

การเปรียบเทียบการประมาณค่าพารามิเตอร์ของการวิเคราะห์การถดถอย
ที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง

A COMPARISON OF PARAMETER ESTIMATION WITH PENALIZED
REGRESSION ANALYSIS UNDER HIGH-DIMENSIONAL DATA



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาคตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติและการวิเคราะห์ธุรกิจ
ภาควิชาสถิติ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2562

KMITL-2019-SC-M-050-038

การเปรียบเทียบการประมาณค่าพารามิเตอร์ของการวิเคราะห์การถดถอย
ที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง

A COMPARISON OF PARAMETER ESTIMATION WITH PENALIZED
REGRESSION ANALYSIS UNDER HIGH-DIMENSIONAL DATA



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติและการวิเคราะห์ธุรกิจ

ภาควิชาสถิติ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2562

KMITL-2019-SC-M-050-038

A COMPARISON OF PARAMETER ESTIMATION WITH PENALIZED
REGRESSION ANALYSIS UNDER HIGH-DIMENSIONAL DATA



A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR
THE DEGREE OF MASTER OF SCIENCE IN STATISTICS AND BUSINESS ANALYTICS

DEPARTMENT OF STATISTICS FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
2019

KMITL-2019-SC-M-050-038



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

COPYRIGHT 2019

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

หัวข้อวิทยานิพนธ์	การเปรียบเทียบการประมาณค่าพารามิเตอร์ของการวิเคราะห์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง
ชื่อนักศึกษา	นางสาวเบญจมาศ รุ่งศรานนท์
รหัสประจำตัว	60605079
ปริญญา	วิทยาศาสตรมหาบัณฑิต (สถิติและการวิเคราะห์ธุรกิจ)
ภาควิชา	สถิติ
พ.ศ.	2562
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผู้ช่วยศาสตราจารย์ ดร.อัชฌา อระวีพร

บทคัดย่อ

งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันลงโทษ 5 วิธี ได้แก่ วิธีการถดถอยริดจ์ วิธีการถดถอยลาสโซ วิธีการถดถอยอิลาสติกเน็ต วิธีการถดถอยลาสโซแบบปรับปรุง และวิธีการถดถอยอิลาสติกเน็ตแบบปรับปรุง ของตัวแบบการถดถอยเชิงเส้นพหุคูณ ซึ่งตัวแบบนี้ประกอบไปด้วยตัวแปรตาม และตัวแปรอิสระ ในกรณีที่จำนวนตัวแปรอิสระมีมากกว่าขนาดตัวอย่าง หรือที่เรียกว่าข้อมูลที่มีมิติสูง ในการเปรียบเทียบประสิทธิภาพทั้ง 5 วิธี ใช้เกณฑ์ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย ข้อมูลที่ใช้ในการศึกษาครั้งนี้เป็นการจำลองข้อมูล โดยกำหนดขนาดตัวอย่าง $n = 5, 10$ และ 15 จำนวนตัวแปรอิสระ เท่ากับ 16 ตัวแปร ขนาดตัวอย่าง $n = 10, 15$ และ 20 จำนวนตัวแปรอิสระ เท่ากับ 30 ตัวแปร ขนาดตัวอย่าง $n = 20, 30$ และ 40 จำนวนตัวแปรอิสระ เท่ากับ 50 ตัวแปร และขนาดตัวอย่าง $n = 60, 70$ และ 80 จำนวนตัวแปรอิสระ เท่ากับ 100 ตัวแปร โดยที่ตัวแปรอิสระสร้างมาจากการแจกแจงปกติ และค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงเส้นพหุคูณสร้างมาจากการแจกแจงปกติ การแจกแจงปกติปลอมปน และการแจกแจงไวบูล กำหนดค่าสัมประสิทธิ์การถดถอยพหุคูณ 3 กรณี ได้แก่ กรณีที่มีค่าเป็น 1 ทุกตัว กรณีที่มีค่าเป็น 0.5 และ 1 และกรณีที่มีค่าเป็น -1 และ 1 โดยข้อมูลจากการจำลองใช้เทคนิคมอนติคาร์โล ซึ่งแต่ละกรณีถูกทำซ้ำ 1,000 ครั้ง ผลการวิจัยพบว่า วิธีการถดถอยอิลาสติกเน็ตแบบปรับปรุงให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ยที่ต่ำที่สุดในทุกกรณี นอกจากนี้ผู้วิจัยยังนำทั้ง 5 วิธีมาประยุกต์ใช้กับข้อมูลจริงที่ขนาดตัวอย่าง $n = 5, 10$ และ 15 จำนวนตัวแปรอิสระ เท่ากับ 16 ตัวแปร ซึ่งวิธีการถดถอยอิลาสติกเน็ตแบบปรับปรุงเป็นวิธีการที่ดีกว่าวิธีอื่นๆ เช่นเดียวกับข้อมูลจำลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆก็ตาม อื่นๆห้ามมิให้คัดลอกเนื้อหา และต้องอ้างถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำสำคัญ : วิธีการถดถอยริดจ์ วิธีการถดถอยลาสโซ วิธีการถดถอยอิลาสติกเน็ต วิธีการถดถอยลาสโซแบบปรับปรุง วิธีการถดถอยอิลาสติกเน็ตแบบปรับปรุง

Thesis Title	A Comparison of Parameter Estimation with Penalized Regression Analysis Under High-Dimensional Data
Student Name	Benjamas Rungsaranon
Student ID	60605079
Degree	Master of Science (Statistics and Business Analytics)
Department	Statistics
Year	2019
Thesis Advisor	Assistant Professor Dr. Autcha Araveeporn

Abstract

The objective of this research is to compare the efficient of coefficient parameter estimation by using penalized regression analysis on five methods, such as the ridge regression, the least absolute shrinkage and selection operator regression, the elastic net regression, the adaptive least absolute shrinkage and selection operator regression, and the adaptive elastic net regression methods. This research uses the multiple linear regression model which is consisted of dependent variable and independent variables. In case the number of independent variables are larger than the number of sample sizes called high-dimensional data. For comparison the efficient of five methods, the criterion is based on the average mean square errors. The data of this research is simulated by the sample sizes $n = 5, 10,$ and 15 when the number of independent variables are specified by 16 . The sample sizes $n = 10, 15,$ and 20 , the number of independent variables are specified by 30 . The sample sizes $n = 20, 30,$ and 40 , the number of independent variables are specified by 50 . Finally the sample sizes $n = 60, 70,$ and 80 , the number of independent variables are defined 100 . The independent variables are generated from normal distribution and the residuals are generated from normal distribution, contaminated normal distribution, and weibull distribution. The coefficients of regression model are defined in 3 cases, such as 1 for all situations, 0.5 and 1 for half of situations, and 1 and -1 for half of situations. The data are obtained through simulation using a Monte Carlo technique with 1,000 replications for each case. The results are found that the adaptive elastic net regression is the minimum average mean square errors in all

cases. Furthermore, we apply five methods for real data based on the sample sizes $n = 5, 10$ and 15 when the number of independent variables are considered on 16 . The results of real data show that the adaptive elastic net regression outperforms the other methods as the simulation data.

Keywords : ridge regression, least absolute shrinkage and selection operator regression, elastic net regression, the adaptive least absolute shrinkage and selection operator regression, the adaptive elastic net regression



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จลุล่วงไปได้ด้วยดีและมีความถูกต้องในเนื้อหา เนื่องด้วยได้รับความอนุเคราะห์จาก ผศ.ดร.อชมา อระวีพร ซึ่งเป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์ เป็นผู้ซึ่งให้คำแนะนำ คำปรึกษา ช่วยเหลือ เอื้อเพื่อเอกสารต่าง ๆ และหนังสืออ้างอิง ที่ใช้ในการวิเคราะห์ข้อมูลและตรวจสอบแก้ไขข้อบกพร่องต่าง ๆ ตลอดจนติดตามผลงานทุกขั้นตอนของการดำเนินงานในการทำวิทยานิพนธ์ฉบับนี้จนกระทั่งมีความถูกต้องสมบูรณ์ จึงขอกราบขอบพระคุณด้วยความเคารพเป็นอย่างสูงไว้ ณ ที่นี้ด้วย

ขอขอบพระคุณ คณะกรรมการพิจารณาสอบป้องกันวิทยานิพนธ์ ได้แก่ ผศ.ดร.สมศรี บัณฑิตวิไล ผู้ซึ่งเป็นกรรมการ และ รศ.ดร.วราฤทธิ์ พานิชกิจโกศลกุล ผู้ซึ่งเป็นกรรมการ (ผู้ทรงคุณวุฒิจากภายนอก) ที่กรุณาให้คำปรึกษา คำแนะนำ ชี้ให้เห็นถึงข้อบกพร่อง ตลอดจนแก้ไขข้อผิดพลาดเพิ่มเติม ทำให้วิทยานิพนธ์เล่มนี้สมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่มอบทุนการศึกษาให้ผู้จัดทำวิทยานิพนธ์ ได้ศึกษาต่อในระดับวิทยาศาสตรมหาบัณฑิตเพื่อเพิ่มพูนความรู้ และนำความรู้ที่ได้มาดำเนินการจัดทำวิทยานิพนธ์เล่มนี้ได้สำเร็จไปได้ด้วยดี

ขอขอบพระคุณ คณาจารย์และบุคลากรภาควิชาสถิติทุกท่าน ที่ได้ประสิทธิ์ประสาทวิชาความรู้ พร้อมทั้งให้คำแนะนำ รวมถึงความช่วยเหลือในเรื่องต่าง ๆ มาโดยตลอด

สุดท้ายนี้ ขอขอบพระคุณบิดามารดา และครอบครัวของข้าพเจ้า ที่ให้การสนับสนุนและเป็นกำลังใจให้เสมอมา และขอขอบคุณเพื่อน ๆ ทุกคนที่ให้คำปรึกษา ช่วยเหลือในการทำงานมาโดยตลอด จนวิทยานิพนธ์เล่มนี้สำเร็จไปได้ด้วยดี

นางสาว เบญจมาศ รุ่งศรานนท์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ง
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์	4
1.3 สมมติฐานของการวิจัย	4
1.4 ขอบเขตของการวิจัย	4
1.5 เกณฑ์ที่ใช้ในการวิจัย	6
1.6 ประโยชน์ที่คาดว่าจะได้รับ	7
1.7 นิยามศัพท์เฉพาะที่ใช้ในงานวิจัย	7
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	8
2.1 การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ	8
2.2 วิธีที่ใช้ในการวิจัย	11
2.2.1 วิธีถดถอยรีดจ์	11
2.2.2 วิธีถดถอยลาสโซ่	12
2.2.3 วิธีถดถอยอิลาสติกเน็ต	13
2.2.4 วิธีถดถอยลาสโซ่แบบปรับปรุง	14
2.2.5 วิธีถดถอยอิลาสติกเน็ตแบบปรับปรุง	14
2.3 วิธีเลือกพารามิเตอร์การปรับ	15
2.3.1 วิธีตรวจสอบไขว้	15
2.3.2 วิธีเกณฑ์ข้อสนเทศของเบส์	16
2.4 การแจกแจงที่ใช้ในการวิจัย	17
2.4.1 การแจกแจงปกติ	17
2.4.2 การแจกแจงปกติปลอมปน	20
2.4.3 การแจกแจงไวบูล	21
2.5 งานวิจัยที่เกี่ยวข้อง	23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งยังมีใบคุ้มครองเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
บทที่ 3 วิธีการดำเนินงานวิจัย	25
3.1 การวางแผนการวิจัย	25
3.2 วิธีการดำเนินการวิจัย	30
3.3 ขั้นตอนของโปรแกรมที่ใช้ในงานวิจัย	30
บทที่ 4 ผลการวิจัยและการอภิปรายผล	32
4.1 ผลจากการจำลองข้อมูล	32
4.1.1 เมื่อค่าความคลาดเคลื่อนมีการแจกแจงปกติ	33
4.1.2 เมื่อค่าความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน	46
4.1.3 เมื่อค่าความคลาดเคลื่อนมีการแจกแจงไวบูล	58
4.2 ผลจากการประยุกต์ใช้กับข้อมูลจริง	71
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	75
5.1 สรุปผลการวิจัย	76
5.2 ข้อเสนอแนะ	76
บรรณานุกรม	77
ภาคผนวก	79
ภาคผนวก ก	80
ภาคผนวก ข	105
ประวัติผู้เขียน	122

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
3.1 แสดงขนาดตัวอย่าง และจำนวนสัมประสิทธิ์การถดถอย	29
3.2 แสดงจำนวนตัวแปรอิสระ และค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	29
3.3 แสดงจำนวนตัวแปรอิสระ และค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	29
4.1 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	33
4.2 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	35
4.3 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	36
4.4 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	37
4.5 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	38
4.6 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวกรณีค่าสัมประสิทธิ์	39

เอกสารนี้เป็นเอกสารถดถอยเป็น 0.5 และ 1 ใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.7 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	40
4.8 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	41
4.9 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	42
4.10 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	43
4.11 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	44
4.12 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	45
4.13 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยที่มีการนำไปใช้เป็น 1 ทั้งหมด	46

เอกสารนี้เป็นเอกสารของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 ไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆก็ตาม

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.14 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปรกติพลอมปนด้วยเปอร์เซ็นต์การพลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการพลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น 1 ทั้งหมด	47
4.15 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปรกติพลอมปนด้วยเปอร์เซ็นต์การพลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการพลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น 1 ทั้งหมด	48
4.16 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปรกติพลอมปนด้วยเปอร์เซ็นต์การพลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการพลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น 1 ทั้งหมด	49
4.17 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปรกติพลอมปนด้วยเปอร์เซ็นต์การพลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการพลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น 0.5 และ 1	50
4.18 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปรกติพลอมปนด้วยเปอร์เซ็นต์การพลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการพลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น 0.5 และ 1	51

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.19 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น 0.5 และ 1	52
4.20 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น 0.5 และ 1	53
4.21 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น -1 และ 1	54
4.22 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น -1 และ 1	55
4.23 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น -1 และ 1	56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.24 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นการปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอย เป็น -1 และ 1	57
4.25 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	58
4.26 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	60
4.27 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	61
4.28 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	62
4.29 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	63
4.30 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อน มีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	64

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของมหาวิทยาลัยราชภัฏวไลยอลงกรณ์
 ไม่ว่ากรณีใดๆทั้งนี้ การนำเอกสารนี้ไปใช้ประโยชน์ด้านการค้า
 ทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.31 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	65
4.32 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	66
4.33 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	67
4.34 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	68
4.35 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	69
4.36 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	70
4.37 รายละเอียดตัวแปรอิสระ และตัวแปรตามของชุดข้อมูลการสร้างอาคารที่อยู่อาศัยในเมืองเตหาราน ประเทศอิหร่าน ปี 2561	71

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.38 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยพหุคูณ ของชุดข้อมูลการสร้างอาคาร ที่อยู่อาศัย ในเมืองเตหาราน ประเทศอิหร่าน ปี 2561 ของวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net)	72
5.1 จำนวนตัวแปรอิสระ (p) และขนาดตัวอย่าง (n)	75
ก.1 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	81
ก.2 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	82
ก.3 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	83
ก.4 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	84
ก.5 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร	85

เอกสารนี้เป็นเอกสารกรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1 เท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
ก.6	จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	86
ก.7	จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	87
ก.8	จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	88
ก.9	จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีสัมประสิทธิ์การถดถอยเป็น -1 และ 1	89
ก.10	จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีสัมประสิทธิ์การถดถอยเป็น -1 และ 1	90
ก.11	จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีสัมประสิทธิ์การถดถอยเป็น -1 และ 1	91
ก.12	จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีสัมประสิทธิ์การถดถอยเป็น -1 และ 1	92

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
 ไม่ว่ากรณีใดๆทั้งนี้ การนำเอกสารนี้ไปใช้โดยไม่ได้รับอนุญาตจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
ก.13 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปโลมปนด้วยเปอร์เซ็นต์การปโลมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปโลมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	93
ก.14 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปโลมปนด้วยเปอร์เซ็นต์การปโลมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปโลมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	93
ก.15 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปโลมปนด้วยเปอร์เซ็นต์การปโลมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปโลมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	94
ก.16 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปโลมปนด้วยเปอร์เซ็นต์การปโลมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปโลมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	94
ก.17 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปโลมปนด้วยเปอร์เซ็นต์การปโลมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปโลมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	95

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
ก.18 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปโลมปนด้วยเปอร์เซ็นต์การปโลมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปโลมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	95
ก.19 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปโลมปนด้วยเปอร์เซ็นต์การปโลมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปโลมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	96
ก.20 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปโลมปนด้วยเปอร์เซ็นต์การปโลมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปโลมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	96
ก.21 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปโลมปนด้วยเปอร์เซ็นต์การปโลมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปโลมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	97
ก.22 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปโลมปนด้วยเปอร์เซ็นต์การปโลมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปโลมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	97

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
ก.23 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติพลอมปนด้วยเปอร์เซ็นต์การพลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการพลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	98
ก.24 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติพลอมปนด้วยเปอร์เซ็นต์การพลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการพลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	98
ก.25 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	99
ก.26 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	99
ก.27 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	100
ก.28 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด	100

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
ก.29 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	101
ก.30 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	101
ก.31 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	102
ก.32 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1	102
ก.33 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	103
ก.34 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	103
ก.35 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	104

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของมหาวิทยาลัยราชภัฏวชิรเวศน์บุรีรัมย์
 ไม่ว่ากรณีใดๆทั้งนี้ การนำเอกสารนี้ไปใช้ประโยชน์ด้านการค้า
 การทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
ก.36 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1	104



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
2.1 แสดงพื้นที่ของการแจกแจงปรกติเมื่อค่า x ห่างจากค่าเฉลี่ยเท่ากับ μ เป็นระยะหนึ่งเท่า สองเท่า และสามเท่าของค่าเบี่ยงเบนมาตรฐาน σ	18
2.2 การแจกแจงความหนาแน่นความน่าจะเป็นปรกติที่มีค่าพารามิเตอร์ (μ, σ^2) เป็น $(10,4)$ $(10,9)$ และ $(10,18)$	19
2.3 การแจกแจงความหนาแน่นความน่าจะเป็นปรกติที่มีค่าพารามิเตอร์ (μ, σ^2) เป็น $(5,9)$ $(10,9)$ และ $(15,9)$	19
2.4 การแจกแจงความหนาแน่นความน่าจะเป็นปรกติปลอมปนเป็น $0.8 * N(0,2) + 0.2 * N(0,10)$	21
2.5 การแจกแจงความหนาแน่นความน่าจะเป็นไวบูลที่มีค่าพารามิเตอร์ $W(\alpha, \beta)$ เป็น $W(1.2,2)$ $W(1.2,4)$ และ $W(1.2,6)$	22
3.1 การแจกแจงความหนาแน่นความน่าจะเป็นปรกติของ $N(\mu, \sigma^2)$ เป็น $N(0,2)$	25
3.2 การแจกแจงความหนาแน่นความน่าจะเป็นปรกติของ $N(\mu, \sigma^2)$ เป็น $N(0,1)$ $N(0,3)$ $N(0,5)$ และ $N(0,9)$	26
3.3 การแจกแจงความหนาแน่นความน่าจะเป็นปรกติปลอมปนเป็น $0.9 * N(0,1) + 0.1 * N(0,25)$	27
3.4 การแจกแจงความหนาแน่นความน่าจะเป็นปรกติปลอมปนเป็น $0.9 * N(0,1) + 0.1 * N(0,100)$	27
3.5 การแจกแจงความหนาแน่นความน่าจะเป็นไวบูลของ $W(\alpha, \beta)$ ดังนี้ $W(1,1.5)$ $W(1,3)$ และ $W(1,5)$	28
3.6 แผนผังแสดงลำดับวิธีการดำเนินการวิจัย	31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การวิเคราะห์การถดถอยเป็นเทคนิคทางสถิติสำหรับการตรวจสอบ และสร้างตัวแบบเพื่อหาความสัมพันธ์ระหว่างตัวแปร ประกอบด้วยตัวแปรตาม (Dependent Variable) และตัวแปรอิสระ (Independence Variable) ในการวิเคราะห์การถดถอยเชิงเส้นอาจแบ่งออกเป็น 2 ประเภท คือ การวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression) เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระเพียงหนึ่งตัว และการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression) เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระตั้งแต่ 2 ตัวแปรขึ้นไป และใช้สำหรับข้อมูลที่มีขนาดตัวอย่างมากกว่าจำนวนตัวแปรอิสระ เรียกว่า ข้อมูลที่มีมิติต่ำ (Low-Dimensional Data) โดยมีข้อกำหนดเบื้องต้นที่สำคัญในการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ คือ ค่าของตัวแปรอิสระแต่ละตัวเป็นค่าคงที่ และตัวแปรอิสระแต่ละตัวต้องไม่มีความสัมพันธ์กับตัวแปรอิสระอื่น ๆ สำหรับการหาค่าประมาณสัมประสิทธิ์การถดถอยในตัวแบบเชิงเส้นพหุคูณ ถ้าข้อมูลเป็นไปตามข้อกำหนดเบื้องต้นวิธีที่นิยมในการประมาณค่า คือ วิธีกำลังสองน้อยที่สุด (Ordinary Least Square Method : OLS) ซึ่งมีคุณสมบัติที่ดีของตัวประมาณ ได้แก่ เป็นตัวประมาณที่ไม่เอนเอียง (Unbiased) มีความแปรปรวนต่ำสุด (Minimum Variance) มีความสอดคล้อง (Consistency) มีความพอเพียง (Sufficiency) และมีประสิทธิภาพสูงสุด (Efficiency) (ประชุม, 2553) แต่ในปัจจุบันข้อมูลมีขนาดใหญ่ และมีความซับซ้อนมากขึ้น กล่าวคือ ข้อมูลเหล่านี้มีจำนวนตัวแปรอิสระมากกว่าขนาดตัวอย่าง เรียกว่า ข้อมูลที่มีมิติสูง (High-Dimensional Data) ทำให้ข้อมูลไม่เป็นไปตามข้อกำหนดเบื้องต้น ซึ่งวิธีกำลังสองน้อยที่สุดจะไม่เป็นวิธีการประมาณค่าที่ดีที่สุด ดังนั้นในงานวิจัยนี้ผู้วิจัยได้ศึกษาวิธีการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ (Penalized Regression) เพื่อใช้สำหรับวิเคราะห์ข้อมูลที่มีมิติสูงในการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ ซึ่งมีนักสถิติหลาย ๆ ท่านได้นำเสนอวิธีการถดถอยที่ปรับด้วยฟังก์ชันลงโทษ

Hoerl and Kennard (1970) ได้เสนอ วิธีการถดถอยริดจ์ (Ridge Regression) เป็นวิธีที่นิยมสำหรับการประมาณค่าสัมประสิทธิ์การถดถอย เพื่อแก้ปัญหาตัวแปรอิสระแต่ละตัวมีความสัมพันธ์กับตัวแปรอิสระอื่น ๆ เรียกว่า พหุสัมพันธ์ (Multicollinearity) โดยวิธีการถดถอยริดจ์คัดเลือกตัวแปรอิสระทุกตัวเข้าสู่ตัวแบบ ดังนั้นจึงเหมาะกับข้อมูลที่ตัวแปรอิสระทุกตัวมีความสัมพันธ์กับตัวแปรตาม หลักการ คือ การดึงค่าประมาณเข้าสู่ศูนย์ ทำให้ค่าสัมประสิทธิ์การถดถอยทุกตัวมีขนาดเล็กแต่ไม่เท่ากับศูนย์ จึงเหมาะสำหรับการวิเคราะห์ข้อมูลที่มีค่าสัมประสิทธิ์มีค่าน้อยและไม่เท่ากับศูนย์จำนวนมาก

Tibshirani (1996) ได้เสนอ วิธีลาสโซ (Least absolute shrinkage and selection operator : Lasso) วิธีนี้ไม่เพียงแต่ประมาณค่าสัมประสิทธิ์การถดถอย แต่ยังสามารถใช้คัดเลือกตัวแปรอิสระที่เกี่ยวข้องเข้าสู่ตัวแบบ แต่เมื่อเกิดปัญหาพหุสัมพันธ์ วิธีลาสโซจะเลือกตัวแปรอิสระเพียงตัวเดียวจากตัวแปรที่มีความสัมพันธ์กันในกลุ่มนั้น โดยไม่คำนึงว่าตัวแปรอิสระนั้นมีความสัมพันธ์กับตัวแปรตามมากที่สุดหรือไม่ และเนื่องจากวิธีลาสโซสามารถเลือกจำนวนตัวแปรอิสระมากที่สุดได้เท่ากับขนาดตัวอย่าง ซึ่งหากข้อมูลมีจำนวนตัวแปรอิสระมากกว่าขนาดตัวอย่างเป็นจำนวนมาก วิธีนี้อาจไม่เหมาะสม

Zou and Hastie (2005) ได้เสนอวิธีลาสติกเน็ต (Elastic Net) เป็นวิธีการวิเคราะห์การถดถอยที่สามารถประมาณค่าสัมประสิทธิ์การถดถอย และคัดเลือกตัวแปรอิสระที่เกี่ยวข้องเข้าสู่ตัวแบบ โดยวิธีลาสติกเน็ตพัฒนาขึ้นมาเพื่อแก้ไขข้อจำกัดของวิธีลาสโซ ซึ่งวิธีลาสติกเน็ตจะรวมระหว่างวิธีริดจ์และวิธีลาสโซเข้าด้วยกัน เหมาะสำหรับการวิเคราะห์ข้อมูลที่มีจำนวนตัวแปรอิสระมากกว่าขนาดตัวอย่างเป็นจำนวนมาก และเกิดปัญหาพหุสัมพันธ์

Zou (2006) ทำการศึกษาและพัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณที่สามารถแก้ปัญหาการคัดเลือกตัวแปรค่าที่เอนเอียงในตัวแบบวิธีลาสโซ ทำให้ตัวแบบมีความคงเส้นคงวา โดยมีการกำหนดค่าถ่วงน้ำหนักให้มีค่าสูงสำหรับค่าสัมประสิทธิ์ที่มีค่าน้อย และกำหนดค่าถ่วงน้ำหนักให้น้อยสำหรับค่าสัมประสิทธิ์ที่มีค่าสูง ซึ่งเรียกวิธีนี้ว่าวิธีลาสโซแบบปรับปรุง (The adaptive least absolute shrinkage and selection operator)

Zou and Zhang (2009) พัฒนาการประมาณค่าสัมประสิทธิ์การถดถอยในการเลือกหรือละเว้นตัวแปรอิสระเมื่อเกิดปัญหาพหุสัมพันธ์กันในกรณีที่ค่าสัมประสิทธิ์ใกล้เคียงกัน และเนื่องจากวิธีการถดถอยลาสติกเน็ตไม่มีคุณสมบัติที่ว่าเมื่อขนาดตัวอย่างมีจำนวนมากพอ จะมีความสามารถในการเลือกตัวแปรได้เสมือนกับว่าทราบตัวแบบที่แท้จริง (True Model) จึงเสนอวิธีการถดถอยลาสติกเน็ตแบบปรับปรุง (The adaptive Elastic Net) โดยรวมระหว่างวิธีถดถอยลาสติกเน็ต และวิธีลาสโซแบบปรับปรุงเข้าด้วยกัน

จากการศึกษางานวิจัยที่เกี่ยวข้อง โดยณัฐพร (2552) ได้เปรียบเทียบวิธีการคัดเลือกตัวแปรและการประมาณค่าสัมประสิทธิ์การถดถอย โดยเปรียบเทียบวิธีการถดถอยริดจ์ วิธีการถดถอยริดจ์แบบทางเลือก (Alternative Ridge Regression) และวิธีการถดถอยริดจ์แบบมีข้อจำกัด (Restricted Ridge Regression) พบว่า วิธีการถดถอยริดจ์ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (Average Mean Squared Error : AMSE) ต่ำที่สุด ถัดมาชุติกัญจน์ และสุปราณี (2561) ได้เปรียบเทียบประสิทธิภาพการวิเคราะห์สัมประสิทธิ์การถดถอย ด้วยวิธีการถดถอยริดจ์ วิธีการถดถอยลาสโซ และวิธีการถดถอยลาสโซแบบปรับปรุง ในกรณีข้อมูลมีมิติสูง และตัวแปรอิสระมีความสัมพันธ์กันค่อนข้างสูง พบว่าวิธีการถดถอยลาสโซแบบปรับปรุง จะให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของค่าพยากรณ์ (Prediction Mean Square Error : PMSE) ต่ำที่สุด เมื่อความสัมพันธ์ของตัวแปรอิสระเพิ่มสูงขึ้น และหลังจากนั้น Algamil and Lee (2015) ประยุกต์ใช้ข้อมูล DNA microarray จากปัญหา

เกี่ยวกับการจำแนกมะเร็งที่มีข้อมูลมิติสูง ซึ่งได้ศึกษาการถดถอยโลจิสติกส์วิธีอีลาสติคเน็ต (Elastic Net Logistic Regression) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (The adaptive Elastic Net Regression) ได้ถูกนำมาประยุกต์ใช้ในการจำแนกประเภทมะเร็งในมิติสูงเพื่อใช้ในการประมาณค่าสัมประสิทธิ์ของยีน และการคัดเลือกเลือกยีนไปพร้อมกัน โดยในงานวิจัยศึกษาเปรียบเทียบการถดถอยโลจิสติกส์วิธีอีลาสติคเน็ต และอีลาสติคเน็ตแบบปรับ เทียบกับวิธีใหม่ที่น่าเสนอ คือวิธีการปรับของการถดถอยอีลาสติคเน็ตแบบปรับ (Adjusted adaptive regularized Logistic Regression : AAElastic) พบว่าวิธีการที่น่าเสนอให้ผลลัพธ์ดีกว่าการถดถอยวิธีอีลาสติคเน็ต และการถดถอยโลจิสติกส์อีลาสติคเน็ตแบบปรับปรุง

จากงานวิจัยที่กล่าวมาข้างต้น แต่ละวิธีสนใจกรณีที่ตัวแปรอิสระเกิดพหุสัมพันธ์กัน แต่ในงานวิจัยนี้ผู้วิจัยได้นำวิธีดังกล่าวมาใช้วิเคราะห์ข้อมูลในกรณีที่ตัวแปรไม่เกิดพหุสัมพันธ์กัน โดยศึกษาการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ ในกรณีที่จำนวนตัวแปรอิสระมากกว่าขนาดตัวอย่าง และตัวแปรตามเป็นข้อมูลเชิงปริมาณ ในลักษณะข้อมูลที่มีมิติสูง นั่นคือ วิเคราะห์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ (Penalized Regression) ซึ่งเป็นวิธีการวิเคราะห์เพื่อประมาณค่าพารามิเตอร์สัมประสิทธิ์การถดถอย (β) ที่ทำให้ฟังก์ชันเป้าหมาย (Objective Function) มีค่าน้อยที่สุด ภายใต้เงื่อนไขที่แตกต่างกัน เรียกว่า ฟังก์ชันการลงโทษ (Penalty Function) โดยฟังก์ชันนี้จะมีด้วยกันหลายรูปแบบ ซึ่งให้ค่าสัมประสิทธิ์การถดถอยที่แตกต่างกัน ในงานวิจัยนี้จะศึกษาวิธีการวิเคราะห์การถดถอยเชิงเส้นแบบปรับด้วยฟังก์ชันการลงโทษ 5 วิธี คือ วิธีการถดถอยริดจ์ วิธีการถดถอยลาสโซ วิธีการถดถอยอีลาสติคเน็ต วิธีการถดถอยลาสโซแบบปรับปรุง และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง เพื่อหาวิธีที่ดีที่สุดในการประมาณค่าพารามิเตอร์ที่อยู่ภายใต้สถานการณ์ที่แตกต่างกันในหลายเงื่อนไข โดยใช้โปรแกรม R ในการจำลองและการวิเคราะห์ข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์

1.2.1 เพื่อศึกษาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง 5 วิธี ได้แก่ วิธีการถดถอยริดจ์ วิธีการถดถอยลาสโซ่ วิธีการถดถอยอิลาสติคเน็ต วิธีการถดถอยลาสโซ่แบบปรับปรุง และวิธีการถดถอยอิลาสติคเน็ตแบบปรับปรุง ในแต่ละสถานการณ์ของการทดลอง

1.2.2 เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง 5 วิธี ได้แก่ วิธีการถดถอยริดจ์ วิธีการถดถอยลาสโซ่ วิธีการถดถอยอิลาสติคเน็ต วิธีการถดถอยลาสโซ่แบบปรับปรุง และวิธีการถดถอยอิลาสติคเน็ตแบบปรับปรุง ในแต่ละสถานการณ์ของการทดลอง

1.3 สมมติฐานของการวิจัย

วิธีการประมาณค่าสัมประสิทธิ์การถดถอยโดยวิธีการถดถอยลาสโซ่แบบปรับปรุง และวิธีการถดถอยอิลาสติคเน็ตแบบปรับปรุง ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำกว่าวิธีการถดถอยริดจ์ วิธีการถดถอยลาสโซ่ และวิธีการถดถอยอิลาสติคเน็ต

1.4 ขอบเขตของการวิจัย

การกำหนดขอบเขตของการวิจัยเป็นดังนี้

1.4.1 ตัวแปรอิสระ (X) สุ่มจากการแจกแจงปกติ ด้วยค่าเฉลี่ย (μ) และความแปรปรวน (σ^2) หรือเขียนได้ว่า $N(\mu, \sigma^2)$ โดยมีฟังก์ชันความหนาแน่นความน่าจะเป็น ดังนี้

$$f(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} & ; \quad -\infty < x < \infty, -\infty < \mu < \infty, \sigma^2 > 0 \\ 0 & ; \quad x \text{ เป็นค่าอื่น ๆ} \end{cases}$$

ในที่นี้ศึกษาการแจกแจงปกติ $N(\mu, \sigma^2)$ ดังนี้ $N(0,4)$

1.4.2 ค่าความคลาดเคลื่อน (ε) สุ่มมาจาก 3 การแจกแจงของข้อมูล ดังนี้

1.4.2.1 การแจกแจงปกติ ที่มีค่าเฉลี่ย (μ) และความแปรปรวน (σ^2) หรือเขียนได้ว่า $N(\mu, \sigma^2)$ โดยมีฟังก์ชันความหนาแน่นความน่าจะเป็น ดังนี้

$$f(\varepsilon) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\varepsilon-\mu}{\sigma}\right)^2} & ; \quad -\infty < \varepsilon < \infty, -\infty < \mu < \infty, \sigma^2 > 0 \\ 0 & ; \quad \varepsilon \text{ เป็นค่าอื่น ๆ} \end{cases}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในงานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่โดยไม่ขออนุญาต โดยเฉพาะอย่างยิ่งในเชิงพาณิชย์

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในที่นี้ศึกษาการแจกแจงปกติ $N(\mu, \sigma^2)$ ดังนี้ $N(0,1)$ $N(0,3)$ $N(0,5)$ และ $N(0,9)$

1.4.2.2 การแจกแจงปรกติปลอมปนระหว่างการแจกแจงปรกติกับการแจกแจงปรกติ โดยมีฟังก์ชันความหนาแน่นความน่าจะเป็น ดังนี้

$$f(\varepsilon) = \begin{cases} (1-p')N(0, \sigma^2) + p'N(0, c^2 \sigma^2) & ; \quad -\infty < \varepsilon < \infty, -\infty < \mu < \infty, \sigma^2 > 0 \\ 0 & ; \quad \varepsilon \text{ เป็นค่าอื่น ๆ} \end{cases}$$

ในที่นี้ศึกษาการแจกแจงปรกติ $N(\mu, \sigma^2)$ เป็น $N(0,1)$ และค่าความน่าจะเป็นที่ข้อมูลมีการปลอมปนการแจกแจงปรกติอื่นเป็น 0.1 ($p' = 0.1$) โดยที่ c คือ สเกลแฟกเตอร์ (Scale Factor) ศึกษาที่ $c = 5$ สำหรับข้อมูลที่มีค่าผิดปกติในระดับไม่รุนแรง และ $c = 10$ สำหรับข้อมูลที่มีค่าผิดปกติในระดับรุนแรง

1.4.2.3 การแจกแจงไวบูลล์ด้วยพารามิเตอร์บอกอัตราส่วน (α) และพารามิเตอร์บอกรูปร่าง (β) หรือเขียนได้ว่า $W(\alpha, \beta)$ โดยมีฟังก์ชันความหนาแน่นความน่าจะเป็น ดังนี้

$$f(\varepsilon) = \begin{cases} \left(\frac{\beta}{\alpha}\right)\left(\frac{\varepsilon}{\alpha}\right)^{\beta-1} e^{-\left(\frac{\varepsilon}{\alpha}\right)^{\beta}} & ; \quad 0 < \varepsilon < \infty, \beta > 0, \alpha > 0 \\ 0 & ; \quad \varepsilon \text{ เป็นค่าอื่น ๆ} \end{cases}$$

ในที่นี้ศึกษาการแจกแจงไวบูลล์ $W(\alpha, \beta)$ ดังนี้ $W(1,1.5)$ $W(1,3)$ และ $W(1,5)$

1.4.3 ศึกษาในกรณีที่จำนวนตัวแปรอิสระ (p) ในตัวแบบมากกว่าขนาดตัวอย่าง (n) ดังนี้

1.4.3.1 กำหนดขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร

1.4.3.2 กำหนดขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร

1.4.3.3 กำหนดขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร

1.4.3.4 กำหนดขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร

1.4.4 ค่าสัมประสิทธิ์การถดถอยพหุคูณ (β) กำหนด 3 กรณี

1.4.4.1 ค่าสัมประสิทธิ์การถดถอยทุกตัวมีค่าเป็น 1

1.4.4.2 ค่าสัมประสิทธิ์การถดถอยมีค่าเป็น 0.5 จำนวนครึ่งหนึ่งของจำนวนตัวแปรอิสระ และที่เหลือมีค่าสัมประสิทธิ์การถดถอยเป็น 1 เช่น ตัวแปรอิสระจำนวน 16 ตัวแปร ค่าสัมประสิทธิ์การถดถอยมีค่าเป็น 0.5 จำนวน 8 ตัวแปร และมีค่าสัมประสิทธิ์การถดถอยเป็น 1 จำนวน 8 ตัวแปร

1.4.4.3 ค่าสัมประสิทธิ์การถดถอยมีค่าเป็น -1 จำนวนครึ่งหนึ่งของจำนวนตัวแปรอิสระ และที่เหลือมีค่าสัมประสิทธิ์การถดถอยเป็น 1 เช่น ตัวแปรอิสระจำนวน 16 ตัวแปร

ค่าสัมประสิทธิ์การถดถอยมีค่าเป็น -1 จำนวน 8 ตัวแปร และมีค่าสัมประสิทธิ์การถดถอยเป็น 1 จำนวน 8 ตัวแปร

1.4.5 ตัวแปรตาม (y) ใช้ตัวแบบการถดถอยเชิงเส้นพหุคูณ ดังนี้ $y = X\beta + \varepsilon$

1.4.6 โปรแกรมที่ใช้ในการวิจัยครั้งนี้ทั้งหมดเขียนด้วยโปรแกรม R เวอร์ชัน 3.5.2 ซึ่งทำการทดลองซ้ำ 1,000 ครั้งในแต่ละสถานการณ์

1.5 เกณฑ์ที่ใช้ในการวิจัย

สำหรับการพิจารณาว่าวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ ภายใต้ข้อมูลมีมิติสูงที่สุด โดยวัดจากความคลาดเคลื่อนค่าประมาณ ซึ่งใช้ตัววัด ดังนี้

1.5.1 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error : MSE) สามารถคำนวณได้ ดังนี้

$$MSE_h = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad ; \quad i = 1, 2, \dots, n, \quad h = 1, 2, \dots, m \quad (1.1)$$

โดยที่ y_i คือ ตัวแปรตามของค่าสังเกตที่ i
 \hat{y}_i คือ ตัวประมาณของค่าสังเกตที่ i
 n คือ ขนาดตัวอย่าง
 m คือ จำนวนรอบของการทดลองซ้ำ

1.5.2 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (Average Mean Squared Error : AMSE) สามารถคำนวณได้ ดังนี้

$$AMSE = \frac{\sum_{h=1}^m MSE_h}{m} \quad (1.2)$$

สำหรับการพิจารณาการเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุด เป็นค่าวัดความถูกต้องที่ได้จากการทดลองซ้ำ (m) 1,000 ครั้งของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) โดยมีค่ามากกว่าหรือเท่ากับศูนย์ ซึ่งหากค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) มีค่าเข้าใกล้ศูนย์มากเท่าใด หมายถึง ตัวประมาณนั้นมีความผันแปรน้อยลงเท่านั้น กล่าวคือหากเอกสารนี้เป็นตัวประมาณความแม่นยำในการประมาณค่ามาก ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ก็จะไม่ว่ากรณีนี้จะเข้าใกล้ศูนย์ ห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นเกณฑ์ที่ใช้ในการตัดสินใจเลือกตัวประมาณค่าที่ดีที่สุด คือ ตัวประมาณ (\hat{y}_i) ที่ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุด

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1.6.1 เพื่อทราบประสิทธิภาพของวิธีประมาณพารามิเตอร์ในการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง ในแต่ละสถานการณ์ของการทดลอง

1.6.2 เพื่อเป็นแนวทางในการตัดสินใจเลือกตัวประมาณค่าพารามิเตอร์สัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง ในแต่ละสถานการณ์ได้อย่างเหมาะสม

1.6.3 เพื่อใช้เป็นแนวทางในการศึกษาวิจัยวิธีการวิเคราะห์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูงในการเปรียบเทียบกับตัวประมาณวิธีอื่น ๆ ต่อไป

1.7 นิยามศัพท์เฉพาะที่ใช้ในงานวิจัย

1.7.1 ข้อมูลที่มีมิติสูง (High-Dimensional Data) หมายถึง ข้อมูลที่มีจำนวนตัวแปรอิสระมากกว่าขนาดตัวอย่าง

1.7.2 การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ (Penalized Regression) หมายถึง การถดถอยเชิงเส้นพหุคูณเมื่อข้อมูลมีมิติสูง โดยมีการปรับด้วยฟังก์ชันการลงโทษ

1.7.3 ฟังก์ชันเป้าหมาย ($\hat{\beta}_p$) (Objective Function) หมายถึง การหาตัวประมาณสัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ ภายใต้ข้อมูลที่มีมิติสูงที่มีค่าน้อยที่สุด

1.7.4 ฟังก์ชันการลงโทษ ($P_\lambda(\beta)$) (Penalized Function) หมายถึง ฟังก์ชันที่เพิ่มเข้ามาเพื่อปรับให้ตัวประมาณสัมประสิทธิ์การถดถอยมีค่าน้อยที่สุด ภายใต้เงื่อนไขที่แตกต่างกันในแต่ละวิธีการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ

1.7.5 พารามิเตอร์ปรับแต่ง (λ) (Tuning Parameter) หมายถึง พารามิเตอร์ที่หดค่าลง (Shrinkage) ของการประมาณค่าสัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ สำหรับข้อมูลที่มีมิติสูง ที่อยู่ในฟังก์ชันการลงโทษ

1.7.6 ประสิทธิภาพ หมายถึง เกณฑ์ที่ใช้ในการตัดสินใจเลือกตัวประมาณค่าที่ดีที่สุด โดยวัดจากค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ที่มีค่าต่ำที่สุดของวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การถดถอยเชิงเส้นพหุคูณ เมื่อข้อมูลเป็นไปตามข้อกำหนดเบื้องต้นจะนิยมใช้วิธีกำลังสองน้อยที่สุด (Ordinary Least Squares Method) สำหรับการประมาณค่าสัมประสิทธิ์การถดถอยในตัวแบบ ในกรณีที่ข้อมูลมีขนาดตัวอย่างมากกว่าจำนวนตัวแปรอิสระ แต่เมื่อจำนวนตัวแปรอิสระมากกว่าขนาดตัวอย่างจะใช้การวิเคราะห์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ (Penalized Regression) ภายใต้ข้อมูลที่มีมิติสูง (High-Dimensional Data) ในงานวิจัยนี้จะกล่าวถึงวิธีการวิเคราะห์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง 5 วิธี ได้แก่ วิธีการถดถอยริดจ์ (Ridge Regression method) วิธีการถดถอยลาสโซ (Least absolute shrinkage and selection operator Regression method : Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net Regression method) วิธีการถดถอยลาสโซแบบปรับปรุง (The adaptive least absolute shrinkage and selection operator Regression method) และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (The adaptive Elastic net Regression method) โดยวิธีการหาค่าพารามิเตอร์ปรับแต่งในแต่ละการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษที่ใช้กันอย่างแพร่หลายมี 2 วิธี ได้แก่ วิธีการตรวจสอบข้าม (Cross-Validation : CV) และวิธีเกณฑ์ข้อสนเทศของเบส์ (Bayesian Information Criteria : BIC) และในการศึกษาครั้งนี้ผู้วิจัยได้ทำการค้นคว้าเอกสาร ตำรา และงานวิจัยที่เกี่ยวข้องกับการแจกแจงต่าง ๆ ของข้อมูล

2.1 การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression)

ตัวแบบการถดถอยเชิงเส้นพหุคูณเป็นตัวแบบที่มีตัวแปรตาม (Dependent Variable) เป็นเชิงปริมาณเพียง 1 ตัว ซึ่งมีความสัมพันธ์เชิงเส้นกับตัวแปรอิสระ (Independent Variable) เป็นเชิงปริมาณหรือเชิงคุณภาพมากกว่า 1 ตัว โดยทั่วไปตัวแปรตาม (y_i) อาจจะสัมพันธ์กับตัวแปรอิสระ (X) จำนวน p ตัว จะสามารถเขียนตัวแบบได้ ดังนี้ (สายชล, 2560)

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i \quad ; \quad i = 1, 2, 3, \dots, n \quad (2.1)$$

โดยที่

y_i	คือ ตัวแปรตามของค่าสังเกตที่ i
$X_{i1}, X_{i2}, \dots, X_{ip}$	คือ ตัวแปรอิสระของค่าสังเกตที่ i
$\beta_0, \beta_1, \beta_2, \dots, \beta_p$	คือ ค่าสัมประสิทธิ์การถดถอย (Coefficients of Regression)
p	คือ จำนวนตัวแปรอิสระ
ε_i	คือ ค่าความคลาดเคลื่อนของค่าสังเกตที่ i
n	คือ ขนาดตัวอย่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาค้นคว้าเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกหรือทำซ้ำโดยไม่ได้รับอนุญาตจากผู้จัดทำเอกสาร

หรือเขียนอยู่ในรูปของเมทริกซ์ได้ดังนี้ (วรภาพร, 2558)

$$\tilde{y} = X\tilde{\beta} + \varepsilon \quad (2.2)$$

เมื่อ

$$\tilde{y}_{n \times 1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$X_{n \times (p+1)} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}$$

$$\tilde{\beta}_{(p+1) \times 1} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

$$\varepsilon_{n \times 1} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

โดยที่	\tilde{y}	แทน	เวกเตอร์ของตัวแปรตามขนาด $n \times 1$
	X	แทน	เมทริกซ์ของค่าสังเกตของตัวแปรอิสระขนาด $n \times (p+1)$
	$\tilde{\beta}$	แทน	เวกเตอร์ของค่าสัมประสิทธิ์การถดถอยพหุคูณขนาด $(p+1) \times 1$
	p	แทน	จำนวนตัวแปรอิสระ
	n	แทน	ขนาดตัวอย่าง
	ε	แทน	เวกเตอร์ของค่าความคลาดเคลื่อนที่เกิดขึ้นขนาด $n \times 1$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยมีข้อตกลงเบื้องต้นของตัวแบบการถดถอยพหุคูณ ดังนี้

1. $\varepsilon \sim N(0, \sigma^2 I)$ เมื่อ I คือ เมทริกซ์เอกลักษณ์ขนาด $n \times n$
2. ค่าของตัวแปรอิสระแต่ละตัวเป็นค่าคงที่
3. ตัวแปรอิสระแต่ละตัวต้องไม่มีความสัมพันธ์กับตัวแปรอิสระอื่น ๆ

สำหรับการประมาณค่าสัมประสิทธิ์การถดถอยหรือค่าพารามิเตอร์ $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ โดยใช้วิธีกำลังสองน้อยที่สุด (Ordinary Least Squares Method : OLS) จากตัวแบบ (2.2) เป็นดังนี้

วัตถุประสงค์:
$$\min \sum_{i=1}^n \varepsilon_i^2$$

$$\begin{aligned} \varepsilon &= (y - \hat{y}) \\ &= (y - X\hat{\beta}) \\ \varepsilon^T \varepsilon &= (y - X\hat{\beta})^T (y - X\hat{\beta}) \\ &= (y^T - X^T \hat{\beta}^T)(y - X\hat{\beta}) \\ &= y^T y - y^T X \hat{\beta} - \hat{\beta}^T X^T y + \hat{\beta}^T X^T X \hat{\beta} \\ &= y^T y - 2\hat{\beta}^T X^T y + \hat{\beta}^T X^T X \hat{\beta} \end{aligned} \quad (2.3)$$

ค่าประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณได้จากการหาอนุพันธ์ (Differentiation) สมการ (2.3) เทียบกับ $\hat{\beta}$ แล้วให้เท่ากับ 0

$$\frac{\partial \varepsilon^T \varepsilon}{\partial \hat{\beta}} = -2X^T y + 2X^T X \hat{\beta} = 0$$

จะได้สมการปกติ (Normal equations) คือ

$$X^T X \hat{\beta} = X^T y$$

ดังนั้นตัวประมาณค่าสัมประสิทธิ์การถดถอย คือ

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (2.4)$$

โดยจะได้ $\hat{\beta}$ ที่สามารถประมาณค่า \hat{y} (Fitted value) ซึ่งคำนวณได้จาก

เอกสารนี้เป็นเอกสารที่ $\hat{y} = X\hat{\beta}$ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ (2.5) ารค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยทั่วไปวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ จะใช้วิธีกำลังสองน้อยที่สุด โดยตัวประมาณค่าสัมประสิทธิ์การถดถอย ดังนี้

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j X_{ij})^2 = (X^T X)^{-1} X^T y \quad (2.6)$$

แต่สำหรับข้อมูลที่มีมิติสูง คือ เมื่อจำนวนตัวแปรอิสระมากกว่าขนาดตัวอย่าง เมทริกซ์ $X^T X$ จะเป็นเมทริกซ์เอกฐาน (Singular Matrix) ซึ่งไม่สามารถหาเมทริกซ์ผกผัน (Inverse of a Matrix) ได้ ทำให้หาค่าประมาณโดยวิธีกำลังสองน้อยที่สุดไม่ได้ จึงหันมาใช้วิธีการวิเคราะห์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ

2.2 วิธีที่ใช้ในการวิจัย

วิธีการวิเคราะห์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ (Penalized Regression) ซึ่งเป็นการวิเคราะห์เพื่อหาค่าประมาณพารามิเตอร์สัมประสิทธิ์การถดถอย ที่ทำให้ฟังก์ชันเป้าหมายเป็นดังสมการ

$$\hat{\beta}_p = \arg \min_{\beta} \left[\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j X_{ij})^2 + P_{\lambda}(\beta) \right] ; \lambda > 0 \quad (2.7)$$

โดยที่ $P_{\lambda}(\beta)$ คือ ฟังก์ชันการลงโทษ (Penalty Function)
 λ คือ พารามิเตอร์ปรับแต่ง (Tuning Parameter)

สมการข้างต้นถือเป็นการคัดกรองตัวแปรเข้าในตัวแบบหากเราเลือกใช้ฟังก์ชันการลงโทษที่เหมาะสมจะสามารถคัดกรองตัวแปรเข้าในตัวแบบได้ กล่าวคือฟังก์ชันการลงโทษจะทำให้ค่าสัมประสิทธิ์บางตัวมีค่าเท่ากับศูนย์ซึ่งในการวิจัยครั้งนี้ ผู้วิจัยสนใจทำการเปรียบเทียบวิธีการคัดเลือกตัวแปรอิสระ 5 วิธีดังนี้

2.2.1 วิธีการถดถอยริดจ์ (Ridge Regression method)

Hoerl and Kennard (1970) เสนอวิธีการถดถอยริดจ์ เพื่อแก้ปัญหาการเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ ซึ่งจะลดความคลาดเคลื่อนกำลังสองของค่าเฉลี่ยให้มีค่าต่ำลง จากหลักการของวิธีกำลังสองน้อยที่สุดพบว่าถ้าต้องการลดความคลาดเคลื่อนกำลังสองเฉลี่ยต้องทำให้ $|X^T X|$ มีค่าเพิ่มขึ้น โดยบวกค่าคงที่ λ ซึ่งมีค่ามากกว่าศูนย์แต่น้อยกว่าหนึ่ง เข้ากับสมาชิกทุกตัวบนเส้นทแยงมุมของเมทริกซ์ $X^T X$ ดังนั้นตัวประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีการถดถอยริดจ์อยู่ในรูปดังต่อไปนี้

$$\hat{\beta}_R = (X^T X + \lambda I_{p+1})^{-1} X^T y \quad ; \quad 0 < \lambda < 1 \quad (2.8)$$

เมื่อ λ เป็นค่าคงที่ ซึ่งอยู่ระหว่าง 0 ถึง 1

และ I เป็นเมทริกซ์เอกลักษณ์ (Identity Matrix)

จากสมการข้างต้นสามารถแปลงตัวประมาณวิธีการถดถอยริดจ์ ให้อยู่ในรูปของตัวประมาณ $\hat{\beta}$ และจากสมการปกติ $X^T X \hat{\beta} = X^T y$ จะประมาณวิธีการถดถอยริดจ์ได้ดังนี้

$$\begin{aligned} \hat{\beta}_R &= (X^T X + \lambda I)^{-1} X^T y \\ &= (X^T X + \lambda I)^{-1} (X^T X) (X^T X)^{-1} X^T y \\ &= (X^T X + \lambda I)^{-1} X^T X [(X^T X)^{-1} X^T y] \\ &= [I + \lambda (X^T X)^{-1}]^{-1} \hat{\beta}; \hat{\beta} = \hat{\beta}_{OLS} \\ &= Z_\lambda \hat{\beta} \end{aligned}$$

$$\text{โดยที่ } Z_\lambda = [I + \lambda (X^T X)^{-1}]^{-1}$$

โดยค่าสัมประสิทธิ์การถดถอยพหุคูณวิธีการถดถอยริดจ์ จะอยู่ในรูปผลบวกระหว่างผลรวมความคลาดเคลื่อนกำลังสองและผลรวมกำลังสองของค่าสัมประสิทธิ์ถ่วงน้ำหนักให้มีค่าต่ำสุด ดังนี้

$$\hat{\beta}_R = \arg \min_{\beta} \left[\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j X_{ij})^2 + \lambda \sum_{j=1}^p \beta_j^2 \right] \quad ; \quad \lambda > 0 \quad (2.9)$$

เมื่อ $\lambda \sum_{j=1}^p \beta_j^2$ คือ ฟังก์ชันการลงโทษ และ λ คือพารามิเตอร์ปรับแต่ง (Tuning Parameter) ซึ่งควบคุมขนาดการหดตัว (Shrinkage) ของ $\hat{\beta}_R$ โดยใช้วิธีการตรวจสอบไขว้ (Cross-Validation) ในที่นี้จะใช้วิธีการตรวจสอบไขว้ (The generalized cross-validation) (Boonstra. et al., 2015) สำหรับการหาค่าพารามิเตอร์ปรับแต่งที่น้อยที่สุด

2.2.2 วิธีการถดถอยลาสโซ่ (Least absolute shrinkage and selection operator Regression method : Lasso)

Tibshirani (1996) เสนอวิธีการวิเคราะห์การถดถอยลาสโซ่ โดยวิธีลาสโซ่เป็นวิธีที่นิยมสำหรับการประมาณค่าและการคัดเลือกตัวแปรอิสระเข้าสู่ตัวแบบในคราวเดียวกัน ในกรณีที่จำนวนตัวแปรอิสระมีความสัมพันธ์กับตัวแปรตามจำนวนไม่มากนักในตัวแบบ จะทำให้วิธีลาสโซ่มีประสิทธิภาพสูงสุดในการพยากรณ์ โดยค่าสัมประสิทธิ์การถดถอยพหุคูณของวิธีลาสโซ่ จะอยู่ในรูปผลบวกระหว่างผลรวมความคลาดเคลื่อนกำลังสองและผลรวมสัมบูรณ์ของค่าสัมประสิทธิ์ถ่วงน้ำหนักให้มีค่าต่ำสุด หากค่าถ่วงน้ำหนักมาก ๆ จะทำให้สัมประสิทธิ์นั้นหดตัวเหลือเท่ากับศูนย์หรือหายไป กล่าวคือวิธีนี้จะ

ทำให้ค่าสัมประสิทธิ์ส่วนใหญ่เป็นศูนย์และค่าสัมประสิทธิ์บางส่วนไม่เท่ากับศูนย์ (Sparse Estimator) โดยตัวประมาณค่าสัมประสิทธิ์การถดถอยวิธีลาสโซ่ ดังนี้

$$\hat{\beta}_L = \arg \min_{\beta} \left[\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j X_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j| \right] ; \lambda > 0 \quad (2.10)$$

เมื่อ $\lambda \sum_{j=1}^p \beta_j$ คือ ฟังก์ชันการลงโทษ และ λ คือพารามิเตอร์ปรับแต่ง (Tuning Parameter)

ซึ่งควบคุมขนาดการหดตัว (Shrinkage) ของ $\hat{\beta}_L$ โดยใช้วิธีการตรวจสอบไขว้ (Cross-Validation) ในที่นี้จะใช้วิธีอัลกอริทึม LARS (LARS algorithm) (Efron. *et al.*, 2004) สำหรับการหาค่าพารามิเตอร์ปรับแต่งที่น้อยที่สุด

2.2.3 วิธีการถดถอยอีลาสติคเน็ต (Elastic Net Regression method)

Zou and Hastie (2005) เสนอวิธีการถดถอยอีลาสติคเน็ต โดยเป็นการรวมกันของ 2 วิธีระหว่างวิธีการถดถอยริดจ์ และวิธีการถดถอยลาสโซ่ จึงสามารถคัดเลือกตัวแปรอิสระและประมาณค่าไปพร้อมกัน วิธีนี้จึงเหมาะสำหรับการวิเคราะห์ข้อมูลที่มีจำนวนตัวแปรอิสระมากกว่าขนาดตัวอย่างมากๆ และตัวแปรอิสระที่มีความสัมพันธ์กันสูง โดยค่าสัมประสิทธิ์การถดถอยพหุคูณของวิธีอีลาสติคเน็ตจะอยู่ในรูปผลบวกของผลรวมความคลาดเคลื่อนกำลังสอง ผลรวมสัมบูรณ์ของค่าสัมประสิทธิ์ถ่วงน้ำหนัก และผลรวมกำลังสองของค่าสัมประสิทธิ์ถ่วงน้ำหนักให้มีค่าต่ำสุด ซึ่งสามารถประมาณค่าสัมประสิทธิ์การถดถอยโดยวิธีอีลาสติคเน็ต ได้ดังนี้

$$\hat{\beta}_E = \arg \min_{\beta} \left[\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j X_{ij})^2 + \lambda_1 \sum_{j=1}^p |\beta_j| + \lambda_2 \sum_{j=1}^p \beta_j^2 \right] ; 0 < \lambda_1 + \lambda_2 < \infty \quad (2.11)$$

ฟังก์ชันการลงโทษของวิธีการถดถอยอีลาสติคเน็ตเป็นการรวมกันของฟังก์ชันการลงโทษของวิธีการถดถอยริดจ์และวิธีการถดถอยลาสโซ่ และเมื่อ $\lambda_1 = 0$ วิธีการถดถอยอีลาสติคเน็ตจะเป็นวิธีการถดถอยริดจ์อย่างง่าย (Simple ridge regression) ซึ่งฟังก์ชันการลงโทษของวิธีการถดถอยอีลาสติคเน็ตจะมี λ_1 และ λ_2 คือพารามิเตอร์ปรับแต่ง (Tuning Parameter) ซึ่งควบคุมขนาดการหดตัว (Shrinkage) ของ $\hat{\beta}_E$ โดยจะใช้วิธีการตรวจสอบไขว้ (Cross-Validation) (Hastie. *et al.*, 2009) ในการหาพารามิเตอร์ปรับแต่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.4 วิธีการถดถอยลาสโซ่แบบปรับปรุง (The adaptive least absolute shrinkage and selection operator Regression method)

วิธีการถดถอยลาสโซ่ยังมีความอ่อนแอในการคัดเลือกตัวประมาณสัมประสิทธิ์การถดถอย Zou (2006) จึงได้เสนอวิธีการถดถอยลาสโซ่แบบปรับปรุง โดยเพิ่มค่าถ่วงน้ำหนัก (Weight) อีกหนึ่งพารามิเตอร์เข้ามาในวิธีการลาสโซ่ ซึ่งการให้ค่าถ่วงน้ำหนัก กับพารามิเตอร์แต่ละตัวแตกต่างกันในฟังก์ชันการลงโทษ $P_\lambda(\beta)$ โดยการกำหนดค่าถ่วงน้ำหนักให้มีค่าสูงสำหรับค่าสัมประสิทธิ์ที่มีค่าน้อย และกำหนดค่าถ่วงน้ำหนักให้น้อยสำหรับค่าสัมประสิทธิ์ที่มีค่าสูง เพื่อลดความไม่คงเส้นคงวาที่ทำให้เกิดความอ่อนแอในการประมาณค่าสัมประสิทธิ์การถดถอยในวิธีการถดถอยลาสโซ่ การใช้ฟังก์ชันการลงโทษนี้ยังมีคุณสมบัติอีก 1 ข้อ คือ เมื่อขนาดตัวอย่างมีจำนวนมากพอ วิธีการถดถอยลาสโซ่แบบปรับปรุง จะมีความสามารถในการเลือกตัวแปรได้เสมือนกับว่าทราบตัวแบบที่แท้จริง (True Model) ซึ่งคุณสมบัตินี้ไม่มีในวิธีการถดถอยลาสโซ่ โดยสามารถประมาณค่าสัมประสิทธิ์การถดถอยวิธีการลาสโซ่แบบปรับปรุง ได้ดังนี้

$$\hat{\beta}_{AL} = \arg \min_{\beta} \left[\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j X_{ij})^2 + \lambda \sum_{j=1}^p \hat{w}_j |\beta_j| \right] \quad (2.12)$$

$$\text{โดยที่ } \hat{w}_j = \frac{1}{|\hat{\beta}_{j(L)}|^\gamma}; j = 1, 2, \dots, p, \gamma > 0$$

เมื่อ $\hat{\beta}_{j(L)}$ คือ ค่าประมาณสัมประสิทธิ์การถดถอยวิธีการลาสโซ่

โดยทั่วไปกำหนดให้ $\gamma = 1$ และ $\lambda \sum_{j=1}^p \hat{w}_j |\beta_j|$ คือ ฟังก์ชันการลงโทษของวิธีการถดถอยลาสโซ่แบบปรับปรุง โดยเป็นฟังก์ชันการลงโทษของการวิธีการถดถอยลาสโซ่ที่เพิ่มค่าถ่วงน้ำหนักขึ้นมา โดยใช้วิธีการตรวจสอบไขว้แบบวิถิเกณฑ์ข้อสนเทศของเบส์ (BIC Cross-Validation) (Zou. *et al.*, 2007) สำหรับการหาค่าพารามิเตอร์ปรับแต่งที่น้อยที่สุด

2.2.5 วิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (The adaptive Elastic net Regression method)

เนื่องจากวิธีการถดถอยอีลาสติคเน็ตไม่มีคุณสมบัติที่ว่าเมื่อขนาดตัวอย่างมีจำนวนมากพอ จะมีความสามารถในการเลือกตัวแปรได้เสมือนกับว่าทราบตัวแบบที่แท้จริง (True Model) จึงเสนอวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง โดย Zou and Zhang (2009) ได้พัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอย เพื่อเพิ่มประสิทธิภาพในการประมาณค่าสัมประสิทธิ์ และการคัดเลือกตัวแปรอิสระ โดยรวมระหว่างวิธีการถดถอยอีลาสติคเน็ต และวิธีการถดถอยลาสโซ่แบบปรับปรุง เข้าด้วยกัน สามารถประมาณค่าสัมประสิทธิ์การถดถอยวิธีการอีลาสติคเน็ตแบบปรับปรุง ได้ดังนี้

$$\hat{\beta}_{AE} = \arg \min_{\beta} \left[\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j X_{ij})^2 + \lambda_1 \sum_{j=1}^p \hat{w}_j |\beta_j| + \lambda_2 \sum_{j=1}^p \beta_j^2 \right] \quad (2.13)$$

$$\text{โดยที่ } \hat{w}_j = \left(|\hat{\beta}_{j(E)}| + \frac{1}{n} \right)^{-\gamma} ; \quad \gamma > 0$$

เมื่อ $\hat{\beta}_{j(E)}$ คือ ค่าประมาณสัมประสิทธิ์การถดถอยวิธีอิลาสติกเน็ต

โดยทั่วไปกำหนดให้ $\gamma = 1$ และฟังก์ชันการลงโทษของวิธีการถดถอยอิลาสติกเน็ตแบบปรับปรุงเป็นการรวมกันของฟังก์ชันการลงโทษของวิธีการถดถอยอิลาสติกเน็ต และวิธีลาสโซ่แบบปรับปรุง ซึ่งหาค่าพารามิเตอร์ปรับแต่งที่น้อยที่สุดโดยใช้วิธีการตรวจสอบไขว้แบบวิธีเกณฑ์ข้อสนเทศของเบส์ (BIC Cross-Validation) (Zou. *et al.*, 2007)

2.3 วิธีเลือกพารามิเตอร์การปรับ (Tuning Parameter)

การหาค่าพารามิเตอร์การปรับที่เหมาะสมที่สำคัญในการวิเคราะห์การถดถอยลาสโซ่ และการประมาณค่าสัมประสิทธิ์ $\hat{\beta}$ สำหรับวิธีการเลือกพารามิเตอร์การปรับที่ใช้กันอย่างแพร่หลายมี 2 วิธี คือ วิธีการตรวจสอบไขว้ (Cross-Validation) และวิธีเกณฑ์ข้อสนเทศของเบส์ (Bayesian Information Criterion) ดังนี้

2.3.1 วิธีการตรวจสอบไขว้ (Cross-Validation : CV) (จุฑาทิพย์, 2554)

เป็นวิธีการวัดประสิทธิภาพของตัวแบบ โดยแบ่งข้อมูลออกเป็น 2 ส่วน คือข้อมูลชุดการเรียนรู้ (Training Data) ใช้สร้างตัวแบบการถดถอย และข้อมูลชุดตรวจสอบ (Testing Data) ใช้ตรวจสอบตัวแบบการถดถอยที่ได้จากชุดการเรียนรู้ โดยพิจารณาจากค่าความผิดพลาดในการทำนาย

วิธีการตรวจสอบไขว้ เป็นวิธีที่นิยมใช้กันโดยทั่วไปในการหาค่าพารามิเตอร์การปรับสำหรับการวิเคราะห์การถดถอยลาสโซ่ โดยการแบ่งกลุ่มของชุดข้อมูลออกเป็นกลุ่มย่อย k กลุ่ม จากนั้นนำข้อมูลเพียง $k-1$ กลุ่ม (Training Data) มาทำนายกลุ่มที่เหลืออีก 1 กลุ่ม (Testing Data) ทำเช่นนี้ไป k ครั้ง แล้วพิจารณาค่าความผิดพลาดในการทำนาย โดยทั่วไปแล้วจะนิยมใช้ $k = 5$ หรือ $k = 10$ มีขั้นตอนในการคำนวณ ดังนี้ (Syed, 2011)

- 1) กำหนด λ ให้ตัวประมาณแบบจุดมีค่าต่างกันไป จะได้ λ ที่เป็นไปได้เป็นเซตที่ประกอบไปด้วย λ ทั้งหมด q ตัว ดังนี้ $\Lambda = \{\lambda_1, \dots, \lambda_q\}$
- 2) แบ่งกลุ่มของข้อมูลทั้งหมดออกเป็น k กลุ่มๆ ละเท่าๆกัน กลุ่มละข้อมูล n ตัว
- 3) สำหรับ $k = 1, 2, \dots, K$ ทำการทำนาย ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ครั้งที่ k ใช้ข้อมูลจำนวน $K-1$ ชุด คือชุดที่ 1 ถึงชุดที่ K ใดๆ ที่ไม่ใช่ชุดที่ k

- a) ในแต่ละ $\lambda_l; l=1,2,\dots,q$ ทำการประมาณค่าพารามิเตอร์ $\{\hat{\beta}_j\}_l$ แล้ว นำไปหาค่าพยากรณ์ $\{\hat{y}_i\}_l$ ของข้อมูลชุดที่ k
- b) ทำการคำนวณหาค่าความผิดพลาดจาก

$$e_k(\lambda_l) = \sum_{i=1}^{n_k} (y_i - \{\hat{y}_i\}_l)^2 \quad (2.14)$$

- 4) ในแต่ละ $\lambda_l; l=1,2,\dots,q$ คำนวณค่าความผิดพลาดเฉลี่ย λ_l จากทั้งหมด k ครั้ง ดังนี้

$$CV(\lambda_l) = \frac{1}{n} \sum_{k=1}^K e_k(\lambda_l) \quad (2.15)$$

- 5) ตัว λ_l ที่มีค่า $CV(\lambda_l)$ น้อยที่สุด จะเป็นพารามิเตอร์การปรับค่าที่เหมาะสมที่สุด

2.3.2 วิธีเกณฑ์ข้อสนเทศของเบส์ (Bayesian Information Criterion : BIC)

(จุฬาทิพย์, 2554)

วิธีเกณฑ์ข้อสนเทศของเบส์ หรือวิธี Schwarz Information Criterion เป็นวิธีที่ถูกคิดค้นโดย Gideon Schwarz ในปี ค. ศ. 1978 เพื่อใช้ในการคัดเลือกตัวแบบของการวิเคราะห์การถดถอยเชิงเส้นทั่วไป ซึ่งเป็นการดัดแปลงจากเกณฑ์ข้อสนเทศของอาไคเคะ (Akaike Information Criterion : AIC) โดยพิจารณาจากค่าสูงสุดของความน่าจะเป็นภายหลัง (Posterior Probability) ของตัวแบบการถดถอยนั้นๆ ต่อมาถูกพัฒนามาใช้กับข้อมูลที่มีมิติสูงโดย Wang. *et al.* (2009) ในการหาค่าพารามิเตอร์การปรับสำหรับการวิเคราะห์การถดถอยในข้อมูลมิติสูงเพื่อให้ได้ตัวแบบการถดถอยที่มีความคงเส้นคงวา (Consistency) มีขั้นตอนในการคำนวณดังนี้

- 1) กำหนด λ ให้เป็นตัวประมาณแบบจุดมีค่าต่างกันไป จะได้ λ ที่เป็นไปได้เป็นเซตที่ประกอบไปด้วย λ ทั้งหมด q ตัว ดังนี้ $\Lambda = \{\lambda_1, \dots, \lambda_q\}$
- 2) ในแต่ละ $\lambda_l; l=1,2,\dots,q$ ทำการประมาณค่าพารามิเตอร์ $\{\hat{\beta}_j\}_l$ และประมาณค่า BIC โดยคำนวณจาก

$$BIC_l = \log(\hat{\sigma}_l^2) + |S_l| \times \frac{\log n}{n} \times C_n \quad (2.16)$$

เมื่อ $\hat{\sigma}_l^2 = SSE_l / n$

$C_n = \log \log p$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในงานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 S_l คือ เซต $\{\hat{\beta}_j\}_l$ ที่ไม่เท่ากับ 0

- 3) เมื่อ λ_l ที่มีค่า BIC_l น้อยที่สุด จะเป็นพารามิเตอร์การปรับที่เหมาะสมที่สุด

2.4 การแจกแจงที่ใช้ในการวิจัย

โดยทั่วไปค่าความคลาดเคลื่อนจะมีลักษณะการแจกแจงแบบใดแบบหนึ่ง เช่น การแจกแจงปกติ เป็นต้น

เนื่องจากในความเป็นจริงข้อมูลที่นำมาใช้ในการวิเคราะห์อาจไม่ได้มีการแจกแจงปกติเพียงอย่างเดียว ดังนั้นในงานวิจัยนี้จึงนำเสนอการแจกแจง 3 การแจกแจง คือ การแจกแจงปกติ การแจกแจงปกติปลอมปน และการแจกแจงไวบูล ซึ่งแต่ละการแจกแจงมีรายละเอียด ดังนี้

2.4.1 การแจกแจงปกติ (Normal Distribution) (อัชฌา, 2561)

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่มีค่าต่อเนื่องที่สำคัญที่สุดคือการแจกแจงปกติ ซึ่งเป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่มีค่าส่วนมากใกล้เคียงกับค่าเฉลี่ยของตัวแปรสุ่มนั้น โดยจะมีค่าของตัวแปรสุ่มที่มากกว่าหรือน้อยกว่าค่าเฉลี่ยอย่างมากเป็นส่วนน้อย

นิยามที่ 2.1 ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงปกติ ด้วยพารามิเตอร์ μ และ σ^2 แล้วฟังก์ชันความหนาแน่นความน่าจะเป็นคือ

$$f(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} & ; -\infty < x < \infty, -\infty < \mu < \infty, \sigma^2 > 0 \\ 0 & ; x \text{ เป็นค่าอื่น ๆ} \end{cases}$$

หรือเขียนแทนด้วยสัญลักษณ์ $X \sim N(\mu, \sigma^2)$

ทฤษฎีที่ 2.1 ถ้าตัวแปรสุ่ม X มีการแจกแจงปกติ ด้วยพารามิเตอร์ μ และ σ^2 แล้วตัวแปรสุ่ม X จะมีค่าเฉลี่ยและความแปรปรวน ดังนี้

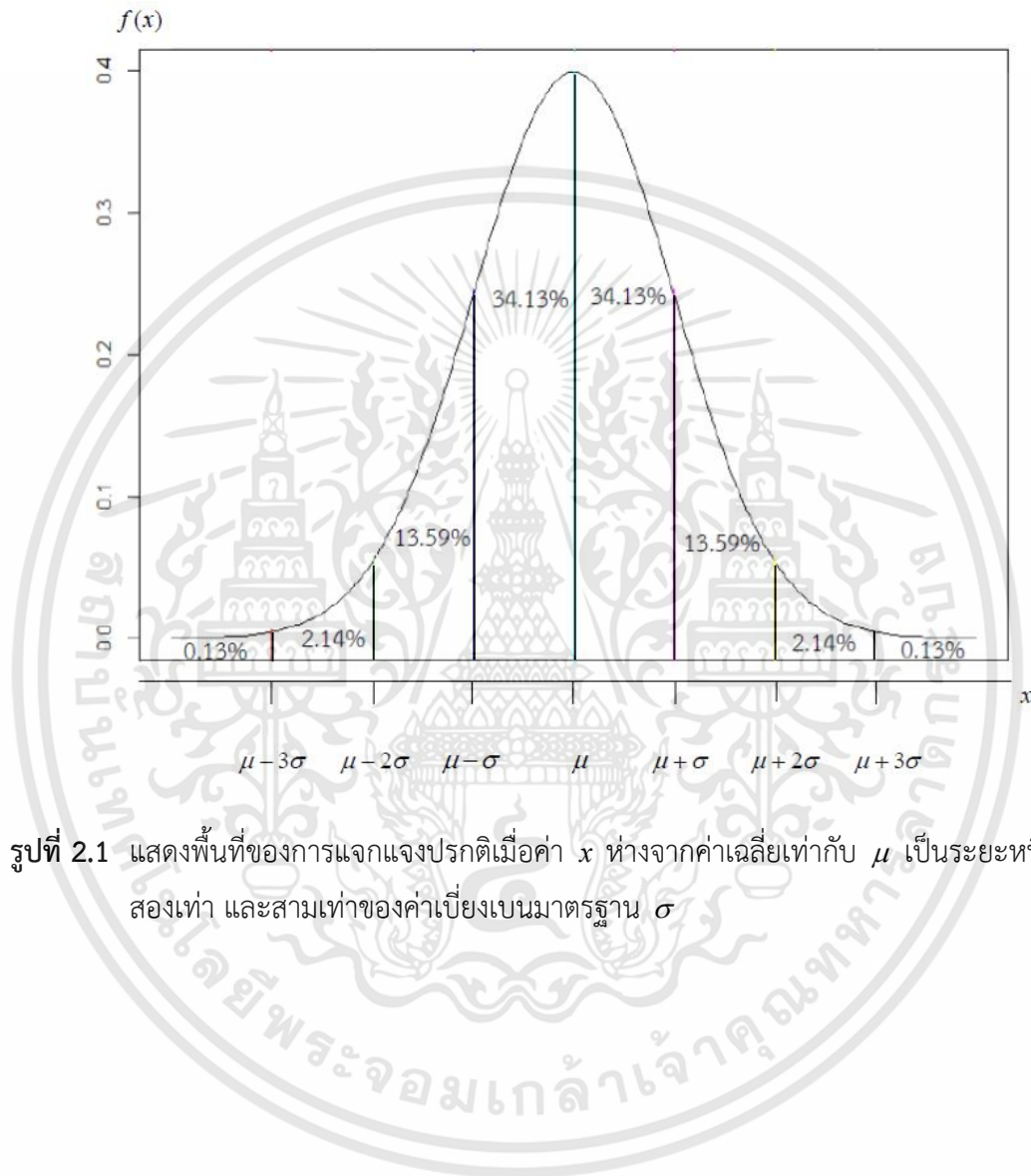
$$\begin{aligned} E(X) &= \mu \\ \text{Var}(X) &= \sigma^2 \end{aligned}$$

คุณสมบัติของโค้งการแจกแจงปกติ

1. มีความสมมาตรรอบค่าเฉลี่ย μ
 2. ค่าเฉลี่ย มัธยฐาน และฐานนิยม มีค่าเท่ากันหมด เท่ากับ μ
 3. พื้นที่ใต้กราฟทั้งหมดเท่ากับ 1
 4. ถ้าลากเส้นตั้งฉากจากแกน x ไปยังเส้นกราฟของการแจกแจงปกติ โดยที่เส้นดังกล่าวห่างจากค่าเฉลี่ย μ ทั้งด้านซ้ายและด้านขวาด้วยระยะหนึ่งเท่าของค่าเบี่ยงเบนมาตรฐาน σ ได้พื้นที่ใต้กราฟ คือ 0.6826 หรือ 68.26% (34.13% + 34.13%) สองเท่าของค่าเบี่ยงเบนมาตรฐาน σ ได้พื้นที่ใต้กราฟ คือ 0.9544 หรือ 95.44% (68.26% + 13.59% + 13.59%) และสามเท่าของค่าเบี่ยงเบนมาตรฐาน σ ได้พื้นที่ใต้กราฟคือ 0.9972 หรือ 99.72% (95.44% + 2.14%+2.14%)
- ดังรูป 2.1

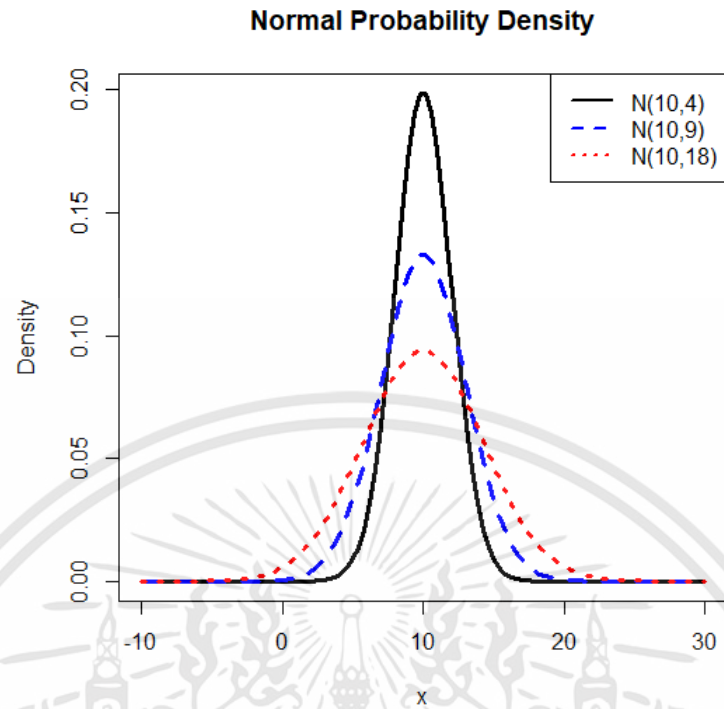
5. ในกรณีที่การแจกแจงปกติมีค่าเฉลี่ย μ เท่ากัน ค่าความแปรปรวน σ^2 จะเป็นตัวกำหนดลักษณะของเส้นกราฟดังรูป 2.2

6. ในกรณีที่การแจกแจงปกติมีค่าความแปรปรวน σ^2 เท่ากัน ค่าเฉลี่ย μ จะเป็นตัวกำหนดจุดกึ่งกลางของเส้นกราฟดังรูป 2.3

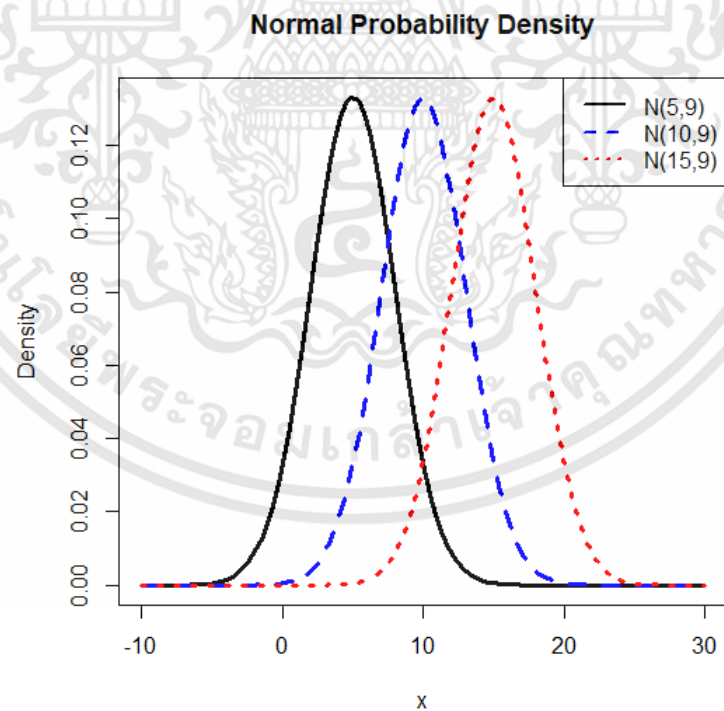


รูปที่ 2.1 แสดงพื้นที่ของการแจกแจงปกติเมื่อค่า x ห่างจากค่าเฉลี่ยเท่ากับ μ เป็นระยะหนึ่งเท่า สองเท่า และสามเท่าของค่าเบี่ยงเบนมาตรฐาน σ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.2 การแจกแจงความหนาแน่นความน่าจะเป็นปกติ ที่มีค่าพารามิเตอร์ (μ, σ^2) เป็น $(10,4)$ $(10,9)$ และ $(10,18)$



รูปที่ 2.3 การแจกแจงความหนาแน่นความน่าจะเป็นปกติ ที่มีค่าพารามิเตอร์ (μ, σ^2) เป็น $(5,9)$ $(10,9)$ และ $(15,9)$ เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อใช้ในการเรียนการสอนเท่านั้น ไม่สามารถนำออกจำหน่ายหรือทำซ้ำโดยไม่ได้รับอนุญาตจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.2 การแจกแจงปรกติปลอมปน (Contaminated Normal Distribution)

การแจกแจงปรกติปลอมปนได้รับการศึกษาโดย John Tukey ในค.ศ. 1960 โดยการแจกแจงปรกติปลอมปนเหมาะสำหรับข้อมูลที่มีการแจกแจงปรกติแต่มีข้อมูลที่มีค่าผิดปกติ (Outliers) รวมอยู่ด้วย กล่าวคือการแจกแจงปรกติที่มีการปลอมปนของการแจกแจงปรกติ โดยค่าความน่าจะเป็นของการผสมระหว่างการแจกแจงปรกติและของการแจกแจงปรกติที่ปลอมปนเป็น p' และ $1-p'$ ตามลำดับ เมื่อ $0 \leq p' \leq 1$

นิยามที่ 2.2 ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงปรกติปลอมปน ด้วยพารามิเตอร์ μ, σ^2 และ c แล้วฟังก์ชันความหนาแน่นความน่าจะเป็น คือ

$$f(x) = \begin{cases} (1-p')N(\mu, \sigma^2) + p'N(\mu, c^2 \sigma^2) & ; -\infty < x < \infty, -\infty < \mu < \infty, \sigma^2 > 0 \\ 0 & ; x \text{ เป็นค่าอื่น ๆ} \end{cases}$$

ซึ่งหมายความว่าค่า x มาจากการแจกแจงแบบ $N(0, \sigma^2)$ ด้วยความน่าจะเป็น $1-p'$ และจากการแจกแจงแบบ $N(0, c^2 \sigma^2)$ ด้วยความน่าจะเป็น p'

เมื่อ $N(0, \sigma^2)$ เมื่อค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ σ^2 โดย σ^2 เป็นพารามิเตอร์ที่กำหนดความแปรปรวนของความคลาดเคลื่อน

เมื่อ p' และ c เป็นค่าคงที่ (Fixed Constant) ที่กำหนดค่าความน่าจะเป็นการปลอมปนและสเกลแฟคเตอร์ (Scale Factor) ตามลำดับ

ทฤษฎีที่ 2.2 ถ้าตัวแปรสุ่ม X มีการแจกแจงปรกติปลอมปน แล้วตัวแปรสุ่ม X จะมีค่าเฉลี่ยและความแปรปรวน ดังนี้

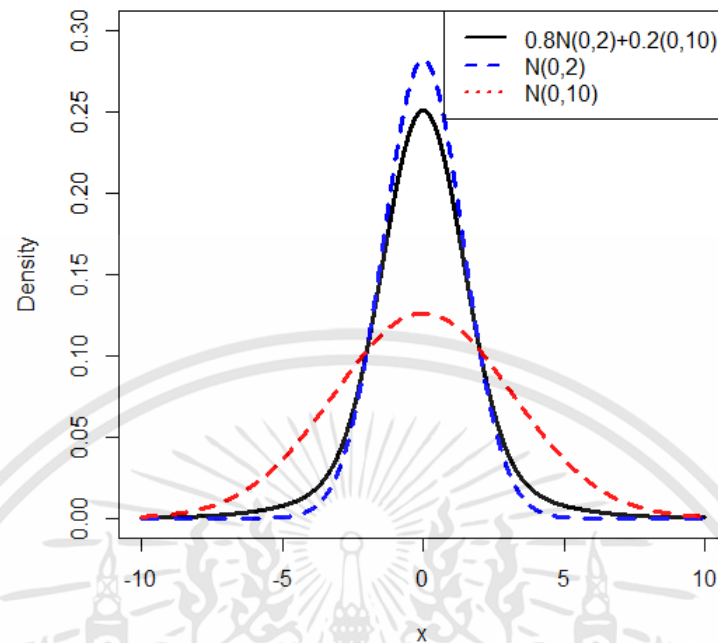
$$\begin{aligned} E(X) &= 0 \\ \text{Var}(X) &= (1-p')\sigma^2 + p'c^2 \sigma^2 \end{aligned}$$

คุณสมบัติของโค้งการแจกแจงปรกติปลอมปน

การแจกแจงปรกติปลอมปนที่ผสมระหว่างการแจกแจง 2 การแจกแจง คือ การแจกแจงปรกติและการแจกแจงปรกติ โดยเส้นโค้งการแจกแจงปรกติปลอมปนจะมีหางยาวกว่าการแจกแจงปรกติ แสดงดังรูปที่ 2.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Contaminated Normal Probability Density



รูปที่ 2.4 การแจกแจงความหนาแน่นความน่าจะเป็นปรกติปลอมปนเป็น $0.8 * N(0,2) + 0.2 * N(0,10)$

2.4.3 การแจกแจงไวบูล (Weibull Distribution)

การแจกแจงไวบูลมักใช้ในการจำลอง เพื่อหาอายุการใช้งานของระบบต่างๆ จนกระทั่งระบบใช้งานไม่ได้ (Wagner. *et al.*, 2010) จึงเหมาะสำหรับข้อมูลที่ต้องมีการทำลายหน่วยตัวอย่าง และลักษณะการแจกแจงไวบูลนั้นมีทั้งสมมาตร เบ้ซ้าย และเบ้ขวา

นิยามที่ 2.3 ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงไวบูล ด้วยพารามิเตอร์ α และ β แล้ว ฟังก์ชันความหนาแน่นความน่าจะเป็น คือ

$$f(x) = \begin{cases} \left(\frac{\beta}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}} & ; 0 < x < \infty, \alpha > 0, \beta > 0 \\ 0 & ; x \text{ เป็นค่าอื่น ๆ} \end{cases}$$

หรือเขียนแทนด้วยสัญลักษณ์ $X \sim W(\alpha, \beta)$

โดยที่

α เป็นพารามิเตอร์มาตราส่วน หรือพารามิเตอร์บอกอัตราส่วน (Scale parameter) ของ

การแจกแจง

β เป็นพารามิเตอร์สัดส่วน หรือพารามิเตอร์บอกรูปร่าง (Shape parameter) ของ

การแจกแจง

ทฤษฎีที่ 2.3 ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงไวบูล ด้วยพารามิเตอร์ α และ β แล้วตัวแปรสุ่ม X จะมีค่าเฉลี่ยและความแปรปรวน ดังนี้

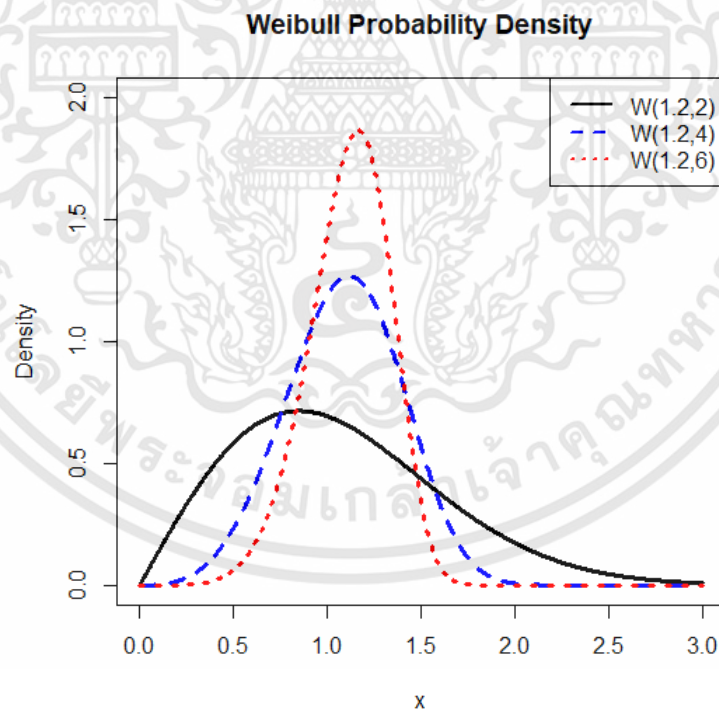
$$E(X) = \alpha \Gamma\left(1 + \frac{1}{\beta}\right)$$

$$Var(X) = \alpha^2 \Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma\left(1 + \frac{1}{\beta}\right)^2$$

คุณสมบัติของโค้งการแจกแจงไวบูล

1. ลักษณะเส้นโค้งของการแจกแจงไวบูลขึ้นอยู่กับค่าพารามิเตอร์ α และ β (นพรัตน์, 2542) โดยที่ α เป็นพารามิเตอร์บอกตำแหน่ง และ β เป็นพารามิเตอร์บอกลักษณะของเส้นโค้ง
2. ค่าพารามิเตอร์ β เปลี่ยนค่าไปจะมีลักษณะเส้นโค้งเป็นดังนี้
 - เมื่อ $3 \leq \beta \leq 4$ การแจกแจงไวบูลเป็นแบบสมมาตร
 - เมื่อ $\beta > 4$ การแจกแจงไวบูลเป็นแบบเบ้ซ้าย
 - เมื่อ $0 < \beta < 3$ การแจกแจงไวบูลเป็นแบบเบ้ขวา

แสดงดังรูป 2.5



รูปที่ 2.5 การแจกแจงความหนาแน่นความน่าจะเป็นไวบูลที่มีค่าพารามิเตอร์ $W(\alpha, \beta)$ เป็น $W(1,2,2)$ $W(1,2,4)$ และ $W(1,2,6)$

2.5 งานวิจัยที่เกี่ยวข้อง

ณัฐพร (2552) ได้เปรียบเทียบวิธีการคัดเลือกตัวแปรที่เหมาะสมสำหรับตัวแบบการถดถอยเชิงเส้น โดยทำการเปรียบเทียบผลที่ได้จากวิธีการคัดเลือกตัวแปรและการประมาณค่าสัมประสิทธิ์การถดถอยที่ได้จากวิธีการที่ต่างกัน โดยเปรียบเทียบวิธีการถดถอยริดจ์ วิธีการถดถอยริดจ์แบบทางเลือก วิธีการถดถอยริดจ์แบบมีข้อจำกัด พบว่าวิธีการถดถอยริดจ์ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด

ทิฆัมพร และนัท (2557) ได้เปรียบเทียบวิธีการคัดเลือกตัวแปรที่เหมาะสมสำหรับตัวแบบการถดถอยเชิงเส้น โดยทำการเปรียบเทียบผลที่ได้จากวิธีการคัดเลือกตัวแปรและการประมาณค่าสัมประสิทธิ์การถดถอยที่ได้จากวิธีการที่ต่างกัน ซึ่งในงานวิจัยนี้จะเปรียบเทียบผลที่ได้จากวิธีเพิ่มตัวแปรอิสระแบบขั้นตอน วิธีการถดถอยลาสโซ่ วิธีการถดถอยลาสโซ่แบบปรับปรุง และวิธีการถดถอยอีลาสติคเน็ต โดยมีการจำลองข้อมูลให้มีขอบเขตที่ต่างกัน ดังนี้ คือ ขนาดตัวอย่างเท่ากับ 40 120 240 และมีอัตราส่วนของจำนวนตัวแปรต่อขนาดตัวอย่างเป็น 0.3 และ 0.7 จำนวนของค่าสัมประสิทธิ์เริ่มต้นที่มีค่าเป็นศูนย์คิดเป็นร้อยละ 10 50 และ 90 ของจำนวนตัวแปรโดยเกณฑ์ที่ใช้ในการวัดประสิทธิภาพของผลที่ได้จากการประมาณค่าของแต่ละวิธี ได้แก่ อัตราความผิดพลาดในการตรวจจับเชิงบวก (FPR) อัตราความผิดพลาดในการตรวจจับเชิงลบ (FNR) และค่าคลาดเคลื่อนการทำนาย (PE) ผลการศึกษาพบว่าอัตราความผิดพลาดในการตรวจจับเชิงบวก (FPR) และค่าคลาดเคลื่อนการทำนาย (PE) นั้นให้ผลไปในทิศทางเดียวกันคือวิธีการถดถอยลาสโซ่แบบปรับปรุงนั้นมีประสิทธิภาพที่เหมาะสมกับข้อมูลที่มีขนาดเล็ก และมีค่าสัมประสิทธิ์บางตัวเป็นศูนย์มากที่สุด

จุฑาทิพย์ (2559) ได้นำเสนอวิธีการเลือกพารามิเตอร์ปรับสำหรับวิธีการถดถอยลาสโซ่ โดยใช้การตรวจสอบข้อบังคับเบื้องต้นของการวิเคราะห์การถดถอยและเปรียบเทียบผลที่ได้กับการเลือกพารามิเตอร์ปรับจากสองวิธีที่ใช้กันอย่างแพร่หลายสำหรับการถดถอยลาสโซ่ ได้แก่ วิธีการตรวจสอบไขว้และวิธีการใช้เกณฑ์ข้อสนเทศของเบส์ ผลการศึกษาจากการจำลองข้อมูลพบว่าวิธีการตรวจสอบข้อบังคับเบื้องต้นของการวิเคราะห์การถดถอยให้อัตราความผิดพลาดในการตรวจจับเชิงบวกต่ำที่สุด วิธีการตรวจสอบไขว้ให้อัตราความผิดพลาดในการตรวจจับเชิงลบต่ำกว่าอีกสองวิธีนี้วิธีการตรวจสอบข้อบังคับเบื้องต้นของการวิเคราะห์การถดถอยและวิธีการตรวจสอบไขว้ไม่มีวิธีใดวิธีหนึ่งที่เหมาะสมกว่าอย่างเด่นชัดกว่ากัน เมื่อพิจารณาจากค่าคลาดเคลื่อนของการพยากรณ์ และสัมประสิทธิ์การถดถอย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชุตติกาญจน์ และสุปราณี (2561) ได้เปรียบเทียบประสิทธิภาพการวิเคราะห์สัมประสิทธิ์การถดถอยบัวซง ด้วยวิธีการวิเคราะห์การถดถอยริดจ์ วิธีการถดถอยลาสโซ่ และวิธีการถดถอยลาสโซ่แบบปรับปรุง ในกรณีข้อมูลมีมิติสูง และตัวแปรอิสระมีความสัมพันธ์กันค่อนข้างสูง ภายใต้สถานการณ์ต่างๆ โดยกำหนดให้พิจารณาตัวแปรอิสระเป็น 2 ลักษณะ หลังจากนั้นทำการจำลองข้อมูลด้วยโปรแกรม R และทำซ้ำ 1,000 ครั้ง เมื่อพิจารณาประสิทธิภาพของการวิเคราะห์ในแต่ละวิธีจากความถูกต้องในการพยากรณ์ พบว่าวิธีการถดถอยลาสโซ่แบบปรับปรุง จะให้ค่า PMSE ต่ำที่สุด เมื่อความสัมพันธ์ของตัวแปรอิสระเพิ่มสูง วิธีการวิเคราะห์การถดถอยริดจ์ จะให้ค่า PMSE ต่ำลง และพิจารณาความผิดพลาดในการคัดเลือกตัวแปรอิสระเข้าสู่ตัวแบบ พบว่า วิธีการถดถอยลาสโซ่เกิดความผิดพลาดในการคัดเลือกตัวแปรอิสระแบบที่ 1 น้อยกว่า วิธีการถดถอยลาสโซ่แบบปรับปรุง แต่เกิดความผิดพลาดในการคัดเลือกตัวแปรอิสระแบบที่ 2 มากกว่าวิธีการถดถอยลาสโซ่แบบปรับปรุง

ปณัช และคณะ (2561) ได้ศึกษาวิธีการประมาณค่าพารามิเตอร์และคัดเลือกตัวแปรในการวิเคราะห์การถดถอยเชิงเส้นกรณีที่มีข้อมูลมิติสูง โดยใช้ขั้นตอนวิธีเชิงพันธุกรรมและนำผลของวิธีที่เสนอไปเปรียบเทียบกับวิธีที่รู้จักกันอย่างแพร่หลาย 3 วิธี ได้แก่ วิธีลาสโซ่ วิธีอีลาสติคเน็ต และวิธีการถดถอยแบบขั้นตอนโดยใช้วิธีการจำลอง เกณฑ์ที่ใช้ในการพิจารณาเปรียบเทียบวิธีที่ศึกษา คือ ร้อยละของการคัดเลือกตัวแปรอิสระได้ถูกต้อง ร้อยละของการคัดเลือกตัวแปรอิสระมากเกินไป ร้อยละของการคัดเลือกตัวแปรอิสระน้อยเกินไป และร้อยละของการคัดเลือกตัวแปรอิสระไม่ถูกต้อง รวมทั้งค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของสมการถดถอย และความถูกต้องของค่าประมาณพารามิเตอร์ของตัวแบบ ผลการศึกษาสรุปได้ว่าขั้นตอนวิธีเชิงพันธุกรรมสามารถประมาณค่าพารามิเตอร์และคัดเลือกตัวแปรอิสระได้ดีที่สุดเมื่อเทียบกับ 3 วิธีดังกล่าวเกือบทุกกรณี

Algamal and Lee (2015) การจำแนกมะเร็งและการคัดเลือกยีนในข้อมูลมิติสูงเป็นหัวข้อวิจัยที่เป็นที่นิยมในด้านพันธุศาสตร์และชีววิทยาระดับโมเลกุล เมื่อเร็ว ๆ นี้การถดถอยโลจิสติกส์วิธีอีลาสติคเน็ต และวิธีอีลาสติคเน็ตแบบปรับ ได้ถูกนำมาประยุกต์ใช้ทั้งในการจำแนกประเภทมะเร็งในมิติสูงเพื่อใช้ในการประมาณค่าสัมประสิทธิ์ของยีน และการคัดเลือกยีนไปพร้อมกัน โดยในงานวิจัยนี้ ศึกษาเปรียบเทียบการถดถอยโลจิสติกส์วิธีอีลาสติคเน็ต และวิธีอีลาสติคเน็ตแบบปรับเทียบกับวิธีใหม่ที่นำเสนอ คือ วิธีการปรับของการถดถอยอีลาสติคเน็ตแบบปรับ (Adjusted adaptive regularized logistic regression : AAElastic) โดยพบว่าวิธีการที่นำเสนอให้ผลลัพธ์ดีกว่าการถดถอยโลจิสติกส์วิธีอีลาสติคเน็ต และวิธีอีลาสติคเน็ตแบบปรับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

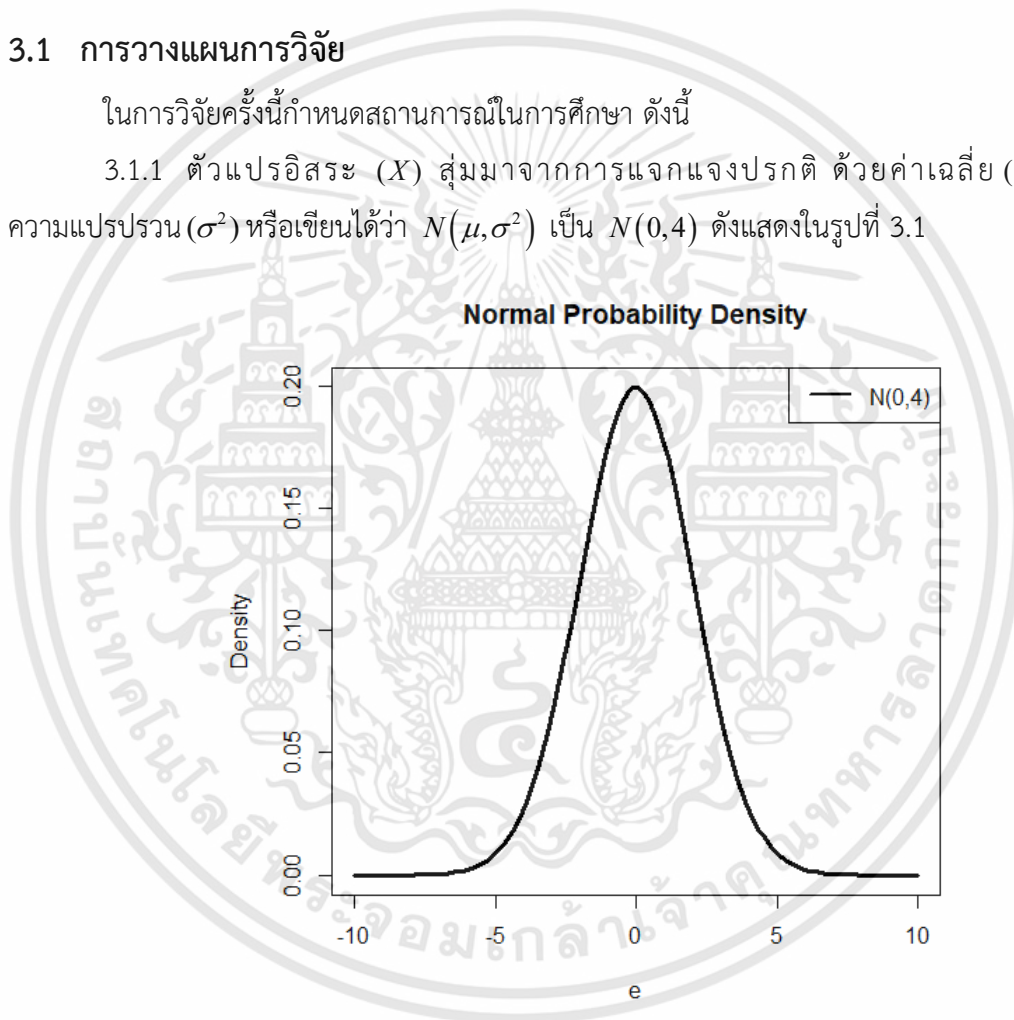
วิธีการดำเนินงานวิจัย

การวิจัยนี้เป็นการวิจัยเชิงจำลองเพื่อศึกษาและเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง โดยมีการวางแผนการวิจัยและวิธีการดำเนินการวิจัย ดังต่อไปนี้

3.1 การวางแผนการวิจัย

ในการวิจัยครั้งนี้กำหนดสถานการณ์ในการศึกษา ดังนี้

3.1.1 ตัวแปรอิสระ (X) สุ่มมาจากการแจกแจงปกติ ด้วยค่าเฉลี่ย (μ) และความแปรปรวน (σ^2) หรือเขียนได้ว่า $N(\mu, \sigma^2)$ เป็น $N(0, 4)$ ดังแสดงในรูปที่ 3.1

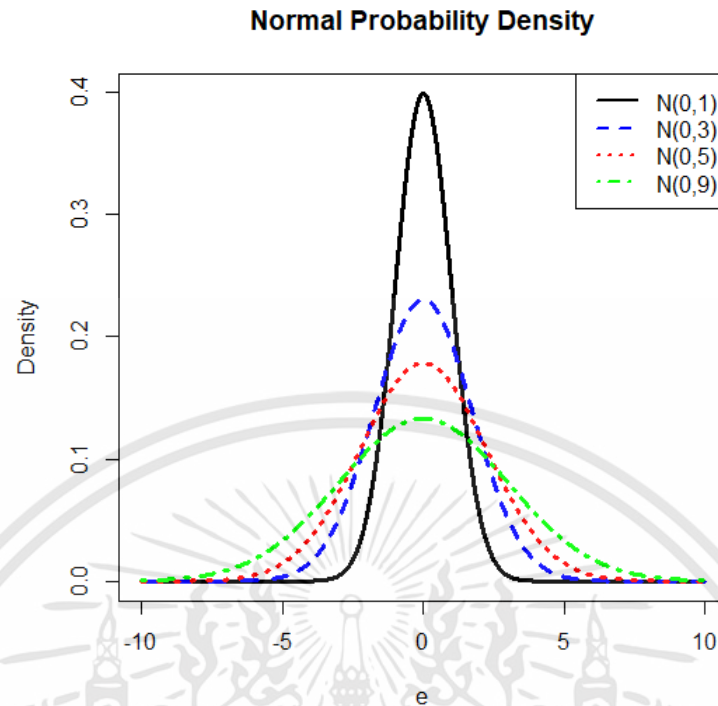


รูปที่ 3.1 การแจกแจงความหนาแน่นความน่าจะเป็นปกติของ $N(\mu, \sigma^2)$ เป็น $N(0, 4)$

3.1.2 ค่าความคลาดเคลื่อน (ε) สุ่มมาจาก 3 การแจกแจงของข้อมูล ดังนี้

3.1.2.1 การแจกแจงปกติของ $N(\mu, \sigma^2)$ เป็น $N(0, 1)$ $N(0, 3)$ $N(0, 5)$ และ $N(0, 9)$ ดังแสดงในรูปที่ 3.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

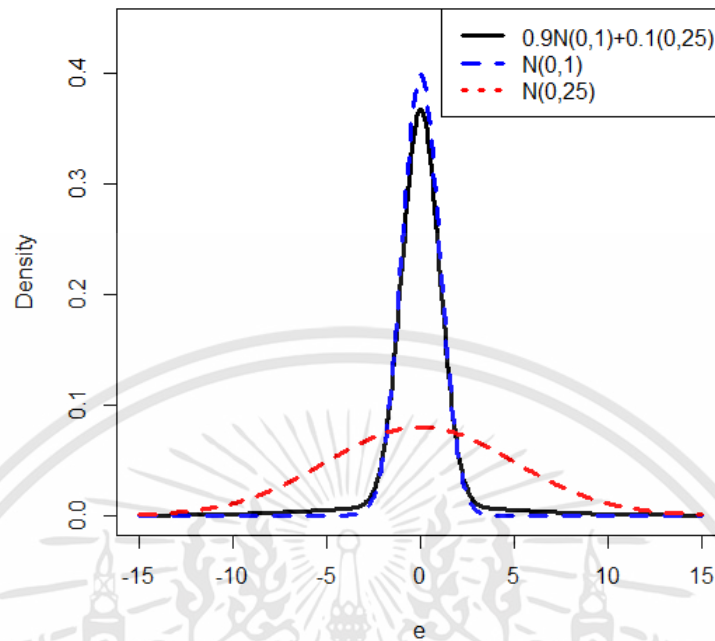


รูปที่ 3.2 การแจกแจงความหนาแน่นความน่าจะเป็นปกติของ $N(\mu, \sigma^2)$ เป็น $N(0,1)$ $N(0,3)$ $N(0,5)$ และ $N(0,9)$

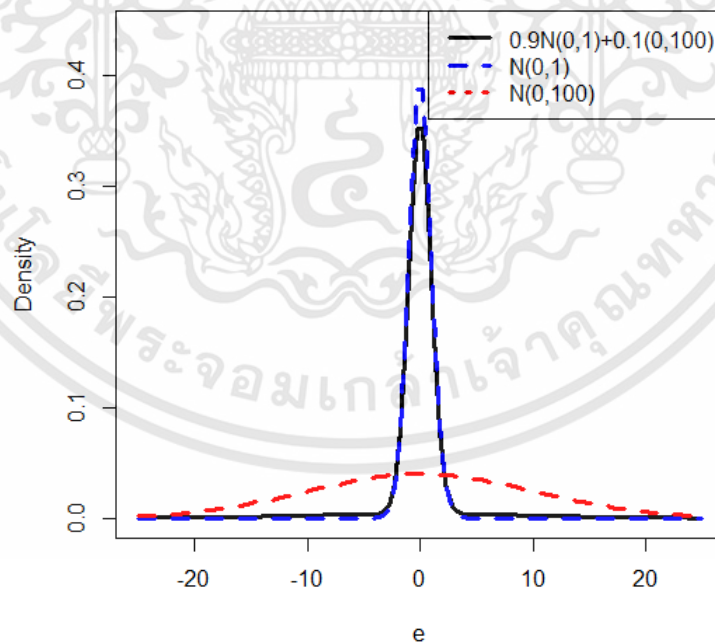
3.1.2.2 การแจกแจงปกติป lom บนระหว่าง การแจกแจงปกติกับการแจกแจงปกติ ในที่นี้ศึกษาการแจกแจงปกติ $N(\mu, \sigma)$ เป็น $N(0,1)$ และค่าความน่าจะเป็นที่ข้อมูลมีการป lom บน การแจกแจงปกติอื่นเป็น 0.1 ($p' = 0.1$) โดยที่ c คือ สเกลแฟกเตอร์ (Scale Factor) ศึกษาที่ $c = 5$ สำหรับข้อมูลที่มีค่าผิดปกติในระดับไม่รุนแรง และ $c = 10$ สำหรับข้อมูลที่มีค่าผิดปกติในระดับรุนแรง ดังแสดงในรูปที่ 3.3 และ 3.4 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Contaminated Normal Probability Density

รูปที่ 3.3 การแจกแจงความหนาแน่นความน่าจะเป็นปรกติปลอมปนเป็น $0.9 * N(0,1) + 0.1 * N(0,25)$

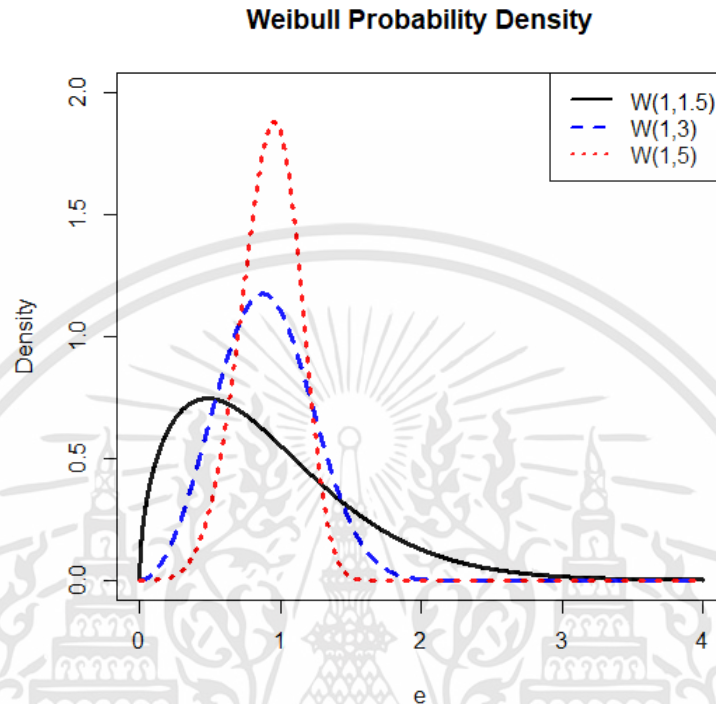
Contaminated Normal Probability Density



รูปที่ 3.4 การแจกแจงความหนาแน่นความน่าจะเป็นปรกติปลอมปนเป็น $0.9 * N(0,1) + 0.1 * N(0,100)$

เอกสารนี้เป็นเอกสารที่เผยแพร่ในนามของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2.3 การแจกแจงไวบูลล์ด้วยพารามิเตอร์บอกอัตราส่วน (α) และพารามิเตอร์บอกรูปร่าง (β) หรือเขียนได้ว่า $W(\alpha, \beta)$ ดังนี้ $W(1,1.5)$ $W(1,3)$ และ $W(1,5)$ ดังแสดงในรูปที่ 3.5



รูปที่ 3.5 การแจกแจงความหนาแน่นความน่าจะเป็นไวบูลล์ ของ $W(\alpha, \beta)$ ดังนี้ $W(1,1.5)$ $W(1,3)$ และ $W(1,5)$

3.1.3 ศึกษาในกรณีที่จำนวนตัวแปรอิสระ (p) ในตัวแบบมากกว่าขนาดตัวอย่าง (n) ดังนี้

3.1.3.1 กำหนดขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร

3.1.3.2 กำหนดขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร

3.1.3.3 กำหนดขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร

3.1.3.4 กำหนดขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถสรุปเป็นตาราง ดังนี้

ตารางที่ 3.1 แสดงขนาดตัวอย่าง และจำนวนสัมประสิทธิ์การถดถอย

ขนาดตัวอย่าง (n)	สัมประสิทธิ์การถดถอยที่ใช้ (β_j) ; $j = 1, 2, \dots, p$
5 10 และ 15	$\beta_0, \beta_1, \beta_2, \dots, \beta_{16}$
10 15 และ 20	$\beta_0, \beta_1, \beta_2, \dots, \beta_{30}$
20 30 และ 40	$\beta_0, \beta_1, \beta_2, \dots, \beta_{50}$
60 70 และ 80	$\beta_0, \beta_1, \beta_2, \dots, \beta_{100}$

3.1.4 ค่าสัมประสิทธิ์การถดถอยพหุคูณ (β_j) กำหนด 3 กรณี

3.1.4.1 ค่าสัมประสิทธิ์การถดถอยทุกตัวมีค่าเป็น 1 ดังนี้

$$\beta_0 = \beta_1 = \beta_2 = \dots = \beta_p = 1 \quad ; \quad j = 0, 1, \dots, p$$

3.1.4.2 ค่าสัมประสิทธิ์การถดถอยมีค่าเป็น 0.5 จำนวนครึ่งหนึ่งของจำนวนตัวแปรอิสระ และที่เหลือมีค่าสัมประสิทธิ์การถดถอยเป็น 1 ดังตารางที่ 3.2

ตารางที่ 3.2 แสดงจำนวนตัวแปรอิสระ และค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

จำนวนตัวแปรอิสระ (p)	ค่าสัมประสิทธิ์การถดถอย
16	$\beta_0 = 1, \beta_1 = \beta_2 = \beta_3 = \dots = \beta_8 = 0.5, \beta_9 = \beta_{10} = \beta_{11} = \dots = \beta_{16} = 1$
30	$\beta_0 = 1, \beta_1 = \beta_2 = \beta_3 = \dots = \beta_{15} = 0.5, \beta_{16} = \beta_{17} = \beta_{18} = \dots = \beta_{30} = 1$
50	$\beta_0 = 1, \beta_1 = \beta_2 = \beta_3 = \dots = \beta_{25} = 0.5, \beta_{26} = \beta_{27} = \beta_{28} = \dots = \beta_{50} = 1$
100	$\beta_0 = 1, \beta_1 = \beta_2 = \beta_3 = \dots = \beta_{50} = 0.5, \beta_{51} = \beta_{52} = \beta_{53} = \dots = \beta_{100} = 1$

3.1.4.3 ค่าสัมประสิทธิ์การถดถอยมีค่าเป็น -1 จำนวนครึ่งหนึ่งของจำนวนตัวแปรอิสระ และที่เหลือมีค่าสัมประสิทธิ์การถดถอยเป็น 1 ดังตารางที่ 3.3

ตารางที่ 3.3 แสดงจำนวนตัวแปรอิสระ และค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

จำนวนตัวแปรอิสระ (p)	ค่าสัมประสิทธิ์การถดถอย
16	$\beta_0 = 1, \beta_1 = \beta_2 = \beta_3 = \dots = \beta_8 = -1, \beta_9 = \beta_{10} = \beta_{11} = \dots = \beta_{16} = 1$
30	$\beta_0 = 1, \beta_1 = \beta_2 = \beta_3 = \dots = \beta_{15} = -1, \beta_{16} = \beta_{17} = \beta_{18} = \dots = \beta_{30} = 1$
50	$\beta_0 = 1, \beta_1 = \beta_2 = \beta_3 = \dots = \beta_{25} = -1, \beta_{26} = \beta_{27} = \beta_{28} = \dots = \beta_{50} = 1$
100	$\beta_0 = 1, \beta_1 = \beta_2 = \beta_3 = \dots = \beta_{50} = -1, \beta_{51} = \beta_{52} = \beta_{53} = \dots = \beta_{100} = 1$

3.1.5 ตัวแปรตามใช้ตัวแบบการถดถอยเชิงเส้นพหุคูณ ดังนี้

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon}$$

3.1.6 โปรแกรมที่ใช้ในการวิจัยครั้งนี้ทั้งหมดเขียนด้วยโปรแกรม R เวอร์ชัน 3.5.2 ซึ่งทำการทดลองซ้ำ 1,000 ครั้งในแต่ละสถานการณ์

3.2 วิธีการดำเนินการวิจัย

ในการวิจัยนี้มีวิธีการดำเนินการวิจัยดังนี้

3.2.1 จำลองข้อมูลของตัวแปรอิสระ (X) ตัวแปรตาม (y) ค่าความคลาดเคลื่อน (ε) และจำลองขนาดตัวอย่าง (n) ที่ใช้ในการวิจัยด้วยโปรแกรมอาร์ (R) เป็นไปตามขอบเขตของการวิจัย

3.2.2 หาค่าพารามิเตอร์ปรับแต่ง (λ) ของการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูงทั้ง 5 วิธี ได้แก่ วิธีการถดถอยริดจ์ วิธีการถดถอยลาสโซ วิธีการถดถอยอิลาสติกเน็ต วิธีการถดถอยลาสโซแบบปรับปรุง และวิธีการถดถอยอิลาสติกเน็ตแบบปรับปรุง ตามการจำลองข้อมูล

3.2.3 นำข้อมูลที่จำลองขึ้น และค่าพารามิเตอร์ปรับแต่ง (λ) ที่ได้จาก 3.2.1 และ 3.2.2 มาทำการคัดกรองตัวแปรอิสระ และหาสมการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูงในแต่ละวิธี และนำสมการถดถอยมาประมาณค่าหาพยากรณ์ (\hat{y})

3.2.4 คำนวณหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของแต่ละวิธีการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง ทำการทดลองซ้ำ 1,000 รอบ

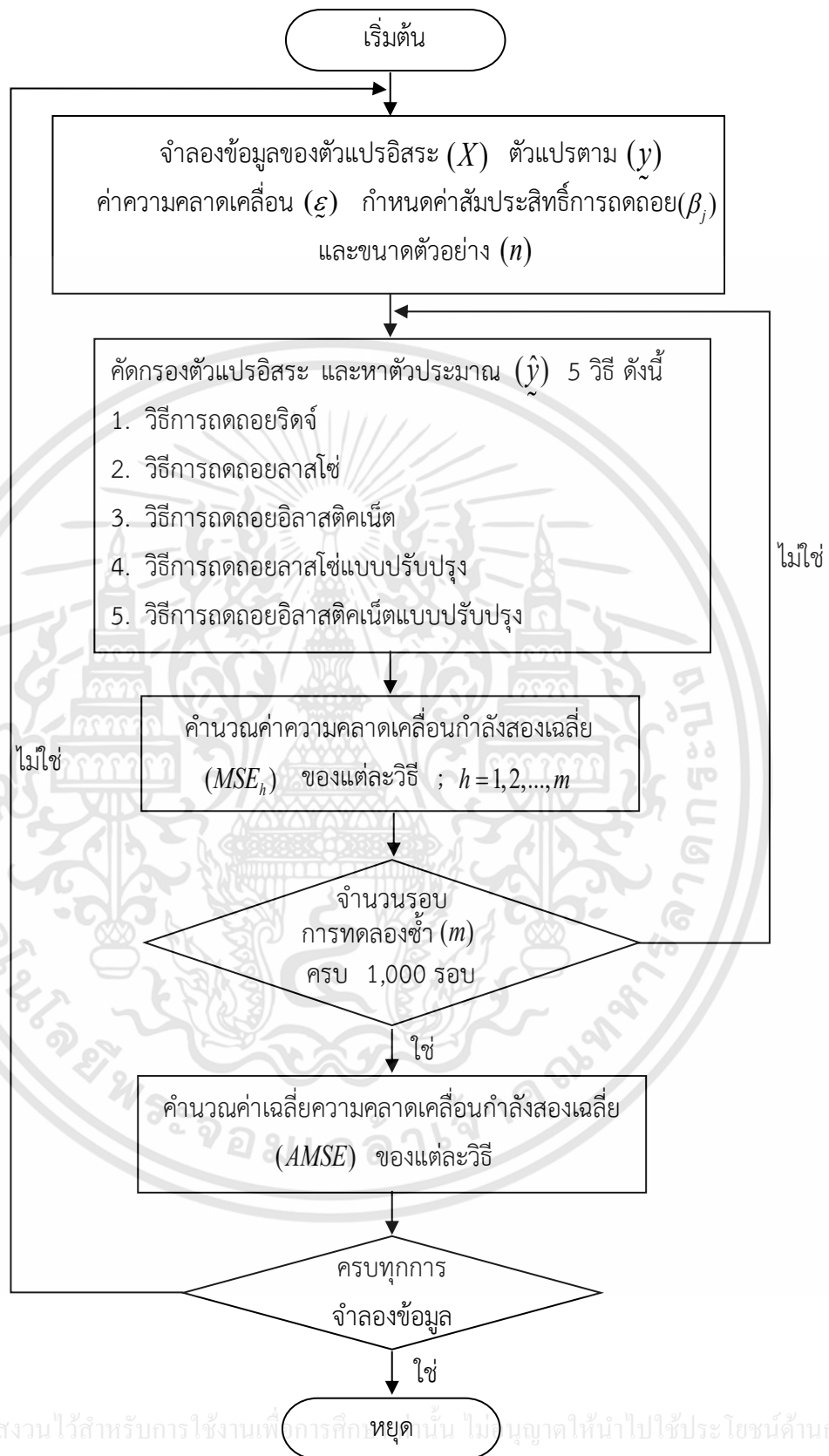
3.2.5 เปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ของแต่ละวิธีการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง โดยค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ที่ต่ำที่สุด เป็นวิธีการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูงที่เหมาะสมสำหรับสถานการณ์นั้นๆ

3.2.6 วิเคราะห์ และสรุปผลทั้ง 5 วิธีการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง

3.3 ขั้นตอนของโปรแกรมที่ใช้ในงานวิจัย

วิธีการดำเนินการวิจัยและการประมวลผลข้อมูลสามารถอธิบายเป็นขั้นตอนได้ดังรูปที่ 3.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานที่การศึกษาหยุดนั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ผู้อื่นนำเอกสารฉบับนี้ไปเผยแพร่หรือใช้ซ้ำโดยไม่ได้รับอนุญาต
รูปที่ 3.6 แผนผังแสดงลำดับวิธีการดำเนินการวิจัย

บทที่ 4

ผลการวิจัยและการอภิปรายผล

การวิจัยนี้เป็นการวิจัยเชิงจำลอง เพื่อศึกษาและเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง 5 วิธี คือ วิธีการถดถอยริดจ์ (Ridge Regression) วิธีการถดถอยลาสโซ (Least absolute shrinkage and selection operator Regression : Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net Regression) วิธีการถดถอยลาสโซแบบปรับปรุง (The adaptive least absolute shrinkage and selection operator Regression) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (The adaptive Elastic net Regression) โดยเกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูง คือ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (Average Mean Squared Error : AMSE) และพิจารณาค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดจะเป็นวิธีการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษภายใต้ข้อมูลที่มีมิติสูงที่ดีที่สุด

โดยกำหนดสัญลักษณ์แทนวิธีการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ ดังนี้

Ridge	หมายถึง	วิธีการถดถอยริดจ์
Lasso	หมายถึง	วิธีการถดถอยลาสโซ
Elastic Net	หมายถึง	วิธีการถดถอยอีลาสติคเน็ต
ALasso	หมายถึง	วิธีการถดถอยลาสโซแบบปรับปรุง
AElastic Net	หมายถึง	วิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง

ผลการวิจัยผู้วิจัยได้ใช้ผลจากการจำลองจากวิธีการดำเนินงานในบทที่ 3 แสดงผลในหัวข้อ 4.1 และนำมาประยุกต์ใช้ข้อมูลจริงแสดงผลในหัวข้อ 4.2 ตามลำดับ

4.1 ผลจากการจำลองข้อมูล

ผลการวิเคราะห์เมื่อค่าความคลาดเคลื่อนมีการแจกแจงปกติ การแจกแจงปกติปลอมปน และการแจกแจงไวบูล โดยค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ของวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.1 เมื่อค่าความคลาดเคลื่อนมีการแจกแจงปรกติ

ผลการวิเคราะห์เมื่อค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงเส้นพหุคูณมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) โดยค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ของวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ดังตารางที่ 4.1 - 4.12

ตารางที่ 4.1 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	5	33.9256	26.9874	22.4558	7.7025	3.3780
	10	37.7946	16.6173	13.5643	6.5365	3.6406
	15	38.9704	7.5894	5.8060	4.1797	2.7336
$N(0,3)$	5	34.9935	29.2523	26.2828	8.3037	3.8562
	10	39.0932	16.4562	13.6789	6.2834	3.5595
	15	40.7725	8.7852	6.2031	3.9689	2.8128
$N(0,5)$	5	36.3529	30.2250	24.5395	8.3435	3.5796
	10	40.5552	19.7997	16.4301	7.1400	4.1159
	15	42.3772	9.3974	7.6466	4.9540	3.5690
$N(0,9)$	5	39.1396	35.3612	29.3117	9.4488	4.1393
	10	43.6695	22.1339	18.5876	8.0446	4.6192
	15	46.2458	12.8152	10.0233	6.0954	4.0198

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.1 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงเส้นพหุคูณที่มีการแจกแจงปรกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) และวิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงปกติเป็น $N(0,3)$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.1 จากภาคผนวก ก



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	10	62.9814	46.5596	38.5948	13.8607	5.2057
	15	64.1520	37.7650	31.6143	10.6831	5.5311
	20	67.2748	24.9838	19.1117	7.5779	4.5425
$N(0,3)$	10	63.6897	47.9835	42.4868	14.2661	5.8937
	15	65.4186	39.0492	31.4786	11.8434	5.9687
	20	67.3045	26.1824	19.5749	7.4505	4.4243
$N(0,5)$	10	63.8888	49.3626	42.8696	14.3071	5.9774
	15	68.7015	40.9951	32.6829	11.3829	5.7850
	20	69.9518	28.8265	21.3332	8.9363	5.4091
$N(0,9)$	10	65.8742	51.6873	45.0598	14.5388	5.9279
	15	69.6923	40.2664	35.6711	12.1898	6.0792
	20	72.2674	30.1931	26.8271	9.9316	5.7014

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.2 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) และวิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อน

กำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงปกติเป็น $N(0,5)$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่

ก.2 จากภาคผนวก ก

ตารางที่ 4.3 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	20	96.2498	70.4524	61.0480	16.5131	7.6294
	30	101.6025	44.2040	34.1240	11.1575	5.6076
	40	106.7399	22.5958	14.0320	7.2192	3.7383
$N(0,3)$	20	96.7310	71.6835	61.7820	16.2152	7.6851
	30	102.7509	44.3214	34.0704	10.4681	6.0766
	40	106.9115	19.9710	14.1245	6.5357	3.2376
$N(0,5)$	20	95.9000	73.1512	60.0148	17.5053	7.4295
	30	101.4797	46.4416	35.5977	10.4087	5.8355
	40	105.9993	22.7376	15.1520	6.4973	3.8568
$N(0,9)$	20	93.8416	70.0297	58.5286	16.8798	7.6230
	30	101.2664	45.8974	34.0036	10.8726	6.0087
	40	105.3065	21.4300	14.2404	6.0439	3.5356

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.3 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.3 จากภาคผนวก ก

ตารางที่ 4.4 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	60	181.1241	86.9848	59.0713	12.8314	7.6452
	70	186.1973	49.0954	35.2171	8.4384	4.8927
	80	190.5646	28.7975	19.8804	6.0043	3.3205
$N(0,3)$	60	181.8536	82.9966	63.4173	12.4916	7.7402
	70	187.1323	57.2230	41.5087	9.1815	5.8639
	80	190.5675	30.9049	18.7061	5.7201	4.0333
$N(0,5)$	60	182.9207	86.4340	63.6522	14.1147	7.8773
	70	187.3913	54.2639	40.7831	8.9504	5.9018
	80	191.9276	33.3979	23.5484	5.9311	3.7920
$N(0,9)$	60	184.7577	93.3526	72.0972	14.2818	8.2960
	70	190.6009	57.6322	37.4456	9.2522	5.4729
	80	194.1427	39.8798	24.6886	6.7811	5.1945

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.4 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.4 จากภาคผนวก ก

ตารางที่ 4.5 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	5	21.5194	17.2703	14.9615	6.1594	2.8052
	10	23.9701	10.1956	8.1738	4.1170	2.6363
	15	24.6762	4.0067	2.9310	2.0876	1.4315
$N(0,3)$	5	22.4466	18.3075	16.1491	6.2116	2.9591
	10	25.2343	10.8501	9.1362	4.2580	2.6647
	15	26.5318	5.6642	4.1384	2.5870	1.8468
$N(0,5)$	5	23.4843	19.7317	16.7650	6.4638	2.9644
	10	26.6087	13.1481	11.0868	4.8021	3.2492
	15	28.1575	6.8305	5.7738	3.5457	2.7074
$N(0,9)$	5	26.8795	23.7903	20.1541	7.2647	3.4981
	10	29.9938	15.1834	13.3462	5.9357	3.7655
	15	31.5710	9.5394	7.9702	4.6638	3.2003

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.5 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) และวิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงปกติเป็น $N(0,1)$ และ $N(0,3)$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.5 จากภาคผนวก ก

ตารางที่ 4.6 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	10	39.5702	26.9029	22.7663	9.2530	4.0476
	15	40.3155	23.0251	19.1610	7.2160	4.1258
	20	42.3726	14.6138	10.6080	4.6851	2.8788
$N(0,3)$	10	40.6452	31.4099	27.5722	10.4279	4.8153
	15	41.7547	25.2455	20.0583	7.8685	4.4981
	20	43.0983	16.6193	12.1137	4.8770	3.1440
$N(0,5)$	10	41.9088	32.2365	27.7592	10.8294	4.7650
	15	43.9359	26.0847	21.6331	8.4743	4.3380
	20	45.0621	17.1135	13.7118	5.3097	3.7226
$N(0,9)$	10	43.4990	33.2617	29.6046	11.1694	4.8689
	15	46.6028	27.9780	23.3665	9.1428	4.8054
	20	47.6592	20.5590	17.7287	6.4897	4.1539

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.6 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) และวิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงปกติเป็น $N(0,3)$ $N(0,5)$ และ $N(0,9)$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.6 จากภาคผนวก ก

ตารางที่ 4.7 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	20	60.1750	40.9192	35.1873	11.4791	5.6885
	30	63.3210	25.3036	19.2580	6.4727	3.8148
	40	67.3502	10.4207	7.0078	3.4994	2.1366
$N(0,3)$	20	60.2346	41.6701	34.0916	11.0273	5.5691
	30	63.9659	25.8535	19.3496	6.0150	3.8089
	40	66.5357	9.8426	7.4309	2.8639	1.7372
$N(0,5)$	20	60.2216	42.5091	34.4437	12.1483	5.6983
	30	63.6227	26.6941	19.6417	6.2711	3.8360
	40	66.2792	11.8723	6.9904	3.2782	2.1236
$N(0,9)$	20	58.2826	40.9299	34.9334	11.2476	5.4554
	30	63.4000	25.0356	17.9936	5.8417	3.7859
	40	66.2365	9.6616	6.6293	2.9602	1.7465

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.7 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.7 จากภาคผนวก ก

ตารางที่ 4.8 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	60	114.4579	45.3313	32.3449	6.5601	4.7585
	70	116.7924	25.2431	16.8827	4.4399	3.0225
	80	119.5957	13.0421	8.8842	2.6155	1.9812
$N(0,3)$	60	114.7146	45.8185	32.7901	7.3613	4.9607
	70	118.3348	30.5285	20.9726	4.8413	3.6450
	80	120.5331	13.2773	9.7171	3.0767	2.2030
$N(0,5)$	60	115.5121	47.5692	36.3351	6.9655	5.7040
	70	118.5523	28.9506	21.2809	5.2594	3.7471
	80	122.0014	16.9092	11.8651	3.7720	2.3793
$N(0,9)$	60	118.2094	52.9125	40.0432	8.0203	5.5698
	70	121.7780	33.9114	23.2094	5.6208	3.9590
	80	124.4003	18.8083	14.4971	3.8158	2.9641

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.8 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.8 จากภาคผนวก ก

ตารางที่ 4.9 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	5	34.3966	29.8505	24.5349	8.5131	4.4944
	10	37.8302	15.0771	12.0927	5.8130	3.2368
	15	39.4942	7.4969	5.4143	3.8412	2.0896
$N(0,3)$	5	35.3487	28.5981	23.5134	8.0302	4.2246
	10	38.2527	17.5447	14.4318	6.6737	3.6617
	15	41.8404	8.8179	6.5501	3.8719	2.7172
$N(0,5)$	5	35.1105	30.5918	23.7662	8.3932	4.1910
	10	39.7464	19.1582	16.2615	7.2466	4.2125
	15	41.8968	10.6856	9.2055	5.2438	3.2277
$N(0,9)$	5	38.5754	33.7197	27.5360	8.8741	4.6403
	10	43.9859	21.9297	19.2255	7.9198	4.9233
	15	47.1443	12.8284	10.9465	6.2145	4.0022

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.9 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) และวิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงปกติเป็น $N(0,1)$ และ $N(0,3)$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.9 จากภาคผนวก ก

ตารางที่ 4.10 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	10	61.0146	45.3961	39.8534	13.1170	6.1158
	15	64.1596	35.4385	29.6715	10.5238	5.3967
	20	66.8414	24.6748	18.4306	7.6167	3.9440
$N(0,3)$	10	62.8183	50.8786	44.7630	14.5273	6.5707
	15	67.5077	39.9062	32.3044	12.1856	5.8357
	20	68.2546	24.5155	19.5706	8.1552	4.1965
$N(0,5)$	10	64.7471	48.7364	43.4044	14.5849	6.4924
	15	68.6399	40.3183	32.0231	11.4036	5.6346
	20	69.6233	28.8106	21.3634	8.4900	4.5905
$N(0,9)$	10	66.2429	52.2426	44.8299	15.3444	6.7946
	15	70.3205	42.5428	35.0709	11.5108	6.5571
	20	71.3362	32.4525	25.9422	9.7669	5.3863

หมายเหตุ **ตัวหนา** หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.10 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.10 จากภาคผนวก ก

ตารางที่ 4.11 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	20	94.7267	68.6009	57.6399	16.2245	7.1447
	30	103.6688	46.1935	38.3521	11.3491	6.1649
	40	106.6999	21.2878	14.1144	5.5407	3.3260
$N(0,3)$	20	94.6125	68.0322	58.9121	16.8326	7.5350
	30	101.4027	45.0964	33.8733	11.3446	5.3586
	40	105.6495	19.2818	13.3955	5.8135	2.8143
$N(0,5)$	20	95.4090	72.6471	62.8489	18.6430	8.3044
	30	101.5639	45.5643	34.2003	10.4875	5.2571
	40	107.1784	21.2385	14.8600	6.0828	2.6813
$N(0,9)$	20	95.0808	68.7847	58.1340	17.5386	7.3458
	30	100.4525	45.1452	33.1612	10.1845	5.1647
	40	104.6581	22.2117	14.2403	6.2367	2.9593

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.11 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.11 จากภาคผนวก ก

ตารางที่ 4.12 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	60	181.0038	86.7783	60.2618	13.1317	6.5538
	70	186.1872	53.4883	37.5971	8.5727	4.3645
	80	188.9433	28.4908	19.3360	5.3754	3.3370
$N(0,3)$	60	181.0752	86.8332	64.6513	12.1141	6.9049
	70	186.5447	58.5442	40.4555	9.3754	5.4826
	80	191.2554	34.8621	22.8852	6.4373	3.4324
$N(0,5)$	60	181.2932	83.2576	63.3583	11.1774	6.7893
	70	186.7820	59.8439	39.4054	11.2617	5.2881
	80	191.6222	32.7752	23.5640	6.8372	3.5858
$N(0,9)$	60	184.1542	89.0276	64.8539	13.7845	7.4268
	70	189.7141	56.6474	41.0700	9.2701	5.0899
	80	194.7137	37.6008	26.4167	7.0393	3.7433

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.12 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติ

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.12 จากภาคผนวก ก

4.1.2 เมื่อค่าความคลาดเคลื่อนมีการแจกแจงปรกติปloomปน

ผลการวิเคราะห์เมื่อค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงเส้นพหุคูณมีการแจกแจงปรกติปloomปน โดยค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ของวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ดังตารางที่ 4.13 - 4.24

ตารางที่ 4.13 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปloomปนด้วยเปอร์เซ็นต์การปloomปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปloomปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	5	35.0811	31.0671	24.7741	8.4406	3.6405
	10	38.9817	16.7417	13.8955	6.2428	3.6534
	15	41.0874	9.1498	6.9239	4.0533	3.1888
10%, $c^2 = 100$	5	38.8782	33.3775	27.6338	8.7207	3.9104
	10	44.2320	22.5650	19.9547	8.0482	4.7798
	15	47.3983	14.2904	11.8839	6.6052	4.7059

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.13 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปloomปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) และวิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และไม่มีผลต่อค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ของวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.13 จากภาคผนวก ก

ตารางที่ 4.14 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	10	64.8490	50.7335	44.1421	13.8509	5.7942
	15	66.3236	39.4853	31.7682	11.3212	5.7866
	20	68.9055	26.6152	19.9842	8.0460	4.9129
10%, $c^2 = 100$	10	69.2137	55.7214	48.2564	16.0951	6.2673
	15	72.0872	47.5633	40.8638	14.0115	6.8587
	20	73.9702	34.1579	27.0825	9.4459	5.7963

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.14 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) และวิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนเป็น $c^2 = 25$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.14 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.15 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	20	96.8406	73.9534	64.6146	17.6593	7.7842
	30	102.9568	44.0060	34.0947	10.2773	6.0346
	40	105.9625	23.7546	17.1172	7.0021	3.8908
10%, $c^2 = 100$	20	96.2471	69.3985	59.6544	17.5066	7.6667
	30	102.8508	46.1801	33.6662	10.8649	5.5765
	40	106.1674	22.1053	14.9190	6.2030	3.8285

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.15 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.15 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.16 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	60	182.8232	83.6694	61.6271	12.4524	8.6602
	70	187.1082	52.7917	38.6351	8.7435	5.6113
	80	190.7047	35.7055	22.9899	6.7027	4.9833
10%, $c^2 = 100$	60	186.8187	96.7751	71.6475	12.9031	9.4370
	70	191.2204	59.5943	42.4976	11.1551	5.9500
	80	195.7240	42.7192	26.4030	7.0348	5.0815

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.16 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.16 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.17 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	5	22.2501	19.5871	16.0087	6.4943	2.9569
	10	25.3211	11.2070	9.0236	4.4986	2.7265
	15	26.7258	6.2952	4.7824	2.6846	2.1236
10%, $c^2 = 100$	5	26.5972	23.4486	19.7274	7.0881	3.3314
	10	30.3202	16.2212	14.1291	6.0887	3.7643
	15	32.8129	11.4474	9.8532	5.1674	3.8101

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.17 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) และวิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนเป็น $c^2 = 25$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.17 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.18 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	10	41.2614	32.2863	27.9464	10.3775	4.6184
	15	42.4920	24.8385	20.0199	7.8413	4.2944
	20	43.9474	15.4943	12.1071	4.6223	3.2696
10%, $c^2 = 100$	10	46.0461	35.3223	30.6588	11.1069	4.8535
	15	47.9161	31.2128	25.4616	9.5652	5.3041
	20	49.6379	23.6486	19.4169	7.2761	4.7288

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.18 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) และวิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนเป็น $c^2 = 25$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.18 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.19 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	20	60.2063	43.1931	37.5632	11.1580	5.6905
	30	64.2455	24.9933	18.1631	6.6705	4.0206
	40	66.5380	11.6128	8.3492	3.3754	2.2798
10%, $c^2 = 100$	20	59.4076	41.9282	35.0900	11.7494	5.5206
	30	64.1454	27.8791	19.6481	6.4772	3.9157
	40	66.4075	9.8383	7.0821	2.9188	2.0209

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.19 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.19 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.20 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	60	115.5315	50.6715	32.9484	6.8301	5.7266
	70	118.5590	28.9540	20.4823	5.2147	3.4380
	80	120.9201	17.8528	11.5104	3.4195	2.5329
10%, $c^2 = 100$	60	119.3722	56.3201	41.1536	8.7802	5.9003
	70	122.4217	32.8938	24.2851	5.9692	3.8191
	80	125.6965	21.9633	15.5407	4.4484	3.3110

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.20 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.20 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.21 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	5	33.9504	29.5246	24.1289	8.1533	4.2364
	10	39.1042	17.7766	15.2729	6.5910	3.9496
	15	41.2070	9.0979	7.1697	4.7477	2.8545
10%, $c^2 = 100$	5	39.2709	35.1249	29.7636	8.7177	4.6607
	10	44.2813	23.1757	19.9465	8.1696	4.9573
	15	47.0885	15.8857	12.8164	6.7852	4.6868

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.21 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) และวิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนเป็น $c^2 = 25$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบดังตารางที่ ก.21 จากภาคผนวก ก

ตารางที่ 4.22 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	10	62.8524	47.0779	40.8570	14.2997	6.0675
	15	65.5771	39.7903	35.3293	12.4518	5.9365
	20	68.1816	27.2172	20.8779	7.5342	4.4423
10%, $c^2 = 100$	10	68.5184	52.2478	45.2451	15.0798	6.6544
	15	71.2850	48.1686	39.3864	13.0875	6.8986
	20	74.1836	32.8009	26.9458	9.7142	5.5982

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.22 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) และวิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนเป็น $c^2 = 25$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบดังตารางที่ ก.22 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.23 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	20	94.4781	71.2791	59.2043	16.8344	7.6946
	30	101.4198	48.2579	38.2683	12.1687	6.0906
	40	106.7324	21.0579	14.7092	5.8156	3.0201
10%, $c^2 = 100$	20	96.9319	72.5785	59.8443	16.5901	8.1155
	30	101.5692	48.3373	37.6520	11.5208	5.8854
	40	106.2152	22.7207	14.8234	6.5616	3.0250

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.23 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.23 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.24 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	60	182.3734	86.1232	63.7044	14.4468	7.3474
	70	188.2114	54.1539	35.6573	8.9292	4.5127
	80	191.0267	29.8862	19.3472	5.5284	3.4317
10%, $c^2 = 100$	60	186.3089	96.6407	68.4813	14.2326	7.4346
	70	190.4794	64.9153	47.1676	10.7143	5.9499
	80	195.1381	35.9044	23.6515	6.2587	3.5167

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.24 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงปรกติปลอมปน

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.24 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.3 เมื่อค่าความคลาดเคลื่อนมีการแจกแจงไวบูล

ผลการวิเคราะห์เมื่อค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงเส้นพหุคูณมีการแจกแจงไวบูล โดยค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ของวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ดังตารางที่ 4.25 - 4.36

ตารางที่ 4.25 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	5	33.0417	28.2946	22.9671	8.0242	3.5182
	10	37.0135	15.6834	12.7651	5.7262	3.4748
	15	39.1220	7.0851	5.1782	3.7130	2.5124
$W(1,3)$	5	33.6305	30.1172	24.1374	8.5887	3.6626
	10	37.0765	16.2172	12.9472	6.0041	3.4910
	15	38.1400	7.3388	4.8485	3.5092	2.1785
$W(1,5)$	5	33.7210	28.4455	23.8239	8.0304	3.5468
	10	36.0683	15.2054	12.9153	5.5274	3.5641
	15	38.6134	6.7608	4.3843	3.2406	2.4917

หมายเหตุ **ตัวหนา** หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.25 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ (Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net) และวิธีการถดถอยลาสโซแบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงไวบูลเป็น $W(1,1.5)$ และ $W(1,3)$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.25 จากภาคผนวก ก



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.26 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1, 1.5)$	10	63.0177	47.6618	42.5847	13.6636	5.8154
	15	63.5768	36.0240	29.6992	11.2561	5.7262
	20	66.5059	24.7193	18.0796	7.5084	4.5559
$W(1, 3)$	10	61.0088	47.1731	40.0455	13.4850	5.3963
	15	64.8496	35.4791	29.8580	10.3363	5.3058
	20	66.4054	24.4857	19.6867	7.4689	4.7576
$W(1, 5)$	10	61.3376	45.6652	38.8046	13.3691	5.2396
	15	64.3087	38.5276	31.5616	11.5290	5.6500
	20	66.0925	25.5427	18.6653	7.7991	4.2131

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.26 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) และวิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงที่ค่าความคลาดเคลื่อนมีการแจกแจงไวบูลเป็น $W(1, 1.5)$ และ $W(1, 3)$ และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.26 จากภาคผนวก ก

ตารางที่ 4.27 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	20	95.5383	71.9751	61.5638	18.3407	7.7880
	30	100.6384	46.5271	35.2927	11.0467	5.9334
	40	106.0123	25.3405	15.4053	6.4533	4.2603
$W(1,3)$	20	95.9776	73.5361	58.7985	18.3416	7.6663
	30	101.5208	43.7559	33.5337	10.2491	5.9679
	40	106.2578	21.6254	15.1425	5.8569	3.4669
$W(1,5)$	20	95.9939	71.1452	58.9340	17.5153	7.3318
	30	101.2850	47.6101	35.0144	11.1925	5.9553
	40	106.5041	24.6008	16.3564	6.8230	4.3425

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.27 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.27 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อการศึกษาเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.28 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	60	181.1241	86.9848	59.0713	12.8314	7.6452
	70	186.1973	49.0954	35.2171	8.4384	4.8927
	80	190.5646	28.7975	19.8804	6.0043	3.3205
$W(1,3)$	60	185.3214	91.2636	68.1079	13.8273	8.1367
	70	190.7263	64.2791	47.5868	10.5184	6.3380
	80	194.3859	35.3487	23.3838	7.1841	4.4780
$W(1,5)$	60	194.6759	107.4386	83.6591	15.2281	8.8849
	70	199.6848	72.4277	55.6476	14.1325	8.3092
	80	204.6196	51.0945	39.2746	10.4549	6.8020

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.28 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.28 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารวิจัยทางวิชาการที่จัดทำขึ้นเพื่อการศึกษาเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.29 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	5	20.5767	16.6039	13.3610	5.6101	2.6474
	10	23.2519	9.5673	7.1771	3.6986	2.2643
	15	24.4890	3.6891	2.3814	1.9229	1.3731
$W(1,3)$	5	20.7409	18.8195	15.0050	6.3998	2.8284
	10	23.0582	9.5413	7.5173	3.8113	2.3962
	15	23.9870	3.1092	2.1820	1.6174	1.1729
$W(1,5)$	5	21.5040	18.0613	15.0932	6.3075	2.8447
	10	22.8231	9.7754	7.5471	4.0239	2.4424
	15	24.4012	3.0507	1.8513	1.5490	1.1474

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.29 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.29 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.30 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	10	38.9342	30.1149	24.9089	9.7604	4.3398
	15	40.1848	22.1463	18.4826	7.7824	4.0734
	20	41.5282	14.5614	10.4763	4.3244	2.7512
$W(1,3)$	10	37.6007	27.7743	23.6452	9.3422	4.2126
	15	40.1988	21.9427	18.2905	7.4958	3.8235
	20	41.6519	13.3839	10.3069	4.3302	3.0680
$W(1,5)$	10	38.6860	28.8219	24.7152	9.4492	4.1812
	15	40.4854	21.0133	16.9065	6.8240	3.9123
	20	41.6220	13.4774	9.9253	4.1292	2.7867

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.30 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.30 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในการเรียนการสอน ไม่สามารถนำไปใช้ประโยชน์ทางการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.31 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	20	59.8324	43.5226	36.8385	11.5503	5.6181
	30	64.1609	25.5404	18.6388	6.2911	3.6022
	40	66.7576	9.9905	5.9275	3.0839	1.7418
$W(1,3)$	20	60.6407	42.2056	34.6090	12.5163	5.7477
	30	63.9570	23.6902	18.0919	6.3170	3.7338
	40	67.2575	10.2680	6.0406	3.0263	1.8368
$W(1,5)$	20	59.8413	40.9158	34.3128	11.1597	5.6922
	30	63.0907	26.7124	19.1893	6.3900	4.0765
	40	66.7588	10.4865	6.8553	3.0682	1.8453

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.31 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.31 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.32 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	60	113.2483	44.7067	32.2421	7.2316	4.8641
	70	115.8791	26.4670	18.5849	4.5200	3.1322
	80	118.6417	14.5323	8.5890	2.5581	1.9531
$W(1,3)$	60	113.1947	45.7621	32.6663	6.3538	4.9933
	70	115.9018	26.4231	17.7514	4.4212	3.2590
	80	118.5424	13.0458	8.8840	2.4489	1.8250
$W(1,5)$	60	112.4388	40.6939	29.7987	6.6807	4.5742
	70	114.8023	26.4176	17.6975	4.3121	2.7607
	80	118.8000	13.1954	8.5542	2.8380	1.9313

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.32 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.32 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ทางการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.33 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	5	33.5441	29.8768	24.4213	8.0613	4.1698
	10	37.3213	17.4127	13.7251	6.4584	3.5850
	15	38.3683	6.9672	4.7053	3.6330	1.9819
$W(1,3)$	5	32.8216	28.1178	21.9330	8.0999	4.1192
	10	36.5834	15.7293	12.7584	6.0504	3.4754
	15	39.0256	6.3499	4.6553	3.0401	1.7990
$W(1,5)$	5	34.8818	29.2668	23.4558	8.1309	4.3380
	10	35.6393	16.2328	13.6012	6.2482	3.5334
	15	38.6205	7.4468	4.5710	3.4206	1.9668

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.33 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.33 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ทางการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.34 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	10	60.6755	46.4156	42.0618	13.4160	6.1829
	15	63.9112	35.4226	30.5215	9.5992	5.6150
	20	67.3257	24.6084	18.5579	7.4421	3.5294
$W(1,3)$	10	59.8137	46.0537	39.0613	14.7041	6.2180
	15	64.6790	36.9564	28.9567	11.3068	5.1485
	20	65.0349	26.4998	20.4411	7.1648	4.0774
$W(1,5)$	10	61.1778	44.9983	39.3523	13.7862	6.0552
	15	64.2676	39.6967	32.9763	11.3991	5.6909
	20	67.3480	25.4199	18.4234	7.1629	3.9298

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.35 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.34 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.35 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	20	95.8233	68.5088	56.3191	15.8312	7.4706
	30	101.8919	45.8838	34.1543	11.5161	5.5821
	40	105.5381	22.3111	14.9280	6.4383	3.2630
$W(1,3)$	20	95.9847	67.4533	60.2498	17.6089	7.6520
	30	100.9023	45.0840	36.4519	10.4869	5.4463
	40	105.5945	21.2180	13.7134	6.4632	3.0323
$W(1,5)$	20	96.0290	72.5832	59.6970	17.0961	7.7063
	30	101.4619	47.9891	36.9525	10.8182	5.7460
	40	107.1836	21.4214	13.8929	6.5749	2.8960

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.35 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.35 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในวงวิชาการเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.36 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เมื่อความคลาดเคลื่อนมีการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	60	180.0729	84.4462	60.6301	12.2112	6.8310
	70	186.1456	51.2521	34.5727	8.7198	4.2363
	80	188.0208	30.5973	19.9257	6.1676	3.3534
$W(1,3)$	60	178.9295	87.8452	62.8575	12.6888	7.2896
	70	184.0288	52.2761	35.2735	8.3631	4.3064
	80	188.7173	28.5659	20.2412	6.3012	3.4539
$W(1,5)$	60	179.8456	77.7993	60.5801	11.5704	6.2886
	70	184.3310	54.3626	38.6149	9.8864	5.3765
	80	188.6251	27.9806	20.0634	5.8415	3.0325

หมายเหตุ ตัวหนา หมายถึง วิธีการถดถอยที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในแต่ละกรณี

จากตารางที่ 4.36 พบว่าวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลงตามลำดับ ส่วนวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุดในทุกขนาดตัวอย่างและทุกค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงพหุคูณที่มีการแจกแจงไวบูล

เมื่อขนาดตัวอย่างมากขึ้น วิธีการถดถอยริดจ์ (Ridge) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สูงขึ้น ส่วนวิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำลง และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณมากขึ้นตามขนาดตัวอย่าง (n) ทุกวิธีการถดถอย ยกเว้นวิธีการถดถอยริดจ์ (Ridge) ที่ไม่ตัดตัวแปรอิสระออกจากตัวแบบ ดังตารางที่ ก.36 จากภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ผลจากการประยุกต์ใช้กับข้อมูลจริง

นำชุดข้อมูลจริงซึ่งเป็นชุดข้อมูลการสร้างอาคารที่อยู่อาศัยในเมืองเตหะราน ประเทศอิหร่าน ปี 2561 (Rafiei and Adeli, 2018) มาประยุกต์ใช้เพื่อเปรียบเทียบวิธีการถดถอยแบบปรับด้วยฟังก์ชันการลงโทษ 5 วิธี ซึ่งข้อมูลที่ใช้ประกอบด้วยตัวแปรอิสระ (X) จำนวน 16 ตัวแปร และตัวแปรตาม (y) ที่มีขนาดตัวอย่าง (n) เป็น 5 10 และ 15 โดยหน่วยของตัวแปรส่วนใหญ่ขึ้นอยู่กับค่า IRR (Internal Rate of Return) คือ อัตราผลตอบแทนภายใน เป็นการประเมินว่าผลตอบแทนให้อัตราผลตอบแทนเท่าใด มีรายละเอียดดังตารางที่ 4.37

ตารางที่ 4.37 รายละเอียดตัวแปรอิสระ และตัวแปรตามของชุดข้อมูลการสร้างอาคารที่อยู่อาศัยในเมืองเตหะราน ประเทศอิหร่าน ปี 2561

ตัวแปร	รายละเอียด	หน่วย
X_1	พื้นที่ทั้งหมดในการก่อสร้าง	ตารางเมตร (m^2)
X_2	การประมาณต้นทุนรวมการก่อสร้างเบื้องต้นตามราคา ณ จุดเริ่มต้นของโครงการ	10,000,000 IRR
X_3	การประมาณต้นทุนการก่อสร้างเบื้องต้นตามราคา ณ จุดเริ่มต้นของโครงการ	10,000 IRR
X_4	การประมาณต้นทุนการก่อสร้างเบื้องต้นโดยที่ขึ้นอยู่กับราคา ณ จุดเริ่มต้นของโครงการ	10,000 IRR
X_5	ระยะเวลาการก่อสร้าง	เดือน
X_6	ราคาต่อหน่วย ณ จุดเริ่มต้นของโครงการต่อตารางเมตร	10,000 IRR
X_7	พื้นที่ทั้งหมดที่มีใบอนุญาตก่อสร้างอาคารที่ออกโดยเมือง / เทศบาล	ตารางเมตร (m^2)
X_8	สภาพคล่องทางการเงิน	10,000,000 IRR
X_9	การลงทุนของภาคเอกชนสำหรับการสร้างอาคารใหม่	10,000,000 IRR
X_{10}	อัตราดอกเบี้ยเงินกู้	ร้อยละ (%)
X_{11}	ต้นทุนเฉลี่ยการก่อสร้างของอาคารโดยภาคเอกชนในขณะที่การก่อสร้างแล้วเสร็จ	10,000 IRR / m^2
X_{12}	ต้นทุนเฉลี่ยการก่อสร้างอาคารโดยภาคเอกชน ณ จุดเริ่มต้นของโครงการ	10,000 IRR / m^2
X_{13}	อัตราแลกเปลี่ยนอย่างเป็นทางการ (ดอลลาร์)	IRR
X_{14}	อัตราแลกเปลี่ยนนอกระบบ (ดอลลาร์)	IRR
X_{15}	ประชากรในเมือง	คน
X_{16}	ราคาทองคำต่อออนซ์	IRR
y	ราคาขายจริงของอาคารที่อยู่อาศัย	10,000 IRR

ผลการวิเคราะห์ของวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) แสดงค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของชุดข้อมูลการสร้างอาคารที่อยู่อาศัย ดังตารางที่ 4.38

ตารางที่ 4.38 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และจำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยพหุคูณ ของชุดข้อมูลการสร้างอาคารที่อยู่อาศัยในเมือง تهران ประเทศอิหร่าน ปี 2561 ของวิธีการถดถอยริดจ์ (Ridge) วิธีการถดถอยลาสโซ่ (Lasso) วิธีการถดถอยอีลาสติคเน็ต (Elastic Net) วิธีการถดถอยลาสโซ่แบบปรับปรุง (ALasso) และวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net)

ขนาดตัวอย่าง (n)	วิธีการ				
	Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$n=5$	104,678.60 (16)	105,560.00 (2)	105,560.00 (5)	39,850.79 (2)	4,053.28 (2)
$n=10$	194,735.60 (16)	152,793.90 (2)	135,290.40 (5)	51,930.01 (2)	3,492.19 (2)
$n=15$	162,579.30 (16)	200,074.90 (2)	207,856.10 (5)	59,951.53 (2)	5,283.75 (2)

หมายเหตุ ค่าในวงเล็บ หมายถึง จำนวนตัวแปรอิสระ (p) ที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ

จากตารางที่ 4.38 จะพบว่าชุดข้อมูลการสร้างอาคารที่อยู่อาศัย วิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ต่ำที่สุดทุกขนาดตัวอย่าง ทั้งยังมีผลลัพธ์สอดคล้องกับผลการวิเคราะห์ของชุดข้อมูลที่จำลองขึ้นในงานวิจัยครั้งนี้ที่ว่าวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net) ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุด โดยแบ่งตามขนาดตัวอย่างได้สมการถดถอย ดังนี้

1. สมการของแต่ละวิธีการถดถอย เมื่อขนาดตัวอย่าง (n) เป็น 5

1.1 สมการถดถอยวิธีการถดถอยริดจ์ (Ridge)

$$\hat{y} = 5,818.4491 - 0.5051X_1 - 0.7993X_2 - 0.4223X_3 - 1.0682X_4$$

$$-134.9568X_5 - 0.1007X_6 + 22.4361X_7 - 0.000038X_8 - 0.0091X_9$$

$$-10.1455X_{10} - 0.0324X_{11} - 0.0338X_{12} - 0.0172X_{13} - 0.0239X_{14}$$

$$-0.0008X_{15} - 0.00003X_{16}$$

1.2 สมการถดถอยวิธีการถดถอยลาสโ (Lasso)

$$\hat{y} = 6,707.6978 - 3.2242X_1 - 9.1133X_2$$

1.3 สมการถดถอยวิธีการถดถอยอีลาสติกเน็ต (Elastic Net)

$$\hat{y} = 8,998.8315 - 3.6686X_1 - 4.1914X_2 - 537.3900X_5 - 0.0346X_{13} - 0.0085X_{14}$$

1.4 สมการถดถอยวิธีการถดถอยลาสโแบบปรับปรุง (ALasso)

$$\hat{y} = 6,771.1964 - 3.2581X_1 - 9.4494X_2$$

1.5 สมการถดถอยวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net)

$$\hat{y} = 6,798.6316 - 3.3362X_1 - 9.3918X_2$$

2. สมการของแต่ละวิธีการถดถอย เมื่อขนาดตัวอย่าง (n) เป็น 10

2.1 สมการถดถอยวิธีการถดถอยริดจ์ (Ridge)

$$\hat{y} = 5,925.605 - 0.5164X_1 - 0.7664X_2 - 0.5051X_3 - 1.1882X_4 - 130.4126X_5 - 0.1037X_6 + 9.9087X_7 - 0.000044X_8 - 0.0062X_9 - 5.9395X_{10} - 0.0350X_{11} - 0.0350X_{12} - 0.0169X_{13} - 0.0233X_{14} - 0.0012X_{15} - 0.000042X_{16}$$

2.2 สมการถดถอยวิธีการถดถอยลาสโ (Lasso)

$$\hat{y} = 6,706.2880 - 3.2118X_1 - 9.1066X_2$$

2.3 สมการถดถอยวิธีการถดถอยอีลาสติกเน็ต (Elastic Net)

$$\hat{y} = 8,887.3106 - 3.7241X_1 - 3.7745X_2 - 489.5465X_5 - 0.0394X_{13} - 0.0235X_{14}$$

2.4 สมการถดถอยวิธีการถดถอยลาสโแบบปรับปรุง (ALasso)

$$\hat{y} = 6,765.9321 - 3.2407X_1 - 9.4650X_2$$

2.5 สมการถดถอยวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (AElastic Net)

$$\hat{y} = 6,793.8757 - 3.3222X_1 - 9.4017X_2$$

3. สมการของแต่ละวิธีการถดถอย เมื่อขนาดตัวอย่าง (n) เป็น 15

3.1 สมการถดถอยวิธีการถดถอยริดจ์ (Ridge)

$$\hat{y} = 5818.449 - 0.5051X_1 - 0.7993X_2 - 0.4223X_3 - 1.0682X_4 - 134.9568X_5 - 0.1001X_6 + 22.4361X_7 - 0.00004X_8 - 0.0091X_9 - 10.1455X_{10} - 0.0324X_{11} - 0.0338X_{12} - 0.0173X_{13} - 0.0240X_{14} - 0.0008X_{15} - 0.00003X_{16}$$

3.2 สมการถดถอยวิธีการถดถอยลาสโ (Lasso)

$$\hat{y} = 6,707.6978 - 3.2242X_1 - 9.1133X_2$$

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกส่งมอบหรือเผยแพร่ข้อมูลอันเนื่องถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 สมการถดถอยวิธีการถดถอยอีลาสติคเน็ต (Elastic Net)

$$\hat{y} = 8,998.8315 - 3.6686X_1 - 4.1914X_2 - 537.3900X_5 - 0.0346X_{13} - 0.0085X_{14}$$

3.4 สมการถดถอยวิธีการถดถอยลาโซ่แบบปรับปรุง (ALasso)

$$\hat{y} = 6,771.1964 - 3.2581X_1 - 9.4494X_2$$

3.5 สมการถดถอยวิธีการถดถอยอีลาสติคเน็ตแบบปรับปรุง (AElastic Net)

$$\hat{y} = 6,798.6316 - 3.3362X_1 - 9.3918X_2$$



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ โดยวิธีการถดถอยแบบปรับด้วยฟังก์ชันการลงโทษ 5 วิธี คือ วิธีการถดถอยริดจ์ (Ridge Regression) วิธีการถดถอยลาสโซ (Least absolute shrinkage and selection operator Regression : Lasso) วิธีการถดถอยอีลาสติกเน็ต (Elastic Net Regression) วิธีการถดถอยลาสโซแบบปรับปรุง (The adaptive least absolute shrinkage and selection operator Regression) และวิธีการถดถอยอีลาสติกเน็ตแบบปรับปรุง (The adaptive Elastic net Regression) ในกรณีที่จำนวนตัวแปรอิสระมากกว่าขนาดตัวอย่าง ในลักษณะข้อมูลที่มีมิติสูง โดยศึกษาและเปรียบเทียบในสถานการณ์ได้ดังนี้

1. กำหนดจำนวนตัวแปรอิสระ (p) ในตัวแบบการถดถอยเชิงเส้นพหุคูณมากกว่าขนาดตัวอย่าง (n) แสดงดังตารางที่ 5.1

ตารางที่ 5.1 จำนวนตัวแปรอิสระ (p) และขนาดตัวอย่าง (n)

p	n
16	5 10 และ 15
30	10 15 และ 20
50	20 30 และ 40
100	60 70 และ 80

2. ค่าสัมประสิทธิ์การถดถอยพหุคูณ (β) กำหนด 3 กรณี

- กำหนดค่าสัมประสิทธิ์การถดถอยทุกตัวมีค่าเป็น 1
- กำหนดค่าสัมประสิทธิ์การถดถอยมีค่าเป็น 0.5 จำนวนครึ่งหนึ่งของจำนวนตัวแปรอิสระ และที่เหลือมีค่าสัมประสิทธิ์การถดถอยเป็น 1
- กำหนดค่าสัมประสิทธิ์การถดถอยมีค่าเป็น -1 จำนวนครึ่งหนึ่งของจำนวนตัวแปรอิสระ และที่เหลือมีค่าสัมประสิทธิ์การถดถอยเป็น 1

ผู้วิจัยทำศึกษาวิธีการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ โดยวิธีการถดถอยแบบปรับด้วยฟังก์ชันการลงโทษ และเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) เพื่อใช้ในการตัดสินใจเลือกวิธีการประมาณค่าที่ดีที่สุด

เอกสารนี้เป็นเอกสารสงวนไว้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมีเหตุผลปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.1 สรุปผลการวิจัย

เมื่อค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงเส้นพหุคูณมีการแจกแจงปกติ การแจกแจงปกติปลอมปน และการแจกแจงไวบูล จะพบว่าวิธีการถดถอยอิลาสติคเน็ตแบบปรับปรุงให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดในทุกการแจกแจงของค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงเส้นพหุคูณ ทุกขนาดตัวอย่าง และทุกจำนวนตัวแปรอิสระ และเมื่อนำมาประยุกต์ใช้กับข้อมูลจริงให้ผลเช่นเดียวกับการจำลองข้อมูล เนื่องจากวิธีการถดถอยอิลาสติคเน็ตแบบปรับปรุงมีพารามิเตอร์การปรับ 3 ค่า ได้แก่ λ_1, λ_2 และ γ และวิธีที่ให้ผลลัพธ์ดีรองลงมาคือวิธีการถดถอยลาสโซ่แบบปรับปรุง และรองลงมาคือวิธีการถดถอยลาสโซ่ และวิธีการถดถอยอิลาสติคเน็ต และวิธีสุดท้ายคือวิธีการถดถอยริดจ์ เนื่องจากวิธีการถดถอยริดจ์ไม่ได้ตัดตัวแปรอิสระออกจากตัวแบบของการถดถอยเชิงเส้นพหุคูณ และผลการวิเคราะห์ยังแสดงให้เห็นว่าวิธีการถดถอยอิลาสติคเน็ตแบบปรับปรุงเมื่อขนาดตัวอย่างเล็กจะไม่มีผลต่อค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย กล่าวคือเมื่อขนาดตัวอย่างเล็กไม่ทำให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำลง แต่เมื่อขนาดตัวอย่างใหญ่ขึ้นจะทำให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำลง

5.2 ข้อเสนอแนะ

5.2.1 ด้านการนำไปใช้ประโยชน์

การวิจัยในครั้งนี้สามารถนำไปประยุกต์ใช้ในการเลือกใช้วิธีการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ โดยวิธีการถดถอยแบบปรับด้วยฟังก์ชันการลงโทษ ในกรณีที่มีจำนวนตัวแปรอิสระมากกว่าขนาดตัวอย่างในลักษณะข้อมูลที่มีมิติสูงได้อย่างเหมาะสม เมื่อตัวแปรอิสระสุ่มมาจากประชากรที่มีการแจกแจงปกติ และเมื่อค่าความคลาดเคลื่อนของตัวแบบการถดถอยเชิงเส้นพหุคูณมีการแจกแจงปกติ การแจกแจงปกติปลอมปน และการแจกแจงไวบูล ควรใช้วิธีการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ โดยวิธีการถดถอยแบบปรับด้วยฟังก์ชันการลงโทษ คือวิธีการถดถอยอิลาสติคเน็ตแบบปรับปรุง

5.2.2 ด้านการศึกษาวิจัย

- ในงานวิจัยครั้งต่อไปผู้วิจัยควรศึกษากรณีที่ตัวแปรอิสระเกิดพหุสัมพันธ์กัน
- ควรศึกษาในกรณีที่ตัวแปรอิสระสุ่มมาจากประชากรที่มีการแจกแจงอื่น ๆ นอกเหนือจากการแจกแจงปกติ
- ควรศึกษาในกรณีที่ค่าความคลาดเคลื่อนมีการแจกแจงอื่น ๆ นอกเหนือจากการที่ศึกษาในงานวิจัยครั้งนี้
- ในงานวิจัยเลือกมาศึกษา 5 วิธี แต่ในความเป็นจริงยังมีวิธีอื่นๆที่น่าสนใจ โดยอาจนำมาใช้พิจารณา เพื่อเปรียบเทียบกับวิธีในงานวิจัย เช่น วิธีการปรับของการถดถอยอิลาสติคเน็ตแบบปรับ (Adjusted adaptive regularized regression : AAELastic) เป็นต้น

บรรณานุกรม

จุฬาทิพย์ นันทสุวรรณ. 2559. “การเลือกพารามิเตอร์การปรับสำหรับวิธีการถดถอยแบบลาสโซ่.”

วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี, จุฬาลงกรณ์มหาวิทยาลัย.

ชุตिकाญจน์ ชูสวัสดิ์ และสุปราณี ลิสวัสดิ์. 2561. “การเปรียบเทียบประสิทธิภาพของวิธีวิเคราะห์การถดถอยแบบบริดจ์ ลาสโซ่ และลาสโซ่แบบปรับปรุง ในตัวแบบการถดถอยปัวซอง ภายใต้ข้อมูลที่มีมิติสูงแบบบางเบา และตัวแปรอิสระมีความสัมพันธ์กันสูง.” หน้า 305-314. ใน **การประชุมวิชาการเสนอผลงานวิจัยระดับบัณฑิตศึกษาแห่งชาติ ครั้งที่ 19**. ขอนแก่น : มหาวิทยาลัยขอนแก่น.

ณัฐพร ภัคดี. 2552. “การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณเมื่อตัวแปรอิสระมีพหุสัมพันธ์กันโดยวิธีการถดถอยบริดจ์.” วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

ทิฆัมพร สารระกอ และนันท กุลวานิช. 2557. “การเปรียบเทียบประสิทธิภาพการพยากรณ์และการคัดเลือกตัวแปรของวิธีเพิ่มลดตัวแปรแบบขั้นตอน วิธีแลสโซ่ วิธีอิลาสติกเน็ต และวิธีแลสโซ่ปรับปรุง สำหรับผลกระทบขนาดเล็ก และมีค่าสัมประสิทธิ์บางตัวเป็นศูนย์.” หน้า 327-343. ใน **การประชุมสัมมนาทางวิชาการ มทร.ตะวันออก มรภ.กลุ่มศรีอยุธยา และราชบัณฑิตยสถาน**. กรุงเทพฯ : มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก.

นพรัตน์ กระต่ายทอง. 2542 “การแปลงข้อมูลให้มีการแจกแจงปกติ.” วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี, จุฬาลงกรณ์มหาวิทยาลัย.

ปณัช อภาวุฒิชัย, วิชิต หล่อจ๊ะระชุมท์กุล และ จิราวัลย์ จิตรถเวช. 2561. “การวิเคราะห์การถดถอยเชิงเส้นกรณีที่มีข้อมูลมิติสูงโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม.” **วารสารวิทยาศาสตร์บูรพา**. 23(1) : 255-270.

ประชุม สุวดี. 2553. **ทฤษฎีการอนุมานเชิงสถิติ**. พิมพ์ครั้งที่ 3. กรุงเทพฯ : โครงการส่งเสริมและพัฒนาเอกสารวิชาการ สถาบันบัณฑิตพัฒนบริหารศาสตร์.

วรภาพร เหลือสินทรัพย์. 2558. **การวิเคราะห์การถดถอยและสหสัมพันธ์**. กรุงเทพฯ : สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

เอกสารนี้เป็นสายชล สิ้นสมบุรณ์ทอง. 2560. **การวิเคราะห์การถดถอย Regression Analysis**. กรุงเทพฯ : ถ้าไม่ว่ากรณีใดๆทั้งสิ้น จามจุรีโปรดักส์. คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อัชฌา อระวีพร. 2561. **ความน่าจะเป็น**. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย

- Algamal, Z.Y. and Lee, M.H. 2015. "Regularized Logistic Regression with Adjusted Adaptive Elastic Net for Gene Selection in High Dimensional Cancer Classification" *Computers in Biology and Medicine Journal*. 67(1) : 136-145.
- Boonstra, P.S. Mukherjee, B. and Taylor, J.M. 2015. "A Small-Sample Choice of the Tuning Parameter in Ridge Regression." *Statistics Sinica*. 25(3) : 1185-1206.
- Efron, B. Hastie, T. Johnstone, I. and Tibshirani, R. 2004. "Least Angle Regression" *The annals of Statistics*. 32(2) : 407-499.
- Hastie, T. Tibshirani, R. and Friedman, J. 2009. **The Elements of Statistical Learning : Data Mining Inference and Prediction**. 2nd ed. California : Springer.
- Hoerl, A.E. and Kennard, R.W. 1970. "Ridge Regression: Biased Estimation for Nonorthogonal Problems." *Technometrics Journal*. 12(1) : 55-67.
- Rafiei, M.H. and Adeli, H. 2018, **Residential Building Data Set**. [Online]. Available : <http://archive.ics.uci.edu/ml/datasets/Residential+Building+Data+Set#>.
- Syed, A.R. 2011. "A Review of Cross Validation and Adaptive Model Selection, in Mathematics and Statistics." Master of Science in the College of Arts and Sciences, Georgia State University.
- Tibshirani R.J. 1996. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society*. 58(1) : 267-288.
- Wagner, B.S. Morais, A.L. and Cordeiro, G.M. 2010. "The Weibull-Geometric distribution." *Journal of Statistical Computation and Simulation*. 81(5) : 645-657.
- Wang, H. Li, B. and Leng, C. 2009. "Shrinkage Tuning Parameter Selection with a Diverging Number of Parameters." *Journal of Royal Statistical Society Series B*. 71(3) : 23-25.
- Zou, H. 2006. "The Adaptive Lasso and Its Oracle Properties." *American Statistical Association*. 101(476) : 1418-1429.
- Zou, H. Hastie, T. and Tibshirani, R. 2007. "On the Degrees of Freedom of Lasso." *Annals of Statistics*. 35(5) : 2173-2192.
- Zou, H. and Hastie, T. 2005. "Regularization and Variable Selection via the Elastic Net." *Journal of Royal Statistical Society Series B : Statistical Methodology*. 67(2) : 301-320.
- Zou, H. and Zhang, T. 2009. "On the Adaptive Elastic Net with a Diverging Number of Parameters." *Annals of Statistics*. 37(4) : 1733-1751.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.1 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	11	10	11
$N(0,3)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	9	11	10	11
$N(0,5)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	9	11	9	11
$N(0,9)$	5	16	2	4	2	3
	10	16	5	6	6	7
	15	16	9	10	9	11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.2 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	10	30	5	7	5	7
	15	30	7	10	8	10
	20	30	12	15	13	15
$N(0,3)$	10	30	4	6	5	7
	15	30	7	10	8	10
	20	30	12	15	13	15
$N(0,5)$	10	30	4	6	5	7
	15	30	7	10	8	10
	20	30	11	15	12	14
$N(0,9)$	10	30	4	6	5	7
	15	30	7	10	8	10
	20	30	11	13	12	14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.3 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	20	50	9	13	11	14
	30	50	18	22	19	22
	40	50	28	33	27	31
$N(0,3)$	20	50	9	13	11	14
	30	50	18	22	19	22
	40	50	29	33	28	31
$N(0,5)$	20	50	9	13	11	14
	30	50	17	22	19	22
	40	50	28	33	28	31
$N(0,9)$	20	50	9	13	11	14
	30	50	17	22	19	22
	40	50	28	33	28	31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.4 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	60	100	36	45	41	45
	70	100	49	56	50	53
	80	100	61	67	60	63
$N(0,3)$	60	100	37	44	41	44
	70	100	48	55	50	53
	80	100	61	68	60	62
$N(0,5)$	60	100	36	44	40	44
	70	100	48	55	50	53
	80	100	60	66	60	63
$N(0,9)$	60	100	35	43	40	44
	70	100	48	56	50	53
	80	100	58	66	59	61

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.5 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5, 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	11	10	11
$N(0,3)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	9	11	9	11
$N(0,5)$	5	16	2	4	2	3
	10	16	5	6	6	7
	15	16	9	10	9	10
$N(0,9)$	5	16	2	4	2	3
	10	16	5	6	5	7
	15	16	8	10	9	10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.6 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	10	30	5	7	5	7
	15	30	7	10	8	10
	20	30	12	15	13	15
$N(0,3)$	10	30	4	6	5	7
	15	30	7	10	8	10
	20	30	12	15	13	15
$N(0,5)$	10	30	4	6	5	7
	15	30	7	10	8	10
	20	30	12	14	12	14
$N(0,9)$	10	30	4	6	5	7
	15	30	7	10	8	10
	20	30	11	13	12	14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.7 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	20	50	10	13	11	14
	30	50	18	22	19	22
	40	50	29	33	28	31
$N(0,3)$	20	50	10	13	11	14
	30	50	18	22	19	22
	40	50	30	33	28	31
$N(0,5)$	20	50	10	13	11	13
	30	50	18	22	19	22
	40	50	29	34	28	31
$N(0,9)$	20	50	10	13	11	14
	30	50	18	23	19	22
	40	50	30	33	28	31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.8 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	60	100	39	46	42	44
	70	100	51	58	51	54
	80	100	63	69	61	63
$N(0,3)$	60	100	38	45	41	44
	70	100	49	56	51	53
	80	100	63	68	60	63
$N(0,5)$	60	100	37	44	41	44
	70	100	49	56	50	53
	80	100	61	67	59	62
$N(0,9)$	60	100	37	44	41	44
	70	100	48	55	50	53
	80	100	60	65	59	62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.9 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5, 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	5	16	2	4	2	3
	10	16	6	7	6	7
	15	16	10	12	10	11
$N(0,3)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	11	10	11
$N(0,5)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	9	10	9	11
$N(0,9)$	5	16	2	4	2	3
	10	16	5	6	6	7
	15	16	9	10	9	11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.10 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	10	30	5	7	5	7
	15	30	8	10	8	11
	20	30	12	15	13	15
$N(0,3)$	10	30	4	6	5	7
	15	30	7	10	8	10
	20	30	12	15	12	15
$N(0,5)$	10	30	4	6	5	7
	15	30	7	10	8	10
	20	30	11	14	12	15
$N(0,9)$	10	30	4	6	5	7
	15	30	7	10	8	10
	20	30	11	14	12	14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.11 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	20	50	9	13	11	13
	30	50	18	21	19	22
	40	50	28	33	28	31
$N(0,3)$	20	50	10	13	11	14
	30	50	17	22	19	22
	40	50	29	33	28	31
$N(0,5)$	20	50	9	12	11	13
	30	50	18	22	19	22
	40	50	29	32	28	31
$N(0,9)$	20	50	10	13	11	14
	30	50	17	22	19	22
	40	50	28	33	28	31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.12 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปรกติ ($N(\mu, \sigma^2)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$N(0,1)$	60	100	37	45	41	45
	70	100	49	56	51	54
	80	100	62	68	60	63
$N(0,3)$	60	100	36	44	41	44
	70	100	47	55	50	53
	80	100	60	67	59	63
$N(0,5)$	60	100	37	44	41	44
	70	100	48	56	50	53
	80	100	60	66	59	63
$N(0,9)$	60	100	36	44	41	44
	70	100	47	55	50	54
	80	100	59	65	59	62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.13 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	11	10	11
10%, $c^2 = 100$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	9	10	9	10

ตารางที่ ก.14 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	10	30	4	7	5	7
	15	30	7	10	8	10
	20	30	12	15	13	15
10%, $c^2 = 100$	10	30	4	6	5	7
	15	30	7	9	8	10
	20	30	11	14	12	14

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ภายใต้การคุ้มครองของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ไม่ควรเผยแพร่โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.15 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	20	50	9	12	11	13
	30	50	18	22	19	22
	40	50	28	32	28	30
10%, $c^2 = 100$	20	50	9	13	11	14
	30	50	18	22	19	22
	40	50	28	32	28	31

ตารางที่ ก.16 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	60	100	37	44	41	44
	70	100	48	55	50	53
	80	100	60	67	60	62
10%, $c^2 = 100$	60	100	35	43	41	43
	70	100	47	54	49	53
	80	100	57	65	59	62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษานานาชาติ ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านอื่น ๆ
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.17 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	9	11	10	11
10%, $c^2 = 100$	5	16	2	4	2	3
	10	16	5	6	6	7
	15	16	8	10	9	10

ตารางที่ ก.18 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	10	30	4	7	5	7
	15	30	7	10	8	10
	20	30	12	15	13	15
10%, $c^2 = 100$	10	30	4	6	5	7
	15	30	7	9	8	10
	20	30	11	13	12	14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.19 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	20	50	9	13	11	13
	30	50	18	23	19	22
	40	50	29	33	28	31
10%, $c^2 = 100$	20	50	9	13	11	14
	30	50	18	22	19	22
	40	50	30	33	28	31

ตารางที่ ก.20 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	60	100	37	46	41	44
	70	100	50	56	51	53
	80	100	61	67	60	62
10%, $c^2 = 100$	60	100	36	44	40	44
	70	100	48	55	50	53
	80	100	59	65	59	61

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.21 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปlomปนด้วยเปอร์เซ็นต์การปlomปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปlomปนที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	9	11	9	11
10%, $c^2 = 100$	5	16	2	4	2	3
	10	16	5	6	6	7
	15	16	8	10	9	11

ตารางที่ ก.22 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปlomปนด้วยเปอร์เซ็นต์การปlomปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปlomปนที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	10	30	4	6	5	7
	15	30	7	10	8	10
	20	30	12	14	13	15
10%, $c^2 = 100$	10	30	4	6	5	7
	15	30	7	9	8	10
	20	30	11	14	12	14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.23 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	20	50	9	13	11	14
	30	50	17	21	18	22
	40	50	29	33	28	31
10%, $c^2 = 100$	20	50	9	13	11	13
	30	50	17	21	19	22
	40	50	28	32	28	31

ตารางที่ ก.24 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปนด้วยเปอร์เซ็นต์การปลอมปน และสเกลแฟกเตอร์ (c) ของข้อมูลที่มีการปลอมปนที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปร กรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
10%, $c^2 = 25$	60	100	36	44	41	44
	70	100	49	56	50	54
	80	100	61	68	60	63
10%, $c^2 = 100$	60	100	35	44	41	44
	70	100	46	53	49	53
	80	100	59	66	59	63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.25 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาด เคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	12	10	11
$W(1,3)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	12	10	11
$W(1,5)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	12	10	11

ตารางที่ ก.26 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาด เคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	10	30	4	6	5	7
	15	30	8	10	8	10
	20	30	12	15	13	15
$W(1,3)$	10	30	4	7	5	7
	15	30	8	10	9	10
	20	30	12	15	13	15
$W(1,5)$	10	30	5	7	5	7
	15	30	7	10	8	10
	20	30	12	15	12	15

เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อใช้ในการศึกษาเท่านั้น ไม่สามารถนำข้อมูลไปใช้เพื่อการอื่นได้

ตารางที่ ก.27 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	20	50	9	12	11	14
	30	50	17	22	19	22
	40	50	28	32	28	31
$W(1,3)$	20	50	9	13	11	14
	30	50	18	22	19	22
	40	50	28	33	28	31
$W(1,5)$	20	50	9	13	11	14
	30	50	17	22	19	22
	40	50	28	32	28	30

ตารางที่ ก.28 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น 1 ทั้งหมด

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	60	100	36	45	41	45
	70	100	49	56	50	53
	80	100	61	67	60	63
$W(1,3)$	60	100	36	43	41	44
	70	100	46	54	50	53
	80	100	60	66	59	62
$W(1,5)$	60	100	34	41	40	43
	70	100	45	52	48	52
	80	100	55	61	57	60

เอกสารนี้เป็นเอกสารงานวิจัยที่จัดทำขึ้นเพื่อการศึกษาเท่านั้น ไม่ควรนำข้อมูลไปใช้ประโยชน์อื่นใดโดยไม่ได้รับอนุญาตจากเจ้าของเอกสาร

ตารางที่ ก.29 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาด เคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	12	10	11
$W(1,3)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	12	10	11
$W(1,5)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	12	10	11

ตารางที่ ก.30 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาด เคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	10	30	4	7	5	7
	15	30	7	10	8	10
	20	30	12	15	13	15
$W(1,3)$	10	30	4	7	5	7
	15	30	8	10	8	10
	20	30	12	15	13	14
$W(1,5)$	10	30	4	7	5	7
	15	30	8	11	8	10
	20	30	12	15	13	15

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ห้ามเผยแพร่โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามทำซ้ำ ผลิต กระจาย หรือเผยแพร่ข้อมูลของเอกสารนี้โดยไม่ได้รับอนุญาต

ตารางที่ ก.31 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	20	50	9	13	11	13
	30	50	18	22	19	22
	40	50	30	34	28	31
$W(1,3)$	20	50	10	13	11	14
	30	50	19	23	19	22
	40	50	29	34	28	31
$W(1,5)$	20	50	10	13	11	13
	30	50	17	22	19	22
	40	50	29	33	28	31

ตารางที่ ก.32 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น 0.5 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	60	100	39	46	42	44
	70	100	51	57	51	54
	80	100	63	69	61	63
$W(1,3)$	60	100	38	46	42	44
	70	100	51	58	51	54
	80	100	63	69	61	63
$W(1,5)$	60	100	40	47	42	44
	70	100	51	57	51	54
	80	100	64	69	61	63

เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อใช้ในการเรียนการสอนเท่านั้น ไม่ควรนำเอกสารนี้ไปเผยแพร่หรือใช้เพื่อวัตถุประสงค์อื่นใดโดยไม่ได้รับอนุญาตจากทางมหาวิทยาลัย

ตารางที่ ก.33 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 5 10 และ 15 จำนวนตัวแปรอิสระ (p) เท่ากับ 16 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาด เคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	12	10	11
$W(1,3)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	12	10	11
$W(1,5)$	5	16	2	4	2	3
	10	16	5	7	6	7
	15	16	10	12	10	11

ตารางที่ ก.34 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 10 15 และ 20 จำนวนตัวแปรอิสระ (p) เท่ากับ 30 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาด เคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	10	30	4	6	5	7
	15	30	7	10	9	10
	20	30	12	15	13	15
$W(1,3)$	10	30	4	6	5	7
	15	30	8	10	8	11
	20	30	11	14	13	15
$W(1,5)$	10	30	4	7	5	7
	15	30	7	10	8	10
	20	30	12	15	13	15

เอกสารนี้เป็นเอกสารงานวิจัยที่จัดทำขึ้นเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามทำซ้ำโดยไม่ได้รับอนุญาตจากเจ้าของเอกสารทุกครั้ง มีการใช้

ตารางที่ ก.35 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 20 30 และ 40 จำนวนตัวแปรอิสระ (p) เท่ากับ 50 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	20	50	10	13	11	14
	30	50	18	22	19	22
	40	50	28	32	28	31
$W(1,3)$	20	50	10	13	11	14
	30	50	18	21	19	22
	40	50	28	33	28	31
$W(1,5)$	20	50	9	13	11	14
	30	50	17	21	19	22
	40	50	29	33	28	31

ตารางที่ ก.36 จำนวนตัวแปรอิสระที่ถูกคัดเลือกเข้าสู่ตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อความคลาดเคลื่อนมีการแจกแจงการแจกแจงไวบูล ($W(\alpha, \beta)$) ที่มีขนาดตัวอย่าง (n) เท่ากับ 60 70 และ 80 จำนวนตัวแปรอิสระ (p) เท่ากับ 100 ตัวแปรกรณีค่าสัมประสิทธิ์การถดถอยเป็น -1 และ 1

ค่าความคลาดเคลื่อน	n	วิธีการ				
		Ridge	Lasso	Elastic Net	ALasso	AElastic Net
$W(1,1.5)$	60	100	37	45	41	44
	70	100	49	57	51	54
	80	100	61	68	60	63
$W(1,3)$	60	100	36	44	41	44
	70	100	49	57	51	54
	80	100	62	67	60	63
$W(1,5)$	60	100	38	45	41	45
	70	100	49	56	50	54
	80	100	61	67	60	63

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี 100 ปี 1942-2022



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข1 คำสั่งโปรแกรมอาร์ สำหรับสร้างกราฟการแจกแจงความหนาแน่นความน่าจะเป็นปกติ

```
e=seq(-10,10,length.out=10000)
plot(e,dnorm(e,0,sqrt(1)),main="NormalProbabilityDensityFunction",type="l",lwd="3",
pch=16,lty=1,ylab="Density",col="black")
lines(e,dnorm(e,0,sqrt(3)),type="l",pch=17,lty=2,lwd="3",col="blue")
lines(e,dnorm(e,0,sqrt(5)),type="l",pch=18,lty=3,lwd="3",col="red")
lines(e,dnorm(e,0,sqrt(9)),type="l",pch=19,lty=4,lwd="3",col="green")
labels=c("N(0,1)","N(0,3)","N(0,5)","N(0,9)")
colors=c("black","blue","red","green")
A=c(1,2,3,4)
legend("topright",inset=0,labels,lwd=2,lty=A,col=colors)
```

คำสั่งโปรแกรมอาร์ สำหรับสร้างกราฟการแจกแจงความหนาแน่นความน่าจะเป็นปกติ

ปลอมปน

```
e=seq(-25,25,length.out=10000)
dnorm1=dnorm(e,0,sqrt(1))
dnorm2=dnorm(e,0,sqrt(100))
Density=0.9*dnorm(e,0,sqrt(1))+0.1*dnorm(e,0,sqrt(100))
plot(e,Density,main="ContaminatedNormalProbabilityDensityFunction",type="l",lwd="3",
pch=16,lty=1,ylim=c(0,0.44),ylab="Density",col="black")
lines(e,dnorm1,ylab="Density",type="l",pch=17,lty=2,lwd="3",col="blue")
lines(e,dnorm2,ylab="Density",type="l",pch=17,lty=2,lwd="3",col="red")
labels=c("0.9N(0,1)+0.1(0,100)","N(0,1)","N(0,100)")
colors=c("black","blue","red")
A=c(1,2,3)
legend("topright",inset=0,labels,lwd=2,lty=A,col=colors)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำสั่งโปรแกรมอาร์ สำหรับสร้างกราฟการแจกแจงความหนาแน่นความน่าจะเป็นไวบูล

```
e=seq(0,4,length.out=10000)
weibull1=dweibull(e,scale=1,shape=1.5)
weibull2=dweibull(e,scale=1,shape=3)
weibull3=dweibull(e,scale=1,shape=5)
plot(e,weibull1,main='WeibullProbabilityDensityFunction',ylab="Density",type='l',lwd='3',
pch=16,lty=1,ylim=c(0,2),col="black")
lines(e,weibull2,ylab="Density",type='l',pch=17,lty=2,lwd='3',col= "blue")
lines(e,weibull3,ylab="Density",type='l',pch=18,lty=3,lwd='3',col= "red")
labels=c("W(1,1.5)","W(1,3)","W(1,5)")
colors=c("black","blue","red")
A=c(1,2,3)
legend("topright",inset=0,labels,lwd=2,lty=A,col=colors)
```

ข2 คำสั่งโปรแกรมอาร์ สำหรับค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ในกรณีข้อมูลสุ่มจากประชากรที่มีการแจกแจงปรกติ

```
library(glmnet)
library(HDeconometrics)
set.seed(9)
n = 15
m = 1000
q = 4
Var = c(1,3,5,9)
mse1 = c(); mse2 = c(); mse3 = c() ; mse4 = c(); mse5 = c()
amse1 = c(); amse2 = c(); amse3 = c() ; amse4 = c(); amse5 = c()
fit_Ridge1 = c() ; fit_Lasso1 = c() ; fit_ELasticNet1 = c()
fit_ALasso1 = c() ; fit_AELasticNet1 = c()
fit1 = c(); fit2 = c(); fit3 = c() ; fit4 = c(); fit5 = c()
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

for (e in 1:q) {
  for (j in 1:m) {
    error = rnorm(n,0,sqrt(Var[e]))
    x1 = rnorm(n,0,2)
    x2 = rnorm(n,0,2)
    x3 = rnorm(n,0,2)
    x4 = rnorm(n,0,2)
    x5 = rnorm(n,0,2)
    x6 = rnorm(n,0,2)
    x7 = rnorm(n,0,2)
    x8 = rnorm(n,0,2)
    x9 = rnorm(n,0,2)
    x10 = rnorm(n,0,2)
    x11 = rnorm(n,0,2)
    x12 = rnorm(n,0,2)
    x13 = rnorm(n,0,2)
    x14 = rnorm(n,0,2)
    x15 = rnorm(n,0,2)
    x16 = rnorm(n,0,2)
    xm=data.frame(x1=x1,x2=x2,x3=x3,x4=x4,x5=x5,x6=x6,x7=x7,x8=x8,x9=x9,x10=x10,
    x11=x11,x12=x12,x13=x13,x14=x14,x15=x15, x16=x16)
    x = as.matrix(xm)
    y = 1 + x[,1] + x[,2] + x[,3] + x[,4] + x[,5] + x[,6] + x[,7] + x[,8] + x[,9] + x[,10] + x[,11]
    + x[,12]+ x[,13] + x[,14] + x[,15] + x[,16] + error
    ##### Ridge #####
    cvfit = cv.glmnet(x,y, alpha = 0)
    ll = cvfit$lambda.min
    fit_Ridge = glmnet(x,y,family="gaussian",lambda=ll, alpha = 0)
    est_Ridge = predict(fit_Ridge, newx = x)
    y_1 <- as.numeric(est_Ridge)
    residual1 = c()
    for (i in 1:n) {
      residual1[i] = (y[i]-y_1[i])^2
    }
  }
}

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

}
mse1[j] = mean(residual1)
fit_Ridge1[j] = fit_Ridge$df
##### Lasso #####
cvfit = cv.glmnet(x,y, alpha = 1)
ll = cvfit$lambda.min
fit_Lasso = glmnet(x,y, family = "gaussian", lambda =ll,alpha = 1)
est_Lasso = predict(fit_Lasso, newx = x)
y_1 <- as.numeric(est_Lasso)
residual2 = c()
for (i in 1:n) {
residual2[i] = (y[i]-y_1[i])^2
}
mse2[j] = mean(residual2)
fit_Lasso1[j] = fit_Lasso$df
##### Elastic Net #####
cvfit = cv.glmnet(x,y, alpha = 0.5)
ll = cvfit$lambda.min
fit_ElasticNet = glmnet(x,y, family = "gaussian", lambda =ll,alpha = 0.5)
est_ElasticNet = predict(fit_ElasticNet, newx = x)
y_1 <- as.numeric(est_ElasticNet)
residual3 = c()
for (i in 1:n) {
residual3[i] = (y[i]-y_1[i])^2
}
mse3[j] = mean(residual3)
fit_ElasticNet1[j] = fit_ElasticNet$df
##### ALasso #####
tau=1
Lasso=ic.glmnet(x,y,crit = "bic")
first.step.coef=coef(Lasso)[-1]
penalty.factor=abs(first.step.coef)^(-tau)
cvfit = cv.glmnet(x,y, alpha = 1)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

ll = cvfit$lambda.min
fit_ALasso = glmnet(x,y, family = "gaussian", alpha = 1 ,lambda =ll, penalty.factor =
penalty.factor)
est_ALasso = predict(fit_ALasso, newx = x)
y_1 <- as.numeric(est_ALasso)
residual4 = c()
for (i in 1:n) {
residual4[i] = (y[i]-y_1[i])^2
}
mse4[j] = mean(residual4)
fit_ALasso1[j] = fit_ALasso$df
##### AElastic Net #####
tau=1
Lasso=ic.glmnet(x,y,crit = "bic")
first.step.coef=coef(Lasso)[-1]
penalty.factor=abs(first.step.coef+1/sqrt(nrow(x)))^(-tau)
cvfit = cv.glmnet(x,y, alpha = 0.5)
ll = cvfit$lambda.min
fit_AElasticNet = glmnet(x,y, family = "gaussian", alpha = 0.5 ,lambda =ll,
penalty.factor = penalty.factor)
est_AElasticNet = predict(fit_AElasticNet, newx = x)
y_1 <- as.numeric(est_AElasticNet)
residual5 = c()
for (i in 1:n) {
residual5[i] = (y[i]-y_1[i])^2
}
mse5[j] = mean(residual5)
fit_AElasticNet1[j] = fit_AElasticNet$df
#####
cat(c("loop :",j),fill=T)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 }
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
##### ค่า AMSE และจำนวนตัวแปรอิสระที่คัดเลือกโดยเฉลี่ย แต่ละวิธี ทั้ง 5 วิธี #####
amse1[e]=mean(mse1)
amse2[e]=mean(mse2)
amse3[e]=mean(mse3)
amse4[e]=mean(mse4)
amse5[e]=mean(mse5)
fit1[e]=mean(fit_Ridge1)
fit2[e]=mean(fit_Lasso1)
fit3[e]=mean(fit_ElasticNet1)
fit4[e]=mean(fit_ALasso1)
fit5[e]=mean(fit_AElasticNet1)
}
#####
amse1<-data.frame(amse1[1],amse2[1],amse3[1],amse4[1],amse5[1])
amse2<-data.frame(amse1[2],amse2[2],amse3[2],amse4[2],amse5[2])
amse3<-data.frame(amse1[3],amse2[3],amse3[3],amse4[3],amse5[3])
amse4<-data.frame(amse1[4],amse2[4],amse3[4],amse4[4],amse5[4])
f_1<-data.frame(fit1[1],fit2[1],fit3[1],fit4[1],fit5[1])
f_2<-data.frame(fit2[2],fit2[2],fit3[2],fit4[2],fit5[2])
f_3<-data.frame(fit3[3],fit2[3],fit3[3],fit4[3],fit5[3])
f_4<-data.frame(fit4[4],fit2[4],fit3[4],fit4[4],fit5[4])
a1 = as.matrix(amse1)
a2 = as.matrix(amse2)
a3 = as.matrix(amse3)
a4 = as.matrix(amse4)
f1 = as.matrix(f_1)
f2 = as.matrix(f_2)
f3 = as.matrix(f_3)
f4 = as.matrix(f_4)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำสั่งโปรแกรมอาร์ สำหรับค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ในกรณี
ข้อมูลสุ่มจากประชากรที่มีการแจกแจงปรกติปดอมปน

```

library(glmnet)
library(HDeconometrics)
set.seed(9)
n=15
m=1000
q=2
cpct = 0.1
mu1 = 0 ; sig1 = 1
mu2 = 0 ; sig2=c(25,100)
mse1 = c(); mse2 = c(); mse3 = c() ; mse4 = c(); mse5 = c()
amse1 = c(); amse2 = c(); amse3 = c() ; amse4 = c(); amse5 = c()
fit_Ridge1 = c() ; fit_Lasso1 = c() ; fit_ELasticNet1 = c()
fit_ALasso1 = c() ; fit_AELasticNet1 = c()
fit1 = c(); fit2 = c(); fit3 = c() ; fit4 = c(); fit5 = c()
#####
for (e in 1:q) {
  for (j in 1:m) {
    x00 = rnorm(n , mean=mu1 , sd=sig1)
    x01 = rnorm(n , mean=mu2 , sd=sqrt(sig2[e]))
    flag0 = rbinom(n , size=1 , prob=cpct)
    error = (x00*(1-flag0)) + (x01*flag0)
    x1 = rnorm(n,0,2)
    x2 = rnorm(n,0,2)
    x3 = rnorm(n,0,2)
    x4 = rnorm(n,0,2)
    x5 = rnorm(n,0,2)
    x6 = rnorm(n,0,2)
    x7 = rnorm(n,0,2)
    x8 = rnorm(n,0,2)
  }
}

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

x9 = rnorm(n,0,2)
x10 = rnorm(n,0,2)
x11 = rnorm(n,0,2)
x12 = rnorm(n,0,2)
x13 = rnorm(n,0,2)
x14 = rnorm(n,0,2)
x15 = rnorm(n,0,2)
x16 = rnorm(n,0,2)

xm=data.frame(x1=x1,x2=x2,x3=x3,x4=x4,x5=x5,x6=x6,x7=x7,x8=x8,x9=x9,x10=x10,
x11=x11,x12=x12,x13=x13,x14=x14,x15=x15, x16=x16)
x = as.matrix(xm)
y = 1 + x[,1] + x[,2] + x[,3] + x[,4] + x[,5] + x[,6] + x[,7] + x[,8] + x[,9] + x[,10] + x[,11]
+ x[,12]+ x[,13] + x[,14] + x[,15] + x[,16] + error
##### Ridge #####
cvfit = cv.glmnet(x,y, alpha = 0)
ll = cvfit$lambda.min
fit_Ridge = glmnet(x,y,family="gaussian",lambda=ll, alpha = 0)
est_Ridge = predict(fit_Ridge, newx = x)
y_1 <- as.numeric(est_Ridge)
residual1 = c()
for (i in 1:n) {
residual1[i] = (y[i]-y_1[i])^2
}
mse1[j] = mean(residual1)
fit_Ridge1[j] = fit_Ridge$df
##### Lasso #####
cvfit = cv.glmnet(x,y, alpha = 1)
ll = cvfit$lambda.min
fit_Lasso = glmnet(x,y, family = "gaussian", lambda =ll,alpha = 1)
est_Lasso = predict(fit_Lasso, newx = x)
y_1 <- as.numeric(est_Lasso)
residual2 = c()
for (i in 1:n) {

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

residual2[i] = (y[i]-y_1[i])^2
}
mse2[j] = mean(residual2)
fit_Lasso1[j] = fit_Lasso$df
##### Elastic Net #####
cvfit = cv.glmnet(x,y, alpha = 0.5)
ll = cvfit$lambda.min
fit_ElasticNet = glmnet(x,y, family = "gaussian", lambda =ll,alpha = 0.5)
est_ElasticNet = predict(fit_ElasticNet, newx = x)
y_1 <- as.numeric(est_ElasticNet)
residual3 = c()
for (i in 1:n) {
residual3[i] = (y[i]-y_1[i])^2
}
mse3[j] = mean(residual3)
fit_ELasticNet1[j] = fit_ElasticNet$df
##### ALasso #####
tau=1
Lasso=ic.glmnet(x,y,crit = "bic")
first.step.coef=coef(Lasso)[-1]
penalty.factor=abs(first.step.coef)^(-tau)
cvfit = cv.glmnet(x,y, alpha = 1)
ll = cvfit$lambda.min
fit_ALasso = glmnet(x,y, family = "gaussian", alpha = 1 ,lambda =ll, penalty.factor =
penalty.factor)
est_ALasso = predict(fit_ALasso, newx = x)
y_1 <- as.numeric(est_ALasso)
residual4 = c()
for (i in 1:n) {
residual4[i] = (y[i]-y_1[i])^2
}
mse4[j] = mean(residual4)
fit_ALasso1[j] = fit_ALasso$df

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 "ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้"

```
##### AElastic Net #####
tau=1
Lasso=ic.glmnet(x,y,crit = "bic")
first.step.coef=coef(Lasso)[-1]
penalty.factor=abs(first.step.coef+1/sqrt(nrow(x)))^(-tau)
cvfit = cv.glmnet(x,y, alpha = 0.5)
ll = cvfit$lambda.min
fit_AElasticNet = glmnet(x,y, family = "gaussian", alpha = 0.5 ,lambda =ll,
penalty.factor = penalty.factor)
est_AElasticNet = predict(fit_AElasticNet, newx = x)
y_1 <- as.numeric(est_AElasticNet)
residual5 = c()
for (i in 1:n) {
residual5[i] = (y[i]-y_1[i])^2
}
mse5[j] = mean(residual5)
fit_AElasticNet1[j] = fit_AElasticNet$df
#####
cat(c("loop :",j),fill=T)
}
##### ค่า AMSE และจำนวนตัวแปรอิสระที่คัดเลือกโดยเฉลี่ย แต่ละวิธี ทั้ง 5 วิธี #####
amse1[e]=mean(mse1)
amse2[e]=mean(mse2)
amse3[e]=mean(mse3)
amse4[e]=mean(mse4)
amse5[e]=mean(mse5)
fit1[e]=mean(fit_Ridge1)
fit2[e]=mean(fit_Lasso1)
fit3[e]=mean(fit_ElasticNet1)
fit4[e]=mean(fit_ALasso1)
fit5[e]=mean(fit_AElasticNet1)
}
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งยังมีให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#####
amse1<-data.frame(amse1[1],amse2[1],amse3[1],amse4[1],amse5[1])
amse2<-data.frame(amse1[2],amse2[2],amse3[2],amse4[2],amse5[2])
amse3<-data.frame(amse1[3],amse2[3],amse3[3],amse4[3],amse5[3])
amse4<-data.frame(amse1[4],amse2[4],amse3[4],amse4[4],amse5[4])
f_1<-data.frame(fit1[1],fit2[1],fit3[1],fit4[1],fit5[1])
f_2<-data.frame(fit2[2],fit2[2],fit3[2],fit4[2],fit5[2])
f_3<-data.frame(fit3[3],fit2[3],fit3[3],fit4[3],fit5[3])
f_4<-data.frame(fit4[4],fit2[4],fit3[4],fit4[4],fit5[4])
a1 = as.matrix(amse1)
a2 = as.matrix(amse2)
a3 = as.matrix(amse3)
a4 = as.matrix(amse4)
f1 = as.matrix(f_1)
f2 = as.matrix(f_2)
f3 = as.matrix(f_3)
f4 = as.matrix(f_4)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำสั่งโปรแกรมอาร์ สำหรับค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ในกรณี
ข้อมูลสุ่มจากประชากรที่มีการแจกแจงไวบูล

```

library(glmnet)
library(HDeconometrics)
set.seed(9)
n=15
m=1000
q=3
sh=c(1.5,3,5)
mse1 = c(); mse2 = c(); mse3 = c() ; mse4 = c(); mse5 = c()
amse1 = c(); amse2 = c(); amse3 = c() ; amse4 = c(); amse5 = c()
fit_Ridge1 = c() ; fit_Lasso1 = c() ; fit_ElasticNet1 = c()
fit_ALasso1 = c() ; fit_AElasticNet1 = c()
fit1 = c(); fit2 = c(); fit3 = c() ; fit4 = c(); fit5 = c()
#####
for (e in 1:q) {
for (j in 1:m) {
error = rweibull(n , scale = 1 , shape=sh[q])
x1 = rnorm(n,0,2)
x2 = rnorm(n,0,2)
x3 = rnorm(n,0,2)
x4 = rnorm(n,0,2)
x5 = rnorm(n,0,2)
x6 = rnorm(n,0,2)
x7 = rnorm(n,0,2)
x8 = rnorm(n,0,2)
x9 = rnorm(n,0,2)
x10 = rnorm(n,0,2)
x11 = rnorm(n,0,2)
x12 = rnorm(n,0,2)
x13 = rnorm(n,0,2)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

x14 = rnorm(n,0,2)
x15 = rnorm(n,0,2)
x16 = rnorm(n,0,2)
xm=data.frame(x1=x1,x2=x2,x3=x3,x4=x4,x5=x5,x6=x6,x7=x7,x8=x8,x9=x9,x10=x10,
x11=x11,x12=x12,x13=x13,x14=x14,x15=x15, x16=x16)
x = as.matrix(xm)
y = 1 + x[,1] + x[,2] + x[,3] + x[,4] + x[,5] + x[,6] + x[,7] + x[,8] + x[,9] + x[,10] + x[,11]
+ x[,12]+ x[,13] + x[,14] + x[,15] + x[,16] + error
##### Ridge #####
cvfit = cv.glmnet(x,y, alpha = 0)
ll = cvfit$lambda.min
fit_Ridge = glmnet(x,y,family="gaussian",lambda=ll, alpha = 0)
est_Ridge = predict(fit_Ridge, newx = x)
y_1 <- as.numeric(est_Ridge)
residual1 = c()
for (i in 1:n) {
residual1[i] = (y[i]-y_1[i])^2
}
mse1[j] = mean(residual1)
fit_Ridge1[j] = fit_Ridge$df
##### Lasso #####
cvfit = cv.glmnet(x,y, alpha = 1)
ll = cvfit$lambda.min
fit_Lasso = glmnet(x,y, family = "gaussian", lambda = ll,alpha = 1)
est_Lasso = predict(fit_Lasso, newx = x)
y_1 <- as.numeric(est_Lasso)
residual2 = c()
for (i in 1:n) {
residual2[i] = (y[i]-y_1[i])^2
}
mse2[j] = mean(residual2)
fit_Lasso1[j] = fit_Lasso$df

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

##### Elastic Net #####
cvfit = cv.glmnet(x,y, alpha = 0.5)
ll = cvfit$lambda.min
fit_ElasticNet = glmnet(x,y, family = "gaussian", lambda =ll,alpha = 0.5)
est_ElasticNet = predict(fit_ElasticNet, newx = x)
y_1 <- as.numeric(est_ElasticNet)
residual3 = c()
for (i in 1:n) {
residual3[i] = (y[i]-y_1[i])^2
}
mse3[j] = mean(residual3)
fit_ELasticNet1[j] = fit_ElasticNet$df
##### ALasso #####
tau=1
Lasso=ic.glmnet(x,y,crit = "bic")
first.step.coef=coef(Lasso)[-1]
penalty.factor=abs(first.step.coef)^(-tau)
cvfit = cv.glmnet(x,y, alpha = 1)
ll = cvfit$lambda.min
fit_ALasso = glmnet(x,y, family = "gaussian", alpha = 1 ,lambda =ll, penalty.factor =
penalty.factor)
est_ALasso = predict(fit_ALasso, newx = x)
y_1 <- as.numeric(est_ALasso)
residual4 = c()
for (i in 1:n) {
residual4[i] = (y[i]-y_1[i])^2
}
mse4[j] = mean(residual4)
fit_ALasso1[j] = fit_ALasso$df
##### AElastic Net #####
tau=1
Lasso=ic.glmnet(x,y,crit = "bic")
first.step.coef=coef(Lasso)[-1]

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

penalty.factor=abs(first.step.coef+1/sqrt(nrow(x)))^(-tau)
cvfit = cv.glmnet(x,y, alpha = 0.5)
ll = cvfit$lambda.min
fit_AElasticNet = glmnet(x,y, family = "gaussian", alpha = 0.5 ,lambda =ll,
penalty.factor = penalty.factor)
est_AElasticNet = predict(fit_AElasticNet, newx = x)
y_1 <- as.numeric(est_AElasticNet)
residual5 = c()
for (i in 1:n) {
residual5[i] = (y[i]-y_1[i])^2
}
mse5[j] = mean(residual5)
fit_AElasticNet1[j] = fit_AElasticNet$df
#####
cat(c("loop :",j),fill=T)
}
##### ค่า AMSE และจำนวนตัวแปรอิสระที่คัดเลือกโดยเฉลี่ย แต่ละวิธี ทั้ง 5 วิธี #####
amse1[e]=mean(mse1)
amse2[e]=mean(mse2)
amse3[e]=mean(mse3)
amse4[e]=mean(mse4)
amse5[e]=mean(mse5)
fit1[e]=mean(fit_Ridge1)
fit2[e]=mean(fit_Lasso1)
fit3[e]=mean(fit_ElasticNet1)
fit4[e]=mean(fit_ALasso1)
fit5[e]=mean(fit_AElasticNet1)
}
#####
amse1<-data.frame(amse1[1],amse2[1],amse3[1],amse4[1],amse5[1])
amse2<-data.frame(amse1[2],amse2[2],amse3[2],amse4[2],amse5[2])
amse3<-data.frame(amse1[3],amse2[3],amse3[3],amse4[3],amse5[3])
amse4<-data.frame(amse1[4],amse2[4],amse3[4],amse4[4],amse5[4])

```

```
f_1<-data.frame(fit1[1],fit2[1],fit3[1],fit4[1],fit5[1])
f_2<-data.frame(fit2[2],fit2[2],fit3[2],fit4[2],fit5[2])
f_3<-data.frame(fit3[3],fit2[3],fit3[3],fit4[3],fit5[3])
f_4<-data.frame(fit4[4],fit2[4],fit3[4],fit4[4],fit5[4])
a1 = as.matrix(amse1)
a2 = as.matrix(amse2)
a3 = as.matrix(amse3)
a4 = as.matrix(amse4)
f1 = as.matrix(f_1)
f2 = as.matrix(f_2)
f3 = as.matrix(f_3)
f4 = as.matrix(f_4)
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ	นางสาวเบญจมาศ รุ่งศรานนท์
วัน เดือน ปีเกิด	18 พฤศจิกายน 2537
ที่อยู่ปัจจุบัน	472 ซอยสะแกงาม 35/3 แยก 10 แขวงแสมดำ เขตบางขุนเทียน กรุงเทพมหานคร 10150
ประวัติการศึกษา	(2559) วิทยาศาสตรบัณฑิต สาขาสถิติประยุกต์ เกรตเฉลี่ย 3.25 (สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง) (2561) วิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติและการวิเคราะห์ธุรกิจ เกรตเฉลี่ย 3.87 (สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง)
ทุนการศึกษาที่ได้รับ	ทุนยกเว้นค่าธรรมเนียมการศึกษา ระดับบัณฑิตศึกษาคณะวิทยาศาสตร์
ผลงานทางวิชาการ	Rungsaranon, B. and Araveeporn, A. 2018. A Comparison of Parameter Estimation with Penalized Regression Analysis on High-Dimensional Data. PROCEEDINGS International Conference on Applied Statistics 2018.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้