บทที่ 2

ทฤษฎีในการหาค่าออปติมัม

ความนำ

ในบทนี้ กล่าวถึงทฤษฎีต่างๆที่ใช้หาค่าออปติมัมในวิทยานิพนธ์นี้ เพื่อให้เวลาที่ใช้ในการคำนวณหาค่าออปติมัมและความแม่นยำดีขึ้น ทฤษฎีในบทนี้แบ่งเป็น 5 ส่วนคือ ส่วนแรกกล่าวถึงรายละเอียดงานวิจัยของ Glasos , Hui และ Zak ที่ใช้เครือข่ายนิวรอลหาค่าออปติมัมแบบมีเงื่อนไข ซึ่งนำมาพัฒนาปรับปรุงในวิทยานิพนธ์นี้ ส่วนที่สองกล่าวถึงอัลกอริทึมในงานวิจัยของ Glasos , Hui และ Zak ส่วนที่สามกล่าวถึงรายละเอียดงานวิจัยที่ใช้เครือข่ายนิวรอลหาค่าออปติมัมแบบมีเงื่อนไขในวิทยานิพนธ์นี้ ส่วนที่สี่กล่าวถึงการพิสูจน์การลู่เข้าหาค่าออปติมัมของวิธี Conjugate Gradient และส่วนสุดท้ายกล่าวถึงอัลกอริทึมของงานวิจัยในวิทยานิพนธ์นี้

งานวิจัยของ Glasos , Hui และ Zak

การหาค่าออปติมัมในงานวิจัยของ Glasos , Hui และ Zak มี 3 ขั้นตอนดังนี้คือ

1. ใช้วิธี Penalty function แปลงจากปัญหาการหาค่าออปติมัมแบบมีเงื่อนไข (**P1**) ให้เป็นปัญหาการหาค่าออปติมัมแบบไม่มีเงื่อนไข (**P2**)

$$\mathbf{P1}: \begin{cases} \text{หา } \mathbf{x} \text{ ซึ่งทำให้ } f(\mathbf{x}) \text{ มีค่าต่ำสุด} \\ \text{โดยที่ } g_i(\mathbf{x}) \le 0 \;,\; i = 1,2,\ldots,m \end{cases} \tag{2.1}$$

$\mathbf{x} \in \Re^n$, $f: \Re^n \to \Re$ และ $g_i: \Re^n \to \Re$, $i = 1,2,\ldots,m$

$$\mathbf{P2}: \{ \text{ หา } \mathbf{x} \text{ ซึ่งทำให้ } p(\mathbf{x}) \text{ มีค่าต่ำสุด} \tag{2.2}$$

ซึ่ง
$$p(\mathbf{x}) = H(\mathbf{g}(\mathbf{x}))\, f(\mathbf{x}) + \mu\, q(\mathbf{x}) \tag{2.3}$$

โดยที่ $H(.)$ เป็น Multidimensional step function คือ

$H(\mathbf{g}(\mathbf{x})) = 1$ เมื่อ $g_i(\mathbf{x}) \le 0$, $i = 1,2,\ldots,m$

และ $H(\mathbf{g}(\mathbf{x})) = 0$ เมื่อ $g_1(\mathbf{x}) > 0$ หรือ $g_2(\mathbf{x}) > 0\ldots$ หรือ $g_m(\mathbf{x}) > 0$

และถ้า $\mathbf{q(x)}$ ในสมการที่ (2.3) เป็น

$$q(\mathbf{x}) = \sum_{i=1}^{m}[U(g_i(\mathbf{x}))\,|\,g_i(\mathbf{x})|\,] \quad \text{สำหรับ Absolute value} \tag{2.4}$$

โดยที่ $U(.)$ เป็น Unidimensional step function คือ

$U(v) = 1$ เมื่อ $v > 0$

และ $U(v) = 0$ เมื่อ $v \leq 0$

ถ้าให้ Ω เป็นเขตพื้นที่ที่ยอมรับ (Feasible region) นั้นคือ

$$\Omega = \bigcap_{i=1}^{m} \{\, \mathbf{x} \mid g_i(\mathbf{x}) \leq 0 \,\}$$

ถ้าให้ Index set เป็น

$$J(\mathbf{x}) = \{\, i \mid g_i(\mathbf{x}) > 0,\ i = 1,2,\ldots,m \,\}$$

2. ใช้วิธี Steepest Descent แก้ปัญหา **P2** โดยแปลงจากปัญหาการหาค่าออปติมัมแบบไม่มีเงื่อนไข (**P2**) ให้เป็นปัญหาของการแก้สมการอนุพันธ์ (differential equation) ของตัวแปรสเตท ซึ่งมี Forcing function เป็นฟังก์ชันของ Steepest descent (**P3**)

$$\textbf{P3 :} \qquad \tau \frac{d\mathbf{x}(t)}{dt} = \mathbf{s}_p(\mathbf{x}(t)) \tag{2.5}$$

โดยที่ $\mathbf{s}_p(\mathbf{x}(t))$ เป็น Steepest descent direction ซึ่งหาได้จากสมการที่ (2.6), พารามิเตอร์ τ และ μ มีค่าคงที่เป็นจำนวนจริงบวก และ $\nabla p(\mathbf{x}) \triangleq \left[\frac{\partial p(\mathbf{x})}{\partial x_1} \ \frac{\partial p(\mathbf{x})}{\partial x_2} \ \cdots \ \frac{\partial p(\mathbf{x})}{\partial x_n} \right]^T$

$$\mathbf{s}_p(\mathbf{x}(t)) = -\nabla p(\mathbf{x}(t)) \tag{2.6}$$

$$\mathbf{s}_p(\mathbf{x}(t)) = \left\{ \begin{array}{ll} -\mu\left(\sum_{i \in J(\mathbf{x}(t))} \nabla g_i(\mathbf{x}(t)) \right) & ,\ \mathbf{x}(t) \notin \Omega \\ \\ -\nabla f(\mathbf{x}(t)) & ,\ \mathbf{x}(t) \in \Omega \end{array} \right\} \tag{2.7}$$

สมการที่ (2.7) หาได้จากสมการที่ (2.3) , (2.4) และ (2.6) ดังจะอธิบายต่อไปนี้คือ

เมื่อแทน $\mathbf{p}(\mathbf{x}(t))$ ในสมการที่ (2.6) ด้วยสมการที่ (2.3) จะได้ว่า

$$\begin{aligned} \mathbf{s}_p(\mathbf{x}(t)) &= -\nabla\left[H(\mathbf{g}(\mathbf{x}(t)))\, f(\mathbf{x}(t)) + \mu\, q(\mathbf{x}(t))\right] \\ &= -H(\mathbf{g}(\mathbf{x}(t)))\, \nabla f(\mathbf{x}(t)) - \mu\, \nabla q(\mathbf{x}(t)) \end{aligned} \tag{2.8}$$

เมื่อแทน $q(\mathbf{x}(t))$ ในสมการที่ (2.8) ด้วยสมการที่ (2.4) จะได้ว่า

$$\mathbf{s}_p(\mathbf{x}(t)) = -H(\mathbf{g}(\mathbf{x}(t)))\nabla f(\mathbf{x}(t)) - \mu\, \nabla \sum_{i=1}^{m}[U(g_i(\mathbf{x}(t)))\, |g_i(\mathbf{x}(t))|] \tag{2.9}$$

ซึ่งถ้ามีการละเมิดเงื่อนไข($\mathbf{x}(t) \notin \Omega$) $H(\mathbf{g}(\mathbf{x}(t)))$ จะเท่ากับศูนย์ นั้นคือพจน์ที่หนึ่งทางขวาของสมการที่ (2.9) เป็นศูนย์ ดังนั้นสมการที่ (2.9) จะได้เป็น

$$\begin{aligned} \mathbf{s}_p(\mathbf{x}(t)) &= -\mu\, \nabla \sum_{i=1}^{m}[U(g_i(\mathbf{x}(t)))\, |g_i(\mathbf{x}(t))|] \\ &= -\mu\left(\sum_{i \in J(\mathbf{x}(t))} \nabla g_i(\mathbf{x}(t)) \right) \quad , \mathbf{x}(t) \notin \Omega \end{aligned} \tag{2.10}$$

และถ้าไม่มีการละเมิดเงื่อนไข($\mathbf{x}(t) \in \Omega$) $U(g_i(\mathbf{x}(t)))$, $i = 1, 2, \ldots, m$ จะเท่ากับศูนย์ นั้นคือพจน์ที่สองทางขวาของสมการที่ (2.9) เป็นศูนย์ ดังนั้นสมการที่ (2.9) จะได้เป็น

$$\begin{aligned} \mathbf{s}_p(\mathbf{x}(t)) &= -H(\mathbf{g}(\mathbf{x}(t)))\nabla f(\mathbf{x}(t)) \\ &= -\nabla f(\mathbf{x}(t)) \qquad\qquad , \mathbf{x}(t) \in \Omega \end{aligned} \tag{2.11}$$

สังเกตได้ว่า สมการที่ (2.10) และ (2.11) คือสมการที่ (2.7) นั่นเอง

จากปัญหา **P3** สามารถเขียนเป็นปัญหา **P4** ซึ่งอยู่ในรูปแบบที่ไม่ต่อเนื่องทางเวลา (discrete time) ได้ดังนี้คือ

$$\mathbf{P4:} \quad \mathbf{x}(k+1) = \mathbf{x}(k) + \frac{T}{\tau}\mathbf{s}_p(\mathbf{x}(k)) \qquad , k = 0,1,2,\ldots \qquad (2.12)$$

โดยที่
$$\mathbf{s}_g(\mathbf{x}(k)) = \frac{1}{\mu}\mathbf{s}_p(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \notin \Omega, k = 0,1,2,\ldots \qquad (2.13)$$

$$\mathbf{s}_f(\mathbf{x}(k)) = \mathbf{s}_p(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \in \Omega, k = 0,1,2,\ldots \qquad (2.14)$$

$$\mathbf{s}_g(\mathbf{x}(k)) \triangleq -\left(\sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k))\right) , \mathbf{x}(k) \notin \Omega, k = 0,1,2,\ldots \qquad (2.15)$$

$$\mathbf{s}_f(\mathbf{x}(k)) \triangleq -\nabla f(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \in \Omega, k = 0,1,2,\ldots \qquad (2.16)$$

โดยที่พารามิเตอร์ τ และ μ มีค่าคงที่เป็นจำนวนจริงบวก , $T = t_{k+1} - t_k$, k=0,1,2,... และสมการที่ (2.12) - (2.16) หาได้จากสมการที่ (2.3) - (2.6) ดังจะอธิบายได้ดังนี้คือ

สมการที่ (2.12) หาได้จากสมการที่ (2.5) ดังนี้คือ

$$(2.5) \Rightarrow \qquad \tau\left\{\frac{\mathbf{x}(k+1) - \mathbf{x}(k)}{T}\right\} = \mathbf{s}_p(\mathbf{x}(k)) \qquad , k = 0,1,2,\ldots \qquad (2.17)$$

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \frac{T}{\tau}\mathbf{s}_p(\mathbf{x}(k)) \qquad , k = 0,1,2,\ldots \qquad (2.18)$$

$$(2.6) \Rightarrow \qquad \mathbf{s}_p(\mathbf{x}(k)) = -\nabla p(\mathbf{x}(k)) \qquad , k = 0,1,2,\ldots \qquad (2.19)$$

เมื่อแทน $p(\mathbf{x}(k))$ ในสมการที่ (2.19) ด้วยสมการที่ (2.3) จะได้ว่า

$$\begin{aligned} \mathbf{s}_p(\mathbf{x}(k)) &= -\nabla\left[H(\mathbf{g}(\mathbf{x}(k)))\, f(\mathbf{x}(k)) + \mu\, q(\mathbf{x}(k))\right] \\ &= -H(\mathbf{g}(\mathbf{x}(k)))\, \nabla f(\mathbf{x}(k)) - \mu\, \nabla q(\mathbf{x}(k)) \qquad , k = 0,1,2,\ldots \qquad (2.20) \end{aligned}$$

เมื่อแทน $q(\mathbf{x}(k))$ ในสมการที่ (2.20) ด้วยสมการที่ (2.4) จะได้ว่า

$$\mathbf{s}_p(\mathbf{x}(k)) = -H(\mathbf{g}(\mathbf{x}(k)))\nabla f(\mathbf{x}(k)) - \mu\, \nabla\sum_{i=1}^{m}[U(g_i(\mathbf{x}(k)))|g_i(\mathbf{x}(k))|] \quad , k = 0,1,2,\ldots \qquad (2.21)$$

สมการที่ (2.13) และ (2.15) หาได้จากสมการที่ (2.21) ดังนี้คือ ถ้ามีการละเมิดเงื่อนไข ($\mathbf{x}(k) \notin \Omega$) $H(\mathbf{g}(\mathbf{x}(k)))$ จะเท่ากับศูนย์ นั้นคือพจน์ที่หนึ่งทางขวาของสมการที่ (2.21) เป็นศูนย์ ดังนั้นสมการที่ (2.21) จะได้เป็น

$$
\begin{aligned}
\mathbf{s}_p(\mathbf{x}(k)) &= -\mu \nabla \sum_{i=1}^{m} [U(g_i(\mathbf{x}(k)))\, |g_i(\mathbf{x}(k))|] \\
&= -\mu \left(\sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k)) \right) \\
&= \mu\, \mathbf{s}_g(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \notin \Omega,\ k = 0,1,2,\ldots \qquad (2.22)
\end{aligned}
$$

$$
\mathbf{s}_g(\mathbf{x}(k)) \triangleq -\left(\sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k)) \right) \qquad , \mathbf{x}(k) \notin \Omega,\ k = 0,1,2,\ldots \qquad (2.23)
$$

สมการที่ (2.14) และ (2.16) หาได้จากสมการที่ (2.21) ดังนี้คือ ถ้าไม่มีการละเมิดเงื่อนไข($\mathbf{x}(k) \in \Omega$) $U(g_i(\mathbf{x}(k)))$, $i = 1,2,\ldots,m$ จะเท่ากับศูนย์ นั้นคือพจน์ที่สองทางขวาของสมการที่ (2.21) เป็นศูนย์ ดังนั้นสมการที่ (2.21) จะได้เป็น

$$
\begin{aligned}
\mathbf{s}_p(\mathbf{x}(k)) &= -H(\mathbf{g}(\mathbf{x}(k)))\nabla f(\mathbf{x}(k)) \\
&= -\nabla f(\mathbf{x}(k)) \\
&= \mathbf{s}_f(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \in \Omega,\ k = 0,1,2,\ldots \qquad (2.24)
\end{aligned}
$$

$$
\mathbf{s}_f(\mathbf{x}(k)) \triangleq -\nabla f(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \in \Omega,\ k = 0,1,2,\ldots \qquad (2.25)
$$

สมการที่ (2.18) และ (2.22) - (2.25) คือปัญหา **P4** หรือคือสมการที่ (2.12) - (2.16) นั่นเอง ซึ่งอยู่ในรูปแบบที่ไม่ต่อเนื่องทางเวลา

3. หาผลตอบของปัญหา **P4** ได้จากการคำนวณด้วยเครื่องคอมพิวเตอร์ จากสมการอนุพันธ์ดังกล่าวเราสามารถใช้วิธีประมวลผลแบบขนานกันโดยใช้เครือข่ายนิวรอลสำหรับหาผลตอบของสมการอนุพันธ์ได้

อัลกอริทึมในงานวิจัยของ Glasos , Hui และ Zak

วิธี Steepest Descent หรือวิธี Gradient เป็นวิธีพื้นฐานง่ายๆ ซึ่งอาจกล่าวได้ว่าเป็นวิธีเชิงตัวเลข (Numerical method) วิธีหนึ่งที่ใช้ในการหาค่าออปติมัมแบบไม่มีเงื่อนไข ทิศทางการค้นหา (Search direction) ของวิธี Steepest Descent เป็นทิศทางที่ทำให้ Cost function ลดมากที่สุด ทิศทางการค้นหานี้เรียกว่า Steepest descent direction ซึ่งค้นพบโดย Cauchy เมื่อปี ค.ศ.1847 วิธี Steepest Descent อาจเรียกว่าเป็นวิธีอันดับหนึ่ง (First-order method) เนื่องจากทิศทางการค้นหาของวิธี Steepest Descent หาจากทิศตรงข้าม Gradient ของ Cost function

Glasos , Hui และ Zak ใช้เครือข่ายนิวรอลในการหาค่าออปติมัมแบบมีเงื่อนไข โดยใช้วิธี penalty function แปลงจากปัญหาการหาค่าออปติมัมแบบมีเงื่อนไขให้เป็นปัญหาการหาค่าออปติมัมแบบไม่มีเงื่อนไข จากนั้นแปลงปัญหาการหาค่าออปติมัมแบบไม่มีเงื่อนไขให้เป็นปัญหาของการแก้สมการอนุพันธ์ของตัวแปรสถานะซึ่งมี Forcing function เป็นฟังก์ชันของ Steepest descent ผลตอบของสมการอนุพันธ์เป็นผลตอบของปัญหาการหาค่าออปติมัมแบบไม่มีเงื่อนไข และผลตอบของสมการอนุพันธ์นี้หาได้จากการคำนวณด้วยเครื่องคอมพิวเตอร์ โดยที่สมการอนุพันธ์ดังกล่าวสามารถทำเป็น Parallel processing ได้ นั้นคือสามารถจัดให้อยู่ในรูปแบบทางคณิตศาสตร์ของเครือข่ายนิวรอล จึงทำให้เวลาที่ใช้ในการคำนวณรวดเร็ว

ถ้า *p*-norm (หรืออาจเรียกว่า L_p-norm) ของเวคเตอร์ $\mathbf{x} = [\ \mathrm{x}_1\ \ \mathrm{x}_2\ \dots\ \mathrm{x}_\mathrm{n}]^\mathrm{T}$ คือ

$$\|\mathbf{x}\|_p = \left(\sum_{\mathrm{i}=1}^{\mathrm{n}} |\mathrm{x}_\mathrm{i}|^p\right)^{1/p}$$

ดังนั้น 2-norm หรือ Euclidean หรือ L_2-norm ของเวคเตอร์ $\mathbf{x}$ จะได้เป็น

$$\|\mathbf{x}\|_2 = (\mathbf{x}^\mathrm{T}\mathbf{x})^{1/2} = \left(\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x}_\mathrm{i}^2\right)^{1/2}$$

โดยทั่วไปมักนิยมเขียน $\|\mathbf{x}\|_2$ ด้วย $\|\mathbf{x}\|$

อัลกอริทึมในงานวิจัยของ Glasos, Hui และ Zak สามารถแสดงขั้นตอนได้ดังนี้คือ

1. เริ่มต้นด้วย $\mathrm{k}=0$ เดาจุดเริ่มต้น $\mathbf{x}(\mathrm{k})$,

พารามิเตอร์ $\mu > 0$ ที่มีค่าคงที่เป็นจำนวนจริงบวก ,

พารามิเตอร์ $\lambda > 0$ หรือ $\dfrac{\mathrm{T}}{\tau} > 0$ ที่มีค่าคงที่เป็นจำนวนจริงบวก

และ ดรรชนีความแม่นยำ $\varepsilon > 0$ ที่มีค่าคงที่เป็นจำนวนจริงบวกค่าน้อยๆ

, เราอาจเดาค่า μ =0.8 และค่า λ = 0.01 เนื่องจากจากการทดลองทั้ง 3 ตัวอย่าง พบว่าเป็นค่าที่เหมาะสมสำหรับการหาค่าออปติมัม

2. ตรวจสอบ $\mathbf{x}(k)$ ว่ามีการละเมิดเงื่อนไขหรือไม่

2.1 ถ้าไม่มีการละเมิดเงื่อนไขคือ $g_i(\mathbf{x}(k)) \le 0$, $i = 1,2,\dots,m$ ให้ไปขั้นตอนที่ 3

2.2 ถ้ามีการละเมิดเงื่อนไขคือ$g_1(\mathbf{x}(k)) > 0$ หรือ $g_2(\mathbf{x}(k)) > 0$...หรือ$g_m(\mathbf{x}(k)) > 0$ เงื่อนไขใดเงื่อนไขหนึ่ง ให้ไปขั้นตอนที่ 4

3. คำนวณดังต่อไปนี้

$$\mathbf{s}_f(\mathbf{x}(k)) \triangleq -\nabla f(\mathbf{x}(k)) ,$$

$$\mathbf{s}_p(\mathbf{x}(k)) = \mathbf{s}_f(\mathbf{x}(k)) ,$$

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \lambda \mathbf{s}_p(\mathbf{x}(k)) ,$$

$$\frac{\|\mathbf{x}(k+1) - \mathbf{x}(k)\|}{\|\mathbf{x}(k)\|}$$

และ $f(\mathbf{x}(k))$

3.1 ถ้า $\dfrac{\|\mathbf{x}(k+1) - \mathbf{x}(k)\|}{\|\mathbf{x}(k)\|} \le \varepsilon$ จบการทำงาน นั่นคือ $\mathbf{x}^* = \mathbf{x}(k)$ ซึ่งเป็นจุดออปติมัม

3.2 ถ้า $\dfrac{\|\mathbf{x}(k+1) - \mathbf{x}(k)\|}{\|\mathbf{x}(k)\|} > \varepsilon$ ไปขั้นตอนที่ 5

4. คำนวณดังต่อไปนี้

$$\mathbf{s}_g(\mathbf{x}(k)) \triangleq -\left(\sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k)) \right)$$

และ $$\mathbf{s}_p(\mathbf{x}(k)) = \mu\, \mathbf{s}_g(\mathbf{x}(k))$$

5. คำนวณจุดใหม่ $\mathbf{x}(k+1) = \mathbf{x}(k) + \lambda \mathbf{s}_p(\mathbf{x}(k))$,

ให้ $k = k+1$ และไปขั้นตอนที่ 2

งานวิจัยในวิทยานิพนธ์นี้

การหาค่าออปติมัมของงานวิจัยในวิทยานิพนธ์นี้มี 3 ขั้นตอนดังนี้คือ

1. ใช้วิธี Penalty function แปลงจากปัญหาการหาค่าออปติมัมแบบมีเงื่อนไข (**P1**) ให้เป็นปัญหาการหาค่าออปติมัมแบบไม่มีเงื่อนไข (**P2**)

$$\mathbf{P1:} \quad \begin{cases} \text{หา } \mathbf{x} \text{ ซึ่งทำให้ } f(\mathbf{x}) \text{ มีค่าต่ำสุด} \\ \text{โดยที่} \quad g_i(\mathbf{x}) \le 0 \ , \ i = 1,2,\dots,m \end{cases} \qquad (2.26)$$

$\mathbf{x} \in \Re^n$, $f: \Re^n \to \Re$ และ $g_i: \Re^n \to \Re$, $i = 1,2,\dots,m$

P2 : $\{$ หา $\mathbf{x}$ ซึ่งทำให้ $p(\mathbf{x})$ มีค่าต่ำสุด (2.27)

ซึ่ง $p(\mathbf{x}) = H(\mathbf{g}(\mathbf{x}))\, f(\mathbf{x}) + \mu\, q(\mathbf{x})$ (2.28)

โดยที่ $H(.)$ เป็น Multidimensional step function คือ

$H(\mathbf{g}(\mathbf{x})) = 1$ เมื่อ $g_i(\mathbf{x}) \le 0$, $i = 1,2,\ldots,m$

และ $H(\mathbf{g}(\mathbf{x})) = 0$ เมื่อ $g_1(\mathbf{x}) > 0$ หรือ $g_2(\mathbf{x}) > 0\ldots$ หรือ $g_m(\mathbf{x}) > 0$

และถ้า $q(\mathbf{x})$ ในสมการที่ (2.28) เป็น

$$q(\mathbf{x}) = \sum_{i=1}^{m}[\, U(g_i(\mathbf{x}))\, |\, g_i(\mathbf{x})|\,]$$ สำหรับ Absolute value (2.29)

โดยที่ $U(.)$ เป็น Unidimensional step function คือ

$U(v) = 1$ เมื่อ $v > 0$

และ $U(v) = 0$ เมื่อ $v \le 0$

ถ้าให้ Ω เป็นเขตพื้นที่ที่ยอมรับ (Feasible region) นั้นคือ

$$\Omega = \bigcap_{i=1}^{m} \{\, \mathbf{x} \mid g_i(\mathbf{x}) \le 0 \,\}$$

ถ้าให้ Index set เป็น

$$J(\mathbf{x}) = \{\, i \mid g_i(\mathbf{x}) > 0 \,,\ i = 1,2,\ldots,m \,\}$$

2. ใช้วิธี Conjugate Gradient แก้ปัญหา **P2** จากเดิมในขั้นตอนที่ 2 ของงานวิจัยของ Glasos, Hui และ Zak นั้นแปลงจากปัญหาการหาค่าออปติมัมแบบไม่มีเงื่อนไข (**P2**) เป็นปัญหาของสมการอนุพันธ์ของตัวแปรสเตท (**P3**) ซึ่งมี Forcing function เป็นฟังก์ชันของ ***Steepest descent*** ซึ่งให้อัตราเร็วของการลู่เข้าหาค่าออปติมัมช้าและจะช้าลงมากเมื่อเริ่มเข้าใกล้จุดออปติมัม จึงทำให้ใช้เวลาในการคำนวณนานมากขึ้น จากข้อเสียดังกล่าวถ้าให้สมการอนุพันธ์ของตัวแปรสเตทใน **P3** มี Forcing function เป็นฟังก์ชันของ ***Conjugate gradient*** จะทำให้เวลาที่ใช้ในการคำนวณรวดเร็วขึ้น

ถ้าให้ $\mathbf{A}$ เป็น nxn symmetric matrix เราจะเรียก $\mathbf{x}$ และ $\mathbf{y}$ ว่า เป็น conjugate ซึ่งกันและกัน ก็ต่อเมื่อ $\mathbf{x}^T\mathbf{A}\mathbf{y} = 0$ จากคุณสมบัติดังกล่าวมีข้อดีคือไม่จำเป็นที่จะต้องก้าวไปในทิศทางที่ตั้งฉากเท่านั้น

$$\textbf{P3 :} \qquad \tau\,\dot{\mathbf{x}}(t_k) = \mathbf{s}_p(\mathbf{x}(t_k)) \;,\; k = 0,1,2,\ldots \tag{2.30}$$

โดยที่ $\mathbf{s}_p(\mathbf{x}(t_k))$ เป็น Conjugate direction ซึ่งหาได้จาก (2.31) และ (2.32) , $\mathbf{t}_{k+1} - \mathbf{t}_k = \mathbf{T}$, k=0,1,2,... และพารามิเตอร์ τ และ μ มีค่าคงที่เป็นจำนวนจริงบวก

$$\mathbf{s}_p(\mathbf{x}(t_k)) = -\nabla p(\mathbf{x}(t_k)) + \frac{\|\nabla p(\mathbf{x}(t_k))\|^2}{\|\nabla p(\mathbf{x}(t_{k-1}))\|^2}\,\mathbf{s}_p(\mathbf{x}(t_{k-1}))\,,k = 0,1,2,\ldots \tag{2.31}$$

$$\mathbf{s}_p(\mathbf{x}(t_{-1})) \triangleq \mathbf{0} \tag{2.32}$$

$$\nabla p(\mathbf{x}(t_k)) = \begin{cases} \mu\left(\displaystyle\sum_{i\in J(\mathbf{x}(t_k))}\nabla g_i(\mathbf{x}(t_k))\right) & ,\ \mathbf{x}(t_k) \notin \Omega\,,\ k = 0,1,2,\ldots \\ \nabla f(\mathbf{x}(t_k)) & ,\ \mathbf{x}(t_k) \in \Omega\,,\ k = 0,1,2,\ldots \end{cases} \tag{2.33}$$

$$\mathbf{s}_g(\mathbf{x}(t_k)) = \frac{1}{\mu}\mathbf{s}_p(\mathbf{x}(t_k)) \qquad ,\ \mathbf{x}(t_k) \notin \Omega\,,\ k = 0,1,2,\ldots \tag{2.34}$$

$$\mathbf{s}_f(\mathbf{x}(t_k)) = \mathbf{s}_p(\mathbf{x}(t_k)) \qquad ,\ \mathbf{x}(t_k) \in \Omega\,,\ k = 0,1,2,\ldots \tag{2.35}$$

$$\mathbf{s}_g(\mathbf{x}(t_k)) \triangleq -\left(\sum_{i\in J(\mathbf{x}(t_k))}\nabla g_i(\mathbf{x}(t_k))\right) + \frac{\mu\left\|\displaystyle\sum_{i\in J(\mathbf{x}(t_k))}\nabla g_i(\mathbf{x}(t_k))\right\|^2}{\|\nabla p(\mathbf{x}(t_{k-1}))\|^2}\mathbf{s}_p(\mathbf{x}(t_{k-1}))$$

$$,\ \mathbf{x}(t_k) \notin \Omega\,,\ k = 0,1,2,\ldots \tag{2.36}$$

$$\mathbf{s}_f(\mathbf{x}(t_k)) \triangleq -\nabla f(\mathbf{x}(t_k)) + \frac{\|\nabla f(\mathbf{x}(t_k))\|^2}{\|\nabla p(\mathbf{x}(t_{k-1}))\|^2}\mathbf{s}_p(\mathbf{x}(t_{k-1}))$$

$$,\ \mathbf{x}(t_k) \in \Omega\,,\ k = 0,1,2,\ldots \tag{2.37}$$

สมการที่ (2.33) - (2.37) หาได้จากสมการที่ (2.28) , (2.29) และ (2.31) ดังจะอธิบายต่อไปนี้คือ

สมการที่ (2.33) หาได้จากสมการที่ (2.28) และ (2.29) ดังนี้คือ จากสมการที่ (2.28) ถ้ามีการละเมิดเงื่อนไข($\mathbf{x}(t_k) \notin \Omega$) $H(\mathbf{g}(\mathbf{x}))$ จะเท่ากับศูนย์ ดังนั้นสมการที่ (2.28) จะได้เป็น

$$p(\mathbf{x}(t_k)) = \mu\, q(\mathbf{x}(t_k)) \qquad , \mathbf{x}(t_k) \notin \Omega,\ k = 0,1,2,\ldots \quad (2.38)$$

และเมื่อแทน $q(\mathbf{x}(t_k))$ ในสมการที่ (2.38) ด้วยสมการที่ (2.29) จะได้ว่า

$$\begin{aligned} p(\mathbf{x}(t_k)) &= \mu \sum_{i=1}^{m}[U(g_i(\mathbf{x}(t_k)))|g_i(\mathbf{x}(t_k))|] \\ &= \mu\left(\sum_{i \in J(\mathbf{x}(t_k))} g_i(\mathbf{x}(t_k))\right) \quad , \mathbf{x}(t_k) \notin \Omega,\ k = 0,1,2,\ldots \\ \nabla p(\mathbf{x}(t_k)) &= \mu\left(\sum_{i \in J(\mathbf{x}(t_k))} \nabla g_i(\mathbf{x}(t_k))\right) \quad , \mathbf{x}(t_k) \notin \Omega,\ k = 0,1,2,\ldots \quad (2.39) \end{aligned}$$

จากสมการที่ (2.28) ถ้าไม่มีการละเมิดเงื่อนไข($\mathbf{x}(t_k) \in \Omega$) $U(g_i(\mathbf{x}))$, $i = 1,2,\ldots,m$ จะเท่ากับศูนย์ ดังนั้นสมการที่ (2.28) จะได้เป็น

$$\begin{aligned} p(\mathbf{x}(t_k)) &= H(\mathbf{g}(\mathbf{x}(t_k)))\, f(\mathbf{x}(t_k)) \\ &= f(\mathbf{x}(t_k)) \qquad , \mathbf{x}(t_k) \in \Omega,\ k = 0,1,2,\ldots \\ \nabla p(\mathbf{x}(t_k)) &= \nabla f(\mathbf{x}(t_k)) \qquad , \mathbf{x}(t_k) \in \Omega,\ k = 0,1,2,\ldots \quad (2.40) \end{aligned}$$

สมการที่ (2.34) และ (2.36) หาได้จากสมการที่ (2.28) , (2.29) และ (2.31) ดังนี้คือ ถ้ามีการละเมิดเงื่อนไข($\mathbf{x}(t_k) \notin \Omega$) $H(\mathbf{g}(\mathbf{x}))$ จะเท่ากับศูนย์ ทำให้สมการที่ (2.28) กลายเป็น $p(\mathbf{x}(t_k)) = \mu\, q(\mathbf{x}(t_k))$ ซึ่งเมื่อแทน $p(\mathbf{x}(t_k))$ ในสมการที่ (2.31) ด้วย $\mu\, q(\mathbf{x}(t_k))$ จะได้สมการดังข้างล่างนี้คือ

$$\mathbf{s}_p(\mathbf{x}(t_k)) = -\mu\nabla q(\mathbf{x}(t_k)) + \frac{\|\mu\nabla q(\mathbf{x}(t_k))\|^2}{\|\nabla p(\mathbf{x}(t_{k-1}))\|^2}\mathbf{s}_p(\mathbf{x}(t_{k-1})) \quad , \mathbf{x}(t_k) \notin \Omega,\ k = 0,1,2,\ldots \quad (2.41)$$

และเมื่อแทน $q(\mathbf{x}(t_k))$ ในสมการที่ (2.41) ด้วยสมการที่ (2.29) จะได้ว่า

$$
\begin{aligned}
\mathbf{s}_p(\mathbf{x}(t_k)) &= -\mu\nabla\sum_{i=1}^{m}[U(g_i(\mathbf{x}(t_k)))|g_i(\mathbf{x}(t_k))|] + \frac{\left\|\mu\nabla\sum_{i=1}^{m}[U(g_i(\mathbf{x}(t_k)))|g_i(\mathbf{x}(t_k))|]\right\|^2}{\|\nabla p(\mathbf{x}(t_{k-1}))\|^2}\mathbf{s}_p(\mathbf{x}(t_{k-1})) \\
&= -\mu\left(\sum_{i\in J(\mathbf{x}(t_k))}\nabla g_i(\mathbf{x}(t_k))\right) + \frac{\left\|\mu\left(\sum_{i\in J(\mathbf{x}(t_k))}\nabla g_i(\mathbf{x}(t_k))\right)\right\|^2}{\|\nabla p(\mathbf{x}(t_{k-1}))\|^2}\mathbf{s}_p(\mathbf{x}(t_{k-1})) \\
&= \mu\left(-\left(\sum_{i\in J(\mathbf{x}(t_k))}\nabla g_i(\mathbf{x}(t_k))\right) + \frac{\mu\left\|\sum_{i\in J(\mathbf{x}(t_k))}\nabla g_i(\mathbf{x}(t_k))\right\|^2}{\|\nabla p(\mathbf{x}(t_{k-1}))\|^2}\mathbf{s}_p(\mathbf{x}(t_{k-1}))\right) \\
&= \mu\,\mathbf{s}_g(\mathbf{x}(t_k)) \qquad , \mathbf{x}(t_k) \notin \Omega,\ k = 0,1,2,\ldots \qquad (2.42)
\end{aligned}
$$

$$
\mathbf{s}_g(\mathbf{x}(t_k)) \triangleq -\left(\sum_{i\in J(\mathbf{x}(t_k))}\nabla g_i(\mathbf{x}(t_k))\right) + \frac{\mu\left\|\sum_{i\in J(\mathbf{x}(t_k))}\nabla g_i(\mathbf{x}(t_k))\right\|^2}{\|\nabla p(\mathbf{x}(t_{k-1}))\|^2}\mathbf{s}_p(\mathbf{x}(t_{k-1}))
$$

$$
, \mathbf{x}(t_k) \notin \Omega,\ k = 0,1,2,.. \qquad (2.43)
$$

สมการที่ (2.35) และ (2.37) หาได้จากสมการที่ (2.28) และ (2.31) ดังนี้คือ ถ้าไม่มีการละเมิดเงื่อนไข($\mathbf{x}(t_k) \in \Omega$) $U(g_i(\mathbf{x}))$, $i = 1,2,\ldots,m$ จะเท่ากับศูนย์ ทำให้สมการที่ (2.28) กลายเป็น $p(\mathbf{x}(t_k)) = H(\mathbf{g}(\mathbf{x}(t_k)))\,f(\mathbf{x}(t_k)) = f(\mathbf{x}(t_k))$ ซึ่งเมื่อแทน $p(\mathbf{x}(t_k))$ ในสมการที่ (2.31) ด้วย $f(\mathbf{x}(t_k))$ จะได้สมการดังข้างล่างคือ

$$
\begin{aligned}
\mathbf{s}_p(\mathbf{x}(t_k)) &= -\nabla f(\mathbf{x}(t_k)) + \frac{\|\nabla f(\mathbf{x}(t_k))\|^2}{\|\nabla p(\mathbf{x}(t_{k-1}))\|^2}\mathbf{s}_p(\mathbf{x}(t_{k-1})) \\
&= \mathbf{s}_f(\mathbf{x}(t_k)) \qquad , \mathbf{x}(t_k) \in \Omega,\ k = 0,1,2,\ldots \qquad (2.44)
\end{aligned}
$$

$$\mathbf{s}_f(\mathbf{x}(t_k)) \triangleq -\nabla f(\mathbf{x}(t_k)) + \frac{\left\|\nabla f(\mathbf{x}(t_k))\right\|^2}{\left\|\nabla p(\mathbf{x}(t_{k-1}))\right\|^2}\mathbf{s}_p(\mathbf{x}(t_{k-1}))$$

$$, \mathbf{x}(t_k) \in \Omega \text{ , } k = 0,1,2,.. \qquad (2.45)$$

สังเกตได้ว่า สมการที่ (2.39) , (2.40) และ (2.42) -(2.45) คือสมการที่ (2.33) - (2.37) ในปัญหา **P3** ซึ่งอยู่ในรูปแบบที่ต่อเนื่องทางเวลา

จากปัญหา **P3** สามารถเขียนเป็นปัญหา **P4** ซึ่งอยู่ในรูปแบบที่ไม่ต่อเนื่องทางเวลาได้ดังนี้คือ

$$\mathbf{P4}: \ \mathbf{x}(k+1) = \mathbf{x}(k) + \frac{T}{\tau}\mathbf{s}_p(\mathbf{x}(k)) \qquad , k = 0,1,2,\ldots \qquad (2.46)$$

โดยที่ $\mathbf{s}_g(\mathbf{x}(k)) = \frac{1}{\mu}\mathbf{s}_p(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \notin \Omega \text{ , } k = 0,1,2,\ldots \qquad (2.47)$

$$\mathbf{s}_f(\mathbf{x}(k)) = \mathbf{s}_p(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \in \Omega \text{ , } k = 0,1,2,\ldots \qquad (2.48)$$

$$\nabla p(\mathbf{x}(k)) = \left\{ \begin{array}{ll} \mu\left(\sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k))\right) & , \mathbf{x}(k) \notin \Omega \text{ , } k = 0,1,2,\ldots \\ \nabla f(\mathbf{x}(k)) & , \mathbf{x}(k) \in \Omega \text{ , } k = 0,1,2,\ldots \end{array} \right\} \qquad (2.49)$$

$$\mathbf{s}_g(\mathbf{x}(k)) \triangleq -\left(\sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k))\right) + \frac{\mu\left\|\sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k))\right\|^2}{\left\|\nabla p(\mathbf{x}(k-1))\right\|^2}\mathbf{s}_p(\mathbf{x}(k-1))$$

$$, \mathbf{x}(k) \notin \Omega \text{ , } k = 0,1,2,\ldots \qquad (2.50)$$

$$\mathbf{s}_f(\mathbf{x}(k)) \triangleq -\nabla f(\mathbf{x}(k)) + \frac{\left\|\nabla f(\mathbf{x}(k))\right\|^2}{\left\|\nabla p(\mathbf{x}(k-1))\right\|^2}\mathbf{s}_p(\mathbf{x}(k-1))$$

$$, \mathbf{x}(k) \in \Omega \text{ , } k = 0,1,2,\ldots \qquad (2.51)$$

$$\mathbf{s}_p(\mathbf{x}(-1)) \triangleq \mathbf{0} \qquad (2.52)$$

โดยที่พารามิเตอร์ τ และ μ มีค่าคงที่เป็นจำนวนจริงบวก , $T = t_{k+1} - t_k$, k=0,1,2,... และสมการที่ (2.46) - (2.52) หาได้จากสมการที่ (2.28) - (2.32) ดังจะอธิบายต่อไปนี้คือ

สมการที่ (2.49) หาได้จากสมการที่ (2.28) และ (2.29) ดังนี้คือ จากสมการที่ (2.28) ถ้ามีการละเมิดเงื่อนไข($\mathbf{x}(k) \notin \Omega$) $H(\mathbf{g}(\mathbf{x}))$ จะเท่ากับศูนย์ ดังนั้น สมการที่ (2.28) จะได้เป็น

$$p(\mathbf{x}(k)) = \mu\, q(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \notin \Omega,\ k = 0,1,2,\ldots \tag{2.53}$$

และเมื่อแทน $q(\mathbf{x}(k))$ ในสมการที่ (2.53) ด้วยสมการที่ (2.29) จะได้ว่า

$$\begin{aligned} p(\mathbf{x}(k)) &= \mu \sum_{i=1}^{m} [U(g_i(\mathbf{x}(k)))|g_i(\mathbf{x}(k))|] \\ &= \mu\left(\sum_{i \in J(\mathbf{x}(k))} g_i(\mathbf{x}(k))\right) \qquad , \mathbf{x}(k) \notin \Omega,\ k = 0,1,2,\ldots \\ \nabla p(\mathbf{x}(k)) &= \mu\left(\sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k))\right) \qquad , \mathbf{x}(k) \notin \Omega,\ k = 0,1,2,.. \end{aligned} \tag{2.54}$$

จากสมการที่ (2.28) ถ้าไม่มีการละเมิดเงื่อนไข($\mathbf{x}(k) \in \Omega$) $U(g_i(\mathbf{x}))$, $i = 1,2,\ldots,m$ จะเท่ากับศูนย์ ดังนั้นสมการที่ (2.28) จะได้เป็น

$$\begin{aligned} p(\mathbf{x}(k)) &= H(\mathbf{g}(\mathbf{x}(k)))\, f(\mathbf{x}(k)) \\ &= f(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \in \Omega,\ k = 0,1,2,\ldots \\ \nabla p(\mathbf{x}(k)) &= \nabla f(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \in \Omega,\ k = 0,1,2,\ldots \end{aligned} \tag{2.55}$$

สมการที่ (2.46) และ (2.52) หาได้จากสมการที่ (2.30) และ (2.32) และจากสมการที่ (2.30) - (2.32) สามารถจัดให้อยู่ในรูปแบบที่ไม่ต่อเนื่องทางเวลาได้ดังนี้คือ

$$(2.30) \Rightarrow \tau\left\{\frac{\mathbf{x}(k+1) - \mathbf{x}(k)}{T}\right\} = \mathbf{s}_p(\mathbf{x}(k)) \qquad , k=0,1,2,\ldots \tag{2.56}$$

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \frac{T}{\tau}\mathbf{s}_p(\mathbf{x}(k)) \qquad , k=0,1,2,\ldots \tag{2.57}$$

$$(2.31) \Rightarrow \quad \mathbf{s}_p(\mathbf{x}(k)) = -\nabla p(\mathbf{x}(k)) + \frac{\|\nabla p(\mathbf{x}(k))\|^2}{\|\nabla p(\mathbf{x}(k-1))\|^2}\mathbf{s}_p(\mathbf{x}(k-1))\ , k=0,1,2,\ldots \tag{2.58}$$

$$(2.32) \Rightarrow \quad \mathbf{s}_p(\mathbf{x}(-1)) \triangleq \mathbf{0} \tag{2.59}$$

สมการที่ (2.47) และ (2.50) หาได้จากสมการที่ (2.28) , (2.29) และ (2.58) ดังนี้คือ ถ้ามีการละเมิดเงื่อนไข($\mathbf{x}(k) \notin \Omega$) $H(\mathbf{g}(\mathbf{x}))$ จะเท่ากับศูนย์ ทำให้สมการที่ (2.28) กลายเป็น $p(\mathbf{x}(k)) = \mu\, q(\mathbf{x}(k))$ ซึ่งเมื่อแทน $p(\mathbf{x}(k))$ ในสมการที่ (2.58) ด้วย $\mu\, q(\mathbf{x}(k))$ จะได้สมการดังข้างล่างนี้คือ

$$\mathbf{s}_p(\mathbf{x}(k)) = -\mu\nabla q(\mathbf{x}(k)) + \frac{\left\|\mu\nabla q(\mathbf{x}(k))\right\|^2}{\left\|\nabla p(\mathbf{x}(k-1))\right\|^2}\mathbf{s}_p(\mathbf{x}(k-1))\,, \mathbf{x}(k) \notin \Omega, k = 0,1,2,\ldots \quad (2.60)$$

และเมื่อแทน $q(\mathbf{x}(k))$ ในสมการที่ (2.60) ด้วยสมการที่ (2.29) จะได้ว่า

$$\begin{aligned}
\mathbf{s}_p(\mathbf{x}(k)) &= -\mu\nabla\sum_{i=1}^{m}[U(g_i(\mathbf{x}(k)))\left|g_i(\mathbf{x}(k))\right|] + \frac{\left\|\mu\nabla\sum_{i=1}^{m}[U(g_i(\mathbf{x}(k)))\left|g_i(\mathbf{x}(k))\right|]\right\|^2}{\left\|\nabla p(\mathbf{x}(k-1))\right\|^2}\mathbf{s}_p(\mathbf{x}(k-1)) \\
&= -\mu\left(\sum_{i\in J(\mathbf{x}(k))}\nabla g_i(\mathbf{x}(k))\right) + \frac{\left\|\mu\left(\sum_{i\in J(\mathbf{x}(k))}\nabla g_i(\mathbf{x}(k))\right)\right\|^2}{\left\|\nabla p(\mathbf{x}(k-1))\right\|^2}\mathbf{s}_p(\mathbf{x}(k-1)) \\
&= \mu\left(-\left(\sum_{i\in J(\mathbf{x}(k))}\nabla g_i(\mathbf{x}(k))\right) + \frac{\mu\left\|\sum_{i\in J(\mathbf{x}(k))}\nabla g_i(\mathbf{x}(k))\right\|^2}{\left\|\nabla p(\mathbf{x}(k-1))\right\|^2}\mathbf{s}_p(\mathbf{x}(k-1))\right) \\
&= \mu\,\mathbf{s}_g(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \notin \Omega\,,\ k = 0,1,2,\ldots
\end{aligned} \quad (2.61)$$

$$\mathbf{s}_g(\mathbf{x}(k)) \triangleq -\left(\sum_{i\in J(\mathbf{x}(k))}\nabla g_i(\mathbf{x}(k))\right) + \frac{\mu\left\|\sum_{i\in J(\mathbf{x}(k))}\nabla g_i(\mathbf{x}(k))\right\|^2}{\left\|\nabla p(\mathbf{x}(k-1))\right\|^2}\mathbf{s}_p(\mathbf{x}(k-1))$$
$$, \mathbf{x}(k) \notin \Omega\,,\ k = 0,1,2,\ldots \quad (2.62)$$

สมการที่ (2.48) และ (2.51) หาได้จากสมการที่ (2.28) และ (2.58) ดังนี้คือ ถ้าไม่มีการละเมิดเงื่อนไข($\mathbf{x}(k) \in \Omega$) $U(g_i(\mathbf{x}))$, $i = 1,2,\ldots,m$ จะเท่ากับศูนย์ ทำให้สมการที่ (2.28) กลายเป็น $p(\mathbf{x}(k)) = H(\mathbf{g}(\mathbf{x}(k)))\, f(\mathbf{x}(k)) = f(\mathbf{x}(k))$ ซึ่งเมื่อแทน $p(\mathbf{x}(k))$ ในสมการที่ (2.58) ด้วย $f(\mathbf{x}(k))$ จะได้สมการดังข้างล่างคือ

$$\begin{aligned} \mathbf{s}_p(\mathbf{x}(k)) &= -\nabla f(\mathbf{x}(k)) + \frac{\|\nabla f(\mathbf{x}(k))\|^2}{\|\nabla p(\mathbf{x}(k-1))\|^2}\, \mathbf{s}_p(\mathbf{x}(k-1)) \\ &= \mathbf{s}_f(\mathbf{x}(k)) \qquad , \mathbf{x}(k) \in \Omega \, , \; k = 0,1,2,\ldots \qquad (2.63) \end{aligned}$$

$$\mathbf{s}_f(\mathbf{x}(k)) \triangleq -\nabla f(\mathbf{x}(k)) + \frac{\|\nabla f(\mathbf{x}(k))\|^2}{\|\nabla p(\mathbf{x}(k-1))\|^2}\, \mathbf{s}_p(\mathbf{x}(k-1))$$
$$, \mathbf{x}(k) \in \Omega \, , \; k = 0,1,2,\ldots \qquad (2.64)$$

จากสมการที่ (2.54) , (2.55) , (2.57) , (2.59) และ (2.61) - (2.64) คือปัญหา **P4** หรือคือสมการที่ (2.46) -(2.52) นั้นเอง ซึ่งอยู่ในรูปแบบที่ไม่ต่อเนื่องทางเวลา

3. หาผลตอบของปัญหา **P4** ได้จากการคำนวณด้วยเครื่องคอมพิวเตอร์ จากสมการอนุพันธ์ดังกล่าวเราสามารถใช้วิธีประมวลผลแบบขนานกันโดยใช้เครือข่ายนิวรอลสำหรับหาผลตอบของสมการอนุพันธ์ได้

การพิสูจน์การลู่เข้าหาค่าออปติมัมของวิธี Conjugate Gradient

วิธี Conjugate Gradient เป็นวิธีที่พัฒนามาจากวิธี Steepest Descent เพื่อให้อัตราการลู่เข้าค่าออปติมัมเร็วขึ้น การพิสูจน์การลู่เข้าหาค่าออปติมัมของวิธี Conjugate Gradient สามารถพิสูจน์ได้ดังนี้คือ ถ้าให้ $\mathbf{s}_p(\mathbf{x}(k))$ เป็น Conjugate direction ก้าวที่ k และ $p(\mathbf{x}(k))$ เป็น Penalty Function จะได้ว่า

จากสมการที่ (2.56) , (2.58) และ (2.59) สามารถเขียนใหม่ได้เป็น

$$\tau\left\{\frac{\mathbf{x}(k+1)-\mathbf{x}(k)}{T}\right\} = \mathbf{s}_p(\mathbf{x}(k)) \qquad , k = 0,1,2,\ldots \quad (2.65)$$

$$\mathbf{s}_p(\mathbf{x}(k)) = -\nabla p(\mathbf{x}(k)) + \frac{\|\nabla p(\mathbf{x}(k))\|^2}{\|\nabla p(\mathbf{x}(k-1))\|^2}\mathbf{s}_p(\mathbf{x}(k-1)) \;, k = 0,1,2,\ldots \quad (2.66)$$

$$\mathbf{s}_p(\mathbf{x}(-1)) \triangleq \mathbf{0} \quad (2.67)$$

จากสมการที่ (2.28) สามารถเขียนใหม่ได้เป็น

$$p(\mathbf{x}(k)) = H(\mathbf{g}(\mathbf{x}(k)))\, f(\mathbf{x}(k)) + \mu\, q(\mathbf{x}(k)) \quad , k = 0,1,2,\ldots \quad (2.68)$$

โดยที่พารามิเตอร์ τ และ μ มีค่าคงที่เป็นจำนวนจริงบวก , $T = t_{k+1} - t_k$, k=0,1,2,...
ถ้าให้ $\lambda = \dfrac{T}{\tau}$ ดังนั้นสมการที่ (2.65) สามารถเขียนใหม่ได้ดังนี้คือ

$$\left\{\frac{\mathbf{x}(k+1)-\mathbf{x}(k)}{\lambda}\right\} = \mathbf{s}_p(\mathbf{x}(k)) \qquad ,k = 0,1,2,\ldots$$

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \lambda\, \mathbf{s}_p(\mathbf{x}(k)) \qquad ,k = 0,1,2,\ldots$$

$$\mathbf{x}(k) = \mathbf{x}(k-1) + \lambda\, \mathbf{s}_p(\mathbf{x}(k-1)) \quad ,k = 1,2,3,\ldots \quad (2.69)$$

และถ้าให้ λ^* เป็น λ ที่ทำให้ $p(\mathbf{x}(k))$ มีค่าต่ำสุด λ^*หาได้จาก $\dfrac{d\,p(\mathbf{x}(k))}{d\,\lambda} = 0$ ดังนั้นจะได้ว่า

$$\frac{d\,p(\mathbf{x}(k))}{d\,\lambda} = \frac{\partial p(\mathbf{x}(k))}{\partial \mathbf{x}(k)}^{T} \frac{d\,\mathbf{x}(k)}{d\,\lambda} = 0 \;, k = 0,1,2,\ldots \quad (2.70)$$

เราสามารถหา $\dfrac{d\,\mathbf{x}(k)}{d\,\lambda}$ ในสมการที่ (2.70) ได้จากสมการที่ (2.69) คือ

$$(2.69) \Rightarrow \qquad \frac{d\mathbf{x}(k)}{d\,\lambda} = \frac{d\left[\mathbf{x}(k-1)+\lambda\,\mathbf{s}_p(\mathbf{x}(k-1))\right]}{d\,\lambda} \quad , \; k = 1,2,3,\ldots$$

$$= \mathbf{s}_p(\mathbf{x}(k-1)) \quad , \; k = 1,2,3,.. \qquad (2.71)$$

และจากสมการที่ (2.70) สามารถเขียนใหม่เป็น

$$(2.70) \Rightarrow \qquad \frac{\partial p(\mathbf{x}(k))}{\partial \mathbf{x}(k)}^{T} \frac{d\,\mathbf{x}(k)}{d\,\lambda} = 0 \quad , \; k = 0,1,2,\ldots$$

$$\nabla p(\mathbf{x}(k))^{T} \frac{d\,\mathbf{x}(k)}{d\,\lambda} = 0 \quad , \; k = 0,1,2,\ldots \qquad (2.72)$$

ใช้สมการที่ (2.71) ในสมการที่ (2.72) และจากสมการที่ (2.67) $\mathbf{s}_p(\mathbf{x}(-1)) \triangleq \mathbf{0}$ จะได้ว่า

$$\nabla p(\mathbf{x}(k))^{T} \mathbf{s}_p(\mathbf{x}(k-1)) = 0 \quad , k = 0,1,2,\ldots \qquad (2.73)$$

จากสมการที่ (2.65) และ (2.66) สามารถเขียนใหม่ได้เป็น

$$\boldsymbol{\tau}\left\{\frac{\mathbf{x}(k+1) - \mathbf{x}(k)}{T}\right\} = -\nabla p(\mathbf{x}(k)) + \frac{\|\nabla p(\mathbf{x}(k))\|^2}{\|\nabla p(\mathbf{x}(k-1))\|^2}\mathbf{s}_p(\mathbf{x}(k-1)) \, , \; k = 0,1,2,\ldots \qquad (2.74)$$

และเมื่อคูณด้านหน้าของสมการที่ (2.74) ด้วย $\nabla p(\mathbf{x}(k))^{T}$ ทั้งสองข้าง จะได้ว่า

$$\nabla p(\mathbf{x}(k))^{T}\boldsymbol{\tau}\left\{\frac{\mathbf{x}(k+1)-\mathbf{x}(k)}{T}\right\} = -\nabla p(\mathbf{x}(k))^{T}\nabla p(\mathbf{x}(k)) + \frac{\|\nabla p(\mathbf{x}(k))\|^2}{\|\nabla p(\mathbf{x}(k-1))\|^2}\nabla p(\mathbf{x}(k))^{T}\mathbf{s}_p(\mathbf{x}(k-1))$$

$$, \; k = 0,1,2,\ldots \qquad (2.75)$$

ใช้สมการที่ (2.73) ในพจน์ที่สองทางด้านขวาของสมการที่ (2.75) ดังนั้นจะได้ว่า

$$\nabla p(\mathbf{x}(k))^T \tau \left\{ \frac{\mathbf{x}(k+1) - \mathbf{x}(k)}{T} \right\} = -\nabla p(\mathbf{x}(k))^T \nabla p(\mathbf{x}(k)) \quad , k = 0,1,2,\dots$$

$$\nabla p(\mathbf{x}(k))^T \left\{ \frac{\mathbf{x}(k+1) - \mathbf{x}(k)}{T} \right\} = -\frac{1}{\tau} \nabla p(\mathbf{x}(k))^T \nabla p(\mathbf{x}(k)) \quad , k = 0,1,2,\dots \qquad (2.76)$$

ในการพิสูจน์การลู่เข้าหาค่าออปติมัม เราสามารถพิสูจน์ได้โดยการพิจารณา $\frac{dp}{dt}$ ถ้า $\frac{dp}{dt}$ น้อยกว่าหรือเท่ากับศูนย์แสดงว่า p มีการเปลี่ยนแปลงในแนวทางที่ลดลงหรือไม่มีการเปลี่ยนแปลงตามลำดับ เหตุการณ์ดังกล่าวแสดงว่า p ลู่เข้าหาค่าๆหนึ่งหรือค่า minimum นั่นเอง ซึ่งเขียนได้ดังนี้คือ

$$\begin{aligned} \frac{dp}{dt} &= \frac{\partial p}{\partial \mathbf{x}}^T \frac{d\mathbf{x}}{dt} \\ &= \nabla p(\mathbf{x})^T \dot{\mathbf{x}} \end{aligned} \qquad (2.77)$$

จากสมการที่ (2.77) สามารถจัดให้อยู่ในรูปแบบที่ไม่ต่อเนื่องทางเวลาได้ดังนี้คือ

$$\left\{ \frac{p(\mathbf{x}(k+1)) - p(\mathbf{x}(k))}{T} \right\} = \nabla p(\mathbf{x}(k))^T \left\{ \frac{\mathbf{x}(k+1) - \mathbf{x}(k)}{T} \right\} \quad , k=0,1,2,\dots \qquad (2.78)$$

และเมื่อแทนพจน์ทางขวาของสมการที่ (2.78) ด้วยสมการที่ (2.76) จะได้ว่า

$$\begin{aligned} \left\{ \frac{p(\mathbf{x}(k+1)) - p(\mathbf{x}(k))}{T} \right\} &= -\frac{1}{\tau} \nabla p(\mathbf{x}(k))^T \nabla p(\mathbf{x}(k)) \quad , k=0,1,2,\dots \\ &= -\frac{1}{\tau} \left\| \nabla p(\mathbf{x}(k)) \right\|^2 \quad , k=0,1,2,\dots \end{aligned} \qquad (2.79)$$

ซึ่ง $\because$ พารามิเตอร์ $\tau > 0$ และ $\left\| \nabla p(\mathbf{x}(k)) \right\|^2 \geq 0, k=0,1,2,\dots$ ดังนั้นจากสมการที่ (2.79) จะได้ว่า

$$\left\{ \frac{p(\mathbf{x}(k+1)) - p(\mathbf{x}(k))}{T} \right\} \leq 0 \qquad (2.80$$

จากสมการที่ (2.80) พิสูจน์ได้ว่า การหาค่าออปติมัมโดยใช้วิธี Conjugate Gradient นั้น สามารถหาค่า minimum ได้ เนื่องจากพจน์ทางด้านขวาน้อยกว่าหรือเท่ากับศูนย์เสมอ จากเหตุการณ์ดังกล่าวอาจแยกวิเคราะห์การลู่เข้าหาค่าออปติมัมเป็น 2 กรณีดังนี้คือ

1. เมื่อ $\mathbf{x}(k)$ มีการละเมิดเงื่อนไขคือ $g_1(\mathbf{x}(k)) > 0$ หรือ $g_2(\mathbf{x}(k)) > 0$ ... หรือ $g_m(\mathbf{x}(k)) > 0$ เงื่อนไขใดเงื่อนไขหนึ่ง $H(\mathbf{g}(\mathbf{x}))$ จะเท่ากับศูนย์ ทำให้สมการที่ (2.28) กลายเป็น $p(\mathbf{x}(k)) = \mu\, q(\mathbf{x}(k))$ หรือ $p(\mathbf{x}(k)) = \mu \sum_{i=1}^{m}[U(g_i(\mathbf{x}(k)))|g_i(\mathbf{x}(k))|]$ นั้นหมายความว่า เมื่อ $p(\mathbf{x}(k))$ ลดลง $g_i(\mathbf{x}(k))$ จะลดลง ($g_i(\mathbf{x}(k))$ ละเมิดเงื่อนไขน้อยลง) และเมื่อ $p(\mathbf{x}(k))$ ลดลงจนกระทั่ง $g_i(\mathbf{x}(k))$ ไม่มีการละเมิดเงื่อนไข $U(g_i(\mathbf{x}))$, $i = 1,2,\ldots,m$ จะเท่ากับศูนย์ ทำให้สมการที่ (28) กลายเป็น $p(\mathbf{x}(k)) = H(\mathbf{g}(\mathbf{x}(k)))\, f(\mathbf{x}(k)) = f(\mathbf{x}(k))$ นั้นคือเมื่อ $p(\mathbf{x}(k))$ ลู่เข้าหาค่าออปติมัมก็คือ $f(\mathbf{x}(k))$ ลู่เข้าหาค่าออปติมัมนั้นเอง หรือเมื่อ $p(\mathbf{x}(k))$ ลดลงจนกระทั่ง $g_i(\mathbf{x}(k))$ มีการละเมิดเงื่อนไข ซึ่งก็จะกระทำเช่นเดิมจนกว่าจะลู่เข้าหาค่าออปติมัม

2. เมื่อ $\mathbf{x}(k)$ไม่มีการละเมิดเงื่อนไขคือ$g_i(\mathbf{x}(k)) < 0, i = 1,2,\ldots,m$ $U(g_i(\mathbf{x})), i = 1,2,\ldots,m$ จะเท่ากับศูนย์ ทำให้สมการที่ (28) กลายเป็น$p(\mathbf{x}(k)) = H(\mathbf{g}(\mathbf{x}(k)))\, f(\mathbf{x}(k)) = f(\mathbf{x}(k))$ นั้นคือ เมื่อ $p(\mathbf{x}(k))$ ลู่เข้าหาค่าออปติมัมก็คือ $f(\mathbf{x}(k))$ ลู่เข้าหาค่าออปติมัมนั้นเอง หรือเมื่อ$p(\mathbf{x}(k))$ ลดลงจนกระทั่ง $g_i(\mathbf{x}(k))$ มีการละเมิดเงื่อนไข ซึ่งก็คือกรณีที่ 1 นั้นเอง

อัลกอริทึมในงานวิจัยของวิทยานิพนธ์นี้

วิธี Conjugate Gradient เป็นวิธีที่ง่ายและพัฒนามาจากวิธี Steepest Descent โดยปรับปรุงอัตราการลู่เข้าหาค่าออปติมัมของวิธี Steepest Descent เพื่อให้ลู่เข้าหาค่าออปติมัมเร็วขึ้น

อัลกอริทึมในงานวิจัยของวิทยานิพนธ์นี้สามารถแสดงขั้นตอนได้ดังนี้คือ

1. เริ่มต้นด้วย k=0 เดาจุดเริ่มต้น $\mathbf{x}(k)$,

$\mathbf{s}_p(\mathbf{x}(-1)) \triangleq \mathbf{0}$,

พารามิเตอร์ $\mu > 0$ ที่มีค่าคงที่เป็นจำนวนจริงบวก ,

พารามิเตอร์ $\lambda > 0$ หรือ $\frac{T}{\tau} > 0$ ที่มีค่าคงที่เป็นจำนวนจริงบวก

และ ดรรชนีความแม่นยำ $\varepsilon > 0$ ที่มีค่าคงที่เป็นจำนวนจริงบวกค่าน้อยๆ

, เราอาจเดาค่า μ =0.8 และค่า λ = 0.01 เนื่องจากจากการทดลองทั้ง 3 ตัวอย่าง พบว่าเป็นค่าที่เหมาะสมสำหรับการหาค่าออปติมัม

2. ตรวจสอบ $\mathbf{x}(k)$ ว่ามีการละเมิดเงื่อนไขหรือไม่

   2.1 ถ้าไม่มีการละเมิดเงื่อนไขคือ $g_i(\mathbf{x}(k)) \le 0$, $i = 1,2,...,m$ ให้ไปขั้นตอนที่ 3

   2.2 ถ้ามีการละเมิดเงื่อนไขคือ $g_1(\mathbf{x}(k)) > 0$ หรือ $g_2(\mathbf{x}(k)) > 0$...หรือ $g_m(\mathbf{x}(k)) > 0$ เงื่อนไขใดเงื่อนไขหนึ่ง ให้ไปขั้นตอนที่ 4

3. คำนวณดังต่อไปนี้

$$\mathbf{s}_f(\mathbf{x}(k)) \triangleq -\nabla f(\mathbf{x}(k)) + \frac{\|\nabla f(\mathbf{x}(k))\|^2}{\|\nabla p(\mathbf{x}(k-1))\|^2}\mathbf{s}_p(\mathbf{x}(k-1)) \quad ,$$

$$\mathbf{s}_p(\mathbf{x}(k)) = \mathbf{s}_f(\mathbf{x}(k)) \quad ,$$

$$\nabla p(\mathbf{x}(k)) = \nabla f(\mathbf{x}(k)) \quad ,$$

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \lambda \mathbf{s}_p(\mathbf{x}(k)) \quad ,$$

$$\frac{\|\mathbf{x}(k+1) - \mathbf{x}(k)\|}{\|\mathbf{x}(k)\|}$$

และ $f(\mathbf{x}(k))$

   3.1 ถ้า $\dfrac{\|\mathbf{x}(k+1) - \mathbf{x}(k)\|}{\|\mathbf{x}(k)\|} \le \varepsilon$ จบการทำงาน นั่นคือ $\mathbf{x}^* = \mathbf{x}(k)$ ซึ่งเป็นจุดออปติมัม

   3.2 ถ้า $\dfrac{\|\mathbf{x}(k+1) - \mathbf{x}(k)\|}{\|\mathbf{x}(k)\|} > \varepsilon$ ไปขั้นตอนที่ 5

4. คำนวณดังต่อไปนี้

$$\mathbf{s}_g(\mathbf{x}(k)) \triangleq -\left(\sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k))\right) + \frac{\mu \left\| \sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k)) \right\|^2}{\|\nabla p(\mathbf{x}(k-1))\|^2} \mathbf{s}_p(\mathbf{x}(k-1)) \quad ,$$

$$\mathbf{s}_p(\mathbf{x}(k)) = \mu\, \mathbf{s}_g(\mathbf{x}(k)) \quad ,$$

และ $$\nabla p(\mathbf{x}(k)) = \mu \left(\sum_{i \in J(\mathbf{x}(k))} \nabla g_i(\mathbf{x}(k))\right)$$

5. คำนวณจุดใหม่ $\mathbf{x}(k+1) = \mathbf{x}(k) + \lambda \mathbf{s}_p(\mathbf{x}(k))$,

   ให้ $k = k+1$ และไปขั้นตอน 2