

การออกแบบระบบเชิงเส้นโดยใช้วิธีสมการเมทริกซ์เชิงเส้น

LINEAR SYSTEM DESIGN: AN LMI APPROACH

นายสิทธิโชค ใจยอด

รหัส 47364146

ห้องสมุดคณะวิศวกรรมศาสตร์
วันที่รับ..... 5 เม.ย. 2553
เลขทะเบียน..... 14999848
เลขเรียกหนังสือ..... ๗/ร.
มหาวิทยาลัยนเรศวร ๙๔๒๒ ๗

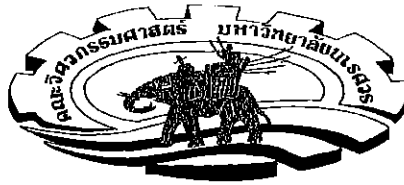
2550

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

ปีการศึกษา 2550



ใบรับรองโครงการวิศวกรรม

หัวข้อโครงการ การออกแบบระบบเชิงเส้นโดยใช้วิธีสมการเมทริกซ์เชิงเส้น
ผู้ดำเนินโครงการ นายสิทธิโชค ใจยอด รหัส 47364146
อาจารย์ที่ปรึกษา ผศ.ดร.ธนิต มาลากร
สาขาวิชา วิศวกรรมไฟฟ้า
ภาควิชา วิศวกรรมไฟฟ้าและคอมพิวเตอร์
ปีการศึกษา 2550

.....
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธนบุรี อนุมัติให้โครงการฉบับนี้เป็นส่วนหนึ่งของ
การศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมไฟฟ้า
คณะกรรมการสอบโครงการวิศวกรรม

.....ประธานกรรมการ
(ผศ.ดร.ธนิต มาลากร)

.....กรรมการ
(ดร.อัศรพันธ์ วงศ์กังแห)

.....กรรมการ
(ดร.แคทรียา สุวรรณศรี)

หัวข้อโครงการ	การออกแบบระบบเชิงเส้น โดยใช้วิธีอสมการเมทริกซ์เชิงเส้น
ผู้ดำเนินโครงการ	นายสิทธิโชค ใจยอด รหัส 47364146
อาจารย์ที่ปรึกษา	ผศ. ดร. ธนิต มาลากร
สาขาวิชา	วิศวกรรมไฟฟ้า
ภาควิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์
ปีการศึกษา	2550

.....

บทคัดย่อ

หลากหลายปัญหาในทฤษฎีระบบควบคุม รวมทั้งในการประยุกต์ต่าง ๆ สามารถพิจารณาเป็นปัญหาค่าเหมาะที่สุดแบบคอนเวกซ์ซึ่งสามารถจัดให้อยู่ในรูปแบบของอสมการเมทริกซ์เชิงเส้นได้ โดยทั่วไป ปัญหาเหล่านี้ไม่สามารถคำนวณหาผลเฉลยเชิงวิเคราะห์ แต่สามารถหาผลเฉลยเชิงตัวเลขได้ด้วยตัวแก้สมการเมทริกซ์เชิงเส้น โครงการฉบับนี้นำเสนอปัญหาที่น่าสนใจบางปัญหาในแวดวงของระบบควบคุม รวมทั้งการแปลงปัญหาเหล่านั้นให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้น โดยการใช้จอร์จคอมพลิเมนต์ ผลการทดลองเชิงตัวเลขพบว่าผลเฉลยที่ได้จากปัญหาดังต้นเหมือนกับผลเฉลยที่ได้จากการแปลงปัญหาให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้น

Project Title	Linear System Design: an LMI Approach		
Name	Mr. Sittichoke	Jaiyod	ID. 47364146
Project Advisor	Asst. Prof. Tanit	Malakorn, PhD	
Major	Electrical Engineering		
Department	Electrical and Computer Engineering		
Academic Year	2007		

.....

ABSTRACT

A wide variety of problems in control system theory including many practical applications can be cast or recast as convex optimization problems involving linear matrix inequalities (LMIs). In general, these problems do not have analytical solutions but can be solved numerically using LMI solvers. This project presents a list of some interesting problems in control literature associated with the reduction of such problems to LMI problems via Schur complement. Numerical experiments show that the solutions of the original problem and the LMI counterpart are identical.

กิตติกรรมประกาศ

หากปราศจากความกรุณาและความช่วยเหลือจากอาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. ธนิต มาลากร โครงการฉบับนี้คงจะไม่สำเร็จลุล่วงไปได้ด้วยดี ผู้จัดทำโครงการจึงใคร่ขอขอบคุณอาจารย์เป็นอย่างสูง ที่ได้ให้ความช่วยเหลือในการค้นหาข้อมูลที่เป็นต่างๆ รวมทั้งคำแนะนำที่มีคุณค่าและข้อคิดที่เป็นประโยชน์อย่างมากในการทำโครงการ ตลอดจนเสียสละเวลาอันมีค่าในการตรวจและแก้ไขข้อบกพร่องต่างๆ ในโครงการฉบับนี้

นอกจากนี้ ผู้จัดทำโครงการใคร่ขอขอบคุณ อ. ดร.อัศรพันธ์ วงศ์กังแห และ อ. ดร. แคทรียา สุวรรณศรี ที่ได้สละเวลามาเป็นกรรมการสอบโครงการ อ่านร่างโครงการ รวมทั้งให้คำแนะนำถึงข้อบกพร่องต่าง ๆ จนทำให้โครงการฉบับนี้เสร็จสมบูรณ์ด้วยดี

สุดท้ายนี้ ขอกราบขอบพระคุณบิดา มารดาและญาติพี่น้องทุกท่านที่ได้อบรมสั่งสอนให้เห็นถึงคุณค่าของการศึกษา รวมไปถึงคุณอาจารย์ทุกท่านที่ได้ให้ความรู้ คอยอบรมสั่งสอน และให้คำปรึกษาที่มีคุณค่ามาโดยตลอด และที่ลืมไม่ได้คือ เพื่อนๆ ทุกคนที่คอยให้กำลังใจเสมอมา ซึ่งผู้จัดทำโครงการขอขอบใจมา ณ ที่นี้

ผู้จัดทำโครงการ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญ	ง
สารบัญรูป	ช
บทที่ 1 บทนำ	
1.1 ที่มาและความสำคัญของโครงการ	1
1.2 วัตถุประสงค์ของโครงการ	3
1.3 ขอบเขตโครงการ	3
1.4 ขั้นตอนการดำเนินงาน	4
1.5 ผลที่คาดว่าจะได้รับ	5
1.6 งบประมาณ	5
บทที่ 2 หลักการและทฤษฎีเบื้องต้น	
2.1 พีชคณิตเชิงเส้น (Linear Algebra)	6
2.1.1 สนาม, ปริภูมิเวกเตอร์และปริภูมิย่อย (Field, Vector space and Subspace)	7
2.1.2 ตัวดำเนินการเชิงเส้นในปริภูมิเวกเตอร์ (Linear operator in vector space)	12
2.1.3 ปริภูมิเวกเตอร์นอร์ม (Normed vector space)	15
2.1.4 ตัวดำเนินการเชิงเส้นในปริภูมิเวกเตอร์นอร์ม (Linear operator in normed vector space)	17
2.1.5 ทฤษฎีเมทริกซ์ (Matrix theory)	19
2.2 ความเป็นคอนเวกซ์ (Convexity)	30
2.2.1 เซตคอนเวกซ์และฟังก์ชันคอนเวกซ์ (Convex set)	31
2.2.2 ฟังก์ชันคอนเวกซ์ (Convex function)	31

สารบัญ(ต่อ)

	หน้า
2.3 ทฤษฎีพื้นฐานของระบบควบคุม (Basic Control System Theory)	34
2.3.1 การวิเคราะห์เชิงปริภูมิสถานะ (State space analysis)	34
2.3.1.1 การหาผลเฉลยของระบบ (System Solution)	36
2.3.1.2 ความมีเสถียรภาพ (Stability)	38
2.3.1.3 ความสามารถควบคุมได้และความสามารถสังเกตได้	40
2.3.2 การวิเคราะห์เชิงฟังก์ชันถ่ายโอน (Transfer function approach)	42
2.3.2.1 การเชื่อมต่อกันของระบบย่อย	43
2.3.2.2 ความมีเสถียรภาพ (Stability)	44
2.3.2.3 การวิเคราะห์ผลตอบสนองของระบบอันดับที่สอง	46
2.3.2.4 คุณลักษณะของผลตอบสนองชั่วขณะของระบบอันดับที่สอง	49
2.3.3 ความสัมพันธ์ระหว่างปริภูมิสถานะกับฟังก์ชันถ่ายโอน	50
บทที่ 3 หลักการและทฤษฎีของอสมการเมทริกซ์เชิงเส้น	
3.1 นิยามของอสมการเมทริกซ์เชิงเส้น	52
3.2 คอนเวกซ์ของอสมการเมทริกซ์เชิงเส้น	52
3.3 อสมการเมทริกซ์เชิงเส้นหลายๆ สมการ	53
3.4 ชัวร์คอมพลีเมนต์ (Schur complement)	54
3.5 การแปลงรูปสมการให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้น	54
3.5.1 สมการของลีอาปูนอฟ (Lyapunov equation)	54
3.5.2 สมการรีคคาติเชิงพีชคณิต (Algebraic Riccati Equation: ARE)	55
3.5.3 อสมการของค่าขอบเขตจริง (Bounded real) ที่ใช้ในระบบควบคุมแบบ H_∞	55
3.5.4 อสมการของความเป็นบวกจริง (Positive real)	56
3.5.5 สมการของพหุนามที่เป็นบวก (Positive polynomials)	56

สารบัญ(ต่อ)

	หน้า
3.6 การใช้ตัวแก้สมการเมทริกซ์เชิงเส้น (LMI Solver) ในโปรแกรมแมทแล็บ	57
3.6.1 รูปแบบคำสั่งที่สำคัญ	58
3.6.2 ตัวอย่างโปรแกรมที่ใช้ตัวแก้สมการเมทริกซ์เชิงเส้น	60
บทที่ 4 การออกแบบระบบเชิงเส้นด้วยวิธีอสมการเมทริกซ์เชิงเส้น	
4.1 การตรวจสอบเสถียรภาพด้วยวิธีลียาปูนอฟ (Lyapunov Stability Testing)	61
4.2 การหาเสถียรภาพของระบบที่มีเมทริกซ์ระบบ A_k หลายตัว	63
4.3 การออกแบบตัวป้อนกลับสถานะ	66
4.4 การวิเคราะห์ระบบควบคุมที่ใช้วิธี H_∞ (H_∞ Control Analysis)	70
บทที่ 5 บทสรุปและข้อเสนอแนะ	
5.1 สรุปผลการดำเนินงาน	75
5.2 ข้อเสนอแนะ	75
เอกสารอ้างอิง	76
ภาคผนวก ก คู่มือการติดตั้งและใช้งานตัวแก้สมการเมทริกซ์เชิงเส้น	77
ประวัติผู้ทำโครงการ	88

สารบัญรูป

รูปที่	หน้า
2.1 ระบบในรูปแบบสัญญาณเข้า-สัญญาณออก	6
2.2 ฟังก์ชันที่ไม่คอนเวกซ์	30
2.3 ฟังก์ชันที่คอนเวกซ์	30
2.4 (a) เซตที่คอนเวกซ์	31
2.4 (b) เซตที่ไม่คอนเวกซ์	31
2.5 ฟังก์ชันคอนเวกซ์	32
2.6 ฟังก์ชันคอนเวกซ์และฟังก์ชันคอนเคฟ	33
2.7 ระบบสปริง-มวล-ตัวหน่วง (Spring-mass-damper system)	34
2.8 ระบบควบคุมที่ต่อแบบอนุกรม (Cascade Control System)	43
2.9 ระบบควบคุมที่ต่อแบบขนาน (Parallel Control System)	43
2.10 ระบบควบคุมที่ต่อแบบป้อนกลับ (Feedback Control System)	44
2.11 ตำแหน่งขั้วและผลตอบสนองของระบบอันดับหนึ่ง	45
2.12 ตำแหน่งขั้วและผลตอบสนองของระบบอันดับสอง	46
2.13 ระบบสปริง-มวล-ตัวหน่วง (Spring-mass-damper system)	46
2.14 ผลตอบสนองของระบบอันดับสองที่ค่าอัตราการหน่วง (ζ) ต่างๆ	48
2.15 ลักษณะการตอบสนองชั่วขณะของระบบอันดับที่สอง $0 < \zeta < 1$	49
3.1 ค่าต่ำสุดของฟังก์ชันพหุนาม	57
4.1 ระบบสปริง-มวล-ตัวหน่วง (Spring-mass-damper system)	63
4.2 ระบบควบคุมแบบป้อนกลับ	66
4.3(a) ขั้ว (Pole) และศูนย์ (Zero) ของระบบก่อนออกแบบ	68
4.3(b) ขั้ว (Pole) และศูนย์ (Zero) ของระบบหลังออกแบบ	69
4.4(a) ผลตอบสนองต่อสัญญาณหนึ่งหน่วยของระบบก่อนออกแบบ	69
4.4(b) ผลตอบสนองต่อสัญญาณหนึ่งหน่วยของระบบหลังออกแบบ	69
4.5 สัญญาณควบคุมของระบบหลังการออกแบบ	70
4.6 ผลตอบสนองของระบบในโดเมนความถี่	72

สารบัญรูป(ต่อ)

รูปที่	หน้า
4.7 ค่าเอกฐานสูงสุด (Maximum singular value) ของระบบ	72
4.8 ระบบทางกายภาพของแขนกลหุ่นยนต์แบบอ่อนตัว (Flexible Robot Arm)	73
4.9 ผลตอบสนองของระบบในโดเมนความถี่	74



บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของโครงการ

อสมการเมทริกซ์เชิงเส้น (Linear matrix inequality: LMI) ได้ถูกนำมาใช้วิเคราะห์ระบบพลวัต (Dynamical Systems) เป็นเวลายาวนาน จากหลักฐานที่ปรากฏ พบว่า ประมาณปี ค.ศ. 1890 ลีอาปูนอฟ (Lyapunov) ได้แสดงว่าสมการอนุพันธ์เชิงเส้นของระบบพลวัต

$$\frac{d}{dt}x(t) = Ax(t) \quad (1.1)$$

จะมีเสถียรภาพภายใน (กล่าวคือ ตัวแปรสถานะ x ทุกตัวจะเคลื่อนที่เข้าสู่จุดสมดุลเมื่อเวลาผ่านไป) ก็ต่อเมื่อสามารถหาเมทริกซ์ที่เป็นบวกแน่นอน¹ $P > 0$ ซึ่งสอดคล้องกับอสมการ

$$A^T P + PA < 0 \quad (1.2)$$

เงื่อนไขความเป็นบวกแน่นอน $P > 0$ และอสมการใน (1.2) ข้างต้น จัดว่าเป็น “อสมการเมทริกซ์เชิงเส้น หรือ LMI” ประเภทหนึ่ง

ลีอาปูนอฟได้แสดงให้เห็นด้วยว่าอสมการเมทริกซ์เชิงเส้นที่ได้นั้นสามารถคำนวณหาผลเฉลยได้โดยการนิยามเมทริกซ์ตัวแปรใหม่ขึ้นมาและเขียนแทนเมทริกซ์ตัวแปรนี้ด้วย Q โดยที่ Q ต้องเป็นเมทริกซ์ที่เป็นบวกแน่นอน ดังนั้นอสมการเมทริกซ์เชิงเส้นในอสมการที่ (1.2) จึงสามารถเขียนใหม่ได้เป็น

$$A^T P + PA = -Q < 0 \quad (1.3)$$

ดังนั้น ระบบพลวัตที่เขียนแทนด้วยสมการที่ (1.1) จะมีเสถียรภาพภายใน (ตัวแปรสถานะทุกตัวเข้าสู่จุดสมดุล เมื่อเวลาผ่านไป) ก็ต่อเมื่อเมทริกซ์ P ที่คำนวณได้จากสมการที่ (1.3) นั้นมีความเป็นบวกแน่นอน สำหรับทุก ๆ เมทริกซ์ที่เป็นบวกแน่นอนใด ๆ Q

ต่อมาในปี 1940 ลูเอ (Lur'e), พอสนิคอฟ (Postnikov) และทีมวิจัยในกลุ่มโซเวียตยูเนียนได้ประยุกต์วิธีการของลีอาปูนอฟมาใช้ในแก้ปัญหาทางวิศวกรรมระบบควบคุม โดยเฉพาะอย่างยิ่งนำมาใช้ในการแก้ปัญหาเกี่ยวกับเสถียรภาพของระบบควบคุมที่มีความไม่เชิงเส้นปรากฏอยู่ในอุปกรณ์การวัด โดยที่เกณฑ์ของความมีเสถียรภาพของระบบนั้น จัดให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นได้ อสมการที่ได้เหล่านี้สามารถลดรูปให้อยู่ในแบบของอสมการพหุนาม (Polynomial inequalities) ซึ่งสามารถคำนวณหาผลเฉลยได้โดยไม่ต้องยากนักสำหรับระบบขนาดเล็ก (ระบบอันดับสอง หรืออันดับสามเท่านั้น)

¹ เมทริกซ์สมมาตรกำลังจริง A ที่มีมิติ $n \times n$ เรียกว่าเป็นเมทริกซ์บวกแน่นอน (Positive definite matrix) (ซึ่งนิยามเขียนแทนด้วย $A > 0$) ถ้า $x^T A x > 0$ ทุกค่า $x \neq 0$ ใน $\mathbb{R}^n$ (ผู้อ่านสามารถอ่านรายละเอียดเกี่ยวกับนิยามทางคณิตศาสตร์เหล่านี้ได้ในบทที่ 2)

ในปี 1960 ยาคุโบวิช (Yakubovich), โปโปฟ (Popov), คาลมาน (Kalman), และทิมวิจัย ประสบผลสำเร็จในการลดรูปผลเฉลยของสมการเมทริกซ์เชิงเส้นที่เป็นปัญหาของลูเอ ลงมาเป็นการทำกระบวนการเชิงกราฟฟีกอย่างง่าย ซึ่งต่อมากระบวนการเชิงกราฟฟีกนี้ถูกเรียกว่า บทตั้งค่าจริงบวก หรือ Positive real (PR) lemma ผลที่ได้จากบทตั้งดังกล่าวถูกนำไปประยุกต์ใช้กับระบบที่มีอันดับสูงๆ ได้เป็นอย่างดี

ต่อมาในปี ค.ศ. 1970 เป็นที่ทราบกันอย่างแพร่หลายว่าสมการเมทริกซ์เชิงเส้นที่ปรากฏในบทตั้งค่าจริงบวกนั้นสามารถหาคำตอบได้ด้วยอีกวิธีหนึ่งที่น่าสนใจจากวิธีการกระบวนการเชิงกราฟฟีก นั่นคือการหาคำตอบผ่านทางผลเฉลยของสมการริคคาติเชิงพีชคณิต (Algebraic Riccati Equation: ARE) ต่อมาในปี ค.ศ. 1971 งานวิจัยของวิลเลม (J.C. Willems) ได้แสดงให้เห็นถึงความสมมูลกัน (ภายใต้เงื่อนไขพิเศษบางประการที่ยอมรับได้) ระหว่างสมการเมทริกซ์เชิงเส้น

$$\begin{bmatrix} A^T P + PA + Q & PB \\ B^T P & R \end{bmatrix} \geq 0 \quad (1.4)$$

และสมการริคคาติเชิงพีชคณิต (ARE)

$$A^T P + PA - PBR^{-1}B^T P + Q = 0 \quad (1.5)$$

ซึ่งการหาผลเฉลยของ ARE นั้น สามารถหาได้จากการแยกค่าเฉพาะ (Eigen-decomposition) ของเมทริกซ์ฮามิลโทเนียน (Hamiltonian matrix) ได้

ดังนั้น ตั้งแต่ปี ค.ศ. 1971 เป็นต้นมา การคำนวณหาผลเฉลยของสมการเมทริกซ์เชิงเส้นสามารถกระทำได้หลายวิธี ได้แก่ การคำนวณหาผลเฉลยโดยตรง (ซึ่งเหมาะสำหรับระบบขนาดเล็กที่มีอันดับไม่เกินสาม) หรือ การคำนวณโดยใช้กระบวนการเชิงกราฟฟีก หรือ โดยใช้การหาผลเฉลยของ ARE ซึ่งกระบวนการทั้งสามที่กล่าวมานี้ จัดว่าเป็นกระบวนการเชิงวิเคราะห์ (Analytic methods) ซึ่งสามารถคำนวณหาผลเฉลยที่อยู่ในรูปทั่วไปได้สำหรับสมการเมทริกซ์เชิงเส้นที่มีลักษณะโครงสร้างเฉพาะ

ปัญหาของสมการเมทริกซ์เชิงเส้นที่ปรากฏในระบบควบคุมนั้นสามารถจัดให้อยู่ในรูปของปัญหาการหาค่าเหมาะที่สุดแบบคอนเวกซ์ (Convex optimization problem) ได้ ยกตัวอย่างเช่น ในปัญหาการลดอันดับของระบบ (Model reduction) ที่ต้องการประมาณระบบเสถียรที่กำหนดให้ด้วยระบบที่เสถียรเช่นกันแต่มีอันดับที่ต่ำกว่า การคำนวณหาระบบที่ดีที่สุดที่ใช้ในการประมาณเช่นนี้ สามารถแปลงให้อยู่ในรูปของปัญหาการหาค่าเหมาะที่สุดของฟังก์ชันตรรกยะ (Rational function) นอกจากนี้ ยังมีปัญหาของการระบุระบบ (System Identification) และการประมวลผลสัญญาณ (Signal Processing) สามารถจัดให้อยู่ในรูปแบบของสมการเมทริกซ์เชิงเส้นซึ่งนำไปสู่การแก้ปัญหาโดยการหาค่าเหมาะที่สุดแบบคอนเวกซ์ได้เช่นกัน

ในการแก้ปัญหาการหาค่าเหมาะที่สุด สิ่งสำคัญที่สุดคือ การบ่งชี้ว่าค่าสุดขีดแบบมหภาค (Global extremum) สามารถเกิดขึ้นได้หรือไม่ และหากเกิดขึ้นได้ จะเกิดขึ้นที่บริเวณใดในปริภูมิที่

กำหนดมาให้ แต่วิธีการแก้ปัญหาค่าเหมาะที่สุดโดยทั่วไปในปัจจุบันนี้ยังไม่สามารถตอบคำถามนี้ได้อย่างชัดเจน กระบวนการแก้ปัญหาคู่ระยะเบี่ยงวิธีเชิงตัวเลขอาจนำไปสู่จุดสุดขีดแบบจุลภาค (Local extremum) ได้ แต่อย่างไรก็ตาม หากทราบว่าปัญหาการหาค่าเหมาะที่สุดนั้นสามารถเขียนอยู่ในรูปของปัญหา Convex-programming ได้ปัญหาดังกล่าวนั้นสามารถหาผลเฉลยที่เป็นค่าสุดขีดแบบมหภาคได้อย่างแน่นอนและมีขั้นตอนวิธี (Algorithm) ที่มีประสิทธิภาพในการหาผลเฉลยค่าเหมาะที่สุดเชิงตัวเลขได้

โครงการนี้ศึกษาการนำกระบวนการของสมการเมทริกซ์เชิงเส้นมาประยุกต์ใช้กับการออกแบบระบบควบคุมแบบต่างๆ กัน โดยการนำตัวอย่างของปัญหาที่สำคัญบางตัวอย่างที่ปรากฏอยู่ในทฤษฎีระบบควบคุมมาจัดให้อยู่ในรูปแบบของสมการเมทริกซ์เชิงเส้นแล้วคำนวณหาผลเฉลยเชิงตัวเลข

1.2 วัตถุประสงค์

1. เพื่อแสดงให้เห็นถึงความสำคัญในการประยุกต์ใช้คณิตศาสตร์ในการแก้ปัญหาในชีวิตประจำวันได้
2. เพื่อศึกษาการแก้ปัญหาระบบควบคุมโดยใช้วิธีสมการเมทริกซ์เชิงเส้น
3. เพื่อศึกษาการใช้โปรแกรม MATLAB ในการหาผลเฉลยของปัญหาทางวิศวกรรมระบบควบคุมได้

1.3 ขอบข่ายของโครงการ

ศึกษาการแปลงปัญหาทางระบบควบคุมบางปัญหาให้อยู่ในรูปของปัญหาสมการเมทริกซ์เชิงเส้น (LMI) แล้วทำการเปรียบเทียบผลเฉลยที่ได้จากปัญหาดังเดิมกับปัญหาที่จัดอยู่ในรูปของสมการเมทริกซ์เชิงเส้นพร้อมทั้งนำระบบทางกายภาพมาเป็นกรณีศึกษา

1.5 ผลที่คาดว่าจะได้รับจากโครงการ

1. เกิดความเข้าใจถึงแนวคิดและหลักการเบื้องต้นของทฤษฎีระบบควบคุม
2. สามารถนำความรู้ด้านทฤษฎีควบคุมไปประยุกต์ใช้ได้กับระบบทางกายภาพที่พบในชีวิตประจำวันได้
3. สามารถนำปัญหาทางระบบควบคุมมาทำการแก้ปัญหาโดยวิธีการของอสมการเมทริกซ์เชิงเส้นได้

1.6 งบประมาณ

1. ค่าถ่ายเอกสารและค่าเช่าเล่มรายงานฉบับสมบูรณ์	เป็นเงิน	500	บาท
2. ค่าวัสดุสำนักงาน	เป็นเงิน	300	บาท
3. ค่าวัสดุคอมพิวเตอร์	เป็นเงิน	200	บาท
รวมเป็นเงินทั้งสิ้น		<u>1,000</u>	บาท
		(หนึ่งพันบาทถ้วน)	

หมายเหตุ ถัวเฉลี่ยทุกรายการ



บทที่ 2

หลักการและทฤษฎีเบื้องต้น

ในการศึกษาทฤษฎีด้านระบบควบคุมนั้น จะต้องมียุทธศาสตร์ความรู้ทางคณิตศาสตร์หลายส่วน เพื่อนำมาใช้ในการวิเคราะห์ระบบหรือสังเคราะห์ตัวควบคุมเพื่อให้ระบบป้อนกลับที่ได้มีประสิทธิภาพ ดังนั้นในบทนี้จึงนำเสนอหลักการและทฤษฎีทางคณิตศาสตร์เบื้องต้นที่จำเป็นต่อโครงงานนี้ สำหรับผู้อ่านที่มีความสนใจในเนื้อหาเกี่ยวกับพีชคณิตเชิงเส้น สามารถศึกษาได้ที่ [7, 12, 13]

2.1 พีชคณิตเชิงเส้น (Linear Algebra)

พิจารณาระบบในรูปที่ 2.1 ซึ่งเป็นระบบที่แสดงอยู่ในรูปแบบของสัญญาณเข้า-สัญญาณออก (Input-output description) $y = G(u)$



รูปที่ 2.1 ระบบในรูปแบบสัญญาณเข้า – สัญญาณออก

โดยที่ u และ y คือสัญญาณเข้า (Input) และสัญญาณออก (Output) ของระบบตามลำดับ และ G คือ การส่งที่แทนพฤติกรรมระหว่างสัญญาณเข้ากับสัญญาณออกของระบบ

นิยาม 2.1.1 ระบบที่แสดงในรูปที่ 2.1 ถูกกล่าวว่าเป็นระบบเชิงเส้นถ้า

$$G(a_1u_1 + a_2u_2) = a_1G(u_1) + a_2G(u_2) \quad (2.1)$$

เมื่อ a_1, a_2 เป็นปริมาณสเกลาร์ใดๆ

กำหนดให้ y_i เป็นสัญญาณออกจากระบบเมื่อป้อนสัญญาณเข้า u_i หากระบบที่พิจารณาเป็นระบบเชิงเส้น จะพบว่า เมื่อป้อนสัญญาณเข้าเป็น $\sum_{i=1}^n a_i u_i$ จะได้สัญญาณออกเป็น $\sum_{i=1}^n a_i y_i$ เสมอ

จากนิยาม 2.2.1 พบว่าการนิยามความเป็นเชิงเส้นของระบบได้นั้น ต้องสามารถนำสัญญาณเข้า u_1, u_2 ใดๆ ของระบบมารวมกันได้ และสามารถคูณสัญญาณเข้า u ใดๆ ด้วยสเกลาร์ได้ ปริมาณสเกลาร์ที่น่ามาคูณนี้จะต้องเป็นสมาชิกของเซตที่มีโครงสร้างพิเศษซึ่งเรียกว่า ฟิลด์ นอกจากนั้น เซตที่มีโครงสร้างทางพีชคณิตซึ่งทำให้สามารถนิยามการบวกของสมาชิกในเซตนั้น และการคูณสมาชิกในเซตนั้นด้วยปริมาณสเกลาร์ได้ เรียกว่า ปริภูมิเวกเตอร์ หรือ ปริภูมิเชิงเส้น

2.1.1 สนาม, ปริภูมิเวกเตอร์และปริภูมิย่อย (Field, Vector space and Subspace)

สนาม หรือ ฟีลด์

นิยาม 2.1.2 สนาม หรือ ฟีลด์ (Field) คือระบบนามธรรมทางคณิตศาสตร์ $(F, +, \cdot)$ ที่ประกอบด้วย

- เซต F ที่ซึ่งสมาชิกของ F เรียกว่า สเกลาร์
- การดำเนินการทวิภาค 2 อย่างคือ $+: F \times F \rightarrow F$ และ $\cdot: F \times F \rightarrow F$ ซึ่งเรียกว่าการบวก (Addition) และการคูณ (Multiplication) ตามลำดับ

โดยที่ $(F, +, \cdot)$ ต้องมีสมบัติดังนี้

- 1) F มีคุณสมบัติปิดภายใต้การดำเนินการทั้งสอง นั่นคือ สำหรับ a และ b ใด ๆ ใน F

$$a + b \in F \text{ และ } a \cdot b \in F$$

- 2) การบวกและการคูณมีคุณสมบัติการสลับที่ นั่นคือ สำหรับ a และ b ใด ๆ ใน F

$$a + b = b + a \text{ และ } a \cdot b = b \cdot a$$

- 3) การบวกและการคูณมีคุณสมบัติการจับหมู่ นั่นคือ สำหรับ a, b, c ใด ๆ ใน F

$$(a + b) + c = a + (b + c) \text{ และ } (a \cdot b) \cdot c = a \cdot (b \cdot c)$$

- 4) การคูณมีคุณสมบัติการกระจายเทียบกับการบวก นั่นคือ สำหรับ a, b, c ใด ๆ ใน F

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

- 5) มีสมาชิกตัวหนึ่งใน F ซึ่งเรียกว่า เอกลักษณ์การบวก (Additive identity) และเขียนแทนด้วย 0 โดยที่ $a + 0 = a = 0 + a$ สำหรับ a ใด ๆ ใน F

- 6) สำหรับแต่ละสมาชิก $a \in F$ จะมีสมาชิก $b \in F$ ซึ่ง $a + b = 0 (= b + a)$ และเรียกสมาชิก b นี้ว่า ตัวผกผันการบวก (Additive inverse) ของ a และเขียนแทนด้วย $-a$

- 7) มีสมาชิกตัวหนึ่งใน F ซึ่งเรียกว่า เอกลักษณ์การคูณ (Multiplicative identity) และเขียนแทนด้วย 1 โดยที่ $a \cdot 1 = a = 1 \cdot a$ สำหรับ a ใด ๆ ใน F

- 8) สำหรับแต่ละสมาชิก $a \in F$ โดยที่ $a \neq 0$ จะมีสมาชิก $b \in F$ ซึ่ง $a \cdot b = 1 (= b \cdot a)$ และเรียกสมาชิก b นี้ว่า ตัวผกผันการคูณ (Multiplicative inverse) ของ a และเขียนแทนด้วย a^{-1}

ตัวอย่างที่ 2.1.1 สนาม หรือ ฟีลด์

1. เซตของจำนวนจริง $\mathbb{R}$ และเซตของจำนวนเชิงซ้อน $\mathbb{C}$ กับ การบวกและการคูณ เป็นสนาม
2. สนามกาลัว (Galois field) หรือสนามจำกัด (Finite field) $GF(n)$ คือสนามที่มีสมาชิกเป็นจำนวนจำกัด และใช้แทนด้วย $GF(n)$ โดยที่ n คือจำนวนสมาชิกของเซต ตัวอย่างเช่น $GF(2) = \{0, 1\}$ ซึ่งมีนิยามการบวกและการคูณตามตารางนี้

$\oplus$	0	1
0	0	1
1	1	0

$\otimes$	0	1
0	0	0
1	0	1

จากตารางข้างต้น พบว่า เอกลักษณ์การบวก และเอกลักษณ์การคูณของ $GF(2)$ คือ 0 และ 1 ตามลำดับ ตัวผกผันการบวกของ 0 คือ 0 ตัวผกผันการบวกของ 1 คือ 1 และ ตัวผกผันการคูณของ 1 คือ 1

ปริภูมิเวกเตอร์ (Vector space)

นิยาม 2.1.3 ปริภูมิเวกเตอร์ (Vector space) หรือปริภูมิเชิงเส้น (Linear space) คือระบบนามธรรมทางคณิตศาสตร์ $(X, F, +, \cdot)$ ที่ประกอบด้วย

- เซต X ที่ซึ่งสมาชิกของ X เรียกว่า เวกเตอร์ (Vector)
- ฟิวด์ F ที่ซึ่งสมาชิกของ F เรียกว่า สเกลาร์ (Scalar)
- การดำเนินการทวิภาค 2 อย่างคือ $+: X \times X \rightarrow X$ และ $\cdot: F \times X \rightarrow X$ ซึ่งเรียกว่า การบวกเวกเตอร์ (Vector addition) และการคูณด้วยสเกลาร์ (Scalar multiplication) ตามลำดับ

โดยที่ $(X, F, +, \cdot)$ ต้องมีคุณสมบัติดังนี้ ให้ $x, y, z \in X$ และ $a, b \in F$

- 1) X มีคุณสมบัติปิดภายใต้การดำเนินการทั้งสอง นั่นคือ

$$x + y \in X \text{ และ } a \cdot x \in X$$

- 2) การบวกเวกเตอร์มีคุณสมบัติการสลับที่และมีคุณสมบัติการจัดหมู่ นั่นคือ

$$x + y = y + x \text{ และ } (x + y) + z = x + (y + z)$$

- 3) การคูณด้วยสเกลาร์มีคุณสมบัติการจัดหมู่ นั่นคือ

$$a \cdot (b \cdot x) = (a \cdot b) \cdot x$$

และมีคุณสมบัติการกระจายเทียบเท่ากับการบวกเวกเตอร์และการบวกสเกลาร์ นั่นคือ

$$a \cdot (x + y) = a \cdot x + a \cdot y \text{ และ } (a + b) \cdot x = a \cdot x + b \cdot x$$

- 4) มีสมาชิกหนึ่งใน X เรียกว่า เวกเตอร์ศูนย์ (Zero vector) เขียนแทนด้วย 0 ซึ่งทำหน้าที่เป็น เอกลักษณ์สำหรับการบวกเวกเตอร์ นั่นคือ

$$x + 0 = x = 0 + x$$

- 5) สำหรับแต่ละเวกเตอร์ $x \in X$ จะมีเวกเตอร์ $y \in X$ ซึ่ง $x + y = 0 (= y + x)$ ซึ่งเวกเตอร์ y นี้เรียกว่า เวกเตอร์ผกผันการบวก (Additive inverse vector) ของ x และเขียนแทนด้วย $-x$

- 6) มีเอกลักษณ์การคูณเวกเตอร์ด้วยปริมาณสเกลาร์จากสนาม F ซึ่งเขียนแทนด้วย 1 ที่
ซึ่ง $1 \cdot x = x (= x \cdot 1)$

ตัวอย่างที่ 2.1.2 ปริภูมิเวกเตอร์

1. ให้ $\mathbb{R}^n$ เป็นเซตของเวกเตอร์ขนาด $n \times 1$ ที่มีแต่ละส่วนประกอบเป็นจำนวนจริง

$$\mathbb{R}^n = \{x = (x_1, x_2, \dots, x_n) : x_i \in \mathbb{R}\}$$

จะได้ว่า $\mathbb{R}^n$ เป็นปริภูมิเวกเตอร์บนสนาม $\mathbb{R}$ ถ้านิยามการบวกเวกเตอร์และการคูณด้วย
สเกลาร์ดังนี้ ให้ $x, y \in \mathbb{R}^n$ และ $c \in \mathbb{R}$

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \text{ และ } c \cdot x = (cx_1, cx_2, \dots, cx_n)$$

ปริภูมิเวกเตอร์ $\mathbb{R}^n$ นี้เรียกว่า ปริภูมิยูคลิดค่าจริง (Real Euclidean space)

ในกรณีที่ $n=1$ จะได้ว่า $\mathbb{R}$ เป็นปริภูมิเวกเตอร์บนสนาม $\mathbb{R}$ ด้วย

2. ให้ $\mathbb{C}^n$ เป็นเซตของเวกเตอร์ขนาด $n \times 1$ ที่มีแต่ละส่วนประกอบเป็นจำนวนเชิงซ้อน

$$\mathbb{C}^n = \{z = (z_1, z_2, \dots, z_n) : z_i \in \mathbb{C}\}$$

จะได้ว่า $\mathbb{C}^n$ เป็นปริภูมิเวกเตอร์บนสนาม $\mathbb{C}$ ถ้านิยามการบวกเวกเตอร์และการคูณ
ด้วยสเกลาร์ดังนี้ ให้ $w, z \in \mathbb{C}^n$ และ $c \in \mathbb{C}$

$$z + w = (z_1 + w_1, z_2 + w_2, \dots, z_n + w_n) \text{ และ } c \cdot z = (cz_1, cz_2, \dots, cz_n)$$

ปริภูมิเวกเตอร์ $\mathbb{C}^n$ นี้เรียกว่า ปริภูมิยูคลิดค่าเชิงซ้อน (Complex Euclidean space)

ในกรณีที่ $n=1$ จะได้ว่า $\mathbb{C}$ เป็นปริภูมิเวกเตอร์บนสนาม $\mathbb{C}$ ด้วย

3. ให้ $\mathbb{R}^{n \times m}$ เป็นเซตของเมทริกซ์ขนาด $n \times m$ ที่มีแต่ละส่วนประกอบเป็นจำนวนจริง

$$\mathbb{R}^{n \times m} = \{A = [a_{ij}] : a_{ij} \in \mathbb{R}; i = 1, \dots, n; j = 1, \dots, m\}$$

จะได้ว่า $\mathbb{R}^{n \times m}$ เป็นปริภูมิเวกเตอร์บนฟิลด์ $\mathbb{R}$ ถ้านิยามการบวกเวกเตอร์และการคูณ
ด้วยสเกลาร์ดังนี้ ให้ $A = [a_{ij}], B = [b_{ij}] \in \mathbb{R}^{n \times m}$ และ $c \in \mathbb{R}$

$$A + B = [a_{ij} + b_{ij}] \text{ และ } c \cdot A = [ca_{ij}]$$

สังเกตว่า “เวกเตอร์” (ในความหมายที่ว่าเป็นสมาชิกของปริภูมิเวกเตอร์) ในกรณีนี้
นิยมเรียกว่า เมทริกซ์

4. ให้ X เป็นเซตของลำดับจำนวนจริง

$$X = \{x = \{x_i\}_{i=1}^{\infty} : x_i \in \mathbb{R}\}$$

จะได้ว่า X เป็นปริภูมิเวกเตอร์บนสนาม $\mathbb{R}$ ถ้านิยามการบวกเวกเตอร์และการคูณ
ด้วยสเกลาร์ดังนี้ สำหรับ $x, y \in X$ และ $c \in \mathbb{R}$

$$x + y = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots) \text{ และ } c \cdot x = (cx_1, cx_2, cx_3, \dots)$$

สังเกตว่า “เวกเตอร์” (ในความหมายที่ว่าเป็นสมาชิกของปริภูมิเวกเตอร์) ในกรณีนี้
นิยมเรียกว่า ลำดับ

5. ให้ $C[a,b]$ เป็นเซตของฟังก์ชันต่อเนื่องค่าจริงบนช่วงปิด $[a,b]$

$$C[a,b] = \{f : [a,b] \rightarrow \mathbb{R} : f \text{ เป็นฟังก์ชันต่อเนื่องบน } [a,b]\}$$

จะได้ว่า $C[a,b]$ เป็นปริภูมิเวกเตอร์บนสนาม $\mathbb{R}$ ถ้านิยามการบวกเวกเตอร์และการคูณด้วยสเกลาร์ ดังนี้ สำหรับ $f, g \in C[a,b]$ และ $c \in \mathbb{R}$

$$(f+g)(t) = f(t) + g(t) \text{ และ } (c \cdot f)(t) = cf(t) \text{ ทุกค่า } t \in [a,b]$$

สังเกตว่า “เวกเตอร์” (ในความหมายที่ว่าเป็นสมาชิกของปริภูมิเวกเตอร์) ในกรณีนี้ นิยมเรียกว่า ฟังก์ชัน

ปริภูมิย่อย (Subspace)

นิยาม 2.1.4 ให้ X เป็นปริภูมิเวกเตอร์ เซต $M \subset X$ ถูกกล่าวว่าเป็นปริภูมิย่อย (Subspace) ของ X ถ้า

$$ax + by \in M \text{ สำหรับเวกเตอร์ } x, y \in M \text{ และ } a, b \in F$$

กล่าวคือเซตย่อย M ของ X จะเป็นปริภูมิย่อยก็ต่อเมื่อ M มีสมบัติการปิดเทียบกับการบวกเวกเตอร์ และการคูณด้วยสเกลาร์ หรือนั่นคือเซต M เองต้องเป็นปริภูมิเวกเตอร์ด้วย

ตัวอย่างที่ 2.1.3 ปริภูมิย่อย

1. ให้ X เป็นปริภูมิเวกเตอร์ของลำดับของจำนวนจริง และให้

$$M_1 = \{x = \{x_n\} : x \text{ เป็นลำดับลู่เข้า}\}$$

$$M_2 = \{x = \{x_n\} : x \text{ เป็นลำดับมีขอบเขต}\}$$

จะได้ว่า M_1 และ M_2 เป็นปริภูมิย่อยของ X

2. ให้ X เป็นปริภูมิเวกเตอร์ของฟังก์ชันค่าจริงที่นิยามบนช่วงปิด $[a,b]$ และให้

$$M_1 = \{f : [a,b] \rightarrow \mathbb{R} : f \text{ เป็นฟังก์ชันต่อเนื่องเอกรูป (uniformly continuous) บน } [a,b]\}$$

$$M_2 = \{f : [a,b] \rightarrow \mathbb{R} : f \text{ เป็นฟังก์ชันหาอนุพันธ์ได้ (differentiable) บน } [a,b]\}$$

$$M_3 = \{f : [a,b] \rightarrow \mathbb{R} : f \text{ เป็นฟังก์ชันหาเมเชอร์ได้ (Measurable) }\}$$

$$M_4 = \{f : [a,b] \rightarrow \mathbb{R} : f \text{ เป็นฟังก์ชันอย่างง่าย}\}$$

$$M_5 = \{f : [a,b] \rightarrow \mathbb{R} : f \text{ เป็นพหุนาม}\}$$

จะได้ว่า

- M_1 เป็นปริภูมิย่อยของ X
- M_2 เป็นปริภูมิย่อยของ X และ M_2 เป็นปริภูมิย่อยของ M_1
- M_3 เป็นปริภูมิย่อยของ X และ M_1 เป็นปริภูมิย่อยของ M_3
- M_4 และ M_5 เป็นปริภูมิย่อยของ X

สังเกตว่า $0 \in M$ แสดงว่าทุกๆ ปริภูมิย่อยต้องมีเวกเตอร์ศูนย์อยู่ และเซตว่าง $\emptyset$ ไม่เป็นปริภูมิย่อยของปริภูมิเวกเตอร์ X ใดๆ ปริภูมิย่อยที่เล็กที่สุดที่เป็นไปได้คือ $\{0\}$ และปริภูมิย่อยที่ใหญ่ที่สุดที่เป็นไปได้คือ X

นิยาม 2.1.5 ให้ M, N เป็นปริภูมิย่อยของ X ผลบวกของ M และ N ซึ่งเขียนแทนด้วย $M + N$ คือเซตของเวกเตอร์ในรูป $m + n$ เมื่อ $m \in M$ และ $n \in N$

$$M + N = \{m + n : m \in M, n \in N\} \quad (2.2)$$

นิยาม 2.1.6 ให้ $\{x_i\}_{i=1}^n$ เป็นเวกเตอร์ในปริภูมิเวกเตอร์ X และให้ $\{a_i\}_{i=1}^n$ เป็นสมาชิกจากสนาม F แล้ว ผลรวม $a_1x_1 + \dots + a_nx_n$ เรียกว่า ผลรวมเชิงเส้น (Linear combination) ของเวกเตอร์ $x_1, \dots, x_n$ นอกจากนี้เซตของผลรวมเชิงเส้นทั้งหมดของ $x_1, \dots, x_n$ เรียกว่า ผลการแผ่ (Span) ของ $x_1, \dots, x_n$ และเขียนแทนด้วย

$$\text{span}\{x_1, \dots, x_n\} = \left\{ x = \sum_{i=1}^n a_i x_i \in X : x_i \in X; a_i \in F \right\} \quad (2.3)$$

หมายเหตุ ผู้อ่านสามารถพิสูจน์ได้โดยไม่ต้องยากว่า $M + N$ ในนิยาม 2.1.5 และ $\text{span}\{x_1, \dots, x_n\}$ ในนิยาม 2.1.6 เป็นปริภูมิย่อยของ X เสมอ

นิยาม 2.1.7 เวกเตอร์ $x_1, x_2, \dots, x_n$ ถูกกล่าวว่ามี ความเป็นอิสระเชิงเส้น (Linearly independent) ถ้าให้ผลรวมเชิงเส้น $a_1x_1 + a_2x_2 + \dots + a_nx_n = 0$ แล้วได้ว่า $a_1 = a_2 = \dots = a_n = 0$ มิฉะนั้นจะกล่าวว่า $x_1, x_2, \dots, x_n$ ไม่เป็นอิสระเชิงเส้น (Linearly dependent)

ตัวอย่างที่ 2.1.4 ใน $\mathbb{R}^2$ ให้ $e_1 = (1, 0)$, $e_2 = (0, 1)$ และ $e_3 = (-1, -1)$ จะพบว่า e_1, e_2 มีความอิสระเชิงเส้นต่อกันเนื่องจาก เมื่อกำหนดให้ $a_1e_1 + a_2e_2 = 0$ จะได้ว่า $a_1 = 0 = a_2$ ในทำนองเดียวกัน e_2, e_3 มีความอิสระเชิงเส้นต่อกัน อย่างไรก็ตาม e_1, e_2, e_3 ไม่อิสระเชิงเส้นต่อกัน เนื่องจาก เมื่อกำหนดให้ $a_1e_1 + a_2e_2 + a_3e_3 = 0$ แล้ว ไม่จำเป็นที่ a_1, a_2 และ a_3 ต้องเท่ากับศูนย์พร้อมกันทุกตัว ยกตัวอย่างเช่น $a_1 = 1 = a_2 = a_3$

นิยาม 2.1.8 ให้ $V = \{e_k : k = 1, 2, 3, \dots\} \subset X$ จะกล่าวว่า V แผ่ทั่ว X (V spans X) หรือ X ถูกแผ่ทั่วด้วย V (X is spanned by V) ถ้าเวกเตอร์ $x \in X$ ทุกตัวสามารถเขียนให้อยู่ในรูปของผลรวมเชิงเส้นของเวกเตอร์ใน V ได้ ซึ่งจะเขียนแทนด้วย $\text{span } V = X$

ตัวอย่างที่ 2.1.5 ใน $\mathbb{R}^2$ ให้ $e_1 = (1, 0)$, $e_2 = (0, 1)$ และ $e_3 = (-1, -1)$ จะพบว่า $\text{span}\{e_1, e_2\} = \mathbb{R}^2$

เนื่องจากเวกเตอร์ $x = (a, b) \in \mathbb{R}^2$ ทุกตัวสามารถเขียนให้อยู่ในรูป $x = a_1 e_1 + a_2 e_2$ โดยที่ $a_1 = a$ และ $a_2 = b$ เช่นเดียวกัน $\text{span}\{e_1, e_2, e_3\} = \mathbb{R}^2$ ด้วย เนื่องจาก $x = (a, b)$ สามารถเขียนได้เป็น $x = a_1 e_1 + a_2 e_2 + a_3 e_3$ ได้โดยที่ $(a_1, a_2, a_3) = (a+1, b+1, 1)$ หรือ $(1, b-a+1, 1-a)$

นิยาม 2.1.9 $E = \{e_k : k = 1, 2, 3, \dots\} \subset X$ เป็นฐานหลัก (Basis) ของ X ถ้า

1. $\{e_k : k = 1, 2, 3, \dots\}$ มีความอิสระเชิงเส้นต่อกัน
2. $\text{span } E = X$

ข้อสังเกต จากตัวอย่างที่ 2.1.4 พบว่า จำนวนเวกเตอร์ที่สามารถมีความอิสระเชิงเส้นต่อกันได้ใน $\mathbb{R}^n$ ต้องมีไม่เกิน n เวกเตอร์ และจากตัวอย่างที่ 2.1.5 พบว่าเซตของเวกเตอร์ที่แผ่ทั่วปริภูมิ $\mathbb{R}^2$ มีได้หลายแบบ และจำนวนเวกเตอร์ที่สามารถแผ่ทั่ว $\mathbb{R}^2$ ต้องมีอย่างน้อย 2 เวกเตอร์ ในกรณีทั่วไปพบว่าเวกเตอร์ที่สามารถแผ่ทั่ว $\mathbb{R}^n$ ต้องมี n เวกเตอร์เป็นอย่างน้อย ดังนั้น นักคณิตศาสตร์จึงนิยามคำว่า “ฐานหลัก” ขึ้นมา ซึ่งทำให้จำนวนเวกเตอร์ที่สามารถแผ่ทั่ว $\mathbb{R}^n$ และมีความอิสระเชิงเส้นต่อกันมีจำนวน n เวกเตอร์

นิยาม 2.1.10 ให้ X เป็นปริภูมิเวกเตอร์และ E เป็นฐานหลักใดๆ ใน X

1. มิติ (Dimension) ของ X คือจำนวนการนับ (Cardinality) ของเซต E
2. X ถูกกล่าวว่าเป็น ปริภูมิเวกเตอร์มิติจำกัด (Finite-dimensional vector space) ถ้า E เป็นเซตจำกัด
3. ถ้า X ไม่ใช่ปริภูมิเวกเตอร์มิติจำกัด จะกล่าวว่า X เป็น ปริภูมิเวกเตอร์มิติอนันต์ (Infinite-dimensional vector space)

2.1.2 ตัวดำเนินการเชิงเส้นในปริภูมิเวกเตอร์ (Linear operator in vector space)

เมื่อมีปริภูมิเวกเตอร์ใด ๆ แล้วประสงค์ที่ต้องการศึกษาความสัมพันธ์ภายในปริภูมินั้น ๆ หรือ ความสัมพันธ์ระหว่างปริภูมิ จึงจำเป็นต้องสร้างนิยามของคำว่า “ตัวดำเนินการ” ซึ่งทำหน้าที่คล้ายกับ ฟังก์ชัน ที่สร้างความสัมพันธ์ระหว่างเซต

นิยาม 2.1.11 ให้ X และ Y เป็นปริภูมิเวกเตอร์บนสนาม F เดียวกัน และให้ $D \subset X$ แล้ว การจับคู่สมาชิก $y \in Y$ เข้ากับแต่ละสมาชิก $x \in D$ ถูกเรียกว่า การแปลง (Transformation) หรือ การส่ง (Mapping) จาก X ไปยัง Y ที่มีโดเมน D ซึ่งนิยามใช้สัญลักษณ์แทนการแปลงด้วย T ดังนั้น จะได้ว่า $T : D \subset X \rightarrow Y$

นิยาม 2.1.12 การแปลง $T: X \rightarrow Y$ ถูกเรียกว่าเป็น การแปลงเชิงเส้น (Linear transformation) หรือ ตัวดำเนินการเชิงเส้น (Linear operator) ถ้า

1. T มีคุณสมบัติเชิงการบวก (Additive property)

$$T(x + y) = T(x) + T(y), \forall x, y \in X$$

2. T มีสมบัติเอกพันธ์ (Homogeneous property)

$$T(ax) = aT(x), \forall x \in X, a \in F$$

หรือสรุปรวมเป็นเงื่อนไขเดียวได้ว่า T จะเป็น ตัวดำเนินการเชิงเส้น ถ้า

$$T(ax + by) = aT(x) + bT(y), \forall x, y \in X, a, b \in F$$

และเขียนแทนเซตของตัวดำเนินการเชิงเส้นจาก X ไป Y ด้วย $\mathcal{L}(X, Y)$

ตัวอย่างที่ 2.1.6 ตัวดำเนินการเชิงเส้นในปริภูมิเวกเตอร์

1. การคูณด้วยเมทริกซ์

ตัวดำเนินการ $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ที่นิยามโดย

$$Tx = Ax$$

เมื่อ A เป็นเมทริกซ์ขนาด $m \times n$ ($A \in \mathbb{R}^{m \times n}$) เป็นตัวดำเนินการเชิงเส้นจาก $\mathbb{R}^n$ ไปยัง $\mathbb{R}^m$

2. ตัวดำเนินการเอกลักษณ์ (Identity operator)

ให้ X เป็นปริภูมิเวกเตอร์ จะได้ว่าตัวดำเนินการ $i: X \rightarrow X$ ที่นิยามโดย

$$ix = x, \forall x \in X$$

เป็นตัวดำเนินการเชิงเส้นบน X

3. ตัวดำเนินการศูนย์ (Zero operator)

ให้ X และ Y เป็นปริภูมิเวกเตอร์ จะได้ว่าตัวดำเนินการ $0: X \rightarrow Y$ ที่นิยามโดย

$$0x = 0, \forall x \in X$$

เป็นตัวดำเนินการเชิงเส้นจาก X ไปยัง Y

4. ตัวดำเนินการเชิงปริพันธ์ (Integral operator)

ให้ $C[a, b]$ เป็นปริภูมิเวกเตอร์ของฟังก์ชันต่อเนื่องที่นิยามบน $[a, b]$ จะได้ว่าตัวดำเนินการเชิงปริพันธ์ $T: C[a, b] \rightarrow C[a, b]$ ที่นิยามโดย

$$Tf(t) = g(t) = \int_a^t f(\tau) d\tau \text{ สำหรับค่า } t \in [a, b]$$

เป็นตัวดำเนินการเชิงเส้นบน $C[a, b]$

5. ตัวดำเนินการเชิงอนุพันธ์ (Differential operator)

ให้ $P[a, b]$ เป็นปริภูมิเวกเตอร์ของพหุนามที่นิยามบน $[a, b]$ จะได้ว่าตัวดำเนินการเชิงอนุพันธ์ $D: P[a, b] \rightarrow P[a, b]$ ที่นิยามโดย

$$Df = g, \quad g(t) = \frac{df}{dt}$$

เป็นตัวดำเนินการเชิงเส้นบน $P[a, b]$

นิยาม 2.1.13 ให้ $T : X \rightarrow Y$ เป็นตัวดำเนินการเชิงเส้น

1. ปริภูมิพิสัย (Range space หรือ image) ของ T ซึ่งเขียนแทนด้วย $\text{im}T$ หรือ $R(T)$ คือ

$$\text{im}T = \{y = Tx : x \in X\} \quad (2.4)$$

2. ปริภูมิศูนย์ (Null space หรือ kernel) ของ T ซึ่งเขียนแทนด้วย $\ker T$ หรือ $N(T)$ คือ

$$\ker T = \{x \in X : Tx = 0\} \quad (2.5)$$

ให้ $T : X \rightarrow Y$ เป็นตัวดำเนินการเชิงเส้น ผู้อ่านสามารถพิสูจน์โดยไมยากว่า $\ker T$ และ $\text{im}T$ เป็นปริภูมิเวกเตอร์ด้วย

นิยาม 2.1.14 ให้ $T : X \rightarrow Y$ เป็นตัวดำเนินการเชิงเส้น

1. จะกล่าวว่า T เป็น ตัวดำเนินการหนึ่งต่อหนึ่ง (One-to-one operator, injective operator หรือ injection) ถ้า $Tx = 0$ แล้ว $x = 0$.
2. จะกล่าวว่า T เป็น ตัวดำเนินการทั่วถึง (Onto operator, surjective operator หรือ surjection) ถ้า สำหรับทุกๆ $y \in Y$ จะมี $x \in X$ ซึ่ง $Tx = y$
3. จะกล่าวว่า T เป็น ตัวดำเนินการหาตัวผกผันได้ (Invertible operator) ถ้ามีตัวดำเนินการ $S : \text{im}T \rightarrow X$ ซึ่ง

$$STx = x, \quad \forall x \in X \quad (2.6)$$

และ

$$TSy = y, \quad \forall y \in \text{im}T \quad (2.7)$$

จะเรียกตัวดำเนินการ S นี้ว่า ตัวดำเนินการผกผันของ T และเขียนแทนด้วย T^{-1}

ทฤษฎีบทที่ 2.1.1 ให้ $T : X \rightarrow Y$ เป็นตัวดำเนินการเชิงเส้น

1. T เป็นตัวดำเนินการหาตัวผกผันได้ ก็ต่อเมื่อ T เป็นตัวดำเนินการหนึ่งต่อหนึ่ง
2. ถ้า T เป็นตัวดำเนินการหาตัวผกผันได้ จะได้ว่า T^{-1} เป็นตัวดำเนินการเชิงเส้นด้วย

2.1.3 ปริภูมิเวกเตอร์นอร์ม (Normed vector space)

นอร์ม เป็นนัยทั่วไปของความยาวของเวกเตอร์ใน $\mathbb{R}^n$ และปริภูมิเวกเตอร์นอร์ม คือปริภูมิเวกเตอร์ที่มีการนิยามนอร์ม ดังนั้นในปริภูมิเวกเตอร์นอร์มจึงสามารถนิยามคุณสมบัติต่างๆ เช่น เซตเปิด ลิมิต การลู่เข้า ความต่อเนื่อง เป็นต้น

นิยาม 2.1.15 ให้ X เป็นปริภูมิเวกเตอร์บนสนาม F นอร์ม (Norm) คือฟังก์ชันค่าจริง $\|\cdot\|: X \rightarrow \mathbb{R}$ ซึ่งมีคุณสมบัติดังนี้ ให้ x, y เป็นเวกเตอร์ใดๆ ใน X และ $a \in F$

1. $\|x\| \geq 0$
2. $\|x\| = 0 \Leftrightarrow x = 0$
3. $\|ax\| = |a|\|x\|$
4. $\|x + y\| \leq \|x\| + \|y\|$

ถ้า X เป็นปริภูมิเวกเตอร์ที่มีนอร์มนิยามบน X จะเรียก X ว่า ปริภูมิเชิงเส้นที่มีนอร์ม (Normed linear space) หรือ ปริภูมิเวกเตอร์นอร์ม (Normed vector space) ซึ่งใช้สัญลักษณ์แทนด้วย $(X, \|\cdot\|)$

ตัวอย่างที่ 2.1.7 ตัวอย่างที่ง่ายที่สุดของปริภูมิเวกเตอร์นอร์มคือเซตของจำนวนจริง $\mathbb{R}$ โดยที่

$$\|x\| = |x|, x \in \mathbb{R}$$

จะเห็นว่า ค่าสัมบูรณ์ $|\cdot|$ สอดคล้องกับนิยามทั้งสี่ข้อคือ

1. $|x| \geq 0$
2. $|x| = 0$ ก็ต่อเมื่อ $x = 0$
3. $|ax| = |a||x|$
4. $|\cdot|$ สอดคล้องกับอสมการสามเหลี่ยม นั่นคือ $|x + y| \leq |x| + |y|, x, y \in \mathbb{R}$

อีกตัวอย่างหนึ่งของปริภูมิเวกเตอร์นอร์มคือ ปริภูมิแบบยูคลิด n มิติ (n -dimensional Euclidean space) $\mathbb{R}^n$ ที่มีนอร์ม-2 (2-norm)

$$\|x\|_2 = \left(\sum_{k=1}^n |x_k|^2 \right)^{1/2}$$

สามารถพิสูจน์ได้ว่า $\|\cdot\|_2$ นี้มีคุณสมบัติสอดคล้องกับนิยามของนอร์มทั้งสี่ข้อคือ

1. $\|x\|_2 \geq 0$ และ $\|x\|_2 = 0$ ก็ต่อเมื่อ $x = 0$ (เวกเตอร์ศูนย์)
2. $\|ax\|_2 = |a|\|x\|_2$
3. $\|\cdot\|_2$ สอดคล้องกับอสมการสามเหลี่ยม นั่นคือ สำหรับ $x, y \in \mathbb{R}^n$

$$\|x + y\|_2 = \left(\sum_{k=1}^n |x_k + y_k|^2 \right)^{1/2} \leq \left(\sum_{k=1}^n |x_k|^2 \right)^{1/2} + \left(\sum_{k=1}^n |y_k|^2 \right)^{1/2} = \|x\|_2 + \|y\|_2$$

ความจริงแล้ว นอร์ม -2 เป็นกรณีพิเศษของ นอร์ม - p (p -norm) ที่นิยามด้วย

$$\|x\|_p = \left(\sum_{k=1}^n |x_k|^p \right)^{1/p}$$

ซึ่งจะเป็นนอร์มบน $\mathbb{R}^n$ ได้เมื่อ $1 \leq p < \infty$

นอกจากนี้ยังสามารถนิยาม นอร์มอนันต์ (∞ -norm) ใน $\mathbb{R}^n$ ได้ดังนี้

$$\|x\|_\infty = \max \{|x_k| : k = 1, \dots, n\}$$

ตัวอย่างที่ 2.1.8 ปริภูมิ ℓ_p

สำหรับ $1 \leq p < \infty$ และพิจารณาปริภูมิเวกเตอร์ X ที่มีสมาชิกเป็นลำดับอนันต์ (ดูตัวอย่างที่ 2.1.2)

หากนิยามนอร์มบนปริภูมิเวกเตอร์ด้วย $\|x\|_p = \left(\sum_{k=1}^\infty |x_k|^p \right)^{1/p}$ สำหรับทุกค่า $x \in X$ จะได้ว่า ปริภูมิ

ดังกล่าวที่นิยามนอร์มข้างต้น จัดว่าเป็นปริภูมิเวกเตอร์นอร์ม หากนอร์มมีค่าจำกัด ปริภูมิเวกเตอร์นอร์มดังกล่าวนี้ นิยมเขียนแทนด้วย ℓ_p นั่นคือ

$$\ell_p = \{x = (x_1, x_2, x_3, \dots) : \|x\|_p < \infty\}$$

และนอร์มที่นิยามนี้ นิยมเรียกว่า นอร์ม - p บนปริภูมิเวกเตอร์นอร์ม ℓ_p

ในกรณีที่ $p = \infty$ จะนิยามนอร์มอีกประเภทหนึ่ง เรียกว่า นอร์มอนันต์ นั่นคือ $\|x\|_\infty = \sup_{k \in \mathbb{N}} |x_k|$ ซึ่งปริภูมิเวกเตอร์นอร์มภายใต้ นอร์มอนันต์ นิยมเขียนแทนด้วย ℓ_∞ หรือนั่นคือ

$$\ell_\infty = \{x = (x_1, x_2, x_3, \dots) : \sup_{k \in \mathbb{N}} |x_k| < \infty\}$$

ตัวอย่างที่ 2.1.9 ปริภูมิฟังก์ชัน

ให้ $X = C[a, b]$ เป็นเซตของฟังก์ชันต่อเนื่องบน $[a, b]$ สามารถนิยามนอร์มบนปริภูมินี้ได้หลายแบบ เช่น นอร์ม - p นิยามว่า

$$\|f\|_p = \left(\int_a^b |f(x)|^p dx \right)^{1/p}$$

และนอร์มอนันต์

$$\|f\|_\infty = \sup_{a \leq x \leq b} |f(x)|$$

ข้อสังเกตประการหนึ่งคือ ถึงแม้ว่าปริภูมิเวกเตอร์ที่พิจารณาจะเป็นแบบเดียวกันคือ X แต่หากมีการนิยามนอร์มที่ต่างกันจะทำให้ได้ปริภูมิเวกเตอร์นอร์มที่แตกต่างกันออกไป ดังนั้น จึงมีความจำเป็นที่ต้องระบุชนิดของนอร์มเสมอ เวลากล่าวถึงปริภูมิเวกเตอร์นอร์ม

2.1.4 ตัวดำเนินการเชิงเส้นในปริภูมิเวกเตอร์นอร์ม (Linear operator in normed vector space)

ในหัวข้อที่ 2.1.2 ได้ศึกษาตัวดำเนินการเชิงเส้นในปริภูมิเวกเตอร์ไปแล้ว ในหัวข้อนี้จะเป็นการศึกษาตัวดำเนินการเชิงเส้นในปริภูมิเวกเตอร์นอร์ม โดยเริ่มจาก ตัวดำเนินการเชิงเส้นมีขอบเขต (Bounded linear operator) ซึ่งจะพบว่าเซตของตัวดำเนินการเชิงเส้นมีขอบเขตนี้สามารถทำให้เป็นปริภูมิเวกเตอร์และปริภูมิเวกเตอร์นอร์มได้ ภายใต้การนิยามนอร์มที่เหมาะสม

นิยาม 2.1.16 ให้ X, Y เป็นปริภูมิเวกเตอร์นอร์มและให้ $A: X \rightarrow Y$ แล้ว A ถูกกล่าวว่าเป็นตัวดำเนินการมีขอบเขต (Bounded operator) ถ้ามี $M > 0$ ซึ่ง

$$\|Ax\|_Y \leq M \|x\|_X, \quad \forall x \in X \quad (2.8)$$

สาเหตุที่เรียกตัวดำเนินการที่มีคุณสมบัติดังกล่าวว่าตัวดำเนินการมีขอบเขต เนื่องจากเป็นตัวดำเนินการที่ส่งเซตที่มีขอบเขตใน X ไปเป็นเซตที่มีขอบเขตใน Y โดยพิสูจน์ได้ดังนี้

ให้ $S \subset X$ เป็นเซตที่มีขอบเขต จะได้ว่ามี $M_1 > 0$ ซึ่ง

$$\|x\|_X \leq M_1, \quad \forall x \in S \quad (2.9)$$

เนื่องจาก $A: X \rightarrow Y$ เป็นตัวดำเนินการมีขอบเขต จะได้ว่า

$$\|Ax\|_Y \leq M \|x\|_X \leq MM_1, \quad \forall x \in S \quad (2.10)$$

นั่นคือ $A(S)$ เป็นเซตที่มีขอบเขตด้วย

นิยาม 2.1.17 ให้ X, Y เป็นปริภูมิเวกเตอร์นอร์มและให้ $A: X \rightarrow Y$ แล้ว A ถูกกล่าวว่าเป็นตัวดำเนินการเชิงเส้นมีขอบเขต (Bounded linear operator) ถ้า A เป็นทั้งตัวดำเนินการเชิงเส้นและตัวดำเนินการมีขอบเขต และนิยมเขียนแทนเซตของตัวดำเนินการเชิงเส้นมีขอบเขตจาก X ไปยัง Y ด้วย $B(X, Y)$ และกรณีที่ $X = Y$ จะเขียนแทน $B(X, Y)$ ด้วย $B(X)$ และเรียก $A \in B(X)$ ว่าตัวดำเนินการเชิงเส้นมีขอบเขตบน X

ให้ X และ Y เป็นปริภูมิเวกเตอร์นอร์มและ $B(X, Y)$ เป็นเซตของตัวดำเนินการเชิงเส้นมีขอบเขต จาก X ไปยัง Y จะพบว่าเซต $B(X, Y)$ สามารถทำให้เป็นปริภูมิเวกเตอร์ได้ กล่าวคือสำหรับตัวดำเนินการเชิงเส้นที่มีขอบเขตใด ๆ $A, B \in B(X, Y)$ และ a เป็นปริมาณสเกลาร์แล้วนิยามการบวกเวกเตอร์และการคูณด้วยสเกลาร์บน $B(X, Y)$ ดังนี้

$$(A + B)x = Ax + Bx \quad \text{และ} \quad (aA)x = aAx \quad (2.11)$$

นอกจากนี้ยังสามารถทำให้เซต $B(X, Y)$ เป็นปริภูมิเวกเตอร์นอร์มได้โดยการนิยามนอร์มใน $B(X, Y)$ ดังต่อไปนี้

นิยาม 2.1.18 ให้ X, Y เป็นปริภูมิเวกเตอร์ นอร์มของตัวดำเนินการ (Operator norm) $A \in \mathcal{B}(X, Y)$ นิยามโดย

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} \quad (2.12)$$

ผู้อ่านสามารถพิสูจน์ได้ว่า $\|A\|$ มีคุณสมบัติของนอร์มทุกประการ คือ

1. $\|A\| \geq 0$
2. $\|A\| = 0$ ก็ต่อเมื่อ $A = 0$
3. $\|\lambda A\| = \sup_{x \neq 0} \left\{ \frac{\|\lambda Ax\|_Y}{\|x\|_X} \right\} = |\lambda| \|A\|$
4. อสมการสามเหลี่ยม

$$\|A + B\| = \sup_{x \neq 0} \frac{\|(A+B)x\|_Y}{\|x\|_X} \leq \sup_{x \neq 0} \frac{\|Ax\|_Y + \|Bx\|_Y}{\|x\|_X} = \|A\| + \|B\|$$

ดังนั้น $\mathcal{B}(X, Y)$ จึงเป็นปริภูมิเวกเตอร์นอร์มภายใต้ นอร์มของตัวดำเนินการที่นิยามข้างต้น นอกจากนิยาม 2.1.18 แล้ว ยังมีนิยามอื่นๆ ของ $\|A\|$ ที่สมมูลกันอีกดังต่อไปนี้

นิยามต่อไปนี้สมมูลกัน

1. $\|A\| = \sup\{\|Ax\|_Y : \|x\|_X \leq 1\}$
2. $\|A\| = \sup\{\|Ax\|_Y : \|x\|_X = 1\}$
3. $\|A\| = \inf\{M : \|Ax\|_Y \leq M \|x\|_X, \forall x \in X\}$

ตัวอย่างที่ 2.1.10 ตัวดำเนินการเชิงเส้นในปริภูมิเวกเตอร์นอร์ม

1. ให้ $X = Y = \mathbb{R}$ และนิยาม $A: \mathbb{R} \rightarrow \mathbb{R}$ โดย $Ax = ax$, $a \in \mathbb{R}$

นอร์มของ A ในกรณีนี้คือ $\|A\| = \sup_{x \neq 0} \frac{|ax|}{|x|} = |a|$

2. ให้ $X = Y = \mathbb{R}^2$ และนิยาม $A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ โดย $Ax = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} x$

ถ้าใช้นอร์ม-2 ใน $\mathbb{R}^2$ จะได้ว่า $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$ สำหรับ $x \in \mathbb{R}^2$ และ

$$\|y\|_2 = \|Ax\|_2 = \sqrt{(2x_1 + x_2)^2 + (x_1 + 2x_2)^2} = \sqrt{5(x_1^2 + x_2^2) + 8x_1x_2}$$

เนื่องจาก $\frac{\|Ax\|_2}{\|x\|_2} = \sqrt{\frac{5(x_1^2 + x_2^2) + 8x_1x_2}{x_1^2 + x_2^2}} = \sqrt{5 + \frac{8x_1x_2}{x_1^2 + x_2^2}}$ มีค่าสูงสุดที่ $x_1 = x_2$ ดังนั้น

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sqrt{5 + \frac{8x_1^2}{2x_1^2}} = \sqrt{9} = 3$$

2.1.5 ทฤษฎีเมทริกซ์ (Matrix Theory)

ในหัวข้อนี้ จะเน้นศึกษาตัวดำเนินการเชิงเส้นมีขอบเขต ที่ส่งปริภูมิยุคลิด $\mathbb{R}^n$ (หรือ $\mathbb{C}^n$) ไปยังปริภูมิยุคลิด $\mathbb{R}^m$ (หรือ $\mathbb{C}^m$) ซึ่งตัวดำเนินการประเภทนี้เรียกว่า เมทริกซ์ (Matrix)

นิยาม 2.1.19 เมทริกซ์ค่าจริง (Real-valued matrix) ที่มี m แถวและ n หลัก คือตัวดำเนินการเชิงเส้นมีขอบเขตที่ส่งปริภูมิเวกเตอร์ $\mathbb{R}^n$ ไปยังปริภูมิเวกเตอร์ $\mathbb{R}^m$ นั่นคือ $A \in B(\mathbb{R}^n, \mathbb{R}^m)$ ดังสมการ $y = Ax$ โดยที่ $x \in \mathbb{R}^n$ และ $y \in \mathbb{R}^m$ โดยที่ $a_{ij} \in \mathbb{R}, i = 1, 2, 3, \dots, m, j = 1, 2, 3, \dots, n$ คือสมาชิกในแถวที่ i และหลักที่ j หรือนั่นคือ

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \quad (2.13)$$

หรือเขียนโดยย่อเป็น $A = [a_{ij}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ ซึ่งมีมิติ (Dimension) เป็น $m \times n$

หมายเหตุ หากเปลี่ยนจากปริภูมิเวกเตอร์ $\mathbb{R}^n$ และ $\mathbb{R}^m$ เป็นปริภูมิเวกเตอร์ $\mathbb{C}^n$ และ $\mathbb{C}^m$ ตามลำดับ แล้ว ตัวดำเนินการที่ได้จะเรียกว่า เมทริกซ์ค่าเชิงซ้อน (Complex-valued matrix) ที่ซึ่ง $a_{ij} \in \mathbb{C}$

เมทริกซ์ที่มีลักษณะพิเศษที่สำคัญ

1. เมทริกซ์จัตุรัส (Square matrix)

นิยาม 2.1.20 เมทริกซ์จัตุรัส หมายถึงเมทริกซ์ที่มีจำนวนแถวและจำนวนหลักเท่ากัน นั่น

คือ $A = [a_{ij}]_{1 \leq i, j \leq n}$ เช่น

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 4 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & -4 \end{bmatrix}, \quad C = \begin{bmatrix} -1+i & 1 \\ -i & 0 \end{bmatrix}$$

2. เมทริกซ์เอกลักษณ์ (Identity matrix)

นิยาม 2.1.21 เมทริกซ์เอกลักษณ์ หมายถึงเมทริกซ์จัตุรัสที่มีสมาชิกในแนวเส้นทแยงมุม

เป็น 1 ส่วนสมาชิกที่เหลือมีค่าเป็น 0 หรือนั่นคือ $A = [a_{ij}]_{1 \leq i, j \leq n}$ โดยที่

$$a_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}, \quad i, j = 1, 2, \dots, n \quad (2.14)$$

และใช้สัญลักษณ์แทนเมทริกซ์เอกลักษณ์มิติ $n \times n$ ด้วย I_n เช่น

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3. เมทริกซ์สลับเปลี่ยนและเมทริกซ์สลับเปลี่ยนสังยุค (Transpose of a matrix and Conjugate Transpose of a matrix)

นิยาม 2.1.22 เมทริกซ์สลับเปลี่ยนของเมทริกซ์ค่าจริง $A = [a_{ij}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ หมายถึงเมทริกซ์ $[a_{ji}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ (นั่นคือ สลับตำแหน่งของแถวและหลัก) ซึ่งใช้สัญลักษณ์แทนด้วย A^T เช่น

$$\text{ถ้า } A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \text{ แล้ว } A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

ในกรณีของเมทริกซ์ค่าเชิงซ้อน $A = [a_{ij}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ โดยที่ $a_{ij} \in \mathbb{C}$ นอกจากมีการสลับเปลี่ยนตำแหน่งระหว่างแถวและหลักแล้ว ยังต้องมีการหาค่าสังยุค (Conjugate) ด้วย ซึ่งผลที่ได้จึงเรียกว่า เมทริกซ์สลับเปลี่ยนสังยุค (Conjugate transpose matrix) และใช้สัญลักษณ์แทนด้วย A^* เช่น

$$\text{ถ้า } A = \begin{bmatrix} 1+i & 3-2i & 5 \\ 2 & 4+3i & 6-i \end{bmatrix} \text{ แล้ว } A^* = \begin{bmatrix} 1-i & 2 \\ 3+2i & 4-3i \\ 5 & 6+i \end{bmatrix}$$

4. เมทริกซ์สมมาตรและเมทริกซ์เฮอร์มิเทียน (Symmetric matrix and Hermitian matrix)

นิยาม 2.1.23 เมทริกซ์จัตุรัสค่าจริง A ที่มีคุณสมบัติ $A = A^T$ เรียกว่า เมทริกซ์สมมาตร เช่น

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & -4 \end{bmatrix}$$

เมทริกซ์จัตุรัสค่าเชิงซ้อน A ที่มีคุณสมบัติ $A = A^*$ เรียกว่า เมทริกซ์เฮอร์มิเทียน เช่น

$$A = \begin{bmatrix} 1 & 3+2i \\ 3-2i & 4 \end{bmatrix} \quad \text{ที่ซึ่ง} \quad A^* = \begin{bmatrix} 1 & 3+2i \\ 3-2i & 4 \end{bmatrix}$$

5. เมทริกซ์เอกฐานและเมทริกซ์ไม่เป็นเอกฐาน (Singular and Non-singular Matrices)

นิยาม 2.1.24 เมทริกซ์ A จะถูกเรียกว่าเมทริกซ์ไม่เป็นเอกฐาน เมื่อสามารถหาเมทริกซ์ B ที่ซึ่ง $AB = I$ (หรือ $BA = I$) โดยเมทริกซ์ B ที่ได้เรียกว่า เมทริกซ์ผกผัน (Inverse) ของ A ซึ่งนิยมเขียนแทน B ด้วย A^{-1} และถ้า A ไม่สามารถหาเมทริกซ์ผกผันได้ A ถูกเรียกว่าเมทริกซ์เอกฐาน

ข้อสังเกต การนิยามความเป็นเอกฐานหรือไม่เอกฐานของเมทริกซ์นั้น ขึ้นอยู่กับว่าเมทริกซ์ดังกล่าวสามารถหาเมทริกซ์ผกผันได้หรือไม่ และการที่เมทริกซ์จะหาตัวผกผันได้หรือไม่ นั้น ขึ้นอยู่กับค่าตัวกำหนด หรือ Determinant ดังนั้น ความเป็นหรือไม่เป็นเอกฐานของเมทริกซ์ คำนิยามข้างต้น สามารถเขียนใหม่ได้ ดังนี้

นิยาม 2.1.25 เมทริกซ์จัตุรัส A ใดๆ ซึ่งค่า $\det A = 0$ จะเรียกว่า เมทริกซ์เอกฐาน หากค่า $\det A \neq 0$ จะเรียกว่า เมทริกซ์ไม่เป็นเอกฐาน

ค่าเฉพาะและเวกเตอร์เฉพาะ (Eigen value and eigen vector)

นิยาม 2.1.26 ให้ A เป็นเมทริกซ์ที่มีมิติ $n \times n$ แล้ว λ ถูกเรียกว่า ค่าเฉพาะของ A ถ้ามีเวกเตอร์ที่ไม่ใช่เวกเตอร์ศูนย์ x ที่ทำให้ $Ax = \lambda x$ ซึ่งเวกเตอร์ $x \neq 0$ ที่สอดคล้องสมการ $Ax = \lambda x$ นี้จะเรียกว่า เวกเตอร์เฉพาะของ A ที่สมนัยกับค่าเฉพาะ λ

ในการคำนวณหาค่าเฉพาะและเวกเตอร์เฉพาะนั้น ให้พิจารณาดังนี้ เนื่องจาก ค่าเฉพาะ λ คือค่าที่สอดคล้องกับสมการ $Ax = \lambda x$ ซึ่งสมมูลกับ $(\lambda I - A)x = 0$ นั่นคือ $x \in \ker(\lambda I - A)$ แต่เนื่องจาก $x \neq 0$ นั้นย่อมแสดงว่า ตัวดำเนินการหรือ Pencil $\lambda I - A$ นี้ต้องสูญเสียลำดับชั้น (rank) หรือนั่นคือ $\det(\lambda I - A) = 0$ ดังนั้นปริมาณสเกลาร์ λ ที่สอดคล้องกับสมการ $\det(\lambda I - A) = 0$ คือค่าเฉพาะของ A ซึ่งสมการ $\det(\lambda I - A) = 0$ เรียกว่า สมการคุณลักษณะ (Characteristic equation) ของ A

ตัวอย่างที่ 2.1.11 การหาค่าเฉพาะและเวกเตอร์เฉพาะของเมทริกซ์

กำหนดให้ $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$ ดังนั้น สมการคุณลักษณะคือ

$$\det(\lambda I - A) = \det \left(\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix} \right) = \det \left(\begin{bmatrix} \lambda - 2 & -2 \\ -1 & \lambda - 3 \end{bmatrix} \right) = \lambda^2 - 5\lambda + 4 = 0$$

ซึ่งจะได้ว่า $\lambda = 1, 4$

ที่ $\lambda_1 = 1$ จะได้ $(I - A)x_1 = \begin{bmatrix} -1 & -2 \\ -1 & -2 \end{bmatrix} x_1 = 0$ ดังนั้น $x_1 = c \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ โดยที่ c เป็นค่าคงที่ใด ๆ

เพื่อความสะดวก เลือก $c = 1$ จะได้ว่า เวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะ $\lambda_1 = 1$ คือ $x_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

ที่ $\lambda_2 = 4$ จะได้ $(4I - A)x_2 = \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix} x_2 = 0$ ดังนั้น $x_2 = c \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ โดยที่ c เป็นค่าคงที่ใด ๆ

เพื่อความสะดวก เลือก $c = 1$ จะได้ว่า เวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะ $\lambda_2 = 4$ คือ $x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

ทฤษฎีบทถัดไป เป็นทฤษฎีบทที่สำคัญที่เกี่ยวข้องกับเมทริกซ์ที่มีคุณสมบัติพิเศษ นั่นคือ เมทริกซ์สมมาตร

ทฤษฎีบทที่ 2.1.2 เมทริกซ์สมมาตรค่าจริง A มิติ $n \times n$ สามารถหาเวกเตอร์เฉพาะจริงที่ตั้งฉากกัน (mutually orthogonal) ได้ n เวกเตอร์เสมอ โดยแต่ละเวกเตอร์ทำหน้าที่เป็นฐานหลักใน $\mathbb{R}^n$ หรือนั่นคือ สามารถหาเมทริกซ์เชิงตั้งฉาก (Orthogonal matrix) M ที่ซึ่ง $M^T A M = D$ เมทริกซ์เส้นทแยงมุมที่มีค่าจะอยู่ในแนวเส้นทแยงมุม

บทพิสูจน์

ใช้การพิสูจน์โดยอุปนัยเชิงคณิตศาสตร์ (Mathematical induction) บนมิติของเมทริกซ์ ซึ่งพบว่า ในกรณีที่ $n=1$ จะได้ $M=1$ สมมติว่าทฤษฎีเป็นจริงสำหรับทุก ๆ เมทริกซ์สมมาตรค่าจริงที่มีมิติ $(n-1) \times (n-1)$

ให้ A เป็นเมทริกซ์สมมาตรค่าจริงมิติ $n \times n$ เนื่องจากเมทริกซ์จัตุรัสที่ไม่ใช่เมทริกซ์ศูนย์สามารถหาเวกเตอร์เฉพาะได้อย่างน้อย 1 เวกเตอร์ได้เสมอ จึงให้เวกเตอร์ดังกล่าวแทนด้วย v_1 โดยที่ $\|v_1\|=1$ และสมมูลกับค่าเฉพาะ λ_1 (นั่นคือ $A v_1 = \lambda_1 v_1$) จากนั้น สร้างเวกเตอร์ที่เหลืออีก $n-1$ เวกเตอร์ ที่ซึ่งทำหน้าที่เป็นฐานหลักเชิงตั้งฉากปกติ (Orthonormal basis) ใน $\mathbb{R}^n$ โดยอาศัย Gram Schmidt process โดยให้ฐานหลักดังกล่าวคือเซต $B = \{v_1, x_2, \dots, x_n\}$ และให้ $M_1 = [v_1 \ x_2 \ \dots \ x_n]$ เนื่องจาก M_1 สร้างจากฐานหลักเชิงตั้งฉากปกติ ดังนั้น $M_1^{-1} = M_1^T$

ให้ $Q = M_1^{-1} A M_1$ และพิจารณา

$$Q^T = (M_1^{-1} A M_1)^T = (M_1^T A M_1)^T = M_1^T A^T M_1 = M_1^{-1} A M_1 = Q \quad (2.15)$$

นั่นคือ $M_1^{-1} A M_1$ มีคุณสมบัติเป็นเมทริกซ์สมมาตรค่าจริง ให้ e_k เป็นเวกเตอร์ฐานหลักมาตรฐานใน $\mathbb{R}^n$ นั่นคือ สมาชิกตัวที่ k ของ e_k คือ 1 ส่วนสมาชิกที่เหลือมีค่า 0 ดังนั้น $(M_1^{-1} A M_1) e_1$ คือ แนวตั้งแรกของเมทริกซ์ $M_1^{-1} A M_1$ ในทำนองเดียวกัน $M_1 e_1$ ย่อมได้แนวตั้งแรกของเมทริกซ์ M_1 นั่นคือเวกเตอร์ v_1 ซึ่งจะได้ว่า

$$(M_1^{-1} A M_1) e_1 = M_1^{-1} A (M_1 e_1) = M_1^{-1} A v_1 = \lambda_1 M_1^{-1} v_1 = \lambda_1 e_1 \quad (2.16)$$

หรือเขียนได้เป็น

$$M_1^{-1} A M_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix} \quad (2.17)$$

โดยที่ A_1 เป็นเมทริกซ์สมมาตรค่าจริงมิติ $(n-1) \times (n-1)$

จากสมมติฐานของอุปนัยเชิงคณิตศาสตร์ จะต้องเกิดเมทริกซ์เชิงตั้งฉาก (Orthogonal matrix) M_2 ที่ซึ่ง $M_2^T A_1 M_2 = D_1$ เมทริกซ์เส้นทแยงมุมมิติ $n-1$ ที่มีค่าจะงอยู่ในแนวเส้นทแยงมุม สร้างเมทริกซ์ใหม่ $M_3 = \begin{bmatrix} 1 & 0 \\ 0 & M_2 \end{bmatrix}$ ซึ่งเป็นเมทริกซ์เชิงตั้งฉากมิติ n นอกจากนี้ พบว่า

$$\begin{aligned} M_3^{-1} (M_1^{-1} A M_1) M_3 &= \begin{bmatrix} 1 & 0 \\ 0 & M_2^T \end{bmatrix} (M_1^{-1} A M_1) \begin{bmatrix} 1 & 0 \\ 0 & M_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & M_2^T \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & M_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & M_2^T A_1 M_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & D_1 \end{bmatrix} \end{aligned}$$

กำหนดให้ $M = M_1 M_3$ จะได้ว่า $M^{-1} A M = D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & D_1 \end{bmatrix}$ ดังนั้น M คือ เมทริกซ์เชิงตั้งฉาก และ D คือเมทริกซ์เส้นทแยงมุมที่มีค่าจะงอยู่ในแนวเส้นทแยงมุมที่ซึ่ง $M^{-1} A M = D$

จากทฤษฎีบทข้างต้น พบว่า $A = M D M^{-1} = M D M^T$ และหากให้ λ_k คือค่าจะง และให้ q_k คือแนวตั้งที่ k ของเมทริกซ์ M (นั่นคือ $q_k = M e_k$) ดังนั้น จะได้ว่า

$$A = M D M^T = \sum_{k=1}^n \lambda_k q_k q_k^T \quad (2.18)$$

ทฤษฎีบทที่ 2.1.3 เมทริกซ์สมมาตรค่าจริง A มิติ $n \times n$ สอดคล้องกับอสมการ

$$\lambda_{\min} x^T x \leq x^T A x \leq \lambda_{\max} x^T x \quad (2.19)$$

บทพิสูจน์

ให้ A เป็นเมทริกซ์สมมาตรค่าจริง ดังนั้นจากทฤษฎีบทที่ 2.1.2 เมทริกซ์ A สามารถเขียนได้เป็น $A = M D M^T = \sum_{k=1}^n \lambda_k q_k q_k^T$ ดังสมการที่ 2.18 และให้พิจารณาความสัมพัทธ์ต่อไปนี้

$$\begin{aligned} x^T A x &= x^T (M D M^T) x = (M^T x)^T D (M x) = \sum_{k=1}^n \lambda_k x^T q_k q_k^T x \\ &\leq \lambda_{\max} \sum_{k=1}^n x^T (q_k q_k^T) x = \lambda_{\max} \sum_{k=1}^n x_k^2 \\ &= \lambda_{\max} \|x\|_2^2 = \lambda_{\max} x^T x \end{aligned}$$

นั่นคือ $x^T A x \leq \lambda_{\max} x^T x$

โดยใช้หลักการพิจารณาเช่นเดียวกับในข้างต้น พบว่า $\lambda_{\min} x^T x \leq x^T A x$ ดังนั้น จึงสรุปได้ว่า $\lambda_{\min} x^T x \leq x^T A x \leq \lambda_{\max} x^T x$

ทฤษฎีบทที่ 2.1.4 ค่าจะงทุกตัวของเมทริกซ์เฮอร์มิเทียน (หรือสมมาตร) จะเป็นค่าจริงเสมอ

บทพิสูจน์

ให้ A เป็นเมทริกซ์เฮอร์มิเทียน (หรือสมมาตร)ใด ๆ และให้ x_k เป็นเวกเตอร์เฉพาะจริงที่สมมูลกับค่าเฉพาะ λ_k นั่นคือ $Ax_k = \lambda_k x_k$ ให้ x_k^* เป็นเวกเตอร์สลับเปลี่ยนสังยุคของ x_k ดังนั้น

$$x_k^* Ax_k = x_k^* \lambda_k x_k = \lambda_k x_k^* x_k \quad (2.20)$$

แต่เนื่องจาก A เป็นเมทริกซ์เฮอร์มิเทียน จะได้ว่า $A = A^*$ ดังนั้น

$$x_k^* Ax_k = x_k^* A^* x_k = (Ax_k)^* x_k = (\lambda_k x_k)^* x_k = \overline{\lambda_k} x_k^* x_k \quad (2.21)$$

จากสองสมการข้างต้น จะพบว่า $\lambda_k = \overline{\lambda_k}$ นั่นคือ λ_k ต้องเป็นค่าจริง และเนื่องจาก λ_k เป็นค่าเฉพาะจริงใด ๆ ของ A จึงสรุปได้ว่า ค่าเฉพาะจริงทุกตัวของเมทริกซ์เฮอร์มิเทียนต้องเป็นค่าจริงเสมอ

ผลบวกเฉียงของเมทริกซ์ (Trace of matrix)

นิยาม 2.1.27 ผลบวกเฉียงของเมทริกซ์ A ที่มีมิติ $n \times n$ คือผลรวมของสมาชิกที่บนเส้นทแยงมุมหลักของเมทริกซ์ A ซึ่งสามารถเขียนได้ดังนี้

$$Tr(A) = a_{11} + a_{22} + \dots + a_{nn} = \sum_{i=1}^n a_{ii} \quad (2.22)$$

ทฤษฎีบทที่ 2.1.5 ถ้า A เป็นเมทริกซ์ค่าจริงหรือค่าเชิงซ้อนที่มีมิติ $n \times n$ และให้ค่าเฉพาะของ A เท่ากับ $\lambda_1, \lambda_2, \dots, \lambda_n$ แล้วจะได้ว่า

$$Tr(A) = \sum_{i=1}^n \lambda_i \quad (2.23)$$

บทพิสูจน์

ให้ x_k แทนเวกเตอร์เฉพาะจริงที่สมมูลกับค่าเฉพาะ λ_k หากค่าเฉพาะที่ได้มีค่าแตกต่างกัน กล่าวคือ $\lambda_i \neq \lambda_j$ เมื่อ $i \neq j$ แล้ว ย่อมสามารถหาเวกเตอร์เฉพาะ x_i และ x_j ที่มีความอิสระเชิงเส้นต่อกันได้เสมอ แต่ในกรณีที่ค่าเฉพาะซ้ำกัน กล่าวคือ $\lambda_i = \lambda_j$ สำหรับบางค่า $i \neq j$ ในกรณีนี้อาจจะหาเวกเตอร์เฉพาะได้ไม่ครบ อันเนื่องจาก มิติของ $\ker(\lambda I - A)$ มีค่าน้อยกว่าจำนวนซ้ำของค่าเฉพาะ ซึ่งในกรณีนี้ แม้ว่าจะหาเวกเตอร์เฉพาะได้ไม่ครบ แต่ยังสามารถหาเวกเตอร์เฉพาะจนนับทั่วไป (Generalized eigen vector) ได้เพื่อทำให้ได้จำนวนเวกเตอร์ n เวกเตอร์เท่ากับจำนวนของค่าเฉพาะ

นำเวกเตอร์เฉพาะ (หรือเวกเตอร์เฉพาะจนับทั่วไป) มาเขียนเรียงกัน เพื่อสร้างเมทริกซ์การแปลง (Transformation matrix) ดังนี้ $M = [x_1, x_2, \dots, x_n]$ ซึ่งเมทริกซ์การแปลงดังกล่าว ทำหน้าที่แปลงเมทริกซ์ A ให้อยู่ในรูปของเมทริกซ์เส้นทแยงมุมโดยที่สมาชิกในแนวเส้นทแยงมุมคือค่าเฉพาะ (หรือในรูปแบบของจอร์แดน หากมีการใช้ เวกเตอร์เฉพาะจนับทั่วไป) นั่นคือ $M^{-1}AM = J$ โดยอาศัยทฤษฎีบทของผลบวกเฉียงที่ว่า

$$Tr(AB) = Tr(BA) \quad (2.24)$$

จะได้ว่า

$$\text{Tr}(J) = \text{Tr}(M^{-1}AM) = \text{Tr}(AM^{-1}M) = \text{Tr}(A) \quad (2.25)$$

แต่ไม่ว่า J จะเป็นเมทริกซ์เส้นทแยงมุมหรืออยู่ในรูปแบบของจอร์แดน สมาชิกในแนวเส้นทแยง

มุมคือค่าเฉพาะเสมอ ดังนั้น $\text{Tr}(A) = \text{Tr}(J) = \sum_{i=1}^n \lambda_i$

๒/๕.

๗๗๒๒๗

ตัวอย่างที่ 2.1.12 ผลบวกเฉลี่ยกับค่าเฉพาะของเมทริกซ์

จากตัวอย่างที่ 2.1.11 ค่าเฉพาะของ $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$ คือ $\lambda = 1, 4$ ดังนั้น $\sum \lambda_i = 5$ หากพิจารณา

ผลบวกเฉลี่ยของ A พบว่า $\text{Tr}(A) = 2 + 3 = 5 = \sum \lambda$

๒๕๕๐

รูปกำลังสองของเมทริกซ์ (Quadratic form of matrix)

นิยาม 2.1.28 ฟังก์ชันที่มีการส่งจาก $\mathbb{R}^n \rightarrow \mathbb{R}$ ที่อยู่ในรูปของ $f(x) = x^T A x = \sum_{i,j=1}^n a_{ij} x_i x_j$

จะเรียกว่า รูปกำลังสอง

หมายเหตุ โดยทั่วไป หากมีการพิจารณารูปกำลังสองของเมทริกซ์ A มักนิยามตั้งสมมติฐานว่า A เป็นเมทริกซ์สมมาตร (นั่นคือ $A = A^T$) เนื่องจาก

$$x^T A x = \frac{1}{2} x^T A x + \frac{1}{2} x^T A x = \frac{1}{2} x^T A x + \frac{1}{2} x^T A^T x = x^T \left(\frac{A + A^T}{2} \right) x = x^T P x \quad (2.26)$$

โดยที่ $P = \frac{A + A^T}{2}$ เรียกว่า ส่วนสมมาตรของ A (Symmetric part of A)

ตัวอย่างที่ 2.1.13 รูปกำลังสองกับส่วนสมมาตรของเมทริกซ์

ให้ $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$ รูปกำลังสองของ A คือ $x^T A x = 2x_1^2 + 3x_1x_2 + 3x_2^2$ หากจัดรูปใหม่ให้อยู่ในรูป

ของเมทริกซ์สมมาตร จะได้ว่า $P = \frac{A + A^T}{2} = \begin{bmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 3 \end{bmatrix}$ ซึ่งมีรูปกำลังสองเท่ากับรูปกำลังสองของ A

นั่นคือ $x^T P x = 2x_1^2 + 3x_1x_2 + 3x_2^2$

ความเป็นบวกแน่นอน ลบแน่นอน กึ่งบวก และ กึ่งลบของเมทริกซ์

เป็นที่ทราบดีว่า ระบบจำนวนจริงมีคุณสมบัติที่สำคัญอย่างหนึ่งที่ระบบแบบอื่นไม่มี กล่าวคือ กฎไตรวิภาค หรือ Trichotomy law ซึ่งกล่าวว่า สำหรับ $a \in \mathbb{R}$ แล้ว a จะสอดคล้องกับเงื่อนไขใดเพียงเงื่อนไขเดียวในสามเงื่อนไข ดังนี้

$$a < 0, \quad a = 0, \quad a > 0$$

ในกรณีของระบบจำนวนเชิงซ้อน หรือ เวกเตอร์ หรือ เมทริกซ์ หรือลำดับ พบว่าระบบเหล่านี้ขาดคุณสมบัติของไตรวิภาค ดังนั้น หากกำหนดให้ A, B เป็นเมทริกซ์ใด ๆ จึงจำเป็นต้องสร้างนิยามเพื่อให้ทำให้ สัญญลักษณ์ $A > B$ มีความหมาย

นิยาม 2.1.29 เมทริกซ์สมมาตรค่าจริง A มิติ $n \times n$ ถูกกล่าวไว้ว่า

1. เป็นเมทริกซ์บวกแน่นอน (Positive definite matrix) เขียนแทนด้วย $A > 0$ ถ้า $x^T A x > 0$ ทุกค่า $x \neq 0$ ใน $\mathbb{R}^n$
2. เป็นเมทริกซ์กึ่งบวก (Positive-semi definite matrix) เขียนแทนด้วย $A \geq 0$ ถ้า $x^T A x \geq 0$ ทุกค่า $x \neq 0$ ใน $\mathbb{R}^n$
3. เป็นเมทริกซ์ลบแน่นอน (Negative definite matrix) เขียนแทนด้วย $A < 0$ ถ้า $-A > 0$
4. เป็นเมทริกซ์กึ่งลบ (Negative-semi definite matrix) เขียนแทนด้วย $A \leq 0$ ถ้า $-A \geq 0$
5. ไม่มีนิยาม (Indefinite matrix) หาก A ไม่เข้าเงื่อนไขใด ๆ ในสี่ข้อข้างต้น

จากนิยามข้างต้น หากกำหนดให้ $A = A^T, B = B^T \in \mathbb{R}^{n \times n}$ จะใช้สัญญลักษณ์ $A \geq B$ ถ้า $A - B \geq 0$ และจะใช้สัญญลักษณ์ว่า $A < B$ ถ้า $B - A > 0$ เป็นเช่นนี้กับความสัมพันธ์อื่น ๆ

คุณสมบัติของอสมการเมทริกซ์

- 1) ถ้า $A \geq B$ และ $C \geq D$ แล้ว $A + C \geq B + D$
- 2) ถ้า $B \leq 0$ แล้ว $A + B \leq A$
- 3) ถ้า $A \geq 0$ และ $\alpha \geq 0$ แล้ว $\alpha A \geq 0$
- 4) ถ้า $A \geq 0$ แล้ว $A^2 \geq 0$
- 5) ถ้า $A > 0$ แล้ว $A^{-1} > 0$

ทฤษฎีบทที่ 2.1.6

1. เมทริกซ์ $A > 0$ ก็ต่อเมื่อ $\lambda_{\min}(A) > 0$
2. เมทริกซ์ $A \geq 0$ ก็ต่อเมื่อ $\lambda_{\min}(A) \geq 0$
3. เมทริกซ์ $A < 0$ ก็ต่อเมื่อ $\lambda_{\max}(A) < 0$
4. เมทริกซ์ $A \leq 0$ ก็ต่อเมื่อ $\lambda_{\max}(A) \leq 0$

บทพิสูจน์

ให้ A เป็นเมทริกซ์สมมาตรค่าจริง ดังนั้นจากทฤษฎีบทที่ 2.1.3 เมทริกซ์ A สอดคล้องกับอสมการ $\lambda_{\min} \|x\|^2 \leq x^T A x \leq \lambda_{\max} \|x\|^2$ หากกำหนดให้ $\lambda_{\min}(A) > 0$ จะได้ว่า $x^T A x > 0$ ทุกค่า $x \neq 0$ ใน $\mathbb{R}^n$ หรือนั่นคือ $A > 0$ ในทำนองกลับกัน ให้ M เป็นเมทริกซ์การแปลงเปลี่ยนฐานหลัก กล่าวคือ $x = Mz$ ดังนั้น รูปกำลังสองของ A นั่นคือ $x^T A x$ เมื่อเปลี่ยนฐานหลักแล้ว จะได้

$z^T M^T A M z = z^T D z$ หาก $A > 0$ นั้นหมายความว่า $x^T A x > 0$ ทุกค่า $x \neq 0$ ใน $\mathbb{R}^n$ ซึ่งสมมูลกับ $z^T D z > 0$ ทุกค่า $z \neq 0$ ใน $\mathbb{R}^n$ แต่เนื่องจาก D เป็นเมทริกซ์ทแยงมุมที่ซึ่งมีค่าเฉพาะจะอยู่ในแนวเส้นทแยงมุม ดังนั้นจากเงื่อนไขดังกล่าวข้างต้น จึงสรุปว่า ค่าเฉพาะของ A ต้องมีค่ามากกว่าศูนย์ ดังนั้น $\lambda_{\min}(A) > 0$ สำหรับการพิสูจน์ความเป็นกึ่งบวก ลบแน่นอน หรือกึ่งลบ ผู้อ่านสามารถใช้หลักการอธิบายในทำนองคล้ายคลึงกับข้างต้นได้

ตัวอย่างที่ 2.1.14 เมทริกซ์บวกแน่นอน

กำหนดให้ $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$ ซึ่งจะได้ว่ารูปกำลังสองของ A มีค่ามากกว่าศูนย์ทุกค่า $0 \neq x \in \mathbb{R}^2$ ดังเห็นได้จาก

$$x^T A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + 3x_2^2 > 0 \quad \text{สำหรับทุกค่า } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq 0 \text{ ใน } \mathbb{R}^2$$

หากพิจารณาความเป็นบวกของเมทริกซ์ A โดยดูจากค่าเฉพาะจะพบว่า ค่าเฉพาะของ A เท่ากับ 1, 3 ซึ่งมีค่ามากกว่าศูนย์ ดังนั้น จึงสรุปได้ว่า A เป็นเมทริกซ์บวกแน่นอน (Positive definite matrix)

ตัวอย่างที่ 2.1.15 เมทริกซ์กึ่งบวก

กำหนดให้ $A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ ซึ่งจะได้ว่ารูปกำลังสองของ A มีค่ามากกว่าหรือเท่ากับศูนย์ทุกค่า $0 \neq x \in \mathbb{R}^2$ ดังเห็นได้จาก

$$x^T A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - 2x_1x_2 + x_2^2 \geq 0 \quad \text{สำหรับทุกค่า } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq 0 \text{ ใน } \mathbb{R}^2$$

กล่าวคือ ถ้ากำหนดให้ $x_1 = 1$ และ $x_2 = 2$ จะได้ว่า $x_1^2 - 2x_1x_2 + x_2^2 = 1$ ซึ่งมีค่ามากกว่า 0แน่นอน และถ้ากำหนดให้ $x_1 = 1$ และ $x_2 = 1$ จะได้ว่า $x_1^2 - 2x_1x_2 + x_2^2 = 0$

หากพิจารณาความเป็นกึ่งบวกของเมทริกซ์ A โดยดูจากค่าเฉพาะจะพบว่าค่าเฉพาะของ $A = 0, 2$ ดังนั้น จึงสรุปได้ว่า A เป็นเมทริกซ์กึ่งบวก

ตัวอย่างที่ 2.1.16 เมทริกซ์ลบแน่นอน

กำหนดให้ $A = \begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix}$ เนื่องจาก $-A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$ เป็นเมทริกซ์บวกแน่นอน (จากตัวอย่างที่ 2.1.14) ดังนั้น จึงสรุปได้ว่า A เป็นเมทริกซ์ลบแน่นอน

ตัวอย่างที่ 2.1.17 เมทริกซ์กึ่งลบ

กำหนดให้ $A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ เนื่องจาก $-A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ เป็นเมทริกซ์กึ่งบวก (จากตัวอย่างที่ 2.1.15)

ดังนั้น จึงสรุปได้ว่า A เป็นเมทริกซ์กึ่งลบ

เมทริกซ์นอร์ม (Matrix norm)

กำหนดให้ $A \in \mathbb{R}^{m \times n}$ (ไม่จำเป็นต้องเป็นเมทริกซ์สมมาตรหรือเมทริกซ์จัตุรัส) และสำหรับแต่ละเวกเตอร์ที่ไม่ใช่เวกเตอร์ศูนย์ $0 \neq x \in \mathbb{R}^n$ ค่าของ $\frac{\|Ax\|_r}{\|x\|_x}$ บ่งบอกอัตราขยายของ A ในทิศทางของ x ดังนั้น เมื่อ x มีค่าต่าง ๆ กัน อัตราขยายที่ได้อาจแตกต่างกันด้วย จึงเกิดคำถามทั่วไปว่า อัตราขยายสูงสุด (หรือ ต่ำสุด) ของ A มีค่าเท่าไร จึงนำไปสู่นิยามของคำว่า เมทริกซ์นอร์ม ดังนี้

เมทริกซ์นอร์ม

นิยาม 2.1.30 อัตราขยายสูงสุด $\max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$ คือ เมทริกซ์นอร์ม หรือ สเปกตรัมนอร์ม (Spectral norm) ของ A และเขียนแทนด้วย $\|A\|$

ข้อสังเกต เนื่องจาก $\|Ax\|^2 = x^T A^T A x \leq \lambda_{\max}(A^T A) \|x\|^2$ ดังนั้น $\frac{\|Ax\|^2}{\|x\|^2} \leq \lambda_{\max}(A^T A)$ และจาก

นิยามของเมทริกซ์นอร์ม จะได้ว่า $\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sqrt{\lambda_{\max}(A^T A)}$ ในทำนองคล้ายกัน พบว่า

อัตราขยายต่ำสุดของ A คือ $\min_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sqrt{\lambda_{\min}(A^T A)}$

ตัวอย่างที่ 2.1.18 เมทริกซ์นอร์ม

ให้ $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ ที่ซึ่ง $A^T A = \begin{bmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{bmatrix}$ มีค่าเฉพาะจริงคือ $\lambda = 0, 0.5973, 90.4027$

ดังนั้น จากข้อสังเกต จะได้ว่า $0 \leq \frac{\|Ax\|}{\|x\|} \leq \sqrt{90.4027}$ เพื่อตรวจสอบ ให้คำนวณหาเวกเตอร์

เฉพาะจริงที่สอดคล้องกับค่าเฉพาะจริงต่ำสุดและสูงสุด ซึ่งพบว่า เวกเตอร์เฉพาะจริงที่สอดคล้องกับ $\lambda = 0$

คือ $x_{\lambda_{\min}} = \begin{bmatrix} -0.4082 \\ 0.8165 \\ -0.4082 \end{bmatrix}$ ที่ซึ่ง $\|x_{\lambda_{\min}}\| = 1$ และ $\|Ax_{\lambda_{\min}}\| = \left\| A \begin{bmatrix} -0.4082 \\ 0.8165 \\ -0.4082 \end{bmatrix} \right\| = 0$ ดังนั้น $\frac{\|Ax\|}{\|x\|} = 0$

และเวกเตอร์เงาจะงที่สอดคล้องกับ $\lambda = 90.4027$ คือ $x_{\lambda_{\max}} = \begin{bmatrix} 0.4287 \\ 0.5663 \\ 0.7039 \end{bmatrix}$ ที่ซึ่ง $\|x_{\lambda_{\max}}\| = 1$ และ

$\|Ax_{\lambda_{\max}}\| = \left\| A \begin{bmatrix} 0.4287 \\ 0.5663 \\ 0.7039 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 3.6731 \\ 8.7699 \end{bmatrix} \right\| = 9.5080$ ดังนั้น $\frac{\|Ax\|}{\|x\|} = 9.5080 = \|A\|$ ดังนั้นจึงสรุป

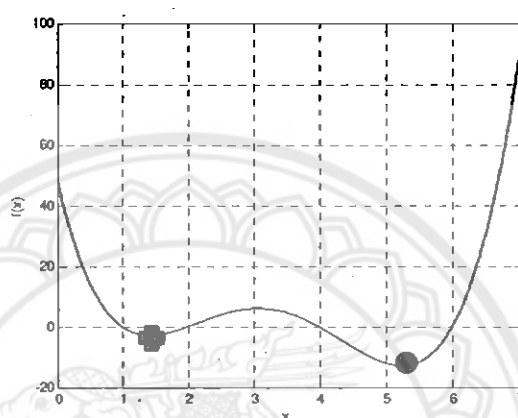
ได้ว่า $0 \leq \frac{\|Ax\|}{\|x\|} \leq \sqrt{90.4027} = 9.5080 = \|A\|$

คุณสมบัติของเมทริกซ์นอร์ม

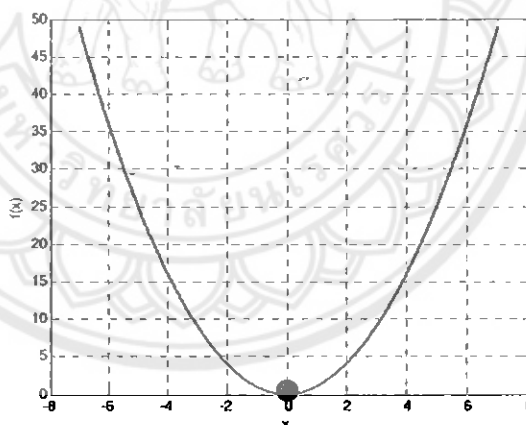
- 1) เมทริกซ์นอร์มของ $a \in \mathbb{R}^{n \times 1}$ (เวกเตอร์นอร์ม) มีค่าเท่ากับ $\sqrt{\lambda_{\max}(a^T a)} = \sqrt{a^T a}$
- 2) สำหรับเวกเตอร์ x ใดๆ พบว่า $\|Ax\| \leq \|A\| \|x\|$
- 3) $\|aA\| = \|a\| \|A\|$
- 4) อสมการสามเหลี่ยม $\|A + B\| \leq \|A\| + \|B\|$
- 5) $\|A\| = 0$ ก็ต่อเมื่อ $A = 0$
- 6) $\|AB\| \leq \|A\| \|B\|$

2.2 ความเป็นคอนเวกซ์ (Convexity) ([1, 11])

ในการหาผลเฉลยของปัญหาการหาค่าที่เหมาะสมที่สุด (Optimization problem) สิ่งที่สำคัญสิ่งหนึ่งคือเรื่องของความเป็นคอนเวกซ์ของเซตหรือฟังก์ชัน ซึ่งถ้าฟังก์ชันที่กำลังพิจารณาอยู่นั้นมี ความเป็นคอนเวกซ์ และถูกนิยามบนเซตคอนเวกซ์แล้ว ฟังก์ชันดังกล่าวจะเกิดค่าสุดขีดแบบจุลภาค (Local extremum) ขึ้นเสมอ ยิ่งไปกว่านั้นค่าสุดขีดที่ได้ยังเป็นค่าสุดขีดแบบมหภาค (Global extremum) ด้วย ดังแสดงในรูปตัวอย่างต่อไปนี้



รูปที่ 2.2 ฟังก์ชันที่ไม่คอนเวกซ์



รูปที่ 2.3 ฟังก์ชันที่คอนเวกซ์

ฟังก์ชันในรูปที่ 2.2 เป็นฟังก์ชันที่ไม่คอนเวกซ์ ดังนั้นหากจุดเริ่มต้นในการคำนวณหาค่าต่ำสุดอยู่ในย่าน $[1, 2]$ จะทำให้ค่าต่ำสุดของฟังก์ชันลู่เข้าสู่จุดที่ทำเครื่องหมายบวกลบในรูป แต่หากจุดเริ่มต้นในการคำนวณอยู่ในย่าน $[4, 6]$ จะทำให้ค่าต่ำสุดของฟังก์ชันลู่เข้าสู่จุดที่ทำเครื่องหมายวงกลม จะพบว่าค่าต่ำสุดที่ได้ขึ้นอยู่กับจุดเริ่มต้นของการคำนวณ ค่าต่ำสุดที่ได้ในแต่ละครั้งของการคำนวณ อาจจะไม่ใช่ค่าต่ำสุดของฟังก์ชันในเซตที่พิจารณา แต่หากพิจารณาฟังก์ชันในรูปที่ 2.3 ซึ่งเป็นฟังก์ชันที่คอนเวกซ์ พบว่า ไม่ว่าจะเริ่มต้นคำนวณที่จุดใด ๆ ในเซต ค่าต่ำสุดของฟังก์ชันจะลู่เข้า

คู่จุดเดียวกันเสมอ นั่นคือจุดที่ทำเครื่องหมายวงกลม ดังนั้น ความเป็นคอนเวกซ์ของฟังก์ชัน จึงมีความสำคัญในเรื่อง การหาค่าเหมาะที่สุด

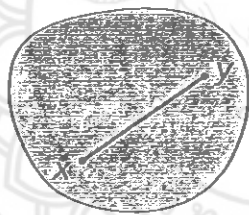
2.2.1 เซตคอนเวกซ์ (Convex set)

เซตย่อยในปริภูมิยูคลิดจัดว่าเป็นเซตคอนเวกซ์ถ้าสามารถสร้างเส้นตรงเชื่อมระหว่างทุก ๆ คู่สมาชิกใด ๆ ในเซตย่อยดังกล่าวได้โดยเส้นตรงนั้นต้องเป็นส่วนหนึ่งในเซตย่อย นั่นคือที่มาของนิยามต่อไปนี้

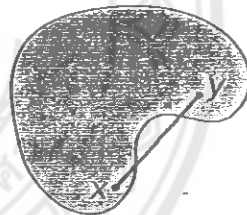
นิยาม 2.2.1 ให้ $S \subseteq \mathbb{R}^n$ (หรือ $\mathbb{C}^n$) แล้ว S ถูกกล่าวว่าเป็น เซตคอนเวกซ์ (Convex set) ถ้าทุกเวกเตอร์ $x, y \in S$ และ $\lambda \in [0, 1]$ แล้ว เวกเตอร์ที่เกิดจากการรวมเชิงคอนเวกซ์ของ x และ y ย่อมเป็นสมาชิกของเซต S นั่นคือ

$$z = \lambda x + (1 - \lambda)y \in S \quad (2.27)$$

ตัวอย่างของเซตคอนเวกซ์ได้แก่ เซตโตน (Singleton) เซตว่าง (Empty set) เซตของ $\mathbb{R}^n$ เซตของช่วง (interval) ในระบบจำนวนจริง $\mathbb{R}$ หรือเซตที่ไม่มีการเว้า (indentation) ใน $\mathbb{R}^2$ เป็นต้น



รูปที่ 2.4 (a) เซตที่คอนเวกซ์



รูปที่ 2.4 (b) เซตที่ไม่คอนเวกซ์

ทฤษฎีบทที่ 2.2.1 การอินเตอร์เซกชันระหว่างเซตคอนเวกซ์ จะได้เซตคอนเวกซ์ด้วย

บทพิสูจน์

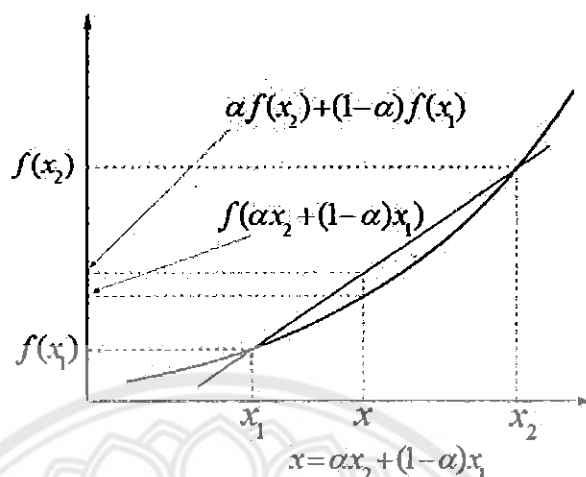
ให้ S_k เป็นเซตคอนเวกซ์ใด ๆ และให้ $x, y \in \cap S_k$ หรือนั่นคือ x, y อยู่ในเซตคอนเวกซ์ทุกเซต ให้ z เป็นการรวมกันเชิงคอนเวกซ์ของ x และ y ซึ่งจะได้ว่า $z \in S_k$ ทุกเซตเช่นกัน หรือนั่นคือ $z \in \cap S_k$ ดังนั้น $\cap S_k$ ย่อมเป็นเซตคอนเวกซ์

2.2.2 ฟังก์ชันคอนเวกซ์ (Convex function)

นิยาม 2.2.2 ให้ $S \subseteq \mathbb{R}^n$ เป็นเซตคอนเวกซ์แล้ว ฟังก์ชัน $f: S \rightarrow \mathbb{R}$ ถูกกล่าวว่าเป็น ฟังก์ชันคอนเวกซ์ (Convex function) ถ้าทุกเวกเตอร์ $x, y \in S$ และ $\lambda \in [0, 1]$ แล้ว

$$f(\alpha x_2 + (1 - \alpha)x_1) \leq \alpha f(x_2) + (1 - \alpha)f(x_1) \quad (2.28)$$

และฟังก์ชัน f ถูกกล่าวว่าเป็น ฟังก์ชันคอนเวกซ์ (Concave function) ถ้า $-f$ เป็นฟังก์ชันคอนเวกซ์



รูปที่ 2.5 ฟังก์ชันคอนเวกซ์

จากนิยามข้างต้น อาจจะพิจารณาได้ว่า ฟังก์ชันคอนเวกซ์ คือฟังก์ชันที่มีลักษณะ โค้งงอขึ้นด้านบน ดังแสดงในรูปที่ 2.5 และฟังก์ชันคอนเคฟ คือฟังก์ชันที่มีลักษณะ โค้งงอลงด้านล่าง

ตัวอย่างที่ 2.2.1 ฟังก์ชันคอนเวกซ์

1. ฟังก์ชันเชิงเส้นในเซตคอนเวกซ์เป็นฟังก์ชันคอนเวกซ์
2. ฟังก์ชันพาราโบลา $f(x) = x^2$ นิยามบนเซตจำนวนจริง เป็นฟังก์ชันคอนเวกซ์
3. ฟังก์ชันไซน์ $f(x) = \sin x$ นิยามบนช่วงปิด $[\pi, 2\pi]$ เป็นฟังก์ชันคอนเวกซ์

คุณสมบัติของฟังก์ชันคอนเวกซ์

1. ฟังก์ชันคอนเวกซ์ f ที่นิยามบนช่วงเปิด $(a, b) \subset \mathbb{R}$ จะมีความต่อเนื่องบนช่วงเปิดดังกล่าว นอกจากนี้ f มีความสามารถหาอนุพันธ์ได้เกือบทุกแห่ง และถ้า f นิยามบนช่วงปิด $[a, b]$ แล้ว f อาจจะขาดคุณสมบัติความต่อเนื่องที่จุดปลาย
2. ฟังก์ชัน f ที่หาอนุพันธ์ได้อย่างต่อเนื่อง (Continuously differentiable) ในหนึ่งตัวแปร เป็นฟังก์ชันคอนเวกซ์บนช่วง $I \subset \mathbb{R}$ ก็ต่อเมื่อ $f(y) \geq f(x) + f'(x)(y-x)$ สำหรับทุกค่า $x, y \in I$ และหากพบว่า $f'(z) = 0$ แล้ว z คือจุดต่ำสุดมหภาค (Global minimum) ของ $f(x)$
3. ฟังก์ชัน f ในหนึ่งตัวแปรที่สามารถหาอนุพันธ์ได้สองครั้ง เป็นฟังก์ชันคอนเวกซ์ภายในช่วง $I \subset \mathbb{R}$ ก็ต่อเมื่อ $f'' \geq 0$ บนช่วงดังกล่าว และหาก $f'' > 0$ จะกล่าวว่า f เป็นฟังก์ชันคอนเวกซ์โดยแท้ (Strictly convex) แต่บทกลับไม่เป็นจริงเสมอไป ยกตัวอย่างเช่น $f(x) = x^4$

4. ฟังก์ชันต่อเนื่องหลายตัวแปร $f(x_1, x_2, \dots, x_n)$ ที่สามารถหาอนุพันธ์ได้อย่างน้อยสองครั้งบนเซตคอนเวกซ์ S เป็นฟังก์ชันคอนเวกซ์ ก็ต่อเมื่อ เมทริกซ์เฮสเซียน (Hessian matrix) ที่สอดคล้องกับฟังก์ชันดังกล่าวมีคุณสมบัติเป็นกึ่งบวกที่ทุกจุดภายใน (Interior) ของ S โดยที่เมทริกซ์เฮสเซียน คือเมทริกซ์ที่เขียนอยู่ในรูป $H(x) = \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right]$

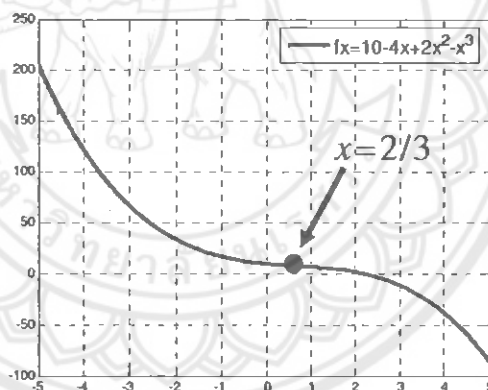
ตัวอย่างที่ 2.2.2 การตรวจสอบความเป็นคอนเวกซ์ของฟังก์ชัน $f(x_1, x_2) = x_1^2 + x_2^2 - 1$

ใช้นิยามของเมทริกซ์เฮสเซียน จะได้

$$H(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

เนื่องจากค่าเฉพาะของ H มีค่า $\lambda = 2, 2$ ดังนั้น ฟังก์ชันที่กำหนดมาเป็นฟังก์ชันคอนเวกซ์

ตัวอย่างที่ 2.2.3 การตรวจสอบความเป็นคอนเวกซ์ของฟังก์ชัน $f(x) = 10 + 4x + 2x^2 - x^3$



รูปที่ 2.6 ฟังก์ชันที่เป็นทั้งคอนเวกซ์และคอนเคฟ

พิจารณาอนุพันธ์อันดับที่สองของฟังก์ชัน ดังนี้

$$\frac{d^2 f(x)}{dx^2} = 4 - 6x$$

เนื่องจาก f จะเป็นฟังก์ชันคอนเวกซ์ก็ต่อเมื่อ $\frac{d^2 f(x)}{dx^2} \geq 0$ ดังนั้นจะได้ว่าฟังก์ชันนี้จะเป็นคอนเวกซ์เมื่อ $4 - 6x \geq 0$ หรือนั่นคือ $x \leq \frac{2}{3}$ และเมื่อ $x \geq \frac{2}{3}$ ฟังก์ชันนี้จะเป็นฟังก์ชันคอนเคฟ

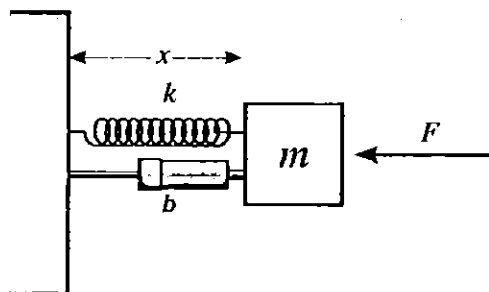
2.3 ทฤษฎีพื้นฐานของระบบควบคุม (Basic Control System Theory) ([3, 4, 8])

ในการวิเคราะห์ระบบควบคุมสามารถวิเคราะห์ได้หลายวิธี เช่น การวิเคราะห์เชิงปริภูมิสถานะ (State space analysis) ซึ่งเป็นการวิเคราะห์ในโดเมนของเวลา (Time domain) หรือการวิเคราะห์เชิงฟังก์ชันถ่ายโอน (Transfer function approach) ซึ่งเป็นการวิเคราะห์ในโดเมนของความถี่ (Frequency domain) โดยอาศัยการแปลงลาปลาซ (Laplace transform) เข้ามาช่วยในการแปลงระบบที่บรรยายในโดเมนของเวลาให้อยู่ในรูปแบบที่บรรยายในโดเมนของความถี่ ซึ่งแต่ละวิธีการสามารถใช้ในการวิเคราะห์ระบบที่ต่างกัน ดังนั้นในการวิเคราะห์ระบบควบคุมจึงควรวิเคราะห์ระบบทั้งในเชิงปริภูมิสถานะและในเชิงฟังก์ชันถ่ายโอนควบคู่กัน

2.3.1 การวิเคราะห์เชิงปริภูมิสถานะ (State space analysis)

ในการบรรยายสมการเชิงอนุพันธ์ทั่วไปนั้นสามารถบรรยายหรือเขียนอยู่ในรูปของระบบสมการเชิงอนุพันธ์อันดับที่หนึ่งโดยที่ผลเฉลยของระบบสมการดังกล่าวเรียกว่า ตัวแปรสถานะ (State variable) ที่ซึ่งตัวแปรสถานะแต่ละตัวทำหน้าที่เป็นฐานหลักในปริภูมิสถานะ (State space) ความคิดในการบรรยายสมการเชิงอนุพันธ์ในรูปแบบปริภูมิสถานะ (State-space description) นี้ก่อให้เกิดวิธีการใหม่ๆ ในการออกแบบระบบควบคุมซึ่งเป็นพื้นฐานของการออกแบบระบบควบคุมที่เรียกว่าการควบคุมสมัยใหม่ (Modern control) ที่มีวิธีการแตกต่างจากวิธีการที่เรียกว่าการควบคุมแบบคลาสสิก (Classical control) โดยวิธีการออกแบบตัวควบคุมแบบคลาสสิกนั้นจะใช้สมการฟังก์ชันถ่ายโอนเป็นหลักในการออกแบบ ในหัวข้อนี้จะกล่าวถึงการวิเคราะห์ระบบควบคุมโดยใช้การพรรณนาแบบปริภูมิสถานะเป็นหลัก ซึ่งจะเรียกวิธีการออกแบบระบบควบคุมนี้ว่าการออกแบบโดยวิธีปริภูมิสถานะ (State space design)

พิจารณาระบบสปริง-มวล-ตัวหน่วงดังแสดงในรูปที่ 2.7 ซึ่งนำมาเป็นตัวอย่างเพื่ออธิบายถึงการเขียนระบบให้อยู่ในรูปของปริภูมิสถานะ



รูปที่ 2.7 ระบบสปริง-มวล-ตัวหน่วง (Spring-mass-damper system)

จากรูปที่ 2.7 สามารถเขียนสมการแทนการเคลื่อนที่ของระบบโดยอาศัยหลักของนิวตันได้ดังนี้

$$m\ddot{x} + b\dot{x} + kx = F(t) \quad (2.29)$$

โดยที่ m คือมวล b, k คือค่าคงที่ในตัวหน่วงและสปริงตามลำดับ และ F คือแรงภายนอก ถ้ากำหนดให้ $x_1 = x$ และ $x_2 = \dot{x}$ จากสมการที่ (2.29) สามารถเปลี่ยนตัวแปร $x, \dot{x}$ ให้อยู่ในรูปของตัวแปร x_1, x_2 ได้ดังนี้คือ

$$\dot{x}_1 = x_2 \quad (2.30)$$

และ

$$\dot{x}_2 = -\frac{b}{m}x_2 - \frac{k}{m}x_1 + \frac{1}{m}F(t) \quad (2.31)$$

หรือเขียนให้อยู่ในรูปของเมทริกซ์ได้ดังนี้

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} F(t) \quad (2.32)$$

รูปแบบของสมการที่ (2.32) นี้เรียกว่ารูปแบบปริภูมิสถานะ ซึ่งจะเห็นว่าได้รูปแบบสมการเชิงอนุพันธ์อันดับที่สอง 1 สมการถูกแปลงให้เป็นสมการเชิงอนุพันธ์อันดับที่หนึ่ง 2 สมการ ในกรณีทั่วไป กระบวนการเปลี่ยนตัวแปรข้างต้น สามารถแปลงสมการเชิงอนุพันธ์สามัญอันดับที่ n ใด ๆ หนึ่งสมการ ให้เป็นระบบของสมการเชิงอนุพันธ์สามัญอันดับหนึ่ง n สมการได้

$$\text{หากกำหนดให้ } x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} \in \mathbb{R}^n \text{ และ } u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_m(t) \end{bmatrix} \in \mathbb{R}^m \text{ แล้วรูปแบบปริภูมิ}$$

สถานะของระบบสมการเชิงเส้นที่ตัวพารามิเตอร์ของระบบไม่ขึ้นกับเวลา (Linear time invariant system) นั้นสามารถเขียนได้ดังนี้

$$\dot{x} = Ax + Bu \quad (2.33)$$

ซึ่งสำหรับตัวอย่างข้างต้นนี้ จะได้ว่า

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} \quad (2.34)$$

และถ้าต้องการให้สัญญาณออกคือ $y = x$ แล้วสามารถเขียนได้ดังนี้ คือ

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (2.35)$$

หรือในรูปแบบ

$$y = Cx \quad (2.36)$$

ดังนั้น รูปแบบทั่วไปของสมการปริภูมิสถานะสามารถสรุปได้ดังนี้

$$\dot{x} = Ax + Bu \quad (2.37)$$

$$y = Cx + Du \quad (2.38)$$

โดยที่ A คือ เมทริกซ์ระบบ (System matrix) ($n \times n$)

B คือ เมทริกซ์สัญญาณอินพุต (Input matrix) ($n \times m$)

C คือ เมทริกซ์สัญญาณเอาต์พุต (Output matrix) ($p \times n$)

D คือ เมทริกซ์สัญญาณป้อนไปข้างหน้า (Feedforward matrix) ($p \times m$)

n คือ อันดับของระบบ (System order)

m คือ จำนวนสัญญาณเข้า (Number of input)

p คือ จำนวนสัญญาณออก (Number of output)

2.3.1.1 การหาผลเฉลยของระบบ (System Solution)

พิจารณาสมการเอกพันธ์อันดับหนึ่ง $\frac{dx(t)}{dt} = ax(t)$ โดยที่ $a \neq 0$ และมีเงื่อนไขเริ่มต้นคือ $x(t_0) = x_0$ ผลเฉลยเฉพาะของสมการดังกล่าวคือ $x(t) = e^{a(t-t_0)}x_0$ แต่ในกรณีทั่วไปที่มีแรงภายนอกมากระทำ กล่าวคือ สมการเชิงอนุพันธ์สามัญที่อยู่ในรูปของ $\frac{dx(t)}{dt} = ax(t) + bu(t)$ โดยที่ a, b เป็นค่าคงที่และ u คือแรงภายนอก (Forcing function) ซึ่งผลเฉลยเฉพาะของสมการดังกล่าวคือ

$$x(t) = e^{a(t-t_0)}x_0 + \int_{t_0}^t e^{-a(t-\tau)}bu(\tau)d\tau \quad (2.39)$$

สังเกตว่าการหาผลเฉลยนี้สามารถกระทำได้ง่ายจากพื้นฐานความรู้ของรายวิชาแคลคูลัส อย่างไรก็ตาม เมื่อ a และ b ไม่ใช่ปริมาณสเกลาร์ การหาผลเฉลยจะต้องอาศัยทฤษฎีบทที่มีรากฐานมาจากรายวิชาการวิเคราะห์จำนวนจริง ดังนี้

ทฤษฎีบทที่ 2.3.1

ให้ A เป็นเมทริกซ์มิตีจักระขนาด $n \times n$ ใดๆ แล้วฟังก์ชันเมทริกซ์ยกกำลัง e^{At} สามารถนิยามได้ด้วยอนุกรมกำลัง

$$e^{A(t-t_0)} = \sum_{n=0}^{\infty} \frac{A^n (t-t_0)^n}{n!} \quad (2.40)$$

ซึ่งจะลู่เข้าเอกรูป (Uniform convergence) และลู่เข้าสัมบูรณ์ (Absolute convergence) บนช่วงปิด $[-T, T]$ สำหรับค่า T ใดๆ ที่มากกว่าศูนย์

บทพิสูจน์

ภายในช่วงปิด $[-T, T]$ ฟังก์ชันเมทริกซ์ที่อยู่ในรูปอนุกรมอนันต์ในสมการที่ (2.40) มีขอบเขตเนื่องจาก

$$\left\| \sum_{n=0}^{\infty} \frac{A^n t^n}{n!} \right\| \leq \sum_{n=0}^{\infty} \left\| \frac{A^n t^n}{n!} \right\| \leq \sum_{n=0}^{\infty} \frac{\|A\|^n t^n}{n!} \leq \sum_{n=0}^{\infty} \frac{\|A\|^n T^n}{n!} \quad (2.41)$$

เนื่องจากอนุกรม $\sum_{n=0}^{\infty} \frac{\|A\|^n T^n}{n!}$ เป็นอนุกรมค่าจริงที่ลู่เข้าสู่ฟังก์ชัน $e^{\|A\|T}$ ดังนั้นโดยใช้ Weierstrass M-test จึงสรุปได้ว่าอนุกรมทางฝั่งขวาของสมการที่ (2.40) ลู่เข้าเอกรูป (Uniform convergence) และลู่เข้าสัมบูรณ์ (Absolute convergence) บนช่วงปิด $[-T, T]$

จากคุณสมบัติของการลู่เข้าในทฤษฎีบทนี้เองส่งผลให้ฟังก์ชันเมทริกซ์ยกกำลังมีคุณสมบัติของการวิเคราะห์ได้ (Analyticity) ภายในช่วงปิดจำกัดใดๆ นอกจากนี้จะพบอีกด้วยว่า ฟังก์ชันเมทริกซ์ยกกำลัง e^{At} ที่ถูกนิยามด้วยอนุกรมกำลัง $\sum_{n=0}^{\infty} \frac{A^n t^n}{n!}$ สามารถหาอนุพันธ์ได้ ดังนี้

$$\frac{d}{dt} e^{At} = \frac{d}{dt} \left[\sum_{n=0}^{\infty} \frac{A^n t^n}{n!} \right] = \sum_{n=0}^{\infty} \left[\frac{d}{dt} \left(\frac{A^n t^n}{n!} \right) \right] = \begin{cases} A \sum_{n=0}^{\infty} \frac{A^n t^n}{n!} := Ae^{At} \\ \sum_{n=0}^{\infty} \frac{A^n t^n}{n!} A := e^{At} A \end{cases} \quad (2.42)$$

เนื่องจากฟังก์ชันเมทริกซ์ยกกำลังที่ถูกนิยามโดยอนุกรมกำลังในสมการที่ (2.40) มีคุณสมบัติคล้ายกับคุณสมบัติของฟังก์ชันยกกำลัง ดังนั้น การหาผลเฉลยของสมการเชิงอนุพันธ์สามัญอันดับหนึ่ง $\frac{dx}{dt} = Ax + Bu$ จึงสามารถกระทำได้เช่นเดียวกับกรณีปริมาณสเกลาร์ ซึ่งจะได้ผลเฉลยของสมการคือ

$$x(t) = e^{A(t-t_0)} x_0 + \int_{t_0}^t e^{A(t-\tau)} B u(\tau) d\tau \quad (2.43)$$

ในการคำนวณหาผลเฉลยของระบบสมการเชิงเส้น นอกจากวิธีที่นำเสนอในข้างต้นแล้ว ยังสามารถใช้วิธีการแปลงลาปลาซ ดังนี้ พิจารณาสมการเชิงเส้น $\frac{dx}{dt} = Ax + Bu$ โดยมีเงื่อนไขเริ่มต้นคือ $x(0) = x_0$ เนื่องจากความเป็นเชิงเส้น จึงสามารถใส่ตัวดำเนินการลาปลาซทั้งสองข้างของสมการ ซึ่งจะได้เป็น

$$sX(s) - x(0) = AX(s) + BU(s)$$

หรือ

$$(sI - A)X(s) = x_0 + BU(s)$$

ดังนั้น

$$\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{x}_0 + (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{B}\mathbf{U}(s)$$

ซึ่งจะได้ว่า

$$\mathbf{x}(t) = \mathbf{L}^{-1} \left[(s\mathbf{I} - \mathbf{A})^{-1} \right] \mathbf{x}_0 + \mathbf{L}^{-1} \left[(s\mathbf{I} - \mathbf{A})^{-1} \mathbf{B}\mathbf{U}(s) \right]$$

โดยที่ $\mathbf{L}^{-1}$ คือตัวดำเนินการลาปลาซผกผัน

แต่เนื่องจากผลเฉลยทั่วไปของสมการ $\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$ เมื่อกำหนดเงื่อนไขเริ่มต้น $\mathbf{x}(0) = \mathbf{x}_0$ คือ

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}_0 + \int_0^t e^{\mathbf{A}(t-\tau)} \mathbf{B}\mathbf{u}(\tau) d\tau$$

และจากทฤษฎีบทที่สำคัญในรายวิชาสมการเชิงอนุพันธ์สามัญกล่าวว่า หากระบบสมการเชิงอนุพันธ์สามัญเชิงเส้นมีผลเฉลยเฉพาะ (Particular solutions) แล้วผลเฉลยเฉพาะดังกล่าวจะมีเพียงผลเฉลยเดียว (Uniqueness property) ดังนั้น จะได้ว่า

$$e^{\mathbf{A}t} = \mathbf{L}^{-1} \left[(s\mathbf{I} - \mathbf{A})^{-1} \right] \text{ และ } \int_0^t e^{\mathbf{A}(t-\tau)} \mathbf{B}\mathbf{u}(\tau) d\tau = \mathbf{L}^{-1} \left[(s\mathbf{I} - \mathbf{A})^{-1} \mathbf{B}\mathbf{U}(s) \right]$$

ยกตัวอย่างเช่น หากกำหนดให้

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & -2 \end{bmatrix}$$

ดังนั้น

$$e^{\mathbf{A}t} = \mathbf{L}^{-1} \left[(s\mathbf{I} - \mathbf{A})^{-1} \right] = \mathbf{L}^{-1} \left[\begin{bmatrix} s & -1 \\ 0 & s+2 \end{bmatrix}^{-1} \right] = \mathbf{L}^{-1} \left[\begin{bmatrix} \frac{1}{s} & \frac{1}{s(s+2)} \\ 0 & \frac{1}{s+2} \end{bmatrix} \right] = \begin{bmatrix} 1 & \frac{1}{2}(1-e^{-2t}) \\ 0 & e^{-2t} \end{bmatrix}$$

2.3.1.2 ความมีเสถียรภาพ (Stability)

ในการพิจารณาความมีเสถียรภาพของระบบควบคุมเชิงเส้น สามารถแยกพิจารณาได้หลายประเภท แต่ที่สำคัญได้แก่ เสถียรภาพภายใน (Internal stability) และ เสถียรภาพแบบสัญญาณเข้าและสัญญาณออกถูกจำกัดขนาด (Bounded Input-Bounded Output stability, BIBO) โดยเสถียรภาพภายในหมายถึง เสถียรภาพอันเกิดจากตัวของตัวเอง ซึ่งการตรวจสอบเสถียรภาพภายในของระบบที่เขียนพรรณนาในเชิงปริภูมิสถานะนั้นดูที่ตำแหน่งของค่าเฉพาะ ดังนี้

ทฤษฎีบทที่ 2.3.2 ระบบเชิงเส้น $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$, $\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}$ ถูกกล่าวว่า

1. มีเสถียรภาพภายใน ถ้าส่วนจริงของค่าเฉพาะของ $\mathbf{A}$ มีค่าติดลบทุกค่า นั่นคือ $\text{Re } \lambda_k(\mathbf{A}) < 0$ สำหรับทุกค่า k
2. มีเสถียรภาพภายในแบบวิกฤติ ถ้าส่วนจริงของค่าเฉพาะของ $\mathbf{A}$ มีค่าติดลบหรือเป็นศูนย์ นั่นคือ $\text{Re } \lambda_k(\mathbf{A}) \leq 0$ สำหรับทุกค่า k

3. ขาดเสถียรภาพ ถ้าส่วนจริงของค่าเฉพาะของ A มีค่าเป็นบวกบางค่า นั่นคือ $\operatorname{Re} \lambda_k(A) > 0$ สำหรับบางค่า k

นอกจากนี้ ยังมีอีกหลายวิธีในการพิจารณาว่าระบบนั้นมีเสถียรภาพภายในหรือไม่ อาทิ เช่น การหาเสถียรภาพของลีโอทูนอฟ ซึ่งสามารถใช้อธิบายได้ทั้งระบบเชิงเส้น และระบบไม่เชิงเส้น เนื่องจากลีโอทูนอฟ พิจารณาพลังงานภายในของระบบ และอาศัยหลักการอย่างง่ายที่ว่า หากพลังงานในระบบลดลง ระบบจะเริ่มเคลื่อนที่เข้าหาจุดสมดุล ยกตัวอย่างเช่น ลูกตุ้ม (Pendulum) ที่แกว่งอยู่ภายใต้แรงดึงดูดของโลก ซึ่งการแกว่งนั้นเกิดจากการเปลี่ยนพลังงานศักย์มาเป็นพลังงานจลน์สลับกันไป และระหว่างการเปลี่ยนพลังงานนั้น หากเกิดการสูญเสียพลังงานออกไปจากระบบ เช่น เกิดแรงเสียดทานระหว่างลูกตุ้มและอากาศ ทำให้พลังงานรวมของระบบลดลง จนในที่สุด ลูกตุ้มมาหยุดนิ่งที่สภาวะคงตัว ลักษณะแบบนี้ จะกล่าวว่า ลูกตุ้มมีเสถียรภาพภายใน ซึ่งลีโอทูนอฟได้เสนอความคิดนี้ขึ้นในปี ค.ศ. 1899

จากหลักการข้างต้น หากกำหนดให้ ฟังก์ชัน $V(x): \mathbb{R}^n \rightarrow \mathbb{R}$ ที่ซึ่ง

- $V(x) > 0$
- $\frac{dV(x)}{dt} \leq 0$

ฟังก์ชันที่มีคุณสมบัติดังกล่าว ถูกเรียกว่า ฟังก์ชันลีโอทูนอฟ (Lyapunov function) ซึ่งมีลักษณะเหมือนฟังก์ชันพลังงานที่ใช้พิจารณาเสถียรภาพของระบบ กล่าวคือ พลังงานย่อมมีค่าเป็นบวก และหากพลังงานรวมของระบบมีค่าลดลง แสดงว่า ระบบย่อมมีเสถียรภาพนั่นเอง

ในกรณีของระบบเชิงเส้น ไม่แปรตามเวลาที่พรรณนาในรูปแบบของตัวแปรสถานะ $\dot{x} = Ax$ ฟังก์ชันลีโอทูนอฟที่เหมาะสมคือ $V(x) = x^T P x$, $P > 0$ ซึ่งพบว่า

$$\frac{dV(x)}{dt} = \dot{x}^T P x + x^T P \dot{x} = (Ax)^T P x + x^T P A x = x^T (A^T P + P A) x \quad (2.44)$$

ดังนั้น $\frac{dV(x)}{dt} \leq 0$ หมายความว่า

$$A^T P + P A \leq 0 \quad (2.45)$$

ให้ $Q > 0$ แล้ว อสมการที่ (2.45) สมมูลกับสมการต่อไปนี้ ซึ่งเรียกว่า สมการของลีโอทูนอฟ (Lyapunov equation)

$$A^T P + P A + Q = 0 \quad (2.46)$$

ทฤษฎีบทที่ 2.3.3 ระบบเชิงเส้นไม่แปรตามเวลา $\dot{x} = Ax$ มีเสถียรภาพภายใน ก็ต่อเมื่อ P ที่สอดคล้องกับสมการลีโอทูนอฟ มีค่าเป็นบวกแน่นอน (Positive definite matrix) สำหรับทุก ๆ เมทริกซ์ $Q > 0$

นอกจากพิจารณาเสถียรภาพภายในของระบบแล้ว ยังมีเสถียรภาพอีกประเภทที่สำคัญ เรียกว่า เสถียรภาพของระบบที่มีสัญญาณเข้าและสัญญาณออกถูกจำกัดขนาด (Bounded Input – Bounded Output: BIBO) กล่าวคือ ถ้าระบบมีเสถียรภาพแบบ BIBO แล้ว สัญญาณออกจากระบบจะมีขอบเขต เมื่อป้อนสัญญาณที่มีขอบเขตเข้าสู่ระบบเสมอ โดยเงื่อนไขที่ทำให้ระบบเชิงเส้นไม่แปรตามเวลามีคุณสมบัติความเสถียรภาพแบบ BIBO คือ ผลตอบสนองต่อสัญญาณอิมพัลส์ (Impulse response) สามารถหาปริพันธ์ได้อย่างสมบูรณ์ (Absolutely integrable) หรือนั่นคือ

$$\|h\|_1 := \int_{-\infty}^{\infty} |h(t)| dt = M < \infty \quad (2.47)$$

ในทางปฏิบัติ สัญญาณเข้าสู่ระบบจะถูกสร้างขึ้น ณ เวลาใดเวลาหนึ่งที่เป็นบวก นั่นคือ $u(t)$ โดยที่ $t \geq 0$ ดังนั้น จากทฤษฎีบทสังวัตนาการ (Convolution theorem) จะได้ว่า สัญญาณออกจากระบบ $y(t)$ ณ เวลา $t > 0$ จะสอดคล้องกับสมการ

$$y(t) = h(t) * u(t) := \int_{-\infty}^t h(\tau) u(t-\tau) d\tau \quad (2.48)$$

โดยที่ $h(t)$ เป็นผลตอบสนองต่อสัญญาณอิมพัลส์ของระบบ

ถ้า $u(t)$ ถูกจำกัดขนาด นั่นคือ $|u(t)| \leq M_1 < \infty$ ดังนั้น

$$|y(t)| = \left| \int_{-\infty}^t h(\tau) u(t-\tau) d\tau \right| \leq \int_{-\infty}^t |h(\tau)| |u(t-\tau)| d\tau \leq M_1 \int_{-\infty}^t |h(t-\tau)| d\tau \quad (2.49)$$

ซึ่งจะเห็นว่าสัญญาณออก $y(t)$ จะถูกจำกัดขอบเขตสำหรับทุกค่า t ใดๆ เมื่อ $\int_{-\infty}^{\infty} |h(t)| dt < \infty$

2.3.1.3 ความสามารถควบคุมได้และความสามารถสังเกตได้

ความสามารถควบคุมได้และความสามารถสังเกตได้เป็นคุณสมบัติที่สำคัญสำหรับระบบควบคุม และมีบทบาทอย่างมากในการออกแบบระบบป้อนกลับหรือระบบควบคุมประเภทอื่นๆ ซึ่งหลักการของความสามารถควบคุมได้นั้น หมายถึงความสามารถในการนำระบบให้เคลื่อนไปยังตำแหน่งต่าง ๆ ในปริภูมิสถานะได้โดยใช้สัญญาณควบคุมที่เหมาะสม สำหรับหลักการของความสามารถสังเกตได้คือ เป็นความสามารถในการวัดว่าตัวแปรสถานะภายในในระบบสามารถระบุได้หรือไม่จากข้อมูลที่ได้จากสัญญาณขาออก สำหรับนิยามในเชิงคณิตศาสตร์ มีดังนี้

ความสามารถควบคุมได้ (Controllability)

นิยาม 2.3.1 ระบบควบคุมถูกกล่าวว่า มีความสามารถควบคุมได้อย่างสมบูรณ์ (Completely Controllable) บนช่วงปิด $[t_0, t_f]$ ถ้าสามารถหาสัญญาณควบคุมที่ทำให้เวกเตอร์ของตัวแปรสถานะ $x(t)$ เคลื่อนที่จากสถานะเริ่มต้น $x(t_0)$ ใดๆ ในปริภูมิสถานะไปยังสถานะสุดท้าย $x(t_f) = 0$ ได้ ในเวลาที่จำกัด (finite time) เมื่อกำหนดเวลาเริ่มต้นและสถานะเริ่มต้น $x(t_0) = x_0$ ใดๆ มาให้

ทฤษฎีบทที่ 2.3.4 ระบบเชิงเส้นที่ไม่แปรตามเวลา

$$\begin{aligned}\dot{x} &= Ax + Bu & ; x(0) &= x_0 \\ y &= Cx + Du\end{aligned}\quad (2.50)$$

โดยที่ $x(t)$ คือตัวแปรสถานะมิติ n และ $u(t)$ คือสัญญาณควบคุมขาเข้ามิติ m เงื่อนไขจำเป็นและเพียงพอที่ทำให้ระบบมีความสามารถควบคุมได้อย่างสมบูรณ์ คือเมทริกซ์ความสามารถควบคุมได้ (Controllability matrix)

$$C = [B \ AB \ A^2B \ \dots \ A^{n-1}B] \quad (2.51)$$

ต้องมีค่าลำดับขั้นเต็ม (Full rank) เป็น n

ความสามารถสังเกตได้ (Observability)

นิยาม 2.3.2 ระบบควบคุมถูกกล่าวว่า มีความสามารถสังเกตได้อย่างสมบูรณ์ (Completely Observable) บนช่วงปิด $[t_0, t_f]$ ถ้าทุกๆ สถานะเริ่มต้นใดๆ $x(t_0) = x_0$ เวกเตอร์ของตัวแปรสถานะ $x(t)$ สามารถระบุได้จากสัญญาณขาออก $y(t)$ บนช่วงปิด $[t_0, t_f]$

ทฤษฎีบทที่ 2.3.5 ระบบเชิงเส้นที่ไม่แปรตามเวลา

$$\begin{aligned}\dot{x} &= Ax + Bu & ; x(0) &= x_0 \\ y &= Cx + Du\end{aligned}\quad (2.52)$$

โดยที่ $x(t)$ คือตัวแปรสถานะมิติ n และ $y(t)$ คือสัญญาณขาออกมิติ p เงื่อนไขจำเป็นและเพียงพอที่ทำให้ระบบสังเกตได้อย่างสมบูรณ์ คือ เมทริกซ์ความสามารถสังเกตได้ (Observability matrix)

$$O = [C^T \ A^T C^T \ (A^2)^T C^T \ \dots \ (A^{n-1})^T C^T]^T \quad (2.53)$$

ต้องมีค่าลำดับขั้นเต็ม (Full rank) เป็น n

ตัวอย่างที่ 2.3.1 ความสามารถควบคุมได้และความสามารถสังเกตได้

กำหนดให้ $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, C = [1 \ 1 \ 0]$

จาก $C = [B \ AB \ A^2B \ \dots \ A^{n-1}B]$

และ $O = [C^T \ A^T C^T \ (A^2)^T C^T \ \dots \ (A^{n-1})^T C^T]^T$

จะได้ $C = [B \ AB \ A^2B] = \begin{bmatrix} 1 & 2 & 10 \\ 0 & 3 & 9 \\ 1 & 2 & 7 \end{bmatrix}$ และ $O = \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 3 & 4 \\ 5 & 9 & 14 \end{bmatrix}$

เนื่องจาก ลำดับชั้นของ C และของ O มีค่าเท่ากับ 3 ดังนั้นระบบที่กำหนดมานี้มีความสามารถควบคุมและสังเกตได้อย่างสมบูรณ์

2.3.2 การวิเคราะห์เชิงฟังก์ชันถ่ายโอน (Transfer function approach)

การวิเคราะห์เชิงฟังก์ชันถ่ายโอนจะสามารถใช้วิเคราะห์กับระบบเชิงเส้นไม่แปรตามเวลาเท่านั้น โดยจะให้เงื่อนไขเริ่มต้นของระบบมีค่าเป็นศูนย์เสมอ การวิเคราะห์เชิงฟังก์ชันถ่ายโอนนี้ จะสามารถทำได้โดยการแปลงลาปลาซของสมการเชิงอนุพันธ์สามัญที่บรรยายพฤติกรรมของระบบพลวัต ซึ่งจะได้ว่า ฟังก์ชันถ่ายโอน (Transfer function) คืออัตราส่วนระหว่างสัญญาณออกต่อสัญญาณเข้าที่ทำการแปลงลาปลาซ โดยกำหนดให้เงื่อนไขเริ่มต้นเป็นศูนย์

ให้สมการเชิงอนุพันธ์สามัญที่บรรยายระบบพลวัตเขียนอยู่ในรูปทั่วไป คือ

$$y^{(n)} + b_{n-1}y^{(n-1)} + \dots + b_1\dot{y} + b_0y = K(u^{(m)} + a_{m-1}u^{(m-1)} + \dots + a_1\dot{u} + a_0u) \quad (2.54)$$

ดังนั้น ฟังก์ชันถ่ายโอนที่ได้คือ

$$\frac{Y(s)}{U(s)} = G(s) = \frac{K(s^m + a_{m-1}s^{m-1} + a_{m-2}s^{m-2} + \dots + a_1s + a_0)}{s^n + b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \dots + b_1s + b_0} \quad (2.55)$$

สมการคุณลักษณะของระบบก็คือสมการที่เป็นส่วนหรือ Denominator นั่นคือ

$$s^n + b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \dots + b_1s + b_0 = 0 \quad (2.56)$$

โดยที่สมการนี้สามารถบอกถึงเสถียรภาพของระบบได้

ตัวอย่างที่ 2.3.2

สมมติแบบจำลองทางคณิตศาสตร์ของระบบที่สนใจสามารถเขียนด้วยสมการเชิงอนุพันธ์อันดับที่สองดังนี้คือ

$$\frac{d^2y}{dt^2} + 3\frac{dy}{dt} + 2y = 2\frac{du}{dt} + u$$

โดยกำหนดว่า u คือสัญญาณเข้าและ y คือสัญญาณออก หากผลการแปลงลาปลาซของสมการนี้ โดยกำหนดว่าค่าเงื่อนไขเริ่มต้น (Initial condition) มีค่าเป็นศูนย์ จะได้ว่า

$$s^2Y(s) + 3sY(s) + 2Y(s) = 2sU(s) + U(s)$$

ดังนั้น ฟังก์ชันถ่ายโอนของระบบนี้คือ

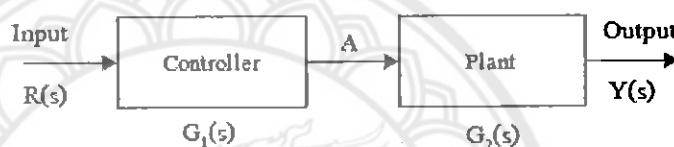
$$G(s) = \frac{Y(s)}{U(s)} = \frac{(2s+1)}{(s^2+3s+2)}$$

2.3.2.1 การเชื่อมต่อกันของระบบย่อย

ระบบทางกายภาพโดยทั่วไปมีขนาดใหญ่ ในการคำนวณจึงจำเป็นต้องแยกระบบออกเป็นองค์ประกอบย่อย ๆ แล้วคำนวณหาฟังก์ชันถ่ายโอนขององค์ประกอบย่อยนั้นเพื่อนำมาเขียนเป็นแผนภาพกล่อง ดังนั้น หนึ่งองค์ประกอบจะมีหนึ่งแผนภาพกล่อง เมื่อนำแผนภาพกล่องทุกกล่องมารวมกันเพื่อพิจารณาทั้งระบบ จึงจำเป็นต้องอาศัยหลักการลดรูปดังนี้

ระบบควบคุมที่ต่อแบบอนุกรม (Cascade Control System)

พิจารณาการต่อแผนภาพกล่องหรือระบบย่อยสองระบบแบบอนุกรม หรือบ่อยครั้งเรียกว่าการนำระบบมา Cascade กัน ตามรูปที่ 2.8



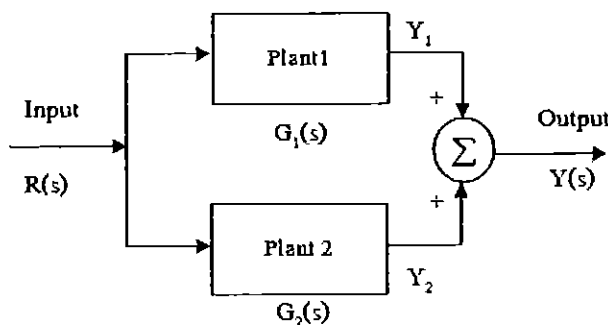
รูปที่ 2.8 ระบบควบคุมที่ต่อแบบอนุกรม

จากแผนภาพกล่องพบว่า $Y = G_2 A$ และ $A = G_1 R$ หรือนั่นคือ $Y = G_2 G_1 R$ ดังนั้นหากมีการต่อแผนภาพกล่องแบบอนุกรม ฟังก์ชันถ่ายโอนรวมเกิดจากการนำฟังก์ชันถ่ายโอนย่อยมาคูณกัน นั่นคือ

$$\frac{Y(s)}{R(s)} = G_1(s) \cdot G_2(s) \quad (2.57)$$

ระบบควบคุมที่ต่อแบบขนาน (Parallel Control System)

พิจารณาการต่อแผนภาพกล่องหรือระบบย่อยสองระบบแบบขนาน ดังรูปที่ 2.9 โดยที่สัญญาณเข้าของทั้งสองระบบคือ $R(s)$ และสัญญาณออกของระบบรวมคือสัญญาณรวมของสัญญาณออกของระบบย่อยทั้งสอง



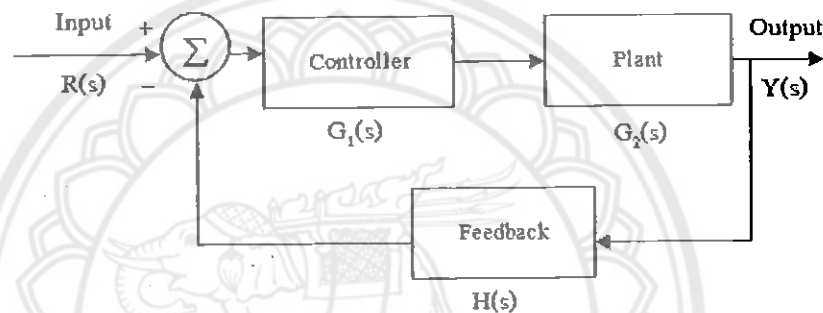
รูปที่ 2.9 ระบบควบคุมที่ต่อแบบอนุกรม

สมมติให้ฟังก์ชันถ่ายโอนของระบบควบคุมทั้งสองคือ $G_1(s)$ และ $G_2(s)$ ตามลำดับ ดังนั้น $Y(s) = Y_1(s) + Y_2(s) = (G_1(s) + G_2(s))R(s)$ จึงสรุปได้ว่าฟังก์ชันถ่ายโอนรวมของระบบที่ต่อแบบขนานมีค่าเท่ากับ

$$\frac{Y(s)}{R(s)} = G_1(s) + G_2(s) \quad (2.58)$$

ระบบควบคุมที่ต่อแบบป้อนกลับ (Feedback Control System)

รูปที่ 2.10 เป็นรูปที่ประกอบด้วยระบบย่อยมาต่อเข้าด้วยกันในลักษณะที่นำสัญญาณออกป้อนกลับมาเปรียบเทียบกับสัญญาณเข้าในลักษณะของระบบปิด



รูปที่ 2.10 ระบบควบคุมที่ต่อแบบป้อนกลับ

สมมติให้ฟังก์ชันถ่ายโอนของตัวควบคุมและระบบที่ต้องการควบคุมเป็นเท่ากับ $G_1(s)$ และ $-G_2(s)$ ตามลำดับ และฟังก์ชันถ่ายโอนของตัวป้อนกลับ (Feedback) เท่ากับ $H(s)$ ดังนั้นสามารถหาความสัมพันธ์ของสัญญาณเข้าและสัญญาณออกของระบบรวมที่เป็นฟังก์ชันถ่ายโอนของระบบได้คือ

$$\frac{Y(s)}{R(s)} = \frac{G_1(s) \cdot G_2(s)}{1 + G_1(s) \cdot G_2(s) \cdot H(s)} \quad (2.59)$$

2.3.2.2 ความมีเสถียรภาพ (Stability)

สิ่งที่สำคัญในการออกแบบระบบควบคุมคือเรื่องเสถียรภาพของระบบ สำหรับระบบควบคุมเชิงเส้นไม่แปรตามเวลาที่พารามิเตอร์ระบบด้วยฟังก์ชันถ่ายโอนนั้น เสถียรภาพของระบบขึ้นอยู่กับตำแหน่งของขั้ววงปิด (Closed loop poles) พิจารณาฟังก์ชันถ่ายโอนวงปิดระบบควบคุมที่อยู่ในรูปทั่วไป ดังนี้

$$G(s) = \frac{K(s^m + a_{m-1}s^{m-1} + a_{m-2}s^{m-2} + \dots + a_1s + a_0)}{s^n + b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \dots + b_1s + b_0} \quad (2.60)$$

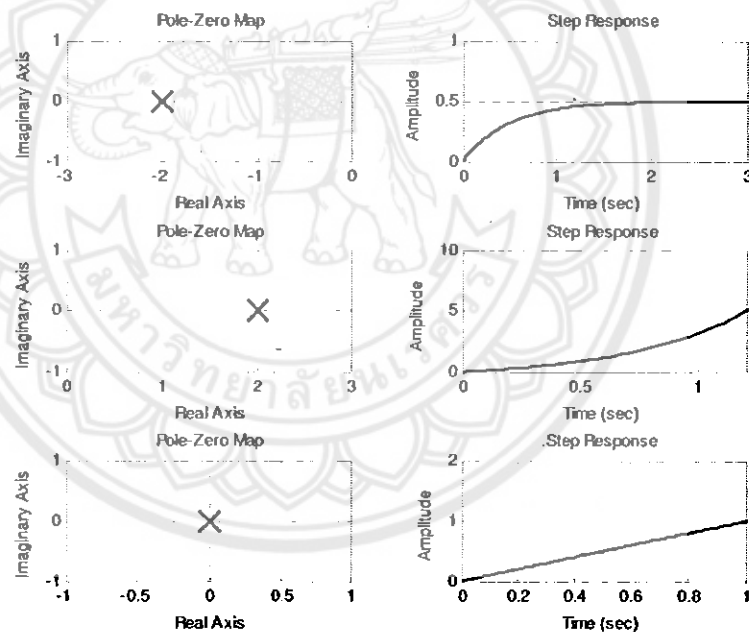
โดยอาศัยทฤษฎีพื้นฐานของพีชคณิต (Fundamental Theorem of Algebra) ฟังก์ชันถ่ายโอน $G(s)$ ใน (2.60) สามารถแยกองค์ประกอบได้เป็น

$$G(s) = \frac{K(s-z_1)(s-z_2)\dots(s-z_m)}{(s-p_1)(s-p_2)\dots(s-p_n)} \quad (2.61)$$

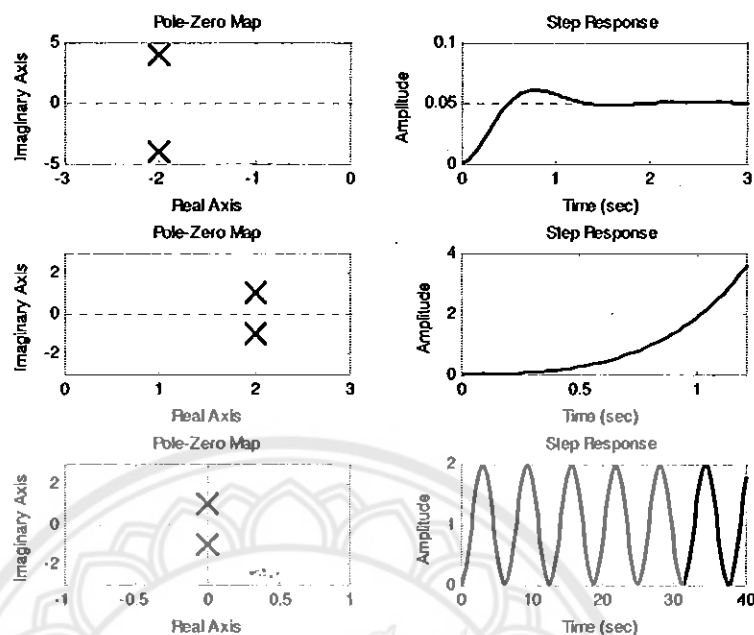
โดย z_i เรียกว่าเป็นศูนย์ (Zeros) ของฟังก์ชันถ่ายโอน ในขณะที่ p_j จะเรียกว่าเป็นขั้ว (Poles) ของฟังก์ชันถ่ายโอน $G(s)$ และ K ซึ่งเป็นค่าคงที่จะเรียกว่าเป็นอัตราขยายของระบบ

ทฤษฎีบทที่ 2.3.6 ระบบควบคุมที่มีฟังก์ชันถ่ายโอน $G(s)$ ดังสมการที่ (2.61) จะมีเสถียรภาพ ก็ต่อเมื่อ ส่วนจริงของขั้ววงปิดมีค่าน้อยกว่าศูนย์ (นั่นคือ $\text{Re } p_j < 0$)

รูปที่ 2.11 และรูปที่ 2.12 แสดงตำแหน่งขั้วและการตอบสนองของระบบอันดับหนึ่ง และระบบอันดับสอง ตามลำดับ โดยแสดงทั้งระบบที่เสถียรและไม่เสถียร



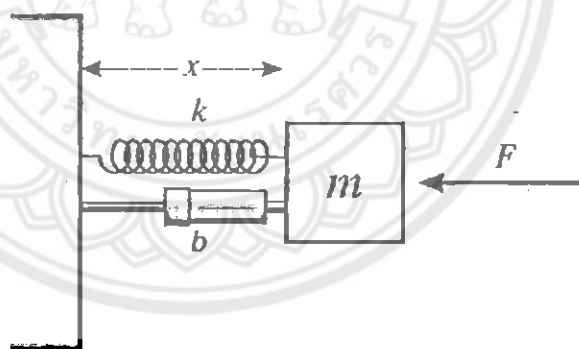
รูปที่ 2.11 ตำแหน่งขั้วและผลตอบสนองของระบบอันดับหนึ่ง



รูปที่ 2.12 ตำแหน่งขั้วและผลตอบสนองของระบบอันดับสอง

2.3.2.3 การวิเคราะห์ผลตอบสนองของระบบอันดับที่สอง (Second Order Systems)

พิจารณาระบบสปริง-มวล-ตัวหน่วง ดังแสดงในรูปที่ 2.13



รูปที่ 2.13 ระบบสปริง-มวล-ตัวหน่วง (Spring-mass-damper system)

หากกำหนดให้ระยะขจัด x เป็นสัญญาณออกที่ต้องการทราบ และ $F(t)$ คือแรงภายนอกที่กระทำต่อระบบ ผู้อ่านสามารถพิสูจน์ได้โดยไม่ต้องยากว่า ฟังก์ชันถ่ายโอนของระบบดังกล่าวนี้คือ

$$\frac{X(s)}{F(s)} = G(s) = \frac{1}{k} \frac{k/m}{s^2 + \frac{b}{m}s + \frac{k}{m}} \quad (2.62)$$

จากสมการที่ (2.62) สามารถปรับให้อยู่ในรูปแบบมาตรฐานที่ใช้พารามิเตอร์อัตราส่วนการหน่วง ζ (Damping ratio) และความถี่ธรรมชาติ ω_n เป็นตัวบ่งบอกถึงลักษณะของการตอบสนองของระบบอันดับที่สอง รูปแบบมาตรฐานสามารถเขียนได้ดังนี้

$$\frac{X(s)}{F(s)} = G(s) = \frac{1}{k} \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (2.63)$$

ในกรณีของระบบสปริง-มวล-ตัวหน่วงนี้ จะได้ว่า

$$2\zeta\omega_n = \frac{b}{m} \quad \text{และ} \quad \omega_n^2 = \frac{k}{m} \quad (2.64)$$

ตัวแปร $x(t)$ สามารถคำนวณหาได้โดยวิธีการหาผลการแปลงลาปลาซผกผัน หากสมมติให้สัญญาณเข้า $F(t)$ เป็นสัญญาณขั้นบันไดหนึ่งหน่วย (Unit step function) จะได้ว่า $F(s) = \frac{1}{s}$ ดังนั้นค่า $x(t)$ สามารถหาได้ดังนี้

$$X(s) = G(s) = \frac{1}{k} \frac{\omega_n^2}{(s^2 + 2\zeta\omega_n s + \omega_n^2)} \times \frac{1}{s} \quad (2.65)$$

อย่างไรก็ตาม ผลตอบสนองของระบบขึ้นกับค่าของ ζ ดังนี้

1. ผลตอบสนองแบบหน่วงต่ำกว่าวิกฤต: Underdamped ($0 < \zeta < 1$)

ในกรณีนี้ สมการ (2.65) สามารถเขียนได้ดังนี้

$$X(s) = \frac{1}{k} \left(\frac{1}{s} - \frac{s + 2\zeta\omega_n}{(s^2 + 2\zeta\omega_n s + \omega_n^2)} \right) \quad (2.66)$$

$$= \frac{1}{k} \left(\frac{1}{s} - \frac{s + \zeta\omega_n}{(s + \zeta\omega_n)^2 + \omega_d^2} - \frac{\zeta\omega_n}{(s + \zeta\omega_n)^2 + \omega_d^2} \right) \quad (2.67)$$

โดยที่ $\omega_d = \omega_n \sqrt{1 - \zeta^2}$ คือ ความถี่ธรรมชาติถูกหน่วง (Damped natural frequency) หลังจากใช้พีชคณิตอย่างง่าย ร่วมกับการแปลงลาปลาซผกผัน จะได้ว่า ผลตอบสนอง $x(t)$ ในกรณีคือ

$$x(t) = \frac{1}{k} \left\{ 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1 - \zeta^2}} \sin \left(\omega_d t + \tan^{-1} \frac{\sqrt{1 - \zeta^2}}{\zeta} \right) \right\}, \quad (t \geq 0) \quad (2.68)$$

2. ผลตอบสนองแบบหน่วงวิกฤต: Critically damp ($\zeta = 1$)

ในกรณีนี้ สมการ (2.65) สามารถเขียนได้เป็น

$$X(s) = \frac{1}{k} \frac{\omega_n^2}{(s + \omega_n)^2} \times \frac{1}{s} = \frac{1}{k} \left[\frac{1}{s} - \frac{\omega_n}{(s + \omega_n)^2} - \frac{1}{(s + \omega_n)} \right] \quad (2.69)$$

จากการหาผลการแปลงกลับลาปลาซ จะได้ว่า

$$x(t) = 1 - e^{-\omega_n t} (1 + \omega_n t), \quad (t \geq 0) \quad (2.70)$$

3. ผลตอบสนองแบบหน่วงสูงกว่าวิกฤต: Overdamped ($\zeta > 1$)

ในกรณีสุดท้ายนี้ สมการ (2.65) สามารถเขียนได้เป็น

$$X(s) = \frac{1}{k} \frac{\omega_n^2}{(s + \zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1})(s + \zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1})} \times \frac{1}{s} \quad (2.71)$$

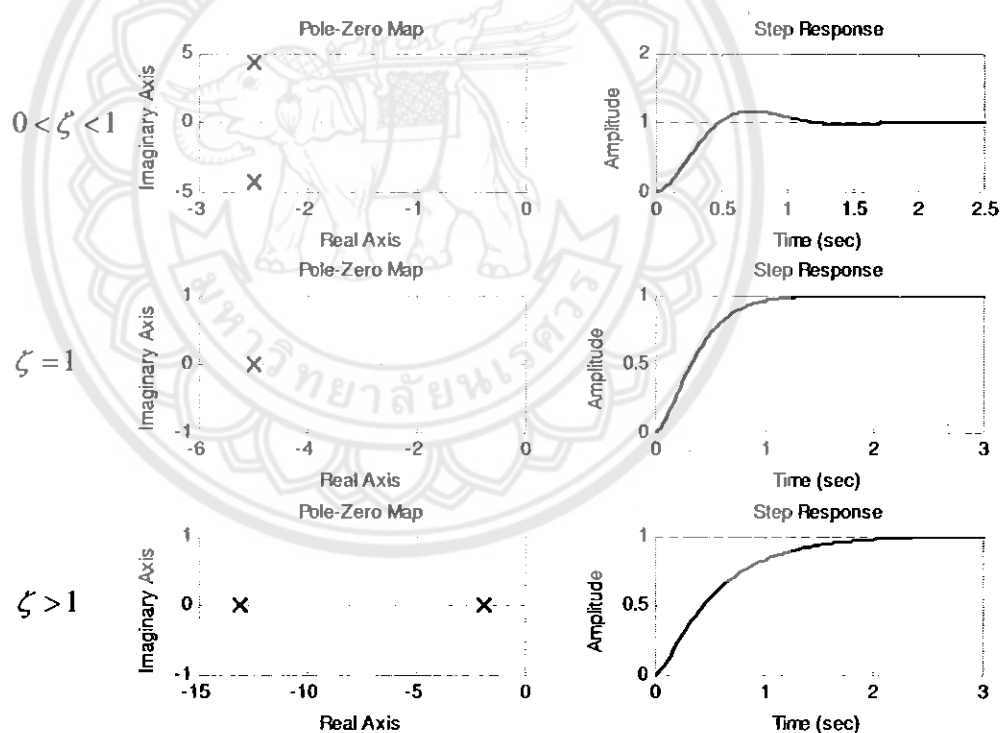
กำหนดให้ $s_1 = \zeta + \sqrt{\zeta^2 - 1}$ และ $s_2 = \zeta - \sqrt{\zeta^2 - 1}$ ดังนั้น สมการข้างต้นเขียนได้เป็น

$$X(s) = \frac{1}{k} \frac{\omega_n^2}{(s + s_1\omega_n)(s + s_2\omega_n)} \times \frac{1}{s} \quad (2.72)$$

จากการหาผลการแปลงกลับลาปลาซ จะได้ว่า

$$x(t) = \frac{1}{k} \left[1 + \frac{1}{2\sqrt{\zeta^2 - 1}} \left(\frac{e^{-s_1\omega_n t}}{s_1} - \frac{e^{-s_2\omega_n t}}{s_2} \right) \right], \quad (t \geq 0) \quad (2.73)$$

ผลตอบสนองของระบบอันดับสองต่อสัญญาณขั้นบันไดขนาดหนึ่งหน่วยแสดงได้ ดังรูปนี้

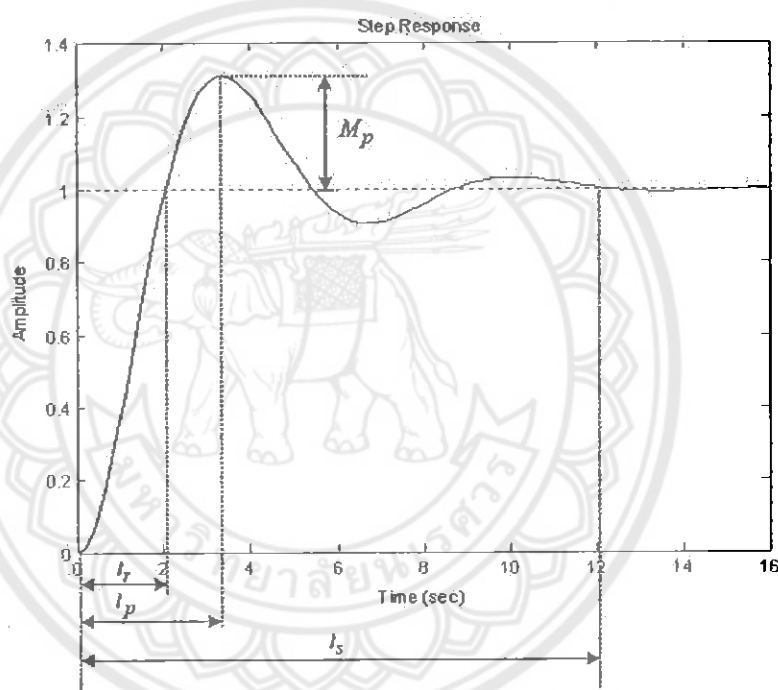


รูปที่ 2.14 ผลตอบสนองของระบบอันดับสองที่ค่าอัตราหน่วง (ζ) ต่างๆ

2.3.2.4 คุณสมบัติของผลตอบสนองชั่วขณะของระบบอันดับที่สอง

เนื่องจากผลตอบสนองของระบบที่เป็นแบบหน่วงต่ำกว่าวิกฤตหรือที่มีค่า $0 < \zeta < 1$ มีผลตอบสนองที่เร็วเนื่องจากระบบมีการแกว่งไกว (Oscillation) ซึ่งช่วยให้ระบบเข้าสู่สภาวะคงตัวได้เร็วขึ้น ดังนั้นในการออกแบบระบบควบคุมจึงให้ความสนใจเกี่ยวกับผลการสนองแบบหน่วงต่ำกว่าวิกฤตมากกว่ากรณีอื่นๆ ซึ่งจากในหัวข้อที่ผ่านมาพบว่าผลตอบสนองต่อสัญญาณเข้าแบบขั้นบันไดขนาดหนึ่งหน่วยในกรณีหน่วงต่ำกว่าวิกฤต ในกรณี $k = 1$ คือ

$$x(t) = 1 - \frac{e^{-\zeta\omega_d t}}{\sqrt{1-\zeta^2}} \sin\left(\omega_d t + \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta}\right), \quad (t \geq 0) \quad (2.74)$$



รูปที่ 2.15 ลักษณะการตอบสนองชั่วขณะของระบบอันดับที่สอง $0 < \zeta < 1$

1. ช่วงเวลาขึ้น (Rise Time, t_r)

เป็นช่วงเวลาที่ผลตอบสนองเมื่อสัญญาณออกเพิ่มขึ้นจาก 10% จนถึง 90% หรือจาก 0% ถึง 100% ซึ่งเป็นช่วงเวลาที่นิยมมากที่สุด ในการคำนวณเวลานี้จะใช้เวลาที่ผลตอบสนองของระบบจาก 0%-100% ของค่าผลตอบสนองในสภาวะคงตัว (Steady state response) โดยที่เวลาขึ้นสามารถประมาณได้ว่า

$$t_r = \frac{\pi - \theta}{\omega_d} = \frac{\pi - \theta}{\omega_n \sqrt{1-\zeta^2}} \quad (2.75)$$

โดยที่ $\theta = \cos^{-1} \zeta$

2. เวลาของค่ายอด (Peak Time, t_p)

เป็นเวลาที่ผลตอบสนองมีค่าสูงสุดค่าแรก ซึ่งเป็นการคำนวณหาค่า t ที่ทำให้อนุพันธ์ของผลตอบสนองตามสมการที่ (2.74) มีค่าเป็นศูนย์ นั่นคือ $\frac{dx(t)}{dt} = 0$ ซึ่งพบว่า

$$t_p = \frac{\pi}{\omega_d} = \frac{\pi}{\omega_n \sqrt{1-\zeta^2}} \quad (2.76)$$

3. โอเวอร์ชูตสูงสุด (Maximum Overshoot, M_p)

เป็นอัตราส่วนของผลต่างระหว่างค่าที่เวลาของค่ายอดกับค่าที่สภาวะคงตัวต่อค่าที่สภาวะคงตัว การบอกค่าโอเวอร์ชูตสูงสุดส่วนมากจะบอกเป็นร้อยละ ซึ่งมีค่าเป็น

$$\%M_p = e^{\frac{-\zeta\pi}{\sqrt{1-\zeta^2}}} \times 100\% \quad (2.77)$$

4. เวลาเข้าสู่สภาวะคงตัว (Settling Time, t_s)

เป็นเวลาที่ผลตอบสนองเข้าใกล้ค่าสภาวะคงตัวภายในช่วงที่กำหนด โดยที่ช่วงที่กำหนดนี้มีการวัดเป็นค่าร้อยละ เช่น หากที่สภาวะคงตัวผลตอบสนองมีค่าเป็น 1 ดังนั้น หากผลตอบสนองเข้าในย่าน 0.99–1.01 จะกล่าวว่า ผลตอบสนองเข้าสู่สภาวะคงตัวในช่วง 1% สำหรับเวลาที่ผลตอบสนองเข้าใกล้ค่าที่สภาวะคงตัวสามารถกำหนดได้ดังนี้

$$\text{สำหรับกรณี 1\%} \quad t_s = \frac{4.6}{\zeta\omega_n} \quad (2.78)$$

$$\text{สำหรับกรณี 2\%} \quad t_s = \frac{4}{\zeta\omega_n} \quad (2.79)$$

$$\text{สำหรับกรณี 5\%} \quad t_s = \frac{3}{\zeta\omega_n} \quad (2.80)$$

2.3.3 ความสัมพันธ์ระหว่างปริภูมิสถานะกับฟังก์ชันถ่ายโอน (Transfer function)

ต่อไปจะเป็นการศึกษาความสัมพันธ์ระหว่างระบบที่เขียนพรรณนาในรูปของปริภูมิสถานะกับระบบที่บรรยายด้วยฟังก์ชันถ่ายโอน ดังนี้

$$\text{จากสมการสถานะ} \quad \dot{x} = Ax + Bu \quad (2.81)$$

$$\text{และสมการขาออก} \quad y = Cx + Du \quad (2.82)$$

การหาผลการแปลงลาปลาซของสมการที่ (2.81) จะได้ว่า

$$sX(s) - x(0) = AX(s) + BU(s) \quad (2.83)$$

จัดรูปใหม่ได้ว่า

$$X(s) = (sI - A)^{-1}BU(s) + (sI - A)^{-1}x(0) \quad (2.84)$$

และจากสมการขาออก จะได้ดังนี้

$$Y(s) = CX(s) + DU(s) \quad (2.85)$$

ดังนั้น

$$Y(s) = C \{ (sI - A)^{-1} BU(s) + (sI - A)^{-1} x(0) \} + DU(s) \quad (2.86)$$

$$= \{ C(sI - A)^{-1} B + D \} U(s) + C(sI - A)^{-1} x(0) \quad (2.87)$$

ให้ $x(0) = 0$ จะได้ฟังก์ชันถ่ายโอนคือ

$$G(s) = \frac{Y(s)}{U(s)} = C(sI - A)^{-1} B + D \quad (2.88)$$

ตัวอย่างที่ 2.3.3 การแปลงระบบปริภูมิสถานะเป็นฟังก์ชันถ่ายโอน

กำหนดให้ $A = \begin{bmatrix} -5 & -6 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $C = [0 \ 1]$

สามารถหาสมการของฟังก์ชันถ่ายโอนได้คือ

$$G(s) = \frac{Y(s)}{U(s)} = C(sI - A)^{-1} B = \frac{1}{(s+2)(s+3)}$$

บทที่ 3

หลักการและทฤษฎีของอสมการเมทริกซ์เชิงเส้น

ในการศึกษาเกี่ยวกับการแก้ปัญหาของระบบควบคุมด้วยวิธีการอสมการเมทริกซ์เชิงเส้น (Linear Matrix Inequality: LMI) นั้นจะต้องมีความรู้ความเข้าใจเกี่ยวกับหลักการและทฤษฎีของอสมการเมทริกซ์เชิงเส้นก่อน เพื่อที่จะสามารถนำไปประยุกต์ใช้ในการแก้ปัญหาของระบบควบคุมได้อย่างถูกต้องและมีประสิทธิภาพ ดังนั้นในบทนี้เน้นศึกษาเกี่ยวกับทฤษฎีพื้นฐานต่างๆ ของอสมการเมทริกซ์เชิงเส้น สำหรับผู้ที่สนใจ สามารถศึกษาเพิ่มเติมได้ที่ [2, 5, 9]

3.1 นิยามของอสมการเมทริกซ์เชิงเส้น

อสมการเมทริกซ์เชิงเส้นหรือ LMI คือเมทริกซ์ที่อยู่ในรูป

$$F(x) = F_0 + \sum_{i=1}^m x_i F_i > 0 \quad (3.1)$$

โดยที่

x เป็นเวกเตอร์ตัวแปรค่าจริง ($x \in \mathbb{R}^m$)

F_i เป็นเมทริกซ์สมมาตรที่ทราบค่า; ($F_i = F_i^T \in \mathbb{R}^{n \times n}$)

$F(x) > 0$ มีความหมายว่าเป็นบวกแน่นอน (Positive definite)

ตัวอย่างที่ 3.1.1 อสมการเมทริกซ์เชิงเส้น

$$1 + x < 0$$

$$1 + x_1 + 2x_2 < 0$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + x_1 \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} + x_2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} < 0$$

3.2 คอนเวกซ์ของอสมการเมทริกซ์เชิงเส้น

คอนเวกซ์เป็นคุณสมบัติที่สำคัญมากของอสมการเมทริกซ์เชิงเส้น ซึ่งพบว่าเซตของ $\{x \mid F(x) > 0\}$ มีความเป็นคอนเวกซ์ นั่นคือ

$$F(\lambda x + (1-\lambda)y) = \lambda F(x) + (1-\lambda)F(y) > 0 \quad (3.2)$$

ซึ่งสามารถพิสูจน์ได้ดังนี้

กำหนดให้ x และ y เป็นเวกเตอร์ที่ทำให้ $F(x) > 0$ และ $F(y) > 0$ และให้ $\lambda \in (0,1)$ แล้วจะได้ว่า

$$\begin{aligned} F(\lambda x + (1-\lambda)y) &= F_0 + \sum_{i=1}^m (\lambda x_i + (1-\lambda)y_i)F_i \\ &= F_0 + \sum_{i=1}^m \lambda x_i F_i + \sum_{i=1}^m (1-\lambda)y_i F_i \\ &= F_0(1-\lambda + \lambda) + \lambda \sum_{i=1}^m x_i F_i + (1-\lambda) \sum_{i=1}^m y_i F_i \\ &= \lambda F_0 + \lambda \sum_{i=1}^m x_i F_i + (1-\lambda)F_0 + (1-\lambda) \sum_{i=1}^m y_i F_i \\ &= \lambda F(x) + (1-\lambda)F(y) > 0 \end{aligned}$$

3.3 อสมการเมทริกซ์เชิงเส้นหลายๆ สมการ

ในระบบควบคุมโดยทั่วไป พบว่ามีเงื่อนไขที่เขียนอยู่ในรูปของอสมการเมทริกซ์เชิงเส้นหลายๆ อสมการ แต่ด้วยคุณสมบัติของอสมการเมทริกซ์เชิงเส้น พบว่าเงื่อนไขของอสมการหลายๆ อสมการ สามารถลดรูปให้อยู่ในรูปของอสมการเพียงอสมการเดียวได้

กำหนดให้มีอสมการเมทริกซ์เชิงเส้นอยู่ k อสมการ คือ

$$F^1(x) > 0, F^2(x) > 0, \dots, F^k(x) > 0 \quad (3.3)$$

ถ้าอสมการที่ (3.3) นี้มีความเป็นคอนเวกซ์แล้วสามารถเขียนใหม่ได้ดังนี้

$$F(x) = F_0 + \sum_{i=1}^m x_i F_i = \begin{bmatrix} F^1(x) & 0 & \dots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & F^k(x) \end{bmatrix} > 0 \quad (3.4)$$

ตัวอย่างที่ 3.3.1 การจัดรูปสมการของลียาปูนอฟ (Lyapunov equation) ให้อยู่ในรูปของ LMI

พิจารณาสมการของลียาปูนอฟ $A^T P + PA + Q = 0$ เนื่องจากระบบมีเสถียรภาพภายใน หากสามารถหาเมทริกซ์สมมาตร $P \geq 0$ ได้สำหรับทุกค่า $Q > 0$ ซึ่งสามารถเขียนให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นได้ดังนี้

$$\begin{bmatrix} -P & 0 \\ 0 & A^T P + PA + Q \end{bmatrix} \leq 0 \quad (3.5)$$

3.4 ชั่วร์คอมพลิเมนต์ (Schur complement)

ชั่วร์คอมพลิเมนต์สามารถเปลี่ยนรูปอสมการไม่เชิงเส้นแบบคอนเวกซ์ (Convex nonlinear inequality) ที่พบในปัญหาของระบบควบคุมให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นได้ ซึ่งอสมการไม่เชิงเส้นแบบคอนเวกซ์ จัดอยู่ในรูปดังนี้

$$R(x) > 0, \quad Q(x) - S(x)R(x)^{-1}S(x)^T > 0 \quad (3.6)$$

โดยที่ $Q(x) = Q^T(x)$, $R(x) = R^T(x)$

จากอสมการไม่เชิงเส้นแบบคอนเวกซ์ในอสมการที่ (3.6) สามารถเปลี่ยนให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นโดยใช้ชั่วร์คอมพลิเมนต์ ได้ดังนี้

$$\begin{bmatrix} Q(x) & S(x) \\ S(x)^T & R(x) \end{bmatrix} > 0 \quad (3.7)$$

ทฤษฎีบท 3.4.1 เมทริกซ์สมมาตร $M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$ มีความเป็นบวกแน่นอน (Positive definite)

ก็ต่อเมื่อ $M_{11} > 0$ และ $M_{22} - M_{21}M_{11}^{-1}M_{12} > 0$

หรือ $M_{22} > 0$ และ $M_{11} - M_{12}M_{22}^{-1}M_{21} > 0$

ซึ่งจะได้ว่า

$$S' := M_{22} - M_{21}M_{11}^{-1}M_{12} \quad (3.8)$$

และ

$$S'' := M_{11} - M_{12}M_{22}^{-1}M_{21} \quad (3.9)$$

ซึ่งเรียกสมการที่ (3.8) และ (3.9) ว่า ชั่วร์คอมพลิเมนต์ (Schur complement)

สำหรับในกรณีที่เป็นระบบแบบไม่เชิงเส้นสามารถทำให้เป็นระบบแบบเชิงเส้นที่อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นได้ ผู้สนใจ สามารถศึกษาได้ที่ [10]

3.5 การแปลงรูปสมการให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้น

ในปัญหาของระบบควบคุมส่วนมากตัวแปรต่างๆ จะอยู่ในรูปของเมทริกซ์ ดังนั้นในวิธีการแก้สมการก่อนข้างที่จะยุ่งยากและบางสมการอาจจะอยู่ในรูปของสมการที่ไม่เป็นเชิงเส้น (Nonlinear) ดังนั้นในหัวข้อนี้จะเป็นการแปลงสมการในรูปทั่วไปให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้น ซึ่งถ้าสมการใดสามารถจัดให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นได้แล้ว คำสุดท้ายที่ได้จัดว่าเป็นคำสุดท้ายแบบมหากอย่างแน่นอน

3.5.1 สมการของลียาปูนอฟ (Lyapunov equation)

สมการของลียาปูนอฟคือ

$$A^T P + PA + Q = 0 \quad (3.10)$$

โดยที่ A เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $n \times n$

Q เป็นเมทริกซ์สมมาตรค่าจริงกึ่งบวกที่ทราบค่า มีมิติ $n \times n$

P เป็นเมทริกซ์สมมาตรที่เป็นบวกแน่นอนที่ไม่ทราบค่า $P = P^T \in \mathbb{R}^{n \times n}$

ซึ่งสามารถเขียนให้อยู่ในรูปของ LMI ดังนี้

$$\begin{bmatrix} -P & 0 \\ 0 & A^T P + PA + Q \end{bmatrix} \leq 0 \quad (3.11)$$

3.5.2 อสมการรีคคาติเชิงพีชคณิต (Algebraic Riccati Equation: ARE)

อสมการของรีคคาติเชิงพีชคณิตคือ

$$A^T P + PA - PBR^{-1}B^T P + Q \geq 0, \quad R > 0 \quad (3.12)$$

โดยที่ A เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $n \times n$

B เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $n \times m$

Q เป็นเมทริกซ์สมมาตรค่าจริงกึ่งบวกที่ทราบค่า มีมิติ $n \times n$

R เป็นเมทริกซ์สมมาตรค่าจริงที่เป็นบวกแน่นอน มีมิติ $m \times m$

P เป็นเมทริกซ์สมมาตรที่เป็นบวกแน่นอนที่ไม่ทราบค่า $P = P^T \in \mathbb{R}^{n \times n}$

เมื่อแปลงอสมการของรีคคาติให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้น โดยใช้ตัวคอมพลิเมนต์ จะได้

$$\begin{bmatrix} A^T P + PA + Q & PB \\ B^T P & R \end{bmatrix} \geq 0 \quad (3.13)$$

3.5.3 อสมการของค่าขอบเขตจริง (Bounded real) ที่ใช้ในระบบควบคุมแบบ H_∞

อสมการของค่าขอบเขตจริง (Bounded real) ที่ใช้ในระบบควบคุมแบบ H_∞ คือ

$$(A^T P + PA + C^T C) + (PB + C^T D)(\gamma^2 I - D^T D)^{-1}(B^T P + D^T C) < 0 \quad (3.14)$$

เมื่อ $\gamma^2 I - D^T D > 0$

โดยที่ A เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $n \times n$

B เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $n \times m$

C เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $p \times n$

D เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $p \times m$

P เป็นเมทริกซ์สมมาตรที่เป็นบวกแน่นอนที่ไม่ทราบค่า $P = P^T \in \mathbb{R}^{n \times n}$

เมื่อแปลงอสมการที่ (3.14) ให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นโดยใช้ตัวประกอบลิเม็นต์ จะได้

$$\begin{bmatrix} -(A^T P + PA + C^T C) & -(PB + C^T D) \\ -(B^T P + D^T C) & \gamma^2 I - D^T D \end{bmatrix} > 0 \quad (3.15)$$

หรือมีค่าเทียบเท่ากับ

$$\begin{bmatrix} A^T P + PA + C^T C & PB + C^T D \\ B^T P + D^T C & D^T D - \gamma^2 I \end{bmatrix} < 0 \quad (3.16)$$

3.5.4 อสมการของความเป็นบวกจริง (Positive real)

อสมการของความเป็นบวกจริง (Positive real) คือ

$$A^T P + PA + (PB - C^T)(D + D^T)^{-1}(B^T P - C) \leq 0 \quad (3.17)$$

เมื่อ $(D + D^T) > 0$

โดยที่ A เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $n \times n$

B เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $n \times m$

C เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $p \times n$

D เป็นเมทริกซ์ค่าจริงที่ทราบค่า มีมิติ $p \times m$

P เป็นเมทริกซ์สมมาตรที่เป็นบวกแน่นอนที่ไม่ทราบค่า $P = P^T \in \mathbb{R}^{n \times n}$

เมื่อแปลงอสมการที่ (3.17) ให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นโดยใช้ตัวประกอบลิเม็นต์ จะได้

$$\begin{bmatrix} A^T P + PA & PB + C^T \\ B^T P - C & -D^T - D \end{bmatrix} \leq 0 \quad (3.18)$$

3.5.5 สมการของพหุนามที่เป็นบวก (Positive polynomials)

จากสมการพหุนามที่มีกำลังสูงสุดเป็นเลขคู่ (Even polynomials) ซึ่งสามารถเขียนได้เป็น

$$p(s) = p_0 + p_1 s + \dots + p_{2n} s^{2n} \quad (3.19)$$

จะได้ว่า $p(s) \geq 0$ ทุกค่าจริง s ก็ต่อเมื่อ

$$p_k = \sum_{i+j=k} X_{ij}, \quad k = 0, 1, \dots, 2n \quad (3.20)$$

สำหรับเมทริกซ์บางตัว ซึ่ง $X = X^T \geq 0$ ซึ่งปัญหาในเรื่องนี้สามารถศึกษาและพิสูจน์ได้จากเรื่องของ Sum-of-squares ผู้สนใจ สามารถศึกษาได้ที่ [6]

ตัวอย่างที่ 3.5.1 การหาค่าต่ำสุด (Global minimization) ของฟังก์ชันพหุนามที่เป็นบวกโดยจัดให้อยู่ในรูปของสมการเมทริกซ์เชิงเส้น

$$\text{สมการที่พิจารณาคือ } p(s) = 48 - 92s + 56s^2 - 13s^3 + s^4$$

ซึ่งสามารถแปลงสมการดังกล่าวนี้ให้อยู่ในรูปของสมการเมทริกซ์เชิงเส้นได้ดังนี้

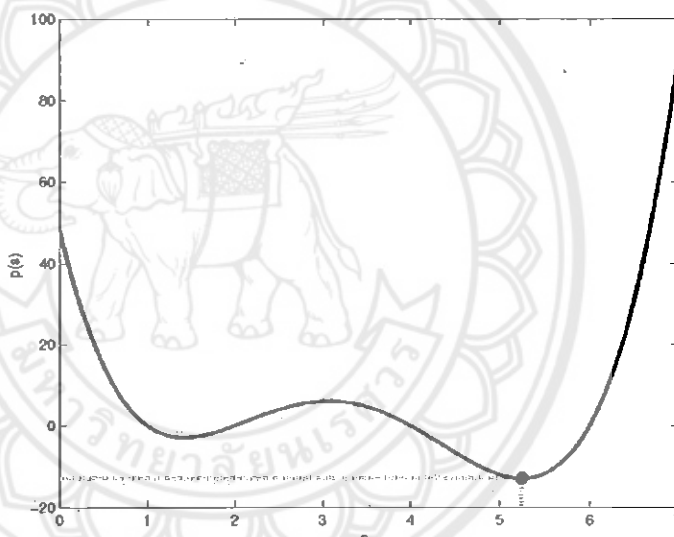
$$\min 48 - 92y_1 + 56y_2 - 13y_3 + y_4$$

$$\text{s.t. } \begin{bmatrix} 1 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \geq 0$$

เมื่อแก้สมการนี้ด้วยตัวแก้สมการเมทริกซ์เชิงเส้นแล้ว ได้ค่าดังนี้

$$S_{\text{Global min}} = 5.25$$

ซึ่งทำให้ได้ค่าต่ำสุดของฟังก์ชันพหุนามมีค่าเท่ากับ -12.89 ดังแสดงในรูปที่ 3.1



รูปที่ 3.1 ค่าต่ำสุดของฟังก์ชันพหุนาม

3.6 การใช้ตัวแก้สมการเมทริกซ์เชิงเส้น (LMI Solver) ในโปรแกรมแมทแลป

ในการแก้ปัญหาของสมการเมทริกซ์เชิงเส้นนั้นจะใช้ตัวแก้สมการเมทริกซ์เชิงเส้นหรือที่เรียกว่า LMI Solver โดยจะใช้งานร่วมกับโปรแกรมแมทแลป ซึ่งในปัจจุบันมีตัวแก้สมการอยู่หลายตัว ในโครงการนี้จะใช้ตัวแก้สมการที่เป็นเครื่องมือซอฟต์แวร์ (Software tool) ที่มีชื่อว่า YALMIP (Yet Another LMI Parser) โดยที่ซอฟต์แวร์ตัวนี้มีข้อดีตรงที่สามารถทำงานร่วมกับตัวแก้สมการได้หลายๆ ตัว อาทิเช่น LMILAB, SeDuMi (Self-Dual-Minimization), CSDP, OOQP, DSDP, GPLK เป็นต้น จึงสามารถใช้แก้ปัญหาในเรื่องของสมการเมทริกซ์เชิงเส้นนี้ได้เป็นอย่างดี และในการใช้งาน YALMIP ก็มีความสะดวกและง่ายต่อการใช้งาน ดังนั้นในโครงการนี้จะเลือกใช้

YALMIP มาเป็นซอฟต์แวร์เพื่อช่วยในการแก้ปัญหาของสมการเมทริกซ์เชิงเส้นและในโครงการนี้จะเลือกตัวแก้สมการเมทริกซ์เชิงเส้นสองตัวที่จะทำงานร่วมกับซอฟต์แวร์ YALMIP คือ LMILAB และ SeDuMi เพื่อที่จะสามารถนำคำตอบที่ได้มาเปรียบเทียบกับว่ามีความถูกต้องมากน้อยเพียงใด ซึ่งรายละเอียดของซอฟต์แวร์ YALMIP สามารถดูได้จากคู่มือของซอฟต์แวร์และสามารถศึกษาข้อมูลเพิ่มเติมเกี่ยวกับซอฟต์แวร์ได้ที่ <http://control.ee.ethz.ch/~joloef/yalmip.php>

3.6.1 รูปแบบคำสั่งที่สำคัญ

3.6.1.1 SDPVAR

คำอธิบาย: เป็นการกำหนดตัวแปรเมทริกซ์ที่ไม่ทราบค่า

รูปแบบคำสั่ง: $x = \text{sdpvar}(n,m,'field','type')$

n : จำนวนแถวของเมทริกซ์

m : จำนวนคอลัมน์ของเมทริกซ์

$field$: ชนิดของสมาชิกของเมทริกซ์ (real, complex)

$type$: ชนิดของเมทริกซ์ (symmetric, full, hermitian, toeplitz, hankel, skew)

ตัวอย่าง:

$P = \text{sdpvar}(2,2,'symmetric')$

3.6.1.2 DOUBLE

คำอธิบาย: เป็นการคืนค่าให้กับตัวแปรเมทริกซ์หลังจากแก้สมการเสร็จ

รูปแบบคำสั่ง: $Y = \text{double}(X)$

X : ตัวแปรแบบ sdp หรือ เซตของออปเจ็ค

ตัวอย่าง:

$\text{solvesdp}(F,obj);$

$\text{optobj} = \text{double}(obj);$

3.6.1.3 SET

คำอธิบาย: เป็นการกำหนดรูปแบบของสมการที่ต้องการจะแก้ปัญหา

รูปแบบคำสั่ง: $F = \text{set}(X,'lag')$

F : เซตของออปเจ็ค

X : ตัวแปรแบบ sdp

lag : ตัวอักษร

ตัวอย่าง:

$$F = \text{set}(P > 0) + \text{set}(\text{trace}(P) < 1);$$

$$G = \text{set}(P \geq 0) + \text{set}(\text{trace}(P) \leq 1);$$

3.6.1.4 CHECKSET

คำอธิบาย: เป็นคำสั่งที่ใช้ตรวจสอบรูปแบบของสมการที่ได้กำหนดไว้

รูปแบบคำสั่ง: `checkset(F)`

F: เซตของออปเจก

ตัวอย่าง:

```
solvesdp(F,objective);
```

```
checkset(F)
```

3.6.1.5 SDPSETTINGS

คำอธิบาย: เป็นการกำหนดค่าพารามิเตอร์ต่างๆ ให้กับตัวแก้สมการ

รูปแบบคำสั่ง: `options = sdpsettings('field','value','field','value',...)`

field: ตัวเลือกที่ต้องการแก้ไข

value: ค่าที่ต้องการกำหนด

ตัวอย่าง:

```
ops = sdpsettings('solver','linlab','verbose',0)
```

3.6.1.6 SOLVESDP

คำอธิบาย: เป็นคำสั่งที่ใช้ในการแก้สมการของปัญหาค่าเหมาะสมที่สุด

รูปแบบคำสั่ง: `solution = solvesdp(F,h,ops)`

solution: คำตอบที่ได้

F: เซตของออปเจกที่ต้องการหาคำตอบ

h: ตัวแปรแบบ sdp ซึ่งต้องเป็นปริมาณสเกลาร์เท่านั้น

ops: ตัวเลือกเพิ่มเติม

ตัวอย่าง:

```
x = sdpvar(length(c),1);
```

```
F = set(A*x < b)
```

```
solvesdp(F,c'*x);
```

3.6.2 ตัวอย่างโปรแกรมที่ใช้ตัวแก้สมการเมทริกซ์เชิงเส้น

```

clear all
A=[-1 2; 0 -2]
n=length(A);
Q=eye(n)
O=zeros(n);
%Check eigenvalue of matrix A%
Eigen_A=eig(A)
%Create unknown symmetric matrix%
P = sdpvar(length(A));
%Define Lyapunov inequalities%
F=set([-P O ; O A'*P+P*A+Q]<0);
% Set solver and its options%
opts=sdpsettings;
opts.solver='sedumi,lmilab';
%Find feasible solution, minimize trace(P)%
solution=solvesdp(F,trace(P),opts);
%Check solution%
P=double(P)
Eigen_P=eig(P);
checkset(F);

```

บทที่ 4

การออกแบบระบบเชิงเส้นด้วยวิธีอสมการเมทริกซ์เชิงเส้น

ปัญหาในทางระบบควบคุมหลายๆ เป็นปัญหาเกี่ยวกับการหาค่าเหมาะที่สุด (Optimization problem) ซึ่งเป็นที่ทราบกันดีว่าการคำนวณทำได้ค่อนข้างยากหรืออาจจะต้องใช้เวลานาน และในปัจจุบันมีการนำทฤษฎีเกี่ยวกับการหาค่าเหมาะที่สุดมาประยุกต์ใช้งานเป็นอย่างมากเพื่อทำให้ระบบควบคุมที่ได้ออกมานั้นมีประสิทธิภาพที่ดียิ่งขึ้น ในโครงการนี้ได้นำปัญหาที่เกี่ยวกับระบบควบคุมที่น่าสนใจมาแก้ปัญหาโดยใช้วิธีการของอสมการเมทริกซ์เชิงเส้น (LMI) ซึ่งมีตัวอย่างของปัญหาที่ได้นำมาศึกษาดังต่อไปนี้

4.1 การตรวจสอบเสถียรภาพด้วยวิธีลียาปูนอฟ (Lyapunov Stability Testing)

จากทฤษฎีเสถียรภาพของลียาปูนอฟในบทที่ 2 จะได้ว่า

$$\frac{d}{dt} x(t) = Ax(t) \quad (4.1)$$

มีเสถียรภาพภายในก็ต่อเมื่อสามารถหาเมทริกซ์ที่เป็นบวกแน่นอน $P > 0$ ซึ่งสอดคล้องกับอสมการ

$$A^T P + PA < 0 \quad (4.2)$$

ให้ Q เป็นเมทริกซ์ที่เป็นบวกแน่นอนใด ๆ แล้วอสมการเมทริกซ์เชิงเส้นในสมการที่ (4.2) สามารถเขียนใหม่ได้เป็น

$$A^T P + PA = -Q < 0 \quad (4.3)$$

ดังนั้นระบบพลวัตที่เขียนแทนด้วยสมการที่ (4.1) จะมีเสถียรภาพภายในก็ต่อเมื่อเมทริกซ์ P ที่คำนวณได้จากสมการที่ (4.3) นั้นมีความเป็นบวกแน่นอน สำหรับทุก ๆ เมทริกซ์ที่เป็นบวกแน่นอนใด ๆ Q และเมื่อจัดรูปสมการของลียาปูนอฟให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นจะได้ดังนี้

$$\begin{bmatrix} -P & 0 \\ 0 & A^T P + PA + Q \end{bmatrix} \leq 0 \quad (4.4)$$

ตัวอย่างศึกษาที่ 1 การตรวจสอบเสถียรภาพภายในของระบบด้วยวิธีของลียาปูนอฟ

กำหนดให้เมทริกซ์สถานะของระบบคือ $A = \begin{bmatrix} -1 & 2 \\ 0 & -2 \end{bmatrix}$ และเลือก $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ซึ่งเป็นบวกแน่นอน ดังนั้นหากสามารถหา P ที่สอดคล้องสมการ $A^T P + PA + Q = 0$ ที่ซึ่ง $P > 0$ แล้วจึงสรุปได้ว่าระบบมีเสถียรภาพภายใน

จากสมการของลียาปูนอฟ

$$A^T P + PA = -Q \quad (4.5)$$

ให้ $P = \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix}$ เป็นเมทริกซ์สมมาตร แทนค่าที่กำหนดมาให้ลงในสมการ (4.5) ได้ดังนี้

$$\begin{bmatrix} -1 & 0 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix} + \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 0 & -2 \end{bmatrix} = - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (4.6)$$

ซึ่งจะได้

$$P = \begin{bmatrix} 0.5 & 0.333 \\ 0.333 & 0.583 \end{bmatrix}$$

เนื่องจาก $P > 0$ (ค่าเฉพาะจะมีค่าเท่ากับ 0.2057, 0.8776) จึงสรุปได้ว่า ระบบนี้มีเสถียรภาพภายใน

ในกรณีที่มีเมทริกซ์ A มีขนาดใหญ่มากขึ้น การคำนวณเชิงวิเคราะห์ (Analytic approach) เช่นนี้ก่อให้เกิดความยุ่งยากมากขึ้น ดังนั้น จึงอาศัยการแก้ปัญหาด้วยโปรแกรมคอมพิวเตอร์ ซึ่งทำการจัดรูปให้อยู่ในรูปแบบของอสมการเมทริกซ์เชิงเส้น แล้วทำการแก้ปัญหาค่าที่เหมาะสมที่สุด ดังนี้

$$\begin{aligned} \min \quad & \text{Tr}(P) \\ \text{s.t.} \quad & \begin{bmatrix} -P & 0 \\ 0 & A^T P + PA + Q \end{bmatrix} \leq 0 \end{aligned} \quad (4.7)$$

เมื่อใช้ตัวแก้สมการเมทริกซ์เชิงเส้นแก้สมการนี้แล้วพบว่าได้

$$P = \begin{bmatrix} 0.5000 & 0.3333 \\ 0.3333 & 0.5833 \end{bmatrix}$$

ซึ่งสอดคล้องกับเมทริกซ์ P ที่ได้จากการคำนวณเชิงวิเคราะห์ใน (4.7) แต่อย่างไรก็ตาม พบว่า

$\tilde{P} = \begin{bmatrix} 1.5221 & 0.8397 \\ 0.8397 & 1.7108 \end{bmatrix}$ สอดคล้องกับสมการลิวาฮอฟด้วยเช่นกัน แต่ค่าของผลรวมเส้นทแยงมุมมีค่าไม่เท่ากัน นั่นคือ $\text{Tr}(P) = 1.0833$ ในขณะที่ $\text{Tr}(\tilde{P}) = 3.2329$ ดังนั้น เมทริกซ์ P ที่ได้จากตัวแก้สมการเมทริกซ์เชิงเส้น (LMI solver) นอกจากจะให้ผลเฉลย $P > 0$ แล้ว ยังเลือก P ที่ซึ่ง $\text{Tr}(P)$ มีค่าต่ำสุดด้วย

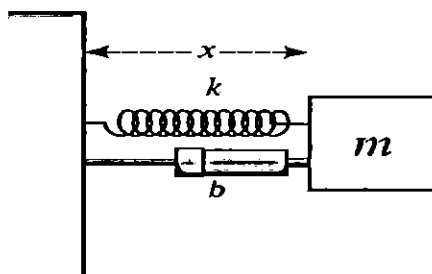
ข้อสังเกต สมการที่ (4.6) สามารถจัดให้อยู่ในรูปทั่วไปของอสมการเมทริกซ์เชิงเส้นได้ ดังนี้

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + p_1 \begin{bmatrix} -2 & 2 \\ 2 & 0 \end{bmatrix} + p_2 \begin{bmatrix} 0 & -3 \\ -3 & -4 \end{bmatrix} + p_3 \begin{bmatrix} 0 & 0 \\ 0 & -4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (4.8)$$

เพราะฉะนั้นจะได้ว่า

$$F_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, F_1 = \begin{bmatrix} -2 & 2 \\ 2 & 0 \end{bmatrix}, F_2 = \begin{bmatrix} 0 & -3 \\ -3 & -4 \end{bmatrix}, \text{ และ } F_3 = \begin{bmatrix} 0 & 0 \\ 0 & -4 \end{bmatrix} \quad (4.9)$$

ตัวอย่างศึกษาที่ 2 การตรวจสอบเสถียรภาพภายในของระบบสปริง-มวล-ตัวหน่วงด้วยวิธีของ ลีอาปูนอฟ



รูปที่ 4.1 ระบบสปริง-มวล-ตัวหน่วง (Spring-mass-damper system)

ถ้ากำหนดให้ $k=1, b=2, m=5$ แล้วระบบในรูปที่ 4.1 สามารถเขียนบรรยายในรูปของสมการสถานะได้เป็น

$$\dot{x} = Ax \quad \text{โดยที่} \quad A = \begin{bmatrix} 0 & 1 \\ -\frac{1}{5} & -\frac{2}{5} \end{bmatrix}$$

จากนั้น เลือก $Q = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ ซึ่งทำการจัดรูปให้อยู่ในรูปแบบของสมการเมทริกซ์เชิงเส้น แล้วทำการแก้ปัญหาค่าเหมาะที่สุด ดังนี้

$$\begin{aligned} \min \quad & \text{Tr}(P) \\ \text{s.t.} \quad & \begin{bmatrix} -P & 0 \\ 0 & A^T P + PA + Q \end{bmatrix} \leq 0 \end{aligned} \quad (4.10)$$

เมื่อใช้ตัวแก้สมการเมทริกซ์เชิงเส้นแก้สมการนี้แล้วพบว่าได้ $P = \begin{bmatrix} 5 & 5 \\ 5 & 15 \end{bmatrix}$ ซึ่ง $P > 0$ ดังนั้นจึง

สรุปว่าระบบนี้มีเสถียรภาพภายใน ผู้อ่านสามารถตรวจสอบได้ว่า ค่าเฉพาะของเมทริกซ์ A มีค่าเป็นลบทั้งหมด เพราะฉะนั้นระบบที่พิจารณาอยู่มีเสถียรภาพ ซึ่งสอดคล้องกับผลที่ได้ข้างต้น

4.2 การหาเสถียรภาพของระบบที่มีเมทริกซ์ระบบ A_k หลายตัว

จากหัวข้อที่ผ่านมาข้างต้นเป็นการศึกษาปัญหาเสถียรภาพของระบบเพียงระบบเดียว หรือกล่าวคือ ระบบที่มีเมทริกซ์ A เพียงเมทริกซ์เดียวเท่านั้น ซึ่งสามารถคำนวณหาได้ด้วยวิธีการคำนวณเชิงวิเคราะห์หรือใช้คำสั่งสำเร็จรูปในโปรแกรมแมทแล็บช่วยหาคำตอบ นั่นคือ โดยการใช้ฟังก์ชัน `LYAP()` ซึ่งเป็นฟังก์ชันที่สามารถใช้แก้สมการของลีอาปูนอฟได้โดยตรง แต่ในบางกรณี ระบบที่พิจารณาอาจจะประกอบด้วยเมทริกซ์ A หลายเมทริกซ์ ต่อไปนี้เป็นตัวอย่างหนึ่งของระบบดังกล่าว

พิจารณาระบบเชิงเส้นแปรตามเวลา เช่น การควบคุมจรวดนำวิถี ซึ่งพบว่าเมื่อเวลาเปลี่ยนไป มวลรวมของระบบย่อมมีการเปลี่ยนแปลงอันเนื่องมาจากเชื้อเพลิงที่ถูกใช้งานไป การที่จะออกแบบตัวควบคุมที่สามารถทำให้ผลตอบสนองของระบบนี้เป็นไปอย่างต้องการได้ สามารถใช้วิธีการทางทฤษฎีระบบควบคุมได้หลายประเภท ซึ่งหลักการควบคุมประเภทหนึ่งที่นิยมใช้คือ Gain scheduling control โดยมีหลักการคือ ทำการแปลงระบบไม่เชิงเส้น หรือระบบเชิงเส้นแปรตามเวลาให้เป็นระบบเชิงเส้นไม่แปรตามเวลา ณ แต่ละจุดปฏิบัติการ ซึ่งทำให้ได้ เซตของระบบเชิงเส้นไม่แปรตามเวลา จากนั้นทำการออกแบบตัวควบคุมเชิงเส้นสำหรับระบบเชิงเส้นไม่แปรตามเวลา ณ ตำแหน่งจุดปฏิบัติการ แล้วตัวควบคุมรวมของระบบทั้งหมดเกิดจากการประมาณค่าในช่วง (Interpolation) ของตัวควบคุมเชิงเส้นที่ได้ในแต่ละจุดปฏิบัติการ

จากตัวอย่างข้างต้น พบว่าเซตของระบบเชิงเส้นไม่แปรตามเวลาที่ได้ของระบบทั้งหมด เขียนได้เป็น

$$\dot{x} = A_k x + B_k u_k, \quad k = 1, 2, \dots, n \quad (4.11)$$

หากต้องการพิจารณาว่าระบบเชิงเส้นไม่แปรตามเวลา ณ แต่ละตำแหน่ง k มีเสถียรภาพภายในได้หรือไม่โดยอาศัยหลักการของลีโอทูนอฟนั้น สามารถทำได้ดังแสดงในตัวอย่างศึกษาที่ 1 และ 2 ข้างต้น กล่าวคือ ณ แต่ละตำแหน่ง k สามารถหาเมทริกซ์ P_k ที่สอดคล้องกับสมการลีโอทูนอฟ $A_k^T P_k + P_k A_k + Q = 0$ โดยที่ $P_k > 0$ ได้หรือไม่ แต่อย่างไรก็ตามหากต้องการทราบว่าเซตของระบบเชิงเส้นไม่แปรตามเวลา (4.11) มีเสถียรภาพภายในเชิงลีโอทูนอฟหรือไม่ หรือพูดอีกนัยได้ว่า เป็นไปได้หรือไม่ที่จะหาเมทริกซ์ P เพียงเมทริกซ์เดียวที่สอดคล้องกับสมการลีโอทูนอฟ $A_k^T P + P A_k + Q = 0$ สำหรับทุกค่า k โดยที่ $P > 0$ หากสามารถหาได้ จะกล่าวว่าเซตของระบบเชิงเส้นไม่แปรตามเวลา (4.11) มีเสถียรภาพภายในเชิงลีโอทูนอฟ

ตัวอย่างศึกษาที่ 3

พิจารณาเซตของระบบเชิงเส้นไม่แปรตามเวลาที่มีเมทริกซ์ระบบ (System matrices) ดังนี้

$$A_1 = \begin{bmatrix} -0.5121 & -0.4455 & 0.0840 \\ -0.4455 & -0.9286 & 0.1941 \\ 0.0840 & 0.1941 & -0.2945 \end{bmatrix}, A_2 = \begin{bmatrix} -0.5243 & 2.6006 & -0.1219 \\ -2.5839 & -0.4815 & 0.5637 \\ -0.3191 & -0.4804 & -1.5747 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} -1.1145 & 1.9928 & -1.4768 \\ -2.4781 & -0.8372 & 0.9643 \\ 0.1046 & -1.7606 & -1.8220 \end{bmatrix} \text{ และกำหนดให้เมทริกซ์ } Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

เมื่อแทนเมทริกซ์ A_k แต่ละตัวลงในฟังก์ชัน LYAP() ในแมทแลป จะได้เมทริกซ์ P_k ออกมา ดังนี้

$$\begin{aligned}
 P \text{ ที่ได้เมื่อ } A_1 &= \begin{bmatrix} -0.5121 & -0.4455 & 0.0840 \\ -0.4455 & -0.9286 & 0.1941 \\ 0.0840 & 0.1941 & -0.2945 \end{bmatrix} \text{ คือ } P_1 = \begin{bmatrix} 1.6781 & -0.8178 & -0.0606 \\ -0.8178 & 1.0231 & 0.4412 \\ -0.0606 & 0.4412 & 1.9713 \end{bmatrix} \\
 P \text{ ที่ได้เมื่อ } A_2 &= \begin{bmatrix} -0.5243 & 2.6006 & -0.1219 \\ -2.5839 & -0.4815 & 0.5637 \\ -0.3191 & -0.4804 & -1.5747 \end{bmatrix} \text{ คือ } P_2 = \begin{bmatrix} 1.0133 & 0.0054 & -0.1417 \\ 0.0054 & 1.0409 & 0.0268 \\ -0.1417 & 0.0268 & 0.3381 \end{bmatrix} \\
 P \text{ ที่ได้เมื่อ } A_3 &= \begin{bmatrix} -1.1145 & 1.9928 & -1.4768 \\ -2.4781 & -0.8372 & 0.9643 \\ 0.1046 & -1.7606 & -1.8220 \end{bmatrix} \text{ คือ } P_3 = \begin{bmatrix} 0.5919 & -0.0732 & -0.2069 \\ -0.0732 & 0.6756 & -0.1201 \\ -0.2069 & -0.1201 & 0.3786 \end{bmatrix}
 \end{aligned}$$

แม้ว่า P_k ที่ได้ทั้งสามจะมีความเป็นบวกแน่นอน ซึ่งส่งผลว่าระบบแต่ละระบบนั้นมีเสถียรภาพภายใน แต่ยังไม่สามารถสรุปได้ชัดเจนว่าระบบโดยรวมนั้นมีเสถียรภาพภายใน เพราะระบบโดยรวมจะมีเสถียรภาพภายในเมื่อสามารถหามเมทริกซ์ P เพียงเมทริกซ์เดียวที่สอดคล้องกับอสมการลียาปูนอฟ $A_k^T P + P A_k + Q \leq 0$ สำหรับทุกค่า k โดยที่ $P > 0$

หากแทนเมทริกซ์ P_2, P_3 แทนลงใน $A_k^T P_k + P_k A_k + Q$ พบว่าค่าเฉพาะที่ได้ มีทั้งมากกว่าศูนย์และน้อยกว่าศูนย์ กล่าวคือ

$$eig(A_1^T P_2 + P_2 A_1 + Q) = -1.5953, 0.5430, 0.8458$$

$$eig(A_1^T P_3 + P_3 A_1 + Q) = -0.7073, 0.6928, 0.8983$$

หรือนั่นคือ $A_k^T P_k + P_k A_k + Q$ ไม่เป็นทั้งลบตามต้องการ หรือแม้กระทั่งเอาค่าเฉลี่ยของ P_k ทั้งสามเมทริกซ์มา นั่นคือ

$$\bar{P} = (P_1 + P_2 + P_3)/3 = \begin{bmatrix} 1.0943 & -0.2951 & -0.1363 \\ -0.2951 & 0.9131 & 0.1159 \\ -0.1363 & 0.1159 & 0.8960 \end{bmatrix} \quad (4.12)$$

เมื่อแทนค่า $\bar{P}$ ลงใน $A_k^T \bar{P} + \bar{P} A_k + Q$ พบว่า $eig(A_1^T \bar{P} + \bar{P} A_1 + Q) = -0.7531, 0.4120, 0.5668$

ดังนั้น $A_k^T \bar{P} + \bar{P} A_k + Q$ จึงไม่เป็นเมทริกซ์ทั้งลบตามต้องการ

เพื่อแก้ปัญหาเหล่านี้จึงต้องใช้วิธีการจัดรูปให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นหลายๆ อสมการแล้วหาคำตอบพร้อมกันในครั้งเดียว ซึ่งวิธีการนี้สามารถทำได้ด้วยการใช้ตัวแก้สมการเมทริกซ์เชิงเส้น โดยสามารถจัดให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นได้ดังนี้

$$\begin{aligned}
 \min \quad & Tr(P) \\
 \text{s.t.} \quad & \begin{bmatrix} -P & 0 \\ 0 & A_k^T P + P A_k + Q \end{bmatrix} \leq 0, \quad k=1,2,3
 \end{aligned} \quad (4.13)$$

เมื่อแก้สมการด้วยตัวแก้สมการเมทริกซ์เชิงเส้น คำตอบที่ได้มีค่าเป็น

$$P_{new} = \begin{bmatrix} 2.1230 & -0.1363 & -0.2056 \\ -0.1363 & 2.0680 & 0.2178 \\ -0.2056 & 0.2178 & 2.0205 \end{bmatrix}$$

เนื่องจากค่าเฉพาะของ P_{new} คือ 1.8053, 1.9615, 2.4433 ซึ่งเป็นบวกทุกค่า ดังนั้น $P_{new} > 0$ และยิ่งไปกว่านั้น เมทริกซ์ P_{new} ที่ได้ยังสอดคล้องกับสมการลียาปูนอฟทุกค่า k นั่นคือ

$$\text{eig}(A_1^T P_{new} + P_{new} A_1 + Q) = -3.8603, 0, 0$$

$$\text{eig}(A_2^T P_{new} + P_{new} A_2 + Q) = -5.4491, -1.9176, 0$$

$$\text{eig}(A_3^T P_{new} + P_{new} A_3 + Q) = -8.3782, -3.2991, -0.5286$$

หรือนั่นคือ $A_k^T P + P A_k \leq -Q$ ทุกค่า $k=1,2,3$ ดังนั้นจึงสรุปได้ว่าการแก้สมการด้วยการจัดรูปให้อยู่ในรูปของสมการเมทริกซ์เชิงเส้นทั้งหมดก่อนแล้วจึงแก้สมการเพื่อหาคำตอบในครั้งเดียวนี้สามารถแก้ปัญหาของสมการของลียาปูนอฟแบบหลายสมการนี้ได้เป็นอย่างดี

4.3 การออกแบบตัวป้อนกลับสถานะ

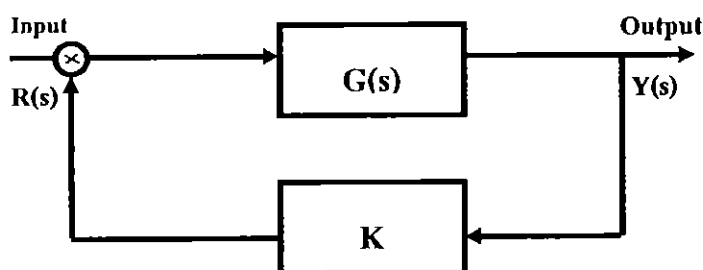
ระบบควบคุมที่ใช้ในการศึกษานี้จะเป็นระบบเชิงเส้นที่ไม่แปรเปลี่ยนตามเวลาซึ่งมีสมการสถานะดังนี้

$$\dot{x} = Ax + Bu \quad (4.14)$$

สมมติว่าผู้ออกแบบทราบตัวแปรสถานะทุกตัวในระบบ ดังนั้น ให้สัญญาณควบคุมการป้อนกลับของตัวแปรสถานะ (State-feedback controller) คือ $u = Kx$ โดยที่ K คืออัตราขยายของการป้อนกลับ ดังนั้นจะได้ระบบควบคุมวงปิดดังนี้

$$\dot{x} = (A + BK)x \quad (4.15)$$

การที่ระบบจะสามารถป้อนกลับได้นั้น ระบบจะต้องมีคุณสมบัติความสามารถควบคุมได้โดยสมบูรณ์ กล่าวคือ $\text{rank } C = \text{rank} \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = n$



รูปที่ 4.2 ระบบควบคุมแบบป้อนกลับ

การคำนวณหาอัตราขยาย K เพื่อให้ระบบวงปิดมีเสถียรภาพนั้น สามารถหาได้หลายวิธี วิธีที่นำเสนอในโครงการนี้ คือการใช้เงื่อนไขเสถียรภาพของลออาปูนอฟมาประยุกต์ใช้กับสมการที่ (4.15) นั่นคือ ระบบวงปิด $\dot{x} = (A + BK)x$ จะมีเสถียรภาพภายใน ถ้าสามารถหาเมทริกซ์สมมาตร $P > 0$ ที่สอดคล้องสมการ

$$(A + BK)^T P + P(A + BK) < 0 \quad (4.16)$$

ผู้อ่านจะเห็นได้จากสมการที่ (4.16) ว่ามีเมทริกซ์ตัวแปรที่ไม่ทราบค่าสองเมทริกซ์ที่คูณกัน จึงทำให้เกิดส่วนที่เป็นสมการเชิงเส้นสองตัวแปร (Bilinear term) โดยทั่วไปแล้วจะอยู่ในรูปของ Bilinear Matrix Inequalities (BMIs) [5, 9] ซึ่งพบว่าไม่มีความเป็นคอนเวกซ์ ดังนั้นจึงต้องทำการแปลงให้อสมการนี้เป็นอสมการเชิงเส้นตัวแปรเดียวก่อน ซึ่งสามารถทำได้ดังนี้

จาก

$$(A + BK)^T P + P(A + BK) < 0 \quad (4.17)$$

เนื่องจาก $P > 0$ จึงสามารถหาตัวผกผัน P^{-1} ได้ นอกจากนี้ $P^{-1} > 0$ ด้วย ดังนั้นคูณสองข้างของสมการที่ (4.17) ด้วย P^{-1} ซึ่งจะได้ว่า

$$P^{-1}[(A + BK)^T P + P(A + BK)]P^{-1} < 0 \quad (4.18)$$

ซึ่งสมมูลกับ

$$P^{-1}A^T + P^{-1}K^T B^T + AP^{-1} + BKP^{-1} < 0 \quad (4.19)$$

กำหนดให้ $Q = P^{-1}$ และ $Y = KP^{-1}$

จะได้

$$AQ + QA^T + BY + Y^T B^T < 0 \quad (4.20)$$

สมการที่ (4.20) ที่ได้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นแล้ว ดังนั้นจึงสามารถคำนวณหาผลเฉลยของสมการนี้ได้ด้วยวิธีการของอสมการเมทริกซ์เชิงเส้นได้ ซึ่งแสดงไว้ในตัวอย่างต่อไปนี้

ตัวอย่างศึกษาที่ 4 การออกแบบระบบควบคุมแบบป้อนกลับตัวแปรสถานะ โดยใช้ทฤษฎีเสถียรภาพ

$$\text{พิจารณาระบบ } \dot{x} = Ax + Bu \text{ โดยที่ } A = \begin{bmatrix} 5 & -1 \\ 1 & -1 \end{bmatrix} \text{ และ } B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

เนื่องจาก ค่าเฉพาะจริงของ A มีค่า 4.8284, -0.8284 ระบบขาดเสถียรภาพ ดังนั้นจึงจำเป็นต้องออกแบบหาอัตราขยายสำหรับการป้อนกลับตัวแปรสถานะ เพื่อให้ระบบวงปิดที่ได้มีเสถียรภาพ สิ่งแรกที่ต้องคำนึงคือเรื่อง ความสามารถควบคุมได้อย่างสมบูรณ์ของระบบ ให้พิจารณาเมทริกซ์ความสามารถควบคุมได้

$$C = [B \quad AB \quad A^2B \quad \dots \quad A^{n-1}B] \quad (4.21)$$

เมื่อแทนค่าเมทริกซ์ต่าง ๆ ลงไป จะได้ว่า $C = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}$ ซึ่งมีลำดับชั้นเท่ากับ 2 ดังนั้นระบบนี้สามารถควบคุมได้โดยสมบูรณ์ จากนั้นทำการออกแบบเพื่อหาอัตราขยาย K ที่ป้อนกลับให้กับระบบใช้สมการที่ (4.20) และเมื่อใช้ตัวแก้สมการเมทริกซ์เชิงเส้นแก้สมการได้คำตอบออกมาคือ

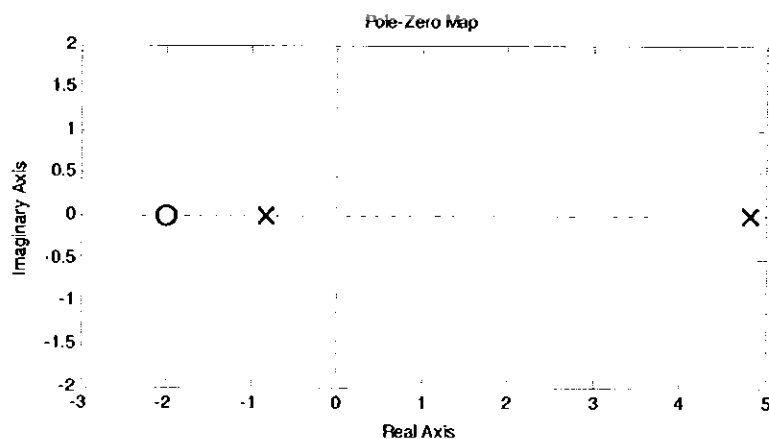
$$P = \begin{bmatrix} 1.0023 & -0.1821 \\ -0.1821 & 1.3664 \end{bmatrix}$$

ซึ่งมีค่าเจาะจงเท่ากับ 0.9269, 1.4418 ดังนั้น $P > 0$ นอกจากนี้ $K = [-5.3398 \quad -0.1199]$

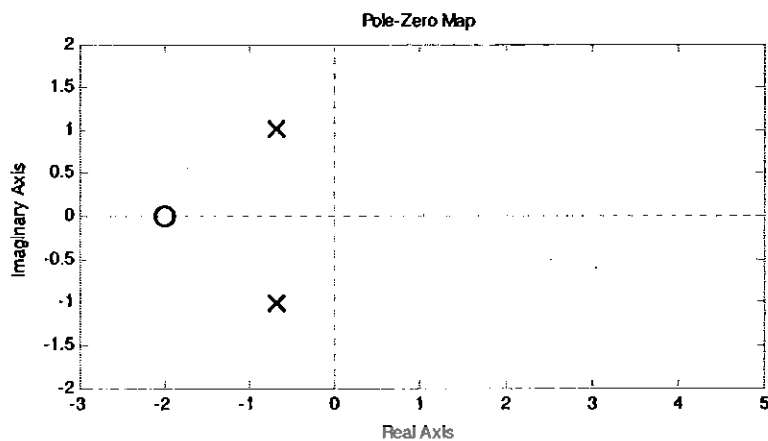
ด้วยอัตราขยายที่ได้นี้ ทำให้ระบบวงปิด $\dot{x} = (A + BK)x$ มีเสถียรภาพ ซึ่งดูได้จากค่าเจาะจงมีส่วนจริงติดลบ

$$\text{eig}(A + BK) = -0.6699 + 1.0055i, -0.6699 - 1.0055i$$

ค่าเจาะจงที่ไม่เสถียรของระบบหลังจากการออกแบบนี้มีการเลื่อนไปจากทางขวามือไปยังทางซ้ายมือ ดังแสดงในรูปที่ 4.3(a) และ 4.3(b) ส่งผลทำให้ระบบที่ทำการออกแบบแล้วนี้มีเสถียรภาพ เมื่อพิจารณาผลตอบสนองต่อสัญญาณขั้นบันได (Step response) ของระบบ พบว่า ก่อนการออกแบบ ระบบขาดเสถียรภาพ ซึ่งทำให้ผลตอบสนองของระบบไม่สามารถเข้าสู่สภาวะคงตัวได้ ซึ่งดูได้จากรูปที่ 4.4(a) ในขณะที่รูปที่ 4.4(b) แสดงผลตอบสนองของระบบที่ผ่านการออกแบบแล้ว จะเห็นว่าผลตอบสนองของระบบสามารถเข้าสู่สภาวะคงตัวได้และค่าพุ่ง (Overshoot) มีค่าเท่ากับ 15.8% ซึ่งมีค่าไม่มากนัก สำหรับรูปที่ 4.5 แสดงสัญญาณควบคุม ซึ่งพบว่าเมื่อระบบเข้าสู่สภาวะคงตัวแล้วจะ ไม่มีการป้อนสัญญาณควบคุมให้กับระบบอีก ดังเห็นได้จากที่เวลาประมาณ 5 วินาที พบว่าสัญญาณควบคุมก็มีค่าเป็น 0 ซึ่งสอดคล้องกับที่ผลตอบสนองเข้าสู่สภาวะคงที่ในเวลาประมาณ 5 วินาทีเช่นกัน (ดูกราฟผลตอบสนองในรูปที่ 4.4(b))



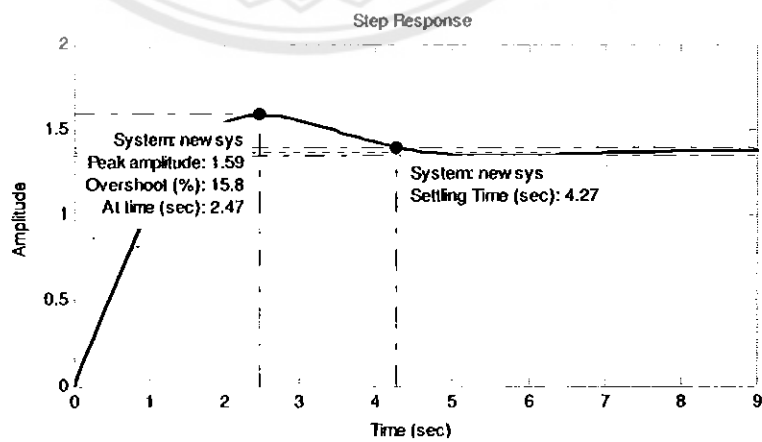
รูปที่ 4.3(a) ขั้ว (Pole) และศูนย์ (Zero) ของระบบก่อนออกแบบ



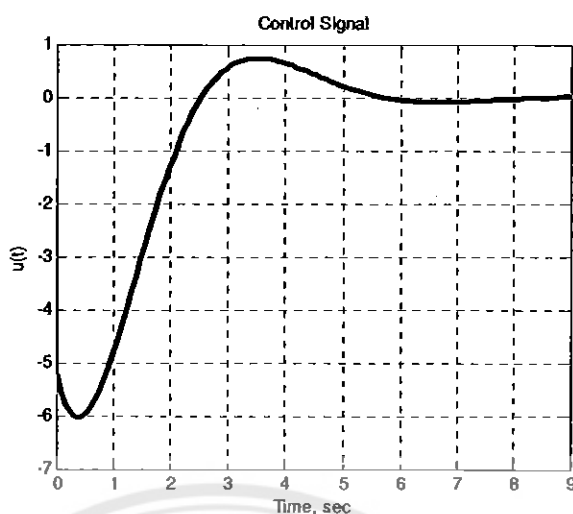
รูปที่ 4.3(b) ขั้ว (Pole) และศูนย์ (Zero) ของระบบหลังออกแบบ



รูปที่ 4.4(a) ผลตอบสนองต่อสัญญาณหนึ่งหน่วยของระบบก่อนออกแบบ



รูปที่ 4.4(b) ผลตอบสนองต่อสัญญาณหนึ่งหน่วยของระบบหลังออกแบบ



รูปที่ 4.5 สัญญาณควบคุมของระบบหลังการออกแบบ

จากตัวอย่างศึกษาที่ได้กล่าวไปแล้วนี้เป็นการศึกษาระบบที่ขาดเสถียรภาพ ซึ่งต้องการออกแบบอัตราขยายป้อนกลับเพื่อให้ระบบวงปิดมีเสถียรภาพเกิดขึ้น ซึ่งสามารถนำทฤษฎีเสถียรภาพของดีอาฟูนอฟมาประยุกต์ใช้ร่วมกับการแก้ปัญหาโดยการใช่วิธีของอสมการเมทริกซ์เชิงเส้นได้เช่นเดียวกัน ซึ่งจากตัวอย่างศึกษานี้แสดงให้เห็นว่าระบบที่ได้ทำการออกแบบขึ้นมาใหม่ด้วยวิธีดังกล่าวนี้มีเสถียรภาพเกิดขึ้นจริง

4.4 การวิเคราะห์ระบบควบคุมที่ใช้วิธี H_∞ (H_∞ Control Analysis)

ในหัวข้อนี้เป็นการวิเคราะห์ระบบควบคุมที่ใช้วิธีของ H_∞ โดยเป็นการออกแบบให้ระบบควบคุมยังสามารถที่จะทำงานอยู่ได้แม้อยู่ในสถานะที่เลวร้ายที่สุด (Worst case scenario) ซึ่งในการวิเคราะห์ระบบควบคุมด้วยวิธีนี้เริ่มต้นพิจารณาจากการวิเคราะห์ด้วยวิธีปริภูมิสถานะ (State space) ของระบบเชิงเส้นซึ่งบรรยายด้วยสมการ

$$\dot{x} = Ax + Bu \text{ และ } y = Cx + Du \text{ โดยกำหนดให้ } x(0) = 0$$

เมื่อ $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{p \times n}$ และ $D \in \mathbb{R}^{p \times p}$ ภายใต้สมมติฐานที่ว่าระบบมีเสถียรภาพ

จากบทที่ 2 พบว่า ฟังก์ชันถ่ายโอนของระบบดังกล่าวนี้มีค่าเป็น

$$G(s) = C(sI - A)^{-1}B + D \quad (4.22)$$

ดัชนีชี้วัดถึงกรณีที่ระบบอยู่ในสถานะที่เลวร้ายที่สุดนั้น ปกตินิยมแสดงโดย H_∞ นอร์ม นั่นคือ

$$\|G(s)\|_\infty = \sup_{\text{Re}(s) > 0} \bar{\sigma}(G(s)) = \sup_{\omega \in \mathbb{R}} \bar{\sigma}(G(j\omega)) \quad (4.23)$$

ซึ่ง H_∞ นอร์มนี้สามารถเขียนให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นได้ จากการศึกษาพบว่า

$\|G(s)\|_\infty < \gamma$ ก็ต่อเมื่อ $\gamma^2 I - D^T D > 0$ และต้องหามเมทริกซ์ $P = P^T$ ที่เป็นบวกแน่นอนที่ซึ่ง

$$(A^T P + PA + C^T C) + (PB + C^T D)(\gamma^2 I - D^T D)^{-1}(B^T P + D^T C) < 0 \quad (4.24)$$

อสมการข้างต้นซึ่งเป็นรูปหนึ่งของ อสมการริคคาติ สามารถเขียนให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้น (LMI) ได้โดยใช้ตัวประกอบพหุนามได้ดังนี้

$$\begin{bmatrix} -(A^T P + PA + C^T C) & -(PB + C^T D) \\ -(B^T P + D^T C) & \gamma^2 I - D^T D \end{bmatrix} > 0 \quad (4.25)$$

ซึ่งสมมูลกับ

$$\begin{bmatrix} A^T P + PA + C^T C & PB + C^T D \\ B^T P + D^T C & D^T D - \gamma^2 I \end{bmatrix} < 0 \quad (4.26)$$

ตัวอย่างศึกษาที่ 5 การวิเคราะห์ระบบควบคุมที่ใช้วิธีการของ H_∞ โดยมี

เมทริกซ์สถานะของระบบคือ $A = \begin{bmatrix} -1 & -4 & 1 \\ 1 & -2 & -2 \\ -3 & 4 & -1 \end{bmatrix}$

เมทริกซ์สัญญาณเข้า คือ $B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

เมทริกซ์สัญญาณออก คือ $C = \begin{bmatrix} 0 & -1 & 1 \end{bmatrix}$

และเมทริกซ์สัญญาณป้อนไปข้างหน้า คือ $D = 0$

ระบบนี้มีเสถียรภาพเนื่องจากค่าเฉพาะของเมทริกซ์ A มีค่าเท่ากับ

$$\text{eig}(A) = -0.7668 + 3.9534i, -0.7668 - 3.9534i, -2.4665$$

ผู้อ่านสามารถตรวจสอบได้ว่า ระบบนี้มีความสามารถควบคุมได้และสังเกตได้อย่างสมบูรณ์และเมื่อใช้ตัวแก้สมการเมทริกซ์เชิงเส้นแก้สมการ

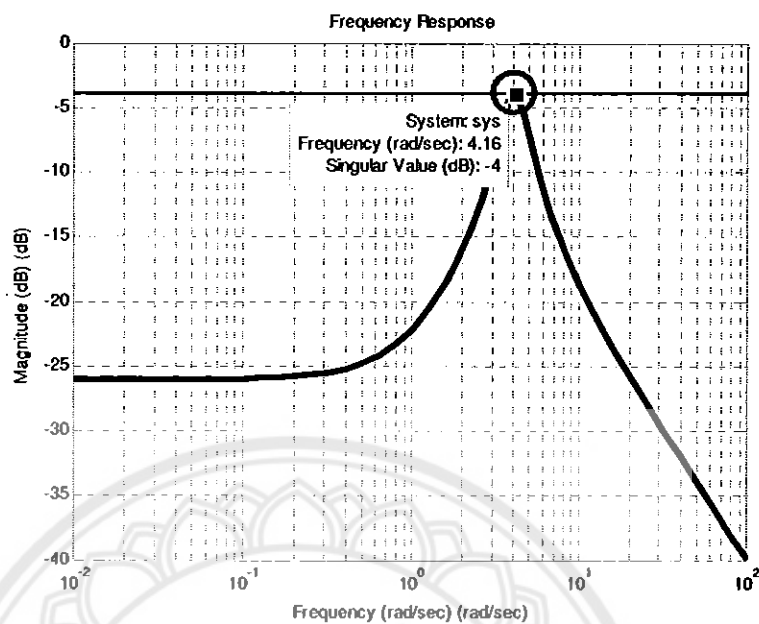
$$\begin{aligned} \min \gamma \\ \text{s.t.} \quad \begin{bmatrix} A^T P + PA + C^T C & PB + C^T D \\ B^T P + D^T C & D^T D - \gamma^2 I \end{bmatrix} < 0 \end{aligned} \quad (4.27)$$

ซึ่งคำตอบออกมาคือ

$$P = \begin{bmatrix} 0.8972 & -0.8450 & -0.5179 \\ -0.8450 & 2.9101 & -0.0446 \\ -0.5179 & -0.0446 & 1.5026 \end{bmatrix}$$

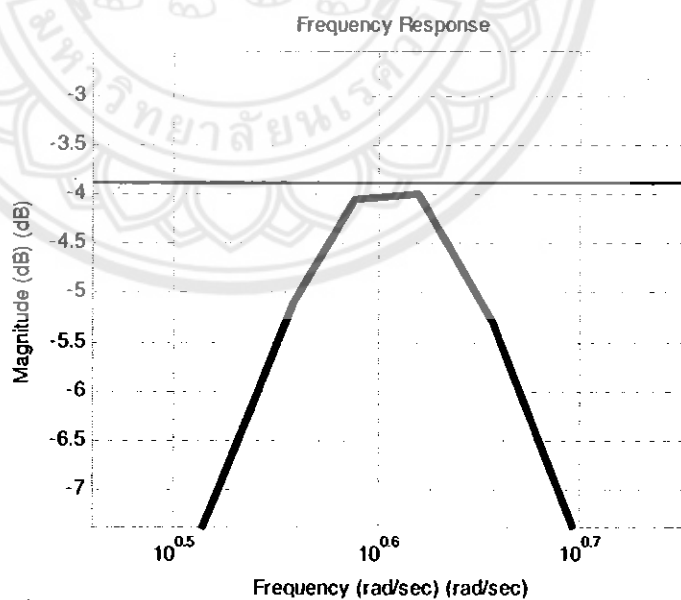
และค่า $\gamma = 0.6397$ หรือในหน่วยเดซิเบลคือ $\gamma_{\text{indB}} = -3.88 \text{ dB}$

เมื่อพิจารณาค่าเอกฐาน (σ) สูงสุดของระบบพบว่ามีความเท่ากับ -4 dB ซึ่งมีค่าน้อยกว่าค่า γ ที่ได้จากการแก้สมการด้วยวิธีอสมการเมทริกซ์เชิงเส้นจริง ส่งผลให้ระบบนี้ยังสามารถทำงานอยู่ได้แม้อยู่ในสภาวะที่เลวร้ายที่สุด ซึ่งสามารถดูได้จากรูปที่ 4.6



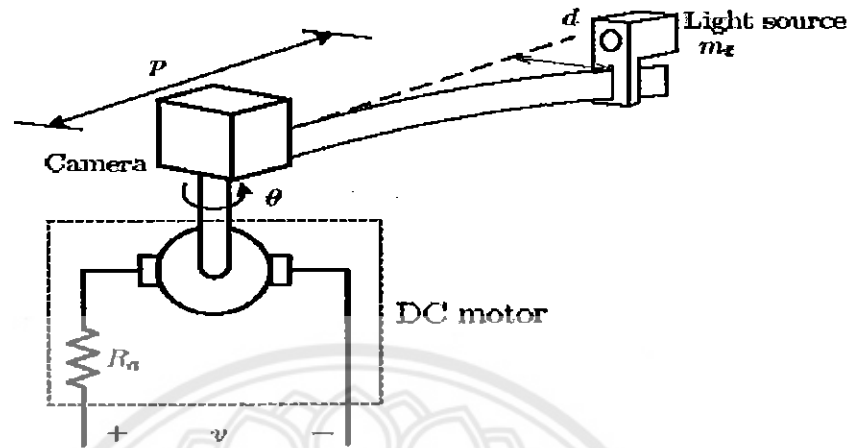
รูปที่ 4.6 ผลตอบสนองของระบบในโดเมนความถี่

จากบริเวณที่ถูกวงกลมไว้ในรูปที่ 4.6 นั้น เมื่อขยายรูปออกมาแล้วจะพบว่าเอกฐาน (σ) สูงสุดมีค่าน้อยกว่าค่า γ ซึ่งดูได้ดังนี้



รูปที่ 4.7 ค่าเอกฐานสูงสุด (Maximum singular value) ของระบบ

ตัวอย่างศึกษาที่ 6 [14] ระบบแขนกลหุ่นยนต์แบบอ่อนตัวโดยวิเคราะห์ระบบควบคุมนี้โดยใช้วิธีการของ H_∞ ซึ่งระบบทางกายภาพแสดงได้ดังนี้



รูปที่ 4.8 ระบบทางกายภาพของแขนกลหุ่นยนต์แบบอ่อนตัว (Flexible Robot Arm)

จากรูปจะพบว่าแขนกลจะทำงานโดยเมื่อป้อนแรงดัน v เข้าไปจะทำให้มอเตอร์กระแสตรงมีการหมุนจากมุม $\theta = \theta_0$ ไปยัง $\theta = \theta_f$ (ตำแหน่งเส้นปะ) โดยใช้อัตราเร็วเชิงมุม $\dot{\theta}$ และจะเห็นว่ามอเตอร์จะต่อกับแขนกลหุ่นยนต์แบบอ่อนตัว ทำให้แขนกลหมุนจากตำแหน่ง $d = d_0$ ไปยังตำแหน่ง $d = d_f$ (ตำแหน่งเส้นปะ) โดยใช้อัตราเร็ว $\dot{d}$ แบบจำลองของแขนกลหุ่นยนต์แบบอ่อนนี้สามารถเขียนให้อยู่ในรูปปริภูมิสถานะ (State Space) ได้ดังนี้

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -4.22 & 0 & -13.81 \\ 0 & 0 & 0 & 1 \\ 0 & -13.81 & -559.78 & -45.21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 7.85 \\ 0 \\ 25.71 \end{bmatrix} u \quad (4.28)$$

$$y = \begin{bmatrix} 1 & 0 & -0.38 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \end{bmatrix} u \quad (4.29)$$

- เมื่อ
- $x_1 = \theta$: มุมที่หมุน (rad)
 - $x_2 = \dot{\theta}$: อัตราเร็วเชิงมุมที่หมุน (rad / s)
 - $x_3 = d$: ตำแหน่งที่แขนกลแกว่ง (m)
 - $x_4 = \dot{d}$: อัตราเร็วที่แขนกลแกว่ง (m / s)
 - u = แรงดันเข้ามอเตอร์ (volt)
 - y = ตำแหน่งเชิงมุมใดๆของโหลด (m)

ระบบจะมีฟังก์ชันถ่ายโอน(Transfer Function) แบบวงปิดเป็นดังนี้

$$\frac{Y(s)}{R(s)} = \frac{-1.92 s^2 - 0.1899 s + 4394}{s^4 + 49.43 s^3 + 559.9 s^2 + 2362 s} \quad (4.30)$$

เมื่อใช้ตัวแก้สมการเมทริกซ์เชิงเส้นแก้สมการ

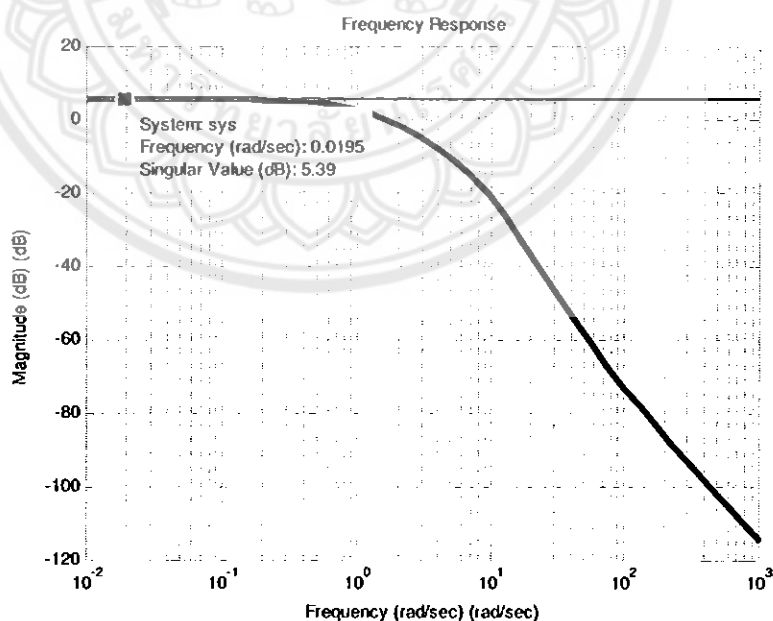
$$\begin{aligned} &\min \gamma \\ &s.t. \begin{bmatrix} A^T P + P A + C^T C & P B + C^T D \\ B^T P + D^T C & D^T D - \gamma^2 I \end{bmatrix} < 0 \end{aligned} \quad (4.31)$$

คำตอบที่ได้คือ

$$P = \begin{bmatrix} 0.4751 & 0.0624 & 1.3561 & -0.0015 \\ 0.0624 & 0.0662 & 0.4027 & 0.0011 \\ 1.3561 & 0.4027 & 21.2954 & 0.0098 \\ -0.0015 & 0.0011 & 0.0098 & 0.0304 \end{bmatrix}$$

และค่า $\gamma = 1.86$ หรือในหน่วยเดซิเบลคือ $\gamma_{dB} = 5.4 \text{ dB}$

เมื่อพิจารณาค่าซิงกูลาร์ (σ) สูงสุดของระบบพบว่ามีความเท่ากับ 5.4 dB ซึ่งมีค่าเท่ากับค่า γ ที่ได้จากการแก้สมการด้วยวิธีสมการเมทริกซ์เชิงเส้นจริง ส่งผลให้ระบบนี้ยังสามารถทำงานอยู่ได้ ซึ่งสามารถดูได้จากรูปที่ 4.9



รูปที่ 4.9 ผลตอบสนองของระบบในโดเมนความถี่

บทที่ 5

บทสรุป

5.1 สรุปผลการดำเนินงาน

จากการศึกษาวิธีการแก้ปัญหาของระบบควบคุมแบบเชิงเส้น (Linear control system) โดยใช้วิธีการของอสมการเมทริกซ์เชิงเส้น (Linear matrix inequality: LMI) พบว่าวิธีการนี้สามารถแก้ปัญหาทางระบบควบคุมได้เป็นอย่างดี จากการวิเคราะห์และออกแบบระบบควบคุมทั้งในรูปแบบคลาสสิก เช่น การป้อนกลับตัวแปรสถานะ (State feedback problem) หรือการวิเคราะห์ระบบโดยใช้วิธีการของ H_∞ เทียบกับการวิเคราะห์และออกแบบโดยการแปลงปัญหานั้น ๆ ให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้น พบว่าผลลัพธ์ที่ได้จากทั้งสองวิธีการนี้มีค่าที่เท่ากัน แต่อย่างไรก็ตาม ด้วยวิธีการแปลงระบบดังกล่าวนี้ นอกจากจะได้ผลลัพธ์ที่ถูกต้อง แม่นยำแล้ว ผลลัพธ์ที่ได้ยังมีค่าเหมาะที่สุดตามที่ผู้ใช้งานต้องการ เช่น ผลลัพธ์ที่ได้ นอกจากจะทำให้ระบบป้อนกลับเสถียรแล้ว ผลลัพธ์นั้นต้องมีผลรวมของค่าจะงค่าที่สุด เป็นต้น

นอกจากนี้ยังพบว่า มีปัญหาทางระบบควบคุมบางปัญหาที่ไม่สามารถหาผลลัพธ์ได้ด้วยวิธีการแก้ปัญหาแบบคลาสสิก (Classical approach) ยกตัวอย่าง เช่น ปัญหาของลิอาปูนอฟแบบหลายอสมการ (Multiple Lyapunov) ซึ่งเมื่อนำปัญหาดังกล่าวมาจัดให้อยู่ในรูปของอสมการเมทริกซ์เชิงเส้นแล้วพบว่าสามารถคำนวณหาผลลัพธ์ที่สอดคล้องกับสมการได้

ซอฟต์แวร์ที่ถูกเลือกนำมาใช้ในการคำนวณหาผลลัพธ์ของปัญหาที่จัดให้อยู่ในรูปแบบของอสมการเมทริกซ์เชิงเส้นนั้นคือ ซอฟต์แวร์ YALMIP (Yet Another LMI Parser) เนื่องจากซอฟต์แวร์นี้สามารถทำงานร่วมกับตัวแก้สมการเมทริกซ์เชิงเส้น (LMI solver) ได้อย่างหลากหลาย ซึ่งตัวแก้สมการแต่ละตัวมีคุณสมบัติในการทำงานที่แตกต่างกันออกไปขึ้นอยู่กับปัญหาที่พิจารณา เนื่องจาก บางครั้งพบว่าตัวแก้สมการที่เลือกใช้ไม่สามารถคำนวณหาผลลัพธ์ได้ (Solver not applicable) จึงควรเลือกตัวแก้สมการอื่นให้มีความเหมาะสมกับปัญหาที่ต้องการด้วย

5.2 ข้อเสนอแนะ

1. ควรมีการเปรียบเทียบผลลัพธ์ที่ได้จากการแก้ปัญหาด้วยวิธีการอื่นๆ
2. ควรมีการนำซอฟต์แวร์ตัวอื่นที่สามารถแก้ปัญหาของ LMI ได้มาใช้งาน เพื่อนำคำตอบนั้นมาเปรียบเทียบกับซอฟต์แวร์ YALMIP
3. ควรมีการนำวิธีการนี้ไปใช้แก้ปัญหาหรือออกแบบระบบควบคุมชนิดอื่นๆ ที่สนใจ

เอกสารอ้างอิง

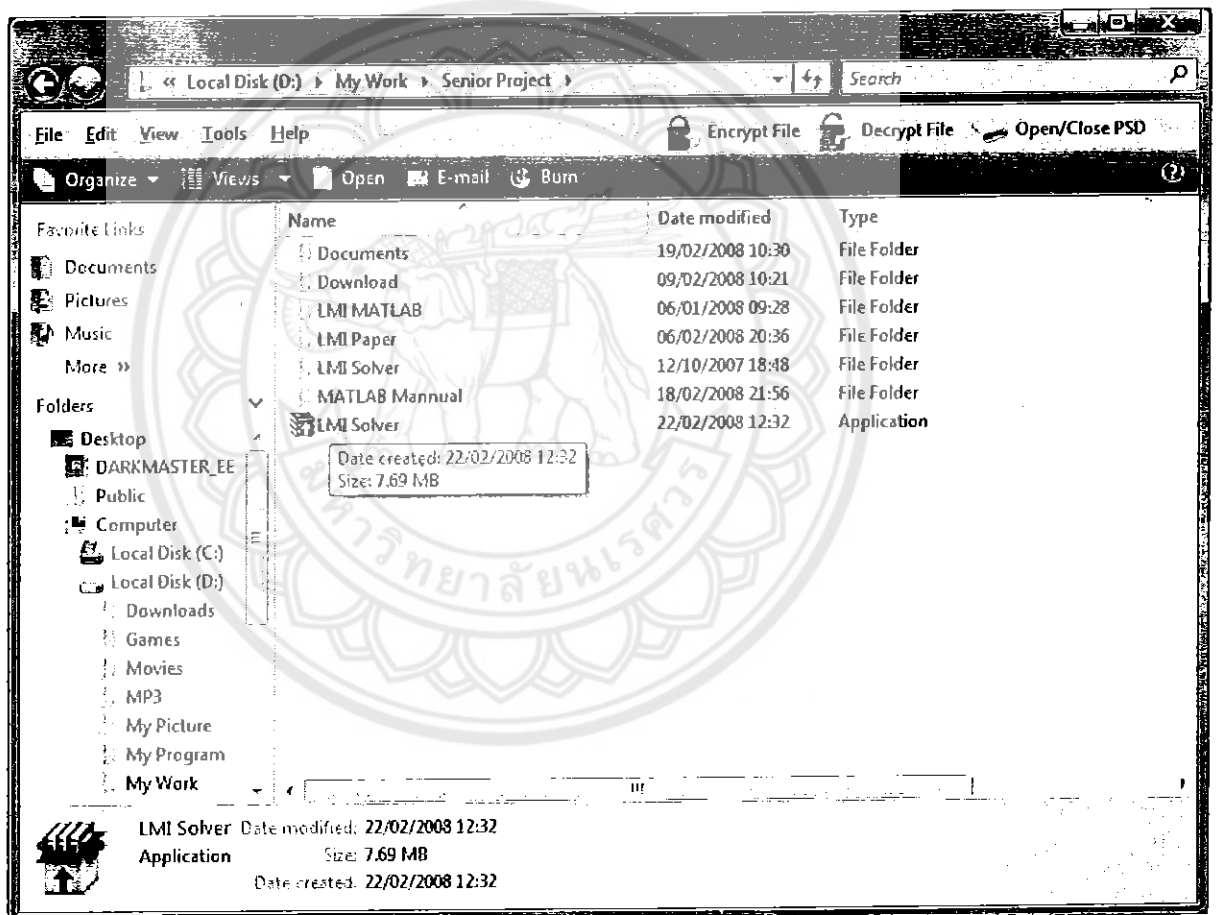
- [1] S. Boyd and L. Vandenberghe. **Convex Optimization**. United Kingdom: Cambridge University Press. 2004.
- [2] S. Boyd, L. Ghaoui, E. Feron and V. Balakrishnan. **Linear Matrix Inequalities in System and Control Theory**. vol 15 of Studies in Applied Mathematics. Philadelphia: SIAM. 1994.
- [3] G. F. Franklin, J. D. Powell and A. Emani-Naeini. **Feedback control of dynamic systems**. 4 th Edition, America: Prentice Hall. 2002.
- [4] M. Gopal. **Control systems principles and design**. 2nd Edition, New York: McGraw-Hill. 2003.
- [5] D. Henrion. "Course on LMI: LMIs in Systems Control (State-Space Methods Control Design)." <http://www.laas.fr/~henrion/events.html>. 2006.
- [6] D. Henrion. "Graduate course on polynomial method for robust control: Robust Analysis with Linear Matrix Inequalities and Polynomial Matrices." <http://www.laas.fr/~henrion/events.html>. 2001.
- [7] E. Kreyszig. **Introductory Functional Analysis with Applications**. Singapore: Wiley. 1987.
- [8] N. S. Nise. **Control Systems Engineering**. 4 th Edition, America: Wiley. 2004.
- [9] J. G. VanAntwerp and R. D. Braatz. "A tutorial on linear and bilinear matrix Inequalities." *Journal of Process Control* 10. 2000. pp. 363-385.
- [10] S. Weiland. "DISC course on Linear Matrix Inequalities in Control." <http://www.cs.ele.tue.nl/sweiland/lmi.html>. 2004.
- [11] "Convex set and convex function." [online]. Available: <http://en.wikipedia.org/wiki/Convex>.
- [12] "Lector 15: Symmetric matrices, quadratic forms, matrix norm and SVD." http://www.eecs.berkeley.edu/~elghaoui/courses_files/ee227a/Fall%202007.
- [13] วัชรพงษ์ ไชวิฑูรกิจ. คณิตศาสตร์วิศวกรรมไฟฟ้าขั้นสูง. กรุงเทพมหานคร: สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย. 2546.
- [14] สมชาย เชื้อบุญมี. การออกแบบระบบควบคุมโดยใช้วิธีการหาค่าเหมาะสมที่สุด (LQ). พิมพ์โลก: มหาวิทยาลัยนเรศวร 2547.

ภาคผนวก ก

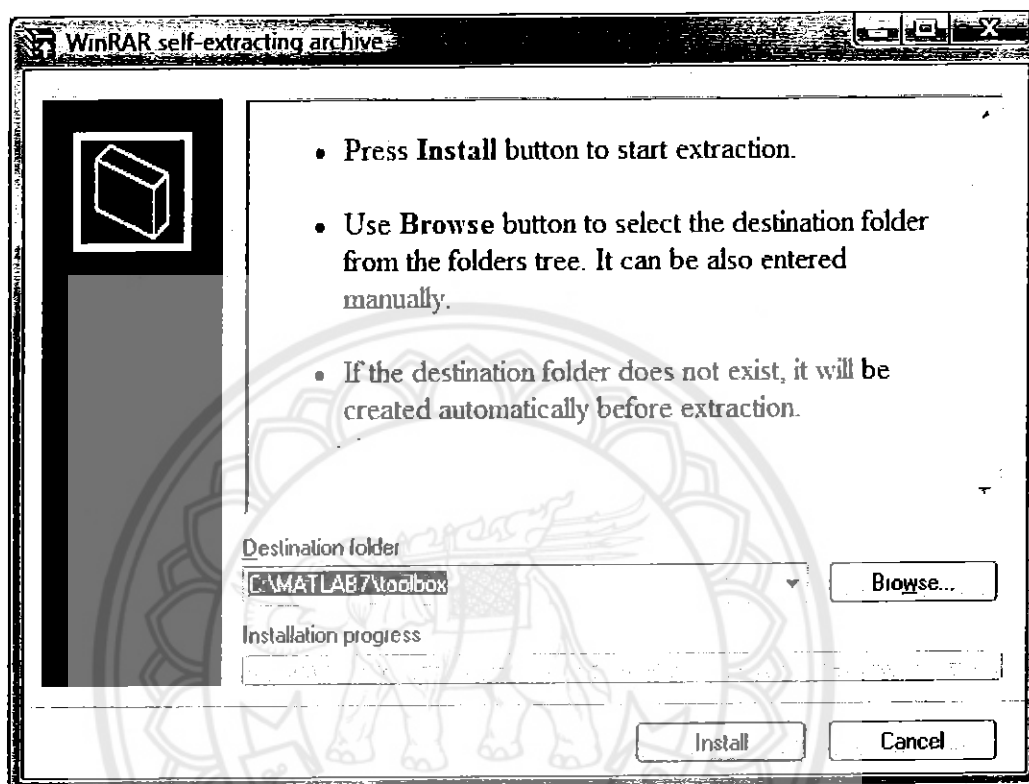
คู่มือการติดตั้งและใช้งานตัวแก้สมการเมทริกซ์เชิงเส้น

การติดตั้งตัวแก้สมการเมทริกซ์เชิงเส้น (LMI solver)

1. ถัดลอกตัวแก้สมการเมทริกซ์เชิงเส้น (LMI solver) จากแผ่นซีดีไปยังไดเรกทอรีที่ต้องการ



2. ดับเบิลคลิก "LMI Solver.exe" จากนั้นกดปุ่ม Install เพื่อทำการติดตั้งลงในโปรแกรม MATLAB



3. รอสักครู่เพื่อให้ทำการติดตั้งจนเสร็จสมบูรณ์

วิธีการใช้งานตัวแก้สมการเมทริกซ์เชิงเส้น (LMI solver) ในโปรแกรม MATLAB

1. เปิดโปรแกรม MATLAB ขึ้นมา

MATLAB®

The Language of Technical Computing

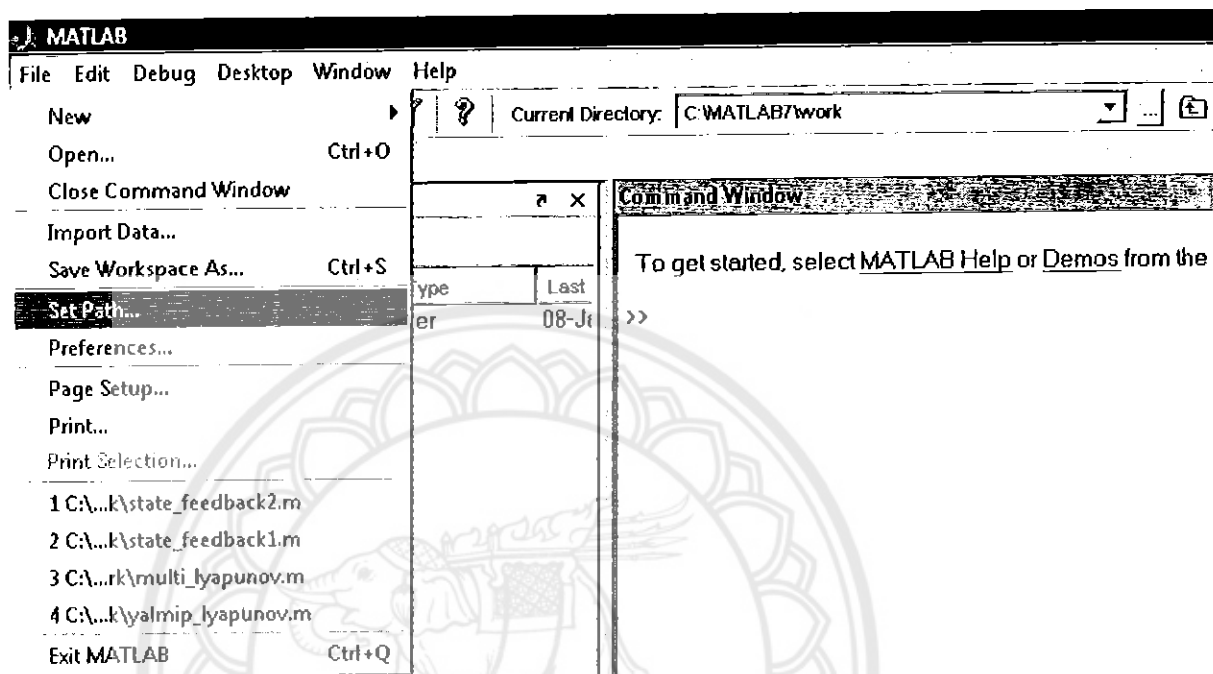
Version 7.0.0.19920 (R14)

May 06, 2004

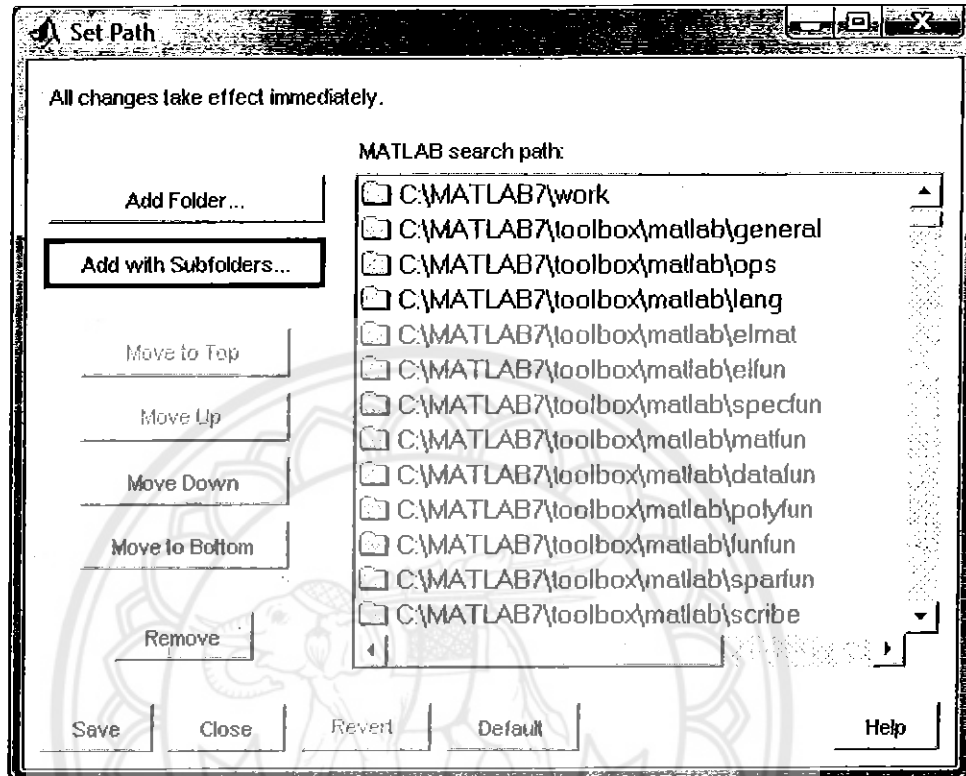
License Number: 216666



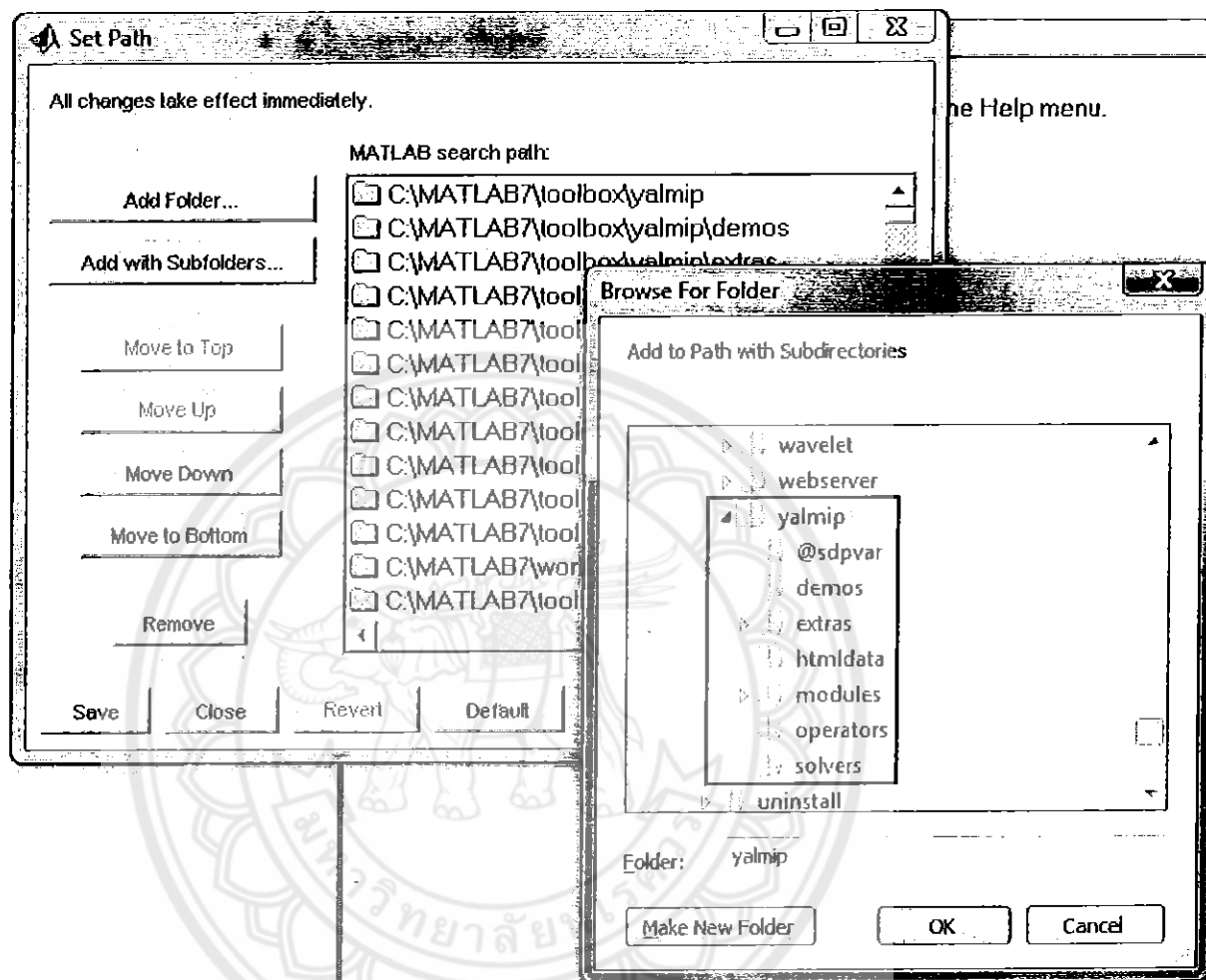
2. จากนั้นต้องมีการเพิ่มหรือติดตั้ง Path หรือไดเรกทอรีของ LMI solver ก่อนเพื่อให้โปรแกรม MATLAB รู้จักโดยไปที่ File -> Set Path...



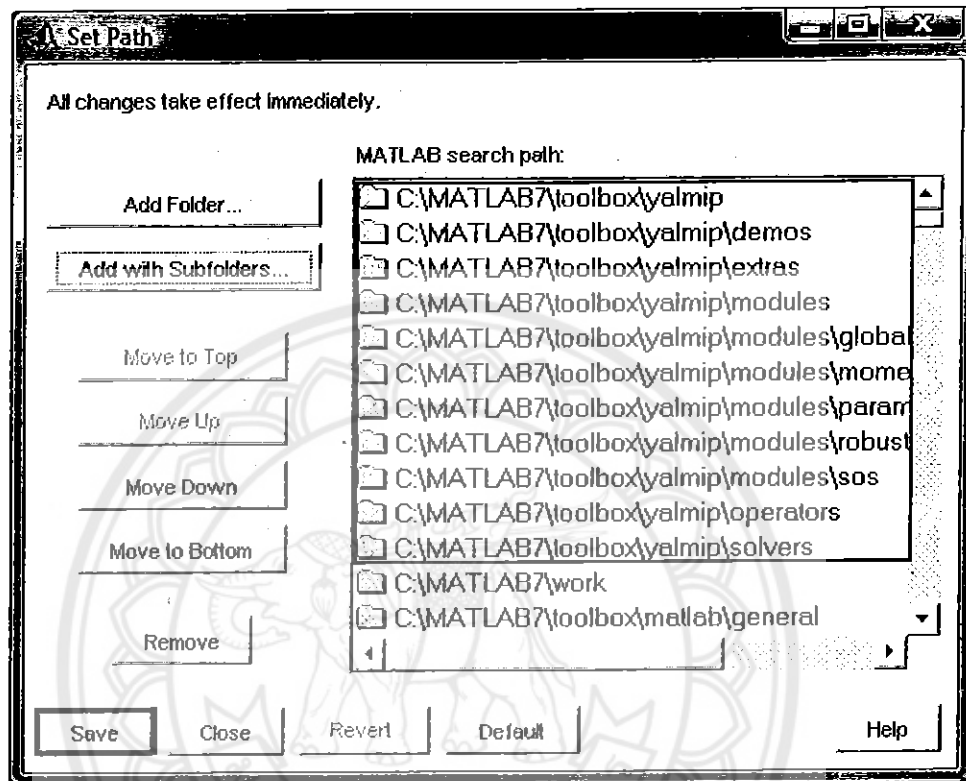
3. จากนั้นกดปุ่ม Add with Subfolders ... เพื่อทำการเพิ่มไดเรกทอรีใหม่เข้าไป



4. เลือกไดเรกทอรีของ LMI solver ซึ่งก็คือ ScDuMi และ YALMIP จากนั้นกดปุ่ม OK



5. หลังจากนั้นจะพบว่า มีไคเรคทอรีใหม่เพิ่มเข้ามาแล้ว ตรวจสอบว่าถูกต้องหรือไม่ แล้วกดปุ่ม Save เพื่อทำการบันทึก



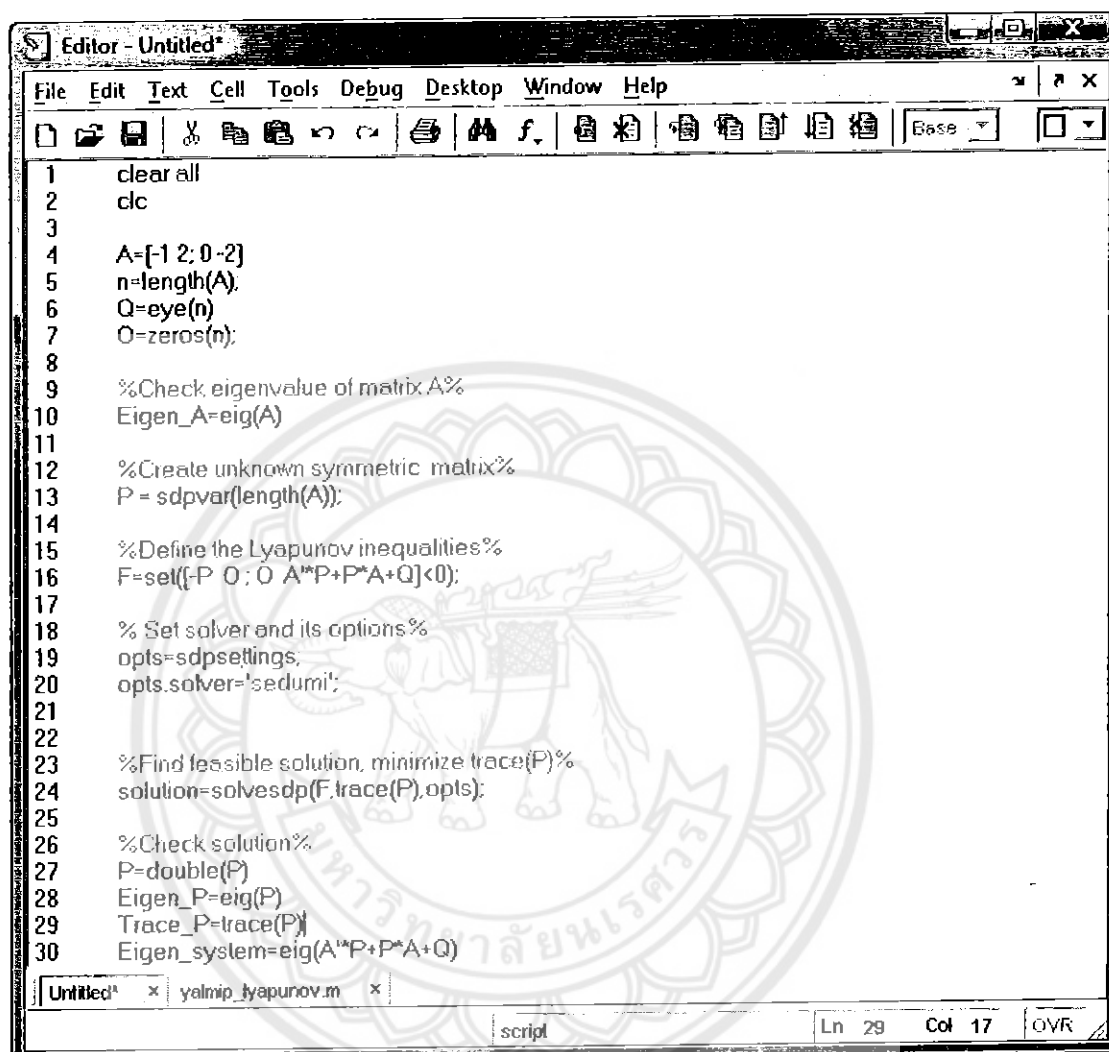
6. จากนั้นพิมพ์คำว่า "Edit" ใน Command Window เพื่อเรียกตัวอีดิตเตอร์ของโปรแกรม MATLAB

Command Window

To get started, select MATLAB Help or Demos from the Help menu.

```
>> edit
>>
```

7. เขียนโปรแกรมเพื่อแก้ LMI อย่างง่ายใน Editor ตามคำสั่งข้างล่างนี้



```

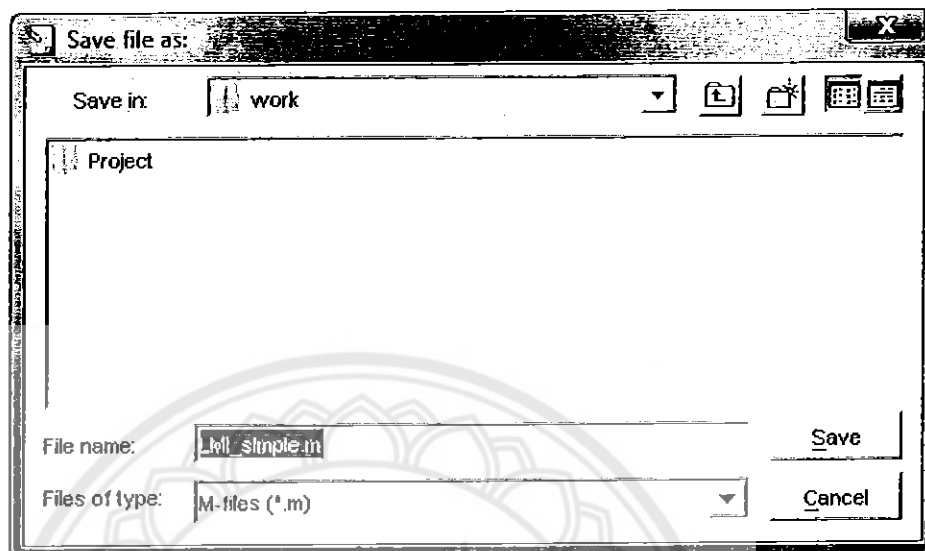
1 clear all
2 clc
3
4 A=[-1 2; 0 -2]
5 n=length(A);
6 Q=eye(n);
7 O=zeros(n);
8
9 %Check eigenvalue of matrix A%
10 Eigen_A=eig(A)
11
12 %Create unknown symmetric matrix%
13 P = sdpvar(length(A));
14
15 %Define the Lyapunov inequalities%
16 F=set([-P O; O A'*P+P*A+Q]<0);
17
18 % Set solver and its options%
19 opts=sdpssettings;
20 opts.solver='sedumi';
21
22
23 %Find feasible solution, minimize trace(P)%
24 solution=solvesdp(F,trace(P),opts);
25
26 %Check solution%
27 P=double(P)
28 Eigen_P=eig(P)
29 Trace_P=trace(P)
30 Eigen_system=eig(A'*P+P*A+Q)

```

Untitled x yalmip_lyapunov.m x

script Ln 29 Col 17 OVR

8. กดปุ่ม F5 เพื่อทำการ Save และ Run โปรแกรม



9. ผลการทำงานของ LMI solver ที่ได้นี้จะแสดงใน Command Window ซึ่งสามารถดูและตรวจคำตอบได้ ดังรูปข้างล่างนี้

```

Command Window

A =
    -1     2
     0    -2

Q =
     1     0
     0     1

Eigen_A =
    -1
    -2

SeDuMi 1.1R3 by AdvOL, 2006 and Jos F. Sturm, 1998-2003.
Alg = 2: xz-corrector, theta = 0.250, beta = 0.500
eqs m = 3, order n = 5, dim = 17, blocks = 2
nnz(A) = 8 + 0, nnz(ADA) = 9, nnz(L) = 6
it:  b*y   gap   delta rate  \IP* \ID* feas cg cg prec
0:   7.20E+000 0.000
1: -1.16E+000 1.96E+000 0.000 0.2720 0.9000 0.9000 1.72 1 1 2.1E+000
2: -1.08E+000 3.79E-001 0.000 0.1934 0.9000 0.9000 1.80 1 1 2.9E-001
3: -1.08E+000 1.54E-002 0.000 0.0407 0.9900 0.9900 1.21 1 1 1.1E-002
4: -1.08E+000 5.02E-007 0.000 0.0000 1.0000 1.0000 1.01 1 1 3.5E-007
5: -1.08E+000 5.14E-014 0.000 0.0000 1.0000 1.0000 1.00 1 1 3.5E-014

iter seconds digits   c*x       b*y
 5    0.4  Inf -1.08333333333e+000 -1.08333333333e+000
|Ax-b| = 4.2e-014, |Ay-c|_+ = 6.1E-015, |x| = 9.0e-001, |y| = 8.4e-001

Detailed timing (sec)
Pre  IPM  Post
2.028E-001 4.368E-001 4.680E-002
Max-norms: ||b||=1, ||c|| = 1,
Cholesky |add|=0, |skip| = 0, ||L.L|| = 1.

```

```
P =
    0.5000    0.3333
    0.3333    0.5833
```

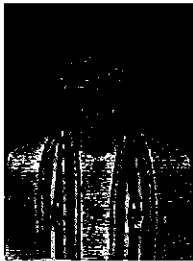
```
Eigen_P =
    0.2057
    0.8776
```

```
Trace_P =
    1.0833
```

```
Eigen_system =
    1.0e-014 *
    0.5055
    0.6048
```

ข้อเสนอแนะ: สำหรับรายละเอียดเพิ่มเติมเกี่ยวกับ LMI solver สามารถศึกษาได้จากการอ่านข้อมูลในไฟล์ `yalmip.html` ซึ่งในโฟลเดอร์ของ `yalmip` หรือใช้คำสั่ง `help yalmip` ในโปรแกรม MATLAB และสำหรับ `SeDuMi` ให้ศึกษาไฟล์ `SeDuMi_Guide_11.pdf` ในโฟลเดอร์ของ `SeDuMi` หรือใช้คำสั่ง `help sedumi` ในโปรแกรม MATLAB ได้เช่นเดียวกัน

ประวัติผู้ทำโครงการ



ชื่อ นายสิทธิโชค ใจยอด

ภูมิลำเนา 16 หมู่ 5 ตำบลแม่ตาว อำเภอแม่สอ ด จังหวัดตาก 63110

ประวัติการศึกษา

- สำเร็จการศึกษาระดับมัธยมศึกษาจากโรงเรียนสรรพวิทยาคม จังหวัดตาก
- ปัจจุบันกำลังศึกษาในระดับปริญญาตรีชั้นปีที่ 4 สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร จังหวัดพิษณุโลก

E-mail: choke_ee@hotmail.com

