

ПРАКТИЧНА РОБОТА 9

АДАПТИВНІ МЕТОДИ ПРОГНОЗУВАННЯ

Завдання №1. Метод ковзних середніх.

На основі вхідних даних щодо динаміки щомісячних обсягів продажу продукції підприємства та за допомогою методу ковзних середніх, виконати наступні етапи:

1. Провести фільтрацію вхідних даних для інтервалів згладжування $m = 2$ та $m = 3$;
2. Для кожного з випадків фільтрації даних, зробити свій прогноз на наступний період. Показати його на графіку;
3. Порівняти отримані результати фільтрації даних та прогнозування для обох інтервалів згладжування, зробити висновки.

Вихідні дані для виконання роботи наведено в таблиці 4.1.

Таблиця 4.1 – Вхідні дані для методу ковзних середніх

Період часу t	Обсяг продажів продукції $Y(t)$, млн. грн.
1	4,6
2	9,8
3	$= 9,4 + 1,1 * N$
4	13,3
5	8,4
6	14,1
7	$= 17,2 + 0,9 * N$
8	15,6
9	13,5
10	15,6
11	20,4
12	22,6
13	$= 23,8 + 0,9 * N$

де N – порядковий номер студента по журналу

Рекомендації до виконання. Метод ковзних середніх – це метод короткострокового прогнозування, який заснований на процедурі згладжування (фільтрації) рівнів часового ряду. Згладжений часовий ряд має меншу варіацію,

тому краще відображає дію основних факторів. Просте ковзне середнє (SMA) розраховується за формулою (4.1).

$$SMA(t) = \frac{y(t - m + 1) + \dots + y(t - 1) + y(t)}{m}, \quad (4.1)$$

Де $SMA(t)$ – просте ковзне середнє протягом періоду часу t ; m – кількість інтервалів згладжування; $y(t)$ – значення часового ряду протягом періоду часу t .

Тоді, прогноз на наступний $(t + 1)$ період буде знаходитись, як:

$$Forecast(t + 1) = SMA(t), \quad (4.2)$$

Де $Forecast(t+1)$ – прогнозне значення часового ряду Y на $(t + 1)$ період.

Відповідно до наведеної вище формули (4.1), якщо інтервал згладжування $m = 3$, то просте ковзне середнє буде розраховуватись за формулою:

$$SMA(t) = \frac{y(t - 2) + y(t - 1) + y(t)}{3} = \frac{1}{3}y(t - 2) + \frac{1}{3}y(t - 1) + \frac{1}{3}y(t)$$

Як бачимо, при розрахунку простої ковзної середньої $SMA(t)$, кожний рівень часового ряду $y(t)$ враховується з однаковим ваговим коефіцієнтом $1/m$. Таке «усереднення» рівнів часового ряду згладжує випадкові коливання вхідного показника.

В таблиці 4.2, для вхідного часового ряду Y , методом ковзних середніх були розраховані значення $SMA(t)$, з інтервалом згладжування $m = 3$.

Таблиця 4.2 – Приклад розрахунку простої ковзної середньої для $m = 3$

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Y	4,6	9,8	9,4	13,3	8,4	14,1	17,2	15,6	13,5	15,6	20,4	22,6	23,8	–
SMA			7,9	10,8	10,4	11,9	13,2	15,6	15,4	14,9	16,5	19,5	22,3	22,3

Прогноз на наступний період, згідно формули (4.2), дорівнював останньому розрахованому значенню *SMA*. Графічно, вхідний та отриманий ряди динаміки мають вигляд, як показано на рис. 4.1.

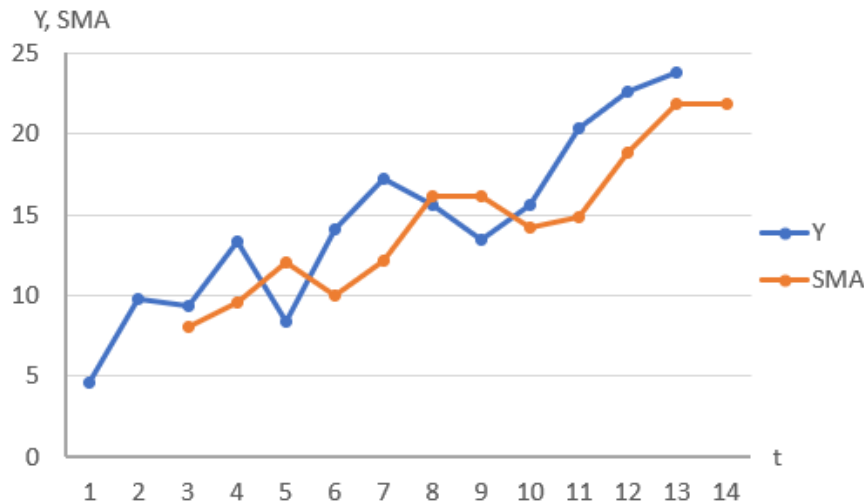


Рис. 4.1. Фільтрація рівнів часового ряду за методом ковзних середніх

Подальший розвиток методу ковзних середніх призводить до використання зважених ковзних середніх. Тобто, кожному рівню ряду, в межах інтервалу згладжування, відповідає певний індивідуальний ваговий коефіцієнт.

Завдання №2. Метод простого експоненційного згладжування.

На основі вхідних даних щодо динаміки щомісячних обсягів продажу продукції підприємства та за допомогою методу експоненційного згладжування, виконати наступні етапи:

1. Провести процедуру простого експоненційного згладжування для параметру адаптації $\alpha = 0,2$;
2. Розрахувати прогнозне значення обсягів продажу на наступний період, показати його на графіку;
3. Провести моделювання параметру адаптації α в діапазоні $(0; 0,3]$ з кроком 0,05 й проаналізувати відповідні зміни прогнозних обсягів випуску продукції.

Вихідні дані для виконання роботи наведено в таблиці 4.1.

Рекомендації до виконання. Метод експоненційного згладжування – це один з найбільш популярних методів прогнозування часових рядів. Як і метод ковзних середніх, він базується на припущенні, що майбутнє значення часового ряду залежить від його попередніх значень. Однак, на відміну від останнього, вагові коефіцієнти кожного попереднього рівня ряду, експоненційно зменшуються з часом.

Розрізняють декілька варіацій даного методу: просте експоненційне згладжування; подвійне, або потрійне експоненційне згладжування тощо. На практиці, найчастіше використовуються перші два методи.

Розглянемо просте експоненційне згладжування, формули якого мають вигляд:

$$\begin{aligned} EMA(0) &= y(1), \\ EMA(t) &= \alpha \times y(t) + (1 - \alpha) \times EMA(t - 1), \end{aligned} \quad (4.3)$$

Де $EMA(t)$ – експоненційне ковзне середнє протягом періоду часу t ; $y(t)$ – значення часового ряду протягом періоду часу t ; α – коефіцієнт згладжування (параметр адаптації), $\alpha \in (0; 1]$.

Тоді, прогноз на наступний $(t + 1)$ період буде знаходитись, як:

$$Forecast(t + 1) = EMA(t), \quad (4.4)$$

Де $Forecast(t+1)$ – прогнозне значення часового ряду Y на $(t + 1)$ період.

Зазвичай, значення параметру згладжування α обирається експериментальним шляхом й найчастіше належить інтервалу $0,01 \dots 0,3$. Також, для його обчислення, може використовуватись формула:

$$\alpha = \frac{2}{m + 1}, \quad (4.5)$$

Де m – кількість інтервалів згладжування.

Незважаючи на свою простоту, даний метод є достатньо ефективним й використовується для прогнозування часових рядів в різних сферах діяльності, зокрема фінансах, логістиці та інших. Однак, прогнозування за допомогою простого експоненційного згладжування може здійснюватися лише на 1 період наперед. Наступний метод буде позбавлений вказаного недоліку.

Завдання №3. Метод подвійного експоненційного згладжування.

На основі вхідних даних щодо динаміки щомісячних обсягів продажу продукції підприємства та за допомогою методу експоненційного згладжування, виконати наступні етапи:

1. Провести процедуру подвійного експоненційного згладжування, обчисливши параметр адаптації α за формулою (4.5);
2. Розрахувати прогнозні значення обсягів продажу на наступні три періоди, показати їх на графіку;
3. Порівняти результати прогнозування з тими, що були отримані за допомогою методу простого експоненційного згладжування.

Вихідні дані для виконання роботи наведено в таблиці 4.1.

Рекомендації до виконання. Метод подвійного експоненційного згладжування складається з наступних етапів:

1. Розрахунок значення параметру адаптації α за формулою (4.5). Прийняти в даному прикладі кількість інтервалів згладжування $m = 5$.
2. На основі вхідних значень рівнів часового ряду $y(t)$, обчислюються параметри a_0 та a_1 лінійної регресії $Y(t) = a_0 + a_1 t$. Для цього можна використовувати вбудовану функцію Microsoft Excel `=ЛИНЕЙН()`, або `=LINEST()`.
3. Обчислені параметри лінійної регресії a_0 та a_1 використовуються для розрахунку початкових значень експоненційних середніх першого $EMA^{[1]}(0)$ та другого $EMA^{[2]}(0)$ порядку, формули (4.6).

$$EMA^{[1]}(0) = a_0 - \frac{1 - \alpha}{\alpha} a_1; \quad EMA^{[2]}(0) = a_0 - \frac{2 \times (1 - \alpha)}{\alpha} a_1, \quad (4.6)$$

4. Обчислення експоненційної середньої першого порядку відбувається за формулою (4.7).

$$EMA^{[1]}(t) = \alpha \times y(t) + (1 - \alpha) \times EMA^{[1]}(t - 1), \quad (4.7)$$

Де $EMA^{[1]}(t)$ – експоненційне ковзне середнє першого порядку протягом періоду часу t ; $y(t)$ – значення часового ряду протягом періоду часу t ; α – коефіцієнт згладжування (параметр адаптації).

5. Обчислення експоненційної середньої другого порядку, для уточнення тенденції, формула (4.8).

$$EMA^{[2]}(t) = \alpha \times EMA^{[1]}(t) + (1 - \alpha) \times EMA^{[2]}(t - 1), \quad (4.8)$$

Де $EMA^{[2]}(t)$ – експоненційне ковзне середнє другого порядку протягом періоду часу t .

6. На основі розрахованих експоненційних середніх $EMA^{[1]}(t)$ та $EMA^{[2]}(t)$ за останній період часу, обчислюються коефіцієнти b_0 та b_1 :

$$b_0 = 2EMA^{[1]} - EMA^{[2]}; \quad b_1 = \frac{\alpha}{1 - \alpha} \times (EMA^{[1]} - EMA^{[2]}), \quad (4.9)$$

Де $EMA^{[1]}$, $EMA^{[2]}$ – відповідно, значення експоненційних ковзних середніх першого та другого порядків, за останній період часу.

7. Прогнозування рівнів часового ряду виконується за формулою (4.10).

$$Y_{\text{розр}}(T) = b_0 + b_1 T, \quad (4.10)$$

Де T – номер періоду прогнозування ($T = 1, 2, \dots$).

Не слід забувати, що хоча метод подвійного експоненційного згладжування дозволяє виконувати прогнозування на довільну кількість періодів наперед, однак, найбільш ефективним він є у випадку короткострокового прогнозування.

Завдання №4. Метод декомпозиції Хольта-Вінтерса при наявності сезонної складової.

На основі вхідних даних щодо динаміки щомісячної кількості переглядів веб-ресурсу за останні чотири роки та за допомогою методу Хольта-Вінтерса, виконати наступні етапи:

1. Побудувати прогноз переглядів веб-ресурсу на наступний рік за допомогою методу Хольта-Вінтерса, для наступних значень параметрів адаптації: $\alpha = \beta = \gamma = 0,17$;

2. Результати прогнозування показати на графіку;

3. Провести моделювання параметрів адаптації α , β та γ в діапазоні $(0; 0,3]$ з кроком 0,05 й проаналізувати відповідні зміни прогнозованої кількості переглядів веб-ресурсу.

Вихідні дані для виконання роботи наведено в таблиці 4.3.

Таблиця 4.3 – Вхідні дані для методу Хольта-Вінтерса

Місяць	2020 рік	2021 рік	2022 рік	2023 рік
Січень	22467	51802	73724	60496
Лютий	26217	51647	66612	61032
Березень	33363	60089	76228	70400
Квітень	29093	61394	74576	75041
Травень	34523	76296	86318	86291
Червень	30767	67486	65759	76115
Липень	14342	31862	31968	42879
Серпень	17553	32387	30086	36962
Вересень	29823	50688	51802	58776
Жовтень	45786	71383	68960	81204
Листопад	53309	81973	72853	98339
Грудень	57896	82804	75019	116717

Рекомендації до виконання. Метод Хольта-Вінтерса включає розклад вхідного часового ряду на три основні компоненти: рівень (Level), тренд (Trend) та сезонність (Seasonality). Їхній економічний зміст показаний на рис. 4.2.

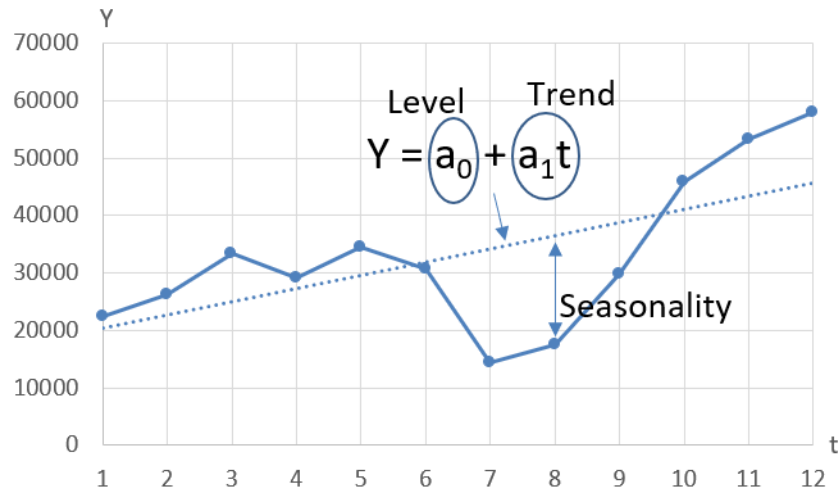


Рис. 4.2. Розклад вхідного часового ряду на основні компоненти за методом Хольта-Вінтерса

В нашому випадку, тривалість сезонного періоду становить $m = 12$ місяців. Тому, вхідні дані за перші 12 місяців, або за весь 2020 рік використовуються для визначення початкових значень складових часового ряду: Level, Trend та Seasonality, колонки E:G, рис. 4.3.

{=ЛИНЕЙН(D2:D13;B2:B13;1;1)}										
	A	B	C	D	E	F	G	H	I	J
	Рік	Місяць	Перегляди веб-ресурсу	Level	Trend	Seasonal				
1										
2	2020	1	Січень	22467	18000,45	2296,58	1,107		$\alpha = 0,17$	
3		2	Лютий	26217	18000,45	4593,17	1,160		$\beta = 0,17$	
4		3	Березень	33363	18000,45	6889,75	1,340		$\gamma = 0,17$	
5		4	Квітень	29093	18000,45	9186,34	1,070			
6		5	Травень	34523	18000,45	11482,92	1,171			
7		6	Червень	30767	18000,45	13779,50	0,968			
8		7	Липень	14342	18000,45	16076,09	0,421		a_1	a_0
9		8	Серпень	17553	18000,45	18372,67	0,483		2296,584	18000,45
10		9	Вересень	29823	18000,45	20669,26	0,771			
11		10	Жовтень	45786	18000,45	22965,84	1,118			
12		11	Листопад	53309	18000,45	25262,42	1,232			
13		12	Грудень	57896	18000,45	27559,01	1,271			

Рис. 4.3. Визначення вхідних значень складових часового ряду: Level, Trend та Seasonality

Для цього, на основі вхідних даних щодо переглядів веб-ресурсу протягом 2020 року, в комірках I9:J9 спочатку розраховуються параметри лінійного тренду a_0 та a_1 , з використанням вбудованої функції Microsoft Excel =ЛИНЕЙН() , або =LINEST() :

– в комірку I9 вводимо формулу:

$\text{=ЛИНЕЙН(D2:D13;B2:B13;1;1)}$, або $\text{=LINEST(D2:D13;B2:B13;1;1)}$.

– виділяємо діапазон комірок I9:J9, починаючи з комірки I9 й натискаємо клавіші <F2>, <CTRL> + <SHIFT> + <ENTER>.

В результаті, в комірці J9 буде міститись розраховане значення a_0 , а в комірці I9 – значення a_1 . рис. 4.3. Далі, на їх основі обчислюються початкові значення складових часового ряду: Level, Trend та Seasonality, комірки E2:G13, рис. 4.4:

	A	B	C	D	E	F	G
1	Рік	Місяць		Перегляди веб-ресурсу	Level	Trend	Seasonal
2	2020	1	Січень	22467	=J\$9	=I\$9*B2	=D2/(E2+F2)
3		2	Лютий	26217	=J\$9	=I\$9*B3	=D3/(E3+F3)
4		3	Березень	33363	=J\$9	=I\$9*B4	=D4/(E4+F4)
5		4	Квітень	29093	=J\$9	=I\$9*B5	=D5/(E5+F5)
6		5	Травень	34523	=J\$9	=I\$9*B6	=D6/(E6+F6)
7		6	Червень	30767	=J\$9	=I\$9*B7	=D7/(E7+F7)
8		7	Липень	14342	=J\$9	=I\$9*B8	=D8/(E8+F8)
9		8	Серпень	17553	=J\$9	=I\$9*B9	=D9/(E9+F9)
10		9	Вересень	29823	=J\$9	=I\$9*B10	=D10/(E10+F10)
11		10	Жовтень	45786	=J\$9	=I\$9*B11	=D11/(E11+F11)
12		11	Листопад	53309	=J\$9	=I\$9*B12	=D12/(E12+F12)
13		12	Грудень	57896	=J\$9	=I\$9*B13	=D13/(E13+F13)

Рис. 4.4. Формули для розрахунку вхідних значень складових часового ряду: Level, Trend та Seasonality

– початкове значення складової часового ряду Level на протязі всього 2020 р. буде дорівнювати знайденому параметру a_0 , рис. 4.2. Тому, в комірках E2:E13 слід ввести формулу $\text{=J$9}$, рис. 4.4;

– економічним змістом складової часового ряду Trend, згідно рис. 4.2, є добуток a_{1t} . Тому, в комірку F2 слід ввести формулу $=I\$9*B2$ й за правий нижній кут розтягнути її вниз до комірки F13;

– економічним змістом складової часового ряду Seasonality є індекс, що визначає співвідношення між фактичним значенням вхідного показника й його сукупною трендовою складовою (Level + Trend). Тому, до комірки G2 слід ввести формулу $=D2/(E2+F2)$ й за правий нижній кут розтягнути її вниз до комірки G13.

Обчислені таким чином за даними 2020 р. вхідні значення складових часового ряду, рис. 4.4, є основою для розрахунку цих компонентів протягом 2021-2023 років. Відповідні формули (4.11)-(4.13) мають наступний вигляд:

1. Рівень (Level):

$$L_t = \alpha \times \frac{Y_t}{S_{t-m}} + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad (4.11)$$

Де m – тривалість сезонного періоду ($m = 12$ місяців)

До комірки E14 слід ввести формулу $=J\$2*D14/G2+(1-J\$2)*(E13+F13)$ й розтягнути її вниз до грудня 2023 р., комірка E49.

2. Тренд (Trend):

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, \quad (4.12)$$

До комірки F14 слід ввести формулу $=J\$3*(E14-E13)+(1-J\$3)*F13$ й так само розтягнути її вниз до грудня 2023 р., комірка F49.

3. Сезонність (Seasonality):

$$S_t = \gamma \times \frac{Y_t}{L_t} + (1 - \gamma)S_{t-m}, \quad (4.13)$$

До комірки G14 слід ввести формулу $=J\$4*D14/E14+(1-J\$4)*G2$ й розтягнути її вниз до комірки G49.

У цих рівняннях використовуються наступні умовні позначення: Y_t – значення часового ряду в момент часу t ; L_t – рівень часового ряду в момент часу t ; T_t – тренд часового ряду в момент часу t ; S_t – індекс сезонності часового ряду в момент часу t ; α , β , γ – коефіцієнти згладжування, які визначають ваги спостережень, рівня, тренду та сезонності, відповідно.

Тоді, прогнозування цільового показника (Forecast) на h періодів наперед буде здійснюватися за формулою:

$$F_{t+h} = (L_t + hT_t) \times S_{t-m+h}, \quad (4.14)$$

Де h – порядковий номер періоду прогнозування (1, 2, ...)

Згідно формули (4.14), до комірки D50 слід ввести формулу $=(\$E\$49+B50*\$F\$49)*G38$ й розтягнути її вниз до комірки D61, рис. 4.5.

D50				fx		=(\$E\$49+B50*\$F\$49)*G38	
	A	B	C	D	E	F	G
38			Січень	60496	=SJS2*D38/G2	=SJS3*(E38-E3)	=SJS4*D38/E3
39			Лютий	61032	=SJS2*D39/G2	=SJS3*(E39-E3)	=SJS4*D39/E3
40			Березень	70400	=SJS2*D40/G2	=SJS3*(E40-E3)	=SJS4*D40/E4
41			Квітень	75041	=SJS2*D41/G2	=SJS3*(E41-E4)	=SJS4*D41/E4
42			Травень	86291	=SJS2*D42/G3	=SJS3*(E42-E4)	=SJS4*D42/E4
43			Червень	76115	=SJS2*D43/G3	=SJS3*(E43-E4)	=SJS4*D43/E4
44			Липень	42879	=SJS2*D44/G3	=SJS3*(E44-E4)	=SJS4*D44/E4
45			Серпень	36962	=SJS2*D45/G3	=SJS3*(E45-E4)	=SJS4*D45/E4
46			Вересень	58776	=SJS2*D46/G3	=SJS3*(E46-E4)	=SJS4*D46/E4
47			Жовтень	81204	=SJS2*D47/G3	=SJS3*(E47-E4)	=SJS4*D47/E4
48			Листопад	98339	=SJS2*D48/G3	=SJS3*(E48-E4)	=SJS4*D48/E4
49	2023		Грудень	116717	=SJS2*D49/G3	=SJS3*(E49-E4)	=SJS4*D49/E4
50		1	Січень	=(\$E\$49+B50*\$F\$49)*G38			
51		2	Лютий	=(\$E\$49+B51*\$F\$49)*G39			
52		3	Березень	=(\$E\$49+B52*\$F\$49)*G40			
53		4	Квітень	=(\$E\$49+B53*\$F\$49)*G41			
54		5	Травень	=(\$E\$49+B54*\$F\$49)*G42			
55		6	Червень	=(\$E\$49+B55*\$F\$49)*G43			
56		7	Липень	=(\$E\$49+B56*\$F\$49)*G44			
57		8	Серпень	=(\$E\$49+B57*\$F\$49)*G45			
58		9	Вересень	=(\$E\$49+B58*\$F\$49)*G46			
59		10	Жовтень	=(\$E\$49+B59*\$F\$49)*G47			
60		11	Листопад	=(\$E\$49+B60*\$F\$49)*G48			
61	2024	12	Грудень	=(\$E\$49+B61*\$F\$49)*G49			

Рис. 4.5. Розрахунок прогнозних значень (Forecast) переглядів веб-ресурсу

Тоді, динаміка вхідного часового ряду й його прогноз на наступний рік для $\alpha = \beta = \gamma = 0,17$, будуть мати вигляд, як показано на рис. 4.6.



Рис. 4.6. Прогноз переглядів веб-ресурсу на 2024 рік

Як видно з рис. 4.6, вхідні дані дійсно мають сезонний характер коливань, тривалістю в один рік. Далі, згідно пункту 3 завдання №4, необхідно самостійно провести моделювання параметрів адаптації α , β та γ в діапазоні $(0; 0,3]$ з кроком 0,05 й проаналізувати відповідні зміни в динаміці прогнозного показника.