

УДК 629.735

Владов С. І., к.т.н.

Шмельов Ю. М., к.т.н.

Кременчуцький льотний коледж Харківського національного університету внутрішніх справ, м. Кременчук, Україна

НАВЧАННЯ НЕЙРОННОЇ МЕРЕЖІ МЕТОДОМ РОЗШИРЕНОГО ФІЛЬТРА КАЛМАНА ДЛЯ ДІАГНОСТИКИ І ПАРИРУВАННЯ ВІДМОВ ДАТЧИКІВ АВІАЦІЙНОГО ДВИГУНА ТВ3-117

Діагностика і парирування відмов датчиків витрати палива авіаційного двигуна ТВ3-117 здійснюються на основі застосування алгоритмів Калман-фільтрації з вбудованою логікою виявлення та локалізації відмови вимірювального каналу. Можливість застосування фільтрів Калмана в LABEM розглянутого класу доведена на основі статистичної обробки даних натурних випробувань двигуна [1].

Фільтр Калмана є ефективним рекурсивним фільтром, який за низкою зашумлених і неповних вимірювань дозволяє оцінити внутрішній стан динамічної системи і застосовується в найширшому спектрі технічних пристроїв, від спідометрів автомобіля до радіоприймачів і радарів. Типовою задачею для фільтра Калмана є оцінка минулих, поточних або майбутніх значень положення, швидкості або прискорення деякої динамічної системи, для якої відома її лінійна або миттєва лінеаризована модель.

Навчання нейронних мереж є досить несподіваним додатком для теорії фільтрації Калмана і, разом з тим, дуже ефективним: з одного боку, якість такого навчання знаходиться на рівні кращих пакетних алгоритмів другого порядку, таких як метод Левенберга-Марквардта або квазіньютонівських методів [2], а з іншого боку – навчання проводиться в режимі онлайн, що є актуальним у разі вибірок значного обсягу і задач управління. Існують різні модифікації цього методу навчання, так чи інакше підвищують його ефективність: багатопоточне навчання (multistream learning) [3], що дозволяє мінімізувати ризик потрапляння в локальний мінімум; пакетна форма навчання (batch form) [4], що дозволяє обробляти кілька останніх вимірювань за один раз; незв'язний розширений фільтр Калмана (Decoupled Extended Kalman Filter) [3], який застосовується для економії обчислювальних ресурсів. Також нещодавно були запропоновані нові реалізації фільтра Калмана, що мають велику обчислювальну точність і внаслідок цього забезпечують поліпшену збіжність: фільтр Калмана на основі квадратного кореня [5, с. 960] і кубатурний фільтр Калмана [4, с. 787]. У цій роботі описується найбільш проста і технологічна реалізація навчання методом розширеного фільтра Калмана – «глобальний розширений фільтр Калмана» (Global Extended Kalman Filter, GEKF).

В усіх цих випадках навчання нейронної мережі розглядається як задача оцінки істинного стану деякої невідомої «ідеальної» нейронної мережі, що забезпечує нульову неузгодженість. Під станами в даному випадку

приймаються значення ваг нейронної мережі $w(k)$, а під неузгодженістю – поточна помилка навчання $e(k)$.

Цей динамічний процес навчання може бути описаний парою рівнянь у просторі станів. Рівняння стану є моделлю процесу, що, у свою чергу, є еволюцією вектора ваг під впливом випадкового процесу $\xi(k)$, який вважається білим шумом з нульовим математичним сподіванням і відомою діагональною коваріаційною матрицею Q :

$$w(k+1) = w(k) + \xi(k); \quad (1)$$

Рівняння виходу становить лінеаризовану модель нейронної мережі $y = g\left(\sum_j w_j^{(2)} f\left(\sum_i w_{ji}^{(1)} x_i\right)\right)$, де $w^{(1)}$ – ваги нейронів прихованого шару; $f(\cdot)$ – активаційні функції нейронів прихованого шару; $w^{(2)}$ – ваги нейронів вихідного шару; $g(\cdot)$ – активаційні функції нейронів вихідного шару на такті k , зашумлений випадковим процесом $\zeta(k)$, який вважається білим шумом з нульовим математичним сподіванням і відомою діагональною коваріаційною матрицею R :

$$h(k) = \frac{\partial y(w(k), v(k), x(k))}{\partial w} + \zeta(k); \quad (2)$$

де $w(k)$ – ваги нейронної мережі; $v(k)$ – постсинаптичні потенціали нейронів; $x(k)$ – вхідні значення мережі.

Обчислення миттєвих значень похідних $\frac{\partial y}{\partial w}$ здійснюється методом зворотного поширення. Неузгодженість $e(k)$ розраховується за математичним виразом:

$$e(k) = t(k) - y(k); \quad (3)$$

де $t(k)$ – цільове значення для нейронної мережі; $y(k)$ – реальний вихід нейронної мережі.

Перед навчанням нейронної мережі відбувається етап ініціалізації. Задаються коваріаційні матриці шуму вимірювання $R = \eta I$ і динамічного шуму навчання $Q = \mu I$, розмір матриць $L \times L$ і $N \times N$ відповідно, де L – кількість вихідних нейронів, N – кількість вагових коефіцієнтів нейронної мережі. Коефіцієнт η має сенс швидкості навчання, у даній задачі $\eta = 0,001$, коефіцієнт μ визначає шум вимірювання, в даній задачі приймається $\mu = 10^{-4}$. Також задаються на етапі ініціалізації одинична матриця коваріації P розміру $N \times N$ і нульова матриця вимірювань H розміру $L \times N$.

Етап навчання проводиться в режимі онлайн, корекція ваг нейронної мережі послідовно виконується для кожного прикладу навчальної вибірки. На такті k виконуються наступні дії.

Етап навчання здійснюється в режимі онлайн, корекція ваг нейронної мережі послідовно виконується для кожного прикладу навчальної вибірки. На такті k виконуються такі дії:

1) Обчислюється нове значення виходу нейронної мережі $y(k)$, виконується «прямий прохід» нейронної мережі.

2) Виконується «зворотний прохід» нейронної мережі: методом зворотного поширення обчислюються похідні $\frac{\partial y}{\partial w_i}$, $i = \overline{1, N}$.

3) Це здійснюється за допомогою такої ж техніки, що і в методі зворотного поширення помилки, але локальні градієнти для вихідних нейронів встановлюються рівними не поточній помилці $e(k)$, а константі 1, що при всіх тих же обчисленнях забезпечує отримання значень якобіанів виходів нейронної

мережі $\frac{\partial y}{\partial w_i}$ замість градієнтів $\frac{\partial (e(k))^2}{\partial w_i}$, оскільки $\frac{\partial (e(k))^2}{\partial w_i} = 2e(k) \frac{\partial y}{\partial w}$.

Формується матриця спостережень $H(k)$:

$$H(k) = \left[\frac{\partial y}{\partial w_1}, \frac{\partial y}{\partial w_2}, \dots, \frac{\partial y}{\partial w_N} \right]^T. \quad (4)$$

4) Визначається поточна помилка роботи мережі $e(k)$ за формулою (4), формується матриця відхилень $E(k)$ розмірності $1 \times L$:

$$E(k) = [e(k)]. \quad (5)$$

5) Обчислюються нові значення ваг нейронної мережі $w(k+1)$ і матриці кореляції $P(k+1)$ за формулами:

$$K(k) = P(k)H(k)^T \left[H(k)P(k)H(k)^T + R \right]^{-1}; \quad (6)$$

$$P(k+1) = P(k) - K(k)H(k)P(k) + Q; \quad (7)$$

$$w(k+1) = w(k) + K(k)e(k); \quad (8)$$

де $K(k)$ – матриця підсилення Калмана, її розмірність $N \times L$. Дії 1–4 виконуються для всіх елементів навчальної вибірки.

Кореляційна матриця P , що оновлюється на кожному такті, містить інформацію другого порядку про поверхні помилок, що забезпечує методу розширеного фільтра Калмана перевагу порівняно з методами навчання першого порядку, такими як градієнтний спуск і його модифікації.

Список літератури

4. VladovS., ShmelovYu., KotliarovK., HrybanovaS., HusarovaO., Derevyanko I., ChyzhovaL. Onboard parameter identification method of the TV3-117 aircraft engine of the neural network technologies. *Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*, issue 5/2019 (118), 2019, pp. 90–96.

5. Alessandri A., Cuneo M., Pagnan S. On the convergence of EKF-based parameters optimization for Neural Networks, *Proceedings of the 42nd IEEE Conference on Decision and Control*, vol. 6, 2001, pp. 6181–6186.

6. Feldkamp L.A. and Puskorius G.V. Training controllers for robustness: multi-stream DEKF, *Proceedings of International Conference on Neural Networks*, 27 June 27 – July 02, 1994, vol. 4, pp. 2377–2382.

7. Haykin S. *Neural Networks and Learning Machines* (Third Edition), New York, Prentice Hall, 2009, 936 p.

8. Haykin S., *Neural networks: full course*, Moscow, Williams, 2006, 1104 p.

УДК 629.735

Владов С. І., к.т.н.

Шмельов Ю. М., к.т.н.

Матусєв А. О., курсант

Плужник З. І., курсант

Кременчуцький льотний коледж Харківського національного університету внутрішніх справ, м. Кременчук, Україна

ВИБІР ІНФОРМАЦІЙНОГО КРИТЕРІЯ ДЛЯ КОНТРОЛЮ І ДІАГНОСТИКИ ТЕХНІЧНОГО СТАНУ АВІАЦІЙНОГО ДВИГУНА ТВЗ-117

Відомо, що технічний стан авіаційного двигуна ТВЗ-117 визначається кількістю інформації, одержуваної системою контролю в залежності від смислового призначення цієї інформації.

З розглянутих положень, наукових підходів в теорії інформації, а також видів ентропії, що характеризують основну властивість невизначеності складних систем, до яких належить авіаційний двигун ТВЗ-117, можна виділити три групи ентропії і віднести їх до наступних категорій: термодинамічна, статистична та інформаційна ентропії (табл. 1).

Таблиця 1 – Порівняльні характеристики різних ентропій

№ з/п	Вид ентропії		Придатність для розв'язку поставлених задач
	Найменування	Вираз	
1	2	3	4
1	Класична термодинамічна ентропія Р. Клазіуса	$S = \frac{\Delta Q}{T}$	Можливо використовувати тільки для оцінки стану речовини в різних температурних умовах
2	Ентропія Л. Больцмана	$S_b = k \ln P$	Оцінюється потенційна мінливість об'єкта (системи)