УДК 517.514

В.С. МУХА

МИНИМАЛЬНЫЙ СРЕДНИЙ РИСК И ЭФФЕКТИВНОСТЬ ОПТИМАЛЬНОГО ПОЛИНОМИАЛЬНОГО МНОГОМЕРНО-МАТРИЧНОГО ПРЕДИКТОРА

Ключевые слова: *теория оптимальных статистических решений, многомерные матрицы, полиномиальная регрессия.*

ВВЕДЕНИЕ

В настоящее время одна из актуальных проблем — прогнозирование явлений с неопределенностью поведения, например метеорологических явлений. Эту проблему можно решить в рамках теории статистических решений. Простейшей является скалярная задача теории статистических решений, состоящая в построении оптимальной прогнозирующей функции (предиктора), позволяющей по наблюдению одной случайной величины принять решение о значении другой случайной величины. В работах [1, 2] эта задача обобщена на многомерные случайные матрицы. Как в скалярном, так и в многомерно-матричном случае оптимальным предиктором при квадратичной функции потерь является апостериорное среднее, и этот предиктор в случае совместного гауссовского распределения линейный по наблюдению. Однако отыскание апостериорного среднего для распределений, отличных от гауссовских, связано с серьезными трудностями. В этом случае можно ограничиться отысканием наилучшего полиномиального предиктора. Последняя задача была поставлена и решена в работе [3]. Однако вопрос о величине среднего риска для оптимального полиномиального многомерно-матричного предиктора в ней не рассматривался. Решение этого вопроса важно с точки зрения выяснения выигрыша от использования того или иного предиктора. Этот вопрос рассматривается в данной работе. Исследуются также свойства оптимального полиномиального многомерно-матричного предиктора и вводится понятие его эффективности.

ОПТИМАЛЬНЫЕ СТАТИСТИЧЕСКИЕ РЕШЕНИЯ ДЛЯ НЕПРЕРЫВНЫХ МНОГОМЕРНО-МАТРИЧНЫХ СОСТОЯНИЙ И НАБЛЮДЕНИЙ

Пусть состояние некоторой системы представляет собой p-мерную матрицу $\eta=(\eta_{j_1,j_2,\ldots,j_p})$, $\eta\in S$, S — пространство состояний, и известна плотность вероятности состояния $f(\eta)$. В результате эксперимента над системой (наблюдения) получаем q-мерную матрицу наблюдений $\xi=(\xi_{i_1},\xi_{i_2},\ldots,\xi_{i_q})\in X$, X — пространство наблюдений. Связь ξ с η выражается условной плотностью вероятности $f(\xi/\eta)$, которую будем считать известной. Необходимо принять решение $\Delta\in D$ о состоянии системы η, где D — пространство решений такое же, как и пространство состояний.

Будем считать, что выносимое решение рандомизированно определяется плотностью вероятности $f(\Delta)$, заданной в пространстве решений D. В данном случае решение зависит от матрицы наблюдений ξ, т.е. $\Delta=\Delta(\xi)$. Функцию $\Delta=\Delta(\xi)$ будем называть прогнозирующей или предиктором. Качество решения охарактеризуем средним риском

$$r = E(W(\eta,\Delta)) = \int_S \int_X \int_D W(\eta,\Delta) f(\Delta) f(\xi/\eta) f(\eta) d\eta d\xi d\Delta\,, \qquad (1)$$

где $W(\eta,\Delta)$ — функция потерь. Задача состоит в определении решающей функции $f(\Delta(\xi))$, минимизирующей средний риск:

$$r = r(f(\Delta)) \to \min_{f(\Delta)}.$$

Данную задачу можно интерпретировать как задачу оценивания многомерной случайной матрицы η по наблюдению другой многомерной случайной матрицы ξ в условиях, когда известна совместная плотность вероятности этих матриц.

Аналогично скалярному случаю [4] можно показать, что оптимальное решение является нерандомизированным и оптимальный предиктор $\Delta = \Delta(\xi)$ определяется из условия

$$r(\Delta) = \int_S W(\eta,\Delta) f(\eta/\xi) d\xi \to \min_{\Delta}, \qquad (2)$$

где $f(\eta/\xi) = \dfrac{f(\eta) f(\xi/\eta)}{\int_{-\infty}^{\infty} f(\eta) f(\xi/\eta) d\eta}$ — апостериорная плотность вероятности состояния η. Уравнение Эйлера для минимизации функционала (2) имеет вид

$$\frac{dr(\Delta)}{d\Delta} = 0\,.$$

При квадратичной функции потерь $W(\eta,\Delta) = {}^{0,p}(\eta-\Delta)^2$, где ${}^{0,p}(\eta-\Delta)^2$ — $(0,p)$-свернутый квадрат матрицы $(\eta-\Delta)$ [5, 6], получим

$$\frac{d}{d\Delta}\, {}^{0,p}(\eta-\Delta)^2 = \frac{d}{d\Delta}\, {}^{0,p}((\eta-\Delta)(\eta-\Delta)) =$$

$$= -{}^{0,p}(E(0,q)(\eta-\Delta))^{H_p} - {}^{0,p}((\eta-\Delta)E(0,q)) = -2(\eta-\Delta)\,,$$

$$\frac{dr(\Delta)}{d\Delta} = -2\int_S (\eta-\Delta) f(\eta/\xi) d\eta = 0\,.$$

Из последнего уравнения следует, что

$$\Delta = \int_S \eta f(\eta/\xi) d\eta = E(\eta/\xi)\,, \qquad (3)$$

т.е., как и в скалярном случае [4], оптимальный предиктор определяется как апостериорное среднее случайной многомерной матрицы η или как функция регрессии η на ξ.

Апостериорное среднее сравнительно легко можно получить для матриц ξ и η с совместным гауссовским распределением. В случае не гауссовского совместного распределения матриц η и ξ задача отыскания оптимального предиктора $\Delta(\xi)$ усложняется. Тогда при постановке задачи можно постулировать полиномиальный характер предиктора и отыскивать коэффициенты этого предиктора из условия минимума среднего риска. Такая задача сформулирована и решена в работе [3]. Постановка задачи и результаты ее решения приводятся в следующем разделе.

Полиномиальный степени m предиктор p-мерной матрицы $\eta=(\eta_{i_{(p)}})$, $i_{(p)}=$ $=(i_1,i_2,\ldots,i_p)$ по наблюдению q-мерной матрицы $\xi=(\xi_{j_{(q)}})$, $j_{(q)}=(j_1,j_2,\ldots$ $\ldots,j_q)$ имеет вид

$$\Delta=\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}\xi^k)=\sum_{k=0}^{m}{}^{0,kq}(\xi^k C_{(kq,p)}),\quad m=0,1,2,\ldots,\tag{4}$$

где $C_{(p,kq)}$ — $(p+kq)$-мерные матрицы коэффициентов,

$$C_{(p,kq)}=(c_{i_{(p)},\bar{j}_{(k)}}),\ i_{(p)}=(i_1,i_2,\ldots,i_p),\ \bar{j}_{(k)}=(j_{(q),0},j_{(q),1},\ldots,j_{(q),k}),$$

симметричные относительно q-мультииндексов $j_{(q),0},j_{(q),1},\ldots,j_{(q),k}$ и удовлетворяющие условиям

$$C_{(p,kq)}=C_{(kq,p)}^{H_{p+kq,kq}},\ C_{(kq,p)}=C_{(p,kq)}^{B_{p+kq,kq}},$$

$H_{p+kq,kq}$ и $B_{p+kq,kq}$ — подстановки транспонирования типа «вперед» и «назад» соответственно [5]. Обозначим

$$z=\eta-\Delta=\eta-\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}\xi^k)=\eta-\sum_{k=0}^{m}{}^{0,kq}(\xi^k C_{(kq,p)})\tag{5}$$

и сформулируем следующую оптимизационную задачу:

$$r=E({}^{0,p}(zz))\to\min_{C_{(p,0)},C_{(p,q)},\ldots,C_{(p,mq)}}.\tag{6}$$

Предиктор (4), коэффициенты которого определяются как решение задачи (6), будем называть наилучшим полиномиальным многомерно-матричным предиктором. Он минимизирует средний риск при квадратичной функции потерь $W(\eta,\Delta)={}^{0,p}(\eta-\Delta)^2=\sum_{i_{(p)}}(\eta_{i_{(p)}}-\Delta_{i_{(p)}})^2$. Матрицу

$$R=E({}^{0,0}(zz))=E(z^2)=(R_{i_{(p)},j_{(p)}})=E((\eta-\Delta)^2)\tag{7}$$

назовем дисперсионной матрицей среднего риска. Легко показать, что средний риск определяется как след дисперсионной матрицы риска:

$$r=\mathrm{tr}(R)=\sum_{i_{(p)}}(R_{i_{(p)},i_{(p)}}).$$

В [3] получено необходимое условие минимума среднего риска (6)

$$\sum_{\lambda=0}^{m}{}^{0,\lambda q}(C_{(p,\lambda q)}\nu_{\xi}^{(\lambda+s)})=\nu_{\eta\xi^s},\quad s=\overline{0,m},\tag{8}$$

которое представляет собой многомерно-матричную систему линейных алгебраических уравнений относительно коэффициентов $C_{(p,kq)}$ полиномиального предиктора (4). В (8) и ниже применяются следующие обозначения для моментов многомерных случайных матриц. Начальные и центральные моменты k-го порядка многомерной случайной матрицы φ обозначаются соответственно как

$$\nu_{\varphi}^{(k)}=E(\varphi^k),\quad \mu_{\varphi}^{(k)}=E((\varphi-E(\varphi))^k).$$

Смешанные начальные и центральные моменты двух многомерных случайных матриц φ и ψ обозначаются соответственно как

$$\nu_{\varphi\psi} = E(\varphi\psi)\,,\ \ \mu_{\varphi\psi} = E((\varphi - E(\varphi))(\psi - E(\psi)))\,.$$

Здесь E — символ математического ожидания, а произведения матриц понимаются как (0,0)-свернутые [5, 6]. В этих обозначениях справедливы следующие равенства:

$$\mu_{\xi^k\xi^l} = E((\xi^k - E(\xi^k))(\xi^l - E(\xi^l))) = \nu_{\xi}^{(k+l)} - \nu_{\xi}^{(k)}\nu_{\xi}^{(l)}, \tag{9}$$

$$\mu_{\xi^k\eta^l} = E((\xi^k - E(\xi^k))(\eta^l - E(\eta^l))) = \nu_{\xi^k\eta^l} - \nu_{\xi}^{(k)}\nu_{\eta}^{(l)}. \tag{10}$$

МИНИМАЛЬНОЕ ЗНАЧЕНИЕ СРЕДНЕГО РИСКА ДЛЯ ПОЛИНОМИАЛЬНОГО МНОГОМЕРНО-МАТРИЧНОГО ПРЕДИКТОРА

Теорема 1. Минимальное значение среднего риска r (6) для полиномиального многомерно-матричного предиктора (4) определяется одним из следующих выражений:

$$r_{\min} = \operatorname{tr}\left(\nu_{\eta}^{(2)} - \sum_{k=0}^{m}\sum_{l=0}^{m} {}^{0,kq}\left({}^{0,lq}\left(C_{(p,lq)}\nu_{\xi}^{(l+k)}\right)C_{(kq,p)}\right)\right), \tag{11}$$

$$r_{\min} = \operatorname{tr}\left(\nu_{\eta}^{(2)} - \sum_{l=0}^{m} {}^{0,lq}\left(\nu_{\eta\xi^l}C_{(lq,p)}\right)\right). \tag{12}$$

Доказательство. Поскольку средний риск определяется как след дисперсионной матрицы риска, будем искать дисперсионную матрицу риска

$$R = E({}^{0,0}(\eta - \Delta)^2) =$$

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

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

Используя указанные выше обозначения начальных моментов многомерных матриц ξ, η, имеем

$$R = \nu_{\eta}^{(2)} - 2\sum_{k=0}^{m} {}^{0,kq}(\nu_{\eta\xi^k}C_{(kq,p)}) + \sum_{k=0}^{m} {}^{0,kq}\left(\sum_{l=0}^{m} {}^{0,lq}(C_{(p,lq)}\nu_{\xi}^{(l+k)})C_{(kq,p)}\right). \tag{13}$$

Подставляя во второе слагаемое выражения (13) вместо моментов $\nu_{\eta\xi^k}$ сумму из левой части выражения (8) при $s = k$, $\lambda = l$, получим

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

$$+ \sum_{k=0}^{m}\sum_{l=0}^{m} {}^{0,kq}({}^{0,lq}(C_{(p,lq)}\nu_{\xi}^{(l+k)})C_{(kq,p)}) =$$

$$= \nu_{\eta}^{(2)} - \sum_{k=0}^{m}\sum_{l=0}^{m} {}^{0,kq}({}^{0,lq}(C_{(p,lq)}\nu_{\xi}^{(l+k)})C_{(kq,p)}). \tag{14}$$

Подставляя в третье слагаемое выражения (13) вместо суммы по переменной l правую часть выражения (8) при $s = k$, $\lambda = l$, получаем

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

$$= \nu_{\eta}^{(2)} - \sum_{k=0}^{m} {}^{0,kq}(\nu_{\eta\xi^k} C_{(kq,p)}). \tag{15}$$

Таким образом, имеем две формы оптимальной дисперсионной матрицы риска (14), (15) и средний риск (11), (12) как их след.

Теорема доказана.

Более экономичным для численных расчетов является выражение минимального риска (12). Формула (11) оказывается полезной в аналитических преобразованиях.

СВОЙСТВА НАИЛУЧШЕГО ПОЛИНОМИАЛЬНОГО МНОГОМЕРНО-МАТРИЧНОГО ПРЕДИКТОРА

Теорема 2. Среднее значение и дисперсионная матрица оптимального полиномиального многомерно-матричного предиктора Δ (4) определяются следующими выражениями:

$$E(\Delta) = \nu_{\Delta}^{(1)} = \nu_{\eta}^{(1)}, \tag{16}$$

$$\mu_{\Delta}^{(2)} = E((\Delta - E(\Delta))^2) = \sum_{k=1}^{m}\sum_{l=1}^{m} {}^{0,kq}({}^{0,lq}(C_{(p,lq)}\mu_{\xi^l\xi^k})C_{(kq,p)}), \tag{17}$$

где $C_{(p,lq)}$, $C_{(kq,p)}$ — оптимальные значения коэффициентов.

Доказательство. Коэффициенты $C_{(p,kq)}$ оптимального полиномиального многомерно-матричного предиктора Δ (4) определяются из системы многомерно-матричных линейных алгебраических уравнений (8). Эту систему можно решить методом исключения Гаусса. При этом коэффициент $C_{(p,0)}$ определяется на последнем шаге обратного хода метода Гаусса, т. е. из первого уравнения системы (8), и имеет вид

$$C_{(p,0)} = \nu_{\eta}^{(1)} - \sum_{k=1}^{m} {}^{0,kq}(C_{(p,kq)}\nu_{\xi}^{(k)}). \tag{18}$$

Представляя предиктор Δ (4) в виде

$$\Delta = C_{(p,0)} + \sum_{k=1}^{m} {}^{0,kq}(C_{(p,kq)}\xi^k) = \sum_{k=1}^{m} {}^{0,kq}(\xi^k C_{(kq,p)}) + C_{(0,p)}$$

и учитывая выражение (18), для оптимального полиномиального многомерно-матричного предиктора получаем следующее выражение:

$$\Delta = \nu_{\eta}^{(1)} + \sum_{k=1}^{m} {}^{0,kq}(C_{(p,kq)}(\xi^k - \nu_{\xi}^{(k)})) = \sum_{k=1}^{m} {}^{0,kq}((\xi^k - \nu_{\xi}^{(k)})C_{(kq,p)}) + \nu_{\eta}^{(1)}, \tag{19}$$

которое позволяет найти среднее значение оптимального полиномиального многомерно-матричного предиктора:

$$E(\Delta) = \nu_{\eta}^{(1)} + \sum_{k=1}^{m} {}^{0,kq}(C_{(p,kq)}\mu_{\xi^k}^{(1)}).$$

Если учесть, что $\mu^{(1)}_{\xi^k} = E(\xi^k - E(\xi^k)) = 0$, то получим $E(\Delta) = \nu^{(1)}_\eta = E(\eta)$. Это означает, что оптимальный многомерно-матричный полиномиальный предиктор Δ (19) является несмещенной оценкой многомерной матрицы η. С учетом несмещенности оптимального предиктора Δ из (19) получаем, что

$$\Delta - E(\Delta) = \sum_{k=1}^{m} {}^{0,kq}(C_{(p,kq)}(\xi^k - \nu^{(k)}_\xi)) = \sum_{k=1}^{m} {}^{0,kq}((\xi^k - \nu^{(k)}_\xi) C_{(kq,p)}),$$

откуда находим дисперсионную матрицу оптимального предиктора:

$$\mu^{(2)}_\Delta = E((\Delta - E(\Delta)^2)) = E\left(\sum_{k=1}^{m} {}^{0,kq}(C_{(p,kq)}(\xi^k - \nu^{(k)}_\xi)) \sum_{l=1}^{m} {}^{0,lq}((\xi^l - \nu^{(l)}_\xi) C_{(lq,p)}) \right).$$

Выполнив в этом выражении умножение и взяв математическое ожидание, получим выражение (17).

Теорема доказана.

Следствие. Матрица $R_{\min} = E(z^2)$ (14), (15) неотрицательно определенная.

Действительно, для оптимального предиктора в силу теоремы 2 $E(z) = E(\Delta - \eta) = 0$ и матрица $R_{\min} = E(z^2)$ (14) представляет собой дисперсионную матрицу матрицы $z = \Delta - \eta$, т.е. неотрицательно-определенную матрицу. $2p$-мерную матрицу назовем неотрицательно-определенной, если $(p,0,p)$-ассоциированная с ней матрица неотрицательно-определенная.

СВОЙСТВА ОПТИМАЛЬНОЙ ДИСПЕРСИОННОЙ МАТРИЦЫ РИСКА

Теорема 3. Оптимальную дисперсионную матрицу $R_{\min}$ (14) оптимального предиктора Δ представима в виде

$$R_{\min} = \mu^{(2)}_\eta - \mu^{(2)}_\Delta, \tag{20}$$

где $\mu^{(2)}_\eta$ — дисперсионная матрица матрицы η, $\mu^{(2)}_\Delta$ — дисперсионная матрица (17) оптимального полиномиального предиктора Δ.

Доказательство. Из формулы для полиномиального предиктора (4) получаем выражение для его начального момента второго порядка:

$$\nu^{(2)}_\Delta = E(\Delta^2) = E\left(\left(\sum_{l=0}^{m} {}^{0,lq}(C_{(p,lq)}\xi^l) \right) \left(\sum_{k=0}^{m} {}^{0,kq}(\xi^k C_{(kq,p)}) \right) \right) =$$

$$= \sum_{k=0}^{m}\sum_{l=0}^{m} {}^{0,kq}({}^{0,lq}(C_{(p,lq)}\nu^{(l+k)}_\xi) C_{(kq,p)}). \tag{21}$$

Обратимся к выражению для $R_{\min}$ (14). Учитыкая выражение (21), вместо (14) получим

$$R_{\min} = \nu^{(2)}_\eta - \nu^{(2)}_\Delta.$$

Поскольку $\nu^{(2)}_\eta = \mu^{(2)}_\eta + (\nu^{(1)}_\eta)^2$, $\nu^{(2)}_\Delta = \mu^{(2)}_\Delta + (\nu^{(1)}_\Delta)^2$, имеем

$$R_{\min} = \mu^{(2)}_\eta + (\nu^{(1)}_\eta)^2 - \mu^{(2)}_\Delta - (\nu^{(1)}_\Delta)^2.$$

Учитывая несмещенность оптимального полиномиального предиктора, т.е. свойство (16), получаем выражение (20).

Теорема 3 доказана.

Замечание. В силу равенства $\nu_\eta^{(1)} = \nu_\Delta^{(1)}$ матрицу $\mu_\Delta^{(2)}$ можно трактовать как дисперсионную матрицу оптимального полиномиального предиктора Δ относительно среднего значения оцениваемой матрицы η. Понятно также, что матрицу $R_{\min}$ можно трактовать как дисперсионную матрицу оптимального предиктора относительно матрицы η.

Следствие. Минимальное значение среднего риска r (6) для полиномиального многомерно-матричного предиктора (4) наряду с выражениями (11), (12) определяется также выражением $r_{\min} = \mathrm{tr}(\mu_\eta^{(2)} - \mu_\Delta^{(2)})$, где $\mu_\Delta^{(2)}$ имеет вид (17).

ЭФФЕКТИВНОСТЬ ОПТИМАЛЬНОГО ПОЛИНОМИАЛЬНОГО МНОГОМЕРНО-МАТРИЧНОГО ПРЕДИКТОРА

Теорема 3 позволяет ввести такое понятие, как эффективность полиномиального многомерно-матричного предиктора. Учитывая неотрицательную определенность матрицы $R_{\min}$, запишем неравенство $\mu_\eta^{(2)} - \mu_\Delta^{(2)} \geq 0$.

Поскольку $\mu_\eta^{(2)}$ — неотрицательно-определенная матрица, то после умножения левой части последнего неравенства на $(0, p)$-обратную матрицу ${}^{0,p}(\mu_\eta^{(2)})^{-1}$ получим также неотрицательно-определенную матрицу, т.е. $E(0, p) -$ $-{}^{0,p}({}^{0,p}(\mu_\eta^{(2)})^{-1} \mu_\Delta^{(2)}) \geq 0$, или

$$E(0, p) \geq {}^{0,p}\,({}^{0,p}(\mu_\eta^{(2)})^{-1} \mu_\Delta^{(2)}),$$

где $E(0, p)$ — $(0, p)$-единичная матрица. Отсюда следует, что

$$\det E(0, p) \geq \det{}^{0,p}({}^{0,p}(\mu_\eta^{(2)})^{-1} \mu_\Delta^{(2)}). \tag{22}$$

Определителем $2p$-мерной матрицы назовем определитель $(p, 0, p)$-ассоциированной с ней матрицы. Обозначим

$$e = \det{}^{0,p}({}^{0,p}(\mu_\eta^{(2)})^{-1}) \det(\mu_\Delta^{(2)}) \tag{23}$$

и назовем эффективностью оптимального полиномиального многомерно-матричного предиктора Δ по прогнозированию (оцениванию) матрицы η. Из (22) с учетом того, что $\det E(0, p) = 1$, получаем соотношение для эффективности: $0 \leq e \leq 1$.

МИНИМАЛЬНЫЙ СРЕДНИЙ РИСК И ЭФФЕКТИВНОСТЬ ПОСТОЯННОГО ПРЕДИКТОРА

Для оптимального постоянного многомерно-матричного предиктора $(m = 0)$ $\Delta = C_{(p,0)}$ в [3] получено выражение $C_{(p,0)} = \nu_\eta^{(1)}$. В этом случае в соответствии с формулой (12) минимальный средний риск равен

$$r_{\min} = \mathrm{tr}(\nu_\eta^{(2)} - {}^{0,0}(C_{(p,0)} \nu_\eta^{(1)})) = \mathrm{tr}(\nu_\eta^{(2)} - \nu_\eta^{(1)} \nu_\eta^{(1)}) = \mathrm{tr}(\mu_\eta^{(2)}).$$

Из формулы (17) следует, что $\mu_\Delta^{(2)} = 0$, так что эффективность e (23) постоянного предиктора равна нулю. Это понятно, поскольку решение о матрице η принимается без наблюдения матрицы ξ, что не уменьшает неопределенность матрицы η.

МИНИМАЛЬНЫЙ СРЕДНИЙ РИСК И ЭФФЕКТИВНОСТЬ ЛИНЕЙНОГО ПРЕДИКТОРА

Для наилучшего линейного многомерно-матричного предиктора ($m=1$)

$$\Delta = C_{(p,0)} +^{0,q} (C_{(p,q)}\xi)$$

в [3] получены следующие значения коэффициентов:

$$C_{(p,q)} =^{0,q} (\mu_{\eta\xi}{}^{0,q} (\mu_{\xi}^{(2)})^{-1}), \tag{24}$$

$$C_{(p,0)} = \nu_{\eta}^{(1)} -^{0,q} (C_{(p,q)}\nu_{\xi}^{(1)}).$$

В соответствии с формулой (12) минимальный средний риск равен

$$r_{\min} = \mathrm{tr}(\nu_{\eta}^{(2)} - C_{(p,0)}\nu_{\eta}^{(1)} -^{0,q} (C_{(p,q)}\nu_{\xi\eta})) =$$

$$= \mathrm{tr}(\nu_{\eta}^{(2)} - ((\nu_{\eta}^{(1)} -^{0,q} (C_{(p,q)}\nu_{\xi}^{(2)}))\nu_{\eta}^{(1)}) -^{0,q} (C_{(p,q)}\nu_{\xi\eta})) =$$

$$= \mathrm{tr}(\nu_{\eta}^{(2)} - \nu_{\eta}^{(1)}\nu_{\eta}^{(1)} +^{0,q} (C_{(p,q)}(\nu_{\xi}^{(1)}\nu_{\eta}^{(1)})) -^{0,q} (C_{(p,q)}\nu_{\xi\eta})) =$$

$$= \mathrm{tr}(\mu_{\eta}^{(2)} -^{0,q} (C_{(p,q)}\mu_{\xi\eta})) = \mathrm{tr}\,(\mu_{\eta}^{(2)} -^{0,q} ({}^{0,q}(\mu_{\eta\xi}{}^{0,q} (\mu_{\xi}^{(2)})^{-1})\mu_{\xi\eta})).$$

Теперь получим формулы для расчета эффективности оптимального линейного многомерно-матричного предиктора. Из формулы (17) следует

$$\mu_{\Delta}^{(2)} =^{0,q} ({}^{0,q}(C_{(p,q)}\mu_{\xi}^{(2)})C_{(q,p)}).$$

Подставляя значение коэффициента $C_{(p,q)}$ (24), получаем необходимую для расчета эффективности (23) матрицу $\mu_{\Delta}^{(2)}$:

$$\mu_{\Delta}^{(2)} =^{0,q} (\mu_{\eta\xi}{}^{0,q} ({}^{0,q}(\mu_{\xi}^{(2)})^{-1}\mu_{\xi\eta})).$$

МИНИМАЛЬНЫЙ СРЕДНИЙ РИСК И ЭФФЕКТИВНОСТЬ КВАДРАТИЧНОГО ПРЕДИКТОРА

В [3] показано, что коэффициенты наилучшего многомерно-матричного квадратичного предиктора

$$\Delta = C_{(p,0)} +^{0,q} (C_{(p,q)}\xi) +^{0,2q} (C_{(p,2q)}\xi^2)$$

рассчитываются по следующим формулам в порядке их написания:

$$B_{(p,2q)} = \mu_{\eta\xi^2} -^{0,q} ({}^{0,q}(\mu_{\eta\xi}{}^{0,q} (\mu_{\xi}^{(2)})^{-1})\mu_{\xi\xi^2}),$$

$$A_{(2q,2q)} = (\mu_{\xi^2}^{(2)} -^{0,q} ({}^{0,q}(\mu_{\xi^2\xi}{}^{0,q} (\mu_{\xi}^{(2)})^{-1})\mu_{\xi\xi^2})),$$

$$C_{(p,2q)} =^{0,2q} (B_{(p,2q)}{}^{0,2q} A_{(2q,2q)}^{-1}),$$

$$C_{(p,q)} =^{0,q} (\mu_{\eta\xi}{}^{0,q} (\mu_{\xi}^{(2)})^{-1}) -^{0,q} ({}^{0,2q}(C_{(p,2q)}\mu_{\xi^2\xi})^{0,q} (\mu_{\xi}^{(2)})^{-1}),$$

$$C_{(p,0)} = \nu_{\eta}^{(1)} -^{0,q} (C_{(p,q)}\nu_{\xi}^{(1)}) -^{0,2q} (C_{(p,2q)}\nu_{\xi}^{(2)}).$$

Заменяя центральные моменты в данных выражениях начальными в соответствии с равенствами (9), (10), получаем следующее выражение для минимального среднего риска оптимального квадратичного многомерно-матричного предиктора:

$$r_{\min} = \mathrm{tr}(\nu_{\eta}^{(2)} -^{0,0} (\nu_{\eta}^{(1)} C_{(0,p)}) -^{0,q} (\nu_{\eta\xi} C_{(q,p)}) -^{0,2q} (\nu_{\eta\xi^2} C_{(2q,p)})) =$$

$$= \mathrm{tr}(\nu_{\eta}^{(2)} -^{0,0} (C_{(p,0)} \nu_{\eta}^{(1)}) -^{0,q} (C_{(p,q)} \nu_{\xi\eta}) -^{0,2q} (C_{(p,2q)} \nu_{\xi^2\eta})) .$$

Из формулы (17) получаем матрицу $\mu_{\Delta}^{(2)}$, необходимую для расчета эффективности оптимального предиктора в этом случае:

$$\mu_{\Delta}^{(2)} =^{0,q} ({}^{0,q}(C_{(p,q)} \mu_{\xi}^{(2)}) C_{(q,p)}) +^{0,q} ({}^{0,2q}(C_{(p,2q)} \mu_{\xi^2\xi}) C_{(q,p)}) +$$

$$+^{0,q} (C_{(p,q)} {}^{0,2q} (\mu_{\xi\xi^2} C_{(2q,p)})) +^{0,2q} ({}^{0,2q}(C_{(p,2q)} \mu_{\xi^2}^{(2)}) C_{(2q,p)}) .$$

Пример. Рассмотрим две скалярные случайные величины: η и ξ с возможными значениями $\eta = (1, 2, 3)$, $\xi = (2, 3, 4)$ и двумя матрицами совместных вероятностей:

$$P_1 = (p_{1,i,j}) = P(\eta = \eta_i, \xi = \xi_j) = \begin{pmatrix} 0{,}1 & 0{,}1 & 0{,}06 \\ 0{,}03 & 0{,}13 & 0{,}22 \\ 0{,}05 & 0{,}08 & 0{,}23 \end{pmatrix},$$

$$P_2 = (p_{2,i,j}) = P(\eta = \eta_i, \xi = \xi_j) = \begin{pmatrix} 0{,}1 & 0{,}5 & 0 \\ 0{,}05 & 0{,}05 & 0 \\ 0 & 0 & 0{,}3 \end{pmatrix}.$$

Для распределения с матрицей вероятностей P_1 (для модели 1) моменты равны:

$$\nu_{\xi}^{(1)} = 3{,}33,\ \mu_{\xi}^{(2)} = 0{,}5811,\ \nu_{\xi}^{(2)} = 11{,}67,\ \mu_{\xi\xi^2} = 3{,}5889,$$

$$\mu_{\xi^2}^{(2)} = 22{,}3611,\ \nu_{\eta}^{(1)} = 2{,}1,\ \mu_{\eta}^{(2)} = 0{,}61,\ \mu_{\eta\xi} = 0{,}187.$$

Для распределения с матрицей вероятностей P_2 (для модели 2) имеем:

$$\nu_{\xi}^{(1)} = 3{,}15,\ \mu_{\xi}^{(2)} = 0{,}4275,\ \nu_{\xi}^{(2)} = 10{,}35,\ \mu_{\xi\xi^2} = 2{,}6475,$$

$$\mu_{\xi^2}^{(2)} = 16{,}6275,\ \nu_{\eta}^{(1)} = 1{,}7,\ \mu_{\eta}^{(2)} = 0{,}81,\ \mu_{\eta\xi} = 0{,}445 .$$

Для этих двух моделей данных рассчитаны минимальный средний риск $r_{\min}$ и эффективность e постоянного, линейного и квадратичного оптимальных предикторов. Результаты расчетов приведены в табл. 1.

Из табл. 1 видно, что как по минимальному значению среднего риска, так и по эффективности наиболее предпочтительным для обеих моделей является квадратичный предиктор. Постоянный предиктор имеет нулевую эффективность для обеих моделей, а линейный и квадратичный более эффективны для второй модели. Вместе с тем для модели 1 эффективность линейного и квадратичного оптимального предикторов весьма мала, причем переход от линейного предиктора к квадратичному не способствует существенному увеличению эффективности. В то же время для модели 2 эффективность линейного предиктора достаточно высока, а использование квадратичного предиктора существенно увеличивает эффективность прогнозирования по сравнению с линейным предиктором.

Таблица 1

Матрицы вероятностей	Постоянный предиктор	Линейный предиктор	Квадратичный предиктор
P_1 (модуль 1)	Средний риск, $r_{\min} = 0{,}6100$	Средний риск, $r_{\min} = 0{,}5498$	Средний риск, $r_{\min} = 0{,}5482$
	Эффективность, $e = 0$	Эффективность, $e = 0{,}0987$	Эффективность, $e = 0{,}1014$
P_2 (модуль 2)	Средний риск, $r_{\min} = 0{,}8100$	Средний риск, $r_{\min} = 0{,}3468$	Средний риск, $r_{\min} = 0{,}0788$
	Эффективность, $e = 0$	Эффективность, $e = 0{,}5719$	Эффективность, $e = 0{,}9027$

ЗАКЛЮЧЕНИЕ

Таким образом, в настоящей работе рассмотрена задача теории оптимальных статистических решений с наблюдениями в случае многомерно-матричных состояний, наблюдений и решений. При наличии трудностей в отыскании оптимального решения (оптимального предиктора), которые неизбежно возникают в случае негауссовских распределений, на этапе постановки задачи предлагается ограничиться полиномиальным предиктором и отыскивать его коэффициенты из условия минимума среднего риска. В более ранней работе [3] получены выражения и алгоритмы оптимальных полиномиальных многомерно-матричных предикторов. В данной статье решены вопросы, связанные с минимальным значением среднего риска оптимального полиномиального многомерно-матричного предиктора и его эффективности. Полученные в статье алгоритмы для расчета этих характеристик позволяют сравнивать различные предикторы между собой, оценивать их возможности и выбирать наиболее приемлемые для использования.

СПИСОК ЛИТЕРАТУРЫ

1. Mukha V.S., Korchyts K.S. Statistical weather forecasting // Computer Data Analysis and Modeling. Robustness and Computer Intensive Methods. Proceeding of the Seventh Intern. Conf. (Minsk, September, 6–10, 2004). — 2004. — **2**. — P. 142–145.
2. Муха В.С. Статистическое векторное прогнозирование количественных характеристик погоды // Информационные системы и технологии (IST'2004). Материалы междунар. конф., Минск, Беларусь, 8–10 ноября 2004. — 2004. — Ч. 2. — С. 195–200.
3. Муха В.С. Наилучшая полиномиальная многомерно-матричная регрессия // Кибернетика и системный анализ. — 2007. — № 3. — С. 138–143.
4. Муха В.С. Статистические методы обработки данных: учеб. пособие. — Минск: Изд. центр БГУ, 2009. — 183 с.
5. Муха В.С. Анализ многомерных данных. — Минск: Технопринт, 2004. — 368 с.
6. Соколов Н.П. Введение в теорию многомерных матриц. — Киев: Наук. думка, 1971. — 167 с.

Поступила 17.03.2009