

УДК 681-51

Л.М. Любчик, В.А. Колбасин

**РЕКУРРЕНТНЫЕ АЛГОРИТМЫ ИДЕНТИФИКАЦИИ
НЕЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ НА ОСНОВЕ
ЯДРОВЫХ МОДЕЛЕЙ**

Рассматривается задача идентификации нелинейных систем в реальном масштабе времени. На основе ядрового подхода получены рекуррентные алгоритмы идентификации с растущим и скользящим окном, предложены соответствующие вычислительные процедуры.

Введение

Проблема идентификации динамических систем является одной из центральных в современной теории управления. Несмотря на обилие публикаций [1] она сохраняет свою актуальность в первую очередь за счет расширения классов задач и привлечения новых математических методов. Если проблемы идентификации линейных систем в значительной мере получили исчерпывающее решение, то поиск эффективных методов построения моделей нелинейных систем по экспериментальным данным является предметом интенсивных исследований. Здесь на первый план выходит проблема структурной идентификации, связанная с поиском структуры модели, оптимальным образом согласованной с имеющейся априорной и текущей информацией [2,3]. Для нелинейных систем, характеризующихся сложным характером зависимостей между переменными, использование традиционного параметрического подхода приводит к существенным трудностям, связанным с вычислительным трудностям оценивания большого числа неизвестных параметров переусложненных моделей (так называемое «проклятие размерности»). Это в свою очередь стимулирует развитие непараметрических методов идентификации и методов «мягкой» идентификации на основе подходов вычислительного интеллекта. Эффективные подходы созданы на основе метода опорных векторов [3] и так называемых ядровых методов [4], обеспечивающих получение нелинейных версий алгоритмов идентификации, пригодных для использования в условиях малых выборок.

© Любчик Л.М., Колбасин В.А., 2010

Ядровые методы, основанные на идее нелинейного преобразования исходных данных в новое пространство высокой размерности (пространство признаков), обеспечивают возможность идентификации нелинейных моделей высокой сложности. В соответствии с теоремой Мерсера, указанное преобразование выбирается таким образом, чтобы скалярные произведения в пространстве признаков имели бы вид положительно определенной функции-ядра. При этом идентифицируемая модель может быть представлена в непараметрической форме в виде взвешенной линейной комбинации функций-ядер, а соответствующие весовые коэффициенты могут быть вычислены без непосредственного использования векторов признаков, что приводит к возможности получения весьма экономных и эффективных вычислительных процедур идентификации нелинейных систем [4, 5].

Классические алгоритмы ядровой идентификации использовали для получения оценок полную выборку наблюдений и характеризовались ростом размерности модели с увеличением ее длины. В настоящее время ведутся интенсивные поиски рекуррентных модификаций ядровых алгоритмов, пригодных для реализации в реальном масштабе времени и обеспечивающих ограничение сложности идентифицируемой модели [6]. В частности, разработаны ядровые аналоги рекуррентного метода наименьших квадратов [7] и метода скользящего окна [8].

В настоящей статье предлагается единый подход к построению рекуррентных алгоритмов ядровой идентификации с использованием как растущего, так и скользящего окна наблюдений.

Постановка задачи

Рассматривается скалярная дискретная нелинейная динамическая система, описываемая уравнением вида

$$x_{k+1} = f(x_k) + \varepsilon_k, \quad k = 0, \dots \quad (1)$$

где $f(\cdot)$ неизвестная нелинейная функция, ε_k - внешнее воздействие, описываемое случайным процессом типа дискретного белого шума $\mathbf{E}\{\varepsilon_k\} = 0$, $\mathbf{E}\{\varepsilon_k^2\} = \sigma^2$.

Задача идентификации состоит в оценивании параметров идентифицируемой модели:

$$y_k = \hat{f}(x_k) = \varphi^T(x_k)w + \varepsilon_k, \quad k = \overline{0, n}, \quad (2)$$

где $\hat{f}(x)$ - неизвестная функциональная зависимость, описывающая модель системы (1), $\varphi: \mathbf{R}^1 \rightarrow \mathbf{R}^M$ нелинейное отображение, трансформирующее состояния системы в многомерный вектор признаков $\varphi(x) \in \mathbf{R}^M$, $\mathbf{w} \in \mathbf{R}^M$ - вектор неизвестных параметров идентифицируемой модели.

Предполагается заданной выборка наблюдений $\{y_k, x_k\}_{k=0}^n$, $y_k = x_{k+1}$.

В матричной форме уравнение (2) имеет вид $y_n = \Phi_{n-1}^T \mathbf{w} + \varepsilon_n$, где $\Phi_{n-1} = (\varphi(x_0) \ \varphi(x_1) \ \dots \ \varphi(x_{n-1}))$ - матрица признаков, $y_n = (y_0 \ y_1 \ \dots \ y_n)^T$ - вектор наблюдений, $\varepsilon_n = (\varepsilon_0 \ \varepsilon_1 \ \dots \ \varepsilon_n)^T$ - вектор возмущений.

В соответствии с общей методологией ядрового подхода [4] выберем вектора признаков так, чтобы их скалярные произведения имели вид положительно определенной функции $\varphi^T(x_i)\varphi(x_j) = \kappa(x_i, x_j)$, $i, j = \overline{1, n}$. Обычно используются полиномиальные ядра вида $\kappa(x, x') = (\mu + x \cdot x')^p$ заданной степени p , либо гауссовские ядра $\kappa(x, x') = \exp\{-\mu(x - x')^2\}$ с параметрами p, μ .

Тогда идентифицируемая модель системы (1) может быть представлена в виде

$$\hat{x}_{k+1} = \hat{f}(x_n) = \varphi^T(x_n)\Phi_{n-1}\mathbf{w}_n = \mathbf{k}_{n-1}^T(x_n)\boldsymbol{\lambda}_n, \quad (3)$$

где $\mathbf{k}_{n-1}^T(x_n) = (\kappa(x_n, x_0) \ \kappa(x_n, x_1) \ \dots \ \kappa(x_n, x_{n-1}))$ - вектор функций-ядер, а $\boldsymbol{\lambda}_n \in \mathbf{R}^n$ - вектор вспомогательных переменных, таких, что $\mathbf{w}_n = \Phi_{n-1}\boldsymbol{\lambda}_n$.

Задача состоит в построении рекуррентных оценок вектора $\boldsymbol{\lambda}_n$ на основе обучающей выборки $\{y_k, x_k\}_{k=0}^{n-1}$, при этом в соответствии с общей идеологией ядровых методов оценки должны строиться на основе матрицы $\mathbf{K}_n = \Phi_n^T \Phi_n$, $\mathbf{K}_n = \|k_{i,j}\|$, $k_{i,j} = \kappa(x_i, x_j)$, $i, j = \overline{1, n}$, (так называемой ядровой матрицы), вычисляемой непосредственно без использования векторов признаков.

Целесообразно рассмотреть два варианта построения рекуррентных алгоритмов идентификации.

а. *Рекуррентная ядровая идентификация с «растущим окном».* При этом оценки $\lambda_{n+1} = F(\lambda_{n+1}, y_{n+1}, \mathbf{K}_n)$ строятся на основе использования полной выборки наблюдений. Этот подход можно рассматривать как рекуррентное обобщение традиционного ядрового подхода. Его очевидным недостатком является рост сложности модели с увеличением длины выборки.

б. *Рекуррентная ядровая идентификация со «скользящим окном».* При этом оценки $\bar{\lambda}_{n+1} = F(\bar{\lambda}_{n+1}, y_{n+1}, \bar{\mathbf{K}}_{n,s}), \bar{\lambda}_n \in \mathbf{R}^s$, используют только ограниченную подвыборку наблюдений, формируемую на основе принципа «скользящего окна» $\{y_k, x_k\}_{k=n-s}^n$ и ядровая матрица $\bar{\mathbf{K}}_{n,s}$ строится на основе соответствующих наблюдений.

Рекуррентный алгоритм идентификации с «растущим окном»

На основе известного подхода метода опорных векторов [3] задача построения оценок параметров модели (2) на основе использования полной обучающей выборки может быть сведена к следующей регуляризованной оптимизационной задаче:

$$J_n(\mathbf{w}) = \mathbf{w}^T \mathbf{w} + \gamma \cdot \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \rightarrow \min_{\mathbf{w}, \boldsymbol{\varepsilon}}, \quad \mathbf{y}_n = \Phi_{n-1}^T \mathbf{w} + \boldsymbol{\varepsilon}_n. \quad (4)$$

Решение этой задачи может быть получено на основе функции Лагранжа $L_n(\mathbf{w}, \boldsymbol{\lambda}) = \mathbf{w}^T \mathbf{w} + \gamma \cdot \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} + \boldsymbol{\lambda}^T (\mathbf{y}_n - \Phi_{n-1}^T \mathbf{w} - \boldsymbol{\varepsilon}_n)$ с двойственными переменными $\boldsymbol{\lambda} \in \mathbf{R}^n$. Используя известные условия оптимальности, получим решение в явном виде $\mathbf{w}_n = \Phi_{n-1} \boldsymbol{\lambda}_n, \boldsymbol{\varepsilon}_n = \gamma^{-1} \boldsymbol{\lambda}_n$, где оценки двойственных переменных $\boldsymbol{\lambda}_n$ в свою очередь выражаются через регуляризованную ядровую матрицу $\mathbf{K}_{n-1}(\gamma)$:

$$\boldsymbol{\lambda}_n = (\gamma^{-1} \mathbf{I}_n + \mathbf{K}_{n-1})^{-1} \mathbf{y}_n = \mathbf{K}_{n-1}^{-1}(\gamma) \mathbf{y}_n, \quad (5)$$

С использованием очевидного представления ядровой матрицы

$$\mathbf{K}_n(\gamma) = \left(\begin{array}{c|c} \mathbf{K}_{n-1}(\gamma) & \mathbf{k}_{n-1}(x_n) \\ \hline \mathbf{k}_{n-1}^T(x_n) & \gamma^{-1} + k_{n,n} \end{array} \right), \quad (6)$$

на основе формулы Шермана-Моррисона-Вудбери [9] нетрудно получить рекуррентное представление для обратной ядровой матрицы

$$\mathbf{K}_n^{-1}(\gamma) = \left(\begin{array}{c|c} \mathbf{K}_{n-1}^{-1}(\gamma) + \delta_n^{-1} \mathbf{K}_{n-1}^{-1}(\gamma) \mathbf{k}_{n-1}(x_n) \mathbf{k}_{n-1}^T(x_n) \mathbf{K}_{n-1}^{-1}(\gamma) & -\delta_n^{-1} \mathbf{K}_{n-1}^{-1}(\gamma) \mathbf{k}_{n-1}(x_n) \\ \hline -\delta_n^{-1} \mathbf{k}_{n-1}^T(x_n) \mathbf{K}_{n-1}^{-1}(\gamma) & \delta_n^{-1} \end{array} \right), \quad (7)$$

где $\delta_n = \gamma^{-1} + k_{n,n} - \mathbf{k}_{n-1}^T(x_n) \mathbf{K}_{n-1}^{-1}(\gamma) \mathbf{k}_{n-1}(x_n)$.

Тогда рекуррентная оценка для для вектора параметров $\lambda_{n+1} = \mathbf{K}_n^{-1}(\gamma) \mathbf{y}_n$ может быть представлена в виде:

$$\lambda_{n+1} = \left(\begin{array}{c} \lambda_n - \delta_n^{-1} [y_{n+1} - \omega_n(\lambda_n)] \mathbf{K}_{n-1}^{-1}(\gamma) \mathbf{k}_{n-1}(x_n) \\ \delta_n^{-1} [y_{n+1} - \omega_n(\lambda_n)] \end{array} \right), \quad (8)$$

где $\omega_n(\lambda_n) = \mathbf{k}_{n-1}^T(x_n) \lambda_n$.

Соотношения (5), (8) с учетом (3) образуют рекуррентный алгоритм идентификации нелинейной системы (1) с «растущим окном». Очевидно, что сложность модели, определяемая размерностью вектора двойственных переменных λ , растет с увеличением длины выборки.

Рекуррентный алгоритм идентификации со «скользящим окном»

Метод «скользящего окна» использует для построения оценки в момент n только $n-s$ последних наблюдений, а именно, $\mathbf{y}_{n,s} = (y_{n-s+1} \dots y_n)^T$. При этом возникает возможность идентификации систем с изменяющимися параметрами.

Уравнение наблюдений в этом случае имеет вид $\mathbf{y}_{n,s} = \Phi_{n-1,s}^T \mathbf{w} + \varepsilon_{n,s} = \mathbf{K}_{n-1,s} \bar{\lambda}_n + \varepsilon_{n,s}$, где $\bar{\lambda}_n \in \mathbf{R}^s$, а ядровая матрица $\bar{\mathbf{K}}_{n,s} = \Phi_{n,s}^T \Phi_{n,s}$ сохраняет фиксированную размерность $(s \times s)$.

Сформируем «скользящий» показатель качества идентификации в момент $n+1$ и априорную информацию в виде оценок в предыдущий момент времени n , а также представление $\mathbf{w} = \Phi_{n,s} \bar{\lambda}_n$:

$$J_{n,s} = \left\| \mathbf{y}_{n+1,s} - \mathbf{K}_{n,s} \bar{\lambda}_n \right\|^2 + \gamma^{-1} (\bar{\lambda}_n - \bar{\lambda}_n)^T \mathbf{K}_{n,s} (\bar{\lambda}_n - \bar{\lambda}_n) \rightarrow \min_{\bar{\lambda}}. \quad (9)$$

где второе слагаемое, имеющее смысл регуляризирующей составляющей, учитывает априорное предположение о медленном характере изменения параметров модели.

Условия оптимальности приводят к системе нормальных уравнений:

$$\left(\bar{\mathbf{K}}_{n,s}^T \bar{\mathbf{K}}_{n,s} + \gamma^{-1} \bar{\mathbf{K}}_{n,s} \right) \bar{\lambda}_n = \bar{\mathbf{K}}_{n,s}^T \mathbf{y}_{n+1,s} + \gamma^{-1} \bar{\mathbf{K}}_{n,s} \bar{\lambda}_n. \quad (10)$$

из которых после ряда преобразований можно получить рекуррентную оценку двойственного вектора

$$\bar{\lambda}_{n+1} = \left(\gamma^{-1} \mathbf{I}_s + \bar{\mathbf{K}}_{n,s} \right)^{-1} \left(\gamma^{-1} \bar{\lambda}_n + \mathbf{y}_{n+1} \right) \quad (11)$$

Для получения рекуррентного алгоритма пересчета обратной регуляризованной ядровой матрицы $\bar{\mathbf{K}}_{n,s}^{-1}(\gamma)$ используется двухшаговый подход, предложенный в [8].

Схема пересчета имеет вид $\bar{\mathbf{K}}_{n-1,s}^{-1}(\gamma) \rightarrow \bar{\mathbf{K}}_{n-1,s-1}^{-1}(\gamma) \rightarrow \bar{\mathbf{K}}_{n,s}^{-1}(\gamma)$, где используется вспомогательная квадратная матрица $\bar{\mathbf{K}}_{n-1,s}(\gamma)$, которую можно представить в виде:

$$\bar{\mathbf{K}}_{n-1,s}(\gamma) = \left(\begin{array}{c|c} \gamma^{-1} + k_{n-s,n-s} & \mathbf{k}_{n-1,s-1}^T(x_{n-s}) \\ \hline \mathbf{k}_{n-1,s-1}(x_{n-s}) & \bar{\mathbf{K}}_{n-1,s-1}(\gamma) \end{array} \right), \quad (12)$$

$$\mathbf{k}_{n-1,s-1}(x_{n-s}) = \left(\kappa_{n-1}(x_{n-s}) \dots \kappa_{n-s+1}(x_{n-s}) \right)^T$$

На первом шаге процедуры рекуррентного обращения вычисляется

$$\bar{\mathbf{K}}_{n-1,s-1}^{-1} = \mathbf{R}_s \bar{\mathbf{K}}_{n-1,s}^{-1} \mathbf{R}_s^T - \left(\mathbf{e}_1^T \bar{\mathbf{K}}_{n-1,s}^{-1} \mathbf{e}_1 \right)^{-1} \mathbf{R}_s \bar{\mathbf{K}}_{n-1,s}^{-1} \mathbf{e}_1 \mathbf{e}_1^T \bar{\mathbf{K}}_{n-1,s}^{-1} \mathbf{R}_s^T, \quad (13)$$

где $\mathbf{R}_s = (0_s : \mathbf{I}_{s-1})$, $\mathbf{e}_1 = (1 \dots 0)^T$.

На втором шаге процедуры рекуррентного обращения с использованием следующего представления регуляризованной ядровой матрицы $\mathbf{K}_{n,s}^{-1}(\gamma)$

$$\bar{\mathbf{K}}_{n,s}(\gamma) = \left(\begin{array}{c|c} \bar{\mathbf{K}}_{n-1,s-1}(\gamma) & \mathbf{k}_{n-1,s-1}(x_n) \\ \hline \mathbf{k}_{n-1,s-1}^T(x_n) & \gamma^{-1} + k_{n,n} \end{array} \right), \quad (14)$$

где $\mathbf{k}_{n-1,s-1}(x_n) = \left(\kappa_{n-1}(x_n) \dots \kappa_{n-s+1}(x_n) \right)^T$, вычисляется обратная ядровая матрица на шаге n :

$$\mathbf{K}_{n,s}^{-1} = \left(\begin{array}{c|c} \mathbf{K}_{n-1,s-1}^{-1}(\gamma) + \delta_n^{-1} \mathbf{K}_{n-1,s-1}^{-1}(\gamma) \mathbf{k}_{n-1}(x_n) \mathbf{k}_{n-1}^T(x_n) \mathbf{K}_{n-1,s-1}^{-1}(\gamma) & -\delta_n^{-1} \mathbf{K}_{n-1,s-1}^{-1}(\gamma) \mathbf{k}_{n-1,s-1}^T(x_n) \\ \hline -\delta_n^{-1} \mathbf{K}_{n-1,s-1}^{-1}(\gamma) \mathbf{k}_{n-1,s-1}^T(x_n) & \delta_n^{-1} \end{array} \right), \quad (15)$$

где $\delta_n = \gamma^{-1} + k_{n,n} - \mathbf{k}_{n-1,s-1}^T(x_n) \bar{\mathbf{K}}_{n-1,s-1}^{-1}(\gamma) \mathbf{k}_{n-1,s-1}(x_n)$.

Окончательно выражения (11), (13), (15), образуют ядровый рекуррентный алгоритм идентификации системы (1) со «скользящим

окном», использующий ограниченный объем текущей информации и модель ограниченной сложности.

Выводы

Предложенные рекуррентные алгоритмы идентификации нелинейных систем обеспечивают возможность оценивания параметров модели в реальном масштабе времени и обладают сравнительной вычислительной простотой и эффективностью. Существенным является то, что сложность модели не увеличивается в ростом длины выборки, а преимущества ядрового подхода позволяют строить модели весьма сложных нелинейных динамических систем.

Дальнейшие направления исследований могут быть связаны с оптимизацией параметров алгоритма, в частности, параметров ядровых функций, параметров регуляризации, длины «скользящего окна». Для этого необходимо использовать дополнительные, так называемые внешние критерии идентификации, в качестве которых в данном случае целесообразно использование оценок типа скользящего контроля. Также важным направлением является разработка робастных модификаций предложенных рекуррентных алгоритмов.

ЛИТЕРАТУРА

1. L. Ljung. System Identification: Theory for the User. Prentice Hall, New Jersey, Second edition, 1999.
2. G. Dorffner. Neural networks for time series processing // Neural Network World, 1996, 6(4), Pp. 447–468.
3. V. Vapnik. Statistical Learning Theory. Wiley, New-York, 1998.
4. B. Scholkopf, A. Smola. Learning with kernels. Cambridge, MA: MIT Press, 2002.
5. M. Espinoza, J.A.K. Suykens, and B. De Moor. Kernel based partially linear models and nonlinear identification // IEEE Transactions on Automatic Control, 2005, 50 (10), Pp. 1602–1606.
6. J. Kivinen, A.J. Smola, R.C. Williamson. Online learning with kernels // IEEE Transactions on Signal Processing, 2004, Vol. 52, Pp. 2165 – 2176.
7. J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle. Least Squares Support Vector Machines. World Scientific, Singapore, 2002.
8. S. Van Vaerenbergh, V. Javier, I. Santamar. Nonlinear System Identification using a New Sliding Window Kernel RLS Algorithm // Journal of Communications, 2007, Vol. 2, No. 3, Pp. 156 -162.
9. G.H. Golub, C.F. Van Loan. Matrix Computations. Johns Hopkins Univ. Press. Baltimore MD, 1989.