วิธีลดเร็วที่สุดสำหรับกำหนดการเชิงเส้น

STEEPEST DESCENT METHOD FOR LINEAR PROGRAMMING

ดุษฎา ปานประสิทธิ์

DUSADA PANPRASIT

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาคณิตศาสตร์ประยุกต์

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2543

ISBN 974-622-838-2

วิธีลดเร็วที่สุดสำหรับกำหนดการเชิงเส้น

STEEPEST DESCENT METHOD FOR LINEAR PROGRAMMING

ดุษยา ปานประสิทธิ์

DUSAYA PANPRASIT

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาคณิตศาสตร์ประยุกต์

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2543

ISBN 974-622-838-2

STEEPEST DESCENT METHOD FOR LINEAR PROGRAMMING

DUSAYA PANPRASIT

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE PROGRAM IN APPLIED MATHEMATICS
SCHOOL OF GRADUATE STUDIES
KING MONGKUT' S INSTITUTE OF TECHNOLOGY LADKRABANG
2000

ISBN 974-622-838-2

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์	วิธีลดเร็วที่สุดสำหรับกำหนดการเชิงเส้น STEEPEST DESCENT METHOD FOR LINEAR PROGRAMMING
ชื่อนักศึกษา	นางสาวดุษยา ปานประสิทธิ์
รหัสประจำตัว	39065367
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	คณิตศาสตร์ประยุกต์
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ.อุบลวรรณา เงินวิจิตร
อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม	รศ.ดร.อำพล ธรรมเจริญ

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
รศ.ผ่องพรรณ	ตันธนาวันต์	
รศ.อุบลวรรณา	เงินวิจิตร	
รศ.ดร.อำพล	ธรรมเจริญ	
ดร.สมศรี	บัณฑิตวิไล	
ผศ.กฤษฎา	ไตรสุรัตน์	

วัน/เดือน/ปี ที่สอบ 23 พฤษภาคม 2543 เวลา 10.00 น. เป็นต้นไป

สถานที่สอบ ณ ห้อง 424 ห้องประชุม-สัมมนา

บัณฑิตวิทยาลัยรับรองแล้ว

(รศ.ดร.วันชัย สังวรสิทธิ์)

คณบดีบัณฑิตวิทยาลัย

วันที่...31.....เดือน...พฤษภาคม...พ.ศ...2543....

หัวข้อวิทยานิพนธ์	วิธีลดเร็วที่สุดสำหรับกำหนดการเชิงเส้น
นักศึกษา	นางสาวดุษยา ปานประสิทธิ์
รหัสประจำตัว	39065367
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	คณิตศาสตร์ประยุกต์
พ.ศ.	2543
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ.อุบลวรรณา เงินวิจิตร
อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม	รศ.ดร.อำพล ธรรมเจริญ

บทคัดย่อ

ปัญหากำหนดการเชิงเส้นในรูปแบบของคาร์มาร์คาร์ (Karmarkar type problem)อยู่ในรูปแบบ

$$\begin{aligned} \text{Minimize} \quad & \mathbf{c}^T\mathbf{x} \\ \text{Subject to} \quad & A\mathbf{x} = 0 \\ & \mathbf{e}^T\mathbf{x} = 1 \\ & \mathbf{x} \geq 0 \end{aligned}$$

เมื่อ A เป็นเมทริกซ์ ขนาด $m \times n$

$\mathbf{e}^T = (1,1,1,\ldots 1)^T$ เป็นเวกเตอร์ของ 1 ขนาด n เวกเตอร์

$\mathbf{c}, \mathbf{x} \in R^n$

ซึ่งเมื่อแปลงเป็นปัญหาคู่กันแล้วจะได้ปัญหาอยู่ในรูปแบบ

$$\min_{\mathbf{x} \in \mathbf{X}} \max \ \{L_1(\mathbf{x}), L_2(\mathbf{x}), \ldots, L_m(\mathbf{x})\}\ , \ \mathbf{x} \in R^n$$

เมื่อ $L_i(\mathbf{x})$; $i = 1,2,\ldots,m$ เป็นฟังก์ชันเชิงเส้น

เมื่อแก้ปัญหานี้ได้เท่ากับการแก้ปัญหากำหนดการเชิงเส้นได้ วิธีที่จะใช้ในการแก้ปัญหานี้เลือกใช้วิธีลดเร็วที่สุด (Steepest Descent Method) ฟังก์ชันเป้าหมายของปัญหาไม่มีอนุพันธ์ที่บางจุด งานวิจัยนี้จึงศึกษาเพื่อหาทิศทางลดเร็วที่สุดที่จุดที่หาอนุพันธ์ไม่ได้ และหาขั้นตอนวิธีในการแก้ปัญหาดังกล่าว พร้อมทั้งเขียนโปรแกรมคอมพิวเตอร์เพื่อช่วยในการทดสอบการแก้ปัญหากำหนดการเชิงเส้นแบบคาร์มาร์คาร์

Thesis Title	Steepest Descent Method for Linear Programming
Student	Miss. Dusaya Panprasit
Student ID.	39065367
Degree	Master of Science
Programme	Applied Mathematics
Year	2000
Thesis Advisor	Assoc.Prof.Ubolwanna Ngernwichit
Thesis Coadvisor	Assoc.Prof.Dr.Ampon Dhamacharoen

ABSTRACT

The Karmarkar type linear programming problem is of the form :

$$\begin{array}{ll} \text{Minimize} & \mathbf{c}^T\mathbf{x} \\ \text{Subject to} & A\mathbf{x} = \mathbf{0} \\ & \mathbf{e}^T\mathbf{x} = 1 \\ & \mathbf{x} \geq \mathbf{0} \end{array}$$

where A is an $m \times n$ dimensional matrix of full row rank

$\mathbf{e}^T = (1,1,1,\ldots 1)^T$ is an n-vector of all ones and

$\mathbf{c}, \mathbf{x} \in R^n$

which its dual programming problem has the form

$$\min_{\mathbf{x}\in\mathbf{X}} \max \ \{L_1(\mathbf{x}), L_2(\mathbf{x}), \ldots, L_m(\mathbf{x})\}\ , \ \mathbf{x} \in R^n$$

where $L_i(\mathbf{x})$; $i = 1,2,\ldots,m$ are the linear functions.

The solution of the latter problem will yield the solution of the former Linear Programming problems. In this research will use the Steepest Descent method to solve this problem. Since the objective functions in this problem are nondifferentiable, the steepest descent direction at the nondifferentiable point will be discussed. The algorithms will be constructed based on derived informations. Along with the algorithms, the computer program will be coded and tested using the Karmarkar type problem.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้เป็นอย่างดี ด้วยคำแนะนำและคำปรึกษาจาก รศ.อุบลวรรณา เงินวิจิตร อาจารย์ผู้ควบคุมวิทยานิพนธ์และ รศ.ดร. อำพล ธรรมเจริญ ภาควิชาคณิตศาสตร์ มหาวิทยาลัยบูรพา ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม ผู้วิจัยรู้สึกซาบซึ้งในความอนุเคราะห์จากท่านทั้งสอง และขอกราบขอบพระคุณเป็นอย่างสูง

ขอขอบคุณเพื่อนๆ นักศึกษาทุกคนที่ช่วยเหลือให้คำแนะนำต่างๆและยังให้กำลังใจต่อผู้วิจัยอย่างใกล้ชิดตลอดมา

คุณค่าและประโยชน์อันพึงมีจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่ผู้มีพระคุณทุกท่าน

ดุษยา ปานประสิทธิ์

สารบัญ

หน้า

สารบัญ (ต่อ)

หน้า

สารบัญภาพ

บทที่ 1
บทนำ

1.1 ความสำคัญและที่มาของวิทยานิพนธ์

กำหนดการเชิงเส้น (linear programming) เป็นแบบทางคณิตศาสตร์ซึ่งเป็นที่รู้จักกันดีและมีการประยุกต์ใช้มากมายในการแก้ปัญหาต่างๆ เช่น ปัญหาการขนส่ง ปัญหาการจัดสรรทรัพยากรที่มีจำกัดให้ได้ประโยชน์สูงสุด ปัญหาการวางแผนผลิตสินค้าให้ได้กำไรสูงสุดหรือให้มีต้นทุนต่ำสุด วิธีทางคณิตศาสตร์วิธีหนึ่งที่ใช้กันทั่วไปเป็นที่แพร่หลายในการแก้ปัญหากำหนดการเชิงเส้น คือ วิธีซิมเพลกซ์ (simplex method) แต่ก็ยังมีผู้วิจัยจำนวนมากพยายามที่จะหาวิธีที่ดีกว่าวิธีซิมเพลกซ์ เช่นวิธีเชิงทรงรี (ellipsoid method) และวิธีจุดภายใน (interior point method) อันประกอบด้วยวิธีของคาร์มาร์คาร์ (Karmarkar method) วิธีปรับมาตราจากภาพฉาย (Projective Scaling) และอื่นๆ [1] วิธีที่มีต้นแบบใหม่นี้จะใช้แก้ปัญหาได้ดีกว่าวิธีซิมเพลกซ์ในบางด้าน กล่าวคือบางวิธีเช่น วิธีของคาร์มาร์คาร์จะมีจำนวนครั้งสูงสุดของการกระทำเป็นฟังก์ชันพหุนาม (polynomial-time) ของขนาดของปัญหาในขณะที่วิธีซิมเพลกซ์จะมีจำนวนครั้งสูงสุดของการกระทำเป็นฟังก์ชันชี้กำลัง (exponential-time) ของขนาดของปัญหา ซึ่งในทางทฤษฎีวิธีคาร์มาร์คาร์ดีกว่าวิธีซิมเพลกซ์ แต่ว่าในขั้นตอนวิธีของคาร์มาร์คาร์มีความยุ่งยากมากกว่า ดังนั้นจึงยังคงนิยมใช้วิธีซิมเพลกซ์สำหรับปัญหาที่มีขนาดเล็ก แต่สำหรับปัญหาที่มีขนาดใหญ่มีการทดลองและรายงานผลว่าวิธีจุดภายในหลายวิธีใช้แก้ปัญหาได้ดีกว่าวิธีซิมเพลกซ์ [2]

ในวิทยานิพนธ์นี้ผู้วิจัยได้ทดลองที่จะใช้วิธีลดเร็วที่สุด (steepest descent method) มาใช้แก้ปัญหากำหนดการเชิงเส้น โดยการแปลงปัญหากำหนดการเชิงเส้นที่มีเงื่อนไขจำกัดไปเป็นปัญหาการหาค่าต่ำสุดของฟังก์ชันไม่เชิงเส้น (nonlinear function) ปัญหาที่แปลงแล้วนี้จะแก้ปัญหาได้ง่ายกว่าปัญหาเดิมเพราะว่าไม่มีเงื่อนไขหรือมีเพียงบางส่วน วิธีแก้ปัญหากำหนดการไม่เชิงเส้น (nonlinear programming) ที่ไม่มีเงื่อนไขมีหลายวิธี เช่น วิธีลดเร็วที่สุด วิธีของนิวตัน (Newton method) วิธีของบรอยเดน (Broyden method) สองวิธีหลังนี้เป็นวิธีที่มีประสิทธิภาพมากแต่ฟังก์ชันที่จะหาค่าสูงสุดต้องเป็นฟังก์ชันที่มีอนุพันธ์ทุกจุด แต่ปัญหาที่แปลงมาจากปัญหากำหนดการเชิงเส้นจะไม่มีอนุพันธ์ที่บางจุด ดังนั้นจึงไม่สามารถจะใช้วิธีของนิวตันได้ ปัญหาที่มีเงื่อนไขบางปัญหาก็อาจใช้วิธีของนิวตันหรือของบรอยเดนไม่ได้ วิธีที่ผู้วิจัยจะเลือกใช้คือใช้วิธีลดเร็วที่สุดซึ่งผู้วิจัยจะต้องปรับวิธีการให้เหมาะกับปัญหาต่อไป

เราทราบว่าในแต่ละขั้นตอนของวิธีซิมเพลกซ์จะเป็นการเคลื่อนจากจุดเดิมไปยังจุดใกล้เคียงเพื่อทำให้ค่าของฟังก์ชันจุดประสงค์มีค่าน้อยลง แต่วิธีลดเร็วที่สุดจะเป็นการเคลื่อนจากจุด

เดิมไปยังจุดใหม่ในทิศที่ค่าของฟังก์ชันจุดประสงค์ลดลงเร็วที่สุด ดังนั้นผู้วิจัยจึงเลือกใช้วิธีนี้ เพราะคาดหวังว่าจะเข้าสู่ผลเฉลยได้เร็วกว่าวิธีซิมเพลกซ์

1.2 ปัญหากำหนดการเชิงเส้นและการแปลง

ปัญหากำหนดการเชิงเส้นที่จะพิจารณาคือปัญหาในแบบของคาร์มาร์คาร์ (Karmarkar type problem)

$$
\text{(P)} \qquad \begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{0} \\ & \mathbf{e}^T\mathbf{x} = 1 \\ & \mathbf{x} \geq \mathbf{0} \end{array}
$$

เมื่อ $\mathbf{A} = [a_{ij}]_{m \times n}$ เป็นเมทริกซ์ขนาด $m \times n$

$\mathbf{e}^T = [1\ 1\ 1\ \dots\ 1]$ เป็น เวกเตอร์ของ 1 ทั้งหมด n จำนวน

$\mathbf{c} = (c_1, c_2, \dots, c_n) \in R^n$

และ $\mathbf{x} = (x_1, x_2, \dots, x_n) \in R^n$

สำหรับปัญหากำหนดการเชิงเส้นทั่วไปสามารถจะแปลงให้อยู่ในแบบปัญหาของคาร์มาร์คาร์ได้ [3] ในที่นี้เราจึงสนใจแก้ปัญหาในแบบของคาร์มาร์คาร์เท่านั้น

จากปัญหาแบบของคาร์มาร์คาร์เราหาปัญหาคู่กัน (dual problem) ของปัญหาคาร์มาร์คาร์ ได้เป็น

$$
\text{(D)} \qquad \begin{array}{ll} \text{maximize} & y_{m+1} \\ \text{subject to} & \sum_{i=1}^{m} a_{ij} y_i + y_{m+1} \leq c_j; \quad j = 1, 2, \dots, n \\ & y_i \in R; \quad i = 1, 2, \dots, m+1 \end{array}
$$

เมื่อให้ $w = -y_{m+1}$ จะได้ปัญหาในรูปแบบที่เราพิจารณา คือ

$$
\text{(A)} \qquad \begin{array}{ll} \text{minimize} & w \\ \text{subject to} & w \geq \sum_{i=1}^{m} a_{ij} y_i - c_j \ ; \ j = 1, 2, \dots, n \\ & y_i \in R \ ; \quad i = 1, 2, \dots, m+1 \end{array}
$$

ให้ $w_j = \sum_{i=1}^{m} a_{ij} y_i - c_j$

$= L_j(\mathbf{y})$

เมื่อ $\mathbf{y} = (y_1, y_2, ..., y_m)$

คือเราหาค่าต่ำสุดของ w เมื่อ w เป็นค่าสูงสุดของ w_j

นั่นคือปัญหา

(A') $$\min_{\mathbf{y} \in Y} \max \{L_1(\mathbf{y}), L_2(\mathbf{y}), ..., L_m(\mathbf{y})\}, \quad \mathbf{y} \in R^n$$

เมื่อ $L_i(\mathbf{y})$; $i=1, 2, ..., m$ เป็นฟังก์ชันเชิงเส้น

เราจะศึกษาวิธีการแก้ปัญหา (A) ซึ่งจะเป็นการแก้ปัญหากำหนดการเชิงเส้น (P) ตามต้องการ

1.3 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

เพื่อหาขั้นตอนวิธีที่จะนำมาใช้ในการแก้ปัญหากำหนดการเชิงเส้นด้วยวิธีลดเร็วที่สุด

1.4 ขอบเขตการวิจัย

ในงานวิจัยนี้กล่าวถึงการแก้ปัญหาในรูปแบบ

$$\min_{\mathbf{x} \in X} \max \{L_1(\mathbf{x}), L_2(\mathbf{x}), ..., L_m(\mathbf{x})\}, \quad \mathbf{x} \in R^n$$

เมื่อ $L_i(\mathbf{x})$; $i=1, 2, ..., m$ เป็นฟังก์ชันเชิงเส้น

โดยมีขอบเขตดังนี้

1. จะดำเนินการหาทิศทางลดเร็ว และทิศทางลดเร็วที่สุด
2. หาขั้นตอนวิธีที่ใช้ในการแก้ปัญหาพร้อมทั้งจะหาทฤษฎีบทที่จะแน่ใจได้ว่าวิธีที่เสนอนี้ทำงานได้ถูกต้องและแสดงตัวอย่างที่ใช้แก้ปัญหา

1.5 ขั้นตอนของการศึกษา

ขั้นตอนที่ 1 ค้นคว้าและศึกษาเอกสารและข้อมูลที่เกี่ยวข้อง

ขั้นตอนที่ 2 หาขั้นตอนวิธีในการแก้ปัญหากำหนดการเชิงเส้นด้วยวิธีลดเร็วที่สุด

ขั้นตอนที่ 3 ทดสอบขั้นตอนวิธีที่ได้กับปัญหา
ขั้นตอนที่ 4 สรุปผลและเขียนวิทยานิพนธ์

บทที่ 2
ความรู้พื้นฐานและงานวิจัยที่เกี่ยวข้อง

2.1 ความรู้พื้นฐาน

2.1.1 ความรู้พื้นฐานเกี่ยวกับปริภูมิเวกเตอร์

กำหนดสัญลักษณ์ R^n เป็นเซตของอันดับ n ตัว $(x_1, x_2, \ldots, x_n)$ เมื่อ $x_1, x_2, \ldots, x_n$ เป็นจำนวนจริงเราใช้สัญลักษณ์ $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ และ R^n เป็นปริภูมิเวกเตอร์ (vector space) ของจำนวนจริง

ถ้าเรากำหนดให้ X เป็นเซตย่อยของ R^n ใดๆซึ่ง $X \neq \phi$ และ d เป็นเมตริก นิยามบน X แล้วเรียกเซต X กับเมตริก d ว่า *ปริภูมิเมตริก* (metric space) แทนด้วยสัญลักษณ์ (X, d)

ถ้ามี x_0 ซึ่ง $x_0 \in X$ และ r เป็นจำนวนจริงบวกใดๆ แล้ว ทรงกลมเปิดที่มีจุดศูนย์กลางที่ x_0 และรัศมีเท่ากับ r แทนด้วย $B(x_0; r)$ กำหนดโดย

$$B(x_0; r) = \{\mathbf{x} \mid \mathbf{x} \in X, d(\mathbf{x}, x_0) < r\}$$

ถ้า A เป็นเซตย่อยของ X และ $\mathbf{x} \in X$ เรากล่าวว่า $\mathbf{x}$ เป็น *จุดภายใน* (interior points) ของ A ก็ต่อเมื่อ มีจำนวนจริงบวก $r > 0$ ซึ่ง $B(\mathbf{x}; r) \subseteq A$ คือ มีทรงกลมเปิดที่มี $\mathbf{x}$ เป็นจุดศูนย์กลาง เป็นเซตย่อยของ A เซตของจุดภายในทั้งหมดของ A แทนด้วยสัญลักษณ์ $Int(A)$ คือ

$$Int(A) = \{\mathbf{x} \mid \mathbf{x} \in A, B(\mathbf{x}; r) \subseteq A\}$$

สำหรับ $r > 0$ บางตัว

สำหรับ A ที่เป็นเซตย่อยของ X เรากล่าวว่าเซต A เป็น *เซตเปิด* (open set) ก็ต่อเมื่อ $A = Int(A)$ กล่าวคือ สำหรับทุกๆ $\mathbf{x} \in A$ จะมีจำนวนจริงบวก $r > 0$ ซึ่ง $B(\mathbf{x}; r) \subseteq A$

ในกรณีที่ $A \subseteq X$ และ $\mathbf{x}_0 \in X$ เราเรียกจุด $\mathbf{x}_0$ ว่าเป็น ***จุดลิมิต*** (limit points) ของเซต A ก็ต่อเมื่อ

$$(B(\mathbf{x}_0; r) - \{\mathbf{x}_0\}) \cap A \neq \phi$$

สำหรับทุก $r > 0$ แสดงว่าทุกๆ ทรงกลมเปิดที่มีจุดศูนย์กลางที่ $\mathbf{x}_0$ มีจุดที่ต่างจาก $\mathbf{x}_0$ อยู่ใน A

เรากล่าวว่าเซต $F \subset X$ เป็น*เซตปิด* (closed set) ก็ต่อเมื่อทุกๆจุดลิมิตของ F เป็นสมาชิกของ F นั่นเอง

เมื่อ $\mathbf{x} \in R^n$ เรากำหนดค่า $\|\mathbf{x}\|$ เป็น *นอร์มแบบยูคลิด* (Euclidean norm) ของ $\mathbf{x}$ คือ

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

ให้ S เป็นเซตของเวกเตอร์ใน R^n ถ้าทุก ๆ สมาชิก $\mathbf{x} \in S$ มีนอร์มเป็นค่าจำกัด คือ $\|\mathbf{x}\| < \infty$ เรากล่าวว่า S เป็นเซตที่มีขอบเขต

เรากล่าวว่า เซตของเวกเตอร์ใน R^n เป็น *เซตปกคลุมแน่น* (compact set) ก็ต่อเมื่อเป็นเซตที่มีขอบเขตและเป็นเซตปิด

ถ้า $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ และ $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ เป็นเวกเตอร์ใน R^n ระยะทางระหว่างเวกเตอร์ คือ

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}$$

เรากำหนด *ผลคูณภายใน* (inner product) ของเวกเตอร์ $\mathbf{x}$ และ $\mathbf{y}$ โดย

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = \sum_{i=1}^{n} x_i y_i$$

และจะเห็นว่า $\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2$

ให้นิยามมุม θ ว่าเป็นมุมระหว่างเวกเตอร์ $\mathbf{x}$ และ $\mathbf{y}$ ดังนี้

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

บางทีเราเขียนเวกเตอร์ในแบบเมทริกซ์คอลัมน์ $\mathbf{x} = [x_1 \; x_2 \; \ldots \; x_n]^T$ และ $\mathbf{y} = [y_1 \; y_2 \; \ldots \; y_n]^T$ ดังนั้นผลคูณภายในของเวกเตอร์ $\mathbf{x}$ และ $\mathbf{y}$ อาจเขียนในแบบ $\mathbf{x}^T \mathbf{y}$

ให้ f เป็นฟังก์ชันจำนวนจริงที่นิยามบนเซต R^n เรากล่าวว่า f ต่อเนื่องที่ $\mathbf{x}$ ถ้า $\mathbf{x}_k \to \mathbf{x}$ แล้ว $f(\mathbf{x}_k) \to f(\mathbf{x})$ หรืออีกนัยหนึ่ง ฟังก์ชัน f ต่อเนื่องที่ $\mathbf{x}$ ถ้ากำหนดให้ $\varepsilon > 0$ จะมี $\delta > 0$ ซึ่งถ้า $|\mathbf{y} - \mathbf{x}| < \delta$ แล้ว $|f(\mathbf{y}) - f(\mathbf{x})| < \varepsilon$

ให้ f เป็นฟังก์ชันต่อเนื่องบนเซต S ถ้า S เป็นเซตปกคลุมแน่นแล้ว f มีค่ามากที่สุด และ f มีค่าน้อยที่สุด

อนุพันธ์ย่อยของฟังก์ชัน f เทียบกับ x_i ที่จุด $\bar{\mathbf{x}}$ เขียนสัญลักษณ์ $\frac{\partial f}{\partial x_i}(\bar{\mathbf{x}})$ หรือ $f_{x_i}(\bar{\mathbf{x}})$ กำหนดโดย

$$\frac{\partial f}{\partial x_i}(\bar{\mathbf{x}}) = \lim_{h \to 0} \frac{f(x_1, x_2, \dots, x_i + h, \dots, x_n) - f(x_1, x_2, \dots, x_n)}{h}$$

และกำหนดสัญลักษณ์ $\nabla f(\bar{\mathbf{x}})$ มีชื่อว่า *เกรเดียนต์* (gradient) ของ f ที่จุด $\bar{\mathbf{x}}$ โดย

$$\nabla f(\bar{\mathbf{x}}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$

ให้ S เป็นเซตไม่ว่างใน R^n กำหนดฟังก์ชัน $f : S \to R$ และ $\alpha : R^n \to R$ และ $\nabla f(\bar{\mathbf{x}})$ เป็น เกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}} \in \operatorname{int} S$

1. เรากล่าวว่าฟังก์ชัน f มีอนุพันธ์ที่จุด $\bar{\mathbf{x}}$ ถ้า

$$f(\mathbf{x}) = f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}}) \cdot (\mathbf{x} - \bar{\mathbf{x}}) + \|\mathbf{x} - \bar{\mathbf{x}}\| \, \alpha(\bar{\mathbf{x}}; \mathbf{x} - \bar{\mathbf{x}})$$

สำหรับแต่ละ $\mathbf{x} \in S$

เมื่อ $\lim_{\mathbf{x} \to \bar{\mathbf{x}}} \alpha(\bar{\mathbf{x}} : \mathbf{x} - \bar{\mathbf{x}}) = 0$

2. เรากล่าวว่าฟังก์ชัน f มีอนุพันธ์ต่อเนื่องบนเซตเปิด $S \subset R^n$ ถ้า f มีอนุพันธ์ต่อเนื่องที่แต่ละจุดใน S

ให้ S เป็นเซตไม่ว่างใน R^n และ $f : S \to R$ ให้ $\bar{\mathbf{x}} \in S$ และ $\mathbf{u}$ เป็นเวกเตอร์ที่ไม่ใช่เวกเตอร์ศูนย์ ซึ่งทำให้ $\bar{\mathbf{x}} + \lambda \mathbf{u} \in S$ เมื่อ $\lambda > 0$ และมีขนาดเล็กพอ *อนุพันธ์ทิศทาง* (directional derivative) ของ f ที่จุด $\bar{\mathbf{x}}$ ในทิศเวกเตอร์ $\mathbf{u}$ กำหนดโดย

$$f'(\bar{\mathbf{x}}; \mathbf{u}) = \lim_{\lambda \to 0^+} \frac{f(\bar{\mathbf{x}} + \lambda \mathbf{u}) - f(\bar{\mathbf{x}})}{\lambda}$$

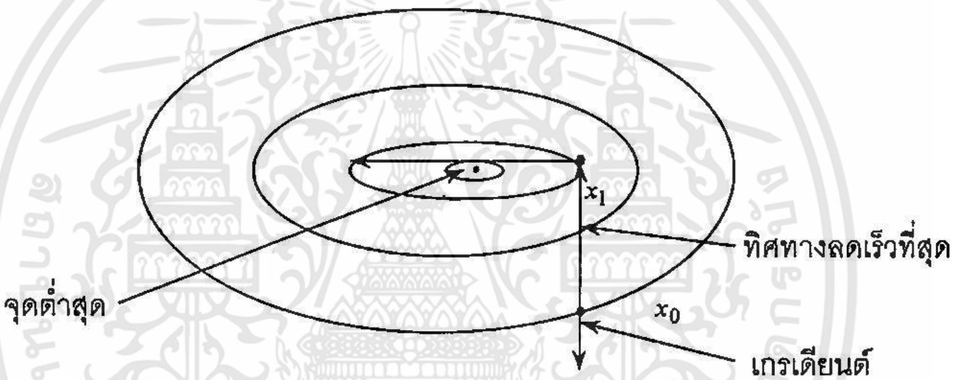
ถ้าฟังก์ชัน f มีอนุพันธ์ที่จุด $\bar{\mathbf{x}}$ เราสามารถแสดงได้ว่า

$$f'(\bar{\mathbf{x}};\mathbf{u}) = \mathbf{u}\cdot\nabla f(\bar{\mathbf{x}})$$

2.1.2 วิธีลดเร็วที่สุด

วิธีลดเร็วที่สุดเป็นวิธีเก่าที่ใช้กันมานานมีหลักการง่ายๆคือ ในแต่ละครั้ง ณ ที่จุดหนึ่ง เราหาทิศทางซึ่งจะทำให้ฟังก์ชันมีค่าลดลงเร็วที่สุด แล้วเปลี่ยนจุดให้ไปตามทิศทางนั้น จนถึงจุดที่ฟังก์ชันมีค่าน้อยที่สุด ณ ที่จุดใหม่นี้จะหาทิศทางลดเร็วที่สุดอันใหม่ต่อไป

โดยปกติแล้วถ้าฟังก์ชันมีอนุพันธ์ทุกจุด ทิศทางลดเร็วที่สุดจะเป็นทิศทางที่ตรงข้ามกับเกรเดียนต์ของฟังก์ชัน ณ จุดนั้น



ภาพที่ 2.1 โค้งระดับทิศทางลดเร็วที่สุด

รายละเอียดเป็นดังนี้

ในการแก้ปัญหา

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

ให้จุด $\mathbf{x}^{(0)}$ เป็นจุดเริ่มต้นเราจะหาทิศทาง $\mathbf{d}$ ที่ทำให้

$$f\left(\mathbf{x}^{(0)}+t\mathbf{d}\right)< f\left(\mathbf{x}^{(0)}\right)$$

เมื่อ $t\ge 0$ ทิศทาง $\mathbf{d}$ มีชื่อว่า ทิศทางลด (descent direction)

ถ้ามี $\delta>0$ และทิศทาง $\mathbf{d}^*$ ซึ่ง $\|\mathbf{d}^*\|=1$ ที่ทำให้

$$f\left(\mathbf{x}^{(0)} + t\mathbf{d}^*\right) \le f\left(\mathbf{x}^{(0)} + t\mathbf{d}\right)$$

สำหรับ $0 < t < \delta$ สำหรับทุก $\mathbf{d}$ ที่เป็นทิศทางลดและ $\|\mathbf{d}\| = 1$ เรากล่าวว่า $\mathbf{d}^*$ เป็นทิศทางลดเร็วที่สุด ของ f ที่จุด $\mathbf{x}^{(0)}$

ขั้นตอนวิธีลดเร็วที่สุด

กำหนดให้มีจุดเริ่มต้น (initial point) $\mathbf{x}^{(0)}$ แล้ว

1. สำหรับ $i = 1,2,\ldots$ ให้ $\mathbf{d}_i$ เป็นทิศทางลดเร็วที่สุด
2. แก้ปัญหาการหาค่าต่ำสุดในแนวเส้น (line search)

$$\min_{t} \quad f(\mathbf{x}_i + t\mathbf{d}_i)$$

ได้ค่า t_k ที่ทำให้ $f(\mathbf{x}_i + t_k\mathbf{d}_i)$ มีค่าต่ำสุดจะได้จุดใหม่

$$\mathbf{x}_{i+1} = \mathbf{x}_i + t_k\mathbf{d}_i$$

2.2 งานวิจัยที่เกี่ยวข้อง

ปัญหากำหนดการเชิงเส้น เป็นรูปแบบของปัญหาที่มีนักคณิตศาสตร์จำนวนมากให้ความสนใจศึกษาหาขั้นตอนวิธีใหม่ๆ เพื่อช่วยในการแก้ปัญหากำหนดการเชิงเส้นนอกเหนือจากวิธีซิมเพลกซ์

ในปี 1978 นักคณิตศาสตร์สองคนชื่อ ซี คาลาลัมบัส (C. Charalambous) และ เอ อาร์ โคนน์ (A.R. Conn) ได้เสนอขั้นตอนวิธีใหม่ขึ้นเพื่อแก้ปัญหามินิแมกซ์ (ค่าต่ำสุดของค่าสูงสุด) ของฟังก์ชันไม่เชิงเส้นเรื่อง " An efficient method to solve the minimax problem directly. " [4] โดยปัญหาการหาค่าเหมาะที่สุดของปัญหาแบบมินิแมกซ์อยู่ในลักษณะ

$$\min_{\mathbf{x}} \quad M_f(\mathbf{x})$$

เมื่อ

$$M_f(\mathbf{x}) = \max_{1 \le i \le m} f_i(\mathbf{x})$$

และ $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ ฟังก์ชันจุดประสงค์(objective function) มีอนุพันธ์อันดับหนึ่งไม่ต่อเนื่องที่จุดซึ่งมีฟังก์ชันมากกว่าหรือเท่ากับสองฟังก์ชันของฟังก์ชัน f_i มีค่าเท่ากันในกรณีที่ $f_i(\mathbf{x})$

เมื่อ $1 \le i \le m$ มีอนุพันธ์อันดับหนึ่งต่อเนื่อง จะใช้วิธีความชัน (gradient method) เพื่อหาค่าต่ำสุดของ $M_f(\mathbf{x})$ได้โดยตรง

ในขั้นตอนวิธีของงานวิจัยดังกล่าวใช้ 2 วิธีในการหาทิศทางแตกต่างกัน วิธีแรกหาทิศทางในแนวนอน (horizontal direction) คือพยายามที่จะลดค่า $M_f(\mathbf{x})$ ลง ในขณะที่พยายามที่จะรักษาค่าของฟังก์ชันให้เข้าใกล้ $M_f(\mathbf{x})$ วิธีที่สอง หาทิศทางในแนวตั้ง (vertical direction) คือพยายามที่จะลดค่าคลาดเคลื่อนภายในฟังก์ชันและใช้วิธีการหาแบบเชิงเส้น (linear search) หลังจากหาทิศทางในแนวนอนแล้ววิธีดังกล่าวนี้เป็นขั้นตอนวิธีที่นำมาใช้แก้ปัญหากำหนดการเชิงเส้นได้

ขั้นตอนวิธีของคาลาลัมบัสและคณะดังที่กล่าวมานี้เป็นการแก้ปัญหามินิแมกซ์โดยวิธีตรง ต่อมาในปี 1992 นักคณิตศาสตร์ชื่อ เอ วาดี (A.Vardi) ได้เสนองานวิจัยเรื่อง "New Minimax Algorithm." [5] โดยแก้ปัญหามินิแมกซ์

$$\min_{\mathbf{x}} \max_{i} \quad f_i(\mathbf{x}), \qquad \mathbf{x} \in R^n, \quad i = 1,2,\ldots,m$$

เมื่อ f_i, $i = 1,2,\ldots,m$ เป็นฟังก์ชันจำนวนจริงที่กำหนดขึ้นบน R^n โดยการแปลงปัญหาดังกล่าวเป็นปัญหาการหาค่าต่ำสุดภายใต้เงื่อนไขของอสมการในรูปแบบ

$$\begin{aligned} &\text{minimize} && t \\ &\text{subject to} && f_i(\mathbf{x}) - t \le 0 \end{aligned}$$

สำหรับ $\forall i, i = 1,2,\ldots,m$

ในปี 1991 เจ อาร์ ราชเสกขรา (J.R. Rajasekera) และ เอส ซี ฟาง (S.C. Fang) ได้เสนองานวิจัยเรื่อง "On the convex programming approach to linear programming." [6] โดยแปลงปัญหาที่เป็นรูปแบบมาตรฐานของคาร์มาร์คาร์ไปเป็นปัญหาคู่กันที่อยู่ในรูปแบบ

$$\begin{aligned} &\text{maximize} && w_{m+1} \\ &\text{subject to} && \sum a_{ij} w_i + w_{i+1} \le c_j,\ j = 1,2,\ldots,n \\ & && w_i \in R, i = 1,2,\ldots,m+1 \end{aligned}$$

เมื่อ $w_i \in R$ สำหรับ $i = 1,2,\ldots,m$

และ

$$w_{m+1} = \min_{j=1,2,\ldots,n} \left(c_j - \sum_{i=1}^{m} a_{ij} w_i \right)$$

เป็นผลเฉลยที่เป็นไปได้ของปัญหาคู่กันและใช้ ฟังก์ชัน entropy-based barried ในการหาเงื่อนไขของปัญหาและผลเฉลย ε-optimal ในการแก้ปัญหา

ต่อมาในปี 1992 เจ อาร์ ราชเสกขรา และ เอส ซี ฟาง ได้นำงานวิจัย [6] มาปรับปรุงเพิ่มเติมและเสนองานวิจัยเรื่อง "Deriving an Unconstrained Convex Program for Linear Programming." [7] โดยพิจารณาปัญหา

$$
(P_{\bar{\mu}}) \qquad \begin{aligned} \text{minimize} \quad & c^T\mathbf{x} + \bar{\mu}\sum_{j=1}^{n}\mathbf{x}_j \log \mathbf{x}_j \\ \text{subject to} \quad & A^T\mathbf{x} = 0 \\ & e^T\mathbf{x} = 1 \\ & \mathbf{x} \geq 0 \end{aligned}
$$

ในปี 1995 เอส เค ซู (Z.K. Xu.) และ เอส ซี ฟาง ได้เสนองานวิจัย เรื่อง "Unconstrained Convex Programming Approach to Linear Programming." [8] โดยการประมาณค่ากำหนดการเชิงเส้นในรูปแบบมาตรฐานของคาร์มาร์คาร์โดยเพิ่มฟังก์ชัน entropic barrier ไปที่ฟังก์ชันจุดประสงค์

บทที่ 3
วิธีการดำเนินการ

ในบทนี้เราจะกล่าวถึงทฤษฎีบทที่เป็นพื้นฐานของขั้นตอนวิธีที่เราพัฒนาขึ้น คือ ทฤษฎีบทที่กล่าวเกี่ยวกับ ซับเกรเดียนต์ เพื่อที่จะอธิบายถึงทิศทางลดและทิศทางลดเร็วที่สุดที่จุดที่ฟังก์ชันไม่มีอนุพันธ์ ทฤษฎีบทต่างๆเหล่านี้มีทั้งส่วนที่เป็นทฤษฎีของเดิมที่ทราบกันดีอยู่แล้วและมีบางทฤษฎีบทที่ผู้วิจัยได้เพิ่มเติมขึ้น ผู้วิจัยจะแสดงการพิสูจน์ทฤษฎีบทที่ไม่ปรากฏอยู่ในตำราที่อ้างอิงและทฤษฎีบทที่ผู้วิจัยพัฒนาขึ้น สำหรับทฤษฎีบทที่มีอยู่ในตำราที่แพร่หลายแล้วจะไม่แสดงการพิสูจน์

3.1 เซตนูนและฟังก์ชันนูน (convex set and convex function)

ในหัวข้อนี้เรากล่าวถึงเซตนูนและฟังก์ชันนูนซึ่งจำเป็นต้องใช้ในเรื่องการหาค่าต่ำสุด

นิยาม 3.1.1 เซต $C \subseteq R^n$ เป็น *เซตนูน* (convex set) ก็ต่อเมื่อ

$$\lambda \mathbf{x} + (1-\lambda)\mathbf{y} \in C$$

สำหรับทุก $\mathbf{x}$ และ $\mathbf{y} \in C$ และ λ ซึ่ง $0 \le \lambda \le 1$

ภาพที่ 3.1 เซตนูน

ทฤษฎีบท 3.1.1 ให้ S เป็นเซตนูนและเป็นเซตปิด และ $\mathbf{y}$ เป็นเวกเตอร์ที่ไม่อยู่ใน S จะมีจุด $\bar{\mathbf{x}}$ ใน S ที่ทำให้ $d(\bar{\mathbf{x}}, \mathbf{y})$ เป็นระยะสั้นที่สุด หรือกล่าวว่า มีจุด $\bar{\mathbf{x}}$ ใน S ที่ทำให้

$$(\mathbf{y}-\bar{\mathbf{x}})^T(\mathbf{s}-\bar{\mathbf{x}}) \le 0$$

สำหรับทุก ๆ $\mathbf{s} \in S$

ถ้าให้ $\xi = \mathbf{y} - \bar{\mathbf{x}}$ และ $\xi^T \bar{\mathbf{x}} = \alpha$ จะได้ว่า $\xi^T \mathbf{s} \le \alpha$ ทุก ๆ $\mathbf{s} \in S$

นิยาม 3.1.2 ให้ $f : S \to R$ เมื่อ S เป็นเซตไม่ว่างใน R^n เรากล่าวว่าฟังก์ชัน f เป็น***ฟังก์ชันนูน*** (convex function) บน S ถ้า

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \le \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

สำหรับแต่ละ $\mathbf{x}, \mathbf{y} \in S$ และ $\lambda \in (0, 1)$ (คือ $0 < \lambda < 1$)

ฟังก์ชันนูน ไม่ใช่ฟังก์ชันนูน

ภาพที่ 3.2 ฟังก์ชันนูน

ทฤษฎีบท 3.1.2 ถ้า f เป็นฟังก์ชันนูนแล้ว f เป็นฟังก์ชันต่อเนื่องที่จุดภายในของ S ทุกจุด

ถ้า f เป็นฟังก์ชันนูน จะได้สมบัติต่าง ๆ เกี่ยวกับอนุพันธ์ทิศทาง ดังนี้

ทฤษฎีบท 3.1.3 กำหนดให้ f เป็นฟังก์ชันนูน ที่กำหนดค่าในเซตนูน S และ $\bar{\mathbf{x}} \in S$

1. ถ้า $\mathbf{u}$ เป็นเวกเตอร์ใด ๆ ที่ไม่ใช่เวกเตอร์ศูนย์ และ $\bar{\mathbf{x}}$ เป็นจุดภายในของ S กำหนดให้

$$g(\lambda) = \frac{f(\bar{\mathbf{x}} + \lambda \mathbf{u}) - f(\bar{\mathbf{x}})}{\lambda}$$

เมื่อ $\lambda > 0$ แล้ว ฟังก์ชัน g เป็นฟังก์ชันเพิ่ม และมีขอบเขตด้านล่าง

2. ถ้า $\bar{\mathbf{x}}$ เป็นจุดภายในของ S อนุพันธ์ทิศทาง $f'(\bar{\mathbf{x}}; \mathbf{u})$ มีเกิดขึ้นสำหรับทุกเวกเตอร์ $\mathbf{u}$
3. $f'(\bar{\mathbf{x}}; \mathbf{u}) \le \dfrac{f(\bar{\mathbf{x}} + \lambda \mathbf{u}) - f(\bar{\mathbf{x}})}{\lambda}$ สำหรับทุกเวกเตอร์ $\mathbf{u}$ และจำนวน $\lambda > 0$
4. $f'(\bar{\mathbf{x}}; k\mathbf{u}) = k\, f'(\bar{\mathbf{x}}; \mathbf{u})$ สำหรับทุกเวกเตอร์ $\mathbf{u}$ และจำนวน $k > 0$
5. อนุพันธ์ทิศทางของฟังก์ชันนูนจะเป็นฟังก์ชันนูน กล่าวคือ

$$f'(\bar{\mathbf{x}}; t\mathbf{u}+(1-t)\mathbf{v}) = t f'(\bar{\mathbf{x}};\mathbf{u}) + (1-t) f'(\bar{\mathbf{x}};\mathbf{v})$$

เมื่อ $0<t<1$

พิสูจน์ 1. จะแสดงว่า ถ้า $0<\lambda_1<\lambda_2$ แล้ว

$$\frac{f(\bar{\mathbf{x}}+\lambda_1\mathbf{u})-f(\bar{\mathbf{x}})}{\lambda_1} \le \frac{f(\bar{\mathbf{x}}+\lambda_2\mathbf{u})-f(\bar{\mathbf{x}})}{\lambda_2}$$

จาก

$$\bar{\mathbf{x}}+\lambda_1\mathbf{u} = \left(1-\frac{\lambda_1}{\lambda_2}\right)\bar{\mathbf{x}}+\frac{\lambda_1}{\lambda_2}(\bar{\mathbf{x}}+\lambda_2\mathbf{u})$$

ดังนั้น โดยที่ f เป็นฟังก์ชันนูน จะได้

$$f(\bar{\mathbf{x}}+\lambda_1\mathbf{u}) \le \left(1-\frac{\lambda_1}{\lambda_2}\right)f(\bar{\mathbf{x}})+\frac{\lambda_1}{\lambda_2}f(\bar{\mathbf{x}}+\lambda_2\mathbf{u})$$

$$f(\bar{\mathbf{x}}+\lambda_1\mathbf{u})-f(\bar{\mathbf{x}}) \le \frac{\lambda_1}{\lambda_2}\left[f(\bar{\mathbf{x}}+\lambda_2\mathbf{u})-f(\bar{\mathbf{x}})\right]$$

หารด้วย λ_1 จะได้ตามต้องการ แสดงว่า g เป็นฟังก์ชันเพิ่ม

เมื่อ λ มีค่าน้อยลง ค่าของ $g(\lambda)$ จะมีค่าน้อยลงด้วย จะแสดงว่า $g(\lambda)\ge k$ เมื่อ k เป็นจำนวนที่ $k>-\infty$

เนื่องจาก $\bar{\mathbf{x}}$ เป็นจุดภายในของ S ดังนั้น จะมีค่า $\delta>0$ ที่ทำให้ฟังก์ชัน $f(\bar{\mathbf{x}}+\lambda\mathbf{u})$ และ $f(\bar{\mathbf{x}}-\lambda\mathbf{u})$ กำหนดค่า เมื่อ $\lambda\in(0,\delta)$

จาก

$$\bar{\mathbf{x}} = \frac{1}{2}(\bar{\mathbf{x}}-\lambda\mathbf{u})+\frac{1}{2}(\bar{\mathbf{x}}+\lambda\mathbf{u})$$

ดังนั้น

$$f(\bar{\mathbf{x}}) \le \frac{1}{2}f(\bar{\mathbf{x}}-\lambda\mathbf{u})+\frac{1}{2}f(\bar{\mathbf{x}}+\lambda\mathbf{u})$$

จะได้

$$-\frac{f(\bar{\mathbf{x}}+\lambda(-\mathbf{u}))-f(\bar{\mathbf{x}})}{\lambda} \le \frac{f(\bar{\mathbf{x}}+\lambda\mathbf{u})-f(\bar{\mathbf{x}})}{\lambda}$$

นิพจน์ทางซ้ายมือมีค่าเพิ่มขึ้นเมื่อ λ มีค่าน้อยลง

ดังนั้นให้

$$k = -\frac{f(\bar{\mathbf{x}}+\delta(-\mathbf{u}))-f(\bar{\mathbf{x}})}{\delta}$$

จะได้ว่า $g(\lambda)\geq k$ เมื่อ k เป็นจำนวนที่ $k>-\infty$ ตามต้องการ ##

พิสูจน์ 2. จากข้อหนึ่ง จะได้ว่า $\lim_{\lambda\to 0^+} g(\lambda)$ มีค่าเกิดขึ้น ซึ่งจะเป็นค่า $f'(\bar{\mathbf{x}};\lambda\mathbf{u})$ นั่นเอง

พิสูจน์ 3. เห็นได้ชัดจากข้อ 1.

พิสูจน์ 4. จากนิยาม

$$f'(\bar{\mathbf{x}};k\mathbf{u}) = \lim_{\lambda\to 0^+}\frac{f(\bar{\mathbf{x}}+\lambda k\mathbf{u})-f(\bar{\mathbf{x}})}{\lambda}$$

$$= \lim_{k\lambda\to 0^+} k\frac{f(\bar{\mathbf{x}}+\lambda k\mathbf{u})-f(\bar{\mathbf{x}})}{k\lambda}$$

$$f'(\bar{\mathbf{x}};k\mathbf{u}) = k\,f'(\bar{\mathbf{x}};\mathbf{u})$$ ##

พิสูจน์ 5. จากนิยาม

$$f'(\bar{\mathbf{x}};t\mathbf{u}+(1-t)\mathbf{v}) = \lim_{\lambda\to 0^+}\frac{f(\bar{\mathbf{x}}+\lambda(t\mathbf{u}+(1-t)\mathbf{v})-f(\bar{\mathbf{x}})}{\lambda}$$

$$\leq \lim_{\lambda\to 0^+} t\frac{f(\bar{\mathbf{x}}+\lambda\mathbf{u})-f(\bar{\mathbf{x}})}{\lambda}+\lim_{\lambda\to 0^+}(1-t)\frac{f(\bar{\mathbf{x}}+\lambda v)-f(\bar{\mathbf{x}})}{\lambda}$$

$$= t\,f'(\bar{\mathbf{x}};\mathbf{u}) + (1-t)\,f'(\bar{\mathbf{x}};\mathbf{v})$$ ##

นิยาม 3.1.3 ให้ S เป็นเซตนูนใน R^n และ $f: S\to R^n$ เป็นฟังก์ชันนูนแล้ว เรากล่าวว่าเวกเตอร์ ξ เป็น *ซับเกรเดียนต์* (subgradient) ของ f ที่จุด $\bar{\mathbf{x}}\in S$ ก็ต่อเมื่อ

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^T(\mathbf{x}-\bar{\mathbf{x}})$$

สำหรับทุก $\mathbf{x}\in S$

เราใช้สัญลักษณ์ $\partial f(\bar{\mathbf{x}})$ (หรือ ∂f) แทนเซตของซับเกรเดียนต์ของฟังก์ชัน f ที่จุด $\bar{\mathbf{x}}$ สำหรับฟังก์ชันนูนจะมีซับเกรเดียนต์ที่ทุกจุด

ทุกจุด $\mathbf{x} \in S$

นั่นคือ $\lambda\xi_1 + (1-\lambda)\xi_2$ เป็นซับเกรเดียนต์ ##

บทแทรก ถ้า $\xi_1, \xi_2, \ldots, \xi_n$ เป็นซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$ และ $\xi = \sum_{i=1}^{n} \lambda_i \xi_i$ เมื่อ $\lambda_i > 0$, $i = 1, 2, \ldots, n$ แล้วผลรวมเชิงนูนของ $\xi_1, \xi_2, \ldots, \xi_n$ จะเป็นซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$ ด้วย

ทฤษฎีบท 3.1.7 ให้ f เป็นฟังก์ชันนูนและ ∂f เป็นเซตของซับเกรเดียนต์ทั้งหมดของ f ที่จุด $\bar{\mathbf{x}}$ แล้ว ∂f เป็นเซตปิด

พิสูจน์ ให้ ξ_1 เป็นเวกเตอร์ซึ่งสำหรับทุก $\varepsilon > 0$ จะมี $\xi \in \partial f$ ซึ่ง $\|\xi - \xi_1\| < \varepsilon$

เราต้องการแสดงว่า $\xi_1 \in \partial f$

สมมติว่า $\xi_1 \notin \partial f$ ดังนั้นจะมีเวกเตอร์หน่วย $\mathbf{u}$ ซึ่งทำให้

$$\xi^T \mathbf{u} \le \frac{f(\bar{\mathbf{x}} + \lambda_1 \mathbf{u}) - f(\bar{\mathbf{x}})}{\lambda_1} < \xi_1^T \mathbf{u}$$

สำหรับบางค่าของ $\lambda_1 > 0$ และสำหรับทุกซับเกรเดียนต์ $\xi \in \partial f$ (โดยนิยามของ ξ)

จาก f เป็นฟังก์ชันนูน โดยทฤษฎีบท 3.1.3 สำหรับทุก $\lambda \in (0, \lambda_1)$ จะได้

$$\xi^T \mathbf{u} \le \frac{f(\bar{\mathbf{x}} + \lambda \mathbf{u}) - f(\bar{\mathbf{x}})}{\lambda} < \xi_1^T \mathbf{u}$$

ดังนั้นเมื่อ $\lambda \to 0^+$ เราได้

$$\xi^T \mathbf{u} \le f'(\bar{\mathbf{x}}, \mathbf{u}) < \xi_1^T \mathbf{u}$$

ให้

$$c = \xi_1^T \mathbf{u} - f'(\bar{\mathbf{x}}; \mathbf{u}) > 0$$

ซึ่งจะได้

$$\xi_1^T \mathbf{u} - \xi^T \mathbf{u} \ge c \qquad \ldots(3.1)$$

ทุกๆ $\xi \in \partial f$

ให้ $\varepsilon = c$ จากที่กำหนดให้ จะมีเวกเตอร์ $\xi \in \partial f$ ซึ่งทำให้

$$\|\xi_1 - \xi\| < \varepsilon$$

นั่นคือ

$$|(\xi_1 - \xi)\mathbf{u}| < \|\xi_1 - \xi\| < \varepsilon$$

ซึ่งขัดแย้งกับ (3.1)
ดังนั้น

$$\xi_1\mathbf{u} \le \frac{f(\bar{\mathbf{x}} + \lambda\mathbf{u}) - f(\bar{\mathbf{x}})}{\lambda}$$

สำหรับทุกเวกเตอร์หน่วย $\mathbf{u}$
เราได้ว่า $\xi_1 \in \partial f$ เพราะฉะนั้น ∂f เป็นเซตปิด ##

ทฤษฎีบท 3.1.8 ให้ f เป็นฟังก์ชันนูนที่กำหนดในเซตนูน S และให้ ∂f เป็นเซตของซับเกรเดียนต์ทั้งหมดของ f ที่จุด $\bar{\mathbf{x}}$ ถ้า $\bar{\mathbf{x}}$ เป็นจุดภายในของ S แล้ว ∂f เป็นเซตที่มีขอบเขต
พิสูจน์ ถ้า $\mathbf{u}$ เป็นเวกเตอร์หน่วย จากนิยามของซับเกรเดียนต์ ξ เราจะได้ว่าสำหรับทุก $\xi \in \partial f$

$$\xi^T\mathbf{u} \le \frac{f(\bar{\mathbf{x}} + \lambda\mathbf{u}) - f(\bar{\mathbf{x}})}{\lambda}$$

ทุกๆค่า $\lambda \in (0, \delta)$ สำหรับ $\delta > 0$ บางจำนวน ดังนั้น สำหรับทุก $\xi \in \partial f$ จะได้

$$\xi^T\mathbf{u} \le f'(\bar{\mathbf{x}}; \mathbf{u})$$

จากทฤษฎีบท 3.1.3 ข้อ 2 อนุพันธ์ทิศทางมีค่าจำกัด จะมีจำนวนคงตัว K ที่

$$f'(\bar{\mathbf{x}}; \mathbf{u}) \le K$$

สำหรับทุก ๆ เวกเตอร์หน่วย $\mathbf{u}$ ดังนั้น

$$\xi^T\mathbf{u} \le K$$

สำหรับทุก $\xi \in \partial f$ และทุก ๆ เวกเตอร์หน่วย $\mathbf{u}$
นั่นคือ ถ้าให้ $\mathbf{u}$ เป็นเวกเตอร์หน่วยทิศเดียวกับ ξ จะได้

$$\|\xi\| = \xi^T \mathbf{u} \quad \le K$$

ซึ่งเป็นจริงสำหรับทุก $\xi \in \partial f$

นั่นคือ ∂f เป็นเซตที่มีขอบเขต ##

เรามีคุณสมบัติที่สำคัญเป็นความสัมพันธ์ของซับเกรเดียนต์กับอนุพันธ์ทิศทางคือที่จุดภายใน $\bar{\mathbf{x}} \in S$ จะมีซับเกรเดียนต์ ξ ที่ทำให้

$$\xi^T \mathbf{d} = f'(\bar{\mathbf{x}};\mathbf{d})$$

ทฤษฎีบท 3.1.9 ให้ f เป็นฟังก์ชันนูนและ ∂f เป็นเซตของซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$ ถ้า $\mathbf{d}$ เป็นเวกเตอร์หน่วยใดๆแล้ว

$$\sup_{\xi \in \partial f} \left\{\xi^T \mathbf{d}\right\} \le f'(\bar{\mathbf{x}};\mathbf{d})$$

พิสูจน์ จากนิยามของซับเกรเดียนต์ ξ เราจะได้ว่าสำหรับทุก $\xi \in \partial f$

$$\xi^T \mathbf{d} \le \frac{f(\bar{\mathbf{x}} + \lambda \mathbf{d}) - f(\bar{\mathbf{x}})}{\lambda}$$

ทุกๆค่า $\lambda \in (0,\delta)$ สำหรับ $\delta > 0$ บางจำนวน

นั่นคือค่าสูงสุดของ $\xi^T \mathbf{d}$ จะไม่เกินค่า $f'(\bar{\mathbf{x}};\mathbf{d})$

ดังนั้น

$$\sup_{\xi \in \partial f} \left\{\xi^T \mathbf{d}\right\} \le f'(\bar{\mathbf{x}};\mathbf{d}) \qquad \#\#$$

ทฤษฎีบท 3.1.10 ให้ f เป็นฟังก์ชันนูนและ ∂f เป็นเซตของซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$ ถ้า $\bar{\mathbf{x}}$ เป็นจุดภายในของ S และ $\mathbf{d}$ เป็นเวกเตอร์หน่วยใดๆแล้ว จะมีซับเกรเดียนต์ ξ ของ f ที่จุด $\bar{\mathbf{x}}$ ที่ทำให้ $\xi^T \mathbf{d} = f'(\bar{\mathbf{x}};\mathbf{d})$ กล่าวคือ

$$\max_{\xi \in \partial f} \left\{\xi^T \mathbf{d}\right\} = f'(\bar{\mathbf{x}};\mathbf{d})$$

พิสูจน์ จากที่ $\bar{\mathbf{x}}$ เป็นจุดภายในของ S ดังนั้นมีจำนวน $\delta > 0$ ที่ทำให้ $\bar{\mathbf{x}} + \lambda \mathbf{d}$ เป็นจุดภายในของ S ทุกค่าของ λ เมื่อ $0 < \lambda < \delta$

ให้

$$\mathbf{y} = \bar{\mathbf{x}} + \lambda \mathbf{d}$$

และให้ $\xi_{\mathbf{y}}$ เป็นซับเกรเดียนต์ของ f ที่จุด $\mathbf{y}$ จะได้

$$\xi_{\mathbf{y}}^{T}(-\mathbf{d}) \leq \frac{f(\mathbf{y}-\lambda\mathbf{d})-f(\mathbf{y})}{\lambda}$$

$$-\xi_{\mathbf{y}}^{T}\mathbf{d} \leq \frac{f(\bar{\mathbf{x}})-f(\bar{\mathbf{x}}+\lambda\mathbf{d})}{\lambda}$$

ดังนั้น

$$\xi_{\mathbf{y}}^{T}\mathbf{d} \geq \frac{f(\bar{\mathbf{x}}+\lambda\mathbf{d})-f(\bar{\mathbf{x}})}{\lambda}$$

และจาก

$$\xi_{\mathbf{y}}^{T}\mathbf{d} \leq \frac{f(\mathbf{y}+\lambda\mathbf{d})-f(\mathbf{y})}{\lambda}$$

$$= \frac{f(\bar{\mathbf{x}}+2\lambda\mathbf{d})-f(\bar{\mathbf{x}}+\lambda\mathbf{d})}{\lambda}$$

$$= 2\frac{f(\bar{\mathbf{x}}+2\lambda\mathbf{d})-f(\bar{\mathbf{x}})}{2\lambda} - \frac{f(\bar{\mathbf{x}}+\lambda\mathbf{d})-f(\bar{\mathbf{x}})}{\lambda}$$

ดังนั้น

$$\frac{f(\bar{\mathbf{x}}+\lambda\mathbf{d})-f(\bar{\mathbf{x}})}{\lambda} \leq \xi_{\mathbf{y}}^{T}\mathbf{d} \leq 2\frac{f(\bar{\mathbf{x}}+2\lambda\mathbf{d})-f(\bar{\mathbf{x}})}{2\lambda} - \frac{f(\bar{\mathbf{x}}+\lambda\mathbf{d})-f(\bar{\mathbf{x}})}{\lambda}$$

จะเห็นว่า เมื่อ $\lambda \to 0^{+}$ จะได้ว่า $\mathbf{y} \to \bar{\mathbf{x}}$ และ $\xi_{\mathbf{y}}^{T}\mathbf{d} \to f'(\bar{\mathbf{x}};\mathbf{d})$

ให้ $\partial f(\mathbf{y})$ เป็นเซตของซับเกรเดียนต์ของ f ที่จุด $\mathbf{y}$ เนื่องจาก $\mathbf{y}$ เป็นจุดภายในของ S ดังนั้น $\partial f(\mathbf{y})$ เป็นเซตที่มีขอบเขต คือ มีจำนวนคงตัว M ที่ทำให้ $\|\xi_{\mathbf{y}}\| < M$ ทุก ๆ ซับเกรเดียนต์ $\xi_{\mathbf{y}}$ และทุก ๆ เวกเตอร์ $\mathbf{y}$ ดังนั้นลำดับ $\{\xi_{\mathbf{y}}^{k}\}$ เป็นลำดับที่มีขอบเขต โดยทฤษฎีบทของโบลซาโน-ไวแยร์สตราสส์ (Bolzano-Weierstrass) จะมีลำดับย่อยของซับเกรเดียนต์ $\{\xi_{\mathbf{y}}^{k}\}$ ที่ลู่เข้าหาเวกเตอร์ ξ ที่มีสมบัติว่า

$$(\xi_{\mathbf{y}}^{k})^{T}\mathbf{d} \to \xi^{T}\mathbf{d} = f'(\bar{\mathbf{x}};\mathbf{d})$$

ต่อไปจะแสดงว่า ξ เป็นซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$

สมมุติว่า ξ ไม่เป็นซับเกรเดียนต์ คือมีเวกเตอร์หน่วย $\mathbf{u}$ และจำนวน $\lambda > 0$ ที่ทำให้

$$\xi^{T}\mathbf{u} > \frac{f(\bar{\mathbf{x}}+\lambda\mathbf{u})-f(\bar{\mathbf{x}})}{\lambda}$$

ให้

$$\xi^T \mathbf{u} = \frac{f(\overline{\mathbf{x}}+\lambda\mathbf{u})-f(\overline{\mathbf{x}})}{\lambda}+k$$

เมื่อ $k>0$

ให้ $\mathbf{y} = \overline{\mathbf{x}} + \mu\mathbf{d}$ และให้ $\xi_{\mathbf{y}}$ เป็นซับเกรเดียนต์ของ f ที่จุด $\mathbf{y}$ จะได้

$$-\xi_{\mathbf{y}}^T\mathbf{u} \geq -\frac{f(\mathbf{y}+\lambda\mathbf{u})-f(\mathbf{y})}{\lambda}$$

ดังนั้น

$$\xi^T\mathbf{u}-\xi_{\mathbf{y}}^T\mathbf{u} \geq \frac{f(\overline{\mathbf{x}}+\lambda\mathbf{u})-f(\overline{\mathbf{x}})}{\lambda}-\frac{f(\mathbf{y}+\lambda\mathbf{u})-f(\mathbf{y})}{\lambda}+k$$

เราได้

$$\xi^T\mathbf{u}-\xi_{\mathbf{y}}^T\mathbf{u} = \frac{1}{\lambda}[(f(\overline{\mathbf{x}}+\lambda\mathbf{u})-f(\mathbf{y}+\lambda\mathbf{u}))-(f(\overline{\mathbf{x}})-f(\mathbf{y}))]+k$$

$$\geq \frac{1}{\lambda}[-|f(\overline{\mathbf{x}}+\lambda\mathbf{u})-f(\mathbf{y}+\lambda\mathbf{u})|-|f(\overline{\mathbf{x}})-f(\mathbf{y})|]+k$$

เนื่องจาก f เป็นฟังก์ชันนูน ดังนั้น f เป็นฟังก์ชันต่อเนื่อง จะมี $\delta>0$ ที่ทำให้

$$|f(\overline{\mathbf{x}}+\lambda\mathbf{u})-f(\mathbf{y}+\lambda\mathbf{u})| < \frac{k\lambda}{3}$$

และ

$$|f(\overline{\mathbf{x}})-f(\mathbf{y})| < \frac{k\lambda}{3}$$

เมื่อ $|\overline{\mathbf{x}}-\mathbf{y}|<\delta$

ดังนั้น

$$\xi^T\mathbf{u}-\xi_{\mathbf{y}}^T\mathbf{u} \geq \frac{1}{\lambda}\left[-\frac{k\lambda}{3}-\frac{k\lambda}{3}\right]+k = \frac{k}{3} \qquad ...(3.2)$$

และเนื่องจาก $\xi_{\mathbf{y}} \to \xi$ เมื่อ $\mathbf{y} \to \overline{\mathbf{x}}$ คือ มี $\delta_1 > 0$ ที่ทำให้

$$\|\xi_{\mathbf{y}}-\xi\| < \frac{k}{3}$$

เมื่อ $|\overline{\mathbf{x}}-\mathbf{y}|<\delta_1$

เราได้ว่า

$$\left|(\xi_{\mathbf{y}}-\xi)^T\mathbf{u}\right| \leq \|\xi_{\mathbf{y}}-\xi\|\|\mathbf{u}\| < \frac{k}{3} \qquad ...(3.3)$$

ดังนั้น เมื่อ $|\overline{\mathbf{x}}-\mathbf{y}| = \alpha < \min\{\delta, \delta_1\}$ จะได้ (3.2) และ (3.3) ซึ่งขัดแย้งกัน
นั่นคือ

$$\xi^T \mathbf{u} \le \frac{f(\overline{\mathbf{x}}+\lambda\mathbf{u})-f(\overline{\mathbf{x}})}{\lambda}$$

ทุกเวกเตอร์หน่วย $\mathbf{u}$ และจำนวน $\lambda > 0$
ดังนั้น ξ เป็นซับเกรเดียนของ f ที่จุด $\overline{\mathbf{x}}$
นั่นคือ มีซับเกรเดียนต์ ξ ของ f ที่จุด $\overline{\mathbf{x}}$ ที่ทำให้ $\xi^T \mathbf{d} = f'(\overline{\mathbf{x}};\mathbf{d})$ และจากทฤษฎีบท 3.1.9
เราได้

$$\max_{\xi\in\partial f} \left\{\xi^T \mathbf{d}\right\} = f'\left(\overline{\mathbf{x}};\mathbf{d}\right)$$ ##

3.2 ทิศทางลด

ในหัวข้อนี้จะกล่าวถึงทิศทางลดซึ่งจะใช้ในขั้นตอนวิธีของการหาค่าต่ำสุด

นิยาม 3.2.1 ให้ f เป็นฟังก์ชันที่กำหนดค่าบน R^n และ $\mathbf{d}$ เป็นเวกเตอร์ใน R^n เรากล่าวว่า $\mathbf{d}$ เป็นทิศทางลด ของ f ที่จุด $\overline{\mathbf{x}}$ ใน R^n ถ้ามี $\delta > 0$ และ $f(\overline{\mathbf{x}}+\lambda\mathbf{d}) < f(\overline{\mathbf{x}})$ สำหรับทุกค่า $\lambda \in (0,\delta)$

ทฤษฎีบท 3.2.1 ถ้า f เป็นฟังก์ชันที่กำหนดค่าบน R^n และ $\mathbf{d}$ เป็นเวกเตอร์ใน R^n จะได้ว่า $\mathbf{d}$ เป็นทิศทางลดของ f ที่จุด $\overline{\mathbf{x}}$ ก็ต่อเมื่อ $f'(\overline{\mathbf{x}};\mathbf{d}) < 0$ ถ้าอนุพันธ์ทิศทางมีค่าเกิดขึ้น

พิสูจน์ ให้ $\mathbf{d}$ เป็นทิศทางลด จะมี δ ที่ทำให้
เมื่อ $0 < \lambda < \delta$ แล้ว

$$f(\overline{\mathbf{x}}+\lambda\mathbf{d}) < f(\overline{\mathbf{x}})$$

นั่นคือ $0 < \lambda < \delta$ แล้ว

$$\frac{f(\overline{\mathbf{x}}+\lambda\mathbf{d})-f(\overline{\mathbf{x}})}{\lambda} < 0$$

เราได้

$$f'(\overline{\mathbf{x}};\mathbf{d}) = \lim_{\lambda\to 0^+} \frac{f(\overline{\mathbf{x}}+\lambda\mathbf{d})-f(\overline{\mathbf{x}})}{\lambda} < 0$$

การพิสูจน์ทางกลับ จากนิยาม

$$f'(\overline{\mathbf{x}};\mathbf{d}) = \lim_{\lambda\to 0^+} \frac{f(\overline{\mathbf{x}}+\lambda\mathbf{d})-f(\overline{\mathbf{x}})}{\lambda}$$

ถ้า $f'(\overline{\mathbf{x}};\mathbf{d}) < 0$ แสดงว่ามี $\delta > 0$ ที่ทำให้

$$\frac{f(\overline{\mathbf{x}}+\lambda\mathbf{d})-f(\overline{\mathbf{x}})}{\lambda} < 0$$

เมื่อ $0<\lambda<\delta$
ดังนั้น

$$f(\bar{x}+\lambda d)<f(\bar{x})$$

นั่นคือ d เป็นทิศทางลด ##

จากทฤษฎีบท 3.2.1 เราเขียนได้อีกอย่างหนึ่งว่า สำหรับฟังก์ชันนูน f จุด $\bar{x}$ จะเป็นจุดต่ำสุดก็ต่อเมื่อ $f'(\bar{x};d)\geq 0$ สำหรับทุกเวกเตอร์หน่วย d

ทฤษฎีบท 3.2.2 ให้ f เป็นฟังก์ชันนูน เวกเตอร์ d จะเป็นทิศทางลดของ f ที่จุด $\bar{x}$ ก็ต่อเมื่อ $d^T\xi<0$ สำหรับทุกซับเกรเดียนต์ ξ ของ f ที่จุด $\bar{x}$
พิสูจน์ ให้ d เป็นเวกเตอร์หน่วย และ ξ เป็นซับเกรเดียนต์
จากนิยามเราได้

$$f(\bar{x}+\lambda d)\geq f(\bar{x})+\lambda\xi^T d$$

ถ้า $d^T\xi\geq 0$ สำหรับบาง ξ จะได้

$$f(\bar{x}+\lambda d)-f(\bar{x})\geq 0$$

d ก็จะไม่เป็นทิศทางลด
ดังนั้น ถ้า d เป็นทิศทางลดแล้ว $d^T\xi<0$ ทุกซับเกรเดียนต์ ξ
การพิสูจน์ทางกลับ ให้ d เป็นเวกเตอร์ ซึ่ง

$$\xi^T d<0$$

สำหรับทุกซับเกรเดียนต์ ξ ของ f ที่ $\bar{x}$
ดังนั้น จากทฤษฎีบท 3.1.10 จะได้ว่า

$$f'(\bar{x};d)=\max_{\xi\in\partial f}\{\xi^T d\}<0$$

ดังนั้น d เป็นทิศทางลด ##

จากทฤษฎีบท 3.2.2 เรากล่าวได้อีกอย่างหนึ่งว่า สำหรับฟังก์ชันนูน f จุด $\bar{x}$ จะเป็นจุดต่ำสุดของ f ก็ต่อเมื่อ สำหรับทุกเวกเตอร์หน่วย d จะได้ $d^T\xi\geq 0$ สำหรับทุกซับเกรเดียนต์ ξ ของ f ที่จุด $\bar{x}$ และเราได้บทแทรก

บทแทรก สำหรับฟังก์ชันนูน f จุด $\bar{\mathbf{x}}$ จะเป็นจุดต่ำสุดของ f ก็ต่อเมื่อ มีเวกเตอร์ $\mathbf{0}$ เป็นซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$

พิสูจน์ ถ้า $\bar{\mathbf{x}}$ จะเป็นจุดต่ำสุดของ f จะได้ว่า สำหรับทุกเวกเตอร์หน่วย $\mathbf{d}$

เราได้

$$\mathbf{d}^T\xi \geq 0$$

สำหรับทุกซับเกรเดียนต์ ξ ของ f ที่จุด $\bar{\mathbf{x}}$

และเราได้ว่า

$$\mathbf{d}^T\mathbf{0} = 0 = \mathbf{d}^T\xi \leq f'(\bar{\mathbf{x}};\mathbf{d})$$

ทุกเวกเตอร์หน่วย $\mathbf{d}$

ดังนั้น $\mathbf{0}$ เป็นซับเกรเดียนต์

ในทางกลับ ถ้า $\mathbf{0}$ เป็นซับเกรเดียนต์

เราได้

$$\mathbf{d}^T\xi \leq 0$$

สำหรับทุกเวกเตอร์หน่วย $\mathbf{d}$

ดังนั้น $\bar{\mathbf{x}}$ จะเป็นจุดต่ำสุดของ f ##

นิยาม 3.2.2 ถ้า f เป็นฟังก์ชันที่กำหนดค่าบน R^n และ $\bar{\mathbf{d}}$ เป็นเวกเตอร์หน่วยใน R^n เรากล่าวว่า $\bar{\mathbf{d}}$ เป็นทิศทางลดเร็วที่สุดของ f ที่จุด $\bar{\mathbf{x}}$ ถ้ามีจำนวน $\delta > 0$ ที่ทำให้

$$f(\bar{\mathbf{x}}+\lambda\bar{\mathbf{d}}) \leq f(\bar{\mathbf{x}}+\lambda\mathbf{d})$$

สำหรับทุกเวกเตอร์หน่วย $\mathbf{d} \in R^n$ และสำหรับทุก $\lambda \in (0,\delta)$

ทฤษฎีบท 3.2.3 ให้ $\bar{\mathbf{d}}$ เป็นเวกเตอร์หน่วยแล้ว $\bar{\mathbf{d}}$ จะเป็นทิศทางลดเร็วที่สุดของฟังก์ชันนูน f ที่จุด $\bar{\mathbf{x}}$ ก็ต่อเมื่อ

$$f'(\bar{\mathbf{x}};\bar{\mathbf{d}}) \leq f'(\bar{\mathbf{x}};\mathbf{d})$$

สำหรับทุกเวกเตอร์หน่วย $\mathbf{d}$

พิสูจน์ ให้ $\bar{\mathbf{d}}$ เป็นทิศทางลดเร็วที่สุดที่จุด $\bar{\mathbf{x}}$ จากนิยามจะได้

$$f(\bar{\mathbf{x}}+\lambda\bar{\mathbf{d}}) \leq f(\bar{\mathbf{x}}+\lambda\mathbf{d})$$

ทุกๆเวกเตอร์หน่วย $\mathbf{d}$ และทุกค่า λ เมื่อ $0<\lambda<\delta$
ดังนั้น

$$\lim_{\lambda\to 0^+} \frac{f(\overline{\mathbf{x}}+\lambda\overline{\mathbf{d}})-f(\overline{\mathbf{x}})}{\lambda} \le \lim_{\lambda\to 0^+} \frac{f(\overline{\mathbf{x}}+\lambda\mathbf{d})-f(\overline{\mathbf{x}})}{\lambda}$$

เราได้

$$f'(\overline{\mathbf{x}};\overline{\mathbf{d}}) \le f'(\overline{\mathbf{x}};\mathbf{d})$$

การพิสูจน์ทางกลับ แสดงได้โดยเขียนกลับจากข้างบน ##

ทฤษฎีบท 3.2.4 ให้ $\overline{\mathbf{d}}$ เป็นเวกเตอร์หน่วยแล้ว $\overline{\mathbf{d}}$ จะเป็นทิศทางลดเร็วที่สุดของฟังก์ชันนูน f ที่จุด $\overline{\mathbf{x}}$ ก็ต่อเมื่อ $\overline{\mathbf{d}} = -\frac{\xi_1}{\|\xi_1\|}$ เมื่อ ξ_1 เป็นซับเกรเดียนต์ที่มีขนาดเล็กที่สุด

พิสูจน์ ก่อนอื่นจะแสดงว่า ถ้า ξ_1 เป็นซับเกรเดียนต์ที่มีขนาดเล็กที่สุด แล้ว

$$\|\xi_1\|^2 \le \xi_1^T \xi$$

ทุก ๆ ซับเกรเดียนต์ ξ

พิจารณาเวกเตอร์ $\lambda\xi_1+(1-\lambda)\xi$ ซึ่งเป็นซับเกรเดียนต์ของฟังก์ชัน ที่จุด $\overline{\mathbf{x}}$ เมื่อ $0<\lambda<1$ จะได้ว่า

$$\|\xi_1\|^2 \le \|\lambda\,\xi_1+(1-\lambda)\,\xi\|^2 = (\lambda\xi_1+(1-\lambda)\xi)(\lambda\xi_1+(1-\lambda)\xi)$$

และเราได้

$$\xi_1^T\,\xi_1 \le \lambda^2\,\xi_1^T\,\xi_1+2(1-\lambda)\lambda\,\xi_1^T\,\xi+(1-\lambda)^2\,\xi^T\,\xi$$

$$(1-\lambda)^2\xi_1^T\,\xi_1 \le 2(1-\lambda)\lambda\,\xi_1^T\,\xi+(1-\lambda)^2\xi^T\,\xi$$

$$(1+\lambda)\xi_1^T\,\xi_1 \le 2\lambda\,\xi_1^T\,\xi+(1-\lambda)\xi^T\,\xi$$

$$(1+\lambda)\xi_1^T\,\xi_1 = (1+\lambda)\xi_1^T\,\xi+(1-\lambda)\xi^T\,\xi-(1-\lambda)\xi_1^T\,\xi$$

$$\xi_1^T\,\xi_1 \le \xi_1^T\,\xi+\frac{1-\lambda}{1+\lambda}(\xi-\xi_1)^T\xi \qquad ...(3.4)$$

สมการ (3.4) เป็นจริงทุกค่าของ λ เมื่อ $0<\lambda<1$ โดยให้ $\lambda\to 1$ จะเห็นว่า

$$\|\xi_1\|^2 \le \xi_1^T\,\xi$$

ต่อไปจะแสดงว่า ถ้าให้ $\overline{\mathbf{d}} = \frac{-\xi_1}{\|\xi_1\|}$ แล้ว $\overline{\mathbf{d}}$ เป็นทิศทางลดเร็วที่สุด

จาก

$$\xi^T \overline{\mathbf{d}} = -\frac{\xi^T \xi_1}{\|\xi_1\|} \leq -\|\xi_1\|$$

เราได้ว่า

$$f'(\overline{\mathbf{x}};\overline{\mathbf{d}}) = -\|\xi_1\|$$

ซึ่งเป็นค่าสูงสุดของ $\xi^T \overline{\mathbf{d}}$

ถ้า $\mathbf{d}$ เป็นเวกเตอร์หน่วยใด ๆ จะได้ว่า

$$\xi_1^T \mathbf{d} \geq -\|\xi_1\|$$

ดังนั้น

$$f'(\overline{\mathbf{x}};\mathbf{d}) \geq -\|\xi_1\|$$

นั่นคือ

$$f'(\overline{\mathbf{x}};\overline{\mathbf{d}}) \leq f'(\overline{\mathbf{x}};\mathbf{d})$$

ทุกเวกเตอร์หน่วย $\mathbf{d}$

ดังนั้น $\overline{\mathbf{d}}$ เป็นทิศทางลดเร็วที่สุดของฟังก์ชัน f ที่จุด $\overline{\mathbf{x}}$ ##

ทฤษฎีบท 3.2.5 สำหรับฟังก์ชันนูน f ที่กำหนดค่าบนเซตนูน S ที่จุด $\overline{\mathbf{x}}$ ที่เป็นจุดภายในของ S ถ้า ξ เป็นซับเกรเดียนต์ และ $\xi_1 = -k\xi$ เมื่อ $k > 0$ เป็นซับเกรเดียนต์ด้วยแล้ว จะได้ว่า เวกเตอร์ $\mathbf{0}$ จะเป็นซับเกรเดียนต์ของ f ที่จุด $\overline{\mathbf{x}}$ ด้วย และในกรณีนี้ ที่จุด $\overline{\mathbf{x}}$ ไม่มีทิศทางลด และ $\overline{\mathbf{x}}$ เป็นจุดต่ำสุดของฟังก์ชัน f

พิสูจน์ เราทราบว่า

$$0 = \frac{k}{k+1}\xi + \frac{1}{k+1}(-k\xi)$$

โดยทฤษฎีบท 3.1.6 เราได้ว่า $\mathbf{0}$ เป็นซับเกรเดียนต์

และสำหรับทุกเวกเตอร์หน่วย $\mathbf{d}$ เราได้

$$\mathbf{0}^T \mathbf{d} = 0$$

โดยทฤษฎีบท 3.2.1 $\mathbf{d}$ ไม่เป็นทิศทางลด

ดังนั้น f ไม่มีทิศทางลด และ $\overline{\mathbf{x}}$ เป็นจุดต่ำสุดของฟังก์ชัน f ##

3.3 ฟังก์ชันค่าสูงสุด

ให้ $f_1, f_2, ..., f_m$ เป็นฟังก์ชันที่กำหนดค่าใน R^n กำหนดฟังก์ชัน

(A) $$f(\mathbf{x}) = \max\{f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_m(\mathbf{x})\}$$

มีชื่อว่า *ฟังก์ชันค่าสูงสุด* ฟังก์ชัน f มีสมบัติดังต่อไปนี้

ทฤษฎีบท 3.3.1 ถ้า $f_1, f_2, ..., f_m$ เป็นฟังก์ชันนูนแล้วฟังก์ชัน f ที่กำหนดโดยสมการ (A) เป็นฟังก์ชันนูน

พิสูจน์ ให้ λ เป็นจำนวนใดๆ ซึ่ง $0 \le \lambda \le 1$

จะได้

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) = \max_{1 \le i \le m}\{f_i(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})\}$$

$$= f_k(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})$$

$$\le \lambda f_k(\mathbf{x}) + (1-\lambda) f_k(\mathbf{y}) \qquad \text{บางค่าของ } k$$

$$\le \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

แสดงว่า f เป็นฟังก์ชันนูน ##

ทฤษฎีบท 3.3.2 ถ้า f_i เป็นฟังก์ชันนูนที่มีอนุพันธ์ที่จุด $\bar{\mathbf{x}}$ และ $\bar{\mathbf{x}}$ เป็นจุดที่ทำให้

$$f(\bar{\mathbf{x}}) = f_1(\bar{\mathbf{x}}) = f_2(\bar{\mathbf{x}}) = ... = f_k(\bar{\mathbf{x}})$$

แล้วจะได้ว่า

(1). $\nabla f_i(\bar{\mathbf{x}})$, $1 \le i \le k$ ต่างก็เป็นซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$

(2). เวกเตอร์ ξ เป็นซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$ ก็ต่อเมื่อ $\xi = \sum_{i=1}^{k} \lambda_i \nabla f_i(\bar{\mathbf{x}})$ เมื่อ $\lambda_i > 0$, $i = 1, 2, ..., n$ และ $\sum_{i=1}^{n} \lambda_i = 1$ (เวกตอร์ ξ เป็นผลรวมเชิงนูนของเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$)

พิสูจน์ (1). จะแสดงว่า $\nabla f_i(\bar{\mathbf{x}})$ เป็นซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$ เมื่อ $1 \le i \le k$

ให้ $\mathbf{u}$ เป็นเวกเตอร์หน่วยใดๆ พิจารณา

$$\nabla f(\bar{\mathbf{x}})^T \mathbf{u} \le \frac{f_i(\bar{\mathbf{x}}+\lambda\mathbf{u}) - f_i(\bar{\mathbf{x}})}{\lambda}$$

เพราะว่า

$$f_i(\bar{\mathbf{x}}+\lambda\mathbf{u}) \le f(\bar{\mathbf{x}}+\lambda\mathbf{u})$$

และ

$$f_i(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}})$$

ดังนั้นเราได้ว่า

$$\nabla f(\bar{\mathbf{x}})^T \mathbf{u} \le \frac{f(\bar{\mathbf{x}}+\lambda\mathbf{u}) - f(\bar{\mathbf{x}})}{\lambda}$$

ทุกค่าของ λ

นั่นคือ $\nabla f_i(\bar{\mathbf{x}})$ เป็นซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$ ##

พิสูจน์ (2). ถ้า
$$\xi = \sum_{i=1}^{k} \lambda_i \nabla f_i(\bar{\mathbf{x}})$$

เมื่อ $\lambda_i > 0$, $i = 1,2,\ldots,n$

และ

$$\sum_{i=1}^{n} \lambda_i = 1$$

ดังนั้นจากบทแทรกของทฤษฎีบท 3.1.6 เราได้ว่า ξ เป็นซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$

การพิสูจน์ทางกลับ ให้ ξ เป็นซับเกรเดียนต์ของ f ที่จุด $\bar{\mathbf{x}}$ ดังนั้น เมื่อ $\mathbf{u}$ เป็นเวกเตอร์ใด ๆ

$$\xi^T \mathbf{u} \quad \le f'(\bar{\mathbf{x}};\mathbf{u}) = \quad \nabla f_i(\bar{\mathbf{x}})^T \mathbf{u}$$

สำหรับ i บางตัว เมื่อ $1 \le i \le k$

และจาก

$$\xi^T(-\mathbf{u}) \quad \le f'(\bar{\mathbf{x}};-\mathbf{u}) = \quad \nabla f_j(\bar{\mathbf{x}})^T(-\mathbf{u})$$

สำหรับ j บางตัว เมื่อ $1 \le j \le k,\ j \ne i$

นั่นคือได้

$$\nabla f_j(\bar{\mathbf{x}})^T \mathbf{u} \quad \le \xi^T \mathbf{u} \le \quad \nabla f_i(\bar{\mathbf{x}})^T \mathbf{u}$$

เราได้

$$\xi^T \mathbf{u} \ = t\nabla f_j(\bar{\mathbf{x}})^T \mathbf{u} \ + (1-t)\nabla f_i(\bar{\mathbf{x}})^T \mathbf{u}$$

เมื่อ $0 \le t \le 1$

คือ ได้ว่า

$$\xi^T \mathbf{u} = \mathbf{d}^T \mathbf{u}$$

เมื่อ

$$\mathbf{d} = t\ \nabla f_j(\bar{\mathbf{x}})^T + (1-t)\ \nabla f_i(\bar{\mathbf{x}})^T \qquad \text{...(3.5)}$$

พิจารณาเซต

$$S = \{\ \mathbf{d} \mid \mathbf{d} = \sum_{i=1}^{k} \lambda_i \nabla f_i(\bar{\mathbf{x}}) \text{ เมื่อ } \sum_{i=1}^{n} \lambda_i = 1\}$$

เป็นเซตนูนปิด

ดังนั้น สำหรับเวกเตอร์ ξ ซึ่งเป็นซับเกรเดียนต์ จะมีเวกเตอร์ $\mathbf{d}$ ในเซต S ที่ทำให้

$$(\mathbf{v}-\mathbf{d})^T(\xi-\mathbf{d}) \le 0 \qquad \text{...(3.6)}$$

ทุกเวกเตอร์ $\mathbf{v}$ ในเซต S

จาก (3.5) ได้

$$\xi^T(\xi-\mathbf{d}) = \mathbf{v}^T(\xi-\mathbf{d})$$

เมื่อ $\mathbf{v}$ เป็นเวกเตอร์ในเซต S

จาก (3.6) เราได้

$$\xi^T(\xi-\mathbf{d}) = \mathbf{v}^T(\xi-\mathbf{d}) \le \mathbf{d}^T(\xi-\mathbf{d})$$

และ

$$(\xi-\mathbf{d})^T(\xi-\mathbf{d}) \le 0$$

ดังนั้น

$$(\xi-\mathbf{d})^T(\xi-\mathbf{d}) = 0$$

สรุปได้ว่า

$$\xi-\mathbf{d} = 0$$

นั่นคือ $\xi = \mathbf{d}$ เป็นผลรวมเชิงนูนของเกรเดียนต์ของ f_i ##

บทที่ 4
ขั้นตอนวิธีลดเร็วที่สุดสำหรับกำหนดการเชิงเส้น

4.1 ปัญหาและขั้นตอนวิธี

ปัญหา minimize $f(\mathbf{x})$

subject to $\mathbf{x} \in \mathbf{R}^n$

เมื่อ

$$f(\mathbf{x}) = \max_{1 \le i \le m} \{f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})\}$$

และ $f_i(\mathbf{x}) = b_i + \mathbf{a}_i\mathbf{x}$; $i = 1,2,\dots,m$ เป็นฟังก์ชันเชิงเส้น

การหาค่าต่ำสุดของฟังก์ชันโดยวิธีลดเร็วที่สุดจะประกอบด้วยการหาทิศทางลดเร็วที่สุด และหาค่าของฟังก์ชัน f ในทิศทางลดเร็วที่สุดนั้น ถ้าให้ $\mathbf{u}$ เป็นทิศทางลดเร็วที่สุดเราจะหาค่าต่ำสุดของ $f(\mathbf{x}+t\mathbf{u})$

เพื่อความสะดวกเรากำหนดให้

$$z_i = f_i(\mathbf{x}) = b_i + \mathbf{a}_i\mathbf{x} \;\; ; \;\; i = 1,2,\dots,m$$

กำหนดให้ $z_i^{(k)} = b_i + \mathbf{a}_i\mathbf{x}_i^{(k)}$ เป็นค่า z_i ที่จุด $\mathbf{x}^{(k)}$ เมื่อ k เป็นลำดับการทำซ้ำ (ครั้งที่ k) เมื่อ $k = 0,1,2,\dots$

ที่จุด

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{u}\,t$$

จะได้

$$\begin{aligned} z_i^{(k+1)} &= b_i + \mathbf{a}_i^T\left(\mathbf{x}^{(k)} + \mathbf{u}\,t\right) \\ &= b_i + \mathbf{a}_i^T\mathbf{x}^{(k)} + \mathbf{a}_i^T\mathbf{u}\,t \end{aligned}$$

ดังนั้น

$$z_i^{(k+1)} = z_i^{(k)} + \mathbf{a}_i^T\mathbf{u}\,t$$

ในการกำหนดจุดเริ่มต้น $\mathbf{x}^{(0)}$ เพื่อให้เป็นการง่ายเราให้ $\mathbf{x}^{(0)} = \mathbf{0}$ จะได้

$$z_i = b_i \;\; ; \;\; i = 1,2,\dots,m$$

แล้วทำการหาค่าต่ำสุดตามขั้นตอนวิธี 3 ขั้นตอนหลักดังนี้

ขั้นตอนที่ 1 หาค่า $f(\mathbf{x}^{(k)})$ แล้วตรวจดูว่า $\mathbf{x}^{(k)}$ เป็นจุดต่ำสุดหรือไม่โดยการหาทิศทาง $\mathbf{u}$ ซึ่งเป็นทิศทางลดเร็วที่สุด ถ้าไม่มีทิศทางลดเร็วที่สุดแสดงว่า $\mathbf{x}^{(k)}$ เป็นจุดต่ำสุดแต่ถ้ามีทิศทางลดเร็วที่สุดให้ทำขั้นตอนที่ 2

ขั้นตอนที่ 2 หาค่า $t^{(k)}$ ที่ทำให้ $f\left(\mathbf{x}^{(k)}+t\mathbf{u}\right)$ เป็นค่าต่ำสุดในทิศทางของ $\mathbf{u}$ โดยวิธี *เลือกเส้นต่ำสุด* ได้ค่า $t=t^{(k)}$

ขั้นตอนที่ 3 ให้ $\mathbf{x}^{(k+1)}=\mathbf{x}^{(k)}+t^{(k)}\mathbf{u}$ และหาค่า $z_i=f_i\left(\mathbf{x}^{(k+1)}\right)$, $i=1,2,...,m$ แล้วทำขั้นตอนที่ 1

ดังแสดงรายละเอียดต่อไปนี้

ขั้นตอนที่ 1

1.1 หาค่า $f(\mathbf{x}^{(k)})$

ให้ k_i เป็นจำนวนเต็มที่ z_{k_i} เป็นค่าสูงสุด

$$z_{k_i} = \max_{1\le j\le m}\{\ z_j\ \} \quad ,i=1,2,...,l$$

อาจมีค่า k_i หลายตัวจึงสมมุติว่า $z_{k_1}=z_{k_2}=...=z_{k_l}$
ได้

$$f(\mathbf{x}^{(k)}) = z_{k_1}=z_{k_2}=...=z_{k_l}$$

กำหนดให้

$$K=\{k_i \mid i=1,2,...,l\}$$

1.2 ตรวจดูว่า $\mathbf{x}^{(k)}$ เป็นจุดต่ำสุดหรือไม่ โดยหาทิศทางลดเร็วที่สุด
เนื่องจาก

$$\nabla f_i(\mathbf{x}) = \mathbf{a}_i \quad , \ i=1,2,...,m$$

ให้

$$\xi = y_1\mathbf{a}_{k_1}+y_2\mathbf{a}_{k_2}+...+y_l\mathbf{a}_{k_l}$$

เมื่อ $y_i\ge 0$, $i=1,2,...,l$

และ $\sum_{i=1}^{l} y_i = 1$

เวกเตอร์ ξ จะเป็นซับเกรเดียนต์ของ f ที่จุด $\mathbf{x}^{(k)}$ (ทฤษฎีบท 3.3.2 (2))

จะหา y_i ที่ทำให้ $\|\xi\|$ มีค่าน้อยที่สุด

แล้วให้

$$\mathbf{u} = -\frac{\xi}{\|\xi\|}$$

จะได้ว่า $\mathbf{u}$ เป็นทิศทางลดเร็วที่สุด (ทฤษฎีบท 3.2.4)

ให้ $\mathbf{A} = [\mathbf{a}_{k_1} \ \ \mathbf{a}_{k_2} \ \ \ldots \ \ \mathbf{a}_{k_l}]^T$ เป็นเมทริกซ์ซึ่งประกอบด้วยเวกเตอร์ $\mathbf{a}_{k_i}$, $i = 1, 2, \ldots, l$

ดังนั้น

$$\xi = \mathbf{A}\mathbf{y}$$

เมื่อ $\mathbf{y} = [y_1 \ \ y_2 \ \ \ldots \ \ y_l]^T$, $y_i \geq 0$, $i = 1, 2, \ldots, l$

นั่นคือแก้ปัญหา

(A) minimize $\|\mathbf{A}\mathbf{y}\|^2$

subject to $\mathbf{y}^T\mathbf{e} = 1$

$\mathbf{y} \geq \mathbf{0}$

เมื่อ $\mathbf{e} = [1\ 1\ 1 \ldots 1]^T$ เป็นเวกเตอร์ที่มี 1 เป็นส่วนประกอบ l ตัว

ถ้าปัญหามีผลเฉลย ให้

$$\xi = \mathbf{A}\mathbf{y}$$

และให้

$$\mathbf{u} = -\frac{\xi}{\|\xi\|}$$

แล้วทำขั้นตอนที่ 2

ในการแก้ปัญหาค่าต่ำสุดข้างบนเราหาค่า $\mathbf{y}$ ที่สอดคล้องกับเงื่อนไขคารุช-คุน-ทักเกอร์ (Karush-Kuhn-Tucker condition) (ดูภาคผนวก ก.)

$$\nabla\varphi(\mathbf{y}) + \lambda\nabla g(\mathbf{y}) + \mu\nabla h(\mathbf{y}) = 0$$

เมื่อ

$$\varphi(\mathbf{y}) = \|\mathbf{A}\mathbf{y}\|^2 = \mathbf{y}^T\mathbf{A}^T\mathbf{A}\mathbf{y}$$

$$g(\mathbf{y}) = \mathbf{y}^T\mathbf{e} - 1 \quad = \quad \sum_{i=1}^{n} y_i - 1$$

$$h_i(\mathbf{y}) = -y_i \ , \quad i = 1,2,\ldots,n$$

ดังนั้น

$$\nabla\varphi(\mathbf{y}) = 2\mathbf{A}^T\mathbf{A}\mathbf{y}$$

$$\nabla g(\mathbf{y}) = \mathbf{e} = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}^T$$

และ

$$\nabla h_i(\mathbf{y}) = \begin{bmatrix} 0 & 0 & -1 & \ldots & 0 \end{bmatrix}^T \ , \ i = 1,2,\ldots,n$$

เป็นเวกเตอร์ที่มี -1 เป็นส่วนประกอบที่ i
นั่นคือเงื่อนไขของ คารูช-คุน-ทักเกอร์ จะเป็น

(B)

$$2\mathbf{A}^T\mathbf{A}\mathbf{y} + \lambda\mathbf{e} - \mu = 0$$

$$\lambda\left(\mathbf{y}^T\mathbf{e} - 1\right) = 0 \ , \ y_i\mu_i = 0 \ , \ i = 1,2,\ldots,n$$

$$\mathbf{y} \geq \mathbf{0} \ \text{ และ } \ \mu \geq \mathbf{0}$$

นั่นคือแก้สมการ (B) โดยเริ่มต้นแก้สมการ

$$\mathbf{A}^T\mathbf{A}\mathbf{y} = \mathbf{e}$$

เมื่อแก้สมการได้ผลเฉลย $\mathbf{y} \geq \mathbf{0}$ แล้ว
เราให้

$$\mathbf{y}_1 = \frac{\mathbf{y}}{\mathbf{y}^T\mathbf{e}}$$

เมื่อ $\mathbf{y}_1$ เป็น $\mathbf{y}$ ตัวใหม่ที่ได้จากการคำนวณ
ก็จะได้ว่า $\mathbf{y}$ สอดคล้องกับเงื่อนไข $\mathbf{y}^T\mathbf{e} = 1$ ตามต้องการ
จากนี้ ได้ค่า $\mu = \mathbf{0}$ และ

$$\lambda = \frac{-2}{\mathbf{y}^T\mathbf{e}}$$

สอดคล้องกับเงื่อนไขของคารูช-คุน-ทักเกอร์ สมการ (B) และได้ผลเฉลยของปัญหา (A)

ในขั้นตอนคำนวณจริง ไม่จำเป็นต้องหาค่า $\mathbf{y}_1$ เพราะว่าในการหาจุด

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{u}\,t$$

เมื่อ $\mathbf{u}$ เป็นทิศทางลดเร็วที่สุด เราต้องหาค่า t ที่เหมาะสมที่จะทำให้ฟังก์ชัน f มีค่าต่ำสุดในทิศทางนี้ ดังนั้นเวกเตอร์ $\mathbf{u}$ จึงไม่จำเป็นต้องเป็นเวกเตอร์หน่วย เมื่อได้ผลเฉลย $\mathbf{y} > 0$ ของสมการ

$$\mathbf{A}^T\mathbf{A}\mathbf{y} = \mathbf{e}$$

แล้ว เราให้

$$\mathbf{u} = -\mathbf{A}\mathbf{y}$$

ก็จะได้ทิศทางลดเร็วที่สุดตามต้องการ

กรณีที่สมการ $\mathbf{A}^T\mathbf{A}\mathbf{y} = \mathbf{e}$ ไม่มีผลเฉลย ซึ่งจะเกิดขึ้นเมื่อ $\operatorname{rank} A^T A < l$ เนื่องมาจากคอลัมน์ของ $\mathbf{A}$ ไม่เป็นอิสระเชิงเส้น (เมื่อ $l > n$ จะเกิดกรณีนี้) ดังนั้นสมการ $\mathbf{A}\mathbf{y} = 0$ อาจมีผลเฉลย $\mathbf{y} > 0$ ซึ่งกรณีนี้แสดงว่าไม่มีทิศทางลด เพราะว่ามี $\mathbf{0}$ เป็นซับเกรเดียนต์ (บทแทรก 3.2.2)

สรุปขั้นตอนการหาทิศทางลดเร็วที่สุด เป็นดังนี้

1.1 ให้ k_i เป็นจำนวนเต็มที่ z_{k_i} เป็นค่าสูงสุด

$$z_{k_i} = \max_{1 \le j \le m} \{ z_j \} \quad , i = 1, 2, \ldots, l$$

และได้

$$f(\mathbf{x}^{(k)}) = z_{k_1} = z_{k_2} = \ldots = z_{k_l}$$

กำหนดให้

$$K = \{k_i \mid i = 1, 2, \ldots, l\}$$

พิจารณาค่าของ l

(1). ถ้า $l > n$ ให้กระทำขั้น 1.2

(2). ถ้า $l \le n$ ให้กระทำขั้น 1.3

1.2 แก้สมการ

$$\mathbf{A}\mathbf{y} = 0, \quad \mathbf{y} \ge 0, \quad \mathbf{y} \ne 0$$

(1). ถ้ามีผลเฉลย แสดงว่าไม่มีทิศทางลด ที่จุด $\mathbf{x}^{(k)})$ เป็นจุดต่ำสุด

(2). ถ้าสมการไม่มีผลเฉลย $\mathbf{y} \ge 0$ แต่มีผลเฉลย $y_i > 0$ สำหรับ i บางค่า และมีผลเฉลย $y_j < 0$ สำหรับ j บางค่า แสดงว่ามีทิศทางลด เราจัดค่า $\mathbf{y}$ ให้จำนวนของ i มากกว่าหรือเท่ากับจำนวน

ของ j (ถ้าจำนวนของ i น้อยกว่าจำนวนของ j ให้คูณ $\mathbf{y}$ ด้วย -1) แล้วให้ตัด $\mathbf{a}_{k_j}$ ออกจากเมทริกซ์ A ทุกค่า j ที่ $y_j < 0$ แล้วดำเนินการขั้น 1.3

1.3 แก้สมการ

$$\mathbf{A}^T\mathbf{A}\mathbf{y} = \mathbf{e}$$

(1). ถ้ามีผลเฉลย $\mathbf{y} > 0$ ให้ $\mathbf{u} = -\mathbf{A}\mathbf{y}$ แล้วดำเนินการในขั้นตอนที่ 2

(2). ถ้ามีผลเฉลย $y_j < 0$ สำหรับ j บางตัว ให้ตัด $\mathbf{a}_{k_j}$ ออกจากเมทริกซ์ A ทุกค่า j ที่ $y_j < 0$ แล้วดำเนินการขั้น 1.3 ใหม่

(3). ถ้าไม่มีผลเฉลย ให้กระทำตามขั้น 1.2

ขั้นตอนที่ 2 (Line Search)

หาค่า $t^{(k)}$ ที่ทำให้

$$f\left(\mathbf{x}^{(k)} + t^{(k)}\mathbf{u}\right) \le f\left(\mathbf{x}^{(k)} + t\mathbf{u}\right)$$

ทุกค่าของ t โดยวิธี*เลือกเส้นต่ำสุด*

ขั้นตอนที่ 3

สำหรับ $i = 1, 2, ..., n$ ให้

$$\mathbf{x}_i^{(k+1)} = \mathbf{x}_i^{(k)} + \mathbf{u}\,t^{(k)}$$

สำหรับ $i = 1, 2, ..., m$ ให้

$$z_i^{(k+1)} = z_i^{(k)} + \mathbf{a}^T\mathbf{u}\,t^{(k)}$$

แล้วกลับไปทำขั้นตอนที่ 1 ใหม่

4.2 รายละเอียดของขั้นตอนวิธี

การแก้ระบบสมการเชิงเส้น

ในการแก้สมการ $\mathbf{A}^T\mathbf{A}\mathbf{y} = \mathbf{e}$ เราใช้วิธีกำจัดออกของเกาส์ (Gaussian elimination method) คือ กระทำการดำเนินการแบบแถวกับเมทริกซ์แต่งเติม $\left[\mathbf{A}^T\mathbf{A} \mid \mathbf{e}\right]$ เพื่อทำให้เป็นเมทริกซ์เป็นขั้น ถ้าในการดำเนินการมีผลให้เมทริกซ์มีแถวใดแถวหนึ่งเป็นศูนย์ทั้งแถว แสดงว่าระบบสมการไม่มีผลเฉลย หรือมีผลเฉลยจำนวนไม่จำกัด ซึ่งแสดงว่าเวกเตอร์คอลัมน์ของเมทริกซ์ A ไม่เป็นอิสระเชิงเส้น

ในการแก้สมการ $\mathbf{A}\mathbf{y} = 0$, $\mathbf{y} \ge 0$ เมื่อทราบแล้วว่าคอลัมน์ของเมทริกซ์ A ไม่เป็นอิสระเชิงเส้น จะมีผลเฉลย $\mathbf{y} \ne 0$ แน่นอน

วิธีเลือกเส้นต่ำสุด

สำหรับหาจุดต่ำสุดของฟังก์ชันนูนในแบบ

$$g(t) = max\{z_1 + m_1 t, z_2 + m_2 t, \dots, z_l + m_l t\}$$

สมมุติว่าที่จุด $t = t_1$ ค่าสูงสุดเท่ากับ $z_1 + m_1 t_1$
เพื่อความสะดวกในการอธิบาย สมมุติว่า $m_1 < 0$
และ

$$m_1 \le m_2 \le \dots \le m_k$$

เริ่มต้นให้ $t_1 = 0$

ขั้นที่ 1 หา i ที่มีสมบัติว่า $m_1 m_i < 0$

กรณีที่ 1 ถ้ามี i ที่มีคุณสมบัติดังกล่าวให้

$$I = \{i \mid m_1 m_i < 0\}$$

หาค่า t_i เมื่อ $i \in I$ ที่ $|t_i - t_1|$ มีค่าน้อยที่สุดที่ทำให้

$$z_i + m_i t_i = z_1 + m_1 t_1$$

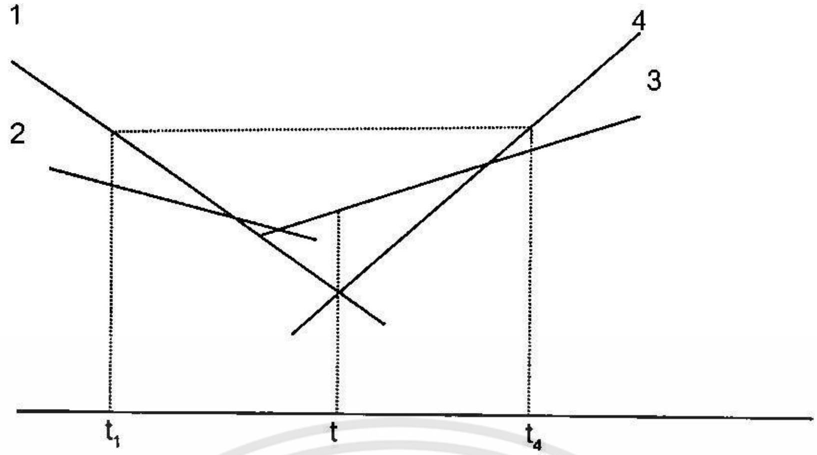
ได้ j แล้วไปทำขั้นที่ 2

กรณีที่ 2 ถ้าไม่มี i ที่มีสมบัติว่า $m_1 m_i < 0$ แต่มี i ที่มีสมบัติว่า $m_i = 0$ ให้

$$I = \{ i \mid m_i = 0\}$$

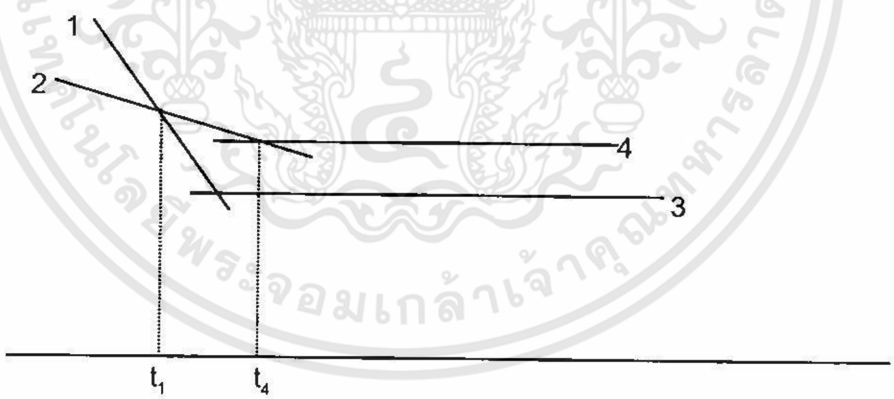
หาค่าสูงสุดของ z_i เมื่อ $i \in I$ ได้ z_j เป็นค่าสูงสุด ได้ค่า j แล้วไปทำขั้นที่ 2

กรณีที่ 3 ถ้าไม่มี i ที่มีสมบัติว่า $m_1 m_i \le 0$ แสดงว่า $g(t)$ มีค่าลดลงโดยไม่มีขอบเขต กรณีนี้ไม่มีผลเฉลย สิ้นสุดขั้นตอนวิธี



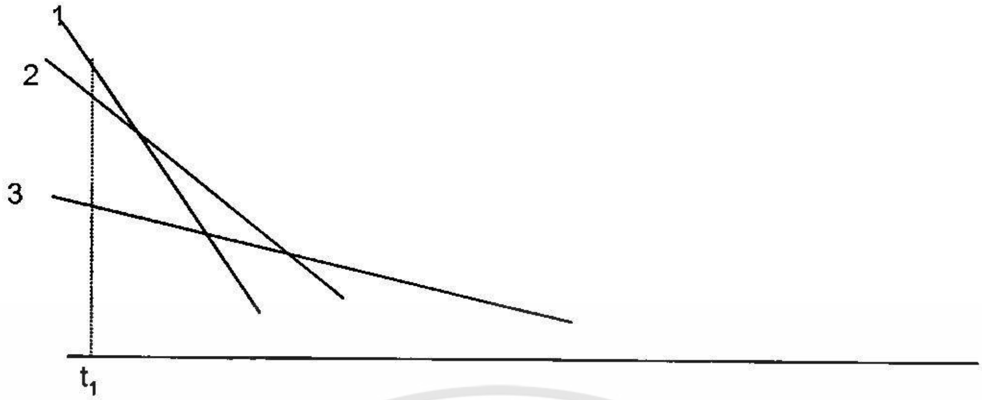
ภาพที่ 4.1 แสดงจุดตัดและค่าของฟังก์ชัน

กรณีที่ 1 $|t_4 - t_1|$ น้อยที่สุดที่ทำให้ $z_1 + m_1 t_1 = z_4 + m_4 t_4$ ได้ค่า t_j และได้ค่า $g(t) = z_3 + m_3 t$ ในขั้นที่ 2



ภาพที่ 4.2 แสดงจุดตัดและค่าของฟังก์ชัน

กรณีที่ 2 $m_3 = 0$ และ $m_4 = 0$ ได้ $g(t) = z_4$



ภาพที่ 4.3 แสดงจุดตัดและค่าของฟังก์ชัน

กรณีที่ 3 $m_1 m_i > 0$ ทุกค่า i ไม่มีค่าต่ำสุด

ขั้นที่ 2 แก้สมการหาค่า t

$$z_1 + m_1 t = z_j + m_j t$$

ได้ค่า t แล้วหาค่า $g(t)$

$$g(t) = \max \{z_1 + m_1 t, z_2 + m_2 t, \ldots, z_k + m_k t\}$$

(ในการหาค่าของ $z_i + m_i t$ เพื่อเปรียบเทียบค่าเพื่อหาค่าสูงสุด ไม่จำเป็นต้องหาค่าเมื่อ $m_i < m_1$ เพราะว่าค่าจะต่ำกว่า $z_1 + m_1 t$ อยู่แล้ว)

ถ้า

$$g(t) = z_1 + m_1 t$$

แสดงว่า t เป็นจุดต่ำสุด จบขั้นตอนนี้

ถ้า

$$g(t) = z_j + m_j t > z_1 + m_1 t$$

สำหรับ j บางตัว แสดงว่า t ยังไม่ใช่จุดต่ำสุด

ให้แทน j ด้วย 1 คือให้ $t_1 = t$ แล้วกลับไปทำขั้นที่ 1

4.3 ตัวอย่างและผลเฉลย

ตัวอย่าง 1 ปัญหาเดิม

(P) minimize $C = -s_1 - 2s_2 + 4s_5$

subject to

$$s_2 - s_3 = 0 \quad ...(1)$$

$$2s_1 - 2s_2 + 4s_3 - 4s_5 = 0 \quad ...(2)$$

$$s_1 + 2s_2 + s_4 - 4s_5 = 0 \quad ...(3)$$

$$s_1 + s_2 + s_3 + s_4 + s_5 = 1 \quad ...(4)$$

$$s_i \geq 0$$

ปัญหานี้มีผลเฉลยเหมาะที่สุดคือ $s_1 = s_4 = 0$, $s_2 = s_3 = 0.4$ และ $s_5 = 0.2$ และค่าต่ำสุดเท่ากับ 0

ปัญหาคู่กันของปัญหาเดิมคือ

(D) maximize x_4

subject to

$$2x_2 + x_3 + x_4 \leq -1$$

$$x_1 - 2x_2 + 2x_3 + x_4 \leq -2$$

$$-x_1 + 4x_2 + x_4 \leq 0$$

$$x_3 + x_4 \leq 0$$

$$-4x_2 - 4x_3 + x_4 \leq 4$$

$$x_i \text{ free}$$

การดำเนินการหาผลเฉลย

จากปัญหาคู่กันให้ $w = -x_4$ จะได้

minimize w

subject to

$$w \geq 1 + 2x_2 + x_3$$

$$w \geq 2 + x_1 - 2x_2 + 2x_3$$

$$w \geq -x_1 + 4x_2$$

$$w \geq x_3$$

$$w \geq -4 - 4x_2 - 4x_3$$

$$x_i \text{ free}$$

ให้

$z_1 = 1 + 2x_2 + x_3$ $\quad a_1 = (0, 2, 1)$

$z_2 = 2 + x_1 - 2x_2 + 2x_3$ $\quad a_2 = (1, -2, 2)$

$z_3 = -x_1 + 4x_2$ $\quad a_3 = (-1, 4, 0)$

$z_4 = x_3$ $\quad a_4 = (0, 0, 1)$

$z_5 = -4 - 4x_2 - 4x_3$ $a_5 = (0, -4, -4)$

เริ่มต้นให้ $x^0 = (x_1, x_2, x_3) = (0, 0, 0)$

$w = \max\{z_1, z_2, z_3\}$

$z_1^0 = 1$

$z_2^0 = 2$ ...ค่าสูงสุด

$z_3^0 = 0$

$z_4^0 = 0$

$z_5^0 = -4$

ครั้งที่ 1

จำนวนระนาบเท่ากับ 1

ได้ z_2 เป็น ค่าสูงสุด และ $a_2 = (1, -2, 2)$

แต่ $u = -a = (-1, 2, -2)$

$x^1 = x^0 + ut = (0, 0, 0) + (1, -2, 2)t$

$= (-t, 2t, -2t)$

หาค่า $z_i^{(k+1)} = z_i^{(k)} + a_i^T ut$

$z_1^1 = 1+2t$

$z_2^1 = 2-9t$

$z_3^1 = 9t$

$z_4^1 = -2t$

$z_5^1 = -4$

เลือกสมการที่แทนค่า $t \geq 0$ แล้วทำให้ w เพิ่มขึ้น นั่นคือหา t ที่น้อยที่สุด

ดังนั้นเลือกสมการที่ $m_2 m_i < 0$ ได้ระนาบที่ 1 และ ระนาบที่ 3

ให้ $z_1^1 = 1+2t = 2$ (ค่าสูงสุด) ได้ t = 0.5

ให้ $z_3^1 = 9t = 2$ (ค่าสูงสุด) ได้ t = 0.22....

เลือก t ที่น้อยที่สุดได้ระนาบที่ 3

ให้ $z_3 = z_2$

แก้สมการ $2-9t = 9t$

ได้ t = 0.1111 แทนค่า t ใน z_i^1

$z_1^1 = 1.2222$ ...ค่าสูงสุด

$z_2^1 = 1$

$z_3^1 = 1$

$z_4^1 = -0.2222$

$z_5^1 = -4$

ให้ $z_1 = z_2$ ได้ $t = 0.090909$ แทนใน z_i^1

$z_1^1 = 1.1818$ ...ค่าสูงสุด

$z_2^1 = 1.1818$ ...ค่าสูงสุด

$z_3^1 = 0.8182$

$z_4^1 = -0.1818$

$z_5^1 = -4$

ฉะนั้น $x^1 = (-t, 2t, -2t) = (-.0909, 0.1818, -0.1818)$

ครั้งที่ 2

จำนวนระนาบเท่ากับ 2

$u = y_1a_1 + y_2a_2$

หา y_i จาก $A^TAy = 1$

ได้ค่า $y_1 = 0.2682927$ และ $y_2 = 0.1707317$

ดังนั้น $u = (-0.1707, -0.1951, -0.6098)$ และหาค่า z_i^2

$z_1^2 = 1.1818 - t$

$z_2^2 = 1.1818 - t$

$z_3^2 = 0.8182 - 0.6098\,t$

$z_4^2 = -0.1818 - 0.6098\,t$

$z_5^2 = -4 + 3.2195\,t$

ให้ $z_1 = z_2 = z_5$ ได้ค่า $t = 1.2281$ แทนใน z_i^2

$z_1^2 = -0.0462$

$z_2^2 = -0.0462$

$z_3^2 = 0.0694$ ...ค่าสูงสุด

$z_4^2 = -0.9306$

$z_5^2 = -0.0462$

z_3 เป็นค่าสูงสุดเลือก $z_3 = z_5$ ได้ $t = 1.2583$ แทนใน z_i^2

$z_1^2 = -0.0764$

$z_2^2 = -0.0764$

$z_3^2 = 0.0510$ ...ค่าสูงสุด

$z_4^2 = -0.9490$

z_5^2 = 0.0510 ...ค่าสูงสุด

ฉะนั้น $x^2 = (-0.3057, -0.0637, -0.9490)$

ครั้งที่ 3

จำนวนระนาบเท่ากับ 2

u = $y_1a_3 + y_2a_5$

แก้สมการหาค่า y_i **ได้** $y_1 = 0.1667$ **และ** $y_2 = 0.1146$

แทนค่า y_i **ได้** u = (0.1667, -0.2083, 0.4583) **และหาค่า** z_i^3

$z_1^3 = -0.0764 + 0.0417\,t$

$z_2^3 = -0.0764 + 1.5\,t$

$z_3^3 = 0.0510 - t$

$z_4^3 = -0.9490 + 0.4583\,t$

$z_5^3 = 0.0510 - t$

ได้ t = 0.0510 **แทนใน** z_i^3

$z_1^3 = -0.0743$

$z_2^3 = 0$ ...ค่าสูงสุด

$z_3^3 = 0$ ...ค่าสูงสุด

$z_4^3 = -0.9257$

$z_5^3 = 0$ ...ค่าสูงสุด

ฉะนั้น $x^3 = (-0.2972, -0.0743, -0.9257)$

ครั้งที่ 4

จำนวนระนาบเท่ากับ 3

หาค่า $u = y_1a_2 + y_2a_3 + y_3a_5$ แก้สมการ $Ay = 0$

จะได้

$$\begin{bmatrix} 1 & -1 & 0 \\ -2 & 4 & -4 \\ 2 & 0 & -4 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

ได้ค่า $y_1 = t$, $y_2 = t$ **และ** $y_3 = 0.5t$

ได้คำตอบ y ที่เป็นบวกทั้งหมดแสดงว่าไม่มีทิศทางลด

ที่จุด $x^3 = (-0.2972, -0.0743, -0.9257)$ **เป็นจุดต่ำสุด และมีค่าต่ำสุดเท่ากับ 0** ##

เมื่อใช้ เงื่อนไขส่วนเติมเต็มของตัวแปรช่วย(complementary slackness condition)

(ดูภาคผนวก ข.) จะได้ผลเฉลย $s_1 = s_4 = 0$, $s_2 = s_3 = 0.4$ และ $s_5 = 0.2$ และค่าต่ำสุดเท่ากับ 0 ตรงกับผลเฉลย ของปัญหาเดิม

ตัวอย่าง 2

(P) minimize $C = -3s_2 + 5s_4 + 10s_5 + 8s_6$

subject to

$$-1.5s_2 - s_3 - s_4 + s_5 + 2s_6 = 0 \quad ...(1)$$
$$-1.25s_1 - 0.5s_2 + s_4 + s_5 - s_6 = 0 \quad ...(2)$$
$$s_1 + s_2 + s_3 + s_4 + s_5 + s_6 = 1 \quad ...(4)$$
$$s_i \ge 0$$

ปัญหานี้มีผลเฉลยเหมาะที่สุดคือ $s_1 = 0.2857$, $s_2 = s_4 = s_6 = 0$, $s_3 = 0.3571$ และ $s_5 = 0.3571$ และค่าต่ำสุดเท่ากับ 3.5714

ปัญหาคู่กันของปัญหาเดิมคือ

(D) maximize x_3

subject to

$$-1.25x_2 + x_3 \le 0$$
$$-1.5x_1 - 0.5x_2 + x_3 \le -3$$
$$-x_1 + x_3 \le 0$$
$$-x_1 + x_2 + x_3 \le 5$$
$$x_1 + x_2 + x_3 \le 10$$
$$2x_1 - x_2 + x_3 \le 8$$
$$x_i \text{ free}$$

การดำเนินการหาผลเฉลย

จากปัญหาคู่กันให้ $w = -x_3$ จะได้

minimize w

subject to

$$w \ge -1.25x_2$$
$$w \ge 3 - 1.5x_1 - 0.5x_2$$
$$w \ge -x_1$$
$$w \ge -5 - x_1 + x_2$$
$$w \ge -10 + x_1 + x_2$$
$$w \ge -8 + 2x_1 - x_2$$
$$x_i \text{ free}$$

ให้ $z_1 = -1.25x_2$ $a_1 = (0, -1.25)$

$z_2 = 3 - 1.5x_1 - 0.5x_2$ $a_2 = (-1.5, -0.5)$

$z_3 = -x_1$ $a_3 = (-1, 0)$

$z_4 = -5 - x_1 + x_2$ $a_4 = (-1, 1)$

$z_5 = -10 + x_1 + x_2$ $a_5 = (1, 1)$

$z_6 = -8 + 2x_1 - x_2$ $a_6 = (2, -1)$

เริ่มต้นให้ $x^0 = (x_1, x_2) = (0, 0)$

$w = \max\{z_1, z_2, \ldots, z_6\}$

$z_1^0 = 0$

$z_2^0 = 3$ ...ค่าสูงสุด

$z_3^0 = 0$

$z_4^0 = -5$

$z_5^0 = -10$

$z_6^0 = -8$

ครั้งที่ 1

จำนวนระนาบเท่ากับ 1

ได้ z_2 เป็น ค่าสูงสุด และ $a_2 = (-1.5, -0.5)$

แต่ $u = -a_2 = (1.5, 0.5)$

$x^1 = x^0 + ut = (0, 0) + (1.5, 0.5)t$

$= (1.5t, 0.5t)$

หาค่า $z_i^{(k+1)} = z_i^{(k)} + \mathbf{a}_i^T \mathbf{u}t$

$z_1^1 = -0.625\,t$

$z_2^1 = 3 - 2.5\,t$

$z_3^1 = -1.5\,t$

$z_4^1 = -5 - t$

$z_5^1 = -10 + 2\,t$

$z_6^1 = -8 + 2.5\,t$

เลือกสมการที่แทนค่า $t \geq 0$ แล้วทำให้ w เพิ่มขึ้น

นั่นคือเลือกสมการที่ $m_2m_i < 0$ ได้ระนาบที่ 5 และระนาบที่ 6

ให้ $z_5^1 = -10 + 2\,t = 3$ (ค่าสูงสุด) ได้ $t = 6.5$

$z_6^1 = -8 + 2.5\,t = 3$ ได้ t = 4.4

เลือก t ที่น้อยที่สุดได้ระนาบที่ 6

ดังนั้น ให้ $z_6 = z_2$

แก้สมการ $3 - 2.5t = -8 + 2.5t$

ได้ t = 2.2 แทนใน z_i^1

$z_1^1 = -1.375$ ...ค่าสูงสุด

$z_2^1 = -2.5$

$z_3^1 = -3.3$

$z_4^1 = -7.2$

$z_5^1 = -5.6$

$z_6^1 = -2.5$

z_2^1 และ z_6^1 ไม่ใช่ค่าสูงสุด

ดังนั้นเลือก $z_6 = z_1$ ได้ค่า t = 2.56 แทนใน z_i^1 ใหม่ได้

$z_1^1 = -1.6$ ...ค่าสูงสุด

$z_2^1 = -3.4$

$z_3^1 = -3.84$

$z_4^1 = -7.56$

$z_5^1 = -4.88$

$z_6^1 = -1.6$ ...ค่าสูงสุด

ฉะนั้น $x^1 = (3.84, 1.28)$

ครั้งที่ 2

จำนวนระนาบเท่ากับ 2

$u = y_1a_1 + y_2a_6$

หาค่า y_i โดยหา $A^TAy = 1$ ได้ค่า $y_1 = 0.6$ และ $y_2 = 0.05$

แทนค่าได้ u = (-0.1, 0.8) และหาค่า z_i^2

$z_1^2 = -1.6 - t$

$z_2^2 = -3.4 - 0.25\,t$

$z_3^2 = -3.84 + 0.1t$

$z_4^2 = -7.56 + 0.9\,t$

$z_5^2 = -4.88 + 0.7\,t$

z_6^2 = -1.6 – t

ได้ค่า t = 1.9294 แทนใน z_i^2

z_1^2 = -3.5294 ...ค่าสูงสุด

z_2^2 = -3.8824

z_3^2 = -3.6471

z_4^2 = -5.8235

z_5^2 = -3.5294 ...ค่าสูงสุด

z_6^2 = -3.5294 ...ค่าสูงสุด

ดังนั้น x^2 = (3.6471, 2.8235)

ครั้งที่ 3

จำนวนระนาบเท่ากับ 3

$u = y_1a_1 + y_2a_5 + y_3a_6$ หา y_i จาก A y = 0

$$\begin{bmatrix} 0 & 1 & 2 \\ -1.25 & 1 & -1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

ได้ $y_1 = 1.6t$, $y_2 = t$ และ $y_3 = -t$ จะเห็นว่า y_3 เป็นลบ

ดังนั้นพิจารณาเฉพาะ y ที่เป็นบวกแล้วแก้สมการใหม่ได้

$$\begin{bmatrix} 0 & 1 \\ -1.25 & 1 \end{bmatrix}^T \begin{bmatrix} 0 & 1 \\ -1.25 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

ได้ค่า $y_1 = 2.08$, $y_2 = 1.8$

ดังนั้น ได้ u = (-1.8, -0.8) และหาค่า z_i^3

z_1^3 = -3.5294 – t

z_2^3 = -3.8824 + 2.3 t

z_3^3 = -3.6471 + 1.8 t

z_4^3 = -5.8235 + 2.6 t

z_5^3 = -3.5294 - t

z_6^3 = -3.5294 – 4.4 t

ให้ $z_1 = z_3 = z_5$ ได้ค่า t = 0.0418 แทนค่าใน z_i^3

z_1^3 = 3.5714 ...ค่าสูงสุด

z_2^3 = -3.7863

z_3^3 = 3.5714 ...ค่าสูงสุด

z_4^3 = -5.7148

z_5^3 = 3.5714 ...ค่าสูงสุด

z_6^3 = -3.7133

ฉะนั้น x^3 = (3.5714, 2.8576)

ครั้งที่ 4

จำนวนระนาบเท่ากับ 3

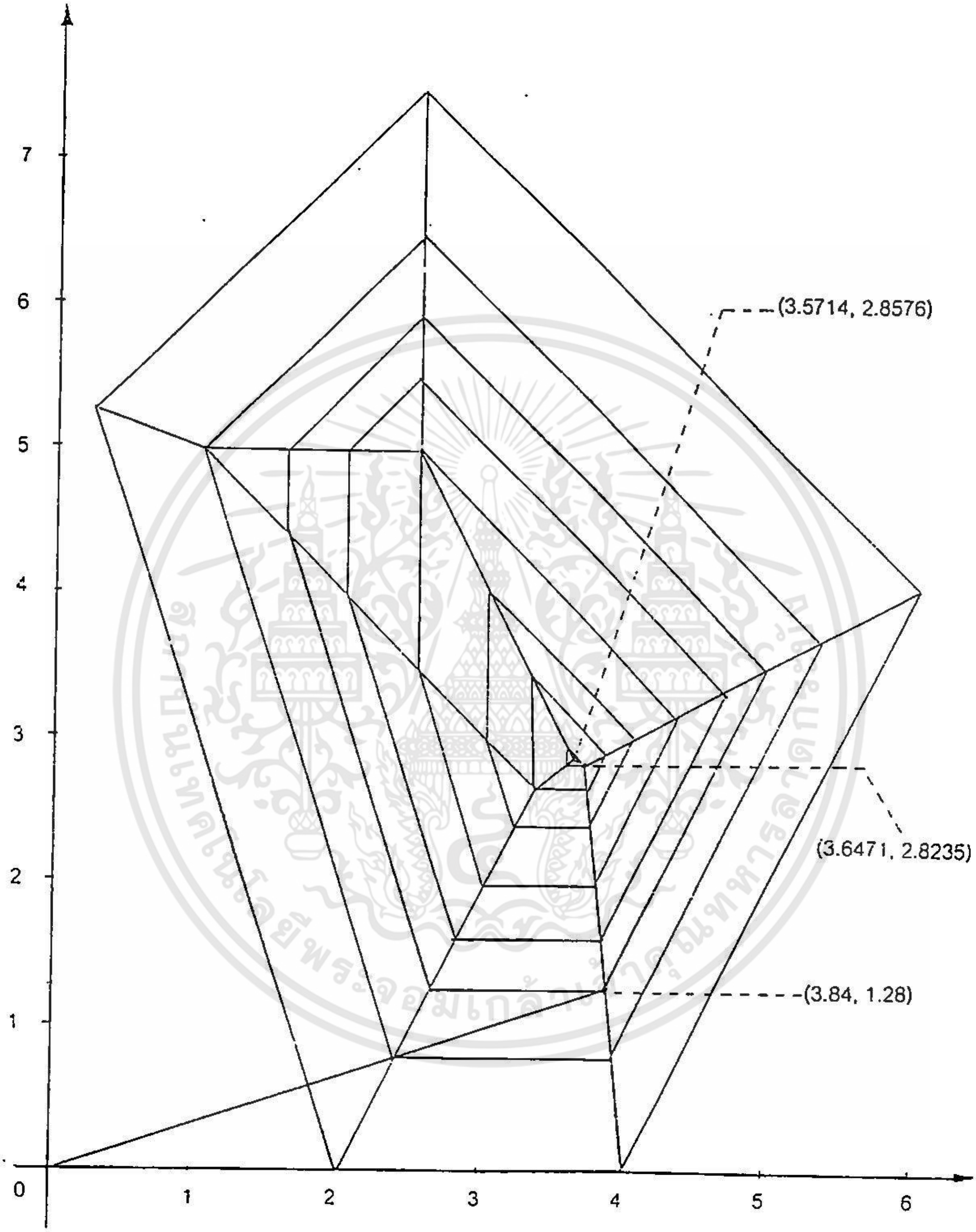
$u = y_1a_1 + y_2a_3 + y_3a_5$ หา y_i จาก $A\,y = 0$

$$\begin{bmatrix} 0 & -1 & 1 \\ -1.25 & 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

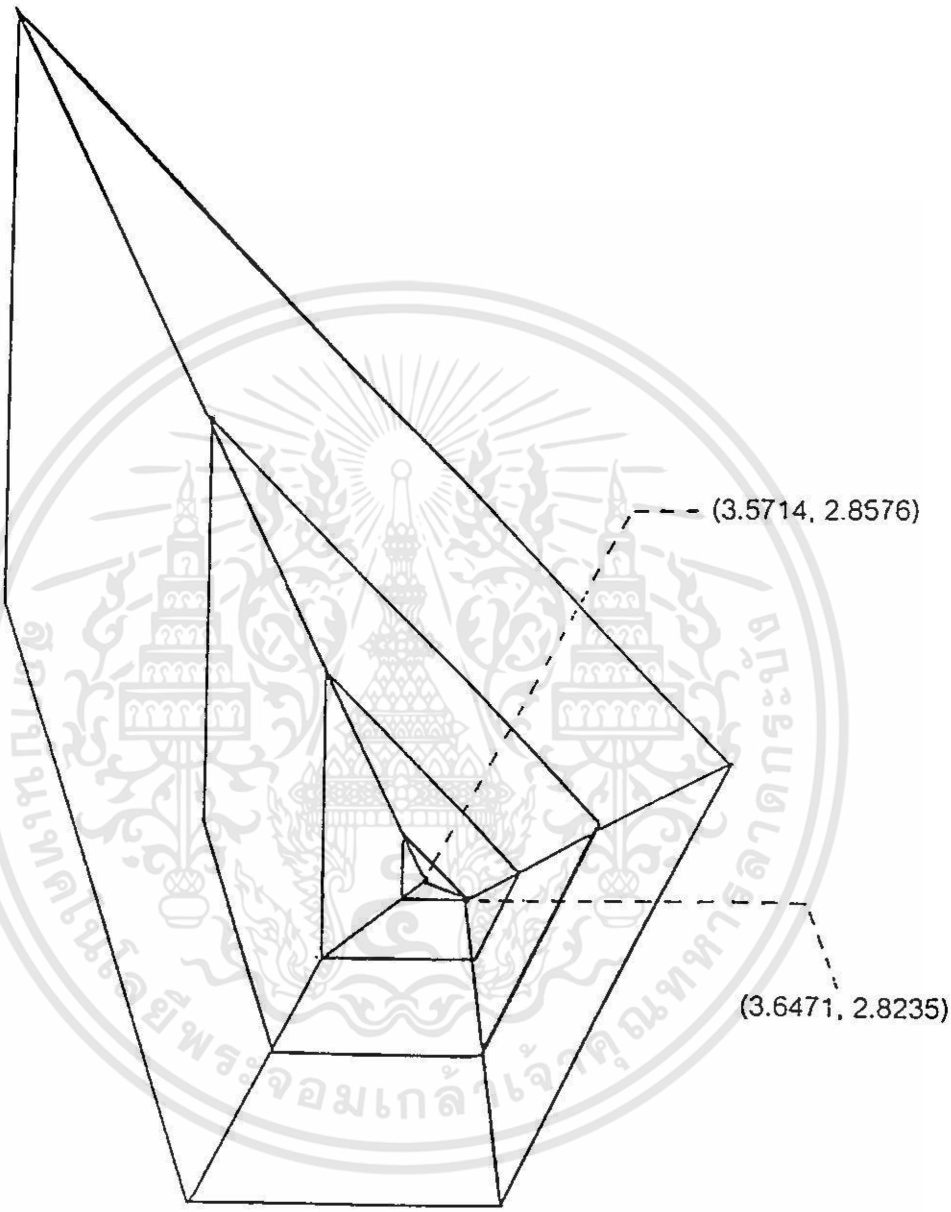
ได้ $y_1 = 0.8\,t$, $y_2 = t$ และ $y_3 = t$

ได้ผลเฉลย y ที่เป็นบวกทั้งหมดแสดงว่าไม่มีทิศทางลด ที่จุด x^3 = (3.5714, 2.8576) เป็นจุดต่ำสุดและมีค่าต่ำสุดเท่ากับ 3.5714 ##

เมื่อใช้ เงื่อนไขส่วนเติมเต็มของตัวแปรช่วย (ดูภาคผนวก ข.) จะได้ผลเฉลย $s_1 = 0.2857$, $s_2 = s_4 = s_6 = 0$, $s_3 = 0.3571$ และ $s_5 = 0.3571$ และค่าต่ำสุดเท่ากับ 3.5714 ตรงกับผลเฉลย ของปัญหาเดิม



ภาพที่ 4.4 โค้งระดับของฟังก์ชันและทิศทางลดเร็วที่สุด



ภาพที่ 4.5 ภาพขยายแสดงจุดต่ำสุดและทิศทางลดเร็วที่สุด

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ปัญหากำหนดการเชิงเส้นที่จะพิจารณาคือปัญหาในแบบของคาร์มาร์คาร์

$$
\text{(P)}\qquad \begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = 0 \\ & \mathbf{e}^T\mathbf{x} = 1 \\ & \mathbf{x} \ge 0 \end{array}
$$

เมื่อ $\mathbf{A} = [a_{ij}]_{m \times n}$ เป็นเมทริกซ์ขนาด $m \times n$

$\mathbf{e}^T = [1\ 1\ 1 \ldots 1]$ เป็น เวกเตอร์ของ 1 ทั้งหมด n จำนวน

$\mathbf{c} = (c_1, c_2, \ldots, c_n) \in R^n$

และ $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in R^n$

จากปัญหาแบบของคาร์มาร์คาร์เราหาปัญหาคู่กัน (dual problem) ของปัญหาคาร์มาร์คาร์ ได้เป็น

$$
\text{(D)}\qquad \begin{array}{ll} \text{maximize} & y_{m+1} \\ \text{subject to} & \sum_{i=1}^{m} a_{ij} y_i + y_{m+1} \le c_j; \quad j = 1, 2, \ldots, n \\ & y_i \in R; \quad i = 1, 2, \ldots, m+1 \end{array}
$$

เมื่อให้ $w = -y_{m+1}$ จะได้ปัญหาในรูปแบบที่เราพิจารณา คือ

$$
\text{(A)}\qquad \begin{array}{ll} \text{minimize} & w \\ \text{subject to} & w \ge \sum_{i=1}^{m} a_{ij} y_i - c_j; \quad j = 1, 2, \ldots, n \\ & y_i \in R; \quad i = 1, 2, \ldots, m+1 \end{array}
$$

ให้ $w_j = \sum_{i=1}^{m} a_{ij} y_i - c_j$

$= L_j(\mathbf{y})$

เมื่อ $\mathbf{y} = (y_1, y_2, \ldots, y_m)$

คือเราหาค่าต่ำสุดของ w เมื่อ w เป็นค่าสูงสุดของ w_j นั่นคือปัญหา

(A') $$\min_{\mathbf{y} \in Y} \max \{L_1(\mathbf{y}), L_2(\mathbf{y}), \ldots, L_m(\mathbf{y})\}, \quad \mathbf{y} \in R^n$$

เมื่อ $L_i(\mathbf{y})$; $i = 1, 2, \ldots, m$ เป็นฟังก์ชันเชิงเส้น

เมื่อแก้ปัญหานี้ได้เท่ากับการแก้ปัญหากำหนดการเชิงเส้นได้ วิธีที่จะใช้ในการแก้ปัญหานี้เลือกใช้วิธีลดเร็วที่สุดหรือวิธีลดซึ่งฟังก์ชันเป้าหมายของปัญหาหาอนุพันธ์ไม่ได้ ดังนั้นผู้วิจัยจึงศึกษาเพื่อปรับวิธีการให้เหมาะสมในการหาทิศทางลดเร็วที่สุดที่จุดที่หาอนุพันธ์ไม่ได้และหาขั้นตอนวิธีในการแก้ปัญหาดังกล่าว

ในการแก้ปัญหาเพื่อความสะดวกเรากำหนดให้จุดเริ่มต้น $\mathbf{x}^0 = (0, 0, \ldots, 0)$ แล้วได้ค่า $f_i(\mathbf{x})$ ซึ่งหาง่าย ค่าของฟังก์ชัน $f_i(\mathbf{x})$ จะอยู่บนระนาบใดระนาบหนึ่งและทิศทางลดเร็วที่สุดเป็นความชันของฟังก์ชันซึ่งเป็นจุดที่หาอนุพันธ์ได้ แล้วหาค่าต่ำสุดในทิศทางลดเร็วที่สุด หลังจากขั้นตอนนี้จะได้จุดต่ำสุดซึ่งเป็นจุดที่ไม่มีอนุพันธ์ โดยที่จุดนี้จะเป็นรอยตัดของ 2 ระนาบหรือมากกว่า 2 ระนาบ ตรงจุดนี้เราจะได้เวกเตอร์หน่วย $\bar{\mathbf{d}}$ ซึ่งเป็นทิศทางลดเร็วที่สุดของฟังก์ชันนูน f ที่จุด $\bar{\mathbf{x}}$ ซึ่งเป็นจุดที่ไม่มีอนุพันธ์ เราทราบว่า $\bar{\mathbf{d}} = -\frac{\xi}{\|\xi\|}$ เมื่อ ξ เป็นซับเกรเดียนต์ที่มีขนาดเล็กที่สุด

ในการแก้ปัญหากำหนดการเชิงเส้นด้วยวิธีลดเร็วที่สุดเราจะดำเนินการตามขั้นตอนหลักๆ 3 ขั้นตอนดังนี้

ขั้นตอนที่ 1 ตรวจดูว่า $\mathbf{x}_k$ เป็นจุดต่ำสุดหรือไม่โดยการหาทิศทาง $\mathbf{u}$ ซึ่งเป็นทิศทางลดเร็วที่สุด ถ้าไม่มีทิศทางลดเร็วที่สุดแสดงว่า $\mathbf{x}_k$ เป็นจุดต่ำสุดแต่ถ้ามีทิศทางลดเร็วที่สุดให้ทำขั้นตอนที่ 2

ขั้นตอนที่ 2 หาค่า t_k ที่ทำให้ $f(\mathbf{x}_k + t\mathbf{u})$ เป็นค่าต่ำสุดในทิศทางของ t โดยวิธีเลือกเส้นต่ำสุด ได้ค่า $t = t_k$

ขั้นตอนที่ 3 ให้ $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k\mathbf{u}$ และหาค่า $f(\mathbf{x}_{k+1})$ แล้วทำขั้นตอนที่ 1

5.2 ข้อเสนอแนะ

เนื่องจากปัญหากำหนดการเชิงเส้นเป็นแบบทางคณิตศาสตร์ซึ่งเป็นที่รู้จักกันดีและมีการประยุกต์ใช้มากมายในการแก้ปัญหาการจัดสรรทรัพยากรที่มีจำกัดให้ได้ประโยชน์สูงสุดหรือปัญหาการวางแผนการผลิตให้ได้กำไรสูงสุดหรือมีต้นทุนต่ำสุด จึงมีผู้วิจัยจำนวนมากพยายามที่จะหาวิธีใหม่ๆที่ง่ายต่อการแก้ปัญหาและลดขั้นตอนวิธีในการแก้ปัญหา เพื่อให้เสียเวลาในการคำนวณน้อยที่สุด

จากการวิจัยนี้ทำให้ได้ความรู้ที่ชัดเจนเกี่ยวกับทิศทางลดและทิศทางลดเร็วที่สุด ณ จุดที่ฟังก์ชันไม่มีอนุพันธ์ ทฤษฎีบทที่เป็นหลักในการพิสูจน์ คือ ทฤษฎีบท 3.1.10 ซึ่งกล่าวเกี่ยวกับความสัมพันธ์ของซับเกรเดียนต์กับอนุพันธ์ทิศทางของฟังก์ชันนูน ผู้วิจัยคิดว่าจะนำไปใช้อ้างอิงในเรื่องอื่นๆได้ (ทฤษฎีบทนี้เท่าที่ศึกษาค้นคว้ายังไม่พบในที่ใด)

อีกเรื่องหนึ่งที่น่ากล่าวถึงก็คือวิธีหาค่าต่ำสุดของฟังก์ชันตัวแปรเดียวที่ฟังก์ชันมีลักษณะเป็นฟังก์ชันเชิงเส้นเป็นช่วงๆ (piecewise linear function) ผู้วิจัยใช้วิธีเลือกเส้นต่ำสุดซึ่งเป็นวิธีที่ผู้วิจัยและอาจารย์ที่ปรึกษาได้คิดขึ้น วิธีนี้มีข้อดีคือได้ผลเฉลยแม่นตรง (exact solution) ในขั้นตอนที่จำกัด

ดังนั้นงานวิจัยเล่มนี้น่าจะเป็นประโยชน์ในการนำความรู้ใหม่ที่ได้ไปใช้ในการแก้ปัญหากำหนดการเชิงเส้นหรือนำขั้นตอนวิธีใหม่ที่ได้ไปประยุกต์ใช้กับงานด้านอุตสาหกรรม วิศวกรรม หรือด้านวิทยาศาสตร์และเทคโนโลยี และที่สำคัญงานวิจัยนี้น่าจะเป็นแนวทางให้กับนักคณิตศาสตร์ในการศึกษากำหนดการเชิงเส้นในรูปแบบอื่นๆ

เอกสารอ้างอิง

[1] Nash, Stephen G. and Sofer, Ariela. Linear and Nonlinear Programming. Singspore : McGraw-Hill International Editions. 1996.

[2] Winston, Wayne L. Operation Research : Applications and Algorithms. 3rd Ed. California : Duxbury Press. 1994.

[3] Fang, Shu-Cheng and Puthenpura, Sarat. Linear Optimization and Extensions : Theory and Algorithms. New jersey : Prentice - Hall,Inc. 1993.

[4] C.Charalambous. and A.R.Conn. " An efficient method to solve the minimax problem directly." Siam Journal Numerical Analysis., vol.15, no.1, February 1978. pp.162-187.

[5] Vardi, A. "New Minimax Algorithm." Journal of Optimization Theory and Applications., vol.75, no.3, December 1992. pp. 613-634.

[6] Rajasekera, J.R., and Fang, S.C. "On the convex programming approach to linear programming." Operation Research Letters., vol.10, no.6, August 1991. pp.309-312.

[7] Rajasekera, J.R., and Fang, S.C. "Deriving an Unconstrained Convex Program for Linear Programming." Journal of Optimization Theory and Applications., vol.75, no.3, December 1992. pp. 603-612.

[8] Xu, Z.K., and Fang, S.C. "Unconstrained Convex Programming Approach to Linear Programming." Journal of Optimization Theory and Applications., vol.86, no.3, September 1995. pp. 745-752.

[9] Luenberger, David G. Linear and Nonlinear Programming. 2nd Ed. Massachusetts : Addisen-Wesley Publishing Company,Inc. 1989.

[10] Gass, Sual I. Linear Programming. 5th Ed. Singspore : McGraw-Hill International Editions. 1986.

[11] Taha, Hamdy A. Operations Research. 5th Ed. New Jersey : Macmillian Publishing Company. 1992.

[12] Pressini, Anthony L. et.al. The Mathematics of Nonlinear Programming. New York : Springer-Verlag New York Inc. 1988.

[13] Bazaraa, Moktar S. and C.M. Shetty. **Nonlinear programming : Theory and Algorithms.** New York : John Wiley & Sons. 1979.

ภาคผนวก ก.

ทฤษฎีบทของการคู่กันเป็นเรื่องที่มีความสำคัญในเรื่องการโปรแกรมเชิงเส้นจึงจะขอกล่าวถึงทฤษฎีบทที่มีความเกี่ยวข้องกับงานวิจัยดังต่อไปนี้

ปัญหาคู่กัน

ให้ปัญหาเดิม (primal problem) คือปัญหา A มีแบบเป็น

$$\begin{aligned} \text{minimize} \quad & P = \mathbf{c}^T\mathbf{x} + d \\ \text{subject to} \quad & A\mathbf{x} \le \mathbf{b} \\ & \mathbf{x} \ge \mathbf{0} \end{aligned}$$

ปัญหาคู่กัน (dual problem) คือปัญหา B มีแบบเป็น

$$\begin{aligned} \text{minimize} \quad & Q = \mathbf{b}^T\mathbf{y} + d \\ \text{subject to} \quad & A^T\mathbf{y} \ge \mathbf{c} \\ & \mathbf{y} \ge \mathbf{0} \end{aligned}$$

เราอาจเรียกกลับกัน คือเรียกปัญหา B เป็นปัญหาเดิมและเรียกปัญหา A ว่าเป็นปัญหาคู่กันก็ได้

ทฤษฎีบทการคู่กันแบบอ่อน

ถ้าปัญหา A และ ปัญหา B ต่างก็ไม่มีข้อขัดแย้งและ เซต S และเซต T เป็นเซตนูนแล้วจะได้

1. $\mathbf{b}^T\mathbf{y}+d \;\ge\; \mathbf{c}^T\mathbf{x}+d$ ทุกจุด $\mathbf{x}\in S$ และ $\mathbf{y}\in T$
2. มีค่า c ที่ทำให้ $\mathbf{c}^T\mathbf{x}+d \le c \le \mathbf{b}^T\mathbf{y}+d$ ทุกจุด $\mathbf{x}\in S$ และ $\mathbf{y}\in T$
3. ปัญหา A และ ปัญหา B ต่างก็มีขอบเขต
4. S^* และ T^* ไม่เป็นเซตว่าง
5. $\mathbf{b}^T\mathbf{y}+d \;=\; \mathbf{c}^T\mathbf{x}+d$ ก็ต่อเมื่อ $\mathbf{x}^T\mathbf{v}=0$ และ $\mathbf{y}^T\mathbf{u}=0$ ซึ่งในกรณีนี้จะได้ว่า $\mathbf{x}\in S^*$ และ $\mathbf{y}\in T^*$ คุณสมบัติข้อนี้เราเรียกว่า " *เงื่อนไขส่วนเติมเต็มของตัวแปรช่วย* "
6. S^* และ T^* เป็นเซตนูน

ทฤษฎีบทการคู่กันแบบเข้ม

ถ้าปัญหา A มีผลเฉลยที่เป็นไปได้ จะได้ว่าปัญหา A มีขอบเขตก็ต่อเมื่อปัญหา B มีผลเฉลยที่เป็นไปได้

เงื่อนไขคารุช-คุน-ทักเกอร์

สำหรับปัญหากำหนดการไม่เชิงเส้นในรูปแบบ

$$\min_{\mathbf{x}\in R^n} \quad f(\mathbf{x})$$

$$\textit{subject to} \quad g_j(\mathbf{x}) \le 0 \quad , \; j=1,2,\dots,l$$

$$h_i(\mathbf{x}) = 0 \quad , \; i=1,2,\dots,m$$

จะได้

ทฤษฎีบทของ คารุช-คุน-ทักเกอร์

สมมุติ $\mathbf{x}^*$ สอดคล้องกับเงื่อนไขบังคับ ถ้า $\mathbf{x}^*$ เป็นจุดต่ำสุดของปัญหากำหนดการไม่เชิงเส้นแล้ว จะมี $\mu^* \in R^l$ และ $\lambda^* \in R^m$ ซึ่งทำให้

ความชัน :

$$\nabla f(\mathbf{x}^*) + \sum_{j=1}^{l} \mu_j^* \nabla g_j(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i^*(\mathbf{x}^*) = 0$$

เงื่อนไขส่วนเติมเต็มของตัวแปรช่วย :

$$\mu_j^* g_j(\mathbf{x}^*) = 0 \quad , \; j=1,2,\dots,l$$

ค่าที่ไม่เป็นลบ (Nonnegativity) :

$$\mu_j \ge 0 \quad , \; j=1,2,\dots,l$$

เงื่อนไขบังคับเริ่มต้น (Original Constraints) :

$$g_j(\mathbf{x}^*) \le 0 \quad , \; j=1,2,\dots,l$$

$$h_i(\mathbf{x}^*) = 0 \quad , \; i=1,2,\dots,m$$

ภาคผนวก ข.

จากตัวอย่าง 1 และ 2 ในบทที่ 4 นั้นเรานำผลเฉลยที่ได้มาใช้คุณสมบัติ เงื่อนไขส่วนเติมเต็มของตัวแปรช่วย เพื่อหาผลเฉลยของปัญหาสมการเชิงเส้นดังนี้

ตัวอย่าง 1 เรามีปัญหากำหนดการเชิงเส้น

(P) minimize $C = -s_1 - 2s_2 + 4s_5$

subject to

$$s_2 - s_3 = 0 \quad ...(1)$$
$$2s_1 - 2s_2 + 4s_3 - 4s_5 = 0 \quad ...(2)$$
$$s_1 + 2s_2 + s_4 - 4s_5 = 0 \quad ...(3)$$
$$s_1 + s_2 + s_3 + s_4 + s_5 = 1 \quad ...(4)$$
$$s_i \geq 0$$

ปัญหาคู่กันของปัญหาเดิมคือ

(D) maximize x_4

subject to

$$2x_2 + x_3 + x_4 \leq -1$$
$$x_1 - 2x_2 + 2x_3 + x_4 \leq -2$$
$$-x_1 + 4x_2 + x_4 \leq 0$$
$$x_3 + x_4 \leq 0$$
$$-4x_2 - 4x_3 + x_4 \leq 4$$
$$x_i \text{ free}$$

จากปัญหาคู่กันให้ $w = -x_4$ จะได้

minimize w

subject to

$$w \geq 1 + 2x_2 + x_3$$
$$w \geq 2 + x_1 - 2x_2 + 2x_3$$
$$w \geq -x_1 + 4x_2$$
$$w \geq x_3$$
$$w \geq -4 - 4x_2 - 4x_3$$
$$x_i \text{ free}$$

จากตัวอย่าง 1 ในบทที่ 4 เราได้ $x^3 = x^* = (-0.2972, -0.0743, -0.9257)$

และ $w = -x_4 = 0$ ดังนั้น $x_4 = 0$

ดังนั้น $x_1 = -0.2972$

$x_2 = -0.0743$

$x_3 = -0.9257$

$x_4 = 0$

ใช้ เงื่อนไขส่วนเติมเต็มของตัวแปรช่วยเราได้

1. $(2x_2 + x_3 + x_4 + 1)\, s_1 = 0$

$(-0.1486 - 0.9257 + 0 + 1)\, s_1 = 0$

$(-0.0743)\, s_1 = 0$ ดังนั้น $s_1 = 0$

2. $(x_1 - 2x_2 + 2x_3 + x_4 + 2)s_2 = 0$

$(-0.2972 + 0.1486 - 1.8514 + 0 + 2)\, s_2 = 0$

$(0)\, s_2 = 0$ ดังนั้น $s_2 \neq 0$

3. $(-x_1 + 4x_2 + x_4)\, s_3 = 0$

$(0.2972 - 0.2972 + 0)\, s_3 = 0$

$(0)\, s_3 = 0$ ดังนั้น $s_3 \neq 0$

4. $(x_3 + x_4)s_4 = 0$

$(-0.9257 + 0)s_4 = 0$ ดังนั้น $s_4 = 0$

5. $(-4x_2 - 4x_3 + x_4 - 4)\, s_5 = 0$

$(0.2972 + 3.7028 + 0 - 4)s_5 = 0$ ดังนั้น $s_5 \neq 0$

เราได้ $s_1 = 0$, $s_2 \neq 0$, $s_3 \neq 0$, $s_4 = 0$ และ $s_5 \neq 0$

แทนค่า ที่ได้ในเงื่อนไขของสมการ (1) - (4)

จะได้

$$s_2 - s_3 = 0$$

$$-2s_2 + 4s_3 - 4s_5 = 0$$

$$2s_2 - 4s_5 = 0$$

$$s_2 + s_3 + s_5 - 1 = 0$$

แก้สมการทั้ง 4 ได้ค่า

$s_1 = s_4 = 0$, $s_2 = s_3 = 0.4$ และ $s_5 = 0.2$

หาค่าต่ำสุด $C = -s_1 - 2s_2 + 4s_5$

ได้ค่าต่ำสุดเท่ากับ 0

ดังนั้นผลเฉลยของสมการคือ $s_1 = s_4 = 0$, $s_2 = s_3 = 0.4$ และ $s_5 = 0.2$ และค่าต่ำสุดเท่ากับ 0 ##

ตัวอย่าง 2 เรามีปัญหากำหนดการเชิงเส้นคือ

(P) minimize $C = -3s_2 + 5s_4 + 10s_5 + 8s_6$

subject to $-1.5s_2 - s_3 - s_4 + s_5 + 2s_6 = 0$ ...(1)

$-1.25s_1 - 0.5s_2 + s_4 + s_5 - s_6 = 0$ ...(2)

$s_1 + s_2 + s_3 + s_4 + s_5 + s_6 = 1$ ...(4)

$s_i \geq 0$

ปัญหาคู่กันของปัญหาเดิมคือ

(D) maximize x_3

subject to $-1.25x_2 + x_3 \leq 0$

$-1.5x_1 - 0.5x_2 + x_3 \leq -3$

$-x_1 + x_3 \leq 0$

$-x_1 + x_2 + x_3 \leq 5$

$x_1 + x_2 + x_3 \leq 10$

$2x_1 - x_2 + x_3 \leq 8$

x_i free

จากปัญหาคู่กันให้ $w = -x_3$ จะได้

minimize w

subject to $w \geq -1.25x_2$

$w \geq 3 - 1.5x_1 - 0.5x_2$

$w \geq -x_1$

$w \geq -5 - x_1 + x_2$

$w \geq -10 + x_1 + x_2$

$w \geq -8 + 2x_1 - x_2$

x_i free

จากตัวอย่าง 2 ในบทที่ 4

เราได้ $x^3 = x^* = (3.5714, 2.8516)$

และ $w = -x_3 = 3.5714$ ดังนั้น $x_3 = 3.5714$

ดังนั้น $x_1 = 3.5714$

$x_2 = 2.8516$

$x_3 = 3.5714$

ใช้ เงื่อนไขส่วนเติมเต็มของตัวแปรช่วยเราได้

1. $(-1.25x_2 + x_3)\, s_1 = 0$

$(0)\, s_1 = 0$ ดังนั้น $s_1 \neq 0$

2. $(-1.5x_1 - 0.5x_2 + x_3 + 3)s_2 = 0$

$(-5.3568 - 1.4288 + 3.5712 + 3)\, s_2 = 0$

$(-0.2144)\, s_2 = 0$ ดังนั้น $s_2 = 0$

3. $(-x_1 + x_3)\, s_3 = 0$

$(0)\, s_3 = 0$ ดังนั้น $s_3 \neq 0$

4. $(-x_1 + x_2 + x_3 - 5)s_4 = 0$

$(-2.1424)s_4 = 0$ ดังนั้น $s_4 = 0$

5. $(x_1 + x_2 + x_3 - 10)\, s_5 = 0$

$(0)s_5 = 0$ ดังนั้น $s_5 \neq 0$

6. $(2x_1 - x_2 + x_3 - 8)\, s_5 = 0$

$(6.9984)s_5 = 0$ ดังนั้น $s_5 = 0$

เราได้

$s_1 \neq 0,\ s_2 = 0,\ s_3 \neq 0,\ s_4 = 0,\ s_5 \neq 0$ และ $s_6 = 0$

แทนค่า ที่ได้ในเงื่อนไขของสมการ (1) - (3)

จะได้

$$-s_3 + s_5 = 0$$

$$-1.25s_1 + s_5 = 0$$

$$s_1 + s_3 + s_5 - 1 = 0$$

แก้สมการทั้ง 4 ได้ค่า $s_1 = 0.2857$, $s_2 = s_4 = s_6 = 0$, $s_3 = 0.3571$ และ $s_5 = 0.3571$

หาค่าต่ำสุด $C = -3s_2 + 5s_4 + 10s_5$ ได้ค่าต่ำสุดเท่ากับ 3.5714

ดังนั้นผลเฉลยของสมการคือ $s_1 = 0.2857$, $s_2 = s_4 = s_6 = 0$, $s_3 = 0.3571$ และ $s_5 = 0.3571$ และค่าต่ำสุดเท่ากับ 3.5714 ##

ประวัติผู้เขียน

นางสาวดุษยา ปานประสิทธิ์ เกิดเมื่อวันที่ 16 เมษายน 2516 ที่จังหวัดระยอง สำเร็จการศึกษาวิทยาศาสตรบัณฑิต สาขาวิชาคณิตศาสตร์ จากมหาวิทยาลัยบูรพา ปีการศึกษา 2538