

УДК 519.7:007.52

Н. С. ЛЕСНАЯ, В. Б. РЕПКА

ОБ ОДНОМ ПОДХОДЕ К ВЫБОРУ ЭФФЕКТИВНОГО МЕТОДА СМЕЩЕННОГО ОЦЕНИВАНИЯ ПАРАМЕТРОВ МОДЕЛЕЙ КВАЗИСТАЦИОНАРНЫХ ПРОЦЕССОВ С ПРИМЕНЕНИЕМ АППАРАТА НЕЙРОННЫХ СЕТЕЙ

Выбор метода построения математических моделей процессов и объектов на сегодняшний день является достаточно сложной задачей, формализация которой связана с трудностями выбора процедуры оценивания коэффициентов модели, а также с анализом теоретических предпосылок и предположений в зависимости статистических характеристик исходных данных. Особенно сложно выбрать метод оценивания коэффициентов модели в условиях, когда исходные данные подвержены влиянию линейной зависимости между независимыми векторами. Для решения задач данного класса разработаны и успешно применяются методы смещенного оценивания, которые характеризуются в первую очередь высокой эффективностью и вычислительной простотой, но при этом имеют множество различных подходов к выбору их фактора деформации с пересекающимися областями эффективности, что затрудняет выбор оптимального метода с точки зрения точности, сложности и быстродействия.

На основе проведенного анализа существующих подходов к выбору наиболее эффективного метода смещенного оценивания параметров модели, в основе которых лежат математические модели оценки эффективности методов смещенного оценивания, анализ обобщенных критериев, основанных на мере рассеивания и индексе ошибки относительных среднеквадратичных критериев и др. [1-3] предлагается иной подход к решению данной проблемы. Указанные подходы подразумевают автоматический выбор метода оценивания и ориентированы на профессионализм эксперта, что значительно сужает область их применения. Интеллектуализация этого процесса позволит не только расширить область применения существующих подходов, повысить эффективность принятия решений в условиях неопределенности, а также упростить выбор метода построения математической модели процесса при высокой коррелируемости исходной информации.

Рассмотрим некоторое множество методов смещенного оценивания (табл. 1), принадлежащих к 4 типам: класс обычных гребневых оценок ($E7, R_0, H$), сжатых оценок (JS), обобщенных оценок ($E1, E2m, E4_1, E4_3$) и оценок дробного ранга (FRN, ST, SG).

Эффективность указанных методов смещенного оценивания зависит от выбора фактора деформации k , σ^2 – дисперсии помех, уровня мультиколлинеарности исходных данных $\alpha = \frac{\delta}{p}$, где

$$\delta = \frac{(p-1)}{1-r_{jl}} + \frac{1}{[1+(p-1)r_{jl}]}, r_{jl} - \text{коэффициент корреляции между } j \text{ и } l \text{ входными переменными;}$$

p – количество входных переменных.

Оценку эффективности данных методов произведем по:

– критерию относительной среднеквадратичной погрешности, характеризующей эффективность метода относительно МНК:

$$E(L_1^2) = \frac{E[(\hat{\beta} - \beta)'(\hat{\beta} - \beta)]}{E[(\beta^* - \beta)'(\beta^* - \beta)]},$$

где $\hat{\beta}^*$ и β^* – оценки и истинное значение коэффициентов по методу МНК; $\hat{\beta}$ и β – оценки и истинное значение коэффициентов по методу смещенного оценивания;

Таблица 1

| Шифр метода | Название алгоритмов |
|---|---|
| Алгоритмы обычного гребневого оценивания | |
| E7 | Итеративная оценка с выбором фактора деформации $(k_{j-1} - k_j) / k_j \leq \delta$ |
| R ₀ | Метод McDonald с выбором фактора деформации так, чтобы $\hat{\beta}_k \hat{\beta} = W$, где $W = \hat{\beta}' \hat{\beta} - \sigma^2 \sum_{i=1}^p \lambda_j^{-1}$. |
| H | Метод Hoerl с выбором фактора деформации $k_j = \sigma^2 / \beta_j^2$ |
| Алгоритмы обобщенного гребневого оценивания | |
| E1 | Метод Hammerle с выбором фактора деформации k – диагональная матрица с неотрицательными диагональными элементами $k_j = \frac{\sigma^2}{\lambda_j \beta_j^2}$. |
| E2m | Метод Hammerle с наложением ограничений на возрастание среднеквадратичной ошибки на m % |
| E4 ₁ | Метод Аллена с выбором фактора деформации $\hat{M}_j = \beta_j^2 [(v_j - l + e_{j(0)})^2]$. |
| E4 ₃ | Метод Аллена с выбором фактора деформации $\hat{M}_j = \beta_j^2 [(v_j l)^2]$. |
| Алгоритмы сжатого оценивания | |
| JS | Метод James and Stein с выбором фактора деформации $a(\lambda) = D^{-1/2} p' (1 - \frac{cQ(y)}{\ g\ ^2}) g$ |
| Алгоритмы дробного ранга | |
| FRN | Метод Фейербразена в котором начальное значение ранга r_0 выбирается из условия $\max_r \sum_{j=r+i}^p (1 - \tau_i^2) \lambda_j$, если среди этих сумм есть положительные, иначе $r_0 = p-1$. |
| ST | Метод Стейна с выбором фактора деформации $l = \max \left(0, 1 - \frac{p-2}{n-p+2} \cdot \frac{1-R^2}{R^2} \right)$, где R^2 – коэффициент множественной корреляции. |
| SG | Метод Смита-Гольдстейна с выбором фактора сжатия вида: $l = 1 - \frac{c}{\sum_{j=1}^p z_j^2}$, где z – определяется при помощи $z = T'Y$; a с выбирается из условия $\lambda_{\min} > \frac{c+4}{2 \sum_{i=1}^p \lambda_i^{-2}}$. |

– по критерию, характеризующему дисперсию оценок коэффициентов модели

$$E(L_2^2) = E[(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))'];$$

– по критерию, характеризующему смещение оценок относительно моделируемых истинных

значений оценок $E(L_3^2) = (E(\hat{\beta}) - \beta)'(E(\hat{\beta}) - \beta)$;

– по критерию среднеквадратичной ошибки модели $E(L_4^2) = E \left[\left(\hat{\beta} - \beta \right)' \left(\hat{\beta} - \beta \right) \right]$;

– по критерию максимальной абсолютной координатной ошибки метода

$$E(L_5^2) = \max_j \left| \hat{\beta}_j - \beta_j \right|^2 ;$$

– по критерию максимальной координатной относительной ошибки метода

$$E(L_6^2) = \max_j \left| \frac{\hat{\beta}_j - \beta_j}{\beta_j} \right|^2 .$$

Рассмотрим задачу классификации методов смещенного оценивания по указанному набору характеристик, используя аппарат нейронных сетей (НС).

При практической работе с НС, как правило, приходится экспериментировать с большим числом различных сетей, порой обучая каждую из них по несколько раз и сравнивая полученные результаты. Главным показателем качества результата является контрольная ошибка [4]. При этом, в соответствии с общенаучным принципом, согласно которому при прочих равных следует предпочесть более простую модель, из двух сетей с приблизительно равными ошибками контроля имеет смысл выбрать ту, которая меньше.

В связи с этим, была промоделирована работа следующих типов НС, применяемых на сегодняшний день для решения задачи классификации с соответствующими им алгоритмами обучения: многослойный перцептрон, сеть встречного распространения, вероятностная нейронная сеть (ВНС), сеть радиальной базисной функции, самоорганизующиеся карты признаков Кохонена (или НС Кохонена). Наилучший результат по показателю контрольной ошибки с поправкой на размер сети был получен вероятностной НС.

Рассмотрим решение задачи классификации выше указанных методов смещенного оценивания коэффициентов модели, используя ВНС. Этот тип сети имеет алгоритм обучения "с учителем". Для ВНС необходима обучающая выборка, которая включает объекты, имеющие набор характеристик и известный класс принадлежности каждого объекта к одному из классов. Вид НС, имеющей 11 входных характеристик (X_1-X_{11}) и 11 классов выше указанных алгоритмов смещенного оценивания представлен на рис. 1.

Входные характеристики методов смещенного оценивания:

X_1 – название класса методов смещенного оценивания, к которому принадлежит алгоритм;

X_2 – возможность нахождения оптимального значения параметра деформации, минимизирующего квадратичное отклонение;

X_3 – достижение минимума дисперсии и смещения оцениваемых коэффициентов модели при высокой степени мультиколлинеарности ($\alpha = 1, 2, 5, 10$) входных данных процесса;

X_4 – достижение минимума дисперсии и смещения коэффициентов модели по алгоритму при средней степени мультиколлинеарности ($\alpha = 50$) входных данных процесса;

X_5 – достижение минимума дисперсии и смещения коэффициентов модели по алгоритму при низкой степени мультиколлинеарности ($\alpha = 100$ и выше) входных данных процесса;

X_6 – достижение минимума дисперсии и смещения коэффициентов модели по алгоритму при высокой степени дисперсии помех ($\alpha = 1, 2, 5, 10$) входных данных процесса;

X_7 – достижение минимума дисперсии и смещения коэффициентов модели по алгоритму при средней степени дисперсии помех ($\alpha = 50$) входных данных процесса;

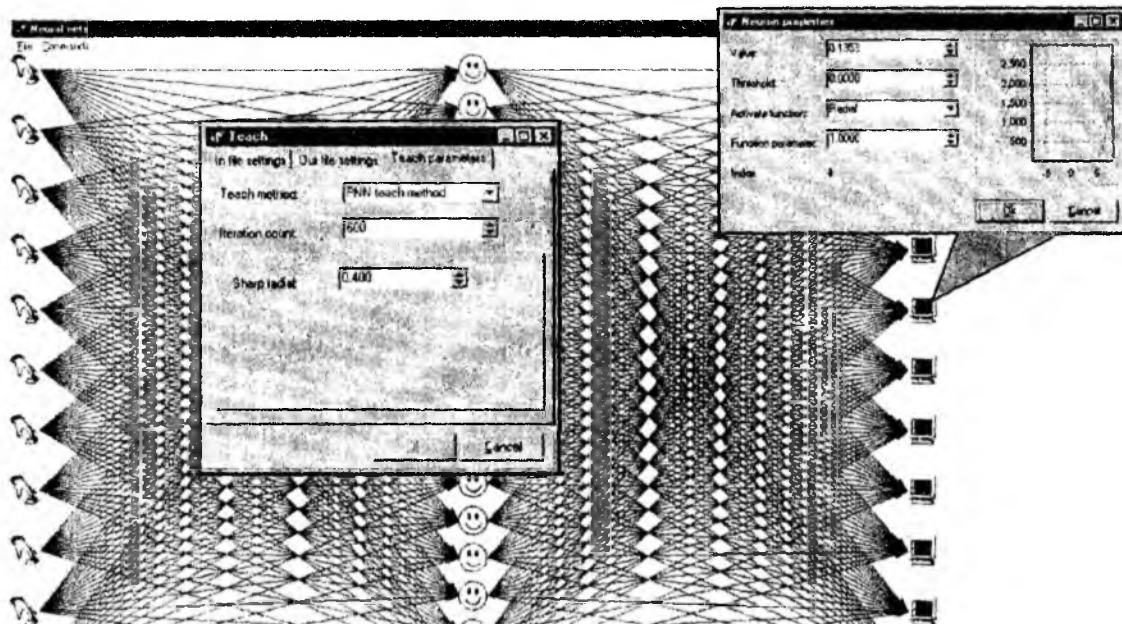


Рис. 1

X_8 – достижение минимума дисперсии и смещения коэффициентов модели по алгоритму при низкой степени дисперсии помех ($\alpha = 100$ и выше) входных данных процесса;

X_9 – для итерационных и последовательных методов смещенного оценивания показатель, характеризующий степень сходимости алгоритма;

X_{10} – точность метода, характеризующая совокупностью описанных выше критериев;

X_{11} – популярность использования метода (вычислительная сложность, наличие в ППП, частота встречаемости в литературе).

После обучения и тестирования сети контрольная ошибка составила всего 3 %, то есть, некоторые алгоритмы были отнесены не к "своим" классам. Например, алгоритм R_0 был отнесен к алгоритму E7. Проанализировав полученную ситуацию, можно заключить, что данные алгоритмы можно объединить в один класс, поскольку они имеют сходные характеристики и для исследователя не существует принципиальной разницы, каким из этих алгоритмов воспользоваться.

Таким образом, с помощью вероятностной НС была проверена возможность классификации указанных методов по приведенным отличительным характеристикам. Но поскольку многие методы смещенного оценивания можно объединить в гораздо меньшее количество классов, предлагается использование иного типа НС, которая бы дала возможность по имеющимся входным характеристикам самостоятельно формировать классы, выявляя саму структуру данных, в частности НС Кохонена.

На рис. 2 представлен фрагмент НС Кохонена, которая имеет 11 входных нейронов по количеству входных характеристик. Эти характеристики – те же, что и в вероятностной НС. Для обучения были применены метод Гаусса и метод выпуклой комбинации [5]. После настройки параметров НС и ее обучения на основании выше указанных характеристик (X_1 - X_{11}) сеть Кохонена объединила методы смещенного оценивания в 4 класса: первый – E7, R_0 , E1; второй – E2M, E4, JS, FRN, ST, SG; третий – H; четвертый – E43.

Контрольная ошибка данной сети составила также 3 %. Проанализировав результаты, выданные НС Кохонена, можно сделать заключения, что по данному набору входных характеристик методы смещенного оценивания можно разделить на четыре непересекающихся класса. Для дальнейшей классификации описываемых методов смещенного оценивания необходим иной набор входных характеристик, который бы включал такие особенности каждого метода, на основании которых можно было четко разграничить каждый отдельный метод.

На основании проведенных исследований можно сделать вывод, что качество классификации методов смещенного оценивания параметров модели существенно зависит от многих факторов, как

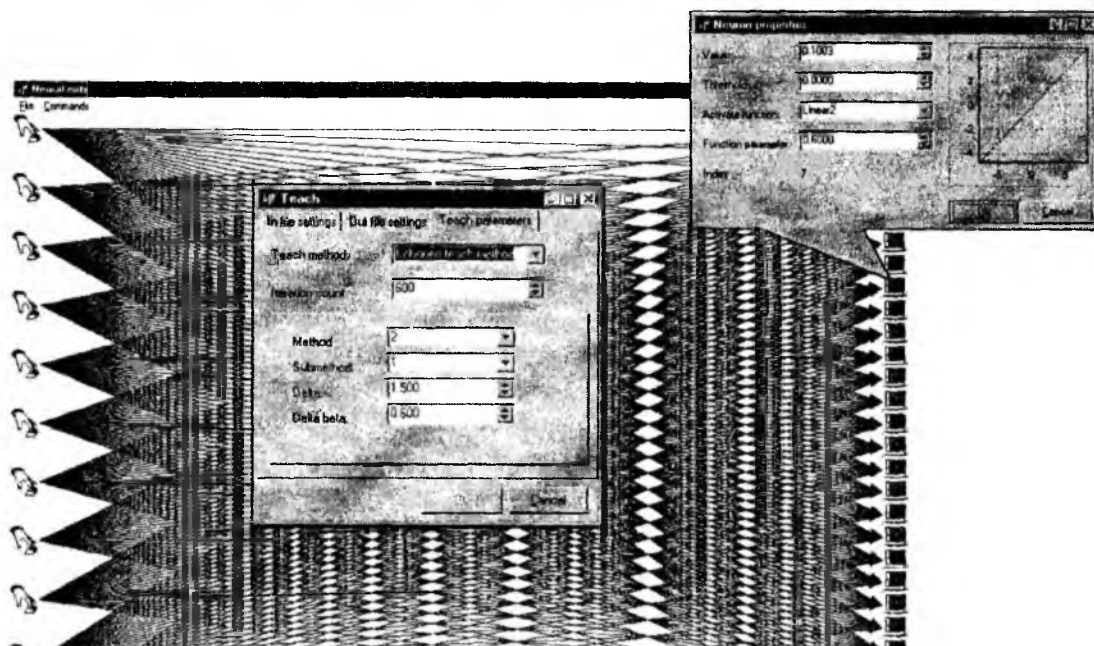


Рис. 2

то: выбора типа нейронной сети, способа ее обучения, качества и объема обучающей выборки. В этой связи промоделирована работа целого ряда архитектур НС, на основании минимальной контрольной ошибки выбрана вероятностная НС, для которой была составлена в качестве «учителя» обучающая выборка, включающая приведенные в статье одиннадцать характеристик и соответствующие этим характеристикам методы смещенного оценивания. А также НС Кохонена, которая самостоятельно выявила закономерности в предъявленных ей данных и сформировала четыре выходных класса. Таким образом, можно заключить, что предложенный нейросетевой подход к выбору наиболее эффективного метода смещенного оценивания параметров модели позволит упростить и ускорить процесс построения математической модели процесса при высокой коррелируемости исходной информации.

Список литературы: 1. Hoerl A.E., Kennard R.W. Ridge regression. Encyclopedia of statistical sciences. Vol. 8. New York: Willey, 1988. 2. Шамина Т.Б. Сравнительный анализ качественных показателей методов смещенного оценивания // Системный анализ, управление и информационные технологии: Вестник Харьковского государственного политехнического университета. 1999. №71. С.181-185. 3. Spiess M., Hamerle A. A comparison of different methods for the estimation of regression models with correlated binary responses // Computational Statistics and Data Analysis. 2000. Vol. 33, №4. P. 439-455. 4. Горбань А., Россиев Д. Нейронные сети на персональном компьютере. Новосибирск: Наука. 1996, 276 с. 5. Kohonen T. The Self-Organizing Maps (2-nd edition). New-York: Springer, 1997.

Поступила в редколлегию 8.06.2000

