УДК 517.514

МИНИМАЛЬНЫЙ РИСК ДЛЯ НАИЛУЧШЕГО ПОЛИНОМИАЛЬНОГО МНОГОМЕРНО-МАТРИЧНОГО ПРЕДИКТОРА

В.С. МУХА, А.Ф. ТРОФИМОВИЧ

Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь

Поступила в редакцию 22 декабря 2007

Рассматривается задача построения оптимальной полиномиальной прогнозирующей функции, позволяющей принимать решение о случайной многомерной матрице по наблюдению другой случайной многомерной матрицы. Получен алгоритм расчета минимального риска для наилучшей полиномиальной прогнозирующей функции при квадратичной функции потерь. Алгоритм предназначен для сравнения различных полиномиальных предикторов между собой. Он может быть использован, в частности, для анализа эффективности нелинейного прогнозирования случайных процессов.

Ключевые слова: риск, многомерные матрицы, полиномиальный предиктор.

Введение

В статье рассматривается задача принятия оптимального статистического решения о случайной многомерной матрице по наблюдению другой случайной многомерной матрицы. Известно [1], что в случае квадратичной функции потерь оптимальное решение (оптимальный предиктор) определяется как апостериорное среднее. Поскольку для негауссовских распределений отыскание апостериорного среднего является сложной задачей, то целесообразно отыскивать оптимальный полиномиальный предиктор [2, 3]. При этом возникает необходимость сравнения различных полиномиальных предикторов между собой. Такое сравнение можно выполнить по величине минимального риска, соответствующего предиктору. В данной статье отыскивается минимальный риск для полиномиального многомерно-матричного предиктора при квадратичной функции потерь.

Постановка задачи

В работах [2, 3] сформулирована задача отыскания наилучшей полиномиальной прогнозирующей функции

$$y = \sum_{k=0}^{m} {}^{0,kq}\left(C_{p,kq}\, \xi^k \right) = \sum_{k=0}^{m} {}^{0,kq}\left(\xi^k C_{kq,p} \right), \quad m = 0,1,2,\ldots, \tag{1}$$

позволяющей по q-мерной матрице $\xi = (\xi_{j_q})$ получить наилучший прогноз p-мерной матрицы $\eta = (\eta_{i_p})$ в смысле минимума риска прогнозирования r при квадратичной функции потерь:

$$r = E\left({}^{0,p}(zz) \right) \to \min_{C_{(p,0)}, C_{(p,q)}, \ldots, C_{(p,mq)}}, \tag{2}$$

где

$$z = \eta - y = (z_{i_p}) = \eta - \sum_{k=0}^{m} {}^{0,kq}\left(C_{p,kq}\, \xi^k \right) = \eta - \sum_{k=0}^{m} {}^{0,kq}\left(\xi^k C_{kq,p} \right).$$

В [2] получена система уравнений для расчета коэффициентов $C_{p,kq}$ полиномиального предиктора (1) произвольной степени, а также выражения коэффициентов постоянного, линейного и квадратичного предиктора. В данной статье отыскивается минимальное значение риска (2), то есть значение, соответствующее оптимальным значениям коэффициентов предиктора (1).

Решение задачи

Матрицу вида

$$R = E\left({}^{0,0}\left(\eta - y \right)^2 \right) \tag{3}$$

назовем ковариационной матрицей риска. Можно показать, что риск r определяется как след ковариационной матрицы риска,

$$r = tr(R).$$

Поэтому будем искать ковариационную матрицу риска R. В соответствии с (3) и (1) получим:

$$R = E\left({}^{0,0}\left(\eta - y \right)^2 \right) = E\left(\eta^2 \right) - 2\sum_{k=0}^{m} {}^{0,kq}\left(E\left(\eta\xi^k \right) C_{kq,p} \right) + \sum_{l=0}^{m}\sum_{k=0}^{m} {}^{0,lq}\left({}^{0,kq}\left(C_{p,kq}\, E\left(\xi^k \xi^l \right) \right) C_{lq,p} \right).$$

Вводя обозначения начальных моментов многомерных матриц ξ, η

$$\nu_{\xi^{\lambda+k}} = E\left(\xi^\lambda \xi^k \right), \quad \nu_{\eta\xi^\lambda} = E\left(\eta\xi^\lambda \right),$$

будем иметь:

$$R = \nu_{\eta^2} - 2\sum_{k=0}^{m} {}^{0,kq}\left(\nu_{\eta\xi^k} C_{kq,p} \right) + \sum_{l=0}^{m}\sum_{k=0}^{m} {}^{0,lq}\left({}^{0,kq}\left(C_{p,kq}\, \nu_{\xi^{k+l}} \right) C_{lq,p} \right). \tag{4}$$

В [2] получено необходимое условие минимума риска (2), имеющее вид:

$$\sum_{\lambda=0}^{m} {}^{0,\lambda q}\left(C_{(p,\lambda q)} \nu_{\xi^{\lambda+s}} \right) = \nu_{\eta\xi^s}, \; s = \overline{0,m}. \tag{5}$$

При выполнении этого условия мы будем иметь минимальную ковариационную матрицу риска $R_{\min}$ и минимальное значение риска $r_{\min}$ будет определяться как след минимальной ковариационной матрицы риска:

$$r_{\min} = tr(R_{\min}). \tag{6}$$

Подставляя во второе слагаемое выражения (4) вместо моментов $\nu_{\eta\xi^k}$ сумму из левой части выражения (5) при $s = k$, $\lambda = l$, получим:

$$R_{\min} = \nu_{\eta^2} - 2\sum_{k=0}^{m} {}^{0,kq}\left(\left(\sum_{l=0}^{m} {}^{0,lq}\left(C_{(p,lq)} \nu_{\xi^{l+k}} \right) \right) C_{kq,p} \right) + \sum_{l=0}^{m}\sum_{k=0}^{m} {}^{0,lq}\left({}^{0,kq}\left(C_{p,kq}\, \nu_{\xi^{k+l}} \right) C_{lq,p} \right) =$$

$$= \nu_{\eta^2} - 2\sum_{l=0}^{m}\sum_{k=0}^{m} {}^{0,lq}\left({}^{0,kq}\; C_{p,kq}\, \nu_{\xi^{k+l}} \;\; C_{lq,p} \right) + \sum_{l=0}^{m}\sum_{k=0}^{m} {}^{0,lq}\left({}^{0,kq}\; C_{p,kq}\, \nu_{\xi^{k+l}} \;\; C_{lq,p} \right) =$$

$$= \nu_{\eta^2} - \sum_{l=0}^{m}\sum_{k=0}^{m} {}^{0,lq}\left({}^{0,kq}\; C_{p,kq}\, \nu_{\xi^{k+l}} \;\; C_{lq,p} \right).$$

Подставляя в третье слагаемое выражения (4) вместо суммы по переменной k правую часть выражения (5), полагая в этом выражении $s = l$, $\lambda = k$, получим:

$$R_{\min} = \nu_{\eta^2} - 2\sum_{k=0}^{m} {}^{0,kq}\; \nu_{\eta\xi^k} C_{kq,p} + \sum_{k=0}^{m} {}^{0,kq}\left(\sum_{l=0}^{m} {}^{0,lq}\; C_{p,lq}\, \nu_{\xi^{l+k}} \;\; C_{kq,p} \right) =$$

$$= \nu_{\eta^2} - 2\sum_{k=0}^{m} {}^{0,kq}\; \nu_{\eta\xi^k} C_{kq,p} + \sum_{k=0}^{m} {}^{0,kq}\; \nu_{\eta\xi^k} C_{kq,p} = \nu_{\eta^2} - \sum_{k=0}^{m} {}^{0,kq}\; \nu_{\eta\xi^k} C_{kq,p} \quad .$$

Таким образом, мы получили две формы оптимальной ковариационной матрицы риска:

$$R_{\min} = \nu_{\eta^2} - \sum_{l=0}^{m}\sum_{k=0}^{m} {}^{0,lq}\left({}^{0,kq}\; C_{p,kq}\, \nu_{\xi^{k+l}} \;\; C_{lq,p} \right), \tag{7}$$

$$R_{\min} = \nu_{\eta^2} - \sum_{k=0}^{m} {}^{0,kq}\; \nu_{\eta\xi^k} C_{kq,p} \quad . \tag{8}$$

Минимальный риск определяется выражением (6), где минимальная ковариационная матрица риска $R_{\min}$ находится по (7) или (8). Более экономичным для численных расчетов является выражение (8). Формула (7) оказывается полезной в аналитических преобразованиях.

Минимальный риск для наилучшего квадратичного многомерно-матричного предиктора

Найдем минимальный риск $r_{\min}$ для квадратичного многомерно-матричного предиктора ($m = 2$)

$$y = C_{p,0} + {}^{0,q}\; C_{p,q}\, \xi \; + {}^{0,2q}\; C_{p,2q}\, \xi^2 \quad .$$

Из (6), (8) получаем:

$$r_{\min} = tr(R_{\min}) = tr\left(\nu_{\eta^2} - {}^{0,0}\; C_{p,0}\, \nu_{\eta} \; - {}^{0,q}\; C_{p,q}\, \nu_{\xi\eta} \; - {}^{0,2q}\; C_{p,2q}\, \nu_{\xi^2\eta} \right). \tag{9}$$

Коэффициенты в выражении (9) получены в работе [2]. Они рассчитываются по формулам:

$$C_{(p,0)} = \nu_{\eta} - {}^{0,p}\; C_{(p,q)} \nu_{\xi} \; - {}^{0,2p}\; C_{(p,2q)} \nu_{\xi^2} \quad , \tag{10}$$

$$C_{p,q} = {}^{0,q}\left(\left(\mu_{\eta\xi} - {}^{0,2q}\; C_{p,2q} \;\; \nu_{\xi^3} - \nu_{\xi^2} \nu_{\xi} \right) \mu^{-1}_{\xi\xi} \right), \tag{11}$$

$$C_{p,2q} = {}^{0,2q}\; B_{p,2q} \; {}^{0,2q}A^{-1}_{2q,2q} \quad , \tag{12}$$

где

$$B_{p,2q} = \nu_{\eta\xi^2} - \nu_\eta \nu_{\xi^2} - {}^{0,q}\ {}^{0,q}\ \mu_{\eta\xi}\mu_{\xi\xi}^{-1}\ \ \nu_{\xi^3} - \nu_\xi \nu_{\xi^2}\ \ , \tag{13}$$

$$A_{2q,2q} = \nu_{\xi^4} - \nu_{\xi^2}\nu_{\xi^2} - {}^{0,q}\left({}^{0,q}\ \ \nu_{\xi^3} - \nu_{\xi^2}\nu_\xi\ \ \mu_{\xi\xi}^{-1}\ \ \nu_{\xi^3} - \nu_\xi \nu_{\xi^2}\ \right), \tag{14}$$

$\mu_{\xi\xi} = \nu_{\xi^2} - \nu_\xi^2$, $\mu_{\eta\xi} = \nu_{\eta\xi} - \nu_\eta \nu_\xi$.

На основе вышеизложенного приведем алгоритм расчета минимального риска для наилучшего квадратичного многомерно-матричного предиктора. В первую очередь рассчитывается $B_{p,2q}$ по (13), затем $A_{2q,2q}$ по (14), далее последовательно определяются коэффициенты $C_{p,2q}$ (12), $C_{p,q}$ (11), $C_{p,0}$ (10) и, наконец, рассчитывается минимальный риск $r_{\min}$ по формуле (9).

Минимальный риск для наилучшего квадратичного скалярного предиктора

Определим минимальный риск $r_{\min}$ для квадратичного предиктора ($m = 2$) при $p = q = 0$, то есть для наилучшего квадратичного скалярного предиктора. Предиктор в этом случае имеет вид

$$y = C_{p,0} + {}^{0,0}\ C_{p,q}\,\xi\ + {}^{0,0}\ C_{p,2q}\,\xi^2\ \ , \tag{15}$$

минимальная ковариационная матрица риска $R_{\min}$ является скалярной величиной

$$R_{\min} = \nu_{\eta^2} - \sum_{k=0}^{2} C_{p,kq}\,\nu_{\eta\xi^k}\ \ ,$$

поэтому минимальный риск определяется выражением:

$$r_{\min} = R_{\min} = \nu_{\eta^2} - C_{p,0}\,\nu_\eta - C_{p,q}\,\nu_{\eta\xi} - C_{p,2q}\,\nu_{\eta\xi^2}. \tag{16}$$

В данном случае оказывается возможным получение конечного выражения для минимального риска $r_{\min}$. Получим это выражение, выполняя расчеты в соответствии с алгоритмом, приведенным в предыдущем разделе. При этом вместо начальных моментов будем использовать центральные моменты случайных величин ξ, η. В частности, для случайных величин φ и ψ введем следующее обозначение центральных моментов:

$$\mu_{\varphi\psi} = E\ \ \ \varphi - E\ \varphi\ \ \ \ \psi - E\ \psi\ \ \ .$$

С учетом этого обозначения для случайных величин ξ, η можно получить следующие формулы связи между центральными и начальными моментами:

$$\nu_{\xi^2} = \mu_{\xi\xi} + \nu_\xi^2,\ \nu_{\eta^2} = \mu_{\eta\eta} + \nu_\eta^2,$$

$$\nu_{\eta\xi} = \mu_{\eta\xi} + \nu_\eta \nu_\xi,$$

$$\nu_{\xi^3} = \mu_{\xi\xi\xi} + 3\mu_{\xi\xi}\nu_\xi + \nu_\xi^3,$$

$$\nu_{\eta\xi^2} = \mu_{\eta\xi\xi} + 2\mu_{\eta\xi}\nu_\xi + \mu_{\xi\xi}\nu_\eta + \nu_\eta \nu_\xi^2,$$

$$\nu_{\xi^4} = \mu_{\xi\xi\xi\xi} + 4\mu_{\xi\xi\xi}\nu_\xi + 6\mu_{\xi\xi}\nu_\xi^2 + \nu_\xi^4.$$

Воспользовавшись этими выражениями, по (10)–(14) получим:

$$B_{p,2q}=\frac{\mu_{\xi\xi}\mu_{\eta\xi\xi}-\mu_{\xi\xi\xi}\mu_{\eta\xi}}{\mu_{\xi\xi}},$$

$$A_{2q,2q}=\frac{\mu_{\xi\xi\xi\xi}\mu_{\xi\xi}-\mu_{\xi\xi\xi}^2-\mu_{\xi\xi}^3}{\mu_{\xi\xi}},$$

$$C_{p,2q}=\frac{\mu_{\xi\xi}\mu_{\eta\xi\xi}-\mu_{\xi\xi\xi}\mu_{\eta\xi}}{\mu_{\xi\xi\xi\xi}\mu_{\xi\xi}-\mu_{\xi\xi\xi}^2-\mu_{\xi\xi}^3},$$

$$C_{p,q}=\frac{\mu_{\xi\xi\xi\xi}\mu_{\eta\xi}-\mu_{\xi\xi\xi}\mu_{\eta\xi\xi}+2\mu_{\xi\xi\xi}\mu_{\eta\xi}\nu_\xi-2\mu_{\xi\xi}\mu_{\eta\xi\xi}\nu_\xi-\mu_{\xi\xi}^2\mu_{\eta\xi}}{\mu_{\xi\xi\xi\xi}\mu_{\xi\xi}-\mu_{\xi\xi\xi}^2-\mu_{\xi\xi}^3},$$

$$C_{p,0}=\nu_\eta-\frac{\mu_{\xi\xi\xi\xi}\mu_{\eta\xi}\nu_\xi-\mu_{\xi\xi\xi}\mu_{\eta\xi\xi}\nu_\xi+\mu_{\xi\xi\xi}\mu_{\eta\xi}\nu_\xi^2-\mu_{\xi\xi\xi}\mu_{\xi\xi}\mu_{\eta\xi}-\mu_{\xi\xi}\mu_{\eta\xi\xi}\nu_\xi^2-\mu_{\xi\xi}^2\mu_{\eta\xi}\nu_\xi+\mu_{\xi\xi}^2\mu_{\eta\xi\xi}}{\mu_{\xi\xi\xi\xi}\mu_{\xi\xi}-\mu_{\xi\xi\xi}^2-\mu_{\xi\xi}^3}.$$

Наконец, по (16) находим минимальный риск $r_{\min}$:

$$r_{\min}=\mu_{\eta\eta}-\frac{\mu_{\xi\xi\xi\xi}\mu_{\eta\xi}^2-2\mu_{\xi\xi\xi}\mu_{\eta\xi\xi}\mu_{\eta\xi}+\mu_{\xi\xi}\mu_{\eta\xi\xi}^2-\mu_{\xi\xi}^2\mu_{\eta\xi}^2}{\mu_{\xi\xi\xi\xi}\mu_{\xi\xi}-\mu_{\xi\xi\xi}^2-\mu_{\xi\xi}^3}. \qquad (17)$$

Минимальный риск для наилучшего квадратичного скалярного предиктора в случае совместного нормального распределения

Из литературы известно (см., например, [5]), что в случае совместного нормального распределения случайных величин η и ξ оптимальный предиктор является линейным, и этому предиктору соответствует минимальный риск, определяемый выражением:

$$r_{\min}=\mu_{\eta\eta}-\frac{\mu_{\eta\xi}^2}{\mu_{\xi\xi}}.$$

Этот результат может быть получен как частный случай выражения (17). Действительно, для совместного нормального распределения случайных величин η и ξ смешанные центральные моменты третьего порядка равны нулю, т.е. $\mu_{\xi\xi\xi}=0$, $\mu_{\eta\xi\xi}=0$, и центральный момент четвертого порядка $\mu_{\xi\xi\xi\xi}=3\mu_{\xi\xi}^2$. Подставляя эти равенства в (17), получаем выражение, приведенное выше:

$$r_{\min}=\mu_{\eta\eta}-\frac{\mu_{\xi\xi\xi\xi}\mu_{\eta\xi}^2-\mu_{\xi\xi}^2\mu_{\eta\xi}^2}{\mu_{\xi\xi\xi\xi}\mu_{\xi\xi}-\mu_{\xi\xi}^3}=\mu_{\eta\eta}-\frac{3\mu_{\xi\xi}^2\mu_{\eta\xi}^2-\mu_{\xi\xi}^2\mu_{\eta\xi}^2}{3\mu_{\xi\xi}^3-\mu_{\xi\xi}^3}=\mu_{\eta\eta}-\frac{\mu_{\eta\xi}^2}{\mu_{\xi\xi}}.$$

Кроме того, в случае совместного нормального распределения коэффициент $C_{p,2q}$ в выражении (15) равен нулю, и наилучший квадратичный предиктор вырождается в известный линейный предиктор [5] вида

$$y=\nu_\eta+\frac{\mu_{\eta\xi}}{\mu_{\xi\xi}}\;\xi-\nu_\xi\;.$$

Результаты данного раздела служат дополнительным подтверждением правильности общих результатов данной статьи.

MINIMAL RISK FOR BEST POLYNOMIAL MULTIDIMENSIONAL-MATRIX PREDICTOR

V.S. MUKHA, A.F. TROFIMOVICH

Abstract

The task of construction of the optimal polynomial predicting function is considered, permitting to make a decision on a casual multidimensional matrix on observation of other casual multidimensional matrix. The algorithm of calculation of minimal risk for the best polynomial predicting function is obtained at a quadratic loss function. The algorithm intended for matching various polynomial predictors among themselves. It can be used, in particular, for the analysis of effectiveness of nonlinear forecasting of casual processes.

Литература

1. *Mukha V.S., Korchyts K.S.* // Proc. of the 7th International Conference. Minsk, 6–10 September, 2004. Vol. 2. P. 142–145.
2. *Муха В.С.* // Кибернетика и системный анализ. 2007. № 3. С. 1–7.
3. *Mukha V.S.* // Proc. of the 8th International Conference. Minsk, 11–15 September, 2007. Vol. 1. P. 166–169.
4. *Муха В.С.* Анализ многомерных данных. Минск, 2004.
5. *Муха В.С.* Теория вероятностей. Минск, 2001.