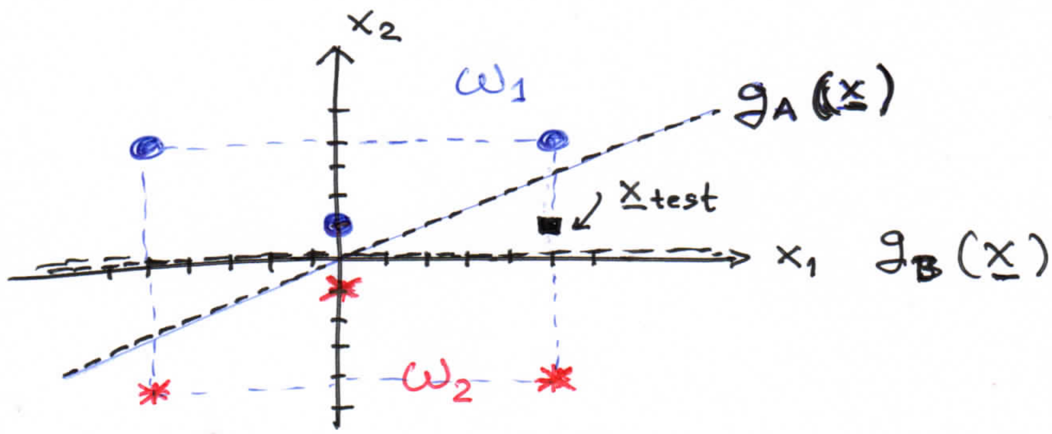
2 a



- Σχεδιάσαμε τα έξι διανύσματα εκπαίδευσης και το διάνυσμα δοκιμής. Είναι προφανές ότι οι κλάσεις είναι γραμμικά διαχωρίσιμες.

- Μια από τις (άπειρες!) δυνατές γραμμικές συναρτήσεις διάκρισης δίνεται στο σχήμα. Είναι η:

$$\boxed{g_A(\underline{x}): \; -x_1 + 2x_2 = 0}, \text{ δηλ } \underline{w} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_0 \end{bmatrix}$$

(περνάει από τα $[0,0]^T$ κ $[5, 2.5]^T$). Παρατηρούμε ότι $g_A(\underline{x}_{test}) = -4+2 = -2 < 0$

δηλαδή ταξινομεί το $\underline{x}_{test}$ στην ω_2 κλάση

2b Είναι προφανές ότι η αρχική λύση $\underline{w}(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ ("degenerate") ταξινομεί όλα τα δείγματα εκπαίδευσης λανθασμένα. Άρα:

$$\underline{w}(1) = \underline{w}(0) - \rho \sum_{i=1}^{6} \delta_{x_i} \underline{x}_i =$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \rho \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \rho \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} + \rho \begin{bmatrix} -5 \\ 4 \\ 1 \end{bmatrix}$$

$$- \rho \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} - \rho \begin{bmatrix} 5 \\ -4 \\ 1 \end{bmatrix} - \rho \begin{bmatrix} -5 \\ -4 \\ 1 \end{bmatrix}$$

$$= \rho \begin{bmatrix} 0 \\ 18 \\ 0 \end{bmatrix} \overset{\rho = 1/6}{=} \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$$

Παρατηρούμε ότι η ευθεία που βρέθηκε είναι η

$$\boxed{g_B(\underline{x}): \; 3x_2 = 0}$$

η οποία ισαι ταξινομεί όλα τα δείγματα εκπαίδευσης σωστά! Άρα ο αλγόριθμος συγκλίνει με 1 επανάληψη.

Εύκολα επίσης βλέπουμε ότι $\underline{x}_{test} \in \omega_1$, καθώς $g_B(\underline{x}_{test}) > 0$.

2c • Παρατηρούμε εύκολα ότι η ευθεία

$$g_c(\underline{x}): \boxed{x_2 = 0} \rightsquigarrow \underline{w} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{array}{l} \to w_1 \\ \to w_2 \\ \to w_0 \end{array}$$

διαχωρίζει σωστά το σύνολο εκπαίδευσης, επιτυγχάνοντας το μέγιστο δυνατό περιθώριο (MARGIN = 2)

- Τα διανύσματα που απέχουν $\frac{2}{2} = 1$ από την ευθεία διαχωρισμού είναι τα <u>SUPPORT VECTORS</u>, δηλαδή τα $[0, 1]^T$ κ $[0, -1]^T$.

- Βλέπουμε επίσης ότι:

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \underbrace{\lambda_1}_{1/2} \underbrace{y_1}_{+1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \underbrace{\lambda_2}_{1/2} \underbrace{y_2}_{-1} \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

με $\lambda_1 y_1 + \lambda_2 y_2 = 0$

- Το ότι $w_0 = 0$ φαίνεται από την εξίσωση

$$[\underbrace{w_1}_{0} \ \underbrace{w_2}_{1}] \begin{bmatrix} 0 \\ 1 \end{bmatrix} + w_0 = +1 \Leftrightarrow w_0 = 0$$

(όμοια προκύπτει αν χρησιμοποιηθεί το $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$).

- Το $\underline{x}_{test}$ ταξινομείται στην κλάση ω_1, καθώς $g_c(\underline{x}_{test}) > 0$.

2d Παρατηρούμε πρώτα ότι η μέση τιμή των δεδομένων εκπαίδευσης είναι μηδενική: $\frac{1}{6}\sum_{i=1}^{6} \underline{x}_i = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

- Στη συνέχεια βρίσκουμε τον πίνακα συνδιασποράς των δεδομένων (ίσο με τον πίνακα ετεροσυσχέτισης λόγω της μηδενικής μέσης τιμής):

$$\frac{1}{6}\sum_{i=1}^{6} \underline{x}_i \underline{x}_i^T = \frac{1}{6}\begin{bmatrix} 0 & 5 & -5 & 0 & 5 & -5 \\ 1 & 4 & 4 & -1 & -4 & -4 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 5 & 4 \\ -5 & 4 \\ 0 & -1 \\ 5 & -4 \\ -5 & -4 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 100 & 0 \\ 0 & 66 \end{bmatrix}$$

- Βλέπουμε ότι ο πίνακας είναι ήδη διαγώνιος με μέγιστη ιδιοτιμή την 100/6. Το αντίστοιχο ιδιοδιάνυσμα είναι το $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

- Παρατηρούμε ότι στην κατεύθυνση αυτή τα διανύσματα εκπαίδευσης της ω_1 προβάλλονται στα μονοδιάστατα σημεία $-5, 1, 5$, ενώ της ω_2 στα $-5, 1, 5$.

- Προφανώς οι δύο κλάσεις δεν είναι διαχωρίσιμες στον 1-D χώρο του PCA, καθώς τα διανύσματα εκπαίδευσης των δύο κλάσεων επικαλύπτονται (πλήρως!)

$\boxed{2e}$ Βρίσκουμε πρώτα τις μέσες τιμές των διανυσμάτων κάθε κλάσης:

$$\underline{\mu}_1 = \frac{1}{3}\sum_{i=1}^{3} \underline{x}_i = \frac{1}{3}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ 4 \end{bmatrix} + \begin{bmatrix} -5 \\ 4 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$$\underline{\mu}_2 = \frac{1}{3}\sum_{i=4}^{6} \underline{x}_i = \frac{1}{3}\left(\begin{bmatrix} 0 \\ -1 \end{bmatrix} + \begin{bmatrix} 5 \\ -4 \end{bmatrix} + \begin{bmatrix} 5 \\ -4 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$$

Στη συνέχεια αφαιρούμε τη μέση τιμή από τα διανύσματα κάθε κλάσης και υπολογίζουμε τους πίνακες Σ_1 & Σ_2

$$\Sigma_1 = \frac{1}{3}\sum_{i=1}^{3} (\underline{x} - \underline{\mu}_1)(\underline{x} - \underline{\mu}_1)^T =$$

$$= \frac{1}{3}\begin{bmatrix} 0 & 5 & -5 \\ -2 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & -2 \\ 5 & 1 \\ -5 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 50 & 0 \\ 0 & 6 \end{bmatrix}$$

$$\Sigma_2 = \frac{1}{3}\sum_{i=4}^{6} (\underline{x} - \underline{\mu}_2)(\underline{x} - \underline{\mu}_2)^T =$$

$$= \frac{1}{3}\begin{bmatrix} 0 & 5 & -5 \\ 2 & -1 & -1 \end{bmatrix}\begin{bmatrix} 0 & 2 \\ 5 & -1 \\ -5 & -1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 50 & 0 \\ 0 & 6 \end{bmatrix}$$

Συνεπώς $$\Sigma_W = \frac{1}{2}\left(\Sigma_1 + \Sigma_2\right) = \frac{1}{3}\begin{bmatrix} 50 & 0 \\ 0 & 6 \end{bmatrix}$$

$P(\omega_1) = P(\omega_2) = \frac{1}{2}$

2e - συνέχεια

... και κατά συνέπεια η κατεύθυνση προβολής σε 1-D

είναι η:

$$\underline{w} = \Sigma_w^{-1}(\underline{\mu}_1 - \underline{\mu}_2) = 3\begin{bmatrix} \frac{1}{50} & 0 \\ 0 & \frac{1}{6} \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 6 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

- Στην κατεύθυνση αυτή, οι προβολές των διανυσμάτων, $\underline{w}^T \underline{x}_i / \|w\|$ είναι: **1, 4, 4** για την ω_1 και **−1, −4, −4** για την κλάση ω_2, προφανώς λοιπόν οι κλάσεις είναι διαχωρίσιμες. Βλέπουμε επίσης ότι το $\underline{x}_{test}$ προβάλλεται στο **+1**, δηλαδή στην περιοχή προβολής της ω_1.