บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการวิจัยเพื่อเปรียบเทียบวิธีการหาตัวประมาณพารามิเตอร์ที่เหมาะสมของสมการถดถอยพหุนามเมื่อพิจารณาความคลาดเคลื่อนในตัวแปรอิสระด้วย โดยที่ผู้วิจัยคาดว่าวิธี WLS จะให้ผลดีที่สุด ดังนั้นผู้วิจัยจึงทำการเปรียบเทียบวิธีการหาตัวประมาณพารามิเตอร์โดยวิธี WLS กับวิธีต่าง ๆ ว่าจะมีประสิทธิภาพดีเพียงพอหรือไม่ การวิจัยจะศึกษาในสถานการณ์ต่าง ๆ ที่กำหนดขึ้นดังนี้

1. การแจกแจงของความคลาดเคลื่อนสุ่มในตัวแปรตาม (ε^{*}) เป็นแบบปกติที่มีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 0.1, 0.3, 0.5, 0.7 และ 1.0
2. การแจกแจงของความคลาดเคลื่อนในตัวแปรอิสระเป็นแบบปกติที่มีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 0.1, 0.3, 0.5 และ 0.7
3. ขนาดตัวอย่างที่ศึกษาเป็น 15, 30, 50, 100 และ 200
4. กำลังสูงสุดของตัวแปรอิสระที่ใช้สำหรับการสร้างตัวแปรตามในตัวแบบถดถอยพหุนามเป็น 2, 3, 4, 5 และ 6

การสรุปผลว่าวิธีใดเป็นวิธีที่ดีที่สุดจะพิจารณาจากค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ (*ARRMSE*) และค่าอัตราส่วนผลต่างของค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ (*RDARRMSE*) โดยที่วิธีใดให้ค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ และค่าอัตราส่วนผลต่างของค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ที่มีค่าต่ำสุดจะเป็นวิธีที่ดีที่สุด ผลการวิจัยมีข้อสรุปดังนี้

ปัจจัยที่มีผลต่อค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ของแต่ละวิธี

1. <u>กำลังสูงสุดของตัวแปรอิสระที่ใช้ในการสร้างตัวแปรตามในตัวแบบถดถอยพหุนาม (MB)</u> เมื่อ MB เพิ่มขึ้น วิธี OLS ALS และ WLS จะได้ค่า *ARRMSE* เพิ่มขึ้นตามไปด้วย เนื่องจากค่า y ที่พยากรณ์ได้จากตัวแบบซึ่งประกอบด้วยตัวแปรที่มีความสัมพันธ์กันเองจึงมีผลทำให้ค่า *ARRMSE* มีค่าเพิ่มขึ้นเมื่อ MB เพิ่มสูงขึ้น

กรณี MB $=2$, $n=15$ เมื่อ $\sigma_{\varepsilon^{*}}$ และ σ_{δ} มีค่าไม่มากนัก ($\sigma_{\varepsilon^{*}} \leq 0.3$ และ $\sigma_{\delta} \leq 0.3$) พบว่าวิธี OLS และวิธี ALS จะให้ค่า *ARRMSE* ต่ำกว่าวิธี WLS เพราะว่าวิธี WLS เป็นวิธีที่ขึ้นอยู่กับค่าพารามิเตอร์ r โดยที่ $r = \dfrac{\sigma_{x^{*}}^{2}}{\sigma_{x^{*}}^{2} + \sigma_{\delta}^{2}}$ และค่า r นี้ควรมีค่าอยู่ในช่วง $0.7 \leq r \leq 0.8$ นั่นคือ ค่า σ_{δ} ควรมีค่าอยู่ในช่วง $0.5 \leq \sigma_{\delta} \leq 0.7$ จึงจะทำให้วิธี WLS เป็นวิธีที่

ดีที่สุด แต่อย่างไรก็ตามเมื่อ MB หรือ n เพิ่มขึ้น วิธี WLS จะยังคงเป็นวิธีที่ดีที่สุดแม้ว่าค่า r จะมีค่าอยู่ในช่วงดังกล่าว ทั้งนี้เนื่องจากการใช้ค่า r ที่เหมาะสมมีผลกระทบต่อค่า $ARRMSE$ น้อยกว่าการเพิ่มขึ้นของ MB หรือ n

2. ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระ (σ_{δ}) ค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์($ARRMSE$) ของทุกวิธีมีแนวโน้มเพิ่มขึ้นเมื่อส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระมีค่ามากขึ้น เพราะ $\hat{y}_i$ มีความสัมพันธ์กับ δ_i นั่นคือหาก δ_i มีการกระจายมากจะมีผลทำให้ $\hat{y}_i$ กระจายมาก(รูปแบบความสัมพันธ์เป็นไปตามที่กล่าวไว้ในบทที่ 2) ในขณะที่ y_i^* ไม่ได้ขึ้นอยู่กับ δ_i และมีการกระจายน้อย ดังนั้นจึงส่งผลให้ค่า $ARRMSE$ มีค่ามากตาม

3. ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่ม (σ_{ε}) ค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ ของทุกวิธีมีแนวโน้มเพิ่มขึ้นเมื่อส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มมีค่ามากขึ้น เพราะค่า $ARRMSE$ แปรผันตรงกับ MSE ซึ่งเป็นตัวประมาณค่าความแปรปรวนของความคลาดเคลื่อนสุ่ม σ^2

การเพิ่มของ σ_{δ} เมื่อ σ_{ε} *คงที่* จะมีผลทำให้ค่า $ARRMSE$ เพิ่มขึ้นมากกว่า *การเพิ่มของ* σ_{ε} เมื่อ σ_{δ} *คงที่* เนื่องจาก δ_i เป็นความคลาดเคลื่อนในตัวแปรอิสระ(x_i) หากเกิดความผันแปรใน δ_i มาก(σ_{δ}^2 มาก) ย่อมมีผลทำให้ $\hat{y}_i$ มีการกระจายมาก จึงทำให้ค่า $ARRMSE$ มีแนวโน้มเพิ่มขึ้นมากกว่าการเพิ่ม σ_{ε}^2.

4. ขนาดตัวอย่าง(n) การเพิ่มขนาดตัวอย่างทำให้ค่า $ARRMSE$ ของแต่ละวิธีมีแนวโน้มลดลงเพราะขนาดตัวอย่างช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุได้ ซึ่งแบ่งการพิจารณาได้ดังนี้

ก) วิธี OLS

ค่า $ARRMSE$ ของวิธี OLS มีแนวโน้มที่จะคงเส้นคงวาเมื่อ $n = 100$ และเมื่อ $n = 200$ ความคงเส้นคงวาจะไม่ค่อยแตกต่างจาก $n = 100$ มากนักหรือกล่าวได้ว่าความคงเส้นคงวาจะไม่ค่อยเด่นชัดนักเมื่อหน่วยตัวอย่างมากกว่า 100

ข) วิธี WLS และ ALS

ค่า $ARRMSE$ ของวิธี OLS มีแนวโน้มที่จะคงเส้นคงวาเมื่อ $n = 50$ และเมื่อ $n = 100$ 200 ความคงเส้นคงวาจะไม่ค่อยแตกต่างจาก $n = 50$ มากนักหรือกล่าวได้ว่าความ คงเส้นคงวาจะไม่ค่อยเด่นชัดนักเมื่อหน่วยตัวอย่างมากกว่า 50

เมื่อเกิดสถานการณ์ดังกล่าวจึงมีผลทำให้ค่า $RDARRMSE$ ของวิธี OLS นั้นมีค่าสูงมาก สำหรับทุกค่าของ MB แต่จะไม่ค่อยชัดเจนมากนักเมื่อ MB = 6

การเปรียบเทียบค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ของแต่ละวิธี

โดยทั่วไปวิธี WLS จะดีกว่าวิธีอื่น ๆ ยกเว้นในบางกรณีที่ MB=2 , n , σ_δ และ σ_{ε} มีค่าน้อย ๆ วิธี OLS และ ALS จะดีกว่า WLS แต่ในกรณีที่ MB>2 , n , σ_δ และ σ_{ε} มีค่ามาก พบว่าวิธี WLS จะเป็นวิธีที่ดีมากและเริ่มคงเส้นคงวาเร็วกว่าวิธี OLS จึงมีผลทำให้ค่าอัตราส่วนผลต่างของค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลิยสัมพัทธ์ของวิธี OLS มีค่ามาก ส่วน ALS นั้นจะเริ่มคงเส้นคงวาเมื่อ $n=50$ เช่นเดียวกับวิธี WLS ดังนั้นค่าอัตราส่วนผลต่างของค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ของวิธี ALS จึงมีค่าต่ำ

ข้อเสนอแนะ

1. การวิจัยครั้งนี้มีขอบเขตการวิจัยที่ $\delta_i \sim N(0,\sigma_\delta^2)$ โดยที่ $\sigma_\delta = 0.1,\ 0.3,\ 0.5$ และ 0.7 ซึ่งผลการวิจัยภายใต้ขอบเขตการวิจัยนี้พบว่าวิธี WLS เป็นวิธีที่ดีที่สุด สำหรับการวิจัยครั้งต่อไปควรมีการเพิ่มค่า μ_δ และ σ_δ

2. การวิจัยครั้งนี้มีขอบเขตการวิจัยที่ $x_i^* \sim N(\mu_{x^*},\sigma_{x^*}^2)$ ซึ่งผลการวิจัยภายใต้ขอบเขตการวิจัยนี้พบว่าวิธี WLS เป็นวิธีที่ดีที่สุด แต่ถ้า x_i^* มีฟังก์ชันการแจกแจงเป็นแบบอื่น ๆ เช่น $F(x^*)=(1-\eta)\phi(x^*)+\eta\phi(\frac{x^*}{1.5})$ เมื่อ $\eta = 0.1,\ 0.3$ และ $\phi(x^*)$ เป็นฟังก์ชันการแจกแจงแบบปกติมาตรฐาน ($N(0,1)$) พบว่าวิธี ALS จะเป็นวิธีที่ดีที่สุดและมีความคงเส้นคงวา ส่วนวิธี WLS และ OLS นั้นจะไม่คงเส้นคงวาทั้งสองวิธี [6]

3. ผลการวิจัยที่ได้พบว่าค่า $ARRMSE$ ของวิธี ALS ที่ได้มีค่าใกล้เคียงกับวิธี WLS มาก ซึ่งในทางปฏิบัตินั้นการหาค่าตัวประมาณโดยวิธี ALS จำเป็นต้องทราบค่าที่แท้จริงของ σ_δ^2 ซึ่งเป็นการยากมาก หรือถ้าทำได้จะต้องเสียค่าใช้จ่ายเป็นจำนวนมากดังนั้นวิธี WLS จึงเป็นวิธีเหมาะสมในทางปฏิบัติมากกว่าวิธี ALS

4. ผู้วิจัยได้ทดลองทำกรณีที่ x^* มีค่าต่ำมาก ($x_i^* \sim N(0,1)$) หรือสูงมาก ($x_i^* \sim N(100,\sigma_{x^*}^2)$) พบว่าให้ค่าค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ที่สูงมากกว่าปกติ ทั้งนี้อาจมีผลเนื่องมาจากความคลาดเคลื่อนในขั้นตอนการคำนวณค่า $(X'X)^{-1}$

5. ผู้วิจัยได้ทดลองทำกรณีที่ไม่พิจารณาความคลาดเคลื่อนในตัวแปรอิสระ พบว่าที่กำลังสูงสุดของตัวแปรอิสระที่ใช้สำหรับการสร้างตัวแปรตามในตัวแบบถดถอยพหุนาม (MB) เท่ากับ 2 วิธี OLS เป็นวิธีที่ดีที่สุดสำหรับทุกค่าของขนาดตัวอย่าง และความคลาดเคลื่อนสุ่มที่เกิด

---

[6] S.Huang & L.Huwang." On the Polynomial Structural Relationship." *The Canadian Journal Statistics*.29.2001

จากตัวแปรตามที่เปลี่ยนแปลงไป และพบว่าที่กำลัง 3, 4, 5 และ 6 วิธี WLS เป็นวิธีที่ดีที่สุดสำหรับทุกค่าของขนาดตัวอย่าง และความคลาดเคลื่อนสุ่มที่เกิดจากตัวแปรตามที่เปลี่ยนแปลงไป รองลงมาเป็นวิธี ALS และวิธี OLS ตามลำดับ

ตารางที่ 5.1 แสดงวิธีที่ให้ค่า $ARRMSE$ ต่ำสุดในสถานการณ์ต่างๆ ซึ่งจำแนกตาม MB σ_δ σ_ε และ n

MB	σ_δ	σ_ε	n				
			15	30	50	100	200
2	0.1	0.1	OLS	WLS	WLS	WLS	WLS
		0.3	OLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.3	0.1	OLS	WLS	WLS	WLS	WLS
		0.3	ALS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.5	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.7	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS

ตารางที่ 5.1(ต่อ) แสดงวิธีที่ให้ค่า $ARRMSE$ ต่ำสุดในสถานการณ์ต่างๆ ซึ่งจำแนกตาม MB σ_{δ} $\sigma_{\varepsilon^{*}}$ และ n

MB	σ_{δ}	$\sigma_{\varepsilon^{*}}$	n				
			15	30	50	100	200
3	0.1	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.3	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.5	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.7	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS

ตารางที่ 5.1(ต่อ) แสดงวิธีที่ให้ค่า $ARRMSE$ ต่ำสุดในสถานการณ์ต่างๆ ซึ่งจำแนกตาม MB σ_δ σ_ε และ n

MB	σ_δ	σ_ε	n				
			15	30	50	100	200
4	0.1	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.3	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.5	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.7	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS

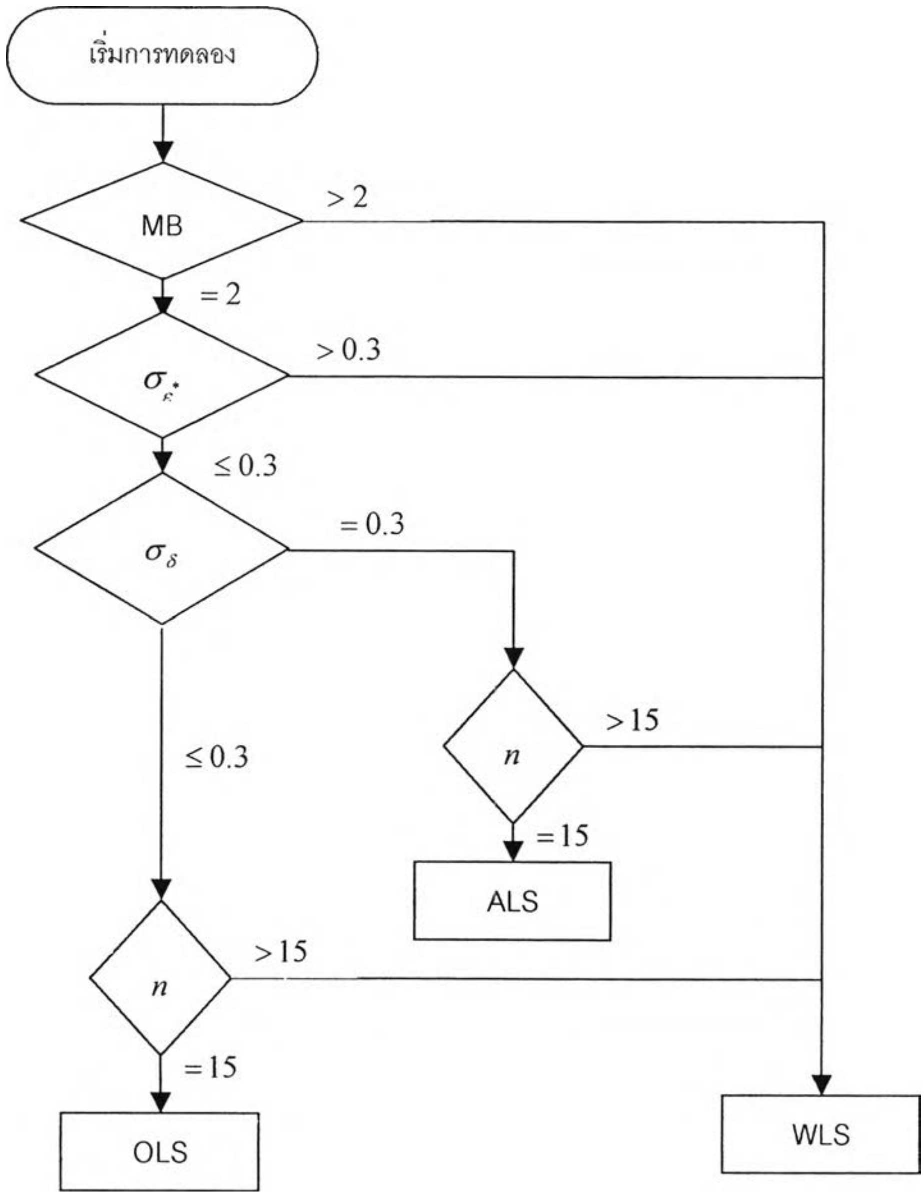
ตารางที่ 5.1(ต่อ) แสดงวิธีที่ให้ค่า $ARRMSE$ ต่ำสุดในสถานการณ์ต่างๆ ซึ่งจำแนกตาม MB σ_{δ} σ_{ε} และ n

MB	σ_{δ}	σ_{ε}	n				
			15	30	50	100	200
5	0.1	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.3	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.5	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.7	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS

ตารางที่ 5.1(ต่อ) แสดงวิธีที่ให้ค่า $ARRMSE$ ต่ำสุดในสถานการณ์ต่างๆ ซึ่งจำแนกตาม MB σ_{δ} σ_{ε} และ n

MB	σ_{δ}	σ_{ε}	n				
			15	30	50	100	200
6	0.1	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.3	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.5	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS
	0.7	0.1	WLS	WLS	WLS	WLS	WLS
		0.3	WLS	WLS	WLS	WLS	WLS
		0.5	WLS	WLS	WLS	WLS	WLS
		0.7	WLS	WLS	WLS	WLS	WLS
		1.0	WLS	WLS	WLS	WLS	WLS

แผนผังแสดงผลสรุปการเลือกวิธีการประมาณค่าพารามิเตอร์ในการวิเคราะห์ความถดถอยพหุนาม กรณีที่มีความคลาดเคลื่อนในตัวแปรอิสระ ตามทฤษฎี

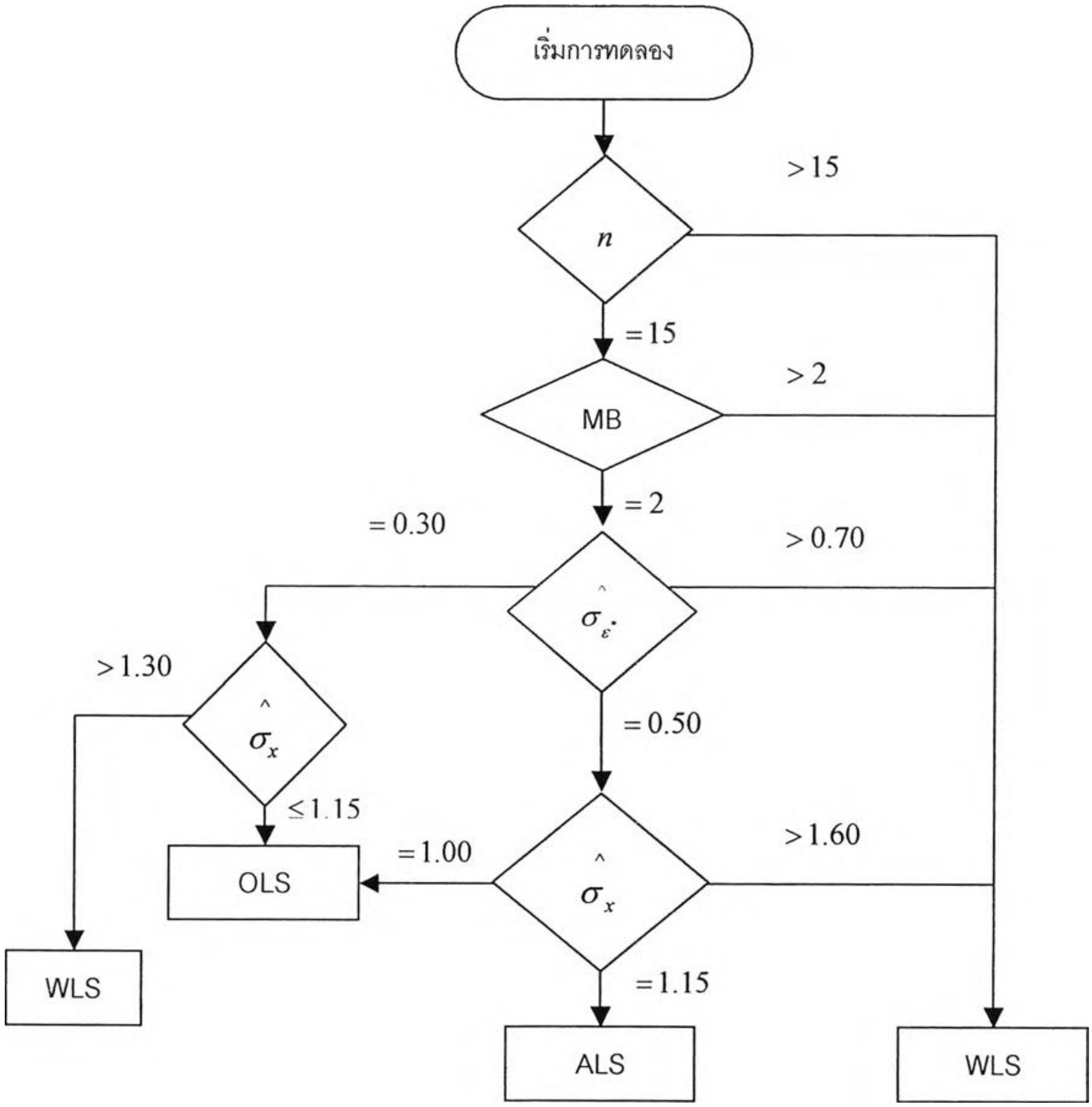


หมายเหตุ

MB หมายถึง กำลังสูงสุดของตัวแปรอิสระที่ใช้สำหรับการสร้างตัวแปรตามในตัวแบบถดถอยพหุนาม

การสรุปการเลือกวิธีการประมาณค่าพารามิเตอร์ดังกล่าวนี้ใช้เป็นแนวทางเมื่อพิจารณาความคลาดเคลื่อนในตัวแปรอิสระ แต่ถ้าข้อมูลไม่มีความคลาดเคลื่อนในตัวแปรอิสระ และ MB = 2 ควรใช้วิธีการประมาณค่าแบบ OLS แต่ถ้า MB = 3,4,5 และ 6 ควรใช้วิธีการประมาณค่าแบบ WLS

แผนผังแสดงผลสรุปการเลือกวิธีการประมาณค่าพารามิเตอร์ในการวิเคราะห์ความถดถอยพหุนาม กรณีที่มีความคลาดเคลื่อนในตัวแปรอิสระ สำหรับใช้ในทางปฏิบัติ



หมายเหตุ

$\hat{\sigma}_{\varepsilon}$ คือ รากที่สองของค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสอง

$\hat{\sigma}_x$ คือ ค่าส่วนเบี่ยงเบนมาตรฐานของข้อมูลตัวอย่างของตัวแปรอิสระ (x)