ДОКЛАДЫ АН ВШ РФ
2017 июль–сентябрь № 3 (36)

ТЕХНИЧЕСКИЕ НАУКИ

УДК 519.23

НАСТРОЙКА ПАРАМЕТРОВ ЯДЕРНЫХ ФУНКЦИЙ В МЕТОДЕ LS-SVM С ИСПОЛЬЗОВАНИЕМ ВНЕШНИХ КРИТЕРИЕВ КАЧЕСТВА МОДЕЛЕЙ

А.П. Суходолов[1], А.А. Попов[2], Ш.А. Бобоев[2],
[1]*Байкальский государственный университет*
[2]*Новосибирский государственный технический университет*

В работе рассматривается задача восстановления регрессионной зависимости по методу опорных векторов с квадратичной функцией потерь (LS-SVM). Данный метод относится к классу ядерных методов. Для настройки ряда внутренних параметров алгоритма LS-SVM обсуждается проблема использования внешних критериев качества моделей. Приведены различные критерии селекции моделей, которые основываются на разбиении выборки на обучающую и тестовую части. Проблема разбиения выборки на тестовую и обучающую части с использованием метода *D*-оптимального планирования эксперимента подробно рассмотрена для случая линейных параметрических регрессионных моделей. Данный метод получения тестовой выборки предложено использовать для метода LS-SVM. Приводится последовательный алгоритм получения обучающей и тестовой частей выборки наблюдений применительно к методу LS-SVM. Для использования критериев в симметричной форме предлагается алгоритм построения бипланов. Приводятся результаты вычислительного эксперимента по анализу возможности использования трех внешних критериев для подбора масштаба гауссовой ядерной функции. В качестве внешних критериев использовались критерий перекрестной проверки, критерий регулярности и критерий стабильности. Параметр масштаба ядерной функции подбирался по минимуму внешнего критерия качества. Окончательно точность получаемых решений проверялась по среднеквадратичной ошибке. Вычислительный эксперимент проводился на модельных данных. В качестве модели, порождающей данные, была выбрана нелинейная зависимость от входного фактора. Дисперсия помехи (уровень шума) определялся в процентах от мощности сигнала. Результаты отдельных проведенных вычислительных экспериментов приведены в таблицах и рисунках. По результатам проведенных вычислительных экспериментов делаются выводы о том, что эффективность использования критерия стабильности в целях получения решения с малой среднеквадратичной ошибкой, как правило, выше, чем при использовании критерия регулярности. Эффективность критерия перекрестной проверки выше эффективности критериев регулярности и стабильности в условиях повышенного шума и использования тестовых выборок малого относительного объема.

Ключевые слова: регрессия, метод LS-SVM, квадратичная функция потерь, тестовая выборка, обучающая выборка, оптимальное планирование эксперимента, *D*-оптимальный план, критерий регулярности, критерий стабильности, критерий скользящего контроля, ядерная функция, среднеквадратичная ошибка.

DOI: 10.17212/1727-2769-2017-3-88-104

Введение

Метод опорных векторов с квадратичной функцией потерь (LS-SVM) является модификацией алгоритма опорных векторов (SVM) с функцией нечувствительности Вапника как с линейными, так и нелинейными ядерными функциями [1–3]. Он является один из наиболее перспективных алгоритмов построения регрессии. Одним из важных этапов построения регрессии с использованием метода опорных векторов является настройка его ряда внутренних параметров. При использовании произвольных значений параметров алгоритма опорных векторов качество работы алгоритма может существенно варьироваться. В работе [4] предлагается эвристический алгоритм априорного выбора параметров алгоритма на основе характеристик имеющейся выборки данных. Однако в общем случае для получения более качественных решений необходимо решать задачу выбора оптимальных значений параметров алгоритма. Ключевым моментом в решении задачи настройки параметров алгоритма опорных векторов, является выбор критерия качества получаемых решений. Простым и в то же время эффективным подходом к выбору модели является метод скользящего контроля (cross-validation, CV) [5–7]. При этом для подбора оптимальных параметров алгоритма используются различные варианты поиска решений на сетке их значений [8, 9]. В целях уменьшения вычислительных затрат вместо критерия CV применяют неполные варианты обобщенного CV – так называемые *K*-FOLD CV *(K-Fold Cross Validation)*. В этом случае исходная выборка разбивается некоторое количество раз на обучающую и контрольную объемом в *K* наблюдений с усреднением результатов [10, 11].

Задача подбора внутренних параметров алгоритма LS SVM по сути относится к проблеме борьбы с эффектом переобучения, которой постоянно уделяют повышенное внимание [12–16]. Применительно к линейному параметрическому регрессионному моделированию эта проблема может быть решена в рамках получения так называемых моделей оптимальной сложности [17]. Известным и активно развиваемым подходом для выбора линейных параметрических моделей оптимальной сложности является использование так называемых внешних критериев. В нашем случае в качестве таковых могут быть использованы различные варианты критериев, связанных с точностью прогноза на тестовой выборке. В данной работе исследуется возможность использования внешних критериев при разбиении выборки на обучающую и тестовую части с привлечением методов оптимального планирования эксперимента для решения задачи выбора параметра гауссовой ядерной функции в алгоритме LS SVM.

1. Внешние критерии селекции моделей

При построении моделей, описывающих поведение отклика от действующих факторов, главной задачей является определение структуры модели, поскольку, как правило, она априори не известна. Исследователь сталкивается с проблемой выбора структуры модели. Для решения этой задачи назначаются определенные критерии «качества», которым должна удовлетворять искомая модель. Будем в дальнейшем называть их критериями селекции моделей. Перечень используемых критериев селекции достаточно широк и подробно представлен в обзорах [18–21].

Критерии селекции моделей можно поделить на две группы: критерии, использующие всю выборку данных, и критерии, основанные на разбиении выборки на части.

Критерии, основанные на разбиении выборки на части.

Пусть модель объекта подчиняется следующему уравнению наблюдения:

$$Y = \widehat{Y} + \varepsilon = \widehat{X}\widehat{\beta} + \varepsilon , \tag{1.1}$$

где $\widehat{Y} - (n \times 1)$ – вектор ненаблюдаемого незашумленного выхода объекта; $\widehat{X} - (n \times m)$ – расширенная матрица плана, соответствующая истинному набору регрессоров $\widehat{x}_1, ..., \widehat{x}_m$; $\varepsilon - (n \times 1)$ – вектор ненаблюдаемых случайных ошибок измерения, относительно которых выполнены предположения $E(\varepsilon) = 0_n$; $E(\varepsilon\varepsilon^T) = \sigma^2 I_n$, где 0_n – вектор, состоящий из нулей, σ^2 – неизвестная дисперсия наблюдения, I_n – единичная матрица размера n. Набор регрессоров $\widehat{x}_1, ..., \widehat{x}_m$

образует множество $\widehat{X}$, о котором известно, что $\widehat{X} \subset \Re$, где $\Re$ – некоторое расширенное множество регрессоров. Пусть в результате наблюдения объекта получена $Z-(n\times p)$ – расширенная матрица плана из n наблюдений над p регрессорами из $\Re$ и требуется определить множество $\widehat{X}$ и получить оценку параметров $\widehat{\beta}$. Для поиска наилучшей аппроксимации для (1.1) воспользуемся каким-либо переборным алгоритмом. Пусть $X-(n\times s)$ – расширенная матрица наблюдений для текущей модели из s регрессоров, образующих множество $L \subset \Re$. Регрессия отклика y по L будет определяться по уравнению наблюдения

$$y = X\theta + e, \tag{1.2}$$

где $e-(n\times 1)$ – вектор ненаблюдаемых случайных ошибок измерения, относительно которых выполнены предположения $E(e)=0_n$, $E(ee^{\mathrm{T}})=\sigma^2 I_n$.

Предположим, что выборка наблюдений W разбита на две части A и B. В методах структурной оптимизации активно используются следующие, так называемые внешние критерии селекции моделей [21–23]:

критерий регулярности:

$$\Delta^2(B) = \Delta^2(B/A) = \left\| y_B - X_B\hat{\theta}_A \right\|^2,$$

где запись $\Delta^2(B/A)$ означает «ошибка» на выборке B модели, коэффициенты, которой получены с использованием выборки A;
критерий симметричной регулярности:

$$d^2 = \Delta^2(B/A) + \Delta^2(A/B) = \left\| y_B - X_B\hat{\theta}_A \right\|^2 + \left\| y_A - X_A\hat{\theta}_B \right\|^2;$$

критерий стабильности:

$$S^2 = \Delta^2(A \cup B/A) = \left\| y_W - X_W\hat{\theta}_A \right\|^2;$$

симметричный критерий стабильности:

$$S_{cc}^2 = \Delta^2(A\cup B/A) + \Delta^2(A\cup B/B) = \left\| y_W - X_W\hat{\theta}_A \right\|^2 + \left\| y_W - X_W\hat{\theta}_B \right\|^2;$$

критерий непротиворечивости:

$$n_{CM}^2 = \left\| X_W\hat{\theta}_A - X_W\hat{\theta}_B \right\|^2;$$

критерий вариативности:

$$V^2 = (X_W\hat{\theta}_A - X_W\hat{\theta}_W)^{\mathrm{T}}(X_W\hat{\theta}_W - X_W\hat{\theta}_B).$$

К рассматриваемой группе критериев относится также критерий «скользящего контроля» (CV-cross validation):

$$\Delta_{ck}^2 = \sum_i (y_i - f^{\mathrm{T}}(x_i)\hat{\theta}_{(i)})^2,$$

где $\hat{\theta}_{(i)}$ – оценка параметров по выборке W с исключенным i-м наблюдением.

Теоретическое обоснование внешних критериев проведено в работах [22–26]. Использование критериев регулярности позволяет отбирать модели оптимальной сложности, ориентированные на работу в режиме прогноза. В условиях действия помех большой интенсивности эти критерии будут указывать, как правило, на модели простой структуры. Использование критериев стабильности позволяет отбирать модели с хорошими свойствами как прогнозирования так и сглаживания.

2. Разбиение выборки для внешних критериев с использованием методов планирования эксперимента

Использование внешних критериев селекции при решении задачи выбора модели оптимальной сложности предполагает разбиение выборки наблюдения на две части: обучающую и проверочную. На обучающей выборке производится оценивание параметров тестируемых моделей, а на проверочной – проверка их прогнозируемых свойств или свойств согласованности решений с обучающей частью выборки.

В данном разделе основное внимание будет уделено критериям качества моделей, связанным с точностью прогнозирования. В силу этого неизбежно встает задача управления разбиением выборки. Некоторые подходы к решению задачи разбиения с использованием методов оптимального планирования эксперимента предложены в работах [27, 28].

Записывая критерий $\Delta^2(B)$ в канонической форме, легко получить его математическое ожидание [29]:

$$E(\Delta^2(B)) = (\hat{X}_B - P_{BA}\hat{X}_A\hat{\theta})^{\mathrm{T}}(\hat{X}_B - P_{BA}\hat{X}_A\hat{\theta}) +$$

$$+\sigma^2\left(n_B + \mathrm{tr}\left(X_A^{\mathrm{T}}X_A\right)^{-1}\left(X_B^{\mathrm{T}}X_B\right)\right), \tag{2.1}$$

где

$$P_{BA} = X_B(X_A^{\mathrm{T}}X_A)^{-1}X_A^{\mathrm{T}}.$$

В [21] рассмотрены условия, при которых оптимальная структура, соответствующая минимуму (2.1), совпадает с истинной структурой $\hat{s} = m$. Эти условия диктуют «квадратично зависимое» разбиение матрицы X:

$$\rho^2 X_A^{\mathrm{T}}X_A = X_B^{\mathrm{T}}X_B, \tag{2.2}$$

где ρ^2 – некоторое произвольное число. Точное квадратичное разбиение (2.2) может иметь место лишь в специально подобранной матрице X, что на практике маловероятно. Кроме того, рекомендации типа (2.2) не учитывают поведение второго слагаемого в (2.1). С учетом (2.2) его можно записать как

$$J_\sigma(s, \sigma) = \sigma^2\left(n_B + s/\rho^2\right). \tag{2.3}$$

Скорость возрастания $J_\sigma(\mathrm{s},\sigma)$ в зависимости от σ определяет помехоустойчивость критерия селекции моделей. Ясно, что необходимо выбирать разбиение с возможно большим значением ρ^2 при малой величине n_B.

В общем случае разбиения X на X_A и X_B величина J_σ в соответствии с (2.1) равна

$$J_{\sigma}(s,\sigma)=\sigma^2\left(n_B+\mathrm{tr}(X_A^{\mathrm{T}}X_A)^{-1}X_B^{\mathrm{T}}X_B)\right). \quad (2.4)$$

Видим, что скорость возрастания $J_{\sigma}(\mathrm{s},\sigma)$ в зависимости от σ определяется средней дисперсией прогноза $\mathrm{tr}\left((X_A^{\mathrm{T}}X_A)^{-1}X_B^{\mathrm{T}}X_B\right)$. Исследуем возможность минимизации J_{σ} (2.4) путем выбора того или иного варианта разбиения X на X_A, X_B при условии, что n_B зафиксировано. Введем следующие обозначения. Пусть ξ есть непрерывный нормированный план, а $M(\xi)$ – информационная матрица, равная $X_A^T X_A / n_A = \sum_{i=1}^{n} p_i x_i x_i^T$. Далее пусть X определяет собой множество точек, среди которых необходимо выбрать n_A точек, присвоив им веса, равные $1/n_A$, а остальным точкам присвоить веса, равные 0. Оптимальный план ξ^* будем находить как решение следующей экстремальной задачи:

$$\xi^* = \mathrm{Arg}\max_{p} \Psi[M(\xi)].$$

В качестве функционала $\Psi[M(\xi)]$ будем рассматривать определитель информационной матрицы, что соответствует D-оптимальному планированию эксперимента [30]. Для рассматриваемого функционала $\Psi[M(\xi)]$ компоненты вектора градиента имеют вид

$$\frac{\partial\Psi[M(\xi)]}{\partial p_j}=d(x_j,\xi)=x_j^T M(\xi)^{-1}x_j=\mathrm{tr}M(\xi)^{-1}x_j x_j^T,$$

где $d(x_j,\xi)$ – дисперсия оценки математического ожидания отклика в точке x_j. Для D-оптимального плана ξ^* будет справедливо

$$\mathrm{tr}M^{-1}(\xi^*)X_B^T X_B < \mathrm{tr}M^{-1}(\xi)X_B^T X_B,$$

где ξ – не D-оптимальный план [28]. Данное утверждение позволяет предложить достаточно простую схему действий: для заданного полного плана эксперимента в виде имеющейся выборки решается задача построения D-оптимального плана ξ^* с n_A точками из X. Не вошедшие в оптимальный план точки выборки образуют собой тестовую ее часть. Отметим, что к такому выводу мы приходим и при рассмотрении критериев стабильности и непротиворечивости [17].

3. Разбиение выборки на обучающую и тестовую части для метода LS-SVM

Рассмотрим задачу восстановления зависимости по зашумленным данным. Дана обучающая выборка $D_n=\{(x_k,y_k):x_k\in X, y_k\in Y; k=1,\dots,n\}$ объема n наблюдений вида

$$y_k=m(x_k)+e_k, k=1,...,n, \quad (3.1)$$

где $e_k\in R$ будем считать независимо и одинаково распределенной ошибкой с $E\left[e_k \mid x=x_k\right]=0$ и $\mathrm{Var}[e_k]=\sigma^2<\infty$; $m(x)$ – неизвестная действительная гладкая функция и $E\left[y_k \mid x = x_k\right] = m(x_k)$. Вместо неизвестной функции $m(x)$ будем использовать ее аппроксимацию в виде $f(x) = \omega^{\mathrm{T}}\varphi(x) + b$. Функционал эмпирического риска использования такой аппроксимации:

$$R_{emp}(\omega, b) = \frac{1}{n}\sum_{k=1}^{n}\left(\left(\omega^{\mathrm{T}}\varphi(x_k) + b\right) - y_k\right)^2. \tag{3.2}$$

Задачу нахождения вектора ω и $b \in R$ можно свести к решению следующей задачи оптимизации [1]:

$$\min_{\omega, b, e} J(\omega, e) = \frac{1}{2}\omega^{\mathrm{T}}\omega + \frac{1}{2}\gamma\sum_{k=1}^{n} e_k^2 \tag{3.3}$$

в предположении, что $y_k = \omega^{\mathrm{T}}\varphi(x_k) + b + e_k$, $k = 1, \ldots, n$. В (3.3) параметр регуляризации γ отвечает за сложность модели, которая в данном случае определяется нормой вектора ω.

Решение задачи (3.3) обычно проводят в двойственном пространстве с использованием функционала Лагранжа:

$$L(\omega, b, e, \alpha) = J(\omega, e) - \sum_{k=1}^{n}\alpha_k\left(\omega^{\mathrm{T}}\varphi(x_k) + b + e_k - y_k\right) \tag{3.4}$$

с лагранжевыми множителями $\alpha_k \in R$.

Условия оптимальности задаются следующим образом:

$$\begin{cases} \dfrac{dL}{d\omega} = 0 \to \omega = \sum\limits_{k=1}^{n}\alpha_k\varphi(x_k),\ k = 1, \ldots, n; \\ \dfrac{dL}{db} = 0 \to \sum\limits_{k=1}^{n}\alpha_k = 0,\ k = 1, \ldots, n; \\ \dfrac{dL}{de_k} = 0 \to \alpha_k = \gamma e_k,\ k = 1, \ldots, n; \\ \dfrac{dL}{d\alpha_k} = 0 \to \omega^{\mathrm{T}}\varphi(x_k) + b + e_k = y_k,\ k = 1, \ldots, n. \end{cases} \tag{3.5}$$

После исключения ω и e получаем решение:

$$\begin{bmatrix} 0 & 1_n^{\mathrm{T}} \\ 1_n & \Omega + \dfrac{1}{\gamma}I_n \end{bmatrix}\begin{bmatrix} \hat{b} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}, \tag{3.6}$$

где $y = (y_1, \ldots, y_n)^{\mathrm{T}}$, $1_n = (1, \ldots, 1)^{\mathrm{T}}$, $\hat{\alpha} = \left(\hat{\alpha}_1, \ldots, \hat{\alpha}_n\right)$ и $\Omega_{kl} = \varphi(x_k)^{\mathrm{T}}\varphi(x_l)$ для $k, l = 1, \ldots, n$. Результирующая LS-SVM модель имеет вид

$$\hat{y}_n(x) = \sum_{k=1}^{n}\hat{\alpha}_k K(x, x_k) + \hat{b}, \tag{3.7}$$

где $K(x, x_k)$ – ядро скалярного произведения,

$$\hat{b} = \frac{1_n^{\mathrm{T}} \left(\Omega + \frac{1}{\gamma} I_n \right)^{-1} y}{1_n^{\mathrm{T}} \left(\Omega + \frac{1}{\gamma} I_n \right)^{-1} 1_n}, \quad \hat{\alpha} = \left(\Omega + \frac{1}{\gamma} I_n \right)^{-1} \left(y - 1_n \hat{b} \right). \tag{3.8}$$

В случае выборок большого размера для получения оценок всех параметров вместо обращения матриц в (3.8) решают систему уравнений (3.6). Точность получаемого решения (3.7), как мы уже отмечали, во многом определяется настройкой внутренних параметров алгоритма LS-SVM, к числу которых относят параметр регуляризации и параметры ядерных функций. Настройку этих параметров будем вести с использованием внешних критериев качества моделей [1, 10].

При рассмотрении точности оценивания модели (3.7) основное внимание будем уделять точности оценивания параметров α.

Обозначим оценки параметров α, полученные на обучающей выборке, как

$$\hat{\alpha}_A = \left(\Omega_A + \frac{1}{\gamma} I_{n_A} \right)^{-1} \left(y_B^* \right),$$

где $\Omega_A = K(x_i, x_j),\ i, j = 1, \ldots, n_A$.

Для удобства различения точек обучающей и тестовой выборок будем обозначать координаты точек обучающей выборки через x, а координаты точек тестовой выборки – через z. С учетом этого элементы ядерной матрицы Φ_B для вычисления прогноза в точки тестовой выборки будем обозначать как

$$(\Phi_B)_{ij} = K(z_i, x_j), i = 1, \ldots, n_B, \ j = 1, \ldots, n_A.$$

Прогнозные значения по модели, полученной на выборке A, рассчитываются как

$$\hat{y}_B = \Phi_B \hat{\alpha}_A + \hat{b}_A.$$

Ковариационная матрица ошибок прогноза на выборку B имеет вид

$$\operatorname{cov}\left(\hat{y}_B\right) = \left(\sigma^2 + \operatorname{cov}(\hat{b}_A)\right) \Phi_B \left(\Omega_A + \frac{1}{\gamma} I_{n_A} \right)^{-2} \Phi_B^{\mathrm{T}} + \operatorname{cov}(\hat{b}_A),$$

где

$$\operatorname{cov}(\hat{b}_A) = \sigma^2 \frac{1_{n_A}^{\mathrm{T}} \left(\Omega + \frac{1}{\gamma} I_{n_A} \right)^{-2} 1_{n_A}}{\left[1_{n_A}^{\mathrm{T}} \left(\Omega + \frac{1}{\gamma} I_{n_A} \right)^{-1} 1_{n_A} \right]^2}.$$

Средняя дисперсия прогноза вычисляется как

$$\overline{\sigma}^2(\hat{y}_B) = (\sigma^2 + \operatorname{cov}(\hat{b}_A)) \operatorname{tr} (\Omega_A + \frac{1}{\gamma} I_{n_A})^{-2} \Phi_B^{\mathrm{T}} \Phi_B / n_B + \operatorname{cov}(\hat{b}_A).$$

Минимизировать среднюю дисперсию $\overline{\sigma}^2(\hat{y}_B)$ будем опосредованно через минимизацию определителя дисперсионной матрицы оценок параметров α. В нашем случае эта дисперсионная матрица имеет вид

$$\operatorname{cov}(\hat{\alpha}_A) = \left(\sigma^2 + \operatorname{cov}(\hat{b}_A)\right)\left(\Omega_A + \frac{1}{\gamma} I_{n_A}\right)^{-2}.$$

Учитывая, что матрица $(\Omega_A + \frac{1}{\gamma} I_{n_A})^{-1}$ положительно определена, будем рассматривать минимизацию определителя $\left|\left(\Omega_A + \frac{1}{\gamma} I_{n_A}\right)^{-1}\right|$ или, что намного проще – максимизацию определителя $\left|\left(\Omega_A + \frac{1}{\gamma} I_{n_A}\right)\right|$. Тем самым мы будем строить дискретный D-оптимальный план объемом в n_A наблюдений, используя все точки имеющейся выборки.

В нашем случае для построения дискретного D-оптимального плана удобно воспользоваться хорошо себя зарекомендовавшими последовательными алгоритмами [31, 32].

Обозначим через G_s матрицу размером $s \times s$ для обучающей выборки объемом в s наблюдений и состоящую из элементов $(G_s)_{ij} = K(x_i, x_j) + \frac{1}{\gamma} I_s$, $i, j = 1, \ldots, s$.

Тогда на шаге $s+1$ матрица G_{s+1} будет иметь вид

$$G_{s+1} = \begin{pmatrix} G_s & F(x_{s+1}) \\ F^{\mathrm{T}}(x_{s+1}) & K(x_{s+1}, x_{s+1}) + \frac{1}{\gamma} \end{pmatrix},$$

где $F^{\mathrm{T}}(x_{s+1}) = \left(K(x_1, x_{s+1}),\ K(x_2, x_{s+1}), \ldots,\ K(x_s, x_{s+1})\right)$.

Определитель окаймленной матрицы легко вычисляется:

$$|G_{s+1}| = |G_s| * \Delta(x_{s+1}),$$

где $\Delta(x_{s+1}) = \left[K(x_{s+1}, x_{s+1}) + \frac{1}{\gamma} - F^{\mathrm{T}}(x_{s+1}) G_s^{-1} F(x_{s+1}) \right]$.

Таким образом, очередная точка, включаемая в обучающую выборку, отыскивается по следующей схеме: $x_{s+1} = \operatorname{Arg}\max_x \Delta(x)$, где аргумент x принимает значения координат точек исходной выборки, еще не включенных в обучающую часть. После проведенного разбиения выборки на части A и B возможно конструировать такие критерии, как критерий регулярности и критерий стабильности. Для использования внешних критериев в симметричной форме необходимо проводить разбиение выборки на две примерно равнозначные части. Для этого можно воспользоваться технологией построения так называемых бипланов [17]. Бипланом $\xi^{(1,2)}$ назовем совокупность $\xi^{(1,2)} = \left\{\xi^{(1)},\ \xi^{(2)}\right\}$, где $\xi^{(1)},\ \xi^{(2)}$ – планы, составленные из точек выборки и различающиеся между собой составом включенных в план точек. Помимо требования о различии планов $\xi^{(1)}$, $\xi^{(2)}$ желательно также, что бы они были примерно равны по информативности. Поэтому для получения оптимального биплана необходимо использовать параллельные конкурирующие между собой за ресурсы последовательные процессы:

$$x_{s+1}^{1} = \operatorname{Arg}\max_{x} \Delta(x)\,, \tag{3.9}$$

$$x_{s+1}^{2} = \operatorname{Arg}\max_{x} \Delta(x)\,, \tag{3.10}$$

где аргумент x принимает значения координат точек исходной выборки еще не включенных в биплан, а процедуры (3.9), (3.10) выполняются поочередно.

4. Вычислительный эксперимент

Настройку параметров ядерных функций, опираясь на тот или иной внешний критерий, как мы уже отмечали, проводят в целях получения решений с хорошей обобщающей способностью. В данном вычислительном эксперименте в качестве внешних критериев использовались критерий регулярности $\Delta^2(B)$, критерий стабильности S^2 и критерий скользящего прогноза $\Delta^2_{\text{ск}}$. Вычисление критерия скользящего прогноза проводилось по всей имеющейся выборке, а для критериев регулярности и стабильности использовалось разбиение выборки на обучающую и тестовую части, выполненное с помощью процедуры D-оптимального планирования.

Для проведения исследования использовалась тестовая функция: $m(x) = 7 / e^{(x+0.75)^2} + 3x$, заданная на отрезке $[-1; 1]$. В качестве ядерной функции использовалось гауссово ядро. В качестве помехи использовались нормально распределенные величины. Уровень помехи (дисперсия случайной величины) выбирался как 5, 10, 15 и 20 % от мощности незашумленного сигнала. Количество наблюдений выбиралось равным 10, 20, 30 и 50. Для получения решений по LS-SVM значение параметра регуляризации γ выбиралось равным 10. Подбор лучшего решения осуществлялся по параметру масштаба RBF ядра, который варьировался от 10^{-5} до 10^{0} с шагом 0,1.

Приведенные ниже в табл. 1, 2 данные отражают только часть полученных в вычислительном эксперименте результатов. Так, в табл. 1, 2 приведены усредненные по 600 реализациям шума значения среднеквадратичной ошибки (MSE), рассчитанной по полученным решениям, выбранным с помощью того или иного внешнего критерия. В таблицах в строках, озаглавленных как CV, REG, STAB, представлены соответственно средние значения MSE, полученные при использовании критерия скользящего контроля, критерия регулярности и критерия стабильности. Условия экспериментов по столбцам различались тем, что использовалось различное количество точек в тестовой части в % от объема полной выборки.

Анализ табл. 1, 2 показывает, что эффективность критерия перекрестной проверки выше эффективности критериев регулярности и стабильности в условиях повышенного шума и использования тестовых выборок малого относительного объема. Эффективность использования критерия стабильности, как правило, выше, чем критерия регулярности.

Таблица 1 / Table 1

Среднее значение MSE при 5 % уровне шума

The average value of the MSE at 5 % noise level

Объем выборки / The sample size	Критерий / Criterion	Количество точек в тестовой части в % / The number of points in the test part in %									
		5	10	15	20	25	30	35	40	45	50
N = 10	CV	0,0152	0,0152	0,0152	0,0152	0,0152	0,0152	0,0152	0,0152	0,0152	0,0152
	REG	0,0113	0,0113	0,0289	0,0289	0,0279	0,0279	0,0144	0,0144	0,0068	0,0068
	STAB	0,0034	0,0034	0,0049	0,0049	0,0049	0,0049	0,0049	0,0049	0,0053	0,0053
N = 20	CV	0,0055	0,0055	0,0055	0,0055	0,0055	0,0055	0,0055	0,0055	0,0055	0,0055
	REG	0,0051	0,0050	0,0041	0,0030	0,0031	0,0015	0,0016	0,0017	0,0019	0,0019
	STAB	0,0015	0,0015	0,0015	0,0015	0,0015	0,0015	0,0015	0,0015	0,0016	0,0016
N = 30	CV	0,0029	0,0029	0,0029	0,0029	0,0029	0,0029	0,0029	0,0029	0,0029	0,0029
	REG	0,0037	0,0035	0,0016	0,0010	0,0011	0,0011	0,0012	0,0012	0,0015	0,0015
	STAB	0,0009	0,0009	0,0009	0,0009	0,0009	0,0009	0,0009	0,0009	0,0010	0,0009
N = 50	CV	0,0013	0,0013	0,0013	0,0013	0,0013	0,0013	0,0013	0,0013	0,0013	0,0013
	REG	0,0021	0,0013	0,0008	0,0008	0,0008	0,0007	0,0007	0,0006	0,0007	0,0007
	STAB	0,0007	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005

Таблица 2 / Table 2

Среднее значение MSE для выборки объема 20 наблюдений при различных уровнях шума

The average value of MSE for the sample of size 20 observations at the different noise levels

Уровень шума / The noise level	Критерий / Criterion	Количество точек в тестовой части в % / The number of points in the test part in %									
		5	10	15	20	25	30	35	40	45	50
5 %	CV	0,0055	0,0055	0,0055	0,0055	0,0055	0,0055	0,0055	0,0055	0,0055	0,0055
	REG	0,0051	0,0050	0,0041	0,0030	0,0031	0,0015	0,0016	0,0017	0,0019	0,0019
	STAB	0,0015	0,0015	0,0015	0,0015	0,0015	0,0015	0,0015	0,0015	0,0016	0,0016
10 %	CV	0,0062	0,0062	0,0062	0,0062	0,0062	0,0062	0,0062	0,0062	0,0062	0,0062
	REG	0,0077	0,0064	0,0057	0,0053	0,0053	0,0035	0,0032	0,0032	0,0032	0,0032
	STAB	0,0034	0,0032	0,0030	0,0029	0,0029	0,0028	0,0027	0,0027	0,0027	0,0027
15 %	CV	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074
	REG	0,0111	0,0094	0,0084	0,0088	0,0081	0,0060	0,0060	0,0059	0,0056	0,0056
	STAB	0,0071	0,0067	0,0061	0,0059	0,0057	0,0056	0,0051	0,0050	0,0049	0,0049
20 %	CV	0,0091	0,0091	0,0091	0,0091	0,0091	0,0091	0,0091	0,0091	0,0091	0,0091
	REG	0,0135	0,0128	0,0119	0,0123	0,0114	0,0096	0,0091	0,0090	0,0083	0,0083
	STAB	0,0127	0,0118	0,0108	0,0102	0,0099	0,0095	0,0085	0,0084	0,0079	0,0078

На рис. 1 представлены достигнутые средние значения MSE при использовании трех внешних критериев в зависимости от величины тестовой выборки для случая 15 % шума и выборки объемом в 30 наблюдений. Средние значения MSE, достигнутые при использовании критерия скользящего контроля, показаны горизонтальной прямой. Видим, что выигрыш от использования критериев регулярности и стабильности, как правило, может быть достигнут при относительно большой тестовой части выборки. На рис. 2 представлены достигнутые средние значения MSE при использовании критерия стабильности для выборки в 30 наблюдений при вариации объема тестовой части и изменении уровня шума от 5 до 20 %. Видим, что при увеличении уровня шума минимум критерия сдвигается вправо. Это говорит о том, что в условиях использования сильно зашумленных выборок целесообразно по возможности использовать критерий стабильности с относительно большой тестовой частью. Аналогичные результаты для критерия регулярности представлены на рис. 3.

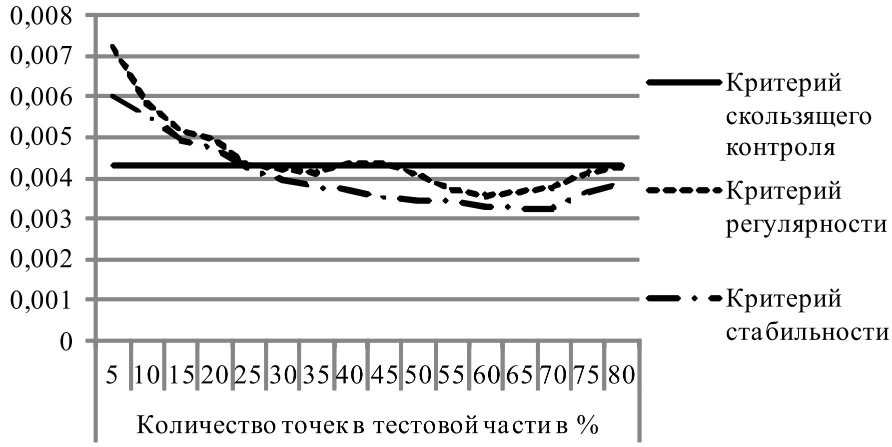

Рис. 1 – График средних значений MSE для выборки объема 30 при 15 % уровне шума

Fig. 1 – The graph of average values of MSE for the sample of size 30 with 15 % noise level

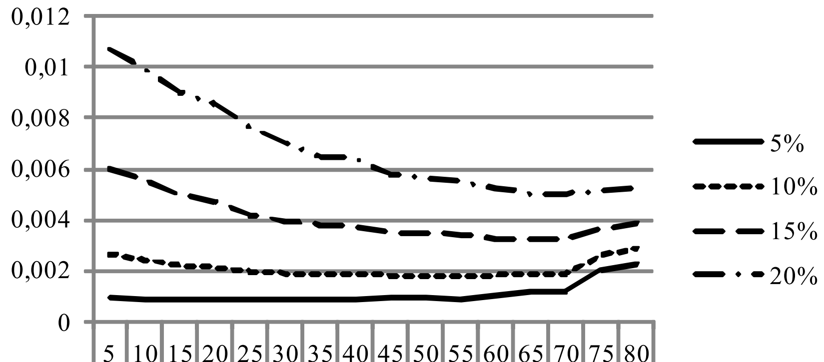

Рис. 2 – График средних значений MSE при использовании критерия стабильности для выборки объема 30 с изменением уровня шума от 5 до 20 %

Fig. 2 – The graph of average values of the MSE when using the stability criterion for the sample of size 30 with the change in noise level from 5 to 20 %

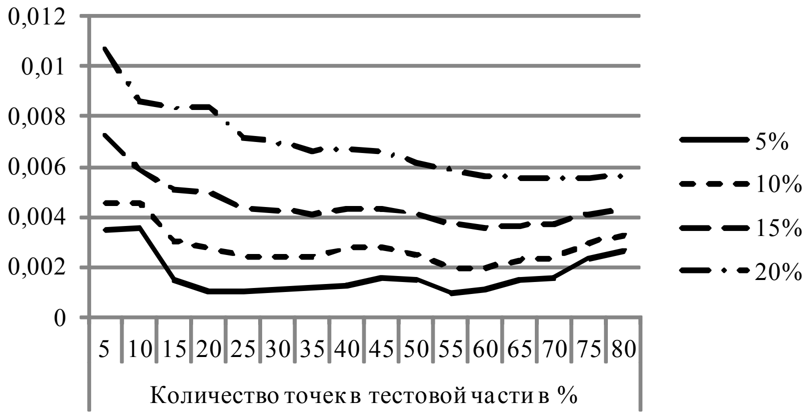

Рис. 3 – График средних значений MSE при использовании критерия регулярности для выборки объема 30 с изменением уровня шума от 5 до 20 %

Fig. 3 – The graph of average values of the MSE when using the regularity criterion for the sample of size 30 with the change in noise level from 5 to 20 %

Заключение

В работе для настройки параметров ядерных функций в методе LS-SVM проведено сравнительное исследование нескольких внешних критериев. Для возможности использования критериев регулярности и стабильности предложен способ разбиения выборки на тестовую и обучающую части на основе метода планирования эксперимента. Предложен соответствующий последовательный алгоритм получения D-оптимального плана. Для использования внешних критериев в симметричной форме предлагается алгоритм построения бипланов.

По результатам проведенных вычислительных экспериментов можно сделать выводы о том, что эффективность использования критериев стабильности и регулярности зависит от объема тестовой части выборки. При достаточном объеме тестовой части их эффективность, как правило, выше эффективности критерия скользящего контроля. В случаях использования выборок малого объема предпочтение следует отдавать критерию скользящего контроля. Перспективность использования критериев стабильности и регулярности связана также с возможностью их использования для получения так называемых разреженных решений.

ЛИТЕРАТУРА

1. Least squares support vector machines / J.A.K. Suykens, T. van Gestel, J. de Brabanter, B. de Moor, J. Vandewalle. – New Jersey; London; Singapore; Hong Kong: World Scientific, 2002. – 290 p.
2. Regularization, optimization, kernels, and support vector machines / ed. by J.A.K. Suykens, M. Signoretto, A. Argyriou. – Boca Raton, FL: CRC Press, 2014. – 525 p. – (Chapman & Hall/CRC Machine Learning & Pattern Recognition Series).
3. **Vapnik V.** Statistical learning theory. – New York: John Wiley, 1998. – 736 p.
4. **Cherkassky V., Ma Y.Q.** Practical selection of SVM parameters and noise estimation for SVM regression // Neural Networks. – 2004. – N 17. – P. 113–126.
5. **Stone M.** Cross-validatory choice and assessment of statistical predictions // Journal of the Royal Statistical Society. Series B. – 1974. – N 36 (2). – P. 111–147.
6. **Wahba G.** A survey of some smoothing problems and the method of generalized cross-validation for solving them // Application of Statistics: Proceedings of the Symposium Held at Wright State University, Dayton, Ohio, 14–18 June 1976. – Amsterdam: North-Holland, 1977. – P. 507–523.

7. **Wahba G.** Support vector machines, reproducing kernel Hilbert spaces and the randomized GACV // Advances in Kernel Methods: Support Vector Learning. – Cambridge: MIT Press, 1999. – P. 69–88.
8. **Попов А.А., Саутин А.С.** Определение параметров алгоритма опорных векторов при решении задачи построения регрессии // Сборник научных трудов НГТУ. – 2008. – № 2 (52). – С. 35–40.
9. **Popov A.A., Sautin A.S.** Selection of support vector machines parameters for regression using nested grids // The Third International Forum on Strategic Technology (IFOST 2008): proceedings, Novosibirsk–Tomsk, Russia, 23–29 June 2008. – Novosibirsk, 2008. – P. 329–331.
10. **Попов А.А., Бобоев Ш.А.** Построение регрессионных зависимостей с использованием квадратичной функции потерь в методе опорных // Сборник научных трудов НГТУ. – 2015. – № 3 (81). – С. 69–78.
11. **Гладкова А.В., Попов А.А.** Выбор настраиваемых параметров алгоритма опорных векторов с квадратичной функцией потерь // Обработка информации и математическое моделирование: материалы Российской научно-технической конференции, Новосибирск, 23–24 апреля 2015 г. – Новосибирск, 2015. – С. 62–66.
12. **Cawley G.C., Talbot N.L.C.** Preventing over-fitting during model selection via Bayesian regularisation of the hyper-parameters // Journal of Machine Learning Research. – 2007. – Vol. 8. – P. 841–861.
13. Leave-one-out cross-validation-based model selection for multi-input multi-output support vector machine / W. Mao, X. Mu, Y. Zheng, G. Yan // Neural Computing and Application. – 2014. – Vol. 24, iss. 2. – P. 441–451.
14. **Rivas-Perea P., Cota-Ruiz J., Rosiles J.-G.** A nonlinear least squares quasi-Newton strategy for LP-SVR hyper-parameters selection // International Journal of Machine Learning and Cybernetics. – 2014. – Vol. 5, iss. 4. – P. 579–597.
15. Optimisation of turning parameters by integrating genetic algorithm with support vector regression and artificial neural networks / A. Gupta, S. Guntuku, R. Desu, A. Balu // The International Journal of Advanced Manufacturing Technology. – 2015. – Vol. 77, iss. 1–4. – P. 331–339.
16. **Гультяева Т.А., Попов А.А., Саутин А.С.** Методы статистического обучения в задачах регрессии и классификации: монография. – Новосибирск: Изд-во НГТУ, 2016. – 322 с.
17. **Попов А.А.** Оптимальное планирование эксперимента в задачах структурной и параметрической идентификации моделей многофакторных систем: монография. – Новосибирск: Изд-во НГТУ, 2013. – 296 с.
18. **Перельман И.И.** Методология выбора структуры модели при идентификации объектов управления // Автоматика и телемеханика. – 1983. – № 11. – С. 5–29.
19. **Романов В.Л.** Выбор наилучшей линейной регрессии: сравнение формальных критериев // Заводская лаборатория. – 1990. – № 1. – С. 90–95.
20. **Себер Дж.** Линейный регрессионный анализ. – М.: Мир, 1980. – 456 с.
21. **Степашко В.С., Кочерга Ю.Л.** Методы и критерии решения задач структурной идентификации // Автоматика. – 1985. –№ 5. – С. 29–37.
22. **Кочерга Ю.Л.** J-оптимальная редукция структуры модели в схеме Гаусса–Маркова // Автоматика. – 1988. – № 4. – С. 34–38.
23. **Сарычев А.П.** Усредненный критерий регулярности метода группового учета аргументов в задаче поиска наилучшей регрессии // Автоматика. – 1990. – № 5. – С. 28–33.
24. **Степашко В.С.** Асимптотические свойства внешних критериев выбора моделей // Автоматика. – 1988. – № 6. – С. 75–82.
25. **Степашко В.С.** Потенциальная помехоустойчивость моделирования по комбинаторному алгоритму МГУА без использования информации о помехах // Автоматика. – 1983. – № 3. – С. 18–28.
26. **Степашко В.С.** Селективные свойства критерия непротиворечивости моделей // Автоматика. – 1986. – № 2. – С. 40–49.
27. **Попов А.А.** Планирование эксперимента в задачах разбиения выборки в МГУА // Сборник научных трудов НГТУ. – 1995. – Вып. 2. – С. 35–40.

28. **Попов А.А.** Разбиение выборки для внешних критериев селекции моделей с использованием методов планирования эксперимента // Заводская лаборатория. Диагностика материалов. – 1997. – № 1. – С. 49–53.
29. **Юрачковский Ю.П., Грошков А.Н.** Применение канонической формы внешних критериев для исследования их свойств // Автоматика. – 1979. – № 3. – С. 85–89.
30. **Федоров В.В.** Активные регрессионные эксперименты // Математические методы планирования эксперимента. – Новосибирск: Наука, 1981. – С. 19–73.
31. **Попов А.А.** Последовательные схемы построения оптимальных планов эксперимента // Сборник научных трудов НГТУ. – 1995. – Вып. 1. – С. 39–44.
32. **Попов А.А.** Последовательные схемы синтеза оптимальных планов эксперимента // Доклады АН ВШ РФ. – 2008. – № 1 (10). – С. 45–55.

KERNEL FUNCTION PARAMETER SETTING IN THE LS-SVM METHOD USING EXTERNAL CRITERIA OF MODEL QUALITY

Sukhodolov A.P., Popov A.A., Boboev Sh.A.
Novosibirsk State Technical University, Novosibirsk, Russia

The problem of regression dependence recovery by the support vector method with a quadratic loss function is considered in the paper. This method belongs to the kernel methods class. To set up a number of internal parameters of the LS-SVM algorithm the problem of using external criteria of model quality is discussed. Various criteria of model selection which are based on partitioning the sample into learning and test parts are given. The problem of partitioning the sample into learning and test parts with the use of the D–optimal experiment design method is considered in detail for the case of linear parametric regression models. The method of obtaining the test sample is proposed for the LS–SVM method. A sequential algorithm is presented for obtaining the learning and test parts of the sample observations applied to the LS–SVM method. To use the criteria in the symmetric form the algorithm of construction biplanes is proposed. The results of computational experiment are presented to consider the possibility of using the three external criteria to select the Gaussian kernel function scale. The cross-validation criterion, the regularity criterion and the stability criterion were used as external criteria. The scale parameter of the kernel function was selected by the minimum of external criterion quality. The final accuracy of solution obtained was tested by the mean-square error. The computational experiment was performed on simulated data. Nonlinear dependence on the input factor was chosen as a model generating data.Noise variance (the noise level) was determined as a percentage of the power signal. The results of some computational experiments are given in tables and figures. Based on the results of computational experiments conclusions are made that the effectiveness of using the stability criterion to obtain solutions with a small mean square error tends to be higher than when the regularity criterion is used. The efficiency of the cross-validation criterion is higher than the efficiency of the regularity and stability criteria under conditions of a high noise level and using test samples of a small relative volume.

Keywords: regression, LS-SVM method, quadratic loss function, test sample, training sample, optimal experiment design, D-optimal design, regularity criterion, stability criterion, cross-validation criterion, kernel function, mean square error.

DOI: 10.17212/1727-2769-2017-3-88-104

REFERENCES

1. Suykens J.A.K., Gestel T. van, Brabanter J. de, Moor B. de, Vandewalle J. *Least square support vector machines*. New Jersey, London, Singapore, Hong Kong, World Scientific, 2002. 290 p.
2. Suykens J.A.K., Signoretto M., Argyriou A., eds. *Regularization, optimization, kernels, and support vector machines*. Boca Raton, FL, CRC Press, 2014. 525 p.
3. Vapnik V. *Statistical learning theory*. New York, John Wiley, 1998. 736 p.

4. Cherkassky V., Ma Y. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks*, 2004, no. 17, pp. 113–126.
5. Stone M.. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society. Series B*, 1974, no. 36 (2), pp. 111–147.
6. Wahba G. A survey of some smoothing problems and the method of generalized cross-validation for solving them. *Application of Statistics: Proceedings of the Symposium Held at Wright State University*, Dayton, Ohio, 14–18 June 1976. Amsterdam, North-Holland, 1977, pp. 507–523.
7. Wahba G. Support vector machines, reproducing kernel Hilbert spaces and the randomized GACV. *Advances in Kernel Methods: Support Vector Learning.* Cambridge, MIT Press, 1999, pp. 69–88.
8. Popov A.A., Sautin A.S. Opredelenie parametrov algoritma opornykh vektorov pri reshenii zadachi postroeniya regressii [Parameters estimation in support vector regression]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Transaction of scientific papers of the Novosibirsk state technical university*, 2008, no. 2 (52), pp. 35–40.
9. Popov A.A., Sautin A.S. Selection of support vector machines parameters for regression using nested grids. *The third international forum on strategic technology (IFOST 2008): proceedings*, Novosibirsk–Tomsk, Russia, 23–29 June 2008, pp. 329–331.
10. Popov A.A., Boboev Sh.A. Postroenie regressionnykh zavisimostei s ispol'zovaniem kvadratichnoi funktsii poter' v metode opornykh [The construction of a regression relationships using least square in support vector machines]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Transaction of scientific papers of the Novosibirsk state technical university*, 2015, no. 3 (81), pp. 69–78.
11. Gladkova A.V., Popov A.A. [The select of the configurable parameters of the algorithm of support vector machine with quadratic loss function]. *Obrabotka informatsii i matematicheskoe modelirovanie*: materialy Rossiiskoi nauchno-tekhnicheskoi konferentsii [Information processing and mathematical modeling: proceedings of Russian scientific and technical conference], Novosibirsk, 24–25 April 2015, pp. 62–66. (In Russian).
12. Cawley G.C., Talbot N.L.C. Preventing over-fitting during model selection via Bayesian regularisation of the hyper-parameters. *Journal of Machine Learning Research*, 2007, vol. 8, pp. 841–861.
13. Mao W., Mu X., Zheng Y., Yan G. Leave-one-out cross-validation-based model selection for multi-input multi-output support vector machine. *Neural Computing and Application*, 2014, vol. 24, iss. 2, pp. 441–451.
14. Rivas-Perea P., Cota-Ruiz J., Rosiles J.-G. A nonlinear least squares quasi-Newton strategy for LP-SVR hyper-parameters selection. *International Journal of Machine Learning and Cybernetics*, 2014, vol. 5, iss. 4, pp. 579–597.
15. Gupta A., Guntuku S., Desu R., Balu A. Optimisation of turning parameters by integrating genetic algorithm with support vector regression and artificial neural networks. *The International Journal of Advanced Manufacturing Technology*, 2015, vol. 77, iss. 1–4, pp. 331–339.
16. Gul'tyaeva T.A., Popov A.A., Sautin A.S. *Metody statisticheskogo obucheniya v zadachakh regressii i klassifikatsii* [The methods of statistical learning in problems of regression and classification]. Novosibirsk, NSTU Publ., 2016. 322 p.
17. Popov A.A. *Optimal'noe planirovanie eksperimenta v zadachakh strukturnoi i parametricheskoi identifikatsii modelei mnogofaktornykh sistem* [The optimal planning of experiment in problems of structural and parametric identification of models of multifactor systems]. Novosibirsk, NSTU Publ., 2013. 296 p.
18. Perel'man I.I. Metodologiya vybora struktury modeli pri identifikatsii ob"ektov upravleniya [A methodology for the selection of the model structure when identification of objects of management]. *Avtomatika i telemekhanika – Automation and Remote Control*, 1983, no. 11, pp. 5–29. (In Russian).
19. Romanov V.L. Vybor nailuchshei lineinoi regressii: sravnenie formal'nykh kriteriev [The select of the best linear regression: a comparison of formal criteria]. *Zavodskaya laboratoriya – Industrial laboratory*, 1990, no. 1, pp. 90–95. (In Russian).
20. Seber J.A.F. *Linear regression analysis*. New York, Wiley, 1977 (Russ. ed.: Seber Dzh. *Lineinyi regressionnyi analiz*. Moscow, Mir Publ., 1980. 456 p.).

21. Stepashko V.S., Kocherga Yu.L. Metody i kriterii resheniya zadach strukturnoi identifikatsii [Methods and criteria of the solving problems of structural identification]. *Avtomatika – Soviet Journal of Automation and Information Sciences*, 1985, no. 5, pp. 29–37. (In Russian).
22. Kocherga Yu.L. J-optimal'naya reduktsiya struktury modeli v skheme Gaussa–Markova [J-optimal reduction of structure of model in the scheme of Gauss–Markov]. *Avtomatika – Soviet Journal of Automation and Information Sciences*, 1988, no. 4, pp. 34–38. (In Russian).
23. Sarychev A.P. Usrednennyi kriterii regulyarnosti metoda gruppovogo ucheta argumentov v zadache poiska nailuchshei regressii [The averaged regularity criterion of group method of accounting arguments in the problem of finding the best regression]. *Avtomatika – Soviet Journal of Automation and Information Sciences*, 1990, no. 5, pp. 28–33. (In Russian).
24. Stepashko V.S. Asimptoticheskie svoistva vneshnikh kriteriev vybora modelei [The asymptotic properties of the external criteria of selection models]. *Avtomatika – Soviet Journal of Automation and Information Sciences*, 1988, no. 6, pp. 75–82. (In Russian).
25. Stepashko V.S. Potentsial'naya pomekhoustoichivost' modelirovaniya po kombinatornomu algoritmu MGUA bez ispol'zovaniya informatsii o pomekhakh [The potential noise immunity of modeling by combinatorial GMDH algorithm without using the interference information]. *Avtomatika – Soviet Automatic Control*, 1983, no. 3, pp. 18–28. (In Russian).
26. Stepashko V.S. Selektivnye svoistva kriteriya neprotivorechivosti modelei [The selective properties of the consistency criterion of models]. *Avtomatika – Soviet Journal of Automation and Information Sciences*, 1986, no. 2, pp. 40–49. (In Russian).
27. Popov A.A. Planirovanie eksperimenta v zadachakh razbieniya vyborki v MGUA [The experiment planning in problems of splitting the sample in GMDH]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Transaction of scientific papers of the Novosibirsk state technical university*, 1995, iss. 2, pp. 35–40.
28. Popov A.A. Razbienie vyborki dlya vneshnikh kriteriev selektsii modelei s ispol'zovaniem metodov planirovaniya eksperimenta [Splitting the sample for ex-ternal criteria of selection models using methods of experiment planning]. *Zavodskaya laboratoriya. Diagnostika materialov – Industrial laboratory. Materials diagnostics*, 1997, no. 1, pp. 49–53. (In Russian).
29. Yurachkovskii Yu.P., Groshkov A.N. Primenenie kanonicheskoi formy vneshnikh kriteriev dlya issledovaniya ikh svoistv [The use of canonical form of external criteria for the research of their properties]. *Avtomatika – Soviet Automatic Control*, 1979, no. 3, pp. 85–89. (In Russian).
30. Fedorov V.V. Aktivnye regressionnye eksperimenty [The active regression experiments]. *Matematicheskie metody planirovaniya eksperimenta* [The mathematical methods of experimental planning]. Novosibirsk, Nauka Publ., 1987, pp. 19–73.
31. Popov A.A. Posledovatel'nye skhemy postroeniya optimal'nykh planov eksperimenta [The sequential schemes constructing of the optimal experiment plans]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Transaction of scientific papers of the Novosibirsk state technical university*, 1995, iss. 1, pp. 39–44.
32. Popov A.A. Posledovatel'nye skhemy sinteza optimal'nykh planov eksperimenta [Sequential schemes of synthesis of optimum plans of experiment]. *Doklady Akademii nauk vysshei shkoly Rossiiskoi Federatsii – Proceedings of the Russian higher school Academy of sciences*, 2008, no. 1 (10), pp. 45–55.

СВЕДЕНИЯ ОБ АВТОРАХ

Суходолов Александр Петрович – родился в 1956 году, д-р экон. наук, профессор, ректор Байкальского государственного университета. Область научных интересов: социально-экономическое развитие регионов Сибири, российско-азиатские научные и образовательные связи, цифровая экономика, математическое моделирование, массовые коммуникации. (Адрес: 664003, Россия, Иркутская область, г. Иркутск, ул. Ленина, д. 11. E-mail: info@bgu.ru).

Sukhodolov Alexander Petrovich (b. 1956), Doctor of Science (Econ.), professor, rector of Baikal State University. Alexander P. Sukhodolov’s research interests are focused on social and economic development of Siberian regions, Russian-Asian scientific and educational cooperation, digital economy, mathematical modelling, and mass communications. (Address: 11 Lenin street, Irkutsk, 664003, Russia. E-mail: info@bgu.ru).

Попов Александр Александрович – родился в 1952 году, д-р техн. наук, профессор, профессор, кафедра теоретической и прикладной информатики Новосибирского государственного технического университета. Область научных интересов: методы анализа данных, оптимальное планирование экспериментов. Опубликовано более 150 научных работ. (Адрес: 630073, Россия, Новосибирск, пр. Карла Маркса, 20. E-mail: a.popov@corp.nstu.ru).

Popov Aleksandr Aleksandrovich (b. 1952) – Doctor of Science (Eng.), professor, professor, Department of Theoretical and Applied Informatics, Novosibirsk State Technical University. His research interests are currently focused on methods of data analysis and optimal experiment planning. He is the author of 150 scientific papers. (Address: 20, Karl Marx Av., Novosibirsk, 630073, Russia. E-mail: a.popov@corp.nstu.ru).

Бобоев Шараф Асрорович – родился в 1987 году, аспирант, кафедра теоретической и прикладной информатики Новосибирского государственного технического университета. Область научных интересов: статистические методы анализа данных. Опубликовано 4 научные работы. (Адрес: 630073, Россия, Новосибирск, пр. Карла Маркса, 20. E-mail: shboboev@mail.ru).

Boboev Sharaf Asrorovich (b. 1987) – post-graduate student, Department of Theoretical and Applied Informatics, Novosibirsk State Technical University. His research interests are currently focused on statistical methods of data analysis. He is the author of 4 scientific papers. (Address: 20, Karl Marx Av., Novosibirsk, 630073, Russia. E-mail: shboboev@mail.ru).

Статья поступила 26 марта 2017 г.
Received March 26, 2017

To Reference:

Sukhodolov A.P., Popov A.A, Boboev Sh.A. Nastroika parametrov yadernykh funktsii v metode LS-SVM s ispol'zovaniem vneshnikh kriteriev kachestva modelei [Kernel function parameter setting in the LS-SVM method using external criteria of model quality]. *Doklady Akademii nauk vysshei shkoly Rossiiskoi Federatsii – Proceedings of the Russian higher school Academy of sciences*, 2017, no. 3 (36), pp. 88–104. doi: 10.17212/1727-2769-2017-3-88-104