УДК 621.391.83

В.С. Михальчан
Одеська національна академія зв’язку ім.О.С. Попова,
кафедра документального електрозв’язку

АЛГОРИТМ НАВЧАННЯ НЕРЕКУРСИВНОГО ФІЛЬТРА

Запропоновано ітераційний алгоритм навчання адаптивного нерекурсивного фільтра для зменшення міжсимвольної інтерференції шляхом мінімізації довжини вектора градієнта, що має більш високу швидкість збіжності порівняно з методом якнайшвидшого спуску.

The iteration algorithm of adaptive nonrecursive filter training for reduction of an intersymbol interference is offered by minimization of gradient vector depth having more high rate of convergence in comparison with a steepest descent method.

Адаптивні нерекурсивні фільтри [1, 2] призначені для зменшення міжсимвольної інтерференції, вони працюють у двох режимах: у режимі навчання за тестовими сигналами і в стохастичному режимі адаптації за інформаційними сигналами. Час, витрачений на режим навчання, для передачі даних загублено, а застосовувані алгоритми на основі методу якнайшвидшого спуску мають дуже повільну швидкість цього процесу [1, 2], тобто потрібен тривалий час для його завершення.

Розглянемо алгоритм навчання адаптивних нерекурсивних фільтрів, заснований на методі мінімальних нев’язок [3], що дозволяє істотно скоротити час навчання.

Введемо такі умови:

1. Імпульсний відгук каналу має кінцеву тривалість імпульсного відгуку;
2. Відліки сигналу x_i на виході каналу (вході фільтра) випадкові взаємно некорельовані величини;
3. Відліки шуму мають нормальний розподіл і взаємно некорельовані.

Задача навчання адаптивного нерекурсивного фільтра полягає у визначенні значень регульованих параметрів φ_i фільтра шляхом вирішення кінцевого дискретного рівняння Віннера-Хопфа в згортках:

$$\sum_i x_{k-i}\varphi_i = d_k \text{, при } i = -N,...,N\text{; } k = -2N,...,2N\text{.} \qquad (1)$$

Останнє рівняння являє собою несумісну систему рівнянь, тому що матриця A, складена з відліків x_i тестового сигналу на вході фільтра, розміром $(2N+1)\times(2N+1)$; φ – вектор невідомих (значення регульованих вагових параметрів фільтра), розміром $(2N+1)$; d – вектор вільних членів (необхідні відліки на виході фільтра, які потрібно отримати в результаті навчання), розміром $(4N+1)$, тобто число рівнянь більше числа невідомих.

Для розв’язання несумісної системи (1) скористаємося методом найменших квадратів, склавши функціонал, що являє собою середньоквадратичну помилку (СКП):

$$\varepsilon = E\left[\varepsilon_N(\varphi)\right] = E\left[\sum_k (a_k - d_k)^2\right] = E\left[\sum_k r_k^2\right], \qquad (2)$$

де $r_k - a_k - d_k$ – відліки сигналу помилки, тобто відхилення відліків сигналу a_k на виході фільтра від відліків необхідного сигналу d_k.

При цьому

$$a_k = \sum_i x_{k-i}\varphi_i \,. \qquad (3)$$

Підставляючи (3) в (2), отримаємо:

$$\varepsilon(\varphi) = \sum_i \varphi_i \sum_j \varphi_j \sum_k x_{k-i}x_{k-j} - 2\sum_i \varphi_i \sum_k x_{k-i}d_k + E\left[\sum_k d_k^2\right],$$
$$i, j = -N,...,N\,;\; k = -2N,...,2N\,. \qquad (4)$$

Для визначення оптимальних значень регульованих вагових параметрів φ_i фільтра, при яких досягається мінімум СКП, застосуємо диференціальне обчислення для цієї мінімізаційної задачі, поклавши похідні від $\varepsilon(\varphi)$ по φ_i рівними нулю. Для $i, j = -N,...,N$ отримаємо компоненти вектора градієнта функціонала $\varepsilon(\varphi)$:

$$\frac{\partial\varepsilon(\varphi)}{\partial\varphi_i} = g_i = 2\sum_k x_{k-j}\sum_i x_{k-i}\varphi_i - 2\sum_k x_{k-i}d_k = 0\,.$$

Звідси отримаємо систему «нормальних» рівнянь, в якій число рівнянь дорівнює числу невідомих:

$$\sum_k x_{k-j}\sum_i x_{k-i}\varphi_i = \sum_k x_{k-i}d_k\,,$$
$$i, j = -N,...,N\,;\; k = -2N,...,2N\,. \qquad (5)$$

З рівнянь (5) можна отримати значення похідних (компоненти вектора градієнта) для довільних значень φ_i:

$$g_i = \sum_k x_{k-j}\sum_i x_{k-i}\varphi_i - \sum_k x_{k-i}d_k = \sum_k x_{k-i}r_k\,. \qquad (6)$$

Введемо до розгляду функціонал, складений з компонентів вектора-градієнта:

$$F(\varphi) = \sum_i g_i^2 = \sum_i\left[\sum_k x_{k-i}r_k\right]^2. \qquad (7)$$

Задача оптимізації для визначення оптимальних значень φ_i^* фільтра буде полягати в мінімізації цільової функції (7). Можна показати, що мінімум функціонала (7), тобто рішення системи (5), перетворює в мінімум і цільову функцію (4).

Для вирішення поставленої оптимізаційної задачі використовуємо ітераційний градієнтний метод, в якому обчислення значень регульованих вагових параметрів φ_i фільтра при зміні від $\varphi_i(n)$ до $\varphi_i(n+1)$ описується таким співвідношенням:

$$\varphi_i(n+1) = \varphi(n) - \alpha(n)\sum_k x_{k-i}r_k(n) = \varphi_i(n) - \alpha(n)g_i(n), \qquad (8)$$

де $\alpha(n)$ – скалярний параметр, поки довільний; $g(n)$ – напрямок спуска, протилежний напрямкові вектора градієнта; n – поточний номер ітерації.

Потім підставляючи (8) у (3), отримані значення відліків сигналу $a_k(n+1)$ на виході фільтра в (7), отримаємо значення цільової функції на $(n+1)$-ї ітерації:

$$F\left[\varphi(n+1)\right] = \sum_i g_i^2(n+1) = \sum_i g_i^2(n) - 2\alpha(n)\sum_i g_i(n)h_i(n) + \alpha^2(n)\sum_i h_i^2(n)\,. \qquad (9)$$

Останній вираз продиференціюємо за параметром $\alpha(n)$ і прирівнявши до нуля результат, остаточно маємо:

$$\alpha(n)=\frac{\sum_i g_i(n)h_i(n)}{\sum_i h_i^2(n)}, \tag{10}$$

де $h_i(n)=\sum_k x_{k-j}\sum_i x_{k-i}g_i(n)=\sum_k x_{k-i}f_k$, $i,j=-N,...,N$; $k=-2N,...,2N$.

Підставляючи (10) у (8), отримуємо структуру ітераційного алгоритму навчання адаптивного нерекурсивного фільтра. При цьому на кожній ітерації мінімізується довжина вектора градієнта.

Як видно з формули (10), розрахункова формула для $\alpha(n)$ отримана в такій формі, що матриця $A^T A$, фактично не обчислюється (елементи цієї матриці можуть обчислюватися за формулою $\sum_k x_{k-i}x_{k-j}$, T – знак транспонування). Всі значення, необхідні для розрахунку $\alpha(n)$, можна безпосередньо отримати у вигляді додаткової згортки $\left(\sum_i x_{k-i}g_i(n)\right)$, і як взаємну кореляцію між відліками сигналу x_i на відводах лінії затримки і помилки $r_k\left(\sum_k x_{k-i}r_k\right)$, а також як взаємну кореляцію між відліками сигналу x_i та допоміжними відліками $f_k\left(\sum_k x_{k-i}f_k\right)$.

Доведемо збіжність запропонованого процесу до оптимального значення φ^*.

Для цього підставимо (10) у (9) і після деяких перетворень отримаємо:

$$\frac{F\left[\varphi(n+1)\right]}{F\left[\varphi(n)\right]}=1-\frac{\left[\sum_i g_i(n)h_i(n)\right]^2}{\left[\sum_i h_i^2\right]\left[\sum_i g_i^2\right]}. \tag{11}$$

Оцінимо знизу від'ємник у правій частині останньої рівності.

Нехай $m=\lambda_1<\lambda_2<...<\lambda_{2N+1}=M$ – власні значення матриці $B=A^T A$, що належать $e_1,e_2,...,e_{2N+1}$, або власні вектори, ортогональні один одному і нормовані так, що $(e_i,e_i)=1$ при $i=1,2,...,2N+1$. Оскільки кореляційна матриця B позитивно визначена, то всі $\lambda_i>0$. Нехай далі вектор $g(n)$ лежить у площині, що натягнута на власні вектори e_1 і e_{2N+1} матриці B. При цьому вектор $g(n)$ розташований під кутом 45^0 до зазначених власних векторів e_1 і e_{2N+1}. Тоді вектор градієнт

$$g(n)=c_1e_1+c_2e_2+...+c_{2N+1}e_{2N+1}=ce_1+ce_{2N+1},$$

і допоміжний вектор

$$h(n)=c_1\lambda_1e_1+c_2\lambda_2e_2+...+c_{2N+1}\lambda_{2N+1}e_{2N+1}=c\lambda_1e_1+c\lambda_{2N+1}e_{2N+1}.$$

Отже,

$$\sum_i g_i^2(n) = c^2 + c^2 = 2c^2,$$

$$\sum_i \left[g_i(n) h_i(n) \right] = c^2 \left(\lambda_1 + \lambda_{2N+1}^2 \right) = c^2 (m + M),$$

$$\sum_i h_i^2(n) = c^2 \left(\lambda_1^2 + \lambda_{2N+1}^2 \right) = c^2 \left(m^2 + M^2 \right),$$

і тому співвідношення (11) набуде вигляду:

$$\frac{F\left[\varphi(n+1)\right]}{F\left[\varphi(n)\right]} \leq \frac{(M-m)^2}{2\left(M^2+m^2\right)} < 1.$$

Звідси

$$F\left[\varphi(n+1)\right] \leq \left[\frac{(M-m)^2}{2\left(M^2+m^2\right)} \right] F\left[\varphi(n)\right],$$

і очевидно, що

$$F\left[\varphi(n)\right] \leq \left[\frac{(M+m)^2}{2\left(M^2+m^2\right)} \right]^n F\left[\varphi(0)\right], \qquad (12)$$

де $\varphi(0)$ – вихідні (початкові) значення регульованих параметрів фільтра.

З виразу (12) видно, що значення цільової функції $F\left[\varphi(n)\right] \to 0$ при $n \to \infty$, тому що знаменник геометричної прогресії $q = (M-m)^2 / 2\left(M^2+m^2\right) < 1$, і тому $\varphi(n) \to \varphi^*$, тобто значення регульованих параметрів адаптивного фільтра встановлюються в оптимальне положення в процесі навчання.

Отже, застосування запропонованого алгоритму навчання дозволяє отримати швидкість збіжності до рішення, що характеризується знаменником геометричної прогресії q. В якості m й M взяті найменше і, відповідно, найбільше власне значення матриці $B = A^T A$. При цьому для навчання фільтра не потрібно знання власних чисел матриці B.

Порівняємо пропонований алгоритм з алгоритмами на основі методу якнайшвидшого спуску. Відомо [3], що швидкість убування значень функціонала (2) у методі якнайшвидшого спуску визначається членами геометричної прогресії зі знаменником

$$q_1 = \frac{(M-m)^2}{(M+m)^2}.$$

Очевидно, що

$$\frac{q}{q_1} = \frac{(M+m)^2}{2\left(M^2+m^2\right)} < 1,$$

Тобто швидкість збіжності алгоритму (10) вище в $2\left(M^2+m^2\right) / (M+m)^2$ раз, ніж швидкість збіжності алгоритму на основі методу якнайшвидшого спуску в оцінці швидкості убування функціонала.

Залежність цільової функції (7) від числа одиночних тестових сигналів навчання з управлінням параметрами фільтра алгоритмом навчання (10), наведена у вигляді зміни типових кривих навчання, тобто у вигляді зміни довжини вектора градієнта від числа ітерацій (криві 1, 2 на рис.1).

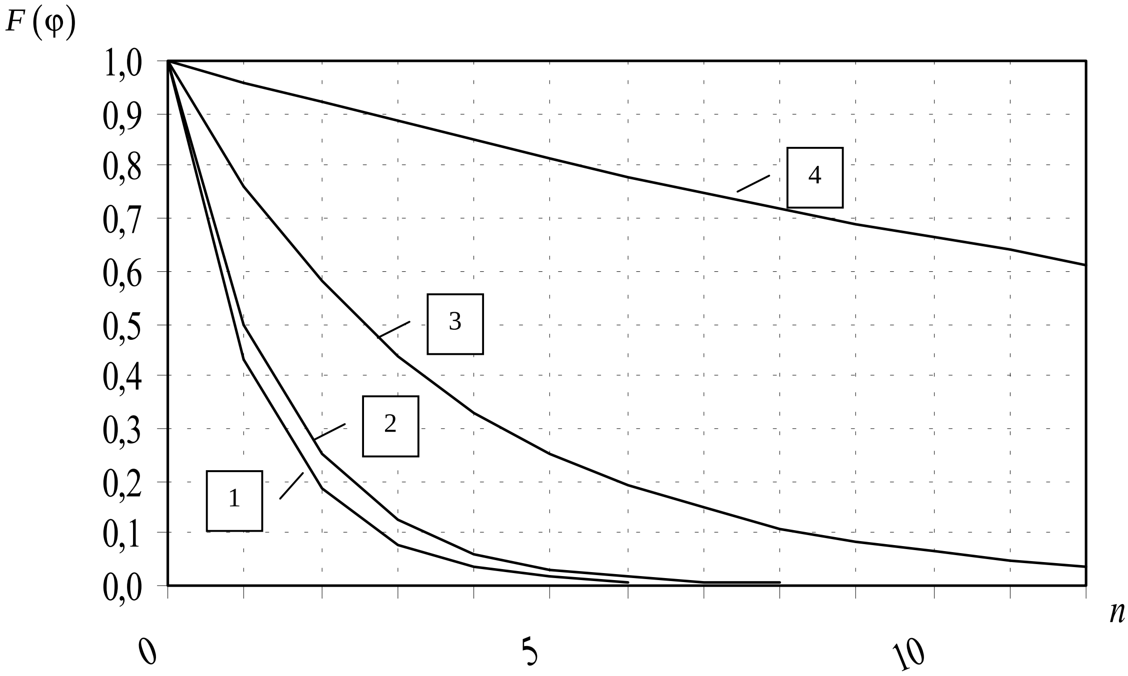

Рис.1. Навчальні криві – графіки залежності цільової функції від числа ітерацій

Для порівняння з алгоритмом якнайшвидшого спуску наведені криві навчання (криві 3 і 4 на рис.1) з використанням зазначеного алгоритму при мінімізації функціонала (2). Кривим 1 і 3 відповідає значення спектральної обумовленості $\rho = M/m = 15$ матриці B. При цьому $q = 0,43$ і $q_1 = 0,76$ відповідно. Кривим 2 і 4 відповідає значення спектральної обумовленості $\rho = 100$ матриці B. При цьому $q = 0,49$ і $q_1 = 0,96$.

Для ілюстрації ефективності управління регульованими ваговими параметрами φ фільтра пропонованим алгоритмом, розглянемо простий приклад.

Нехай задана функція $f(\varphi_1, \varphi_2) = \dfrac{1}{2}\varphi_1^2 + \dfrac{15}{2}\varphi_2^2$. Це позитивно визначена функція з точкою мінімуму при значеннях $\varphi^* = [0\ \ 0]^T$. Градієнт даної функції для довільних φ, $g = [\varphi_1\ 15\varphi_2]^T$, а матриця B:

$$B = \frac{\partial^2 f(\varphi)}{\partial \varphi^2} = \begin{bmatrix} 1 & 0 \\ 0 & 15 \end{bmatrix} \text{ для усіх } \varphi,$$

де $m = 1$, $M = 15$, $\rho = 15$.

Задамо початкове значення $\varphi(0) = [15\ \ 1]^T$ та організуємо навчання фільтра з двома регульованими ваговими параметрами запропонованим алгоритмом (10) і алгоритмом на основі методу якнайшвидшого спуска.

Послідовність значень $\varphi(n)$ при заданих $f(\varphi)$, $g(\varphi)$ і B, керованих алгоритмом на основі методу якнайшвидшого спуску, наведена на рис. 2. Послідовність значень $\varphi(n)$ при управлінні ваговими параметрами алгоритмом (10) ілюструє рис. 3.

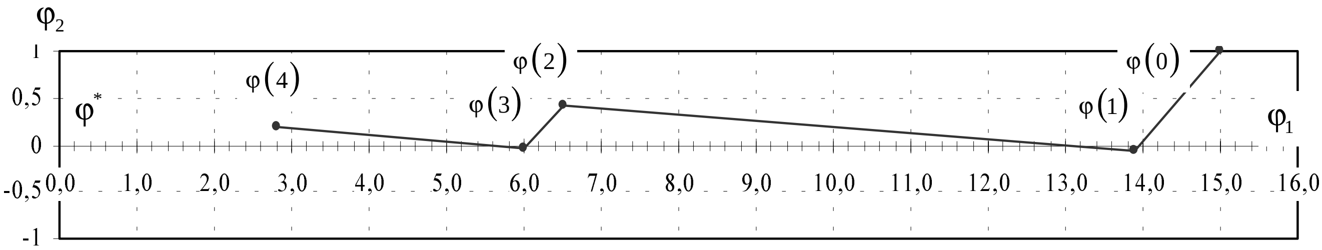

Рис.2 Послідовність значень $\varphi(n)$ *при управлінні алгоритмом якнайшвидшого спуска*

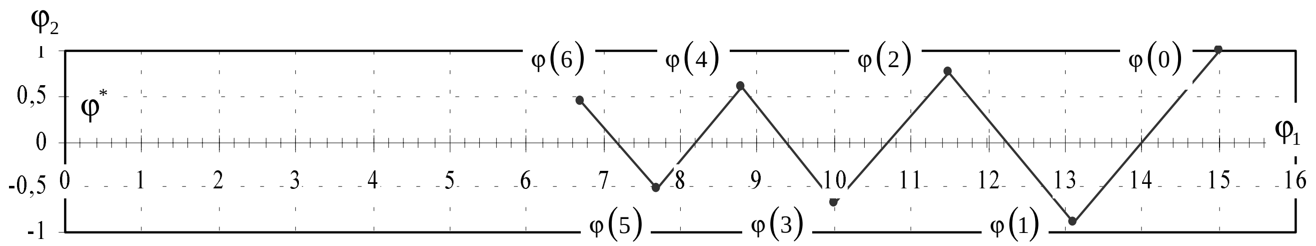

Рис.3. Послідовність значень $\varphi(n)$ *при управлінні алгоритмом мінімальних градієнтів*

Застосування запропонованого алгоритму навчання адаптивних нерекурсивних фільтрів у кілька разів скорочує час навчання порівняно з алгоритмом навчання на основі методу якнайшвидшого спуска.

1. Уидроу Б., Стирнз С. Адаптивная обработка сигналов: Пер. с англ. – М.: Радио и связь, 1989. – 440 с. 2. Михальчан В.С. Разработка и исследование алгоритмов и устройств настройки гармонических корректоров для высокоскоростных систем связи. Автореф. дис... канд. техн. наук. – Одесса, 1984. – 24 с. 3. Красносельский М.А., Крейн С.Г. Итерационный процесс с минимальными невязками. – В кн.: Математический сборник. Новая серия. – 1952 – Т. 31. – вып. 2. – С. 315-334.