การเรียนรู้แบบปรับตัวเองด้วยวิธีเคอร์เนลแบบออนไลน์

โดยใช้ส่วนต่อประสานกราฟิกกับผู้ใช้

ADAPTIVE LEARNING WITH ONLINE KERNEL METHOD USING GUI

นางสาวสุวภา ทองพงษ์เนียม รหัส 50362801

นางสาวภูษณิศา อุภาดี รหัส 50364737

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

ปีการศึกษา 2553

ใบรับรองปริญญานิพนธ์

ชื่อหัวข้อโครงงาน การเรียนรู้แบบปรับตัวเองด้วยวิธีเคอร์เนลแบบออนไลน์โดยใช้ส่วนต่อประสานกราฟิกกับผู้ใช้

ผู้ดำเนินโครงงาน นางสาวสุวภา ทองพงษ์เนียม รหัส 50362801

นางสาวภูษณิศา อุภาดี รหัส 50364737

ที่ปรึกษาโครงงาน ดร. ศุภวรรณ พลพิทักษ์ชัย

สาขาวิชา วิศวกรรมไฟฟ้า

ภาควิชา วิศวกรรมไฟฟ้าและคอมพิวเตอร์

ปีการศึกษา 2553

....

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร อนุมัติให้ปริญญานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า

....ที่ปรึกษาโครงงาน

(ดร. ศุภวรรณ พลพิทักษ์ชัย)

N Jantharamin...กรรมการ

(ดร. นิพัทธ์ จันทรมินทร์)

....กรรมการ

(ดร. มุฑิตา สงฆ์จันทร์)

ชื่อหัวข้อโครงงาน	**การเรียนรู้แบบปรับตัวเองด้วยวิธีเคอร์เนลแบบออนไลน์โดยใช้ส่วนต่อประสานกราฟิกกับผู้ใช้**
ผู้ดำเนินโครงงาน	**นางสาวสุวภา ทองพงษ์เนียม รหัส 50362801**
	นางสาวภูษณิศา อุภาดี รหัส 50364737
ที่ปรึกษาโครงงาน	**ดร. ศุภวรรณ พลพิทักษ์ชัย**
สาขาวิชา	**วิศวกรรมไฟฟ้า**
ภาควิชา	**วิศวกรรมไฟฟ้าและคอมพิวเตอร์**
ปีการศึกษา	2553

บทคัดย่อ

ปริญญานิพนธ์ฉบับนี้ได้ศึกษาการปรับค่าอัตราการเรียนรู้สำหรับการเรียนรู้แบบปรับตัวเองด้วยวิธีเคอร์เนลแบบออนไลน์โดยใช้ส่วนต่อประสานกราฟิกกับผู้ใช้ เพื่อหาฟังก์ชันการประมาณค่าซึ่งเป็นตัวแทนฟังก์ชันที่เราไม่รู้ค่าจากข้อมูลขาเข้าและขาออกที่เก็บค่าได้ เพื่อให้สามารถจำลองฟังก์ชันที่เปลี่ยนแปลงตามเวลาได้ จึงต้องให้อัตราการเรียนรู้มีการปรับค่าตลอดเวลา ซึ่งจะมี 4 วิธีที่ใช้แต่ละวิธีจะมีค่าพารามิเตอร์ที่เกี่ยวข้อง การทดลองจะแสดงให้เห็นถึงประสิทธิภาพของการปรับค่าอัตราการเรียนรู้ในแต่ละแบบว่าค่าความผิดพลาดที่ได้มีค่าเป็นอย่างไร นอกจากนั้นยังแสดงให้เห็นถึงผลของค่าพารามิเตอร์ในการปรับอัตราการเรียนรู้แต่ละวิธีที่มีต่อค่าความผิดพลาดกำลังสอง

Project title	Adaptive Learning with Online Kernel Method using GUI
Name	Ms. Suwapa Tongpongneim ID. 50362801
	Ms. Pusanisa Upadee ID. 50364737
Project advisor	Ms. Supawan Ponpitakchai, Ph.D.
Major	Electrical Engineering
Department	Electrical and Computer Engineering
Academic year	2010

Abstract

This project studies the methods to adapt learning rate for online learning with the kernel method using GUI (Graphic User Interface). Approximation function which is the representation of underlying function from input and output data can be investigated through kernel method. In order to track a non-stationary system, the learning rate needs to be adjusted all the time. This project presents 4 methods to adapt learning rate. There are correspondence parameters to be adjusted in each method. The experiments show the performance of the methods and there parameters through MSE (Mean Square Error).

กิตติกรรมประกาศ

โครงงานนี้สำเร็จลุล่วงไปด้วยความกรุณาเป็นอย่างยิ่งจาก ดร.ศุภวรรณ พลพิทักษ์ชัย ซึ่งเป็นอาจารย์ที่ปรึกษาโครงงาน ที่ได้คอยชี้แนะแนวทางตลอดการทำโครงงาน คณะผู้ดำเนินโครงงานขอกราบขอบพระคุณเป็นอย่างสูงและขอระลึกถึงความกรุณาของท่านไว้ตลอดไป

นอกจากนี้ยังต้องขอขอบคุณคณาจารย์ทุกท่านที่ประสิทธิ์ประสาทวิชาความรู้ให้กับคณะผู้ดำเนินโครงงาน อีกทั้งภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ที่ให้ยืมอุปกรณ์และเครื่องมือต่างๆ จนทำให้โครงงานนี้สำเร็จลุล่วงไปได้ด้วยดี

เหนือสิ่งอื่นใด คณะผู้ดำเนินโครงงานขอกราบขอบพระคุณบิดามารดา ผู้มอบความรักเมตตา สติปัญญา รวมทั้งเป็นผู้ให้ทุกสิ่งทุกอย่างตั้งแต่วัยเยาว์จวบจนถึงปัจจุบัน คอยเป็นกำลังใจที่ทำให้ได้รับความสำเร็จอย่างทุกวันนี้ และต้องขอบคุณทุก ๆ คนในครอบครัวของคณะผู้ดำเนินโครงงานที่ไม่ได้กล่าวไว้ ณ ที่นี้ด้วย

นางสาวสุวภา ทองพงษ์เนียม

นางสาวภูษณิศา อุภาดี

สารบัญ

หน้า

สารบัญ (ต่อ)

หน้า

สารบัญ (ต่อ)

หน้า

สารบัญตาราง

ตารางที่ **หน้า**

สารบัญรูป

รูปที่ หน้า

สารบัญรูป (ต่อ)

รูปที่ หน้า

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของโครงงาน

การเรียนรู้ระบบ (Learning System) คือ การประมาณค่าความสัมพันธ์ทางคณิตศาสตร์ที่ขึ้นอยู่กับตัวเลข หรือ บางความสัมพันธ์ขึ้นอยู่กับตัวอย่างของข้อมูล ในที่นี้การเรียนรู้ระบบสามารถพิจารณาได้จากการประมาณค่าความสัมพันธ์ทางคณิตศาสตร์โดยวิเคราะห์เปรียบเทียบระบบว่าสิ่งที่ไม่ทราบค่ามาก่อนกับสิ่งที่ทราบค่าอยู่แล้ว คือสามารถทำนายผลข้อมูลขาออก (Output) ได้ เมื่อทราบค่าข้อมูลขาเข้า (Input) ดังนั้นการเรียนรู้คือการหาความสัมพันธ์ระหว่างข้อมูลขาเข้าและขาออก ซึ่งข้อมูลที่ได้นั้นจะต้องมีค่าความผิดพลาดน้อยที่สุดและเป็นค่าที่ดีที่สุดที่สามารถนำไปใช้งานได้จริง

เมื่อไม่นานมานี้การประยุกต์ใช้การเรียนรู้ระบบแบบออนไลน์เพิ่มความสำคัญขึ้นในหลายๆด้าน เช่น กระบวนการทางเคมี กระบวนการคิด การออกแบบ ระบบส่งออก ทางด้านการแพทย์และด้านหุ่นยนต์

การเรียนรู้โดยทั่วไปจะเป็นแบบทั้งหมด (Batch learning) คือการเรียนรู้จากการใช้ข้อมูลทั้งหมดในการสอนเพียงหนึ่งครั้ง ข้อเสียของการเรียนรู้แบบทั้งหมด คือ ข้อมูลมีจำนวนมาก เวลาส่งข้อมูลก็จะต้องส่งในจำนวนมากทำให้สามารถเกิดข้อผิดพลาดได้ง่าย

การเรียนรู้แบบออนไลน์ (Online learning) คือการเรียนรู้โดยการส่งข้อมูลไปทีละหนึ่งเพราะฉะนั้นจึงใช้ความจำเพียงเล็กน้อยและมีความแน่นอนมากกว่าการเรียนรู้แบบทั้งหมด ในแต่ละข้อมูลที่ส่งไปจะมีการสร้างแบบจำลอง (Model) ของแต่ละข้อมูลเอาไว้เพื่อจะหาค่าความผิดพลาดว่ามีค่ามากน้อยเพียงใด และจะทำให้คอมพิวเตอร์ของเราทำงานได้อย่างมีประสิทธิภาพ

เครื่องการเรียนรู้ (Machine learning) เป็นงานวิชาหนึ่งของการเรียนรู้แบบออนไลน์ซึ่งจะมีการเรียนรู้ 2 แบบคือ แบบมีผู้สอน (Supervised learning) และแบบไม่มีผู้สอน (Unsupervised learning) การเรียนรู้แบบแรกจะเริ่มด้วยการส่งสิ่งเร้าที่ใช้ในการสอนเข้าไปเป็นข้อมูลขาเข้าของระบบ เมื่อสร้างแบบจำลองขึ้นมาเพื่อทำนายข้อมูลขาออก แล้วจะมีการเปรียบเทียบผลความคลาดเคลื่อนที่เกิดขึ้นเพื่อนำไปคำนวณการปรับแต่งค่าต่างๆ เพื่อลดความคลาดเคลื่อนลงให้เหลือน้อยที่สุด ในทางกลับกันการเรียนรู้แบบไม่มีผู้สอน คือไม่จำเป็นต้องมีค่าเป้าหมายของแต่ละข้อมูลตัวอย่าง ในระหว่างการเรียนรู้จะได้รับข้อมูลกระตุ้นในรูปแบบต่างๆ และจะทำการจัดกลุ่มรูปแบบต่างๆ เหล่านั้นเองตามต้องการ ผลลัพธ์ของการเรียนรู้แบบไม่มีผู้สอนนี้จะเป็นการระบุ

กลุ่มของข้อมูลที่ใส่เข้าไป เช่น การคำนวณค่าความหนาแน่นของข้อมูล (Density estimation), การกระจายและการรวมกลุ่มข้อมูล (Distribution and Clustering) เป็นต้น

การเรียนรู้แบบออนไลน์จะอาศัยทฤษฎีบทของเคอร์เนลที่เรียกว่า RKHS (Reproducing Kernel Hilbert Spaces) ซึ่งมีข้อแตกต่างกับวิธีการเรียนรู้แบบเดิมเช่น วิธีโครงข่ายประสาท (Neural Network) หรือโครงข่ายใยประสาทเสมือน (Artificial Neural Network: ANN) ในแง่ของการหาค่าที่เหมาะสมที่สุด (Optimization) ที่สามารถทำได้ด้วยวิธีเคอร์เนลและยังไม่มีปัญหาด้านค่าที่ต่ำสุดเฉพาะที่ (Local minimum) ในปัจจุบันนี้การเรียนรู้วิธีเคอร์เนลถูกนำไปใช้ในงานต่างๆ อาทิเช่น เอสวีเอ็ม (SVM-support vector machine)

โครงงานนี้เป็นการศึกษาทฤษฎีของการเรียนรู้แบบออนไลน์ด้วยวิธีเคอร์เนล โดยจะมุ่งเน้นการหาฟังก์ชันการประมาณค่าจากข้อมูลจริง เพื่อที่จะหาค่าที่เหมาะสมที่สุด ซึ่งประสิทธิภาพของการเรียนรู้จะถูกแสดงให้เห็นว่าขึ้นอยู่กับการปรับค่าพารามิเตอร์ต่างๆ โดยเน้นไปที่อัตราการเรียนรู้ โดยใช้โปรแกรมแมทแลป (MatLab) และติดต่อกับผู้ใช้งานผ่าน GUI (Graphic User Interface)

1.2 วัตถุประสงค์

เพื่อศึกษาทฤษฎีและประสิทธิภาพของการเรียนรู้แบบออนไลน์ด้วยวิธีเคอร์เนลและศึกษาผลการปรับอัตราการเรียนรู้แบบปรับตัวเองโดยใช้โปรแกรมแมทแลปและติดต่อกับผู้ใช้งานผ่านส่วนต่อประสานกราฟิกกับผู้ใช้

1.3 ขอบเขตของโครงงาน

1) ศึกษาทฤษฎีบทของการเรียนรู้แบบออนไลน์ด้วยวิธีเคอร์เนล
2) ศึกษาการสร้างโปรแกรมแมทแลปสำหรับการเรียนรู้แบบออนไลน์ด้วยวิธีเคอร์เนล
3) นำข้อมูลมาวิเคราะห์ด้วยวิธีเคอร์เนลและศึกษาผลกระทบที่ได้จากการปรับอัตราการเรียนรู้แบบปรับตัวเอง
4) สร้างส่วนต่อประสานกราฟิกกับผู้ใช้

1.4 ขั้นตอนและแผนการดำเนินงาน

รายละเอียด	ปี 2553							ปี 2554		
	มิ.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.
1) รวบรวมข้อมูลการเรียนรู้แบบออนไลน์ด้วยวิธีเคอร์เนล										
2) ศึกษาทฤษฎีทั่วไปของ การเรียนรู้แบบออนไลน์ด้วยวิธีเคอร์เนล										
3) ดำเนินการวิเคราะห์ผลที่ได้จากวิธีการเรียนรู้แบบออนไลน์ด้วยวิธีเคอร์เนล										
4) สรุปผลการดำเนินงานและจัดทำปริญญานิพนธ์ฉบับสมบูรณ์										

1.5 ผลที่คาดว่าจะได้รับ

1) เข้าใจทฤษฎีบทของการเรียนรู้แบบออนไลน์ด้วยวิธีเคอร์เนล
2) เข้าใจการสร้างโปรแกรมแมทแลปสำหรับการเรียนรู้แบบออนไลน์ด้วยวิธีเคอร์เนล
3) สามารถนำข้อมูลมาวิเคราะห์ด้วยวิธีเคอร์เนลและทราบถึงผลกระทบที่ได้จากการปรับอัตราการเรียนรู้แบบปรับตัวเอง
4) สามารถสร้างส่วนต่อประสานกราฟิกกับผู้ใช้

1.6 งบประมาณ

1) ค่าถ่ายเอกสารและค่าเข้าเล่มปริญญานิพนธ์ฉบับสมบูรณ์	1000 บาท
2) ค่าพิมพ์เอกสาร	500 บาท
3) ค่าวัสดุคอมพิวเตอร์	500 บาท
รวมทั้งสิ้น (สองพันบาทถ้วน)	2,000 บาท

หมายเหตุ: ถัวเฉลี่ยทุกรายการ

บทที่ 2
การเรียนรู้แบบตัวเองด้วยวิธีเคอร์เนลแบบออนไลน์ โดยใช้ส่วนต่อประสานกราฟิกกับผู้ใช้

ในบทนี้จะนำเสนอแนวทางในการแก้ไขปัญหาที่สามารถใช้ในงานแตกต่างกัน อาทิเช่น การประมวลผลสัญญาณ การควบคุมเครื่องกลและการประมาณค่าฟังก์ชัน โดยส่วนใหญ่แล้วเราจะมองถึงปัญหาเหล่านี้ในรูปแบบของการประมาณค่าระบบที่ไม่รู้จักโดยอาศัยข้อมูลตัวอย่าง ในปัจจุบันนี้วิธีอาร์เคเอชเอส (RKHS – Reproducing Kernel Hilbert Spaces) ได้ถูกศึกษาและนำไปใช้กับปัญหาต่าง ๆที่ได้กล่าวมาแล้ว ดังนั้นในบทนี้จะเป็นการแสดงทฤษฎีที่เกี่ยวข้องกับการแก้ปัญหาด้วย อาร์เคเอชเอสแบบออนไลน์ นั่นคือ เป็นการหาฟังก์ชันการประมาณค่าที่เป็นตัวแทนของระบบที่เราไม่รู้จักจากจำนวนข้อมูลจำกัด ซึ่งสามารถเก็บค่าได้จากระบบนั้นๆ

2.1 การประมาณค่าฟังก์ชั่นในอาร์เอชเอส [1]

สมมติว่าฟังก์ชันไม่รู้ค่า f ซึ่งสามารถสังเกตค่าจำนวนจำกัดได้ว่าเป็นส่วนหนึ่งของอาร์เคเอชเอส $\mathcal{F}$ และ f นิยามบนเซต $\mathcal{X}$ สามารถถือได้ว่าเป็นข้อมูลขาเข้าโดยที่ $x \in \mathcal{X}$ ดังนั้น $f(x)$ จะสามารถแสดงถึงการประมาณค่าของ f ที่ x และเซตข้อมูลขาเข้า x จะถูกมองว่าเป็นซับเซตปริภูมิยูคลิเดียน (Euclidian space) นั่นคือ $\mathcal{X} \subseteq R^n$ โดยที่ x แต่ละตัวเป็นเวกเตอร์มิติ n

เซตจำกัดของการสังเกตหรือข้อมูลขาออกของฟังก์ชัน $\{z_i\}_{i=1}^N$ จะสอดคล้องกับข้อมูลขาเข้า $\{x_i\}_{i=1}^N$ โดยเราสามารถกล่าวได้ว่า เซตของข้อมูลขาออกอยู่ในปริภูมิ z ถ้าไม่มีข้อผิดพลาดค่าจากการสังเกตจะแสดงได้ดังนี้

$$z_i = L_i f \tag{2.1}$$

เมื่อ $\{L_i\}_{i=1}^N$ เป็นเซตของการประเมินผลฟังก์ชันนัลแบบ (Linear evaluation functional) ถ้ากำหนดบน $\mathcal{F}$ ซึ่งสอดคล้องกับ f เราสามารถแสดงค่าทั้งหมดของการสังเกต $\{z_i\}_{i=1}^N$ ได้ดังนี้

$$z_i = L_i f = \sum_{i=1}^N (L_i f) s_i \tag{2.2}$$

โดยที่ $s_i \subseteq R^n$ เป็นเวกเตอร์มาตรฐาน (Standard basis vector) ที่ลำดับ i โดยในทางปฏิบัติสามารถเขียนได้เป็น

$$z_i = f(x_i) \tag{2.3}$$

ซึ่งสามารถนำไปใช้ในแก้ปัญหาการประมาณค่าปัญหาฟังก์ชันการประมาณค่า ซึ่งเป็นตัวแทนฟังก์ชันที่ไม่รู้ค่า สามารถกระทำได้โดยกำหนดคลาสของ $\mathcal{F}$ ของฟังก์ชัน และค่าสังเกต $\{z_i\}_{i=1}^{N}$ ของฟังก์ชันนัล L_i ซึ่งถูกกำหนดบน $\mathcal{F}$ ดังนั้นจะสามารถฟังก์ชัน f บน $\mathcal{F}$ ซึ่งสอดคล้องกับสมการที่ (2.1) และสมการที่ (2.3) ตามหลัก เราสามารถกำหนดค่าอาร์เคเอชเอสให้เป็นปริภูมิฮิลเบิร์ต (Hilbert space) ของฟังก์ชันบน $\mathcal{X}$ ด้วยคุณสมบัติที่ว่าในแต่ละ $x \in \mathcal{X}$ ฟังก์ชันนัล L_i ซึ่งเชื่อมโยง f ด้วย $f(x_i)$ นั่นคือ $L_i \to f(x_i)$ จะเป็นฟังก์ชันนัลแบบเส้นตรงที่ถูกกำหนดขอบเขต และการมีขอบเขต หมายถึง การมีของค่า M โดยที่

$$|L_i f| = |f(x_i)| \leq M||f|| \quad \text{สำหรับทุกๆ } f \text{ ใน RKHS}$$

เมื่อ$||\cdot||$ เป็นค่านอร์มในอาร์เคเอชเอส

ฟังก์ชันนัล L_i ถูกกำหนดขอบเขต ตามทฤษฎีบทของรีซ (Riesz representationtheorem) ซึ่งเราสามารถแสดงค่าการสังเกตได้ดังนี้

$$L_i f = \langle f, k_i \rangle \quad i = 1, \ldots, N \tag{2.4}$$

เมื่อ $\langle \cdot,\cdot \rangle$ หมายถึงผลคูณภายใน (Inner product) ใน $\mathcal{F}$ และ $\{k_i\}_{i=1}^{N}$ เป็นชุดของฟังก์ชันที่เรียกว่ารีโพรดิวซิ่งเคอร์เนล (Reproducing kernels) ซึ่งเป็นส่วนหนึ่งของ $\mathcal{F}$ และจะไม่ถูกกำหนดซ้ำกัน โดยฟังก์ชันนัล L_i

ดังนั้นการหาปัญหาการประมาณจะสามารถกำหนดใหม่ได้อีกแบบ นั่นคือ การกำหนดปริภูมิฮิลเบิร์ต ของเซตของฟังก์ชัน $\{k_i\}_{i=1}^{N} \subset \mathcal{F}$ และค่าการสังเกต $\{z_i\}_{i=1}^{N}$ จากเงื่อนไขดังกล่าวมาแล้ว สามารถหาฟังก์ชันการประมาณค่า $f \in \mathcal{F}$ ที่สอดคล้องกับสมการที่ (2.4) ได้ในทุกอาร์เคเอชเอสจะเป็นฟังก์ชันบวกแน่นอน (Positive − definite function) ที่เรียกว่า รีโพรดิวซิ่งเคอร์เนล (k) ที่กำหนดบน $\mathcal{X} \times \mathcal{X}$

(i) $k(\cdot, x') \in \mathcal{F}$;

(ii) $\langle f, k(\cdot, x') \rangle_{\mathcal{F}} = f(x')$

สำหรับทุก f ใน $\mathcal{F}$

จากคุณสมบัติข้างต้นแสดงให้เห็นว่า ฟังก์ชันรีโพรดิวซิ่งเคอร์เนล k_i เป็นฟังก์ชันบน $\mathcal{X} \times \mathcal{X}$ แต่อย่างไรก็ตาม เราสามารถแสดงฟังก์ชันเคอร์เนลรูปแบบของฟังก์ชันบน $\mathcal{X}$ เท่านั้น ได้ด้วยการพิจารณาฟังก์ชันเคอร์เนล $k(x, x_i)$ เมื่อ x_i เป็นจุดที่ไม่เปลี่ยนแปลงค่าและ $x_i \in \mathcal{X}$ ดังนั้นการเขียน $k(x, x_i) = k(x_i)$ (หรือ k_i) จะสามารถสรุปได้ว่า $k_i \in \mathcal{F}$ บน $\mathcal{X}$ มีจุดศูนย์กลางบน x_i โดยส่วนใหญ่แล้วจุด x_i มักจะถูกกำหนดให้เป็นศูนย์กลางของฟังก์ชันเคอร์เนลและยังเป็นไฮเปอร์พารามิเตอร์ (Hyperparameter) ของฟังก์ชันเคอร์เนลอีกด้วย ซึ่งโดยทั่วไปแล้วศูนย์กลางเหล่านี้จะสอดคล้องกับข้อมูลขาเข้า ดังนั้นเราจึงสามารถแสดงฟังก์ชันทั่วๆไป f บนอาร์เคเอชเอส $\mathcal{F}$ ด้วยรีโพรดิวซิ่งเคอร์เนล k ได้ดังนี้

$$f(x) \;=\; \sum_i \alpha_i k(x, x_i) \;=\; \sum_i \alpha_i k_i(x) \tag{2.5}$$

เมื่อ $\alpha_i \in \mathbb{R}, i \in \square$ เพราะฉะนั้นการแก้ปัญหาการประมาณค่าฟังก์ชันจะเปลี่ยนรูปเป็นการประมาณค่าที่เหมาะสมสำหรับตัวแปร α_i ในสมการที่ (2.5)

เราสนใจการกำหนดนิยามของ RKHS ซึ่งเป็นฟังก์ชัน สามารถหาค่าได้ดังนี้

$$f(x) \;=\; \sum_{i=1}^{p} \alpha_i k_i(x) \tag{2.6}$$

ที่ $p \in \square\mathbb{N}$

ฟังก์ชันนี้เป็นฟังก์ชันเคอร์เนลซึ่งมีหลายชนิดที่นิยมใช้ ตัวอย่างเช่น ฟังก์ชันเกาเสียนเรเดียลเบสิคที่ใช้ในระบบโครงข่ายประสาทเทียม

$$k(x, x^{\square}) \;=\; exp(-\beta||x - x\square||^2) \tag{2.7}$$

โดยที่ $\beta > 0$

ฟังก์ชันโพลิเนเมียลของชุดของค่า x กำลัง d แสดงได้โดย

$$k(x, x^{\square}) \;=\; \sum_{k=0}^{d} \frac{1}{k!\rho_k}(xx^{\square})^k \tag{2.8}$$

เมื่อ ρ_k คือเซตของค่าคงที่ที่เรียกว่าค่าน้ำหนัก มีค่าเท่ากับ

$$k(x, x^{\square}) \;=\; (1 + xx^{\square})^d \tag{2.9}$$

ฟังก์ชันเคอร์เนลสามารถแสดงได้ด้วยปริภูมิ พาเลย์ – เวียเนอร์ ของฟังก์ชันจำกัดแถบ (Band limited function)

$$k(x, x^{\square}) \;=\; \frac{sin\pi(x - x^{\square})}{\pi(x - x^{\square})} \tag{2.10}$$

2.2 การประมาณค่าฟังก์ชั่นในอาร์เคเอชเอสแบบออนไลน์ [3]

ในกรณีนี้ได้ตั้งสมมติฐานว่าในแต่ละรอบของการทำซ้ำ เราจะทราบค่าสังเกตเพียงแค่ค่าเดียว นั่นคือ $\mathrm{z_n}$ ดังนั้น

$$L_n f \;=\; z_n \tag{2.11}$$

จากนั้นเราจะได้ฟังก์ชันนัล $\hat{\mathrm{g}}_{\mathrm{reg}}$ ที่เวลา n ที่เป็นฟังก์ชันไม่เป็นค่าลบ โดยที่ $\hat{g}_{reg} : Z \to R$

$$\hat{g}_{reg}(f_n) \;=\; \frac{1}{2}||L_{n+1}f_n - z_{n+1}||^2 + \frac{\rho}{2}||f_n||^2 \tag{2.12}$$

ตั้งค่าเริ่มต้น f_1 ซึ่งโดยส่วนใหญ่จะให้มีค่าเท่ากับศูนย์ จากนั้นเราจะหาค่า f_n ที่จะทำให้สมการที่ (2.12) มีค่าต่ำที่สุดโดยอาศัยวิธีสโตแคสติกเกรเดียนเคสเซนต์ (Stochastic gradient descent, SGD) ที่แสดงโดย

$$f_{n+1} = f_n - \eta_n \nabla\, \hat{g}_{reg}(f_n) \tag{2.13}$$

เมื่อ $\nabla \hat{g}_{reg}(f_n)$ เป็นเกรเดียน ณ เวลาปัจจุบันของ $\hat{g}_{reg}$ เทียบกับ f_n ดังนั้น

$$\begin{aligned}\hat{g}_{reg}(f_n) &= L^*_{n+1}L_{n+1}f_n - L^*_{n+1}z_{n+1} + \rho f_n \\ &= L^*_{n+1}(L_{n+1}f_n - z_{n+1}) + \rho f_n\end{aligned}$$

แทนค่าในสมการที่ (2.13) จะได้ว่า

$$f_{n+1} = (1-\eta_n\rho)f_n - \eta_n L^*_{n+1}(L_{n+1}f_n - z_{n+1}) \tag{2.14}$$

โดยที่ค่าคงที่ใดๆ a จะสามารถแสดงได้ว่า $L^*_{n+1}a = k_{n+1}a$ และ $L_{n+1}f_n = f_n(x_{n+1})$ ดังนั้น

$$\nabla \hat{g}_{reg}(f_n) = k_{n+1}[f_n(x_{n+1}) - z_{n+1}] + \rho f_n \tag{2.15}$$

เพราะฉะนั้นในสมการที่ (2.13) สามารถเขียนได้เป็น

$$f_{n+1} = (1-\eta_n\rho)f_n - \eta_n k_{n+1}[f_n(x_{n+1}) - z_{n+1}] \tag{2.16}$$

สมมติว่าการประมาณค่าฟังก์ชัน ณ เวลาปัจจุบันมีเคอร์เนลจำนวน p เทอม ดังนั้นฟังก์ชันการประมาณค่าเหล่านี้สามารถเขียนใหม่เป็น

$$f_{n+1}(x) = (1-\eta_n\rho)\sum_{i=1}^{p} \alpha_n^i k_i(x) - \eta_n e_{n+1}\, k_{n+1}(x) \tag{2.17}$$

เมื่อ α_{n+1}^i ถูกคำนวณมาก่อนแล้วดังสมการต่อไปนี้

$$\alpha_{n+1}^i = (1-\eta_n\rho)\alpha_n^i \qquad \text{โดยที่ } i \le p \tag{2.18}$$

และ

$$\alpha_{n+1}^i = -\eta_n e_{n+1} \qquad \text{โดยที่ } i = p+1 \tag{2.19}$$

จากสมการที่ผ่านมาเราสามารถสรุปได้ว่า เมื่อค่าน้ำหนัก (α) ค่าใหม่ถูกเพิ่มให้กับแบบจำลอง (f_{n+1}) โดยการคูณเข้ากับฟังก์ชันเคอร์เนล ค่าน้ำหนักค่าเก่าจะถูกปรับปรุงโดยการคูณด้วยเทอม $(1-\eta_n\rho)$ ซึ่งหมายถึงการลดลงของค่าน้ำหนักค่านั้นๆ

$$f_{n+1} = \sum_{i=1}^{p+1}(1-\eta_n\rho)^{n+1-i}\,\eta e_i k_i \tag{2.20}$$

2.3 อัตราการเรียนรู้ชนิดปรับตัวได้ [4]

หลักการทำงานของวิธีนี้คืออัตราการเรียนรู้จะมีการปรับในทุกๆรอบของการคำนวณ เพื่อที่จะติดตามฟังก์ชันที่เราไม่ทราบค่าที่เปลี่ยนแปลงตามเวลาให้ได้ค่าที่เหมาะสมที่สุด โดยที่ค่าพารามิเตอร์ ρ จะถูกกำหนดใช้ให้เป็นค่าคงที่ ฟังก์ชันการประมาณค่า f_{n+1} จะแสดงดังนี้

$$f_{n+1} = f_n - \eta_n \frac{\partial \hat{g}_{reg}(f_n)}{\partial f_n} \tag{2.21}$$

การปรับอัตราการเรียนรู้จะมี 3 วิธี อาศัยแนวคิดของการปรับค่าอัตราการเรียนรู้ในงานด้านสโตแคสติกออพติไมเซชัน (Stochastic optimisation)

2.3.1 การปรับอัตราการเรียนรู้วิธีที่ 1 [4]

สมมติว่าฟังก์ชันที่ไม่ทราบค่าของเรามีการเปลี่ยนแปลงอย่างช้าๆ ดังนั้นเราจึงต้องการให้อัตราการเรียนรู้มีการปรับตัวอย่างช้าๆด้วยโดยกระบวนการปรับค่าอัตราการเรียนรู้คือ

$$\eta_n = \eta_{n-1} - \gamma \frac{\partial \hat{g}_{reg}(f_{n+1})}{\partial \eta_n} \tag{2.22}$$

เมื่อ □ $\gamma \in \mathbb{R}^+$ เป็นค่าแฟคเตอร์ของอัตราการเรียนรู้ที่ปรับตัวได้ เมื่อ

$$\frac{\partial \hat{g}_{reg}(f_{n+1})}{\partial \eta_n} = \frac{\partial \hat{g}_{reg}(f_{n+1})}{\partial f_{n+1}} \cdot \frac{\partial f_{n+1}}{\partial \eta_n} \tag{2.23}$$

จะได้

$$\begin{aligned}\frac{\partial f_{n+1}}{\partial \eta_n} &= \frac{\partial\left\{f_n - \eta_n \frac{\partial \hat{g}_{reg}(f_n)}{\partial f_n}\right\}}{\partial \eta_n} \\ &= -\frac{\partial \hat{g}_{reg}(f_n)}{\partial f_n}\end{aligned} \tag{2.24}$$

แทนค่าลงในสมการที่ (2.23) จะได้

$$\frac{\partial \hat{g}_{reg}(f_{n+1})}{\partial □_n} = -\frac{\partial \hat{g}_{reg}(f_{n+1})}{\partial f_{n+1}} \cdot \frac{\partial \hat{g}_{reg}(f_n)}{\partial f_n} \tag{2.25}$$

ใช้สมการด้านบนแทนค่าในสมการที่ (2.22) ดังนั้น

$$\eta_n = \eta_{n-1} - \gamma \left\langle \frac{\partial \hat{g}_{reg}(f_{n+1})}{\partial f_{n+1}}, \frac{\partial \hat{g}_{reg}(f_n)}{\partial f_n} \right\rangle \tag{2.26}$$

สมการข้างต้นเรียกว่ากฎการปรับตัว (Update rule) ซึ่งจะมีการใช้ค่าที่เวลา $n+1$ ที่เป็นค่าในอนาคตไม่สามารถหาได้ เราจะใช้สมมติฐานที่กล่าวมาแล้ว ว่าระบบที่เราสนใจมีการเปลี่ยนแปลงอย่างช้าๆ ดังนั้นจึงสามารถใช้ค่าที่เวลาก่อนหน้านี้ในสมการที่ (2.26) ได้แสดงดังสมการต่อไปนี้

$$\begin{aligned}\eta_n &= \eta_{n-1} - \gamma \left\langle \frac{\partial \hat{g}_{reg}(f_{n+1})}{\partial f_{n+1}}, \frac{\partial \hat{g}_{reg}(f_n)}{\partial f_n} \right\rangle \\ &= \eta_{n-1} + \gamma \left\langle \nabla \hat{g}_{reg}(f_n), \nabla \hat{g}_{reg}(f_{n-1}) \right\rangle\end{aligned} \tag{2.27}$$

การหาค่า η_n จะต้องอาศัยการคำนวณผลคูณภายในของ $\langle \nabla\hat{g}_{reg}(f_n), \nabla\hat{g}_{reg}(f_{n-1})\rangle$ โดยอาศัยสมการดังต่อไปนี้

$$\begin{aligned}\nabla \hat{g}_{reg}(f_n) &= L^*_{n+1}(f_n(x_{n+1}) - z_{n+1})\rho f_n \\ &= \eta_{n-1} + \gamma \left\langle \nabla \hat{g}_{reg}(f_n), \nabla \hat{g}_{reg}(f_{n-1}) \right\rangle\end{aligned} \tag{2.28}$$

และ

$$\begin{aligned}\nabla \hat{g}_{reg}(f_{n-1}) &= L^*_n(f_{n-1}(x_n) - z_n) + \rho f_{n-1} \\ &= L^*_{n+1} e_{n+1} + \rho f_n\end{aligned} \tag{2.29}$$

ซึ่งจะได้ผลคูณภายใน

$$\langle\nabla\hat{g}_{reg}(f_n),\nabla\hat{g}_{reg}(f_{n-1})\rangle = \langle L^{*}_{n+1}e_{n+1} + \square f_n, L^{*}_{n}e_n + \square f_{n-1}$$
$$= \langle L^{*}_{n}e_n, L^{*}_{n+1}e_{n+1}\rangle + \langle L^{*}_{n}e_n, \rho f_n\rangle$$
$$+\langle \rho f_{n-1},, L^{*}_{n+1}e_{n+1}\rangle + \rho^2\langle f_{n-1}, f_n\rangle \tag{2.30}$$

สามารถเขียนได้เป็น

$$\langle\nabla\hat{g}_{reg}(f_n),\nabla\hat{g}_{reg}(f_{n-1})\rangle = \langle L_{n+1}L^{*}_{n}e_n, e_{n+1}\rangle + \rho\langle e_n, L_n f_n\rangle$$
$$+\rho\langle L_{n+1}f_{n-1}, e_{n+1}\rangle + \rho^2\langle f_{n-1}, f_n\rangle \tag{2.31}$$

แปลงอยู่ในรูปเมตริกซ์ได้ดังนี้

$$\langle\nabla\hat{g}_{reg}(f_n),\nabla\hat{g}_{reg}(f_{n-1})\rangle = e_n k(x_n, x_{n+1}) + \square\big(e_n f_n(x_n) + e_{n+1}f_{n-1}(x_{n+1})\big)$$
$$+ \rho^2\alpha^T_{n-1}K_{n-1,n}\alpha_n \tag{2.32}$$

เมื่อ $e_{n+1} = f_n(x_{n+1}) - z_{n+1}, e_n = f_n(x_n) - z_n,\ K_{n-1,n} \in \mathbb{R}^{n\times n+1}$,

$K_{ij} = k(x_n, x_j)$ และ α_{n-1}, α_n เป็นเวกเตอร์ของพารามิเตอร์ที่เวลา $n-1$ และ n ตามลำดับ

2.3.2 การปรับอัตราการเรียนรู้วิธีที่ 2 [4]

การปรับค่าด้วยวิธีการนี้ ค่า η_n จะถูกทำให้มีค่ามากๆได้อย่างรวดเร็ว โดยอาศัยการปรับค่าด้วยวิธีเรขาคณิต นั่นคือ สามารถกระทำได้โดยการเลือกค่าแฟคเตอร์เป็น $\gamma\eta_{n-1}$ ดังนั้นเราจะได้กฎการปรับตัว

$$\eta_n = \eta_{n-1} + \gamma\eta_{n-1}\langle\nabla\hat{g}_{reg}(f_n),\nabla\hat{g}_{reg}(f_{n-1})\rangle$$
$$= \eta_{n-1}\left\{1 + \gamma\langle\nabla\hat{g}_{reg}(f_n),\nabla\hat{g}_{reg}(f_{n-1})\rangle\right\} \tag{2.33}$$

อย่างไรก็ตามวิธีการปรับค่าวิธีนี้ค่อนข้างจะขึ้นอยู่กับฟังก์ชั่น $\nabla\hat{g}_{reg}(f_n)$ ดังนั้นวิธีต่อไปจะเป็นการกำจัดการขึ้นอยู่กับฟังก์ชันที่ได้กล่าวมาแล้ว

2.3.3 การปรับอัตราการเรียนรู้วิธีที่ 3 [4]

ปกติการปรับจะไม่ขึ้นอยู่กับการอนุพันธ์บางส่วนซึ่ง $\hat{g}_{reg}$ เป็นผลมาจากการใช้อัตราการเรียนรู้แบบสมดุล $\frac{\gamma\eta_{n-1}}{u_n}$ ที่ μ_n ได้มาโดย

$$u_n = \mu\eta_{n-1} + (1-\mu)\|\nabla\hat{g}_{reg}(f_n)\|^2 \tag{2.34}$$

ค่าของ μ_0 เป็นค่าเริ่มต้น รูปแบบของเมตริก $\|\nabla\hat{g}_{reg}(f_n)\|^2$ มีข้อกำหนดโดย $\langle\nabla\hat{g}_{reg}(f_n) = L^{*}_{n+1}e_{n+1} + \rho f_n\rangle$ ตามที่

$$\|\nabla\hat{g}_{reg}(f_n)\|^2 = \langle\nabla\hat{g}_{reg}(f_n),\nabla\hat{g}_{reg}(f_n)\rangle$$
$$= \langle L^{*}_{n+1}e_{n+1} + \rho f_n, L^{*}_{n+1}e_{n+1} + \rho f_n\rangle$$

$$
\begin{aligned}
&= \langle L^*_{n+1}e_{n+1}, L^*_{n+1}e_{n+1}\rangle + 2\langle L^*_{n+1}e_{n+1}, \square f_n\rangle + \langle \square f_n, \square f_n\rangle \\
&= \langle L_{n+1}L^*_{n+1}e_{n+1}, e_{n+1}\rangle + 2\square\langle e_{n+1}, L_{n+1}f_n\rangle + \rho\square^2\langle f_n, f_n\rangle \\
&= k(x_{n+1}, x_{n+1})e^2_{n+1} + 2e_{n+1}\rho f_n(x_{n+1}) + \rho^2\alpha_n^T K_{n,n}\alpha_n
\end{aligned}
\tag{2.35}
$$

ปัจจุบันเรากำหนดกฎการปรับค่าใหม่เหมือนกฎการปรับค่าตามปกติโดย

$$\eta_n = \eta_{n-1}\{1 + \frac{\gamma}{u_n}\langle\nabla\hat{g}_{reg}(f_n), \nabla\hat{g}_{reg}(f_{n-1})\rangle \tag{2.36}$$

2.4 การปรับค่าอัตราการเรียนรู้วิธีเอสเอ็มดี [5]

ถึงแม้ว่าการปรับอัตราการเรียนรู้วิธีที่ 1 วิธีที่ 2 และวิธีที่ 3 จะปรับอัตราการเรียนรู้ได้ค่อนข้างดี แต่วิธีการเหล่านี้มีข้อเสียอยู่หลายประการ เช่น อัตราการเรียนรู้ที่ถูกปรับค่าโดยใช้ผลคูณของอนุพันธ์จะสามารถเป็นค่าลบได้ ยิ่งไปกว่านั้น กฎการปรับตัวของวิธีที่ผ่านมาจะใช้ข้อมูลเวลา n และ $n-1$ เท่านั้น ดังนั้นจึงมีการเสนอวิธีสำหรับการปรับค่าอัตราการเรียนรู้ที่เรียกว่า วิธีเอสเอ็มดี ซึ่งวิธีการนี้หลีกเลี่ยงค่าที่เป็นลบของอัตราการเรียนรู้ที่ผ่านการปรับค่าผ่านฟังก์ชันเอกซ์โปเนนเชียล นอกจากนี้ค่าทั้งหมดที่ผ่านมาของข้อมูลจะมีส่วนร่วมในการปรับค่าอัตราการเรียนรู้ ณ เวลาปัจจุบัน

จากสมการที่ (2.23) การปรับค่าอัตราการเรียนรู้จะถูกกระทำโดยฟังก์ชันเอ็กโปเนนเชียลเพื่อที่จะสามารถครอบคลุมการเปลี่ยนแปลงช่วงกว้างโดยที่ค่าที่ได้มีค่าบวกเสมอ ซึ่งจะแสดงดังสมการต่อไปนี้

$$
\begin{aligned}
\ln\eta_n &= \ln\eta_{n-1} - \mu\frac{\nabla\hat{g}_{reg}(f_n)}{\partial\ln\eta_{n-1}} \\
&= \ln\eta_{n-1} - \mu\frac{\nabla\hat{g}_{reg}(f_n)}{\partial(f_n)}\cdot\frac{\partial(f_n)}{\partial\ln\eta_{n-1}}
\end{aligned}
\tag{2.37}
$$

เพราะฉะนั้น

$$\eta_n = \eta_{n-1}\cdot exp(-\mu\langle h_n, v_n\rangle h_n \tag{2.38}$$

เมื่อ $h_n \equiv \frac{\nabla\hat{g}_{reg}(f_n)}{\partial(f_n)}$, $v_n \equiv \frac{\partial(f_n)}{\partial\ln\eta_{n-1}}$ และ μ เป็นค่าโกลบอลของอัตราการเรียนรู้ เพื่อให้คำนวณง่ายขึ้นจะใช้การประมาณค่า $e^u \approx 1 + u$ ดั้งนั้นจะได้กฎการปรับตัว

$$\eta_n = \eta_{n-1}\cdot\max(\tfrac{1}{2}, 1 - \mu\langle h_n, v_n\rangle) \tag{2.39}$$

เราจะเรียก v ว่าเป็นเกรเดียนเทรซ (Gradient trace) ใช้วัดผลกระทบที่มีต่อฟังก์ชัน f_n ที่มีการปรับปรุง ในที่นี้เราพิจารณาผลกระทบในการเปลี่ยนแปลงเวลาทั้งหมดจนถึงเวลา n ในรูปแบบของค่าเฉลี่ยแบบเอ็กซ์โปเนนเชียล ดังนั้นการหาค่าเกรเดียนเทรซจะเป็นการบวกค่าเกรเดียนจนถึงที่เวลา n นั้นคือ

$$v_{n+1} \equiv \sum_{i=0}^{n} \square^{i} \frac{\partial(f_n)}{\partial \ln \square_{n-i}} \tag{2.40}$$

กำหนดให้ $\lambda \in [0,1]$ เป็นค่าฟอร์เกตติ้ง (Forgetting factor) แทนค่า f_{n+1} จะได้

$$\begin{aligned} v_{n+1} &\equiv \sum_{i=0}^{n} \lambda^{i} \frac{\partial(f_n)}{\partial\, ln\, \eta_{n-i}} - \sum_{i=0}^{n} \lambda^{i} \frac{\partial(f_n)}{\partial\, ln\, \eta_{n-i}} \\ &\approx \lambda v_n - \eta_n \cdot h_n - \eta_n \sum_{i=0}^{n} \lambda^{i} \frac{\partial h_n}{\partial f_n} \cdot \frac{\partial f_n}{\partial\, ln\, \eta_{n-i}} \\ &= \lambda v_n - \eta_n \cdot h_n - \eta_n H_n \lambda v_n \\ &= \lambda v_n - \eta_n (h_n + H_n \lambda v_n) \end{aligned} \tag{2.41}$$

เมื่อ H_n หมายถึงค่าเฮสเซียน (Hessian) ของ $\hat{g}_{reg}$ ที่เวลา n

2.5 การใช้วิธีเอสเอ็มดีในการปรับค่าการเรียนรู้แบบออนไลน์บนอาร์เคเอชเอส [5]

ในการใช้วิธีเอสเอ็มดี สำหรับการปรับค่าอัตราการเรียนรู้ที่แสดงในหัวข้อที่แล้ว จะถูกปรับให้เข้ากับการทำงานบนอาร์เคเอชเอส ดังนั้นสูตรที่ถูกปรับปรุงในการหาค่าในสมการที่ (2.41) ซึ่งต้องการ h_n และ v_n ดังนี้

$$\begin{aligned} h_n &= \frac{\nabla \hat{g}_{reg}(f_n)}{\partial f_n} = \nabla \hat{g}_{reg}(f_n) \\ &= L_{n+1}^{*}(L_n + f_n - z_{n+1}) + \rho_n f_n \end{aligned} \tag{2.42}$$

สำหรับค่า v แสดงในสมการที่ (2.40) จะได้ค่าเฮสเซียนเมตริก H คือ

$$H_n = \frac{\partial h_n}{\partial f_n} = \frac{\partial^2 g_{reg}(f_n)}{{f_n}^2} = L_{n+1}^{*} + L_{n+1} + \rho \tag{2.43}$$

เพราะฉะนั้นการปรับค่าในสมการที่ (2.43) จะเท่ากับ

$$\begin{aligned} v_{n+1} &= \lambda v_n - \eta_n (L_{n+1}^{*} e_{n+1} + \rho_n) - \eta_n \lambda (L_{n+1}^{*} e_{n+1} + \rho) v_n \\ &= (1 - \rho\eta_n)\lambda v_n - \eta_n \lambda L_{n+1}^{*} L_{n+1} v_n - \eta_n (L_{n+1}^{*} e_{n+1} + \rho) \\ &= (1 - \rho\eta_n)\lambda v_n - \eta_n \lambda L_{n+1}^{*} L_{n+1} v_n - \eta_n (L_{n+1}^{*} e_{n+1} + \rho f_n) \\ &= (1 - \rho\eta_n)\lambda v_n - \eta_n \rho f_n - \eta_n \lambda k_{n+1}(x_{n+1}) - \eta_n k_{n+1} e_{n+1} \end{aligned} \tag{2.44}$$

เมื่อ $L_{n+1}^{*} a = k_{n+1} a$ เราสามารถแสดงฟังก์ชัน v ใส่ในรูปของฟังก์ชันเคอร์เนลนั้นคือ

$$v_{n+1} = \sum_{i=1}^{n+1} \beta_{n+1}^{i} k_i \tag{2.45}$$

โดยที่ $\beta_i \in \mathbb{R}$ จากสมการที่ (2.46) และสมการที่ (2.47) จะได้

$$\begin{aligned} v_{n+1} = (1 - \rho\eta_n)\lambda \sum_{i=1}^{n+1} \beta_n^{i} k_i - \eta_n \rho \sum_{i=1}^{n+1} \alpha_n^{i} k_i \\ -\eta_n k_{n+1}(\lambda v_n(x_{n+1}) - e_{n+1}) \end{aligned} \tag{2.46}$$

ขณะที่ตัวแปร β_{n+1} มีการปรับค่า

$$\beta_{n+1}^{i} = \begin{cases} (1-\eta_{\mathrm{n}}\rho)\lambda\beta_{n}^{i} - \eta_{\mathrm{n}}\rho\alpha_{n}^{i} & \text{เมื่อ } i \le n \\ -\eta_{\mathrm{n}}\lambda v_{n}(x_{n+1}) - \eta_{\mathrm{n}e_{n+1}} & \text{เมื่อ } i = n+1 \end{cases} \tag{2.47}$$

ดังนั้นเราจะได้กฎการปรับค่าสำหรับปรับปรุงอัตราการเรียนรู้ในสมการที่ (2.39) โดยคำนวณค่าผลคูณภายใน $\langle h_n, v_n \rangle$ จากค่า h_{n} ในสมการที่ (2.42) จะได้

$$\begin{aligned} \langle h_n, v_n \rangle &= \langle L_{n+1}^{*} e_{n+1} + \rho f_n, v_n \rangle \\ &= \langle L_{n+1}^{*} e_{n+1}, v_n \rangle + \langle \rho f_n, v_n \rangle \\ &= \langle e_{n+1} L_{n+1}^{*} v_n \rangle + \rho \langle f_n, v_n \rangle \\ &= e_{n+1} v_n(x_{n+1}) + \rho\alpha_n^T K \beta_{\mathrm{n}} \end{aligned} \tag{2.48}$$

เมื่อ $\langle f_n, v_n \rangle = \sum_{i=1}^{n+1} \alpha_n^i k_i, \sum_{i=1}^{n+1} \beta_n^i k_i = \rho\alpha_n^T K \beta_n$ ดังนั้นกฎการปรับค่าสำหรับอัตราการเรียนรู้คือ

$$\eta_{\mathrm{n}} = \eta_{\mathrm{n-1}} \cdot \max(\tfrac{1}{2}, 1 - \mu(e_{n+1} v_n(x_{n+1}) + \rho\alpha_n^T K \beta_{\mathrm{n}})) \tag{2.49}$$

2.6 ส่วนต่อประสานกราฟิกกับผู้ใช้ [2]

GUI ย่อมาจาก Graphical User Interface ซึ่งเป็นโปรแกรมโต้ตอบกับผู้ใช้แบบกราฟ ซึ่งถูกสร้างจากวัตถุแบบกราฟ ได้แก่ ปุ่มกด (Push buttons) ตัวเลือกแบบเมนู (Pop-up menu) กราฟ (Axes) ฯลฯ การที่ผู้ใช้ป้อนค่าให้กับโปรแกรม GUI ได้แก่ การคลิกเมาส์บนปุ่มกดการป้อนข้อความลงในช่องแก้ข้อความ (Edit box) เป็นต้น การทำงานของโปรแกรมจึงขึ้นอยู่กับผู้ใช้ว่ามีการป้อนข้อมูลหรือสั่งการกับองค์ประกอบ GUI หรือไม่ หากโปรแกรมโต้ตอบกับผู้ใช้ได้รับการออกแบบที่ดีจะทำให้ผู้ใช้สามารถเข้าถึงการ ทำงานหรือใช้งานได้อย่างไม่ต้องทราบขั้นตอนการทำงานมาก่อนเลย

การสร้างโปรแกรม GUI ในแมทแลปมีส่วนประกอบหลัก 2 ส่วนที่ต้องพิจารณา มีดังนี้

1) องค์ประกอบของ GUI บนหน้าต่างที่ใช้ติดต่อโต้ตอบกับผู้ใช้ เช่น ปุ่มกด ช่องแก้ข้อความ แกน องค์ประกอบของ GUI เป็นต้น เหล่านี้ล้วนแล้วแต่ให้พิจารณาว่าเป็นออบเจกต์ชนิดหนึ่งในโปรแกรมแมทแลปทั้งสิ้น

2) ฟังก์ชันที่ทำงานเพื่อตอบสนองต่อเหตุการณ์ตามที่ผู้ใช้โต้ตอบกับองค์ประกอบแต่ละส่วนของโปรแกรม GUI โดยฟังก์ชันเหล่านี้มีชื่อเฉพาะว่า ฟังก์ชัน callback

2.7 การสร้างโปรแกรม GUI โดยใช้ guide [2]

ในโปรแกรมแมทแลปได้พัฒนาเครื่องมือที่เรียกว่า guide (Graphical User Interface Development Environment) ขึ้นเพื่อช่วยในการสร้างโปรแกรม GUI โดยเฉพาะ เครื่องมือดังกล่าวนี้ทำให้ผู้ใช้โปรแกรมสามารถจัดวางองค์ประกอบของ GUI ต่างๆบนหน้าต่างกราฟิกที่ใช้ติดต่อโต้ตอบกับผู้ใช้งานได้โดยง่าย เช่น ปุ่มกด ช่องแก้ข้อความ แกนวาดกราฟ หลังจากที่จัดวางองค์ประกอบของ GUI บนหน้าต่างจนได้ตามที่ต้องการแล้ว เมื่อกดปุ่มสั่งให้โปรแกรมบันทึกการออกแบบดังกล่าว แมทแลปจะสร้างไฟล์ 2 ไฟล์ ไฟล์แรกนามสกุล .fig มีหน้าที่เก็บข้อมูลของการจัดวางองค์ประกอบที่สร้างขึ้นไว้ทั้งหมด และไฟล์ที่สองนามสกุล .m บรรจุฟังก์ชันที่เริ่มต้นการทำงานของโปรแกรม GUI ซึ่งในไฟล์นี้มีการสร้างฟังก์ชันย่อยที่สัมพันธ์กับองค์ประกอบที่ผู้ใช้วางบนหน้าต่าง โดยอัตโนมัติด้วย ในส่วนนี้เองที่ผู้เขียนโปรแกรม GUI สามารถแทรกชุดคำสั่งของตนเพื่อให้โปรแกรมทำงานตามที่ต้องการ

บทที่ 3
การออกแบบฟังก์ชันการประมาณค่าด้วยวิธีการปรับค่าอัตราการเรียนรู้

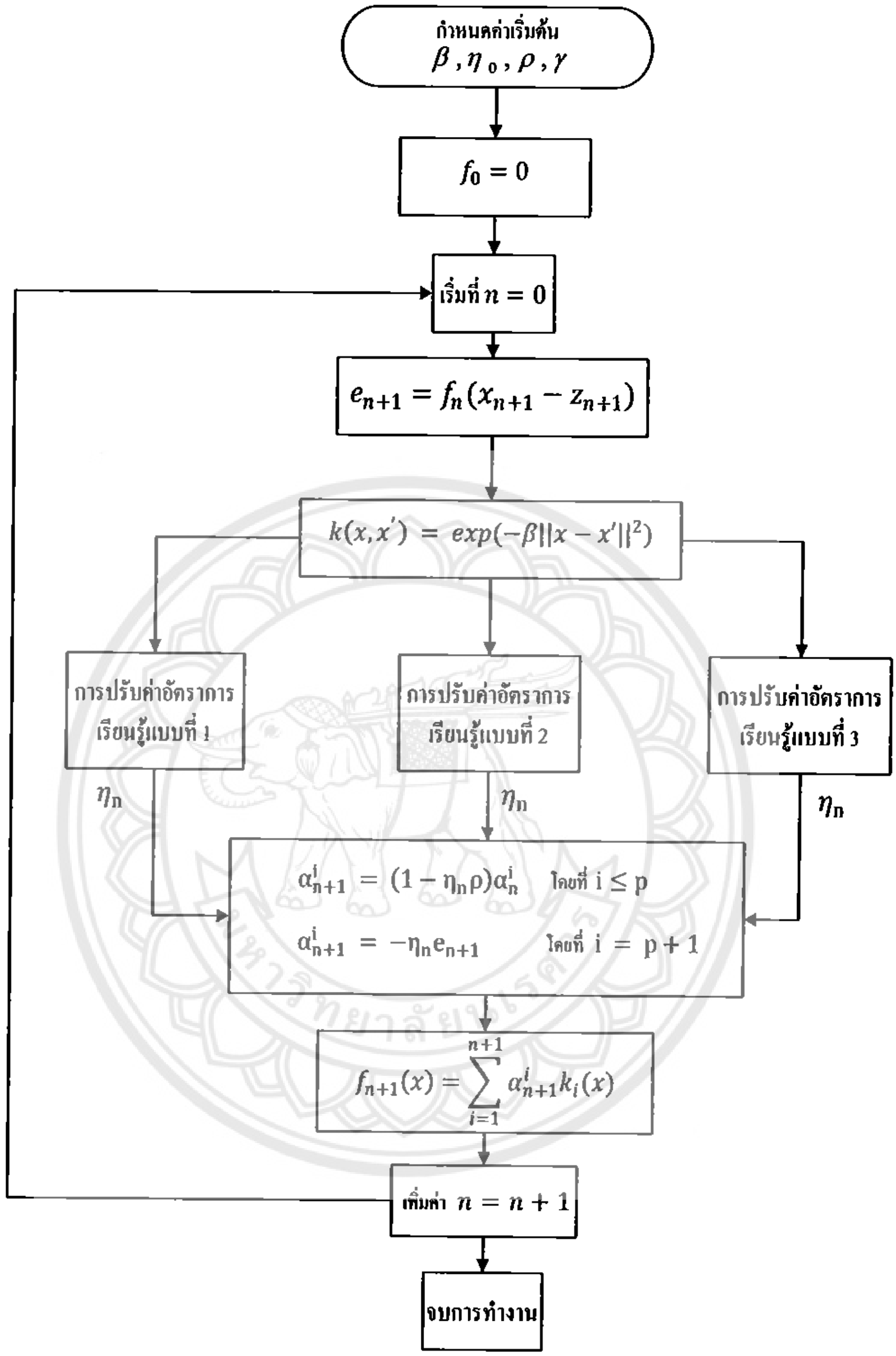
โครงงานนี้ได้มุ่งเน้นศึกษาการออกแบบระบบหรือการหาฟังก์ชันการประมาณค่าที่เราไม่ทราบจากข้อมูลขาเข้าและขาออกที่เก็บค่าได้โดยใช้วิธีเคอร์เนล เพื่อให้สามารถจำลองฟังก์ชันที่เปลี่ยนแปลงตามเวลาได้ จึงต้องให้อัตราการเรียนรู้มีการปรับค่าตลอดเวลาดังทฤษฎีที่เสนอไปแล้วในบทที่ 2

3.1 หลักการออกแบบระบบด้วยวิธีการประมาณค่าฟังก์ชัน

การประมาณค่าฟังก์ชันที่ไม่ทราบค่าที่เปลี่ยนแปลงตามเวลาด้วย**วิธีเคอร์เนล ซึ่งจะใช้การ**ปรับอัตราการเรียนรู้แบบที่ 1-3 จะมีขั้นตอนดังต่อไปนี้

1) กำหนดค่าเริ่มต้นค่า $\beta, \eta_0 . \rho, \gamma , \mu$ และฟังก์ชัน $f_0 = 0$
2) กำหนดค่า n เริ่มที่ $n = 0$
3) หาค่าความผิดพลาดจากสมการ $e_{n+1} = f_n(x_{n+1} - z_{n+1})$
4) หาค่า k จาก สมการที่ (2.7)
5) หาค่า α จากสมการที่ (2.18) และ (2.19) โดยที่อัตราการเรียน**รู้ที่ใช้ในสมการ**ดังกล่าวมีการปรับค่า 3 วิธี
6) หาค่าฟังก์ชันจากสมการ $f_{n+1}(x) = \sum_{i=1}^{n+1} \alpha_{n+1}^{i} k_i(x)$
7) เพิ่มค่า n ทำซ้ำตามขั้นตอนที่ 3 ถึง 6

ขั้นตอนการหาฟังก์ชันการประมาณค่า แสดงดังรูปที่ 3.1



รูปที่ 3.1 ขั้นตอนการออกแบบระบบด้วยวิธีการประมาณค่าฟังก์ชัน

การประมาณค่าฟังก์ชันที่ไม่ทราบค่าที่เปลี่ยนแปลงตามเวลาด้วยวิธีเคอร์เนล จะใช้การปรับอัตราการเรียนรู้แบบที่ 4 (วิธีเอสเอ็มดี) จะมีขั้นตอนดังต่อไปนี้

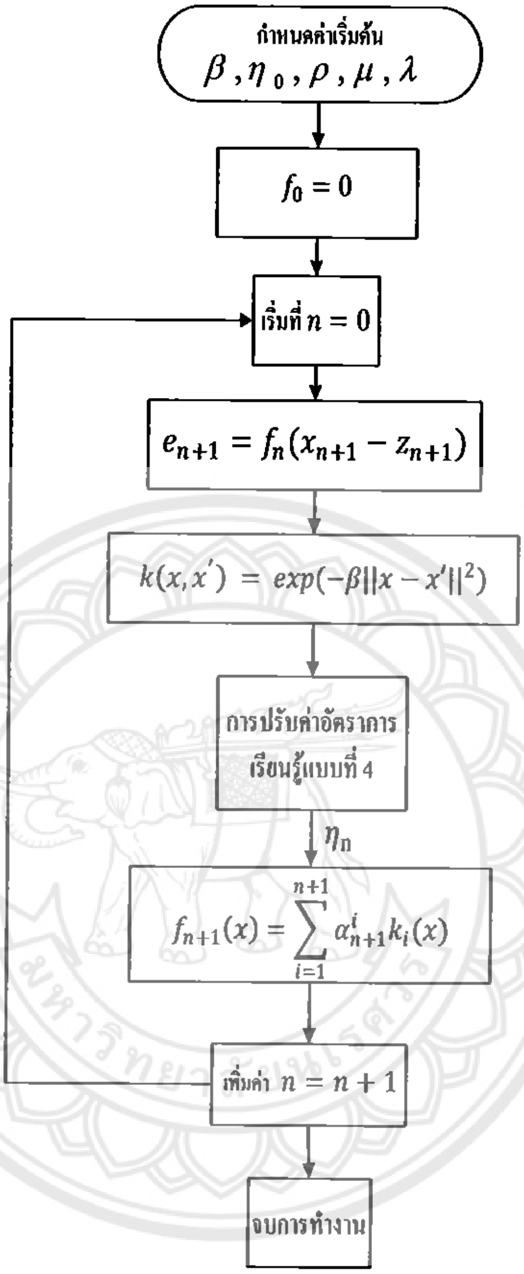
1) กำหนดค่าเริ่มต้นค่า β, η, ρ, μ และฟังก์ชัน $f_0 = 0$
2) กำหนดค่า n เริ่มที่ $n = 0$
3) หาค่าความผิดพลาดจากสมการ $e_{n+1} = f_n(x_{n+1} - z_{n+1})$
4) หาค่า k จาก สมการที่ (2.7)
5) หาค่า v_{n+1} จากสมการ $v_{n+1} \equiv \sum_{i=0}^{n} \lambda^{\mathrm{i}} \frac{\partial(\mathrm{f_n})}{\partial \ln \eta_{\mathrm{n-i}}}$ และ หาค่า β_{n+1}^{i} จากสมการ

$$\beta_{n+1}^{i} = \begin{cases} (1 - \eta_{\mathrm{n}}\rho)\lambda \beta_n^i - \eta_{\mathrm{n}}\rho\alpha_n^i & \text{เมื่อ } i \leq n \\ -\eta_{\mathrm{n}}\lambda v_n(x_{n+1}) - \eta_{\mathrm{n}e_{n+1}} & \text{เมื่อ } i = n+1 \end{cases}$$

ไปแทนในการปรับค่าอัตราการเรียนรู้แบบที่ 4 (วิธีเอสเอ็มดี)

6) หาค่า η_{n} จากสมการที่ (2.49)
7) หาค่า α จากสมการที่ (2.18) และ (2.19)
8) หาค่าฟังก์ชันจากสมการ $f_{n+1}(x) = \sum_{i=1}^{n+1} \alpha_{n+1}^{i} k_i(x)$
9) เพิ่มค่า n ทำซ้ำตามขั้นตอนที่ 3 ถึง 7

ขั้นตอนการหาฟังก์ชันการประมาณค่า แสดงดังรูปที่ 3.2



รูปที่ 3.2 ขั้นตอนการออกแบบระบบด้วยวิธีการประมาณค่าฟังก์ชันแบบ SMD

3.2 วิธีการปรับค่าอัตราการเรียนรู้

จากขั้นตอนการหาฟังก์ชันการประมาณค่า ในการที่จะปรับปรุงค่า α ในข้อที่ 4 จะต้องใช้ค่าอัตราการเรียนรู้ซึ่งต้องเป็นค่าที่ถูกปรับตลอดเวลาเพื่อให้ทันต่อฟังก์ชันที่เราไม่ทราบค่าที่เปลี่ยนแปลงตลอดเวลา ซึ่งจะมี 4 วิธีที่ใช้

1) การปรับค่าอัตราการเรียนรู้แบบที่ 1

กฎการปรับตัวจะแสดงดังสมการที่ (2.27) นั่นคือ

$$\begin{aligned}\eta_n &= \eta_{n-1} - \gamma\left\langle\frac{\partial\hat{g}_{reg}(f_{n+1})}{\partial f_{n+1}}, \frac{\partial\hat{g}_{reg}(f_n)}{\partial f_n}\right\rangle \\ &= \eta_{n-1} + \gamma\langle\nabla\,\hat{g}_{reg}(f_n), \nabla\,\hat{g}_{reg}(f_{n-1})\rangle\end{aligned}$$

2) การปรับค่าอัตราการเรียนรู้แบบที่ 2

กฎการปรับตัวจะแสดงดังสมการที่ (2.33) นั่นคือ

$$\begin{aligned}\eta_n &= \eta_{n-1} + \gamma\eta_{n-1}\langle\nabla\,\hat{g}_{reg}(f_n), \nabla\,\hat{g}_{reg}(f_{n-1})\rangle \\ &= \eta_{n-1}\{1 + \gamma\langle\nabla\,\hat{g}_{reg}(f_n), \nabla\,\hat{g}_{reg}(f_{n-1})\rangle\}\end{aligned}$$

3) การปรับค่าอัตราการเรียนรู้แบบที่ 3

กฎการปรับตัวจะแสดงดังสมการที่ (2.36) และ (2.34) นั่นคือ

$$\eta_n = \eta_{n-1}\left\{1 + \frac{\gamma}{u_n}\langle\nabla\hat{g}_{reg}(f_n), \nabla\hat{g}_{reg}(f_{n-1})\rangle\right\}$$

และ

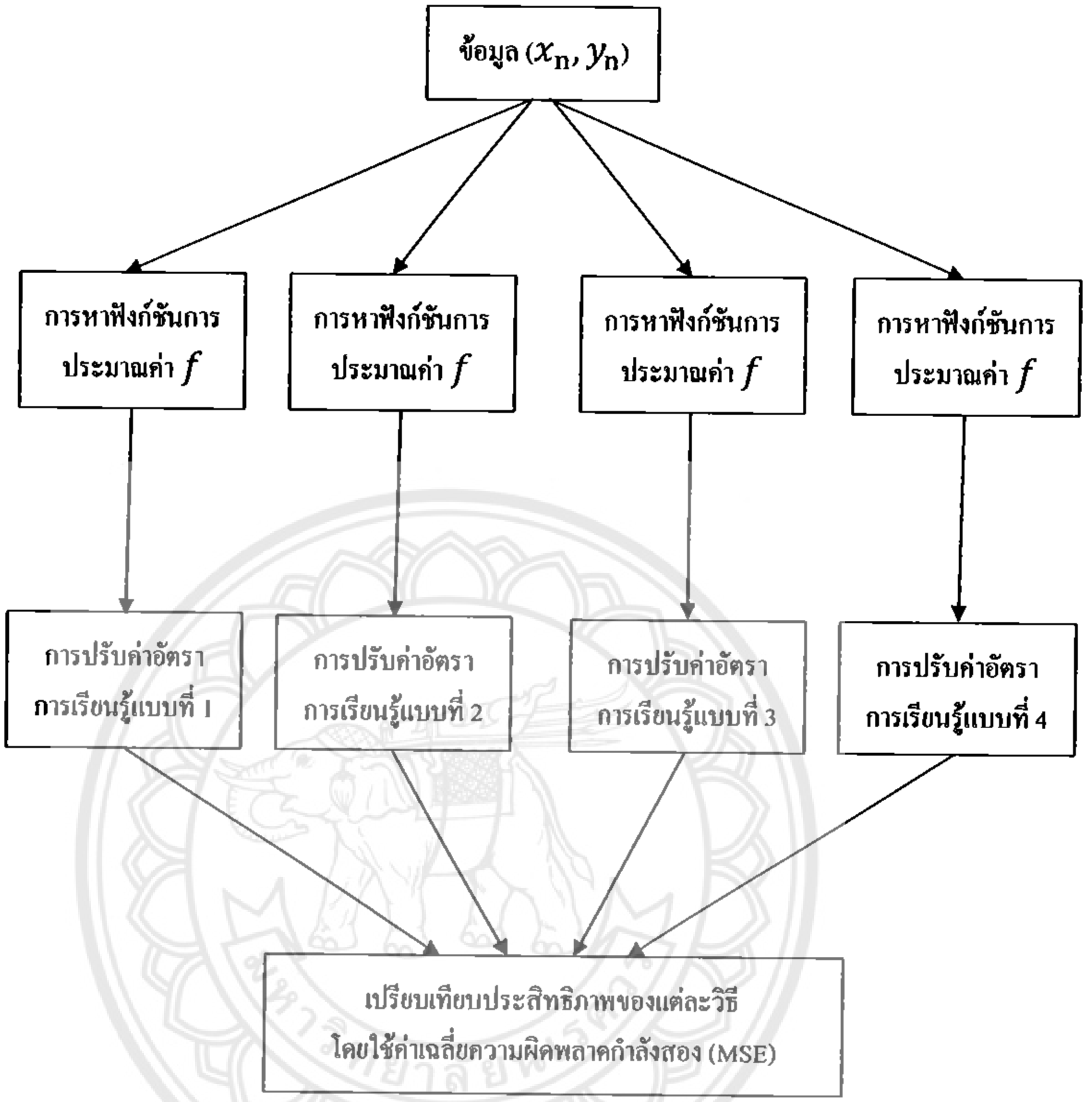
$$u_n = \mu\eta_{n-1} + (1-\mu)\|\nabla\hat{g}_{reg}(f_n)\|^2$$

4) การปรับค่าอัตราการเรียนรู้แบบที่ 4 (วิธีเอสเอ็มดี)

กฎการปรับตัวจะแสดงดังสมการที่ (2.49) นั่นคือ

$$\eta_n = \eta_{n-1}\cdot\max\left(\frac{1}{2}, 1 - \mu(e_{n+1}v_n(x_{n+1}) + \rho\alpha_n^T K\beta_n)\right)$$

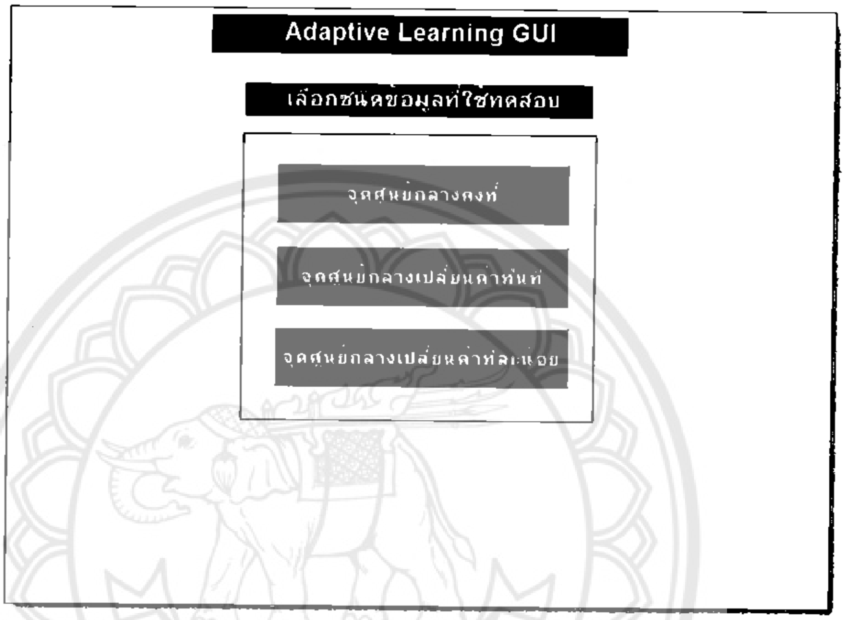
ขั้นตอนการหาฟังก์ชันการประมาณค่าทั้ง 4 วิธี แสดงดังรูปที่ 3.3



รูปที่ 3.3 ขั้นตอนแสดงการทดลองการหาค่าความผิดพลาดของแต่ละวิธี

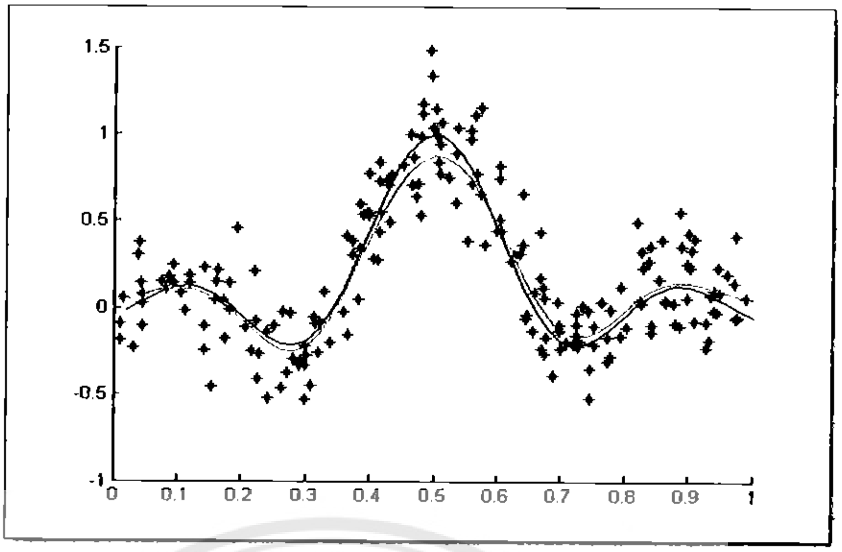
3.3 การออกแบบและการใช้โปรแกรม

ในการใช้งานโดยใช้งานผ่านส่วนต่อประสานกราฟิกกับผู้ใช้ โดยเราออกแบบหน้าต่างการทำงานของโปรแกรม 4 หน้าต่าง ซึ่งหน้าต่างแรกของการใช้งาน คือ การเลือกชนิดข้อมูลที่ใช้ทดสอบ แสดงดังรูปที่ 3.4



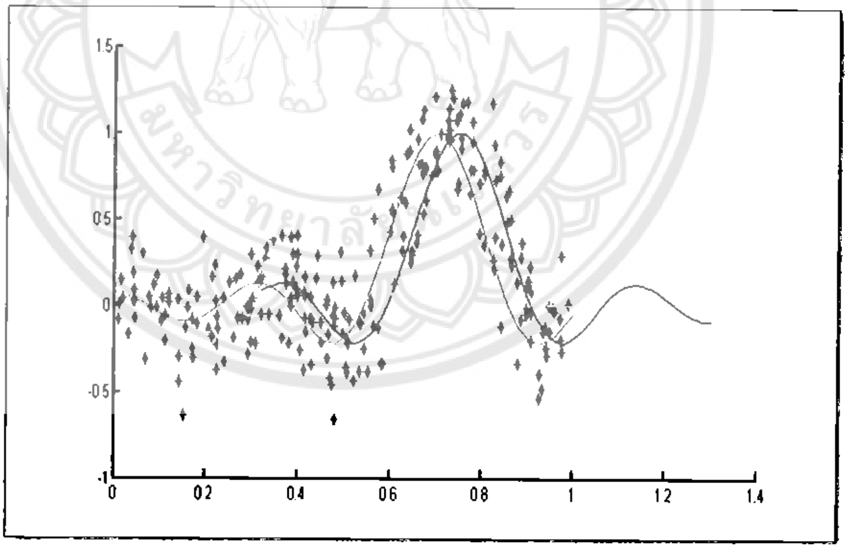
รูปที่ 3.4 หน้าต่างการทำงานของโปรแกรมเมื่อเลือกชนิดข้อมูลที่ใช้ทดสอบ

จากรูปที่ 3.4 เมื่อเริ่มการทำงานของโปรแกรมแล้ว จะมีปุ่มการทำงานหลักๆ อยู่ 3 ปุ่ม ซึ่งจะเป็นการเลือกรูปแบบข้อมูลของฟังก์ชันซิงค์โดยมีทั้งหมด 3 แบบคือ จุดศูนย์กลางคงที่ จุดศูนย์กลางเปลี่ยนค่าทันที และจุดศูนย์กลางเปลี่ยนค่าทีละน้อย เมื่อทำการเลือกปุ่ม จุดศูนย์กลางคงที่ ข้อมูลที่ใช้ทดสอบไม่เปลี่ยนแปลงตามเวลา มีลักษณะข้อมูลที่คงที่ ไม่เปลี่ยนแปลง แสดงดังรูปที่ 3.5



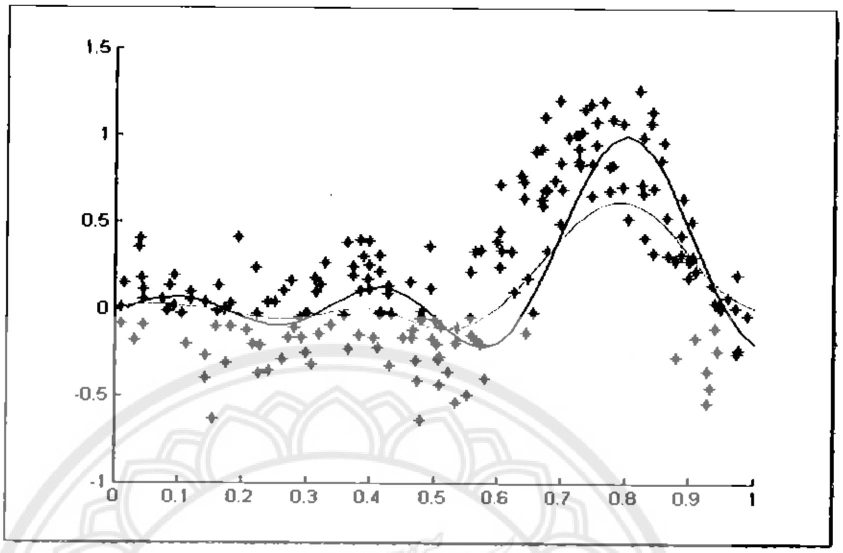
รูปที่ 3.5 ข้อมูลที่จุดศูนย์กลางคงที่

หากทำการเลือกปุ่ม จุดศูนย์กลางเปลี่ยนค่าทันที ข้อมูลที่ใช้ทดสอบเปลี่ยนแปลงตามเวลา มีลักษณะข้อมูลที่มีการเปลี่ยนแปลงอย่างรวดเร็ว แสดงดังรูปที่ 3.6



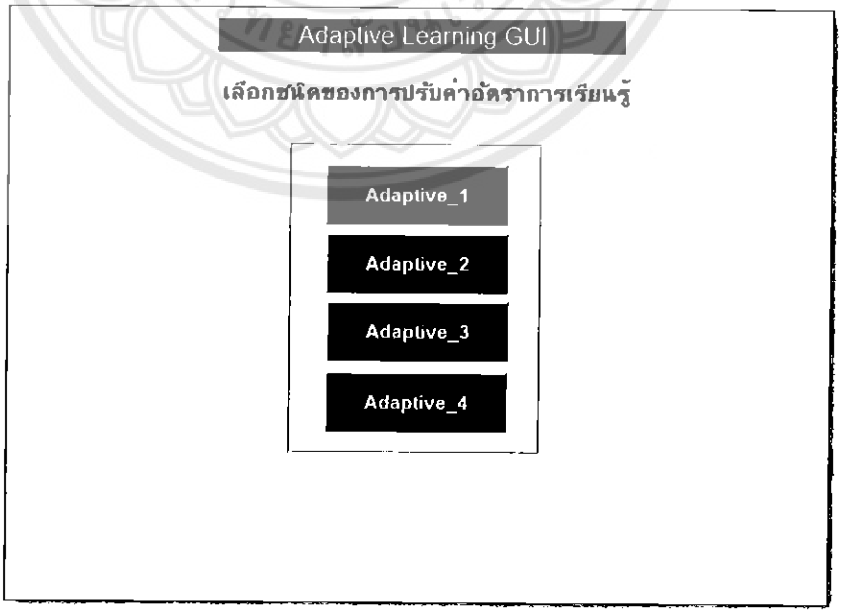
รูปที่ 3.6 ข้อมูลที่จุดศูนย์กลางเปลี่ยนค่าทันที

หากทำการเลือกปุ่ม จุดศูนย์กลางเปลี่ยนค่าทีละน้อย ข้อมูลที่ใช้ทดสอบเปลี่ยนแปลงตามเวลา มีลักษณะข้อมูลที่การเปลี่ยนแปลงอย่างช้าๆ แสดงดังรูปที่ 3.7



รูปที่ 3.7 ข้อมูลที่จุดศูนย์กลางเปลี่ยนค่าทีละน้อย

เมื่อทำการเลือกชนิดของข้อมูลที่ใช้ทดสอบแล้ว หน้าต่างถัดไปทำการเลือกชนิดของการปรับค่าอัตราการเรียนรู้ทั้ง 4 ชนิด แสดงดังรูปที่ 3.8



รูปที่ 3.8 หน้าต่างการทำงานของโปรแกรมเมื่อเลือกชนิดของการปรับค่าอัตราการเรียนรู้

จากรูปที่ 3.8 หน้าต่างนี้ใช้เลือกชนิดของการปรับค่าอัตราการเรียนรู้ มีชนิดข้อมูลที่ใช้ทดสอบให้เลือก 4 แบบซึ่งสามารถอธิบายส่วนต่างๆ ของโปรแกรม ได้ดังต่อไปนี้

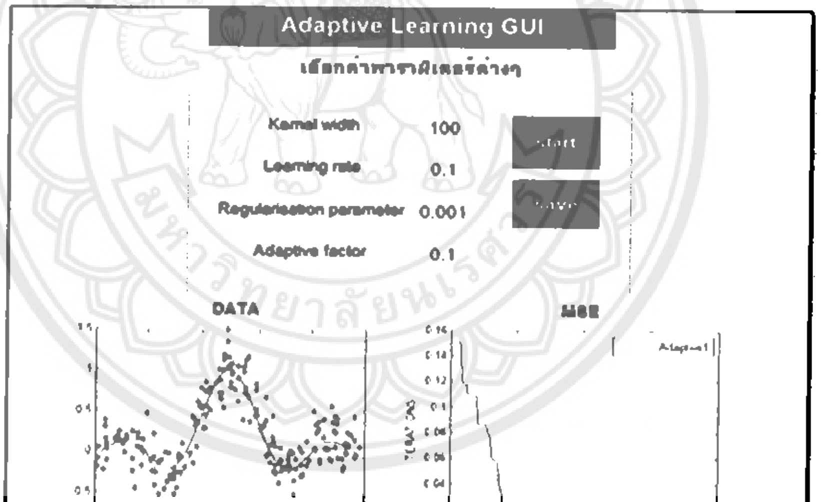
ทำการเลือกปุ่ม Adaptive_1 การปรับค่าอัตราการเรียนรู้ของฟังก์ชันที่ไม่ทราบค่าของเราที่มีการเปลี่ยนแปลงอย่างช้าๆ

ทำการเลือกปุ่ม Adaptive_2 การปรับค่าอัตราการเรียนรู้ของฟังก์ชันที่ไม่ทราบค่าของเราที่มีการเปลี่ยนแปลงอย่างอย่างรวดเร็ว โดยอาศัยการปรับค่าด้วยวิธีเรขาคณิต

ทำการเลือกปุ่ม Adaptive_3 การปรับค่าอัตราการเรียนรู้ของฟังก์ชันที่ไม่ทราบค่าของเรา

ทำการเลือกปุ่ม Adaptive_4 การปรับค่าอัตราการเรียนรู้ของฟังก์ชันที่ไม่ทราบค่าของเราแบบ เอสเอ็มดี

เมื่อทำการเลือกชนิดของข้อมูล หน้าต่างถัดไปจะเป็นการเลือกค่าพารามิเตอร์ต่างๆ แสดงดังรูปที่ 3.8



รูปที่ 3.9 หน้าต่างการทำงานของโปรแกรมเมื่อเลือกค่าพารามิเตอร์ข้อมูลชนิดจุดศูนย์กลางคงที่

จากรูปที่ 3.9 หน้าต่างการทำงานของโปรแกรม GUI ที่สร้างขึ้นมาเพื่อเลือกค่าพารามิเตอร์ต่างๆ ซึ่งสามารถอธิบายส่วนต่างๆ ของโปรแกรม ได้ดังต่อไปนี้

กราฟแสดงค่า MSE เป็นกราฟที่ใช้สำหรับแสดงค่าเฉลี่ยความผิดพลาดกำลังสองที่ได้จากการประมาณค่าฟังก์ชัน ที่เกิดจากการรันโปรแกรมของค่าพารามิเตอร์ต่างๆที่ได้ถูกกำหนดไว้

กราฟด้านล่างแสดงค่าข้อมูลที่ใช้สอน ข้อมูลที่ใช้ทดสอบ จากกราฟจุดสีแดง คือข้อมูลที่ใช้สอน เส้นสีน้ำเงิน คือข้อมูลที่ใช้ทดสอบ ส่วนเส้นสีเขียวคือฟังก์ชันการประมาณค่าที่โมเดลสร้างขึ้นซึ่งเป็นค่าที่ผู้ใช้ได้กำหนดให้กับโปรแกรมก่อนที่โปรแกรมจะเริ่มทำงาน

พารามิเตอร์ต่างๆที่ใช้สำหรับรูปที่ 3.9 คือค่า beta, learnrate, rho, gamma, mu และ lamda หากเลือกค่าพารามิเตอร์ที่เหมาะสม เมื่อทำการกดปุ่ม start โปรแกรมจะได้ค่าที่เหมาะสม เมื่อเทียบจากข้อมูลที่ใช้สอนและข้อมูลที่ใช้ทดสอบ และเมื่อต้องการบันทึกค่าพารามิเตอร์ สามารถทำได้โดยการกดปุ่ม save

บทที่ 4
ผลการทดลองและการวิเคราะห์ผล

ในบทนี้จะกล่าวถึงผลการทดลองของวิธีการปรับค่าอัตราการเรียนรู้ทั้ง 4 แบบที่ข้อมูลมีจุดศูนย์กลางคงที่ จุดศูนย์กลางเปลี่ยนค่าทันที และจุดศูนย์กลางเปลี่ยนค่าเล็กน้อย เป็นฟังก์ชันซิงค์ที่จะทำการเปลี่ยนค่าพารามิเตอร์ต่างๆและดูผลที่เกิดขึ้น โดยแต่ละค่าพารามิเตอร์ที่แตกต่างกันจะให้ผลของค่าความผิดพลาด (ค่าความผิดพลาดเฉลี่ยกำลังสองหรือ MSE) ที่แตกต่างกันออกไป แต่สิ่งที่ต้องการคือมีการลดลงของค่าความผิดพลาดและการลดลงต้องค่อยๆลดลง

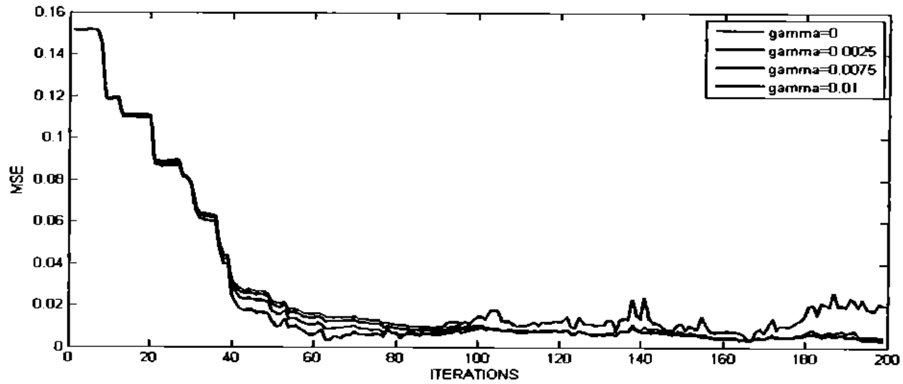
4.1 การปรับค่าอัตราการเรียนรู้วิธีที่ 1

4.1.1 ข้อมูลที่มีจุดศูนย์กลางคงที่

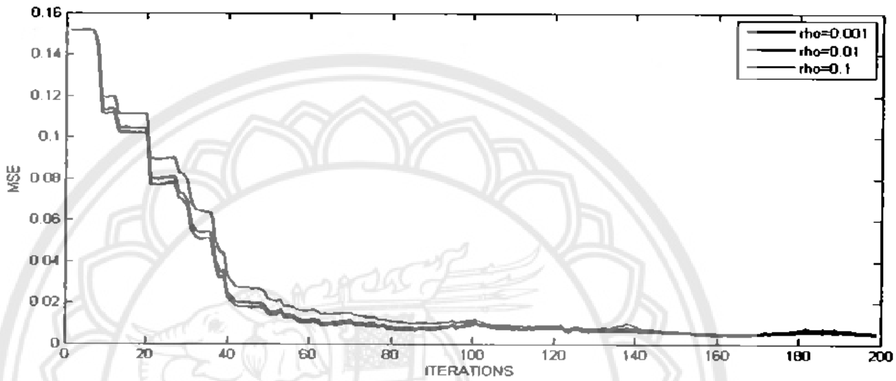
ข้อมูลที่ใช้จะมีชุดสอนและชุดทดสอบซึ่งในที่นี้ชุดสอนจะมี 200 ค่าและชุดทดสอบมี 50 ค่าและจะกำหนดให้ค่าอัตราการเรียนรู้เท่ากับ

$$\begin{aligned}\eta_n &= \eta_{n-1} - \gamma \left\langle \frac{\partial \hat{g}_{reg}(f_{n+1})}{\partial f_{n+1}}, \frac{\partial \hat{g}_{reg}(f_n)}{\partial f_n} \right\rangle \\ &= \eta_{n-1} + \gamma \langle \nabla \hat{g}_{reg}(f_n), \nabla \hat{g}_{reg}(f_{n-1}) \rangle\end{aligned}$$

เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.1 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแกมมาและค่าเรกกูลาไรเซชัน เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด สามารถแสดงดังรูปที่ 4.1



(ก)



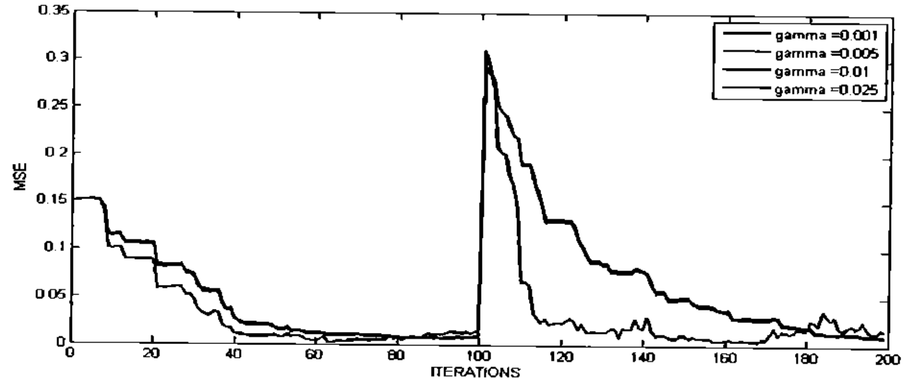
(ข)

รูปที่ 4.1 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 1 แบบข้อมูลที่มีจุดศูนย์กลางคงที่

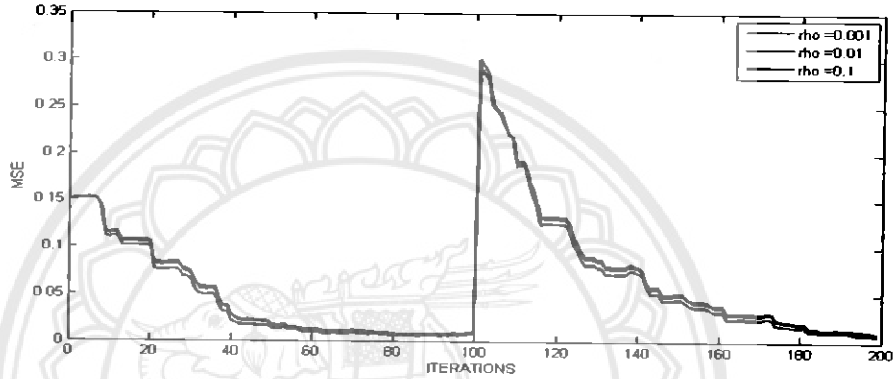
จากรูปที่ 4.1 ค่าแกมมาและค่าเรกกูลาไรเซชันแต่ละค่าทำให้ได้ค่าความผิดพลาดที่ค่อนข้างใกล้เคียงกัน แต่ค่าที่ดีที่สุดคือ $\gamma = 0.01$ และ $\rho = 0.001$ โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ ค่าแกมมาและค่าเรกกูลาไรเซชันอื่นๆ มีค่าความผิดพลาดที่ได้จะมีค่ามากและการลดลงของค่าความผิดพลาดไม่ดี

4.1.2 ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทันที

เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.1 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแกมมาและค่าเรกกูลาไรเซชัน เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด สามารถแสดงดังรูปที่ 4.2



(ก)



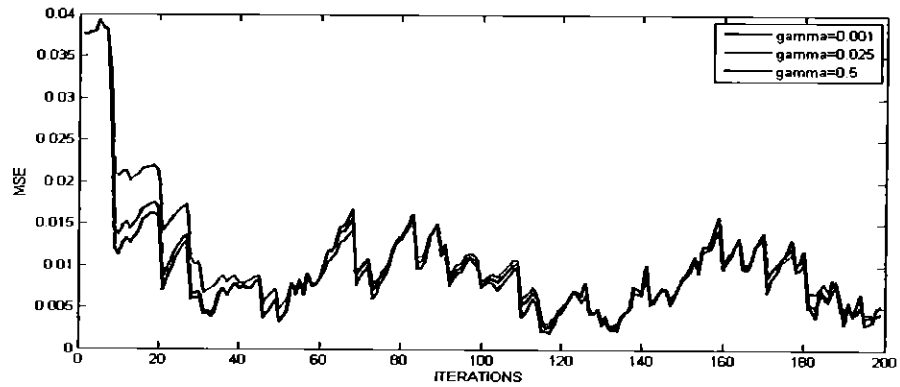
(ข)

รูปที่ 4.2 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 1 แบบข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทันที

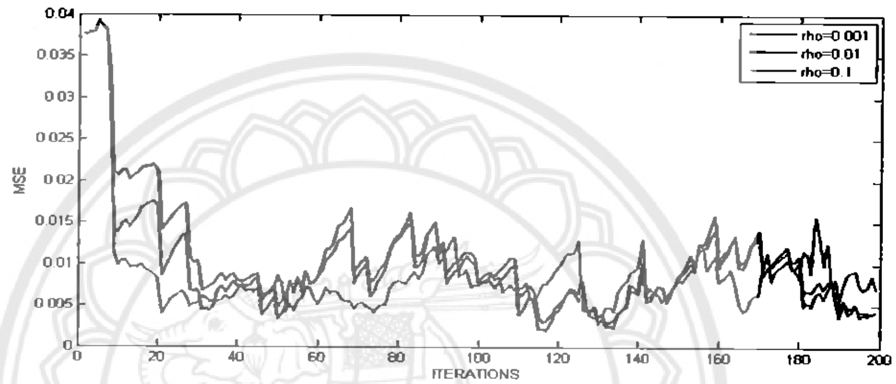
จากรูปที่ 4.2 ค่าความผิดพลาดที่ได้จากการปรับค่าแกมมาและค่าเรกกูลาไรเซชันจะเห็นว่าค่าความผิดพลาดที่ $\gamma = 0.025$ และ $\rho = 0.01$ มีการลดลงของค่าความผิดพลาดที่ดี เมื่อเทียบกับค่าค่าพารามิเตอร์อื่นๆ โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ ค่าแกมมาและค่าเรกกูลาไรเซชันอื่นๆ ค่าความผิดพลาดที่ได้จะมีค่ามากและการลดลงของค่าความผิดพลาดไม่ดี

4.1.3 ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทีละน้อย

เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.4 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแกมมาและค่าเรกกูลาไรเซชัน เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด สามารถแสดงดังรูปที่ 4.3



(ก)



(ข)

รูปที่ 4.3 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 1 แบบข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทีละน้อย

จากรูปที่ 4.3 ค่าความผิดพลาดที่ได้จากการปรับค่าแกมมาและค่าเรกกูลาไรเซชันจะเห็นว่าค่าความผิดพลาดที่ค่า $\gamma = 0.001$ และ $\rho = 0.1$ มีการลดลงของค่าความผิดพลาดที่ดี เมื่อเทียบกับค่าค่าพารามิเตอร์อื่นๆ โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ ค่าแกมมาและค่าเรกกูลาไรเซชันอื่นๆ มีค่าความผิดพลาดที่ได้จะมีค่ามากและการลดลงของค่าความผิดพลาดไม่ดี

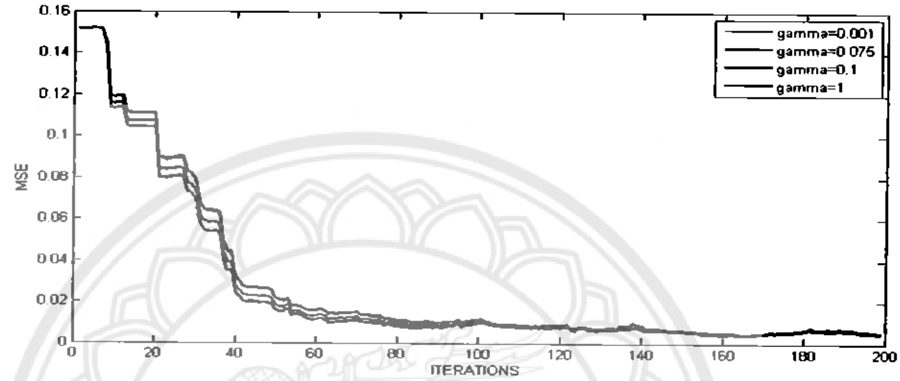
4.2 การปรับค่าอัตราการเรียนรู้วิธีที่ 2

4.2.1 ข้อมูลที่มีจุดศูนย์กลางคงที่

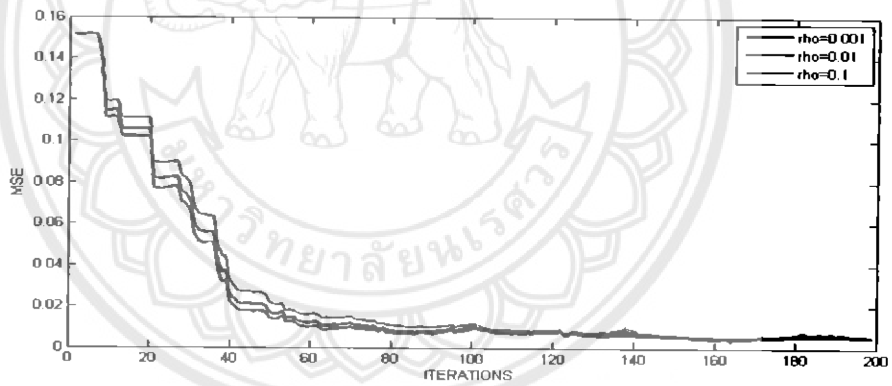
ข้อมูลที่ใช้จะมีชุดสอนและชุดทดสอบซึ่งในที่นี้ชุดสอนจะมี 200 ค่า และชุดทดสอบมี 50 ค่า และจะกำหนดให้ค่าอัตราการเรียนรู้เท่ากับ

$$\eta_n = \eta_{n-1} + \gamma\eta_{n-1}\langle \nabla\, \hat{g}_{reg}(f_n), \nabla\, \hat{g}_{reg}(f_{n-1})\rangle$$
$$= \eta_{n-1}\{1 + \gamma\langle \nabla\, \hat{g}_{reg}(f_n), \nabla\, \hat{g}_{reg}(f_{n-1})\rangle\}$$

เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.1 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแกมมาและค่าเรกกูลาไรเซชัน เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด สามารถแสดงดังรูปที่ 4.4



(ก)



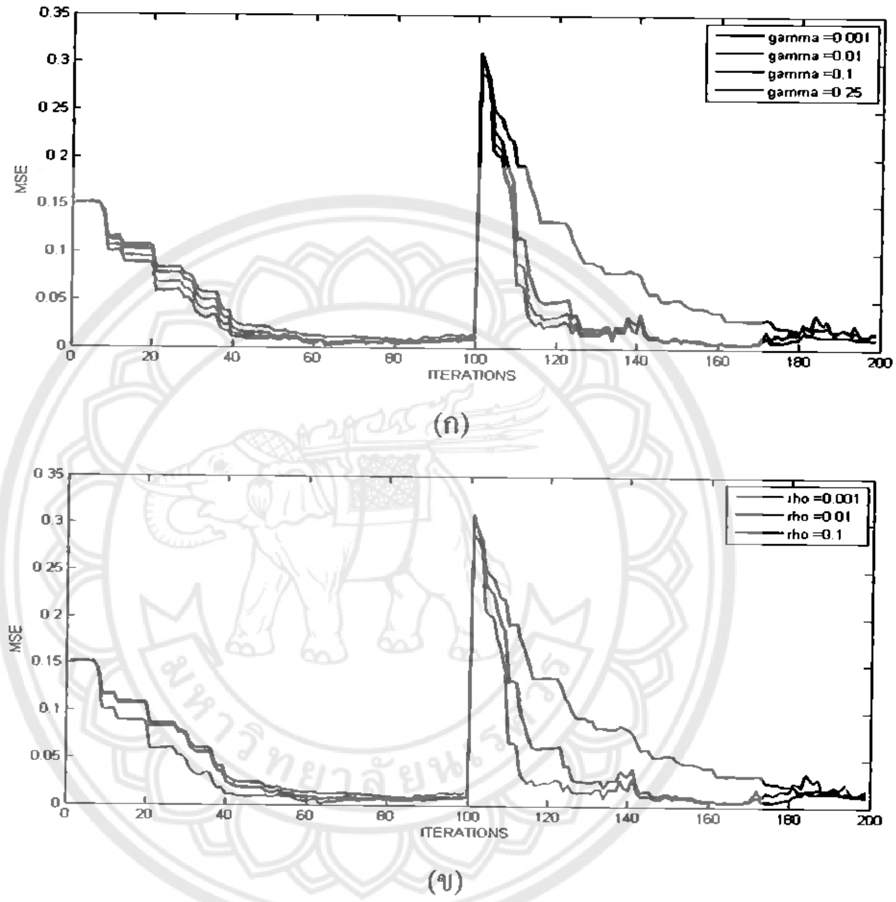
(ข)

รูปที่ 4.4 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 2 แบบข้อมูลที่มีจุดศูนย์กลางคงที่

จากรูปที่ 4.4 ค่าความผิดพลาดที่ได้จากการปรับค่าแกมมาและค่าเรกกูลาไรเซชันจะเห็นว่าค่าความผิดพลาดที่ $\gamma = 0.1$ และ $\rho = 0.001$ มีการลดลงของค่าความผิดพลาดที่ดี เมื่อเทียบกับค่าพารามิเตอร์อื่นๆ โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ ค่าแกมมาและค่าเรกกูลาไรเซชันอื่นๆ มีค่าความผิดพลาดที่ได้จะมีค่ามากและการลดลงของค่าความผิดพลาดไม่ดี

4.2.2 ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทันที

เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.1 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแกมมาและค่าเรกกูลาไรเซชัน เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด สามารถแสดงดังรูปที่ 4.5

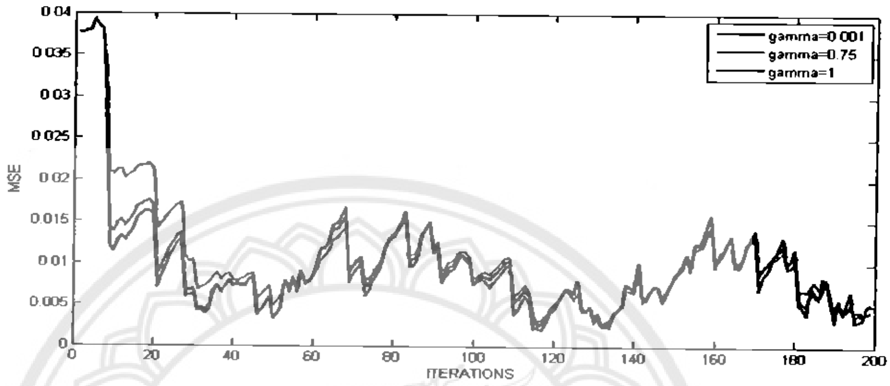


รูปที่ 4.5 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 2 แบบข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทันที

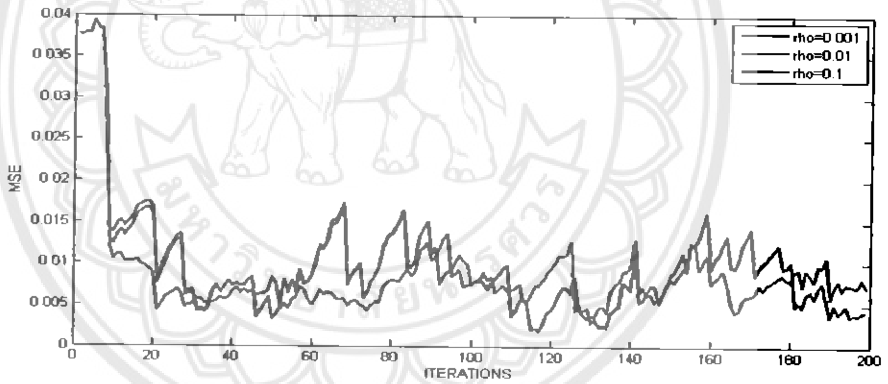
จากรูปที่ 4.5 ความผิดพลาดที่ $\gamma = 0.25$ และ $\rho = 0.001$ มีการลดลงของค่าความผิดพลาดที่ดี โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ ค่าแกมมาและค่าเรกกูลาไรเซชันค่าอื่นๆ มีค่าความผิดพลาดที่ได้จะมีค่ามากและการลดลงของค่าความผิดพลาดไม่ดี

4.2.3 ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทีละน้อย

เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.4 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแกมมาและค่าเรกกูลาไรเซชัน เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด สามารถแสดงดังรูปที่ 4.6



(ก)



(ข)

รูปที่ 4.6 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 2 แบบข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทีละน้อย

จากรูปที่ 4.6 ค่าความผิดพลาดที่ได้จากการปรับค่าแกมมาและค่าเรกกูลาไรเซชันจะเห็นว่าค่าความผิดพลาดที่ $\gamma = 0.001$ และ $\rho = 0.001$ มีการลดลงของค่าความผิดพลาดที่ดี โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ ค่าแกมมาและค่าเรกกูลาไรเซชัน ค่าความผิดพลาดที่ได้จะมีค่ามากและการลดลงของค่าความผิดพลาดไม่ดี

4.3 การปรับค่าอัตราการเรียนรู้วิธีที่ 3

4.3.1 ข้อมูลที่มีจุดศูนย์กลางค่าคงที่

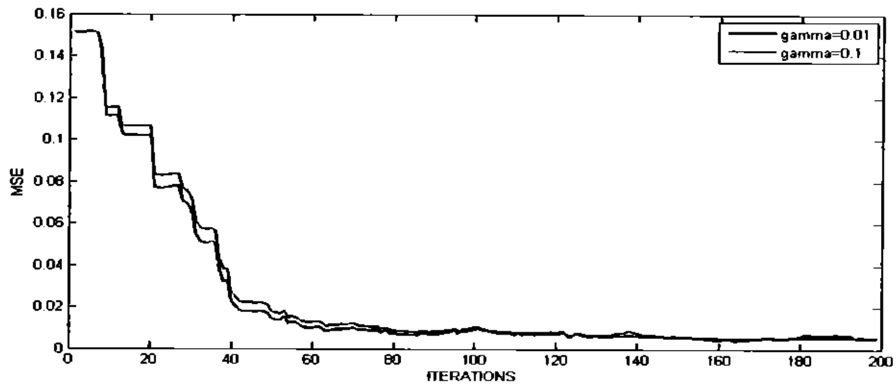
ข้อมูลที่ใช้จะมีชุดสอนและชุดทดสอบซึ่งในที่นี้ชุดสอนจะมี 200 ค่าและชุดทดสอบมี 50 ค่าและจะกำหนดให้ค่าอัตราการเรียนรู้เท่ากับ

$$\eta_{\mathrm{n}} = \eta_{n-1}\{1 + \frac{\gamma}{u_n}\langle \nabla \hat{g}_{reg}(f_n), \nabla \hat{g}_{reg}(f_{n-1})\rangle\}$$

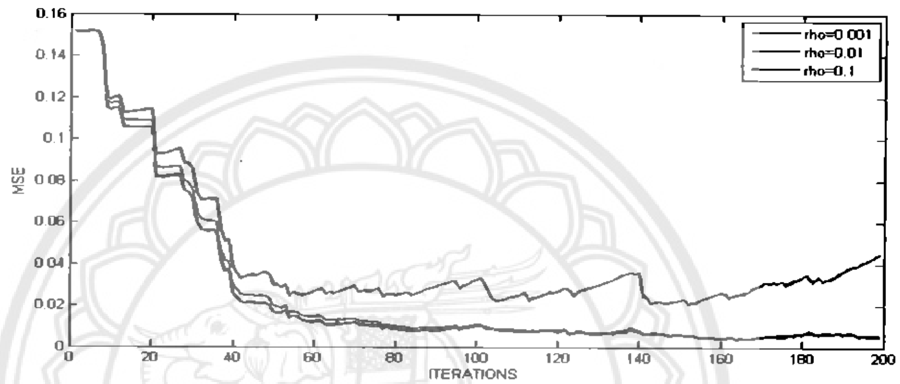
$$u_{\mathrm{n}} = \mu u_{n-1} + (1-\mu)\|\nabla \hat{g}_{reg}(f_n)\|^2$$

ในการปรับค่าพารามิเตอร์ เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.1 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแกมมา ค่าเรกกูลาไรเซชัน และค่ามิว เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด

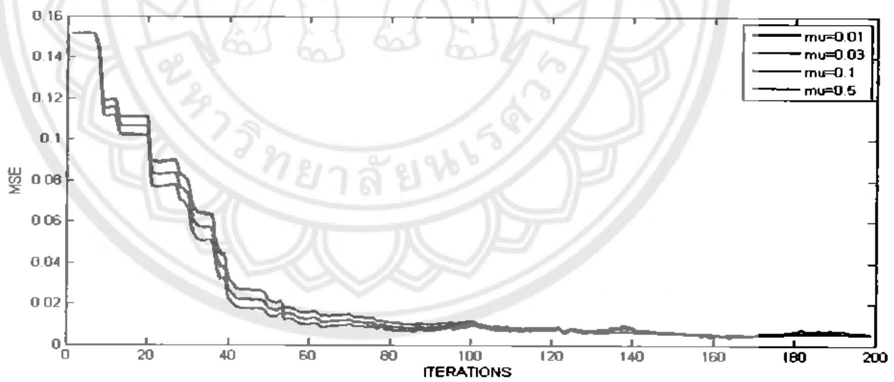
ค่าความผิดพลาดที่ได้จากการปรับค่าแกมมา ค่าเรกกูลาไรเซชัน และค่ามิว จะเห็นว่าค่าความผิดพลาดที่ได้มีการลดลงของค่าความผิดพลาดที่ดี โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ โดยค่าความผิดพลาดที่ $\gamma = 0.01$ มีค่าที่ดีที่สุด ส่วนค่าแกมมาค่าอื่นๆ มีค่าความผิดพลาดที่มากและการลดลงของค่าความผิดพลาดไม่ดี ค่าความผิดพลาดที่ $\rho = 0.001$ มีค่าที่ดีที่สุด ส่วนค่าเรกกูลาไรเซชันค่าอื่นๆ มีค่าความผิดพลาดที่มากและการลดลงของค่าความผิดพลาดไม่ดี ค่าความผิดพลาดที่ $\mu = 0.1$ มีค่าที่ดีที่สุด ส่วนค่ามิวค่าอื่นๆ มีค่าความผิดพลาดที่มากและการลดลงของค่าความผิดพลาดไม่ดี สามารถแสดงดังรูปที่ 4.7



(ก)



(ข)

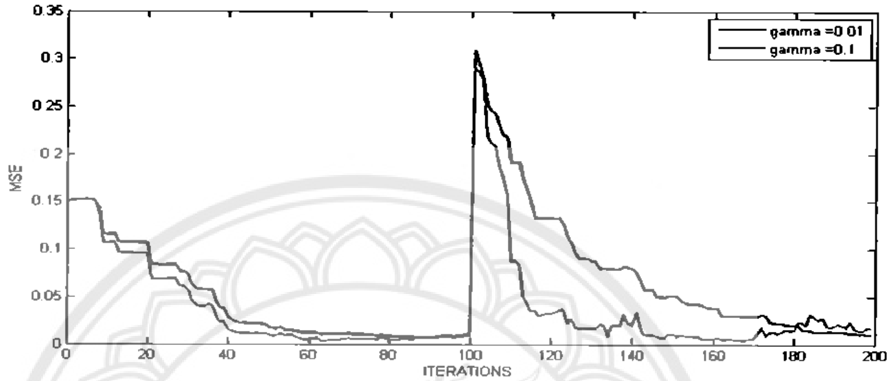


(ค)

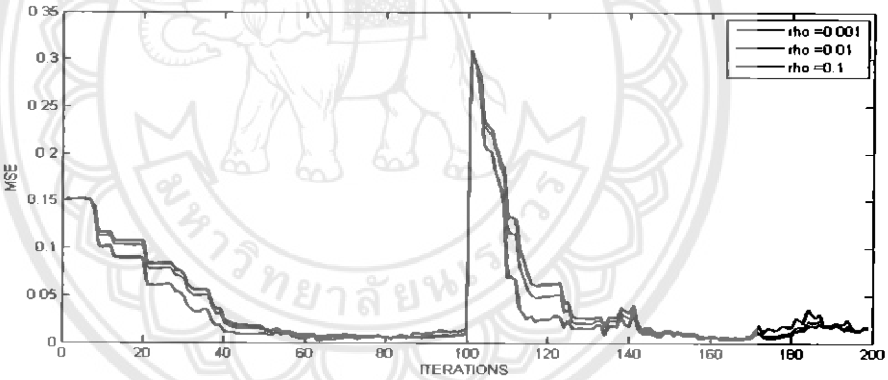
รูปที่ 4.7 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 3 แบบข้อมูลที่มีจุดศูนย์กลางคงที่

4.3.2 ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทันที

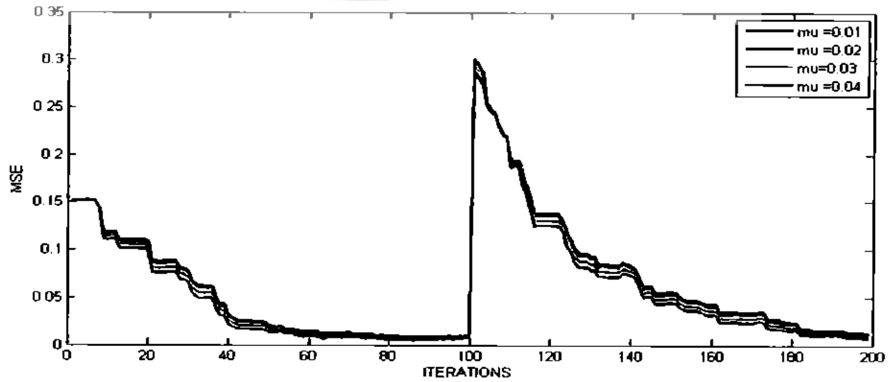
ผลที่ได้จากการปรับค่าแกมมา ค่าเรกกูลาไรเซชัน และค่ามิว แล้วหาค่าความผิดพลาด โดยเราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.1 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าค่าพารามิเตอร์ต่างๆ เพื่อให้ค่าเฉลี่ยความผิดพลาดมีค่าน้อยที่สุด สามารถแสดงดังรูปที่ 4.8



(ก)



(ข)



(ค)

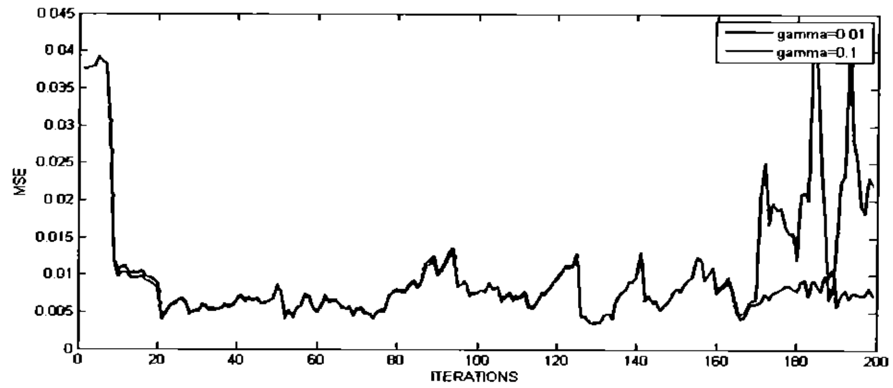
รูปที่ 4.8 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 3 แบบข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทันที

จากรูปที่ 4.8 ค่าความผิดพลาดที่ได้จากการปรับค่าแกมมา ค่าเรกกูลาไรเซชัน และค่ามิว จะเห็นว่าค่าความผิดพลาดที่ $\gamma = 0.1$, $\rho = 0.01$ และ $\mu = 0.04$ มีการลดลงของค่าความผิดพลาดที่ดี โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ ค่าแกมมา ค่าเรกกูลาไรเซชัน และค่ามิว ค่าความผิดพลาดที่ได้จะมีค่ามากและการลดลงของค่าความผิดพลาดไม่ดี

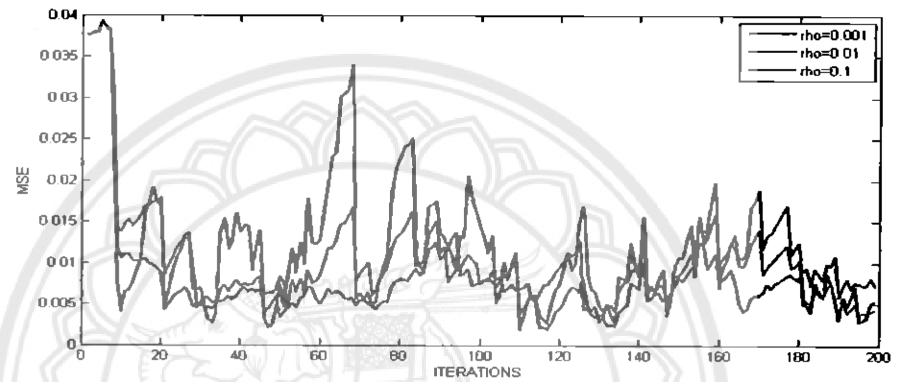
4.3.3 ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทีละน้อย

เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.4 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแกมมา ค่าเรกกูลาไรเซชัน และค่ามิว เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด

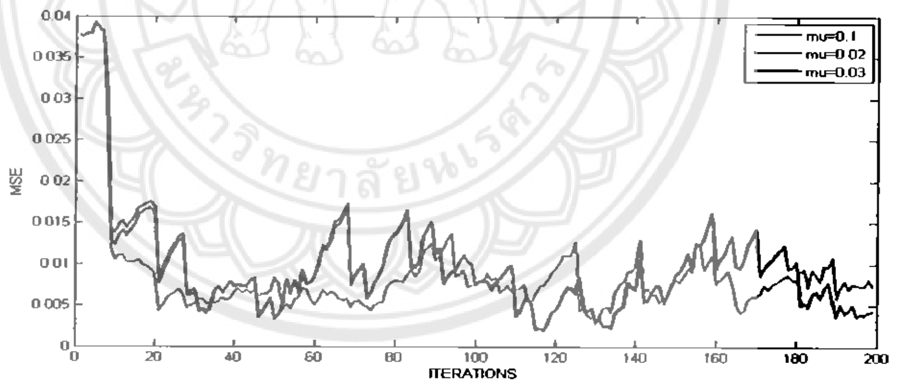
ค่าความผิดพลาดที่ได้จากการปรับค่าแกมมา ค่าเรกกูลาไรเซชัน และค่ามิว จะเห็นว่าค่าความผิดพลาดที่ได้มีการลดลงของค่าความผิดพลาดที่ดี โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ โดยค่าความผิดพลาดที่ $\gamma = 0.01$ มีค่าที่ดีที่สุด ส่วนค่าแกมมาค่าอื่นๆ มีค่าความผิดพลาดที่มากและการลดลงของค่าความผิดพลาดไม่ดี ค่าความผิดพลาดที่ $\rho = 0.1$ มีค่าที่ดีที่สุด ส่วนค่าเรกกูลาไรเซชันค่าอื่นๆ มีค่าความผิดพลาดที่มากและการลดลงของค่าความผิดพลาดไม่ดี ค่าความผิดพลาดที่ $\mu = 0.1$ มีค่าที่ดีที่สุด ส่วนค่ามิวค่าอื่นๆ มีค่าความผิดพลาดที่มากและการลดลงของค่าความผิดพลาดไม่ดี สามารถแสดงดังรูปที่ 4.9



(ก)



(ข)



(ค)

รูปที่ 4.9 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 3 แบบข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทีละน้อย

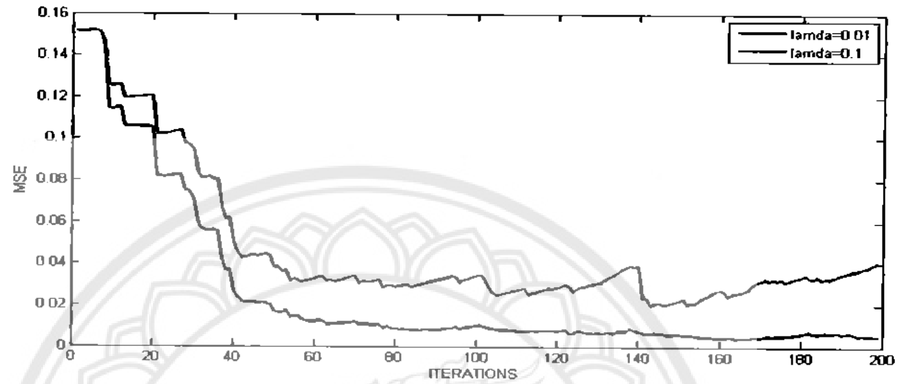
4.4 การปรับค่าอัตราการเรียนรู้วิธีที่ 4

4.4.1 ข้อมูลที่มีจุดศูนย์กลางค่าคงที่

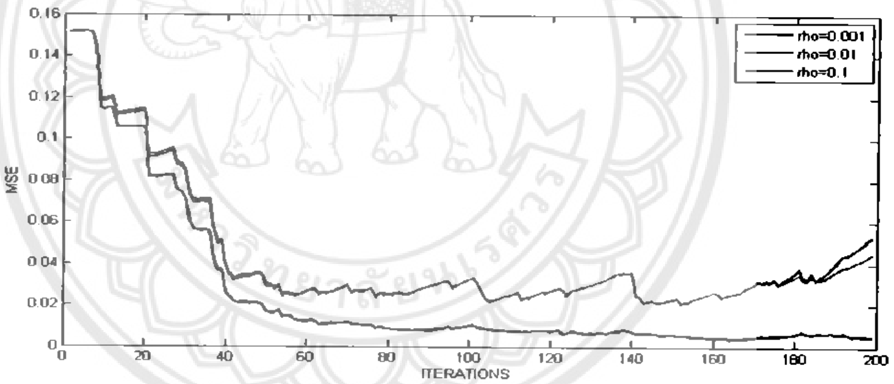
ข้อมูลที่ใช้จะมีชุดสอนและชุดทดสอบซึ่งในที่นี้ชุดสอนจะมี 200 ค่าและชุดทดสอบมี 50 ค่าและจะกำหนดให้ค่าอัตราการเรียนรู้เท่ากับ

$$\eta_{\mathbf{n}} = \eta_{\mathbf{n-1}} \cdot \max(\frac{1}{2}, 1 - \mu(e_{n+1}v_n(x_{n+1}) + \rho\alpha_n^T K\beta_{\mathrm{n}}))$$

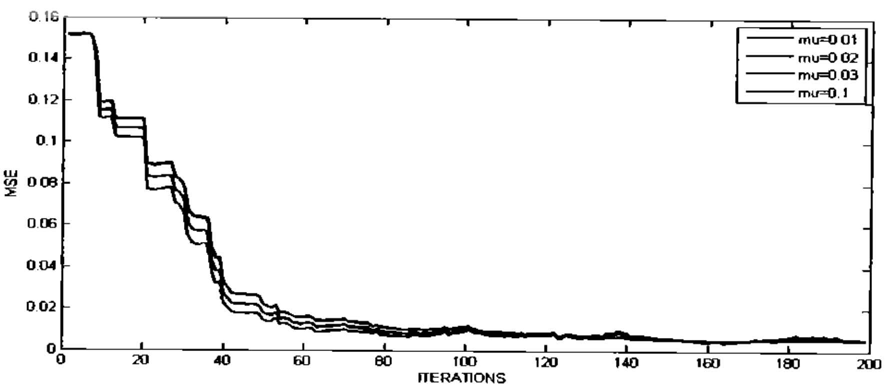
เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.4 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแลมดา ค่าเรกกูลาไรเซชัน และค่ามิว เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด สามารถแสดงดังรูปที่ 4.10



(ก)



(ข)



(ค)

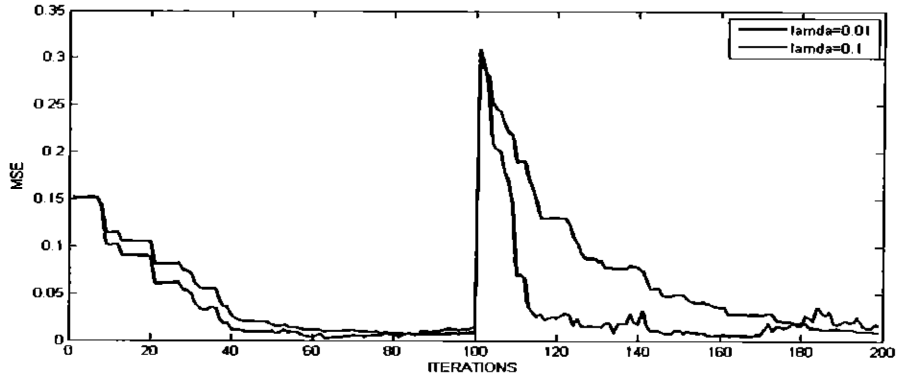
รูปที่ 4.10 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 4 แบบข้อมูลที่มีจุดศูนย์กลางคงที่

จากรูปที่ 4.10 ค่าความผิดพลาดที่ได้จากการปรับค่าแลมดา ค่าเรกกูลาไรเซชัน และค่ามิว จะเห็นว่าค่าความผิดพลาดที่ $\lambda = 0.01, \rho = 0.001$ และ $\mu = 0.1$ มีการลดลงของค่าความผิดพลาดที่ดี โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ และค่าพารามิเตอร์ที่ทำให้ค่าความผิดพลาดมีค่าที่น้อยที่สุด คือ $\gamma = 0.01, \rho = 0.001$ และ $\mu = 0.1$

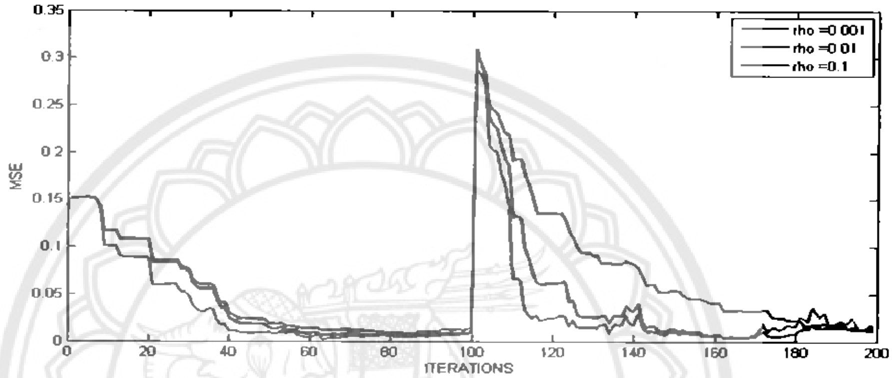
4.4.2 ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทันที

เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.4 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแลมดา ค่าเรกกูลาไรเซชัน และมิว เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด สามารถแสดงดังรูปที่ 4.11

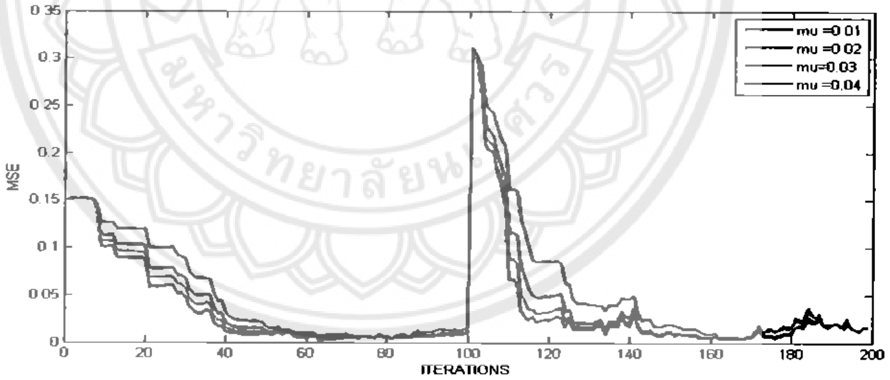
ค่าความผิดพลาดที่ได้จากการปรับค่าแลมดา ค่าเรกกูลาไรเซชัน และค่ามิว จะเห็นว่าค่าความผิดพลาดที่ได้มีการลดลงของค่าความผิดพลาดที่ดี โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ โดยค่าความผิดพลาดที่ $\gamma = 0.01$ มีค่าที่ดีที่สุด ส่วนค่าแกมมาค่าอื่นๆ มีค่าความผิดพลาดที่มากและการลดลงของค่าความผิดพลาดไม่ดี ค่าความผิดพลาดที่ $\rho = 0.001$ มีค่าที่ดีที่สุด ส่วนค่าเรกกูลาไรเซชันค่าอื่นๆ มีค่าความผิดพลาดที่มากและการลดลงของค่าความผิดพลาดไม่ดี ค่าความผิดพลาดที่ $\mu = 0.01$ มีค่าที่ดีที่สุด ส่วนค่ามิวค่าอื่นๆ มีค่าความผิดพลาดที่มากและการลดลงของค่าความผิดพลาดไม่ดี สามารถแสดงดังรูปที่ 4.11



(ก)



(ข)

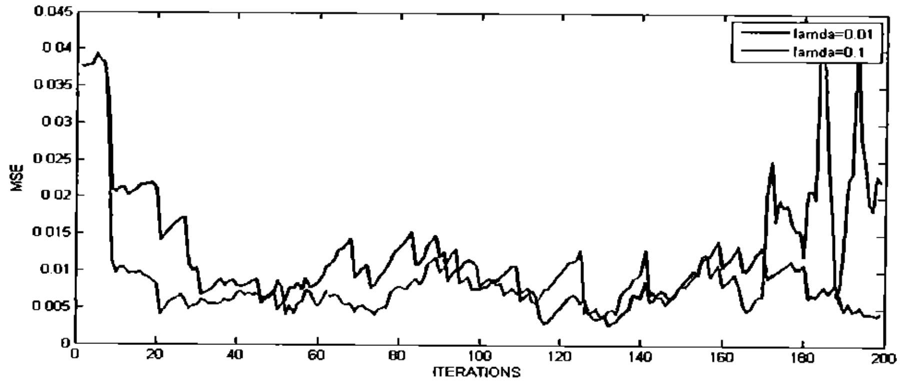


(ค)

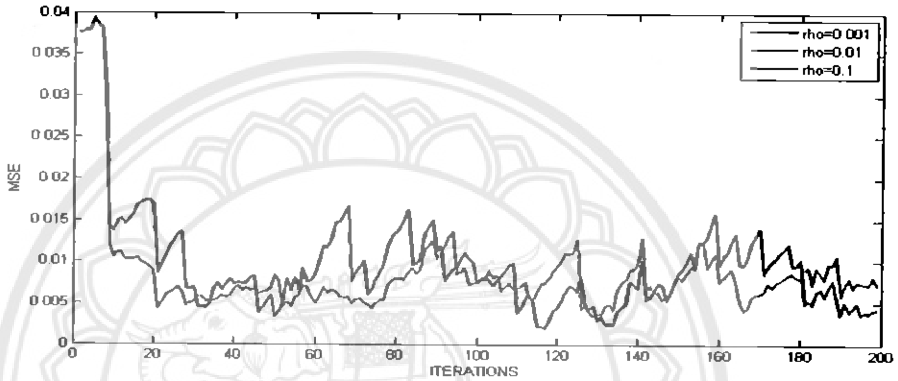
รูปที่ 4.11 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 4 แบบข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทันที

4.4.3 ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทีละน้อย

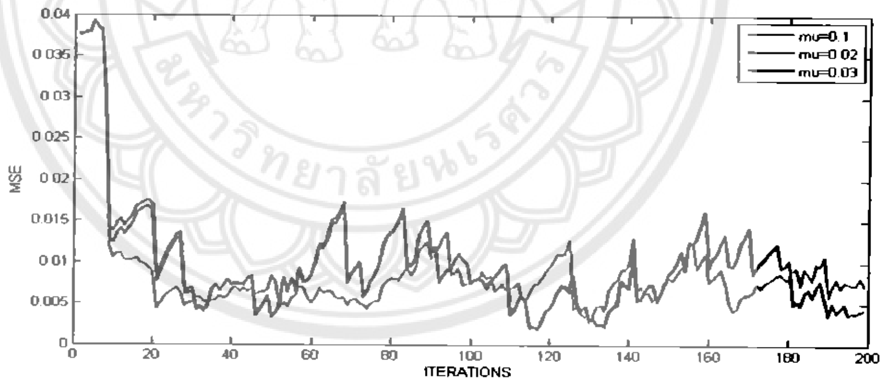
เราจะให้ค่าอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.4 ค่าเบต้าเท่ากับ 100 และเราจะทำการปรับค่าแลมดา ค่าเรกกูลาไรเซชัน และค่ามิว เพื่อให้ได้ค่าเฉลี่ยความผิดพลาดมีค่าลดลงมากที่สุด สามารถแสดงดังรูปที่ 4.12



(ก)



(ข)



(ค)

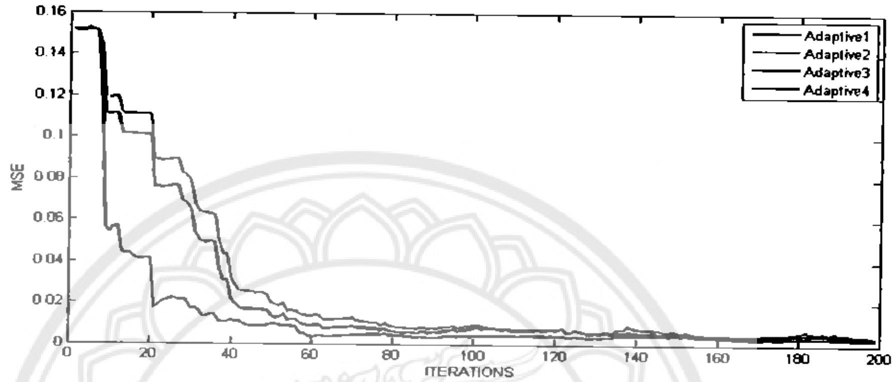
รูปที่ 4.12 ผลของค่าพารามิเตอร์ต่างๆ สำหรับการปรับค่าอัตราการเรียนรู้วิธีที่ 4 แบบข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทีละน้อย

จากรูปที่ 4.12 ค่าความผิดพลาดที่ได้จากการปรับค่าแลมดา ค่าเรกกูลาไรเซชัน และค่ามิว จะเห็นว่าค่าความผิดพลาดที่ $\lambda = 0.01$, $\rho = 0.01$ และ $\mu = 0.01$ มีการลดลงของค่าความผิดพลาดที่ดี โดยในช่วงแรกจะลดลงอย่างฉับพลันตั้งแต่รอบที่ 0–50 แล้วหลังจากนั้นจะค่อยๆ ลดลงจนค่าความผิดพลาดนั้นมีลักษณะคงที่ ค่าแลมดา ค่าเรกกูลาไรเซชัน และค่ามิว ค่าความผิดพลาดที่ได้จะมีค่ามากและการลดลงของค่าความผิดพลาดไม่ดี

4.5 เปรียบเทียบการปรับค่าอัตราการเรียนรู้ทั้ง 4 วิธี

4.5.1 ข้อมูลที่มีจุดศูนย์กลางคงที่

ค่าพารามิเตอร์ที่ดีที่สุดของชนิดข้อมูลจุดศูนย์กลางคงที่ของการปรับค่าอัตราการเรียนรู้ทั้ง 4 วิธี เมื่อนำมาเปรียบเทียบกันทั้ง 4 วิธี แสดงดังรูปที่ 4.13

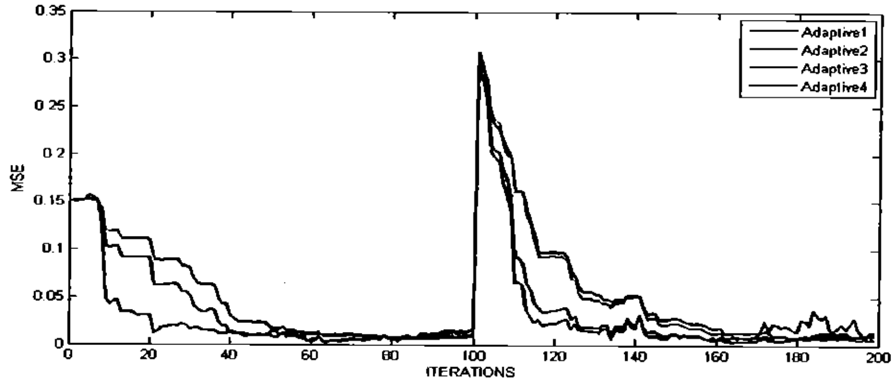


รูปที่ 4.13 กราฟแสดงการเปรียบเทียบค่าความผิดพลาดทั้ง 4 วิธี

จากรูปที่ 4.13 ค่าความผิดพลาดที่ลดลงมากที่สุดของแต่ละวิธี เมื่อนำมาค่าความผิดพลาดที่ดีที่สุดแต่ละวิธีมาเปรียบเทียบกัน จะเห็นว่าค่าความผิดพลาดของวิธีที่ 4 ดีที่สุด เนื่องจากมีค่าความผิดพลาดลดลงมากกว่าวิธีอื่น รองลงมา คือ ค่าความผิดพลาดของวิธีที่ 3

4.5.2 ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทันที

ค่าพารามิเตอร์ที่ดีที่สุดของชนิดข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีของการปรับค่าอัตราการเรียนรู้ทั้ง 4 วิธี เมื่อนำมาเปรียบเทียบกันทั้ง 4 วิธี แสดงดังรูปที่ 4.14

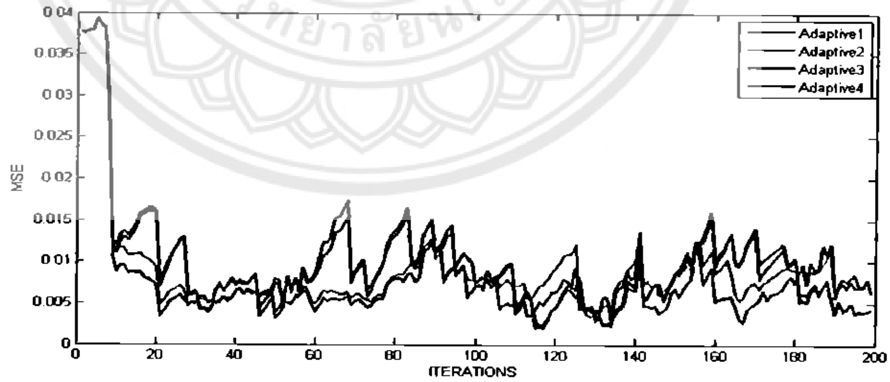


รูปที่ 4.14 กราฟแสดงการเปรียบเทียบค่าความผิดพลาดทั้ง 4 วิธี

จากรูปที่ 4.14 ค่าความผิดพลาดที่ลดลงมากที่สุดของแต่ละวิธี เมื่อนำมาค่าความผิดพลาดที่ดีที่สุดแต่ละวิธีมาเปรียบเทียบกัน จะเห็นว่าค่าความผิดพลาดของวิธีที่ 4 ดีที่สุด เนื่องจากมีค่าความผิดพลาดลดลงมากกว่าวิธีอื่น รองลงมา คือ ค่าความผิดพลาดของวิธีที่ 2

4.5.3 ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทีละน้อย

ค่าพารามิเตอร์ที่ดีที่สุดของชนิดข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยของการปรับค่าอัตราการเรียนรู้ทั้ง 4 วิธี เมื่อนำมาเปรียบเทียบกันทั้ง 4 วิธี แสดงดังรูปที่ 4.15



รูปที่ 4.15 กราฟแสดงการเปรียบเทียบค่าความผิดพลาดทั้ง 4 วิธี

จากรูปที่ 4.15 ค่าความผิดพลาดที่ลดลงมากที่สุดของแต่ละวิธี เมื่อนำมาค่าความผิดพลาดที่ดีที่สุดแต่ละวิธีมาเปรียบเทียบกัน จะเห็นว่าค่าความผิดพลาดของวิธีที่ 4 ดีที่สุด เนื่องจากมีค่าความผิดพลาดลดลงมากกว่าวิธีอื่น รองลงมา คือ ค่าความผิดพลาดของวิธีที่ 2

บทที่ 5
สรุปผลการทดลอง

ในบทนี้จะกล่าวถึงผลของการทดลองที่เกิดขึ้นในการหาฟังก์ชันการประมาณค่าด้วยวิธีเคอร์เนลที่มีการปรับค่าอัตราการเรียนรู้ โดยข้อมูลที่ใช้ทดสอบเป็นซิงค์ฟังก์ชันที่มีจุดศูนย์กลางคงที่ จุดศูนย์กลางเปลี่ยนค่าทันที และจุดศูนย์กลางเปลี่ยนค่าทีละนิด ค่าความผิดพลาดที่มีการเปลี่ยนแปลงเนื่องจากการปรับค่าพารามิเตอร์ต่างๆ

5.1 การหาค่าพารามิเตอร์ที่ดีที่สุดของการประมาณค่าฟังก์ชัน

จากการทดลองในบทที่ 4 จะมีการปรับค่าพารามิเตอร์ทั้งหมด คือ ค่าอัตราการเรียนรู้ ค่าเรกกูลาไรเซชัน ค่าเบต้า ค่ามิว ค่าแกมมา และค่าแลมดา เพื่อหาค่าความผิดพลาดที่ดีที่สุด ซึ่งผลการทดลองที่ได้จะเห็นว่าแต่ละชนิดของข้อมูล การปรับค่าอัตราการเรียนรู้แต่ละวิธีมีค่าพารามิเตอร์ที่แตกต่างกัน

1) การปรับค่าอัตราการเรียนรู้วิธีที่ 1

ตารางที่ 5.1 ค่าพารามิเตอร์ดีที่สุดแต่ละชนิดข้อมูลในการปรับค่าอัตราการเรียนรู้วิธีที่ 1

ชนิดของข้อมูลที่ใช้ทดสอบ	ค่าพารามิเตอร์ที่เหมาะสม			
	β	η	ρ	γ
ข้อมูลที่จุดศูนย์กลางคงที่	100	0.1	0.001	0.01
ข้อมูลที่จุดศูนย์กลางเปลี่ยนค่าทันที	100	0.1	0.01	0.025
ข้อมูลที่จุดศูนย์กลางเปลี่ยนค่าทีละน้อย	100	0.4	0.1	0.001

2) การปรับค่าอัตราการเรียนรู้วิธีที่ 2

ตารางที่ 5.2 ค่าพารามิเตอร์ดีที่สุดแต่ละชนิดข้อมูลในการปรับค่าอัตราการเรียนรู้วิธีที่ 2

ชนิดของข้อมูลที่ใช้ทดสอบ	ค่าพารามิเตอร์ที่เหมาะสม			
	β	η	ρ	γ
ข้อมูลที่จุดศูนย์กลางคงที่	100	0.1	0.001	0.1
ข้อมูลที่จุดศูนย์กลางเปลี่ยนค่าทันที	100	0.1	0.001	0.25
ข้อมูลที่จุดศูนย์กลางเปลี่ยนค่าทีละน้อย	100	0.4	0.001	0.001

3) การปรับค่าอัตราการเรียนรู้วิธีที่ 3

ตารางที่ 5.3 ค่าพารามิเตอร์ดีที่สุดแต่ละชนิดข้อมูลในการปรับค่าอัตราการเรียนรู้วิธีที่ 3

ชนิดของข้อมูลที่ใช้ทดสอบ	ค่าพารามิเตอร์ที่เหมาะสม				
	β	η	ρ	γ	μ
ข้อมูลที่จุดศูนย์กลางคงที่	100	0.1	0.001	0.01	0.01
ข้อมูลที่จุดศูนย์กลางเปลี่ยนค่าทันที	100	0.1	0.01	0.1	0.04
ข้อมูลที่จุดศูนย์กลางเปลี่ยนค่าทีละน้อย	100	0.4	0.1	0.01	0.1

4) การปรับค่าอัตราการเรียนรู้วิธีที่ 4

ตารางที่ 5.4 ค่าพารามิเตอร์ดีที่สุดแต่ละชนิดข้อมูลในการปรับค่าอัตราการเรียนรู้วิธีที่ 4

ชนิดของข้อมูลที่ใช้ทดสอบ	ค่าพารามิเตอร์ที่เหมาะสม				
	β	η	ρ	λ	μ
ข้อมูลที่จุดศูนย์กลางคงที่	100	0.1	0.001	0.01	0.1
ข้อมูลที่จุดศูนย์กลางเปลี่ยนค่าทันที	100	0.1	0.001	0.01	0.01
ข้อมูลที่จุดศูนย์กลางเปลี่ยนค่าทีละน้อย	100	0.4	0.01	0.01	0.01

5) **การปรับค่าอัตราการเรียนรู้ที่ดีที่สุดแต่ละชนิดข้อมูล**

ตารางที่ 5.5 ค่าพารามิเตอร์ดีที่สุดแต่ละชนิดข้อมูล

ชนิดของข้อมูลที่ใช้ทดสอบ	การปรับค่าอัตราการเรียนรู้ที่ดีที่สุด
ข้อมูลที่มีจุดศูนย์กลางคงที่	การปรับค่าอัตราการเรียนรู้วิธีที่ 4
ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทันที	การปรับค่าอัตราการเรียนรู้วิธีที่ 4
ข้อมูลที่มีจุดศูนย์กลางเปลี่ยนค่าทีละน้อย	การปรับค่าอัตราการเรียนรู้วิธีที่ 4

5.2 ข้อเสนอแนะ

ในการทดลองการเรียนรู้แบบปรับตัวเองด้วยวิธีเคอร์เนลแบบออนไลน์โดยใช้ส่วนต่อประสานกราฟิกกับผู้ใช้นั้น หากเราเลือกค่าพารามิเตอร์ที่เหมาะสม จะได้ค่าความผิดพลาดที่ดี และสามารถนำค่าพารามิเตอร์ที่เหมาะสมนี้ไปประยุกต์ใช้งานได้ต่อไป

เอกสารอ้างอิง

[1] N.Aronszajn, "Theory of reproducing kernels," Transactions of the American Mathematical society, vol. 68, pp. 337-404, 1950.

[2] สัญจร วุฒิสิทธิกุลกิจ และคณะ, "MATLAB การประยุกต์ใช้งานทางวิศวกรรมไฟฟ้า", พิมพ์ครั้งที่ 3, สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2551.

[3] T. J. Dodd, "Gradient descent approach to approximation in reproducing kernel Hilbertspaces," Department of Automatic Control and Systems Engineering, University of Sheffield, UK, Tech. Rep. 821, 2002.

[4] S. Phonphitakchai and T. J. Dodd, "Sparse learning and adaption in online kernel methods," In proceeding of the second Mahasarakham International Workshop on AI (MIWAI), R. Boothand and C. Sombattheera, Eds., 2008, PP.20-29.

[5] "Stochastic meta descent in online kernel methods," in sixth annual international conference organized by Electrical Engineering/Electronics, Computer, Telecommunication and Information Technology (ECTI) Association, 2009.

[6] นางสาววารี คำแสน, นายสุนทร ศิลปะวงศ์, นายเทวา ตาเปี้ยสืบ. (2552), การปรับค่าอัตราการเรียนรู้สำหรับการเรียนรู้แบบออนไลน์ด้วยวิธีเคอร์เนล, ปริญญานิพนธ์ วศ.บ., มหาวิทยาลัยนเรศวร, พิษณุโลก

ภาคผนวก

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางคงที่วิธีที่ 1

η	ρ	γ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง		
				50	100	200
0.1	0.001	0	-	0.0856	0.0498	0.0285
0.1	0.001	0.0025	-	0.0855	0.0495	0.0283
0.1	0.001	0.005	-	0.0853	0.0491	0.0281
0.1	0.001	0.0075	-	0.0852	0.0488	0.0279
0.1	0.001	0.01	-	0.085	0.0486	0.0277
0.1	0.001	0.025	-	0.0842	0.0472	0.0271
0.1	0.001	0.05	-	0.0830	0.0457	0.0268
0.1	0.001	0.075	-	0.0819	0.0448	0.027
0.1	0.001	0.1	-	0.081	0.0443	0.0274
0.1	0.001	0.25	-	0.0771	0.0444	0.0315
0.1	0.001	0.5	-	0.0730	0.0482	0.0397
0.1	0.001	0.75	-	0.0701	0.0528	0.0473
0.1	0.001	1	-	0.0677	0.0577	0.0533
0.1	0.01	0	-	0.0869	0.0502	0.0285
0.1	0.01	0.0025	-	0.0868	0.0498	0.0283
0.1	0.01	0.005	-	0.0866	0.0498	0.0281
0.1	0.01	0.0075	-	0.0865	0.0492	0.0279
0.1	0.01	0.01	-	0.0864	0.0489	0.0277
0.1	0.01	0.025	-	0.0856	0.0476	0.0271
0.1	0.01	0.05	-	0.0844	0.0461	0.0268
0.1	0.01	0.075	-	0.0833	0.0452	0.0270

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางคงที่วิธีที่ 1 (ต่อ)

η	ρ	γ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง		
				50	100	200
0.1	0.01	0.1	-	0.0824	0.0446	0.0274
0.1	0.01	0.25	-	0.0784	0.0447	0.0315
0.1	0.01	0.5	-	0.0741	0.0483	0.0397
0.1	0.01	0.75	-	0.0710	0.0528	0.0473
0.1	0.01	1	-	0.0684	0.0576	0.0533
0.1	0.1	0	-	0.0910	0.0592	0.0438
0.1	0.1	0.0025	-	0.0909	0.0589	0.0437
0.1	0.1	0.005	-	0.0909	0.0589	0.0437
0.1	0.1	0.0075	-	0.0909	0.0589	0.0437
0.1	0.1	0.01	-	0.0909	0.0589	0.0437
0.1	0.1	0.025	-	0.0898	0.0575	0.0441
0.1	0.1	0.05	-	0.0887	0.0566	0.0450
0.1	0.1	0.075	-	0.0887	0.0562	0.0461
0.1	0.1	0.1	-	0.0869	0.0560	0.0472
0.1	0.1	0.25	-	0.0832	0.0567	0.0530
0.1	0.1	0.5	-	0.0795	0.0602	0.0617
0.1	0.1	0.75	-	0.0770	0.0648	0.0694
0.1	0.1	1	-	0.0752	0.0702	0.0759

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางคงที่วิธีที่ 2

η	ρ	γ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง		
				50	100	200
0.1	0.001	0	-	0.0852	0.0490	0.0277
0.1	0.001	0.0025	-	0.0852	0.0490	0.0276
0.1	0.001	0.005	-	0.0852	0.0489	0.0276
0.1	0.001	0.0075	-	0.0852	0.0489	0.0276
0.1	0.001	0.01	-	0.0852	0.0489	0.0276
0.1	0.001	0.025	-	0.0851	0.0487	0.0274
0.1	0.001	0.05	-	0.0849	0.0483	0.0272
0.1	0.001	0.075	-	0.0848	0.0480	0.0271
0.1	0.001	0.1	-	0.0846	0.0477	0.0269
0.1	0.001	0.25	-	0.0838	0.0463	0.0265
0.1	0.001	0.5	-	0.0827	0.0450	0.0271
0.1	0.001	0.75	-	0.0818	0.0448	0.0281
0.1	0.001	1	-	0.0812	0.0457	0.0319
0.1	0.01	0	-	0.0869	0.0502	0.0285
0.1	0.01	0.0025	-	0.0869	0.0502	0.0285
0.1	0.01	0.005	-	0.0869	0.0502	0.0285
0.1	0.01	0.0075	-	0.0869	0.0502	0.0285
0.1	0.01	0.01	-	0.0869	0.0500	0.0284
0.1	0.01	0.025	-	0.0866	0.0498	0.0283
0.1	0.01	0.05	-	0.0866	0.0498	0.0283
0.1	0.01	0.075	-	0.0865	0.0492	0.0279

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางคงที่วิธีที่ 2 (ต่อ)

η	ρ	γ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง		
				50	100	200
0.1	0.01	0.1	-	0.0864	0.0489	0.0277
0.1	0.01	0.25	-	0.0856	0.0474	0.0272
0.1	0.01	0.5	-	0.0845	0.0460	0.0278
0.1	0.01	0.75	-	0.0837	0.0457	0.0297
0.1	0.01	1	-	0.0831	0.0464	0.0329
0.1	0.1	0	-	0.0910	0.0592	0.0438
0.1	0.1	0.0025	-	0.0910	0.0592	0.0438
0.1	0.1	0.005	-	0.0910	0.0592	0.0438
0.1	0.1	0.0075	-	0.0910	0.0592	0.0438
0.1	0.1	0.01	-	0.0910	0.0592	0.0438
0.1	0.1	0.025	-	0.0909	0.0589	0.0437
0.1	0.1	0.05	-	0.0909	0.0589	0.0437
0.1	0.1	0.075	-	0.0906	0.0592	0.0438
0.1	0.1	0.1	-	0.0906	0.0592	0.0438
0.1	0.1	0.25	-	0.0898	0.0575	0.0448
0.1	0.1	0.5	-	0.0889	0.0570	0.0481
0.1	0.1	0.75	-	0.0882	0.0573	0.0528
0.1	0.1	1	-	0.0878	0.0583	0.0599

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางคงที่วิธีที่ 3

ρ	γ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง		
			50	100	200
0.001	0.01	0.01	0.0980	0.0646	0.0236
0.001	0.01	0.02	0.0900	0.0582	0.0245
0.001	0.01	0.03	0.0887	0.0572	0.0250
0.001	0.01	0.04	0.0883	0.0570	0.0254
0.001	0.01	0.1	0.0889	0.0574	0.0265
0.001	0.01	0.2	0.0899	0.0581	0.0270
0.001	0.01	0.3	0.0904	0.0585	0.0273
0.001	0.01	0.4	0.0907	0.0587	0.0273
0.001	0.1	0.01	9.6190	5.4796	2.630
0.001	0.1	0.02	0.0902	0.0548	0.0275
0.001	0.1	0.03	0.0889	0.0601	0.0306
0.001	0.1	0.04	12.7391	9.802	0.0854
0.001	0.1	0.1	0.0940	0.0516	0.6592
0.001	0.1	0.2	0.0823	0.0530	0.0260
0.001	0.1	0.3	0.0859	0.0550	0.0272
0.001	0.1	0.4	0.0880	0.0562	0.0275
0.01	0.01	0.01	0.0755	0.0417	0.0236
0.01	0.01	0.02	0.0778	0.0433	0.0245
0.01	0.01	0.03	0.0793	0.0443	0.0250
0.01	0.01	0.04	0.0802	0.0450	0.0254
0.01	0.01	0.1	0.0827	0.0469	0.0265

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางคงที่วิธีที่ 3 (ต่อ)

ρ	γ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง		
			50	100	200
0.01	0.01	0.2	0.0840	0.0478	0.0270
0.01	0.01	0.3	0.0846	0.0483	0.0272
0.01	0.01	0.4	0.0848	0.0485	0.0273
0.01	0.1	0.01	0.0887	0.0483	9.619
0.01	0.1	0.02	0.0670	0.0372	0.0275
0.01	0.1	0.03	0.0646	0.0361	0.0306
0.01	0.1	0.04	0.0647	0.0360	12.739
0.01	0.1	0.1	0.0700	0.0383	0.659
0.01	0.1	0.2	0.0761	0.0414	0.0260
0.01	0.1	0.3	0.0799	0.0437	0.0272
0.01	0.1	0.4	0.0820	0.0450	0.0275
0. 1	0.01	0.01	0.0778	0.0431	0.0245
0. 1	0.01	0.02	0.0799	0.0446	0.0253
0. 1	0.01	0.03	0.0812	0.0455	0.0258
0. 1	0.01	0.04	0.0821	0.0462	0.0262
0. 1	0.01	0.1	0.0845	0.0480	0.0273
0. 1	0.01	0.2	0.0857	0.0490	0.0278
0. 1	0.01	0.3	0.0863	0.0494	0.0281
0. 1	0.01	0.4	0.0866	0.0496	0.0282
0. 1	0.1	0.01	0.0778	0.0677	1.4374e3
0. 1	0.1	0.02	0.0799	0.0397	2.4475e9

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางคงที่วิธีที่ 3 (ต่อ)

ρ	γ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง		
			50	100	200
0. 1	0.1	0.03	0.0812	0.0376	6.9767e10
0. 1	0.1	0.04	0.0821	0.0373	1.1928e11
0. 1	0.1	0.1	0.0845	0.0392	248.1849
0. 1	0.1	0.2	0.0857	0.0424	0.0259
0. 1	0.1	0.3	0.0863	0.0448	0.0278
0. 1	0.1	0.4	0.0866	0.0460	0.0283

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางคงที่วิธีที่ 4

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง		
				50	100	200
0.1	0.001	0.01	0.01	0.0803	0.0445	0.0251
0.1	0.001	0.01	0.02	0.0744	0.0406	0.0232
0.1	0.001	0.01	0.03	0.0691	0.0374	0.0216
0.1	0.001	0.01	0.04	0.0643	0.0346	0.0203
0.1	0.001	0.01	0.1	0.0442	0.0242	0.0139
0.1	0.001	0.01	0.2	0.0529	0.0306	0.0178
0.1	0.001	0.01	0.3	0.3886	NAN	NAN
0.1	0.001	0.01	0.4	NAN	NAN	NAN
0.1	0.001	0.1	0.01	0.0803	0.0445	0.0251
0.1	0.001	0.1	0.02	0.0744	0.0406	0.0232
0.1	0.001	0.1	0.03	0.0691	0.0374	0.0216
0.1	0.001	0.1	0.04	0.0643	0.0346	0.0203
0.1	0.001	0.1	0.1	0.0442	0.0242	0.0139
0.1	0.001	0.1	0.2	0.0529	0.0306	0.0178
0.1	0.001	0.1	0.3	0.4089	NAN	NAN
0.1	0.001	0.1	0.4	NAN	NAN	NAN
0.1	0.01	0.01	0.01	0.0807	0.0453	0.0258
0.1	0.01	0.01	0.02	0.075	0.0413	0.0238
0.1	0.01	0.01	0.03	0.0697	0.0381	0.0223
0.1	0.01	0.01	0.04	0.0649	0.0353	0.0209
0.1	0.01	0.01	0.1	0.0446	0.0246	0.0150

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางคงที่วิธีที่ 4 (ต่อ)

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง		
				50	100	200
0.1	0.01	0.01	0.2	0.0487	0.0276	0.0158
0.1	0.01	0.01	0.3	0.3266	NAN	NAN
0.1	0.01	0.01	0.4	NAN	NAN	NAN
0.1	0.01	0.1	0.01	0.0807	0.055	0.0423
0.1	0.01	0.1	0.02	0.0750	0.0413	0.0238
0.1	0.01	0.1	0.03	0.0697	0.0381	0.0223
0.1	0.01	0.1	0.04	0.0649	0.0353	0.0209
0.1	0.01	0.1	0.1	0.0446	0.0246	0.0150
0.1	0.01	0.1	0.2	0.0487	0.0277	0.0158
0.1	0.01	0.1	0.3	0.49	NAN	NAN
0.1	0.01	0.1	0.4	NAN	NAN	NAN
0.1	0.1	0.01	0.01	0.0855	0.055	0.0423
0.1	0.1	0.01	0.02	0.0803	0.0518	0.0418
0.1	0.1	0.01	0.03	0.0755	0.0493	0.0421
0.1	0.1	0.01	0.04	0.0712	0.0475	0.0429
0.1	0.1	0.01	0.2	0.0514	NAN	NAN
0.1	0.1	0.01	0.3	NAN	NAN	NAN
0.1	0.1	0.01	0.4	NAN	NAN	NAN
0.1	0.1	0.1	0.01	0.0855	0.055	0.0423
0.1	0.1	0.1	0.02	0.0803	0.0518	0.0418
0.1	0.1	0.1	0.03	0.0755	0.0493	0.042

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางคงที่วิธีที่ 4 (ต่อ)

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง		
				50	100	200
0.1	0.1	0.1	0.04	0.0712	0.0475	0.0428
0.1	0.1	0.1	0.1	0.0353	0.0557	0.0763
0.1	0.1	0.1	0.2	0.0485	0.3713	NAN
0.1	0.1	0.1	0.3	NAN	NAN	NAN
0.1	0.1	0.1	0.4	NAN	NAN	NAN

หมายเหตุ: NAN คือค่าที่ได้จากค่าพารามิเตอร์ที่ไม่เหมาะสม ทำให้ให้ไม่สามารถหาค่าได้

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีวิธีที่ 1

η	ρ	γ	ค่า MSE ต่อจำนวนรอบการทดลอง
			200
0.1	0.001	0	0.0691
0.1	0.001	0.001	0.0679
0.1	0.001	0.0025	0.0664
0.1	0.001	0.005	0.0641
0.1	0.001	0.0075	0.0621
0.1	0.001	0.01	0.0603
0.1	0.001	0.025	0.0531
0.1	0.001	0.05	0.0469
0.1	0.001	0.075	0.0435
0.1	0.001	0.1	0.0415
0.1	0.001	0.25	0.0390
0.1	0.001	0.5	0.0436
0.1	0.001	0.75	0.0499
0.1	0.001	1	0.0562
0.1	0.01	0	0.0675
0.1	0.01	0.001	0.0664
0.1	0.01	0.0025	0.0649
0.1	0.01	0.005	0.0626
0.1	0.01	0.0075	0.0607
0.1	0.01	0.01	0.0590
0.1	0.01	0.025	0.0519

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีวิธีที่ 1 (ต่อ)

η	ρ	γ	ค่า MSE ต่อจำนวนรอบการทดลอง
			200
0.1	0.01	0.05	0.0458
0.1	0.01	0.075	0.0425
0.1	0.01	0.1	0.0406
0.1	0.01	0.25	0.0381
0.1	0.01	0.5	0.0418
0.1	0.01	0.75	0.0472
0.1	0.01	1	0.0526
0.1	0.1	0	0.0638
0.1	0.1	0.001	0.0630
0.1	0.1	0.0025	0.0620
0.1	0.1	0.005	0.0604
0.1	0.1	0.0075	0.0591
0.1	0.1	0.01	0.0580
0.1	0.1	0.025	0.0535
0.1	0.1	0.05	0.0501
0.1	0.1	0.075	0.0486
0.1	0.1	0.1	0.0478
0.1	0.1	0.25	0.0482
0.1	0.1	0.5	0.0527
0.1	0.1	0.75	0.0577
0.1	0.1	1	0.0624

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีวิธีที่ 2

η	ρ	γ	ค่า MSE ต่อจำนวนรอบการทดลอง
			200
0.1	0.001	0	0.0691
0.1	0.001	0.001	0.0690
0.1	0.001	0.0025	0.0688
0.1	0.001	0.005	0.0685
0.1	0.001	0.0075	0.0682
0.1	0.001	0.01	0.0679
0.1	0.001	0.025	0.0663
0.1	0.001	0.05	0.0637
0.1	0.001	0.075	0.0614
0.1	0.001	0.1	0.0593
0.1	0.001	0.25	0.0502
0.1	0.001	0.5	0.0434
0.1	0.001	0.75	0.0415
0.1	0.001	1	0.0423
0.1	0.01	0	0.0675
0.1	0.01	0.001	0.0674
0.1	0.01	0.0025	0.0672
0.1	0.01	0.005	0.0669
0.1	0.01	0.0075	0.0666
0.1	0.01	0.01	0.0664
0.1	0.01	0.025	0.0648

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีวิธีที่ 2 (ต่อ)

η	ρ	γ	ค่า MSE ต่อจำนวนรอบการทดลอง
			200
0.1	0.01	0.05	0.0623
0.1	0.01	0.075	0.06
0.1	0.01	0.1	0.0580
0.1	0.01	0.25	0.0492
0.1	0.01	0.5	0.0426
0.1	0.01	0.75	0.0405
0.1	0.01	1	0.0409
0.1	0.1	0	0.0638
0.1	0.1	0.001	0.0637
0.1	0.1	0.0025	0.0636
0.1	0.1	0.005	0.0634
0.1	0.1	0.0075	0.0632
0.1	0.1	0.01	0.0630
0.1	0.1	0.025	0.0619
0.1	0.1	0.05	0.0602
0.1	0.1	0.075	0.0587
0.1	0.1	0.1	0.0574
0.1	0.1	0.25	0.0522
0.1	0.1	0.5	0.0492
0.1	0.1	0.75	0.0491
0.1	0.1	1	0.0507

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีวิธีที่ 3

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.1	0.001	0.01	0.01	0.057
0.1	0.001	0.01	0.02	0.0586
0.1	0.001	0.01	0.03	0.0598
0.1	0.001	0.01	0.04	0.0606
0.1	0.001	0.01	0.1	0.0629
0.1	0.001	0.01	0.2	0.0643
0.1	0.001	0.01	0.3	0.0649
0.1	0.001	0.01	0.4	0.0652
0.1	0.001	0.1	0.01	0.0515
0.1	0.001	0.1	0.02	0.0377
0.1	0.001	0.1	0.03	0.0359
0.1	0.001	0.1	0.04	0.0358
0.1	0.001	0.1	0.1	0.0383
0.1	0.001	0.1	0.2	0.0417
0.1	0.001	0.1	0.3	0.0444
0.1	0.001	0.1	0.4	0.0459
0.1	0.01	0.01	0.01	0.0558
0.1	0.01	0.01	0.02	0.0573
0.1	0.01	0.01	0.03	0.0584
0.1	0.01	0.01	0.04	0.0592
0.1	0.01	0.01	0.1	0.0615

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีวิธีที่ 3 (ต่อ)

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.1	0.01	0.01	0.2	0.0628
0.1	0.01	0.01	0.3	0.0634
0.1	0.01	0.01	0.4	0.0637
0.1	0.01	0.1	0.01	0.0645
0.1	0.01	0.1	0.02	0.0389
0.1	0.01	0.1	0.03	0.0367
0.1	0.01	0.1	0.04	0.0361
0.1	0.01	0.1	0.1	0.0377
0.1	0.01	0.1	0.2	0.0409
0.1	0.01	0.1	0.3	0.0434
0.1	0.01	0.1	0.4	0.0449
0.1	0.1	0.01	0.01	0.0687
0.1	0.1	0.01	0.02	0.0606
0.1	0.1	0.01	0.03	0.0593
0.1	0.1	0.01	0.04	0.0590
0.1	0.1	0.01	0.1	0.0596
0.1	0.1	0.01	0.2	0.0605
0.1	0.1	0.01	0.3	0.0610
0.1	0.1	0.01	0.4	0.0612
0.1	0.1	0.1	0.01	5.5645e3
0.1	0.1	0.1	0.02	1.534e49

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีวิธีที่ 3 (ต่อ)

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.1	0.1	0.1	0.03	2.1304e13
0.1	0.1	0.1	0.04	0.0724
0.1	0.1	0.1	0.1	0.0473
0.1	0.1	0.1	0.2	0.0478
0.1	0.1	0.1	0.3	0.0492
0.1	0.1	0.1	0.4	0.0501

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีวิธีที่ 4

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.4	0.001	0.01	0.01	0.0306
0.4	0.001	0.01	0.02	0.0337
0.4	0.001	0.01	0.03	0.0449
0.4	0.001	0.01	0.04	0.0626
0.4	0.001	0.01	0.1	0.1607
0.4	0.001	0.01	0.2	NAN
0.4	0.001	0.01	0.3	NAN
0.4	0.001	0.01	0.4	NAN
0.4	0.001	0.1	0.01	0.0306
0.4	0.001	0.1	0.02	0.0337
0.4	0.001	0.1	0.03	0.0449
0.4	0.001	0.1	0.04	0.0626
0.4	0.001	0.1	0.1	0.1608
0.4	0.001	0.1	0.2	NAN
0.4	0.001	0.1	0.3	NAN
0.4	0.001	0.1	0.4	NAN
0.4	0.01	0.01	0.01	0.0289
0.4	0.01	0.01	0.02	0.0298
0.4	0.01	0.01	0.03	0.0354
0.4	0.01	0.01	0.04	0.0423
0.4	0.01	0.01	0.1	0.1255

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีวิธีที่ 4 (ต่อ)

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.4	0.01	0.01	0.2	NAN
0.4	0.01	0.01	0.3	NAN
0.4	0.01	0.01	0.4	NAN
0.4	0.01	0.1	0.01	0.0289
0.4	0.01	0.1	0.02	0.0289
0.4	0.01	0.1	0.03	0.0355
0.4	0.01	0.1	0.04	0.0424
0.4	0.01	0.1	0.1	0.1267
0.4	0.01	0.1	0.2	NAN
0.4	0.01	0.1	0.3	NAN
0.4	0.01	0.1	0.4	NAN
0.4	0.1	0.01	0.01	0.0389
0.4	0.1	0.01	0.02	0.0404
0.4	0.1	0.01	0.03	0.0439
0.4	0.1	0.01	0.04	0.0505
0.4	0.1	0.01	0.1	NAN
0.4	0.1	0.01	0.2	NAN
0.4	0.1	0.01	0.3	NAN
0.4	0.1	0.01	0.4	NAN
0.4	0.1	0.1	0.01	0.0388
0.4	0.1	0.1	0.02	0.0402

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทันทีวิธีที่ 4 (ต่อ)

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.4	0.1	0.1	0.03	0.0432
0.4	0.1	0.1	0.04	0.0483
0.4	0.1	0.1	0.1	NAN
0.4	0.1	0.1	0.2	NAN
0.4	0.1	0.1	0.3	NAN
0.4	0.1	0.1	0.4	NAN

หมายเหตุ: NAN คือค่าที่ได้จากค่าพารามิเตอร์ที่ไม่เหมาะสม ทำให้ให้ไม่สามารถหาค่าได้

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยวิธีที่ 1

η	ρ	γ	ค่า MSE ต่อจำนวนรอบการทดลอง
			200
0.1	0.001	0	0.0089
0.1	0.001	0.001	0.0089
0.1	0.001	0.0025	0.0089
0.1	0.001	0.005	0.0089
0.1	0.001	0.0075	0.0089
0.1	0.001	0.01	0.0088
0.1	0.001	0.025	0.0087
0.1	0.001	0.05	0.0086
0.1	0.001	0.075	0.0086
0.1	0.001	0.1	0.0086
0.1	0.001	0.25	0.0091
0.1	0.001	0.5	0.0109
0.1	0.001	0.75	0.0130
0.1	0.001	1	0.0151
0.1	0.01	0	0.0083
0.1	0.01	0.001	0.0083
0.1	0.01	0.0025	0.0083
0.1	0.01	0.005	0.0083
0.1	0.01	0.0075	0.0083
0.1	0.01	0.01	0.0082
0.1	0.01	0.025	0.0082

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยวิธีที่ 1 (ต่อ)

η	ρ	γ	ค่า MSE ต่อจำนวนรอบการทดลอง
			200
0.1	0.01	0.05	0.0080
0.1	0.01	0.075	0.0080
0.1	0.01	0.1	0.0080
0.1	0.01	0.25	0.0084
0.1	0.01	0.5	0.0098
0.1	0.01	0.75	0.0114
0.1	0.01	1	0.0129
0.1	0.1	0	0.0096
0.1	0.1	0.001	0.0096
0.1	0.1	0.0025	0.0096
0.1	0.1	0.005	0.0096
0.1	0.1	0.0075	0.0096
0.1	0.1	0.01	0.0096
0.1	0.1	0.025	0.0096
0.1	0.1	0.05	0.0096
0.1	0.1	0.075	0.0096
0.1	0.1	0.1	0.0097
0.1	0.1	0.25	0.0103
0.1	0.1	0.5	0.0115
0.1	0.1	0.75	0.0126
0.1	0.1	1	0.0133

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยวิธีที่ 2

η	ρ	γ	ค่า MSE ต่อจำนวนรอบการทดลอง
			200
0.1	0.001	0	0.0089
0.1	0.001	0.001	0.0096
0.1	0.001	0.0025	0.0096
0.1	0.001	0.005	0.0096
0.1	0.001	0.0075	0.0096
0.1	0.001	0.01	0.0096
0.1	0.001	0.025	0.0088
0.1	0.001	0.05	0.0088
0.1	0.001	0.075	0.0080
0.1	0.001	0.1	0.0080
0.1	0.001	0.25	0.0080
0.1	0.001	0.5	0.0086
0.1	0.001	0.75	0.0087
0.1	0.001	1	0.0118
0.1	0.01	0	0.0083
0.1	0.01	0.001	0.0083
0.1	0.01	0.0025	0.0083
0.1	0.01	0.005	0.0083
0.1	0.01	0.0075	0.0083
0.1	0.01	0.01	0.0083
0.1	0.01	0.025	0.0082

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยวิธีที่ 2 (ต่อ)

η	ρ	γ	ค่า MSE ต่อจำนวนรอบการทดลอง
			200
0.1	0.01	0.05	0.0082
0.1	0.01	0.075	0.0081
0.1	0.01	0.1	0.0081
0.1	0.01	0.25	0.0080
0.1	0.01	0.5	0.0084
0.1	0.01	0.75	0.0093
0.1	0.01	1	0.0105
0.1	0.1	0	0.0096
0.1	0.1	0.001	0.0096
0.1	0.1	0.0025	0.0096
0.1	0.1	0.005	0.0096
0.1	0.1	0.0075	0.0096
0.1	0.1	0.01	0.0096
0.1	0.1	0.025	0.0096
0.1	0.1	0.05	0.0096
0.1	0.1	0.075	0.0096
0.1	0.1	0.1	0.0096
0.1	0.1	0.25	0.0098
0.1	0.1	0.5	0.0104
0.1	0.1	0.75	0.0113
0.1	0.1	1	0.0120

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยวิธีที่ 3

η	ρ	γ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.1	0.001	0.01	0.01	0.0094
0.1	0.001	0.01	0.02	0.0087
0.1	0.001	0.01	0.03	0.0085
0.1	0.001	0.01	0.04	0.0085
0.1	0.001	0.01	0.1	0.0085
0.1	0.001	0.01	0.2	0.0081
0.1	0.001	0.01	0.3	0.0080
0.1	0.001	0.01	0.4	0.0080
0.1	0.001	0.1	0.01	1.8226e140
0.1	0.001	0.1	0.02	2.1516e87
0.1	0.001	0.1	0.03	2.7965e67
0.1	0.001	0.1	0.04	5.6594e55
0.1	0.001	0.1	0.1	4.8837
0.1	0.001	0.1	0.2	0.0245
0.1	0.001	0.1	0.3	0.0170
0.1	0.001	0.1	0.4	0.0098
0.1	0.01	0.01	0.01	0.0306
0.1	0.01	0.01	0.02	0.0093
0.1	0.01	0.01	0.03	0.0085
0.1	0.01	0.01	0.04	0.0083
0.1	0.01	0.01	0.1	0.0081

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยวิธีที่ 3 (ต่อ)

η	ρ	γ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.1	0.01	0.01	0.2	0.0081
0.1	0.01	0.01	0.3	0.0081
0.1	0.01	0.01	0.4	0.0082
0.1	0.01	0.1	0.01	6.2133e59
0.1	0.01	0.1	0.02	4.5928e50
0.1	0.01	0.1	0.03	5.5988e10
0.1	0.01	0.1	0.04	1.2479e6
0.1	0.01	0.1	0.1	0.0399
0.1	0.01	0.1	0.2	0.0125
0.1	0.01	0.1	0.3	0.0157
0.1	0.01	0.1	0.4	0.0338
0.1	0.1	0.01	0.01	0.3621
0.1	0.1	0.01	0.02	0.0106
0.1	0.1	0.01	0.03	0.0092
0.1	0.1	0.01	0.04	0.0089
0.1	0.1	0.01	0.1	0.0087
0.1	0.1	0.01	0.2	0.0086
0.1	0.1	0.01	0.3	0.0087
0.1	0.1	0.01	0.4	0.0087
0.1	0.1	0.1	0.01	0.2475
0.1	0.1	0.1	0.02	5.6094e3

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยวิธีที่ 3 (ต่อ)

η	ρ	γ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.1	0.1	0.1	0.03	8.6612e5
0.1	0.1	0.1	0.04	1.0501
0.1	0.1	0.1	0.1	0.0118
0.1	0.1	0.1	0.2	0.0120
0.1	0.1	0.1	0.3	21.5565
0.1	0.1	0.1	0.4	2.0491e6

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยวิธีที่ 4

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.4	0.001	0.01	0.01	0.0079
0.4	0.001	0.01	0.02	0.0086
0.4	0.001	0.01	0.03	0.0098
0.4	0.001	0.01	0.04	0.0106
0.4	0.001	0.01	0.1	NAN
0.4	0.001	0.01	0.2	NAN
0.4	0.001	0.01	0.3	NAN
0.4	0.001	0.01	0.4	NAN
0.4	0.001	0.1	0.01	0.0079
0.4	0.001	0.1	0.02	0.0086
0.4	0.001	0.1	0.03	0.00.98
0.4	0.001	0.1	0.04	0.0106
0.4	0.001	0.1	0.1	NAN
0.4	0.001	0.1	0.2	NAN
0.4	0.001	0.1	0.3	NAN
0.4	0.001	0.1	0.4	NAN
0.4	0.01	0.01	0.01	0.0074
0.4	0.01	0.01	0.02	0.0083
0.4	0.01	0.01	0.03	0.0107
0.4	0.01	0.01	0.04	0.1301
0.4	0.01	0.01	0.1	NAN

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยวิธีที่ 4 (ต่อ)

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.4	0.01	0.01	0.2	NAN
0.4	0.01	0.01	0.3	NAN
0.4	0.01	0.01	0.4	NAN
0.4	0.01	0.1	0.01	0.0074
0.4	0.01	0.1	0.02	0.0083
0.4	0.01	0.1	0.03	0.0107
0.4	0.01	0.1	0.04	0.0167
0.4	0.01	0.1	0.1	NAN
0.4	0.01	0.1	0.2	NAN
0.4	0.01	0.1	0.3	NAN
0.4	0.01	0.1	0.4	NAN
0.4	0.1	0.01	0.01	0.0096
0.4	0.1	0.01	0.02	0.0116
0.4	0.1	0.01	0.03	0.0184
0.4	0.1	0.01	0.04	NAN
0.4	0.1	0.01	0.1	NAN
0.4	0.1	0.01	0.2	NAN
0.4	0.1	0.01	0.3	NAN
0.4	0.1	0.01	0.4	NAN
0.4	0.1	0.1	0.01	0.0096
0.4	0.1	0.1	0.02	0.0114

การปรับค่าอัตราการเรียนรู้ของข้อมูลจุดศูนย์กลางเปลี่ยนค่าทีละน้อยวิธีที่ 4 (ต่อ)

η	ρ	λ	μ	ค่า MSE ต่อจำนวนรอบการทดลอง
				200
0.4	0.1	0.1	0.03	0.0170
0.4	0.1	0.1	0.04	0.1298
0.4	0.1	0.1	0.1	NAN
0.4	0.1	0.1	0.2	NAN
0.4	0.1	0.1	0.3	NAN
0.4	0.1	0.1	0.4	NAN

หมายเหตุ: NAN คือค่าที่ได้จากค่าพารามิเตอร์ที่ไม่เหมาะสม ทำให้ให้ไม่สามารถหาค่าได้

ประวัติผู้ดำเนินโครงงาน

ชื่อ นางสาวสุวภา ทองพงษ์เนียม

ภูมิลำเนา 460/2 หมู่ 4 ต.บ้านน้อยซุ้มขี้เหล็ก อ.เนินมะปราง จ.พิษณุโลก

ประวัติการศึกษา

– จบระดับมัธยมศึกษาจากโรงเรียนเนินมะปรางศึกษาวิทยา จ.พิษณุโลก

– ปัจจุบันกำลังศึกษาระดับปริญญาตรี ชั้นปีที่ 4 สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

E-mail: ee_zazaa@hotmail.co.th

ชื่อ นางสาวภูษณิศา อุภาดี

ภูมิลำเนา 815 หมู่ 10 ต.แม่สาย อ.แม่สาย จ.เชียงราย

ประวัติการศึกษา

– จบระดับมัธยมศึกษาจากโรงเรียนแม่สายประสิทธิ์ศาสตร์ จ.เชียงราย

– ปัจจุบันกำลังศึกษาระดับปริญญาตรี ชั้นปีที่ 4 สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

E-mail: kik-minnie@hotmail.com