สำนักหอสมุด

สัญญาเลขที่ R2559C098

รายงานวิจัยฉบับสมบูรณ์

เรื่อง

การประมาณพารามิเตอร์ของแบบจำลองความผันผวนเชิงเฟ้นสุ่ม : หลักฐานเชิงประจักษ์จากอัตราแลกเปลี่ยนสกุลเงินดอลลาร์สหรัฐต่อเงินบาทไทย

สำนักหอสมุด มหาวิทยาลัยนเรศวร
วันลงทะเบียน 11 มิ.ย. 2564
เลขทะเบียน 10 38695
เลขเรียกหนังสือ ว TA
345
ป262ร
2560

โดย

รองศาสตราจารย์ ดร. ธนิต มาลากร

ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

สนับสนุนโดยงบประมาณรายได้มหาวิทยาลัยนเรศวร ปีงบประมาณ 2559

ประกาศคุณูปการ

การทำวิจัยครั้งนี้ไม่สามารถสำเร็จลุล่วงไปได้ หากปราศจากการสนับสนุนจากหน่วยงานต่าง ๆ โดยเฉพาะอย่างยิ่งกองทุนวิจัยมหาวิทยาลัยนเรศวรซึ่งสนับสนุนด้านงบประมาณ หน่วยงานการเงินและพัสดุ คณะวิศวกรรมศาสตร์ ที่ให้การช่วยเหลือและให้คำแนะนำในการเบิกจ่ายเงิน รวมทั้งหน่วยงานวิจัย คณะวิศวกรรมศาสตร์ ที่ช่วยให้คำแนะนำเกี่ยวกับการจัดทำรูปเล่มรายงานฉบับสมบูรณ์ พร้อมทั้งคอยชี้แจงเกี่ยวกับหลักเกณฑ์ต่าง ๆ ที่เกี่ยวข้องกับงานวิจัย

ผู้จัดทำใคร่ขอขอบคุณบิดา มารดา ย่า ป้า และ อา สำหรับคำแนะนำที่ดีเยี่ยมและให้การสนับสนุนอย่างเต็มที่ด้วยความเต็มใจอย่างเสมอมา รวมทั้งมอบความรักความอบอุ่นให้กับผู้จัดทำอยู่โดยตลอด ขอขอบคุณเพื่อน ๆ ที่คอยให้กำลังใจอย่างสม่ำเสมอ และขอบใจนิสิตหลายคนที่ช่วยเป็นธุระให้ในหลายด้าน

ธนิต มาลากร

บทคัดย่อ

งานวิจัยชิ้นนี้มุ่งศึกษาการประยุกต์ใช้ขั้นตอนวิธี EM ร่วมกับตัวกรองอนุภาคเกาส์เซียนในการประมาณพารามิเตอร์แบบความควรจะเป็นสูงสุดสำหรับแบบจำลองความผันผวนเชิงเฟ้นสุ่มโดยใช้ข้อมูล 2 ชุดในการทดลอง กล่าวคือข้อมูลที่ได้จากการจำลองและข้อมูลของอัตราแลกเปลี่ยนเงินตราระหว่างประเทศ ผลการทดลองเมื่อใช้ข้อมูลที่ได้จากการจำลองแสดงให้เห็นว่าแนววิถีของพารามิเตอร์ที่ประมาณขึ้นมานั้นมีแนวโน้มลู่เข้าไปยังพารามิเตอร์จริงของแบบจำลอง เมื่อทำการทดลองกับข้อมูลของอัตราแลกเปลี่ยนโดยใช้ทั้งวิธีที่นำเสนอในงานวิจัยนี้และวิธีตามแบบฉบับซึ่งใช้ตัวกรองอนุภาคเพื่อทำการเปรียบเทียบปรากฏว่าพารามิเตอร์ที่ประมาณได้จากทั้ง 2 วิธีมีค่าสอดคล้องกัน อย่างไรก็ตามเมื่อพิจารณาเวลาในการประมวลผลพบว่าการเลือกใช้ตัวกรองที่นำเสนอในงานวิจัยนี้ใช้เวลาน้อยกว่าการใช้ตัวกรองอนุภาคในวิธีตามแบบฉบับ

ABSTRACT

This research primarily studies the application of the expectation-maximization (EM) algorithm coupled with a Gaussian particle filter for maximum likelihood parameter estimation of stochastic volatility models. Two data sets are provided for demonstration purposes: simulated data and daily foreign exchange rates data. Simulation studies illustrate that the parameter estimate trajectories are likely to converge to the true ones. When comparing the empirical results obtained from the proposed method and the typical method based on the particle filter, it can be seen that the parameter estimates from both methods are consistent with each other; however, the computational time of the former is significantly reduced as compared to the latter one.

สารบัญ

บทที่		หน้า
1	**บทนำ**	1
	ความเป็นมาของปัญหา	1
	จุดมุ่งหมายของการวิจัย	6
	ประโยชน์และความสำคัญ	7
	ขอบเขตของการวิจัย	7
2	**ความรู้เบื้องต้น ขั้นตอนวิธีและงานวิจัยที่เกี่ยวข้อง**	8
	ความผันผวนและแบบจำลอง	8
	แบบจำลองมาร์คอฟซ่อน	11
	การอนุมานแบบเบย์	14
	วิธีมอนติคาร์โลในกระบวนการเชิงเฟ้นสุ่ม	16
	ตัวกรองอนุภาค ตัวกรองการปลุกเครื่องและตัวกรองอนุภาคเกาส์เซียน . . .	19
	ตัวปรับเรียบอนุภาคและตัวปรับเรียบอนุภาคแบบย้อนกลับ	23
	การประมาณพารามิเตอร์และขั้นตอนวิธี EM	25
3	**วิธีดำเนินการวิจัย**	33
	ภาพรวมของขั้นตอนดำเนินการวิจัย	33
	การประมาณพารามิเตอร์ของแบบจำลอง SV ที่บรรยายในรูปแบบไม่เชิงเส้น	34
	การประมาณพารามิเตอร์ของแบบจำลอง SV ที่บรรยายในรูปแบบเชิงเส้น . .	38
4	**ผลการทดลอง**	41
	ผลการทดลองของแบบจำลอง SV ที่บรรยายในรูปแบบไม่เชิงเส้น	41
	ผลการทดลองของแบบจำลอง SV ที่บรรยายในรูปแบบเชิงเส้น	45
5	**บทสรุป**	50
	สรุปงานวิจัย	50

ข้อเสนอแนะและแนวทางการพัฒนางานวิจัย 51

บรรณานุกรม 53

ภาคผนวก 58

สารบัญตาราง

ตาราง		หน้า
1	$\hat{\theta} = \{ \hat{\phi}, \hat{Q}, \hat{\beta} \}$ สำหรับ Σ_{NL} ของอัตราแลกเปลี่ยน	42
2	คุณสมบัติด้านโมเมนต์ของฟังก์ชันการแจกแจงสำหรับ Σ_{NL}	44
3	$\hat{\theta} = \{ \hat{\phi}, \hat{Q}, \hat{\alpha} \}$ สำหรับ Σ_{L} ของอัตราแลกเปลี่ยน	46
4	คุณสมบัติด้านโมเมนต์ของฟังก์ชันการแจกแจงสำหรับ Σ_{L}	49

สารบัญภาพ

ภาพ		หน้า
1	ขั้นตอนวิธีของตัวกรองการปลุกเครื่อง : BF	21
2	ขั้นตอนวิธีของตัวกรองอนุภาคเกาส์เซียน : GPF	23
3	ขั้นตอนวิธีของตัวปรับเรียบอนุภาคแบบย้อนกลับ : BS-PS	26
4	ผังงานแสดงการทำงานของโปรแกรม	35
5	การลู่เข้าของ $\hat{\theta} = \{\hat{\phi}, \hat{Q}, \hat{\beta}\}$ สำหรับ Σ_{NL}	43
6	เวลาในการประมวลผลระหว่าง BF และ GPF ของ USD/THB สำหรับ Σ_{NL} .	43
7	ฟังก์ชันการแจกแจงของผลตอบแทนจริงและผลตอบแทนจากการจำลอง Σ_{NL}	44
8	ค่าความผันผวนและผลตอบแทนของ USD/THB สำหรับ Σ_{NL}	45
9	การลู่เข้าของ $\hat{\theta} = \{\hat{\phi}, \hat{Q}, \hat{\alpha}\}$ สำหรับ Σ_L	47
10	เวลาในการประมวลผลระหว่าง BF และ GPF ของ USD/THB สำหรับ Σ_L . .	47
11	ฟังก์ชันการแจกแจงของผลตอบแทนจริงและผลตอบแทนจากการจำลอง Σ_L	48
12	ค่าความผันผวนและผลตอบแทนของ USD/THB สำหรับ Σ_L	49

อักษรย่อ

AR	การถดถอยอัตตะ (Autoregression)
BF	ตัวกรองการปลุกเครื่อง (Bootstrap filter)
BS-PS	ตัวปรับเรียบอนุภาคแบบย้อนกลับ (Backward-simulation particle smoother)
DV	ความผันผวนเชิงกำหนด (Deterministic volatility)
EM	ขั้นตอนวิธีอีเอ็ม (Expectation Maximization method)
FV	ราคาที่เหมาะสม (Fair value)
GPF	ตัวกรองอนุภาคเกาส์เซียน (Gaussian particle filter)
HMM	แบบจำลองมาร์คอฟซ่อน (Hidden Markov model)
IS	การชักตัวอย่างสำคัญ (Importance sampling)
KF	ตัวกรองคาลมาน (Kalman filter)
KM	วิธีคาลมาน (Kalman method)
KS	ตัวปรับเรียบคาลมาน (Kalman smoother)
MAP	ความน่าจะเป็นภายหลังสูงสุด (Maximum a posteriori)
MCM	วิธีมอนติคาร์โล (Monte Carlo method)
ML	ความควรจะเป็นสูงสุด (Maximum likelihood)
PF	ตัวกรองอนุภาค (Particle filter)
QML	ความควรจะเป็นเสมือนสูงสุด (Quasi maximum likelihood)
SIR	การชักตัวอย่างสำคัญซ้ำโดยลำดับ (Sequential importance resampling)
SIR-PS	ตัวปรับเรียบอนุภาคแบบ SIR (SIR Particle smoother)
SIS	การชักตัวอย่างสำคัญโดยลำดับ (Sequential importance sampling)
SV	ความผันผวนเชิงเฟ้นสุ่ม (Stochastic volatility)

บทที่ 1

บทนำ

ความเป็นมาของปัญหา

สถานการณ์ด้านธุรกิจพลังงานของประเทศไทยในปัจจุบันไม่จำกัดแต่เฉพาะการดำเนินธุรกิจในประเทศเพียงเท่านั้น เป็นที่ทราบกันทั่วไปว่าหน่วยงานหลักที่มีความเกี่ยวข้องโดยตรงกับพลังงานของไทยมีการดำเนินธุรกิจระหว่างประเทศ เช่น การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย (กฟผ.) มีการทำสัญญาซื้อไฟฟ้าจากประเทศเพื่อนบ้าน และบริษัท ปตท. จำกัด (มหาชน) มีการทำสัญญาซื้อน้ำมันและแก๊สธรรมชาติจากต่างประเทศ ดังนั้นจึงเป็นสิ่งที่หลีกเลี่ยงไม่ได้ที่ต้องอาศัยการแลกเปลี่ยนสกุลเงินต่าง ๆ ระหว่างประเทศคู่ค้าเพื่อใช้ในการดำเนินธุรกิจ

อย่างไรก็ตาม อัตราแลกเปลี่ยนเงินตราระหว่างประเทศมีการเปลี่ยนแปลงอยู่โดยตลอด ดังนั้นเพื่อเป็นการป้องกันความเสี่ยงจากความผันผวนของค่าเงิน การทำธุรกิจระหว่างประเทศจึงนิยมทำสัญญาซื้อขายสกุลเงินล่วงหน้าซึ่งเป็นสัญญาที่คู่สัญญาทั้ง 2 ฝ่าย (อาจจะเป็นระหว่างบริษัทคู่ค้าด้วยกันหรือระหว่างบริษัทและสถาบันการเงิน) ตกลงที่จะซื้อหรือขายสกุลเงินใดสกุลเงินหนึ่งโดยมีการระบุจำนวนสกุลเงินที่ต้องการ อัตราแลกเปลี่ยนรวมทั้งเวลาที่ต้องการซื้อหรือขายไว้ในสัญญาที่จัดทำขึ้น ณ วันทำสัญญา เมื่อเวลาที่ระบุไว้ในสัญญามาถึง คู่สัญญาทั้ง 2 ฝ่ายยินยอมที่จะมีการซื้อหรือขายสกุลเงินในจำนวนและอัตราแลกเปลี่ยนที่ได้ระบุไว้ในสัญญาโดยไม่ขึ้นกับว่าอัตราแลกเปลี่ยน ณ เวลานั้นมีค่าเป็นเท่าใด สัญญาในลักษณะดังกล่าวเรียกว่า **สัญญาซื้อขายล่วงหน้า** (Forward contract) แต่หากต้องการนำสัญญามาซื้อขายผ่านตลาดที่มีการจัดตั้งอย่างเป็นทางการ สัญญานั้นจะถูกเรียกว่า **สัญญาฟิวเจอร์ส** (Futures contract) ซึ่งสัญญาทั้ง 2 ชนิดนั้นสามารถใช้สินทรัพย์อื่นนอกเหนือจากอัตราแลกเปลี่ยนเงินตราระหว่างประเทศมาใช้เป็นสินทรัพย์อ้างอิงได้ เช่น ดัชนีราคาหลักทรัพย์ หลักทรัพย์ พันธบัตรรัฐบาล อัตราดอกเบี้ย ทองคำ น้ำมัน เครดิตคาร์บอน เป็นต้น

แม้ว่าสัญญาซื้อขายล่วงหน้าและสัญญาฟิวเจอร์สจะช่วยลดความเสี่ยงอันเกิดจากความผันผวนของค่าเงินไปได้ในระดับหนึ่ง แต่หากมีการกำหนดอัตราแลกเปลี่ยนเงินตราระหว่างประเทศล่วงหน้าผิดพลาดย่อมส่งผลให้คู่สัญญาฝ่ายหนึ่งได้ประโยชน์ ในขณะที่คู่สัญญาอีกฝ่ายหนึ่งเสียประโยชน์เมื่อวันสิ้นสุดสัญญามาถึง เพราะสัญญาในลักษณะนี้ถือว่าเป็นข้อผูกมัดระหว่างคู่

สัญญา ต่อมาได้มีการพัฒนาเครื่องมือทางการเงินอีกประเภทหนึ่งเรียกว่า **ตราสารสิทธิ** (Option) ซึ่งเป็นสัญญาที่ให้สิทธิกับคู่สัญญาฝ่ายที่เป็นผู้ซื้อหรือผู้ถือตราสารสิทธิในการตัดสินใจว่าจะซื้อหรือจะขายสกุลเงินในอัตราแลกเปลี่ยนที่ถูกระบุไว้ในสัญญานั้นหรือไม่ กล่าวคือ หากผู้ซื้อตราสารสิทธิเห็นว่าอัตราแลกเปลี่ยนที่ระบุไว้ในสัญญาเอื้อประโยชน์แก่ฝ่ายตนจึงควรเลือกใช้สิทธิในการแลกเปลี่ยนสกุลเงินในอัตราดังกล่าว แต่หากเห็นว่าตนเองเสียประโยชน์จากอัตราแลกเปลี่ยนเงินตราระหว่างประเทศที่ระบุไว้ในสัญญา ผู้ซื้อตราสารสิทธิย่อมมีสิทธิที่จะยกเลิกสัญญาโดยไม่ทำการแลกเปลี่ยนสกุลเงินนั้นได้ ทั้งนี้ผู้ขายหรือผู้ออกตราสารสิทธิต้องปฏิบัติตามความต้องการของผู้ซื้อหรือผู้ถือตราสารสิทธิโดยไม่สามารถหลีกเลี่ยงได้ จากคุณสมบัติของตราสารสิทธิที่ให้สิทธิแก่ผู้ซื้อหรือผู้ถือตราสารสิทธิที่จะปฏิบัติตามข้อตกลงในสัญญาหรือไม่ก็ได้ ผู้ซื้อหรือผู้ถือตราสารสิทธิต้องชำระเงินจำนวนหนึ่งที่เรียกว่าค่าพรีเมียม (Premium) ให้กับผู้ขายหรือผู้ออกตราสารสิทธิเพื่อให้ได้มาซึ่งสิทธินั้น

ทั้งสัญญาฟิวเจอร์สและตราสารสิทธิต้องทำการซื้อขายผ่านตลาดที่มีการจัดตั้งอย่างเป็นทางการเท่านั้น ไม่สามารถซื้อขายระหว่างคู่สัญญากันเองในลักษณะ OTC (Over the counter) ได้ ตลาดที่เป็นสื่อกลางในการซื้อขายสัญญาฟิวเจอร์สและตราสารสิทธิในประเทศไทยคือ **ตลาดสัญญาซื้อขายล่วงหน้า** (TFEX) โดยสินทรัพย์อ้างอิงที่ทำการซื้อขายได้ในตลาดดังกล่าว ได้แก่ อัตราดอกเบี้ย ทองคำ น้ำมัน หลักทรัพย์และดัชนีราคาหลักทรัพย์ ทั้งนี้มีเพียงดัชนีราคาหลักทรัพย์ใน SET 50 เท่านั้นที่มีการซื้อขายแบบตราสารสิทธิได้ในขณะนี้ [1]

แม้ว่าราคาของสัญญาที่ทำการซื้อขายในตลาดจะขึ้นกับอุปสงค์และอุปทานของฝ่ายผู้เสนอซื้อและฝ่ายผู้เสนอขายแล้ว ราคาที่เหมาะสม (Fair value: FV) ของสัญญามีบทบาทสำคัญต่อการตัดสินใจในการซื้อขายสัญญาในตลาด การคำนวณค่า FV ของสัญญาฟิวเจอร์สค่อนข้างตรงไปตรงมา ในขณะที่การหาค่า FV ของตราสารสิทธิมีความซับซ้อนมาก โดยทั่วไปนิยมเลือกใช้สูตรแบลค์-โชวส์ (Black-Scholes formula) [2, 3] ซึ่งเป็นผลเฉลยของสมการเชิงอนุพันธ์เชิงเฟ้นสุ่มในการคำนวณหาค่า FV ของตราสารสิทธิ อย่างไรก็ตามสูตรแบลค์-โชวส์ที่ถูกพัฒนาขึ้นมาในช่วงแรกนั้นกำหนดให้ค่าความผันผวน (Volatility) ของสินทรัพย์อ้างอิงเป็นค่าคงที่ σ เนื่องจากตราสารสิทธิที่มีการซื้อขายในช่วงเวลานั้นมีอายุสั้นราว 2-3 เดือน ต่อมาเมื่อมีการพัฒนารูปแบบของตราสารสิทธิเพื่อให้มีอายุยาวนานมากขึ้นซึ่งตราสารสิทธิบางชนิดมีอายุมากกว่า 10 ปี การแทน σ ซึ่งเป็นค่าคงที่ลงในสูตรแบลค์-โชวส์อาจส่งผลให้ค่า FV ที่คำนวณได้แตกต่างจากค่า FV ที่แท้จริงของตราสารสิทธิ ดังนั้นเพื่อให้การคำนวณหาค่า FV ของตราสารสิทธิมีความแม่นยำมาก

ขึ้นจึงจำเป็นต้องสร้างแบบจำลองสำหรับความผันผวนที่มีการแปรเปลี่ยนตามเวลา σ_k แทนการใช้ค่าคงที่ตามแบบเดิม

แบบจำลองความผันผวนของสินทรัพย์ที่มีใช้แพร่หลายในปัจจุบันได้ถูกพัฒนาขึ้นมาอย่างหลากหลาย ทั้งในกลุ่มของ GARCH ซึ่งจัดว่าเป็นแบบจำลองความผันผวนเชิงกำหนด (Deterministic volatility model : DV) ได้แก่ EGARCH, IGARCH, NGARCH, QGARCH, TGARCH, EGARCH, GJR-GARCH และในกลุ่มของแบบจำลองความผันผวนเชิงเฟ้นสุ่ม (Stochastic volatility model : SV) เช่น แบบจำลอง Taylor แบบจำลอง Heston แบบจำลอง SABR และแบบจำลอง Chen สำหรับในรายงานวิจัยฉบับนี้เลือกใช้แบบจำลอง SV เนื่องจากมีโครงสร้างที่เขียนบรรยายในรูปแบบของปริภูมิสถานะซึ่งเป็นแบบจำลองที่นิยมใช้อย่างแพร่หลายในงานด้านสาขาวิศวกรรมควบคุมอัตโนมัติโดยใช้อัตราแลกเปลี่ยนสกุลเงินดอลลาร์สหรัฐต่อบาทไทย (USD/THB) มาเป็นกรณีศึกษาเนื่องจากตราสารสิทธิที่มีอัตราแลกเปลี่ยนเงินตราระหว่างประเทศเป็นสินทรัพย์อ้างอิงมีโอกาสที่จะเปิดทำการซื้อขายได้ในประเทศไทย อีกทั้งอัตราแลกเปลี่ยนเงินตราระหว่างประเทศมีการเปลี่ยนแปลงค่อนข้างซ้ากว่าราคาหลักทรัพย์หรือสินค้าโภคภัณฑ์

สิ่งสำคัญในการสร้างแบบจำลองคือการประมาณพารามิเตอร์ของแบบจำลองโดยอาศัยข้อมูลที่ได้จากการเก็บตัวอย่าง ในงานวิจัยนี้เลือกใช้ขั้นตอนวิธี EM (Expectation- Maxmization : EM) ที่ถูกพัฒนาขึ้นโดย Dempster, Laird, and Rubin [4] ในปี ค.ศ. 1977 มาใช้ในการประมาณพารามิเตอร์โดยมีวัตถุประสงค์เพื่อให้ฟังก์ชันควรจะเป็น (Likelihood function) มีค่าสูงสุด ขั้นตอนวิธี EM เป็นวิธีทำซ้ำ (Iterative method) ประกอบด้วย 2 ขั้นตอนย่อย นั่นคือขั้นตอน E (Expectation step) และขั้นตอน M (Maximization step) โดยขั้นตอนทั้งสองจะทำงานสลับกันเพื่อทำให้ขอบเขตล่าง (Lower bound) ของฟังก์ชันควรจะเป็นมีค่าสูงขึ้นในแต่ละการวนรอบ การทำซ้ำจะสิ้นสุดลงเมื่อทำงานครบตามจำนวนรอบที่กำหนดหรือเมื่อฟังก์ชันควรจะเป็นในรอบติดกันมีค่าใกล้เคียงกันมาก

เมื่อนำขั้นตอนวิธี EM มาประยุกต์ใช้ในการประมาณหาพารามิเตอร์ของแบบจำลอง SV ซึ่งบรรยายในรูปแบบปริภูมิสถานะโดยมีตัวแปรสถานะ x_k ที่ไม่สามารถวัดหรือสังเกตได้โดยตรง จึงจำเป็นต้องสังเคราะห์ตัวประมาณปรับเรียบ x_k^s ของตัวแปรสถานะที่ได้จากการแจกแจงปรับเรียบ (Smoothing distribution) โดยอาศัยข้อมูล y_k ที่รวบรวมมาได้หรือที่วัดได้จากกระบวนการเพื่อนำมาคำนวณหาค่าคาดหมายความควรจะเป็นแบบสมบูรณ์ (Complete likelihood) ในขั้นตอน E จากนั้นจึงนำค่าคาดหมายที่ได้มาคำนวณหาพารามิเตอร์เพื่อทำให้ฟังก์ชันควรจะเป็นมีค่า

สูงสุดในขั้นตอน M

การสังเคราะห์ตัวประมาณปรับเรียบ x_k^s มีหลายวิธี สำหรับกระบวนการเชิงเฟ้นสุ่มเชิงเส้นที่มีการแจกแจงของสัญญาณรบกวนเป็นแบบเกาส์เซียน (Gaussian distribution) **วิธีคาลมาน (Kalman method : KM)** ซึ่งเป็นขั้นตอนวิธีแบบข้างหน้า-ย้อนกลับ (Forward-backward algorithm) โดยมีการทำงานร่วมกันระหว่างตัวกรองคาลมาน (Kalman filter : KF) และตัวปรับเรียบคาลมาน (Kalman smoother : KS) เป็นวิธีที่นิยมเลือกใช้เนื่องจากมีขั้นตอนการคำนวณที่เรียบง่ายและสามารถเขียนบรรยาย x_k^s ในรูปแบบปิด (Closed form) ได้ด้วย ยิ่งไปกว่านั้นตัวประมาณปรับเรียบ x_k^s ที่สังเคราะห์ได้จะมีค่าใกล้เคียงกับตัวแปรสถานะจริง x_k ของกระบวนการมากที่สุด โดยให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยระหว่าง x_k^s และ x_k ต่ำสุดเมื่อเทียบกับวิธีสังเคราะห์ตัวประมาณด้วยวิธีเชิงเส้นชนิดอื่น ดังนั้นวิธี KM จึงจัดว่าเป็นวิธีประมาณค่าเหมาะที่สุด (Optimal estimation) สำหรับกระบวนการเชิงเฟ้นสุ่มเชิงเส้น [5]

ในกรณีของกระบวนการเชิงเฟ้นสุ่มเชิงเส้นที่มีการแจกแจงของสัญญาณรบกวนในรูปแบบอื่นที่ไม่ใช่การแจกแจงเกาส์เซียนหรือกระบวนการเชิงเฟ้นสุ่มไม่เชิงเส้นซึ่งทำให้ฟังก์ชันการแจกแจงของตัวแปรสถานะ x_k ไม่อยู่ในรูปแบบของการแจกแจงเกาส์เซียนแม้ว่าสัญญาณรบกวนภายนอกมีฟังก์ชันการแจกแจงเกาส์เซียน ในกรณีเช่นนี้วิธี KM ไม่จัดว่าเป็นวิธีประมาณค่าเหมาะที่สุด การใช้วิธี KM ในขั้นตอนวิธี EM จะนำไปสู่ความควรจะเป็นเสมือนสูงสุด (Quasi maximum likelihood : QML) เท่านั้น [6, 7, 8] ต่อมาได้มีการพัฒนาวิธี KM ให้สามารถรองรับกับกระบวนการเชิงเฟ้นสุ่มไม่เชิงเส้น ได้แก่วิธีคาลมานแบบขยาย (Extended Kalman method) [9, 10, 11, 12] วิธีคาลมานแบบแปลงจุด (Unscented Kalman method) [13, 14, 15] และวิธีคาลมานคิวเบเจอร์ (Cubature Kalman method) [16]

นอกเหนือจากวิธีการออกแบบโดยอาศัยหลักการพื้นฐานของคาลมานแล้ว **วิธีมอนติคาร์โล (Monte Carlo method : MCM)** เป็นอีกวิธีหนึ่งซึ่งเป็นที่ยอมรับแพร่หลายในการนำมาออกแบบตัวกรองและตัวปรับเรียบโดยอาศัยวิธีการแก้ปัญหาแบบศึกษาสำนึก (Heuristic approach) ด้วยการชักตัวอย่างจากฟังก์ชันการแจกแจงแล้วทำการประมาณปริมาณที่ต้องการทราบด้วยค่าเฉลี่ยทางสถิติแทนการคำนวณในรูปแบบปิดเช่นในวิธีของคาลมาน ด้วยคุณลักษณะเช่นนี้จึงทำให้การออกแบบด้วยวิธี MCM ถูกนำไปประยุกต์ใช้กับกระบวนการเชิงเฟ้นสุ่มได้อย่างหลากหลายโดยไม่มีข้อจำกัดของฟังก์ชันการแจกแจงของกระบวนการ [17, 18]

แนวคิดในการออกแบบด้วยวิธี MCM คือการสร้าง**ฟังก์ชันการแจกแจงสำคัญ** (Impor-

tant distribution) $q(\cdot)$ ที่มีลักษณะใกล้เคียงกับฟังก์ชันแจกแจงจริงของกระบวนการ $p(\cdot)$ แต่ง่ายต่อการชักตัวอย่าง จากนั้นจึงคำนวณหาค่าถ่วงน้ำหนักเพื่อทำหน้าที่ปรับตัวอย่างที่ได้จากการแจกแจง $q(\cdot)$ ให้มีค่าใกล้เคียงกับการแจกแจงจริง $p(\cdot)$ วิธีการชักตัวอย่างจาก $q(\cdot)$ ในลักษณะเช่นนี้เรียกว่า **การชักตัวอย่างสำคัญ** (Importance sampling : IS) ทั้งนี้ตัวอย่างที่ได้จากการชักตัวอย่างนิยมเรียกว่า **อนุภาค** (Particle)

เมื่อนำวิธี IS มาประยุกต์ใช้กับกระบวนการเชิงเฟ้นสุ่มจะทำให้เกิดเซตของอนุภาคตามแนวแกนเวลาซึ่งกระบวนการสร้างเซตของอนุภาคนี้เรียกว่า **การชักตัวอย่างสำคัญโดยลำดับ** (Sequential importance sampling : SIS) จากการศึกษาพบว่าวิธี SIS นำไปสู่ปัญหาใหม่ซึ่งเรียกว่า **ปัญหาการลดลง** (Degeneracy problem) กล่าวคือเมื่อเวลาผ่านไประยะหนึ่ง ค่าถ่วงน้ำหนักของอนุภาคส่วนใหญ่จะมีค่าลดลงจนมีค่าน้อยมากหรือเป็นศูนย์จึงส่งผลให้อนุภาคที่สมนัยกับค่าถ่วงน้ำหนักเหล่านั้นหายไปจากกระบวนการสร้างเซตของอนุภาคจึงทำให้วิธี SIS ถูกละเลยไปเป็นเวลาหลายปี

ต่อมาได้มีแนวคิดในการแก้ปัญหาการลดลงของวิธี SIS โดยอาศัยกระบวนการปรับอนุภาคใหม่ (Resampling process) ด้วยการกำจัดอนุภาคที่มีค่าถ่วงน้ำหนักน้อยมากออกไปแล้วนำอนุภาคที่มีค่าถ่วงน้ำหนักมากเข้าไปแทนที่ วิธี SIS ที่อาศัยกระบวนการปรับอนุภาคใหม่ร่วมด้วยนี้เรียกว่า การชักตัวอย่างสำคัญซ้ำโดยลำดับ (Sequential importance resampling : SIR) หรือส่วนใหญ่รู้จักกันในชื่อว่า **ตัวกรองอนุภาค** (Particle filter : PF) ผู้อ่านที่สนใจสามารถค้นคว้าเพิ่มเติมได้จาก [19, 20, 21]

มีงานวิจัยเป็นจำนวนมากที่มุ่งศึกษาแนวทางการออกแบบ PF โดยการปรับเปลี่ยนวิธีการชักตัวอย่าง วิธีการเลือกฟังก์ชันการแจกแจงสำคัญ วิธีการเลือกค่าถ่วงน้ำหนัก หรือวิธีการปรับอนุภาคใหม่เพื่อให้เกิดความเหมาะสมกับปัญหาแต่ละประเภท ยกตัวอย่างเช่น ในราวปี ค.ศ. 1993 Gordon, Salmond, and Smith [22] ได้นำเสนอวิธีการออกแบบ PF โดยนำการแจกแจงแบบมีเงื่อนไขของตัวแปรสถานะมาใช้เป็นฟังก์ชันการแจกแจงสำคัญและให้ค่าถ่วงน้ำหนักมาจากการแจกแจงแบบมีเงื่อนไขของข้อมูลที่วัดได้จากกระบวนการซึ่งทำให้มีการนำตัวกรองชนิดนี้ไปใช้งานได้ง่ายยิ่งขึ้น การออกแบบ PF ในวิธีนี้บ่อยครั้งนิยมเรียกว่า **ตัวกรองการปลุกเครื่อง** (Bootstrap filter : BF)

ต่อมา Kotecha and Djuric [23] ได้ตั้งข้อสังเกตว่ากระบวนการปรับอนุภาคใหม่ของตัวกรอง PF ใช้เวลาในการประมวลผลค่อนข้างนานเนื่องจากมีการทำงานแบบเรียงลำดับ (Se-

rial implementation) จึงมีแนวคิดในการพัฒนาตัวกรองชนิดใหม่เรียกว่า **ตัวกรองอนุภาคเกาส์เซียน** (Gaussian Particle Filter : GPF) ขึ้นในปี ค.ศ. 2003 เพื่อเพิ่มประสิทธิภาพการทำงานโดยลดเวลาในการประมวลผลลงด้วยการใช้วิธีการชักตัวอย่างจากฟังก์ชันการแจกแจงเกาส์เซียน ซึ่งมีค่ามัชฌิม (Mean) และค่าความแปรปรวน (Variance) คำนวณมาจากเซตของอนุภาคและค่าถ่วงน้ำหนักแทนวิธีการปรับอนุภาคใหม่ จากการสืบค้นพบว่า Kotecha and Djuric มุ่งพัฒนาตัวกรองเพียงอย่างเดียวและเท่าที่ทราบไม่ปรากฏว่ามีการนำ GPF มาใช้ในงานด้านการประมาณพารามิเตอร์มาก่อนหน้า

ในการออกแบบตัวปรับเรียบนั้น Kitagawa [24] ได้นำเสนอการออกแบบกระบวนการปรับเรียบด้วยวิธี MCM ผ่านทางวิธี SIS โดยการเก็บอนุภาคทุกตัวไว้แล้วทำการปรับอนุภาคทุกตัวด้วยกระบวนการปรับอนุภาคใหม่ อย่างไรก็ตามวิธีการของ Kitagawa ยังคงมีปัญหาการลดลง เพื่อช่วยลดปัญหาดังกล่าว Godsill, Doucet, and West [25] จึงได้พัฒนา ตัว**ปรับเรียบอนุภาคแบบย้อนกลับ** (Backward-simulation particle smoother : BS-PS) โดยอาศัยอนุภาคและค่าถ่วงน้ำหนักทุกตัวที่ได้จากกระบวนการกรองมาใช้ในการคำนวณด้วยวิธีการแบบเรียกซ้ำย้อนกลับ ซึ่งแนวคิดดังกล่าวคล้ายคลึงกับการออกแบบตัวปรับเรียบ KS นั่นเอง

จุดมุ่งหมายของการวิจัย

1. เพื่อเข้าใจหลักการทำงานของขั้นตอนวิธี EM ร่วมกับการสังเคราะห์ตัวแปรสถานะเพื่อนำมาใช้ประมาณพารามิเตอร์ของแบบจำลอง SV ได้

2. เพื่อศึกษาการประยุกต์ใช้ตัวกรอง BF และตัวกรอง GPF ร่วมกับขั้นตอนวิธี EM ในการประมาณหาพารามิเตอร์ของแบบจำลอง SV ได้

3. เพื่อสามารถประมาณพารามิเตอร์ของแบบจำลอง SV สำหรับข้อมูลของอัตราแลกเปลี่ยน USD/THB ได้

4. เพื่อสามารถนำความรู้จากการศึกษาไปใช้ในการประมาณแบบจำลองประเภทอื่นได้

5. เพื่อเป็นการเรียนรู้แบบบูรณาการทั้งการประมวลผลสัญญาณ การกรอง การปรับเรียบ เศรษฐมิติและสถิติ

6. เพื่อก่อให้เกิดองค์ความรู้ใหม่อันนำไปสู่การพัฒนาที่ยั่งยืน

ประโยชน์และความสำคัญ

1. เพื่อช่วยลดเวลาในการประมวลผลเมื่อเลือกใช้ตัวกรอง GPF แทนการใช้ตัวกรอง BF (หรือตัวกรอง PF) ในการประมาณพารามิเตอร์ของแบบจำลอง

2. เพื่อนำพารามิเตอร์ที่ประมาณได้มาประมาณหาความผันผวนผ่านทางแบบจำลอง SV แล้วนำไปใช้คำนวณหาค่า FV ของตราสารสิทธิแทนสูตรแบลค์-โชวส์แบบดั้งเดิม

3. นำความรู้ที่ได้จากงานวิจัยนี้ไปใช้ในประมาณพารามิเตอร์ของแบบจำลองที่บรรยายในรูปแบบอื่นได้

4. เป็นการบูรณาการความรู้ในหลายสาขาเพื่อให้เกิดองค์ความรู้ใหม่

5. เพื่อเป็นพื้นฐานในการทำวิจัยต่อยอดอันนำไปสู่การพัฒนาที่ยั่งยืน

ขอบเขตของการวิจัย

1. ใช้แบบจำลอง SV ที่พัฒนามาจากแบบจำลองของ Taylor

2. พิจารณาเฉพาะในกรณีของปริมาณสเกลาร์เท่านั้น

3. ใช้ขั้นตอนวิธี EM ร่วมกับการสังเคราะห์ตัวประมาณปรับเรียบ 2 วิธีนั่นคือ

   3.1 ตัวกรองการปลุกเครื่อง BF ร่วมกับตัวปรับเรียบอนุภาคแบบย้อนกลับ BS

   3.2 ตัวกรองอนุภาคเกาส์เซียน GPF ร่วมกับตัวปรับเรียบอนุภาคแบบย้อนกลับ BS

4. ใช้อัตราแลกเปลี่ยน USD/THB ตั้งแต่วันที่ 4 มกราคม พ.ศ. 2554 จนถึง 29 เมษายน พ.ศ. 2559 มาใช้เป็นกรณีศึกษา

บทที่ 2

ความรู้เบื้องต้น ขั้นตอนวิธีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีการประมาณเป็นสาขาหนึ่งของทฤษฎีการควบคุมและการประมวลผลสัญญาณ ซึ่งศึกษาเกี่ยวกับการประมาณพารามิเตอร์ของแบบจำลองในกระบวนการเชิงเฟ้นสุ่มโดยอาศัยข้อมูลจากการวัดหรือจากผลการทดลองที่อาจมีการปนเปื้อนจากสัญญาณรบกวนวิธีการประมาณที่ศึกษาในงานวิจัยนี้คือขั้นตอนวิธี EM ซึ่งนำมาใช้ในการประมาณค่าความผันผวนของอัตราแลกเปลี่ยนเงินตราระหว่างประเทศ

ความผันผวนและแบบจำลอง

ความผันผวน (Volatility) เริ่มเข้ามามีบทบาทสำคัญต่อการลงทุนโดยเฉพาะอย่างยิ่งสำหรับนักลงทุนและนักเก็งกำไรในตลาดการเงิน (Financial market) ซึ่งพิจารณาความผันผวนที่เกิดขึ้นคือความเสี่ยงหรือความไม่แน่นอนของราคาสินทรัพย์ในตลาด ในขณะที่นักวิเคราะห์ในตลาดอนุพันธ์ (Derivative market) ศึกษาและพยากรณ์ความผันผวนเพื่อใช้คำนวณหามูลค่าที่เหมาะสมสำหรับตราสารสิทธิ (Option)

แบบจำลองความผันผวนสามารถจำแนกออกเป็น 2 ประเภทใหญ่ [26] นั่นคือ

1. แบบจำลองความผันผวนเชิงกำหนด (Deterministic volatility model : DV)

แบบจำลองประเภทนี้พิจารณาค่าความผันผวนแบบมีเงื่อนไขเป็นฟังก์ชันของค่าที่สังเกตได้ก่อนหน้า ตัวอย่างของแบบจำลองประเภทนี้ ได้แก่ แบบจำลอง Autoregressive conditional heteroskedasticity (ARCH) ซึ่งถูกนำเสนอโดย Engle [27] ในปี ค.ศ. 1982 โดยมีแบบจำลองของผลตอบแทนดังนี้

$$r_k = m_k + \eta_k = m_k + \sigma_k \epsilon_k, \qquad \epsilon_k \sim \text{iid}\, \mathcal{N}(0,1) \tag{2.1}$$

โดยที่ m_k คือค่าเฉลี่ยของผลตอบแทนซึ่งโดยทั่วไปนิยมสมมติให้มีค่าเป็นศูนย์เพื่อความสะดวกต่อการวิเคราะห์และ σ_k^2 คือความแปรปรวนแบบมีเงื่อนไข (Conditional variance) ซึ่งสอดคล้องกับ

สมการดังนี้

$$\sigma_k^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \eta_{k-i}^2 \tag{2.2}$$

ภายใต้เงื่อนไข $q > 0$, $\alpha_0 > 0$ และ $\alpha_j \geq 0$ สำหรับ $j = 1, \ldots, q$ เพื่อเป็นการรับประกันได้ว่า $\sigma_k^2 > 0$ และค่าความผันผวนของสินทรัพย์ที่ได้จากการประมาณด้วยแบบจำลองนี้คือ σ_k

สัญกรณ์ $\mathcal{N}(\mu, \sigma^2)$ แทนฟังก์ชันการแจกแจงปกติที่มีค่ามัชฌิมเป็น μ และมีค่าความแปรปรวนเป็น σ^2 ซึ่งในสมการที่ (2.1) $\mu = 0$ และ $\sigma^2 = 1$ สำหรับ iid แทนคำว่า Independent and identically distributed ซึ่งหมายความว่าตัวแปรสุ่มภายใต้การพิจารณานั้นต้องมีฟังก์ชันการแจกแจงเหมือนกันทุกประการและต้องมีความเป็นอิสระเชิงสถิติซึ่งกันและกันด้วย

ต่อมาในปี ค.ศ. 1986 Bollerslev [28] ได้นำแบบจำลอง ARCH มาพัฒนาให้มีความเป็นนัยทั่วไปมากขึ้นเรียกว่าแบบจำลอง GARCH (Generalized ARCH) โดยยังคงมีแบบจำลองของผลตอบแทนดังเดิมหากแต่เปลี่ยนแบบจำลองของความแปรปรวนแบบมีเงื่อนไขในสมการที่ (2.2) มาเป็น

$$\sigma_k^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \eta_{k-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{k-j}^2 \tag{2.3}$$

จากนั้นได้มีการนำแบบจำลอง GARCH ไปพัฒนาเพื่อให้ได้แบบจำลองที่มีความเหมาะสมกับการประยุกต์ใช้งาน เช่น NGARCH, IGARCH, QGARCH, TGARCH, EGARCH, GARCH-M, GJR-GARCH และ fGARCH ดังเห็นได้จากในงานของ [29, 30, 31, 32, 33, 34]

จากการที่แบบจำลองในกลุ่มของ GARCH ได้ถูกนำไปประยุกต์ใช้อย่างกว้างขวางและก่อให้เกิดประโยชน์อย่างมากทางด้านการเงินและเศรษฐกิจจึงทำให้ Engle ซึ่งเป็นผู้ริเริ่มในการนำเสนอแบบจำลองในกลุ่มนี้ได้รับรางวัลโนเบลสาขาเศรษฐศาสตร์ในปี ค.ศ. 2003 [35]

2. **แบบจำลองความผันผวนเชิงเฟ้นสุ่ม** (Stochastic volatility model : SV)

แบบจำลองประเภทนี้พิจารณาค่าความผันผวนแบบมีเงื่อนไขเป็นเสมือนตัวแปรซ่อนของกระบวนการเชิงเฟ้นสุ่ม ตัวอย่างของแบบจำลองประเภทนี้ได้แก่ แบบจำลองเทย์เลอร์ แบบจำลองเฮสตัน แบบจำลองอัลฟา-เบตา-โร เชิงเฟ้นสุ่มและแบบจำลองเชน

แบบจำลอง SV ที่ถูกนำมาใช้ในงานด้านเศรษฐมิติเป็นแบบจำลองแรกคือแบบจำลองเทย์เลอร์ซึ่งถูกนำเสนอโดย Taylor [36] ในปี ค.ศ. 1982 โดยพิจารณาให้ผลตอบแทนถูกบรรยาย

โดย

$$r_k = \sigma_k \epsilon_k, \qquad \epsilon_k \sim \text{iid}\, \mathcal{N}(0,1) \tag{2.4}$$

ซึ่งมีลักษณะคล้ายกับแบบจำลอง ARCH ในสมการที่ (2.1) เมื่อกำหนดให้ $m_k = 0$ และให้ค่าลอการิทึมของความแปรปรวนแบบมีเงื่อนไข $\log \sigma_k^2$ เขียนอยู่ในรูปของแบบจำลองการถดถอยอัตตะแบบคงที่ (Stationary Autoregression) AR(1) นั่นคือ

$$\log \sigma_k^2 = \mu + \phi(\log \sigma_{k-1}^2 - \mu) + w_k, \qquad w_k \sim \text{iid}\, \mathcal{N}(0, Q) \tag{2.5}$$

โดยที่ $|\phi| < 1$ และ σ_k คือความผันผวนของสินทรัพย์ที่ต้องการประมาณ

หากกำหนดให้ $x_k = \log \sigma_k^2$ หรือนั่นคือ $\sigma_k = e^{x_k/2}$ ดังนั้นสมการที่ (2.4) และสมการที่ (2.5) จึงเขียนใหม่ได้เป็น

$$r_k = e^{x_k/2}\epsilon_k \quad \text{และ} \quad x_k = \mu + \phi(x_{k-1} - \mu) + w_k \tag{2.6}$$

โดยที่ x_k เรียกว่า ความผันผวนแบบลอการิทึม (Log volatility) ทำหน้าที่เป็นตัวแปรซ่อน (Hidden variables) ในแบบจำลองเนื่องจากไม่สามารถวัดค่าได้โดยตรง จากนิยามของ x_k ซึ่งเป็นค่าลอการิทึมของความผันผวน จึงเป็นการรับประกันได้ว่าค่าความแปรปรวนของกระบวนการเป็นบวกเสมอโดยไม่จำเป็นต้องมีเงื่อนไขอื่นมาประกอบ ทั้งนี้ x_k มีการแจกแจงปกติ ดังนั้นความแปรปรวนแบบมีเงื่อนไข σ_k^2 จึงมีการแจกแจงแบบลอการิทึมปกติ (Log-normal distribution)

จากแบบจำลองความผันผวนในสมการที่ (2.6) ด้านขวาพบว่ามีค่าคงที่ร่วมอยู่ด้วย มีงานวิจัยหลายชิ้นที่ปรับแบบจำลองเทย์เลอร์โดยการใส่ตัวประกอบมาตราส่วน (Scaling factor) β ทำหน้าที่เสมือนค่าความผันผวนที่เกิดขึ้นขณะหนึ่ง (Instantaneous volatility) เข้าไปในแบบจำลองผลตอบแทนในสมการที่ (2.6) ด้านซ้ายเพื่อทำการกำจัดค่าคงที่ดังกล่าวเช่นในงานวิจัยของ [37, 38, 39, 40, 41] ดังนั้นแบบจำลอง SV ใน (2.6) จึงเขียนใหม่ได้เป็น

$$\Sigma_{\text{NL}} := \begin{cases} x_k &= \phi x_{k-1} + w_k \\ r_k &= \beta e^{x_k/2}\epsilon_k \end{cases} \tag{2.7}$$

สังเกตว่า Σ_{NL} เป็นแบบจำลอง SV แบบไม่เชิงเส้นเนื่องจากมีฟังก์ชันเลขชี้กำลังร่วมด้วย หากนำสมการของผลตอบแทน r_k มายกกำลังสองแล้วนำมาใส่ฟังก์ชันลอการิทึมจะได้ว่า

$$\log r_k^2 = x_k + \log \beta^2 + \log \epsilon_k^2 \tag{2.8}$$

เนื่องจาก $\epsilon_k \sim \text{iid } \mathcal{N}(0,1)$ ดังนั้น $\log \epsilon_k^2$ จึงมีการแจกแจงไคกำลังสองแบบลอการิทึม (Log-Chi-squared distribution) เขียนแทนด้วย $\log \chi^2$ ซึ่งมีค่ามัชฌิมประมาณ -1.27 และมีความผันผวน $0.5\pi^2$ โดยทั่วไปนิยมสมมติให้ตัวแปรสุ่มมีค่ามัชฌิมเป็น 0 จึงกำหนดให้

$$v_k := \log \epsilon_k^2 - \mathbb{E}[\log \epsilon_k^2] \tag{2.9}$$

ดังนั้นสมการที่ (2.8) จึงเขียนใหม่ได้เป็น

$$\log r_k^2 = x_k + \log \beta^2 + \mathbb{E}[\log \epsilon_k^2] + v_k \tag{2.10}$$

ซึ่งอยู่ในรูปแบบของสมการเชิงเส้น ดังนั้นแบบจำลอง Σ_{NL} ในสมการที่ (2.7) จึงเขียนให้อยู่ในรูปของแบบจำลองเชิงเส้นได้ดังนี้

$$\Sigma_{\text{L}} := \begin{cases} x_k &= \phi x_{k-1} + w_k \\ y_k &= x_k + \alpha + v_k \end{cases} \tag{2.11}$$

เมื่อ $y_k = \log r_k^2$ และ $\alpha = \log \beta^2 + \mathbb{E}[\log \epsilon_k^2]$

แบบจำลองมาร์คอฟซ่อน

แบบจำลองซึ่งเป็นที่รู้จักอย่างแพร่หลายในทฤษฎีการควบคุมคือแบบจำลองที่บรรยายในรูปแบบปริภูมิสถานะ (State space model) แบบจำลองประเภทนี้มีความยืดหยุ่นสูง สามารถใช้ได้กับกระบวนการหลายชนิด ทั้งกระบวนการเชิงเส้นและไม่เชิงเส้น กระบวนการไม่แปรตามเวลาและแปรเปลี่ยนตามเวลา กระบวนการเชิงกำหนดและเชิงเฟ้นสุ่ม กระบวนการในเวลาต่อเนื่องและในเวลาวิยุต สำหรับแบบจำลองในรูปแบบปริภูมิสถานะของกระบวนการเชิงเฟ้นสุ่มไม่เชิงเส้น

ในเวลาวิยุตมีรูปแบบทั่วไปคือ

$$\Sigma := \begin{cases} x_k &= f(x_{k-1}, u_{k-1}, w_{k-1}; \theta_k) \\ y_k &= g(x_k, u_k, v_k; \theta_k) \end{cases} \tag{2.12}$$

โดยที่

- $x_k \in \mathbb{R}^{n_x}$ คือตัวแปรสถานะ (State variable) ซึ่งพิจารณาเป็นสัญญาณที่เกิดขึ้นภายในกระบวนการ ในงานด้านสถิติประยุกต์นิยมเรียกตัวแปรนี้ว่า ตัวแปรแฝง (Latent variable) หรือ ตัวแปรซ่อน (Hidden variable)
- $u_k \in \mathbb{R}^{n_u}$ คือสัญญาณเข้า (Input signal) ซึ่งเป็นสัญญาณเชิงกำหนด (Deterministic signal) ทำหน้าที่ควบคุมให้กระบวนการแสดงพฤติกรรมตามที่ผู้ออกแบบต้องการหรือเพื่อใช้เป็นสัญญาณอ้างอิงให้กับกระบวนการ
- $w_k \in \mathbb{R}^{n_w}$ คือสัญญาณรบกวนกระบวนการ (Process noise) ซึ่งเป็นสัญญาณเชิงเฟ้นสุ่มที่เข้ามารบกวนกระบวนการซึ่งอาจเกิดจากพฤติกรรมของตัวกระบวนการเองหรือเกิดจากความไม่แน่นอนที่เกิดขึ้นภายในกระบวนการนั้น
- $y_k \in \mathbb{R}^{n_y}$ คือสัญญาณออก (Output signal) ซึ่งอาจเป็นได้ทั้งผลตอบสนองของกระบวนการ (Process response) หรือเป็นสัญญาณที่วัดได้จากกระบวนการ (Measurement signal) เมื่อเกิดการกระตุ้นจากสัญญาณเข้าและ/หรือสัญญาณรบกวน
- $v_k \in \mathbb{R}^{n_v}$ คือสัญญาณรบกวนการวัด (Measurement noise) ซึ่งเป็นสัญญาณเชิงเฟ้นสุ่มที่เกิดจากความผิดพลาดในการวัดหรือเป็นสัญญาณรบกวนที่เกิดขึ้นในระหว่างการวัด
- θ_k คือพารามิเตอร์ของกระบวนการ

สังเกตว่าในแบบจำลอง Σ มีสัญญาณเชิงเฟ้นสุ่ม w_k และ v_k เข้ามาเกี่ยวข้องด้วย โดยทั่วไปนิยมบรรยายพฤติกรรมของสัญญาณเหล่านี้ด้วยฟังก์ชันการแจกแจง ดังนั้นแบบจำลอง Σ จึงสามารถบรรยายด้วย**แบบจำลองมาร์คอฟซ่อน**[1] (Hidden Markov Model : HMM) ผ่านทาง

[1]กระบวนการเชิงเฟ้นสุ่มที่มีคุณสมบัติมาร์คอฟถูกเรียกว่า **กระบวนการมาร์คอฟ** (Markov process) สำหรับแบบจำลองของกระบวนการมาร์คอฟที่มีตัวแปรซ่อนถูกเรียกว่า **แบบจำลองมาร์คอฟซ่อน** ซึ่งตั้งขึ้นเพื่อเป็นเกียรติแด่ Andrey Markov (ค.ศ. 1856 – ค.ศ. 1922) นักคณิตศาสตร์ชาวรัสเซียผู้ทุ่มเทศึกษากระบวนการเชิงเฟ้นสุ่ม

ฟังก์ชันการแจกแจงดังนี้

$$\Sigma_{\mathsf{HMM}} := \begin{cases} \mathsf{x}_k & \sim \quad p(\mathsf{x}_k|\mathsf{x}_{k-1}) \\ \mathsf{y}_k & \sim \quad p(\mathsf{y}_k|\mathsf{x}_k) \end{cases} \tag{2.13}$$

โดยที่

- $p(\mathsf{x}_k|\mathsf{x}_{k-1})$ คือแบบจำลองความน่าจะเป็นของตัวแปรสถานะซึ่งเป็นฟังก์ชันการแจกแจงแบบมีเงื่อนไขของตัวแปรสถานะ ณ เวลา k เมื่อทราบค่าของตัวแปรสถานะ ณ เวลา $k-1$
- $p(\mathsf{y}_k|\mathsf{x}_k)$ คือแบบจำลองความน่าจะเป็นของสัญญาณออกซึ่งเป็นฟังก์ชันการแจกแจงแบบมีเงื่อนไขของสัญญาณออก ณ เวลา k เมื่อทราบค่าของตัวแปรสถานะ ณ เวลา k

ทั้งนี้แบบจำลองมาร์คอฟซ่อนคือเป็นกระบวนการเชิงเฟ้นสุ่มที่มีคุณสมบัติดังต่อไปนี้

1. **คุณสมบัติมาร์คอฟสำหรับตัวแปรสถานะ** (Markov property of states) กล่าวว่าเซตของตัวแปรสถานะ $\{\mathsf{x}_k\}$ เมื่อ $k = 0, 1, 2, \ldots$ เป็นลำดับมาร์คอฟ (Markov sequence) **หรือห่วงโซ่มาร์คอฟ** (Markov chain) นั่นคือข้อมูลของกระบวนการที่จำเป็นต่อการคำนวณหาค่าของตัวแปรสถานะ ณ เวลา k ถูกรวบรวมไว้ภายในตัวแปรสถานะ ณ เวลา $k-1$ ด้วยเหตุนี้จึงไม่มีความจำเป็นที่ต้องเก็บหรือจดจำข้อมูลของกระบวนการในอดีตซึ่งเขียนในรูปของคณิตศาสตร์ได้ว่า

$$\boxed{p(\mathsf{x}_k|\mathsf{x}_{0:k-1}, \mathsf{y}_{1:k-1}) = p(\mathsf{x}_k|\mathsf{x}_{k-1}) \quad \text{เมื่อ } k = 1, 2, \ldots, N} \tag{2.14}$$

โดยสัญกรณ์ $\mathsf{z}_{a:b} = \{\mathsf{z}_a, \mathsf{z}_{a+1}, \ldots, \mathsf{z}_b\}$ คือลำดับของตัวแปร z ตั้งแต่เวลา $k = a$ จนถึงเวลา $k = b$ และกำหนดให้ $\mathsf{z}_{a:b} = 0$ ถ้า $b < a$

นอกจากนี้ข้อมูลของกระบวนการที่จำเป็นต่อการคำนวณหาค่าของตัวแปรสถานะ ณ เวลา $k-1$ ถูกรวบรวมไว้ภายในตัวแปรสถานะ ณ เวลา k ด้วยเหตุนี้จึงไม่มีความจำเป็นที่ต้องเก็บหรือจดจำข้อมูลของกระบวนการในอนาคตซึ่งเขียนในรูปของคณิตศาสตร์ได้ว่า

$$\boxed{p(\mathsf{x}_{k-1}|\mathsf{x}_{k:N}, \mathsf{y}_{k:N}) = p(\mathsf{x}_{k-1}|\mathsf{x}_k) \quad \text{เมื่อ } k = N, N-1, \ldots, 1} \tag{2.15}$$

2. **ความเป็นอิสระแบบมีเงื่อนไขของสัญญาณออก** (Conditional independence of outputs) กล่าวว่าค่าของตัวแปรสถานะในอดีตและสัญญาณออกในอดีตไม่ส่งผลต่อสัญญาณออก ณ เวลาปัจจุบัน y_k มีเพียงแต่ตัวแปรสถานะ ณ เวลาปัจจุบัน x_k เท่านั้นที่ส่งผลกระทบต่อ y_k ซึ่งเขียนเป็นสมการคณิตศาสตร์ได้ดังนี้

$$p(\mathsf{y}_k|\mathsf{x}_{0:k}, \mathsf{y}_{1:k-1}) = p(\mathsf{y}_k|\mathsf{x}_k) \quad \text{เมื่อ } k = 1, 2, \ldots, N \tag{2.16}$$

การอนุมานแบบเบย์

การอนุมานแบบเบย์มีความเชื่อมโยงกับการประมาณความควรจะเป็นสูงสุดดังนี้ ให้ $\mathsf{y}_{1:N} = \{\mathsf{y}_1, \ldots, \mathsf{y}_N\}$ คือลำดับของสัญญาณออกซึ่งพิจารณาให้เป็นตัวแปรสุ่มที่มีความเป็นอิสระเชิงสถิติและ θ คือพารามิเตอร์ที่ไม่ทราบค่า ให้ $p(\mathsf{y}_{1:N}|\theta)$ เป็นการแจกแจงร่วมแบบมีเงื่อนไขซึ่งพิจารณาให้เป็นฟังก์ชันของพารามิเตอร์ θ นั่นคือ

$$\mathcal{L}(\theta) = p(\mathsf{y}_{1:N}|\theta) = \prod_{k=1}^{N} p(\mathsf{y}_k|\theta) \tag{2.17}$$

แล้ว $\hat{\theta}$ ที่ทำให้ $\mathcal{L}(\theta)$ มีค่าสูงสุดคำนวณได้จาก

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}\, \mathcal{L}(\theta) \tag{2.18}$$

ทั้งนี้ฟังก์ชัน $\mathcal{L}(\theta)$ เรียกว่า **ฟังก์ชันควรจะเป็น** (Likelihood function) และวิธีการประมาณหา $\hat{\theta}$ เพื่อทำให้ $p(\mathsf{y}_{1:N}|\theta)$ มีค่าสูงสุดเรียกว่า **วิธีความควรจะเป็นสูงสุด** (Maximum likelihood : ML)

สำหรับการอนุมานแบบเบย์จะพิจารณาให้ θ เป็นตัวแปรสุ่มแล้วคำนวณหาการแจกแจงภายหลัง (Posterior distribution) โดยอาศัยกฎของเบย์ (Bayes' rule)

$$p(\theta|\mathsf{y}_{1:N}) = \frac{p(\mathsf{y}_{1:N}|\theta)p(\theta)}{p(\mathsf{y}_{1:N})} \tag{2.19}$$

โดยที่

◇ $p(\mathsf{y}_{1:N}|\theta)$ คือความควรจะเป็นของสัญญาณออกเมื่อกำหนดพารามิเตอร์ θ มาให้

◇ $p(\theta)$ คือการแจกแจงก่อน (Prior distribution) ซึ่งเป็นแบบจำลองข้อมูลของพารามิเตอร์ θ ก่อนการทดลองหรือก่อนการวัดสัญญาณออกของกระบวนการ

◇ $p(\mathsf{y}_{1:N})$ คือการแจกแจงร่วมของ $\mathsf{y}_{1:N}$ ซึ่งไม่ขึ้นกับพารามิเตอร์ θ

แนวคิดของการอนุมานแบบเบย์คือการเลือกตัวแปรสุ่ม θ ซึ่งทำให้การแจกแจงภายหลัง $p(\theta|\mathsf{y}_{1:N})$ มีค่าสูงสุด วิธีการนี้จึงเรียกว่า **วิธีความน่าจะเป็นภายหลังสูงสุด** (Maximum a posteriori : MAP) สังเกตว่าการประมาณแบบ ML คือการประมาณแบบ MAP เมื่อให้การแจกแจงก่อนของพารามิเตอร์ θ เป็นแบบเอกรูป (Uniform distribution) นั่นคือ $p(\theta) \propto 1$

พิจารณาแบบจำลอง HMM ในสมการที่ (2.13) ซึ่งมีตัวแปรสถานะ x_k เป็นตัวแปรซ่อน หากต้องการประมาณ $\hat{\mathsf{x}}_k$ โดยใช้การอนุมานแบบเบย์จึงแทน θ ในสมการที่ (2.19) ด้วย $\mathsf{x}_{0:N} = \{\mathsf{x}_0, \ldots, \mathsf{x}_N\}$ จึงได้การแจกแจงภายหลังของตัวแปรสถานะดังนี้

$$\boxed{p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}) = \frac{p(\mathsf{x}_{0:N})p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N})}{p(\mathsf{y}_{1:N})}} \tag{2.20}$$

ทั้งนี้โดยอาศัยคุณสมบัติของมาร์คอฟ การแจกแจงร่วมของตัวแปรสถานะ $\mathsf{x}_{0:N}$ และความควรจะเป็นของสัญญาณออก $\mathsf{y}_{1:N}$ จึงสามารถเขียนได้เป็น

$$p(\mathsf{x}_{0:N}) = p(\mathsf{x}_0)\prod_{k=1}^{N} p(\mathsf{x}_k|\mathsf{x}_{k-1}) \tag{2.21}$$

$$p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N}) = \prod_{k=1}^{N} p(\mathsf{y}_k|\mathsf{x}_k) \tag{2.22}$$

การคำนวณหาการแจกแจงภายหลังในสมการที่ (2.20) เป็นการคำนวณแบบต่อเนื่องเป็นชุด (Batch data processing) ซึ่งเป็นไปได้ยากในทางปฏิบัติเนื่องจากจำนวนของการคำนวณเพิ่มขึ้นตามจำนวนของข้อมูลที่ได้รับในแต่ละรอบ เมื่อมีข้อมูลใหม่เพิ่มเติมขึ้นจากการวัดสัญญาณออก y_{N+1} จำเป็นต้องคำนวณการแจกแจงใหม่ทั้งหมด เพื่อแก้ไขปัญหาดังกล่าวจึงนิยมเลือกใช้วิธีการคำนวณในลักษณะของการเรียกซ้ำ (Recursive data processing) ซึ่งมีจำนวนของการคำนวณคงที่ในแต่ละรอบ

สำหรับการคำนวณแบบเรียกซ้ำอาศัยการแจกแจงการกรอง การแจกแจงการทำนายและการแจกแจงปรับเรียบซึ่งมีนิยามดังนี้

1. การแจกแจงการกรอง (Filtering distribution) ซึ่งนำมาใช้ในขั้นตอนการกรองของตัวกรองแบบเบย์คือการแจกแจงภายหลังตามขอบ (Marginal posterior distribution) ของตัวแปรสถานะ ณ เวลาปัจจุบัน x_k เมื่อกำหนดสัญญาณออกตั้งแต่เริ่มต้นจนถึงเวลาปัจจุบัน $\mathsf{y}_{1:k}$ มาให้ นั่นคือ

$$p(\mathsf{x}_k|\mathsf{y}_{1:k}), \quad k = 1, \ldots, N \tag{2.23}$$

2. การแจกแจงการทำนาย (Prediction distribution) ซึ่งนำมาใช้ในขั้นตอนการทำนายของตัวกรองแบบเบย์คือการแจกแจงภายหลังตามขอบของตัวแปรสถานะในเวลาอนาคต x_{k+n} เมื่อกำหนดสัญญาณออกตั้งแต่เริ่มต้นจนถึงเวลาปัจจุบัน $\mathsf{y}_{1:k}$ มาให้และ n คือจำนวนขั้นของเวลา (time steps) โดยนับจากเวลาปัจจุบัน นั่นคือ

$$p(\mathsf{x}_{k+n}|\mathsf{y}_{1:k}), \quad k = 1, \ldots, N, \quad n = 1, 2, \ldots \tag{2.24}$$

3. การแจกแจงปรับเรียบ (Smoothing distribution) ซึ่งนำมาใช้ในตัวปรับเรียบแบบเบย์คือการแจกแจงภายหลังตามขอบของตัวแปรสถานะ ณ เวลาปัจจุบัน x_k เมื่อกำหนดสัญญาณออกทั้งหมดมาให้ นั่นคือ

$$p(\mathsf{x}_k|\mathsf{y}_{1:N}), \quad k = 1, \ldots, N \tag{2.25}$$

วิธีมอนติคาร์โลในกระบวนการเชิงเฟ้นสุ่ม

สิ่งหนึ่งที่จำเป็นต้องใช้ทั้งในขั้นตอนวิธีของตัวกรองและขั้นตอนวิธีของตัวปรับเรียบคือการหาค่าคาดหมายเมื่อกำหนดการแจกแจงแบบมีเงื่อนไขมาให้ เช่น

$$\mathbb{E}[g(\mathsf{x})|\mathsf{y}_{1:N}] = \int g(\mathsf{x})p(\mathsf{x}|\mathsf{y}_{1:N})\, d\mathsf{x} \tag{2.26}$$

เมื่อ $g : \mathbb{R}^{n_x} \to \mathbb{R}^{n_g}$ เป็นฟังก์ชันในตัวแปร $\mathsf{x} \in \mathbb{R}^{n_x}$ ซึ่งพบว่ามีเพียงบางกรณีเท่านั้นที่สามารถคำนวณค่าปริพันธ์ข้างต้นด้วยวิธีวิเคราะห์ได้ ด้วยเหตุนี้วิธีเชิงตัวเลขจึงเข้ามามีบทบาทสำคัญในการประมาณค่าปริพันธ์ดังกล่าว

วิธีมอนติคาร์โล (Monte Carlo method : MCM) เป็นวิธีเชิงตัวเลขประเภทหนึ่งที่สามารถใช้คำนวณหาค่าคาดหมายที่อยู่ในรูปแบบของปริพันธ์ดังในสมการที่ (2.26) ได้โดยอาศัยวิธีการชัก

ตัวอย่างจากฟังก์ชันการแจกแจงแล้วทำการประมาณหาปริมาณที่ต้องการทราบด้วยค่าเฉลี่ยทางสถิติแทนการคำนวณด้วยวิธีวิเคราะห์

ในทางปฏิบัติพบว่าการชักตัวอย่างจากฟังก์ชันการแจกแจง $p(\mathsf{x}|\mathsf{y}_{1:N})$ โดยตรงค่อนข้างยุ่งยากมากอันเนื่องจากโครงสร้างรูปแบบของฟังก์ชันซึ่งขึ้นกับพฤติกรรมของแต่ละกระบวนการจึงเกิดแนวคิดในการใช้ฟังก์ชันการแจกแจงอื่น $q(\mathsf{x}|\mathsf{y}_{1:N})$ ที่ง่ายต่อการชักตัวอย่างโดยเรียกฟังก์ชันการแจกแจงนั้นว่า **ฟังก์ชันการแจกแจงสำคัญ** (Importance distribution) ทั้งนี้ฟังก์ชันที่เลือกมาเป็นฟังก์ชันการแจกแจงสำคัญนั้นต้องอยู่ภายใต้เงื่อนไขสำคัญคือ $q(\cdot)$ ต้องไม่เป็นศูนย์เมื่อไรก็ตามที่ $p(\cdot)$ ไม่เป็นศูนย์หรือเขียนในรูปแบบคณิตศาสตร์ได้คือ

$$\boxed{\text{supp}\, q(\mathsf{x}|\mathsf{y}_{1:N}) \geq \text{supp}\, p(\mathsf{x}|\mathsf{y}_{1:N})} \tag{2.27}$$

1. การชักตัวอย่างสำคัญ

การชักตัวอย่างจากฟังก์ชันการแจกแจงสำคัญ $q(\cdot)$ เรียกว่า **การชักตัวอย่างสำคัญ** (Importance sampling : IS) ซึ่งอาศัยฟังก์ชันการแจกแจงสำคัญ $q(\cdot)$ ร่วมกับกฎของเบย์ในการเปลี่ยนค่าคาดหมายเมื่อเทียบกับฟังก์ชันการแจกแจง $p(\cdot)$ ในสมการที่ (2.26) ให้มาเป็นค่าคาดหมายเมื่อเทียบกับฟังก์ชันการแจกแจงสำคัญ $q(\cdot)$ ดังนี้

$$\begin{aligned}
\mathbb{E}^p[g(\mathsf{x})|\mathsf{y}_{1:N}] &= \int g(\mathsf{x})p(\mathsf{x}|\mathsf{y}_{1:N})\, d\mathsf{x} \quad = \int g(\mathsf{x}) \left[\frac{p(\mathsf{y}_{1:N}|\mathsf{x})p(\mathsf{x})}{p(\mathsf{y}_{1:N})}\right] d\mathsf{x} \\
&= \frac{\int g(\mathsf{x})p(\mathsf{y}_{1:N}|\mathsf{x})p(\mathsf{x})\, d\mathsf{x}}{\int p(\mathsf{y}_{1:N}|\mathsf{x})p(\mathsf{x})\, d\mathsf{x}} = \frac{\int \left[\frac{p(\mathsf{y}_{1:N}|\mathsf{x})p(\mathsf{x})}{q(\mathsf{x}|\mathsf{y}_{1:N})} g(\mathsf{x})\right] q(\mathsf{x}|\mathsf{y}_{1:N})\, d\mathsf{x}}{\int \left[\frac{p(\mathsf{y}_{1:N}|\mathsf{x})p(\mathsf{x})}{q(\mathsf{x}|\mathsf{y}_{1:N})}\right] q(\mathsf{x}|\mathsf{y}_{1:N})\, d\mathsf{x}} \\
&= \frac{\mathbb{E}^q\left[\frac{p(\mathsf{y}_{1:N}|\mathsf{x})p(\mathsf{x})}{q(\mathsf{x}|\mathsf{y}_{1:N})} g(\mathsf{x})\right]}{\mathbb{E}^q\left[\frac{p(\mathsf{y}_{1:N}|\mathsf{x})p(\mathsf{x})}{q(\mathsf{x}|\mathsf{y}_{1:N})}\right]}
\end{aligned} \tag{2.28}$$

ทำการประมาณ $\mathbb{E}^q[\,\cdot\,]$ โดยการชักตัวอย่างจากฟังก์ชันการแจกแจงสำคัญ $q(\mathsf{x}|\mathsf{y}_{1:N})$

$$\mathsf{x}^{(i)} \sim q(\mathsf{x}|\mathsf{y}_{1:N}), \quad \text{เมื่อ } i = 1, \ldots, n_f \tag{2.29}$$

ดังนั้นค่าคาดหมายในสมการที่ (2.28) จึงประมาณได้ดังนี้

$$\mathbb{E}^p[g(\mathsf{x})|\mathsf{y}_{1:N}] \approx \frac{\frac{1}{n_f}\sum_{i=1}^{n_f}\frac{p(\mathsf{y}_{1:N}|\mathsf{x}^{(i)})p(\mathsf{x}^{(i)})}{q(\mathsf{x}^{(i)}|\mathsf{y}_{1:N})}g(\mathsf{x}^{(i)})}{\frac{1}{n_f}\sum_{j=1}^{n_f}\frac{p(\mathsf{y}_{1:N}|\mathsf{x}^{(j)})p(\mathsf{x}^{(j)})}{q(\mathsf{x}^{(j)}|\mathsf{y}_{1:N})}}$$

$$= \sum_{i=1}^{n_f}\left[\frac{\frac{p(\mathsf{y}_{1:N}|\mathsf{x}^{(i)})p(\mathsf{x}^{(i)})}{q(\mathsf{x}^{(i)}|\mathsf{y}_{1:N})}}{\sum_{j=1}^{n_f}\frac{p(\mathsf{y}_{1:N}|\mathsf{x}^{(j)})p(\mathsf{x}^{(j)})}{q(\mathsf{x}^{(j)}|\mathsf{y}_{1:N})}}\right]g(\mathsf{x}^{(i)}) := \sum_{i=1}^{n_f} w^{(i)}g(\mathsf{x}^{(i)}) \qquad (2.30)$$

เมื่อ

$$w^{(i)} := \frac{\frac{p(\mathsf{y}_{1:N}|\mathsf{x}^{(i)})p(\mathsf{x}^{(i)})}{q(\mathsf{x}^{(i)}|\mathsf{y}_{1:N})}}{\sum_{j=1}^{n_f}\frac{p(\mathsf{y}_{1:N}|\mathsf{x}^{(j)})p(\mathsf{x}^{(j)})}{q(\mathsf{x}^{(j)}|\mathsf{y}_{1:N})}} \qquad (2.31)$$

คือค่าถ่วงน้ำหนักซึ่งทำหน้าที่ปรับตัวอย่างที่ได้จากฟังก์ชันการแจกแจงสำคัญ $q(\mathsf{x}|\mathsf{y}_{1:N})$ ให้มีค่าเข้าใกล้กับฟังก์ชันการแจกแจงจริง $p(\mathsf{x}|\mathsf{y}_{1:N})$ นั่นเอง อนึ่ง ในงานวิจัยด้านการออกแบบตัวกรองด้วยวิธี MCM นิยมเรียกตัวอย่างที่ได้จากการชักตัวอย่างสำคัญ $\mathsf{x}^{(i)}$ ว่า อนุภาค (Particles)

2. การชักตัวอย่างสำคัญโดยลำดับ

การชักตัวอย่างสำคัญเป็นการดำเนินการ ณ เวลาใดเวลาหนึ่ง แต่เมื่อนำมาประยุกต์กับกระบวนการเชิงเฟ้นสุ่มซึ่งเป็นกระบวนการพลวัตที่เปลี่ยนแปลงตามเวลาจึงต้องทำการชักตัวอย่างสำคัญตลอดช่วงเวลา ลักษณะดังกล่าวนี้เรียกว่า **การชักตัวอย่างสำคัญโดยลำดับ** (Sequential importance sampling : SIS) ซึ่งใช้ในการประมาณฟังก์ชันการแจกแจงการกรองของกระบวนการเชิงเฟ้นสุ่มที่มีแบบจำลอง HMM ดังในสมการที่ (2.13)

ขั้นตอนวิธี SIS จะอาศัยทั้งอนุภาคที่เกิดจากการชักตัวอย่างของฟังก์ชันการแจกแจงสำคัญและค่าถ่วงน้ำหนักในการประมาณการแจกแจงการกรอง $p(\mathsf{x}_k|\mathsf{y}_{1:k})$ เพื่อให้ค่าคาดหมายแบบมีเงื่อนไขของฟังก์ชัน $g(\cdot)$ พิจารณา ณ เวลา k สามารถประมาณได้จาก

$$\mathbb{E}[g(\mathsf{x}_k)|\mathsf{y}_{1:k}] \approx \sum_{i=1}^{n_f} w_k^{\mathsf{f}(i)}g(\mathsf{x}_k^{\mathsf{f}(i)}) \qquad (2.32)$$

ดังนั้นจึงอาจกล่าวได้ว่า ขั้นตอนวิธี SIS สามารถนำมาใช้ในการประมาณการแจกแจงการกรองได้ดังนี้

$$p(\mathrm{x}_k|\mathrm{y}_{1:k}) \approx \sum_{i=1}^{n_f} w_k^{\mathrm{f}(i)} \delta(\mathrm{x}_k - \mathrm{x}_k^{\mathrm{f}(i)}) \tag{2.33}$$

ในการเลือกฟังก์ชันการแจกแจงสำคัญ $q(\cdot)$ เพื่อนำมาใช้ในการประมาณการแจกแจงการกรองในแบบจำลอง HMM นิยมเลือก $q(\cdot)$ ที่มีคุณสมบัติของมาร์คอฟ กล่าวคือ

$$q(\mathrm{x}_k|\mathrm{x}_{0:k-1}^{\mathrm{f}(i)}, \mathrm{y}_{1:k}) = q(\mathrm{x}_k|\mathrm{x}_{k-1}^{\mathrm{f}(i)}, \mathrm{y}_{1:k}) \tag{2.34}$$

ด้วยคุณสมบัติของมาร์คอฟจึงทำให้ไม่จำเป็นต้องเก็บข้อมูลของอนุภาคทั้งหมดตั้งแต่เริ่มต้นจนถึงเวลาปัจจุบัน $\mathrm{x}_{0:k}^{\mathrm{f}(i)}$ มีเพียงเฉพาะอนุภาค ณ เวลาปัจจุบัน $\mathrm{x}_k^{\mathrm{f}(i)}$ เท่านั้นที่ใช้ในการคำนวณ

3. การปรับอนุภาคใหม่

ขั้นตอนวิธี SIS มีปัญหาการลดลง (Degenerary problem) นั่นคือเมื่อผ่านขั้นตอนการทำซ้ำไประยะหนึ่ง ค่าถ่วงน้ำหนักของอนุภาคส่วนใหญ่จะมีค่าเป็นศูนย์หรือใกล้ศูนย์จึงทำให้อนุภาคเหล่านั้นหายไปจากกระบวนการสร้างเซตของอนุภาค วิธีการหนึ่งที่ใช้แก้ไขปัญหาดังกล่าวคืออาศัยกระบวนการปรับอนุภาคใหม่ (Resampling process) โดยพิจารณค่าถ่วงน้ำหนักที่ได้เป็นเสมือนการแจกแจงวิยุตแล้วจึงทำการชักตัวอย่างจากการแจกแจงดังกล่าวเพื่อให้ได้อนุภาคชุดใหม่ จากนั้นจึงนำอนุภาคชุดใหม่นี้ไปแทนอนุภาคชุดเดิมที่มีค่าถ่วงน้ำหนักน้อย กระบวนนี้เป็นการกำจัดอนุภาคที่มีค่าถ่วงน้ำหนักน้อยออกไปแล้วนำอนุภาคที่มีค่าถ่วงน้ำหนักมากเข้ามาแทนที่

ตัวกรองอนุภาค ตัวกรองการปลุกเครื่องและตัวกรองอนุภาคเกาส์เซียน

หัวข้อนี้ศึกษาขั้นตอนวิธีของตัวกรองอนุภาค ตัวกรองการปลุกเครื่องและตัวกรองอนุภาคเกาส์เซียนโดยอาศัยหลักการที่ได้ศึกษาในหัวข้อที่ผ่านมาแล้วโดยมีรายละเอียดดังนี้

1. ตัวกรองอนุภาค

การชักตัวอย่างสำคัญซ้ำโดยลำดับ (Sequential importance resampling : SIR) หรือบ่อยครั้งนิยมเรียกว่า ตัวกรองอนุภาค (Particle filter : PF) คือขั้นตอนวิธี SIS ที่มีการ

ใช้กระบวนการปรับอนุภาคใหม่ร่วมด้วย ทั้งนี้เพื่อให้อนุภาคและค่าถ่วงน้ำหนักสามารถประมาณการแจกแจงของกระบวนการได้ใกล้เคียงความเป็นจริงมากที่สุดจึงไม่ควรเลือกใช้กระบวนการปรับอนุภาคใหม่ในทุกรอบของการทำซ้ำเพราะอาจทำให้สูญเสียอนุภาคบางตัวไป กระบวนการปรับอนุภาคใหม่นั้นจะถูกนำมาใช้ก็ต่อเมื่อค่าถ่วงน้ำหนักของอนุภาคส่วนใหญ่กลายเป็นศูนย์หรือมีค่าน้อยมากเท่านั้น อย่างไรก็ตามเพื่อความสะดวก นิยมกำหนดให้มีการปรับอนุภาคใหม่เมื่อผ่านขั้นตอนการทำซ้ำไปเป็นจำนวน N_r รอบ

สำหรับขั้นตอนวิธีในการออกแบบตัวกรอง PF นำมาสรุปไว้ดังต่อไปนี้

ขั้นตอนวิธีที่ 2.1 *ขั้นตอนวิธีของตัวกรองอนุภาค : PF*

1. *ทำตามขั้นตอนวิธี SIS*
2. *เข้ากระบวนการปรับอนุภาคใหม่เมื่อทำขั้นตอนการทำซ้ำครบ N_r รอบ*
3. *สังเคราะห์ตัวประมาณการกรองจาก* $\mathsf{x}_k' = \sum_{i=1}^{n_f} w_k^{\mathsf{f}(i)} \mathsf{x}_k^{\mathsf{f}(i)}$

2. ตัวกรองการปลุกเครื่อง

ในปี ค.ศ. 1993 Gordon et al. [22] ได้นำเสนอ **ตัวกรองการปลุกเครื่อง** (Bootstrap filter : BF) โดยใช้แบบจำลองความน่าจะเป็นของตัวแปรสถานะ $p(\mathsf{x}_k|\mathsf{x}_{k-1})$ เป็นฟังก์ชันการแจกแจงสำคัญและให้ค่าถ่วงน้ำหนักคำนวณจากแบบจำลองความน่าจะเป็นของสัญญาณออก $p(\mathsf{y}_k|\mathsf{x}_k)$ จึงทำให้สามารถนำไปใช้งานได้ง่ายขึ้น ทั้งนี้เพื่อให้ตัวกรอง BF มีประสิทธิภาพสูงจึงควรเลือกจำนวนอนุภาค n_f เป็นจำนวนมากและต้องมีการปรับอนุภาคใหม่ทุกครั้งในขั้นตอนการทำซ้ำ

สำหรับขั้นตอนวิธีในการออกแบบตัวกรอง BF นำมาสรุปไว้ดังต่อไปนี้

ขั้นตอนวิธีที่ 2.2 ***ขั้นตอนวิธีของตัวกรองการปลุกเครื่อง : BF***

ขั้นตอนเริ่มต้น *สุ่มเลือกอนุภาคเริ่มต้นจากฟังก์ชันการแจกแจงก่อน* $\mathsf{x}_0^{\mathsf{f}(i)} \sim p(\mathsf{x}_0)$ *และให้ค่าถ่วงน้ำหนักเริ่มต้นของแต่ละอนุภาคเป็น* $w_0^{\mathsf{f}(i)} = 1/n_f$ *เมื่อ* $i = 1, \ldots, n_f$

ขั้นตอนการทำซ้ำ เมื่อ $k = 1, \ldots, N$

1. *สุ่มเลือกอนุภาคตัวใหม่* $\mathsf{x}_k^{\mathsf{f}(i)}$ *จากแบบจำลองความน่าจะเป็นของตัวแปรสถานะ*

$$\mathsf{x}_k^{\mathsf{f}(i)} \sim p(\mathsf{x}_k|\mathsf{x}_{k-1}^{\mathsf{f}(i)}), \quad \text{เมื่อ } i = 1, \ldots, n_f \tag{2.35}$$

2. *คำนวณค่าถ่วงน้ำหนัก* $\tilde{w}_k^{f(i)}$ *จากแบบจำลองความน่าจะเป็นของสัญญาณออก*

$$\tilde{w}_k^{f(i)} \propto p(y_k|x_k^{f(i)}), \quad \text{เมื่อ } i = 1, \ldots, n_f \tag{2.36}$$

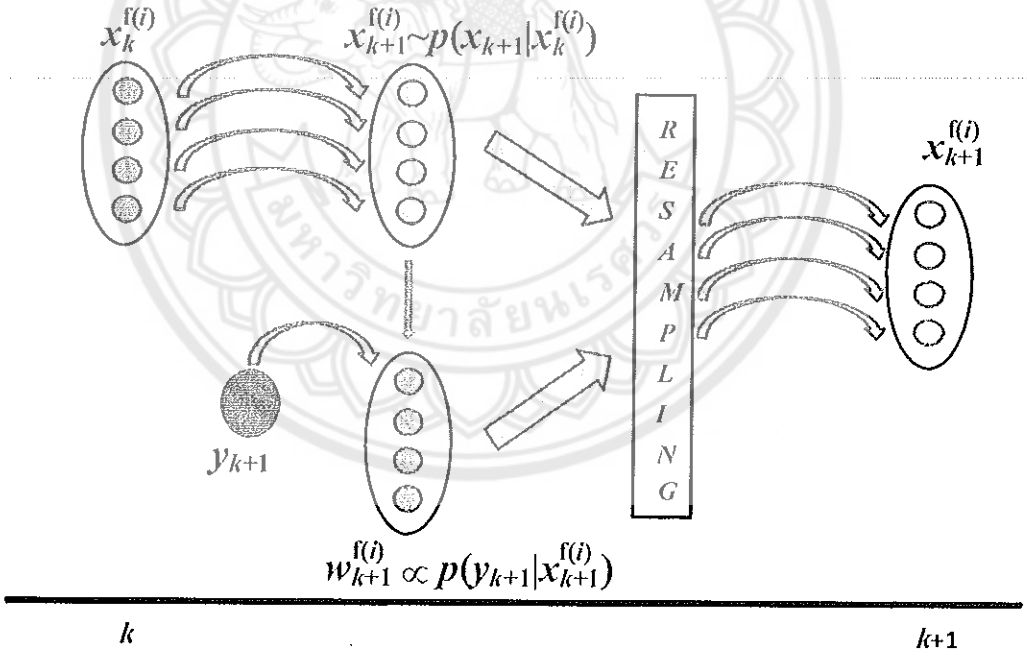
แล้วทำให้เป็นบรรทัดฐานเพื่อให้ผลรวมของค่าถ่วงน้ำหนักเป็น 1 นั่นคือ

$$w_k^{f(i)} = \frac{\tilde{w}_k^{f(i)}}{\sum_{i=1}^{n_f} \tilde{w}_k^{f(i)}} \tag{2.37}$$

3. *เข้ากระบวนการปรับอนุภาคใหม่*

4. *สังเคราะห์ตัวประมาณการกรองจาก* $x_k^f = \sum_{i=1}^{n_f} w_k^{f(i)} x_k^{f(i)}$

ขั้นตอนวิธีของตัวกรองการปลุกเครื่องนำมาแสดงไว้ในภาพ 1



ภาพ 1 ขั้นตอนวิธีของตัวกรองการปลุกเครื่อง : BF

3. ตัวกรองอนุภาคเกาส์เซียน

งานวิจัยของ Kotecha and Djuric [23] ได้ตั้งข้อสังเกตว่า กระบวนการปรับอนุภาคใหม่ในขั้นตอนวิธีของ PF และ BF มีการทำงานแบบเรียงลำดับ (Serial implementation) ซึ่งต้องใช้เวลาในการประมวลผลค่อนข้างนาน จึงมีแนวคิดในการพัฒนาตัวกรองชนิดใหม่ขึ้นในปี ค.ศ. 2003

เรียกว่า **ตัวกรองอนุภาคเกาส์เซียน** (Gaussian particle filter : GPF) ซึ่งเป็นการผสมผสานการทำงานระหว่างขั้นตอนวิธีของ KF และของ PF กล่าวคือ GPF ยังคงอาศัยขั้นตอนวิธี SIS โดยการสุ่มเลือกอนุภาคและคำนวณค่าถ่วงน้ำหนักเช่นเดียวกับใน PF แต่เปลี่ยนจากการใช้กระบวนการปรับอนุภาคใหม่ในการแก้ปัญหาการลดลงโดยเลือกใช้วิธีการชักตัวอย่างจากฟังก์ชันการแจกแจงเกาส์เซียนที่มีค่ามัชฌิมและและค่าความแปรปรวนคำนวณจากเซตของอนุภาคและค่าถ่วงน้ำหนักแทน วิธีการทำงานเช่นนี้คล้ายคลึงกับการทำงานของ KF ตรงที่มีการเลือกใช้ฟังก์ชันการแจกแจงเกาส์เซียนแทนการใช้ฟังก์ชันการแจกแจงจริงนั่นเอง

สำหรับขั้นตอนวิธีในการออกแบบตัวกรอง GPF นำมาสรุปไว้ดังต่อไปนี้

ขั้นตอนวิธีที่ 2.3 *ขั้นตอนวิธีของตัวกรองอนุภาคเกาส์เซียน : GPF*

ขั้นตอนเริ่มต้น *สุ่มเลือกอนุภาคเริ่มต้นจากฟังก์ชันการแจกแจงก่อน* $\mathsf{x}_0^{\mathsf{f}(i)} \sim p(\mathsf{x}_0)$ *และให้ค่าถ่วงน้ำหนักเริ่มต้นของแต่ละอนุภาคเป็น* $w_0^{\mathsf{f}(i)} = 1/n_f$ *เมื่อ* $i = 1, \ldots, n_f$

ขั้นตอนการทำซ้ำ *เมื่อ* $k = 1, \ldots, N$

1. *สุ่มเลือกอนุภาคตัวใหม่* $\mathsf{x}_k^{\mathsf{f}(i)}$ *จากแบบจำลองความน่าจะเป็นของตัวแปรสถานะ*

$$\mathsf{x}_k^{\mathsf{f}(i)} \sim p(\mathsf{x}_k|\mathsf{x}_{k-1}^{\mathsf{f}(i)}), \quad \text{เมื่อ } i = 1, \ldots, n_f \tag{2.38}$$

2. *คำนวณค่าถ่วงน้ำหนัก* $\tilde{w}_k^{\mathsf{f}(i)}$ *จากแบบจำลองความน่าจะเป็นของสัญญาณออก*

$$\tilde{w}_k^{\mathsf{f}(i)} \propto p(\mathsf{y}_k|\mathsf{x}_k^{\mathsf{f}(i)}), \quad \text{เมื่อ } i = 1, \ldots, n_f \tag{2.39}$$

แล้วทำให้เป็นบรรทัดฐานเพื่อให้ผลรวมของค่าถ่วงน้ำหนักเป็น 1 *นั่นคือ*

$$w_k^{\mathsf{f}(i)} = \frac{\tilde{w}_k^{\mathsf{f}(i)}}{\sum_{i=1}^{n_f} \tilde{w}_k^{\mathsf{f}(i)}} \tag{2.40}$$

3. *สุ่มเลือกอนุภาคใหม่โดยการชักตัวอย่างจากฟังก์ชันการแจกแจงเกาส์เซียน*

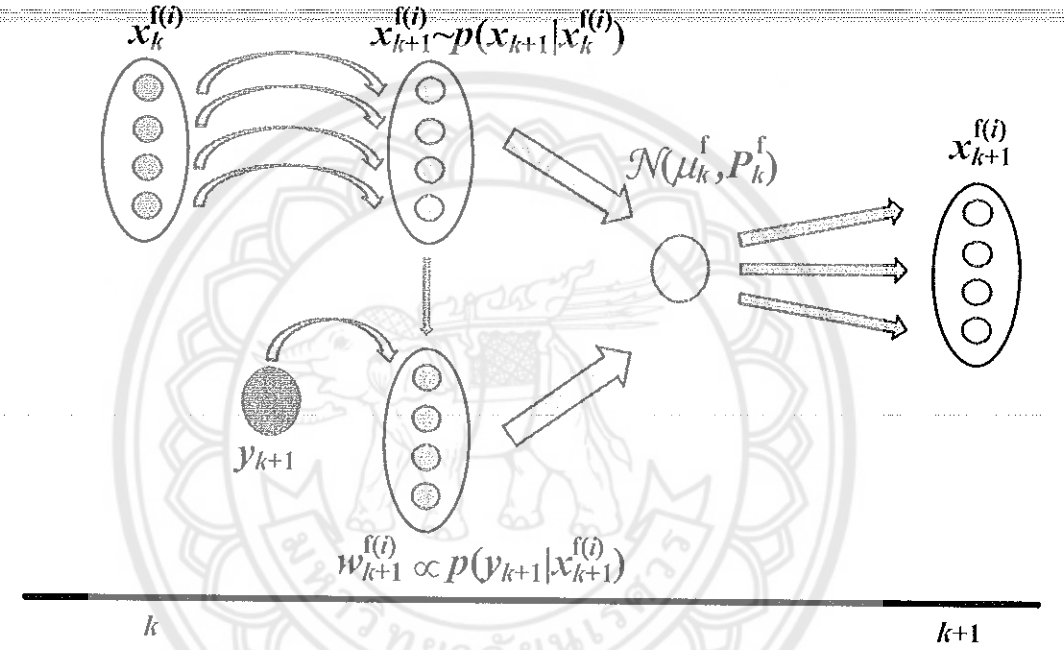
$$\mathsf{x}_k^{\mathsf{f}(i)} \sim \mathcal{N}(\mu_k^{\mathsf{f}}, P_k^{\mathsf{f}}), \quad \text{เมื่อ } i = 1, \ldots, n_f \tag{2.41}$$

เมื่อ

$$\mu_k^{\mathrm{f}} = \sum_{i=1}^{n_f} w_k^{\mathrm{f}(i)} \mathrm{x}_k^{\mathrm{f}(i)} \quad \text{และ} \quad P_k^{\mathrm{f}} = \sum_{i=1}^{n_f} w_k^{\mathrm{f}(i)} (\mathrm{x}_k^{\mathrm{f}(i)} - \mu_k^{\mathrm{f}})^2 \qquad (2.42)$$

4. สังเคราะห์ตัวประมาณการกรองจาก $\mathrm{x}_k^{\mathrm{f}} = \sum_{i=1}^{n_f} w_k^{\mathrm{f}(i)} \mathrm{x}_k^{\mathrm{f}(i)} = \mu_k^{\mathrm{f}}$

ขั้นตอนวิธีของตัวกรองอนุภาคเกาส์เซียนนำมาแสดงไว้ในภาพ 2



ภาพ 2 ขั้นตอนวิธีของตัวกรองอนุภาคเกาส์เซียน : GPF

ตัวปรับเรียบอนุภาคและตัวปรับเรียบอนุภาคแบบย้อนกลับ

หัวข้อนี้นำเสนอวิธีการออกแบบตัวประมาณปรับเรียบ $\mathrm{x}_k^{\mathrm{s}}$ โดยใช้วิธี MCM มาทำการประมาณฟังก์ชันการแจกแจงปรับเรียบ ในที่นี้จะนำเสนอ 2 วิธี นั่นคือ ขั้นตอนวิธีของตัวปรับเรียบอนุภาคแบบ SIR และขั้นตอนวิธีของตัวปรับเรียบอนุภาคแบบย้อนกลับโดยมีรายละเอียดดังนี้

1. **ตัวปรับเรียบอนุภาคแบบ SIR**

ในปี ค.ศ. 1996 Kitagawa [24] ได้นำเสนอวิธีการออกแบบ **ตัวปรับเรียบอนุภาคแบบ SIR** (SIR Particle smoother : SIR-PS) ซึ่งใช้หลักการเดียวกับการออกแบบตัวกรอง PF โดยสังเกตว่าขั้นตอนวิธีของ SIS จะประมาณการแจกแจงภายหลังทั้งหมดของ $\mathrm{x}_{0:k}^{\mathrm{f}(i)}$ แต่การทำงาน

ของตัวกรองจะเลือกเพียงตัวประมาณการกรอง ณ เวลาปัจจุบัน $\mathsf{x}_k^{f(i)}$ มาใช้โดยไม่พิจารณาเซตของอนุภาคที่เกิดขึ้นก่อนหน้า $\mathsf{x}_{0:k-1}^{f(i)}$ เนื่องจากในวิธีการกรองอาศัยเฉพาะการแจกแจงการกรองเท่านั้น เมื่อต้องการออกแบบตัวปรับเรียบจึงจำเป็นต้องเก็บข้อมูลย้อนหลังทั้งหมดที่ได้จากขั้นตอนวิธีของ SIS จากนั้นจึงนำเซตของอนุภาคทั้งหมดมาเข้ากระบวนการปรับอนุภาคใหม่ซึ่งต่างจากขั้นตอนวิธีของตัวกรอง PF ที่ใช้เฉพาะ $\mathsf{x}_k^{f(i)}$ มาปรับอนุภาคใหม่

2. ตัวปรับเรียบอนุภาคแบบย้อนกลับ

ในการออกแบบตัวปรับเรียบแบบ SIR-PS พบว่า ถ้า $N \gg k$ ยังคงมีปัญหาการลดลงแม้ว่าจะใช้กระบวนการปรับอนุภาคใหม่มาช่วย ดังนั้นในปี ค.ศ. 2004 Godsill et al. [25] จึงได้นำเสนอวิธีการแก้ปัญหาการลดลงโดยการใช้ทั้งเซตของอนุภาคและเซตของค่าถ่วงน้ำหนักทั้งหมดที่เกิดจากขั้นตอนวิธีการกรองมาคำนวณด้วยวิธีการแบบเรียกซ้ำย้อนกลับในลักษณะที่คล้ายคลึงกับการออกแบบตัวปรับเรียบ KS การออกแบบตัวปรับเรียบอนุภาคในลักษณะเช่นนี้เรียกว่า **ตัวปรับเรียบอนุภาคแบบย้อนกลับ** (Backward-simulation particle smoother : BS-PS)

ในการออกแบบตัวปรับเรียบ BS-PS ต้องเริ่มจากการใช้ขั้นตอนวิธีของตัวกรองเพื่อสร้างเซตของอนุภาคและค่าถ่วงน้ำหนัก $\{\mathsf{x}_k^{f(i)}, w_k^{f(i)}\}$ เมื่อ $k = 1, \ldots, N$ และ $i = 1, \ldots, n_f$ จากนั้นจึงเข้าสู่ขั้นตอนวิธีของตัวปรับเรียบ BS-PS ดังนี้

ขั้นตอนวิธีที่ 2.4 *ขั้นตอนวิธีของตัวปรับเรียบอนุภาคแบบย้อนกลับ : BS-PS*

ให้ $j = 1, \ldots, n_s$ เมื่อ n_s คือจำนวนแนววิถีของตัวประมาณปรับเรียบ $\mathsf{x}_{0:N}^{s(j)}$ ที่ต้องการ

ขั้นตอนเริ่มต้น *พิจารณาค่าถ่วงน้ำหนัก $w_N^{f(i)}$ เป็นความน่าจะเป็นในการเลือกอนุภาคการกรองที่ตำแหน่ง i ณ เวลาสุดท้ายให้มาเป็นอนุภาคปรับเรียบในแนววิถี j ณ เวลาสุดท้าย นั่นคือ*

$$\mathsf{x}_N^{s(j)} = \mathsf{x}_N^{f(i)} \quad \text{ด้วยความน่าจะเป็น } w_N^{f(i)} \tag{2.43}$$

ขั้นตอนการทำซ้ำ เมื่อ $k = N-1, \ldots, 0$

1. คำนวณค่าถ่วงน้ำหนักใหม่ $w_k^{s(i)}$ จาก

$$w_k^{s(i)} \propto w_k^{f(i)} p(\mathsf{x}_{k+1}^{s(j)} | \mathsf{x}_k^{f(i)}), \quad \text{เมื่อ } i = 1, \ldots, n_f \tag{2.44}$$

วิ ทส
วิทร
อ๒๖๖
ร๕๖๐
1038685

2. *พิจารณาค่าถ่วงน้ำหนัก $w_k^{s(i)}$ เป็นความน่าจะเป็นในการเลือกอนุภาคการกรอง ที่ตำแหน่ง i ณ เวลา k มาเป็นอนุภาคปรับเรียบในแนววิถี j ณ เวลา k นั่นคือ*

$$x_k^{s(j)} = x_k^{f(i)} \quad \text{ด้วยความน่าจะเป็น } w_k^{s(i)} \tag{2.45}$$

สังเกตว่าในแต่ละรอบของ j จะได้แนววิถีของตัวประมาณปรับเรียบมา 1 แนว นั่นคือ $x_{0:N}^{s(j)}$ ซึ่งเมื่อทำจนครบจำนวนรอบจะได้แนววิถีของตัวประมาณปรับเรียบมาทั้งหมด n_s แนว จากนั้นจึงทำการประมาณการแจกแจงปรับเรียบจาก

$$p(x_{0:N}|y_{1:N}) \approx \frac{1}{n_s}\sum_{j=1}^{n_s}\delta(x_{0:N} - x_{0:N}^{s(j)}) \tag{2.46}$$

และการแจกแจงปรับเรียบ ณ เวลา k เมื่อกำหนดข้อมูลทั้งหมดมาให้ประมาณได้จาก

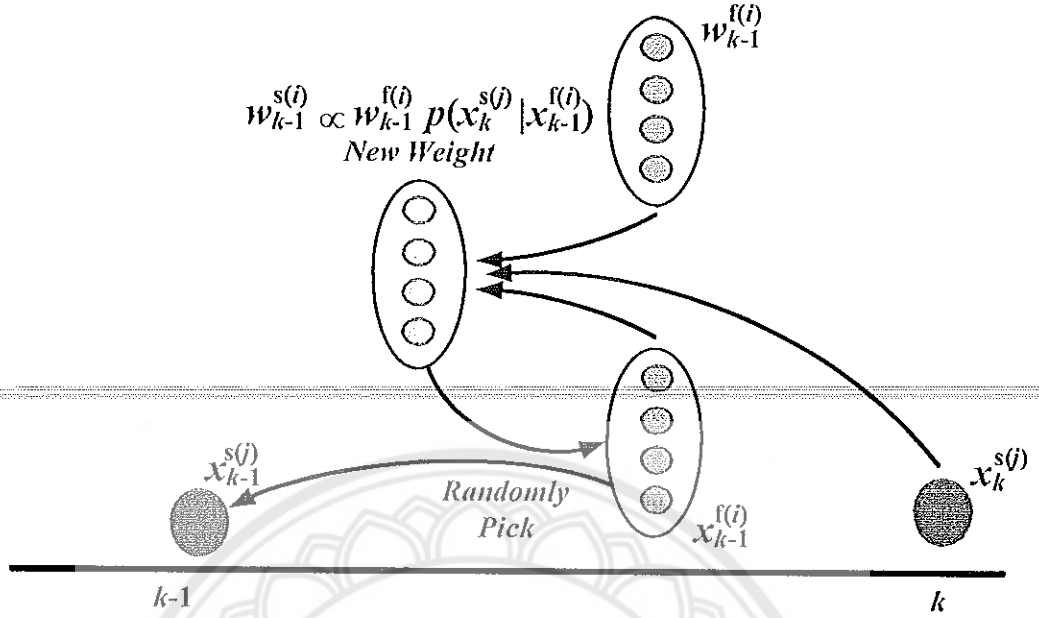
$$p(x_k|y_{1:N}) \approx \frac{1}{n_s}\sum_{j=1}^{n_s}\delta(x_k - x_k^{s(j)}) \tag{2.47}$$

เมื่อ $x_k^{s(j)}$ คือเซตของอนุภาคที่เกิดขึ้น ณ เวลา k ของ $x_{0:N}^{s(j)}$

จากการแจกแจงปรับเรียบที่ประมาณได้ข้างต้นจึงนำมาสังเคราะห์ตัวประมาณปรับเรียบได้ดังนี้ $x_k^s = \frac{1}{n_s}\sum_{j=1}^{n_s} x_k^{s(j)}$ สำหรับขั้นตอนวิธีของตัวปรับเรียบอนุภาคแบบย้อนกลับนำมาแสดงไว้ในภาพ 3

การประมาณพารามิเตอร์และขั้นตอนวิธี EM

ในหัวข้อที่ผ่านมาเป็นการศึกษาวิธีการประมาณการแจกแจงการกรองและการแจกแจงปรับเรียบเพื่อให้ได้มาซึ่งตัวประมาณการกรอง x_k^f และตัวประมาณปรับเรียบ x_k^s ตามลำดับ ภายใต้สมมติฐานที่ว่าผู้ออกแบบทราบพารามิเตอร์ θ ทุกตัวของแบบจำลอง อย่างไรก็ตามสิ่งที่ผู้ออกแบบทราบในการวิเคราะห์อนุกรมเวลาคือเซตของข้อมูลซึ่งเป็นสัญญาณออกที่วัดได้จากกระบวนการเท่านั้น ด้วยเหตุนี้จึงจำเป็นต้องประมาณพารามิเตอร์ $\hat{\theta}$ ของแบบจำลองขึ้นมาเป็นอันดับแรกโดยอาศัยข้อมูลที่ได้จากการวัดแล้วจึงทำการสังเคราะห์ตัวประมาณการกรองและ/หรือตัวประมาณ



ภาพ 3 ขั้นตอนวิธีของตัวปรับเรียบอนุภาคแบบย้อนกลับ : BS-PS

ปรับเรียบในลำดับถัดไป

หลักการประมาณพารามิเตอร์มีหลายวิธี สำหรับในงานวิจัยนี้เลือกใช้วิธีความควรจะเป็นสูงสุด (ML) ในการประมาณ $\hat{\theta}$ อย่างไรก็ตามการคำนวณด้วยวิธีวิเคราะห์ผ่านทางการหาปริพันธ์ตามกฎของเบย์นั้นมีความยุ่งยากซับซ้อนโดยเฉพาะอย่างยิ่งเมื่อมีการเก็บข้อมูลมาเป็นจำนวนมาก ดังนั้นวิธีเชิงตัวเลขจึงถูกนำมาใช้โดยในที่นี้เลือกใช้ขั้นตอนวิธี EM เนื่องจากเป็นวิธีทำซ้ำที่เรียบง่าย สะดวกต่อการนำไปประยุกต์ใช้งานและลู่เข้าสู่ผลเฉลยไว สำหรับเนื้อหาในบทนี้จึงมุ่งเน้นศึกษาหลักการประมาณพารามิเตอร์พร้อมนำเสนอขั้นตอนวิธี EM โดยมีรายละเอียดดังนี้

1. การประมาณพารามิเตอร์แบบเบย์

พิจารณาแบบจำลอง HMM ของกระบวนการเชิงเฟ้นสุ่มโดยมี θ เป็นพารามิเตอร์ที่ไม่ทราบค่าดังนี้

$$\Sigma_{\mathrm{HMM};\theta} := \begin{cases} \mathsf{x}_k & \sim \quad p(\mathsf{x}_k|\mathsf{x}_{k-1},\theta) \\ \mathsf{y}_k & \sim \quad p(\mathsf{y}_k|\mathsf{x}_k,\theta) \end{cases} \tag{2.48}$$

โดยให้การแจกแจงความน่าจะเป็นก่อนของพารามิเตอร์ θ และของตัวแปรสถานะเริ่มต้น x_0 คือ

$$\theta \sim p(\theta) \qquad \text{และ} \qquad \mathsf{x}_0 \sim p(\mathsf{x}_0|\theta) \tag{2.49}$$

หัวใจหลักของการประมาณคือการคำนวณหาการแจกแจงภายหลังเมื่อกำหนดเซตของข้อมูลมาให้ซึ่งคำนวณได้จากกฎของเบย์ดังนี้

$$p(\mathsf{x}_{0:N}, \boldsymbol{\theta}|\mathsf{y}_{1:N}) = \frac{p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N}, \boldsymbol{\theta})p(\mathsf{x}_{0:N}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathsf{y}_{1:N})} \tag{2.50}$$

โดยอาศัยคุณสมบัติมาร์คอฟจะได้ว่า

$$p(\mathsf{x}_{0:N}|\boldsymbol{\theta}) = p(\mathsf{x}_0|\boldsymbol{\theta}) \prod_{k=1}^{N} p(\mathsf{x}_k|\mathsf{x}_{k-1}, \boldsymbol{\theta}) \tag{2.51}$$

$$p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N}, \boldsymbol{\theta}) = \prod_{k=1}^{N} p(\mathsf{y}_k|\mathsf{x}_k, \boldsymbol{\theta}) \tag{2.52}$$

ดังนั้นการแจกแจงภายหลังในสมการที่ (2.50) จึงเขียนใหม่ได้เป็น

$$p(\mathsf{x}_{0:N}, \boldsymbol{\theta}|\mathsf{y}_{1:N}) = \frac{p(\boldsymbol{\theta})p(\mathsf{x}_0|\boldsymbol{\theta}) \prod_{k=1}^{N} p(\mathsf{x}_k|\mathsf{x}_{k-1}, \boldsymbol{\theta}) \prod_{k=1}^{N} p(\mathsf{y}_k|\mathsf{x}_k, \boldsymbol{\theta})}{p(\mathsf{y}_{1:N})} \tag{2.53}$$

แต่เนื่องจาก $p(\mathsf{y}_{1:N})$ เป็นการแจกแจงของข้อมูลที่ทราบค่า ดังนั้นความสัมพันธ์ข้างต้นจึงเขียนอยู่ในรูปของสัดส่วนตรงได้ดังนี้

$$p(\mathsf{x}_{0:N}, \boldsymbol{\theta}|\mathsf{y}_{1:N}) \propto p(\boldsymbol{\theta})p(\mathsf{x}_0|\boldsymbol{\theta}) \prod_{k=1}^{N} p(\mathsf{x}_k|\mathsf{x}_{k-1}, \boldsymbol{\theta}) \prod_{k=1}^{N} p(\mathsf{y}_k|\mathsf{x}_k, \boldsymbol{\theta}) \tag{2.54}$$

หากต้องการทราบเฉพาะการแจกแจงของ $\boldsymbol{\theta}$ สามารถทำได้โดยการหาการแจกแจงภายหลังตามขอบต่อไปนี้

$$p(\boldsymbol{\theta}|\mathsf{y}_{1:N}) = \int p(\mathsf{x}_{0:N}, \boldsymbol{\theta}|\mathsf{y}_{1:N})\, d\mathsf{x}_{0:N} \tag{2.55}$$

สังเกตว่าการหาปริพันธ์ข้างต้นทำได้ยากโดยเฉพาะอย่างยิ่งเมื่อมีการเก็บข้อมูลมากขึ้น ดังนั้นเพื่อเป็นการหลีกเลี่ยงวิธีการหาปริพันธ์โดยตรง จึงเลือกใช้วิธีการเชิงตัวเลขในการประมาณการแจกแจงภายหลังตามขอบของพารามิเตอร์เช่นเดียวกับวิธีการกรองดังในหัวข้อที่ผ่านมาโดยอาศัย

ความสัมพันธ์ที่ว่า

$$p(\boldsymbol{\theta}|\mathsf{y}_{1:N}) \propto p(\mathsf{y}_{1:N}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{2.56}$$

จากนั้นจึงใช้ขั้นตอนวิธีแบบเรียกซ้ำเพื่อหาค่าของ $p(\boldsymbol{\theta}|\mathsf{y}_{1:N})$ โดยตรงโดยไม่จำเป็นต้องคำนวณหาการแจกแจงภายหลังร่วมดังเช่นในสมการที่ (2.54)

จากความสัมพันธ์ใน (2.56) พบว่า ถ้าทราบความควรจะเป็นตามขอบของพารามิเตอร์ $p(\mathsf{y}_{1:N}|\boldsymbol{\theta})$ และการแจกแจงก่อน $p(\boldsymbol{\theta})$ จะทำให้ทราบการแจกแจงภายหลังตามขอบ $p(\boldsymbol{\theta}|\mathsf{y}_{1:N})$ ได้ โดยทั่วไปผู้ออกแบบมักนิยมเลือกการแจกแจงก่อน $p(\boldsymbol{\theta})$ ให้ง่ายต่อการคำนวณ สำหรับความควรจะเป็นตามขอบ $p(\mathsf{y}_{1:N}|\boldsymbol{\theta})$ สามารถแยกตัวประกอบได้เป็น

$$p(\mathsf{y}_{1:N}|\boldsymbol{\theta}) = \prod_{k=1}^{N} p(\mathsf{y}_k|\mathsf{y}_{1:k-1}, \boldsymbol{\theta}) \tag{2.57}$$

สำหรับแต่ละตัวประกอบคำนวณจาก

$$p(\mathsf{y}_k|\mathsf{y}_{1:k-1}, \boldsymbol{\theta}) = \int p(\mathsf{y}_k|\mathsf{x}_k, \boldsymbol{\theta}) p(\mathsf{x}_k|\mathsf{y}_{1:k-1}, \boldsymbol{\theta})\, d\mathsf{x}_k \tag{2.58}$$

โดยในที่นี้กำหนดให้ $p(\mathsf{y}_1|\mathsf{y}_{1:0}, \boldsymbol{\theta}) := p(\mathsf{y}_1|\boldsymbol{\theta})$

สังเกตว่า $p(\mathsf{y}_k|\mathsf{x}_k, \boldsymbol{\theta})$ คือแบบจำลองความน่าจะเป็นของสัญญาณออกซึ่งคำนวณได้จาก $\Sigma_{\mathsf{HMM};\boldsymbol{\theta}}$ ในสมการที่ (2.48) สำหรับ $p(\mathsf{x}_k|\mathsf{y}_{1:k-1}, \boldsymbol{\theta})$ คือการแจกแจงการทำนายของตัวแปรสถานะซึ่งคำนวณโดยใช้วิธีการทำซ้ำ

เมื่อคำนวณหาความควรจะเป็นตามขอบของพารามิเตอร์ $p(\mathsf{y}_{1:N}|\boldsymbol{\theta})$ เป็นที่เรียบร้อยแล้ว จึงนำมาแทนลงในความสัมพันธ์ที่ (2.56) เพื่อทำการประมาณการแจกแจงภายหลังตามขอบของพารามิเตอร์ $p(\boldsymbol{\theta}|\mathsf{y}_{1:N})$ เป็นลำดับถัดไป จากนั้นจึงเข้าสู่กระบวนการประมาณพารามิเตอร์ซึ่งมีอยู่หลายวิธี วิธี MAP เป็นวิธีการหนึ่งที่ใช้ประมาณหาพารามิเตอร์ที่ทำให้การแจกแจงภายหลังตามขอบเกิดค่าสูงสุด นั่นคือ

$$\boxed{\hat{\boldsymbol{\theta}}^{\mathsf{MAP}} = \underset{\boldsymbol{\theta}}{\mathsf{argmax}}[p(\boldsymbol{\theta}|\mathsf{y}_{1:N})] = \underset{\boldsymbol{\theta}}{\mathsf{argmax}}[p(\mathsf{y}_{1:N}|\boldsymbol{\theta})p(\boldsymbol{\theta})]} \tag{2.59}$$

ในกรณีที่เลือกการแจกแจงก่อนของพารามิเตอร์เป็นแบบเอกรูป $p(\theta) \propto 1$ การประมาณแบบ MAP จะลดรูปลงเหลือเป็นการประมาณแบบ ML กล่าวคือเป็นกระบวนการหาพารามิเตอร์ที่ทำให้ความควรจะเป็นของพารามิเตอร์เกิดค่าสูงสุด

$$\boxed{\hat{\theta}^{\mathrm{ML}} = \underset{\theta}{\operatorname{argmax}}[p(\mathbf{y}_{1:N}|\theta)]} \tag{2.60}$$

เพื่อให้สะดวกต่อการคำนวณหา $\hat{\theta}^{\mathrm{ML}}$ จึงนิยมพิจารณาความควรจะเป็นแบบลอการิทึม ซึ่งอยู่ในรูปของผลบวกแทนความควรจะเป็นในสมการที่ (2.57) ซึ่งอยู่ในรูปของผลคูณ นั่นคือ

$$L(\theta) = \log p(\mathbf{y}_{1:N}|\theta) = \sum_{k=1}^{N} \log p(\mathbf{y}_k|\mathbf{y}_{1:k-1}, \theta) \tag{2.61}$$

ทั้งนี้ $\hat{\theta}^{\mathrm{ML}}$ ที่คำนวณจากความควรจะเป็นในสมการที่ (2.57) และ (2.61) เป็นค่าเดียวกันเนื่องจากคุณสมบัติของฟังก์ชันลอการิทึมซึ่งเป็นฟังก์ชันเพิ่มทางเดียว (Monotonically increasing function) นั่นเอง นอกจากนี้ฟังก์ชันลอการิทึมจัดว่าเป็น ฟังก์ชันเว้า (Concave function) จึงสามารถหาค่าสูงสุดได้

2. ขั้นตอนวิธี EM

ขั้นตอนในการคำนวณหา $\hat{\theta}^{\mathrm{MAP}}$ ในสมการที่ (2.59) หรือ $\hat{\theta}^{\mathrm{ML}}$ ในสมการที่ (2.60) โดยตรงผ่านทางวิธีการหาค่าเหมาะที่สุด (Optimization method) มีหลายวิธี ในหลายกรณีที่วิธีการเหล่านั้นไม่สามารถนำมาประยุกต์ได้อันอาจเกิดจากความยุ่งยากของฟังก์ชันการแจกแจงภายหลัง ในปี ค.ศ. 1977 Dempster et al. [4] จึงได้พัฒนาขั้นตอนวิธี EM (Expectation-Maximization algorithm : EM) เพื่อใช้ในการประมาณหาค่าพารามิเตอร์ที่ทำให้เกิดความควรจะเป็นสูงสุดโดยอาศัยวิธีการทำซ้ำแทนการคำนวณเชิงวิเคราะห์ ต่อมาขั้นตอนวิธี EM ได้ถูกนำไปประยุกต์ใช้อย่างกว้างขวางดังเห็นได้จาก [42, 43, 44, 45] รวมทั้งเอกสารอ้างอิงภายใน

หลักการพื้นฐานของขั้นตอนวิธี EM คือการคำนวณหา**ขอบเขตล่าง** (Lower bound) ของ $L(\theta)$ แทนการคำนวณหา $L(\theta)$ โดยตรง จากนั้นจึงพยายามทำให้ขอบเขตล่างของ $L(\theta)$ มีค่ามากขึ้นในแต่ละรอบของการทำซ้ำจึงส่งผลให้ $L(\theta)$ มีค่ามากขึ้นด้วยเช่นกัน โดยทั่วไปค่าสูงสุดที่ได้จากขั้นตอนวิธี EM จะเป็นค่าสูงสุดเฉพาะที่ (Local maxima) ดังนั้นการกำหนดค่าเริ่มต้นของ

การทำซ้ำจึงมีบทบาทสำคัญต่อการลู่เข้า สำหรับวิธีการคำนวณหาขอบเขตล่างของ $L(\theta)$ มีดังนี้

ให้ θ แทนพารามิเตอร์ใด ๆ และให้ $\theta^{(\kappa)}$ แทนพารามิเตอร์ที่ประมาณได้ในแต่ละรอบของการทำซ้ำแล้วพิจารณาผลต่างของความควรจะเป็นต่อไปนี้

$$\begin{aligned}
L(\theta) - L(\theta^{(\kappa)}) &= \log p(\mathsf{y}_{1:N}|\theta) - \log p(\mathsf{y}_{1:N}|\theta^{(\kappa)}) \\
&= \log \left[\int p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N}, \theta) p(\mathsf{x}_{0:N}|\theta) \, d\mathsf{x}_{0:N} \right] - \log p(\mathsf{y}_{1:N}|\theta^{(\kappa)}) \\
&= \log \left[\int \frac{p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N}, \theta) p(\mathsf{x}_{0:N}|\theta)}{p(\mathsf{y}_{1:N}|\theta^{(\kappa)})} \, d\mathsf{x}_{0:N} \right] \\
&= \log \left[\int p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) \left[\frac{p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N}, \theta) p(\mathsf{x}_{0:N}|\theta)}{p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) p(\mathsf{y}_{1:N}|\theta^{(\kappa)})} \right] d\mathsf{x}_{0:N} \right] \\
&= \log \mathbb{E}^{\mathsf{x}|\mathsf{y},\theta^{(\kappa)}} \left[\frac{p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N}, \theta) p(\mathsf{x}_{0:N}|\theta)}{p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) p(\mathsf{y}_{1:N}|\theta^{(\kappa)})} \right] \qquad (2.62)
\end{aligned}$$

เนื่องจากฟังก์ชันลอการิทึมเป็นฟังก์ชันเว้าจึงสามารถประยุกต์ใช้ทฤษฎีบทต่อไปนี้ได้

ทฤษฎีบทที่ 2.1 (อสมการของเจนเสน (Jensen's inequality)) *ให้ x เป็นตัวแปรสุ่มใด ๆ และ $f(x)$ เป็นฟังก์ชันเว้าแล้ว*

$$f[\mathbb{E}(x)] \geq \mathbb{E}[f(x)] \qquad (2.63)$$

เมื่อประยุกต์อสมการของเจนเสนกับสมการที่ (2.62) จะได้ว่า

$$\begin{aligned}
L(\theta) - L(\theta^{(\kappa)}) &= \log \left[\mathbb{E}^{\mathsf{x}|\mathsf{y},\theta^{(\kappa)}} \left(\frac{p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N}, \theta) p(\mathsf{x}_{0:N}|\theta)}{p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) p(\mathsf{y}_{1:N}|\theta^{(\kappa)})} \right) \right] \\
&\geq \mathbb{E}^{\mathsf{x}|\mathsf{y},\theta^{(\kappa)}} \left[\log \left(\frac{p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N}, \theta) p(\mathsf{x}_{0:N}|\theta)}{p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) p(\mathsf{y}_{1:N}|\theta^{(\kappa)})} \right) \right]
\end{aligned}$$

ดังนั้น

$$\begin{aligned}
L(\theta) - L(\theta^{(\kappa)}) &\geq \mathbb{E}^{\mathsf{x}|\mathsf{y},\theta^{(\kappa)}} \left[\log \frac{p(\mathsf{y}_{1:N}|\mathsf{x}_{0:N}, \theta) p(\mathsf{x}_{0:N}|\theta)}{p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) p(\mathsf{y}_{1:N}|\theta^{(\kappa)})} \right] \\
&= \int p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) \log \frac{p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta)}{p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta^{(\kappa)})} \, d\mathsf{x}_{0:N} \\
&\triangleq \ell(\theta, \theta^{(\kappa)}) \qquad (2.64)
\end{aligned}$$

นั่นคือ

$$\boxed{L(\theta) \geq L(\theta^{(\kappa)}) + \ell(\theta, \theta^{(\kappa)})} \tag{2.65}$$

สังเกตว่าถ้าแทน θ ด้วย $\theta^{(\kappa)}$ ลงใน $\ell(\theta, \theta^{(\kappa)})$ จะได้

$$\ell(\theta^{(\kappa)}, \theta^{(\kappa)}) = \int p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) \log \frac{p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta^{(\kappa)})}{p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta^{(\kappa)})}\, d\mathsf{x}_{0:N} = 0 \tag{2.66}$$

ดังนั้น $L(\theta^{(\kappa)}) + \ell(\theta, \theta^{(\kappa)})$ จึงเป็นขอบเขตล่างของ $L(\theta)$ และจะมีค่าเท่ากับ $L(\theta)$ ก็ต่อเมื่อเลือกพารามิเตอร์ θ ให้เท่ากับพารามิเตอร์ที่ประมาณได้ในแต่ละรอบ นั่นคือ $\theta = \theta^{(\kappa)}$ ด้วยเหตุนี้หากสามารถหา $\theta^{(\kappa)}$ ทำให้ $\ell(\theta, \theta^{(\kappa)})$ มีค่ามากขึ้นในแต่ละรอบของการทำซ้ำย่อมเป็นการรับประกันได้ว่า $L(\theta)$ ต้องมีค่ามากขึ้นเช่นกัน เพื่อบรรลุวัตถุประสงค์ดังกล่าวจึงเลือกพารามิเตอร์ในการทำซ้ำรอบถัดไป $\theta^{(\kappa+1)}$ ดังนี้

$$\begin{aligned}\theta^{(\kappa+1)} &= \underset{\theta}{\operatorname{argmax}}[\ell(\theta, \theta^{(\kappa)})] \\ &= \underset{\theta}{\operatorname{argmax}} \left[\int p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) \log \frac{p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta)}{p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta^{(\kappa)})}\, d\mathsf{x}_{0:N}\right]\end{aligned} \tag{2.67}$$

เนื่องจาก $p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta^{(\kappa)})$ ไม่เป็นฟังก์ชันของ θ จึงไม่มีผลต่อการคำนวณหา θ ดังนั้นเงื่อนไขข้างต้นจึงลดรูปเหลือ

$$\theta^{(\kappa+1)} = \underset{\theta}{\operatorname{argmax}} \left[\int p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) \log p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta)\, d\mathsf{x}_{0:N}\right] \tag{2.68}$$

ให้ $\mathcal{Q}(\cdot)$ คือค่าคาดหมายของค่าลอการิทึมของความควรจะเป็นแบบสมบูรณ์ (Complete likelihood) เมื่อเทียบกับการแจกแจงภายหลังร่วมเมื่อกำหนดพารามิเตอร์ $\theta^{(\kappa)}$ มาให้ นั่นคือ

$$\begin{aligned}\mathcal{Q}(\theta, \theta^{(\kappa)}) &= \mathbb{E}^{\mathsf{x}|\mathsf{y}, \theta^{(\kappa)}} [\log p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta)] \\ &= \int p(\mathsf{x}_{0:N}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) \log p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta)\, d\mathsf{x}_{0:N}\end{aligned} \tag{2.69}$$

ดังนั้น

$$\theta^{(\kappa+1)} = \underset{\theta}{\operatorname{argmax}} \left[Q(\theta, \theta^{(\kappa)}) \right] \tag{2.70}$$

ภายใต้สมมติฐานของแบบจำลอง HMM จึงอาศัยคุณสมบัติมาร์คอฟเพื่อกระจายฟังก์ชันการแจกแจงของ $p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta)$ ได้ดังนี้

$$p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta) = p(\mathsf{x}_0|\theta) \prod_{k=1}^{N} p(\mathsf{x}_k|\mathsf{x}_{k-1}, \theta) \prod_{k=1}^{N} p(\mathsf{y}_k|\mathsf{x}_k, \theta) \tag{2.71}$$

ดังนั้น

$$\log p(\mathsf{x}_{0:N}, \mathsf{y}_{1:N}|\theta) = \log p(\mathsf{x}_0|\theta) + \sum_{k=1}^{N} \log p(\mathsf{x}_k|\mathsf{x}_{k-1}, \theta) + \sum_{k=1}^{N} \log p(\mathsf{y}_k|\mathsf{x}_k, \theta) \tag{2.72}$$

เมื่อแทนกลับลงในสมการที่ (2.69) จะได้ว่า

$$Q(\theta, \theta^{(\kappa)}) = Q_{\mathsf{x}_0}(\theta, \theta^{(\kappa)}) + Q_{\mathsf{x}_k}(\theta, \theta^{(\kappa)}) + Q_{\mathsf{y}_k}(\theta, \theta^{(\kappa)}) \tag{2.73}$$

โดยที่

$$Q_{\mathsf{x}_0}(\theta, \theta^{(\kappa)}) = \int p(\mathsf{x}_0|\mathsf{y}_{1:N}, \theta^{(\kappa)}) \log p(\mathsf{x}_0|\theta) \, d\mathsf{x}_0 \tag{2.74}$$

$$Q_{\mathsf{x}_k}(\theta, \theta^{(\kappa)}) = \sum_{k=1}^{N} \int p(\mathsf{x}_k, \mathsf{x}_{k-1}|\mathsf{y}_{1:N}, \theta^{(\kappa)}) \log p(\mathsf{x}_k|\mathsf{x}_{k-1}, \theta) \, d\mathsf{x}_k \, d\mathsf{x}_{k-1} \tag{2.75}$$

$$Q_{\mathsf{y}_k}(\theta, \theta^{(\kappa)}) = \sum_{k=1}^{N} \int p(\mathsf{x}_k|\mathsf{y}_{1:N}, \theta^{(\kappa)}) \log p(\mathsf{y}_k|\mathsf{x}_k, \theta) \, d\mathsf{x}_k \tag{2.76}$$

จุดเด่นของแบบจำลอง HMM คือการลดขั้นตอนในการคำนวณหาการแจกแจงภายหลังตามขอบทั้งหมดดังในสมการที่ (2.69) เหลือแต่เพียงการคำนวณหาการแจกแจงปรับเรียบดังในสมการที่ (2.74) ถึง (2.76) เท่านั้น

บทที่ 3

วิธีดำเนินการวิจัย

ในบทนี้จะเป็นการนำทฤษฎีในบทที่ 2 มาประยุกต์ใช้กับการประมาณพารามิเตอร์ของแบบจำลอง SV ด้วยขั้นตอนวิธี EM ซึ่งเนื้อหาในบทนี้จะกล่าวถึงภาพรวมของขั้นตอนดำเนินการวิจัย ข้อมูลที่ใช้ในงานวิจัยและวิธีการประมาณพารามิเตอร์โดยการประยุกต์ใช้ตัวกรองและตัวปรับเรียบร่วมกับขั้นตอนวิธี EM โดยมีรายละเอียดดังนี้

ภาพรวมของขั้นตอนดำเนินการวิจัย

ในการทำนายความผันผวนของสินทรัพย์จำเป็นต้องอาศัยแบบจำลองซึ่งในที่นี้เลือกใช้แบบจำลอง SV ทั้งในรูปแบบไม่เชิงเส้น Σ_{NL} ในสมการที่ (2.7) และในรูปแบบเชิงเส้น Σ_{L} ในสมการที่ (2.11) โดยในแบบจำลองดังกล่าวมีพารามิเตอร์ที่ไม่ทราบค่า ดังนั้นจึงจำเป็นต้องทำการประมาณหาพารามิเตอร์ของแบบจำลองเป็นอันดับแรก จากนั้นจึงนำแบบจำลองที่ได้มาสังเคราะห์หาความผันผวนเป็นลำดับถัดไป

วัตถุประสงค์หลักของงานวิจัยชิ้นนี้คือการพัฒนาโปรแกรมใน MATLAB® เพื่อใช้ในการประมาณพารามิเตอร์ของแบบจำลอง SV ด้วยขั้นตอนวิธี EM โดยทำการเก็บตัวอย่างข้อมูลที่ได้จากสัญญาณออกของแบบจำลองและข้อมูลของอัตราแลกเปลี่ยน USD/THB เพื่อนำมาใช้ในการสังเคราะห์ตัวประมาณปรับเรียบซึ่งถูกนำมาใช้หาความควรจะเป็นแบบสมบูรณ์ซึ่งเป็นขั้นตอนหนึ่งในขั้นตอนวิธี EM ดังนั้นในการออกแบบงานวิจัยจึงมีลำดับขั้นตอนดังนี้

ขั้นตอนการทดสอบโปรแกรม

1. สร้างแบบจำลอง SV โดยกำหนดพารามิเตอร์ $\boldsymbol{\theta}$ ทุกตัว
2. สังเคราะห์หาสัญญาณออก $y_{1:N}$ จากแบบจำลอง
3. นำ $y_{1:N}$ มาป้อนเข้าสู่โปรแกรมที่ได้พัฒนาขึ้นเพื่อทำการประมาณ $\hat{\boldsymbol{\theta}}$
4. ตรวจสอบค่าความควรจะเป็นซึ่งควรมีแนวโน้มสูงขึ้นในทุก ๆ การวนรอบ
5. ให้ค่าความผิดพลาด $e = \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|$
   - 5.1 ถ้า $e \to 0$ จึงนำโปรแกรมไปใช้งานจริง
   - 5.2 ถ้า $e \not\to 0$ ให้ปรับแก้โปรแกรม

ขั้นตอนการนำไปใช้งาน

1. นำข้อมูลของอัตราแลกเปลี่ยน USD/THB มาป้อนเข้าสู่โปรแกรมที่ได้พัฒนาขึ้นเพื่อทำการประมาณ $\hat{\theta}$

2. สร้างแบบจำลอง SV โดยใช้พารามิเตอร์ที่ประมาณได้ $\hat{\theta}$

3. สังเคราะห์หาตัวแปรสถานะ x_k และความผันผวน $\sigma_k = e^{x_k/2}$

4. ใช้การคำนวณแบบทำซ้ำเพื่อทำนายความผันผวนของอัตราแลกเปลี่ยน

ทั้งนี้ข้อมูลของอัตราแลกเปลี่ยน USD/THB ที่เลือกใช้ในงานวิจัยชิ้นนี้นำมาจากเวปไซต์ของธนาคารแห่งประเทศไทย [46] โดยใช้ข้อมูลตั้งแต่ 4 มกราคม 2554 ถึง 29 เมษายน 2559

ขั้นตอนทั้งหมดนำมาแสดงไว้ในรูปของผังงานได้ดังในภาพ 4

การประมาณพารามิเตอร์ของแบบจำลอง SV ที่บรรยายในรูปแบบไม่เชิงเส้น

หัวข้อนี้เป็นการนำเสนอขั้นตอนในการประมาณพารามิเตอร์ของแบบจำลอง SV ที่บรรยายในรูปแบบไม่เชิงเส้น Σ_{NL} ในสมการที่ (2.7) ซึ่งนำมาเขียนใหม่ได้ดังนี้

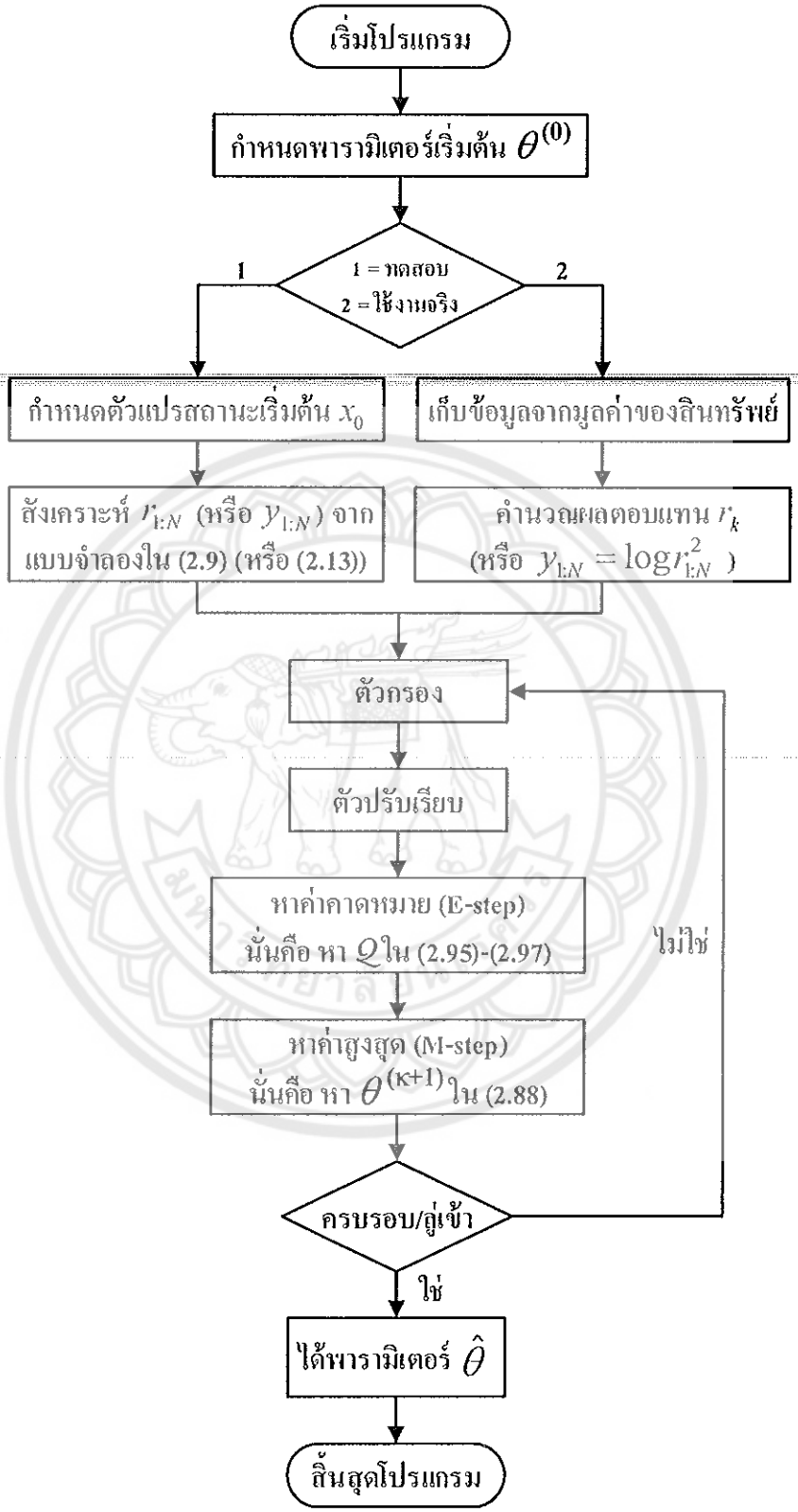
$$\Sigma_{\text{NL}} := \begin{cases} x_k &= \phi x_{k-1} + w_k, \\ r_k &= \beta e^{x_k/2} \epsilon_k \end{cases} \tag{3.1}$$

เมื่อ $w_k \sim \text{iid}\, \mathcal{N}(0, Q)$ และ $\epsilon_k \sim \text{iid}\, \mathcal{N}(0, 1)$

จากแบบจำลองข้างต้นเมื่อนำมาเขียนบรรยายด้วยแบบจำลอง HMM ผ่านทางฟังก์ชันการแจกแจงจะได้ว่า

$$\Sigma_{\text{HMM, NL}} := \begin{cases} x_k \sim p(x_k|x_{k-1}) = \mathcal{N}(\phi x_{k-1}, Q) \\ r_k \sim p(r_k|x_k) = \dfrac{1}{\beta\sqrt{2\pi}} \exp\left(-\dfrac{1}{2}\left[x_k + \dfrac{r_k^2}{\beta^2} e^{-x_k} \right] \right) \end{cases} \tag{3.2}$$

สังเกตว่าพารามิเตอร์ที่ไม่ทราบค่าในแบบจำลอง Σ_{NL} คือ $\theta = \{ \phi, Q, \beta \}$ นอกจากนี้ ในแบบจำลองดังกล่าวมีตัวแปรสถานะ x_k ซึ่งเป็นตัวแปรซ่อนจึงจำเป็นต้องทำการสังเคราะห์ขึ้นมาโดยอาศัยขั้นตอนวิธีของตัวกรองและขั้นตอนวิธีของตัวปรับเรียบเพื่อนำมาใช้คำนวณหาความควรจะเป็นแบบสมบูรณ์ซึ่งเป็นส่วนหนึ่งในขั้นตอนวิธี EM ดังนั้นในการออกแบบโปรแกรมจึงแบ่ง



ภาพ 4 ผังงานแสดงการทำงานของโปรแกรม

งานออกเป็น 2 ขั้นตอน กล่าวคือขั้นตอนการทดสอบโปรแกรมและขั้นตอนการนำไปใช้งานโดยมีรายละเอียดดังนี้

ในขั้นตอนการทดสอบโปรแกรม ให้กำหนดพารามิเตอร์ของแบบจำลอง $\boldsymbol{\theta} = \{\phi, Q, \beta\}$ พร้อมทั้งกำหนดตัวแปรสถานะเริ่มต้น $x_0 \sim \mathcal{N}(0,1)$ เพื่อคำนวณหา $r_{1:N} = \{r_1, \ldots, r_N\}$ โดยการคำนวณแบบทำซ้ำดังในสมการที่ (3.1) จากนั้นจึงนำ $r_{1:N}$ มาใช้เป็นข้อมูลสำหรับประมาณพารามิเตอร์โดยสมมติให้พารามิเตอร์เริ่มต้นสำหรับการวนซ้ำในขั้นตอนวิธี EM คือ $\boldsymbol{\theta}^{(0)}$

ขั้นแรกให้สังเคราะห์ตัวประมาณการกรองโดยในที่นี้เลือกใช้ขั้นตอนวิธีของตัวกรอง BF ดังแสดงไว้ในขั้นตอนวิธีที่ 2.2 หน้า 20 ในขั้นตอนเริ่มต้นให้สุ่มเลือกอนุภาคเริ่มต้น $x_0^{\mathrm{f}(i)} \sim \mathcal{N}(0,1)$ ซึ่งเป็นการแจกแจงก่อนของตัวแปรสถานะและให้ค่าถ่วงน้ำหนักเริ่มต้น $w_0^{\mathrm{f}(i)} = 1/n_f$ เมื่อ n_f คือจำนวนอนุภาคการกรองที่แต่ละเวลา

ในขั้นตอนการทำซ้ำให้สุ่มเลือกอนุภาคตัวใหม่ $x_k^{\mathrm{f}(i)}$ และคำนวณหาค่าถ่วงน้ำหนัก $\tilde{w}_k^{\mathrm{f}(i)}$ จากฟังก์ชันการแจกแจงแบบมีเงื่อนไขในแบบจำลอง HMM ดังในสมการที่ (3.2) กล่าวคือ

$$x_k^{\mathrm{f}(i)} \sim p(x_k|x_{k-1}^{\mathrm{f}(i)}) = \frac{1}{\sqrt{2\pi Q}} \exp\left(-\frac{1}{2}\frac{(x_k - \phi x_{k-1}^{\mathrm{f}(i)})^2}{Q}\right) \tag{3.3}$$

$$\tilde{w}_k^{\mathrm{f}(i)} \propto p(r_k|x_k^{\mathrm{f}(i)}) = \frac{1}{\beta\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[x_k^{\mathrm{f}(i)} + \frac{r_k^2}{\beta^2} e^{-x_k^{\mathrm{f}(i)}}\right]\right) \tag{3.4}$$

แล้วทำให้เป็นบรรทัดฐาน $w_k^{\mathrm{f}(i)}$ เพื่อป้อนเข้าสู่กระบวนการปรับอนุภาคใหม่ เมื่อคำนวณจนครบจำนวนรอบ N จะได้เซตของอนุภาคและค่าถ่วงน้ำหนัก $\{\mathrm{x}_k^{\mathrm{f}(i)}, w_k^{\mathrm{f}(i)}\}$ เมื่อ $k = 1, \ldots, N$ และ $i = 1, \ldots, n_f$

จากนั้นจึงเข้าสู่ขั้นตอนวิธีของตัวปรับเรียบ BS-PS ดังแสดงไว้ในขั้นตอนวิธีที่ 2.4 หน้า 24 โดยในขั้นตอนการทำซ้ำให้เลือกค่าถ่วงน้ำหนักจากการแจกแจงดังนี้

$$w_k^{\mathrm{s}(i)} \propto w_k^{\mathrm{f}(i)} p(x_{k+1}^{\mathrm{s}(j)}|x_k^{\mathrm{f}(i)}) = \frac{w_k^{\mathrm{f}(i)}}{\sqrt{2\pi Q}} \exp\left(-\frac{1}{2}\frac{(x_{k+1}^{\mathrm{s}(j)} - \phi x_k^{\mathrm{f}(i)})^2}{Q}\right) \tag{3.5}$$

เมื่อทำจนครบจำนวนรอบ N จะได้แนววิถีของตัวประมาณปรับเรียบมา 1 แนว จากนั้นให้ทำซ้ำจนได้จำนวนแนววิถีครบ n_s แนวตามที่ออกแบบไว้ ทั้งนี้จำนวน n_s ไม่จำเป็นต้องเท่ากับจำนวนอนุภาค n_f ในขั้นตอนวิธีของตัวกรอง

ให้ $x_{0:N}^{s(j)}$ เมื่อ $j = 1, \ldots, n_s$ แทนแนววิถีของตัวประมาณปรับเรียบ จากนั้นจึงเข้าสู่ขั้นตอน E (Expectation step) ในการคำนวณค่าคาดหมายโดยประยุกต์ใช้สมการที่ (2.74) ถึง (2.76) เมื่อได้ฟังก์ชันค่าคาดหมายข้างต้นเป็นที่เรียบร้อยแล้ว จึงนำเข้าสู่ขั้นตอน M (Maximization step) โดยการนำ $\mathcal{Q}_{x_0}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(\kappa)})$, $\mathcal{Q}_{x_k}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(\kappa)})$ และ $\mathcal{Q}_{y_k}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(\kappa)})$ มาหาอนุพันธ์ย่อยเทียบกับพารามิเตอร์ $\boldsymbol{\theta} = \{ \phi, Q, \beta \}$ ซึ่งจะได้ว่า

$$\phi^{(\kappa+1)} = \frac{1}{n_s} \sum_{j=1}^{n_s} \left[\sum_{k=1}^{N} x_k^{s(j)} x_{k-1}^{s(j)} \right] \Bigg/ \frac{1}{n_s} \sum_{j=1}^{n_s} \left[\sum_{k=1}^{N} \left(x_{k-1}^{s(j)} \right)^2 \right] \tag{3.6}$$

$$Q^{(\kappa+1)} = \frac{1}{n_s} \sum_{j=1}^{n_s} \left[\frac{1}{N} \sum_{k=1}^{N} \left(x_k^{s(j)} - \phi^{(\kappa+1)} x_{k-1}^{s(j)} \right)^2 \right] \tag{3.7}$$

$$\beta^{(\kappa+1)} = \sqrt{\frac{1}{n_s} \sum_{j=1}^{n_s} \left[\frac{1}{N} \sum_{k=1}^{N} \left(r_k^2 e^{-x_k^{s(j)}} \right) \right]} \tag{3.8}$$

ดังนั้นพารามิเตอร์ชุดใหม่ซึ่งจะนำไปเป็นพารามิเตอร์เริ่มต้นในการวนซ้ำรอบถัดไปคือ

$$\boldsymbol{\theta}^{(\kappa+1)} = \{ \phi^{(\kappa+1)}, Q^{(\kappa+1)}, \beta^{(\kappa+1)} \}$$

การทำงานของโปรแกรมจะสิ้นสุดลงเมื่อทำค่าความควรจะเป็นเข้าใกล้กันมากจนเป็นที่น่าพอใจหรือทำงานครบจำนวนรอบที่กำหนดไว้ในตอนแรกโดยจะได้พารามิเตอร์ที่ทำให้ฟังก์ชันควรจะเป็นมีค่าสูงสุดคือ $\hat{\theta}$ จากนั้นจึงนำมาเปรียบเทียบกับพารามิเตอร์จริง θ ที่ได้กำหนดไว้ในตอนแรก หากการทำงานของโปรแกรมถูกต้อง $\hat{\theta}$ และ θ จะมีค่าใกล้เคียงกัน ไม่เช่นนั้นแล้วให้ตรวจสอบโปรแกรมอีกครั้งจนได้ผลเป็นที่น่าพึงพอใจ

เมื่อผ่านขั้นตอนการทดสอบโปรแกรมเป็นที่เรียบร้อยแล้ว ถัดมาคือขั้นตอนการนำไปใช้งานโดยการรับข้อมูลที่เป็นมูลค่าของสินทรัพย์ซึ่งในที่นี้คืออัตราแลกเปลี่ยน USD/THB นำอัตราแลกเปลี่ยนมาคำนวณหาผลตอบแทนแบบลอการิทึม $r_{1:N}$ เมื่อ N คือจำนวนข้อมูลทั้งหมด จากนั้นจึงนำ $r_{1:N}$ มาเข้าสู่ขั้นตอนวิธี EM แล้วดำเนินการตามขั้นตอนการกรอง การปรับเรียบ การหาค่าคาดหมายและการหาค่าสูงสุดดังเช่นในขั้นตอนการทดสอบโปรแกรมจนกระทั่งได้พารามิเตอร์ $\hat{\theta}$ ที่ต้องการ

ในกรณีที่เลือกใช้ขั้นตอนวิธีของตัวกรอง GPF ดังแสดงไว้ในขั้นตอนวิธีที่ 2.3 หน้า 22 จะ

มีขั้นตอนการทำงานเช่นเดียวกับที่กล่าวไว้ข้างต้นยกเว้นในส่วนของกระบวนการปรับอนุภาคใหม่ โดยขั้นตอนวิธีของตัวกรอง GPF จะนำอนุภาค $x_k^{(i)}$ ในสมการที่ (3.3) และค่าถ่วงน้ำหนัก $w_k^{(i)}$ ที่ทำเป็นบรรทัดฐานแล้วซึ่งได้จาก $\tilde{w}_k^{(i)}$ ในสมการที่ (3.4) มาคำนวณหา μ_k^l และ P_k^l ดังในสมการที่ (2.42) แล้วจึงสุ่มเลือกหาอนุภาคชุดใหม่จาก $x_k^{(i)} \sim \mathcal{N}(\mu_k^l, P_k^l)$

เมื่อหาพารามิเตอร์ของแบบจำลองเรียบร้อยแล้ว จึงทำการคำนวณแบบวนซ้ำโดยอาศัยสมการที่ (3.1) เพื่อคำนวณหาตัวแปรสถานะ x_k ซึ่งในที่นี้คือความผันผวนแบบลอการิทึม ดังนั้นความผันผวนที่ต้องการทราบคือ $\sigma_k = e^{x_k/2}$ นั่นเอง

การประมาณพารามิเตอร์ของแบบจำลอง SV ที่บรรยายในรูปแบบเชิงเส้น

หัวข้อนี้เป็นการนำเสนอขั้นตอนในการประมาณพารามิเตอร์ของแบบจำลอง SV ที่บรรยายในรูปแบบเชิงเส้น Σ_L ในสมการที่ (2.11) ซึ่งนำมาเขียนใหม่ได้ดังนี้

$$\Sigma_L := \begin{cases} x_k &= \phi x_{k-1} + w_k, \\ y_k &= x_k + \alpha + v_k \end{cases} \tag{3.9}$$

เมื่อ $w_k \sim \text{iid } \mathcal{N}(0, Q)$ และ $v_k \sim \log \chi^2 - \mathbb{E}[\log \epsilon_k^2]$

จากแบบจำลองข้างต้นเมื่อนำมาเขียนบรรยายด้วยแบบจำลอง HMM ผ่านทางฟังก์ชันการแจกแจงจะได้ว่า

$$\Sigma_{\text{HMM, L}} := \begin{cases} x_k \sim p(x_k|x_{k-1}) = \mathcal{N}(\phi x_{k-1}, Q) \\ y_k \sim p(y_k|x_k) = \dfrac{1}{\sqrt{2\pi}} \exp\left[-\dfrac{1}{2}\left(\exp(\gamma_k) - \gamma_k\right)\right] \end{cases} \tag{3.10}$$

เมื่อ $\gamma_k := y_k - x_k - \alpha + \mathbb{E}[\log \epsilon_k^2]$

สำหรับพารามิเตอร์ที่ไม่ทราบค่าในแบบจำลอง Σ_L คือ $\boldsymbol{\theta} = \{\phi, Q, \alpha\}$ และมี x_k เป็นตัวแปรซ่อนในลักษณะเช่นเดียวกับแบบจำลอง Σ_{NL} ดังนั้นการประมาณพารามิเตอร์ในกรณีนี้จึงมีขั้นตอนคล้ายคลึงกับที่ได้ศึกษามาในหัวข้อที่ผ่านมาหากแต่ต้องมีการปรับเปลี่ยนฟังก์ชันการแจกแจงตามความเหมาะสม

ในขั้นตอนการทดสอบโปรแกรม ให้กำหนดพารามิเตอร์ของแบบจำลอง $\boldsymbol{\theta} = \{\phi, Q, \alpha\}$

พร้อมทั้งกำหนดตัวแปรสถานะเริ่มต้น $x_0 \sim \mathcal{N}(0,1)$ เพื่อคำนวณหา $y_{1:N} = \{y_1, \ldots, y_N\}$ โดยการคำนวณแบบทำซ้ำดังในสมการที่ (3.9) จากนั้นจึงนำ $y_{1:N}$ มาใช้เป็นข้อมูลสำหรับประมาณพารามิเตอร์โดยสมมติให้พารามิเตอร์เริ่มต้นสำหรับการวนซ้ำในขั้นตอนวิธี EM คือ $\theta^{(0)}$

ขั้นแรกให้สังเคราะห์ตัวประมาณการกรองโดยใช้ขั้นตอนวิธีของตัวกรอง BF ในขั้นตอนเริ่มต้นให้สุ่มเลือกอนุภาคเริ่มต้น $x_0^{\mathrm{f}(i)} \sim \mathcal{N}(0,1)$ และให้ค่าถ่วงน้ำหนักเริ่มต้น $w_0^{\mathrm{f}(i)} = 1/n_f$ จากนั้นจึงเข้าสู่ขั้นตอนการทำซ้ำโดยสุ่มเลือกอนุภาคตัวใหม่ $x_k^{\mathrm{f}(i)}$ และคำนวณค่าถ่วงน้ำหนัก $\tilde{w}_k^{\mathrm{f}(i)}$ จากฟังก์ชันการแจกแจงแบบมีเงื่อนไขในแบบจำลอง HMM ดังในสมการที่ (3.10) กล่าวคือ

$$x_k^{\mathrm{f}(i)} \sim p(x_k | x_{k-1}^{\mathrm{f}(i)}) = \frac{1}{\sqrt{2\pi Q}} \exp\left(-\frac{1}{2} \frac{(x_k - \phi x_{k-1}^{\mathrm{f}(i)})^2}{Q} \right) \tag{3.11}$$

$$\tilde{w}_k^{\mathrm{f}(i)} \propto p(y_k | x_k^{\mathrm{f}(i)}) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\exp\left(\gamma_k^{\mathrm{f}(i)} \right) - \gamma_k^{\mathrm{f}(i)} \right) \right] \tag{3.12}$$

เมื่อ $\gamma_k^{\mathrm{f}(i)} := y_k - x_k^{\mathrm{f}(i)} - \alpha + \mathbb{E}[\log \epsilon_k^2]$

นำ $\tilde{w}_k^{\mathrm{f}(i)}$ มาทำให้เป็นบรรทัดฐาน $w_k^{\mathrm{f}(i)}$ เพื่อป้อนเข้าสู่กระบวนการปรับอนุภาคใหม่ เมื่อคำนวณจนครบจำนวนรอบ N จะได้เซตของอนุภาคและค่าถ่วงน้ำหนัก $\{x_k^{\mathrm{f}(i)}, w_k^{\mathrm{f}(i)}\}$ เมื่อ $k = 1, \ldots, N$ และ $i = 1, \ldots, n_f$ ถัดมาเป็นขั้นตอนวิธีของตัวปรับเรียบ BS-PS โดยในขั้นตอนการทำซ้ำให้เลือกค่าถ่วงน้ำหนักจากการแจกแจงดังนี้

$$w_k^{\mathrm{s}(i)} \propto w_k^{\mathrm{f}(i)} p(x_{k+1}^{\mathrm{s}(j)} | x_k^{\mathrm{f}(i)}) = \frac{w_k^{\mathrm{f}(i)}}{\sqrt{2\pi Q}} \exp\left(-\frac{1}{2} \frac{(x_{k+1}^{\mathrm{s}(j)} - \phi x_k^{\mathrm{f}(i)})^2}{Q} \right) \tag{3.13}$$

เมื่อทำจนครบจำนวนรอบจะได้แนววิถีของตัวประมาณปรับเรียบ $x_{0:N}^{\mathrm{s}(j)}$ เมื่อ $j = 1, \ldots, n_s$ จากนั้นจึงเข้าสู่ขั้นตอน E ในการคำนวณค่าคาดหมายโดยใช้สมการที่ (2.74) ถึง (2.76) นำฟังก์ชันค่าคาดหมายที่ได้มาหาอนุพันธ์ย่อยเทียบกับพารามิเตอร์ $\theta = \{ \phi, Q, \alpha \}$ จะได้

$$\phi^{(\kappa+1)} = \frac{1}{n_s} \sum_{j=1}^{n_s} \left[\sum_{k=1}^{N} x_k^{\mathrm{s}(j)} x_{k-1}^{\mathrm{s}(j)} \right] \Bigg/ \frac{1}{n_s} \sum_{j=1}^{n_s} \left[\sum_{k=1}^{N} \left(x_{k-1}^{\mathrm{s}(j)} \right)^2 \right] \tag{3.14}$$

$$Q^{(\kappa+1)} = \frac{1}{n_s} \sum_{j=1}^{n_s} \left[\frac{1}{N} \sum_{k=1}^{N} \left(x_k^{\mathrm{s}(j)} - \phi^{(\kappa+1)} x_{k-1}^{\mathrm{s}(j)} \right)^2 \right] \tag{3.15}$$

$$\alpha^{(\kappa+1)} = \log\left(\frac{1}{n_s} \sum_{j=1}^{n_s} \left[\frac{1}{N} \sum_{k=1}^{N} \exp\left(y_k - x_k^{\mathrm{s}(j)} + \mathbb{E}[\log \epsilon_k^2] \right) \right] \right) \tag{3.16}$$

ดังนั้นพารามิเตอร์ชุดใหม่ซึ่งจะนำไปเป็นพารามิเตอร์เริ่มต้นในการวนซ้ำรอบถัดไปคือ

$$\theta^{(\kappa+1)} = \{\, \phi^{(\kappa+1)},\, Q^{(\kappa+1)},\, \alpha^{(\kappa+1)} \,\}$$

โปรแกรมจะสิ้นสุดการทำงานเมื่อค่าความควรจะเป็นเข้าใกล้กันมากจนเป็นที่น่าพอใจ หรือทำงานครบจำนวนรอบที่กำหนดไว้ในตอนแรกโดยจะได้พารามิเตอร์ที่ทำให้ฟังก์ชันควรจะเป็นมีค่าสูงสุดคือ $\hat{\theta}$ จากนั้นจึงนำมาเปรียบเทียบกับพารามิเตอร์จริง θ ที่ได้กำหนดไว้ในตอนแรก หากการทำงานของโปรแกรมถูกต้อง $\hat{\theta}$ และ θ จะมีค่าใกล้เคียงกัน ไม่เช่นนั้นแล้วให้ตรวจสอบโปรแกรมอีกครั้งจนได้ผลเป็นที่น่าพึงพอใจ

เมื่อผ่านขั้นตอนการทดสอบโปรแกรมเป็นที่เรียบร้อยแล้ว ถัดมาคือขั้นตอนการนำไปใช้งานโดยการรับข้อมูลที่เป็นมูลค่าของสินทรัพย์ซึ่งในที่นี้เลือกใช้อัตราแลกเปลี่ยน USD/THB นำอัตราแลกเปลี่ยนมาคำนวณหาผลตอบแทนแบบลอการิทึม $r_{1:N}$ เมื่อ N คือจำนวนข้อมูลทั้งหมด จากนั้นจึงคำนวณหา $y_k = \log r_k^2$ เมื่อ $k = 1, \ldots, N$ ก่อนนำมาเข้าสู่ขั้นตอนวิธี EM แล้วดำเนินการตามขั้นตอนการกรอง การปรับเรียบ การหาค่าคาดหมายและการหาค่าสูงสุดดังเช่นในขั้นตอนการทดสอบโปรแกรมจนกระทั่งได้พารามิเตอร์ $\hat{\theta}$ ที่ต้องการ

ในกรณีที่เลือกใช้ขั้นตอนวิธีของตัวกรอง GPF ดังแสดงไว้ในขั้นตอนวิธีที่ 2.3 หน้า 22 จะมีขั้นตอนการทำงานเช่นเดียวกับที่กล่าวไว้ข้างต้นยกเว้นในส่วนของกระบวนการปรับอนุภาคใหม่ โดยขั้นตอนวิธีของตัวกรอง GPF จะนำอนุภาค $x_k^{\mathrm{f}(i)}$ ในสมการที่ (3.11) และค่าถ่วงน้ำหนัก $w_k^{\mathrm{f}(i)}$ ที่ทำเป็นบรรทัดฐานแล้วซึ่งได้จาก $\tilde{w}_k^{\mathrm{f}(i)}$ ในสมการที่ (3.12) มาคำนวณหา μ_k^{f} และ P_k^{f} ดังในสมการที่ (2.42) แล้วจึงสุ่มเลือกหาอนุภาคชุดใหม่จาก $\mathrm{x}_k^{\mathrm{f}(i)} \sim \mathcal{N}(\mu_k^{\mathrm{f}}, P_k^{\mathrm{f}})$

เมื่อหาพารามิเตอร์ของแบบจำลองเป็นที่เรียบร้อยแล้ว จึงทำการคำนวณแบบวนซ้ำโดยอาศัยสมการที่ (3.9) เพื่อคำนวณหาตัวแปรสถานะ x_k ซึ่งคือความผันผวนแบบลอการิทึม ดังนั้นความผันผวนที่ต้องการทราบคือ $\sigma_k = e^{x_k/2}$ นั่นเอง

บทที่ 4

ผลการทดลอง

เนื้อหาในบทนี้เป็นการนำเสนอผลการประมาณพารามิเตอร์ของแบบจำลองเชิงเฟ้นสุ่มด้วยขั้นตอนวิธี EM ทั้งในกรณีแบบจำลองไม่เชิงเส้นและกรณีแบบจำลองเชิงเส้นดังนี้

ผลการทดลองของแบบจำลอง SV ที่บรรยายในรูปแบบไม่เชิงเส้น

1. ผลการทดสอบโปรแกรม

การเลือกพารามิเตอร์เริ่มต้น $\theta^{(0)}$ จำนวนอนุภาคการกรอง n_f และจำนวนแนววิถีปรับเรียบ n_s มีความสำคัญต่อการประมาณด้วยวิธี MCM ซึ่งเป็นวิธีการแก้ปัญหาแบบศึกษาสำนึก ดังนั้นในการทดสอบโปรแกรมนี้จึงทำการทดลองโดยการปรับค่า $\theta^{(0)}$, n_f และ n_s เพื่อเลือกค่าที่เหมาะสมสำหรับการประมาณพารามิเตอร์ของแบบจำลองไม่เชิงเส้น Σ_{NL} ทั้งนี้สมมติให้พารามิเตอร์จริงของแบบจำลองคือ

$$\theta = \{\phi, Q, \beta\} = \{0.9, 0.5, 0.0022\}$$

ซึ่งค่าดังกล่าวอ้างอิงจากงานวิจัยของ [?]

1. เมื่อกำหนดพารามิเตอร์เริ่มต้น $\theta^{(0)}$ ใน 2 ลักษณะ นั่นคือ

แบบที่ 1 ให้ $\theta_1^{(0)} = \{0.81, 0.45, 0.00198\}$ ซึ่งห่างจาก θ ในอัตราร้อยละ 10

แบบที่ 2 ให้ $\theta_2^{(0)} = \{0.45, 0.25, 0.0011\}$ ซึ่งห่างจาก θ ในอัตราร้อยละ 50

จากการทดลองพบว่าค่าเฉลี่ยของพารามิเตอร์ที่ประมาณได้ $\hat{\theta}$ เมื่อกำหนด $\theta^{(0)}$ ที่แตกต่างกันสามารถลู่เข้าสู่พารามิเตอร์จริง θ ของแบบจำลองโดยใช้เวลาไม่แตกต่างกันมากนัก แต่หากพิจารณาเวลาในขั้นตอนวิธีการกรองพบว่า ตัวกรอง GPF ใช้เวลาในการสังเคราะห์ตัวประมาณการกรองน้อยกว่าตัวกรอง BF อย่างชัดเจน

2. การทดลองถัดมาคือการเลือกจำนวนอนุภาคการกรอง n_f โดยกำหนดให้พารามิเตอร์

จริง θ และพารามิเตอร์เริ่มต้น $\theta^{(0)}$ คือ

$$\theta = \{\, \phi,\, Q,\, \beta \,\} = \{\, 0.9,\, 0.5,\, 0.0022 \,\}$$

$$\theta^{(0)} = \{\, \phi^{(0)},\, Q^{(0)},\, \beta^{(0)} \,\} = \{\, 0.45,\, 0.25,\, 0.0011 \,\}$$

โดยเริ่มต้นกำหนดให้ $n_f = 400 = n_s$ จากนั้นจึงลดค่า n_f เป็นดังนี้ $n_f = 200, 100, 50$ จากการทดลองพบว่า $n_f = 200$ มีค่าเหมาะสมของแบบจำลองนี้เนื่องจากพารามิเตอร์ลู่เข้าเร็วกว่าและให้ค่าพารามิเตอร์ใกล้เคียงกับค่าจริงมากที่สุด

3. สุดท้ายกำหนดให้ $n_f = 200$ จากนั้นจึงปรับจำนวนแนววิถีปรับเรียบ n_s ดังนี้ $n_s = 200, 100, 50, 25, 12$ ผลการทดลองพบว่า $n_s = 50$ ให้ค่าเหมาะสมเนื่องจากมีจำนวนรอบในการลู่เข้าน้อยที่สุด

2. **ผลการนำไปใช้งานกับข้อมูลของอัตราแลกเปลี่ยน**

การทดลองในส่วนนี้เป็นการหาพารามิเตอร์ของแบบจำลอง Σ_{NL} โดยใช้ข้อมูลของอัตราแลกเปลี่ยน USD/THB ระหว่างวันที่ 4 ม.ค. 2554 จนถึง 29 เม.ย. 2559 มาเป็นกรณีศึกษา โดยกำหนดให้พารามิเตอร์เริ่มต้น $\theta^{(0)} = \{0.45, 0.25, 0.0011\}$ จำนวนอนุภาคการกรอง $n_f = 200$ และจำนวนแนววิถีปรับเรียบ $n_s = 50$ ซึ่งเป็นค่าที่ได้มาจากการทดสอบโปรแกรมในหัวข้อที่ผ่านมาแล้วโดยทำการทดลองซ้ำทั้งหมด 20 ครั้งเพื่อหาค่าเฉลี่ยทางสถิติ

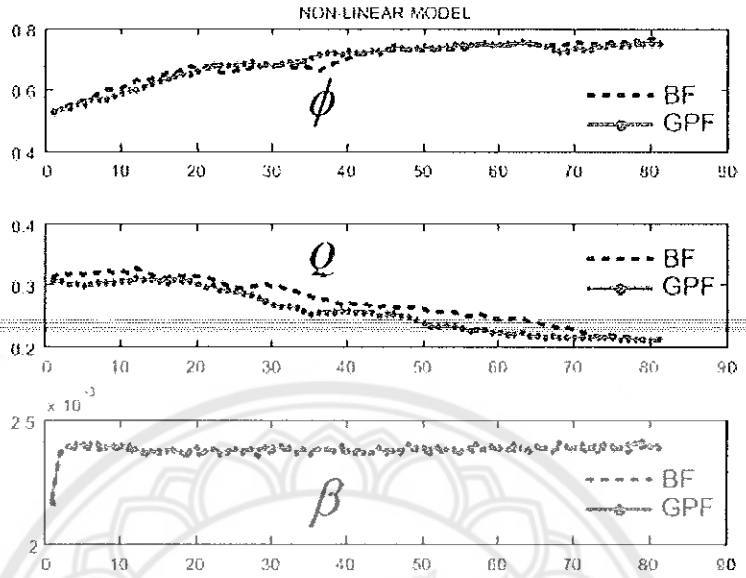
ผลการประมาณพารามิเตอร์ที่ได้แสดงไว้ในตาราง 1 โดยค่าในวงเล็บคือส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์ที่ได้จากการทดลองซ้ำ

ตาราง 1 $\hat{\theta} = \{\, \hat{\phi},\, \hat{Q},\, \hat{\beta} \,\}$ สำหรับ Σ_{NL} ของอัตราแลกเปลี่ยน

สกุลเงิน	BF			GPF		
	$\hat{\phi}$	$\hat{Q}$	$\hat{\beta}$	$\hat{\phi}$	$\hat{Q}$	$\hat{\beta}$
USD	0.7920 (0.0259)	0.1808 (0.0299)	0.00240 (0.00002)	0.7876 (0.0171)	0.1886 (0.0240)	0.00240 (0.00002)

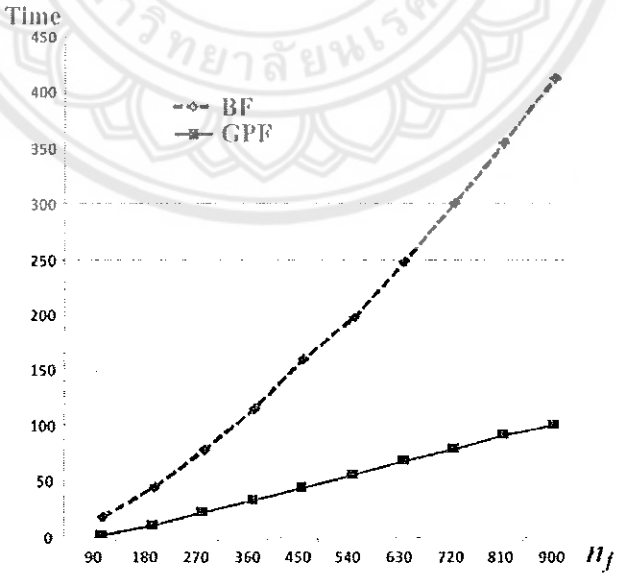
สำหรับกราฟการลู่เข้าของพารามิเตอร์ $\hat{\theta} = \{\, \hat{\phi},\, \hat{Q},\, \hat{\beta} \,\}$ ที่ได้จากขั้นตอนวิธี EM ของอัตราแลกเปลี่ยน USD/THB แสดงไว้ในภาพ 5 ซึ่งพบว่าค่าของพารามิเตอร์ที่ประมาณได้โดยใช้ตัวกรอง BF และตัวกรอง GPF ลู่เข้าสู่ค่าที่ใกล้เคียงกันแต่หากพิจารณาในแง่ของเวลาพบว่าตัวกรอง

GPF ใช้เวลาในการประมวลผลน้อยกว่า ตัวกรอง BF



ภาพ 5 การลู่เข้าของ $\hat{\theta} = \{ \hat{\phi}, \hat{Q}, \hat{\beta} \}$ สำหรับ Σ_{NL}

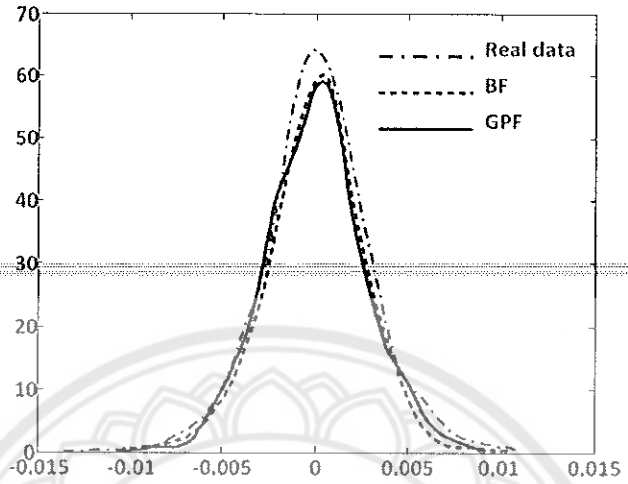
ภาพ 6 แสดงเวลาที่ใช้ประมวลผลในขั้นตอนวิธีการกรองโดยเปรียบเทียบระหว่างการใช้ตัวกรอง BF และตัวกรอง GPF ของอัตราแลกเปลี่ยน USD/THB ซึ่งจะเห็นว่าเมื่อเพิ่มจำนวนอนุภาคการกรองมากขึ้น ตัวกรอง BF จะใช้เวลาประมวลผลมากกว่าตัวกรอง GPF อย่างชัดเจน



ภาพ 6 เวลาในการประมวลผลระหว่าง BF และ GPF ของ USD/THB สำหรับ Σ_{NL}

สุดท้ายให้นำพารามิเตอร์ $\hat{\theta} = \{ \hat{\phi}, \hat{Q}, \hat{\beta} \}$ ในตาราง 1 มาจำลองหาผลตอบแทนผ่านทางแบบจำลอง Σ_{NL} แล้วนำมาวาดฟังก์ชันการแจกแจงเพื่อตรวจสอบความคล้ายคลึงกันระหว่าง

ฟังก์ชันการแจกแจงของผลตอบแทนจริงและฟังก์ชันการแจกแจงของผลตอบแทนที่เกิดจากการจำลอง



ภาพ 7 ฟังก์ชันการแจกแจงของผลตอบแทนจริงและผลตอบแทนจากการจำลอง Σ_{NL}

ในภาพ 7 เป็นการเปรียบเทียบฟังก์ชันการแจกแจงของผลตอบแทนจริงที่เกิดจากอัตราแลกเปลี่ยน USD/THB กับฟังก์ชันการแจกแจงที่ได้จากการจำลองโดยเลือกใช้พารามิเตอร์ดังนี้

$$\hat{\theta}_{BF} = \{\, 0.7920,\ 0.1808,\ 0.00240 \,\} \quad \text{และ} \quad \hat{\theta}_{GPF} = \{\, 0.7876,\ 0.1886,\ 0.00240 \,\}$$

จะเห็นว่ารูปร่างเค้าโครงของฟังก์ชันการแจกแจงมีลักษณะคล้ายคลึงกัน หากพิจารณาคุณสมบัติด้านโมเมนต์ของฟังก์ชันการแจกแจงได้แก่ ค่ามัชฌิม (Mean) ค่าความแปรปรวน (Variance) ค่าความเบ้ (Skewness) และค่าภาวะยอดมน (Kurtosis) ซึ่งเป็นโมเมนต์ที่ 1, 2, 3 และ 4 ตามลำดับ ผลที่ได้แสดงไว้ในตาราง 2

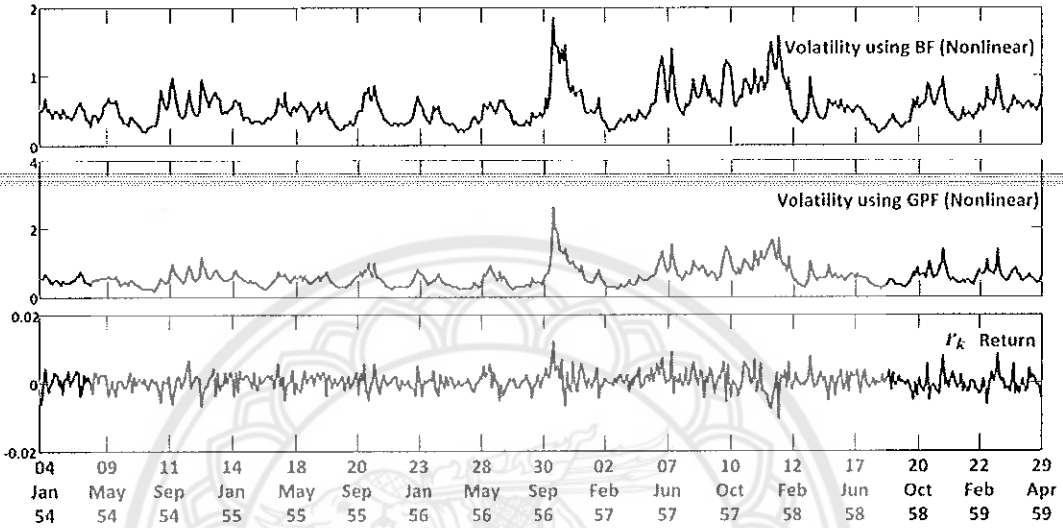
ตาราง 2 คุณสมบัติด้านโมเมนต์ของฟังก์ชันการแจกแจงสำหรับ Σ_{NL}

	ค่ามัชฌิม	ค่าความแปรปรวน	ค่าความเบ้	ค่าภาวะยอดมน
USD/THB	0.0001	9.12×10^{-6}	-0.2544	5.8528
BF	0.0000	6.02×10^{-6}	-0.0506	4.1009
GPF	0.0000	7.31×10^{-6}	-0.0451	4.5365

จากคุณสมบัติด้านโมเมนต์ที่ได้พบว่าค่าโมเมนต์ทั้งสี่ของฟังก์ชันการแจกแจงของผลตอบแทนจริงและฟังก์ชันการแจกแจงที่ได้จากการจำลองมีความใกล้เคียงกัน จึงอนุมานได้ว่าผลการประมาณพารามิเตอร์ที่ได้มีความถูกต้อง

3. **ความผันผวนของอัตราแลกเปลี่ยน**

เมื่อการทำงานของโปรแกรมสิ้นสุดลง ให้นำตัวประมาณปรับเรียบ x_k^s ที่ได้ในรอบสุดท้ายมาคำนวณหาความผันผวน $\sigma_k = e^{x_k^s/2}$



ภาพ 8 ค่าความผันผวนและผลตอบแทนของ USD/THB สำหรับ Σ_{NL}

ผลการทดลองของแบบจำลอง SV ที่บรรยายในรูปแบบเชิงเส้น

1. ผลการทดสอบโปรแกรม

ในหัวข้อนี้ใช้วิธีการทดลองเช่นเดียวกับในกรณีของแบบจำลอง Σ_{NL} โดยทำการทดลองเพื่อเลือกพารามิเตอร์เริ่มต้น $\theta^{(0)}$ จำนวนอนุภาคการกรอง n_f และจำนวนแนววิถีปรับเรียบ n_s ที่เหมาะสมโดยในที่นี้สมมติให้พารามิเตอร์จริงของแบบจำลองคือ

$$\boxed{\theta = \{\phi, Q, \alpha\} = \{0.9, 0.5, -13.5\}}$$

1. เมื่อกำหนดพารามิเตอร์เริ่มต้น $\theta^{(0)}$ ใน 2 ลักษณะ นั่นคือ

แบบที่ 1 ให้ $\theta_1^{(0)} = \{0.81, 0.45, -12.15\}$ ซึ่งห่างจาก θ ในอัตราร้อยละ 10

แบบที่ 2 ให้ $\theta_2^{(0)} = \{0.45, 0.25, -6.75\}$ ซึ่งห่างจาก θ ในอัตราร้อยละ 50

จากการทดลองพบว่าค่าเฉลี่ยของพารามิเตอร์ที่ประมาณได้ $\hat{\theta}$ เมื่อกำหนด $\theta^{(0)}$ ที่แตกต่างกันสามารถลู่เข้าสู่พารามิเตอร์จริง θ ของแบบจำลองโดยใช้เวลาไม่แตกต่างกันมากนัก

แต่หากพิจารณาเวลาในขั้นตอนวิธีการกรองพบว่า ตัวกรอง GPF ใช้เวลาในการสังเคราะห์ตัวประมาณการกรองน้อยกว่าตัวกรอง BF อย่างชัดเจน

2. การทดลองถัดมาคือการเลือกจำนวนอนุภาคการกรอง n_f โดยกำหนดให้พารามิเตอร์จริง θ และพารามิเตอร์เริ่มต้น $\theta^{(0)}$ คือ

$$\theta = \{\phi,\ Q,\ \alpha\} = \{0.9,\ 0.5,\ -13.5\}$$

$$\theta^{(0)} = \{\phi^{(0)},\ Q^{(0)},\ \alpha^{(0)}\} = \{0.45,\ 0.25,\ -6.75\}$$

โดยเริ่มต้นกำหนดให้ $n_f = 400 = n_s$ จากนั้นจึงลดค่า n_f เป็นดังนี้ $n_f = 200, 100, 50$ จากการทดลองพบว่า $n_f = 200$ มีค่าเหมาะสมของแบบจำลองนี้เนื่องจากพารามิเตอร์ลู่เข้าเร็วกว่าและให้ค่าพารามิเตอร์ใกล้เคียงกับค่าจริงมากที่สุด

3. สุดท้ายกำหนดให้ $n_f = 200$ จากนั้นจึงปรับจำนวนแนววิถีปรับเรียบ n_s ดังนี้ $n_s = 200, 100, 50, 25, 12$ ผลการทดลองพบว่า $n_s = 100$ ให้ค่าเหมาะสมเนื่องจากมีจำนวนรอบในการลู่เข้าน้อยที่สุด

2. ผลการนำไปใช้งานกับข้อมูลของอัตราแลกเปลี่ยน

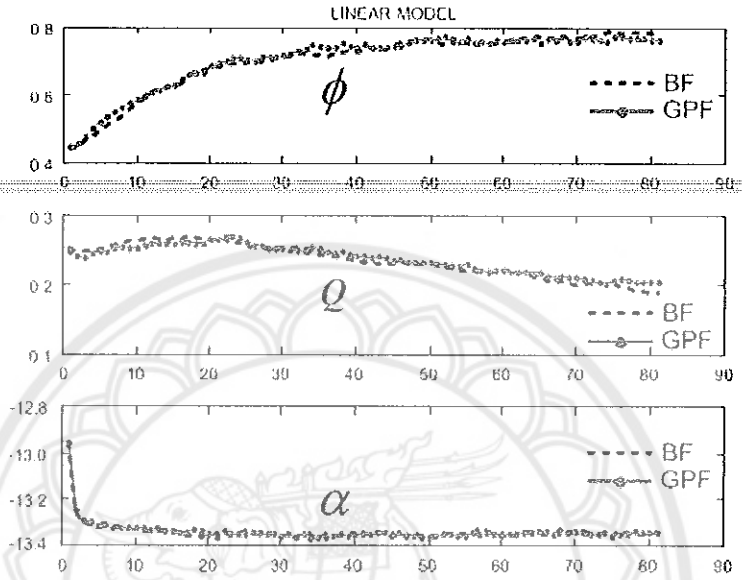
การทดลองในส่วนนี้เป็นการหาพารามิเตอร์ของแบบจำลอง Σ_L โดยใช้ข้อมูลของอัตราแลกเปลี่ยน USD/THB ระหว่างวันที่ 4 ม.ค. 2554 จนถึง 29 เม.ย. 2559 มาเป็นกรณีศึกษาโดยกำหนดให้พารามิเตอร์เริ่มต้น $\theta^{(0)} = \{0.45, 0.25, -6.75\}$ จำนวนอนุภาคการกรอง $n_f = 200$ และจำนวนแนววิถีปรับเรียบ $n_s = 100$ ซึ่งเป็นค่าที่ได้มาจากการทดสอบโปรแกรมในหัวข้อที่ผ่านมาแล้วโดยทำการทดลองซ้ำทั้งหมด 20 ครั้งเพื่อหาค่าเฉลี่ยทางสถิติ

ผลการประมาณพารามิเตอร์ที่ได้แสดงไว้ในตาราง 3 โดยค่าในวงเล็บคือส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์ที่ได้จากการทดลองซ้ำ

ตาราง 3 $\hat{\theta} = \{\hat{\phi},\ \hat{Q},\ \hat{\alpha}\}$ สำหรับ Σ_L ของอัตราแลกเปลี่ยน

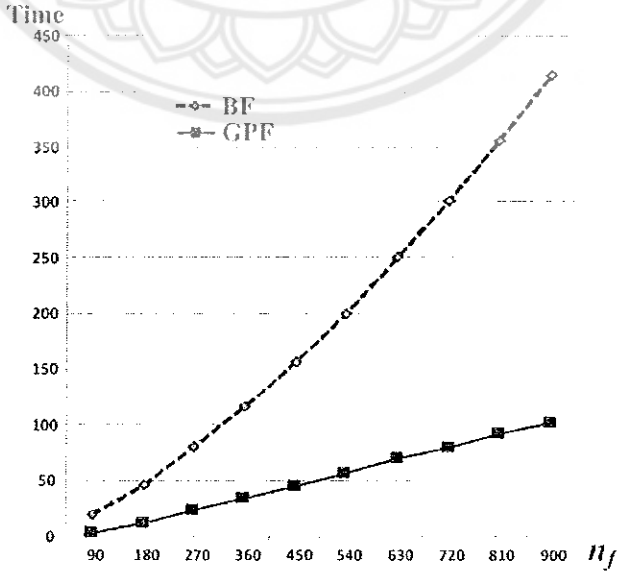
สกุลเงิน	BF			GPF		
	$\hat{\phi}$	$\hat{Q}$	$\hat{\alpha}$	$\hat{\phi}$	$\hat{Q}$	$\hat{\alpha}$
USD	0.7959 (0.0153)	0.1793 (0.0225)	−13.3360 (0.0087)	0.7818 (0.0157)	0.1868 (0.0141)	−13.3406 (0.0068)

กราฟการลู่เข้าของพารามิเตอร์ $\hat{\theta} = \{\hat{\phi}, \hat{Q}, \hat{\alpha}\}$ สำหรับข้อมูลของอัตราแลกเปลี่ยน USD/THB แสดงไว้ในภาพ 9 ซึ่งเห็นได้ว่าพารามิเตอร์ที่ประมาณได้จากตัวกรอง BF และตัวกรอง GPF มีค่าใกล้เคียงกันและลู่เข้าสู่ค่าคงที่ค่าหนึ่ง



ภาพ 9 การลู่เข้าของ $\hat{\theta} = \{\hat{\phi}, \hat{Q}, \hat{\alpha}\}$ สำหรับ Σ_L

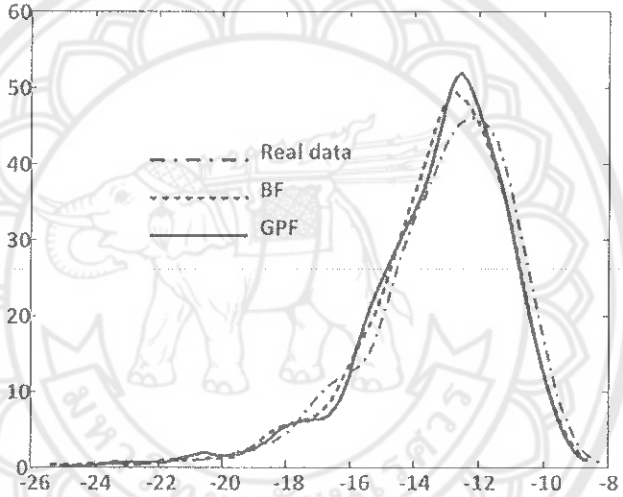
แม้ว่าพารามิเตอร์ที่ประมาณได้จากตัวกรองทั้งสองไม่แตกต่างกันอย่างมีนัยสำคัญแต่หากพิจารณาในแง่ของเวลาพบว่าตัวกรอง GPF ใช้เวลาในการประมวลผลน้อยกว่า ตัวกรอง BF



ภาพ 10 เวลาในการประมวลผลระหว่าง BF และ GPF ของ USD/THB สำหรับ Σ_L

ภาพ 10 แสดงเวลาที่ใช้ประมวลผลในขั้นตอนวิธีการกรองโดยเปรียบเทียบระหว่างการใช้ตัวกรอง BF และตัวกรอง GPF ของอัตราแลกเปลี่ยน USD/THB ซึ่งจะเห็นว่าเมื่อเพิ่มจำนวนอนุภาคการ กรอง n_f มากขึ้น ตัวกรอง BF จะใช้เวลาประมวลผลมากกว่าตัวกรอง GPF อย่างชัดเจน

เมื่อนำพารามิเตอร์ $\hat{\theta} = \{\hat{\phi}, \hat{Q}, \hat{\alpha}\}$ ที่ได้จากตาราง 3 มาจำลองหาผลตอบแทนผ่านทางแบบจำลอง Σ_L แล้วนำมาวาดฟังก์ชันการแจกแจงเพื่อตรวจสอบความคล้ายคลึงกันระหว่างฟังก์ชันการแจกแจงของผลตอบแทนจริงและฟังก์ชันการแจกแจงของผลตอบแทนที่เกิดจากการจำลอง



ภาพ 11 ฟังก์ชันการแจกแจงของผลตอบแทนจริงและผลตอบแทนจาก**การจำลอง** Σ_L

ภาพ 11 แสดงการเปรียบเทียบฟังก์ชันการแจกแจงของผลตอบแทนจริงที่เกิดจากอัตราแลกเปลี่ยน USD/THB กับฟังก์ชันการแจกแจงที่ได้จากการจำลองโดยเลือกใช้พารามิเตอร์ดังนี้

$$\hat{\theta}_{\text{BF}} = \{0.7959,\ 0.1793,\ -13.3360\} \quad \text{และ} \quad \hat{\theta}_{\text{GPF}} = \{0.7818,\ 0.1868,\ -13.3406\}$$

จะเห็นว่ารูปร่างเค้าโครงของฟังก์ชันการแจกแจงมีลักษณะคล้ายคลึงกัน หากพิจารณาคุณสมบัติด้านโมเมนต์ของฟังก์ชันการแจกแจงได้แก่ ค่ามัชฌิม (Mean) ค่าความแปรปรวน (Variance) ค่าความเบ้ (Skewness) และค่าภาวะยอดมน (Kurtosis) ซึ่งเป็นโมเมนต์ที่ 1, 2, 3 และ 4 ตามลำดับ ผลที่ได้แสดงไว้ในตาราง 4

จากคุณสมบัติด้านโมเมนต์ที่ได้พบว่าค่าโมเมนต์ทั้งสี่ของฟังก์ชันการแจกแจงของผล

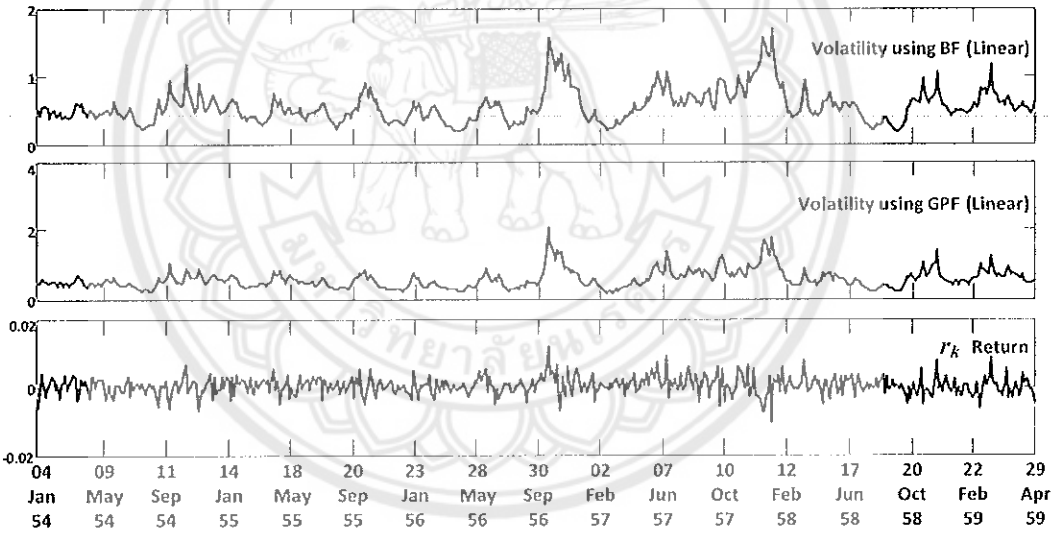
ตาราง 4 คุณสมบัติด้านโมเมนต์ของฟังก์ชันการแจกแจงสำหรับ Σ_L

	ค่ามัชฌิม	ค่าความแปรปรวน	ค่าความเบ้	ค่าภาวะยอดมน
USD/THB	-13.1502	5.4222	-1.2635	5.8110
BF	-13.3230	5.1553	-1.3712	6.5932
GPF	-13.3223	5.1269	-1.4001	6.4573

ตอบแทนจริงและฟังก์ชันการแจกแจงที่ได้จากการจำลองมีความใกล้เคียงกัน จึงอนุมานได้ว่าผลการประมาณพารามิเตอร์ที่ได้มีความถูกต้อง

3. ความผันผวนของอัตราแลกเปลี่ยน

เมื่อการทำงานของโปรแกรมสิ้นสุดลง ให้นำตัวประมาณปรับเรียบ x_k^s ที่ได้ในรอบสุดท้ายมาคำนวณหาความผันผวน $\sigma_k = e^{x_k^s/2}$



ภาพ 12 ค่าความผันผวนและผลตอบแทนของ USD/THB สำหรับ Σ_L

ภาพ 12 แสดงผลตอบแทน r_k และความผันผวนของอัตราแลกเปลี่ยน USD/THB ที่คำนวณได้เมื่อเลือกใช้ตัวกรอง BF และตัวกรอง GPF ซึ่งสังเกตว่าเมื่อผลตอบแทนมีอัตราการเปลี่ยนแปลงค่อนข้างสูง เช่นในช่วงเดือนตุลาคม 2556 หรือช่วงเดือนพฤศจิกายน 2557 ความผันผวนที่สมนัยกับผลตอบแทนในช่วงดังกล่าวมีค่าสูงมากด้วยเช่นกัน ในทางตรงข้ามหากผลตอบแทนมีการเปลี่ยนแปลงเพียงเล็กน้อย ความผันผวนที่เกิดขึ้นในช่วงดังกล่าวค่อนข้างน้อย

บทที่ 5

บทสรุป

สรุปงานวิจัย

งานวิจัยนี้ศึกษาวิธีการประมาณพารามิเตอร์ของแบบจำลอง SV ด้วยขั้นตอนวิธี EM ร่วมกันตัวกรอง GPF และตัวกรอง BF โดยใช้อัตราแลกเปลี่ยน USD/THB ตั้งแต่วันที่ 4 มกราคม พ.ศ. 2554 จนถึง 29 เมษายน พ.ศ. 2559 มาใช้เป็นกรณีศึกษาโดยมีบทสรุปดังนี้

1. จากการทดลองโดยใช้ข้อมูลจากการจำลองพบว่า ทั้งตัวกรอง GPF และตัวกรอง BF สามารถประมาณพารามิเตอร์ของแบบจำลอง SV ได้เป็นอย่างดี

2. เมื่อทดลองโดยใช้ข้อมูลของอัตราแลกเปลี่ยน USD/THB พบว่า พารามิเตอร์ที่ได้จากตัวกรอง GPF และจากตัวกรอง BF ไม่แตกต่างกันมากนักและเมื่อนำพารามิเตอร์ที่ประมาณได้แทนกลับลงในแบบจำลอง SV แล้วทำการจำลองหาฟังก์ชันการแจกแจงของผลตอบแทนพบว่าฟังก์ชันการแจกแจงของผลตอบแทนจริงและฟังก์ชันการแจกแจงของผลตอบแทนที่ได้จากการจำลองคล้ายคลึงกันดังเห็นได้จากภาพ 7 สำหรับแบบจำลองไม่เชิงเส้น หรือภาพ 11 สำหรับแบบจำลองเชิงเส้น นอกจากนี้ค่าโมเมนต์ที่ 1, 2, 3 และ 4 ของฟังก์ชันการแจกแจงทั้งสองไม่แตกต่างกันมากนักดังแสดงไว้ในตารางที่ 2 สำหรับแบบจำลองไม่เชิงเส้น หรือตารางที่ 4 สำหรับแบบจำลองเชิงเส้น

3. แม้ว่าทั้งตัวกรอง GPF และตัวกรอง BF สามารถนำมาใช้ร่วมกับขั้นตอนวิธี EM ในการประมาณพารามิเตอร์ของแบบจำลอง SV ได้เป็นอย่างดีแต่การเลือกใช้ตัวกรอง GPF จะลดเวลาที่ใช้ในการประมวลผลลงได้มากดังเห็นได้จากภาพ 6 สำหรับแบบจำลองไม่เชิงเส้น หรือภาพ 10 สำหรับแบบจำลองเชิงเส้น

4. โดยทั่วไปนิยมเลือกจำนวนแนววิถีปรับเรียบ n_s ให้เท่ากับจำนวนอนุภาคการกรอง n_f สำหรับงานวิจัยชิ้นนี้ได้ทำการทดลองพบว่าวิธีการเลือกดังกล่าวไม่มีความจำเป็น การลด n_s ลงจะช่วยลดเวลาในการประมวลผลลงได้มากโดยไม่กระทบต่อผลการประมาณพารามิเตอร์ที่ได้

แนวทางการพัฒนางานวิจัย

1. ขอบเขตของงานวิจัยนี้เป็นการศึกษากรณีของปริมาณสเกลาร์ นั่นคือโปรแกรมจะทำการประมวลผลได้เมื่อป้อนข้อมูลเพียงชุดเดียว จึงเป็นที่น่าสนใจหากทำการพัฒนาโปรแกรมให้รองรับกับการทำงานของเวกเตอร์-เมทริกซ์ นั่นคือให้โปรแกรมสามารถพิจารณาข้อมูลหลายชุดพร้อมกัน

2. งานวิจัยนี้มุ่งเน้นศึกษาและออกแบบขั้นตอนวิธี EM สำหรับประมาณพารามิเตอร์ของแบบจำลองความผันผวนเชิงเฟ้นสุ่มเท่านั้น อย่างไรก็ดีขั้นตอนวิธีดังกล่าวสามารถนำไปประยุกต์ใช้ได้กับแบบจำลองชนิดอื่นที่มีข้อมูลแบบอนุกรมเวลา เช่น สัญญาณแผ่นดินไหว สัญญาณคลื่นสมอง หรือสัญญาณการเต้นของหัวใจ จึงเป็นแนวทางหนึ่งในการพัฒนางานวิจัยต่อไป

3. โดยทั่วไปนิยมเลือกให้ทั้งจำนวนอนุภาคการกรอง n_f และจำนวนแนววิถีปรับเรียบ n_s มีค่ามากและเลือกให้ $n_f = n_s$ แต่ในงานวิจัยชิ้นนี้พบว่าการลด n_s ลงจะช่วยลดเวลาในการประมวลผลลงได้มาก อย่างไรก็ตามเท่าที่ผู้วิจัยทราบยังไม่พบวิธีการเลือก n_f และ n_s ในเชิงวิเคราะห์ แนวทางการพัฒนางานวิจัยชิ้นนี้คือการหาหลักเกณฑ์เชิงวิเคราะห์ในการเลือก n_f และ n_s ที่เหมาะสมสำหรับแต่ละปัญหา

4. ในงานวิจัยนี้เลือกใช้ทั้งขั้นตอนวิธีของตัวกรอง BF และของ GPF ในการสังเคราะห์ตัวประมาณการกรองซึ่งพบว่า การสังเคราะห์ตัวประมาณการกรอง x_k^i ด้วยตัวกรอง GPF ช่วยลดเวลาในการประมวลผลลงได้อย่างมากเมื่อเทียบกับการใช้ตัวกรอง BF อย่างไรก็ตามเมื่อได้ x_k^i เป็นที่เรียบร้อยแล้วจึงนำมาป้อนเข้าสู่ขั้นตอนวิธีของตัวปรับเรียบ BS-PS เพื่อสังเคราะห์แนววิถีปรับเรียบซึ่งมีลักษณะการทำงานแบบเรียงลำดับ (Serial) หากต้องการจำนวนแนววิถีปรับเรียบ n_s มาก จะต้องใช้เวลาในการประมวลผลมากตามไปด้วย แนวทางในการพัฒนางานวิจัยในส่วนนี้คือการสังเคราะห์ตัวประมาณปรับเรียบ x_k^s ที่ใช้เวลาในการประมวลผลลดลงซึ่งอาจออกแบบให้มีการทำงานแบบขนาน (Parallel) หรือพัฒนาขั้นตอนวิธีของตัวปรับเรียบ SIR-PS ให้ดีขึ้นเพื่อแก้ปัญหาการลดลง (Degeneracy problem) ซึ่งขั้นตอนวิธีนี้ไม่จำเป็นต้องทำการประมวลผลแบบเรียกซ้ำย้อนกลับจึงช่วยลดเวลาในการประมวลผลลงได้

5. การตรวจสอบความถูกต้องของโปรแกรมในงานวิจัยนี้อาศัยการทดสอบสมมติฐานรวมทั้งพิจารณาค่ามัชฌิม (Mean) ค่าความแปรปรวน (Variance) ค่าความเบ้ (Skewness) และค่าภาวะยอดมน (Kurtosis) ซึ่งเป็นการพิจารณาค่าโมเมนต์ของฟังก์ชันการแจกแจง แนวทางหนึ่ง

ในการพัฒนางานวิจัยนี้คือการพิจารณาความแตกต่างระหว่างฟังก์ชันการแจกแจงของข้อมูลจริง และฟังก์ชันการแจกแจงที่เกิดจากการจำลองด้วยโปรแกรมคอมพิวเตอร์ นั่นคือพิจารณาเอนโทรปีสัมพัทธ์ (Relative Entropy : RE) หรือบางครั้งเรียกว่า การลู่เข้าของ Kullback–Leibler (Kullback–Leibler divergence) ซึ่งมีความน่าเชื่อถือมากเนื่องจากเป็นการตรวจสอบในเชิงคณิตศาสตร์

บรรณานุกรม

[1] ตลาดสัญญาซื้อขายล่วงหน้า. (2559). *ตราสารสิทธิ*. สืบค้น 1 พฤษภาคม 2559, จาก http://www.tfex.co.th/th/products/set50futures-spec.html

[2] Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economics, 81*, 637–654.

[3] Merton, R. C. (1973). The theory of rational option pricing. *Bell Journal of Economics, 4*, 141–183.

[4] Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society, 39*(1), 1–38.

[5] Anderson, B. D. O., & Moore, J. B. (2005). *Optimal Filtering*. New York, USA: Dover.

[6] Fridman, M., & Harris, L. (1998). A maximum likelihood approach for non-Gaussian stochastic volatility models. *Journal of Business and Economics Statistics, 1*(3), 284–291.

[7] Shephard, N., & Pitt, M. K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika, 84*(3), 653–667.

[8] Chatterjee, S. (2005). *Application of the Kalman filter for estimating continuous time term structure models–the case of UK and Germany* (Report No. G128RT). Glasgow: Scotland, Department of Economics, University of Glasgow.

[9] Reif, K., Günther, S., Yaz, E., & Unbehauen, R. (1999). Stochastic stability of the discrete-time extended kalman filter. *IEEE Transactions on Automatic Control, 44*(4), 714–728.

[10] Reif, K., & Unbehauen, R. (1999). The extended Kalman filter as an exponential observer for nonlinear systems. *IEEE Transactions on Signal Processing, 47*(8), 2324–2328.

[11] Becerra, V. M., Roberts, P. D., & Griffiths, G. W. (2001). Applying the extended Kalman filter to systems described by nonlinear differential-algebraic equations. *Control Engineering Practice, 9*(3), 267–281.

[12] Julier, S. J., Uhlmann, J. K., & Durrant-Whyte, H. F. (1995). A new approach for filtering nonlinear systems. *Proceeding of the American Control Conference* (pp. 1628–1632). Washington, USA.

[13] Gannot, S. and Moonen, M. (2003). On the application of the unscented Kalman filter to speech processing. *Proceeding of International Workshop on Acoustic Echo Noise Control* (pp. 27–30). Kyoto, Japan.

[14] Kolossa, D., & Haeb-Umbach, R. (2011). *Robust speech recognition of uncertian or missing data: theory and application.* Berlin, Germany: Springer.

[15] Haitao, Z., Gang, D., Junxin, S., & Yujiao, Z. (2013). Unscented Kalman filter and its nonlinear application for tracking a moving target. *Optik-International Journal for Light and Electron Optics, 124*(20), 4468–4471.

[16] Ito, K., & Xiong, K. (2000). Gaussian filter for nonlinear filtering problems. *IEEE Transactions on Automatic Control, 45*(5), 910–927.

[17] Doucet, A., Godsill, S., & Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing, 10*(3), 197–208.

[18] Sankarakrishnan, A., & Billinton, R. (1995). Sequential Monte Carlo simulation for composite power system reliability analysis with time varying loads. *IEEE Transactions on Power Systems, 10*(3), 1540–1545.

[19] Doucet, A., Freitas, N. D., & Gordon, N. (2001). *Sequential monte carlo methods in practice.* New York, USA: Springer.

[20] Moral, P. D., Doucet, A., & Jasra, A. (2006). Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society, 68*(3),411–436.

[21] Sakka, S. (2013). *Bayesian filtering and smoothing*. Cambridge, England: Cambridge University Press.

[22] Gordon, N. J., Salmond, D. J., & Smith, A. F. M. (1993). Novel approach to nonlinear/ non-Gaussian Bayesian state estimation. *IEEE Proceedings F-Radar and Signal Processing, 140*(2), 107–113.

[23] Kotecha, J. H., & Djuric, P. M. (2003). Gaussian Particle Filtering. *IEEE Transactions on Signal Processing, 51*(10), 2592–2601.

[24] Kitagawa, G. (1996). Monte carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics, 5*(1), 1–25.

[25] Godsill, S. J., Doucet, A., & West, M. (2004). Monte Carlo smoothing for nonlinear time series. *Journal of American Statistical Association, 99*(465), 156–168.

[26] Engle, R. F., & Patton, A. J. (2001). What good is a volatility model? *Quantitative Finance, 1*, 237–245.

[27] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econometrica, 50*(4), 987–1008.

[28] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*(3), 307–327.

[29] Sabiruzzaman, Md., Huq, M. Md., Beg, A. R., & Anwar, S. (2010). Modeling and forecasting trading volume index: GARCH versus TGARCH approach. *The Quarterly Review of Economics and Finance, 50*(2), 141–145.

[30] Su, C. (2010). *Application of EGARCH Model to estimate financial volatility of daily returns* (Master's Thesis). Gothenburg, Sweden: University of Gothenburg.

[31] Su, Y. C., Huang, H. C., & Lin, Y. J. (2011). GJR-GARCH model in value-at-risk of financial holdings. *Applied Financial Economics, 21*(4), 1819–1829.

[32] Kim, W., & Linton, O. (2011). Estimation of a Semiparametric IGARCH(1,1) Model. *Econometric Theory, 27*(3), 639–661.

[33] Ali, G. (2013). EGARCH, GJR-GARCH, TGARCH, AVGARCH, NGARCH, IGARCH and APARCH Models for Pathogens at Marine Recreational Sites. *Journal of Statistical and Econometric Methods, 2*(3), 57–73.

[34] Wang, P. (2009). *Financial Econometrics* (2nd ed.). New York, USA: Taylor & Francis.

[35] Diebold, F. X. (2004). The Nobel Memorial Prize for Robert F. Engle. *The Scandinavian Journal of Economics, 106*(2), 165–185.

[36] Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes–a study of daily sugar prices, 1961–79. *In Time Series Analysis: Theory and Practice*, (pp. 203–226) Amsderdam, North-Holland: O. D. Anderson.

[37] Broto, C., & Ruiz, E. (2004). Estimation Methods for Stochastic Volatility Models: A Survey, *Journal of Economic Surveys, 18*(5), 613–649.

[38] Platanioti, K., McCoy, E. J., & Stephens, D. A. (2005). *A Review of Stochastic Volatility: univariate and multivariate models*. London, UK: Imporial College.

[39] Kim, J. (2005). *Parameter estimation in stochastic volatility models with missing data using particle method and em algorithm* (Doctoral Dissertation). Pittsburgh, USA: University of Pittsburgh.

[40] Kim, J., & Stoffer, D. S. (2006). Fitting stochastic volatility models in the presence of irregular sampling via Particle methods and the EM algorithm, *Journal of Time Series Analysis, 29*(5), 811–833.

[41] Huang, S. J., & Yu, J. (2008). An efficient method for maximum likelihood estimation of a stochastic volatility model. *Statistics and Its Interface, 1*, 289–296.

[42] Darrell, B. R., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46*(4), 443--459.

[43] Levitan, E., & Herman, G. T. (1987). A Maximum a Posteriori Probability Expectation Maximization Algorithm for Image Reconstruction in Emission Tomography. *IEEE Transactions on Medical Imaging, 6*(3), 185–192.

[44] Liu, C., & Rubin, D. B. (1994). The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence. *Biometrika, 81*, 633–648.

[45] Ollinger, J. M. (1994). Maximum-likelihood reconstruction of transmission images in emission computed tomography via the EM algorithm. *IEEE Transactions on Medical Imaging, 13*(1), 89–101.

[46] ธนาคารแห่งประเทศไทย. (2559). *อัตราแลกเปลี่ยน*. สืบค้น 1 พฤษภาคม 2559, จาก https://www.bot.or.th

ภาคผนวก

NUEJ
NARESUAN UNIVERSITY
ENGINEERING JOURNAL

การประมาณพารามิเตอร์สำหรับแบบจำลองความผันผวนเชิงเฟ้นสุ่มด้วยขั้นตอนวิธี EM
Parameter Estimation for the Stochastic Volatility Model using the EM Algorithm

ธนภัทร เอี่ยมตาล[1] และ ธนิต มาลากร[2*]

Thanapat Iamtan[1], Tanit Malakorn[2*]

[1]นิสิตปริญญาเอก สาขาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร E-mail: i.thanapat@yahoo.com
[2*]ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร E-mail: tanitm@nu.ac.th

บทคัดย่อ

งานวิจัยฉบับนี้ศึกษาการประยุกต์ใช้ขั้นตอน วิธี Expectation-Maximization Algorithm (EM) ในการประมาณพารามิเตอร์ของแบบจำลองความผันผวนเชิงเฟ้นสุ่ม จากการทดลองพบว่า การใช้วิธีมอนติคาร์โลในขั้นตอนวิธี EM ให้พารามิเตอร์มีค่าใกล้เคียงกับการใช้วิธีคาลมานโดยพารามิเตอร์ที่ได้มีค่าไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ 0.05 เมื่อเทียบกับพารามิเตอร์จริง จากนั้นจึงนำขั้นตอนวิธี EM ร่วมกับทั้งวิธีมอนติคาร์โลและวิธีคาลมานมาใช้ในการประมาณหาพารามิเตอร์ของแบบจำลองความผันผวนเชิงเฟ้นสุ่ม โดยใช้อัตราแลกเปลี่ยนเงินตราต่างประเทศ 5 อัตราเป็นกรณีศึกษา แบบจำลองที่ได้สามารถนำมาใช้ในการพยากรณ์ค่าความผันผวน ซึ่งเป็นตัวแปรต้นในการคำนวณมูลค่าของตราสารสิทธิที่มีอัตราแลกเปลี่ยนเป็นสินทรัพย์อ้างอิงได้

คำสำคัญ: การประมาณพารามิเตอร์, ความผันผวนเชิงเฟ้นสุ่ม,ขั้นตอนวิธี EM, วิธีคาลมาน, วิธีมอนติคาร์โล.

Abstract

This paper presents the application of the Expectation-Maximization Algorithm (EM) in estimating the parameters of the stochastic volatility model. The experimental results showed that the parameters computed by the Monte Carlo method and by the Kalman method in the EM algorithm are slightly different at the 0.05 significance level compared to the true parameters. We then applied the EM algorithm coupled with the Monte Carlo method and the Kalman method to estimate the parameters of the stochastic volatility model using 5 foreign exchange rates as a case study.

The proper model can be used to forecast the volatility of the foreign exchange rate which is one major variable in calculating the value of the options having the exchange rate as an underlying asset.

Keywords: Parameter Estimation, Stochastic Volatility, EM Algorithm, Kalman Method, Monte Carlo Method.

1. บทนำ

ในงานด้านการจัดการโครงการทางวิศวกรรมขนาดใหญ่ต้องมีความเกี่ยวข้องกับการบริหารการเงินแทบทั้งสิ้นการบริหารความเสี่ยงที่เหมาะสมจะช่วยลดความเสี่ยงจากการขาดทุนของโครงการได้เป็นอย่างดี วิธีการหนึ่งของการบริหารความเสี่ยงคือการซื้อสัญญาซื้อขายล่วงหน้า ได้แก่ ฟิวเจอร์ส (Futures) หรือตราสารสิทธิ (Options) เพื่อเป็นการป้องกันการเปลี่ยนแปลงของราคาในอนาคต รวมทั้งสามารถรับประกันได้ว่าจะมีการส่งมอบสินทรัพย์อ้างอิงได้ภายในช่วงเวลาที่กำหนด

การซื้อขายฟิวเจอร์สหรือตราสารสิทธิในประเทศไทยสามารถทำได้ในตลาดสัญญาซื้อขายล่วงหน้า (TFEX) โดยสินทรัพย์อ้างอิงที่ทำการซื้อขายได้ในตลาดดังกล่าว ได้แก่ ดัชนีราคาหลักทรัพย์ หลักทรัพย์ อัตราดอกเบี้ย ทองคำ และ น้ำมัน ทั้งนี้มีเพียงดัชนีราคาหลักทรัพย์ใน SET 50 เท่านั้นที่มีการซื้อขายแบบตราสารสิทธิได้ในขณะนี้ [1]

การคำนวณหามูลค่าที่เหมาะสมของฟิวเจอร์สค่อนข้างง่าย ในขณะที่การคำนวณหามูลค่าที่เหมาะสมของตราสารสิทธิมีความยุ่งยากซับซ้อน ต้องอาศัยสูตรแบลค์-โชวส์ [2,3] (Black-Scholes formula) ซึ่งเป็นผลเฉลยของสมการเชิงอนุพันธ์เชิงเฟ้นสุ่ม การพัฒนาสูตรแบลค์-โชวส์ในช่วงแรกสมมติให้ค่าความผันผวน (volatility) ของสินทรัพย์อ้างอิงเป็นค่าคงที่ σ ต่อมามีการค้นพบว่าค่าความผันผวนมีการแปรเปลี่ยนตามเวลา σ_k ดังนั้นเพื่อให้การคำนวณหามูลค่าที่เหมาะสมของตราสารสิทธิมี

ความแม่นยำมากขึ้นจึงจำเป็นต้องสร้างแบบจำลองสำหรับความผันผวนดังกล่าว

แบบจำลองที่ใช้ในงานวิจัยนี้คือแบบจำลองความผันผวนเชิงเฟ้นสุ่ม (Stochastic volatility model: SV) โดยใช้อัตราแลกเปลี่ยนเงินตราต่างประเทศต่อสกุลเงินบาทไทยที่สำคัญ 5 อัตรา ได้แก่สกุลเงินดอลลาร์สหรัฐ (USD) สกุลเงินปอนด์อังกฤษ (GBP) สกุลเงินยูโร (EUR) สกุลเงินหยวนจีน (CNY) และสกุลเงินดอลลาร์สิงคโปร์ (SGD) มาเป็นกรณีศึกษา เนื่องจากตราสารสิทธิที่มีอัตราแลกเปลี่ยนเงินตราต่างประเทศเป็นสินทรัพย์อ้างอิงมีโอกาสที่จะเปิดทำการซื้อขายได้ในประเทศไทย

สิ่งสำคัญในการสร้างแบบจำลองคือการประมาณพารามิเตอร์ของแบบจำลองโดยอาศัยข้อมูลที่ได้จากการเก็บตัวอย่าง งานวิจัยชิ้นนี้เลือกใช้ขั้นตอนวิธี EM (EM algorithm) ที่ถูกพัฒนาขึ้นโดย Dempster et al. [4] ในปี ค.ศ. 1977 ในการประมาณพารามิเตอร์โดยมีวัตถุประสงค์เพื่อให้ฟังก์ชันควรจะเป็น (Likelihood function) มีค่าสูงสุด

ขั้นตอนวิธี EM เป็นวิธีทำซ้ำประกอบด้วย 2 ขั้นตอนย่อย นั่นคือขั้นตอน E (Expectation step) ซึ่งเป็นขั้นตอนในการหาค่าคาดหมายความควรจะเป็นแบบสมบูรณ์ (Complete likelihood) เมื่อเทียบกับการแจกแจงปรับเรียบ (Smooth distribution) และขั้นตอน M (Maximization step) ซึ่งเป็นขั้นตอนในการหาพารามิเตอร์เพื่อทำให้ฟังก์ชันควรจะเป็นมีค่าสูงสุด

เมื่อนำขั้นตอนวิธี EM มาประยุกต์ใช้ในการประมาณหาพารามิเตอร์ของแบบจำลอง SV ซึ่งบรรยายในรูปแบบปริภูมิสถานะ จึงจำเป็นต้องทำการประมาณตัวแปรสถานะปรับเรียบ x_k^s ที่ได้จากการแจกแจงปรับเรียบเพื่อนำมาคำนวณหาฟังก์ชันควรจะเป็นแบบสมบูรณ์ ในกรณีของแบบจำลองเชิงเส้นที่มีการแจกแจงแบบปกติ นิยมเลือกใช้วิธีคาลมาน (Kalman Method : KM) ซึ่งเป็นการทำงานร่วมกันระหว่างตัวกรองคาลมาน (Kalman filter : KF) และตัวปรับเรียบคาลมาน (Kalman smoother : KS) ในการประมาณ x_k^s เนื่องจากวิธี KM จัดได้ว่าเป็นวิธีการประมาณเหมาะที่สุด (Optimal estimation) [5,6] ซึ่งทำให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยระหว่างตัวแปรสถานะจริงและตัวแปรสถานะที่ประมาณขึ้นมีค่าต่ำสุด นอกจากนี้วิธี KM ยังมีความเรียบง่ายและมีผลเฉลยในรูปแบบปิด (Closed form)

สำหรับแบบจำลองไม่เชิงเส้นหรือแบบจำลองที่มีการแจกแจงแบบอื่นที่ไม่ใช่การแจกแจงปกติ การใช้วิธี KM ในขั้นตอนวิธี EM จะนำไปสู่ความควรจะเป็นเสมือนสูงสุด (Quasi maximum likelihood) แทน [7,8] ด้วยเหตุนี้วิธีมอนติคาร์โล (Monte Carlo method : MCM) ซึ่งอาศัยวิธีการแก้ปัญหาแบบศึกษำนึก (Heuristic approach) จึงเข้ามามีบทบาทในการประมาณ x_k^s แทนวิธี KM โดยใช้หลักการชักตัวอย่างจากฟังก์ชันการแจกแจง หาค่าถ่วงน้ำหนักพร้อมทำการปรับตัวอย่างใหม่ ทั้งนี้ตัวอย่างที่ได้จากการชักตัวอย่างนิยมเรียกว่าอนุภาค (Particle) ด้วยเหตุนี้ตัวกรองและตัวปรับเรียบที่ถูกออกแบบด้วยวิธี MCM จึงเรียกว่าตัวกรองอนุภาค (Particle Filter : PF) และตัวปรับเรียบอนุภาค (Particle Smoother : PS) ตามลำดับ

วัตถุประสงค์ของงานวิจัยนี้คือการประมาณหาพารามิเตอร์ในแบบจำลอง SV จากข้อมูลที่สร้างขึ้นด้วยขั้นตอนวิธี EM โดยใช้ทั้งวิธี KM และวิธี MCM ในการประมาณตัวแปรสถานะปรับเรียบ x_k^s จากนั้นจึงนำมาประยุกต์ใช้ประมาณพารามิเตอร์ในแบบจำลอง SV ของอัตราแลกเปลี่ยนเงินตราต่างประเทศทั้ง 5 อัตรา

2. แบบจำลองความผันผวนเชิงเฟ้นสุ่ม

ให้ R_k เป็นอัตราแลกเปลี่ยน ณ เวลาที่ k และให้ r_k เป็นผลตอบแทนแบบลอการิทึมของ R_k นั่นคือ $r_k = \log\left(\frac{R_k}{R_{k-1}}\right)$ ซึ่งสอดคล้องกับแบบจำลอง SV ที่นำเสนอโดย Taylor [9] ดังนี้

$$r_k = \sigma_k \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0,1) \tag{1}$$

เมื่อ σ_k คือความผันผวนของ R_k โดยที่ σ_k^2 มีการแจกแจงเป็นแบบปกติเชิงลอการิทึม (Log-normally distribution) นั่นคือ ต้องมีตัวแปรสุ่ม x_k ที่มีการแจกแจงแบบปกติโดยที่ $x_k = \log\sigma_k^2$ ดังนั้นแบบจำลองในสมการที่ (1) จึงเขียนใหม่ได้เป็น

$$r_k = \exp\left(\frac{x_k}{2}\right)\varepsilon_k \tag{2}$$

โดยทั่วไปนิยมกำหนดให้ x_k สอดคล้องสมการการถดถอยอัตตะ (Autoregression) AR (1) ที่มีสัญญาณรบกวนแบบปกติดังนี้

$$x_k = \phi x_{k-1} + c + w_k, \; w_k \sim \mathcal{N}(0,Q) \tag{3}$$

โดยที่ $|\phi| < 1$ และ c เป็นค่าคงที่

เพื่อกำจัดค่าคงที่ c ออกจากสมการที่ (3) จึงแทรก β เข้าไปในสมการที่ (2) ดังนั้นแบบจำลอง SV จึงเขียนใหม่ได้เป็น

$$\Sigma_{NL} := \begin{cases} x_k = \phi x_{k-1} + w_k \\ r_k = \beta \exp\left(\frac{x_k}{2}\right)\varepsilon_k \end{cases} \text{ เมื่อ } k = 1,\ldots,N \tag{4}$$

เนื่องจาก Σ_{NL} เป็นแบบจำลองแบบไม่เชิงเส้นซึ่งไม่สามารถประยุกต์ใช้วิธี KM ได้จึงทำการแปลงแบบจำลองดังกล่าวให้อยู่ในรูปแบบเชิงเส้นโดยอาศัยฟังก์ชันลอการิทึมจึงได้ว่า

$$\Sigma_L := \begin{cases} x_k = \phi x_{k-1} + w_k \\ y_k = \alpha + x_k + v_k \end{cases} \text{ เมื่อ } k = 1,\ldots,N \tag{5}$$

เมื่อ $y_k = \log r_k^2$, $\alpha = \log\beta^2 + E[\log\varepsilon_k^2]$ และ $v_k = \log\varepsilon_k^2 - E[\log\varepsilon_k^2]$

แม้ว่าแบบจำลองที่ได้จะเป็นแบบจำลองเชิงเส้นแต่ยังคงไม่สามารถประยุกต์ใช้วิธี KM ได้เนื่องจากผลจากการแปลงทำให้ v_k มีการแจกแจงแบบลอการิทึมไคกำลังสองที่มีค่ามัชฌิมเป็น

และมีความแปรปรวนเท่ากับ $0.5\pi^2$ อย่างไรก็ตามพบว่ามีงานวิจัยหลายชิ้นที่นำวิธี KM มาใช้กับแบบจำลองในสมการที่ (5) โดยสมมติให้ v_k มีการแจกแจงปกติที่มีค่ามัชฌิมและความแปรปรวนเป็น 0 และ $0.5\pi^2$ ตามลำดับเช่นในงานวิจัยของ [10-13] หากแต่พารามิเตอร์ที่คำนวณได้จะทำให้เกิดความควรจะเป็นเสมือนสูงสุดเท่านั้น

เพื่อช่วยแก้ปัญหาดังกล่าว วิธี MCM จึงถูกเลือกนำมาใช้แทนวิธี KM เนื่องจากวิธี MCM มีความยืดหยุ่นสูง อีกทั้งไม่อาศัยสมมติฐานของความเป็นเกาส์เซียนและ/หรือความเป็นเชิงเส้นของแบบจำลองอีกด้วย ในงานวิจัยนี้จึงเลือกตัวกรองปลุกเครื่อง (Bootstrap filter: BF) ร่วมกับตัวปรับเรียบอนุภาคแบบย้อนกลับ (Backward-simulation particle smoother: BS) ซึ่งได้พัฒนามาจาก PF และ PS มาใช้แทนวิธี KM ในการหาพารามิเตอร์ $\theta = (\phi, Q, \alpha)$ ของแบบจำลองเชิงเส้น Σ_L ดังเห็นได้จากในงานวิจัยของ [14-16]

3. วิธีคาลมาน

วิธี KM เป็นกระบวนการในการประมาณหาตัวแปรสถานะหรือตัวแปรแฝง (Latent variable) ของแบบจำลองเชิงเส้นที่มีการแจกแจงปกติ เพื่อให้สามารถประยุกต์ใช้วิธี KM กับแบบจำลอง Σ_L ในสมการที่ (5) ซึ่งการแจกแจงของ v_k เป็นแบบลอการิทึมไคกำลังสองจึงต้องประมาณการแจกแจงของ v_k ด้วยการแจกแจงปกติ นั่นคือ $v_k \sim \mathcal{N}(0, 0.5\pi^2)$

ในที่นี้สมมติให้ตัวแปรสถานะเริ่มต้นมีการแจกแจงปกติ $x_0 \sim \mathcal{N}(\mu_0^{f-}, P_0^{f-})$ และให้ $y_{1:N} = \{y_1, \ldots, y_N\}$ แทนเซตของชุดข้อมูล

3.1 ตัวกรองคาลมาน

KF เป็นกระบวนการเรียกซ้ำแบบไปข้างหน้าซึ่งประกอบด้วย 2 ขั้นตอนดังนี้

1. ขั้นตอนเริ่มต้น: เมื่อ $k = 0$ เป็นการกำหนดเงื่อนไขเริ่มต้นจากการแจกแจงก่อน (Prior distribution)

$$x_0 \sim \mathcal{N}(\mu_0^{f-}, P_0^{f-}) \quad (6)$$

2. ขั้นตอนการทำซ้ำ: เมื่อ $k = 1, \ldots, N$ เป็นการทำงานสลับกันระหว่างขั้นตอนการทำนายและขั้นตอนการปรับค่าดังนี้

ขั้นตอนการทำนาย (Prediction Step) เป็นการหาการแจกแจงการทำนาย

$$x_k^{f-} \sim p(x_k \mid y_{1:k-1}) = \mathcal{N}(\mu_k^{f-}, P_k^{f-}) \quad (7)$$

โดยที่

$$\mu_k^{f-} = \phi \mu_{k-1}^{f+} \quad (8)$$

$$P_k^{f-} = \phi^2 P_{k-1}^{f+} + Q \quad (9)$$

ขั้นตอนการปรับค่า (Update Step) เป็นการหาการแจกแจงการกรอง

$$x_k^{f+} \sim p(x_k \mid y_{1:k}) = \mathcal{N}(\mu_k^{f+}, P_k^{f+}) \quad (10)$$

โดยที่

$$K_k = P_k^{f-} / (P_k^{f-} + 0.5\pi^2) \quad (11)$$

$$\mu_k^{f+} = \mu_k^{f-} + K_k (y_k - \mu_k^{f-} - \alpha) \quad (12)$$

$$P_k^{f+} = (1 - K_k) P_k^{f-} \quad (13)$$

ทั้งนี้ตัวแปรสถานะการกรองคือ

$$\hat{x}_k^f = E[x_k \mid y_{1:k}] = \mu_k^{f+} \quad (14)$$

3.2 ตัวปรับเรียบคาลมาน

KS เป็นกระบวนการเรียกซ้ำแบบย้อนกลับโดยอาศัยค่าที่คำนวณได้จาก KF ซึ่งประกอบด้วย 2 ขั้นตอนดังนี้

1. ขั้นตอนสุดท้าย: เมื่อ $k = N$ เป็นการกำหนดการแจกแจงที่ได้มาจาก KF

$$x_N^s \sim p(x_N) = \mathcal{N}(\mu_N^{f+}, P_N^{f+}) = \mathcal{N}(\mu_N^s, P_N^s) \quad (15)$$

2. ขั้นตอนการทำซ้ำ: เมื่อ $k = N-1, \ldots, 0$ เป็นการคำนวณหาการแจกแจงปรับเรียบ

$$x_k^s \sim p(x_k \mid y_{1:N}) = \mathcal{N}(\mu_k^s, P_k^s) \quad (16)$$

โดยที่

$$J_k = \phi P_k^{f+} / (P_{k+1}^{f-}) \quad (17)$$

$$\mu_k^s = \mu_k^{f+} + J_k (\mu_{k+1}^s - \mu_{k+1}^{f-}) \quad (18)$$

$$P_k^s = P_k^{f+} + J_k^2 (P_{k+1}^s - P_{k+1}^{f-}) \quad (19)$$

ทั้งนี้ตัวแปรสถานะการปรับเรียบคือ

$$\hat{x}_k^s = E[x_k \mid y_{1:N}] = \mu_k^s \quad (20)$$

4. วิธีมอนติคาร์โล

วิธี MCM อาศัยการแก้ปัญหาแบบศึกษาสำนึกโดยการชักตัวอย่างจากฟังก์ชันการแจกแจงและหาค่าถ่วงน้ำหนักเพื่อทำการประมาณปริมาณที่ต้องการทราบด้วยการหาค่าเฉลี่ยทางสถิติแทนการคำนวณดังเช่นในวิธี KM ในที่นี้จะนำเสนอ BF และ BS ซึ่งมีโครงสร้างการทำงานคล้ายคลึงกับ KF และ KS ตามลำดับโดยมีรายละเอียดดังนี้

4.1 ตัวกรองการปลุกเครื่อง

BF ได้ถูกพัฒนามาจาก PF ซึ่งประกอบด้วย 2 ขั้นตอนดังนี้

1. ขั้นตอนเริ่มต้น: เมื่อ $k = 0$ ทำการชักตัวอย่างอนุภาคเริ่มต้นจากการแจกแจงก่อน

$$x_0^{f(i)} \sim \mathcal{N}(\mu_0^f, P_0^f) \text{ เมื่อ } i = 1, \ldots, M \quad (21)$$

และให้ค่าถ่วงน้ำหนักเริ่มต้นของแต่ละอนุภาคเป็น

$$\omega_0^{f(i)} = 1/M \text{ เมื่อ } i = 1, \ldots, M \quad (22)$$

เมื่อ M คือจำนวนของอนุภาค

2. ขั้นตอนการทำซ้ำ: เมื่อ $k = 1, \ldots, N$ เป็นการทำงานสลับกันระหว่างขั้นตอนการทำนาย ขั้นตอนการหาค่าถ่วงน้ำหนักและขั้นตอนการปรับอนุภาคใหม่ดังนี้

ขั้นตอนการทำนายเป็นการหาอนุภาคที่เวลาถัดไปโดยอาศัยสมการสถานะ นั่นคือ

$$x_k^{f(i)} = \phi x_{k-1}^{f(i)} + w_k^{f(i)} \text{ เมื่อ } i = 1,\ldots,M \quad (23)$$

ขั้นตอนการหาค่าถ่วงน้ำหนักเป็นการหาการแจกแจงแบบมีเงื่อนไขของสัญญาณออก นั่นคือ

$$\begin{aligned}\tilde{\omega}_k^{f(i)} &\sim p(y_k \mid x_k^{f(i)}) \\ &\propto \exp\left(\frac{\gamma_k}{2}\right)\exp\left(-\frac{1}{2}\exp(\gamma_k)\right)\end{aligned} \quad (24)$$

โดยที่ $\gamma_k = y_k - \alpha - x_k^{f(i)} + E\left[\log \varepsilon_k^2\right]$

จากนั้นจึงนำมาทำให้เป็นบรรทัดฐานเพื่อให้ผลรวมค่าถ่วงน้ำหนักมีค่าเป็น 1 ดังนี้

$$\omega_k^{f(i)} = \frac{\tilde{\omega}_k^{f(i)}}{\sum_{i=1}^{M} \tilde{\omega}_k^{f(i)}} \quad (25)$$

ขั้นตอนการปรับอนุภาคใหม่เป็นขั้นตอนที่ใช้เพื่อกำจัดอนุภาค $x_k^{f(i)}$ ที่มีค่าถ่วงน้ำหนักน้อยมากออกไปแล้วนำอนุภาคที่มีค่าถ่วงน้ำหนักมากมาใส่แทนที่ซึ่งเป็นวิธีลดปัญหาการลดลง (Degeneracy problem)

ทั้งนี้ตัวแปรสถานะการกรองคำนวณได้จากผลรวมเชิงเส้นของอนุภาคการกรอง นั่นคือ

$$x_k^f = \sum_{i=1}^{M} \omega_k^{f(i)} x_k^{f(i)} \quad (26)$$

4.2 ตัวปรับเรียบอนุภาคแบบย้อนกลับ

BS เป็นกระบวนการเรียกซ้ำแบบย้อนกลับโดยอาศัยค่าที่คำนวณได้จาก BF ซึ่งประกอบด้วย 2 ขั้นตอนดังนี้

1. ขั้นตอนสุดท้าย: เมื่อ $k = N$ ให้พิจารณาค่าถ่วงน้ำหนัก $\omega_N^{f(i)}$ เป็นความน่าจะเป็นในการเลือกอนุภาคการกรองที่ตำแหน่ง i มาเป็นอนุภาคการปรับเรียบที่ตำแหน่ง j นั่นคือ

$$x_N^{s(j)} = x_N^{f(i)} \text{ ด้วยความน่าจะเป็น } \omega_N^{f(i)} \quad (27)$$

สำหรับทุกค่า $i, j = 1,\ldots,M$

2. ขั้นตอนการทำซ้ำ: เมื่อ $k = N-1,\ldots,0$ เป็นการคำนวณค่าถ่วงน้ำหนักใหม่ $\omega_k^{s(i)}$ จาก

$$\omega_k^{s(i)} \propto \omega_k^{f(i)} p(x_{k+1}^{s(j)} \mid x_k^{f(i)}) = \mathcal{N}(\phi x_k^{f(i)}, Q) \quad (28)$$

เมื่อ $i, j = 1,\ldots,M$

จากนั้นจึงพิจารณาค่าถ่วงน้ำหนัก $\omega_k^{s(i)}$ ที่ได้เป็นความน่าจะเป็นในการเลือกอนุภาคการกรองที่ตำแหน่ง i ให้มาเป็น

$$x_k^{s(j)} = x_k^{f(i)} \text{ ด้วยความน่าจะเป็น } \omega_k^{s(i)} \quad (29)$$

สำหรับทุกค่า $i, j = 1,\ldots,M$

ทั้งนี้ตัวแปรสถานะการปรับเรียบคำนวณได้จากผลรวมเชิงเส้นของอนุภาคการปรับเรียบ นั่นคือ

$$x_k^s = \sum_{i=1}^{M} \omega_k^{s(i)} x_k^{s(i)} \quad (30)$$

5. ขั้นตอนวิธี EM

วิธีการกรองและการปรับเรียบในหัวข้อที่ผ่านมาต้องอาศัยสมมติฐานที่ว่าผู้ออกแบบทราบพารามิเตอร์ทุกตัวของแบบจำลอง อย่างไรก็ตามสิ่งที่ผู้ออกแบบทราบมีเพียงข้อมูลที่วัดได้เท่านั้น ด้วยเหตุนี้จึงจำเป็นต้องประมาณพารามิเตอร์ขึ้นมาเป็นลำดับแรก จากนั้นจึงทำการประมาณตัวแปรสถานะในลำดับถัดไป

เกณฑ์ในการชี้วัดถึงความแม่นยำในการประมาณพารามิเตอร์ที่เลือกใช้ในงานวิจัยนี้คือวิธีความควรจะเป็นสูงสุด (Maximum likelihood: ML) โดยอาศัยขั้นตอนวิธี EM โดยมีรายละเอียดดังนี้

ให้ θ แทนพารามิเตอร์ของระบบแล้ว ความควรจะเป็นของพารามิเตอร์แบบลอการิทึมคือ

$$\log p(y_{1:N} \mid \theta) := \mathcal{L}(\theta) \quad (31)$$

วัตถุประสงค์ของวิธี ML คือการหา $\hat{\theta}^{ML}$ ที่ทำให้ $\mathcal{L}(\hat{\theta}^{ML})$ ในสมการที่ (26) มีค่าสูงสุด นั่นคือ

$$\hat{\theta}^{ML} = \arg\max_{\theta} \log p(y_{1:N} \mid \theta) \quad (32)$$

โดยทั่วไปการคำนวณหาค่าเหมาะที่สุดข้างต้นต้องอาศัยวิธีเชิงตัวเลข Dempster et al. [1] จึงได้พัฒนาขั้นตอนวิธี EM เพื่อใช้ในการประมาณหา $\hat{\theta}^{ML}$ อย่างไรก็ตามขั้นตอนวิธี EM ไม่ได้คำนวณหาค่าเหมาะที่สุดดังในสมการที่ (32) โดยตรง หากแต่อาศัยการทำให้ขอบเขตล่างของ $\mathcal{L}(\theta)$ มีค่ามากขึ้นในแต่ละรอบของการทำซ้ำจึงส่งผลทำให้ $\mathcal{L}(\theta)$ มีค่ามากสูงขึ้นตามไปนั่นเอง

ดังนั้นในขั้นตอนวิธี EM สมการที่ (32) จึงเปลี่ยนเป็น

$$\hat{\theta}^{(k+1)} = \arg\max_{\theta} E[\, \mathcal{L}(\theta, \hat{\theta}^{(k)})] \quad (33)$$

เมื่อ

$$\mathcal{L}(\theta, \hat{\theta}^{(k)}) = \log p(x_{0:N}^s, y_{1:N} \mid \theta) \quad (34)$$

จากคุณสมบัติของมาร์คอฟจะได้ว่า

$$\begin{aligned}\mathcal{L}(\theta, \hat{\theta}^{(k)}) &= \log p(x_0^s \mid \theta) + \sum_{k=1}^{N} \log p(x_k^s \mid x_{k-1}^s, \theta) \\ &\quad + \sum_{k=1}^{N} \log p(y_k \mid x_k^s, \theta)\end{aligned} \quad (35)$$

ดังนั้นขั้นตอนวิธี EM จึงนำมาสรุปไว้ดังนี้

1. ขั้นตอนเริ่มต้น: กำหนดค่าเริ่มต้น $\hat{\theta}^{(0)}$

2. ขั้นตอนการทำซ้ำ: เมื่อ $k = 1,\ldots,N$

ขั้นตอน E คำนวณหา $(\theta, \hat{\theta}^{(k)})$ ตามสมการที่ (34)

ขั้นตอน M คำนวณหา $\hat{\theta}^{(k+1)}$ ตามสมการที่ (33)

5.1 ขั้นตอนวิธี EM ร่วมกับวิธี KM

การประยุกต์ขั้นตอนวิธี EM ร่วมกับวิธี KM ในการประมาณพารามิเตอร์ $\theta = (\phi, Q, \alpha)$ ของแบบจำลอง Σ_L ในสมการที่ (5) ต้องพิจารณา v_k ให้มีการแจกแจงเป็นแบบปกติ

นั่นคือ $v_k \sim \mathcal{N}(0,0.5\pi^2)$ ดังนั้น $\mathcal{L}(\theta,\hat{\theta}^{(k)})$ ในสมการที่ (35) จึงเขียนได้เป็น

$$\mathcal{L}(\theta,\hat{\theta}^{(k)}) = -\frac{1}{2}\log P_0 - \frac{1}{2}\frac{(x_0^s-\mu_0)^2}{P_0} - \frac{1}{2}N\log Q - \frac{1}{2}\sum_{k=1}^{N}\frac{(x_k^s-\phi x_{k-1}^s)^2}{Q} - \frac{1}{2}\sum_{k=1}^{N}\frac{(y_k-x_k^s-\alpha)^2}{0.5\pi^2} - C \quad (36)$$

โดยที่ C เป็นผลรวมของพจน์ของค่าคงที่ซึ่งไม่มีผลต่อการคำนวณหาค่าสูงสุด

จากนั้นจึงคำนวณหา $\hat{\theta}$ ที่ทำให้ $E[\mathcal{L}(\theta,\hat{\theta}^{(k)})]$ มีค่าสูงสุดโดยการหาอนุพันธ์ย่อยเทียบกับพารามิเตอร์แต่ละตัวที่ต้องการทราบจึงได้ว่า

$$\hat{\phi} = \frac{\sum_{k=1}^{N}\left[\mu_k^s\mu_{k-1}^s + P_{k,k-1}\right]}{\sum_{k=1}^{N}\left[(\mu_{k-1}^s)^2 + P_{k-1}\right]} \quad (37)$$

$$\hat{Q} = \frac{\sum_{k=1}^{N}\left[(\mu_k^s)^2 + P_k\right]}{N} - \frac{\sum_{k=1}^{N}\hat{\phi}\left[\mu_k^s\mu_{k-1}^s + P_{k,k-1}\right]}{N} \quad (38)$$

$$\hat{\alpha} = \frac{1}{N}\sum_{k=1}^{N}\left[y_k - \mu_k^s\right] \quad (39)$$

โดยที่ $\mu_k^s = E[x_k^s \mid y_{1:N}]$, $P_k = E\left[(x_k^s-\mu_k^s)^2 \mid y_{1:N}\right]$ และ $P_{k,l} = E\left[(x_k^s-\mu_k^s)(x_l^s-\mu_l^s) \mid y_{1:N}\right]$

5.2 ขั้นตอนวิธี EM ร่วมกับวิธี MCM

เนื่องจาก v_k ของแบบจำลองในสมการที่ (5) มีการแจกแจกแบบลอการิทึมไคกำลังสอง ดังนั้น $\mathcal{L}(\theta,\hat{\theta}^{(k)})$ ในสมการที่ (35) จึงเขียนได้เป็น

$$\mathcal{L}(\theta,\hat{\theta}^{(k)}) = \frac{1}{M}\sum_{j=1}^{M}\left\{-\frac{1}{2}\log P_0 - \frac{1}{2}\frac{(x_0^s-\mu_0)^2}{P_0} - \frac{1}{2}N\log Q - \frac{1}{2}\sum_{k=1}^{N}\frac{(x_k^s-\phi x_{k-1}^s)^2}{Q} - \frac{1}{2}\sum_{k=1}^{N}[\exp(\gamma_k)-\gamma_k] - C\right\} \quad (40)$$

โดยที่ C เป็นผลรวมของพจน์ของค่าคงที่ซึ่งไม่มีผลต่อการคำนวณหาค่าสูงสุด

จากนั้นจึงคำนวณหา $\hat{\theta}$ ที่ทำให้ $E[\mathcal{L}(\theta,\hat{\theta}^{(k)})]$ มีค่าสูงสุดโดยการหาอนุพันธ์ย่อยเทียบกับพารามิเตอร์แต่ละตัวที่ต้องการทราบจึงได้ว่า

$$\hat{\phi} = \frac{\sum_{k=1}^{N}E\left[x_k^s x_{k-1}^s \mid y_{1:N}\right]}{\sum_{k=1}^{N}E\left[(x_{k-1}^s)^2 \mid y_{1:N}\right]} \quad (41)$$

$$\hat{Q} = \frac{\sum_{k=1}^{N}E\left[(x_{k+1}^s - \hat{\phi}x_k^s)^2 \mid y_{1:N}\right]}{N} \quad (42)$$

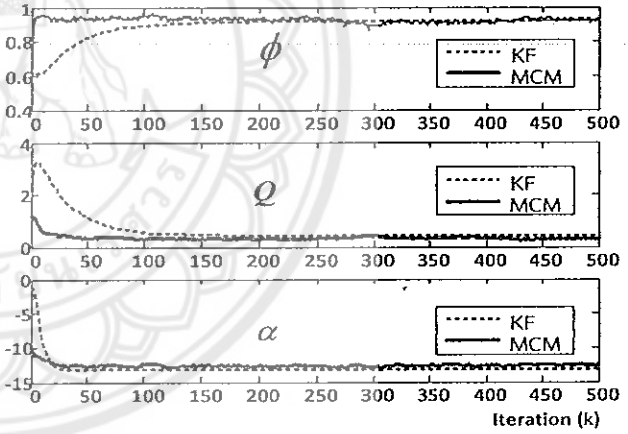
$$\hat{\alpha} = \log\left[\frac{\sum_{k=1}^{N}E\left[\exp\left(y_k - \mu_k^s + E\left[\log\varepsilon_k^2\right]\right) \mid y_{1:N}\right]}{N}\right] \quad (43)$$

6. ผลการทดลอง

ในส่วนนี้เป็นการทดลองเพื่อคำนวณหาพารามิเตอร์ของแบบจำลอง SV ในสมการที่ (5) โดยใช้ชุดข้อมูล 2 ชนิด คือ 1. ข้อมูลที่สร้างขึ้น (Simulated Data) ด้วยแบบจำลอง SV เพื่อใช้ในการทดสอบโปรแกรมที่ได้พัฒนาขึ้นและ 2. ข้อมูลของอัตราแลกเปลี่ยนเงินตราต่างประเทศ 5 อัตรา

6.1 ข้อมูลที่สร้างขึ้น

ในส่วนนี้เป็นการทดสอบความแม่นยำของโปรแกรมที่ได้พัฒนาขึ้นโดยเริ่มจากการสร้างข้อมูลด้วยสมการที่ (5) เมื่อกำหนดค่าพารามิเตอร์ ϕ, Q, α เป็น 0.9, 0.5 และ -13.5 ตามลำดับ จากนั้นจึงนำข้อมูลที่ได้ป้อนเข้าสู่โปรแกรมโดยทำการทดลองซ้ำทั้งหมด 30 ครั้งเพื่อหาค่าเฉลี่ยของพารามิเตอร์ จากการทดลองพบว่าพารามิเตอร์ที่ได้จากการประมาณด้วยวิธี KM และวิธี MCM ลู่เข้าใกล้ค่าพารามิเตอร์จริงดังรูปที่ 1 โดยเส้นประจุดและเส้นทึบแสดงแนววิถีของพารามิเตอร์ที่ได้จากวิธี KM และวิธี MCM ตามลำดับ



รูปที่ 1 การลู่เข้าของพารามิเตอร์

จากนั้นจึงทำการทดสอบทางสถิติด้วยวิธีการทดสอบแบบที (t-test) พบว่าพารามิเตอร์ที่ได้จากวิธีการทั้งสองให้คำตอบที่ไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ 0.05 เมื่อเทียบกับพารามิเตอร์จริงซึ่งผลทดลองที่ได้แสดงไว้ในตารางที่ 1

ตารางที่ 1 ค่าพารามิเตอร์ที่ได้จากการทดสอบ

วิธีการ	ϕ	Q	α
MCM	0.8958	0.4884	-13.5286
(S.D.)	(0.0210)	(0.1017)	(0.0990)
P-Value	0.2806	0.5354	0.3848
KM	0.8992	0.4907	-13.4962
(S.D.)	(0.0401)	(0.2154)	(0.3999)
P-Value	0.9109	0.8149	0.9591

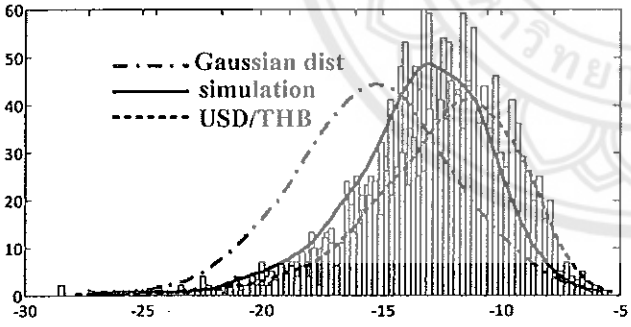
6.2 ข้อมูลอัตราแลกเปลี่ยน

ในงานวิจัยนี้เลือกใช้อัตราแลกเปลี่ยนเงินตราระหว่างประเทศที่สำคัญ 5 อัตรามาหาพารามิเตอร์โดยใช้ทั้งวิธี MCM และวิธี KM โดยเก็บข้อมูลตั้งแต่วันที่ 4 ม.ค. 2554 จนถึง 29 เม.ย. 2559 ค่าพารามิเตอร์ที่ประมาณได้แสดงไว้ในตารางที่ 2

ตารางที่ 2 ค่าพารามิเตอร์ของอัตราแลกเปลี่ยน 5 อัตรา

สกุลเงิน	วิธีการ	ϕ	Q	α
USD	MCM	0.7959	0.1793	-13.3360
	KM	0.3263	0.2886	-13.3625
GBP	MCM	0.6455	0.1893	-12.1771
	KM	0.4282	0.2851	-12.1362
EUR	MCM	0.8362	0.2022	-11.9219
	KM	0.9232	0.0852	-11.8932
CNY	MCM	0.5627	0.4125	-13.3066
	KM	0.5964	0.3413	-13.3292
SGD	MCM	0.7456	0.3338	-13.4691
	KM	0.3685	1.6690	-13.3897

จากผลการทดลองพบว่าพารามิเตอร์ที่ได้จากวิธี KM ค่อนข้างแตกต่างจากค่าที่ได้ด้วยวิธี MCM นั่นเป็นเพราะผลตอบแทนของอัตราแลกเปลี่ยนเงินตราที่นำมาใช้ในการทดลองมีฟังก์ชันการแจกแจงที่ต่างไปจากการแจกแจงปกติค่อนข้างมากดังแสดงไว้ในรูปที่ 2 ในขณะที่การประมาณด้วยด้วยวิธี MCM อาศัยการใช้อนุภาคร่วมกับค่าถ่วงน้ำหนักในการประมาณการแจกแจงจริงของข้อมูลจึงทำให้ค่าพารามิเตอร์ที่ประมาณได้มีค่าความควรจะเป็นมากกว่าวิธี KM



รูปที่ 2 การแจกแจงของข้อมูล

7. สรุปผลการทดลอง

งานวิจัยฉบับนี้นำเสนอวิธีการประมาณหาค่าพารามิเตอร์ของแบบจำลอง SV ในรูปแบบเชิงเส้นโดยใช้วิธี MCM และวิธี KM ในขั้นตอนวิธี EM จากผลการทดลองพบว่าค่าพารามิเตอร์ที่ได้จากวิธีการทั้งสองไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ 0.05 เมื่อเทียบกับพารามิเตอร์จริง จากนั้นจึงนำทั้งวิธี MCM และวิธี KM มาประยุกต์ใช้ในการหาพารามิเตอร์ของแบบจำลอง SV โดยใช้อัตราแลกเปลี่ยนเงินตราต่างประเทศต่อสกุลเงินบาทไทยที่สำคัญ 5 อัตราเป็นกรณีศึกษา ผลการทดลองพบว่าพารามิเตอร์ที่ได้จากทั้ง 2 วิธีมีค่าเฉลี่ยต่างกันนั่นเป็นเพราะวิธี KM อาศัยการประมาณการแจกแจงของข้อมูลจริงด้วยการแจกแจงปกติ ในขณะที่วิธี MCM ใช้การประมาณการแจกแจงจริงของข้อมูลด้วยอนุภาคและค่าถ่วงน้ำหนัก

8. เอกสารอ้างอิง

[1] Thailand Futures Exchange. Retrieved from http://www.tfex.co.th/th/products/set50futures-spec.html.

[2] Merton, R. C. (1973). The theory of rational option pricing. *Bell Journal of Economics*, *4*, 141-183.

[3] Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economics*, *81*, 637–654.

[4] Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society*, *39*(1), 1-38.

[5] Anderson, B. D. O., & Moore, J. B. (2005). *Optimal Filtering*, New York, USA: Dover.

[6] Faragher, R. (2012). Understanding the basis of Kalman filter via a simple and intuitive derivation. *IEEE Signal Processing Magazine*, *5*(29), 128–132.

[7] Fridman, M., & Harris, L. (1998). A maximum likelihood approach for non-Gaussian stochastic volatility models. *Journal of Business and Economics Statistics*, *1*(3), 284-291.

[8] Breidt, F. J., & Carriquiry A. L., (1996). *Improvement quasi maximum likelihood estimation for stochastic volatility models*, New York, USA: Springer.

[9] Taylor, S. J. (1994). *Financial returns modeled by the product of two stochastic processes a study of daily sugar prices 1961–79. In Time Series Analysis: Theory and Practice* (vol. 1). New York, USA: Elsevier Science Publishing.

[10] Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of Econometrics*, *63*, 289-306.

[11] Chatterjee, S. (2005). Application of the Kalman filter for estimating continuous time term structure models - the case of UK and Germany. Retrieved from: http://www. gla.ac.uk/media/media_22193_en.pdf.

[12] Racicot, F. E., & Theoret, R. (2010). Forecasting stochastic volatility using the Kalman filter: An application to Canadian interest rates and price-earning ratio. *AESTI-MATIO, the IEB International Journal of Finance*, *1*, 28-47.

[13] Shephard, N., & Xiu D. (2012). Econometric analysis of multivariate realized QML: estimation of the covariation of equity prices under asynchronous trading. Retrieved from: http://faculty.chicagobooth.edu/dacheg.xiu/research/KFQMLE.pdf.

[14] Malakorn, T., & Iamtan, T. (2015). Parameter Estimation of Stochastic Volatility Models using Particle Method and EM Algorithm. *Proceeding of the 38th Electrical Engineering Conference.* Woraburi Hotel and Resort, Phra Nakhon Si Ayutthaya.

[15] Kim, J., & Stoffer, D. S. (2006). Fitting stochastic volatility models in the presence of irregular sampling via Particle Methods and the EM algorithm. *Journal of Time Series Analysis, 29*(5), 811-833.

[16] Nkemnole, E. B., & Abass, O. (2015). A t-distribution based particle filter for univariate and multivariate stochastic volatility models. *Journal of the Nigerian Mathematical Society, 34*(2), 227-242.

9. ประวัติผู้วิจัย

ธนิต มาลากร สำเร็จการศึกษาระดับปริญญาเอกในสาขาวิศวกรรมไฟฟ้าจาก Virginia Tech (VPI & SU) ประเทศสหรัฐอเมริกาในปี พ.ศ. 2546 ปัจจุบันดำรงตำแหน่งรองศาสตราจารย์ประจำสาขาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร งานวิจัยที่สนใจได้แก่ คณิตศาสตร์การเงิน (Mathematical Finance) ทฤษฎีหลายมาตราส่วน (Multiscale theory) ระบบเชิงเส้นหลายมิติ (Multidimensional linear system)

ธนภัทร เอี่ยมตาล สำเร็จการศึกษาระดับปริญญาตรี สาขาวิศวกรรมไฟฟ้าสื่อสารและระดับปริญญาโท สาขาวิศวกรรมการจัดการในปี พ.ศ. 2549 และ พ.ศ. 2552 ตามลำดับ ปัจจุบันกำลังศึกษาในระดับปริญญาเอกสาขาวิศวกรรมไฟฟ้า มหาวิทยาลัยนเรศวร งานวิจัยที่สนใจได้แก่ การค้นหาแบบศึกษาสำนึก (Heuristic search) ตัวกรองและกระบวนการเชิงเฟ้นสุ่ม (Filter and Stochastic process)

CT2

วันพฤหัสบดีที่ 3 พฤศจิกายน 2559 เวลา 13:00-14:40 น. ห้องชะอำ 1

Session Chair : ผศ. ดร.วันจักร์ เล่นวารี

CT006

ขั้นตอนวิธี EM ร่วมกับตัวกรองอนุภาคเกาส์เซียนสำหรับการประมาณพารามิเตอร์

ธนิต มาลากร และ ธนภัทร เอี่ยมดวล

งานวิจัยนี้นำเสนอการประยุกต์ใช้ขั้นตอนวิธี Expectation-Maximization (EM) ร่วมกับตัวกรองอนุภาคเกาส์เซียนในการประมาณพารามิเตอร์ของแบบจำลองไม่เชิงเส้นโดยใช้แบบจำลองความผันผวนเชิงเส้นสุ่มเป็นกรณีศึกษา ผลการทดลองพบว่าพารามิเตอร์ที่ได้จากการใช้ตัวกรองดังกล่าวเมื่อเทียบกับตัวกรองอนุภาคมีค่าไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ 0.05 แต่ตัวกรองอนุภาคเกาส์เซียนใช้เวลาในการประมวลผลน้อยกว่าตัวกรองอนุภาคอย่างมีนัยสำคัญ

CT007

การเรียนรู้เหตุและผลแบบเกรนเจอร์สำหรับข้อมูลอนุกรมเวลาด้วยวิธีทดสอบแบบวาลด์บนแบบจำลองที่มีเสถียรภาพ

นันทนัช รักษาศรี และ จิตโกมุท ส่งศิริ

บทความนี้จะอธิบายความสัมพันธ์ระหว่างตัวแปรของข้อมูลอนุกรมเวลาผ่านเงื่อนไขความสัมพันธ์เชิงเหตุและผลแบบเกรนเจอร์ (GC) บนแบบจำลองถดถอยตัวเอง และหารูปแบบความสัมพันธ์นั้นผ่านการทดสอบวาลด์ (Wald) บนพารามิเตอร์ของแบบจำลอง ผลลัพธ์ของบทความนี้คือ รูปแบบปัญหาการประมาณแบบจำลองด้วยวิธีกำลังสองต่ำสุดที่รวมเงื่อนไข GC และเสถียรภาพของแบบจำลอง โดยเงื่อนไขเพียงหอของเสถียรภาพอยู่ในรูปนอร์มแบบอนันต์ของเมทริกซ์พลวัตของแบบจำลองมีค่าน้อยกว่าหนึ่ง ปัญหาที่ได้จะอยู่ในรูปคอนเวกซ์แบบกำลังสองและผลเฉลยของปัญหาคือแบบจำลองที่การันตีว่ามีเสถียรภาพ อย่างไรก็ตามแบบจำลองจะมีค่าความคลาดเคลื่อนที่มากเกินความจำเป็นเนื่องจากเงื่อนไขเสถียรภาพเป็นแค่เงื่อนไขเพียงหอเท่านั้น

CT008

การประมาณค่าความเร็วรอบของมอเตอร์เหนี่ยวนำสามเฟสโดยใช้ตัวกรองคาลมาน

ศศิยา ยุคมสุข, กองพล อารีรักษ์, ธิดารัตน์ อารีรักษ์ และ กองทัน อารีรักษ์

บทความนี้นำเสนอวิธีตัวกรองคาลมาน เพื่อประมาณค่าความเร็วรอบของมอเตอร์เหนี่ยวนำสามเฟส โดยวิธีการดังกล่าวมีตัวแปรสถานะคือ กระแสสเตเตอร์ และฟลักซ์โรเตอร์ ส่วนความเร็วรอบของมอเตอร์ที่ต้องการประมาณค่าจะถูกพิจารณาเป็นตัวแปรสถานะเพิ่มเติมด้วยการคำนวณของวิธีตัวกรองคาลมานทำให้ได้โดยอาศัยปริมาณที่สามารถวัดได้คือ แรงดัน และกระแสสเตเตอร์ การตรวจสอบความถูกต้องของการประมาณค่าความเร็วรอบของมอเตอร์นี้ วิธีตัวกรองคาลมันจะทำการคำนวณเทียบกับการจำลองสถานการณ์ระบบขับเคลื่อนมอเตอร์เหนี่ยวนำสามเฟส ซึ่งผลจากการเปรียบเทียบแสดงให้เห็นว่าค่าความเร็วรอบที่ประมาณจากวิธีนำเสนอมีความใกล้เคียงและสู่ตามค่าความเร็วรอบจริงถึงแม้ว่าโหลดจะมีการเปลี่ยนแปลง

CT009

อัลกอริทึม RGBDSLAM แบบปรับปรุงสำหรับรองรับสิ่งแวดล้อมแบบพลวัต

คุภมร ศรีบุญแก้ว และ พจน์ ตั้งงามจิรศ์

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาการระบุตำแหน่งและสร้างแผนที่ 3 มิติด้วยอัลกอริทึม RGBDSLAM ที่รองรับกับสภาพแวดล้อมที่มีวัตถุเคลื่อนที่ โดยมีแนวคิดให้คัดแยกภาพที่มีการเคลื่อนที่ของวัตถุออกก่อนที่จะนำภาพเข้าสู่กระบวนการระบุตำแหน่งและสร้างแผนที่ โดยใช้วิธีการคำนวณการติดตามการเคลื่อนที่ของภาพของ Lucas-Kanade ในการคัดแยกลักษณะเวกเตอร์การเคลื่อนที่ของภาพที่ถูกนำมาระบุตำแหน่งและสร้างแผนที่แบ่งออกเป็น 2 ชนิด คือ ภาพที่มีการเคลื่อนที่ของวัตถุและภาพที่มีการเคลื่อนที่ของกล้องในวิธีการที่นำเสนอนี้ภาพที่มีการเคลื่อนที่ของวัตถุจะถูกกรองออกไปก่อนเข้าสู่กระบวนการ RGBDSLAM ทำให้ผลการทดลองที่ประเมินความถูกต้องและแม่นยำของการระบุตำแหน่งและสร้างแผนที่ 3 มิติ จากการเปรียบเทียบค่าความผิดพลาดเฉลี่ย (RMSE) ดีขึ้นเมื่อเทียบกับวิธีการเดิมโดยผลที่ได้จากวิธีการเดิมมีค่าความผิดพลาดเฉลี่ย 19.943% ส่วนวิธีที่ได้พัฒนาขึ้นมีค่าความผิดพลาดเฉลี่ย 4.416%

CT010

การประมาณขอบเขตของการมีเสถียรภาพแบบเชิงเส้นกำกับจากฟังก์ชันเลียปูนอฟที่คำนวณได้ด้วยวิธีการของทาคากิและซูจิโน

อภิชัย สุยะพันธ์, กองทัน อารีรักษ์ และ กองพล อารีรักษ์

บทความนี้นำเสนอการประมาณขอบเขตของการมีเสถียรภาพ แบบเชิงเส้นกำกับของระบบไฟฟ้ากำลังดีซีอย่างง่ายที่มีโหลดกำลังไฟฟ้า คงตัวแบบอุดมคติ โดยอาศัยการคำนวณหาฟังก์ชันเลียปูนอฟด้วยวิธีการที่นำเสนอโดยทาคากิและซูจิโน ผลการประมาณที่ได้จากวิธีการที่นำเสนอในบทความนี้จะนามาเปรียบเทียบกับผลการประมาณที่ได้จากฟังก์ชันเลียปูนอฟที่คำนวณได้ด้วยวิธีการที่นำเสนอโดยเบรย์ทันและมอเซอร์ รวมทั้งเปรียบเทียบกับขอบเขตของการมีเสถียรภาพแบบเชิงเส้นกำกับที่แท้จริง การตรวจสอบความถูกต้องของขอบเขตของการมีเสถียรภาพแบบเชิงเส้นกำกับที่ได้จากวิธีที่นำเสนอในบทความนี้ จะอาศัยการจำลองสถานการณ์บนคอมพิวเตอร์ด้วยโปรแกรม MATLAB ผลการตรวจสอบแสดงให้เห็นว่าขอบเขตของการมีเสถียรภาพแบบเชิงเส้นกำกับที่ประมาณได้นอกจากจะมีขนาดใกล้เคียงกับขอบเขตของการมีเสถียรภาพแบบเชิงเส้นกำกับที่แท้จริง มากกว่าขอบเขตที่ประมาณได้จากฟังก์ชันเลียปูนอฟที่คำนวณได้ด้วยวิธีการที่นำเสนอโดยเบรย์ทันและมอเซอร์แล้ว ยังสามารถรับประกันได้อีกว่า การเริ่มต้นการทำงานหรือการเปลี่ยนแปลงโหลดแบบทันทีทันใด ด้วยจุดใด ๆ ที่อยู่ภายในขอบเขต จะไม่ส่งผลต่อเสถียรภาพของระบบอย่างแน่นอน

ขั้นตอนวิธี EM ร่วมกับตัวกรองอนุภาคเกาส์เซียนสำหรับการประมาณพารามิเตอร์

The EM Algorithm Coupled with Gaussian Particle Filter for Parameter Estimation

ธนิต มาอาคร[1] และ ธนภัทร เอี่ยมตาล[2]

[1]ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร E-mail: tanitm@nu.ac.th

[2]นิสิตปริญญาเอก สาขาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

บทคัดย่อ

งานวิจัยนี้นำเสนอการประยุกต์ใช้ขั้นตอนวิธี Expectation-Maximization (EM) ร่วมกับตัวกรองอนุภาคเกาส์เซียนในการประมาณพารามิเตอร์ของแบบจำลองไม่เชิงเส้นโดยใช้แบบจำลองความผันผวนเชิงเฟ้นสุ่มเป็นกรณีศึกษา ผลการทดลองพบว่าพารามิเตอร์ที่ได้จากการใช้ตัวกรองดังกล่าวเมื่อเทียบกับตัวกรองอนุภาคมีค่าไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ 0.05 แต่ตัวกรองอนุภาคเกาส์เซียนใช้เวลาในการประมวลผลน้อยกว่าตัวกรองอนุภาคอย่างมีนัยสำคัญ

คำสำคัญ: การประมาณพารามิเตอร์ ขั้นตอนวิธี EM ตัวกรองอนุภาค ตัวกรองอนุภาคเกาส์เซียน

Abstract

This paper presents the application of the EM algorithm coupled with the Gaussian Particle Filters (GPF) in estimating the parameters of nonlinear models using the Stochastic Volatility (SV) model as a case study. The experimental results showed that the parameters estimated by such filters and the conventional Particle Filters (PF) are slightly different at the 0.05 significance level, but the time complexity of the GPF is reduced dramatically as compared with the PF.

Keywords: Parameter Estimation, EM Algorithm, Particle Filter, Gaussian Particle Filter

1. บทนำ

ในงานวิจัยด้านการวิเคราะห์อนุกรมเวลา (Time series analysis) การสร้างแบบจำลองที่เหมาะสมโดยอาศัยข้อมูลจากการเก็บตัวอย่างนับว่ามีบทบาทสำคัญอย่างมากต่อการพยากรณ์ Dempster et al. [1] ได้พัฒนาขั้นตอนวิธี EM ขึ้นในปี ค.ศ. 1977 เพื่อใช้ในการประมาณพารามิเตอร์ ทั้งนี้ขั้นตอนวิธีดังกล่าวเป็นวิธีทำซ้ำประกอบด้วย 2 ขั้นตอนย่อย นั่นคือ ขั้นตอน E (Expectation step) ซึ่งเป็นขั้นตอนในการหาค่าคาดหมายความควรจะเป็นแบบสมบูรณ์ (Complete likelihood) เมื่อเทียบกับการแจกแจงปรับเรียบ (Smooth distribution) และขั้นตอน M (Maximization step) ซึ่งเป็นขั้นตอนในการหาพารามิเตอร์เพื่อทำให้ฟังก์ชันควรจะเป็น (Likelihood function) มีค่าสูงสุด

เมื่อนำขั้นตอนวิธี EM มาประยุกต์ใช้ในการประมาณพารามิเตอร์ของแบบจำลองที่ถูกบรรยายในรูปแบบปริภูมิสถานะจำเป็นต้องสังเคราะห์ตัวแปรสถานะปรับเรียบ x_k^s ที่ได้จากการแจกแจงปรับเรียบเพื่อนำมาคำนวณหาฟังก์ชันควรจะเป็นแบบสมบูรณ์ วิธีการของคาลมาน (Kalman Method : KM) เป็นวิธีที่นิยมเลือกใช้อย่างแพร่หลายสำหรับแบบจำลองเชิงเส้นที่มีการแจกแจงเกาส์เซียนเนื่องจากวิธีการดังกล่าวมีความเรียบง่ายและมีผลเฉลยในรูปแบบปิด (Closed form) อีกทั้งยังจัดได้ว่าเป็นวิธีประมาณเหมาะที่สุด (Optimal estimation) [2,3] กล่าวคือความคลาดเคลื่อนกำลังสองเฉลี่ยระหว่างตัวแปรสถานะจริงและตัวแปรสถานะที่สังเคราะห์ขึ้นด้วยวิธี KM มีค่าต่ำสุดเมื่อเทียบกับวิธีการสังเคราะห์ตัวแปรสถานะด้วยวิธีอื่น ๆ

ในกรณีของแบบจำลองไม่เชิงเส้นหรือแบบจำลองที่มีการแจกแจงแบบอื่นที่ไม่ใช่การแจกแจงเกาส์เซียน การใช้วิธี KM ในขั้นตอนวิธี EM จะนำไปสู่ความควรจะเป็นเสมือนสูงสุด (Quasi maximum likelihood) [4,5] ด้วยเหตุนี้ วิธีมอนติคาร์โล [6-8] ซึ่งอาศัยวิธีการแก้ปัญหาแบบศึกษาลำนึก (Heuristic approach) จึงเข้ามามีบทบาทในการสังเคราะห์ x_k^s แทนวิธี KM โดยใช้หลักการชักตัวอย่างจากฟังก์ชันการแจกแจง หาค่าถ่วงน้ำหนักพร้อมทำการปรับตัวอย่างใหม่ ทั้งนี้ตัวอย่างที่ได้จากการชักตัวอย่างนิยมเรียกว่าอนุภาค (Particle) ด้วยเหตุนี้ตัวกรองและตัวปรับเรียบที่ถูกออกแบบด้วยวิธีมอนติคาร์โล จึงเรียกว่าตัวกรองอนุภาค (Particle Filter : PF) และตัวปรับเรียบอนุภาค (Particle Smoother : PS) ตามลำดับ

จุดด้อยของ PF อยู่ที่ขั้นตอนการกรองซึ่งใช้วิธีการปรับตัวอย่างใหม่ วิธีการดังกล่าวใช้เวลาในการประมวลผลค่อนข้างนานโดยเฉพาะอย่างยิ่งเมื่อมีอนุภาคเป็นจำนวนมากเนื่องจากเป็นวิธีที่ใช้การทำงานแบบเรียงลำดับ ในปี ค.ศ. 2003 Kotecha และ Djurić [9] จึงได้นำเสนอตัวกรองอนุภาคเกาส์เซียน (Gaussian Particle Filter : GPF) เพื่อลดเวลาในการประมวลผลโดยใช้วิธีการชักตัวอย่างจากการแจกแจงเกาส์เซียนแทนวิธีการปรับตัวอย่างใหม่ อย่างไรก็ตาม ในงานวิจัยดังกล่าวมุ่งพัฒนาตัวกรองเพียงอย่างเดียวและเท่าที่ผู้วิจัยทราบไม่ปรากฏว่ามีการนำ GPF มาใช้ในงานด้านการประมาณเหาพารามิเตอร์มาก่อนหน้านี้

ดังนั้นงานวิจัยชิ้นนี้จึงศึกษาการทำงานของขั้นตอนวิธี EM ร่วมกับ GPF ในการประมาณเหาพารามิเตอร์ของแบบจำลองไม่เชิงเส้นซึ่งในที่นี้เลือกใช้แบบจำลองความผันผวนเชิงเฟ้นสุ่มที่นำเสนอโดย Taylor [10] มาเป็นกรณีศึกษาเพื่อให้มีความต่อเนื่องจากงานใน [11,12]

2. ตัวกรองและตัวปรับเรียบ

2.1. ตัวกรองอนุภาค (PF) [6]

การทำงานของ PF มี 2 ขั้นตอนดังนี้

1. ขั้นตอนเริ่มต้น: เมื่อ $k=0$ ทำการชักตัวอย่างอนุภาคเริ่มต้นจากการแจกแจงก่อน (Prior distribution)

$$x_0^{f(i)} \sim \mathcal{N}(\mu_0^f, P_0^f) \text{ เมื่อ } i=1,\ldots,N_f \quad (1)$$

และให้ค่าถ่วงน้ำหนักเริ่มต้นของแต่ละอนุภาคเป็น

$$\omega_0^{f(i)} = 1/N_f \quad \text{เมื่อ } i=1,\ldots,N_f \quad (2)$$

โดยที่ N_f คือจำนวนของอนุภาค

2. ขั้นตอนการทำซ้ำ: เมื่อ $k=1,\ldots,N$ เป็นการทำงานแบบวนซ้ำระหว่างขั้นตอนการทำนาย ขั้นตอนการหาค่าถ่วงน้ำหนักและขั้นตอนการกรองดังนี้

<u>ขั้นตอนการทำนาย</u>เป็นการหาอนุภาคการทำนายที่เวลาถัดไปโดยอาศัยสมการสถานะ นั่นคือ

$$x_k^{p(i)} \sim p(x_k, x_{k-1}^{f(i)}) \text{ เมื่อ } i=1,\ldots,N_f \quad (3)$$

<u>ขั้นตอนการหาค่าถ่วงน้ำหนัก</u>เป็นการหาการแจกแจงแบบมีเงื่อนไขของสัญญาณออก นั่นคือ

$$\tilde{\omega}_k^{f(i)} \propto p(y_k \mid x_k^{p(i)}) \text{ เมื่อ } i=1,\ldots,N_f \quad (4)$$

จากนั้นจึงนำมาทำให้เป็นบรรทัดฐานเพื่อให้ผลรวมค่าถ่วงน้ำหนักมีค่าเป็น 1 ดังนี้

$$\omega_k^{f(i)} = \frac{\tilde{\omega}_k^{f(i)}}{\sum_{i=1}^{N_f} \tilde{\omega}_k^{f(i)}} \quad (5)$$

<u>ขั้นตอนการกรอง</u>เป็นการหาอนุภาคการกรอง $x_k^{f(i)}$ โดยอาศัยวิธีการปรับอนุภาคใหม่ (Resampling) เพื่อกำจัดอนุภาค $x_k^{p(i)}$ ที่มีค่าถ่วงน้ำหนักน้อยมากออกไปแล้วนำอนุภาค $x_k^{p(j)}$ ที่มีค่าถ่วงน้ำหนักมากมาใส่แทนที่เพื่อเป็นวิธีลดปัญหาการลดลง (Degeneracy problem)

ทั้งนี้ตัวแปรสถานะการกรองคำนวณได้จากผลรวมเชิงเส้นของอนุภาคการกรอง นั่นคือ

$$x_k^f = \sum_{i=1}^{N_f} \omega_k^{f(i)} x_k^{f(i)} \quad (6)$$

2.2. ตัวกรองอนุภาคเกาส์เซียน (GPF) [7]

การทำงานของ GPF มี 2 ขั้นตอนเช่นเดียวกับ PF แต่ในการหาอนุภาคการกรองนั้น ให้เปลี่ยนจากวิธีการปรับอนุภาคใหม่เป็นวิธีการชักตัวอย่างจากการแจกแจงเกาส์เซียน

$$x_k^{f(i)} \sim \mathcal{N}(\mu_k^f, P_k^f) \text{ เมื่อ } i=1,\ldots,N_f \quad (7)$$

โดยที่

$$\mu_k^f = \sum_{i=1}^{N_f} \omega_k^{f(i)} x_k^{p(i)} \quad (8)$$

$$P_k^f = \sum_{i=1}^{N_f} \omega_k^{f(i)} (x_k^{p(i)} - \mu_k^f)^2 \quad (9)$$

และตัวแปรสถานะการกรองคำนวณจากผลรวมเชิงเส้นของอนุภาคการกรอง

$$x_k^f = \sum_{i=1}^{N_f} \omega_k^{f(i)} x_k^{f(i)} \quad (10)$$

2.3. ตัวปรับเรียบอนุภาค (PS)

PS เป็นกระบวนการเรียกซ้ำแบบย้อนกลับโดยอาศัยค่าที่คำนวณได้จาก PF หรือ GPF การทำงานของ PS ประกอบด้วย 2 ขั้นตอนดังนี้

1. ขั้นตอนสุดท้าย: เมื่อ $k=N$ ให้พิจารณาค่าถ่วงน้ำหนัก $\omega_N^{f(i)}$ เป็นความน่าจะเป็นในการเลือกอนุภาคการกรองที่ตำแหน่ง i มาเป็นอนุภาคการปรับเรียบ นั่นคือ

$$x_N^s = x_N^{f(i)} \text{ ด้วยความน่าจะเป็น } \omega_N^{f(i)} \quad (11)$$

2. ขั้นตอนการทำซ้ำ: เมื่อ $k=N-1,\ldots,0$ เป็นการคำนวณค่าถ่วงน้ำหนักใหม่ $\omega_k^{s(i)}$ จาก

$$\omega_k^{s(i)} \propto \omega_k^{f(i)} p(x_{k+1}^s \mid x_k^{f(i)}) \quad (12)$$

จากนั้นจึงพิจารณาค่าถ่วงน้ำหนัก $\omega_k^{s(i)}$ ที่ได้เป็นความน่าจะเป็นในการเลือกอนุภาคการกรองที่ตำแหน่ง i ให้มาเป็นอนุภาคการปรับเรียบ

$$x_k^s = x_k^{f(i)} \text{ ด้วยความน่าจะเป็น } \omega_k^{s(i)} \quad (13)$$

สำหรับทุกค่า $i=1,\ldots,N_f$

ให้ N_s แทนจำนวนแนววิถีของอนุภาคการปรับเรียบที่ต้องการ ให้ $x_k^{s(j)}$ และ $\omega_{j,k}^{s(i)}$ แทนแนววิถีที่ j ของอนุภาคการปรับเรียบและของค่าถ่วงน้ำหนักตามลำดับ แล้วตัวแปรสถานะการปรับเรียบคำนวณจากผลรวมเชิงเส้นของอนุภาคการปรับเรียบ นั่นคือ

$$x_k^s = \sum_{j=1}^{N_s} \omega_{j,k}^{s(i)} x_k^{s(j)} \quad (14)$$

ตัวแปรสถานะการปรับเรียบจะถูกนำไปใช้ในขั้นตอนวิธี EM ซึ่งจะอธิบายในหัวข้อถัดไป

3. ขั้นตอนวิธี EM

ในงานวิจัยนี้เลือกใช้ขั้นตอนวิธี EM ในการประมาณหาพารามิเตอร์ของกระบวนการเชิงเส้นสุ่มโดยขั้นตอนวิธีดังกล่าวเป็นวิธีทำซ้ำเพื่อทำให้ฟังก์ชันควรจะเป็นมีค่าสูงสุดโดยมีรายละเอียดดังนี้

ให้ θ แทนพารามิเตอร์ของระบบและให้ฟังก์ชันควรจะเป็นของพารามิเตอร์แบบลอการิทึมคือ

$$\log p(y_{1:N} \mid \theta) := \mathcal{L}(\theta) \quad (15)$$

แล้ว $\hat{\theta}$ ที่ทำให้ $\mathcal{L}(\hat{\theta})$ ในสมการที่ (15) มีค่าสูงสุดคือ

$$\hat{\theta} = \arg\max_{\theta} \log p(y_{1:N} \mid \theta) \quad (16)$$

โดยทั่วไปการคำนวณหาค่าเหมาะที่สุดข้างต้นต้องอาศัยวิธีเชิงตัวเลข สำหรับขั้นตอนวิธี EM ไม่ได้คำนวณหาค่าเหมาะที่สุดโดยตรง หากแต่อาศัยการทำให้ขอบเขตล่างของ $\mathcal{L}(\theta)$ มีค่ามากขึ้นในแต่ละรอบของการทำซ้ำจึงส่งผลทำให้ $\mathcal{L}(\theta)$ มีค่ามากสูงขึ้นตามไปนั่นเอง

ดังนั้นในขั้นตอนวิธี EM สมการที่ (16) จึงเปลี่ยนเป็น

$$\hat{\theta}^{(k+1)} = \arg\max_{\theta} \mathrm{E}[\, \mathcal{L}(\theta, \hat{\theta}^{(k)})] \qquad (17)$$

เมื่อ

$$\mathrm{E}[\, \mathcal{L}(\theta, \hat{\theta}^{(k)})] = \mathrm{E}[\log p(x_{0:N}^s, y_{1:N} \mid \theta)] \qquad (18)$$

จากคุณสมบัติของมาร์คอฟจะได้ว่า

$$\mathcal{L}(\theta, \hat{\theta}^{(k)}) = \log p(x_0^s \mid \theta) + \sum_{k=1}^{N} \log p(x_k^s \mid x_{k-1}^s, \theta) + \sum_{k=1}^{N} \log p(y_k \mid x_k^s, \theta) \qquad (19)$$

สำหรับขั้นตอนวิธี EM มีดังนี้

1. ขั้นตอนเริ่มต้น: กำหนดค่าเริ่มต้น $\hat{\theta}^{(0)}$

2. ขั้นตอนการทำซ้ำ: เมื่อ $k = 1, \ldots, N$

ขั้นตอน E คำนวณหา $(\theta, \hat{\theta}^{(k)})$ ตามสมการที่ (18) และ (19)

ขั้นตอน M คำนวณหา $\hat{\theta}^{(k+1)}$ ตามสมการที่ (17)

4. ผลการทดลอง

งานวิจัยชิ้นนี้เลือกใช้แบบจำลอง SV มาเป็นตัวอย่างซึ่งมีรูปแบบดังนี้

$$\Sigma := \begin{cases} x_k = \phi x_{k-1} + w_k \\ r_k = \beta \exp\left(\frac{x_k}{2}\right)\varepsilon_k \end{cases} \quad \text{เมื่อ } k = 0, \ldots, N \qquad (20)$$

โดยที่ r_k คือผลตอบแทนแบบลอการิทึมของสินทรัพย์ที่มีความผันผวนเป็น σ_k และ $x_k = \log \sigma_k^2$ ให้ $w_k \sim \mathcal{N}(0, Q)$, $\varepsilon_k \sim \mathcal{N}(0,1)$ โดยสมมติให้ทั้ง w_k และ ε_k มีความอิสระเชิงสถิติต่อกัน

ให้ $\theta = (\phi, Q, \beta)$ เป็นพารามิเตอร์ของแบบจำลองใน (20) ที่ต้องการทราบโดยสมมติว่าทราบข้อมูล r_k จากนั้นจึงประยุกต์ใช้ขั้นตอนวิธี EM โดยเลือกใช้ทั้ง PF และ GPF มาเป็นตัวกรอง ทั้งนี้ฟังก์ชันควรจะเป็นของตัวอย่างนี้คือ

$$\mathcal{L}(\theta, \hat{\theta}^{(k)}) = \frac{1}{N_s}\sum_{j=1}^{N_s}\left\{-\frac{1}{2}\log P_0 - \frac{1}{2}\frac{(x_0^s - \mu_0)^2}{P_0} - \frac{N}{2}\log Q - \frac{1}{2}\sum_{k=1}^{N}\frac{(x_k^s - \phi x_{k-1}^s)^2}{Q} - N\log\beta - \frac{1}{2}\sum_{k=1}^{N}\left[x_k^s + \frac{r_k^2}{\beta^2}\exp(-x_k^s)\right] - C\right\} \qquad (21)$$

โดยที่ C เป็นผลรวมของพจน์ของค่าคงที่ซึ่งไม่มีผลต่อการหาค่าสูงสุด

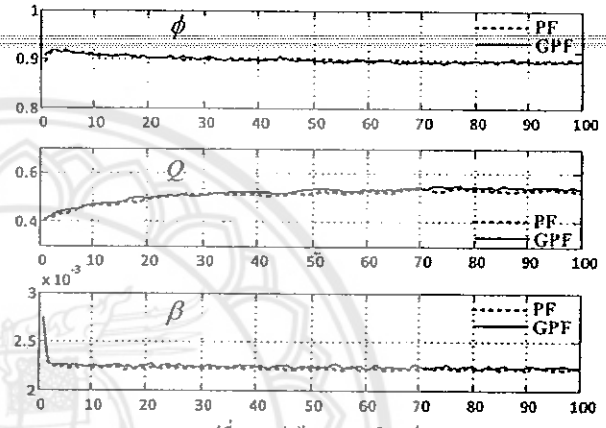
ลำดับต่อมาคือการคำนวณหา $\hat{\theta}$ ตามสมการที่ (17) โดยการหาอนุพันธ์ย่อยเทียบกับพารามิเตอร์แต่ละตัวที่ต้องการทราบจึงได้ว่า

$$\hat{\phi} = \frac{\sum_{k=1}^{N}\mathrm{E}\left[x_k^s x_{k-1}^s \mid r_{1:N}\right]}{\sum_{k=1}^{N}\mathrm{E}\left[(x_{k-1}^s)^2 \mid r_{1:N}\right]} \qquad (22)$$

$$\hat{Q} = \frac{\sum_{k=1}^{N}\mathrm{E}\left[(x_{k+1}^s - \hat{\phi}x_k^s)^2 \mid r_{1:N}\right]}{N} \qquad (23)$$

$$\hat{\beta} = \sqrt{\frac{\sum_{k=1}^{N}\mathrm{E}\left[r_k^2 \exp(-x_k^s) \mid r_{1:N}\right]}{N}} \qquad (24)$$

ในการทดลอง สมมติให้ $\theta = (\phi, Q, \beta) = (0.9, 0.5, 0.0022)$ จากนั้นจึงใช้แบบจำลอง Σ มาสร้างข้อมูล r_k โดยการจำลองด้วยโปรแกรม MATLAB® นำข้อมูลที่ได้มาเข้าสู่ขั้นตอนวิธี EM โดยทำการทดลองซ้ำทั้งหมด 30 ครั้งเพื่อใช้วิเคราะห์หาค่าทางสถิติของพารามิเตอร์ โดยสมมติให้พารามิเตอร์เริ่มต้นคือ $\theta_0 = (0.85, 0.4, 0.0127)$ และให้ $N = 500, N_f = 400, N_s = 100$ ค่าพารามิเตอร์ที่ประมาณได้โดยใช้ PF และ GPF แสดงไว้ในรูปที่ 1



รูปที่ 1 การลู่เข้าของพารามิเตอร์

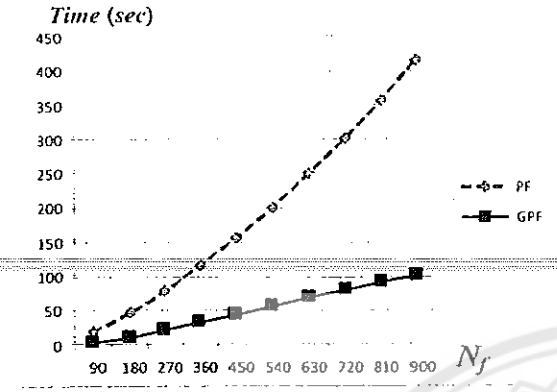
สังเกตว่าพารามิเตอร์ที่ได้จากตัวกรองทั้งคู่ลู่เข้าสู่พารามิเตอร์จริงที่กำหนดไว้ เมื่อทำการทดสอบทางสถิติด้วยวิธีการทดสอบแบบที (t-test) พบว่าพารามิเตอร์ที่ได้จากตัวกรองทั้งสองชนิดให้ค่าที่ไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ 0.05 เมื่อเทียบกับพารามิเตอร์จริงซึ่งผลทดลองที่ได้แสดงไว้ในตารางที่ 1

ตารางที่ 1 พารามิเตอร์ที่ประมาณโดยใช้ PF และ GPF

ตัวกรอง	ϕ	Q	β
PF	0.8964	0.4839	0.0021
(S.D.)	(0.0220)	(0.1038)	(0.000371)
P-Value	0.3784	0.4041	0.1087
Log Likelihood	-1230.99		
GPF	0.8947	0.4955	0.0021
(S.D.)	(0.0221)	(0.1040)	(0.000373)
P-Value	0.2025	0.8140	0.2013
Log Likelihood	-1238.50		

จากการวิเคราะห์ข้างต้นพบว่า พารามิเตอร์ที่ประมาณได้จากการเลือกใช้ PF หรือ GPF ร่วมกับขั้นตอนวิธี EM มีค่าที่ไม่แตกต่างกันอย่างมีนัยสำคัญ อย่างไรก็ตาม หากพิจารณาในเรื่องของความซับซ้อนในการ

คำนวณ (Computational complexity) พบว่า PF ใช้เวลาในการประมวลผลมากกว่า GPF อย่างชัดเจนโดยเฉพาะอย่างยิ่งเมื่อมีจำนวนอนุภาคเพิ่มมากขึ้นดังแสดงไว้ในรูปที่ 2



รูปที่ 2 เวลาในการประมวลผลระหว่าง PF และ GPF

5. สรุปผลการทดลอง

งานวิจัยฉบับนี้นำเสนอวิธีการประมาณหาค่าพารามิเตอร์ของแบบจำลองไม่เชิงเส้นด้วยขั้นตอนวิธี EM ร่วมกับ GPF และ PF โดยใช้แบบจำลอง SV เป็นกรณีศึกษา จากการทดลองพบว่าค่าพารามิเตอร์ที่ได้จากวิธีการทั้งสองไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ 0.05 เมื่อเทียบกับพารามิเตอร์จริงแต่การประมาณโดยใช้ GPF ใช้เวลาในการประมวลผลน้อยกว่า PF อย่างมีนัยสำคัญ

เอกสารอ้างอิง

[1] A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximum likelihood from incomplete data via EM algorithm," *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. 1–38, 1977.

[2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, New York: Dover, 2005.

[3] R. Faragher, "Understanding the basis of Kalman filter via a simple and intuitive derivation," *IEEE Signal Processing Magazine*, pp. 128–132, 2012.

[4] M. Fridman and L. Harris, "A maximum likelihood approach for non-Gaussian stochastic volatility models," *Journal of Business and Economics Statistics*, vol. 1, no. 3, pp. 284–291, 1998.

[5] F. J. Breidt, and A. L. Carriquiry, *Improvement quasi maximum likelihood estimation for stochastic volatility models*, New York: Springer, 1996, pp. 228–247.

[6] S. Särkkä. *Bayesian Filtering and Smoothing*. Cambridge: CUP, 2013.

[7] J. Kim and D. S Stoffer, "Fitting stochastic volatility models in the presence of irregular sampling via Particle Methods and the EM algorithm," *Journal of Time Series Analysis*, vol. 29, no. 5, pp. 811-833, 2006.

[8] E.B. Nkemnole and O. Abass, "A t-distribution based particle filter for univariate and multivariate stochastic volatility models," *Journal of the Nigerian Mathematical Society*, vol. 34, no. 2, pp. 227-242, 2015.

[9] J. H. Kotecha and P. M. Djurić, "Gaussian Particle Filtering," *IEEE Trans. Signal Processing*, vol. 51, pp. 2592–2601, Oct. 2003.

[10] S. J. Taylor, *Financial returns modeled by the product of two stochastic processes–a study of daily sugar prices 1961–79 In Time Series Analysis. Theory and Practice*, vol. 1. New York: Elsevier Science Publishing, pp. 203–226, 1982.

[11] T. Malakorn and T. Iamtan, "Parameter Estimation of Stochastic Volatility Models using Particle Method and EM Algorithm," in *Proc. of the 38th Electrical Engineering Conf.*, Nov., pp. 233-236, 2015.

[12] ธนภัทร เอี่ยมคาล และ ธนิต มาลากร, "การประมาณพารามิเตอร์สำหรับแบบจำลองความผันผวนเชิงเฟ้นสุ่มด้วยขั้นตอนวิธี EM," *วิศวกรรมสาร มหาวิทยาลัยนเรศวร*, ปีที่ 11, ฉบับที่ 2, กรกฎาคม – ธันวาคม 2559.

ธนิต มาลากร สำเร็จการศึกษาระดับปริญญาเอกในสาขาวิศวกรรมไฟฟ้าจาก Virginia Tech (VPI & SU) ประเทศสหรัฐอเมริกาในปี พ.ศ. 2546 ปัจจุบันดำรงตำแหน่งรองศาสตราจารย์ประจำสาขาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร งานวิจัยที่สนใจได้แก่ คณิตศาสตร์การเงิน (Mathematical Finance) ทฤษฎีหลายมาตราส่วน (Multiscale theory) ระบบเชิงเส้นหลายมิติ (Multidimensional linear system)

ธนภัทร เอี่ยมคาล สำเร็จการศึกษาระดับปริญญาตรีสาขาวิศวกรรมไฟฟ้าสื่อสารและระดับปริญญาโทสาขาวิศวกรรมการจัดการในปี พ.ศ. 2549 และ พ.ศ. 2552 ตามลำดับ ปัจจุบันกำลังศึกษาในระดับปริญญาเอก สาขาวิศวกรรมไฟฟ้า มหาวิทยาลัยนเรศวร งานวิจัยที่สนใจได้แก่ การจัดเรียงเครื่องจักร (Machine layout) การค้นหาแบบศึกษาสำนึก (Heuristic serach) ตัวกรองและกระบวนการเชิงเฟ้นสุ่ม (Filter and Stochastic process)

Parameter Estimation of Stochastic Volatility Models using Particle Method and EM Algorithm

Tanit Malakorn and Thanapat Iamtan

Department of Electrical and Computer Engineering, Faculty of Engineering,

Naresuan University, THAILAND, tanitm@nu.ac.th

GN05

Abstract

Stochastic volatility models can be regarded as a particular type of linear state-space models corrupted by log chi-square noise processes. For estimation purposes, the Kalman filter is no longer optimal in general when the assumptions of linearity and Gaussianity are dropped; particle filters, on the other hand, can be applied to any state-space model. This paper presents one application of the EM algorithm coupled with the particle method for parameter estimation of SV models using simulated data and real data from the Foreign Exchange Market. The numerical results illustrate the effectiveness of the proposed method.

Keywords: Stochastic Volatility, Parameter Estimation, Particle Filter, EM Algorithm, Currency Exchange Rate

1 Introduction

In the last two decades, demand for electrical energy has gradually increased every year along with economic growth [1]. Approximately 70 percent of the electricity produced in Thailand is generated using natural gas which is nonrenewable. As reserves decline, domestic production of natural gas will continually decrease and import will steadily increase. Therefore, the currency exchange rate will play a central role in the calculation of the value of the trade and therefore the import costs.

Depending on the demand & supply of currency situation, exchange rate fluctuates continuously and randomly. In the time-series literature, exchange rates can be modelled as geometric Brownian motions (GBM) assuming constant volatility σ and drift μ (sometimes called mean); i.e., it follows a stochastic differential equation (SDE) of the form :

$$dS_t = \mu S_t\, dt + \sigma S_t\, dW_t, \qquad (1)$$

where S_t and W_t denote the exchange rate and a random walk, respectively (see, e.g., [2] –[7] and references therein).

Black and Scholes [6] derived European-style option pricing based on the assumption that asset prices follow SDE as in (1), and published their groundbreaking paper in 1973. However, empirical studies strongly indicate that the underlying volatility *is not* constant.

Stochastic volatility (SV) models are one approach to resolving the drawback of the Black-Scholes model. By assuming that the volatility is a stochastic process rather than a constant, it becomes possible to model derivatives more accurately. Once the model is chosen, it is subsequently fitted to the empirical data using the parameter estimation method. There are several estimation methods proposed in the literature; one of which is the Expectation-Maximization (EM) algorithm.

The EM algorithm, proposed by Dempster *et al.* [8] in 1977, is an iterative method used to compute parameters for Maximum Likelihood (ML) estimation. Basically, the EM algorithm consists of two alternating steps: the Expectation step (E-step) and the Maximization step (M-step). The E-step computes the best estimate of the likelihood function using the current knowledge of the ML estimates and the observed data. This step may require state filtering and smoothing with current parameter estimates to get state estimates. Then the M-step uses the filtered or smoothed state estimates to re-estimate the parameters so that the new likelihood function is maximized.

For the linear model with additive Gaussian noise, the classical *Kalman Filter (KF)* is an optimal state estimator in the sense that it minimizes the mean square error of the state estimates. When the assumptions of linearity and Gaussianity are dropped, the optimality does not hold in general. Due to the simplicity of the KF algorithm, however, it is still being used in reseach; see, e.g., [9]–[12]. Apart from the KF particle filters (PF) are sequential Monte Carlo methods based on point mass representations of probability densities, which can apply to any state-space model, and which generalize the classical KF methods. In this paper, we utilize the EM algorithm with particle methods to estimate SV parameters using the simulated data and the USD/THB daily exchange rates from the market. It is worth noting that the PF can apply in many different directions; interested readers may consult, e.g., [13]–[16] and references therein.

The remainder of this paper is organized as follows: Section 2 presents the SV modelling. The EM algorithm and the particle method are provided in Section 3 and Section 4. Section 5 demonstrates the empirical examples using two data sets: (1) the simulated data, and (2) the USD/THB daily exchange rates. We then draw some conclusions in Section 6.

2 Stochastic Volatility Modelling

Let E_k denote the exchange rate at time k. The log return[1] of E_k is then defined by $r_k := \log\left(\frac{E_k}{E_{k-1}}\right)$, which admits the univariate SV model introduced by Taylor [17].

$$r_k = \sigma_k \epsilon_k, \quad \epsilon_k \sim \text{i.i.d. } \mathcal{N}(0,1) \tag{2}$$

where σ_k is the volatility of r_k. According to the empirical results, σ_k^2 is the log-normally distribution; this means that there must exist a normal random variable x_k so that $x_k = \log\sigma_k^2$, and hence (2) becomes

$$r_k = \exp(x_k/2)\epsilon_k. \tag{3}$$

Traditionally, the log volatility x_k is assumed to follow the stationary AR(1) process with Gaussian innovation noise; i.e.,

$$x_k = \phi x_{k-1} + c + w_k, \quad w_k \sim \text{i.i.d. } \mathcal{N}(0,Q) \tag{4}$$

where $|\phi| < 1$.

Rather than (3) and (4), one may introduce a scaling factor β in (3) to remove the constant term c in (4). Hence the SV model for the log return is given by

$$\Sigma_{NL} := \begin{cases} x_k &= \phi x_{k-1} + w_k, \\ r_k &= \beta \exp(x_k/2)\epsilon_k. \end{cases} \tag{5}$$

Obviously this is a nonlinear state-space model. By taking logarithms of the squares of r_k, the nonlinear SV model in (5) can be transformed into a linear model as

$$\Sigma_L := \begin{cases} x_k &= \phi x_{k-1} + w_k, \\ y_k &= \alpha + x_k + v_k, \end{cases} \tag{6}$$

where $y_k = \log r_k^2$, $\alpha = \log\beta^2 + \mathbb{E}[\log\epsilon_k^2]$, and $v_k = \log\epsilon_k^2 - \mathbb{E}[\log\epsilon_k^2]$. We shall use the linear state-space model of the form (6) in this research. Note that v_k follows the $\log\chi_1^2$ distribution whose mean and variance are 0 and $\pi^2/2$, respectively.

3 EM Algorithm

Given the observed data $Y_N = \{y_0, \ldots, y_N\}$, the likelihood function is a measure of the plausibility of the data under parameter θ; i.e.,

$$L(\theta) = p(Y_N|\theta), \tag{7}$$

where $\theta \in \Theta$, the parameter space. The main goal of the ML estimation is to determine $\hat{\theta}^{ML}$ which makes $L(\hat{\theta}^{ML})$ achieve maximum:

$$\hat{\theta}^{ML} = \arg\max_{\theta\in\Theta} L(\theta) = \arg\max_{\theta\in\Theta} p(Y_N|\theta) \tag{8}$$

Due to the fact that

$$\arg\max_{\theta\in\Theta} L(\theta) = \arg\max_{\theta\in\Theta} \log L(\theta), \tag{9}$$

it is more convenient to maximize $\log L(\theta)$ rather than maximizing $L(\theta)$ itself.

Consider a nonlinear state-space model with additive noise processes of the form:

$$x_k = f(x_{k-1};\theta) + w_k. \tag{10}$$

$$y_k = h(x_k) + v_k. \tag{11}$$

Let θ_0 be an initial guess for estimating parameters. The EM algorithm iteratively generates a sequence of estimates θ_k for $k = 1, \ldots, N$ as follows:

1) E-Step : Evaluate auxiliary (complete) likelihood

$$Q(\theta_k,\theta) = \mathbb{E}\{\log\, p(X_N, Y_N;\theta|Y_N,\theta_{k-1})\}, \tag{12}$$

where $X_N = \{x_1, \ldots, x_N\}$.

2) M-Step : Maximize auxiliary (complete) likelihood; i.e, to compute

$$\theta_{k+1} = \arg\max_{\theta\in\Theta} Q(\theta_k,\theta) \tag{13}$$

The main advantage of the EM algorithm is that the monotonic increasing property holds; i.e., in each consecutive iteration, $L(\theta_{k+1}) \geq L(\theta_k)$.

In the case of the SV model in (6) where $\theta = (\phi, Q, \alpha)$, the (complete) likelihood function is then given by

$$\begin{aligned} L(\theta) &= p(x_{s,0})\prod_{k=1}^{N} p(x_{s,k}|x_{s,k-1})\prod_{k=0}^{N} p(y_k|x_{s,k}) \\ &= \frac{1}{\sqrt{2\pi\sigma_0^2}}\exp\left(-\frac{(x_{s,0}-\mu_0)^2}{2\sigma_0^2}\right)\times \\ &\quad \prod_{k=1}^{N}\frac{1}{\sqrt{2\pi Q}}\exp\left(-\frac{(x_{s,k}-\phi x_{s,k-1})^2}{2Q}\right)\times \\ &\quad \prod_{k=0}^{N}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{\gamma_k}{2}\right)\exp\left(-\frac{1}{2}\exp(\gamma_k)\right), \end{aligned} \tag{14}$$

where $\gamma_k := y_k - \alpha - x_{s,k} + \mathbb{E}[\log\epsilon_k^2]$, and $x_{s,k}$ is the particle smoother from the Particle Smoothing process (see in the next section). By taking the partial derivative of $\mathbb{E}[\log\, L(\theta|Y_N)]$ with respect to each parameter and setting them all to zero, the estimates $\hat{\theta}$ are then calcuated as follows:

[1] Financial time-series data, such as asset prices, oil prices, gold prices, GDP, inflations, including currency exchange rates are in general non-stationary. It is customary to transform non-stationary data to stationary ones using the logarithm of the ratio of consecutive data points, called *log return* : r_k

$$\hat{\phi} = \frac{\sum_{k=1}^{N} \left[\hat{x}_k \hat{x}_{k-1} + P_{k,k-1}\right]}{\sum_{k=1}^{N} \left[\hat{x}_{k-1}^2 + P_{k-1}\right]} \tag{15}$$

$$\hat{Q} = \frac{\sum_{k=1}^{N} \left[\hat{x}_k^2 + P_k\right]}{N} - \frac{1}{N} \frac{\left(\sum_{k=1}^{N} \left[\hat{x}_k \hat{x}_{k-1} + P_{k,k-1}\right]\right)^2}{\sum_{k=1}^{N} \left[\hat{x}_{k-1}^2 + P_{k-1}\right]} \tag{16}$$

$$\hat{\alpha} = \log\left[\frac{\sum_{k=0}^{N} \mathbb{E}\left[\exp\left(y_k - \hat{x}_k + \mathbb{E}[\log \epsilon_k^2]\right) | Y_N\right]}{N+1}\right], \tag{17}$$

where $\hat{x}_k = \mathbb{E}[x_{s,k}|Y_N]$, $P_k = \mathbb{E}[(x_{s,k} - \hat{x}_k)^2|Y_N]$ and $P_{k,l} = \mathbb{E}[(x_{s,k} - \hat{x}_k)(x_{s,l} - \hat{x}_l)|Y_N]$

4 Particle Method

Particle filters (PF) are sequential Monte Carlo methods based on a set of random particles with associated weights to approximate the conditional densities, $f(x_k|Y_t = \{y_0, \ldots, y_t\})$ rather than a single estimate as in the KF methods.

4.1 Particle Filtering

Let M and N be the number of particles and the length of the data, respectively.

1) Initialization : Generate random particles $x_{f,0}^{(j)} \sim \mathcal{N}(\mu_0, \sigma_0^2)$ for $j = 1, \ldots, M$

2) Recursive Step : For $k = 1, \ldots, N$

1. Generate noises $w_k^{(j)} \sim \mathcal{N}(0, Q)$ for $j = 1, \ldots, M$
2. Update estimated states $\hat{x}_{f,k}^{(j)} = \Phi x_{f,k-1}^{(j)} + w_k^{(j)}$ for $j = 1, \ldots, M$
3. Compute weights from the conditional density of y_k given $\hat{x}_{f,k}^{(j)}$; i.e,

$$w_k^{(j)} = p(y_k|\hat{x}_{f,k}^{(j)}) \propto \exp\left(\frac{\gamma_k}{2}\right) \exp\left(-\frac{1}{2}\exp(\gamma_k)\right)$$

where $\gamma_k := y_k - \alpha - \hat{x}_{f,k}^{(j)} + \mathbb{E}[\log \epsilon_k^2]$

4. Resampling $\hat{x}_{f,k}^{(j)}$ with weights $w_k^{(j)}$ to get $x_{f,k}^{(j)}$

4.2 Particle Smoothing

Collect $x_{f,k}$ for all $k = 0, \ldots, N$ from the Filtering step to determine the particle smoother.

1) Initialization : Choose the smooth particle $x_{s,N}^{(j)} = x_{f,N}^{(i)}$ with probability $1/M$ for $j = 1, \ldots, M$

2) Recursive Step : For $k = N-1, \ldots, 0$

1. Calculate $w_{k|k+1}^{(i)} \propto p(x_{s,k+1}^{(j)}|x_{f,k}^{(i)}) = \mathcal{N}(\Phi x_{f,k}^{(j)}, Q)$ for $i = 1, \ldots, M$
2. Choose $x_{s,k}^{(j)} = x_{f,k}^{(i)}$ with probability $w_{k|k-1}^{(i)}$

Repeat *1)* and *2)* for $j = 1, \ldots, M$ to obtain the particle smoother $x_{s,k}$ for $k = 0, \ldots, N$. These values along with the data set Y_k will be used to estimate the parameters in the EM algorithm.

5 Examples and Results

To demonstrate the procedure for estimating the parameters of the SV model, this section presents two examples using the two data sets: one from the SV model in (6), the other from the Foreign Exchange market.

5.1 Simulated Data

The data set is generated from the SV model in (6) where the true parameters are $(\phi, Q, \alpha) = (0.9, 0.5, -13.5)$. The trajectories of the estimated parameters are shown in Fig. 1 with final estimates are $(\hat{\phi}, \hat{Q}, \hat{\alpha}) = (0.8738, 0.4667, -13.1068)$. Apparently, the estimated parameters are close to the true ones.

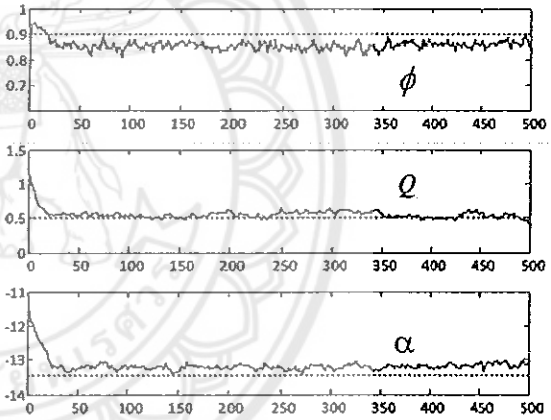


Fig. 1: $(\hat{\phi}, \hat{Q}, \hat{\alpha})$ from the simulated data

5.2 US Dollar/Thai Baht Daily Exchange Rates

The data set used in this example is the USD/THB daily exchange rates from August 1st, 2012 to June 15th, 2015. The trajectories of the estimated parameters are shown in Fig. 2 with final estimates are $(\hat{\phi}, \hat{Q}, \hat{\alpha}) = (0.9353, 0.0375, -13.1216)$.

6 Conclusions

In this paper, we present the SV model in the linear state-space form with additive log chi-square noise. We apply the EM algorithm coupled with the particle method to estimate the parameters of the model from the observed data. Empirical studies show that the estimates are close to the true ones.

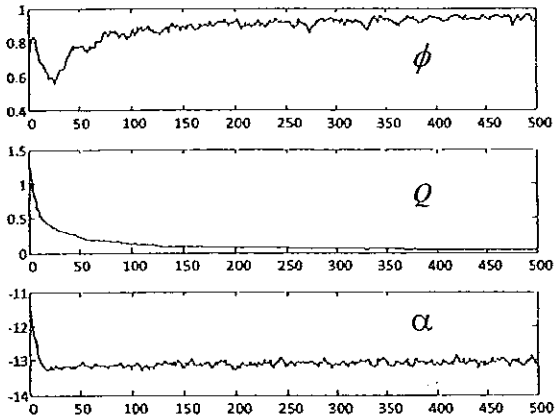


Fig. 2: (ϕ, Q, α) from the USD/THB daily exchange rates

7 Acknowledgments

The authors would like to thank Assist. Prof. Dr. Katechan Jampachaisri for valuable discussions on the log χ^2 distribution. This work was fully supported by Naresuan University under Grant no. P2558C445.

References

[1] KUMNOONSATE, S., "Get Electric Energy Ready for a Sustainable Future," *EPPO Journal*, vol. 103, pp. 7–15, 2014. (in Thai)

[2] AMIN, K. I. & JARROW, R. A., "Pricing foreign currency options under stochastic interest rates," *J. Int. Money Fin.*, vol. 10, pp. 310–329, 1991.

[3] GÖZGÖR, G., MEMIÇ, C. & KARABULUT, G., "The application of stochastic processes in currency exchange rate forecasting and benchmarking for USD-TL and EURO-TL exchange rates," *Proc. Int. Conf. Applied Economics 2010*, Athens, Greece, August, 26–28, 2010.

[4] HULL, J. C. (2005), *Options, Futures, and other Derivatives* (6th Ed.), Pearson Prentice Hall, NJ.

[5] VAN EMMERICH, C., "Modelling correlation as a stochastic process," *Working paper*, Department of Mathematics, University of Wuppertal, Germany.

[6] BLACK, F. & SCHOLES, M., "The pricing of options and corporate liabilities," *J. Pol. Econ.*, vol. 81, pp. 637-654, 1973.

[7] MERTON, R. C., "The theory of rational option pricing," *Bell J. Econ.*, vol. 4, pp. 141–183, 1973.

[8] DEMPSTER, A. P., LAIRD, N. M., & RUBIN, D. B., "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 39, No. 1, pp. 1–38, 1977.

[9] CHATTERJEE, S., "Application of the Kalman Filter for Estimating Continuous Time Term Structure Models–the Case of UK and Germany," Department of Economics, University of Glasgow, January 2005.

[10] RACICOT, F. É. & THÉORET, R., "Forecasting Stochastic Volatility Using the Kalman Filter: An Application to Candaian Interest Rates and Price-Earning Ratio," *AESTIMATIO, the IEB Int. J. of Fin.*, vol. 1, pp. 28–47, 2010.

[11] RUIZ, E., "Quasi-maximum likelihood estimation of stochastic volatility models," *J. Econometrics*, vol 63, pp. 289–306, 1994.

[12] SHEPHARD, N. & XIU, D., "Econometric analysis of multivariate realised QML: efficient positive semi-definite estimators of the covariation of equity prices," *Discussion paper series*, Department of Economics, University of Oxford, UK., Apr. 2012.

[13] ABABSA, F., MALLEM, M. & ROUSSEL, D., "Comparison Between Particle Filter Approach and Kalman Filter-Based Technique For Head Tracking in Augmented Reality Systems," *Proc. IEEE Int. Conf. Robot.*, New Orleans, LA, April 2004.

[14] DOWD, M. & JOY, R., "Estimating behavioral parameters in animal movement models using a state-augmented particle filter," *Ecology*, vol. 92, No. 3, pp. 568–575, 2011.

[15] PAPILLI, M. & CALWAY, A., "Real-time camera tracking using a particle filter," In Proc. British Machine Vision Conference (BMVC'05), pp. 519–528, Oxford, September 2005.

[16] WALKER, E. A., "Comparison of a particle filter and other state estimation methods for prognostics of Lithumion batteries," *Master's thesis*, College of Engineering and Computing, University of South Carolina, Columbia, SC, 2013.

[17] TAYLOR, S. J. (1986), *Modelling Financial Time Series*, John Wiley: Chichester, UK.

Tanit Malakorn received the B.Eng. (Hons.) degree in Control Engineering from KMITL, Thailand in 1995, and the M.S. and Ph.D. degrees in Electrical Engineering from Virginia Tech (VPI & SU), USA in 1999 and 2003, respectively. He joined the faculty at Naresuan University, where he is currently Asscociate Professor of Electrical Engineering. His research interests include H_∞ control, multidimensional linear systems, multiscale theory and mathematical finance.

Thanapat Iamtan received the B.Eng. degree in Electrical Engineering and the M.Eng. degree in Engineering Management from Naresuan University, in 2006 and 2009, respectively. He is currently pursuing the Ph.D. degree in Electrical Engineering at Naresuan University. He won the best paper awards from the OR Network conference and the IE Network conference in 2009 and 2010, respectively. His research interests include machine layout design, Heuristic search and Stochastic process.