

4. Чикирдин Э. Г., Стольцер С. М., Астраханцев Ф. А. Рентгеновские томографические аппараты.— М.: Медицина, 1976.
 5. Бердонос В. А., Горбунов В. И., Стоянов А. К. Голографический синтез 3-мерных рентгеновских изображений // Материалы 7-й Всесоюз. школы по голографии.— Л.: ЛИЯФ, 1975.

Поступила в редакцию 2 ноября 1983 г.

УДК 62-501.4 : 518.5

О. Г. РУДЕНКО

(Харьков)

ОБ ОДНОМ РЕДУЦИРОВАННОМ МНОГОШАГОВОМ АЛГОРИТМЕ ИДЕНТИФИКАЦИИ ЛИНЕЙНЫХ ОБЪЕКТОВ

Введение. В ряде работ [1—5] для идентификации объекта, описываемого уравнением

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

где y_n — наблюдаемый выходной сигнал; $x_n = (x_{1n}, \dots, x_{Nn})^T$ — вектор входных сигналов; $c^* = (c_1^*, c_2^*, \dots, c_N^*)^T$ — вектор искомых параметров; ξ_n — случайная помеха, используются многошаговые адаптивные алгоритмы, которые путем увеличения объема учитываемой информации позволяют в ряде случаев добиться существенного ускорения процесса идентификации.

В [2, 4, 5] изучалась следующая многошаговая процедура:

$$c_n = c_{n-1} + X_n^{(S)} [X_n^{(S)T} X_n^{(S)}]^{-1} E_n^{(S)}, \quad (2)$$

здесь $X_n^{(S)} = (x_n, x_{n-1}, \dots, x_{n-s+1})$ — матрица $N \times S$; $E_n^{(S)} = (y_n - c_{n-1}^T x_n, y_{n-1} - c_{n-1}^T x_{n-1}, \dots, y_{n-s+1} - c_{n-1}^T x_{n-s+1})^T$ — вектор $S \times 1$; S — число учитываемых шагов в алгоритме (глубина памяти).

При исследовании свойств алгоритма (2) предполагается, что: 1) $\{x_{ik}\}$ — независимые гауссовы случайные величины с $M\{x_{ik}\} = 0$ и одинаковой дисперсией $M\{x_{ik}x_{jm}\} = \sigma_x^2 \delta_{ij} \delta_{km}$; 2) ξ_n — гауссова величина с $M\{\xi_n\} = 0$ и $M\{\xi_n^2\} = \delta_\xi^2$.

Некоторые свойства проекционных алгоритмов. Если ввести в рассмотрение ошибку идентификации $\Theta_i = c_i - c^*$, то (при учете, что $E_n^{(S)} = X_n^{(S)T} \Theta_{n-1}$) алгоритм (2) может быть записан следующим образом:

$$\Theta_n = (I_N - P_n^{(S)}) \Theta_{n-1} + B_n^{(S)} \xi_n^{(S)}, \quad (3)$$

где I_N — единичная матрица $N \times N$; $P_n^{(S)} = X_n^{(S)} [X_n^{(S)T} X_n^{(S)}]^{-1} X_n^{(S)T}$ — матрица ортогонального проектирования на линейную оболочку векторов $x_n, x_{n-1}, \dots, x_{n-s+1}$; $B_n^{(S)} = X_n^{(S)} [X_n^{(S)T} X_n^{(S)}]^{-1}$ — матрица $N \times S$; $\xi_n^{(S)} = (\xi_n, \xi_{n-1}, \dots, \xi_{n-s+1})$ — вектор $S \times 1$.

Учет принятых допущений о X_n и ξ_n , свойств операторов проектирования [6], а также следующих свойств матриц $P_n^{(S)}$ и $B_n^{(S)}$ [5]:

$$M_{x_n/x_{n-1}, \dots, x_{n-s+1}, \xi} \{P_n^{(S)}\} = \left(1 - \frac{1}{m}\right) P_{n-1}^{(S-1)} + \frac{1}{m} I_N;$$

$$M_{(x)} \{B_n^{(S)T} B_n^{(S)}\} = M_{(x)} \{[X_n^{(S)T} X_n^{(S)}]^{-1}\} = \frac{1}{(m-2)\sigma_x^2} I_S;$$

$$M_{(x)} \{B_n^{(S)T} P_n^{(S-1)} B_n^{(S)}\} = \frac{1}{N-S} I_S^{(S-1)},$$

где $m = N - S + 1$; $I_S^{(S-1)} = \begin{pmatrix} I_{S-1} & 0 \\ 0 & 0 \end{pmatrix}$,

позволяет исследовать алгоритм (2). При этом справедливо [5]:

$$\lim_{n \rightarrow \infty} M \{\Theta_n\} = 0; \quad (4)$$

$$M \{\|\Theta_n\|^2\} = k_s \|\Theta_0\|^2 \left(1 - \frac{1}{m}\right)^{n-S} + \frac{N+S-1}{N-S-1} \frac{\sigma_{\xi}^2}{\sigma_x^2} \left(1 - \left(1 - \frac{1}{m}\right)^{n-S}\right). \quad (5)$$

Здесь k_s — коэффициент, зависящий от способа итерирования на первых S шагах. Например, если на этих шагах происходит уточнение оценок по алгоритму Качмажа, то $k_s = (1 - (1/N))^s$; если же на первом шаге используется одношаговый алгоритм (Качмажа), на втором — двухшаговый и т. д., то $k_s = (1 - (1/N))(1 - 2/N), \dots$. В случае, если на первых S шагах информация накапливается, а оценка строится только после прихода x_s, x_{s-1}, \dots, x_1 , то $k_s = (1 - (s/N))$.

Из выражения (4) следует, что оценка, получаемая с помощью (2), асимптотически не смещена. Формула (5) устанавливает область, к которой сходится алгоритм, т. е.

$$\lim_{n \rightarrow \infty} M \{\|\Theta_n\|^2\} = \frac{N+S-1}{N-S-1} \frac{\sigma_{\xi}^2}{\sigma_x^2}. \quad (6)$$

Из (5) и (6) видно, что с ростом S не только увеличивается скорость сходимости (первое слагаемое в (5)), но и возрастает размер области сходимости (6).

Редуцированные S-шаговые алгоритмы и их свойства. Один из способов уменьшения области сходимости (6) — переход от несмещенных оценок к смещенным, так как известно, что в случае нормальной выборки в классе всех оценок может быть получена оценка, имеющая наименьший средний квадрат ошибки, однако являющаяся смещенной в отличие от обычной. Такой переход целесообразен и для повышения стабильности оценок, получаемых с помощью (2), при наличии мультиколлинеарности. (Отказ от требования несмещенности позволяет получать устойчивые оценки даже при вырожденных матрицах наблюдений.) По аналогии с редуцированными оценками, построенными на основе метода наименьших квадратов и хорошо изученными в литературе (см., например, [7]), рассмотрим оценки, основой которых являются многошаговые алгоритмы (2).

Редуцированный S -шаговый алгоритм имеет вид

$$c_n^{\lambda} = c_{n-1}^{\lambda} + X_n^{(S)} [X_n^{(S)T} X_n^{(S)}]^{-1} \bar{E}_n^{(S)}, \quad (7)$$

где $\bar{E}_n^{(S)} = (\lambda y_n - c_{n-1}^{\lambda T} x_n, \dots, \lambda y_{n-S+1} - c_{n-1}^{\lambda T} x_{n-S+1})^T$ — вектор $S \times 1$; $0 < \lambda \leq 1$ — коэффициент редукции.

При изучении свойств данного алгоритма используются сделанные выше предположения относительно статистических свойств полезных сигналов и помех.

Вычитая из обеих частей (7) c^* , запишем алгоритм относительно ошибок идентификации Θ_n . Так как выходные сигналы измеряются с помехами, то (7) можно представить следующим образом:

$$\Theta_n^{\lambda} = (I_N - P_n^{(S)}) \Theta_{n-1}^{\lambda} - (1 - \lambda) P_n^{(S)} c^* + \lambda B_n^{(S)} \xi_n^{(S)}. \quad (8)$$

Определим $M \{\Theta_n^{\lambda}\}$. С учетом статистических свойств операторов $P_n^{(S)}$ и статистических свойств полезных сигналов и помех после усреднения (8) и итерирования имеем

$$M \{\Theta_n^{\lambda}\} = k_s \Theta_0 \left(1 - \frac{1}{m}\right)^{n-S} - (1 - \lambda) \left(1 - \left(1 - \frac{1}{m}\right)^{n-S}\right) c^*,$$

откуда следует, что с ростом времени идентификации

$$\lim_{n \rightarrow \infty} M \{\Theta_n^{\lambda}\} = -(1 - \lambda) c^*,$$

т. е.

$$\lim_{n \rightarrow \infty} M \{c_n^{\lambda}\} = \lambda c^*. \quad (9)$$

Таким образом, оценка c_n^{λ} смещена.

Умножив обе части (8) слева на $\Theta_n^{\lambda T}$, после усреднения полученного выражения имеем

$$M \{ \|\Theta_n^\lambda\|^2 \} = M \{ \Theta_{n-1}^{\lambda T} (I_N - P_n^{(S)}) \Theta_{n-1}^\lambda \} + (1 - \lambda)^2 M \{ c^{*T} P_n^{(S)} c^* \} + \lambda^2 M \{ \xi_n^{(S)T} B_n^{(S)T} B_n^{(S)} \xi_n^{(S)} \}.$$

Поступая по аналогии с [5], приходим к

$$M \{ \|\Theta_n^\lambda\|^2 \} = k_S \|\Theta_0\|^2 \left(1 - \frac{1}{m}\right)^{n-S} + (1 - \lambda)^2 \left(1 - \left(1 - \frac{1}{m}\right)^{n-S}\right) \|c^*\|^2 + \lambda^2 \frac{N + S - 1}{N - S - 1} \frac{\sigma_\xi^2}{\sigma_x^2} \left(1 - \left(1 - \frac{1}{m}\right)^{n-S}\right),$$

откуда следует, что

$$\lim_{n \rightarrow \infty} M \{ \|\Theta_n^\lambda\|^2 \} = (1 - \lambda)^2 \|c^*\|^2 + \lambda \frac{N + S - 1}{N - S - 1} \frac{\sigma_\xi^2}{\sigma_x^2}. \quad (10)$$

Из (10) видно, что, выбирая в (7) коэффициент редукции $0 < \lambda < 1$, получим оценку c_n^λ с $M \{ \|\Theta_n^\lambda\|^2 \} < M \{ \|\Theta_n\|^2 \}$. Более точно определить границы изменения λ можно следующим образом. Так как требуется, чтобы $M \{ \|\Theta_n^\lambda\|^2 \} < M \{ \|\Theta_n\|^2 \}$, то, учитывая (10) и (6), имеем неравенство

$$(1 - \lambda)^2 \|c^*\|^2 + \lambda^2 \frac{N + S - 1}{N - S - 1} \frac{\sigma_\xi^2}{\sigma_x^2} < \frac{N + S - 1}{N - S - 1} \frac{\sigma_\xi^2}{\sigma_x^2},$$

которое, как легко убедиться, справедливо при:

- а) $0 < \lambda < 1$, если $\|c^*\|^2 < \frac{N + S - 1}{N - S - 1} \frac{\sigma_\xi^2}{\sigma_x^2}$,
- б) $\frac{(N - S - 1) \sigma_x^2 \|c^*\|^2 - (N + S - 1) \sigma_\xi^2}{(N - S - 1) \sigma_x^2 \|c^*\|^2 + (N + S - 1) \sigma_\xi^2} < \lambda < 1$ в противном случае.

Оптимальное значение коэффициента редукции $\lambda^{\text{опт}}$ определим, минимизируя (10). После дифференцирования выражения (10) по λ и приравнивания производной нулю, получаем линейное относительно искомого параметра уравнение, решая которое находим

$$\lambda^{\text{опт}} = \frac{(N - S - 1) \sigma_x^2 \|c^*\|^2}{(N - S - 1) \sigma_x^2 \|c^*\|^2 + (N + S - 1) \sigma_\xi^2}. \quad (11)$$

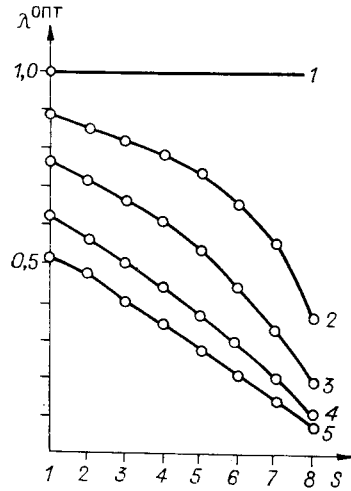
Так как $\partial^2 M \{ \|\Theta_n^\lambda\|^2 \} / \partial \lambda^2 > 0$, выражение (11) минимизирует (10). Этому значению $\lambda^{\text{опт}}$ соответствует

$$M \{ \|\Theta_\infty^\lambda\|^2 \}_{\min} = \frac{(N - S - 1) \|c^*\|^2 \sigma_\xi^2 (3 \|c^*\|^2 (N + S - 1) + (N - S - 1) \sigma_\xi^2)}{(N + S - 1) \sigma_x^2 (N - S - 1) \sigma_\xi^2 + (N + S - 1) \|c^*\|^2 \sigma_x^2}.$$

Следует отметить, что использование (11) на практике не представляется возможным, так как значение $\lambda^{\text{опт}}$ зависит от неизвестных параметров c^* , σ_x^2 , σ_ξ^2 . Поэтому при практическом применении (7), вместо входящих в выражение (11) c^* , σ_x^2 , σ_ξ^2 , следует воспользоваться их оценками \hat{c}_n^λ , $\hat{\sigma}_x^2$, $\hat{\sigma}_\xi^2$, т. е. перейти от $\lambda^{\text{опт}}$, определяемого выражением (11), к стохастическому коэффициенту редукции

$$\hat{\lambda}_n^{\text{опт}} = \frac{(N - S - 1) \hat{\sigma}_x^2 \|\hat{c}_n^\lambda\|^2}{(N - S - 1) \hat{\sigma}_x^2 \|\hat{c}_n^\lambda\|^2 + (N + S - 1) \hat{\sigma}_\xi^2}.$$

В этом случае $\hat{\lambda}_n^{\text{опт}}$ зависит от времени n . Поэтому данное значение $\hat{\lambda}_n^{\text{опт}}$ следует применять для нахождения оценки c_{n+1}^λ , по которой определяется $\hat{\lambda}_{n+1}^{\text{опт}}$ и т. д.



Зависимость оптимального значения коэффициента редукции от глубины памяти алгоритма

Из вышеприведенных формул видно, что при отсутствии помех ($\sigma_{\xi}^2 \equiv 0$) $\lambda^{\text{опт}} = 1$, а $M\{\|\Theta_{\infty}\|^2\} = 0$, т. е. параметры определяются точно.

На рисунке показаны зависимости $\lambda^{\text{опт}}$ от S для $N = 10$, $\|c^*\|^2 = 1$ и различных соотношений $\sigma_{\xi}^2/\sigma_x^2$. Кривая 1 соответствует случаю $\sigma_{\xi}^2/\sigma_x^2 = 0$; 2 — $\sigma_{\xi}^2/\sigma_x^2 = 0,1$; 3 — $\sigma_{\xi}^2/\sigma_x^2 = 0,25$; 4 — $\sigma_{\xi}^2/\sigma_x^2 = 0,5$; 5 — $\sigma_{\xi}^2/\sigma_x^2 = 0,75$. Приведенные зависимости свидетельствуют о том, что значение $\lambda^{\text{опт}}$ уменьшается с ростом S и с увеличением соотношения $\sigma_{\xi}^2/\sigma_x^2$. Аналогичный характер имеют зависимости и для других $\|c^*\|^2$.

Рекуррентная форма редуцированных S-шаговых алгоритмов. Применение алгоритмов (2), (7) требует вычисления на каждом шаге итерационного процесса идентификации обратной матрицы $[X_n^{(S)T} X_n^{(S)}]^{-1}$. Более удобен с вычислительной точки зрения рекуррентный пересчет этой матрицы. Для получения рекуррентной формы алгоритма (7) (алгоритм (2) является частным случаем (7) с $\lambda = 1$) разобьем матрицу $X_n^{(S)}$ на блоки $[x_n : X_{n-1}^{(S-1)}]$, где $X_{n-1}^{(S-1)} = (x_{n-1}, \dots, x_{n-S+1})$ — матрица $N \times (S-1)$. С учетом правил умножения и обращения блочных матриц после несложных преобразований имеем

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

где $R_{n-1}^{(S-1)} = I_N - X_n^{(S)} [X_n^{(S)T} X_n^{(S)}]^{-1} X_n^{(S)T}$, $\alpha_n = x_n^T R_{n-1}^{(S-1)} x_n$.

Использование блочного представления матрицы $X_{n-1}^{(S-1)} = x_{n-1} : X_{n-2}^{(S-2)}$ позволяет получить рекуррентную форму вычисления $R_{n-1}^{(S-1)}$. Опуская несложные преобразования, окончательно запишем

$$R_{n-S+i}^{(i)} = R_{n-S+i-1}^{(i-1)} - \frac{R_{n-S+i-1}^{(i-1)} x_{n-S+i} x_{n-S+i}^T R_{n-S+i-1}^{(i-1)}}{\alpha_{n-S+i}}, \quad (13)$$

где $\alpha_{n-S+i} = x_{n-S+i}^T R_{n-S+i-1}^{(i-1)} x_{n-S+i}$, $i = \overline{1, S-1}$, $R_{n-S}^{(0)} = I$.

Таким образом, непосредственное обращение матрицы $X_n^{(S)T} X_n^{(S)}$ заменяется ее рекуррентным вычислением, связанным с необходимостью проведения на каждом шаге процесса идентификации $(S-1)$ -й итерации в соответствии с (13). Рекуррентная форма алгоритма с $\lambda = 1$ приведена в [2].

Следует отметить, что входящие в (12), (13) коэффициенты α представляют собой отношение определителей Грама, свидетельствующих о степени линейной зависимости входных векторов. Если на каком-либо шаге, например i -м, α_i обращается в нуль, то данное измерение x_i в алгоритм не включают, а переходят к следующему измерению x_{i+1} .

Повышение вычислительной устойчивости алгоритма (7) может быть осуществлено путем введения в него регуляризирующей добавки $\delta^2 > 0$ следующим образом:

$$c_n^{\lambda} = c_{n-1}^{\lambda} + X_n^{(S)} [X_n^{(S)T} X_n^{(S)} + \delta^2 I_S]^{-1} \widehat{E}_n^{(S)}. \quad (14)$$

Нетрудно убедиться, что такое введение δ^2 дополнительного смещения оценки не вызывает, а приводит к тому, что коэффициенты α в нуль не обращаются. Это позволяет корректировать оценку c_n на каждой итерации, используя все наблюдения.

Для получения рекуррентной формы (14) следует вместо исходных векторов x размерностью $N \times 1$ ввести расширенные векторы \tilde{x} размерностью $(N + S) \times 1$. При этом векторы $\tilde{x}_n, \tilde{x}_{n-1}, \dots, \tilde{x}_{n-S+1}$ имеют следующий вид: $\tilde{x}_n^T = (x_n^T, 0, 0, \dots, \delta)$; $\tilde{x}_{n-1}^T = (x_{n-1}^T, 0, 0, \dots, \delta, 0), \dots, \tilde{x}_{n-S+1}^T = (x_{n-S+1}^T, \delta, 0, \dots, 0)$, т. е. каждый измеренный вектор дополняется $(S - 1)$ -м нулевым компонентом и одним ненулевым — δ . Соответствующие матрицы $\tilde{X}_n^{(S)}$, составленные из векторов $\tilde{x}_n, \tilde{x}_{n-1}, \dots, \tilde{x}_{n-S+1}$, имеют вид

$$\tilde{X}_n^{(S)} = \begin{bmatrix} X_n^{(S)} \\ \dots \\ \delta I_S \end{bmatrix}.$$

Поступая как и ранее, т. е. разбивая $X_n^{(S)}$ и $\tilde{X}_n^{(S)}$ на блоки и т. д., после несложных преобразований получаем

$$c_n^\lambda = c_{n-1}^\lambda + \frac{1}{\alpha_n} (x_n - \tilde{P}_{n-1}^{(S-1)} \tilde{x}_n) (e_n - \tilde{x}_n^{\lambda T} \tilde{c}_{n-1}^\lambda); \quad (15)$$

$$\tilde{c}_n^\lambda = \tilde{c}_{n-1}^\lambda + \frac{1}{\alpha_n} \tilde{R}_{n-1}^{(S-1)} \tilde{x}_n (\lambda y_n - \tilde{x}_n^T \tilde{c}_{n-1}^\lambda), \quad (16)$$

где $\tilde{\alpha}_n = \tilde{x}_n^T \tilde{R}_{n-1}^{(S-1)} \tilde{x}_n$; $e_n = \lambda y_n - \tilde{x}_n^T \tilde{c}_{n-1}^\lambda$; $\tilde{P}_{n-1}^{(S-1)} = X_{n-1}^{(S-1)} [\tilde{X}_{n-1}^{(S-1)T} \tilde{X}_{n-1}^{(S-1)}]^{-1} \times X_{n-1}^{(S-1)T}$ — матрица $N \times (N + S)$; $\tilde{R}_{n-1}^{(S-1)} = I_{N+S} - \tilde{X}_{n-1}^{(S-1)} [\tilde{X}_{n-1}^{(S-1)} \tilde{X}_{n-1}^{(S-1)T}]^{-1} \tilde{X}_{n-1}^{(S-1)T}$ — матрица $(N + S) \times (N + S)$, которые могут быть вычислены рекуррентно следующим образом:

$$\tilde{P}_{n-S+i}^{(i)} = \tilde{P}_{n-S+i-1}^{(i-1)} + \frac{1}{\alpha_{n-S+i}} (x_{n-S+i} - \tilde{P}_{n-S+i-1}^{(i-1)} \tilde{x}_{n-S+i}) \tilde{x}_{n-S+i}^T \tilde{R}_{n-S+i-1}^{(i-1)}; \quad (17)$$

$$\tilde{R}_{n-S+i}^{(i)} = \tilde{R}_{n-S+i-1}^{(i-1)} - \frac{1}{\alpha_{n-S+i}} \tilde{R}_{n-S+i-1}^{(i-1)} \tilde{x}_{n-S+i} \tilde{x}_{n-S+i}^T \tilde{R}_{n-S+i-1}^{(i-1)}, \quad (18)$$

здесь $\tilde{\alpha}_{n-S+i} = \tilde{x}_{n-S+i}^T \tilde{R}_{n-S+i-1}^{(i-1)} \tilde{x}_{n-S+i}$, $\tilde{R}_{n-S}^{(0)} = I$, $\tilde{P}_{n-S}^{(0)} = 0$, $i = \overline{1, S-1}$.
Вычисление оценок в соответствии с (12), (13) и (15) — (18) возможно только после момента времени $n \geq S$. До этого же момента оценки определяются по данным формулам, у которых, однако, параметр S переменный.

Другая рекуррентная форма алгоритма (7) может быть получена на основании результатов работы [8]. Использование псевдообращения матриц по Муру — Пенроузу позволяет изменить процедуру вычисления матрицы $[X_n^{(S)T} X_n^{(S)}]^{-1}$ при поступлении нового входного вектора x_{n+1} . В этом случае становится возможным не пересчет матриц $(S - 1)$ раз, как это было реализовано в рассмотренных выше рекуррентных формах, а расширение матрицы путем добавления $(n + 1)$ -го наблюдения (накопления информации) и последующего исключения из нее $(n - S + 1)$ -го наблюдения (сброса устаревшей информации).

ЛИТЕРАТУРА

1. Пышкин Я. З. Адаптация и обучение в автоматических системах. — М.: Наука, 1968.
2. Аведьян Э. Д. Модифицированные алгоритмы Качмажа для оценки параметров линейных объектов // Автоматика и телемеханика. — 1978. — № 5.
3. Бодня В. Г., Руруа А. А., Чадеев В. М. Двухшаговый алгоритм идентификации линейных объектов // Автоматика и телемеханика. — 1982. — № 8.
4. Ищенко Л. А., Руденко О. Г. Многошаговые адаптивные алгоритмы идентификации линейных объектов // Автоматика. — 1983. — № 6.
5. Ищенко Л. А., Либероль Б. Д., Руденко О. Г. Проекционные алгоритмы идентификации линейных объектов // ДАН УССР. Сер. А. — 1985. — № 7.
6. Себер Дж. Линейный регрессионный анализ. — М.: Мир, 1980.
7. Демиденко Е. З. Линейная и нелинейная регрессии. — М.: Финансы и статистика, 1981.
8. Алберт А. Регрессия, псевдоинверсия и рекуррентное оценивание. — М.: Наука, 1977.

Поступила в редакцию 14 ноября 1985 г.