

УДК 517.977.12

СРАВНЕНИЕ МЕТОДА СКОРОСТНОГО ГРАДИЕНТА И ПРОЦЕДУРЫ DREM В ЗАДАЧЕ ИДЕНТИФИКАЦИИ СЕТИ ИЗ МОДЕЛЕЙ НЕЙРОНОВ ФИТЦХЬЮ-НАГУМО

А.В. Рыбалко

Институт проблем машиноведения РАН

Россия, 199178, Санкт-Петербург, Большой проспект В.О., 61

E-mail: alexandrorybalco21@gmail.com

Ключевые слова: идентификация, метод скоростного градиента, DREM, нейронные сети, модель ФитцХью-Нагумо

Аннотация: Работа посвящена задаче идентификации модели нейронной сети, составленной из моделей нервных клеток ФитцХью-Нагумо. Задача решается в предположениях о неизмеряемости части переменных и всех производных, а также с учетом возникающих погрешностей. Сначала исходная модель с помощью математических преобразований и фильтрации приводится к виду линейной регрессии, параметры которой требуется оценить. Затем применяются классический метод скоростного градиента и процедура динамического расширения и смешивания регрессора (DREM). Полученные решения промоделированы в Simulink, выявлены недостатки и преимущества использования каждого из методов в рассматриваемой задаче.

1. Введение

В данной работе рассматривается задача идентификации сети из N моделей нейронов ФитцХью-Нагумо (ФХН) [1, 2]. Модель ФХН – это упрощенная модель Ходжкина-Хаксли, имеющая множество приложений благодаря своей простоте и универсальности. Обычно только одна переменная ФХН доступна измерению, что усложняет не только идентификацию модели, но и ее применение в целом.

В [3, 4] теоретически и численно обосновано решение задачи на основе метода скоростного градиента [5]. В отличие от существующих работ, посвященных этой задаче, в [4] решалась задача идентификации не одной модели, а N моделей ФХН, связанных между собой. Также подход [3, 4] учитывает погрешность при измерении переменной мембранного потенциала. Целью этой работы является сравнение результата [4] с результатами, которые можно получить, применяя альтернативный подход к идентификации, процедуру DREM. Она была предложена всего несколько лет назад [6], однако, успела зарекомендовать себя в качестве альтернативы классического градиентного подхода, работающей при выполнении более слабых условий на исходную модель [7].

2. Построение алгоритма идентификации

Рассмотрим модель нейронной сети, составленную из N моделей ФитцХью-Нагумо:

$$(1) \quad \begin{cases} u'_k = u_k - \frac{u_k^3}{3} - v_k + I_{ext} + \sigma \sum_{j=1}^N A_{kj} [B_{uu}(u_j - u_k) + B_{uv}(v_j - v_k)], \\ v'_k = \varepsilon(u_k - a - bv_k) + \sigma \sum_{j=1}^N A_{kj} [B_{vu}(u_j - u_k) + B_{vv}(v_j - v_k)], \end{cases}$$

где $k \in 1 : N$, u_k – мембранные потенциалы нейронов, v_k – переменные совокупного действия всех медленных ионных токов, отвечающих за восстановление потенциала покоя. $a, b, \varepsilon > 0$, I_{ext} – параметры модели ФХН [1, 2], σ – общая сила связи, A_{kj} – коэффициенты матрицы смежности сети (предполагаем, что связи ненаправленные), $B_{uu}, B_{uv}, B_{vu}, B_{vv}$ – некоторые коэффициенты масштабирования.

Предполагается, что измеряемы только переменные u_k , что соответствует биологической интерпретации модели. Более того, при измерении u_k неизбежно возникнет погрешность. Учтем ее мультипликативно, введя замену: $y_k = cu_k$, где c – априори неизвестный коэффициент масштабирования, y_k – новые переменные. С помощью математических преобразований [4] система (1) приводится к уравнению:

$$(2) \quad \left(\sum_{k=1}^N y_k \right)'' = \theta_1^* \left(\sum_{k=1}^N y_k \right)' + \theta_2^* \left(\sum_{k=1}^N y_k^3 \right)' + \theta_3^* \left(\sum_{k=1}^N y_k \right) + \theta_4^* \left(\sum_{k=1}^N y_k^3 \right) + \theta_5^*,$$

в котором отсутствуют неизменяемые переменные v_k , $\theta^* = (1 - \varepsilon b \quad -\frac{1}{3c^2} \quad \varepsilon(b-1) \quad -\varepsilon b/3c^2 \quad Nc\varepsilon(a + bI_{ext}))^T$ – вектор истинных значений параметров.

Теперь нам осталось убрать неизменяемые производные y'_k и y''_k из уравнения (2). Для этой цели будем использовать фильтр, двойной реальный дифференциатор: $(\tau_1 p + 1)(\tau_2 p + 1)\xi = p^2 \eta$, где ξ – оценка η'' , $p = d/dt$, $\tau_i > 0$ $i = 1, 2$. Теперь применим фильтр с передаточной функцией $W(p) = ((\tau_1 p + 1)(\tau_2 p + 1))^{-1}$ к обеим частям (2) и введем новые отфильтрованные переменные:

$$(3) \quad \begin{aligned} x_1 &= p^2 ((\tau_1 p + 1)(\tau_2 p + 1))^{-1} \sum_{k=1}^N y_k, \\ x_2 &= p ((\tau_1 p + 1)(\tau_2 p + 1))^{-1} \sum_{k=1}^N y_k, \quad x_3 = p ((\tau_1 p + 1)(\tau_2 p + 1))^{-1} \sum_{k=1}^N y_k^3, \\ x_4 &= ((\tau_1 p + 1)(\tau_2 p + 1))^{-1} \sum_{k=1}^N y_k, \quad x_5 = ((\tau_1 p + 1)(\tau_2 p + 1))^{-1} \sum_{k=1}^N y_k^3. \end{aligned}$$

Тогда уравнение (2) принимает вид:

$$(4) \quad x_1 = \theta_1^* x_2 + \theta_2^* x_3 + \theta_3^* x_4 + \theta_4^* x_5 + \theta_5^*.$$

Задача состоит в том, чтобы получить алгоритм настройки оценок параметров, $\theta_i(t)$ $i \in 1 : 5$, модели (4), который бы обеспечивал достижение цели идентификации:

$$(5) \quad \begin{aligned} 1) \quad & \hat{x}_1(t) - x_1(t) \rightarrow 0 \text{ для } t \rightarrow \infty, \\ 2) \quad & \theta(t) - \theta^* \rightarrow 0 \text{ для } t \rightarrow \infty, \end{aligned}$$

где \hat{x}_1 – оценка выходного сигнала x_1 модели (4), $\theta(t) = (\theta_1(t) \quad \theta_2(t) \quad \theta_3(t) \quad \theta_4(t) \quad \theta_5(t))^T$.

2.1. Метод скоростного градиента

Построим дополнительную адаптивную систему с $\hat{x}_1(t)$ в качестве выходной переменной и $\theta(t)$ в качестве параметров:

$$(6) \quad \hat{x}_1 = \theta_1 x_2 + \theta_2 x_3 + \theta_3 x_4 + \theta_4 x_5 + \theta_5.$$

К ней можно применить метод скоростного градиента для интегрального целевого функционала Q_t , явно зависящего от настраиваемых параметров [5]:

$$(7) \quad Q_t = \int_0^t \frac{1}{2} \delta^2(x(s), \theta(s), s) ds,$$

где $\delta(x, \theta, t) = \hat{x}_1(t) - x_1(t)$. В таком случае алгоритм скоростного градиента принимает вид:

$$(8) \quad \theta'(t) = -\Gamma \delta(x, \theta, t) z(x, t),$$

где $\Gamma = \text{diag}\{\gamma_1, \dots, \gamma_5\}$ – матрица положительных коэффициентов усиления, $z(x, t) = (x_2 \ x_3 \ x_4 \ x_5 \ 1)^T$ – вектор наблюдаемых величин (регрессор из уравнения (4)).

Определение 1. Для вектор-функции $\phi(t)$ выполняется условие *постоянного возбуждения*, если функция ограничена при $t \geq 0$ и существуют положительные L, α, t_0 такие, что для любого $t > t_0$ выполняется:

$$(9) \quad \int_t^{t+L} \phi(s) \phi(s)^T ds \geq \alpha I_m.$$

Теорема 1. О достижении цели идентификации для сети из N моделей ФХН [4]. Если вектор-функция $z(x(t), t) = (x_2 \ x_3 \ x_4 \ x_5 \ 1)^T$ удовлетворяет условию постоянного возбуждения (9), то в системе (2), (3), (6), (8) цель идентификации (5) достигается при интегральном целевом функционале (7).

2.2. Процедура DREM

Альтернативным способом идентификации уравнений вида (4) является процедура DREM [6, 7]. Перед определением закона настройки параметров уравнение регрессии преобразуется в два этапа. На первом шаге проводится процедура расширения регрессора: к уравнению (4) применяются линейные устойчивые фильтры. Результатом такой фильтрации будет новое многомерное уравнение регрессии:

$$(10) \quad X_1(t) = Z(t)\theta^*,$$

где $X_1(t) \in \mathbb{R}^5$, $Z(t) \in \mathbb{R}^{5 \times 5}$. Одним из основных подходов к расширению регрессора является подход, предложенный Крейссельмейером [8], в котором $X_1(t)$ и $Z(t)$ из (10) получаются, как решения дифференциальных уравнений:

$$(11) \quad \begin{aligned} X_1'(t) &= -lX_1(t) + z(t)x_1(t), X_1(t_0) = 0_5, \\ Z'(t) &= -lZ(t) + z(t)z^T(t), Z(t_0) = 0_{5 \times 5}, \end{aligned}$$

где $l > 0$ – настраиваемый коэффициент. Второй шаг – процедура смешивания регрессора: обе части уравнения (10) умножаются на присоединенную к $Z(t)$ матрицу:

$$(12) \quad \tilde{x}_1(t) = \Delta(t)\theta^*,$$

где $\tilde{x}_1(t) = \text{adj}\{Z(t)\}X_1(t)$, $\Delta(t) = \det\{Z(t)\}$. Затем для (12) задается градиентный закон настройки параметров аналогичный (8):

$$(13) \quad \theta'(t) = -\Gamma\Delta(t)(\Delta(t)\theta(t) - \tilde{x}_1(t)).$$

Основным преимуществом закона (13) является то, что достижение цели идентификации будет обеспечено при выполнении в некотором смысле более слабого условия, чем условие постоянного возбуждения $z(x, t)$, а именно $\Delta(t) \notin L_2$. Более того, для процедуры расширения (11) было доказано в [9], что если условие постоянного возбуждения выполнено для $z(x, t)$, то оно будет выполнено и для $\Delta(t)$.

3. Результаты компьютерного моделирования

В Simulink были промоделированы системы (2), (3), (8) и (2), (3), (11), (13) в первую очередь для одной модели ФХН с начальными данными и параметрами:

$$(14) \quad y(0) = 0.7, y'(0) = 0.4, \theta(0) = (-0.9 \quad 0.02 \quad 0.8 \quad -0.1 \quad 0.15)^T$$

$$(15) \quad I = 1, a = 0.7, b = 0.1, \epsilon = 1/12.5, c = 0.9, \tau_1 = \tau_2 = 0.01, \gamma_i = 1 \quad i \in 1 : 5.$$

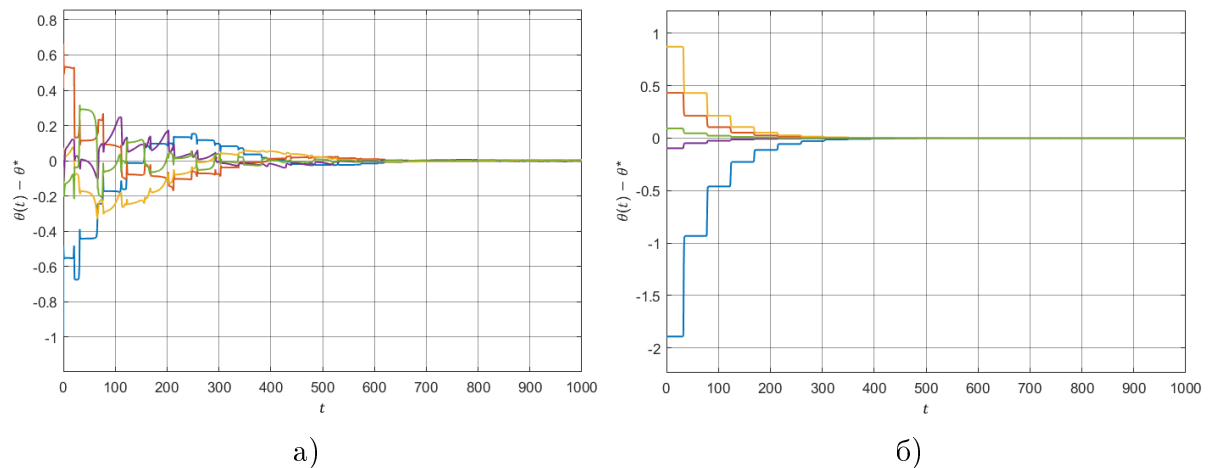


Рис. 1. Графики ошибок оценивания параметров θ^* одной модели ФХН с начальными данными (14) и параметрами (15) в случае применения а) метода скоростного градиента б) DREM с $l = 0.6$

На рис.1 видно, что время сходимости параметров к их истинным значениям меньше для алгоритма, полученного с помощью DREM, чем для классического алгоритма скоростного градиента. Однако, с точки зрения объема вычислений DREM значительно уступает: для достижения точности оценивания каждого из параметров равной 10^{-2} в случае а) Simulink потребовалось около трех секунд, тогда как в случае б) такая же точность достигается почти за четыре минуты.

Такое увеличение времени моделирования объясняется усложнением моделируемой системы: применение DREM вводит дополнительную фильтрацию, дополнительные арифметические операции (вычисление определителя и сопряженной матрицы) и увеличивает размерность исследуемой системы (если представить решения в виде систем уравнений в пространстве состояний, то (3), (8) имеет размерность 9, тогда как (3), (11), (13) – размерность 39). Это может стать серьезным препятствием при моделировании активности нейронов в реальном времени, как и необходимость выбирать настраиваемый коэффициент l из (11). Незначительное его изменение приводило к возникновению сингулярностей в решении, либо наоборот к практически моментальному обращению решения в ноль.

4. Заключение

Полученные результаты показывают, что эффективным способом оценивания параметров сети из N моделей нейронов ФХН в случае неизмеряемости переменных совокупного действия медленных ионных токов и производных каждой из переменных является приведение исходной системы к уравнению линейной регрессии с помощью фильтров-дифференциаторов и последующая идентификация полученной регрессии с помощью метода скоростного градиента или с помощью процедуры DREM. Преимуществами применения DREM в этой задаче являются ослабление условия постоянного возбуждения, которое, однако, всегда выполнялось во время наших численных экспериментов с моделями ФХН, а также уменьшение времени, за которое достигается цель идентификации. Его существенный на практике недостаток по сравнению с методом скоростного градиента – значительное увлечение вычислительной сложности.

Список литературы

1. FitzHugh R. Impulses and physiological states in theoretical models of nerve membrane // Biophysical Journal. 1961. Vol. 1. P. 445–466.
2. Nagumo J., Arimoto S., Yoshizawa S. An Active Pulse Transmission Line Simulating Nerve Axon // Proceedings of the IRE. 1962. Vol. 50, No. 10. P. 2061–2070.
3. Rybalko A., Fradkov A. Identification of two-neuron FitzHugh–Nagumo model based on the speed-gradient and filtering // Chaos: An Interdisciplinary Journal of Nonlinear Science. 2023. Vol. 33, No. 8. P. 083126.
4. Rybalko A., Fradkov A. Identification of Group of FitzHugh–Nagumo Neuron Models Based on the Speed-Gradient and Filtering // 7th Scientific School Dynamics of Complex Networks and their Applications (DCNA). Kalinigrad, Russia, 2023. IEEE, 2023. P. 233–236.
5. Фрадков А. Л. Адаптивное управление в сложных системах: беспойсковые методы. СПб.: Наука, 1990. 296 с.
6. Aranovskiy S., Bobtsov A., Ortega R., Pyrkin A. Parameters estimation via dynamic regressor extension and mixing // ACC. Boston, USA, 2016. IEEE, 2016. P. 6971–6976.
7. Glushchenko A., Lastochkin K. Exact asymptotic estimation of unknown parameters of regression equations with additive perturbations // arXiv preprint arXiv:2310.14073. 2023.
8. Kreisselmeier G. Adaptive observers with exponential rate of convergence // IEEE Transactions on Automatic Control. 1977. Vol. 22, No. 1. P. 2–8.
9. Aranovskiy S., Ushirobira R., Vedyakov M., Korotinaand A. On preserving-excitation properties of Kreisselmeier’s regressor extension scheme // IEEE transactions on automatic control. 2022. Vol. 68, No. 2. P. 1296–1302.