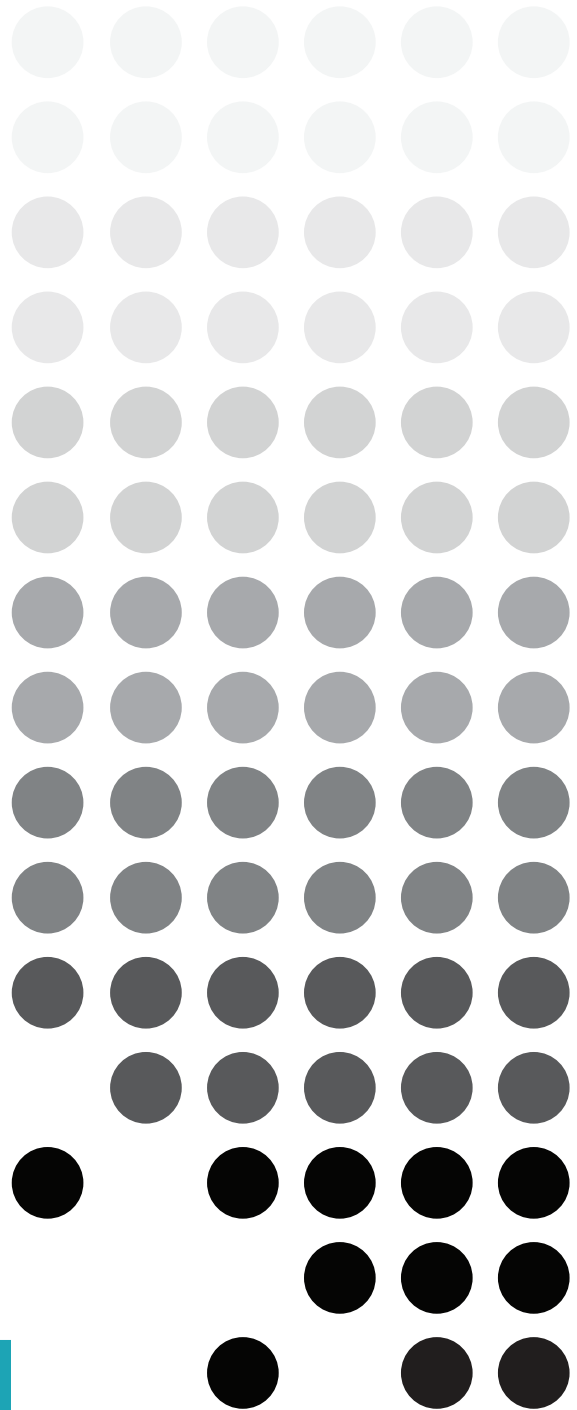




การหาเอกลักษณ์
ของระบบ



SYSTEM IDENTIFICATION

จิตโกมุต สงศิริ

การหาเอกลักษณ์ของระบบ

System Identification

การหาเอกลักษณ์ของระบบ System Identification

จิตโกมุท ส่งศิริ

ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย

จิตโกมุท ส่งศิริ

การหาเอกลักษณ์ของระบบ / จิตโกมุท ส่งศิริ

1. เอกลักษณ์ของระบบ.

003.1

ISBN (e-Book) 978-616-590-946-4

สพจ. Q22/09/2565



สรรคุณค่าวิชาการ สู่สังคม
Knowledge to All
www.cupress.chula.ac.th

หนังสือเล่มนี้ผ่านการตรวจคุณภาพวิชาการจากผู้ทรงคุณวุฒิของสำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย

สงวนลิขสิทธิ์เนื้อหาและภาพประกอบของหนังสือเล่มนี้โดย จิตโกมุท ส่งศิริ © พ.ศ. 2565

ห้ามคัดลอก จัดพิมพ์ หรือทำซ้ำ ก่อนได้รับอนุญาตเป็นลายลักษณ์อักษร

จัดทำโดย จิตโกมุท ส่งศิริ

ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

ครั้งที่ 1 พ.ศ. 2564

บรรณาธิการอำนวยการ : นางอรทัย นันทนาดีชัย รองศาสตราจารย์ ดร.อรัญ หาญสืบสาย

กองบรรณาธิการฝ่ายวิชาการ : ศาสตราจารย์กิตติคุณ ดร.ปิยนาด บุนนาค

รองศาสตราจารย์ ดร.พิมพ์พันธ์ เดชะคุปต์

รองศาสตราจารย์ นายแพทย์ชิษณุ พันธุ์เจริญ

รองศาสตราจารย์ ดร.วิมลวรรณ พิมพ์พันธุ์

ผู้ประสานงาน : วาสนา ชำเซ็น

ผู้พิสูจน์อักษร : บุศรินทร์ สรวัตร

ออกแบบปก : กิตติภูมิ ส่งศิริ และอาทิมา อุทัยชลาชนนท์

ออกแบบรูปเล่ม : จิตโกมุท ส่งศิริ

สั่งซื้อได้ที่ ศูนย์หนังสือจุฬาลงกรณ์มหาวิทยาลัย

ถนนพญาไท เขตปทุมวัน กรุงเทพฯ 10330

<http://www.chulabook.com> โทร. 08-6323-3703-4

customer@cubook.chula.ac.th, info@cubook.chula.ac.th

Apps: CU-eBook Store

กิตติกรรมประกาศ

ผู้เขียนขอขอบพระคุณ ผศ. ดร.มานพ วงศ์สายสุวรรณ สำหรับคำแนะนำและข้อมูลที่เป็นประโยชน์ เมื่อผู้เขียนเริ่มสอนวิชาการหาเอกลักษณ์ของระบบในปีพ.ศ.2554 ขอขอบพระคุณ รศ. ดร.วัชรพงษ์ ไขว้ทูลกิจ ผศ. ดร.สุชิน อรุณสวัสดิ์วงศ์ และ ศ. ดร.เดวิด บรรเจิดพงษ์ชัย คณาจารย์ในสาขาระบบควบคุม ภาควิชาวิศวกรรมไฟฟ้า ที่ได้ให้พื้นฐานความรู้ที่เป็นรากฐานอันสำคัญต่อผู้เขียนขณะเป็นนิสิต และให้คำแนะนำอันเป็นประโยชน์ ต่อการสอนและพัฒนาเอกสารทางวิชาการ ในตลอดเวลาที่ผู้เขียนปฏิบัติหน้าที่ในภาควิชาวิศวกรรมไฟฟ้าตั้งแต่ปีพ.ศ. 2554

ขอขอบพระคุณ Professor Lieven Vandenberghe ซึ่งเป็นอาจารย์ที่ปรึกษาของผู้เขียนในการเรียนระดับปริญญาเอก ณ มหาวิทยาลัยแคลิฟอร์เนีย ลอสแอนเจลิส ประเทศสหรัฐอเมริกา อาจารย์ได้จุดประกายให้ผู้เขียนสนใจทำวิจัยด้านเทคนิคการหาค่าเหมาะที่สุดและการเรียนรู้ของเครื่อง ท่านเป็นแบบอย่างที่ดีในด้านการสอนอันมีเนื้อหาที่ทันสมัยจากงานวิจัยเสมอ การมีปรัชญาในการเขียนผลงานในรูปแบบที่ง่ายต่อผู้อ่านเป็นสำคัญ สิ่งเหล่านี้ ผู้เขียนได้เรียนรู้และนำมาปรับใช้จนถึงปัจจุบัน

ขอขอบคุณคุณฉวีวรรณศาสตร จุฬาลงกรณ์มหาวิทยาลัย ที่ให้โอกาสผู้เขียนปฏิบัติงานเพิ่มพูนความรู้ทางวิชาการ ตั้งแต่ 1 สิงหาคม พ.ศ. 2563 ถึง 31 กรกฎาคม พ.ศ. 2564 ขอขอบคุณคณาจารย์ในสาขาระบบควบคุมและในภาควิชา ที่ช่วยแบ่งเบาภาระการสอนตลอดระยะเวลาที่ผู้เขียนปฏิบัติงาน เพื่อเพิ่มพูนความรู้ทางวิชาการ ขอขอบคุณ ภาควิชาวิศวกรรมไฟฟ้า ที่สนับสนุนการทำงานบุคลากรอย่างต่อเนื่อง รศ.ดร.แนบบุญ หุ่นเจริญ รศ.ดร.สุพัฒนา เอื้อทวีเกียรติ ผศ.ดร.อภิวัฒน์ เล็กอุทัย และ รศ.ดร.นิศาชล ตั้งเสถียรวิสัย ที่ให้คำแนะนำอันเป็นประโยชน์ต่อการทำวิจัยอันเป็นเนื้อหาที่เกี่ยวข้องในการเขียนหนังสือเล่มนี้ รวมถึงเพื่อนคณาจารย์ในจุฬาลงกรณ์มหาวิทยาลัย ที่แลกเปลี่ยนกำลังใจซึ่งกันและกันในการทำงานตลอดมา ขอขอบคุณนิสิตในที่ปรึกษาระดับป.ตรี ป.โท และป.เอกทุกคน นิสิตในสาขาระบบควบคุมและที่เรียนวิชา 2102531 การหาเอกลักษณ์ของระบบในอดีตจนถึงปัจจุบัน ที่ได้ให้ความเห็นอันเป็นประโยชน์ต่อการสอนและการทำวิจัยร่วมกัน ครูบาอาจารย์ระดับประถมศึกษาจนถึงมหาวิทยาลัยที่ได้อบรมสั่งสอนและเป็นแบบอย่างที่ดีต่อผู้เขียน

ขอขอบคุณกองบรรณาธิการแห่งสำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัยที่ให้คำปรึกษาอย่างดีในการจัดพิมพ์ โดยเฉพาะอย่างยิ่งผู้ทรงคุณวุฒิที่กรุณาอ่านต้นฉบับและให้คำแนะนำที่เป็นประโยชน์ และคุณบุศรินทร์ สรวัตร ที่ตรวจพิสูจน์อักษรอย่างถี่ถ้วน

ขอขอบคุณ รศ. ดร.เบญจมาศ พนมรัตน์รักษ์ และผศ. ดร.สุดชาย บุญโต มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี และรศ.ดร.พีระยศ แสนโกชณ์ มหาวิทยาลัยเกษตรศาสตร์ ที่ให้คำแนะนำอันเป็นประโยชน์ในการเรียบเรียงเนื้อหา ขอขอบคุณนายณภัทร สามเसार นิสิตที่เรียนวิชาการหาเอกลักษณ์ของ

ระบบ ที่ช่วยตรวจทานแก้ไข สิ่งที่มีผิดพลาดในหนังสือเล่มนี้ และนายปรินทร์ มโนมัยเสาวภาคย์ ที่ได้ช่วยเตรียมรูปผลการทดลองบางส่วนเพื่ออธิบายประกอบเนื้อหา

ผู้เขียนขอขอบคุณโครงการสนับสนุนการเขียนตำรา/หนังสือ/คู่มือ ของคณาจารย์และบุคลากรสายสนับสนุนคณะวิศวกรรมศาสตร์ที่ให้การสนับสนุน และหนังสือ “การหาเอกลักษณ์ของระบบ” เล่มนี้เป็นหนังสือเล่มที่ 32 ในโครงการสนับสนุนการเขียนตำรา/หนังสือ/คู่มือ ของคณาจารย์และบุคลากรสายสนับสนุน คณะวิศวกรรมศาสตร์

ขอขอบคุณแบรนด์ KILTT จากบริษัทคูคิดดีไซน์ #kilttdesign ที่เอื้อเฟื้อชุดเฟอร์นิเจอร์ทำงานไม้โอ๊ค ให้ผู้เขียนใช้ทำงานที่บ้าน ตั้งแต่เริ่มมีสถานการณ์โควิด-19 เมื่อเดือนเมษายน พ.ศ. 2562 ทำให้ผู้เขียนมีเครื่องมือที่พร้อมในการเขียนหนังสือและทำงานวิจัยในตลอดเวลาที่ลาปฏิบัติหน้าที่ ขอขอบคุณดีไซน์เนอร์ของ KILTT และคุณอาทิมา อุทัยลานนท์ ที่ออกแบบกราฟิกของหนังสือเล่มนี้

สุดท้ายนี้ ขอขอบพระคุณ คุณพ่ออภิกิจ ส่งศิริ คุณแม่ละเมียด ส่งศิริ คุณพิมพ์ใจ ส่งศิริ และคุณกิตติภูมิ ส่งศิริ ที่เป็นกำลังใจซึ่งกันและกันในครอบครัว โดยเฉพาะคุณแม่ที่ช่วยเหลือจัดการธุระในบ้าน น้องสาว น้องชายและภรรยาที่ช่วยเกื้อกูลดูแลความเป็นอยู่และสุขภาพสมาชิกทุกคน ตลอดเวลาสถานการณ์โควิด-19 คุณพ่อที่ได้ช่วยตรวจสอบการเขียนภาษาไทยในหนังสือนี้ ครอบครัวได้เกื้อกูลให้ผู้เขียนได้มีเวลาทุ่มเทกับหนังสือเล่มนี้จนสำเร็จ



จิตโกมุต ส่งศิริ

ภาควิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

พฤษภาคม พ.ศ. 2565

คำนำ

ข้อมูลรอบตัวที่เติบโต. ในยุคที่ข้อมูลรอบตัวเข้าถึงได้ง่ายและมีปริมาณมากขึ้น การใช้ข้อมูลอันมหาศาลที่สนใจจากงานประยุกต์หนึ่งๆ ด้วยหลักการทางคณิตศาสตร์ จะเป็นเครื่องมือหลักในการตัดสินใจนโยบายสำคัญของงานประยุกต์นั้น การใช้ข้อมูลสัญญาณเข้า/ออกของระบบพลวัตเพื่อประมาณแบบจำลองระบบ ทำให้ใช้แบบจำลองนั้นในการออกแบบตัวควบคุม การเรียนรู้รูปแบบการใช้พลังงานของยูนิตต่างๆ ร่วมกับการพยากรณ์แหล่งจ่ายพลังงานสะอาด เพื่อจัดสรรกำลังการจ่ายพลังงานจากแหล่งจ่ายต่างๆ ในช่วงเวลาหนึ่งให้มีประสิทธิภาพมากขึ้น การใช้สัญญาณชีพ เช่นสัญญาณสมองของคนไข้ เพื่อวิเคราะห์ภาวะผิดปกติทางสมองเบื้องต้น ก็จะพัฒนาระบบคัดกรองให้มีความเป็นอัตโนมัติมากขึ้น และลดการใช้ทรัพยากรบุคคลไปได้มาก จากตัวอย่างเหล่านี้แสดงให้เห็นว่า ทฤษฎีด้านการประมาณแบบจำลอง การใช้เครื่องมือทางคณิตศาสตร์เพื่อเรียนรู้เชิงสถิติจากข้อมูล กลายเป็นสิ่งจำเป็นในหลายภาคส่วนของงานประยุกต์

ศาสตร์ที่เกี่ยวข้อง. การหาเอกลักษณ์ของระบบคือ เทคนิควิธีการประมาณแบบจำลองพลวัตเพื่อนำไปใช้อธิบายระบบกายภาพจริง มีวัตถุประสงค์ร่วมกันกับศาสตร์ด้านการเรียนรู้เชิงสถิติ ในแง่ที่เป็นการหาแบบจำลอง เพื่ออธิบายความสัมพันธ์ระหว่างตัวแปรต้นและตัวแปรตอบสนอง ด้วยการใช้ฟังก์ชันคณิตศาสตร์ ทั้งสองศาสตร์นี้ อาศัยรากฐานเดียวกันคือ ทฤษฎีการประมาณ เพื่อที่จะวิเคราะห์ข้อมูลด้วยหลักการทางสถิติและประมาณพารามิเตอร์ของแบบจำลองนั้น โดยส่วนใหญ่ ปัญหาการประมาณแบบจำลองมักจะอาศัยขั้นตอนเชิงตัวเลขที่ใช้แก้ปัญหาการหาค่าเหมาะที่สุด หนังสือเล่มนี้เน้นที่การใช้แบบจำลองพลวัต เชื่อมโยงกับหลักการเรียนรู้เชิงสถิติในด้านปัญหาถดถอย และให้พื้นฐานที่จำเป็นในด้านทฤษฎีการประมาณ เพื่อให้ผู้อ่านสามารถนำทักษะเหล่านี้ไปสร้างแบบจำลองจากข้อมูลที่ได้รับในงานประยุกต์ที่สนใจ ถึงแม้ว่าเทคนิคการหาค่าเหมาะที่สุดจะเป็นสิ่งที่จำเป็นอย่างยิ่ง แต่หนังสือเล่มนี้จะกล่าวได้พอสังเขปในแง่การใช้เท่านั้น

กลุ่มผู้อ่านที่คาดหวัง. หนังสือเล่มนี้สามารถใช้ประกอบการเรียนวิชาการหาเอกลักษณ์ของระบบ (System Identification) อันเป็นวิชาเลือกสำหรับนิสิตระดับปริญญาตรีปี 4 และนิสิตระดับบัณฑิตศึกษา ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เนื้อหาบางส่วนด้านทฤษฎีการประมาณ การถดถอย และการประมาณเชิงสถิติเป็นพื้นฐานสำคัญที่ใช้ในวิชาการอนุมานเชิงสถิติและแบบจำลอง (ภาควิชาวิศวกรรมไฟฟ้า) และวิชาเทคนิคเชิงสถิติสำหรับวิศวกรรมการเงิน (ของหลักสูตรวิศวกรรมการเงิน) เนื้อหานางานประยุกต์ของปัญหาที่ต่อยอดมาจากวิธีกำลังสองน้อยสุดเชิงเส้นได้มีกล่าวถึงในวิชา การหาค่าเหมาะที่สุดแบบคอนเวกซ์ (ภาควิชาวิศวกรรมไฟฟ้า) หนังสือเล่มนี้ได้เริ่ม

พัฒนาจากเอกสารคำสอนที่ได้ใช้มาตั้งแต่ปีพ.ศ. 2554 ซึ่งอาศัยพื้นฐานด้านพีชคณิตเชิงเส้น เมทริกซ์ ความน่าจะเป็นและสถิติ กระบวนการสุ่ม และระบบเชิงเส้น ผู้อ่านที่มีพื้นฐานเหล่านี้จึงติดตามเนื้อหาได้ไม่ยากนัก ผู้เขียนจัดเนื้อหาหลักการประมาณด้านทฤษฎี บทพิสูจน์ที่สำคัญ เพื่อให้มีความรัดกุมในเชิงคณิตศาสตร์สำหรับวิศวกร ส่วนผู้อ่านที่เน้นการใช้งานภาคปฏิบัติ จะมีตัวอย่างเชิงตัวเลขแสดงด้วยโปรแกรม MATLAB เพื่อจำลองข้อมูลด้วยคอมพิวเตอร์ ซึ่งความเชื่อมโยงของหลักการทฤษฎีกับการเรียกใช้งานชุดคำสั่ง และแสดงการแปลผล

เนื้อหาของหนังสือ แบ่งเป็น 3 ส่วนหลัก ได้แก่ หนึ่ง การแจกแจงแบบจำลองและการเลือกสัญญาณ กระตุ้นระบบ สอง เทคนิควิธีการประมาณแบบจำลองต่างๆ อันได้แก่ วิธีการกำลังสองน้อยสุดเชิงเส้นและปัญหาต่อขยาย วิธีปริภูมิย่อย วิธีความผิดพลาดในการทำนาย การประมาณภาวะน่าจะเป็นสูงสุด และการประมาณภายหลังสูงสุด จากนั้นส่วนสุดท้ายที่สาม จะกล่าวถึง การเลือกและตรวจสอบความสมเหตุสมผลของแบบจำลอง เนื้อหาเหล่านี้ได้นำมาจากหนังสือด้านการหาเอกลักษณ์ของระบบ หรือการเรียนรู้เชิงสถิติหลายเล่ม ตัวอย่างเช่น

- L. Ljung, *System Identification: Theory for the User*, Springer, 1999
- T. Söderström และ P. Stoica, *System Identification*, Prentice Hall, 1989
- T. Hastie and R. Tibshirani and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, Springer, 2009
- G. James, D. Witten, T. Hastie, และ R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*, Springer, 2013

เนื้อหาของแบบจำลองในบทที่ 2 จะเสนอกลุ่มแบบจำลองพลวัตเชิงเส้นเวลาวิฤตและประโยชน์เพื่อไปใช้ด้านการทำนายผลตอบ บทที่ 3 จะอธิบายถึงสัญญาณเข้าที่ดีเพื่อวัตถุประสงค์การหาเอกลักษณ์ระบบนั้น ต้องมีสมบัติกระตุ้นระบบที่เพียงพอ ซึ่งการเลือกสัญญาณเข้านี้จะแตกต่างกับวัตถุประสงค์เพื่อวิเคราะห์ผลตอบระบบ หรือเพื่อออกแบบตัวควบคุม ก่อนที่จะอธิบายเรื่องวิธีประมาณแบบจำลองต่างๆ บทที่ 4 จึงปูพื้นฐานด้านทฤษฎีการประมาณให้ทราบสมบัติของตัวประมาณที่ควรรู้ ได้แก่ ความไม่เอนเอียง ความคงเส้นคงวา และประสิทธิภาพ ผลลัพธ์สำคัญที่ใช้กันแพร่หลาย คือสมบัติการแจกแจงเชิงเส้นกำกับแบบปรกติของตัวประมาณเมื่อเรามีข้อมูลมาก ผู้อ่านอาจจะอ่านข้ามบทที่ 4 ก่อน หากต้องการศึกษาเรื่องขั้นตอนการประมาณแบบจำลองโดยตรง แต่การแปลผลเพื่อเปรียบเทียบกับสถิติเชิงสถิติของแบบจำลอง จำเป็นต้องมีพื้นฐานในบทที่ 4 ที่สามารถศึกษาได้ภายหลัง

บทที่ 5-9 กล่าวถึงรายละเอียดเทคนิคการประมาณต่างๆ เริ่มต้นด้วยวิธีพื้นฐานคือปัญหากำลังสองน้อยสุดเชิงเส้น เพราะในงานวิศวกรรม แบบจำลองเชิงเส้นยังมีการใช้งานอย่างแพร่หลาย บทที่ 5 จึงเริ่มด้วยการจัดรูปแบบปัญหา การหาผลเฉลย สมรรถนะของแบบจำลองถดถอยเชิงเส้น และแง่มุมที่สำคัญในการประยุกต์ใช้งาน การศึกษาในบท 5 นี้อาศัยพีชคณิตเชิงเส้นและการทดสอบสถิติ บทที่ 6 ได้ขยายต่อปัญหากำลังสองน้อยสุดเชิงเส้นที่มีเงื่อนไขบังคับ ฟังก์ชันลงโทษ หรือมีพารามิเตอร์ของรูปแบบปัญหาที่มีความไม่แน่นอน รวมถึงปัญหากำลังสองน้อยสุดไม่เชิงเส้น การอธิบายเนื้อหาในบทนี้ จะเน้นการจัดรูปแบบปัญหาการประมาณ ให้อยู่ในรูปปัญหาการหาค่าเหมาะที่สุดที่ส่วนใหญ่จะเป็นกลุ่มปัญหากำลังสอง (หรือไม่ก็เป็นปัญหาไม่เชิงเส้น) การแก้ปัญหาลักษณะเชิงตัวเลขจึงเป็นการแนะนำชุดคำสั่งใน MATLAB จากนั้น เพื่อความต่อเนื่องของการใช้พื้นฐานของวิธีการกำลังสองน้อยสุด จึงเสนอวิธีปริภูมิย่อยในบทที่ 7 ถัดมาอันเป็นวิธีที่ใช้ประมาณแบบจำลองปริภูมิสถานะเชิงเส้นที่ประมาณทั้งตัวแปรสถานะ

และเมทริกซ์ของระบบ จากข้อมูลสัญญาณเข้า/ออกเท่านั้น การหาคำตอบเชิงตัวเลขใช้วิธีแบบไม่ทำซ้ำ เทคนิคของวิธีนี้อาศัยพื้นฐานด้านพีชคณิตเชิงเส้น การฉายภาพเชิงตั้งฉากและการแยกตัวประกอบของเมทริกซ์ บทที่ 8 เสนอวิธีความผิดพลาดในการทำนาย ที่สามารถประยุกต์ใช้กับแบบจำลองเชิงเส้นในรูปทั่วไป ไม่ว่าจะเป็นรูปฟังก์ชันถ่ายโอน หรือปริภูมิสถานะ หลักการของวิธีนี้คือการใช้ฟังก์ชันสูญเสียที่แบ่งค่าคลาดเคลื่อนของแบบจำลอง พารามิเตอร์ของแบบจำลองจึงถูกเลือกเพื่อให้ฟังก์ชันสูญเสียมีค่าต่ำสุด จัดเป็นปัญหาการหาค่าเหมาะที่สุดและการหาผลเฉลยเชิงตัวเลขจึงใช้ขั้นตอนวิธีทำซ้ำ หลักการทางทฤษฎีของวิธีนี้มีความเป็นทั่วไปสูง โดยเฉพาะในด้านสมบัติเชิงเส้นกำกับของตัวประมาณที่อาศัยพื้นฐานจากบทที่ 4 ผู้เขียนจึงได้ยกตัวอย่างในกรณีเฉพาะเพื่ออธิบายประกอบให้เกิดความเข้าใจได้ง่าย

วิธีประมาณในบทที่ 5-8 อยู่ในกลุ่มที่ไม่ได้อาศัยสมบัติเชิงสถิติของตัวแปรในแบบจำลองในขั้นตอนการประมาณ (แต่เราจำเป็นต้องใช้สมบัติเชิงสถิติในการวิเคราะห์สมรรถนะของแบบจำลอง) บทที่ 9 จะนำเสนอวิธีในกลุ่มที่เรียกว่า การประมาณเชิงสถิติ เพราะว่าแบบจำลองที่สมมติขึ้นจะมีตัวแปร หรือมีพารามิเตอร์ที่สมมติให้มีความไม่แน่นอน การทราบถึงการแจกแจง หรือโมเมนต์บางอันดับของตัวแปรเหล่านั้น จึงนำมาใช้เป็นสมมติฐานส่วนหนึ่งในการจัดรูปแบบปัญหาประมาณ เทคนิควิธีในกลุ่มเหล่านี้คือ การประมาณความคลาดเคลื่อนกำลังสองเฉลี่ยน้อยสุด ภาวะน่าจะเป็นสูงสุด และภายหลังสูงสุด อันเป็นวิธีพื้นฐานในหลายศาสตร์ ทั้งในการเรียนรู้เชิงสถิติ หรือการเรียนรู้ของเครื่อง หลักการของวิธีกลุ่มนี้อาศัยความรู้ทางตัวแปรสุ่มและสถิติ สมบัติเชิงสถิติของค่าประมาณพารามิเตอร์ที่ประมาณได้ขึ้นกับจำนวนตัวอย่างข้อมูลและสมมติฐานที่ใช้ การศึกษาสมบัติเชิงเส้นกำกับของตัวประมาณจะใช้เนื้อหาพื้นฐานทฤษฎีการประมาณจากบทที่ 4

เมื่อผู้อ่านได้เข้าใจหลักการประมาณแบบจำลองหลายวิธีแล้ว บทที่ 10 จึงเสนอประเด็นที่สำคัญเพื่อเลือกแบบจำลองจากหลายแบบในกลุ่มตัวเลือก หลายแบบจำลองนั้นอาจแปรจากกลุ่มโครงสร้างแบบจำลองต่างชนิดกัน หรือแปรจากอันดับของแบบจำลองในโครงสร้างหนึ่งๆ กระบวนการเลือกแบบจำลองอิงกับทฤษฎีสำคัญทางสถิติ คือการค่านกันระหว่างความเอนเอียงและความแปรปรวนของแบบจำลอง แบบจำลองอย่างง่ายหรือแบบซับซ้อน จะมีทั้งข้อดีและข้อเสียในสองมุมดังกล่าว ผู้เขียนกล่าวถึงกระบวนการเลือก 2 วิธีได้แก่ การใช้เกณฑ์การเลือกแบบจำลองและการตรวจสอบไขว้ อันเป็นวิธีพื้นฐานในศาสตร์ด้านการเรียนรู้เชิงสถิติ เมื่อได้แบบจำลองที่เลือกแล้ว การตรวจสอบแบบจำลองเป็นขั้นตอนสุดท้ายที่จะทดสอบว่าแบบจำลองที่เลือกนั้น นำไปใช้อธิบายข้อมูลชุดใหม่ได้ดีเพียงใด

ผู้เขียนได้คัดเลือกตัวอย่างงานวิจัยของการหาเอกลักษณ์ระบบในด้านการอนุมานพารามิเตอร์ของแบบจำลอง เพื่ออธิบายความสัมพันธ์ของตัวแปรที่มีจำนวนมากในระบบซับซ้อน บทที่ 11 เสนอการใช้แบบจำลองแบบเบาบาง (sparse models) และเทคนิควิธีการประมาณที่ใช้ฟังก์ชันนอร์ม-1 เป็นส่วนหนึ่งของฟังก์ชันสูญเสียของรูปแบบปัญหาประมาณ เนื่องจากการใช้นอร์ม-1 มีสมบัติสนับสนุนให้พารามิเตอร์ของแบบจำลองมีค่าศูนย์เป็นจำนวนมาก งานวิจัยในกลุ่มนี้เป็นที่ใช้แพร่หลายในการเรียนรู้เชิงสถิติและการเรียนรู้ของเครื่องจนถึงปัจจุบัน และมีการพัฒนาอย่างมากในด้านขั้นตอนวิธีเชิงเลขที่ใช้แก้ปัญหาที่เหมาะสมที่สุดขนาดใหญ่ ผู้เขียนได้เสนองานประยุกต์ด้านการเรียนรู้ความเชื่อมโยงการทำงานของสมองในบริเวณต่างๆ จากหลายรูปแบบปัญหาประมาณด้วยกัน

โครงสร้างของหนังสือจะเริ่มจากการอธิบายหลักการและทฤษฎีบทที่สำคัญ บทพิสูจน์ที่ไม่ยากต่อการทำความเข้าใจจะนำเสนอหลังจากทฤษฎีบท จากนั้นจะแสดงตัวอย่างการประยุกต์เพื่อจัดรูปแบบปัญหา และการคำนวณเชิงตัวเลขที่อิงจากการใช้ชุดคำสั่ง MATLAB ด้วยกล่องเครื่องมือด้านสถิติ และการหาเอกลักษณ์ของระบบ ท้ายบทจะมีแบบฝึกหัดให้ผู้อ่านได้ทบทวนความเข้าใจ บางแบบฝึกหัดจะเป็นการพิสูจน์ผลลัพธ์ของขั้นตอนระหว่างที่อธิบายในหนังสือ บางแบบฝึกหัดเป็นการทดลองของการ

จำลองด้วยคอมพิวเตอร์จากข้อมูลที่ให้ อย่างที่ได้กล่าวไว้ว่าเนื้อหาของหนังสืออาศัยพื้นฐานด้านระบบเชิงเส้น เมทริกซ์ กระบวนการสุ่ม ตัวแปรสุ่ม และสถิติ ผลลัพธ์พื้นฐานเฉพาะที่นำมาใช้ในหนังสือเล่มนี้ จึงได้รวบรวมไว้ในภาคผนวก ผู้อ่านสามารถศึกษาเพิ่มเติมได้จากตำราพื้นฐานทั่วไปที่มีแนะนำไว้ในรายการอ้างอิง

โปรแกรมและเว็บไซต์. เนื้อหาที่เกี่ยวข้องกับการคำนวณด้วยคอมพิวเตอร์ จะมีตัวอย่างการใช้ฟังก์ชันใน MATLAB เพื่อให้ผู้อ่านได้ทดลองเขียนโปรแกรมประกอบ โดยการใช้กล่องเครื่องมือ System Identification, Statistics and Machine Learning, และบางส่วนจาก Econometric ผู้อ่านสามารถดาวน์โหลดชุดข้อมูลเพื่อทำการทดลอง สไลด์ประกอบการเรียน และวิดีโอการบรรยายเนื้อหาได้ที่

http://jitkomut.eng.chula.ac.th/book_sysiden.html

สิ่งที่หนังสือนี้ไม่ครอบคลุม. การหาเอกลักษณ์ของระบบด้วยวิธีไม่อิงแบบจำลอง เช่น การวิเคราะห์สหสัมพันธ์ หรือ ผลตอบสนองเชิงความถี่ การวิเคราะห์ฟูเรียร์หรือสเปกตรัม เทคนิคการเรียนรู้สถิติ ปัญหาด้านการจำแนก หรือการใช้แบบจำลองอื่นในศาสตร์การเรียนรู้ด้วยเครื่อง จะไม่ได้ถูกนำเสนอในตำรานี้ เทคนิคการวิเคราะห์ข้อมูล เช่น การแปลงข้อมูลเป็นโดเมนอื่น การสกัดคุณลักษณะ การลดอันดับของตัวแปร ก่อนนำข้อมูลเข้าไปใช้ในแบบจำลอง ก็เป็นสิ่งสำคัญเช่นเดียวกัน แต่เนื้อหาเหล่านี้สามารถศึกษาได้จากตำราด้านการเรียนรู้เชิงสถิติ หรือการเรียนรู้ของเครื่องเพิ่มเติม ศาสตร์ในด้านการประมาณระบบเกี่ยวข้องอย่างมากกับการใช้วิธีเชิงตัวเลขเพื่อแก้ปัญหาค่าเหมาะที่สุด หนังสือจะกล่าวถึงมุมมองจากขั้นตอนเชิงตัวเลขโดยย่อเท่านั้น นอกจากนี้ เราจะอธิบายวิธีประมาณแบบจำลองพลวัตด้วยชุดข้อมูลโดเมนเวลาเป็นหลัก ถึงแม้ว่าการหาเอกลักษณ์ของระบบจะใช้ข้อมูลโดเมนความถี่ได้เช่นกัน

ผู้เขียนหวังว่าหนังสือเล่มนี้จะเป็นประโยชน์ต่อผู้อ่านในด้านพื้นฐานการประมาณแบบจำลอง และการทำวิจัยในงานประยุกต์ต่อไป

จิตโกมุท ส่งศิริ

ภาควิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

พฤษภาคม พ.ศ. 2565

สารบัญ

กิตติกรรมประกาศ	v
คำนำ	vii
สารบัญ	xi
สารบัญรูป	xiv
สารบัญทฤษฎีบท	xviii
สารบัญรหัสคำสั่ง	xx
สัญลักษณ์และตัวแปร	xxi
อักษรย่อ	xxiv
1 บทนำ	1
1.1 ความสำคัญของการหาเอกลักษณ์ของระบบ	2
1.2 องค์ประกอบที่สำคัญ	4
1.3 ขั้นตอนการหาเอกลักษณ์ของระบบ	7
1.4 ตัวอย่างเบื้องต้น	8
1.5 บทสรุป	9
2 แบบจำลอง	13
2.1 การจำแนกแบบจำลอง	13
2.2 ฟังก์ชันและตัวดำเนินการถ่ายโอน	15
2.3 โครงสร้างแบบจำลองทั่วไป	16
2.4 การใช้แบบจำลอง	17
2.5 ตัวอย่างแบบจำลอง	23
2.6 เงื่อนไขของแบบจำลอง	32
2.7 รูปแบบของตัวแปรแบบจำลองใน MATLAB	32
แบบฝึกหัด	36
3 สัญญาณเข้าเพื่อการหาเอกลักษณ์ระบบ	37
3.1 สัญญาณเข้าและคุณสมบัติ	39

3.2	ระดับการกระตุ้นคงอยู่	49
3.3	การเลือกสัญญาณเข้าสำหรับการทดลองวงเปิด	55
3.4	ตัวอย่างคำสั่ง MATLAB	59
	แบบฝึกหัด	60
4	พื้นฐานทฤษฎีการประมาณ	63
4.1	สมบัติของตัวประมาณ	63
4.2	สมบัติของแบบจำลอง	66
4.3	ตัวประมาณสุดขีด (Extremum estimator)	68
	แบบฝึกหัด	72
5	การประมาณกำลังสองน้อยสุดเชิงเส้น	75
5.1	รูปแบบปัญหาและตัวอย่างงานประยุกต์	76
5.2	การหาคำตอบเชิงตัวเลข	81
5.3	ปัญหากำลังสองน้อยสุดถ่วงน้ำหนัก	83
5.4	สมรรถนะของแบบจำลองถดถอยเชิงเส้น	84
5.5	คุณสมบัติเชิงสถิติ	85
5.6	การทดสอบนัยสำคัญของตัวประมาณ	88
5.7	การเลือกตัวแปรทำนายที่สำคัญต่อตัวแปรตอบสนอง	91
5.8	ประเด็นสำคัญเชิงปฏิบัติ	94
	แบบฝึกหัด	97
6	การขยายปัญหากำลังสองน้อยสุดเชิงเส้น	101
6.1	ปัญหากำลังสองน้อยสุดเชิงเส้นที่มีเงื่อนไขบังคับ	102
6.2	ปัญหากำลังสองน้อยสุดเชิงเส้นที่มีฟังก์ชันลงโทษ	105
6.3	ปัญหากำลังสองน้อยสุดเชิงเส้นเมื่อไม่ทราบค่าพารามิเตอร์	111
6.4	วิธีกำลังสองน้อยสุดไม่เชิงเส้น	117
6.5	ตัวอย่างงานประยุกต์	120
6.6	เครื่องมือในการแก้ปัญหาค่าเหมาะที่สุด	124
	แบบฝึกหัด	129
7	วิธีปริภูมิย่อย	135
7.1	ภาพรวมของวิธีปริภูมิย่อย	135
7.2	ความสัมพันธ์ของสัญญาณเข้าและสัญญาณออก	137
7.3	การกำจัดผลจากสัญญาณเข้า	139
7.4	การกำจัดผลจากสัญญาณรบกวน	139
7.5	การประมาณ A และ C	141
7.6	การประมาณ B และ D	144
7.7	วิธีปริภูมิย่อยในมุมมองของปัญหา LS	144
7.8	ตัวอย่างคำสั่ง MATLAB	149
7.9	ประเด็นต่อยอดและงานประยุกต์	153
	แบบฝึกหัด	156
8	วิธีความผิดพลาดในการทำนาย	157
8.1	องค์ประกอบของตัวทำนาย	157
8.2	การทำนายแบบเหมาะสมที่สุด	160
8.3	คุณสมบัติของตัวประมาณ PEM	162

8.4	ประเด็นการคำนวณเชิงตัวเลข	169
8.5	ตัวอย่างชุดคำสั่งจาก MATLAB	171
	แบบฝึกหัด	176
9	การประมาณเชิงสถิติ	179
9.1	การประมาณความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด	179
9.2	การประมาณภาวะน่าจะเป็นสูงสุด	184
9.3	การประมาณภายหลังสูงสุด	193
9.4	สมบัติของตัวประมาณ	196
	แบบฝึกหัด	202
10	การเลือกและตรวจสอบแบบจำลอง	207
10.1	ความเอนเอียงและแปรปรวนของแบบจำลอง	209
10.2	การเลือกแบบจำลองเหมาะสมที่สุด	214
10.3	การตรวจสอบแบบจำลอง	227
10.4	ตัวอย่างคำสั่ง MATLAB	232
	แบบฝึกหัด	235
11	งานประยุกต์ด้านการเรียนรู้ความสัมพันธ์ของตัวแปรในระบบ	239
11.1	โครงสร้างระบบซับซ้อน	239
11.2	ความสัมพันธ์ของตัวแปรนิยามจากอะไร?	240
11.3	การประยุกต์ใช้ฟังก์ชันลงโทษแบบนอร์ม-1	244
11.4	แบบจำลองเชิงกราฟของตัวแปรเกาส์เซียน	244
11.5	การจำลองแบบสมการเชิงโครงสร้าง	246
11.6	แบบจำลองเวกเตอร์ถดถอยในตัว	248
11.7	กลุ่มแบบจำลองเวกเตอร์ถดถอยในตัว	252
11.8	ข้อคิดเห็นและบทสรุป	257
	รายการภาคผนวก	261
A	แคลคูลัส	263
A.1	อนุพันธ์ของฟังก์ชันหลายตัวแปร	263
A.2	ฟังก์ชันของเมทริกซ์	264
B	เมทริกซ์	267
C	ตัวแปรสุ่มและกระบวนการสุ่ม	273
C.1	การสุ่มของตัวแปรสุ่ม	273
C.2	ตัวแปรเกาส์เซียน	277
C.3	กระบวนการสุ่ม	279
C.4	ความหนาแน่นเชิงสเปกตรัมกำลังของสัญญาณรายคาบ	281
	บรรณานุกรม	283
	อภิธานศัพท์	291
	ดัชนี	297
	ประวัติผู้แต่ง	303

สารบัญรูป

1.1	ความสัมพันธ์ของสัญญาณเข้ากับสัญญาณออกผ่านระบบที่มีสัญญาณรบกวน	1
1.2	ตัวอย่างการใช้แบบจำลองที่ประมาณได้ในการทำนายค่าสัญญาณออก y	3
1.3	การหาเอกลักษณ์ของระบบเพื่อนำไปใช้ร่วมกับการออกแบบตัวควบคุม	3
1.4	ความรู้เกี่ยวกับสมบัติของสัญญาณรบกวนที่นำไปใช้ร่วมกับการประมาณแบบจำลอง	6
1.5	ขั้นตอนการหาเอกลักษณ์ของระบบ	8
1.6	ตัวอย่างการประมาณจุดข้อมูลด้วยฟังก์ชันพหุนามอันดับต่างๆ	8
1.7	ผลตอบสนอง y ต่อ u ในระบบพลวัต	9
1.8	ผลการทดสอบแบบจำลองบนข้อมูลชุดตรวจสอบ สำหรับอันดับแบบจำลองต่างๆ	10
1.9	ตำแหน่งของศูนย์ (o) และโพล (x) ในแบบจำลอง ARMAX ที่ประมาณได้	10
2.1	ตัวอย่างผลตอบสนองจากระบบ $y = Gu + He$ โดยกราฟเส้นหนา คือ ผลตอบ Gu และกราฟเส้นประคือผลตอบที่มาจาก He ทั้ง 10 ตัวอย่าง	18
2.2	ตัวอย่างการทำนาย 1 ก้าวล่วงหน้า ด้วยการใช้แบบจำลอง ARMAX	21
2.3	อนุกรมเวลาที่มีผลของฤดูกาล ค่าเฉลี่ยของกระบวนการนี้เป็นฟังก์ชันเป็นคาบ	30
2.4	กระบวนการ AR(1) และกระบวนการผลรวมอันดับต่างๆ	30
2.5	อนุกรมเวลาของ log CPI เมื่อผ่านการแปลงหาผลต่าง	31
2.6	ระบบผกผัน ระบบต่อเรียง และระบบต่อขนาน	35
3.1	ลักษณะของสัญญาณออกของระบบ $G(s)$ ที่มีต่อสัญญาณเข้าชนิดต่างๆ	38
3.2	การหาค่าพารามิเตอร์ของแบบจำลองอันดับหนึ่งที่มีค่าประวิงเวลาจากกราฟผลตอบสนองของกระบวนการเมื่อเข้าสู่ภาวะอยู่ตัว	40
3.3	สัญญาณไซน์หลายความถี่: $u(t) = \sin(0.4t) + 2\sin(0.8t) + 3\sin(2t)$ สำหรับ $t = 0, 1, 2, \dots, 1000$	43
3.4	สัญญาณรบกวนขาวเกาส์เซียนที่มีความแปรปรวนเป็นหนึ่ง	43
3.5	สัญญาณ ARMA ที่สร้างจากกระบวนการ ARMA 1: $G(z) = \frac{1+0.3z^{-1}}{1-0.7z^{-1}-0.2z^{-2}}$ ARMA 2: $G(z) = \frac{1-0.5z^{-1}}{1-1.4z^{-1}+0.8z^{-2}}$ และ ARMA 3: $G(z) = 1 - 1.4z^{-1} + 0.8z^{-2}$ และใช้สัญญาณรบกวนขาวมีความแปรปรวนเท่ากับหนึ่ง	44
3.6	แผนผังการสร้างลำดับฐานสองสุ่มเทียม (PRBS)	45
3.7	สัญญาณ PRBS ที่ใช้ $n = 5$ มีคาบเท่ากับ $M = 31$ และมีขนาด ± 10	48
3.8	ฟังก์ชันสหสัมพันธ์ของ $y(t)$ เมื่อกระตุ้นด้วยสัญญาณเข้า PRBS และสัญญาณรบกวนขาว ผ่านระบบอันดับหนึ่ง $G(z) = 1/(1 - az^{-1})$ เมื่อ $a = 0.8$	50

3.9	ตัวอย่าง $n = 5$ ที่มี $H_1(z) = 2z^{-1} - 7z^{-2} + 8z^{-3} - 4.5z^{-4} + z^{-5}$, $H_2(z) = 10z^{-1} - 18.76z^{-2} + 18.97z^{-3} - 7.3z^{-4} + z^{-5}$, $H_3(z) = z^{-1} - 3.21z^{-2} + 4.53z^{-3} - 3.21z^{-4} + z^{-5}$ ทั้งสามพหุนามนี้มีจำนวนศูนย์ของ $H(z)H(z^{-1})$ ที่อยู่บนวงกลมหนึ่งหน่วยที่ต่างกัน	51
3.10	สัญญาณไซน์ความถี่เดียวที่ 11 Hz จำนวน 58 ตัวอย่างที่ให้สมรรถนะประมาณแบบจำลองค่าใกล้เคียงกันที่ 76% แต่ไม่สามารถแยกแยะแบบจำลองแต่ละแบบได้ <i>แฉวบน</i> แผนภาพโบเดของ $\hat{G}$ <i>แฉวกลาง</i> ตำแหน่งศูนย์และสัมประสิทธิ์ของพหุนามเศษของ $\hat{G}$ โดยมีสีที่ต่างกันสำหรับแต่ละค่าศูนย์และสัมประสิทธิ์เศษ <i>แฉวล่าง</i> ตำแหน่งโพลและสัมประสิทธิ์ของพหุนามส่วนของ $\hat{G}$ โดยมีสีที่ต่างกันสำหรับแต่ละค่าโพลและสัมประสิทธิ์ส่วน	55
3.11	สัญญาณไซน์หลายความถี่ (multisine)	57
3.12	ความไวของแบบจำลองต่อการเปลี่ยนแปลงของพารามิเตอร์และผลการประมาณ	58
5.1	ผลเฉลยกำลังสองน้อยสุดทำให้ $X\beta_{ls}$ ใกล้เวกเตอร์ y มากที่สุดในนอร์ม-2 และตั้งฉากกับ e	78
5.2	การประมาณฟังก์ชันพหุนามอันดับ 6 เข้ากับชุดข้อมูล	79
5.3	การใช้ตัวประมาณกำลังสองน้อยสุดถ่วงน้ำหนักเพื่อหาเส้นตรงที่ประมาณจุดข้อมูล	83
5.4	การเลือกเซตย่อยที่ดีที่สุดของตัวแปรทำนายในปัญหาถดถอย	93
5.5	การเลือกตัวแปรทำนายแบบเดินหน้าและถอยหลังในปัญหาถดถอย	93
5.6	ลักษณะข้อมูลที่ไม่เป็นไปตามสมมติฐานของปัญหาถดถอยเชิงเส้น	96
5.7	ค่าความเข้มแสงอาทิตย์และกำลังไฟฟ้าที่ผลิตได้ในวันที่ท้องฟ้าโปร่งใสและมีเมฆมาก	98
6.1	ตัวอย่างรูปร่างของเซต C แบบต่างๆ ใน $\mathbf{R}^2$ อันเป็นเซตเงื่อนไขของปัญหา (6.1) และภาพของ C ภายใต้การแปลง A	103
6.2	ตัวอย่างรูปแบบผลเฉลยเมื่อมีเทอมลงโทษ $ x $	106
6.3	<i>แฉวบน</i> ฮิสโตแกรมของผลเฉลยจากปัญหา LS ที่ใช้ฟังก์ชันลงโทษนอร์ม-2 และนอร์ม-1 <i>แฉวกลาง</i> ทางเดินของผลเฉลยปัญหา LS ที่ใช้ฟังก์ชันลงโทษนอร์ม-2 และนอร์ม-1 เมื่อแปรค่าพารามิเตอร์ลงโทษ (แต่ละสีคือแต่ละค่า $x_i(\gamma)$) <i>แฉวล่าง</i> ปัญหา (6.10) เมื่อมีเงื่อนไขบังคับเป็น ℓ_2 -นอร์มบอล และ ℓ_1 -นอร์มบอล โดยที่คอนทัวร์ของ $\ Ax - y\ _2^2$ ที่ค่า (A, y) สองชุดแสดงด้วยสีฟ้า และสีม่วง	107
6.4	ทางเดินผลเฉลยปัญหาตาข่ายยืดหยุ่นเมื่อใช้ $\alpha = 0.1$ (ซ้าย) และ $\alpha = 1$ คือผลเฉลยแลสโซ (ขวา)	109
6.5	(ซ้าย) รูปแบบศูนย์ของผลเฉลยแลสโซกลุ่ม ในแนวตั้งที่ค่า γ หนึ่ง จุดสี่เดียวกัน 4 จุดหมายถึง $x_i \in \mathbf{R}^4$ ที่ไม่เท่ากับศูนย์ทั้งบล็อก และมีจำนวน 5 บล็อก เมื่อ γ มากขึ้น จำนวนบล็อกใน x ที่เป็นศูนย์มีมากขึ้น (ขวา) ฟังก์ชันลงโทษแบบประกอบนอร์ม-1 กับนอร์ม-2 เมื่อแปรค่า γ	110
6.6	ทางเดินผลเฉลยปัญหาแลสโซเชื่อม เมื่อ $\gamma_1 = 0$ (ซ้าย) และเมื่อใช้ $\gamma_1 = 2$ (ขวา)	111
6.7	สมรรถนะของผลเฉลยต่างๆ ในปัญหา LS เมื่อ $A \in \{A_1, A_2, A_3\}$	112
6.8	นอร์ม-2 ของความคลาดเคลื่อนเลวร้ายที่สุด (6.20) เมื่อค่า x ที่ใช้เป็นค่าต่างๆ	115
6.9	ผลเปรียบเทียบนอร์ม-2 ของค่าคลาดเคลื่อนเมื่อใช้ผลเฉลยกำลังสองน้อยสุดทั้ง 3 แบบ	117

6.10	ผลการประมาณฟังก์ชันเกาส์เซียนด้วยวิธีกำลังสองน้อยสุดไม่เชิงเส้น โดยเปรียบเทียบข้อมูลวัด (ที่มีสัญญาณรบกวน) และเปรียบเทียบระหว่างฟังก์ชันที่ใช้ค่าพารามิเตอร์จริงกับค่าพารามิเตอร์ที่ประมาณได้	119
6.11	การจำลองแหล่งกำเนิดสัญญาณคลื่นไฟฟ้าสมอง	120
7.1	ภาพรวมการหาเอกลักษณ์ของระบบด้วยวิธีปริภูมิย่อย	136
7.2	การประมาณแบบจำลองปริภูมิสถานะด้วยวิธี n4sid	150
7.3	การประมาณระบบมวล-สปริงด้วยแบบจำลองปริภูมิสถานะที่กำหนดโครงสร้างบางส่วน	152
8.1	ผลการประมาณระบบ ARMAX ด้วยการใช้แบบจำลอง ARMA และ ARX	172
8.2	ผลการประมาณข้อมูลของมอเตอร์กระแสตรงด้วยแบบจำลองปริภูมิสถานะและ ARX	174
8.3	การประมาณแบบจำลอง ARMA ด้วยชุดคำสั่งจากกล่องเครื่องมือ Econometric และ System Identification	175
9.1	ตัวอย่างฟังก์ชันแจกแจงความน่าจะเป็นสำหรับพารามิเตอร์ค่าต่างๆ	185
9.2	ค่าพารามิเตอร์ที่เลือกในปัญหา การประมาณภาวะน่าจะเป็นสูงสุด Maximum likelihood estimation (MLE) คือค่าที่ทำให้ $\mathcal{L}(\theta)$ มีค่าสูงสุด	188
9.3	ฟังก์ชันความหนาแน่นก่อนและภายหลัง	194
9.4	ฟังก์ชันความหนาแน่นของตัวแปรเกาส์เซียน (ซ้าย) และลาปลาซ (ขวา)	195
10.1	ความยืดหยุ่นและความสามารถในการตีความของแบบจำลอง [HTF09]	212
10.2	การประมาณแบบจำลอง AR เมื่อแปรอันดับของแบบจำลอง ค่าฟังก์ชันสูญเสียจุดหักที่ประมาณ $p = 5$	212
10.3	ค่าคลาดเคลื่อนของแบบจำลองเมื่อประเมินบนข้อมูลชุดประมาณและชุดทดสอบ	213
10.4	ภาวะเข้ากันเกินเมื่อ $f(x)$ มีความไม่เชิงเส้นสูง และ $f(x)$ ใกล้เคียงกับฟังก์ชันเชิงเส้น	214
10.5	ผลการเลือกตัวแปรทำนายของแบบจำลองถดถอยเชิงเส้นด้วยเกณฑ์การเลือกแบบจำลองเมื่อใช้ $N = 50$ เกณฑ์ AIC AICc BIC และ C_p เลือกจำนวนตัวแปรทำนายเท่ากับ 15, 4, 4 และ 7 ตามลำดับ Adjusted R^2 เลือก 16 ตัวแปร และ R^2 เลือกใช้ 20 ตัวแปร	220
10.6	ผลการเลือกอันดับของแบบจำลอง AR ด้วยเกณฑ์การเลือกแบบจำลอง	221
10.7	การแบ่งข้อมูลสำหรับการตรวจสอบไขว้	225
10.8	การตรวจสอบแบบจำลองพหุนามด้วยวิธี 1) การใช้ข้อมูลตรวจสอบ (ที่มีการทำซ้ำ 7 ครั้ง) 2) การตรวจสอบไขว้แบบ 5 ส่วน (ที่มีการทำซ้ำ 7 ครั้ง) และ 3) การตรวจสอบไขว้แบบ 1 ส่วน	226
10.9	การเลือกพารามิเตอร์ลงโทษของปัญหาแลสโซ ในตัวอย่างนี้ที่ใช้จำนวนข้อมูล $N = 200$ วิธี k -fold และ LOOCV เลือกค่า γ ที่ใกล้เคียงกัน เกณฑ์ AIC และ AICc เลือกค่า γ ที่เท่ากัน ส่วน BIC เลือก γ ที่สูงขึ้น ส่งผลให้ผลเฉลยมีศูนย์มากขึ้น	227
10.10	ผลการวิเคราะห์ค่าส่วนเหลือด้วยการทดสอบความขาวและฟังก์ชันสหสัมพันธ์ไขว้	231
10.11	ตำแหน่งของโพล (x) และศูนย์ (o) ของแบบจำลองอันดับ 3 และ 6 ด้วยคำสั่ง iopzplot วรสีจึงหมายถึงบริเวณความเชื่อมั่นของค่าประมาณโพลและศูนย์	231
11.1	โครงสร้างความสัมพันธ์เชิงกราฟของตัวแปรในระบบขับเคลื่อน	240
11.2	ตัวอย่างงานประยุกต์การแสดงความสัมพันธ์ของตัวแปรในระบบที่สนใจ	241
11.3	การนำผลประมาณแบบจำลองเพื่ออนุมานโครงสร้างความสัมพันธ์ของตัวแปรในระบบ	243

11.4	รูปแบบศูนย์ในอินเวอร์สของเมทริกซ์ความแปรปรวนร่วมของตัวแปรเกาส์เซียน: ค่าที่แท้จริง, ค่าที่คำนวณจากตัวอย่าง, ผลเฉลยของปัญหาแลสโซเชิงกราฟ เมื่อใช้ค่าพารามิเตอร์ลงโทษเท่ากับ $\gamma = 0.02$ และ $\gamma = 0.15$ ตามลำดับ	245
11.5	โครงสร้างร่วมของความสัมพันธ์ของตัวแปรในระบบมลภาวะจากสถานีต่างๆ เมื่อตีความความคล้ายเปลี่ยน 3 ระดับ (ค่ามากหมายถึงโครงสร้างดังกล่าวพบมากในหลายสถานี) หมายเหตุ: รูปนี้ได้รับอนุญาตให้พิมพ์ซ้ำจาก Elsevier วารสาร Signal Processing ฉบับที่ 166 ปีค.ศ. 2020 ผู้แต่ง A. Pruttiakaranich และ J. Songsiri บทความเรื่อง Convex Formulation for Regularized Estimation of Structural Equation Models	248
11.6	เครือข่ายสมองของกลุ่มผู้ป่วยออทิสติกและกลุ่มปกติ ลิงก์สีคือการเชื่อมโยงที่พบในทั้ง 2 กลุ่ม ลิงก์สีแดง คือการเชื่อมโยงที่พบเฉพาะ กลุ่มออทิสติก และ ลิงก์สีน้ำเงิน คือการเชื่อมโยงที่พบเฉพาะใน กลุ่มปกติ หมายเหตุ: ได้รับอนุญาตให้พิมพ์ซ้ำจาก Elsevier วารสาร Signal Processing ฉบับที่ 166 ปีค.ศ. 2020 ผู้แต่ง A. Pruttiakaranich และ J. Songsiri บทความเรื่อง Convex Formulation for Regularized Estimation of Structural Equation Models	249
11.7	ตัวอย่างรูปแบบศูนย์ในเมทริกซ์สัมประสิทธิ์ของกระบวนการถดถอยในตัวที่ประมาณจากปัญหา sparse VAR (11.12)	250
11.8	ข้อมูลแนวโน้มเคสไข้หวัดของกูเกิล (Google Flu Trends) หมายเหตุ: ได้รับอนุญาตให้พิมพ์ซ้ำจาก IEEE บทความที่ประชุมวิชาการ 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) ปีค.ศ. 2013 ผู้แต่ง Jitkomut Songsiri เรื่อง Sparse autoregressive model estimation for learning Granger causality in time series	251
11.9	โครงสร้างเชื่อมโยงการทำงานของสมองที่เรียนรู้จากข้อมูลสมอง 2 กลุ่ม	253
11.10	การจัดเรียงพารามิเตอร์ของ K แบบจำลองเพื่อออกแบบฟังก์ชันลงโทษ g	253
11.11	ตัวอย่างแบบจำลอง VAR 2 แบบที่ให้โครงสร้างเชิงกราฟแบบเกรนเจอร์เหมือนกัน (CGN) หรือต่างกันบางส่วน (DGN, FGN) เมทริกซ์ขนาด $n \times n$ มีกริด (i, j) ที่แสดงความเข้มของ $\ B_{ij}\ _2$	254
11.12	การเชื่อมโยงการทำงานของสมองเมื่อใช้ปัญหาการประมาณกลุ่มแบบจำลองกับข้อมูล ADHD และเลือกแสดงลิงก์สำคัญจากบริเวณ orbitofrontal และระบบลิมบิก หมายเหตุ: ได้รับอนุญาตให้รูป/ทำซ้ำ/ดัดแปลงจาก Elsevier วารสาร Neural Networks ฉบับที่ 149 ปีค.ศ. 2022 ผู้แต่ง P. Manomaisaowapak และ J. Songsiri บทความเรื่อง Joint estimation of multiple Granger causal networks: Inference of group-level brain connectivity	256
11.13	ขั้นตอนการเลือกพารามิเตอร์ลงโทษในปัญหาการประมาณแบบเบาบาง: การใช้เกณฑ์การเลือกแบบจำลอง และวิธีตรวจสอบไขว้	258
B.1	วงรี $x^T P^{-1} x = \alpha$	269
B.2	การแยกตัวประกอบ QR และ RQ	270
C.1	วงรีความเชื่อมั่นของเวกเตอร์เกาส์เซียน	278

สารบัญทฤษฎีบท

2.1	ประพจน์ (ระบบเชื่อมโยงกัน)	35
3.1	ทฤษฎีบท (บทตั้ง C5.3.1 ของ [SS89])	46
3.2	ทฤษฎีบท (Theorem 13.1 ของ [Lju99])	52
4.1	นิยาม (unbiased estimator)	64
4.2	นิยาม (consistent estimator)	64
4.3	นิยาม (Efficient estimator)	65
4.4	นิยาม (Asymptotic distribution of estimators)	65
4.5	นิยาม (System identifiable)	67
4.6	นิยาม (Parameter identifiable)	68
4.1	ทฤษฎีบท (Consistency of global minimum)	69
4.2	ทฤษฎีบท (Consistency of local minimum)	69
4.3	ทฤษฎีบท (Limit distribution of local minimum)	70
5.1	ทฤษฎีบท (ทฤษฎีบทของเกาส์-มาร์คอฟ หรือ Gauss-Markov Theorem) . .	87
7.1	นิยาม (ตัวแปรเชิงเครื่องมือ)	139
9.1	ทฤษฎีบท (ตัวประมาณ MMSE)	181
9.2	ทฤษฎีบท (สมบัติของตัวประมาณ MLE)	196
9.4	ทฤษฎีบท (อสมการของคราเมอร์-ราว Cramér-Rao inequality)	200
10.1	ทฤษฎีบท (Bias-Variance decomposition)	209
B.1	ประพจน์ (สมบัติของเมทริกซ์สมมาตร)	267
B.1	นิยาม (เมทริกซ์บวกกึ่งแน่นอน)	267
B.2	นิยาม (รากที่สองของเมทริกซ์)	267
B.3	นิยาม (การเปรียบเทียบ 2 เมทริกซ์บวกกึ่งแน่นอน)	268
B.4	นิยาม (ส่วนเติมเต็มชัวร์)	268
B.1	บทตั้ง (การหาอินเวอร์สของเมทริกซ์ที่แบ่งเป็นบล็อก)	268
B.2	บทตั้ง (สูตรอินเวอร์สของ Sherman-Morrison-Woodbury)	269
B.5	นิยาม (วงรีใน $\mathbf{R}^n$)	269
B.2	ประพจน์ (ลักษณะของวงรี)	269

C.1	นิยาม (การลู่เข้าเชิงความน่าจะเป็น)	273
C.2	นิยาม (การลู่เข้าเชิงการแจกแจง)	273
C.1	ทฤษฎีบท (กฎจำนวนมากอย่างอ่อน (weak law of large numbers (WLLN)))	274
C.2	ทฤษฎีบท (กฎของลิมิตเชิงความน่าจะเป็น)	275
C.3	ทฤษฎีบท (ทฤษฎีบทการส่งต่อเนื่อง—continuous mapping theorem) . .	275
C.4	ทฤษฎีบท (ทฤษฎีบทของ Slutsky—Slutsky’s Theorem)	275
C.5	ทฤษฎีบท (กฎลิมิตผลคูณปกติ—product limit normal rule)	276
C.6	ทฤษฎีบท (ทฤษฎีบทลิมิตกลางของ Lindeberg-Levy)	276
C.7	ทฤษฎีบท (ทฤษฎีบทลิมิตกลางของ Lindeberg-Feller)	276
C.1	ประพจน์ (การแปลงสัมพรรคของเกาส์เซียน—affine mapping of Gaussian)	277
C.2	ประพจน์ (การแจกแจงตามขอบของเกาส์เซียน)	277
C.3	ประพจน์ (สหสัมพันธ์กับความเป็นอิสระต่อกัน)	277
C.4	ประพจน์ (ตัวแปรเกาส์เซียนแบบมีเงื่อนไข)	277
C.5	ประพจน์ (เกาส์เซียนมาตรฐาน)	278
C.6	ประพจน์ (กำลังสองของเกาส์เซียน)	278
C.7	ประพจน์ (วงรีเชื่อกัน)	278
C.3	นิยาม (กระบวนการคงที่เชิงกว้าง)	279
C.4	นิยาม (สมบัติกระบวนการคงที่เชิงกว้าง)	279
C.5	นิยาม (กระบวนการกึ่งคงที่)	279
C.8	ทฤษฎีบท (ทฤษฎีบทของวินเนอร์-คินชิน Wiener-Khinchin Theorem) . .	280
C.9	ทฤษฎีบท (ความหนาแน่นเชิงสเปกตรัมกำลังของสัญญาณออกจากระบบเชิงเส้น)	280
C.10	ทฤษฎีบท (ฟังก์ชันแปรปรวนร่วมของสัญญาณออกจากระบบเชิงเส้น)	280

สารบัญรหัสคำสั่ง

2.1	ตัวอย่างผลการทำนายจากระบบที่กำหนดพารามิเตอร์เองกับระบบที่ประมาณ	21
3.1	การสร้างสัญญาณเข้า	59
3.2	การเก็บข้อมูลเพื่อหาเอกลักษณ์ระบบแบบวงเปิด	60
5.1	การใช้คำสั่ง fitlm	82
6.1	ปัญหา LS ที่มีเงื่อนไขบังคับ	125
6.2	ปัญหา LS ที่มีเงื่อนไขเป็นนอร์มบอล	126
6.3	ปัญหาแลสโซ	126
6.4	ปัญหาดาข่ายยืดหยุ่น	126
6.5	ปัญหาแลสโซกลุ่ม	127
6.6	ปัญหาแลสโซเชื่อม	127
6.7	ปัญหากำลังสองน้อยสุดแบบเกร็ง	127
6.8	การประมาณฟังก์ชันด้วยวิธีกำลังสองน้อยสุดไม่เชิงเส้น	128
7.1	การประมาณแบบจำลองปริภูมิสถานะ	149
7.2	การประมาณแบบจำลองปริภูมิสถานะที่มีโครงสร้าง	151
8.1	การประมาณ ARMAX ด้วยวิธี PEM	173
8.2	การประมาณแบบจำลองปริภูมิสถานะด้วยวิธี PEM	174
8.3	การประมาณแบบจำลอง ARMA	175
10.1	การประมาณแบบจำลองด้วยคำสั่ง fit	232
10.2	การประมาณแบบจำลองด้วยคำสั่ง fitlm	232
10.3	การเลือกพารามิเตอร์ลงโทษในปัญหาแลสโซ	233
10.4	การสร้างข้อมูลอนุกรมเวลาและประมาณแบบจำลอง ARMAX	233
10.5	การประมาณแบบจำลอง ARMAX และ ARX และตรวจสอบแบบจำลอง	234
10.6	การประมาณแบบจำลอง ARMA ด้วยชุดคำสั่ง Econometric	234

สัญลักษณ์และตัวแปร

สัญลักษณ์	คำอธิบาย
$\mathbf{R}$	เซตจำนวนจริง
$\mathbf{R}^n$	ปริภูมิเวกเตอร์ค่าจริงอันดับ n
$\mathbf{R}^{m \times n}$	ปริภูมิเมทริกซ์ค่าจริงอันดับ $m \times n$
$\mathbf{S}^n$	ปริภูมิเมทริกซ์สมมาตรค่าจริงขนาด $n \times n$
$\mathbf{C}$	เซตจำนวนเชิงซ้อน
i	$\sqrt{-1}$
$\text{Re}(z)$	ค่าจริงของจำนวนเชิงซ้อน z
$\text{Im}(z)$	ค่าจินตภาพของจำนวนเชิงซ้อน z
$\mathbf{C}^{m \times n}$	ปริภูมิเมทริกซ์ค่าเชิงซ้อนอันดับ $m \times n$
$\mathbf{S}^n$	เซตของเมทริกซ์สมมาตรขนาด $n \times n$
$\mathcal{N}(A)$	ปริภูมิศูนย์ (nullspace) ของเมทริกซ์ A
$\mathcal{R}(A)$	ปริภูมิพิสัย (range space) ของเมทริกซ์ A (หรือปริภูมิคอลัมน์)
$\mathcal{M}^\perp$	ส่วนเติมเต็มตั้งฉากของปริภูมิ $\mathcal{M}$
$\text{tr}(A)$	ผลรวมแนวทแยงมุมของ A
$\text{diag}(A)$	สมาชิกแนวทแยงมุมของเมทริกซ์ A
$\text{rank}(A)$	ลำดับชั้นของเมทริกซ์ A
$\sigma(A)$	ค่าเอกฐานของ A
$A^\dagger$	ตัวผกผันเทียม (pseudo inverse) ของ A
A^*	สังยุคเฮอร์มิเทียนของ $A \in \mathbf{C}^{n \times n}$ (Hermittian conjugate) มีค่าเท่ากับ $A^* = \bar{A}^T$
$\mathbf{E}[X]$	ค่าคาดหวังของตัวแปรสุ่ม X
$\text{var}[X]$	ความแปรปรวนของตัวแปรสุ่ม X เมื่อ X เป็นสเกลาร์
$\text{cov}(X)$	เมทริกซ์ความแปรปรวนร่วมของ X
$\text{dist}(x, S)$	ระยะระหว่างจุด x ใดๆ ไปยังเซต S

หนังสือเล่มนี้ใช้ตัวแปรสำหรับเวกเตอร์และเมทริกซ์เป็นหลัก ดังนั้น $x \in \mathbf{R}^n$ จะหมายถึงเวกเตอร์ที่มี n สมาชิก (เราจะไม่ใช้ $\vec{x}$) ดังนั้นการแยกแยะระหว่างสเกลาร์กับเวกเตอร์จึงพิจารณาอันดับของ x ได้ตามบริบท เช่น xy^T เป็นเมทริกซ์ $n \times n$ ส่วน $x^T y$ เป็นค่าสเกลาร์ หรือ พจน์ $(I + x^T y)^{-1}$ แท้ที่จริงแล้วเป็นสเกลาร์ จึงมีค่าเท่ากับ $1/(1 + x^T y)$ สัญลักษณ์ $\mathbf{1}$ คือเวกเตอร์ที่ทุกสมาชิกเท่ากับหนึ่ง นอร์มต่างๆ ของเวกเตอร์ x ใช้สัญลักษณ์เป็น $\|x\|_1, \|x\|_2, \|x\|_\infty$ หมายถึงนอร์ม-1 นอร์ม-2 และนอร์มอนันต์ ตามลำดับ เราใช้ $\|x\|$ สำหรับนอร์มใดๆ (ขอให้ดูตามบริบท) การอ้างถึงสมาชิกตัว

ที่ i ของ x ใช้ x_i ส่วนสัญลักษณ์ $x \succeq y$ หมายถึง $x_i \geq y_i$ สำหรับทุก i ฟังก์ชันของเวกเตอร์ที่ดำเนินการทีละสมาชิกของ x (elementwise operator) จะใช้สัญลักษณ์ที่คล้ายกับคำสั่ง MATLAB เช่น $\text{sign}(x)$ ให้ค่าเครื่องหมายของแต่ละสมาชิกใน x เป็น ± 1 หรือ $\max(0, x) = x^+$ คือการให้ค่า x_i หาก $x_i \geq 0$ และให้ค่าเป็นศูนย์ถ้าหาก $x_i < 0$

สัญลักษณ์ $X \in \mathbf{R}^{m \times n}$ หมายถึงเมทริกซ์ค่าจริง และ x_{ij} จะหมายถึงตำแหน่ง (i, j) ของ X เมื่อเราใช้ X^T จะหมายถึงเมทริกซ์สลับเปลี่ยนของ X (transpose) ในพีชคณิตเมทริกซ์ เรามักจะแบ่ง $X \in \mathbf{R}^{m \times n}$ ออกเป็นแถวหรือคอลัมน์ ดังนี้

$$X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_m^T \end{bmatrix}, \quad X = [x_1 \ x_2 \ \cdots \ x_n]$$

สัญลักษณ์ $X \succeq 0$ และ $X \succ 0$ หมายถึง X เป็นเมทริกซ์บวกกึ่งแน่นอน (positive semidefinite) และเมทริกซ์บวกแน่นอน (positive definite) ตามลำดับ หาก $X \not\succeq 0$ จะหมายถึง X เป็นเมทริกซ์ไม่แน่นอน (indefinite matrix) สำหรับ $X \succeq 0$ เราใช้ $X^{1/2}$ ว่าเป็นรากที่สองของเมทริกซ์ X (ดูนิยาม B.2) ถ้าหาก $X \succeq 0$ สัญลักษณ์ $X^{-1/2}$ หมายถึงตัวผกผันของ $X^{1/2}$

ในการอธิบายฟังก์ชัน $f: X \rightarrow Y$ เป็นฟังก์ชัน f ที่แปลงค่าจากสมาชิกในเซต X ไปยังภาพฉายที่อยู่ในเซต Y สัญลักษณ์ $\text{dom } f$ คือโดเมนของฟังก์ชัน f

ในแต่ละศาสตร์จะมีแนวปฏิบัติการใช้ตัวแปรที่ต่างกันไป เช่น

- ด้านพีชคณิตเชิงเส้นใช้ตัวพิมพ์เล็ก x เป็นเวกเตอร์ ตัวพิมพ์ใหญ่ X เป็นเมทริกซ์ หรือบางครั้งเป็น เซต/ปริภูมิ
- ด้านระบบควบคุมใช้ G, H, F แทนฟังก์ชันถ่ายโอน และ g, h, f แทนผลตอบสนองอิมพัลส์ที่สมนัย
- ด้านความน่าจะเป็น ใช้ตัวแปรพิมพ์ใหญ่ X เพื่อแทนตัวแปรสุ่ม ตัวพิมพ์เล็ก x เป็นค่าที่เป็นได้ค่าหนึ่งของ X หากเป็นเรื่องกระบวนการสุ่ม เราเลือกใช้สัญลักษณ์ $x(t)$ แทนกระบวนการสุ่ม (ถึงแม้ว่าหากจะกล่าวให้รัดกุมต้องใช้ $\{x(t)\}_{t \in T}$ อันหมายถึงเซตของ $x(t)$ ทุกเวลา โดยทั่วไปจะกำหนดให้ $x(t) = (x_1(t), x_2(t), \dots, x_n(t))$ เป็นเวกเตอร์กระบวนการสุ่ม

เนื่องจากตำราเล่มนี้ มีเนื้อหาที่ประยุกต์จากหลายศาสตร์ จึงยังไม่ได้ปรับการใช้สัญลักษณ์ให้เป็นไปตามแนวปฏิบัติของแต่ละศาสตร์ อย่างสม่ำเสมอกับทุกบท เพราะอาจทำให้ผู้อ่านสับสน (เช่น การใช้ตัวแปร X ที่อาจจะหมายถึงเมทริกซ์ หรือตัวแปรสุ่ม) จึงขอให้ผู้อ่านตีความจากบริบทเป็นสำคัญ

อักษรย่อ

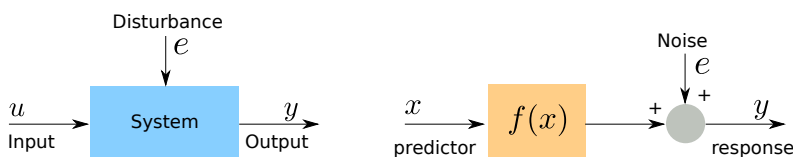
AIC	เกณฑ์สารสนเทศของอะกะอิเกะ Akaike Information Criterion.
AR	ถดถอยในตัว Autoregressive.
ARIMA	รวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่ Autoregressive Integrated Moving Average.
ARMA	เฉลี่ยเคลื่อนที่ที่ถดถอยในตัว Autoregressive Moving Average.
ARMAX	เฉลี่ยเคลื่อนที่ที่ถดถอยในตัวที่มีสัญญาณนอกระบบ Autoregressive Moving Average with Exogeneous Input.
ARX	ถดถอยในตัวที่มีสัญญาณนอกระบบ Autoregressive with Exogeneous Input.
BIC	เกณฑ์สารสนเทศของเบส์ Bayesian Information Criterion.
BLUE	ตัวประมาณค่าไม่เอนเอียงเชิงเส้นดีที่สุด Best Linear Unbiased Estimator.
CLT	ทฤษฎีบทลิมิตกลาง Central Limit Theorem.
CR	อสมการคราเมอร์-ราว Cramér-Rao inequality.
CV	การตรวจสอบไขว้ Cross validation.
CVA	Canonical Variate Analysis.
DCM	การจำลองเชิงเหตุผลพลวัต Dynamic causal modeling.
DGP	กระบวนการก่อกำเนิดข้อมูล Data Generating Process.
EEG	อีอีจี Electroencephalogram.
ESS	ผลบวกกำลังสองที่ถูกลบอธิบาย Explained Sum of Squares.
FIR	ผลตอบสนองอิมพัลส์อันดับจำกัด Finite Impulse Response.
fMRI	เอฟเอ็มอาร์ไอ Functional magnetic resonance imaging.
i.i.d.	ไอไอดี Independently and Identically Distributed.
KL	การลู่ออกของคัลล์แบก-ลีบเลอร์ Kullback–Leibler divergence.
LMSE	การประมาณเชิงเส้นไม่เอนเอียงที่มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด Linear Unbiased MMSE Estimation.
LS	กำลังสองน้อยสุดเชิงเส้น Linear Least-Squares.

MA	การเฉลี่ยเคลื่อนที่.
MAE	ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย Mean Absolute Error.
MAP	การประมาณภายหลังสูงสุด Maximum a posteriori estimation.
MEG	เอ็มอีจี Magnetoencephalography.
MIMO	หลายสัญญาณเข้าและหลายสัญญาณออก Multiple Input Multiple Output.
MLE	การประมาณภาวะน่าจะเป็นสูงสุด Maximum likelihood estimation.
MMSE	ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด Minimum mean square error.
MSE	ค่าคลาดเคลื่อนกำลังสองเฉลี่ย Mean Square Error.
MVUE	ตัวประมาณไม่เอนเอียงแปรปรวนต่ำสุด Minimum Variance Unbiased Estimator.
N4SID	Numerical Algorithms for Subspace State Space System Identification.
NLS	กำลังสองน้อยสุดไม่เชิงเส้น Nonlinear Least Squares.
OE	ค่าคลาดเคลื่อนที่สัญญาณออก Output Error.
OLS	กำลังสองน้อยสุดสามัญ Ordinary Least-Squares.
PEM	วิธีความผิดพลาดในการทำนาย Prediction Error Method.
PO-MOESP	Past outputs Multivariable Output-Error State-space.
PRBS	ลำดับฐานสองสุ่มเทียม Pseudo Random Binary Sequence.
QP	กำหนดการกำลังสอง Quadratic programming.
RMSE	รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย Root Mean Square Error.
RSE	ความคลาดเคลื่อนส่วนเหลือมาตรฐาน Residual Standard Error.
RSS	ผลบวกกำลังสองของส่วนเหลือ Residual Sum of Squares.
SE	ความคลาดเคลื่อนมาตรฐานของสัมประสิทธิ์ Standard Error of the Coefficient.
SEM	การจำลองแบบสมการเชิงโครงสร้าง Structural equation models.
SISO	สัญญาณเข้าเดียวและสัญญาณออกเดียว Single Input Single Output.
TSS	ผลบวกกำลังสองทั้งหมด Total Sum of Squares.
VAR	เวกเตอร์กระบวนการถดถอยในตัว Vector autoregressive process.
WLLN	กฎจำนวนมากอย่างอ่อน Weak law of large numbers.
WLS	กำลังสองน้อยสุดถ่วงน้ำหนัก Weighted Least-squares.
WSS	คงที่เชิงกว้าง Wide-sense stationary.

บทที่ 1

บทนำ

การหาเอกลักษณ์ของระบบ คือการสร้างแบบจำลองเพื่อเป็นตัวแทนของระบบนั้นจากข้อมูลของสัญญาณเข้าและสัญญาณออกที่วัดได้ ด้วยการกระตุ้นสัญญาณเข้าไปยังระบบ จากรูป 1.1 (ซ้าย) แสดงถึงระบบพลวัตทั่วไปที่มีสัญญาณออก (output) y อันเป็นผลตอบสนองของระบบที่มีต่อสัญญาณเข้า (input) u และสัญญาณรบกวน (disturbance) e ผสมกัน เนื่องจากโดยทั่วไป สัญญาณรบกวนเป็นสิ่งที่ไม่ทราบค่า ส่วนสัญญาณเข้า u นั้นจะเป็นสิ่งที่ผู้ใช้กำหนด หรือ u สามารถหาค่าได้ รูปแบบปัญหาการหาเอกลักษณ์ของระบบ คือ การมีเซตข้อมูลของ (u, y) อันจะเรียกว่าข้อมูล I/O ที่พึงตระหนักเสมอว่า y นั้นจะมีทั้งส่วนที่เป็นผลตอบสนองต่อ u และผลตอบสนองต่อสัญญาณรบกวน (อันเป็นส่วนไม่พึงประสงค์) ดังนั้น การจะหาแบบจำลอง หรือคำจำกัดความของระบบ จึงต้องอาศัยทฤษฎีด้านการประมาณแบบจำลอง ตัวอย่างง่าย เช่น ให้ระบบ $G = 2$ เป็นระบบค่าคงตัว หากเราใส่ $u = 2$ และมีสัญญาณรบกวน $e = 0.01$ โดยที่สัญญาณรบกวนนั้นเกิดในรูปแบบการบวก (additive noise) กล่าวคือ $y = Gu + e$ เราจะได้ว่า $y = 4.01$ ดังนั้นหากเราถือว่า จะหา $G = y/u$ เราจึงประมาณได้ว่า $\hat{G} = 2.005$ ซึ่งเกิดความคลาดเคลื่อนของระบบที่ประมาณได้เมื่อเทียบกับระบบจริง อันเป็นผลมาจาก e



รูป 1.1: ความสัมพันธ์ของสัญญาณเข้ากับสัญญาณออกผ่านระบบที่มีสัญญาณรบกวน

รูป 1.1 (ขวา) แสดงถึงการอธิบายระบบในรูปแบบจำลองทางคณิตศาสตร์ในศาสตร์การเรียนรู้เชิงสถิติ สัญญาณออก y เรียกว่า ตัวแปรตอบสนอง (response) ซึ่งเกิดจากตัวแปรทำนาย (predictor) x อันส่งผลถึง y ผ่านทางความสัมพันธ์ที่อธิบายด้วยฟังก์ชัน f และโดนรบกวนด้วย e ที่พิจารณาให้เป็นตัวแปรสุ่มอันเป็นลักษณะของสัญญาณรบกวน (noise) การประมาณระบบ จึงหมายถึงการประมาณฟังก์ชัน f จากชุดข้อมูล (x, y) หรือ หมายถึงระบบ G จากชุดข้อมูล (u, y) นั่นเอง

การหาเอกลักษณ์ของระบบ หรือการประมาณแบบจำลองนั้น มีความเกี่ยวข้องในงานประยุกต์ที่หลากหลายสาขา บทนี้จะกล่าวถึง ความสำคัญของการเข้าใจพื้นฐานหลักการประมาณแบบจำลอง และ

องค์ประกอบ ขั้นตอนที่สำคัญเพื่อบรรลุวัตถุประสงค์ในการหาแบบจำลอง ในตอนท้ายของบทนี้จะเสนอตัวอย่างผลลัพธ์เบื้องต้นเพื่อประกอบการอธิบายขั้นตอนเหล่านั้น

1.1 ความสำคัญของการหาเอกลักษณ์ของระบบ

กำหนดให้แบบจำลองทางคณิตศาสตร์ที่แสดงลักษณะของระบบ คือ

$$y = Gu + e \quad (1.1)$$

หรือ

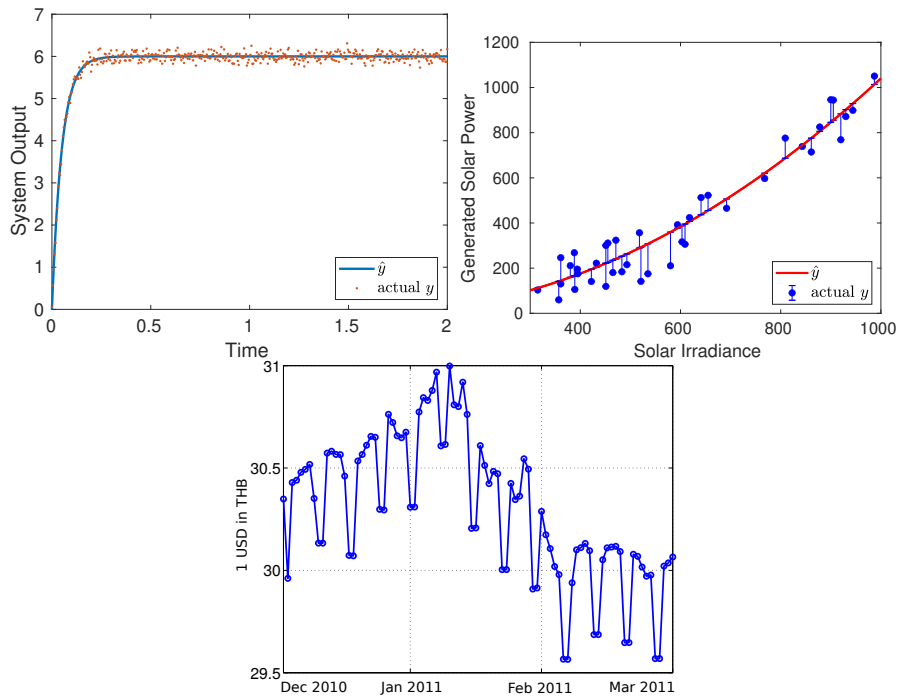
$$y = f(x) + e \quad (1.2)$$

หลังจากการประมาณระบบ นั่นคือ หา $\hat{G}$ หรือ $\hat{f}$ ได้ เราสามารถแบ่งผลการนำไปใช้ของการประมาณแบบจำลอง ได้เป็น 3 กลุ่มใหญ่ดังนี้

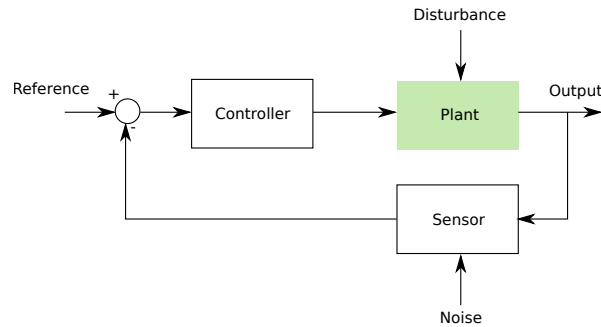
การทำนาย ปัญหากลุ่มนี้ คือเราจะมีข้อมูลของ u หรือ x แต่เราไม่ทราบค่า y หรืออาจจะหาได้ยาก ดังนั้น หากเราทราบค่าประมาณของระบบ เราจึงประมาณค่า y ได้ด้วย $\hat{y} = \hat{G}u$ หรือ $\hat{y} = \hat{f}(x)$ ตัวอย่างเช่น หาก G แสดงถึงระบบมอเตอร์กระแสตรงที่ y คือ ความเร็วเชิงมุมของมอเตอร์ (สมมติให้หน่วยเป็นโวลต์ หรือ V เมื่อแปลงเป็นสัญญาณทางไฟฟ้า) และ u คือ สัญญาณเข้ามอเตอร์ (V) เมื่อเราประมาณได้ว่า $\hat{G}(s) = \frac{3}{(0.05s+1)}$ เราก็ประมาณได้ว่า เมื่อใส่ $u(t) = 2$ V สำหรับ $t \geq 0$ เข้ามอเตอร์ เราจะประมาณความเร็วเชิงมุมของมอเตอร์ในภาวะอยู่ตัวได้เป็น 6 V โดยที่ค่าจริงของ y ที่วัดได้นั้น อาจจะเบี่ยงเบนจาก 6 จากผลของสัญญาณรบกวนใน (1.1) หรือเราอาจตั้งคำถามว่า เราจะประมาณค่า ความเร็วเชิงมุม ณ วินาทีที่ 0.05 ได้เป็นเท่าใด เราก็สามารถใช้แบบจำลอง $\hat{G}$ เพื่อประมาณค่า $\hat{y}(0.05)$ ได้ ดังรูป 1.2 (ซ้าย) ตัวอย่างที่สอง คือความสัมพันธ์ระหว่าง y ที่เป็นกำลังผลิตไฟฟ้าจากแผงเซลล์สุริยะหนึ่งกับ x ที่เป็น ความเข้มแสงอาทิตย์ หากเราประมาณความสัมพันธ์นั้นได้เป็น $y = a_1x + a_2x^2$ โดยที่ a_1, a_2 คือค่าคงตัวที่ประมาณได้ เราจึงสามารถประมาณ $\hat{y}$ ได้ หากเราทราบค่าความเข้มแสงอาทิตย์และอุณหภูมิแผง ณ เวลาหนึ่งของวันที่สนใจ ดังรูป 1.2 (ขวา) ปัญหาการทำนายยังรวมถึงการพยากรณ์ค่า y ในอนาคต เช่นรูป 1.2 (ล่าง) หากเราสมมติแบบจำลอง เป็น $\hat{y}_{Apr} = a_1y_{Mar} + a_2y_{Feb}$ และเราประมาณค่า a_1, a_2 ได้ เมื่อทราบค่าอัตราแลกเปลี่ยนเงินบาทในอดีตสองเดือนก่อนหน้า เราจึงสามารถทำนายค่าอัตราแลกเปลี่ยนในเดือนถัดไปได้

การออกแบบตัวควบคุม เมื่อพิจารณารูป 1.3 การออกแบบตัวควบคุมโดยทั่วไปนั้น จะอาศัยแบบจำลองคณิตศาสตร์ของระบบ เช่น ฟังก์ชันถ่ายโอน หรือสมการปริภูมิสถานะ เพื่อให้ผู้ออกแบบเข้าใจสมบัติพื้นฐานของระบบ เช่น เสถียรภาพของระบบวงเปิด ตำแหน่งโพล ก่อนที่จะออกแบบตัวควบคุมเพื่อแก้ปัญหาคควบคุมใดที่สนใจ ในภาคปฏิบัติ หากเรามีระบบ G ทางกายภาพ เช่น มอเตอร์กระแสตรง แต่ผู้ใช้ไม่ทราบข้อกำหนดรายละเอียดของฮาร์ดแวร์ที่นำมาประกอบเป็นมอเตอร์นั้น จึงสามารถกำหนดโครงสร้างของแบบจำลอง เช่น การพิจารณาแบบจำลองเชิงเส้นอันดับหนึ่ง และนำข้อมูล (u, y) มาประมาณ G เมื่อประมาณได้ $\hat{G}$ แล้ว จึงออกแบบตัวควบคุมที่ใช้ค่าพารามิเตอร์ของ $\hat{G}$

การอนุมาน เมื่อเราใช้แบบจำลองคณิตศาสตร์ในการบรรยายระบบ พารามิเตอร์ของแบบจำลอง หรือฟังก์ชันของพารามิเตอร์นั้น สามารถบ่งชี้บางอย่างทางการเรียนรู้เชิงสถิติ หรือบ่งชี้รูปแบบใดในข้อมูลได้ ตัวอย่างหนึ่งสำหรับ (1.2) เมื่อ f เป็นฟังก์ชันเชิงเส้น $y = \beta_1x_1 + \beta_2x_2$ เราสามารถทดสอบ



รูป 1.2: ตัวอย่างการใช้แบบจำลองที่ประมาณได้ในการทำนายค่าสัญญาณออก y



รูป 1.3: การหาเอกลักษณ์ของระบบเพื่อนำไปใช้ร่วมกับการออกแบบตัวควบคุม

สมมติฐานเชิงสถิติว่า $\beta_1 = 0$ หรือไม่ ซึ่งผลการทดสอบนั้นจะอธิบายว่า ตัวแปรทำนาย x_1 มีความสำคัญต่อการอธิบาย y หรือไม่ กล่าวคือ รูปแบบของศูนย์ในพารามิเตอร์ β ใช้เพื่ออนุมานความสำคัญของตัวแปรทำนาย อันเป็นวิธีพื้นฐานหนึ่งของการคัดเลือกคุณลักษณะ (feature selection) ในปัญหาการเรียนรู้เชิงสถิติ ตัวอย่างที่สอง คือ ระบบพลวัตเชิงเส้นในรูปแบบฟังก์ชันถ่ายโอน (1.1) ที่มีหลายสัญญาณเข้าและหลายสัญญาณออก Multiple Input Multiple Output (MIMO) เช่น

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) & \cdots & G_{1m}(s) \\ G_{21}(s) & G_{22}(s) & \cdots & G_{2m}(s) \\ \vdots & \vdots & \ddots & \vdots \\ G_{r1}(s) & G_{r2}(s) & \cdots & G_{rm}(s) \end{bmatrix}$$

หากสัมประสิทธิ์ทั้งหมดของพหุนามเศษในฟังก์ชันถ่ายโอนย่อย $G_{ij}(s)$ เป็นศูนย์ จึงสรุปได้ว่า ถึงแม้จะใส่สัญญาณเข้า $u_j \neq 0$ ในระบบนี้ แต่ y_i จะไม่เกิดผลตอบสนองต่อ u_j แต่อย่างใด กล่าวคือ รูปแบบศูนย์ของฟังก์ชันถ่ายโอนแบบ MIMO จะบอกถึงโครงสร้างความสัมพันธ์ระหว่าง u กับ y

การนำ $\hat{G}$ ไปใช้เป็นตัวแทนของ G จะมีสมรรถนะดีเพียงใด ก็ขึ้นกับความถูกต้องในการประมาณแบบจำลอง ที่จะเกิดจากปัจจัยอันสำคัญ คือ หนึ่ง การใช้ความรู้เกี่ยวกับสัญญาณรบกวน e ในการประมาณแบบจำลอง สอง การมีสมมติฐานของโครงสร้างแบบจำลองว่าจะใกล้เคียงกับระบบจริงเพียงใด และ สาม เทคนิควิธีในการประมาณแบบจำลองนั้น ในตำราจะกล่าวถึงรายละเอียดต่อไปสำหรับผลกระทบของปัจจัยทั้งสาม และเทคนิคการประมาณต่างๆ สำหรับระบบเชิงเส้น

1.2 องค์ประกอบที่สำคัญ

การประมาณแบบจำลองเกี่ยวข้องกับองค์ประกอบที่สำคัญ 3 ส่วน และมีขั้นตอน ดังนี้

ส่วนแรก คือ ข้อมูล (u, y) สัญญาณเข้า-สัญญาณออก หรือ (x, y) ตัวแปรทำนาย-ตัวแปรตอบสนอง หากเราประมาณระบบทางกายภาพ เช่น ระบบอุณหภูมิในห้องปิด เราจำเป็นต้องมีตัววัดค่าพลังงานเข้า u และตัววัดอุณหภูมิ y ในห้องนั้น การออกแบบสัญญาณเข้าเพื่อไปกระตุ้นระบบให้เกิดสัญญาณออก ก็เป็นหัวข้อสำคัญที่จะกล่าวต่อไป ว่าควรเลือก u เป็นสัญญาณลักษณะอย่างไร ในภาคปฏิบัติที่ข้อมูลจริงนั้นจะมีความไม่เป็นอุดมคติ เช่น ค่าวัดหายไปในช่วงเวลา ค่า y คลาดเคลื่อนไปจากความจริง เช่น วัดอุณหภูมิบรรยากาศในกรุงเทพในหน่วยเซลเซียสออกมาได้เป็นค่าลบ หรือ การวัดสัญญาณไฟฟ้าของคลื่นสมอง (EEG) ก็สามารถคลาดเคลื่อนไปในระดับสูงหากเรากระพริบตา หรือ ขยับกล้ามเนื้อในส่วนใบหน้า ดังนั้น ขั้นตอนสำคัญก่อนที่จะนำข้อมูลไปใช้ประมาณระบบ มีตัวอย่างเช่น การเติมข้อมูล (missing-value imputation) การจำกัดสิ่งแปลกปน (artifacts) นอกจากนี้ในการพิจารณาแบบจำลองหนึ่งๆ หรือ การใช้เทคนิควิธีเชิงตัวเลขในการประมาณ อาจมีสมมติฐานเกี่ยวกับช่วงของข้อมูล ที่จะส่งผลกับรูปแบบการประมาณแบบจำลอง หรืออาจจะส่งผลกับสมรรถนะของเทคนิควิธีเชิงตัวเลขนั้น ขั้นตอนการจัดการข้อมูลจึงอาจรวมถึง การจัดเรียงข้อมูลให้เข้ากับรูปแบบจำลอง หรือ การแปลงค่าข้อมูลให้อยู่ช่วงมาตรฐาน (normalization) เช่น การกำจัดแนวโน้ม (detrend) และแปลงให้ข้อมูลมีความแปรปรวนเป็นหนึ่ง เป็นต้น ทั้งนี้ ในหลายงานประยุกต์ x อาจจะเป็นปริมาณที่แปลงมาจากตัวแปรต้นตัวอื่น เช่น การจะอธิบายสัญญาณกะพริบของตาจาก EEG เราจะสกัดองค์ประกอบของสัญญาณที่ความถี่ใกล้เคียงกับการกะพริบตา (9-13 ครั้งต่อวินาที) ดังนั้น ในที่นี้เราอาจจะพิจารณา กำหนดให้ x คือพลังงานของสัญญาณในช่วงความถี่ 0.15-0.22 Hz ซึ่งต้องอาศัยการแปลงเชิงสเปกตรัมจากสัญญาณทางเวลา ให้ไปเป็นสัญญาณเชิงความถี่ก่อน กล่าวโดยสรุปนั้น การจัดการข้อมูลในขั้นต้น (pre-processing) จะเกี่ยวข้องกับ การจัดการความผิดพลาดของข้อมูล การเติมข้อมูล การแปลงข้อมูลให้เป็นเวกเตอร์คุณลักษณะ (feature) ที่สนใจ (หากมี) การจัดเรียงและแปลงข้อมูลในช่วงมาตรฐาน

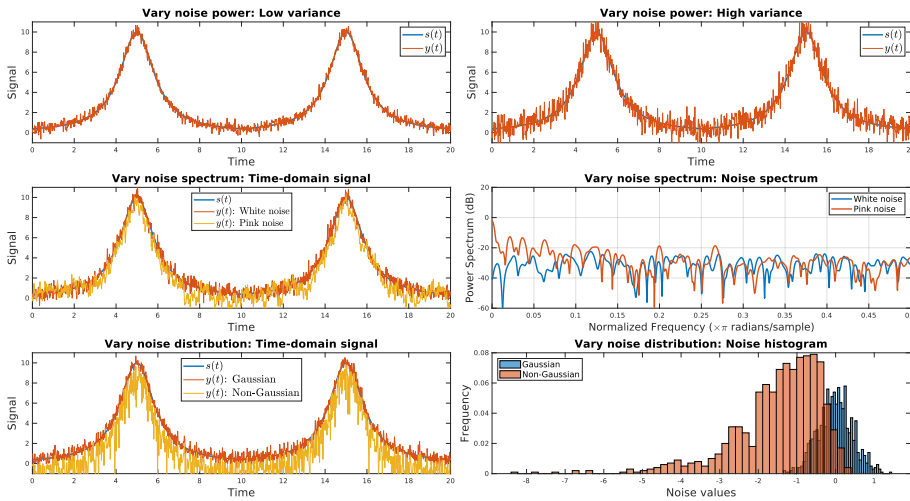
ส่วนที่สอง คือ แบบจำลองคณิตศาสตร์ที่เป็นส่วนสำคัญอันส่งผลถึงสมรรถนะของการใช้แบบจำลองอันดับแรก ผู้ออกแบบต้องเลือกโครงสร้างแบบจำลองที่เหมาะสมกับระบบจริงที่ต้องการประมาณ เช่น การอธิบายระบบกายภาพที่ทราบว่ามีคุณสมบัติเชิงเส้นในจุดทำงานหนึ่ง เราเลือกใช้ระบบพลวัตแบบเชิงเส้นที่มีอันดับจำกัดเหมาะสม การเลือกแบบจำลองไม่เชิงเส้นเป็นฟังก์ชันคณิตศาสตร์ที่ได้จากการวิเคราะห์สมการลักษณะเฉพาะของระบบ หรือ การใช้แบบจำลองเชิงความน่าจะเป็น (probabilistic model) สำหรับการอธิบายปรากฏการณ์หนึ่งๆ ของตัวแปรสุ่ม เป็นต้น เราจะเลือกแบบจำลองได้ดี

และเหมาะสม หากเรามีความเข้าใจในตัวแปร y อย่างถ่องแท้ เพราะการทราบค่า y มีผลมาจากตัวแปรใด นั่นคือ กำหนดให้ชุดว่าตัวแปรกายภาพใดควรเป็นสัญญาณเข้า (หรือเป็นตัวแปรทำนาย) จะทำให้เราพิจารณาสมการคณิตศาสตร์ที่อธิบายความสัมพันธ์ของ (u, y) หรือ (x, y) ได้เหมาะสม ตัวอย่างเช่น เราต้องการหาแบบจำลองเพื่ออธิบายกำลังผลิตไฟฟ้าจากเซลล์สุริยะ เมื่อเราศึกษาหลักการฟิสิกส์ของงานประยุกต์นี้ จะพบว่ากำลังผลิตไฟฟ้าจะขึ้นกับความเข้มแสงอาทิตย์ และอุณหภูมิของแผงเซลล์ แต่เนื่องจากการแผ่รังสีของดวงอาทิตย์เป็นปรากฏการณ์ธรรมชาติ ค่าของความเข้มแสงอาทิตย์จึงขึ้นกับตัวแปรสภาพอากาศหลายตัวแปร เช่น ปริมาณเมฆ (เพราะเมฆจะมาบังแดดได้) ขนาดและทิศทางลม (เพราะลมพัดก้อนเมฆให้ลอยไปจากจุดที่แผงเซลล์ตั้งอยู่) อุณหภูมิ หรือความชื้นสัมพัทธ์ในอากาศ ตัวแปรเหล่านี้จึงกำหนดให้เป็น x ในการอธิบาย y ที่เป็นกำลังผลิตไฟฟ้า เป็นต้น นอกจากนี้ การหาแบบจำลองที่เหมาะสมกับชุดข้อมูล ไม่ได้ถูกจำกัดที่ผู้ใช้จะต้องเลือกแค่ประเภทเดียว โดยที่เราสามารถพิจารณาแบบจำลองหลายประเภท เช่น ฟังก์ชันถ่วงโอน ปริภูมิสถานะแบบไม่เชิงเส้น และสามารถทดสอบสมรรถนะของแต่ละแบบจำลอง เพื่อผ่านกระบวนการคัดเลือกแบบจำลอง อันจะได้กล่าวต่อไปในหัวข้อ 10.2

ส่วนที่สาม คือ เทคนิคการประมาณแบบจำลอง อันเป็นเนื้อหาหลักของตำรานี้ วิธีการประมาณตั้งอยู่บนหลักการเลือกพารามิเตอร์ ที่ทำให้แบบจำลองมีความเหมาะสมที่สุดด้วยตัวชี้วัดบางอย่าง เช่น วิธีการประมาณแบบกำลังสองน้อยสุด จะเลือกค่าพารามิเตอร์ที่ทำให้ $\hat{y} = Gu$ นั้นมีความผิดพลาดจาก y (ผลตอบระบบจริง) น้อยที่สุด เมื่อวัดความผิดพลาดนั้นด้วยผลรวมกำลังสองของค่าความคลาดเคลื่อนในรูป $y - \hat{y}$ เทคนิคในการประมาณที่เราสามารถเลือกมาใช้ได้นั้น จะขึ้นกับสมมติฐานที่ใช้ของแบบจำลองหนึ่งๆ ด้วย เช่น หากแบบจำลองที่เราเลือก มีการกำหนดเทอมที่ส่งผลจากสัญญาณรบกวน e มายัง y ในลักษณะที่เป็นพลวัต กล่าวคือ y ขึ้นกับค่า e ในอดีต ดังนั้น การใช้เทคนิควิธีความผิดพลาดในการทำนาย Prediction Error Method (PEM) หรือ วิธีการหาเอกลักษณ์แบบจำลองปริภูมิสถานะ จึงเป็นตัวเลือกที่ใช้ได้กับแบบจำลองนี้ แต่เราไม่สามารถประยุกต์ใช้วิธีกำลังสองน้อยสุดได้ เป็นต้น ในขั้นตอนการประมาณแบบจำลองนั้น วิธีการประมาณจะแบ่งเป็น 2 กลุ่มใหญ่ คือกลุ่มที่ใช้ข้อมูลของสัญญาณรบกวน e ในแบบจำลอง กับกลุ่มที่ไม่ต้องใช้สมมติฐานใดเลยจาก e หากเรามีความรู้เกี่ยวกับ e มาก เราอาจจะใช้สิ่งนี้ในการช่วยประมาณแบบจำลองได้ดีขึ้น เช่น หาก e มีความแปรปรวนมาก เราอาจจะเชื่อมั่นได้น้อยลงเกี่ยวกับค่าวัด y ที่นำมาประมาณแบบจำลอง ความเชื่อมั่นใน y ที่น้อยลงจะถูกชดเชยอย่างไรในวิธีการประมาณแบบจำลอง ก็จะเป็นรายละเอียดที่จะอภิปรายในหัวข้อ 5.3 และบทที่ 9 สำหรับระดับขั้นความรู้เกี่ยวกับ e ซึ่งเป็นตัวแปรสุ่ม (หรืออาจจะจำลองเป็นกระบวนการสุ่ม) นั้น เราจะทราบเป็นเพียงแค่บางโมเมนต์ (เช่น ค่าเฉลี่ย ความแปรปรวน) หรือเราจะทราบถึงรูปแบบการแจกแจง (เช่น ฟังก์ชันความหนาแน่นความน่าจะเป็น) ก็จะเป็นปัจจัยที่สอดคล้องต่อหลักการของวิธีประมาณแบบจำลองที่ต่างกันไป

การหาเอกลักษณ์ของระบบเกี่ยวข้องกับทฤษฎีการประมาณ ตัวแปรสุ่ม และกระบวนการสุ่ม ด้วยสาเหตุที่ว่า ข้อมูล (u, y) ที่เราเก็บมาถูกสมมติว่ามีสัญญาณรบกวนปนมาด้วย จากองค์ประกอบทั้งสามที่กล่าวมาข้างต้นพบว่า ข้อมูลความรู้เกี่ยวกับสัญญาณรบกวนนั้น สามารถนำไปใช้ร่วมกับขั้นตอนการจัดการข้อมูล การเลือกแบบจำลอง หรือ การเลือกเทคนิคการประมาณแบบจำลองได้ ขึ้นกับว่าเราทราบเกี่ยวกับสัญญาณรบกวนนั้นมากเพียงใด หากสมมติว่า $y(t) = s(t) + e(t)$ โดยกำหนดให้ $s = Gu$ คือสัญญาณจากระบบที่แท้จริง และ e คือสัญญาณรบกวน และ y คือสัญญาณออกจากระบบที่เราวัดได้ ในแต่ละแถวของรูป 1.4 แสดงตัวอย่างของการทราบสมบัติเกี่ยวกับ e ดังนี้

- การทราบกำลังของ e (คือทราบความแปรปรวน) เป็นตัวอย่างที่พบได้ทั่วไปในทางปฏิบัติ เพราะ



รูป 1.4: ความรู้เกี่ยวกับสมบัติของสัญญาณรบกวนที่นำไปใช้ร่วมกับการประมาณแบบจำลอง

ความรู้เกี่ยวกับบางโมเมนต์ของตัวแปรสุ่ม (เช่น ค่าเฉลี่ย ความแปรปรวน) เป็นสิ่งที่ประมาณได้ง่ายจากค่าตัวอย่าง ตัวอย่างเช่น หาก e แทนสัญญาณรบกวนที่มาจกเครื่องมือวัด (sensor noise) การวัดสัญญาณพื้นหลังและดูค่า RMS จึงเสมือนเป็นการประมาณค่าเบี่ยงเบนมาตรฐานของ e ในแถวที่ 1 ของรูป 1.4 แสดงตัวอย่างการวัด y จาก e ที่มีความแปรปรวนมากกับน้อย หากการเก็บข้อมูล y ในแต่ละค่าตัวอย่างนั้น e มีความแปรปรวนไม่คงที่ (เช่น อาจเกิดจากการเปลี่ยนเครื่องมือวัด) ในหัวข้อ 5.3 จะอธิบายหลักการว่า เราควรนำความแปรปรวนของ e ที่ไม่เท่ากันนั้นมาชดเชยได้อย่างไร ในขั้นตอนการประมาณแบบจำลองด้วยวิธีกำลังสองน้อยสุด

- การทราบช่วงความถี่ของ e คือคุณสมบัติในโดเมนความถี่ของ e ว่ามีการแจกแจงพลังงานของ e อยู่ในช่วงใด ตัวอย่างในแถวสองของรูป 1.4 คือ e เป็นสัญญาณรบกวนขาวที่มีความหนาแน่นพลังงานในทุกย่านความถี่พอๆ กัน หรือเป็นสัญญาณรบกวนชมพู ที่มีความหนาแน่นสเปกตรัมเป็นส่วนกลับกับความถี่ นั่นคือ มีพลังงานสูงในย่านความถี่ต่ำ คุณสมบัติเชิงความถี่ของ e ทำให้เราสามารถจัดการข้อมูลได้เหมาะสม เช่น การใช้ตัวกรองเพื่อกำจัด e ออก ต้องออกแบบย่านความถี่ในตัวกรอง เพื่อไม่ให้บิดเบือน s ที่ต้องการวัด แต่ต้องกำจัด e ออกไป
- การทราบการแจกแจงของ e เช่นแถวสามของรูป 1.4 ว่า e มีการแจกแจงแบบเกาส์เซียนหรือแบบอื่น เช่น เบต้า แกมมา และเป็นการทราบว่าการแจกแจงนั้น มีพารามิเตอร์อย่างไร เช่น การทราบว่า e มีการแจกแจงแบบเกาส์เซียนที่มีพารามิเตอร์ (μ, σ^2) เป็น $(0, 1)$ หรือเป็น $(2, 3)$ ก็ถือว่าเป็นการแจกแจงที่ต่างกัน สมบัตินี้จะส่งผลโดยตรงต่อ เทคนิคการประมาณแบบจำลองที่เลือกใช้ ดังจะอธิบายในบทที่ 9

สำหรับความรู้เกี่ยวกับ e นั้น หากทราบการแจกแจงของ e ก็จะหมายถึง การทราบกำลังของ e ด้วย (แต่บทกลับไม่จริง) เราจึงคาดเดาได้ว่า การทราบการแจกแจงเป็นความรู้เชิงสถิติที่สมบูรณ์ที่สุดของ e (เทียบกับการทราบเพียงแค่บางโมเมนต์) ผลการประมาณที่ใช้ความรู้การแจกแจงด้วย น่าจะทำให้ผลที่ดีกว่า เมื่อเทียบกับการใช้ความรู้บางส่วนของ e เท่านั้น