

การเปรียบเทียบการประมาณค่า ในการวิเคราะห์ความถดถอยพหุ โดยวิธีรีดจ์รีเกรสชัน

รีเกรสชันพหุคูณ, ปีลคอมโพเนนท์ และวิธีกำลังสองน้อยที่สุด ในกรณีที่เกิด

พหุสัมพันธ์ระหว่างตัวแปรอิสระ



นางสาว ดวงพร ชูรักษา

ศูนย์วิทยพัชการ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาตรีศึกษาศาสตร์มหาบัณฑิต

จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชาสถิติ

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย


พ.ศ. 2529

ISBN 974-566-593-2

013302

I 15590628

A COMPARISON OF RIDGE REGRESSION, REGRESSION PRINCIPAL COMPONENT
AND MULTIPLE REGRESSION IN THE CASE OF EXISTING
MULTICOLLINEARITY AMONG INDEPENDENT VARIABLES.



Miss Doungporn Choorugsa

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science
Department of Statistics
Graduate School

Chulalongkorn University

1986

ISBN 974-566-593-2

หัวข้อวิทยานิพนธ์ การเปรียบเทียบการประมาณค่าในการวิเคราะห์ความถดถอยพหุ โดยวิธี
รีดจ์รีเกรสชั่น, รีเกรสชั่นพรีนซ์ปีลคอมโพเนนท์และวิธีกำลังสองน้อยที่สุด
ในกรณีที่เกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ

โดย นางสาว ดวงพร ชูรักษา

ภาควิชา สถิติ

อาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร. สร้อย พิศาลบุตร



บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้มหาวิทยาลัยราชภัฏฉบับนี้เป็นส่วนหนึ่งของการ
การศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....
(รองศาสตราจารย์ ดร. สร้อย พิศาลบุตร)

รักษาการในตำแหน่งรองคณบดีฝ่ายวิชาการ
ปฏิบัติราชการแทนรักษาการในตำแหน่งคณบดีบัณฑิตวิทยาลัย

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร. สุชาติ กระเนินหม)

..... กรรมการ
(รองศาสตราจารย์ ดร. สร้อย พิศาลบุตร)

..... กรรมการ
(รองศาสตราจารย์ สุภาพ เตชะรินทร์)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ผกาวัต ศิริรังษี)

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

หัวข้อวิทยานิพนธ์	การเปรียบเทียบการประมาณค่าในการวิเคราะห์ความถดถอยพหุ โดยวิธี ริดจ์ รีเกรสชัน รีเกรสชันพรีนซีเปิ้ลคอมโพเนนท์และวิธีกำลังสองน้อยที่สุด ในกรณีที่เกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ
ชื่อนิสิต	นางสาว ดวงพร ชูรักษ์
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ ดร. สรชัย พิศาลบุตร
ภาควิชา	สถิติ
ปีการศึกษา	2528



บทคัดย่อ

ในกรณีที่เกิดสภาพไม่เหมาะสมในตัวแปรอิสระ การประมาณค่าสัมประสิทธิ์ความถดถอย
ด้วยวิธีกำลังสองน้อยที่สุด จะทำให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองสูง ซึ่งวิธีการนี้อาจทำให้
ค่าประมาณพารามิเตอร์ β ที่ได้มีคุณภาพไม่ดีเท่าที่ควร การวิจัยครั้งนี้ ทำการศึกษาเพื่อเปรียบเทียบ
การประมาณค่าเมื่อเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ โดยใช้วิธีริดจ์ รีเกรสชัน วิธีรีเกรสชัน
พรีนซีเปิ้ลคอมโพเนนท์ และวิธีกำลังสองน้อยที่สุด ข้อมูลที่ใช้ในการวิจัยเรื่องนี้ได้มาจากวิธีกาลอง
ของ Wichern and Churchill (1984 : 304) ซึ่งจะได้ตัวแปรอิสระที่มีความสัมพันธ์กันใน
ระดับต่าง ๆ

จากการศึกษาการประมาณค่าสัมประสิทธิ์ความถดถอยพหุ สามารถสรุปได้ว่า เมื่อข้อมูล
มีความแปรปรวน 0.01 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองโดยวิธีริดจ์ รีเกรสชัน และวิธี
รีเกรสชันพรีนซีเปิ้ลคอมโพเนนท์จะมีค่ามากกว่าวิธีกำลังสองน้อยที่สุด เสมอไม่ว่าตัวแปร
อิสระจะมีพหุสัมพันธ์กันมากหรือน้อย เมื่อความแปรปรวนของข้อมูลมีค่าเท่ากับ 1 ค่าเฉลี่ยความ
คลาดเคลื่อนกำลังสอง โดยวิธีริดจ์ รีเกรสชัน และวิธีรีเกรสชันพรีนซีเปิ้ลคอมโพเนนท์ จะมีค่า
น้อยกว่าวิธีกำลังสองน้อยที่สุด ในกรณีที่ตัวแปรอิสระมีพหุสัมพันธ์กันมาก แต่เมื่อความแปรปรวนของ
ข้อมูลมีค่าเท่ากับ 25 และ 100 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง โดยวิธีริดจ์ รีเกรสชัน
วิธีรีเกรสชันพรีนซีเปิ้ลคอมโพเนนท์จะมีค่ามากกว่าวิธีกำลังสองน้อยที่สุด ไม่ว่าตัวแปรอิสระจะมี
พหุสัมพันธ์กันมากหรือน้อยก็ตาม

จากการศึกษาค่าประมาณของตัวแปรตาม ผลการวิจัยปรากฏว่าในทุกกรณีค่าเฉลี่ย
ความคลาดเคลื่อนกำลังสองของค่าประมาณด้วยวิธีกำลังสองน้อยที่สุดจะมีค่าน้อยที่สุด ค่าเฉลี่ย
ความคลาดเคลื่อนกำลังสองของวิธีดัลท์ ไรเกอร์สกี จะใกล้เคียงกับวิธีกำลังสองน้อยที่สุดมากกว่า
วิธีไรเกอร์สกี-พริ้นซ์-บีลคอมโพเนนท์ เมื่อเพิ่มขนาดตัวอย่างให้มากขึ้น ค่าเฉลี่ยความคลาดเคลื่อน
กำลังสองโดยวิธีดัลท์ ไรเกอร์สกีและวิธีไรเกอร์สกี-พริ้นซ์-บีลคอมโพเนนท์ มีค่าใกล้เคียงกับวิธีกำลัง
สองน้อยที่สุดมากขึ้น



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Thesis Title : A comparison of Ridge regression, Regression principal component and Multiple Regression in the case of existing multicollinearity among independent variables.

Name : Miss Doungporn Choorugsa

Advisor : Associate Professor Sorachai Bhisalbutra, Ph.D.

Department : Statistics

Year : 1985



ABSTRACT

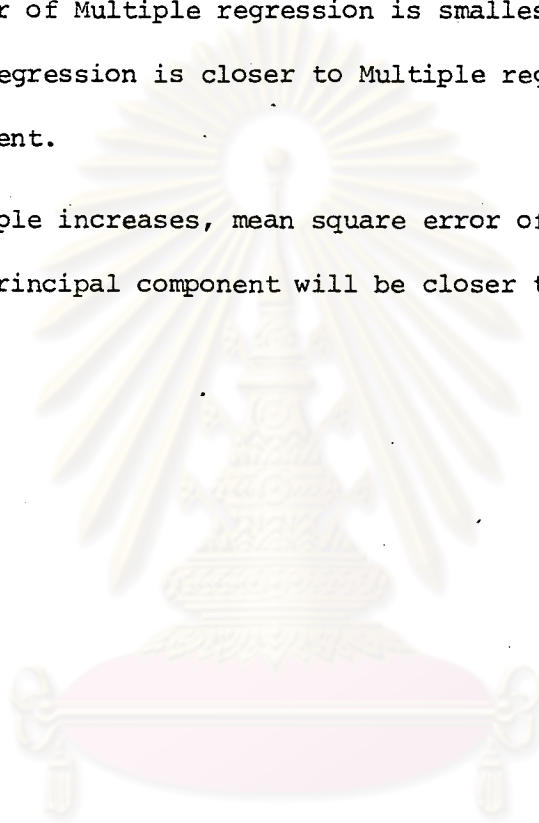
The maximum mean square error will result when the multiple regression is used to determine the regression coefficients of illcondition of the independent variables. Infact, the approximate value of parameter β is not good enough in quality. This research studies a comparison of Ridge regression, Regression principal component and Multiple regression analysis in the case of existing multicollinearity among independent variables. The data for each experiment were obtained through simulation from Wichern and Churchill (1984 : 304).

The conclusion obtained from the studies of regression coefficients estimation when the variance of data is 0.01, mean square error of Ridge regression and Regression principal component are larger than Multiple regression analysis for every level of degree of multicollinearity .. When variance of data is 1 , mean square error of Ridge regression and Regression principal component are smaller than Multiple regression in the case that the degree of multicollinearity of the independent variable is small. When variance of the data is 25 and 100, the mean

square error of Ridge regression and Regression principal component are larger than Multiple regression analysis for every level of degree of multicollinearity.

The studies of dependent variable estimation result that the mean square error of Multiple regression is smallest. The mean square error of Ridge regression is closer to Multiple regression than Regression principal component.

When sample increases, mean square error of Ridge regression and Regression principal component will be closer to Multiple regression.



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความกรุณาของ รองศาสตราจารย์ ดร. สรชัย พิศาลบุตร คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ที่ให้คำแนะนำ ปรึกษา ตลอดจนแก้ไขข้อบกพร่องต่าง ๆ เป็นอย่างดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณท่านอาจารย์ทุก ๆ ท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้ให้แก่ ผู้วิจัยมาโดยตลอด

ขอกราบขอบพระคุณ คุณพ่อ คุณแม่ ที่ช่วยส่งเสริมและสนับสนุนการเรียนของผู้วิจัย ตลอดมา

สุดท้ายนี้ขอขอบคุณ เพื่อน ๆ ทุกคนที่ให้ความช่วยเหลือ และคำแนะนำต่าง ๆ

ดวงพร ชูรัักษ์

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



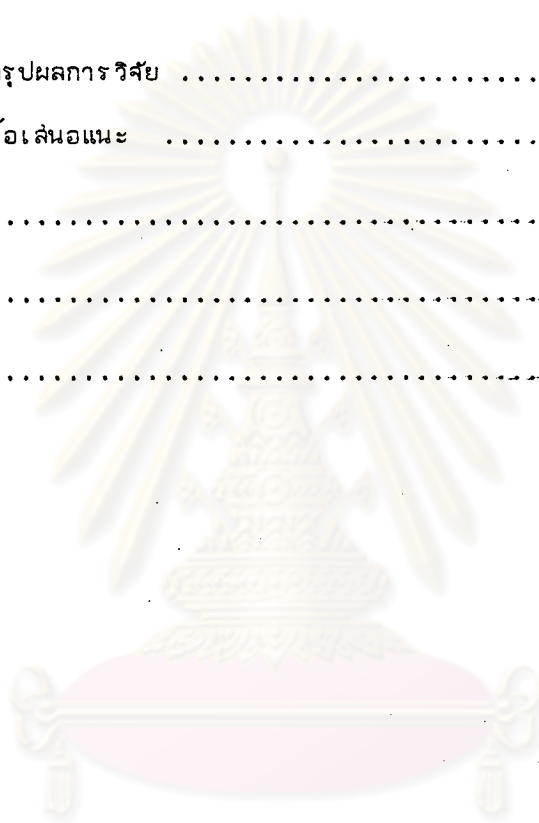
สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ค
บทคัดย่อภาษาอังกฤษ	ง
กิตติกรรมประกาศ	ช
รายการตารางประกอบ	ญ
รายการรูปประกอบ	ฎ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์ของการวิจัย	3
1.3 สัมมติฐานของการวิจัย	3
1.4 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัยนี้	3
บทที่ 2 ทฤษฎีที่เกี่ยวข้องกับการวิจัย	5
2.1 คุณลักษณะของค่าประมาณสัมประสิทธิ์ความถดถอยพหุด้วยวิธีกำลังสองน้อยที่สุด	10
2.2 การประมาณค่าสัมประสิทธิ์ความถดถอยพหุ โดยวิธีรีดจ์ รีเกรสชั่น	13
2.3 การประมาณค่าสัมประสิทธิ์ความถดถอยพหุ โดยวิธีรีเกรสชั่นพหุขั้นฮิลเบิร์ตคอมโพเนนท์	16
2.4 ดัชนีพหุสัมพันธ์	23
บทที่ 3 วิธีดำเนินการวิจัย	25
บทที่ 4 ผลการวิจัย	31

สารบัญ

หน้า

บทที่ 5	สรุปผลการวิจัยและข้อเสนอแนะ	69
	5.1 สรุปผลการวิจัย	69
	5.2 ข้อเสนอแนะ	70
บรรณานุกรม	71
ภาคผนวก	72
ประวัติ	93



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รายการตารางประกอบ

ตารางที่	หน้า
4.1 แสดงค่าขอบเขตบนโดยเฉลี่ยของ PPC และ PR เมื่อจำนวนตัวแปร อิสระเท่ากับ 4	34
4.2 แสดงค่าขอบเขตล่างโดยเฉลี่ยของ PPC และ PR เมื่อจำนวนตัวแปร อิสระเท่ากับ 4	35
4.3 แสดงค่าขอบเขตบนโดยเฉลี่ยของ VPC และ VR เมื่อจำนวนตัวแปร อิสระเท่ากับ 4	36
4.4 แสดงค่าขอบเขตล่างโดยเฉลี่ยของ VPC และ VR เมื่อจำนวนตัวแปร อิสระเท่ากับ 4	37
4.5 แสดงค่าขอบเขตบนโดยเฉลี่ยของ PPC และ PR เมื่อจำนวนตัวแปร อิสระเท่ากับ 6	38
4.6 แสดงค่าขอบเขตล่างโดยเฉลี่ยของ PPC และ PR เมื่อจำนวนตัวแปร อิสระเท่ากับ 6	39
4.7 แสดงค่าขอบเขตบนโดยเฉลี่ยของ VPC และ VR เมื่อจำนวนตัวแปร อิสระเท่ากับ 6	40
4.8 แสดงค่าขอบเขตล่างโดยเฉลี่ยของ VPC และ VR เมื่อจำนวนตัวแปร อิสระเท่ากับ 6	41
4.9 แสดงค่าขอบเขตบนโดยเฉลี่ยของ PPC และ PR เมื่อจำนวนตัวแปร อิสระเท่ากับ 9	42
4.10 แสดงค่าขอบเขตล่างโดยเฉลี่ยของ PPC และ PR เมื่อจำนวนตัวแปร อิสระเท่ากับ 9	43
4.11 แสดงค่าขอบเขตบนโดยเฉลี่ยของ VPC และ VR เมื่อจำนวนตัวแปร อิสระเท่ากับ 9	44
4.12 แสดงค่าขอบเขตล่างโดยเฉลี่ยของ VPC และ VR เมื่อจำนวนตัวแปร อิสระเท่ากับ 9	45

รายการรูปประกอบ

หน้า

รูปที่		
4.15 ค	แสดงค่าขอบเขตบนและขอบเขตล่างโดยเฉลี่ยของ PPC และ PR เมื่อใช้จำนวนตัวแปรอิสระเท่ากับ 9 ขนาดตัวอย่างเท่ากับ 100 ความแปรปรวนเท่ากับ 25	66
4.16 ก	แสดงค่าขอบเขตบนและขอบเขตล่างโดยเฉลี่ยของ PPC และ PR เมื่อใช้จำนวนตัวแปรอิสระเท่ากับ 9 ขนาดตัวอย่างเท่ากับ 30 ความแปรปรวนเท่ากับ 100	67
4.16 ข	แสดงค่าขอบเขตบนและขอบเขตล่างโดยเฉลี่ยของ PPC และ PR เมื่อใช้จำนวนตัวแปรอิสระเท่ากับ 9 ขนาดตัวอย่างเท่ากับ 50 ความแปรปรวนเท่ากับ 100	67
4.16 ค	แสดงค่าขอบเขตบนและขอบเขตล่างโดยเฉลี่ยของ PPC และ PR เมื่อใช้จำนวนตัวแปรอิสระเท่ากับ 9 ขนาดตัวอย่างเท่ากับ 100 ความแปรปรวนเท่ากับ 100	68

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย