

$$r(n) = d(n) \bar{x}(n)$$

ดังนั้นค่าน้ำหนักที่ต้องการจึงได้มาจากสมการ

$$\begin{aligned} w(n+1) &= w(n) + \mu \bar{x}(n) [d(n) - \bar{x}^T(n) w(n)] \\ &= w(n) + \mu \bar{x}(n) e(n) \end{aligned}$$

โดยที่ ค่า $w(0)$ จะเป็นเวกเตอร์น้ำหนักที่กำหนดขึ้น เมื่อ $n = 0$ ที่จะทำให้ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (MSE) มีค่าน้อย

วิจารณ์ผล

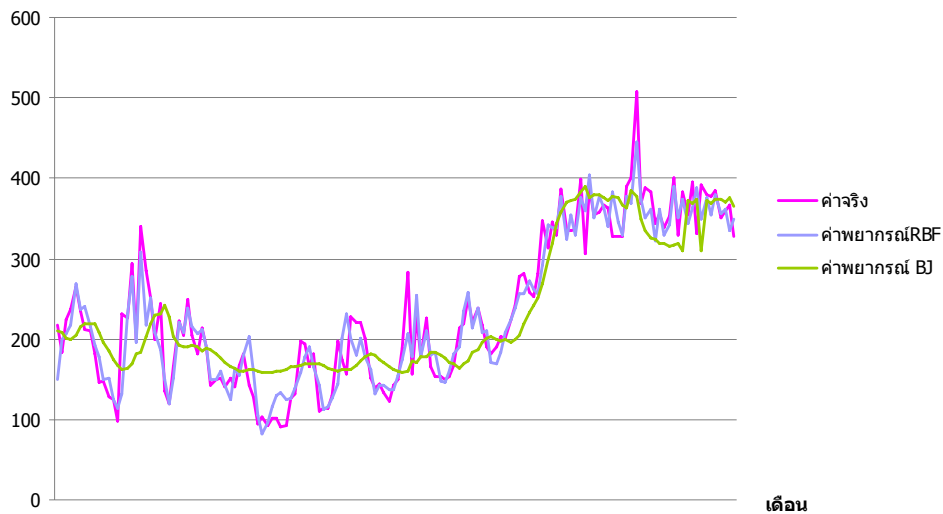
ผลการพยากรณ์ข้อมูลจะนำเสนอด้วยกราฟ และผลการเปรียบเทียบประสิทธิภาพของการวิเคราะห์อนุกรมเวลาด้วยเทคนิคบ็อกซ์และเจนกินส์ และ เทคนิควิธีโครข่ายเรเดียลเบชีสฟังก์ชัน

ในการวัดค่าความถูกต้องระหว่างค่าจริงและค่าพยากรณ์โดยใช้ตัวสถิติ RMSE ในการตรวจสอบความถูกต้องได้ผลการวิเคราะห์ข้อมูลดังนี้

1. ผลการพยากรณ์ด้วยกราฟของข้อมูลอนุกรมเวลาทั้ง 4 แบบด้วย

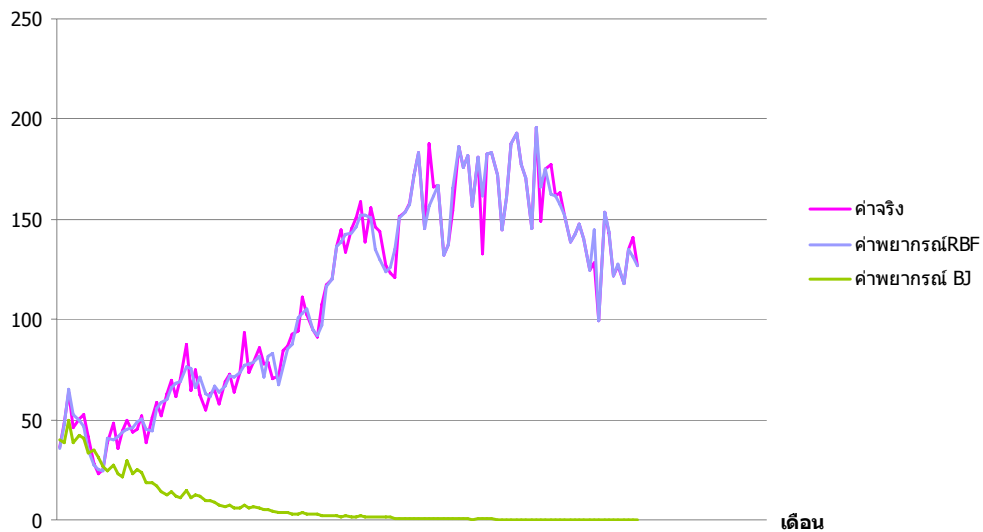
1.1 ผลการเปรียบเทียบพยากรณ์ด้วยกราฟของข้อมูลอนุกรมเวลาทั้ง 4 แบบด้วยวิธีโครข่ายเรเดียลเบชีสฟังก์ชัน

กราฟเปรียบเทียบค่าจริงและค่าพยากรณ์ของการจำหน่ายคอมพิวเตอร์ในประเทศไทย



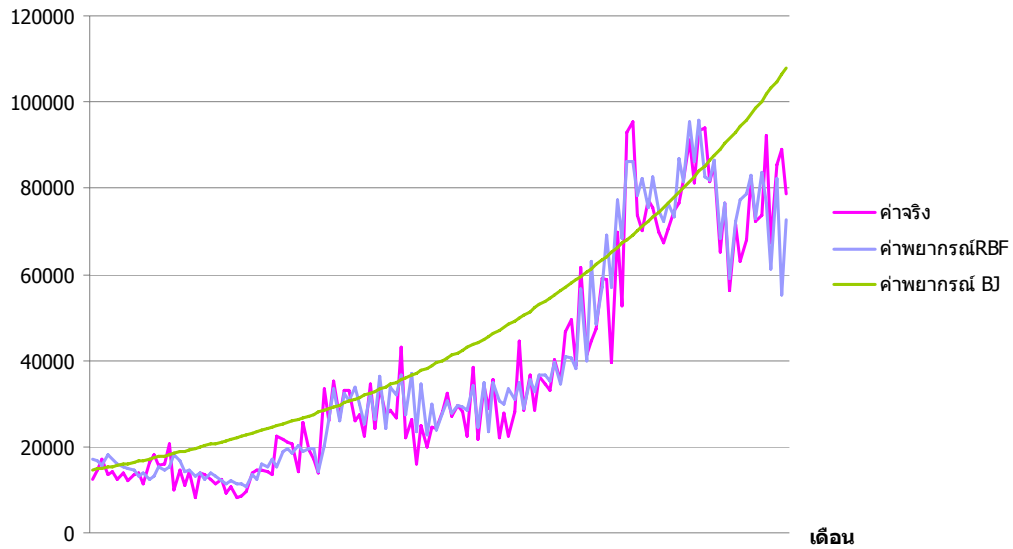
รูปที่ 7 ผลการพยากรณ์ปริมาณการจำหน่ายคอมพิวเตอร์ในประเทศไทย

กราฟเปรียบเทียบค่าจริงและค่าพยากรณ์ของปริมาณการ
จำหน่ายรถจักรยานยนต์ในประเทศไทย



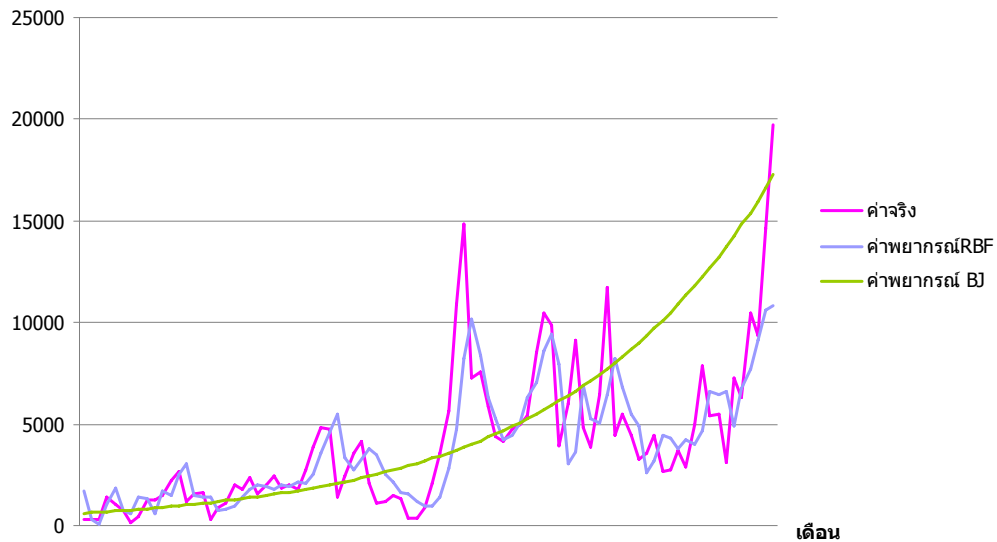
รูปที่ 8 ผลการพยากรณ์ปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศไทย

กราฟเปรียบเทียบค่าจริงและค่าพยากรณ์ของปริมาณการนำเข้า
เชื้อเพลิงธรรมชาติและสารหล่อลื่นในประเทศไทย



รูปที่ 9 ผลการพยากรณ์ปริมาณการนำเข้าเชื้อเพลิงธรรมชาติและสารหล่อลื่นในประเทศไทย

กราฟเปรียบเทียบค่าจริงและค่าพยากรณ์ของปริมาณการส่งออกสุกรมชีวิตในประเทศไทย



รูปที่ 10 ผลการพยากรณ์ปริมาณการส่งออกสุกรมชีวิตในประเทศไทย

1.2 ผลการเปรียบเทียบประสิทธิภาพของการวิเคราะห์อนุกรมเวลาด้วย เทคนิคบ็อกซ์และเจนกินส์ และ เทคนิควิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน

ตารางที่ 1 แสดงการเปรียบเทียบประสิทธิภาพของการพยากรณ์ด้วยเทคนิคบ็อกซ์และเจนกินส์ และ วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน ของข้อมูลปริมาณการจำหน่ายคอมพิวเตอร์ในประเทศไทย

ค่า	ตัวแบบ		อัตราส่วนเปรียบเทียบ ค่าความคลาดเคลื่อน ของทั้ง 2 วิธี
	เทคนิคบ็อกซ์ และเจนกินส์	วิธีโครงข่าย เรเดียลเบชีสฟังก์ชัน	
RMSE	45.855	25.940	1.77 : 1

ตารางที่ 1 แสดงผลการเปรียบเทียบค่าพยากรณ์ของปริมาณการจำหน่ายคอมพิวเตอร์ในประเทศไทย ระหว่างวิธีบ็อกซ์และเจนกินส์ และ วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน พบว่า วิธีโครงข่ายเรเดียลเบชีสฟังก์ชันมีประสิทธิภาพในการพยากรณ์มากกว่าเทคนิคบ็อกซ์และเจนกินส์โดยให้ค่า

RMSE ต่ำกว่าวิธีเทคนิคบ็อกซ์และเจนกินส์ และเมื่อพิจารณาอัตราส่วนเปรียบเทียบค่าความคลาดเคลื่อนระหว่างทั้ง 2 วิธี พบว่า เทคนิคบ็อกซ์และเจนกินส์ ให้ค่าความคลาดเคลื่อนสูงกว่า วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน 1.77 เท่า

ตารางที่ 2 แสดงการเปรียบเทียบประสิทธิภาพของการพยากรณ์ด้วย เทคนิคบ็อกซ์และเจนกินส์ และ วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน ของข้อมูลปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศไทย

ค่า	ตัวแบบ		อัตราส่วนเปรียบเทียบ ค่าความคลาดเคลื่อน ของทั้ง 2 วิธี
	เทคนิคบ็อกซ์และเจนกินส์	วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน	
RMSE	119.007	6.907	17.23 : 1

ตารางที่ 2 แสดงผลการเปรียบเทียบค่าพยากรณ์ของปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศไทย ระหว่างวิธีบ็อกซ์และเจนกินส์ และ วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน พบว่า วิธีโครงข่ายเรเดียลเบชีสฟังก์ชันมีประสิทธิภาพในการพยากรณ์มากกว่าวิธีเทคนิคบ็อกซ์และเจนกินส์โดยให้ค่า

RMSE ต่ำกว่าเทคนิคบ็อกซ์และเจนกินส์ และเมื่อพิจารณาอัตราส่วนเปรียบเทียบค่าความคลาดเคลื่อนระหว่างทั้ง 2 วิธี พบว่า เทคนิคบ็อกซ์และเจนกินส์ ให้ค่าความคลาดเคลื่อนสูงกว่า วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน 17.23 เท่า

ตารางที่ 3 แสดงการเปรียบเทียบประสิทธิภาพของการพยากรณ์ด้วย เทคนิคบ็อกซ์และเจนกินส์ และ วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน ของข้อมูลปริมาณการนำเข้าเชื้อเพลิงธรรมชาติและสารหล่อลื่นในประเทศไทย

ค่า	ตัวแบบ		อัตราส่วนเปรียบเทียบ ค่าความคลาดเคลื่อน ของทั้ง 2 วิธี
	เทคนิคบ็อกซ์และเจนกินส์	วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน	
RMSE	12,974.734	6,333.984	2.05 : 1

ตารางที่ 3 แสดงผลการเปรียบเทียบค่าพยากรณ์ของข้อมูลปริมาณการนำเข้าเชื้อเพลิงธรรมชาติและสารหล่อลื่นในประเทศไทยระหว่างวิธีบ็อกซ์และเจนกินส์ และ วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน พบว่า วิธีโครงข่ายเรเดียลเบชีสฟังก์ชันมีประสิทธิภาพในการพยากรณ์มากกว่าวิธีเทคนิคบ็อกซ์และเจนกินส์โดยให้ค่า RMSE ต่ำกว่าวิธีเทคนิคบ็อกซ์และเจนกินส์ และเมื่อพิจารณาอัตราส่วนเปรียบเทียบค่าความคลาดเคลื่อนระหว่างทั้ง 2 วิธี พบว่า เทคนิคบ็อกซ์และเจนกินส์ ให้ค่าความคลาดเคลื่อนสูงกว่า วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน 2.05 เท่า

ตารางที่ 4 แสดงการเปรียบเทียบประสิทธิภาพของการพยากรณ์ด้วย เทคนิคบ็อกซ์และเจนกินส์ และ วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน ข้อมูลปริมาณการส่งออกสุกรมี่ชีวิตของประเทศไทย

ค่า	ตัวแบบ		อัตราส่วนเปรียบเทียบ ค่าความคลาดเคลื่อน ของทั้ง 2 วิธี
	เทคนิคบ็อกซ์และเจนกินส์	วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน	
RMSE	4696.204	2179.583	2.15 : 1

ตารางที่ 4 แสดงผลการเปรียบเทียบค่าพยากรณ์ของข้อมูลปริมาณการส่งออกสุกรมี่ชีวิตของประเทศไทยระหว่างวิธีบ็อกซ์และเจนกินส์ และ วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน พบว่า วิธีโครงข่ายเรเดียลเบชีสฟังก์ชันมีประสิทธิภาพในการพยากรณ์มากกว่าวิธีเทคนิคบ็อกซ์และเจนกินส์โดยให้ค่า RMSE ต่ำกว่าวิธีเทคนิคบ็อกซ์และเจนกินส์ และเมื่อพิจารณาอัตราส่วนเปรียบเทียบค่าความคลาดเคลื่อนระหว่างทั้ง 2 วิธี พบว่า เทคนิคบ็อกซ์และเจนกินส์ ให้ค่าความคลาดเคลื่อนสูงกว่า วิธีโครงข่ายเรเดียลเบชีสฟังก์ชัน 2.15 เท่า

สรุปและอภิปรายผล

จากผลการศึกษาเห็นได้ว่าวิธีโครงข่ายเรเดียลเบชีสฟังก์ชันมีประสิทธิภาพในการพยากรณ์ข้อมูลอนุกรมเวลาที่มีรูปแบบการเคลื่อนไหวแบบไม่เชิงเส้นเนื่องจากวิธีดังกล่าวให้ค่าสถิติ RMSE ต่ำกว่า เทคนิคบ็อกซ์และเจนกินส์ ในทุกรูปแบบการ

เคลื่อนไหวของข้อมูล แต่ทั้งนี้วิธีโครงข่ายเรเดียลเบสส์ฟังก์ชัน ยังมีข้อจำกัดในด้านความสะดวกและรวดเร็วในการวิเคราะห์ข้อมูล และเป็นวิธีที่ต้องใช้เวลาในการวิเคราะห์ข้อมูลมากกว่าเทคนิคบ็อกซ์และเจนกินส์ เพราะจะต้องทำการฝึกหัดโครงข่ายหลาย ๆ โครงข่ายจนกว่าจะได้โครงข่ายที่เหมาะสม ประกอบกับโปรแกรมที่ใช้ในการประมวลผลก็ยังไม่ค่อยเป็นที่รู้จักเท่าใดนัก ส่วนเทคนิคบ็อกซ์และเจนกินส์ อาจเป็นวิธีการพยากรณ์ที่มีประสิทธิภาพต่ำกว่า แต่ก็เป็วิธีที่สะดวกและรวดเร็วในการวิเคราะห์ข้อมูล เนื่องจากเป็นวิธีที่เป็นที่รู้จักกันอย่างแพร่หลาย อีกทั้งยังมีซอฟต์แวร์ในการวิเคราะห์ข้อมูลและประมวลผลหลายตัว เช่น โปรแกรม SPSS MINITAP ฯลฯ และเมื่อพิจารณาจากค่าอัตราส่วนเปรียบเทียบค่าความคลาดเคลื่อนจะเห็นว่าข้อมูลปริมาณการจำหน่ายคอมพิวเตอร์ในประเศไทย ซึ่งมีลักษณะการเคลื่อนไหวแบบควอดเรติก ข้อมูลปริมาณการนำเข้าเชื้อเพลิงธรรมชาติและสารหล่อลื่น ซึ่งมีการ

เคลื่อนไหวแบบเอ็กโปเนนเชียล และ ข้อมูลปริมาณการส่งออกสุกรมี่ชีวิตของประเทศไทย ซึ่งมีลักษณะการเคลื่อนไหวแบบไม่มีแบบแผน ซึ่งข้อมูลทั้ง 3 แบบ นี้ให้ค่า RMSE จากวิธีบ็อกซ์และเจนกินส์สูงกว่าวิธีโครงข่ายเรเดียลเบสส์ฟังก์ชัน เพียง 1.77 2.05 และ 2.15 เท่า ตามลำดับเท่านั้น แสดงให้เห็นว่าในบางครั้ง หากต้องการความสะดวกและรวดเร็วในการวิเคราะห์ข้อมูลที่มีรูปแบบการเคลื่อนไหวแบบไม่เชิงเส้นทั้ง 3 แบบดังกล่าว อาจใช้เทคนิคบ็อกซ์และเจนกินส์แทนวิธีโครงข่ายเรเดียลเบสส์ฟังก์ชันได้ ส่วนปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศไทย ซึ่งมีลักษณะการเคลื่อนไหวแบบคิวบิก จะเห็นว่ามีค่า RMSE จากวิธีบ็อกซ์และเจนกินส์สูงกว่าวิธีโครงข่ายเรเดียลเบสส์ฟังก์ชัน ถึง 17.23 เท่า แสดงว่าอนุกรมเวลาที่มีรูปแบบการเคลื่อนไหวแบบคิวบิกเหมาะที่ใช้เทคนิคการพยากรณ์ด้วยวิธีโครงข่ายเรเดียลเบสส์ฟังก์ชัน และไม่เหมาะที่จะใช้เทคนิคบ็อกซ์และเจนกินส์ในการพยากรณ์

บรรณานุกรม

- กัญญารัตน์ บุษบรณ. 2540. **"ตัวแบบสำหรับการพยากรณ์จำนวนคนไข้ในของโรงพยาบาลขนาดใหญ่ 3 โรงพยาบาล."** วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาสถิติประยุกต์ บัณฑิตวิทยาลัย, มหาวิทยาลัยเกษตรศาสตร์.
- สมเกียรติ เกตุเอี่ยม. (2548) **เทคนิคการพยากรณ์.** มหาวิทยาลัยทักษิณ
สำนักงานเศรษฐกิจการเกษตร. **"ข้อมูลปริมาณการส่งออกสุกรมีชีวิต"** แหล่งที่มา : <http://www.oae.go.th>
ธนาคารแห่งประเทศไทย. **"ข้อมูลปริมาณการจำหน่ายคอมพิวเตอร์"** หมวดสถิติเศรษฐกิจและการเงิน
แหล่งที่มา : [http:// bot.or.th](http://bot.or.th)
ธนาคารแห่งประเทศไทย. **"ข้อมูลปริมาณการจำหน่ายรถจักรยานยนต์"** หมวดสถิติเศรษฐกิจและการเงิน
แหล่งที่มา : <http:// bot.or.th>
ธนาคารแห่งประเทศไทย. **"ข้อมูลปริมาณการจำหน่ายเชื้อเพลิงธรรมชาติและสารหล่อลื่น"** หมวดสถิติเศรษฐกิจและการเงิน แหล่งที่มา : <http:// bot.or.th>
- C.H.Wei.(2008) **"Chapter 4 Least-Mean-Square Algorithm."** Adaptive Signal Processing. From <http://cwww.ee.nctu.edu.tw/course/asp/>
- K. Warwick and R. Craddock.(1996) **An Introduction to Radial Basis Function for system identification : a comparison with other neural network methods.** Conference on Decision and Control, pp 464-469.
- Makridakis, S., S.C. Wheelwright and R.J. Hyndman. 1998. **Forecasting : Method and Applications.** 3rd ed. New York : John Wiley and Sons.