



ОБ ОДНОМ РЕКУРРЕНТНОМ РЕГУЛЯРИЗОВАННОМ АЛГОРИТМЕ ИДЕНТИФИКАЦИИ

АГАДЖАНОВ С.Г., АКСАК Н.Г., ТЕРЕНКОВСКИЙ И.Д.

Предлагаются рекуррентные формы оценок, позволяющих упростить операцию обращения матриц и ускорить процесс идентификации.

1. Основная часть

Наиболее часто оценивание параметров линейной регрессионной модели вида

$$y_n = c^*T x_n + \xi_n, \quad (1)$$

где y_n – наблюдаемый выходной сигнал; $x_n = (x_{n1}, x_{n2}, \dots, x_{nN})^T$ – вектор входных сигналов $N \times 1$; $c^* = (c_1^*, c_2^*, \dots, c_N^*)^T$ – вектор искомых параметров $N \times 1$; x_n – помеха; $n=0, 1, 2, \dots$ – дискретное время, при числе наблюдений большем числа оцениваемых параметров ($n \gg N$) успешно решается с помощью метода наименьших квадратов (МНК). Если же $n < N$, то при $x_n \neq 0$ получаемая путём минимизации квадратичного функционала линейная несмещённая оценка с минимальной среднеквадратичной ошибкой имеет вид [1]

$$c_n = X_n (X_n^T X_n)^{-1} Y_n, \quad (2)$$

а при $\xi_n \sim N(0, \sigma_\xi^2)$ определяется следующим образом:

$$c_n = X_n (X_n^T X_n + \sigma_\xi^2 I)^{-1} Y_n. \quad (3)$$

Здесь $Y_n = (y_n, y_{n-1}, \dots, y_1)^T$ – вектор $n \times 1$; $X_n = (x_n, x_{n-1}, \dots, x_1)$ – матрица $N \times n$; I – единичная матрица $n \times n$.

Общеизвестны неудобства, возникающие при использовании нерекуррентных алгоритмов [2]. Поэтому целью данной работы является получение рекуррентных форм оценок (2) и (3), позволяющих упростить операцию обращения матриц и ускорить процесс идентификации. Однако переход от обычной к рекуррентной форме алгоритма, замена обращения матрицы обращением числа не снимает проблемы численной устойчивости решения плохо обусловленных систем. Как отмечается в [3], чёткое различие между вырожденными и невырожденными матрицами существует только в теории вещественных чисел, а поскольку действия над матрицами на ЭВМ производятся с округлениями, то это различие становится неопределённым и зависит от разрядности ЭВМ. Поэтому наличие в обрабатываемой матрице оценки (3) положительного слагаемого, определяемого дисперсией помехи, должно привести к повы-

шению устойчивости процесса идентификации, т.е. σ_ξ^2 играет регуляризующую роль.

Так как оценка (2) является частным случаем оценки (3), остановимся на получении рекуррентной формы оценки (3), для чего воспользуемся подходом, предложенным в [4].

Вместо исходных векторов x_n, x_{n-1}, \dots размерности $N \times 1$ рассмотрим расширенные векторы $\tilde{x}_n, \tilde{x}_{n-1}, \dots$ размерности $(N+n) \times 1$, которые имеют следующий вид:

$$\tilde{x}_n^T = (x_n^T, \sigma, 0, \dots, 0), \tilde{x}_{n-1}^T = (x_{n-1}^T, 0, \sigma, \dots, 0), \dots, \tilde{x}_1^T = (x_1^T, 0, 0, \dots, \sigma),$$

т.е. каждый измеренный вектор входных воздействий дополняется $(n-1)$ -й нулевыми компонентами и одной ненулевой. Соответствующая матрица наблюдений, составленная из векторов $\tilde{x}_n, \tilde{x}_{n-1}, \dots, \tilde{x}_1$, выглядит так:

$$\tilde{X}_n = \begin{pmatrix} X_n \\ \Lambda \\ \sigma I \end{pmatrix}.$$

С учётом этого оценка (3) может быть переписана в следующем виде:

$$c_n = X_n (\tilde{X}_n^T \tilde{X}_n)^{-1} Y_n. \quad (4)$$

Разобьём матрицы X_n, \tilde{X}_n и вектор Y_n на блоки.

Тогда, обозначив $P_n = (\tilde{X}_n^T \tilde{X}_n)^{-1}$, получим

$$P_n = (\tilde{X}_n^T \tilde{X}_n)^{-1} = \begin{pmatrix} \frac{1}{\alpha_n} & M & \frac{1}{\alpha_n} \\ K & & \Lambda \\ -\frac{1}{\alpha_n} P_{n-1} \tilde{X}_{n-1}^T \tilde{x}_n & M P_{n-1} + \frac{1}{\alpha_n} P_{n-1} \tilde{X}_{n-1}^T \tilde{x}_n \tilde{x}_n^T \tilde{X}_{n-1} P_{n-1} & \end{pmatrix}, \quad (5)$$

где $P_{n-1} = (\tilde{X}_{n-1}^T \tilde{X}_{n-1})^{-1}$;

$$X_n = (x_n M_{n-1}), \tilde{X}_n = (\tilde{x}_n \tilde{M}_{n-1}),$$

$$Y_n = \begin{pmatrix} y_n \\ \Lambda \\ Y_{n-1} \end{pmatrix};$$

$$\alpha_n = \tilde{x}_n^T (I - \tilde{X}_{n-1} (\tilde{X}_{n-1}^T \tilde{X}_{n-1})^{-1} \tilde{X}_{n-1}^T) \tilde{x}_n.$$

Следует отметить, что входящая в выражение для α_n матрица $R_{n-1} = I - \tilde{X}_{n-1} (\tilde{X}_{n-1}^T \tilde{X}_{n-1})^{-1} \tilde{X}_{n-1}^T$ может вычисляться рекуррентно. Для этого также воспользуемся блочным представлением матрицы

$$\tilde{X}_{n-1} = (\tilde{x}_{n-1} \tilde{M}_{n-2}).$$

Подставляя в R_{n-1} выражение для P_{n-1} , аналогичное (5), производя умножение и несложные преобразования, получаем

$$R_{n-1} = R_{n-2} - \frac{R_{n-2} \tilde{x}_{n-1} \tilde{x}_{n-1}^T R_{n-2}}{\tilde{x}_{n-1}^T R_{n-2} \tilde{x}_{n-1}}.$$

После подстановки (5) в (4) с учётом того, что

$$c_{n-1} = X_{n-1} (\tilde{X}_{n-1}^T \tilde{X}_{n-1})^{-1} Y_{n-1},$$

имеем

$$c_n = c_{n-1} + \frac{1}{\alpha_n} \left[x_n - X_{n-1} (\tilde{X}_{n-1}^T \tilde{X}_{n-1})^{-1} \tilde{X}_{n-1}^T \tilde{x}_n \right] \times \\ \times \left[y_n - \tilde{x}_n^T \tilde{X}_{n-1} (\tilde{X}_{n-1}^T \tilde{X}_{n-1})^{-1} Y_{n-1} \right],$$

где $\alpha_n = \tilde{x}_n^T R_{n-1} \tilde{x}_n$.

Обозначим $\tilde{X}_{n-1} (\tilde{X}_{n-1}^T \tilde{X}_{n-1})^{-1} Y_{n-1} = \tilde{c}_{n-1}$.

Тогда, воспользовавшись блочным представлением

$$\tilde{X}_{n-1} = (\tilde{x}_{n-1} \tilde{M}_{n-2}), \quad Y_n = \begin{pmatrix} Y_{n-1} \\ \Lambda \\ Y_{n-2} \end{pmatrix},$$

после несложных преобразований получаем, что вспомогательный вектор \tilde{c}_{n-1} может вычисляться рекуррентно в соответствии с алгоритмом

$$\tilde{c}_{n-1} = \tilde{c}_{n-2} + \frac{1}{\alpha_n} R_{n-2} \tilde{x}_{n-1} (y_{n-1} - \tilde{c}_{n-2}^T \tilde{x}_{n-1}). \quad (6)$$

Рассмотрим входящую в (6) матрицу $X_{n-1} (\tilde{X}_{n-1}^T \tilde{X}_{n-1})^{-1} \tilde{X}_{n-1}^T$. Обозначим её M_{n-1} . Снова воспользуемся блочным представлением матриц X_{n-1} и \tilde{X}_{n-1} . После несложных преобразований имеем

$$M_{n-1} = M_{n-2} - \frac{1}{\alpha_{n-1}} (x_{n-1} - M_{n-2} \tilde{x}_{n-1}) \tilde{x}_{n-1}^T R_{n-2},$$

где $\alpha_{n-1} = \tilde{x}_{n-1}^T R_{n-2} \tilde{x}_{n-1}$.

Таким образом, в окончательном виде рекуррентный алгоритм вычисления оценки (3) выглядит следующим образом (здесь для простоты вычисляемые векторы и матрицы выписаны для n -го шага процесса идентификации (такта)):

$$c_n = c_{n-1} + \frac{1}{\alpha_n} (x_n - M_{n-1} \tilde{x}_n) (y_n - c_{n-1}^T x_n); \quad (7)$$

$$M_n = M_{n-1} - \frac{1}{\alpha_n} (x_n - M_{n-1} \tilde{x}_n) \tilde{x}_n^T R_{n-1}; \quad (8)$$

$$R_n = R_{n-1} - \frac{1}{\alpha_n} R_{n-1} \tilde{x}_n \tilde{x}_n^T R_{n-1}; \quad (9)$$

$$\tilde{c}_n = \tilde{c}_{n-1} + \frac{1}{\alpha_n} R_{n-1} \tilde{x}_n (y_n - \tilde{c}_{n-1}^T \tilde{x}_n); \quad (10)$$

$$\alpha_n = \tilde{x}_n^T R_{n-1} \tilde{x}_n, \quad R_0 = I. \quad (11)$$

Нетрудно видеть, что при отсутствии помех ($\sigma_\xi^2 = 0$) и алгоритм (7) - (11), т.е. рекуррентное выражение для вычисления оценки (2), принимает вид [2]:

$$c_n = c_{n-1} + \frac{1}{\alpha_n} R_{n-1} x_n (y_n - c_{n-1}^T x_n); \quad (12)$$

$$R_n = R_{n-1} - \frac{1}{\alpha_n} R_{n-1} x_n x_n^T R_{n-1}; \quad (13)$$

$$\alpha_n = x_n^T R_{n-1} x_n, \quad R_0 = I. \quad (14)$$

Отметим, что стоящие в знаменателях приведенных соотношений величины $\alpha_n = x_n^T R_{n-1} x_n$ являются отношением определителей Грама на двух соседних шагах процесса идентификации и характеризуют степень линейной зависимости вектора x_n с векторами x_{n-1}, \dots . Поэтому, если на некоторых шагах процесса идентификации появляется вектор измерений входных переменных, являющийся линейной комбинацией предыдущих векторов, α_n будет равно нулю. Следовательно, в этом случае настройка коэффициентов c_n производиться не будет, так как

полученная предыдущая оценка удовлетворяет последним уравнениям $y_i - c_{n-1}^T x_i = 0$ ($i=n-1, \dots$), а вновь поступившее измерение x_n не несёт новой информации. Поэтому вычисление α_n позволяет исследовать матрицу наблюдений X_n на "информативность", т.е. находить в ней столбцы, линейно зависящие от других (при заданной точности вычислений).

2. Заключение

Сравнивая величины α_n , входящие в (7) - (11) и (12) - (14), можно увидеть следующее. При $n=1$ величина α_n , определяемая (14), будет равна $\alpha_1 = \|x_1\|^2$, т.е. алгоритм (12) - (14) будет соответствовать алгоритму Качмажа. При этом α_1 может оказаться равной нулю, т.е. оценка будет неустойчивой, если компоненты вектора x_1 находятся вблизи машинного нуля.

Для алгоритма (7) - (11) величина $\alpha_1 = \|x_1\|^2 + \sigma_\xi^2$, а так как $\sigma_\xi^2 > 0$, оценка будет более устойчивой.

Далее, при $n=2$ для алгоритма (12) - (14)

$$\alpha_2 = \|x_1\|^2 \|x_2\|^2 - (x_1^T x_2)^2 = \|x_1\|^2 \|x_2\|^2 \sin^2 \varphi,$$

где φ - угол между векторами x_1 и x_2 . Поэтому, как было сказано выше, в случае линейной зависимости x_1 и x_2 $\alpha_2 = 0$.

Для алгоритма (7) - (11)

$$\alpha_2 = \|x_1\|^2 \|x_2\|^2 \sin^2 \varphi + \sigma_\xi^2 (\|x_1\|^2 + \|x_2\|^2), \quad \text{т.е.}$$

при $\varphi=0$, $\alpha_2 \neq 0$ и алгоритм является более устойчивым. Таким образом, в данном алгоритме величина дисперсии помехи играет регуляризующую роль. В случае, когда дисперсия помехи не известна, в качестве σ_ξ^2 следует брать оценку этой дисперсии. Однако даже при отсутствии помех целесообразно применять алгоритм (3) при (7) - (11) с некоторым положительным σ^2 , что позволяет получить более устойчивые оценки.

Литература. 1. Браммер К., Зиффлинг Г. Фильтр Калмана-Бьюси. М.: Наука, 1982. 200с. 2. Алберт Р. Регрессия, псевдоинверсия и рекуррентное оценивание. М.: Наука, 1977. 224с. 3. Браверман В.Я., Жуковский Е.Л. Использование обобщённых обратных матриц в методе группового учёта аргументов // Автоматика, 1985. №1. С.83-86 4. Руденко О.Г. Об одном редуцированном многошаговом алгоритме идентификации линейных объектов // Автоматика, 1987. №4. С.7-11.

Поступила в редколлегию 4.01.99

Рецензент: д-р техн. наук, проф. Руденко О. Г.

Агаджанов Семен Грантович, канд. техн. наук, младший научный сотрудник кафедры ЭВМ ХТУРЭ. Научные интересы: программирование. Увлечения и хобби: туризм. Адрес: Украина, 130076, Харьков, пр. Ленина, 14, тел. 45-20-20.

Аксак Наталья Георгиевна, канд. техн. наук, старший научный сотрудник кафедры ЭВМ ХТУРЭ. Научные интересы: системы идентификации. Увлечения и хобби: спорт. Адрес: Украина, 130076, Харьков, пр. Ленина, 14, тел. 40-93-54.

Теренковский Иван Дмитриевич, аспирант кафедры ЭВМ ХТУРЭ. Научные интересы: программирование баз данных. Увлечения и хобби: компьютер. Адрес: Украина, 130076, Харьков, пр. Ленина, 14, тел. 40-93-54.