Πανεπιστήμιο Πατρών
Τμήμα Μηχανικών Η/Υ και Πληροφορικής

Στατιστική Επεξεργασία Σήματος και Μάθηση

Πρώτη Εργαστηριακή Άσκηση

Ακαδημαϊκό Έτος 2023/24

Σούρλας Ζήσης *

10 Φεβρουαρίου 2024

*AM: 1072477 Email: sourlas.zisis@ac.upatras.gr

Περιεχόμενα

1 Θεωρητικός υπολογισμός βέλτιστου φίλτρου Weiner **2**

1.1 Υπολογισμός βέλτιστων συντελεστών 2

1.2 Υπολογισμός ελάχιστου σφάλματος 3

2 Εφαρμογή των αλγορίθμων LMS, wLMS και RLS **4**

2.1 Επιλογή βήματος για τον LMS/wLMS 4

2.2 Πειραματική σύγκριση αλγορίθμων 4

2.3 Ειδική περίπτωση του wLMS 6

3 Περιγραφή αγνώστου συστήματος **7**

1 Θεωρητικός υπολογισμός βέλτιστου φίλτρου Weiner

1.1 Υπολογισμός βέλτιστων συντελεστών

Γνωρίζουμε ότι:

$$h(n) = 51.0\delta(n) - 41.6\delta(n-1) - 30.73\delta(n-2) + 200.15\delta(n-3)$$

άρα:

$$h = \begin{bmatrix} 51.0 \\ -41.6 \\ -30.73 \\ 200.15 \end{bmatrix} \tag{1}$$

Αφού η είσοδος μας είναι λευκός Gaussian θόρυβος με μέση τιμή 0 και διασπορά 1, το μητρώο αυτοσυσχέτισης R_x της εισόδου θα είναι ίσο με:

$$R_x = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \sigma^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & 1 \end{bmatrix} = I \tag{2}$$

Γνωρίζουμε επίσης ότι:

$$d(n) = x^T(n)h = h^T x(n) \tag{3}$$

Η ετεροσυσχέτιση r_{dx} μεταξύ του σήματος εισόδου και του σήματος εξόδου δίνεται από τον τύπο:

$$r_{dx} = E\{d(n)x(n)\} = E\{x(n)d(n)\} \tag{4}$$

Ο οποίος αξιοποιώντας την (3) γίνεται:

$$r_{dx} = E\{x(n)x^T(n)h\} = E\{x(n)x^T(n)\}h = R_x h \tag{5}$$

και άρα λόγω της (2):

$$r_{dx} = Ih = h \tag{6}$$

Οπότε η εξίσωση Weiner-Hopf λόγω των (2),(6) δίνεται ως εξής:

$$R_x w = r_{dx} \implies Iw = r_{dx} \implies w = r_{dx} \implies w = h \tag{7}$$

Για το οποίο ανάλογα με το μήκος φίλτρου L που θα θεωρήσουμε και λόγω της (1) μας δίνει:

- Για $L = 3$:

$$w = \begin{bmatrix} 51.0 \\ -41.6 \\ -30.73 \end{bmatrix}$$

- Για $L = 4$:

$$w = \begin{bmatrix} 51.0 \\ -41.6 \\ -30.73 \\ 200.15 \end{bmatrix}$$

- Για $L = 5$:

$$w = \begin{bmatrix} 51.0 \\ -41.6 \\ -30.73 \\ 200.15 \\ 0 \end{bmatrix}$$

1.2 Υπολογισμός ελάχιστου σφάλματος

Γνωρίζουμε ότι:

$$r_d(0) = E\{d^2(n)\} \tag{8}$$

το οποίο λόγω των (3,2) γίνεται:

$$r_d(0) = E\{h^T x(n) x^T(n) h\} = h^T E\{x(n) x^T(n)\} h = h^T R_x h = h^T I h = h^T h = ||h||^2 \tag{9}$$

Το ελάχιστο σφάλμα δίνεται από τον τύπο:

$$\xi_{min} = r_d(0) - r_{dx}^H w \tag{10}$$

το οποίο από τις (9,6) δίνεται ως:

$$\xi_{min} = ||h||^2 - ||w||^2 \tag{11}$$

Οπότε ανάλογα με το μήκος φίλτρου L που θεωρήσαμε έχουμε:

- Για $L = 3$:

$$\xi_{min} = ((51.0)^2 + (-41.6)^2 + (-30.73)^2 + (200.15)^2) - ((51.0)^2 + (-41.6)^2 + (-30.73)^2) =$$
$$= (200.15)^2 = 40060.0225$$

- Για $L = 4$:

$$\xi_{min} = ((51.0)^2 + (-41.6)^2 + (-30.73)^2 + (200.15)^2) - ((51.0)^2 + (-41.6)^2 + (-30.73)^2 + (200.15)^2) = 0$$

- Για $L = 5$:

$$\xi_{min} = ((51.0)^2 + (-41.6)^2 + (-30.73)^2 + (200.15)^2) - ((51.0)^2 + (-41.6)^2 + (-30.73)^2 + (200.15)^2 + 0^2) = 0$$

2 Εφαρμογή των αλγορίθμων LMS, wLMS και RLS

2.1 Επιλογή βήματος για τον LMS/wLMS

Γνωρίζουμε ότι για να συγκλίνει ο LMS πρέπει να ισχύει ότι:

$$0 < \mu < \frac{2}{\lambda_{max}}$$

όπου λ_{max} η μέγιστη ιδιοτιμή του πίνακα αυτοσυσχέτισης της εισόδου. Αφού η είσοδος είναι λευκός Gaussian θόρυβος με μέση τιμή 0 και διασπορά 1, ισχύει ότι $\lambda_{max} = 1$. Συνεπώς,

$$\mu_{max} = 2$$

άρα

$$\mu = 0.01\mu_{max} = 0.02$$

2.2 Πειραματική σύγκριση αλγορίθμων

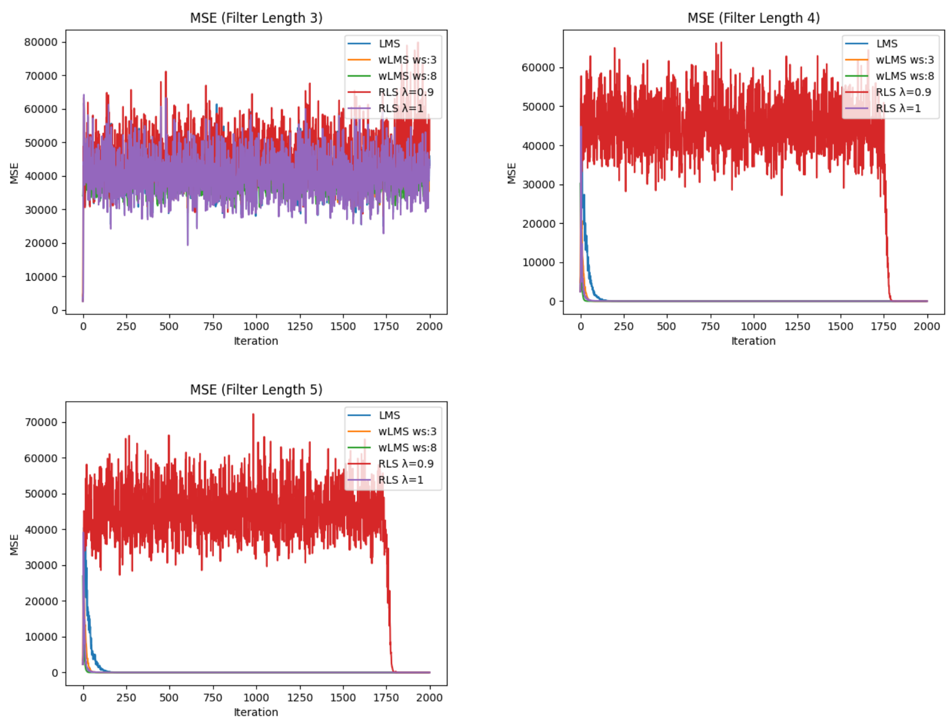

Figure 1: Καμπύλες μάθησης των αλγορίθμων για διαφορετικά μήκη φίλτρων

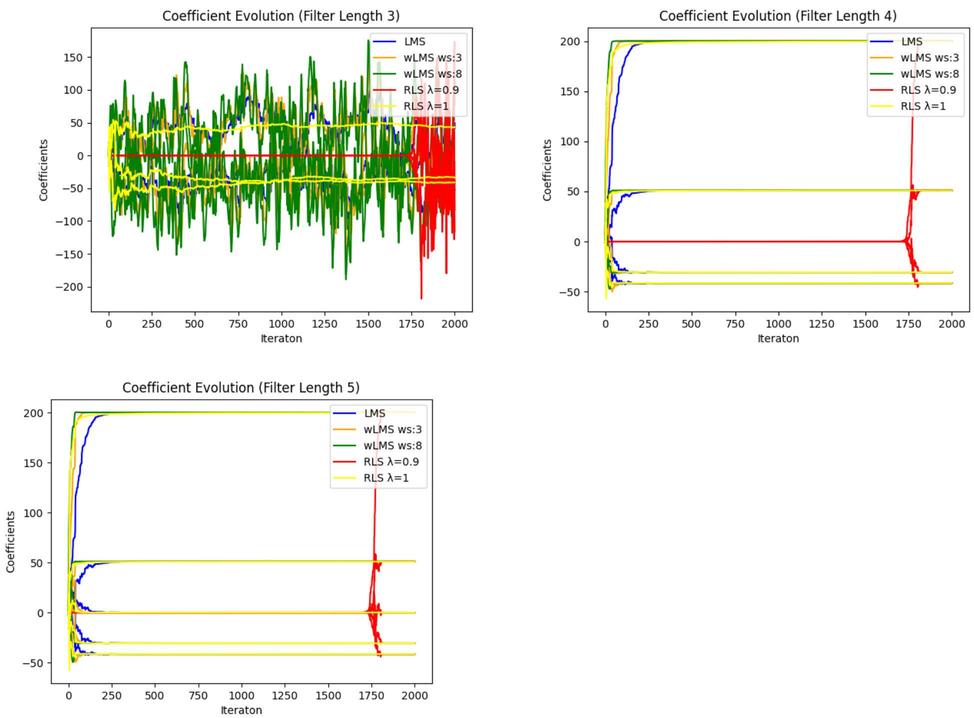

Figure 2: Εξέλιξη των εκτιμώμενων συντελεστών για διαφορετικά μήκη φίλτρων

Τόσο από τις καμπύλες μάθησης όσο και από την εξέλιξη των εκτιμώμενων συντελεστών παρατηρούμε ότι και οι τρεις αλγόριθμοι (ανεξαρτήτως παραμέτρων) συγκλίνουν όταν δεν υποεκτιμάται το πλήθος των συντελεστών του συστήματος. Στην περίπτωση υπερεκτίμησης των συντελεστών δεν παρατηρείται αλλαγή στην ταχύτητα σύγκλισης απλώς ο πλεονάζων συντελεστής συγκλίνει (όπως αναμένεται) στην μηδενική τιμή. Παρατηρείται επίσης (καθαρότερα στα διαγράμματα εξέλιξης των συντελεστών) η αναμενόμενη ταχύτερη σύγκλιση του RLS σε σχέση με τον LMS. Ενδιαφέρον παρουσιάζει επίσης η περίπτωση του wLMS ο οποίος γενικά συγκλίνει ταχύτερα από τον απλό LMS λόγω της αξιοποίησης πολλαπλών δειγμάτων ανά επανάληψη. Η ταχύτητα σύγκλισής του είναι αντίστοιχη με αυτή του RLS ή και υψηλότερη. Η αύξηση του μήκους του παραθύρου φαίνεται να οδηγεί σε ταχύτερη σύγκλιση.

Ειδική μνεία αξίζει να γίνει στον RLS. Παρατηρούμε ότι για παράγοντα λήθης $\lambda = 1$ και αρχικοποίηση του πίνακα P με $\delta = 0.15$ ο αλγόριθμος λειτουργεί κανονικά και συγκλίνει ταχύτατα. Κάτι τέτοιο δεν είναι εφικτό για $\lambda = 0.9$. Σε αυτή την περίπτωση ο αλγόριθμος αδυνατεί εντελώς να συγκλίνει εκτός της περίπτωσης που χρησιμοποιείται πολύ μικρό δ (της τάξεως του 10^{-82}), κάτι το οποίο οδηγεί, ωστόσο, σε υπερβολικά αργή σύγκλιση. Αυτή η περίπτωση είναι που απεικονίζεται και στα διαγράμματα. Επιπλέον σε αντίθεση με τους άλλους δύο αλγορίθμους ο RLS καταφέρνει να κάνει μια προσέγγιση των 3 εκ των 4 συντελεστών στην περίπτωση υποεκτίμησης των παραμέτρων του συστήματος. Αυτό θεωρούμε ότι οφείλεται στον αναδρομικό υπολογισμό του πίνακα αυτοσυσχέτισης της εισόδου που πραγματοποιεί ο αλγόριθμος. Με αυτό τον τρόπο καταφέρνει να προσεγγίσει την βέλτιστη θεωρητική λύση Weiner όπου, όπως δείξαμε πιο πάνω το σφάλμα εξαρτάται αποκλειστικά από τον συντελεστή που απουσιάζει. Πράγματι βλέπουμε στην καμπύλη μάθησης (για μήκος φίλτρου 3) ότι το μέσο τετραγωνικό σφάλμα του RLS κινείται μεταξύ του 50.000 και του 30.000 δηλαδή με μια μέση τιμή κοντά στο 40.000, πολύ κοντά δηλαδή στο θεωρητικό σφάλμα της εξίσωσης Weiner που υπολογίσαμε πιο πάνω.

2.3 Ειδική περίπτωση του wLMS

Στην περίπτωση που το μήκος παραθύρου του wLMS είναι ίσο με τον αριθμό των δειγμάτων εισόδου ισχύουν τα εξής:

Ο αλγόριθμος θα κάνει μόνο μία επανάληψη οπότε περιγράφεται πλήρως από την εξίσωση ανανέωσης των βαρών του:

$$w = w_{n-1} + \mu x^T (d - y_{n-1}) \tag{12}$$

αφού η αρχικοποίηση των βαρών είναι μηδενική και άρα η πρώτη εκτίμηση της εξόδου επίσης μηδενική, η (12) γίνεται:

$$w = \mu x^T (d - y_{n-1}) = \mu x^T d =$$

$$= \mu \begin{bmatrix} x(0)d(0) + \cdots + x(N-1)d(N-1) \\ 0 + x(0)d(1) + \cdots + x(N-2)d(N-1) \\ \vdots \\ 0 + 0 + \cdots + x(0)d(L-1) + \cdots + x(N-L+1)d(N) \end{bmatrix} =$$

$$= \mu \begin{bmatrix} r_{dx}(0) \\ r_{dx}(1) \\ \vdots \\ r_{dx}(L-1) \end{bmatrix} = \mu r_{dx} \tag{13}$$

όπου N το πλήθος των δειγμάτων και L το πλήθος των συντελεστών.
Παρατηρούμε ότι αν θέσουμε $\mu = 1$ (πράγμα ορθό όταν ο αλγόριθμος πραγματοποιεί μόνο μία επανάληψη), η (13) γίνεται:

$$w = r_{dx} \tag{14}$$

δηλαδή πρόκειται για την λύση της εξίσωσης Weiner για τη συγκεκριμένη είσοδο όπως αυτή εξηγήθηκε παραπάνω.

3 Περιγραφή αγνώστου συστήματος

Mean Weiner-Hopf Results	
w1	-0.012141145052683353
w2	-41.37560059833335
w3	-30.541243646681977

Table 1: Μέσος όρος αποτελεσμάτων της εξίσωσης Weiner-Hopf για τα δοθέντα δεδομένα

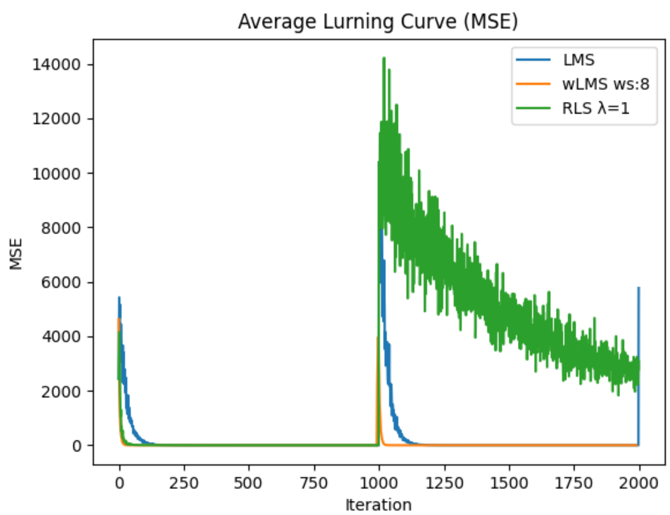

Figure 3: Μέση καμπύλη μάθησης για τους τρεις αλγορίθμους

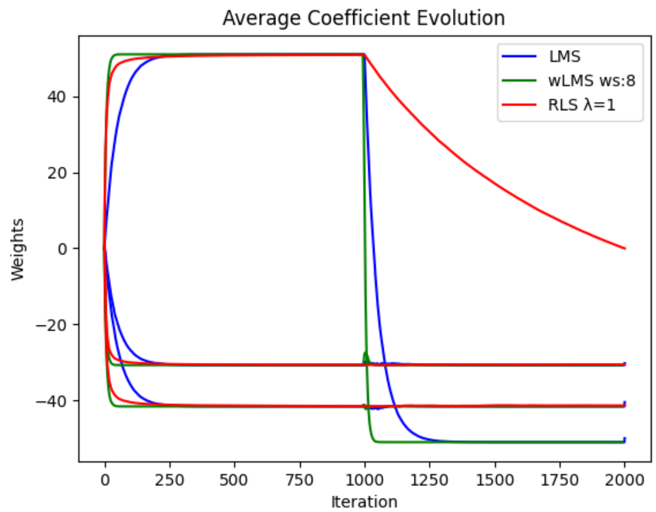

Figure 4: Μέση εξέλιξη συντελεστών για τους τρεις αλγορίθμους

Όπως φαίνεται στα παραπάνω διαγράμματα ένας εκ των συντελεστών του συστήματος αλλάζει τιμή τη χρονική στιγμή 1000. Οι τρεις αλγόριθμοι εντοπίζουν την αλλαγή αυτή καθώς εκτοξεύεται το μέσο τετραγωνικό σφάλμα (όπως φαίνεται στην καμπύλη μάθησης) και επανυπολογίζουν τον συντελεστή. Ο LMS και ο wLMS επανασυγκλίνουν πολύ γρήγορα με τον δεύτερο να είναι αναμενόμενα ταχύτερος. Ο RLS έχει πολύ βραδύτερο ρυθμό σύγκλισης πράγμα αναμενόμενο καθώς, σε αντίθεση με τους άλλους δύο αλγορίθμους δεν βασίζεται μόνο στο στιγμιαίο σφάλμα αλλά διατηρεί "μνήμη". Επιπλέον λόγω του μεγάλου παράγοντα λήθης αποτυγχάνει, ουσιαστικά, να συγκλίνει.

Η λύση της εξίσωσης Weiner-Hopf προσεγγίζει σωστά τους δύο σταθερούς συντελεστές ενώ για τον μεταβαλλόμενο προσεγγίζει την μέση τιμή του, δηλαδή 0 (αφού ο συντελεστής μεταβάλλεται από 51 σε -51). Είναι απόλυτα αναμενόμενο η λύση της εξίσωσης Weiner να μην προσεγγίζει σωστά τον μεταβαλλόμενο συντελεστή καθώς δεν πρόκειται για κάποιο προσαρμοστικό αλγόριθμο αλλά για μια στατική εξίσωση που αντιμετωπίζει το σήμα ως όλον.

Από το διάγραμμα εξέλιξης των συντελεστών μπορούμε να συμπεράνουμε πως η κρουστική απόκριση του χρονικά μεταβαλλόμενου συστήματος δίνεται από τον τύπο:

$$h(n) = \begin{cases} 51\delta(n) + 41.6\delta(n-1) + 30.73\delta(n-2) & n < 1000 \\ -51\delta(n) + 41.6\delta(n-1) + 30.73\delta(n-2) & n \geq 1000 \end{cases}$$