УДК 004.852 DOi 10.30837/ bi.2021.1(96).11

О.Г. Руденко[1], О.О. Безсонов[2], Н.М. Сердюк[3], О.Г. Лебедєв[4], В.О. Лебедєв[5]

[1]Доктор технічних наук, завідувач кафедри комп'ютерних інтелектуальних технологій та систем, Харківський національний університет радіоелектроніки, Україна, oleh.rudenko@nure.ua, ORCID iD: 0000-0003-0859-2015

[2]Доктор технічних наук, професор кафедри комп'ютерних інтелектуальних технологій та систем, Харківський національний університет радіоелектроніки, Україна, oleksandr.bezsonov@nure.ua, ORCID iD: 0000-0001-6104-4275

[3] Кандидат технічних наук, доцент кафедри комп'ютерних інтелектуальних технологій та систем, Харківський національний університет радіоелектроніки, Україна, nataliya.serdyuk@nure.ua, ORCID iD: 0000-0002-0107-4365

[4] Кандидат технічних наук, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Україна, oleh.lebediev@nure.ua, ORCID iD: 0000-0001-5998-0136

[5] Аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Україна, valentyn.lebediev@nure.ua ORCID id: 0000-0002-0095-7481

БАГАТОКРОКОВИЙ МЕТОД НАВЧАННЯ АДАЛІНИ ЗА НАЯВНІСТЮ СТАЦІОНАРНИХ КОРЕЛЬОВАНИХ ЗАВАД

Розглянуто задачу побудови рекурентної форми багатокрокового алгоритму навчання АДАЛІНИ при наявності стаціонарних корельованих завад. Процес навчання АДАЛІНИ можна значно прискорити, якщо застосувати багатокроковий алгоритм, який використовує обмежене число вимірів, тобто має обмежену пам'ять, та побудований на базі алгоритму поточного регресійного аналізу. Розглянуто задачу побудови рекурентної форми алгоритму поточного регресійного аналізу, що дозволяє здійснювати оцінювання невідомих параметрів при наявності стаціонарних корельованих завад. Отримано основні співвідношення, які описують процеси накопичення нової і скидання застарілої інформації. Показано, що алгоритми, які розглядаються, використовують перетворення випадкового вектора з корельованими складовими в випадковий вектор з некорельованими складовими.

КОРЕЛЯЦІЯ, НАВЧАННЯ, РЕКУРЕНТНИЙ АЛГОРИТМ, КОРЕЛЯЦІЙНА МАТРИЦЯ, ЦЕНТРУВАННЯ, ЗАВАДА, ВКЛЮЧЕННЯ ІНФОРМАЦІЇ, СКИДАННЯ ІНФОРМАЦІЇ

Руденко О.Г., Бессонов А.А., Сердюк Н.Н., Лебедев О.Г., Лебедев В.О. Многошаговый метод обучения АДАЛИНЫ при наличии стационарных коррелированных помех. Рассмотрена задача построения рекуррентной формы многошагового алгоритма обучения АДАЛИН при наличии стационарных коррелированных помех. Процесс обучения АДАЛИН можно значительно ускорить, если применить многошаговый алгоритм, который использует ограниченное число измерений, то есть имеет ограниченную память и построен на базе алгоритма текущего регрессионного анализа. Рассмотрена задача построения рекуррентной формы алгоритма текущего регрессионного анализа, что позволяет осуществлять оценку неизвестных параметров при наличии стационарных коррелированных помех. Получены основные соотношения, описывающие процессы накопления новой и сброса устаревшей информации. Показано, что рассматриваемые алгоритмы используют преобразования случайного вектора с коррелированными составляющими в случайный вектор с некоррелированными составляющими.

КОРРЕЛЯЦИЯ, ОБУЧЕНИЕ, РЕКУРРЕНТНЫЙ АЛГОРИТМ, КОРРЕЛЯЦИОННАЯ МАТРИЦА, ЦЕНТРОВКА, ПОМЕХА, ВКЛЮЧЕНИЕ ИНФОРМАЦИИ, СБРОС ИНФОРМАЦИИ.

Rudenko O.G., Bezsonov O.O., Serdiuk N.M., Lebediev O.G., Lebediev V.O. Multi-step method of ADALINE learning in the presence of stationary correlated noises. The problem of constructing a recurrent form of a multistep learning algorithm ADALIN in the presence of stationary correlated noise is considered. The ADALIN learning process can be significantly accelerated if we apply a multi-step algorithm that uses a limited number of measurements, that is, it has a limited memory and is based on the current regression analysis algorithm. The problem of constructing a recurrent form of the current regression analysis algorithm is considered, which makes it possible to estimate unknown parameters in the presence of stationary correlated noise. The main relationships are obtained that describe the processes of accumulation of new and dumping of obsolete information. It is shown that the algorithms under consideration use transformations of a random vector with correlated components into a random vector with uncorrelated components.

CORRELATION, LEARNING, RECURRENT ALGORITHM, CORRELATION MATRIX, CENTERING, INTERFERENCE, INCLUDING INFORMATION, RESETTING INFORMATION.

Вступ

АДАЛІНА (адаптивний лінійний елемент) став першою лінійною нейронною мережею, запропонованою Уїдроу Б. і Хоффом М.Є., яка представляла альтернативу персептрону [1]. Згодом даний елемент і алгоритм його навчання знайшли досить широке застосування в задачах ідентифікації, управління, фільтрації і т.п. Алгоритм навчання Уїдроу – Хоффа є алгоритмом Качмажа рішення систем лінійних алгебраїчних рівнянь. Властивості даного алгоритму при вирішенні задачі ідентифікації досить повно описані в [2]. В роботі [3] регуляризований алгоритм Качмажа (Уїдроу-Хоффа) застосовувався для навчання АДАЛІНИ в завданні оцінювання нестаціонарних параметрів і отримано відповідні оцінки швидкості збіжності та точності.

Слід зазначити, що процес навчання АДАЛІНИ можна значно прискорити, якщо замість одно крокового алгоритму Уїдроу – Хоффа (Качмажа) застосувати багатокроковий алгоритм, який використовує обмежене число вимірів, тобто має обмежену пам'ять, та побудований на базі алгоритму поточного регресійного аналізу [4-6].

Зазвичай при побудові нейромережевих моделей і їх дослідженні передбачається, що присутні в вимірах завади не корельовані. Наявність же корельованих завад призводить до ускладнення процедури оцінювання. Тому часто корельованістю завад нехтують, отримуючи при цьому свідомо неоптимальні результати.

Рекурентне оцінювання при корельованих завадах розглядалося в [7-9]. В [7] запропонована рекурентна форма МНК, в [8] вивчалася процедура типу стохастичної апроксимації, коефіцієнти якої на кожному кроці обираються оптимальними в сенсі мінімізації суми апостеріорної дисперсії оцінок. В роботі [9] показано, що обидва ці способи мають багато спільного і було проведено їх експериментальне порівняння, яке показало, що в разі дуже сильної кореляції РМНК має більшу обчислювальну стійкість.

Метою даної роботи є отримання рекурентної форми багатокрокового алгоритму навчання за умов вистаціонарних корельованих завад.

1. Побудова рекурентної форми алгоритму поточного регресійного аналізу

Розглянемо задачу ідентифікації об'єкта, що описується рівнянням

$$Y_n = X_n c^* + \Xi_n, \tag{1}$$

де $Y_n = \left(y_1, y_2, \ldots y_n\right)^T$ – вектор вихідних сигналів;

$X_n^T = \left(x_1, x_2, \ldots, x_n\right)^T = \left(X_{n-1}^T x_n\right)$ – матриця вхідних сигналів;

$c^* = (c_1^*, c_2^*, \ldots c_N^*)^T$ – вектор параметрів, що оцінюються;

$\Xi_n = \left(\xi_1, \xi_2, \ldots, \xi_n\right)^T$ – вектор завад.

Коваріаційна матриця D_n порядку n завади ξ_{n+1} має вид

$$D_n = M\left\{\Xi_n \Xi_n^T\right\} = \begin{bmatrix} d_{1,1} & d_{1,2}\ldots & d_{1,n-1} & d_{1,n} \\ d_{2,1} & d_{2,2}\ldots & d_{2,n-1} & d_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ d_{n,1} & d_{n,2\ldots} & d_{n,n-1} & d_{n,n} \end{bmatrix} = \begin{bmatrix} D_{n-1} & d_{n-1} \\ d_{n-1}^T & d_{nn} \end{bmatrix}, \tag{2}$$

де $d_{ij} = M\left\{\xi_i \xi_j\right\}; \ldots$

$$d_{n-1}^T = \left(d_{n,1}, d_{n,2}, \ldots, d_{n,n-1}\right) = M\left\{\xi_n \Xi_{n-1}^T\right\}.$$

Як відомо [10], застосування оцінки

$$c_n = \left(X_n^T X_n\right)^{-1} X_n^T Y_n \tag{3}$$

до моделі з корельованими завадами дає оцінки, дисперсії яких будуть занижені.

Гаусівська-марківська оцінка (МНК), що отримується мінімізацією квадратичного функціоналу, має вигляд

$$c_n = \left(X_n^T D_n^{-1} X_n\right)^{-1} X_n^T D_n^{-1} Y_n \tag{4}$$

Алгоритм поточного регресійного аналізу, що має вигляд

$$c_{n|L} = (X_{n|L}^T X_{n|L})^{-1} X_{n|L}^T Y_{n|L}, \tag{5}$$

$$Y_{n|L} = \begin{pmatrix} Y_{n-1|L-1} \\ --- \\ y_n \end{pmatrix} = \begin{pmatrix} y_{n-L+1} \\ --- \\ Y_{n|L-1} \end{pmatrix}, \tag{6}$$

де $L \times 1$ – вектор;

$$X_{n|L} = \begin{pmatrix} X_{n-1|L-1} \\ --- \\ x_n^T \end{pmatrix} = \begin{pmatrix} x_{n-L+1}^T \\ --- \\ X_{n|L-1} \end{pmatrix}, \tag{7}$$

$(L \times 1) \times N$ – матриця;

був запропонований в роботі [4], а в [5] розглянута модифікація цього алгоритму, що використовує механізм забування минулого інформації (згладжування). Тут $L = const (L \geq N)$ - пам'ять алгоритму.

Особливістю алгоритмів з $L = const$ є те, що використовувані при побудові оцінок матриці і вектори спостережень на кожному кроці оцінювання формуються наступним чином: в них включається інформація про новоприбулі виміри і виключається інформація про найбільш старі. Залежно від того, як формуються ці матриці і вектори (додається спочатку нова інформація, а потім виключається застаріла, або ж спочатку виключається застаріла, а потім додається нова) можливі дві форми оцінки.

Зупинимося на цьому докладніше.

За аналогією з гаусівсько-марківською оцінкою (4) може бути запропонована оцінка

$$c_{n|L} = (X_{n|L}^T D_{n|L}^{-1} X_{n|L})^{-1} X_{n|L}^T D_{n|L}^{-1} Y_{n|L}, \tag{8}$$

де

$$D_{n|L}=\begin{bmatrix} d_{n-L+1,n-L+1} & d_{n-L,n-L+2}\dots & d_{n-L+1,n-1} & d_{n-L+1,n} \\ d_{n-L+2,n-L+1} & d_{n-L+2,n-L+2} & d_{n-L+2,n-1} & d_{n-L+2,n} \\ \dots & \dots & \dots & \dots \\ d_{n,n-L+1} & d_{n,n-L+2} & d_{n,n-1} & d_{n,n} \end{bmatrix}=$$

$$=\begin{bmatrix} D_{n-1,L-1+1} & \vdots & d_{n-1} \\ - & - & - \\ d_{n-1}^T & \vdots & d_{n,n} \end{bmatrix}, \qquad (9)$$

$$d_{n-1}^T=\left(d_{n,n-L+1},d_{n,n-L+2},\dots,d_{n,n-1}\right)=M\left\{\xi_n\Xi_{n-1|L-1}^T\right\}$$

У зв'язку з тим, що матриця $D_{n|L}$ має блокове представлення, то справедливо [11]

$$D_{n|L}^{-1}=$$

$$=\begin{bmatrix} D_{n-1,L-1}^{-1}+\dfrac{D_{n-1,L-1}^{-1}d_{n-1}d_{n-1}^T D_{n-1,L}^{-1}}{\alpha_n} & \vdots & -\dfrac{D_{n-1,L-1}^{-1}d_{n-1}}{\alpha_n} \\ \dots & \dots & \dots \\ -\dfrac{d_{n-1}^T D_{n-1,L-1}^{-1}}{\alpha_n} & \vdots & \dfrac{1}{\alpha_n} \end{bmatrix}, \qquad (10)$$

де $\alpha_n=d_{nn}-d_{n-1}^T D_{n-1,L-1}^{-1}d_{n-1}$.

Аналогічні співвідношення і використовуються в [7] для отримання рекурентної форми МНК.

Розглянемо випадок стаціонарних завад ξ.

Нормована коваріаційна матриця для стаціонарної завади з дисперсією σ_ξ^2 має вигляд

$$R_{n|L}=D_{n|L}/\sigma_\xi^2=\begin{bmatrix} 1 & \rho_1\dots & \rho_{n-L} & \rho_{n-L+1} \\ \rho_1 & 1\dots & \rho_{n-3} & \rho_{n-L} \\ \dots & \dots & \dots & \dots \\ \rho_{n-L+1} & \rho_{n-L}\dots & \rho_1 & 1 \end{bmatrix}, \quad (11)$$

де ρ_i — значення нормованої кореляційної функції процесу ξ.

С урахуванням цього оцінка (8) набирає вигляду

$$c_{n|L}=(X_{n|L}^T R_{n|L}^{-1}X_{n|L})^{-1}X_{n|L}^T R_{n|L}^{-1}Y_{n|L}, \qquad (12)$$

Введемо $(n-m+1)$ — вимірний вектор

$$r_{m,n}^T=\left(\rho_n,\rho_{n-1},\dots\rho_m\right),\quad (m\le n).$$

Тоді матриця $R_{n|L}$ може бути представленою в блочному виді

$$R_{n|L}=D_{n|L}/\sigma_\xi^2=\begin{bmatrix} R_{n-1|L-1} & r_{n-1,1} \\ r_{n-1,1}^T & 1 \end{bmatrix}, \qquad (13)$$

де $r_{n-1,1}^T=\left(\rho_{n-1},\rho_{n-2},\dots\rho_1\right)$, а зворотна матриця $R_{n|L}^{-1}$ буде обчислюватися так:

$$R_{n|L}^{-1}=$$

$$=\begin{bmatrix} R_{n-1|L-1}^{-1}+\dfrac{R_{n-1|L-1}^{-1}r_{n-1,1}r_{n-1,1}^T R_{n-1|L-1}^{-1}}{1-r_{n-1,1}^T R_{n-1|L-1}^{-1}r_{n-1,1}} & -\dfrac{R_{n-1|L-1}^{-1}r_{n-1,1}}{1-r_{n-1,1}^T R_{n-1|L-1}^{-1}r_{n-1,1}} \\ -\dfrac{r_{n-1,1}^T R_{n-1}^{-1}}{1-r_{n-1,1}^T R_{n-1|L-1}^{-1}r_{n-1,1}} & \dfrac{1}{1-r_{n-1,1}^T R_{n-1|L-1}^{-1}r_{n-1,1}} \end{bmatrix}. \quad (14)$$

2. Процедура додавання нової інформації

Припустимо, що на $(n-1)$-му такті отримана оцінка

$$c_{n-1|L}=\left(X_{n-1|L}^T R_{n-1|L}^{-1}X_{n-1|L}\right)^{-1}X_{n-1|L}^T R_{n-1|L}^{-1}Y_{n-1|L} \qquad (15)$$

Прихід нової інформації (додавання нового виміру) призводить до обчислення оцінки, яка за аналогією з (15) може бути записана наступним чином:

$$c_{n|L+1}=(X_{n|L+1}^T R_{n|L+1}^{-1}X_{n|L+1})^{-1}X_{n|L+1}^T R_{n|L+1}^{-1}Y_{n|L+1}, \qquad (16)$$

де

$$Y_{n|L+1}=\begin{pmatrix} Y_{n-1|L} \\ --- \\ y_n \end{pmatrix}=\begin{pmatrix} y_{n-L+1} \\ --- \\ Y_{n|L} \end{pmatrix} \text{ — вектор } (L+1)\times 1. \qquad (17)$$

Введемо позначення

$$K_{n|L+1}^{-1}=\left(X_{n|L+1}^T R_{n|L+1}^{-1}X_{n|L+1}\right);$$
$$K_{n-1|L}^{-1}=\left(X_{n-1|L}^T R_{n-1|L}^{-1}X_{n-1|L}\right); \qquad (19)$$
$$K_{n|L}^{-1}=\left(X_{n|L}^T R_{n|L}^{-1}X_{n|L}\right).$$

$$\alpha_{j,n}=\rho_j-r_{j+1,n+j-1}R_{n-1|L-1}^{-1}r_{1,n-1}\quad (j=0,1,\dots) \qquad (20)$$

та обчислимо $K_{n|L+1}^{-1}$

$$K_{n|L+1}^{-1}=X_{n-1|L}^T R_{n-1|L}^{-1}X_{n-1|L}+$$
$$+\frac{X_{n-1|L}^T R_{n-1|L}^{-1}r_{n-1,1}r_{n-1,1}^T R_{n-1|L}^{-1}X_{n-1|L}}{\alpha_{n,0}}-$$
$$-\frac{x_n r_{n-1,1}^T R_{n-1|L}^{-1}X_{n-1|L}}{\alpha_{n,0}}-\frac{X_{n-1|L}^T R_{n-1|L}^{-1}r_{n-1,1}x_n^T}{\alpha_{n,0}}+ \qquad (21)$$
$$+\frac{x_n x_n^T}{\alpha_{n,0}}=K_{n-1|L}^{-1}+x_n^* x_n^{*T},$$

де $x_n^*=\dfrac{x_n-X_{n-1|L}^T R_{n-1|L}^{-1}r_{n-1,1}}{\sqrt{\alpha_{n,0}}}$.

Аналогічно обчислимо

$$X_{n|L+1}^T R_{n|L+1}^{-1}Y_{n|L+1}=X_{n-1|L}^T R_{n-1|L}^{-1}Y_{n-1|L}+x_n^* y_n^*, \qquad (22)$$

де $y_n^*=\dfrac{y_n-Y_{n-1|L}R_{n-1|L}^{-1}r_{n-1,1}}{\sqrt{\alpha_{n,0}}}$.

Додаючи до обох частин (15) $x_n^* x_n^{*T}c_{n-1|L}$

$$K_{n-1|L}^{-1}c_{n-1|L}+x_n^* x_n^{*T}c_{n-1|L}=$$
$$=X_{n-1|L}^T R_{n-1|L}^{-1}Y_{n-1|L}+x_n^* x_n^{*T}c_{n-1|L}$$

і віднімаючи (15) з (16) (з урахуванням властивостей $K_{n-1|L}^{-1}$ и $X_{n|L+1}^T R_{n|L+1}^{-1}Y_{n|L+1}$), отримуємо

$$K_{n|L+1}^{-1}\left(c_{n|L+1}-c_{n-1|L}\right)=x_n^*\left(y_n^*-c_{n-1|L}^T x_n^*\right)$$

або

$$c_{n|L+1}=c_{n-1|L}+K_{n|L+1}x_n^*\left(y_n^*-c_{n-1|L}^T x_n^*\right), \qquad (23)$$

де

$$K_{n|L+1}=K_{n-1|L}-\frac{K_{n-1|L}x_n^* x_n^{*T}K_{n-1|L}}{1+x_n^{*T}K_{n-1|L}x_n^*}. \qquad (24)$$

3. Процедура відкиданні застарілої інформації

При відкиданні застарілої інформації, отриманої на $n-L+1$-у кроці, ми приходимо від оцінки $c_{n|L+1}$ до оцінки $c_{n|L}$. Для отримання відповідних правил корекції оцінки поступимо таким чином.

Скористаємося блоковим поданням коваріаційної матриці $R_{n|L+1}$

$$R_{n|L+1}=\begin{bmatrix}1&\rho_1&\rho_2&\rho_{n-L+1}\\ \rho_1&1&\rho_1&\rho_{n-L+1,n}\\ \dots&\dots&\dots&\dots\\ \rho_{n-L+1}&\rho_{n-L}&\rho_{n-L-1}&1\end{bmatrix}=\begin{bmatrix}1&r^T_{n-1,1}\\ r_{n-1,1}&R_{n|L}\end{bmatrix}, \quad (25)$$

де $r^T_{n-1,1}=\left(\rho_{n-L+1},\rho_{n-L},\dots,\rho_{n-1}\right)$, та представленням зворотної матриці $R^{-1}_{n|L+1}$ у вигляді

$$R^{-1}_{n|L+1}=\begin{bmatrix}\dfrac{1}{\alpha_{n-L+1,0}}&-\dfrac{r^T_{n-L+1}R^{-1}_{n,L}}{\alpha_{n-L+1,0}}\\ -\dfrac{R^{-1}_{n|L}r_{n-L+1}}{\alpha_{n-L+1,0}}&R^{-1}_{n,L}+\dfrac{R^{-1}_{n,L}r_{n-L+1}r^T_{n-L+1}R^{-1}_{n,L}}{\alpha_{n-L+1,0}}\end{bmatrix}, \quad (26)$$

де $\alpha_{n-L+1,0}=1-r^T_{n-L+1,1}R^{-1}_{n,L}r_{n-L+1,1}$.

В цьому випадку

$$\begin{aligned}K^{-1}_{n|L+1}&=\left(X^T_{n|L+1}R^{-1}_{n|L+1}X_{n|L+1}\right)=\\ &=\frac{x_{n-L+1}x^T_{n-L+1}}{\alpha_{n-L}}-+\frac{X^T_{n|L}R^{-1}_{n|L}d_{n-L+1}d^T_{n-L+1}R^{-1}_{n|L}X_{n|L}}{\alpha_{n-L}}-\\ &-\frac{x_{n-L+1}d^T_{n-L+1}R^{-1}_{n|L}X_{n|L}}{\alpha_{n-L+1}}-\frac{X^T_{n|L}R^{-1}_{n|L}d_{n-L+1}x^T_{n-L+1}}{\alpha_{n-L+1}}+\\ &+X^T_{n|L}R^{-1}_{n|L}X_{n|L}=K^{-1}_{n|L}+x^*_{n-L+1}x^{*T}_{n-L+1},\end{aligned} \quad (27)$$

де $x^*_{n-L+1}=\dfrac{x_{n-L+1}-X^T_{n|L}D^{-1}_{n|L}r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}$.

Аналогічно маємо

$$X^T_{n|L+1}R^{-1}_{n|L+1}Y_{n|L+1}=X^T_{n|L}R^{-1}_{n|L}Y_{n|L}+x^*_{n-L+1}y^*_{n-L+1}, \quad (28)$$

де $y^*_{n-L+1}=\dfrac{y_{n-L+1}-Y_{n|L}D^{-1}_{n|L}r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}$.

Віднімемо з обох частин (16) $x^*_{n-L+1}x^{*T}_{n-L+1}c_{n|L+1}$. Тоді

$$\begin{aligned}&\left(\left(X^T_{n|L+1}R^{-1}_{n|L+1}Y_{n|L+1}\right)^{-1}-x^*_{n-L+1}x^{*T}_{n-L+1}\right)c_{n|L+1}=\\ &=X^T_{n|L+1}R^{-1}_{n|L+1}Y_{n|L+1}-x^*_{n-L+1}x^{*T}_{n-L+1}c_{n|L+1}.\end{aligned} \quad (29)$$

Враховуючи, що

$$\left(X^T_{n|L}R^{-1}_{n|L}X_{n|L}\right)c_{n|L}=X^T_{n|L}R^{-1}_{n|L}Y_{n|L}, \quad (30)$$

віднімемо з (30) співвідношення (16) (с урахуванням виразів для $K^{-1}_{n|L}$ та $X^T_{n|L}R^{-1}_{n|L}Y_{n|L}$)

$$K^{-1}_{n|L}\left(c_{n|L}-c_{n|L+1}\right)=x^*_{n-L+1}x^{*T}_{n-L+1}c_{n|L+1}-x^*_{n-L+1}y^*_{n-L+1},$$

звідки

$$c_{n|L}=c_{n|L+1}-K_{n|L}x^*_{n-L+1}\left(y^*_{n-L+1}-c^T_{n|L+1}x^*_{n-L+1}\right), \quad (31)$$

але $K^{-1}_{n|L}=K^{-1}_{n|L+1}-x^*_{n-L+1}x^{*T}_{n-L+1}$.

Тому

$$K_{n|L}=K_{n|L+1}+\frac{K_{n|L+1}x^*_{n-L+1}x^{*T}_{n-L+1}K_{n|L+1}}{1-x^{*T}_{n-L+1}K_{n|L+1}x^*_{n-L+1}}. \quad (32)$$

Таким чином, алгоритм буде мати вигляд (перші два співвідношення описують включення нової інформації, а наступні два - відкидання застарілої)

$$c_{n|L+1}=c_{n-1|L}+K_{n|L+1}x^*_n\left(y^*_n-c^T_{n-1|L}x^*_n\right), \quad (33)$$

$$K_{n|L+1}=K_{n-1|L}-\frac{K_{n-1|L}x^*_nx^{*T}_nK_{n-1|L}}{1+x^{*T}_nK_{n-1|L}x^*_n}, \quad (34)$$

$$c_{n|L}=c_{n|L+1}-K_{n|L}x^*_{n-L+1}\left(y^*_{n-L+1}-c^T_{n|L+1}x^*_{n-L+1}\right), \quad (35)$$

$$K_{n|L+1}=K_{n|L}+\frac{K_{n|L+1}x^*_{n-L+1}x^{*T}_{n-L+1}K_{n|L+1}}{1-x^{*T}_{n-L+1}K_{n|L+1}x^*_{n-L+1}}. \quad (36)$$

Якщо ж спочатку відкидається застаріла інформація, а потім включається нова, то алгоритм набуває вигляду

$$\begin{aligned}c_{n-1|L-1}&=c_{n-1|L}-\\ &-K_{n-1|L-1}x^*_{n-L+1}\left(y^*_{n-L+1}-c^T_{n-1|L}x^*_{n-L+1}\right);\end{aligned} \quad (37)$$

$$K_{n-1|L-1}=K_{n-1|L}+\frac{K_{n-1|L}x^*_{n-L+1}x^{*T}_{n-L+1}K_{n-1|L}}{1-x^{*T}_{n-L+1}K_{n-1|L}x^*_{n-L+1}}; \quad (38)$$

$$c_{n|L}=c_{n-1|L-1}+K_{n|L}x^*_n\left(y^*_n-c^T_{n-1|L-1}x^*_n\right); \quad (39)$$

$$K_{n|L}=K_{n-1|L-1}-\frac{K_{n-1|L-1}x^*_nx^{*T}_nK_{n-1|L-1}}{1+x^{*T}_nK_{n-1|L-1}x^*_n}; \quad (40)$$

Тут

$$x^*_{n-L+1}=\frac{x_{n-L+1}-X^T_{n|L}D^{-1}_{n|L}r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}; \quad (41)$$

$$x^*_n=\frac{x_n-X^T_{n-1|L}D^{-1}_{n-1|L}r_{n-1,1}}{\sqrt{\alpha_{n,0}}}; \quad (42)$$

$$y^*_{n-L+1}=\frac{y_{n-L+1}-Y_{n|L}D^{-1}_{n|L}r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}; \quad (43)$$

$$y^*_n=\frac{y_n-Y_{n-1|L}D^{-1}_{n-1|L}d_{n-1}}{\sqrt{\alpha_n}}. \quad (44)$$

4. Особливості оцінювання для випадку стаціонарних завад

Розглянемо деякі особливості оцінювання для випадку стаціонарних завад.

З урахуванням введених позначень перетворений вихідний сигнал y^*_n може бути записаний так

$$y^*_n=y_n-r^T_{1,n-1}R^{-1}_{n-1|L-1}Y_{n-1|L-1}. \quad (45)$$

Визначимо добуток $r^T_{1,n-1}R^{-1}_{n-1|L-1}Y_{n-1|L-1}$, користуючись блоковим поданням векторів і матриць, що входять до нього.

У зв'язку з тим, що

$$r^T_{n-1,1}=\left(\rho_{n-1}\vdots r^T_{n-2,1}\right),\qquad Y^T_{n-1}=\left(Y^T_{n-2|L-1}\vdots y_n\right),$$

$$R_{n-1|L-1}=\begin{bmatrix}R_{n-2|L-2}&r^T_{n-2,1}\\ r_{n-2,1}&1\end{bmatrix},$$

а зворотна матриця обчислюється аналогічно (14)

$$R^{-1}_{n-1|L-1} = \tag{46}$$

$$= \begin{bmatrix} R^{-1}_{n-2|L-2} + \dfrac{R^{-1}_{n-2|L-2} r_{n-2,1} r^T_{n-2,1} R^{-1}_{n-2|L-2}}{1 - r^T_{n-2,1} R^{-1}_{n-2|L-2} r_{n-2,1}} & -\dfrac{R^{-1}_{n-2|L-2} r_{n-2,1}}{1 - r^T_{n-2,1} R^{-1}_{n-2|L-2} r_{n-2,1}} \\ -\dfrac{r^T_{n-1,1} R^{-1}_{n-1}}{1 - r^T_{n-2,1} R^{-1}_{n-2|L-2} r_{n-2,1}} & \dfrac{1}{1 - r^T_{n-2,1} R^{-1}_{n-2|L-2} r_{n-2,1}} \end{bmatrix},$$

то

$$r^T_{1,n-1} R^{-1}_{n-1|L-1} Y_{n-1|L-1} = \left(\rho_{n-1} \vdots r^T_{n-2,1}\right) \times$$

$$\times \begin{bmatrix} R^{-1}_{n-2|L-2} + \dfrac{R^{-1}_{n-2|L-2} r_{n-2,1} r^T_{n-2,1} R^{-1}_{n-2|L-2}}{\alpha_{n-1,0}} & -\dfrac{R^{-1}_{n-2|L-2} r_{n-2,1}}{\alpha_{n-1,0}} \\ -\dfrac{r^T_{n-2,1} R^{-1}_{n-2|L-2}}{\alpha_{n-1,0}} & \dfrac{1}{\alpha_{n-1,0}} \end{bmatrix} \times$$

$$\times \begin{pmatrix} Y_{n-2|L-2} \\ \dots \\ y_{n-1} \end{pmatrix} = r^T_{n-2,1} R^{-1}_{n-2|L-2} Y_{n-2|L-2} + \tag{47}$$

$$+ \left(\frac{\rho_{n-1} - r^T_{n-2,1} R^{-1}_{n-2|L-2} r_{n-2,1}}{\alpha_{n-1,0}} \right) \times \left(y_{n-1} - r^T_{n-2,1} R^{-1}_{n-2|L-2} Y_{n-2|L-2} \right),$$

де $\alpha_{n-1,0} = 1 - r^T_{n-2,1} R^{-1}_{n-2|L-2} r_{n-2,1}$. Але

$$\rho_{n-1} - r^T_{n-2,1} R^{-1}_{n-2|L-2} r_{n-2,1} = \alpha_{n-1,1},$$

$$y_{n-1} - Y^T_{n-2|L-2} R^{-1}_{n-2|L-2} r_{n-2,1} = y^*_{n-1}.$$

Тому остаточно маємо

$$Y^T_{n-1,1} R^{-1}_{n-1|L-1} r^T_{n-1,1} = Y^T_{n-2,L-2} R^{-1}_{n-2|L-2} r^T_{n-2,1} + \frac{\alpha_{n-1,1}}{\alpha_{n-1,0}} y^*_{n-1}. \tag{48}$$

Після підстановки (48) в (46) отримуємо

$$y^*_n = y_n - \frac{\alpha_{n-1,1}}{\alpha_{n-1,0}} y^*_{n-1} - Y^T_{n-2,2} R^{-1}_{n-2|L-2} r_{n-2,1}. \tag{49}$$

Добуток $Y^T_{n-2|L-2} R^{-1}_{n-2|L-2} r_{n-2,1}$, в свою чергу, обчислюється рекурентно за формулою (47). Таким чином, змінна y^*_n обчислюється наступним чином:

$$y^*_n = y_n - \sum_{k=1}^{n-1} \frac{\alpha_{n-k,k}}{\alpha_{n-k,0}} y^*_{n-k}. \tag{50}$$

Коефіцієнти $\alpha_{n-k,k}$, що входять в даний вираз, допускають рекурентне обчислення. Для цього визначимо величини $r^T_{i,j} R^{-1}_l r_{k,m}$.

У зв'язку з тим, що згідно з прийнятими позначеннями

$$r^T_{i,j} = \left(\rho_i \vdots r^T_{i-1,j}\right),\ r^T_{k,m} = \left(\rho_k \vdots r^T_{k-1,m}\right),$$

а матриця R^{-1}_l обчислюється відповідно до (43), то

$$r^T_{i,j} R^{-1}_l r_{k,m} = \left(\rho_i \vdots r^T_{i-1,j}\right) \times$$

$$\times \begin{bmatrix} \dfrac{1}{\alpha_{l-1,0}} & -\dfrac{r^T_{l-1,1} R^{-1}_{l-1}}{\alpha_{n-1,0}} \\ -\dfrac{R^{-1}_{l-1} r_{l-1,1}}{\alpha_{l-1,0}} & R^{-1}_{l-1} + \dfrac{R^{-1}_{l-1} r_{l-1,1} r^T_{l-1,1} R^{-1}_{l-1}}{\alpha_{l-1,0}} \end{bmatrix} \begin{pmatrix} \rho_k \\ \dots \\ r_{k-1,m} \end{pmatrix} = \tag{50}$$

$$= r^T_{i-1,j} R^{-1}_{l-1} r_{k-1,m} + \frac{\left(\rho_i - r^T_{i-1,j} R^{-1}_{l-1} r_{l-1,1}\right)\left(\rho_k - r^T_{k-1,m} R^{-1}_{l-1} r_{l-1,1}\right)}{\alpha_{l-1,0}}.$$

Тут $\alpha_{l-1,0} = 1 - r^T_{l-1,0} R^{-1}_{l-1} r_{l-1,0}$.

Це дозволяє записати для любого моменту часу

$$r^T_{n+k-1,k} R^{-1}_{n-1|L-1} r_{n+m-1,m} =$$

$$= r^T_{n+k-2,k+1} R^{-1}_{n-2|L-2} r_{n+m-2,m+1} -$$

$$- \frac{\left(\rho_{n+k-2} - r^T_{n+k-2,k} R^{-1}_{n-2|L-2} r_{n-1,1}\right)\left(\rho_{n+m-2} - r^T_{n+m-2,m} R^{-1}_{n-2|L-2} r_{n-1,1}\right)}{\alpha_{n-2,0}} = \tag{51}$$

$$= r^T_{n+k-2,k} R^{-1}_{n-2|L-2} r_{n+m-2,m} - \frac{\alpha_{n-1,k} \alpha_{n-1,m}}{\alpha_{n-2,0}}.$$

Підставляючи (51) в (47) та ітеруючи, отримуємо

$$\alpha_{n,j} = \rho_j - \sum_{k=1}^{n-1} \frac{\alpha_{n-k,k} \alpha_{n-k,j+k}}{\alpha_{n-k,0}}. \tag{52}$$

Таким чином, змінна y^*_n, використовувана в алгоритмі, може бути визначена за формулами (50), (52) без обчислення зворотної матриці R^{-1}_{n-1}.

Зупинимося детальніше на властивостях перетвореного процесу y^*_n. Розглянемо центровану складову процесу y^*_n.

$$\bar{y}^*_n = y^*_n - M\left\{y^*_n\right\} = \xi_n - \Xi^T_{n-1|L-1} R^{-1}_{n-1|L-1} r_{n-1,1}. \tag{53}$$

Обчислимо $M\left\{\left(\bar{y}^*_n\right)^2\right\}$:

$$M\left\{\left(\bar{y}^*_n\right)^2\right\} = M\left\{\left(\xi_n - \Xi^T_{n-1|L-1} R^{-1}_{n-1|L-1} r_{n-1,1}\right)^2\right\} =$$

$$= M\left\{\xi^2_n\right\} - 2M\left\{\xi_n \Xi^T_{n-1|L-1} R^{-1}_{n-1|L-1} r_{n-1,1}\right\} + \tag{54}$$

$$+ M\left\{r^T_{n-1,1} R^{-1}_{n-1|L-1} \Xi_{n-1|L-1} \Xi^T_{n-1|L-1} R^{-1}_{n-1|L-1} r_{n-1,1}\right\}.$$

Внаслідок симетричності матриці R^{-1}_{n-1}

$$r^T_{n-1,1} R^{-1}_{n-1|L-1} \Xi_{n-1|L-1} = \Xi^T_{n-1|L-1} R^{-1}_{n-1|L-1} r_{n-1,1}.$$

Так як $M\left\{\Xi_{n-1}\xi_n\right\} = r_{n-1,1}, M\left\{\Xi_{n-1}\Xi^T_{n-1}\right\} = R^{-1}_{n-1}$, то після взяття операції математичного сподівання отримуємо

$$M\left\{\left(\bar{y}^*_n\right)^2\right\} = 1 - r^T_{n-1,1} R^{-1}_{n-1|L-1} r_{n-1,1} = \alpha_{n-1,0}, \tag{55}$$

тобто параметр $\alpha_{n-1,0}$ визначає дисперсію процесу y^*_n.

Аналогічно визначимо значення взаїмної кореляційної функції між завадою ξ_{n-j} $(j = 1, 2, ..., n-1)$ та центрованою складовою y^*_n

$$M\left\{\xi_{n+j} \bar{y}^*_n\right\} =$$

$$= M\left\{\xi_{n+j}\xi_n - \Xi^T_{n-1|L-1} R^{-1}_{n-1|L-1} r_{n-1,1} \xi_{n+j}\right\} =$$

$$= \rho_j - r^T_{n-1,1} R^{-1}_{n-1|L-1} r_{n-j+1,j+1}.$$

Звідси слід, що коефіцієнти $\alpha_{n,j}$ являють собою значення взаємної кореляційної функції центрованої складової сигналу y^*_n і завади ξ_n.

Обчислимо $M\left\{\bar{y}^*_n \bar{y}^*_{n-1}\right\}$.

Скориставшись виразом (54) і тим, що $\Xi^T_{n|L} = \left(\Xi^T_{n-1|L-1} \vdots \xi_n\right)$, отримуємо

$$M\left\{\bar{y}^*_n \bar{y}^*_{n-1}\right\} = M\left\{\left(\xi_n - \Xi^T_{n-1|L-1} R^{-1}_{n-1|L-1} r_{n-1,1}\right) \bar{y}^*_{n-1}\right\}.$$

Визначимо рекурентну формулу для перерахунку вектора $\Xi^T_{n-1|L-1} R^{-1}_{n-1|L-1} r_{n-1,1}$, що входить в даний вираз, використовуючи блочне представлення співмножників, які входять до нього. В цьому випадку

$$\Xi_{n-1|L-1}^{T} R_{n-1|L-1}^{-1} r_{n-1,1} = \left(\Xi_{n-1|L-1}^{T} \vdots \xi_n\right) R_{n-1|L-1}^{-1} \begin{pmatrix} \rho_{n-1} \\ \dots \\ r_{n-2,1} \end{pmatrix},$$

де $R_{n-1|L-1}^{-1}$ визначається формулою (46). Після нескладних перетворень маємо

$$\Xi_{n-1|L-1}^{T} R_{n-1|L-1}^{-1} r_{n-1,1} = \Xi_{n-2|L-2}^{T} R_{n-2|L-2}^{-1} r_{n-2,1} +$$

$$+\left(\xi_{n-1} - \Xi_{n-2|L-2}^{T} R_{n-2|L-2}^{-1} r_{n-2,1}\right) \frac{\alpha_{n-1,1}}{\alpha_{n-1,0}}.$$

В наших позначеннях

$\xi_{n-1} - \Xi_{n-2|L-2}^{T} R_{n-2|L-2}^{-1} r_{n-2,1} = \bar{y}_{n-1}^{*}$. Таким чином,

$$\Xi_{n-1|L-1}^{T} R_{n-1|L-1}^{-1} r_{n-1,1} = \Xi_{n-2|L-2}^{T} R_{n-2|L-2}^{-1} r_{n-2,1} + \bar{y}_{n-1}^{*} \frac{\alpha_{n-2,1}}{\alpha_{n-2,0}}. \quad (56)$$

Тоді

$$M\left\{\bar{y}_n^{*} \bar{y}_{n-1}^{*}\right\} =$$

$$= M\left\{\left(\xi_{n-1} - \Xi_{n-2|L-2}^{T} R_{n-2|L-2}^{-1} r_{n-2,1} - \bar{y}_{n-1}^{*} \frac{\alpha_{n-1,1}}{\alpha_{n-1,0}}\right) \bar{y}_{n-1}^{*}\right\} =$$

$$= M\left\{\left(\xi_n \bar{y}_{n-1}^{*} - \Xi_{n-2|L-2}^{T} R_{n-2|L-2}^{-1} r_{n-2,1} \bar{y}_{n-1}^{*} - \left(\bar{y}_{n-1}^{*}\right)^2 \frac{\alpha_{n-1,1}}{\alpha_{n-1,0}}\right)\right\} =$$

$$= -M\left\{\xi_n \Xi_{n-2|L-2}\right\} R_{n-2|L-2}^{-1} r_{n-2,1} + \quad (57)$$

$$+ r_{n-2,1}^{T} R_{n-2|L-2}^{-1} M\left\{\Xi_{n-2|L-2} \Xi_{n-2|L-2}^{T}\right\} R_{n-2|L-2}^{-1} r_{n-2,1} =$$

$$= -r_{n-2,1}^{T} R_{n-2|L-2}^{-1} r_{n-2,1} + r_{n-2,1}^{T} R_{n-2|L-2}^{-1} r_{n-2,1} = 0.$$

Звідси впливає, що

$$M\left\{\left(\Xi_{n-2|L-2}^{T} R_{n-2|L-2}^{-1} r_{n-2,1} \bar{y}_{n-1}^{*}\right)\right\} = 0,$$

тобто

$$M\left\{\Xi_{n-2|L-2} \bar{y}_{n-1}^{*}\right\} = 0. \quad (58)$$

Враховуючи (46), (48), (58), маємо

$$M\left\{\bar{y}_n^{*} \bar{y}_{n-2}^{*}\right\} =$$

$$= M\left\{\left(\xi_n - \Xi_{n-2|L-2}^{T} R_{n-2|L-2}^{-1} r_{n-2,1} - \bar{y}_{n-1}^{*} \frac{\alpha_{n-2,1}}{\alpha_{n-1,0}}\right) \bar{y}_{n-2}^{*}\right\} =$$

$$= M\left\{\left(\xi_n \bar{y}_{n-1}^{*} - \Xi_{n-3|L-3}^{T} R_{n-3|L-3}^{-1} r_{n-3,1} - \bar{y}_{n-2}^{*} \frac{\alpha_{n-2,1}}{\alpha_{n-2,0}} - \bar{y}_{n-1}^{*} \frac{\alpha_{n-1,1}}{\alpha_{n-1,0}}\right) \bar{y}_{n-2}^{*}\right\}.$$

Після взяття операції математичного сподівання так, як було зроблено вище, отримуємо

$$M\left\{\bar{y}_{n-1}^{*} \bar{y}_{n-2}^{*}\right\} = 0,$$

$$M\left\{\xi_n \bar{y}_{n-2}^{*}\right\} = \alpha_{n-2,1},$$

$$M\left\{\left(\bar{y}_{n-2}^{*}\right)^2\right\} = \alpha_{n-2,0}.$$

Внаслідок (57) $r_{n-3,1}^{T} R_{n-3}^{-1}$ $M\left\{\Xi_{n-3} \bar{y}_{n-2}^{*}\right\} = 0$ і остаточно маємо

$$M\left\{\bar{y}_n^{*} \bar{y}_{n-2}^{*}\right\} = 0.$$

Обчислюючи $M\left\{\bar{y}_n^{*} \bar{y}_{n-3}^{*}\right\}$, $M\left\{\bar{y}_n^{*} \bar{y}_{n-4}^{*}\right\}$ і т.д., переконуємося, що всі ці математичні сподівання добутків дорівнюють нулю. Отже, центрована складова перетвореного процесу, яка визначається співвідношенням (57), являє собою дискретний білий шум, тобто співвідношення (44) описує перетворення випадкового вектора з корельованими складовими в випадковий вектор з некорельованими складовими.

Висновки

Таким чином, розглянуто задачу побудови рекурентної форми алгоритму поточного регресійного аналізу, що дозволяє здійснювати оцінювання невідомих параметрів при наявності стаціонарних корельованих завад.

Отримано основні співвідношення, які описують процеси накопичення нової і скидання застарілої інформації. Показано, що алгоритми, які розглядаються, використовують перетворення випадкового вектора з корельованими складовими в випадковий вектор з некорельованими складовими.

Конфлікт інтересів

Автори не заявляють конфлікту інтересів.

References

[1] Widrow B. Adaptive switching circuits, / B. Widrow, M. Hoff // IRE WESCON Convention Record. Part 4. New York: Institute of Radio Engineers, 1960. – P.96–104.

[2] Либероль Б.Д. Исследование сходимости одношаговых адаптивных алгоритмов идентификации. / Б.Д. Либероль, О.Г. Руденко, А.А. Бессонов // Проблемы управления и информатики. – 2018. –, 5. – С.19 – 32.

[3] Руденко О.Г. Регуляризованный алгоритм обучения адалины в задаче оценивания нестационарных параметров / О.Г. Руденко, А.А. Бессонов // Управляющие системы и машины. – 2019. –1. – С.22 – 30.

[4] Перельман И.И. Оперативная идентификация объектов управления./ И.И. Перельман //–М.: Энергоиздат,1982. –272 с.

[5] Руденко О.Г. Модифицированный алгоритм текущего регрессионного анализа в задачах идентификации и прогнозирования. / О.Г. Руденко, И.Д. Теренковский, А. Штефан, Г.А. Ода // Радиоэлектроника и информатика. – 1998. –№ 4(05). – С.58–61

[6] Rudenko O.G. Adaline robust multistep training algorithm / O.G. Rudenko, O.O. Bezzonov // Control Systems and Computers. – 2020. – № 3. – P.15–27

[7] Аведьян Э.Д. Рекуррентный метод наименьших квадратов при коррелированных помехах / Э.Д. Аведьян // Автоматика и телемеханика . –1975.-5.-С.67-75

[8] Медведев Г.А. Рекуррентное оценивание при помощи коррелированных наблюдений / Г.А. Медведев // Автоматика и телемеханика . 1974.-5.-С.110-116

[9] Медведев Г.А. О рекуррентных оценках по коррелированным наблюдениям / Г.А. Медведев, Г.А. Хацкевич // Автоматика и телемеханика .-1979.-8.-С.69-75

[10] Джонстон Дж. Эконометрические методы./ Дж. Джонстон - М.: Статистика,1980. -444с.

[11] Фаддеев Д.К. Вычислительные методы линейной алгебры. / Д.К. Фаддеев, В.Н. Фаддеева – М.: Физматгиз,1963. – 735 с.

Надійшла до редколегії 29.01.2021