Метод определения связей между элементами стохастической нейронной сети

А.В. ВОЛИК[1], С.Н. ГЕРАСИН[1], Е.В. СЛИПЧЕНКО[2]

[1]Харьковский национальный университет радиоэлектроники,
[2]Национальный технический университет «Харьковский политехнический институт»

В статье рассматривается метод оценивания связей между элементами стохастической нейронной сети при наличии внешних воздействий. Для решения данной задачи предложена модификация метода Кокса, основанная на теории точечных процессов. Получены характеристики данного метода, отвечающие максимуму функции правдоподобия.

У статті розглядається метод оцінювання зв'язків між елементами стохастичної нейронної мережі за наявності зовнішніх дій. Для вирішення даного завдання запропонована модифікація методу Коксу, заснована на теорії точкових процесів. Отримані характеристики даного методу, що відповідають максимуму функції правдоподібності.

In the article the method of evaluation of connections is examined between the elements of stochastic neuron network at presence of external influences. For the decision of this task modification of method of Coke, based on the theory of processes of points, is offered. Descriptions of this method, answering a maximum of function of verisimilitude, are got.

Введение. Эксперименты по исследованию нейронных сетей на моделях, аккумулирующих тот или иной набор известных параметров и свойств реальных нейронов, позволяют воспроизводить некоторые особенности их активности, анализировать возможные механизмы и формулировать предположения, доступные экспериментальной проверке. Одним из направлений исследований в данной области является изучение организации нейронных структур как многомерных сетей и определение влияния свойств отдельных клеток на динамику активности и свойства всей сети. Микроподход к изучению взаимодействий элементов нейронной сети означает, что будут рассматриваться следующие задачи. Анализируя импульсную активность отдельных элементов сети (двух или трех), определить наличие взаимосвязей между этими элементами, определить относительные величины этих связей, изучить динамику изменения анализируемых связей при различных внешних воздействиях на сеть, т.е. провести анализ взаимосвязей между элементами в условиях, когда величины связей могут, изменяться. Такая постановка вопроса является актуальной в связи с целым рядом нейрофизиологических исследований. Сюда относятся вопросы о характере взаимосвязей в различных структурах мозга, вопросы о нейрофизиологических аспектах выработки условного рефлекса, запоминания информации и т.д. [1]

Известные статистические методы, по тем или иным причинам непригодны для решения сформулированных задач. Поэтому в работе разработан новый статистический метод для определения связей между элементами нейронной сети. Метод специально приспособлен для анализа нейронной импульсной активности и учитывает специфику генерации импульсов элементами сети. Новый метод является, в некотором математическом смысле, наилучшим из возможных методов, поскольку статистические оценки этого метода, получаются при максимизации условной функции правдоподобия. Он свободен от большинства недостатков, присущих другим методам, в частности, не предполагается стационарность исследуемых потоков, имеются возможности для учета нелинейности связей, легко обобщается на случай трех и более потоков.

Метод определения зависимости между элементами нейронной сети. В основе нового метода лежит статистический анализ зависимости точечных процессов, предложенный Д. Коксом [2]. Рассмотрим точечный процесс, подверженный внешнему воздействию. Основное предположение Кокса состоит в том, что риск срабатывания точечного процесса является произведением собственного риска (риска срабатывания при отсутствии воздействия) и функции, зависящей от воздействия:

$$\varphi(t) = \lambda_0(t)\exp(\beta z(t)).$$

Оцениваемый параметр β характеризует степень воздействия, а функция z(t) отражает структуру воздействия, $\lambda_0(t)$ - собственный риск. В частности, β =0 означает отсутствие воздействия. Кокс [2] предложил способ оценивания параметра β (вообще говоря, многомерного в случае многих воздействий), исходя из максимума условной функции правдоподобия, что позволяет получить хорошие асимптотические свойства оценки.

Статистический анализ зависимости точечных процессов. Пусть М - точечный процесс, вообще говоря, нестационарный и пусть в момент t* произошло событие в этом процессе. Функция $\varphi^{t^*}(t)$ называется функцией риска процесса М, если

$$\varphi^{t^*}(t) = \lim_{\Delta t \to 0} \frac{P_r\left\{M(t, t+\Delta t) = 1 \left| \begin{array}{l} M\{t^*\} = 1 \\ M(t^*, t) = 0 \end{array} \right.\right\}}{\Delta t},$$

то есть $\varphi^{t^*}(t)$ условная плотность вероятности срабатывания, при условии, что в момент t* было событие и от t* до t срабатываний не происходило.

Очевидно, что для пуассоновского процесса

$$\varphi^{t^*}(t) = \varphi(u(t)) = P_M ,$$

где P_M - интенсивность пуассоновского процесса, $u(t)$ - обратное время возвращения

Аналогично, для процесса восстановления риск в момент t зависит лишь от обратного времени возвращения

$$\varphi^{t^*}(t) = \varphi(u(t)) = \lambda_0(u(t)).$$

Предположим, что для нестационарного точечного процесса M функция риска $\varphi^{t^*}(t)$ имеет следующий вид:

$$\varphi^{t^*}(t) = \lambda_0(u(t))e^{\beta_1 Z_1(t)+\beta_2 Z_2(t)+...+\beta_k Z_k(t)}, \qquad (1)$$

где $Z_1(t)$, $Z_2(t)$, ... , $Z_k(t)$ - заданные функции; $\beta_1, \beta_2, ..., \beta_k$ - неизвестные параметры.

Процесс M, удовлетворяющий этому условию, называют модулированным процессом восстановления, поскольку собственный риск процесса $\lambda_0(u(t))$ как бы модулируется воздействующим вектором $Z_1(t)$, $Z_2(t)$, ... , $Z_k(t)$. В частном случае $\lambda_0(u(t)) = P_M$, процесс М называется модулированным пуассоновским процессом.

Пример I. Тренд. Пусть К=1 и $Z_1(t)=t$, тогда

$$\varphi^{t^*}(t) = \lambda_0(u(t))e^{\beta t}.$$

При малых t риск срабатывания процесса примерно равен собственному риску $\lambda_0(u(t))e^{\beta t}$ и экспоненционально возрастает с постом t (т.е. срабатывания процесса становятся все более частыми при увеличении t).

Пример 2. Синусоидальные изменения. Пусть k=1, $Z_1(t) = sin(\omega_0 t)$, тогда:

$$\varphi^{t^*} = \lambda_0(u(t))e^{\beta\, sin(\omega_0 t)}$$

В этом примере риск процесса изменяется синусоидально с известной частотой ω_0

Пример 3. Зависимость двух точечных процессов. Рассмотрим двумерный точечный процесс (А, В). Обозначим $u_8(t)$- обратное время возвращения для процесса В. Пусть $k=1, Z_1(t)=Y(u_B(t))$, тогда функция риска для процесса А будет иметь вид:

$$\varphi^{t^*}(t) = \lambda_0(u(t))e^{\beta Y(u_B(t))}.$$

Этот пример иллюстрирует простейшую возможность введения зависимости двух точечных процессов. Риск процесса А зависит от обратного времени возвращения в процессе В, которое предполагается известным и неслучайным. Можно показать, что для модельной нейронной сети риск срабатывания элемента А зависит от обратного времени возвращения $u_B(t)$ в процессе, генерируемом воздействующим на А элементом В, причем

$$Y(u_B(t)) = e^{-\frac{1}{2}u_B(t)}.$$

Значения неизвестных параметров β_i характеризуют зависимость основного процесса А от модулирующих воздействий. Если все β_i равны нулю, то процесс не зависит от воздействующей на него модуляции и риск срабатывания равен собственному риску, если же величина β_i отлична от нуля, то она характеризует степень воздействия модулирующего процесса.

Перейдем к построению оценок для параметров β_i методом условной функции правдоподобия [3]. Рассмотрим сначала случай модулированного пуассоновского процесса, то есть $\varphi(t) = pe^{\beta Z(t)}$, где для простоты возьмем случай К=1.

Пусть на интервале наблюдения $(0,T]$ события процесса появились в моменты t_1, t_2, ..., t_n. Тогда функция правдоподобия имеет вид:

$$L(p,\beta,t_1,t_2,...,t_n) = f^0(t_1)f^{t1}(t_2-t_1)...$$
$$...f^{t_{n-1}}(t_n - t_{n-1})(1-F^{tn}(T-t_n)),$$

где $F^{t^*}(x)$ - функция распределения длины интервала до очередного срабатывания (t* - момент предыдущего срабатывания); $f^*(x)$ - плотность этого распределения:

$$f(x) = \frac{\partial}{\partial x}F^{t^*}(x)$$

Поскольку:

$$\varphi^{t^*}(t^*+x) = \frac{f^{t^*}(x)}{1-F^{t^*}(x)}, \text{ то}$$

$$\frac{\partial}{\partial x}\left[ln(1-F^{t^*}(x)\right] = -\varphi^{t^*}(t+x)$$

и решая это уравнение, находим:

$$F^{t^*}(x) = 1 - e^{-S_{t^*}^{t^*+x}\varphi^{t^*}(u)du},$$

$$f^{t^*}(x) = \frac{\partial}{\partial x}F^{t^*}(x) = \varphi^{t^*}(t^*+x)e^{-S_{t^*}^{t^*+x}\varphi^{t^*}(u)du}.$$

Учитывая предположение о виде функции $\varphi^{t^*}(t^*+x)$ и подставляя полученные соотношения в формулу для функции правоподобия, получим

$$\alpha(p,\beta,t_1,t_2,...t_n) = p^n e^{\beta(Z(t_1)+Z(t_2)+...+Z(t_n))}\times$$
$$\times e^{-p\int_0^T e^{\beta Z(u)du}}$$

В функцию правдоподобия кроме параметра, который мы хотим оценить, входит неизвестный мешающий параметр р. Чтобы исключить его, воспользуемся тем, что число срабатываний процесса – n на интервале наблюдения (О,Т] является достаточной статистикой для р, и перейдем к условной функции правдоподобия [3]:

$$L(p,\beta;t_1,t_2,...,t_n|n) = \frac{\alpha(p,\beta;t_1,t_2,...,t_n)}{P_r\{A(O,T]=n\}}.$$

Известно, что для стационарного пуассоновского процесса число событий на интервале (О, Т] имеет распределение Пуассона [4], поэтому:

$$P_r\{A(O,T]=n\} = \frac{p^n\left[\int_O^T e^{\beta Z(u)}du\right]^{nT} e^{-p\int_O^T e^{\beta Z(u)}du}}{n!}.$$

Подставляя полученное выражение, логарифмируя и отбрасывая постоянные сомножители, получим

$$L_p(\beta)=\beta S-n\ln\left[\int_O^T e^{\beta Z(t)}dt\right],$$

где $S=\Sigma_{i=1}^{n}Z(t_i)$.

Обозначим

$$U_p(\beta)=\frac{\partial L_p(\beta)}{\partial\beta}=S-n\frac{\int_O^T Z(t)e^{\beta Z(t)}dt}{\int_O^T e^{\beta Z(t)}dt}=S-nA_p(Z,\beta)\ ,$$

$$I_p(\beta)=-\frac{\partial^2 L_p(\beta)}{\partial\beta^2}=$$

$$=n\frac{\left(\int_O^T Z^2(t)e^{\beta Z(t)}dt\right)\left(\int_O^T e^{\beta Z(t)}dt\right)-\left(\int_O^T Z(t)e^{\beta Z(t)}dt\right)^2}{\left(\int_O^T e^{\beta Z(t)}dt\right)^2}=$$

$$=n\left[A_p(Z^2,\beta)-A_p^2(Z,\beta)\right],$$

где $$A_p(Z,\beta)=\frac{\int_O^T Z(t)e^{\beta Z(t)}dt}{\int_O^T e^{\beta Z(t)}dt}\ .$$

Известно, что $U_p(\beta)$ имеет асимптотически нормальное распределение со средним ноль и дисперсией $I_p(\beta)$ [3]. Отсюда находим условия для вычисления $(1-\alpha)$ доверительного интервала параметра β

$$\left\{\beta:\left|S-nA_p(Z,\beta)\right|\prec K_{\alpha/2}\sqrt{I_p(\beta)}\right\},$$

где $$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-k_\alpha}e^{-\frac{u^2}{2}}du=\alpha\ .$$

С вычислительной точки зрения это условие не является простым. Можно рассчитать требуемую статистику $U_p(\beta)$ при различных β и, интерполируя, найти такие β при которых выполняется неравенство.

Рассмотрим один частный случай, для которого полученная формула имеет простой вид. Предположим, что $Z(t)=u_\beta(t)$ и найдем критическое соотношение для проверки гипотезы о равенстве нулю параметра β. Для этого случая

$$A_p(Z,0)=\frac{1}{2}\cdot\frac{\Sigma_{j=1}^{m}(x_j^B)^2}{T}=\frac{1}{2}m_2^B\ ,$$

$$A_p(Z^2,0)=\frac{1}{3}\cdot\frac{\Sigma_{j=1}^{m}(x_j^B)^3}{T}=\frac{1}{3}m_3^B$$

и величина

$$\xi=\frac{\Sigma_{i=1}^{n}u_B(t_i)-\frac{1}{2}nm_2^B}{\sqrt{n\left[\frac{1}{3}m_3^B-\frac{1}{4}(m_2^B)^2\right]}}$$

имеет нормальное распределение со средним 0 и дисперсией I. Здесь мы обозначили через $x_1^B,x_2^B,\ldots,x_m^B$ интервалы между моментами появления событий в процессе В, кроме того, для простоты предполагается, что интервал (О,Т] начинается и заканчивается событием процесса В.

Предполагаем, что эти моменты являются фиксированными, а не случайными, и анализируем зависимость процесса А от процесса В при условии известных зафиксированных моментов появления событий процесса В на интервале наблюдения (О, Т].

Рассмотрим теперь случай модулированного процесса восстановления А, для которого функция риска имеет вид

$$\varphi^{t^*}(t)=\lambda_0(u(t))e^{\beta Z(t)}\ .$$

Здесь предполагается, что $Z(t)$ - модулирующее воздействие - фиксированная (не случайная) функция текущего времени t, если $Z(t)$ связана с реализацией другого случайного процесса, то всю процедуру оценивания проводим при условии, что задана реализация воздействующего процесса и обращаемся с $Z(t)$ как с неслучайной величиной; $\lambda_0(u(t))$ - неизвестный собственный риск, который является мешающей функцией и должен быть исключен из окончательных оценок; β - неизвестный параметр, который требуется оценить (для простоты рассматривается одномерный случай).

Для оценивания β используется не безусловная функция правдоподобия, содержащая неизвестную мешающую функцию $\lambda_0(\cdot)$, а условная функция правдоподобия, при условии, что известны длины межимпульсных интервалов $x_1, x_2, \ldots, x_n$ в процессе А. Это условие позволяет исключить мешающую функцию и получить оценку для параметра β, обладающую хорошими асимптотическими свойствами.

Длины межимпульсных интервалов, упорядоченные по возрастанию, обозначим $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$, так что $x_{(1)}\prec x_{(2)}\prec\ldots\prec x_{(n)}$, причем предполагается, что все интервалы различны. Обозначим Z_{ij} следующую величину: пусть $i\geq j$, тогда интервал $x(j)\leq x_{(i)}$, величина Z_{ij} зависит от интервала $x_{(j)}$ но вычисляется в момент $x_{(i)}$,от начала интервала $x_{(i)}$. Другими словами: к началу большего интервала $x_{(i)}$ прикладываем меньший интервал $x_{(j)}$ и в момент времени, соответствующий концу интервала $x_{(j)}$ вычисляем величину Z_{ij}. Величина Z_{ij} вычисляется в момент, соответствующий концу интервала $x_{(i)}$.

Аналогично тому, как это было сделано для модулированного пуассоновского процесса, можно получить функцию правдоподобия для модулированного процесса восстановления

$$\alpha\left(\lambda_0(u(t)),\beta;x_1,x_2,\ldots,x_n\right)=$$

$$\lambda_0(x_{(1)})e^{\beta Z(x_{(1)})}e^{-\int_o^{x(1)}\lambda_o(u)e^{\beta Z(u)}du}\times$$

$$\times\lambda_o(x_{(2)})e^{\beta Z(x_{(1)}+x_{(2)})}e^{-\int_o^{x(2)}\lambda_o(u)e^{\beta Z(x(1)+u)}}\times\ldots$$

$$\times\lambda_o(x_{(n)})e^{\beta Z(x_{(1)}+x_{(2)}+\ldots+x_{(n)})}e^{-\int_o^{x(n)}\lambda_o(u)e^{\beta Z(x(1)+\ldots+x(n-1)+u)}du}$$

Для того чтобы написать условное правдоподобие, нужно безусловную функцию правдоподобия разделить на вероятность условия:

$$
P_r\left\{\begin{array}{l}\text{длины интервалов}\\ \text{равны } x_{(1)}, x_{(2)}, \ldots, x_{(n)}\end{array}\right\} = P_r\left\{\begin{array}{l}\text{минимальный интервал}\\ \text{равен } x_{(1)}\end{array}\right\}\times
$$

$$
\times P_r\left\{\begin{array}{l}\text{второй по величине}\\ \text{интервал равен } x_{(2)}\end{array}\middle|\begin{array}{l}\text{минимальный}\\ \text{равен } x_{(1)}\end{array}\right\}\times\ldots\times
$$

$$
\times P_r\left\{\begin{array}{l}\text{максимальный по вели –}\\ \text{чине интервал равен } x_{(n)}\\ \\ \\ \end{array}\middle|\begin{array}{l}\text{минимальный равен } x_{(1)},\\ \text{второй равен } x_{(2)}, \ldots,\\ (n-1)-\text{й по величине}\\ \text{равен } x_{(n-1)}\end{array}\right\}.
$$

Самый короткий интервал $x_{(1)}$ дает следующий вклад в эту вероятность:

$$
\lambda_o(x_{(1)})e^{-\int_o^{x(1)}\lambda_o(u)e^{\beta Z_{11}}du}\left(e^{\beta Z_{11}}+e^{\beta Z^{21}}+\ldots+e^{\beta Z_{n1}}\right).
$$

Вклад от следующего по величине интервала $x_{(2)}\succ x_{(1)}$ при условии, что минимальный интервал $x_{(1)}$ известен, будет

$$
\lambda_o(x_{(2)})e^{-\int_o^{x(2)}\lambda_o(u)e^{\beta Z22}du}\left(e^{\beta Z_{22}}+e^{\beta Z^{22}}+\ldots+e^{\beta Z_{n2}}\right).
$$

и т. д.

При выводе формул используется соображение, что информация о β не может быть получена кроме как из наблюдаемых интервалов между появлениями событий, то есть $x_1, x_2, \ldots, x_n$; поэтому для всех других значений аргумента, функция $\lambda_o(\cdot)$ тождественно равна нулю. Подставляя полученные выражения в формулу для условной функции правдоподобия, получаем

$$
L(\beta; x_1, x_2, \ldots, x_n) = \frac{e^{\beta\Sigma_{i=1}^{n}Zii}}{Z_{i=1}^{n}e^{\beta Z_{i2}}\times\ldots\times\Sigma_{i=1}^{n}e^{\beta Z_{in}}}.
$$

Логарифмируя, получим

$$
L_R(\beta) = \beta S - \Sigma_{i=1}^{n}\log\left(\Sigma_{l=i}^{n}e^{\chi Z_{li}}\right),
$$

где S -то же самое, что и в пуассоновском случае.

Обозначим:

$$
U_R(\beta) = \frac{\partial L_R(\beta)}{\partial\beta} = S - nA_R(Z,\beta),
$$

$$
I_R(\beta) = -\frac{\partial^2 L_R(\beta)}{\partial\beta^2} = n\left[A_R(Z^2,\beta) - \frac{1}{n}\Sigma_{i=1}^{n}\frac{\left(\Sigma_{l=i}^{n}Z_{li}e^{\beta Z_{li}}\right)^2}{\left(\Sigma_{l=i}^{n}e^{\beta Z_{li}}\right)^2}\right],
$$

ãäå $A_R(Z,\beta) = \frac{1}{n}\Sigma_{i=1}^{n}\left[\frac{\Sigma_{l=i}^{n}Z_{li}e^{\beta Z_{li}}}{\Sigma_{l=i}^{n}e^{\beta Z_{li}}}\right]$.

И совершенно аналогично тому, как это делается в пуассоновском случае, строится доверительный интервал для параметра β и выписывается критическое отношение для проверки гипотезы $\beta=0$. В случае, когда β есть вектор размерности $k(k\ge 2)$, все формулы получаются аналогично. Для проверки гипотезы $\vec{\beta}=0$ рассматривается величина

$$
\eta = U_R^T(0)[I_R(0)]^{-1}U_R(0),
$$

где U_R^T транспонированный вектор первых производных условной функции правдоподобия; $[I_R(O)]^{-1}$ - матрица, обратная к матрице вторых производных.

Величина η имеет распределение X^2 с k степенями свободы. Поэтому для построения границы доверительной области нужно найти такие значения параметров β, для которых $\eta = \chi^2_{крит.}$, где $\chi^2_{крит.}$, - критическое значение, соответствующее выбранной доверительной вероятности.

Выводы. В статье предложен метод для определения взаимосвязи между элементами нейронной сети, он не предполагает стационарности входного потока, учитывает нелинейный характер связей между элементами нейронной сети. Он может быть использован при анализе активности сети и позволяет производить оценку динамических характеристик нейронной сети.

ЛИТЕРАТУРА

1. *Круглов В.В., Борисов В.В.* Искусственные нейронные сети. -М.: Горячая линия – Телеком, 2001. – С. 324-325.
2. *Cox D. R.* The statistical analysis of dependencies in point processes // Stochastic point processes (P. A. Lewis ed.).- N.-Y.: Willey, 1972.-P. 55-66.
3. *Кокс Д.*, Льюис П. Статистический анализ последовательностей событий. - М.: Мир.- 1969.- 310 с.
4. *Петунин. Ю. И.* Приложение теории случайных процессов в биологии и медицине. – К.: Наукова думка, 1981.- 320 с.

пост. 09.03.06.