วารสารวิชาการพระจอมเกล้าพระนครเหนือ ปีที่ 27 ฉบับที่ 3 ก.ค.–ก.ย. 2560
The Journal of KMUTNB., Vol. 27, No. 3, Jul.–Sep. 2017

บทความวิชาการ

ปัญหาผกผันและการประยุกต์ในด้านการเงิน

นิฟาตมะห์ มะกาเจ และ อาทิตย์ อินทรสิทธิ์*
ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี

* ผู้นิพนธ์ประสานงาน โทรศัพท์ 0–7331–2179 อีเมล: arthit.i@psu.ac.th DOI: 10.14416/j.kmutnb.2017.06.010
รับเมื่อ 7 เมษายน 2559 ตอบรับเมื่อ 5 สิงหาคม 2559 เผยแพร่ออนไลน์ 27 มิถุนายน 2560

บทคัดย่อ

ปัญหาผกผันเป็นการอธิบายความสัมพันธ์จากผลลัพธ์ไปสู่เหตุของระบบซึ่งตรงกันข้ามกับปัญหาตรง โดยทั่วไป การหาผลเฉลยของปัญหาผกผันด้วยวิธีเชิงตัวเลขจะทำให้เกิดผลเฉลยที่ไร้เสถียรภาพ จึงไม่จัดเป็นปัญหาที่สร้างขึ้นอย่างดีตามหลักเกณฑ์ของ Hadamard การผ่อนปรนกระบวนการหาผลเฉลยด้วยทฤษฎีเรกูลาร์ไรเซชันจะทำให้ผลเฉลยมีเสถียรภาพ ในบทความประมวลความรู้เชิงวิเคราะห์ฉบับนี้ได้กล่าวถึงการวิเคราะห์เชิงคณิตศาสตร์ในการแก้ปัญหาผกผัน ตลอดจนประมวลทฤษฎีเรกูลาร์ไรเซชันที่สำคัญ ท้ายสุดได้ยกตัวอย่างการหาความผันผวนเฉพาะถิ่นในตราสารอนุพันธ์ทางการเงิน ซึ่งเป็นการแก้ปัญหาผกผันในคณิตศาสตร์ประยุกต์ด้านการเงินที่สำคัญปัญหาหนึ่ง

คำสำคัญ: ปัญหาผกผัน, ทฤษฎีเรกูลาร์ไรเซชัน, การตั้งราคาออปชัน, ความผันผวนเฉพาะถิ่น

การอ้างอิงบทความ: นิฟาตมะห์ มะกาเจ และ อาทิตย์ อินทรสิทธิ์, "ปัญหาผกผันและการประยุกต์ในด้านการเงิน," *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 27, ฉบับที่ 3, หน้า 589–601, ก.ค.–ก.ย. 2560

Academic Article

Inverse Problems and Its Applications in Financial Problems

Nifatamah Makaje and Arthit Intarasit*
Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani, Thailand

* Corresponding Author, Tel. 0–7331–2179, E–mail: arthit.i@psu.ac.th DOI: 10.14416/j.kmutnb.2017.06.010
Received 7 April 2016; Accepted 5 August 2016; Published online: 27 June 2017

Abstract

Inverse problems could be described as the relationship between system inputs and outputs. In general, a solution obtained from numerical solving for an inverse problem is said to be unstable. So the inverse problems can be ill-posed problems according to the Hadamand's criteria. By relaxing the process of solving the problem with the corresponding regularization theorem, the stability of the solution will be obtained. In this review article, the mathematical analysis for solving inverse problems and a comprehensive account of regularization theorem are presented. Finally, we show some examples for solving local volatility in financial derivatives which is deemed an important aspect of inverse problems concerning the field of applied mathematics for finance.

Keywords: Inverse Problems, Regularization Theory, Option Pricing, Local Volatility

Please cite this article as: N. Makaje and A. Intarasit, "Inverse problems and its applications in financial problems," *The Journal of KMUTNB.*, vol. 27, no. 3, pp. 589–601, Jul.–Sep. 2017 (in Thai).

1. ปัญหาผกผัน

การอธิบายปรากฏการณ์ตามธรรมชาติหรือการทดลองด้วยตัวแบบเชิงคณิตศาสตร์ก่อให้เกิดประโยชน์และความก้าวหน้าอย่างมากในการศึกษาค้นคว้าและวิจัย การสร้างตัวแบบทางคณิตศาสตร์เป็นการวิเคราะห์เพื่อหาความสัมพันธ์ระหว่างเหตุและผลลัพธ์โดยการกำหนดตัวแปรต่างๆ ถึงแม้ว่าบ่อยครั้งจำเป็นต้องกำหนดข้อสมมติในเชิงอุดมคติที่ไม่เป็นจริงบางประการก็ตาม แต่ตัวแบบที่สร้างขึ้นอย่างดี (Well-posed) จะสามารถอธิบายปรากฏการณ์จริงได้ภายใต้ข้อสมมติเหล่านั้น

แท้จริงแล้วตัวแบบอธิบายถึง 1) ระบบที่มีกระบวนการดำเนินการ 2) เงื่อนไขของการดำเนินการ และ 3) ความสัมพันธ์เชิงปริมาณของอินพุต พารามิเตอร์ และเอาต์พุต รูปที่ 1 แสดงความสัมพันธ์ดังกล่าวนี้

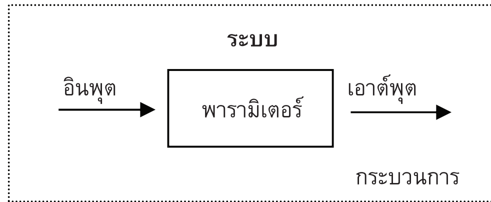
ปัญหาทางกายภาพใดๆ อาจถูกจำแนกออกตามลักษณะของกระบวนการดำเนินการได้เป็น 3 แบบ คือ 1) ปัญหาตรง (Direct Problem) 2) ปัญหาย้อนกลับ (Backward Problem) และ 3) ปัญหาการหาพารามิเตอร์ที่เหมาะสมของตัวแบบ (Parameter Identification Problem)

เมื่อกำหนดให้ X แทนปริภูมิของอินพุต Y แทนปริภูมิของเอาต์พุต P แทนปริภูมิพารามิเตอร์ของระบบ และ $A(p)$ แทนตัวดำเนินการ (Operator) ของระบบจาก X ไปยัง Y ที่เกี่ยวข้องกับพารามิเตอร์ $p \in P$ แล้วจะสามารถอธิบายความสัมพันธ์เชิงปริมาณของ X, Y, P และ $A(p)$ ที่สอดคล้องกับปรากฏการณ์ที่ศึกษาได้โดยอาศัยหลักการทางคณิตศาสตร์ เช่น การวิเคราะห์เชิงจริง พีชคณิต การวิเคราะห์เชิงฟังก์ชัน และการวิเคราะห์เชิงสโตแคสติก เป็นต้น

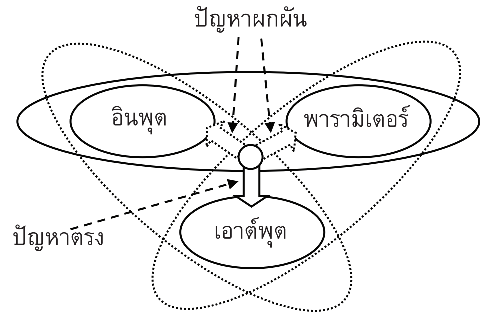
ประพจน์เชิงคณิตศาสตร์สำหรับปัญหาทั้ง 3 แบบข้างต้นตามลำดับ เป็นดังนี้

1) กำหนดให้ $x \in X$ และ $p \in P$ ต้องการหา $y \in Y$ จากความสัมพันธ์ $y = A(p)\,(x)$

2) กำหนดให้ $y \in Y$ และ $p \in P$ ต้องการหา $x \in X$ ที่ทำให้ $y = A(p)\,(x)$



รูปที่ 1 ความสัมพันธ์เชิงปริมาณในกระบวนการ



รูปที่ 2 ธรรมชาติของปัญหาตรงและปัญหาผกผัน

3) กำหนดให้ $x \in X$ และ $y \in Y$ ต้องการหา $p \in P$ ที่ทำให้ $y = A(p)\,(x)$

ปัญหาตรงรู้จักกันในอีกชื่อหนึ่งว่า *ปัญหาไปข้างหน้า* (Forward Problem) เพราะเป็นปัญหาที่มีกระบวนการดำเนินการจากเหตุไปสู่ผลลัพธ์ ส่วนปัญหาย้อนกลับและปัญหาการหาพารามิเตอร์ที่เหมาะสมของตัวแบบ เป็นปัญหาที่มีลักษณะตรงกันข้ามกับปัญหาตรง โดยมีกระบวนการดำเนินการจากผลไปสู่เหตุ โดยพิจารณาให้เหตุเป็นตัวแปรไม่ทราบค่า และให้ผลลัพธ์เป็นตัวทราบค่า จึงเรียกปัญหาลักษณะดังกล่าวนี้ว่า *ปัญหาผกผัน* (Inverse Problem) ในรูปที่ 2 ได้แสดงธรรมชาติของปัญหาตรงและปัญหาผกผัน

ผู้อ่านอาจคิดว่าการหาผลเฉลยของปัญหาตรงแบบ 1) ปัญหาตรง อาจดูง่ายกว่าปัญหาผกผันทั้งแบบ 2) ปัญหาย้อนกลับ และ 3) ปัญหาการหาพารามิเตอร์ที่เหมาะสมของตัวแบบ อย่างไรก็ตามการหาผลเฉลย y ที่สอดคล้องกับ $y = A(p)\,(x)$ ในปัญหาแบบ 1) ปัญหาตรง นั้นอาจต้องหาผลเฉลยของสมการเชิงอนุพันธ์และปริพันธ์ที่ซับซ้อนจนอาจเกิดความยุ่งยากไม่น้อยไปกว่าการแก้ปัญหาผกผันก็ได้

ส่วนใหญ่แล้วในการแก้ปัญหาผกผันใดๆ เป็นการหาผลเฉลยโดยประมาณจากข้อมูลที่เก็บรวบรวมจากการสังเกตโดยปัญหาแบบ 2) ปัญหาย้อนกลับ และ 3) ปัญหาการหาพารามิเตอร์ที่เหมาะสมของตัวแบบ จะแทนข้อมูลที่เก็บรวบรวมจากการสังเกตด้วย y, p และ x, y ตามลำดับ ทั้งนี้การแก้ปัญหาแบบ 2) ปัญหาย้อนกลับ ในภาคปฏิบัตินั้นหากมีข้อมูล y จากการสังเกตเพียงค่าเดียวจะไม่เพียงพอที่ประมาณค่า p ได้

ต่อไปนี้เป็นตัวอย่างของปัญหาการคำนวณปริพันธ์ของฟังก์ชัน $e^{-pt}x(t)$ ในรูปแบบของปัญหาทั้ง 3 แบบดังนี้ โดยที่สัญกรณ์ $\Re$ แทนเซตของจำนวนจริง

1) ปัญหาตรง กำหนดให้ $x:[0,1]\to\Re$ เป็นฟังก์ชันต่อเนื่องและ $p>0$ เป็นพารามิเตอร์ แล้วต้องการคำนวณหา $y(t) = \int_0^t e^{-ps}x(s)\,ds$ สำหรับทุก $t \in [0,1]$

2) ปัญหาย้อนกลับ กำหนดให้ $y : [0,1] \to \Re$ เป็นฟังก์ชันที่หาอนุพันธ์ได้และ $p > 0$ แล้วต้องการหา $x : [0,1] \to \Re$ ซึ่งทำให้

$$y(t) = \int_0^t e^{-ps}x(s)\,ds \text{ สำหรับทุก } t \in [0,1] \qquad (1)$$

เรียกนิพจน์ e^{-ps} ในสมการที่ (1) ว่า *แฟกเตอร์คิดลด* (Discount Factor)

3) ปัญหาการหาพารามิเตอร์ที่เหมาะสม กำหนดให้ $y:[0,1]\to\Re$ เป็นฟังก์ชันที่หาอนุพันธ์ได้และ $x:[0,1]\to\Re$ เป็นฟังก์ชันต่อเนื่อง แล้วต้องการหาพารามิเตอร์ $p > 0$ ที่ทำให้

$$y(t) = \int_0^t e^{-ps}x(s)\,ds \text{ สำหรับทุก } t \in [0,1] \qquad (2)$$

เราสามารถหาผลเฉลยของปัญหาแบบ 1) ปัญหาตรงได้ทั้งในเชิงวิเคราะห์หรือเชิงตัวเลข แต่ในที่นี้จะสนใจการแก้ปัญหาผกผันทั้งแบบ 2) ปัญหาย้อนกลับ และ 3) ปัญหาการหาพารามิเตอร์ที่เหมาะสมเป็นสำคัญ โดยกำหนดให้ $y(0) = 0$ เป็นเงื่อนไขเริ่มต้น

ถ้า $x(t)$ เป็นฟังก์ชันที่มีความต่อเนื่องบนช่วง $[0,1]$ และ $y(t)$ เป็นฟังก์ชันที่นิยามบนช่วง $[0,1]$ โดยทฤษฎีบทพื้นฐานแคลคูลัส ได้ว่า $y(t)$ เป็นฟังก์ชันที่มีความต่อเนื่องบนช่วง $[0,1]$ และหาอนุพันธ์ได้บนช่วง $(0,1)$ และ $y'(t) = e^{-pt}x(t)$ สำหรับทุก $t \in (0,1)$ โดยที่ $y'(t) = dy/dt$ และเมื่อจัดพจน์ในรูปของฟังก์ชัน x ได้ว่าปัญหา 2) ปัญหาย้อนกลับสมมูลกับ

$$x(t) = y'(t)e^{pt},\ t \in (0,1) \qquad (3)$$

ขอให้สังเกตว่าถ้าต้องการคำนวณหาค่าเชิงตัวเลขของ x โดยใช้สมการที่ (3) จะต้องทราบค่า $y'(t)$ ก่อน

ในทำนองเดียวกันนี้ ปัญหาการหาพารามิเตอร์ที่เหมาะสม จะสมมูลกับการหาพารามิเตอร์ p ที่ทำให้

$$y'(t) = e^{-pt}x(t) \text{ สำหรับทุก } t \in [0,1] \qquad (4)$$

2. ปัญหาผกผันในนามธรรม

ในคณิตศาสตร์นามธรรม ตัวแบบเชิงคณิตศาสตร์ หมายถึง *การส่ง* (Mapping) จากเหตุไปยังผลลัพธ์

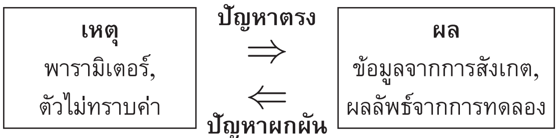
$F : X \to Y$

โดยที่ X และ Y เป็นเซตของเหตุและเซตของผลลัพธ์ตามลำดับส่วนคำว่า “ตัวดำเนินการ” ที่ได้กล่าวแล้วก่อนหน้านี้ หมายถึง การส่งจากปริภูมิหนึ่งไปยังอีกปริภูมิหนึ่ง

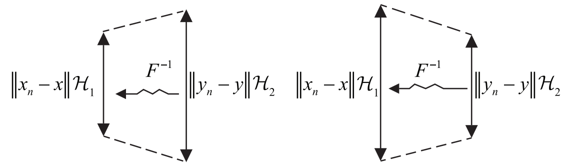
ปัญหาตรง หมายถึง การหาค่า y โดยการแทนค่า $x \in X$ ลงใน $F(x)$ โดยที่ F เป็นตัวดำเนินการที่ส่งจากปริภูมิทอพอโลยี (Topological Space) X ไปยังปริภูมิทอพอโลยี Y โดยทั่วไปแล้ว F มีความต่อเนื่องและ X และ Y เป็นปริภูมิฮิลเบิร์ต (Hilbert Space) ที่มี $\|\cdot\|_X$ และ $\|\cdot\|_Y$ เป็นนอร์มบน X และ Y ตามลำดับ

ปัญหาผกผัน หมายถึง การหาค่า $x \in X$ เมื่อกำหนด $y \in Y$ ที่ทำให้ $F(x) = y$ ลักษณะของปัญหาตรงและปัญหาผกผันแสดงในรูปที่ 3

เพื่อความสะดวกจะพิจารณาสมการเชิงเส้นในปริภูมินอร์มเท่านั้น สำหรับสมการไม่เชิงเส้นในปริภูมิอิงระยะทาง (Metric Space) หรือปริภูมิทอพอโลยีใดๆ จะนิยามได้ในทำนองเดียวกัน



รูปที่ 3 ปัญหาตรงและปัญหาผกผัน



รูปที่ 4 เสถียรภาพของผลเฉลย

พิจารณาสมการตัวดำเนินการ ดังนี้

$$F(x) = y \tag{5}$$

โดยที่ F เป็นตัวดำเนินการจากปริภูมิฮิลเบิร์ต $\mathcal{H}_1$ ไปยังปริภูมิฮิลเบิร์ต $\mathcal{H}_2$ บนฟิลด์ (Field) $\mathbb{F}$

ผลเฉลยของปัญหาผกผันของสมการที่ (5) อยู่ในรูป

$$x = F^{-1}(y) \tag{6}$$

โดยที่ F^{-1} เป็นตัวดำเนินการผกผันของ F

จะกล่าวว่า ปัญหาผกผันของสมการที่ (5) หรือการหาผลเฉลยผกผันในรูปสมการที่ (6) เป็นปัญหาที่ตั้งขึ้นมาอย่างดีตามหลักเกณฑ์ของ Hadamard ถ้าสอดคล้องเงื่อนไขทั้ง 3 ประการต่อไปนี้

1) ผลเฉลย x มีจริง (Existence) หรือตัวดำเนินการผกผัน F^{-1} มีจริง นั่นคือ สำหรับแต่ละค่าของ $y \in \mathcal{H}_2$ แล้วจะมี $x \in \mathcal{H}_1$ ที่สอดคล้องกับสมการที่ (5)

2) ผลเฉลยมีเพียงหนึ่งเดียว (Uniqueness) นั่นคือสำหรับแต่ละค่าของ $y \in \mathcal{H}_2$ แล้วจะมี $x \in \mathcal{H}_1$ เพียงตัวเดียวที่สอดคล้องกับสมการที่ (5) กล่าวคือ ถ้ามี $x_1 = F^{-1}(y)$ และ $x_2 = F^{-1}(y)$ เมื่อ $x_1, x_2 \in \mathcal{H}_1$ แล้ว $x_1 = x_2$

3) ผลเฉลยมีเสถียรภาพ (Stability) นั่นคือ ถ้า $y_n \to y$ ลู่เข้าใน $\mathcal{H}_2$ โดยที่ $F(x_n) = y_n$ และ $F(x) = y$ แล้ว $x_n \to x$ ลู่เข้าใน $\mathcal{H}_1$ ดังนั้นผลเฉลย $x \in \mathcal{H}_1$ ของสมการที่ (5) ขึ้นกับ $y \in \mathcal{H}_2$ อย่างต่อเนื่อง

การลู่เข้าของ $x_n \to x$ ใน $\mathcal{H}_1$ หมายถึง $\| x_n - x \|_{\mathcal{H}_1} \to 0$ และการลู่เข้าของ $y_n \to y$ ใน $\mathcal{H}_2$ หมายถึง $\| y_n - y \|_{\mathcal{H}_2} \to 0$

แท้จริงแล้วเกณฑ์ข้อที่ 2 ของ Hadamard จะเกิดขึ้นเมื่อตัวดำเนินการ F เป็นการส่งแบบหนึ่งต่อหนึ่ง (Injective Mapping) หลักเกณฑ์ข้อที่ 1 และ 2 ทำให้ได้ว่าตัวดำเนินการผกผัน F^{-1} มีจริงและมีโดเมน $D(F^{-1})$ (หรือพิสัย $R(F)$ ของตัวดำเนินการ F) เป็นสับเซตของ $\mathcal{H}_2$ ซึ่งจะได้ว่า F เป็นตัวดำเนินการแบบทั่วถึง (Bijective Operator)

เกณฑ์ข้อที่ 3 มีความหมายว่า การส่งผกผัน $F^{-1} : \mathcal{H}_2 \to \mathcal{H}_1$ มีความต่อเนื่อง ซึ่งมีความหมายว่า หากมีการเปลี่ยนแปลงเพียงเล็กน้อยใน y จะส่งผลให้ผลเฉลย x เกิดการเปลี่ยนแปลงเพียงเล็กน้อยเท่านั้น

ถ้ามีเงื่อนไขข้อใดข้อหนึ่งไม่เป็นจริง จะกล่าวว่าปัญหาผกผันนั้นตั้งขึ้นอย่างบกพร่อง (Ill-posed)

รูปที่ 4 แสดงแนวคิดเรื่องเสถียรภาพของผลเฉลย (ด้านซ้าย) และไร้เสถียรภาพของผลเฉลย (ด้านขวา) เมื่อขนาดของลูกศรแทนขนาดของความคลาดเคลื่อนที่เกิดขึ้น

ผลเฉลยจะมีเสถียรภาพหรือไร้เสถียรภาพนั้นขึ้นกับปริภูมิ $\mathcal{H}_1$ ของผลเฉลยด้วย ซึ่งรวมถึงการเลือกนอร์มในปริภูมินั้น ด้วยเหตุนี้จึงกล่าวได้ว่า ปัญหาผกผันอาจเป็นปัญหาที่ตั้งขึ้นอย่างดีหรือบกพร่องขึ้นกับนอร์มในปริภูมิของผลเฉลย โดยปกติการเลือกนอร์มจะขึ้นกับการประยุกต์ในปัญหาที่ศึกษา

3. ตัวอย่างการพิจารณาหลักเกณฑ์ Hadamard

ตัวดำเนินการเชิงเส้น $A : \mathcal{H}_1 \to \mathcal{H}_2$ หมายถึง การส่งเชิงเส้นจาก $\mathcal{H}_1$ ไปยัง $\mathcal{H}_2$ ที่ซึ่ง $A(u+v) = \alpha u + \beta v$ สำหรับทุก $u, v \in \mathcal{H}_1$ และ $\alpha, \beta \in \mathbb{F}$ ซึ่งเป็นฟิลด์ แล้วนอร์มของตัวดำเนินการ (Operator Norm) ของ A นิยามดังนี้

$$\| A \| = \sup_{\|x\| \le 1} \| Ax \|$$

ถ้า $\| A \| < \infty$ แล้วจะกล่าวว่า A เป็นตัวดำเนินการเชิงเส้นที่มีขอบเขต นั่นคือถ้ามีค่าคงที่ C ที่ทำให้ $\| Ax \| \le C \| x \|$ สำหรับทุกค่า x

กำหนดให้ $A : l_2 \to l_1$ เป็นการส่งเชิงเส้นโดยที่ l_2 แทนเซตของลำดับของจำนวนจริง ซึ่งนิยามดังนี้

$$l_2 = \{(x_n)_{n\in\mathbb{N}} \,|\, \|(x_n)\|_2 = \left(\sum_{n=1}^{\infty} |x_n|^2\right)^{\frac{1}{2}} < \infty\}$$

พิจารณาการแก้ปัญหา

$A(x) = y$

โดยกำหนด $x = (\xi_j) \in l_2$ และ $y = (\eta_j) \in l_2$ โดยที่ $\eta_j = \xi_j / j$ สำหรับ $j \in \mathbb{N}$ เมื่อ $x \in l_2$ ได้ว่า

$$Ax = (\xi_1, \frac{1}{2}\xi_2, \frac{1}{3}\xi_3, \ldots, \frac{1}{k}\xi_k, \ldots)_{k\in\mathbb{N}}$$

เพราะว่า

$$\| Ax \|_2 = \left(\sum_{k=1}^{\infty} \frac{1}{k^2}\xi_k^{\,2}\right)^{\frac{1}{2}} \le \left(\sum_{k=1}^{\infty} \xi_k^2\right)^{\frac{1}{2}} = \| x \|_2$$

จึงทำให้ได้ว่า $\| Ax \|_2 < \infty$ เพราะว่า $x \in l_2$ ดังนั้น $\| x \|_2 < \infty$ สำหรับทุก $x \in l_2$ ดังนั้น A จึงเป็นตัวดำเนินการที่มีขอบเขต

พิจารณาตัวดำเนินการผกผันของ A สำหรับ $y = (\eta_j) \in l_2$ ที่นิยามดังนี้

$$x = A^{-1}(y) = (\eta_1, 2\eta_2, 3\eta_3, \ldots) \tag{7}$$

ต่อไปจะพิจารณาว่า ปัญหาผกผันสมการที่ (7) เป็นปัญหาที่สร้างขึ้นมาอย่างดีตามหลักเกณฑ์ของ Hadamard หรือไม่ดังนี้

1) พิจารณาการมีจริงของผลเฉลยดังนี้ เมื่อเลือก $y = (1, \frac{1}{2}, \frac{1}{3}, \ldots) \in l_2$ และแทนลงในสมการที่ (7) ได้ว่า

$$x = A^{-1}(y) = A^{-1}(1, \frac{1}{2}, \frac{1}{3}, \ldots) = (1,1,1,\ldots) \notin l_2$$

เพราะว่า $\sum_{k=1}^{\infty} |x_k|^2$ เป็นอนุกรมลู่ออกที่ไม่สามารถหาผลบวกได้ ดังนั้น สำหรับ $y \in l_2$ บางตัวแล้วปัญหาสมการที่ (7) ไม่มีผลเฉลย $x \in l_2$

2) ถ้าผลเฉลยมีจริงแล้วความเป็นหนึ่งเดียวของผลเฉลยพิจารณาได้ดังนี้ สมมติว่ามี $\tilde{x} = (\tilde{\xi}_j) \in l_2$ และ $\hat{x} = (\hat{\xi}_j) \in l_2$ เป็นผลเฉลยของสมการที่ (7) ที่ซึ่ง

$\tilde{x} = A^{-1}(y)$ และ $\hat{x} = A^{-1}(y)$

จึงได้ว่า $\tilde{\xi}_j = j\eta_j$ และ $\hat{\xi}_j = j\eta_j$ สำหรับแต่ละ $j \in \mathbb{N}$ ดังนั้น $\tilde{x} = \hat{x}$ นั่นคือ ผลเฉลยของสมการที่ (7) มีเพียงหนึ่งเดียว

3) พิจารณาเสถียรภาพของผลเฉลยดังนี้

กำหนดให้ $y = (\eta_j)$ โดยที่ $\eta_j = 0$ ทุกค่า j และให้ $y^{\delta} = (\eta_j^{\delta})$ โดยที่ $\eta_j^{\delta} = 0$ เมื่อ $j \neq n$ และ $\eta_j^{\delta} = 1/\sqrt{n}$ เมื่อ $j = n$

เมื่อแทนค่า y^{δ} ลงในสมการที่ (7) ได้ว่า

$$\begin{aligned} x^{\delta} &= A^{-1}(0,0,\ldots,0,\tfrac{1}{\sqrt{n}},0\ldots,0,\ldots) \\ &= (0,0,\ldots,0,\sqrt{n},0\ldots,0,\ldots) \end{aligned}$$

และแทนค่า y ลงในสมการที่ (7) ได้ว่า

$$x = A^{-1}(0,0,\ldots,0,\ldots,0,\ldots) = (0,0,\ldots,0,\ldots,0,\ldots)$$

คำนวณ

$\| y - y^{\delta} \|_2 = \frac{1}{\sqrt{n}}$ จะมีค่าลดลงเมื่อ n มีค่ามากขึ้น และ

$\| x - x^{\delta} \|_2 = \sqrt{n}$ จะมีค่ามากขึ้นเมื่อ n มีค่ามากขึ้น ดังนั้น

$$\lim_{n\to\infty} \| y - y^{\delta} \|_2 = \lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0 \tag{8}$$

และ

$$\lim_{n\to\infty} \| x - x^{\delta} \|_2 = \lim_{n\to\infty} \sqrt{n} = \infty \tag{9}$$

การเปลี่ยนแปลงค่าเพียงเล็กน้อยสมการที่ (8) ของอินพุต y^{δ} ในตัวดำเนินการ A^{-1} ก่อให้เกิดการเปลี่ยนแปลงอย่างมากสมการที่ (9) ของเอาต์พุต x^{δ} กล่าวคือ ผลเฉลยที่ไร้เสถียรภาพ

เมื่อพิจารณาตามหลักเกณฑ์ของ Hadamard แล้วปัญหาผกผันสมการที่ (7) เป็นปัญหาที่ตั้งขึ้นอย่างบกพร่อง อย่างไรก็ตามหากว่ามีเงื่อนไขข้อใดข้อหนึ่งไม่เป็นจริงตามหลักเกณฑ์ของ Hadamard จะเป็นการเพียงพอที่จะกล่าวได้ว่า เป็นปัญหาที่ตั้งขึ้นอย่างบกพร่อง

4. ความไร้เสถียรภาพในปัญหาผกผัน

Engl และคณะ [1] ได้กล่าวว่า ถ้าในขั้นตอนการหาผลเฉลยของปัญหาผกผันมีการคำนวณอนุพันธ์ของ

ข้อมูลนำเข้าแล้วปัญหาผกผันนั้นอาจเป็นปัญหาที่ตั้งขึ้นอย่างบกพร่องในแง่ที่ผลเฉลยที่ไร้เสถียรภาพ กล่าวคือ ผลเฉลยจะมีการเปลี่ยนแปลงอย่างไร้ขอบเขตเมื่อมีการเปลี่ยนแปลงเพียงเล็กน้อยในข้อมูล

ต่อไปจะพิจารณาเสถียรภาพของผลเฉลยในปัญหาสมการที่(3) และสมการที่ (4) โดยแบ่งเป็นสองกรณี กรณีแรกเมื่อข้อมูลนำเข้า y เป็นฟังก์ชันค่าจริง และกรณีที่สองเมื่อข้อมูลนำเข้า y เป็นข้อมูลแบบดีสครีต ดังนี้

กรณีที่ 1 พิจารณา $y \in C^1[0,1]$ ซึ่งเป็นเซตของฟังก์ชันค่าจริงที่ต่อเนื่องบน [0,1] และอนุพันธ์ของฟังก์ชัน $y(t)$ หาค่าได้และมีความต่อเนื่องบนช่วง [0,1] และ $x \in C[0,1]$

กำหนดให้ $\delta \in (0,1)$, $k \in \mathbb{N}$ และกำหนดให้ y^δ แทนข้อมูลที่มีความคลาดเคลื่อนซึ่งเกิดจากสัญญาณรบกวน δ ดังนี้

$$y^\delta(t) = y(t) + \sqrt{2}\delta\sin(2\pi kt), t \in [0,1]$$

จึงได้

$$(y^\delta)'(t) = y'(t) + 2\sqrt{2}\delta\pi k\cos(2\pi kt), t \in [0,1]$$

เมื่อพิจารณาค่าคลาดเคลื่อนของข้อมูลภายใต้นอร์ม

$$\| \phi; C^m[0,1] \|_{C^m} = \max_{0\le|\alpha|\le m} \sup_{t\in(0,1)} \left| \frac{d^\alpha}{dt^\alpha}\phi(t) \right|$$

โดยที่ α เป็นจำนวนเต็ม ได้ว่า

$$\begin{aligned} \| y^\delta - y \|_{C^1} &= \max\{\sup_{t\in(0,1)} | y^\delta - y |, \sup_{t\in(0,1)} | y^{\delta\prime} - y' |\} \\ &= \max\{\sup_{t\in(0,1)} |\sqrt{2}\delta\sin(2\pi kt)|, \\ &\quad \sup_{t\in(0,1)} |2\sqrt{2}\delta\pi k\cos(2\pi kt)|\} \\ &= 2\sqrt{2}\pi k\delta \end{aligned} \tag{10}$$

และเมื่อพิจารณาค่าคลาดเคลื่อนของผลเฉลยภายใต้นอร์ม

$$\| x; C[0,1] \|_C = \sup_{t\in(0,1)} | x(t) |$$

ได้ว่า

$$\begin{aligned} \| x^\delta - x \|_C &= \sup_{t\in(0,1)} | y^{\delta\prime}(t) - y'(t) \| e^{pt} | \\ &= 2\sqrt{2}\pi k\delta M \end{aligned} \tag{11}$$

โดยที่ $M \ge |e^{pt}|$, $t \in [0,1]$ ซึ่ง M เป็นค่าคงตัวที่อิสระไม่ขึ้นกับข้อมูลแต่ M จะมีค่ามากหรือน้อยขึ้นกับค่าพารามิเตอร์ p

จากสมการที่ (10) และสมการที่ (11) ได้ว่า

$$\| x^\delta - x \|_C = M \| y^\delta - y \|_{C^1}$$

เนื่องจาก M เป็นค่าคงที่ ซึ่งขึ้นกับค่าพารามิเตอร์ p ดังนั้น $\|x^\delta - x\|_C$ จึงมีความสัมพันธ์โดยตรงกับ $\|y^\delta - y\|_{C^1}$ กล่าวคือ หากมีการเปลี่ยนแปลงของข้อมูล ($\| y^\delta - y \|_{C'}$) แต่เพียงเล็กน้อยแล้วจะเกิดเปลี่ยนแปลงของผลเฉลย ($\| x^\delta - x \|_C$) แต่เพียงเล็กน้อยด้วยเช่นกัน นั่นคือ เสถียรภาพของผลเฉลยโดยประมาณขึ้นกับค่า M จะเรียกเสถียรภาพลักษณะเช่นนี้ว่า ความเสถียรแบบมีเงื่อนไข (Conditionally Stable)

กรณีที่ 2 ในทางปฏิบัติ y^δ เป็นข้อมูลจากการสังเกตที่เก็บรวบรวมไว้นั้นเป็นข้อมูลแบบดีสครีต เมื่อพิจารณา $y \in L^2(\Omega)$ ทั้งนี้ปริภูมิ L^p [2] มีนิยามดังนี้

สำหรับ $0 < p < \infty$ กำหนดให้ $L^p(\mu)$ เป็นเซตของฟังก์ชันค่าเชิงซ้อน f ที่นิยามบน X โดยที่ f เป็นฟังก์ชันเมเชอเรเบิล ซึ่ง $\int_X | f |^p d\mu < \infty$ ถ้า $f \in L^p(\mu)$ โดยนิยาม $\| f \|_p = (\int_X | f |^p d\mu)^{1/p}$ ซึ่งเรียกว่า แอลพีนอร์มของ f

ถ้า μ เป็นเมเชอร์การนับ (Counting Measure) บนเซต A ซึ่งเป็นเซตนับได้ จะใช้สัญลักษณ์ l_p แทน $L^p(\mu)$ โดยสมาชิกของ l_p เป็นลำดับของจำนวนเชิงซ้อน $x = (x_n)$ ที่ซึ่ง $\sum_{n=1}^\infty | x_n |^p < \infty$

เมื่อพิจารณาค่าคลาดเคลื่อนของข้อมูลในปัญหาสมการที่ (3) และสมการที่ (4) ในปริภูมิ $L^2[0,1]$ พบว่า

$$\begin{aligned} \| y^\delta - y \|_{L^2} &= \| \sqrt{2}\delta\sin(2\pi kt) \|_{L^2} \\ &= \sqrt{2}\,|\delta| \left(\int_0^1 \sin^2(2\pi kt)dt \right)^{\frac{1}{2}} \\ &= \frac{|\delta|}{\sqrt{\pi k}} \left(\frac{2\pi k}{2} - \frac{\sin(4\pi k)}{4} \right)^{\frac{1}{2}} = |\delta| \end{aligned}$$

ทั้งนี้เพราะ $\sin(4\pi k) = 0$ เมื่อ $k = 1,2,3,...$ และถ้าให้ $\| e^{pt} \|_{L^2} = M$ จะได้ว่า

$$\begin{aligned} \| x(t) - \tilde{x}(t) \|_{L^2} &= \| y'(t)e^{pt} - y^{\delta\prime}(t)e^{pt} \|_{L^2} \\ &= \| y'(t) - y^{\delta\prime}(t) \| \cdot \| e^{pt} \|_{L^2} \\ &= \| \sqrt{2}(2\pi\delta k)\cos(2\pi kt) \|_{L^2}\, M \end{aligned}$$

$$
\begin{aligned}
&= 2\sqrt{2}\pi k\,|\delta|\left(\int_0^1 \cos^2(2\pi kt)\,dt\right)^{\frac{1}{2}} M \\
&= 2\sqrt{\pi k}\,|\delta|\left(\left(\frac{2\pi k}{2}+\frac{\sin 4\pi k}{4}\right)-\left(\frac{0}{2}+\frac{0}{4}\right)\right)^{\frac{1}{2}} M \\
&= 2\pi k\,|\delta|\,M
\end{aligned}
$$

เห็นได้ว่า ถึงแม้ว่า δ เป็นจำนวนจริงที่มีค่าน้อยมาก แต่ $k \in \mathbb{N}$ ซึ่งอาจทำให้ δk เป็นจำนวนจริงที่มีค่ามาก ซึ่งทำให้สมการที่ (1) และสมการที่ (2) เป็นปัญหาบกพร่องในแง่ที่ว่า ผลเฉลยที่ไร้เสถียรภาพ กล่าวคือการเปลี่ยนแปลงของข้อมูล ($\|y - y^{\delta}\|_{L^2} = \delta$) แต่เพียงเล็กน้อย จะทำให้เกิดการเปลี่ยนแปลงของผลเฉลย ($\|x - \tilde{x}\|_{L^2} = 2\pi k|\delta|M$) อย่างมากมายนั้นเอง

5. การหาผลเฉลยของปัญหาผกผัน

เราทราบแล้วว่าปัญหาผกผันต้องการหาเหตุ x ที่ทำให้เกิดผลลัพธ์ y นั้นคือ ต้องการหาค่า x จาก

$$y = F(x) \tag{12}$$

เมื่อทราบค่าของ y ในความเป็นจริงแล้วผลลัพธ์ที่บันทึกเก็บข้อมูลได้ย่อมมีค่าคลาดเคลื่อนแฝงอยู่ด้วยเสมอ เมื่อให้ δ แทนสัญญาณรบกวนที่ทำให้ข้อมูลเกิดความคลาดเคลื่อนแล้วปัญหาสมการที่ (12) จะอยู่ในรูปดังนี้

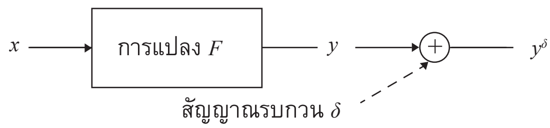
$$y^{\delta} = F(x) + \delta \tag{13}$$

โดยที่ทราบค่า y^{δ} รูปที่ 5 แสดงความสัมพันธ์ของตัวแปรต่างๆ ในปัญหาสมการที่ (13)

ในที่นี้จะพิจารณาการแก้ปัญหาผกผันของสมการที่ (13) ในกรณีเมื่อ F^{-1} เป็นการส่งที่มีความต่อเนื่อง อย่างไรก็ตามในการแก้ปัญหาเชิงตัวเลขจะแปลงปัญหาสมการที่ (13) ไปเป็นปัญหาที่มีมิติจำกัด การแปลงปัญหาดังกล่าวนี้เรียกว่า การดีสครีไทเซชัน (Discretization)

กำหนดให้ $x = (x_1, \dots, x_n)$ และ $y^{\delta} = (y_1^{\delta}, \dots, y_n^{\delta})$ แล้วปัญหาสมการที่ (13) จะอยู่ในรูปดังนี้

$$F\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) + \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_n \end{pmatrix} = \begin{pmatrix} y_1^{\delta} \\ \vdots \\ y_n^{\delta} \end{pmatrix}$$



รูปที่ 5 ความสัมพันธ์ของตัวแปรในปัญหาสมการที่ (13)

หรืออยู่ในรูป

$$F(x^T) + \delta^T = (y^{\delta})^T$$

โดยที่ x^T หมายถึงทรานสโพสของเมทริกซ์ x และ $\delta = (\delta_1, \dots, \delta_n)$ ซึ่งจะสามารถหาผลเฉลยโดยประมาณของปัญหาได้โดยวิธีกำลังสองน้อยสุดแบบไม่เชิงเส้น ดังนี้ $x^* = \min_x \sum_{i=1}^{n} |F(x_i) - y_i^{\delta}|^2$ นั่นคือ จะหาผลเฉลย x ที่ทำให้นิพจน์ $\sum_{i=1}^{n} |F(x_i) - y_i^{\delta}|^2$ มีค่าน้อยที่สุดเมื่อทราบค่า $y_i^{\delta}, i = 1, \dots, n$

ในกรณีทั่วไป ถ้าผลเฉลยของปัญหาผกผันไม่มีจริงนั้น เรายังคงสนใจหาผลเฉลยโดยประมาณที่มีความหมายบางประการ ถ้าผลเฉลยของปัญหาผกผันไม่มีเพียงหนึ่งเดียว เราอาจต้องเพิ่มเติมเงื่อนไขบางประการเพื่อหาผลเฉลยที่มีลักษณะเฉพาะ ในกรณีทั่วไปบทนิยามของ*ผลเฉลยโดยประมาณที่ดีที่สุด* (Best Approximate Solution) ของตัวดำเนินการผกผันเป็นดังนี้

จะกล่าวว่า $x \in X$ เป็นผลเฉลยกำลังสองน้อยที่สุด (Least-Squares Solution) ของสมการ $Ax = y$ ก็ต่อเมื่อ $\|Ax - y\| = \inf\{\|Az - y\| \mid z \in X\}$

ผลเฉลยกำลังสองน้อยที่สุด $x \in X$ ของสมการ $Ax = y$ เป็นผลเฉลยโดยประมาณที่ดีที่สุด ถ้ามีนอร์มต่ำสุดดังนี้ $\|x\| = \inf\{\|z\| \mid z$ คือผลเฉลยที่ให้กำลังสองน้อยที่สุด$\}$

ตัวดำเนินการที่ส่ง y ไปเป็นผลเฉลยโดยประมาณที่ดีที่สุดของสมการ $Ax = y$ คือ ตัวดำเนินการผกผัน Moore – Penrose ซึ่งมีนิยามดังนี้

กำหนดให้ $A: X \to Y$ เป็นตัวดำเนินการเชิงเส้นโดยที่ $N(A) = \{x \in X \mid Ax = 0\}$ เป็นปริภูมินัล (Null Space) และ $R(A) = \{Ax \mid x \in X\}$ เป็นปริภูมิเรนจ์ (Range Space) ของตัวดำเนินการ A แล้วตัวดำเนินการผกผัน Moore-Penrose คือ

$$A^{\dagger} : R(A) \oplus R(A)^{\perp} \to N(A)^{\perp}$$

โดยที่ $N(A)^{\perp} = \{z \in X \mid \langle x, z \rangle = 0,\ \forall x \in N(A)\}$ คือส่วนเติมเต็มเชิงตั้งฉากของปริภูมินัลของ A และ

$R(A)^{\perp} = \{w \in Y \mid \langle y, w \rangle = 0,\ \forall y \in R(A)\}$

คือส่วนเติมเต็มเชิงตั้งฉากของปริภูมิเรนจ์ ของ A โดยที่ $\oplus$ คือ ผลบวกตรง (Direct Sum) และ $\langle \cdot, \cdot \rangle$ แทนผลคูณภายใน (Inner Product)

ทฤษฎีบทต่อไปนี้จะยืนยันได้ว่า จะสามารถหาผลเฉลยที่มีค่านอร์มต่ำสุดได้จากตัวดำเนินการ $A^{\dagger}$

ทฤษฎีบท [1] แต่ละ $y \in D(A^{\dagger})$ ปัญหา $Ax = y$ มีผลเฉลยที่ค่านอร์มมีค่าต่ำสุดเพียงหนึ่งเดียว คือ $x^{\dagger} := A^{\dagger} y$ และ $\{x^{\dagger}\} \cup N(A)$ เป็นเซตของผลเฉลยกำลังสองน้อยที่สุด

สำหรับปัญหาที่ไม่สอดคล้องกับเงื่อนไขข้อที่สามตามหลักเกณฑ์ Hadamard มักเกิดขึ้นในการแก้ปัญหาเชิงตัวเลขโดยเฉพาะ วิธีเชิงตัวเลขพื้นฐานที่รู้จักกันอย่างดีอาจทำให้เกิดผลเฉลยที่ไร้เสถียรภาพ ในการผ่อนปรนความไร้เสถียรภาพของผลเฉลยจะอาศัยวิธีเรกูลาร์ไรเซชันซึ่งเป็นการประมาณปัญหาที่ตั้งขึ้นอย่างบกพร่องด้วยวงศ์ของปัญหาที่ตั้งขึ้นมาอย่างดี

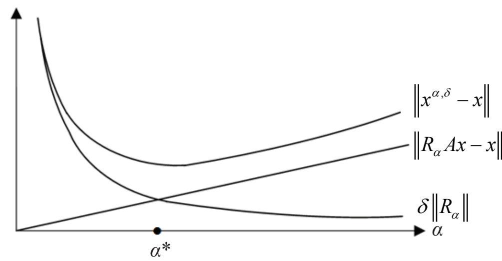
6. วิธีเรกูลาร์ไรเซชัน

กำหนดให้ $A: X \to Y$ เป็นตัวดำเนินการเชิงเส้นที่มีขอบเขตระหว่างปริภูมิฮิลเบิร์ต $\mathcal{H}_1$ ไปยัง $\mathcal{H}_2$ บนฟิลด์ $\mathbb{F}$ เดียวกัน เรกูลาร์ไรเซชัน คือวงศ์ของตัวดำเนินการเชิงเส้นที่มีขอบเขต $R_\alpha: X \to Y, \alpha > 0$ ที่ทำให้ $R_\alpha y^{\delta} = x^{\alpha,\delta} \to x$ เมื่อ $\delta \to 0$ สำหรับทุก x โดยที่ $\| y - y^{\delta} \| \le \delta$ ทั้งนี้ $\| R_\alpha \| \to \infty$

เราสามารถประมาณความคลาดเคลื่อนระหว่างผลเฉลยแม่นตรงกับผลเฉลยที่คำนวณได้โดยอาศัยสมบัติอสมการอิงรูปสามเหลี่ยมและสมบัติของนอร์มดังนี้

$$\begin{aligned} \| x^{\alpha,\delta} - x \| &\le \| R_\alpha y^{\delta} - R_\alpha y \| + \| R_\alpha y - x \| \\ &\le \| R_\alpha \| \cdot \| y^{\delta} - y \| + \| R_\alpha Ax - x \| \\ &\le \delta \| R_\alpha \| + \| R_\alpha Ax - x \| \end{aligned}$$

ในพจน์แรกเป็นค่าคลาดเคลื่อนของข้อมูล δ คูณกับ $\| R_\alpha \|$ ซึ่งจะลู่เข้าสู่อนันต์เมื่อ $\alpha \to 0$ ส่วนในพจน์ที่สองเป็นค่าคลาดเคลื่อนโดยประมาณ $\| R_\alpha Ax - x \| = \| (R_\alpha - A^{-1}) y \|$ ซึ่งจะลู่เข้าสู่ศูนย์เมื่อ $\alpha \to 0$ รูปที่ 6 แสดงลักษณะการลู่เข้าดังกล่าวนี้



รูปที่ 6 ลักษณะการลู่เข้าของความคลาดเคลื่อน

จากเหตุผลที่กล่าวข้างต้น การกำหนดค่าของพารามิเตอร์ α มีความสำคัญ กล่าวคือจะต้องเลือกใช้ α ที่ทำให้เกิดค่าคลาดเคลื่อนระหว่างผลเฉลยมีค่าต่ำสุด ในรูปที่ 6 ค่า α^* เป็นค่าที่เหมาะสมที่สุด

พิจารณาปัญหา $Ax = y$ ที่มี (σ_n, u_n, v_n) เป็นระบบเอกฐาน (Singular System) เมื่อพิจารณาตัวดำเนินการ $A^{\dagger}$ และ R_α ในรูปแบบการแยกค่าเอกฐาน (Singular Value Decomposition: SVD) แล้วเขียนผลเฉลยโดยประมาณของปัญหานี้ ได้ดังนี้

$$x^{\dagger} = A^{\dagger} y = \sum_{n=1}^{\infty} \frac{\langle y, v_n \rangle}{\sigma_n} u_n$$

และ

$$x_\alpha = R_\alpha y = \sum_{n=1}^{\infty} g_\alpha(\sigma_n) \langle y, v_n \rangle u_n$$

เนื่องจากเราประมาณค่าผลเฉลย $x^{\dagger}$ ด้วยผลเฉลยที่ได้จากตัวดำเนินการเรกูลาร์ไรเซชัน x_α ดังนั้นเงื่อนไขใด ๆ ที่ใช้เพื่อนิยาม $g_\alpha(\sigma_n)$ จะเป็นเงื่อนไขที่ส่งผลให้ x_α ลู่เข้าสู่ $x^{\dagger}$ ด้วย

เมื่อพิจารณาการลู่เข้าแบบรายจุด (Pointwise Convergence) ถ้า $g_\alpha(\sigma) \to 1/\sigma$ แล้ว $R_\alpha \to A^{\dagger}$ Engl และคณะ [1] ได้พิสูจน์ให้เห็นว่า ถ้ากำหนดให้ $g_\alpha : \mathbb{R}^{+} \to \mathbb{R}^{+}$ โดยที่ $g_\alpha(\sigma) \to 1/\sigma$ สำหรับ $\sigma > 0$ เมื่อ $\alpha \to 0$ และ $g_\alpha(\sigma) \le C_\alpha < \infty$ สำหรับจำนวนจริง σ ทุกค่าแล้ว (R_α, α) เป็นวิธีเรกูลาร์ไรเซชันที่ลู่เข้า

ต่อไปจะให้ตัวอย่างการสร้างวิธีเรกูลาร์ไรเซชันที่เป็นที่รู้จักกันดี 3 วิธี ดังนี้

วิธีที่ 1 วิธีการแยกค่าเอกฐานแบบตัดปลาย (Truncated Singular Value Decomposition) แนวคิดหลักของวิธีการนี้คือ การเพิกเฉยต่อค่าเอกฐานที่มีค่าน้อยกว่าค่าของ α โดยกำหนดให้

$$g_\alpha(\sigma) = \begin{cases} 1/\sigma & \text{ถ้า } \sigma \geq \alpha \\ 0 & \text{ถ้า } \sigma < \alpha \end{cases}$$

โดยวิธีการนี้จะหาผลเฉลยโดยประมาณได้จาก

$$x_\alpha = R_\alpha y = \sum_{\sigma_n \geq \alpha} \frac{\langle y, v_n \rangle}{\sigma_n} u_n$$

วิธีที่ 2 วิธีเรกูลาร์ไรเซชันแบบ Lavrentiev (Lavrentiev Regularization) แนวคิดหลักของวิธีการนี้คือ การเลื่อนค่าเอกฐานด้วยค่า α โดยนิยามให้

$$g_\alpha(\sigma) = \frac{1}{\sigma + \alpha}$$

โดยวิธีการนี้จะหาผลเฉลยโดยประมาณได้จาก

$$x_\alpha = R_\alpha y = \sum_{n=1}^{\infty} \frac{1}{\sigma_n + \alpha} \langle y, v_n \rangle u_n$$

หรือ x_α เป็นผลเฉลยของสมการ $(A + \alpha I)x_\alpha = y$

วิธีที่ 3 วิธีเรกูลาร์ไรเซชันแบบทิกโฮนอฟ (Tikhonov Regularization) แนวคิดหลักของวิธีการนี้คือ การพิจารณาสมการเกาส์เซียนเชิงปกติ

$$A^*Ax = A^*y$$

โดยที่ A^* คือตัวดำเนินการผูกผัน (Adjoint Operator) ที่ซึ่ง $\langle Ax, y\rangle_Y = \langle x, A^*y\rangle_X$ ทุกค่า $x \in X, y \in Y$ โดยที่ $\langle \cdot,\cdot \rangle_X$ และ $\langle \cdot,\cdot \rangle_Y$ คือผลคูณภายในปริภูมิฮิลเบิร์ต X และปริภูมิฮิลเบิร์ต Y ตามลำดับ

เมื่อเขียน x ในรูป SVD ได้ว่า

$$x = \sum_{n=1}^{\infty} \frac{\sigma_n}{\sigma_n^2} \langle y, v_n \rangle u_n$$

เมื่ออาศัยหลักการเลื่อนค่าเอกฐาน (Shifting Singular Value) ด้วยค่า α เช่นเดียวกับวิธีของ Lavrentiev ทำให้ได้ว่า

$$g_\alpha(\sigma) = \frac{\sigma}{\sigma^2 + \alpha}$$

โดยวิธีนี้ จะสามารถหาผลเฉลยโดยประมาณได้จาก

$$x_\alpha = R_\alpha y = \sum_{n=1}^{\infty} \frac{\sigma_n}{\sigma_n^2 + \alpha} \langle y, v_n \rangle u_n$$

หรือ x_α คือผลเฉลยของสมการ

$$(A^*A + \alpha I)x_\alpha = A^*y$$

ถ้าแทนที่ y ด้วย y^δ แล้วเป้าหมายของเราคือการหาผลเฉลยโดยประมาณที่ดีที่สุด x_α^δ ของปัญหา $Ax = y$ จากสมการ

$$(A^*A + \alpha I)x_\alpha^\delta = A^*y^\delta$$

ซึ่งปัญหาดังกล่าวนี้จะสมมูลกับปัญหาการหา x ที่ทำให้ $J_\alpha(x)$ มีค่าต่ำสุด โดยที่

$$J_\alpha(x) = \| Ax + y^\delta \|_Y^2 + \alpha \| x \|_X^2, \ x \in X \qquad (14)$$

นั่นคือ ต้องการหา

$$\begin{aligned} x_\alpha &= \arg(\min_x J_\alpha(x)) \\ &= \arg\left(\min_x \left[\| Ax + y^\delta \|^2 + \alpha \| x \|^2\right]\right) \qquad (15) \end{aligned}$$

เมื่อไรก็ตามที่ค่าต่ำสุดมีจริง

เราเรียกพจน์ $\alpha \| x \|^2$ ในสมการที่ (15) ว่า พจน์การปรับลด (Penalty Term) และเรียกสมการที่ (14) ว่า ทิกโฮนอฟฟังก์ชันนอล (Tikhonov Functional) และเรียกพารามิเตอร์ $\alpha > 0$ ว่า พารามิเตอร์เรกูลาร์ไรเซชัน (Regularization Parameter) โดยทั่วไปจะสามารถหา α ได้จาก Discrepancy Principle ของ Morozov ในกรณีของเราจะเลือก $\alpha = \alpha(\delta, y^\delta)$ และ x_α ที่ทำให้

$$\rho_1 \delta \leq \| Ax + y^\delta \| \leq \rho_2 \delta$$

โดยที่ ρ_1, ρ_2 เป็นค่าคงที่ซึ่ง $1 < \rho_1 < \rho_2$

สำหรับปัญหาไม่เชิงเส้น (Nonlinear) โดยวิธีทิกโฮนอฟมีรูปแบบ ดังนี้

$$\min_p \| F(p) - y \|^2 + \alpha \| p - p^0 \|^2$$

โดยที่ $\alpha > 0$ วิธีการนี้มีชื่อเรียกเฉพาะว่า *เรกูลาร์ไรเซชัน*

ที่ปรับเลื่อนด้วยสเปกตรัม (Regularization by Shifting the Spectrum) [1], [3] และ [4]

7. ปัญหาผกผันในคณิตศาสตร์การเงิน

ในหัวข้อนี้จะได้ยกตัวอย่างการหาค่าความผันผวนเฉพาะถิ่น (Local Volatility) $\sigma(S,t)$ ในสมการเชิงอนุพันธ์ย่อยแบล็ค-โชลของมูลค่าออปชัน V ที่ขึ้นกับราคาหุ้น S ณ เวลา t ดังนี้

$$\frac{\partial V}{\partial t}+(r-d)S\frac{\partial V}{\partial S}+\frac{1}{2}\sigma^2(S,t)S^2\frac{\partial^2 V}{\partial S^2}=rV \quad (16)$$

โดยที่ r และ d เป็นค่าคงที่ ซึ่งแทนอัตราดอกเบี้ย และอัตราการจ่ายเงินปันผลตามลำดับ

กำหนดให้ $V(S,t,\sigma(S,t))$ ซึ่งแทนด้วย $V(\sigma)$ เป็นผลเฉลยของสมการที่ (16) โดยพิจารณาให้ $V(\sigma)$ เป็นตัวดำเนินการไม่เชิงเส้นเมื่อพิจารณาเทียบกับ $\sigma(S,t)$ (แทนด้วย σ) ดังนี้

$$L^2(\Omega)\supseteq D\ni\sigma\xrightarrow{V}V(\sigma)\in L^2(\Omega)$$

ค่าออปชัน V จะต้องมีค่าอยู่ในช่วงระหว่างราคาเสนอซื้อกับราคาเสนอขาย (Bid-ask Spread) ดังนี้

$$V_{ij}^b\le V(S_0,0,K_{ij},T_i,\sigma)\le V_{ij}^a \quad (17)$$

เราต้องการหาค่า σ ที่สอดคล้องกับเงื่อนไขสมการที่ (17) ซึ่งทำให้

$$G(\sigma)=\sum_{i=1}^{N}\sum_{j=1}^{M_i}(V(S_0,0,K_{ij},T_i,\sigma)-V_{ij})^2$$

มีค่าต่ำสุด โดยที่ $V_{ij}=\frac{1}{2}(V_{ij}^a+V_{ij}^b)$

เมื่ออาศัยวิธีเรกูลาร์ไรเซชันแบบทิกโฮนอฟในการแก้ปัญหาข้างต้นนี้จะได้ว่า เป็นการหาค่า σ ที่ทำให้

$$F_\alpha(\sigma)=G(\sigma)+\alpha\|\sigma\|^2 \quad (18)$$

มีค่าต่ำสุด ซึ่งปัญหาข้างต้นจะสมมูลกับปัญหาการหาค่าต่ำสุด

$$\tilde{F}_\lambda(\sigma)=\lambda G(\sigma)+\|\sigma\|^2$$

เมื่อ λ เป็นค่าคงตัวที่มากกว่าศูนย์

งานวิจัยเช่น [5]–[8] ได้ศึกษาวิธีการแก้ปัญหาเพื่อหาค่าความผันผวนเฉพาะถิ่น ในปี ค.ศ. 1997 Lagnado และ Osher [9] ได้เสนอวิธีการหาค่า σ โดยใช้วิธีเรกูลาร์ไรเซชันแบบทิกโฮนอฟ โดยได้แทนที่ $\|\sigma\|^2$ ในสมการที่ (18) ด้วย $\|\nabla\sigma\|^2$

นอกจากวิธีการเลือกแทนที่พจน์ที่สองในสมการที่ (18) ที่แตกต่างกันแล้ว ในปี ค.ศ. 2002 Chiarella และคณะ [10] ได้แทนที่ $G(\sigma)$ ในสมการที่ (18) ด้วย $H(\sigma)$ โดยที่

$$H(\sigma)=\sum_{i=1}^{N}\sum_{j=1}^{M_i}\int_0^{T_{\mathrm{cur}}}\int_0^\infty(V(S_t,t,K_{ij},T_i,\sigma)-V_{ij})^2\,dSdt$$

เมื่อ T_{cur} แทนเวลาปัจจุบันโดย Chiarella และคณะได้พิจารณาปัญหาการหาค่า σ ที่ทำให้

$$J_\lambda(\sigma)=\lambda H(\sigma)+\|\nabla\sigma\|^2$$

มีค่าต่ำสุด โดยที่

$$\|\nabla\sigma\|^2=\iint\frac{\partial^2\sigma}{\partial S^2}+\frac{\partial^2\sigma}{\partial t^2}\,d\Omega_1 d\Omega_2$$

โดยที่ $\Omega_1=(0,\infty)$ และ $\Omega_2=(0,T_{\mathrm{cur}})$

ในปี ค.ศ. 2014 Zeng และคณะ [11] ได้ศึกษาและวิจัยวิธีการหา σ โดยวิธีทิกโฮนอฟจากงานวิจัยของ Lagnado และ Osher [9] วิธีของ Chia-rella และคณะ [12] และศึกษาวิธี Total Variation Regularization ที่คิดค้นโดย Rudin และ คณะ [13] ในการแก้ปัญหาที่ผลเฉลยเป็นฟังก์ชันที่ไม่มีความต่อเนื่อง เช่น ปัญหาการทำรูปขึ้นใหม่ เป็นต้น Zeng และคณะ [11] ได้นำวิธีข้างต้นมาประยุกต์ใช้กับการหาค่าความผันผวนเฉพาะถิ่นในกรณีที่ความผันผวนมีความไม่ต่อเนื่อง (ซึ่งจะเกิดขึ้นในกรณีของที่มีการกระโดด (Jump) เป็นต้น) โดยพิจารณาปัญหาการหาค่า σ ที่ทำให้ $\breve{J}(\sigma)$ มีค่าต่ำสุด โดยที่

$$\breve{J}(\sigma)=\frac{1}{2}\|V(\sigma)-v\|_{L^2(\Omega)}^2+\alpha J_\beta(\sigma)$$

โดยที่ v เป็นเวกเตอร์ของราคาตลาดของออปชัน $J_\beta(\sigma) = \int \sqrt{|\nabla\sigma|^2 + \beta^2}\, dSdt$ และ β คือพารามิเตอร์ที่เป็นค่าคงตัว ซึ่งเป็นจำนวนจริงบวก เช่น $\beta = 10^{-6}$

อีกทั้ง Zeng และคณะ [11] ได้แสดงให้เห็นว่า ผลลัพธ์ที่ได้จากวิธีที่เขาเสนอดีกว่าผลลัพธ์ที่ได้จากวิธีทิกโฮนอฟของ Lagnado และ Osher นอกจากนี้พวกเขายังได้ศึกษาเงื่อนไขของการมีผลเฉลย การลู่เข้าและเสถียรภาพของผลเฉลยอีกด้วย

8. สรุป

บทความประมวลความรู้เชิงวิเคราะห์ (Review Article) ฉบับนี้ได้อภิปรายถึงปัญหาผกผันซึ่งเป็นปัญหาที่มีลักษณะตรงกันข้ามกับปัญหาตรง ปัญหาผกผันเป็นการอธิบายความสัมพันธ์จากผลลัพธ์ไปสู่เหตุ ในตอนต้นของบทความนี้ได้อธิบายถึงลักษณะปัญหาต่าง ๆ ได้แก่ ปัญหาตรง ปัญหาย้อนกลับ และปัญหาการหาพารามิเตอร์ที่เหมาะสมของตัวแบบ ปัญหาที่สร้างขึ้นอย่างดีหมายถึงปัญหาที่ต้องสอดคล้องเงื่อนไขทั้ง 3 ประการตามหลักเกณฑ์ของ Hadamard ที่ว่า ผลเฉลยของปัญหานั้นมีจริง มีอยู่เพียงหนึ่งเดียว และเสถียรภาพของผลเฉลย

ในปัญหาจริงเชิงวิจัยส่วนใหญ่แล้วเป็นปัญหาผกผันที่มีผลเฉลยที่ไร้เสถียรภาพ จึงจัดเป็นปัญหาที่ตั้งขึ้นอย่างบกพร่อง การแก้ปัญหาในลักษณะเช่นนี้จะอาศัยทฤษฎีเรกูลาร์ไรเซชัน ในบทความฉบับนี้ได้อธิบายแนวคิดของทฤษฎีเรกูลาร์ไรเซชัน 3 วิธี ได้แก่ วิธีการตัดค่าเอกฐาน วิธีเรกูลาร์ไรเซชันแบบ Lavrentiev และวิธีเรกูลาร์ไรเซชันแบบทิกโฮนอฟ ท้ายสุดได้ประมวลสรุปงานวิจัยที่ศึกษาการหาความผันผวนเฉพาะถิ่นในทางคณิตศาสตร์การเงินด้วยวิธีเรกูลาร์ไรเซชัน

9. กิตติกรรมประกาศ

ผู้เขียนขอขอบคุณอาจารย์ภาขวัญ ริยาพันธ์ สังกัดภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี และขอขอบพระคุณคณะกรรมการผู้ทรงคุณวุฒิทุกท่านที่ได้กรุณาให้คำแนะนำตลอดจนแก้ไขข้อบกพร่องของบทความฉบับนี้ให้สมบูรณ์ที่สุด

เอกสารอ้างอิง

[1] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*. Dordrecht, Netherlands: Kluwer Academic Publishers, 1996.

[2] E. Kreyszig, *Introductory Functional Analysis: With Applications to Boundary Value Problems and Finite Elements*. Springer Science & Business Media, 1978.

[3] T. I. Seidman and C. R. Vogel, "Well posedness and convergence of some regularisation methods for non-linear ill posed problems," *Inverse Problems*, vol. 5, no. 2, pp. 227–238, 1989.

[4] M. Hanke, "Regularizing properties of a truncated Newton–CG algorithm for nonlinear inverse problems," *Numerical Functional Analysis Optimization*, vol. 18, pp. 971–993, May 1997.

[5] N. Jackson, E. Suli, and S. Howison, "Computation of deterministic volatility surfaces," *Journal of Computational Finance*, vol. 2, 1998.

[6] J. Lishang and T. Youshan, "Identifying the volatility of underlying assets from option prices," *Inverse Problems*, vol. 17, no. 1, pp. 137–155, 2001.

[7] S. Crépey, "Calibration of the local volatility in a generalized black--scholes model using tikhonov regularization," *SIAM on Journal of Mathematical Analysis*, vol. 34, no. 5, 2003.

[8] H. Egger and H. Engl, "Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates," *Inverse Problems*, vol. 21, no. 3, April 2005.

[9] R. Lagnodo and S. Osher, "A technique for

calibrating derivative security pricing models: numerical solution of an inverse problem," *The Journal of Computational Finance*, no. 1, pp. 13–25, 1997.

[10] C. Chiarella, M. Craddock, and N. El-Hassan, "A short time expansion of the volatility function for the calibration of option pricing models," *Computing in Economics and Finance*, no. 261, 2002.

[11] Y. Zeng, S. Wang, and Y. Yang, "Calibration of the volatility in option pricing using the total variation regularization," *Journal of Applied Mathematics*, vol. 2014, pp. 1–9, 2014.

[12] C. Chiarella, M. Craddock, and N. El-Hassan, "The calibration of stock option pri-cing models using inverse problem methodology," *Quantitative Finance Research Centre*, *University of Technology*, Sydney, no. 39, April 2000.

[13] L. I. Rudin. S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992.