

УДК 629.488

В. А. Четвергов

Омский государственный университет путей сообщения

В. М. Бочаров, А. И. Мишин, П. А. Сиряк, В. А. Мельниченко

ООО «Производственное конструкторско-технологическое предприятие «Транспорт»

ПРИМЕНЕНИЕ МЕТОДА ЛЕВЕНБЕРГА–МАРКВАРДА ДЛЯ АНАЛИЗА ДАННЫХ БОРТОВОЙ ДИАГНОСТИЧЕСКОЙ СИСТЕМЫ ЛОКОМОТИВА

Для создания автоматизированных информационных систем необходимы формирование информационных потоков и разработка методик и алгоритмов обработки этой информации. По результатам исследования предложена методика аппроксимации диагностических параметров, измеренных бортовой системой АПК «Борт», на примере тока тягового генератора в режиме разгона тепловоза методом наименьших квадратов с использованием нелинейных моделей. Изложенная методика ориентирована на применение в области диагностирования тягового подвижного состава по результатам анализа информации, получаемой от бортовых систем.

экспериментальные статистические данные, диагностирование, аппроксимация, матрица, метод наименьших квадратов, метод Левенберга-Марквардта, средняя ошибка аппроксимации.

Введение

Целью данной работы является разработка методики аппроксимации результатов измерения бортовой диагностической системой тока тягового генератора, в режиме разгона тепловоза, методом наименьших квадратов с использованием нелинейных моделей. Данная методика может быть использована для автоматизации процессов оценки технического состояния локомотивов и формирования рекомендаций по объему и периодичности технического обслуживания и ремонта на основе результатов измерения диагностических параметров бортовой системой.

Изложенная в настоящей статье методика ориентирована на применение в области диагностирования тягового подвижного состава по результатам анализа информации, получаемой от бортовых систем. При разработке использовались исходные экспериментальные данные из архивов результатов измерений аппаратно-программным комплексом «Борт» (АПК «Борт»).

1 Методика применения метода Левенберга-Марквардта

В статье [1] описана методика оценки неизвестных параметров регрессионных моделей по выборочным данным методом наименьших модулей (далее – МНМ). Выборками являлись результаты измерения АПК «Борт» тока тягового генератора в режиме разгона маневрового тепловоза. Данные одного переходного процесса были названы реализацией, требования к которой были определены. В качестве модели использовалась функция:

$$\tilde{F}_1(t_i, \mathbf{w}) = y_{\min} + \frac{(y_{\max} - y_{\min})y_{\max}}{at_i^3 + bt_i^2 + ct_i + y_{\max}}, \quad (1)$$

где $\tilde{F}_1(t_i, \mathbf{w})$ – аппроксимирующая функция; t_i – время измерения тока в реализации; y_{\min}, y_{\max} – наименьшее и наибольшее значения тока в одной реализации; $i = 1, 2, \dots, n$ – номер измерения в одной реализации; $\mathbf{w} = (a, b, c)$ – вектор неизвестных коэффициентов.

Далее по тексту обозначение w_1 соответствует a , $w_1 - b$ и т. п.

Модель (1) подбиралась экспериментально. С учетом особенностей полиномов 3-й степени возможно наличие локальных экстремумов в области определения ($t_i \in [0, t_{\max}]$). Поэтому для дальнейших исследований дополнительно подобраны функции:

$$\tilde{F}_2(t_i, \mathbf{w}) = \exp\left(\frac{1}{w_1 t_i^3 + w_2 t_i^2 + w_3 t_i + w_4}\right); \quad (2)$$

$$F_3(t_i, \mathbf{w}) = w_1 \exp(w_2 t_i) + w_4; \quad (3)$$

$$\tilde{F}_4(t_i, \mathbf{w}) = w_1 \exp(t_i^{w_2} - t_i^{w_3}) + w_4; \quad (4)$$

$$\begin{aligned} \tilde{F}_5(t_i, \mathbf{w}) = \\ = w_1 \exp(w_2 t_i) (w_3 t_i^3 + w_4 t_i^2 + w_5 t_i + 1). \end{aligned} \quad (5)$$

Регрессионный анализ в основном выполняется методом наименьших квадратов (далее – МНК), основанным на минимизации суммы квадратов некоторых функций искомого переменных.

С этой точки зрения измерения тока тягового генератора локомотивов образуют генеральную совокупность A . Переходный процесс представляет собой выборку $D = \{t_i, y_i\}_{i=1}^n$, где $D \subset A$, $t_i \in \mathbf{N}$ – время измерения тока в реализации D , $y_i \in \mathbf{Z}$ – соответствующее времени значение тока. Целью статистической обработки этих данных является нахождение коэффициентов функции регрессии $\tilde{F}_z(t_i, \mathbf{w})$. При таких условиях формализация задачи о МНК имеет вид:

$$S(\mathbf{w}) = \sum_{i=1}^n (\tilde{F}_z(t_i, \mathbf{w}) - y_i)^2 \rightarrow \min, \quad (6)$$

где $z = 1, \dots, 5$ – номер функции.

То есть требуется найти такие параметры \mathbf{w} , при которых функция $S(\mathbf{w})$ достигает минимального значения. Для этого решается система уравнений (7), где частные производные функции (6) приравниваются к нулю:

$$\begin{cases} \frac{\partial S}{\partial w_1} = 0, \\ \frac{\partial S}{\partial w_2} = 0, \\ \dots \\ \frac{\partial S}{\partial w_k} = 0, \end{cases} \quad (7)$$

где k – количество коэффициентов \mathbf{w} .

Если в формуле (6) функция $S(\mathbf{w})$ линейна или сводится к этому виду, то система уравнений (7) также будет линейной (функции (1), (2)). Такую систему можно решить любым численным прямым методом. Для функций (3), (4), (5) необходимо применять итерационные методы решения.

После анализа существующих методов многомерной оптимизации (градиентный спуск, покоординатный спуск, метод Ньютона и т. п.) выбран метод Левенберга–Марквардта (МЛМ), сочетающий методы наискорейшего спуска (т. е. минимизации вдоль градиента) и Ньютона (т. е. использование квадратичной модели для ускорения поиска минимума функции). От метода наискорейшего спуска МЛМ позаимствовал стабильность работы, от метода Ньютона – ускоренную сходимость [2].

Для реализации МЛМ вычисляется начальный вектор коэффициентов \mathbf{w}^0 . Общего правила нахождения \mathbf{w}^0 для функций (3), (4), (5) не существует [3], поэтому для определения \mathbf{w}^0 в функцию (3), в частности, подставляются значения $(0, y_{\max})$. Получается условие $w_1 + w_4 = y_{\max}$, которое можно использовать для подбора начального значения коэффициентов.

Приняты следующие значения начальных коэффициентов:

$$w_4^0 = y_{\min}, \quad w_1^0 = y_{\max} - w_4^0. \quad (8)$$

Коэффициенты w_2^0 и w_3^0 находятся из системы уравнений (9) по двум совокупностям данных (t, y) , равномерно взятым из выборки D :

$$\begin{cases} y_p = (y_{\max} - y_{\min}) \exp(w_2^0 t_p^{w_3^0}) + y_{\min}, \\ y_q = (y_{\max} - y_{\min}) \exp(w_2^0 t_q^{w_3^0}) + y_{\min}, \end{cases} \quad (9)$$

где p, q – числа из ряда $1, 2, \dots, n$, делящие отрезок $[1, n]$ на равные интервалы.

Выражая коэффициенты из системы (9), получаем уравнения для вычисления w_2^0 и w_3^0 :

$$\begin{aligned} w_3^0 &= \\ &= \frac{\ln\left(-\ln\left(\frac{y_q - w_4^0}{w_1^0}\right)\right) - \ln\left(-\ln\left(\frac{y_p - w_4^0}{w_1^0}\right)\right)}{\ln(t_q) - \ln(t_p)}, \\ w_2^0 &= \frac{\ln\left(\frac{y_p - w_4^0}{w_1^0}\right)}{t_p^{w_3^0}}. \end{aligned} \quad (10)$$

В результате начальный вектор \mathbf{w}^0 для функции (3) определяется выражениями (8), (10).

Для функции (4) применялся иной подход нахождения компонентов вектора \mathbf{w}^0 . Коэффициенты w_1^0, w_4^0 подбирались вручную и равны соответственно:

$$w_1^0 = 0,8y_{\max}, \quad w_4^0 = 0,2y_{\max}. \quad (11)$$

Такие соотношения взяты из результата работы МЛМ по обучающей выборке $D_{\text{обуч}}$, являющейся одним из характерных переходных процессов.

Для нахождения w_2^0 и w_3^0 использовался метод секущих, применяемый после определенных преобразований [4]. После выбора p, q и подстановки w_1^0, w_4^0 в (4) была получена система:

$$\begin{cases} y_p = w_1^0 \exp(t_p^{w_2^0} - t_p^{w_3^0}) + w_4^0, \\ y_q = w_1^0 \exp(t_q^{w_2^0} - t_q^{w_3^0}) + w_4^0. \end{cases} \quad (12)$$

Она (12) приводится к виду:

$$\begin{cases} \ln\left(\frac{y_p - w_4^0}{w_1^0}\right) = t_p^{w_2^0} - t_p^{w_3^0}, \\ \ln\left(\frac{y_q - w_4^0}{w_1^0}\right) = t_q^{w_2^0} - t_q^{w_3^0}. \end{cases} \quad (13)$$

Заменим переменные следующим образом: $t_p^{w_2^0} = u, t_p^{w_3^0} = v, t_q^{w_2^0} = u^m, t_q^{w_3^0} = v^m$. Это позволит вычислить показатель степени m , используя систему уравнений:

$$\begin{cases} t_p^{w_2^0} = u, \\ t_q^{w_2^0} = u^m, \end{cases} \quad (14)$$

$$\text{откуда } m = \frac{\ln(t_q)}{\ln(t_p)}.$$

Система (13) принимает вид:

$$\begin{cases} \ln\left(\frac{y_p - w_4^0}{w_1^0}\right) = t_p^{w_2^0} - t_p^{w_3^0}, \\ \ln\left(\frac{y_q - w_4^0}{w_1^0}\right) = (t_p^{w_2^0})^{\frac{\ln(t_q)}{\ln(t_p)}} - (t_p^{w_3^0})^{\frac{\ln(t_q)}{\ln(t_p)}}. \end{cases} \quad (15)$$

Результатом преобразований является уравнение с одной переменной:

$$\begin{aligned} f(w_2^0) &= \ln\left(\frac{y_q - w_4^0}{w_1^0}\right) - \\ &- (t_p^{w_2^0})^{\frac{\ln(t_q)}{\ln(t_p)}} + \left| \ln\left(\frac{y_p - w_4^0}{w_1^0}\right) - t_p^{w_2^0} \right|^{\frac{\ln(t_q)}{\ln(t_p)}}. \end{aligned} \quad (16)$$

Функция (16) нелинейная и является гладкой на отрезке $[-1, 1]$, то есть она имеет непрерывную первую производную на этом отрезке (рис. 1). Решение данного уравнения выполняется итерационным численным методом, который должен обладать устойчивостью (нечувствительностью к неточностям в исходных данных) и сходимостью (близостью получаемого численного решения задачи к истинному решению) [5]. Рас-

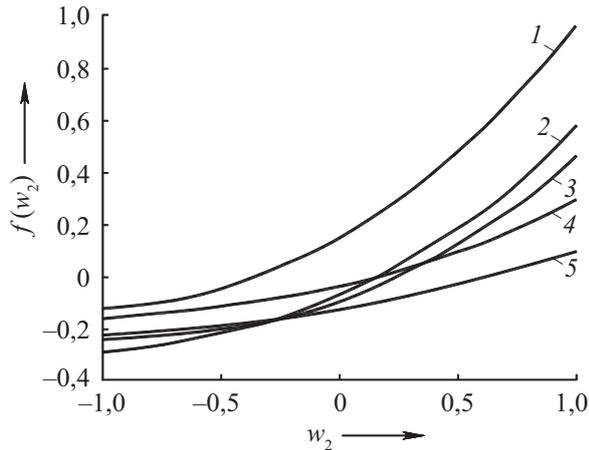


Рис. 1. Зависимость $f(w_2)$ от w_2 :
1, 2, 3, 4, 5 – реализации с n равным
соответственно 8, 11, 16, 24, 32

ширенным понятием сходимости является порядок сходимости α . Для наиболее распространенных одинаково устойчивых методов хорд, дихотомии, секущих α имеет значение 1, 1, 1,618 соответственно. Поэтому в данной работе предпочтение отдано последнему методу, итерационная формула которого имеет вид [6]:

$$w_2^{j+1} = w_2^j - \frac{w_2^j - w_2^{j-1}}{f(w_2^j) - f(w_2^{j-1})} f(w_2^j), \quad (17)$$

где $j = 1, 2, \dots$ – номер итерации.

Метод секущих является модификацией метода Ньютона. Суть модификации заключается в замене производной $f'(w_2^j)$ на конечно-разностную аппроксимацию $\frac{f(w_2^j) - f(w_2^{j-1})}{w_2^j - w_2^{j-1}}$. В процессе итераций но-

вое приближение w_2^{j+1} определяется двумя предыдущими значениями w_2^j и w_2^{j-1} [6].

На основе анализа совокупности A начальные значения w_2^{j-1} и w_2^j принимались равными 0 и 0,5. Критерий остановки итерационного процесса – $|f(w_2^j)| < 0,0001$. После нахождения w_2^0 значение w_3^0 равно:

$$w_3^0 = \frac{\ln(y_p + t_p^{w_2^0})}{\ln(t_p)}. \quad (18)$$

Таким образом, начальный вектор \mathbf{w}^0 для функции (4) определяется выражениями (11), (17), (18).

Для функции (5) длительность переходного процесса приведена к единому отрезку времени $t_i \in [0, 1]$. Коэффициенты $\mathbf{w}^0 = [1, -1.5, 1, 1.5, 0.5]$ подобраны экспериментально.

Далее вектор начального приближения \mathbf{w}^0 заменяется на вектор $\mathbf{w}^0 + \Delta\mathbf{w}$ и реализация итерационной формулы МЛМ имеет вид:

$$\mathbf{w}^{j+1} = \mathbf{w}^j + \Delta\mathbf{w}^j, \quad (19)$$

где $j = 1, 2, \dots$ – номер итерации.

Приращение $\Delta\mathbf{w}$ находится по формуле [2]:

$$\Delta\mathbf{w} = \left(J^T J + \lambda \cdot \text{diag}(J^T J) \right)^{-1} J^T (\mathbf{y} - \tilde{\mathbf{F}}), \quad (20)$$

где J – матрица Якоби; $\lambda \geq 0$ – параметр регуляризации; $\text{diag}(J^T J)$ – диагональная матрица с элементами $J^T J$; \mathbf{y} , $\tilde{\mathbf{F}}$ – векторы экспериментальных и регрессионных значений.

Матрица J составляется из частных производных функции $\tilde{F}_z(t_i, \mathbf{w})$ по всем компонентам вектора \mathbf{w} :

$$J = \begin{bmatrix} \frac{\partial \tilde{F}_z(t_1, \mathbf{w})}{\partial w_1} & \dots & \frac{\partial \tilde{F}_z(t_1, \mathbf{w})}{\partial w_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial \tilde{F}_z(t_n, \mathbf{w})}{\partial w_1} & \dots & \frac{\partial \tilde{F}_z(t_n, \mathbf{w})}{\partial w_k} \end{bmatrix}. \quad (21)$$

Для вычисления частных производных используется центральная конечно-разностная аппроксимация [6]:

$$\frac{\partial \tilde{F}_z(t_i, \mathbf{w})}{\partial w_k} = \frac{\tilde{F}_z(t_i, \mathbf{w} + \eta) - \tilde{F}_z(t_i, \mathbf{w} - \eta)}{2\eta}, \quad (22)$$

где η – конечно-разностный интервал.

Значение η определяется методом подбора для каждой конкретной функции. При отладке программной реализации МЛМ использо-

ваны значения $\eta_3 = 10^{-6}$, $\eta_4 = 10^{-6}$, $\eta_5 = 10^{-3}$ для функций (3), (4), (5) соответственно.

Матрица частных производных J транспонируется в J^T . Результатом произведения $J^T J$ является прямоугольная матрица размером $k \times k$. Сумма $J^T J + \lambda \cdot \text{diag}(J^T J)$ представляет собой матрицу Гессе H . Обращение H выполняется методом Гаусса–Жордана [7], он заключается в том, что если с единичной матрицей E провести элементарные преобразования, которыми невырожденная матрица H приводится к E , то получается обратная матрица H^{-1} .

Наличие параметра регуляризации λ является отличительной особенностью МЛМ от прочих методов многомерной оптимизации. Монотонное убывание минимизируе-

мой функции $S(\mathbf{w})$ достигается в МЛМ за счет подбора значений λ [6].

Критерием оценки качества аппроксимации и окончания итерации принимается средняя ошибка аппроксимации:

$$\bar{\varepsilon} = \frac{\sum |y_i - \tilde{F}_z(t_i, \mathbf{w})| / y_i}{n} 100\%.$$

При исследовании пяти реализаций (1... 5) с использованием функций (3), (4), (5) сделан вывод, что зависимость $\bar{\varepsilon} = f(\lambda)$ нелинейная (рис. 2). Различные значения $\bar{\varepsilon}$ на начальной стадии адаптивной коррекции объясняются особенностями (в частности, точностью) определения \mathbf{w}^0 .

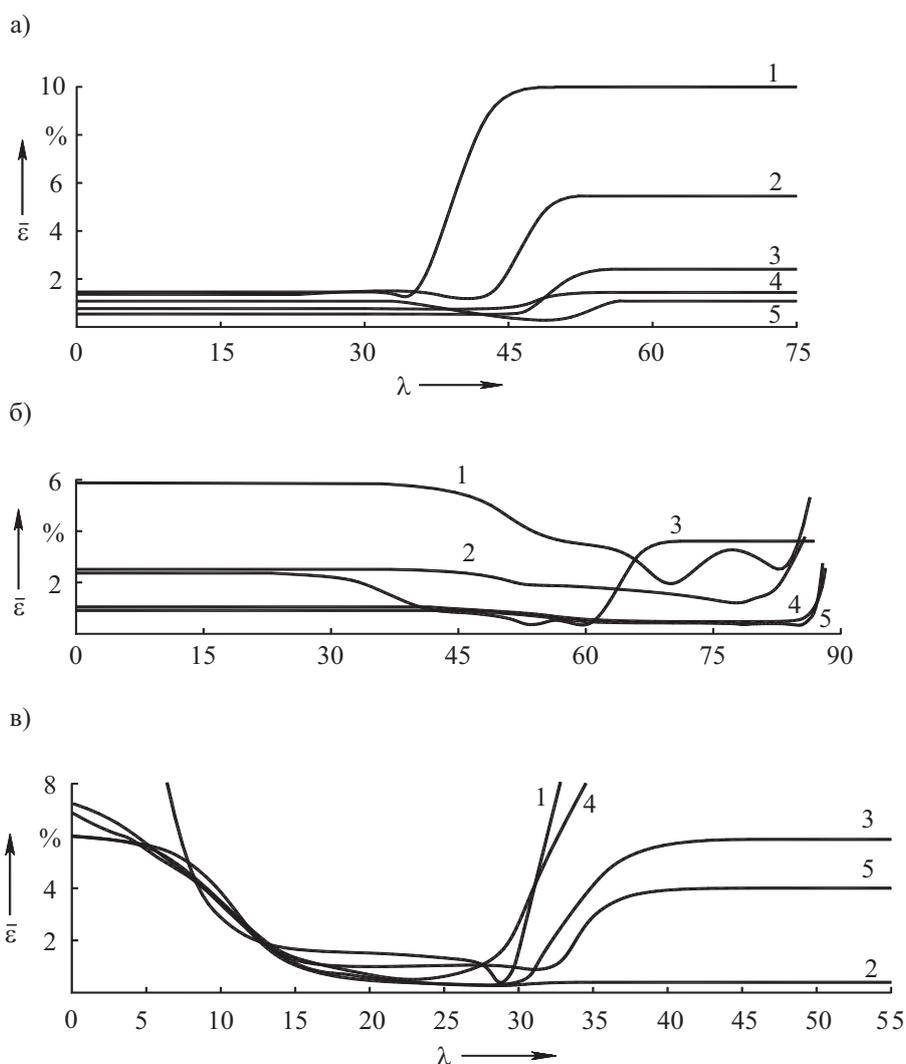


Рис. 2. Зависимость $\bar{\varepsilon}$ от коэффициента регуляризации (λ) для функций: а) – (3), б) – (4), в) – (5)

2 Адаптивная коррекция

$$\bar{\varepsilon}_1^j < \bar{\varepsilon}_2^j, \quad (24)$$

Цель адаптивной коррекции – найти значение λ на текущем шаге итерации, при котором $\bar{\varepsilon} \rightarrow \min$. Так как аналитическое выражение $\bar{\varepsilon} = f(\lambda)$ отсутствует, необходимо реализовывать алгоритм итерационного метода подбора соответствующего значения λ (рис. 3).

Для решения этой задачи начальное λ_i^0 принимается равным максимальному значению из элементов матрицы $J^T J$ (блок 7). Это приводит к тому, что МЛМ ведет себя как метод градиентного спуска с медленной сходимостью. При таких условиях начинается внешний цикл по j , вычисляются поправки $\Delta \mathbf{w}_1^j$, вектор \mathbf{w}_1^j (блок 11) и средняя ошибка аппроксимации (далее – СОА) $\bar{\varepsilon}_1$ (блок 12).

Полученное значение $\bar{\varepsilon}_1^j$ сравнивается с $\bar{\varepsilon}_0^j$ (блок 14). При увеличении СОА λ_i^0 последовательно уменьшается в 1,618 и повторно выполняются действия блоков 8–15.

Если в результате достигается состояние, при котором

$$\lambda_i^j < \tau \quad (23)$$

(где τ – заданное малое число), то программа завершается, \mathbf{w}_0^j считается лучшим приближением, а коэффициенты – найденными (блок 30). Выход из цикла по i осуществляется при выполнении условия $\bar{\varepsilon}_0^j > \bar{\varepsilon}_1^j$.

Цикл по k начинается после присвоения $\lambda_k^j = \lambda_i^j / 1,618$ (блок 16). При неверности неравенства (блок 23):

где $\bar{\varepsilon}_1^j$ – СОА при \mathbf{w}_1^j , которому соответствует λ_k^j ; $\bar{\varepsilon}_2^j$ – СОА при \mathbf{w}_2^j , которому соответствует $\lambda_k^j = \lambda_k^j / 1,618$; продолжается уменьшение параметра регуляризации (блок 25) для отыскания минимального значения $\bar{\varepsilon}_2^j$ (блоки 17–25). Уменьшение λ_k^j и окончание цикла по k происходит по двум критериям:

1) выполнение условия $\lambda_k^j < \tau$, при этом программа завершается, а \mathbf{w}_1^j считается лучшим приближением;

2) увеличение $\bar{\varepsilon}_2^j$ по сравнению с предыдущим $\bar{\varepsilon}_1^j$ (блок 23), при этом \mathbf{w}_1^j считается лучшим приближением на текущем j шаге и выполняются завершающие операции цикла j .

Критерием окончания внешнего итерационного процесса по j , то есть нахождения наилучшего приближения вектора w , является выполнение условия $|\bar{\varepsilon}_0 - \bar{\varepsilon}_1| < 10^{-6}$ в течение трех итераций (блок 27). В этом случае $\mathbf{w} = \mathbf{w}_1^j$ и МЛМ считается реализованным.

Заключение

Аппроксимация переходных процессов D функциями (3), (4), (5) методом наименьших квадратов с применением МЛМ обеспечивает получение значений искомых коэффициентов, при которых средняя ошибка аппроксимации не превышает 1,4%, (рис. 4). Такие значения меньше рекомендуемых предельных значений [8].

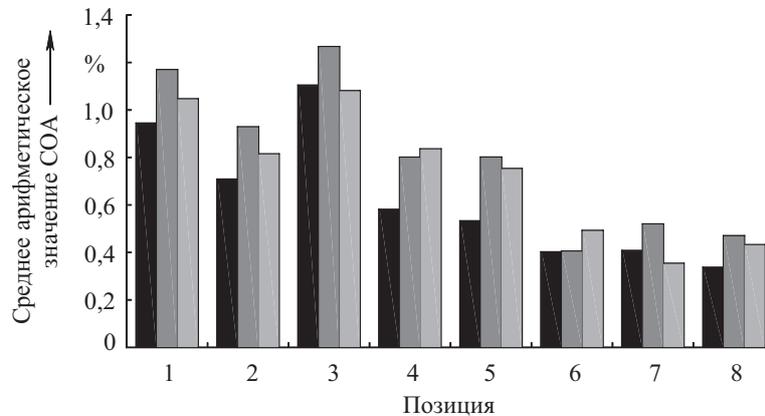


Рис. 3. Среднее значение СОА функции аппроксимации (3), (4), (5) в зависимости от номера позиции контроллера машиниста

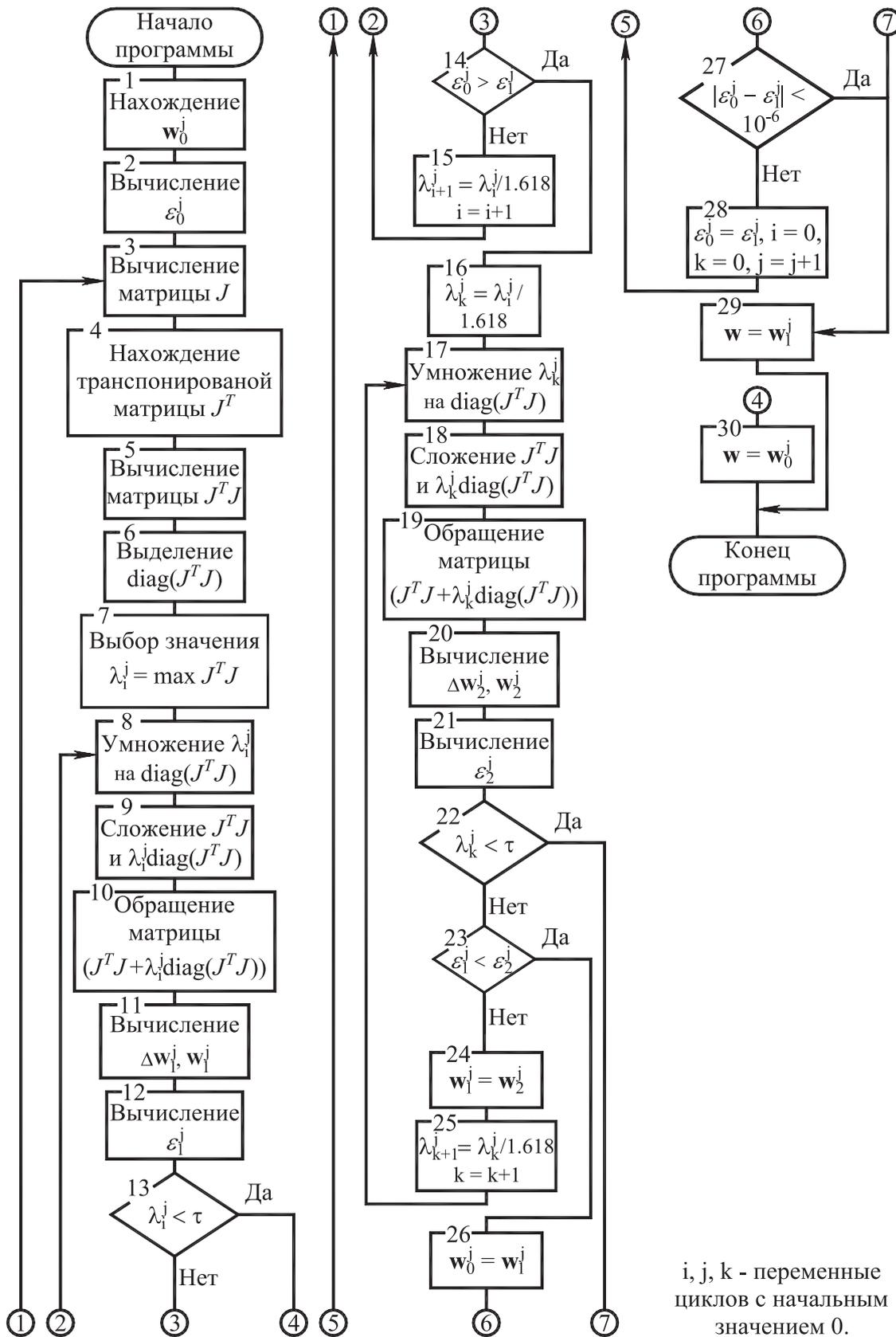


Рис. 4. Алгоритм реализации МЛМ

На позициях 1–5 контроллера машиниста заметна менее качественная аппроксимация, чем на позициях 6–8. Влияние выбросов и их конкретные причины можно установить после статистического анализа и проведения эксперимента.

В работе описана методика аппроксимации получаемых с бортовой системы АПК «Борт» данных изменения тока тягового генератора в режиме разгона тепловоза. Методика распространена на линейные и нелинейные аппроксимирующие функции. Полученные результаты свидетельствуют о приемлемой погрешности аппроксимации и возможности использования их для контроля технического состояния локомотивов.

Библиографический список

1. **Математическое** описание экспериментальных данных контроля технического состояния локомотивов / В. А. Четвергов, В. М. Боча-

ров, А. И. Мишин, О. В. Гателюк, П. А. Сиряк // Современные проблемы науки и образования. – 2013. – № 4. [Электронный ресурс]. – Режим доступа: <http://www.science-education.ru/110-9767>.

2. **Метод** Левенберга-Марквардта [Электронный ресурс]. – Режим доступа: <http://alglib.sources.ru/optimization/levenbergmarquardt.php>.

3. **Прикладная** статистика / С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин – Москва : Финансы и статистика, 1985. – 487 с., ил.

4. **Численные** методы безусловной оптимизации и решения нелинейных уравнений / Дж. мл. Деннис, Р. Шнабель ; пер. с англ. – Москва : Мир, 1988. – 440 с., ил.

5. **Основы** численных методов : учеб. пособие / Л. И. Турчак. – Москва : Наука, 1987. – 320 с.

6. **Практическая** оптимизация / Ф. Гилл, У. Мюррей, М. Райт ; пер. с англ. – Москва : Мир, 1985. – 509 с., ил.

7. **Метод** Гаусса-Жордана [Электронный ресурс]. – Режим доступа: <http://goo.gl/r1RtpQ>.

8. **Метод** наименьших модулей / В. И. Мудров, В. Л. Кушко. – Москва : Знание, 1971. – 61 с.

УДК 629.45/.46.02

Н. А. Чурков

Петербургский государственный университет путей сообщения
Императора Александра I

О ПЕРСПЕКТИВНОМ ГРУЗОВОМ ВАГОНЕ

Перевод экономики на рыночные отношения сформировал новые условия хозяйствования, которым должен удовлетворять и железнодорожный транспорт. Концепция грузовых вагонов, сформулированная в 20-х гг. XX в. и с тех пор не изменявшаяся, устарела, в связи с чем в статье поставлен вопрос о необходимости создания новой концепции. Показана целесообразность использования специализированных контейнеров, созданных на базе универсальных крупнотоннажных ISO контейнеров. Перевозку грузов в контейнерах предложено организовать в грузовых нерасцепляемых поездах, курсирующих по определенным направлениям и по расписанию, подобно пассажирским. В качестве перспективного грузового вагона предложена платформа для специализированных контейнеров.

товарные вагоны, перспективы развития, контейнеры, железнодорожный транспорт.

Введение

В 1991 г. наша страна – социалистическое общенародное государство трудящихся всех

наций и народностей СССР – была преобразована в Российскую Федерацию – федеративное государство с республиканской формой правления. В результате изменилось