ліку [Електронний ресурс] / Н.О. Ткач //Вісник Національного університету водного господарства та природокористування. – 2012. – № 2 (58). – Режим доступу: http://old. nuwm.rv.ua/metods/asp/vd1/Ve5828.pdf

7. Сенів Б.Г. Шляхи удосконалення оцінки ефективності інноваційної діяльності підприємства [Електронний ресурс] / Б.Г. Сенів // Інноваційна економіка. – 2013. – №7 (45). – Режим доступу: http://nbuv.gov.ua/UJRN/inek_2013_7_18

8. Фомичев А. Индексация основных средств: новые методы налоговиков для выполнения показателей при проверке [Электронный ресурс] / А. Фомичев // Бухгалтер 911. – 2014. – Режим доступа: http:/buhgalter911. com/ShowArticle. aspx?a=107208

9. Ущаповський К.В. Удосконалення науково-практичних підходів до оцінки вартості основних засобів ДП «НЕК «Укренерго» [Електронний ресурс] / К.В. Ущаповський//Європейський вектор економічного розвитку. – 2015. – №1(18). – Режим доступу: http://duan.edu.ua/uploads/vidavnitstvo14-15/12139.pdf

---

А.О. СІГАЙОВ,
д.е.н., професор НТУУ «Київський політехнічний інститут імені Ігоря Сікорського»
А.В. ВОЛОВИК,
аспірант НТУУ «Київський політехнічний інститут імені Ігоря Сікорського»

Статистичні моделі нейронних мереж для прогнозування вартості медичного страхування

Нейронні мережі вже давно стали популярним методом нелінійного статистичного прогнозування, отже уявляється актуальним застосувати цю методику до прогнозування вартості медичного страхування. В цій роботі досліджується мережа, заснована на стохастичній моделі, що має багаторівневу архітектуру прямого зв'язку з випадковими зв'язками між модулями і частотною характеристикою з перешкодами. Отримана байєсівська методика, виведена шляхом логічного розв'язку для цієї моделі, базується на фільтрі Калмана. Отриманий при цьому алгоритм навчання узагальнює так званий одновимірний метод Ньютона, що втілює алгоритм, популярний нині в літературі з нейронних мереж. У статті представлений чисельний метод вивчення прогнозування вартості медичного страхування у вигляді хаотичних часових рядів з похибками і показана більш висока точність прогнозування нового алгоритму порівняно з існуючими.

Ключові слова: *витрати медичного страхування, статистичне прогнозування, нейронні мережі.*

А.А. СИГАЁВ,
д.э.н., профессор НТУУ «Киевский политехнический институт имени Игоря Сикорского»
А.В. ВОЛОВИК,
аспирант НТУУ «Киевский политехнический институт имени Игоря Сикорского»

Статистические модели нейронных сетей для прогнозирования стоимости медицинского страхования

Нейронные сети уже давно стали популярным методом нелинейного статистического прогнозирования, следовательно, представляется актуальным применение этой методики к прогнозированию стоимости медицинского страхования. В этой работе исследуется сеть, основанная на стохастической модели, имеющая многоуровневую архитектуру прямой связи со случайными связями между модулями и частотной характеристикой с препятствиями. Полученная байесовская методика, выведенная путем логического решения этой модели, базируется на фильтре Калмана. Полученный при этом алгоритм обучения обобщает так называемый одномерный метод Ньютона, популярный сейчас в литературе по нейронным сетям. В статье представлен численный метод изучения прогнозирования стоимости медицинского страхования в виде хаотических временных рядов с погрешностями и показана более высокая точность прогнозирования нового алгоритма по сравнению с существующими.

Ключевые слова: *стоимость медицинского страхования, статистическое прогнозирование, нейронные сети*

© А.О. СІГАЙОВ, А.В. ВОЛОВИК, 2017

A. SIGAYOV,
Professor, D.Sc. Igor Sikorsky Kyiv Polytechnic Institute
A. VOLOVYK,
Doctoral student Igor Sikorsky Kyiv Polytechnic Institute

Statistical models of neural networks for medical insurance cost forecasting

Neural networks have become a popular method of non-linear statistical forecasting therefore it seems important to apply this method to forecasting the cost of health insurance. In this paper the network based on a stochastic model that has direct connection tiered architecture with casual connections between modules and frequency characteristics with obstacles. The resulting Bayesian method, derived by logical solution for this model is based on the Kalman filter. Thus resulting learning algorithm generalizes the so-called one-dimensional Newton method that implements the algorithm is now popular in the literature on neural networks. The article presents numerical prediction method for studying relative cost of health insurance in the form of chaotic time series of errors and shows higher forecasting accuracy of the new algorithm compared to the existing ones.

Keywords: *medical insurance cost, statistical forecasting, neural networks*

Постановка проблеми. Ця стаття описує новий підхід до проблеми прогнозування нелінійних часових рядів. Іноді нелінійність в еволюції часу демонструє характерні риси, які можуть бути досить добре описані за допомогою існуючих параметризованих класів нелінійних моделей, таких як білінійні або граничні моделі. Частіше, нелінійність є результатом вкрай нерегулярної поведінки зі складними часовими залежностями, і визначення специфічної нелінійної моделі може бути важким.

Аналіз останніх досліджень та публікацій. Підхід, описаний у цій статті, ґрунтується на стратегії, запропонованій у нещодавній роботі з теорії нейронних систем. У цій теорії широкий клас моделей нейронної мережі представлений у вигляді пристроїв з обробки та розподілу інформації, основним завданням яких є складання моделей на базі потоків вхідних даних (див. Герц, Крог та Палмер [8]; Румельхарт та Макклеланд [17]).

Дотримуючись теорії Лапедеза та Фарбера [11], а також Джонса та ін. [9], ми застосовуємо цю стратегію до проблеми прогнозування нелінійних часових рядів. Ми моделюємо доступні для огляду часові ряди у вигляді систем обробки інформації, де інформація, у вигляді системи екстраполяторів, перетворюється та розподіляється через систему зв'язувань серед обмеженого числа модулів (нейронів), щоб одержати відповідь системи. Точніше, ми вводимо стохастичну модель нейронної мережі, засновану на багатошаровій архітектурі прямого зв'язку з випадковими зв'язками між модулями та реакціями зашумленої системи. Представляючи цю модель у вигляді просторового стану, ми надалі розробляємо статистичну методику прогнозування, виведену на основі фільтра Калмана.

Метою написання ***статті*** є презентація ефективної методики статистичного моделювання за допомогою нейронно-мережевих моделей. Мета досягається наступним чином. У першому підрозділі основного змісту описується архітектура прямого зв'язку прогнозуючої мережі Connectionist Normalized Local Spline (CNLS) (нормалізований локальний сплайн). У другому підрозділі представлене стохастичне динамічне узагальнення цієї мережі, виходячи з випадкової структури, як у зв'язках, так і в реакції системи. У третьому підрозділі представлений байєсівський алгоритм дослідження для стохастичної мережі. У четвертому підрозділі представлений метод оптимізації Ньютона, популярний сьогодні оновлюваний алгоритм, використовуваний у мережі CNLS як особливий випадок нашого байєсівського алгоритму. У п'ятому розділі використовується стохастична мережа для прогнозування зашумлених хаотично часових рядів і аналізується, як помилка прогнозування залежить від рівня шуму в системі та параметрів навчального алгоритму. Зокрема, продемонстровано, що можна суттєво покращити метод Ньютона. Нарешті, у шостому підрозділі показано, що стохастична мережа включає, як особливий випадок, деякі стандартні моделі часових рядів, у тому числі лінійні авторегресивні моделі, граничні моделі та моделі, що залежать від стану в авторегресивній форми.

Виклад основного матеріалу.

АРХІТЕКТУРА МЕРЕЖІ

Щоб створити модель прогнозної нейронної мережі, ми вибираємо багатошарову структуру з динамікою прямого зв'язку. Зокрема, ми ухвалюємо прогнозну мережу CNLS, введену Джонсом та ін. [9]. Вона складається з трьох упорядкованих шарів вузлів: кінцевого вхідного шару з p вузлами, одного схованого або проміжного шару з q вузлами і кінцевого вихідного шару. Хоча, в принципі, мережа могла б мати довільне число пристроїв висновку, у цій статті ми обмежимося випадком з одним пристроєм висновку.

Мережа має просту структуру прямого зв'язку. Кожен пристрій введення повідомляє про свій стан усім схованим модулям. Ці вузли обчислюють свій стан, обробляючи інформацію, отриману від пристроїв введення, а потім повідомляють про свій стан пристрою висновку. Пристрій висновку далі використовує цю інформацію, щоб обчислювати відповідь системи. Жодна інформація не проходить «назад» або між вузлами в тому самому шарі системи.

Стан пристрою введення і позначено як X_i, а чинність зв'язку між i та схованим вузлом j буде кодуватися сполучним параметром $v_{ji} = (\mu_{ji}, \sigma_{jt}^2)$. Стан схованого вузла j складається з двох компонентів, $\tilde{X}$ і H_j, обумовлених як $X = (X_1,...,X_p)$ і $v_j = (v_{ji}, ...,V_{jP})$ у такий спосіб:

$$\tilde{X}_j = X - \mu_j$$

і

$$H_j = u\left(X, v_j\right) \equiv u_j\left(X\right) = \lambda e^{\frac{-\left\|\tilde{X}_j\right\|^2}{2\lambda_j^2}}, \qquad (1)$$

де λ — нормувальний множник, який дорівнює

$$\frac{1}{\lambda} = \sum_s e^{\frac{-\left\|\tilde{X}_j\right\|^2}{2\lambda_S^2}},$$

$$s = 1, ..., q.$$

Таким чином, згідно з (1) компонент H стану кожного схзованого вузла виходить з локалізованого рецепторного поля в p–вимірному вхідному просторі, полі, зосередженому на μ_j, з розмірним співвідношенням до σ_j^2.

СТОХАСТИЧНА МОДЕЛЬ НЕЙРОННОЇ МЕРЕЖІ

У цьому розділі ми побудуємо динамічну стохастичну модель, яка узагальнює детерміновану структуру введення–висновок, надану в (1). Ця модель має два джерела довільності. По–перше, на реакцію Y будуть впливати випадкові збурення, відповідні до спостережуваного шуму. По–друге, чинність зв'язку між схованими модулями та пристроєм висновку порушується випадковим збурюванням параметрів β_j, відповідних до системних шумів.

Ми зараз описуємо стохастичну модель для мережі, використовуючи примітки розділу 2. Для ряду вхідних значень $X_1, X_2, ...$, визначимо ряд послідовних реакцій $Y_1, Y_2, ...$ за допомогою наступних трьох динамічних рівнянь:

$$Y_t = u\left(X_t\right) w\left(X_t\right) + a_i,$$

$$w\left(X_{t+1}\right) = w\left(X_t\right) + \left(\tilde{X}_{j,t+1} - \tilde{X}_{j,t}\right) \qquad (2)$$

$$\beta_{j,t+1} = \beta_j + \xi_{j,t+1}$$

де a_t і ξ_i означають спостережувані шуми та системні шуми, з $t = 1,2,...$ і $j = 1, ..., q$.

Шумові змінні відповідають наступним припущенням розподілу: $\{a_t\}$ і $\{\xi_{i,t+1}\}$ незалежні, О означає гаусіан для білих шумів; a_t має дисперсію R_t та ξ_t має коваріантність Q_t. В (3) параметри β_{jt} мають випадкові значення. Це також особливість залежних від стану моделей, описаних Прістлі (1988).

Щоб отримати методи оцінки і прогнозування параметра зі стохастичної моделі, описаної раніше, і зв'язати цю модель з іншими в літературі часових рядів, корисно представити це в наступному компактному вигляді:

$$Y_t = h\left(X_t\right) \vartheta\left(X_t\right) + a_t, \qquad (4)$$

і

$$\vartheta\left(X_{t+1}\right) = G\left(X_{t+1}\right) \vartheta\left(X_t\right) + b_{t+1}, \qquad (5)$$

де $h(X_t)$ є $q(p+1)$–вектором, вираженим як

$$h\left(X_t\right) = \left(u_1\left(X_t\right), \ldots, u_q\left(X_t\right); 0; \ldots; 0\right), \qquad (6)$$

$\vartheta\left(X_t\right)$ є $q(p+1)$–вектором, вираженим як

$$\vartheta\left(X_t\right) = \left(w_1\left(X_t\right), \ldots, w_q\left(X_t\right); \beta_{1t}; \ldots; \beta_{qt}\right), \qquad (7)$$

і матриця $G\left(X_{t+1}\right)$ визначається як

$$G\left(X_{t+1}\right) = \begin{pmatrix} I_q & \vdots & D\left(X_{t+1}\right) & 0 & 0 \\ & & 0 & D\left(X_{t+1}\right) & 0 \\ & & 0 & 0 & D\left(X_{t+1}\right) \\ \cdots & \cdots & \cdots & & \\ 0 & \vdots & vI_{pq} & & \end{pmatrix},$$

де I_q — $(q \times q)$ одинична матриця і

$$D\left(X_{t+1}\right) = \left(X_{1,t+1} - X_{1,t}, \ldots, X_{p,t+1} - X_{p,t}\right).$$

Шумовий компонент в (5) представлений є $q(p+1)$–вектором–стовпчиком b_{t+1}, де $b_t = \left(0, \ldots, 0; \xi_{1t}, \ldots, \xi_{qt}\right)$, $\xi_{jt} = \left(\xi_{1jt}, \ldots, \xi_{pjt}\right)$, $j = 1, \ldots, q$, $\xi_t = \left(\xi_{1t}, \ldots, \xi_{qt}\right)$.

У цій формулі ми бачимо, що стохастична модель мережі – це особливий випадок моделі в просторі станів, описаної, наприклад, Аокі [1]: (4)

є рівнянням спостереження та (5) є системним рівнянням, де $\theta(X_t)$ – вектор стану системи. Таким чином, узагальнення мережі та результати, описані в цій статті, можна розглядати в дусі роботи, представленої Де Джонгом [4], Гордоном та Смітом [7], Мейнхольдом та Сінгпурвалом [12], Тодіні [18], і Вестом, Харрісоном та Мігоном [20].

Модель також посилається на роботи, опубліковані в літературі з нейронних мереж. Локально настроєна мережа, представлена Муді та Даркеном [13], і мережа регуляризації Погіо та Жироссі [14] – це особливі випадки моделі в просторі станів, описаної тут, у яких функція з'єднання $w(X_t)$ – константа, w_t. Для цих мереж представлення моделі в просторі станів має вигляд простої формули

$$Y_t = h(X_t)\, w_t + a_t$$

$$\text{та } w_{t+1} = G_{t+1} W_t + \xi_{t+1}$$

де G_{t+1} – відома (q x q) матриця.

У результаті стохастичний алгоритм навчання, описаний у наступному розділі, також міг бути застосований до цих мереж (Полі та Джонса 1990, 1991).

СТОХАСТИЧНИЙ НАВЧАЛЬНИЙ АЛГОРИТМ

Ми зараз переходимо до проблеми прогнозування майбутнього значення ряду Y_{n+1}, виходячи з набору отриманих даних Y_t, $t = 1, ..., n$. Як введення в наш прогнозний алгоритм потім ми будемо використовувати вектор $X_{n+1} = (Y_n, Y_{n-1}, \ldots Y_{n-m},)$. Алгоритм буде демонструватися на наборі прикладів за формулою $(Y_t, X_t, t = 1, \ldots, n)$. Екстраполятор, який мінімізує MSE (середньоквадратичну погрішність) – це умовний очікуваний результат Y_{n+1} при X_{n+1} (див. Прістлі 1988).

Таким чином, наше завдання – визначити, що $Y^*_{n+1} = E\left[\{Y_{n+1} \mid X_{n+1}\}\right]$ на підставі моделі, описаної в (4) і (5). Для цього ми надамо кожному t гаусівський апріорний розподіл за вектором θ, який є незалежним від a_t і ξ_t. Потім, після одержання Y_t, стандартний байєсовський розрахунок дає апостеріорний розподіл для θ. Середнє значення і коваріантність цього розподілу, при $t = n$, мають форму (Полі та Джонс 1990):

$$\hat{\theta}_{n+1} = G_n \hat{\theta}_n + K_{n+1} e_{n+1}$$

$$\Sigma_{n+1} = \left(I - K_{n+1} h_{n+1}\right) P_{n+1}$$

при $K_{n+1} = P_n h'_n \left(h_n P_n h'_n + R_n\right)^{-1}$

де $P_n = G_n \Sigma_{n-1} G'_n + Q_n$ та $e_n = Y_n - h_n G_n \hat{\theta}_{n-1}$. Для полегшення демонстрації ми пропустили залежність θ, G і h від X_n. Визначення $\hat{\theta}_{n+1}$ в (8) досягається шляхом відновлення екстраполятора параметра $G_n \hat{\theta}_{n+1}$ за нового одержання Y_{n+1}. Екстраполятор $G_n \hat{\theta}_{n+1}$ фактично коригується, або поновлюється прогнозованою погрішністю e_{n+1}. Вплив цієї прогнозованої погрішності на $\hat{\theta}_{n+1}$ коректується формулою K_{n+1}, яка може розглядатися як швидкість навчання алгоритму. Також коваріаційна матриця Σ_{n+1} в (9) актуалізує відповідні значення прогнозованого розподілу, одержувані від P_{n+1}, за швидкості навчання K_{n+1} і елементів виміру h_{n+1}.

Випереджувальний екстраполятор Y_n потім приймає форму $Y_n = h_{n+1} G_n \hat{\theta}_n$, з прогнозованою дисперсією, що виражається як $K_{n+1} = h_{n+1} P_{n+1} h'_{n+1} + R_n$.

Оскільки рекурсії (8) і (9) представляють актуалізуючі рівняння методу фільтрації Калмана, що навчає стохастичний алгоритм, описаний вище, має звичайні властивості цього методу. Важливо відзначити, що оптимальність фільтра Калмана залежить від правильної специфікації моделі; тобто, що запропонована модель описує процес, який генерує дані (Берк [2]). Це звичайно не є метою моделі, наданої в розділі 3, мета якої – просто мотивувати представлення ефективного алгоритму прогнозування. Успіх моделі має оцінюватися з погляду успіху алгоритму прогнозування, до якого він веде.

МЕТОД ОПТИМІЗАЦІЇ НЬЮТОНА

Метод оптимізації Ньютона для оцінки параметра і прогнозування в детермінованих багаторівневих мережах є на даний момент одним із найпопулярніших методів у літературі з нейронних мереж (Герц, Крог, і Палмер [8]). Щоб визначити зв'язок цього методу з методом фільтрації Калмана, розглянутим у цій статті, ми зараз розглянемо часову еволюцію, змодельовану за відсутності шуму.

Використовуючи ту ж мережну архітектуру, описану в (2), ми розглянемо наступну модель у просторі станів:

$$Y_{t+1} = u(X_t)\, w(X_t)$$

$$\text{і} \quad w(X_{t+1}) = G_{t+1} w(X_{t+1}),$$

де $u(X_t) = \left(u_1(X_t), \ldots, u_q(X_t)\right)$ – перші q елементів $h(X_t)$ в (6), $w(X_t) = \left(w_1(X_t), \ldots, w_q(X_t)\right)$ – перші q елементи вектора $\theta(X_t)$ в (7), і $G_{t+1} = I_q$ – одинична матриця. Потім ми приймемо гаусівський апріорний розподіл для с із середнім $\hat{w}_1$ та одиничною матрицею коваріантності.

Оновлена формула для $w(X_t)$, отримана за рекурсивним алгоритмом, описаним у попередньому розділі, для $t = n$, дає нам

$$w(X_{n+1}) = w(X_n) + K_{n+1}e_{n+1} \qquad (10)$$

при

$$K_{n+1} = \left(u_1(X_{n+1}) \Big/ \left[\Sigma_j u_j^2(X_{n+1}) \right], \ldots, u_q(X_{n+1}) \Big/ \left[\Sigma_j u_j^2(X_{n+1}) \right] \right)$$

Ця формула ідентична рекурсивній оновлюючій формулі, отриманій Джонсом та ін. [9], який використовував одномірний метод оптимізації Ньютона. Таким чином, для негучних даних і одиничної матриці переходів метод оптимізації Ньютона виводить правило відновлення, яке є особливим випадком стохастичного навчального алгоритму, отриманого раніше для стохастичної моделі мережі. Для порівняння з методом алгоритму зворотного поширення, див. Вайт [21].

ВИКОРИСТАННЯ СТОХАСТИЧНОЇ МЕРЕЖІ ДЛЯ ПРОГНОЗУВАННЯ НЕЛІНІЙНИХ ЧАСОВИХ РЯДІВ

З метою ілюстрації ми представляємо результати чисельного аналізу, у якому використовувалася стохастична нейронна мережа для прогнозування нелінійних часових рядів. У цьому прикладі ми аналізуємо хаотичні часові ряди, значення яких визначаються за логістичною схемою. Складна проблема прогнозування хаотичних часових рядів нещодавно досліджувалася теоретиками динамічних систем, які ввели великий і цікавий клас методологій (Касдаглі [3]; Фармер та Сидорович [5]; Гібсон та ін. [6]; див. також Лапедес та Фарбер [11] для прогнозування за допомогою детерміністичної мережі).

Набір даних, за якими ми прагнемо дати прогноз, визначається за логістичною схемою $Y_{t+1} = 4Y_t(1 - Y_t)$, забрудненою шумами спостереження $a_t = N(0,1)$.

Для прогнозування цього ряду ми використовуємо прогнозуючу модель нейронної мережі, представлену в (4) і (5), з одним пристроєм введення і п'ятьма рівномірно розташованими схованими вузлами. Функція активації в схованих вузлах прийнята у вигляді нормалізованого гаусіана, тоді $u_j(X) = \lambda \exp\left\{ -\left\| \tilde{X}_j \right\|^2 \Big/ 2\lambda_j^2 \right\}$ як в (1), при $1 \big/ \sigma_j^2 = 4$.

Ми порівняли роботу двох різних оновлюючих алгоритмів: специфічної версії навчального стохастичного алгоритму (LSA) розділу 4, описаної в попередньому параграфі, і одномірного методу Ньютона (1DN). Прогнози, вироблені цими двома алгоритмами, були порівняні за допомогою нормалізованої помилки прогнозування

$$\varepsilon = E_\Omega\left(Y_{n+\tau} - Y_{n+\tau}^* \right)^{1/2} \Big/ \left\{ E_\Omega\left(Y_{n+\tau} - E_\Omega\left(Y_{n+\tau}^* \right) \right)^{1/2} \right\},$$

де $Y_{n+\tau}^*$ є прогнозними значеннями, отриманими шляхом навчання мережі за допомогою навчальної методики, $E_\Omega(Y_{n+\tau})$ є середнім значенням даних, τ – час випередження прогнозу, і E_Ω означає оператор усереднення для прогнозованих даних.

Висновки

1. Моделі нейронної мережі для прогнозування, розглянуті в цій статті, включають, у якості особливих випадків, такі знайомі часові ряди, як: лінійні авторегресивні моделі, граничні моделі та моделі, що залежать від стану модельованої системи в авторегресивній формі. Усі ці моделі можуть застосовуватися з використанням базової архітектури нейронної мережі, як було представлено тут, але без схованих вузлів, так званої простої персептронної архітектури.

Щоб переконатися в цьому, див. стохастичну мережу, представлену в (4), з використанням минулих значень рядів як вхідних даних; тобто, $X_t = (Y_{t-1}, Y_{t-1}, \ldots Y_{t-p})$. Якщо функції зв'язку є константою і відсутні сховані шари, то стохастична мережа зводиться до $Y_t = X_t w + a_t$, тобто до лінійної авторегресивної моделі порядку p.

А тепер припустимо, що функції зв'язку приймають вигляд

$w(X_t) = w_t^{(j)}$, якщо $X_t \in C^{(j)}, \quad j = 1, 2, \ldots, q,$

де $w_t^{(j)}$ – константи, а $C^{(j)}$ – задані області в R^p. Припустимо також, що немає жодного схованого шару. Тоді модель стохастичної нейронної мережі зводиться до граничної моделі (Полі 1987; Тонг 1980).

І нарешті, якщо ми тільки припустимо відсутність схованих вузлів і будемо підтримувати залежність функцій зв'язки від вхідних значень, як у (4), то модель буде відповідати моделі, що залежить від стану модельованої системи з авторегресивною схемою (Прістлі [16]).

У літературі з нейронних мереж є численні підтвердження того, що, на відміну від простих персептронів, сховані вузли можуть суттєво поліпшувати роботу мережі з цілого ряду завдань. Тому цілком доцільно досліджувати, чи можуть варіанти представленої тут загальної стохастичної мережі зі схованими вузлами запропонувати аналогічні переваги перед моделями з часовими рядами, згаданими в попередніх параграфах.

2. У нашій стохастичній моделі параметри $\left(\mu_j, \sigma_j^2\right)$, що визначають локальні рецепторні поля, через які сховані вузли реагують на введені дані, незмінні. Дуже бажано було б також змоделювати ці параметри як випадкові таким чином, щоб вони (і локальні поля, які вони визначають) могли бути адаптивно скореговані залежно від даних, що вводяться. Таке розширення моделі потребувало б нелінійної версії фільтра Калмана, щоб здійснювати відновлення параметрів і прогнозу. На даний ведеться робота над таким розширенням даної моделі.

Список використаних джерел

1. Aoki M. State–Space Modeling of Time Series. – Berlin: Springer–Verlag. – 1987.
2. Berk R. H. Limiting Behavior of Posterior Distribution When the Model is Incorrect. – The Annals of Mathematical Statistics. – Vol. 37. – 1966. – P. 51–58.
3. Casdagli M.L. Nonlinear Prediction of Chaotic Time Series. – Physica D. – Vol. 35. – 1989. – P. 335–356.
4. De Jong P. The Diffuse Kalman Filter. – The Annals of Statistics. – Vol. 19. – 1991. – P. 1073–1083.
5. Farmer J. D., Sidorowich J. J. Predicting chaotic time series. – Physics Review Letters. – Vol. 59. – 1987. – P. 845–847.
6. Gibson J. F., Farmer J. D., Casdagli M., Eubank S. An Analytic Approach to State Space Reconstruction. – Physica D. – Vol. 57. – 1992. – P. 1–30.
7. Gordon K., Smith A.F. Modeling and Monitoring Discontinuous Changes in Time Series // Bayesian Analysis of Time Series and Dynamic Models, ed. J. C. Spall. – New York: Marcel Dekker. – 1988.
8. Hertz J., Krogh A., Palmer R. Introduction to the Theory of Neural Computation. – Santa Fe Institute Studies in the Science of Complexity, Amsterdam: Addison–Wesley. – 1991.
9. Jones R. D, Lee Y.C., Barnes C W., Flake G.W., Lee K., Lewis P.S., Qian S. Function Approximation and Time Series Prediction With Neural Networks. –Working Paper LA–UR 9021. – Los Alamos, NM: Los Alamos National Laboratory. – 1989.
10. Jones R.D., Lee Y.C. Nonlinear Adaptive Computation. – Los Alamos, NM: Los Alamos National Laboratory. – 1991.
11. Lapedes A. S., Farber R.M. Nonlinear Signal Processing Using Neural Networks: Prediction and System Modeling. – Working Parer LA–UR 872662. – Los Alamos, NM: Los Alamos National Laboratory. – 1987.
12. Meinhold H. J., Singpurwalla N. D. Robustification of Kalman Filter Models. – Journal of the American Statistical Association. – Vol. 84. – 1989. – P. 479–486.
13. Moody J., Darken C.J. Fast Learning in Networks of Locally Tuned Processing Units. – Neural Computation. – Vol. 1. – 1989. – P. 281–294.
14. Poggio T., Girosi F. Regularization Algorithms for Learning That are Equivalent to Multilayer Networks. – Science. – Vol. – 247. – 1990. – P. 978–982.
15. Poli I., Jones R.D. A Predictive Algorithm for Noisy Feed–forward Neural Nets. – Technical Report. Bologna: Dip. di Statistica P. Fortunati. – 1990.
16. Priestley M. B. State–Dependent Models: A General Approach to Nonlinear Time Series Analysis. – Journal of Time Series Analysis. – Vol. 1. – 1980. –P. 47–71.
17. Rumelhart D.E., McClelland J.L. Parallel Distributed Processing: Explorations in the Microstructures of Cognition. – Cambridge, MA: MIT Press. – 1986.
18. Todini E. Mutually Interactive State Parameter (MISP) Estimation // Proc. AGU Chapman Conf. Applied Kalman Filtering Techniques – 1988. – P. 58–72.
19. Tong H., Lim K.S. Threshold Autoregression, Limit Cycles, and Cyclical Data. – Journal of the Royal Statistical Society. – Ser. B. – Vol. 42. – 1980. P. 245–292.
20. West M., Harrison P.J., Migon H. S. Dynamic Generalized Linear Models and Bayesian Forecasting. – Journal of the American Statistical Association. – Vol. 80. 1985. – P. 73–97.
21. White H. Some Asymptotic Results for Learning in Single Hidden–Layer Feedforward Network Models. – Journal of the American Statistical Association. – Vol. 84. – 1989. – P. 1003–1013.