

Jacek Szandula

OPTYMALIZACJA WAG W MODELU ŚREDNIEJ RUCHOMEJ WAŻONEJ

1. Szacowanie parametrów

Model średniej ruchomej ważonej zakłada, że prognozę na okres t określają odpowiednio ważone ostatnie k wartości szeregu oraz wahania przypadkowe:

$$y_t^* = \omega'_1 y_{t-k} + \omega'_2 y_{t-k+1} + \dots + \omega'_k y_{t-1} + \xi, \quad (1)$$

gdzie: y_t^* – wartość zmiennej Y w prognozowanym okresie t ,

y_{t-i} – wartość zmiennej w $(t-i)$ -tym okresie,

ω'_i – i -ty parametr modelu,

ξ – składnik losowy.

Zastępując w równaniu (1) wartości rzeczywiste teoretycznymi, otrzymuje się:

$$\hat{y}_t^* = w'_1 y_{t-k} + w'_2 y_{t-k+1} + \dots + w'_{k-1} y_{t-2} + w'_k y_{t-1}, \quad (2)$$

gdzie: $\hat{y}_t^*$ – prognoza zmiennej Y na okres t ,

w_i – i -ta waga,

$i = 1, 2, \dots, k$.

Na wagi nałożone są warunki:

$$w'_1 + w'_2 + \dots + w'_k = 1, \quad (3a)$$

$$0 \leq w'_1 \leq w'_2 \leq \dots \leq w'_k \leq 1. \quad (3b)$$

Warunek (3b) oznacza, że wszystkie wagi w modelu mogą być równe, co oznacza średnią ruchomą prostą. Poszczególne wagi można rozpisać w następujący sposób:

$$w'_1 = w_1, \quad (4a)$$

$$w'_2 = w_1 + w_2, \quad (4b)$$

$$\vdots$$

$$w'_k = w_1 + w_2 + \dots + w_k, \quad (4k)$$

przy czym:

$$w_i \geq 0, \quad (5)$$

$$kw_1 + (k-1)w_2 + \dots + w_k = 1. \quad (6)$$

Podstawiając równania (4a)-(4k) do równania (2), otrzymuje się:

$$\hat{y}_t^* = w_1 y_{t-k} + (w_1 + w_2) y_{t-k+1} + \dots + (w_1 + w_2 + \dots + w_k) y_{t-1}. \quad (7)$$

Wymnażając wyrażenia w nawiasach przez czynniki stojące przed nawiasami i dokonując odpowiednich przestawień, dochodzi się do równania:

$$\hat{y}_t^* = w_1 y_{t-k} + w_1 y_{t-k+1} + \dots + w_1 y_{t-1} + w_2 y_{t-k+1} + \dots + w_2 y_{t-1} + w_k y_{t-1}, \quad (8)$$

które jest tożsame z:

$$\hat{y}_t^* = w_1 \sum_{i=t-k}^{t-1} y_i + w_2 \sum_{i=t-k+1}^{t-1} y_i + \dots + w_k y_{t-1} \quad (9)$$

Aby uprościć zapis, sumy przy wagach można zastąpić w następujący sposób:

$$x_{1,t} = \sum_{i=t-k}^{t-1} y_i, \quad (10a)$$

$$x_{2,t} = \sum_{i=t-k+1}^{t-1} y_i, \quad (10b)$$

$$\vdots$$

$$x_{k,t} = y_{t-1}. \quad (10k)$$

W takim razie:

$$\hat{y}_t^* = w_1 x_{1,t} + w_2 x_{2,t} + \dots + w_k x_{k,t}. \quad (11)$$

Równanie (11) i ograniczenia (3a), (3b) można zapisać w postaci macierzowej:

$$\hat{y}_t^* = \mathbf{X}_t \mathbf{w}, \quad (12)$$

$$\mathbf{w} \geq \mathbf{0}, \quad (13a)$$

$$\mathbf{k}\mathbf{w} = 1, \quad (13b)$$

gdzie:

$$\mathbf{X}_t = [x_{1,t} \ x_{2,t} \ \dots \ x_{k,t}], \quad (14)$$

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix}, \quad (15)$$

$$\mathbf{k} = [k \ k-1 \ \dots \ 1]. \quad (16)$$

Przyjmując, że szereg czasowy składa się z n obserwacji, można wyznaczyć $n-k$ prognoz wygasłych. Odnosząc równanie (12) do całego wektora prognoz wygasłych, otrzymuje się:

$$\hat{\mathbf{y}}^* = \mathbf{X}\mathbf{w}, \quad (17)$$

gdzie:

$$\hat{\mathbf{y}}^* = \begin{bmatrix} y_{k+1}^* \\ y_{k+2}^* \\ \vdots \\ y_n^* \end{bmatrix}, \quad (18)$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{k+1} \\ \mathbf{X}_{k+2} \\ \vdots \\ \mathbf{X}_n \end{bmatrix}, \quad (19)$$

czyli:

$$\mathbf{X} = \begin{bmatrix} \sum_{i=1}^k y_i & \sum_{i=2}^k y_i & \dots & y_k \\ \sum_{i=2}^{k+1} y_i & \sum_{i=3}^{k+1} y_i & \dots & y_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=n-k}^{n-1} y_i & \sum_{i=n-k+1}^{n-1} y_i & \dots & y_{n-1} \end{bmatrix}. \quad (20)$$

Ponieważ prognozy wygasłe dotyczą okresów przeszłych, można je traktować jako wartości teoretyczne modelu. To znaczy z równania (17) można wyeliminować „gwiazdkę” oznaczającą prognozę:

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{w}. \quad (21)$$

Różnice między wartościami rzeczywistymi zmiennej a teoretycznymi (prognozami wygasłymi) stanowią reszty modelu:

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}, \quad (22)$$

czyli:

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\mathbf{w}. \quad (23)$$

Przyjmując średni kwadratowy błąd prognoz wygasłych dany wzorem:

$$s^{*2} = \frac{\mathbf{e}^T \mathbf{e}}{n - k} \quad (24)$$

za funkcję celu

$$f(\mathbf{w}) = s^{*2}, \quad (25)$$

$$f(\mathbf{w}) = \frac{1}{n - k} (\mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{X}\mathbf{w} + \mathbf{w}^T \mathbf{X}^T \mathbf{X}\mathbf{w}), \quad (26)$$

i szukając jej minimum

$$f(\mathbf{w}) \rightarrow \min, \quad (27)$$

można określić optymalny zestaw wag w_i .

Ponieważ na wektor $\mathbf{w}$ nałożone są ograniczenia dane wzorami (13a) i (13b), minimalizacja funkcji (26) wymaga zastosowania programowania kwadratowego, np. algorytmu Beale'a [1]. Znajac $\mathbf{w}$ i korzystając z równań (4a)-(4k) można ustalić ostateczne wartości wag modelu.

2. Dopuszczalność prognozy

Prognozę wyznaczoną na okres t można uznać za dopuszczalną, gdy względny błąd prognozy *ex ante* jest nie większy od pewnej wartości progowej:

$$\eta_t \leq \eta^*, \quad (28)$$

gdzie: η_t – względny błąd prognozy *ex ante* na okres t ,

η^* – wartość progowa błędu prognozy *ex ante*.

Wartość względnego błędu prognozy *ex ante* wyznacza się jako stosunek bezwzględnego błędu prognozy *ex ante* i prognozy zmiennej:

$$\eta_t = \frac{v_t}{\hat{y}_t^*} \times 100, \quad (29)$$

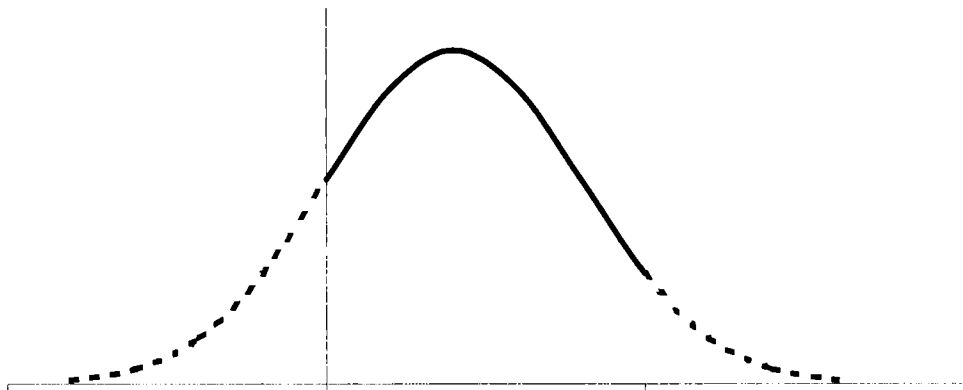
gdzie: v_t – bezwzględny błąd prognozy *ex ante*.

Bezwzględny błąd prognozy *ex ante* jest pierwiastkiem z wariancji błędu prognozy:

$$v_t = \sqrt{\text{var}(y_t^* - \hat{y}_t^*)}. \quad (30)$$

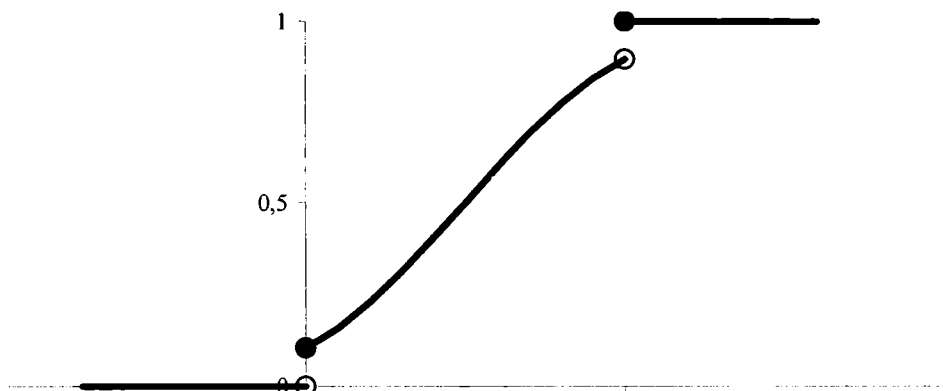
Problem wyznaczenia wariancji błędu prognozy nie nastęrczałby trudności, gdyby wektor estymatorów parametrów strukturalnych $\mathbf{w}$ (wektor wag) był nie-obciążony. Niestety, nieliniowe ograniczenia narzucone na wektor $\mathbf{w}$ powodują, że jego rozkład (zakładając, że składnik losowy ma rozkład normalny) ma charakter rozkładu dyskretno-ciągłego (zob. rys. 1 i 2). Oznacza to, że obszar na lewo od początku układu współrzędnych zastępuje masa prawdopodobieństwa zgromadzona w punkcie 0, a obszar na prawo od 1 reprezentuje masa prawdopodobieństwa w punkcie 1 [3, s. 360]. W większości przypadków rozkłady te nie będą symetryczne. Z tego względu estymator $\mathbf{w}$ parametrów $\boldsymbol{\omega}$ jest estymatorem obciążonym, tzn. $E(\mathbf{w}) \neq \boldsymbol{\omega}$.

J. Kmenta [2, s. 497] zauważa, że jeżeli prawdziwa wartość parametru ω_i znajduje się wewnątrz przedziału (0, 1), to ograniczenia nakładane na estymator w_i asymptotycznie będą stawać się nieistotne. Spostrzeżenie to ma zasadnicze znaczenie dla dalszego postępowania.



Rys. 1. Funkcja gęstości estymatorów przy warunkach pobocznych w postaci nierówności

Źródło: [3, s. 360].



Rys. 2. Dystrybuenta estymatorów przy warunkach pobocznych w postaci nierówności

Źródło: opracowanie własne.

Bez zmniejszenia ogólności rozważań wektor wag otrzymany za pomocą algorytmu Beale'a po przestawieniu jego elementów można podzielić na dwa, z których pierwszy zawiera wagi większe od zera, a drugi składa się z samych zer:

$$\mathbf{w}'' = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix}, \quad (31)$$

gdzie $\mathbf{w}''$ – pewna permutacja elementów wektora $\mathbf{w}$, przy czym:

$$\mathbf{w}_1 > \theta, \quad (32a)$$

$$\mathbf{w}_2 = \theta. \quad (32b)$$

Analogicznie postępuje się z wektorem ograniczeń:

$$\mathbf{k}'' = [\mathbf{k}_1 \quad \mathbf{k}_2], \quad (33)$$

gdzie $\mathbf{k}''$ – permutacja elementów wektora $\mathbf{k}$, przy czym:

$$\mathbf{k}_1 \mathbf{w}_1 = 1. \quad (34)$$

Liczba elementów wektorów $\mathbf{k}_1$ oraz $\mathbf{w}_1$: $m \in [1; k]$.

Dokonując przestawień odpowiednich kolumn wektora $\mathbf{X}$, równanie (21) można zapisać następująco:

$$\hat{\mathbf{y}} = [\mathbf{X}_1 \quad \mathbf{X}_2] \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix} = \mathbf{X}_1 \mathbf{w}_1. \quad (35)$$

Druga równość w powyższym równaniu wynika z tego, że powyższy iloczyn jest skalar. Ponieważ wszystkie wartości wektora $\mathbf{w}_1$ są większe od zera, ograniczenia w postaci nierówności (13a) stają się ograniczeniami nieaktywnymi. Oznacza to, że gdyby z góry znano niezerowe elementy wektora $\mathbf{w}$, nie trzeba by było korzystać z algorytmu Beale'a. Wartość wag wektora $\mathbf{w}_1$ można bowiem znaleźć, rozwiązując następujące zadanie:

$$f(\mathbf{w}_1) \rightarrow \min, \quad (36)$$

gdzie:

$$f(\mathbf{w}_1) = (\mathbf{y} - \mathbf{X}_1 \mathbf{w}_1)^T (\mathbf{y} - \mathbf{X}_1 \mathbf{w}_1), \quad (37)$$

przy czym:

$$\mathbf{k}_1 \mathbf{w}_1 = 1. \quad (38)$$

Niech:

$$L = (\mathbf{y} - \mathbf{X}_1 \mathbf{w}_1)^T (\mathbf{y} - \mathbf{X}_1 \mathbf{w}_1) + \lambda(1 - \mathbf{k}_1 \mathbf{w}_1), \quad (39)$$

będzie funkcją Lagrange'a. Wyznaczając pochodne cząstkowe:

$$\mathbf{L}'_{\mathbf{w}_1} = -2\mathbf{X}_1^T \mathbf{y} + 2\mathbf{X}_1^T \mathbf{X}_1 \mathbf{w}_1 - \lambda \mathbf{k}_1, \quad (40a)$$

$$L'_\lambda = 1 - \mathbf{k}_1 \mathbf{w}_1, \quad (40b)$$

a następnie przyrównując je do zera, otrzymuje się układ równań:

$$\mathbf{0} = -2\mathbf{X}_1^T \mathbf{y} + 2\mathbf{X}_1^T \mathbf{X}_1 \mathbf{w}_1 - \lambda \mathbf{k}_1, \quad (41a)$$

$$0 = 1 - \mathbf{k}_1 \mathbf{w}_1. \quad (41b)$$

Z równania (41a) wyznacza się $\mathbf{w}_1$:

$$\mathbf{w}_1 = (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{X}_1^T \mathbf{y} + 1/2 (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{k}_1^T \lambda. \quad (42)$$

Podstawiając wektor $\mathbf{w}$ do równania (41a), otrzymuje się:

$$\mathbf{k}_1 (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{X}_1^T \mathbf{y} + \frac{1}{2} \lambda \mathbf{k}_1 (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{k}_1^T - 1 = 0, \quad (43)$$

$$\lambda = 2 \left(\mathbf{k}_1 (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{k}_1^T \right)^{-1} - 2 \mathbf{k}_1 (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{X}_1^T \mathbf{y} \left(\mathbf{k}_1 (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{k}_1^T \right). \quad (44)$$

Następnie należy wstawić λ do równania (43):

$$\begin{aligned} \mathbf{w}_1 = & (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{X}_1^T \mathbf{y} + (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{k}_1^T \left(\mathbf{k}_1 (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{k}_1^T \right)^{-1} - \\ & - (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{k}_1^T \left(\mathbf{k}_1 (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{k}_1^T \right)^{-1} \mathbf{k}_1 (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{X}_1^T \mathbf{y}. \end{aligned} \quad (45)$$

lub

$$\mathbf{w}_1 = \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{y} + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} (1 - \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{y}), \quad (46)$$

gdzie:

$$\mathbf{S} = \mathbf{X}_1^T \mathbf{X}_1, \quad (47)$$

$$\mathbf{P} - \text{skalar}; \mathbf{P} = \mathbf{k}_1 (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{k}_1^T. \quad (48)$$

Na podstawie otrzymanego oszacowania oraz wykorzystując równości:

$$\mathbf{y} = \mathbf{X}_1 \omega_1 + \xi, \quad (49)$$

$$\mathbf{k}_1 \omega_1 = 1, \quad (50)$$

$$E(\xi) = \theta, \quad (51)$$

gdzie: $E(\xi)$ – wartość oczekiwana składnika losowego,
można wyznaczyć wartość oczekiwaną wektora wag:

$$\begin{aligned} E(\mathbf{w}_1) &= E \left[\mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{y} + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} (1 - \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{y}) \right] = \\ &= E \left[\mathbf{S}^{-1} \mathbf{X}_1^T (\mathbf{X}_1 \omega_1 + \xi) + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} (1 - \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T (\mathbf{X}_1 \omega_1 + \xi)) \right] = \\ &= E \left[\omega_1 + \mathbf{S}^{-1} \mathbf{X}_1^T \xi + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} (1 - \mathbf{k}_1 \omega_1) + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \xi \right] = \omega_1. \end{aligned} \quad (52)$$

Równanie (52) oznacza, że wektor $\mathbf{w}_1$ jest wektorem nieobciążonym.

Znając wartość oczekiwaną $\mathbf{w}_1$, z definicji można wyznaczyć jego macierz wariancji i kowariancji:

$$\text{var}(\mathbf{w}_1) = E(\mathbf{w}_1 - E(\mathbf{w}_1))^T E(\mathbf{w}_1 - E(\mathbf{w}_1)) = E(\mathbf{w}_1 - \omega_1)^T E(\mathbf{w}_1 - \omega_1). \quad (53)$$

Ponieważ:

$$\begin{aligned} \mathbf{w}_1 - \omega_1 &= \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{y} + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{y} - \omega_1 = \\ &= \mathbf{S}^{-1} \mathbf{X}_1^T (\mathbf{X}_1 \omega_1 + \xi) + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T (\mathbf{X}_1 \omega_1 + \xi) - \omega_1 = \\ &= \omega_1 + \mathbf{S}^{-1} \mathbf{X}_1^T \xi + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \omega_1 - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \xi - \omega_1 = \end{aligned}$$

$$\begin{aligned}
&= \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} (1 - \mathbf{k}_1 \omega_1) + (\mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) \xi = \\
&= (\mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) \xi.
\end{aligned} \tag{54}$$

Następnie, wykorzystując, że $\mathbf{S}^{-1}$ jest macierzą symetryczną, czyli jest równa własnej transpozycji:

$$(\mathbf{w}_1 - \omega_1)(\mathbf{w}_1 - \omega_1)^T = (\mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) \xi \xi^T (\mathbf{X}_1 \mathbf{S}^{-1} - \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1}) \tag{55}$$

oraz:

$$E(\xi \xi^T) = E((\xi - \theta)(\xi - \theta)^T) = E((\xi - E(\xi))(\xi - E(\xi))^T) = \text{var}(\xi) = \sigma^2 \mathbf{I}, \tag{56}$$

otrzymujemy:

$$\begin{aligned}
\text{var}(\mathbf{w}_1) &= E((\mathbf{w}_1 - \omega_1)(\mathbf{w}_1 - \omega_1)^T) = \\
&= E((\mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) \xi \xi^T (\mathbf{X}_1 \mathbf{S}^{-1} - \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1})) = \\
&= \sigma^2 (\mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} - \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} + \\
&\quad + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1}) = \\
&= \sigma^2 (\mathbf{S}^{-1} - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1}) = \\
&= \sigma^2 (\mathbf{S}^{-1} - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1}),
\end{aligned} \tag{57}$$

bo:

$$\mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 = \mathbf{S}^{-1} \mathbf{S} = \mathbf{I}, \tag{58}$$

$$\mathbf{k}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} = \mathbf{P} \mathbf{P}^{-1} = \mathbf{I}. \tag{59}$$

Analogicznie do (49) wektor prognoz można zapisać:

$$\mathbf{y}^* = \mathbf{X}_1^* \omega_1 + \xi^*, \tag{60}$$

gdzie: $\mathbf{y}^*$ – wektor prognoz,

$\mathbf{X}_1^*$ – macierz zmiennych objaśniających w okresach prognozowanych,

ξ^* – wektor składników losowych w okresach prognozowanych.

Błąd prognozy rozumiany jako różnica między wartościami rzeczywistymi a teoretycznymi wynosi więc:

$$\begin{aligned} \mathbf{y}^* - \mathbf{X}_1^* \mathbf{w}_1 &= \mathbf{X}_1^* \boldsymbol{\omega}_1 + \boldsymbol{\xi}^* - \mathbf{X}_1^* \mathbf{w}_1 = \mathbf{X}_1^* (\boldsymbol{\omega}_1 - \mathbf{w}_1) + \boldsymbol{\xi}^* = \\ &= (\mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^T) \boldsymbol{\xi} + \boldsymbol{\xi}^*. \end{aligned} \quad (61)$$

Trzeci znak powyższej równości wynika z (54).

Przy wyznaczaniu wariancji błędu predykcji użyteczne jest obliczenie wartości wyrażenia:

$$\begin{aligned} (\mathbf{y}^* - \mathbf{X}_1^* \mathbf{w}_1)(\mathbf{y}^* - \mathbf{X}_1^* \mathbf{w}_1)^T &= \left((\mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^T) \boldsymbol{\xi} + \boldsymbol{\xi}^* \right) \times \\ &\times \left(\boldsymbol{\xi}^T (\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) + \boldsymbol{\xi}^{*T} \right) = \\ &= (\mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^T) \boldsymbol{\xi} \boldsymbol{\xi}^T (\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \\ &- \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) + \boldsymbol{\xi}^* \boldsymbol{\xi}^{*T} + (\mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^T) \boldsymbol{\xi} \boldsymbol{\xi}^{*T} + \\ &+ \boldsymbol{\xi}^* \boldsymbol{\xi}^T (\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T). \end{aligned} \quad (62)$$

Z założenia, że składniki losowe z różnych okresów nie są ze sobą skorelowane:

$$E(\boldsymbol{\xi}^* \boldsymbol{\xi}^T) = E(\boldsymbol{\xi}^*) E(\boldsymbol{\xi}^T) = \boldsymbol{\theta} \times \boldsymbol{\theta}^T, \quad (63)$$

$$E(\boldsymbol{\xi} \boldsymbol{\xi}^{*T}) = E(\boldsymbol{\xi}) E(\boldsymbol{\xi}^{*T}) = \boldsymbol{\theta} \times \boldsymbol{\theta}^T, \quad (64)$$

otrzymuje się:

$$\begin{aligned} E \left[(\mathbf{y}^* - \mathbf{X}_1^* \mathbf{w}_1)(\mathbf{y}^* - \mathbf{X}_1^* \mathbf{w}_1)^T \right] &= \sigma^2 (\mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \\ &- \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T} - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T} + \\ &+ \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T} + \mathbf{I}) = \sigma^2 (\mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T} - \\ &- \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T} - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T} + \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^{*T} + \mathbf{I}) = \\ &= \sigma^2 (\mathbf{I} + \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^{*T} - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T}). \end{aligned} \quad (65)$$

Tak wyznaczona macierz wariancji i kowariancji zakłada posiadanie wiedzy o wartości parametru σ^2 . Niestety wartość ta najczęściej pozostaje nieznana, można ją jednak estymować. Niech:

$$\begin{aligned} \mathbf{e} = \mathbf{y} - \mathbf{X}_1 \mathbf{w}_1 &= \mathbf{X}_1 \boldsymbol{\omega}_1 + \boldsymbol{\xi} - \mathbf{X}_1 \mathbf{w}_1 = \mathbf{X}_1 (\boldsymbol{\omega}_1 - \mathbf{w}_1) + \boldsymbol{\xi} = \\ &= (\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{I}) \boldsymbol{\xi} = \mathbf{M} \boldsymbol{\xi}, \end{aligned} \quad (66)$$

gdzie: $\mathbf{e}$ – reszty modelu,

$\mathbf{M}$ – macierz idempotentna:

$$\mathbf{M}^T \mathbf{M} = \mathbf{M} = (\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{I}). \quad (67)$$

Wobec tego:

$$\mathbf{e}^T \mathbf{e} = \xi^T \mathbf{M}^T \mathbf{M} \xi = \xi^T \mathbf{M} \xi = \xi^T (\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{I}) \xi \quad (68)$$

Ponieważ $\mathbf{e}^T \mathbf{e}$ jest skalar, zatem uzyskany wynik równy jest swojemu śladowi:

$$\begin{aligned} \mathbf{e}^T \mathbf{e} &= \text{tr}(\xi^T (\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{I}) \xi) = \\ &= \xi^T \xi \text{tr}(\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{I}) = \\ &= \xi^T \xi \left[\text{tr}(\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) - \text{tr}(\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{I}) \right] = \\ &= \xi^T \xi \left[\text{tr}(\mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{I}) - \text{tr}(\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) - \text{tr} \mathbf{I} \right] = \xi^T \xi (n - m + 1). \end{aligned} \quad (69)$$

Czwarta równość powyższego równania wynika z tego, że:

$$\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA}), \quad (70)$$

o ile wymiary macierzy pozwalają na wykonanie obu iloczynów. W takim razie:

$$E(\mathbf{e}^T \mathbf{e}) = E(\xi^T \xi (n - m + 1)) = \sigma^2 (n - m + 1). \quad (71)$$

Nieobciążonym estymatorem wariancji składnika losowego jest więc:

$$S_e^2 = \frac{\mathbf{e}^T \mathbf{e}}{n - m + 1}. \quad (72)$$

Na podstawie (65) i (72) wariancja błędu prognozy wynosi:

$$v^2 = S_e^2 (\mathbf{I} + \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^{*T} - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T}). \quad (73)$$

Ostatecznie bezwzględny błąd prognozy *ex ante* można wyznaczyć ze wzoru:

$$v = S_e \sqrt{\mathbf{I} + \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^{*T} - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T}}. \quad (74)$$

3. Prognoza przedziałowa

Gdy wyznaczana jest jedna prognoza, równanie (60) można zapisać następująco:

$$\mathbf{y}^* = \mathbf{X}_1^* \omega_1 + \xi^*. \quad (75)$$

Jeżeli $\xi^T \xi^*$ ma $(n+1)$ -wymiarowy rozkład normalny, to:

$$\frac{\mathbf{X}_1^* \mathbf{w}_1 - \mathbf{y}^*}{\sigma \sqrt{1 + \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^{*T} - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T}}} \quad (76)$$

ma rozkład normalny standaryzowany, ponieważ wyrażenie w mianowniku jest pierwiastkiem z wariancji błędu predykcji [3, s. 152].

Z kolei:

$$(n-m+1)S_e^2 = \xi^T \mathbf{M} \xi, \quad (77)$$

gdzie macierz $\mathbf{M}$ jest macierzą idempotentną (zob. wzór (67)) rzędu $n-m+1$, ma rozkład χ^2 o $n-m+1$ stopniach swobody [3, s. 104-105, 147]. Zatem:

$$\frac{(n-m+1)S_e^2}{\sigma^2} \quad (78)$$

ma rozkład χ^2 o $n-m+1$ stopniach swobody. Ponadto zmienne dane wzorami (76) i (78) są niezależne. Aby to udowodnić, wystarczy wykazać, że iloczyn macierzy $\mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T$ z wyrażenia $\mathbf{w}_1 - \omega_1 = (\mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) \xi$ oraz macierzy $\mathbf{M}$ równa się zeru:

$$\begin{aligned} & (\mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) \mathbf{M} = \\ & = (\mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T) (\mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{I}) = \\ & = \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{S}^{-1} \mathbf{X}_1^T - \\ & \quad - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \cdot \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T \mathbf{X}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \\ & \quad - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T = \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{S}^{-1} \mathbf{X}_1^T - \\ & \quad - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T + \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T - \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^T = 0. \end{aligned} \quad (79)$$

Z powyższych rozważań oraz definicji rozkładu t -Studenta wynika, że:

$$\begin{aligned} & \frac{\mathbf{X}_1^* \mathbf{w}_1 - \mathbf{y}^*}{\sigma \sqrt{1 + \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^{*T} - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T}}} \sqrt{n-m+1} = \\ & \quad \sqrt{\frac{(n-m+1)S_e^2}{\sigma^2}} = \\ & = \frac{\mathbf{X}_1^* \mathbf{w}_1 - \mathbf{y}^*}{S_e \sqrt{1 + \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{X}_1^{*T} - \mathbf{X}_1^* \mathbf{S}^{-1} \mathbf{k}_1^T \mathbf{P}^{-1} \mathbf{k}_1 \mathbf{S}^{-1} \mathbf{X}_1^{*T}}} = \frac{\mathbf{X}_1^* \mathbf{w}_1 - \mathbf{y}^*}{v} \end{aligned} \quad (80)$$

ma rozkład t -Studenta o $n - m + 1$ stopniach swobody.

Na podstawie powyższego równania oraz wiedząc, że:

$$P\{-t_{\alpha} \leq t_{n-m+1} \leq t_{\alpha}\} = 1 - \alpha, \quad (81)$$

gdzie t_{α} – wartość statystyki odczytana z tablic t -Studenta z dwustronnym obszarem krytycznym dla poziomu istotności równego α i $n - m + 1$ stopni swobody, można wyznaczyć przedział ufności dla prognozy:

$$P\{\mathbf{X}_1^* \mathbf{w}_1 - t_{\alpha} v \leq \mathbf{y}^* \leq \mathbf{X}_1^* \mathbf{w}_1 + t_{\alpha} v\} = 1 - \alpha. \quad (82)$$

Literatura

- [1] Beale E.M.L., *On Minimizing a Convex Function Subject to Linear Inequalities*, „Journal of the Royal Statistical Society. Series B” 1955, 17, 173-184.
- [2] Kmenta J., *Elements of Econometrics*, Second Edition, The University of Michigan Press, 2000.
- [3] Theil H., *Zasady ekonometrii*, PWN, Warszawa 1979.

WEIGHTS OPTIMIZATION IN WEIGHTED MOVING AVERAGE MODEL

Summary

The paper presents a way to optimize weights in weighted moving average model. The minimum of the ex post forecasts mean-square error was chosen as a criterion of optimization. The author also presents the way of estimating ex ante forecasts errors. They are useful both for evaluating the admissibility of forecasts calculated with the use of estimated models and for calculating forecasts confidence intervals.

Dr Jacek Szandula jest adiunktem w Katedrze Prognoz i Analiz Gospodarczych Akademii Ekonomicznej we Wrocławiu.