Θεόδωρος Αλεξόπουλος, Αναπληρωτής Καθηγητής
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ - ΤΟΜΕΑΣ ΦΥΣΙΚΗΣ
ΗΡΩΩΝ ΠΟΛΥΤΕΧΝΕΙΟΥ 9
ΑΘΗΝΑ 157 80
Τηλ: 210 772-3019, Fax: 210 772-3025
e-mail: theoalex@central.ntua.gr

Theodoros Alexopoulos, Associate Professor
NATIONAL TECHNICAL UNIVERSITY
DEPARTMENT OF PHYSICS
ZOGRAFOU CAMPUS
157 80 ATHENS - GREECE
Phone : +30 210 772-3019, Fax: +30 210 772-3025
e-mail: Theodoros.Alexopoulos@cern.ch
html://www.physics.ntua.gr/Faculty/theoalex

Αναγνώριση Προτύπων & Νευρωνικά Δίκτυα

Προβλήματα 2

(Επιστροφή: 12 Ιανουαρίου 2004)

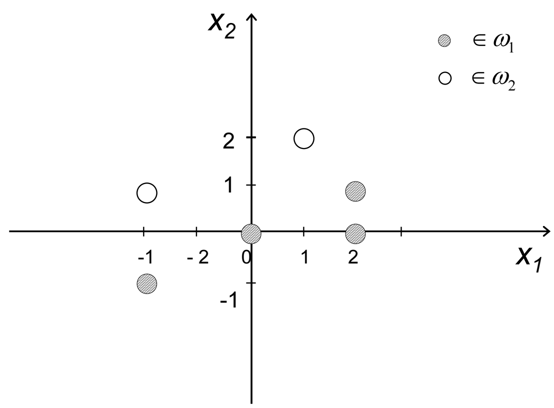
0.1

Υπάρχουν διάφοροι τρόποι για να γενικεύσουμε την ιδέα των συναρτήσεων διάκρισης δύο κλάσεων σε C κλάσεις $\omega_1, \omega_2, \ldots, \omega_k$. Ένας τρόπος θα είναι να χρησιμοποιήσουμε $(C-1)$ συναρτήσεις διάκρισης, έτσι ώστε αν $g_k(\overrightarrow{x}) > 0$ τότε το δείγμα $\overrightarrow{x} \epsilon \omega_k$, ενώ αν $g_k(\overrightarrow{x}) < 0$ τότε $\overrightarrow{x} \notin \omega_k$. Με τη βοήθεια ενός παραδείγματος σε δύο διαστάσεις για $C = 3$ (τρεις κλάσεις), να δείξετε ότι αυτός ο τρόπος ταξινόμησης μπορεί να οδηγήσει σε περιοχές στο δειγματοχώρο των $\overrightarrow{x}$ για τις οποίες η ταξινόμηση να είναι ασαφής. Ένας άλλος τρόπος είναι να χρησιμοποιήσουμε μια συνάρτηση διάκρισης $g_{jk}(\overrightarrow{x})$ για κάθε δυνατό ζεύγος κλάσεων ω_j και ω_k, έτσι ώστε για $g_{jk}(\overrightarrow{x}) > 0$ το $\overrightarrow{x} \epsilon \omega_j$ και για $g_{jk}(\overrightarrow{x}) < 0$ το $\overrightarrow{x} \epsilon \omega_k$. Για C κλάσεις, απαιτούνται $C(C-1)/2$ συναρτήσεις διάκρισης. Με ένα απλό παράδειγμα σε δύο διαστάσεις για $C = 3$, να δείξετε ότι και αυτός ο τρόπος θα οδηγήσει σε ασαφείς περιοχές ταξινόμησης.

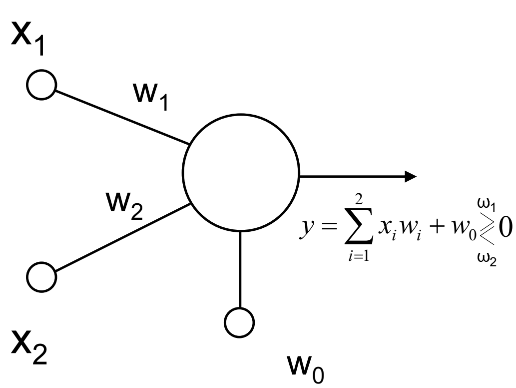
0.2

Έστω τα 6 δείγματα εκπαίδευσης που ανήκουν σε δύο κλάσεις ω_1, ω_2, όπως φαίνονται στο σχήμα (1):

$$S = \left\{ \underbrace{\begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}}_{\omega_2}, \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}}_{\omega_1} \right\}.$$



Σχήμα 1:



Σχήμα 2:

Να ορίσετε ένα νευρώνα (perceptron) που να ορίσει την ταξινόμηση αυτών των δύο κλάσεων. Δηλαδή, ορίστε τις συνάψεις w_1, w_2 και το κατώφλι w_0.

0.3

Θεωρείστε τον αλγόριθμο της απότομης πτώσης (steepest descent) όπου η συνάρτηση κόστους $J(w) = k_1(w - w_0)^2 + k_2$.
(α) Βρείτε τη βέλτιστη λύση w^* που ελαχιστοποιεί τη συνάρτηση κόστους $J(w)$.
(β) Με τη βοήθεια του αλγορίθμου της απότομης πτώσης:

$$w(i+1) = w(i) + \varrho \frac{dJ}{dw},$$

να βρείτε μια αναλυτική έκφραση του $w(i)$. Θυμηθείτε πώς επιλύαμε διαφοροεξισώσεις από την Ανάλυση Σήματος! Να βρείτε τη συνθήκη που πρέπει να υπακούει η παράμετρος του ρυθμού εκμάθησης ϱ, ώστε ο αλγόριθμος της απότομης πτώσης να συγκλίνει.
(γ) Να βρείτε το όριο του $w(i)$ για $i \to \infty$.

0.4

Έστω $\overrightarrow{x}$ είναι μια τυχαία μεταβλητή με μέση τιμή $\overrightarrow{\mu}$ και πίνακα διασποράς $\overline{\overline{\Sigma}}$. Να δείξετε ότι ο $\overline{\overline{\Sigma}}$ είναι θετικά ορισμένος, δηλαδή να δείξετε ότι:

$$\overrightarrow{y}^t \overline{\overline{\Sigma}} \overrightarrow{y} > 0 \ \forall \ \overrightarrow{y} \neq 0.$$

0.5

Έστω το διάνυσμα $\overrightarrow{x} = (x_1, x_2)^t$ είναι κατανεμημένο κατά Gauss με πυκνότητες πιθανότητας $p(\overrightarrow{x}/\omega_i) \sim N(\overrightarrow{\mu}_i, \sigma^2 I)$. Να σχεδιάσετε τις επιφάνειες απόφασης για ένα ταξινομητή ελάχιστης ευκλείδειας απόστασης αν έχουμε 5 κλάσεις με τις ακόλουθες μέσες τιμές:

$$\overrightarrow{\mu}_1 = (0\ 0)^t,$$

$$\overrightarrow{\mu}_2 = (2\ 0)^t,$$

$$\overrightarrow{\mu}_3 = (-2\ 0)^t,$$

$$\overrightarrow{\mu}_4 = (2\ 2)^t,$$

$$\overrightarrow{\mu}_5 = (2 \ \ -2)^t.$$

0.6

Δίνονται τα παρακάτω διανύσματα εκπαίδευσης:

$$S = \Bigg\{ \underbrace{\begin{pmatrix} 0,1 \\ 0,2 \end{pmatrix}, \begin{pmatrix} 0,2 \\ 0,1 \end{pmatrix}, \begin{pmatrix} -0,15 \\ 0,2 \end{pmatrix}, \begin{pmatrix} 1,1 \\ 0,8 \end{pmatrix}, \begin{pmatrix} 1,2 \\ 1,1 \end{pmatrix}}_{\omega_1},$$

$$\underbrace{\begin{pmatrix} 1,1 \\ -0,1 \end{pmatrix}, \begin{pmatrix} 1,25 \\ 0,15 \end{pmatrix}, \begin{pmatrix} 0,9 \\ 0,1 \end{pmatrix}, \begin{pmatrix} 0,1 \\ 1,2 \end{pmatrix}, \begin{pmatrix} 0,2 \\ 0,9 \end{pmatrix}}_{\omega_2} \Bigg\}.$$

Να δείξετε γραφικά ότι αυτά είναι γραμμικά διαχωρίσιμα γεγονότα, και να σχεδιαστεί μια κατάλληλη αρχιτεκτονική Perceptron που να τα διαχωρίζει.

0.7

Υποθέστε ότι δύο σύνολα S_1 και S_2 στο χώρο $\mathbb{R}^l$ είναι γραμμικά διαχωρίσιμα. Δηλαδή, υπάρχει ένα διάνυσμα $\overrightarrow{w}' \in \mathbb{R}^l$ και ένα βαθμωτό $w_0 \in R$, έτσι ώστε:

$$\overrightarrow{w}'^t \overrightarrow{x} + w_0 = \begin{cases} > 0 & \forall \ \overrightarrow{x} \in S_1 \\ < 0 & \forall \ \overrightarrow{x} \in S_2 \end{cases}$$

Θεωρείστε τα νέα διανύσματα ως ακολούθως:

$$\overrightarrow{x} = \begin{pmatrix} \overrightarrow{x} \\ 1 \end{pmatrix}, \qquad \overrightarrow{w} = \begin{pmatrix} \overrightarrow{w}' \\ w_0 \end{pmatrix}.$$

(α) Να δείξετε ότι $\overrightarrow{w}^t \hat{x} > 0 \ \forall \hat{x}$, αν το πρόσημο των διανυσμάτων $\hat{x}$ του συνόλου S_2 αλλάζει.
(β) Να περιγράψετε γραφικά τί συμβαίνει στο ερώτημα (α) όταν $l = 1, S_1 = \{-3, -2, -1\}$, και $S_2 = \{5, 6, 7\}$.

0.8

Στον αλγόριθμο του νευρώνα (Perceptron) έχουμε:

$$\overrightarrow{w}(i+1) = \overrightarrow{w}(i) + \varrho(i)\overrightarrow{x}_i,$$

όταν το $\overrightarrow{x}_i$ έχει ταξινομηθεί λανθασμένα, δηλαδή,

$$\overrightarrow{w}^t\overrightarrow{x}_i < 0.$$

Να βρείτε τη συνθήκη που διέπει την παράμετρο του ρυθμού εκμάθησης $\varrho(i)$, έτσι ώστε το $\overrightarrow{x}$ να ταξινομηθεί ορθά στο βήμα $i+1$.

0.9

Έστω ότι τα σύνολα S_1, S_2 είναι γραμμικά διαχωρίσιμα, και έχουμε γεγονότα εκπαίδευσης $x_1 = -2 \in S_1$, $x_2 = -1 \in S_2$ και $x_3 = 2 \in S_2$. Η αρχική τιμή του διανύσματος $\overrightarrow{w}$ είναι $\overrightarrow{w}(0) = (1,1)^t$, και η παράμετρος του ρυθμού εκμάθησης είναι $\varrho = 0,5$. Να προσδιοριστεί το διάνυσμα $\overrightarrow{w}$ ώστε ο αλγόριθμος του Perceptron να μας δώσει την ορθή ταξινόμηση.

0.10

(α) Να εφαρμόσετε τον απλό αλγόριθμο εύρεσης των προτύπων για τα διανύσματα:

$$S = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -5 \\ -4 \end{pmatrix}, \begin{pmatrix} -5 \\ -5 \end{pmatrix}, \begin{pmatrix} -4 \\ -5 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}.$$

Θεωρείστε ότι το κατώφλι $T = 3$.
(β) Να εφαρμόσετε τη μέθοδο $MaxMin$ στο δειγματοχώρο S του ερωτήματος (α).
(γ) Να εφαρμόσετε τη μέθοδο K-μέσου (K-means) στο δειγματοχώρο S του ερωτήματος (α).