УДК 681.5, 004.8

ВЫЧИСЛЕНИЕ ДОПУСТИМОЙ СКОРОСТИ ОБУЧЕНИЯ НЕЙРОСЕТЕВОГО РЕГУЛЯТОРА В ЗАДАЧЕ СТАБИЛИЗАЦИИ БАЛАНСИРУЮЩЕГО РОБОТА

А.И. Глущенко
Старооскольский технологический институт им. А.А. Угарова (филиал) ФГАОУ ВО «Национальный исследовательский технологический университет «МИСиС»
Россия, 309516, Старый Оскол, Макаренко, 42
E-mail: a.glushchenko@sf-misis.ru

В.А. Петров
Старооскольский технологический институт им. А.А. Угарова (филиал) ФГАОУ ВО «Национальный исследовательский технологический университет «МИСиС»
Россия, 309516, Старый Оскол, Макаренко, 42
E-mail: 79040882508@ya.ru

К.А. Ласточкин
Старооскольский технологический институт им. А.А. Угарова (филиал) ФГАОУ ВО «Национальный исследовательский технологический университет «МИСиС»
Россия, 309516, Старый Оскол, Макаренко, 42
E-mail: lastconst@yandex.ru

Ключевые слова: балансирующий робот; нейросетевое управление; устойчивость; скорость оперативного обучения.

Аннотация: В работе решается задача управления балансирующим роботом на основе применения нейронной сети. Они выступает в роли регулятора и формирует на своем выходном слое управляющее воздействие для объекта (напряжения для левого и правого двигателей). Оперативное обучение такой нейронной сети необходимо для улучшения качества управления роботом в условиях изменения его параметров или смены режима работы. При реализации такого обучения актуальным является вопрос о выборе моментов времени, когда оно необходимо, и величины его шага. Именно поэтому в работе была рассмотрена проблема выбора предельной скорости оперативного обучения, непосредственно связанная с оценкой устойчивости изучаемой системы управления, поскольку излишне высокие скорости обучения могут привести к переходу объекта в неустойчивое состояние. В работе предложен подход, основанный на втором методе Ляпунова и позволяющий, не имея модели объекта управления, определять верхний допустимый предел для скорости обучения нейронной сети в текущий момент времени в различных ситуациях.

1. Введение

Ранее, на основе анализа недостатков существующих методов управления двухколесным балансирующим роботом, коллективом авторов данной работы был предложен собственный алгоритм прямого нейросетевого управления [1, 2]. Причем обучение такого нейросетевого регулятора происходит не автономно, а оперативно в моменты времени, строго определяемые системой правил и ограничений, учитывающей особенности функционирования объекта.

Недостатком разработанного нейросетевого регулятора является постоянство скоростей обучения скрытого и выходного слоя нейронной сети и отсутствие правил, останавливающих оперативное обучение в моменты времени, когда оно может привести к неустойчивости балансирующего робота.

В данном исследовании предлагается решение этих проблем за счет оценки устойчивости балансирующего робота по второму методу Ляпунова в оперативном режиме, а далее – вычисления на основании этой оценки максимально допустимых скоростей обучения нейронной сети.

2. Оценка устойчивости балансирующего робота

Функция Ляпунова для линейных систем является квадратичной формой фазовых координат системы [3].

В предлагаемой схеме управления регулятор, в роли которого выступает нейросеть, принимает на вход вектор рассогласований из семи элементов $E = [e_{\theta int}\, e_\theta\, e_\psi\, e_{\theta'}\, e_{\psi'}\, e_\varphi\, e_{\varphi'}]^T$. θ – угол поворота колес робота (θ_{int} – интеграл θ, θ’ – производная θ), ψ – угол отклонения робота от нормали в вертикальной плоскости (ψ’ – производная ψ), φ – угол поворота робота в горизонтальной плоскости (φ’ – производная φ). Сеть имеет семь входов, семь скрытых нейронов, два выходных нейрона. Матрица весовых коэффициентов скрытого слоя является единичной.

Обучение каждого из весовых коэффициентов каждого из двух нейронов выходного слоя (ввиду наличия двух двигателей у робота) – размер матрицы весов 7x2 – производится методом обратного распространения ошибки пропорционально рассогласованию по соответствующей координате, то есть одной из семи величин $e_{\theta int}$, e_θ, e_ψ, $e_{\theta'}$, $e_{\psi'}$, e_φ, $e_{\varphi'}$. И именно эти величины одновременно являются и фазовыми координатами системы, и элементами, вносящими нелинейность, поскольку если они равны нулю, то обучение сети не произойдет, а весовая матрица выходного слоя не изменится.

На основании вышеизложенного предлагается следующий вид (1) функций Ляпунова для оценки устойчивости нейросетевой системы управления балансирующим роботом по всем семи его фазовым координатам:

(1) $$V(e_i) = \frac{1}{2} e_i^2(t), \quad i = 1,..,7$$

Такая функция является знакоопределенной. Согласно теореме Ляпунова, если существует знакоопределенная функция $V(e_i)$, полная производная которой по времени в силу дифференциальных уравнений является знакопостоянной функцией противоположного с $V(e_i)$ знака (не положительной, в данном случае), то невозмущенное движение устойчиво [4]. Вычислим производную функции V (2) по времени:

(2) $$\Delta V_i(t) = \frac{V_i(t) - V_i(t-\Delta t)}{\Delta t} = \frac{1}{2\Delta t}\left(e_i^2(t) - e_i^2(t-\Delta t)\right) = \\ = \frac{1}{2\Delta t}\cdot\left(\left(y_i(t) - r_i(t)\right)^2 - \left(y_i(t-\Delta t) - r_i(t-\Delta t)\right)^2\right).$$

Используя указанное выше выражение, в любой момент работы системы в реальном времени производится проверка, выполняется ли достаточное условие устойчивости объекта по каждой из координат состояния. Для этого по интересующей координате берутся известные значения задания $r_i(t)$ и выхода балансирующего робота $y_i(t)$ в текущий момент и предыдущий момент $r_i(t - \Delta t)$ и $y_i(t - \Delta t)$.

Далее, имея подобную оценку устойчивости, возможно разработать метод определения предельно допустимого значения скорости обучения каждого из весовых коэффициентов каждого из нейронов выходного слоя нейросетевого регулятора.

3. Вычисление предельно допустимой скорости оперативного обучения нейронной сети

Рассмотрим алгоритм работы контура управления в зависимости от знака производной функции Ляпунова.

Сначала вызывается нейросетевой регулятор. Далее по разработанному ранее алгоритму [2] выбирается знак коррекций весов выходного слоя.

Для всех координат (i = 1, …, 7) вычисляется $\Delta V_i(t)$. Дальнейшие оценки ведутся для каждой из $\Delta V_i(t)$ (i = 1, …, 7). Здесь возможны два варианта.

1) Значение $\Delta V_i(t)$ положительно или равно нулю. То есть система на границе устойчивости или неустойчива по Ляпунову по координате i. Для рассматриваемого ОУ в данном случае необходимо вернуть нейросетевому регулятору весовые коэффициенты, использовавшиеся в течение последнего устойчивого переходного процесса по данной координате.

2) Значение $\Delta V_i(t)$ отрицательно. Скорость обучения выбранного весового коэффициента выбранного нейрона выходного слоя, отвечающего за рассогласование по рассматриваемой координате, должна определяться исходя из имеющегося запаса устойчивости. Его предлагается определять следующим образом. Необходимо решить квадратное неравенство (2) относительно $y_i(t)$.

2.1) Дискриминант квадратного уравнения D, полученного из неравенства (2), является положительным (так как текущее значение $\Delta V_i(t) < 0$), т.е. будет получен интервал значений ($y_{i.min}$; $y_{i.max}$). При этом $y_i(t)$ принадлежит интервалу ($y_{i.min}$; $y_{i.max}$). При этом, если задание по рассматриваемой координате равно нулю (режим стабилизации робота), то $y_{i.min}$ и $y_{i.max}$, исходя из решения квадратного неравенства, являются равными по модулю величинами, которые отличаются знаком. Поэтому в качестве желаемого значения $y_{i.st}$ рассматриваемой координаты среди значений $y_{i.min}$ и $y_{i.max}$ выбирается то, которое в текущий момент времени совпадает по знаку с рассматриваемой координатой y_i.

2.2) Далее предположим, что выходной сигнал по всем координатам y_i был бы равен $y_{i.st}$. Тогда изменились бы значения рассогласований по каждой координате $e_{i\,st} = r_i - y_{i.st}$. Подадим на вход нейронной сети сначала реальный вектор рассогласований E и вычислим выходы сети $O_k^{(2)}$ (k =1, 2). А затем подадим на ту же сеть вектор E_{st}, где на каждой i-ой позиции вместо e_i будет стоять $e_{i\,st}$, вычислив тем самым на выходе сети «желаемое» управляющее воздействие $O_{k.st}^{(2)}$ (k =1, 2). Далее вычислим разницу $\Delta O_k^{(2)} = (O_k^{(2)} - O_{k.st}^{(2)})$. Эта величина является максимально допустимой коррекцией управляющего воздействия.

2.3) Зная значение $\Delta O_k^{(2)}$, максимально допустимая коррекция весового коэффициента между k-м нейроном выходного слоя и j-м нейроном скрытого слоя $\omega_{kj}^{(2)}$ рассчитывается на основе выражений (3) и (4).

(3) $$\Delta\omega_{kj}^{(2)} = \eta_{kj}^{(2)}\delta_k^{(2)}O_j^{(1)}, \quad k = \overline{1,2}.$$

(4) $$O_k^{(2)} = \sum_{j=1}^{N_{hidden}} \omega_{kj}^{(2)} \cdot O_j^{(1)}, \quad k = \overline{1,2}.$$

Здесь $\eta_{kj}^{(2)}$ – скорость обучения для веса, стоящего на связи между j-м нейроном скрытого слоя и k-м нейроном выходного слоя; $\delta_k^{(2)}$ – совокупная ошибка k-го нейрона выходного слоя (при условии линейности функции активации выходного слоя она совпадает со значением разницы между «желаемым» и реальным управлением $\Delta O_k^{(2)}$), $O_j^{(1)}$ – выход j-го нейрона скрытого слоя, N_{hidden} – число нейронов скрытого слоя.

Добавим в уравнение (4) знак «дельта» слева и справа от знака равенства – (5).

(5) $$\Delta O_k^{(2)} = \sum_{j=1}^{N_{hidden}} \Delta\omega_{kj}^{(2)} \cdot O_j^{(1)}, \quad k = \overline{1,2}.$$

Рассмотрим каждую i-ю ($j = i$ – данное утверждение правомерно, так как в скрытом слое сети матрица весовых коэффициентов является единичной) координату по отдельности. Тогда выражение (5) можно рассмотреть как сумму коррекций управления $\Delta O_{kj}^{(2)}$, вносимых каждой i-й ($j = i$) координатой (6).

(6) $$\Delta O_{kj}^{(2)} = \Delta\omega_{kj}^{(2)} \cdot O_j^{(1)}, \quad j = i, k = \overline{1,2}.$$

$\Delta O_{kj}^{(2)}$ вычисляется с помощью (7) ($j = i$, k =1, 2).

(7) $$\Delta O_{kj}^{(2)} = O_{kj}^{(2)} - O_{kj\,st}^{(2)} = \omega_{kj}^{(2)} \cdot O_j^{(1)} - \omega_{kj}^{(2)} \cdot O_{j\,st}^{(1)} = \omega_{kj}^{(2)} \cdot e_j - \omega_{kj}^{(2)} \cdot e_{j\,st}.$$

Подставим уравнение (3) в (6) и получим (8) ($j = i$, k =1, 2).

(8) $$\Delta O_{kj}^{(2)} = \eta_{kj}^{(2)} \delta_k^{(2)} O_j^{(1)} \cdot O_j^{(1)} = \eta_{kj}^{(2)} \delta_k^{(2)} (O_j^{(1)})^2 = \eta_{kj}^{(2)} \Delta O_k^{(2)} (O_j^{(1)})^2.$$

Выразим $\eta_{kj}^{(2)}$ из (8), не учитывая знак $\delta_k^{(2)}$, так как он определяется в соответствии с алгоритмом, описанным в работе [2]. Получим предельную скорость обучения, которую можно обозначить как $\eta_{kj\,max}^{(2)}$ и вычислять по формуле (9). Именно формула (9) будет использоваться в эксперименте на балансирующем роботе:

(9) $$\eta_{kj\,\max}^{(2)} = \frac{\Delta O_{kj}^{(2)}}{\delta_k^{(2)} (O_j^{(1)})^2} = \frac{\Delta O_{kj}^{(2)}}{\Delta O_k^{(2)} (O_j^{(1)})^2}, \quad k = \overline{1,2}.$$

4. Апробация разработанного метода на реальном роботе

Экспериментальная проверка разработанного метода вычисления скорости проводилась в режиме стабилизации (нулевые задания по всем координатам состояния) на двухколесном роботе на платформе Lego EV3 (рис. 1). Оценка качества переходных процессов в этом режиме работы выражается в сравнении окончательных значений координаты θ_{int} для систем управления с обычным LQ-регулятором и нейронной сетью.

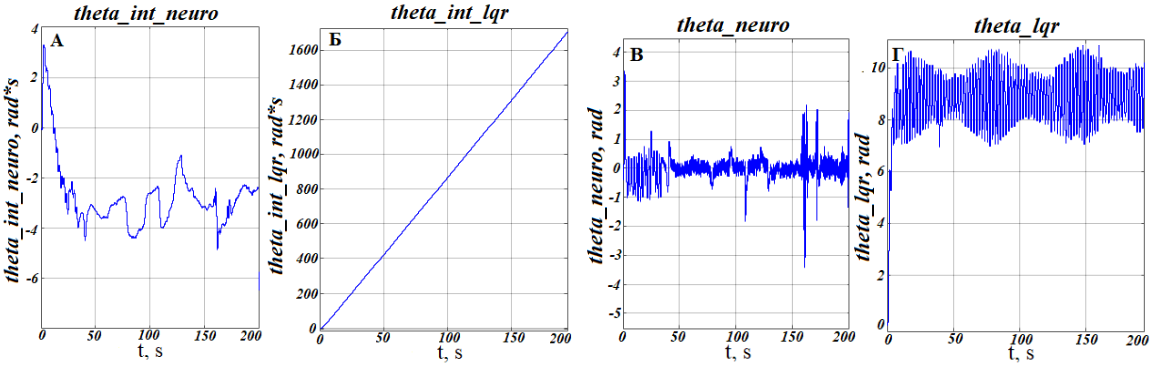

Рис. 1. Переходные процессы балансирующего робота по координатам θ_{int} и θ (а, в – нейронная сеть, б, г – LQ-регулятор).

На рис. 1 представлены переходные процессы балансирующего робота по координатам θ_{int} и θ под управлением нейронной сети (а, в) и LQ регулятора (б, г). Из характеристик заметно, что при управлении с помощью нейросети статическая ошибка по положению (координате θ) отсутствует, а также за все время эксперимента объект отъехал от исходной точки на меньшее расстояние (2,3 радиан×секунду / 200 секунд = 0,012 радиан), чем при управлении LQ регулятором (1650 радиан*секунду / 200 секунд = 8,25 радиан).

На рис. 2 изображен график предельно допустимой скорости обучения весового коэффициента, отвечающего за рассогласование по координате θ_{int} для первого выходного нейрона. По остальным координатам характер изменения скорости имеет вид, схожий с представленным. Отличие заключается в амплитудных значениях скоростей.

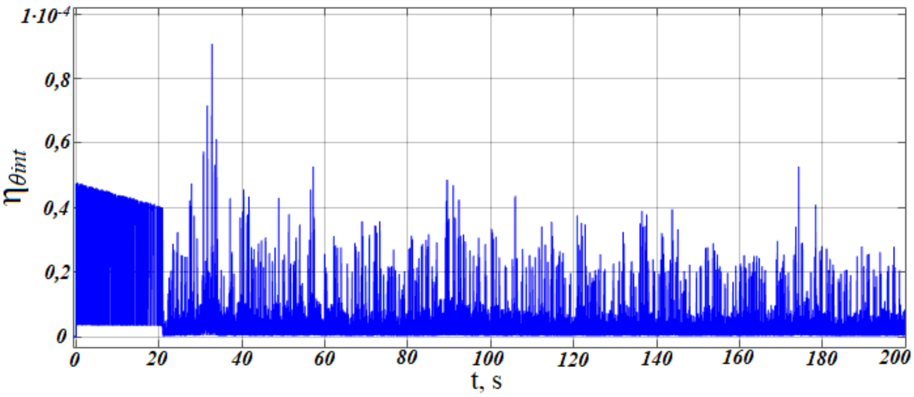

Рис. 2. Предельная допустимая скорость оперативного обучения веса, отвечающего за рассогласование по координате θ_{int} для первого выходного нейрона сети.

5. Заключение

В работе предложен подход к вычислению верхней допустимой скорости оперативного обучения нейронной сети, выполняющей роль регулятора в контуре управления балансирующим роботом, основанный на использовании функции Ляпунова и оценке устойчивости такой системы управления.

Результаты эксперимента в режиме реального времени позволяют сделать выводы об эффективности предложенного подхода. Разработанный метод позволяет повысить качество управления балансирующим роботом в режиме стабилизации по сравнению с LQ регулятором. При этом в течение всего эксперимента система управления оставалась устойчивой с точки зрения выбранных критериев.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (18-47-310003-р_а).

Список литературы

1. Glushchenko A.I., Petrov V.A., Lastochkin K.A. Adaptive Neural Network Based Control of Balancing Robot in Real Time Mode // CEUR Workshop Proceedings. 2018. Vol. 2258. P. 168-178.
2. Glushchenko A.I., Petrov V.A., Lastochkin K.A. On Development of Neural Network Controller with Online Training to Control Two-Wheeled Balancing Robot // Proceedings of 2018 International Russian Automation Conference (RusAutoCon). IEEE, 2018. P. 1-6.

3. Song Y., Guo J., Huang X. Smooth Neuroadaptive PI Tracking Control of Nonlinear Systems with Unknown and Non-smooth Actuation Characteristics // IEEE Trans. on Neural Networks and Learning Systems. 2016. Vol. 99. P. 1-13.
4. Lyapunov A.M. The general problem of stability of motion // Int. J. of Control. 1992. Vol. 55. P. 531-534.