

Восстановление параметров геополей с учетом шумов измерений. Ч. 2. Методы идентификации, регуляризации и мягких вычислений (soft computing)

А. М. ПАШАЕВ, Р. А. САДЫХОВ, Ф. Г. РАМАЗАНОВ, Ф. С. НАЗИРЛИ

Национальная академия авиации Азербайджана, Баку, Азербайджан, e-mail: sadixov@mail.ru

Проанализированы различные методы восстановления параметров геополей, учитывающие шумы измерений. Показаны возможности применения программных пакетов пользователями-неспециалистами в области теории программирования. Даны рекомендации по правильному выбору метода и его управляющих параметров для аппроксимации и моделирования геополей. Выявлены преимущества и недостатки этих методов на различных стадиях восстановления параметров геополей. Предложено на ранней стадии, когда объем информации обладает свойством неопределенности и ограниченности, воспользоваться нечеткими нейронными сетями.

Ключевые слова: цифровые модели геополей, методы обработки зашумленной измерительной информации, технология soft computing, нейронная сеть.

Various methods of geofield parameters restoration considering the measurements of the noises have been analyzed. Availability of programs for users without profound knowledge of theory and programming have been shown. Recommendation on correct method choice and its control parameters while preparing for approximation and modeling of geofields have been given. Advantages and disadvantages of these methods at different stages of geofield parameters restoration have been revealed. For instance, at an early stage fuzzy neural networks are more preferable if information content is uncertain and limited.

Key words: numerical geofield models, processing methods of noisy measuring information, soft computing technology, neural networks.

Различные методы восстановления параметров геополей с учетом шумов измерений (кусочно-полиномиальная аппроксимация, сглаживающие сплайны, непрерывный кригинг и др.) и без их учета (фильтры, интерполяционные сплайны, дискретный кригинг и др.) рассмотрены в [1, 2]. Выявлены преимущества и недостатки этих методов и даны рекомендации по правильному выбору метода и его управляющих параметров.

Ниже представлены результаты исследования, позволяющие пользователю-неспециалисту в области теории приближений и информационных технологий обоснованно выбрать метод оценивания (наименьших квадратов рекуррентной формы и с регуляризацией, нечетких нейронных сетей и др.), критерий качества, управляющие параметры и провести аппроксимацию поверхностей геополей на основе обработки зашумленной измерительной информации.

Под цифровой моделью геополя (ЦМГ) здесь будем понимать форму представления исходных данных и способ их структурного описания, позволяющий вычислять (восстанавливать) значения функции, определяющей поверхность, ее интерполированием в заданной области [1, 2] или аппроксимацией, т. е. путем восстановления поверхности с учетом шумов измерений (погрешностей как входных, так и выходных координат модели). При разных способах аппроксимации можно получать различные результаты, оцениваемые лишь с точки зрения практических приложений [2—24].

Разграничение задач интерполяции и аппроксимации вызвано тем, что чаще всего они алгоритмически различны по своему решению. Например, для обучающей выборки

$\left((x_1, y_1, \hat{h}_1), (x_2, y_2, \hat{h}_2), \dots, (x_n, y_n, \hat{h}_n) \right)$, которая генерируется неизвестной функцией, искаженной шумом, задача аппроксимации состоит в нахождении оценки этой функции.

Отметим, что ни одна из существующих моделей геополей не удовлетворяет в полной мере изложенным требованиям, и, по-видимому, не приходится ожидать создания универсальной ЦМГ, пригодной для любых поверхностей. Реальной перспективой является разработка моделей, включающих математический аппарат, учитывающий специфику задач. Поэтому целесообразно создавать автоматизированные банки данных различных методов восстановления геополей и их картирования, содержащие соответствующие программные пакеты аппроксимации (ППА) [1, 2].

В отличие от дедуктивного подхода [2] индуктивный подразумевает, что выбор функций, аппроксимирующих поверхности геополей, из некоторого существующего набора остается за пользователем.

В настоящее время для решения проблем, связанных с моделированием геополей, разработано и доступно достаточное количество различных ППА. В большинстве случаев они рассчитаны на пользователя, обладающего соответствующими знаниями в области обработки экспериментальных данных и программирования. В то же время существующие ППА не выдают требуемого предупреждения, если анализируемые данные плохо обусловлены или применяется не тот метод. Поэтому распространена практика самостоятельного выбора метода и его оптимальных параметров. К тому же ППА обладают многими общими функциями, но и во

многим разнятся. Эффективное решение — совместное использование этих программ, что требует от пользователя хороших знаний данных функций. При этом трудности определения оптимальных значений параметров тесно связаны с применяемыми методами регуляризации и фильтрации. Используя фильтры, приходится решать те же проблемы, что и при сглаживании и аппроксимации: какие частоты должен пропускать (удерживать) фильтр, как выбрать шаг фильтрации с наименьшей трудоемкостью вычислительных процедур и т. п.

В настоящее время все более широкое применение находит способ построения (идентификации) математической модели геополя на основе обработки текущей информации [1, 3, 5, 7, 8, 10, 19, 24]. Принцип работы алгоритмов идентификации этих моделей базируется на измерении входных и выходных координат и статистических характеристиках проведенных измерений, т. е. целенаправленная обработка текущей информации компенсирует недостаток априорной информации (одно из преимуществ статистических методов обработки).

Методы наименьших квадратов рекуррентной формы, учитывающие шумы измерений (фильтр типа Калмана). При восстановлении параметров геополей в предположении, что значения $\hat{h}_i, x_i, y_i, 1 \leq i \leq n$ известны точно, используют рекуррентный алгоритм оценивания (фильтр) линейной модели регрессии [1, 3—5]. Рассмотренный класс моделей важен, поскольку для определения регрессионного вектора можно применять хорошо известные [3], эффективные и простые методы оценивания. Однако использование метода наименьших квадратов (МНК) для оценки коэффициентов при погрешностях измерений входных параметров системы даже в линейной модели приводит к смещенным, неэффективным и несостоятельным оценкам.

Поскольку на практике \hat{h}_i, x_i, y_i определяются (измеряются или оцениваются) с погрешностью, т. е. имеются шумы измерений, то приведенный в [1, 3—5] фильтр для восстановления параметров геополей применять нельзя, так как он не учитывает погрешности входных координат и приводит к неточным (смещенным) результатам. Кроме того, фильтр будет давать неправильные оценки погрешностей этих координат.

В [6] разработаны итерационные алгоритмы, учитывающие погрешности измерений как выходных, так и входных координат системы. Однако вследствие сложности и громоздкости вычислительных процедур они не были практически реализованы. Для обработки данных в подобном случае целесообразно использовать методы конъюнктного анализа [5]. Выбор метода зависит от вида математической модели и априорной информации о значениях аргументов и функций.

Ниже предложен простой и эффективный рекуррентный алгоритм идентификации параметров геополя, учитывающий шумы измерений [7, 8]. Пусть ЦМГ описывается с помощью модели регрессии

$$h(x, y) = \sum_{j=0}^m \sum_{k=0}^{m-j} c_{jk} x^j y^k. \quad (1)$$

Для определения коэффициентов c_{jk} из (1) необходимо, чтобы число точек удовлетворяло неравенству

$$n \geq (m + 1)(m + 2) / 2. \quad (2)$$

Рассмотрим полином — частный случай (1) при $m = 2$ — уравнение регрессии второго порядка

$$h(x, y) = c_{00} + c_{10}x + c_{01}y + c_{20}x^2 + c_{11}xy + c_{02}y^2. \quad (3)$$

Для нахождения его коэффициентов должно выполняться условие (2), т. е. $n \geq 6$.

Представим (3) в виде линейной модели регрессии в векторной форме

$$H(k) = X^T(k) \theta, \quad (4)$$

где H — высота поля в точке поверхности с координатами (x, y) — выходной параметр системы; $\theta = \|c_{00}, c_{10}, c_{01}, c_{20}, c_{11}, c_{02}\|$ — вектор искомых (оцениваемых) коэффициентов; $X(k) = \|1, x(k), y(k), x^2(k), x(k)y(k), y^2(k)\|$ — вектор входных координат; k — номер итерации.

Входные и выходные координаты регистрируются измерительной аппаратурой. Уравнения их измерений имеют вид, соответственно

$$Z_H(k) = H(k) + \zeta_H(k);$$

$$Z_X(k) = X(k) + \zeta_X(k),$$

где ζ_H, ζ_X — случайные погрешности измерений, распределенные по гауссовому закону.

Линейную модель (4) представим как

$$\begin{aligned} \theta(k+1) &= \theta(k); \\ Z_H(k) &= X^T(k) \theta(k) + \zeta_H(k). \end{aligned} \quad (5)$$

В (5) кроме погрешности измерения выходной координаты $H(k)$ входит погрешность $\zeta_X(k)$ параметров входных координат $X(k)$. Тогда

$$\begin{aligned} Z_H(k) &= [X(k) + \zeta_X(k)]^T \theta(k) + \zeta_H(k) = \\ &= X^T(k) \theta(k) + \zeta_X^T(k) \theta(k) + \zeta_H(k) = X^T(k) \theta(k) + \zeta_{np}(k), \end{aligned}$$

где $\zeta_{np}(k) = \zeta_X^T(k) \theta(k) + \zeta_H(k)$ — приведенные погрешности, их дисперсия определяется выражением

$$D_{Hnp} = D_H + \theta^T D_X \theta, \quad (6)$$

D_X — дисперсионная матрица погрешностей входных координат.

Применив к модели (5) фильтр Калмана, с учетом (6) получим

$$\left. \begin{aligned} \hat{\theta}(k) &= \hat{\theta}(k-1) + K(k) [Z_H - X^T(k) \hat{\theta}(k-1)]; \\ K(k) &= \frac{D(k-1)X(k)}{D_H(k) + \hat{\theta}^T(k-1)D_X(k)\hat{\theta}(k-1) + X^T(k)D(k-1)X(k)}; \\ D(k) &= D(k-1) - \\ &- \frac{D(k-1)X(k)X^T(k)D(k-1)}{D_H(k) + \hat{\theta}^T(k-1)D_X(k)\hat{\theta}(k-1) + X^T(k)D(k-1)X(k)}. \end{aligned} \right\} (7)$$

Здесь $\hat{\theta}(0) = 0$; $K(k)$ — коэффициент усиления фильтра; $D(k)$ — дисперсионная матрица ошибок оценок.

Полученная система (7) представляет собой рекуррентную процедуру и позволяет оценить линейную модель (4) с учетом погрешности параметров входных координат $X(k)$. При этом шаг итерации k указывает порядковый номер проведенного эксперимента.

Следует отметить, что предлагаемый подход можно использовать лишь при достаточно простом характере приближаемой функции, поскольку в противном случае приходится конструировать полином слишком больших степеней. Это приводит к вычислительным трудностям, связанным с плохой обусловленностью систем нормальных уравнений высоких порядков, и неустойчивому поведению полинома между узлами и на всей плоскости [1, 2]. К тому же возможна значительная потеря информации из-за появления систематических отклонений $|h_i - \hat{h}_i|$ в экстремумах аномалий, особенно в тех случаях, когда амплитуда помехи велика, а размеры аномалии малы. Положение усугубляется на хаотических сетях, так как необходимая степень сглаживания может зависеть и от структуры расположения узлов.

Двухступенчатая идентификация параметров модели геополя. При решении задачи идентификации параметров модели (4) в определенных условиях при больших значениях n точек наблюдений могут возникнуть достаточно высокий уровень помех, плохая обусловленность информационной матрицы, коррелированность входных параметров.

В [4] показано, что относительное смещение оценки зависит не только от относительной погрешности измерения (наблюдения) переменных, но и от их коррелированности между собой. Поэтому в случае плохо обусловленной матрицы небольшие погрешности наблюдения переменных могут давать значительное смещение оценок коэффициентов модели. Тогда целесообразно использовать двухступенчатые методы оценивания. Для этого векторы X, θ разобьем на блоки

$$X^T = (X_{(1)}^T, X_{(2)}^T); \theta^T = (\theta_{(1)}^T, \theta_{(2)}^T),$$

где $X_{(1)}, \theta_{(1)} \in R^p$; $X_{(2)}, \theta_{(2)} \in R^q$; $p + q = n$, и (4) представим в виде

$$H(k) = X_{(1)}^T(k)\theta_{(1)} + X_{(2)}^T(k)\theta_{(2)}.$$

Предлагаемый двухступенчатый алгоритм оценивания параметров модели (4) заключается в следующем. Рассматривая сначала $X_{(2)}^T(k)\theta_{(2)}$ как помеху, оцениваем $\theta_{(1)}$, далее, зная $\hat{\theta}_{(1)}$ и полагая

$$\tilde{H}(k) = H(k) - X_{(1)}^T\theta_{(1)},$$

оцениваем $\theta_{(2)}$. Рекуррентные выражения для оценок $\hat{\theta}_{(1)}, \hat{\theta}_{(2)}$ даны в [9].

Отметим, что двухступенчатые оценки имеют меньшую среднюю квадратическую погрешность, чем оценки, полученные с применением МНК, который при этом носит регуляризующий характер.

Метод наименьших квадратов с регуляризацией. Система уравнений (4) может иметь множество различных решений, поскольку связь между выходными и входными координатами модели недостаточно жесткая из-за шумов в них, в связи с чем не обеспечивается единственность. Поэтому на входные параметры накладываются дополнительные ограничения. Регуляризация решения достигается благодаря минимизации функционала

$$E = |X\theta - H|^2 + \alpha^2|\theta|^2 =$$

$$= (X\theta - H)^T (X\theta - H) + \alpha^2\theta^T\theta = \theta^T X^T X \theta - \theta^T X^T H + H^T H + \alpha^2\theta^T\theta.$$

В отличие от стандартной процедуры МНК, решение при таком подходе будет иметь вид (с учетом условия $\partial E / \partial \theta^T = 0$)

$$\hat{\theta} = (X^T X + \alpha^2 I)^{-1} (X^T Z_H), \quad (8)$$

где I — единичная матрица.

Выбор параметра α регуляризации зависит от уровня шумов в данных. Присутствие α гарантирует существование обратной матрицы и препятствует получению сильно зашумленного решения.

На практике матрицу $(X^T X + \alpha^2 I)$ из правой части (8) непосредственно не обращают. Итеративный способ типа Гаусса—Зейделя обычно обеспечивает достаточную сходимость за 4—5 итераций. При этом время вычислений по сравнению с требуемым для непосредственного обращения матрицы сокращается в несколько раз.

Отметим, что при плохой обусловленности системы (4) в [10] предложено воспользоваться аналогичным подходом (метод ридж-регрессии или гребневая регрессия).

Достоинством МНК с регуляризацией является простота его вычислительной реализации, а недостатком — отсутствие четких критериев в выборе наилучшего значения параметра α регуляризации.

В целях автоматизации процессов решения задач восстановления необходимо выработать некоторые формальные правила, применяемые без участия человека. В этом случае целесообразно использовать методы теории нечетких множеств [12] (субъективизм проявляется при выборе вида функций принадлежности). Нечеткий подход обеспечивает большую гибкость и учет множества нюансов при

формализации априорной информации, поскольку возможности варьирования вида функций принадлежности намного богаче, чем возможности варьирования одного параметра α .

Статистические методы регуляризации. В оптимальных методах фильтрации Калмана [13] и Колмогорова—Винера [14] среди известных устойчивых методов (в отличие от детерминистских) используется наибольшее количество априорной информации о решении. В методе Калмана — ковариация ошибок и математического ожидания правой части и решения, а Колмогорова—Винера — спектральные плотности шума правой части и решения, например, интегрального уравнения 1-го рода типа свертки.

Сравним методы Калмана (одношаговый) [13] и Тихонова (метод регуляризации 0-го порядка) [11]. Если положить $\alpha = \delta^2 / \varepsilon^2$, где ε, δ — априорные средние квадратические погрешности решения и правой части, то эти методы дают одинаковые решения. Фильтр Калмана обычно применяется лишь для многошаговых процессов, а методы детерминистской регуляризации эффективны, в первую очередь, для одношаговых процессов.

Сравнение методов Тихонова и Колмогорова—Винера показывает, что их регуляризованные решения переходят одно в другое [2].

Группа методов Калмана и Колмогорова—Винера может обеспечить высокую эффективность решения (меньшую неопределенность, большую точность) по сравнению с группой детерминистских методов, однако требуемая для реализации этих методов исходная информация (например, ковариация ошибок решения) часто неизвестна. Вместе с тем, входящий в метод Тихонова параметр α регуляризации однозначно определяется по значению погрешности правой части способом невязки.

Неустойчивость всегда обуславливается тем, что решение задачи ищется в слишком широком классе. Ее можно восстановить некоторым сужением класса возможных решений благодаря привлечению априорной информации качественного или количественного характера. В результате получаются различные алгоритмы регуляризации [11]. Таким образом, разные методы решения некорректных (по Адамару) задач отличаются видом априорной информации.

Как известно [3, 7, 10, 13, 15], в геонаучных исследованиях для аппроксимации и моделирования геополей широко применяются методы регрессионного анализа. Однако при практическом использовании этих методов обнаруживается, что отсутствуют объективные критерии выбора наилучшего результата количественной обработки экспериментальных данных, полученных с нерегулярной сети наблюдений. В большинстве случаев регрессионная задача принадлежит к классу некорректно поставленных. Их можно регуляризовать за счет экзогенной информации. Но выбор параметра регуляризации является трудноформализуемой процедурой, имеющей субъективный характер. Появление же новых более современных алгоритмов [3, 7, 10, 13, 15] (метод всех регрессий, шаговый, ступенчатый, Гарсайда и др.) также не позволяет получать единственный и однозначный ответ. Поэтому распространена практика, когда пользователю предлагается самому (т. е. субъективно) находить наилучшее решение из множества. Для устранения субъективизма в результатах существует большое число различных и часто взаимоисключающих рекомендаций [10].

Необходимо создавать регуляризирующие помехоустойчивые алгоритмы идентификации математических моделей геополей, так как из-за неизбежных погрешностей измере-

ний (шумов как входных, так и выходных координат модели) появляется неустойчивость решения. Регуляризирующие помехоустойчивые алгоритмы оказываются особенно ценными, поскольку анализ экспериментальных данных, по сути дела, и есть анализ шумов.

Применение вероятностно-статистических методов требует проведения предварительного анализа данных на нормальность выборочного распределения, что предполагает выполнение некоторых условий [1]. Проведенный анализ результатов исследований [15], где используются аналогичные ЦМГ, показал, что выборочное распределение противоречит предположению о нормальности [1].

В [4] показано, что в отсутствие предположения о нормальности статистические процедуры, основанные на моментах 1-го и 2-го порядков, быстро теряют свои оптимальные свойства. Там же приведены устойчивые методы оценивания коэффициентов линейного и полиномиального уравнений регрессии для распределения, отличного от нормального. Однако эти методы применяются, когда имеется некоторая априорная информация, например, точно известны класс распределения или его близость к нормальному, или при небольших значениях погрешностей и т. д. Многие из этих методов являются видоизмененными стандартными процедурами оценивания коэффициентов уравнения регрессии. Относительно их управляющих параметров делаются лишь некоторые предположения. Таким образом, указанные алгоритмы не универсальны, зависят от специфики обрабатываемых экспериментальных данных.

Проблема становится еще более трудоемкой с увеличением сложности объекта, особенно на ранней стадии моделирования геополей, когда объем информации не только ограничен и не определен, но и обладает свойством нечеткости (входные и выходные координаты системы изменяются в определенных диапазонах, а их значения измеряются с определенной погрешностью). В связи с этим возникла необходимость идентификации параметров математической модели многомерного нечеткого объекта, описываемого нечетким регрессионным уравнением

$$\tilde{H}_m = \sum_{j=0}^m \sum_{k=0}^{m-j} \tilde{c}_{jk} \otimes \tilde{x}^j \otimes \tilde{y}^k; \quad j=0, m; k=0, m; j+k \leq m, \quad (9)$$

где \tilde{c}_{jk} — искомые нечеткие параметры; \otimes — знак нечеткого умножения.

Ставится задача определения нечетких значений параметров \tilde{c}_{jk} уравнения (9) на основе статистических нечетких экспериментальных данных, т. е. входных \tilde{x} , \tilde{y} и выходных \tilde{H} координат модели. Рассмотрим решение этой задачи с использованием нечеткой математики и искусственных нейронных сетей [17—19].

Нечеткий нейросетевой метод. Нечеткая нейронная сеть состоит из соединенных между собой наборов нечетких нейронов. При ее использовании для решения нечеткого регрессионного уравнения (9) входными сигналами сети являются нечеткие значения переменной $\tilde{V} = (\tilde{x}, \tilde{y})$, а выходным — значение \tilde{H} . В качестве параметров сети выступают нечеткие значения параметров \tilde{c}_{jk} .

При решении задачи идентификации параметров \tilde{c}_{jk} из (9) с применением нечеткой нейронной сети основной проблемой становится их обучение. Для обучения нечетких значений параметров воспользуемся α -срезом [17—19].

Нечеткие переменные будем представлять в виде нечетких чисел LR-типа, например в треугольной форме, функции принадлежности которых вычисляются по формуле

$$\mu(x) = \begin{cases} 1 - (x - \bar{x})/\alpha, & \text{если } \bar{x} - \alpha < x < \bar{x}, \alpha > 0; \\ 1 - (\bar{x} - x)/\beta, & \text{если } \bar{x} < x < \bar{x} + \beta, \beta > 0; \\ 0. & \end{cases} \quad (10)$$

Допустим, имеются экспериментальные статистические нечеткие входные и выходные данные, из которых составляем пары (\tilde{B}, \tilde{T}) для обучения нечеткой нейронной сети. Для построения модели процесса на вход нечеткой нейронной сети подаются входные сигналы \tilde{B} (рис. 1), а ее выходные сигналы сравниваются с эталонными \tilde{T} .

Нечеткая ошибка (отклонение) между текущим нечетким выходом нечеткой нейронной сети \tilde{H}_i и желаемым нечетким выходом \tilde{T}_i определяется как

$$\tilde{E} = \frac{1}{2} \sum_{i=1}^l (\tilde{H}_i \ominus \tilde{T}_i)^2, \quad (11)$$

где \ominus — знак нечеткого вычитания.

С применением α -среза для левой и правой частей отклонения вычисляются по формулам

$$E_1 = \frac{1}{2} \sum_{i=1}^l [h_{i1}(\alpha) - t_{i1}(\alpha)]^2;$$

$$E_2 = \frac{1}{2} \sum_{i=1}^l [h_{i2}(\alpha) - t_{i2}(\alpha)]^2; \quad E = E_1 + E_2,$$

где $\tilde{H}_i(\alpha) = [h_{i1}(\alpha), h_{i2}(\alpha)]$; $\tilde{T}_i(\alpha) = [t_{i1}(\alpha), t_{i2}(\alpha)]$.

Если для всех обучающих пар отклонения E меньше заданных, то обучение (коррекция) параметров сети заканчивается (рис. 2). В противном случае оно продолжается до тех пор, пока значение E не достигнет минимума.

Таким образом, алгоритм обучения нечеткой нейронной сети заключается в минимизации (11),

т. е. в приближении \tilde{E} к нулю. Однако в (11) используется нечеткое вычитание, поэтому даже при $\tilde{H}_i = \tilde{T}_i$ для всех $i = \overline{1, l}$ \tilde{E} не будет равно нулю.

Следовательно, надо уточнить правило останова вычислений в обучении нечеткой нейронной сети.

Представим, что носителем нечеткого множества \tilde{H}_i является интервал $[h_{i1}, h_{i2}]$, $1 \leq i \leq l$, т. е.

$\tilde{H}_i(0) = [h_{i1}, h_{i2}]$. Тогда при $\tilde{H}_i = \tilde{T}_i$ носителем для

\tilde{E} будет интервал $[-\lambda, \lambda]$, т. е. $\tilde{E}(0) = [-\lambda, \lambda]$, где

$$\lambda = \frac{1}{2} \sum_{i=1}^l (t_{i1} - t_{i2})^2.$$

Если $\varepsilon > 0$ — приемлемое отклонение от \tilde{E} (приемлемое качество обучения) при $\tilde{H}_i = \tilde{T}_i$, то правило останова может быть сформулировано следующим образом: итерации по поиску \tilde{c}_{jk} необходимо прекратить, когда \tilde{E} находится внутри множества

$$\Omega = [-\lambda - \varepsilon, \lambda + \varepsilon] \times [0, 1].$$

Коррекция параметров сети для левой и правой частей осуществляется согласно выражениям

$$c_{jk1}^n = c_{jk1}^0 + \gamma \frac{\partial E}{\partial c_{jk}}; \quad c_{jk2}^n = c_{jk2}^0 + \gamma \frac{\partial E}{\partial c_{jk}}, \quad (12)$$

где $c_{jk1}^0, c_{jk1}^n, c_{jk2}^0, c_{jk2}^n$ — предыдущие и последующие значения левой и правой частей параметров нечеткой нейронной сети; $c_{jk} = [c_{jk1}, c_{jk2}]$; γ — скорость обучения.

Следует отметить, что впервые нечеткие множества в нейронных сетях были рассмотрены в [17]. Обзор исследований по объединению нечетких множеств и нейронных сетей приведен в [18].

Численный пример 1. Рассмотрим модель, описываемую нечетким уравнением регрессии вида (частный случай (9) при $m = 2$)

$$\tilde{H} = \tilde{c}_{00} + \tilde{c}_{10}\tilde{x} + \tilde{c}_{01}\tilde{y} + \tilde{c}_{20}\tilde{x}^2 + \tilde{c}_{11}\tilde{x}\tilde{y} + \tilde{c}_{02}\tilde{y}^2. \quad (13)$$

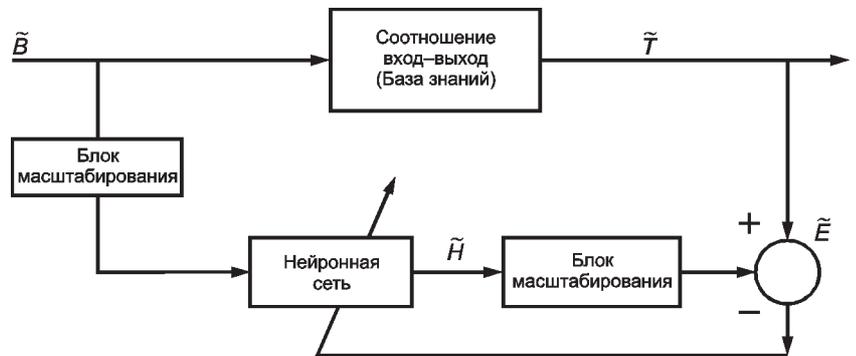


Рис. 1. Схема нечеткой нейронной идентификации

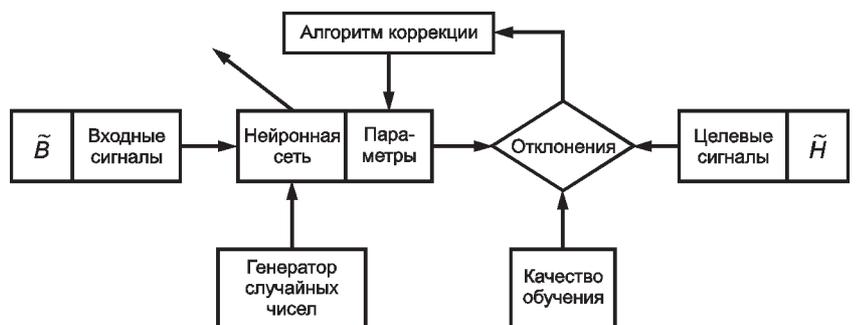


Рис. 2. Схема обучения с обратной связью нечетких параметров (порогов, весов) сети

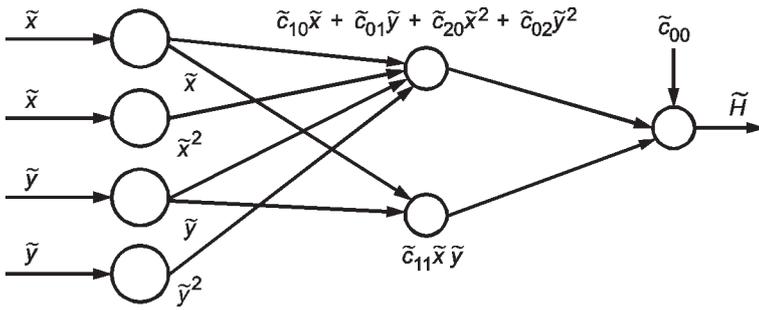


Рис. 3. Структура нечеткой нейронной сети для нечеткого уравнения регрессии 2-го порядка

Построим нейронную структуру для решения (13), где в качестве параметров сети выступают нечеткие коэффициенты $\tilde{c}_{00}, \tilde{c}_{10}, \tilde{c}_{01}, \tilde{c}_{20}, \tilde{c}_{11}, \tilde{c}_{02}$. При этом структура сети будет иметь четыре нечетких входа и один нечеткий выход (рис. 3). Осуществим обучение параметров сети с использованием (12). Для $\alpha = 0$ получим выражения

$$\begin{aligned} \frac{\partial E_1}{\partial c_{001}} &= \sum_{i=1}^l (h_{i1} - t_{i1}); & \frac{\partial E_2}{\partial c_{002}} &= \sum_{i=1}^l (h_{i2} - t_{i2}); \\ \frac{\partial E_1}{\partial c_{101}} &= \sum_{i=1}^l (h_{i1} - t_{i1}) x_{i1}; & \frac{\partial E_2}{\partial c_{102}} &= \sum_{i=1}^l (h_{i2} - t_{i2}) x_{i2}; \\ \frac{\partial E_1}{\partial c_{011}} &= \sum_{i=1}^l (h_{i1} - t_{i1}) y_{i1}; & \frac{\partial E_2}{\partial c_{012}} &= \sum_{i=1}^l (h_{i2} - t_{i2}) y_{i2}; \\ \frac{\partial E_1}{\partial c_{201}} &= \sum_{i=1}^l (h_{i1} - t_{i1}) x_{i1}^2; & \frac{\partial E_2}{\partial c_{202}} &= \sum_{i=1}^l (h_{i2} - t_{i2}) x_{i2}^2; \\ \frac{\partial E_1}{\partial c_{111}} &= \sum_{i=1}^l (h_{i1} - t_{i2}) x_{i1} y_{i1}; & \frac{\partial E_2}{\partial c_{112}} &= \sum_{i=1}^l (h_{i2} - t_{i2}) x_{i2} y_{i2}; \\ \frac{\partial E_1}{\partial c_{021}} &= \sum_{i=1}^l (h_{i1} - t_{i1}) y_{i1}^2; & \frac{\partial E_2}{\partial c_{022}} &= \sum_{i=1}^l (h_{i2} - t_{i2}) y_{i2}^2. \end{aligned} \quad (14)$$

При значениях параметров $\tilde{c}_{jk} < 0$ формулы вычисления, включающие \tilde{c}_{jk} из (13) и их коррекцию (14), будут изменять свою форму. Например, допустим $\tilde{c}_{11} < 0$, тогда формула вычисления пятого выражения (14), которая содержит \tilde{c}_{11} из (13), будет иметь вид $h_{51} = c_{111} x_2 y_2$; $h_{52} = c_{112} x_1 y_1$, а формулы коррекции запишутся как

$$\frac{\partial E_1}{\partial c_{111}} = \sum_{i=1}^l (h_{i1} - t_{i1}) x_2 y_2; \quad \frac{\partial E_2}{\partial c_{112}} = \sum_{i=1}^l (h_{i2} - t_{i2}) x_1 y_1.$$

Для $\alpha = 1$ имеем формулы коррекции

$$\frac{\partial E_3}{\partial c_{003}} = \sum_{i=1}^l (h_{i3} - t_{i3}); \quad \frac{\partial E_3}{\partial c_{103}} = \sum_{i=1}^l (h_{i3} - t_{i3}) x_3;$$

$$\begin{aligned} \frac{\partial E_3}{\partial c_{013}} &= \sum_{i=1}^l (h_{i3} - t_{i3}) y_3; & \frac{\partial E_3}{\partial c_{113}} &= \sum_{i=1}^l (h_{i3} - t_{i3}) x_3 y_3; \\ \frac{\partial E_3}{\partial c_{203}} &= \sum_{i=1}^l (h_{i3} - t_{i3}) x_3^2; & \frac{\partial E_3}{\partial c_{023}} &= \sum_{i=1}^l (h_{i3} - t_{i3}) y_3^2. \end{aligned} \quad (15)$$

В результате обучения по (14), (15) находим параметры сети, удовлетворяющие базе знаний с требуемым качеством обучения.

Перед проведением компьютерного моделирования в результате экспериментов были собраны нечеткие статистические данные (табл. 1), которые учитывают диапазоны изменения значений входных и выходного параметров модели. Используя построенную структуру нечеткой нейронной сети и нечеткие экспериментальные обучающие данные, проведем обучение нечетких параметров сети:

$$\begin{aligned} \tilde{c}_{00} &= (1,4124; 1,4223; 1,4275); & \tilde{c}_{10} &= (1,9884; 2,1131; 2,2339); \\ \tilde{c}_{01} &= (-2,5353; -2,5349; -2,5326); \\ \tilde{c}_{20} &= (-1,1043; -1,1042; -1,1036); \\ \tilde{c}_{11} &= (-0,8845; -0,8741; -0,8639); \\ \tilde{c}_{02} &= (1,3158; 1,3162; 1,3166). \end{aligned} \quad (16)$$

Эти данные получены при максимальной абсолютной погрешности, равной 0,09. Сравним (16) с результатами, полученными при использовании четкой модели (3) и МНК [15]:

$$\begin{aligned} \hat{c}_{00} &= 1,45; & \hat{c}_{10} &= 2,26; & \hat{c}_{01} &= -2,55; & \hat{c}_{20} &= -1,28; \\ & & \hat{c}_{11} &= -0,85; & \hat{c}_{02} &= 1,33. \end{aligned}$$

Отметим, что применение нечетких нейронных сетей в решении проблем, связанных с идентификацией параметров математических моделей различных по своей природе геополей, имеет определенные преимущества по сравнению с традиционными вероятностно-статистическими подходами. Прежде всего, это обусловлено тем, что предложенную методику можно использовать независимо от вида выборочного распределения параметров ЦМГ. При моделировании геополей возможности нейронных сетей усиливаются вследствие применения технологии обработки информации, основанной на нечетких множествах, — soft computing. Предлагаемый подход понижает требования к точности как входных, так и выходных координат модели в процессе ее построения и становится универсальным инструментом для создания ЦМГ.

Наиболее уязвимый элемент нечетких множеств — произвольное назначение вида функции принадлежности. Выше был указан класс функций принадлежности (10), достаточный при решении практических задач, для которого выделен подкласс нечетких чисел LR-типа, достаточно обширный для описания ситуаций, возникающих при восстановлении параметров геополей с учетом погрешностей измерений входных и выходных координат системы, и дающий возможность просто вводить бинарные арифметические операции.

Нечеткие экспериментальные статистические данные

\tilde{x}	\tilde{y}				
	3; 7; 11	17; 21; 25	31; 35; 39	45; 49; 53	59; 63; 67
28; 31; 35	0,77; 0,81; 0,85	1,08; 1,13; 1,17	1,28; 1,33; 1,44	1,43; 1,47; 1,51	1,49; 1,53; 1,57
50; 54; 58	0,48; 0,52; 0,56	0,68; 0,72; 0,76	0,81; 0,85; 0,89	0,89; 0,93; 0,97	0,93; 0,97; 1,01
68; 72; 76	0,37; 0,41; 0,45	0,53; 0,57; 0,61	0,63; 0,67; 0,71	0,69; 0,73; 0,77	0,72; 0,76; 0,80
82; 86; 90	0,30; 0,34; 0,38	0,43; 0,47; 0,51	0,52; 0,58; 0,60	0,57; 0,61; 0,65	0,60; 0,64; 0,68
92; 96; 100	0,27; 0,31; 0,35	0,39; 0,43; 0,47	0,46; 0,50; 0,54	0,51; 0,55; 0,59	0,54; 0,58; 0,62
96; 100; 104	0,23; 0,27; 0,31	0,34; 0,38; 0,42	0,41; 0,45; 0,49	0,46; 0,40; 0,54	0,47; 0,51; 0,55

Если два нечетких числа LR-типа обозначить $\tilde{m}=(m, \alpha, \beta)$ и $\tilde{n}=(n, \gamma, \delta)$, то нечеткие сложение и вычитание определяются формулами

$$\tilde{m} \oplus \tilde{n} = (m + n, \alpha + \gamma, \beta + \delta); \ominus \tilde{m} = (-m, \beta, \alpha);$$

$$\tilde{m} \ominus \tilde{n} = (m - n, \alpha + \delta, \beta + \gamma).$$

Вследствие того, что нечеткие числа LR-типа не образуют группу по умножению, операции умножения и деления не приводят к числам того же типа, тогда можно использовать простые приближенные формулы

$$\tilde{m} \otimes \tilde{n} \approx (mn, m\gamma + n\alpha, m\delta + n\beta); (\tilde{m})^{-1} \approx (m^{-1}, \beta m^{-2}, \alpha m^{-2}),$$

$$\tilde{m}, \tilde{n} > 0, \quad (17)$$

т. е. носители \tilde{m}, \tilde{n} расположены на полуоси $(0 + \infty)$,

$$\tilde{m} \oplus \tilde{n} \approx (mn^{-1}, (\delta m + \alpha n)n^{-2}, (\gamma m + \beta n)n^{-2}). \quad (18)$$

Формулы (17), (18) зависят от знаков операндов. При этом носители нечетких чисел \tilde{m}, \tilde{n} не должны содержать ноль.

Отметим, что помимо операций с числами LR-типа существуют и другие методы выполнения операций с нечеткими числами. Среди них есть методы, основанные на использовании уровненых множеств, отличающиеся значительным упрощением вычислений по сравнению с операциями на базе принципа обобщения [20].

Таким образом, возникла необходимость разработки ППА с использованием технологии soft computing ППА, позволяющая эксперту (пользователю-неспециалисту) без специальных знаний в области математики и информационных технологий работать на компьютере в интерактивном режиме как при формировании, так и при поиске решений трудноформализуемых задач моделирования геополей. Причем в подобных системах разумно сочетаются достоинства интерактивных систем с эвристическими правилами (стратегия здравого смысла и интуитивного суждения, базирующаяся на теоретических и экспериментальных знаниях), вследствие чего они приобретают ряд важных свойств: представление знания экспертов на языке, близком к естественному, и обоснование правильности принятого решения.

Применяемый алгоритм обратного распространения ошибки (ОРО) — это самый распространенный алгоритм обучения, при реализации которого необходимо помнить следующее:

при минимизации средней квадратической погрешности (ошибки) в течение эпохи (один полный цикл предъявления набора примеров обучения) может произойти преждевременный останов, поскольку интенсивность ее изменения — достаточно малая величина (0,01—1 %);

для предотвращения застревания ОРО в нежелательных локальных минимумах можно увеличить количество скрытых нейронов или использовать методы тяжелого шарика, модельной закалки, инерции [22], последний более предпочтителен, так как позволяет быстрее проходить плоские участки;

обучение сети можно рассматривать как задачу оптимальной фильтрации (фильтр Калмана) [21, 23]. Рекуррентный МНК — фильтр Калмана работает на порядок быстрее ОРО;

процесс обучения можно обобщить, добавив обучение по подсказке (априорную информацию о неизвестной функции — восстанавливаемой поверхности) [24].

Графические модели геополей (карты) являются определенной формой представления исходных данных (цифровая картографическая информация) — расположением точек:

регулярным на прямоугольных, треугольных или геоксagonalных сетках;

нерегулярным по структурным линиям, профилям, центрам площадей, локальным точкам, случайным точкам и т. д.;

полурегулярным — изолинейным (уровневым), учитывающим сложность изолинейного рисунка.

В повседневной практике геодезических, геологических, геофизических, картографических и других работ регулярно приходится строить различные карты изолиний: изобар, изогипс, изокон, изопохит, изотерм, изохрон и пр. Поэтому основным картографическим средством визуализации реальных геополей служат изолинии, хорошо передающие плавность, постепенность изменения количественных признаков и позволяющие проводить их качественный анализ.

Ниже для построения изолиний предлагаем воспользоваться новым методом (о других методах см. [16]), основанном на фундаментальной теореме Пуанкаре «О возвращаемости» замкнутой физической системы, ранее предложенным в [2]. Для этого функцию Гамильтона представим в виде

$$H(x, y) = \frac{1}{2}(u_1, u_2) + R(h(x, y) - h_0(x, y))^2,$$

где $h_0(x, y)$ — заданное значение восстанавливаемого параметра в некоторой точке x^0, y^0 ; $u = (u_1, u_2)$ — скорость движения.

Тогда задача построения изолиний сводится к решению системы дифференциальных уравнений

$$\left. \begin{aligned} dx/dt &= u_1; \\ dy/dt &= u_2; \\ du_1/dt &= -2R(h(x, y) - h_0(x, y))\partial h(x, y)/\partial x; \\ du_2/dt &= -2R(h(x, y) - h_0(x, y))\partial h(x, y)/\partial y. \end{aligned} \right\} \quad (19)$$

Эта система решается методом Рунге—Кутты 4-го порядка с начальным условием

$$M_0 = (x^0, y^0, u_1^0, u_2^0) \quad (20)$$

на отрезке $[0, 1]$ с шагом 0,01, где $u^0 = (u_1^0, u_2^0)$ — произвольный вектор; (x^0, y^0) удовлетворяет условию $h(x^0, y^0) = h_0$; $h(x, y)$, предварительно находят одним из рассмотренных выше методов; значения $h_0(x, y)$ задает картограф.

Итак, найдя решение (19), (20) с помощью точек (x_i, y_i) , $i = \overline{1, 101}$, можно нарисовать изолинию со значением h_0 в любой ее точке. Для зарисовки изолиний с уменьшенным количеством расчетных точек можно воспользоваться кубическими сплайнами [16].

Численный пример 2. Для иллюстрации возможностей предлагаемого метода рассмотрим задачу построения единичной окружности $x^2 + y^2 = 1$, т. е. $h(x, y) = x^2 + y^2$, $h_0 = 1$. Имеем

$$\begin{aligned} dx/dt &= u_1; \\ dy/dt &= u_2; \\ du_1/dt &= -4R(x^2 + y^2 - 1)x; \\ du_2/dt &= -4R(x^2 + y^2 - 1)y; \end{aligned}$$

$$(x^0, y^0, u_1^0, u_2^0) = (0, 1, 1, 0).$$

Выберем начальную точку

$$M_0 = \left(0, 1, 1, -\frac{\partial h(x, y)}{\partial x} \Big|_{x=0, y=1} / \frac{\partial h(x, y)}{\partial y} \Big|_{x=0, y=1} \right).$$

Результаты численных расчетов приведены в табл. 2, где \bar{x}, \bar{y} — расчетные значения координат единичной окружности.

Как следует из табл. 2, отклонение $E = \left| \bar{x}^2 + \bar{y}^2 - 1 \right| (x^2 + y^2)^{-1}$ очень мало, что подтверждает эффективность данного метода. Его можно использовать для построения изолиний различных геополей.

Результаты расчета координат единичной окружности

Номер точки	\bar{x}	x	\bar{y}	y	E
1	0,836	0,8	0,551	0,6	0,02
2	0,921	0,9	-0,395	-0,43	0,03
3	0,481	0,48	-0,921	-0,87	0,04
4	-0,727	-0,73	-0,693	-0,7	0,02
5	-0,979	-0,96	0,225	0,25	0,04
6	-0,351	-0,35	0,942	0,92	0,03
7	0,583	0,58	0,812	0,8	0,001
8	0,98	0,97	-0,276	-0,251	0,03
9	0,514	0,5	-0,865	-0,861	0,01
10	-0,43	-0,4	-0,901	-0,915	0,03

Отметим, что при построении карт необходимо проводить проверку детерминированной компоненты картируемого параметра:

если детерминированная компонента отсутствует, построение карты не имеет смысла, поскольку в этом случае оценка дисперсии отклонений прогнозируемых параметров

от точных $D(h_i - \hat{h}_i)$ будет примерно равна исходной, определяемой по наблюдаемым значениям Dh_i ;

если применяется метод аппроксимации, не имеющий эффективной регуляризации и дающий неустойчивые результаты, дисперсия может существенно возрасти, и никакие процедуры восстановления функции не смогут улучшить получаемые оценки.

Выводы. Из приведенного анализа методов восстановления параметров геополей с учетом шумов измерений (МНК рекуррентной формы — фильтр Калмана и МНК с регуляризацией, двухступенчатый метод идентификации, фильтры Калмана и Колмогорова—Винера, нечеткие нейронные сети) на различных стадиях выявлены их преимущества и недостатки. Рекомендуется на ранней стадии восстановления параметров геополей, когда объем информации обладает свойствами неопределенности и ограниченности, воспользоваться нечеткими нейронными сетями. Для повышения скорости их обучения предлагается объединить сеть с фильтром Калмана.

В связи с тем, что существуют различные методы (учитывающие погрешности входных и выходных координат модели) аппроксимации поверхностей геополей, а также программные пакеты аппроксимации, которые позволяют неспециалисту проводить необходимые расчеты для восстановления параметров геополей, правильный выбор метода и его управляющих параметров является одним из основных этапов при подготовке к моделированию поверхностей и картированию параметров геополей. При этом рекомендуется совместное использование этих программных пакетов, т. е. экспорт результатов из одного в другой.

Л и т е р а т у р а

1. Пашаев А. М. и др. Восстановление параметров геополей // Измерительная техника. 2005. № 12. С. 3—9; Pashaev A. M. e. a. Geofield parameter recovery // Measurement Techniques. 2005. V. 48. N 12. P. 1149—1158.

2. Пашаев А. М. и др. Восстановление параметров геоплей с учетом шумов измерений. Ч. 1. Методы сглаживания, фильтрации и регуляризации // Измерительная техника. 2010. № 8. С. 3—11.
3. Льюинг Л. Идентификация систем: Теория для пользователя. М.: Наука, 1991.
4. Петрович М. Л. Регрессионный анализ и его математическое обеспечение на ЕС ЭВМ. М.: Финансы и статистика, 1982.
5. Грановский В. А., Синая Т. Н. Методы обработки экспериментальных данных при измерениях. Л.: Энергоатомиздат, 1990.
6. Федоров В. В. Теория оптимального эксперимента. М.: Наука, 1971.
7. Садыхов Р. А., Гаджиев Ч. М. Идентификация параметров математической модели нефтяного пласта // Метрология. 1996. № 4. С. 3—14.
8. Гаджиев Ч. М. Рекуррентный алгоритм оценивания линейной регрессионной модели с учетом погрешностей в регрессионном векторе. // Заводская лаборатория. 2000. Т. 66. № 7. С. 66—71.
9. Абдуллаев Ф. М., Гейдаров Э. Х. // Автоматика и телемеханика. 1985. № 1. С. 41—47.
10. Чебраков Ю. В. Теория оценивания параметров в измерительных экспериментах. СПб.: Изд-во Санкт-Петербургского гос. ун-та, 1997.
11. Самарский А. А., Вабишевич П. Н. Численные методы решения обратных задач математической физики. М.: Едиториал УРСС, 2004.
12. Садыхов Р. А., Дадашев А. М. Применение нечеткой логики и Soft Computing технологии. Баку: Nafta-Press, 2000.
13. Mohinder S. G., Angus P. A. Kalman Filtering. Theory and Practice. New Jersey: Prentice Hall, Upper Saddle River, 1993.
14. Дмитриев В. И. и др. Вычислительные математика и техника в разведочной геофизике. М.: Недра, 1990.
15. Буряковский Л. А., Джафаров И. С., Джаваншир Р. Д. Моделирование систем нефтегазовой геологии. М.: Недра, 1990.
16. Садыхов Р. А. Математические модели и эффективные алгоритмы решения оптимальных задач АСУ ТП добычи нефти: Автореф. дис. на соиск. учен. степ. докт. техн. наук. М., 1995.
17. Yager R. R., Zadeh L. A. (Eds). Fuzzy sets, neural networks and soft computing. N. Y.: VAN Nostrand Reinhold, 1994.
18. Алиев Р. А., Алиев Р. Р. Теория интеллектуальных систем и ее применение. Баку: Casioglu, 2001.
19. Sadigov R. A. e. a. Identification of fuzzy models of geofiled // Eighth International conf. appl. of fuzzy systems and soft computing. Helsinki, Finland, 2008. P. 187—192.
20. Bohme G. Fuzzy-logic. N.Y. Springer — Verlag, 1994.
21. Хайкин С. Нейронные сети. М.: Изд-во ООО «И. Д. Вильямс», 2006.
22. Круглов В. В., Борисов В. В. Искусственные нейронные сети. Теория и практика. М. Горячая линия — Телеком, 2002.
23. Рутковская Т., Пилиньский М., Рутковский Л. Нейронные сети и нечеткие системы. М.: Горячая линия — Телеком, 2006.
24. Pashayev A. M. e. a. Identification System for the Technical Condition of Gas Turbine Engines of Aircraft // Aviation. 2008. N 12(4). P. 101—112.

Дата принятия 13.05.2010 г.

621.398

Регуляризация задачи контроля технического состояния сложных технических объектов с использованием программных комплексов для систем нечеткого условного вывода

С. А. ГАЛАЕВ, А. Ю. ПОТЮПКИН

Военная академия ВВСН им. Петра Великого, Москва, Россия
e-mail: galaevsergey1@rambler.ru

На конкретном примере показано применение методики регуляризации задачи контроля технического состояния сложных технических объектов в нештатных ситуациях с привлечением дополнительной семантической информации.

Ключевые слова: сложные технические объекты, бортовая информационно-телеметрическая система, контроль технического состояния, нештатная ситуация, регуляризация, программное обеспечение, семантическая информация.

The application of technical condition control problem regularization method for complex technical objects in emergency situations by more semantic information attraction is shown concrete example.

Key words: complex technical objects, onboard information-telemetric system, technical condition control problem, the monitoring system, emergency situation, regularization software, semantic information.

В настоящее время актуальной является задача определения технического состояния (ТС) сложных технических объектов (СТО) при возникновении нештатных ситуаций, обус-

ловленных недостатком измерительной информации, сбоями и недоработками программного обеспечения и технической документации. В [1] приведен алгоритм определе-