ΒΕΣ 06 – Προσαρμοστικά Συστήματα στις Τηλεπικοινωνίες

Προσαρμοστικοί Αλγόριθμοι Υλοποίησης Βέλτιστων Ψηφιακών Φίλτρων:

Ο αλγόριθμος καθόδου κατά την μέγιστη κλίση (Steepest-descent)

- Εισαγωγή
- Αναδρομή βασισμένη στη κλίση
- Μέθοδος κατάβασης κατά τη μέγιστη κλίση
- Μέθοδος Newton
- Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Βιβλιογραφία Ενότητας

- *Benvenuto [2002]: Κεφάλαιο 3*
- *Widrow [1985]: Chapter 3*
- *Haykin [2001]: Chapter 8*
- *Sayed [2003]: Chapter 3*
- *Boroujeny [1999]: Chapter 3*
- *Bose [2003]: Chapter 8*
- *Chassaing [2004]: Chapter 7*

Εισαγωγή

- Οι προσαρμοστικοί αλγόριθμοι δεν είναι αναδρομικές τεχνικές αναζήτησης της λύσης ενός προβλήματος
  - Ο όρος αναδρομικός αναφέρεται στην προσέγγιση της ζητούμενης λύσης ξεκινώντας από μια τυχαία αρχικοποίηση και βελτιώνοντας διαδοχικά την προσέγγιση μας
  - Στα προσαρμοστικά συστήματα το πρόβλημα είναι το βέλτιστο γραμμικό φιλτράρισμα και η ζητούμενη λύση είναι η λύση Wiener
  - Η ανάγκη για προσαρμοστική αναζήτηση της λύσης Wiener προέρχεται από την μη ευστάθεια του πίνακα αυτοσυσχέτισης R ή της προσέγγισης τους (δεδομένου ότι στην πλειονότητα των περιπτώσεων δεν είναι γνωστός).
  - Η αναδρομική αναζήτηση της λύσης του βέλτιστου γραμμικού φιλτραρίσματος είναι ιδιαίτερα σημαντική σε περιπτώσεις στις οποίες η στοχαστική διεργασία εισόδου δεν είναι στάσιμη αλλά μεταβάλλεται σχετικά αργά
- Οι αναδρομικοί αλγόριθμοι βασισμένοι στην κλίση προϋποθέτουν τη γνώση του πίνακα αυτοσυσχέτισης της στοχαστικής διεργασίας εισόδου. Εντούτοις αποτελούν τη βάση για τον πολύ διαδεδομένο αλγόριθμο LMS

Αναδρομή βασισμένη στη κλίση

- Έχουμε ήδη δει ότι η λύση στο πρόβλημα του βέλτιστου γραμμικού φιλτραρίσματος δίνεται από τις εξισώσεις Wiener-Hoph:

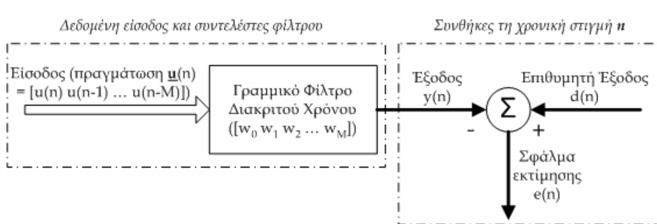
$$\mathbf{R}_u \mathbf{w} = \mathbf{p}_{du}$$

  - Η ανωτέρω λύση ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα:

$$J(\mathbf{w}) = E[\{d(n) - \mathbf{w}^T \mathbf{u}(n)\}^2] = \sigma_d^2 - 2\mathbf{w}^T \mathbf{p}_{du} + \mathbf{w}^T \mathbf{R}_u \mathbf{w}$$

  - Το ελάχιστο μέσο τετραγωνικό σφάλμα δίνεται από τη σχέση

$$J_{\min} = \sigma_d^2 - \mathbf{p}_{du}^T \mathbf{w}_o = \sigma_d^2 - \mathbf{p}_{du}^T \mathbf{R}_u^{-1} \mathbf{p}_{du}$$



☑ Εισαγωγή
★ Αναδρομή βασισμένη στη κλίση
☐ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
☐ Μέθοδος Newton
☐ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Αναδρομή βασισμένη στη κλίση (II)

- Για την αναδρομική εκτίμηση της λύσης Wiener $\mathbf{w}_o = \mathbf{R}_u^{-1}\mathbf{p}_{du}$ το διάνυσμα **w** των συντελεστών του φίλτρου επανεκτιμάται σε κάθε χρονική στιγμή:
  - **w**(n+1) = **w**(n) + δ**w**
  - Το βασικό ζητούμενο είναι η εύρεση του διανύσματος μεταβολής δ**w** ώστε το διάνυσμα **w**(n+1) να αποτελεί καλύτερη προσέγγιση στη λύση Wiener από ότι το διάνυσμα **w**(n) να ισχύει δηλαδή

    $$\|\mathbf{w}(n+1)-\mathbf{w}_o\| < \|\mathbf{w}(n)-\mathbf{w}_o\|$$

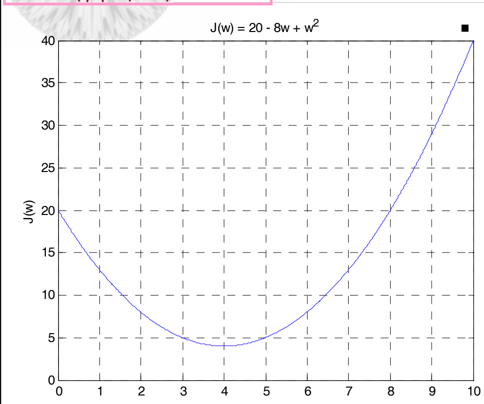
- Στους προσαρμοστικούς αλγόριθμους βασισμένους στη κλίση το διάνυσμα δ**w** υπολογίζεται με βάση τη παράγωγο (ανάδελτα) της διανυσματικής συνάρτησης:

  $$J(\mathbf{w}) = \sigma_d^2 - 2\mathbf{w}^T\mathbf{p}_{du} + \mathbf{w}^T\mathbf{R}_u\mathbf{w}$$

  - Για να είναι εφικτό αυτό η συνάρτηση J(**w**) πρέπει να εκφράζεται αναλυτικά, επομένως τόσο ο πίνακας αυτοσυσχέτισης $\mathbf{R}_\mathbf{u}$ όσο και το διάνυσμα ετεροσυσχέτισης $\mathbf{p}_\mathbf{du}$πρέπει να είναι γνωστά.

☑ Εισαγωγή
★ Αναδρομή βασισμένη στη κλίση
☐ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
☐ Μέθοδος Newton
☐ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Αναδρομή βασισμένη στη κλίση (III)



- Για κατανόηση της αναδρομής βασισμένης στη κλίση θεωρούμε τη μονοδιάστατη περίπτωση (φίλτρο με ένα μόνο συντελεστή *w*)
  - Ζητούμενο: ή εύρεση με αναδρομικό τρόπο του w που ελαχιστοποιεί τη συνάρτηση J(w)
  - Αρχική τιμή για το w: w(0) = 1
  - Ποια θα πρέπει να είναι η επόμενη εκτίμηση w(1) ώστε να πλησιάσουμε προς τη βέλτιστη λύση w_o=4;
- Παρατηρούμε ότι για να κινηθούμε προς τη βέλτιστη λύση πρέπει να κινηθούμε αντίθετα (δw) από το πρόσημο της παραγώγου της συνάρτησης J(w) στο σημείο w(0): $\left.\dfrac{dJ(w)}{dw}\right|_{w=w(0)}$

☑ Εισαγωγή
★ Αναδρομή βασισμένη στη κλίση
☐ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
☐ Μέθοδος Newton
☐ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Παράδειγμα

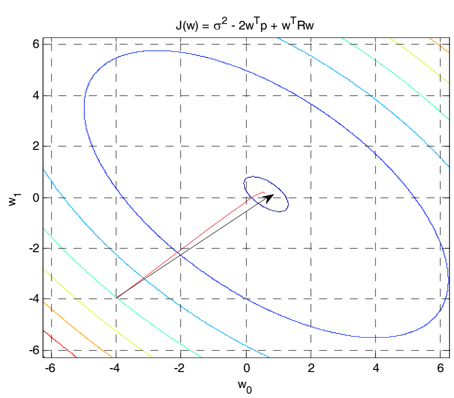
- Η παράγωγος της συνάρτησης $J(w) = 20 - 8w + w^2$ στο σημείο w(0)=1 είναι: $\left.\dfrac{dJ(w)}{dw}\right|_{w=w(0)} = 2w - 8\Big|_{w=1} = -6$
- Επομένως το δw θα πρέπει να είναι θετικό (δεδομένου ότι η παράγωγος της J(w) στο w(0) = 1 έχει αρνητική τιμή)
  - Άρα w(1) = w(0)+δw και προφανώς ισχύει w(1)>w(0)
  - Η τιμή w(2) θα είναι w(2)=w(1)+δw αλλά το δw θα πρέπει να έχει πρόσημο αντίθετο από την παράγωγο στο σημείο *w(1)*.
  - Πρέπει να σημειωθεί ότι εκτός από το πρόσημο του δw θα πρέπει να ορισθεί και το μέγεθος (τιμή) του. Μεγάλη τιμή του δw μπορεί να μην οδηγήσει στη βέλτιστη λύση ενώ πολύ μικρή τιμή μπορεί να οδηγήσει στη λύση μεν αλλά με πολύ αργό ρυθμό δε.
- Αν η αρχική τιμή για το w ήταν w(0)=5 θα είχαμε:

$$\left.\frac{dJ(w)}{dw}\right|_{w=w(0)} = 2w - 8\Big|_{w=5} = 2$$

και προφανώς το δw θα ήταν αρνητικό (άρα w(1)<w(0))

☑ Εισαγωγή
★ Αναδρομή βασισμένη στη κλίση
☐ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
☐ Μέθοδος Newton
☐ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Παράγωγος διανυσματικής συνάρτησης



- Στην περίπτωση διανυσματικών συναρτήσεων η παράγωγος ως προς το διάνυσμα (γνωστή ως κλίση ή ανάδελτα) των ανεξάρτητων μεταβλητών είναι ένα διάνυσμα
  - Η μεταβολή του διανύσματος δ**w** (**w**(n+1) = **w**(n)+ δ**w**) γίνεται με κατεύθυνση αντίστροφη προς τη κλίση της συνάρτησης J(**w**) στο σημείο **w**(n).

☑ Εισαγωγή
★ Αναδρομή βασισμένη στη κλίση
☐ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
☐ Μέθοδος Newton
☐ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

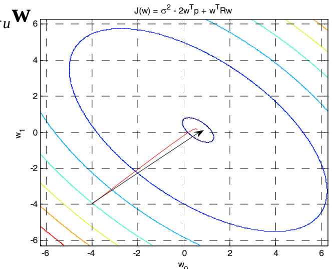
Παράδειγμα

- Έστω η συνάρτηση $J(\mathbf{w}) = \sigma_d^2 - 2\mathbf{w}^T\mathbf{p}_{du} + \mathbf{w}^T\mathbf{R}_u\mathbf{w}$

  με $\sigma_d^2=2$ και

  $$\mathbf{R}_u = \begin{bmatrix} 1 & 0.6 \\ 0.6 & 1 \end{bmatrix} \quad \mathbf{p}_{du} = \begin{bmatrix} 0.7 \\ 0.5 \end{bmatrix}$$

  Αν $\mathbf{w}(0) = \begin{bmatrix} -4 \\ -4 \end{bmatrix}$



  να βρεθεί ένα πιθανό διάνυσμα $\mathbf{w}(1)$ σύμφωνα με βάση τη κλίση της συνάρτησης J(**w**)

- Έχουμε: $\dfrac{\partial J(\mathbf{w})}{\partial \mathbf{w}} = -2\mathbf{p}_{du} + 2\mathbf{R}_u\mathbf{w}\Big|_{\mathbf{w}=\mathbf{w}(0)} = -2\mathbf{p}_{du} + 2\mathbf{R}_u\mathbf{w}\Big|_{\mathbf{w}=\begin{bmatrix} -4 \\ -4 \end{bmatrix}} = \begin{bmatrix} -7.1 \\ -6.9 \end{bmatrix}$

  Επομένως το δ**w** θα πρέπει να είναι αντίθετο με το διάνυσμα $\begin{bmatrix} -7.1 \\ -6.9 \end{bmatrix}$

☑ Εισαγωγή
☑ Αναδρομή βασισμένη στη κλίση
★ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
☐ Μέθοδος Newton
☐ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Μέθοδος κατάβασης κατά τη μέγιστη κλίση

- Η μέθοδος κατάβασης κατά τη μέγιστη κλίση επιδιώκει την εύρεση, με αναδρομικό τρόπο, του ελάχιστου της συνάρτησης:

  $$J(\mathbf{w}) = \sigma_d^2 - 2\mathbf{w}^T\mathbf{p}_{du} + \mathbf{w}^T\mathbf{R}_u\mathbf{w}$$

  μεταβάλλοντας το διάνυσμα των συντελεστών του φίλτρου **w**(n) κατά την αντίστροφη κατεύθυνση από την κλίση της συνάρτησης J(**w**).
  - Με τον τρόπο αυτό η μεταβολή του διανύσματος **w**(n) γίνεται προς την κατεύθυνση προς την οποία μειώνεται περισσότερο το μέσο τετραγωνικό σφάλμα (εξ ου και το όνομα κατάβαση κατά τη μέγιστη κλίση ή steepest descent)
  - Η μεταβολή του διανύσματος των συντελεστών δίνεται από τη σχέση:

    $$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu\nabla J(\mathbf{w})$$

  - Η κλίση (ανάδελτα) της συνάρτησης J(**w**) δίνεται από τη σχέση:

    $$\nabla J(\mathbf{w}) = -2\mathbf{p}_{du} + 2\mathbf{R}_u\mathbf{w}$$

  - Η παράμετρος μ καθορίζει το μέγεθος της μεταβολής του διανύσματος των συντελεστών και καθορίζει την ταχύτητα εύρεσης της βέλτιστης λύσης

Αλγόριθμος κατάβασης κατά τη μέγιστη κλίση

1. Ξεκινάμε την αναζήτηση της βέλτιστης λύσης $\mathbf{w}_o$ από ένα τυχαίο διάνυσμα συντελεστών **w**(0). Αν δεν υπάρχει κάποια εκ των προτέρων πληροφορία για την τιμή του διανύσματος $\mathbf{w}_o$ τότε το **w**(0) τίθεται ίσο με το μηδενικό διάνυσμα **w**(0) = **0**
2. Υπολογίζουμε το διάνυσμα της κλίσης της συνάρτησης J(**w**) στο σημείο **w** = **w**(0)
3. Υπολογίζουμε τη νέα εκτίμηση του διανύσματος $\mathbf{w}_o$ μεταβάλλοντας διάνυσμα **w**(0) κατά την αντίστροφη κατεύθυνση του διανύσματος της κλίσης $$\mathbf{w}(1) = \mathbf{w}(0) - \mu \nabla J(\mathbf{w})\Big|_{\mathbf{w} = \mathbf{w}(0)}$$
4. Επαναλαμβάνουμε τα βήματα 2-3 με τις τιμές n = 1, 2, … (θέτοντας όπου **w**(0) = **w**(n) και όπου **w**(1) = **w**(n+1))

Παράδειγμα

- Να εφαρμόσετε τον αλγόριθμο κατάβασης κατά τη μέγιστη κλίση για την εύρεση του βέλτιστου ψηφιακού φίλτρου 2 συντελεστών. Δίνονται:

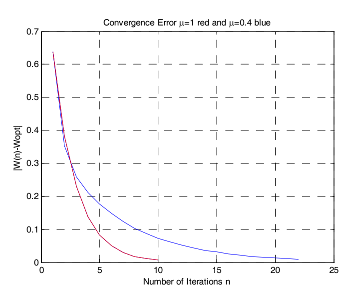
$\sigma_d^2 = 2$

$$\mathbf{R}_u = \begin{bmatrix} 1 & 0.6 \\ 0.6 & 1 \end{bmatrix} \qquad \mathbf{p}_{du} = \begin{bmatrix} 0.7 \\ 0.5 \end{bmatrix}$$

Χρησιμοποιείστε μ = 1 και μ = 0.4 και συγκρίνεται τα αποτελέσματα.

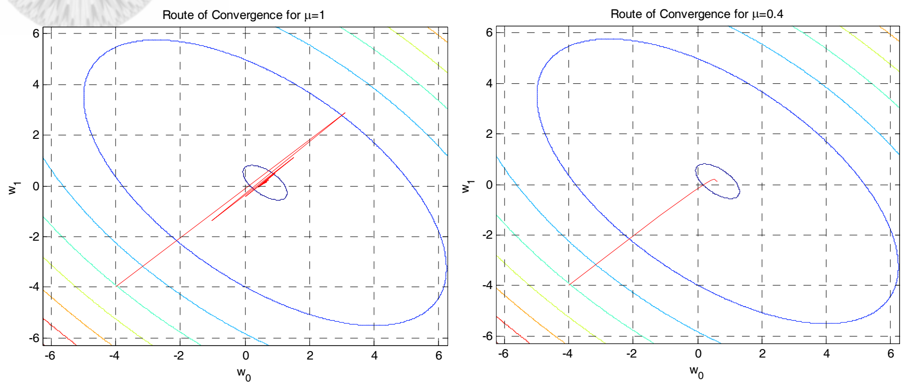
- Για μ=1 παίρνουμε τις επόμενες εκτιμήσεις:

$$W = \begin{bmatrix} 0 & 0.70 & 0.40 & 0.65 & 0.54 & 0.64 & 0.60 & 0.63 & 0.61 & 0.626 \\ 0 & 0.50 & 0.08 & 0.26 & 0.11 & 0.17 & 0.12 & 0.14 & 0.12 & 0.131 \end{bmatrix}$$



☑ Εισαγωγή
☑ Αναδρομή βασισμένη στη κλίση
★ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
☐ Μέθοδος Newton
☐ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Παράδειγμα (συν.)



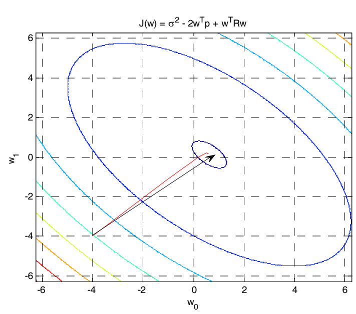
☑ Εισαγωγή
☑ Αναδρομή βασισμένη στη κλίση
☑ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
★ Μέθοδος Newton
☐ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Μέθοδος Newton

- Στη μέθοδο Newton ακολουθεί τις αρχές της μεθόδου κατάβασης κατά τη μέγιστη κλίση, η μεταβολή όμως του διανύσματος των συντελεστών δίνεται από τη σχέση:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu R_u^{-1} \nabla J(\mathbf{w})$$

  - Ο πολλαπλασιασμός του διανύσματος της κλίσης (ανάδελτα) με τον αντίστροφο του πίνακα αυτοσυσχέτισης R_u^{-1} στρέφει το διάνυσμα της μεταβολής προς την κατεύθυνση της βέλτιστης λύσης (βλέπε βέλος στο διπλανό σχήμα)



Η μέθοδος Newton συγκλίνει ταχύτερα στη βέλτιστη λύση από ότι η μέθοδος κατάβασης κατά τη μέγιστη κλίση. Απαιτεί όμως ο πίνακας αυτοσυσχέτισης R_u να είναι αντιστρέψιμος και ευσταθής. Επειδή συνήθως ο πίνακας αυτοσυσχέτισης R_u δεν είναι γνωστός αλλά εκτιμάται από τα δεδομένα εισόδου η μέθοδος Newton δεν εφαρμόζεται συχνά στην πράξη.

☑ Εισαγωγή
☑ Αναδρομή βασισμένη στη κλίση
☑ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
★ Μέθοδος Newton
☐ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

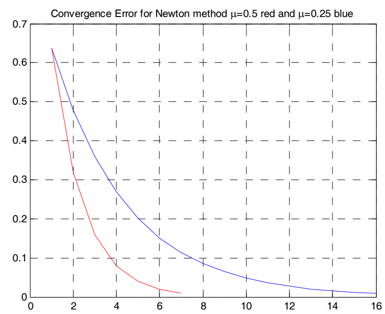
Παράδειγμα

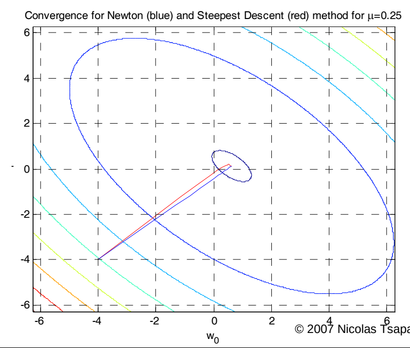
- Να εφαρμόσετε τον αλγόριθμο Newton για την εύρεση του βέλτιστου ψηφιακού φίλτρου 2 συντελεστών. Δίνονται:

  $\sigma_d^2=2$

$$\mathbf{R}_u = \begin{bmatrix} 1 & 0.6 \\ 0.6 & 1 \end{bmatrix} \quad \mathbf{p}_{du} = \begin{bmatrix} 0.7 \\ 0.5 \end{bmatrix}$$

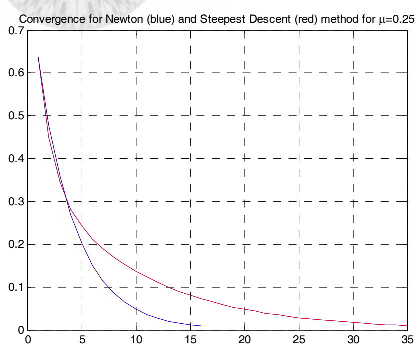
Χρησιμοποιείστε μ = 0.5 και μ = 0.25 και συγκρίνεται τα αποτελέσματα.

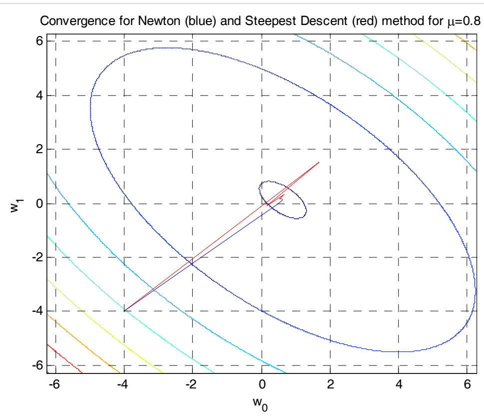




☑ Εισαγωγή
☑ Αναδρομή βασισμένη στη κλίση
☑ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
★ Μέθοδος Newton
☐ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Παράδειγμα (συν)





☑ Εισαγωγή
☑ Αναδρομή βασισμένη στη κλίση
☑ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
☑ Μέθοδος Newton
★ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Σύγκλιση αλγορίθμου κατάβασης

- Με τον όρο σύγκλιση εννοούμε την προοδευτική προσέγγιση της βέλτιστης λύσης $\mathbf{w}_0$ από το διάνυσμα $\mathbf{w}(n)$ όσο αυξάνεται το *n*.
    - Η σύγκλιση του αλγορίθμου κατάβασης κατά τη μέγιστη κλίση εξαρτάται από τη επιλογή της παραμέτρου *μ* (step size parameter)

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu \nabla J(\mathbf{w}) = \mathbf{w}(n) + 2\mu(\mathbf{p}_{du} - \mathbf{R}_u \mathbf{w}(n))$$
$$= \mathbf{w}(n) + 2\mu(\mathbf{R}_u \mathbf{w}_o - \mathbf{R}_u \mathbf{w}(n))$$

- Για να έχουμε σύγκλιση χρειάζεται: $\|\mathbf{w}(n+1) - \mathbf{w}_o\| < \|\mathbf{w}(n) - \mathbf{w}_o\|$

$$\mathbf{w}(n+1) - \mathbf{w}_o = \mathbf{w}(n) - \mathbf{w}_o + 2\mu(\mathbf{R}_u \mathbf{w}_o - \mathbf{R}_u \mathbf{w}(n)) = (\mathbf{I} - 2\mu \mathbf{R}_u)(\mathbf{w}(n) - \mathbf{w}_o)$$

από την προηγούμενη σχέση είναι φανερό ότι για να έχουμε σύγκλιση χρειάζεται: $\|\mathbf{I} - 2\mu \mathbf{R}_u\| < 1$

η παραπάνω σχέση μεταφράζεται στη σχέση $0 < \mu < \frac{2}{\lambda_{\max}}$

όπου λ_{max} είναι η μεγαλύτερη ιδιοτιμή του πίνακα αυτοσυσχέτισης R_u.

☑ Εισαγωγή
☑ Αναδρομή βασισμένη στη κλίση
☑ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
☑ Μέθοδος Newton
★ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Παράδειγμα

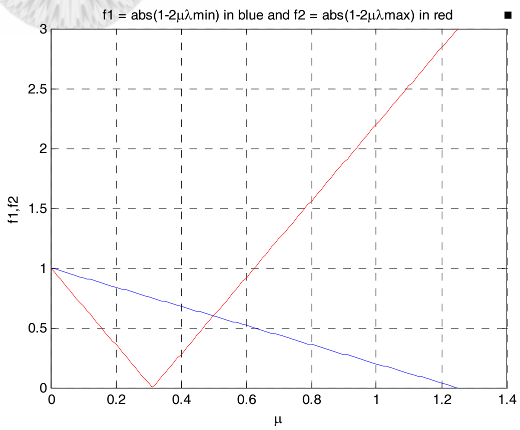
- Να βρεθούν οι τιμές του *μ* για τις οποίες ο αλγόριθμος κατάβασης κατά τη μέγιστη κλίση συγκλίνει προς τη λύση Wiener στην περίπτωση που έχουμε:

$$\mathbf{R}_u = \begin{bmatrix} 1 & 0.6 \\ 0.6 & 1 \end{bmatrix} \quad \mathbf{p}_{du} = \begin{bmatrix} 0.7 \\ 0.5 \end{bmatrix}$$

- Να απεικονίσετε για διάφορες τιμές του *μ* τα διαγράμματα $|1\text{-}2\mu\lambda_{min}|$ και $|1\text{-}2\mu\lambda_{max}|$ και να βρείτε την τιμή του *μ* για την οποία επιτυγχάνεται η ταχύτερη σύγκλιση στη βέλτιστη λύση (μ_{opt}). Η ταχύτητα σύγκλισης καθορίζεται από τον παράγοντα $a = |1-2\mu_{opt}\lambda_{min}|$. Όσο μικρότερο είναι το *a* τόσο ταχύτερη είναι η σύγκλιση.

☑ Εισαγωγή
☑ Αναδρομή βασισμένη στη κλίση
☑ Μέθοδος κατάβασης κατά τη μέγιστη κλίση
☑ Μέθοδος Newton
★ Σύγκλιση αλγορίθμου κατάβασης κατά τη μέγιστη κλίση

Παράδειγμα (συν)



- Το σημείο τομής των δύο διαγραμμάτων δείχνει τη βέλτιστη τιμή για το μ:
  - Αποδεικνύεται ότι η τιμή αυτή είναι ίση με:

$$\mu_{opt} = \frac{1}{\lambda_{\min} + \lambda_{\max}}$$

  - Ο παράγοντας *a* δίνεται τότε από τη σχέση

$$a = \frac{\frac{\lambda_{\max}}{\lambda_{\min}} - 1}{\frac{\lambda_{\max}}{\lambda_{\min}} + 1}$$