ΚΕΣ 03 – Αναγνώριση Προτύπων και Ανάλυση Εικόνας

Γραμμικοί Ταξινομητές

Τμήμα Επιστήμης και Τεχνολογίας Τηλεπικοινωνιών

Πανεπιστήμιο Πελοποννήσου

Περιεχόμενα – Βιβλιογραφία

- Περιεχόμενα Ενότητας
  - Εισαγωγή
  - Γραμμικές συναρτήσεις διαχωρισμού
  - Ο αλγόριθμος Perceptron
  - Μέθοδοι Ελαχίστων Τετραγώνων
  - Support Vector Machines

- Βιβλιογραφία:
  - Duda [2004]: Chapter 5
  - Theodoridis [2002]: Chapter 3
  - Bow [2002]: Chapter 3

Εισαγωγή

- Η εφαρμογή της ταξινόμησης σύμφωνα με τον κανόνα του Bayes απαιτεί τη γνώση των κατανομών πιθανότητας $p(\underline{x}|\omega_i)$, $i = 1,2, \ldots, M$
  - Αν οι ανωτέρω κατανομές δεν είναι γνωστές πρέπει να εκτιμηθούν από ένα σύνολο διανυσμάτων $\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N\}$ για τα οποία είναι γνωστές οι κλάσεις στις οποίες ανήκουν
- Εναλλακτικά μπορούν να εφαρμοστούν μέθοδοι ταξινόμησης, οι οποίες προφανώς δεν είναι βέλτιστες, χωρίς να απαιτείται η γνώση ή η εκτίμηση των ανωτέρω κατανομών πιθανότητας παρά μόνο οι τιμές ενός συνόλου διανυσμάτων εκπαίδευσης $\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N\}$ για τα οποία είναι γνωστές οι κλάσεις στις οποίες ανήκουν.
  - Στις μεθόδους αυτές πολλές φορές χρησιμοποιούμε γραμμικές συναρτήσεις (των διανυσμάτων $\underline{x}$) για το διαχωρισμό των διανυσμάτων $\underline{x}_i$, $i=1,2,\ldots,N$ στις κλάσεις ω_j , $j=1,2,\ldots,M$.
  - Η χρήση γραμμικών συναρτήσεων γίνεται για απλότητα υπολογισμών. Η εκτίμηση τους πραγματοποιείται από τα διανύσματα $\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N\}$ για τα οποία είναι γνωστές οι κλάσεις στις οποίες ανήκουν

Εισαγωγή (II)

- Η εύρεση γραμμικών συναρτήσεων διαχωρισμού των διανυσμάτων $\underline{x}_i$ στις κλάσεις ω_j, $j=1,2,\ldots,L$ απαιτεί στην ουσία την εύρεση L συναρτήσεων $g_j(\underline{x})$, $j=1,2,\ldots,L$ για τις οποίες να ισχύει:

$$\left.\begin{cases} g_j(\mathbf{x}_i) > 0 \\ g_k(\mathbf{x}_i) < 0 \quad \forall k \neq j \end{cases}\right. \quad \text{όταν } \mathbf{x}_i \text{ ανήκει στη κλάση } \omega_j$$

Γραμμικές Συναρτήσεις Διαχωρισμού

- Έστω ένα πρόβλημα ταξινόμησης σε δύο κλάσεις ω_1, ω_2
  - Αν τα διανύσματα $\underline{x}_i \epsilon R^l$ $i=1,2,\ldots,N$ είναι γραμμικά διαχωρίσιμα τότε υπάρχει μια γραμμική συνάρτηση των l-στοιχείων του $\underline{x}$ της μορφής
  $$g(\underline{x}) = \underline{w}^T \underline{x} + w_0 = 0 =$$
  $$w_1 x_1 + w_2 x_2 + \ldots + w_l x_l + w_0$$
  η οποία διαχωρίζει τις δύο κλάσεις.
  - Το διάνυσμα παραμέτρων $\underline{w} = [w_1, w_2, \ldots, w_l]^T$ καθώς και το κατώφλι w_0 προσδιορίζονται από μια διαδικασία μάθησης με βάση τα διανύσματα εκπαίδευσης
  - Έστω ότι τα διανύσματα $\underline{x}_1$, $\underline{x}_2$ ανήκουν στην επιφάνεια (υπερεπίπεδο) διαχωρισμού των κλάσεων ω_1, ω_2 τότε ισχύει:

$$0 = \underline{w}^T \underline{x}_1 + w_0 = \underline{w}^T \underline{x}_2 + w_0 \Rightarrow \underline{w}^T (\underline{x}_1 - \underline{x}_2) = 0 \quad \forall \underline{x}_1, \underline{x}_2$$

Γραμμικές Συναρτήσεις Διαχωρισμού (II)

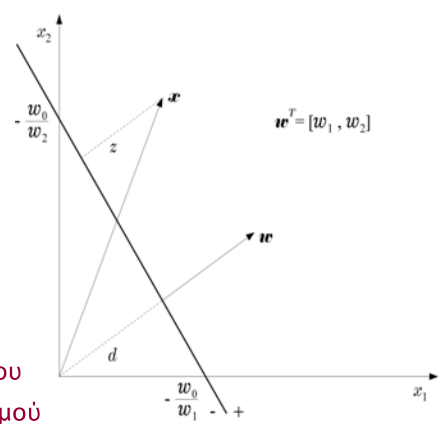
- Επομένως το διάνυσμα των βαρών $\underline{w}$ είναι κάθετο στο υπερεπίπεδο διαχωρισμού που ορίζεται από τη σχέση:

$$g(\underline{x}) = \underline{w}^T \underline{x} + w_0 = 0$$

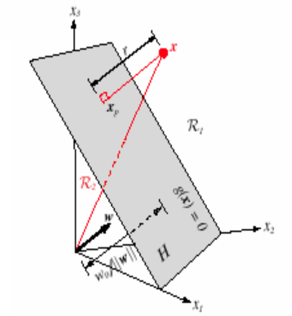
- Από το διπλανό σχήμα (απόσταση σημείου από ευθεία) προκύπτουν οι σχέσεις:

$$d = \frac{|w_0|}{\sqrt{w_1^2 + w_2^2}}, \quad z = \frac{|g(\underline{x})|}{\sqrt{w_1^2 + w_2^2}}$$

- Δηλαδή η τιμή $g(\underline{x})$ είναι ένα μέτρο της απόστασης του σημείου $\underline{x}$ από το υπερεπίπεδο διαχωρισμού



Γεωμετρική ερμηνεία



- Μπορούμε να γράψουμε το διάνυσμα x ως:

$$x = x_p + \frac{r.w}{\|w\|}$$

(Το διάνυσμα των βαρών w είναι συνευθειακό με το διάνυσμα x-x_p και επιπλέον

$\frac{w}{\|w\|} = 1$ είναι το μοναδιαίο διάνυσμα των βαρών)

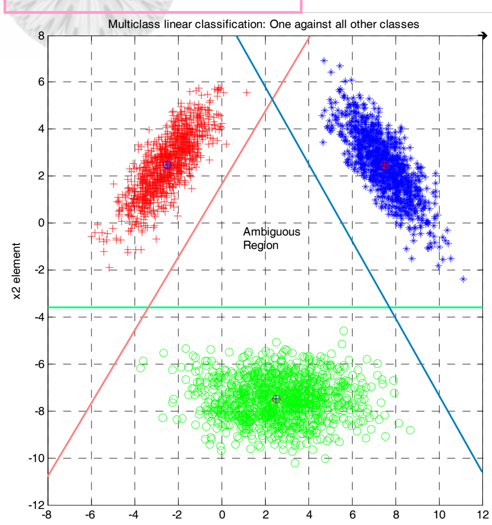
Επιπλέον: $g(\mathbf{x}_p) = 0 \quad \text{και} \quad w^T w = \|w\|^2$

- Οπότε: $r = \frac{g(x)}{\|w\|}$ Ισχύει επίσης: $d(0,H) = \frac{w_0}{\|w\|}$

Συμπέρασμα:

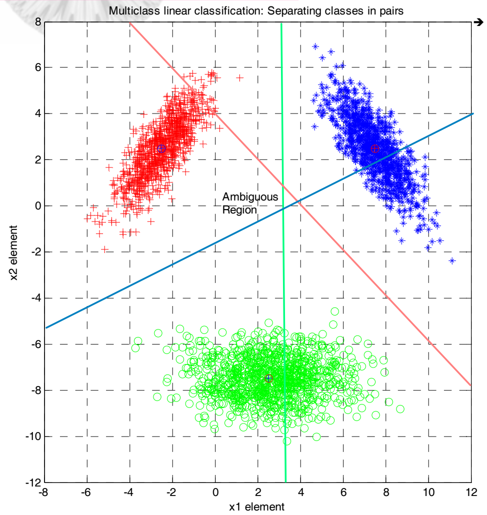
- Μια γραμμική συνάρτηση διαχωρισμού διαιρεί το χώρο των προτύπων μέσω μιας υπερεπιφάνειας διαχωρισμού
- Η κατεύθυνση της υπερεπιφάνειας ορίζεται από το διάνυσμα των βαρών w (κάθετο στην υπερεπιφάνεια) ενώ η θέση της ορίζεται από το κατώφλι w_0.

Πολλαπλές κλάσεις



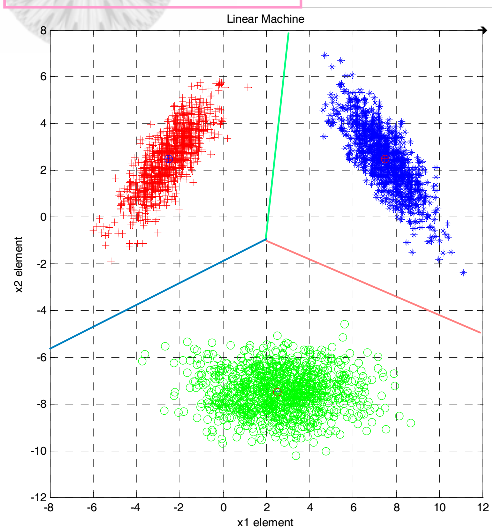
- Όταν έχουμε διαχωρισμό σε πολλές κλάσεις ω_1, ω_2, …, ω_L τότε μπορούμε να επεκτείνουμε την προηγούμενη μεθοδολογία με δύο τρόπους:
  - Για κάθε κλάση ω_i, βρίσκουμε την επιφάνεια διαχωρισμού η οποία διαχωρίζει τα πρότυπα της κλάσης ω_i, από όλα τα υπόλοιπα.
  - Η μεθοδολογία αυτή παράγει μεγάλες περιοχές ασάφειας (ambiguous regions), δηλαδή περιοχές για τις οποίες δεν μπορούμε να αποφασίσουμε σε ποια κλάση θα ταξινομήσουμε ένα πρότυπο, όπως φαίνεται στο σχήμα

Πολλαπλές κλάσεις (ΙΙ)



- Η δεύτερη μεθοδολογία είναι:
  - Για κάθε ζεύγος κλάσεων ω_i, ω_j βρίσκουμε την επιφάνεια διαχωρισμού η οποία διαχωρίζει τα των δύο κλάσεων
  - Επαναλαμβάνουμε τη διαδικασία για όλα τα δυνατά ζεύγη κλάσεων
  - Η μεθοδολογία αυτή πιο υπολογιστικά πολύπλοκη από την προηγούμενη αλλά παράγει περιοχές ασάφειας (ambiguous regions) σαφώς μικρότερες, όπως φαίνεται στο σχήμα

Γραμμικές μηχανές



- Το πρόβλημα της ταξινόμησης σε πολλές κλάσεις χωρίς τη δημιουργία περιοχών ασάφειας λύνεται με τη βοήθεια των γραμμικών μηχανών:
  - Τα όρια τα περιοχών καθορίζονται έτσι ώστε το διάνυσμα $\underline{x}$ να ταξινομείται στη κλάση ω_i, εφόσον ισχύει:

$$g_i(\underline{x}) > g_j(\underline{x}) \quad \forall j \neq i$$

$$g_i(\underline{x}) = \underline{w_i}^T \underline{x} + w_{i0} = 0 \quad i = 1,2,\ldots,M$$

Ο αλγόριθμος Perceptron

- Μέχρι τώρα είδαμε ότι το διάνυσμα των βαρών <u>w</u> είναι κάθετο στο υπερεπίπεδο διαχωρισμού που ορίζεται από τη συνάρτηση διαχωρισμού: $g(\underline{x})=\underline{w}^T\underline{x}+w_0=0$
  - Το διάνυσμα των βαρών $\underline{w}^* = [\underline{w}^T\ w_0]^T$ δεν είναι γνωστό και χρειάζεται να εκτιμηθεί από τα διανύσματα $\underline{x}_i$, $i=1,2,\ldots,N$ για τα οποία είναι γνωστές οι κλάσεις στις οποίες ανήκουν.
  - Στην περίπτωση των πολλαπλών κλάσεων χρειάζεται να υπολογιστούν τα διανύσματα βαρών $\underline{w}_i^* = [\underline{w}_i^T\ w_{i0}]^T$ για να οριστούν οι συναρτήσεις διαχωρισμού:

$$g_i(\underline{x}) = \underline{w}_i^T\underline{x} + w_{i0} = 0 \quad i=1,2,\ldots,M$$

- Για τον υπολογισμό των διανυσμάτων $\underline{w}_i^*$ συνήθως χρησιμοποιείται κάποια μεθοδολογία βελτιστοποίησης.
  - Κάθε μεθοδολογία βελτιστοποίησης απαιτεί: (α) το ορισμό μιας συνάρτησης κόστους, και (β) το ορισμό ενός αλγορίθμου για τη βελτιστοποίηση της συνάρτησης αυτής
  - Μια μεθοδολογία βελτιστοποίησης είναι ο αλγόριθμος Perceptron.

ΚΕΣ 03: Αναγνώριση Προτύπων και Ανάλυση Εικόνας

☑ Εισαγωγή
☑ Γραμμικές Συναρτήσεις Διαχωρισμού
★ Ο αλγόριθμος Perceptron
☐ Μέθοδοι Ελαχίστων Τετραγώνων
☐ Support Vector Machines

Ο αλγόριθμος Perceptron (ΙΙ)

- Έστω οι γραμμικά διαχωρίσιμες κλάσεις ω$_1$, ω$_2$: $\exists\ \underline{w}^*: \ \underline{w}^{*T}\underline{x} > 0\ \ \forall \underline{x}\in\omega_1$
  $\underline{w}^{*T}\underline{x} < 0\ \ \forall \underline{x}\in\omega_2$
  - Η περίπτωση $\underline{w}^{*T}\underline{x}+w_0^*$ υπάγεται στην πιο πάνω περίπτωση δεδομένου ότι:

$$\underline{w}' \equiv \begin{bmatrix} \underline{w}^* \\ w_0^* \end{bmatrix},\ \underline{x}' = \begin{bmatrix} \underline{x} \\ 1 \end{bmatrix} \qquad \underline{w}^{*T}\underline{x} + w_0^* = \underline{w}'^T\underline{x}' = 0$$

- Στόχος: Υπολογισμός ενός διανύσματος <u>w</u> ώστε: $\underline{w}^T\underline{x} > 0\ \ \underline{x}\in\omega_1$, $\underline{w}^T\underline{x} < 0\ \ \underline{x}\in\omega_2$

- Βήματα:
  - Ορισμός μιας συνάρτησης κόστους
  - Επιλογή ενός αλγορίθμου για ελαχιστοποίηση της συνάρτησης κόστους
  - Επιλογή του <u>w</u> το οποίο να ελαχιστοποιεί τη συνάρτηση κόστους

Ο αλγόριθμος Perceptron (III)

- Συνάρτηση κόστους του αλγορίθμου Perceptron:

$$J(\underline{w}) = \sum_{\underline{x} \in Y} (\delta_x \underline{w}^T \underline{x})$$

Όπου Y είναι το υποσύνολο των διανυσμάτων εκπαίδευσης τα οποία ταξινομούνται εσφαλμένα από το τρέχον διάνυσμα $\underline{w}$.

$$\delta_x = -1 \text{ if } \underline{x} \in Y \text{ and } \underline{x} \in \omega_1$$
$$\delta_x = +1 \text{ if } \underline{x} \in Y \text{ and } \underline{x} \in \omega_2$$

$$J(\underline{w}) \geq 0$$

- Όταν το σύνολο Y είναι κενό τότε η ζητούμενη λύση έχει επιτευχθεί και επομένως ισχύει: $J(\underline{w}) = 0$

- Η συνάρτηση J($\underline{w}$) είναι κατά τμήματα γραμμική (βλέπε σχήμα)

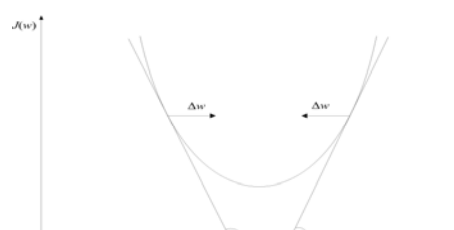
Ο αλγόριθμος Perceptron (IV)

- Ο αλγόριθμος Perceptron βασίζεται στη λογική του αλγορίθμου κατάβασης κατά τη μέγιστη κλίση (gradient descend):

$$\underline{w}(t+1) = \underline{w}(t) + \Delta \underline{w}$$

$\underline{w}(t)$: διάνυσμα βαρών στην t επανάληψη

$$\Delta \underline{w} = -\mu \frac{\partial J(\underline{w})}{\partial \underline{w}} \Big|_{\underline{w} = \underline{w}(t)}$$



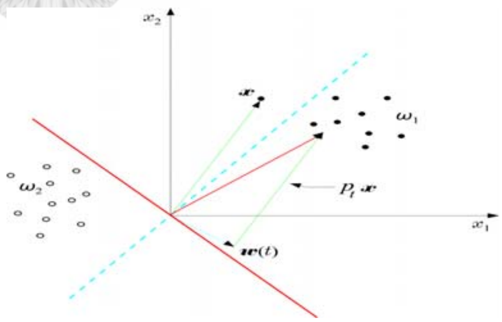
- Στις περιοχές που ορίζεται η μερική παράγωγός έχουμε:

$$\frac{\partial J(\underline{w})}{\partial \underline{w}} = \frac{\partial}{\partial \underline{w}} \left(\sum_{\underline{x} \in Y} \delta_x \underline{w}^T \underline{x} \right) = \sum_{\underline{x} \in Y} \delta_x \underline{x}$$

- Οπότε έχουμε τη βασική σχέση βελτιστοποίησης της συνάρτησης J($\underline{w}$) σύμφωνα με τη μεθοδολογία Perceptron:

$$\boxed{\underline{w}(t+1) = \underline{w}(t) - \rho_t \sum_{\underline{x} \in Y} \delta_x \underline{x}}$$

Παράδειγμα



$$\underline{w}(t+1) = \underline{w}(t) + \rho_t \underline{x}$$
$$= \underline{w}(t) - \rho_t \delta_x \underline{x} \quad (\delta_x = -1)$$

- Ο αλγόριθμος Perceptron συγκλίνει σε πεπερασμένο αριθμό βημάτων όταν τα βήματα προσαρμογής ρ_t πληρούν τις κάτωθι συνθήκες:

$$\lim_{t\to\infty} \sum_{k=0}^{t} \rho_k \to \infty, \qquad \lim_{t\to\infty} \sum_{k=0}^{t} \rho_k^{\,2} < +\infty$$

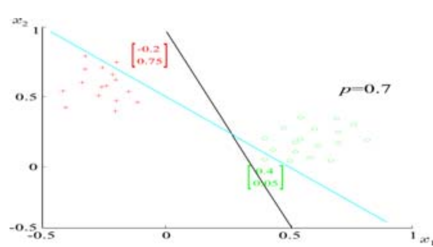
$$\text{Ενδεικτική επιλογή} : \rho_t = \frac{c}{t}$$

Παράδειγμα (II)

- Έστω ότι σε κάποιο στάδιο του αλγορίθμου Perceptron έχουμε:

  $w_1 = 1,\ w_2 = 1,\ w_0 = -0.5$

  - Επομένως η συνάρτηση διαχωρισμού είναι: $g(\mathbf{x}) = x_1 + x_2 - 0.5 = 0 \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
  - Η επιφάνεια διαχωρισμού είναι:



  - Το νέο διάνυσμα των βαρών θα είναι:

$$\underline{w}(t+1) = \begin{bmatrix} 1 \\ 1 \\ -0.5 \end{bmatrix} - 0.7(-1)\begin{bmatrix} 0.4 \\ 0.05 \\ 1 \end{bmatrix} - 0.7(+1)\begin{bmatrix} -0.2 \\ 0.75 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.42 \\ 0.51 \\ -0.5 \end{bmatrix}$$

Παραλλαγές του αλγόριθμου Perceptron

- Μια πολύ χρήσιμη παραλλαγή του αλγορίθμου Perceptron χρησιμοποιεί σταθερό (ως προς τον αριθμό της επανάληψης) βήμα προσαρμογής και επιπλέον βασίζεται μόνο στη ταξινόμηση του τρέχοντος διανύσματος εισόδου $\underline{x}_{(t)}$:

$$\underline{w}(t+1) = \underline{w}(t) + \rho \underline{x}_{(t)}, \quad \begin{array}{c} \underline{w}^T(t)\underline{x}_{(t)} \le 0 \\ \underline{x}_{(t)} \in \omega_1 \end{array}$$

$$\underline{w}(t+1) = \underline{w}(t) - \rho \underline{x}_{(t)}, \quad \begin{array}{c} \underline{w}^T(t)\underline{x}_{(t)} \ge 0 \\ \underline{x}_{(t)} \in \omega_2 \end{array}$$

$$\underline{w}(t+1) = \underline{w}(t) \quad \text{otherwise}$$

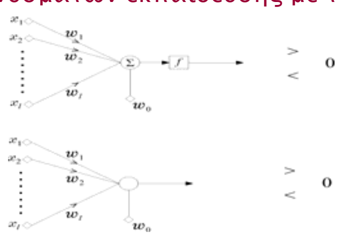
  - Ο ανωτέρω αλγόριθμος υπάγεται σε μια γενικότερη κατηγορία αλγορίθμων οι οποίοι είναι γνωστοί ως αλγόριθμοι «ανταμοιβής και τιμωρίας» (reward and punishment)

ΚΕΣ 03: Αναγνώριση Προτύπων και Ανάλυση Εικόνας

☑ Εισαγωγή
☑ Γραμμικές Συναρτήσεις Διαχωρισμού
★ Ο αλγόριθμος Perceptron
☐ Μέθοδοι Ελαχίστων Τετραγώνων
☐ Support Vector Machines

Μηχανή Perceptron

- Οι Μηχανή Perceptron είναι ένα δίκτυο της μορφής του σχήματος. Πρόκειται για μια εκπαιδευόμενη μηχανή η οποία μαθαίνει από ένα σύνολο διανυσμάτων εκπαίδευσης με τη βοήθεια του αλγορίθμου Perceptron.



w_i's συνάψεις ή συναπτικά βάρη

w_0 κατώφλι

  - Η μηχανή Perceptron είναι πιο γνωστή ως νευρώνας (neuron)
  - Συνδυασμός πολλών νευρώνων δημιουργεί επίσης ένα εκπαιδευόμενο δίκτυο το οποίο είναι γνωστό ως Νευρωνικό Δίκτυα (Neural Networks)
  - Υπάρχουν πολλές παραλλαγές Νευρωνικών Δικτύων ανάλογα με το αλγόριθμο μάθησης που χρησιμοποιούν αλλά και ανάλογα με τη διάταξη και συνδεσμολογία των νευρώνων.

Μέθοδος Ελαχίστων Τετραγώνων

- Όταν οι κλάσεις δεν είναι γραμμικά διαχωρίσιμες τότε σε πολλές περιπτώσεις εξακολουθεί να είναι χρήσιμο να βρούμε μια επιφάνεια διαχωρισμού η οποία να ελαχιστοποιεί τα σφάλματα ταξινόμησης
  - Μια μέθοδος για την εύρεση της καταλληλότερης επιφάνειας διαχωρισμού προκύπτει με την εφαρμογή της μεθόδου των ελαχίστων τετραγώνων
  - Η επιφάνεια διαχωρισμού τοποθετείται έτσι ώστε να ελαχιστοποιείται η απόσταση όλων των διανυσμάτων εκπαίδευσης από την επιφάνεια
- Το πρόβλημα τίθεται ως ακολούθως:
  - Υπολόγισε το διάνυσμα βαρών $\underline{w}$ έτσι ώστε η διαφορά ανάμεσα στην επιθυμητή έξοδο y (κλάσεις $\begin{matrix} +1 \text{ if } \underline{x} \in \omega_1 \\ -1 \text{ if } \underline{x} \in \omega_2 \end{matrix}$) και στην πραγματική έξοδο $y' = \underline{w}^T \underline{x}$ να ελαχιστοποιείται.

ΚΕΣ 03: Αναγνώριση Προτύπων και Ανάλυση Εικόνας

Κριτήριο κόστους για τη μέθοδο των ελαχίστων τετραγώνων

- Ελαχιστοποίηση πραγματοποιείται υπό την έννοια του μέσου τετραγώνου (mean square), επομένως το κριτήριο κόστους είναι:

$$J(\underline{w}) \equiv E[(y - \underline{w}^T \underline{x})^2]$$

όπου y είναι η επιθυμητή έξοδος

  - Το διάνυσμα βαρών που ελαχιστοποιεί το μέσω τετραγωνικό σφάλμα δίνεται από τη σχέση:

$$\hat{\underline{w}} = \arg\min_{\underline{w}} J(\underline{w})$$

Ελαχιστοποίηση μέσου τετραγωνικού σφάλματος

- Η ελαχιστοποίηση του κριτήριου κόστους επιτυγχάνεται με μηδενισμό των μερικών παραγώγων:

$$\frac{\partial J(\underline{w})}{\partial \underline{w}} = \frac{\partial}{\partial \underline{w}} E[(y - \underline{w}^T x)^2] = 0$$

$$= 2E[\underline{x}(y - \underline{x}^T \underline{w})] \Rightarrow E[\underline{x}\underline{x}^T]\underline{w} = E[\underline{x}y] \Rightarrow$$

$$\hat{\underline{w}} = R_x^{-1} E[\underline{x}y]$$

Όπου ο πίνακας Rx είναι ο πίνακας αυτοσυσχέτισης:

$$R_x \equiv E[\underline{x}\underline{x}^T] = \begin{bmatrix} E[x_1x_1] & E[x_1x_2]\ldots & E[x_1x_l] \\ \ldots\ldots\ldots . & \ldots\ldots\ldots \ldots & \ldots\ldots\ldots . \\ E[x_lx_1] & E[x_lx_2]\ldots & E[x_lx_l] \end{bmatrix}$$

και $E[\underline{x}y] = \begin{bmatrix} E[x_1y] \\ \ldots \\ E[x_ly] \end{bmatrix}$ είναι το διάνυσμα ετεροσυσχέτισης εισόδου εξόδου

Παράδειγμα

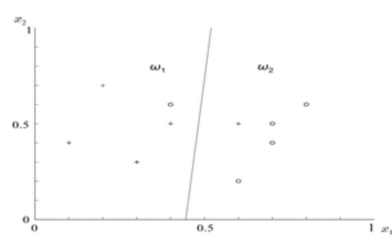
- Στην πράξη οι πίνακες R_x και R_{xy} = E[<u>x</u>y] εκτιμούνται από τα διανύσματα εκπαίδευσης:

$$\hat{R}_x = \sum_{i=1}^{N}\sum_{j=1}^{N} \mathbf{x}_i \mathbf{x}_j^T \qquad \hat{R}_{xy} = \sum_{i=1}^{N} \mathbf{x}_i y_i$$

- Να υπολογιστεί η συνάρτηση διαχωρισμού των κλάσεων ω_1, ω_2, με βάση τα διανύσματα εκπαίδευσης που δίνονται παρακάτω, με τη μέθοδο των ελαχίστων τετραγώνων

$$\omega_1 : \begin{bmatrix} 0.4 \\ 0.5 \end{bmatrix}, \begin{bmatrix} 0.6 \\ 0.5 \end{bmatrix}, \begin{bmatrix} 0.1 \\ 0.4 \end{bmatrix}, \begin{bmatrix} 0.2 \\ 0.7 \end{bmatrix}, \begin{bmatrix} 0.3 \\ 0.3 \end{bmatrix}$$

$$\omega_2 : \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}, \begin{bmatrix} 0.6 \\ 0.2 \end{bmatrix}, \begin{bmatrix} 0.7 \\ 0.4 \end{bmatrix}, \begin{bmatrix} 0.8 \\ 0.6 \end{bmatrix}, \begin{bmatrix} 0.7 \\ 0.5 \end{bmatrix}$$



Παράδειγμα (συν.)

- Σχηματίζουμε τον πίνακα X για να εκτιμήσουμε τον πίνακα αυτοσυσχέτισης:

$$\hat{R}_x = \sum_{i=1}^{N}\sum_{j=1}^{N} \mathbf{x}_i \mathbf{x}_j^T = X^T X$$

και τον πίνακα ετεροσυσχέτισης

$$\hat{R}_{xy} = \sum_{i=1}^{N} \mathbf{x}_i y_i = X^T y$$

$$X = \begin{bmatrix} 0.4 & 0.5 & 1 \\ 0.6 & 0.5 & 1 \\ 0.1 & 0.4 & 1 \\ 0.2 & 0.7 & 1 \\ 0.3 & 0.3 & 1 \\ 0.4 & 0.6 & 1 \\ 0.6 & 0.2 & 1 \\ 0.7 & 0.4 & 1 \\ 0.8 & 0.6 & 1 \\ 0.7 & 0.5 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} = \underline{y}$$

$$\hat{R}_x = X^T X = \begin{bmatrix} 2.8 & 2.24 & 4.8 \\ 2.24 & 2.41 & 4.7 \\ 4.8 & 4.7 & 10 \end{bmatrix}, \hat{R}_{xy} = X^T \underline{y} = \begin{bmatrix} -1.6 \\ 0.1 \\ 0.0 \end{bmatrix}$$

$$\underline{w} = (X^T X)^{-1} X^T \underline{y} = \hat{R}_x^{-1} \hat{R}_{xy} = \begin{bmatrix} -3.13 \\ 0.24 \\ 1.34 \end{bmatrix}$$