Στοχαστικά Σήματα και Τηλεπικοινωνιές

Ενότητα 5: Προσαρμοστική Επεξεργασία

Καθηγητής Κώστας Μπερμπερίδης
Πολυτεχνική Σχολή
Τμήμα Μηχανικών Η/Υ και Πληροφορικής

Σκοποί ενότητας

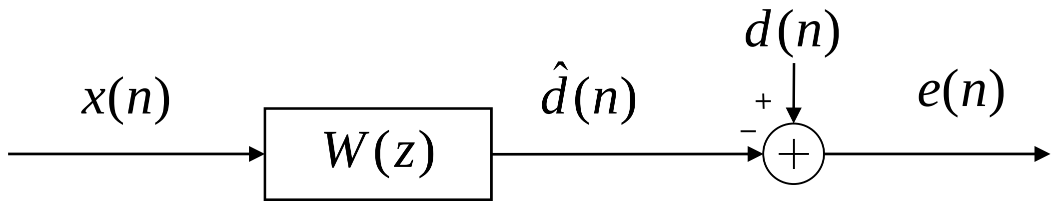
- Παρουσίαση των βασικών εννοιών των προσαρμοστικών αλγορίθμων καθώς και η περαιτέρω ανάλυση των κυριότερων από αυτούς.

Περιεχόμενα ενότητας

- Εισαγωγικές έννοιες
- Μέθοδος μέγιστης καθόδου
- Αλγόριθμος LMS
- Αλγόριθμος NLMS
- Αλγόριθμος RLS

Εισαγωγή (1/4)

- Υπολογισμός FIR φίλτρου Wiener σε στάσιμο περιβάλλον: Οι διαδικασίες $x(n)$ και $d(n)$ είναι από κοινού WSS.



$$\min_{w(k)} \xi = \min_{w(k)} E\left\{ \left| e(n) \right|^2 \right\} \quad \Longrightarrow \quad \sum_{l=0}^{p-1} w(l) r_x(k-l) = r_{dx}(k) \quad \Longrightarrow \quad \boldsymbol{R}_x \boldsymbol{w} = \boldsymbol{r}_{dx}$$

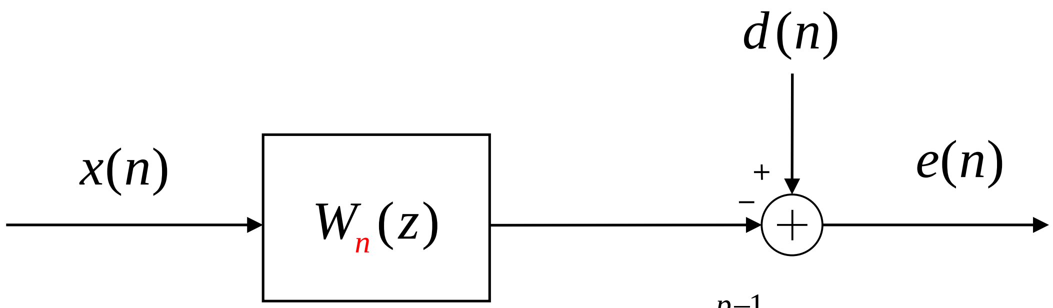
$$\text{για} \quad k = 0, 1, \ldots, p-1$$

$$r_x(k) = E\left\{ x(n) x^*(n-k) \right\} \longrightarrow \hat{r}_x(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) x^*(n-k)$$

$$r_{dx}(k) = E\left\{ d(n) x^*(n-k) \right\} \longrightarrow \hat{r}_{dx}(k) = \frac{1}{N} \sum_{n=0}^{N-1} d(n) x^*(n-k)$$

Εισαγωγή (2/4)

- Υπολογισμός FIR προσαρμοστικού φίλτρου σε μη στάσιμο περιβάλλον: Οι διαδικασίες $x(n)$ και $d(n)$ δεν είναι στάσιμες.



$$\hat{d}(n) = \sum_{k=0}^{p-1} w_n(k)x(n-k)$$

οι εξισώσεις εξαρτώνται από το n

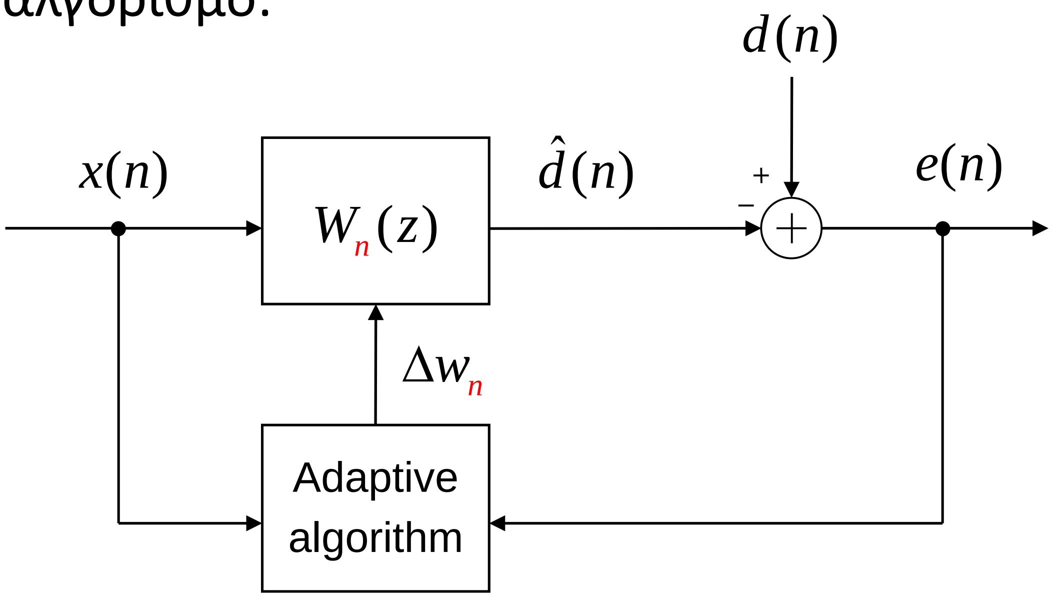
$$\min_{w_n(k)} \xi(n) = \min_{w_n(k)} E\left\{|e(n)|^2\right\} \Rightarrow$$

$$\sum_{l=0}^{p-1} w_n(l)E\left\{x(n-l)x^*(n-k)\right\} = E\left\{d(n)x^*(n-k)\right\} \Rightarrow \boldsymbol{R}_x(n)\boldsymbol{w}_n = \boldsymbol{r}_{dx}(n)$$

για $k = 0,1,\ldots,p-1$

Εισαγωγή (3/4)

- Αναζητούμε έναν επαναληπτικό προσαρμοστικό (adaptive) αλγόριθμο:



$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \Delta \boldsymbol{w}_n$$

- Το $\boldsymbol{\Delta w}_n$ είναι ένας διορθωτικός όρος, ο οποίος ανανεώνει τους συντελεστές τη χρονική στιγμή $n+1$.

- Ο αλγόριθμος χρησιμοποιεί το σήμα σφάλματος $e(n)$ για να μετρήσει την απόδοση του φίλτρου και να ανανεώσει του συντελεστές.

Εισαγωγή (4/4)

- Το προσαρμοστικό φίλτρο θέλουμε να έχει τις ακόλουθες ιδιότητες:
  - Σε ένα στάσιμο περιβάλλον, το φίλτρο θα πρέπει να παράγει μια ακολουθία διορθώσεων Δwn, κατά τέτοιο τρόπο ώστε το διάνυσμα wn να συγκλίνει στη λύση των εξισώσεων Wiener-Hopf:

$$\lim_{n\to\infty} \boldsymbol{w}_n = \boldsymbol{R}_x^{-1}\boldsymbol{r}_{dx}$$

  - Για να υπολογίσουμε τη διόρθωση $\boldsymbol{\Delta w}_n$ δε θα πρέπει να απαιτείται η εκ των προτέρων γνώση των στατιστικών $r_x(k)$ και $r_{dx}(k)$, αλλά η εκτίμηση των παραμέτρων αυτών θα πρέπει να είναι ενσωματωμένη στον προσαρμοστικό αλγόριθμο.
  - Για μη στάσιμα σήματα, το φίλτρο θα πρέπει να προσαρμόζεται στα μεταβαλλόμενα στατιστικά και να παρακολουθεί τη λύση καθώς αυτή εξελίσσεται στο χρόνο.

Μέθοδος μέγιστης καθόδου (1/7)

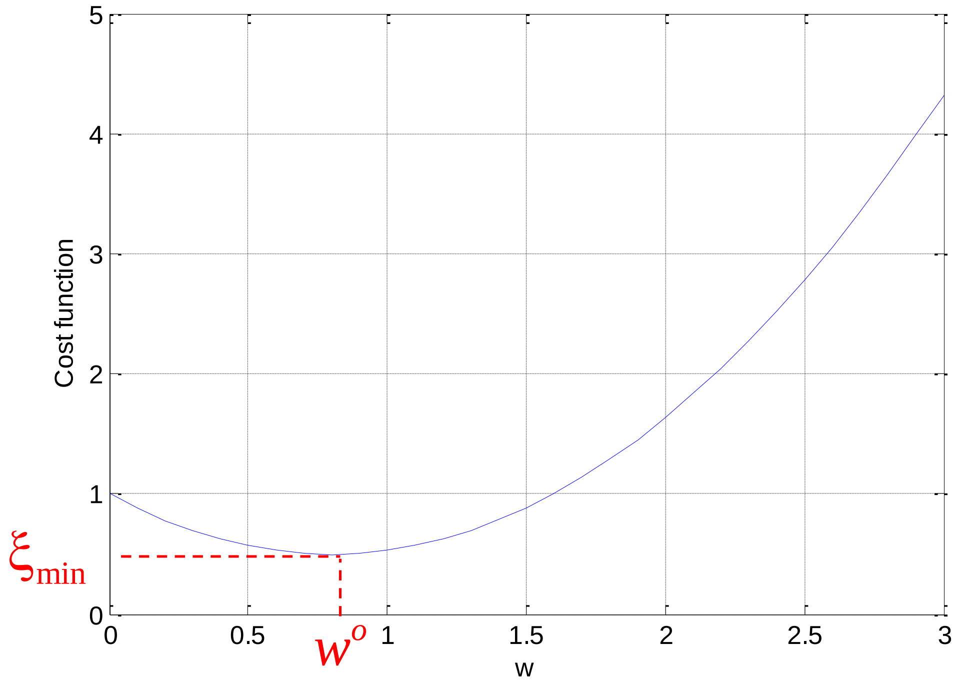
- Θεωρείστε τη συνάρτηση κόστους για το πρόβλημα βέλτιστου φιλτραρίσματος με πραγματικές τιμές για FIR φίλτρο Wiener με ένα συντελεστή:

$$\xi(\boldsymbol{w}) = r_d(0) - 2\boldsymbol{w}^T \boldsymbol{r}_{dx} + \boldsymbol{w}^T \boldsymbol{R}_x \boldsymbol{w} \quad \Rightarrow \quad \xi(w(0)) = r_d(0) - 2r_{dx}(0)w(0) + r_x(0)w^2(0)$$

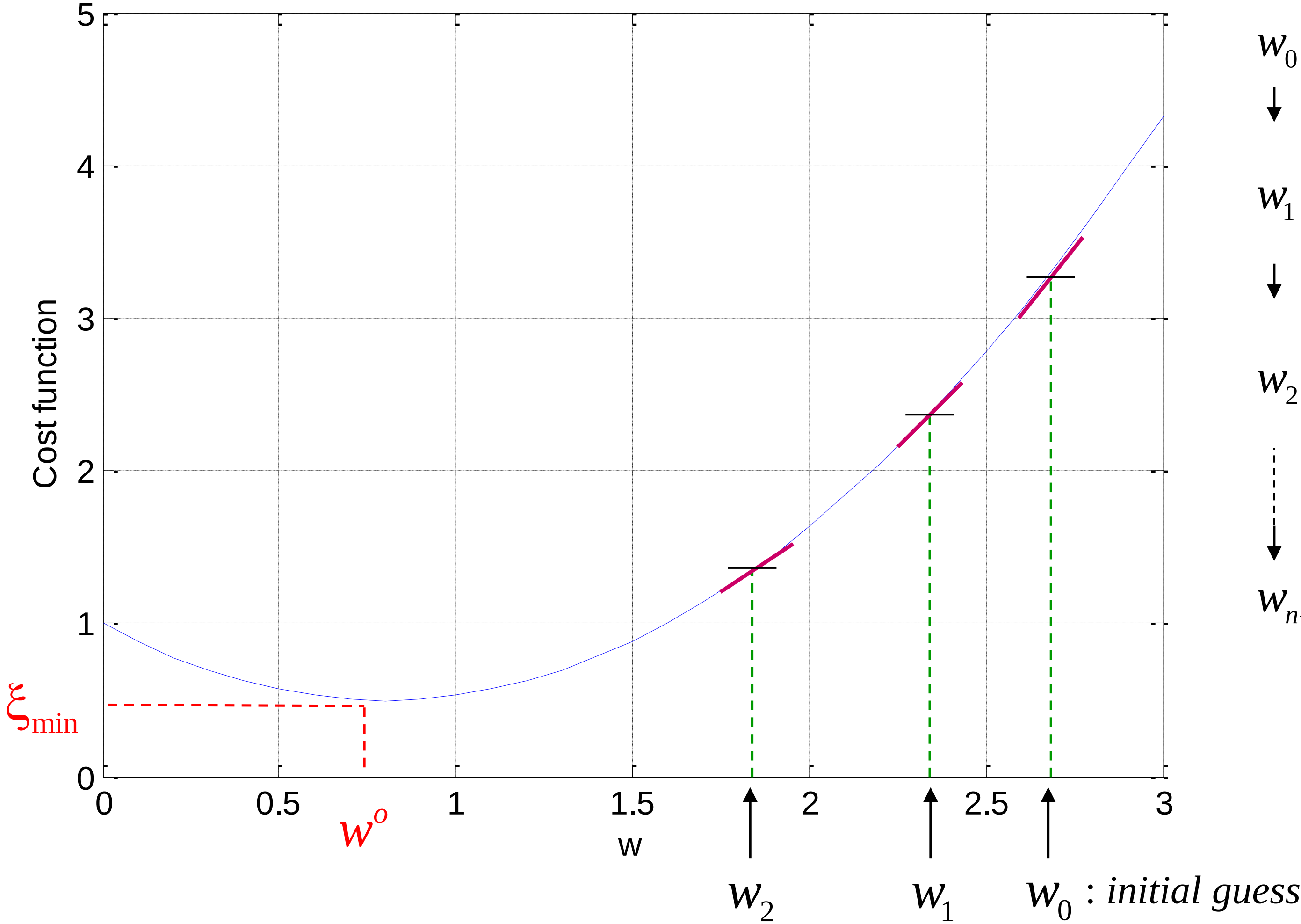
$$\frac{\partial \xi}{\partial w(0)} = -2r_{dx}(0) + 2r_x(0)w(0)$$

$$\frac{\partial^2 \xi}{\partial^2 w(0)} = 2r_x(0) > 0$$

$$\frac{\partial \xi}{\partial w(0)} = 0 \quad \Rightarrow w^o(0) = \frac{r_{dx}(0)}{r_x(0)}$$



Μέθοδος μέγιστης καθόδου (2/7)



$$w_0$$

$$\downarrow$$

$$w_1 = w_0 + \mu\left[-\left.\frac{\partial \xi(w)}{\partial w}\right|_{w=w_0} \right]$$

$$\downarrow$$

$$w_2 = w_1 + \mu\left[-\left.\frac{\partial \xi(w)}{\partial w}\right|_{w=w_1} \right]$$

$$\downarrow$$

$$w_{n+1} = w_n + \mu\left[-\left.\frac{\partial \xi(w)}{\partial w}\right|_{w=w_n} \right]$$

$$\boxed{\mu > 0}$$

Μέθοδος μέγιστης καθόδου (3/7)

- Θεωρείστε τη συνάρτηση κόστους για το πρόβλημα βέλτιστου φιλτραρίσματος με πραγματικές τιμές για FIR φίλτρο Wiener με δύο συντελεστές:

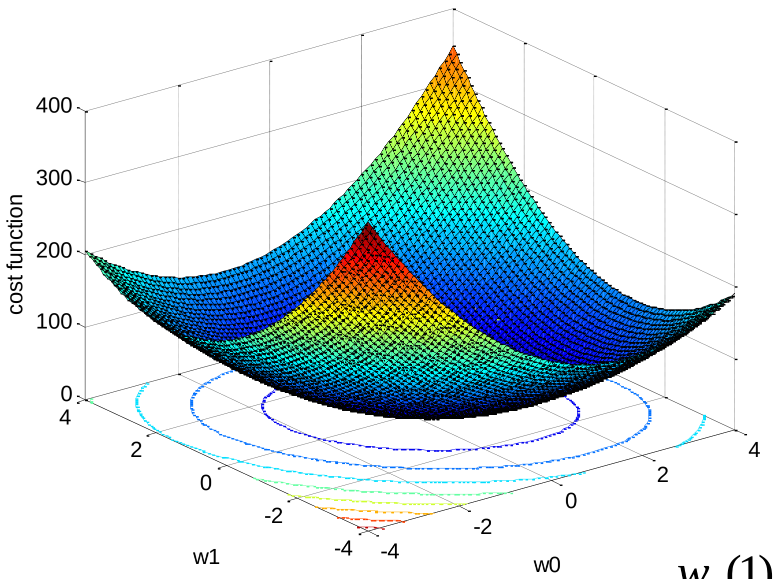
$$\xi(\boldsymbol{w}) = r_d(0) - 2\boldsymbol{w}^T \boldsymbol{r}_{dx} + \boldsymbol{w}^T \boldsymbol{R}_x \boldsymbol{w} \quad \Rightarrow$$

$$\xi(\boldsymbol{w}) = r_d(0) - 2r_{dx}(0)w(0) - 2r_{dx}(1)w(1) + r_x(0)\left[w^2(0) + w^2(1)\right] + 2r_x(1)w(0)w(1)$$

$$\nabla\xi = \begin{bmatrix} \dfrac{\partial \xi}{\partial w(0)} \\ \dfrac{\partial \xi}{\partial w(1)} \end{bmatrix} = \begin{bmatrix} -2r_{dx}(0) + 2r_x(0)w(0) + 2r_x(1)w(1) \\ -2r_{dx}(1) + 2r_x(0)w(1) + 2r_x(1)w(0) \end{bmatrix} \qquad \nabla^2\xi = \begin{bmatrix} \dfrac{\partial^2 \xi}{\partial^2 w(0)} \\ \dfrac{\partial^2 \xi}{\partial^2 w(1)} \end{bmatrix} = \begin{bmatrix} 2r_x(0) \\ 2r_x(0) \end{bmatrix} > 0$$

$$\nabla\xi = \mathbf{O} \;\; \Rightarrow \begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} w^o(0) \\ w^o(1) \end{bmatrix} = \begin{bmatrix} r_{dx}(0) \\ r_{dx}(1) \end{bmatrix}$$

Μέθοδος μέγιστης καθόδου (4/7)



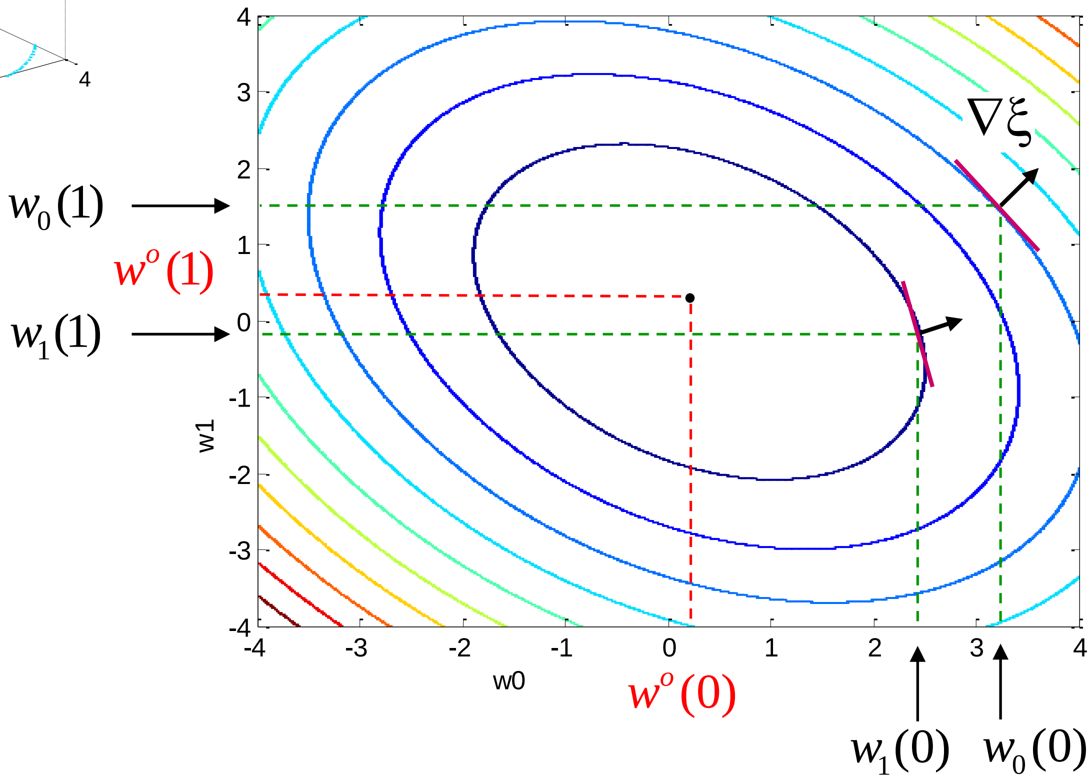
$$\boldsymbol{w}_0$$

$$\downarrow$$

$$\boldsymbol{w}_1 = \boldsymbol{w}_0 + \mu\left[-\nabla\xi\big|_{\boldsymbol{w}=\boldsymbol{w}_0}\right]$$

$$\downarrow$$

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu\left[-\nabla\xi\big|_{\boldsymbol{w}=\boldsymbol{w}_n}\right]$$



Μέθοδος μέγιστης καθόδου (5/7)

- Γενικά, ο αλγόριθμος μέγιστης καθόδου (steepest descent) συνοψίζεται ως εξής:
    1. Αρχικοποίηση του διανύσματος των συντελεστών: $\boldsymbol{w}_{n=0}$
    2. Υπολογισμός του διανύσματος $\nabla\boldsymbol{\xi}(n)$ τη χρονική στιγμή n.
    3. Ανανέωση του διανύσματος των συντελεστών:
    $$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n - \mu\, \nabla\boldsymbol{\xi}(n)$$
    4. Επιστροφή στο βήμα 2 για $n = n + 1$.
- Η ποσότητα μ ονομάζεται μέγεθος βήματος (step size) ή απλά βήμα. Είναι θετικός αριθμός και επηρεάζει το ρυθμό με τον οποίο το διάνυσμα των συντελεστών κινείται προς το ελάχιστο MSE.

Μέθοδος μέγιστης καθόδου (6/7)

- Υπολογισμός του gradient:

$$\nabla\xi(n)\Big|_{\mathbf{w}=\mathbf{w}_n} = \begin{bmatrix} \left.\dfrac{\partial\xi}{\partial w^*(0)}\right|_{w^*(0)=w_n^*(0)} \\ \vdots \\ \left.\dfrac{\partial\xi}{\partial w^*(p-1)}\right|_{w^*(p-1)=w_n^*(p-1)} \end{bmatrix} = \begin{bmatrix} -E\left\{e(n)x^*(n)\right\} \\ -E\left\{e(n)x^*(n-1)\right\} \\ \vdots \\ -E\left\{e(n)x^*(n-p+1)\right\} \end{bmatrix} = -E\left\{e(n)\mathbf{x}^*(n)\right\}$$

$$\Uparrow$$

$$\frac{\partial\xi}{\partial w^*(k)} = \frac{\partial E\left\{e(n)e^*(n)\right\}}{\partial w^*(k)} = E\left\{e(n)\frac{\partial e^*(n)}{\partial w^*(k)}\right\} = -E\left\{e(n)x^*(n-k)\right\} \qquad \text{για} \quad k=0,1,\ldots,p-1$$

$$\uparrow$$

$$e(n) = d(n) - \sum_{l=0}^{p-1} w_n(l)x(n-l)$$

Μέθοδος μέγιστης καθόδου (7/7)

- Τελικά, ο αναδρομικός τύπος μέγιστης καθόδου γράφεται:

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n - \mu \nabla \xi(n) \quad \Longrightarrow \quad \boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu E\left\{ e(n)\boldsymbol{x}^*(n) \right\}$$

- Στη συνέχεια θεωρούμε ότι οι διαδικασίες $x(n)$ και $d(n)$ είναι από κοινού WSS:

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu E\left\{ e(n)\boldsymbol{x}^*(n) \right\} \Rightarrow \boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu \left[E\left\{ \left[d(n) - \boldsymbol{w}_n^T \boldsymbol{x}(n) \right] \boldsymbol{x}^*(n) \right\} \right]$$

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu \left[E\left\{ d(n)\boldsymbol{x}^*(n) \right\} - E\left\{ \boldsymbol{w}_n^T \boldsymbol{x}(n)\boldsymbol{x}^*(n) \right\} \right] \Rightarrow$$

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu \left[E\left\{ d(n)\boldsymbol{x}^*(n) \right\} - E\left\{ \boldsymbol{x}(n)\boldsymbol{x}^H(n) \right\} \boldsymbol{w}_n \right] \Rightarrow$$

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu \left[\boldsymbol{r}_{dx} - \boldsymbol{R}_x \boldsymbol{w}_n \right] \quad \Longrightarrow \quad \text{Αν } \boldsymbol{w}_n = \boldsymbol{w}^o \text{ τότε } \boldsymbol{w}_{n+1} = \boldsymbol{w}_n$$

Μέθοδος μέγιστης καθόδου - Σύγκλιση (1/9)

- Συνέχεια: $\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu\left[\boldsymbol{r}_{dx} - \boldsymbol{R}_x \boldsymbol{w}_n\right] \Rightarrow$

$$\boldsymbol{w}_{n+1} = \left(I - \mu \boldsymbol{R}_x\right)\boldsymbol{w}_n + \mu \boldsymbol{r}_{dx} \Rightarrow$$

$$\boldsymbol{R}_x \boldsymbol{w}^0 = \boldsymbol{r}_{dx}$$

$$\boldsymbol{w}_{n+1} - \boldsymbol{w}^0 = \left(I - \mu \boldsymbol{R}_x\right)\boldsymbol{w}_n + \mu \boldsymbol{r}_{dx} - \boldsymbol{w}^0 \Rightarrow$$

$$\boldsymbol{w}_{n+1} - \boldsymbol{w}^0 = \left(I - \mu \boldsymbol{R}_x\right)\boldsymbol{w}_n - \left(I - \mu \boldsymbol{R}_x\right)\boldsymbol{w}^0 \Rightarrow$$

$$\boldsymbol{w}_{n+1} - \boldsymbol{w}^0 = \left(I - \mu \boldsymbol{R}_x\right)\left(\boldsymbol{w}_n - \boldsymbol{w}^0\right)$$

- Ορίζουμε το διάνυσμα σφάλματος των συντελεστών: $\boldsymbol{c}_n = \boldsymbol{w}_n - \boldsymbol{w}^o$
- Άρα: $\boldsymbol{c}_{n+1} = (\boldsymbol{I} - \mu \boldsymbol{R}_x)\boldsymbol{c}_n$
- Ο παραπάνω τύπος μας δείχνει πως εξελίσσεται το διάνυσμα σφάλματος των συντελεστών. Παρατηρούμε ότι αν ο πίνακας $\boldsymbol{R}_x$ δεν είναι διαγώνιος, τότε υπάρχει μια αλληλεξάρτηση μεταξύ των σφαλμάτων κάθε συντελεστή.

Μέθοδος μέγιστης καθόδου - Σύγκλιση (2/9)

- Ο πίνακας αυτοσυσχέτισης $\boldsymbol{R}_x$ είναι ερμιτιανός και μη αρνητικά ορισμένος. Από Spectral Theorem για ερμιτιανούς πίνακες μπορούμε να γράψουμε:

$$\boldsymbol{R}_x = \boldsymbol{V}\cdot\boldsymbol{\Lambda}\cdot\boldsymbol{V}^{-1} = \boldsymbol{V}\cdot\boldsymbol{\Lambda}\cdot\boldsymbol{V}^{H} = \sum_{k=0}^{p-1}\lambda_k \boldsymbol{v}_k \boldsymbol{v}_k^H$$

- Ο $\boldsymbol{V}$ είναι ο πίνακας με στήλες τα ιδιοδιανύσματα του $\boldsymbol{R}_x$ ενώ ο $\boldsymbol{\Lambda}$ είναι διαγώνιος πίνακας με τις ιδιοτιμές του $\boldsymbol{R}_x$, δηλαδή

$$\boldsymbol{\Lambda} = \mathrm{diag}\left\{\lambda_0, \lambda_1, \ldots, \lambda_{p-1}\right\}$$

- Επίσης, οι ιδιοτιμές είναι πραγματικές και μη αρνητικές $\lambda_k \geq 0$ και τα ιδιο-διανύσματα μπορούν να επιλεγούν ως ορθοκανονικά:

$$\left\langle \boldsymbol{v}_i \cdot \boldsymbol{v}_j \right\rangle = \begin{cases} 1 & \text{αν} \quad i \neq j \\ 0 & \text{αν} \quad i = j \end{cases} \Rightarrow \boldsymbol{V}\cdot\boldsymbol{V}^H = \boldsymbol{V}^H\cdot\boldsymbol{V} = \boldsymbol{I} \Rightarrow \boldsymbol{V}^{-1} = \boldsymbol{V}^H$$

Μέθοδος μέγιστης καθόδου - Σύγκλιση (3/9)

- Συνεπώς μπορούμε να γράψουμε:

$$\boldsymbol{c}_{n+1} = \left(I - \mu \boldsymbol{R}_x\right)\boldsymbol{c}_n \Rightarrow \boldsymbol{c}_{n+1} = \left(I - \mu \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H\right)\boldsymbol{c}_n$$

$$\Rightarrow \boldsymbol{c}_{n+1} = \left(\boldsymbol{V}\boldsymbol{V}^H - \mu \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H\right)\boldsymbol{c}_n$$

$$\Rightarrow \boldsymbol{c}_{n+1} = \boldsymbol{V}\left(\boldsymbol{I} - \mu \boldsymbol{\Lambda}\right)\boldsymbol{V}^H\boldsymbol{c}_n$$

$$\Rightarrow \boldsymbol{V}^H\boldsymbol{c}_{n+1} = \underbrace{\left(\boldsymbol{V}^H\boldsymbol{V}\right)}_{\boldsymbol{I}}\left(\boldsymbol{I} - \mu \boldsymbol{\Lambda}\right)\boldsymbol{V}^H\boldsymbol{c}_n$$

- Ορίζουμε το γραμμικό μετασχηματισμό: $\boldsymbol{u}_n = \boldsymbol{V}^H\boldsymbol{c}_n$
- Άρα: $\boldsymbol{u}_{n+1} = (I - \mu\boldsymbol{\Lambda})\boldsymbol{u}_n \Rightarrow \boldsymbol{u}_n = (\boldsymbol{I} - \mu\boldsymbol{\Lambda})^n\boldsymbol{u}_0$
- Σημειώνεται πως: $(I - \mu\Lambda)^n = diag\left\{(1 - \mu\lambda_0)^n, \dots, \left(1 - \mu\lambda_{p-1}\right)^n\right\}$

Μέθοδος μέγιστης καθόδου - Σύγκλιση (4/9)

- Τέλικα:

$$\boldsymbol{u}_n = \begin{bmatrix} (1-\mu\lambda_0)^n u_0(0) \\ \vdots \\ (1-\mu\lambda_{p-1})^n u_0(p-1) \end{bmatrix}$$

- Άρα, για να έχουμε σύγκλιση του διανύσματος $\boldsymbol{w}_n$ στη λύση $\boldsymbol{w}^o$, έχοντας ξεκινήσει από μια αρχική τιμή $\boldsymbol{w}_0$ (δηλαδή $\boldsymbol{u}_0$), θα πρέπει το διάνυσμα $\boldsymbol{c}_n$ να συγκλίνει στο μηδέν ή ισοδύναμα το διάνυσμα $\boldsymbol{u}_n$ να συγκλίνει στο μηδέν. Για οποιαδήποτε τιμή $\boldsymbol{u}_0$ (άρα και $\boldsymbol{w}_0$), αυτό θα συμβεί όταν:

$$|1-\mu\lambda_k| < 1 \qquad \text{για} \quad k = 0,1,\ldots,p-1$$

$$\Downarrow$$

$$0 < \mu < \frac{2}{\lambda_k} \qquad \text{για} \quad k = 0,1,\ldots,p-1 \qquad \Rightarrow \qquad 0 < \mu < \frac{2}{\lambda_{\max}}$$

Μέθοδος μέγιστης καθόδου - Σύγκλιση (5/9)

- Ισχύει λοιπόν η παρακάτω ιδιότητα για τη σύγκλιση:

Για από κοινού WSS διαδικασίες $x(n)$ και $d(n)$, το προσαρμοστικό φίλτρο μέγιστης καθόδου συγκλίνει στη λύση των εξισώσεων Wiener-Hopf, δηλαδή

$$\lim_{n\to\infty} \boldsymbol{w}_n = \boldsymbol{R}_x^{-1}\boldsymbol{r}_{dx}$$

αν το μέγεθος βήματος ικανοποιεί τη συνθήκη $0 < \mu < \frac{2}{\lambda_{max}}$,

όπου λ_{max} είναι η μέγιστη ιδιοτιμή του πίνακα αυτοσυσχέτισης $\boldsymbol{R}_x$.

Μέθοδος μέγιστης καθόδου - Σύγκλιση (6/9)

- Για το διάνυσμα των συντελεστών γράφουμε:

$$\boldsymbol{w}_n = \boldsymbol{w}^0 + \boldsymbol{c}_n = \boldsymbol{w}^0 + \boldsymbol{V}\boldsymbol{u}_n = \boldsymbol{w}^0 + \begin{bmatrix} \boldsymbol{v}_0 & \boldsymbol{v}_1 & \dots & \boldsymbol{v}_{p-1} \end{bmatrix} \begin{bmatrix} (1-\mu\lambda_0)^n u_0(0) \\ (1-\mu\lambda_1)^n u_0(1) \\ \vdots \\ (1-\mu\lambda_{p-1})^n u_0(p-1) \end{bmatrix}$$

$$\Downarrow$$

$$\boldsymbol{w}_n = \boldsymbol{w}^0 + \sum_{k=0}^{p-1} (1-\mu\lambda_k)^n u_0(k)\boldsymbol{v}_k$$

- Δηλαδή, το διάνυσμα των συντελεστών είναι ένας γραμμικός συνδυασμός των ιδιοδιανυσμάτων (ονομάζονται modes του φίλτρου). Άρα, ο χρόνος σύγκλισης του διανύσματος $\boldsymbol{w}_n$ στη λύση Wiener-Hopf εξαρτάται από το πιο αργό mode.
- Ο ρυθμός μείωσης κάθε mode είναι: $(1-\mu\lambda_k)^n$
- Για κάθε mode, ορίζουμε τη χρονική σταθερά τ_k ως το χρόνο που απαιτείται ώστε να μειωθεί στο $1/e$ της αρχικής του τιμής:

$$(1-\mu\lambda_k)^{\tau_k} = 1/e \Rightarrow \tau_k = -1/\ln(1-\mu\lambda_k)$$

Μέθοδος μέγιστης καθόδου - Σύγκλιση (7/9)

- Ορίζουμε τη συνολική χρονική σταθερά τ ως το χρόνο που απαιτείται ώστε το πιο αργό mode να μειωθεί στο $1/e$ της αρχικής του τιμής:

$$\tau = \max\left\{\tau_k\right\}$$

- Για μικρές τιμές του μ, δηλαδή όταν $\mu\lambda_k \ll 1$, μπορούμε να γράψουμε:

$$\tau = \max\left\{-1/\ln(1-\mu\lambda_k)\right\} \approx \max\left\{1/(\mu\lambda_k)\right\} = \frac{1}{\mu\lambda_{\min}}$$

- Θέτουμε $\mu = \alpha\left(\frac{2}{\lambda_{max}}\right)$ όπου $0 < \alpha < 1$ (προσέξτε ότι ικανοποιείται η συνθήκη σύγκλισης). Άρα, η συνολική χρονική σταθερά γίνεται:

$$\tau = \frac{1}{2a}\frac{\lambda_{\max}}{\lambda_{\min}} \longrightarrow$$ αριθμός κατάστασης του πίνακα $\boldsymbol{R}_x$

- Συνεπώς, ο ρυθμός σύγκλισης εξαρτάται από τη διασπορά των ιδιοτιμών του πίνακα $\boldsymbol{R}_x$.

Μέθοδος μέγιστης καθόδου - Σύγκλιση (8/9)

- Για το μέσο τετραγωνικό σφάλμα γράφουμε:

$$
\begin{aligned}
\xi(\boldsymbol{w}_n) &= r_d(0) - \boldsymbol{r}_{dx}^H \boldsymbol{w}_n - \boldsymbol{w}_n^H \boldsymbol{r}_{dx} + \boldsymbol{w}_n^H \boldsymbol{R}_x \boldsymbol{w}_n \\
&= r_d(0) - \boldsymbol{r}_{dx}^H \left(\boldsymbol{w}^o + \boldsymbol{c}_n\right) - \left(\boldsymbol{w}^o + \boldsymbol{c}_n\right)^H \boldsymbol{r}_{dx} + \left(\boldsymbol{w}^o + \boldsymbol{c}_n\right)^H \boldsymbol{R}_x \left(\boldsymbol{w}^o + \boldsymbol{c}_n\right) \\
&= r_d(0) - \boldsymbol{r}_{dx}^H \boldsymbol{w}^o - \boldsymbol{r}_{dx}^H \boldsymbol{c}_n - \left(\boldsymbol{w}^o\right)^H \boldsymbol{r}_{dx} - \boldsymbol{c}_n^H \boldsymbol{r}_{dx} + \\
&\qquad + \left(\boldsymbol{w}^o\right)^H \boldsymbol{R}_x \boldsymbol{w}^o + \left(\boldsymbol{w}^o\right)^H \boldsymbol{R}_x \boldsymbol{c}_n + \boldsymbol{c}_n^H \boldsymbol{R}_x \boldsymbol{w}^o + \boldsymbol{c}_n^H \boldsymbol{R}_x \boldsymbol{c}_n \\
&= r_d(0) - \boldsymbol{r}_{dx}^H \boldsymbol{w}^o - \cancel{\boldsymbol{r}_{dx}^H \boldsymbol{c}_n} - \cancel{\left(\boldsymbol{w}^o\right)^H} \boldsymbol{r}_{dx} - \cancel{\boldsymbol{c}_n^H \boldsymbol{r}_{dx}} + \\
&\qquad + \cancel{\left(\boldsymbol{w}^o\right)^H} \boldsymbol{r}_{dx} + \cancel{\boldsymbol{r}_{dx}^H \boldsymbol{c}_n} + \cancel{\boldsymbol{c}_n^H \boldsymbol{r}_{dx}} + \boldsymbol{c}_n^H \boldsymbol{R}_x \boldsymbol{c}_n
\end{aligned}
$$

- Άρα:

$$
\xi(\boldsymbol{w}_n) = \underbrace{r_d(0) - \boldsymbol{r}_{dx}^H \boldsymbol{w}^o}_{\xi_{\min}} + \boldsymbol{c}_n^H \boldsymbol{R}_x \boldsymbol{c}_n \quad \Rightarrow \quad \xi(\boldsymbol{w}_n) = \xi_{\min} + \boldsymbol{c}_n^H \boldsymbol{R}_x \boldsymbol{c}_n
$$

Μέθοδος μέγιστης καθόδου - Σύγκλιση (9/9)

- Τελικά:

$$\xi(\boldsymbol{w}_n) = \xi_{\min} + \boldsymbol{c}_n^H \boldsymbol{R}_x \boldsymbol{c}_n = \xi_{\min} + \boldsymbol{c}_n^H \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H \boldsymbol{c}_n$$

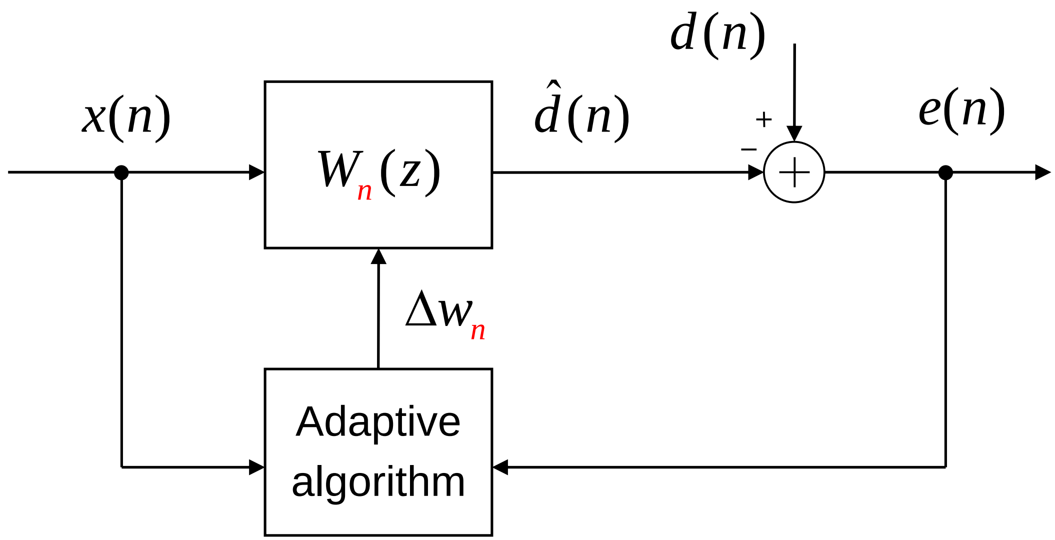
$$= \xi_{\min} + \boldsymbol{u}_n^H \boldsymbol{\Lambda} \boldsymbol{u}_n = \xi_{\min} + \sum_{k=0}^{p-1} \lambda_k \left| u_n(k) \right|^2$$

$$= \xi_{\min} + \sum_{k=0}^{p-1} \lambda_k \left(1 - \mu\lambda_k\right)^{2n} \left| u_0(k) \right|^2$$

- Συνεπώς, αν το μέγεθος βήματος ικανοποιεί τη συνθήκη σύγκλισης, τότε η συνάρτηση κόστους φθίνει εκθετικά στην ελάχιστη τιμή της λύσης Wiener-Hopf.

- Η καμπύλη μεταβολής της συνάρτησης κόστους ως προς n ονομάζεται καμπύλη εκμάθησης (learning curve).

Αλγόριθμος LMS (1/16)



- Είδαμε, ότι ο αλγόριθμος μέγιστης καθόδου χρησιμοποιεί τον αναδρομικό τύπο:

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n - \mu \nabla \xi(n) \quad \Rightarrow \quad \boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu E\left\{ e(n) \boldsymbol{x}^*(n) \right\}$$

- Για από κοινού WSS διαδικασίες $x(n)$ και $d(n)$, το προσαρμοστικό φίλτρο μέγιστης καθόδου συγκλίνει στη λύση των εξισώσεων Wiener-Hopf, δηλαδή $\lim_{n\to\infty} \boldsymbol{w}_n = \boldsymbol{R}_x^{-1} \boldsymbol{r}_{dx}$, αν το μέγεθος βήματος ικανοποιεί τη συνθήκη: $0 < \mu < 2/\lambda_{max}$.

Αλγόριθμος LMS (2/16)

- Στην πράξη, για τον υπολογισμό της ποσότητας $E\{e(n)\boldsymbol{x}^*(n)\}$ χρησιμοποιείται μια εκτίμηση από τα δεδομένα:

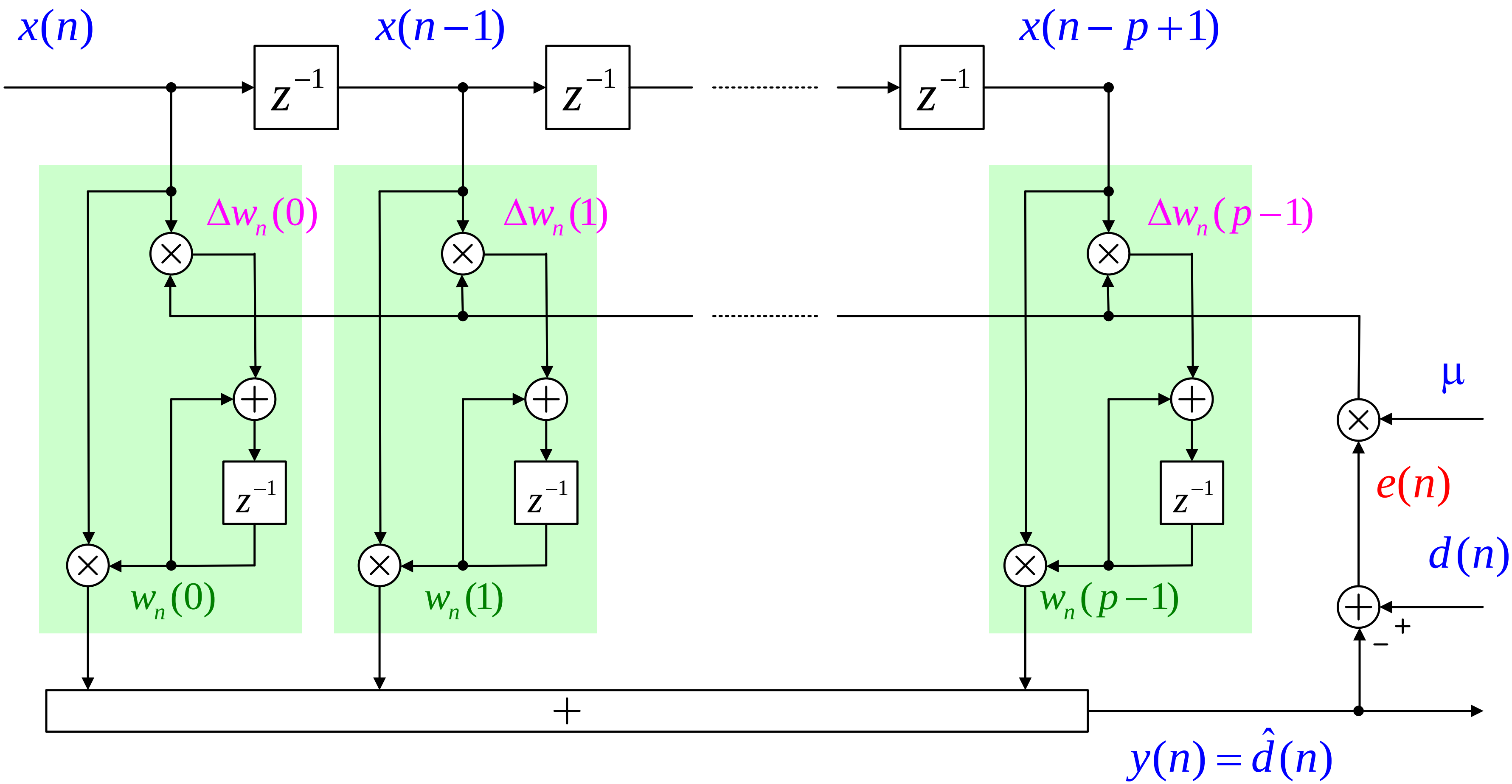
$$\hat{E}\left\{e(n)\boldsymbol{x}^*(n)\right\} = \frac{1}{L}\sum_{l=0}^{L-1} e(n-l)\boldsymbol{x}^*(n-l)$$

- Αν επιλέξουμε $L = 1$, τότε ο αναδρομικός τύπος ανανέωσης των συντελεστών γίνεται:

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu e(n)\boldsymbol{x}^*(n)$$

- Ο παραπάνω τύπος ανανέωσης των συντελεστών του προσαρμοστικού φίλτρου ονομάζεται αλγόριθμος Ελαχίστων Μέσων Τετραγώνων (LMS: least mean squares).

Αλγόριθμος LMS (3/16)



$$w_{n+1}(k)=w_n(k)+\mu e(n)x(n-k) \Rightarrow w_n(k)+\mu\left[d(n)-\underbrace{\sum_{l=0}^{p-1}w_n(l)x(n-l)}_{y(n)}\right]x(n-k)$$

Αλγόριθμος LMS (4/16)

- Γενικά, ο αλγόριθμος LMS συνοψίζεται ως εξής:
  1. Αρχικοποίηση του διανύσματος των συντελεστών: $\boldsymbol{w}_{n=0}$
  2. Ανανέωση των συντελεστών κατά την n-οστή επανάληψη:

$$y(n) = \boldsymbol{w}_n^T \boldsymbol{x}(n)$$

$$e(n) = d(n) - y(n)$$

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu e(n) \boldsymbol{x}^*(n)$$

  3. Επιστροφή στο βήμα 2 για $n = n + 1$.

Αλγόριθμος LMS (5/16)

- Για τη σύγκλιση του LMS, θεωρούμε ότι $x(n)$ και $d(n)$ είναι από κοινού WSS:

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu e(n)\boldsymbol{x}^*(n) \Rightarrow \boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu\left[d(n) - \boldsymbol{w}_n^T \boldsymbol{x}(n)\right]\boldsymbol{x}^*(n)$$

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu\left[d(n)\boldsymbol{x}^*(n)\right] - \mu\left[\boldsymbol{x}(n)\boldsymbol{x}^H(n)\boldsymbol{w}_n\right] \Rightarrow$$

$$E\left\{\boldsymbol{w}_{n+1}\right\} = E\left\{\boldsymbol{w}_n + \mu\left[d(n)\boldsymbol{x}^*(n)\right] - \mu\left[\boldsymbol{x}(n)\boldsymbol{x}^H(n)\boldsymbol{w}_n\right]\right\} \Rightarrow$$

$$E\left\{\boldsymbol{w}_{n+1}\right\} = E\left\{\boldsymbol{w}_n\right\} + \mu E\left\{d(n)\boldsymbol{x}^*(n)\right\} - \mu E\left\{\boldsymbol{x}(n)\boldsymbol{x}^H(n)\boldsymbol{w}_n\right\}$$

- Υπόθεση ανεξαρτησίας: Τα δεδομένα $x(n)$ και το διάνυσμα των συντελεστών $\boldsymbol{w}_n$ είναι στατιστικά ανεξάρτητα.

$$E\left\{\boldsymbol{w}_{n+1}\right\} = E\left\{\boldsymbol{w}_n\right\} + \mu E\left\{d(n)\boldsymbol{x}^*(n)\right\} - \mu E\left\{\boldsymbol{x}(n)\boldsymbol{x}^H(n)\right\} E\left\{\boldsymbol{w}_n\right\} \Rightarrow$$

$$E\left\{\boldsymbol{w}_{n+1}\right\} = E\left\{\boldsymbol{w}_n\right\} + \mu \boldsymbol{r}_{dx} - \mu \boldsymbol{R}_x E\left\{\boldsymbol{w}_n\right\} \Rightarrow$$

$$E\left\{\boldsymbol{w}_{n+1}\right\} = \left(\boldsymbol{I} - \mu \boldsymbol{R}_x\right) E\left\{\boldsymbol{w}_n\right\} + \mu \boldsymbol{r}_{dx}$$

Αλγόριθμος LMS (6/16)

- Συνέχεια:

$$E\{\boldsymbol{w}_{n+1}\} = (I - \mu \boldsymbol{R}_x) E\{\boldsymbol{w}_n\} + \mu \boldsymbol{r}_{dx} \Rightarrow$$

$$E\{\boldsymbol{c}_{n+1} + \boldsymbol{w}^0\} = (I - \mu \boldsymbol{R}_x) E\{\boldsymbol{c}_n + \boldsymbol{w}^0\} + \mu \boldsymbol{r}_{dx} \Rightarrow$$

$$E\{\boldsymbol{c}_{n+1}\} + \boldsymbol{w}^0 = (I - \mu \boldsymbol{R}_x) E\{\boldsymbol{c}_n\} + \boldsymbol{w}^0 - \mu \boldsymbol{R}_x \boldsymbol{w}^0 + \mu \boldsymbol{r}_{dx} \Rightarrow$$

$$\boldsymbol{R}_x \boldsymbol{w}^0 = \boldsymbol{r}_{dx}$$

$$E\{\boldsymbol{c}_{n+1}\} = (I - \mu \boldsymbol{R}_x) E\{\boldsymbol{c}_n\} \Rightarrow$$

$$E\{\boldsymbol{c}_{n+1}\} = \left(I - \mu \boldsymbol{V} \boldsymbol{\Lambda} \boldsymbol{V}^H\right) E\{\boldsymbol{c}_n\} \Rightarrow$$

$$\boldsymbol{V}^H E\{\boldsymbol{c}_{n+1}\} = (I - \mu \boldsymbol{\Lambda}) \boldsymbol{V}^H E\{\boldsymbol{c}_n\} \Rightarrow$$

$$E\{\boldsymbol{u}_{n+1}\} = (I - \mu \boldsymbol{\Lambda}) E\{\boldsymbol{u}_n\} \quad \text{όπου} \quad \boldsymbol{u}_n = \boldsymbol{V}^H \boldsymbol{c}_n$$

- Συνεπώς:

$$E\{\boldsymbol{u}_n\} = (I - \mu \boldsymbol{\Lambda})^n \cancel{E\{\boldsymbol{u}_0\}} \quad \boldsymbol{u}_0$$

Αλγόριθμος LMS (7/16)

- Τελικά:

$$E\{\boldsymbol{u}_n\} = \begin{bmatrix} (1-\mu\lambda_0)^n u_0(0) \\ \vdots \\ (1-\mu\lambda_{p-1})^n u_0(p-1) \end{bmatrix}$$

- Ισχύει η παρακάτω ιδιότητα για τη σύγκλιση:

Για από κοινού WSS διαδικασίες $x(n)$ και $d(n)$, το προσαρμοστικό φίλτρο LMS συγκλίνει ως προς τη μέση τιμή στη λύση των εξισώσεων Wiener-Hopf, δηλαδή

$$\lim_{n\to\infty} E\{\boldsymbol{w}_n\} = \boldsymbol{R}_x^{-1}\boldsymbol{r}_{dx}$$

αν το μέγεθος βήματος ικανοποιεί τη συνθήκη: $0 < \mu < 2/\lambda_{max}$

και ικανοποιείται η υπόθεση ανεξαρτησίας.

Αλγόριθμος LMS (8/16)

- Στην πράξη, ο προσδιορισμός του άνω φράγματος του μ για τη σύγκλιση του φίλτρου ως προς τη μέση τιμή, μπορεί να γίνει ως εξής: Αφού ο πίνακας $\boldsymbol{R}_x$ είναι ερμιτιανός και μη αρνητικά ορισμένος, ισχύει η ιδιότητα:

$$\mathrm{Trace}\left\{\boldsymbol{R}_x\right\} = \sum_{k=0}^{p-1} \lambda_k$$

- Άρα: $$\lambda_{\max} \le \sum_{k=0}^{p-1} \lambda_k \Rightarrow \lambda_{\max} \le \mathrm{Trace}\left\{\boldsymbol{R}_x\right\} \Rightarrow \lambda_{\max} \le \sum_{k=0}^{p-1} r_x(0) = p r_x(0)$$

$$\Rightarrow \lambda_{\max} \le pE\left\{|x(n)|^2\right\} = p\underbrace{\frac{1}{N}\sum_{k=0}^{N-1}|x(n-k)|^2}_{\hat{E}\left\{|x(n)|^2\right\}}$$

- Συνεπώς η συνθήκη σύγκλισης γράφεται: $0 < \mu < \dfrac{2}{p\hat{E}\{|x(n)|^2\}}$

Αλγόριθμος LMS (9/16)

- Το σφάλμα του αλγορίθμου LMS κάθε χρονική στιγμή είναι:

$$e(n) = d(n) - \boldsymbol{w}_n^T \boldsymbol{x}(n) = d(n) - \left[\boldsymbol{w}^o + \boldsymbol{c_n}\right]^T \boldsymbol{x}(n)$$

$$= \underbrace{d(n) - \left(\boldsymbol{w}^o\right)^T \boldsymbol{x}(n)}_{e_{\min}(n)} - \boldsymbol{c}_{\boldsymbol{n}}^T \boldsymbol{x}(n) = e_{\min}(n) - \boldsymbol{c}_{\boldsymbol{n}}^T \boldsymbol{x}(n)$$

- Θεωρώντας ότι το φίλτρο έχει συγκλίνει, δηλαδή $E\{\boldsymbol{c}_n\} = 0$, η συνάρτηση κόστους γράφεται:

$$\xi(n) = E\left\{\left|e(n)\right|^2\right\} = \xi_{\min} + \xi_{\mathrm{ex}}(n)$$

- Σημειώνεται ότι το $\xi_{ex}(n)$ ονομάζεται πλεονάζων μέσο τετραγωνικό σφάλμα (Excess MSE or Misadjustment).

Αλγόριθμος LMS (10/16)

- Ισχύει η παρακάτω ιδιότητα:

Το MSE $\xi(n)$ συγκλίνει σε μια τιμή σταθερής κατάστασης (steady state)

$$\xi(\infty) = \xi_{\min} + \xi_{ex}(\infty) = \xi_{\min} \frac{1}{1 - \mu \sum_{k=0}^{p-1} \frac{\lambda_k}{2 - \mu \lambda_k}}$$

και ο αλγόριθμος LMS λέγεται ότι συγκλίνει ως προς τη μέση τετραγωνική τιμή αν το μέγεθος βήματος ικανοποιεί τις συνθήκες:

$$0 < \mu < \frac{2}{\lambda_{\max}} \qquad \text{και} \qquad \mu \sum_{k=0}^{p-1} \frac{\lambda_k}{2 - \mu \lambda_k} < 1$$

και ικανοποιείται η υπόθεση ανεξαρτησίας.

Αλγόριθμος LMS (11/16)

- Λύνουμε ως προς $\xi_{ex}(\infty)$:

$$\xi_{\text{ex}}(\infty) = \xi_{\min}\left(\mu\sum_{k=0}^{p-1}\frac{\lambda_k}{2-\mu\lambda_k}\right)\Big/\left(1-\mu\sum_{k=0}^{p-1}\frac{\lambda_k}{2-\mu\lambda_k}\right)$$

- Γενικά, στην πράξη προκύπτει $\mu \ll 2/\lambda_{max}$, οπότε η δεύτερη συνθήκη σύγκλισης ως προς τη μέση τετραγωνική τιμή γίνεται:

$$\mu\sum_{k=0}^{p}\frac{\lambda_k}{\underbrace{2-\mu\lambda_k}_{2-\mu\lambda_k \simeq 2}} < 1 \Rightarrow \mu\sum_{k=0}^{p}\frac{\lambda_k}{2} < 1 \Rightarrow \mu < \frac{1}{\frac{1}{2}\text{Trace}\left\{\boldsymbol{R}_x\right\}} = \frac{2}{\text{Trace}\left\{\boldsymbol{R}_x\right\}}$$

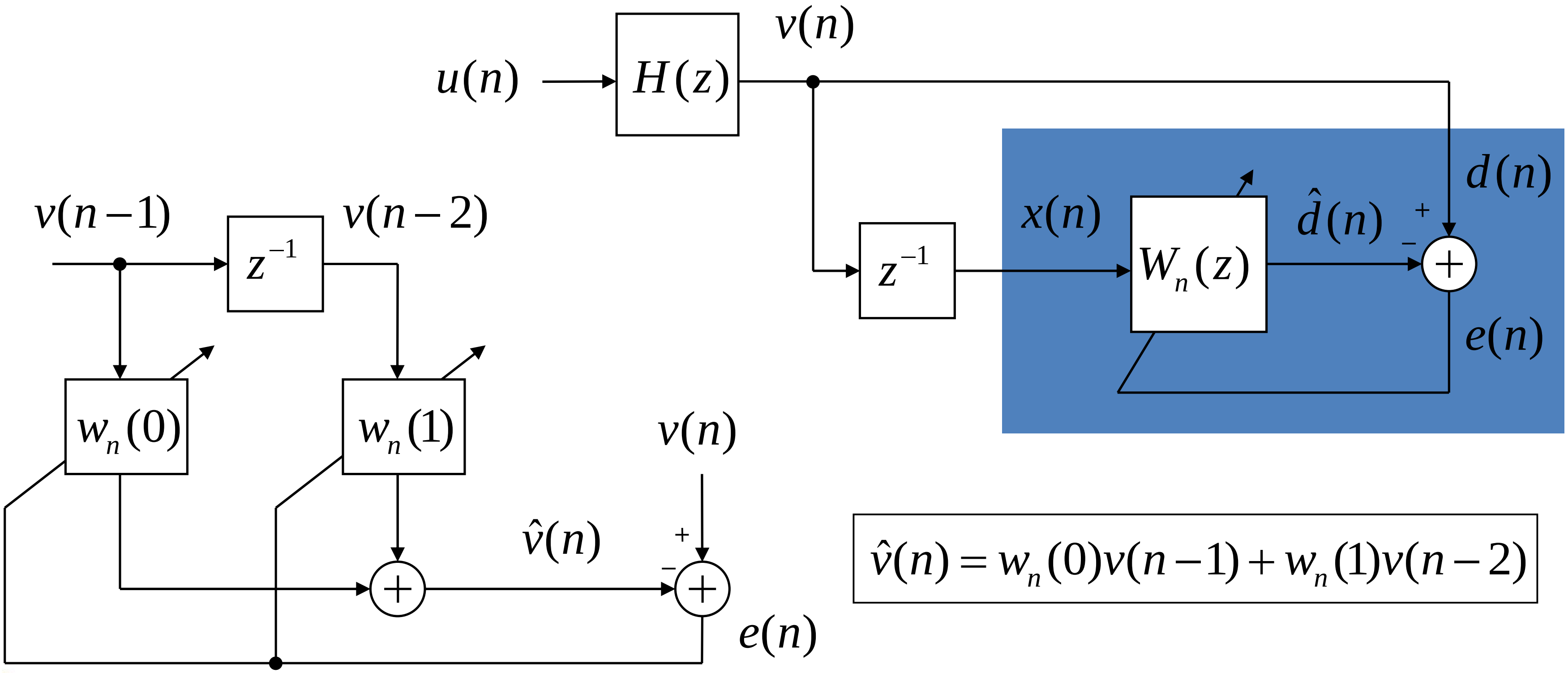
- Επιπλέον:

$$\xi(\infty) = \xi_{\min}\frac{1}{1-\mu\frac{1}{2}\text{Trace}\left\{\boldsymbol{R}_x\right\}} \qquad \xi_{\text{ex}}(\infty) = \frac{\mu}{2}\xi_{\min}\text{Trace}\left\{\boldsymbol{R}_x\right\}$$

- Ορίζουμε το κανονικοποιημένο πλεονάζον MSE σταθερής κατάστασης και το ονομάζουμε misadjustment: $\mathrm{M} = \xi_{ex}(\infty)/\xi_{min}$

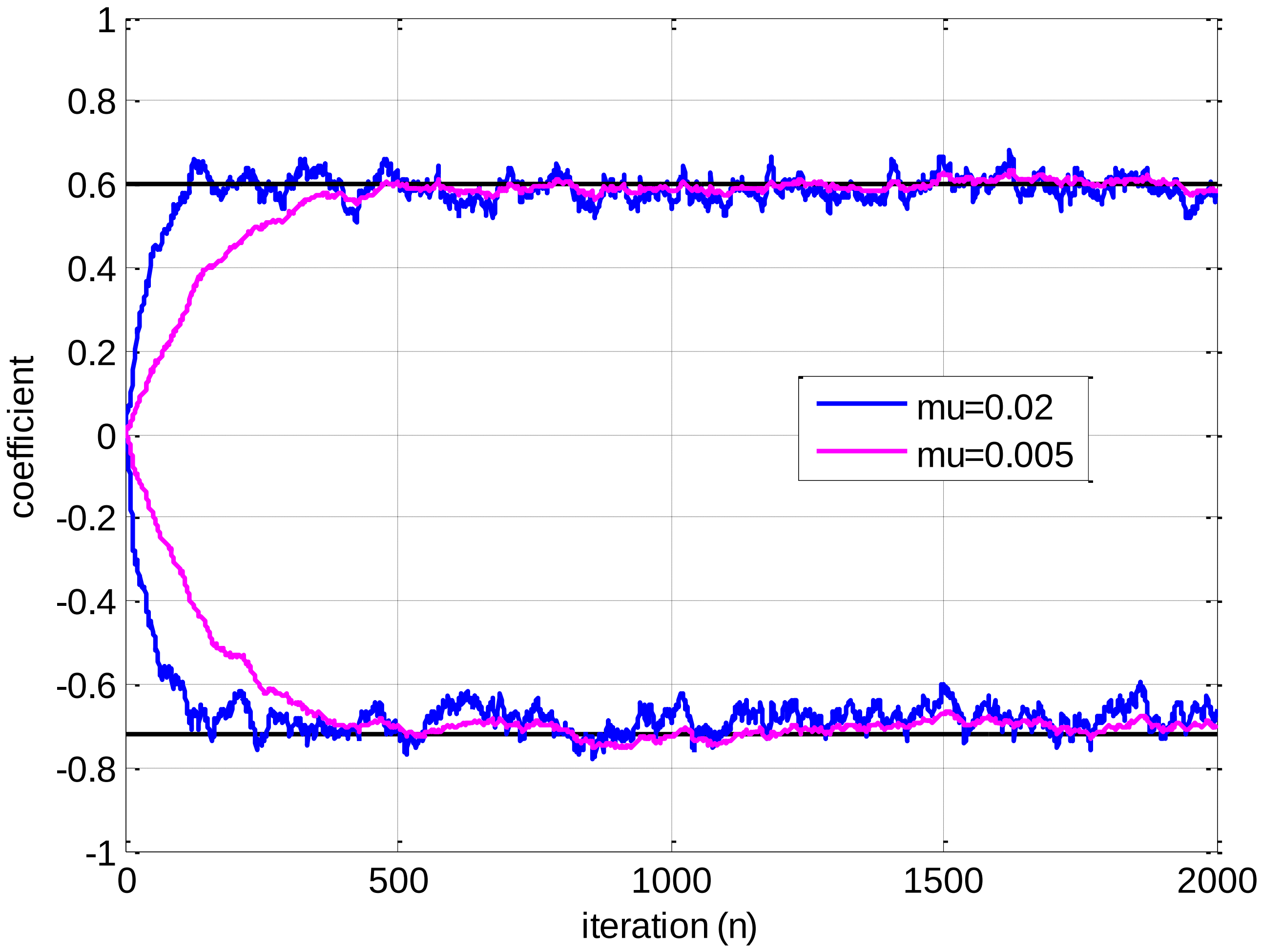
Αλγόριθμος LMS (12/16)

- Το παράδειγμα της γραμμική πρόβλεψη (Linear Prediction). Η διαδικασία $v(n)$ είναι μια AR(2) διαδικασία με την ακόλουθη εξίσωση διαφορών, όπου $u(n)$ είναι λευκός θόρυβος με μοναδιαία διασπορά: $v(n) = 0.6v(n-1) - 0.72v(n-2) + u(n)$

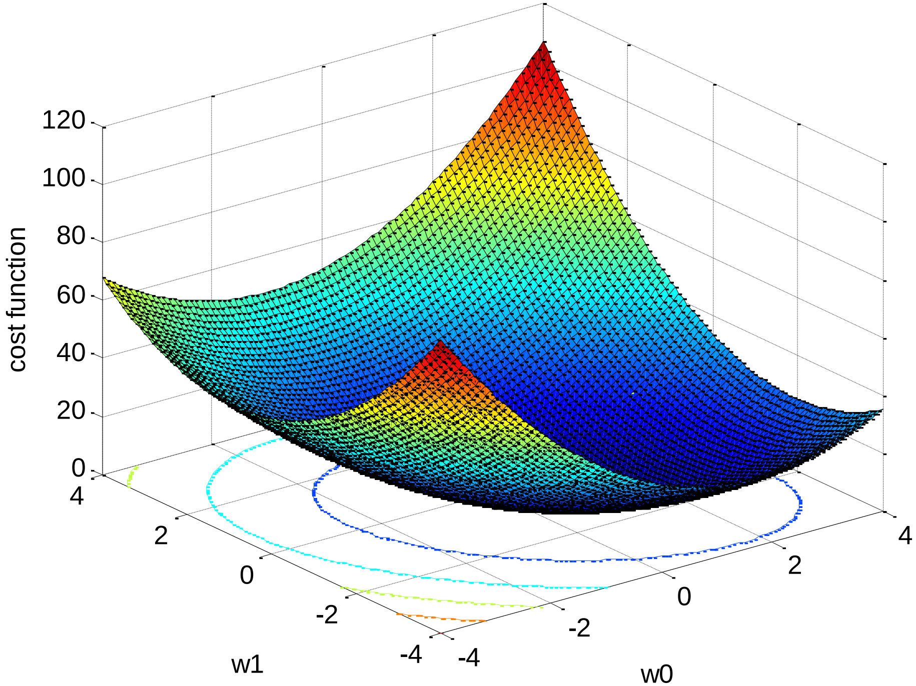


$$\hat{v}(n) = w_n(0)v(n-1) + w_n(1)v(n-2)$$

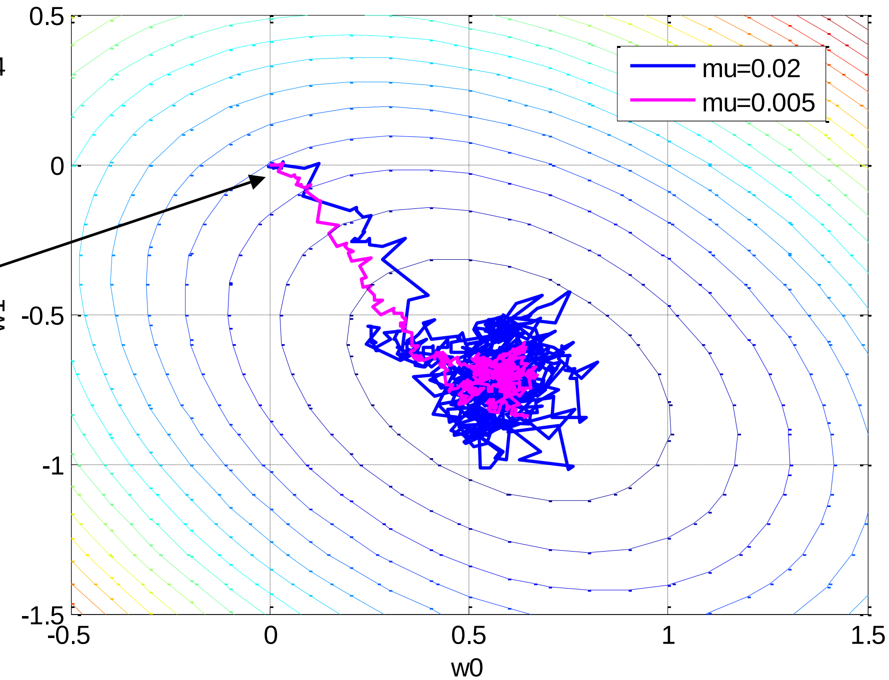
Αλγόριθμος LMS (13/16)



Αλγόριθμος LMS (14/16)

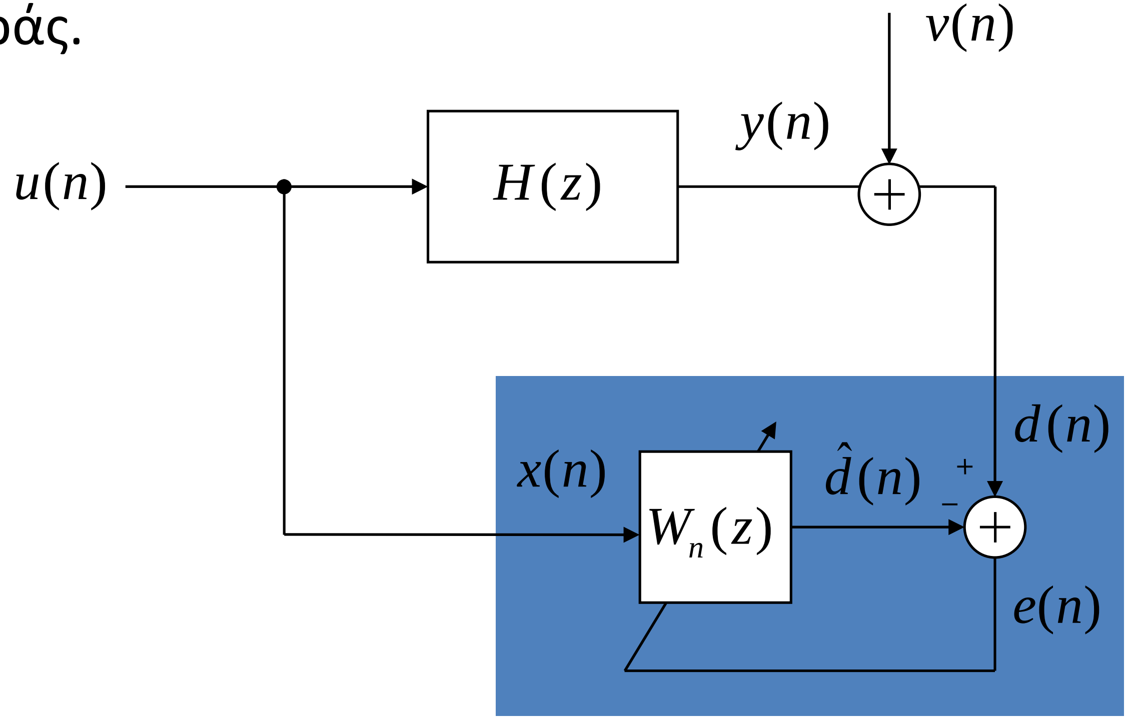


Αρχικοποίηση
των συντελεστών

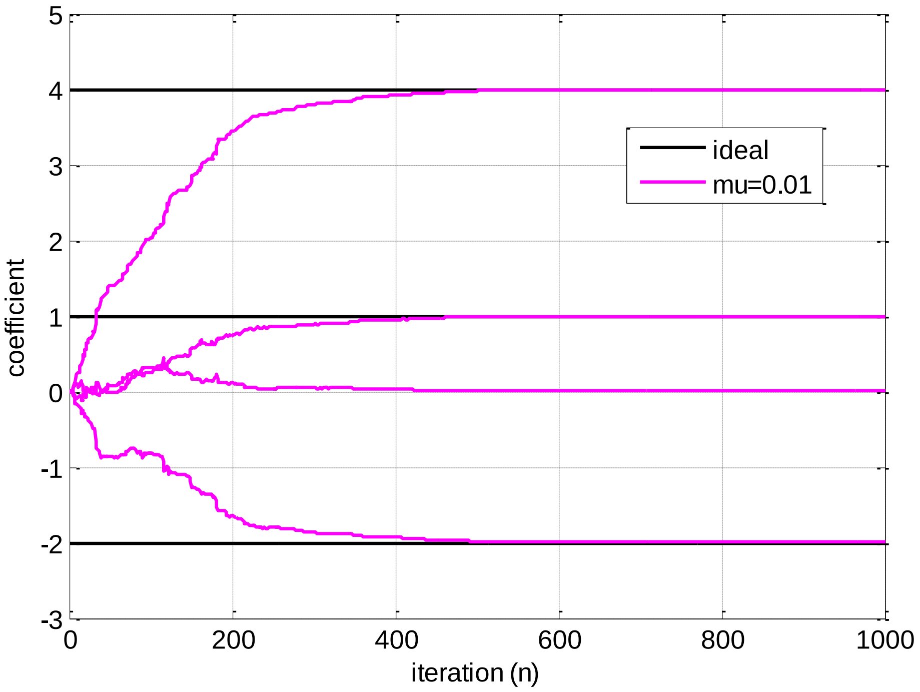


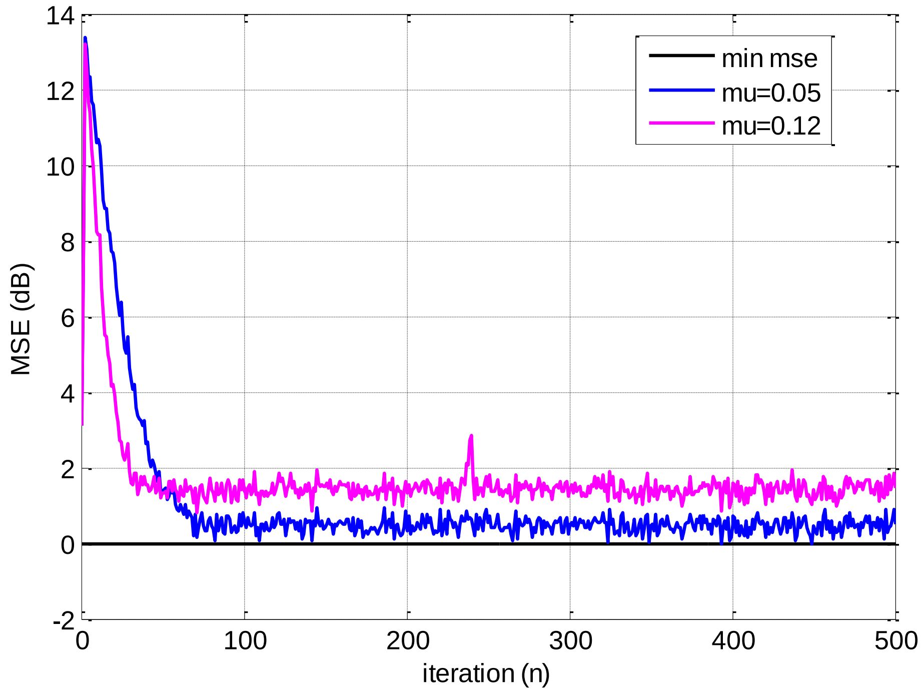
Αλγόριθμος LMS (15/16)

- Παράδειγμα: Αναγνώριση Συστήματος (System Identification). Το άγνωστο σύστημα έχει συνάρτηση μεταφοράς $H(z) = 1 - 2z^{-1} + 4z^{-2}$. Το σήμα $u(n)$ είναι διαδικασία λευκού θορύβου μοναδιαίας διασποράς. Το φίλτρο LMS έχει τέσσερις συντελεστές. Ο προσθετικός θόρυβος $v(n)$ είναι λευκός θόρυβος μοναδιαίας διασποράς.



Αλγόριθμος LMS (16/16)





Αλγόριθμος NLMS (1/3)

- Για τη σύγκλιση του αλγορίθμου LMS, η παράμετρος μ (μέγεθος βήματος) πρέπει να ικανοποιεί τις συνθήκες:
  - σύγκλιση ως προς τη μέση τιμή: $0 < \mu < 2/\lambda_{max}$
  - σύγκλιση ως προς τη μέση τετραγωνική τιμή:

$$\mu \sum_{k=0}^{p-1} \frac{\lambda_k}{2-\mu\lambda_k} < 1 \quad \xrightarrow{\mu \ll 2/\lambda_{\max}} \quad \mu < \frac{2}{\mathrm{Trace}\{\boldsymbol{R}_x\}}$$

- Στην πράξη ο υπολογισμός του πίνακα $\boldsymbol{R}_x$ γίνεται μέσω εκτίμησης της στιγμιαίας ενέργειας του σήματος $x(n)$:

$$\mathrm{Trace}\{\boldsymbol{R}_x\} = pE\left\{|x(n)|^2\right\} \approx p\left[\underbrace{\frac{1}{N}\sum_{k=0}^{N-1}|x(n-k)|^2}_{\hat{E}\left\{|x(n)|^2\right\}}\right] \overset{N=p}{=} \boldsymbol{x}^H(n)\boldsymbol{x}(n) \quad \Longrightarrow$$

$$0 < \mu < \frac{2}{\boldsymbol{x}^H(n)\boldsymbol{x}(n)}$$

Αλγόριθμος NLMS (2/3)

- Ορίζουμε το (χρονικά μεταβαλλόμενο) μέγεθος βήματος ως:

$$\mu(n) = \frac{\beta}{\boldsymbol{x}^H(n)\boldsymbol{x}(n)} = \frac{\beta}{\left\|\boldsymbol{x}(n)\right\|^2}$$

- Η παράμετρος β ονομάζεται κανονικοποιημένο μέγεθος βήματος (normalized step size) και ισχύει $0 < \beta < 2$.

- Ο αναδρομικός τύπος ανανέωσης των συντελεστών γίνεται:

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \beta e(n) \frac{\boldsymbol{x}^*(n)}{\left\|\boldsymbol{x}(n)\right\|^2}$$

- Ο παραπάνω τύπος ανανέωσης των συντελεστών του προσαρμοστικού φίλτρου ονομάζεται αλγόριθμος NLMS: normalized least mean squares. Για τη σύγκλιση ως προς τη μέση τετραγωνική τιμή αποδεικνύεται ότι αρκεί: $0 < \beta < 2$

Αλγόριθμος NLMS (3/3)

- Συγκρίνοντας τους αναδρομικούς τύπους LMS και NLMS παρατηρούμε ότι:

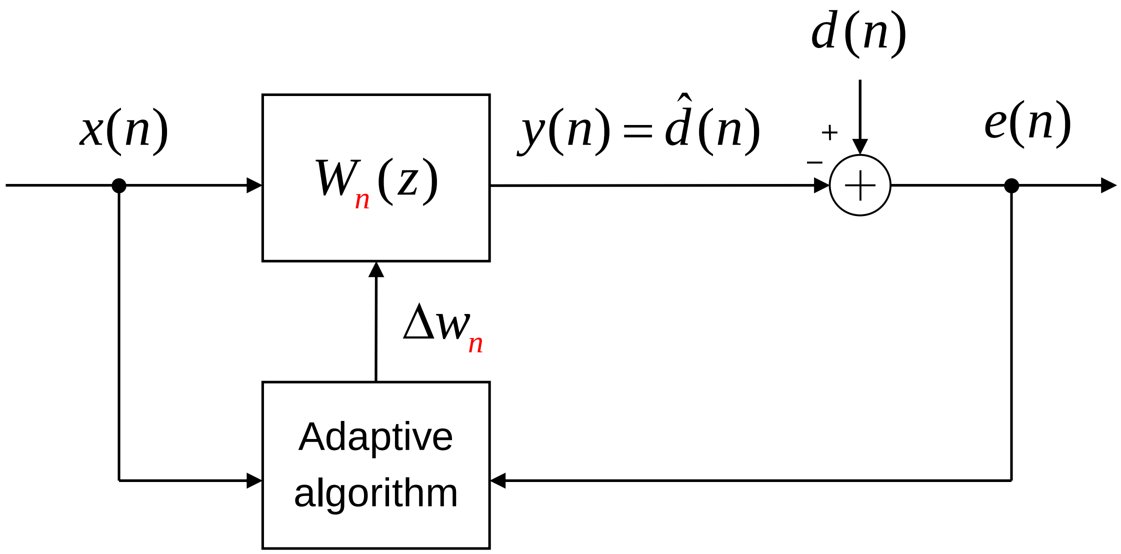
LMS: $\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu e(n)\boldsymbol{x}^*(n)$ NLMS: $\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \beta e(n)\dfrac{\boldsymbol{x}^*(n)}{\left\|\boldsymbol{x}(n)\right\|^2}$

- Στον αλγόριθμο LMS, ο διορθωτικός όρος είναι μια (αμερόληπτη) εκτίμηση του $\nabla\xi(n)$. Συνεπώς η εκτίμηση $\nabla\hat{\xi}(n)$ περιέχει θόρυβο. Όταν, οι τιμές του διανύσματος $x(n)$ είναι μεγάλες, τότε εμφανίζεται ενίσχυση του θορύβου στην στιγμιαία εκτίμηση $\nabla\hat{\xi}(n)$.

- Ο αλγόριθμος NLMS μειώνει την ενίσχυση του θορύβου μέσω του παράγοντα κανονικοποίησης $\|\boldsymbol{x}(n)\|^2$. Ταυτόχρονα, όμως αντιμετωπίζει το ίδιο πρόβλημα όταν η τιμή $\|\boldsymbol{x}(n)\|^2$ είναι πολύ μικρή. Για το λόγο αυτό, τροποποιούμε τον αναδρομικό τύπο ως εξής:

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \beta e(n)\frac{\boldsymbol{x}^*(n)}{\varepsilon + \left\|\boldsymbol{x}(n)\right\|^2}$$ (όπου ε μια μικρή θετική σταθερά)

Αλγόριθμος RLS (1/17)



- Θεωρούμε ένα FIR προσαρμοστικό φίλτρο με p συντελεστές, οι οποίοι τη χρονική στιγμή n ελαχιστοποιούν το σφάλμα των εκθετικά ζυγισμένων τετραγώνων: $E(n) = \sum_{\{i=0\}}^{n} 1^{n-i}|e(i)|^2$

- Σημειώνεται πως το $0 < i \leq 1$ ονομάζεται εκθετικός παράγοντας λήθης και το στιγμιαίο σφάλμα $e(i)$ ορίζεται ως:

$$e(i) = d(i) - y(i) = d(i) - \boldsymbol{w}_n^T \boldsymbol{x}(i)$$

Αλγόριθμος RLS (2/17)

- Υπολογισμός βέλτιστων συντελεστών:

$$\min_{w_n(k)} \mathrm{E}(n) = \min_{w_n(k)} \sum_{i=0}^{n} 1^{\,n-i} \left|e(i)\right|^2 \quad \Longrightarrow \quad \frac{\partial \mathrm{E}(n)}{\partial w_n^*(k)} = 0 \quad \text{για } k = 0,1,\ldots,p-1$$

συνάρτηση κόστους (cost function)

- Το σφάλμα γράφεται: $e(i) = d(i) - \boldsymbol{w}_n^T \boldsymbol{x}(i) = d(i) - \sum_{l=0}^{p-1} w_n(l) x(i-l)$

- Άρα:

$$\frac{\partial \mathrm{E}(n)}{\partial w_n^*(k)} = 0 \Rightarrow \sum_{i=0}^{n} 1^{\,n-i} \frac{\partial \left|e(i)\right|^2}{\partial w_n^*(k)} = 0 \Rightarrow \sum_{i=0}^{n} 1^{\,n-i} e(i) \frac{\partial e^*(i)}{\partial w_n^*(k)} = 0$$

$$\Rightarrow \sum_{i=0}^{n} 1^{\,n-i} e(i) \left[-x^*(i-k) \right] = 0$$

$$\Rightarrow \sum_{i=0}^{n} 1^{\,n-i} e(i) x^*(i-k) = 0 \quad \text{για} \quad k = 0,1,\ldots,p-1$$

Αλγόριθμος RLS (3/17)

- Συνέχεια (για $k = 0,1,\ldots,p-1$):

$$\sum_{i=0}^{n} \mathrm{l}^{\,n-i} e(i) x^*(i-k) = 0 \Rightarrow$$

$$\sum_{i=0}^{n} \mathrm{l}^{\,n-i} \left[d(i) - \sum_{l=0}^{p-1} w_n(l) x(i-l) \right] x^*(i-k) = 0 \Rightarrow$$

$$\sum_{i=0}^{n} \mathrm{l}^{\,n-i} d(i) x^*(i-k) - \sum_{i=0}^{n} \mathrm{l}^{\,n-i} \left[\sum_{l=0}^{p-1} w_n(l) x(i-l) \right] x^*(i-k) = 0 \Rightarrow$$

$$\sum_{l=0}^{p-1} w_n(l) \underbrace{\sum_{i=0}^{n} \mathrm{l}^{\,n-i} x(i-l) x^*(i-k)}_{r_x(k,l;n)} = \underbrace{\sum_{i=0}^{n} \mathrm{l}^{\,n-i} d(i) x^*(i-k)}_{r_{dx}(k;n)}$$

ντετερμινιστική αυτοσυσχέτιση

ντετερμινιστική ετεροσυσχέτιση

Αλγόριθμος RLS (4/17)

- Τελικά, για κάθε χρονική στιγμή n, καταλήγουμε σε ένα σύστημα γραμμικών εξισώσεων ως προς τους συντελεστές, το οποίο είναι γνωστό ως ντετερμινιστικές κανονικές εξισώσεις:

$$\sum_{l=0}^{p-1} w_n(l) r_x(k,l;n) = r_{dx}(k;n) \quad \text{για} \quad k = 0,1,\ldots,p-1$$

$$\Downarrow$$

$$\boldsymbol{R}_x(n)\boldsymbol{w}_n = \boldsymbol{r}_{dx}(n)$$

όπου:

$$\boldsymbol{R}_x(n) = \sum_{i=0}^{n} 1^{\,n-i} \boldsymbol{x}^*(i)\boldsymbol{x}^T(i)$$

$$\boldsymbol{r}_{dx}(n) = \sum_{i=0}^{n} 1^{\,n-i} d(i)\boldsymbol{x}^*(i)$$

Αλγόριθμος RLS (5/17)

- Διερεύνηση της συνάρτησης κόστους:

$$\mathrm{E}(\boldsymbol{w}_n)=\sum_{i=0}^{n} l^{\,n-i}\left|e(i)\right|^2=\sum_{i=0}^{n} l^{\,n-i}e(i)e^*(i)=\sum_{i=0}^{n} l^{\,n-i}e(i)\left[d^*(i)-\sum_{k=0}^{p-1}w_n^*(k)x^*(i-k)\right]$$

$$=\sum_{i=0}^{n} l^{\,n-i}e(i)d^*(i)-\sum_{k=0}^{p-1}w_n^*(k)\sum_{i=0}^{n} l^{\,n-i}e(i)x^*(i-k)$$

$$=\sum_{i=0}^{n} l^{\,n-i}\left[d(i)-\boldsymbol{x}^T(i)\boldsymbol{w}_n\right]d^*(i)-\sum_{k=0}^{p-1}w_n^*(k)\sum_{i=0}^{n} l^{\,n-i}\left[d(i)-\boldsymbol{x}(i)^T\boldsymbol{w}_n\right]x^*(i-k)$$

$$=\sum_{i=0}^{n} l^{\,n-i}d(i)d^*(i)-\sum_{i=0}^{n} l^{\,n-i}\left[\boldsymbol{x}^T(i)\boldsymbol{w}_n\right]d^*(i)-\sum_{k=0}^{p-1}w_n^*(k)\sum_{i=0}^{n} l^{\,n-i}d(i)x^*(i-k)+$$

$$+\sum_{k=0}^{p-1}w_n^*(k)\sum_{i=0}^{n} l^{\,n-i}\left[\boldsymbol{x}(i)^T\boldsymbol{w}_n\right]x^*(i-k)$$

$$=\sum_{i=0}^{n} l^{\,n-i}d(i)d^*(i)-\left[\sum_{i=0}^{n} l^{\,n-i}\boldsymbol{x}^T(i)d^*(i)\right]\boldsymbol{w}_n-\boldsymbol{w}_n^H\sum_{i=0}^{n} l^{\,n-i}d(i)\boldsymbol{x}^*(i)+$$

$$+\boldsymbol{w}_n^H\left[\sum_{i=0}^{n} l^{\,n-i}\boldsymbol{x}^*(i)\boldsymbol{x}^T(i)\right]\boldsymbol{w}_n$$

$$=\left\|\boldsymbol{d}(n)\right\|_l^2-\boldsymbol{r}_{dx}^H(n)\boldsymbol{w}_n-\boldsymbol{w}_n^H\boldsymbol{r}_{dx}(n)+\boldsymbol{w}_n^H\boldsymbol{R}_x(n)\boldsymbol{w}_n$$

Αλγόριθμος RLS (6/17)

- Όταν οι συντελεστές $\boldsymbol{w}_n$ ικανοποιούν τις ντετερμινιστικές κανονικές εξισώσεις, δηλαδή $\boldsymbol{w}_n = \boldsymbol{R}_x^{-1}(n)\boldsymbol{r}_{dx}(n)$, τότε:

$$\sum_{i=0}^{n} \lambda^{n-i} e(i) x^*(i-k) = 0 \quad \Rightarrow \quad \mathrm{E}_{\min} = \left\| \boldsymbol{d}(n) \right\|_\lambda^2 - \boldsymbol{r}_{dx}^H(n)\boldsymbol{w}_n$$

$$\text{για} \quad k = 0, 1, \ldots, p-1$$

- Οι ντετερμινιστικές κανονικές εξισώσεις εξαρτώνται από το n. Αντί να λύνουμε τις εξισώσεις απευθείας κάθε χρονική στιγμή, δηλαδή $\boldsymbol{w}_n = \boldsymbol{R}_x^{-1}(n)\boldsymbol{r}_{dx}(n)$, θα αναπτύξουμε μια αναδρομική μέθοδο:

$$\boldsymbol{w}_n = \boldsymbol{w}_{n-1} + \Delta \boldsymbol{w}_{n-1}$$

Αλγόριθμος RLS (7/17)

- Ισχύει:

$$\boldsymbol{r}_{dx}(n) = \sum_{i=0}^{n} \lambda^{n-i} d(i)\boldsymbol{x}^*(i) = \sum_{i=0}^{n-1} \lambda^{n-i} d(i)\boldsymbol{x}^*(i) + d(n)\boldsymbol{x}^*(n)$$

$$= \left(\lambda\lambda^{-1}\right)\sum_{i=0}^{n-1} \lambda^{n-i} d(i)\boldsymbol{x}^*(i) + d(n)\boldsymbol{x}^*(n)$$

$$= \lambda \sum_{i=0}^{n-1} \lambda^{(n-1)-i} d(i)\boldsymbol{x}^*(i) + d(n)\boldsymbol{x}^*(n)$$

- Άρα:

$$\boldsymbol{r}_{dx}(n) = \lambda\,\boldsymbol{r}_{dx}(n-1) + d(n)\boldsymbol{x}^*(n)$$

αναδρομικός τύπος για το διάνυσμα ετεροσυσχέτισης

- Ομοίως:

$$\boldsymbol{R}_x(n) = \lambda\,\boldsymbol{R}_x(n-1) + \boldsymbol{x}^*(n)\boldsymbol{x}^T(n)$$

αναδρομικός τύπος για τον πίνακα αυτοσυσχέτισης

Αλγόριθμος RLS (8/17)

- Χρησιμοποιούμε την ταυτότητα του Woodbury:

$$\left(\boldsymbol{A}+\boldsymbol{u}\boldsymbol{v}^H\right)^{-1}=\boldsymbol{A}^{-1}-\frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^H\boldsymbol{A}^{-1}}{1+\boldsymbol{v}^H\boldsymbol{A}^{-1}\boldsymbol{u}}$$

και θέτουμε:

$$\boldsymbol{A}=\text{l}\,\boldsymbol{R}_x(n-1)$$

$$\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{x}^*(n)$$

$$\Downarrow$$

$$\Big(\underbrace{\text{l}\,\boldsymbol{R}_x(n-1)+\boldsymbol{x}^*(n)\boldsymbol{x}^T(n)}_{\boldsymbol{R}_x(n)}\Big)^{-1}=\text{l}^{-1}\boldsymbol{R}_x^{-1}(n-1)-\frac{\left[\text{l}^{-1}\boldsymbol{R}_x^{-1}(n-1)\right]\left[\boldsymbol{x}^*(n)\boldsymbol{x}^T(n)\right]\left[\text{l}^{-1}\boldsymbol{R}_x^{-1}(n-1)\right]}{1+\boldsymbol{x}^T(n)\left[\text{l}^{-1}\boldsymbol{R}_x^{-1}(n-1)\right]\boldsymbol{x}^*(n)}$$

$$\Downarrow$$

$$\boldsymbol{R}_x^{-1}(n)=\text{l}^{-1}\boldsymbol{R}_x^{-1}(n-1)-\frac{\text{l}^{-2}\boldsymbol{R}_x^{-1}(n-1)\boldsymbol{x}^*(n)\boldsymbol{x}^T(n)\boldsymbol{R}_x^{-1}(n-1)}{1+\text{l}^{-1}\boldsymbol{x}^T(n)\boldsymbol{R}_x^{-1}(n-1)\boldsymbol{x}^*(n)} \qquad (*)$$

αναδρομικός τύπος για τον αντίστροφο πίνακα αυτοσυσχέτισης

Αλγόριθμος RLS (9/17)

- Στη συνέχεια θέτουμε: $\boldsymbol{P}(n) = \boldsymbol{R}_x^{-1}(n)$

$$(*) \quad \Rightarrow \quad \boldsymbol{P}(n) = l^{-1}\boldsymbol{P}(n-1) - l^{-1}\frac{l^{-1}\boldsymbol{P}(n-1)\boldsymbol{x}^*(n)\boldsymbol{x}^T(n)\boldsymbol{P}(n-1)}{1 + l^{-1}\boldsymbol{x}^T(n)\boldsymbol{P}(n-1)\boldsymbol{x}^*(n)}$$

- Επιπλέον, ορίζουμε το διάνυσμα κέρδους (gain vector):

$$\boldsymbol{g}(n) = \frac{l^{-1}\boldsymbol{P}(n-1)\boldsymbol{x}^*(n)}{1 + \boldsymbol{x}^T(n)\left[l^{-1}\boldsymbol{P}(n-1)\boldsymbol{x}^*(n)\right]} \qquad (**)$$

$$(*) \quad \Rightarrow \quad \boldsymbol{P}(n) = l^{-1}\left[\boldsymbol{P}(n-1) - \boldsymbol{g}(n)\boldsymbol{x}^T(n)\boldsymbol{P}(n-1)\right]$$

$$(**) \quad \Rightarrow \quad \boldsymbol{g}(n) = l^{-1}\left[\boldsymbol{P}(n-1) - \boldsymbol{g}(n)\boldsymbol{x}^T(n)\boldsymbol{P}(n-1)\right]\boldsymbol{x}^*(n) = \boldsymbol{P}(n)\boldsymbol{x}^*(n)$$

Αλγόριθμος RLS (10/17)

- Το αρχικό σύστημα των ντετερμινιστικών κανονικών εξισώσεων:

$$\boldsymbol{R}_x(n)\boldsymbol{w}_n = \boldsymbol{r}_{dx}(n)$$

- Με κατάλληλα βήματα, οδηγηθήκαμε, στο παρακάτω σύστημα:

$$\boldsymbol{g}(n) = \boldsymbol{P}(n)\boldsymbol{x}^*(n) \quad \Rightarrow \quad \boldsymbol{R}_x(n)\boldsymbol{g}(n) = \boldsymbol{x}^*(n)$$

- Ο αναδρομικός τύπος ανανέωσης των συντελεστών γράφεται:

$$\boldsymbol{w}_n = \boldsymbol{R}_x^{-1}(n)\boldsymbol{r}_{dx}(n) = \boldsymbol{P}(n)\boldsymbol{r}_{dx}(n) = \boldsymbol{P}(n)\left[l\, \boldsymbol{r}_{dx}(n-1) + d(n)\boldsymbol{x}^*(n) \right]$$

$$= l\, \boldsymbol{P}(n)\boldsymbol{r}_{dx}(n-1) + d(n)\boldsymbol{P}(n)\boldsymbol{x}^*(n)$$

$$= l\left[l^{-1}\left[\boldsymbol{P}(n-1) - \boldsymbol{g}(n)\boldsymbol{x}^T(n)\boldsymbol{P}(n-1) \right] \right]\boldsymbol{r}_{dx}(n-1) + d(n)\boldsymbol{g}(n)$$

$$= \boldsymbol{P}(n-1)\boldsymbol{r}_{dx}(n-1) - \boldsymbol{g}(n)\boldsymbol{x}^T(n)\boldsymbol{P}(n-1)\boldsymbol{r}_{dx}(n-1) + d(n)\boldsymbol{g}(n)$$

$$= \boldsymbol{w}_{n-1} - \boldsymbol{g}(n)\boldsymbol{x}^T(n)\boldsymbol{w}_{n-1} + d(n)\boldsymbol{g}(n)$$

Αλγόριθμος RLS (11/17)

- Τελικά: $$\boldsymbol{w}_n = \boldsymbol{w}_{n-1} - \boldsymbol{g}(n)\boldsymbol{x}^T(n)\boldsymbol{w}_{n-1} + d(n)\boldsymbol{g}(n)$$
$$= \boldsymbol{w}_{n-1} + \boldsymbol{g}(n)\left[d(n) - \boldsymbol{x}^T(n)\boldsymbol{w}_{n-1}\right]$$
$$= \boldsymbol{w}_{n-1} + \boldsymbol{g}(n)\underbrace{\left[d(n) - \boldsymbol{w}_{n-1}^T\boldsymbol{x}(n)\right]}_{a(n)\,:\,\text{scalar}} \Rightarrow \boldsymbol{w}_n = \boldsymbol{w}_{n-1} + a(n)\boldsymbol{g}(n)$$

- Ο παραπάνω τύπος ανανέωσης των συντελεστών του προσαρμοστικού φίλτρου ονομάζεται αλγόριθμος Αναδρομικών Ελαχίστων Τετραγώνων (RLS: recursive least squares). Γενικά, αναφερόμαστε στον εκθετικά ζυγισμένο αλγόριθμο RLS. Ειδικά, όταν $l = 1$ ο αλγόριθμος RLS καλείται αλγόριθμος αυξανομένου παραθύρου (growing window RLS).

- Η ποσότητα $a(n) = d(n) - \boldsymbol{w}_{n-1}^T\boldsymbol{x}(n)$ ονομάζεται a priori σφάλμα, ενώ η $\mathrm{e}(n) = d(n) - \boldsymbol{w}_n^T\boldsymbol{x}(n)$ ονομάζεται a posteriori σφάλμα.

Αλγόριθμος RLS (12/17)

- Γενικά, ο αλγόριθμος RLS συνοψίζεται ως εξής:
    1. Αρχικοποίηση του διανύσματος των συντελεστών: $\boldsymbol{w}_{n=0} = \boldsymbol{0}$
    2. Αρχικοποίηση του πίνακα $\boldsymbol{P}(n)$: $\boldsymbol{P}(n = 0) = \delta^{-1}\boldsymbol{I}$
    3. Ανανέωση των συντελεστών κατά την επανάληψη $n = 1, 2, \ldots$:

$$\boldsymbol{z}(n) = \boldsymbol{P}(n-1)\boldsymbol{x}^{*}(n)$$

$$\boldsymbol{g}(n) = \frac{1}{1 + \boldsymbol{x}^{T}(n)\boldsymbol{z}(n)}\boldsymbol{z}(n)$$

$$a(n) = d(n) - \boldsymbol{w}_{n-1}^{T}\boldsymbol{x}(n)$$

$$\boldsymbol{w}_{n} = \boldsymbol{w}_{n-1} + a(n)\boldsymbol{g}(n)$$

$$\boldsymbol{P}(n) = 1^{-1}\left[\boldsymbol{P}(n-1) - \boldsymbol{g}(n)\boldsymbol{z}^{H}(n)\right]$$

4. Επιστροφή στο βήμα 2 για $n = n + 1$.

Αλγόριθμος RLS (13/17)

- Αποδεικνύεται ότι το πλεονάζον MSE σταθερής κατάστασης είναι:

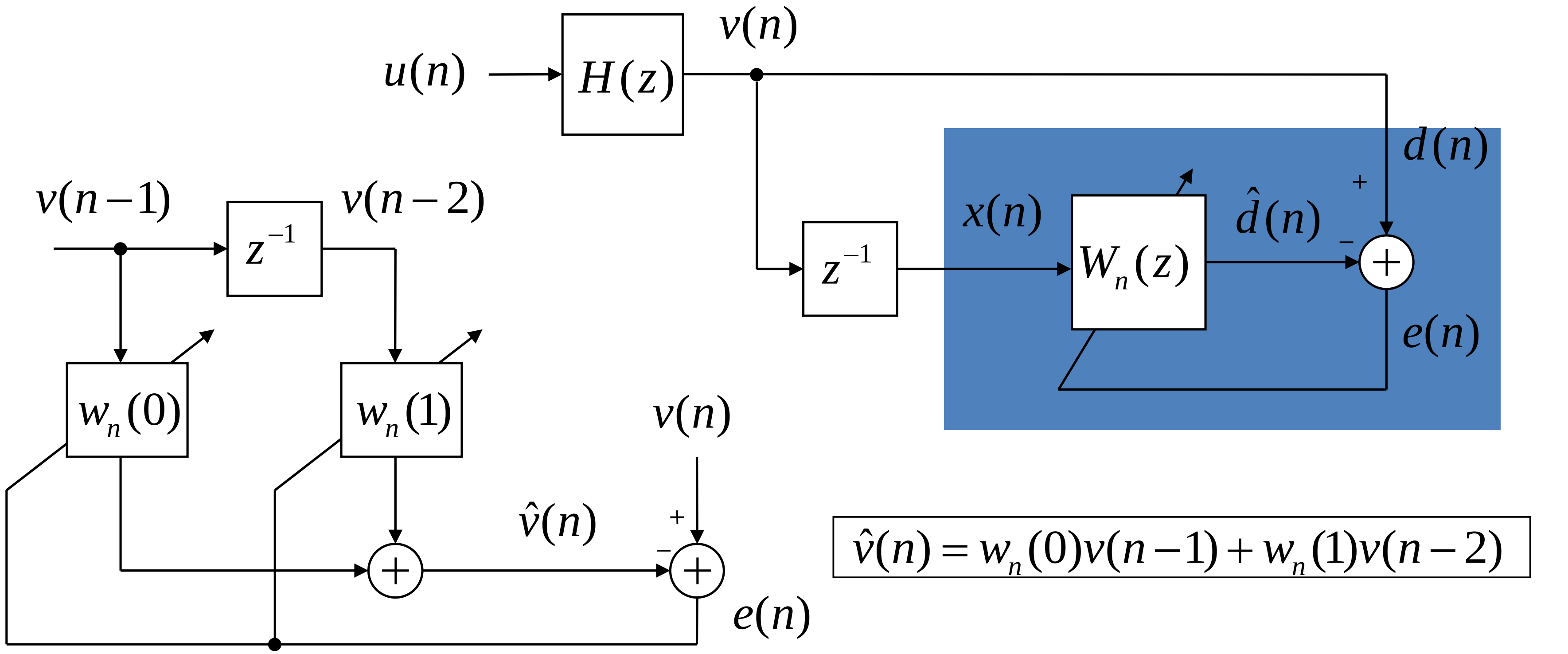
$$\xi_{\mathrm{ex}}(\infty) = p\left(\frac{1-l}{1+l}\right)\xi_{\min}$$

- Σημειώενται πως ως σήμα σφάλματος χρησιμοποιείται το a priori σφάλμα

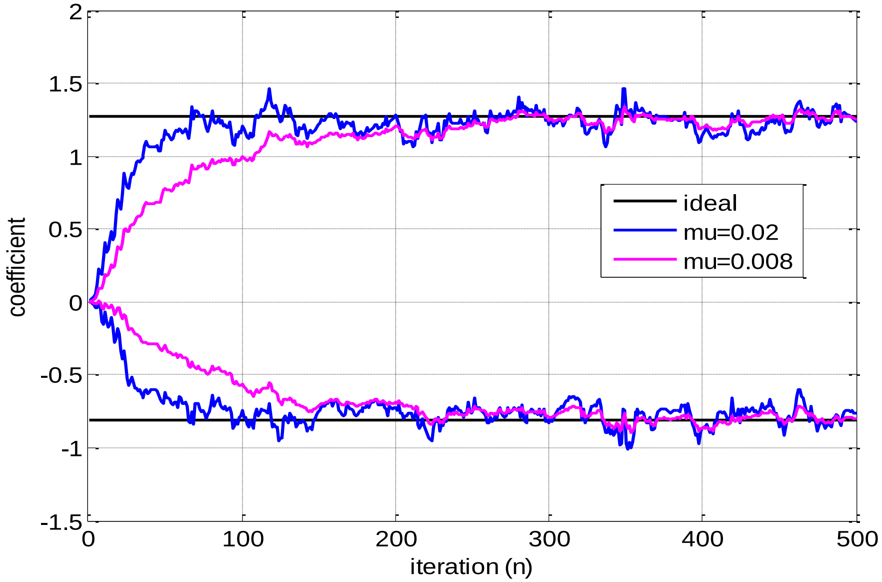
$$a(n) = d(n) - \boldsymbol{w}_{n-1}^{T}\boldsymbol{x}(n)$$

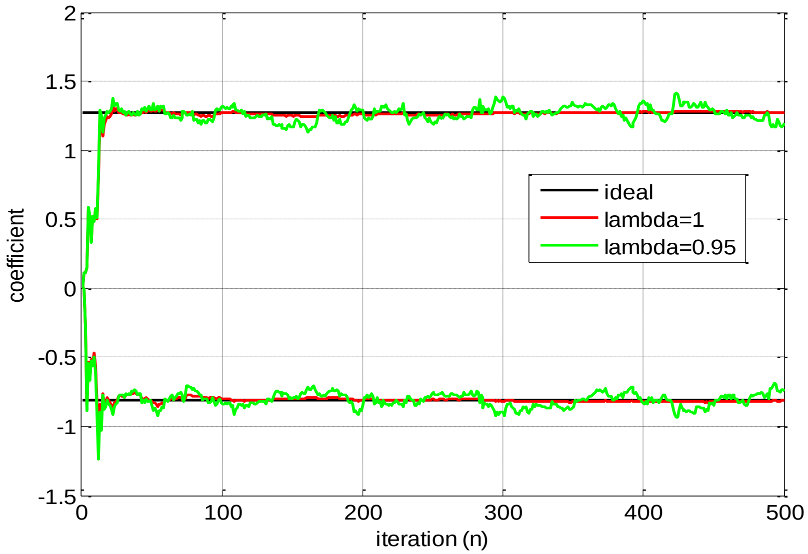
Αλγόριθμος RLS (14/17)

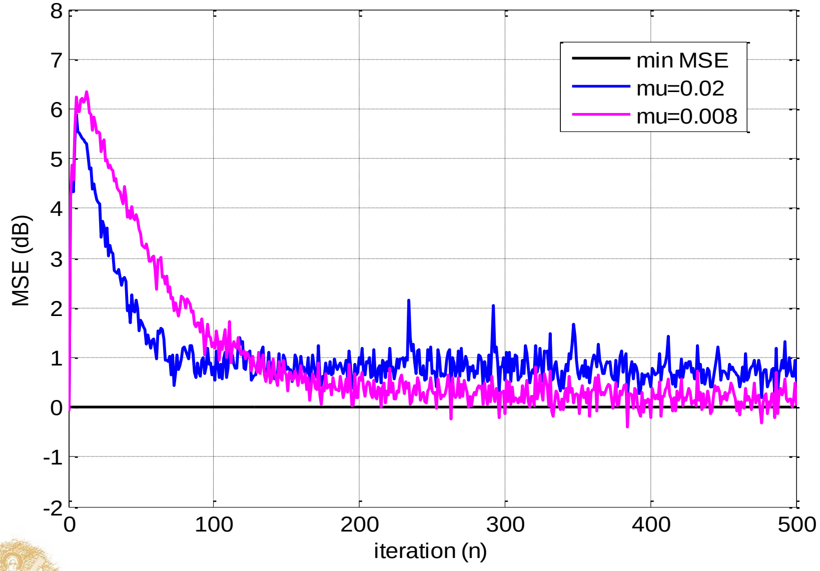
- Γραμμική Πρόβλεψη (Linear Prediction): Η διαδικασία $v(n)$ είναι AR(2) διαδικασία με την ακόλουθη εξίσωση διαφορών, όπου $u(n)$ είναι λευκός θόρυβος μοναδιαίας διασποράς: $v(n) = 1.2728v(n -$

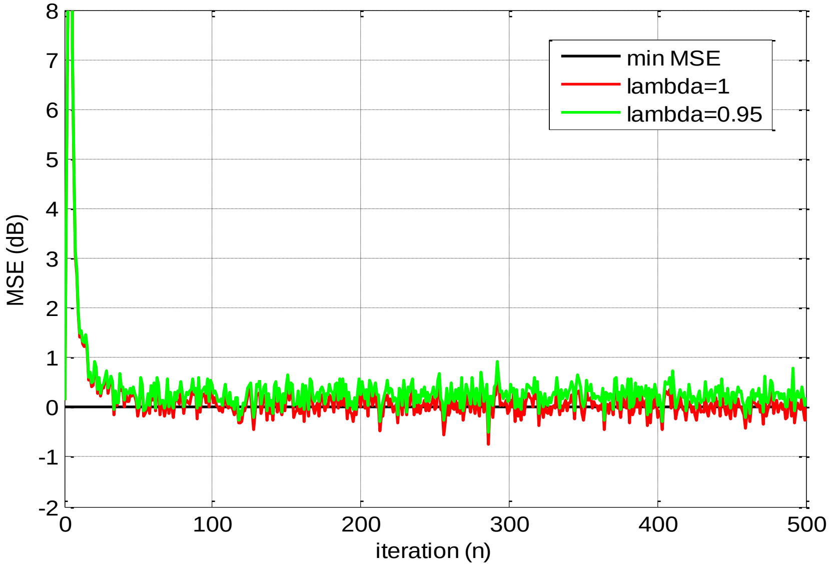


Αλγόριθμος RLS (15/17)

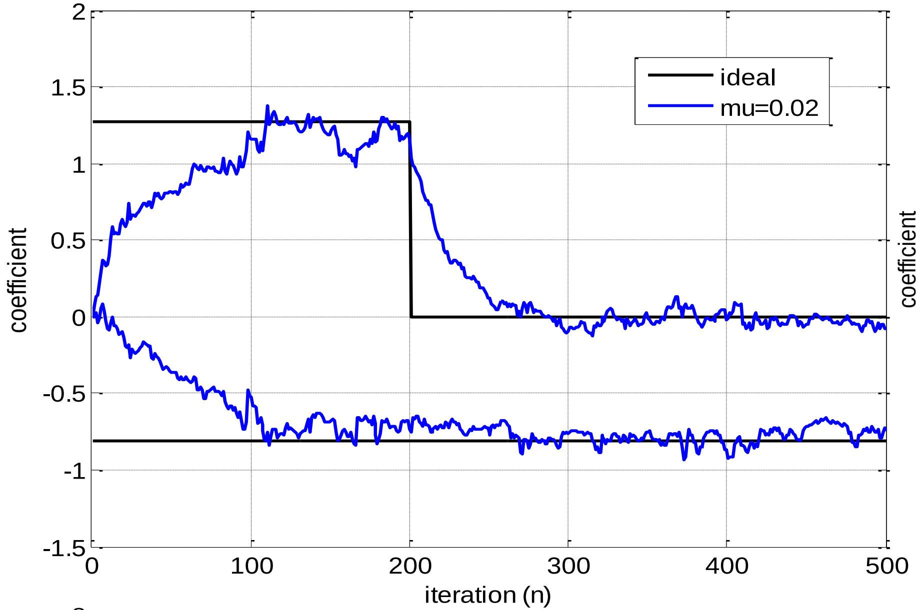


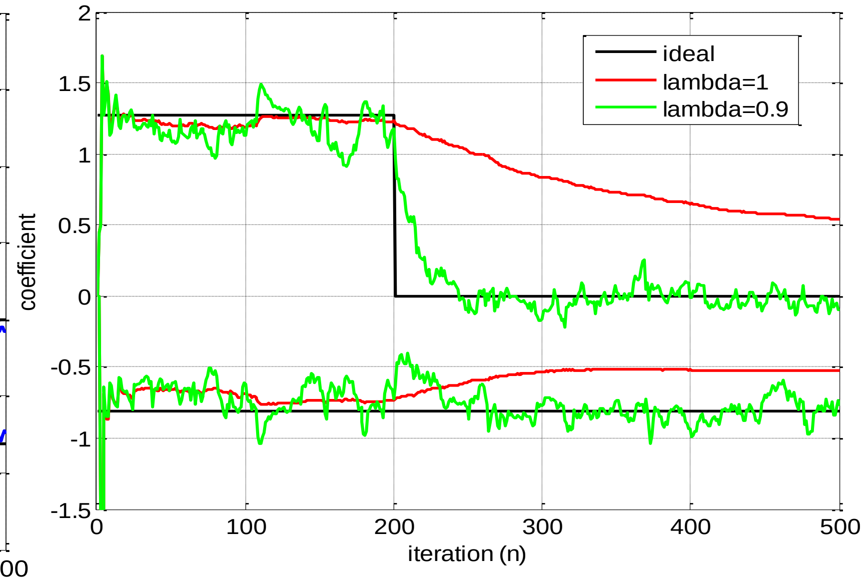


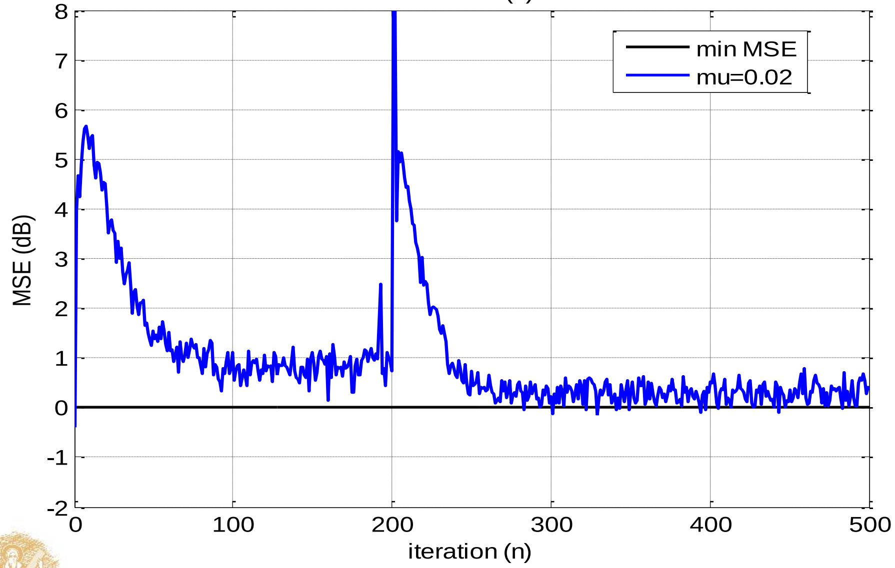


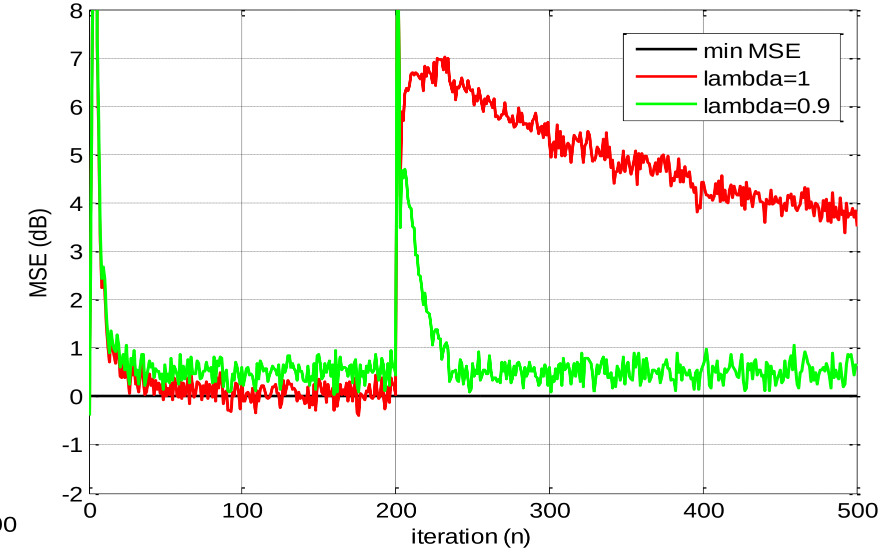


Αλγόριθμος RLS (16/17)

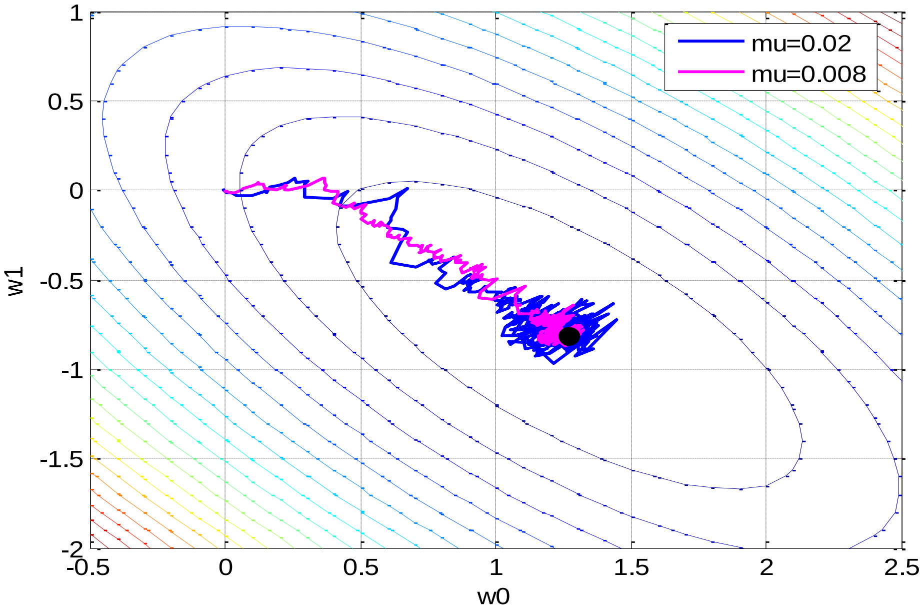


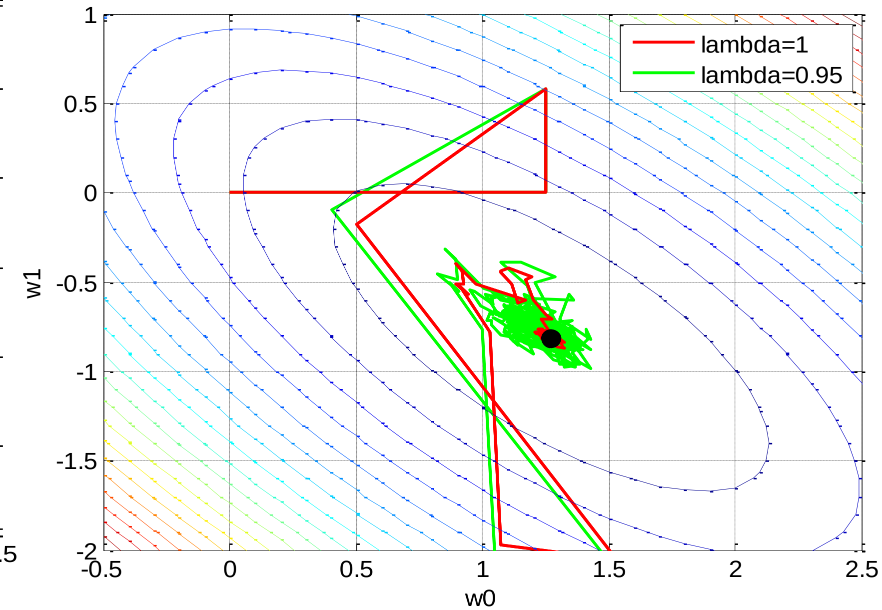


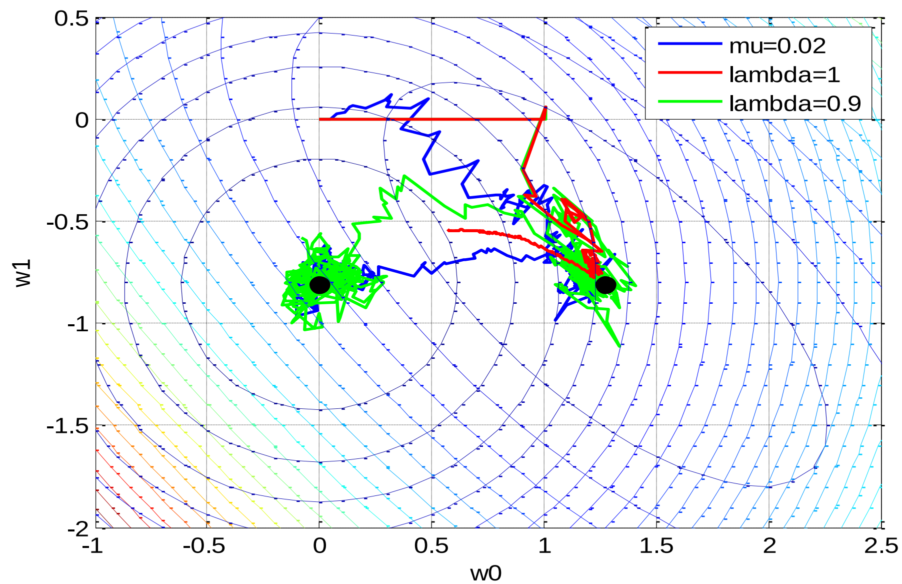




Αλγόριθμος RLS (17/17)







Τέλος Ενότητας 5

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Πατρών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σημειώματα

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.00.

Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Πατρών, Κώστας Μπερμπερίδης. «Στοχαστικά Σήματα και Τηλεπικοιωνίες». Έκδοση: 1.0. Πάτρα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση: https://eclass.upatras.gr/courses/CEID1111/

Σημείωμα Αδειοδότησης

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.