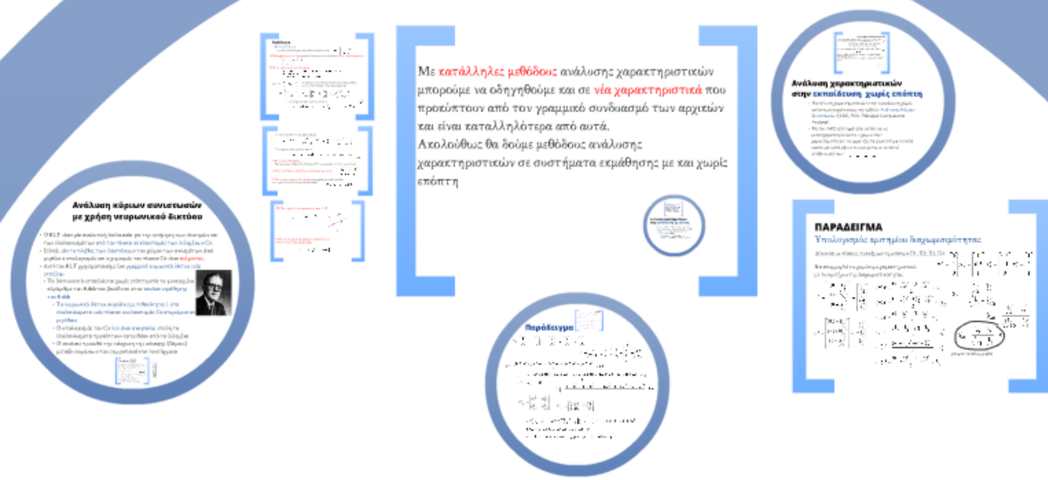


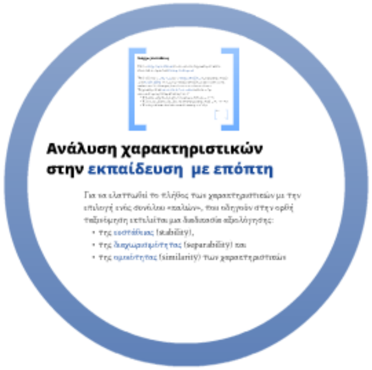
Ανάλυση χαρακτηριστικών

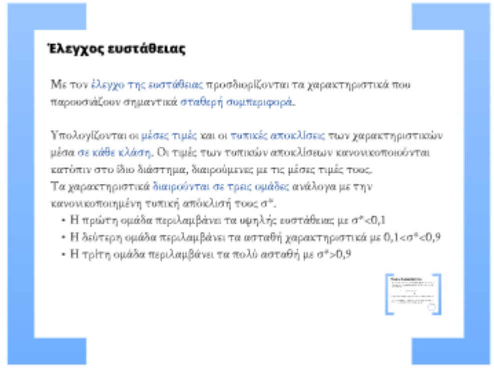
επιτυγχάνεται

- αξιολόγηση των χαρακτηριστικών
- επιλογή των καταλληλότερων
- ελάττωση του πλήθους των χαρακτηριστικών

Με κατάλληλες μεθόδους ανάλυσης χαρακτηριστικών μπορούμε να οδηγηθούμε και σε νέα χαρακτηριστικά που προκύπτουν από τον γραμμικό συνδυασμό των αρχικών και είναι καταλληλότερα από αυτά.

Ακολούθως θα δούμε μεθόδους ανάλυσης χαρακτηριστικών σε συστήματα εκμάθησης με και χωρίς επόπτη





Έλεγχος ευστάθειας

Με τον έλεγχο της ευστάθειας προσδιορίζονται τα χαρακτηριστικά που παρουσιάζουν σημαντικά σταθερή συμπεριφορά.

Υπολογίζονται οι μέσες τιμές και οι τυπικές αποκλίσεις των χαρακτηριστικών μέσα σε κάθε κλάση. Οι τιμές των τυπικών αποκλίσεων κανονικοποιούνται κατόπιν στο ίδιο διάστημα, διαιρούμενες με τις μέσες τιμές τους.
Τα χαρακτηριστικά διαιρούνται σε τρεις ομάδες ανάλογα με την κανονικοποιημένη τυπική απόκλισή τους σ^*.

- Η πρώτη ομάδα περιλαμβάνει τα υψηλής ευστάθειας με $\sigma^*<0,1$
- Η δεύτερη ομάδα περιλαμβάνει τα ασταθή χαρακτηριστικά με $0,1<\sigma^*<0,9$
- Η τρίτη ομάδα περιλαμβάνει τα πολύ ασταθή με $\sigma^*>0,9$

Ανάλυση χαρακτηριστικών στην εκπαίδευση με επόπτη

Για να ελαττωθεί το πλήθος των χαρακτηριστικών με την επιλογή ενός συνόλου «καλών», που οδηγούν στην ορθή ταξινόμηση εκτελείται μια διαδικασία αξιολόγησης:

- της ευστάθειας (stability),
- της διαχωρισιμότητας (separability) και
- της ομοιότητας (similarity) των χαρακτηριστικών

Έλεγχος ευστάθειας

Με τον έλεγχο της ευστάθειας προσδιορίζονται τα χαρακτηριστικά που παρουσιάζουν σημαντικά σταθερή συμπεριφορά.

Υπολογίζονται οι μέσες τιμές και οι τυπικές αποκλίσεις των χαρακτηριστικών μέσα σε κάθε κλάση. Οι τιμές των τυπικών αποκλίσεων κανονικοποιούνται κατόπιν στο ίδιο διάστημα, διαιρούμενες με τις μέσες τιμές τους.
Τα χαρακτηριστικά διαιρούνται σε τρεις ομάδες ανάλογα με την κανονικοποιημένη τυπική απόκλισή τους σ^*.

- Η πρώτη ομάδα περιλαμβάνει τα υψηλής ευστάθειας με $\sigma^*<0,1$
- Η δεύτερη ομάδα περιλαμβάνει τα ασταθή χαρακτηριστικά με $0,1<\sigma^*<0,9$
- Η τρίτη ομάδα περιλαμβάνει τα πολύ ασταθή με $\sigma^*>0,9$

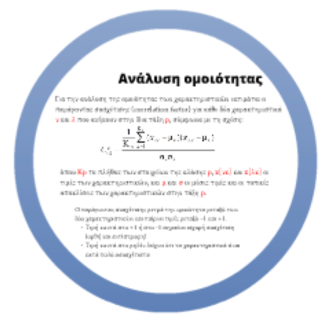
Έλεγχος διαχωρισιμότητας

Για τον έλεγχο της ικανότητας διαχωρισμού καθορίζεται ο παράγοντας διαχωριστικότητας (seperability factor) μεταξύ δύο κλάσεων για κάθε χαρακτηριστικό:

$$S_{\nu} = \frac{\left|\mu_{\nu}^{i} - \mu_{\nu}^{j}\right|}{\sqrt{(\sigma_{\nu}^{i})^{2} + (\sigma_{\nu}^{j})^{2}}}$$

όπου μ και σ είναι οι μέσες τιμές και οι τυπικές αποκλίσεις σε κάθε κλάση.

Μεγάλη διαχωρισιμότητα σημαίνει ότι το χαρακτηριστικό έχει μεγάλη ικανότητα να διαχωρίζει τις δύο κλάσεις μεταξύ τους.



Ανάλυση ομοιότητας

Για την ανάλυση της ομοιότητας των χαρακτηριστικών εκτιμάται ο παράγοντας συσχέτισης (correlation factor) για κάθε δύο χαρακτηριστικά ν και λ που ανήκουν στην ίδια τάξη p, σύμφωνα με τη σχέση:

$$C^{p}_{\nu\lambda} = \frac{\frac{1}{K_p}\sum_{\kappa=1}^{K_p}(x_{\nu\kappa}-\mu_\nu)(x_{\lambda\kappa}-\mu_\lambda)}{\sigma_\nu\sigma_\lambda}$$

όπου Kp το πλήθος των στοιχείων της κλάσης p, x[νκ] και x[λκ] οι τιμές των χαρακτηριστικών, και μ και σ οι μέσες τιμές και οι τυπικές αποκλίσεις των χαρακτηριστικών στην τάξη p.

Ο παράγοντας συσχέτισης μετρά την ομοιότητα μεταξύ των δύο χαρακτηριστικών και παίρνει τιμές μεταξύ –1 και +1.

- Τιμή κοντά στο +1 ή στο -1 σημαίνει ισχυρή συσχέτιση (ορθή και αντίστροφη)
- Τιμή κοντά στο μηδέν δείχνει ότι τα χαρακτηριστικά είναι κατά πολύ ασυσχέτιστα

ΠΑΡΑΔΕΙΓΜΑ

Υπολογισμός κριτηρίου διαχωρισιμότητας

Δίνονται οι πίνακες τεσσάρων προτύπων Π1, Π2, Π3, Π4

$$x_1^A = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, x_2^A = \begin{bmatrix} 1 \\ 0.5 \\ 1 \end{bmatrix}, x_3^B = \begin{bmatrix} 0.5 \\ -1 \\ 0 \end{bmatrix}, x_4^B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Να απορριφθεί το χειρότερο χαρακτηριστικό με το κριτήριο της διαχωριστικότητας

$$\mu^A = \begin{bmatrix} (0+1)/2 \\ (1+0.5)/2 \\ (0+1)/2 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 3/4 \\ 1/2 \end{bmatrix}$$

$$\mu^B = \begin{bmatrix} (0.5+1)/2 \\ (-1+0)/2 \\ (0+1)/2 \end{bmatrix} = \begin{bmatrix} 3/4 \\ -1/2 \\ 1/2 \end{bmatrix}$$

$$(\sigma_1^A)^2 = \frac{(0-1/2)^2+(1-1/2)^2}{2} = 1/4$$

$$(\sigma_2^A)^2 = \frac{(1-3/4)^2+(1/2-3/4)^2}{2} = 1/16$$

$$(\sigma_3^A)^2 = \frac{(0-1/2)^2+(1-1/2)^2}{2} = 1/4$$

$$(\sigma_1^B)^2 = \frac{(1/2-3/4)^2+(1-3/4)^2}{2} = 1/16$$

$$(\sigma_2^B)^2 = \frac{(-1+1/2)^2+(0-1/2)^2}{2} = 1/4$$

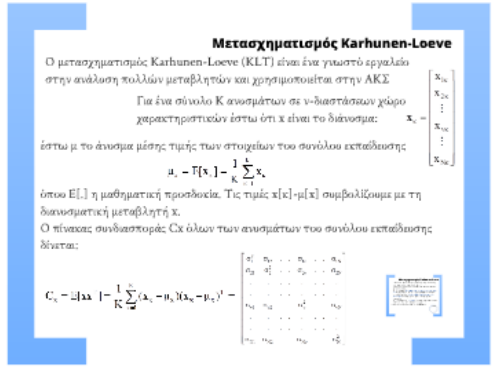
$$(\sigma_3^B)^2 = \frac{(0-1/2)^2+(1-1/2)^2}{2} = 1/4$$

$$S_{AB}^1 = \frac{\left|1/2 - 3/4\right|}{\sqrt{1/4 + 1/16}}$$

$$S_{AB}^2 = \frac{\left|3/4 + 1/2\right|}{\sqrt{1/16 + 1/4}}$$

$$S_{AB}^3 = \frac{\left|1/2 - 1/2\right|}{\sqrt{1/4 + 1/4}} = 0$$

μπορεί να απορριφθεί



Μετασχηματισμός Karhunen-Loeve

Ο μετασχηματισμός Karhunen-Loeve (KLT) είναι ένα γνωστό εργαλείο στην ανάλυση πολλών μεταβλητών και χρησιμοποιείται στην ΑΚΣ

Για ένα σύνολο Κ ανυσμάτων σε ν-διαστάσεων χώρο χαρακτηριστικών έστω ότι x είναι το διάνυσμα:

$$\mathbf{x}_\kappa = \begin{bmatrix} x_{1\kappa} \\ x_{2\kappa} \\ \vdots \\ x_{\nu\kappa} \\ \vdots \\ x_{N\kappa} \end{bmatrix}$$

έστω μ το άνυσμα μέσης τιμής των στοιχείων του συνόλου εκπαίδευσης

$$\mu_x = E[x_\kappa] = \frac{1}{K}\sum_{\kappa=1}^{K} x_\kappa$$

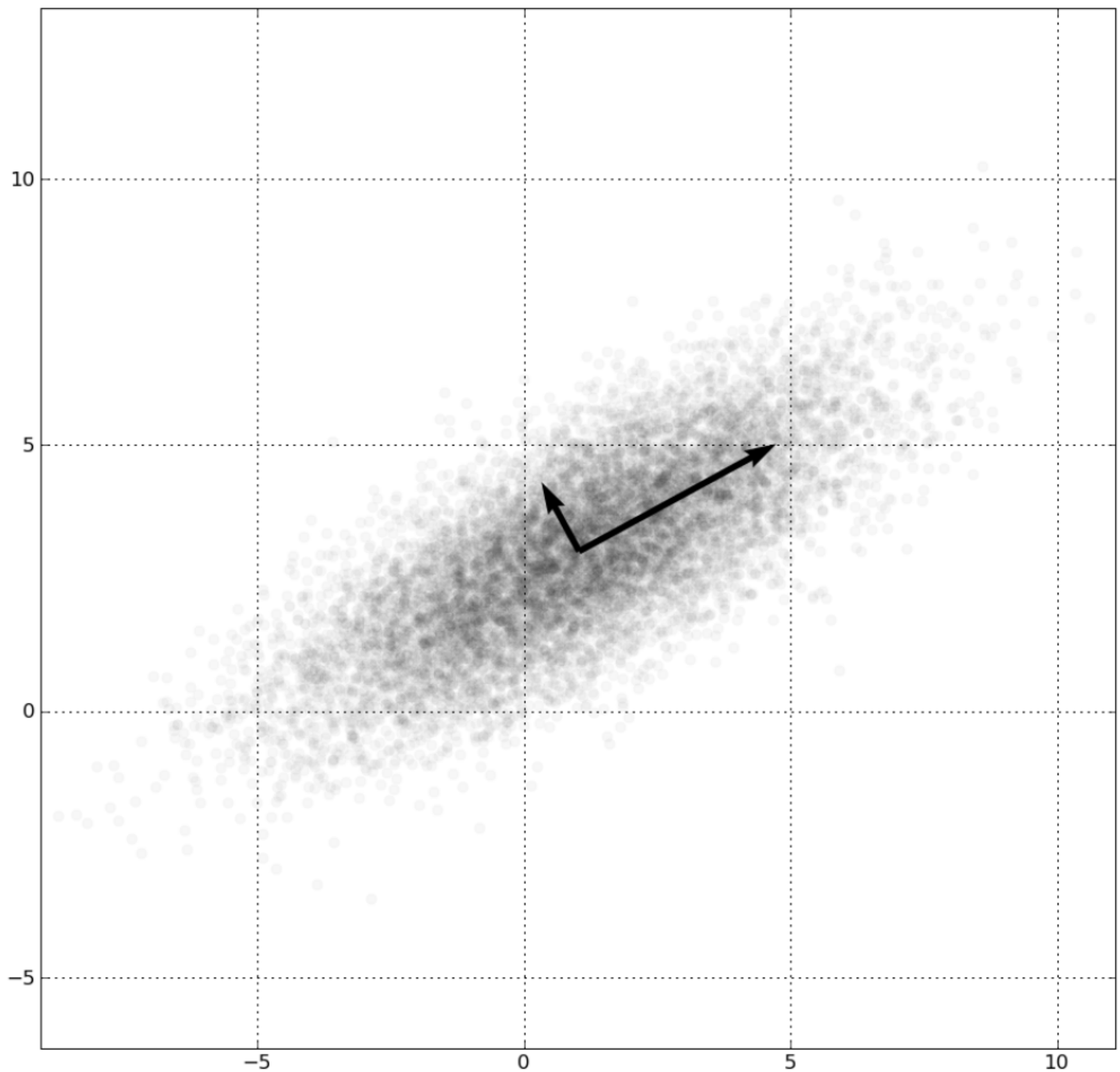
όπου E[.] η μαθηματική προσδοκία. Τις τιμές x[κ]-μ[x] συμβολίζουμε με τη διανυσματική μεταβλητή x.

Ο πίνακας συνδιασποράς Cx όλων των ανυσμάτων του συνόλου εκπαίδευσης δίνεται:

$$C_x = E[xx^T] = \frac{1}{K}\sum_{\kappa=1}^{K}(x_\kappa - \mu_x)(x_\kappa - \mu_x)^T = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1\nu} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2\nu} & \cdots & \sigma_{2N} \\ \vdots & & & & & \vdots \\ \sigma_{\nu 1} & \sigma_{\nu 2} & & \sigma_\nu^2 & & \sigma_{\nu N} \\ \vdots & & & & & \vdots \\ \sigma_{N1} & \sigma_{N2} & & \sigma_{N\nu} & & \sigma_N^2 \end{bmatrix}$$

Ανάλυση χαρακτηριστικών στην εκπαίδευση χωρίς επόπτη

- Η ανάλυση χαρακτηριστικών στην εκπαίδευση χωρίς επόπτη, πετυχαίνεται με την μέθοδο Ανάλυσης Κύριων Συνιστωσών (ΑΚΣ, PCA: Principal Components Analysis).
- Με την ΑΚΣ επιτυγχάνεται κατάλληλος μετασχηματισμός ώστε ο χώρος των χαρακτηριστικών να εμφανίζει τη μεγαλύτερη δυνατή διασπορά κατά μήκος του μικρότερου δυνατού πλήθους αξόνων



10
5
0
−5
−5
0
5
10

Μετασχηματισμός Karhunen-Loeve

Ο μετασχηματισμός Karhunen-Loeve (KLT) είναι ένα γνωστό εργαλείο στην ανάλυση πολλών μεταβλητών και χρησιμοποιείται στην ΑΚΣ

Για ένα σύνολο K ανυσμάτων σε ν-διαστάσεων χώρο χαρακτηριστικών έστω ότι x είναι το διάνυσμα:

$$\mathbf{x}_\kappa = \begin{bmatrix} x_{1\kappa} \\ x_{2\kappa} \\ \vdots \\ x_{\nu\kappa} \\ \vdots \\ x_{N\kappa} \end{bmatrix}$$

έστω μ το άνυσμα μέσης τιμής των στοιχείων του συνόλου εκπαίδευσης

$$\mu_x = E[\mathbf{x}_\kappa] = \frac{1}{K}\sum_{\kappa=1}^{K} \mathbf{x}_\kappa$$

όπου E[.] η μαθηματική προσδοκία. Τις τιμές x[κ]-μ[x] συμβολίζουμε με τη διανυσματική μεταβλητή x.

Ο πίνακας συνδιασποράς Cx όλων των ανυσμάτων του συνόλου εκπαίδευσης δίνεται:

$$\mathbf{C}_x = E[\mathbf{x}\mathbf{x}^T] = \frac{1}{K}\sum_{\kappa=1}^{K}(\mathbf{x}_\kappa - \mu_x)(\mathbf{x}_\kappa - \mu_x)^T = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & . & . & \sigma_{1\nu} & . & . & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & . & . & \sigma_{2\nu} & . & . & \sigma_{2N} \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ \sigma_{\nu 1} & \sigma_{\nu 2} & . & . & \sigma_\nu^2 & . & . & \sigma_{\nu N} \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ \sigma_{N1} & \sigma_{N2} & . & . & \sigma_{N\nu} & . & . & \sigma_N^2 \end{bmatrix}$$

Μετασχηματισμός Karhunen-Loeve

- Κάθε διαγώνια τιμή σ^2 του πίνακα συνδιασποράς εκφράζει τη διασπορά των ανυσμάτων x για τη μεταβλητή-άξονα ν
- Οι υπόλοιπες τη συνδιασπορά των δεδομένων μεταξύ δύο διαφορετικών μεταβλητών-αξόνων
- Για την ταξινόμηση είναι επιθυμητές μεγάλες τιμές των διαγωνίων τιμών σ^2 διότι μαρτυρούν ένα μεγάλο "άπλωμα" (spreading) των δεδομένων κατά μήκος του κ άξονα, ενώ είναι επιθυμητές μηδενικές τιμές για τις συνδιασπορές ώστε οι μεταβλητές - άξονες να είναι μεταξύ τους ασυσχέτιστες.
- Για να επιτύχουμε αυτό αναζητούμε ένα τέτοιο μετασχηματισμό W που θα εκφράζεται από ένα πίνακα pxp ώστε τα διανύσματα $\mathbf{y} = \mathbf{W}^T \cdot \mathbf{x}$

να έχουν πίνακα συνδιασποράς Cy με την ακόλουθη μορφή

$$\mathbf{C}_y = \begin{bmatrix} \lambda_1 & 0 & . & . & 0 & . & . & 0 \\ 0 & \lambda_2 & . & . & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & \lambda_v & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & 0 & . & . & \lambda_N \end{bmatrix} = \mathbf{\Lambda},\ \lambda_1 > \lambda_2 > \ldots > \lambda_v > \ldots > \lambda_N$$

Μετασχηματισμος K-L

Ακόμη ισχύει ότι

$$\boldsymbol{C}_y = E[yy^T] = E[\boldsymbol{W}^T\boldsymbol{x}(\boldsymbol{W}^T\boldsymbol{x})^T] = E[\boldsymbol{W}^T\boldsymbol{x}\boldsymbol{x}^{\mathbf{T}}\boldsymbol{W}] = \boldsymbol{W}^T E[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}]\boldsymbol{W} = \boldsymbol{W}^T\boldsymbol{C}_x\boldsymbol{W}$$

Είναι φανερό ότι η τιμή $\lambda_\nu = \sigma_\nu'^2$ είναι η διασπορά του y κατά τον άξονα ν, ο οποίος εκφράζεται από τη στήλη ν του πίνακα W.

Η στήλη w[ν] είναι "αποτελεσματικό" χαρακτηριστικό ενώ η τιμή λ[ν] αξιολογεί τη σπουδαιότητά του.

Ο πίνακας Cx είναι συμμετρικός και ο πίνακας Λ διαγώνιος.

Η σχέση $\boldsymbol{W}^T\boldsymbol{C}_x\boldsymbol{W} = \Lambda$ ικανοποιείται αν w1,…,wN είναι τα ιδιοανύσματα του πίνακα Cx, W=[w1,…,wN].

Οι ιδιοτιμές του Cx είναι οι τιμές λ1,…,λN.

Επειδή ο πίνακας Cx είναι συμμετρικός τα ιδιοδιανύσματα είναι κάθετα μεταξύ τους.

Οι ιδιοτιμές είναι ρίζες του χαρακτηριστικού πολυωνύμου όπως αυτό προκύπτει από τη σχέση $Det(\boldsymbol{C}_x - \boldsymbol{\lambda}\mathrm{I}) = 0$

Μετασχηματισμος K-L

Για κάθε ιδιοτιμή λν το αντίστοιχο ιδιοδιάνυσμα w[ν] είναι μία από τις άπειρες λύσεις του αόριστου συστήματος

$$(\boldsymbol{C}_x - \lambda_\mathbf{v}\mathbf{I})\boldsymbol{w}_\nu = 0$$

Είναι σκόπιμο να επιλέγουμε τα μοναδιαία ιδιοδιανύσματα $\|\mathbf{w}_\nu\| = 1$

Σε περίπτωση που απαιτείται επιπλέον μείωση του πλήθους N των ιδιοδιανυσμάτων μπορούμε να επιλέξουμε τα M από αυτά (M<N) με τις μεγαλύτερες αντίστοιχες ιδιοτιμές, αποδεχόμενοι ένα μέσο τετραγωνικό σφάλμα

$$\overline{\varepsilon^2}(\mathrm{m}) = \sum_{\nu=M+1}^{N} \lambda_\nu$$

Παράδειγμα

$$x_1 = \begin{bmatrix}1\\0\end{bmatrix},\ x_2 = \begin{bmatrix}1\\1\end{bmatrix},\ x_3 = \begin{bmatrix}2\\0\end{bmatrix},\ x_4 = \begin{bmatrix}4\\3\end{bmatrix},\ x_5 = \begin{bmatrix}5\\2\end{bmatrix}$$

$$\mu = \begin{bmatrix}(1+1+2+4+5)/5\\(0+1+0+3+4)/5\end{bmatrix} = \begin{bmatrix}\frac{13}{5}\\ \frac{6}{5}\end{bmatrix} = \begin{bmatrix}2{,}6\\1{,}2\end{bmatrix}$$

$$\sigma_1^2 = \frac{(1-2.6)^2 + (1-2.6)^2 + (2-2.6)^2 + (4-2.6)^2 + (5-2.6)^2}{5} = 2.64$$

$$\sigma_{21}^2 = \frac{(0-1.2)(1-2.6)+(1-1.2)(1-2.6)+(0-1.2)(2-2.6)+(3-1.2)(4-2.6)+(2-1.2)(5-2.6)}{5} = 1.48$$

$$\sigma_{12}^2 = \sigma_{21}^2 = 1.48$$

$$\sigma_2^2 = \frac{(0-1.2)^2 + (1-1.2)^2 + (0-1.2)^2 + (3-1.2)^2 + (2-1.2)^2}{5} = 1.36$$

$$C_x = \begin{vmatrix}\sigma_1^2 & \sigma_{12}^2\\ \sigma_{21}^2 & \sigma_{22}^2\end{vmatrix} \qquad C_x = \begin{bmatrix}2.64 & 1.48\\1.48 & 1.36\end{bmatrix}$$

$$\det(C_x - \lambda \cdot I) = \det\begin{bmatrix}2.64-\lambda & 1.48\\1.48 & 1.36-\lambda\end{bmatrix} = (2.64-\lambda)\cdot(1.36-\lambda) - 1.48^2 =$$

$$1.36 \cdot 2.64 - 1.36\cdot\lambda - 2.64\cdot\lambda + \lambda^2 - 1.48^2 =$$

$\lambda^2 - 4\cdot\lambda - 1.4$ που είναι το χαρακτηριστικό πολυώνυμο.

ΠΑΡΑΔΕΙΓΜΑ

Για $\lambda^2 - 4\cdot\lambda + 1.4 = 0$

προκύπτουν οι ρίζες: λ1 = 0.39 και λ2 = 3.61 που είναι οι ιδιοτιμές του Cx

Τα αντίστοιχα ιδιοδυανύσματα w1, w2 ικανοποιούν τις σχέσεις

$$(\boldsymbol{C}_x - \lambda_1 \cdot \mathrm{I})\cdot \boldsymbol{w}_1 = 0 \; \kappa\alpha\iota \; (\boldsymbol{C}_x - \lambda_2 \cdot \mathrm{I})\cdot \boldsymbol{w}_2 = 0$$

Για το w1 θα έχουμε αναλυτικά

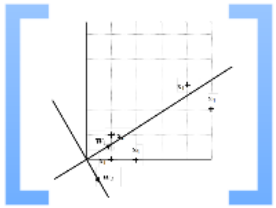
$$\begin{vmatrix} 2.64-3.61 & 1.48 \\ 1.48 & 1.36-3.61 \end{vmatrix} \cdot \begin{vmatrix} \mathrm{w}_{11} \\ \mathrm{w}_{21} \end{vmatrix} = 0 \Rightarrow \begin{aligned} -0.97\cdot \mathrm{w}_{11} + 1.48\cdot \mathrm{w}_{21} &= 0 \\ 1.48\cdot \mathrm{w}_{11} - 2.25\cdot \mathrm{w}_{21} &= 0 \end{aligned}$$

από το παραπάνω αόριστο σύστημα λύση για w[11]=α είναι w1=[α, 0.65α]'

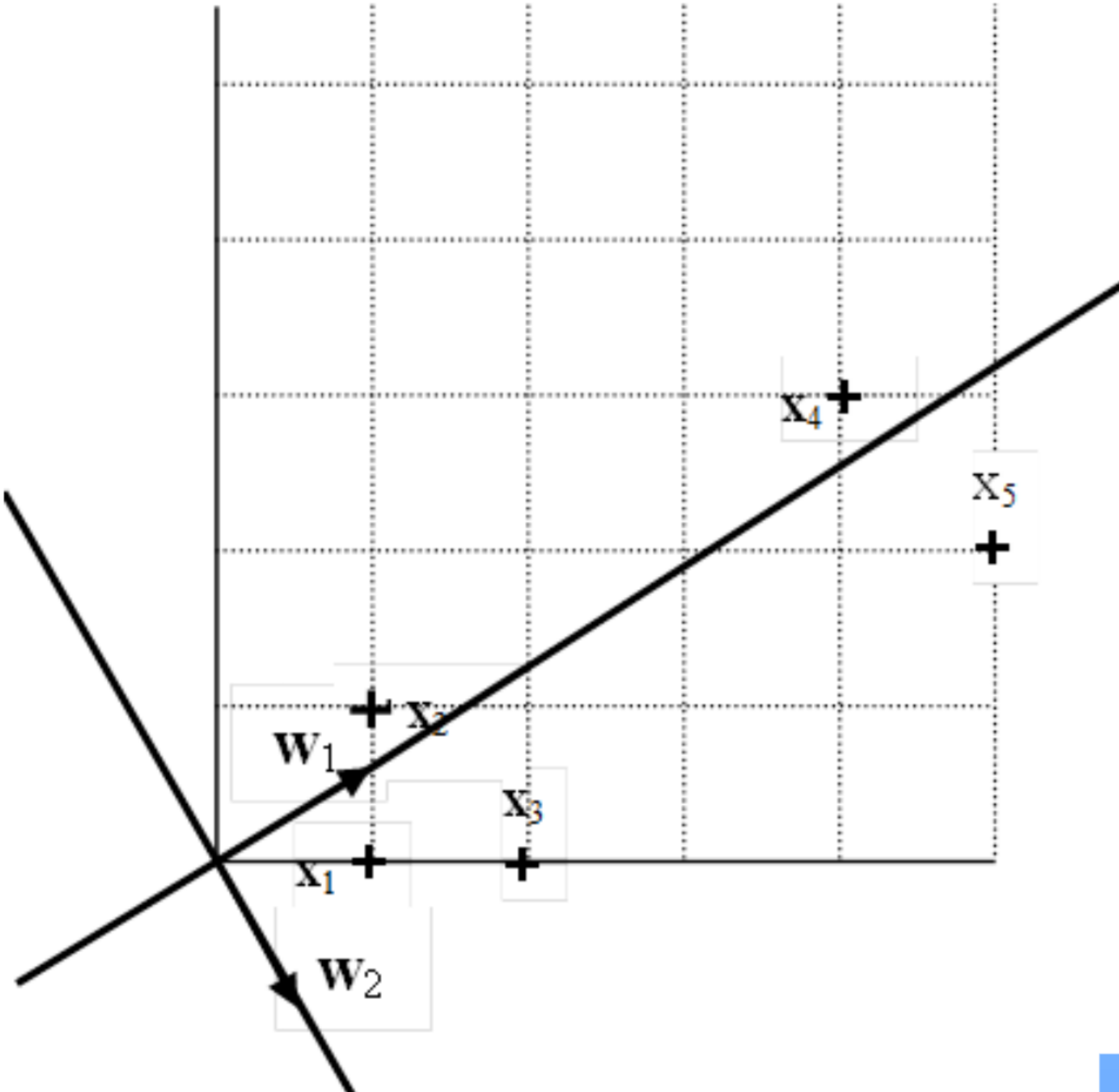
Ακόμη για να ισχύει $\|\mathbf{w}_1\| = 1 \Leftrightarrow \sqrt{\alpha^2 + (\alpha\cdot 0.65)^2} = 1 \Leftrightarrow \alpha = \pm 1/\sqrt{1+0.65^2} \approx .84$

μοναδιαίο ιδιοδιάνυσμα είναι το w1=[0.84,0.55]'

Ό μοια για την ιδιοτιμή λ2 προκύπτει w2=[0.55,-0.84]'



Τα διανύσματα είναι κάθετα μεταξύ τους όπως αναμενόταν



x4
x5
x2
W1
x3
x1
W2

Παράδειγμα

...από την άλλη όψη

Ο πίνακας συνδιασποράς ενός συνόλου προτύπων είναι $\mathbf{C}_x = \begin{bmatrix} 7/2 & \sqrt{3} \\ \sqrt{3} & 3/2 \end{bmatrix}$

Α) Πόσα χαρακτηριστικά Ν περιγράφουν διανυσματικά τα δεδομένα και ποια η συνδιασπορά τους;

$$N = 2 \quad \text{αφού} \quad C_{x(2X2)}$$

$$\sigma_{12} = \sigma_{21} = \sqrt{3}$$

Β) Ποιες οι κύριες συνιστώσες του χώρου;

$$\det(C_x - \lambda I) = (7/2 - \lambda(3/2 - \lambda) - 3 = \lambda^2 - 5\lambda + 9/4 = 0$$
$$\Delta = 25 - 4(9/4) = 16$$
$$\lambda_1 = (5+4)/2 = 9/2 \; \kappa\alpha\iota \; \lambda_2 = (5-4)/2 = 1/2$$

Αν $\mathbf{w}_1 = \begin{bmatrix} w_{11} \\ w_{21} \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} w_{12} \\ w_{22} \end{bmatrix}$ τα μοναδιαία ιδιοδιανύσματα ορίζουν τους κύριους άξονες, για λ1>λ2 θα προκύψει ο πρωτεύον άξονας w1

$$\mathbf{C}_x = \begin{bmatrix} 7/2 - 9/2 & \sqrt{3} \\ \sqrt{3} & 3/2 - 9/2 \end{bmatrix} \cdot \begin{bmatrix} w_{11} \\ w_{21} \end{bmatrix} = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} -w_{11} + \sqrt{3}w_{21} = 0 \\ \sqrt{3}w_{11} - 3w_{21} = 0 \end{cases}$$

Για το αόριστο σύστημα έστω w[11]=α

$$w_{21} = \alpha/\sqrt{3} \text{ επειδή } \|\mathbf{w}_1\| = 1 \Rightarrow \alpha^2 + \frac{\alpha^2}{3} = \frac{4\alpha^2}{3} = 1 \Rightarrow \alpha = \frac{\sqrt{3}}{2}, \mathbf{w}_1 = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$$

Για λ2<λ1 προκύπτει ο δευτερεύων άξονας

$$\mathbf{C}_x = \begin{bmatrix} 7/2-1/2 & \sqrt{3} \\ \sqrt{3} & 3/2-1/2 \end{bmatrix} \cdot \begin{bmatrix} w_{12} \\ w_{22} \end{bmatrix} = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} 3w_{12} + \sqrt{3}w_{22} = 0 \\ \sqrt{3}w_{12} + 3w_{22} = 0 \end{cases}$$

Για το αόριστο σύστημα έστω w12=α

$$w_{22} = -3\alpha/\sqrt{3} \text{ επειδή } \|\mathbf{w}_1\| = 1 \Rightarrow \alpha^2 + 3\alpha^2 = 4\alpha^2 = 1 \Rightarrow \alpha = \frac{1}{2}, \mathbf{w}_2 = \begin{bmatrix} \frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{bmatrix}$$

Γ) Ποιος ο πρωτεύον άξονας;

Πρωτεύων είναι ο άξονας που προέκυψε από την μεγαλύτερη ιδιοτιμή, δηλαδή ο w1

Δ) Ποια η διασπορά των προτύπων σε κάθε κύρια συνιστώσα;

$$(\sigma_1')^2 = \lambda_1 = 9/2$$

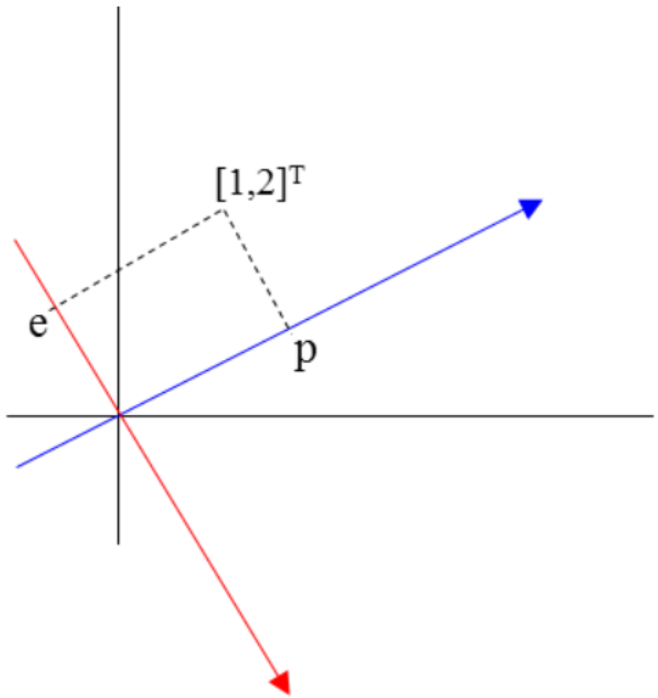
$$(\sigma_2')^2 = \lambda_2 = 1/2$$

Ε) Ποιο το μέσο τετραγωνικό σφάλμα αν περιγράψουμε τα πρότυπα μόνο με τις προβολές τους στον πρωτεύοντα άξονα;

$$\overline{\varepsilon}^2 = (\sigma_2')^2 = \lambda_2 = 1/2$$

ΣΤ) Ποια η προβολή του προτύπου με πίνακα [1, 2]';

$$p = [1,2]\begin{bmatrix} \sqrt{3}/2 \\ 1/2 \end{bmatrix} = \frac{\sqrt{3}+2}{2}$$



Ζ) Ποιο το σφάλμα αν περιγράψουμε το πρότυπο με την προβολή του στο πρωτεύοντα άξονα;

$$|e| = \left|[1,2]\begin{bmatrix} 1/2 \\ -\sqrt{3}/2 \end{bmatrix}\right| = \left|1/2 - \sqrt{3}\right|$$

Ανάλυση κύριων συνιστωσών με χρήση νευρωνικού δικτύου

- Ο KLT είναι μία αναλυτική διαδικασία για την εκτίμηση των ιδιοτιμών και των ιδιοδιανυσμάτων από τον πίνακα συνδιασποράς των δεδομένων Cx.
- Ειδικά, εάν το πλήθος των διαστάσεων του χώρου των ανυσμάτων είναι μεγάλο ο υπολογισμός και ο χειρισμός του πίνακα Cx είναι ανέφικτος.
- Αντί του KLT χρησιμοποιούμε ένα γραμμικό νευρωνικό δίκτυο ενός επιπέδου
  - Το δίκτυο αυτό εκπαιδεύεται χωρίς επόπτη από το γενικευμένο αλγόριθμο του Hebb που βασίζεται στον κανόνα εκμάθησης του Hebb
    - Το νευρωνικό δίκτυο συγκλίνει με πιθανότητα 1 στα ιδιοδιανύσματα ενός πίνακα συνδιασποράς Cx απεριόριστου μεγέθους
    - Ο υπολογισμός του Cx δεν είναι αναγκαίος επειδή τα ιδιοδιανύσματα προκύπτουν κατευθείαν από τα δεδομένα
    - Ο κανόνας προωθεί την ενίσχυση της σύναψης (βάρους) μεταξύ νευρώνων που ενεργοποιούνται ταυτόχρονα

Γ) Ποιος ο πρωτεύον άξ
Πρωτεύων είναι ο άξο
Δ) Ποια η διασπορά τω
Ε) Ποιο το μέσο τετραγ
με τις προβολές τους στ
ΣΤ) Ποια η προβολ
Ζ) Ποιο το σφάλμα
προβολή του στο π

Εκπαίδευση

t=1,2,3,...	μεταβλητή για μέτρηση της επανάληψης της διαδικασίας
x(t)	το διάνυσμα εισόδου στο νευρωνικό δίκτυο τη χρονική στιγμή t, με συνιστώσες $x_\nu(t)$, ν=0,1,…,N
K	το πλήθος των νευρώνων (K ≤ N) που ως εκ τούτου είναι και το πλήθος των επιθυμητών κύριων συνιστωσών
κ	δείκτης που αποδίδεται στους νευρώνες κ=1,…,K
$w_{\kappa\nu}(t)$	η τιμή του βάρους της σύναψης που συνδέει τον κ νευρώνα με την ν είσοδο κατά την επανάληψη t
$W_{K \times N}(t)$	ο πίνακας των $w_{\kappa\nu}(t)$
y(t)	το άνυσμα εξόδου του δικτύου κατά την επανάληψη t, με συνιστώσες $y_\kappa(t)$

Βήμα 1. Αφαιρείται πρώτα το διάνυσμα μ[x] της μέσης τιμής από κάθε στοιχείο του συνόλου εκπαίδευσης. Με τον τρόπο αυτό οι τιμές που προκύπτουν έχουν μηδενικό διάνυσμα μέσης τιμής

Βήμα 2. Αποδίδονται αρχικά (t=0) στα βάρη w[κν](0) των συνάψεων μικρές τυχαίες τιμές και στην παράμετρο γ του ρυθμού εκπαίδευσης μικρή θετική τιμή (π.χ. γ=0.007).

Βήμα 3. Υπολογίζεται το διάνυσμα εξόδου y(t) με συνιστώσες y[κ](t) για ν=1,...,N, κ=1,...,K από τη σχέση

$$y_k(t) = \sum_{\nu=1}^{N} w_{\kappa\nu}(t)\cdot x_\nu(t) \text{ ή } y(t)=W(t)\, x(t)$$

και η ποσότητα Δw[κν](t) από την σχέση

$$\Delta w_{\kappa\nu}(t) = \gamma\left(\underbrace{y_\kappa(t)\cdot x_\nu(t)}_{(\alpha)} - \underbrace{y_\kappa(t)\cdot \sum_{\lambda=1}^{\kappa} w_{\lambda\nu}(t)\cdot y_\lambda(t)}_{(\beta)}\right)$$

- Ο όρος (α) εκφράζει τον απλό κανόνα του Hebb, που λέει ότι εάν δύο νευρώνες που βρίσκονται στα άκρα μιας σύναψης (σύνδεσης), ενεργοποιούνται συγχρόνως, τότε η ισχύς της σύναψης αυξάνεται. Αλλιώς εξασθενεί ή εκφυλίζεται
- Ο όρος (β) επιβάλλει ένα όριο στην αύξηση της σύναψης

Η τιμή του w[κν](t+1) προσαρμόζεται σύμφωνα με τη σχέση $W_{\kappa\nu}(t+1) = w_{\kappa\nu}(t) + \Delta w_{\kappa\nu}$

Βήμα 4. Αύξηση της μεταβλητής t κατά 1 και επανάληψη από το βήμα 3

- φέως ότου τα βάρη των συνάψεων w[κν] φθάσουν στη σταθερή κατάσταση

t=1,2,3,…	μεταβλητή για μέτρηση της επανάληψης της διαδικασίας
x(t)	το διάνυσμα εισόδου στο νευρωνικό δίκτυο τη χρονική στιγμή t, με συνιστώσες $x_{\nu}(t)$, ν=0,1,…,N
K	το πλήθος των νευρώνων (K ≤ N) που ως εκ τούτου είναι και το πλήθος των επιθυμητών κύριων συνιστωσών
K	δείκτης που αποδίδεται στους νευρώνες κ=1,…,K
$w_{\kappa\nu}(t)$	η τιμή του βάρους της σύναψης που συνδέει τον κ νευρώνα με την ν είσοδο κατά την επανάληψη t
$W_{K \times N}(t)$	ο πίνακας των $w_{\kappa\nu}(t)$
y(t)	το άνυσμα εξόδου του δικτύου κατά την επανάληψη t, με συνιστώσες $y_{\kappa}(t)$

Εκπαίδευση

t=1,2,3,...	μεταβλητή για μέτρηση της επανάληψης της διαδικασίας
x(t)	το διάνυσμα εισόδου στο νευρωνικό δίκτυο τη χρονική στιγμή t, με συνιστώσες $x_v(t)$, v=0,1,...,N
K	το πλήθος των νευρώνων (K ≤ N) που εις εκ ταύτου είναι και το πλήθος των επιθυμητών κύριων συνιστωσών
K	δείκτης που αποδίδεται στους νευρώνες κ=1,...,K
$w_{\kappa\nu}(t)$	η τιμή του βάρους της σύναψης που συνδέει τον κ νευρώνα με την ν είσοδο κατά την επανάληψη t
$W_{K\times N}(t)$	ο πίνακας των $w_{\kappa\nu}(t)$
y(t)	το άνυσμα εξόδου του δικτύου κατά την επανάληψη t, με συνιστώσες $y_\kappa(t)$

Βήμα 1. Αφαιρείται πρώτα το διάνυσμα μ[x] της μέσης τιμής από κάθε στοιχείο του συνόλου εκπαίδευσης. Με τον τρόπο αυτό οι τιμές που προκύπτουν έχουν μηδενικό διάνυσμα μέσης τιμής

Βήμα 2. Αποδίδονται αρχικά (t=0) στα βάρη w[κν](0) των συνάψεων μικρές τυχαίες τιμές και στην παράμετρο γ του ρυθμού εκπαίδευσης μικρή θετική τιμή (π.χ. γ=0.007).

Βήμα 3. Υπολογίζεται το διάνυσμα εξόδου y(t) με συνιστώσες y[κ](t) για ν=1,...,N, κ=1,...,K από τη σχέση

$$y_k(t) = \sum_{\nu=1}^{N} w_{\kappa\nu}(t)\cdot x_\nu(t) \text{ ή } y(t)=W(t)\,x(t)$$

και η ποσότητα Δw[κν](t) από την σχέση

$$\Delta w_{\kappa\nu}(t) = \gamma\left(\underbrace{y_\kappa(t)\cdot x_\nu(t)}_{(\alpha)} - \underbrace{y_\kappa(t)\cdot\sum_{\lambda=1}^{\kappa} w_{\lambda\nu}(t)\cdot y_\lambda(t)}_{(\beta)}\right)$$

- Ο όρος (α) εκφράζει τον απλό κανόνα του Hebb, που λέει ότι εάν δύο νευρώνες που βρίσκονται στα άκρα μιας σύναψης (σύνδεσης), ενεργοποιούνται συγχρόνως, τότε η ισχύς της σύναψης αυξάνεται. Αλλιώς εξασθενεί ή εκφυλίζεται
- Ο όρος (β) επιβάλλει ένα όριο στην αύξηση της σύναψης

Η τιμή του w[κν](t+1) προσαρμόζεται σύμφωνα με τη σχέση $W_{\kappa\nu}(t+1) = w_{\kappa\nu}(t) + \Delta w_{\kappa\nu}$

Βήμα 4. Αύξηση της μεταβλητής t κατά 1 και επανάληψη από το βήμα 3

- φέως ότου τα βάρη των συνάψεων w[κν] φθάσουν στη σταθερή κατάσταση

Μετά τη φάση της εκπαίδευσης

- τα βάρη w[κν] του κ νευρώνα συγκλίνουν στη ν συνιστώσα του ιδιοδιανύσματος που αντιστοιχεί στην κ ιδιοτιμή του πίνακα συνδιασποράς Cx
- με άλλα λόγια οι γραμμές του πίνακα W έχουν προσεγγίσει τα πρώτα K ιδιοδιανύσματα του Cx, ταξινομημένες σε φθίνουσα σειρά