Algoritmul proiecţiei afine (Affine Projection Algorithm – APA)

Este cel mai nou dintre algoritmii adaptivi care s-au impus (Ozeki şi Umeda, 1984) şi poate fi privit ca o extindere a algoritmului NLMS. Ca şi în cazul acestuia, reactualizarea coeficienţilor se face pornind de la minimizarea normei euclidiene a variaţiei

$$\delta \boldsymbol{w}(n+1) = \boldsymbol{w}(n+1) - \boldsymbol{w}(n)$$

dar cu condiţiile

$$d(n-k) = \boldsymbol{w}^{\boldsymbol{H}}(n+1)x(n-k), \qquad k = 0, \cdots, M-1$$

Semnificaţia acestora ar fi aceea că noii coeficienţi ar fi anulat eroarea la M momente anterioare $n, n-1, \cdots, n-M+1, \quad M < N$ (nu doar la momentul n, ca în cazul LMS). M poartă numele de *ordin al proiecţiei*. Vom defini deci o funcţie cost reală

$$J(n) = \|\boldsymbol{w}(n+1) - \boldsymbol{w}(n)\|^2 + \sum_{k=0}^{M-1} \mathrm{Re}\left\{\lambda_k^*\left[d(n-k) - \boldsymbol{w}^{\boldsymbol{H}}(n+1)x(n-k)\right]\right\}$$

Pentru a sistematiza scrierea vom defini următoarele structuri

- Vectorii semnalelor din linia de întârziere a filtrului, pentru cele M momente de timp

$$\boldsymbol{x}(n-k) = [x(n-k), x(n-k-1), \cdots, x(n-k-N+1)]^T, \;\; k = 0, \cdots, M-1$$

- Matricea având drept coloane aceşti vectori

$$\boldsymbol{A}^H(n) = [\boldsymbol{x}(n), \boldsymbol{x}(n-1), \cdots, \boldsymbol{x}(n-M+1)]$$

- Vectorul semnalelor dorite

$$\boldsymbol{d}^H(n) = [d(n), d(n-1), \cdots, d(n-M+1)]$$

- Vectorul erorilor

$$\boldsymbol{e}^H(n) = [e(n), e(n-1), \cdots, e(n-M+1)]$$

- Vectorul ieşirilor

$$\boldsymbol{y}^H(n) = [y(n), y(n-1), \cdots, y(n-M+1)] = \boldsymbol{A}(n)\boldsymbol{w}(n)$$

  Evident,

$$\boldsymbol{e}(n) = \boldsymbol{d}(n) - \boldsymbol{y}(n) = \boldsymbol{d}(n) - \boldsymbol{A}(n)\boldsymbol{w}(n)$$

- Vectorul multiplicatorilor Lagrange

$$\boldsymbol{\lambda}^H = [\lambda_0, \lambda_1, \cdots, \lambda_{M-1}]$$

  Cu aceste notaţii, setul de condiţii devine

$$\boldsymbol{d}(n) = \boldsymbol{A}(n)\boldsymbol{w}(n+1)$$

  iar funcţia cost

$$J(n) = \|\boldsymbol{w}(n+1) - \boldsymbol{w}(n)\|^2 + \mathrm{Re}\left\{\left(\boldsymbol{d}(n) - \boldsymbol{A}(n)\boldsymbol{w}(n+1)\right)^H \boldsymbol{\lambda}\right\}$$

Minimul funcţiei cost se obţine odată cu minimul cantităţii $\|\boldsymbol{w}(n+1)-\boldsymbol{w}(n)\|^2$, dacă este îndeplinit setul de condiţii $\boldsymbol{d}(n)=\boldsymbol{A}(n)\boldsymbol{w}(n+1)$. Va trebui deci să anulăm gradientul complex al funcţiei cost calculat în raport cu setul de coeficienţi $\boldsymbol{w}(n+1)$. Componentele acestuia sunt

$$\nabla_{\boldsymbol{w}^*(n+1)}\{\|\boldsymbol{w}(n+1)-\boldsymbol{w}(n)\|^2\}=\nabla_{\boldsymbol{w}^*(n+1)}\left\{\left(\boldsymbol{w}(n+1)-\boldsymbol{w}(n)\right)^{\mathrm{H}}\left(\boldsymbol{w}(n+1)-\boldsymbol{w}(n)\right)\right\}$$

$$\nabla_{\boldsymbol{w}^*(n+1)}\{\|\boldsymbol{w}(n+1)-\boldsymbol{w}(n)\|^2\}=\boldsymbol{w}(n+1)-\boldsymbol{w}(n)$$

deoarece

$$\nabla_{\boldsymbol{w}^*(n+1)}\{\boldsymbol{w}^H(n+1)\boldsymbol{w}(n+1)\}=\boldsymbol{w}(n+1)$$

iar

$$\nabla_{\boldsymbol{w}^*(n+1)}\{\boldsymbol{w}^H(n+1)\boldsymbol{w}(n)+\boldsymbol{w}^H(n)\boldsymbol{w}(n+1)\}=\boldsymbol{w}(n)$$

şi

$$\nabla_{\boldsymbol{w}^*(n+1)}\left\{\mathrm{Re}\left\{\left(\boldsymbol{d}(n)-\boldsymbol{A}(n)\boldsymbol{w}(n+1)\right)^H\boldsymbol{\lambda}\right\}\right\}$$

$$=\nabla_{\boldsymbol{w}^*(n+1)}\left\{\frac{1}{2}\left\{\left(\boldsymbol{d}(n)-\boldsymbol{A}(n)\boldsymbol{w}(n+1)\right)^H\boldsymbol{\lambda}+\boldsymbol{\lambda}^{\boldsymbol{H}}\left(\boldsymbol{d}(n)-\boldsymbol{A}(n)\boldsymbol{w}(n+1)\right)\right\}\right\}$$

care conduce la

$$\nabla_{\boldsymbol{w}^*(n+1)}\left\{\mathrm{Re}\left\{\left(\boldsymbol{d}(n)-\boldsymbol{A}(n)\boldsymbol{w}(n+1)\right)^H\boldsymbol{\lambda}\right\}\right\}=-\frac{1}{2}\boldsymbol{A}^{\boldsymbol{H}}(n)\boldsymbol{\lambda}$$

În consecinţă,

$$\nabla_{\boldsymbol{w}^*(n+1)}\{J(n)\}=\boldsymbol{w}(n+1)-\boldsymbol{w}(n)-\frac{1}{2}\boldsymbol{A}^{\boldsymbol{H}}(n)\boldsymbol{\lambda}$$

şi condiţia $\nabla_{\boldsymbol{w}^*(n+1)}\{J(n)\}=\boldsymbol{0}$ conduce la

$$\boldsymbol{w}(n+1)=\boldsymbol{w}(n)+\frac{1}{2}\boldsymbol{A}^{\boldsymbol{H}}(n)\boldsymbol{\lambda}$$

$\boldsymbol{\lambda}$ urmează a fi determinat din setul de condiţii $\boldsymbol{d}(n)=\boldsymbol{A}(n)\boldsymbol{w}(n+1)$. Pentru aceasta, se înmulţeşte relaţia precedentă la stânga cu $\boldsymbol{A}(n)$ şi rezultă

$$\boldsymbol{d}(n)=\boldsymbol{A}(n)\boldsymbol{w}(n)+\frac{1}{2}\boldsymbol{A}(n)\boldsymbol{A}^{\boldsymbol{H}}(n)\boldsymbol{\lambda}$$

Introducând vectorul erorilor

$$\boldsymbol{e}(n)=\boldsymbol{d}(n)-\boldsymbol{A}(n)\boldsymbol{w}(n)$$

rezultă

$$\boldsymbol{\lambda}=2\left(\boldsymbol{A}(n)\boldsymbol{A}^{\boldsymbol{H}}(n)\right)^{-\mathbf{1}}\boldsymbol{e}(n)$$

Relaţia de reactualizare a coeficienţilor devine

$$\boldsymbol{w}(n+1)=\boldsymbol{w}(n)+\bar{\mu}\boldsymbol{A}^H(n)\left(\boldsymbol{A}(n)\boldsymbol{A}^{\boldsymbol{H}}(n)\right)^{-\mathbf{1}}\boldsymbol{e}(n)$$

unde, ca şi în cazul algoritmului NLMS a mai fost introdus coeficientul de ponderare $\bar{\mu}$ (*pas normalizat*). Înlocuind vectorul $\boldsymbol{e}(n)$ în relaţia de mai sus, rezultă

$$\boldsymbol{w}(n+1)=\left[\mathbf{I}-\bar{\mu}\boldsymbol{A}^H(n)\left(\boldsymbol{A}(n)\boldsymbol{A}^H(n)\right)^{-\mathbf{1}}\boldsymbol{A}(n)\right]\boldsymbol{w}(n)+\bar{\mu}\boldsymbol{A}^H(n)\left(\boldsymbol{A}(n)\boldsymbol{A}^H(n)\right)^{-\mathbf{1}}\boldsymbol{d}(n)$$

Matricea

$$\boldsymbol{A}^{+}=\boldsymbol{A}^H(n)\left(\boldsymbol{A}(n)\boldsymbol{A}^H(n)\right)^{-\mathbf{1}}$$

este pseudoinversa matricei $\boldsymbol{A}(n)$, iar

$$\boldsymbol{P} = \boldsymbol{A}^H(n)\left(\boldsymbol{A}(n)\boldsymbol{A}^H(n)\right)^{-\mathbf{1}}\boldsymbol{A}(n)$$

este matricea de proiecţie pe spaţiul definit de vectorii coloană ai matricei $\boldsymbol{A}(n)$.

În consecinţă relaţia de reactualizare a coeficienţilor se mai poate scrie

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \bar{\mu}\boldsymbol{A}^+(n)\boldsymbol{e}(n)$$

Ca şi în cazul algoritmului NLMS, există riscul ca matricea $\boldsymbol{A}(n)$ să devină nulă (dacă datele de la intrare sunt nule), şi matricea $\left(\boldsymbol{A}(n)\boldsymbol{A}^{\boldsymbol{H}}(n)\right)$ să devină neinversabilă. Pentru a evita o asemenea situaţie, se introduce şi aici un *coeficient de regularizare*, aşa încât formula de reactualizare a coeficienţilor devine

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \bar{\mu}\boldsymbol{A}^H(n)\left[\boldsymbol{A}(n)\boldsymbol{A}^{\boldsymbol{H}}(n) + \delta\boldsymbol{I}\right]^{-\mathbf{1}}\boldsymbol{e}(n)$$

unde δ este o constantă mică pozitivă.

$$\begin{gathered}
\mathbf{w}(0) = \mathbf{0} \\
\text{for} \quad n = 0,1,2,\cdots \\
y(n) = \mathbf{w}^H(n)\mathbf{x}(n) \\
\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{A}(n)\mathbf{w}(n) \\
\mathbf{w}(n+1) = \mathbf{w}(n) + \bar{\mu}\mathbf{A}^H(n)\left[\mathbf{A}(n)\mathbf{A}^H(n) + \delta\mathbf{I}\right]^{-1}\mathbf{e}(n) \\
\text{end}
\end{gathered}$$

Observaţii

- Algoritmul NLMS poate fi privit ca un caz particular al algoritmului APA, pentru $M = 1$.
- O analiză a convergenţei conduce la faptul că stabilitatea algoritmului necesită ca pasul să fie ales la fel ca la algoritmul NLMS, $\bar{\mu} \in (0,2)$.
- Viteza de convergență și dezadaptarea variază cu $\bar{\mu}$ în același mod ca în cazul NLMS.
- Viteza de convergenţă, pentru același $\bar{\mu}$, este mai mare ca în cazul NLMS, pentru semnale de intrare corelate şi creşte cu numărul M, dar creşterea scade pe măsură ce creşte M. În principiu, ordinul optim al proiecţiei (până la care se obţine o creştere semnificativă) este dependent de durata corelaţiei. Mărirea lui M nu conduce însă la o creştere semnificativă a dezadaptării.
- Complexitatea aritmetică, în mod evident va fi mai ridicată decât la NLMS, o pondere importantă având-o calculul inversei. Pentru $N \gg M$, acesta este aproximativ proporţională cu NM.