บทที่ 2
งานวิจัยที่เกี่ยวข้อง

ตัวกรองคาลมานเป็นเทคนิคการประมาณค่าวิธีหนึ่ง ซึ่งได้รับความสนใจและมีการนำไปประยุกต์ใช้กับสาขาต่าง ๆ มากมาย ดังนั้นเพื่อให้เห็นถึงบทบาทและความสำคัญของตัวกรองคาลมาน เนื้อหาของบทนี้จึงกล่าวถึงความเป็นมาและวัตถุประสงค์ของตัวกรองคาลมาน รวมไปถึงแนวคิดและทฤษฎีเกี่ยวกับการประมาณค่า สำหรับเนื้อหาเกี่ยวกับทฤษฎีและขั้นตอนวิธีตัวกรองคาลมานจะกล่าวถึงในบทถัดไป

นอกจากนี้ ได้ทำการรวบรวมงานวิจัยต่าง ๆ ที่เกี่ยวข้องกับการคาดคะเนความต้องการซึ่งถูกนำไปประยุกต์ใช้กับกระบวนการต่าง ๆ

เทคนิคตัวควบคุมแบบไม่เชิงเส้นส่วนใหญ่มักจะสมมติให้มีการวัดหรือประมาณค่าของตัวแปรสถานะทั้งหมด แต่ในทางปฏิบัตินั้น การวัดค่าตัวแปรสถานะทั้งหมดเป็นไปไม่ได้และค่าที่ได้จากการวัดจะมีสัญญาณรบกวนและ/หรืออาจเกิดความคลาดเคลื่อนขึ้นได้ นอกจากนี้แบบจำลองมักจะไม่ทราบและต้องพบกับปัญหาในการเลือกแบบจำลองกระบวนการที่มีอยู่อย่างมากมาย ซึ่งแบบจำลองต่าง ๆ ทั้งหมดอาจจะไม่ถูกต้องก็ได้ ดังนั้นจึงได้มีการนำเทคนิคการประมาณค่ามาประยุกต์ใช้เพื่อประมาณค่าตัวแปรสถานะที่ไม่ได้วัดหรือวัดไม่ได้และเพื่อลดผลกระทบของสัญญาณรบกวน เทคนิคการประมาณค่าหรือเรียกกันทั่วไปว่า “ตัวกรอง” จะให้ค่าประมาณของค่ากระบวนการจริงจากการวัดกระบวนการที่มีสัญญาณรบกวน และค่าที่คำนวณได้จากแบบจำลองกระบวนการที่เหมาะสม ด้วยเหตุนี้ การประมาณค่าตัวแปรสถานะที่ไม่ได้วัดหรือวัดไม่ได้จึงไม่ใช่ปัญหาเล็ก ๆ แล้ว โดยเฉพาะการนำเทคนิคการประมาณค่าพารามิเตอร์ที่ไม่ทราบค่ามาใช้ ทั้งนี้เนื่องจากต้องอาศัยความรู้ความชำนาญทางคณิตศาสตร์เป็นอย่างมากเพื่อแก้ไขแบบจำลอง และการเลือกสุ่มค่าเริ่มต้นที่เหมาะสมก่อนทำการประมาณค่าพารามิเตอร์เพื่อให้การประมาณค่าประสบผลสำเร็จ

ขั้นตอนวิธีการประมาณค่าวิธีหนึ่งที่ถูกนำมาใช้กันอย่างกว้างขวาง คือตัวกรองคาลมานซึ่งเป็นขั้นตอนวิธีที่ให้ค่าประมาณของตัวแปรระบบที่กำลังถูกควบคุมโดยค่าการวัดของตัวตรวจจับที่เหมาะสม ทฤษฎีตัวกรองคาลมานเป็นเครื่องมือที่สำคัญสำหรับการวิเคราะห์และการแก้ไขปัญหาการประมาณค่าต่าง ๆ สมการตัวกรองคาลมานหรือขั้นตอนวิธีคาลมานช่วยลดขนาดของหน่วยความจำคอมพิวเตอร์โดยให้ค่าประมาณของสัญญาณที่ทันสมัยในช่วงเวลาการวัดโดยปราศจากการเก็บข้อมูลการวัดในอดีต นั่นคือตัวกรองคาลมานมีความยืดหยุ่นเนื่องจากมันสามารถจัดการกับค่าการวัดได้ครั้งละหนึ่งค่าหรือครั้งละหนึ่งชุด ตัวกรองคาลมานเป็นขั้นตอนวิธีการดำเนินการกับข้อมูลแบบเรียกตนเองที่เหมาะสมที่สุด

ทฤษฎีตัวกรองคาลมานเป็นทฤษฎีตัวกรองที่ได้รับการพัฒนาและมีการนำไปประยุกต์ใช้กันอย่างกว้างขวางในระบบเชิงเส้น โดยเฉพาะ การนำเทคนิคการกรองคาลมานไปประยุกต์ใช้กับการกำหนดเส้นทางการบิน, การควบคุมทิศทางและการควบคุม ซึ่งเป็นสาขาแรกที่ได้นำมาใช้ นอกจากนั้น ได้มีการนำไปประยุกต์ใช้กับสาขาต่าง ๆ อาทิเช่น การหาวงโคจร, การติดตามเครื่องบิน, การเคลื่อนที่ของเรือ, อุตสาหกรรมรถยนต์, การควบคุมกระบวนการทางด้านเคมี, การวัดอัตราการไหลในทันทีทันใดและการประมาณค่าและการทำนายตัวแปรที่วัดไม่ได้ในกระบวนการอุตสาหกรรม

ตัวกรองคาลมานมาตรฐานเป็นตัวกรองที่ให้ค่าประมาณเชิงเส้นไม่เอนเอียงตามลำดับที่ดีที่สุด (หรือค่าประมาณที่เหมาะสมที่สุดทั้งหมด) โดยมีสัญญาณรบกวนเป็นแบบเกาส์ ด้วยเหตุนี้กระบวนการสโตแคสติคที่ซับซ้อนจึงมักสมมติให้สัญญาณรบกวนเป็นแบบเกาส์ ซึ่งทำให้การคำนวณทางคณิตศาสตร์ของปัญหาการประมาณค่าง่ายขึ้น วิธีการประมาณค่าที่รู้จักกันดีมี 3 วิธี คือ (1) วิธีกำลังสองน้อยสุด (Ieast-square) , (2) วิธีภาวะน่าจะเป็นสูงสุด (maximum-likelihood) และ (3) วิธีเบส์ (Bayesian method) โดยทั้งสามวิธีให้ค่าประมาณที่ใกล้เคียงกัน ถึงแม้ว่าความไว้วางใจ (reliability) จะแตกต่างกัน

ปัจจัยหนึ่งที่สำคัญและทำให้ทฤษฎีการประมาณค่าและการควบคุมประสบความสำเร็จ คือ เครื่องคอมพิวเตอร์ดิจิตอลที่มีความเร็วสูงและมีขนาดหน่วยความจำมากสำหรับใช้ในการแก้ไขปัญหา

ทฤษฎีสมัยใหม่ของการประมาณค่ามีรากฐานมาจากผลงานชิ้นแรกของ A.W. Kolmogorov และ N. Wiener Kolmogorov, 1941 และ Wiener, 1942 ได้พัฒนาและกำหนดสูตรทฤษฎีการประมาณค่าขั้นพื้นฐานที่สำคัญของการประมาณค่ากำลังสองเฉลี่ยน้อยที่สุดเชิงเส้น (linear minimum mean-square) ซึ่งผลเฉลยของปัญหา Wiener-Kolmogorov เดิมเป็นลักษณะเฉพาะของการกรองสัญญาณรบกวนที่มีเวลาต่อเนื่อง ขณะที่เทคนิคการถดถอยแบบเชิงเส้น (linear regression) ซึ่งอาศัยกำลังสองน้อยสุดแบบถ่วงน้ำหนัก (weighted least-square) หรือหลักการภาวะน่าจะเป็นสูงสุด เป็นลักษณะเฉพาะของการจัดการกับปัญหาการกรองที่มีเวลาไม่ต่อเนื่อง แต่อย่างไรก็ตาม ผลเฉลยของ Wiener Kolmogorov ที่อยู่ในรูปสมการอินทิกรัลนั้นใช้ได้กับกระบวนการคงที่ (stationary) เท่านั้น จนกระทั่งต้นปี 1960 เมื่อ R.E. Kalman และต่อมา Kalman และ R. Bucy ได้ทำการพัฒนาต่อจนทำให้ตัวกรอง Kalman-Bucy ถูกนำไปประยุกต์ใช้กับระบบเชิงเส้นต่าง ๆ ได้อย่างรวดเร็วและได้พยายามที่จะนำไปประยุกต์กับระบบไม่เชิงเส้น ซึ่งผลของความพยายามทำให้มีตัวกรองคาลมานแบบยืดขยาย (extended Kalman filters) เกิดขึ้น

ตัวกรองคาลมานเป็นขั้นตอนวิธีการดำเนินการกับข้อมูลแบบเรียกตนเองที่เหมาะสมที่สุด ซึ่งให้ค่าประมาณตัวแปร (หรือสถานะ) ของระบบที่กำลังถูกควบคุมโดยการใช้ประโยชน์จากแบบจำลองของระบบ และอาศัยความแตกต่างระหว่างการทำนายจากแบบจำลองกับค่าการวัด ตัวกรองคาลมานจะจัดการกับความคลาดเคลื่อนของระบบและใช้ค่าการวัดที่เหมาะสมทั้งหมดโดย

ไม่คำนึงถึงความถูกต้องของมันในการประมาณค่าปัจจุบันของตัวแปรที่สนใจ โดยอาศัยข้อเท็จจริงดังต่อไปนี้คือ (1) ความรู้เกี่ยวกับแบบจำลองระบบและพลศาสตร์ของเครื่องมือการวัด, (2) ค่าสถิติของสัญญาณรบกวนและความไม่แน่นอน และ (3) ข้อมูลเกี่ยวกับเงื่อนไขตอนเริ่มต้นของตัวแปร หน้าที่หลักของตัวกรองคาลมาน คือการประมาณค่าเวกเตอร์สถานะโดยใช้ตัวตรวจจับระบบและข้อมูลการวัดที่มีสัญญาณรบกวน ขั้นตอนวิธีตัวกรองคาลมานถูกนำไปใช้กับปัญหาการประมาณค่า 2 ชนิด คือ (1) การกรอง (หรือการปรับปรุงให้ทันสมัย) และ (2) การทำนาย (หรือการแผ่กระจาย) เมื่อเวลาของค่าประมาณของเวกเตอร์ที่ต้องการตรงกับค่าการวัดในอดีต เราเรียกปัญหาการประมาณค่านี้ว่า *การกรอง (filtering)* หรืออาจกล่าวได้ว่า การกรอง คือการประมาณค่าเวกเตอร์

วิทยานิพนธ์ฉบับนี้ปิดท้ายด้วยภาคผนวก 5 บท ภาคผนวก ก ประกอบด้วยพื้นฐานความรู้ทางคณิตศาสตร์เกี่ยวกับทฤษฎีระบบ ภาคผนวกนี้นำมาใส่ไว้เพื่อเป็นการทบทวนความรู้สำหรับผู้อ่านที่สนใจ ซึ่งจำเป็นสำหรับการทำความเข้าใจในระบบสโตแคสติคพลวัต หัวข้อที่กล่าวถึงในภาคผนวกนี้เป็นพื้นฐานทางคณิตศาสตร์เกี่ยวกับเวกเตอร์, เมทริกซ์, ปริภูมิเวกเตอร์และทฤษฎีเซต, ระบบเชิงเส้น และผลเฉลยของสมการอนุพันธ์เชิงเส้น ภาคผนวก ข กล่าวถึงสัญญาณสุ่ม ซึ่งประกอบด้วยพื้นฐานที่จำเป็นสำหรับสัญญาณสุ่ม (สโตแคสติค) โดยเริ่มต้นด้วยคำจำกัดความเกี่ยวกับปริมาณต่าง ๆ เกี่ยวกับสัญญาณสุ่มและการรวมกันของสัญญาณสุ่ม ภาคผนวก ค บทพิสูจน์ ภาคผนวก ง แสดงวิธีการใช้โปรแกรม MATHLAB ซึ่งอธิบายขั้นตอนในการใช้งานและช่วยให้การใช้งานสะดวกขึ้น ภาคผนวก จ. ผลการวิเคราะห์ด้วยโปรแกรม SPSS

2.1 เอกสารและผลงานที่เกี่ยวข้องกับตัวกรองคาลมาน

Kolmogorov, 1930 ได้กำหนดสูตรทฤษฎีของสัจพจน์ความน่าจะเป็น (axiomatic theory of probability) ซึ่งเป็นการค้นพบทฤษฎีของกระบวนการสโตแคสติค และในปี 1941 เขาและ Wiener ได้สร้างทฤษฎีการประมาณค่า (estimation theory) สำหรับระบบพลวัต ในต้นศตวรรษที่ 18 ตระกูลของ Bernoulli, Bayes, Gauss และคนอื่น ๆ ได้ตีพิมพ์บทความเรื่องทฤษฎีความน่าจะเป็นและวิธีกำลังสองน้อยสุดของการประมาณค่าพารามิเตอร์

Japan International Cooperation Agency (JICA) ทำการศึกษา โครงข่ายโทรคมนาคม (telecommunication Network) ในเขตนครหลวงของไทยในปี 1992 ในส่วนของการคาดคะเนปริมาณการใช้โทรศัพท์แยกเป็น 2 ขั้นตอน คือ

1. การคาดคะเนภาพรวม (macro forecast) คาดคะเนปริมาณการใช้โทรศัพท์รวมในเขตนครหลวง นครปฐม สมุทรสาคร และอยุธยา แบ่งเป็น

1.1 ปริมาณการใช้โทรศัพท์ภายในท้องถิ่น (local traffic)

1.2 ปริมาณการใช้โทรศัพท์ทางไกลภายในประเทศ (long distance traffic)

1.3 ปริมาณการใช้โทรศัพท์ทางไกลระหว่างประเทศ (international traffic)

พบว่าปริมาณการใช้โทรศัพท์ในเขตดังกล่าว มีความสัมพันธ์กับอัตราการเจริญเติบโตของผู้เช่าในเขตนั้นๆในรูปสมการ logaritm และปริมาณการใช้โทรศัพท์ทางไกลในประเทศและทางไกลระหว่างประเทศ มีความสัมพันธ์กับค่าอัตราการเจริญเติบโตในรูปสมการเส้นตรง

2. การคาดคะเนรายชุมสาย (micro forecast)

พิจารณาปริมาณการใช้โทรศัพท์ต่อผู้เช่า (traffic per subscriber) ในแต่ละชุมสาย พบว่าอยู่ในช่วง 0.02-0.09 Erang/sub แล้วแต่ว่าอยู่ในเขตธุรกิจ หรือ เขตที่อยู่อาศัย พบว่า ถ้าอยู่ในเขตธุรกิจ ค่าปริมาณการใช้โทรศัพท์ต่อผู้เช่าจะมีค่าสูง

การศึกษาของ JICA ให้แนวคิดในการคาดคะเน ทั้งด้านภาพรวมและรายชุมสายโดยใช้ Logistic Model ในการคาดคะเนอัตราการเจริญเติบโตของทราฟฟิค

2. SCANLAN สัมมนาเรื่อง Modern Methods of Telecommunication Planning ของ ASIA-PASIFIC TELECOMMUNITY (APT) ในปี 1989 ได้สรุปกระบวนการคาดคะเนปริมาณการใช้โทรศัพท์ว่าปริมาณการใช้โทรศัพท์ ขึ้นอยู่กับ ผลิตภัณฑ์ประชาชาติภายในประเทศ,ปริมาณการค้า,การอพยพย้ายถิ่น,ค่าบริการโทรศัพท์ และอื่นๆ ซึ่งทำให้ผู้ศึกษามองเห็นภาพรวมของปริมาณการใช้โทรศัพท์ว่า ขึ้นอยู่กับปัจจัยทางเศรษฐกิจ และสามารถนำไปใช้ในการวางแผนต่อไป

Chambers และคณะ ในปี 1971 ศึกษาใน Harvard Business Review Planning Seris : Part IV เดือน กรกฎาคม-สิงหาคม 1971 หัวข้อเรื่อง “ How to Choose the Right Forecasting Technique” ซึ่งสรุปเทคนิคการคาดคะเนโดยแสดง คุณลักษณะ (characteristics) ของแต่ละวิธี , การแสดงจุดวกกลับ (turning points), รายละเอียดการนำไปใช้ (Typical applications) ข้อมูลที่ต้องการ,ค่าใช้จ่ายในการคาดคะเน,ระยะเวลาในการคาดคะเน พบว่าวิธีการที่เหมาะสมต่อการคาดคะเนปริมาณการใช้โทรศัพท์ คือ ใช้วิธีการทางเศรษฐมิติ ซึ่งเป็นระบบของสมการถดถอยซึ่งขึ้นอยู่กับกันและกัน ค่าพารามิเตอร์ จะถูกประมาณการไปพร้อมๆกัน โดยใช้ข้อมูลการวิเคราะห์ในอดีต ซึ่งเป็นการวิเคราะห์เปรียบเทียบค่าอัตราการเจริญเติบโตจากจุดเริ่มต้น โดยการคาดคะเนจะขึ้นอยู่กับความคล้ายคลึงกันของรูปแบบข้อมูล แล้วนำข้อมูลเหล่านั้นมาทำการวิเคราะห์สมการถดถอย ซึ่งเป็นการแสดงความสัมพันธ์ระหว่างตัวแปรตาม และตัวแปรทางเศรษฐกิจ และคาดคะเนโดยใช้วิธีการกำลังสองน้อยที่สุด ซึ่งแสดงโดยค่าทางสถิติ ความสัมพันธ์จะถูกเลือกโดยการทดสอบอย่างมีเหตุผล ซึ่งการศึกษาของ Chambers และคณะ ทำให้ผู้ศึกษาวิธีการคาดคะเนคัดเลือกวิธีการคาดคะเนได้อย่างเหมาะสมกับลักษณะงาน และข้อมูลที่มีอยู่ โดยมีความถูกต้องแม่นยำ

นิตย์ ฝาม (2528) ศึกษาเกี่ยวกับตัวแบบพยากรณ์อุปสงค์การบริการโทรศัพท์ในประเทศไทยในอนาคต เพื่อดูแนวโน้มความต้องการบริการทางโทรศัพท์ว่ามีการเปลี่ยนแปลงอย่างไร โดยใช้ปัจจัยที่เป็นรายปีตั้งแต่ พ.ศ. 2521-2527 การพยากรณ์เป็นแบบ break down method คือ เลือกหาตัวแบบที่เหมาะสมที่สุดในแต่ละเขตย่อย ทั้งนครหลวงและภูมิภาคแล้วนำค่าพยากรณ์แต่ละเขตย่อยมารวมกันเป็นค่าพยากรณ์ทั่วประเทศ โดยวิธีที่ใช้พยากรณ์แต่ละเขตประกอบด้วย วิธีวิเคราะห์ความถดถอย วิธีค่าเฉลี่ยเคลื่อนที่แบบเส้นตรง เทคนิคทำให้เรียบแบบเอกซ์โปเนนเชียล 2 ขั้น เทคนิคการทำให้เรียบแบบเอกซ์โปเนนเชียลแบบ 2 พารามิเตอร์(Holt) ผลการคาดคะเนพบว่าในปี 2533 มีความต้องการโทรศัพท์ทั้งสิ้น 1,989,961 เลขหมาย คิดเป็นอัตราการขยายตัวภายใน 6 ปี ร้อยละ 67.2 โดยผลการศึกษาชุดนี้สามารถนำมาใช้ประโยชน์ในการศึกษาตัวแบบพยากรณ์รูปแบบต่าง ๆ

กิตติ สุตันติวณิชยกุล ศึกษาปัญหาการตั้งแบบจำลองในการประมาณสมการเศรษฐมิติในปี 2532 เพื่อศึกษาผลที่เกิดขึ้นในทางปฏิบัติในการเลือกตัวแปรอิสระที่ถูกต้องและไม่ถูกต้องในการพยากรณ์

ประภาศรี ปทุมรัตน์ ศึกษาปัจจัยทางเศรษฐกิจที่มีผลกระทบต่อปริมาณการใช้โทรศัพท์ในเขตนครหลวง และทำการคาดคะเนปริมาณการใช้โทรศัพท์ในเขตนครหลวงทั้งภาพรวมและภาพรายชุมสาย ในปี 2537 ผลการศึกษาพบว่า วิธีการปรับข้อมูลที่เหมาะสม คือ เติมค่าปริมาณการใช้โทรศัพท์โดยทดสอบการเปลี่ยนแปลงโครงสร้างของข้อมูล ปริมาณการใช้โทรศัพท์ ขึ้นอยู่กับดัชนีราคาผู้บริโภคในเขตนครหลวง ปริมาณรถยนต์ที่จดทะเบียน และปริมาณที่อยู่อาศัยที่เพิ่มขึ้น

วารุณี ตรีบำรุงศักดิ์ (2538) ศึกษาการพยากรณ์ด้วยวิธีการถดถอยเชิงเส้นพหุเมื่อตัวแปรตามมีค่าสูญหาย การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการประมาณค่าสูญหายของตัวแปรตาม ด้วยวิธีค่าเฉลี่ย วิธีสมการถดถอย วิธีอีเอ็ม (EM Algorithm) และวิธีการของฮันท์ (Hunt's Method) ข้อมูลที่ใช้ในการวิจัยได้จากการจำลองด้วยเทคนิคมอนติคาร์โล และทำการทดลองซ้ำๆกัน 200 รอบ สำหรับแต่ละสถานการณ์ที่กำหนดเพื่อประมาณค่าที่สูญหาย และหารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง(RMSE) ของค่าพยากรณ์ด้วยวิธีการทั้ง 4 วิธี ซึ่งผลการวิจัยสรุปได้ว่าถ้าข้อมูลมีขนาดเล็ก (10-20) ค่าเบี่ยงเบนมาตรฐานขนาดไม่ใหญ่นัก วิธีการของฮันท์จะให้ค่าคลาดเคลื่อนต่ำกว่าวิธีอื่นๆ แต่ถ้าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีขนาดเพิ่มขึ้น วิธีค่าเฉลี่ยจะให้ค่าความคลาดเคลื่อน RMSE ของค่าพยากรณ์ต่ำกว่าวิธีอื่น ส่วนในสถานการณ์ที่ขนาดตัวอย่างขนาดใหญ่วิธีสูญหายจะเหมาะสมเกือบทุกกรณี นั่นคือถ้าขนาดตัวอย่างใหญ่พอ การตัดชุดข้อมูลสูญหายทิ้งจะมีผลกระทบน้อยมากกับผลการวิเคราะห์การถดถอยเชิงเส้นพหุ โดยใช้วิธีกำลังสองน้อยที่สุด

สหพร กลัดนิ่ม (2541) ศึกษาปัจจัยทางเศรษฐกิจที่มีผลกระทบต่อปริมาณการใช้โทรศัพท์ทางไกลระหว่างประเทศ โดยใช้วิธีการ Multiple Linear Regression การศึกษาวิเคราะห์ได้ใช้

ข้อมูลปัจจัยต่างๆเพื่อหาตัวแปรที่เหมาะสมในการสร้างสมการแบบจำลองสำหรับการคาดคะเนค่าทราฟฟิคในอนาคตล่วงหน้า 4 ปี เพื่อนำมาคำนวณหาจำนวนวงจรและจำนวนอุปกรณ์ DTI ที่เหมาะสม

จากงานวิจัยที่ผ่านมา จะเห็นว่าเทคนิคการพยากรณ์ความต้องการใช้โทรศัพท์โดยใช้ตัวกรองคาลมานไม่มีผู้ใดศึกษา ดังนั้นผู้ศึกษาจึงได้เล็งเห็นถึงความสำคัญของการนำตัวกรองคาลมานไปประยุกต์ใช้เนื่องจากมีขบวนการในการคำนวณง่ายกว่าและใช้ตัวแปรไม่มากนัก ตัวกรองคาลมานได้ถูกพัฒนามากขึ้นสามารถนำไปประยุกต์ใช้กับระบบไม่เชิงเส้นได้ แต่เนื่องจากตัวกรองคาลมานเป็นเทคนิคการประมาณค่าสถานะและพารามิเตอร์แบบเรียกตัวเอง ซึ่งเป็นการทำงานแบบวนซ้ำและมีการคำนวณทางคณิตศาสตร์ที่ค่อนข้างซับซ้อน และจำเป็นต้องอาศัยเครื่องคอมพิวเตอร์เข้ามาช่วยในการคำนวณ

ตัวกรองทำงานได้ดีที่สุดในระบบเชิงเส้น และสมมติให้สัญญาณรบกวนระบบที่มีผลต่อตัวกรองเป็นสีขาวและแบบเกาส์

คุณสมบัติของตัวกรองที่นำไปใช้กับแบบจำลองการประมาณค่าสรุปได้ดังนี้

1. ที่เวลา t ตัวกรองให้ค่าประมาณ $\hat{X}(t)$ ที่ไม่เอนเอียงของเวกเตอร์สถานะ X นั่นคือค่าที่คาดหมายของค่าประมาณเป็นค่าของเวกเตอร์สถานะที่เวลา t

2. ค่าประมาณ เป็นค่าประมาณที่มีความแปรปรวนน้อยที่สุด

3. ตัวกรองเป็นแบบเรียกตนเอง (recursive) หมายถึงไม่มีการเก็บข้อมูลในอดีต

4. ตัวกรองเป็นเชิงเส้นหรือต้องทำให้เป็นเชิงเส้น ซึ่งการทำให้เป็นเชิงเส้นช่วยให้การคำนวณง่ายขึ้นและสะดวกต่อการคำนวณด้วยเครื่องคอมพิวเตอร์

ในการประยุกต์ใช้ทฤษฎีตัวกรองคาลมาน เราสร้างแบบจำลองตามข้อสมมติต่อไปนี้

1. เวกเตอร์สถานะ X(t) คงอยู่ในสภาวะแวดล้อมแบบสุ่ม (คือ พลศาสตร์ระบบ) ซึ่งเป็นแบบเกาส์มีค่าเฉลี่ยเป็นศูนย์ และเมทริกซ์ความแปรปรวนร่วม Q(t) ที่เวลา t

2. เวกเตอร์สถานะที่ไม่ทราบค่าสามารถประมาณค่าได้โดยใช้การสังเกตหรือตัวอย่างข้อมูลที่เป็นฟังก์ชันของเวกเตอร์สถานะ

3. การสังเกตที่จุดเวลา t ถูกรบกวนโดยสัญญาณรบกวนแบบเกาส์ที่ไม่สหสัมพันธ์กัน (uncorrelated) มีค่าเฉลี่ยเป็นศูนย์ และเมทริกซ์ความแปรปรวนร่วม R(t) ในระบบเชิงพลวัต

สมการเวลาไม่ต่อเนื่องเชิงเส้นมีรูปแบบเหมือนกับแบบจำลองสถานะของตัวกรองคาลมานที่มีเวลาต่อเนื่อง สามารถหาได้จากแผนแบบต่าง ๆ คือ (1) เทคนิคออยเลอร์, (2) ระเบียบวิธีซิมป์สัน และ (3) ระเบียบวิธีสี่เหลี่ยมคางหมู (trapezoidal method) ในที่นี้ เราขอกล่าวถึงเฉพาะเทคนิคประมาณค่าออยเลอร์แบบข้างหน้า โดย

$$x(t_k) \approx \frac{x(t_{k+1}) - x(t_k)}{t_{k+1} - t_k}$$

เมื่อ $x(t_k)$ เป็นค่าข้อมูลสุ่มตัวอย่างของสัญญาณเวลาต่อเนื่อง $x(t)$

แทนสมการข้างต้นลงในแบบจำลองสถานะเวลาต่อเนื่อง จะได้

$$\frac{x(t_{k+1}) - x(t_k)}{t_{k+1} - t_k} \approx A(t_k)x(t_k) + B(t_k)u(t_k)$$

ให้ $h = t_{k+1} - t_k = \Delta t$ แล้วแบบจำลองเวลาไม่ต่อเนื่องจะมีรูปแบบเป็น

$$x(t_{k+1}) \approx [I + hA(t_k)]x(t_k) + hB(t_k)u(t_k)$$

ให้ $x(t_k)$ แทน $x(k)$ แล้วเมทริกซ์ A คือ

$$A(k) = I + hA(t_k)$$

และเมทริกซ์ B เป็น

$$B(k) = hB(t_k)$$

ดังนั้นเราสามารถเขียนแบบจำลองสถานะที่มีเวลาไม่ต่อเนื่องแปรเปลี่ยนตามเวลาแบบเชิงเส้นเป็น

$$x(k+1) = A(k)x(k) + B(k)u(k)$$
$$y(k) = C(k)x(k) + D(k)u(k)$$

สำหรับระบบที่แปรเปลี่ยนตามเวลา เราสามารถเขียนตัวกรองคาลมานในรูปแบบไม่ต่อเนื่องแบบเรียกตนเองได้ดังนี้

$$x(k+1) = A(k+1,k)x(k) + T(k)\xi(k),\ \text{k=0,1,2,...} \tag{2.1}$$

$\xi(k) \sim \mathrm{N}(0, Q(k))$ (ลำดับแบบเกาส์สีขาวที่มี)

$$y(k) = C(k)x(k) + n(k) \tag{2.2}$$

$n(k) \sim N(0, R(k))$ (ลำดับแบบเกาส์สีขาว)

$$\hat{X}(k+1|k) = A(k+1,k)\hat{x}(k|k) \tag{2.3}$$

$$P(k+1|k) = A(k+1,k)P(k|k)A^T(k+1,k) + \Gamma(k)Q(k+1)\Gamma^T(k) \tag{2.4}$$

$$\hat{X}(k|k) = \hat{X}(k|k-1) + K(k)\left[y(k) - C(k)\hat{x}(k|k-1)\right] \tag{2.5}$$

$$P(k|k) = [I - K(k)C(k)]P(k|k-1) \tag{2.6}$$

หรือ

$$P(k|k) = P(k|k-1) - K(k)C(k)P(k|k-1) \tag{2.7}$$

หรือ

$$P(k|k) = [I - K(k)C(k)]P(k|k-1)[I - K(k)C(k)]^T + K(k)RK^T(k) \tag{2.8}$$

$$K(k) = P(k|k-1)C^T(k)\left[C(k)P(k|k-1)C^T(k) + R(k)\right]^{-1} \tag{2.9}$$

เมื่อดัชนีบน T แทนทรานสโพสของเมทริกซ์

สมการ (2.1) เรียกว่าสมการสถานะเป็นสมการผลต่างสืบเนื่องอันดับหนึ่งที่เวลา k ซึ่ง x ค่าหนึ่งคือ x(k) สัมพันธ์กับค่า x(k+1) ถัดไป เวกเตอร์ แทนพารามิเตอร์หรือเวกเตอร์สถานะที่ขึ้นอยู่กับเวลา สมการ (2.1) เป็นแบบจำลองของสถานะจริง สมการ (2.2) เรียกว่าสมการการวัดเป็นแบบจำลองของกระบวนการการวัด y(k) แทนเวกเตอร์การวัดซึ่งประกอบด้วยค่าการวัดสเกลาร์แต่ละค่าที่เวลา k นอกจากนี้ สมการ (2.2) สัมพันธ์กับค่าการวัดเหล่านี้และเวกเตอร์สถานะโดยผ่านเมทริกซ์การวัด C(k) เมื่อ $k \geq 1$ และแทนสัญญาณรบกวนแบบสุ่มด้วย $\eta(k)$ เมื่อ $k \geq 1$

สมการ (2.4) แทนการแผ่กระจายความแปรปรวนร่วม พิจารณาสมการ (2.5) สมการกล่าวว่าค่าประมาณที่ดีที่สุดที่ t=k เท่ากับค่าประมาณทำนายบวกกับความคลาดเคลื่อน (หรือความแตกต่าง) ระหว่างค่าสังเกตและค่าสังเกตประมาณคูณด้วยแฟกเตอร์ถ่วงน้ำหนัก (หรือเกน) K(k)

K(k) เปลี่ยนแปลงตามเวลาและต้องการให้มีค่าเหมาะสมที่สุดเสมอ ซึ่งขึ้นอยู่กับลักษณะของสัญญาณรบกวนการวัด R(k) อย่างมาก ด้วยเหตุนี้ ถ้า R(k)=1 แล้วสมาชิกของ K(k) จะมีค่าสัมบูรณ์ลดลงอย่างแน่นอน ถ้าการวัดที่ได้มีสัญญาณรบกวนการวัดเล็กน้อย เมื่อเทียบกับความคลาดเคลื่อนครั้งก่อน (นั่นคือ ถ้า R มีค่าน้อยเมื่อเทียบกับ $C(kP(k|k-1)C^T(k))$) แล้วเกนจะมีค่ามาก และถ้าการวัดถูกรบกวนมากเมื่อเทียบกับความคลาดเคลื่อนครั้งก่อน (นั่นคือ ถ้า R มากเมื่อเทียบกับ $C(kP(k|k-1)C^T(k))$) แล้วเกนมีค่าน้อยและตัวกรองจะสนใจการวัดเล็กน้อย

การประยุกต์ใช้ตัวกรองเชิงเส้นกับระบบจำเพาะ เราต้องระบุแบบจำลองเชิงพลวัต (A, Γ, C) ค่าสถิติของสัญญาณรบกวน (Q,R) และข้อมูลค่าก่อน $\left(\hat{x}(0), P(0)\right)$ ในกรณีเวลาต่อเนื่องสมมติให้ค่าสถิติครั้งก่อนของกระบวนการสัญญาณรบกวน $\xi(t)$ และ $\eta(t)$ มีค่าเฉลี่ยศูนย์และสีขาว

2.2 ระเบียบวิธีการพยากรณ์

เพื่อให้เห็นถึงแนวความคิดของระเบียบวิธีการพยากรณ์เชิงปริมาณ และเป็นพื้นฐานที่จะศึกษาในรายละเอียดในบทต่อไป จะขอกล่าวถึงระเบียบวิธีการพยากรณ์ที่นิยมใช้กันในปัจจุบันอย่างย่อ ๆ ดังนี้

2.2.1 เทคนิคการทำให้เรียบ (Smoothing Techniques)

ระเบียบวิธีการพยากรณ์ส่วนใหญ่ภายใต้เทคนิคการทำให้เรียบได้รับการพัฒนาขึ้นจากตัวแบบการพยากรณ์ 2 ตัวแบบ คือ ตัวแบบการพยากรณ์แบบค่าเฉลี่ยคงที่ (constant mean model) และตัวแบบพยากรณ์แบบแนวโน้มเชิงเส้น (linear trend model) ตัวแบบการพยากรณ์ทั้ง 2 แบบดังกล่าวอาจมีหรือไม่มีฤดูกาลก็ได้ ในตัวแบบการพยากรณ์แบบค่าเฉลี่ยคงที่ วิธีการทำให้เรียบที่นิยมใช้กันมากคือวิธีค่าเฉลี่ยเคลื่อนที่ (moving average method) และการทำให้เรียบแบบเอกซ์โปเนนเชียล (exponential smoothing) เป็นวิธีการที่เหมาะสำหรับใช้ในการพยากรณ์ระยะสั้น และเหมาะสำหรับพยากรณ์ค่าของตัวแปรที่มักจะมีการเปลี่ยนแปลงไม่มากนักในหนึ่งหน่วยเวลา ค่าเฉลี่ยเคลื่อนที่ คือค่าเฉลี่ยเชิงเลขคณิต (arithmetic mean) ของข้อมูลในอดีต n ข้อมูล และใช้ค่านี้เป็นค่าพยากรณ์ซึ่งเขียนเป็นสมการคณิตศาสตร์ได้ดังนี้

$$\hat{X}_t(1) = \frac{X_t + X_{t-1} + \ldots + X_{t-n+1}}{n} \tag{2.10}$$

โดย $\hat{X}_t(1)$ เป็นค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าของ x โดยทำการพยากรณ์ ณ t เวลา และ X_t เป็นค่าสังเกต (observed value) ของ ณ เวลา t

เมื่อสิ้นสุดเวลา $t+1$ ค่าของตัวแปร x ณ เวลา $t+1$ จะเข้ามาโดยจะนำมาใช้แทน X_{t-n+1} ซึ่งเป็นข้อมูลที่ล้าสมัยที่สุด ดังนั้นค่าพยากรณ์ครั้งต่อไปคือ

$$\hat{X}_{t+1}(1) = \frac{X_{t+1} + X_t + X_{t-1} + \ldots + X_{t-n+2}}{n} \tag{2.11}$$

จะเห็นว่าการพยากรณ์โดยวิธีค่าเฉลี่ยเคลื่อนที่ใช้ข้อมูลในอดีตที่เพิ่งผ่านมา n หน่วยเวลาเท่านั้น และค่าพยากรณ์เป็นค่าเฉลี่ยเชิงเลขคณิต จึงทำให้วิธีการนี้มีจุดอ่อนที่สำคัญ 2 ประการ คือ

1. ข้อมูลในอดีตที่ผ่านมามากกว่า n หน่วยเวลา จะไม่ได้รับการพิจารณาเลย ดังนั้นค่าของ n จึงควรจะมากหน่อย เพื่อจะได้ลักษณะการเปลี่ยนแปลงที่สำคัญอยู่ในขบวนการพยากรณ์ด้วย แต่ในขณะเดียวกันถ้าใช้จำนวนข้อมูลในอดีตมากเกินไปก็จะทำให้ความเฉื่อย (inertia) ของการพยากรณ์สูง ซึ่งหมายความว่าค่าพยากรณ์อาจปรับตัวไม่ทันต่อการเปลี่ยนแปลงที่เกิดขึ้น

2. การที่ใช้ค่าเฉลี่ยเชิงเลขคณิตเป็นค่าพยากรณ์อาจมีผู้ท้วงติงได้มากเพราะโดยทั่ว ๆ ไป “สาระ” (information) ของข้อมูล ณ เวลา t อาจมีมากกว่า “สาระ” ของข้อมูล ณ เวลา $t-n+1$

เมื่อทำการพยากรณ์หนึ่งหน่วยเวลาล่วงหน้า ณ เวลา t ดังนั้น การให้น้ำหนักในการเฉลี่ยเท่ากันหมดอาจไม่ถูกต้องก็เป็นได้

การทำให้เรียบแบบเอกซ์โปเนนเชียลได้แก้ไขจุดอ่อนบางอย่างของวิธีค่าเฉลี่ยเคลื่อนที่ โดยใช้ค่าพยากรณ์เป็นค่าเฉลี่ยของข้อมูลในอดีตทั้งหมดด้วยน้ำหนักในการเฉลี่ยที่ค่อย ๆ ลดลงเมื่อเวลาของข้อมูลห่างออกไปจากปัจจุบันมากขึ้น ค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าซึ่งกระทำ ณ เวลา t จะเขียนเป็นสมการเชิงคณิตศาสตร์ได้ดังนี้

$$\hat{X}_t(1) = \alpha X_t + \alpha(1-\alpha)X_{t-1} + \alpha(1-\alpha)^2 X_{t-2} + \alpha(1-\alpha)^3 X_{t-3} + \ldots \qquad (2.12)$$

โดย α เป็นค่าคงที่การทำให้เรียบ (smoothing constant) มีค่าระหว่าง 0 กับ 1 ดังนั้น น้ำหนักการเฉลี่ยของข้อมูล X_{t-i} ซึ่งเท่ากับ $\alpha(1-\alpha)^i$ จะมีค่าลดลงเมื่อ i มีค่าสูงขึ้น เพื่อให้เกิดความเข้าใจเกี่ยวกับวิธีการทำให้เรียบแบบเอกซ์โปเนนเชียลมากขึ้นสมการ (2.12) อาจเขียนในรูปใหม่ได้เป็น

$$\begin{aligned}\hat{X}_t(1) &= \alpha X_t + (1-\alpha)\hat{X}_{t-1}(1) \\ &= \hat{X}_{t-1}(1) + \alpha\left(X_t - \hat{X}_{t-1}(1)\right) \qquad (2.13)\end{aligned}$$

ซึ่งจะเห็นว่าค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าที่กระทำ ณ เวลา t มีค่าเท่ากับค่าเฉลี่ยโดยน้ำหนักระหว่างข้อมูลกับค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าที่ได้กระทำ ณ เวลา $t-1$ หรือเท่ากับค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าที่กระทำ ณ เวลา $t-1$ บวกกับผลคูณระหว่าง α กับความคลาดเคลื่อนในการพยากรณ์ที่กระทำ ณ เวลา $t-1$

หาก α มีค่าใกล้ 1 ค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าจะมีค่าใกล้เคียงกับค่าปัจจุบันของตัวแปร และหาก α มีค่าใกล้ 0 ค่าพยากรณ์จะไม่คำนึงถึงความผิดพลาดในการพยากรณ์ ณ เวลาที่เพิ่งผ่านมามากนัก ถึงแม้การกำหนดค่าของ α ในการพยากรณ์ไม่อาจกระทำได้โดยง่ายแต่วิธีทำให้เรียบแบบเอกซ์โปเนนเชียลเป็นวิธีที่นิยมใช้กันมากในวงการธุรกิจเพราะเป็นวิธีที่ง่าย และผู้รับผิดชอบอาจใช้ประสบการณ์ช่วยกำหนดค่าของ α ก็ได้

ไม่ว่าจะเป็นการทำให้เรียบแบบค่าเฉลี่ยเคลื่อนที่หรือแบบเอกซ์โปเนนเชียล เมื่อข้อมูลมีแนวโน้ม ค่าพยากรณ์จะมีลักษณะที่สำคัญดังนี้ ค่าพยากรณ์จะมีแนวโน้มให้ค่าต่ำกว่าความเป็นจริงในกรณีที่ข้อมูลมีแนวโน้มสูงขึ้น และค่าพยากรณ์จะมีแนวโน้มให้ค่าสูงกว่าความเป็นจริงในกรณีที่ข้อมูลมีแนวโน้มลดลง หากพิจารณากันอย่างผิวเผินแล้ว อาจกล่าวว่าการทำให้เรียบแบบค่าเฉลี่ยเคลื่อนที่หรือแบบเอกซ์โปเนนเชียล มิอาจนำไปประยุกต์ใช้กับข้อมูลที่มีแนวโน้มได้ แต่หากพิจารณากันให้ลึกซึ้งแล้วจะเห็นว่า จุดอ่อนเหล่านี้กลับกลายเป็นจุดเด่นของระเบียบวิธีที่อาจใช้ในการชี้แนะว่า ข้อมูลมีแนวโน้มสูงขึ้นหรือลดลงจริง มิใช่เป็นการเปลี่ยนแปลงชั่วคราวเนื่องจากการ

รบกวน เช่น ในกรณีการทำให้เรียบแบบค่าเฉลี่ยเคลื่อนที่ เส้นพยากรณ์เมื่อ n มีค่ามาก จะมีความเฉื่อยสูงกว่าเส้นพยากรณ์เมื่อ n มีค่าน้อย ดังนั้น ในขณะที่ข้อมูลมีแนวโน้มเพิ่มขึ้น ในบางครั้งข้อมูลอาจมีค่าต่ำกว่าเส้นพยากรณ์ที่ใช้จำนวนพจน์น้อย แต่ถ้ามีข้อมูลมีค่าลดลงมาต่ำกว่าเส้นพยากรณ์ที่ใช้จำนวนเทอมมาก จะเป็นการชี้แนะว่าข้อมูลกำลังมีแนวโน้มลดลงและในทำนองเดียวกันในขณะที่ข้อมูลมีแนวโน้มลดลง แต่ข้อมูลมีค่าสูงกว่าเส้นพยากรณ์ที่ใช้จำนวนเทอมมาก จะเป็นการชี้แนะว่าแนวโน้มของข้อมูลได้เปลี่ยนทิศทางแล้วกลับกลายเป็นเพิ่มขึ้น

อย่างไรก็ตาม หากข้อมูลที่มีแนวโน้มเป็นเชิงเส้น ตัวแบบการพยากรณ์ที่จะใช้อธิบายข้อมูลก็ควรเป็นตัวแบบการพยากรณ์แบบแนวโน้มเชิงเส้น การพยากรณ์อาจกระทำโดยเทคนิคการทำให้เรียบแบบเอกซ์โปเนนเชียลสองชั้น (double exponential smoothing) หรือเทคนิคการเฉลี่ยเคลื่อนที่สองชั้น (double moving average) นอกจากนี้ยังมีระเบียบวิธีของ Holt ซึ่งให้ความถูกต้องของค่าพยากรณ์สูงในกรณีที่ข้อมูลมีแนวโน้ม โดยใช้แนวความคิดของการทำให้เรียบแบบเอกซ์โปเนนเชียล การพยากรณ์ระดับ (level) ของข้อมูล และความชัน (slope) ซึ่งอาจเขียนได้ดังนี้

$$\hat{b}_{0t} = \alpha X_t + (1-\alpha)\hat{X}_{t-1}(1) \tag{2.14}$$

$$\hat{b}_{1t} = \beta\left[\hat{b}_{0t} - \hat{b}_{0(t-1)}\right] + (1+\beta)\hat{b}_{1(t-1)} \tag{2.15}$$

$$\hat{X}_t(1) = \hat{b}_{0t} + \hat{b}_{1t} \tag{2.16}$$

จะเห็นได้ว่า ทั้งระดับ $\hat{b}_{0t}$ และความชัน $\hat{b}_{1t}$ จะถูกปรับเปลี่ยนแปลงไปตามข้อมูลล่าสุดที่เข้ามาในกรณีที่ข้อมูลมีฤดูกาล Winters ได้แนะนำให้วิเคราะห์โดยใช้ตัวแบบเชิงคูณ

$$X_t = (b_0 + b_1 t)C_t + \varepsilon_t \tag{2.17}$$

หรือตัวแบบเชิงบวก

$$X_t = b_0 + b_1 t + C_t + \varepsilon_t \tag{2.18}$$

โดย b_0 เป็นส่วนประกอบคงที่ (permanent component) b_1 เป็นส่วนประกอบแนวโน้มเชิงเส้น (linear trend component) C_t เป็นปัจจัยฤดูกาลเชิงคูณ (multiplicative seasonal factor) หรือปัจจัยฤดูกาลเชิงบวก (additive seasonal factor) แล้วแต่กรณี และ ε_t เป็นความรบกวนเชิงสุ่ม แนวทางที่จะเลือกตัวแบบเชิงคูณหรือเชิงบวกนั้น อาจพิจารณาได้โดยสังเกตลักษณะของข้อมูลว่าการเปลี่ยนแปลงในแต่ละฤดูกาลมีขนาดคงที่หรือแปรเปลี่ยนไปตามปัจจัยแนวโน้ม ถ้าการเปลี่ยนแปลงภายในแต่ละฤดูกาลมีขนาดที่ค่อนข้างคงที่ ก็ควรเลือกตัวแบบเชิงบวก แต่ถ้าลักษณะการเปลี่ยนแปลงภายในฤดูกาลแปรเปลี่ยนไปตามปัจจัยแนวโน้ม ก็ให้เลือกตัวแบบเชิงคูณ

การปรับค่าส่วนประกอบคงที่ ส่วนประกอบแนวโน้มเชิงเส้น และปัจจัยฤดูกาลยังคงใช้เทคนิคการทำให้เรียบแบบเอกซ์โปเนนเชียล ซึ่งจะเห็นได้ว่ามีพารามิเตอร์อยู่ทั้งหมด 3 ตัวที่เกี่ยวข้องกับการทำให้เรียบที่จะต้องพิจารณาค่ารายละเอียดเกี่ยวกับระเบียบวิธีการพยากรณ์ของ Winters

2.2.2 การพยากรณ์แบบปรับได้ (Adaptive Forecasting)

ในการพยากรณ์แบบการทำให้เรียบแบบเอกซ์โปเนนเชียล จะเห็นว่าค่าคงที่การทำให้เรียบเมื่อได้ถูกกำหนดค่าที่เหมาะสมแล้วจะมีค่าคงที่ การพยากรณ์จะใช้ค่าพารามิเตอร์นี้ไปคำนวณค่าพยากรณ์จนกว่าจะมีการกำหนดค่าใหม่ ซึ่งอาจทำให้ค่าพยากรณ์ไม่สามารถปรับตนเองไปตามการเปลี่ยนแปลงในสภาพแวดล้อมที่อาจเกิดขึ้นในช่วงที่ยังไม่มีการปรับเปลี่ยนค่าพารามิเตอร์ จึงได้มีนักวิชาการเสนอแนวความคิดที่จะปรับค่าพารามิเตอร์ตามลักษณะความคลาดเคลื่อนที่ได้เกิดขึ้นทำให้ค่าพยากรณ์มีความอ่อนไหวไปตามการเปลี่ยนแปลงในสภาพแวดล้อมได้

Chow เสนอให้สร้างตัวแบบการพยากรณ์แบบการทำให้เรียบแบบเอกซ์โปเนนเชียลขึ้นมา 3 ตัวแบบ โดยใช้ค่าคงที่การทำให้เรียบ 3 ค่า ดังนี้ α_0 เป็นค่าคงที่การทำให้เรียบสำหรับค่าพยากรณ์ที่จะนำไปใช้งาน ส่วนค่าคงที่การทำให้เรียบอีก 2 ค่า จะกำหนดให้มีค่าต่างจาก α_0 เล็กน้อย

$$\alpha_u = \alpha_0 + \delta \tag{2.19}$$

$$\alpha_\ell = \alpha_0 - \delta \tag{2.20}$$

โดย δ เป็นค่าบวกซึ่งมักจะกำหนดให้มีค่าเท่ากับ 0.05 ค่าคงที่การทำให้เรียบจะปรับค่าไปยังทิศทางที่ความคลาดเคลื่อนสัมบูรณ์ที่ทำให้เรียบแล้ว (smoothed absolute error) Δ_t มีค่าน้อยที่สุด

$$\Delta_t = \gamma|e_t| + (1-\gamma)\Delta_{t-1} \tag{2.21}$$

โดย

$$e_t = X_t - X_{t-1}(1) \tag{2.22}$$

และ γ เป็นค่าคงที่การทำให้เรียบสำหรับความคลาดเคลื่อนสัมบูรณ์ ซึ่งไม่เกี่ยวข้องกับค่าคงที่การทำให้เรียบทั้ง 3 ค่าข้างต้น ในหน่วยเวลาที่ t Δ_t จะมีค่า 3 ค่า ตามความคลาดเคลื่อนที่เกิดขึ้นจากค่าพยากรณ์จาก α_0 α_u และ α_ℓ สำหรับการพยากรณ์ในหน่วยเวลาที่ $t+1$ ค่าคงที่การทำให้เรียบ α_0 จะกำหนดให้เท่ากับค่าคงที่การทำให้เรียบของ Δ_t ที่มีค่าน้อยที่สุด และ α_u และ α_ℓ สำหรับการพยากรณ์ในหน่วยเวลาที่ $t+1$ จะคำนวณใหม่ตามสมการ (2.19) และ (2.20) จะเห็นได้ว่าค่าคงที่การทำให้เรียบในการพยากรณ์แบบการทำให้เรียบแบบเอกซ์โปเนนเชียล ปรับค่าตนเอง

ได้ตามความคลาดเคลื่อนที่เกิดขึ้น

ส่วน Trigg-Leach เสนอให้กำหนดค่าคงที่การทำให้เรียบในการพยากรณ์แบบการทำให้เรียบแบบเอกซ์โปเนนเชียล ณ หน่วยเวลา t ให้เท่ากับค่าสัมบูรณ์ของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว (smoothed error tracking signal) ซึ่งมีค่าเท่ากับ

$$\alpha_t = \left| Q_t / \Delta_t \right| \tag{2.23}$$

โดย

$$Q_t = \gamma e_t + (1-\gamma) Q_{t-1} \tag{2.24}$$

วิธีของ Trigg-Leach เป็นที่รู้จักกันในอีกนามว่า ARRSES (adaptive response rate simple exponential smoothing)

2.2.3 อนุกรมเวลาแบบคลาสสิค

อนุกรมเวลาแบบคลาสสิคเป็นวิธีการพยากรณ์ที่อาศัยส่วนประกอบ 4 ประการ คือ ส่วนประกอบแนวโน้ม (trend) ซึ่งเป็นการฉายภาพ (projection) ระยะยาว ส่วนประกอบวัฏฏะ (cycle) ซึ่งเป็นลักษณะของการเปลี่ยนแปลงตามวงจร ส่วนประกอบฤดูกาล (seasonal factor) ซึ่งเป็นการเปลี่ยนแปลงในช่วงเวลาของวงจร และความรบกวนสุ่ม (random disturbance) ดังนั้น ค่าของอนุกรมเวลาอาจเขียนเป็นเชิงบวกได้เท่ากับ

$$X = T+C+S+I \tag{2.25}$$

ซึ่งมีค่าพยากรณ์เท่ากับ

$$\hat{X} = \hat{T} + \hat{C} + \hat{S} \tag{2.26}$$

หรือค่าของอนุกรมเวลาอาจเขียนเป็นเชิงคูณได้เท่ากับ

$$X=TCS+I \tag{2.27}$$

ซึ่งมีค่าพยากรณ์เท่ากับ

$$\hat{X} = \hat{T}\hat{C}\hat{S} \tag{2.28}$$

เมื่อ T คือ ส่วนประกอบแนวโน้ม C เป็นส่วนประกอบวัฏฏะ S เป็นดัชนีฤดูกาล (seasonal index) และ I เป็นความรบกวนสุ่ม วิธีการหาส่วนประกอบแนวโน้ม ส่วนประกอบวัฏฏะ และดัชนีฤดูกาลอาจหาศึกษาได้จากหนังสืออนุกรมเวลาแบบคลาสสิค

ผู้ที่เคยทำการพยากรณ์ด้วยอนุกรมเวลาแบบคลาสสิคนั้น ย่อมทราบถึงความลำบากใจในการเลือกรูปแบบฟังก์ชันของแนวโน้มและการหาส่วนประกอบวัฏฏะ ผู้ที่จะใช้รูปแบบฟังก์ชันของแนวโน้มเป็นโพลิโนเมียล ควรจะทราบถึงความผิดพลาดที่อาจเกิดขึ้นได้ และการหาส่วนประกอบวัฏฏะจะต้องมีข้อมูลในอดีตเป็นจำนวนมากเพื่อให้เห็นการซ้ำของวัฏฏะ จุดอ่อนที่สำคัญอีกอย่างหนึ่งของอนุกรมเวลาแบบคลาสสิคคือ โดยเนื้อแท้แล้ววิธีการนี้ไม่ใช่วิธีการเชิงสถิติ การทดสอบนัยสำคัญ (test of significance) ไม่อาจกระทำได้โดยตรง

2.2.4 การวิเคราะห์การถดถอย ซึ่งจะกล่าวในบทที่ 3 ต่อไป

2.2.5 การพยากรณ์เชิงเศรษฐมิติ

เพื่อให้เกิดความเข้าใจในการพยากรณ์เชิงเศรษฐมิติ (econometric forecasting) ใคร่จะขอกล่าวถึงการวิเคราะห์การถดถอย (regression analysis) สักเล็กน้อย ในสมการถดถอยนั้น ตัวแปรอิสระจะต้องอยู่ภายนอกอิทธิพลของตัวแปรตามและจะเป็นอิสระต่อกันด้วย ข้อสมมุตินี้ใช้ได้และมีเหตุมีผลในหลาย ๆ สถานการณ์ แต่ในบางสถานการณ์ข้อสมมุตินี้ก็อาจใช้ไม่ได้ เช่น สมมุติว่าปริมาณขายเป็นฟังก์ชันของ GDP (Gross Domestle Produet) ราคาขาย และการโฆษณา ถ้าจะพยากรณ์การขายในรูปแบบนี้โดยวิธีการถดถอย ตัวแปรอิสระในที่นี้คือ GDP ราคาขายและการโฆษณา ซึ่งจะต้องอยู่ภายนอกอิทธิพลของการขายและเป็นอิสระระหว่างกันด้วย จะเห็นว่า GDP เป็นตัวแปรอิสระซึ่งไม่ถูกกระทบกระเทือนจากปริมาณขาย ราคาขายหรือการโฆษณา แต่ข้อสมมุติดังกล่าวข้างต้นไม่น่าจะใช้ได้กับตัวแปรอิสระ ราคาขายและการโฆษณาเพราะถ้าต้นทุนต่อหน่วยอยู่ในรูปฟังก์ชันกำลังสองระดับการขายที่แตกต่างกันย่อมได้ต้นทุนต่อหน่วยที่แตกต่างกันด้วย และค่าใช้จ่ายในการโฆษณาย่อมมีผลกระทบต่อราคาขายแน่ ในทางกลับกันราคาขายย่อมขึ้นอยู่กับปริมาณขายและปริมาณขายนี้อาจมีผลต่อระดับการโฆษณาก็ได้ จะเห็นว่าในสถานการณ์เช่นนี้แสดงให้เห็นว่าตัวแปรเหล่านี้มีความพึ่งพิงซึ่งกันและกัน (interdependency) และหากความพึ่งพิงซึ่งกันและกันนี้มีมาก การวิเคราะห์การถดถอยก็อาจใช้ไม่ได้ ดังนั้น ในรูปแบบเศรษฐมิติปัญหาดังกล่าวจึงอาจเขียนได้ดังนี้

ปริมาณขาย = f(GDP,ราคาขาย,การโฆษณา)

ต้นทุน = f (ระดับการผลิต, ค่าใช้จ่ายอื่น ๆ)

ค่าใช้จ่ายในการขาย = f(การโฆษณา,ค่าใช้จ่ายอื่น ๆ)

ราคาขาย = f(ต้นทุน,ค่าใช้จ่ายในการขาย)

จะเห็นได้ว่า แทนที่จะมีเพียงหนึ่งสมการอย่างในการวิเคราะห์การถดถอย แต่มีถึง 4 สมการ ขั้นตอนจากนี้ไปจะเหมือนกับการวิเคราะห์การถดถอยคือ กำหนดรูปแบบฟังก์ชันในสมการทั้ง 4 แล้ว ประมาณค่าของพารามิเตอร์ทั้งหมดพร้อม ๆ กัน แล้วทดสอบนัยสำคัญของผลลัพธ์และความถูกต้องของข้อสมมุติ จะเห็นว่าความได้เปรียบของการพยากรณ์เศรษฐมิตินี้ เมื่อเปรียบเทียบกับ

การวิเคราะห์การถดถอยที่สำคัญคือ ค่าของตัวแปรอิสระบางตัวถูกกำหนดขึ้นเองภายใต้ตัวแบบทำให้ผู้ใช้ไม่ต้องไปคาดคะเนค่าของตัวแปรอิสระเหล่านั้น หากพิจารณาในแง่นี้ อาจกล่าวได้ว่าความถูกต้องของการพยากรณ์เศรษฐมิติควรจะดีกว่าการวิเคราะห์การถดถอย แต่ปัญหาที่ต้องกำหนดรูปแบบฟังก์ชันต่าง ๆ ซึ่งเป็นความยุ่งยากใจของผู้พยากรณ์ด้วยวิธีวิเคราะห์การถดถอยรู้สึกจะยังคงเป็นปัญหาอยู่เช่นเดิม

2.2.6 อนุกรมเวลา Box-Jenkins

ความลำบากในการพยากรณ์ด้วยวิธีอนุกรมเวลาและแบบคลาสสิคเป็นที่ทราบกันดี โดยเฉพาะเมื่ออนุกรมเวลามิได้มีรูปแบบของแนวโน้ม วัฏฏะ หรือฤดูกาลที่เด่นชัด การวิเคราะห์อนุกรมเวลาแบบ Box-Jenkins ได้ขจัดปัญหาดังกล่าวให้หมดไป นอกจากนี้ Box-Jenkins ได้เปลี่ยนปรัชญาในการพยากรณ์ใหม่ในแง่ที่ว่า ในการพยากรณ์ด้วยวิธีอื่น ๆ ตามที่ได้กล่าวมาแล้ว ผู้ที่สร้างตัวแบบพยากรณ์จะต้องกำหนดรูปแบบของความสัมพันธ์ขึ้นก่อน จึงจะทำการวิเคราะห์ต่อไปได้ เช่น ในการวิเคราะห์การถดถอยจะต้องกำหนดรูปแบบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ ในการวิเคราะห์อนุกรมเวลาแบบคลาสสิครูปแบบของแนวโน้มจะต้องกำหนดขึ้น เป็นต้น แต่วิธี Box-Jenkins ไม่มีการกำหนดรูปแบบตายตัวขึ้นก่อนทำการวิเคราะห์ ในระหว่างการวิเคราะห์รูปแบบจะถูกกำหนดขึ้นมาเอง ขั้นตอนของวิธี Box-Jenkins พอสรุปได้ ดังนี้

1. วิเคราะห์สหสัมพันธ์ในตัวเอง (autocorrelation) และสหสัมพันธ์ในตัวเองบางส่วน (partial autocorrelation) ของข้อมูลในอนุกรมเวลาและผลต่างของข้อมูล

2. ผลการวิเคราะห์ในขั้นตอนที่ 1 จะใช้ในการเลือกตัวแบบพยากรณ์

3. ประมาณค่าพารามิเตอร์ในตัวแบบและตรวจสอบความถูกต้องของตัวแบบ

4. ถ้าผลการตรวจสอบปรากฏว่า ตัวแบบไม่มีความเหมาะสม ผู้วิเคราะห์จะต้องกลับไปขั้นตอนที่ 2 โดยเลือกตัวแบบพยากรณ์ใหม่ โดยอาศัยผลการวิเคราะห์ในขั้นตอนที่ 1 และการตรวจสอบในขึ้นตอนที่ 3 แล้วทำซ้ำขั้นตอนที่ 3 ใหม่ ถ้าผลการตรวจสอบปรากฏว่ามีความเหมาะสมในตัวแบบ ตัวแบบนี้ก็จะนำไปใช้ในการพยากรณ์

จากขั้นตอนดังกล่าวข้างต้นจะเห็นว่า สหสัมพันธ์ในตัวเองมีความสำคัญต่อวิธีการของ Box-Jenkins เป็นอย่างมาก สหสัมพันธ์ในตัวเองเป็นมาตรการวัดความสัมพันธ์ในค่าของข้อมูลที่เกิดขึ้น ณ เวลาต่าง ๆ ซึ่งจะบ่งถึงโครงสร้างของข้อมูลและลักษณะการเปลี่ยนแปลงในข้อมูล ในกรณีที่ข้อมูลเป็นข้อมูลสุ่มอย่างสมบูรณ์ (completely random) สหสัมพันธ์จะมีค่าเท่ากับศูนย์หมด แต่ถ้าข้อมูลมีความพึ่งพิงต่อกัน สหสัมพันธ์จะมีค่าสูง ดังนั้น เมื่อคำนวณสหสัมพันธ์ในตัวเองของข้อมูลเรียบร้อยแล้ว จะทำให้ทราบถึงโครงสร้างของข้อมูล ซึ่งจะเป็นประโยชน์มากในการเลือกตัวแบบพยากรณ์ และผลการวิเคราะห์ค่าคงเหลือในการตรวจสอบจะแนะนำการปรับปรุงตัวแบบด้วย

ดังนั้น ถึงแม้ผู้วิเคราะห์ขาดประสบการณ์ แต่ถ้ามีความเข้าใจในเรื่องสหสัมพันธ์ ก็อาจได้ตัวแบบพยากรณ์ที่ดีได้ อนึ่ง ช่วงความเชื่อมั่นของค่าพยากรณ์ก็อาจสร้างขึ้นได้ในวิธีการ Box-

Jenkins เมื่อได้ตัวแบบพยากรณ์แล้ว การพยากรณ์จะเป็นไปได้อย่างง่ายดาย และสามารถกระทำด้วยมือไม่ต้องอาศัยคอมพิวเตอร์ ในปัจจุบันนี้วิธีการ Box-Jenkins เป็นเทคนิคการพยากรณ์ที่ให้ความถูกต้องสูงสุดวิธีหนึ่ง

2.3 เทคนิคการทำให้เรียบ

เทคนิคการทำให้เรียบ (smoothing techniques) เป็นระเบียบวิธีการพยากรณ์ที่ได้รับความสนใจอย่างกว้างขวางในวงการวิชาการ ได้พัฒนาตัวแบบการพยากรณ์มากมายภายใต้แนวความคิดของเทคนิคการทำให้เรียบทำให้มีตัวแบบการพยากรณ์สำหรับข้อมูลที่มีการเคลื่อนไหวในลักษณะปราศจากแนวโน้ม ในลักษณะที่มีแนวโน้มและในลักษณะที่มีฤดูกาลทั้งที่มีและไม่มีแนวโน้ม จึงอาจกล่าวได้ว่า ไม่ว่าข้อมูลจะมีการเคลื่อนไหวในลักษณะใดจะสามารถหาตัวแบบการพยากรณ์ที่เหมาะสมภายใต้แนวความคิดของเทคนิคการทำให้เรียบสำหรับข้อมูลได้

เทคนิคการทำให้เรียบเป็นการพยากรณ์ภายใต้แนวความคิดที่ว่าพฤติกรรมในอดีตของตนเองมีความเพียงพอที่จะพยากรณ์พฤติกรรมในอนาคตของตนเองได้ ในบทนี้จะจำแนกตัวแบบการพยากรณ์ออกเป็น 3 ประเภทใหญ่ ๆ ตัวแบบการพยากรณ์ประเภทที่ 1 เป็นตัวแบบค่าเฉลี่ยคงที่ (constant mean model) ซึ่งมีความเหมาะสมสำหรับข้อมูลที่มีการเคลื่อนไหวแบบปราศจากแนวโน้มหรือแบบที่ระดับข้อมูลมีการเปลี่ยนแปลงค่อนข้างช้า ตัวแบบการพยากรณ์ประเภทที่ 2 เป็นตัวแบบแนวโน้มเชิงเส้น (linear trend model) ตัวแบบการพยากรณ์ประเภทนี้มีแนวโน้มเพิ่มขึ้น และบางครั้งอาจมีแนวโน้มลดลง ตัวแบบการพยากรณ์ประเภทที่ 3 เป็นตัวแบบฤดูกาล (seasonality model) ซึ่งเหมาะสมสำหรับข้อมูลที่มีการเคลื่อนไหวแบบมีฤดูกาล

2.3.1 ตัวแบบค่าเฉลี่ยคงที่ (Constant Mean Model)

ตัวแบบประเภทนี้พัฒนาขึ้นสำหรับข้อมูลที่เคลื่อนไหวรอบ ๆ ค่าคงที่ค่าหนึ่ง ค่าคงที่ดังกล่าวอาจจะคงที่ตลอดช่วงระยะเวลายาวนาน หรือคงที่เฉพาะในช่วงระยะเวลาหนึ่ง ๆ และเปลี่ยนแปลงอย่างเชื่องช้าจากช่วงระยะเวลาหนึ่งไปยังอีกช่วงระยะเวลาหนึ่ง ดังนั้น ตัวแบบค่าเฉลี่ยคงที่จึงจำแนกออกเป็น 2 ตัวแบบ คือ ตัวแบบค่าเฉลี่ยคงที่ระยะยาว (global constant mean model) และตัวแบบค่าเฉลี่ยคงที่ระยะสั้น (local constant mean model)

1. ตัวแบบค่าเฉลี่ยคงที่ระยะยาว (Global Constant Mean Model)

ตัวแบบนี้อาจเขียนเป็นสมการได้ดังนี้

$$X_t = \mu + \varepsilon_t \qquad t = 1, 2, 3, \ldots \qquad (2.29)$$

ค่าของตัวแปร X ณ เวลา t ประกอบด้วยค่าคงที่ μ และตัวแปรสุ่ม ε ณ เวลา t ε_t เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน นั่นคือ ε_t และ ε_k เมื่อ $t \neq k$ ไม่มีความพึ่งพิงต่อกัน แต่มีการแจกแจงแบบเดียวกัน โดยมีค่าคาดหมายเท่ากับศูนย์ และมีค่าความแปรปรวนเป็นค่าคงที่เท่ากับ σ^2

เนื่องจากการพยากรณ์ค่า X_{t+k} อาจกระทำ ณ จุดเวลา $t, t+1, t+2, \ldots$ จนถึง $t+k-1$ ซึ่งอาจได้ค่าพยากรณ์ที่แตกต่างกัน จึงควรสร้างสัญลักษณ์ของค่าพยากรณ์ให้ชัดเจนถึงจุดที่กระทำการพยากรณ์และจำนวนหน่วยเวลาล่วงหน้าที่พยากรณ์ ดังนั้น ค่าพยากรณ์ kหน่วยเวลาล่วงหน้าที่กระทำ ณ จุดเวลา t ของ X_{t+k} จะเขียนแทนด้วย $\hat{X}_t(k)$ เมื่อพิจารณาตัวแบบนี้ให้ถ่องแท้ จากสมการ (2.29) จะพบว่ามีพารามิเตอร์ที่ไม่ทราบค่าอยู่ 2 ตัว คือ ค่าคงที่ μ ซึ่งเป็นพารามิเตอร์โดยตรงของตัวแบบ และค่าความแปรปรวน σ^2 ของตัวแปรสุ่ม ε_t ณ จุดเวลา t ข้อมูลที่มีอยู่จะเป็น $X_1, X_2, \ldots, X_t$ ซึ่งในค่า X แต่ละค่าจะมีค่า μ ผสมอยู่กับค่าของตัวแปรสุ่ม ε ในการคำนวณค่า $\hat{X}_t(k)$ จะเห็นได้จากสมการ (2.29) ว่าจะต้องประมาณค่าคงที่ μ และค่าของตัวแปรสุ่ม ε_{t+k} เนื่องจาก ε_{t+k} เป็นอิสระกับค่าในอดีตของตัวแปรสุ่ม $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_t$ จึงมิอาจพยากรณ์ค่าของตัวแปรสุ่ม ε_{t+k} ได้ ดังนั้น จึงกำหนดให้ค่า ε_{t+k} เท่ากับค่าคาดหมายของตัวแปรสุ่ม ซึ่งมีค่าเท่ากับศูนย์ จากสมการ (2.1) ค่าพยากรณ์ของ X_{t+k} ที่กระทำ ณ จุดเวลา จึงอาจเขียนได้เป็น

$$\hat{X}_t(k) = \mu \qquad (2.30)$$

การพยากรณ์ค่า X_{t+k} ไม่ว่า k จะมีค่าใด ๆ จะได้ค่า μ เสมอ เนื่องจาก μ เป็นค่าคงที่ไม่เปลี่ยนแปลงไปตามเวลา คำถามต่อไปก็คือ จะประมาณ μ ได้อย่างไร จากข้อมูลที่มีอยู่ คือ $X_1, X_2, \ldots, X_t$ ค่าประมาณที่เหมาะสมที่สุดของ μ ก็คือ ค่าเฉลี่ยของข้อมูลที่มีอยู่ทั้งหมดในขณะนั้น ดังนั้น

$$\hat{X}\, t(1) = \frac{\sum_{i=1}^{t} X_i}{t} \qquad k = 1,2,3,\ldots \qquad (2.31)$$

รูปแบบดังกล่าวข้างต้นเรียกว่า explicit form ซึ่งภาระการคำนวณจะเพิ่มขึ้นตามปริมาณข้อมูลที่เพิ่มขึ้น เพื่อมิให้ภาระการคำนวณจะต้องเพิ่มขึ้นตามปริมาณข้อมูลสมการ (2.31) ควรปรับปรุงใหม่ให้อยู่ในรูปแบบ recurrence ซึ่งจะทำให้สามารถหาค่าพยากรณ์ใหม่จากค่าพยากรณ์เดิมกับข้อมูลใหม่ที่เพิ่งเกิดขึ้น ทำให้ภาระการคำนวณไม่ขึ้นอยู่กับปริมาณข้อมูล จากสมการ (2.31) $\hat{X}_t(k)$ อาจเขียนใหม่ได้เป็น

$$\hat{X}_t(1) = \frac{\sum_{i=1}^{t} X_i}{t} \tag{2.32}$$

$$= \frac{(t-1)\hat{X}_{t-1}(k) + X_t}{t} \quad ; \; k = 1,2,3,\ldots$$

ซึ่งจะเห็นว่าค่าพยากรณ์ k หน่วยเวลาล่วงหน้า $\hat{X}_t(k)$ ที่กระทำ ณ เวลา t เป็นค่าเฉลี่ยแบบถ่วงน้ำหนักระหว่างค่าพยากรณ์ k หน่วยเวลาล่วงหน้าที่กระทำ ณ เวลา t-1 กับค่า X_t จึงทำให้ภาระการคำนวณไม่ขึ้นอยู่กับปริมาณข้อมูลเลย

สูตรที่จะใช้คำนวณค่าพยากรณ์ k หน่วยเวลาล่วงหน้า ซึ่งกระทำการพยากรณ์ ณ จุดเวลา t อาจเขียนอยู่ในรูปแบบ error correction ซึ่งจะใช้ความคลาดเคลื่อนที่เกิดขึ้น ณ จุดเวลา t มาปรับแก้ค่าพยากรณ์ $\hat{X}_t(k)$ ที่ได้กระทำการพยากรณ์ หน่วยเวลาล่วงหน้า ณ เวลา t-1 ค่าความคลาดเคลื่อนที่เกิดขึ้นในค่าพยากรณ์ ณ จุดเวลา t อาจเขียนได้เป็น

$$e_t = X_t - \hat{X}_{t-1}(1)$$

หรือ $X_t = \hat{X}_{t-1}(1) + e_t$

เมื่อนำค่า X_t ไปแทนค่าในสมการ (2.32) จะได้ผลดังนี้

$$\hat{X}_t(1) = \hat{X}_{t-1}(1) + e_t \tag{2.33}$$

จะเห็นได้ว่า ถ้า t มีค่ามาก ซึ่งหมายถึงมีข้อมูลที่จะนำมาใช้ในการพยากรณ์เป็นจำนวนมาก จะทำให้ความคลาดเคลื่อนในการพยากรณ์ (forecast error) มีผลที่จะมาปรับค่าพยากรณ์น้อยมาก ทำให้ค่าพยากรณ์มีความเฉื่อยสูงขึ้น ไม่เปลี่ยนแปลงมากนักแม้จะมีข้อมูลใหม่ ๆ เข้ามา

2. ตัวแบบค่าเฉลี่ยคงที่ระยะสั้น (Local Constant Mean)

ตัวแบบเฉลี่ยคงที่ระยะสั้นมีแนวความคิดคล้ายคลึงกับตัวแบบค่าเฉลี่ยที่ระยะยาว ในแง่ที่ว่า ข้อมูลอนุกรมเวลาเคลื่อนไหวรอบค่าคงที่ค่าหนึ่ง แต่ตัวแบบค่าเฉลี่ยคงที่ ระยะสั้นเห็นว่าในบางสภาพการณ์ค่าเฉลี่ยคงที่มิได้มีค่าคงที่ตลอดระยะเวลายาวนาน แต่มีค่าคงที่ในช่วงระยะสั้น ๆ (locality) และมีค่าเปลี่ยนแปลงที่ไม่รวดเร็วนัก ดังนั้น ตัวแบบค่าเฉลี่ยคงที่ระยะสั้น จึงอธิบายได้สมการ 2.34

$$X_t = \mu_t + \varepsilon_t \tag{2.34}$$

โดย μ_t เป็นค่าคงที่ในช่วงระยะสั้น และอาจมีค่าเปลี่ยนแปลงได้อย่างเชื่องช้า ส่วน ε_t เป็นความรบกวนสุ่มซึ่งเป็นอิสระต่อกัน แต่มีการแจกแจงความน่าจะเป็นแบบเดียวกัน โดยมีค่าคาดหมายเท่ากับศูนย์ และความแปรปรวนเท่ากับ σ^2 การพยากรณ์ค่า X_t ในสมการ (2.34) จึงเป็นการพยากรณ์ค่าคงที่ ซึ่งอาจเปลี่ยนแปลงค่าได้คือ μ_t ซึ่งในที่นี้จะกล่าวถึงวิธีการพยากรณ์โดยค่าเฉลี่ยเคลื่อนที่ชั้นเดียว (single moving average) และการทำให้เรียบแบบเอกซ์โปแนนเชียลชั้นเดียว (single exponential smoothing)

- ค่าเฉลี่ยเคลื่อนที่ชั้นเดียว (Single Moving Average)

จากแนวความคิดที่ว่าในช่วงระยะเวลาสั้น ๆ อนุกรมเวลาเคลื่อนไหวรอบค่าคงที่ค่าหนึ่ง จึงนำไปสู่การพยากรณ์ค่าคงที่นั้น ๆ โดยเฉลี่ยของข้อมูลในช่วงระยะเวลาสั้น ดังกล่าว ดังนั้น ถ้าในช่วงระยะเวลาดังกล่าว มีจำนวนข้อมูลอยู่ n พจน์ค่าพยากรณ์หนึ่งหน่วยเวลา ล่วงหน้าที่ทำการพยากรณ์ ณ เวลา t จึงเท่ากับค่าเฉลี่ยของข้อมูล n พจน์

$$\hat{X}\,t(1) = \frac{X_{t-n+1} + X_{t-n+2} + \ldots + X_t}{n} \qquad (2.35)$$

เมื่อเวลาได้เคลื่อนที่ไปถึงหน่วยเวลา $t+1$ ค่าจริงของ x_{t+1} ก็ปรากฏค่าพยากรณ์ ที่กระทำการ ณ เวลา $t+1$ ก็ยังเป็นค่าเฉลี่ยเลขคณิตของข้อมูล n พจน์ แต่เป็น n พจน์ล่าสุดโดยเอา x_{t+1} ซึ่งเป็นข้อมูลที่เพิ่งเข้ามาไปแทน x_{t-n+1} ซึ่งได้เข้ามาแล้วตั้งแต่ $t+1$ หน่วยเวลา

$$\hat{X}_{t+1}(1) = \frac{X_{t-n+2} + X_{t-n+3} + \ldots + X_t + X_{t+1}}{n}$$

ซึ่งเขียนใหม่ในรูป recurrence ได้เป็น

$$\hat{X}_{t+1}(1) = \hat{X}_t(1) + \frac{X_{t+1} - X_{t-n+1}}{n} \qquad (2.36)$$

สมการ (2.36) แสดงความสัมพันธ์ระหว่างค่าพยากรณ์ใหม่กับค่าพยากรณ์เก่าโดยปรับค่าพยากรณ์เก่าด้วยผลต่างของข้อมูลล่าสุดกับข้อมูลเก่าสุดหารด้วยจำนวนพจน์ที่ใช้ในการหาค่าเฉลี่ย ซึ่งจะเห็นได้ว่า ถ้าจำนวนพจน์ที่ใช้ในการหาค่าเฉลี่ยมีจำนวนมาก ส่วนที่จะปรับค่าพยากรณ์เก่าจะมีค่าน้อยทำให้ค่าพยากรณ์ใหม่มีค่าใกล้เคียงกับค่าพยากรณ์เก่า จึงมีการกล่าวว่าค่าพยากรณ์จะมีค่าเฉื่อยสูง (high inertia) ถ้าพจน์ที่ใช้ในการหาค่าเฉลี่ยมีจำนวนมาก สมการ (2.36) อาจเขียนใหม่ในรูปแบบที่ใช้หาค่าความคลาดเคลื่อนในการพยากรณ์ มาปรับแก้ (error correction) ได้ดังนี้

$$\hat{X}_{t+1}(1) = \frac{n+1}{n}\hat{X}_t(1) + \frac{e_{t+1} - X_{t-n+1}}{n} \qquad (2.37)$$

- การทำให้เรียบแบบเอกซ์โปเนนเชียลชั้นเดียว (Single Exponential Smoothing)

การหาค่าเฉลี่ยด้วยวิธีการเฉลี่ยเคลื่อนที่นั้นได้ให้น้ำหนักของข้อมูลที่หน่วยเวลาต่างๆ เท่า ๆ กัน ซึ่งอาจมีข้อโต้แย้งได้มาก ในกรณีที่ข้อมูลไม่มีฤดูกาล ข้อมูลที่หน่วยเวลาใกล้เคียงกับปัจจุบันน่าจะมีสาระที่เป็นประโยชน์ในการประมาณค่าพยากรณ์มากกว่าข้อมูลในอดีตที่ห่างไกล ดังนั้น วิธีการทำให้เรียบแบบเอกซ์โปเนนเชียล จึงให้น้ำหนักต่าง ๆ กันสำหรับข้อมูลที่หน่วยเวลาต่าง ๆ โดยจะให้น้ำหนักของข้อมูลลดลงแบบเอกซ์โปเนนเชียลตามหน่วยเวลาที่ห่างออกไปในอดีต ลักษณะการให้น้ำหนักกับข้อมูลจะเป็นดังนี้

ข้อมูล	X_1	X_2	X_3	…	X_{t-2}	X_{t-1}	X_t
น้ำหนัก	a^{t-1}	a^{t-2}	a^{t-3}	…	a^2	a^1	a^0

เมื่อ a คือค่าคงที่เรียกว่า ปัจจัยส่วนลด (discounting factor) และ $0 < a < 1$ ค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้า เมื่อทำการพยากรณ์ ณ เวลา t ของวิธีการทำให้ เรียบแบบเอกซ์โปเนนเชียลชั้นเดียว จึงเป็นค่าเฉลี่ยแบบถ่วงน้ำหนัก ดังนี้

$$\hat{X}_t(1) = \frac{\sum_{r=0}^{t-1} a^r X_{t-r}}{\sum_{r=0}^{t-1} a^r} \qquad (2.38)$$

ภาระการคำนวณการพยากรณ์ในสมการ (2.38) จะเพิ่มอย่างรวดเร็ว เมื่อหน่วยเวลาเพิ่มขึ้น ดังนั้น เพื่อมิให้ภาระการคำนวณมากขึ้นตามหน่วยเวลาที่เพิ่มขึ้น

ให้

$$S_t = \sum_{r=0}^{t-1} a^r X_{t-r} \qquad (2.39)$$

$$W_t = \sum_{r=0}^{t-1} a^r \qquad (2.40)$$

จากคำนิยามในสมการ (2.39) ความสัมพันธ์ระหว่าง s_t กับ s_{t-1} อาจเขียนได้เป็น

$$S_t = a\ S_{t-1} + X_t \qquad (2.41)$$

ในทำนองเดียวกัน

$$W_t = a\ W_{t-1} + 1 \qquad (2.42)$$

โดยกำหนดให้ s_0 และ w_0 มีค่าเท่ากับศูนย์ ดังนั้น สูตรในรูปแบบ recurrence สำหรับค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้า เมื่อทำการพยากรณ์ ณ เวลา t จึงเขียนได้เป็น

$$\hat{X}\,t(1) = S_t \;/ W_t \tag{2.43}$$

พิจารณาสมการ (2.40) ให้ละเอียด อาจกล่าวได้ว่า w_t เป็นผลบวกอนุกรม เรขาคณิต ค่า w_t เมื่อ t มีค่ามาก จะมีค่าเท่ากับ

$$\lim_{t \to \infty} W_t \;= \frac{1}{1-a} \tag{2.44}$$

ดังนั้น กรณีที่ t มีค่าสูง ค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าในสมการ (2.44) จะประมาณได้ดังนี้

ให้
$$\begin{aligned}\hat{X}\,t(1) &= (1-a) \sum_{r=0}^{t-1} a^r X_{t-r} \\ &= (1-a)(X_t + aX_{t-1} + a^2 X_{t-2} + \ldots) \\ &= (1-a)X_t + a(1-a)(X_{t-1} + a^2 X_{t-2} + \ldots) \\ &= (1-a)X_t + \hat{X}\,t-1(1)\end{aligned} \tag{2.45}$$

ส่วนค่าคาดหมายของค่าพยากรณ์นั้น เมื่อได้พิจารณาสมการ (2.43) แล้ว จะมีค่าเท่ากับอัตราส่วนของค่าคาดหมายของ s_t กับ w_t ซึ่งค่าคาดหมายของ s_t อาจเขียนได้จากสมการ (2.39) ได้เท่ากับ

$$E(S_t\,) = \sum_{r=0}^{t-1} a^r \mu_{t-r}$$

ดังนั้น ค่าคาดหมายของค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าจึงมีค่าเท่ากับ

$$\begin{aligned}E\left[\;\hat{X}\,t(1)\;\right] &= E(S_t\,)/W_t \\ &= \frac{\sum_{r=0}^{t-1} a^r \mu_{t-r}}{\sum_{r=0}^{t-1} a^r}\end{aligned} \tag{2.46}$$

ซึ่งเป็นค่าเฉลี่ยแบบถ่วงน้ำหนักของค่าคงที่ μ_t ที่อาจเปลี่ยนแปลงได้อย่างเชื่องช้า ส่วนค่าคาดหมายของ x_t ซึ่งเขียนได้จากสมการ (2.34) เป็น

$$E(X_t) = \mu_t \tag{2.47}$$

จะเห็นได้ว่า ภายใต้สมมติฐานที่ว่าคงที่ μ_t อาจเปลี่ยนแปลงค่าได้อย่างเชื่องช้าและในช่วงระยะเวลาสั้น ๆ ค่าเฉลี่ยของค่าที่ μ_t มีค่าคงที่ μ ค่าเฉลี่ยแบบถ่วงน้ำหนักของค่าคงที่ μ_t ในอดีตจึงเป็นค่าประมาณของ μ ที่เหมาะสมและมีค่าใกล้เคียงกับความเป็นจริง อย่างไรก็ตาม หากค่าคงที่ μ_t มีค่าคงที่จริงเท่ากับ μ ค่าคาดหมายของค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้า ในสมการ(2.46) จะกลายเป็น

$$E[\ \hat{X}\, t(1)\ ] = \mu$$

ซึ่งเท่ากับค่าคาดหมายของข้อมูลในสมการ (2.48)

ความแปรปรวนของ $\hat{X}\, t(1)$ อาจเขียนได้จากสมการ (2.42) เท่ากับ

$$\begin{aligned} Var[\ \hat{X}\, t(1)\ ] &= Var\left(\frac{S_t}{W_t}\right) \\ &= \frac{Var(S_t)}{W_t^2} \end{aligned} \tag{2.48}$$

จากสมการ(2.34) และ (2.38)

$$\begin{aligned} Var(S_t) &= E[S_t - E(S_t)]^2 \\ &= E\left[\sum_{r=0}^{t-1} a^r \varepsilon_{t-r}\right]^2 \\ &= \sum_{r=0}^{t-1} a^{2r}\sigma^2 \\ &= \frac{1-a^{2t}}{1-a^2}\sigma^2 \end{aligned} \tag{2.49}$$

โดยคำนิยามในสมการ (2.39) w_t เป็นผลรวมของอนุกรมเราขาคณิต t พจน์ จึงอาจเขียนได้เท่ากับ

$$W_t = \frac{1-a^t}{1-a}$$

ดังนั้น ความแปรปรวนของค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าในสมการ (2.48) จึงกลายเป็น

$$Var[\hat{X}\, t(1)] = \frac{(1-a)^2}{(1-a^t)^2} \cdot \frac{1-a^{2t}}{1-a^2}\sigma^2$$

$$= \frac{(1-a)}{(1+a)} \cdot \frac{1+a^t}{1-a^t}\sigma^2 \qquad (2.50)$$

และเมื่อ t มีค่ามากขึ้น ความแปรปรวนในสมการ (2.50) จะมีค่าเข้าใกล้

$$\lim_{t \to \infty} Var\left(\hat{X}_t(1)\right) = \frac{1-a}{1+a}\sigma^2 \qquad (2.51)$$

ซึ่งจะเห็นว่าความแปรปรวนของค่าพยากรณ์จากวิธีการทำให้เรียบแบบเอกซ์โปเนนเชียลชั้นเดียว จะเข้าไปสู่ค่าคงที่เมื่อดำเนินการพยากรณ์ไปได้ระยะหนึ่ง การที่ความแปรปรวนของค่าพยากรณ์มีค่าคงที่สามารถอธิบายได้จากแนวความคิดของวิธีการพยากรณ์วิธีนี้นั่นคือ เมื่อมีข้อมูลใหม่เข้ามา ข้อมูลเก่าจะมีความสำคัญน้อยลง ดังนั้นถึงแม้ว่าจำนวนข้อมูลในอดีตที่มีอยู่จะมีจำนวนมาก (t มีค่ามาก) ก็ไม่ทำให้สาระของข้อมูลเหล่านั้นที่มาใช้ในการพยากรณ์มีมากขึ้น เนื่องจากน้ำหนักแบบเอกซ์โปเนนเชียล ที่ใช้ในการคำนวณค่าเฉลี่ยจะมีค่าน้อยมากสำหรับข้อมูลในอดีตที่ห่างไกลไปจากปัจจุบัน สาระของข้อมูลในอดีตที่อยู่ใกล้กับปัจจุบัน จะมีอิทธิพลมากต่อค่าพยากรณ์ จึงอาจกล่าวได้ว่าการที่ใช้ข้อมูลจำนวนมากในการพยากรณ์ จะไม่ทำให้เส้นพยากรณ์ของตัวแบบการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซ์โปแนนเชียลมีความเฉื่อยเพิ่มขึ้น จากสมการ (2.51) จะเห็นว่าเมื่อกำหนดค่าปัจจุบันส่วนลด (discounting factor) a มีค่าใกล้ 1 ความแปรปรวนของค่าพยากรณ์จะมีค่าใกล้ศูนย์ ทั้งนี้เนื่องจากน้ำหนักแบบเอกซ์โปแนนเชียลจะมีค่าลดลงในอัตราที่ช้ามากขึ้นเมื่อ a มีค่าใกล้ 1 มากขึ้น ทำให้สาระของข้อมูลในอดีตที่นำมาใช้ในการพยากรณ์ยังคงมีอยู่มาก ดังนั้น การเพิ่ม a มากขึ้นในวิธีการทำให้เรียบแบบเอกซ์โปแนนเชียล จึงเสมือนหนึ่งเป็นการเพิ่มจำนวนพจน์ n ในการคำนวณค่าเฉลี่ยเคลื่อนที่ และในทางตรงกันข้ามการที่ลดค่า a ในวิธีการทำให้เรียบแบบเอกซ์โปแนนเชียล ก็เสมือนหนึ่งเป็นการลดจำนวนพจน์ n ในการคำนวณค่าเฉลี่ยเคลื่อนที่

อนึ่ง การที่ความแปรปรวนของค่าพยากรณ์ แปรตรงข้ามกับค่าปัจจัยส่วนลด a ย่อมหมายความถึงความเรียบ (smoothness) ของเส้นพยากรณ์แปรตามค่าปัจจัยส่วนลด ดังนั้น จึงได้มีการนิยามค่าคงที่การทำให้เรียบ (smoothing constant) α ให้มีค่าเท่ากับ 1-a เพื่อที่จะได้บ่งชี้ความเรียบของเส้นพยากรณ์ เมื่อนำค่าคงที่การทำให้เรียบตามคำนิยามดังกล่าว ไปแทนค่าปัจจัยส่วนลดในสมการ (2.44) ค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าจะเขียนได้เป็น

$$\hat{X}_t(1) = \alpha X_t + (1-a)\hat{X}_{t-1}(1) \tag{2.52}$$

ซึ่งเป็นที่รู้จักกันแพร่หลายในเอกสารวิชาการมากกว่าสมการ (2.44)

เมื่อย้อนกลับไปพิจารณาตัวแบบค่าเฉลี่ยคงที่ระยะยาว (global constant mean) จะเห็นว่าความแปรปรวนของค่าพยากรณ์ เมื่อมีข้อมูลในอดีตมาก จะมีค่าเข้าสู่ศูนย์ทำเส้นพยากรณ์มีความเฉื่อยสูงขึ้น มีการเคลื่อนไหวเปลี่ยนแปลงช้าลง แต่ตัวแบบค่าเฉลี่ยคงที่ระยะสั้น (local constant mean) ทั้ง 2 ตัวแบบที่ได้กล่าวมาแล้ว ต่างก็มีค่าความแปรปรวนของค่าพยากรณ์คงที่มิได้ลดลงเข้าสู่ศูนย์เมื่อมีข้อมูลในอดีตมากขึ้น เส้นพยากรณ์ของตัวแบบค่าเฉลี่ยคงที่ระยะสั้นจึงมีการเคลื่อนไหวเปลี่ยนแปลงค่าได้ตามปกติ มิได้ขึ้นอยู่กับจำนวนข้อมูลในอดีตแต่ประการใด

ค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าจากตัวแบบการทำให้เรียบแบบเอกซ์โปแนนเชียลในทางปฏิบัติ จะคำนวณโดยใช้สมการ (2.52) แต่สมการนี้จะมีปัญหาที่จุดเริ่มต้น ณ หน่วยเวลาที่หนึ่ง เพราะไม่มีค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้าที่จะกระทำ ณ หน่วยเวลาที่ศูนย์ ดังนั้น เพื่อให้สามารถเริ่มต้นการคำนวณได้ จึงนิยมกำหนดให้ค่าพยากรณ์ $\hat{X}_1(1)$ มีค่าเท่ากับ X_1

2.3.2 ตัวแบบแนวโน้มเชิงเส้น (Linear Trend Model)

ตัวแบบแนวโน้มเชิงเส้นถูกพัฒนาขึ้นมาเพื่อใช้กับข้อมูลที่มีแนวโน้มเพิ่มขึ้นและ/หรือลดลง แนวโน้มจะมีลักษณะค่อนข้างคงที่ในช่วงระยะเวลายาว หรือลักษณะของแนวโน้มมีการเปลี่ยนแปลงน้อย เฉพาะในช่วงระยะเวลาสั้น ๆ เท่านั้นก็ได้ ดังนั้น ตัวแบบแนวโน้มเชิงเส้นที่จะกล่าวถึงต่อไปนี้ จึงอาจใช้ได้กับข้อมูลที่มีแนวโน้มที่ไม่ใช่เชิงเส้นก็ได้ เพราะแนวโน้มที่ไม่ใช่เชิงเส้นก็อาจประกอบขึ้นจากแนวโน้มเชิงเส้นหลาย ๆ เส้น ในที่นี้จะกล่าวถึงตัวแบบแนวโน้มเชิงเส้นใน 2 แบบเช่นเดียวกับตัวแบบค่าเฉลี่ยคงที่ คือตัวแบบแนวโน้มเชิงเส้นระยะยาว (global linear trend model) และตัวแบบแนวโน้มเชิงเส้นระยะสั้น (local linear trend model)

1. ตัวแบบแนวโน้มเชิงเส้นระยะยาว (Global Linear Trend Model)

ตัวแบบนี้อาจอธิบายได้ด้วยสมการ

$$X_t = b_0 + b_1 t + \varepsilon_t \tag{2.53}$$

โดย $b_0 =$ ค่าคาดหมายของข้อมูลที่หน่วยเวลาเริ่มต้นที่ $t = o$

$b_1 =$ ความลาดชันของแนวโน้ม

$\varepsilon_t =$ ความรบกวนสุ่มซึ่งมีค่าคาดหมายเป็นศูนย์ และความแปรปรวนเท่ากับค่าคงที่ σ^2 และเป็นอิสระต่อกัน

จากตัวแบบจะเห็นว่า b_1 เป็นความลาดชันที่มีค่าคงที่ นั่นคือ อัตราการเพิ่มขึ้นหรืออัตราการลดลงของข้อมูล จะไม่เปลี่ยนแปลงในช่วงเวลาที่ศึกษาและไม่ทราบว่าจะต้องประมาณค่า b_0 เป็นค่าคงที่อีกค่าหนึ่ง ซึ่งไม่ทราบค่า และจะต้องประมาณค่าให้ $\hat{b}_0$ และ $\hat{b}_1$ เป็นค่าประมาณของ b_0 และ b_1 ตามลำดับ เส้นพยากรณ์จะเขียนขึ้น จาก

$$\hat{X}\,t = \hat{b}_0 + \hat{b}_1\,t \qquad (2.54)$$

ความแตกต่างระหว่างข้อมูลจริงกับค่าในเส้นพยากรณ์เรียกว่า ส่วนที่เหลือ (residual) ซึ่งมีค่าเท่ากับ

$$\begin{aligned} \hat{\varepsilon}_t &= X_t - \hat{X}_t \\ &= X_t - \hat{b}_0 - \hat{b}_1\,t \end{aligned} \qquad (2.55)$$

$\hat{\varepsilon}_t$ อาจพิจารณาเป็นค่าประมาณของความรบกวนสุ่ม ε_t ก็ได้ ซึ่งมักจะเป็นเกณฑ์ในการประมาณค่า b_0 และ b_1 ในสมการ (2.53) โดยหลักการแล้ว การประมาณค่า b_0 และ b_1 อาจกระทำได้หลายวิธี แต่วิธีที่นิยมใช้กันคือ วิธีกำลังสองมีค่าน้อยที่สุด (least square method) นั่นคือ หาค่า $\hat{b}_0$ และ $\hat{b}_t$ ที่ทำให้ผลบวกของ $\hat{\varepsilon}_t^{\,2}$ ทั้ง พจน์มีค่าต่ำสุด

ให้ S

$$\begin{aligned} S &= \sum_{i=1}^{t} \hat{\varepsilon}_i^{\,2} \\ &= \sum_{i=1}^{t} \left(X_i - \hat{b}_0 - \hat{b}_1\,i \right)^2 \end{aligned} \qquad (2.56)$$

เงื่อนไขที่ S จะมีค่าต่ำสุดก็คือ

$$\frac{\partial S}{\partial \hat{b}_0} = -2\sum_{i=1}^{t} \left(X_i - \hat{b}_0 - \hat{b}_1\,i \right) = 0$$

$$\sum_{i=1}^{t} X_i = \hat{b}_0\,t + \hat{b}_1 \sum_{i=1}^{t} i \qquad (2.57)$$

$$\frac{\partial S}{\partial \hat{b}_0} = -2\sum_{i=1}^{t} \left(X_i - \hat{b}_0 - \hat{b}_1\,i \right) i = 0$$

$$\sum_{i=1}^{t} iX_i = \hat{b}_0 \sum_{i=1}^{t} i + \hat{b}_1 \sum_{i=1}^{t} i^2 \qquad (2.58)$$

สมการที่ (2.57) และ (2.58) เป็นที่รู้จักกันในนามสมการปกติ (normal equations) เพื่อให้การแก้สมการดังกล่าวเพื่อหาค่า $\hat{b}_0$ และ $\hat{b}_1$ ง่ายขึ้น จะทำการปรับค่าของหน่วยเวลาใหม่เพื่อให้ผลบวกของค่าของหน่วยเวลาใหม่เป็นศูนย์ ดังนั้น สมการ (2.57) และ (2.58) จะกลายเป็น

$$\sum_j X_j = \hat{b}_0 t \tag{2.59}$$

$$\sum_j jX_j = \hat{b}_0 \sum_j j^2 \tag{2.60}$$

ซึ่งค่าประมาณของ b_0 และ b_1 เขียนได้เท่ากับ

$$\begin{aligned} \hat{b}_0 &= \sum_j X_j / t \\ &= \overline{X} \end{aligned} \tag{2.61}$$

$$\hat{b}_1 = \sum_j jX_j / \sum_j j^2 \tag{2.62}$$

ดังนั้น เส้นพยากรณ์จึงเขียนได้เป็น

$$\hat{X}_{T-1}(1) = \hat{b}_0 + \hat{b}_1 T \tag{2.63}$$

โดย T เป็นหน่วยเวลาที่นับจากจุดเริ่มต้นใหม่

เพื่อศึกษาประสิทธิภาพของการประมาณค่า b_0 และ b_1 จึงต้องศึกษาค่าคาดหมายและความแปรปรวนของ $\hat{b}_0$ และ $\hat{b}_t$ จากสมการ (2.61) ค่าคาดหมายของ $\hat{b}_0$ อาจเขียนได้เป็น

$$\begin{aligned} E\left(\hat{b}_0\right) &= E\left(\overline{X}\right) \\ &= E\left[\frac{\sum_j \left(b_0 + b_1 j + \varepsilon_j\right)}{t}\right] \\ &= b_0 \end{aligned} \tag{2.64}$$

จากสมการ (2.62) ค่าคาดหมายของ $\hat{b}_1$ อาจเขียนได้เท่ากับ

$$E\left(\hat{b}_1\right) = E\left[\sum_j j\left(b_0 + b_1 j + \varepsilon\right) / \sum_j j^2\right] = b_1 \tag{2.65}$$

ดังนั้น จึงกล่าวได้ว่า ตัวประมาณ $\hat{b}_0$ และ $\hat{b}_t$ เป็นค่าประมาณที่ไม่เอียงเฉ ความแปรปรวนของ $\hat{b}_0$ อาจเขียนได้จากสมการ (2.61) และ (2.64) ได้เท่ากับ

$$
\begin{aligned}
Var\left(\hat{b}_0\right) &= E\left(\hat{b}_0 - b_0\right)^2 \\
&= E\left[\frac{\sum_j \left(b_0 + b_1 j + \varepsilon_j\right)}{t} - b_0\right]^2 \\
&= \frac{1}{t^2} E\left[\sum_j \varepsilon_j\right]^2 \\
&= \frac{1}{t^2} \sum_j E\left[\varepsilon_j^2\right] \\
&= \sigma^2 / t \qquad (2.66)
\end{aligned}
$$

ความแปรปรวนของ $\hat{b}_t$ อาจเขียนได้จากสมการ (2.62) และ (2.65) ได้เท่ากับ

$$
\begin{aligned}
Var\left(\hat{b}_1\right) &= E\left(\hat{b}_1 - b_1\right)^2 \\
&= E\left[\frac{\sum_j j\left(b_0 + b_1 j + \varepsilon_j\right)}{\sum_j j^2} - b_1\right]^2 \\
&= \frac{1}{\left(\sum_j j^2\right)^2} E\left[\sum_j j\varepsilon_j\right]^2 \\
&= \frac{\sigma^2}{\sum_j j^2} \qquad (2.67)
\end{aligned}
$$

ซึ่งจะเห็นได้ว่าหากข้อมูลมีจำนวนมาก ความแปรปรวนของ $\hat{b}_0$ และ $\hat{b}_t$ จะลดน้อยลงยังผลให้ ค่าคะเน $\hat{b}_0$ และ $\hat{b}_t$ มีความแม่นยำสูงขึ้น ซึ่งจะสะท้อนต่อไปยังเส้นพยากรณ์มีความแม่นยำสูงขึ้นด้วย

ค่าคาดหมายของค่าพยากรณ์ k หน่วยเวลาล่วงหน้าที่กระทำ ณ เวลา t ซึ่งจุดเวลา $t+k$ คือจุด T ในจุดเริ่มต้นใหม่มีค่าเท่ากับ

$$
\begin{aligned}
E\left(\hat{X}_t(k)\right) &= E\left(\hat{b}_0 + \hat{b}_1 T\right) \\
&= b_0 + b_1 T \qquad (2.68)
\end{aligned}
$$

ซึ่งมีค่าเท่ากับค่าคาดหมายของข้อมูล $E(X_T)$ ดังนั้น ค่าคาดหมายของความคลาดเคลื่อนของค่าพยากรณ์ k หน่วยเวลาล่วงหน้าที่กระทำ ณ เวลา t จึงมีค่าเท่ากับศูนย์

$$E[e_t(k)] = 0 \tag{2.69}$$

ค่าเฉลี่ยของความคลาดเคลื่อนยกกำลังสอง อาจเขียนได้กับค่าคาดหมายของความคลาดเคลื่อนยกกำลังสอง

$$\begin{aligned} \text{MSE} &= E\left[e_t^2(k)\right] \\ &= Var\left[e_t(k)\right] \end{aligned} \tag{2.70}$$

ทั้งนี้เพราะค่าคาดหมายของ $e_t(k)$ มีค่าเท่ากับศูนย์ สมการ (2.70) อาจเขียนใหม่ในรูปของผลต่างระหว่างค่าจริงของข้อมูลกับค่าพยากรณ์ได้เป็น

$$\begin{aligned} \text{MSE} &= Var\left[\left(b_0 - \hat{b}_0\right) + \left(b_1 - \hat{b}_1\right)T + \varepsilon_T\right] \\ &= Var\left(\hat{b}_0\right) + T^2 Var\left(\hat{b}_1\right) + \sigma^2 \\ &= \frac{\sigma^2}{t} + \frac{T^2\sigma^2}{\sum_j j^2} + \sigma^2 \\ &= \sigma^2\left(\frac{1}{t} + \frac{T^2}{\sum_j j^2} + 1\right) \end{aligned} \tag{2.71}$$

เมื่อกระทำการพยากรณ์หลายหน่วยเวลาล่วงหน้า ซึ่งหมายความว่า T มีค่ามากขึ้น จะเห็นได้ว่า MSE ในสมการ (2.71) จะมีค่ามากขึ้น ดังนั้น ช่วงความเชื่อมั่นของค่าพยากรณ์เมื่อพยากรณ์ล่วงหน้าหลายหน่วยเวลา จะกว้างกว่าช่วงความเชื่อมั่นของค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้า แต่หากข้อมูลมีจำนวนมากขึ้น ซึ่งหมายถึง t มีค่ามากขึ้น MSE ของค่าพยากรณ์ที่พยากรณ์ล่วงหน้า k หน่วยเวลา จะมีค่าลดน้อยลง ซึ่งยังผลให้ช่วงความเชื่อมันแคบลง

การวิเคราะห์ดังกล่าวข้างต้นเป็นการวิเคราะห์ในรูปแบบ explicit จึงควรปรับการวิเคราะห์ใหม่ให้อยู่ในรูปแบบ recurrence เพื่อความสะดวกในการคำนวณ $\hat{b}_0$ และ $\hat{b}_1$ ในสมการปกติ (2.57) และ (2.58) ซึ่งเป็นค่าประมาณการที่คำนวณ ณ หน่วยเวลา t จึงควรระบุให้ชัดเจน โดยเปลี่ยนแปลงสัญลักษณ์ $\hat{b}_0$ และ $\hat{b}_1$ ในสมการปกติ (2.57) และ (2.58) เป็น $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ ตามลำดับ
โดยกำหนดให้

$$\hat{b}_{0t} = \hat{b}_{0(t-1)} + \Delta\hat{b}_{0t} \tag{2.72}$$

$$\hat{b}_{1t} = \hat{b}_{1(t-1)} + \Delta\hat{b}_{1t} \tag{2.73}$$

ดังนั้น สมการปกติ (2.57) และ (2.58) ในรูปแบบ recurrence อาจเขียนได้เป็น

$$\sum_{i=1}^{t-1} X_i + X_t = \left(\hat{b}_{0(t-1)} + \Delta\hat{b}_{0t}\right)t + \left(\hat{b}_{1(t-1)} + \Delta\hat{b}_{1t}\right)\left(\sum_{i=1}^{t-1} i + t\right) \tag{2.74}$$

$$\sum_{i=1}^{t-1} iX_i + tX_t = \left(\hat{b}_{0(t-1)} + \Delta\hat{b}_{0t}\right)\left(\sum_{i=1}^{t-1} i + t\right) + \left(\hat{b}_{1(t-1)} + \Delta\hat{b}_{1t}\right)\left(\sum_{i=1}^{t-1} i^2 + t^2\right) \tag{2.75}$$

เนื่องจากค่า $\hat{b}_{0(t-1)}$ และ $\hat{b}_{1(t-1)}$ เป็นค่าที่ทำให้ผลบวกของ $\hat{\varepsilon}_i^2$ ตั้งแต่ $i=1$ ถึง $i=(t-1)$ มีค่าต่ำสุด สมการ (2.74) และ (2.75) จึงกลายเป็น

$$X_t = \hat{b}_{0(t-1)} + t\Delta\hat{b}_{0t} + t\Delta\hat{b}_{1(t-1)} + \Delta\hat{b}_{1t}\sum_{i=1}^{t} i \tag{2.76}$$

$$tX_t = t\hat{b}_{0(t-1)} + \Delta\hat{b}_{0t}\sum_{i=1}^{t} i + t^2\hat{b}_{1(t-1)} + \Delta\hat{b}_{1t}\sum_{i=1}^{t} i^2 \tag{2.77}$$

ค่า $\hat{b}_{0(t-1)}$+t$\hat{b}_{1(t-1)}$ เป็นค่าพยากรณ์ X_t ที่กระทำ ณ หน่วยเวลา $(t-1)$ สมการ (2.76) และ (2.77) จึงอาจเขียนในรูปของความคลาดเคลื่อนได้เป็น

$$t\Delta\hat{b}_{0t} + \sum_{i=1}^{t} i\Delta\hat{b}_{1t} = e_t \tag{2.78}$$

$$\sum_{i=1}^{t} i\Delta\hat{b}_{0t} + \sum_{i=1}^{t} i^2\Delta\hat{b}_{1t} = te_t \tag{2.79}$$

ค่า $\Delta\hat{b}_{0t}$ และ $\Delta\hat{b}_{1t}$ จึงเขียนได้เท่ากับ

$$\begin{aligned}\Delta\hat{b}_{0t} &= \frac{e_t\sum_{i=1}^{t} i^2 - te_t\sum_{i=1}^{t} i}{D_t} \\ &= \frac{t\left(1-t^2\right)e_t}{6D_t}\end{aligned} \tag{2.80}$$

$$\Delta\hat{b}_{1t} = \frac{t(t-1)e_t}{2D_t} \tag{2.81}$$

โดย

$$
\begin{aligned}
D_t &= t\sum_{i=1}^{t} i^2 - \left(\sum_{i=1}^{t} i^2\right)^2 \\
&= t\left(\sum_{i=1}^{t-1} i^2 + t^2\right) - \left(\sum_{i=1}^{t-1} i + 1\right)^2 \\
&= D_{t-1} + \frac{t(t-1)(2t-1)}{6} \qquad (2.82)
\end{aligned}
$$

ค่าพยากรณ์ k หน่วยเวลาล่วงหน้าที่กระทำ ณ เวลา t อาจเขียนได้เท่ากับ

$$
\begin{aligned}
\hat{X}_t(k) &= \hat{b}_{0t} + \hat{b}_{1t}(t+k) \\
&= \hat{X}_{t-1}(1) + \Delta\hat{b}_{0t} + k\hat{b}_{1t} + t\Delta\hat{b}_{1t} \qquad (2.83)
\end{aligned}
$$

ซึ่งจะเห็นได้ว่า เส้นพยากรณ์ที่กระทำ ณ หน่วยเวลา t เมื่อลากย้อนกลับไปในอดีต อาจไม่ผ่านค่าพยากรณ์ที่กระทำ ณ เวลาต่าง ๆ ที่มิใช่เวลา t เพราะจุดตัดที่แกน เมื่อ $t=0$ และความลาดชันที่ประมาณค่า ณ หน่วยเวลา t จะถูกปรับเปลี่ยนด้วย $\Delta\hat{b}_{0t}$ และ $\Delta\hat{b}_{1t}$ ในสมการ (2.80) และ (2.81) ตามลำดับ

การคำนวณตามสมการ (2.72) และ (2.73) จะต้องมีค่าเริ่มต้นสำหรับ $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ ในการปฏิบัติ จะกำหนดให้

$$
\begin{aligned}
\hat{b}_{0t} &= X_1 \\
\hat{b}_{0t} &= 0 \\
D_1 &= 0
\end{aligned}
$$

2. ตัวแบบแนวโน้มเชิงเส้นระยะสั้น (Local Linear Trend Model)

ตัวแบบแนวโน้มเชิงเส้นระยะยาวนั้นต้องอยู่บนสมมติฐานที่ว่าแนวโน้มของข้อมูลมีลักษณะเชิงเส้นและไม่เปลี่ยนแปลงเป็นระยะเวลานาน ซึ่งอาจไม่เป็นจริงในบางกรณีที่เกิดขึ้นในโลกแห่งความเป็นจริง แนวโน้มของข้อมูลอาจไม่เป็นเชิงเส้นแต่อาจประมาณได้ด้วยแนวโน้มเชิงเส้นหลาย ๆ เส้นที่มีความลาดชันที่แตกต่างกัน จึงอาจกล่าวได้ว่าในช่วงระยะเวลาสั้น (Locality of time) แนวโน้มของข้อมูลมีลักษณะเชิงเส้นและคงที่ เพื่อตอบสนองต่อข้อมูลที่มีแนวโน้มที่อาจไม่เป็นเชิงเส้นดังกล่าว จึงได้พัฒนาตัวแบบแนวโน้มเชิงเส้นระยะสั้นขึ้นโดยความลาดชันของเส้นพยากรณ์สามารถปรับเปลี่ยนได้ ในที่นี้จะกล่าวถึงตัวแบบแนวโน้มเชิงเส้นระยะสั้น โดยมีวิธีการ 4 วิธี คือ วิธีการของ Holt วิธีกำลังสองต่ำสุดแบบถ่วงน้ำหนัก (Discounted Least

Squares) วิธีการเฉลี่ยเคลื่อนที่สองชั้น (Double Moving Average) และวิธีการทำให้เรียบแบบเอกซ์โปแนนเชียล สองชั้น (Double Exponential Smoothing)

2.1) วิธีการของ Holt (Holt's Method)

ในตัวแบบแนวโน้มเชิงเส้นระยะยาวพารามิเตอร์ b_0 และ b_1 จะกำหนดค่าให้เท่ากับ $\hat{b}_0$ และ $\hat{b}_1$ ตามสมการ (2.61) และ(2.62) ซึ่งมีค่าคงที่ วิธีการของ Holt มีแนวความคิดที่จะปรับค่า $\hat{b}_0$ และ $\hat{b}_1$ ตามลักษณะของข้อมูลที่ไหลเข้ามาโดยวิธีการทำให้เรียบแบบเอกซ์โปแนนเชียล เนื่องจากข้อมูล อาจมีความลาดชันแตกต่างกันที่หน่วยเวลาต่าง ๆ การนิยาม b_0 ดังในตัวแบบแนวโน้มเชิงเส้นระยะยาว ซึ่งเป็นค่าของอนุกรมที่จุดเริ่มต้นของการนับเวลา (time origin) อาจมีความไม่เหมาะสมและเกิดความสับสนได้ง่าย เมื่อหน่วยเวลาเคลื่อนห่างไกลออกจากจุดเริ่มต้น การเปลี่ยนแปลงความลาดชัน อาจทำให้ค่าของ b_0 เปลี่ยนแปลงอย่างมากมายได้ จึงเห็นควรที่จะนิยาม b_{0_t} เป็นระดับ (level) ของข้อมูล ณ จุดเวลา t และ b_{1_t} เป็นความลาดชันของข้อมูล ณ จุดเวลา t ดังนั้น ตัวแบบ ณ จุดเวลา t จึงเขียนได้เป็น

$$X_{t+k} = b_{ot} + b_{1t}k + \varepsilon_{t+k} \quad (2.84)$$

ตามวิธีการของ Holt ระดับของข้อมูลจะประมาณค่าโดยวิธีการทำให้เรียบแบบเอกซ์โปแนนเชียลด้วยค่าคงที่การทำให้เรียบ α

$$\hat{b}_{0t} = \alpha X_t + (1-\alpha)\hat{X}_{t-1}(1) \quad (2.85)$$

แนวความคิดที่จะประมาณความลาดชันของข้อมูล ณ จุดเวลา t อาจมีได้หลายแนวความคิด แนวความคิดหนึ่งจะตั้งอยู่บนพื้นฐานของสมการ (2.84) กล่าวคือ เมื่อ $k = 0$ และ $k = -1$ สมการ (2.84) จะกลายเป็น

$$X_t = b_{0t} + \varepsilon_t$$

$$X_{t-1} = b_{0t} + b_{1t} + \varepsilon_{t-1}$$

ผลต่างของสมการทั้ง 2 จะเขียนได้เท่ากับ

$$X_t - X_{t-1} = b_{1t} + \varepsilon_t - \varepsilon_{t-1} \quad (2.86)$$

ซึ่งมีค่าคาดหมายเท่ากับ

$$E(X_t - X_{t-1}) = b_{1t} \quad (2.87)$$

ดังนั้น ผลต่าง $X_t - X_{t-1}$ น่าจะใช้เป็นค่าประมาณของ b_{1t} ได้ แต่เมื่อพิจารณาสมการ (2.86)ให้ถ่องแท้ จะเห็นว่าผลต่าง $X_t - X_{t-1}$ มีความแปรปรวนเป็น 2 เท่าของความแปรปรวนของความรบกวนสุ่ม จึงอาจทำให้การประมาณค่า b_{1t} ด้วยผลต่าง $X_t - X_{t-1}$ มีความคลาดเคลื่อนได้มาก

แนวความคิดในการประมาณค่า b_{1t} อีกแนวหนึ่งก็คือ ใช้คำนิยามของ b_{0t} และ b_{1t} ซึ่งทำให้อาจเขียนความสัมพันธ์ต่อไปนี้ได้

$$b_{0t} \quad = \quad b_{0(t-1)} + b_{1(t-1)} \tag{2.88}$$

ดังนั้น ความลาดชันล่าสุดของข้อมูลที่อาจทราบได้ ณ จุดเวลา t ก็คือความลาดชันที่เกิดขึ้น ณ จุดเวลา t -1 $b_{t(t-1)}$ ซึ่งอาจประมาณค่าได้จาก

$$\hat{b}_{1(t-1)} = \hat{b}_{01} - \hat{b}_{0(t-1)} \tag{2.89}$$

การประมาณค่า b_{1t} จึงอาจใช้แนวความคิดของการทำให้เรียบแบบเอกซ์โปเนนเชียล ด้วยค่าคงที่การทำให้เรียบ

$$\hat{b}_{1t} = \beta\left(\hat{b}_{0t} - \hat{b}_{0(t-1)}\right) + (1-\beta)\hat{b}_{1(t-1)} \tag{2.90}$$

เนื่องจาก b_{1t} เป็นความลาดชันซึ่งมักจะไม่เปลี่ยนแปลงอย่างรวดเร็ว ค่าคงที่การทำให้เรียบ β จึงไม่ควรมีค่ามากให้ค่า $\hat{b}_{1t}$ ในสมการ(2.90) เปลี่ยนแปลงช้า

เมื่อสามารถประมาณค่า $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ ตามสมการ (2.85) และ (2.90) ตามลำดับ k ค่าพยากรณ์ k หน่วยเวลาล่วงหน้าที่พยากรณ์ ณ เวลา t จึงอาจเขียนได้เป็น

$$\hat{x}_t(k) = \hat{b}_{0t} + \hat{b}_{1t}\, k \tag{2.91}$$

สูตรที่ใช้ประมาณค่าระดับของข้อมูลในสมการ (2.85) อาจเขียนได้ใหม่ในรูปของการปรับแก้ด้วยความคลาดเคลื่อน (error correction form) ได้เป็น

$$\hat{b}_{0t} = \hat{X}_{t-1}(1) + \alpha e_t \tag{2.92}$$

ซึ่งเมื่อแทน $\hat{X}_{t-1}(1)$ ด้วยสมการ (2.91) ได้เป็น

$$\hat{b}_{0t} = \hat{b}_{0(t-1)} + \hat{b}_{1(t-1)} + \alpha e_t \tag{2.93}$$

เมื่อนำค่าผลต่าง $\hat{b}_{0t}-\hat{b}_{0(t-1)}$ ในสมการ(2.93) ไปแทนในสมการ (2.90) จะได้สูตรที่ใช้ประมาณค่าความลาดชันในรูปการปรับแก้ด้วยความคลาดเคลื่อน ดังนี้

$$\hat{b}_{1t} = \hat{b}_{1(t-1)} + \alpha\beta e_t \tag{2.94}$$

การเริ่มต้นการคำนวณ $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ ตามสมการ (2.93) และ (2.94) ตามลำดับจะต้องทราบค่า $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ เมื่อ $t = 0$ ซึ่งค่าทั้ง 2 นี้ อาจประมาณค่าโดยวิธีการที่กล่าวในตัวแบบแนวโน้มเชิงเส้นระยะยาว แต่ควรกระทำด้วยความรอบคอบ โดยการคัดเลือกข้อมูลจำนวนหนึ่งในช่วงแรกจากข้อมูลทั้งหมดเท่านั้น มิใช่นำข้อมูลทั้งหมดประมาณค่า $\hat{b}_0$ และ $\hat{b}_1$ จากตัวแบบแนวโน้มเชิงเส้นระยะยาว ทั้งนี้เพราะแนวโน้มระยะยาวอาจมีทิศทางตรงข้ามกับแนวโน้มระยะสั้นก็ได้ ตามสมมติฐานของตัวแบบแนวโน้มระยะสั้นที่ยอมให้ความลาดชันในช่วงระยะเวลาหนึ่งอาจเปลี่ยนแปลงไปจากช่วงระยะเวลาเพิ่งผ่านมาได้ ดังนั้น ในทางปฏิบัติมักจะนิยมที่กำหนดให้

$$\begin{aligned}
\hat{b}_{01} &= X_1 \\
\hat{b}_{11} &= 0 \\
\hat{X}_1(1) &= \hat{b}_{01} + \hat{b}_{11} \\
&= X_1
\end{aligned}$$

เพื่อมิต้องใช้ตัวแบบแนวโน้มเชิงเส้นระยะยาวประมาณค่าเริ่มต้น อนึ่งหากข้อมูลมีจำนวนพอสมควรเพื่อใช้ทดสอบตัวแบบพยากรณ์ อิทธิพลของค่าเริ่มต้นที่มีผลต่อค่าพยากรณ์ ณ หน่วยเวลาสุดท้ายจะน้อยมาก เพราะน้ำหนักที่ใช้ในการถ่วงเป็นแบบเอกซ์โปเนนเชียล

- วิธีกำลังสองต่ำสุดแบบถ่วงน้ำหนัก (Discounted Least Squares)

ในตัวแบบ Global Trend การหาค่าประมาณของพารามิเตอร์ b_0 และ b_1 ใช้วิธีที่ทำให้ผลบวกของความคลาดเคลื่อนของค่าพยากรณ์กำลังสองมีค่าต่ำสุด นั่นคือ

$$\text{minimize} \qquad S = \sum_{i=1}^{t} \hat{\varepsilon}_i^2$$

โดย $\hat{\varepsilon}_i$ เป็นผลต่างระหว่างข้อมูลจริง ณ เวลา i กับค่าบนเส้นพยากรณ์ ณ จุดเวลาเดียวกัน วิธีการประมาณค่าพารามิเตอร์ b_0 และ b_1 ดังกล่าว ให้น้ำหนักแก่ความคลาดเคลื่อน ณ เวลาต่าง ๆ เท่ากันหมด ซึ่งอาจจะไม่เหมาะสมสำหรับข้อมูลแนวโน้มมีลักษณะคงที่ในระยะสั้น ๆ และอาจเปลี่ยนทิศทางได้ สำหรับข้อมูลที่มีลักษณะเช่นนั้น น้ำหนักที่ให้แก่ความคลาดเคลื่อนที่ใกล้กับ

ปัจจุบัน ควรมากกว่าน้ำหนักที่ให้แก่ความคลาดเคลื่อนที่ห่างไกลจากปัจจุบัน ดังนั้น จึงได้มีการนำน้ำหนักแบบเอกซ์โปเนนเชียลมาใช้ในการกำหนดน้ำหนักให้แก่ความคลาดเคลื่อนเกณฑ์ที่ใช้ในการประมาณค่า b_0 และ b_1 จึงกลายเป็น

$$\min \quad S = \sum_{r=0}^{t} a^r \hat{\varepsilon}_{t-r}^{2} \tag{2.95}$$

โดย a มีค่าระหว่าง 0 กับ 1 (0< a <1) เกณฑ์ในสมการ (2.95) เป็นที่รู้จักกันในนามเกณฑ์กำลังสองต่ำสุดแบบถ่วงน้ำหนัก (discounted least squares criterion)
จากสมการ (2.84) ε_{t-r} อาจเขียนได้เป็น

$$\hat{\varepsilon}_{t-r} = x_{t-r} - \hat{b}_{0t} + \hat{b}_{1t}\, r \tag{2.96}$$

เมื่อนำค่า $\hat{\varepsilon}_{t-r}$ จากสมการ (2.96) ไปแทนค่าในสมการ (2.95) จะได้ผลดังนี้

$$S = \sum_{r=0}^{t} a^r \left(X_{t-r} - \hat{b}_{0t} + \hat{b}_{1t}\, r \right)^2 \tag{2.97}$$

เงื่อนไขที่ทำให้ S มีค่าต่ำสุด อาจเขียนได้เป็น

$$\frac{\partial S}{\partial \hat{b}_{0t}} = -2 \sum_{r=0}^{t-1} a^r \left(X_{t-r} - \hat{b}_{0t} + \hat{b}_{1t}\, r \right) = 0$$

$$\hat{b}_{0t} \sum_{r=0}^{t-1} a^r - \hat{b}_{1t} \sum_{r=0}^{t-1} r a^r = \sum_{r=0}^{t-1} a^r X_{t-r} \tag{2.98}$$

$$\frac{\partial S}{\partial \hat{b}_{1t}} = 2 \sum_{r=0}^{t-1} a^r \left(X_{t-r} - \hat{b}_{0t} + \hat{b}_{1t}\, r \right) r = 0$$

$$\hat{b}_{0t} \sum_{r=0}^{t-1} r a^r - \hat{b}_{1t} \sum_{r=0}^{t-1} r^2 a^r = \sum_{r=0}^{t-1} r a^r X_{t-r} \tag{2.99}$$

จากสมการ (2.98) และ (2.99) $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ อาจเขียนได้เท่ากับ

$$\hat{b}_{0t} = \frac{-\sum_{r=0}^{t-1} a^r X_{t-r} \sum_{r=0}^{t-1} r^2 a^r X_{t-r} + \sum_{r=0}^{t-1} ra^r X_{t-r} \sum_{r=0}^{t-1} ra^r}{-\sum_{r=0}^{t-1} a^r \sum_{r=0}^{t-1} r^2 a^r + \left(\sum_{r=0}^{t-1} ra^r\right)^2} \tag{2.100}$$

$$\hat{b}_{1t} = \frac{\sum_{r=0}^{t-1} a^r \sum_{r=0}^{t-1} ra^r X_{t-r} - \sum_{r=0}^{t-1} ra^r \sum_{r=0}^{t-1} a^r X_{t-r}}{-\sum_{r=0}^{t-1} a^r \sum_{r=0}^{t-1} r^2 a^r + \left(\sum_{r=0}^{t-1} ra^r\right)^2} \tag{2.101}$$

การคำนวณค่า $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ ตามสมการ (2.100) และ (2.101) ตามลำดับ มีความไม่เหมาะสมและเป็นภาระมาก เมื่อต้องคำนวณค่า $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ หลาย ๆ หน่วยเวลา จึงควรที่จะปรับสูตรการคำนวณค่า $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ เป็นสูตรแบบ recurrence เพื่อมิให้ภาระการคำนวณเพิ่มขึ้นตามปริมาณข้อมูล

ให้
$$W_t = \sum_{r=0}^{t-1} a^r \tag{2.102}$$

ซึ่งอาจเขียนใหม่ได้เป็น

$$W_t = 1 + aW_{t-1} \tag{2.103}$$

ให้
$$A_t = \sum_{r=0}^{t-1} ra^r \tag{2.104}$$

ซึ่งกระจายออกได้เป็น

$$\begin{aligned} A_t &= a + 2a^2 + 3a^2 + 4a^4 + \ldots + (t-1)a^{t-1} \\ &= a(1 + 2a + 3a + 4a^3 + \ldots + (t-1)a^{t-2}) \\ &= a(1 + a + a^2 + a^3 + \ldots a^{t-2}) + a + 2a^2 + 3a^2 + \ldots + (t-2)a^{t-2}) \\ &= a(W_{t-1} + A_{t-1}) \end{aligned} \tag{2.105}$$

ให้
$$B_t = \sum_{r=0}^{t-1} r^2 a^r \tag{2.106}$$

$$= a + 2^2 a^2 + 3^2 a^3 + 4^2 a^4 + ... + (t-1)^2 a^{t-1}$$
$$= a + a^2 + a^3 + a^4 + ... + a^{t-1} + a^2 + 2^2 a^3 + 3^2 a^4 + ... + (t-2)^2 a^{t-1}$$
$$= aW_{t-1} + 2aA_{t-1} + aB_{t-1}$$
$$= a(W_{t-1} + A_{t-1}) + (A_{t-1} + B_{t-1})$$
$$= A_t + a(A_{t-1} + B_{t-1}) \qquad (2.107)$$

ให้ $$Y_t = \sum_{r=0}^{t-1} a^r X_{t-r} \qquad (2.108)$$

$$= X_t + aX_{t-1} + a^2 X_{t-2} + ... + a^{t-1} X_1$$
$$= X_t + a(X_{t-1} + a^2 X_{t-2} + ... + a^{t-2} X_1)$$
$$= X_t + aY_{t-1} \qquad (2.109)$$

ให้ $$Z_t = \sum_{r=0}^{t-1} ra^r X_{t-r} \qquad (2.110)$$

$$= aX_{t-1} + 2a^2 X_{t-2} + 3a^3 X_{t-3} + ... + (t-1)a^{t-1} X_1$$
$$= a(X_{t-1} + aX_{t-2} + a^2 X_{t-3} + ... + a^{t-2} X_1 + aX_{t-2} + 2a^2 X_{t-3} + ... + (t-2)a^{t-2} X_1)$$
$$= a(Y_{t-1} + Z_{t-1}) \qquad (2.111)$$

ค่าเริ่มต้นของสัญลักษณ์เหล่านี้ตามคำจำกัดความ คือ

$W_1 = 1$

$A_1 = 0$

$B_1 = 0$

$Y_1 = X_1$

$Z_1 = 0$

ดังนั้น เมื่อนำ w_t , A_t , B_t , Y_t และ Z_t ไปแทนค่าในสมการปกติ (2.98) และ (2.99) จะได้

$$W_t \hat{b}_{0t} - A_t \hat{b}_{1t} = Y_t \qquad (2.112)$$
$$A_t \hat{b}_{0t} - A_t \hat{b}_{1t} = Z_t$$

ซึ่งจากการแก้สมการ 2 สมการนี้ ได้ผลได้ดังนี้

$$\hat{b}_{0t} = C_t\,(B_t Y_t - A_t Z_t) \tag{2.113}$$

$$\hat{b}_{1t} = C_t\,(A_t Y_t - W_t Z_t)$$

โดย

$$C_t = 1/(B_t W_t - A_t^2) \tag{2.114}$$

จะเห็นได้ว่า ณ จุดเวลา t มีข้อมูล X_t เข้ามาใหม่ 1 ค่า การคำนวณค่า w_t ตามสมการ (2.103) A_t ตามสมการ (2.105) B_t ตามสมการ (2.107) Y_t ตามสมการ (2.109) และ Z_t ตามสมการ (2.111) สามารถกระทำได้โดยง่าย เมื่อระยะเวลาผ่านไปนาน ค่า w_t, A_t, B_t และ C_t จะเข้าใกล้ค่าดังนี้

$$\begin{aligned} \mathrm{W} &= \lim_{t\to\infty} W_t \\ &= \lim_{t\to\infty} \sum_{r=0}^{t-1} a^r \\ &= \frac{1}{1-a} \end{aligned} \tag{2.115}$$

$$\begin{aligned} \mathrm{A} &= \lim_{t\to\infty} A_t \\ &= \lim_{t\to\infty} \sum_{r=0}^{t-1} r a^r = \frac{a}{(1-a)^2} \end{aligned} \tag{2.116}$$

$$\begin{aligned} \mathrm{B} &= \lim_{t\to\infty} B_t \\ &= \lim_{t\to\infty} \sum_{r=0}^{t-1} r^2 a^r = \frac{\mathrm{a}}{(1-\mathrm{a})^2} \end{aligned} \tag{2.117}$$

$$\begin{aligned} \frac{1}{C} &= \lim_{t\to\infty} \left(B_t W_t - A_t^2\right) \\ &= \frac{a(1+a)}{(1-a)^3}\cdot\frac{1}{1-a} - \frac{a^2}{(1-a)^4} = \frac{a}{(1-a)^4} \end{aligned} \tag{2.118}$$

- ค่าเฉลี่ยเคลื่อนที่สองชั้น (Double Moving Average)

ดังที่ได้กล่าวมาแล้วว่า หากข้อมูลมีแนวโน้มขึ้น ค่าที่คำนวณจากวิธีการเฉลี่ยเคลื่อนที่ชั้นเดียว จะมีค่าต่ำกว่าข้อมูลจริง แต่ก็มีแนวโน้มเพิ่มขึ้นเช่นกัน และหากข้อมูลมีแนวโน้มลดลง ค่าที่คำนวณจากวิธีการเฉลี่ยเคลื่อนที่ชั้นเดียว จะมีค่าสูงกว่าข้อมูลจริง โดยมีแนวโน้มลดลงเช่นเดียวกับข้อมูล ดังนั้น ในกรณีที่ข้อมูลมีแนวโน้มเป็นเส้นตรงหากนำค่าที่คำนวณได้จากวิธีการเฉลี่ยเคลื่อนที่ชั้นเดียวมาคำนวณด้วยวิธีการเฉลี่ยเคลื่อนที่อีกครั้งในลักษณะเดียวกัน ค่าที่คำนวณด้วยวิธีการเฉลี่ยเคลื่อนที่อีกครั้งเมื่อเปรียบเทียบกับค่าที่คำนวณด้วยวิธีการเฉลี่ยเคลื่อนที่ชั้นเดียว จะมีลักษณะเดียวกันกับค่าที่คำนวณด้วยวิธีการเฉลี่ยเคลื่อนที่ชั้นเดียวเมื่อเปรียบเทียบกับข้อมูลจริง จึงมีแนวความคิดที่จะนำผลต่างระหว่างที่คำนวณด้วยวิธีการเฉลี่ยเคลื่อนที่สองชั้นกับค่าที่คำนวณด้วยวิธีการเฉลี่ยเคลื่อนที่ชั้นเดียวมาปรับแก้ค่าพยากรณ์

ตัวแบบข้อมูลที่มีแนวโน้มเชิงเส้นอาจอธิบายได้ด้วยสมการ

$$X_t = b_0 + b_1 t + \varepsilon_t \tag{2.119}$$

ค่าที่คำนวณด้วยวิธีการเฉลี่ยเคลื่อนที่ชั้นเดียว ณ เวลา t อาจเขียนได้เป็น

$$M_{1t} = \frac{X_{t-n+1} + \ldots + X_t}{n}$$

$$= \frac{1}{n}\left[\{b_0 + b_1(t-n+1) + \varepsilon_{t-n+1}\} + \{b_0 + b_1(t-n+2) + \varepsilon_{t-n+2}\} + \ldots + \{b_0 + b_1 t + \varepsilon_t\}\right]$$

$$= b_0 + b_1 t - \frac{n-1}{2} b_1 + \frac{1}{n}\sum_{r=0}^{n-1} \varepsilon_{t-r} \tag{2.120}$$

เนื่องจากค่าคาดหมายของการรบกวนสุ่มมีค่าเท่ากับศูนย์ พจน์สุดท้ายทางด้านขวามือของสมการ (2.120) จึงอาจกำหนดให้ค่าเท่ากับศูนย์ได้ ดังนั้นค่าประมาณของ b_0 และ b_1 จึงอาจเขียนได้จากสมการ (2.120) ได้เป็น

$$M_{1t} = \hat{b}_0 + \hat{b}_1 t - \frac{n-1}{2}\hat{b}_1 \tag{2.121}$$

ค่าเฉลี่ยเคลื่อนที่สองชั้น ณ เวลา t ซึ่งเป็นค่าเฉลี่ยของค่าพยากรณ์ด้วยวิธีการเฉลี่ยเคลื่อนที่ชั้นเดียวล่าสุด n ค่า จึงเขียนได้เท่ากับ

$$M_{2t} = \frac{M_{1(t-n+1)} + M_{1(t-n+2)} \ldots + M_{1t}}{n}$$

$$= \frac{1}{n}\left[\left\{\hat{b}_0 + \hat{b}_1(t-n+1) - \frac{n-1}{2}\hat{b}_1\right\} + \left\{\left\{\hat{b}_0 + \hat{b}_1(t-n+2)\right) - \frac{n-1}{2}\hat{b}_1\right\} + \ldots + \left\{\hat{b}_0 + \hat{b}_1 - \frac{n-1}{2}\hat{b}_1\right\}\right]$$

$$= \hat{b}_0 + \hat{b}_1 t - (n-1)\hat{b}_1 \qquad (2.122)$$

เมื่อลบสมการ (2.121) ด้วยสมการ (2.122) จะได้ผลลัพธ์

$$\hat{b}_1 = \frac{2}{n-1}(M_{1t} - M_{2t}) \qquad (2.123)$$

ค่าพยากรณ์ k หน่วยเวลาล่วงหน้าที่พยากรณ์ ณ เวลา t อาจเขียนได้จากสมการ (2.119) เป็น

$$\hat{X}_t(k) = \hat{b}_0 + \hat{b}_1(t+k) \quad = \hat{X}_t(0) + k\hat{b}_1$$

โดย $\hat{X}_t(0)$ คือ ค่าพยากรณ์ระดับของข้อมูล ณ เวลา t หรืออีกนัยหนึ่งคือจุดเริ่มต้นของเส้นพยากรณ์ที่พยากรณ์ ณ เวลา t นั่นเอง ดังนั้น เพื่อมิให้เกิดความสับสนในสัญลักษณ์จึงให้

$$\hat{b}_{0t} = \hat{X}_t(0) \quad = \hat{b}_0 + \hat{b}_1 t \qquad (2.124)$$

เมื่อแทนค่า $\hat{b}_0 + \hat{b}_1 t$ จากสมการ (2.122) ลงในสมการ (2.124) จะได้ผลลัพธ์

$$\hat{b}_{0t} = M_{2t} + (n-1)\hat{b}_1$$

ซึ่งเมื่อนำค่า $\hat{b}_1$ จากสมการ (2.123) มาแทนแล้ว สมการจะกลายเป็น

$$\hat{b}_{0t} = 2M_{1t} - M_{2t} \qquad (2.125)$$

เนื่องจากความลาดชันในสมการ (2.123) มีค่าเปลี่ยนแปลงตามค่าเฉลี่ยเคลื่อนที่ชั้นเดียว และสองชั้นที่คำนวณ ณ เวลา t จึงเห็นควรที่จะปรับสัญลักษณ์ $\hat{b}_1$ เพื่อให้แสดงถึงการเปลี่ยนแปลงที่อาจเกิดขึ้นได้ตามหน่วยเวลาเป็น $\hat{b}_{1t}$

$$\hat{b}_{1t} = \frac{2}{n-1}(M_{1t} - M_{2t}) \qquad (2.126)$$

ดังนั้น ค่าพยากรณ์ k หน่วยเวลาล่วงหน้าพยากรณ์ ณ เวลา t จึงอาจเขียนได้เป็น

$$\hat{X}_t(k) = \hat{b}_{0t} + k\hat{b}_{1t} \qquad (2.127)$$

- การทำให้เรียบแบบเอกซ์โปเนนเชียลสองชั้น (Double Exponential Smoothing)

วิธีนี้อาจรู้จักกันในอีกชื่อว่าการทำให้เรียบเชิงเส้นแบบเอกซ์โปเนนเชียลของ Brown (Brown's Linear Exponential Smoothing) ซึ่งมีแนวความคิดทำนองเดียวกันกับวิธีการเฉลี่ยเคลื่อนที่สองชั้น ในกรณีที่ข้อมูลมีแนวโน้มที่อาจเปลี่ยนแปลงได้ดังในสมการ (2.72) ค่าที่คำนวณด้วยการทำให้เรียบชั้นเดียว ณ เวลา t อาจเขียนได้เท่ากับ

$$S_{1t} = \frac{\sum_{r=0}^{t-1} a^r X_{t-r}}{\sum_{r=0}^{t-1} a^r} = \frac{\sum_{r=0}^{t-1} a^r (b_{0t} - b_{1t} r + \varepsilon_{t-r})}{\sum_{r=0}^{t-1} a^r}$$

ซึ่งมีค่าคาดหมายเท่ากับ

$$E(S_{1t}) = b_{0t} - b_{1t} \frac{\sum_{r=0}^{t-1} r a^r}{\sum_{r=0}^{t-1} a^r} \qquad (2.128)$$

เมื่อ t มีค่ามาก สัมประสิทธิ์ของ b_{1t} จะมีค่าเข้าใกล้ $a/(1-a)$ ดังนั้น เมื่อ t มีค่ามาก สมการ (2.128) จึงกลายเป็น

$$E(S_{1t}) = b_{0t} - \frac{a}{1-a} b_{1t} \qquad (2.129)$$

จึงอาจกล่าวได้ว่าค่าที่คำนวณด้วยการทำให้เรียบชั้นเดียวในกรณีที่ข้อมูลมีแนวโน้มเพิ่มขึ้นเป็นเส้นตรง จะมีค่าต่ำกว่าจุดเริ่มต้นของเส้นพยากรณ์ด้วยจำนวน $ab_{1t}/(1-a)$ และในกรณีที่ข้อมูลมีแนวโน้มลดลงเป็นเส้นตรง ค่าที่คำนวณด้วยการทำให้เรียบชั้นเดียว จะมีค่าสูงกว่าจุดเริ่มต้นของเส้นพยากรณ์ด้วยจำนวน $ab_{1t}/(1-a)$ เช่นกัน ดังนั้น การประมาณ b_{0t} จึงอาจเขียนได้จากสมการ (2.129) เป็น

$$\hat{b}_{0t} = S_{1t} + \frac{a}{1-a} b_{1t} \tag{2.130}$$

เมื่อนำค่าที่ได้คำนวณไว้ทำให้เรียบอีกครั้ง จะได้ค่าที่คำนวณด้วยการทำให้เรียบสองชั้นเป็น

$$S_{1t} = \frac{\sum_{r=0}^{t-1} a^r S_{1(1-r)}}{\sum_{r=0}^{t-1} a^r} \tag{2.131}$$

ตามที่ได้กล่าวมาแล้วในกรณีที่ข้อมูลมีแนวโน้มเป็นเส้นตรง ค่าพยากรณ์ก็มีแนวโน้มเป็นเส้นตรงเช่นกัน แต่จะสูงกว่าหรือต่ำกว่าข้อมูลจริงด้วยจำนวนหนึ่งขึ้นอยู่กับแนวโน้มของข้อมูลจะลดลงหรือเพิ่มขึ้น ดังนั้น S_{2t} เมื่อเปรียบเทียบกับ S_{1t} จึงเหมือนกับ S_{1t} เมื่อเปรียบเทียบกับข้อมูลจริง

$$S_{1t} = S_{2t} + \frac{a}{1-a} \hat{b}_{1t}$$

ให้ $$\hat{b}_{1t} = \frac{1-a}{a} (S_{1t} - S_{2t}) \tag{2.132}$$

เมื่อแทนค่า $\hat{b}_{1t}$ ในสมการ (2.132) ลงในสมการ (2.130) จะได้ผลลัพธ์

$$\hat{b}_{0t} = 2S_{1t} - S_{2t} \tag{2.133}$$

ค่าพยากรณ์ k หน่วยเวลาล่วงหน้าที่พยากรณ์ ณ เวลา t จึงเขียนได้เป็น

$$\hat{X}_t(k) = \hat{b}_{0t} + k\hat{b}_{1t} \tag{2.134}$$

การคำนวณค่า $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ เมื่อ $t = 0$ ตามสมการ (2.133) และ (2.132) ตามลำดับ จะต้องทราบค่า S_{10} และ S_{20} ซึ่ง Brown ได้เสนอแนะให้กำหนดค่า S_{10} และ S_{20} ดังนี้

$$S_{10} = \hat{b}_0 - \frac{a}{1-a} \hat{b}_1 \tag{2.135}$$

$$S_{10} = \hat{b}_0 - \frac{2a}{1-a} \hat{b}_1 \tag{2.136}$$

โดย $\hat{b}_0$ และ $\hat{b}_1$ เป็นค่าที่ได้จากตัวแบบแนวโน้มเชิงเส้นระยะยาว อนึ่งการใช้ตัวแบบแนวโน้มเชิงเส้นระยะยาว มีประมาณค่าเริ่มต้นของการคำนวณค่าพยากรณ์ของตัวแบบการทำให้เรียบแบบเอกซ์โปเนนเชียลสองชั้น ซึ่งเป็นตัวแบบแนวโน้มเชิงเส้นระยะสั้น ต้องกระทำด้วยความรอบคอบ ดังที่ได้กล่าวไว้ในกรณีวิธีการ Holt เมื่อหลีกเลี่ยงปัญหาที่อาจเกิดขึ้นจากการเปลี่ยนแปลงทิศทางของแนวโน้มในกรณีที่มีข้อมูลจำนวนมาก ค่าเริ่มต้นการพยากรณ์มักนิยมกำหนดดังนี้

$$S_{10} = X_1 \tag{2.137}$$

$$S_{21} = S_{11} \tag{2.138}$$

สูตรที่นิยมใช้กันมากจะอยู่ในรูปของค่าที่การทำให้เรียบ α ซึ่งมีค่าเท่ากับ 1-a สมการ (2.132) จึงเขียนใหม่ได้เป็น

$$\hat{b}_{1t} = \frac{\alpha}{1-\alpha}(S_{1t} - S_{2t}) \tag{2.139}$$

$$S_{1t} = \alpha X_t + (1-\alpha)S_{1(t-1)} \tag{2.140}$$

$$S_{2t} = \alpha S_{1t} + (1-\alpha)S_{2(t-1)} \tag{2.141}$$

เนื่องจากวิธีการทำให้เรียบแบบเอกซ์โปเนนเชียลสองชั้นกับวิธีการของ Holt ต่างก็ตั้งอยู่บนพื้นฐานของการทำให้เรียบแบบเอกซ์โปเนนเชียล โดยวิธีการของ Holt ใช้ค่าคงที่การทำให้เรียบ 2 ค่า เป็นอิสระต่อกัน ค่าคงที่การทำให้เรียบค่าหนึ่งสำหรับระดับข้อมูล ส่วนอีกค่าหนึ่งสำหรับความลาดชัน แต่วิธีการทำให้เรียบแบบเอกซ์โปเนนเชียลสองชั้นใช้ค่าคงที่การทำให้เรียบเพียงค่าเดียว จึงควรที่จะศึกษาความสัมพันธ์ระหว่าง 2 วิธีการนี้

สมการ (2.139) อาจเขียนได้ใหม่ในรูปของ error-correction ได้เป็น

$$\begin{aligned}
\hat{b}_{1t} &= \frac{\alpha}{1-\alpha}\left[\alpha X_t + (1-\alpha)S_{1(t-1)}\right] - \frac{\alpha}{1-\alpha}\left[\alpha S_{1t} + (1-\alpha)S_{2(t-1)}\right] \\
&= \frac{\alpha^2}{1-\alpha}X_t + \alpha(S_{1(t-1)} - S_{2(t-1)} - \frac{\alpha^2}{1-\alpha}S_{1t} \\
&= \hat{b}_{1(t-1)} + \alpha^2(X_t - \hat{b}_{0(t-1)} - \hat{b}_{1(t-1)}) \\
&= \hat{b}_{1(1-1)} + \alpha^2 e_t
\end{aligned} \tag{2.142}$$

สมการ (2.130) อาจเขียนใหม่ในรูปของ error-correction ได้เป็น

$$\hat{b}_{0t} = S_{1t} + \frac{1-\alpha}{\alpha}\hat{b}_{1t}$$

$$= \alpha X_t + (1-\alpha)S_{1(t-1)} + \frac{1-\alpha}{\alpha}\left(\hat{b}_{1(t-1)} + \alpha^2 e_t\right)$$

เมื่อนำค่า $S_{1(t-1)}$ ในสมการ (2.130) มาแทนแล้ว $\hat{b}_{0t}$ จะกลายเป็น

$$\hat{b}_{0t} = \alpha X_t + (1-\alpha)\left(\hat{b}_{0(t-1)} - \frac{1-\alpha}{\alpha}\hat{b}_{1(t-1)}\right) + \frac{1-\alpha}{\alpha}\hat{b}_{1(t-1)} + (1-\alpha)\alpha e_t$$

$$= \hat{b}_{0(t-1)} + \hat{b}_{1(t-1)} + \alpha\left[X_t - \hat{b}_{0(t-1)} - \hat{b}_{1(t-1)}\right] + (1-\alpha)\alpha e_t$$

$$= \hat{b}_{0(t-1)} + \hat{b}_{1(t-1)} + \alpha(2-\alpha)e_t \tag{2.143}$$

เพื่อเปรียบเทียบ $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ ในสมการ (2.143) และสมการ (2.142) ตามลำดับกับ $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ ในกรณีวิธีการของ Holt ในสมการ (2.93) และ (2.90) จะเห็นว่าได้ผลการพยากรณ์จะเหมือนกันถ้า

$$\alpha_h = \alpha(2-\alpha) \tag{2.144}$$

$$\beta_h = \frac{\alpha}{2-\alpha} \tag{2.145}$$

โดย α_h และ β_h เป็นค่าคงที่การทำให้เรียบในวิธีการของ Holt ดังนั้น จึงอาจกล่าวได้ว่า วิธีการทำให้เรียบแบบเอกซ์โปเนนเชียลสองชั้นเป็นกรณีพิเศษของวิธีการ Holt โดยกำหนดค่าคงที่การทำให้เรียบ α_h และ β_h ต้องเป็นไปตามสมการ (2.144) และ (2.145) ตามลำดับ

2.3.3 ตัวแบบฤดูกาล (Seasonal Model)

วิธีการพยากรณ์ที่ได้กล่าวมาข้างต้น นี้ เป็นวิธีการพยากรณ์สำหรับข้อมูลที่ไม่มีฤดูกาล แต่ในโลกแห่งความเป็นจริง ในบางกรณีข้อมูลมีฤดูกาล ซึ่งหมายความว่าข้อมูลที่อยู่ห่างไกลจากปัจจุบันอาจมีสาระที่เป็นประโยชน์ต่อการพยากรณ์มากกว่าข้อมูลที่อยู่ใกล้กับปัจจุบัน ซึ่งทำให้การพึ่งพิงของข้อมูลมิได้เป็นไปตามข้อสมมติของตัวแบบที่ได้กล่าวมาแล้ว ในบทนี้จึงจำเป็นต้องพัฒนาตัวแบบฤดูกาลขึ้นมา เพื่อใช้กับข้อมูลที่มีฤดูกาล

ข้อมูลที่มีฤดูกาลอาจมีหรือไม่มีแนวโน้มก็ได้เช่นเดียวกันกับกรณีข้อมูลที่ไม่มีฤดูกาล และการเปลี่ยนแปลงของข้อมูลภายในฤดูกาลอาจมีขนาดคงที่ เป็นอิสระกับเวลา หรือลักษณะการเปลี่ยนแปลงของข้อมูลภายในฤดูกาลอาจแปรตามแนวโน้มก็ได้ Winters ได้พัฒนาวิธีการพยากรณ์

สำหรับข้อมูลที่มีฤดูกาลด้วยวิธีการทำให้เรียบชั้นเดียว โดยใช้ค่าคงที่การทำให้เรียบ 3 ค่า ค่าหนึ่งสำหรับพยากรณ์ระดับข้อมูล อีกค่าสำหรับพยากรณ์ความลาดชัน และค่าสุดท้ายสำหรับพยากรณ์ดัชนีฤดูกาล (seasonal index) ซึ่งเป็นที่รู้จักกันในปัจจุบันในนามวิธีการของ Winters เนื่องจากการเปลี่ยนแปลงของข้อมูลภายในฤดูกาลอาจแปรตามแนวโน้มหรือไม่แปรตามแนวโน้ม จึงจำเป็นที่จะต้องมีตัวแบบเชิงคูณ (multiplicative model) และตัวแบบเชิงบวก (additive model) เพื่อรองรับการเปลี่ยนแปลงของข้อมูลภายในฤดูกาลดังกล่าว

1. ตัวแบบเชิงคูณ (Multiplicative Model)

ในกรณีที่การเปลี่ยนแปลงของข้อมูลภายในฤดูกาลแปรตามแนวโน้มของข้อมูลตัวแบบเชิงคูณของ Winters ซึ่งเขียนได้เป็น

$$X_t = (b_0 + b_1 t)C_t + \varepsilon_t \tag{2.146}$$

มีความเหมาะสมที่จะใช้อธิบายความเคลื่อนไหวของข้อมูล พจน์ $b_0 + b_1 t$ จะอธิบายส่วนที่เป็นแนวโน้มที่มีอยู่ในข้อมูล C_t ซึ่งเป็นดัชนีฤดูกาลจะอธิบายการเปลี่ยนแปลงของข้อมูลภายในฤดูกาล และ ε_t คือ ความรบกวนสุ่มที่เป็นอิสระต่อกัน แต่มีการแจกแจงความน่าจะเป็นแบบเดียวกัน โดยมีค่าคาดหมายเท่ากับศูนย์ และความแปรปรวนคงที่ σ^2 ถ้าฤดูกาลของข้อมูลมีความยาวเท่ากับ L หน่วยเวลา ผลบวกของดัชนีฤดูกาลทั้ง L ค่า จะกำหนดให้มีค่าเท่ากับ L

$$\sum_{j=1}^{L} C_j = L \tag{2.147}$$

เมื่อแนวโน้มเพิ่มขึ้นระดับของข้อมูลก็สูงขึ้นด้วย ยังผลให้การเปลี่ยนแปลงภายในฤดูกาลของตัวแบบเชิงคูณก็จะมีขนาดใหญ่ขึ้นตามไปด้วย และเมื่อแนวโน้มลดลง ทำให้ $b_0 + b_1 t$ มีค่าลดลง ตัวแบบเชิงคูณจึงมีขนาดการเปลี่ยนแปลงภายในฤดูกาลเล็กลงด้วย

Winters ได้แยกการพยากรณ์ออกเป็น 3 ส่วน คือ ส่วนที่หนึ่งเป็นระดับข้อมูล ซึ่งใช้แทนด้วยสัญลักษณ์ b_{0t} ระดับข้อมูลนี้เป็นส่วนของข้อมูลที่ไม่มีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง ค่าพยากรณ์ของระดับข้อมูลหนึ่งหน่วยเวลาล่วงหน้าที่พยากรณ์ ณ เวลา t-1 คือ $\hat{b}_{0(t-1)} + \hat{b}_{1(t-1)}$ โดย $\hat{b}_{0(t-1)}$ เป็นค่าพยากรณ์ของระดับข้อมูลและ $\hat{b}_{0(t-1)}$ เป็นค่าพยากรณ์ความลาดชัน ทั้ง 2 ค่าพยากรณ์นี้ พยากรณ์ ณ เวลาt-1 และเมื่อเวลาได้เคลื่อนมาอยู่ ณ เวลา t จะทราบค่าข้อมูล X_t แต่ข้อมูล X_t นี้ยังอยู่ภายใต้อิทธิพลของฤดูกาล จึงจำเป็นต้องขจัดอิทธิพลของฤดูกาลออกเพื่อทราบถึงระดับของข้อมูล โดยจะใช้ดัชนีฤดูกาลที่จุดเวลาเดียวกันในฤดูกาลที่แล้ว ดังนั้น ภายใต้แนวความคิดการทำให้เรียบแบบเอกซ์โปเนนเชียล ค่าพยากรณ์ของระดับข้อมูลที่พยากรณ์ ณ เวลา t จึงเขียนได้เป็น

$$\hat{b}_{0t} = \alpha \frac{X_t}{C_{t-L}} + (1-\alpha)\left[\hat{b}_{0(t-1)} + \hat{b}_{1(t-1)}\right] \tag{2.148}$$

โดย α เป็นค่าคงที่การทำให้เรียบ

ส่วนที่สองเป็นการพยากรณ์ความลาดชัน ซึ่งจะใช้วิธีการของ Holt คือ เป็นค่าเฉลี่ยแบบถ่วงน้ำหนักระหว่างผลต่างของระดับข้อมูลกับความลาดชันที่พยากรณ์ในหน่วยเวลาที่ผ่านมา

$$\hat{b}_{1t} = \beta\left(\hat{b}_{0t} - \hat{b}_{0(t-1)}\right) + (1-\beta)\hat{b}_{1(t-1)} \tag{2.149}$$

อีกนัยหนึ่งอาจกล่าวได้ว่าการพยากรณ์ความลาดชันใช้วิธีการทำให้เรียบแบบเอกซ์โปเนนเชียล โดยประมาณความลาดชันปัจจุบันด้วยผลต่างระหว่าง $\hat{b}_{0t}$ กับ $\hat{b}_{0(t-1)}$ ค่าคงที่การทำให้เรียบ β ในสมการ (2.149) ไม่ควรที่จะมีค่ามากนัก เพราะการเปลี่ยนแปลงความลาดชันของข้อมูลมักจะเป็นไปอย่างเชื่องช้า

ส่วนที่สาม คือ ดัชนีฤดูกาล ซึ่งการพยากรณ์ยังคงใช้วิธีการทำให้เรียบแบบเอกซ์โปเนนเชียล

$$\hat{C}_t = \gamma \frac{X_t}{\hat{b}_{0t}} + (1-\gamma)\hat{C}_{t-L} \tag{2.150}$$

อัตราส่วน $X_t / \hat{b}_{0t}$ เป็นสาระเกี่ยวกับดัชนีฤดูกาล ณ เวลา t และ γ เป็นค่าคงที่การทำให้เรียบ ซึ่งมักจะมีค่าน้อยกว่า β

ค่าพยากรณ์ τ หน่วยเวลาล่วงหน้าที่พยากรณ์ ณ เวลา t จึงอาจเขียนได้เป็น

$$\hat{X}_t(\tau) = \left(\hat{b}_{0t} + \hat{b}_{1t}\,\tau\right)\hat{C}_{t-L+\tau} \tag{2.151}$$

การคำนวณค่าพยากรณ์ระดับข้อมูลในสมการ (2.148) ความลาดชันในสมการ (2.149) และดัชนีฤดูกาล ในสมการ (2.150) จะต้องทราบค่าเริ่มต้น $\hat{b}_{00}$ และ $\hat{b}_{10}$ และดัชนีฤดูกาล ทั้ง L ค่า

สมมุติมีข้อมูลอยู่ mL หน่วยเวลา กล่าวคือ มีข้อมูลอยู่ m ฤดูกาล และฤดูกาลมีความยาว L หน่วยเวลา ถ้า $\overline{X}_i$ เป็นค่าเฉลี่ยของข้อมูลในฤดูกาลที่ i จะเห็นว่าในช่วงระยะเวลา $(m-1)L$ หน่วยเวลา การเปลี่ยนแปลงของค่าเฉลี่ยของฤดูกาล จะเท่ากับ $\overline{X}_m - \overline{X}_1$ Winters จึงได้แนะนำให้ใช้ ค่าเริ่มต้นของ $\hat{b}_{1t}$ เท่ากับ

$$\hat{b}_{10} = \frac{\overline{X}_m - \overline{X}_1}{(m-1)L} \tag{2.152}$$

โดยสมมติให้ค่าเฉลี่ยของข้อมูลในฤดูกาลหนึ่งเป็นค่าที่กึ่งกลางฤดูกาล ระดับข้อมูล ณ จุดเริ่มต้น $\hat{b}_{00}$ จึงเขียนได้เท่ากับ

$$\hat{b}_{00} = \overline{X}_1 - \frac{L+1}{2}\hat{b}_{10} \tag{2.153}$$

ค่าประมาณเบื้องต้นของดัชนีฤดูกาลที่หน่วยเวลา t, t = 1, 2, …, mL จะคำนวณจากอัตราส่วนของ X_t กับเส้นแนวโน้มที่ผ่านค่าเฉลี่ยของข้อมูลในฤดูกาลหนึ่งที่จุดกึ่งกลางฤดูกาล โดยมีความลาดชันเท่ากับ $\hat{b}_{10}$

$$\tilde{C}_{ij} = \frac{X_t}{\overline{X}_i - [(L+1)/2 - j]\hat{b}_{10}} \qquad t = 1, 2, \ldots, \mathrm{mL} \tag{2.154}$$

โดยความสัมพันธ์ระหว่าง t กับ i และ j เป็นดังนี้

$$t = (i-1)L + j \qquad i = 1, 2, \ldots, \mathrm{m} \tag{2.155}$$
$$j = 1, 2, \ldots, \mathrm{L}$$

ค่าเฉลี่ยของค่าประมาณของดัชนีฤดูกาลที่หน่วยเวลา j นับจากต้นฤดูกาล อาจเขียนได้เท่ากับ

$$\overline{C}_j = \frac{1}{m}\sum_{i=1}^{m}\tilde{C}_{(i-1)L+j} \qquad j = 1, 2, \ldots, \mathrm{L} \tag{2.156}$$

ซึ่งเมื่อปรับให้ผลบวกของดัชนีฤดูกาลทั้ง L ค่ามีค่าเท่ากับ L ค่าดัชนีฤดูกาลที่จะใช้ในการพยากรณ์จึงเท่ากับ

$$\hat{C}_j = \overline{C}_j \frac{L}{\sum_{j=1}^{L}\overline{C}_j} \qquad j = 1, 2, \ldots, \mathrm{L} \tag{2.157}$$

2. ตัวแบบเชิงบวก (Additive Model)

ตัวแบบเชิงบวกของ Winters อาจเขียนได้เป็น

$$X_t = b_0 + b_1 t + C_t + \varepsilon_t \tag{2.158}$$

จะเห็นได้ว่าการเปลี่ยนแปลงภายในฤดูกาลในตัวแบบเชิงบวกมีขนาดคงที่ มิได้พึ่งพิงต่อแนวโน้มของข้อมูล ซึ่งแตกต่างไปจากตัวแบบเชิงคูณที่ได้กล่าวมาแล้ว ในกรณีนี้จะกำหนดให้ผลบวกของดัชนีฤดูกาลทั้ง L ค่าในหนึ่งฤดูกาล มีค่าเท่ากับศูนย์

$$\sum_{i=1}^{L} C_{ij} = 0 \qquad j = 1, 2, \ldots, m \tag{2.159}$$

โดย C_{ij} เป็นค่าดัชนีฤดูกาลในฤดูที่ i และหน่วยเวลาที่ j ในฤดูกาลนั้น ซึ่งอาจเขียนเป็นค่าดัชนีฤดูกาลที่หน่วยเวลา t ได้ดังนี้

$$C_t = C_{ij}$$

โดยที่ $t = (i-1)L + j \qquad i = 1, 2, \ldots, m$

$$j = 1, 2, \ldots, L$$

การพยากรณ์ยังคงแยกออกเป็นสามส่วนเช่นเดียวกันกับกรณีตัวแบบเชิงคูณ โดยการพยากรณ์ระดับของข้อมูลจะเขียนได้เท่ากับ

$$\hat{b}_{0t} = \alpha\left[X_t - \hat{C}_{t-L}\right] + (1-\alpha)\left[\hat{b}_{0(t-1)} + \hat{b}_{1(t-1)}\right] \tag{2.160}$$

ความลาดชันที่พยากรณ์ ณ จุดเวลา t ยังคงเหมือนกับกรณีตัวแบบเชิงคูณ

$$\hat{b}_{1t} = \beta\left(\hat{a}_t - \hat{a}_{t-1}\right) + (1-\beta)\hat{b}_{1(t-1)} \tag{2.161}$$

ส่วนการพยากรณ์ดัชนีฤดูกาล ณ เวลา t ในกรณีนี้จะเขียนได้เป็น

$$\hat{C}_t = \gamma\left(X_t - \hat{b}_{0t}\right) + (1-\gamma)\hat{C}_{t-L} \tag{2.162}$$

ค่าพยากรณ์ τ หน่วยเวลาล่วงหน้าที่พยากรณ์ ณ เวลา t จึงเขียนได้เท่ากับ

$$\hat{X}_t(\tau) = \hat{b}_{0t} + \hat{b}_{1t}\,\tau + \hat{C}_{t-L+\tau} \tag{2.163}$$

เช่นเดียวกันกับกรณีตัวแบบเชิงคูณ การคำนวณค่าพยากรณ์ $\hat{b}_{0t}$ ในสมการ (2.160) ในสมการ $\hat{b}_{1t}$ (2.161) และ $\hat{C}_t$ ในสมการ (2.162) จะต้องทราบค่าเริ่มต้น ซึ่งในกรณีนี้จะประมาณค่าโดยวิธีกำลังสองน้อยที่สุด (least squares method)

สมมติมีข้อมูลอยู่ m ฤดูกาล และความยาวของฤดูกาลเท่ากับ L หน่วยเวลา ดัชนีฤดูกาลที่จะใช้ในการพยากรณ์ตามสมการ (2.162) นั้น จะประมาณค่าจากข้อมูลทั้ง mL จำนวน โดยกำหนดให้ดัชนีฤดูกาลที่หน่วยเวลาเดียวกัน เมื่อนับจากต้นฤดูกาลมีค่าเท่ากัน

$$C_j = C_{L+j} = \ldots = C_{(m-1)L+j} \qquad j = 1, 2, \ldots, L \tag{2.164}$$

ดังนั้น ข้อมูล X_t ในสมการ (2.158) จึงกลายเป็น

$$X_t = b_0 + b_1 t + C_{(i-1)L+j} + \varepsilon_t \qquad t = 1, 2, \ldots, \mathrm{mL} \tag{2.165}$$

โดยความสัมพันธ์ระหว่าง t กับ i และ j ยังคงเป็นไปตามสมการ (2.155)

แนวทางการประมาณค่าเริ่มต้นก็คือ การกำหนดค่า $\hat{b}_0, \hat{b}_1$ และ $\hat{C}_j,\ j = 1,2,\ldots,L$ เพื่อทำให้ผลบวกกำลังสองของความคลาดเคลื่อนทั้ง mL ค่า มีค่าต่ำสุด

$$\begin{aligned} \mathrm{E} &= \sum_{t=1}^{mL}\left(X_t - \hat{b}_0 - \hat{b}_1 t - \hat{C}_{(i-1)L+j}\right)^2 \\ &= \sum_{t=1}^{mL}\left(X_t - \hat{b}_0 - \hat{b}_1 t - \hat{C}_t\right)^2 \end{aligned} \tag{2.166}$$

ซึ่งเงื่อนไขที่ค่าต่ำสุดของ E จะเกิดขึ้นอาจจะเขียนได้เป็น

$$mL\hat{b}_0 - b_1\sum_{t=1}^{mL} t + m\sum_{j=1}^{mL}\hat{C}_j = \sum_{t=1}^{mL} X_t \tag{2.167}$$

$$\hat{b}_0\sum_{t=1}^{mL} t + \hat{b}_1\sum_{t=1}^{mL} t^2 + \sum_{t=1}^{mL}\hat{C}_j = \sum_{t=1}^{mL} tX_t^{\ 2} \tag{2.168}$$

$$m\hat{b}_0 - b_1\sum_{t=1}^{mL}\left[(i-1)L + j\right] + m\hat{C}_j = \sum_{i=1}^{m} X_{(i-1)L+j} \qquad j = 1, 2, \ldots, \mathrm{L} \tag{2.169}$$

จากสมการ (2.169) $\hat{C}_j$ เขียนได้เป็น

$$\hat{C}_j = \frac{1}{m}\sum_{i=1}^{m} X_{(i-1)L+j} - \hat{b}_0 - \hat{b}_1\left\{j + (m-1)L/2\right\} \qquad j = 1, 2, \ldots, \mathrm{L} \tag{2.170}$$

เนื่องจากผลบวกของดัชนีฤดูกาลทั้ง L ค่าในหนึ่งฤดูกาล มีค่าเท่ากับศูนย์ ดังในสมการ (2.159) สมการ (2.167) จึงกลายเป็น

$$mL\hat{b}_0 + \frac{mL(mL+1)}{2}\hat{b}_1 = \sum_{t=1}^{mL} X_t \tag{2.171}$$

และสมการ (2.168) จะเหลือเพียง

$$\frac{mL(mL+1)}{2}\hat{b}_0 + \frac{mL(mL+1)(2mL+1)}{6}\hat{b}_1 + \sum_{j=1}^{L} mj\hat{C}_j = \sum_{t=1}^{mL} tX_t \tag{2.172}$$

พจน์ $\sum_{j=1}^{L} mj\hat{C}_t$ อาจเขียนได้จากสมการ (2.170) มีค่าเท่ากับ

$$\sum_{j=1}^{L} mj\hat{C}_t = \sum_{j=1}^{L}\sum_{j=1}^{m} jX_{(i-1)L+j} - \frac{mL(L+1)}{2}\hat{b}_0 - \frac{mL(L+1)(3mL+L+2)}{12}\hat{b}_1 \tag{2.173}$$

ซึ่งเมื่อนำไปแทนค่าในสมการ (2.172) แล้ว จะได้ผลลัพธ์

$$\frac{m(m-1)L^2}{2}\hat{b}_0 + \frac{m(m-1)L^2[(4m+1)L+3]}{12}\hat{b}_1 = \sum_{t=1}^{mL} tX_t - \sum_{j=1}^{L}\sum_{j=1}^{m} jX_{(i-1)L+j} \tag{2.174}$$

เมื่อพิจารณาสมการ (2.174) โดยถ่องแท้แล้ว จะพบว่าพจน์สุดท้ายขวามืออาจเขียนใหม่ในรูปของ X_{ij} ซึ่งเป็นค่า X_t ในฤดูกาลที่ i ณ หน่วยเวลา j ในฤดูกาลนั้น

$$\frac{m(m-1)L^2}{2}\hat{b}_0 + \frac{m(m-1)L^2[(4m+1)L+3]}{12}\hat{b}_1 = \sum_{t=1}^{mL} tX_t - \sum_{i=1}^{L}\sum_{j=1}^{m} jX_{ij} \tag{2.175}$$

หากหารทั้งสองข้างของสมการ (2.171) ด้วย mL แล้ว สมการ (2.171) จะกลายเป็น

$$\hat{b}_0 = \overline{X}_t - \frac{mL+1}{2}\hat{b}_1 \tag{2.176}$$

โดยที่ $\overline{X}_t$ เป็นค่าเฉลี่ย X_t ทั้ง mL ค่า เมื่อนำค่า $\hat{b}_0$ ในสมการ (2.176) ไปแทนในสมการ (2.174) แล้ว จะพบว่า $\hat{b}_1$ มีค่าเท่ากับ

$$\hat{b}_1 = \frac{12}{m(m^2-1)L^3}\left[\sum_{t=1}^{mL} tX_t - \sum_{i=1}^{m}\sum_{j=1}^{L} jX_{ij}\right] - \frac{6\overline{X}_t}{(m+1)L} \tag{2.177}$$

เมื่อนำค่า $\hat{b}_1$ ในสมการ (2.177) กลับไปแทนค่าในสมการ (2.176) $\hat{b}_0$ มีค่าเท่ากับ

$$\hat{b}_0 = \frac{(4mL+L+3)\overline{X}_t}{(m+1)L} - \frac{6(mL+1)}{m(m^2-1)L^3}\left[\sum_{t=1}^{mL} tX_t - \sum_{i=1}^{m}\sum_{j=1}^{L} jX_{ij}\right] \tag{2.178}$$

ดังนั้น เมื่อทราบค่า $\hat{b}_0$ และ $\hat{b}_1$ จากสมการ (2.178) และ (2.177) ตามลำดับแล้ว ดัชนีฤดูกาลเชิงบวก $\hat{C}_j : j = 1,2,...,L$ ก็สามารถคำนวณได้จากสมการ (2.170) ซึ่งจะกำหนดให้เป็นค่าดัชนีฤดูกาลที่เพิ่งผ่านไป เพื่อใช้ในการพยากรณ์ในสมการ (2.163) ค่า $\hat{b}_0$ และ $\hat{b}_1$ ที่คำนวณได้จะกำหนดเป็นค่าเริ่มต้นของระดับข้อมูล $\hat{b}_{00}$ และค่าเริ่มต้นของความลาดชัน $\hat{b}_{10}$ เพื่อใช้ในการพยากรณ์ค่า $\hat{b}_{0t}$ และ $\hat{b}_{1t}$ ตามสมการ (2.160) และ (2.161) ตามลำดับ

อนึ่ง อาจพิสูจน์ได้ว่า ผลบวกของค่าดัชนีฤดูกาลที่คำนวณจากสมการ (2.170) ทั้ง L มีค่าเท่ากับศูนย์ โดยใช้ผลลัพธ์ของสมการ (2.171) จึงไม่จำเป็นต้องปรับค่าดัชนีดังเช่นในกรณีตัวแบบเชิงคูณ

2.4 การพยากรณ์แบบปรับได้

ในตัวแบบต่าง ๆ ภายใต้แนวความคิดการทำให้เรียบ จะเห็นว่าพารามิเตอร์ของตัวแบบการพยากรณ์เมื่อได้กำหนดค่าแล้ว จะมีค่าคงที่มิอาจปรับเปลี่ยนค่าได้ จนกว่าจะมีการพิจารณาค่าพารามิเตอร์ใหม่ ซึ่งจะกระทำเป็นครั้งคราวโดยอาจใช้ความคลาดเคลื่อนในการพยากรณ์ เป็นเกณฑ์ที่จะให้มีการทบทวนค่าพารามิเตอร์ของตัวแบบการพยากรณ์ การทบทวนค่าพารามิเตอร์ของตัวแบบในลักษณะนี้ หมายความว่าจะต้องมีความคลาดเคลื่อนที่มีนัยสำคัญเกิดขึ้นเป็นช่วงระยะเวลาหนึ่ง หรือมีความคลาดเคลื่อนที่มีนัยสำคัญมากเกิดขึ้น จึงจะมีการทบทวนพิจารณาค่าพารามิเตอร์ของตัวแบบ ความยืดหยุ่นของตัวแบบการพยากรณ์ที่เกี่ยวกับพารามิเตอร์จึงไม่มีจนกว่าจะมีการเปลี่ยนแปลงค่าโดยผู้ที่กระทำการพยากรณ์ ในกรณีที่ผู้ทำการพยากรณ์มีความรับผิดชอบในตัวแบบการพยากรณ์ หลาย ๆ ตัวแบบ ภาระของผู้ที่กระทำการพยากรณ์ในการที่จะต้องติดตามคอยทบทวนค่าพารามิเตอร์ของตัวแบบจะมีมาก อาจทำให้การทบทวนไม่ทันการ ดังนั้น จึงมีแนวความคิดที่จะสร้างตัวแบบการพยากรณ์ที่มีความสามารถปรับค่าพารามิเตอร์เองได้ (adaptive forecasting model) ซึ่งในที่นี้กล่าวถึงการปรับค่าคงที่การทำให้เรียบโดยวิธีการของ Chow และวิธีการของ Trigg & Leach และการพยากรณ์แบบการกรองที่ปรับได้ (adaptive filtering) นอกจากนี้ ยังได้เสนอวิธีค่าเฉลี่ยเคลื่อนที่ชั้นเดียวแบบปรับจำนวนพจน์ (adaptive moving average) และวิธีการพยากรณ์ที่ปรับปรุงจากวิธีการของ Chow ซึ่งทั้ง 2 วิธีเป็นวิธีการพยากรณ์ใหม่ที่ได้พัฒนาขึ้น

2.4.1 การปรับค่าคงที่การทำให้เรียบ

เมื่อกล่าวถึงตัวแบบค่าเฉลี่ยคงที่ระยะสั้น (local constant mean model)

$$X_t = \mu_t + \varepsilon_t \qquad (2.179)$$

โดยค่าเฉลี่ย μ_t อาจมีการเปลี่ยนแปลงค่า แต่การเปลี่ยนแปลงเป็นไปอย่างเชื่องช้า และ ε_t คือ ความรบกวนสุ่มซึ่งเป็นอิสระต่อกัน และมีค่าคาดหมายเท่ากับศูนย์ ภายใต้ตัวแบบนี้ระเบียบวิธีที่นิยมใช้กันอย่างแพร่หลายวิธีหนึ่ง คือ เทคนิคการทำให้เรียบแบบเอกซ์โปเนนเชียล ซึ่งมีค่าพยากรณ์ดังนี้

$$\hat{X}_t(1) = \alpha X_t + (1-\alpha)\hat{X}_{t-1}(1) \tag{2.180}$$

โดย α เป็นค่าคงที่การทำให้เรียบ (smoothing constant) ซึ่งมีค่าอยู่ระหว่าง 0 กับ 1 จะเห็นได้ว่าการกำหนดค่าคงที่การทำให้เรียบ มีอิทธิพลสูงต่อความถูกต้องของค่าพยากรณ์ เมื่อกำหนดค่าคงที่การทำให้เรียบมีค่าน้อย คือ ใกล้ศูนย์ ค่าพยากรณ์จะพึ่งพิงข้อมูลในอดีตมาก ทำให้มีความเฉื่อยสูง จึงเหมาะสมกับข้อมูลที่ไม่มีการเปลี่ยนแปลงอย่างรวดเร็ว และเมื่อกำหนดค่าคงที่การทำให้เรียบมีค่ามาก คือ ใกล้หนึ่ง ค่าพยากรณ์จะมีการปรับค่าอย่างรวดเร็วตามข้อมูลที่เพิ่งเกิดขึ้น แต่จะต้องระวังไม่ให้ค่าพยากรณ์ปรับค่าไปตามความรบกวนสุ่ม ε_t มิฉะนั้นจะทำให้ค่าพยากรณ์ขาดเสถียรภาพในทางปฏิบัติ การกำหนดค่าคงที่การทำให้เรียบ มักจะใช้ข้อมูลในอดีตมาเป็นแนวทางในการกำหนดค่า และเมื่อได้กำหนดค่าแล้ว ค่าคงที่การทำให้เรียบจะมีค่าคงที่ และจะใช้ค่านี้ในการพยากรณ์ไประยะหนึ่ง แล้วจึงพิจารณาค่าคงที่การทำให้เรียบที่ใช้อยู่ว่ายังคงมีความเหมาะสมหรือไม่

Chow และ Trigg & Leach มีแนวความคิดที่คล้ายกันคือต้องการให้ค่าคงที่การทำให้เรียบปรับค่าตนเองไปตามการเปลี่ยนแปลงในข้อมูล แต่การปรับค่าคงที่การทำให้เรียบกระทำภายใต้หลักการที่แตกต่างกัน

1. การปรับค่าคงที่การทำให้เรียบตามวิธีของ Chow

Chow ได้แนะนำให้คำนวณค่าพยากรณ์การทำให้เรียบแบบเอกซ์โปเนนเชียล ตามสมการ (2.180) โดยใช้ค่าคงที่การทำให้เรียบ 3 ค่า คือ α_0 เป็นค่าคงที่การทำให้เรียบกรณีฐาน (nominal value) α_u เป็นค่าคงที่การทำให้เรียบกรณีสูง (upper value) ซึ่งมีค่าเท่ากับ

$$\alpha_u = \alpha_0 + \delta \tag{2.181}$$

และ α_ℓ เป็นค่าคงที่การทำให้เรียบกรณีต่ำ (lower value) ซึ่งมีค่าเท่ากับ

$$\alpha_\ell = \alpha_0 - \delta \tag{2.182}$$

โดยกำหนดให้ค่า δ เท่ากับ 0.05 ดังนั้น วิธีการของ Chow จึงมีค่าพยากรณ์อยู่ 3 ค่าตามค่าคงที่การทำให้เรียบ และค่าพยากรณ์ที่คำนวณจากค่าคงที่การทำให้เรียบกรณีฐานจะเป็นค่าพยากรณ์ที่นำไปใช้งาน ในกรณีนี้จะมีความคลาดเคลื่อนสัมบูรณ์ 3 ค่า คือ

$$\left|e_{t+1}(\alpha_0)\right| = \left|X_{t+1} - \hat{X}_t(\alpha_0,1)\right| \tag{2.183}$$

$$\left|e_{t+1}(\alpha_u)\right| = \left|X_{t+1} - \hat{X}_t(\alpha_u,1)\right| \tag{2.184}$$

$$\left|e_{t+1}(\alpha_\ell)\right| = \left|X_{t+1} - \hat{X}_t(\alpha_\ell,1)\right| \tag{2.185}$$

เพื่อให้ค่าพยากรณ์มีเสถียรภาพพอสมควร การปรับค่าคงที่การทำให้เรียบจะกระทำบนพื้นฐานของค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ ซึ่งในที่นี้จะใช้ความคลาดเคลื่อนสัมบูรณ์ที่ทำให้เรียบแล้ว (smoothed absolute error)

$$\Delta_t = \gamma\left|e_t\right| + (1-\gamma)\Delta_{t-1} \tag{2.186}$$

โดย γ เป็นค่าคงที่การทำให้เรียบสำหรับความคลาดเคลื่อนสัมบูรณ์ มีค่าระหว่างศูนย์กับหนึ่ง ความคลาดเคลื่อนสัมบูรณ์ที่ทำให้เรียบแล้ว จะมี 3 ค่าตามความคลาดเคลื่อนที่เกิดขึ้น จากค่าคงที่การปรับให้เรียบ α_0, α_u และ α_ℓ ความคลาดเคลื่อนสัมบูรณ์ที่ทำให้เรียบตามคำนิยามในสมการ (2.186) อาจเขียนใหม่ในรูปของความคลาดเคลื่อนสัมบูรณ์ในอดีตได้เป็น

$$\Delta_t = \gamma\sum_{j=0}^{t-1}(1-\gamma)^j\left|e_{t-j}\right| + (1-\gamma)^t\left|e_0\right| \tag{2.187}$$

ซึ่งจะเห็นว่า ความคลาดเคลื่อนสัมบูรณ์ที่ทำให้เรียบแล้ว เป็นค่าเฉลี่ยถ่วงน้ำหนักของความคลาดเคลื่อนสัมบูรณ์ในอดีต ค่าคงที่การทำให้เรียบ γ ไม่ควรกำหนดให้มีค่าสูงมากนักเพื่อให้ Δ_t สะท้อนความคลาดเคลื่อนที่เกิดขึ้นในอดีตย้อนหลังพอสมควร

ค่าคงที่การทำให้เรียบจะปรับค่าไปยังทิศทางที่ความคลาดเคลื่อนสัมบูรณ์ที่ทำให้เรียบแล้วมีค่าน้อยที่สุดตามหลักการดังนี้

1. ถ้า $\Delta_t(\alpha_0)$ มีค่าไม่มากกว่า $\Delta_t(\alpha_u)$ และ $\Delta_t(\alpha_\ell)$ ให้ยังคงใช้ค่าคงที่การทำให้เรียบปัจจุบันในการพยากรณ์ครั้งต่อไป

2. ถ้า $\Delta_t(\alpha_0)$ มีค่ามากกว่า $\Delta_t(\alpha_u)$ แต่ไม่มากกว่า $\Delta_t(\alpha_\ell)$ ให้ปรับค่าคงที่ทำให้เรียบในการพยากรณ์ครั้งต่อไป ดังนี้

$$\alpha_0 = \alpha_u \tag{2.188}$$

$$\alpha_u = \alpha_0 + \delta \tag{2.189}$$

$$\alpha_\ell = \alpha_0 - \delta \tag{2.190}$$

3. ถ้า $\Delta_t(\alpha_0)$ มีค่ามากกว่า $\Delta_t(\alpha_\ell)$ แต่ไม่มากกว่า $\Delta_t(\alpha_u)$ ค่าคงที่ทำให้เรียบในการพยากรณ์ครั้งต่อไป ดังนี้

$$\alpha_0 = \alpha_\ell \tag{2.191}$$

$$\alpha_u = \alpha_0 + \delta \tag{2.192}$$

$$\alpha_\ell = \alpha_0 - \delta \tag{2.193}$$

4. ถ้า $\Delta_t(\alpha_0)$ มีค่ามากกว่าทั้ง $\Delta_t(\alpha_u)$ และ $\Delta_t(\alpha_\ell)$ ให้พิจารณาดังนี้ ถ้า $\Delta_t(\alpha_u)$ มีค่าน้อยกว่า $\Delta_t(\alpha_\ell)$ ให้ α_0, α_u และ α_ℓ เป็นไปตามสมการ (2.188), (2.189), และ (2.190) ตามลำดับ มิฉะนั้นให้ α_0, α_u และ α_ℓ เป็นไปตามสมการ (2.191), (2.192), และ (2.193) ตามลำดับ

จากประสบการณ์ที่นำวิธีการพยากรณ์ของ Chow ไปใช้ในการพยากรณ์ พบว่า การกำหนดค่าสูงสุดของค่าคงที่การทำให้เรียบ $\alpha_{\max}$ และค่าต่ำสุดของค่าคงที่การทำให้เรียบ $\alpha_{\min}$ ไว้เพื่อป้องกันมิให้มีการปรับค่าคงที่การทำให้เรียบจนกระทั่ง α_u มีค่าเท่ากับหนึ่ง หรือ α_ℓ มีค่าเท่ากับศูนย์ จะทำให้ค่า MSE ของค่าพยากรณ์ลดลง หาก α_u มีค่าเท่ากับ $\alpha_{\max}$ แล้ว และการปรับค่าคงที่การทำให้เรียบ ยังคงเรียกร้องให้ปรับตามสมการ (2.188), (2.189), และ (2.190) อีก ซึ่งจะทำให้ α_u มีค่าเกินค่า $\alpha_{\max}$ จะไม่มีการปรับค่าคงที่การทำให้เรียบ ให้สูงขึ้นอีก การพยากรณ์จะคำนวณจากค่าเดิมทั้ง 3 ค่า ของค่าคงที่การทำให้เรียบ และในทำนองเดียวกัน หาก α_ℓ มีค่าเท่ากับ $\alpha_{\min}$ แล้ว และการปรับค่าคงที่การทำให้เรียบ ยังต้องกระทำตามสมการ (2.191), (2.192) และ (2.193) อีกก็จะไม่มีการปรับค่าคงที่การทำให้เรียบให้ลดลงอีก และการพยากรณ์จะคำนวณจากค่าเดิมทั้ง 3 ค่า ของค่าคงที่การทำให้เรียบเช่นกัน การกำหนดค่าเหมาะสมให้แก่ $\alpha_{\max}$ และ $\alpha_{\min}$ อาจเป็นหนทางหนึ่งที่ลดความเคลื่อนในค่าพยากรณ์ได้อีกด้วย

2. การทำให้เรียบแบบเอกซ์โปเนนเชียลด้วยอัตราการตอบสนองที่ปรับได้ (adaptive response rate)

การปรับค่าคงที่การทำให้เรียบโดยวิธีการของ Trigg & Leach นั้น จะกระทำในหลักการที่พยายามเพิ่มค่าให้แก่ค่าคงที่การทำให้เรียบ เมื่อมีความคลาดเคลื่อนสูง ทั้งนี้เพื่อให้ค่าพยากรณ์มีโอกาสในการปรับตนเองได้เร็วขึ้นโดยพิจารณาจากความคลาดเคลื่อนที่ทำให้เรียบแล้ว (smoothed error) Q_t ซึ่งนิยามให้มีค่าเท่ากับ

$$Q_t = \gamma e_t + (1-\gamma)Q_{t-1} \tag{2.194}$$

โดย e_t เป็นค่าคลาดเคลื่อนของค่าพยากรณ์ ณ เวลา t และ γ เป็นค่าคงที่การทำให้เรียบ ซึ่งมีค่าอยู่ระหว่าง 0 กับ 1 สมการ (2.194) อาจเขียนใหม่ในรูป

$$Q_t = \gamma \sum_{j=0}^{t-1} (1-\gamma)^j e_{t-j} + (1-\gamma)^t e_0 \qquad (2.195)$$

ซึ่งจะเห็นได้ว่าค่า Q_t เป็นค่าเฉลี่ยถ่วงน้ำหนักของความคลาดเคลื่อน ดังนั้น ถ้าค่าพยากรณ์ไม่เอียงเฉ (unbias) Q_t จะมีค่าเข้าใกล้ศูนย์ และเพื่อให้ Q_t สะท้อนความคลาดเคลื่อนในอดีตที่ผ่านมาค่าคงที่การทำให้เรียบ γ ไม่ควรกำหนดให้มีค่าสูงมากนัก

Trigg & Leach ได้แนะนำให้ปรับค่าคงที่การทำให้เรียบที่ใช้ในการพยากรณ์บนพื้นฐานของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว (smoothed error tracking signal) ซึ่งนิยามให้เท่ากับ

$$Q_t / \Delta_t \qquad (2.196)$$

โดย Δ_t เป็นความคลาดเคลื่อนสัมบูรณ์ที่ทำให้เรียบแล้ว (smoothed absolute error) ตามที่นิยามไว้ในสมการ (2.196) จะเห็นได้ว่า Q_t / Δ_t มีค่าอยู่ระหว่าง -1 กับ 1 ทั้งนี้เพราะค่าสัมบูรณ์ของ Q_t จะมีค่าเกิน Δ_t ไม่ได้ ถ้าค่าพยากรณ์มีค่าสูงกว่าค่าจริงเป็นระยะเวลานานอย่างต่อเนื่อง หรือมีค่าต่ำกว่าค่าจริงเป็นระยะเวลาหนึ่งอย่างต่อเนื่อง จะมีผลให้ Q_t มีค่าสัมบูรณ์เข้าใกล้ Δ_t ซึ่งจะทำให้ $|Q_t / \Delta_t|$ มีค่าสูงขึ้น สภาพที่ค่าพยากรณ์สูงกว่าหรือต่ำกว่าค่าจริงเป็นระยะเวลานานอย่างต่อเนื่องไม่ใช่สภาพของค่าพยากรณ์ที่ดี ในกรณีนี้ค่าคงที่การทำให้เรียบควรปรับปรุงให้มีค่าสูงขึ้น เพื่อให้ค่าพยากรณ์ตามค่าจริงได้รวดเร็วมากขึ้น ค่าพยากรณ์ที่ดีควรมีค่าสูงกว่าค่าจริงสลับอย่างสุ่มกับมีค่าต่ำกว่าค่าจริง ในสภาพเช่นนี้ Q_t จะมีค่าสัมบูรณ์น้อย และจะถอยห่างออกจาก Δ_t จึงทำให้ $|Q_t / \Delta_t|$ มีค่าลดลง ในกรณีนี้ค่าคงที่การทำให้เรียบควรมีค่าลดลงเพื่อให้ค่าพยากรณ์มีเสถียรภาพสูงขึ้น ดังนั้น Trigg & Leach จึงแนะนำให้กำหนดค่าคงที่การทำให้เรียบในการพยากรณ์ ณ จุดเวลา t มีค่าเท่ากับ

$$\alpha_t = |Q_t / \Delta_t| \qquad (2.197)$$

ซึ่งจะทำให้ค่าคงที่การทำให้เรียบมีการปรับค่าตนเองตามการเปลี่ยนแปลงในข้อมูล การปรับค่าคงที่การทำให้เรียบในลักษณะนี้ จึงเป็นที่รู้จักกันในนามการปรับแบบอัตราการตอบสนองที่ปรับได้ (adaptive response rate) เพื่อป้องกันมิให้ค่าพยากรณ์ปรับตนเองตามข้อมูลที่เพิ่งเกิดขึ้นเร็วเกินควร จึงมีการกำหนดค่าสูงสุดของค่าคงที่การทำให้เรียบไว้เท่ากับ $\alpha_{\max}$ และให้ค่าคงที่การทำให้เรียบปรับเปลี่ยนค่าดังนี้

$$\alpha_t = \alpha_{\max} \left| Q_t / \Delta_t \right| \tag{2.198}$$

อย่างไรก็ตาม การปรับค่าคงที่การทำให้เรียบตามสมการ (2.198) นั้น จะไม่ปรับเต็มที่ตามสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว แม้ในกรณีที่สัญญาณมีค่าต่ำกว่า $\alpha_{\max}$ จึงอาจเป็นการเพิ่มความเฉื่อยในค่าพยากรณ์เกินความจำเป็น และยิ่งไปกว่านั้น ในกรณีที่สัญญาณมีค่าต่ำอยู่แล้ว ค่าคงที่การทำให้เรียบที่ปรับตามสมการ (2.198) จะต้องมีค่าต่ำกว่าค่าสัญญาณอีก ดังนั้นเพื่อแก้ไขจุดอ่อนในการปรับค่าคงที่การทำให้เรียบตามสมการ (2.198) ค่าคงที่การทำให้เรียบจะปรับค่าโดยกำหนดค่า $\alpha_{\max}$ และ $\alpha_{\min}$ ไว้ในทำนองเดียวกันกับที่ได้อธิบายไว้ในกรณีของ Chow ดังนี้

1. ถ้า $\left| Q_t / \Delta_t \right|$ มีค่าอยู่ระหว่างค่าต่ำสุด $\alpha_{\min}$ และค่าสูงสุด $\alpha_{\max}$ ให้ปรับค่า α_t ตามสมการ (3.19)

2. ถ้า $\left| Q_t / \Delta_t \right|$ มีค่าน้อยกว่า $\alpha_{\min}$ ให้ $\alpha_t = \alpha_{\min}$

3. ถ้า $\left| Q_t / \Delta_t \right|$ มีค่ามากกว่า $\alpha_{\max}$ ให้ $\alpha_t = \alpha_{\max}$

และหากกำหนดค่าเหมาะสมให้แก่ $\alpha_{\max}$ และ $\alpha_{\min}$ อาจพบว่าความคลาดเคลื่อนในค่าพยากรณ์จะมีค่าลดลงได้

3. การปรับค่าคงที่การทำให้เรียบในกรณีที่มีมากกว่า 1 ตัว

การปรับค่าคงที่การทำให้เรียบตามวิธีการของ Chow และตามวิธีของ Trigg & Leach เป็นการปรับค่าคงที่การทำให้เรียบในกรณีที่ตัวแบบการพยากรณ์มีพารามิเตอร์เพียงตัวเดียว แต่ตัวแบบการพยากรณ์อาจมีพารามิเตอร์มากกกว่าหนึ่งตัวก็เป็นได้ เช่น ตัวแบบการพยากรณ์ของ Holt มีค่าคงที่การทำให้เรียบตัวหนึ่ง สำหรับระดับของข้อมูล และอีกตัวสำหรับความลาดชันของข้อมูล ตัวแบบการพยากรณ์ของ Winters มีพารามิเตอร์สามตัวพารามิเตอร์สองตัวจะใช้สำหรับระดับและความลาดชันของข้อมูลเหมือนกับกรณีตัวแบบการพยากรณ์ของ Holt ส่วนค่าคงที่การทำให้เรียบตัวที่สามใช้สำหรับฤดูกาลของข้อมูล เป็นต้น การที่จะปรับเปลี่ยนค่าคงที่การทำให้เรียบที่ใช้สำหรับความลาดชันนั้น จะต้องกระทำด้วยความระมัดระวังอย่างยิ่ง เพราะการเปลี่ยนแปลงความลาดชันรวดเร็วเกินไปอาจทำให้ค่าพยากรณ์ไม่มีเสถียรภาพ ดังนั้นในทางปฏิบัติอาจจะมีความเพียงพอแล้วที่ปรับค่าคงที่การทำให้เรียบสำหรับระดับของข้อมูลเพียงตัวเดียวตามวิธีการของ Chow หรือวิธีการของ Trigg & Leach ส่วนค่าคงที่การทำให้เรียบสำหรับความลาดชันหรือฤดูกาลนั้น อาจไม่จำเป็นต้องมีการปรับค่าด้วยตนเอง

อย่างไรก็ตาม ในกรณีที่ประสงค์จะปรับค่าคงที่การทำให้เรียบหลายค่าก็อาจใช้ เทคนิคการพยากรณ์ที่ปรับตนเองได้ (Self – Adaptive Forecasting Technique - SAFT) ซึ่ง Robert & Reed เป็นผู้พัฒนาขึ้นโดยขยายแนวความคิดของ Chow มาประยุกต์ใช้ในกรณีที่ตัวแบบการพยากรณ์มีพารามิเตอร์หลายตัว

ระเบียบวิธีของ Robert & Reed จะกำหนดให้ค่าคงที่การทำให้เรียบแต่ละตัวมี 3 ค่า คือ ค่าฐาน (nominal) ค่าสูง (upper value) และค่าต่ำ (lower value) ในลักษณะเดียวกันกับวิธีของ Chow ดังนั้น เมื่อมีพารามิเตอร์ k ตัว ค่าพยากรณ์ที่จะต้องพิจารณาจะมีอยู่ทั้งหมด 2^k+1 ค่าโดยเป็นค่าพยากรณ์ที่คำนวณจากค่าสูงและ/หรือค่าต่ำของพารามิเตอร์ 2^k ค่าและค่าพยากรณ์อีกหนึ่งค่าจะเป็นค่าพยากรณ์ที่คำนวณจากค่าฐานของพารามิเตอร์ เพื่อให้เห็นการปรับค่าคงที่การทำให้เรียบในลักษณะที่เป็นรูปธรรมยิ่งขึ้น ให้พิจารณากรณีมีพารามิเตอร์ 2 ตัว เช่น ในระเบียบวิธีการที่พยากรณ์ของ Holt ซึ่งมีค่าคงที่การทำให้เรียบ α สำหรับการพยากรณ์ระดับ b_{0t} และค่าคงที่การทำให้เรียบ β สำหรับการพยากรณ์ความลาดชัน b_{1t} ในการพยากรณ์ครั้งหนึ่ง ๆ จะมีค่าพยากรณ์ที่จะต้องพิจารณาอยู่ 5 ค่า ดังนี้คือ

1. $\hat{X}_t(\tau,\alpha_0,\beta_0)$ = ค่าพยากรณ์ที่ 1 คำนวณจากค่าฐานของพารามิเตอร์
2. $\hat{X}_t(\tau,\alpha_\ell,\beta_\ell)$ = ค่าพยากรณ์ที่ 2 คำนวณจากค่าต่ำของพารามิเตอร์ทั้ง 2 ค่า
3. $\hat{X}_t(\tau,\alpha_\ell,\beta_u)$ = ค่าพยากรณ์ที่ 3 คำนวณจากค่าต่ำของ α และค่าสูงของ β
4. $\hat{X}_t(\tau,\alpha_u,\beta_\ell)$ = ค่าพยากรณ์ที่ 4 คำนวณจากค่าสูงของ α และค่าต่ำของ β
5. $\hat{X}_t(\tau,\alpha_u,\beta_u)$ = ค่าพยากรณ์ที่ 5 คำนวณจากค่าสูงของพารามิเตอร์ทั้ง 2 ค่า

ในระเบียบวิธีของ Robert & Reed นั้น เมื่อกำหนดค่าชุดหนึ่งสำหรับพารามิเตอร์แล้ว จะใช้ค่าเหล่านี้ในการพยากรณ์ไประยะหนึ่งก่อนจึงจะพิจารณาความเหมาะสมของค่าพารามิเตอร์ ทั้งนี้เพื่อให้ค่าพยากรณ์มีเสถียรภาพมากขึ้น ให้ n เป็นจำนวนครั้งที่ใช้ค่าพารามิเตอร์ชุดหนึ่งในการพยากรณ์ก่อนที่จะมีการพิจารณาเปลี่ยนแปลงค่าพารามิเตอร์ และให้ E_{ij} เป็นความคลาดเคลื่อนสัมบูรณ์ของค่าพยากรณ์ที่ i ในการพยากรณ์ครั้งที่ j สำหรับค่าพารามิเตอร์ชุดหนึ่ง ๆ ดังนั้น ค่าเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์ของค่าพยากรณ์ที่ i สำหรับพารามิเตอร์ ชุดนี้ จึงมีค่าเท่ากับ

$$\overline{E}_i = \frac{1}{n}\sum_{j=1}^{n} E_{ij} \qquad i = 1, 2, 3, 4, 5 \tag{2.199}$$

ผลกระทบต่อความคลาดเคลื่อนสัมบูรณ์จากการเปลี่ยนค่าคงที่การทำให้เรียบ α จากค่าสูง α_u เป็นค่าต่ำ α_ℓ อาจเขียนได้โดยประมาณเท่ากับ

$$E_\alpha = \frac{1}{2}\left[\left(\overline{E}_4 + \overline{E}_5\right) - \left(\overline{E}_2 + \overline{E}_3\right)\right] \tag{2.200}$$

และในทำนองเดียวกัน ผลกระทบต่อความคลาดเคลื่อนสัมบูรณ์จากการเปลี่ยนค่าคงที่ทำให้เรียบ β จากค่าสูง β_u เป็นค่าต่ำ β_ℓ จะเท่ากับ

$$E_{\beta} = \frac{1}{2}\left[\left(\overline{E}_3 + \overline{E}_5\right) - \left(\overline{E}_2 + \overline{E}_4\right)\right] \tag{2.201}$$

จะเห็นได้ว่า ถ้า E_{α} มีค่าเป็นบวก ย่อมเป็นการแสดงให้เห็นว่า ค่าพยากรณ์เมื่อใช้ค่าต่ำของ α มีค่าใกล้เคียงกับค่าจริงมากกว่าค่าพยากรณ์เมื่อใช้ค่าสูงของ α และถ้า E_{α} มีค่าเป็นลบ ค่าพยากรณ์เมื่อใช้ค่าสูงของ α มีค่าใกล้เคียงกับค่าจริงมากกว่าค่าพยากรณ์เมื่อใช้ค่าต่ำของ α ข้อความในลักษณะทำนองเดียวกันกับที่ได้กล่าวเกี่ยวกับ α ข้างต้นนี้สามารถใช้ได้กับกรณีค่าคงที่การทำให้เรียบ β ด้วย ดังนั้น E_{α} และ E_{β} จึงอาจใช้เป็นแนวทางให้มีการเปลี่ยนค่าคงที่การทำให้เรียบได้ ในทางปฏิบัติเกณฑ์ที่จะใช้ในการตัดสินใจจะปรับเปลี่ยนค่าคงที่การทำให้เรียบ คือช่วงความเชื่อมั่น 99% ของผลกระทบ ซึ่งโดยประมาณเท่ากับ

$$\left[-3\hat{\sigma}_e / \sqrt{n}, 3\hat{\sigma}_e / \sqrt{n}\right]$$

โดย $\hat{\sigma}_e$ เป็นค่าประมาณของความเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน ถ้า E_{α} มีค่าน้อยกว่า $-3\hat{\sigma}_e / \sqrt{n}$ ค่าคงที่การทำให้เรียบเพิ่มค่าตามวิธีการของ Chow คือ

$$\begin{aligned} \alpha_0 &= \alpha_u \\ \alpha_u &= \alpha_0 + \delta \\ \alpha_{\ell} &= \alpha_0 - \delta \end{aligned}$$

และถ้า E_{α} มีค่ามากกว่า $-3\hat{\sigma}_e / \sqrt{n}$ ค่าคงที่การทำให้เรียบจะลดค่าลง คือ

$$\begin{aligned} \alpha_0 &= \alpha_{\ell} \\ \alpha_u &= \alpha_0 + \delta \\ \alpha_{\ell} &= \alpha_0 - \delta \end{aligned}$$

การปรับเปลี่ยนค่าคงที่การทำให้7เรียบ β ก็กระทำในลักษณะทำนองเดียวกันกับกรณีค่าคงที่การทำให้เรียบ α แต่ใช้ E_{β} แทน E_{α}

ในกรณีที่ตัวแบบการพยากรณ์มีพารามิเตอร์ 3 ตัว เช่น ในระเบียบวิธีของ Winters ในการพยากรณ์รอบหนึ่ง ๆ จะมีค่าพยากรณ์ทั้งหมด $2^3 + 1 = 9$ ตัว ที่จะต้องพิจารณา

1. $\hat{X}_t(\tau, \alpha_0, \beta_0, \gamma_0)$ ค่าพยากรณ์ที่ 1 คำนวณจากค่าฐานของพารามิเตอร์ทั้ง 3 ตัว
2. $\hat{X}_t(\tau, \alpha_{\ell}, \beta_{\ell}, \gamma_{\ell})$ ค่าพยากรณ์ที่ 2 คำนวณจากค่าต่ำของพารามิเตอร์ทั้ง 3 ตัว
3. $\hat{X}_t(\tau, \alpha_{\ell}, \beta_{\ell}, \gamma_u)$ ค่าพยากรณ์ที่ 3 คำนวณจากค่าต่ำของ α และ β และค่าสูงของ γ
4. $\hat{X}_t(\tau, \alpha_{\ell}, \beta_u, \gamma_{\ell})$ ค่าพยากรณ์ที่ 4 คำนวณจากค่าต่ำของ α และ γ และค่าสูงของ β

5. $\hat{X}_t(\tau,\alpha_\ell,\beta_u,\gamma_u)$ ค่าพยากรณ์ที่ 5 คำนวณจากค่าต่ำของ α และค่าสูงของ β และ γ
6. $\hat{X}_t(\tau,\alpha_u,\beta_\ell,\gamma_\ell)$ ค่าพยากรณ์ที่ 6 คำนวณจากค่าสูงของ α และค่าต่ำของ β และ γ
7. $\hat{X}_t(\tau,\alpha_u,\beta_\ell,\gamma_u)$ ค่าพยากรณ์ที่ 7 คำนวณจากค่าสูงของ α และ γ และค่าต่ำของ β
8. $\hat{X}_t(\tau,\alpha_u,\beta_u,\gamma_\ell)$ ค่าพยากรณ์ที่ 8 คำนวณจากค่าสูงของ α และ β และค่าต่ำของ γ
9. $\hat{X}_t(\tau,\alpha_u,\beta_u,\gamma_u)$ ค่าพยากรณ์ที่ 9 คำนวณจากค่าสูงของพารามิเตอร์ทั้ง 3 ตัว

การปรับค่าคงที่การทำให้เรียบในกรณีพารามิเตอร์ 3 ตัว จะกระทำในลักษณะทำนองเดียวกันกับกรณีพารามิเตอร์ 2 ตัว โดยคำนวณผลกระทบต่อความคลาดเคลื่อนจากการปรับค่าคงที่การทำให้เรียบจากค่าสูงเป็นค่าต่ำดังนี้

$$E_\alpha = \frac{1}{4}\left[\left(\overline{E}_6 + \overline{E}_7 + \overline{E}_8 + \overline{E}_9\right) - \left(\overline{E}_2 + \overline{E}_3 + \overline{E}_4 + \overline{E}_5\right)\right] \tag{2.202}$$

$$E_\beta = \frac{1}{4}\left[\left(\overline{E}_4 + \overline{E}_5 + \overline{E}_8 + \overline{E}_9\right) - \left(\overline{E}_2 + \overline{E}_3 + \overline{E}_6 + \overline{E}_7\right)\right] \tag{2.203}$$

$$E_\gamma = \frac{1}{4}\left[\left(\overline{E}_3 + \overline{E}_5 + \overline{E}_7 + \overline{E}_9\right) - \left(\overline{E}_2 + \overline{E}_4 + \overline{E}_6 + \overline{E}_8\right)\right] \tag{2.204}$$

และความเชื่อมั่น 99% ของผลกระทบในกรณีนี้ จะเท่ากับ

$$\left[-3\hat{\sigma}_e / \sqrt{2n}, 3\hat{\sigma}_e / \sqrt{2n}\right]$$

รายละเอียดการปรับค่าคงที่การทำให้เรียบ α, β และ γ จะกระทำในทำนองเดียวกันกับที่ได้กล่าวมาแล้วในกรณีที่ตัวแบบมีพารามิเตอร์ 2 ตัว

4. ตัวแบบค่าเฉลี่ยเคลื่อนที่ชั้นเดี่ยวแบบปรับจำนวนพจน์ (Adaptive Moving Average)

เมื่อศึกษาแนวความคิดของ Chow เกี่ยวกับข้อเสนอแนะในการปรับค่าคงที่การทำให้เรียบในการพยากรณ์ตามระเบียบวิธีการทำให้เรียบแบบเอกซ์โปเนนเชียลแล้ว ก็อาจใช้แนวความคิดของ Chow ดังกล่าวในการสร้างตัวแบบค่าเฉลี่ยเคลื่อนที่ชั้นเดียวแบบปรับจำนวนพจน์ได้ โดยสร้างตัวแบบค่าเฉลี่ยเคลื่อนที่ชั้นเดียว 3 ตัวแบบ ตัวแบบที่ใช้พยากรณ์ใช้จำนวนพจน์เท่ากับ N ตัวแบบที่ใช้เปรียบเทียบอีก 2 ตัวแบบ ใช้จำนวนพจน์เท่ากับ $N_\ell = N - \delta_N$ และ $N_u = N + \delta_N$ ค่า δ_N จะเป็นค่าที่ใช้ปรับจำนวนพจน์ N ในตัวแบบที่ใช้พยากรณ์ จึงต้องเป็นค่าตัวเลขเต็ม และไม่ควรมีค่ามาก ในกรณีที่ N มีค่าน้อย δ_N จะกำหนดให้มีค่าเท่ากับ 1 ทั้งนี้ เพื่อให้ค่าพยากรณ์มีเสถียรภาพ

ความคลาดเคลื่อนของตัวแบบค่าเฉลี่ยเคลื่อนที่ชั้นเดียวทั้ง 3 ตัวแบบ คือ

$$e_{t+1}(N) = \hat{X}_{t+1} - \hat{X}_t(N,1) \tag{2.205}$$

$$e_{t+1}(N_u) = \hat{X}_{t+1} - \hat{X}_t(N_u, 1) \tag{2.206}$$

$$e_{t+1}(N_\ell) = \hat{X}_{t+1} - \hat{X}_t(N_\ell, 1) \tag{2.207}$$

จะนำไปคำนวณความคลาดเคลื่อนสัมบูรณ์ที่ทำให้เรียบแล้ว ตามสมการ (2.186) ซึ่งจะใช้เป็นฐานในการปรับจำนวนพจน์ในตัวแบบค่าเฉลี่ยเคลื่อนที่ชั้นเดียว ดังนี้

1. ถ้า $\Delta_t(N)$ มีค่าไม่มากกว่า $\Delta_t(N_u)$ และ $\Delta_t(N_\ell)$ ให้ยังคงใช้จำนวนพจน์ N ในการพยากรณ์ครั้งต่อไป

2. ถ้า $\Delta_t(N)$ มีค่ามากกว่า $\Delta_t(N_u)$ แต่ไม่มากกว่า $\Delta_t(N_\ell)$ ให้ปรับจำนวนพจน์ในการพยากรณ์ครั้งต่อไปนี้

$$N = N_u \tag{2.208}$$

$$N_u = N + \delta_N \tag{2.209}$$

$$N_\ell = N - \delta_N \tag{2.210}$$

3. ถ้า $\Delta_t(N)$ มีค่ามากกว่า $\Delta_t(N_\ell)$ แต่ไม่มากกว่า $\Delta_t(N_u)$ ให้ปรับจำนวนพจน์ในการพยากรณ์ครั้งต่อไปดังนี้

$$N = N_\ell \tag{2.211}$$

$$N_u = N + \delta_N \tag{2.212}$$

$$N_\ell = N - \delta_N \tag{2.213}$$

4. ถ้า $\Delta_t(N)$ มีค่ามากกว่าทั้ง $\Delta_t(N_u)$ และ $\Delta_t(N_\ell)$ ให้พิจารณาดังนี้ ถ้า $\Delta_t(N_u)$ มีค่าน้อยกว่า $\Delta_t(N_\ell)$ ให้ปรับค่า N, N_u และ N_ℓ ตามสมการ (2.208), (2.209) และ (2.210) ตามลำดับ มิฉะนั้น ให้ปรับค่า N, N_u และ N_ℓ ตามสมการ (2.211), (2.212) และ (2.213) ตามลำดับ

เพื่อป้องกันมิให้มีการปรับจำนวนพจน์ที่ใช้ในการคำนวณค่าเฉลี่ยมากเกินไป หรือน้อยเกิน ดังที่ได้เสนอแนะไว้ในการปรับค่าคงที่การทำให้เรียบตามวิธีของ Chow และตามวิธีของ Trigg & Leach จึงควรกำหนดจำนวนพจน์มากที่สุด $N_{\max}$ และจำนวนพจน์น้อยที่สุด $N_{\min}$ ที่จะใช้ในการคำนวณค่าเฉลี่ยเคลื่อนที่ $N_{\max}$ และ $N_{\min}$ ควรมีค่าซึ่งทำให้ผลต่าง ($N_{\max}$ - N) และ (N-$N_{\min}$) หารด้วย δ_N ลงตัว โดย N คือ จำนวนพจน์เริ่มต้นที่ใช้ในการคำนวณเมื่อตัวแบบค่าเฉลี่ยเคลื่อนที่ชั้นเดียวมีการปรับจำนวนพจน์ที่ใช้คำนวณค่าเฉลี่ยสูงขึ้นเรื่อย ๆ เพื่อเพิ่มความเฉื่อยในค่า

พยากรณ์จนถึงค่า $N_{\max} - \delta_N$ และยังคงเรียกร้องให้ปรับจำนวนพจน์ให้สูงขึ้นอีก ในกรณีนี้จะปรับค่า N, N_u และ N_ℓ ดังนี้

$$N = N_{\max} \tag{2.214}$$

$$N_u = N_{\max} \tag{2.215}$$

$$N_\ell = N - \delta_N \tag{2.216}$$

ซึ่งจะทำให้จำนวนพจน์ในตัวแบบพยากรณ์เท่ากับจำนวนพจน์มากที่สุดที่กำหนดไว้ และเท่ากับจำนวนพจน์ในตัวแบบเปรียบเทียบทางสูงอีกด้วย ในทางตรงข้าม เมื่อตัวแบบค่าเฉลี่ยเคลื่อนที่ชั้นเดียวมีการปรับจำนวนพจน์ที่ใช้คำนวณค่าเฉลี่ยลดลงเรื่อย ๆ เพื่อให้ค่าพยากรณ์มีการปรับค่ารวดเร็วขึ้นจนถึงค่า $N_{\min} - \delta_N$ และยังคงเรียกร้องให้ปรับจำนวนพจน์ลงอีก ในกรณีนี้จะปรับค่า N, N_u และ N_ℓ ดังนี้

$$N = N_{\min} \tag{2.217}$$

$$N_u = N + \delta_N \tag{2.218}$$

$$N_\ell = N_{\min} \tag{2.219}$$

ซึ่งจะทำให้จำนวนพจน์ในตัวแบบพยากรณ์เท่ากับจำนวนพจน์น้อยที่สุดที่กำหนดไว้ และเท่ากับจำนวนพจน์ในตัวแบบเปรียบเทียบทางต่ำอีกด้วย ในกรณีที่ N ถูกปรับลดลงมาจนเท่ากับ $N_{\min}$ และกำหนดค่า $N_{\min}$ ให้เท่ากับหนึ่ง ตัวแบบการพยากรณ์จะกลายเป็นตัวแบบการพยากรณ์แบบ naïve คือ ค่าพยากรณ์หนึ่งหน่วยเวลาล่วงหน้า มีค่าเท่ากับค่าปัจจุบัน ส่วน $N_{\max}$ ไม่ควรกำหนดให้มีค่ามากนัก เพราะจะทำให้ค่าพยากรณ์มีความเฉื่อยมากเกินไป และเมื่อค่าจริงมีการเปลี่ยนแปลงอย่างฉับพลัน ตัวแบบจะต้องใช้ระยะเวลาหนึ่งในการปรับค่า N ที่เหมาะสม โดยปรับลดได้ δ_N ต่อหน่วยเวลาเท่านั้น แต่ถ้ากำหนดให้ δ_N มีค่ามาก ตัวแบบจะสามารถปรับตัวได้อย่างรวดเร็วจนบางครั้งอาจทำให้ค่าพยากรณ์ขาดเสถียรภาพได้ ดังนั้น การกำหนดค่าเหมาะสมสำหรับ $N_{\min}, N_{\max}$ และ δ_N จะต้องคำนึงถึงลักษณะการเคลื่อนไหวของข้อมูลด้วย

2.4.2 วิธีการพยากรณ์แบบปรับได้ที่ปรับปรุงใหม่

การพยากรณ์ตามวิธีการของ Chow ต้องใช้ตัวแบบการทำให้เรียบแบบเอกซ์โปเนนเชียลถึง 3 ตัวแบบ ทำให้มีภาระมากในการคำนวณเมื่อเปรียบเทียบกับวิธีการของ Trigg & Leach ซึ่งใช้ตัวแบบการทำให้เรียบแบบเอกซ์โปเนนเชียลเพียงตัวแบบเดียว โดยให้ค่าคงที่การทำให้เรียบปรับ

ค่าเท่ากับค่าสัมบูรณ์ของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว จุดเด่นของวิธีการของ Chow คือ ความมีเสถียรภาพของค่าคงที่การทำให้เรียบ โดยให้ค่อย ๆ ปรับครั้งละ 0.05 ดังนั้น หากนำแนวความคิดของ Trigg & Leach มาปรับปรุงในวิธีการของ Chow ก็จะสามารถพยากรณ์ด้วยตัวแบบการทำให้เรียบแบบเอกซ์โปเนนเชียลเพียงตัวแบบเดียว การเพิ่มขึ้นและการลดลงในค่าสัมบูรณ์ของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว ทำให้ค่าคงที่การทำให้เรียบในตัวแบบการพยากรณ์การทำให้เรียบแบบเอกซ์โปเนนเชียลมีค่าสูงขึ้นและลดลงตามลำดับ เมื่อเป็นเช่นนั้น การปรับค่าคงที่การทำให้เรียบตามวิธีการเดิมของ Chow ก็อาจสามารถปรับปรุงใหม่เพื่อลดภาระการคำนวณได้ดังนี้

1. กรณีที่ค่าสัมบูรณ์ของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว ณ เวลา t มีค่าสูงขึ้น ค่าคงที่การทำให้เรียบสำหรับการพยากรณ์ที่กระทำ ณ เวลา t จะถูกปรับให้มีค่าสูงขึ้น 0.05

2. กรณีที่ค่าสัมบูรณ์ดังกล่าวมีค่าลดลง ค่าคงที่การทำให้เรียบสำหรับการพยากรณ์ที่กระทำ ณ เวลา t จะถูกปรับให้มีค่าลดลง 0.05

3. กรณีที่ไม่มีการเปลี่ยนแปลงในค่าสัมบูรณ์ของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว ณ เวลา t จะไม่มีการปรับค่าคงที่การทำให้เรียบสำหรับการพยากรณ์ที่กระทำ ณ t เวลา แต่ประการใด

วิธีการพยากรณ์ดังกล่าวข้างต้น ซึ่งเป็นการบูรณาการแนวความคิดของ Chow และ Trigg & Leach จึงขอเรียกสั้น ๆ ว่า วิธีการ CTL (Chow, Trigg & Leach) การใช้ตัวแบบการทำให้เรียบแบบเอกซ์โปเนนเชียลเพียงตัวแบบเดียว ทำให้ภาระการคำนวณลดลงได้อย่างมีนัยสำคัญ การใช้การเพิ่มขึ้นและลดลงในค่าสัมบูรณ์ของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว มาปรับค่าคงที่การทำให้เรียบครั้งละ 0.05 ตามวิธีการของ Chow ทำให้ค่าคงที่การทำให้เรียบมีเสถียรภาพ ไม่ปรับค่าอย่างรวดเร็วดังในวิธีการ Trigg & Leach

แนวความคิดการใช้การเพิ่มขึ้นและลดลงในค่าสัมบูรณ์ของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว ยังอาจนำมาประยุกต์ในกรณีของตัวแบบการพยากรณ์ด้วยค่าเฉลี่ยเคลื่อนที่ชั้นเดียวแบบปรับจำนวนพจน์ได้ด้วย กล่าวคือ

1. กรณีที่ค่าสัมบูรณ์ของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว ณ เวลา t มีค่าสูงขึ้น จำนวนพจน์ที่ใช้ในการพยากรณ์ที่กระทำ ณ เวลา t จะถูกปรับลดลงด้วย δ_N

2. กรณีที่ค่าสัมบูรณ์ดังกล่าวมีค่าลดลง จำนวนพจน์ที่ใช้ในการพยากรณ์ที่กระทำ ณ เวลา t จะถูกปรับเพิ่มขึ้นด้วย δ_N

3. กรณีที่ไม่มีการเปลี่ยนแปลงในค่าสัมบูรณ์ของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว ณ เวลา t จะไม่มีการปรับจำนวนพจน์ที่ใช้ในการพยากรณ์ที่กระทำ ณ เวลา t

วิธีการพยากรณ์ด้วยค่าเฉลี่ยเคลื่อนที่แบบปรับจำนวนพจน์ โดยใช้การเพิ่มขึ้นและลดลงในค่าสัมบูรณ์ของสัญญาณติดตามความคลาดเคลื่อนที่ทำให้เรียบแล้ว จะขอเรียกสั้น ๆ ว่า วิธีการ AMA-TL (Adaptive Moving Average-Trigg & Leach)

2.4.3 การพยากรณ์แบบการกรองที่ปรับได้ (Adaptive Filtering)

ในตัวแบบต่าง ๆ ภายใต้หลักการการทำให้เรียบ อาจสรุปได้ว่า ค่าพยากรณ์จะเขียนได้ในพจน์ของข้อมูลในอดีต n ตัว ได้ดังนี้

$$\hat{X}_t(1) = W_0 X_t + W_1 X_{t-1} + ... + W_{n-1} X_{t-n+1} \tag{2.220}$$

โดย W_i เป็นสัมประสิทธิ์ที่มีค่าคงที่ เช่น ในตัวแบบค่าเฉลี่ยเคลื่อนที่ W_i มีค่าเท่ากับ $\frac{1}{n}$ เป็นต้น ตัวแบบการพยากรณ์แบบการกรองที่ปรับได้ จะมีค่าพยากรณ์อยู่ในรูปสมการ (2.220) เช่นกัน แต่โดยสัมประสิทธิ์ W_i ไม่เป็นค่าคงที่ แต่สามารถปรับค่าให้สอดคล้องกับการเปลี่ยนแปลงของข้อมูลได้ จึงทำให้ค่าพยากรณ์จากตัวแบบการพยากรณ์แบบการกรองที่ปรับได้มีความยืดหยุ่นมาก

Wheelwright & Makridakis ได้นำแนวความคิดของ Widrow มาใช้ในการปรับค่าสัมประสิทธิ์ W_i บนพื้นฐานของความคลาดเคลื่อนที่เกิดขึ้น เนื่องจากค่าสัมประสิทธิ์ W_i อาจปรับเปลี่ยนค่าได้ตามหน่วยเวลาที่เคลื่อนไป จึงควรจะปรับสัญลักษณ์ของสัมประสิทธิ์ในสมการ (2.220) ใหม่ เพื่อให้สะท้อนถึงการเปลี่ยนค่าของ W_i ตามหน่วยเวลา ดังนี้

$$\hat{X}_t(1) = W_{0t} X_t + W_{1t} X_{t-1} + ... + W_{(n-1)t} X_{t-n+1} \tag{2.221}$$

โดย W_{it} เป็นสัมประสิทธิ์ของ X_{t-i} ในการพยากรณ์ที่กระทำ ณ เวลา t สัมประสิทธิ์ในสมการ (3.43) จะถูกปรับค่าเพื่อให้ความคลาดเคลื่อนกำลังสองมีค่าน้อยที่สุด ซึ่ง Widrow ได้สร้างอัลกอริทึมในการปรับค่าสัมประสิทธิ์เพื่อบรรลุเป้าหมายดังกล่าว โดยให้ค่าสัมประสิทธิ์ $W_{i(1+1)}$ ที่จะใช้ในการพยากรณ์ที่กระทำ ณ เวลา $t+1$ มีค่าเท่ากับ

$$\begin{aligned} W_{i(t+1)} &= W_{it} - k\frac{\partial e_{t+1}^2}{\partial W_{it}} \\ &= W_{it} + 2ke_{t+1}X_{t-1} \end{aligned} \tag{2.222}$$

โดย k เป็นค่าคงที่ในการเรียน (learning constant) ซึ่งมีค่าเป็นบวก การกำหนดค่า จะมีอิทธิพลมากต่อการลู่เข้าสู่ค่าอุตตมะของสัมประสิทธิ์ ถ้ากำหนดค่าต่ำเกินไป การปรับค่าของสัมประสิทธิ์จะเชื่องช้า ทำให้ลู่เข้าสู่ค่าอุตตมะช้า แต่ถ้ากำหนดค่าสูงเกินไป การปรับค่าสัมประสิทธิ์อาจรวดเร็ว

เกินไปจนขาดเสถียรภาพ อาจไม่ลู่เข้าสู่ค่าอุตตมะได้ Makridakis & Wheelwright ได้แนะนำให้กำหนดค่าคงที่ในการเรียนอยู่ระหว่างศูนย์กับ $\frac{1}{n}$ เพื่อให้เกิดเสถียรภาพในการลู่เข้าสู่ค่าอุตตมะ

$$0 < k < \frac{1}{n} \tag{2.223}$$

โดยปรับข้อมูลและความคลาดเคลื่อนเป็นข้อมูลมาตรฐาน และความคลาดเคลื่อนมาตรฐาน (standardized data and standardized error) ดังนี้

$$\begin{aligned} W_{i(t+1)} &= W_{it} + 2k \frac{e_{t+1}}{\sqrt{\sum_{i=0}^{n-1} X_{t-1}^2}} \frac{X_{t-i}}{\sqrt{\sum_{i=0}^{n-1} X_{t-1}^2}} \\ &= W_{it} + \frac{2ke_{t+1} X_{t-i}}{\sum_{i=0}^{n-1} X_{t-1}^2} \end{aligned} \tag{2.224}$$

เมื่อพิจารณาสมการ (2.222) และ (2.223) ให้ละเอียด อาจกล่าวได้ว่าในกรณีที่ไม่ได้ปรับข้อมูลให้เป็นข้อมูลมาตรฐาน ค่าคงที่ในการเรียนควรมีอยู่ระหว่างช่วง

$$0 < k < \frac{1}{n\left[\max_{t \geq n} \sum_{i=0}^{n-1} X_{t-1}^2\right]} \tag{2.225}$$

ซึ่งในทางปฏิบัติ มักจะกำหนดค่าคงที่ในการเรียน k ให้มีค่าใกล้เคียงกับค่าสูงสุดในสมการ (2.225) เพื่อให้ค่าสัมประสิทธิ์ลู่เข้าสู่ค่าอุตตมะอย่างรวดเร็ว และมีเสถียรภาพ และใช้สมการ (2.222) ในการปรับค่าสัมประสิทธิ์ของค่าพยากรณ์

การคำนวณค่าอุตตมะของสัมประสิทธิ์ เพื่อใช้ในการพยากรณ์จะกระทำเป็นรอบ ๆ (iteration) ก่อนที่จะเริ่มต้นการคำนวณจะต้องกำหนดค่าคงที่ในการเรียน k จำนวนพจน์ที่ใช้ในการพยากรณ์ ตามสมการ (2.220) และค่าเริ่มต้นของสัมประสิทธิ์ทั้ง n ค่า เมื่อได้กำหนดค่าเริ่มต้นเหล่านี้เสร็จเรียบร้อยแล้ว การคำนวณในรอบหนึ่ง ๆ จะเป็นดังนี้

1. คำนวณค่าพยากรณ์โดยใช้สมการ (2.221)
2. เคลื่อนหน่วยเวลาไป 1 หน่วยเวลา
3. ถ้าหน่วยเวลายังไม่เกินหน่วยเวลาสุดท้ายของข้อมูลให้ไปคำนวณในขั้นตอนที่ 4 มิฉะนั้นเป็นการสิ้นสุดของรอบ
4. คำนวณความคลาดเคลื่อนของค่าพยากรณ์
5. คำนวณค่าสัมประสิทธิ์ใหม่ทั้ง n ค่าด้วยสมการ (2.222)

6. กลับไปคำนวณในขั้นตอนที่ 1

เมื่อสิ้นสุดการคำนวณในรอบหนึ่ง ๆ จะต้องประเมินผลเพื่อกำหนดว่าสมควรที่จะเริ่มต้นคำนวณใหม่ในรอบถัดไปหรือไม่ เกณฑ์ที่ใช้ในการตัดสินใจที่จะเริ่มต้นคำนวณใหม่หรือไม่ อาจมีได้หลายเกณฑ์ เช่น เปรียบเทียบค่า MSE ของค่าพยากรณ์ในรอบปัจจุบันกับรอบที่แล้วลดลง อย่างมีนัยสำคัญหรือไม่ เปรียบเทียบค่าสัมประสิทธิ์เมื่อสิ้นสุดรอบกับค่าสัมประสิทธิ์เมื่อเริ่มต้นรอบทั้ง n ค่าว่ามีความแตกต่างกันอย่างมีนัยสำคัญหรือไม่ หรือเปรียบเทียบค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อนสัมพัทธ์

$$ARE = \frac{\sum_t |e_t|}{\sum_t |X_t|} \qquad (2.226)$$

ในรอบปัจจุบันกับรอบที่แล้วลดลงอย่างมีนัยสำคัญหรือไม่ เป็นต้น ถ้าเปรียบเทียบค่าเหล่านี้ปรากฏว่ามีความแตกต่างอย่างมีนัยสำคัญให้ใช้ค่าสัมประสิทธิ์ที่คำนวณ ได้เมื่อสิ้นสุดรอบเป็นค่าเริ่มต้นของสัมประสิทธิ์ในรอบใหม่ และดำเนินการคำนวณซ้ำตามขั้นตอนที่ได้กล่าวมาแล้ว มิฉะนั้นจะถือว่าค่าสัมประสิทธิ์ที่คำนวณได้เมื่อสิ้นสุดรอบเป็นค่าอุตตมะและจะใช้เพื่อการพยากรณ์ต่อไป