บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ในปัจจุบันนี้การจำลอง (Simulation) ได้มีบทบาทเข้ามาเกี่ยวข้องกับการวิเคราะห์ระบบต่างๆเป็นจำนวนมาก ทั้งนี้เนื่องมาจากผู้วิจัยสามารถใช้การจำลองในการทดสอบประสิทธิภาพของระบบได้โดยไม่จำเป็นต้องทดสอบกับระบบจริงๆ ไม่ว่าระบบจะมีความซับซ้อนมากน้อยเพียงใดก็ตาม ซึ่งเป็นการประหยัดเวลาและค่าใช้จ่าย อีกทั้งผลการทดลองที่ได้ยังมีความใกล้เคียงกับสภาพความเป็นจริงอีกด้วย

การวิเคราะห์ระบบโดยวิธีการจำลองนั้น เริ่มต้นด้วยการกำหนดวัตถุประสงค์ของการศึกษาระบบนั้นให้สอดคล้องกับปัญหาที่เกิดขึ้น แล้วจึงทำการสร้างตัวแบบขึ้นมาโดยการสร้างตัวแบบนั้น จะเป็นการสร้างตัวแปรต่างๆที่เกี่ยวข้องกับระบบ มีการกำหนดเงื่อนไขและข้อจำกัดรวมไปถึงการกำหนดความสัมพันธ์ระหว่างตัวแปรต่างๆด้วย เมื่อผู้วิจัยได้ตัวแบบที่มีความถูกต้องแล้วจึงนำตัวแบบนั้นมาทำการทดลองเพื่อหาข้อสรุปเกี่ยวกับปัญหาเหล่านั้นซึ่งประสิทธิผลของการวิเคราะห์ระบบนั้นจะมีความถูกต้องใกล้เคียงกับความเป็นจริงมากน้อยเพียงใดนั้นขึ้นอยู่กับตัวแบบที่สร้างขึ้นมา ดังนั้นผู้วิจัยจึงควรให้ความสำคัญกับการสร้างตัวแบบเป็นอย่างยิ่ง

เนื่องจากระบบที่ทำการจำลองนั้นมักจะประกอบไปด้วยปัจจัยที่มีความไม่แน่นอนซึ่งมีลักษณะแตกต่างกัน แต่มีความสัมพันธ์กัน ดังนั้นการจำลองตัวแบบให้ใกล้เคียงสภาพความเป็นจริงมากที่สุดจึงเป็นการจำลองตัวแปรสุ่มร่วมที่มีลักษณะการแจกแจงร่วม (Joint Distribution) ตรงตามของปัจจัยทั้งหมด แต่เนื่องจากความหลากหลายในลักษณะของปัจจัยจึงทำให้การจำลองตัวแปรสุ่มร่วมตามลักษณะการแจกแจงร่วมนั้นทำได้ยาก ดังนั้นจึงมีการเปลี่ยนมาใช้การจำลองตัวแปรสุ่มร่วมที่มีลักษณะการแจกแจงส่วนริมและมีสหสัมพันธ์ให้ตรงตามลักษณะต่างๆของปัจจัยในระบบแทน เพราะสามารถทำการจำลองได้ง่ายกว่าอีกทั้งยังคงให้ข้อมูลใกล้เคียงกับสภาพความจริงอยู่

Cario และ Nelson[1] ได้เสนอวิธีการ NORTA (NORmal To Anything) เพื่อจำลอง (Generate) ตัวแปรสุ่มร่วม $\mathbf{X} = (X_1, X_2, \ldots, X_k)'$ โดยที่แต่ละ X_i มีการแจกแจงส่วนริม (Marginal Distribution) คือ F_i และมีเมตริกซ์สหสัมพันธ์(Correlation Matrix) = Σ_X ซึ่งเป็นเมตริกซ์ที่มีองค์ประกอบที่ (i, j) เป็น $\rho_{x_i x_j}$ และองค์ประกอบที่ $(i,i) = 1$

วิธีการ NORTA มีขั้นตอนดังนี้

1. จำลองตัวแปรสุ่มร่วมที่มีการแจกแจงร่วมแบบปกติ (Multivariate Normal Vector) d ตัว

$$\mathbf{Z} = (Z_1, Z_2, \ldots, Z_d)'$$

ซึ่งมีเวกเตอร์ค่าเฉลี่ย(Mean Vector) = 0

และมีเมตริกซ์สหสัมพันธ์ = Σ_Z โดยที่มีองค์ประกอบที่ $(i,i) = 1$

2. แปลงตัวแปรสุ่มร่วมที่มีการแจกแจงร่วมแบบปกติ $\mathbf{Z}$ ให้เป็นตัวแปรสุ่มร่วมที่มีการแจกแจงส่วนริมแบบสม่ำเสมอ $\mathbf{U}$ (Multivariate Uniform Vector)

$$\mathbf{U} = (\Phi(Z_1), \Phi(Z_2), \ldots, \Phi(Z_d))'$$

โดยที่ Φ คือฟังก์ชันการแจกแจงแบบสะสม (Cumulative Distribution Function) ของตัวแปรสุ่มปกติและ $\mathbf{U}$ คือ Gaussian หรือ Normal Copula

3. คำนวณหา $\mathbf{X} = (X_1, X_2, \ldots, X_d)'$ ผ่านทาง

$$\mathbf{X} = \begin{pmatrix} F_{X_1}^{-1}[\Phi(Z_1)] \\ F_{X_2}^{-1}[\Phi(Z_2)] \\ \vdots \\ F_{X_k}^{-1}[\Phi(Z_d)] \end{pmatrix}$$

โดยที่ $F_i^{-1}(u) = \inf\{x : F_i(x) \geq u\}$ คือ ฟังก์ชันผกผันของฟังก์ชันการแจกแจงแบบสะสมจากผลการแปลงทำให้ได้ X_i ที่มีการแจกแจงส่วนริมดังที่ต้องการ

ส่วนเมตริกซ์สหสัมพันธ์ $\mathbf{X}$ (Σ_X) นั้นมีความสัมพันธ์กับ Σ_Z ดังนี้

$$\rho_{x_i x_j} = \frac{1}{\sigma_{x_i}\sigma_{x_2}}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_i^{-1}(\Phi(z_i))F_j^{-1}(\Phi(z_j)) \cdot \varphi_{ij}(z_i, z_j)\,dz_i\,dz_j - E(X_i)E(X_j)\right)$$

โดยที่ φ_{ij} คือ ฟังก์ชันความหนาแน่นร่วม (Joint Density Function) ของ (Z_i, Z_j)

ซึ่งความสัมพันธ์นี้สามารถส่ง $\rho_{z_i z_j}$ ไปยัง $\rho_{x_i x_j}$ ได้โดยใช้ฟังก์ชัน c_{ij}

---

[1]Cario,M.C.,and Nelson.B.L ,“ Modeling and Generating Random Vectors with Arbitrary Marginal Distributions and Correlation Matrix,” Technical Report Department of Industrial Engineering and Management Sciences Northwestern University (Evanston, IL.,1997)

$$c_{ij} : [-1,1] \rightarrow \Re$$

การวิเคราะห์โดยฟังก์ชัน c_{ij} นี้ไม่มีรูปแบบทั่วไป แต่เมื่ออยู่ภายใต้เงื่อนไขที่ว่า c_{ij} เป็นฟังก์ชันต่อเนื่องไม่ลดของ $\rho_{z_i z_j}$ จะได้ผลการคำนวณหาค่า Σ_Z จาก

$$c_{ij}(\rho_{z_i z_j}) = \rho_{x_i x_j} \qquad \text{สำหรับ } i < j$$

ซึ่งมีคำตอบเพียงหนึ่งเดียว (Unique) ภายใต้ข้อจำกัดข้อสมมติบนการแจกแจงส่วนริม และถ้า Σ_Z ที่คำนวณหาค่าได้เป็นเมตริกซ์ที่มีคุณสมบัติเป็นเมตริกซ์กึ่งบวกแน่นอนที่มีแต่ละองค์ประกอบมีค่าระหว่าง -1 ถึง 1 แล้วจึงจะสามารถใช้ Σ_Z ในการจำลองตัวแปรสุ่ม $\mathbf{X}$ โดยวิธี NORTA ได้

Li และ Hammond[1] ได้ทำการศึกษาการจำลองตัวแปรสุ่มร่วมมีการแจกแจงส่วนริมแบบสม่ำเสมอ (Uniform) ด้วยวิธี NORTA โดยศึกษาถึงลักษณะการแปลง (Transformation) เมตริกซ์สหสัมพันธ์ที่ใช้ในการจำลอง จากผลการศึกษาทำให้ได้รูปแบบของฟังก์ชันความสัมพันธ์ระหว่าง $\rho_{z_i z_j}$ และ $\rho_{x_i x_j}$ จาก

$$\rho_{x_i x_j} = 12 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi(z_i)\Phi(z_j) \frac{1}{2\pi\sqrt{1-\rho_{z_i z_j}^2}} \exp\left(-\frac{z_i^2 - 2\rho_{z_i z_j} z_i^2 z_j^2 + z_j^2}{2\left(1-\rho_{z_i z_j}^2\right)}\right) \, dz_i \, dz_j \; -3$$

เป็นรูปแบบที่ง่ายต่อการคำนวณยิ่งขึ้นคือ

$$\rho_{z_i z_j} = 2\sin[\frac{\pi}{6}\rho_{x_i x_j}]$$

อีกทั้ง Li และ Hammond ได้ยกตัวอย่างการสร้างตัวแปรสุ่มร่วม $\mathbf{X} = (X_1, X_2, X_3)$ โดยที่แต่ละ X_i มีการแจกแจงส่วนริมแบบสม่ำเสมอ(Uniform) และมีเมตริกซ์สหสัมพันธ์คือ

$$\Sigma_X = \begin{pmatrix} 1 & -0.4 & 0.2 \\ -0.4 & 1 & 0.8 \\ 0.2 & 0.8 & 1 \end{pmatrix}$$

ซึ่งสามารถคำนวณหาค่า Σ_Z ได้จาก

$$\rho_{z_i z_j} = 2\sin[\frac{\pi}{6}\rho_{x_i x_j}]$$

ได้ค่า Σ_z ที่ได้มีเพียงหนึ่ง (Unique) คือ

$$\Sigma_z = \begin{pmatrix} 1 & -0.4158 & 0.2091 \\ -0.4158 & 1 & 0.8135 \\ 0.2091 & 0.8135 & 1 \end{pmatrix}$$

---

[1] Li,S.T,and Hammaond,J.L., " Generation of Pseudorandom Numbers with Specified Univariate Distributions and Correlation Coefficients, " IEEE transactions on Systems,Man,and Cybernetics 5 (1975) : 557-561.

แต่เมตริกซ์ Σ_Z ที่ได้ไม่ใช่เมตริกซ์กึ่งบวกแน่นอน (Non - positive Semidefinite) แสดงว่า Σ_z ไม่ใช่เมตริกซ์สหสัมพันธ์ ดังนั้นจึงไม่สามารถจำลองตัวแปรสุ่มร่วม $\mathbf{X}$ ดังกล่าวด้วยวิธีNORTA ได้

จากตัวอย่างแสดงให้เห็นว่าการจำลองด้วยวิธี NORTA ไม่สามารถทำได้ในทุกกรณีและสิ่งที่ทำให้วิธีการจำลอง NORTA มีปัญหาไม่สามารถใช้ในการจำลองตัวแปรสุ่มร่วมได้นั้นก็คือเมตริกซ์สหสัมพันธ์ที่ไม่เหมาะสม

Ghosh และ Henderson[1] ได้สร้างวิธี Onion ซึ่งเป็นวิธีการสุ่มเมตริกซ์ที่มีการแจกแจงแบบสม่ำเสมอบนเซตของเมตริกซ์กึ่งบวกแน่นอนขึ้นมา เพื่อใช้ทำการทดสอบหาค่าความน่าจะเป็นที่ไม่สามารถใช้วิธี NORTA ในการจำลองตัวแปรสุ่มในมิติต่างๆ ซึ่งผลการทดสอบสรุปได้ว่าเมื่อตัวแปรสุ่มร่วมมีมิติเพิ่มขึ้น ความน่าจะเป็นที่ไม่สามารถใช้วิธี NORTA ในการจำลองก็มีค่ามากขึ้นด้วย

แม้ว่าวิธี NORTA จะเป็นวิธีได้รับการยอมรับเป็นอย่างมากในการจำลองตัวแปรสุ่มร่วมเมื่อทราบการแจกแจงส่วนริมและสหสัมพันธ์ เพราะเป็นวิธีที่ทำได้ง่ายและสะดวก แต่วิธี NORTA ก็ยังไม่สามารถจำลองตัวแปรสุ่มร่วมได้ในทุกกรณี และพบข้อสังเกตว่า ถ้ามีการกำหนดตัวแปรสุ่มร่วมที่ต้องการจำลองให้มีการแจกแจงส่วนริมแบบสม่ำเสมอในทุกๆ มิติ วิธีการ NORTA ก็คือการจำลองโดยใช้เทคนิค Gaussian Copula นั่นเอง และเนื่องจาก Copula มีหลายรูปแบบ ดังนั้นการเปลี่ยนรูปแบบของ Copula จาก Gaussian Copula เป็นรูปแบบอื่นอาจจะสามารถขยายขอบเขตการจำลองตัวแปรสุ่มร่วมเมื่อทราบการแจกแจงส่วนริมและสหสัมพันธ์ได้

Student's t Copula เป็นรูปแบบของ Copula ชนิดหนึ่งที่น่าสนใจว่ามีขอบเขตของการจำลองตัวแปรสุ่มร่วมเป็นอย่างไร ทั้งนี้เนื่องจาก Student's t Copula มีคุณสมบัติความสัมพันธ์ส่วนหาง (Tail Dependent) ซึ่งเป็นลักษณะที่เหมาะสมสำหรับข้อมูลบางชนิด ดังนั้นผู้วิจัยจึงมีความสนใจที่จะศึกษาลักษณะของ Student's t Copula เพื่อตรวจสอบว่า Student's t Copula มีความเหมาะสมสำหรับใช้ในการจำลองตัวแปรสุ่มร่วมมากน้อยเพียงใดรวมถึงการเปรียบเทียบขอบเขตของเมตริกซ์สหสัมพันธ์ที่สามารถใช้ในการจำลองตัวแปรสุ่มร่วมระหว่างเทคนิค Gaussian Copula และ Student's t Copula เพื่อนำมาเป็นแนวคิดสำหรับใช้พัฒนาและขยายขอบเขตการจำลองตัวแปรสุ่มร่วมให้กว้างขึ้น

---

[1]Ghosh,S., and Henderson,S.G., " Properties of the NORTA Method in Higher Dimensions," Proceedings of the Winter Simulation Conference E.Yucesen,C-H Chen,J.L.Snowdon, and J.M. Charnes, eds.(2003)

1.2 วัตถุประสงค์

1. เพื่อศึกษาลักษณะของ Gaussian Copula และ Student's t Copula
2. ทำการจำลองตัวแปรสุ่มร่วมเมื่อทราบการแจกแจงส่วนริมและสหสัมพันธ์โดยใช้เทคนิค Copula ได้
3. เพื่อเปรียบเทียบขอบเขตของเมตริกซ์สหสัมพันธ์ที่สามารถใช้ในการจำลองตัวแปรสุ่มร่วมระหว่างเทคนิค Gaussian Copula และ Student's t Copula

1.3 ขอบเขตการวิจัย

1. ศึกษาเฉพาะตัวแปรสุ่มร่วมที่มีการแจกแจงส่วนริมแบบสม่ำเสมอ
2. ศึกษาตัวแปรสุ่มร่วมมิติตั้งแต่ 2 ถึง 10
3. ศึกษา Student's t Copula ที่มีองศาความเป็นอิสระ 3,4,5,10

1.4 วิธีดำเนินการวิจัย

1. ศึกษาลักษณะของเมตริกซ์สหสัมพันธ์
2. ศึกษาขอบเขตของเมตริกซ์สหสัมพันธ์ที่ใช้สามารถใช้ในการจำลองตัวแปรสุ่มร่วมด้วยเทคนิค Gaussian Copula
3. ศึกษาขอบเขตของเมตริกซ์สหสัมพันธ์ที่ใช้สามารถใช้ในการจำลองตัวแปรสุ่มร่วมด้วยเทคนิค Student's t Copula
4. ศึกษาวิธีการสุ่มเมตริกซ์ที่มีการแจกแจงแบบสม่ำเสมอบนเซตของเมตริกซ์สหสัมพันธ์และบนบริเวณส่วนผิวของเซตของเมตริกซ์สหสัมพันธ์
5. ทำการประมาณค่าสัดส่วนของเมตริกซ์สหสัมพันธ์ที่สามารถทำการจำลองตัวแปรสุ่มร่วมด้วยเทคนิค Gaussian Copula และ Student's t Copula
6. เปรียบเทียบขอบเขตของเมตริกซ์สหสัมพันธ์ที่สามารถใช้ในการจำลองตัวแปรสุ่มร่วมระหว่างเทคนิค Gaussian Copula และ Student's t Copula
7. สรุปผลการวิจัย

1.5 ประโยชน์ที่ได้รับ

1. เกิดความเข้าใจที่ชัดเจนเกี่ยวกับลักษณะของ Copula
2. สามารถเลือกใช้ Copulaในรูปแบบต่างๆให้เหมาะสมกับลักษณะตัวแปรสุ่มร่วมที่ต้องการจำลองได้
3. สามารถประยุกต์เทคนิค Copula ในการจำลองตัวแบบทางสถิติได้