ΒΕΣ 06 – Προσαρμοστικά Συστήματα στις Τηλεπικοινωνίες

Προσαρμοστικοί Αλγόριθμοι Υλοποίησης Βέλτιστων Ψηφιακών Φίλτρων:

Ο αναδρομικός αλγόριθμος ελάχιστων τετραγώνων (RLS – Recursive Least Squares)

ΒΕΣ 06: Προσαρμοστικά Συστήματα στις Τηλεπικοινωνίες

Βιβλιογραφία Ενότητας

- *Benvenuto [2002]:*
  - *Κεφάλαιο 2, Ενότητα 2.3*
  - *Κεφάλαιο 3, Ενότητα 3.2*
- *Widrow [1985]: Chapter 5*
- *Haykin [2001]: Chapters 11 & 13*
- *Sayed [2003]: Chapter 5*
- *Boroujeny [1999]: Chapter 5*
- *Bose [2003]: Chapter 9*
- *Chassaing [2004]: Chapter 7*

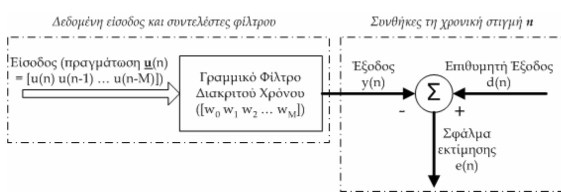
Εισαγωγή

- Ο αλγόριθμος καθόδου κατά τη μέγιστη κλίση (Steepest Descent), ο αλγόριθμος LMS και οι παραλλαγές του (FLMS, sign-LMS κλπ) αποτελούν αναδρομικές τεχνικές για την επίλυση των εξισώσεων Wiener-Hoph
- Οι εξισώσεις Wiener-Hoph προέκυψαν από ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος: $J = E[e^2(n)]$ με την υπόθεση ότι η είσοδος στο σύστημα μας είναι μια υπό την ευρεία έννοια στάσιμη στοχαστική διεργασία
- Υπάρχουν πολλές περιπτώσεις που η είσοδος σε ένα σύστημα δεν είναι υπό την ευρεία έννοια στάσιμη (έχει δηλαδή χρονικά μεταβαλλόμενα στατιστικά χαρακτηριστικά).
- Σε αυτές τις περιπτώσεις επιχειρείται η ελαχιστοποίηση όχι του μέσου τετραγωνικού σφάλματος (στιγμιαία τιμή του σφάλματος e(n) για πολλές πραγματώσεις) αλλά η απόκλιση της εκτιμούμενης εξόδου y(n) από την επιθυμητή d(n) για το σύνολο των παρατηρήσεων (n = 0,1, …)

Ο Αλγόριθμος Ελάχιστων Τετραγώνων (Least Squares)

- Έστω ότι έχουμε καταγράψει τα *N+1* πρώτα δείγματα [*u(0) u(1) … u(N)*] της πραγμάτωση *u(n)* μιας στοχαστικής διεργασίας.
- Θεωρούμε ότι οι αντίστοιχες τιμές της επιθυμητής εξόδου *d(n)* είναι [*d(0) d(1) … d(N)*]
- Ο αλγόριθμος Ελάχιστων Τετραγώνων προσπαθεί να βρει τις τιμές εκείνες $\mathbf{w} = [w_0\ w_1\ \dots\ w_M]$ για το προσαρμοστικό φίλτρο ώστε να ταιριάξει όσο το δυνατόν καλύτερα **το σύνολο** των τιμών *d(n)* και *y(n)*, *n = 0,1,…,N* όπου $y(n) = \mathbf{w}^T \mathbf{u}(n)$ και $\mathbf{u}(n) = [\mathrm{u(n)\ u(n-1) \dots u(n-M)}]^{\mathrm{T}}$
- Επομένως το κριτήριο κόστους ορίζεται ως:

$$J(\mathbf{w}) = \sum_{n=0}^{N} \{d(n) - \mathbf{w}^T \mathbf{u}(n)\}^2$$



Ο Αλγόριθμος Ελάχιστων Τετραγώνων (II)

- Ορίζουμε τα πιο κάτω διανύσματα:

$$\mathbf{e}(N) = \mathbf{d}(N) - \mathbf{y}(N) = \begin{bmatrix} d(0) \\ d(1) \\ \dots \\ d(N) \end{bmatrix} - \begin{bmatrix} y(0) \\ y(1) \\ \dots \\ y(N) \end{bmatrix} = \begin{bmatrix} d(0) \\ d(1) \\ \dots \\ d(N) \end{bmatrix} - \begin{bmatrix} \mathbf{w}^T\mathbf{u}(0) \\ \mathbf{w}^T\mathbf{u}(1) \\ \dots \\ \mathbf{w}^T\mathbf{u}(N) \end{bmatrix} = \begin{bmatrix} e(0) \\ e(1) \\ \dots \\ e(N) \end{bmatrix}$$

όπου

$$\mathbf{w}^T = [w_0 \quad w_1 \quad \dots \quad w_M]$$

$$\mathbf{u}(n) = [u(n) \quad u(n-1) \quad \dots \quad u(n-M)]^T$$

- Το κριτήριο κόστους $J(\mathbf{w}) = \sum_{n=0}^{N}\{d(n) - y(n)\}^2 = \sum_{n=0}^{N}\{d(n) - \mathbf{w}^T\mathbf{u}(n)\}^2$ μπορεί να εκφρασθεί ως:

$$J(\mathbf{w}) = \mathbf{e}^H(N)\mathbf{e}(N) = (\mathbf{d}(N) - \mathbf{y}(N))^H(\mathbf{d}(N) - \mathbf{y}(N))$$

☑ Εισαγωγή
★ Ο αλγόριθμος ελάχιστων τετραγώνων (LS)
☐ Η αρχή της ορθογωνιότητας
☐ Ο αλγόριθμος RLS
☐ Σύγκλιση αλγορίθμου RLS
☐ Παραδείγματα

Ο Αλγόριθμος Ελάχιστων Τετραγώνων (III)

- Ορίζουμε τον πίνακα δεδομένων εισόδου (ή παρατηρήσεων):

$$\mathbf{A}(N) = \begin{bmatrix} u(0) & u(-1) & \dots & u(-M) \\ u(1) & u(0) & \dots & u(1-M) \\ \dots & \dots & \dots & \dots \\ u(N) & u(N-1) & \dots & u(N-M) \end{bmatrix}$$

- Με τη βοήθεια του ανωτέρω πίνακα το διάνυσμα εξόδου **y**(n) μπορεί να εκφρασθεί ως:

$$\mathbf{y}(N) = \mathbf{A}(N)\mathbf{w}$$

Οπότε το κριτήριο κόστους γίνεται:

$$J(\mathbf{w}) = (\mathbf{d}(N) - \mathbf{y}(N))^H(\mathbf{d}(N) - \mathbf{y}(N)) = (\mathbf{d}(N) - \mathbf{A}(N)\mathbf{w})^H(\mathbf{d}(N) - \mathbf{A}(N)\mathbf{w})$$

Ελαχιστοποίηση κριτηρίου κόστους

- Για την εύρεση των συντελεστών **w** του προσαρμοστικού φίλτρου που ελαχιστοποιούν το κριτήριο κόστους J(**w**) υπολογίζουμε τις μερικές παραγωγούς ως προς w_i (i = 0,…,M) και εξισώνουμε με μηδέν:

$$\nabla J(\mathbf{w}) = \nabla\{\mathbf{d}^H\mathbf{d} - \mathbf{w}^H\mathbf{A}^H\mathbf{d} - \mathbf{d}^H\mathbf{A}\mathbf{w} + \mathbf{w}^H\mathbf{A}^H\mathbf{A}\mathbf{w}\} =$$
$$= -2\mathbf{A}^H\mathbf{d} + 2\mathbf{A}^H\mathbf{A}\mathbf{w} = \mathbf{0}$$

- Οι εξισώσεις: $\mathbf{A}^H\mathbf{A}\mathbf{w} = \mathbf{A}^H\mathbf{d}$

  που προκύπτουν από την πιο πάνω διαδικασία ονομάζονται **κανονικές εξισώσεις**.
- Η λύση των κανονικών εξισώσεων ως προς **w** μας δίνει την επιθυμητή λύση που ελαχιστοποιεί το κριτήριο J(**w**):

$$\mathbf{w}_o = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{d}$$

ΒΕΣ 06: Προσαρμοστικά Συστήματα στις Τηλεπικοινωνίες

☑ Εισαγωγή
★ Ο αλγόριθμος ελάχιστων τετραγώνων (LS)
☐ Η αρχή της ορθογωνιότητας
☐ Ο αλγόριθμος RLS
☐ Σύγκλιση αλγορίθμου RLS
☐ Παραδείγματα

Πίνακας συσχέτισης

- Πολύ συχνά οι κανονικές εξισώσεις εκφράζονται με τη βοήθεια του πίνακα συσχέτισης δειγμάτων εισόδου **Φ** = $\mathbf{A}^H\mathbf{A}$ και του διανύσματος ετεροσυσχέτισης επιθυμητής εξόδου και εισόδου **z** = $\mathbf{A}^H\mathbf{d}$:

  $\mathbf{w}_o = \mathbf{\Phi}^{-1}\mathbf{z}$
- Ο πίνακας συσχέτισης:
  - Είναι ερμιτιανός ($\mathbf{\Phi}=\mathbf{\Phi}^H$)
  - Είναι θετικά ημιορισμένος και επομένως σχεδόν πάντοτε αντιστρέψιμος
  - Έχει πραγματικές και μη αρνητικές ιδιοτιμές
  - Μπορεί να εκφραστεί ως: $\mathbf{\Phi} = \sum_{n=0}^{N}\mathbf{u}(n)\mathbf{u}^T(n)$
- Το ελάχιστο σφάλμα προκύπτει με αντικατάσταση της βέλτιστης λύσης στην τιμή του κριτηρίου κόστους:

$$J_{\min} = J(\mathbf{w}_o) = \mathbf{d}^H\mathbf{d} - \mathbf{z}^H\mathbf{w}_o = \mathbf{d}^H\mathbf{d} - \mathbf{z}^H\mathbf{\Phi}^{-1}\mathbf{z} = \mathbf{d}^H\mathbf{d} - \mathbf{d}^H\mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{d}$$

  - Η ποσότητα $\mathbf{d}^H\mathbf{d}$ εκφράζει την ενέργεια της επιθυμητής εισόδου.

Παράδειγμα

- Δίνονται τα πρώτα 8 δείγματα μιας πραγμάτωσης u(n) μιας στοχαστικής διεργασίας:

  u = [4 3 2 1 2 3 4 5]
- Αν η επιθυμητή έξοδος είναι:

  d = [3 2 1 2 3 4 5 4]
  - Να εφαρμοσθεί ο αλγόριθμος LS και να βρεθεί η βέλτιστη λύση για φίλτρο 3 συντελεστών (**w** = [w_0 w_1 w_2]).
- Λύση:
  - Ο πίνακας δεδομένων εισόδου είναι $\mathbf{A}(8) = \begin{bmatrix} 4 & 0 & 0 \\ 3 & 4 & 0 \\ 2 & 3 & 4 \\ 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \\ 4 & 3 & 2 \\ 5 & 4 & 3 \end{bmatrix}$

$$\mathbf{w}_o = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{d} = \begin{bmatrix} 0.956 \\ -0.136 \\ 0.192 \end{bmatrix}$$

Η αρχή της ορθογωνιότητας

- Η συνάρτηση κόστους μπορεί να γραφεί:

$$J(\mathbf{w}) = \mathbf{e}^H(N)\mathbf{e}(N) = \sum_{n=0}^{N} e(n)e^*(n)$$

  - Δεδομένου ότι: $e(n) = d(n) - w^H\mathbf{u}(n)$
  - Εξίσωση των μερικών παραγώγων του κριτήριου κόστους J(**w**) ως προς w_i (i = 0,…,M) μας οδηγεί στις εξισώσεις:

$$\sum_{n=0}^{N} u(n-k)e^*_{\min}(n) = 0 \qquad k = 0,1,\ldots,M$$

$$\Leftrightarrow \mathbf{e}^H(N)\mathbf{A}(N) = \mathbf{0}$$

- Οι ανωτέρω εξισώσεις εκφράζουν την αρχή της ορθογωνιότητας: *«Όταν το προσαρμοστικό φίλτρο λειτουργεί με τους βέλτιστους συντελεστές το διάνυσμα του σφάλματος $\mathbf{e}_{min}(N)$ = $\mathbf{d}(N)$-$\mathbf{A}(N)\mathbf{w}_o$ είναι κάθετο προς τις παρατηρήσεις εισόδου και τις καθυστερημένες έως και Μ εκδοχές τους»*

Παράδειγμα

- Εφαρμόζοντας την αρχή της ορθογωνιότητας στο προηγούμενο παράδειγμα έχουμε:

$$\mathbf{e}_{\min}(8) = \mathbf{d}(N) - \mathbf{A}(N)\mathbf{w} =$$

=[-0.8241 -0.3231 -1.2710 0.7407 0.8404 1.2125 1.2008 -0.8109]^T

και $\mathbf{e}^H\mathbf{A} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Ο αναδρομικός αλγόριθμος ελάχιστων τετραγώνων (RLS)

- Η λύση των κανονικών εξισώσεων για N+1 παρατηρήσεις εισόδου βρήκαμε ότι υπολογίζεται από τη σχέση:

$$\mathbf{w}_o = \mathbf{\Phi}^{-1}\mathbf{z} \;\; \Leftrightarrow \mathbf{w}_o = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{d}$$

- Τι γίνεται αν έχουμε μια νέα παρατήρηση *u(N+1)* στην είσοδο (δηλαδή αν έχουμε N+2 στοιχεία);
  - Είναι προφανές ότι οι κανονικές εξισώσεις μπορεί να διαμορφωθούν ανάλογα προσθέτοντας την επιθυμητή απόκριση για το νέο στοιχείο στο διάνυσμα **d**(N) καθώς και μια νέα στήλη γραμμή στον πίνακα δεδομένων εισόδου **A**(N).

$$\mathbf{d}(N+1) = \begin{bmatrix} \mathbf{d}(N) \\ d(N+1) \end{bmatrix} = \begin{bmatrix} d(0) \\ \ldots \\ d(N) \\ d(N+1) \end{bmatrix}$$

$$\mathbf{A}(N+1) = \begin{bmatrix} \mathbf{A}(N) \\ \mathbf{u}^T(N+1) \end{bmatrix}$$

όπου

$$\mathbf{u}^T(N+1) = [u(N+1) \quad u(N) \quad \ldots \quad u(N-M+1)]$$

Ο αναδρομικός αλγόριθμος ελάχιστων τετραγώνων (ΙΙ)

- Επομένως η λύση των κανονικών εξισώσεων για N+2 υπολογίζεται από τη σχέση:

$$\mathbf{w}_o(N+1) = \mathbf{\Phi}^{-1}(N+1)\mathbf{z}(N+1)$$
$$\Leftrightarrow \mathbf{w}_o(N+1) = (\mathbf{A}(N+1)^H \mathbf{A}(N+1))^{-1}\mathbf{A}(N+1)^H \mathbf{d}(N+1)$$

- Ο αναδρομικός αλγόριθμος ελάχιστων τετραγώνων υπολογίζει τους συντελεστές **w**(N+1) τροποποιώντας κατάλληλα τους συντελεστές **w**(N) χωρίς να χρειάζεται να αντιστρέψει τον πίνακα συσχέτισης **Φ**(N+1) (η αντιστροφή του **Φ**(N+1) αυξανομένου του N είναι ιδιαίτερα χρονοβόρα)
  - Οι παρακάτω σχέσεις μπορούν εύκολα να αποδειχτούν:

$$\mathbf{\Phi}(N+1) = \mathbf{\Phi}(N) + \mathbf{u}(N+1)\mathbf{u}^H(N+1)$$
$$\mathbf{z}(N+1) = \mathbf{z}(N) + d^*(N+1)\mathbf{u}(N+1)$$

  - Για τον αναδρομικό υπολογισμό των **w**(N+1) η πραγματική δυσκολία έγκειται στην αντιστροφή του πίνακα $\mathbf{\Phi}(N+1) = \mathbf{\Phi}(N) + \mathbf{u}(N+1)\mathbf{u}^H(N+1)$ ως συνάρτηση του $\mathbf{\Phi}^{-1}(N)$.
  - Το λήμμα αντιστροφής πινάκων δίνει τη λύση στο πρόβλημα αυτό

Λήμμα αντιστροφής πινάκων

- Έστω ο πίνακας: $A = B^{-1}+CD^{-1}C^H$, όπου οι πίνακες A, B, D είναι θετικά ορισμένοι πίνακες (άρα αντιστρέψιμοι).
- Τα λήμμα αντιστροφής πινάκων εκφράζεται από την κατωτέρω εξίσωση πινάκων:

$$\mathbf{A}^{-1} = \mathbf{B} - \mathbf{BC}\left(\mathbf{D} + \mathbf{C}^H\mathbf{BC}\right)^{-1}\mathbf{C}^H\mathbf{B}$$

  - Αν θέσουμε στη ανωτέρω σχέση:

$$\mathbf{A} = \mathbf{\Phi}(N+1)$$
$$\mathbf{B}^{-1} = \mathbf{\Phi}(N)$$
$$\mathbf{C} = \mathbf{u}(N+1)$$
$$\mathbf{D} = 1$$

  Έχουμε:

$$\mathbf{\Phi}^{-1}(N+1) = \mathbf{\Phi}^{-1}(N) - \frac{\mathbf{\Phi}^{-1}(N)\mathbf{u}(N+1)\mathbf{u}^H(N+1)\mathbf{\Phi}^{-1}(N)}{1+\mathbf{u}^H(N+1)\mathbf{\Phi}^{-1}(N)\mathbf{u}(N+1)}$$

- Ορίζοντας

$$\mathbf{P}(N+1) = \mathbf{\Phi}^{-1}(N+1)$$
$$\mathbf{k}(N+1) = \frac{\mathbf{P}(N)\mathbf{u}(N+1)}{1+\mathbf{u}^H(N+1)\mathbf{P}(N)\mathbf{u}(N+1)}$$

παίρνουμε την αναδρομική σχέση υπολογισμού του αντίστροφου πίνακα συσχέτισης $\mathbf{P}(N+1) = \mathbf{P}(N) - \mathbf{k}(N+1)\mathbf{u}^H(N+1)\mathbf{P}(N)$

Ο αλγόριθμος RLS

- Με τη βοήθεια των προηγούμενων σχέσεων προκύπτει η αναδρομική σχέση υπολογισμού των συντελεστών **w**(N+1):

$$\mathbf{w}(N+1)=\mathbf{w}(N)-\mathbf{k}(N+1)\mathbf{u}^H(N+1)\mathbf{w}(N)+\mathbf{P}(N)\mathbf{u}(N+1)d^*(N+1)$$

- Συνήθως στο αλγόριθμο RLS χρησιμοποιείται μια παράμετρος μνήμης *λ* (0<λ<1) ώστε η ενημέρωση των συντελεστών **w**(N+1) να γίνεται δίνοντας μεγαλύτερη βαρύτητα στις πιο καινούργιες παρατηρήσεις
  - Αυτό έχει ως αποτέλεσμα να ελαχιστοποιείται τα σφάλμα

    $$J(\mathbf{w})=\sum_{n=0}^{N}\lambda^{N-n}e(n)e^*(n)$$

    αντί για το σφάλμα

    $$J(\mathbf{w})=\mathbf{e}^H(N)\mathbf{e}(N)=\sum_{n=0}^{N}e(n)e^*(n)$$

Ο αλγόριθμος RLS (II)

- Στους αρχικούς υπολογισμούς των συντελεστών **w**(N+1) (N = 0,1,…,M-1) ο πίνακας **Φ**(N) δεν είναι αντιστρέψιμος.
- Για να μπορεί να εκτελεστεί ο αναδρομικός υπολογισμός λαμβάνεται αρχικά: $\mathbf{P}(0)=\delta^{-1}\mathbf{I}$

  Όπου **I** είναι μοναδιαίος πίνακας διαστάσεων (M+1)x(M+1) και *δ* μια θετική σταθερά με πολύ μικρή τιμή.
- Επομένως ο αλγόριθμος RLS συνοψίζονται τελικά ως:

$$\begin{aligned}
&\text{Αρχικοποίηση:}\\
&\mathbf{P}(0)=\delta^{-1}\mathbf{I} \qquad (\delta \text{ μικρή θετική σταθερά})\\
&\mathbf{w}(0)=\mathbf{0}\\
&\text{Για κάθε n, } n=1,2....\\
&\mathbf{k}(n)=\frac{\lambda^{-1}\mathbf{P}(n-1)\mathbf{u}(n)}{1+\lambda^{-1}\mathbf{u}^H(n)\mathbf{P}(n-1)\mathbf{u}(n)}\\
&\xi(n)=d(n)-\mathbf{w}^H(n-1)\mathbf{u}(n)\\
&\mathbf{w}(n)=\mathbf{w}(n-1)+\mathbf{k}(n)\xi^*(n)\\
&\mathbf{P}(n)=\lambda^{-1}\mathbf{P}(n-1)-\lambda^{-1}\mathbf{k}(n)\mathbf{u}^H(n)\mathbf{P}(n-1)
\end{aligned}$$

Παράδειγμα

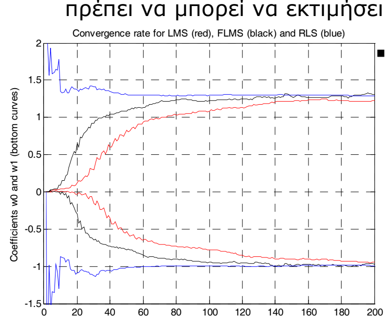
- Δίνονται τα πρώτα 8 δείγματα μιας πραγμάτωσης u(n) μιας στοχαστικής διεργασίας:

  u = [4 3 2 1 2 3 4 5]
- Αν η επιθυμητή έξοδος είναι:

  d = [3 2 1 2 3 4 5 4]
  - Να εφαρμοσθεί ο αλγόριθμος RLS και να βρεθεί η βέλτιστη λύση για φίλτρο 3 συντελεστών (**w** = [w_0 w_1 w_2]) και να δοθούν οι διαδοχικές εκτιμήσεις για τους συντελεστές **w** (χρησιμοποιήστε δ = 0.1).
  - Να συγκριθεί η τελική λύση (n = 8) με τη λύση των ελάχιστων τετραγώνων. Που οφείλονται οι διαφορές στους συντελεστές;

Παραδείγματα

- Δίνεται το σήμα x(n) το οποίο αποτελεί πραγμάτωση μιας στοχαστικής διεργασίας. Να βρεθεί γραμμικός προβλέπτης δύο συντελεστών ([w_1 w_2]) για πρόβλεψη της τιμής x(n+1) με τη βοήθεια των αλγορίθμων LMS, FLMS και RLS
  - Έστω $x(n) = a_1x(n-1)+a_2x(n-2)+v(n)$, όπου v(n) είναι λευκός θόρυβος με μέση τιμή μ_v=0 και διασπορά σ_v^2. Ο γραμμικός προβλέπτης (συντελεστές **w**) πρέπει να μπορεί να εκτιμήσει τις τιμές a_1 και a_2.

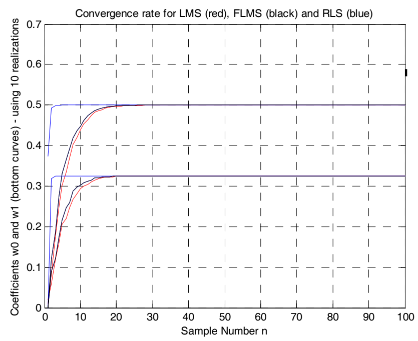


- Στο διπλανό σχήμα φαίνεται η σταδιακή προσέγγιση των τιμών a_1 (=1.3) και a_2 (=-0.995) από τους συντελεστές του φίλτρου [w1 w2] με τη βοήθεια των αλγορίθμων
  - LMS (μ = 0.001) – κόκκινη καμπύλη
  - FLMS (μ = 0.001) – μαύρη καμπύλη
  - RLS (λ = 0.99, δ=0.01) – μπλε καμπύλη

Αναγνώριση συστήματος

- Έστω ότι το άγνωστο σύστημα περιγράφεται από τη συνάρτηση μεταφοράς: $H(z) = \dfrac{0.5 + 0.2z^{-1} + 0.1z^{-2}}{1 - 0.25z^{-1}}$
- Για μοντελοποίηση του ανωτέρω συστήματος με FIR φίλτρο τάξης 5 (6 συντελεστών) η βέλτιστη λύση (λύση Wiener) είναι:

$$\mathbf{w}_o = [0.5 \quad 0.325 \quad 0.1812 \quad 0.0453 \quad 0.0113 \quad 0.0028]^T$$



Στο διπλανό σχήμα φαίνεται η σταδιακή προσέγγιση των τιμών w_0 (=0.5) και w_1 (=0.325) με τη βοήθεια των αλγορίθμων LMS (κόκκινες καμπύλες), FLMS (μαύρες καμπύλες) και RLS (μπλε καμπύλες).