ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΕΡΓΑΣΤΗΡΙΟ ΕΠΕΞΕΡΓΑΣΙΑΣ ΣΗΜΑΤΩΝ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

Εφαρμογες Της Ψηφιακης Επεξεργασιας Σηματων

Μοντελοποιηση Χαμηλης Ταξης ΙδιοΦιλτρα - PCA - ΙδιοΠροσωπα

Διδάσκων Επίκουρος Καθηγητής Εμμανουήλ Ζ. Ψαράκης
Επικουρικό έργο Μαρία Ραμπότα
Πάτρα Οκτώβριος 2015

Μοντελοποιηση Χαμηλης Ταξης

Ας υποθέσουμε ότι έχουμε ένα διάνυσμα $\mathbf{u}(n)$ μήκους M που αναπαριστά μια συγκεκριμένη υλοποίηση μιας στάσιμης με την ευρεία έννοια μηδενικής μέσης τιμής, στοχαστικής διαδικασίας και θέλουμε να μεταδώσουμε αυτό το διάνυσμα σε ένα κανάλι με θόρυβο χρησιμοποιώντας ένα καινούριο σύνολο P διακριτών τιμών με $P < M$.
Ας υποθέσουμε επιπλέον ότι γνωρίζουμε το μητρώο συνδιασπορών της διαδικασίας:

$$C = \mathbb{E}\{\mathbf{u}(n)\mathbf{u}^h(n)\} = Q\Lambda Q^H$$

όπου:

$$Q = [\mathbf{q}_1, ..., \mathbf{q}_M]$$

ένα $M \times M$ μητρώο του οποίου η στήλη $\mathbf{q}_i$ είναι το i-στό ιδιοδιανύσματα του μητρώου C, και

$$\Lambda = diag(\lambda_1, ..., \lambda_M)$$

ένα $M \times M$ διαγώνιο μητρώο που περιέχει τις ιδιοτιμές λ_i , $i = 1, 2, ..., M$ του.

Τότε το αρχικό διάνυσμα μπορεί να γραφεί ως:

$$\mathbf{u}(n) = \sum_{i=1}^{M} c_i(n)\mathbf{q}_i = Q\mathbf{c}(n)$$

και εύκολα μπορούμε να αποδείξουμε ότι ισχύουν οι ακόλουθες σχέσεις:

$$\begin{aligned}
\mathbb{E}\{c_i(n)\} &= 0 && (1)\\
\mathbb{E}\{|c_i(n)|^2\} &= \lambda_i && (2)\\
\mathbb{E}\{\mathbf{u}(n)\mathbf{u}^h(n)\} &= Q\,\underbrace{\mathbb{E}\{\mathbf{c}(n)\mathbf{c}^h(n)}_{\Lambda}\,Q^T = C. && (3)
\end{aligned}$$

Είναι προφανές ότι το διάνυσμα $\mathbf{u}(n)$ μπορεί να ανακατασκευαστεί χρησιμοποιώντας τα P πρώτα ιδιοδιανύσματα του C ως

$$\hat{\mathbf{u}}(n) = \sum_{i=1}^{p} c_i(n)\mathbf{q}_i \text{ , } P < M$$

με σφάλμα

$$\mathbf{e}(n) = u(n) - \hat{\mathbf{u}}(n) = \sum_{i=P+1}^{M} c_i(n)\mathbf{q}_i$$

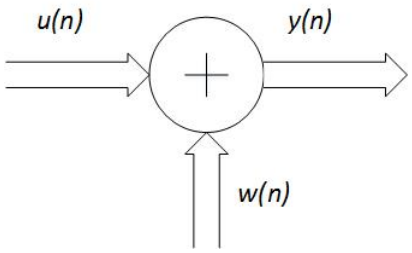
και ενέργεια ίση με την l_2 στάθμη του παραπάνω διανύσματος, δηλαδή:

$$\mathbb{E}\{\|\mathbf{e}(n)\|_2^2\} = \sum_{i=P+1}^{M} \lambda_i.$$

Η παραπάνω σχέση μας επιτρέπει να πούμε ότι η ανακατασκευή $\hat{\mathbf{u}}(n)$ του διανύσματος είναι καλή, με την προϋπόθεση βέβαια ότι οι ιδιοτιμές λ_i , $i = P+1, ...M$ είναι αρκετά μικρές.

Παραδειγμα

Ας υποθέσουμε ότι θέλουμε να μεταδώσουμε σε ενθόρυβο κανάλι ένα σήμα μηδενικής μέσης τιμής. Το διάνυσμα εξόδου $\mathbf{y}(n)$ ενός συστήματος δίνεται από το διάνυσμα εισόδου $\mathbf{u}(n)$, μήκους M, στο οποίο προστίθεται λευκός γκαουσιανός θόρυβος $\mathbf{w}(n)$ ισχύος $\sigma^2_{\mathcal{W}}$ όπως φαίνεται στο σχήμα που ακολουθεί.



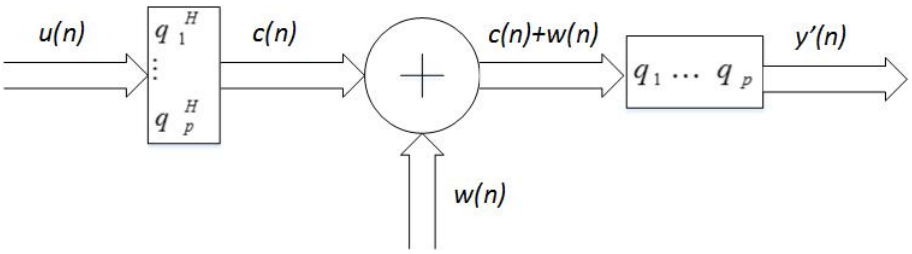
Σχήμα 1: Το προσθετικό μοντέλο θορύβου

Θεωρόντας ότι ο προσθετικός θόρυβος είναι ασυσχέτιστος με το σήμα εισόδου, μπορούμε εύκολα να βρούμε ότι το μητρώο συνδιασπορών εξόδου-θορύβου είναι $C_{\mathcal{YW}} = \sigma^2 I_M$ όπου I_M το μοναδιαίο $M \times M$ μητρώο, ενώ το μητρώο συνδιασπορών εισόδου-θορύβου θα είναι $R_{\mathcal{YW}} = 0$.

Το σφάλμα μετάδοσης είναι $\mathbf{e}(n) = \mathbf{y}(n) - \mathbf{u}(n)$ με

$$\mathbb{E}\{\|\mathbf{e}(n)\|_2^2\} = \mathbb{E}\{\|\mathbf{w}(n)\|_2^2\} = M\sigma^2_{\mathcal{W}} \tag{4}$$

Κάνοντας χρήση των ιδιοδιανυσμάτων του πίνακα συνδιασποράς όπως φαίνεται στο παρακατω σχήμα, μειώνουμε τη διάσταση του διανύσματος που μεταδίδεται σε P.



Σχήμα 2: Μετάδοση ¨ συμπιεσμένου ¨ σήματος

Στην περίπτωση αυτή το σφάλμα μετάδοσης είναι

$$\begin{aligned}\mathbb{E}\{\|\hat{\mathbf{y}}(n)-\mathbf{u}(n)\|_2^2\} &= \mathbb{E}\{\left\|Q_P Q_P^H \mathbf{u}(n)+Q_P^H \mathbf{w}(n)-\mathbf{u}(n)\right\|_2^2\} \\ &= \sum_{i=P+1}^{M} \lambda_i + P\sigma_{\mathcal{W}}^2 \end{aligned} \tag{5}$$

Από τις Σχέσεις (4) και (5) εύκολα συμπεραίνουμε ότι ο ακόλουθος κανόνας

$$(M-p)\sigma_{\mathcal{W}}^2 \lesseqgtr \sum_{i=p+1}^{M} \lambda_i$$

μπορεί να χρησιμοποιηθεί για να δούμε σε ποιές περιπτώσεις η αποστολή $P < M$ τιμών αποδίδει καλύτερα.

Ερωτησεις

1 Αποδείξτε τις Σχέσεις (1-3).

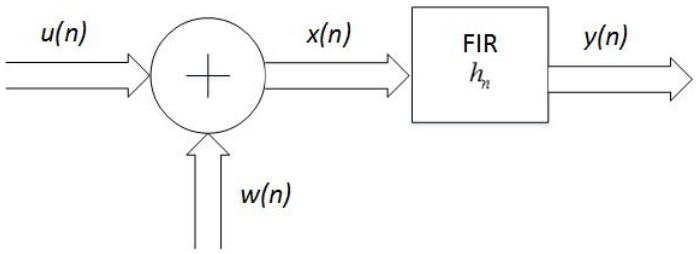
2 Αποδείξτε τις Σχέσεις (4) και (5).

3 Σε ποιές περιπτώσεις η αποστολή P τιμών αποδίδει καλύτερα; Καταγράψτε την απάντησή σας.

4 Πειραματιστείτε στο Matlab και καταγράψτε τα αποτελέσματα των πιο σημαντικών πειραμάτων σας.

Ιδιοφιλτρα (Eigenfilters)

Ας θεωρήσουμε και πάλι το προσθετικό μοντέλο θορύβου της προηγούμενης παραγράφου. Ας υποθέσουμε δηλαδή ότι το σήμα εισόδου $u(n)$, είναι μία στάσιμη δεύτερης τάξης με την ευρεία έννοια διαδικασία, με μέση τιμή μηδέν και μητρώο συνδιασπορών C, στην οποία προστίθεται λευκός γκαουσιανός θόρυβος $\sqsupseteq(n)$ με διασπορά $\sigma_{\sqsupseteq}^2$ και $\mathbb{E}\{u(n)w(n)\}=0$, όπως φαίνεται στο Σχήμα 3.



Σχήμα 3: Γραμμικό φιλτράρισμα για την απομάκρυνση προσθετικού λευκού θορύβου

Η έξοδος του συστήματος θα δίνεται από την ακόλουθη σχέση:

$$\begin{aligned} y(n) &= h(n) * x(n) \\ &= h(n) * u(n) + h(n) * w(n) \end{aligned}$$

και η μέση ισχύς της εξόδου είναι

$$P_0 = \mathbf{h}^H C \mathbf{h}$$

ενώ η μέση ισχύς του θορύβου είναι

$$N_0 = \sigma_{\mathcal{W}}^2 \mathbf{h}^H \mathbf{h} = \sigma_{\mathcal{W}}^2 ||\mathbf{h}||_2^2.$$

Θέλουμε να μεγιστοποιήσουμε το λόγο του σήματος εξόδου προς το θόρυβο δηλαδή τον λόγο:

$$SNR = P_0/N_0,$$

ή ισοδύναμα [1]:

$$\max_{\mathbf{h}} \frac{\mathbf{h}^h C \mathbf{h}}{\sigma_w^2 \mathbf{h}^h \mathbf{h}}. \tag{6}$$

Σε αυτή την περίπτωση μπορεί να αποδειχθεί πολύ εύκολα ότι οι συντελεστές της κρουστικής απόκρισης του βέλτιστου φίλτρου είναι οι ακόλουθοι:

$$\mathbf{h}^* = \mathbf{q}_{max}$$

όπου $\mathbf{q}_{max}$ το ιδιοδιάνυσμα που αντιστοιχεί στη μέγιστη ιδιοτιμή λ_{max}.

Αποδειξη με χρηση πολλαπλασιαστη Lagrange

Το πρόβλημα βελτιστοποίησης (6) μπορούμε να το γράψουμε ισοδύναμα ως ακολούθως:

$$\begin{aligned} & \max_{\mathbf{h}} \mathbf{h}^h C \mathbf{h} \\ \text{με συνθήκη:} \quad & ||\mathbf{h}||_2^2 = 1. \end{aligned} \tag{7}$$

Ορίζοντας τώρα την ακόλουθη επαυξημένη συνάρτηση κόστους :

$$\mathcal{L}(\mathbf{h}) = \mathbf{h}^h C \mathbf{h} - \lambda(\|\mathbf{h}\|_2^2 - 1)$$

και μεγιστοποιώντας την ως προς τους συντελεστές της κρουστικής απόκρισης του φίλτρου, καταλήγουμε στις ακόλουθες κανονικές εξισώσεις οι οποίες αποτελούν ένα πρόβλημα ιδιοτιμών:

$$C\mathbf{h} - \lambda \mathbf{h} = 0,$$

[1] Ο λογος που υπεισέρχεται στο πρόβλημα βελτιστοποίησης είναι γνωστός ως ¨λόγος του Rayleigh¨. Η γενική μορφή του λόγου Rayleigh είναι η ακόλουθη:

$$\mathcal{R}(\mathbf{h}) = \frac{\mathbf{h}^h A \mathbf{h}}{\mathbf{h}^h B \mathbf{h}},$$

όπου A, B θετικά ημιορισμένα τετραγωνικά μητρώα.

ή ισοδύναμα:

$$C\mathbf{h} = \lambda\mathbf{h}.$$

Από την παραπάνω σχέση είναι πολύ εύκολο να αποδείξουμε ότι η βέλτιστη λύση του προβλήματος βελτιστοποίησης (7), είναι το ιδιοδιάνυσμα που αντιστοιχεί στην μέγιστη ιδιοτιμή του μητρώου C και ότι ο πολλαπλασιαστής του Langrange ισούται με τη μέγιστη ιδιοτιμή του μητρώου.

Λογος Rayleigh - Βελτιστοποιηση με Συνθηκες

Ας υποθέσουμε ότι οι ιδιοτιμές του μητρώου C που υπεισέρχεται στο λόγο του Rayleigh είναι ταξινομημένες κατά αύξουσα σειρά, δηλαδή:

$$\lambda_{min} \equiv \lambda_1,\ \lambda_2, ..., \lambda_M \equiv \lambda_{max}.$$

Όπως είδαμε στην προηγούμενη παράγραφο, αν περιορίσουμε τον χώρο των λύσεων στην υπερσφαίρα:

$$\mathcal{H}_M = \{\mathbf{h} \in \mathbb{R}^M :\ ||\mathbf{h}||_2^2 = 1\},$$

ο λόγος του Rayleigh $\mathcal{R}(\mathbf{h})$ παίρνει τη μέγιστη τιμή του αν επιλέξουμε ως $\mathbf{h}$ το ιδιοδιάνυσμα $\mathbf{q}_M$ που αντιστοιχεί στη μέγιστη ιδιοτιμή του μητρώου, δηλαδή:

$$\begin{aligned} \lambda_{max} \equiv \lambda_M &= \max_{\mathbf{h}\in\mathcal{H}_M} \mathcal{R}(\mathbf{h}) \equiv \mathcal{R}(\mathbf{q}_M) \\ \mathbf{q}_M &= arg \max_{\mathbf{h}\in\mathcal{H}_M} \mathcal{R}(\mathbf{h}). \end{aligned}$$

Αντίστοιχα εύκολα μπορούμε να δούμε ότι ο παραπάνω λόγος παίρνει την ελάχιστη τιμή που μπορεί να πάρει αν χρησιμοποιήσουμε το ιδιοδιάνυσμα $\mathbf{q}_1$, που αντιστοιχεί στην ελάχιστη ιδιοτιμή, δηλαδή:

$$\begin{aligned} \lambda_{min} \equiv \lambda_1 &= \min_{\mathbf{h}\in\mathcal{H}_M} \mathcal{R}(\mathbf{h}) \equiv \mathcal{R}(\mathbf{q}_1) \\ \mathbf{q}_1 &= arg \min_{\mathbf{h}\in\mathcal{H}_M} \mathcal{R}(\mathbf{h}). \end{aligned}$$

Επομένως, οι δύο ακραίες ιδιοτιμές αποτελούν την μέγιστη και ελάχιστη τιμή του λόγου και προκύπτουν από τις λύσεις των αντίστοιχων προβλημάτων βελτιστοποίησης που αναφέρθηκαν παραπάνω. Το ερώτημα που τίθεται εδώ είναι αν μπορούν να προκύψουν και οι υπόλοιπες ιδιοτιμές του μητρώου ως οι βέλτιστες τιμές προβλημάτων βελτιστοποίησης. Η απάντηση, όπως θα δούμε στη συνέχεια, στο παραπάνω ερώτημα είναι καταφατική. Για το σκοπό αυτό, θα περιορίσουμε περαιτέρω το χώρο αναζήτησης των λύσεων, είτε στους χώρους :

$$\mathcal{V}^i_{M-k+1} = \{\mathbf{h} \in \mathcal{H}_M \text{ και } Q^h_{k-1}\mathbf{h} = 0\},\ i = 1,\ \cdots,\ \begin{pmatrix} M \\ M-k+1 \end{pmatrix},\ k = 2,\ 3, \cdots,\ M-1,$$

είτε :

$$\mathcal{V}_k = \{\mathbf{h} \in \mathcal{H}_M \text{ και } Q^h_{M-k}\mathbf{h} = 0\},\ i = 1,\ \cdots,\ \begin{pmatrix} M \\ k-1 \end{pmatrix},\ k = 2,\ 3, \cdots,\ M-1$$

όπου Q_l το μητρώο που προκύπτει από το μητρώο Q αν κρατήσουμε l από τις M στήλες του και $\begin{pmatrix} n \\ m \end{pmatrix}$ είναι οι συνδυασμοί των n ανά m. Θα πρέπει να τονίσουμε εδώ ότι, όπως μπορούμε να παρατηρήσουμε από τις παραπάνω σχέσεις, για κάθε τιμή του k, το πλήθος, για παράδειγμα των διαφορετικών χώρων $\mathcal{V}_k^i$ που μπορούμε να ορίσουμε από τις παραπάνω σχέσεις είναι $\begin{pmatrix} M \\ k-1 \end{pmatrix}$.
Τότε, η $k-$ *στη* ιδιοτιμή του μητρώου Q μπορεί να χαρακτηριστεί ως η βέλτιστη τιμή των ακόλουθων προβλημάτων βελτιστοποίησης :

$$\lambda_k = \max_{\mathcal{V}_{M-k+1}} \min_{\mathbf{h}} \mathcal{R}(\mathbf{h})$$

είτε:

$$\lambda_k = \min_{\mathcal{V}_k} \max_{\mathbf{h}} \mathcal{R}(\mathbf{h}).$$

Ανάλυση Κυριων Συνιστωσων-PCA

Έστω η διανυσματική μεταβλητή $\mathbf{h} \in \mathcal{C}^M$ και $\mathcal{X} \in \mathcal{C}^{N\times M}$ ένα μητρώο τυχαίων μεταβλητών. Χρησιμοποιώντας τις παραπάνω ποσότητες ορίζουμε την διανυσματική τυχαία μεταβλητή $\mathcal{Y} = \mathcal{X}\mathbf{h}$. Ο πίνακας συνδιασπορών του $\mathcal{Y}$ δίνεται από την ακόλουθη σχέση:

$$\begin{aligned} C_{\mathcal{Y}} &= \mathbb{E}\{(\mathcal{Y} - \mathbb{E}\{\mathcal{Y}\})(\mathcal{Y} - \mathbb{E}\{\mathcal{Y}\})^h\} \\ &= \mathbb{E}\{(\mathcal{X} - \mathbb{E}\{\mathcal{X}\})\mathbf{h}\mathbf{h}^h(\mathcal{X} - \mathbb{E}\{\mathcal{X}\})^h\} \end{aligned}$$

και

$$tr\{C_{\mathcal{Y}}\} = \mathbf{h}^h\mathbb{E}\{(\mathcal{X} - \mathbb{E}\{\mathcal{X}\})^h(\mathcal{X} - \mathbb{E}\{\mathcal{X}\})\}\mathbf{h} = \mathbf{h}^h G_{\mathcal{X}}\mathbf{h}$$

όπου $G_{\mathcal{X}} = \mathbb{E}\{(\mathcal{X} - \mathbb{E}\{\mathcal{X}\})^h(\mathcal{X} - \mathbb{E}\{\mathcal{X}\})\}$ [2]

Γενικευμενο Κριτηριο Διασπορας

Αν $J(\mathbf{h}) = tr\{C_{\mathcal{Y}}\} = \mathbf{h}^h G_{\mathcal{X}}\mathbf{h}$ στόχος μας είναι να λύσουμε το ακόλουθο πρόβλημα βελτιστοποίησης :

$$\max_{\mathbf{h}\in\mathcal{H}_M} J(\mathbf{h})$$

Η λύση του παραπάνω προβλήματος, όπως είδαμε σε προηγούμενη παράγραφο, είναι το ιδιοδιάνυσμα του μητρώου $G_{\mathcal{X}}$ που αντιστοιχεί στη μέγιστη ιδιοτιμή.

[2] Στην πράξη, ποτέ δεν μπορούμε να υπολογίσουμε το ντετερμινιστικό μητρώο $G_{\mathcal{X}}$. Αντί αυτού στην πράξη χρησιμοποιούμε το ακόλουθο δειγματικό μητρώο:

$$\bar{G}_X = \frac{1}{K}\sum_{k=1}^{K}(X_k - \bar{X})^t(X_k - \bar{X})$$

όπου $\bar{X} = \frac{1}{K}\sum_{k=1}^{K} X_k$ η *δειγματική μέση εικόνα.*

Σε περίπτωση που χρειαζόμαστε περισσότερες από μία συνιστώσες τότε, για d συνιστώσες θέλουμε να επιλύσουμε το ακόλουθο πρόβλημα βελτιστοποίησης :

$$\max_{\mathbf{h}\in\mathcal{H}_M,\ d} J(\mathbf{h}) \text{ με συνθήκες } \mathbf{h}_i^h\mathbf{h}_j = 0\,,\ i,j = 1,2,...,d\,,\ i \neq j$$

και η λύση είναι τα d ιδιοδιανύσματα $\mathbf{h}_1,\ \mathbf{h}_2, \cdots, \mathbf{h}_d$ που αντιστοιχούν στις d μεγαλύτερες ιδιοτιμές του μητρώου $G_{\mathcal{X}}$ [3].

Η PCA έφαρμόζεται σε πολλά και διαφορετικά προβλήματα της επεξεργασίας σημάτων. Στις παράγραφους που ακολουθούν θα δούμε πώς εφαρμόζεται η PCA στην Κατηγοριοποίηση και Ανακατασκεύη εικόνων.

- **Εξαγωγη Χαρακτηριστικων**
  Ας θεωρήσουμε ότι $A = V\Sigma U^T$ είναι η αποσύνθεση ιδιαζουσών τιμών του μητρώου A, που ας θεωρήσουμε ότι είναι μία εικόνα. Έχουμε τότε:

  $$\mathbf{y}_k = A\mathbf{x}_k\,,\ k = 1,2,...,d$$

  όπου τα $\mathbf{y}_k$ είναι οι κύριες συνιστώσες (διανύσματα) της εικόνας A. Δημιουργούμε το $B \in C^{N\times d}$, $B = [\mathbf{y}_1\,\mathbf{y}_2 ... \mathbf{y}_d]$ που είναι το μητρώο χαρακτηριστικών (feature matrix) ή εικόνα χαρακτηριστικών (feature image) της εικόνας A.

- **Κατηγοριοποιηση**
  Η απόσταση μεταξύ δύο μητρώων χαρακτηριστικών $B_i,\ B_j$ ορίζεται ως ακολούθως:

  $$d(B_i, B_j) = \sum_{k=1}^{d} ||\mathbf{y}_k^{(i)} - \mathbf{y}_k^{(j)}||_2^2$$

  Έστω $B_1, B_2, ..., B_K$ τα μητρώα χαρακτηριστικών του συνόλου εκπαίδευσης και $\mathcal{C}_k$ η κατηγορία ανάθεσης. Για να προσδιορίσουμε σε ποιά κατηγορία ανήκει ένα δείγμα B, αν:

  $$d(B, B_l) = \min_j d(B_j, B_l)$$

  και $B_l \in \mathcal{C}_k$ τότε η τελική απόφαση είναι ότι $B \in \mathcal{C}_k$.

- **Ανακατασκευη**
  Αν $B = [\mathbf{y}_1\,\mathbf{y}_2 ... \mathbf{y}_d]$ και $U = [\mathbf{x}_1\,\mathbf{x}_2 ... \mathbf{x}_d]$ τότε, $B = AU$ και ας ορίσουμε το

[3] Εναλλακτικά μπορούμε να πούμε ότι στόχος είναι να ελαχιστοποιήσουμε την ακόλουθη συνάρτηση κόστους :

$$E_{PCA}(B) = \sum_{i=1}^{d} \|\mathbf{x}_i - BB^h\mathbf{x}_i\|_2^2$$

που εκφράζει το σφάλμα ανακατασκευής του X, ενώ ο B είναι ο πίνακας που περιέχει τα ιδιοδιανύσματα του X.

ακόλουθο μητρώο:

$$\hat{A} = BU^h = \sum_{k=1}^{d} \mathbf{y}_k \mathbf{x}_k^h = \sum_{k=1}^{d} A_k.$$

Το μητρώο $\hat{A}$ θα είναι η d τάξης ανακατασκευή της εικόνας A.

Διαδικασια

Σκοπός της άσκησης είναι η χρήση των τεχνικών:

- Αποσύνθεσης Ιδιαζουσών Τιμών (Singular Value Decomposition (SVD)) και
- Ανάλυσης Κύριων Συνιστωσών (Principal Component Analysis (PCA)).

που είναι πολύ γνωστές τεχνικές αποσύνθεσης μητρώων από το μάθημα της Γραμμικής Άλγεβρας και χρησιμοποιούνται στην επεξεργασία σημάτων.

Στην συγκεκριμένη άσκηση οι παραπάνω τεχνικές θα εφαρμοστούν για την εξαγωγή του φόντου από τα πλαίσια (frames) μίας ακολουθίας εικόνων (video). Συγκεκριμένα θα δοθούν τρία αρχεία mat (αρχεία που περιέχουν δεδομένα για να επεξεργαστείτε στην Matlab) που περιέχουν τρεις (3) διαφορετικές σκηνές που διαπραγματεύονται διαφορετικά πρακτικά προβλήματα.

Για την εφαρμογή των παραπάνω τεχνικών θα πρέπει να αλλάξει η αναπαράσταση των δεδομένων ώστε να περιγράφονται από ένα μητρώο που θα έχει ως γραμμές του τα εικονοστοιχεία ενός πλαισίου της ακολουθίας. Έστω $\mathbf{X}$ το μητρώο διαστάσεων $T \times N$ που αναπαριστά ολόκληρη την ακολουθία εικόνων, όπου N είναι ο αριθμός των εικονοστοιχείων ενός πλαισίου και T ο αριθμός των πλαισίων που θα επεξεργαστούμε.

1 Η SVD του μητρώου $\mathbf{X}^T$ εκφράζεται ως:

$$\mathbf{X}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

όπου $\mathbf{U}$ είναι ένα $N \times N$ μητρώο, το $\mathbf{\Sigma}$ είναι ένα $N \times T$ θετικό διαγώνιο μητρώο που περιέχει τις ιδιάζουσες τιμές του μητρώου και το $\mathbf{V}$ είναι ένα $T \times T$ μητρώο.

2 Εφαρμόστε SVD στα πρώτα 800 πλαίσια των τριών σκηνών και κρατώντας τις πρώτες 10 ιδιάζουσες τιμές μηδενίζοντας τις υπόλοιπες ανακτήστε την ακολουθία.

1. Τι παρατηρείτε ότι έχει αλλάξει στη σκηνή;
2. Η συμπεριφορά στα τρία βίντεο είναι η ίδια; Αν όχι καταγράψτε τις διαφορές.

3 Επαναλάβατε το προηγούμενο βήμα και κρατήστε μόνο τις πρώτες 3 ιδιάζουσες τιμές.

4 Επαναλάβατε τα προηγούμενα δύο βήματα αλλά αυτή τη φορά να γίνετε η επεξεργασία κατά ορμαθούς των 100 πλαισίων. Καταγράψετε όλες τις διαφορές που παρατηρείτε.

5 Για να μπορέσετε να χρησιμοποιήσετε την τεχνική της ανάλυσης κύριων συνιστωσών, αφαιρέστε την αριθμητική μέση τιμή από κάθε στήλη του μητρώου $\mathbf{X}$. Ας συμβολίσουμε με $\mathbf{X}_o$ το μητρώο που προέκυψε μετά την εφαρμογή της παραπάνω διαδικασίας.

6 Για να υπολογίσουμε την Ανάλυση Κύριων Συνιστωσών του μητρώου, δημιουργούμε το μητρώο $\mathbf{X}_o^T\mathbf{X}_o$ μεγέθους $N \times N$ και κάνουμε Αποσύνθεση Ιδιαζουσών Τιμών.

7 Ακολουθώντας τα παραπάνω, επαναλάβετε όλα τα παραπάνω ερωτήματα, εφαρμόζοντας τώρα αντί της SVD PCA και καταγράψτε τα συμπεράσματά σας.

Για τον υπολογισμό της SVD να χρησιμοποιήσετε την ομώνυμη συνάρτηση της Matlab.

Ιδιοπροσωπα-(Eigenfaces)

Έστω $\mathbf{c}$ ένα διάνυσμα διάστασης $NM \times 1$ που έχει προκύψει από μια $N \times M$ εικόνα I. Θέλουμε να αναπαραστήσουμε το διάνυσμα $\mathbf{c}$ σε χώρο μικρότερης διάστασης. Συγκεκριμένα αν:

$$\mathbf{v} = \mathbf{c} - \mathbf{m}_c$$

με $\mathbf{m}_c$ να παριστά την ¨*μέση εικόνα*¨ τότε, το παραπάνω διάνυσμα αποκλίσεων μπορεί να εκφραστεί ως ακολούθως:

$$\mathbf{v} = \sum_{i=1}^{k} \alpha_i \mathbf{q}_i \; k \ll NM$$

όπου $\alpha_i,\ \mathbf{q}_i,\ i = 1, 2, \cdots, k$ κατάλληλες βαθμωτές και κατάλληλα ορθοκανονικά διανύσματα τα οποία θα ορίσουμε στη συνέχεια.

Διαδικασια Εκπαιδευσης

1. Έστω $I_1, I_2, ..., I_K$ οι εικόνες που θα χρησιμοποιήσουμε στη φάση εκπαίδευσης του συστήματος. Οι εικόνες αυτές θα πρέπει να έχουν κοινό κέντρο και το ίδιο μέγεθος.

2. Αναπαριστούμε την κάθε εικόνα ως ένα διάνυσμα $\mathbf{c}_i$

3. Υπολογίζουμε τη μέση εικόνα

$$\mathbf{c}_m = \frac{1}{K} \sum_{i=1}^{K} \mathbf{c}_i$$

4. Αφαιρούμε από κάθε διάνυσμα τη μέση εικόνα:

$$\mathbf{v}_i = \mathbf{c}_i - \mathbf{c}_m.$$

5. Υπολογίζουμε το ακόλουθο $NM \times NM$ μητρώο συνδιασπορών:

$$C = \frac{1}{K} \sum_{i=1}^{K} \mathbf{v}_i \mathbf{v}_i^t = VV^T$$

   όπου $V = [\mathbf{v}_1\, \mathbf{v}_2 \ldots \mathbf{v}_K]$.

6. Θέλουμε τώρα να υπολογίσουμε τα ιδιοδιανύσματα $\mathbf{q}_i\, i = 1, 2, ..., NM$ του C. Όμως ο πίνακας C είναι πολύ μεγάλος και ο απευθείας υπολογισμός των ιδιοδιανυσμάτων του έχει πολύ μεγάλη πολυπλοκότητα.
   Αντί αυτού, θεωρούμε το $K \times K$ μητρώο $\bar{C} = V^T V$ και υπολογίζουμε τα ιδιοδιανύσματά του $\mathbf{p}_i$. Εύκολα μπορούμε να αποδείξουμε (δείξτε το) ότι ο C και ο

C_M έχουν τις ίδιες ιδιοτιμές, ενώ οποιοδήποτε από τα K ιδιοδιανύσματα $\mathbf{q}_i,\ i = 1, 2, \cdots, K$ του C προκύπτει από το αντίστοιχο ιδιοδιάνυσμα $\mathbf{p}_i,\ i = 1, 2, \cdots, K$ του $\bar{C}$ ως:

$$\mathbf{q}_i = \frac{1}{\lambda_i} V \mathbf{p}_i,\ i = 1, 2, \cdots, K.$$

7. Διατηρούμε τα k ($k \leq K$) ιδιοδιανύσματα που αντιστοιχούν στις k μεγαλύτερες ιδιοτιμές, τα οποία ονομάζονται ιδιοπρόσωπα, και δημιουργούμε το ακόλουθο $NM \times k$ μητρώο:

$$Q_k = [\mathbf{q}_1\ \mathbf{q}_2 \cdots \mathbf{q}_k],\ k \leq K.$$

8. Χρησιμοποιώντας αυτό το μητρώο, μπορούμε τώρα να εκφράσουμε σε συμπιεσμένη μορφή κάθε διάνυσμα αποκλίσεων $\mathbf{v}_i$, που καθένα σχετίζεται με την αντίστοιχη εικόνα $\mathbf{c}_i$, ως:

$$\mathbf{w}_i = Q_k^T \mathbf{v}_i,\ i = 1, 2, \cdots, K$$

όπου $\mathbf{w}_i$ ένα διάνυσμα μήκους k που κάθε στοιχείο του εκφράζει (ισούται με) την προβολή της απόκλισης της εικόνας στο αντίστοιχο ιδιοδιάνυσμα του υπόχωρου $\mathbb{R}^k$ του $\mathbb{R}^{NM}$.

Ερωτηση: Αν θεωρήσουμε ότι κάθε εικόνα είναι μεγέθους 512×512 και το πλήθος των εικόνων που χρησιμοποιούμε κατά την διαδικασία της εκπαίδευσης είναι 64, δώστε μία εκτίμηση του συντελεστή συμπίεσης των δεδομένων σας.

Αναγνωριση με χρηση Ιδιοπροσωπων

Για μια εικόνα $\mathbf{c}$, σε διανυσματική μορφή, που έχει κοινό κέντρο με τις εικόνες του συνόλου εκπαίδευσης, η διαδικασία αναγνώρισης είναι η ακόλουθη:

1. Υπολογίζουμε το $\mathbf{v} = \mathbf{c} - \mathbf{c}_m$

2. Το διάνυσμα αυτό το προβάλουμε στον υπόχωρο $\mathbb{R}^k$, δηλαδή:

$$\mathbf{w} = Q_k^T \mathbf{v}$$

Είναι φανερό ότι το διάνυσμα:

$$\hat{\mathbf{v}} = Q_k \mathbf{w}$$

εκφράζει την ανακατασκευή (προσέγγιση) του διανύσματος $\mathbf{v}$.

3. Υπολογίζουμε την απόσταση:

$$e_{l^\star} = \min_{l \in \{1, 2, \cdots, K\}} \|\mathbf{w}_l - \mathbf{w}\|_p \tag{8}$$

όπου με $\|\cdot\|_p$ συμβολίζουμε την l_p στάθμη του ορίσματος.

Στους ταξινομητές *ελάχιστης απόστασης* σημαντικό ρόλο έχει η επιλογή της συνάρτησης απόστασης που θα χρησιμοποιήσουμε. Στις παραπάνω συναρτήσεις απόστασης η Ευκλείδεια αντιστοιχεί στην επιλογή του $p = 2$.

Η εικόνα $\mathbf{c}$ ταξινομείται ως η εικόνα $\mathbf{c}_{l^\star}$ από το σύνολο εκπαίδευσης.

4. Υποθέστε ότι κανονικοποιούμε τα διανύσματα $\mathbf{w}_i,\ i = 1, 2, \cdots, k$ [4], όπως και το προς ταξινόμηση διάνυσμα $\mathbf{w}$, και στη συνέχεια λύνουμε το πρόβλημα βελτιστοποίησης (8). Αποδείξτε ότι αν δημιουργήσουμε το $K \times k$ μητρώο:

$$W = \begin{bmatrix} \bar{\mathbf{w}}_1^t \\ \bar{\mathbf{w}}_2^t \\ \vdots \\ \bar{\mathbf{w}}_K^t \end{bmatrix}$$

το παραπάνω πρόβλημα βελτιστοποίησης, είναι ισοδύναμο με το ακόλουθο:

$$\rho_{l^\star} = \max_{l \in \{1,2,\cdots,K\}} \mathbf{p} \tag{9}$$

όπου το μήκους K διάνυσμα $\mathbf{p}$ ορίζεται ως ακολούθως:

$$\mathbf{p} = W\bar{\mathbf{w}}.$$

Διαδικασια

Στο πλαίσιο της άσκησης αυτής θα χρησιμοποιήσετε μία βάση εικόνων που θα την βρείτε στο συμπιεσμένο αρχείο eigenfaces.zip.

1. Αποσυμπιέστε το αρχείο.

2. Χρησιμοποιείστε τις εικόνες που υπάρχουν στον φάκελο training και ακολουθείστε την διαδικασία εκπαίδευσης που περιγράψαμε σε προηγούμενη ενότητα, για την εκπαίδευση του συστήματος.

3. Χρησιμοποιείστε τις εικόνες που υπάρχουν στον φάκελο testing και ακολουθείστε την διαδικασία αναγνώρισης που περιγράψαμε στην προηγούμενη ενότητα.

4. Χρησιμοποιείστε κατάλληλες ποσότητες από αυτές που ορίζονται στο σύνδεσμο $https://en.wikipedia.org/wiki/Sensitivity_and_specificity$, για να καταγράψετε την απόδοση των τεχνικών που εξετάσατε.

[4] Το κανονικοποιημένο διάνυσμα $\mathbf{x}$ θα το συμβολίζουμε με $\bar{\mathbf{x}}$.