УДК 62-50

В.А. КОЛБАСИН, аспирант каф. КММM НТУ «ХПИ»

РЕКУРРЕНТНАЯ ИДЕНТИФИКАЦИЯ НЕЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ НА ОСНОВЕ МЕТОДА ОПОРНЫХ ВЕКТОРОВ

В статті пропонується ітераційний метод ідентифікації нелінійної динамічної системи. Для побудови моделі системи використовується метод опорних векторів. Для ітеративного обчислення матриці Грамма використовуються прямі та зворотні формули Гревіля.

В статье предлагается итерационный метод идентификации нелинейно динамической системы. Для построения модели системы используется метод опорных векторов. Для итеративного вычисления матрицы Грамма используются прямые и обратные формулы Гревиля.

In the article there have been proposed iterative method of nonlinear dynamic system identification. Support vector machines are used for designing dynamic system model. And Greville formulas are used for iterative evaluation Gramm matrix.

Введение. Для управления динамической системой в условиях неопределенности в первую очередь необходимо провести идентификацию этой системы с целью построения модели системы и оценка её параметров.

В последнее время было разработано и предложено множество методов идентификации сложных динамических систем. Многие из них основаны на использовании универсальных аппроксиматоров - искусственных нейронных сетей. Однако их практическое применение связано с рядом трудностей, из которых наиболее существенными являются:

- трудность выбора необходимого количества элементов обучающей выборки (при малом количестве наблюдается слишком грубая оценка модели - underfitting, при слишком большом количестве образцов получаем существенное увеличение времени обучения, а кроме того, наблюдается эффект переобучения - overfitting);
- медленная сходимость алгоритмов обучения нейронных сетей, что создает проблемы при идентификации в реальном масштабе времени (при сложной нелинейном характере идентифицируемой модели возможно наличие множества локальных минимумов, которые существенно затрудняют процесс обучения).

В то же время хорошо себя зарекомендовал метод опорных векторов (support vector machine - SVM), который эффективен при идентификации сложных нелинейных зависимостей в условиях малых выборок, а также обеспечивает возможность получения аналитического выражения для оценки неизвестных параметров модели [1].

Обычно метод опорных векторов применяется к стационарным системам, параметры которых существенно не меняются со временем. Это означает, что

достаточно один раз по первым N наблюдениям построить модель системы, и далее оценки модели будут применимы для всех последующих входных векторов. Однако в случае, когда система существенно нестационарна, необходимо строить модель по некоторому скользящему окну наблюдений. При большой размерности вектора состояния и вектора управления такой подход нереализуем в масштабе реального времени из-за значительного количества вычислений.

В данной работе предложен итерационный алгоритм для идентификации нестационарной нелинейной динамической системы по скользящему окну наблюдений. Для ускорения идентификации используется рекуррентное вычисление обратной матрицы Грамма.

Постановка задачи. Рассмотрим задачу идентификации дискретной нелинейной нестационарной динамической системы:

$$\begin{cases} x(k+1) = f(\{x\}_{k-n+1}^{k}, \{u\}_{k-m+1}^{k}) \\ y(k+1) = x(k+1) + e(k+1) \end{cases}, \tag{1}$$

где $\{x\}_{k-n+1}^{k}$ - предыдущие значения вектора состояния $\{x(k-n+1),...,x(k)\}$; $\{u\}_{k-n+1}^{k}$ - последовательность управляющих сигналов $\{u(k-n+1),...,u(k)\}$; y - выход системы; $f(x)$ - гладкая нелинейная функция; e - аддитивный нормально распределенный шум с нулевым средним и заданной дисперсией σ_e^2.

Для идентификации рассматриваемой динамической системы будем использовать следующую модель системы:

$$y = w^T \varphi(z) + e \tag{2}$$

где $\varphi(z)$ - вектор координатных функций, вектор $w \in \mathbf{R}^m$ состоит из подлежащих определению параметров модели, а вектор z включает в себя как предыдущие значения вектора состояния, так и множество управляющих сигналов в прошедшие моменты времени: $z = [\{x\}_{k-n+1}^{k}, \{u\}_{k-n+1}^{k}]$.

Структура модели определяется априори выбранным вектором координатных функций $\varphi(z) \in \mathbf{R}^m$. Для обеспечения эффективной аппроксимации достаточно сложных моделей вектор-функция $\varphi(z)$ является нелинейным преобразованием, которое переводит входной вектор z в пространство признаков большей размерности, чем размерность входного вектора z.

Поскольку рассматриваемая система нестационарна, параметр w будем оценивать для каждой совокупности наблюдений $I_k = \{y_j, x_j, u_j\}_{j=k-n}^{k-1}$, соответствующих текущему скользящему окну.

Оценка параметра $\hat{w}(k)$ зависит от выбора вектора координатных функций $\varphi(z)$ и в случаях высокой размерности требует весьма большого количества вычислений. Для сокращения объема вычислений в соответствии с известной теоремой Мерцера координатные функции выбирают так, чтобы скалярное произведение векторов $\varphi^T(z_i)\varphi(z_j)$ можно было представить в виде положительно определенной функции $K(z_i, z_j)$ (обычно называемую ядерной функцией). В прикладных задачах распознавания образов и прогнозирования обычно используются радиально-базисные функции (RBF - radial basis functions) вида $K(z_i, z) = \exp(-\mu\|z - z_i\|_2^2)$, где $\mu > 0$ - некоторый настроечный параметр.

Таким образом, для выбранной структуры модели оценка выходной переменной y будет иметь вид:

$$\hat{y}(k) = (\hat{w}(k-1))^T \varphi(z_{k-1}). \tag{3}$$

Решение задачи идентификации. Для решения задачи идентификации модели (2) воспользуемся байесовской схемой оценивания параметров динамической системы [1,2]. При применении этой схемы оценки неизвестных параметров модели (2) определяются путем решения задачи условной минимизации функционала:

$$J(w,\varepsilon) = \frac{\sigma_w^2}{2} w^T w + \frac{\sigma_\varepsilon^2}{2} \sum_{i=1}^{n} \varepsilon_i^2, \quad \varepsilon_i^2 = (y_i - w^T \varphi(z_i))^2. \tag{4}$$

Для решения этой задачи перейдем к эквивалентной сопряженной оптимизационной задаче, которая формулируется с использованием функции Лагранжа для задачи (4):

$$L(w,\varepsilon,\lambda) = J(w,\varepsilon) + \sum_{i=1}^{n} \lambda_i [y_i - w^T \varphi(z_i) - \varepsilon_i], \tag{5}$$

где $\lambda \in \mathbf{R}^n$ - вектор сопряженных (опорных) переменных.

Решая оптимизационную задачу (5), получим систему уравнений:

$$(\gamma^{-1} I_n + \Phi_n)\lambda = Y_n, \tag{6}$$

где $Y_n = (y_1, y_2, ..., y_n)^T$ – вектор наблюдений прогнозируемого параметра, $\Phi_n = \|\Phi_{ij}\|$ – матрица Грамма размером $n \times n$ с элементами вида $\Phi_{ij} = \varphi^T(z_i)\varphi(z_j) = K(z_i, z_j)$ $i, j = \overline{1,n}$, зависящая от предшествующих наблюдений, I_n – единичная матрица соответствующей размерности,

$\gamma = \sigma_\varepsilon^2 / \sigma_w^2$ – отношение "шум - сигнал", величина которого определяется доступной априорной информацией.

Решением задачи (5) являются выражения:

$$\lambda = G_n^{-1}(\gamma) \cdot Y_n, \quad \hat{w}(i) = \sum_{j=1}^{n} \lambda_j(i) \varphi(x_i), \tag{7}$$

где $G_n(\gamma) = (\gamma^{-1} I_n + \Phi_n)$.

С использованием формул (7) окончательно получим выражение для оценки выходной переменной y на k-м шаге:

$$\hat{y}(k+1) = \sum_{j=1}^{n} \lambda_j(k) K(z_j, z_k). \tag{8}$$

Рекуррентный алгоритм идентификации по скользящему окну. Поскольку процедура вычисления обратной матрицы вычислительно сложная емкая, то применение полученного алгоритма идентификации в режиме реального времени достаточно сложно. Для получения итерационных формул будем использовать аппарат псевдообращения матриц, а именно, прямые и обратные формулы Гревиля [3].

Прямые формулы – это формулы определяющие вид псевдообращения матрицы при её дополнении строкой или столбцом. Они определяются соотношениями, в которых используется блочное представление псевдообращения расширенной матрицы:

1) Расширение матрицы A строкой a^T:

$$\begin{pmatrix} A \\ a^T \end{pmatrix}^+ = (P \vdots q) \tag{9}$$

где P - $n \times m$ - матрица и q - $n \times 1$ - вектор;

2) Дополнение матрицы A столбцом a:

$$(A \vdots a)^+ = \begin{pmatrix} Q \\ r^T \end{pmatrix} \tag{10}$$

где Q - $n \times m$ - матрица и r - $m \times 1$ - вектор.

Тогда матрицы P, Q и вектора q, r определяются по формулам:

$$\begin{cases} P = (E - qa^T)A^+ \\ q = \begin{cases} \dfrac{Z(A)a}{a^{\mathrm{T}}Z(A)a}, & a^{\mathrm{T}}Z(A)a > 0 \\ \dfrac{R(A)a}{1 + a^T R(A)a}, & a^{\mathrm{T}}Z(A)a = 0 \end{cases} \end{cases}, \tag{11}$$

$$\begin{cases} Q = A^+(E - ar^T) \\ r = \begin{cases} \dfrac{Z(A^T)a}{a^{\mathrm{T}}Z(A)a}, & a^{\mathrm{T}}Z(A^T)a > 0 \\ \dfrac{R(A^T)a}{1 + a^T R(A^T)a}, & a^{\mathrm{T}}Z(A^T)a = 0 \end{cases} \end{cases}, \tag{12}$$

где $R(A)$ и $Z(A)$ - операторы, определенные формулами:

$$R(A) = A^+ A^{T^+}, \quad R(A^T) = A^{+^T} A^+, \tag{13}$$

$$Z(A) = E_n - A^+ A, \quad Z(A^T) = E_m - AA^+. \tag{14}$$

Обратные формулы Гревиля определяют вид псевдообращения матрицы при вычеркивании в матрице строки или столбца. Как и в прямых формулах Гревиля, вид выражений, связывающих псевдообращения исходной и преобразованной матрицы, выписывается в рамках блочных представлений, аналогичных выражениям (9) и (10), а именно:

$$\begin{pmatrix} a^T \\ A \end{pmatrix}^+ = \left(q \vdots P\right), \tag{15}$$

$$\left(a \vdots A\right)^+ = \begin{pmatrix} r^T \\ Q \end{pmatrix}. \tag{16}$$

Тогда при вычеркивании строки a^T с учетом (15) матрица A^+ примет следующий вид:

$$A^+ = \begin{cases} \left(I_n - qq^T / \|q\|^2\right)P, & a^T q = 1 \\ \left(I_n - qa^T \,/\, (1 - a^T q)\right)P, & a^T q < 1 \end{cases}. \tag{17}$$

При вычеркивании столбца a с учетом (16) A^+ примет следующий вид:

$$A^+ = \begin{cases} Q\left(I_m - \dfrac{rr^T}{\|r\|^2}\right), & a^T r = 1 \\ Q\left(I_m - \dfrac{ra^T}{1 - a^T r}\right), & a^T r < 1 \end{cases}. \tag{18}$$

Для рекуррентного вычисления параметра $\lambda(k+1)$ будем использовать следующий алгоритм:

1. Сформируем выборку для оценки параметров на следующей итерации: $I_{k+1} = \{y_j, x_j, u_j\}_{j=k-n+1}^{k}$

2. Используя вычисленную на предыдущей итерации матрицу $G_k^{-1}(\gamma)$, вычислим матрицу $G_{k+1}^{-1}(\gamma)$: последовательно выполним преобразования (17), (18), (11) и (12) – т.е. получим обратную матрицу от матрицы, полученной следующим образом: вычеркнем в матрице $G_k(\gamma)$ первый столбец и первую строку, затем добавим строку c^T и столбец d, которые равны:

$$c^T = (K(z_{k-n+1}, z_k), K(z_{k-n+2}, z_k), ..., K(z_{k-1}, z_k)), \tag{19}$$

$$d^T = (K(z_k, z_{k-n+1}), K(z_k, z_{k-n+2}), ..., K(z_k, z_k)). \tag{20}$$

3. Для получения оценки параметра $\lambda(k+1)$ выполним преобразование (7) с новой матрицей $G_{k+1}^{-1}(\gamma)$.

Заключение. Таким образом, в работе предложена общая схема рекуррентной идентификации нелинейной динамической системы на основе метода опорных векторов. Практическая реализация метода требует формирования методики оценки настроечных параметров μ, γ, а также решения дополнительных вопросов, связанных с выбором оптимальной сложности модели, определяемой ее размерностью n.

Список литературы: ***1.*** Cristianini N., Shawe-Tailor J. An Introduction to Support Vector Machines. Cambridge, Univ. Press, 2000. - 320 p. ***2.*** *Гринберг Г. Л., Любчик Л. М.* Идентификация и прогнозирование финансовых временных рядов на основе метода опорных векторов.– Вестник НТУ «ХПИ». Сб. научных трудов. Тем. вып. «Системный анализ, управление и информационные технологии».– Харьков: НТУ «ХПИ».– 2003.– № 18.– С. 29-32. ***3.*** *Кириченко Н., Донченко В.* Гиперплоскости в «множествах и расстояниях соответствия»: кластеризация. - International Book Series “Information Science and Computing”.– 2008.– № 2.– С. 25–36.

Поступила в редколлегию 08.01.08