ขั้นตอนวิธีในการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับ

ALGORITHMS FOR FINDING A FEASIBLE POINT OF CONSTRAINTS

เสกสรร สิริทรัพย์ทวี

SEKSON SIRISUBTAWEE

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาคณิตศาสตร์ประยุกต์

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2545

ISBN 974-648-892-9

ALGORITHMS FOR FINDING A FEASIBLE POINT OF CONSTRAINTS

SEKSON SIRISUBTAWEE

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE IN APPLIED MATHEMATICS
SCHOOL OF GRADUATE STUDIES
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2002

ISBN 974-648-892-9

หัวข้อวิทยานิพนธ์	**ขั้นตอนวิธีในการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับ**
นักศึกษา	**นายเสกสรร สิริทรัพย์ทวี**
รหัสประจำตัว	43065307
ปริญญา	**วิทยาศาสตรมหาบัณฑิต**
สาขาวิชา	**คณิตศาสตร์ประยุกต์**
พ.ศ.	2545
อาจารย์ผู้ควบคุมวิทยานิพนธ์	**ผศ.ไพรบูลย์ พันธรักษ์พงษ์**
อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม	**รศ.ดร.อำพล ธรรมเจริญ**

บทคัดย่อ

วิธีการแก้ปัญหาค่าเหมาะที่สุดแบบมีเงื่อนไขบังคับหลายวิธีต้องใช้จุดที่เป็นไปได้ในการเริ่มต้นแก้ปัญหา แต่ในกรณีทั่วไปเป็นการยากที่จะหาจุดเริ่มต้นดังกล่าว ในวิทยานิพนธ์นี้ได้เสนอขั้นตอนวิธีในการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับที่อยู่ในรูปแบบ

(1). $h_j(\mathbf{x}) = 0 \quad , \quad j = 1,\ldots,r$

(2). $g_i(\mathbf{x}) \le 0 \quad , \quad i = 1,\ldots,m$

(3). $g_i(\mathbf{x}) \le 0 \quad , \quad i = 1,\ldots,m \; ,$
$h_j(\mathbf{x}) = 0 \quad , \quad j = 1,\ldots,r$

โดยใช้วิธีฟังก์ชันทัณฑ์และฟังก์ชันกั้นเขต

Thesis Title	Algorithms for Finding a Feasible Point of Constraints
Student	Mr. Sekson Sirisubtawee
Student ID.	43065307
Degree	Master of Science
Programme	Applied Mathematics
Year	2002
Thesis Advisor	Assist.Prof.Praiboon Pantaragphong
Thesis Co-Advisor	Assoc.Prof.Dr.Ampon Dhamacharoen

ABSTRACT

Several methods for solving the constrained optimization problems require a feasible initial point. Generally, it is difficult to obtain that point. This thesis proposes the algorithms for finding a feasible point of constraints in the form :

(1). $h_j(\mathbf{x}) = 0 \quad , \ j = 1,\dots,r$

(2). $g_i(\mathbf{x}) \leq 0 \quad , \ i = 1,\dots,m$

(3). $g_i(\mathbf{x}) \leq 0 \quad , \ i = 1,\dots,m \ ,$
$h_j(\mathbf{x}) = 0 \quad , \ j = 1,\dots,r$

using the penalty and barrier function methods.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้เป็นอย่างดี ด้วยคำแนะนำและคำปรึกษาที่มีคุณค่าต่องานวิจัยนี้จาก ผศ.ไพรบูลย์ พันธรักษ์พงษ์ อาจารย์ผู้ควบคุมวิทยานิพนธ์ และ รศ.ดร.อำพล ธรรมเจริญ ภาควิชาคณิตศาสตร์ มหาวิทยาลัยบูรพา ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม ผู้วิจัยรู้สึกซาบซึ้งในความอนุเคราะห์จากท่านทั้งสอง และขอกราบขอบพระคุณเป็นอย่างสูง

ขอขอบคุณครอบครัวและญาติ ๆ ที่ได้ให้กำลังใจและความช่วยเหลือตลอดมา

ขอขอบคุณเพื่อน ๆ นักศึกษาทุกคนที่ช่วยเหลือให้คำแนะนำต่าง ๆ และยังให้กำลังใจต่อผู้วิจัยอย่างใกล้ชิดเสมอมา

สุดท้ายนี้ขอขอบคุณ มหาวิทยาลัยบูรพา ที่ได้ให้ทุนในการศึกษาในระดับปริญญาโทและทุนสนับสนุนการทำวิทยานิพนธ์

คุณค่าและประโยชน์อันพึงมีจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่ผู้มีพระคุณทุกท่าน

เสกสรร สิริทรัพย์ทวี

สารบัญ

หน้า

สารบัญ (ต่อ)

หน้า

สารบัญ (ต่อ)

หน้า

สารบัญตาราง

สารบัญรูป

รูปที่ หน้า

บทที่ 1
บทนำ

โดยทั่วไปแล้ว การแก้ปัญหาค่าเหมาะที่สุด (optimization problem) แบบมีเงื่อนไขบังคับ (constraints) เช่น ปัญหาในตัวอย่างนี้

$$
\begin{aligned}
\text{minimize} \quad & f(x) = x_1^2 + e^{(x_2 - x_3)} \\
\text{subject to} \quad & x_1^2 + x_2^2 + x_3^2 = 4 \\
& -x_1 + x_2 + x_3 = 2 \\
& -x_2^2 - x_3 \le 0 \\
& -e^{x_1} + x_3 \le 4
\end{aligned}
$$

สามารถทำได้หลายวิธี แต่วิธีทั้งหลายนั้นต้องการใช้จุดที่เป็นไปได้ (feasible point) ในการเริ่มต้นแก้ปัญหา ในบางครั้งเป็นการง่ายที่จะหาจุดเริ่มต้นที่เป็นไปได้ (initial feasible point) แต่ในกรณีทั่วไปไม่เป็นการง่ายที่จะหาจุดดังกล่าว เช่น ในปัญหาที่ยกมาข้างต้นนั้น การหาจุดเริ่มต้นให้สอดคล้องกับเงื่อนไขบังคับทั้งหมดเป็นไปได้ยาก ถึงแม้จะทำการทดลองสุ่มก็ต้องสูญเสียเวลาในการหา ซึ่งอาจทำให้ขบวนการแก้ปัญหาติดขัดและไม่สามารถดำเนินการต่อไปได้

1.1 ความเป็นมาและความสำคัญของวิทยานิพนธ์

ในการแก้ปัญหาค่าที่เหมาะที่สุดแบบมีเงื่อนไขบังคับนั้น ได้มีผู้วิจัยหลายท่านได้คิดค้นและพัฒนาขั้นตอนการแก้ปัญหาดังกล่าวซึ่งเป็นการแก้ปัญหาโดยตรง (เช่น เอกสาร[1]) เนื่องจากความหลากหลายของวิธีแก้ปัญหาโดยตรงที่แตกต่างกันและในแต่ละวิธีก็มีรูปแบบและข้อจำกัดเฉพาะของตัวเอง สิ่งนี้เป็นสาเหตุที่ทำให้ผู้วิจัยหันมาสนใจวิธีการแก้ปัญหาที่มีความเป็นทั่วไปมากกว่าวิธีการแก้ปัญหาโดยตรง โดยผู้วิจัยเลือกใช้วิธีฟังก์ชันทัณฑ์ (penalty function method) และวิธีฟังก์ชันกั้นเขต (barrier function method) ซึ่งได้มีนักคณิตศาสตร์หลายท่าน ทำการวิจัยต่อเนื่องกันมา (เช่น เอกสาร[2-7]) วิธีฟังก์ชันทัณฑ์และวิธีฟังก์ชันกั้นเขตเป็นวิธีที่ใช้ในการแปลงปัญหาค่าเหมาะที่สุดแบบมีเงื่อนไขบังคับไปเป็นปัญหาค่าเหมาะที่สุดแบบไม่มีเงื่อนไขบังคับหรือมีเงื่อนไขบังคับอย่างง่าย แล้วใช้วิธีการแก้ปัญหาค่าเหมาะที่สุดแบบไม่มีเงื่อนไขบังคับในการแก้ปัญหาต่อไป เมื่อผู้วิจัยได้ทำการศึกษาแนวคิดและขั้นตอนวิธีการแก้ปัญหาด้วยวิธีฟังก์ชันกั้นเขตพบว่าวิธีนี้ต้องการจุดเริ่มต้นในการแก้ปัญหาที่เป็นจุดภายใน (interior point) ของบริเวณที่เป็นไปได้ (feasible region) ซึ่งในบางครั้งปัญหาที่มีเงื่อนไขบังคับหลายเงื่อนไขและแต่ละเงื่อนไขบังคับ

มีหลายตัวแปร การหาจุดที่เป็นไปได้ซึ่งเป็นจุดภายในของบริเวณที่เป็นไปได้กลับเป็นเรื่องยากและสร้างปัญหาให้กับผู้วิจัยเพิ่มขึ้นไปอีก

นอกจากนี้ในการแก้ปัญหาค่าเหมาะที่สุดแบบมีเงื่อนไขบังคับโดยวิธีตรง เช่น วิธีในเอกสารอ้างอิง [1] ต้องการจุดเริ่มต้น x_0 ซึ่งเป็นจุดที่เป็นไปได้ (feasible point) ในการแก้ปัญหาแทบทั้งสิ้น ดังนั้น ปัญหาในการหาจุดเริ่มต้นที่เป็นจุดภายในของบริเวณที่เป็นไปได้และจุดเริ่มต้นที่เป็นจุดที่เป็นไปได้ (ในกรณีไม่มีจุดภายใน) จึงเป็นสิ่งจูงใจให้ผู้วิจัยหันกลับมาคิดและหากลวิธีในการแก้ปัญหาดังกล่าว เพราะถ้าสามารถแก้ปัญหาเบื้องต้นนี้ได้ จะทำให้เราดำเนินการแก้ปัญหาค่าเหมาะที่สุดต่อไปได้

1.2 วัตถุประสงค์ของงานวิจัย

เพื่อหาขั้นตอนวิธีสำหรับการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับ ซึ่งเป็นจุดภายในของบริเวณที่เป็นไปได้และจุดที่เป็นไปได้ (ในกรณีที่ไม่มีจุดภายใน)

1.3 ขอบเขตของงานวิจัย

เนื่องจากผู้วิจัยต้องการพัฒนาขั้นตอนวิธีการหาจุดที่เป็นไปได้ในเซตที่เป็นไปได้ (feasible set) ดังนั้นผู้วิจัยจึงทำการแบ่งขอบเขตของปัญหาออกเป็น 3 รูปแบบคือ

(1). การหาจุดที่เป็นไปได้ในเซตที่เป็นไปได้ในกรณีที่ไม่มีจุดภายใน โดยที่รูปแบบของเงื่อนไขบังคับเป็นแบบสมการ

$$h_j(x) = 0 \quad ; \; j = 1,\ldots,r \tag{1.1}$$

เราต้องการพัฒนาขั้นตอนวิธีในการหาจุดที่เป็นไปได้ $x^* \in \{x \in R^n \mid h_j(x) = 0 \quad ; \; j = 1,\ldots,r\}$

(2). การหาจุดที่เป็นไปได้ ซึ่งเป็นจุดภายในของบริเวณที่เป็นไปได้ โดยที่รูปแบบของเงื่อนไขบังคับเป็นแบบอสมการ

$$g_i(x) \le 0 \quad ; \; i = 1,\ldots,m \tag{1.2}$$

เราต้องการพัฒนาขั้นตอนวิธีในการหาจุดที่เป็นไปได้ $x^* \in \{x \in R^n \mid g_i(x) < 0 \quad ; \; i = 1,\ldots,m\}$

(3). การหาจุดที่เป็นไปได้ในกรณีเงื่อนไขบังคับแบบผสม โดยที่รูปแบบของเงื่อนไขบังคับเป็นดังนี้

$$g_i(x) \le 0 \quad , i = 1,\ldots,m$$
$$h_j(x) = 0 \quad ; j = 1,\ldots,r \tag{1.3}$$

เราต้องการพัฒนาขั้นตอนวิธีในการหาจุดที่เป็นไปได้ $x^* \in \{x \in R^n \mid g_i(x) < 0 \;\; ; \; i = 1,\ldots,m$ และ $h_j(x) = 0 \;\; ; \; j = 1,\ldots,r\}$

1.4 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

ได้ขั้นตอนวิธีที่พัฒนาขึ้น เพื่อใช้หาจุดภายในหรือจุดที่เป็นไปได้ของเงื่อนไขบังคับในรูปแบบต่าง ๆ ตาม (1.1) – (1.3) เพื่อใช้เป็นจุดเริ่มต้นซึ่งเป็นจุดที่เป็นไปได้ในการแก้ปัญหาค่าเหมาะที่สุดที่ต้องการจุดเริ่มต้นดังกล่าว

1.5 ขั้นตอนการวิจัย

(1). ค้นคว้าและศึกษาในเอกสารและข้อมูลที่เกี่ยวข้อง

(2). ทำการพัฒนาขั้นตอนวิธีในการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับตาม (1.1)–(1.3)

(3). เขียนโปรแกรมคอมพิวเตอร์เพื่อตรวจสอบความถูกต้องของขั้นตอนวิธีที่ได้

(4). เรียบเรียงและเขียนวิทยานิพนธ์

1.6 นิยามศัพท์ที่ใช้ในวิทยานิพนธ์

ในส่วนนี้จะให้นิยามของคำศัพท์ที่ใช้ในวิทยานิพนธ์โดยจัดเรียงนิยามศัพท์ไว้เป็นหมวดหมู่เดียวกันเพื่อง่ายต่อการอ้างถึง อนึ่งตลอดวิทยานิพนธ์นี้จะใช้ศัพท์บัญญัติและศัพท์ที่แปลขึ้นมาโดยมักไม่ระบุศัพท์ภาษาอังกฤษไว้ด้วย ดังนั้นหากต้องการทราบคำในภาษาอังกฤษของศัพท์บัญญัติใดให้ดูได้ที่หัวข้อนี้

(1). เซตและคำที่เกี่ยวข้อง

นิยาม 1.1 กำหนดจุด $x \in R^n$ และ $\varepsilon > 0$ และเรียกบอล (ball) $N_\varepsilon(x) = \{y \mid \|y - x\| < \varepsilon\}$ ว่าเป็น **ย่านใกล้เคียงขนาด** ε (ε - neighborhood) ของ x

นิยาม 1.2 ให้ $S \subset R^n$ และ $x \in S$ จะเรียก x ว่าเป็น **จุดภายใน** (interior point) ของ S ถ้ามีย่านใกล้เคียงขนาด ε ของ x บรรจุอยู่ใน S นั่นคือ ถ้ามี $\varepsilon > 0$ ซึ่ง $\|y - x\| \le \varepsilon$ จะได้ว่า $y \in S$, เซตของจุดดังกล่าวทั้งหมดจะเรียกว่า **เซตภายในของ** S (interior of S) และแทนด้วย $\operatorname{int} S$ หรือ S^0 , ถ้า $S = \operatorname{int} S$ จะเรียก S ว่า **เซตเปิด** (open set)

นิยาม 1.3 เรียก $S \subset R^n$ ว่า **เซตมีขอบเขต** (bounded set) ถ้า S ถูกบรรจุอยู่ภายในบอลที่มีรัศมีจำกัด

นิยาม 1.4 ให้ $S \subset R^n$, **ส่วนปิดคลุม** (closure) ของ S แทนด้วย $\mathrm{cl}\,S$ คือเซตของจุด x ที่มีสมบัติว่า สำหรับแต่ละ $\varepsilon > 0, S \cap N_\varepsilon(x) \neq \phi$, ถ้า $S = \mathrm{cl}\,S$ จะเรียก S ว่า **เซตปิด** (closed set)

นิยาม 1.5 ให้ $S \subset R^n$ เรียก x ว่า **จุดขอบ** (boundary point) ของ S ถ้าสำหรับแต่ละ $\varepsilon > 0$, $N_\varepsilon(x)$ จะบรรจุจุดที่อยู่ใน S และจุดที่ไม่อยู่ใน S ซึ่งเซตของจุดทั้งหมดเรียกว่า **ขอบของ** S (boundary of S) และแทนด้วย ∂S

นิยาม 1.6 เรียก $S \subset R^n$ ว่า **เซตกระชับ** (Compact set) ถ้า S เป็นเซตปิดและมีขอบเขต

สำหรับทุก ๆ ลำดับ $\{x_k\}$ ในเซตกระชับ S จะได้ว่ามีลำดับย่อยที่ลู่เข้าซึ่งลิมิตอยู่ใน S

(2). ผลเฉลยเหมาะที่สุด

ให้ $f : R^n \to R^1$ และพิจารณาปัญหา

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in S \end{array}$$

นิยาม 1.7 ถ้า $x^* \in S$ จะเรียก x^* ว่าเป็น **ผลเฉลยที่เป็นไปได้** (feasible solution) ของปัญหา

นิยาม 1.8 ถ้า $x^* \in S$ และ $f(x) \geq f(x^*)$ สำหรับทุก ๆ $x \in S$ จะเรียก x^* ว่า **ผลเฉลยเหมาะที่สุด** (optimal solution) หรือ **ผลเฉลยเหมาะที่สุดวงกว้าง** (global optimal solution) หรือ **ผลเฉลย** (solution) ของปัญหา

นิยาม 1.9 ถ้า $x^* \in S$ และถ้ามีย่านใกล้เคียง $N_\varepsilon(x^*)$ รอบจุด x^* ซึ่ง $f(x) \geq f(x^*)$ สำหรับทุก ๆ จุด $x \in S \cap N_\varepsilon(x^*)$ แล้วจะเรียก x^* ว่า **ผลเฉลยเหมาะที่สุดเฉพาะที่** (local optimal solution)

นิยาม 1.10 ถ้า $x^* \in S$ และถ้า $f(x) > f(x^*)$ สำหรับทุก ๆ $x \in S \cap N_\varepsilon(x^*), x \neq x^*$ สำหรับบาง $\varepsilon > 0$ แล้วจะเรียก x^* ว่า **ผลเฉลยเหมาะที่สุดเฉพาะที่โดยแท้** (strict local optimal solution)

(3). เวกเตอร์และเมตริกซ์

นิยาม 1.11 เซตของเวกเตอร์ $x_1,\ldots,x_k$ ใน R^n เป็น **อิสระเชิงเส้น** (linearly independent) ถ้า $\sum_{j=1}^{k}\lambda_j x_j = 0$ จะได้ว่า $\lambda_j = 0$ สำหรับ $j = 1,\ldots,k$

นิยาม 1.12 นอร์มของเวกเตอร์ x ใน R^n แทนด้วย $\|x\|$ ซึ่ง $\|x\| = (x^T x)^{1/2} = (\sum_{j=1}^{n} x_j^2)^{1/2}$ เมื่อ $x = (x_1,\ldots,x_n)^T$

นิยาม 1.13 เมตริกซ์จัตุรัส A เป็น **เมตริกซ์สมมาตร** (symmetric matrix) ถ้า $A = A^T$

นิยาม 1.14 ให้ A เป็นเมตริกซ์ขนาด $m \times n$ **ค่าลำดับชั้น** (rank) ของ A คือจำนวนที่มากที่สุดของแถวที่เป็นอิสระเชิงเส้น หรือ จำนวนที่มากที่สุดของหลักที่เป็นอิสระเชิงเส้นของเมตริกซ์ A , ถ้าค่าลำดับชั้นของ A เท่ากับ minimum $\{m,n\}$ แล้ว จะเรียก A ว่ามีค่าลำดับชั้นเต็ม (full rank)

นิยาม 1.15 ให้ A เป็นเมตริกซ์ขนาด $n \times n$, สเกลาร์ λ และเวกเตอร์ $x(\neq 0)$ ที่สอดคล้องกับสมการ $Ax = \lambda x$ จะเรียกว่า **ค่าเฉพาะ** (eigenvalue) และ **เวกเตอร์เฉพาะ** (eigenvector) ของ A ตามลำดับ

การคำนวณหาค่าเฉพาะของ A ทำได้โดยการแก้สมการ $\det[A - \lambda I] = 0$ ซึ่งเป็นสมการพหุนามของ λ ถ้า A เป็นเมตริกซ์สมมาตร แล้วค่าเฉพาะทั้งหมดของ A เป็นจำนวนจริงและสามารถเขียน $P^T AP = D$ เมื่อ D เป็นเมตริกซ์ทแยงมุม (diagonal matrix) ซึ่งมีค่าเฉพาะของ A ได้แก่ $\lambda_1,\ldots,\lambda_n$ อยู่ในแนวทแยงมุมหลักและ P เป็นเมตริกซ์ขนาด $n \times n$ ซึ่งเป็นเมตริกซ์ตั้งฉาก (orthogonal matrix) ดังนั้น $PP^T = I$ หรือ $P^T P = I$ นั่นคือ $P^{-1} = P^T$ (คุณสมบัติเหล่านี้ดูได้จากเอกสารอ้างอิง [8])

นิยาม 1.16 ให้ A เป็นเมตริกซ์สมมาตรขนาด $n \times n$

ถ้า $x^T Ax > 0$ สำหรับทุก $x(\neq 0)$ ใน R^n แล้วเรียก A ว่าเป็น **เมตริกซ์ที่เป็นบวกแน่นอน** (positive definite matrix)

ถ้า $x^T Ax \geq 0$ สำหรับทุก ๆ x ใน R^n แล้วเรียก A ว่าเป็น **เมตริกซ์ที่เป็นบวกกึ่งแน่นอน** (positive semidefinite matrix)

ถ้า $x^T Ax < 0$ สำหรับทุก $x(\neq 0)$ ใน R^n แล้วเรียก A ว่าเป็น **เมตริกซ์ที่เป็นลบแน่นอน** (negative definite matrix)

ถ้า $x^T Ax \le 0$ สำหรับทุก ๆ x ใน R^n แล้วเรียก A ว่าเป็น **เมตริกซ์ที่เป็นลบกึ่งแน่นอน** (negative semidefinite matrix)

ถ้า $x^T Ax > 0$ สำหรับบาง $x \in R^n$ และ $x^T Ax < 0$ สำหรับ x อื่น ๆ ใน R^n แล้วเรียก A ว่าเป็น **เมตริกซ์ที่ไม่แน่นอน** (indefinite matrix)

สำหรับการบอกชนิดของเมตริกซ์ A จากค่าเฉพาะของ A นั้น สามารถดูได้จากทฤษฎีบทที่ 2.6 ในบทที่ 2

(4). เซตนูนและความนูนของฟังก์ชัน

นิยาม 1.17 เซต S ใน R^n เป็น **เซตนูน** (convex set) ถ้าสำหรับแต่ละ $x_1, x_2 \in S$ แล้ว ส่วนของเส้นตรง $\lambda x_1 + (1-\lambda)x_2$ สำหรับ $\lambda \in [0,1]$ อยู่ใน S นั่นคือ ถ้า $x_1, x_2 \in S$ แล้ว $\lambda x_1 + (1-\lambda)x_2 \in S$ สำหรับ $\lambda \in [0,1]$

รูปที่ 1.1 แสดงเซตนูนและเซตที่ไม่เป็นเซตนูน

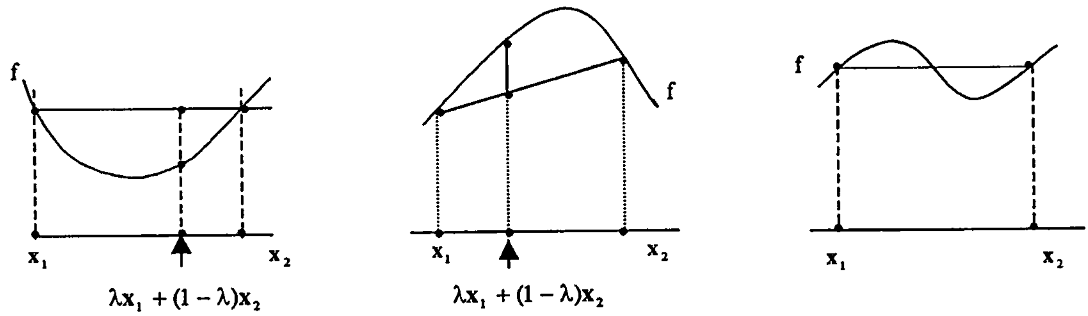
นิยาม 1.18 ให้ $S(\neq \phi)$ เป็นเซตนูนใน R^n , ฟังก์ชัน $f: S \to R^1$ เป็น **ฟังก์ชันนูน** บน S ถ้า

$$f[\lambda x_1 + (1-\lambda)x_2] \le \lambda f(x_1) + (1-\lambda)f(x_2)$$

สำหรับทุก ๆ $x_1, x_2 \in S$ และสำหรับทุก ๆ $\lambda \in [0,1]$, เรียกฟังก์ชัน f ว่าเป็น **ฟังก์ชันนูนโดยแท้** (strictly convex function) บน S ถ้า

$$f[\lambda x_1 + (1-\lambda)x_2] < \lambda f(x_1) + (1-\lambda)f(x_2)$$

สำหรับแต่ละจุด x_1 , x_2 ที่แตกต่างกันซึ่งอยู่ใน S และสำหรับแต่ละ $\lambda \in (0,1)$, เรียกฟังก์ชัน f ว่าเป็น **ฟังก์ชันเว้า** (concave function) ถ้า $-f$ เป็นฟังก์ชันนูน , เรียกฟังก์ชัน f ว่าเป็น **ฟังก์ชันเว้าโดยแท้** (strictly concave function) ถ้า $-f$ เป็นฟังก์ชันนูนโดยแท้



(ก). ฟังก์ชันนูน (ข). ฟังก์ชันเว้า (ค). ไม่ใช่ทั้งฟังก์ชันนูนและฟังก์ชันเว้า

รูปที่ 1.2 แสดงตัวอย่างของฟังก์ชันนูนและฟังก์ชันเว้า

นิยาม 1.19 ให้ $S(\neq \phi)$ เป็นเซตนูนใน R^n

ฟังก์ชัน $f: S \to R^1$ เป็น **ฟังก์ชันกึ่งนูน** (quasiconvex function) บน S ถ้าสำหรับแต่ละ $x_1, x_2 \in S$ แล้วอสมการต่อไปนี้เป็นจริง

$f[\lambda x_1 + (1-\lambda)x_2] \leq \text{maximum } \{f(x_1), f(x_2)\}$ สำหรับแต่ละ $\lambda \in (0,1)$

ฟังก์ชัน f เป็น **ฟังก์ชันกึ่งนูนโดยแท้** (strictly quasiconvex function) บน S ถ้า $f[\lambda x_1 + (1-\lambda)x_2] < \text{maximum } \{f(x_1), f(x_2)\}$ สำหรับ $f(x_1) \neq f(x_2)$

ฟังก์ชัน f เป็น **ฟังก์ชันกึ่งนูนแบบเข้ม** (strongly quasiconvex function) บน S ถ้า $f[\lambda x_1 + (1-\lambda)x_2] < \text{maximum } \{f(x_1), f(x_2)\}$ สำหรับ $x_1 \neq x_2$

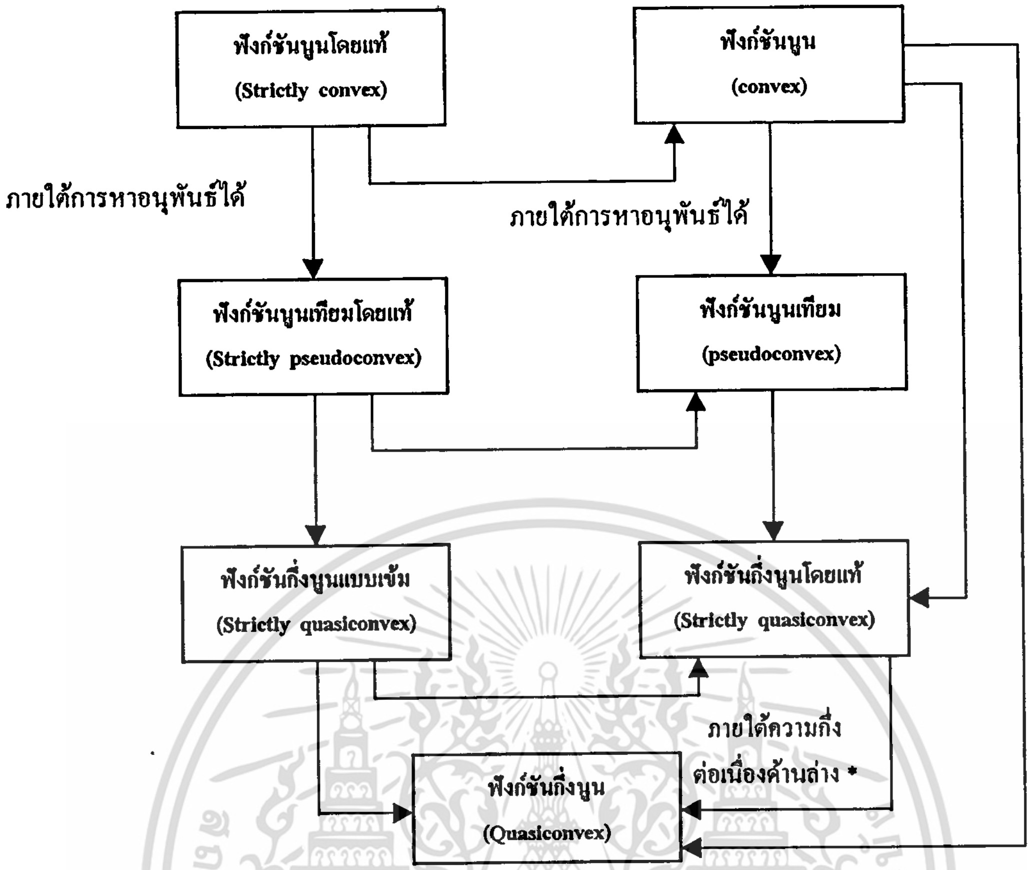
นิยาม 1.20 ให้ $S(\neq \phi)$ เป็นเซตนูนเปิดใน R^n

ฟังก์ชัน $f: S \to R^1$ เป็น **ฟังก์ชันนูนเทียม** (pseudoconvex function) ถ้าสำหรับแต่ละ $x_1, x_2 \in S$ ซึ่ง $\nabla f(x_1)^T(x_2 - x_1) \geq 0$ จะได้ว่า $f(x_2) \geq f(x_1)$

ฟังก์ชัน f เป็น **ฟังก์ชันนูนเทียมโดยแท้** (strictly pseudoconvex function) บน S ถ้าเมื่อ x_1 และ x_2 เป็นจุดที่แตกต่างกันใน S ซึ่ง $\nabla f(x_1)^T(x_2 - x_1) \geq 0$ จะได้ว่า $f(x_2) > f(x_1)$

(ก). เป็นทั้งฟังก์ชันกึ่งนูนและฟังก์ชันนูนเทียม (ข). ฟังก์ชันกึ่งนูนแต่ไม่เป็นฟังก์ชันนูนเทียม (ค). ไม่เป็นทั้งฟังก์ชันกึ่งนูนและฟังก์ชันนูนเทียม

รูปที่ 1.3 แสดงความเป็นกึ่งนูนและนูนเทียมของฟังก์ชัน



รูปที่ 1.4 แสดงความสัมพันธ์ระหว่างความนูนชนิดต่าง ๆ ของฟังก์ชัน

นิยาม 1.21 ให้ S เป็นเซตไม่ว่างใน R^n, ฟังก์ชัน $f:S \rightarrow R^1$ เป็น ฟังก์ชันกึ่งต่อเนื่องด้านล่าง* (lower semicontinuous) ที่ $x^* \in S$ ถ้าสำหรับแต่ละ $\varepsilon > 0$, มี $\delta > 0$ ซึ่ง $x \in S$ และ $\|x - x^*\| < \delta$ จะได้ว่า $f(x) - f(x^*) > -\varepsilon$

(5). เกรเดียนต์ของฟังก์ชันและเมตริกซ์เฮซเซียน

นิยาม 1.22 ให้ f เป็นฟังก์ชันค่าจริงของ n ตัวแปร นั่นคือ $f(x) = f(x_1,\ldots,x_n)$, เวกเตอร์ของอนุพันธ์อันดับหนึ่งของ f เรียกว่า เกรเดียนต์ของ f (gradient of f) ซึ่งแทนด้วย

$$\nabla f(x) = \begin{bmatrix} \dfrac{\partial f(x)}{\partial x_1} \\ \dfrac{\partial f(x)}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

เมตริกซ์ของอนุพันธ์อันดับสองของ f เรียกว่า เมตริกซ์เฮซเซียน (Hessian matrix) และแทนด้วย $\nabla^2 f(x)$ หรือ $Hf(x)$ โดยที่

$$Hf(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

สำหรับฟังก์ชัน f ซึ่งมีอนุพันธ์อันดับสองต่อเนื่อง จะได้ว่า

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}$$

ดังนั้นในกรณีนี้ $Hf(x)$ เป็นเมตริกซ์สมมาตร

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะให้ทฤษฎีพื้นฐานเกี่ยวกับฟังก์ชันนูน (convex function), ฟังก์ชันนูนโดยแท้ (strictly convex function) และเงื่อนไขความเหมาะที่สุด (optimality condition) ในส่วนกลางของบทจะกล่าวถึงวิธีการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ (unconstrained minimization problem) ซึ่งจะนำไปใช้ในการแก้ปัญหาหลังการแปลงปัญหาเริ่มต้นที่เป็นปัญหาค่าต่ำสุดแบบมีเงื่อนไขบังคับ (constrained minimization problem) โดยวิธีฟังก์ชันกั้นเขต (barrier function method) และวิธีฟังก์ชันทัณฑ์ (penalty function method) สำหรับในส่วนสุดท้ายของบทจะกล่าวถึงงานวิจัยที่เกี่ยวข้องและมีความสำคัญต่องานวิจัยของผู้วิจัย อนึ่งนิยามศัพท์ที่กล่าวถึงในบทนี้จะอยู่ในบทที่ 1 หัวข้อ 1.7 จึงไม่ขอกล่าวไว้ในส่วนนี้อีก

2.1 ทฤษฎีพื้นฐานและเงื่อนไขความเหมาะที่สุด

ในหัวข้อนี้จะอ้างถึงทฤษฎีพื้นฐานที่มีความสำคัญต่องานวิจัยนี้ โดยจะยกข้อความของทฤษฎีมากล่าวอ้าง โดยไม่ทำการพิสูจน์ (ซึ่งดูได้จากเอกสารอ้างอิงต่าง ๆ ที่กำกับไว้)

2.1.1 ทฤษฎีพื้นฐาน

ทฤษฎีต่อไปนี้จะให้ผลลัพธ์ที่น่าสนใจ สำหรับจุดต่ำสุดของฟังก์ชันนูน

ทฤษฎีบทที่ 2.1 จุดต่ำสุดเฉพาะที่ใด ๆ ของฟังก์ชันนูน $f(x)$ ซึ่งนิยามบนเซตย่อยนูน S ของ R^n จะเป็นจุดต่ำสุดวงกว้างด้วย

จุดต่ำสุดเฉพาะที่ใด ๆ ของฟังก์ชันนูนโดยแท้ $f(x)$ ซึ่งนิยามบนเซตย่อยนูน S ใน R^n จะเป็นจุดต่ำสุดวงกว้างโดยแท้เพียงจุดเดียวของ $f(x)$ บน S

พิสูจน์ ดูเอกสารอ้างอิง [8]

ทฤษฎีบทที่ 2.2 ที่จะกล่าวต่อไปนี้จะให้วิธีการสร้างฟังก์ชันเพื่อให้เป็นฟังก์ชันนูนหรือฟังก์ชันนูนโดยแท้

ทฤษฎีบทที่ 2.2

(1). ถ้า $f_1(x),\dots,f_k(x)$ เป็นฟังก์ชันนูนบนเซตนูน S ใน R^n แล้ว $f(x) = f_1(x) +$

$f_2(x)+\ldots+f_k(x)$ เป็นฟังก์ชันนูน

ถ้ามีฟังก์ชัน $f_i(x)$ อย่างน้อยหนึ่งฟังก์ชันเป็นฟังก์ชันนูนโดยแท้บน S แล้ว ฟังก์ชันผลรวม $f(x)$ เป็นฟังก์ชันนูนโดยแท้

(2). ถ้า $f(x)$ เป็นฟังก์ชันนูน (ฟังก์ชันนูนโดยแท้) บนเซตนูน S ใน R^n และถ้า α เป็นจำนวนบวก แล้ว $\alpha f(x)$ เป็นฟังก์ชันนูน (ฟังก์ชันนูนโดยแท้) บน S

(3). ถ้า $f(x)$ เป็นฟังก์ชันนูน (ฟังก์ชันนูนโดยแท้) ซึ่งนิยามบนเซตนูน S ใน R^n และถ้า $g(y)$ เป็นฟังก์ชันเพิ่ม (ฟังก์ชันเพิ่มโดยแท้) ซึ่งเป็นฟังก์ชันนูนที่นิยามบนเรนจ์ของ $f(x)$ ใน R^1 แล้ว $g[f(x)]$ เป็นฟังก์ชันนูน (ฟังก์ชันนูนโดยแท้) บน S

พิสูจน์ ดูเอกสารอ้างอิง [8]

ทฤษฎีบท 2 ทฤษฎีที่จะกล่าวต่อไปนี้ จะให้ความสัมพันธ์ระหว่างฟังก์ชัน $f(x)$ และเมตริกซ์เฮซเซียน $Hf(x)$ ของฟังก์ชัน

ทฤษฎีบทที่ 2.3 ให้ $S(\neq \phi)$ เป็นเซตนูนเปิดใน R^n และให้ $f: S \to R^1$ เป็นฟังก์ชันที่หาอนุพันธ์ได้สองครั้งบน S แล้ว f เป็นฟังก์ชันนูนก็ต่อเมื่อเมตริกซ์เฮซเซียนเป็นบวกกึ่งแน่นอนที่ทุก ๆ จุดใน S

พิสูจน์ ดูเอกสารอ้างอิง [1]

ทฤษฎีบทที่ 2.4 ให้ $S(\neq \phi)$ เป็นเซตนูนเปิดใน R^n และให้ $f: S \to R^1$ เป็นฟังก์ชันที่หาอนุพันธ์ได้สองครั้งบน S ถ้าเมตริกซ์เฮซเซียนเป็นบวกแน่นอนที่ทุก ๆ จุดใน S แล้ว f เป็นฟังก์ชันนูนโดยแท้ ในทางกลับ ถ้า f เป็นฟังก์ชันนูนโดยแท้แล้วเมตริกซ์เฮซเซียนเป็นบวกกึ่งแน่นอนที่ทุก ๆ จุดใน S อย่างไรก็ตาม ถ้า f เป็นฟังก์ชันนูนโดยแท้และเป็นฟังก์ชันกำลังสองแล้ว เมตริกซ์เฮซเซียนของ f จะเป็นบวกแน่นอน

พิสูจน์ ดูเอกสารอ้างอิง [1]

การพิจารณาเมตริกซ์สมมาตร A ว่าเป็นเมตริกซ์ที่เป็นบวกแน่นอนหรือไม่นั้นอาจพิจารณาจากทฤษฎีข้างล่างนี้ แต่ก่อนอื่นต้องทำการนิยามสัญลักษณ์ที่เกี่ยวข้องก่อน ให้ Δ_k เป็นดีเทอร์มิแนนต์ของเมตริกซ์ย่อยขนาด $k \times k$ มุมซ้ายมือบนของ A สำหรับ $1 \le k \le n$ เมื่อ

A เป็นเมตริกซ์สมมาตรขนาด $n \times n$ เราเรียก Δ_k ว่าไมเนอร์มุขสำคัญที่ k (k-th principal minor) ของ A นั่นคือ

$$A = \begin{array}{c} \begin{array}{ccccc} \Delta_1 & \Delta_2 & \Delta_3 & & \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{12} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nn} \end{bmatrix} \end{array}$$

$$\Delta_1 = a_{11} \quad , \quad \Delta_2 = \det\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \ ,\ldots,\ \Delta_n = \det A$$

ทฤษฎีบทที่ 2.5 ถ้า A เป็นเมตริกซ์สมมาตรขนาด $n \times n$ และถ้า Δ_k เป็นไมเนอร์มุขสำคัญที่ k ของ A สำหรับ $1 \le k \le n$ แล้ว

(1). A เป็นเมตริกซ์ที่เป็นบวกแน่นอน ก็ต่อเมื่อ $\Delta_k > 0$ สำหรับ $k = 1,2,\ldots,n$

(2). A เป็นเมตริกซ์ที่เป็นลบแน่นอน ก็ต่อเมื่อ $(-1)^k \Delta_k > 0$ สำหรับ $k = 1,2,\ldots,n$

(3). ถ้า $\Delta_1 > 0, \Delta_2 > 0, \ldots, \Delta_{n-1} > 0, \Delta_n = 0$ แล้ว A เป็นเมตริกซ์ที่เป็นบวกกึ่งแน่นอน

(4). ถ้า $(-1)^k \Delta_k > 0$ สำหรับ $k = 1,\ldots,n-1$ และ $\Delta_n = 0$ แล้ว A เป็นเมตริกซ์ที่เป็นลบกึ่งแน่นอน

พิสูจน์ ดูเอกสารอ้างอิง [8]

นอกจากการพิจารณา Δ_k แล้ว เราอาจพิจารณาชนิดของเมตริกซ์ A ได้จากค่าเฉพาะของเมตริกซ์ A ดังทฤษฎีต่อไปนี้

ทฤษฎีบทที่ 2.6 ถ้า A เป็นเมตริกซ์สมมาตรแล้ว

(1). เมตริกซ์ A เป็นบวกแน่นอน (เป็นลบแน่นอน) ก็ต่อเมื่อ ค่าเฉพาะทุกตัวของ A เป็นบวก (เป็นลบ)

(2). เมตริกซ์ A เป็นบวกกึ่งแน่นอน (เป็นลบกึ่งแน่นอน) ก็ต่อเมื่อ ค่าเฉพาะทุกตัวของ A ไม่เป็นลบ (ไม่เป็นบวก)

(3). เมตริกซ์ A เป็นเมตริกซ์ไม่กำหนด ก็ต่อเมื่อ A มีค่าเฉพาะบางค่าที่เป็นบวกและบางค่าที่เป็นลบ

พิสูจน์ ดูเอกสารอ้างอิง [8]

สมมติว่า $f(x)$ เป็นฟังก์ชันค่าจริงที่นิยามบนเซตย่อย D ของ R^n แล้วเราจะเรียกจุด $x^* \in D$ ว่าเป็นจุดวิกฤต (critical point) สำหรับ $f(x)$ ถ้ามีอนุพันธ์ย่อยอันดับหนึ่งของ $f(x)$ ที่ x^* และ $\dfrac{\partial f(x^*)}{\partial x_i} = 0$, $i = 1,2,\ldots,n$

ทฤษฎีบทที่ 2.7 สมมติว่า x^* เป็นจุดวิกฤตของฟังก์ชัน $f(x)$ ซึ่งมีอนุพันธ์ย่อยอันดับหนึ่งและสองต่อเนื่องบน R^n และ $Hf(x)$ เป็นเมตริกซ์เฮซเซียนของ $f(x)$ แล้ว x^* เป็น

(1). จุดต่ำสุดวงกว้างสำหรับ $f(x)$ ถ้า $Hf(x)$ เป็นบวกกึ่งแน่นอนบน R^n

(2). จุดต่ำสุดวงกว้างโดยแท้สำหรับ $f(x)$ ถ้า $Hf(x)$ เป็นบวกแน่นอนบน R^n

(3). จุดต่ำสุดเฉพาะที่โดยแท้ของ $f(x)$ ถ้า $Hf(x^*)$ เป็นบวกแน่นอน

พิสูจน์ ดูเอกสารอ้างอิง [8]

2.1.2 เงื่อนไขความเหมาะที่สุด

2.1.2.1 ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ

บทแทรกและทฤษฎีต่อไปนี้จะให้เงื่อนไขความเหมาะที่สุดที่จำเป็นสำหรับปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ

บทแทรกที่ 2.8 สมมติว่า $f: R^n \rightarrow R^1$ สามารถหาอนุพันธ์ได้ที่ x^* ถ้า x^* เป็นจุดต่ำสุดเฉพาะที่แล้ว $\nabla f(x^*) = 0$

พิสูจน์ ดูเอกสารอ้างอิง [1]

เราเรียกบทแทรกที่ 2.8 ว่าเป็นเงื่อนไขอันดับหนึ่ง และทฤษฎีที่ 2.9 ที่จะกล่าวต่อไปนี้เป็นเงื่อนไขอันดับสอง

ทฤษฎีบทที่ 2.9 สมมติว่า $f:R^n \to R^1$ สามารถหาอนุพันธ์ได้สองครั้งที่จุด x^* ถ้า x^* เป็นจุดต่ำสุดเฉพาะที่ แล้ว $\nabla f(x^*) = \mathbf{0}$ และ $Hf(x^*)$ เป็นบวกกึ่งแน่นอน

พิสูจน์ ดูเอกสารอ้างอิง [1]

ทฤษฎีบทที่ 2.10 และ 2.11 ที่จะให้ต่อไปนี้จะเป็นเงื่อนไขที่เพียงพอสำหรับความเหมาะที่สุด

ทฤษฎีบทที่ 2.10 สมมติว่า $f:R^n \to R^1$ สามารถหาอนุพันธ์ได้สองครั้งที่ x^* ถ้า $\nabla f(x^*) = \mathbf{0}$ และ $Hf(x^*)$ เป็นบวกแน่นอน แล้ว x^* เป็นจุดต่ำสุดเฉพาะที่โดยแท้ (สังเกตว่าทฤษฎีนี้ก็คือทฤษฎีบทที่ 2.7 (3). นั่นเอง)

พิสูจน์ ดูเอกสารอ้างอิง [1]

ทฤษฎีบทที่ 2.11 ให้ $f:R^n \to R^1$ เป็นฟังก์ชันนูนเทียมที่ x^* แล้ว x^* เป็นจุดต่ำสุดวงกว้างก็ต่อเมื่อ $\nabla f(x^*) = \mathbf{0}$

พิสูจน์ ดูเอกสารอ้างอิง [1]

2.1.2.2 ปัญหาค่าต่ำสุดแบบมีเงื่อนไขบังคับแบบอสมการและสมการ

ทฤษฎีบทที่ 2.12 (เงื่อนไขที่จำเป็นแบบคารุช-คุห์น-ทักเกอร์)

ให้ $X(\neq \phi)$ เป็นเซตเปิดใน R^n และให้ $f:R^n \to R^1$, $g_i:R^n \to R^1$ สำหรับ $i = 1,\ldots,m$ และ $h_j:R^n \to R^1$ สำหรับ $j = 1,\ldots,r$

พิจารณาปัญหา (P)

$$\left.\begin{array}{lll} \text{minimize} & f(\mathbf{x}) & \\ \text{subject to} & g_i(\mathbf{x}) \leq 0 & ;\quad i = 1,\ldots,m \\ & h_j(\mathbf{x}) = 0 & ;\quad j = 1,\ldots,r \\ & \mathbf{x} \in X & \end{array}\right\} \qquad (2.1)$$

ให้ $\mathbf{x}^*$ เป็นผลเฉลยที่เป็นไปได้และให้ $I = \{i \mid g_i(\mathbf{x}^*) = 0\}$ สมมติว่า f และ g_i สำหรับ $i \in I$ สามารถหาอนุพันธ์ได้ที่ $\mathbf{x}^*$ สมมติว่าแต่ละ g_i สำหรับ $i \in I$ เป็นฟังก์ชันต่อเนื่องที่ $\mathbf{x}^*$ และแต่ละ h_j สำหรับ $j = 1,\ldots,r$ เป็นฟังก์ชันที่มีอนุพันธ์ต่อเนื่องที่ $\mathbf{x}^*$ สมมติเพิ่มเติมว่า $\nabla g_i(\mathbf{x}^*)$ สำหรับ $i \in I$ และ $\nabla h_j(\mathbf{x}^*)$ สำหรับ $j = 1,\ldots,r$ เป็นอิสระเชิงเส้น

ถ้า $\mathbf{x}^*$ แก้ปัญหา (P) ซึ่งเป็นจุดต่ำสุดเฉพาะที่ แล้วจะมีสเกลาร์ u_i สำหรับ $i \in I$ และ v_j สำหรับ $j = 1,\ldots,r$ เพียงหนึ่งเดียวที่ทำให้

$$\left.\begin{aligned} \nabla f(\mathbf{x}^*) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{r} v_j \nabla h_j(\mathbf{x}^*) &= 0 \\ u_i \ge 0 \quad \text{สำหรับ } i \in I \end{aligned}\right\} \quad (2.2)$$

ถ้าเพิ่มข้อสมมติให้กับข้อสมมติข้างบนว่า แต่ละ g_i สำหรับ $i \notin I$ เป็นฟังก์ชันที่หาอนุพันธ์ได้ที่ $\mathbf{x}^*$ แล้วเงื่อนไขแบบคารุช-คุห์น-ทักเกอร์ เขียนในรูปแบบที่สมมูลกันได้ดังนี้

$$\left.\begin{aligned} \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{r} v_j \nabla h_j(\mathbf{x}^*) &= \mathbf{0} \\ u_i g_i(\mathbf{x}^*) &= 0 \quad ; i = 1,\ldots,m \\ u_i &\ge 0 \quad ; i = 1,\ldots,m \end{aligned}\right\} \quad (2.3)$$

พิสูจน์ ดูเอกสารอ้างอิง [1]

ทฤษฎีบทที่ 2.13 (เงื่อนไขที่เพียงพอแบบคารุช-คุห์น-ทักเกอร์)

ให้ $X(\neq \phi)$ เป็นเซตเปิดใน R^n และให้ $f: R^n \to R^1$, $g_i: R^n \to R^1$ สำหรับ $i = 1,\ldots,m$ และ $h_j: R^n \to R^1$ สำหรับ $j = 1,\ldots,r$ พิจารณาปัญหา (P) ใน (2.1) ให้ $\mathbf{x}^*$ เป็นผลเฉลยที่เป็นไปได้และให้ $I = \{i \mid g_i(\mathbf{x}^*) = 0\}$

สมมติว่าเงื่อนไขแบบคารุช-คุห์น-ทักเกอร์ เป็นจริงที่ $\mathbf{x}^*$ นั่นคือ มีสเกลาร์ $u_i \ge 0$ สำหรับ $i \in I$ และ v_j^* สำหรับ $j = 1,\ldots,r$ ที่ทำให้

$$\nabla f(\mathbf{x}^*) + \sum_{i \in I} u_i^* \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^{r} v_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0}$$

ให้ $J = \{j \mid v_j^* > 0\}$ และ $K = \{j \mid v_j^* < 0\}$ สมมติเพิ่มเติมว่า f เป็นฟังก์ชันนูนเทียมที่ x^* , g_i เป็นฟังก์ชันกึ่งนูนที่ x^* สำหรับ $i \in I$, h_j เป็นฟังก์ชันกึ่งนูนที่ x^* สำหรับ $j \in J$ และ h_j เป็นฟังก์ชันกึ่งเว้าที่ x^* สำหรับ $j \in K$ แล้ว x^* จะเป็นผลเฉลยเหมาะที่สุดวงกว้างของปัญหา (P) โดยเฉพาะถ้าข้อสมมติเกี่ยวกับความนูน (convexity) บนฟังก์ชันจุดประสงค์และเงื่อนไขบังคับถูกจำกัดบนโดเมน $N_\varepsilon(x^*)$ สำหรับบาง $\varepsilon > 0$ แล้ว x^* จะเป็นจุดต่ำสุดเฉพาะที่ของ (P)

พิสูจน์ ดูเอกสารอ้างอิง [1]

ในบางครั้งเงื่อนไขที่จำเป็นของความเหมาะที่สุดที่จุดเป็นไปได้ x^* (2.3) อาจเขียนได้ใหม่เมื่อกำหนดตัวคูณ $\lambda_i = -u_i \le 0$ และ $\mu_j = -v_j$ เงื่อนไข (2.3) จะกลายเป็น

$$\left.\begin{aligned} \nabla f(x^*) - \sum_{i=1}^{m}\lambda_i \nabla g_i(x^*) - \sum_{j=1}^{r}\mu_j \nabla h_j(x^*) &= 0 \\ \lambda_i g_i(x^*) = 0 \quad &; i = 1,\ldots,m \\ \lambda_i \le 0 \quad &; i = 1,\ldots,m \end{aligned}\right\} \tag{2.4}$$

ซึ่งเป็นอีกรูปแบบหนึ่งที่นิยมเขียนกัน

2.1.2.3 เงื่อนไขที่จำเป็นและเพียงพออันดับสองสำหรับปัญหาแบบมีเงื่อนไขบังคับ

เรานิยามฟังก์ชันลากรองจ์แบบจำกัด (restricted Lagrangian function) ได้ดังนี้

$$L(x) = \Phi(x, u^*, v^*) = f(x) + \sum_{i \in I} u_i^* g_i(x) + \sum_{j=1}^{r} v_j^* h_j(x) \tag{2.5}$$

เมื่อ u^*, v^* เป็นตัวคูณลากรองจ์ (Lagrangian multiplier) ที่สมนัยกับเงื่อนไขบังคับแบบอสมการและสมการตามลำดับ โดยที่ u^*, v^* ได้ที่จุด x^* ซึ่งเป็นจุดคารุช-คุห์น-ทักเกอร์ ต่อไปจะให้ทฤษฎีสำหรับเงื่อนไขที่เพียงพอและจำเป็นอันดับสองโดยจะใช้ $L(x)$

ทฤษฎีบทที่ 2.14 (เงื่อนไขที่พอเพียงอันดับสองแบบคารุช-คุห์น-ทักเกอร์)

พิจารณาปัญหา (P) ใน (2.1) ซึ่งฟังก์ชันจุดประสงค์และเงื่อนไขบังคับเป็นฟังก์ชันที่สามารถหาอนุพันธ์ได้สองครั้งและ $X(\neq \phi)$ เป็นเซตเปิดใน R^n ให้ $\mathbf{x}^*$ เป็นจุดคารุช-คุห์น-ทักเกอร์ สำหรับปัญหา P ด้วยตัวคูณลากรองจ์ $\mathbf{u}^*$ และ $\mathbf{v}^*$ ให้ $I = \{i \mid g_i(\mathbf{x}^*) = 0\}$ และ $I^+ = \{i \in I \mid u_i^* > 0\}$ และ $I^0 = \{i \in I \mid u_i^* = 0\}$ นิยามฟังก์ชันลากรองจ์แบบจำกัด $L(\mathbf{x})$ ใน (2.5) และแทนเมตริกซ์เฮซเซียนที่ $\mathbf{x}^*$ โดย

$$\nabla^2 L(\mathbf{x}^*) = \nabla^2 f(\mathbf{x}^*) + \sum_{i\in I} u_i^* \nabla^2 g_i(\mathbf{x}^*) + \sum_{j=1}^{r} v_j^* \nabla^2 h_j(\mathbf{x}^*) \qquad (2.6)$$

ซึ่ง $\nabla^2 f(\mathbf{x}^*), \nabla^2 g_i(\mathbf{x}^*)$ สำหรับ $i \in I$ และ $\nabla^2 h_j(\mathbf{x}^*)$ สำหรับ $j=1,\ldots,r$ เป็นเมตริกซ์เฮซเซียนของ f, g_i $(i \in I)$ และ h_j $(j=1,\ldots,r)$ ตามลำดับเมื่อทั้งหมดคำนวณที่ $\mathbf{x}^*$ นิยามกรวย $C = \{\mathbf{d} \neq \mathbf{0} \mid \nabla g_i(\mathbf{x}^*)^T \mathbf{d} = 0$ สำหรับ $i \in I^+$, $\nabla g_i(\mathbf{x}^*)^T \mathbf{d} \leq 0$ สำหรับ $i \in I^0$ และ $\nabla h_j(\mathbf{x}^*)^T \mathbf{d} = 0$ สำหรับ $j = 1,\ldots,r\}$ แล้ว ถ้า $\mathbf{d}^T \nabla^2 L(\mathbf{x}^*) \mathbf{d} > 0$ สำหรับทุก ๆ $\mathbf{d} \in C$ จะได้ว่า $\mathbf{x}^*$ เป็นจุดต่ำสุดเฉพาะที่โดยแท้สำหรับ (P)

พิสูจน์ ดูเอกสารอ้างอิง [1]

ทฤษฎีบทที่ 2.15 (เงื่อนไขที่จำเป็นอันดับสองแบบคารุช-คุห์น-ทักเกอร์)

พิจารณาปัญหา (P) ใน (2.1) ซึ่งฟังก์ชันจุดประสงค์และเงื่อนไขบังคับเป็นฟังก์ชันที่สามารถหาอนุพันธ์ได้สองครั้งและ $X(\neq \phi)$ เป็นเซตเปิดใน R^n ให้ $\mathbf{x}^*$ เป็นจุดต่ำสุดเฉพาะที่สำหรับปัญหา (P) และแทน $I = \{i \mid g_i(\mathbf{x}^*) = 0\}$ นิยามฟังก์ชันลากรองจ์แบบจำกัด $L(\mathbf{x})$ ใน (2.5) และแทนเมตริกซ์เฮซเซียนที่ $\mathbf{x}^*$ โดย (2.6) ซึ่ง $\nabla^2 f(\mathbf{x}^*), \nabla^2 g_i(\mathbf{x}^*)$ สำหรับ $i \in I$ และ $\nabla^2 h_j(\mathbf{x}^*)$ สำหรับ $j=1,\ldots,r$ เป็นเมตริกซ์เฮซเซียนของ f, g_i $(i \in I)$ และ h_j $(j = 1,\ldots,r)$ ตามลำดับเมื่อทั้งหมดคำนวณที่ $\mathbf{x}^*$ สมมติว่า $\nabla g_i(\mathbf{x}^*)$ $(i \in I)$ และ $\nabla h_j(\mathbf{x}^*)$ $(j = 1,\ldots, r)$ เป็นอิสระเชิงเส้น แล้ว $\mathbf{x}^*$ จะเป็นจุดคารุช-คุห์น-ทักเกอร์ และ $\mathbf{d}^T \nabla^2 L(\mathbf{x}^*) \mathbf{d} \geq 0$ สำหรับทุก ๆ $\mathbf{d} \in C' = \{\mathbf{d} \neq \mathbf{0} \mid \nabla g_i(\mathbf{x}^*)^T \mathbf{d} \leq 0$ $(i \in I)$ และ $\nabla h_j(\mathbf{x}^*)^T \mathbf{d} = 0$ $(j = 1,\ldots,r)\}$

พิสูจน์ ดูเอกสารอ้างอิง [1]

2.2 แนวคิดและขั้นตอนวิธีในการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ

ในหัวข้อนี้จะได้กล่าวถึงแนวคิดในการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ ซึ่งในปัจจุบันมีวิธีการแก้ปัญหาหลายวิธี เช่น วิธีของนิวตัน (Newton's Method) , วิธีลดเร็วที่สุด (Steepest Descent Method) และวิธีแนวนิวตัน (Quasi-Newton Method) ซึ่งเป็นวิธีที่ผู้วิจัยเลือกใช้ สำหรับแนวคิดในการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับมีหลักการคือ พยายามหาทิศทางลด (descent direction) ซึ่งกระทำในขั้นตอนการหาทิศทาง (search direction) ต่อจากนั้นจะหาความยาวในการเดินทางตามทิศทางลดที่หามาได้ ซึ่งกระทำในขั้นตอนการค้นหาแนวเส้น (line search) เมื่อได้ทิศทางและความยาวที่จะเดินทางแล้ว เราจะได้จุดใหม่ $\mathbf{x}_{k+1}$ จากจุดเก่า $\mathbf{x}_k$ ซึ่งดีกว่าเดิมในแง่ของค่าฟังก์ชันจุดประสงค์ที่ลดลง ที่กล่าวมานี้คือ จากจุด $\mathbf{x}_k$ เราหาทิศทางลด $\mathbf{p}_k$ และความยาวในการเดินทาง α_k และกำหนดให้

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \quad ; \quad \alpha_k > 0 \tag{2.7}$$

กระทำดังนี้ต่อไปเรื่อย ๆ จนกว่าจะถึงจุดที่เหมาะที่สุด เพื่อให้กระบวนการข้างต้นเป็นไปอย่างมีประสิทธิภาพ จึงได้มีการกำหนดเกณฑ์สำหรับการสร้างวิธีที่ดีในการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับดังจะกล่าวในหัวข้อต่อไปนี้

2.2.1 เกณฑ์ที่ใช้ในการสร้างวิธีการที่ดีในการแก้ปัญหาแบบไม่มีเงื่อนไขบังคับ

กำหนด $f(x)$ เป็นฟังก์ชันที่มีอนุพันธ์ย่อยอันดับหนึ่งที่ต่อเนื่องบน R^n

เกณฑ์ 1 $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$ เมื่อ $\nabla f(\mathbf{x}_k) \neq \mathbf{0}$

เกณฑ์นี้เป็นการกำหนดว่าวิธีการที่ดี ควรเป็นวิธีการลด (descent method) ในค่าของฟังก์ชันจุดประสงค์ $f(x)$

เกณฑ์ 2 $\mathbf{p}_k^T \nabla f(\mathbf{x}_k) < 0$

พิจารณา $F_k(\alpha) = f(\mathbf{x}_k + \alpha \mathbf{p}_k)$

ซึ่ง $F_k'(\alpha) = \mathbf{p}_k^T \nabla f(\mathbf{x}_k + \alpha \mathbf{p}_k)$

และ $F_k'(0) = \mathbf{p}_k^T \nabla f(\mathbf{x}_k) < 0$ (ถ้าเกณฑ์ 2 เป็นจริง)

ดังนั้น สำหรับ $\alpha > 0$ ซึ่งเล็กเพียงพอ จะได้ว่า

$$f(\mathbf{x}_k + \alpha \mathbf{p}_k) = F_k(\alpha) < F_k(0) = f(\mathbf{x}_k)$$

เป็นจริง เพราะความต่อเนื่องของอนุพันธ์ย่อยอันดับหนึ่งของ f ดังนั้นถ้าเกณฑ์ 2 เป็นจริงและสำหรับ $\alpha_k > 0$ ซึ่งเล็กเพียงพอแล้ว $f(x_{k+1}) = f(x_k + \alpha_k p_k) < f(x_k)$ นั่นคือจะสอดคล้องกับเกณฑ์ 1 ด้วย

ถ้าเราดำเนินต่อไป โดยให้ α_k เป็นค่าบวกเล็กมาก ๆ แล้วก็คงเป็นไปไม่ได้ที่ x_k จะเคลื่อนไปหาจุดต่ำสุด x^* ได้อย่างรวดเร็วเพราะขนาดของการเดินทางนั้นเล็กเกินไป เพื่อหลีกเลี่ยงการเดินทางที่สั้นเกินไปจึงต้องกำหนดเกณฑ์ 3 ขึ้นมา

เกณฑ์ 3 มี η ซึ่ง $0 < \eta < 1$ ที่ทำให้

$$p_k^T \nabla f(x_{k+1}) \geq \eta p_k^T \nabla f(x_k) \tag{2.8}$$

ต่อไปนี้จะแสดงว่าเกณฑ์ 2 และ 3 ช่วยป้องกันการเลือก α_k ที่เล็กเกินไปได้ สมมติว่า η ถูกเลือกให้คงที่ในช่วง $0 < \eta < 1$ และ α_k ถูกเลือกให้สอดคล้องกับเกณฑ์ 3 แล้วจะได้ว่า

$$p_k^T \nabla f(x_k + \alpha_k p_k) \geq \eta p_k^T \nabla f(x_k) > p_k^T \nabla f(x_k)$$

เพราะว่า $\eta < 1$ และ $p_k^T \nabla f(x_k) < 0$ (จากเกณฑ์ 2) ดังนั้น

$$p_k^T \nabla f(x_k + \alpha_k p_k) - p_k^T \nabla f(x_k) \geq (\eta - 1) p_k^T \nabla f(x_k) > 0 \tag{2.9}$$

จะเห็นว่าถ้าเลือกให้ $\alpha_k \to 0$ ข้างซ้ายของอสมการ (2.9) จะเข้าใกล้ 0 ในขณะที่ข้างขวาของ (2.9) ยังคงเป็นค่าคงที่ $(\eta - 1) p_k^T \nabla f(x_k) > 0$ ซึ่งเป็นไปไม่ได้ ดังนั้นจึงไม่สามารถเลือกให้ α_k เล็กเกินไปได้ เราเรียกเกณฑ์ 3 นี้ว่าเงื่อนไขของวูล์ฟ (Wolfe Condition)

สำหรับกรณีที่ α_k ใหญ่เกินไป แต่ทำให้ค่าของฟังก์ชันจุดประสงค์ลดลงไปได้พอ ๆ กับ α_k ที่มีขนาดเล็กกว่า ดังนั้นปัญหาของ α_k ที่ใหญ่เกินไปจึงควรมีเกณฑ์ในการควบคุมไม่ให้เกิดเหตุการณ์ดังกล่าวซึ่งตรงกับเกณฑ์ 4

เกณฑ์ 4 มี μ ซึ่ง $0 < \mu < \eta < 1$ ที่ทำให้

$$f(x_{k+1}) \leq f(x_k) + \mu \alpha_k p_k^T \nabla f(x_k) \tag{2.10}$$

ในทางปฏิบัติมักเลือกให้ $\mu \in (0, \frac{1}{2})$ และ $\eta \in (\mu, 1)$ เพื่อให้เข้าใจในความต้องการของเกณฑ์ 4 ให้พิจารณาอสมการ

$$\frac{f(x_k) - f(x_{k+1})}{\alpha_k} \geq \mu[-p_k^T \nabla f(x_k)]$$

เทอม $[-p_k^T \nabla f(x_k)]$ ทางขวาเป็นจำนวนบวกโดยเกณฑ์ 2

ถ้าเราให้
$$M = \frac{-\mu p_k^T \nabla f(x_k)}{\|p_k\|}$$

แล้วเกณฑ์ 4 จะเป็นอสมการต่อไปนี้

$$\frac{f(x_k) - f(x_{k+1})}{\|x_k - x_{k+1}\|} = \frac{f(x_k) - f(x_{k+1})}{\alpha_k \|p_k\|} \geq M$$

นั่นคือ

$$f(x_k) - f(x_{k+1}) \geq M\|x_k - x_{k+1}\| > 0 \tag{2.11}$$

ดังนั้นถ้าระยะในการเดินทางจาก x_k ไป x_{k+1} ใหญ่ (นั่นคือ $\|x_k - x_{k+1}\|$ ใหญ่) แล้วการลดลงในค่าฟังก์ชันจุดประสงค์ $f(x_k) - f(x_{k+1})$ จะเป็นสัดส่วนที่ใหญ่ตามไปด้วยและจะสังเกตเห็นจาก (2.11) ว่าเกณฑ์ 1 เป็นจริงโดยอัตโนมัติเมื่อเกณฑ์ 2 และ 4 เป็นจริง เราเรียกเกณฑ์ 4 นี้ว่าเงื่อนไขของอาร์มิโจ (Armijo Condition) รูปที่ 2.1 แสดงการเลือก α_k เพื่อให้สอดคล้องกับเกณฑ์ 3 และ 4

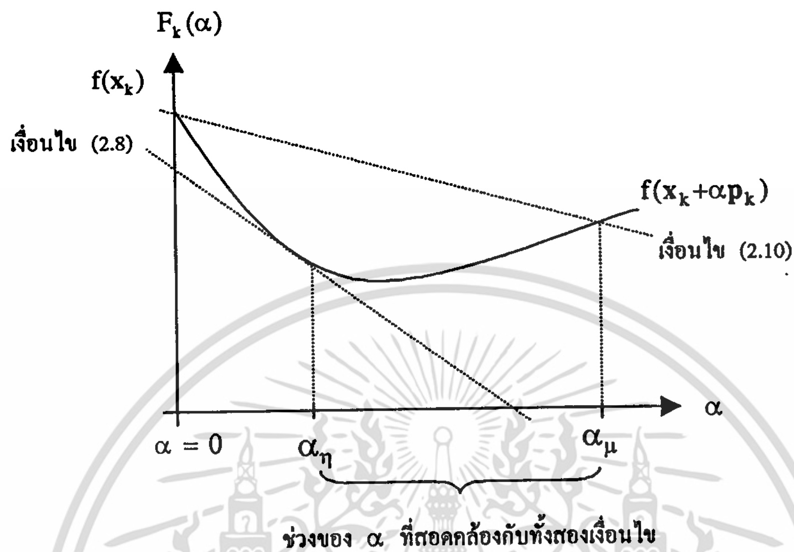
ทฤษฎีบทที่ 2.16 สมมติว่า $f(x)$ มีอนุพันธ์ย่อยอันดับหนึ่งต่อเนื่องและเป็นฟังก์ชันที่มีขอบเขตล่างบน R^n ให้ μ, η ถูกเลือกให้คงที่ โดยที่ $0 < \mu < \eta < 1$ ถ้า p_k และ x_k เป็นเวกเตอร์ใน R^n ที่สอดคล้องกับเงื่อนไข $p_k^T \nabla f(x_k) < 0$ แล้วจะมีจำนวนจริง α_η และ α_μ ซึ่ง $0 \leq \alpha_\eta < \alpha_\mu$ และ

(1). เกณฑ์ 4 เป็นจริงสำหรับทุก ๆ $\alpha_k \in (0, \alpha_\mu)$

(2). เกณฑ์ 3 เป็นจริงสำหรับทุก ๆ $\alpha_k \in (\alpha_\eta, \alpha_\mu)$

นอกจากนี้ จะได้ว่าทั้งเกณฑ์ 3 และ 4 เป็นจริงพร้อมกัน สำหรับทุก ๆ $\alpha_k \in (\alpha_\eta, \alpha_\mu)$

พิสูจน์ ดูเอกสารอ้างอิง [8]



รูปที่ 2.1 แสดงเงื่อนไขการค้นหาแนวเส้นทั้งสองเงื่อนไข

2.2.2 ทิศทางลด

ในการแก้ปัญหา

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in R^n \end{array}$$

ให้จุด x_k เป็นจุดเริ่มต้นในขั้นตอนที่ k เราจะหาทิศทาง p_k ที่ทำให้

$$f(x_k + \alpha_k p_k) < f(x_k) \qquad \text{เมื่อ } \alpha_k > 0 \tag{2.12}$$

อสมการ (2.12) จะเป็นไปได้ถ้าทิศทาง p_k เป็นทิศทางลด นั่นคือ

$$p_k^T \nabla f(x_k) < 0 \tag{2.13}$$

จะเห็นว่าอสมการ (2.13) ก็คือเกณฑ์ 2 นั่นเอง ทฤษฎีต่อไปนี้จะแสดงว่า (2.13) เป็นเงื่อนไขที่ทำให้ p_k เป็นทิศทางลด

ทฤษฎีบทที่ 2.17 สมมติว่า $f : R^n \to R^1$ เป็นฟังก์ชันที่หาอนุพันธ์ได้ที่จุด $\mathbf{x}_k$ ถ้ามีเวกเตอร์ $\mathbf{p}_k$ ที่ทำให้ $\mathbf{p}_k^T \nabla f(\mathbf{x}_k) < 0$ แล้วจะมี $\delta > 0$ ที่ทำให้ $f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) < f(\mathbf{x}_k)$ สำหรับทุก ๆ $\alpha_k \in (0, \delta)$ และจะเรียก $\mathbf{p}_k$ ว่าทิศทางลดของ f ที่จุด $\mathbf{x}_k$

พิสูจน์ ดูเอกสารอ้างอิง [1]

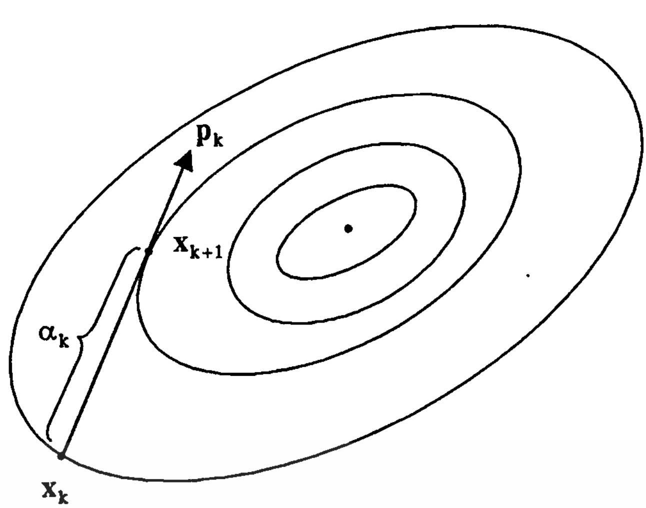
เมื่อวิธีการที่ใช้ดำเนินการหาจุด $\mathbf{x}^*$ ซึ่งเป็นจุดต่ำสุด ต้องการทิศทางลด $\mathbf{p}_k$ ดังนั้นวิธีการที่ดีจึงต้องสามารถหาทิศทางลด $\mathbf{p}_k$ ได้และไม่ควรซับซ้อน ในหัวข้อ 2.2.4 จะให้วิธีกำเนิด $\mathbf{p}_k$ ซึ่งเป็นทิศทางลด

2.2.3 การค้นหาแนวเส้น

เมื่อได้ $\mathbf{p}_k$ ซึ่งเป็นทิศทางลดแล้วจะหาจุด $\mathbf{x}_{k+1}$ จาก (2.7) ได้นั้นจะต้องหา α_k ในขั้นตอนที่เรียกว่า การค้นหาแนวเส้น เรานิยามฟังก์ชัน $F_k(\alpha) = f(\mathbf{x}_k + \alpha \mathbf{p}_k)$ ซึ่งเป็นฟังก์ชันของ α การค้นหาแนวเส้นก็คือ การหาความยาวในการเดินทาง α_k จากจุด $\mathbf{x}_k$ ไปยังจุด $\mathbf{x}_{k+1}$ ตามทิศทางลด $\mathbf{p}_k$ ซึ่งจุด $\mathbf{x}_{k+1}$ จะเป็นจุดที่ค่าฟังก์ชันจุดประสงค์ต่ำที่สุดในทิศทาง $\mathbf{p}_k$ ในขณะเดียวกันถ้าเราใช้ระยะทางในการเดินทาง α มากกว่าระยะทางที่ดีที่สุด α_k มากเกินไป อาจทำให้ $f(\mathbf{x}_{k+1}) > f(\mathbf{x}_k)$ ได้

การหาระยะทางในการเดินทางที่ดีที่สุด α_k คือ การค้นหาแนวเส้นเหมาะที่สุดแท้จริง (exact line search) นั่นคือ $F'_k(\alpha_k) = 0$ แต่การค้นหาแนวเส้นเหมาะที่สุดแท้จริงเป็นการกระทำที่สิ้นเปลืองมากในทางปฏิบัติ ดังนั้นจึงเป็นไปได้ที่จะใช้เกณฑ์ 3 และ 4 ในการค้นหาแนวเส้นทำให้ได้ α_k ที่น่าพอใจและไม่สิ้นเปลืองแรงงาน การหา α_k ที่ยอมรับได้คือการหา α_k ที่สอดคล้องกับเกณฑ์ 3 และ 4 คือไม่เล็กและไม่ใหญ่เกินไป วิธีการที่อาจใช้หา α_k ได้ เช่น วิธีภาคตัดทอง (golden section method) แต่วิธีนี้จะค้นหา α_k ภายในช่วงปิด $[a_0, b_0]$ โดยที่ปัญหาที่เกิดขึ้นคือการหาจุดปลาย b_0 ซึ่ง $F'_k(b_0) > 0$ (ปกติให้ $a_0 = 0$ เพราะ $F'_k(0) = \mathbf{p}_k^T \nabla f(\mathbf{x}_k) < 0$ เนื่องจาก $\mathbf{p}_k$ เป็นทิศทางลด ดังนั้นที่จุด $a_0 = 0$ จึงเป็นจุดที่ความชันของ $F_k(\alpha)$ ลดลง) เพราะลักษณะของฟังก์ชัน $F(\alpha)$ อาจเป็นฟังก์ชันที่ยากแก่การหาจุด b_0 ดังนั้นผู้วิจัยจึงเลือกใช้วิธีถอยหลังกลับ (backtracking method) โดยวิธีนี้สามารถหา α_k ให้สอดคล้องกับเกณฑ์ 4 ได้ โดยเริ่มให้ $\alpha = 1$ แล้วทำการลดค่า α ลงจนไปสอดคล้องกับเกณฑ์ 4 นั่นคือเราตรวจสอบในทางที่เป็นเท็จของเกณฑ์ 4 ว่า

$$f(\mathbf{x}_k + \alpha \mathbf{p}_k) > f(\mathbf{x}_k) + \mu \alpha \mathbf{p}_k^T \nabla f(\mathbf{x}_k)$$



รูปที่ 2.2 แสดงการค้นหาแนวเส้น α_k ซึ่งเป็นระยะที่ดีที่สุดจากจุด x_k ไปยังจุด x_{k+1} บนเส้นโค้งระดับ (level curve)

เป็นจริงหรือไม่ ถ้าเป็นจริงให้เราลด α ลงเป็น $\gamma\alpha$ ซึ่งเลือก $\gamma < 1$ ที่เหมาะสมโดยวิธีการถอยหลังกลับนี้จะหลีกเลี่ยงการได้มาซึ่งค่า α_k ที่น้อยเกินไป ผู้วิจัยจะได้ให้ขั้นตอนวิธีการถอยหลังกลับที่ใช้ในงานวิจัยนี้ต่อไปในบทที่ 3

2.2.4 ขั้นตอนวิธีแนวนิวตัน

ในหัวข้อนี้จะให้แนวคิดและขั้นตอนวิธีในการแก้ปัญหาค่าต่ำแบบไม่มีเงื่อนไขบังคับ หลังจากได้แนวคิดของการหาทิศทางลด p_k และการหาระยะในการเดินทาง α_k โดยวิธีที่กล่าวถึงนี้เรียกว่าวิธีแนวนิวตันซึ่งวิธีนี้อยู่บนพื้นฐานของการประมาณเมตริกซ์เฮซเซียน $Hf(x_k) = \nabla^2 f(x_k)$ โดยเมตริกซ์ B_k อื่น ๆ ซึ่งสามารถหาได้ไม่ยาก ดังนั้นวิธีแนวนิวตันจึงแตกต่างกันที่สูตรในการประมาณ $Hf(x_k)$

เมื่อพิจารณาวิธีนิวตัน โดยที่ทิศทาง p_k สามารถกำหนดได้โดย

$$p_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k) \tag{2.14}$$

ถ้า p_k เป็นทิศทางลดที่จุด x_k แล้ว จะได้ว่า

$$p_k^T \nabla f(x_k) = \nabla f(x_k)^T p_k = -\nabla f(x_k)^T [\nabla^2 f(x_k)]^{-1} \nabla f(x_k) < 0$$

เพราะ p_k เป็นทิศทางลดที่จุด x_k ดังนั้น $p_k^T \nabla f(x_k) < 0$ จากอสมการข้างบน จะได้ว่า

$$\nabla f(\mathbf{x}_k)^T[\nabla^2 f(\mathbf{x}_k)]^{-1}\nabla f(\mathbf{x}_k) > 0 \qquad (2.15)$$

เงื่อนไข (2.15) จะเป็นจริงถ้า $[\nabla^2 f(\mathbf{x}_k)]^{-1}$ เป็นเมตริกซ์ที่เป็นบวกแน่นอน นั่นคือ $\nabla^2 f(\mathbf{x}_k)$ ต้องเป็นเมตริกซ์ที่เป็นบวกแน่นอน (การพิสูจน์จะได้แสดงให้เห็นในบทที่ 3) ดังนั้นถ้าสูตรในการประมาณ $\nabla^2 f(\mathbf{x}_k)$ โดย B_k แล้วยังคงทำให้ B_k เป็นเมตริกซ์ที่เป็นบวกแน่นอนแล้ว ทิศทาง $\mathbf{p}_k = -[B_k]^{-1}\nabla f(\mathbf{x}_k)$ จะเป็นทิศทางลด ในทางปฏิบัติเราจะไม่คำนวณหา $\mathbf{p}_k$ จาก $[B_k]^{-1}$ เพราะการหาเมตริกซ์ผกผันเป็นเรื่องที่ยุ่งยากในทางปฏิบัติ ดังนั้นจะหา $\mathbf{p}_k$ จากการแก้ระบบสมการเชิงเส้น

$$B_k\mathbf{p} = -\nabla f(\mathbf{x}_k) \qquad (2.16)$$

ข้อดีของการใช้วิธีแนวนิวตัน คือ

1. เมตริกซ์ประมาณ B_k สามารถหาได้โดยใช้เพียงแต่ข้อมูลอนุพันธ์อันดับหนึ่ง

2. การค้นหาทิศทาง $\mathbf{p}_k$ สามารถคำนวณได้โดยใช้การกระทำ $O(n^2)$ (ในขณะที่วิธีนิวตันใช้การกระทำ $O(n^3)$)

วิธีแนวนิวตันนั้นมีที่มาจากวิธีเซแคนต์ (secant method) โดยที่วิธีเซแคนต์ประมาณ $f''(\mathbf{x}_k)$ จากสูตร

$$f''(\mathbf{x}_k) \approx \frac{f'(\mathbf{x}_k) - f'(\mathbf{x}_{k-1})}{\mathbf{x}_k - \mathbf{x}_{k-1}} \qquad (2.17)$$

(2.17) เป็นกรณี 1 มิติ แต่ในกรณีหลายมิติจะไม่สามารถใช้ได้เพราะติดขัดกับการหารของเวกเตอร์ซึ่งไม่นิยาม ดังนั้นในกรณีหลายมิติ (2.17) จะกลายเป็น

$$\nabla^2 f(\mathbf{x}_k)(\mathbf{x}_k - \mathbf{x}_{k-1}) \approx \nabla f(\mathbf{x}_k) - \nabla f(\mathbf{x}_{k-1})$$

ดังนั้นถ้าเราจะประมาณ $\nabla^2 f(\mathbf{x}_k)$ โดย B_k แล้ว B_k จะต้องสอดคล้องกับเงื่อนไขต่อไปนี้

$$B_k(\mathbf{x}_k - \mathbf{x}_{k-1}) = \nabla f(\mathbf{x}_k) - \nabla f(\mathbf{x}_{k-1}) \qquad (2.18)$$

เรียกเงื่อนไข (2.18) ว่า เงื่อนไขเซแคนต์ (secant condition)

เมื่อกำหนดเวกเตอร์

$$\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k \tag{2.19}$$

และ

$$\mathbf{y}_k = \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k) \tag{2.20}$$

จากเงื่อนไขเซแคนต์ ได้ว่า $B_k \mathbf{s}_{k-1} = \mathbf{y}_{k-1}$ เพื่อความสะดวกเราจะใช้

$$B_{k+1}\mathbf{s}_k = \mathbf{y}_k \tag{2.21}$$

สำหรับปัญหา n – มิติ เงื่อนไข (2.21) แทนระบบสมการ n สมการซึ่งต้องการหาเมตริกซ์ B_{k+1} ที่มีสมาชิก n^2 ตัว ดังนั้น B_{k+1} อาจมีได้มากกว่า 1 เมตริกซ์ที่สอดคล้องกับ (2.21) วิธีแนวนิวตันจะกำเนิด B_{k+1} จาก B_k ดังนั้นรูปแบบการกำเนิดคือ $B_{k+1} = B_k + [$ บางสิ่งบางอย่าง $]$

เมื่อเราใช้ B_k ในการประมาณ $\nabla^2 f(\mathbf{x}_k)$ ดังนั้น B_k ควรมีคุณสมบัติเช่นเดียวกับเมตริกซ์เฮซเซียน เพราะว่า $\nabla^2 f(\mathbf{x}_k)$ เป็นเมตริกซ์สมมาตร ดังนั้นสูตรในการกำเนิด B_k จาก B_{k-1} (เป็นเมตริกซ์สมมาตร) ควรจะได้ว่า B_k เป็นเมตริกซ์สมมาตรด้วย สูตรต่อไปนี้เป็นรูปแบบหนึ่งในการหาเมตริกซ์ B_{k+1} จาก B_k (เป็นเมตริกซ์สมมาตร) โดยที่ B_{k+1} ยังคงเป็นเมตริกซ์สมมาตรต่อไป ซึ่งการพิสูจน์สามารถดูได้จากเอกสารอ้างอิง [9]

$$B_{k+1} = B_k + \frac{(\mathbf{y}_k - B_k\mathbf{s}_k)(\mathbf{y}_k - B_k\mathbf{s}_k)^T}{(\mathbf{y}_k - B_k\mathbf{s}_k)^T\mathbf{s}_k} \tag{2.22}$$

เรียก (2.22) ว่า สูตรเวียนบังเกิดค่าลำดับชั้น-1 ที่สมมาตร (symmetric rank-one update formula) และสามารถพิสูจน์ได้ว่า (2.22) สอดคล้องกับเงื่อนไขเซแคนต์ (2.21) (ดูจากเอกสารอ้างอิง[9])

นอกจากคุณสมบัติความสมมาตรที่เกิดกับ $\nabla^2 f(\mathbf{x}_k)$ แล้ว เมื่อพิจารณาเมตริกซ์เฮซเซียนที่ผลเฉลย $\mathbf{x}^*$ จะได้ว่า $\nabla^2 f(\mathbf{x}^*)$ มักเป็นเมตริกซ์ที่เป็นบวกแน่นอน ดังนั้นเมตริกซ์ B_k ที่ใช้ประมาณเมตริกซ์เฮซเซียนควรเป็นเมตริกซ์ที่เป็นบวกแน่นอนด้วย ซึ่งจะรับประกันว่าทิศทาง $\mathbf{p}_k$ ที่ได้เป็นทิศทางลดซึ่งสอดคล้องกับเกณฑ์ 2 ในหัวข้อ 2.2.1 เนื่องจากไม่มีสูตรเวียนบังเกิดค่าลำดับชั้น-1 แบบใด ๆ ที่จะรักษาทั้งความสมมาตรและความเป็นบวกแน่นอนของเมตริกซ์ที่ใช้ประมาณ ดังนั้นจึงได้มีการกำเนิดสูตรเวียนบังเกิดค่าลำดับชั้น-2 (rank-two update formula) ซึ่งเป็นสูตรที่สามารถทำสิ่งเหล่านี้ได้ ซึ่งสูตรที่ใช้กันอย่างแพร่หลายและเป็นวิธีที่มีประสิทธิภาพมากคือ สูตร

เวียนบังเกิด BFGS (BFBS update formula) ซึ่งมีผู้พัฒนาต่อเนื่องกันมาจำนวน 4 คน ได้แก่ Broyden, Fletcher, Goldfarb และ Shanno สูตร BFGS เป็นดังนี้

$$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \tag{2.23}$$

s_k และ y_k เป็นดัง (2.19) และ (2.20) ตามลำดับ สูตร (2.23) สอดคล้องกับเงื่อนไขเซแคนต์ (2.21) สามารถดูพิสูจน์ได้ในบทที่ 3 บทตั้งต่อไปนี้จะให้เงื่อนไขที่จำเป็นและเพียงพอสำหรับการกำเนิด B_{k+1} จากสูตร BFGS เพื่อให้ B_{k+1} เป็นเมตริกซ์ที่เป็นบวกแน่นอน

บทตั้งที่ 2.18 ให้ B_k เป็นเมตริกซ์ที่เป็นบวกแน่นอนและสมมาตร สมมติว่า B_{k+1} ได้จาก B_k โดยใช้สูตร BFGS (2.23) แล้ว B_{k+1} เป็นเมตริกซ์ที่เป็นบวกแน่นอนก็ต่อเมื่อ $y_k^T s_k > 0$

พิสูจน์ ดูเอกสารอ้างอิง [9]

สำหรับทฤษฎีการลู่เข้าหาผลเฉลยเหมาะที่สุด โดยวิธีแนวนิวตันเมื่อใช้สูตร BFGS จะให้ไว้ในหัวข้อ 2.4 งานวิจัยที่เกี่ยวข้อง

สำหรับขั้นตอนวิธีแนวนิวตัน (เมื่อใช้สูตร BFGS) เพื่อแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับเป็นดังนี้

ขั้นเริ่มต้น เลือกจุดเริ่มต้น x_0 และเมตริกซ์เริ่มต้นที่ใช้ประมาณเมตริกซ์เฮซเซียน B_0 ซึ่งเลือกให้เป็นเมตริกซ์สมมาตรและเป็นบวกแน่นอน โดยปกติมักเลือก $B_0 = I$

ขั้นหลัก สำหรับ $k = 0,1,\ldots,n$

1. ถ้า x_k เป็นผลเฉลยเหมาะที่สุดให้หยุด
   ถ้าเป็นอย่างอื่นให้แก้ระบบสมการเชิงเส้น $B_k p = -\nabla f(x_k)$ เพื่อหา p_k
2. ใช้การค้นหาแนวเส้นหา $x_{k+1} = x_k + \alpha_k p_k$
3. คำนวณ $s_k = x_{k+1} - x_k$ และ $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$
   คำนวณ $B_{k+1} = B_k - \dfrac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \dfrac{y_k y_k^T}{y_k^T s_k}$

ในการใช้สูตร BFGS เพื่อหาทิศทาง p_k จะต้องผ่านขั้นตอนการแก้ระบบสมการ (2.16) ดังนั้นในงานวิจัยนี้จะใช้สูตรตัวผกผันของ B_k ให้แทนด้วย H_k ดังนั้น

$$p_k = -[B_k]^{-1}\nabla f(x_k) = -H_k \nabla f(x_k)$$

ซึ่งเป็นการคูณกันของเมตริกซ์กับเวกเตอร์ จึงเป็นการลดขั้นตอนการทำงานในการแก้สมการ โดยที่สูตรในการกำเนิด H_{k+1} จะได้จาก $H_k = B_k^{-1}$ และสามารถพิสูจน์ได้ว่า $H_{k+1} = B_{k+1}^{-1}$ ซึ่งตัวสูตรในการกำเนิดและการพิสูจน์จะอยู่ในบทที่ 3

2.3 แนวคิดและวิธีการแปลงปัญหาค่าต่ำสุดแบบมีเงื่อนไขบังคับไปเป็นปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ

ปัญหากำหนดการเชิงคณิตศาสตร์ (mathematical programming problem) คือ การหาเวกเตอร์ $x^* = (x_1^*,\ldots,x_n^*)^T$ ซึ่งแก้ปัญหา (A) ดังต่อไปนี้

$$\left.\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \le 0 & ; i = 1,\ldots,m \\ & h_j(x) = 0 & ; j = 1,\ldots,r \end{array}\right\} \quad (2.24)$$

ถ้ามีฟังก์ชันใดฟังก์ชันหนึ่งใน $f, \{g_i\}, \{h_j\}$ เป็นฟังก์ชันไม่เชิงเส้น เราเรียกปัญหา (A) ว่าปัญหากำหนดการไม่เชิงเส้น (nonlinear programming problem) จะเห็นว่าปัญหา (A) เป็นปัญหาแบบมีเงื่อนไขบังคับ ดังนั้นแนวทางในการแก้ปัญหา (A) จะอยู่บนพื้นฐานของการแปลงปัญหาแบบมีเงื่อนไขบังคับไปเป็นลำดับของปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ การแปลงปัญหาดังกล่าวจะทำได้สำเร็จถ้ามีฟังก์ชันช่วย (auxiliary function) ที่เหมาะสม ซึ่งฟังก์ชันช่วยดังกล่าวจะเป็นฟังก์ชันของเงื่อนไขบังคับทั้งหมดและฟังก์ชันจุดประสงค์เดิม จะเห็นว่าขณะนี้ปัญหาที่ได้เป็นปัญหาแบบไม่มีเงื่อนไขบังคับเมื่อทำการหาค่าต่ำสุดและกำจัดผลของเงื่อนไขบังคับในฟังก์ชันช่วยโดยการควบคุมผ่านค่าพารามิเตอร์ จึงสามารถคาดหวังได้ว่าลำดับของปัญหาแบบไม่มีเงื่อนไขบังคับที่กำเนิดขึ้นมาจะมีผลเฉลยลู่เข้าสู่ผลเฉลยของปัญหาแบบมีเงื่อนไขบังคับเริ่มต้นซึ่งก็คือปัญหา (A) นั่นเอง

วิธีการแปลงปัญหาที่จะใช้ในงานวิจัยนี้เป็นวิธีที่ใช้กันอย่างกว้างขวางและมีพัฒนาการต่อเนื่องมาเป็นลำดับ วิธีการดังกล่าวได้แก่ วิธีฟังก์ชันกั้นเขต (barrier function method) ซึ่งใช้กับปัญหา (A) ในกรณีที่มีเงื่อนไขบังคับเป็นแบบอสมการเท่านั้นและวิธีฟังก์ชันทัณฑ์ (penalty function method) ซึ่งมักใช้กับปัญหา (A) ในกรณีที่มีเงื่อนไขบังคับแบบสมการ นอกจากนี้ยังได้มีผู้พัฒนาวิธีการทั้งสองมารวมกัน เพื่อให้สามารถแก้ปัญหาค่าต่ำสุดแบบมีเงื่อนไขบังคับได้ทั้ง

เงื่อนไขแบบอสมการและสมการ เรียกวิธีแก้ปัญหาดังกล่าวว่า วิธีผสม (mixed method)

2.3.1 วิธีฟังก์ชันกั้นเขต

ในงานวิจัยหลายชิ้นเรียกวิธีฟังก์ชันกั้นเขตนี้ว่า วิธีจุดภายใน (interior point method) เพราะวิธีการนี้ต้องการจุดเริ่มต้น x_0 ที่เป็นจุดภายใน (interior) ของบริเวณที่เป็นไปได้ (feasible region) ของเงื่อนไขบังคับแบบอสมการทั้งหมด เพื่อใช้ในการเริ่มต้นแก้ปัญหาแล้วทำการกำเนิดจุดภายใน x_1 ต่อไปและจะใช้ x_1 แก้ปัญหาเป็นเช่นนี้เรื่อยไปจนถึงจุดเหมาะที่สุด x^* ดังนั้น วิธีฟังก์ชันกั้นเขตจึงใช้ได้กับการแก้ปัญหาค่าต่ำสุดแบบมีเงื่อนไขบังคับเป็นแบบอสมการเท่านั้น โดยที่เซตของจุดภายในของบริเวณที่เป็นไปได้ $\{x \mid g_i(x) < 0 \; ; i = 1,\ldots,m\}$ ไม่เป็นเซตว่าง

พิจารณาปัญหา (B)

$$\left.\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \le 0 \;\; ; i = 1,\ldots,m \end{array}\right\} \qquad (2.25)$$

แนวคิดพื้นฐานของวิธีฟังก์ชันกั้นเขตเป็นดังนี้ ให้ β เป็นฟังก์ชันค่าจริงของ x ซึ่งมีคุณสมบัติ 2 ข้อต่อไปนี้

คุณสมบัติ 1 $\beta(x)$ เป็นฟังก์ชันต่อเนื่องในบริเวณ $G_0 = \{x \mid g_i(x) < 0 \; ; i = 1,\ldots,m\}$

คุณสมบัติ 2 ถ้า $\{x_k\}$ เป็นลำดับอนันต์ของจุดใน G_0 ซึ่งลู่เข้าหา x_B ซึ่ง $g_i(x_B) = 0$ สำหรับบาง $i \in \{1,\ldots,m\}$ แล้ว $\lim\limits_{k\to\infty} \beta(x_k) = +\infty$ เราเรียก $\beta(x)$ ว่าฟังก์ชันกั้นเขต (barrier function)

ให้ $s(\mu)$ เป็นฟังก์ชันค่าจริงของตัวแปรเดียว μ ซึ่งมีคุณสมบัติต่อไปนี้ ถ้า $\mu_1 > \mu_2 > 0$ แล้ว $s(\mu_1) > s(\mu_2) > 0$, ถ้า $\{\mu_k\}$ เป็นลำดับอนันต์ของจุดซึ่ง $\lim\limits_{k\to\infty} \mu_k = 0$ แล้ว $\lim\limits_{k\to\infty} s(\mu_k) = 0$ เราเรียก $s(\mu)$ ว่าพารามิเตอร์กั้นเขต (barrier parameter)

สำหรับกระบวนการแก้ปัญหาโดยวิธีฟังก์ชันกั้นเขตเป็นลำดับดังนี้

1. กำหนดฟังก์ชัน $U(x,\mu_1) = f(x) + s(\mu_1)\beta(x)$ เรียก U ว่าฟังก์ชันช่วยและ μ_1 เป็นจำนวนบวก เราจะใช้จุด $x_0 \in G_0$ เป็นจุดเริ่มต้น

2. ดำเนินกระบวนการจาก x_0 ไปยังจุด $x_1 = x(\mu_1)$ ซึ่งเป็นจุดต่ำสุดเฉพาะที่ของ U ในบริเวณที่เป็นไปได้ $G = \{x \mid g_i(x) \le 0 \; ; i = 1,\ldots,m\}$ ซึ่งคาดหวังได้ว่า $x_1 \in G_0$ เพราะถ้าเป็นอย่างอื่นแล้ว $U = +\infty$ ซึ่งขัดแย้งกับที่สมมติว่า x_1 เป็นจุดต่ำสุดแบบไม่มีเงื่อนไขบังคับเฉพาะที่ (local unconstrained minimum) ของ U ใน G

3. เริ่มจากจุด x_1 หาจุดต่ำสุดเฉพาะที่ของ $U(x,\mu_2)$ ซึ่ง $\mu_1 > \mu_2 > 0$

4. กระทำแบบเดียวกับที่กล่าวข้างต้น จนสามารถหาจุดต่ำสุดเฉพาะที่ $x_k = x(\mu_k)$ สำหรับ $U(x, \mu_k)$ โดยเริ่มจากจุด $x_{k-1} = x(\mu_{k-1})$ สำหรับลำดับลดทางเดียวโดยแท้ (strictly monotonically decreasing sequence) $\{\mu_k\}$

ภายใต้ข้อสมมติที่เหมาะสม ลำดับของจุดต่ำสุดแบบไม่มีเงื่อนไขบังคับเฉพาะที่จะเกิดขึ้น และมีจุดลิมิตเป็นจุดต่ำสุดเฉพาะที่ของปัญหา (B) นั่นเอง ถ้าเราพิจารณา $\{\mu_k\}$ โดยที่ $\mu_1 > \mu_2 > \ldots > \mu_k > \ldots > 0$ จะเห็นว่าลำดับ μ_k สอดคล้องกับเงื่อนไขการเป็นฟังก์ชัน $s(\mu)$ ดังนั้นเพื่อความง่ายในงานวิจัยนี้ จะใช้ $\{\mu_k\}$ เป็นลำดับที่แทนฟังก์ชัน $s(\mu)$ โดยที่ $\{\mu_k\}$ เป็นลำดับลดลงทางเดียวโดยแท้

เมื่อพิจารณาเทอม $\mu\beta(x)$ ซึ่งเป็นส่วนที่เพิ่มเข้าไปให้กับฟังก์ชันจุดประสงค์ $f(x)$ เทอม $\mu\beta(x)$ นี้จะรับรองว่าจุดต่ำสุดของฟังก์ชัน $U(x,\mu)$ ที่ได้จะอยู่ภายในบริเวณที่เป็นไปได้ เช่น เมื่อเริ่มด้วย x_0 ซึ่ง $g_i(x_0) < 0$; $i = 1,\ldots,m$ และสมมติว่าใช้เส้นทางลดเร็วสุด (steepest descent) ในการหาค่าต่ำสุดของ $U(x, \mu_1)$ จะต้องมีจุด x_1 ซึ่งเป็นจุดต่ำสุดที่ทำให้ $U(x, \mu_1)$ มีค่าจำกัดเพราะทิศทางที่เดินจะทำให้ U ลดลงอย่างต่อเนื่อง ดังนั้นจะไม่มีจุดบนเส้นทางที่เดินให้ค่าของ U มากกว่า $U(x_0, \mu_1)$ ได้ ในเมื่อ $U(x_1, \mu_1) < \infty$ ดังนั้น x_1 จึงไม่ใช่จุดขอบ (boundary point) นั่นคือ x_1 เป็นจุดภายในบริเวณที่เป็นไปได้

ในความเป็นจริง เราต้องการฟังก์ชัน $\sigma(x)$ ที่มีคุณสมบัติว่า

$$\sigma(x) = \begin{cases} 0 & \text{ถ้า } x \in G \\ +\infty & \text{ถ้า } x \notin G \end{cases}$$

เป็นฟังก์ชันกั้นเขต เพราะ $\sigma(x)$ ไม่มีผลใด ๆ ต่อการหาค่าต่ำสุดของ $U(x) = f(x) + \sigma(x)$ ถ้า x เป็นจุดที่เป็นไปได้และ $\sigma(x)$ เพิ่มค่าเป็น $+\infty$ ให้กับ $f(x)$ ในกรณีที่ x อยู่นอกบริเวณที่เป็นไปได้ซึ่งไม่สูญเสียอะไร เพราะขบวนการจะหาค่าต่ำสุดของ $U(x)$ ต่อไป แต่ปัญหาที่เกิดขึ้นคือความยากในการแก้ปัญหาค่าต่ำสุดของ $U(x)$ เนื่องจากความไม่ต่อเนื่องของฟังก์ชัน $\sigma(x)$ ดังนั้นจึงได้มีความพยายามหาฟังก์ชัน $\mu\beta(x)$ เพื่อใช้แทน $\sigma(x)$ โดยการควบคุมผ่านพารามิเตอร์กั้นเขต μ เมื่อ μ มีค่าน้อยมากแล้ว $\mu\beta(x) \approx \sigma(x)$ ได้โดยที่ $\mu\beta(x)$ เป็นฟังก์ชันต่อเนื่อง

รูปแบบที่เป็นไปได้ของฟังก์ชัน $\beta(x)$ คือ

$$\beta(x) = \sum_{i=1}^{m} \Phi[g_i(x)]$$

เมื่อ Φ เป็นฟังก์ชันของตัวแปรเดียวซึ่งต่อเนื่องบน $\{y \mid y < 0\}$ และสอดคล้องกับ $\lim_{y \to 0^-} \Phi(y) = \infty$ ฟังก์ชัน Φ ที่เป็นไปได้มี 2 รูปแบบคือ ฟังก์ชันลอการิทึมและฟังก์ชันผกผัน ดังนั้น $\beta(x)$ จะอยู่ในรูปแบบ

$$\beta(x) = -\sum_{i=1}^{m} \ln[-g_i(x)] \tag{2.26}$$

และ

$$\beta(x) = \sum_{i=1}^{m} -\frac{1}{g_i(x)} \tag{2.27}$$

ตัวอย่างต่อไปนี้แสดงให้เห็นว่า $\mu\beta(x)$ สามารถประมาณ $\sigma(x)$ ได้ดีเมื่อ μ มีขนาดเล็ก ๆ พิจารณาเงื่อนไขบังคับ $1 \le x \le 5$ เมื่อจัดรูปแบบจะได้

$$g_1(x) = -x + 1 \le 0$$
$$g_2(x) = \ x - 5 \le 0$$

ดังนั้น $\mu\beta(x) = -\mu\left[\frac{1}{1-x} + \frac{1}{x-5}\right]$ เมื่อทำการพล็อตกราฟจะได้ดังรูปที่ 2.3 จะเห็นว่าเมื่อเราใช้ $\mu\beta(x)$ แทน $\sigma(x)$ จะทำให้เราควบคุมค่า $\mu\beta(x)$ ให้ใกล้ศูนย์ได้เมื่ออยู่ในบริเวณ $\{x \mid g_i(x) < 0 \ ; i = 1,\ldots,m\}$ และให้มีค่ามาก ๆ (ถือว่าใกล้ $+\infty$) เมื่อ $g_i(x) = 0$ สำหรับบาง $i \in \{1,\ldots,m\}$ การกำหนด $\mu\beta(x)$ ในลักษณะเช่นนี้จึงเป็นการป้องกันการหนีออกของจุดจากบริเวณ $\{x \mid g_i(x) \le 0 \ ; i = 1,\ldots,m\}$

ต่อไปนี้จะให้ทฤษฎีบทการลู่เข้าสู่จุดต่ำสุดของวงกว้างของปัญหา (B) เมื่อใช้วิธีฟังก์ชันกั้นเขต ส่วนทฤษฎีการลู่เข้าสู่จุดต่ำสุดเฉพาะที่ของปัญหา (B) สามารถดูได้จากเอกสารอ้างอิง [10] ในที่นี้กำหนดให้ $\theta(\mu) = \inf\{f(x) + \mu\beta(x) \mid g_i(x) < 0 \ ; i = 1,\ldots,m, x \in R^n\}$

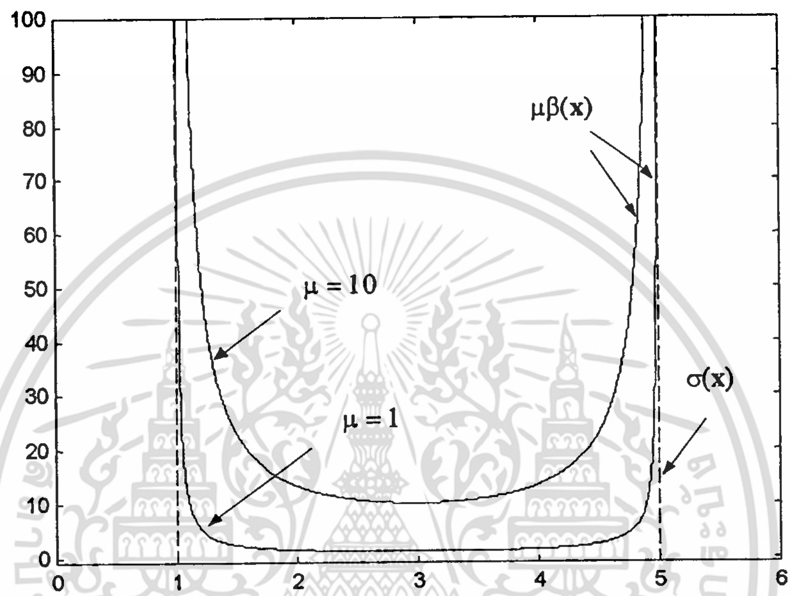
บทตั้งที่ 2.19 ให้ $f, g_1,\ldots,g_m$ เป็นฟังก์ชันต่อเนื่องบน R^n สมมติว่า $\{x \in R^n \mid g_i(x) < 0 \ ; i = 1,\ldots,m\} \ne \phi$ และ β เป็นฟังก์ชันกั้นเขตที่กำหนดโดย (2.26) หรือ (2.27) และ β ต่อเนื่องบน $\{x \mid g_i(x) < 0 \ ; i = 1,\ldots,m\}$ สมมติเพิ่มเติมว่าสำหรับ $\mu (> 0)$ ใด ๆ ที่กำหนดให้ ถ้า $\{x_k\}$ ใน R^n สอดคล้องกับ $g_i(x_k) < 0 \ ; i = 1,\ldots,m$ และ $f(x_k) + \mu\beta(x_k) \to \theta(\mu)$ แล้ว $\{x_k\}$ มีลำดับย่อยที่ลู่เข้า (ข้อสมมตินี้เป็นจริง ถ้า $\{x \in R^n \mid g_i(x) \le 0 \ ; i = 1,\ldots,m\}$ เป็นเซตกระชับ) แล้ว

(1). สำหรับแต่ละ $\mu > 0$ จะมี $x_\mu \in R^n$ ซึ่ง $g_i(x_\mu) < 0 \ ; i = 1,\ldots,m$ ที่ทำให้ $\theta(\mu) = f(x_\mu) + \mu\beta(x_\mu) = \inf\{(x) + \mu\beta(x) \mid g_i(x) < 0 \ ; i = 1,\ldots,m$ และ $x \in R^n\}$

(2). $\inf\{f(\mathbf{x}) \mid g_i(\mathbf{x}) \le 0 \;\; ; i = 1,\ldots,m, \mathbf{x} \in R^n\} \le \inf\{\theta(\mu) \mid \mu > 0\}$

(3). สำหรับ $\mu > 0$ จะได้ว่า $f(\mathbf{x}_\mu)$ และ $\theta(\mu)$ เป็นฟังก์ชันไม่ลดของ μ และ $\beta(\mathbf{x}_\mu)$ เป็นฟังก์ชันไม่เพิ่มของ μ

พิสูจน์ ดูเอกสารอ้างอิง [1]



รูปที่ 2.3 แสดงการใช้ฟังก์ชัน $\mu\beta(x)$ แทน $\sigma(x)$ โดยใช้ μ ในการควบคุม

ทฤษฎีบทที่ 2.20 ให้ $f : R^n \to R^1$ และ $g_i : R^n \to R^m (i = 1,\ldots, m)$ เป็นฟังก์ชันต่อเนื่องและสมมติว่า $\{\mathbf{x} \in R^n \mid g_i(\mathbf{x}) < 0 \;\; ; i = 1,\ldots, m\} \ne \phi$ สมมติเพิ่มเติมว่าปัญหาเดิม minimize $f(\mathbf{x})$ subject to $g_i(\mathbf{x}) \le 0$ $(i = 1,\ldots, m), \mathbf{x} \in R^n$ มีผลเฉลยเหมาะที่สุด $\mathbf{x}^*$ ซึ่งมีคุณสมบัติว่าเมื่อกำหนดย่านใกล้เคียง N รอบ $\mathbf{x}^*$ จะมี $\mathbf{x} \in R^n \cap N$ ซึ่ง $g_i(\mathbf{x}) < 0$ $(i = 1,\ldots,m)$ แล้ว

$$\text{minimum } \{f(\mathbf{x}) \mid g_i(\mathbf{x}) \le 0 \;\; ; i = 1,\ldots,m, \mathbf{x} \in R^n\} = \lim_{\mu \to 0^+} \theta(\mu) = \inf_{\mu > 0} \theta(\mu)$$

เมื่อกำหนดให้ $\theta(\mu) = f(\mathbf{x}_\mu) + \mu\beta(\mathbf{x}_\mu)$ ซึ่ง $\mathbf{x}_\mu \in R^n$ และ $g_i(\mathbf{x}_\mu) < 0$ $(i = 1,\ldots,m)$ แล้วลิมิตของทุก ๆ ลำดับย่อยที่ลู่เข้าของ $\{\mathbf{x}_\mu\}$ จะเป็นผลเฉลยเหมาะที่สุดของปัญหาเดิมและ $\mu\beta(\mathbf{x}_\mu) \to 0$ เมื่อ $\mu \to 0^+$

พิสูจน์ ดูเอกสารอ้างอิง [1]

บทตั้งต่อไปนี้จะให้เงื่อนไขว่า สำหรับแต่ละ μ จะมี x_μ ซึ่งเป็นจุดต่ำสุดวงกว้างของ $U(x,\mu)$

บทตั้งที่ 2.21 สมมติว่า $G = \{x \mid g_i(x) \le 0 \ \ ; i = 1,\ldots,m \ , \ x \in R^n\}$ เป็นเซตกระชับแล้ว สำหรับค่า μ ใด ๆ ที่เป็นบวกจะมีจุด $x_\mu \in G_0 = \{x \mid g_i(x) < 0 \ \ ; i = 1,\ldots,m, x \in R^n\}$ ที่ทำให้ $U(x,\mu) = f(x) + \mu\beta(x)$ มีค่าต่ำสุด

พิสูจน์ ดูเอกสารอ้างอิง [9]

สำหรับขั้นตอนวิธีฟังก์ชันกั้นเขตเป็นดังนี้

ขั้นเริ่มต้น ให้ $\varepsilon > 0$ เป็นค่าสเกลาร์ในการหยุดและเลือกจุด $x_0 \in R^n$ ซึ่ง $g_i(x_0) < 0$ $(i = 1,\ldots,m)$, ให้ $\mu_0 > 0, \gamma \in (0,1)$, ให้ $k = 0$ และไปยังขั้นหลัก

ขั้นหลัก

1. เริ่มต้นด้วยจุด x_k แล้วแก้ปัญหาต่อไปนี้

$$\begin{array}{ll} \text{minimize} & f(x) + \mu_k\beta(x) \\ \text{subject to} & g_i(x) < 0 \ \ ; i = 1,\ldots,m \\ & x \in R^n \end{array}$$

ให้ x_{k+1} เป็นผลเฉลยเหมาะที่สุดของปัญหาข้างต้นและไปยังขั้นที่ 2

2. ถ้า $\mu_k\beta(x_{k+1}) < \varepsilon$ ให้หยุด ถ้าเป็นอย่างอื่นให้ $\mu_{k+1} = \gamma\mu_k$ แล้วเปลี่ยน k เป็น $k+1$ และไปทำซ้ำที่ขั้นที่ 1

2.3.2 วิธีฟังก์ชันทัณฑ์

แนวคิดของวิธีฟังก์ชันทัณฑ์มีความคล้ายกับวิธีฟังก์ชันกั้นเขตในลักษณะของการเพิ่มพจน์บางพจน์เข้าไปในฟังก์ชันจุดประสงค์เดิมในกรณีที่จุดที่กำเนิดได้ไม่สอดคล้องกับเงื่อนไขบังคับ และในทำนองเดียวกันกับวิธีฟังก์ชันกั้นเขต ถ้าจุดที่กำเนิดได้สอดคล้องกับเงื่อนไขบังคับพจน์ที่เพิ่มเข้าไปจะไม่มีผลใด ๆ แต่ความแตกต่างของวิธีทั้งสองคือ วิธีฟังก์ชันทัณฑ์อาจเริ่มต้นด้วยจุด x_0 ที่ไม่ได้อยู่ในเซตที่เป็นไปได้แล้วทำการกำเนิดจุด $x_1, x_2, \ldots$ ซึ่งเป็นจุดที่เป็นไปไม่ได้ ในที่สุดแล้วจะลู่เข้าหาผลเฉลยเหมาะที่สุด x^* ซึ่งเป็นจุดที่เป็นไปได้ ดังนั้นในบางครั้งจะเรียกวิธีฟังก์ชันทัณฑ์นี้ว่า วิธีจุดภายนอก (exterior point method)

พิจารณาปัญหาค่าต่ำสุด (A) ใน (2.24)

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && g_i(x) \le 0 \quad ; i = 1,\ldots,m \\ & && h_j(x) = 0 \quad ; j = 1,\ldots,r \end{aligned}$$

แนวคิดพื้นฐานของวิธีฟังก์ชันทัณฑ์เป็นดังนี้ ให้ $\alpha(x)$ เป็นฟังก์ชันต่อเนื่องของ x ซึ่งมีคุณสมบัติว่า $\alpha(x) = 0$ ถ้า $x \in F = \{x \in R^n \mid g_i(x) \le 0 ; i = 1,\ldots,m$ และ $h_j(x) = 0 ; j = 1,\ldots,r\}$ และ $\alpha(x) > 0$ ถ้า $x \notin F$ ให้ $s(\rho)$ เป็นฟังก์ชันค่าจริงของตัวแปรเดียว ρ ซึ่งมีคุณสมบัติว่า ถ้า $0 < \rho_1 < \rho_2$ แล้ว $0 < s(\rho_1) < s(\rho_2)$ และถ้า $\{\rho_k\}$ เป็นลำดับเพิ่มขึ้นทางเดียวของค่าบวก ซึ่ง $\lim_{k\to\infty} \rho_k = +\infty$ แล้ว $\lim_{k\to\infty} s(\rho_k) = +\infty$ กำหนดฟังก์ชัน $P(x,\rho) = f(x) + s(\rho)\alpha(x)$ เรียกฟังก์ชัน P ว่าฟังก์ชันช่วย, $s(\rho)$ ว่าพารามิเตอร์ทัณฑ์ (parameter), $\alpha(x)$ ว่าฟังก์ชันทัณฑ์

กระบวนการแก้ปัญหาโดยวิธีฟังก์ชันทัณฑ์เป็นดังนี้

(1). สำหรับ $\rho_1 > 0$ หาจุดต่ำสุดเฉพาะที่แบบไม่มีเงื่อนไขบังคับของ $P(x,\rho_1)$ แทนด้วย $x_1 = x(\rho_1)$

(2). ใช้จุด x_1 เพื่อกำเนิดจุดต่ำสุดเฉพาะที่แบบไม่มีเงื่อนไขบังคับของ $P(x,\rho_2)$ ซึ่ง $\rho_2 > \rho_1$

(3). ทำกระบวนการซ้ำเช่นเดียวกับที่กล่าวข้างต้น ด้วยการหาค่าต่ำสุดของ $P(x,\rho_k)$ สำหรับลำดับเพิ่มขึ้นทางเดียว $\{\rho_k\}$

เมื่อ $\rho_k \to \infty$ คาดหวังว่าลำดับของจุดต่ำสุดเฉพาะที่แบบไม่มีเงื่อนไขบังคับจะลู่เข้าหาจุดต่ำสุดเฉพาะที่แบบมีเงื่อนไขบังคับ x^*

พิจารณา $\{\rho_k\}$ ซึ่ง $\rho_1 < \rho_2 < \ldots < \rho_k < \ldots + \infty$ จะเห็นว่า $\{\rho_k\}$ สอดคล้องกับเงื่อนไขของฟังก์ชัน $s(\rho)$ ดังนั้นเพื่อความง่ายเราจะใช้ $\{\rho_k\}$ แทนฟังก์ชัน $s(\rho)$ การกำหนดฟังก์ชันทัณฑ์ $\alpha(x)$ ที่เหมาะสมจะต้องกำหนดให้เป็นฟังก์ชันของเงื่อนไขบังคับทั้งแบบอสมการ $(g_i(x) \; ; i = 1,\ldots,m)$ และสมการ $(h_j(x) \; ; j = 1,\ldots,r)$ ดังนั้นจึงต้องกำหนดลักษณะฟังก์ชัน $\alpha(x)$ เป็นผลบวกของฟังก์ชันสองกลุ่มเพื่อให้สอดคล้องกับเงื่อนไขบังคับดังนี้

$$\alpha(x) = \sum_{i=1}^{m} \Phi[g_i(x)] + \sum_{j=1}^{r} \Psi[h_j(x)]$$

ซึ่ง Φ และ Ψ เป็นฟังก์ชันต่อเนื่องและสอดคล้องกับเงื่อนไขต่อไปนี้

$$\Phi(y) = 0 \text{ ถ้า } y \le 0 \text{ และ } \Phi(y) > 0 \text{ ถ้า } y > 0$$
$$\Psi(y) = 0 \text{ ถ้า } y = 0 \text{ และ } \Psi(y) > 0 \text{ ถ้า } y \ne 0$$

ดังนั้นรูปแบบของ Φ และ Ψ อาจกำหนดโดย

$$\Phi(y) = [\max\text{imum}\{0, y\}]^p$$
$$\Psi(y) = |y|^p$$

ซึ่ง p เป็นจำนวนเต็มบวก ดังนั้นฟังก์ชันทัณฑ์ $\alpha(\mathbf{x})$ ที่นำไปใช้จะอยู่ในรูปแบบ

$$\alpha(\mathbf{x}) = \sum_{i=1}^{m}[\max\text{imum}\{0, g_i(\mathbf{x})\}]^p + \sum_{j=1}^{r}\left|h_j(\mathbf{x})\right|^p \tag{2.28}$$

ในทางปฏิบัติเราเลือกให้ $p = 2$ ดังนั้นจะเรียกฟังก์ชัน $\alpha(\mathbf{x})$ ซึ่งใช้ $p = 2$ ว่าฟังก์ชันสูญเสียกำลังสอง (quadratic-loss function)

ถ้าพิจารณาปัญหาที่มีเฉพาะเงื่อนไขบังคับในรูปแบบอสมการเท่านั้น (ปัญหา (B) ใน (2.25)) จะได้ฟังก์ชันสูญเสียกำลังสองในกรณีนี้เป็น

$$\alpha(\mathbf{x}) = \sum_{i=1}^{m}[\max\text{imum}\{0, g_i(\mathbf{x})\}]^2$$

อนุพันธ์อันดับหนึ่งของ $\alpha(\mathbf{x})$ คือ

$$\nabla\alpha(\mathbf{x}) = 2\sum_{i=1}^{m}[\max\text{imum}\{0, g_i(\mathbf{x})\}]\nabla g_i(\mathbf{x})$$

ซึ่งเป็นฟังก์ชันต่อเนื่องและอนุพันธ์อันดับสองของ $\alpha(\mathbf{x})$ คือ

$$\nabla^2\alpha(\mathbf{x}) = 2[\sum_{i=1}^{m}[\max\text{imum}\{0, g_i(\mathbf{x})\}]\nabla^2 g_i(\mathbf{x}) + \sum_{i=1}^{m}\nabla g_i(\mathbf{x})(\nabla g_i(\mathbf{x}))^T]$$

เป็นฟังก์ชันที่ไม่ต่อเนื่องที่จุด $\mathbf{x}$ ซึ่ง $g_i(\mathbf{x}) = 0$ สำหรับบาง $i \in \{1,\ldots,m\}$ เนื่องจากอนุพันธ์อันดับสองของฟังก์ชันมีความสำคัญต่อการใช้วิธีนิวตันและวิธีแนวนิวตันในการแก้ปัญหาค่าต่ำสุดของฟังก์ชัน ดังนั้นเรามักใช้วิธีฟังก์ชันทัณฑ์กับปัญหาแบบมีเงื่อนไขบังคับเป็น

แบบสมการเท่านั้น สำหรับทฤษฎีที่เกี่ยวข้องกับการลู่เข้าหาจุดต่ำสุดเฉพาะที่โดยวิธีฟังก์ชันทัณฑ์ สามารถดูได้จากเอกสารอ้างอิง [10] ในที่นี้จะได้กล่าวถึงทฤษฎีของการลู่เข้าหาจุดต่ำสุดวงกว้างเท่านั้นและกำหนดให้ $\theta(\rho) = \inf\{f(x) + \rho\alpha(x) \mid x \in R^n\}$

บทตั้งที่ 2.22 สมมติว่า $f, h_1, \ldots, h_r$ เป็นฟังก์ชันต่อเนื่องบน R^n และให้ α เป็นฟังก์ชันต่อเนื่องบน R^n ซึ่งกำหนดโดย $\alpha(x) = \sum_{j=1}^{r} \Psi[h_j(x)]$ เมื่อ $\Psi(y) = y^2$ และสมมติว่าสำหรับแต่ละ ρ มี $x_\rho \in R^n$ ที่ทำให้ $\theta(\rho) = f(x_\rho) + \rho\alpha(x_\rho)$ แล้วข้อความต่อไปนี้เป็นจริง

(1). $\inf\{f(x) \mid x \in R^n, h_j(x) = 0 \ ; \ j = 1,\ldots,r\} \geq \sup_{\rho \geq 0} \theta(\rho)$ ซึ่ง $\theta(\rho) = \inf\{f(x) + \rho\alpha(x) \mid x \in R^n\}$

(2). $f(x_\rho)$ และ $\theta(\rho)$ เป็นฟังก์ชันไม่ลดของ ρ และ $\alpha(x_\rho)$ เป็นฟังก์ชันไม่เพิ่มของ ρ

พิสูจน์ ดูเอกสารอ้างอิง [1]

ทฤษฎีบทที่ 2.23 พิจารณาปัญหาต่อไปนี้

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & h_j(x) = 0 \quad ; \quad j = 1,\ldots,r \\ & x \in R^n \end{array}$$

ซึ่ง $f, h_1, \ldots, h_r$ เป็นฟังก์ชันต่อเนื่องบน R^n สมมติว่าปัญหามีผลเฉลยที่เป็นไปได้และให้ α เป็นฟังก์ชันต่อเนื่องที่กำหนดโดย $\alpha(x) \sum_{j=1}^{r} \Psi\,[h_j(x)]$ ซึ่ง $\Psi(y) = y^2$ สมมติเพิ่มเติมว่าสำหรับแต่ละ ρ จะมีผลเฉลย $x_\rho \in R^n$ ของปัญหา minimize $f(x) + \rho\alpha(x)$ subject to $x \in R^n$ และสมมติว่า $\{x_\rho\}$ เป็นเซตย่อยในเซตย่อยกระชับของ R^n แล้ว

$$\inf\{f(x) \mid h_j(x) = 0 \ ; \ j = 1,,\ldots,r, \ x \in R^n\} = \sup_{\rho \geq 0} \theta(\rho) = \lim_{\rho \to \infty} \theta(\rho)$$

ซึ่ง $\theta(\rho) = \inf\{f(x) + \rho\alpha(x) \mid x \in R^n\} = f(x_\rho) + \rho\alpha(x_\rho)$ นอกจากนั้นลิมิต x^* ของลำดับย่อยใด ๆ ที่ลู่เข้าของ $\{x_\rho\}$ จะเป็นผลเฉลยเหมาะที่สุดของปัญหาเดิม และ $\rho\alpha(x_\rho) \to 0$ เมื่อ $\rho \to \infty$

พิสูจน์ ดูเอกสารอ้างอิง [1]

บทแทรกที่ 2.24 ถ้า $\alpha(\mathbf{x}_\rho) = 0$ สำหรับบางค่าของ ρ แล้ว $\mathbf{x}_\rho$ จะเป็นผลเฉลยเหมาะที่สุดของปัญหา

พิสูจน์ ดูเอกสารอ้างอิง [1]

สำหรับขั้นตอนวิธีฟังก์ชันทัณฑ์เป็นดังนี้

ขั้นเริ่มต้น ให้ $\varepsilon > 0$ เป็นค่าสเกลาร์ในการหยุดและทำการเลือกจุดเริ่มต้น $\mathbf{x}_0$, ค่าพารามิเตอร์ทัณฑ์ $\rho_0 > 0$ และสเกลาร์ $\beta > 1$ ให้ $k = 0$ และไปยังขั้นหลัก

ขั้นหลัก

(1). เริ่มด้วยจุด $\mathbf{x}_k$ แล้วแก้ปัญหาต่อไปนี้

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) + \rho_k \alpha(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in R^n \end{array}$$

ให้ $\mathbf{x}_{k+1}$ เป็นผลเฉลยเหมาะที่สุดของปัญหาและไปยังขั้นที่ 2

(2). ถ้า $\rho_k \alpha(\mathbf{x}_{k+1}) < \varepsilon$ ให้หยุด ถ้าเป็นอย่างอื่นให้ $\rho_{k+1} = \beta\rho_k$ และเปลี่ยน k เป็น $k+1$ และไปยังขั้นที่ 1

2.3.3 วิธีผสมระหว่างวิธีฟังก์ชันกั้นเขตและวิธีฟังก์ชันทัณฑ์

เพื่อให้การแก้ปัญหาค่าต่ำสุดแบบมีเงื่อนไขบังคับครอบคลุมทั้งเงื่อนไขแบบอสมการและสมการและเป็นไปอย่างมีประสิทธิภาพ ฟีแอคโคและแม็คคอร์มิค(Fiacco and McCormick) [10]จึงได้ให้รูปแบบของการแก้ปัญหาซึ่งเป็นลักษณะผสมระหว่างวิธีฟังก์ชันกั้นเขตและวิธีฟังก์ชันทัณฑ์ ดังต่อไปนี้

พิจารณาปัญหา (A) ใน (2.24)

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}) & \\ \text{subject to} & g_i(\mathbf{x}) \le 0 & ; \; i = 1,\ldots,m \\ & h_j(\mathbf{x}) = 0 & ; \; j = 1,\ldots,r \end{array}$$

เมื่อทำการแปลงปัญหาโดยอาศัยแนวคิดของวิธีฟังก์ชันกั้นเขตและวิธีฟังก์ชันทัณฑ์ผสมกัน จะได้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับในรูปแบบ

$$\text{minimize} \quad M(x,\mu,\rho) = f(x) + \mu\beta(x) + \rho\alpha(x) \text{ เมื่อ } x \in R^n$$

โดยที่ $\beta(x)$ คือ ฟังก์ชันกั้นเขตซึ่งสมนัยกับเงื่อนไขอสมการ และ $\alpha(x)$ คือ ฟังก์ชันทัณฑ์ซึ่งสมนัยกับเงื่อนไขสมการ จากลำดับ $\{\mu_k\}$ และ $\{\rho_k\}$ เราอาจให้ความสัมพันธ์ระหว่าง μ_k และ ρ_k เป็น $\rho_k = \dfrac{1}{\mu_k}$ โดยไม่สูญเสียอะไร

ภายใต้ข้อกำหนดที่เหมาะสม โดยเฉพาะอย่างยิ่งแต่ละ μ , ฟังก์ชัน $M(x,\mu,\frac{1}{\mu})$ จะต้องมีจุดต่ำสุดวงกว้าง $x_\mu = x(\mu)$ จะได้ว่า

$$\inf\{f(x) \mid g_i(x) \le 0 \ (i=1,\dots,m), h_j(x) = 0 \ (j=1,\dots,r \text{ และ } x \in R^n\} = \lim_{\mu\to 0^+} \sigma(\mu) \tag{2.29}$$

และ $\quad \mu\beta(x_\mu) \to 0 \, , \, \frac{1}{\mu}\alpha(x_\mu) \to 0 \quad$ เมื่อ $\mu \to 0^+$

ซึ่ง
$$\begin{aligned}\sigma(\mu) &= \inf\{f(x) + \mu\beta(x) + \frac{1}{\mu}\alpha(x) \mid x \in R^n, g_i(x) < 0 \quad ; \; i=1,\dots,m\} \\ &= f(x_\mu) + \mu\beta(x_\mu) + \frac{1}{\mu}\alpha(x_\mu)\end{aligned} \tag{2.30}$$

การพิสูจน์ผลลัพธ์ข้างบนดูได้จากเอกสารอ้างอิง [10] จาก (2.30) จะเห็นว่าในการแก้ปัญหาแบบไม่มีเงื่อนไขบังคับในแต่ละรอบ เราต้องควบคุมให้จุดที่กำเนิดได้อยู่ภายในบริเวณที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการเสมอที่เป็นเช่นนี้ เนื่องจากผลของวิธีฟังก์ชันกั้นเขตและสิ่งที่สำคัญในการแก้ปัญหาผสมคือ ต้องใช้จุด x_0 ซึ่ง $g_i(x_0) < 0 \;\; ; i = 1,\dots,m$ เป็นจุดเริ่มต้นในการแก้ปัญหาด้วย

2.4 งานวิจัยที่เกี่ยวข้อง

ในหัวข้อนี้จะได้นำเสนอผลงานวิจัยที่เกี่ยวข้องซึ่งมีความสำคัญต่อวิทยานิพนธ์ซึ่งในส่วนแรกจะกล่าวถึง ทฤษฎีการลู่เข้าหาผลเฉลยเหมาะที่สุดโดยวิธีแนวนิวตันเมื่อใช้สูตร BFGS สำหรับการแก้ปัญหาค่าเหมาะที่สุดแบบไม่มีเงื่อนไขบังคับและส่วนหลังจะกล่าวถึงงานวิจัยที่เกี่ยวข้องกับกำหนดการนูน (convex programming)

ในปี ค.ศ. 1973 นักคณิตศาสตร์ชื่อ บรอยเดน , เดนนิส และมัวร์ (C.G. Broyden , J.E. Dennis , J.J. Mor'e) ได้เสนองานวิจัยเรื่อง "On the local and superlinear convergence of quasi-Newton methods" [11] โดยใจความสำคัญกล่าวถึงทฤษฎีบทการลู่เข้าจุดต่ำสุดเฉพาะที่โดยวิธีแนวนิวตัน ซึ่งใช้สูตร BFGS แต่ยังไม่มีการเชื่อมโยงกับการหาค่า α_k ในขั้นตอนวิธีค้นหาแนวเส้น (นั่นคือ ใช้ $\alpha_k = 1$) และพวกเขายังได้แสดงว่าลำดับของเมตริกซ์ที่ใช้ประมาณเมตริกซ์เฮซเซียน $\{H_k\}$ จะลู่เข้าหาเมตริกซ์เฮซเซียนที่จุดต่ำสุด $\nabla^2 f(x^*)$ ในทางปฏิบัติพวกเขาพบว่าเมื่อเปรียบเทียบกับวิธีเซแคนต์สมมาตร (symmetric secant method) ที่ได้ค้นพบมาก่อนหน้าจะพบว่าขั้นตอนการทำงานช่วงกลางของการทำซ้ำ วิธี BFGS จะให้การประมาณเมตริกซ์เฮซเซียนได้ถูกต้องมากกว่าวิธีการประมาณจากวิธีเซแคนต์สมมาตรเมื่อเทียบกับ $\nabla^2 f(x^*)$ ดังนั้นวิธี BFGS ในกรณี $\alpha_k = 1$ นี้จะลู่เข้าสู่ x^* ได้เร็วกว่าวิธีเซแคนต์สมมาตร

นิยาม 2.25 ให้ $m,n > 0, G: R^n \to R^{m\times n}, x \in R^n$, ให้ $\|\bullet\|$ เป็นนอร์มบน R^n , และ $\|\bullet\|$ เป็นนอร์มบน $R^{m\times n}$ จะเรียก G ว่าเป็นฟังก์ชันต่อเนื่องลิปชิตซ์ (Lipschitz continuous) ที่ **x** ถ้ามีเซตเปิด $D \subset R^n, x \in D$ และค่าคงตัว γ ที่ทำให้สำหรับทุก ๆ $v \in D$ แล้ว

$$\|G(v) - G(x)\| \le \gamma \|v - x\| \tag{2.31}$$

เรียกค่าคงตัว γ ว่าค่าคงตัวลิปชิตซ์ (Lipschitz constant) สำหรับ G ที่ **x** ถ้า (2.31) เป็นจริงสำหรับทุก ๆ $x \in D$ แล้ว $G \in Lip_\gamma(D)$

ทฤษฎีบทที่ 2.26 (Broyden-Dennis-Mor'e)

ให้ $f: R^n \to R^1$ เป็นฟังก์ชันที่หาอนุพันธ์ได้ต่อเนื่องสองครั้งในเซตนูนเปิด $D \subset R^n$ และให้ $\nabla^2 f \in Lip_\gamma(D)$ สมมติว่ามี $x^* \in D$ ซึ่ง $\nabla f(x^*) = 0$, $\nabla^2 f(x^*)$ เป็นเมตริกซ์ที่มิใช่เอกฐาน (nonsingular) และเป็นบวกแน่นอน และ H_0 เป็นเมตริกซ์สมมาตร แล้วจะมีค่าคงตัวบวก ε, δ ซึ่งถ้า $\|x_0 - x^*\|_2 \le \varepsilon$ และ $\|H_0 - \nabla^2 f(x^*)\|_2 \le \delta$ แล้ว วิธี BFGS ซึ่งกำเนิด $\{x_k\}$ โดยที่

$$x_{k+1} = x_k - B_k^{-1}\nabla f(x_k) \tag{2.32}$$

$$s_k = x_{k+1} - x_k \ , \ y_k = \nabla f(x_{k+1}) - \nabla f(x_k) \tag{2.33}$$

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} \tag{2.34}$$

จะแจ่มชัด (well-defined) , $\{x_k\}$ ยังคงอยู่ใน D และลู่เข้าสู่จุด x^* แบบเหนือเชิงเส้น (superlinearly)

พิสูจน์ ดูเอกสารอ้างอิง [11]

ในลำดับถัดมาได้มีการพัฒนาสูตรการกำเนิด B_{k+1} ซึ่งยังรักษาความเป็นเมตริกซ์ที่เป็นบวกแน่นอน โดยกำหนดสูตร

$$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} + \Phi(s_k^T B_k s_k)\, u_k u_k^T \qquad (2.35)$$

โดยที่ Φ เป็นสเกลาร์ และ

$$u_k = \frac{y_k}{y_k^T s_k} - \frac{B_k s_k}{s_k^T B_k s_k} \qquad (2.36)$$

ในบางครั้งเรียกสูตรการกำเนิดใน (2.35) และ (2.36) ว่าเป็นกลุ่มของบรอยเดน(Broyden class)จะเห็นว่าสูตร BFGS ได้โดยการให้ $\Phi = 0$ และสูตร BFGS จะรักษาความเป็นบวกแน่นอนของเมตริกซ์ ก็ต่อเมื่อ $y_k^T s_k > 0$

ในปี ค.ศ. 1987 บริด , โนเคดอล และหยวน (Richard H. Byrd , Jorge Nocedal และ Ya-Xiang Yuan) ได้เสนองานวิจัยเกี่ยวกับการลู่เข้าแบบวงกว้างของวิธีแนวนิวตันบนปัญหานูน "Global convergence of a class of quasi-Newton methods on convex problems" [12] โดยพวกเขาสามารถเชื่อมโยงขั้นตอนวิธีในการค้นหาแนวเส้น α_k ที่สอดคล้องกับเกณฑ์ 3 และ 4 ((2.8) และ (2.10)) และ สูตร (2.35) ได้เป็นผลสำเร็จเมื่อประยุกต์กับฟังก์ชันจุดประสงค์ที่เป็นฟังก์ชันนูนโดยแท้

ทฤษฎีบทที่ 2.27 ให้ f เป็นฟังก์ชันค่าจริงของ n ตัวแปร ให้ x_0 เป็นจุดเริ่มต้นที่กำหนดและให้ $\{x_k\}$ นิยามโดย $x_{k+1} = x_k + \alpha_k p_k$ ซึ่ง p_k เป็นเวกเตอร์ที่มีมิติ n และ $\alpha_k \geq 0$ เป็นสเกลาร์ สมมติว่า

(1). เซต $S = \{x \mid f(x) \leq f(x_0)\}$ เป็นเซตมีขอบเขต

(2). f , ∇f และ $\nabla^2 f$ เป็นฟังก์ชันต่อเนื่องทุกจุด $x \in S$

(3). $\nabla^2 f(x)$ เป็นเมตริกซ์ที่เป็นบวกแน่นอน สำหรับทุก ๆ x

(4). ทิศทางค้นหา $\{p_k\}$ คำนวณโดย

$$B_k p_k = -\nabla f(x_k)$$

ซึ่ง $B_0 = I$ และเมตริกซ์ $\{B_k\}$ สามารถกำเนิดโดยใช้สูตรเวียนบังเกิดจาก (2.35) ด้วยพารามิเตอร์ $0 \le \Phi < 1$

(5). ความยาวในการเดินทาง $\{\alpha_k\}$ สอดคล้องกับเงื่อนไข

$$\begin{aligned} f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) &\le f(\mathbf{x}_k) + \mu\alpha_k \mathbf{p}_k^T \nabla f(\mathbf{x}_k) \\ \mathbf{p}_k^T \nabla f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) &\ge \eta \mathbf{p}_k^T \nabla f(\mathbf{x}_k) \end{aligned}$$

ซึ่ง $0 < \mu < \eta < 1$ และขั้นตอนวิธีค้นหาแนวเส้นจะใช้ $\alpha_k = 1$ ได้เมื่อสอดคล้องกับเงื่อนไขทั้งสอง แล้ว $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}^*$ ซึ่ง $\mathbf{x}^*$ เป็นจุดต่ำสุดวงกว้างเพียงจุดเดียวของ f บน S และอัตราการลู่เข้าของ $\{\mathbf{x}_k\}$ เป็นแบบเหนือเชิงเส้น

พิสูจน์ ดูเอกสารอ้างอิง [12]

ในปี ค.ศ. 1968 ฟีแอคโคและแม็คคอร์มิค (Anthony V. Fiacco และ Garth P. Mc Cormick) [10] ได้เสนองานวิจัยในการใช้วิธีฟังก์ชันกั้นเขตกับปัญหาที่เรียกว่า ปัญหากำหนดการนูน (convex programming problem) โดยเฉพาะ ซึ่งจะได้อธิบายและยกทฤษฎีที่สำคัญมากล่าวอ้างดังนี้

นิยาม 2.28 ปัญหากำหนดการนูน (C) เขียนได้เป็น

$$\begin{aligned} &\text{minimize} && f(\mathbf{x}) \\ &\text{subject to} && g_i(\mathbf{x}) \le 0 \quad , \ i = 1,\ldots,m \end{aligned}$$

ซึ่ง $f(\mathbf{x})$, $\{g_i(\mathbf{x})\}$ เป็นฟังก์ชันต่อเนื่องแบบนูน

จะเห็นว่าเราสามารถใช้ทฤษฎีบทที่ 2.1 กับปัญหา (C) ได้เพราะ $\{\mathbf{x} \mid g_i(\mathbf{x}) \le 0$, $i = 1,\ldots,m\}$ เป็นเซตนูน ดังนั้นจุดต่ำสุดเฉพาะที่ของปัญหา (C) ก็คือจุดต่ำสุดวงกว้างนั่นเอง

เพื่อให้เป็นไปตามการพิสูจน์ในเอกสารอ้างอิง [10] จะทำการกำหนดให้ $q_i = -g_i$, $i = 1,\ldots,m$ และ $\mathbf{q} = (q_1,\ldots,q_m)^T$ และให้

$$R = \{\mathbf{x} \mid \mathbf{q}(\mathbf{x}) \ge 0\} \ , \ R^0 = \{\mathbf{x} \mid \mathbf{q}(\mathbf{x}) > 0\} \tag{2.37}$$

ดังนั้นฟังก์ชันกั้นเขต $\beta(g) = -\sum_{i=1}^{m}\frac{1}{g_i}$ และ $\beta(g) = -\sum_{i=1}^{m}\ln(-g_i)$ สามารถแทนได้โดยฟังก์ชัน I ของ q โดยที่

$$I(q) = \sum_{i=1}^{m}\frac{1}{q_i} \quad \text{และ} \quad I(q) = -\sum_{i=1}^{m}\ln(q_i) \tag{2.38}$$

จะเห็นได้ง่ายว่า $\beta(g) = I(q)$ เมื่อเป็นเช่นนี้สามารถนิยามปัญหา (C) ได้ใหม่เป็น

$$\left.\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & q_i(x) \geq 0 \quad , i = 1,\ldots,m \end{array}\right\} \tag{2.39}$$

เมื่อ $f(x)$, $\{-q_i(x)\}$ เป็นฟังก์ชันต่อเนื่องแบบนูน

บทตั้งที่ 2.29 ถ้า I เป็นฟังก์ชันลดลงแบบนูน (convex decreasing function) ของ $\mathbf{q} = (q_1,\ldots,q_m)^T$ สำหรับ $\mathbf{q} > \mathbf{0}$ ซึ่งแต่ละ q_i เป็นฟังก์ชันเว้า แล้ว $I[\mathbf{q}(\mathbf{x})]$ เป็นฟังก์ชันนูนของ $\mathbf{x}$ ใน R^0 ที่กำหนดโดย (2.37) และถ้า f เป็นฟังก์ชันนูนของ $\mathbf{x}$ แล้ว $U(\mathbf{x},\mu_k) = f(\mathbf{x}) + \mu_k I[\mathbf{q}(\mathbf{x})]$ จะเป็นฟังก์ชันนูนของ $\mathbf{x}$ ใน R^0 ด้วย เมื่อ $\mu_k > 0$

พิสูจน์ ดูเอกสารอ้างอิง [10]

เห็นได้ชัดว่า $I(\mathbf{q})$ ที่กำหนดโดย (2.38) เป็นฟังก์ชันลดลงแบบนูนใน $\mathbf{q}$ เมื่อ $\mathbf{q} > \mathbf{0}$ ต่อไปนี้จะให้เงื่อนไขสำหรับการมี $\mathbf{x}(\mu_k)$ ที่ทำให้ $U(\mathbf{x},\mu_k)$ ต่ำสุดสำหรับปัญหา (C)

บทตั้งที่ 2.30 ถ้าเซตของจุดที่แก้ปัญหา (C) (2.39) เป็นเซตมีขอบเขต (ดังนั้นเป็นเซตกระชับ) , $R^0 \neq \phi$ และ $I(\mathbf{q})$ มีคุณสมบัติว่า

(1). $\lim_{l\to\infty} I(\mathbf{q}^l) = +\infty$ สำหรับทุก ๆ ลำดับอนันต์ $\{\mathbf{q}^l\}$ ซึ่ง $\mathbf{q}^l > \mathbf{0}$ สำหรับทุก ๆ l และ $\lim_{l\to\infty} q_i^l = 0$ สำหรับบาง i

(2). I เป็นฟังก์ชันลดลงใน $\mathbf{q}$ เมื่อ $\mathbf{q} > \mathbf{0}$

(3). เมื่อ $\mathbf{q}^0 > \mathbf{0}$, $\mathbf{d}^0 \geq \mathbf{0}$ และ $\{\lambda^l\}$ เป็นลำดับที่เป็นบวกโดยแท้ ซึ่งลู่เข้าหาศูนย์โดยที่

$$\lim_{l\to\infty} \lambda^l I\left[\mathbf{q}^0 + \frac{d^0}{\lambda^l}\right] = 0 \qquad (2.40)$$

แล้วใน R^0 , สำหรับทุก ๆ $\mu_k > 0$ จะได้ว่าเส้นชั้นความสูงที่เท่ากัน (isocontour) ของ $U(x,\mu_k) = f(x) + \mu_k I[q(x)]$ เป็นเซตที่มีขอบเขต

พิสูจน์ ดูเอกสารอ้างอิง [10]

จากบทตั้งที่ 2.30 จะได้สิ่งที่ตามมาโดยปริยายคือ ทุก ๆ จุดต่ำสุดของ $U(x,\mu_k)$ ใน R^0 จะต้องเป็นจุดจำกัด (finite point) ฟีแอคโคและแม็คคอร์มิคยังกล่าวว่า ฟังก์ชัน I ที่กำหนดโดย (2.38) จะสอดคล้องกับ (2.40) ดังนั้นจึงมั่นใจได้ว่าสำหรับการหาค่าต่ำสุดของ $U(x,\mu_k)$ จะต้องเกิด $x(\mu_k)$ ซึ่งเป็นจุดต่ำสุดและเป็นจุดจำกัด

ทฤษฎีบทที่จะกล่าวถึงต่อไปนี้ เป็นทฤษฎีของการลู่เข้าหาผลเฉลยของปัญหา (C) โดยวิธีฟังก์ชันกั้นเขต

ทฤษฎีบทที่ 2.31 ถ้า f และ $-\{q_i\}$ $(i = 1,\ldots,m)$ เป็นฟังก์ชันนูน , เซตของจุดที่แก้ปัญหา (C) เป็นเซตมีขอบเขต (ดังนั้นจะเป็นเซตกระชับ) , R^0 ที่กำหนดโดย (2.37) ไม่เป็นเซตว่างและกำหนดฟังก์ชัน I ของ **q** ให้มีคุณสมบัติดังนี้

1. $I(\mathbf{q})$ เป็นฟังก์ชันลดลงแบบนูนใน **q** เมื่อ $\mathbf{q} > 0$ และ $\lim_{l\to\infty} \lambda^l I\left[\mathbf{q}^0 + \frac{d^0}{\lambda^l}\right] = 0$ เมื่อ $\mathbf{q}^0 > 0$, $d^0 \geq 0$ และสำหรับลำดับที่เป็นบวกโดยแท้ซึ่งลู่เข้าศูนย์ $\{\lambda^l\}$

2. สำหรับลำดับที่เป็นบวกโดยแท้ $\{\mathbf{q}^l\}$ ใด ๆ ซึ่ง $\lim_{l\to\infty} I\{\mathbf{q}^l\} = +\infty$ ถ้า $\lim_{l\to\infty} q_i^l = 0$ สำหรับบาง i และให้ $\{\mu_k\}$ เป็นลำดับลดลงโดยแท้ซึ่งลู่เข้าศูนย์ แล้ว

(1). ฟังก์ชัน $U(x,\mu_k) = f(x) + \mu_k I[q(x)]$ มีจุดต่ำสุดแบบไม่มีเงื่อนไขบังคับซึ่งเป็นจุดจำกัด $x(\mu_k) \in R^0$ สำหรับทุก ๆ $\mu_k > 0$

(2). $U(x,\mu_k)$ เป็นฟังก์ชันนูนใน R^0

(3). ทุก ๆ จุดต่ำสุดแบบไม่มีเงื่อนไขบังคับเฉพาะที่ของ $U(x,\mu_k)$ ใน R^0 จะเป็นจุดต่ำสุดแบบไม่มีเงื่อนไขบังคับวงกว้างของ $U(x,\mu_k)$ ใน R^0

(4). $\lim_{k\to\infty} f[x(\mu_k)] = \lim_{k\to\infty} U[x(\mu_k),\mu_k] = v^* = \min_R f(x)$

(5). ทุก ๆ จุดลิมิตของ $\{x(\mu_k)\}$ จะแก้ปัญหา (C) ได้

(6). $\{U[x(\mu_k),\mu_k]\}$ เป็นลำดับลดลงทางเดียว ถ้า $I(\mathbf{q}) > 0$ สำหรับ $\mathbf{q} > \mathbf{0}$

(7). $\{f[x(\mu_k)]\}$ เป็นลำดับลดลงทางเดียว

(8). $\{ I[x(\mu_k)]\}$ เป็นลำดับเพิ่มขึ้นทางเดียว

พิสูจน์ ดูเอกสารอ้างอิง [10]

บทที่ 3
ทฤษฎีและขั้นตอนวิธีที่ใช้ในการแก้ปัญหา

ในบทนี้จะกล่าวถึงทฤษฎีบทที่มีความสำคัญต่อขั้นตอนวิธีที่ใช้ในการแก้ปัญหา โดยแสดงการพิสูจน์ทฤษฎีบทที่ไม่ปรากฏอยู่ในตำราที่อ้างอิงและทฤษฎีบทที่ผู้วิจัยพัฒนาขึ้น สำหรับส่วนสุดท้ายในบทจะเสนอขั้นตอนวิธีที่ใช้ในการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบต่าง ๆ ตาม (1.1)-(1.3)

3.1 ทฤษฎีที่มีความสำคัญต่อขั้นตอนวิธีที่ใช้แก้ปัญหา

บทตั้งที่ 3.1 ถ้า B_k ในสูตร BFGS (2.23)

$$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}$$

เป็นเมตริกซ์สมมาตรแล้ว สูตร BFGS (2.23) สอดคล้องกับเงื่อนไขเซแคนต์ $B_{k+1}s_k = y_k$

พิสูจน์ จากสูตร BFGS ใน (2.23)

$$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}$$

เมื่อ s_k และ y_k กำหนดไว้ใน (2.19) และ (2.20) ตามลำดับ

พิจารณา

$$\begin{aligned}
B_{k+1}s_k &= B_k s_k - \frac{(B_k s_k)(B_k s_k)^T s_k}{s_k^T B_k s_k} + \frac{y_k y_k^T s_k}{y_k^T s_k} \\
&= B_k s_k - \frac{(s_k^T B_k^T s_k)(B_k s_k)}{(s_k^T B_k s_k)} + \frac{(y_k^T s_k) y_k}{(y_k^T s_k)} \\
&= B_k s_k - \frac{(s_k^T B_k s_k)(B_k s_k)}{(s_k^T B_k s_k)} + \frac{(y_k^T s_k) y_k}{(y_k^T s_k)} \quad ; \; B_k^T = B_k \\
&= B_k s_k - B_k s_k + y_k \\
&= y_k
\end{aligned}$$

ดังนั้น สูตร BFGS สอดคล้องกับเงื่อนไขเซแคนต์ (2.21) #

ทฤษฎีบทที่จะกล่าวต่อไปนี้มีประโยชน์ในแง่ของการประหยัดการทำงานในการหาทิศทาง $\mathbf{p}_k$ เพราะเดิม $\mathbf{p}_k$ คำนวณจากการแก้ระบบสมการเชิงเส้น $B_k\mathbf{p}_k = -\nabla f(\mathbf{x}_k)$ โดยที่ B_k สามารถกำเนิดจาก B_{k-1} โดยสูตร BFGS (2.23) แต่ถ้าเราพิจารณา

$$\mathbf{p}_k = -[B_k]^{-1}\nabla f(\mathbf{x}_k)$$

ให้ $H_k = B_k^{-1}$ ดังนั้น

$$\mathbf{p}_k = -H_k\nabla f(\mathbf{x}_k)$$

ซึ่งการคำนวณ $\mathbf{p}_k$ จะง่ายขึ้นเพราะเป็นการคูณระหว่างเมตริกซ์กับเวกเตอร์ ดังนั้นถ้าเราสามารถหาสูตรการกำเนิด H_k จาก H_{k-1} โดยที่ $H_k = B_k^{-1}$ และ $H_{k-1} = B_{k-1}^{-1}$ แล้วเราสามารถลดขั้นตอนการแก้ระบบสมการดังกล่าวได้ เพื่อความสะดวกและให้สอดคล้องกับ (2.23) เราจะกำเนิด H_{k+1} จาก H_k โดยที่ $H_{k+1} = B_{k+1}^{-1}$ และ $H_k = B_k^{-1}$ เราเรียกสูตรการกำเนิด H_{k+1} จาก H_k ว่าสูตรการกำเนิด BFGS ผกผัน ซึ่งเอกสารอ้างอิง [13] ได้ให้ตัวสูตร BFGS ผกผันนี้ไว้ โดยปราศจากการพิสูจน์ ดังนั้นเราจะได้ทำการพิสูจน์ในทฤษฎีบทต่อไป

$$H_{k+1} = H_k - \frac{\mathbf{s}_k\mathbf{y}_k^TH_k + H_k\mathbf{y}_k\mathbf{s}_k^T}{\mathbf{y}_k^T\mathbf{s}_k} + \left[1 + \frac{\mathbf{y}_k^TH_k\mathbf{y}_k}{\mathbf{y}_k^T\mathbf{s}_k}\right]\frac{\mathbf{s}_k\mathbf{s}_k^T}{\mathbf{s}_k^T\mathbf{y}_k} \tag{3.1}$$

ทฤษฎีบทที่ 3.2 ถ้า B_k ในสูตร BFGS (2.23)

$$B_{k+1} = B_k - \frac{(B_k\mathbf{s}_k)(B_k\mathbf{s}_k)^T}{\mathbf{s}_k^TB_k\mathbf{s}_k} + \frac{\mathbf{y}_k\mathbf{y}_k^T}{\mathbf{y}_k^T\mathbf{s}_k}$$

เป็นเมตริกซ์สมมาตรและ $B_k^{-1} = H_k$ แล้ว $B_{k+1}^{-1} = H_{k+1}$ ใน (3.1)

พิสูจน์ ใช้สูตรของเชอร์แมน-มอร์ริสัน-วู๊ดเบอรี (Sherman-Morrison-Woodbury formula) ดังนี้

$$[A + \mathbf{ab}^T]^{-1} = A^{-1} - \frac{A^{-1}\mathbf{ab}^TA^{-1}}{1 + \mathbf{b}^TA^{-1}\mathbf{a}} \tag{3.2}$$

เราประยุกต์สูตร (3.2) สองครั้งเพื่อแสดงว่า $B_{k+1}^{-1} = H_{k+1}$

จากสูตร BFGS ให้ $C = B_k + \dfrac{y_k y_k^T}{y_k^T s_k}$

ดังนั้น
$$C^{-1} = \left[B_k + \frac{y_k y_k^T}{y_k^T s_k}\right]^{-1} = B_k^{-1} - \frac{\dfrac{B_k^{-1} y_k y_k^T B_k^{-1}}{y_k^T s_k}}{1 + \dfrac{y_k^T B_k^{-1} y_k}{y_k^T s_k}}$$

$$= B_k^{-1} - \frac{B_k^{-1} y_k y_k^T B_k^{-1}}{y_k^T s_k + y_k^T B_k^{-1} y_k} \tag{3.3}$$

พิจารณา $B_{k+1} = C + \dfrac{(B_k s_k)(B_k s_k)^T}{-s_k^T B_k s_k}$ ดังนั้น ประยุกต์สูตร (3.2) อีกครั้ง

$$B_{k+1}^{-1} = \left[C + \frac{(B_k s_k)(B_k s_k)^T}{-s_k^T B_k s_k}\right]^{-1}$$

$$= C^{-1} - \left[\frac{\left(\dfrac{C^{-1}(B_k s_k)(B_k s_k)^T C^{-1}}{-s_k^T B_k s_k}\right)}{\left(1 + \dfrac{(B_k s_k)^T C^{-1}(B_k s_k)}{-s_k^T B_k s_k}\right)}\right]$$

$$= C^{-1} - \frac{C^{-1}(B_k s_k)(B_k s_k)^T C^{-1}}{[-s_k^T B_k s_k + (B_k s_k)^T C^{-1}(B_k s_k)]}$$

$$= C^{-1} + \frac{C^{-1}(B_k s_k)(B_k s_k)^T C^{-1}}{[s_k^T B_k s_k - (B_k s_k)^T C^{-1}(B_k s_k)]} \tag{3.4}$$

พิจารณา $C^{-1}(B_k s_k)(B_k s_k)^T C^{-1}$ ใน (3.4) จะได้ว่า

$$C^{-1}(B_k s_k)(B_k s_k)^T C^{-1} = \left[B_k^{-1} - \frac{B_k^{-1} y_k y_k^T B_k^{-1}}{y_k^T s_k + y_k^T B_k^{-1} y_k}\right](B_k s_k)(B_k s_k)^T \left[B_k^{-1} - \frac{B_k^{-1} y_k y_k^T B_k^{-1}}{y_k^T s_k + y_k^T B_k^{-1} y_k}\right]$$

$$= \left[B_k^{-1}(B_k s_k)(B_k s_k)^T - \frac{B_k^{-1} y_k y_k^T B_k^{-1}(B_k s_k)(B_k s_k)^T}{y_k^T s_k + y_k^T B_k^{-1} y_k}\right]\left[B_k^{-1} - \frac{B_k^{-1} y_k y_k^T B_k^{-1}}{y_k^T s_k + y_k^T B_k^{-1} y_k}\right]$$

$$= \left[s_k (B_k s_k)^T - \frac{B_k^{-1} y_k y_k^T s_k (B_k s_k)^T}{y_k^T s_k + y_k^T B_k^{-1} y_k}\right]\left[B_k^{-1} - \frac{B_k^{-1} y_k y_k^T B_k^{-1}}{y_k^T s_k + y_k^T B_k^{-1} y_k}\right]$$

$$= s_k s_k^T B_k^T B_k^{-1} - \frac{s_k s_k^T B_k^T B_k^{-1} y_k y_k^T B_k^{-1}}{y_k^T s_k + y_k^T B_k^{-1} y_k} - \frac{B_k^{-1} y_k y_k^T s_k s_k^T B_k^T B_k^{-1}}{y_k^T s_k + y_k^T B_k^{-1} y_k} + \frac{B_k^{-1} y_k y_k^T s_k s_k^T B_k^T B_k^{-1} y_k y_k^T B_k^{-1}}{(y_k^T s_k + y_k^T B_k^{-1} y_k)^2}$$

$$=s_k s_k^T - \frac{(s_k^T y_k) s_k y_k^T B_k^{-1}}{(y_k^T s_k + y_k^T B_k^{-1} y_k)} - \frac{(y_k^T s_k) B_k^{-1} y_k s_k^T}{(y_k^T s_k + y_k^T B_k^{-1} y_k)} + \frac{(y_k^T s_k s_k^T y_k) B_k^{-1} y_k y_k^T B_k^{-1}}{(y_k^T s_k + y_k^T B_k^{-1} y_k)^2} ; B_k^T = B_k$$

$$=s_k s_k^T - \frac{(y_k^T s_k) s_k y_k^T B_k^{-1}}{(y_k^T s_k + y_k^T B_k^{-1} y_k)} - \frac{(y_k^T s_k) B_k^{-1} y_k s_k^T}{(y_k^T s_k + y_k^T B_k^{-1} y_k)} + \frac{(y_k^T s_k)^2 B_k^{-1} y_k y_k^T B_k^{-1}}{(y_k^T s_k + y_k^T B_k^{-1} y_k)^2} \; ; y_k^T s_k = (y_k^T s_k)^T = s_k^T y_k \tag{3.5}$$

พิจารณาเทอม $s_k^T B_k s_k - (B_k s_k)^T C^{-1}(B_k s_k)$ ใน (3.4) จะได้ว่า

$$\begin{aligned} s_k^T B_k s_k - (B_k s_k)^T C^{-1}(B_k s_k) &= s_k^T B_k s_k - (B_k s_k)^T \left[B_k^{-1} - \frac{B_k^{-1} y_k y_k^T B_k^{-1}}{y_k^T s_k + y_k^T B_k^{-1} y_k} \right] (B_k s_k) \\ &= s_k^T B_k s_k - s_k^T B_k^T B_k^{-1} B_k s_k + \frac{s_k^T B_k^T B_k^{-1} y_k y_k^T B_k^{-1} B_k s_k}{y_k^T s_k + y_k^T B_k^{-1} y_k} \\ &= s_k^T B_k s_k - s_k^T B_k s_k + \frac{s_k^T y_k y_k^T s_k}{y_k^T s_k + y_k^T B_k^{-1} y_k} \\ &= \frac{(y_k^T s_k)^2}{(y_k^T s_k + y_k^T B_k^{-1} y_k)} \end{aligned} \tag{3.6}$$

ดังนั้น (3.5)/(3.6) จะได้

$$\begin{aligned} \frac{C^{-1}(B_k s_k)(B_k s_k)^T C^{-1}}{s_k^T B_k s_k - (B_k s_k)^T C^{-1}(B_k s_k)} &= \left[\frac{(y_k^T s_k + y_k^T B_k^{-1} y_k)}{(y_k^T s_k)^2} \right] \left[s_k s_k^T - \frac{(y_k^T s_k) s_k y_k^T B_k^{-1}}{(y_k^T s_k + y_k^T B_k^{-1} y_k)} - \right. \\ &\qquad \left. \frac{(y_k^T s_k) B_k^{-1} y_k s_k^T}{(y_k^T s_k + y_k^T B_k^{-1} y_k)} + \frac{(y_k^T s_k)^2 B_k^{-1} y_k y_k^T B_k^{-1}}{(y_k^T s_k + y_k^T B_k^{-1} y_k)^2} \right] \\ &= \frac{(y_k^T s_k + y_k^T B_k^{-1} y_k) s_k s_k^T}{(y_k^T s_k)^2} - \frac{s_k y_k^T B_k^{-1}}{(y_k^T s_k)} - \frac{B_k^{-1} y_k s_k^T}{(y_k^T s_k)} + \\ &\qquad \frac{B_k^{-1} y_k y_k^T B_k^{-1}}{(y_k^T s_k + y_k^T B_k^{-1} y_k)} \end{aligned} \tag{3.7}$$

แทน (3.7) ลงใน (3.4) จะได้ว่า

$$\begin{aligned} B_{k+1}^{-1} &= B_k^{-1} - \frac{B_k^{-1} y_k y_k^T B_k^{-1}}{y_k^T s_k + y_k^T B_k^{-1} y_k} + \frac{(y_k^T s_k + y_k^T B_k^{-1} y_k) s_k s_k^T}{(y_k^T s_k)^2} - \frac{s_k y_k^T B_k^{-1}}{(y_k^T s_k)} - \\ &\qquad \frac{B_k^{-1} y_k s_k^T}{(y_k^T s_k)} + \frac{B_k^{-1} y_k y_k^T B_k^{-1}}{(y_k^T s_k + y_k^T B_k^{-1} y_k)} \\ &= B_k^{-1} - \frac{s_k y_k^T B_k^{-1} + B_k^{-1} y_k s_k^T}{y_k^T s_k} + \left[\frac{y_k^T s_k + y_k^T B_k^{-1} y_k}{y_k^T s_k} \right] \frac{s_k s_k^T}{s_k^T y_k} \end{aligned}$$

$$= H_k - \frac{s_k y_k^T H_k + H_k y_k s_k^T}{y_k^T s_k} + \left[1 + \frac{y_k^T H_k y_k}{y_k^T s_k}\right] \frac{s_k s_k^T}{s_k^T y_k} \qquad ; B_k^{-1} = H_k$$

$$= H_{k+1}$$ #

จากบทตั้งที่ 3.1 แสดงว่าสูตร BFGS ใน (2.23) ทำให้เงื่อนไขเซแคนต์ $B_{k+1}s_k = y_k$ เป็นจริง จากสูตร BFGS ผกผันใน (3.1) จะทำให้ $B_{k+1}^{-1}y_k = H_{k+1}y_k = s_k$ นั่นคือสอดคล้องกับเงื่อนไขเซแคนต์เช่นกัน

บทตั้งที่ 3.3 ถ้า A เป็นเมตริกซ์สมมาตรขนาด $n \times n$ และเป็นบวกแน่นอนแล้ว A^{-1} จะเป็นเมตริกซ์สมมาตรและเป็นบวกแน่นอน

พิสูจน์ พิจารณา $(A^{-1})^T = (A^T)^{-1}$

$$= (A)^{-1} \qquad ; A = A^T \text{ เพราะ } A \text{ เป็นเมตริกซ์สมมาตร}$$

ดังนั้น A^{-1} เป็นเมตริกซ์สมมาตร จากเอกสารอ้างอิง [8] กล่าวว่า เมื่อ A เป็นเมตริกซ์สมมาตรแล้วค่าเฉพาะทั้งหมดของ A เป็นจำนวนจริงและสามารถเขียน $P^TAP = D$ ได้ เมื่อ P เป็นเมตริกซ์ขนาด $n \times n$ และ $P^TP = I$ หรือ $PP^T = I$ และ D เป็นเมตริกซ์ทแยงมุมซึ่งมีสมาชิกในแนวทแยงมุมหลักเป็นค่าเฉพาะของ A ซึ่งได้แก่ $\lambda_1,\ldots,\lambda_n$ และเนื่องจาก A เป็นเมตริกซ์ที่เป็นบวกแน่นอน ดังนั้น $\lambda_1,\ldots,\lambda_n$ เป็นจำนวนจริงบวกทั้งหมด (จากทฤษฎีบทที่ 2.6)

จาก $$P^TAP = D$$

ดังนั้น
$$\begin{aligned} A &= P^{-T}DP^{-1} \\ &= (P^T)^{-1}DP^T && ; P^{-1} = P^T \\ &= (P^{-1})^{-1}DP^T && ; P^T = P^{-1} \\ &= PDP^T && (3.8) \end{aligned}$$

พิจารณา
$$\begin{aligned} A(PD^{-1}P^T) &= PDP^T(PD^{-1}P^T) \\ &= PDD^{-1}P^T && ; P^TP = I \\ &= PP^T \\ &= I && ; PP^T = I \end{aligned}$$

จะเห็นว่า $$A^{-1} = PD^{-1}P^T \qquad (3.9)$$

เนื่องจาก A^{-1} เป็นเมตริกซ์สมมาตร ดังนั้นเราสามารถแยก A^{-1} ได้ดัง (3.9) เพราะว่า

$PP^T = I$ หรือ $P^TP = I$ และ D^{-1} เป็นเมตริกซ์ทแยงมุม ซึ่งสมาชิกในแนวทแยงมุมหลักของ D^{-1} ก็คือ ค่าเฉพาะของ A^{-1} และมีค่าเป็นส่วนกลับของค่าเฉพาะของ A ในตำแหน่งที่สมนัยกัน ซึ่งได้แก่ $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}$ จะเห็นว่ามีค่าเป็นบวกทั้งหมด ดังนั้น A^{-1} จึงเป็นเมตริกซ์ที่เป็นบวกแน่นอนตามทฤษฎีบทที่ 2.6 #

จากบทตั้งที่ 3.3 จะได้ว่าสูตร BFGS ผกผัน (3.1) นั้นจะกำเนิด $H_{k+1} = B_{k+1}^{-1}$ เป็นเมตริกซ์สมมาตรและเป็นบวกแน่นอน เนื่องจาก B_{k+1} ที่ได้จากสูตร BFGS (2.23) เป็นเมตริกซ์สมมาตรที่เป็นบวกแน่นอน

บทตั้งที่ 3.4 ให้ H_k เป็นเมตริกซ์ที่เป็นบวกแน่นอนและสมมาตร สมมติว่า H_{k+1} ได้จาก H_k โดยใช้สูตร BFGS ผกผัน (3.1) แล้ว H_{k+1} เป็นเมตริกซ์ที่เป็นบวกแน่นอนก็ต่อเมื่อ $y_k^T s_k > 0$

พิสูจน์ เมื่อ H_{k+1} ในสูตร BFGS ผกผันเป็นเมตริกซ์ที่เป็นบวกแน่นอน ดังนั้น $B_{k+1} = H_{k+1}^{-1}$ เป็นเมตริกซ์ที่เป็นบวกแน่นอน (ตามบทตั้งที่ 3.3) ทำให้ได้ว่า $y_k^T s_k > 0$ ตามบทตั้งที่ 2.18 (เพราะ B_{k+1} กำเนิดจาก $B_k = H_k^{-1}$ ซึ่งเป็นเมตริกซ์สมมาตรและเป็นบวกแน่นอน)

ในทางกลับกัน ถ้า $y_k^T s_k > 0$ และจากที่กำหนดว่า H_k เป็นเมตริกซ์สมมาตรที่เป็นบวกแน่นอน ดังนั้น $B_k = H_k^{-1}$ จะเป็นเมตริกซ์สมมาตรที่เป็นบวกแน่นอน เมื่อกำเนิด B_{k+1} จาก B_k โดยสูตร BFGS จะได้ว่า B_{k+1} เป็นเมตริกซ์สมมาตรและเป็นบวกแน่นอน ดังนั้น $H_{k+1} = B_{k+1}^{-1}$ เป็นเมตริกซ์สมมาตรและเป็นบวกแน่นอน (ตามบทตั้งที่ 3.3) #

บทตั้งที่ 3.5 ถ้า G_i $(i = 1,\ldots,m)$ เป็นเซตนูนแล้ว $G = \bigcap_{i=1}^{m} G_i$ เป็นเซตนูน

พิสูจน์ ดูเอกสารอ้างอิง [10]

บทตั้งที่ 3.6 ถ้า $g(x)$ เป็นฟังก์ชันนูนในเซตนูน G แล้ว สำหรับค่าคงที่ k ใด ๆ จะได้ว่า $S_k = \{x \mid g(x) < k, x \in G\}$ และ $\overline{S}_k = \{x \mid g(x) \le k, x \in G\}$ เป็นเซตนูน

พิสูจน์ ดูเอกสารอ้างอิง [10]

บทตั้งที่ 3.7 ถ้า $g_i(x)$ เป็นฟังก์ชันนูนในเซตนูน G_i สำหรับ $i = 1,\ldots,m$ แล้ว $G_0 = \{x \mid g_i(x) < 0, i = 1,\ldots,m$ และ $x \in \bigcap_{i=1}^{m} G_i\}$ และ $G = \{x \mid g_i(x) \le 0, i = 1,\ldots,m$ และ

$\mathbf{x} \in \bigcap_{i=1}^{m} G_i\}$ เป็นเซตนูน

พิสูจน์ ให้ $S_{0i} = \{\mathbf{x} \mid g_i(\mathbf{x}) < 0,\ \mathbf{x} \in G_i\}$ และ $S_i = \{\mathbf{x} \mid g_i(\mathbf{x}) \le 0,\ \mathbf{x} \in G_i\}$ สำหรับ $i = 1,\ldots,m$ จะได้ว่า S_{0i} และ S_i เป็นเซตนูนตามบทตั้งที่ 3.6 ดังนั้น $G_0 = \bigcap_{i=1}^{m} S_{0i}$ และ $G = \bigcap_{i=1}^{m} S_i$ จะเป็นเซตนูนตามบทตั้งที่ 3.5 #

ตามทฤษฎีบทที่ 2.29 ได้กล่าวว่า เมื่อ $f(\mathbf{x})$ เป็นฟังก์ชันนูนของ $\mathbf{x}$ แล้ว ฟังก์ชันช่วย $U(\mathbf{x}, \mu_k) = f(\mathbf{x}) + \mu_k \beta(\mathbf{x})$ ที่ได้จากการแปลงด้วยวิธีฟังก์ชันกั้นเขตจะเป็นฟังก์ชันนูน แต่เนื่องจากเป็นการกล่าวตามเอกสารอ้างอิง [10] ซึ่งมีการกำหนดลักษณะของฟังก์ชันเงื่อนไขบังคับเป็น $g_i(\mathbf{x}) \ge 0$ ซึ่งกลับกันกับที่กล่าวในวิทยานิพนธ์นี้ ดังนั้นในตอนนี้ผู้วิจัยจะพิสูจน์ว่าการกำหนดลักษณะฟังก์ชันเงื่อนไขบังคับ $g_i(\mathbf{x}) \le 0$ ในวิทยานิพนธ์นี้ก็ให้ผลลัพธ์เช่นเดียวกัน

บทตั้งที่ 3.8 ให้ β เป็นฟังก์ชันของ $\mathbf{g} = (g_1, g_2, \ldots, g_m)^T$ สำหรับ $\mathbf{g} < \mathbf{0}$ ซึ่งแต่ละ g_i เป็นฟังก์ชันนูน แล้ว $\beta(\mathbf{g}) = -\sum_{i=1}^{m} \frac{1}{g_i}$ เป็นฟังก์ชันลดลงแบบนูน

พิสูจน์ กำหนดให้ $\mathbf{g} = (g_1, g_2, \ldots, g_m)^T$ ซึ่ง $\mathbf{g} < \mathbf{0}$ และ $\mathbf{h} = (h_1, h_2, \ldots, h_m)^T$ ซึ่ง $\mathbf{h} < \mathbf{0}$

สมมติว่า
$$\mathbf{g} \le \mathbf{h}$$

ดังนั้น
$$g_i \le h_i \qquad ; i = 1,\ldots,m$$
$$-g_i \ge -h_i > 0 \qquad ; i = 1,\ldots,m$$
$$0 < -\frac{1}{g_i} \le -\frac{1}{h_i} \qquad ; i = 1,\ldots,m$$

จะได้ว่า
$$0 < \sum_{i=1}^{m} -\frac{1}{g_i} \le \sum_{i=1}^{m} -\frac{1}{h_i}$$

และ
$$\beta(\mathbf{g}) \le \beta(\mathbf{h})$$

ดังนั้น $\beta(\mathbf{g})$ เป็นฟังก์ชันเพิ่ม

ต่อไปจะทำการพิสูจน์ว่า $\beta(\mathbf{g})$ เป็นฟังก์ชันนูน

พิจารณา
$$\beta_i(g_i) = -\frac{1}{g_i} \qquad ; i = 1,\ldots,m$$

ดังนั้น
$$\beta_i'(g_i) = (g_i)^{-2} \qquad ; i = 1,\ldots,m$$

และ
$$\beta_i''(g_i) = -2(g_i)^{-3} \qquad ; i = 1,\ldots,m$$

จากที่กำหนดว่า $\mathbf{g} < \mathbf{0}$ ดังนั้น $g_i < 0, \forall i$ ทำให้ $\beta_i''(g_i) > 0, \forall i$

จึงได้ว่า $\beta_i(g_i)$ เป็นฟังก์ชันนูนสำหรับทุก ๆ i

เมื่อพิจารณา
$$\sum_{i=1}^{m}\beta_i(g_i) = -\frac{1}{g_1}-\frac{1}{g_2}-\dots-\frac{1}{g_m}$$
$$= -\sum_{i=1}^{m}\frac{1}{g_i} \quad \text{; เป็นฟังก์ชันนูนตามทฤษฎีบทที่ 2.2}$$
$$= \beta(g)$$

ดังนั้น $\beta(g)$ เป็นฟังก์ชันนูน #

บทตั้งที่ 3.9 กำหนดให้ $\beta(g) = -\sum_{i=1}^{m}\frac{1}{g_i}$ ซึ่ง $g = (g_1, g_2, \dots, g_m)^T < 0$ และแต่ละ g_i เป็นฟังก์ชันนูนแล้ว $\beta[g(x)]$ เป็นฟังก์ชันนูนของ x ใน $G_0 = \{x \mid g_i(x) < 0; i = 1,2,\dots,m\}$ และถ้าสมมติว่า f เป็นฟังก์ชันนูนของ x แล้ว $U(x,\mu_k) = f(x) + \mu_k\beta[g(x)]$ เป็นฟังก์ชันนูนของ x ใน G_0 เมื่อ $\mu_k > 0$

พิสูจน์ สมมติว่า $x_1; x_2 \in G_0$ แล้วสำหรับ $0 \le \lambda \le 1$ จะได้ว่า

$$g[\lambda x_1 + (1-\lambda)x_2] \le \lambda g(x_1) + (1-\lambda)g(x_2) < 0 \quad \text{; ทุก ๆ ส่วนประกอบของ } g \text{ เป็นฟังก์ชันนูน}$$

ดังนั้น

$$\beta\{g[\lambda x_1 + (1-\lambda)x_2]\} \le \beta\{\lambda g(x_1) + (1-\lambda)g(x_2)\} \quad \text{; } \beta \text{ เป็นฟังก์ชันเพิ่มใน } g \text{ ตามบทตั้งที่ 3.8}$$
$$\le \lambda\beta[g(x_1)] + (1-\lambda)\beta[g(x_2)]\text{; } \beta \text{ เป็นฟังก์ชันนูนใน } g \text{ ตามบทตั้งที่ 3.8}$$

ดังนั้น $\beta[g(x)]$ เป็นฟังก์ชันนูนของ x ใน G_0

และจะได้ว่า $U(x,\mu_k) = f(x) + \mu_k\beta[g(x)]$ เป็นฟังก์ชันนูนตามทฤษฎีบทที่ 2.2 (1), (2) #

สำหรับกำหนดการนูน (C) ที่ได้นิยามไว้ในบทที่ 2 เมื่อใช้วิธีฟังก์ชันกั้นเขตแปลงปัญหา จะได้ฟังก์ชันช่วยที่ $\mu_k (> 0)$ ใด ๆ

$$U(x,\mu_k) = f(x) + \mu_k\left[-\sum_{i=1}^{m}\frac{1}{g_i(x)}\right]$$

เป็นฟังก์ชันนูนของ $x \in G_0 = \{x \mid g_i(x) < 0; i = 1,\ldots,m\}$ ตามบทตั้งที่ 3.9 และนอกจากนี้ G_0 เป็นเซตนูนตามบทตั้งที่ 3.7 ดังนั้นถ้าปัญหาค่าต่ำสุดต่อไปนี้

$$\left.\begin{array}{ll} \text{minimize} & U(x,\mu_k) \\ \text{subject to} & x \in G_0 \end{array}\right\} \quad (3.10)$$

มีจุดต่ำสุดเฉพาะที่ x^* แล้ว x^* จะเป็นจุดต่ำสุดวงกว้างของปัญหาด้วยตามทฤษฎีบทที่ 2.1 และถ้าสมมติว่าทุก ๆ μ_k มีจุด $x(\mu_k) \in G_0$ เป็นจุดต่ำสุดของฟังก์ชัน $U(x,\mu_k)$ แล้วลำดับ $\{x(\mu_k)\}$ จะลู่เข้าหาจุดต่ำสุดวงกว้าง x^* ของปัญหากำหนดการนูน (C) (ถ้าสมมติว่าปัญหา (C) มีผลเฉลย) ผู้วิจัยได้แสดงให้เห็นในตัวอย่างที่ 1 ของหัวข้อ 3.2.2 ว่าในกรณีที่ประยุกต์วิธีฟังก์ชันกั้นเขตกับปัญหาที่ไม่ใช่กำหนดการนูนเพื่อหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการ การสรุปว่าไม่มีจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับจะกระทำไม่ได้เพราะวิธีฟังก์ชันกั้นเขตอาจแก้ปัญหาได้จุดต่ำสุดเฉพาะที่

ในหัวข้อที่ 3.2.2 จะได้กล่าวถึงแนวคิดของขั้นตอนในการหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการซึ่งอยู่บนพื้นฐานของการทำซ้ำตัวเองของวิธีฟังก์ชันกั้นเขต ดังนั้นถ้าสมมติว่า $G_0 = \{x \mid g_i(x) < 0; i = 1,\ldots,m\} \neq \phi$ และแต่ละ g_i เป็นฟังก์ชันนูนแล้ว ในการทำซ้ำวิธีฟังก์ชันกั้นเขตแต่ละครั้งจะกระทำกับปัญหากำหนดการนูนซึ่งในที่สุดแล้วจะสามารถหา $x^* \in G_0$ ได้ ในส่วนต่อไปนี้จะได้ให้ทฤษฎีบทเพื่อแสดงว่าเมื่อทำซ้ำวิธีฟังก์ชันกั้นเขตได้สำเร็จเราจะสามารถหาจุด $x^* \in G_0$

บทตั้งที่ 3.10 ให้ $S(\neq \phi)$ เป็นเซตนูนใน R^n และให้ $f : S \rightarrow R^1$ เป็นฟังก์ชันนูน แล้ว f เป็นฟังก์ชันต่อเนื่องบนเซตภายในของ S

พิสูจน์ ดูเอกสารอ้างอิง [1]

บทตั้งที่ 3.11 กำหนดให้ $g_i, \forall i \in I = \{1,2,\ldots,m\}$ เป็นฟังก์ชันนูนและ $G_0 = \{x \mid g_i(x) < 0, \forall i \in I\} \neq \phi$ ให้ $J(\neq I) \subset I$

พิจารณาปัญหากำหนดการนูน (C_p)

$$\begin{array}{lll} \text{minimize} & g_p(x) & \text{สำหรับบาง } p \in I - J \\ \text{subject to} & g_j(x) \le 0 & ; \ \forall j \in J \\ & x \in R^n & \end{array}$$

ถ้าปัญหา (C_p) มีจุดต่ำสุด x_p^* แล้วโดยวิธีฟังก์ชันกั้นเขตจะกำเนิดจุด x_p' ซึ่ง $g_p(x_p^*) \le g_p(x_p') < 0$

พิสูจน์ เนื่องจากปัญหา (C_p) เป็นปัญหาการหาค่าต่ำสุดของฟังก์ชัน $g_p(x)$ บนเซต $G_J = \{x \mid g_j(x) \le 0; \forall j \in J\}$ ซึ่งเป็นเซตนูน (ตามบทตั้งที่ 3.7) และจากที่กำหนดว่า ปัญหา (C_p) มีจุดต่ำสุด x_p^* ดังนั้นจะได้ว่า x_p^* เป็นจุดต่ำสุดวงกว้างของปัญหา (C_p) บน G_J ตามทฤษฎีบทที่ 2.1

เราสามารถใช้วิธีฟังก์ชันกั้นเขตแก้ปัญหา (C_p) เพื่อหาจุด x_p^* ได้ตามทฤษฎีบทที่ 2.31 เพราะ ปัญหา (C_p) มีจุดต่ำสุด x_p^* และ $G_{0J} = \{x \mid g_j(x) < 0; \forall j \in J\} \ne \phi$ (เพราะ $G_0 \ne \phi$) ดังนั้นถ้า

1. $x_p^* \in G_{0J}$ จะได้ว่าวิธีฟังก์ชันกั้นเขตสามารถกำเนิดจุด x_p^* ได้ตามทฤษฎีบทที่ 2.31 ซึ่ง $g_p(x_p^*) < 0$ เพราะ $G_{0Jp} = \{x \mid g_j(x) < 0; \forall j \in J \cup \{p\}\} \ne \phi$ (เพราะ $G_0 \ne \phi$) และ x_p^* เป็นจุดต่ำสุดวงกว้าง

2. x_p^* เป็นจุดขอบของบริเวณที่เป็นไปได้ G_J แล้ววิธีฟังก์ชันกั้นเขตจะสามารถกำเนิดจุด $x_p' \to x_p^*$ เมื่อพารามิเตอร์กั้นเขต $\mu_k \to 0$ ดังนั้น $x_p' \in N(x_p^*, \varepsilon)$ สำหรับ $\varepsilon > 0$ และ $x_p' \in G_{0J}$ เนื่องจาก $G_{0Jp} = \{x \mid g_j(x) < 0; \forall j \in J \cup \{p\}\} \ne \phi$ (เพราะ $G_0 \ne \phi$) ดังนั้น $g_p(x_p^*) < 0$ เพราะ x_p^* เป็นจุดต่ำสุดวงกว้างของปัญหา (C_p) และโดยความต่อเนื่องของฟังก์ชัน $g_p(x)$ บน G_{0J} จะทำให้ $g_p(x_p^*) \le g_p(x_p') < 0$ #

บทตั้งที่ 3.11 แสดงว่าในขณะการทำซ้ำวิธีฟังก์ชันกั้นเขตเพื่อหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับ $g_i \le 0, \forall i \in I = \{1,2,\ldots,m\}$ เมื่อ $G_0 = \{x \mid g_i(x) < 0, \forall i \in I\} \ne \phi$ และแต่ละ $g_i, \forall i \in I$ เป็นฟังก์ชันนูนแล้ว ถ้าการทำซ้ำของวิธีฟังก์ชันกั้นเขตครั้งใด(สมมติว่าเป็นครั้งที่ p) เป็นปัญหาที่มีจุดต่ำสุด นั่นคือ สามารถหาผลเฉลยของปัญหาได้ โดยวิธีฟังก์ชันกั้นเขตในครั้งที่ p จะกำเนิดจุด x_p' ซึ่งเป็นจุดภายในของบริเวณที่เป็นไปได้และทำให้ค่าของฟังก์ชันจุดประสงค์มีค่าเป็นลบซึ่งจุด x_p' นั้นอาจจะไม่ใช่จุดต่ำสุดของปัญหาก็ได้ ดังนั้นการแก้ปัญหาโดยวิธีฟังก์ชันกั้นเขตในครั้งนี้สามารถหยุดดำเนินการได้ก่อนถึงจุดต่ำสุด

แต่ในบางครั้งของการทำซ้ำวิธีฟังก์ชันกั้นเขตเพื่อหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับอสมการ $g_i \le 0, \forall i \in I = \{1,2,\ldots,m\}$ และแต่ละ $g_i, \forall i \in I$ เป็นฟังก์ชันนูน อาจพบกับปัญหากำหนดการที่ไม่มีจุดต่ำสุดดังนั้นทฤษฎีบทที่ 2.31 ซึ่งเกี่ยวข้องกับการแก้ปัญหากำหนดการนูนโดยวิธีฟังก์ชันกั้นเขตจะไม่รับรองในกรณีนี้ แต่เนื่องจากเราไม่มีความจำเป็นที่จะต้องหาจุดต่ำสุดของปัญหาให้ได้ เราเพียงแต่หาจุดภายในของบริเวณที่เป็นไปได้ที่ทำให้ฟังก์ชันจุดประสงค์

มีค่าเป็นลบ ดังนั้นในการทำซ้ำครั้งใดก็ตามที่ได้ปัญหาแบบไม่มีจุดต่ำสุดก็ถือว่าไม่เป็นอุปสรรคต่อกระบวนการ เพราะถ้ามีข้อกำหนดที่เหมาะสมแล้วเราสามารถหาจุดภายในของปัญหาในกรณีนี้ได้

นิยาม 3.12 กำหนดให้ $g_i, \forall i \in I = \{1,2,\ldots,m\}$ เป็นฟังก์ชันต่อเนื่องและให้ $J(\neq I) \subset I$

พิจารณาปัญหากำหนดการ (C_p)

$$\begin{array}{lll} \text{minimize} & g_p(\mathbf{x}) & \text{สำหรับบาง } p \in I - J \\ \text{subject to} & g_j(\mathbf{x}) \le 0 & ; \ \forall j \in J \\ & \mathbf{x} \in R^n & \end{array}$$

จะเรียกปัญหา (C_p) ว่า "ปัญหาไม่มีจุดต่ำสุดแบบไม่มีขอบเขตล่าง" ถ้ามี $\mathbf{x}' \in G_J = \{\mathbf{x} \mid g_j(\mathbf{x}) \le 0, \forall j \in J\}$ ซึ่งทำให้ $g_p(\mathbf{x}') \to -\infty$

กำหนด $C = \{c \in R^1 \mid c < g_p(\mathbf{x}), \forall \mathbf{x} \in G_J$ และ c เป็นค่าจำกัด$\}$ เราจะเรียกปัญหา (C_p) ว่า "ปัญหาไม่มีจุดต่ำสุดแบบมีขอบเขตล่าง" ถ้ามี $c_{max} \in C$ ซึ่งทำให้ $c_{max} = \sup_{c \in C} c < g_p(\mathbf{x})\ , \forall \mathbf{x} \in G_J$

บทตั้งที่ 3.13 กำหนดให้ $g_i, \forall i \in I = \{1,2,\ldots,m\}$ เป็นฟังก์ชันนูนและ $G_{0J} = \{\mathbf{x} \mid g_j(\mathbf{x}) < 0, \forall j \in J\} \neq \phi$ ให้ $J(\neq I) \subset I$

พิจารณาปัญหากำหนดการนูน (C_p)

$$\begin{array}{lll} \text{minimize} & g_p(\mathbf{x}) & \text{สำหรับบาง } p \in I - J \\ \text{subject to} & g_j(\mathbf{x}) \le 0 & ; \ \forall j \in J \\ & \mathbf{x} \in R^n & \end{array}$$

ถ้าปัญหา (C_p) เป็นปัญหาไม่มีจุดต่ำสุดแบบไม่มีขอบเขตล่างแล้ว ในระหว่างขบวนการแก้ปัญหา (C_p) โดยวิธีฟังก์ชันกั้นเขตจะได้ว่า สำหรับพารามิเตอร์กั้นเขต μ_k คงที่ใด ๆ วิธีการจะกำเนิดจุด $\mathbf{x}_p(\mu_k)$ ที่ทำให้ $g_p(\mathbf{x}_p(\mu_k)) < 0$

พิสูจน์ เนื่องจากปัญหา (C_p) เป็นปัญหาไม่มีจุดต่ำสุดแบบไม่มีขอบเขตล่างและ $g_p(\mathbf{x})$ เป็นฟังก์ชันต่อเนื่องบน G_{0J}(เพราะ $g_p(\mathbf{x})$ เป็นฟังก์ชันนูนที่นิยามบน $G_J = \{\mathbf{x} \mid g_j(\mathbf{x}) \le 0; \forall j \in J\}$ ซึ่งเป็นเซตนูน) ดังนั้น $g_p(\mathbf{x})$ เป็นฟังก์ชันที่ไม่มีขอบเขตล่างบน G_{0J} สามารถใช้วิธีฟังก์ชันกั้นเขตแก้ปัญหาได้เพราะ $G_{0J} \neq \phi$ ดังนั้นเมื่อแปลงปัญหา (C_p) โดยวิธีฟังก์ชันกั้นเขตแล้วจะได้ปัญหาใหม่เป็น

$$\begin{array}{lll} \text{minimize} & U_p(\mathbf{x},\mu_k) = g_p(\mathbf{x}) + \mu_k\beta_p(\mathbf{x}) & \\ \text{subject to} & g_j(\mathbf{x}) < 0 \quad ; \forall j \in J & \\ & \mathbf{x} \in R^n & \end{array}$$

พิจารณาปัญหาใหม่ที่ μ_k คงที่ใด ๆ จะได้ว่า $U_p(\mathbf{x},\mu_k)$ เป็นฟังก์ชันที่ไม่มีขอบเขตล่างบน G_{0J} เพราะ $\mu_k\beta_p(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in G_{0J}$ และ $g_p(\mathbf{x})$ เป็นฟังก์ชันที่ไม่มีขอบเขตล่างบน G_{0J} ดังนั้นเมื่อดำเนินการหาค่าต่ำสุดของ $U_p(\mathbf{x},\mu_k)$ ที่ μ_k คงที่ใด ๆ $U_p(\mathbf{x},\mu_k)$ จะลดค่าลงเรื่อย ๆ ในขณะที่ $\mu_k\beta_p(\mathbf{x})$ มีขอบเขตล่างที่ศูนย์ เมื่อ $U_p(\mathbf{x},\mu_k)$ ลดค่าลงจนถึงจุด $\mathbf{x}_p(\mu_k)$ ที่ทำให้ $U_p(\mathbf{x}_p(\mu_k),\mu_k) < 0$ แล้วจะได้ว่า $g_p(\mathbf{x}_p(\mu_k)) < 0$ (เพราะ $\mu_k\beta_p(\mathbf{x}_p(\mu_k)) \geq 0$) #

บทตั้งที่ 3.14 กำหนดให้ $g_i , \forall i \in I = \{1,2,\ldots,m\}$ เป็นฟังก์ชันนูนและ $G_0 = \{x \mid g_i(x) < 0, \forall i \in I\} \neq \phi$ ให้ $J(\neq I) \subset I$

พิจารณาปัญหากำหนดการนูน (C_p)

$$\begin{array}{lll} \text{minimize} & g_p(\mathbf{x}) & \text{สำหรับบาง } p \in I - J \\ \text{subject to} & g_j(\mathbf{x}) \leq 0 & ; \forall j \in J \\ & \mathbf{x} \in R^n & \end{array}$$

ถ้าปัญหา (C_p) เป็นปัญหาไม่มีจุดต่ำสุดแบบมีขอบเขตล่างแล้ว ในระหว่างขบวนการแก้ปัญหา (C_p) โดยวิธีฟังก์ชันกั้นเขตจะได้ว่า สำหรับพารามิเตอร์กั้นเขต μ_k ใด ๆ ที่เล็กเพียงพอให้แทนด้วย μ^* วิธีการจะกำเนิดจุด $\mathbf{x}_p(\mu^*)$ ที่ทำให้ $g_p(\mathbf{x}_p(\mu^*)) < 0$

พิสูจน์ เนื่องจากคุณสมบัติของเทอมกั้นเขต $\mu_k\beta(\mathbf{x})$ จะมีค่าเข้าใกล้เทอมกั้นเขตอุดมคติ $\sigma(\mathbf{x})$ เมื่อ $\mu_k \to 0$ นั่นคือถ้า $\mu_k = \mu^*$ มีค่าเล็กเพียงพอแล้ว

$$\mu^*\beta(\mathbf{x}) \to 0 \ , \forall \mathbf{x} \in G_{0J} = \{\mathbf{x} \mid g_j(\mathbf{x}) < 0, \forall j \in J\}$$

และจะได้ว่า $U_p(\mathbf{x},\mu^*) = g_p(\mathbf{x}) + \mu^*\beta(\mathbf{x}) \to g_p(\mathbf{x})$, $\forall \mathbf{x} \in G_{0J}$ จากที่กำหนดว่า $G_0 \neq \phi$ ดังนั้น $G_{0Jp} = \{\mathbf{x} \mid g_j(\mathbf{x}) < 0, \forall j \in J \cup \{p\}\} \neq \phi$ แสดงว่าจะมี $\mathbf{x}_p(\mu^*) \in G_{0J}$ ที่ทำให้ $g_p(\mathbf{x}_p(\mu^*)) < 0$ เพราะ $\mu^*\beta(\mathbf{x}_p(\mu^*)) \to 0$ #

3.2 แนวคิดที่ใช้ในการหาจุดที่เป็นไปได้ตามเงื่อนไขบังคับ (1.1)-(1.3)

3.2.1 แนวคิดในการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการ

การหาจุดที่เป็นไปได้ที่สอดคล้องกับเงื่อนไข (1.1)

$$h_j(x) = 0 \quad ; \; j = 1,\ldots,r$$

ซึ่ง $x \in R^n$ ก็คือ การแก้ระบบสมการไม่เชิงเส้นเพื่อหาผลเฉลยของระบบสมการนั่นเอง โดยในที่นี้ไม่มีข้อกำหนดของ r และ n ซึ่งอาจจะเท่ากันหรือไม่เท่ากันก็ได้ ในกรณีที่ $r = n$ นั้น ได้มีนักวิจัยหลายท่านได้ทำการศึกษาการหาผลเฉลยในกรณีนี้เป็นจำนวนมากซึ่งในปัจจุบันนี้มีเครื่องมือและซอฟแวร์หลายตัวที่สามารถหาผลเฉลยในกรณีนี้ได้ เช่น MATLAB

สำหรับกรณีที่ $r \neq n$ นั้น ผู้วิจัยได้ประยุกต์ใช้ขั้นตอนวิธีฟังก์ชันทัณฑ์เพื่อหาผลเฉลยของระบบสมการดังกล่าว เพราะในการแก้ปัญหาค่าต่ำสุดแบบมีเงื่อนไขบังคับแบบสมการ ถ้ากำหนดการนั้นมีผลเฉลย ผลเฉลยย่อมสอดคล้องกับเงื่อนไขบังคับด้วย สำหรับงานวิจัยนี้จะประยุกต์ใช้วิธีฟังก์ชันทัณฑ์ในการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการในกรณี

1. จำนวนสมการเท่ากับจำนวนตัวแปร $(r = n)$
2. จำนวนสมการน้อยกว่าจำนวนตัวแปร $(r < n)$
3. จำนวนสมการมากกว่าจำนวนตัวแปร $(r > n)$

โดยกำหนดให้ฟังก์ชันจุดประสงค์ที่ต้องการหาค่าต่ำสุดเป็นฟังก์ชันคงที่ $f(x) = 0$ นั่นคือแก้ปัญหาต่อไปนี้

$$\begin{aligned} &\text{minimize} && f(x) = 0 \\ &\text{subject to} && h_j(x) = 0 \quad ; \; j = 1,\ldots,r \\ & && x \in R^n \end{aligned}$$

ด้วยวิธีฟังก์ชันทัณฑ์

3.2.2 แนวคิดในการหาจุดภายในของบริเวณที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการ

จากเงื่อนไขบังคับแบบอสมการ (1.2)

$$g_i(x) \leq 0 \quad ; \; i = 1,\ldots,m$$

เราต้องการหาจุด $x^* \in \{x \in R^n \mid g_i(x) < 0 \;\; ; \; i = 1,\ldots, m\}$ จะเห็นว่า x^* เป็นจุดภายในของบริเวณที่เป็นไปได้ของเงื่อนไขบังคับ ดังนั้นแนวคิดที่จะหาจุด x^* ดังกล่าวเราจะประยุกต์วิธี

จุดภายในหรือวิธีฟังก์ชันกั้นเขตนั่นเอง โดยใช้วิธีการทำซ้ำวิธีฟังก์ชันกั้นเขตไปเรื่อย ๆ จนกว่าจะได้จุด x^* ตัวอย่างที่สมมติขึ้นต่อไปนี้จะช่วยให้เกิดความเข้าใจในการทำซ้ำวิธีฟังก์ชันกั้นเขตเพื่อให้ได้จุด x^* เช่นสมมติว่า $m = 4$ และกำหนดจุด x_0 ซึ่งทำให้ $g_1(x_0) < 0$ แต่ $g_i(x_0) \geq 0$ สำหรับ $i = 2,3,4$ ดังนั้นเราจะเลือก $g_2(x)$ เป็นฟังก์ชันจุดประสงค์และ $g_1(x) \leq 0$ เป็นเงื่อนไขบังคับแล้วทำการแก้ปัญหาต่อไปนี้ โดยวิธีฟังก์ชันกั้นเขตซึ่งใช้จุด x_0 เป็นจุดเริ่มต้น

$$\left.\begin{array}{ll} \text{minimize} & g_2(x) \\ \text{subject to} & g_1(x) \leq 0 \\ & x \in R^n \end{array}\right\} \qquad (3.11)$$

ถ้ากระบวนการแก้ปัญหา (3.11) กำเนิดจุด x_1 ซึ่งอาจจะเป็นหรือไม่เป็นจุดต่ำสุดวงกว้างของปัญหา แต่ทำให้ $g_2(x_1) < 0$ ได้ ขบวนการจะสามารถดำเนินการต่อได้ ถ้าในกรณีที่ x_1 เป็นจุดต่ำสุดวงกว้างของปัญหา (3.11) และ $g_2(x_1) \geq 0$ ดังนั้น $\{x \in R^n \mid g_i(x) < 0 \ ; i = 1,\ldots,4\} = \phi$ และให้หยุดดำเนินการ ในตอนนี้เราสมมติว่าได้จุด x_1 ที่ทำให้ $g_2(x_1) < 0$ เราจะหาค่า $g_3(x_1)$ (ไม่ต้องหาค่า $g_1(x_1)$ เพราะวิธีฟังก์ชันกั้นเขตจะได้ว่า $g_1(x_1) < 0$ เสมอ) สมมติโดยไม่สูญเสียความเป็นทั่วไปว่า $g_3(x_1) < 0$ ดังนั้นในขณะนี้เราจะเลือก $g_4(x)$ เป็นฟังก์ชันจุดประสงค์และให้ $g_i(x) \leq 0, i = 1,2,3$ เป็นเงื่อนไขบังคับและใช้จุด x_1 เป็นจุดเริ่มต้นในการแก้ปัญหาต่อไปนี้โดยวิธีฟังก์ชันกั้นเขต

$$\left.\begin{array}{ll} \text{minimize} & g_4(x) \\ \text{subject to} & g_1(x) \leq 0 \\ & g_2(x) \leq 0 \\ & g_3(x) \leq 0 \\ & x \in R^n \end{array}\right\} \qquad (3.12)$$

เมื่อดำเนินการแก้ปัญหา (3.12) จนได้จุด x_2 (อาจจะเป็น/ไม่เป็นจุดต่ำสุดวงกว้าง) ซึ่งทำให้ $g_4(x_2) < 0$ แล้ว เราจะได้ว่า x_2 เป็นจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับทั้งสี่นั่นคือ $x_2 \in \{x \in R^n \mid g_i(x) < 0 \ ; i = 1,\ldots,4\}$ แต่ถ้า x_2 เป็นจุดต่ำสุดวงกว้างของปัญหา (3.12) และทำให้ $g_4(x_2) \geq 0$ แล้วเราสามารถสรุปได้ว่า $\{x \in R^n \mid g_i(x) < 0 \ ; i = 1,\ldots,4\} = \phi$

สิ่งที่เป็นข้อจำกัดทางทฤษฎีของวิธีฟังก์ชันกั้นเขต คือ ในกรณีที่เราต้องการหาจุดต่ำสุดวงกว้างของปัญหาที่มีเงื่อนไขบังคับในรูปแบบต่อไปนี้

$$\begin{aligned} &\text{minimize} && f(\mathbf{x}) \\ &\text{subject to} && g_i(\mathbf{x}) \le 0 \quad ; \, i = 1,\dots,m \\ & && \mathbf{x} \in R^n \end{aligned}$$

เราจะต้องหาจุดต่ำสุดวงกว้างของปัญหาแบบไม่มีเงื่อนไขบังคับต่อไปนี้

$$\begin{aligned} &\text{minimize} && U(\mathbf{x}, \mu_k) = f(\mathbf{x}) - \mu_k[\sum_{i=1}^{m} \frac{1}{g_i(\mathbf{x})}] \\ &\text{subject to} && g_i(\mathbf{x}) < 0 \quad ; \, i = 1,\dots,m \\ & && \mathbf{x} \in R^n \end{aligned}$$

ซึ่งเป็นปัญหาที่ได้จากการแปลงปัญหาข้างบนด้วยวิธีฟังก์ชันกั้นเขตให้ได้สำหรับแต่ละ $\mu_k > 0$ เมื่อ $\{\mu_k\}$ เป็นลำดับลดลงทางเดียว

ตัวอย่างต่อไปนี้จะแสดงให้เห็นว่าในกรณีที่การหาค่าต่ำสุดของ $U(\mathbf{x}, \mu_k)$ แล้วได้จุดต่ำสุดเฉพาะที่สำหรับแต่ละค่าของ μ_k แล้วปัญหาแบบมีเงื่อนไขบังคับดังกล่าวจะได้ผลเฉลยที่เป็นจุดต่ำสุดเฉพาะที่

ตัวอย่างที่ 1 หาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับต่อไปนี้

$$g_1(x) = x \le 0$$
$$g_2(x) = \frac{x^4}{4} + \frac{11}{3}x^3 + 17x^2 + 24x + 1 \le 0$$

เมื่อกำหนด $x_0 = -7$ จะเห็นว่า $g_1(x_0) < 0$ ดังนั้นเราจะใช้ $g_2(x)$ เป็นฟังก์ชันจุดประสงค์ และ $g_1(x) \le 0$ เป็นเงื่อนไขบังคับและดำเนินการแก้ปัญหา

$$\begin{aligned} &\text{minimize} && g_2(x) \\ &\text{subject to} && g_1(x) \le 0 \\ & && \mathbf{x} \in R^1 \end{aligned}$$

โดยวิธีฟังก์ชันกั้นเขตแล้วจะได้ปัญหาใหม่เป็น

$$\text{minimize} \qquad U(x,\mu_k) = g_2(x) - \frac{\mu_k}{g_1(x)}$$

$$= \frac{x^4}{4} + \frac{11}{3}x^3 + 17x^2 + 24x + 1 - \mu_k x^{-1}$$

$$\text{subject to} \qquad g_1(x) = x < 0$$

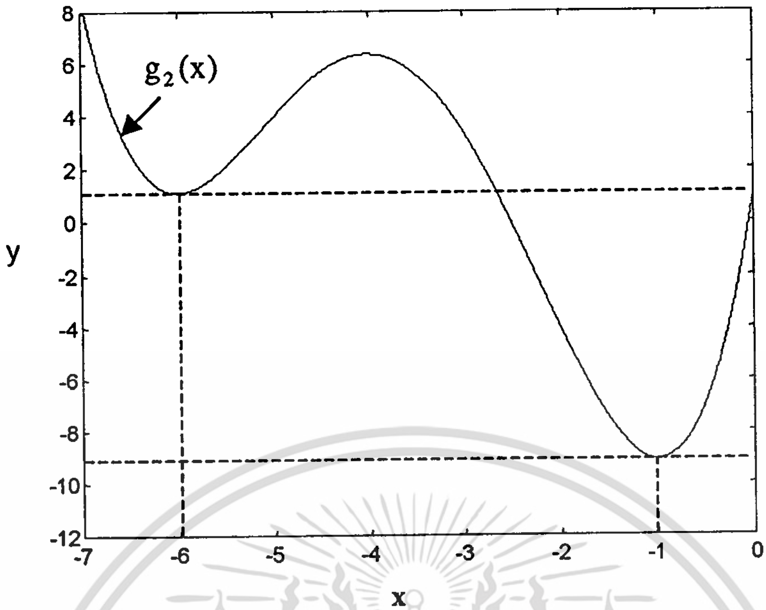
$$x \in R^1$$

เมื่อแก้ปัญหาใหม่สำหรับแต่ละ μ_k ซึ่ง $\{\mu_k\}$ เป็นลำดับลดลงทางเดียวเมื่อ $\mu_k \to 0$ จะได้จุดต่ำสุด $x_1^* = -6$ และ $g_2(x_1^*) = 1$ ดังผลลัพธ์ที่แสดงในตารางที่ 3.1

ตารางที่ 3.1 แสดงผลการคำนวณสำหรับตัวอย่างที่ 1

k	μ_k	x_{k+1}	$g_2(x_{k+1})$	$\mu_k\beta(x_{k+1})$	$g_1(x_{k+1})$
0	10.000000	(-6.0270164)	1.0036955769	1.6591957471	-6.0270164130
1	1.000000	(-6.0027698)	1.0000384097	0.1665897621	-6.0027698423
2	0.100000	(-6.0002777)	1.0000003856	0.0166658953	-6.0002776986
3	0.010000	(-6.0000278)	1.0000000039	0.0016666590	-6.0000277660
4	0.001000	(-6.0000028)	1.0000000000	0.0001666666	-6.0000027777
5	0.000100	(-6.0000003)	1.0000000000	0.0000166667	-6.0000002778
6	0.000010	(-6.0000003)	1.0000000000	0.0000016667	-6.0000002778

จากตัวอย่าง 1 ถ้าเราสรุปว่า $\{x \in R^1 \mid g_i(x) < 0 \ ; \ i = 1,2\} = \phi$ จะเป็นการสรุปที่ผิดเพราะเมื่อวาดกราฟ $g_2(x) = \frac{x^4}{4} + \frac{11}{3}x^3 + 17x^2 + 24x + 1$ จะได้ดังรูปที่ 3.1 และจากการวิเคราะห์จะเห็นว่าเมื่อ $x \in (-\infty,0]$ (สอดคล้องกับเงื่อนไข $g_1(x) \le 0$) เราได้ว่า $g_2(x)$ มีจุดต่ำสุดอีกจุดหนึ่งคือ $x_2^* = -1$ (เป็นจุดต่ำสุดวงกว้าง) และ $g_2(x_2^*) \cong -9.4167 < 0$ ดังนั้น $x_2^* \in \{x \in R^1 \mid g_i(x) < 0 \ ; i = 1,2\}$ #



รูปที่ 3.1 แสดงกราฟ $g_2(x) = \frac{x^4}{4} + \frac{11}{3}x^3 + 17x^2 + 24x + 1$

3.2.3 แนวคิดในการหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบผสม

จากเงื่อนไขบังคับ (1.3)

$$g_i(x) \le 0 \quad ; i = 1,\ldots,m$$
$$h_j(x) = 0 \quad ; j = 1,\ldots,r$$

เราต้องการหาจุด $x^* \in \{x \in R^n \mid g_i(x) < 0 \; ; i = 1,\ldots,m$ และ $h_j(x) = 0 \; ; j = 1,\ldots,r\}$ การแก้ปัญหาเพื่อหาจุด x^* จะอยู่บนพื้นฐานของแนวคิดในหัวข้อ 3.2.1 และ 3.2.2 โดยการทำงานจะแบ่งออกเป็นสองส่วนคือ

ส่วนแรก จะหาจุด $x' \in \{x \in R^n \mid g_i(x) < 0 \; ; i = 1,\ldots,m\}$ ในส่วนนี้จะใช้การทำซ้ำตัวเองของวิธีฟังก์ชันกั้นเขตตามแนวคิด 3.2.2

ส่วนที่สอง จะหาจุด $x^* \in \{x \in R^n \mid g_i(x) < 0 \; ; i = 1,\ldots,m$ และ $h_j(x) = 0 \; ; j = 1,\ldots,r\}$ ในส่วนนี้จะใช้วิธีผสมและแนวคิด 3.2.1 นั่นคือแก้ปัญหาต่อไปนี้

$$\left.\begin{array}{lll} \text{minimize} & f(x) = 0 & \\ \text{subject to} & g_i(x) \le 0 & ; i = 1,..,m \\ & h_j(x) = 0 & ; j = 1,\ldots,r \\ & x \in R^n & \end{array}\right\} \quad (3.13)$$

โดยใช้วิธีผสมแปลงปัญหา (3.13) เป็นปัญหาใหม่คือ

$$\left.\begin{array}{ll} \text{minimize} & M(\mathbf{x},\mu_k) = 0 - \mu_k\left[\sum_{i=1}^{m}\frac{1}{g_i(\mathbf{x})}\right] + \frac{1}{\mu_k}\left[\sum_{j=1}^{r}(h_j(\mathbf{x}))^2\right] \\ \text{subject to} & g_i(\mathbf{x}) < 0 \quad ; i = 1,\ldots,m \\ & \mathbf{x} \in R^n \end{array}\right\} \quad (3.14)$$

แก้ปัญหา (3.14) ด้วยจุดเริ่มต้น $\mathbf{x}'$ (จากส่วนแรก) จนได้จุดต่ำสุดซึ่งจุดต่ำสุดที่ได้จากการแก้ปัญหาใหม่นี้ก็คือจุด $\mathbf{x}^*$ นั่นเอง

3.3 ขั้นตอนวิธีที่ใช้ในการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบต่าง ๆ

3.3.1 ขั้นตอนวิธีการค้นหาแนวเส้นและการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับโดยใช้สูตร BFGS ผกผัน

ในงานวิจัยนี้ใช้วิธีถอยหลังกลับ (backtracking method) ในการค้นหาแนวเส้นเพื่อหา α_k ที่สอดคล้องกับเกณฑ์ 4 (เงื่อนไขของอาร์มิโจ)

$$f(\mathbf{x}_{k+1}) = f(\mathbf{x}_k + \alpha\mathbf{p}_k) \le f(\mathbf{x}_k) + \mu\alpha\mathbf{p}_k^T\nabla f(\mathbf{x}_k)$$

ปกติมักเลือกให้ $\mu \in (0,\frac{1}{2})$ วิธีถอยหลังกลับที่ใช้ในงานวิจัยนี้เป็นดังนี้

ให้เริ่มต้นด้วย $\alpha_0 = 1$ แล้วทำการลดค่า α ลงทีละครึ่ง ดังนั้นจะได้ลำดับเรขาคณิตของค่า α เป็น $\{1,\frac{1}{2},\frac{1}{4},\frac{1}{8},\ldots\}$ ซึ่ง α_k ที่ต้องการก็คือ $\alpha \in \{1,\frac{1}{2},\frac{1}{4},\frac{1}{8},\ldots\}$ และเป็นค่าแรกที่ทำให้เกณฑ์ 4 (เงื่อนไขอาร์มิโจ) เป็นจริง โดยวิธีถอยหลังกลับนี้ ถ้า f เป็นฟังก์ชันที่มีขอบเขตล่างและ $\mathbf{p}_k$ เป็นทิศทางลดแล้ว จะมี α_k ที่สอดคล้องกับเกณฑ์ 4 เสมอตามทฤษฎีบทที่ 2.16 เพราะถ้า $\alpha_0 = 1 \le \alpha_\mu$ (ในทฤษฎีบทที่ 2.16) แล้ว $\alpha_0 = 1$ จะเป็น α_k ที่ต้องการ แต่ถ้า $\alpha_\mu < \alpha_0 = 1$ แล้วเราสามารถลดค่า α_0 ลงไปทีละครึ่งจนได้ค่า α ตัวแรกซึ่ง $\alpha \le \alpha_\mu$ โดยค่า α นี้จะเป็น α_k ที่เราต้องการจะเห็นว่าในกรณีนี้เราทำการลดค่า $\alpha_0 = 1$ เป็นจำนวนครั้งที่จำกัดและจะได้ α_k ที่สอดคล้องกับเกณฑ์ 4 เสมอ ดังนั้นการลดค่า α_0 ลงไปเรื่อย ๆ แบบไม่สิ้นสุดจะไม่เกิดขึ้น ทำให้ α_k ที่ได้ไม่เล็กจนเกินไป

จากที่กล่าวมาข้างต้นจะเห็นว่า ถ้าใช้วิธีถอยหลังกลับเพื่อหา α_k ที่ทำให้เกณฑ์ 4 เป็นจริงซึ่ง α_k ที่ได้จะไม่เล็กเกินไป ดังนั้นจึงไม่มีความจำเป็นที่จะใช้เกณฑ์ 3 ในการหา α_k

<u>ขั้นตอนวิธีที่ 3.1 การค้นหาแนวเส้น</u>

ขั้นเริ่มต้น รับค่าจุด $\mathbf{x}_k$,ทิศทางลด $\mathbf{p}_k$, $\nabla f(\mathbf{x}_k)$ และ $f(\mathbf{x}_k)$ ให้ $\mu \in (0,\frac{1}{2})$ และ $\alpha_k = 1$ จะได้ $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{p}_k$ คำนวณ $f(\mathbf{x}_{k+1})$ และไปยังขั้นหลัก

ขั้นหลัก

(1). ถ้า $f(\mathbf{x}_{k+1}) > f(\mathbf{x}_k) + \mu\alpha_k \mathbf{p}_k^T \nabla f(\mathbf{x}_k)$ แล้ว

ให้ $\alpha_k = \alpha_k / 2$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$$

คำนวณ $f(\mathbf{x}_{k+1})$ และไปยังขั้นที่ 1

ถ้าเป็นอย่างอื่นจะได้ α_k และจุด $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$ ที่ต้องการ

จากขั้นตอนวิธีที่ 3.1 สามารถเขียนแผนผังการทำงานได้ดังรูปที่ 3.2

สำหรับการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in R^n \end{array}$$

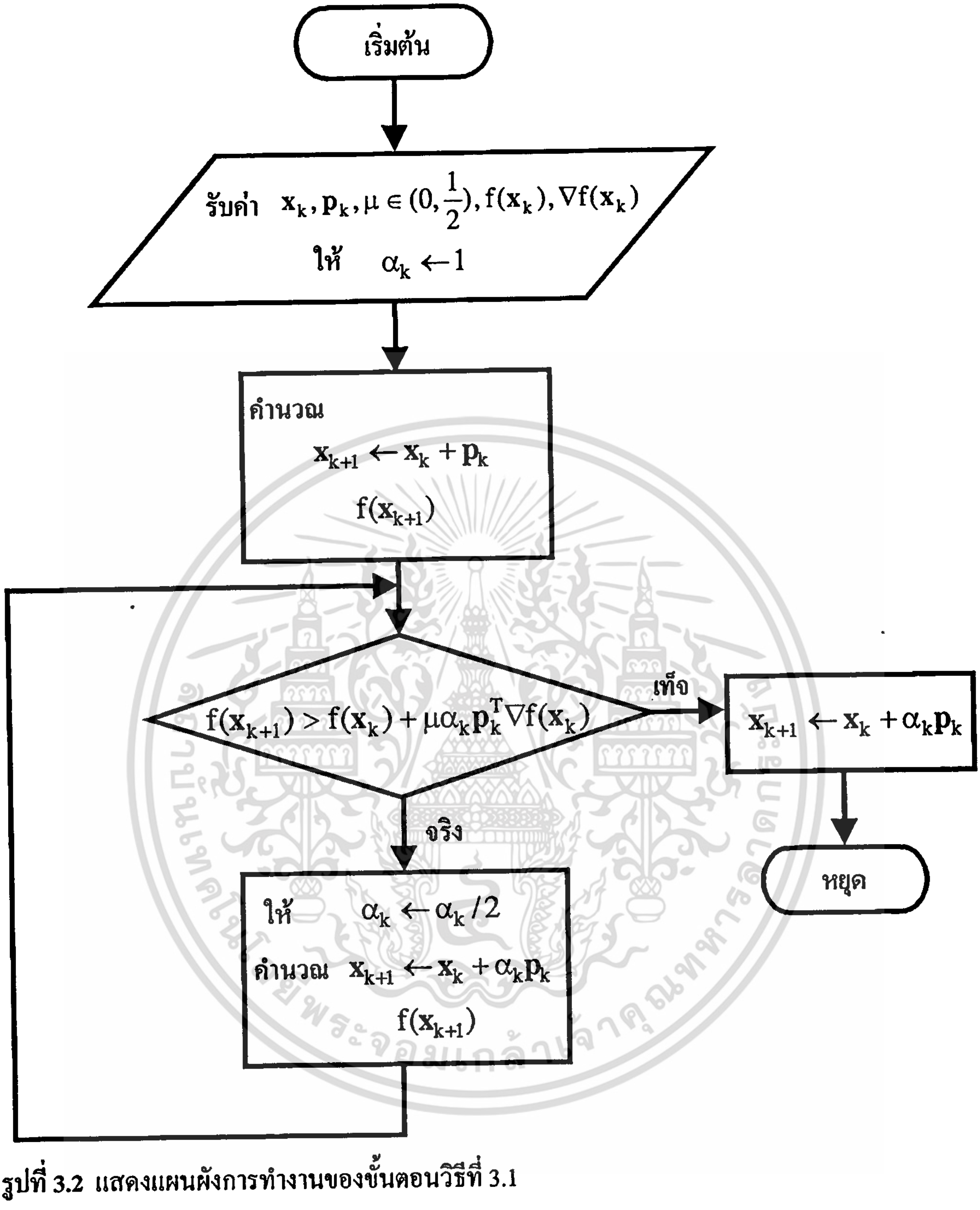
เราจะใช้ขั้นตอนวิธีดังต่อไปนี้

<u>ขั้นตอนวิธีที่ 3.2 การแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับโดยใช้สูตร BFGS ผกผัน</u>

ขั้นเริ่มต้น เลือกจุดเริ่มต้น $\mathbf{x}_0 = (x_1,\ldots,x_n)^T$, เลือก H_0 เป็นเมตริกซ์สมมาตรที่เป็นบวกแน่นอน โดยปกติมักเลือก $H_0 = I_n$ และเลือก $\varepsilon_2 > 0$ (ดัชนี 2 ที่กำหนดขึ้นเพราะเราอาจจะใช้ขั้นตอนวิธีที่ 3.2 นี้เป็นลูปภายในของลูปการทำงานใหญ่อีกลูปหนึ่ง) เป็นค่าสเกลาร์ในการหยุดและคำนวณ $\nabla f(\mathbf{x}_0)$ แล้วไปยังขั้นหลัก

ขั้นหลัก

1. ถ้า $\|\nabla f(\mathbf{x}_0)\| < \varepsilon_2$ จะได้ $\mathbf{x}_0$ เป็นจุดต่ำสุดของปัญหาและให้หยุดดำเนินการ ถ้าเป็นอย่างอื่น คำนวณ $f(\mathbf{x}_0)$,ให้ $k = 0$ และไปยังขั้นที่ 2
2. ดำเนินการดังนี้
   - 2.1 คำนวณ $\mathbf{p}_k$ จาก $\mathbf{p}_k = -H_k \nabla f(\mathbf{x}_k)$
   - 2.2 ใช้ขั้นตอนวิธีที่ 3.1 ค้นหาแนวเส้นจะได้ α_k ,จุด $\mathbf{x}_{k+1}$ และ $f(\mathbf{x}_{k+1})$
   - 2.3 คำนวณ $\nabla f(\mathbf{x}_{k+1})$ แล้วไปยังขั้นที่ 3
3. ถ้า $\|\nabla f(\mathbf{x}_{k+1})\| < \varepsilon_2$ จะได้จุด $\mathbf{x}_{k+1}$ เป็นจุดต่ำสุดของปัญหาและให้หยุดดำเนินการ ถ้าเป็นอย่างอื่น คำนวณ $\mathbf{s}_k$ จาก $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$.



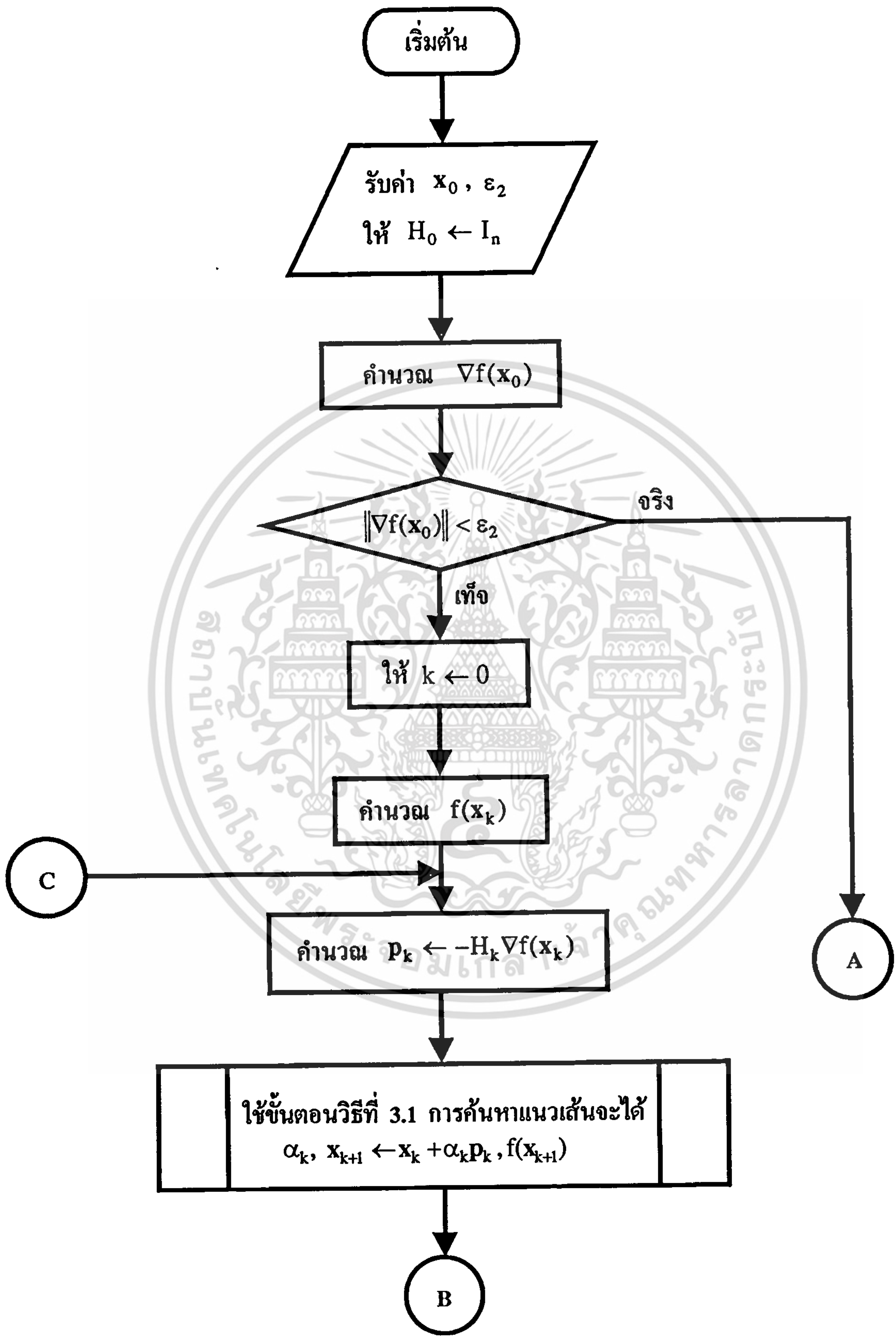
รูปที่ 3.2 แสดงแผนผังการทำงานของขั้นตอนวิธีที่ 3.1

คำนวณ $\mathbf{y}_k$ จาก $\mathbf{y}_k = \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)$

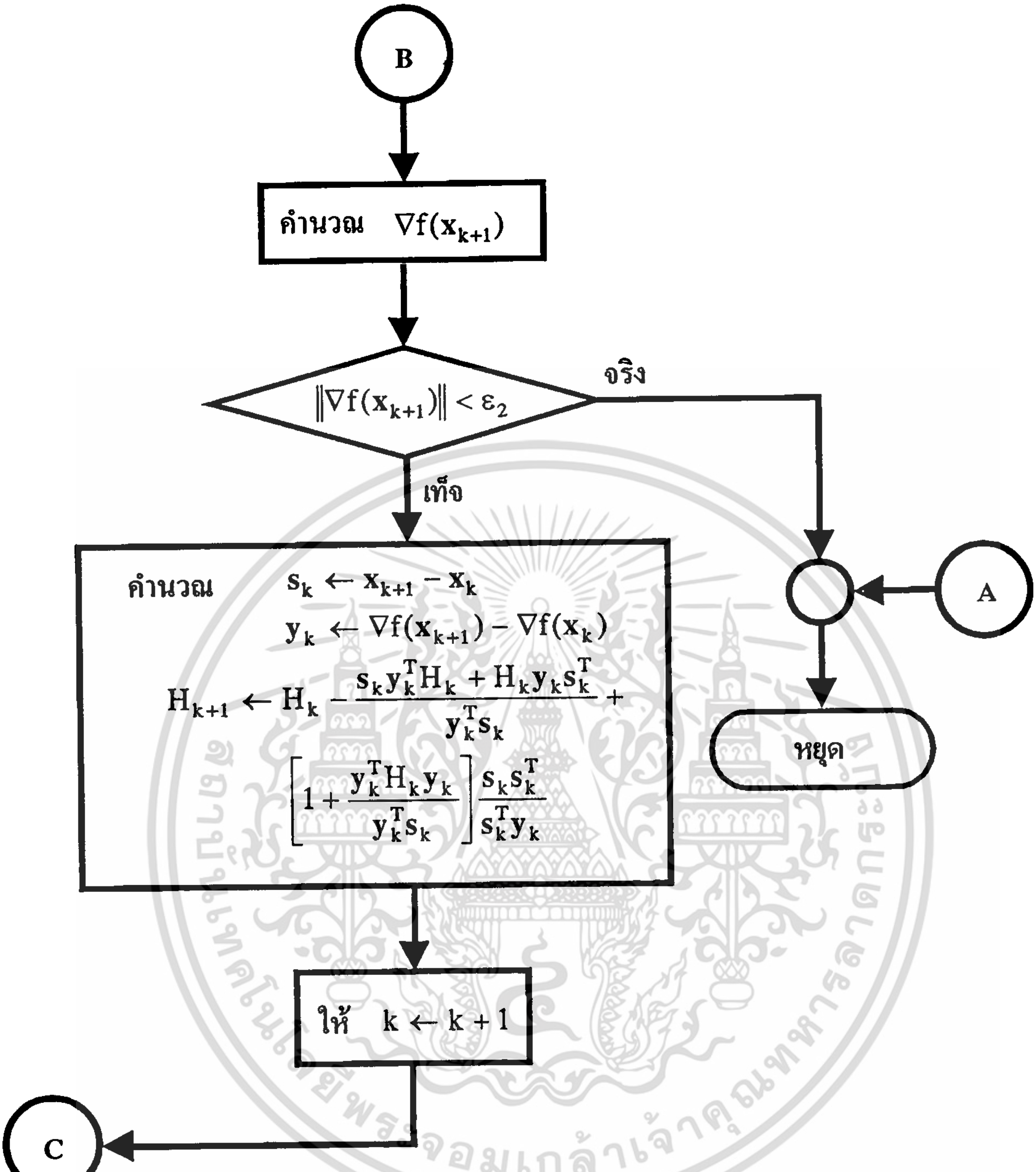
คำนวณ H_{k+1} จาก H_k โดยสูตร (3.1)

ให้ $k = k+1$ และไปยังขั้นที่ 2

จากขั้นตอนวิธีที่ 3.2 สามารถเขียนแผนผังการทำงานได้ดังรูปที่ 3.3



รูปที่ 3.3 แสดงแผนผังการทำงานของขั้นตอนวิธีที่ 3.2



รูปที่ 3.3 (ต่อ)

สำหรับขั้นตอนวิธีที่จะได้เสนอในส่วนหลังนี้ จะใช้ขั้นตอนวิธีที่ 3.1 และ 3.2 เป็นลูปภายใน ดังนั้นเพื่อให้สอดรับกับขั้นตอนวิธีและทฤษฎีที่ได้อธิบายมาแล้ว และป้องกันการสับสนของเลขดัชนี ผู้วิจัยจะอธิบายการทำงานลูปใหญ่ภายนอกด้วยดัชนี t และค่าคงตัวการหยุด ε_1 ส่วนลูปภายในจะใช้ดัชนี k และค่าคงตัวการหยุด ε_2 ตามที่กล่าวมาแล้ว แต่สำหรับโปรแกรมการคำนวณและผลลัพธ์ที่แสดงในบทที่ 4 และ 5 ผู้วิจัยเลือกใช้ดัชนี i แทนการทำงานลูปภายใน , ε_2 เป็นค่าคงตัวการหยุดและดัชนี k แทนการทำงานของลูปใหญ่ , ε_1 เป็นค่าคงตัวในการหยุด

3.3.2 ขั้นตอนวิธีในการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการ (1.1)

จากหัวข้อ 3.2.1 การหาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการ เราจะทำการแก้ปัญหาต่อไปนี้

$$\begin{aligned} &\text{minimize} && f(\mathbf{x}) = 0 \\ &\text{subject to} && h_j(\mathbf{x}) = 0 \quad ; \; j = 1,..,r \\ &&& \mathbf{x} \in R^n \end{aligned}$$

ด้วยวิธีฟังก์ชันทัณฑ์ นั่นคือให้เราแก้ปัญหา

$$\begin{aligned} &\text{minimize} && P(\mathbf{x}, \rho_k) = \rho_k \alpha(\mathbf{x}) = \rho_k\{[h_1(\mathbf{x})]^2 + \ldots + [h_r(\mathbf{x})]^2\} \\ &\text{subject to} && \mathbf{x} \in R^n \end{aligned}$$

สำหรับแต่ละค่าของ ρ_k ซึ่ง $\{\rho_k\}$ เป็นลำดับเพิ่มขึ้นทางเดียว ขั้นตอนวิธีที่ 3.3 ที่จะกล่าวต่อไปนี้จะใช้หาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการซึ่งเรียกใช้ขั้นตอนวิธีที่ 3.2 เป็นขั้นตอนวิธีย่อยโดยไม่มีการเปลี่ยนแปลงใด ๆ ดังนั้นเพื่อความกระชับจะไม่กล่าวถึงรายละเอียดของขั้นตอนวิธีที่ 3.2 ซ้ำอีกรวมทั้งจะไม่แสดงการส่งและรับค่าตัวแปรของขั้นตอนวิธีย่อยกับขั้นตอนวิธีหลักด้วย

<u>**ขั้นตอนวิธีที่ 3.3 การหาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการ (1.1)**</u>

ขั้นเริ่มต้น ให้ $\varepsilon_1 > 0$ เป็นค่าสเกลาร์ในการหยุด โดยเลือกจุดเริ่มต้น $\mathbf{x}_0$ และค่าพารามิเตอร์ทัณฑ์ $\rho_0 > 0$ และ สเกลาร์ $\beta > 1$ ให้ $t = 0$ และไปยังขั้นหลัก

ขั้นหลัก

1. ใช้จุดเริ่มต้น $\mathbf{x}_t$ ในการแก้ปัญหาต่อไปนี้

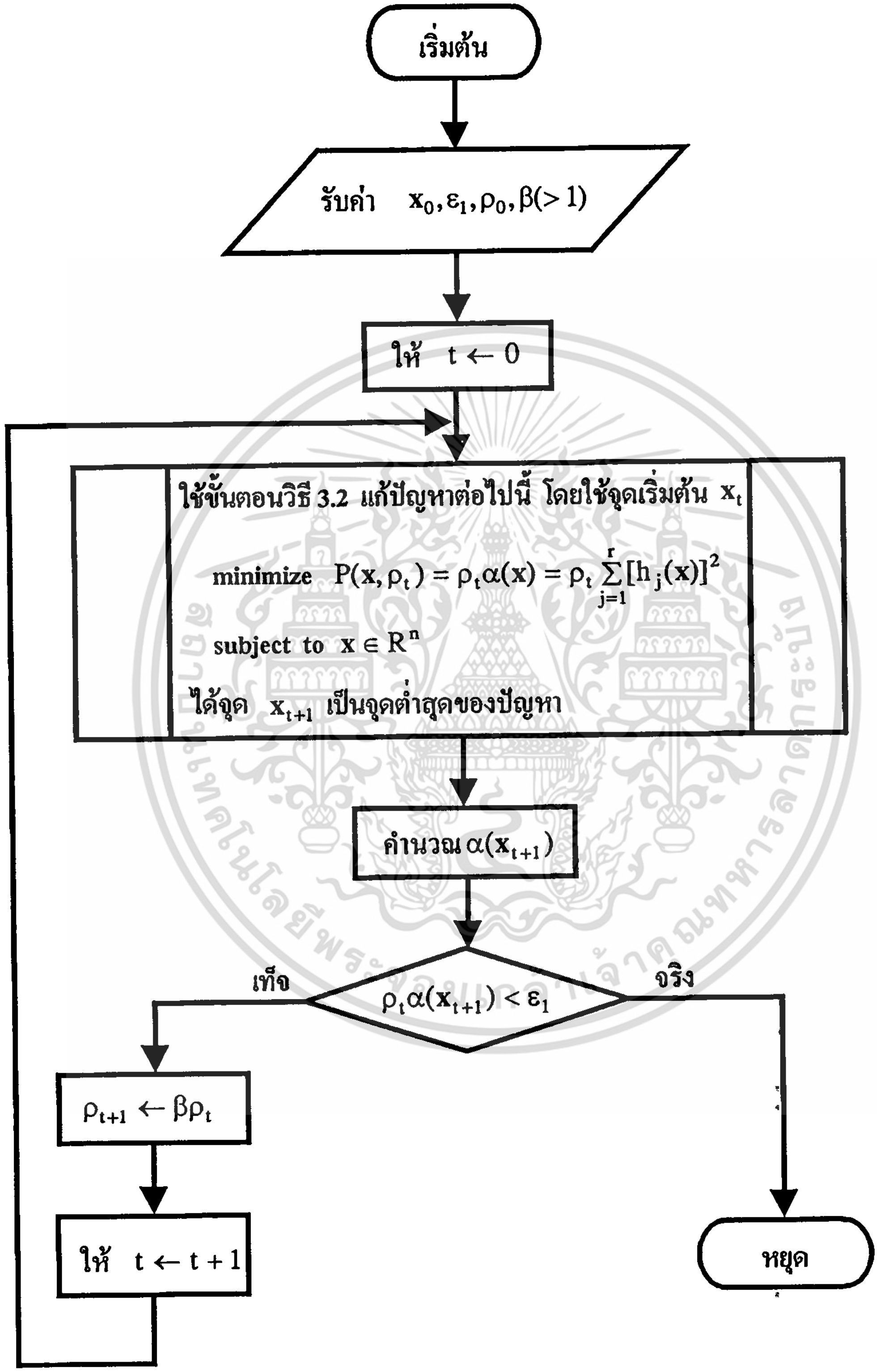
$$\begin{aligned} &\text{minimize} && P(\mathbf{x}, \rho_t) = \rho_t \alpha(\mathbf{x}) = \rho_t\{[h_1(\mathbf{x})]^2 + \ldots + [h_r(\mathbf{x})]^2\} \\ &\text{subject to} && \mathbf{x} \in R^n \end{aligned}$$

โดยใช้ขั้นตอนวิธีที่ 3.2 จะได้ $\mathbf{x}_{t+1}$ เป็นจุดต่ำสุดของปัญหาข้างบน

2. ถ้า $\rho_t \alpha(\mathbf{x}_{t+1}) < \varepsilon_1$ จะได้ $\mathbf{x}_{t+1}$ เป็นจุดที่เป็นไปได้ (สอดคล้องกับสมการทั้งหมด) และให้หยุดดำเนินการ

ถ้าเป็นอย่างอื่น ให้ $\rho_{t+1} = \beta\rho_t$, ให้ $t = t + 1$ และไปยังขั้นที่ 1

จากขั้นตอนวิธีที่ 3.3 สามารถเขียนแผนผังการทำงานได้ดังรูปที่ 3.4



รูปที่ 3.4 แสดงแผนผังการทำงานของขั้นตอนวิธีที่ 3.3

3.3.3 ขั้นตอนวิธีในการหาจุดภายในของบริเวณที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการ (1.2)

ก่อนที่จะให้ขั้นตอนวิธีที่ 3.4 สำหรับการหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการ (1.2)

$$g_i(\mathbf{x}) \leq 0 \quad ; i = 1,\ldots,m$$

ผู้วิจัยจะเสนอขั้นตอนวิธี OPTIM 1 และ OPTIM 2 ที่จะใช้ในขั้นตอนวิธีที่ 3.4 ซึ่งขั้นตอน OPTIM 1 เป็นขั้นตอนที่ใช้แก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ โดยมีการตรวจสอบเครื่องหมายของค่าฟังก์ชันจุดประสงค์ ถ้าฟังก์ชันจุดประสงค์มีค่าเป็นลบที่จุดใด ๆ ให้สามารถออกจากขั้นตอน OPTIM 1 ได้โดยไม่ต้องดำเนินการจนถึงจุดต่ำสุด ส่วนขั้นตอน OPTIM 2 เป็นขั้นตอนที่ใช้แก้ปัญหาค่าต่ำสุดในกรณีที่มีเงื่อนไขบังคับเป็นแบบอสมการโดยวิธีฟังก์ชันกั้นเขต และทำการตรวจสอบเช่นเดียวกับ OPTIM 1

<u>ขั้นตอนวิธี OPTIM 1</u>

เหมือนกับขั้นตอนวิธีที่ 3.2 เพียงแต่เปลี่ยนการดำเนินการ 2.3 ในขั้นตอนวิธีที่ 3.2 เป็น

"2.3 ถ้า $f(\mathbf{x}_{k+1}) < 0$ จะได้ $\mathbf{x}_{k+1}$ เป็นจุดภายในของเงื่อนไขบังคับ $f(\mathbf{x}) \leq 0$ และให้หยุดดำเนินการ

ถ้าเป็นอย่างอื่น คำนวณ $\nabla f(\mathbf{x}_{k+1})$ และไปยังขั้นที่ 3"

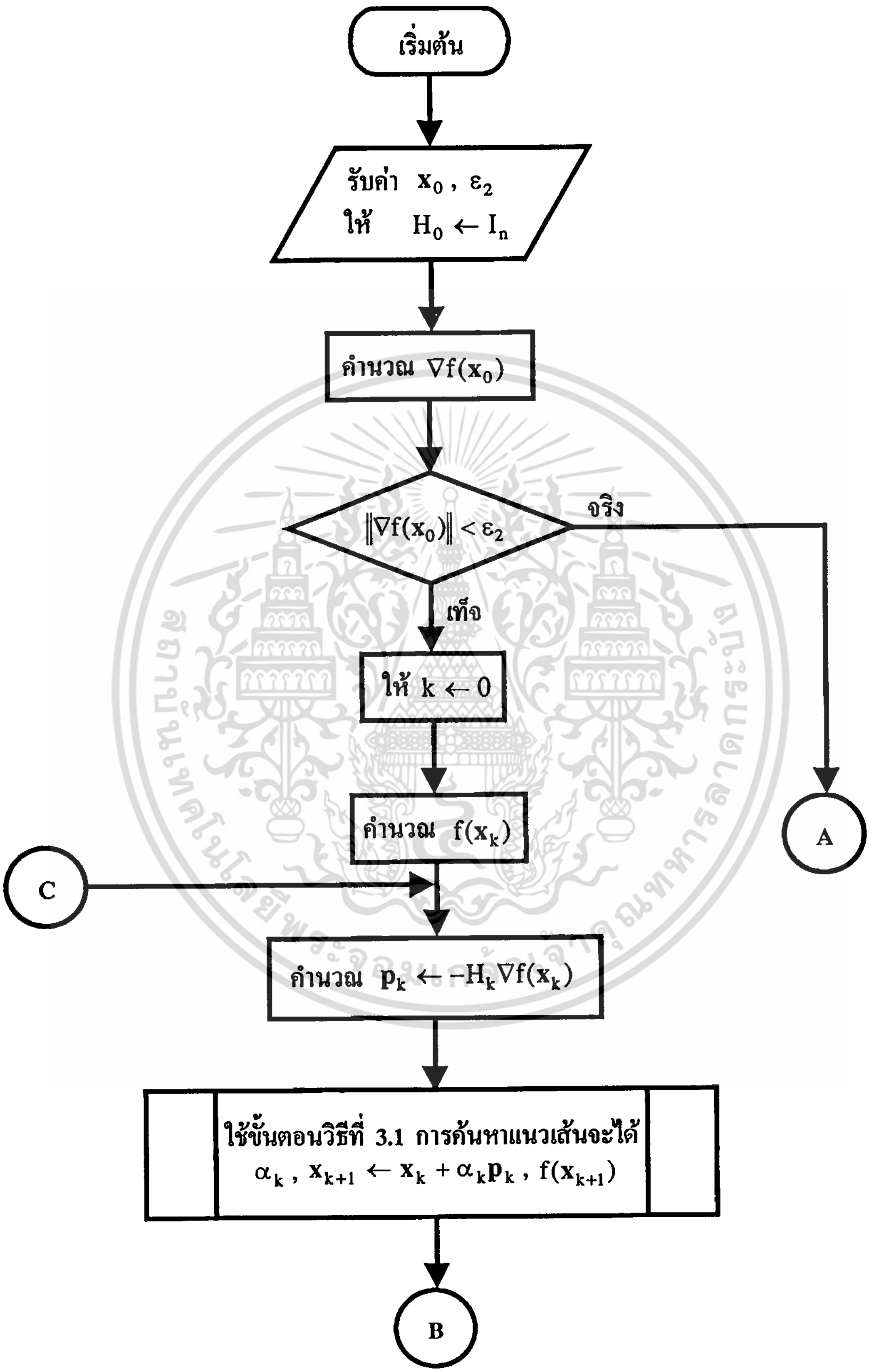
เพื่อแสดงรายละเอียดของขั้นตอนวิธี OPTIM 1 ให้ชัดเจน จึงเขียนแผนผังการทำงานได้ดังรูปที่ 3.5

ขั้นตอนวิธี OPTIM 2 ใช้แก้ปัญหาที่อยู่ในรูปแบบต่อไปนี้ เมื่อ T เป็นเซตดัชนีของเงื่อนไขบังคับ

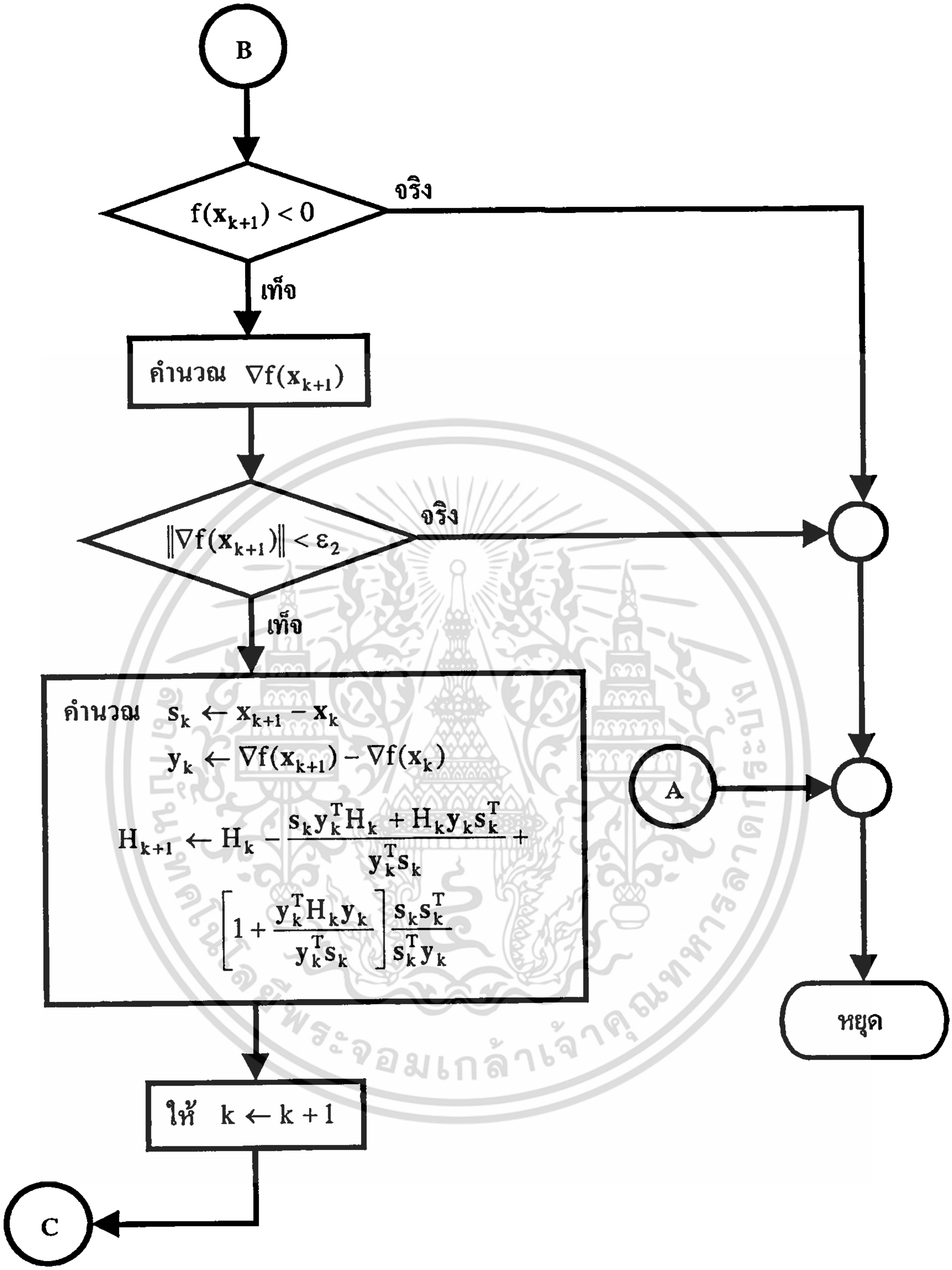
$$\begin{aligned} &\text{minimize} && f_1(\mathbf{x}) \\ &\text{subject to} && g_i(\mathbf{x}) \leq 0 \quad , \forall i \in T \\ &&& \mathbf{x} \in R^n \end{aligned}$$

โดยใช้วิธีฟังก์ชันกั้นเขตแปลงปัญหาเป็น

$$\begin{aligned} &\text{minimize} && U(\mathbf{x}, \mu_k) = f_1(\mathbf{x}) - \mu_k \left[\sum_{i \in T} \frac{1}{g_i(\mathbf{x})} \right] \\ &\text{subject to} && g_i(\mathbf{x}) < 0 \quad , \forall i \in T \\ &&& \mathbf{x} \in R^n \end{aligned}$$



รูปที่ 3.5 แสดงแผนผังการทำงานของขั้นตอนวิธี OPTIM 1



รูปที่ 3.5 (ต่อ)

ดังนั้นขั้นตอนวิธี OPTIM 2 จะเป็นดังนี้

<u>ขั้นตอนวิธี OPTIM 2</u>

ขั้นเริ่มต้น ให้ $\varepsilon_1 > 0$ เป็นค่าสเกลาร์ในการหยุด กำหนดจุดเริ่มต้น x_0 ที่ทำให้

$g_i(\mathbf{x}_0) < 0$, $\forall i \in T$ และค่าพารามิเตอร์กั้นเขต $\mu_0 > 0$ และสเกลาร์ $\gamma \in (0,1)$ ให้ $t = 0$ และไปยังขั้นหลัก

ขั้นหลัก

1. ใช้จุดเริ่มต้น $\mathbf{x}_t$ ในการแก้ปัญหาต่อไปนี้

$$\left.\begin{array}{ll} \text{minimize} & U(\mathbf{x},\mu_t) = f_1(\mathbf{x}) + \mu_t\beta(\mathbf{x}) = f_1(\mathbf{x}) + \mu_t\left[-\sum_{i\in T}\frac{1}{g_i(\mathbf{x})}\right] \\ \text{subject to} & g_i(\mathbf{x}) < 0 \quad , \forall i \in T \\ & \mathbf{x} \in R^n \end{array}\right\} \quad (3.15)$$

โดยใช้ขั้นตอนวิธีที่ 3.2* (ให้ $U(\mathbf{x},\mu_t)$ เป็นฟังก์ชันจุดประสงค์ $f(\mathbf{x})$ ในขั้นตอนวิธีที่ 3.2) ให้ $\mathbf{x}_{t+1}$ เป็นจุดที่ได้หลังจากการเรียกใช้ขั้นตอนวิธีที่ 3.2*

ถ้าจุด $\mathbf{x}_{t+1}$ ทำให้ $f_1(\mathbf{x}_{t+1}) < 0$ จะได้ว่า $\mathbf{x}_{t+1}$ เป็นจุดภายในของเงื่อนไขบังคับ $f_1(\mathbf{x}) \le 0$ และ $g_i(\mathbf{x}) \le 0$, $\forall i \in T$ และให้หยุดดำเนินการ

ถ้าเป็นอย่างอื่น จะได้ว่า $\mathbf{x}_{t+1}$ เป็นจุดต่ำสุดของปัญหา (3.15) และไปยังขั้นที่ 2

2. ถ้า $\mu_t\beta(\mathbf{x}_{t+1}) < \varepsilon_1$ จะได้จุด $\mathbf{x}_{t+1}$ เป็นจุดต่ำสุดของปัญหาเริ่มต้นและให้หยุดดำเนินการ

ถ้าเป็นอย่างอื่น ให้ $\mu_{t+1} = \gamma\mu_t$, $t = t + 1$ และไปยังขั้นที่ 1

* เนื่องจากการแก้ปัญหา (3.15) มีลักษณะพิเศษคือ ต้องการรักษาความเป็นจุดภายในของเงื่อนไขบังคับเอาไว้ ดังนั้นจึงทำการเปลี่ยนแปลง 2.3 ในขั้นตอนวิธีที่ 3.2 ดังนี้

"2.3 ถ้า $g_i(\mathbf{x}_{k+1}) < 0$, $\forall i \in T$ ให้ไป 2.4

ถ้าเป็นอย่างอื่น ให้ $\alpha_k = \alpha_k / 2$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{p}_k$$

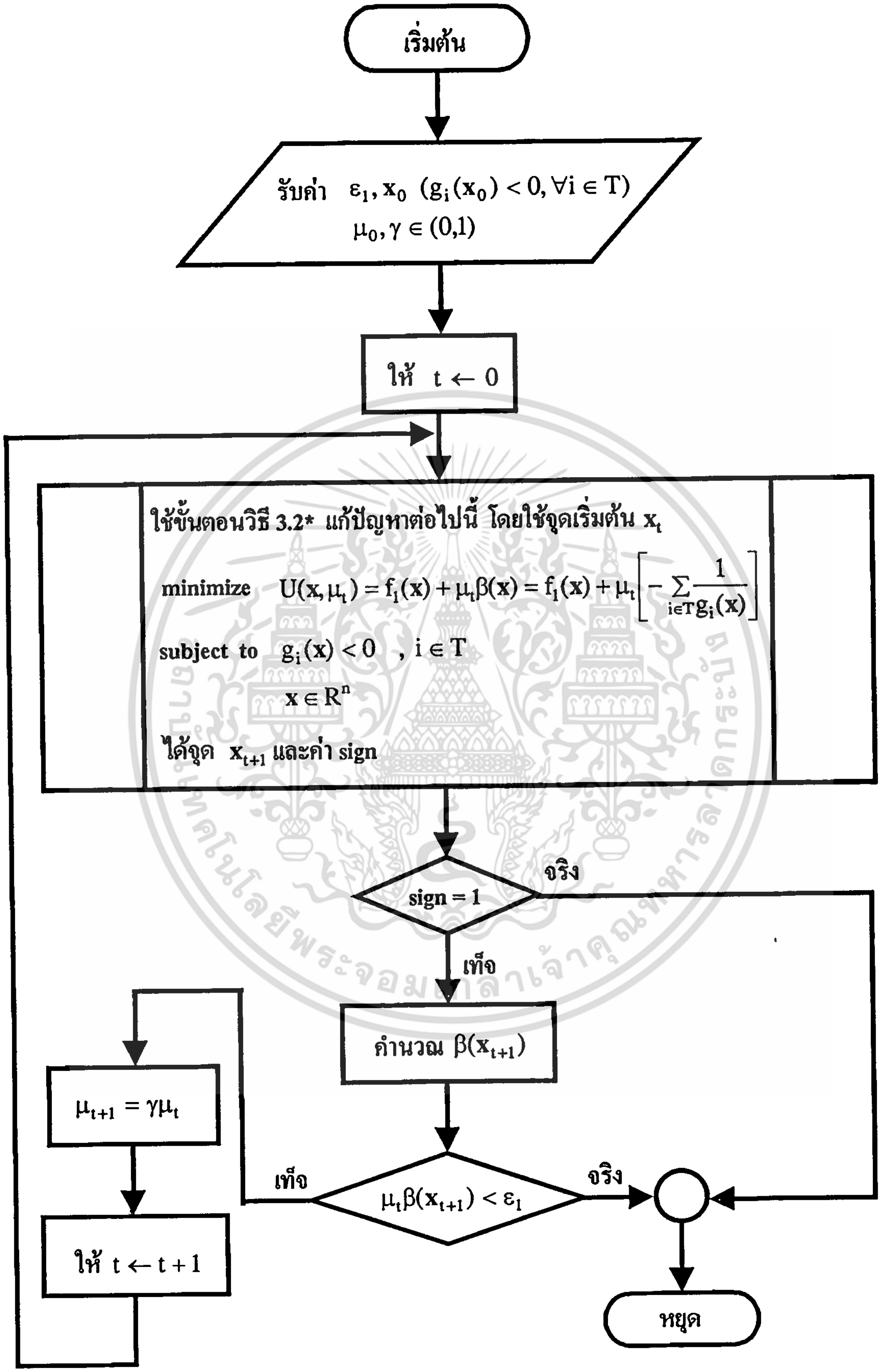
คำนวณ $g_i(\mathbf{x}_{k+1})$, $\forall i \in T$ และไปยัง 2.3

2.4 คำนวณ $f_1(\mathbf{x}_{k+1})$

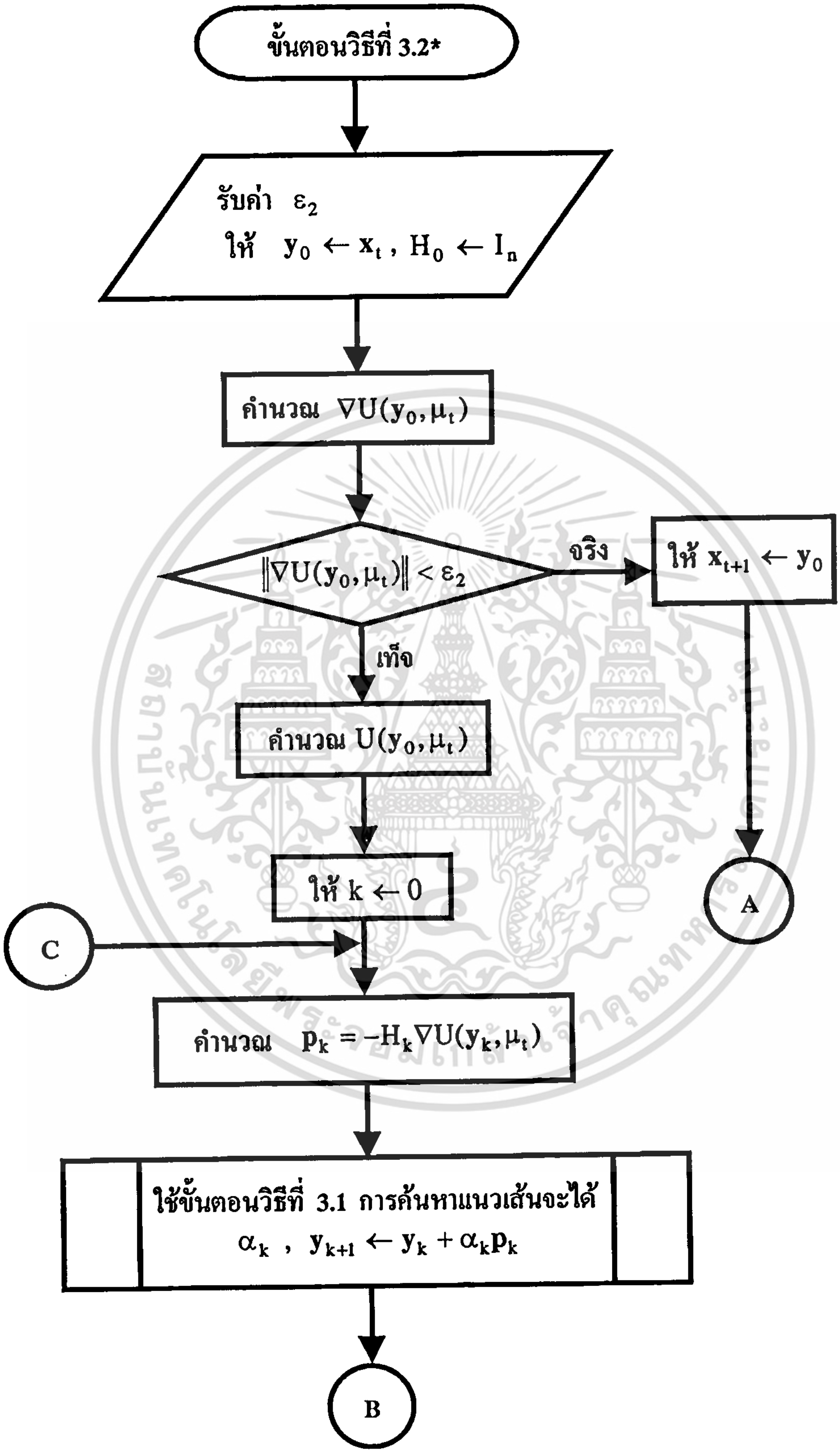
ถ้า $f_1(\mathbf{x}_{k+1}) < 0$ จะได้ว่า $\mathbf{x}_{k+1}$ เป็นจุดภายในของเงื่อนไขบังคับ $f_1(\mathbf{x}) \le 0$ และ $g_i(\mathbf{x}) \le 0$, $\forall i \in T$ และให้หยุดดำเนินการ

ถ้าเป็นอย่างอื่น คำนวณ $\nabla U(\mathbf{x}_{k+1},\mu_t), U(\mathbf{x}_{k+1},\mu_t)$ และไปยังขั้นที่ 3"

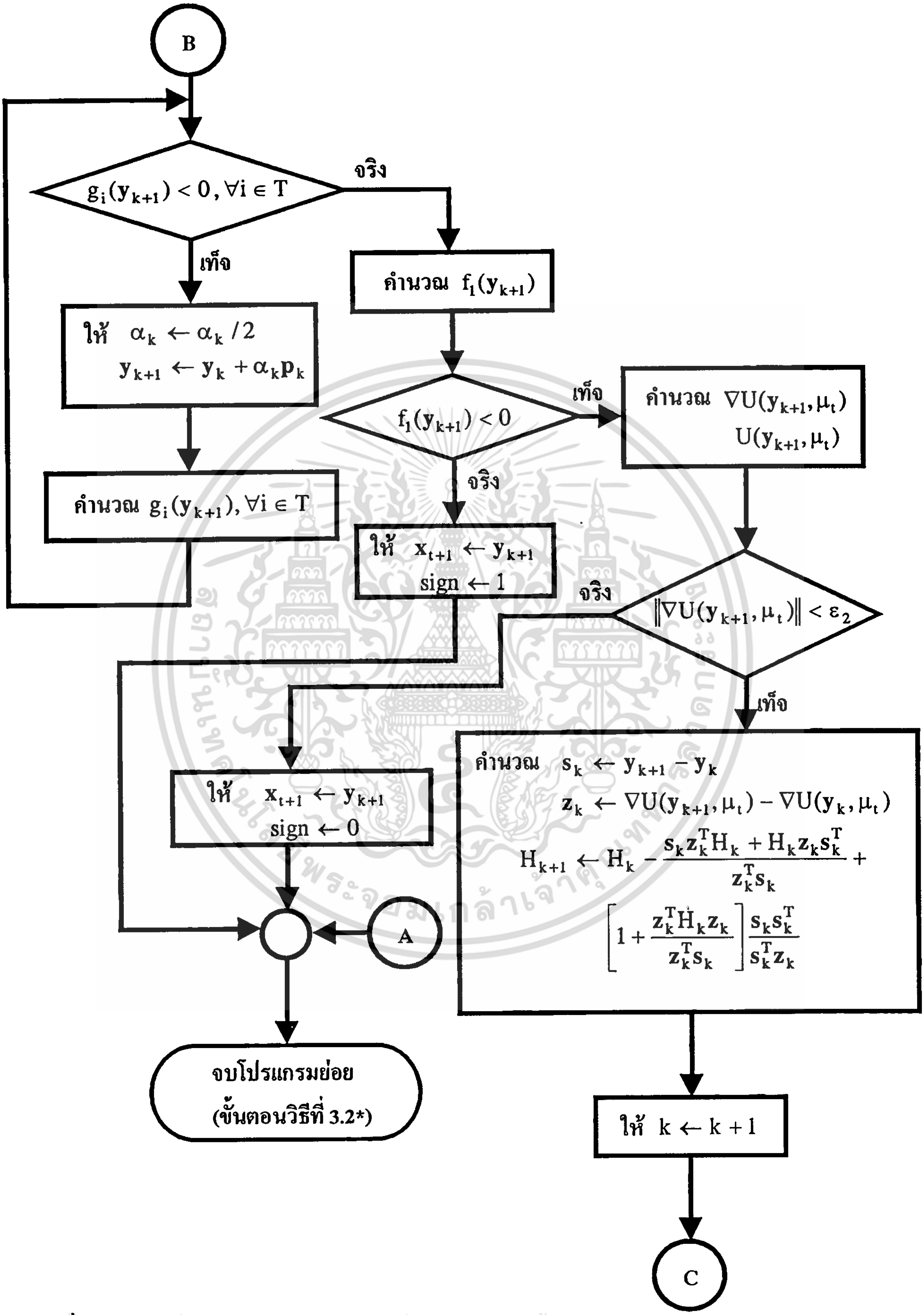
ในรูปที่ 3.6 แสดงแผนผังการทำงานของขั้นตอนวิธี OPTIM 2 (ขั้นตอนวิธีหลัก) และเนื่องจากมีบางส่วนของขั้นตอนวิธีที่ 3.2* (ขั้นตอนวิธีย่อย) ซึ่งแตกต่างจากขั้นตอนวิธีที่ 3.2 ดังนั้นจึงทำการแสดงแผนผังการทำงานของขั้นตอนวิธีที่ 3.2* ไว้ในรูปที่ 3.7 จากรูปที่ 3.6 และ 3.7 ได้แสดงให้เห็นถึงการส่งและรับค่าตัวแปรระหว่างขั้นตอนวิธีหลัก(ลูปใหญ่) กับขั้นตอนวิธีย่อย(ลูปเล็ก) และ



รูปที่ 3.6 แสดงแผนผังการทำงานของขั้นตอนวิธี OPTIM 2



รูปที่ 3.7 แสดงแผนผังการทำงานของขั้นตอนวิธีที่ 3.2*


B
$g_i(y_{k+1}) < 0, \forall i \in T$
จริง
เท็จ
คำนวณ $f_1(y_{k+1})$
ให้ $\alpha_k \leftarrow \alpha_k / 2$
$y_{k+1} \leftarrow y_k + \alpha_k p_k$
$f_1(y_{k+1}) < 0$
เท็จ
คำนวณ $\nabla U(y_{k+1}, \mu_t)$
$U(y_{k+1}, \mu_t)$
จริง
คำนวณ $g_i(y_{k+1}), \forall i \in T$
ให้ $x_{t+1} \leftarrow y_{k+1}$
sign $\leftarrow 1$
จริง
$\|\nabla U(y_{k+1}, \mu_t)\| < \varepsilon_2$
เท็จ
ให้ $x_{t+1} \leftarrow y_{k+1}$
sign $\leftarrow 0$
คำนวณ $s_k \leftarrow y_{k+1} - y_k$
$z_k \leftarrow \nabla U(y_{k+1}, \mu_t) - \nabla U(y_k, \mu_t)$
$H_{k+1} \leftarrow H_k - \dfrac{s_k z_k^T H_k + H_k z_k s_k^T}{z_k^T s_k} + \left[1 + \dfrac{z_k^T H_k z_k}{z_k^T s_k}\right] \dfrac{s_k s_k^T}{s_k^T z_k}$
A
จบโปรแกรมย่อย
(ขั้นตอนวิธีที่ 3.2*)
ให้ $k \leftarrow k + 1$
C

รูปที่ 3.7 (ต่อ)

เพื่อไม่ให้สับสนในตัวแปรจะใช้ตัวแปร y ในลูปเล็กและตัวแปร x ในลูปใหญ่ จะเห็นว่าในขั้นตอนวิธี OPTIM 1 และ OPTIM 2 เราไม่มีความจำเป็นที่จะดำเนินการจนได้จุดต่ำสุดของปัญหา ถ้าจุดที่กำเนิดขึ้นในระหว่างกระบวนการ ทำให้ค่าฟังก์ชันจุดประสงค์ $f(x)$ ของปัญหาเดิมเป็นลบได้ จุดที่กำเนิดได้นั้นก็จะเป็นจุดภายในของเงื่อนไข $f(x) \le 0$ และเงื่อนไขบังคับเดิมที่มีอยู่ ขั้นตอนวิธีที่ 3.4 ที่จะกล่าวต่อไปนี้เป็นขั้นตอนการทำซ้ำตัวเองของวิธีฟังก์ชันกั้นเขตเพื่อหาจุดภายในของบริเวณที่เป็นไปได้ของเงื่อนไขบังคับ $g_i(x) \le 0 \;\; ; i = 1,\ldots,m$ และเนื่องจากเป็นการเรียกใช้ขั้นตอนวิธีย่อย OPTIM 1 และOPTIM 2 ดังนั้นจะไม่กล่าวถึงการส่งและรับค่าตัวแปรและรายละเอียดของขั้นตอนวิธีดังกล่าวซ้ำอีก รวมทั้งจะไม่คำนึงถึงการซ้ำกันของตัวแปรของเลขดัชนีในขั้นตอนวิธีต่าง ๆ ที่ได้กล่าวมาแล้ว

สำหรับขั้นตอนวิธีในการหาจุด $x^* \in \{x \in R^n \mid g_i(x) < 0 \;\; ; i = 1,\ldots,m\}$ เป็นดังนี้

<u>ขั้นตอนวิธีที่ 3.4 การหาจุดภายในของบริเวณที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการ (1.2)</u>

ขั้นเริ่มต้น

กำหนดจุด $x_0 \in R^n$ และให้

$T_{x_0} = \{i \mid g_i(x_0) < 0 \;\; ; 1 \le i \le m\}$ ซึ่ง $n(T_{x_0}) = m_{x_0}$

$F_{x_0} = \{j \mid g_j(x_0) \ge 0 \;\; ; 1 \le j \le m\}$ ซึ่ง $n(F_{x_0}) = n_{x_0}$

$f_{x_0} = \{f(r) \in F_{x_0} \mid r = 1,\ldots,n_{x_0}$ และ $f(1) < f(2) < \ldots < f(r) < \ldots < f(n_{x_0})\}$

ถ้า $m_{x_0} = m$ จะได้จุด x_0 เป็นจุดภายในของเงื่อนไขบังคับ (1.2) และหยุดดำเนินการ

ถ้าเป็นอย่างอื่น ให้ $k = 0$ และไปยังขั้นหลัก

ขั้นหลัก

1. ใช้จุด x_k แก้ปัญหาต่อไปนี้

$$\left.\begin{array}{lll} \text{minimize} & g_{f(1)}(x) & \text{ซึ่ง } f(1) \in f_{xk} \\ \text{subject to} & g_i(x) \le 0 & ; \quad i \in T_{xk} \\ & x \in R^n & \end{array}\right\} \quad (P_k)$$

1.1. ถ้า $T_{xk} = \phi$ ให้ใช้ขั้นตอนวิธี OPTIM 1 กำเนิดจุด x_{k+1}

1.1.1 ถ้า x_{k+1} ไม่ใช่จุดต่ำสุดวงกว้างของปัญหา (P_k) แต่ทำให้ $g_{f(1)}(x_{k+1}) < 0$ ให้ไปขั้นที่ 2

1.1.2 ถ้า x_{k+1} เป็นจุดต่ำสุดวงกว้างของปัญหา (P_k)

1.1.2.1 ถ้า $g_{f(1)}(x_{k+1}) \ge 0$ แล้ว ไม่มีจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการ (1.2) และให้หยุดดำเนินการ

1.1.2.2 ถ้า $g_{f(1)}(x_{k+1}) < 0$ ให้ไปขั้นที่ 2

1.2. ถ้า $T_{xk} \neq \phi$ ให้ใช้ขั้นตอนวิธี OPTIM 2 กำเนิดจุด x_{k+1}

1.2.1 ถ้า x_{k+1} ไม่ใช่จุดต่ำสุดวงกว้างของปัญหา (P_k) แต่ทำให้ $g_{f(1)}(x_{k+1}) < 0$ ให้ไปขั้นที่ 2

1.2.2 ถ้า x_{k+1} เป็นจุดต่ำสุดวงกว้างของปัญหา (P_k)

1.2.2.1 ถ้า $g_{f(1)}(x_{k+1}) \geq 0$ แล้ว ไม่มีจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการ (1.2) และให้หยุดดำเนินการ

1.2.2.2 ถ้า $g_{f(1)}(x_{k+1}) < 0$ ให้ไปขั้นที่ 2

2. ให้ $T_{xk+1} = T_{xk} \cup \{f(1)\}$ และ $n(T_{xk+1}) = m_{xk+1}$

$F_{xk+1} = F_{xk} - \{f(1)\}$ และ $n(F_{xk+1}) = n_{xk+1}$

$f_{xk+1} = \{f(r) \in F_{xk+1} \mid r = 1,\ldots,n_{xk+1}$ และ $f(1) < f(2) < \ldots < f(r) < \ldots < f(n_{xk+1})\}$ ไปยังขั้นที่ 3

3.

3.1 ถ้า $m_{xk+1} = m$ จะได้จุด x_{k+1} เป็นจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับอสมการ (1.2) และให้หยุดดำเนินการ

3.2 ถ้า $m_{xk+1} \neq m$ หาค่าของ $g_{f(1)}(x_{k+1})$ ซึ่ง $f(1) \in f_{xk+1}$

3.2.1 ถ้า $g_{f(1)}(x_{k+1}) < 0$ ให้ $k = k+1$ และไปยังขั้นที่ 2

3.2.2 ถ้า $g_{f(1)}(x_{k+1}) \geq 0$ ให้ $k = k+1$ และไปยังขั้นที่ 1

จากขั้นตอนวิธีที่ 3.4 สามารถเขียนแผนผังการทำงานได้ดังรูปที่ 3.8

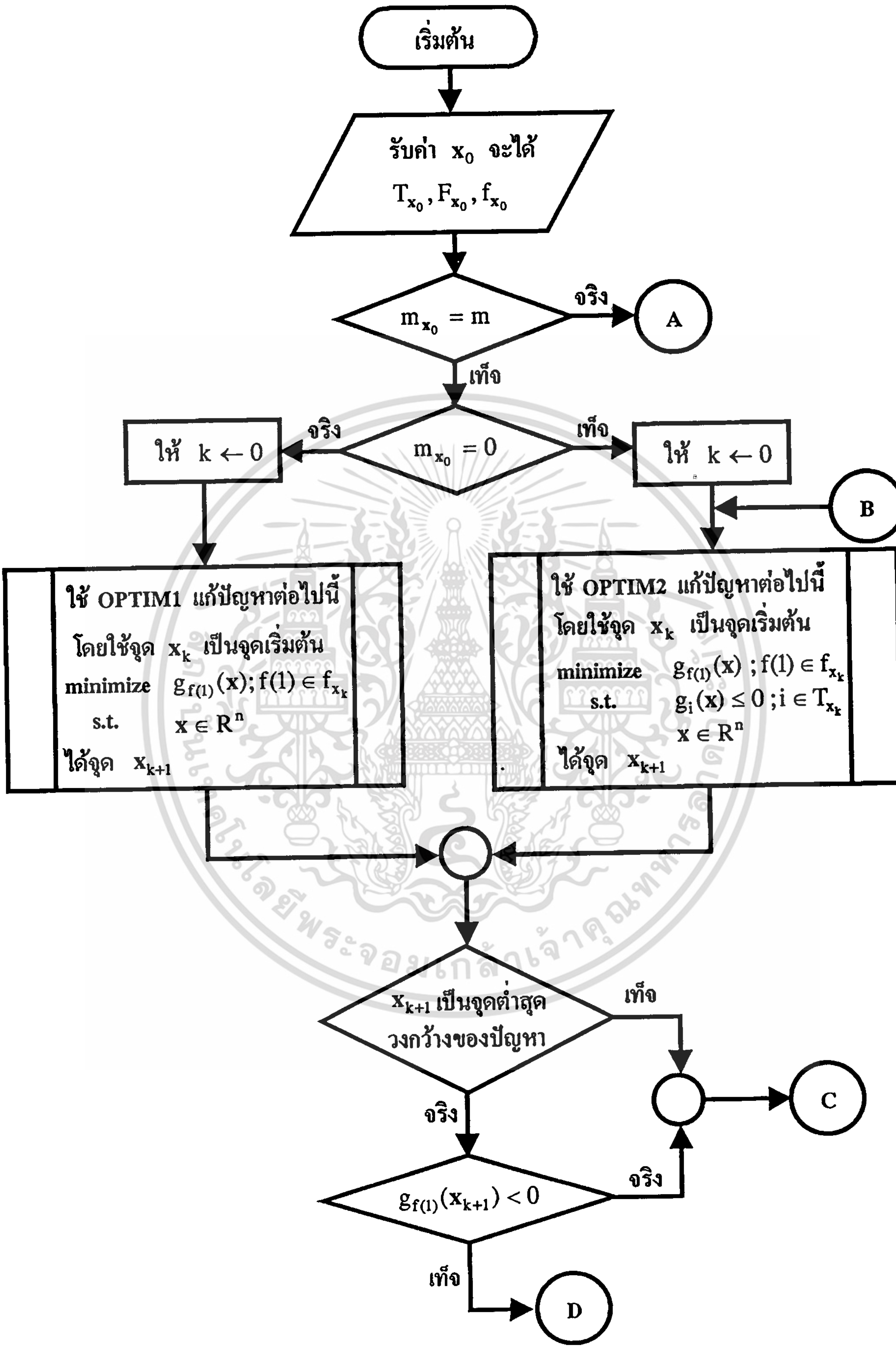
3.3.4 ขั้นตอนวิธีในการหาจุดภายที่เป็นไปได้ของเงื่อนไขบังคับแบบผสม (1.3)

ในการหา $x^* \in \{x \in R^n \mid g_i(x) < 0 \;;\; i = 1,\ldots,m$ และ $h_j(x) = 0 \;;\; j = 1,\ldots,r\}$ นั้น เราจะแบ่งการทำงานออกเป็น 2 ส่วน คือ ส่วนที่ 1 ดำเนินการหาจุด $x' \in \{x \in R^n$ $g_i(x) < 0 \;;\; i = 1,\ldots,m\}$ โดยใช้ขั้นตอนวิธีที่ 3.4 แล้วใช้จุด x' เป็นจุดเริ่มต้นในส่วนที่ 2 ซึ่งในส่วนนี้จะใช้วิธีผสมในการหาจุด x^* ดังนั้นก่อนที่จะให้ขั้นตอนวิธีที่ 3.5 เพื่อหาจุด x^* ดังกล่าวผู้วิจัยจะให้ขั้นตอนวิธีย่อย OPTIM 3 สำหรับการทำงานในส่วนที่ 2 นี้

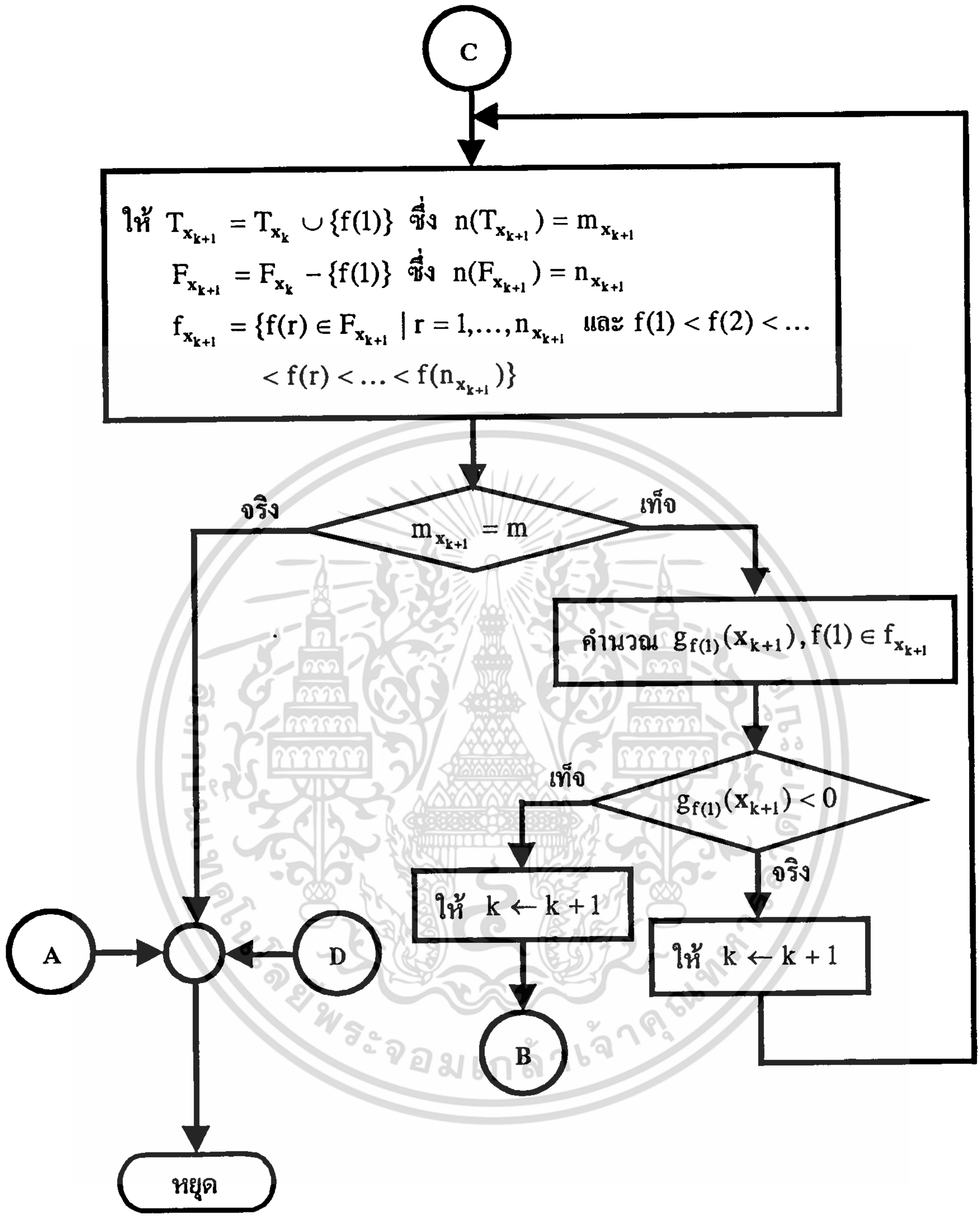
ขั้นตอนวิธี OPTIM 3 จะใช้แก้ปัญหาที่อยู่ในรูปแบบ

$$\text{minimize} \qquad f(x) = 0$$

$$\text{subject to} \qquad g_i(x) \leq 0 \;;\; i = 1,\ldots,m$$



รูปที่ 3.8 แสดงแผนผังการทำงานของขั้นตอนวิธีที่ 3.4



รูปที่ 3.8 (ต่อ)

$$h_j(\mathbf{x}) = 0 \quad ; \; j = 1,\dots,r$$
$$\mathbf{x} \in R^n$$

โดยใช้วิธีผสมแปลงปัญหาจะได้ปัญหาใหม่เป็น

$$\text{minimize} \qquad M(\mathbf{x},\mu_k) = f(\mathbf{x}) + \mu_k\beta(\mathbf{x}) + \frac{1}{\mu_k}\alpha(\mathbf{x}) \quad ; \; f(\mathbf{x}) = 0$$

$$= \mu_k\left[-\sum_{i=1}^{m}\frac{1}{g_i(\mathbf{x})}\right] + \frac{1}{\mu_k}\left[\sum_{j=1}^{r}(h_j(\mathbf{x}))^2\right]$$

subject to $\quad g_i(\mathbf{x}) < 0 \quad ; \; i = 1,\ldots,m$ และ $\mathbf{x} \in R^n$

<u>**ขั้นตอนวิธี OPTIM 3**</u>

ขั้นเริ่มต้น ให้ ε_1 เป็นค่าสเกลาร์ในการหยุด กำหนดจุดเริ่มต้น $\mathbf{x}_0$ ที่ทำให้ $g_i(\mathbf{x}_0) < 0$, $i = 1,\ldots,m$ และค่าพารามิเตอร์กั้นเขต $\mu_0 > 0$ และสเกลาร์ $\gamma \in (0,1)$ ให้ $t = 0$ และไปยังขั้นหลัก

ขั้นหลัก

1. ใช้จุดเริ่มต้น $\mathbf{x}_t$ ในการแก้ปัญหาต่อไปนี้

$$\left.\begin{array}{ll}\text{minimize} & M(\mathbf{x},\mu_t) = \mu_t\left[-\sum_{i=1}^{m}\frac{1}{g_i(\mathbf{x})}\right] + \frac{1}{\mu_t}\left[\sum_{j=1}^{r}(h_j(\mathbf{x}))^2\right] \\ \text{subject to} & g_i(\mathbf{x}) < 0 \quad ; \; i = 1,\ldots,m \\ & \mathbf{x} \in R^n \end{array}\right\} \quad (3.16)$$

โดยใช้ขั้นตอนวิธีที่ 3.2** (โดยให้ $M(\mathbf{x},\mu_t)$ เป็นฟังก์ชันจุดประสงค์ $f(\mathbf{x})$ ในขั้นตอนวิธีที่ 3.2) ให้ $\mathbf{x}_{t+1}$ เป็นจุดต่ำสุดของปัญหาข้างบน

2. ถ้า $\mu_t\beta(\mathbf{x}_{t+1}) < \varepsilon_1$ และ $\frac{1}{\mu_t}\alpha(\mathbf{x}_{t+1}) < \varepsilon_1$ จะได้ว่าจุด $\mathbf{x}_{t+1}$ เป็นจุดต่ำสุดของปัญหาเริ่มต้นและให้หยุดดำเนินการ

ถ้าเป็นอย่างอื่นให้ $\mu_{t+1} = \gamma\mu_t$, ให้ $t = t+1$ และไปยังขั้นที่ 1

**เนื่องจากการแก้ปัญหา (3.16) มีลักษณะพิเศษคือ การรักษาความเป็นจุดภายในของเงื่อนไขบังคับแบบอสมการเอาไว้ ดังนั้นจึงทำการเปลี่ยนแปลง 2.3 ในขั้นตอนวิธีที่ 3.2 ดังนี้

"2.3 ถ้า $g_i(\mathbf{x}_{k+1}) < 0$, $\forall i = 1,\ldots,m$ ให้คำนวณ $\nabla M(\mathbf{x}_{k+1},\mu_t), M(\mathbf{x}_{k+1},\mu_t)$ และไปยังขั้นที่ 3

ถ้าเป็นอย่างอื่น ให้ $\quad \alpha_k = \alpha_k / 2$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{p}_k$$

คำนวณ $g_i(\mathbf{x}_{k+1})$, $\forall i = 1,\ldots,m$ และไปยัง 2.3 "

จากขั้นตอนวิธี OPTIM 3 (ขั้นตอนวิธีหลัก) และขั้นตอนวิธีที่ 3.2** (ขั้นตอนวิธีย่อย) สามารถเขียนแผนผังการทำงานได้ดังรูปที่ 3.9 และ 3.10 ตามลำดับ ซึ่งแผนผังการทำงานดังกล่าวได้แสดงให้เห็นถึงการส่งและรับค่าตัวแปรของขั้นตอนวิธีหลักกับขั้นตอนวิธีย่อย เพื่อป้องกันความสับสนของตัวแปรจะใช้ตัวแปร $\mathbf{y}$ ในขั้นตอนวิธีย่อยและตัวแปร $\mathbf{x}$ ในขั้นตอนวิธีหลัก ต่อไปเราจะนำขั้นตอนวิธี OPTIM 3 นี้ไปใช้เป็นขั้นตอนวิธีย่อย ดังนั้นหากมีการเรียกใช้ขั้นตอนวิธี OPTIM 3 ก็จะไม่แสดงรายละเอียดของขั้นตอนวิธีและการส่ง-รับค่าตัวแปรซ้ำอีก

ขั้นตอนวิธีที่ 3.5 ที่จะกล่าวถึงต่อไปนี้ เป็นขั้นตอนวิธีที่ต้องการหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบผสม (อสมการและสมการ) โดยจะมีการเรียกใช้ขั้นตอนวิธีที่ 3.4 เป็นขั้นตอนย่อย เพื่อหาจุดภายในของเงื่อนไขบังคับแบบอสมการทั้งหมดก่อน ถ้าไม่สามารถหาจุดภายในได้ให้หยุดดำเนินการขั้นตอนวิธีที่ 3.5 ถ้าสามารถหาจุดภายในดังกล่าวได้ให้ดำเนินการต่อโดยเรียกใช้ขั้นตอนวิธีย่อย OPTIM 3 เพื่อหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับทั้งหมด

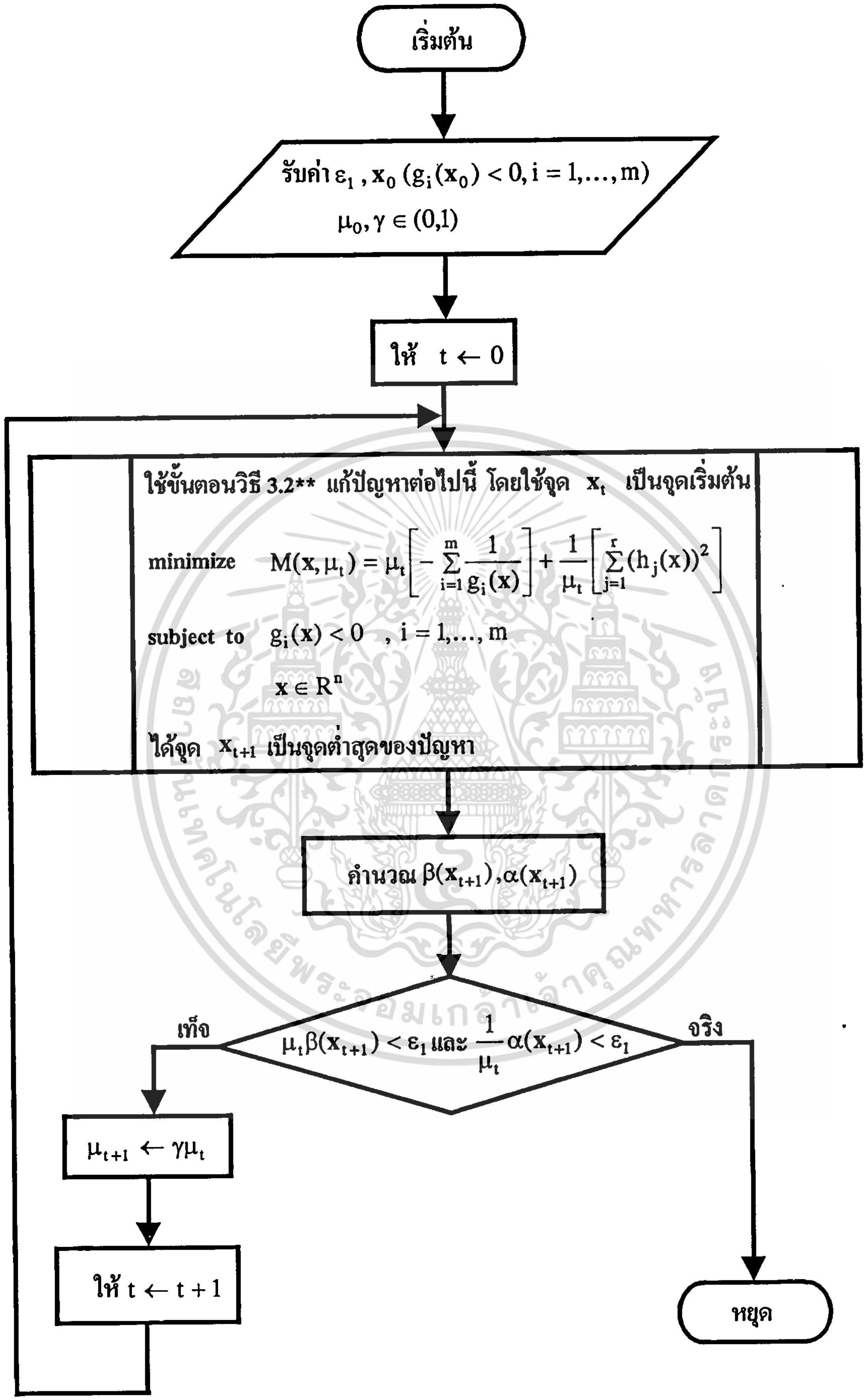
<u>ขั้นตอนวิธีที่ 3.5 การหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบผสม (1.3)</u>

1. กำเนิดจุด $\mathbf{x}' \in \{\mathbf{x} \in R^n \mid g_i(\mathbf{x}) < 0 \quad ; i = 1,\ldots,m\}$ โดยใช้ขั้นตอนวิธีที่ 3.4
   ถ้าไม่สามารถหาจุด $\mathbf{x}'$ ได้ ให้หยุดดำเนินการ
   ถ้าสามารถหาจุด $\mathbf{x}'$ ได้ ให้ไปยังขั้นที่ 2
2. ใช้ขั้นตอนวิธี OPTIM 3 ในการแก้ปัญหา

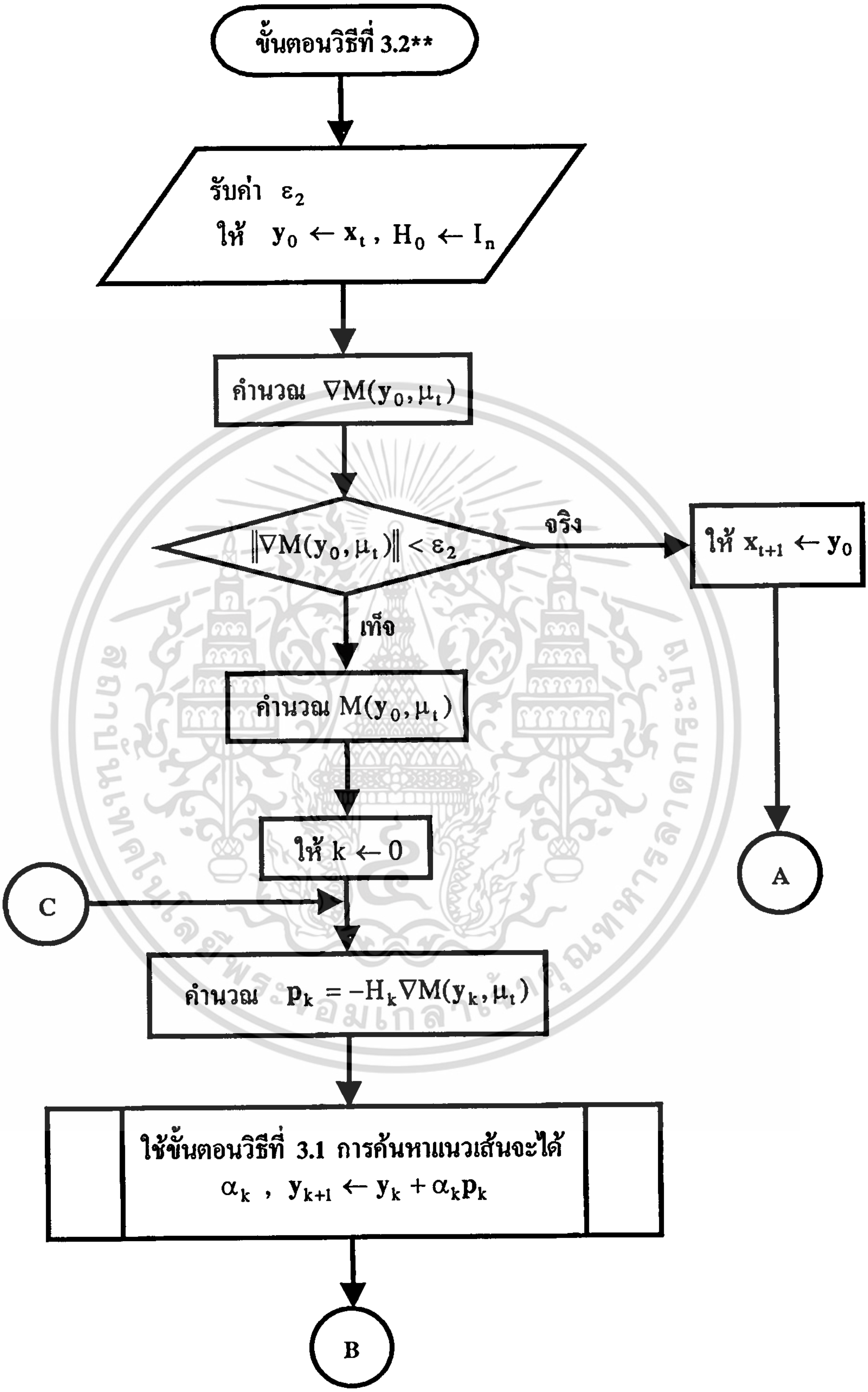
$$
\begin{array}{ll}
\text{minimize} & f(\mathbf{x}) = 0 \\
\text{subject to} & g_i(\mathbf{x}) \le 0 \quad ; i = 1,\ldots,m \\
 & h_j(\mathbf{x}) = 0 \quad ; j = 1,\ldots,r \\
 & \mathbf{x} \in R^n
\end{array}
$$

โดยกำหนดให้ $\mathbf{x}_0 = \mathbf{x}'$ (จากขั้นตอนที่ 1) เป็นจุดเริ่มต้นในการแก้ปัญหา ให้จุด $\mathbf{x}^*$ เป็นผลเฉลยที่ได้จากขั้นตอนวิธี OPTIM 3 แล้ว $\mathbf{x}^* \in \{\mathbf{x} \in R^n \mid g_i(\mathbf{x}) < 0 \quad ; i = 1,\ldots,m$ และ $h_j(\mathbf{x}) = 0 \quad ; j = 1,\ldots,r\}$

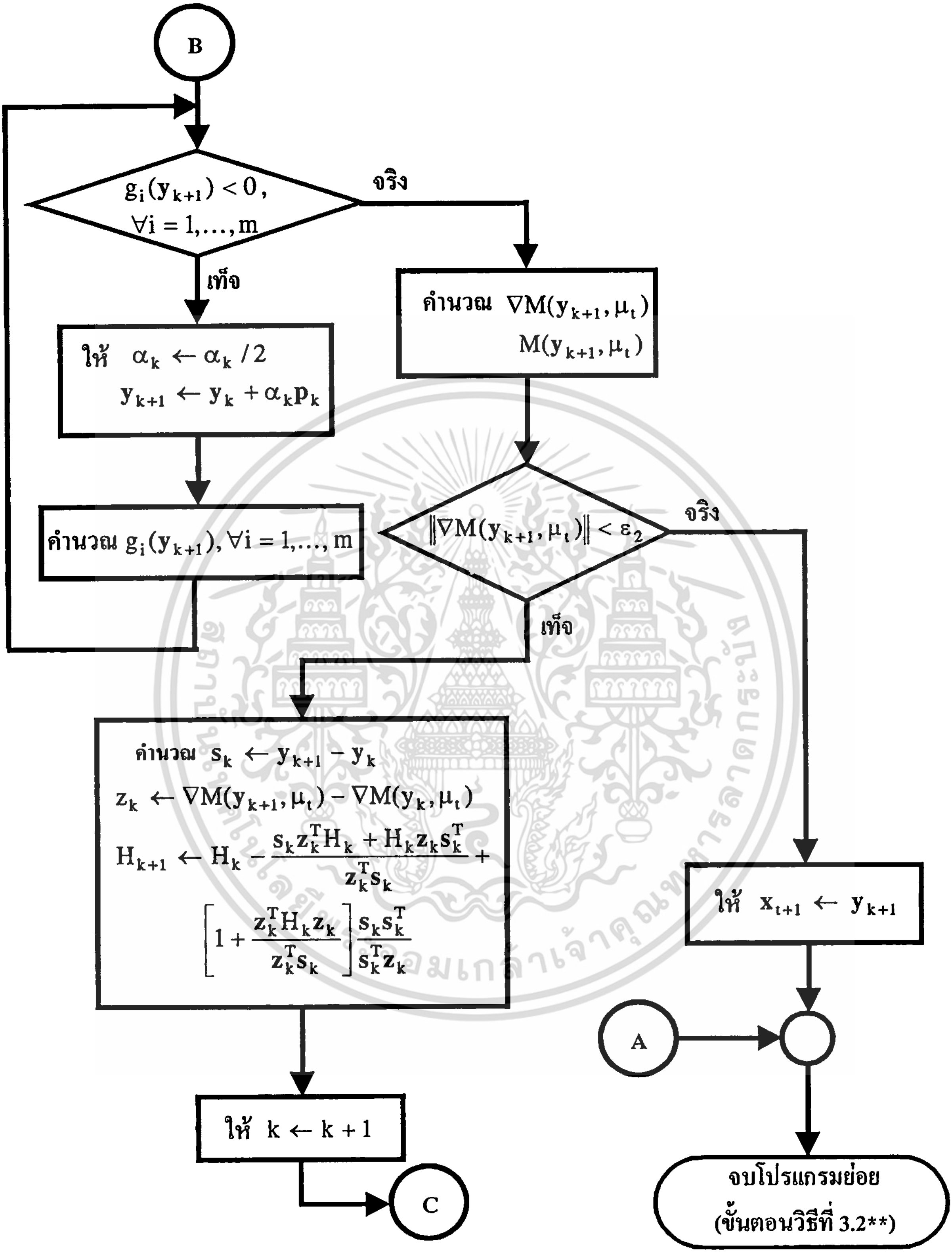
จากขั้นตอนวิธีที่ 3.5 สามารถเขียนแผนผังการทำงานได้ดังรูปที่ 3.11 จะเห็นว่าขั้นตอนวิธีที่ 3.5 เป็นขั้นตอนวิธีที่เรียกใช้ขั้นตอนวิธีย่อยที่ได้กล่าวมาแล้วในตอนต้น ดังนั้นเพื่อให้เห็นความเชื่อมโยงของขั้นตอนวิธีทั้งหมด ผู้วิจัยจึงเลือกซอสโค้ดของโปรแกรมที่ใช้ขั้นตอนวิธีที่ 3.5 มาแสดงไว้ในภาคผนวก ข.



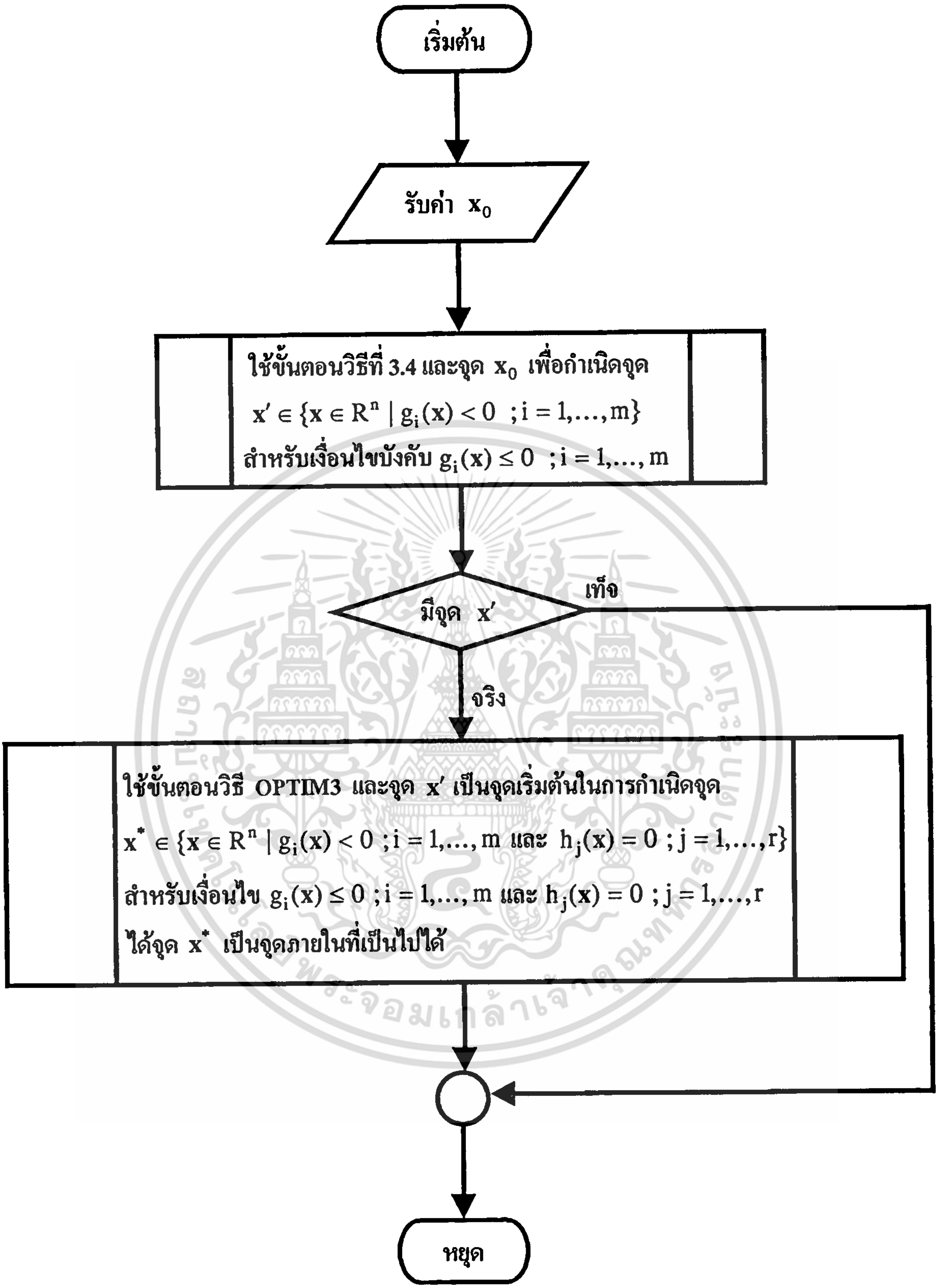
รูปที่ 3.9 แสดงแผนผังการทำงานของขั้นตอนวิธี OPTIM 3



รูปที่ 3.10 แสดงแผนผังการทำงานของขั้นตอนวิธีที่ 3.2**


B
$g_i(y_{k+1}) < 0,$
$\forall i = 1,\ldots,m$
จริง
เท็จ
ให้ $\alpha_k \leftarrow \alpha_k / 2$
$y_{k+1} \leftarrow y_k + \alpha_k p_k$
คำนวณ $g_i(y_{k+1}), \forall i = 1,\ldots,m$
คำนวณ $\nabla M(y_{k+1}, \mu_t)$
$M(y_{k+1}, \mu_t)$
$\|\nabla M(y_{k+1}, \mu_t)\| < \varepsilon_2$
จริง
เท็จ
คำนวณ $s_k \leftarrow y_{k+1} - y_k$
$z_k \leftarrow \nabla M(y_{k+1}, \mu_t) - \nabla M(y_k, \mu_t)$
$H_{k+1} \leftarrow H_k - \frac{s_k z_k^T H_k + H_k z_k s_k^T}{z_k^T s_k} + \left[1 + \frac{z_k^T H_k z_k}{z_k^T s_k}\right] \frac{s_k s_k^T}{s_k^T z_k}$
ให้ $k \leftarrow k+1$
C
ให้ $x_{t+1} \leftarrow y_{k+1}$
A
จบโปรแกรมย่อย
(ขั้นตอนวิธีที่ 3.2**)

รูปที่ 3.10 (ต่อ)



รูปที่ 3.11 แสดงแผนผังการทำงานของขั้นตอนวิธีที่ 3.5

บทที่ 4
ตัวอย่างปัญหาและขั้นตอนดำเนินการ

ในบทนี้จะให้ตัวอย่างปัญหาและขั้นตอนดำเนินการเพื่อหาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบต่าง ๆ ซึ่งได้แก่

1.ปัญหาการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการ

$$h_j(x) = 0 \quad , \quad j = 1,\dots,r$$

เมื่อ $\mathbf{x} \in R^n$ ซึ่งผู้วิจัยได้แบ่งปัญหาเป็น

1.1 ปัญหาในกรณีที่จำนวนสมการเท่ากับจำนวนตัวแปร $(r = n)$

1.2 ปัญหาในกรณีที่จำนวนสมการน้อยกว่าจำนวนตัวแปร $(r < n)$

1.3 ปัญหาในกรณีที่จำนวนสมการมากกว่าจำนวนตัวแปร $(r > n)$

โดยทั้ง 3 กรณี ผู้วิจัยได้แสดงปัญหาและขั้นตอนดำเนินการแก้ปัญหาไว้ในปัญหาที่ 1 , 2 และ 3 ตามลำดับ

2.ปัญหาการหาจุดภายในของบริเวณที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการ

$$g_i(x) \le 0 \quad , \quad i = 1,\dots,m$$

ผู้วิจัยได้แสดงปัญหาและขั้นตอนดำเนินการแก้ปัญหาไว้ในปัญหาที่ 4

3.ปัญหาการหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบผสม

$$g_i(x) \le 0 \quad , \quad i = 1,\dots,m \; ,$$
$$h_j(x) = 0 \quad , \quad j = 1,\dots,r$$

โดยแบ่งปัญหาเป็น

3.1 กรณีเงื่อนไขบังคับแบบผสมทั้งหมดเป็นฟังก์ชันเชิงเส้น

3.2 กรณีเงื่อนไขบังคับแบบผสมทั้งหมดเป็นฟังก์ชันนูน

3.3 กรณีเงื่อนไขบังคับแบบอสมการเป็นฟังก์ชันนูนและเงื่อนไขบังคับแบบสมการเป็นฟังก์ชันแบบใด ๆ

3.4 กรณีเงื่อนไขบังคับทั้งหมดเป็นฟังก์ชันแบบใด ๆ

โดยทั้ง 4 กรณี ผู้วิจัยได้แสดงปัญหาและขั้นตอนดำเนินการแก้ปัญหาไว้ในปัญหาที่ 5 , 6 , 7 และ 8 ตามลำดับ

4.1 ตัวอย่างปัญหาการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการ

ปัญหาที่ 1 เป็นปัญหากรณีที่จำนวนสมการเท่ากับจำนวนตัวแปร (5 สมการ , 5 ตัวแปร) หาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการต่อไปนี้ [14]

$$\left.\begin{aligned}
h_1(x) &= 2x_1 \sin x_2 - 7\cos x_2 = 0 \\
h_2(x) &= 2x_1 \sin x_3 - 5\cos x_3 = 0 \\
h_3(x) &= 2x_1 \sin x_4 - 3\cos x_4 = 0 \\
h_4(x) &= 2x_1 \sin x_5 - \cos x_5 \quad = 0 \\
h_5(x) &= \cos x_2 + \cos x_3 + \cos x_4 + \cos x_5 - 3 = 0
\end{aligned}\right\} \tag{4.1}$$

ขั้นตอนดำเนินการ

เราจะใช้วิธีฟังก์ชันทัณฑ์ในการหาจุดที่เป็นไปได้ โดยทำการแก้ปัญหาค่าต่ำสุดต่อไปนี้

$$\begin{aligned}
\text{minimize} \quad & f(x) = 0 \\
\text{subject to} \quad & h_1(x) = 2x_1 \sin x_2 - 7\cos x_2 = 0 \\
& h_2(x) = 2x_1 \sin x_3 - 5\cos x_3 = 0 \\
& h_3(x) = 2x_1 \sin x_4 - 3\cos x_4 = 0 \\
& h_4(x) = 2x_1 \sin x_5 - \cos x_5 \quad = 0 \\
& h_5(x) = \cos x_2 + \cos x_3 + \cos x_4 + \cos x_5 - 3 = 0 \\
& x \in R^5
\end{aligned}$$

เมื่อทำการแปลงปัญหาข้างต้นด้วยวิธีฟังก์ชันทัณฑ์ จะได้ปัญหาแบบไม่มีเงื่อนไขบังคับดังนี้

$$\begin{aligned}
\text{minimize} \quad & P(x,\rho_k) = f(x) + \rho_k\{\sum_{j=1}^{5}[h_j(x)]^2\} = \rho_k\{\sum_{j=1}^{5}[h_j(x)]^2\} \\
\text{subject to} \quad & x \in R^5
\end{aligned} \tag{4.2}$$

ทำการแก้ปัญหา (4.2) ด้วยวิธีแนวนิวตัน (สูตร BFGS ผกผัน) โดยทำการแก้ปัญหา (4.2) สำหรับแต่ละ ρ_k ซึ่งขั้นตอนวิธีที่ใช้แก้ปัญหาคือ ขั้นตอนวิธีที่ 3.3 โดยกำหนดเงื่อนไขเริ่มต้นดังนี้

ให้ $\varepsilon_1 = 0.000005$ เป็นค่าสเกลาร์ในการหยุด , เลือกจุดเริ่มต้น $x_0 = (0,1,0.5,0,1)^T$ และเลือก $\rho_0 = 0.1, \beta = 10$ ให้ $k = 0$ และเริ่มดำเนินการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับสำหรับแต่ละค่าของ ρ

พิจารณาการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับต่อไปนี้ที่ ρ_k ใด ๆ

$$\text{minimize} \qquad P(\mathbf{x},\rho_k) = \rho_k\{\sum_{j=1}^{5}[h_j(\mathbf{x})]^2\}$$

$$\text{subject to} \qquad \mathbf{x} \in R^5$$

โดยขั้นตอนวิธีที่ 3.2 ทำการกำหนดให้ $\varepsilon_2 = 0.000005$ เป็นค่าสเกลาร์ในการหยุด ต่อไปนี้จะกำหนดให้ $v_{k,i}$ แทนตัวแปรใด ๆในลูปภายใน i ซึ่งอยู่ในลูปใหญ่ k และ u_k แทนตัวแปรใด ๆ ในลูปใหญ่ k (สำหรับปัญหาอื่น ๆ ทั้งหมดก็จะทำการกำหนดตัวแปรในลักษณะเดียวกันนี้) และกำหนดให้ $H_{k,0} = I_5$ (เมตริกซ์เริ่มต้นสำหรับ ρ_k ใด ๆ)

ผู้วิจัยจะได้แสดงผลการคำนวณโดยละเอียดสำหรับที่ $k = 0, i = 0$ เพื่อทำการกำเนิดจุด $\mathbf{x}_1$ (สำหรับลูปภายใน $i = 1,2,\ldots$(ซึ่งใช้แก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ) และลูปใหญ่ $k = 1,2,\ldots$ สามารถทำได้ในทำนองเดียวกัน)

ที่ $k = 0, i = 0$

$$\begin{aligned}P(\mathbf{x}) = P(\mathbf{x},\rho_0) = 0.1[&(2x_1\sin x_2 - 7\cos x_2)^2 + (2x_1\sin x_3 - 5\cos x_3)^2 \\ &+ (2x_1\sin x_4 - 3\cos x_4)^2 + (2x_1\sin x_5 - \cos x_5)^2 \\ &+ (\cos x_2 + \cos x_3\cos x_4 + \cos x_5 - 3)^2]\end{aligned}$$

โดยที่ $$\nabla P(\mathbf{x}) = \left(\frac{\partial P(\mathbf{x})}{\partial x_1}, \frac{\partial P(\mathbf{x})}{\partial x_2}, \frac{\partial P(\mathbf{x})}{\partial x_3}, \frac{\partial P(\mathbf{x})}{\partial x_4}, \frac{\partial P(\mathbf{x})}{\partial x_5}\right)^T$$

ซึ่ง

$$\begin{aligned}\frac{\partial P(\mathbf{x})}{\partial x_1} &= 0.1[2(2x_1\sin x_2 - 7\cos x_2)(2\sin x_2) + 2(2x_1\sin x_3 - 5\cos x_3)(2\sin x_3) \\ &\quad + 2(2x_1\sin x_4 - 3\cos x_4)(2\sin x_4) + 2(2x_1\sin x_5 - \cos x_5)(2\sin x_5)] \\ &= 0.1[4\sin x_2(2x_1\sin x_2 - 7\cos x_2) + 4\sin x_3(2x_1\sin x_3 - 5\cos x_3) \\ &\quad + 4\sin x_4(2x_1\sin x_4 - 3\cos x_4) + 4\sin x_5(2x_1\sin x_5 - \cos x_5)]\end{aligned}$$

$$\begin{aligned}\frac{\partial P(\mathbf{x})}{\partial x_2} &= 0.1[2(2x_1\sin x_2 - 7\cos x_2)(2x_1\cos x_2 + 7\sin x_2) \\ &\quad + 2(\cos x_2 + \cos x_3 + \cos x_4 + \cos x_5 - 3)(-\sin x_2)]\end{aligned}$$

$$\begin{aligned}\frac{\partial P(\mathbf{x})}{\partial x_3} &= 0.1[2(2x_1\sin x_3 - 5\cos x_3)(2x_1\cos x_3 + 5\sin x_3) \\ &\quad + 2(\cos x_2 + \cos x_3 + \cos x_4 + \cos x_5 - 3)(-\sin x_3)]\end{aligned}$$

$$\frac{\partial P(x)}{\partial x_4} = 0.1[2(2x_1 \sin x_4 - 3\cos x_4)(2x_1 \cos x_4 + 3\sin x_4) + 2(\cos x_2 + \cos x_3 + \cos x_4 + \cos x_5 - 3)(-\sin x_4)]$$

$$\frac{\partial P(x)}{\partial x_5} = 0.1[2(2x_1 \sin x_5 - \cos x_5)(2x_1 \cos x_5 + \sin x_5) + 2(\cos x_2 + \cos x_3 + \cos x_4 + \cos x_5 - 3)(-\sin x_5)]$$

ให้ $x_{k,0}$ คือจุดเริ่มต้นเมื่อพิจารณาที่ ρ_k ผลการคำนวณที่จุด x_0 (ให้ $x_{0,0} = x_0$ เพราะกำลังคำนวณในลูป $i = 0$ ตามขั้นตอนวิธีที่ 3.2) เป็นดังนี้

$$\nabla P(x_{0,0}) = (-2.2963,-4.4485,-2.0997,-1.8000,-0.0839)^T$$

$$P(x_{0,0}) = 4.2852 \;,\; p_{0,0} = (2.2963,4.4485,2.0997,1.8000,0.0839)^T$$

เลือกใช้ $\mu = 1/3$ ในการหา $\alpha_{0,0}$ โดยขั้นตอนวิธีที่ 3.1 จะได้ $\alpha_{0,0} = 0.1250$

ดังนั้น $x_{0,1} = x_{0,0} + \alpha_{0,0}p_{0,0} = (0.2870,1.5561,0.7625,0,2250,1.0105)^T$

$\nabla P(x_{0,1}) = (-0.9650,0.8112,-2.3864,-1.9147,0.1177)^T$ และ

คำนวณ $H_{0,1}$ จากสูตร (3.1) ได้

$$H_{0,1} = \begin{bmatrix} 1.0244 & -0.1928 & 0.1568 & 0.1227 & -0.0128 \\ -0.1928 & 0.1615 & 0.0843 & 0.0495 & -0.0336 \\ 0.1568 & 0.0843 & 1.2664 & 0.2176 & -0.0068 \\ 0.1227 & 0.0495 & 0.2176 & 1.1774 & -0.0062 \\ -0.0128 & -0.0336 & -0.0068 & -0.0062 & 0.9990 \end{bmatrix}$$

เพราะว่า $\|\nabla P(x_{0,1})\| = 3.3114 > \varepsilon_2$ ดังนั้นให้ดำเนินการต่อตามขั้นตอนที่กล่าวข้างต้น เพื่อกำเนิด $x_{0,2}, x_{0,3}, \ldots$ จนได้จุด $x_{0,18} = x_1 = (2.2121,1.0072,0.8464,0.5959,0.2223)^T$ ซึ่ง $\|\nabla P(x_1)\| = 1.4673 \times 10^{-6} < \varepsilon_2$ และได้ค่า $\alpha(x_1) = \sum_{j=1}^{5}[h_j(x_1)]^2 = 1.1178 \times 10^{-12}$ ดังนั้น $\rho_0\alpha(x_1) = 1.1178 \times 10^{-13} < \varepsilon_1$ ดังนั้นขบวนการทั้งหมดจะหยุดดำเนินการ ณ. ค่า $\rho_0 = 0.1$ นี้

เมื่อคำนวณค่าของเงื่อนไขบังคับที่จุด x_1 จะได้

$$h_1(x_1) = -\ 4.9\ \times 10^{-7}$$
$$h_2(x_1) = -7.592 \times 10^{-7}$$
$$h_3(x_1) = \ 2.569 \times 10^{-7}$$
$$h_4(x_1) = \ 9.16\ \times 10^{-8}$$
$$h_5(x_1) = \ 4.764 \times 10^{-7}$$

ส่วนรายละเอียดของผลการคำนวณการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับที่ ρ_0 เป็นดังตารางที่ 4.1

ตารางที่ 4.1 แสดงผลการคำนวณโดยละเอียดสำหรับปัญหาที่ 1

i	$x^T_{0,(i+1)} = (x_1, x_2, x_3, x_4, x_5)$	$\lVert\nabla P(x_{0,(i+1)})\rVert$
0	(0.2870434, 1.5560651, 0.7624585, 0.2250000, 1.0104866)	3.3114475533
1	(0.5065981, 1.5538976, 1.2027784, 0.5815922, 0.9941321)	2.0482516922
2	(0.7725619, 1.4119086, 1.1637158, 1.1206275, 0.8134025)	1.0316347075
3	(0.8449785, 1.3386804, 1.2625971, 1.0533882, 0.6197581)	0.5121021218
4	(0.9478420, 1.2687417, 1.1908899, 0.9833617, 0.3577960)	0.3351683356
5	(1.1368504, 1.2282441, 1.1436949, 0.8865188, 0.3110149)	0.2803638107
6	(1.6164002, 1.1272612, 0.9767545, 0.6966716, 0.3250641)	0.2567477500
7	(1.7324674, 1.1071188, 0.9476395, 0.6976360, 0.2700233)	0.1202291737
8	(1.9446817, 1.0611988, 0.9069901, 0.6622319, 0.2459965)	0.0475142024
9	(2.0570910, 1.0371858, 0.8791459, 0.6287963, 0.2279279)	0.0454272679
10	(2.1217993, 1.0248476, 0.8650232, 0.6135873, 0.2309946)	0.0167165952
11	(2.1876303, 1.0118530, 0.8514632, 0.5996484, 0.2244639)	0.0094008960
12	(2.2047240, 1.0085901, 0.8478412, 0.5975638, 0.2231348)	0.0022594441
13	(2.2100903, 1.0075582, 0.8468146, 0.5963925, 0.2225157)	0.0008695198
14	(2.2125079, 1.0070706, 0.8463243, 0.5957893, 0.2222537)	0.0001225946
15	(2.2121635, 1.0071512, 0.8464153, 0.5958525, 0.2222872)	0.0001140976
16	(2.2120958, 1.0071591, 0.8464204, 0.5958694, 0.2222946)	0.0000079893
17	(2.2121247, 1.0071536, 0.8464147, 0.5958634, 0.2222918)	0.0000014673

สรุปที่ $\rho_0 = 0.1$ ได้

$$\mathbf{x}_1 = (2.2121247455, 1.0071536049, 0.8464146665, 0.5958633829, 0.2222917951)^T$$
$$\rho_0\alpha(\mathbf{x}_1) = 1.12 \times 10^{-13}$$
$$h_1(\mathbf{x}_1) = 4.9 \times 10^{-7}$$
$$h_2(\mathbf{x}_1) = -7.592 \times 10^{-7}$$
$$h_3(\mathbf{x}_1) = 2.596 \times 10^{-7}$$
$$h_4(\mathbf{x}_1) = 9.16 \times 10^{-8}$$
$$h_5(\mathbf{x}_1) = 4.764 \times 10^{-7}$$

#

<u>ปัญหาที่ 2</u> เป็นปัญหากรณีที่จำนวนสมการน้อยกว่าจำนวนตัวแปร (7 สมการ , 8 ตัวแปร) ซึ่งปัญหาที่ยกขึ้นมานี้เป็นปัญหาที่ผู้วิจัยดัดแปลงจากปัญหาใน [15]

หาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการต่อไปนี้

$$\left.\begin{aligned}
h_1(\mathbf{x}) &= x_3 + x_4 + x_5 - 1 = 0 \\
h_2(\mathbf{x}) &= x_6 + x_7 + x_8 - 1 = 0 \\
h_3(\mathbf{x}) &= x_1 + x_2 - 1 = 0 \\
h_4(\mathbf{x}) &= x_1x_6 + x_2x_3 + x_5 - 0.05 = 0 \\
h_5(\mathbf{x}) &= x_1x_7 + x_2x_4 + x_5 - 0.25 = 0 \\
h_6(\mathbf{x}) &= \frac{1370}{760}x_6 - x_3 = 0 \\
h_7(\mathbf{x}) &= \frac{550}{760}x_7 - x_4 = 0
\end{aligned}\right\} \qquad (4.3)$$

<u>ขั้นตอนดำเนินการ</u>

เราจะใช้วิธีฟังก์ชันทัณฑ์ในการหาจุดที่เป็นไปได้ โดยทำการแก้ปัญหาค่าต่ำสุดต่อไปนี้

$$\begin{aligned}
\text{minimize} \quad & f(\mathbf{x}) = 0 \\
\text{subject to} \quad & h_1(\mathbf{x}) = x_3 + x_4 + x_5 - 1 \\
& h_2(\mathbf{x}) = x_6 + x_7 + x_8 - 1 = 0 \\
& h_3(\mathbf{x}) = x_1 + x_2 - 1 = 0 \\
& h_4(\mathbf{x}) = x_1x_6 + x_2x_3 + x_5 - 0.05 = 0 \\
& h_5(\mathbf{x}) = x_1x_7 + x_2x_4 + x_5 - 0.25 = 0 \\
& h_6(\mathbf{x}) = \frac{1370}{760}x_6 - x_3 = 0 \\
& h_7(\mathbf{x}) = \frac{550}{760}x_7 - x_4 = 0 \\
& \mathbf{x} \in R^8
\end{aligned}$$

เมื่อทำการแปลงปัญหาข้างต้นด้วยวิธีฟังก์ชันทัณฑ์ จะได้ปัญหาแบบไม่มีเงื่อนไขบังคับดังนี้

$$\text{minimize} \quad P(\mathbf{x}, \rho_k) = f(\mathbf{x}) + \rho_k\{\sum_{j=1}^{7}[h_j(\mathbf{x})]^2\} = \rho_k\{\sum_{j=1}^{7}[h_j(\mathbf{x})]^2\} \qquad (4.4)$$

$$\text{subject to} \quad \mathbf{x} \in R^8$$

ทำการแก้ปัญหา (4.4) ด้วยวิธีแนวนิวตัน (สูตร BFGS ผกผัน) โดยทำการแก้ปัญหา (4.4) สำหรับแต่ละ ρ_k ซึ่งขั้นตอนวิธีที่ใช้แก้ปัญหาคือ ขั้นตอนวิธีที่ 3.3 ทำการกำหนดเงื่อนไขเริ่มต้นดังนี้

ให้ $\varepsilon_1 = 0.0000005$ เป็นค่าสเกลาร์ในการหยุด , เลือกจุดเริ่มต้น $\mathbf{x}_0 = (-5,5,0,-1,0,0,3,-2)^T$ และเลือก $\rho_0 = 0.1$, $\beta = 10$ ให้ $k = 0$ และเริ่มดำเนินการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับสำหรับแต่ละค่าของ ρ

พิจารณาการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับที่ ρ_k ใด ๆ

$$\text{minimize} \quad P(\mathbf{x}, \rho_k) = \rho_k\{\sum_{j=1}^{7}[h_j(\mathbf{x})]^2\}$$

$$\text{subject to} \quad \mathbf{x} \in R^8$$

โดยขั้นตอนวิธีที่ 3.2 ทำการกำหนดให้ $\varepsilon_2 = 0.0000005$ เป็นค่าสเกลาร์ในการหยุดและ $H_{k,0} = I_8$ (เมตริกซ์เริ่มต้นสำหรับ ρ_k ใด ๆ)

ผู้วิจัยจะได้แสดงผลการคำนวณโดยละเอียดสำหรับที่ $k = 0, i = 0$ เพื่อทำการกำเนิดจุด $\mathbf{x}_1$ (สำหรับลูปภายใน $i = 1,2,\ldots$(ซึ่งใช้แก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ) และลูปใหญ่ $k = 1,2,\ldots$ สามารถทำได้ในทำนองเดียวกัน)

ที่ $k = 0, i = 0$

$$\begin{aligned} P(\mathbf{x}) = P(\mathbf{x}, \rho_0) = 0.1[&(x_3 + x_4 + x_5 - 1)^2 + (x_6 + x_7 + x_8 - 1)^2 + (x_1 + x_2 - 1)^2 \\ &+ (x_1x_6 + x_2x_3 - 0.05)^2 + (x_1x_7 + x_2x_4 - 0.25)^2 \\ &+ (\frac{1370}{760}x_6 - x_3)^2 + (\frac{550}{760}x_7 - x_4)^2] \end{aligned}$$

โดยที่

$$\nabla P(\mathbf{x}) = \left(\frac{\partial P(\mathbf{x})}{\partial x_1}, \frac{\partial P(\mathbf{x})}{\partial x_2}, \frac{\partial P(\mathbf{x})}{\partial x_3}, \frac{\partial P(\mathbf{x})}{\partial x_4}, \frac{\partial P(\mathbf{x})}{\partial x_5}, \frac{\partial P(\mathbf{x})}{\partial x_6}, \frac{\partial P(\mathbf{x})}{\partial x_7}, \frac{\partial P(\mathbf{x})}{\partial x_8}\right)^T$$

ซึ่ง

$$\frac{\partial P(x)}{\partial x_1} = 0.1[2(x_1 + x_2 - 1) + 2(x_1x_6 + x_2x_3 - 0.05)x_6 + 2(x_1x_7 + x_2x_4 - 0.25)x_7]$$

$$\frac{\partial P(x)}{\partial x_2} = 0.1[2(x_1 + x_2 - 1) + 2(x_1x_6 + x_2x_3 - 0.05)x_3 + 2(x_1x_7 + x_2x_4 - 0.25)x_4]$$

$$\frac{\partial P(x)}{\partial x_3} = 0.1[2(x_3 + x_4 + x_5 - 1) + 2(x_1x_6 + x_2x_3 - 0.05)x_2 + 2(\frac{1370}{760}x_6 - x_3)(-1)]$$

$$\frac{\partial P(x)}{\partial x_4} = 0.1[2(x_3 + x_4 + x_5 - 1) + 2(x_1x_7 + x_2x_4 - 0.25)x_2 + 2(\frac{550}{760}x_7 - x_4)(-1)]$$

$$\frac{\partial P(x)}{\partial x_5} = 0.1[2(x_3 + x_4 + x_5 - 1)]$$

$$\frac{\partial P(x)}{\partial x_6} = 0.1[2(x_6 + x_7 + x_8 - 1) + 2(x_1x_6 + x_2x_3 - 0.05)x_1 + 2(\frac{1370}{760}x_6 - x_3)(\frac{1370}{760})]$$

$$\frac{\partial P(x)}{\partial x_7} = 0.1[2(x_6 + x_7 + x_8 - 1) + 2(x_1x_7 + x_2x_4 - 0.25)x_1 + 2(\frac{550}{760}x_7 - x_4)(\frac{550}{760})]$$

$$\frac{\partial P(x)}{\partial x_8} = 0.1[2(x_6 + x_7 + x_8 - 1)]$$

ผลการคำนวณที่จุด x_0 (ให้ $x_{0,0} = x_0$ เพราะกำลังคำนวณในลูป $i = 0$ ตามขั้นตอนวิธีที่ 3.2) เป็นดังนี้

$$\nabla P(\mathbf{x}_{0,0}) = (-112.4500, 3.8500, -54.0553, -21.2842, -0.4000, 58.5490,$$
$$22.7090, 2.0000)^T$$
$$P(\mathbf{x}_{0,0}) = 335.5069$$
$$\mathbf{p}_{0,0} = (112.4500, -3.8500, 54.0553, 21.2842, 0.4000, -58.5490, -22.7090, -2.0000)^T$$

เลือกใช้ $\mu = 1/3$ ในการหา $\alpha_{0,0}$ โดยขั้นตอนวิธีที่ 3.1 จะได้ $\alpha_{0,0} = 0.03125$ ดังนั้น

$$\mathbf{x}_{0,1} = \mathbf{x}_{0,0} + \alpha_{0,0}\mathbf{p}_{0,0} = (-1.4859375, 4.8796875, 1.6892270, -0.3348684, 0.0125000,$$
$$8.1703450, 2.2903447, -2.0625000)^T$$
$$\nabla P(\mathbf{x}_{0,1}) = (-8.3941, -0.5009, -6.3871, -5.4852, 0.0734, 7.3537, 3.3393,$$
$$1.4796)^T$$

และคำนวณ $H_{0,1}$ จากสูตร (3.1) ได้

$$H_{0,1} = \begin{bmatrix} 0.3605 & 0.0268 & -0.2928 & -0.0968 & -0.0029 & 0.3144 & 0.1189 & 0.0031 \\ 0.0268 & 0.9989 & 0.0124 & 0.0042 & 0.0001 & -0.0133 & -0.0051 & -0.0002 \\ -0.2928 & 0.0124 & 0.8663 & -0.0438 & -0.0014 & 0.1435 & 0.0542 & 0.0012 \\ -0.0968 & 0.0042 & -0.0438 & 0.9863 & -0.0005 & 0.0469 & 0.0176 & 0.0002 \\ -0.0029 & 0.0001 & -0.0014 & -0.0005 & 1.0000 & 0.0015 & 0.0006 & 0.0000 \\ 0.3144 & -0.0133 & 0.1435 & 0.0469 & 0.0015 & 0.8460 & -0.0582 & -0.0013 \\ 0.1189 & -0.0051 & 0.0542 & 0.0176 & 0.0006 & -0.0582 & 0.9781 & -0.0004 \\ 0.0031 & -0.0002 & 0.0012 & 0.0002 & 0.0000 & -0.0013 & -0.0004 & 1.0001 \end{bmatrix}$$

เนื่องจาก $\|\nabla P(x_{0,1})\| = 14.4574471 > \varepsilon_2$ ดังนั้นให้ดำเนินการต่อตามขั้นตอนที่กล่าวข้างต้น เพื่อกำเนิด $x_{0,2}, x_{0,3}, \ldots$ จนได้จุด $x_{0,44} = x_1 = (-3.5612075, 4.5612075, 0.0193375, -0.6949598, 1.6756223, 0.0107274, -0.9603080, 1.9495807)^T$ ซึ่ง $\|\nabla P(x_1)\| < \varepsilon_2$ และได้ค่า $\alpha(x_1) = \sum_{j=1}^{7}[h_j(x_1)]^2 = 1.2707 \times 10^{-14}$ ดังนั้น $\rho_0\alpha(x_1) = 1.2707 \times 10^{-15} < \varepsilon_1$ ดังนั้นขบวนการทั้งหมดจะหยุดดำเนินการ ณ. ค่า $\rho_0 = 0.1$ นี้

เมื่อคำนวณค่าของเงื่อนไขบังคับที่จุด x_1 จะได้

$$\begin{aligned} h_1(x_1) &= -4 \times 10^{-10} \\ h_2(x_1) &= 1.49 \times 10^{-8} \\ h_3(x_1) &= 4.36 \times 10^{-8} \\ h_4(x_1) &= -4.3 \times 10^{-9} \\ h_5(x_1) &= 4.6 \times 10^{-8} \\ h_6(x_1) &= -1.84 \times 10^{-8} \\ h_7(x_1) &= 9.01 \times 10^{-8} \end{aligned}$$

ส่วนรายละเอียดของผลการคำนวณการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับที่ ρ_0 เมื่อใช้จุดเริ่มต้น $x_0 = (-5, 5, 0, -1, 0, 10, 3, -2)^T$ เป็นดังตารางที่ 4.2

เมื่อเปลี่ยนจุดเริ่มต้นเป็น $x_0 = (-2, 3, 0.2, 1, -0.1, 0.1, 1.5, -0.5)^T$ จะได้จุดที่เป็นไปได้จุดใหม่คือ $x^* = (-2.0038951, 3.0038973, 0.0264237, 1.0643215, -0.0907449$ $0.0146584, 1.4706927, -0.4853507)^T$

ตารางที่ 4.2 แสดงผลการคำนวณโดยละเอียดสำหรับปัญหาที่ 2

i	$\mathbf{x}_{0,(i+1)}^{T} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$	$\lVert \nabla P(\mathbf{x}_{0,(i+1)}) \rVert$
0	(-1.4859375, 4.8796875, 1.6892270, -0.3348684, 0.0125000, 8.1703450, 2.2903447, -2.0625000)	14.4574470503
1	(-2.5230821, 5.3510529, 2.4912248, 1.6233283, -0.0482699, 7.0608101, 1.5906533, -2.7796585)	11.9252928469
2	(-3.7547177, 5.9035257, 3.9087446, 1.8615911, -0.2397551, 5.2865520, 1.5517793, -3.5866301)	10.8840684749
3	(-4.7376361, 4.2987184, 4.1214752, 1.5212852, -0.6630947, 3.1142033, 1.7566230, -4.5105012)	4.6901495758
4	(-4.4889025, 4.0785635, 3.7710905, 1.2152283, -1.0236844, 3.4218466, 1.1809910, -4.4994185)	1.1912770667
5	(-4.4947450, 4.3100351, 3.2077047, 0.9923937, -1.8376869, 2.9040663, 0.6757697, -4.3649732)	2.0104493405
6	(-4.1147209, 4.5813390, 2.5558633, 0.7734960, -2.6891584, 2.8251474, 0.9111185, -3.8602427)	1.0404992168
7	(-4.2592726, 4.4554466, 2.6870239, 0.9180947, -2.6527458, 2.6685017, 0.9733727, -3.6496920)	0.3507526653
8	(-4.3699953, 4.3326615, 2.6158285, 1.0890040, -2.8320813, 2.3665800, 1.0702762, -3.0658552)	0.6849562968
9	(-4.3921109, 4.2624651, 2.4474954, 1.2237340, -2.9433933, 2.1561570, 1.1568522, -2.4938220)	0.5983533965
10	(-4.3585463, 4.2396439, 2.2055580, 1.3243764, -2.9356290, 1.9883350, 1.2295256, -2.0374081)	0.2155794628
11	(-4.3194984, 4.2381441, 1.9522518, 1.3953651, -2.8127163, 1.8116059, 1.2959331, -1.7803244)	0.2493259524
12	(-4.2283619, 4.2502541, 1.2739157, 1.5682035, -2.3483911, 1.2948542, 1.4833755, -1.2964889)	0.7755071463
13	(-4.1842062, 4.2175134, 0.6369868, 1.7492801, -1.7551615, 0.7105904, 1.6836232, -0.9612875)	0.8443169403
14	(-4.1851666, 4.1536251, 0.0596621, 1.9151323, -1.0508963, 0.0820956, 1.8684670, -0.8258498)	0.4301789370
15	(-4.1291766, 4.2733234, -0.1764829, 1.9083601, -0.7819555, -0.0732681, 1.9155432, -0.9274641)	0.6626520571
16	(-4.1392194, 4.3482908, -0.0028659, 1.8442426, -0.8716413, 0.0623515, 1.8866967, -1.1022179)	0.4907731866
17	(-4.1325445, 4.6030150, 0.2826799, 1.7216400, -1.0100117, 0.2946319, 1.8864248, -1.4016488)	0.2190161721
18	(-4.0966901, 4.8721820, 0.2110546, 1.7128230, -0.9023562, 0.2038677, 2.0105643, -1.3956501)	0.2208156440
19	(-4.0674910, 5.0066404, 0.0139772, 1.7617090, -0.7475243, 0.0100978, 2.1263074, -1.2484379)	0.0733048442
20	(-4.0525851, 5.1303216, 0.0136009, 1.7664583, -0.7828288, 0.0031657, 2.1818868, -1.1673752)	0.0132916669
21	(-4.0536109, 5.1168901, 0.0179852, 1.7622157, -0.7820427, 0.0097464, 2.1722103, -1.1842392)	0.0063252963
22	(-4.0533246, 5.1165794, 0.0194697, 1.7609689, -0.7825542, 0.0120143, 2.1704598, -1.1857881)	0.0044005033
23	(-4.0526308, 5.1163953, 0.0228526, 1.7580513, -0.7835088, 0.0168026, 2.1665368, -1.1868629)	0.0053192264
24	(-4.0521008, 5.1162084, 0.0238477, 1.7559872, -0.7825242, 0.0183774, 2.1638856, -1.1856130)	0.0075585962
25	(-4.0495804, 5.1143717, 0.0276639, 1.7449088, -0.7755007, 0.0240369, 2.1496862, -1.1761899)	0.0152419536
26	(-4.0442459, 5.1095812, 0.0319793, 1.7206339, -0.7557019, 0.0307473, 2.1187824, -1.1508052)	0.0262812787
27	(-4.0296688, 5.0945629, 0.0398093, 1.6514986, -0.6946411, 0.0425956, 2.0309352, -1.0725588)	0.0440890449
28	(-3.9946376, 5.0568125, 0.0497406, 1.4826125, -0.5357887, 0.0581020, 1.8169929, -0.8708477)	0.0684473425
29	(-3.9132174, 4.9680799, 0.0607724, 1.0853352, -0.1492562, 0.0756114, 1.3149093, -0.3816956)	0.0920277239

ตารางที่ 4.2 (ต่อ)

i	$x^T_{0,(i+1)} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$	$\lVert \nabla P(x_{0,(i+1)}) \rVert$
30	(-3.7592042, 4.7986110, 0.0618673, 0.3268861, 0.6100829, 0.0795881, 0.3550060, 0.5784819)	0.0982439279
31	(-3.5944767, 4.5792485, 0.0478045, -0.5159988, 1.4712699, 0.0585448, -0.7354950, 1.6902222)	0.0828632257
32	(-3.5522633, 4.5419360, 0.0390923, -0.7230797, 1.6872492, 0.0460913, -0.9947106, 1.9605921)	0.0603454394
33	(-3.5458376, 4.5411133, 0.0221899, -0.7654582, 1.7444781, 0.0177613, -1.0494417, 2.0361813)	0.0198310255
34	(-3.5658117, 4.5656002, 0.0196158, -0.6724344, 1.6526988, 0.0108688, -0.9316403, 1.9211528)	0.0028623432
35	(-3.5618832, 4.5615375, 0.0193073, -0.6915563, 1.6722621, 0.0106493, -0.9557587, 1.9451550)	0.0003431791
36	(-3.5621653, 4.5618597, 0.0192997, -0.6900991, 1.6708128, 0.0106792, -0.9539103, 1.9432444)	0.0001282686
37	(-3.5621840, 4.5619049, 0.0193224, -0.6900134, 1.6706966, 0.0107070, -0.9538068, 1.9430994)	0.0001134138
38	(-3.5620598, 4.5619144, 0.0193760, -0.6906644, 1.6712666, 0.0107752, -0.9546876, 1.9438598)	0.0000586455
39	(-3.5617273, 4.5616919, 0.0193914, -0.6923530, 1.6729318, 0.0107946, -0.9569103, 1.9460547)	0.0000267988
40	(-3.5613311, 4.5613516, 0.0193664, -0.6943471, 1.6749641, 0.0107636, -0.9595164, 1.9487223)	0.0000147783
41	(-3.5612087, 4.5612192, 0.0193436, -0.6949566, 1.6756093, 0.0107350, -0.9603064, 1.9495653)	0.0000048345
42	(-3.5612044, 4.5612058, 0.0193378, -0.6949755, 1.6756375, 0.0107278, -0.9603288, 1.9496007)	0.0000005866
43	(-3.5612075, 4.5612075, 0.0193375, -0.6949598, 1.6756223, 0.0107274, -0.9603080, 1.9495807)	0.0000000294

สรุปที่ $\rho_0 = 0.1$ **ได้**

$$x_1 = (-3.5612075, 4.5612075, 0.0193375, -0.6949598, 1.6756223, 0.0107274, -0.9603080, 1.9495807)^T$$

$$\rho_0 \alpha(x_1) = 1.3 \times 10^{-15}$$
$$h_1(x_1) = -4 \times 10^{-10}$$
$$h_2(x_1) = 1.49 \times 10^{-8}$$
$$h_3(x_1) = 4.36 \times 10^{-8}$$
$$h_4(x_1) = -4.3 \times 10^{-9}$$
$$h_5(x_1) = 4.6 \times 10^{-8}$$
$$h_6(x_1) = 1.84 \times 10^{-8}$$
$$h_7(x_1) = 9.01 \times 10^{-8}$$

#

<u>ปัญหาที่ 3</u> เป็นปัญหากรณีที่จำนวนสมการมากกว่าจำนวนตัวแปร (4 สมการ , 3 ตัวแปร) ซึ่งเป็นปัญหาที่ผู้วิจัยสร้างขึ้นเอง

หาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบสมการต่อไปนี้

$$\left.\begin{aligned} h_1(x) &= (x_1 + x_3)^{-1} - x_2^2 + 4.5 = 0 \\ h_2(x) &= 5\ln(x_1^2) + \sin[(x_2 + x_3)\pi] + 2x_2 - 4 = 0 \\ h_3(x) &= x_1x_2 - x_2x_3 + x_1x_3 - 5 = 0 \\ h_4(x) &= 10\log(x_1^2 + x_3^2) - x_2^{-2} + x_1x_3 + \cos(x_2\pi) - 7.75 = 0 \end{aligned}\right\} \quad (4.5)$$

<u>ขั้นตอนดำเนินการ</u>

เราจะใช้วิธีฟังก์ชันทัณฑ์ในการหาจุดที่เป็นไปได้ โดยทำการแก้ปัญหาค่าต่ำสุดต่อไปนี้

$$\begin{aligned} \text{minimize} \quad & f(x) = 0 \\ \text{subject to} \quad & h_1(x) = (x_1 + x_3)^{-1} - x_2^2 + 4.5 = 0 \\ & h_2(x) = 5\ln(x_1^2) + \sin[(x_2 + x_3)\pi] + 2x_2 - 4 = 0 \\ & h_3(x) = x_1x_2 - x_2x_3 + x_1x_3 - 5 = 0 \\ & h_4(x) = 10\log(x_1^2 + x_3^2) - x_2^{-2} + x_1x_3 + \cos(x_2\pi) - 7.75 = 0 \\ & x \in R^3 \end{aligned}$$

เมื่อทำการแปลงปัญหาด้วยวิธีฟังก์ชันทัณฑ์ จะได้ปัญหาแบบไม่มีเงื่อนไขบังคับดังนี้

$$\left.\begin{aligned} \text{minimize} \quad & P(x,\rho) = f(x) + \rho_k\{\sum_{j=1}^{4}[h_j(x)]^2\} = \rho_k\{\sum_{j=1}^{4}[h_j(x)]^2\} \\ \text{subject to} \quad & x \in R^3 \end{aligned}\right\} \quad (4.6)$$

ทำการแก้ปัญหา (4.6) ด้วยวิธีแนวนิวตัน (สูตร BFGS ผกผัน) โดยทำการแก้ปัญหา (4.6) สำหรับแต่ละ ρ_k ซึ่งขั้นตอนวิธีที่ใช้แก้ปัญหาคือ ขั้นตอนวิธีที่ 3.3 ทำการกำหนดเงื่อนไขเริ่มต้นดังนี้

ให้ $\varepsilon_1 = 5 \times 10^{-10}$ เป็นค่าสเกลาร์ในการหยุด , เลือกจุดเริ่มต้น $x_0 = (3,3,-2)^T$ และเลือก $\rho_0 = 0.1, \beta = 10$ ให้ $k = 0$ และเริ่มดำเนินการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับสำหรับแต่ละค่าของ ρ

พิจารณาการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับต่อไปนี้ที่ ρ_k ใด ๆ

$$\text{minimize} \qquad P(\mathbf{x},\rho_k) = \rho_k\{\sum_{j=1}^{4}[\bar{h}_j(\mathbf{x})]^2\}$$

$$\text{subject to} \qquad \mathbf{x} \in R^3$$

โดยขั้นตอนวิธีที่ 3.2 จะกำหนดให้ $\varepsilon_2 = 5\times10^{-6}$ เป็นค่าสเกลาร์ในการหยุดและ $H_{k,0} = I_3$ ผู้วิจัยจะได้แสดงผลการคำนวณโดยละเอียดสำหรับที่ $k = 0, i = 0$ เพื่อทำการกำเนิดจุด $\mathbf{x}_1$ (สำหรับลูปภายใน $i = 1,2,\ldots$(ซึ่งใช้แก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับ) และลูปใหญ่ $k = 1,2,\ldots$ สามารถทำได้ในทำนองเดียวกัน)

ที่ $k = 0, i = 0$

$$P(\mathbf{x}) = P(\mathbf{x},\rho_0) = 0.1[\{(x_1+x_3)^{-1} - x_2^2 + 4.5\}^2 + \{5\ln(x_1^2) + \sin[(x_2+x_3)\pi] + 2x_2 - 4\}^2 + \{x_1x_2 - x_2x_3 + x_1x_3 - 5\}^2 + \{10\log(x_1^2+x_3^2) - x_2^{-2} + x_1x_3 + \cos(x_2\pi) - 7.75\}^2]$$

โดยที่ $$\nabla P(\mathbf{x}) = \left(\frac{\partial P(\mathbf{x})}{\partial x_1}, \frac{\partial P(\mathbf{x})}{\partial x_2}, \frac{\partial P(\mathbf{x})}{\partial x_3}\right)^T$$

ซึ่ง

$$\frac{\partial P(\mathbf{x})}{\partial x_1} = 0.1[2\{(x_1+x_3)^{-1} - x_2^2 + 4.5\}\{-(x_1+x_3)^{-2}\} + 2\{5\ln(x_1^2) + \sin[(x_2+x_3)\pi] + 2x_2 - 4\}\{10/x_1\} + 2\{x_1x_2 - x_2x_3 + x_1x_3 - 5\}\{x_2 + x_3\} + 2\{10\log(x_1^2+x_3^2) - x_2^{-2} + x_1x_3 + \cos(x_2\pi) - 7.75\}\{20(\log e)x_1/(x_1^2+x_3^2)\}]$$

$$\frac{\partial P(\mathbf{x})}{\partial x_2} = 0.1[2\{(x_1+x_3)^{-1} - x_2^2 + 4.5\}\{-2x_2\} + 2\{5\ln(x_1^2) + \sin[(x_2+x_3)\pi] + 2x_2 - 4\}\{\pi\cos[(x_2+x_3)\pi] + 2\} + 2\{x_1x_2 - x_2x_3 + x_1x_3 - 5\}\{x_1 - x_3\} + 2\{10\log(x_1^2+x_3^2) - x_2^{-2} + x_1x_3 + \cos(x_2\pi) - 7.75\}\{2x_2^{-3} - \pi\sin(x_2\pi)\}]$$

$$\frac{\partial P(\mathbf{x})}{\partial x_3} = 0.1[2\{(x_1+x_3)^{-1} - x_2^2 + 4.5\}\{-(x_1+x_3)^{-2}\} + 2\{5\ln(x_1^2) + \sin[(x_2+x_3)\pi] + 2x_2 - 4\}\{\pi\cos[(x_2+x_3)\pi]\} + 2\{x_1x_2 - x_2x_3 + x_1x_3 - 5\}\{x_1 - x_2\} + 2\{10\log(x_1^2+x_3^2) - x_2^{-2} + x_1x_3 + \cos(x_2\pi) - 7.75\}\{x_1 + \frac{20(\log e)x_3}{(x_1^2+x_3^2)}\}]$$

ให้ $\mathbf{x}_{k,0}$ คือจุดเริ่มต้นเมื่อพิจารณาที่ ρ_k ,$\alpha_{k,0}$ คือระยะทางในการเดินทางจากจุด $\mathbf{x}_{k,0}$ ไปยัง $\mathbf{x}_{k,1}$ ในทิศ $\mathbf{p}_{k,0}$ ผลการคำนวณที่จุด $\mathbf{x}_0$ (ให้ $\mathbf{x}_{0,0} = \mathbf{x}_0$ เพราะกำลังคำนวณในลูป $i = 0$ ตามขั้นตอนวิธีที่ 3.2) เป็นดังนี้

$$\nabla P(\mathbf{x}_{0,0}) = (8.6654, 5.1799, -8.6978)^T$$

$$P(\mathbf{x}_{0,0}) = 21.0740\,,\ \mathbf{p}_{0,0} = (-8.6654, -5.1799, 8.6978)^T$$

เลือกใช้ $\mu = 1/3$ ในการหา $\alpha_{0,0}$ โดยขั้นตอนวิธี 3.1 จะได้ $\alpha_{0,0} = 0.1250$

ดังนั้น $\mathbf{x}_{0,1} = \mathbf{x}_{0,0} + \alpha_{0,0}\mathbf{p}_{0,0} = (1.9168197, 2.3525136, -0.9127776)^T$

$$\nabla P(\mathbf{x}_{0,1}) = (4.4904, 1.6559, -0.8060)^T$$

และคำนวณ $H_{0,1}$ จากสูตร (3.1) ได้

$$H_{0,1} = \begin{bmatrix} 0.9451 & -0.1052 & 0.3157 \\ -0.1052 & 0.8938 & 0.2614 \\ 0.3157 & 0.2614 & 0.4215 \end{bmatrix}$$

ซึ่ง $\|\nabla P(\mathbf{x}_{0,1})\| = 4.8533853 > \varepsilon_2$ ดังนั้นให้ดำเนินการต่อตามขั้นตอนที่กล่าวข้างต้น เพื่อกำเนิด $\mathbf{x}_{0,2}, \mathbf{x}_{0,3}, \ldots$ จนได้จุด $\mathbf{x}_{0,19} = \mathbf{x}_1 = (1.0000002, 2.0000002, -2.9999994)^T$ ซึ่ง $\|\nabla P(\mathbf{x}_1)\| = 2.9994 \times 10^{-13} < \varepsilon_2$ นอกจากนี้เมื่อคำนวณ $\alpha(\mathbf{x}_1) = \sum_{j=1}^{4}[h_j(\mathbf{x}_1)]^2$ จะได้ $\alpha(\mathbf{x}_1) = 2.9994 \times 10^{-12}$ และ $\rho_0\alpha(\mathbf{x}_1) = 2.9994 \times 10^{-13} < \varepsilon_1$ ดังนั้นขบวนการทั้งหมดจะหยุดดำเนินการ ณ. ค่า $\rho_0 = 0.1$ นี้ และเมื่อคำนวณค่าของเงื่อนไขบังคับทั้งหมดที่จุด $\mathbf{x}_1$ จะได้

$$\begin{aligned} h_1(\mathbf{x}_1) &= -9.747 \times 10^{-7} \\ h_2(\mathbf{x}_1) &= -3.04 \times 10^{-8} \\ h_3(\mathbf{x}_1) &= -8.5 \times 10^{-8} \\ h_4(\mathbf{x}_1) &= -1.4287 \times 10^{-6} \end{aligned}$$

ส่วนรายละเอียดของผลการคำนวณการแก้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับที่ ρ_0 เป็นดังตารางที่ 4.3

ตารางที่ 4.3 แสดงผลการคำนวณโดยละเอียดสำหรับปัญหาที่ 3

i	$\mathbf{x}_{0,(i+1)}^T = (x_1, x_2, x_3)$	$\lVert \nabla P(\mathbf{x}_{0,(i+1)}) \rVert$
0	(1.9168197, 2.3525136, -0.9127776)	4.8533852565
1	(0.9630647, 2.1533058, -1.2904956)	9.8331560250
2	(0.8751106, 2.1612968, -1.2916591)	7.2489182887
3	(0.8751106, 2.1612968, -1.2916591)	7.2489182887
4	(0.9973027, 2.1233073, -1.6464444)	3.4238990365
5	(1.0184666, 1.9622878, -1.8110001)	2.0917386275
6	(1.1020726, 1.8331991, -2.3336140)	1.8436277061
7	(1.1629077, 1.9816238, -2.7763441)	1.4911286251
8	(1.1230848, 2.0593391, -2.6940797)	0.6119006970
9	(1.1196261, 2.0554181, -2.7254782)	0.6191507642
10	(1.0807639, 2.0334332, -2.8722815)	0.6183434507
11	(1.0359276, 2.0184086, -2.9479937)	0.3126500567
12	(1.0001807, 2.0023781, -2.9944799)	0.0435718163
13	(1.0011756, 2.0007695, -2.9966467)	0.0032766742
14	(1.0003101, 2.0001130, -2.9995575)	0.0030607835
15	(0.9999582, 1.9999490, -3.0001613)	0.0004583825
16	(0.9999863, 1.9999921, -3.0000335)	0.0000431254
17	(1.0000012, 2.0000012, -2.9999952)	0.0000129433
18	(1.0000002, 2.0000002, -2.9999994)	0.0000008793

สรุปที่ $\rho_0 = 0.1$ ได้

$$\mathbf{x}_1 = (1.0000002, 2.0000002, -2.9999994)^T$$
$$\rho_0\alpha(\mathbf{x}_1) = 2.999 \times 10^{-13}$$
$$h_1(\mathbf{x}_1) = -9.747 \times 10^{-7}$$
$$h_2(\mathbf{x}_1) = -3.04 \times 10^{-8}$$
$$h_3(\mathbf{x}_1) = -8.57 \times 10^{-8}$$
$$h_4(\mathbf{x}_1) = -1.4287 \times 10^{-6}$$

#

4.2 ตัวอย่างปัญหาการหาจุดภายในของบริเวณที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการ

<u>ปัญหาที่ 4</u> เป็นปัญหาที่ผู้วิจัยสร้างขึ้นเพื่อให้สอดคล้องกับเงื่อนไขบังคับในรูปแบบ (1.2) หาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการต่อไปนี้

$$\left.\begin{aligned} g_1(x) &= x_1^2 + 5x_2 + x_3^2 + 5 \quad \le 0 \\ g_2(x) &= -2x_1 + x_2 - x_3 + 10 \le 0 \\ g_3(x) &= x_1x_2 + x_2x_3 + 23 \quad \le 0 \\ g_4(x) &= e^{(x_3 - x_1)} + 7x_2 + 10 \quad \le 0 \end{aligned}\right\} \qquad (4.7)$$

<u>ขั้นตอนดำเนินการ</u>

เราจะใช้การทำซ้ำตัวเองของวิธีฟังก์ชันกั้นเขต โดยใช้จุดเริ่มต้น $\mathbf{x}_0 = (1,1,1)^T$ เมื่อคำนวณ $g(\mathbf{x}_0)$ เมื่อ $g = (g_1(x),\ldots, g_4(x))^T$ จะได้ $g(\mathbf{x}_0) = (12,8,25,18)^T$ ดังนั้นจุด $\mathbf{x}_0$ จึงไม่เป็นจุดภายในที่เป็นไปได้ของเงื่อนไขทั้งหมด ให้ t และ f เป็นอาร์เรย์ของเลขดัชนีที่จุด $\mathbf{x}$ ใด ๆ ซึ่งทำให้ $g_i(x) < 0$ และ $g_i(x) \ge 0$ สำหรับ $i = 1,\ldots,4$ ตามลำดับ (กำหนดเพิ่มเติมว่าให้ f เป็นอาร์เรย์ของเลขดัชนีที่เรียงจากน้อยไปหามาก) ดังนั้นที่จุด $\mathbf{x}_0$ จะได้ว่า t ไม่มีสมาชิกอยู่เลย (t = [ ]) และ f = [1,2,3,4] ดังนั้นเราจะเลือกสมาชิกตัวแรกของอาร์เรย์ f ซึ่งได้แก่ f(1) = 1 เป็นฟังก์ชันจุดประสงค์ แล้วทำการหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับโดยเริ่มต้นด้วยจุด $\mathbf{x}_0$

$$\left.\begin{aligned} &\text{minimize} && g_1(x) = x_1^2 + 5x_2 + x_3^2 + 5 \\ &\text{subject to} && x \in R^3 \end{aligned}\right\} \qquad (4.8)$$

ใช้ขั้นตอนวิธี OPTIM 1 แก้ปัญหา(4.8) เพื่อหาจุด $\mathbf{x}_1$ ซึ่งเป็นจุดต่ำสุดหรือจุดที่ทำให้ $g_1(\mathbf{x}_1) < 0$ ในที่นี้จะได้แสดงผลการคำนวณของ OPTIM 1 ในขั้นตอนต่าง ๆ ดังนี้

(1). **OPTIM 1** กำหนด $\varepsilon_2 = 0.005$ เป็นค่าสเกลาร์ในการหยุด สำหรับขั้นตอนค้นหาแนวเส้นกำหนด $\mu = 1/3$ ในที่นี้กำหนดให้ $H_0 = I_3$ และเกรเดียนต์ของฟังก์ชัน g_1 คือ $\nabla g_1(x) = (2x_1,5,2x_3)^T$ ทำการคำนวณค่าต่าง ๆ ดังนี้

$$\begin{aligned} \nabla g_1(\mathbf{x}_0) &= (2,5,2)^T \\ g_1(\mathbf{x}_0) &= 12 \\ \mathbf{p}_0 &= (-2,-5,-2)^T \\ \alpha_0 &= 1 \\ \mathbf{x}_1 &= \mathbf{x}_0 + \alpha_0\mathbf{p}_0 = (-1,-4,-1)^T \end{aligned}$$

ซึ่งที่จุด x_1 จะได้ $\nabla g_1(x_1) = (-2,5,-2)^T$ ซึ่ง $\|\nabla g_1(x_1)\| = 5.744563 > \varepsilon_2$ แต่ได้ค่า $g_1(x_1) = -13 < 0$ ดังนั้นขั้นตอนวิธี OPTIM 1 จะหยุด ณ. จุด x_1 นี้ และอาร์เรย์ t , f มีการเปลี่ยนแปลงดังนี้ $t = [1], f = [2,3,4]$ จากนั้นทำการใช้ขั้นตอนวิธี OPTIM 2 ในการแก้ปัญหาแบบมีเงื่อนไขบังคับ โดยเลือกสมาชิกตัวแรกของอาร์เรย์ f คือ $f(1) = 2$ เป็นดัชนีของฟังก์ชันจุดประสงค์และให้สมาชิกทั้งหมดของ t เป็นดัชนีของเงื่อนไขบังคับ และใช้วิธีฟังก์ชันกั้นเขตแก้ปัญหาต่อไปนี้

$$\left.\begin{array}{ll} \text{minimize} & g_2(\mathbf{x}) = -2x_1 + x_2 - x_3 + 10 \\ \text{subject to} & g_1(\mathbf{x}) = x_1^2 + 5x_2 + x_3^2 + 5 \le 0 \\ & \mathbf{x} \in R^3 \end{array}\right\} \quad (4.9)$$

โดยใช้จุด x_1 ใน OPTIM 1 เป็นจุดเริ่มต้น

(2). **OPTIM 2** เพื่อแก้ปัญหา (4.9) กำหนดให้ $\varepsilon_1 = \varepsilon_2 = 5 \times 10^{-6}$ เป็นค่าสเกลาร์ในการหยุดในลูปใหญ่ k และลูปเล็ก i ตามลำดับ สำหรับลูป k กำหนดให้ $\beta = 0.1$ และที่ $k = 0$ กำหนดให้ $\mu_0 = 10$ และสำหรับลูป i กำหนดให้ $\mu = 1/3$ เป็นค่าที่ใช้ในการค้นหาแนวเส้น

ที่ $k = 0$, $i = 0$

เมื่อใช้ฟังก์ชันกั้นเขตแก้ปัญหาจะได้ปัญหาใหม่เป็น

$$\begin{array}{ll} \text{minimize} & U_1 = g_2(\mathbf{x}) - \dfrac{\mu_k}{g_1(\mathbf{x})} \\ \text{subject to} & g_1(\mathbf{x}) < 0 \\ & \mathbf{x} \in R^3 \end{array}$$

ในที่นี้จะได้แสดงผลการคำนวณโดยละเอียดดังนี้

กำหนด $\quad \mathbf{x}_{0,0} = \mathbf{x}_1 = (-1,-4,-1)^T$

$H_{0,0} = I_3$

คำนวณ $\quad \nabla U_1(\mathbf{x}_{0,0}) = (-2.1183, 1.2959, -1.1183)^T$

$\mathbf{p}_{0,0} = -H_{0,0}\nabla U_1(\mathbf{x}_{0,0}) = (2.1183, -1.2959, 1.1183)^T$

$U_1(\mathbf{x}_{0,0}) = 9.7692$

$$\alpha_{0,0} = 1$$

$$\mathbf{x}_{0,1} = (1.1183,-5.2959,0.1183)^T$$

$$\nabla U_1(\mathbf{x}_{0,1}) = (-1.9453,1.1224,-0.9942)^T \;,\; \|\nabla U_1(\mathbf{x}_{0,1})\| \doteq 2.45605 > \varepsilon_2$$

และ $g_2(\mathbf{x}_{0,1}) = 2.34911 > 0$ ดังนั้นให้ดำเนินการต่อไป

ที่ $k = 0, i = 1$

คำนวณ $H_{0,1}$ จาก $H_{0,0}$ โดยสูตร (3.1) ได้

$$H_{0,1} = \begin{bmatrix} 6.7755 & -3.3369 & 2.9540 \\ -3.3369 & 2.9213 & -1.7035 \\ 2.9540 & -1.70365 & 2.5094 \end{bmatrix}$$

$$U_1(\mathbf{x}_{0,1}) = 2.8438$$

$$\mathbf{p}_{0,1} = -H_{0,1}\nabla U_1(\mathbf{x}_{0,1}) = (19.8622,-11.4636,10.1532)^T$$

$$\alpha_{0,1} = 0.1250$$

$$\mathbf{x}_{0,2} = \mathbf{x}_{0,1} + \alpha_{0,1}\mathbf{p}_{0,1} = (3.6011,-6.7588,1.3875)^T$$

$$\nabla U_1(\mathbf{x}_{0,2}) = (-1.6191,1.2644,-0.8532)^T \;, \|\nabla U_1(\mathbf{x}_{0,2})\| = 2.22448 > \varepsilon_2$$

แต่ $g_2(\mathbf{x}_{0,2}) = -5.31854 < 0$ ดังนั้นขบวนการใน OPTIM 2 ในครั้งนี้จะหยุดดำเนินการ ณ. จุด $\mathbf{x}_{0,2}$ นี้ และอาร์เรย์ t , f เป็นดังนี้ $t = [1,2]$, $f = [3,4]$ และเมื่อทำการตรวจสอบ $g_3(\mathbf{x}_{0,2})$ และ $g_4(\mathbf{x}_{0,2})$ พบว่า

$$g_3(\mathbf{x}_{0,2}) = -10.56742$$

$$g_4(\mathbf{x}_{0,2}) = -36.99233$$

ขณะนี้ $t = [1,2,3,4]$ และ f ไม่มีสมาชิกอยู่แล้ว ดังนั้นจึงไม่ต้องการการเรียกใช้ขั้นตอนวิธี OPTIM 2 ซ้ำอีก

สรุป ที่จุด $\mathbf{x}_{0,2} \in \{\mathbf{x} \in R^3 \mid g_i(\mathbf{x}) < 0 \;\; ; i = 1,2,3,4\}$ ดังนั้น $\mathbf{x}_{0,2}$ เป็นจุดภายในที่เป็นไปได้ ตารางที่ 4.4 จะแสดงผลการคำนวณโดยรวมของปัญหาที่ 4 #

ตารางที่ 4.4 แสดงผลการคำนวณโดยรวมของปัญหาที่ 4

OPTIM 1

k	$x_{k+1}^T = (x_1, x_2, x_3)$	$\|\nabla g_1(x_{k+1})\|$	ค่าฟังก์ชันจุดประสงค์ที่ x_{k+1}
0	(-1.0000000, -4.0000000, -1.0000000)	5.74456	g(1)=-13.00000

OPTIM 2

ครั้งที่	k	μ_k	$x_{k+1}^T = (x_1, x_2, x_3)$	ค่าฟังก์ชันจุดประสงค์ที่ x_{k+1}	ค่าเงื่อนไขบังคับที่ x_{k+1}			
1	0	10	(3.6011228, -6.7288052, 1.3874918)	g(2)=-5.31854	g(1)=-13.7508074	-	-	-
	0	10	(3.6011228, -6.7288052, 1.3874918)	g(3)=-10.56742	g(1)=-13.7508074	g(2)=-5.3185426	-	-
	0	10	(3.6011228, -6.7288052, 1.3874918)	g(4)=-36.99233	g(1)=-13.7508074	g(2)=-5.3185426	g(3)=-10.5674159	-

4.3 ตัวอย่างปัญหาการหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบผสม

<u>ปัญหาที่ 5</u> เป็นปัญหาที่ผู้วิจัยสร้างขึ้นเพื่อให้สอดคล้องกับเงื่อนไขบังคับในรูปแบบ (1.3)โดยที่เงื่อนไขบังคับทั้งหมดเป็นฟังก์ชันเชิงเส้น

หาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบผสมต่อไปนี้

$$\left.\begin{aligned}
g_1(\mathbf{x}) &= x_1 - x_2 + x_3 - x_4 + 20 \le 0\\
g_2(\mathbf{x}) &= 3x_1 + 6x_2 - 7x_3 - 2x_4 + 8 \le 0\\
g_3(\mathbf{x}) &= -2x_1 - 4x_2 - 3x_3 - x_4 + 1 \le 0\\
g_4(\mathbf{x}) &= x_1 + 2x_2 + 150x_3 + x_4 + 4 \le 0\\
g_5(\mathbf{x}) &= -7x_1 + 6x_2 + 2x_3 - x_4 + 15 \le 0\\
h_1(\mathbf{x}) &= x_1 + x_2 + x_3 + x_4 - 35.5 = 0\\
h_2(\mathbf{x}) &= -4x_1 + 3x_2 - 2x_3 + x_4 - 25.4 = 0\\
h_3(\mathbf{x}) &= 13x_1 - 17x_2 - 142x_3 + 3x_4 - 108 = 0
\end{aligned}\right\} \quad (4.10)$$

<u>ขั้นตอนดำเนินการ</u>

เราแบ่งการทำงานออกเป็น 2 ส่วนคือส่วนแรกกระทำกับส่วนของเงื่อนไขอสมการ เราจะใช้วิธีฟังก์ชันกั้นเขตทำซ้ำตัวเองเพื่อให้ได้จุด $x' \in \{x \in R^4 \mid g_i(x) < 0; i = 1,\ldots,5\}$ และส่วนที่สองจะกระทำกับเงื่อนไขบังคับทั้งหมดโดยใช้จุด x' เป็นจุดเริ่มต้น เพื่อหาจุดที่สอดคล้องกับเงื่อนไขทั้งหมด (อสมการและสมการ)

ส่วนที่ 1

เริ่มต้นใช้จุด $\mathbf{x}_0 = (10,20,30,40)^T$ คำนวณ $g(\mathbf{x}_0)$ เมื่อ $g(g_1(\mathbf{x}),\ldots,g_5(\mathbf{x}))^T$ จะได้ $g(\mathbf{x}_0) = (0,-132,-229,4594,85)^T$ ให้ t และ f เป็นอาร์เรย์ของเลขดัชนีที่จุด $\mathbf{x}$ ใด ๆ ซึ่ง $g_i(\mathbf{x}) < 0$ และ $g_i(\mathbf{x}) \ge 0$ สำหรับ $i = 1,\ldots,5$ ตามลำดับ (กำหนดเพิ่มเติมว่าให้ f เป็นอาร์เรย์ของเลขดัชนีที่เรียงจากน้อยไปหามาก)ดังนั้นที่จุด $\mathbf{x}_0$ จะได้ว่า $t = [2,3]$, $f = [1,4,5]$ ดังนั้นจะเลือกสมาชิกตัวแรกของอาร์เรย์ f นั่นคือ $f(1) = 1$ เป็นดัชนีของฟังก์ชันจุดประสงค์ และให้สมาชิกทั้งหมดใน t เป็นดัชนีของเงื่อนไขบังคับแล้วแก้ปัญหาต่อไปนี้

$$\left.\begin{array}{ll}
\text{minimize} & g_1(\mathbf{x}) = x_1 - x_2 + x_3 - x_4 + 20\\
\text{subject to} & g_i(\mathbf{x}) \le 0 \ ; \ i = 2,3\\
 & \mathbf{x} \in R^4
\end{array}\right\} \quad (4.11)$$

เมื่อใช้วิธีฟังก์ชันกั้นเขตแปลงปัญหาจะได้ปัญหาใหม่เป็น

$$\left.\begin{array}{ll} \text{minimize} & U_1(\mathbf{x},\mu_k) = g_1(\mathbf{x}) - \mu_k\left[\dfrac{1}{g_2(\mathbf{x})} + \dfrac{1}{g_3(\mathbf{x})}\right] \\ \text{subject to} & g_i(\mathbf{x}) < 0 \quad ; \; i = 2,3 \\ & \mathbf{x} \in R^4 \end{array}\right\} \quad (4.12)$$

ใช้ขั้นตอนวิธี OPTIM 2 ในการแก้ปัญหา (4.12) โดยใช้จุด $\mathbf{x}_0$ เป็นจุดเริ่มต้น

(1.1) OPTIM 2 กำหนดให้ $\varepsilon_1 = \varepsilon_2 = 5\times10^{-6}$ เป็นค่าสเกลาร์ในการหยุดสำหรับลูปใหญ่ k และ ลูปภายใน i ตามลำดับ สำหรับลูป k กำหนดให้ $\beta = 0.1$ และที่ $k = 0$ ให้ $\mu_0 = 10$ สำหรับลูป i กำหนด $\mu = 1/3$ เป็นค่าที่ใช้ในการค้นหาแนวเส้น

ต่อไปนี้จะแสดงผลการคำนวณที่ $k = 0$ และ $i = 0$ อย่างละเอียด

ที่ $k = 0, i = 0$

กำหนด $\mathbf{x}_{0,0} = \mathbf{x}_0 = (10,20,30,40)^T$

$H_{0,0} = I_4$

คำนวณ $\nabla U_1(\mathbf{x}_{0,0}) = (1.0013,-0.9973,0.9973,-1.0013)^T$

$$\mathbf{p}_{0,0} = (-1.0013,0.9973,-0.9954,1.0013)^T$$

$$U_1(\mathbf{x}_{0,0}) = 0.1194$$

$$\alpha_{0,0} = 1$$

$$\mathbf{x}_{0,1} = \mathbf{x}_{0,0} + \alpha_{0,0}\mathbf{p}_{0,0} = (8.9987,20.9973,29.0046,41.0013)^T$$

$$\nabla U_1(\mathbf{x}_{0,1}) = (1.0016,-0.9969,0.9949,-1.0015)^T$$

$$\|\nabla U_1(\mathbf{x}_{0,1})\| = 1.9974094 > \varepsilon_2$$

แต่ $g_1(\mathbf{x}_{0,1}) = -3.99541 < 0$ ในขณะที่ $g_2(\mathbf{x}_{0,1}) < 0$ และ $g_3(\mathbf{x}_{0,1}) < 0$ ดังนั้นขบวนการใน OPTIM 2 ในครั้งนี้จะหยุดดำเนินการ ณ. จุด $\mathbf{x}_{0,1}$ นี้ และได้อาร์เรย์ t, f ใหม่เป็นดังนี้ $t = [2,3,1], f = [4,5]$ เมื่อทำการตรวจสอบค่า $g_4(\mathbf{x}_{0,1}) = 404467\times10^3 > 0$ ดังนั้นต้องทำการแก้ปัญหาค่าต่ำสุดต่อไป โดยให้ $g_4(\mathbf{x})$ เป็นฟังก์ชันจุดประสงค์และ $g_i(\mathbf{x}) \le 0$, $i = 1,2,3$ เป็นเงื่อนไขบังคับ

$$\begin{array}{ll} \text{minimize} & g_4(\mathbf{x}) \\ \text{subject to} & g_i(\mathbf{x}) \le 0 \quad ; \; i = 1,2,3 \\ & \mathbf{x} \in R^4 \end{array}$$

เมื่อใช้วิธีฟังก์ชันกั้นเขตแปลงปัญหาจะได้

$$
\left.\begin{array}{ll}
\text{minimize} & U_2(\mathbf{x},\mu_k) = g_4(\mathbf{x}) - \mu_k\left[\dfrac{1}{g_1(\mathbf{x})}+\dfrac{1}{g_2(\mathbf{x})}+\dfrac{1}{g_3(\mathbf{x})}\right] \\
\text{subject to} & g_i(\mathbf{x}) < 0 \quad ; i = 1,2,3 \\
 & \mathbf{x} \in R^4
\end{array}\right\} \quad (4.13)
$$

ใช้ขั้นตอนวิธี OPTIM 2 อีกครั้งในการแก้ปัญหา (4.13) โดยใช้จุด $\mathbf{x}_{0,0} = \mathbf{x}_{0,1} = (8.9987, 20.9973, 29.0046, 41.0013)^T$ เป็นจุดเริ่มต้น

(1.2) OPTIM 2 ทำการกำหนดค่าต่าง ๆ ตาม 1.1 และดำเนินการเช่นเดียวกับ 1.1 จนได้ $\mathbf{x}_{0,5} = (1.5044, 2.5688, -8.8281, 44.9688)^T$ ซึ่ง $g_4(\mathbf{x}_{0,5}) = -1268.6069 < 0$ ขณะนี้ $t = [2,3,1,4]$, $f = [5]$ เมื่อทำการตรวจสอบค่า $g_5(\mathbf{x}_{0,5}) = -42.7429 < 0$ ดังนั้น t และ f จะเปลี่ยนแปลงอีกครั้งเป็น $t = [2,3,1,4,5]$ และ f ไม่มีสมาชิกอยู่เลย ดังนั้นขบวนการส่วนที่ 1 ทั้งหมดจะหยุดดำเนินการ ซึ่งได้จุด

$$\mathbf{x}_{0,5} \in \{\mathbf{x} \in R^4 \mid g_i(\mathbf{x}) < 0 \quad ; i = 1,2,\ldots,5\}$$

ส่วนที่ 2

ในส่วนนี้ต้องการหาจุดภายในที่เป็นไปได้และสอดคล้องกับเงื่อนไขบังคับทั้งหมด โดยใช้วิธีผสมในการแก้ปัญหา โดยจะใช้ $\mathbf{x}_{0,0} = \mathbf{x}_{0,5}$ เป็นจุดเริ่มต้นในการแก้ปัญหาและทำการเรียกใช้ขั้นตอนวิธี OPTIM 3 แก้ปัญหาต่อไปนี้

$$
\left.\begin{array}{ll}
\text{minimize} & M(\mathbf{x},\mu_k) = 0 - \mu_k\left[\dfrac{1}{g_1(\mathbf{x})}+\ldots+\dfrac{1}{g_5(\mathbf{x})}\right] \\
 & \quad + \dfrac{1}{\mu_k}\left[(h_1(\mathbf{x}))^2 + (h_2(\mathbf{x}))^2 + (h_3(\mathbf{x}))^2\right] \\
\text{subject to} & g_i(\mathbf{x}) < 0 \quad ; i = 1,\ldots,5 \\
 & \mathbf{x} \in R^4
\end{array}\right\} \quad (4.14)
$$

(2.1) OPTIM 3 กำหนดให้ $\varepsilon_1 = \varepsilon_2 = 5 \times 10^{-6}$ เป็นค่าสเกลาร์ในการหยุดลูปใหญ่ k และลูปภายใน i ตามลำดับ สำหรับลูป k กำหนดให้ $\beta = 0.1$ และที่ $k = 0$ กำหนดให้ $\mu_0 = 0.001$ สำหรับลูป i กำหนด $\mu = 1/3$ เป็นค่าที่ใช้ในการค้นหาแนวเส้น ในที่นี้จะได้

แสดงผลการคำนวณโดยละเอียดที่ $k = 0, i = 0$ เท่านั้น สำหรับที่ $i = 1,2,\ldots$ และ $k = 1,2,\ldots$ สามารถทำได้ในทำนองเดียวกัน

ที่ $k = 0, i = 0$

กำหนด $\mathbf{x}_{0,0} = \mathbf{x}_{0,5}$ (จากส่วนที่ 1) และ $H_{0,0} = I_4$

คำนวณ $\nabla M(\mathbf{x}_{0,0}) = 10^8(0.3236, -0.4247, -3.5696, 0.0763)^T$

$$\mathbf{p}_{0,0} = -H_{0,0}\nabla M(\mathbf{x}_{0,0}) = 10^8(-0.3236, 0.4247, 3.5696, -0.0763)^T$$

$$M(\mathbf{x}_{0,0}) = 1.58 \times 10^9$$

$$\alpha_{0,0} = 1.4901 \times 10^{-8}$$

$$\mathbf{x}_{0,1} = (1.0221502, 3.2017604, -3.5090005, 44.8552613)^T$$

$$\nabla M(\mathbf{x}_{0,1}) = 10^8(0.1234, -0.1623, -1.3748, 0.0299)^T$$

$$\|\nabla M(\mathbf{x}_{0,1})\| = 139015320.28207$$

ทำการกำเนิด $H_{0,1}$ จากสูตร (3.1) จะได้

$$H_{0,1} = \begin{bmatrix} 0.9919 & 0.0107 & 0.0892 & -0.0019 \\ 0.0107 & 0.9860 & -0.1169 & 0.0025 \\ 0.0892 & -0.1169 & 0.0225 & 0.0207 \\ -0.0019 & 0.0025 & 0.0207 & 0.9996 \end{bmatrix}$$

จะเห็นว่า $\|\nabla M(\mathbf{x}_{0,1})\| > \varepsilon_2$ ดังนั้นให้ดำเนินการซ้ำดังที่กล่าวข้างต้น เพื่อกำเนิดจุด $\mathbf{x}_{0,2}, \mathbf{x}_{0,3}, \ldots$ จนได้ $\mathbf{x}_{0,59} = (3.981044, 4.390537, -0.341381, 27.469800)^T$ ซึ่ง $\|\nabla M(\mathbf{x}_{0,59})\| = 6.1843 \times 10^{-7} < \varepsilon_2$ เราให้ $\mathbf{x}_1 = \mathbf{x}_{0,59}$ และเมื่อคำนวณ $\frac{1}{\mu_0}\alpha(\mathbf{x}_1) = \frac{1}{\mu_0}[\sum_{j=1}^{3}(h_j(\mathbf{x}_1))^2]$ $= 2.3992 \times 10^{-12} < \varepsilon_1$ แต่ $\mu_0\beta(\mathbf{x}_1) = \mu_0\left[-\sum_{i=1}^{5}\frac{1}{g_i(\mathbf{x}_1)}\right] = 5.1242 \times 10^{-4} > \varepsilon_1$

เมื่อเป็นเช่นนี้ เราจะดำเนินการต่อสำหรับ $k = 1$ และใช้จุด $\mathbf{x}_1$ เป็นจุดเริ่มต้นทำการกำเนิด $\mathbf{x}_2, \ldots$ ภายหลังการคำนวณที่ $k = 3$ ซึ่ง $\mu_3 = 0.000001$ จะได้จุด $\mathbf{x}_4 = (3.9810438, 4.3905374, -0.3413814, 27.4698002)^T$ และ $\mu_3\beta(\mathbf{x}_4) = 5.1242 \times 10^{-7} < \varepsilon_1$ และ $\frac{1}{\mu_3}\alpha(\mathbf{x}_4) = 2.5244 \times 10^{-21} < \varepsilon_1$ ดังนั้นจึงหยุดดำเนินการขั้นตอนวิธี OPITM 3 ที่จุด $\mathbf{x}_4$ นี้ เมื่อคำนวณค่าของเงื่อนไขบังคับแบบสมการที่จุด $\mathbf{x}_4$ จะได้

$$h_1(x_4) = -0.0711 \times 10^{-13}$$
$$h_2(x_4) = 0.4974 \times 10^{-13}$$
$$h_3(x_4) = 0$$

ผลการคำนวณจากการเขียนโปรแกรมคอมพิวเตอร์เป็นดังตารางที่ 4.5 (รายละเอียดการคำนวณสำหรับแต่ละ k จะให้ไว้ในภาคผนวก ก.) #

<u>ปัญหาที่ 6</u> เป็นปัญหาที่ผู้วิจัยสร้างขึ้นเพื่อให้สอดคล้องกับเงื่อนไขบังคับในรูปแบบ (1.3)โดยที่เงื่อนไขบังคับทั้งหมดเป็นฟังก์ชันนูน

หาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบผสมต่อไปนี้

$$\left.\begin{aligned}
g_1(x) &= 5x_1^2 + x_2^2 + 2x_1x_2 - x_1 + 2x_2 + 15x_3 + 3 \le 0 \\
g_2(x) &= 2x_1^2 + x_2^2 - 2x_2 + 6x_3 + 2 \le 0 \\
g_3(x) &= 5x_1 + 3x_2 + 4x_3 + 4 \le 0 \\
g_4(x) &= 4e^{(2x_1 - x_3)} + 5e^{x_2^2} + 30x_3 \le 0 \\
h_1(x) &= e^{(2x_1 + 5x_2)} + 3x_3 + 29 = 0 \\
h_2(x) &= x_1^4 + 2x_2^2 + 3x_3^2 - 4x_1 - 4x_2x_3 - 1033 = 0 \\
h_3(x) &= 10x_1 + 7x_2 - 3x_3 + 6 = 0
\end{aligned}\right\} \quad (4.15)$$

ขั้นตอนดำเนินการ

เราแบ่งการทำงานออกเป็น 2 ส่วน เช่นเดียวกับปัญหาที่ 5

ส่วนที่ 1

หาจุด $x' \in \{x \in R^3 \mid g_i(x) < 0; i = 1,\ldots,4\}$ ใช้จุดเริ่มต้น $x_0 = (-0.35, 6.9, 4.8)^T$ คำนวณ $g(x_0)$ เมื่อ $g(x) = (g_1(x),\ldots,g_4(x))^T$ จะได้ $g(x_0) = (132.5425, 64.855, 42.15, 2.375 \times 10^{21})^T$

ให้ t และ f เป็นอาร์เรย์ของเลขดัชนีที่จุด x ใด ๆ ซึ่ง $g_i(x) < 0$ และ $g_i(x) \ge 0$ สำหรับ $i = 1,\ldots,4$ ตามลำดับ (กำหนดเพิ่มเติมว่าให้ f เป็นอาร์เรย์ของเลขดัชนีที่เรียงจากน้อยไปหามาก) ที่ x_0 จะได้ t ไม่มีสมาชิกอยู่เลย (t = [ ]) และ f = [1,2,3,4] ดังนั้นจะเลือกสมาชิกตัวแรกของอาร์เรย์ f คือ f(1) = 1 เป็นดัชนีของฟังก์ชันจุดประสงค์ แล้วทำการแก้ปัญหาต่อไปนี้

$$\left.\begin{array}{ll}
\text{minimize} & g_1(x) = 5x_1^2 + x_2^2 + 2x_1x_2 - x_1 + 2x_2 + 15x_3 + 3 \\
\text{subject to} & x \in R^3
\end{array}\right\} \quad (4.16)$$

ตารางที่ 4.5 แสดงผลการคำนวณโดยรวมของปัญหาที่ 5

OPTIM 2

ครั้งที่	k	μ_k	$x_{k+1}^T = (x_1, x_2, x_3, x_4)$	ค่าฟังก์ชันจุดประสงค์ที่ x_{k+1}	ค่าเงื่อนไขบังคับที่ x_{k+1}			
1	0	10	(8.9986596, 20.9973192, 29.0045895, 41.0013385)	g(1)=-3.99541	g(2)=-124.0549094	g(3)=-229.0017033	-	-
2	0	10	(1.5044140, 2.5688449, -8.8281198, 44.9688913)	g(4)=-1268.60698	g(2)=-0.2146323	g(3)=-30.7687394	g(1)=-34.8614420	-
	0	10	(1.5044140, 2.5688449, -8.8281198, 44.9688913)	g(5)=-42.74296	g(2)=-0.2146323	g(3)=-30.7687394	g(1)=-34.8614420	g(4)=-1268.6069777

OPTIM 3

k	μ_k	$x_{k+1}^T = (x_1, x_2, x_3, x_4)$	$\mu_k \beta(x_{k+1})$	$\frac{1}{\mu_k}\alpha(x_{k+1})$	ค่าเงื่อนไขบังคับที่ x_{k+1} สมการ	ค่าเงื่อนไขบังคับที่ x_{k+1} อสมการ
0	0.001	(3.9810438, 4.3905374, -0.3413815, 27.4698002)	0.00051242	2.3992e-012	h(1)= -3.5754e-008 , h(2)= 3.2072e-008 , h(3)=9.6023e-009	g(1)=-8.2207, g(2)= -6.2636 , g(3)=-50.9699 , g(4)=-6.9753 , g(5)=-14.6766
1	0.0001	(3.9810438, 4.3905374, -0.3413814, 27.4698002)	0.00005124	2.4018e-015	h(1)= -3.5774e-010 , h(2)=3.2088e-010 , h(3)= 9.6065e-011	g(1)=-8.2207, g(2)= -6.2636 , g(3)=-50.9699 , g(4)=-6.9753 , g(5)=-14.6766
2	0.00001	(3.9810438, 4.3905374, -0.3413814, 27.4698002)	0.00000512	2.4046e-018	h(1)= -3.5811e-012 , h(2)=3.2117e-012 , h(3)= 9.5213e-013	g(1)=-8.2207, g(2)= -6.2636 , g(3)=-50.9699 , g(4)=-6.9753 , g(5)=-14.6766
3	0.000001	(3.9810438, 4.3905374, -0.3413814, 27.4698002)	0.00000051	2.5244e-021	h(1)= -7.1054e-015 , h(2)=4.9738e-014 , h(3)=0	g(1)=-8.2207, g(2)= -6.2636 , g(3)=-50.9699 , g(4)=-6.9753 , g(5)=-14.6766

ใช้ขั้นตอนวิธี OPTIM 1 ในการแก้ปัญหา (4.16) โดยใช้จุด $\mathbf{x}_0$ เป็นจุดเริ่มต้น

(1.1) OPTIM 1 กำหนดค่า $\varepsilon_2 = 5\times10^{-3}$ เป็นค่าสเกลาร์ในการหยุดและในขั้นตอนการค้นหาแนวเส้นกำหนดให้ $\mu = 1/3$ สำหรับเมตริกซ์เริ่มต้นกำหนดให้ $H_0 = I_3$ และเวกเตอร์ $\nabla g_1(\mathbf{x}) = (10x_1 + 2x_2 - 1, 2x_2 + 2x_1 + 2, 15)^T$ ทำการคำนวณค่าต่าง ๆ ดังนี้

$$\nabla g_1(\mathbf{x}_0) = (9.3, 15.1, 15)^T$$
$$g_1(\mathbf{x}_0) = 132.5425$$
$$\mathbf{p}_0 = (-9.3, -15.1, -15)^T$$
$$\alpha_0 = 0.25$$
$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0\mathbf{p}_0 = (-2.675, 3.125, 1.050)^T$$

ซึ่งที่จุด $\mathbf{x}_1$ จะได้ $\nabla g_1(\mathbf{x}_1) = (-21.5, 2.9, 1.5)^T$ และ $\|\nabla g_1(\mathbf{x}_1)\| = 26.375367 > \varepsilon_2$ และ $g_1(\mathbf{x}_1) = 56.5 > 0$ ดังนั้นดำเนินตามขั้นตอนที่กล่าวข้างต้นซ้ำอีก โดยกำเนิด H_1 จาก H_0 โดยใช้ (3.1) จะได้

$$H_1 = \begin{bmatrix} 0.2573 & -0.4589 & -0.2164 \\ -0.4589 & 1.4679 & 0.8536 \\ -0.2164 & 0.8536 & 2.2342 \end{bmatrix}$$

$$\mathbf{p}_1 = (10.1074, -26.9277, -40.6411)^T$$
$$\alpha_1 = 0.5$$
$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha_1\mathbf{p}_1 = (2.3787, -10.3388, -19.2706)^T$$

ที่จุด $\mathbf{x}_2$ จะได้ $\nabla g_1(\mathbf{x}_2) = (2.1093, -13.9203, 15)^T$ และ $\|\nabla g_1(\mathbf{x}_2)\| = 20.5724 > \varepsilon_2$ แต่ $g_1(\mathbf{x}_2) = -223.11828 < 0$ ดังนั้นขั้นตอนวิธี OPTIM 1 จะหยุดดำเนินการที่จุด $\mathbf{x}_2$ นี้และมีการเปลี่ยนแปลงอาร์เรย์ดังนี้ $t = [1], f = [2,3,4]$ เมื่อตรวจสอบ $g_2(\mathbf{x}_2) = 25.2621 > 0$ ดังนั้นดำเนินการแก้ปัญหาต่อไปนี้

$$\left.\begin{array}{ll} \text{minimize} & g_2(\mathbf{x}) \\ \text{subject to} & g_1(\mathbf{x}) \le 0 \\ & \mathbf{x} \in R^3 \end{array}\right\} \quad (4.17)$$

โดยใช้จุด x_2 ใน OPTIM 1 เป็นจุดเริ่มต้นทำการแปลงปัญหา (4.17) โดยวิธีฟังก์ชันกั้นเขต จะได้ปัญหา

$$\left.\begin{array}{ll} \text{minimize} & U_1(x,\mu_k) = g_2(x) - \dfrac{\mu_k}{g_1(x)} \\ \text{subject to} & g_1(x) < 0 \\ & x \in R^3 \end{array}\right\} \quad (4.18)$$

และใช้ขั้นตอนวิธี OPTIM 2 ในการแก้ปัญหา (4.18)

(1.2) OPTIM 2 เพื่อแก้ปัญหา (4.18) กำหนดให้ $\varepsilon_1 = \varepsilon_2 = 5\times10^{-6}$ เป็นค่าสเกลาร์ในการหยุดสำหรับลูปใหญ่ k และลูปภายใน i ตามลำดับ สำหรับลูป k กำหนดให้ $\beta = 0.1$ และที่ $k = 0$ ให้ $\mu_0 = 10$ สำหรับลูป i กำหนด $\mu = 1/3$ เป็นค่าที่ใช้ในการค้นหาแนวเส้น

ต่อไปนี้จะแสดงผลการคำนวณที่ $k = 0$ และ $i = 0$ อย่างละเอียดดังนี้

ที่ $k = 0, i = 0$

กำหนดให้ $x_{0,0} = x_2$ (ใน OPTIM1)

$H_{0,0} = I_3$

คำนวณ $\nabla U_1(x_{0,0}) = (9.5152,-22.6804,6.0030)^T$

$$p_{0,0} = (-9.5152,22.6804,-6.0030)^T$$

$$U_1(x_{0,0}) = 25.30696$$

$$\alpha_{0,0} = 0.5$$

$$x_{0,1} = x_{0,0} + \alpha_{0,0}p_{0,0} = (-2.3789,1.0014,-22.27207)^T$$

$$\nabla U_1(x_{0,1}) = (-9.5181,0.0027135,6.00164)^T$$

$$\|\nabla U_1(x_{0,1})\| = 11.2523 > \varepsilon_2$$

แต่ $g_2(x_{0,1}) = -121.31405 < 0$ ในขณะที่ $g_1(x_{0,1}) = -302.16509 < 0$ ขณะนี้ได้ $t = [1,2]$ และ $f = [3,4]$ และเมื่อตรวจสอบ $g_3(x)$ ที่จุด $x_{0,1}$ จะได้ $g_3(x_{0,1}) = -93.97863 < 0$ ดังนั้นไม่ต้องทำ OPTIM 2 ในรอบนี้และ $t = [1,2,3]$, $f = [4]$

เมื่อคำนวณ $g_4(x)$ ที่จุด $x_{0,1}$ จะได้ $g_4(x_{0,1}) = 1.6158\times10^8 > 0$ ดังนั้น ดำเนินขบวนการ OPTIM 2 อีกครั้ง โดยแก้ปัญหาค่าต่ำสุดของ $g_4(x)$ โดยมี $g_i(x) \le 0$, $i = 1,2,3$ เป็นเงื่อนไขบังคับ และใช้วิธีฟังก์ชันกั้นเขตแปลงปัญหาจะได้

$$
\left.\begin{array}{lll}
\text{minimize} & U_2(\mathbf{x},\mu_k) = g_4(\mathbf{x}) - \mu_k\left[\dfrac{1}{g_1(\mathbf{x})} + \dfrac{1}{g_2(\mathbf{x})} + \dfrac{1}{g_3(\mathbf{x})}\right] \\
\text{subject to} & g_i(\mathbf{x}) < 0 \quad ; i = 1,2,3 \\
& \mathbf{x} \in R^3
\end{array}\right\} \quad (4.19)
$$

(1.3) OPTIM 2 ใช้จุด $\mathbf{x}_{0,0} = \mathbf{x}_{0,1}$ (ใน 1.2 OPTIM 2) เป็นจุดเริ่มต้น การกำหนดค่าและการดำเนินการต่าง ๆ จะเหมือนกับ 1.2 ต่อไปนี้จะแสดงผลการคำนวณที่ $k = 0$ และ $i = 0$ ดังนี้

ที่ $k = 0, i = 0$

กำหนดให้ $\mathbf{x}_{0,0} = \mathbf{x}_{0,1}$ (ใน 1.2 OPTIM 2) และ $H_{0,0} = I_3$

คำนวณ

$$\nabla U_2(\mathbf{x}_{0,0}) = (3.23174 \times 10^8, 27.3004, -1.61587 \times 10^8)^T$$

$$\mathbf{p}_{0,0} = (-3.23174 \times 10^8, -27.3004, 1.61587 \times 10^8)^T$$

$$U_2(\mathbf{x}_{0,0}) = 1.61586526 \times 10^8$$

$$\alpha_{0,0} = 1.49011 \times 10^{-8}$$

$$\mathbf{x}_{0,1} = (-7.19458, 1.00139, -19.86423)^T$$

$$\nabla U_2(\mathbf{x}_{0,1}) = (1907.583, 27.236, -924.302)^T$$

$$\|\nabla U_2(\mathbf{x}_{0,1})\| = 2119.89382 > \varepsilon_2$$

และ $g_4(\mathbf{x}_{0,1}) = 372.37899 > 0$ ดังนั้นต้องดำเนินการต่อสำหรับ $i = 1,2,\ldots$

โดยคำนวณ

$$H_{0,1} = \begin{bmatrix} 0.2000 & -1.5946 \times 10^{-10} & 0.3999 \\ -1.5946 \times 10^{-10} & 1.0000 & 7.9739 \times 10^{-11} \\ 0.3999 & 7.9739 \times 10^{-11} & 0.8000 \end{bmatrix}$$

ทำการกำเนิด $\mathbf{x}_{0,2}, \mathbf{x}_{0,3}, \ldots$ จนถึง $\mathbf{x}_{0,24} = (-2.81395, 0.52644, -9.52733)^T$ ซึ่ง $g_4(\mathbf{x}_{0,24}) = -81.75706 < 0$ ดังนั้นขบวนการ OPTIM 2 ในครั้งนี้ จะหยุดดำเนินการ (ผลการคำนวณโดยละเอียดสำหรับ $k = 0$ นี้จะอยู่ในภาคผนวก ก.) ขณะนี้ได้

$$\mathbf{x}_{0,24} \in \{\mathbf{x} \in R^3 \mid g_i(\mathbf{x}) < 0 \quad ; i = 1,\ldots,4\}$$

และดำเนินการส่วนที่ 2 ต่อไป

ส่วนที่ 2

ในส่วนนี้จะใช้ขั้นตอนวิธี OPTIM 3 เพื่อแก้ปัญหาต่อไปนี้

$$\left.\begin{aligned} &\text{minimize} && M(\mathbf{x},\mu_k) = 0 - \mu_k\left[\sum_{i=1}^{4}\frac{1}{g_i(\mathbf{x})}\right] + \frac{1}{\mu_k}\left[\sum_{j=1}^{3}[h_j(\mathbf{x})]^2\right] \\ &\text{subject to} && g_i(\mathbf{x}) < 0 \quad ; i = 1,\ldots,4 \\ &&& \mathbf{x} \in R^3 \end{aligned}\right\} \quad (4.20)$$

(2.1) OPTIM 3 กำหนดให้ $\varepsilon_1 = 5\times10^{-4}$ เป็นค่าสเกลาร์ในการหยุดสำหรับลูปใหญ่ k และ $\varepsilon_2 = 5\times10^{-5}$ เป็นค่าสเกลาร์ในการหยุดสำหรับลูปภายใน i สำหรับลูป k กำหนดให้ $\beta = 0.1$ และที่ $k = 0$ ให้ $\mu_0 = 0.001$ สำหรับลูป i กำหนดให้ $\mu = 1/3$ เป็นค่าที่ใช้ในการค้นหาแนวเส้น ในที่นี้จะได้แสดงผลการคำนวณสำหรับ $k = 0$ และ $i = 0$ (สำหรับ $i = 1,2,\ldots$ สามารถดูผลการคำนวณได้ในภาคผนวก ก.) กำหนดให้ $\mathbf{x}_0 = \mathbf{x}_{0,0} = \mathbf{x}_{0,24}$

ที่ $k = 0$ และ $i = 0$

ให้ $\mathbf{x}_{0,0} = \mathbf{x}_{0,24}$ (ใน 1.3 OPTIM 2)

$H_{0,0} = I_3$

คำนวณ $\nabla M(\mathbf{x}_{0,0}) = (1.2427\times10^8, -5.3434\times10^7, 7.8903\times10^7)^T$

$$\mathbf{p}_{0,0} = (-1.2427\times10^8, 5.3434\times10^7, 7.8903\times10^7)^T$$

$$M(\mathbf{x}_{0,0}) = 4.438157\times10^8$$

$$\alpha_{0,0} = 1.490116\times10^{-8}$$

$$\mathbf{x}_{0,1} = (-4.66572, -1.32266, -10.70307)^T$$

$$\nabla M(\mathbf{x}_{0,1}) = (1.12146\times10^8, -1.31393\times10^7, 1.89753\times10^7)^T$$

$$\|\nabla M(\mathbf{x}_{0,1})\| = 114497013.455 > \varepsilon_2$$

ดังนั้นดำเนินการต่อสำหรับ $i = 1$ โดยคำนวณ

$$H_{0,1} = \begin{bmatrix} 1.8176 & 0.1681 & -0.2546 \\ 0.1681 & 0.7042 & 0.4394 \\ -0.2546 & 0.4394 & 0.3470 \end{bmatrix}$$

เพื่อหา $\mathbf{x}_{0,2}$ และต่อ ๆ ไปจนได้ $\mathbf{x}_1 = \mathbf{x}_{0,27} = (-5,2,-10)^T$ ซึ่ง $\|\nabla M(\mathbf{x}_{0,27})\| = 0.0000123 < \varepsilon_2$ ดังนั้นจะหลุดลูปภายใน i และเมื่อคำนวณ

$$\left(\frac{1}{\mu_0}\right)\alpha(\mathbf{x}_1) = \left(\frac{1}{\mu_0}\right)[\sum_{j=1}^{3}(h_j(\mathbf{x}_1))^2] = 1.6966\times10^{-11} < \varepsilon_1 \quad \text{และ}$$

$$\mu_0\beta(\mathbf{x}_1) = \mu_0\left[-\sum_{i=1}^{4}\frac{1}{g_i(\mathbf{x}_1)}\right] = 2.2113\times10^{-4} < \varepsilon_1$$

ดังนั้นที่ $k=0$ ซึ่ง $\mu_0 = 0.001$ จะได้จุด $\mathbf{x}_1 = (-5,2,-10)^T$ ซึ่งสอดคล้องกับเงื่อนไขบังคับทั้งหมด โดยที่

$$h_1(\mathbf{x}_1) = -1.197924\times10^{-7} \cong 0$$
$$h_2(\mathbf{x}_1) = -1.8503669\times10^{-9} \cong 0$$
$$h_3(\mathbf{x}_1) = -5.1112117\times10^{-8} \cong 0$$

ดังนั้นจะหยุดดำเนินการ OPTIM 3 ที่จุด $\mathbf{x}_1$ นี้ ผลการคำนวณโดยสรุปของปัญหาที่ 6 นี้ แสดงได้ดังตารางที่ 4.6 (ส่วนผลการคำนวณโดยละเอียดแสดงในภาคผนวก ก.) #

<u>ปัญหาที่ 7</u> เป็นปัญหาที่ผู้วิจัยสร้างขึ้นเพื่อให้สอดคล้องกับเงื่อนไขบังคับในรูปแบบ (1.3)โดยที่เงื่อนไขบังคับแบบอสมการเป็นฟังก์ชันนูนและเงื่อนไขบังคับแบบสมการเป็นฟังก์ชันแบบใด ๆ

หาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบผสมต่อไปนี้

$$\left.\begin{aligned}
g_1(\mathbf{x}) &= 5x_1^2 + x_2^2 + 2x_1x_2 - x_1 + 2x_2 + 15x_3 + 3 \le 0\\
g_2(\mathbf{x}) &= 2x_1^2 + x_2^2 - 2x_2 + 6x_3 + 2 \le 0\\
g_3(\mathbf{x}) &= 5x_1 + 3x_2 + 4x_3 + 4 \le 0\\
g_4(\mathbf{x}) &= 4e^{(2x_1 - x_3)} + 5e^{x_2^2} + 30x_3 \le 0\\
h_1(\mathbf{x}) &= -x_1^2 + 3x_2^3 + \sin(\pi x_3) + 1 = 0\\
h_2(\mathbf{x}) &= -e^{x_1+5} - [\cos(\pi x_2)]^2 - x_3 - 8 = 0\\
h_3(\mathbf{x}) &= 10x_1 + 7x_2 - 3x_3 + 6 = 0\\
h_4(\mathbf{x}) &= -x_1^4 + 2x_2^3 - 3x_3^2 + 909 = 0
\end{aligned}\right\} \quad (4.21)$$

<u>ขั้นตอนดำเนินการ</u>

เนื่องจากส่วนของเงื่อนไขอสมการในปัญหาที่ 7 นี้เหมือนกับปัญหาที่ 6 ดังนั้นเราเริ่มต้นด้วยจุด $\mathbf{x}_0 = (-0.35,6.9,4.8)^T$ เช่นเดียวกับปัญหาที่ 6 ซึ่งขบวนการจะมี 2 ส่วนเช่นเดียวกับปัญหาที่ 6 โดยที่ส่วนที่ 1 จะไม่แสดงซ้ำในที่นี้อีก

ตารางที่ 4.6 แสดงผลการคำนวณโดยรวมของปัญหาที่ 6

OPTIM 1

k	$x_{k+1}^T = (x_1, x_2, x_3)$	$\lVert \nabla g_1(x_{k+1}) \rVert$	ค่าฟังก์ชันจุดประสงค์ที่ x_{k+1}
0	(-2.6750000, 3.1250000, 1.0500000)	26.375367	g(1)=56.50000
1	(2.3786942, -10.3388346, -19.2705666)	20.572390	g(1)=-223.11828

OPTIM 2

ครั้งที่	k	μ_k	$x_{k+1}^T = (x_1, x_2, x_3)$	ค่าฟังก์ชันจุดประสงค์ที่ x_{k+1}	ค่าเงื่อนไขบังคับที่ x_{k+1}			
1	0	10	(-2.3789061, 1.0013981, -22.2720732)	g(2)=-121.31405	g(1)=-302.1650905	-	-	-
	0	10	(-2.3789061, 1.0013981, -22.2720732)	g(3)=-93.97863	g(1)=-302.1650905	g(2)=-121.3140488	-	-
2	0	10	(-2.8139522, 0.5264390, -9.5273287)	g(4)=-81.72706	g(1)=-99.1370752	g(2)=-40.1030580	g(3)=-46.5997591	-

OPTIM 3

k	μ_k	$x_{k+1}^T = (x_1, x_2, x_3)$	$\mu_k\beta(x_{k+1})$	$\frac{1}{\mu_k}\alpha(x_{k+1})$	ค่าเงื่อนไขบังคับที่ x_{k+1}	
					สมการ	อสมการ
0	0.001	(-5.0000000, 2.0000000, -10.0000000)	0.00022113	1.6966e-011	h(1)= -1.1979e-007 , h(2)=-1.8504e-009 , h(3)=-5.1112e-008	g(1)=-29.0000002 , g(2)=-8.0000001 , g(3)=-55.0000001 , g(4)=-23.0092665

ส่วนที่ 1

ได้จุด $\mathbf{x}' = (-2.81395, 0.52644, -9.52733)^T \in \{\mathbf{x} \in R^3 \mid g_i(\mathbf{x}) < 0 \ ; \ i = 1, \ldots, 4\}$ และผลการคำนวณเป็นเช่นเดียวกับผลการคำนวณในส่วนที่ 1 ของปัญหาที่ 6

ส่วนที่ 2

ในส่วนนี้จะใช้วิธีผสมโดยขั้นตอนวิธี OPTIM 3 เพื่อแก้ปัญหาต่อไปนี้

$$\left.\begin{array}{ll} \text{minimize} & M(\mathbf{x}, \mu_k) = 0 - \mu_k\left[\sum\limits_{i=1}^{4} \dfrac{1}{g_i(\mathbf{x})}\right] + \dfrac{1}{\mu_k}\left[\sum\limits_{j=1}^{4} (h_j(\mathbf{x}))^2\right] \\ \text{subject to} & g_i(\mathbf{x}) < 0 \quad ; \ i = 1, .., 4 \\ & \mathbf{x} \in R^3 \end{array}\right\} \qquad (4.22)$$

(2.1) OPTIM 3 กำหนดให้ $\varepsilon_1 = 5 \times 10^{-5}$ เป็นค่าสเกลาร์ในการหยุด สำหรับลูปใหญ่ k และ $\varepsilon_2 = 5 \times 10^{-3}$ เป็นสเกลาร์ในการหยุดสำหรับลูปภายใน i สำหรับลูป k กำหนดให้ $\beta = 0.1$ และที่ $k = 0$ ให้ $\mu_0 = 0.001$ สำหรับลูป i กำหนด $\mu = 1/3$ เป็นค่าที่ใช้ในการค้นหาแนวเส้น ในที่นี้จะแสดงผลการคำนวณโดยละเอียดสำหรับ $k = 0$ และ $i = 0$ เท่านั้น

ที่ $k = 0$ และ $i = 0$

ให้ $\mathbf{x}_0 = \mathbf{x}_{0,0} = \mathbf{x}'$ (จากส่วนที่ 1)

$H_{0,0} = I_3$

คำนวณ

$$\nabla M(\mathbf{x}_{0,0}) = (1.0264 \times 10^8, 0.0203 \times 10^8, 0.6561 \times 10^8)^T$$

$$\mathbf{p}_{0,0} = (-1.0264 \times 10^8, -0.0203 \times 10^8, -0.6561 \times 10^8)^T$$

$$M(\mathbf{x}_{0,0}) = 3.2999 \times 10^8$$

$$\alpha_{0,0} = 1.4901 \times 10^{-8}$$

$$\mathbf{x}_{0,1} = \mathbf{x}_{0,0} + \alpha_{0,0}\mathbf{p}_{0,0} = (-4.3434, 0.4962, -10.5050)^T$$

$$\nabla M(\mathbf{x}_{0,1}) = (1.4534 \times 10^8, 5.4 \times 10^5, 2.804 \times 10^7)^T$$

$$\|\nabla M(\mathbf{x}_{0,1})\| = 148019083.19233 > \varepsilon_2$$

$$H_{0,1} = \begin{bmatrix} 5.7249 & 0.2188 & 6.4975 \\ 0.2188 & 1.0068 & 0.2087 \\ 6.4975 & 0.2087 & 7.3759 \end{bmatrix}$$

และทำต่อในรอบที่ $i = 1,2,\ldots$ เพื่อกำเนิด $\mathbf{x}_{0,2}, \mathbf{x}_{0,3}, \ldots$ จนได้ $\mathbf{x}_1 = \mathbf{x}_{0,34} = (-5,2,-10)^T$ ซึ่ง $\|\nabla M(\mathbf{x}_{0,34})\| = 0.0011730 < \varepsilon_2$

และ $$\left(\frac{1}{\mu_0}\right)\alpha(\mathbf{x}_1) = \left(\frac{1}{\mu_0}\right)\left[\sum_{j=1}^{4}(h_j(\mathbf{x}_1))^2\right] = 1.1187 \times 10^{-12} < \varepsilon_1$$

แต่ $$\mu_0\beta(\mathbf{x}_1) = \mu_0\left[-\sum_{i=1}^{4}\frac{1}{g_i(\mathbf{x}_1)}\right] = 0.00022113 > \varepsilon_1$$

ดังนั้นดำเนินการต่อสำหรับ $k = 1$ ซึ่ง $\mu_1 = 0.0001$ จนได้จุด $\mathbf{x}_2 = \mathbf{x}_{0,27} = (-5,2,-10)^T$ เช่นเดิม ซึ่ง $\|\nabla M(\mathbf{x}_{0,27})\| = 0.0004499 < \varepsilon_2$

และ $$\left(\frac{1}{\mu_1}\right)\alpha(\mathbf{x}_2) = \left(\frac{1}{\mu_1}\right)\left[\sum_{j=1}^{4}(h_j(\mathbf{x}_2))^2\right] = 3.4294 \times 10^{-16} < \varepsilon_1$$

และ $$\mu_1\beta(\mathbf{x}_2) = \mu_1\left[-\sum_{i=1}^{4}\frac{1}{g_i(\mathbf{x}_2)}\right] = 2.2113 \times 10^{-5} < \varepsilon_1$$

ดังนั้นที่จุด $\mathbf{x}_2$ จะเป็นจุดภายในที่เป็นไปได้ นั่นคือ $\mathbf{x}_2 \in \{x \in R^3 \mid g_i(x) < 0 \; ; i = 1,\ldots,4$ และ $h_j(x) \cong 0 \; ; j = 1,\ldots,4\}$ โดยที่

$$\begin{aligned}
h_1(\mathbf{x}_2) &= -1.54815 \times 10^{-10} \\
h_2(\mathbf{x}_2) &= 2.1922019 \times 10^{-11} \\
h_3(\mathbf{x}_2) &= 8.710543 \times 10^{-11} \\
h_4(\mathbf{x}_2) &= 4.752109 \times 10^{-11}
\end{aligned}$$

ซึ่งผลการคำนวณโดยสรุปสำหรับปัญหาที่ 7 นี้ แสดงในตารางที่ 4.7 และผลการคำนวณโดยละเอียดจะอยู่ในภาคผนวก ก. #

<u>ปัญหาที่ 8</u> เป็นปัญหาที่ผู้วิจัยสร้างขึ้นเพื่อให้สอดคล้องกับเงื่อนไขบังคับ (1.3) โดยที่เงื่อนไขบังคับแบบอสมการและเงื่อนไขบังคับแบบสมการเป็นฟังก์ชันแบบใด ๆ

หาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบผสมต่อไปนี้

ตารางที่ 4.7 แสดงผลการคำนวณโดยรวมของปัญหาที่ 7

OPTIM 1

k	$x_{k+1}^T = (x_1, x_2, x_3)$	$\lVert \nabla g_1(x_{k+1}) \rVert$	ค่าฟังก์ชันจุดประสงค์ที่ x_{k+1}
0	(-2.6750000, 3.1250000, 1.0500000)	26.375367296	g(1)=56.50000
1	(2.3786942, -10.3388346, -19.2705666)	20.572390442	g(1)=-223.11828

OPTIM 2

ครั้งที่	k	μ_k	$x_{k+1}^T = (x_1, x_2, x_3)$	ค่าฟังก์ชันจุดประสงค์ที่ x_{k+1}	ค่าเงื่อนไขบังคับที่ x_{k+1}			
1	0	10	(-2.3789061, 1.0013981, -22.2720732)	g(2)=-121.31405	g(1)=-302.1650905	-	-	-
	0	10	(-2.3789061, 1.0013981, -22.2720732)	g(3)=-93.97863	g(1)=-302.1650905	g(2)=-121.3140488	-	-
2	0	10	(-2.8139522, 0.5264390, -9.5273287)	g(4)=-81.72706	g(1)=-99.1370752	g(2)=-40.1030580	g(3)=-46.5997591	-

OPTIM 3

k	μ_k	$x_{k+1}^T = (x_1, x_2, x_3)$	$\mu_k \beta(x_{k+1})$	$\frac{1}{\mu_k}\alpha(x_{k+1})$	ค่าเงื่อนไขบังคับที่ x_{k+1} สมการ	ค่าเงื่อนไขบังคับที่ x_{k+1} อสมการ
0	0.001	(-5.0000000, 2.0000000, -10.0000000)	0.0002211253570006	1.1187e-012	h(1)= -3.1439e-008 , h(2)= 3.1566e-009 , h(3)= 1.0796e-008 , h(4)= 1.9349e-009	g(1)=-29.0000001 , g(2)= -8.0000000, g(3)= -55.0000000 , g(4)=-23.0092507
1	0.0001	(-5.0000000, 2.0000000, -10.0000000)	0.0000221125359181	3.4294e-016	h(1)= -1.5482e-010 , h(2)= 2.1922e-011, h(3)= 8.7105e-011 , h(4)= 4.7521e-011	g(1)=-29.0000000 , g(2)= -8.0000000, g(3)= -55.0000000 , g(4)=-23.0092498

$$\left.\begin{aligned}
g_1(\mathbf{x}) &= x_1^4 + 2x_2^2 - 3x_3 - 4x_1 - 4x_1x_3 + 390 \le 0\\
g_2(\mathbf{x}) &= 2x_1^2 + x_2^2 + 2x_2x_3 - x_4^{1/2} + 1330 \le 0\\
g_3(\mathbf{x}) &= x_1 + 2x_2 + 3x_3 + x_4 - 285 \le 0\\
g_4(\mathbf{x}) &= e^{x_2} - x_3 + x_4 + 95 \le 0\\
g_5(\mathbf{x}) &= \ln(x_1^2 + 0.75) + \cos(x_2 + x_3) - x_4 \le 0\\
h_1(\mathbf{x}) &= x_1^2 + x_2 + x_3^2 - x_4 - 9794.25 = 0\\
h_2(\mathbf{x}) &= -e^{0.5-x_1} - x_2x_3 + 5x_4 - 692 = 0\\
h_3(\mathbf{x}) &= (x_1 + 0.5)^3 + \sin(13x_2 + x_3 - 8) + \ln(x_4^2 + 1) - 1 = 0
\end{aligned}\right\} \qquad (4.23)$$

<u>**ขั้นตอนดำเนินการ**</u>

แบ่งการทำงานออกเป็น 2 ส่วนคือ ส่วนที่ 1 ใช้การทำซ้ำของวิธีฟังก์ชันกั้นเขตเพื่อหาจุด $\mathbf{x}' \in \{\mathbf{x} \in R^4 \mid g_i(\mathbf{x}) < 0 \;;\; i = 1,\ldots,5\}$ และส่วนที่ 2 หาจุดภายในที่เป็นไปได้ $\mathbf{x}^* \in \{\mathbf{x} \in R^4 \mid g_i(\mathbf{x}) < 0 \;;\; i = 1,\ldots,5$ และ $h_j(\mathbf{x}) = 0 \;;\; j = 1,2,3\}$

ส่วนที่ 1

เริ่มต้นใช้จุด $\mathbf{x}_0 = (-2,5,0,10)^T$ คำนวณ $\mathbf{g}(\mathbf{x}_0)$ เมื่อ $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}),\ldots,g_5(\mathbf{x}))^T$ จะได้ $\mathbf{g}(\mathbf{x}_0) = (464,1359.8377,-267,253.4131,-8.1582)^T$ ให้ t และ f เป็นอาร์เรย์ของเลขดัชนีที่จุด $\mathbf{x}$ ใด ๆ ซึ่ง $g_i(\mathbf{x}) < 0$ และ $g_i(\mathbf{x}) \ge 0$ สำหรับ $i = 1,\ldots,5$ ตามลำดับ (โดยกำหนดเพิ่มเติมว่า f เป็นอาร์เรย์ที่เรียงเลขดัชนีจากน้อยไปหามาก) และที่จุด $\mathbf{x}_0$ จะได้ว่า $t = [3,5], f = [1,2,4]$ ดังนั้นจะเลือกให้ $g_1(\mathbf{x})$ เป็นฟังก์ชันจุดประสงค์ และ $g_3(\mathbf{x}) \le 0$, $g_5(\mathbf{x}) \le 0$ เป็นเงื่อนไขบังคับ แล้วดำเนินการแก้ปัญหาค่าต่ำสุดต่อไปนี้โดยวิธีฟังก์ชันกั้นเขต

$$\left.\begin{aligned}
&\text{minimize} && U_1(\mathbf{x},\mu_k) = g_1(\mathbf{x}) - \mu_k\left[\frac{1}{g_3(\mathbf{x})} + \frac{1}{g_5(\mathbf{x})}\right]\\
&\text{subject to} && g_i(\mathbf{x}) < 0 \;;\; i = 3,5\\
& && \mathbf{x} \in R^4
\end{aligned}\right\} \qquad (4.24)$$

ใช้ขั้นตอนวิธี OPTIM 2 ในการแก้ปัญหา (4.24) โดยใช้จุด $\mathbf{x}_0$ เป็นจุดเริ่มต้น

(1.1) OPTIM 2 กำหนดให้ $\varepsilon_1 = \varepsilon_2 = 5 \times 10^{-6}$ เป็นค่าสเกลาร์ในการหยุดสำหรับลูปใหญ่ k และลูปภายใน i ตามลำดับ สำหรับลูป k กำหนดให้ $\beta = 0.1$ และที่ $k = 0$ ให้ $\mu_0 = 10$ สำหรับลูป i กำหนด $\mu = 1/3$ เป็นค่าที่ใช้ในการค้นหาแนวเส้น ในที่นี้จะ

แสดงผลการคำนวณที่ $k = 0$ และ $i = 0$ อย่างละเอียด เพื่อกำเนิด $\mathbf{x}_{0,1}$ (สำหรับ $i = 1,2,\ldots$ และ $k = 1,2,\ldots$ สามารถดูผลการคำนวณได้ในภาคผนวก ก.)

ที่ $k = 0$ และ $i = 0$

กำหนด $\mathbf{x}_{0,0} = \mathbf{x}_0 = (-2,5,0,10)^T$

$H_{0,0} = I_4$

คำนวณ $\nabla U_1(\mathbf{x}_{0,0}) = (-36.1264,20.1444,5.1445,-0.1501)^T$

$\mathbf{p}_{0,0} = (36.1264,-20.1444,-5.1445,0.1501)^T$

$U_1(\mathbf{x}_{0,0}) = 465.2632$

$\alpha_{0,0} = 0.0625$

$\mathbf{x}_{0,1} = (0.257899,3.740977,-0.321531,10.009382)^T$

$\nabla U_1(\mathbf{x}_{0,1}) = (-2.5945,14.9862,-4.0092,-0.0800)^T$

$\|\nabla U_1(\mathbf{x}_{0,1})\| = 15.72885 > \varepsilon_2$

เมื่อคำนวณ $g_1(\mathbf{x}_{0,1}) = 418.2894$ ในขณะที่ $g_3(\mathbf{x}_{0,1}) = -268.2153 < 0$ และ $g_5(\mathbf{x}_{0,1}) = -11.17374 < 0$ คำนวณ

$$H_{0,1} = \begin{bmatrix} 0.1497 & 0.1150 & 0.2371 & -0.0016 \\ 0.1150 & 1.1360 & -0.0811 & -0.0005 \\ 0.2371 & -0.0811 & 0.9496 & 0.0007 \\ -0.0016 & -0.0005 & 0.0007 & 1.0000 \end{bmatrix}$$

ดังนั้นที่ $i = 1$ จะใช้ $H_{0,1}$ คำนวณ $\mathbf{p}_{0,1}$ และทำการคำนวณเช่นเดียวกับที่กล่าวข้างต้น เพื่อกำเนิด $\mathbf{x}_{0,2}, \mathbf{x}_{0,3}, \ldots$ จนได้ $\mathbf{x}_1 = \mathbf{x}_{0,14} = (2.5619837,-0.5846846,32.7256545, 10.6215132)^T$ ซึ่ง $g_1(\mathbf{x}_1) = -10.02861 < 0$ ดังนั้นให้หยุดดำเนินการ OPTIM 2 ในครั้งนี้ ขณะนี้ $t = [3,5,1]$, $f = [2,4]$ เมื่อหาค่า $g_2(\mathbf{x}_1)$ จะได้ $g_2(\mathbf{x}_1) = 1301.9419 > 0$ ดังนั้น ให้เรียกใช้ขั้นตอนวิธี OPTIM 2 ซ้ำอีกกระทำกับปัญหาต่อไปนี้

$$\left.\begin{array}{ll} \text{minimize} & U_2(\mathbf{x},\mu_k) = g_2(\mathbf{x}) - \mu_k\left[\dfrac{1}{g_1(\mathbf{x})} + \dfrac{1}{g_3(\mathbf{x})} + \dfrac{1}{g_5(\mathbf{x})}\right] \\ \text{subject to} & g_1(\mathbf{x}) < 0 \\ & g_3(\mathbf{x}) < 0 \\ & g_5(\mathbf{x}) < 0 \\ & \mathbf{x} \in R^4 \end{array}\right\} \quad (4.25)$$

(1.2) OPTIM 2 ทำการกำหนดค่าต่าง ๆ ตาม 1.1 และดำเนินการเช่นเดียวกันโดยให้ $\mathbf{x}_0 = \mathbf{x}_{0,0} = \mathbf{x}_{0,14}$ (ใน 1.1 OPTIM 2) เป็นจุดเริ่มต้น เมื่อ $k = 0$ และ $i = 64$ จะได้จุด $\mathbf{x}_1 = \mathbf{x}_{0,65} = (3.9551480, -14.5462338, 55.385148, 89.009525)^T$ ซึ่งทำให้ $g_2(\mathbf{x}_{0,65}) = -47.84581 < 0$ ดังนั้นให้หยุดดำเนินการ OPTIM 2 ในครั้งนี้ ขณะนี้จะได้ $t = [3,5,1,2]$ และ $f = [4]$ เมื่อคำนวณ $g_4(\mathbf{x}_{0,65}) = 128.6243 > 0$ ดังนั้นใช้ OPTIM 2 อีกครั้งในการแก้ปัญหาต่อไปนี้

$$\left.\begin{array}{ll} \text{minimize} & U_3(\mathbf{x},\mu_k) = g_4(\mathbf{x}) - \mu_k\left[\dfrac{1}{g_1(\mathbf{x})} + \dfrac{1}{g_2(\mathbf{x})} + \dfrac{1}{g_3(\mathbf{x})} + \dfrac{1}{g_5(\mathbf{x})}\right] \\ \text{subject to} & g_1(\mathbf{x}) < 0 \\ & g_2(\mathbf{x}) < 0 \\ & g_3(\mathbf{x}) < 0 \\ & g_5(\mathbf{x}) < 0 \\ & \mathbf{x} \in R^4 \end{array}\right\} \quad (4.26)$$

(1.3) OPTIM 2 ทำงานเช่นเดียวกับ 1.1 และ 1.2 โดยใช้จุด $\mathbf{x}_0 = \mathbf{x}_{0,0} = \mathbf{x}_{0,65}$ (ใน 1.2 OPTIM 2) เป็นจุดเริ่มต้นในการแก้ปัญหา สำหรับที่ $k = 0$ และ $i = 43$ จะได้จุด $\mathbf{x}_1 = \mathbf{x}_{0,44} = (5.06365, -26.5166, 104.8598, 5.3728)^T$ ซึ่ง $g_4(\mathbf{x}_{0,44}) = -4.48703 < 0$ ดังนั้นให้หยุดดำเนินการ OPTIM 2 ในครั้งนี้

ขณะนี้เราได้จุด $\mathbf{x}' = \mathbf{x}_{0,44}$ (ใน 1.3 OPTIM 3) $\in \{\mathbf{x} \in R^4 \mid g_i(\mathbf{x}) < 0 \;;\; i = 1,\ldots,5\}$

ส่วนที่ 2

ใช้ขั้นตอนวิธี OPTIM 3 ในการแก้ปัญหาต่อไปนี้ โดยใช้ $\mathbf{x}'$ (ซึ่งเป็นผลลัพธ์จากส่วนที่ 1) เป็นจุดเริ่มต้น

$$\text{minimize} \qquad M(x,\mu_k) = 0 - \mu_k\left[\sum_{i=1}^{5}\frac{1}{g_i(x)}\right] + \frac{1}{\mu_k}\left[\sum_{j=1}^{3}(h_j(x))^2\right]$$

$$\text{subject to} \qquad g_i(x) < 0 \quad ; \; i = 1,\ldots,5$$

$$x \in R^4$$

สำหรับ OPTIM 3 กำหนดให้ $\varepsilon_1 = 0.09$ เป็นค่าสเกลาร์ในการหยุดลูปใหญ่ k และ $\varepsilon_2 = 0.0005$ เป็นค่าสเกลาร์ในการหยุดสำหรับลูปภายใน i สำหรับลูป k กำหนดให้ $\beta = 0.1$ และที่ $k = 0$ ให้ $\mu_0 = 0.001$ สำหรับลูป i กำหนด $\mu = 1/3$ เป็นค่าที่ใช้ในการค้นหาแนวเส้น จากผลลัพธ์การคำนวณโดยโปรแกรมที่ใช้ขั้นตอนวิธี OPTIM 3 จะได้ว่าที่ $\mu_0 = 0.001$ และ $\mu_1 = 0.0001$ จะให้ผลลัพธ์เป็นจุด x_1 และ x_2 ตามลำดับ โดยที่ $\mathbf{x}_2 = (0.47742, -6.99558, 99.00054, 0.091202)^T$ ซึ่ง

$$\begin{aligned} h_1(x_2) &= 0.0000062 \\ h_2(x_2) &= 0.0000115 \\ h_3(x_2) &= -0.0000361 \end{aligned}$$

ผลการคำนวณโดยละเอียดได้ให้ไว้ในภาคผนวก ก. ดังนั้น $\mathbf{x}^* = \mathbf{x}_2 \in \{x \in R^4 \mid g_i(x) < 0 \; ; \; i = 1,\ldots,5$ และ $h_j(x) \cong 0 \; ; \; j = 1,2,3\}$ สำหรับผลการคำนวณโดยสรุปของปัญหาที่ 8 นี้เป็นดังตารางที่ 4.8 #

ตารางที่ 4.8 แสดงผลการคำนวณโดยรวมของปัญหาที่ 8

OPTIM 2

ครั้งที่	k	μ_k	$x_{k+1}^T = (x_1, x_2, x_3, x_4)$	ค่าฟังก์ชันจุดประสงค์ที่ x_{k+1}	ค่าเงื่อนไขบังคับที่ x_{k+1}			
1	0	10	(2.5619837, -0.5846846, 32.7256545,10.6215132)	g(1)=-10.02861	g(3)=-174.8089089	g(5)=-7.8832850	-	-
2	0	10	(3.9551480,-14.5462338,55.3851483,89.0095256)	g(2)=-47.84581	g(3)=-54.9723489	g(5)=-87.2126577	g(1)=-0.3064651	-
3	0	10	(5.0636544,-26.5116631,104.8598613,5.3728284)	g(4)=-4.48703	g(3)=-13.0072593	g(5)=-3.0815180	g(1)=-5.5538725	g(2)=-3478.1870981

OPTIM 3

k	μ_k	$x_{k+1}^T = (x_1, x_2, x_3, x_4)$	$\mu_k \beta(x_{k+1})$	$\frac{1}{\mu_k}\alpha(x_{k+1})$	ค่าเงื่อนไขบังคับที่ x_{k+1}	
					สมการ	อสมการ
0	0.001	(0.4800408, -6.9575236, 99.0066201, 1.4139170)	1.16364523	1.0825e+004	h(1)= -8.0172e-002 , h(2)= 2.8903e+000, h(3)= 1.5698e+000	g(1)= -2.1815071 , g(2)= -0.0028675 , g(3)= -0.0012291 , g(4)=2.5917517, g(5)= -2.0211076
1	0.0001	(0.4774200, -6.9955858, 99.0005498, 0.0912020)	0.00261112	1.4736e-005	h(1)= 6.1729e-006 , h(2)= 1.1531e-005, h(3)= -3.6091e-005	g(1)= -0.0423235, g(2)= -6.0415929, g(3)= -1.4209000, g(4)= -3.9084319 , g(5)=-0.7360868

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ในงานวิจัยนี้ได้เสนอขั้นตอนวิธีในการหาจุดที่เป็นไปได้ของเงื่อนไขบังคับในรูปแบบต่าง ๆ คือ เงื่อนไขบังคับแบบสมการ , เงื่อนไขบังคับแบบอสมการ , เงื่อนไขบังคับแบบผสม (ทั้งสมการและอสมการ) โดยที่ขั้นตอนวิธีที่ใช้แก้ปัญหาเพื่อหาจุดที่เป็นไปได้ดังกล่าวได้เสนอไว้ในบทที่ 3 แล้ว ดังนั้นในส่วนนี้ผู้วิจัยจะขอกล่าวถึงปัญหาที่เกิดขึ้นกับงานวิจัยดังต่อไปนี้

1.ในการแก้ปัญหาค่าต่ำสุดแบบมีเงื่อนไขบังคับ ผู้วิจัยใช้วิธีฟังก์ชันทัณฑ์หรือฟังก์ชันกั้นเขตในการแปลงปัญหาไปเป็นปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับหรือมีเงื่อนไขบังคับอย่างง่าย เมื่อได้ปัญหาค่าต่ำสุดแบบไม่มีเงื่อนไขบังคับแล้ว ผู้วิจัยดำเนินการแก้ปัญหาดังกล่าวโดยวิธีแนวนิวตันโดยใช้สูตร BFGS ผกผันซึ่งทฤษฎีบทที่ 2.27 รับประกันว่าวิธีการดังกล่าวจะกำเนิดจุดซึ่งเป็นผลเฉลยของปัญหาแบบไม่มีเงื่อนไขบังคับได้(ในกรณีที่ปัญหาแบบไม่มีเงื่อนไขบังคับนั้นมีผลเฉลย)ถ้าเฮซเซียนเมตริกซ์ $\nabla^2 f(x)$ ของฟังก์ชันจุดประสงค์ $f(x)$ เป็นเมตริกซ์ที่เป็นบวกแน่นอนสำหรับทุก ๆ x ในโดเมน แต่ในกรณีทั่วไปฟังก์ชันจุดประสงค์แบบไม่มีเงื่อนไขบังคับ $f(x)$ ที่ได้จากการแปลงปัญหาโดยวิธีฟังก์ชันกั้นเขตหรือฟังก์ชันทัณฑ์บางฟังก์ชันอาจมีเมตริกซ์เฮซเซียน $\nabla^2 f(x)$ ที่ไม่เป็นบวกแน่นอนสำหรับบาง x ดังนั้นในกรณีนี้อาจไม่ได้ผลเฉลยของปัญหา

2.ปัญหาการเลือกจุดเริ่มต้น x_0 เพื่อใช้หาจุดที่เป็นไปได้ของเงื่อนไขบังคับแบบต่าง ๆ ตามขั้นตอนวิธีที่ได้กล่าวไว้ในบทที่ 3 เนื่องจากมีปัจจัยหลายอย่างที่ทำให้จุดเริ่มต้น x_0 บางจุดไม่สามารถไปถึงจุดที่เป็นไปได้ที่เราต้องการ เช่น

2.1ลักษณะของฟังก์ชันเงื่อนไขบังคับแต่ละเงื่อนไข เช่น เป็นฟังก์ชันรากที่สองหรือฟังก์ชันลอการิทึม ดังนั้นการเลือกจุด x_0 ในการเริ่มต้นแก้ปัญหาก็จะมีจุดให้เลือกอยู่ภายในโดเมนที่ทำให้ฟังก์ชันดังกล่าวนิยามได้ ถ้าเลือกจุด x_0 อยู่นอกโดเมนดังกล่าวก็ไม่มีทางเป็นไปได้ที่จะดำเนินการแก้ปัญหา จะเห็นว่าปัญหาในข้อนี้เราสามารถควบคุมได้โดยพิจารณาจากฟังก์ชันเงื่อนไขบังคับ

2.2จุดเริ่มต้น x_0 บางจุดอาจกำเนิดจุด x_k ที่อยู่นอกโดเมนของฟังก์ชันเงื่อนไขบางฟังก์ชันทำให้ไม่สามารถดำเนินการแก้ปัญหาต่อไปได้

2.3เนื่องจากวิธีแก้ปัญหาค่าต่ำสุดโดยวิธีนิวตันไม่รับประกันการลู่เข้าของจุดเริ่มต้น x_0 บางจุด ดังนั้นวิธีแนวนิวตันซึ่งอยู่บนพื้นฐานของวิธีนิวตันจึงไม่รับประกันการลู่เข้าผลเฉลยของจุด x_0 บางจุดเช่นกัน

5.2 ข้อเสนอแนะ

ในการหาจุดภายในที่เป็นไปได้ของเงื่อนไขบังคับแบบอสมการ

$$g_i(x) \le 0 \quad ; i = 1,\ldots,m$$

โดยการทำซ้ำวิธีฟังก์ชันกั้นเขต ผู้วิจัยจะกำหนดจุดเริ่มต้น x_0 และทำการแบ่งเงื่อนไขบังคับออกเป็น 2 ส่วนโดยหาค่า $g_i(x_0)$ สำหรับ $i = 1,\ldots,m$ โดยให้

$$T = \{t \mid g_t(x) < 0 \quad ; 1 \le t \le m\}$$

และ

$$F = \{f \mid g_f(x) \ge 0 \quad ; 1 \le f \le m\}$$

ดังนั้นในตอนเริ่มต้นจะมีเซต T และ F สำหรับจุด x_0 และผู้วิจัยจะให้ $g_{f_1}(x)$ เป็นฟังก์ชันจุดประสงค์สำหรับบาง $f_1 \in F$ และใช้ $g_t(x) \le 0$ สำหรับทุก ๆ $t \in T$ เป็นเงื่อนไขบังคับแล้วดำเนินการหาค่าต่ำสุดโดยวิธีฟังก์ชันกั้นเขต ซึ่งขบวนการนี้จะหยุดก็ต่อเมื่อได้จุด x_1 ซึ่งเป็นจุดต่ำสุดวงกว้างของปัญหา

$$\begin{aligned}&\text{minimize } g_{f_1}(x)\\&\text{subject to } g_t(x) \le 0 \quad ; \forall t \in T \ , \ x \in R^n\end{aligned}$$

และ $g_{f_1}(x_1) \ge 0$ สำหรับกรณีนี้จะสรุปว่า $\{x \in R^n \mid g_i(x) < 0 \quad ; i = 1,\ldots, m\} = \phi$ แต่ถ้าขบวนการฟังก์ชันกั้นเขตกำเนิดจุด x_1 ซึ่งอาจจะเป็น / ไม่เป็นจุดต่ำสุดวงกว้างของปัญหาดังกล่าว แต่ทำให้ $g_{f_1}(x_1) < 0$ ในกรณีนี้เราจะเพิ่มดัชนี f_1 ให้กับเซต T และหักดัชนี f_1 ออกจากเซต F ดังนั้นขณะนี้ F จะมีจำนวนสมาชิกลดลงและเราจะดำเนินการต่อโดยเลือก $g_{f_2}(x)$ เป็นฟังก์ชันจุดประสงค์ สำหรับบาง $f_2 \in F$ (ใหม่) และ $g_t(x) \le 0$ สำหรับทุก ๆ $t \in T$ (ใหม่) เป็นเงื่อนไขบังคับ แล้วดำเนินการหาค่าต่ำสุดโดยวิธีฟังก์ชันกั้นเขตซ้ำอีกครั้ง ถ้าเซตภายในของบริเวณที่เป็นไปได้ของเงื่อนไขบังคับไม่เป็นเซตว่างแล้วการดำเนินการจะซ้ำไปเรื่อย ๆ จนในที่สุดจำนวนสมาชิกของ T มีค่าเท่ากับ m นั่นคือ $n(T) = m$ และได้ $x^* \in \{x \in R^n \mid g_i(x) < 0 \quad ; i = 1,\ldots, m\}$

จากวิธีดังกล่าวข้างต้น จะเห็นว่าการหาจุดภายใน $x^* \in \{x \in R^n \mid g_i(x) < 0 \quad ; i = 1,\ldots,m\} \ne \phi$ สามารถหาได้โดยการทำซ้ำของวิธีฟังก์ชันกั้นเขตเป็นจำนวนครั้งมากที่สุด

เท่ากับจำนวนเงื่อนไขบังคับนั่นคือ m ครั้ง

มีข้อเสนอแนะจาก [10] ซึ่งอาจทำให้การทำซ้ำวิธีฟังก์ชันกั้นเขตใช้จำนวนครั้งน้อยลงกว่าวิธีที่ผู้วิจัยใช้โดยมีแนวคิดคือ พยายามหาจุดภายในที่เป็นไปได้โดยวิธีฟังก์ชันกั้นเขตเพียงครั้งเดียว แต่ในทางปฏิบัติอาจเป็นไปไม่ได้ จึงต้องมีการทำซ้ำวิธีฟังก์ชันกั้นเขตเช่นกัน แต่จำนวนครั้งที่ใช้อาจน้อยลง วิธีที่จะเสนอแนะเป็นดังนี้

ที่จุด $\mathbf{x}_0$ แบ่งเงื่อนไขบังคับออกเป็น 2 ส่วนคือ $T = \{t \mid g_t(\mathbf{x}_0) < 0 \quad ; 1 \le t \le m\}$ และ $F = \{f \mid g_f(\mathbf{x}_0) \ge 0 \quad ; 1 \le f \le m\}$ แล้วใช้จุด $\mathbf{x}_0$ แก้ปัญหาต่อไปนี้

$$\left.\begin{array}{ll} \text{minimize} & \sum_{f\in F} g_f(\mathbf{x}) \\ \text{subject to} & g_t(\mathbf{x}) \quad ; \forall t \in T \\ & \mathbf{x} \in R^n \end{array}\right\} \qquad (5.1)$$

แก้ปัญหา (5.1) โดยวิธีฟังก์ชันกั้นเขต ดังนั้นจะได้ปัญหาใหม่เป็น

$$\left.\text{minimize} \qquad U(\mathbf{x},\mu_k) = \sum_{f\in F} g_f(\mathbf{x}) + \mu_k\left[-\sum_{t\in T}\frac{1}{g_t(\mathbf{x})}\right]\right\} \qquad (5.2)$$

สำหรับลำดับที่ลดลงทางเดียว $\{\mu_k\}$

เมื่อแก้ปัญหา (5.2) จะได้จุดใหม่ซึ่งทำให้เงื่อนไขบังคับ $(g_f(\mathbf{x}), f \in F)$ ที่เคยมีค่าเป็นบวกในการทำงานครั้งที่แล้วมีค่าเป็นลบในการทำงานครั้งนี้ได้หนึ่งเงื่อนไขหรือมากกว่าหนึ่งเงื่อนไข ดังนั้นสมาชิกใน F จะถูกถ่ายไปไว้ใน T ได้มากกว่า 1 สมาชิก ให้ F' เป็นเซตใหม่ที่ได้จากการถ่ายสมาชิกบางตัวใน F ให้กับ T และให้ T' เป็นเซตใหม่ที่รับสมาชิกของ F รวมกับสมาชิกเดิมของ T ดังนั้นในการทำงานขั้นต่อไปคือแก้ปัญหาต่อไปนี้

$$\text{minimize} \qquad U(\mathbf{x},\mu_k) = \sum_{f\in F'} g_f(\mathbf{x}) + \mu_k\left[-\sum_{t\in T'}\frac{1}{g_t(\mathbf{x})}\right]$$

โดยใช้จุดที่ได้จากการแก้ปัญหา (5.2) เป็นจุดเริ่มต้น ดำเนินการในทำนองเดียวกันนี้เพื่อกำเนิดจุดใหม่ต่อ ๆ ไป การทำงานจะหยุดลงได้ 2 กรณีคือ เมื่อได้จุดภายในที่เป็นไปได้ที่สอดคล้องกับเงื่อนไขทั้งหมดและในกรณีที่ไม่สามารถหาจุดภายในที่สอดคล้องกับเงื่อนไขบังคับทั้งหมดได้

อนึ่งปัญหาที่อาจจะเกิดขึ้นกับขั้นตอนวิธีที่เสนอแนะนี้คือ ความยุ่งยากในการจัดการกับฟังก์ชันจุดประสงค์ $U(\mathbf{x},\mu_k)$ เพราะในทุก ๆ ครั้งที่แก้ปัญหาด้วยวิธีฟังก์ชั่นกั้นเขตเราต้องใช้เงื่อน

ไขบังคับทั้งหมดมาเป็นส่วนประกอบของฟังก์ชันจุดประสงค์ $U(x,\mu_k)$ ซึ่งลักษณะของฟังก์ชัน $U(x,\mu_k)$ ดังกล่าวจะมีความซับซ้อนมากถ้าหากเงื่อนไขบังคับเป็นฟังก์ชันแบบใด ๆและมีจำนวนเงื่อนไขมาก

เอกสารอ้างอิง

[1] Bazaraa, Mokhtar S. et.al. **Nonlinear Programming : Theory and Algorithms.** 2^{nd} Ed. Singapore : John Wiley & Sons, Inc. , 1979.

[2] Fiacco, A.V. and McCormick, G.P. "Computational Algorithm for the Sequential Unconstrained Minimization Technique for Nonlinear Programming." Management Science., vol.10, 1964. pp. 601 – 617.

[3] Bellmore, M. et.al. "Generalized Penalty Function Concepts in Mathematical Optimization." Operations Research. Vol.18, 1970. pp. 229 – 252.

[4] McCormick, Garth P. "Penalty Function Versus Non-Penalty Function Methods for Constrained Nonlinear Programming Problems." Mathematical Programming., vol.1, February 1971. pp. 217-238.

[5] Lasdon Leon S. "An Efficient Algorithm for Minimizing Barrier and Penalty Functions." Mathematical Programming., vol.2, July 1972. pp. 65 – 106.

[6] Gould, NIM. "On the Convergence of a Sequential Penalty Function Method for Constrained Minimization." Siam Journal on Numerical Analysis., vol.26, no.1, February 1989. pp. 107-128.

[7] Goldfarb, D. et.al. "A Logarithmic Barrier Function Algorithm for Quadratically Constrained Convex Quadratic Programming." Siam Journal on Optimization., vol.1, no.2, May 1991. pp. 252 –267.

[8] Peressini, Anthony L. et.al. **The Mathematics of Nonlinear Programming.** New York : Springer-Verlag New York Inc. 1988.

[9] Nash, Stephen G. and Sofer, Ariela. **Linear and Nonlinear Programming.** Singapore : McGraw-Hill International Editions. 1996.

[10] Fiacco, Anthony V. and McCormick, Garth P. **Nonlinear Programming : Sequential Unconstrained Minimization Techniques.** SIAM edition. Philadelphia. : the Society for Industrial and Applied Mathematics. 1990.

[11] Dennis, J.E.Jr. and Schnabel, Robert B. **Numerical Methods for Unconstrained Optimization and Nonlinear Equations.** SIAM edition. Philadelphia. : the Society for Industrial and Applied Mathematics. 1996.

[12] Byrd, Richard H. et.al. "Global Convergence of a Class of Quasi-Newton Methods on Convex Programming." Siam Jounal Numerical Analysis., vol.24, no.5, October 1987. pp. 1171-1190.

[13] Luenberger, David G. **Linear and Nonlinear Programming.** 2nd Ed. Massachusetts. : Addison-Wesley Publishing Company. 1984

[14] Dennis, S. "RF 21 Feature Nonlinear function minimizer." [Online].Available : http://members.aol.com/rf21exe/solver.htm. 1983.

[15] Pozrikidis, C. **Numerical Computation in Science and Engineering.** Oxford University Press Inc. 1998.

ภาคผนวก ก.

ในส่วนนี้จะให้ผลการคำนวณโดยละเอียดของตัวอย่างปัญหาในบทที่ 4 เพื่อความสมบูรณ์ของผลลัพธ์ โดยแสดงผลการคำนวณของปัญหาที่ 5 , 6 ,7 และ 8 ซึ่งผลการคำนวณของแต่ละปัญหาจะแบ่งเป็น 3 ส่วนย่อยตามขั้นตอนวิธีที่ใช้เขียนโปรแกรมได้แก่ OPTIM 1 , OPTIM 2 และ OPTIM 3 โดยแสดงผลการคำนวณทั้งลูปใหญ่ (ลูป k) และลูปภายใน (ลูป i)

<u>ผลการคำนวณโดยละเอียดจากโปรแกรมสำหรับปัญหาที่ 5</u>

x0 = (10 , 20 , 30 , 40)

g(x0) = (0 , -132 , -229 , 4594 , 85)

-----OPTIM5_2(ครั้งที่ 1)-----

i	x(i+1)	norm(∇U1(x(i+1)))	
0	(8.9986596,20.9973192,29.0045895,41.0013385)	1.9974094	g(1)=-3.99541

k=0 μ(k)=10.000000 x(k+1)=(8.9986596,20.9973192,29.0045895,41.0013385)

μ(k)β(x(k+1))=0.12427726 g(1)=-3.99541: g(2)=-124.0549094 g(3)=-229.0017033

-----OPTIM5_2(ครั้งที่ 2)-----

i	x(i+1)	norm(∇U2(x(i+1)))	
0	(8.8969093,20.9112756,19.5907572,40.9780840)	150.0542887	g(4)=3034.31113
1	(9.1170805,20.5646235,14.0412774,40.6792796)	149.8912070	g(4)=2201.11721
2	(9.2025460,20.3345829,11.4094680,40.5463559)	144.0002078	g(4)=1805.83827
3	(8.9627543,19.7408959,10.5183205,40.6763086)	127.2732268	g(4)=1670.86894
4	(1.5044140,2.5688449,-8.8281198,44.9688913)	2046.9823386	g(4)=-1268.60698

k=0 μ(k)=10.000000 x(k+1)=(1.5044140,2.5688449,-8.8281198,44.9688913)

μ(k)β(x(k+1))=47.20316266 g(4)=-1268.60698: g(2)=-0.2146323 g(3)=-30.7687394

g(1)=-34.8614420

-----OPTIM5_2-----

k=0 μ(k)=10.000000 x(k+1)=(1.5044140,2.5688449,-8.8281198,44.9688913) g(5)=-42.74296:

g(2)=-0.2146323 g(3)=-30.7687394 g(1)=-34.8614420 g(4)=-1268.6069777

-----OPTIM5_3-----

i	x(i+1)	norm(∇M(x(i+1)))
0	(1.0221502,3.2017604,-3.5090005,44.8552613)	139015320.2820677
1	(4.0492812,0.1804321,-2.8977567,43.6157411)	139013102.6198058
2	(0.8717270,-2.8617389,-2.8938176,40.2895829)	138995648.5271797
3	(0.6430711,-3.3117506,-0.7579541,39.6458649)	52642450.2909261
4	(0.5078874,-3.4779071,-0.4321711,39.4416064)	39481837.7279800
5	(0.4571961,-3.5402117,-0.3100027,39.3650070)	34546608.0104469
6	(0.4350291,-3.5674474,-0.2565565,39.3314642)	32387444.9687162
7	(0.4299910,-3.5734914,-0.2440660,39.3231418)	31870828.3548154
8	(2.8570759,1.6674406,-0.7889429,32.1932339)	31432928.3210861
9	(4.0572257,4.2628414,-1.0464082,28.6452962)	30684888.0331507
10	(4.1533452,4.4707552,-1.0670354,28.3610718)	30624956.6196913
11	(4.2013089,4.5745047,-1.0773284,28.2192431)	30595049.5994570
12	(4.2252556,4.6263039,-1.0824673,28.1484313)	30580111.5621248
13	(4.2371109,4.6519542,-1.0850109,28.1133592)	30572741.3541609
14	(2.2219689,0.8413089,-0.6168047,32.7040011)	26516006.2576958
15	(1.5253645,-0.6722936,-0.4160331,34.7290011)	24861200.8527376
16	(1.1994322,-1.3826570,-0.3216684,35.6813044)	24084288.2119015
17	(1.0415701,-1.7267366,-0.2759594,36.1425924)	23707971.2002990
18	(0.9638725,-1.8960881,-0.2534620,36.3696326)	23522752.0594340
19	(0.9542368,-1.9170897,-0.2506721,36.3977875)	23499760.0854104
20	(0.9530530,-1.9196467,-0.2503324,36.4011934)	23496803.2571156
21	(0.9517450,-1.9209501,-0.2501550,36.4014618)	23490371.1318489
22	(1.2174686,1.0054169,-0.6129126,30.2151486)	19671398.9593462
23	(2.3740115,3.5139277,-0.8218876,26.7395395)	17268882.8185755
24	(2.6801235,4.1088053,-0.8679529,25.9779734)	16729241.6591700
25	(2.8327007,4.3999059,-0.8901896,25.6107792)	16467800.9667615
26	(2.9081793,4.5434280,-0.9011253,25.4302383)	16339141.2640941
27	(2.9269182,4.5790383,-0.9038374,25.3854657)	16307229.1357855
28	(2.9362595,4.5967913,-0.9051893,25.3631430)	16325426.7409019
29	(2.5867599,4.5551340,-0.9017019,24.7342115)	14508949.6905716
30	(2.5919096,4.5704417,-0.8821406,24.6881732)	13616384.0887816
31	(2.5940769,4.5712010,-0.8813428,24.6900349)	13593310.3476248
32	(3.4007365,4.4578368,-0.5114404,26.2709679)	3416175.8850040

i	x(i+1)	norm(∇M(x(i+1)))
33	(3.8954485,4.4542238,-0.3631617,27.1591859)	9085.9501287
34	(3.9757607,4.4665375,-0.3546223,27.2876302)	7350.5783238
35	(4.0194519,4.4733877,-0.3501590,27.3573212)	49.4940201
36	(4.0194518,4.4733889,-0.3501604,27.3573198)	0.3560077
37	(4.0194518,4.4733889,-0.3501604,27.3573197)	0.0002091
38	(4.0194518,4.4733889,-0.3501604,27.3573197)	0.0002086
39	(4.0194516,4.4733886,-0.3501604,27.3573200)	0.0025055
40	(4.0194513,4.4733880,-0.3501603,27.3573208)	0.0051587
41	(4.0194504,4.4733860,-0.3501601,27.3573233)	0.0104704
42	(4.0194479,4.4733809,-0.3501596,27.3573301)	0.0184371
43	(4.0194412,4.4733668,-0.3501581,27.3573490)	0.0317128
44	(4.0194237,4.4733296,-0.3501542,27.3573991)	0.0529391
45	(4.0193774,4.4732306,-0.3501438,27.3575327)	0.0873700
46	(4.0192564,4.4729708,-0.3501163,27.3578842)	0.1427238
47	(4.0189406,4.4722918,-0.3500445,27.3588041)	0.2312229
48	(4.0181270,4.4705399,-0.3498591,27.3611795)	0.3697120
49	(4.0160862,4.4661426,-0.3493935,27.3671449)	0.5749756
50	(4.0113020,4.4558287,-0.3483010,27.3811416)	0.8368911
51	(4.0017379,4.4352025,-0.3461157,27.4091396)	1.0432019
52	(3.9883279,4.4062726,-0.3430502,27.4484176)	0.9318081
53	(3.9795858,4.3874032,-0.3410500,27.4740454)	0.4550753
54	(3.9793584,4.3869038,-0.3409966,27.4747313)	0.0876654
55	(3.9807381,4.3898780,-0.3413116,27.4706954)	0.0005478
56	(3.9810263,4.3904995,-0.3413774,27.4698517)	0.0007529
57	(3.9810437,4.3905373,-0.3413814,27.4698004)	0.0000400
58	(3.9810438,4.3905374,-0.3413815,27.4698002)	0.0000006

k=0 μ(k)=0.001000 x(k+1)=(3.9810438,4.3905374,-0.3413815,27.4698002)

μ(k)β(x(k+1))= 0.00051242 (1/μ(k))α(x(k+1))=0.00000000 : h1(x(k+1))=-0.0000000358

h2(x(k+1))=0.0000000321 h3(x(k+1))=0.0000000096: g1(x(k+1))=-8.2207 g2(x(k+1))=-6.2636

g3(x(k+1))=-50.9699 g4(x(k+1))=-6.9753 g5(x(k+1))=-14.6766

i	x(i+1)	norm(∇M(x(i+1)))
0	(3.9810438,4.3905374,-0.3413815,27.4698002)	0.0076139
1	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0072214
2	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0071358
3	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0000001

k=1 μ(k)= 0.000100 x(k+1)= (3.9810438,4.3905374,-0.3413814,27.4698002)

μ(k)β(x(k+1))= 0.00005124 (1/μ(k))α(x(k+1))= 0.00000000: h1(x(k+1))= -0.0000000004

h2(x(k+1))= 0.0000000003 h3(x(k+1))= 0.0000000001: g1(x(k+1))=-8.2207

g2(x(k+1))=-6.2636 g3(x(k+1))=-50.9699 g4(x(k+1))=-6.9753 g5(x(k+1))=-14.6766

i	x(i+1)	norm(∇M(x(i+1)))
0	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0003843
1	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0002383
2	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0001930
3	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0000002

k=2 μ(k)=0.000010 x(k+1)=(3.9810438,4.3905374,-0.3413814,27.4698002)

μ(k)β(x(k+1))= 0.00000512 (1/μ(k))α(x(k+1))= 0.00000000 : h1(x(k+1))= -0.0000000000

h2(x(k+1))= 0.0000000000 h3(x(k+1))= 0.0000000000: g1(x(k+1))=-8.2207 g2(x(k+1))=-6.2636

g3(x(k+1))=-50.9699 g4(x(k+1))=-6.9753 g5(x(k+1))=-14.6766

i	x(i+1)	norm(∇M(x(i+1)))
0	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0000678
1	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0000540
2	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0000491
3	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0000418
4	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0000109
5	(3.9810438,4.3905374,-0.3413814,27.4698002)	0.0000028

k=3 μ(k)=0.000001 x(k+1)=(3.9810438,4.3905374,-0.3413814,27.4698002)

μ(k)β(x(k+1))= 0.00000051 (1/μ(k))α(x(k+1))= 0.00000000 :h1(x(k+1))= -0.0000000000 h2(x(k+1))= 0.0000000000 h3(x(k+1))= 0.0000000000: g1(x(k+1))= -8.2207 g2(x(k+1))=-6.2636

g3(x(k+1))=-50.9699 g4(x(k+1))=-6.9753 g5(x(k+1))= -14.6766

$\mu(3)\beta(x(4))$=5.1242e-007

$(1/\mu(3))\alpha(x(4))$=2.5244e-021

Solution is (3.9810438,4.3905374,-0.3413814,27.4698002)

<u>ผลการคำนวณโดยละเอียดจากโปรแกรมสำหรับปัญหาที่ 6</u>

x0 =(-3.5000e-001 , 6.9000e+000 , 4.8000e+000)

g(x0) =(1.3254e+002 , 6.4855e+001 , 4.2150e+001 , 2.3754e+021)

-----OPTIM6_1-----

k	x(k+1)	norm(∇g1(x(k+1)))	
0	(-2.6750000,3.1250000,1.0500000)	26.375367	g(1)=56.50000
1	(2.3786942,-10.3388346,-19.2705666)	20.572390	g(1)=-223.11828

----OPTIM6_2(ครั้งที่ 1)-----

i	x(i+1)	norm(∇U1(x(i+1)))	
0	(-2.3789061,1.0013981,-22.2720732)	11.25230	g(2)=-121.31405

k=0 μ(k)=10.000000 x(k+1)=(-2.3789061,1.0013981,-22.2720732) $\mu(k)\beta(x(k+1))$=0.03309449

g(2)=-121.31405 g(1)=-302.1650905

-----OPTIM6_2-----

k=0 μ(k)= 10.000000 x(k+1)= (-2.3789061,1.0013981,-22.2720732) g(3)=-93.97863:

g(1)=-302.1650905 g(2)=-121.3140488

-----OPTIM6_2(ครั้งที่ 2)-----

i	x(i+1)	norm(∇U2(x(i+1)))	
0	(-7.1945793,1.0013977,-19.8642370)	2119.89382	g(4)=372.37899
1	(-7.5632002,0.1502729,-20.6014766)	2112.51514	g(4)=341.74441
2	(-7.8544966,0.1870187,-21.1840673)	1253.52934	g(4)=324.32866
3	(-7.8371257,0.2430128,-21.1493173)	1954.07189	g(4)=325.48948
4	(-7.8734947,0.3868557,-21.2220204)	1933.31917	g(4)=323.77807
5	(-7.8838058,0.4467424,-21.2425515)	1940.16353	g(4)=323.37235
6	(-7.8840462,0.5155273,-21.2427161)	1981.75898	g(4)=323.48337
7	(-7.8743875,0.5591888,-21.2229550)	2018.25493	g(4)=323.96631
8	(-7.8652653,0.6016818,-21.2041923)	2039.11262	g(4)=324.38053

i	x(i+1)	norm(∇U2(x(i+1)))	
9	(-7.8556978,0.6335175,-21.1844875)	2052.27595	g(4)=324.71669
10	(-7.8436165,0.6585319,-21.1595636)	2062.62339	g(4)=324.98469
11	(-7.8230581,0.6839257,-21.1170142)	2072.69079	g(4)=325.16581
12	(-7.7834193,0.7098397,-21.0346544)	2080.18578	g(4)=325.00435
13	(-7.6967166,0.7319282,-20.8539081)	2077.53241	g(4)=323.76256
14	(-7.4737059,0.7404077,-20.3880674)	2044.88691	g(4)=319.38185
15	(-6.7825111,0.7217394,-18.9429081)	1922.93271	g(4)=306.38636
16	(-5.1942980,0.6716298,-15.6204430)	1660.45257	g(4)=287.78937
17	(-4.3041759,0.6465516,-13.7564181)	1526.02222	g(4)=283.29578
18	(-3.9788782,0.6388451,-13.0728590)	1476.13193	g(4)=281.40560
19	(-3.3824829,0.6260241,-11.8124418)	1378.77899	g(4)=275.54239
20	(-2.6661147,0.6107654,-10.2795767)	1246.19810	g(4)=262.07722
21	(-1.8274670,0.5894745,-8.4260635)	1042.80605	g(4)=226.50561
22	(-1.3839380,0.5618084,-7.2355626)	766.58889	g(4)=138.40819
23	(-2.8139522,0.5264390,-9.5273287)	428.99716	g(4)=-81.72706

k=0 μ(k)=10.000000 x(k+1)=(-2.8139522,0.5264390,-9.5273287) μ(k)β(x(k+1))=0.56482136

g(4)=-81.72706 g(1)=-99.1370752 g(2)=-40.1030580 g(3)=-46.5997591

--- OPTIM6_3-------

i	x(i+1)	norm(∇M(x(i+1)))
0	(-4.6657194,1.3226666,-10.7030793)	114497013.4552791
1	(-5.3988597,1.2558157,-10.5997150)	296142862.5712274
2	(-5.3972356,1.9762418,-10.1156415)	296185228.2255320
3	(-5.1170959,1.7080850,-10.4418945)	84627095.4805875
4	(-5.0157200,1.6536206,-10.4504485)	22910584.4066042
5	(-4.9790605,1.6636262,-10.4246228)	2478301.0004360
6	(-4.9745588,1.6951439,-10.4006183)	79194.8339249
7	(-4.9741140,1.7276876,-10.3782504)	120592.2576183
8	(-4.9727099,1.9561261,-10.2220201)	550433.3687278
9	(-4.9731044,2.0107478,-10.1847719)	272179.0563911
10	(-4.9737879,1.9894961,-10.1991381)	30899.6904207
11	(-4.9737769,1.9941574,-10.1955974)	13063.8186253
12	(-4.9738676,1.9989397,-10.1912175)	23846.8759163
13	(-4.9743166,2.0071663,-10.1815438)	67366.6949363

i	x(i+1)	norm(∇M(x(i+1)))
14	(-4.9756926,2.0183359,-10.1628150)	134118.7828529
15	(-4.9766321,2.0216218,-10.1533998)	157232.9456122
16	(-4.9773203,2.0227471,-10.1475176)	166061.8669613
17	(-4.9776009,2.0229268,-10.1453417)	167762.2650320
18	(-5.0002208,2.0117863,-9.9882714)	124192.5857871
19	(-5.0009610,2.0098889,-9.9850814)	56909.2441277
20	(-5.0005167,2.0024812,-9.9943864)	2217.5939024
21	(-5.0000489,2.0001762,-9.9995363)	1620.8981688
22	(-5.0000022,1.9999983,-9.9999872)	139.5102501
23	(-5.0000000,1.9999994,-10.0000002)	3.9637143
24	(-5.0000000,2.0000000,-10.0000000)	0.0051341
25	(-5.0000000,2.0000000,-10.0000000)	0.0013830
26	(-5.0000000,2.0000000,-10.0000000)	0.0000123

k=0 μ(k)=0.001000 x(k+1)=(-5.0000000,2.0000000,-10.0000000)

μ(k)β(x(k+1))=0.00022113 (1/μ(k))α(x(k+1))=0.00000000 : h1(x(k+1))=-0.0000001

h2(x(k+1))=-0.0000000 h3(x(k+1))=-0.0000001: g1(x(k+1))=-29.0000002 g2(x(k+1))=-8.0000001

g3(x(k+1))=-55.0000001 g4(x(k+1))=-23.0092665

μ(0)β(x(1))=2.2113e-004

(1/μ(0))α(x(1))=1.6966e-011

Solution is (-5.0000000,2.0000000,-10.0000000)

<u>ผลการคำนวณโดยละเอียดจากโปรแกรมสำหรับปัญหาที่ 7</u>

x0 =(-3.5000e-001 , 6.9000e+000 , 4.8000e+000)

g(x0) =(1.3254e+002 , 6.4855e+001 , 4.2150e+001 , 2.3754e+021)

-----OPTIM7_1-----

k	x(k+1)	norm(∇g1(x(k+1)))	
0	(-2.6750000,3.1250000,1.0500000)	26.375367296	g(1)=56.50000
1	(2.3786942,-10.3388346,-19.2705666)	20.572390442	g(1)=-223.11828

----OPTIM7_2(ครั้งที่ 1)-----

i	x(i+1)	norm(∇U1(x(i+1)))	
0	(-2.3789061,1.0013981,-22.2720732)	11.2523034	g(2)=-121.31405

k=0 μ(k)=10.000000 x(k+1)=(-2.3789061,1.0013981,-22.2720732) μ(k)β(x(k+1))=0.03309449

g(2)=-121.31405: g(1)=-302.1650905

-----OPTIM7_2-----

k=0 μ(k)=10.000000 x(k+1)=(-2.3789061,1.0013981,-22.2720732) g(3)=-93.97863:

g(1)=-302.1650905 g(2)=-121.3140488

-----OPTIM7_2(ครั้งที่ 2)-----

i	x(i+1)	norm(∇U2(x(i+1)))	
0	(-7.1945793,1.0013977,-19.8642370)	2119.8938227	g(4)=372.37899
1	(-7.5632002,0.1502729,-20.6014766)	2112.5151351	g(4)=341.74441
2	(-7.8544966,0.1870187,-21.1840673)	1253.5293442	g(4)=324.32866
3	(-7.8371257,0.2430128,-21.1493173)	1954.0718879	g(4)=325.48948
4	(-7.8734947,0.3868557,-21.2220204)	1933.3191652	g(4)=323.77807
5	(-7.8838058,0.4467424,-21.2425515)	1940.1635288	g(4)=323.37235
6	(-7.8840462,0.5155273,-21.2427161)	1981.7589750	g(4)=323.48337
7	(-7.8743875,0.5591888,-21.2229550)	2018.2549268	g(4)=323.96631
8	(-7.8652653,0.6016818,-21.2041923)	2039.1126191	g(4)=324.38053
9	(-7.8556978,0.6335175,-21.1844875)	2052.2759515	g(4)=324.71669
10	(-7.8436165,0.6585319,-21.1595636)	2062.6233945	g(4)=324.98469
11	(-7.8230581,0.6839257,-21.1170142)	2072.6907851	g(4)=325.16581
12	(-7.7834193,0.7098397,-21.0346544)	2080.1857789	g(4)=325.00435
13	(-7.6967166,0.7319282,-20.8539081)	2077.5324124	g(4)=323.76256
14	(-7.4737059,0.,-107404077,-20.3880674)	2044.8869086	g(4)=319.38185
15	(-6.7825111,0.7217394,-18.9429081)	1922.9327075	g(4)=306.38636
16	(-5.1942980,0.6716298,-15.6204430)	1660.4525665	g(4)=287.78937
17	(-4.3041759,0.6465516,-13.7564181)	1526.0222174	g(4)=283.29578
18	(-3.9788782,0.6388451,-13.0728590)	1476.1319261	g(4)=281.40560
19	(-3.3824829,0.6260241,-11.8124418)	1378.7789946	g(4)=275.54239
20	(-2.6661147,0.6107654,-10.2795767)	1246.1980993	g(4)=262.07722
21	(-1.8274670,0.5894745,-8.4260635)	1042.8060486	g(4)=226.50561
22	(-1.3839380,0.5618084,-7.2355626)	766.5888925	g(4)=138.40819
23	(-2.8139522,0.5264390,-9.5273287)	428.9971639	g(4)=-81.72706

k=0 $\mu(k)$=10.000000 x(k+1)=(-2.8139522,0.5264390,-9.5273287) $\mu(k)\beta$(x(k+1))= 0.56482136

g(4)=-81.72706: g(1)=-99.1370752 g(2)=-40.1030580 g(3)=-46.5997591

-----OPTIM7_3-----

i	x(i+1)	norm(∇M(x(i+1)))
0	(-4.3434138,0.4961605,-10.5049568)	148019083.1923387
1	(-5.2880901,0.4605867,-11.5771350)	328133736.8425017
2	(-5.2849350,1.9327526,-11.6319881)	312540070.2020298
3	(-4.9721418,1.6835600,-11.1997354)	68626556.3532787
4	(-4.8791477,1.6047873,-11.0992403)	18064704.6463685
5	(-4.8449614,1.5244436,-11.0610943)	2072533.2499087
6	(-4.8447725,1.5206026,-11.0606181)	2017060.3872748
7	(-4.8593062,1.9866821,-11.1116768)	2946015.8880921
8	(-4.8531113,1.8801680,-11.0975712)	1740010.8433487
9	(-4.8515345,1.9147163,-11.0994835)	469619.9529798
10	(-4.8507701,1.9250203,-11.1003714)	37711.2913184
11	(-4.8507646,1.9249287,-11.1003898)	37292.0702179
12	(-4.8511141,1.9283203,-11.0987280)	57103.0046439
13	(-4.8519667,1.9333645,-11.0940159)	92050.4044731
14	(-4.8543982,1.9427779,-11.0795577)	154862.1726590
15	(-4.8606586,1.9594219,-11.0407219)	245974.2407743
16	(-4.8767748,1.9903310,-10.9378387)	327182.6616245
17	(-4.8952446,2.0164928,-10.8164166)	286710.1076659
18	(-4.9032468,2.0247846,-10.7620066)	256466.1699063
19	(-4.9127513,2.0315987,-10.6954468)	219838.6890520
20	(-4.9138620,2.0320757,-10.6874659)	216262.6670056
21	(-4.9149443,2.0323147,-10.6795582)	214172.5221585
22	(-4.9154232,2.0322874,-10.6759991)	215332.4150777
23	(-4.9655843,1.9953169,-10.2872086)	547246.2076708
24	(-4.9743984,1.9922307,-10.2114341)	302274.0868453
25	(-4.9966906,1.9893630,-10.0272821)	252981.4780713
26	(-4.9989165,1.9929842,-10.0042953)	111777.0854411
27	(-5.0017545,1.9951823,-9.9826942)	41017.8853454
28	(-5.0014320,1.9982932,-9.9871355)	12622.3599868
29	(-5.0002082,2.0000011,-9.9982683)	304.3556735
30	(-5.0000156,2.0000119,-9.9998786)	226.4481670

i	x(i+1)	norm(∇M(x(i+1)))
31	(-5.0000000,2.0000004,-10.0000006)	5.7439306
32	(-5.0000000,2.0000000,-10.0000000)	0.0161418
33	(-5.0000000,2.0000000,-10.0000000)	0.0011730

k=0 μ(k)=0.001000 x(k+1)=(-5.0000000,2.0000000,-10.0000000)

μ(k)β(x(k+1))= 0.0002211253570006 (1/μ(k))α(x(k+1))=0.0000000000011187 :

h1(x(k+1))=-0.0000000314 h2(x(k+1))=0.0000000032 h3(x(k+1))=0.0000000108

h4(x(k+1))=0.0000000019: g1(x(k+1))=-29.0000001 g2(x(k+1))=-8.0000000 g3(x(k+1))=-55.0000000

g4(x(k+1))=-23.0092507

i	x(i+1)	norm(∇M(x(i+1)))
0	(-5.0009452,2.0012492,-9.9999782)	4444950.6576469
1	(-5.0011052,2.0052234,-9.9995811)	4010067.9625165
2	(-4.4093966,-0.3726936,-9.1784793)	1926632392.2420216
3	(-4.0437894,-0.6335921,-11.4344385)	1356881192.9757674
4	(-4.5880412,1.8186959,-12.7912108)	100390380.9318341
5	(-4.5641504,1.8261777,-12.6962621)	29906060.4578502
6	(-4.5667637,1.8182235,-12.7276184)	2509156.3414541
7	(-4.5666030,1.8180578,-12.7282772)	2564055.5871589
8	(-4.6646097,1.9155139,-12.2959264)	30966012.3114756
9	(-4.6715860,1.9075585,-12.2404993)	22353791.4728749
10	(-4.8244492,1.9201402,-11.3623124)	52595132.9977653
11	(-4.8104653,1.8833994,-11.3864792)	18038208.9497016
12	(-4.8433528,1.8653031,-11.1409330)	5975674.6108056
13	(-4.8633524,1.9035218,-11.0063269)	830941.6165905
14	(-4.9209524,1.9835283,-10.6344419)	9972565.8089800
15	(-4.9497303,2.0247374,-10.4373086)	13760085.9282177
16	(-4.9509099,2.0268393,-10.4193486)	7862735.2048709
17	(-4.9716418,2.0187884,-10.2416697)	1362450.5895316
18	(-4.9913176,2.0093436,-10.0798626)	2453585.9498874
19	(-4.9934814,2.0021244,-10.0563094)	805527.1677197
20	(-5.0010173,1.9998280,-9.9925887)	695299.9013524
21	(-4.9998771,1.9997913,-10.0008381)	60831.6933997
22	(-4.9999878,1.9999837,-10.0000992)	2425.5303088

i	x(i+1)	norm(∇M(x(i+1)))
23	(-4.9999994,1.9999994,-10.0000056)	501.8604370
24	(-5.0000000,2.0000000,-10.0000000)	14.5993538
25	(-5.0000000,2.0000000,-10.0000000)	0.1535622
26	(-5.0000000,2.0000000,-10.0000000)	0.0004499

k=1 μ(k)=0.000100 x(k+1)= (-5.0000000,2.0000000,-10.0000000)

μ(k)β(x(k+1))= 0.0000221125359181 (1/μ(k))α(x(k+1))=0.0000000000000003 :

h1(x(k+1))=-0.0000000002 h2(x(k+1))=0.0000000000 h3(x(k+1))=0.0000000001

h4(x(k+1))=0.0000000000: g1(x(k+1))=-29.0000000 g2(x(k+1))=-8.0000000 g3(x(k+1))=-55.0000000

g4(x(k+1))=-23.0092498

μ(1)β(x2))=2.2113e-005

(1/μ(k))α(x(k+1))=3.4294e-016

Solution is (-5.0000000,2.0000000,-10.0000000)

<u>ผลการคำนวณโดยละเอียดจากโปรแกรมสำหรับปัญหาที่ 8</u>

x0 = (-2 , 5 , 0 , 10)

g (x0)= (4.6400e+002 , 1.3598e+003 , -2.6700e+002 , 2.5341e+002 , -8.1582e+000)

-----OPTIM8_2(ครั้งที่ 1)-----

i	x(i+1)	norm(∇U1(x(i+1)))	
0	(0.2578991,3.7409776,-0.3215311,10.0093818)	15.728847733	g(1)=418.25894
1	(0.1616605,-0.5219100,1.0881545,10.0310358)	9.330977327	g(1)=385.93071
2	(2.3052630,-0.8520259,7.1273987,10.1405725)	20.920204514	g(1)=323.36765
3	(24.8959491,-4.4671829,73.6561593,11.3908180)	61424.384277664	g(1)=376936.70806
4	(13.6043711,-2.4713303,40.5680032,10.8113858)	9905.497295443	g(1)=32272.68931
5	(11.4659435,-1.8449538,35.8226466,10.7545972)	5882.569264970	g(1)=15884.31225
6	(8.3840365,-1.3988608,30.2253360,10.2768298)	2232.778002808	g(1)=4197.03510
7	(6.5460363,-1.1792577,27.8221131,9.9777364)	1007.240439843	g(1)=1390.80571
8	(5.0866879,-0.7834392,26.6530795,9.8638777)	416.638914975	g(1)=418.10222
9	(4.1027397,-0.7157828,26.5395602,9.9513892)	167.371829460	g(1)=142.78756
10	(3.4641464,-0.6236391,27.0432080,10.1061615)	56.890478206	g(1)=65.07256
11	(3.0963429,-0.5542528,27.9338726,10.2431337)	15.967351220	g(1)=40.37309
12	(2.8441498,-0.5416948,29.3644051,10.3685861)	32.722454092	g(1)=22.48495
13	(2.5619837,-0.5846846,32.7256545,10.6215132)	68.875909215	'g(1)=-10.02861

k=0 μ(k)= 10.000000 x(k+1)=(2.5619837,-0.5846846,32.7256545,10.6215132) g(1)=-10.02861 g(3)=-174.8089089 g(5)=-7.8832850

-----OPTIM8_2(ครั้งที่ 2)-----

i	x(i+1)	norm(∇U2(x(i+1)))	
0	(2.5051749,-1.5837993,32.7661599,10.6264195)	143.814014806	g(2)=1238.01036
1	(2.5678381,-2.2815801,32.8147507,10.6299041)	189.212075595	g(2)=1195.39387
2	(2.4642177,-2.8675317,33.7516016,10.6396372)	3532.522398941	g(2)=1153.53805
3	(2.4602649,-3.3892918,34.2500352,10.6521132)	660392.92658320	g(2)=1118.16262
4	(2.4660620,-3.5745212,34.4138409,10.6778953)	732054.90628447	g(2)=1105.64641
5	(2.4301120,-3.5836443,34.6594799,10.7142342)	426230.00454713	g(2)=1102.96564
6	(2.5664834,-3.7046814,33.9853286,10.9634313)	9734.203207621	g(2)=1101.77760
7	(2.5626115,-3.7031761,34.0078478,11.2499896)	7463.684903118	g(2)=1101.61927
8	(2.5518844,-3.7013767,34.0730000,12.3169024)	4314.092970514	g(2)=1100.98085
9	(2.5401386,-3.7018196,34.1471414,13.7269087)	2886.924213686	g(2)=1100.08998
10	(2.5219030,-3.7049927,34.2651284,16.1224331)	1963.584539733	g(2)=1098.52758
11	(2.4963774,-3.7119902,34.4332470,19.6534423)	1542.259811207	g(2)=1096.17770
12	(2.4617764,-3.7241076,34.6642168,24.6043571)	1644.149092336	g(2)=1092.84284
13	(2.4420290,-3.7337431,34.7992710,27.6163005)	2082.903470316	g(2)=1090.74965
14	(2.4448644,-3.7357062,34.7838847,27.4282135)	1984.878998101	g(2)=1090.78828
15	(2.4655543,-3.7853524,34.7138778,28.6261236)	1504.891919401	g(2)=1088.32795
16	(2.5091319,-3.9680857,34.6599792,36.8338522)	1422.405964825	g(2)=1077.20057
17	(2.5384946,-4.1513752,34.6957399,46.7562297)	2644.323732888	g(2)=1065.21390
18	(2.5291945,-4.2809407,34.9093580,57.4053216)	4151.943434983	g(2)=1054.65369
19	(2.4995574,-4.3139564,35.1347796,63.4934151)	4508.573123602	g(2)=1049.99770
20	(2.4926456,-4.2831399,35.1415519,62.2917621)	3661.382157885	g(2)=1051.84697
21	(2.4867210,-4.2906927,35.1883958,63.9284933)	3909.832781105	g(2)=1050.81689
22	(2.4798190,-4.3038054,35.2495934,66.9458580)	4138.338472551	g(2)=1049.22492
23	(2.4641310,-4.3536864,35.4181222,78.2069616)	4704.732027883	g(2)=1043.85620
24	(2.4485669,-4.4438627,35.6470389,99.4398896)	4867.954519916	g(2)=1034.94583
25	(2.4352910,-4.6401371,36.0234281,147.1220006)	3899.728750747	g(2)=1016.95548
26	(2.4349479,-4.7722437,36.2262099,179.9805424)	3053.884891344	g(2)=1005.45597
27	(2.4402638,-4.7475093,36.1592312,175.2202266)	2171.547690753	g(2)=1007.87897
28	(2.4593161,-4.7603563,36.0658448,179.1716573)	1060.868418754	g(2)=1007.99942
29	(2.4739731,-4.7861763,36.0058283,179.1666592)	987.924545199	g(2)=1007.10277

i	x(i+1)	norm(∇U2(x(i+1)))	
30	(2.5267804,-4.8905484,35.8029630,179.4791478)	1381.944567718	g(2)=1003.09747
31	(2.5158147,-4.8654456,35.8424190,179.5151026)	1109.676006031	g(2)=1004.15420
32	(2.5182630,-4.8683304,35.8320874,179.5439995)	1073.163467844	g(2)=1004.09966
33	(2.5542881,-4.9176951,35.6926790,179.3179342)	754.950228253	g(2)=1002.79011
34	(2.5849990,-4.9690454,35.5943628,179.3390694)	608.454745025	g(2)=1000.92409
35	(2.6387746,-5.0731344,35.4593076,179.4191693)	466.690224853	g(2)=996.48855
36	(2.7102941,-5.2343897,35.3492646,179.4609659)	358.825043821	g(2)=988.63027
37	(2.8192438,-5.5250801,35.3297436,179.2234962)	267.959067007	g(2)=972.63601
38	(2.9827264,-6.0606294,35.6442985,177.8956553)	185.751766137	g(2)=939.13302
39	(3.2390867,-7.1443817,36.9996229,173.0317586)	221.505237058	g(2)=860.19254
40	(3.4058435,-8.1132508,38.8403703,166.0841763)	45.013586726	g(2)=775.89368
41	(3.4678409,-8.4339251,39.3759028,164.2259460)	182.171372489	g(2)=748.18103
42	(3.5140913,-8.7228287,39.9448584,162.3885470)	772.675253084	g(2)=721.17793
43	(3.5753088,-9.1591475,40.8784589,159.5184891)	3152.134853126	g(2)=678.00191
44	(3.6228336,-9.5498345,41.7660095,157.0919071)	3291.284556486	g(2)=637.19859
45	(3.6262358,-9.6479922,42.0300122,157.0834815)	308.805484631	g(2)=625.83917
46	(3.6755997,-10.1544739,43.2045549,155.6065544)	775.343259811	g(2)=570.22012
47	(3.6975363,-10.3902877,43.7545668,155.0217041)	3953.961812371	g(2)=543.60578
48	(3.7067057,-10.5810764,44.2272550,155.4532271)	1586.047309596	g(2)=521.02649
49	(3.7213435,-10.8125600,44.7899461,155.5496797)	3781.089854329	g(2)=493.54833
50	(3.7313760,-11.0107032,45.2836319,155.6889568)	4745.393852731	g(2)=469.39512
51	(3.7506338,-11.4837634,46.4959230,155.6749995)	3212.295853110	g(2)=409.63800
52	(3.7613776,-11.7267350,47.1408203,154.1751525)	1895.360829310	g(2)=377.77969
53	(3.7743815,-11.9927308,47.8368462,152.6373707)	10550.518594620	g(2)=342.57402
54	(3.7893122,-12.2844342,48.6374002,148.9125324)	6749.180340708	g(2)=302.45624
55	(3.7918147,-12.2688916,48.6244311,146.8745786)	904.319267351	g(2)=304.02649
56	(3.8156867,-12.6419873,49.6628681,140.6705640)	1411.751001112	g(2)=251.40361
57	(3.8301329,-12.8645517,50.2846744,136.9143198)	4360.857770513	g(2)=219.35590
58	(3.8517350,-13.1592180,51.1341017,130.2264478)	15302.744748105	g(2)=175.65548
59	(3.8940124,-13.6739934,52.6734457,114.9419174)	24712.777581204	g(2)=96.07097
60	(3.8981789,-13.7178029,52.8242661,112.5282266)	2736.032138196	g(2)=88.69604
61	(3.9129452,-13.9189044,53.4365353,106.5168671)	3248.213315578	g(2)=56.48143
62	(3.9292842,-14.1444501,54.1240729,99.8657941)	9157.187703439	g(2)=19.84023
63	(3.9351457,-14.2375597,54.4143430,97.2949911)	7856.944494900	g(2)=4.36011
64	(3.9551480,-14.5462338,55.3851483,89.0095256)	6858.567623744	g(2)=-47.84581

k=0 $\mu(k)$ =10.000000 x(k+1)=(3.9551480,-14.5462338,55.3851483,89.0095256) g(2)=-47.84581 g(3)=-54.9723489 g(5)=-87.2126577 g(1)=-0.3064651

-----OPTIM8_2(ครั้งที่ 3)-----

i	x(i+1)	norm(∇U3(x(i+1)))	
0	(3.8125387,-14.1681369,55.5075235,89.0094645)	2.235239128	g(4)=128.50194
1	(3.5474669,-14.8668517,57.3593819,88.0083968)	8.933890430	g(4)=125.64902
2	(4.5609484,-16.7226597,64.2813372,84.2188934)	427.321709833	g(4)=114.93756
3	(4.5005089,-16.6880193,64.1040633,84.2363996)	25.171496598	g(4)=115.13234
4	(4.4821424,-18.0466860,68.2889605,80.6311629)	239.982124843	g(4)=107.34220
5	(4.4814239,-17.8713258,67.7469706,80.6924275)	93.717369678	g(4)=107.94546
6	(4.4632665,-17.9873228,68.0947422,78.1124989)	71.407018288	g(4)=105.01776
7	(4.3272891,-18.9415980,70.9744580,56.3695370)	3070.400090101	g(4)=80.39508
8	(4.3252390,-18.8753284,70.7747449,54.9416343)	353.922205745	g(4)=79.16689
9	(4.2572898,-19.0445655,71.2895608,39.3297407)	795.562931312	g(4)=63.04018
10	(4.2273325,-19.1655054,71.7138804,23.1033839)	1622.034067506	g(4)=46.38950
11	(4.2128999,-19.2715286,72.1216497,5.5726507)	1287.392646121	g(4)=28.45100
12	(4.2228109,-19.2670119,72.1311662,4.2349135)	514.586162085	g(4)=27.10375
13	(4.2274486,-19.2556571,72.0968799,5.4502559)	410.386715245	g(4)=28.35338
14	(4.2346548,-19.2545074,72.1098019,5.5644045)	289.309481827	g(4)=28.45460
15	(4.2498552,-19.2614260,72.1729618,5.4815812)	167.555029594	g(4)=28.30862
16	(4.2686908,-19.2780571,72.2817800,5.2208596)	105.278776368	g(4)=27.93908
17	(4.2963389,-19.3078175,72.4612910,4.9288884)	67.056725593	g(4)=27.46760
18	(4.3337853,-19.3503415,72.7121330,4.8243248)	46.375609030	g(4)=27.11219
19	(4.3855695,-19.4102705,73.0630203,4.9261007)	36.411560011	g(4)=26.86308
20	(4.4379815,-19.4721052,73.4230877,5.1277932)	35.668096340	g(4)=26.70471
21	(4.4651406,-19.5050817,73.6140852,5.3061638)	38.056158248	g(4)=26.69208
22	(4.4717754,-19.5147040,73.6678718,5.4009218)	38.601039276	g(4)=26.73305
23	(4.4832503,-19.5372433,73.7870093,5.6357110)	38.529768570	g(4)=26.84870
24	(4.4959818,-19.5751718,73.9758970,5.9721944)	36.250085626	g(4)=26.99630
25	(4.5135090,-19.6536800,74.3507212,6.5073365)	29.985315992	g(4)=27.15662
26	(4.5378020,-19.7940107,75.0067035,7.1881437)	21.680037886	g(4)=27.18144
27	(4.6024403,-20.1505672,76.6750543,8.5148773)	13.904905391	g(4)=26.83982
28	(4.5868894,-20.0831559,76.3541529,8.2804226)	14.218368575	g(4)=26.92627
29	(4.5754624,-20.0408920,76.1502321,8.1463757)	14.238174867	g(4)=26.99614

i	x(i+1)	norm(∇U3(x(i+1)))	
30	(4.5657673,-20.0081037,75.9906363,8.0384108)	14.187295262	g(4)=27.04777
31	(4.5564409,-19.9781896,75.8442070,7.9316629)	14.123174774	g(4)=27.08746
32	(4.5471862,-19.9495185,75.7033196,7.8206288)	14.068369494	g(4)=27.11731
33	(4.5377076,-19.9208381,75.5620170,7.7016000)	14.036304473	g(4)=27.13958
34	(4.5274748,-19.8903261,75.4114598,7.5681884)	14.042946776	g(4)=27.15673
35	(4.5152120,-19.8539502,75.2319127,7.4037702)	14.127314234	g(4)=27.17186
36	(4.4955403,-19.7951599,74.9421158,7.1351681)	14.490571123	g(4)=27.19305
37	(4.8029367,-20.7369373,79.5684017,11.3437009)	77.505345888	g(4)=26.77530
38	(4.7471288,-20.5895869,78.8314904,10.6980008)	31.334417937	g(4)=26.86651
39	(4.7565269,-20.7013687,79.3319899,10.9808443)	25.418503171	g(4)=26.64885
40	(4.8072027,-21.4928430,82.8325490,11.1420657)	9.049690363	g(4)=23.30952
41	(4.8787810,-22.9796871,89.3492969,9.2736962)	4.394854815	g(4)=14.92440
42	(4.9892179,-25.0318589,98.3671914,7.0939028)	6.013695338	g(4)=3.72671
43	(5.0636544,-26.5116631,104.8598613,5.3728284)	47.503794661	g(4)=-4.48703

k=0 μ(k)=10.000000 x(k+1)=(5.0636544,-26.5116631,104.8598613,5.3728284) g(4)=-4.48703

g(3)=-13.0072593 g(5)=-3.0815180 g(1)=-5.5538725 g(2)=-3478.1870981

-----OPTIM8_3-----

i	x(i+1)	norm(∇M(x(i+1)))
0	(4.6418803,-23.1937622,100.2920148,5.2321331)	380430894.739827870
1	(4.4822286,-18.5634282,101.4995288,4.9802851)	354143993.491724910
2	(2.5274953,-18.6317684,101.6173310,4.9774812)	356268603.412875530
3	(2.5226425,-18.7109799,101.6064015,3.4567268)	355675970.075525580
4	(2.5286678,-18.6945137,101.5870100,2.3630344)	353603019.021899280
5	(2.5289138,-18.6934218,101.5862557,2.3207924)	353517094.992949190
6	(2.5289210,-18.6933918,101.5862316,2.3194109)	353514378.730866130
7	(2.5289110,-18.6935186,101.5862327,2.3188911)	353515640.467124880
8	(2.5959131,-17.9459116,101.5225198,1.7407865)	335476623.608389560
9	(2.6124176,-17.6248870,101.4534671,1.5150002)	325477951.000442450
10	(2.6106379,-17.4528417,101.4209125,1.3872967)	319727160.585359990
11	(2.6111556,-17.3888952,101.4030819,1.3404288)	317046567.516669630
12	(2.6100493,-17.3754063,101.3984853,1.3281416)	316446757.540987730
13	(2.6097875,-17.3735590,101.3979334,1.3261432)	316370202.273522140
14	(2.6094471,-17.3716675,101.3973867,1.3242273)	316289733.422677220

i	x(i+1)	norm(∇M(x(i+1)))
15	(1.1945374,-9.3821049,99.1692560,2.4966795)	52221995.233013816
16	(1.0345832,-8.1738737,99.1192055,2.3223845)	27720179.620124482
17	(0.9906523,-7.8485029,99.0942274,2.2904819)	20730590.813937906
18	(0.9818059,-7.7890484,99.0890822,2.2848538)	19524553.464294646
19	(0.9796630,-7.7738158,99.0878722,2.2833810)	19209718.158800893
20	(0.9787295,-7.7665867,99.0872722,2.2827820)	19057965.530987307
21	(0.9783137,-7.7630461,99.0869727,2.2825274)	18983104.761995368
22	(0.9781252,-7.7612945,99.0868226,2.2824163)	18945731.366757069
23	(0.9780152,-7.7604375,99.0867445,2.2823520)	18919168.892227124
24	(0.6853015,-7.4808401,99.0186477,2.1002228)	11603844.547965335
25	(0.5397698,-7.3508613,99.0120095,1.9736745)	8908194.440309448
26	(0.5084546,-7.3234193,99.0117297,1.9450156)	8351985.654867226
27	(0.5010730,-7.3169680,99.0117014,1.9382202)	8220114.007683980
28	(0.5001742,-7.3161827,99.0116984,1.9373931)	8204178.132465224
29	(0.5001526,-7.3161612,99.0116980,1.9373721)	8204771.227467800
30	(0.4925498,-7.2089382,99.0010214,1.8899364)	6012968.681215596
31	(0.4912089,-7.1906789,99.0009868,1.8791466)	5637322.165966207
32	(0.4908948,-7.1864271,99.0010753,1.8764866)	5549161.155680033
33	(0.4908175,-7.1853819,99.0011018,1.8758243)	5524821.311106794
34	(0.4907809,-7.1848969,99.0011071,1.8755105)	5454284.791599286
35	(0.4907163,-7.1842964,99.0009488,1.8750014)	4914318.060572864
36	(0.4906967,-7.1841262,99.0008920,1.8748489)	4696597.185390180
37	(0.4903084,-7.1733011,98.9972057,1.8646718)	154647261.356157870
38	(0.4871325,-7.0229164,98.9449942,1.7224336)	5280378.732810376
39	(0.4840967,-6.9822470,98.9627282,1.5908263)	3650386.760291170
40	(0.4827949,-6.9692096,98.9733263,1.5342590)	2761455.345443978
41	(0.4818136,-6.9598983,98.9816630,1.4916127)	2068469.921028501
42	(0.4817675,-6.9594642,98.9820555,1.4896133)	2035858.308808296
43	(0.4817449,-6.9592517,98.9822487,1.4886313)	2019362.543485524
44	(0.4817399,-6.9592148,98.9822981,1.4884145)	2012034.858605157
45	(0.4801826,-6.9621920,99.0076267,1.4201242)	643346.391377017
46	(0.4801439,-6.9600913,99.0067980,1.4184310)	593319.530295969
47	(0.4800852,-6.9585443,99.0066400,1.4158600)	559869.515439490
48	(0.4800497,-6.9577229,99.0066205,1.4143067)	542096.300407949
49	(0.4800408,-6.9575236,99.0066201,1.4139170)	515562.865667016

k=0 $\mu(k)$=0.001000 x(k+1)=(0.4800408,-6.9575236,99.0066201,1.4139170)

$\mu(k)\beta$(x(k+1))=1.16364523 (1/$\mu(k)$)α(x(k+1))=10824.74334909 :h1(x(k+1))=-0.0801717

h2(x(k+1))=2.8903242 h3(x(k+1))=1.5698222: g1(x(k+1))=-2.1815071 g2(x(k+1))=-0.0028675

g3(x(k+1))=-0.0012291 g4(x(k+1))=-2.5917517 g5(x(k+1))=-2.0211076

i	x(i+1)	norm(∇M(x(i+1)))
0	(0.4800454,-6.9575652,99.0066172,1.4142974)	5402837.967618135
1	(0.4799902,-6.9575327,99.0068860,1.4142648)	5399735.450002393
2	(0.4758393,-6.9576523,99.0070214,1.4118839)	5409974.893079556
3	(0.4750551,-6.9576174,99.0070267,1.4127040)	5409911.502861264
4	(0.4747313,-6.9575804,99.0070293,1.4129470)	5404578.634826384
5	(0.4746029,-6.9575642,99.0070303,1.4130147)	5401906.049230619
6	(0.4754378,-6.9574759,99.0070211,1.4112440)	5368295.884484630
7	(0.4755796,-6.9574502,99.0070187,1.4107162)	5354420.720385976
8	(0.4756522,-6.9574453,99.0070176,1.4103033)	5328519.838046928
9	(0.5704239,-6.9631263,99.0063615,0.8569162)	2470848.384611745
10	(0.5850504,-6.9651775,99.0036624,0.6491903)	925751.551313647
11	(0.5598965,-6.9652858,99.0027469,0.6333670)	611015.868760972
12	(0.4880651,-6.9651326,99.0021761,0.6428200)	997173.538904449
13	(0.4756472,-6.9650917,99.0022398,0.6469609)	997336.374902548
14	(0.4753118,-6.9650906,99.0022430,0.6470910)	996366.802638463
15	(0.4752342,-6.9650904,99.0022439,0.6471219)	996025.978851925
16	(0.4751793,-6.9650905,99.0022450,0.6471430)	995417.423735502
17	(0.4751690,-6.9650918,99.0022485,0.6471417)	992839.599106683
18	(0.4758125,-6.9656676,99.0036614,0.6445868)	623371.862091304
19	(0.4756356,-6.9658867,99.0034307,0.6413044)	538329.463457556
20	(0.4753815,-6.9664528,99.0031940,0.6319466)	489571.237482410
21	(0.4753095,-6.9666521,99.0031412,0.6285316)	483016.329625197
22	(0.4752806,-6.9667415,99.0031234,0.6269728)	480550.126876211
23	(0.4752735,-6.9667676,99.0031192,0.6265139)	478726.773738459
24	(0.4752739,-6.9668398,99.0031100,0.6252640)	9735550.680860983
25	(0.4766131,-6.9855807,99.0007573,0.3027220)	637317.809478320
26	(0.4768679,-6.9891889,99.0004877,0.2223917)	645512.039727073
27	(0.4770222,-6.9913911,99.0004436,0.1753684)	1321056.147647204

i	x(i+1)	norm(∇M(x(i+1)))
28	(0.4772712,-6.9949481,99.0004958,0.1022923)	227751.961186602
29	(0.4773187,-6.9956299,99.0005256,0.0899061)	177807.990779457
30	(0.4773214,-6.9956697,99.0005376,0.0894619)	149521.111295620
31	(0.4773186,-6.9956285,99.0005463,0.0903859)	125272.582425627
32	(0.4773139,-6.9955604,99.0005550,0.0917653)	98985.705759669
33	(0.4773123,-6.9955362,99.0005569,0.0922249)	86071.272704045
34	(0.4773110,-6.9955143,99.0005580,0.0926198)	65083.148141394
35	(0.4773111,-6.9955096,99.0005573,0.0926826)	43154.789895855
36	(0.4773133,-6.9955292,99.0005549,0.0922924)	24503.717712347
37	(0.4773165,-6.9955599,99.0005524,0.0917154)	13847.180012845
38	(0.4773191,-6.9955792,99.0005511,0.0913580)	8577.827421313
39	(0.4773210,-6.9955866,99.0005506,0.0912226)	5407.425671283
40	(0.4773230,-6.9955839,99.0005509,0.0912736)	3152.227423013
41	(0.4773253,-6.9955739,99.0005516,0.0914595)	1759.956391970
42	(0.4773283,-6.9955648,99.0005522,0.0916266)	1015.032069277
43	(0.4773326,-6.9955608,99.0005525,0.0916992)	617.586133420
44	(0.4773390,-6.9955620,99.0005523,0.0916744)	373.407530141
45	(0.4773481,-6.9955674,99.0005518,0.0915704)	208.056850733
46	(0.4773599,-6.9955739,99.0005513,0.0914460)	116.765921201
47	(0.4773745,-6.9955790,99.0005507,0.0913448)	94.425639574
48	(0.4773937,-6.9955827,99.0005503,0.0912688)	81.117919609
49	(0.4774200,-6.9955858,99.0005498,0.0912020)	48.346900670

k=1 μ(k)0.000100 x(k+1)=(0.4774200,-6.9955858,99.0005498,0.0912020)

μ(k)β(x(k+1))=0.00261112 (1/μ(k))α(x(k+1))=0.00001474 :h1(x(k+1))=0.0000062

h2(x(k+1))= 0.0000115 h3(x(k+1))=-0.0000361: g1(x(k+1))=-0.0423235 g2(x(k+1))=-6.0415929

g3(x(k+1))=-1.4209000 g4(x(k+1))=-3.9084319 g5(x(k+1))=-0.7360868

μ(1)β(x(2))=2.6111e-003

(1/μ(1))α(x(2))=1.4736e-005

Solution is (0.4774200,-6.9955858,99.0005498,0.0912020)

ภาคผนวก ข.

ในส่วนนี้จะได้ให้ซอสโค้ด (source code) ที่ผู้วิจัยได้เขียนขึ้นโดยใช้ส่วนการโปรแกรมของ MATLAB เวอร์ชั่น 5.3 ในการเขียน ผู้วิจัยเลือกแสดงซอสโค้ดของปัญหาที่ 6 ในบทที่ 4 โดยมีโปรแกรมหลัก (I_E_6) ที่ทำหน้าที่ดำเนินการซึ่งจะเรียกใช้โปรแกรมย่อยหลัก (optim 1 , 2 และ 3)โดยที่โปรแกรมย่อยหลักดังกล่าวยังสามารถเรียกใช้โปรแกรมย่อย (fx_6 , gf_6 , gx_6 , hx_6) อื่น ๆ มาใช้งานได้ซึ่งมีผลลัพธ์ของการคำนวณดังแสดงในภาคผนวก ก. สำหรับปัญหาข้ออื่น ๆ ในบทที่ 4 ก็สามารถทำการโปรแกรมได้ในทำนองเดียวกัน

```
**********I_E_6**********
x0=[-0.35;6.9;4.8];
g=gx_6(x0)
m=0;n=0;
for i=1:4
   if g(i) < 0
      m=m+1;
      t(m)=i;
   else
      n=n+1;
      f(n)=i;
   end
end
while m <= 4
   if m==4
      x1=x0;
      fprintf('x0 is interior fesible solution : (%0.7f,%0.7f,%0.7f)\n',...
              x1(1),x1(2),x1(3));
      a=1;
      break
   end
   if m == 0
      a=1;
      t=[0;0;0;0];
      c=[0;0;0];
      fprintf('-----OPTIM6_1-----\n');
      x1=optim6_1(x0,a,t,c);
      x0=x1;
      g=gx_6(x1);
      if g(a) >= 0
         fprintf('No has feasible solution !!!\n');
         break
      end
      t=[1];f=[2;3;4];
   end
   i=1;
   while i <=length(f)
      a=f(i);
      c=[0;0;0];
      fprintf('\n\n');
      fprintf('-----OPTIM6_2-----\n');
      x1=optim6_2(x0,a,t,c);
      x0=x1;
```

```
      g=gx_6(x1);
      if g(a) >= 0
         fprintf('\n');
         fprintf('No has feasible solution !!!\n');
         break
      end
      t(length(t)+1)=f(i);
      i=i+1;
   end
   if g(a) >= 0
      break
   end
   m=length(t)+1;
   fprintf('\n\n');
end
eff1=0.0005;
if g(a) >= 0
   fprintf('No interior  feasible  point  for  inequality  constraints !!!\n');
else
   hx=hx_6(x1);
   if all(abs(hx)< eff1)==1
      fprintf('Solution is (%2.7f,%2.7f,%2.7f)\n',x1(1),x1(2),x1(3));
   else
      a=0;
      t=[1;2;3;4];
      c=[1;1;1];
      fprintf('-----OPTIM6_3-----\n');
      [x1,Bx1,alx1]=optim6_3(x0,a,t,c);
      if (Bx1 < eff1) & (alx1 < eff1)
         fprintf('\n\n');
         fprintf('Solution is (%2.7f,%2.7f,%2.7f)\n',x1(1),x1(2),x1(3));
      else
         fprintf('\n\n');
         fprintf('No feasible  solution  for  all constraints\n');
      end
   end
end

**********fx_6**********

function fx_6=fx_6(x,a,t,c,muek)
a1=[5*x(1)^2+x(2)^2+2*x(1)*x(2)-x(1)+2*x(2)+15*x(3)+3;
   2*x(1)^2+x(2)^2-2*x(2)+6*x(3)+2;
   5*x(1)+3*x(2)+4*x(3)+4;
   4*exp(2*x(1)-x(3))+5*exp(x(2)^2)+30*x(3)];
b1=[a1(1)^(-1);
   a1(2)^(-1);
   a1(3)^(-1);
   a1(4)^(-1)];
c1=[(exp(2*x(1)+5*x(2))+3*x(3)+29)^2;
   (x(1)^4+2*x(2)^2+3*x(3)^2-4*x(1)-4*x(2)*x(3)-1033)^2;
   (10*x(1)+7*x(2)-3*x(3)+6)^2];

if a==0
   a_1=0;
else
   a_1=a1(a);
end

if all(t==0)==1
   b_1=0;
else
   b_1=0;
```

```
   for i=1:length(t)
      b_1=b_1+b1(t(i));
   end
end

if all(c==0)==1
   c_1=0;
else
   c_1=0;
   for i=1:length(c)
      c_1=c_1+c1(i);
   end
end

fx_6=a_1-muek*b_1+(1/muek)*c_1;

**********gf_6**********

function gf_6=gf_6(x,a,t,c,muek)
a1=[10*x(1)+2*x(2)-1 2*x(2)+2*x(1)+2 15;
   4*x(1) 2*x(2)-2 6;
   5 3 4;
   8*exp(2*x(1)-x(3)) 10*x(2)*exp(x(2)^2) 30-4*exp(2*x(1)-x(3))];
b11=5*x(1)^2+x(2)^2+2*x(1)*x(2)-x(1)+2*x(2)+15*x(3)+3;
b21=2*x(1)^2+x(2)^2-2*x(2)+6*x(3)+2;
b31=5*x(1)+3*x(2)+4*x(3)+4;
b41=4*exp(2*x(1)-x(3))+5*exp(x(2)^2)+30*x(3);
b1=[-(b11^(-2))*a1(1,1) -(b11^(-2))*a1(1,2) -(b11^(-2))*a1(1,3);
   -(b21^(-2))*a1(2,1) -(b21^(-2))*a1(2,2) -(b21^(-2))*a1(2,3);
   -(b31^(-2))*a1(3,1) -(b31^(-2))*a1(3,2) -(b31^(-2))*a1(3,3);
   -(b41^(-2))*a1(4,1) -(b41^(-2))*a1(4,2) -(b41^(-2))*a1(4,3)];
c11=exp(2*x(1)+5*x(2))+3*x(3)+29;
c21=x(1)^4+2*x(2)^2+3*x(3)^2-4*x(1)-4*x(2)*x(3)-1033;
c31=10*x(1)+7*x(2)-3*x(3)+6;
c1=[2*c11*2*exp(2*x(1)+5*x(2)) 2*c11*5*exp(2*x(1)+5*x(2)) 2*c11*3;
   2*c21*(4*x(1)^3-4) 2*c21*(4*x(2)-4*x(3)) 2*c21*(6*x(3)-4*x(2));
   2*c31*10 2*c31*7 2*c31*-3];

if a==0
   for j=1:3
      a_1(j)=0;
   end
else
   for j=1:3
      a_1(j)=a1(a,j);
   end
end

if all(t==0)==1
   for j=1:3
      b_1(j)=0;
   end
else
   for j=1:3
      b_1(j)=0;
      for i=1:length(t)
         b_1(j)=b_1(j)+b1(t(i),j);
      end
   end
end

if all(c>0)==1
   for j=1:3
```

```
      c_1(j)=0;
      for i=1:length(c)
         c_1(j)=c_1(j)+c1(i,j);
      end
   end
else
   for j=1:3
      c_1(j)=0;
   end
end

for i=1:3
   gf(i)=a_1(i)-muek*b_1(i)+(1/muek)*c_1(i);
end

gf_6=[gf(1);gf(2);gf(3)];
```

************gx_6************

```
function gx_6=gx_6(x)
gx_6=[5*x(1)^2+x(2)^2+2*x(1)*x(2)-x(1)+2*x(2)+15*x(3)+3;
   2*x(1)^2+x(2)^2-2*x(2)+6*x(3)+2;
   5*x(1)+3*x(2)+4*x(3)+4;
   4*exp(2*x(1)-x(3))+5*exp(x(2)^2)+30*x(3)];
```

************hx_6************

```
function hx_6=hx_6(x)
hx_6=[exp(2*x(1)+5*x(2))+3*x(3)+29;
   x(1)^4+2*x(2)^2+3*x(3)^2-4*x(1)-4*x(2)*x(3)-1033;
   10*x(1)+7*x(2)-3*x(3)+6];
```

************optim6_1************

```
function x1=optim6_1(x0,a,t,c)
i=0;
l=0;
eff1=0.005;
muek=1;
alpha1=1/3;
x1=x0;
gf0=gf_6(x0,a,t,c,muek);
fx0=fx_6(x0,a,t,c,muek);
H0=eye(3);
while (norm(gf0) >= eff1)
   p0=-H0*gf0;
   al=1;
   x1=x0+al*p0;
   fx1=fx_6(x1,a,t,c,muek);
   while  fx1 > fx0+alpha1*al*dot(p0,gf0)
      al=al/2;
      x1=x0+al*p0;
      fx1=fx_6(x1,a,t,c,muek);
      if l==4
         break;
      end
      l=l+1;
   end
   l=0;
   gf1=gf_6(x1,a,t,c,muek);
   g=gx_6(x1);
   if (g(a) < 0)
```

```
      fprintf('%0.0f        (%2.7f,%2.7f,%2.7f)   %0.6f    %2.5f\n',...
              i,x1(1),x1(2),x1(3),norm(gf1),g(a));
      break;
   end
   fprintf('%0.0f        (%2.7f,%2.7f,%2.7f)   %0.6f    %2.5f\n',...
           i,x1(1),x1(2),x1(3),norm(gf1),g(a));
   i=i+1;
   p0=x1-x0;
   q0=gf1-gf0;
   H0=H0+(1+q0'*H0*q0/(q0'*p0))*(p0*p0'/(p0'*q0))-((p0*q0'*H0+H0*q0*p0')/
      (q0'*p0));
   gf0=gf1;
   x0=x1;
   fx0=fx1
end

**********optim6_2**********

function x1=optim6_2(x0,a,t,c)
i=0;
k=0;
l=0;
eff1=0.000005;
eff2=0.000005;
muek=10;
beta=0.1;
Bx=1;
alpha1=1/3;
x1=x0;
g=gx_6(x0);
if g(a) < 0
   fprintf('%0.0f  %0.6f          (%2.7f,%2.7f,%2.7f)   %2.5f:',...
           k,muek,x1(1),x1(2),x1(3),g(a));
   for j=1:length(t)
      fprintf('     %2.7f',g(t(j)));
   end
end
fprintf('\n\n');
while (muek*Bx/beta >= eff1) & (g(a) >=0)
   gf0=gf_6(x0,a,t,c,muek);
   H0=eye(3);
   while (norm(gf0) >= eff2)
      fx0=fx_6(x0,a,t,c,muek);
      p0=-H0*gf0;
      al=1;
      x1=x0+al*p0;
      fx1=fx_6(x1,a,t,c,muek);
      while fx1 > fx0+alpha1*al*dot(p0,gf0)
         al=al/2;
         x1=x0+al*p0;
         fx1=fx_6(x1,a,t,c,muek);
         if l==4
            break;
         end
         l=l+1;
      end
      l=0;
      g=gx_6(x1);
      for j=1:length(t)
         g1(j)=g(t(j));
      end
      while all(g1 < 0)==0
         al=al/2;
```

```
        x1=x0+al*p0;
        g=gx_6(x1);
        for j=1:length(t)
           g1(j)=g(t(j));
        end
     end
     gf1=gf_6(x1,a,t,c,muek);
     fprintf('%0.0f   (%2.7f,%2.7f,%2.7f)      %0.5f   g(%0.0f)=%0.5f\n',i,x1
           (1),x1(2),x1(3),norm(gf1),a,g(a));
     if (g(a) < 0)
        break;
     end
     i=i+1;
     p0=x1-x0;
     q0=gf1-gf0;
     H0=H0+(1+q0'*H0*q0/(q0'*p0))*(p0*p0'/(p0'*q0))-((p0*q0'*H0+H0*q0*p0')/
        (q0'*p0));
     gf0=gf1;
     x0=x1;
  end
  i=0;
  Bx=0;
  for j=1:length(t)
     Bx=Bx-(1/g(t(j)));
  end
  fprintf('%0.0f  %0.6f          (%2.7f,%2.7f,%2.7f)   %2.8f g(%0.0f)=%2.7f:',...
          k,muek,x1(1),x1(2),x1(3),muek*Bx,a,g(a));
  for j=1:length(t)
     fprintf('    g(%0.0f)=%2.7f',j,g(t(j)));
  end
  muek=muek*beta;
  k=k+1;
end

**********optim6_3**********

function [x1,Bx1,alx1]=optim6_3(x0,a,t,c)
i=0;
k=0;
l=0;
eff1=0.0005;
eff2=0.00005;
muek=0.001;
beta=0.1;
Bx=1;
alx=10;
alpha1=1/3;
x1=x0;
while (muek*Bx/beta >= eff1) | (alx*beta/muek >= eff1)
   gf0=gf_6(x0,a,t,c,muek);
   H0=eye(3);
   while (norm(gf0) >= eff2)
      fx0=fx_6(x0,a,t,c,muek);
      p0=-H0*gf0;
      al=1;
      x1=x0+al*p0;
      fx1=fx_6(x1,a,t,c,muek);
      while fx1 > fx0+alpha1*al*dot(p0,gf0)
         al=al/2;
         x1=x0+al*p0;
         fx1=fx_6(x1,a,t,c,muek);
         if l==4
            break;
```

```
        end
        l=l+1;
     end
     l=0;
     g=gx_6(x1);
     while all(g < 0)==0
        al=al/2;
        x1=x0+al*p0;
        g=gx_6(x1);
     end
     gf1=gf_6(x1,a,t,c,muek);
     fprintf('%0.0f    (%2.7f,%2.7f,%2.7f)       %2.7f\n',i,x1(1),x1(2),x1
            (3),norm(gf1));
     i=i+1;
     p0=x1-x0;
     q0=gf1-gf0;
     H0=H0+(1+q0'*H0*q0/(q0'*p0))*(p0*p0'/(p0'*q0))-((p0*q0'*H0+H0*q0*p0')/
        (q0'*p0));
     gf0=gf1;
     x0=x1;
   end
   i=0;
   hx=hx_6(x1);
   alx=hx(1)^2+hx(2)^2+hx(3)^2;
   y=0;
   Bx=0;
   for j=1:length(g)
      Bx=Bx-(1/g(j));
   end
   fprintf('%0.0f  %0.6f         (%2.7f,%2.7f,%2.7f)    %0.5f    %2.8f    %2.8f
          :%2.7f    %2.7f    %2.7f:',...k,muek,x1(1),x1(2),x1(3),y,muek*Bx,
          (1/muek)*alx,hx(1),hx(2),hx(3));
   for j=1:length(g)
      fprintf('      g(%0.0f)=%2.7f',j,g(j));
   end
   fprintf('\n');
   k=k+1;
   muek=muek*beta;
end
Bx1=muek*Bx/beta;
alx1=alx*beta/muek;
```

ประวัติผู้เขียน

นายเสกสรร สิริทรัพย์ทวี เกิดเมื่อวันที่ 15 มีนาคม 2521 ที่จังหวัดสุพรรณบุรี สำเร็จการศึกษาวิศวกรรมศาสตรบัณฑิต สาขาอิเล็กทรอนิกส์ จากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังในปีการศึกษา 2542 และได้รับทุนโครงการพัฒนาอาจารย์สาขาคณิตศาสตร์ประยุกต์ จากมหาวิทยาลัยบูรพาในปี พ.ศ. 2544