บทที่ 4
ADAPTIVE FILTERING

ในบทที่ 3 เทคนิคการพยากรณ์ 2 เทคนิคสามารถใช้ได้กว้างขวางสำหรับสถานการณ์ช่วงระยะเวลาสั้น สำหรับสถานการณ์การพยากรณ์หลาย ๆ แบบต้องเตรียมการเทคนิคการพยากรณ์ทั้ง moving average และ exponential smoothing ค่าพยากรณ์จะอยู่บนพื้นฐานค่าข้อมูลเดิม ในบทนี้มีอีกวิธีคือ adaptive filtering ซึ่งค่าพยากรณ์ก็จะอยู่บนพื้นฐานค่าข้อมูลเดิมของอนุกรมเวลา เช่นเดียวกัน แต่ให้ผลลัพธ์ที่ดีกว่า 2 เทคนิคดังกล่าว โดยเฉพาะสถานการณ์ภายใต้รูปแบบพื้นฐานของข้อมูลที่ยุ่งยากซับซ้อนกว่า ก่อนที่จะอธิบายถึงเทคนิค adaptive filtering ทำงานอย่างไรและสามารถใช้อย่างไร เราควรจะพิจารณาประโยชน์ของความสัมพันธ์ระหว่างเทคนิควิธี moving average และ exponential smoothing และ หลักเกณฑ์ของ adaptive filtering การพยากรณ์ของ adaptive filtering ค่าพยากรณ์เขียนอยู่ในรูปผลรวมถ่วงน้ำหนักของค่าที่เกิดขึ้นจริง ในคาบเวลาก่อน ๆ คือ

$$F_{t+1} = w_1 X_t + w_2 X_{t-1} + w_3 X_{t-2} + \ldots + w_N X_{t-N+1}$$

หรือ $$F_{t+1} = \sum_{i=1}^{N} w_i X_{t-i+1}$$

เมื่อ t เป็นคาบเวลา $t = N, N+1, N+2, \ldots, n$ และ $i = 1, 2, \ldots, N$

F เป็นค่าพยากรณ์สำหรับคาบเวลาที่ $t+1$

w_i เป็นค่าถ่วงน้ำหนักที่สอดคล้องกับค่าที่เกิดขึ้นจริงในคาบเวลา $t-i+1$

X_t เป็นค่าที่เกิดขึ้นจริงที่คาบเวลา t

N เป็นจำนวนตัวถ่วงน้ำหนัก

<u>ขั้นตอนของการพยากรณ์วิธี adaptive filtering</u>

1. เก็บรวบรวมข้อมูลค่าที่เกิดขึ้นจริงใน n คาบเวลาที่ผ่านมา คือ $X_1, X_2, X_3, \ldots, X_n$

2. นำค่าข้อมูล N ค่าแรก คือ $X_1, X_2, X_3, \ldots, X_N$ (เมื่อ N เป็นจำนวนตัวถ่วงน้ำหนัก) เมื่อนำไปคำนวณหาค่าพยากรณ์ในคาบเวลาที่ N+1 (F_{N+1})

3. กำหนดค่าถ่วงน้ำหนักที่เหมาะสมเป็น $w_1, w_2, w_3, \ldots, w_N$ จะได้

$$F_{N+1} = w_1 X_N + w_2 X_{N-1} + w_3 X_{N-2} + \ldots + w_N X_1$$

4. คำนวณค่าความคลาดเคลื่อน $e_{N+1} = X_{N+1} - F_{N+1}$ แล้วนำค่า e_{N+1} ใช้ปรับค่า $w_1, w_2, w_3, \ldots, w_N$ ใหม่ โดยปรับค่าถ่วงน้ำหนักได้จาก

$w_i' = w_i + 2k\, e_{t+1} X_{t-i+1}$

โดย $i = 1,2,\ldots,N$ และ $t = N+1, N+2, N+3, \ldots, n$

เมื่อ w_i' = ค่าถ่วงน้ำหนักตัวที่ i ที่ได้ปรับค่าแล้ว

w_i = ค่าถ่วงน้ำหนักตัวที่ i ก่อนทำการปรับค่า

k = ค่าคงที่ (learning constant)

e_{t+1} = ค่าความคลาดเคลื่อนจากการพยากรณ์ในคาบเวลาที่ t + 1

X_{t-i+1} = ค่าข้อมูลจริงในคาบเวลาที่ t - i + 1

5. นำค่า $w_1', w_2', w_3', \ldots, w_N'$ ซึ่งเป็นค่าถ่วงน้ำหนักที่ปรับแล้ว โดยเปลี่ยนเป็นสัญลักษณ์ $w_1, w_2, w_3, \ldots, w_N$ เพื่อใช้เป็นค่าถ่วงน้ำหนักในการหาค่าพยากรณ์ F_{N+2} โดย

$F_{N+2} = w_1 X_{N+1} + w_2 X_N + w_3 X_{N-1} + \ldots + w_N X_2$

6. คำนวณหาค่าความคลาดเคลื่อน $e_{N+2} = X_{N+2} - F_{N+2}$ จากนั้นนำค่า e_{N+2} ใช้ปรับค่า $w_1, w_2, w_3, \ldots, w_N$ ใหม่ โดยใช้สมการของการถ่วงน้ำหนักดำเนินการจนกระทั่งได้ค่าพยากรณ์ F_{n+1} ได้จาก $F_{n+1} = w_1 X_n + w_2 X_{n-1} + w_3 X_{n-2} + \ldots + w_N X_{n-N+1}$

7. เมื่อได้ค่าพยากรณ์ครบในรอบที่ 1 ให้คำนวณหาค่า E ซึ่งมีค่าเท่ากับ

$\sum_{i=N+1}^{n} e_i^2$ แต่ถ้า ค่า E มีค่ามาก ต้องดำเนินการกลับไปหาค่า F_{n+1} ในรอบที่สองใหม่ โดยใช้ค่าถ่วงน้ำหนัก $w_1, w_2, w_3, \ldots, w_N$ ชุดสุดท้ายในรอบแรกเป็นชุดค่าถ่วงน้ำหนักเริ่มต้นในรอบที่สองใหม่ จนกระทั่งได้ค่า F_{n+1} ในรอบที่สองหรืออาจดำเนินการ ซ้ำ ๆ กันหลาย ๆ รอบได้

8. เมื่อได้ E ที่มีค่าน้อยที่สุดในรอบใด จะได้ว่า ค่าถ่วงน้ำหนัก $w_1^*, w_2^*, w_3^*, \ldots, w_N^*$ ลู่เข้า (converge) หาค่าที่ optimal แล้ว

9. นำค่าถ่วงน้ำหนักชุดที่ optimal ไปคำนวณหาค่าพยากรณ์ F_{n+1}, F_{n+2} ได้คือ

$F_{n+1} = w_1^* X_n + w_2^* X_{n-1} + w_3^* X_{n-2} + \ldots + w_N^* X_{n-N+1}$

และ $F_{n+2} = w_1^* X_{n+1} + w_2^* X_n + w_3^* X_{n-1} + \ldots + w_N^* X_{n-N+2}$

เมื่อทราบค่าที่เกิดขึ้นจริงในคาบเวลาที่ n+1 (X_{n+1})

ตัวอย่างที่ 4.1 ค่าข้อมูลที่เก็บรวบรวมถึงคาบที่ 20 โดยมี N = 5 ต้องการหาค่าพยากรณ์ของคาบเวลาที่ 6,7 และ 21 พร้อมทั้งแสดงวิธีการปรับค่าถ่วงน้ำหนัก

คาบที่	1	2	3	4	5	6 7 8 9 10 11 12 13 14 15	16	17	18	19	20
ข้อมูล	X_1	X_2	X_3	X_4	X_5	X_6 X_7 X_8 X_9 X_{10} X_{11} X_{12} X_{13} X_{14} X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}
ค่าถ่วงน้ำหนัก	w_5	w_4	w_3	w_2	w_1		w_5	w_4	w_3	w_2	w_1

$$F_6 = w_1 X_5 + w_2 X_4 + w_3 X_3 + w_4 X_2 + w_5 X_1$$

$e_6 = X_6 - F_6$ นำค่า e_6 ไปใช้ปรับค่า w_1, w_2, w_3, w_4, w_5 ใหม่ได้

$$w_1' = w_1 + 2k\, e_6 X_5 \qquad w_2' = w_2 + 2k\, e_6 X_4$$

$$w_3' = w_3 + 2k\, e_6 X_3 \qquad w_4' = w_4 + 2k\, e_6 X_2$$

$$w_5' = w_5 + 2k\, e_6 X_1$$

นำค่า $w_1', w_2', w_3', w_4', w_5'$ ไปใช้ในการหาค่าพยากรณ์ของคาบเวลาที่ 7 คือ

$$F_7 = w_1' X_6 + w_2' X_5 + w_3' X_4 + w_4' X_3 + w_5' X_2$$

จนกระทั่งได้ค่าตัวถ่วงน้ำหนักชุดสุดท้ายเป็น w_1, w_2, w_3, w_4, w_5 เพื่อใช้ในการพยากรณ์ค่าข้อมูลของคาบเวลาที่ 21 คือ

$$F_{21} = w_1 X_{20} + w_2 X_{19} + w_3 X_{18} + w_4 X_{17} + w_5 X_{16}$$

ตัวอย่างที่ 4.2 ข้อมูลที่ต้องการพยากรณ์โดยวิธี adaptive filtering 10 คาบเวลามีตัวถ่วงน้ำหนัก คือ $w_1 = 0.5$, $w_2 = 0.5$ โดยมี learning constant (k) = 0.9 จงหาค่าพยากรณ์ของคาบเวลาที่ 11 โดยให้มีค่าถ่วงน้ำหนักที่ optimal ใช้ในการพยากรณ์เท่านั้น

คาบที่	1	2	3	4	5	6	7	8	9	10
ค่าข้อมูล	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0

$$F_3 = w_1 X_2 + w_2 X_1 = (0.5)(0.2) + (0.5)(0.1) = 0.15$$

$$e_3 = X_3 - F_3 = 0.3 - 0.15 = 0.15$$

นำค่า e_3 ไปใช้ปรับค่า w_1 และ w_2 ใหม่ได้

$$w_1' = w_1 + 2k\, e_3 X_2 = 0.5 + 2(0.9)(0.15)(0.2) = 0.554$$

$$w_2' = w_2 + 2k\, e_3 X_1 = 0.5 + 2(0.9)(0.15)(0.1) = 0.527$$

$$F_4 = (0.554)(0.3) + (0.527)(0.2) = 0.27$$

<u>รอบที่ 1</u>

คาบที่	X_t	F_{t+1}	ค่าถ่วงน้ำหนัก		e_t	e_t^2
			w_1	w_2		
1	0.1					
2	0.2		0.500	0.500		
3	0.3	0.15	0.554	0.527	0.15	0.0225
4	0.4	0.27	0.624	0.574	0.13	0.0169

คาบที่ t	X_t	F_{t+1}	ค่าถ่วงน้ำหนัก		e_t	e_t^2
			w_1	w_2		
5	0.5	0.42	0.682	0.617	0.08	0.0064
6	0.6	0.59	0.691	0.624	0.01	0.0001
7	0.7	0.73	0.659	0.597	-0.03	0.0009
8	0.8	0.82	0.634	0.575	-0.02	0.0004
9	0.9	0.91	0.620	0.562	-0.01	0.0001
10	1.0	1.01	0.604	0.548	-0.01	0.0001
11		1.10				E = 0.0474

ค่า w_1' และ w_2' สำหรับค่าพยากรณ์ในคาบเวลาที่ 11 ได้จาก

$$0.604 = 0.620 + 2(0.9)(-0.01)(0.9)$$

และ $0.548 = 0.562 + 2(0.9)(-0.01)(0.8)$

ได้ $F_{11} = (0.604)(1) + (0.548)(0.9) = 1.10$

โดยมี $E = 0.0474$ และค่าถ่วงน้ำหนัก $w_1 = 0.604, w_2 = 0.548$ ใช้เป็นค่าถ่วงน้ำหนักเริ่มต้นของรอบที่สองใหม่

<u>รอบที่ 2</u>

คาบที่	X_t	F_{t+1}	ค่าถ่วงน้ำหนัก		e_t	e_t^2
			w_1	w_2		
1	0.1					
2	0.2		0.604	0.548		
3	0.3	0.18	0.647	0.570	0.12	0.0144
4	0.4	0.31	0.696	0.602	0.09	0.0081
5	0.5	0.46	0.725	0.624	0.04	0.0016
6	0.6	0.61	0.716	0.617	-0.01	0.0001
7	0.7	0.74	0.673	0.581	-0.04	0.0016
8	0.8	0.82	0.648	0.559	-0.02	0.0004
9	0.9	0.91	0.634	0.546	-0.01	0.0001
10	1.0	1.01	0.618	0.532	-0.01	0.0001
11		1.10				E = 0.0264

ได้ $F_3 = (0.604)(0.2) + (0.548)(0.1) = 0.18$

$e_3 = 0.3 - 0.18 = 0.12$

ปรับค่า w_1' และ w_2' ได้ดังนี้

ค่า $w_1' = 0.604 + 2(0.9)(0.12)(0.2) = 0.647$

$w_2' = 0.548 + 2(0.9)(0.12)(0.1) = 0.570$

ได้ $F_4 = (0.647)(0.3) + (0.570)(0.2) = 0.31$

คำเนินการเช่นนี้เรื่อยไปจนกระทั่งได้ค่าพยากรณ์ในคาบเวลาที่ 11 (F_{11}) คือ

$F_{11} = (0.618)(1.0) + (0.532)(0.9) = 1.10$

ในรอบที่สองได้ค่า $E = 0.0264$

จากตัวอย่างที่ 4.2 จะเห็นว่าแต่ละรอบหรือแต่ละครั้งที่ทำการพยากรณ์ จะได้ค่า w_1, w_2 ชุดใหม่เสมอ ค่า w_1, w_2 ชุดสุดท้ายของการพยากรณ์รอบที่สองจะใช้เป็นค่าเริ่มต้นในรอบที่สาม ได้ ซึ่งถ้าดำเนินการกระทำซ้ำ ๆ กันเช่นนี้ต่อไป ค่า w_1, w_2 ชุดสุดท้ายในรอบที่ k จะเป็นค่าที่ทำให้ค่าพยากรณ์ดีขึ้นเรื่อย ๆ และเมื่อทำจำนวนรอบมากพอที่จะทำให้ค่า w_1, w_2 converge เข้าหาค่าที่ optimal คือเป็นค่าที่นำมาใช้ในการพยากรณ์ แล้ว จะได้ผลรวมของกำลังสองของความคลาดเคลื่อนต่ำที่สุด คือ

$E = \sum_{i=N+1}^{n} e_i^2$ และในกรณีที่การพยากรณ์นั้น ๆ ต้องการให้มีอัตรา converge สูงขึ้น อาจปรับค่าถ่วงน้ำหนักและค่าความคลาดเคลื่อนใหม่ ได้ดังนี้ คือ

$$w_i' = w_i + 2k\, e^*_{t+1}\, X^*_{t-i+1}$$

โดย $e^*_{t+1} = e_{t+1} / \sqrt{\sum X^2_{t-i+1}}$

$$X^*_{t-i+1} = X_{t-i+1} / \sqrt{\sum_{i=1}^{N} X^2_{t-i+1}}$$

ในกรณีดังกล่าวอาจใช้ค่า $k = 1/N$ เป็นค่าที่เหมาะสมในการพยากรณ์และการเลือกใช้ค่า w_i สำหรับชุดเริ่มต้นอาจเป็น $w_i = 1/N$ ซึ่งเทคนิคการพยากรณ์วิธี adaptive filtering มีข้อเสียเปรียบตรง ต้องกำหนดค่า k ให้เหมาะสมกับข้อมูลอนุกรมเวลา ดังนั้น จึงต้องอาศัยผู้มีประสบการณ์ในการทำการพยากรณ์อาจช่วยลดจำนวนรอบของการคำนวณได้ อีกจุดหนึ่งที่ต้องคำนึงถึง คือการกำหนดค่า w_i สำหรับชุดเริ่มต้นต้องเลือกใช้เพื่อให้เกิดการลู่เข้าได้ง่าย

แบบฝึกหัด

1. ข้อมูลต่อไปนี้จงหาค่าพยากรณ์ F_5, F_6 และ F_7 โดยวิธี adaptive filtering เมื่อมีค่า optimal weights เป็น $w_1 = 0.54$, $w_2 = 0.45$, $w_3 = -0.46$ และ $w_4 = 0.45$ และหาค่าถ่วงน้ำหนักที่ปรับแล้วเพื่อใช้ในการหาค่า F_6 (คือ ค่า w_1', w_2', w_3', w_4') เมื่อ $k = 8x10^{-6}$ และจงพยากรณ์ค่าข้อมูลในคาบที่ 13

คาบที่	1	2	3	4	5	6	7	8	9	10	11	12
ค่าข้อมูล	19	24	32	63	99	120	144	191	243	280	320	383

2. ยอดขายแคมเปญรายเดือนมีหน่วยเป็นพันขวด โดยให้ค่า $w_i = 1/N = 0.083$; $i = 1,2,...,12$ และ $k = 0.08$ จงพยากรณ์ยอดขายแคมเปญของเดือนที่ 25 แสดงค่าถ่วงน้ำหนักที่ปรับแล้ว และหาค่า Mean Square Error ของแต่ละรอบในการพยากรณ์

ปี 1962

เดือน	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
ยอดขาย	2.851	2.672	2.755	2.721	2.946	3.036	2.282	2.212	2.922	4.301	5.764	7.132

ปี 1963

เดือน	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
ยอดขาย	2.541	2.475	3.031	3.266	3.776	3.230	3.028	1.759	3.595	4.474	6.838	8.357