บทที่ 1

บทนำ

ความเป็นมาและความสำคัญ

วิธีเกรเดียนต์ใกล้เคียง เป็นวิธีการหาค่าเหมาะสมที่สุด เป็นหนึ่งวิธีที่นิยมใช้ในการหาค่าน้อยสุดของฟังก์ชันไม่ราบเรียบ ประกอบด้วยการใช้ขั้นตอนเกรเดียนต์ต่อเนื่อง ในฟังก์ชันแรกตามด้วยตัวดำเนินการความใกล้เคียงของฟังก์ชันที่สอง เป็นที่ทราบกันดีว่าอัตราการลู่เข้าทางทฤษฎี และทางปฏิบัติของวิธีนี้ สามารถพัฒนาได้โดยความเฉื่อย หรือที่เรียกว่าการเร่งความเร็วของ Nesterov (1983 : 372-376)

สมมติ H เป็นปริภูมิฮิลเบิร์ตเชิงจริงด้วยผลคูณภายใน $\langle \cdot, \cdot \rangle$ และนอร์ม $\| \cdot \|$ ให้ $B : H \to 2^H$ เป็นตัวดำเนินการทางเดียวมากที่สุดและ $A : H \to H$ เป็นการดำเนินการทางเดียว ต่อเนื่อง ลิพชิทซ์ พิจารณาปัญหาร่วมกันดังนี้ หาค่า $v \in H$ ซึ่ง

$$0 \in (A + B)v \tag{1.1}$$

เซตผลเฉลยของปัญหา (1.1) แสดงโดย

$$(A + B)^{-1}(0) \tag{1.2}$$

ปัญหา (1.1) ประกอบด้วย ที่เป็นกรณีพิเศษได้แก่ ปัญหาการผกผันเชิงเส้นมีเงื่อนไขบังคับแบบคอนเวกซ์ ปัญหาความเป็นไปได้แบบแยก ปัญหาการหาค่าน้อยสุดมีเงื่อนไขบังคับแบบคอนเวกซ์ ปัญหาจุดตรึง และปัญหาอื่นแบบจำลองทางคณิตศาสตร์ยอดนิยมซึ่งสามารถกำหนดได้ภายใต้ ปัญหา (1.1) คือ ปัญหาการหาค่าเหมาะสมที่สุดของผลรวมของสองฟังก์ชัน ดังต่อไปนี้

$$\min_{v \in H} \left(g(v) + h(v) \right) \tag{1.3}$$

โดยที่ $h : H \to \mathbb{R} \cup \infty$ เป็นฟังก์ชันกึ่งต่อเนื่องด้านล่างและคอนเวกซ์ ซึ่ง $g : H \to \mathbb{R}$ เป็นฟังก์ชันคอนเวกซ์ สามารถหาอนุพันธ์ได้ และ h หาอนุพันธ์ย่อยได้ ในกรณีนี้ ปัญหาการหาค่าเหมาะสมที่สุด (1.2) จะสมมูลกับปัญหา (1.1) ด้วยการแทน $A = \partial h$ และ $B = \nabla g$ วิธีที่นิยมในการแก้ปัญหา (1.1) ในปริภูมิฮิลเบิร์ตเชิงจริง คือ วิธีการแยกไปข้างหน้า-ข้างหลัง ซึ่งนำเสนอโดย Passty (1979 : 383-390) และ Lions & Mercier (1979 : 964-979) มีรูปแบบดังนี้

$$v_{i+1} = J_{\rho A}(v_i - \rho B(v_i)) \tag{1.4}$$

โดยที่ $\rho > 0$ เป็นพารามิเตอร์ที่เหมาะสม และ $J_{\rho A} = (I + \rho A)^{-1}$ การลู่เข้าของวิธีนี้เกิดขึ้นภายใต้การดำเนินการ c ทางเดียวแบบเข้มผกผันของการดำเนินการ B นั่นคือ เมื่อ $c > 0$ จะมี

$$\langle B(v) - B(y), v - y \rangle \geq c\|B(v) - B(y)\|, \forall\, v, y \in H \tag{1.5}$$

หลังจากนั้นก็มีการพัฒนาวิธีการแก้ปัญหา (1.1) อย่างต่อเนื่องจนกระทั่งในปี 2000 Tseng (2000 : 431-446) ได้นำเสนอวิธีการแยกไปข้างหน้า-ไปข้างหลัง-ไปข้างหน้า ดังนี้

$$\begin{aligned} y_i &= J_{\rho_i A}(v_i - \rho_i B(v_i)) \\ v_{i+1} &= y_i - \rho_i(B(y_i) - B(v_i)) \end{aligned} \tag{1.6}$$

โดยที่ $\{\rho_i\} \subset (0, \infty)$ เป็นลำดับที่เหมาะสม ลำดับนี้สามารถกำหนดได้ดังนี้ หากทราบค่าคงที่ L ลิพชิทซ์ของ B แล้ว $\{\rho_i\} \subset [a, b] \subset (0, 1/L)$ และในกรณีผกผัน จะพบว่าบางส่วนในขั้นตอนการค้นหาเส้นโดยมีเกณฑ์การหยุดแบบจำกัดในแต่ละการวนซ้ำ อย่างไรก็ตามค่าคงที่ลิพชิทซ์ของการดำเนินการค่อนข้างยากที่จะคำนวณได้ในปัญหาที่ไม่เป็นเชิงเส้น นอกจากนี้การใช้การค้นหาเส้นนั้นอาจจะไม่ถูกต้องและใช้เวลานาน เนื่องจากการค้นหาเส้นมักจะต้องใช้การคำนวณจำนวนมากในการวนซ้ำแต่ละครั้ง

Malitsky & Tam (2020 : 383-390) ได้นำเสนออัลกอริทึมไปข้างหน้า-สะท้อน-ไปข้างหลัง เมื่อ $\rho_0,\ \rho_1 > 0, \gamma \in \{1, \beta^{-1}\},\ \beta \in (0, 1),\ \delta \in (0, 1)$ และ $v_0,\ v_1 \in H$ จะได้ว่า

$$v_{i+1} = \mathrm{prox}_{\rho_i h}(v_i - \rho_i \nabla g(v_i) - \rho_{i-1}(\nabla g(v_i) - \nabla g(v_{i-1}))) \tag{1.7}$$

ซึ่งขนาดขั้นตอน $\rho_i = \gamma \rho_{i-1} \beta^n$ ด้วย n จำนวนเต็มที่น้อยที่สุดที่ไม่เป็นลบ ที่สอดคล้องกับ

$$\rho_i \|\nabla g(v_{i+1}) - \nabla g(v_i)\| \leq \frac{\delta}{2}\|v_{i+1} - v_i\|$$

ต่อมา Hieu, Anh & Muu (2021 : 127-151) ได้นำเสนอวิธีไปข้างหน้า-สะท้อน-ไปข้างหลังแบบปรับปรุง ดังนี้

$$\begin{aligned} v_{i+1} &= \mathrm{prox}_{\rho_i h}(v_i - \rho_i \nabla g(v_i) - \rho_{i-1}(\nabla g(v_i) - \nabla g(v_{i-1}))) \\ \rho_{i+1} &= \min\{\mu \frac{\|v_{i+1} - v_i\|}{\|\nabla g(v_{i+1}) - \nabla g(v_i)\|}, \rho_i\} \end{aligned} \tag{1.8}$$

สำหรับทุก $n \geq 1,\ v_0, v_1 \in H, \rho_0,\ \rho_1 > 0$ และ $\delta \in \left(0, \frac{1}{2}\right)$ ซึ่งไม่ต้องหาค่าคงที่ลิพชิทซ์ของการดำเนินการ

ในโครงงานวิจัยนี้ คณะผู้วิจัยได้รับแรงบันดาลใจจากผลลัพธ์ของ Tseng และ Hieu ซึ่งวิธีที่ คณะผู้วิจัยสนใจที่จะพัฒนา คือ วิธีเกรเดียนต์ใกล้เคียงด้วยขั้นตอนเฉื่อยสลับกันนั้นไม่ต้อง ทราบค่าคงที่ L ลิพชิทซ์ของการดำเนินการและไม่มีขั้นตอนการค้นหาเส้น ตลอดจนนำทฤษฎีนี้ ไปประยุกต์ใช้ในการปรับปรุงภาพที่เบลอให้มีความคมชัดขึ้น

วัตถุประสงค์ของการวิจัย

1. เพื่อสร้างวิธีเกรเดียนต์ใกล้เคียงด้วยขั้นตอนเฉื่อยสลับกันโดยไม่ใช้ค่าคงที่ L ลิพชิทซ์และไม่มีขั้นตอนการค้นหาเส้นสำหรับแก้ไขปัญหาในการหาค่าน้อยสุด

2. เพื่อสร้างทฤษฎีบทและพิสูจน์การลู่เข้าของวิธีเกรเดียนต์ใกล้เคียงด้วยขั้นตอนเฉื่อยสลับกันโดยไม่ใช้ค่าคงที่ L ลิพชิทซ์และไม่มีขั้นตอนการค้นหาเส้นสำหรับแก้ไขปัญหาในการหาค่าน้อยสุด

3. แสดงให้เห็นถึงพฤติกรรมการลู่เข้าเพื่อแสดงประสิทธิภาพและการดำเนินการตามวิธีการที่ได้นำเสนอ

ประโยชน์ของการวิจัย

1. ได้วิธีเกรเดียนต์ใกล้เคียงด้วยขั้นตอนเฉื่อยสลับกันโดยไม่ใช้ค่าคงที่ L ลิพชิทซ์และไม่มีขั้นตอนการค้นหาเส้นสำหรับแก้ไขปัญหาในการหาค่าน้อยสุด

2. ได้ทฤษฎีบทและพิสูจน์การลู่เข้าของวิธีเกรเดียนต์ใกล้เคียงด้วยขั้นตอนเฉื่อย สลับกันโดยไม่ใช้ค่าคงที่ L ลิพชิทซ์และไม่มีขั้นตอนการค้นหาเส้นสำหรับแก้ไขปัญหาในการหาค่าน้อยสุด

3. ได้ยกระดับคุณภาพการศึกษาและเผยแพร่ผลงานในระดับชาติหรือนานาชาติ

ขอบเขตของการวิจัย

ขอบเขตของการวิจัยคือ การสร้างทฤษฎีบทและพิสูจน์การลู่เข้าของวิธีเกรเดียนต์ใกล้เคียงด้วยขั้นตอนเฉื่อยสลับกัน สำหรับแก้ไขปัญหาในการหาค่าน้อยสุดในปริภูมิฮิลเบิร์ตจริง

นิยามศัพท์เฉพาะ

1. H แทนปริภูมิฮิลเบิร์ต หมายถึง ปริภูมิผลคูณภายในบริบูรณ์ซึ่งประกอบด้วยนอร์มที่ถูกนิยามโดยผลคูณภายใน คือ $\|v\| = \sqrt{\langle v, v\rangle}$ สำหรับ $v \in H$ เมื่อ $\langle \cdot, \cdot \rangle$ คือ สัญลักษณ์ของผลคูณภายในและมีสมบัติดังนี้

1.1 $\langle v, v\rangle \geq 0$

1.2 $\langle v, v\rangle = 0$ ก็ต่อเมื่อ $v = 0$

1.3 $\langle v, y\rangle = \langle y, v\rangle$

1.4 $\langle \alpha v + \beta y, z\rangle = \alpha \langle v, z\rangle + \beta \langle y, z\rangle$ สำหรับ $v, y, z \in H$ และ $\alpha, \beta \in \mathbb{R}$

2. $Fix(\mathcal{D}) = \{v \in \mathcal{H} \mid \mathcal{D}v = v\}$ แทนเซตจุดตรึงของการส่ง $\mathcal{D}$

3. ให้ $g : H \to (-\infty, \infty)$ เป็นฟังก์ชันเฉพาะที่เป็นฟังก์ชันกึ่งต่อเนื่องล่างและคอนเวกซ์ โดเมนของ g โดย dom g $= \{v \in H \,|\, g(v) < \infty\}$ สำหรับ $x \in$ dom g ใด ๆ อนุพันธ์ย่อยของ g ที่ v

ถูกกำหนดโดย

$$\partial g(v) = \{z \in H \mid \langle z, y - v\rangle \leq g(y) - g(v),\, y \in H\}$$

4. $\mathcal{P}_{\mathcal{D}}$ แทนการฉายเมตริกของ H ทั่วถึง $\mathcal{D}$ หมายถึง สำหรับทุก $v \in H$ และ $\mathcal{D}$ เซตไม่ว่าง ปิด และคอนเวกซ์ มี $y \in \mathcal{D}$ ซึ่ง

$$\|v - y\| = d(v, \mathcal{D}) = \inf\{\|v - w\| : w \in \mathcal{D}\}$$