

ห้องสมุดงานวิจัย สำนักงานคณะกรรมการวิจัยแห่งชาติ



247846

การเปรียบเทียบการรู้จำของลูกโน้มนားคอมพิวเตอร์โรตัมสำหรับการอ่านเขียน
เมื่อการแสดงผลภาพหลังเป็นแบบปกติที่ลูกทิศทาง

นายพงศ์ศักดิ์ สติรุ่งพรชัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติ ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2553

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

b00052107

ห้องสมุดงานวิจัย สำนักงานคณะกรรมการวิจัยแห่งชาติ



247846

การเปรียบเทียบการละลายของลูกโซ่มาร์คอฟมอนติคาร์โลสำหรับการอนุมานเชิงเบย์
เมื่อการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง



นายพงศ์ศักดิ์ สติรุ่งพรชัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติ ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2553

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย



5 1 8 1 8 6 2 4 2 6

A COMPARISON OF MARKOV CHAIN MONTE CARLO CONVERGENCE FOR
BAYESIAN INFERENCE WITH TRUNCATED NORMAL POSTERIOR DISTRIBUTION

Mr.Pongsak Sathitrunpornchai

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Statistics

Department of Statistics

Faculty of Commerce and Accountancy

Chulalongkorn University

Academic Year 2010

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การเปรียบเทียบการรู้เข้าของลูกโซ่มาร์คอฟมอนติคาร์โล
สำหรับการอนุมานเชิงเบย์ เมื่อการแจกแจงภายหลังเป็น
แบบปกติที่ถูกตัดหาง

โดย

นายพงศ์ศักดิ์ สถิตรุ่งพรชัย

สาขาวิชา

สถิติ

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ผู้ช่วยศาสตราจารย์ ดร.เสกสรร เกียรติสุไพบูลย์

คณะพาณิชย์ศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์
ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาโทบริหารธุรกิจ

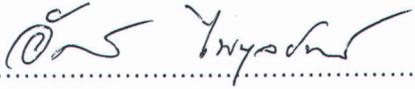

..... คณบดีคณะพาณิชย์ศาสตร์และการบัญชี
(รองศาสตราจารย์ ดร.อรรณพ ต้นละม้าย)

คณะกรรมการสอบวิทยานิพนธ์


..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.ธีระพร วีระถาวร)


..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.เสกสรร เกียรติสุไพบูลย์)


..... กรรมการ
(รองศาสตราจารย์ ดร.สุพล ชรุงศ์วัฒนา)


..... กรรมการภายนอกมหาวิทยาลัย
(อาจารย์ ดร.อัชรินทร์ ไพบูลย์พานิช)

พงศ์ศักดิ์ สถิตรุ่งพรชัย : การเปรียบเทียบการลู่เข้าของลูกโซ่มาร์คอฟมอนติคาร์โล สำหรับการอนุมานเชิงเบย์ เมื่อการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง. (A COMPARISON OF MARKOV CHAIN MONTE CARLO CONVERGENCE FOR BAYESIAN INFERENCE WITH TRUNCATED NORMAL POSTERIOR DISTRIBUTION) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ.ดร.เสกสรร เกียรติสุไพบูลย์, 53 หน้า.

247846

การอนุมานเชิงเบย์เมื่อการแจกแจงก่อนเป็นแบบปกติ และ พารามิเตอร์ที่เป็นค่าคาดหวังมีการเรียงอันดับอย่างสมบูรณ์ จะมีการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง การอนุมานสามารถทำได้โดยใช้วิธีในกลุ่มลูกโซ่มาร์คอฟมอนติคาร์โล การวิจัยนี้จึงทำการเปรียบเทียบประสิทธิภาพของเทคนิคในกลุ่มลูกโซ่มาร์คอฟมอนติคาร์โล 2 วิธี คือ การสุ่มตัวอย่างแบบฮิตแอนด์รัน และ การสุ่มตัวอย่างแบบกิบส์ โดยนำเทคนิคทั้งสองมาแก้ปัญหาการอนุมานเชิงเบย์ที่มีการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง ประสิทธิภาพนิยามโดยค่าครึ่งช่วงความเชื่อมั่นที่คำนวณจากวิธีค่าเฉลี่ยกลุ่ม และค่า MPSRF ของบรู๊คส์-เกลแมน

จากการศึกษาพบว่า ในกรณีที่จำนวนมิติของพารามิเตอร์ที่เป็นค่าคาดหวังมีค่าต่ำ การสุ่มตัวอย่างแบบฮิตแอนด์รัน จะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์ แต่ในกรณีที่จำนวนมิติของพารามิเตอร์ที่เป็นค่าคาดหวังมีค่าสูง การสุ่มตัวอย่างแบบกิบส์ จะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน

ภาควิชาสถิติ..... ลายมือชื่อนิสิต..... พงศ์ศักดิ์ สถิตรุ่งพรชัย
สาขาวิชาสถิติ..... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก..... เสกสรร เกียรติสุไพบูลย์
ปีการศึกษา2553.....

5181862426 : MAJOR STATISTICS

KEYWORDS : CONVERGENCE DIAGNOSTIC / GIBBS SAMPLER / HIT-AND-RUN SAMPLER / MARKOV CHAIN MONTE CARLO

PONGSAK SATHITRUNGPORNCHAI : A COMPARISON OF MARKOV CHAIN MONTE CARLO CONVERGENCE FOR BAYESIAN INFERENCE WITH TRUNCATED NORMAL POSTERIOR DISTRIBUTION.THESIS ADVISOR : ASST. PROF. SEKSAN KIATSUPAIBUL, Ph.D., 53 pp.

217846

In a Bayesian inference, when the prior distribution is a normal distribution and the expectation parameter is completely ordered, the posterior distribution becomes a truncated normal distribution in high dimension. The inference can be performed by employing methods in the class of Markov chain Monte Carlo. The objective of this study is to compare the performances between the two Markov chain Monte Carlo techniques, the hit-and-run sampler and Gibbs sampler, by applying the two techniques to the Bayesian inference problem with a truncated normal posterior distribution. The performance are defined as the half width of the confidence interval formed by the batch means approach and the Brooks-Gelman MPSRF.

From our study, we find that the hit-and-run sampler and the Gibbs sampler perform differently depending on the dimension of the expectation parameter. More specifically, when the dimension of the expectation parameter is low, the hit-and-run sampler works better. On the other hand, the Gibbs sampler performs more effectively when the dimension of the expectation is high.

Department : Statistics Student's Signature Pongsak Sathitrungpornchai
Field of Study : Statistics Advisor's Signature [Signature]
Academic Year : 2010

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความกรุณาและเอาใจใส่อย่างดียิ่งของผู้ช่วยศาสตราจารย์ ดร. เสกสรร เกียรติสุไพบูลย์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งได้ให้คำปรึกษา คำแนะนำ และช่วยเหลือแก้ไขข้อบกพร่อง จนกระทั่งวิทยานิพนธ์ฉบับนี้เสร็จสมบูรณ์ ผู้วิจัยจึงขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.ธีระพร วีระถาวร รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา และ อาจารย์ ดร.อัครินทร์ ไพบูลย์พานิช ในฐานะประธานกรรมการและกรรมการสอบวิทยานิพนธ์ ที่กรุณาตรวจสอบและให้คำแนะนำอันเป็นประโยชน์ในการแก้ไขวิทยานิพนธ์ให้สมบูรณ์ยิ่งขึ้น และขอกราบขอบพระคุณคณาจารย์ทุกท่านที่ได้ให้ความรู้แก่ผู้วิจัยจนกระทั่งสำเร็จการศึกษา

ผู้วิจัยขอกราบขอบพระคุณ บิดา มารดา และครอบครัว ที่ช่วยสนับสนุนในทุกๆ ด้านให้ผู้วิจัยได้มีโอกาสทางการศึกษาและคอยให้กำลังใจเสมอมาจนสำเร็จการศึกษา สุดท้ายนี้ขอขอบคุณเพื่อนๆ ทุกท่าน ที่คอยให้คำแนะนำ และให้กำลังใจในการทำวิทยานิพนธ์ตลอดมา

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ญ
สารบัญภาพ.....	ฎ
บทที่	
1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 ขอบเขตของการวิจัย.....	2
1.4 คำจำกัดความที่ใช้ในการวิจัย.....	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	3
1.6 วิธีดำเนินการวิจัย.....	4
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 แนวคิดและทฤษฎี.....	5
2.1.1 การสุ่มตัวอย่างแบบฮิตแอนดรีน.....	5
2.1.2 การสุ่มตัวอย่างแบบกิบส์.....	6
2.1.3 วิธีค่าเฉลี่ยกลุ่ม.....	6
2.1.4 ค่า MPSRF ของ บรูกซ์-เกลแมน	7
2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	9
3 วิธีดำเนินการวิจัย.....	10
3.1 การจำลองข้อมูลด้วยการสุ่มตัวอย่างแบบฮิตแอนดรีน.....	10
3.2 การจำลองข้อมูลด้วยการสุ่มตัวอย่างแบบกิบส์.....	11
3.3 การคำนวณค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม.....	12
3.4 การคำนวณค่า MPSRF ของ บรูกซ์-เกลแมน.....	13
3.5 การวิเคราะห์และสรุปผลการวิจัย.....	13

บทที่	หน้า
4 ผลการวิจัย.....	14
4.1 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดร์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม.....	14
4.1.1 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดร์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.....	15
4.1.2 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดร์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.5.....	19
4.1.3 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดร์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.9.....	24
4.2 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดร์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรุกซ์-เกลแมน.....	29
4.2.1 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดร์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรุกซ์-เกลแมน เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.....	29
4.2.2 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดร์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรุกซ์-เกลแมน เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.5.....	33
4.2.3 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดร์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรุกซ์-เกลแมน เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.9.....	38
5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ.....	44
5.1 สรุปผลการวิจัย.....	44
5.1.1 สรุปผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดร์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม.....	44

บทที่	หน้า
5.1.2 สรุปผลการเปรียบเทียบประสิทธิภาพของการสู่มั่วอย่างแบบฮิตแอนด์ รันกับการสู่มั่วอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรุกซ์-เกล แมน.....	44
5.2 อภิปรายผลการวิจัย.....	46
5.3 ข้อเสนอแนะ.....	47
รายการอ้างอิง.....	48
ภาคผนวก.....	49
ประวัติผู้เขียนวิทยานิพนธ์.....	53

สารบัญตาราง

ตารางที่		หน้า
4.1	แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนดรีนและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0	15
4.2	แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนดรีนและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.5	20
4.3	แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนดรีนและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.9	24
4.4	แสดงค่า MPSRF ของ บรุกซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนดรีนและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.....	29
4.5	แสดงค่า MPSRF ของ บรุกซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนดรีนและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.5	34
4.6	แสดงค่า MPSRF ของ บรุกซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนดรีนและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.9	39
5.1	แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 และค่า MPSRF ของบรุกซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนดรีนและแบบกิบส์.....	45

สารบัญภาพ

ภาพที่		หน้า
4.1	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 2 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 10,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดรีน และแบบกิบส์	15
4.2	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 3 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 15,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดรีน และแบบกิบส์	16
4.3	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 4 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 20,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดรีน และแบบกิบส์	17
4.4	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 5 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 25,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดรีน และแบบกิบส์	18
4.5	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 7 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 35,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดรีน และแบบกิบส์	18
4.6	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 10 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 50,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดรีน และแบบกิบส์	19
4.7	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 2 มิติ และ $\rho = 0.5$ เมื่อทำการจำลองข้อมูล 20,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดรีน และแบบกิบส์	20
4.8	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 3 มิติ และ $\rho = 0.5$ เมื่อทำการจำลองข้อมูล 45,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดรีน และแบบกิบส์	21
4.9	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 4 มิติ และ $\rho = 0.5$ เมื่อทำการจำลองข้อมูล 80,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดรีน และแบบกิบส์	22

4.30	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 300,000 รอบต่อลูกโซ่ ในกรณี 10 มิติ และ $\rho = 0.5$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดร์ัน และแบบกิบส์.....	38
4.31	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 4,000 รอบต่อลูกโซ่ ในกรณี 2 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดร์ัน และแบบกิบส์.....	40
4.32	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 18,000 รอบต่อลูกโซ่ ในกรณี 3 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดร์ัน และแบบกิบส์.....	40
4.33	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 48,000 รอบต่อลูกโซ่ ในกรณี 4 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดร์ัน และแบบกิบส์.....	41
4.34	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 100,000 รอบต่อลูกโซ่ ในกรณี 5 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดร์ัน และแบบกิบส์.....	42
4.35	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 294,000 รอบต่อลูกโซ่ ในกรณี 7 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดร์ัน และแบบกิบส์.....	42
4.36	แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 300,000 รอบต่อลูกโซ่ ในกรณี 10 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนดร์ัน และแบบกิบส์.....	43