สำนักงานคณะกรรมการวิจัยแห่งชาติ ห้องสมุดงานวิจัย

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 แนวคิดและทฤษฎี

2.1.1 การสุ่มตัวอย่างแบบฮิตแอนด์รัน (Hit-and-Run Sampler)

ขั้นตอนวิธีของการสุ่มตัวอย่างแบบฮิตแอนด์รัน ถูกนำเสนอโดย Smith R.L. โดยเป็นการสร้างลำดับของจุดตัวอย่างซึ่งลู่เข้าสู่การแจกแจงแบบสม่ำเสมอ (uniform distribution) บนบริเวณเปิดที่มีขอบเขตจำกัด $\Re^k$ ใดๆ ซึ่งต่อมาได้ถูกทำให้สามารถสร้างลำดับของจุดตัวอย่างที่มีการแจกแจงที่ต่อเนื่องของหลายตัวแปรใดๆ ได้ (Bélisle C.J.P., Romeijn H.E. and Smith R.L.(1993) และ Romeijn H.E. and Smith R.L.(1994)) โดยการสุ่มตัวอย่างแบบฮิตแอนด์รันนี้จัดว่าเป็นหนึ่งในวิธีที่เร็วที่สุด ในการสร้างจุดตัวอย่างในหลายมิติบนบริเวณรูปร่างนูน (Lovász L.(1998))

สำหรับการแจกแจงเป้าหมาย (target distribution) ทั่วไป การสุ่มตัวอย่างแบบฮิตแอนด์รันจะเป็นดังนี้

ให้ $S \subset \Re^k$ และให้การแจกแจงเป้าหมายมีฟังก์ชันความน่าจะเป็น π บน S

สร้าง $\underset{\sim}{X}_n$ ด้วยขั้นตอนวิธี ดังนี้

กำหนด $n = 0$ และ $\underset{\sim}{X}_n = \underset{\sim}{x}_0 \in S$

1.) ที่ $\underset{\sim}{x} = \underset{\sim}{x}_n \in S$ สุ่มเลือกทิศทาง $\underset{\sim}{d}$ บนพื้นผิวทรงกลมรัศมี 1 หน่วย ลากเส้นตรง L_n ผ่าน $\underset{\sim}{x}$ และตัดกับ S

$$L_n = S \cap \left\{ \underset{\sim}{l} : \underset{\sim}{l} = \underset{\sim}{x} + \lambda \underset{\sim}{d}, \lambda \in \Re \right\}$$

2.) สุ่มเลือกจุด $\underset{\sim}{x}_{n+1}$ บนเส้นตรง L_n ด้วยการแจกแจงแบบ π โดยมีเงื่อนไขบน L_n

3.) เพิ่ม n เป็น $n+1$ และกลับไป 1.)

2.1.2 การสุ่มตัวอย่างแบบกิบส์ (Gibbs Sampler)

การสุ่มตัวอย่างแบบกิบส์ เป็นขั้นตอนวิธีในการสร้างลำดับของจุดตัวอย่างจากฟังก์ชันการแจกแจงร่วม ของตัวแปรสุ่มตั้งแต่ 2 ตัวขึ้นไป ขั้นตอนวิธีของการสุ่มตัวอย่างแบบกิบส์ถูกอธิบายโดย Stuart and Donald Geman ในปีค.ศ.1984

สำหรับการแจกแจงเป้าหมายทั่วไป การสุ่มตัวอย่างแบบกิบส์จะเป็นดังนี้

ให้ S เป็นปริภูมิสถานะของ $\underset{\sim}{X} = \left(X^{(1)},..., X^{(k)}\right)$

ให้ π เป็นการแจกแจงเป้าหมาย ซึ่งมีฟังก์ชันความน่าจะเป็น $\pi'\left(x^{(1)},..., x^{(k)}\right)$

ให้ $\underset{\sim}{X}^{(-i)} = \left(X^{(1)},..., X^{(i-1)}, X^{(i+1)},..., X^{(k)}\right)$ เป็นเวกเตอร์ $\underset{\sim}{X}$ ซึ่งไม่มี $X^{(i)}$

สมมติว่า เราสามารถสุ่มตัวอย่างจาก full conditional distribution $\pi'\left(x^{(i)} \mid \underset{\sim}{x}^{(-i)} \right)$

ขั้นตอนวิธีของการสุ่มตัวอย่างแบบกิบส์มีดังนี้

กำหนด $n = 0$ และเริ่มต้นที่ $\underset{\sim}{X}_0 = \left(X_0^{(1)},..., X_0^{(k)}\right)$

1.) สุ่มตัวอย่าง j แจกแจงแบบสม่ำเสมอ บน $\{1,..., k\}$

2.) สุ่มตัวอย่าง $X_{n+1}^{(j)}$ จาก $\pi'\left(x_n^{(j)} \mid \underset{\sim}{x}_n^{(-j)} \right)$

3.) ให้ $\underset{\sim}{X}_{n+1} = \left(X_n^{(1)},..., X_{n+1}^{(j)},..., X_n^{(k)}\right)$ เพิ่ม n เป็น $n+1$ และกลับไป 1.)

2.1.3 วิธีค่าเฉลี่ยกลุ่ม (Batch Means)

วิธีค่าเฉลี่ยกลุ่ม เป็นวิธีการหนึ่งที่ใช้ประมาณค่าคาดหวัง โดยใช้การคำนวณ Monte Carlo standard error (MCSE) สมมติว่าเราสนใจประมาณค่าคาดหวัง $E_\pi\left(g\left(X\right)\right)$ เมื่อ X มีการแจกแจง π

ขั้นตอนของวิธีค่าเฉลี่ยกลุ่ม มีดังนี้

1.) จำลองลูกโซ่มาร์คอฟ $\{X_i, i = 1,2,..., n\}$ จำนวน $n = ab$ รอบ โดยที่ a และ b เป็นจำนวนเต็ม a เป็นจำนวนกลุ่ม และแต่ละกลุ่มมีจำนวน b

2.) หาค่าเฉลี่ยของตัวอย่าง ในกลุ่มที่ j

$$\overline{y}_j = \frac{\sum_{i=(j-1)b+1}^{jb} g(x_i)}{b}$$

3.) ให้ $\hat{\sigma}_g^2 = \frac{b}{a-1}\sum_{j=1}^{a}\left(\overline{Y}_j - \overline{g}_n\right)^2$ โดยที่ $\overline{g}_n = \frac{\sum_{j=1}^{a}\overline{Y}_j}{a}$

จะได้ว่า ค่าประมาณของ Monte Carlo standard error คือ $\frac{\hat{\sigma}_g}{\sqrt{n}}$

4.) ทำการจำลองจนกระทั่ง $t_{\frac{\alpha}{2},(a-1)}\frac{\hat{\sigma}_g}{\sqrt{n}}$ มีค่าน้อยกว่าความคลาดเคลื่อนที่กำหนด

2.1.4 ค่า MPSRF ของ บรูกซ์-เกลแมน (Brooks–Gelman MPSRF)

การวิเคราะห์ของเกลแมนกับรูบิน (The Gelman-Rubin Diagnostic หรือ GRD) ถูกนำเสนอโดย Gelman and Rubin (1992) เป็นหนึ่งในวิธีที่ได้รับความนิยมที่สุดในการวิเคราะห์การลู่เข้าของลูกโซ่มาร์คอฟมอนติคาร์โล ได้ถูกปรับปรุงใหม่และนำเสนอเพิ่มเติมโดย Brooks and Gelman (1998)

ขั้นตอนของวิธีของเกลแมนกับรูบิน มีดังนี้

1.) ทำการจำลองลูกโซ่มาร์คอฟที่อิสระกันจำนวน m ลูกโซ่(m independent parallel Markov Chains) $m \geq 2$ ซึ่งมีการแจกแจง π แต่ละลูกโซ่มีความยาว $2l$ ดังนั้นต้องทำการจำลองทั้งหมด $2lm$ รอบ Gelman and Rubin (1992) ได้แนะนำว่า การจำลอง l ค่าแรก ให้ตัดทิ้งไป และทำการอนุมานบนการจำลอง l ค่าสุดท้าย สำหรับลูกโซ่ที่ j แทนด้วย $\{X_{1j}, X_{2j},..., X_{lj}\}$ โดยที่ $j = 1,2,..., m$

2.) ให้ $Y_{ij} = g(X_{ij})$ ทำการคำนวณค่า

ความแปรปรวนระหว่างลูกโซ่ (between-chain variance)

$$\frac{B}{l} = \frac{1}{m-1}\sum_{j=1}^{m}\left(\bar{Y}_{.j} - \bar{Y}_{..}\right)^2$$

ค่าเฉลี่ยของความแปรปรวนภายในลูกโซ่ (the average of the m within-chain variance)

$$W = \frac{1}{m}\sum_{j=1}^{m} s_j^2$$

โดยที่ $\bar{Y}_{.j} = \frac{1}{l}\sum_{i=1}^{l} Y_{ij}$, $\bar{Y}_{..} = \frac{1}{m}\sum_{j=1}^{m}\bar{Y}_{.j}$ และ $s_j^2 = \frac{1}{(l-1)}\sum_{i=1}^{l}\left(Y_{ij} - \bar{Y}_{.j}\right)^2$

$\bar{Y}_{..}$ เป็นผลการประมาณแบบจุดของ $E_{\pi}g$

3.) ประมาณค่าความแปรปรวนเป้าหมาย (target variance) โดยการเฉลี่ยถ่วงน้ำหนัก(weighted average) ค่า W และ B ได้เป็นดังนี้

$$\hat{\sigma}_{+}^2 = \frac{l-1}{l}W + \frac{1}{l}B$$

4.) ให้ $\hat{V} = \hat{\sigma}_{+}^2 + \frac{B}{ml} = \frac{l-1}{l}W + \frac{(m+1)B}{ml}$, $d \approx \frac{2\hat{V}^2}{\hat{\text{var}}\left(\hat{V}\right)}$

เมื่อ $\hat{\text{var}}\left(\hat{V}\right) = \left(\frac{l-1}{l}\right)^2 \frac{1}{m}\hat{\text{var}}\left(s_j^2\right) + \left(\frac{m+1}{ml}\right)^2 \frac{2}{m-1}B^2$

$$+ 2\frac{(m+1)(l-1)}{ml^2}\frac{l}{m}\left[\hat{\text{cov}}\left(s_j^2, \bar{Y}_{.j}^2\right) - 2\bar{Y}_{..}\hat{\text{cov}}\left(s_j^2, \bar{Y}_{.j}\right)\right]$$

นิยาม scale reduction factor

$$R = \frac{\hat{V}}{\sigma^2}$$

และประมาณค่า R โดย

$$\hat{R} = \frac{\hat{V}}{W}$$

นิยาม potential scale reduction factor ดังนี้

$$\frac{d}{d-2}\frac{\hat{V}}{W}$$

ต่อมา Brooks and Gelman (1998) ได้ปรับปรุงค่า potential scale reduction factor ใหม่เป็น

$$\hat{R}_c = \frac{d+3}{d+1}\hat{R} = \frac{d+3}{d+1}\frac{\hat{V}}{W}$$

ซึ่งเรียกว่า corrected scale reduction factor

เนื่องจาก ค่า $\hat{R}_c$ เป็นค่าอัตราส่วนระหว่างตัวประมาณของความแปรปรวน 2 ตัว โดยที่ตัวเศษเป็นตัวประมาณของความแปรปรวนที่มีค่าสูงเกินไป (overestimate) และตัวส่วนเป็นตัวประมาณของความแปรปรวนที่มีค่าต่ำเกินไป (underestimate) ดังนั้นการจำลองจะลู่เข้าเมื่อค่า $\hat{R}_c$ มีค่าเข้าใกล้(และมากกว่า) 1

สำหรับในกรณีหลายตัวแปร ซึ่งเราสนใจที่จะประมาณค่าของเวกเตอร์ของพารามิเตอร์ $\underset{\sim}{Y}_{ij}$ บรูกซ์และเกลแมน (Brooks and Gelman (1998)) ได้นำเสนอค่า multivariate potential scale reduction factor หรือ MPSRF ซึ่งเท่ากับ $\max_{\underset{\sim}{a}} \frac{\underset{\sim}{a}^T \hat{V} \underset{\sim}{a}}{\underset{\sim}{a}^T W \underset{\sim}{a}}$ โดยที่

$$\hat{V} = \frac{l-1}{l} W + \left(1 + \frac{1}{m}\right) \frac{B}{l}$$

$$W = \frac{1}{m(l-1)} \sum_{j=1}^{m} \sum_{i=1}^{l} \left(\underset{\sim}{Y}_{ij} - \underset{\sim}{\bar{Y}}_{.j} \right) \left(\underset{\sim}{Y}_{ij} - \underset{\sim}{\bar{Y}}_{.j} \right)^T$$

$$\frac{B}{l} = \frac{1}{m-1} \sum_{j=1}^{m} \left(\underset{\sim}{\bar{Y}}_{.j} - \underset{\sim}{\bar{Y}}_{..} \right) \left(\underset{\sim}{\bar{Y}}_{.j} - \underset{\sim}{\bar{Y}}_{..} \right)^T$$

ซึ่งใช้ตรวจสอบในลักษณะเดียวกับค่า potential scale reduction factor แต่เป็นในรูปแบบของหลายตัวแปร การจำลองจะลู่เข้าเมื่อค่า MPSRF มีค่าเข้าใกล้(และมากกว่า) 1 (ในทางปฏิบัติเราอาจทำการจำลองจนค่า MPSRF มีค่าน้อยกว่า 1.2 จึงจะถือว่าลู่เข้า)

2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

หมิงฮุ่ย เชน และ บรูซ ชไมเซอร์ (Chen M.-H. and Schmeiser B.W., 1993) ได้ศึกษาเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ในกรณีที่เป็นการแจกแจงแบบปกติของ 2 ตัวแปร