

# Εφαρμογές της Εκμάθησης Πολυπτυγμάτων Riemann

Η Μεταπτυχιακή Εργασία Εξειδίκευσης

υποβάλλεται στην ορισθείσα

από τη Γενική Συνέλευση Ειδικής Σύγκλησης  
του Τμήματος Μηχανικών Η/Υ και Πληροφορικής  
Εξεταστική Επιτροπή

από τον

**Θεόφιλο Κατσιγιάννη**

ως μέρος των υποχρεώσεων για την απόκτηση του

**ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ**

**ΣΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ**

**ΜΕ ΕΞΕΙΔΙΚΕΥΣΗ**

**ΣΤΙΣ ΤΕΧΝΟΛΟΓΙΕΣ - ΕΦΑΡΜΟΓΕΣ**

Πανεπιστήμιο Ιωαννίνων

**Φεβρουάριος 2018**

# ΠΕΡΙΕΧΟΜΕΝΑ

---

Κατάλογος Σχημάτων	iii
Κατάλογος Πινάκων	iv
Κατάλογος Αλγορίθμων	v
Extended Abstract	vi
<b>1 Εισαγωγή στα Πολυπύγματα</b>	<b>1</b>
1.1 Ορισμός του Πολυπύγματος . . . . .	1
1.2 Εκθετική και λογαριθμική συνάρτηση . . . . .	2
1.3 Εκμάθηση μετρικής . . . . .	3
<b>2 Η Γεωμετρία των Πολυπτυγμάτων Stiefel</b>	<b>5</b>
2.1 Ορισμοί . . . . .	5
2.2 Η γεωμετρία των χώρων Stiefel . . . . .	6
2.2.1 Ο εφαπτόμενος χώρος των Stiefel πολυπτυγμάτων . . . . .	6
<b>3 Συσταδοποίηση και υποβιβασμός διάστασης σε ένα Riemann Πολύπτυγμα</b>	<b>13</b>
3.1 Εισαγωγή . . . . .	13
3.2 Μέθοδοι τοπικού μη γραμμικού υποβιβασμού της διάστασης σε Ευ- κλείδειους χώρους . . . . .	14
3.2.1 Τοπικά Γραμμική Εμφύτευση (LLE) . . . . .	14
3.2.2 Λαπλασιανές Ιδιοσυναρτήσεις LE . . . . .	15
3.2.3 Εσσιανή LLE . . . . .	16
3.3 Συσταδοποίηση σε ένα Riemann Πολύπτυγμα . . . . .	17
3.3.1 Ο Riemann αλγόριθμος kNN . . . . .	17
3.3.2 Ο αλγόριθμος LLE σε Riemann πολύπτυγμα . . . . .	18

3.3.3	Ο Riemann αλγόριθμος για τον αλγόριθμο LE . . . . .	18
3.3.4	Ο Riemann αλγόριθμος HLLE . . . . .	19
<b>4</b>	<b>Εκμάθηση Προβολικής Μετρικής σε Grassmann Πολυπύγματα</b>	<b>20</b>
4.1	Εισαγωγή . . . . .	20
4.2	Βασικές έννοιες . . . . .	22
4.3	Μοντελοποίηση του προβλήματος . . . . .	22
4.4	Η προβολική μετρική εκμάθησης . . . . .	23
4.5	Συνάρτηση διακριτοποίησης . . . . .	23
4.6	Βελτιστοποίηση χρησιμοποιώντας τον αλγόριθμο RCG . . . . .	24
<b>5</b>	<b>Ευσταθής Μάθηση Διακριτής Προβολής για Πολλαπλή Κατηγοριοποίηση σε Stiefel Πολυπύγματα</b>	<b>27</b>
5.1	Εισαγωγή . . . . .	27
5.2	Το προτεινόμενο πλαίσιο εκμάθησης . . . . .	28
5.3	Πολυπαραγοντική εμφύτευση . . . . .	28
5.4	Το Lasso . . . . .	29
5.5	Βελτιστοποίηση στα πολυπύγματα Stiefel . . . . .	30
	<b>Βιβλιογραφία</b>	<b>32</b>

# ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

---

1.1	Ο εφαπτόμενος χώρος στο $x$ ενός πολυπύγματος $M^n$ . . . . .	2
1.2	Η εκθετική και η λογαριθμική συνάρτηση . . . . .	3
4.1	Η κύρια μεθοδολογία για αναγνώριση προσώπου σε εικόνες. (a)-(b)-(d)-(e) αρχικά εμφυτεύουμε το πολύπτυγμα Grassmann $\mathcal{G}(q, D)$ (b) σε ένα υψηλής διάστασης χώρο Hilbert $\mathcal{H}$ (d) και μετά εκτελούμε υποβιβασμό διάστασης, αφού πρώτα εκπαιδεύσουμε κατάλληλα τις παραμέτρους μιας απεικόνισης από το χώρο Hilbert υψηλής διάστασης σε ένα χώρο Hilbert χαμηλότερης διάστασης $\mathbb{R}^d$ (e). Σε αντίθεση, η μέθοδος που μελετάμε στο παρόν κεφάλαιο ακολουθεί τα βήματα (a)-(b)-(c) για να εκπαιδεύσει μια απεικόνιση που υποβιβάζει τη διάσταση των δεδομένων από ένα πολύπτυγμα Grassmann μεγαλύτερης διάστασης $\mathcal{G}(q, D)$ σε ένα πολύπτυγμα $\mathcal{G}(q, d)$ (c) μικρότερης διάστασης στο οποίο τα δεδομένα διαχωρίζονται καλύτερα . . . . .	21
5.1	Σχέδιο για την εκμάθηση του προβολικού πίνακα με επίβλεψη . . . . .	28

# ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

---

# ΚΑΤΑΛΟΓΟΣ ΑΛΓΟΡΙΘΜΩΝ

---

4.1	Εκμάθηση Προβολικής Μετρικής (PML) . . . . .	26
4.2	Riemannian Conjugate Gradient (RCG) . . . . .	26

# EXTENDED ABSTRACT

---

Theophilos Katsigiannis, M.Sc. in Computer Science, Department of Computer Science and Engineering, University of Ioannina, Greece, February 2018.

Applications of Manifold Learning in Computer Vision.

Advisor: Christophoros Nikou, Associate Professor.

Σκοπός της διατριβής είναι να αναδείξει την συνεισφορά της γεωμετρίας Riemann στον χώρο της μηχανικής μάθησης και της υπολογιστικής όρασης.

Με τον όρο πολύπτυγμα Riemann  $M$  διάστασης  $n$  εννοούμε ένα σύνολο εφοδιασμένο με τοπολογική και διαφορίσιμη δομή, τέτοιο ώστε για κάθε σημείο του  $M$  να υπάρχει μια τοπικά ορισμένη, 1-1, συνεχής και διαφορίσιμη συνάρτηση σε ένα ανοιχτό υποσύνολο του χώρου  $R^n$ . Επίσης για κάθε σημείο ορίζεται ένα εσωτερικό γινόμενο τέτοιο ώστε να υπάρχει συμβατότητα με την τοπολογική και διαφορίσιμη δομή. Πρακτικά το πολύπτυγμα Riemann είναι μια γενίκευση της έννοιας της επιφάνειας. Πολλά προβλήματα μηχανικής μάθησης καταλήγουν σε προβλήματα βελτιστοποίησης επί πολυπτυγμάτων Riemann. Τα πιο χρησιμοποιήσιμα πολυπύγματα στο χώρο της μηχανικής μάθησης είναι ο χώρος των  $k$ -πλαισίων στον πραγματικό  $n$ -διάστατο χώρο, γνωστός και ως χώρος Stiefel και ο χώρος των  $k$ -υποχώρων στον πραγματικό  $n$ -διάστατο χώρο, γνωστός και ως χώρος Grassmann. Παρουσιάζουμε τους χώρους αυτούς αναλυτικά δίνοντας αλγεβρικές εκφράσεις, για όλες τις βασικές γεωμετρικές δομές καθώς και τη μεταγραφή γνωστών κλασικών αλγορίθμων βελτιστοποίησης στα πολυπύγματα αυτά. Τα πολυπύγματα είναι ένα βασικό εργαλείο για να λυθούν προβλήματα υπολογιστικής όρασης όπως πχ η αναγνώριση προσώπου, τα οποία καταλήγουν σε μη γραμμικά προβλήματα βελτιστοποίησης. Με τη χρήση των πολυπύγματος, μπορούν να μοντελοποιηθούν τέτοια προβλήματα χωρίς να χρειαστεί προβολή σε κάποιο γραμμικό χώρο. Ένα τέτοιο πρόβλημα αναγνώρισης προσώπου παρουσιάζεται αναλυτικά καθώς και ο αλγόριθμος που τον

λύνει ο Riemann Congugate Gradient (RCG). Επίσης παρουσιάζεται ένας τρόπος μοντελοποίησης και επίλυσης του γνωστού προβλήματος Multivariate Lasso Regression χρησιμοποιώντας πολυπύγματα Stieffel και έναν αλγόριθμο υποβιβασμού διάστασης μεταξύ πολυπτυγμάτων Stieffel. Τέλος παρουσιάζουμε ένα τρόπο επίλυσης του προβλήματος Multivariate Lasso Regression σαν πρόβλημα βελτιστοποίησης σε χώρους Grassmann και επίλυσης με χρήση του αλγορίθμου RCG.



# ΚΕΦΑΛΑΙΟ 1

## ΕΙΣΑΓΩΓΗ ΣΤΑ ΠΟΛΥΠΤΥΓΜΑΤΑ

---

1.1 Ορισμός του Πολυπτύγματος

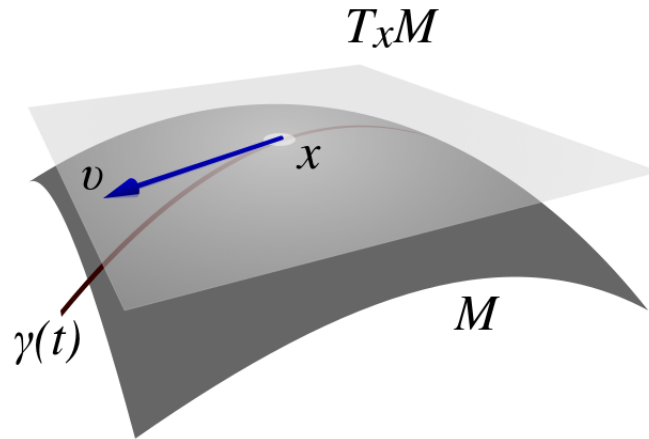
1.2 Εκθετική και λογαριθμική συνάρτηση

1.3 Εκμάθηση μετρικής

---

### 1.1 Ορισμός του Πολυπτύγματος

Ένα πολύπτυγμα Riemman [1] είναι ένα μη κενό σύνολο  $M^n$  εφοδιασμένο με μια πεπερασμένη τοπολογική δομή από ανοιχτά σύνολα  $\{U_i\}$  τα οποία καλύπτουν το  $M^n$  τέτοια ώστε  $\cup_i U_i = M$ . Αν  $x$  ένα σημείο του πολυπτύγματος υπάρχει περιοχή  $U_i$  και συνάρτηση  $f_i : U_i \rightarrow \mathbb{R}^n$  ο οποίος είναι διαφορομορφισμός, δηλαδή 1-1, επί, σε ένα ανοιχτό σύνολο του  $\mathbb{R}^n$  διαφορίσιμη και η αντίστροφη της με τις ίδιες ιδιότητες. Για κάθε σημείο  $x$  ενός πολυπτύγματος ορίζεται ο εφαπτόμενος χώρος  $T_x$ .



Σχήμα 1.1: Ο εφαπτόμενος χώρος στο  $x$  ενός πολυπύγματος  $M^n$

Το σύνολο όλων των εφαπτόμενων χώρων ονομάζεται εφαπτόμενη δέσμη  $TM$ . Επί της  $TM$  ορίζεται μια μετρική  $g : TM \times TM \rightarrow \mathbb{R}$  τέτοια ώστε  $g_x : T_x \times T_x \rightarrow \mathbb{R}$  να είναι ένα εσωτερικό γινόμενο επί του  $_x M$ .

Επί της  $TM$  υπάρχει μοναδική διαφορίσιμη δομή  $\nabla : TM \times TM \rightarrow TM$  τέτοια ώστε:

$$g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

$\nabla$  έχει τη ιδιότητα της κατευθυνόμενης παραγώγου στο πολύπτυγμα.

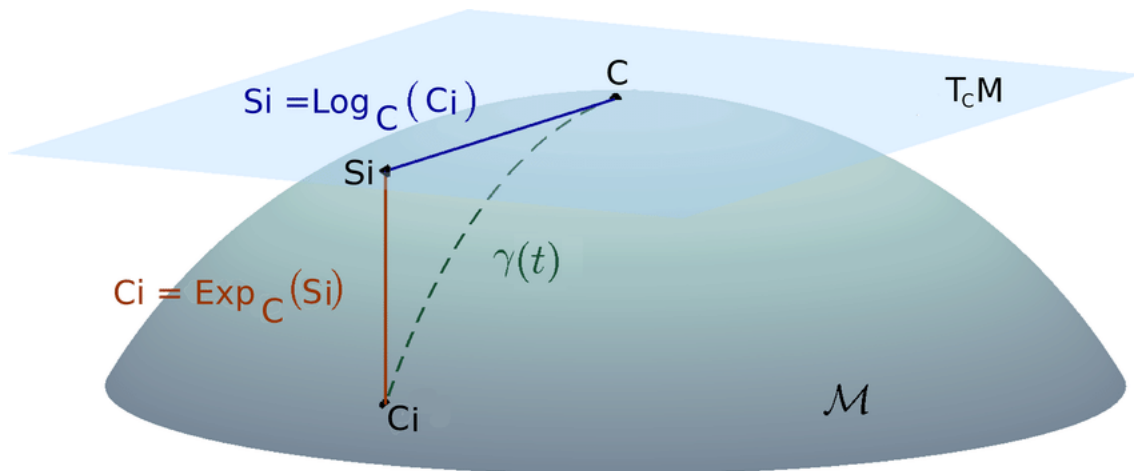
Αν  $\gamma$  είναι μια διαφορίσιμη καμπύλη επί του  $M^n$ . Η  $\gamma$  ονομάζεται γεωδαισιακή στο  $M$  αν

$$\nabla_{\dot{\gamma}} \dot{\gamma} = 0$$

## 1.2 Εκθετική και λογαριθμική συνάρτηση

Αν  $C$  είναι ένα σημείο του πολυπύγματος, ορίζουμε τη συνάρτηση  $\exp_C : T_C M \rightarrow M$  τέτοια ώστε  $\exp_C(S_i) = \gamma(1)$ , όπου  $\gamma$  είναι η μοναδική γεωδαισιακή τέτοια ώστε  $C$

και  $\dot{\gamma}(S_i) = v$  για κάθε διάνυσμα  $S_i \in T_C M$ . Η αντίστροφη συνάρτηση ονομάζεται λογαριθμική.



Σχήμα 1.2: Η εκθετική και η λογαριθμική συνάρτηση

Συνεπώς η λογαριθμική συνάρτηση ορίζεται

$$\text{Log}_C(C_i) = S_i$$

### 1.3 Εκμάθηση μετρικής

Έστω ένα σύνολο δεδομένων που έχουν αναπαράσταση σε διανυσματική μορφή και τα οποία επιθυμούμε να διαχωρίσουμε σε συστάδες. Αν η  $L_2$  μετρική δεν είναι κατάλληλη για να επιτευχθεί ο διαχωρισμός, τότε επιλέγουμε μια καινούργια μετρική, συνήθως τύπου Mahalanobis απόστασης  $dM(x, x') = (x-x')^T M(x-x')$  για να δημιουργήσουμε τις συστάδες. Ο προσδιορισμός των παραμέτρων γίνεται με τη χρήση κάποιου αλγορίθμου εκμάθησης με βάση τα διαθέσιμα παραδείγματα.

Οι περισσότερες μέθοδοι που ανήκουν σε αυτή την κατηγορία, μαθαίνουν τις παραμέτρους της μετρικής κάνοντας χρήση εκπαίδευσης με επίβλεψη και με περιορισμούς της παρακάτω μορφής:  $S = \{(x_i, x_j) : x_i \text{ και } x_j \text{ πρέπει να είναι όμοια}\}$ ,

$D = \{(x_i, x_j) : x_i \text{ και } x_j \text{ πρέπει να είναι ανόμοια}\}$ . Παρόμοιοι περιορισμοί είναι και  $R = \{(x_i, x_j, x_k) : x_i \text{ πρέπει να είναι πιο όμοια } x_j \text{ σε σχέση με } x_k\}$ . Ένας αλγόριθμος εκμάθησης μετρικής βασικά έχει σκοπό να βρει τις παραμέτρους της μετρικής, έτσι ώστε να συμφωνεί με τους παραπάνω περιορισμούς.

## ΚΕΦΑΛΑΙΟ 2

# Η ΓΕΩΜΕΤΡΙΑ ΤΩΝ ΠΟΛΥΠΤΥΓΜΑΤΩΝ STIEFEL

---

### 2.1 Ορισμοί

### 2.2 Η γεωμετρία των χώρων Stiefel

---

### 2.1 Ορισμοί

Εστιάζουμε στη γεωμετρία των πολυπτυγμάτων των πραγματικών πινάκων. Οι σημαντικότερες κατηγορίες αυτών είναι:

1. Η ομάδα των ορθογωνίων πινάκων  $O_n$  που αποτελείται of  $n \times n$  ορθογωνίων πινάκων.
2. Τα Stiefel πολυπτώγματα  $V_{n,p}$  που αποτελούνται από τους  $n \times p$  ορθογώνιους πίνακες.
3. Τα Grassmann πολυπτώγματα  $G_{n,p}$  τα οποία κάθε στοιχείο τους είναι μια κλάση ορθογωνίων πινάκων του  $V_{n,p}$  που παράγουν τον ίδιο υπόχωρο του  $\mathbb{R}^{n \times p}$ . Δηλαδή είναι είναι ένα πολύπτυγμα πηλίκο.

Ισοδύναμα, δυο πίνακες  $A, B \in V_{n,p}$  αναπαριστούν το ίδιο στοιχείο στο χώρο Grassmann, αν υπάρχει ορθογώνιος πίνακας  $P \in O_p$  τέτοιος ώστε  $A = B \cdot P$ . Επομένως,  $G_{n,p} = V_{n,p}/O_p$ .

Τα πολυπτύγματα Stiefel μπορούν να οριστούν σαν χώροι πηλίκα ορθογωνίων πολυπτυγμάτων. Κάθε πίνακας  $Q \in V_{n,p}$  μπορεί να γραφτεί στη μορφή  $\begin{pmatrix} I & 0 \\ 0 & Q \end{pmatrix}$ , όπου  $I$  είναι ο  $p \times p$  ταυτοτικός πίνακας. Επομένως  $V_{n,p} = O_n/O_{n-p}$ . Είναι φανερό, τότε ότι και ο χώρος Grassmann, μπορεί να γραφτεί στη μορφή  $G_{n,p} = O_n/(O_p \times O_{n-p})$ .

## 2.2 Η γεωμετρία των χώρων Stiefel

### 2.2.1 Ο εφαπτόμενος χώρος των Stiefel πολυπτυγμάτων

Τα Stiefel πολυπτυγμάτα  $V_{n,p}$  μπορούν να εμφυτευτούν στον  $np$ -διάστατο Ευκλείδειο χώρο  $n \times p$  των πινάκων. Αν θέσουμε  $p = 1$  τότε το πολύπτυγμα  $V_{n,1}$  είναι το πολύπτυγμα των  $n$ -διάστατων διανυσμάτων μήκους 1. Συνεπώς είναι η  $n - 1$ -διάστατη σφαίρα. Αν θέσουμε  $p = n$  τότε  $V_{n,n} = O_n$ .

Έστω  $Y(t)$  μια καμπύλη του  $V_{n,p}$ . Τότε από τον ορισμό των χώρων Stiefel έχουμε

$$Y^T Y = I$$

Παραγωγίζοντας προκύπτει

$$\frac{dY^T}{dt} Y + Y^T \frac{dY}{dt} = 0$$

Θέτουμε  $\Delta = \frac{dY}{dt}$ , επομένως

$$\Delta^T Y + Y^T \Delta = 0$$

ή

$$Y^T \Delta = - (Y^T \Delta)^T$$

ή

$Y^T \Delta$  είναι αντισυμμετρικός πίνακας

Τα στοιχεία της κύριας διαγωνίου και κάτω από αυτήν, εξαρτώνται από τα άλλα στοιχεία του πίνακα  $\Delta$ . Έχουμε  $1 + \dots + p - 1$  εξαρτημένες μεταβλητές, επιπλέον και τα στοιχεία της κύριας διαγωνίου. Συνεπώς  $\frac{p(p+1)}{2}$  εξαρτημένες μεταβλητές. Επομένως είναι φανερό ότι η διάσταση του εφαπτόμενου χώρου είναι

$$np - \frac{p(p+1)}{2} = \frac{p(p-1)}{2} + p(n-p) \quad (2.1)$$

Μπορούμε να δούμε τα πολυπτυγμάτα Stiefel εμφυτευμένα σε Ευκλείδειους χώρους, και για αυτό το λόγο μπορούν να επάγουμε την Ευκλείδεια μετρική

$$d_e(\Delta_1, \Delta_2) = \text{tr} \Delta_1^T \Delta_2,$$

δεν είναι άλλη από την Frobenius μετρική στους  $n \times p$  πίνακες. Ο κάθετος χώρος σε ένα σημείο  $Y$  του πολυπύγματος αποτελείται από όλους τους πίνακες  $N$  με την ιδιότητα

$$\text{tr} \Delta^T N = 0$$

για κάθε  $\Delta$  του εφαπτόμενου χώρου.

**Λήμμα 2.1.** Ο κάθετος χώρος σε ένα σημείο  $Y$  του πολυπύγματος  $V_{n,p}$  αποτελείται από το σύνολο των πινάκων της μορφής  $N = YS$ , όπου  $S$  είναι οποιοσδήποτε συμμετρικός  $p \times p$  πίνακας.

*Απόδειξη.* Έστω  $N = YS$ . Και  $\Delta$  ένα οποιοσδήποτε διάνυσμα του  $V_{n,p}$ . Συνεπώς

$$g_e(\Delta, YS) = 0,$$

ή

$$\text{tr}(\Delta^T YS) = 0.$$

Από την άλλη πλευρά

$$(\Delta^T Y S)^T = S^T Y^T \Delta = -S \Delta^T Y.$$

Επομένως

$$\text{tr} (\Delta^T Y S)^T = -\text{tr} S \Delta^T Y.$$

Από το οποίο προκύπτει

$$\text{tr} (\Delta^T Y S) = -\text{tr} S \Delta^T Y,$$

ή

$$\text{tr} (\Delta^T Y S) = -\text{tr} (\Delta^T Y S).$$

Έτσι έχουμε

$$\text{tr} (\Delta^T Y S) = 0$$

Επειδή η διάσταση του χώρου είναι  $p(p+1)/2$ , συμπεραίνουμε άμεσα ότι ο κάθετος αποτελείται ακριβώς από το σύνολο των πινάκων  $\{YS\}$ , όπου  $S$  είναι τυχαίος  $p \times p$  συμμετρικός πίνακας.  $\square$

Αν  $X \in V_{n,p}$ , οι στήλες του αποτελούν ένα σύνολο  $p$ -το πλήθος, σύνολο ορθοκανονικών διανυσμάτων του  $\mathbb{R}^n$ . Με τη χρήση του αλγορίθμου Gram-Schmidt μπορούμε να συμπληρώσουμε τα διανύσματα-στήλες του  $X$  με  $n-p$  ορθοκανονικά διανύσματα, κατασκευάζοντας ένα πίνακα  $X_\perp$  με αυτά. Θεωρούμε τον πίνακα  $[XX_\perp]$  που είναι η συνένωση των στηλών του  $X$  με αυτές του  $X_\perp$ . Εύκολα αποδεικνύεται το παρακάτω λήμμα.

**Λήμμα 2.2.** Ο πίνακας  $[XX_\perp]$  είναι ένας γραμμικός αυτομορφισμός του  $\mathbb{R}^{n \times p}$ .

Επομένως  $\mathbb{R}^{n \times p} = [XX_\perp] \mathbb{R}^{n \times p}$ . Συνεπώς για κάθε στοιχείο  $U \in \mathbb{R}^{n \times p}$  υπάρχει μοναδικός πίνακας  $C \in \mathbb{R}^{n \times p}$  τέτοιος ώστε  $U = [XX_\perp] C$ . Διασπώντας τον  $C$



$$C = \begin{bmatrix} A \\ B \end{bmatrix}$$

όπου  $A$  είναι ένας  $p \times p$  πίνακας και ο  $B$  είναι ένας  $(n - p) \times p$  πίνακας, τότε  $U = XA + X_{\perp}B$ .

Για κάθε πίνακα  $U \in \mathbb{R}^{n \times p}$  για δεδομένο  $X$  έχουμε

$$\text{tr}(U^T U) = \text{tr}\left((XA + X_{\perp}B)^T (XA + X_{\perp}B)\right) = \text{tr}(A^T A + B^T B) = \sum_{i,j} a_{ij}^2 + b_{ij}^2$$

Συνεπώς ο πίνακας  $C$  έχει στοιχεία τις ορθογώνιες συντεταγμένες του  $U$ , αν θεωρήσουμε το  $U$  σαν διάνυσμα του  $\mathbb{R}^{n \times p}$  με αρχή το σημείο  $X$ .

**Λήμμα 2.3.** Ο πίνακας  $Z = XA + X_{\perp}B$  ανήκει στον εφαπτόμενο χώρο  $\mathcal{T}_X V_{n,p}$  αν και μόνο αν  $A$  είναι αντισυμμετρικός.

*Απόδειξη.* Έστω  $Z = XA + X_{\perp}B$  το οποίο ανήκει στο  $\mathcal{T}_X V_{n,p}$  τότε  $Z^T X + X^T Z = 0$  το οποίο δίνει  $A^T + A = 0$ . Επομένως ο  $A$  είναι αντισυμμετρικός.

Άρα  $\mathcal{T}_X V_{n,p} \subseteq S$  όπου  $S = \{Z \in \mathbb{R}^{n \times p} \mid Z = XA + X_{\perp}B \text{ όπου } A \text{ αντισυμμετρικός}\}$ . Για κάθε αντισυμμετρικό πίνακα  $p \times p$  έχουμε  $\frac{p(p-1)}{2}$  ανεξάρτητες μεταβλητές. Έτσι

$$\dim(S) = \frac{p(p-1)}{2} + (n-p)p = np - \frac{p(p+1)}{2} = \dim(\mathcal{T}_X V_{n,p})$$

Από το οποίο συνεπάγεται  $S = \mathcal{T}_X V_{n,p}$ . □

Έστω  $Z_1, Z_2 \in \mathcal{T}_X V_{n,p}$ . Ορίζουμε το Ευκλείδειο εσωτερικό γινόμενο

$$\langle Z_1, Z_2 \rangle_e = \text{tr}(Z_1^T Z_2)$$

Με αυτό το εσωτερικό γινόμενο το πολύπτυγμα Stiefel  $V_{n,p}$  γίνεται ένα Riemannian υποπολύπτυγμα του  $\mathbb{R}^{n \times p}$ . Για κάθε  $Z \in \mathcal{T}_X V_{n,p}$ ,  $Z = XA + X_{\perp}B$  όπου  $A$  είναι  $p \times p$  αντισυμμετρικός πίνακας και  $B$  είναι  $(n-p) \times p$  πίνακας.

$$\begin{aligned}
\langle Z, Z \rangle_e &= \text{tr} (B^T X_{\perp}^T + A^T X^T) (XA + X_{\perp} B) \\
&= \text{tr} (B^T X_{\perp}^T XA + B^T X_{\perp}^T X_{\perp} B + A^T X^T XA + A^T X^T X_{\perp} B) \\
&= \text{tr} (A^T A) + (B^T B) = \text{tr} (A^T A) + \text{tr} (B^T B)
\end{aligned}$$

όπου  $X^T X = I$ ,  $X_{\perp}^T X_{\perp} = I$ ,  $X^T X_{\perp} = 0$ . Επειδή ο  $A$  είναι αντισυμμετρικός, έχουμε  $\text{tr} (A^T A) = \sum_{i>j} 2a_{ij}^2$  και  $\text{tr} (B^T B) = \sum_{i,j} b_{ij}^2$ . Παρατηρούμε ότι η Ευκλείδεια μετρική συμπεριφέρεται ανισοβαρώς. Οι “Α” συντεταγμένες έχουν διπλάσιο βάρος από τις “Β” συντεταγμένες.

Ορίζουμε μια καινούργια μετρική

$$\langle Z_1, Z_2 \rangle_c = \text{tr} \left( Z_1 \left( I - \frac{1}{2} X X^T \right) Z_2 \right)$$

Έχουμε

$$\begin{aligned}
\langle Z, Z \rangle_c &= \text{tr} \left( Z^T \left( I - \frac{1}{2} X X^T \right) Z \right) \\
&= \text{tr} (XA + X_{\perp} B)^T \left( Z - \frac{1}{2} X X^T Z \right) \\
&= \text{tr} (XA + X_{\perp} B)^T \left( XA + X_{\perp} B - \frac{1}{2} XA \right) \\
&= \text{tr} (XA + X_{\perp} B)^T \left( \frac{1}{2} XA + X_{\perp} B \right) \\
&= \frac{1}{2} \text{tr} A^T A + \text{tr} B^T B \\
&= \sum_{i>j} a_{ij}^2 + \sum_{i,j} b_{ij}^2
\end{aligned}$$

το οποίο δίνει ίσα βάρη για τις “Α” και τις “Β” συντεταγμένες.

**Λήμμα 2.4.** Ο πίνακας  $Z = XC + X_{\perp} D$  είναι στοιχείο του εφαπτόμενου χώρου  $\mathcal{T}_X V_{n,p}$  αν και μόνο αν  $C$  είναι συμμετρικός και  $D = 0$ .

*Απόδειξη.* Έστω  $Z = XC$  όπου  $C$  είναι ένας συμμετρικός πίνακας. Για κάθε  $W = XA + X_{\perp} B$  στον εφαπτόμενο χώρο

$$\begin{aligned}
\langle Z, W \rangle_e &= \text{tr} (C^T X^T (XA + X_\perp B)) \\
&= \text{tr} (C^T X^T XA + C^T X^T X_\perp B) \\
&= \text{tr} (C^T A) \\
&= \sum_{i>j} a_{ij} c_{ij} + \sum_{i<j} a_{ji} c_{ji} \\
&= \sum_{i>j} a_{ij} c_{ij} - \sum_{i>j} a_{ij} c_{ij} \\
&= 0
\end{aligned}$$

Θέτουμε  $S = \{XC \mid C \text{ συμμετρικός}\}$  έχει  $\frac{p(p+1)}{2}$ , διάσταση  $S \subseteq N(V_{n,p})$ , όπου  $N(V_{n,p})$  είναι ο κάθετος χώρος  $V_{n,p}$ . Και

$$\dim N(V_{n,p}) = \dim \mathbb{R}^{n \times p} - \dim V_{n,p} = np - \left( np - \frac{p(p+1)}{2} \right) = \frac{p(p+1)}{2}$$

Έτσι

$$N(V_{n,p}) = \{XC \mid C \text{ συμμετρικός}\}$$

□

**Λήμμα 2.5.** Έστω  $Z$  ένα διάνυσμα  $\mathbb{R}^{n \times p}$  το οποίο θεωρούμε σαν ένα διάνυσμα του χώρου  $X \in V_{n,p}$ . Τότε

$$Z = \pi_T(Z) + \pi_N(Z)$$

όπου  $\pi_T(Z)$  είναι η προβολή του  $Z$  στο  $\mathcal{T}_X V_{n,p}$  και  $\pi_N(Z)$  είναι η προβολή στον κάθετο χώρο  $\mathcal{N}_X V_{n,p}$ . Και

$$\pi_T(Z) = X \text{skew}(X^T Z) + (I - XX^T) Z$$

$$\pi_N(Z) = X \text{sym}(X^T Z)$$

where  $\text{sym}(A) = \frac{A+A^T}{2}$  and  $\text{skew}(A) = \frac{A-A^T}{2}$ .

Απόδειξη. Σύμφωνα με τα λήμματα (2.3), (2.4) έχουμε

$$Z = (XA + X_{\perp}B) + XC,$$

όπου  $A$  είναι αντισυμμετρικός και  $C$  είναι συμμετρικός. Έτσι

$$\begin{aligned}(I - XX^T)Z &= Z - XX^T Z \\ &= XA + X_{\perp}B - XX^T XA - XX^T X_{\perp}B \\ &= XA + X_{\perp}B - XA \\ &= X_{\perp}B.\end{aligned}\tag{2.2}$$

Επίσης

$$\begin{aligned}X^T Z &= X^T XA + X^T X_{\perp}B + X^T XC \\ &= A + C\end{aligned}$$

Συνεπώς,  $X^T Z = A + C$ . Από αυτό συνεπάγεται

$$\text{skew}(X^T Z) = \text{skew}(A) + \text{skew}(C) = A,$$

και

$$\text{sym}(X^T Z) = \text{sym}(A) + \text{sym}(C) = C.$$

Άρα,

$$Z = X\text{skew}(X^T Z) + (I - XX^T)Z + X\text{sym}(X^T Z)$$

□

## ΚΕΦΑΛΑΙΟ 3

# ΣΥΣΤΑΔΟΠΟΙΗΣΗ ΚΑΙ ΥΠΟΒΙΒΑΣΜΟΣ ΔΙΑΣΤΑΣΗΣ ΣΕ ΕΝΑ RIEMANN ΠΟΛΥΠΤΥΓΜΑ

---

### 3.1 Εισαγωγή

### 3.2 Μέθοδοι τοπικού μη γραμμικού υποβιβασμού της διάστασης σε Ευκλείδειους χώρους

### 3.3 Συσταδοποίηση σε ένα Riemann Πολύπτυγμα

---

## 3.1 Εισαγωγή

Στην εργασία [2] οι ερευνητές προτείνουν ένα καινοτόμο αλγόριθμο για συσταδοποίηση δεδομένων που βρίσκονται σε ένα Riemman πολύπτυγμα κάνοντας χρήση πολλαπλών υποπολλαπλοτήτων Stiefel πολυπτυγμάτων.

Αρχικά, εκπαιδεύουν μια αναπαράσταση των δεδομένων σε χώρους Stiefel, κάνοντας χρήση γενικευμένων μορφών αλγορίθμων, Riemman πολυπτυγμάτων, που αποτελούν μεταγραφές αντίστοιχων αλγορίθμων από τους Ευκλείδειους χώρους. Στη συνέχεια υποθέτοντας ότι τα δεδομένα σαν σημεία ενός χώρου Stiefel είναι διαχωρίσιμα στη συνέχεια δείχνουν ότι ο null χώρος ενός πίνακα που κατασκευάζεται από την τοπική αναπαράσταση, δίνει ένα κατακερματισμό των δεδομένων.

## 3.2 Μέθοδοι τοπικού μη γραμμικού υποβιβασμού της διάστασης σε Ευκλείδειους χώρους

Έστω  $X = \{x_i \in \mathbb{R}^D\}_{i=1}^n$  να είναι ένα σύνολο από  $n$  δεδομένα ενός  $d$ -διάστατου πολύπτυγματος, εμφυτευμένου στο χώρο  $\mathbb{R}^D$  και  $d \ll D$ . Υποθέτουμε ότι τα  $n$  σημεία είναι  $k$ -συνεκτικά δηλαδή για κάθε δύο σημεία  $x_i, x_j \in X$  υπάρχει μια ακολουθία διατεταγμένων σημείων του  $X$  έχοντας το  $x_i$  και το  $x_j$  σαν τελικά σημεία και τέτοια ώστε κάθε δύο διαδοχικά σημεία έχουν μια γειτονιά  $k$ -πλησιέστερων σημείων κοινή. Ο στόχος του υποβιβασμού διάστασης είναι να βρεθεί ένα σύνολο από διανύσματα  $\{y_i \in \mathbb{R}^d\}_{i=1}^n$ , τέτοιο ώστε τα κοντινά σημεία να παραμένουν κοντά και αυτά που πρέπει να είναι μακριά να παραμένουν μακριά.

### 3.2.1 Τοπικά Γραμμική Εμφύτευση (LLE)

Αυτή η μέθοδος υποθέτει ότι η τοπική γειτονιά ενός σημείου σε ένα πολύπτυγμα μπορεί να προσεγγιστεί από έναν γραμμικό υπόχωρο που παράγεται από τα  $k$ -κοντινότερη γειτονιά του σημείου στο πολύπτυγμα και βρίσκει μια χαμηλής διάστασης εμφύτευση των δεδομένων βασισμένη σε αυτήν την προσέγγιση.

1. *Αναζήτηση κοντινότερου γείτονα:* Βρες την  $k$ -κοντινότερη γειτονιά (kNN) για κάθε  $x_i$  σύμφωνα με την Ευκλείδεια απόσταση.
2. *Πίνακας Βαρών:* Υπολόγισε τα βάρη  $W_{ij}$  τέτοια ώστε να βρεθούν σημεία  $x_i$  στις γειτονιές τους, που ελαχιστοποιούν το κόστος

$$\varepsilon(W) = \sum_{i=1}^n \left\| \sum_{j=1}^n W_{ij} x_j - x_i \right\|^2$$

υπό τις συνθήκες  $W_{ij} = 0$  εαν  $x_j$  δεν είναι στην  $k$ -κοντινότερη γειτονιά και υπό τη συνθήκη κυρτότητας  $\sum_{j=1}^n W_{ij} = 1$ .

Για κάθε σημείο  $x_i$  τότε η  $i$ -γραμμή το  $W$  δίνεται

$$W_i = \frac{1^\top C_i^{-1}}{1^\top C_i^{-1} \mathbf{1}},$$

όπου  $C_i \in \mathbb{R}^{k \times k}$  είναι ο τοπικός Gramm πίνακας στο  $x_i$ , δηλαδή  $C_i(j, l) = (x_j - x_i) \cdot (x_l - x_i)$ , και  $1 \in \mathbb{R}^k$  είναι ένα διάνυσμα με όλα τα στοιχεία μονάδες.

3. *Αντικειμενική Συνάρτηση*: Το τρίτο και τελικό βήμα του αλγορίθμου, κάθε  $x_i$  το οποίο έχει υψηλή διάσταση απεικονίζεται σε ένα μικρότερης διάστασης σημείο το  $y_i$  το οποίο διατηρεί τις ολικές συντεταγμένες του πολυπτύγματος. Αυτό πραγματοποιείται, επιλέγοντας τις  $d$ -διάστατες συντεταγμένες για κάθε ένα από τα σημεία  $y_i$  τέτοιες ώστε να ελαχιστοποιείται το σφάλμα.

$$\phi(Y) = \sum_{i=1}^n \|y_i - \sum_{j=1}^n W_{ij} y_j\|^2.$$

Το κόστος είναι αναλλοίωτο από παράλληλες μεταφορές μιας τιμής  $y_i$  κατά μια διανυσματική σταθερά. Έτσι μπορεί να εφαρμοστεί μια παράλληλη μεταφορά ώστε οι τιμές  $y_i$  να είναι κεντραρισμένες στην αρχή των συντεταγμένων με σταθερά  $\sum_{i=1}^n y_i = 0$ . Επίσης μπορούμε να περιστρέψουμε τα  $y_i$  ώστε να μην επηρεαστεί το κόστος. Περιστρέφουμε με τέτοιο τρόπο για πάρουμε τη σωστή μεγένθυση των δεδομένων ώστε

$$\frac{1}{n} \sum_{i=1}^n y_i y_i^\top = I.$$

Με τη χρήση των παραπάνω περιορισμών το πρόβλημα γίνεται καλά ορισμένο. Συνεπώς η συνάρτηση κόστους έχει μοναδικό ελάχιστο.

### 3.2.2 Λαπλασιανές Ιδιοσυναρτήσεις LE

Ο αλγόριθμος LE δίνεται παρακάτω

1. *Αναζήτηση του κοντινότερου γείτονα*: Βρες τον  $k$ -κοντινότερους γείτονες (kNN) για κάθε  $x_i$  σύμφωνα με την Ευκλείδεια μετρική.
2. *Πίνακας Βαρών*: Κατασκεύασε τον πίνακα βαρών  $W \in \mathbb{R}^{n \times n}$

$$W_{ij} = \exp(-\|x_i - x_j\|^2/\sigma^2)$$

3. *Αντικειμενική Συνάρτηση*: Οι ιδιοσυναρτήσεις του Λαπλασιανού τελεστή  $\Delta f = \text{divgrad}f$  όπου  $\text{div}$  είναι η απόκλιση και  $\text{grad}$  είναι η κλίση του πολυπύγματος. Αναζητούμε διανύσματα  $\{y_i \in \mathbb{R}^d\}_{i=1}^n$  που ελαχιστοποιούν την παρακάτω αντικειμενική συνάρτηση:

$$\phi(Y) = \sum_{i,j} \frac{\|y_i - y_j\|^2 W_{ij}}{\sqrt{D_{ii}D_{jj}}}$$

όπου  $Y = [y_1, \dots, y_n]^\top \in \mathbb{R}^{n \times d}$ ,  $D$  είναι διαγώνιος πίνακας με την ιδιότητα  $D_{ii} = \sum_j W_{ij}$ , and  $L = I - D^{-\frac{1}{2}}WD^{-\frac{1}{2}}$ .

### 3.2.3 Εσσιανή LLE

Ο HLLE αλγόριθμος δίνεται παρακάτω

1. *Αναζήτηση γείτονα*: Βρες τους  $k$ -κοντινότερους γείτονες (kNN) για κάθε  $x_i$  σύμφωνα με την Ευκλείδεια μετρική. Έστω  $N_i = \{x_{ij}\}$  να είναι το σύνολο των  $k$ -κοντινότερων γειτόνων  $x_i$ .
2. *Εφαπτομενικές συντεταγμένες*: Θεώρησε τον πίνακα

$$\text{cov}(x_i) = \frac{1}{k} \sum_{j=1}^k (x_{i,j} - \bar{x}_i)(x_{i,j} - \bar{x}_j)^\top$$

όπου  $\bar{x}_i$  είναι η μέση τιμή kNN. Εκτέλεσε ιδιοανάλυση κάνοντας τη χρήση του αλγορίθμου  $SVD$  για τον πίνακα  $\text{cov}(x_i)$  για τον υπολογισμό των  $d$  ιδιοδιανυσμάτων  $\{u_q \in \mathbb{R}^D\}_{q=1}^d$ . Οι εφαπτομενικές συντεταγμένες του kNN δίνονται από τις  $d$  στήλες του  $k \times d$  πίνακα  $V$  που δίνεται παρακάτω, όπου  $p = 1, \dots, k$  και  $q = 1, \dots, d$



$$V_{pq} = (x_{i,p} - \bar{x}_i)^\top u_q = \langle x_{i,p} - \bar{x}_i, u_q \rangle.$$

3. *Αντικειμενική συνάρτηση*: Αρχικά υπολόγισε τον Ερσιανό τελεστή  $h^i$  για κάθε σημείο  $x_i$  στο πολύπτυγμα με τη μέθοδο των ελαχίστων τετραγώνων. Θεώρησε μια λεία συνάρτηση  $f : M \rightarrow \mathbb{R}$ . Υπολόγισε τη συνάρτηση για όλα τα  $k$ NN σε ένα σημείο  $x_i$  του πολυπύγματος και τοποθέτησε τα σε ένα διάνυσμα  $f_i$ . Μπορεί ναδειχθεί  $h^i f_i$  ότι το σύνολο των διανυσμάτων προσεγγίζει τις συνιστώσες της Ερσιανής, όπου  $\frac{\partial^2 f}{\partial V_p \partial V_q}$ . Αυτοί είναι τοπικοί υπολογισμοί, και χρησιμοποιούνται για να δωθεί ένας εμπειρικός υπολογισμός της  $(i, j)$  στήλης του  $H$

$$H_{i,j} = \sum_l \sum_r ((h^l)_{r,i} (h^l)_{r,j}).$$

4. *Το πρόβλημα των αραιών ιδιοτιμών*: Πραγματοποίησε μια ιδιοανάλυση του  $H$  και κατασκεύασε τα  $(d+1)$  ιδιοδιανύσματα που συσχετίζονται με τις  $(d+1)$ -ιδιοδιανύσματα. Τα διανύσματα που έχουν σε όλες τις θέσεις μονάδες,  $1 \in \mathbb{R}^n$ , είναι τα ιδιοδιανύσματα του  $H$  που αντιστοιχούν στην ιδιοτιμή 0. Τα  $d$  ιδιοδιανύσματα του πίνακα  $H$  που αντιστοιχούν στις δεύτερες  $(d+1)$  μικρότερες ιδιοτιμές, παράγουν α  $d$ -διάστατο χώρο το οποίο περιέχει τις εμφυτευμένες συντεταγμένες των στηλών του  $Y$ . Επειδή ο  $H$  είναι συμμετρικός, κάποιος μπορεί να επιλέξει τα εμφυτευμένα διανύσματα που είναι ορθογώνια με το διάνυσμα 1.

### 3.3 Συσταδοποίηση σε ένα Riemann Πολύπτυγμα

Δίνουμε τους αντίστοιχους που ισχύουν σε ένα πολύπτυγμα Riemann

#### 3.3.1 Ο Riemann αλγόριθμος kNN

Είναι ο ίδιος αλγόριθμος όπως περιγράφεται στην προηγούμενη παράγραφο μόνο που αντικαθιστούμε την Ευκλείδεια μετρική με την αντίστοιχη μετρική του πο-

λυπτύγματος Riemann. Πιο συγκεκριμένα ορίζουμε τον  $k$ NN του  $x_i$  να είναι τα  $k$  σημεία  $x_j$  που ελαχιστοποιούν την σχέση  $\|\log_{x_i}(x_j)\|_{x_i}$ .

### 3.3.2 Ο αλγόριθμος LLE σε Riemann πολύπτυγμα

Πρέπει να ελαχιστοποιήσουμε στο πολύπτυγμα Riemann το ολικό σφάλμα,

$$\varepsilon(W) = \sum_{i=1}^n \|\log_{x_i}(\hat{x}_{Riem,i})\|_{x_i}^2$$

υπό τον περιορισμό  $W_{ij} = 0$ , αν  $x_j$  δεν είναι  $k$ NN του  $x_i$  και  $\sum_j W_{ij} = 1$ .  $\hat{x}_{Riem,j}$ , είναι η γεωδαισιακό γραμμική παρεμβολή το  $x_j$  στο δικό του  $k$ NN.

$$\hat{x}_{Riem,i} = \exp_{x_i} \left( \sum_{j=1}^n W_{ij} \log_{x_i}(x_j) \right).$$

Επειδή οι συναρτήσεις  $\exp$  και  $\log$  είναι αντίστροφες μεταξύ τους συναρτήσεις, έχουμε

$$\varepsilon_{Riem}(W) = \sum_{i=1}^n \left\| \sum_{j=1}^n W_{ij} \log(x_j) \right\|_{x_i}^2.$$

Χρησιμοποιώντας την ίδια λογικό όπως και στην περίπτωση των Ευκλείδειων χώρων, τα βέλτιστα βάρη προκύπτουν με τη χρήση του τοπικού Gram πίνακα  $C_i \in \mathbb{R}^{k \times k}$  που ορίζεται ως

$$C_i(j, l) = \langle \log_{x_i}(x_j), \log_{x_i}(x_l) \rangle_{x_i}$$

### 3.3.3 Ο Riemann αλγόριθμος για τον αλγόριθμο LE

Αντί να προσπαθούμε τα γράφουμε κάθε σημείο σαν γραμμικό συνδυασμο των  $k$ NN του, βρίσκουμε τον πίνακα των βαρών  $W \in \mathbb{R}^{n \times n}$  των οποίων οι τιμές  $W_{ij}$  υπολογίζουν το πόσο κοντά είναι τα  $x_i$  σε σχέση με τα  $x_j$ . Επομένως μετασχηματίζοντας τις Laplace ιδιοσυναρτήσεις σε ένα Riemann πολύπτυγμα ώστε οι υπολογισμοί αντί

να χρησιμοποιούν την  $\exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right)$ , να χρησιμοποιείται η αντίστοιχη Riemann μετρική. Συνεπώς

$$W_{ij} = \exp\left(-\frac{\text{dist}_{Riem}(x_i, x_j)^2}{\sigma^2}\right) = \exp\left(-\frac{\|\log_{x_i}(x_j)\|_{x_i}^2}{\sigma^2}\right)$$

και με περιορισμούς  $W_{ij} = 0$  αν  $x_j$  δεν ανήκει στην  $k$ NN του  $x_i$ . Και  $\tilde{M} = L = I - D^{-\frac{1}{2}}WD^{-\frac{1}{2}}$  και  $D$  είναι τέτοιος ώστε  $D_{ii} = \sum_j W_{ij}$ .

### 3.3.4 Ο Riemann αλγόριθμος HLLE

Στο δεύτερο βήμα του αλγορίθμου HLLE πραγματοποιείται ο υπολογισμός των συντεταγμένων, για κάθε  $x_i$  εφαρμόζοντας την ευκλείδεια PCA για κάθε έναν από τους γείτονες του με την παραδοχή ότι τα τοπικά σημεία ανήκουν στο ίδιο υπόχωρο. Αυτή η υπόθεση στην περίπτωση των Riemann πολυπτυγμάτων δεν ισχύει πλέον. Για να υπολογιστούν οι κύριες γεωδαισιακές συνιστώσες δεν είναι απλό διαδικασία. Για το λόγο αυτό πρέπει να χρησιμοποιηθεί η κατάλληλη Riemann μετρική, ο σωστός μέσος και ο αντίστοιχος συναλλοίωτος πίνακας.

Έστω  $\{x_{i,j}\}_{j=1}^k$  είναι το σύνολο των  $k$ NN του  $x_i$ . Πρώτα υπολογίζουμε τον εσωτερικό μέσο  $\bar{x}_i$  του  $k$ NN. Στη συνέχεια υπολογίζουμε το εφαπτόμενο διάνυσμα  $v_j = \log_{\bar{x}_i}(x_{i,j})$  σε σχέση με το  $\bar{x}_i$  και οι γεωδαισιακές συνιστώσες είναι η ορθοκανονική βάση για το  $T_{\bar{x}_i}\mathcal{M}$ . Θα γράψουμε τον προβολικό τελεστή χρησιμοποιώντας τον Riemann πίνακα. Επομένως οι εφαπτόμενες συντεταγμένες του  $k$ NN δίνονται από τον  $k \times d$  πίνακα  $V$ , όπου

$$V_{pq} = \langle \log_{\bar{x}_i}(x_{i,p}), u_q \rangle_{\bar{x}_i}, \quad p = 1, \dots, k, \quad q = 1, \dots, d.$$

## ΚΕΦΑΛΑΙΟ 4

# ΕΚΜΑΘΗΣΗ ΠΡΟΒΟΛΙΚΗΣ ΜΕΤΡΙΚΗΣ ΣΕ GRASSMANN ΠΟΛΥΠΤΥΓΜΑΤΑ

- 
- 4.1 Εισαγωγή
  - 4.2 Βασικές έννοιες
  - 4.3 Μοντελοποίηση του προβλήματος
  - 4.4 Η προβολική μετρική εκμάθησης
  - 4.5 Συνάρτηση διακριτοποίησης
  - 4.6 Βελτιστοποίηση χρησιμοποιώντας τον αλγόριθμο RCG
- 

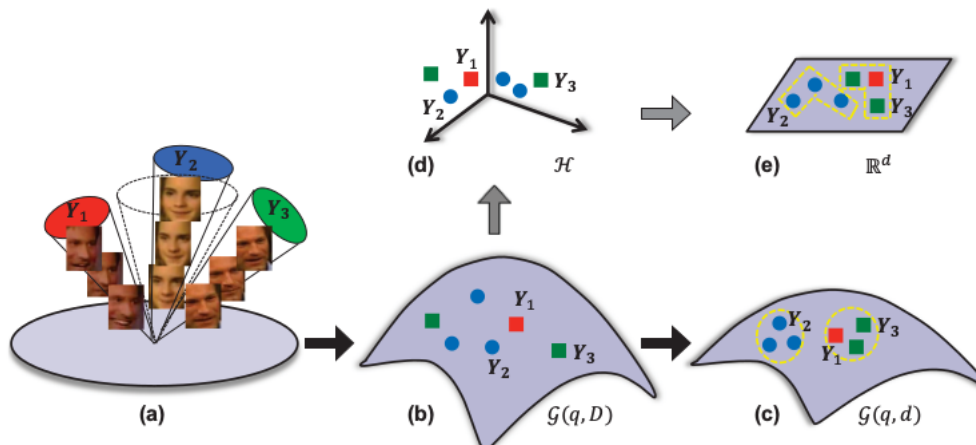
### 4.1 Εισαγωγή

Ένα βασικό εργαλείο στην αναπαράσταση εικόνων video στο πρόβλημα αναγνώρισης προσώπων, είναι οι γραμμικοί υπόχωροι [3],[4],[5],[6]. Στο πρόβλημα της αναγνώρισης προσώπου σε εικόνες video, μεγάλη επιτυχία είχε η μέθοδος στην οποία κάθε video αναπαρίσταται σαν γραμμικός υπόχωρος ο οποίος είναι στοιχείο ενός πολυπτύγματος Grassmann κατάλληλης διάστασης.

Σε αυτό το κεφάλαιο παρουσιάζουμε αναλυτικά την εργασία [7], στην οποία οι ερευνητές, με μιά πρωτοποριακή μέθοδο επιχειρούν να λύσουν το πρόβλημα αναγνώρισης των προσώπου σε video αφού πρώτα αναπαριστούν κάθε video σε μορφή γραμμικού χώρου για να πετύχουν υποβιβασμό διάστασης κάνοντας χρήση ενός πο-

λυπτύγματος Grassmann.

Η μέθοδος επιχειρεί να δώσει καλύτερα αποτελέσματα σε σχέση με τις κλασικές μεθόδους [8],[9],[10],[11],[12] στις οποίες οι ερευνητές επιχειρούν να λύσουν το πρόβλημα κατασκευάζοντας κατάλληλους μετασχηματισμούς από τους χώρους Grassmann σε Ευκλείδειους χώρους υψηλής διάστασης και στη συνέχεια υποβιβάζοντας τη διάσταση των αναπαραστάσεων των δεδομένων σε χαμηλότερης διάστασης Ευκλείδειους χώρους. Πιο συγκεκριμένα, κατασκεύασαν συναρτήσεις θετικά ορισμένων πυρήνων, με πεδίο ορισμού ένα Grassmann πολύπτυγμα και σύνολο αφίξεως έναν υψηλής διάστασης χώρο Hilbert, οι οποίοι είναι εφοδιασμένοι με την Ευκλείδεια γεωμετρία. Έτσι κάθε πρόβλημα αρχικά ορισμένο σε κάποιο χώρο Grassmann εμφυτεύεται σε κάποιο Ευκλείδειο χώρο υψηλής διάστασης. Στη συνέχεια εφαρμόζεται ένας αλγόριθμος για τον υποβιβασμό σε χαμηλότερης διάστασης Ευκλείδειο χώρο.



Σχήμα 4.1: Η κύρια μεθοδολογία για αναγνώριση προσώπου σε εικόνες. (a)-(b)-(d)-(e) αρχικά εμφυτεύουμε το πολύπτυγμα Grassmann  $\mathcal{G}(q, D)$  (b) σε ένα υψηλής διάστασης χώρο Hilbert  $\mathcal{H}$  (d) και μετά εκτελούμε υποβιβασμό διάστασης, αφού πρώτα εκπαιδεύσουμε κατάλληλα τις παραμέτρους μιας απεικόνισης από το χώρο Hilbert υψηλής διάστασης σε ένα χώρο Hilbert χαμηλότερης διάστασης  $\mathbb{R}^d$  (e). Σε αντίθεση, η μέθοδος που μελετάμε στο παρόν κεφάλαιο ακολουθεί τα βήματα (a)-(b)-(c) για να εκπαιδεύσει μια απεικόνιση που υποβιβάζει τη διάσταση των δεδομένων από ένα πολύπτυγμα Grassmann μεγαλύτερης διάστασης  $\mathcal{G}(q, D)$  σε ένα πολύπτυγμα  $\mathcal{G}(q, d)$ (c) μικρότερης διάστασης στο οποίο τα δεδομένα διαχωρίζονται καλύτερα

## 4.2 Βασικές έννοιες

Στο κεφάλαιο αυτό θεωρούμε κάθε video σαν μια ακολουθία εικόνων χωρίς να λαμβάνουμε υπόψη πληροφορίες που σχετίζονται με το χρόνο. Είναι γνωστό ότι μια τέτοια αναπαράσταση video το οποίο περιέχει ένα μόνο πρόσωπο, μπορεί να αναπαρασταθεί σαν ένα γραμμικό χώρο μικρής διάστασης [3]. Με μια τέτοια αναπαράσταση κερδίζουμε το χαμηλό κόστος επεξεργασίας σε σχέση με τα μεγάλα σε μέγεθος δεδομένα [13] που θα πρέπει να επεξεργαστούμε. Σαν συνέπεια αυτού, πολλές τεχνικές έχουν σχεδιαστεί για υποβιβασμό διάστασης αλλά δεν είναι άμεσα εκμεταλεύσιμες από δεδομένα που προέρχονται από Grassmann πολυπύγματα. Στη συνέχεια παραθέτουμε μια σειρά από εργασίες που μελετάνε τη Riemann γεωμετρία των πολυπτυγμάτων Grassmann [14],[15],[16],[17]. Μόνο μερικές από αυτές [14], [15], [17] χρησιμοποιούν την προβολική απεικόνιση για να αναπαραστήσουν κάθε στοιχείο των πολυπτυγμάτων Grassmann με τον προβολικό τελεστή του.

Η προβολική απόσταση σχετίζεται άμεσα με την ακριβή γεωδαισιακή απόσταση σε ένα Grassmann πολύπτυγμα σε κάθε τάξη μεγέθους του πολυπύγματος. Όλοι οι παραδοσιακοί αλγόριθμοι που έχουν σχεδιαστεί για τον Ευκλείδειο χώρο μπορούν να μεταγραφούν και σε μη γραμμικά πολυπύγματα.

## 4.3 Μοντελοποίηση του προβλήματος

Υποθέτουμε ότι έχουμε  $m$  μια ακολουθία από videos  $\{X_1, X_2, \dots, X_m\}$ . Για κάθε ένα από αυτά  $X_i$  αποτελεί μια ακολουθία εικόνων ενός προσώπου. Γενικά  $X_i \in \mathbb{R}^{D \times n_i}$ . Δηλαδή κάθε ένα από τα videos περιέχουν  $n_i$  εικόνες, και κάθε εικόνα μπορεί να περιγραφεί σαν ένα  $D$ -διάστατο διάνυσμα κάποιων περιγραφέντων. Κάθε ένα από αυτά τα videos ανήκουν σε μια κατηγορία που συμβολίζεται με  $C_i$ . Το  $X_i$  video μπορεί να αναπαρασταθεί με ένα  $q$  διάστατο γραμμικό υπόχωρο που παράγεται από μια ορθοκανονική βάση με πίνακα  $Y_i \in \mathbb{R}^{D \times q}$ ,  $X_i X_i^T \simeq Y_i \Lambda_i Y_i^T$ , όπου  $\Lambda_i, Y_i$  αντιστοιχεί σε πίνακες των  $q$  μεγαλύτερων ιδιοτιμών και ιδιοδιανυσμάτων αντίστοιχα.

Δοθέντως ενός γραμμικού υπόχωρου  $\text{span}(Y_i)$  σε ένα Grassmann πολύπτυγμα (ορίζουμε  $Y_i Y_i^T$  την αναπαράσταση κάθε στοιχείου πάνω στο πολύπτυγμα). Επιθυμούμε να εκπαιδευσουμε μια απεικόνιση  $f : \mathcal{G}(q, D) \rightarrow \mathcal{G}(q, d)$  υπολογίζοντας την παράμετρο  $d$  μέσω παραδειγμάτων. Η συνάρτηση  $f$  έχει μορφή

$$f(Y_i Y_i^T) = W^T Y_i Y_i^T W = (W^T Y_i) (W^T Y_i)^T$$

όπου  $W \in \mathbb{R}^{D \times d}$  ( $q \leq D$ ), είναι ένας πίνακας μετασχηματισμού με στήλες μέγιστης βαθμίδας. Με τη χρήση αυτής της απεικόνισης μετασχηματίζουμε το αρχικό πολύπτυγμα Grassmann  $\mathcal{G}(q, D)$  μετασχηματίζεται σε ένα μικρότερης διάστασης Grassman πολύπτυγμα  $\mathcal{G}(q, d)$ . Εκτός από την περίπτωση που ο  $W$  είναι ορθογώνιος πίνακας, ο  $W^T Y_i$  δεν είναι γενικά ορθογώνιος. Μόνο όταν ο γραμμικός υπόχωρος παράγεται από ορθογώνιο πίνακα είναι πραγματικό σημείο ενός πολυπύγματος Grassmann. Οι ερευνητές χρησιμοποιούν τις ορθογώνιες προβολές του  $W^T Y_i$  που ορίζονται από  $W^T Y'$  για να αναπαραστήσουν μια ορθογώνια βάση του πίνακα των μετασχηματισμένων προβολικών πινάκων.

#### 4.4 Η προβολική μετρική εκμάθησης

Η προβολική μετρική για κάθε ένα ζεύγος μετασχηματισμένων προβολικών τελεστών  $W^T Y_i Y_i'^T W$ ,  $W^T Y_j Y_j'^T W$  ορίζεται ως:

$$\begin{aligned} d_p^2(W^T Y_i Y_i'^T W, W^T Y_j Y_j'^T W) &= \\ 2^{-\frac{1}{2}} \|W^T Y_i Y_i'^T W - W^T Y_j Y_j'^T W\|_F^2 &= \\ 2^{-\frac{1}{2}} \text{tr}(P A_{ij} A_{ij}^T P) &. \end{aligned}$$

όπου  $A_{ij} = Y_i Y_i'^T - Y_j Y_j'^T$  και  $P = W W^T$ . Επειδή ο  $W$  απαιτείται να είναι πίνακας μέγιστης βαθμίδας ως προς τις στήλες, ο  $P$  είναι βαθμίδας- $d$  συμμετρικός θετικά ορισμένος (PSD) πίνακας μεγέθους  $D \times D$ , έχει τη δομή ενός Mahalanobis πίνακα.

#### 4.5 Συνάρτηση διακριτοποίησης

Η συνάρτηση διακριτοποίησης είναι σχεδιασμένη να ελαχιστοποιεί τις προβολικές αποστάσεις για κάθε δυο στοιχεία που ανήκουν στην ίδια κλάση και μεγιστοποιεί τις προβολικές αποστάσεις στοιχείων που βρίσκονται σε διαφορετικές κλάσεις. Ο πίνακας  $P$  υπολογίζεται μέσω της αντικειμενικής συνάρτησης  $J(P)$ :

$$P^* = \arg \min_P J(P) = \arg \min_P (J_w(P) - \alpha J_b(P)). \quad (4.1)$$

όπου  $\alpha$  σχετίζεται με την αντιστάθμιση μεταξύ του πόσο συμπαγείς  $J_w(p)$  πρέπει να είναι οι κλάσεις σε σχέση με τις αποστάσεις μεταξύ των κλάσεων  $J_b(P)$ . Όπου  $J_w(p)$  μετρά το μέσο όρο των αποστάσεων των σημείων μέσα στις κλάσεις και  $J_b(p)$  μετρά το μέσο όρο των αποστάσεων των σημείων που αντιστοιχούν σε διαφορετικές κλάσεις. Συνεπώς έχουμε:

$$J_w(P) = \frac{1}{N_w} \sum_{i=1}^m \sum_{j:C_i=C_j} 2^{-\frac{1}{2}} \text{tr} (PA_{ij}A_{ij}^T P),$$

$$J_b(P) = \frac{1}{N_b} \sum_{i=1}^m \sum_{j:C_i \neq C_j} 2^{-\frac{1}{2}} \text{tr} (PA_{ij}A_{ij}^T P).$$

όπου  $N_w$  είναι ο αριθμός των ζευγών των δειγμάτων από την ίδια κλάση και  $N_b$  είναι ο αριθμός των ζευγών σημείων που ανήκουν σε διαφορετικές κλάσεις, και  $A_{ij} = Y_i Y_i^T - Y_j Y_j^T$  και  $P$  είναι ένας θετικά ημιορισμένος που πρέπει να υπολογιστεί μέσω της μάθησης.

## 4.6 Βελτιστοποίηση χρησιμοποιώντας τον αλγόριθμο RCG

Για όλα τα  $i$  ο πίνακας  $Y_i$  πρέπει να κανονικοποιηθεί σε ένα πίνακα  $Y_i'$  για δεδομένο  $P = WW^T$  έτσι οι στήλες του  $W^T Y_i'$  να είναι ορθοκανονικές. Ειδικά οι ερευνητές εκτελούν QR-διάσπαση του  $W^T Y_i$  έτσι ώστε  $W^T Y_i = Q_i R_i$ , όπου  $Q_i \in \mathbb{R}^{D \times q}$  είναι ο ορθοκανονικός πίνακας που αποτελείται από τις  $q$  πρώτες στήλες και  $R_i \in \mathbb{R}^{d \times d}$  είναι ο αντιστρέψιμος πίνακας του άνω τριγωνικού πίνακα. Επειδή ο  $R_i$  είναι αντιστρέψιμος και ο  $Q_i$  είναι ορθογώνιος. Πρέπει να οριστούν οι  $Q_i = W^T (Y_i R_i^{-1})$ . Συνεπώς

$$Y_i' = Y_i R_i^{-1} \quad (4.2)$$

Η εξίσωση (4.1) μετασχηματίζεται στην

$$P^* = \arg \min_P \text{tr}(PS_w P) - \alpha \text{tr}((PS_w P)). \quad (4.3)$$



όπου

$$S_w = \frac{1}{N_w} \sum_{i=1}^m \sum_{j:C_i=C_j} 2^{-1/2} \text{tr}(A_{ij} A_{ij}^T) \quad (4.4)$$

$$S_b = \frac{1}{N_b} \sum_{i=1}^m \sum_{j:C_i \neq C_j} 2^{-1/2} \text{tr}(A_{ij} A_{ij}^T) \quad (4.5)$$

Μπορούμε να περιγράψουμε τον αλγόριθμο ως εξής:

στην  $k$  επανάληψη, βρες το  $P_k$  υπολογίζοντας το ελάχιστο της  $J$  κάνοντας αναζήτηση κατά μήκος της γεωδαισιακής  $\gamma$  που ορίζεται από την διεύθυνση  $H_{k-1}$  και διέρχεται από το σημείο  $P_{k-1} = \gamma(k-1)$ . Υπολόγισε την Riemann ποσότητα  $\nabla_p J(P_k)$  στο σημείο. Στη συνέχεια υπολόγισε τη νέα κατεύθυνση, τέτοια ώστε να αποτελεί συνδυασμό της παλιάς κατεύθυνσης και του νέου gradient. Επανάλαβε μέχρι τη σύγκλιση.

$$H_k \leftarrow -\nabla_p J(P_k) + \eta \tau(H_{k-1}, P_{k-1}, P_k),$$

Το Gradient  $\nabla_p J(P_k)$  μπορεί να προσεγγιστεί με τη βοήθεια του αντίστοιχου του στον Ευκλείδειο χώρο  $D_p J(P_k)$  με  $\nabla_p J(P_k) = D_P J(P_k) - P_k P_k^T D_P J(P_k)$ , και το  $\tau(H_{k-1}, P_{k-1}, P_k)$  συμβολίζει την παράλληλη μεταφορά του εφαπτόμενου διανύσματος  $H_{k-1}$  από  $P_{k-1}$  στο  $P_k$ . Ο υπολογισμός του  $D_p J(P_k)$ :

$$D_P J(P_k) = 1(S_w - \alpha S_b) P_k.$$

Δίνουμε τον παρακάτω αλγόριθμο.

---

**Αλγόριθμος 4.1** Εκμάθηση Προβολικής Μετρικής (PML)

---

**Require:** Για όλους τους γραμμικούς υπόχωρους  $\text{span}(Y) \in \mathcal{G}(q, D)$

- 1:  $W \leftarrow I_{D \times d}, P \leftarrow I_d$
  - 2: **Επανάλαβε:**
  - 3: Κανονικοποίησε το  $Y_i$  χρησιμοποιώντας 4.2 για όλα τα  $i$
  - 4: Υπολόγισε  $S_w$  και  $S_b$  χρησιμοποιώντας 4.4 και 4.5
  - 5: Βελτιστοποίησε  $P$  με 4.3 χρησιμοποιώντας τον παρακάτω αλγόριθμο.
  - 6: Ενημέρωσε το  $W$  υπολογίζοντας την τετραγωνική ρίζα του πίνακα  $P$ .
  - 7: **Τέλος**
  - 8: Έξοδος: Ο πίνακας  $P$ .
- 

---

**Αλγόριθμος 4.2** Riemannian Conjugate Gradient (RCG)

---

**Require:** Ο αρχικός PSD πίνακας  $P_0$

- 1:  $H_0 \leftarrow 0, P \leftarrow P_0$
  - 2: Επανάλαβε
  - 3:  $H_k \leftarrow -\nabla_P J(P_k) + \eta \tau(H_{k-1}, P_{k-1}, P_k)$ .
  - 4: Αναζήτησε γραμμικά κατά μήκος της γεωδαισιακής  $\gamma$  με κατεύθυνση το  $H_k$  από το  $P_{k-1} = \gamma(k-1)$  για να βρεθεί το  $P_k = \text{argmin}_P J(P)$ .
  - 5:  $H_{k-1} \leftarrow H_k, P_{k-1} \leftarrow P_k$ .
  - 6: Μέχρι τη σύγκλιση
  - 7: Έξοδος: Ο πίνακας  $P$ .
-

## ΚΕΦΑΛΑΙΟ 5

# ΕΥΣΤΑΘΗΣ ΜΑΘΗΣΗ ΔΙΑΚΡΙΤΗΣ ΠΡΟΒΟΛΗΣ ΓΙΑ ΠΟΛΛΑΠΛΗ ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΕ STIEFEL ΠΟΛΥΠΤΥΓΜΑΤΑ

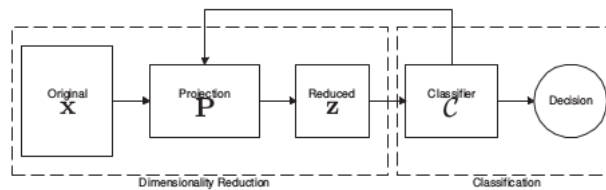
- 
- 5.1 Εισαγωγή
  - 5.2 Το προτεινόμενο πλαίσιο εκμάθησης
  - 5.3 Πολυπαραγοντική εμφύτευση
  - 5.4 Το Lasso
  - 5.5 Βελτιστοποίηση στα πολυπτώγματα Stiefel
- 

### 5.1 Εισαγωγή

Στο κεφάλαιο αυτό θα παρουσιάσουμε τις βασικές μεθόδους της εργασίας [18]. Για την εκμάθηση μιας προβολής, χρησιμοποιώντας μόνο ένα μικρό δείγμα εικόνων που περιέχουν πρόσωπα είναι ένα δύσκολο πρόβλημα στην αναγνώριση προσώπου. Ειδικά όταν τα πρόσωπα που θα πρέπει να αναγνωριστούν έχουν πολλές παραλλαγές ως προς τη θέση στην εικόνα, μεταβλητή φωτεινότητα, και έκφραση προσώπου. Για να μελετήσουν ένα τέτοιο πρόβλημα οι ερευνητές αυτής της εργασίας προτείνουν ένα πλαίσιο μεθόδων που στηρίζεται στην θεωρία της στατιστικής εκμάθησης, που είναι καταλληλότερη για την εκμάθησης μια διακριτής προβολής. Ο υποβιβασμός

διάστασης κάνοντας χρήση επαναληπτικά του προβολικού πίνακα συνδυάζεται με ένα γραμμικό ταξινομητή στο κανονικοποιημένο πλαίσιο της Lasso μεθόδου. Ο προβολικός πίνακας σε συνδυασμό με τις παραμέτρους του ταξινομητή μαθαίνονται, λύνοντας ένα πρόβλημα βελτιστοποίησης επί των πολυπτυγμάτων Stiefel.

## 5.2 Το προτεινόμενο πλαίσιο εκμάθησης



Σχήμα 5.1: Σχέδιο για την εκμάθηση του προβολικού πίνακα με επίβλεψη

Θεωρούμε τα αρχικά δεδομένα, να αναπαρίστανται σε μια διανυσματική δομή  $x \in \mathbb{R}^n$ . Μια γραμμική προβολή  $P \in \mathbb{R}^{n \times k}$  προβάλλει τα αρχικά δεδομένα  $x$  στα  $z = P^T x$ . Αυτή η προβολή όχι μόνο υποβιβάζει τη διάσταση στο  $k \ll n$  αλλά επίσης μειώνει τον θόρυβο των αρχικών δεδομένων πρώτου παραχθεί το  $z$  στον ταξινομητή  $C$ . Στην περίπτωση της μη επιβλέψιμης εκμάθησης, η προβολή  $P$  μπορεί να επιλεγεί ανεξάρτητα του ταξινομητή  $C$ . Στην περίπτωση εκμάθησης με επίβλεψη, η προβολή  $P$  επιλέγεται έτσι ώστε να εξασφαλιστεί, η ταξινόμηση των δεδομένων  $x$  τα οποία διατηρούν τον προβολικό χώρο  $z$ .

## 5.3 Πολυπαραγοντική εμφύτευση

Για να κατασκευαστεί μια τεχνική ταξινόμησης πολλών κατηγοριών κάνοντας χρήση μιας τεχνικής παλινδρόμησης, ο απλούστερος τρόπος είναι να κατασκευαστεί ένας δυαδικός ταξινομητής για κάθε κατηγορία σε σχέση με όλες τις άλλες. Η άλλη μέθοδος είναι να γενικευτεί ένας ταξινομητής για να κάνει πολυπαραγοντική παλινδρόμηση. Τα ονόματα των κλάσεων εμφυτεύονται σε αυτό τον πολυπαραγοντικό χώρο παλινδρόμησης. Για κάθε κορυφή  $y_i = [y_i^1, \dots, y_i^{k-1}]^T$  είναι το κέντρο της κλάσης. Ένα εμφυτευμένο σημείο κοντά στο  $y_i$  θα ταξινομείται στην κλάση με όνομα  $i$ .

## 5.4 Το Lasso

Θεωρούμε το γενικό γραμμικό μοντέλο παλινδρόμησης:

$$y_i = w^T z_i + \beta_0, \quad i = 1, \dots, L.$$

Ο αλγόριθμος lasso αναζητεί της παραμέτρους παλινδρόμησης  $w$  μέσω της συνθήκης

$$\arg \min_w \sum_{i=1}^L (y_i - w^T z_i - \beta_0)^2, \quad \text{subject to } \|w\|_1 \leq s.$$

Μπορεί ναδειχθεί ότι το πρόβλημα είναι ισοδύναμο με το

$$\arg \min_w \sum_{i=1}^L (y_i - w^T z_i - \beta_0)^2 + \lambda \|w\|_1,$$

για μια κατάλληλη τιμή του  $\lambda$ , η οποία υπολογίζεται με τη χρήση της cross-validation τεχνικής. Η διαφορά αυτής της παλινδρόμησης με την οριακή παλινδρόμησης, είναι ο τρόπος επιλογής του  $l_1$  σε σχέση με την κλασική μετρική  $l_2$ . Η επιλογή αυτής της μετρικής παράγει αραιούς συντελεστές.

Στην περίπτωση της επέκτασης του αλγορίθμου lasso στην πολυπαραγοντική περίπτωση με  $y \in \mathbb{R}^{k-1}$ , τότε το μοντέλο της γραμμικής παλινδρόμησης λαμβάνει τη μορφή

$$y_i = W^T z_i + b, \quad i = 1, \dots, L,$$

όπου  $W = [w_1, \dots, w_{k-1}]$ . Όταν η εμφύτευση του  $y_i$  είναι συμμετρική και τα δεδομένα είναι κεντρικοποιημένα, μπορούμε να ορίσουμε  $b = 0$ . Έτσι η βελτιστοποίηση μπορεί να γραφτεί σε μορφή πίνακα

$$\min_w \|Y - W^T Z\|_F^2 + \lambda \|W\|_1,$$

όπου  $Y = [y_1, \dots, y_L]$  και  $Z = [z_1, \dots, z_L]$ . Τότε και το  $Z$  και το  $Y$  είναι γνωστά, και

το πρόβλημα παλινδρόμησης lasso μπορεί να λυθεί αριθμητικά μια που δεν υπάρχει ακριβής μέθοδος.

## 5.5 Βελτιστοποίηση στα πολυπτύγματα Stiefel

Έστω  $Z = P^T X$  ο στόχος αυτού του αλγόριθμου είναι να βρει την προβολή  $P$  και να επιλέξει ένα κατάλληλο γραμμικό ταξινομητή για το παρακάτω πρόβλημα βελτιστοποίησης

$$\arg \min_{P \in \mathcal{S}_{n,k}, W \in \mathbb{R}^{k \times m}} \|Y - W^T P^T X\|_F^2 + \lambda \|W\|_1,$$

όπου

$$\mathcal{S}_{n,k} = \{P \in \mathbb{R}^{n \times k} : P^T P = I\}$$

είναι το πραγματικό Stiefel πολύπτυγμα. Η βελτιστοποίηση λύθηκε επί των Stiefel πολυπτυγμάτων  $\mathcal{S}_{n,k}$  για το  $P$  επειδή η αντικειμενική συνάρτηση δεν είναι αναλλοίωτη στους δεξιούς ορθογώνιους μετασχηματισμούς βαθμίδας  $k$  του προβολικού πίνακα  $P$ . Η αντικειμενική συνάρτηση είναι κυρτή σε σχέση με το  $W$  ή με το  $P$ , αλλά όχι και με τα δύο ταυτόχρονα. Μια υποβέλτιστη λύση παράγεται για να λύσει το πρόβλημα για κάθε μια μεταβλητή μέχρι τη σύγκλιση της. Τότε το  $P$  είναι δεδομένο και η αριθμητική λύση για το  $W$  μπορεί να παραχθεί κάνοντας χρήση του  $l_1$ -επιλυτή για το παρακάτω πρόβλημα:

$$\arg \min_{W \in \mathbb{R}^{k \times m}} \|Y - W^T Z\|_F^2 + \lambda \|W\|_1.$$

Τότε το  $W$  διατηρείται σταθερό, και πρέπει να λύσουμε το πρόβλημα

$$\arg \min_{P \in \mathcal{S}_{n,k}} g(P) = \|Y - W^T P^T X\|_F^2$$

Οι ερευνητές χρησιμοποιούν το παρακάτω θεώρημα [19].

**Θεώρημα 5.1.** (Προβολή στο χώρο Stiefel) Έστω  $X \in \mathbb{R}^{n \times k}$  να είναι ο βαθμίδα- $k$  πίνακας. Ορίζουμε τον προβολικό τελεστή  $\pi : \mathbb{R}^{n \times k} \rightarrow \mathcal{S}_{n,k}$  ως

$$\pi(X) = \arg \min_{Q \in \mathcal{S}_{n,k}} \|X - Q\|_F^2$$

Αν η SVD ανάλυση του πίνακα  $X$  είναι  $X = U\Sigma V^T$  τότε

$$\pi(X) = UI_{n,k}V^T$$

όπου  $I_{n,k}$  είναι οι πρώτες  $k$  του ταυτοτικού πίνακα  $I \in \mathbb{R}^{n \times n}$ .

Δίνουμε τον αλγόριθμο

**Αρχικοποίηση** Αρχικοποίηση  $P = P_0$  (από PCA, για παράδειγμα).

**Επανάληψη i**

$$F = \frac{\partial g(P)}{\partial P} = 2 (XX^T PWW^T - XY^T W^T)$$

**Βήμα 1** Όριση  $G = F - PF^T P, H = -G$ .

**Βήμα 2** Υπολόγισε

$$\langle H, H \rangle = \text{tr} \left[ H^T \left( I - \left( \frac{1}{2} PP^T \right) \right) H \right]$$

**Βήμα 3** Έλεγχξε αν  $\langle H, H \rangle \leq \varepsilon$  τότε σταμάτησε

**Βήμα 4** If  $g(P) - g(\pi(P + 2\gamma H)) \geq \gamma \langle H, H \rangle$  όρισε  $\gamma \rightarrow 2\gamma$  τότε επανέλαβε το βήμα 4

## ΒΙΒΛΙΟΓΡΑΦΙΑ

---

- [1] I. Chavel, *Riemannian Geometry*. 2006.
- [2] A. Goh and R. Vidal, “Clustering and dimensionality reduction on riemannian manifolds,” *CVPR08*, 2008.
- [3] O. Yamaguchi, K. Jukui, and K. Maeda, “Face recognition using temporal image sequence,” *FG,1998*, 1998.
- [4] L. Wolf and A. Shashua, “Learning over sets using kernel principal angles,” *JMLR*, vol. 4, pp. 913–931, 2003.
- [5] K. Fukui and O. Yamaguchi, “Face recognition using multi-viewpoint patterns for robot vision,” *Robotics Research*, pp. 192–201, 2005.
- [6] M. Nishiyama, O. Yamaguchi, and K. Fukui, “Face recognition with the multiple constrained mutual subspace method,” *Audio-and Video-Based Biometric Person Authentication*, pp. 71–80, 2005.
- [7] H. Zhiw, S. Ruiping, W. and Shiguang, and C. Xilin, “Projection metric learning on grassmann manifold with application to video based face recognition,” *CVPR*, 2015.
- [8] H. J. and L. D. D., “Grassmann discriminant analysis: a unifying view on subspace-based learning,” *ICML*, 2008.
- [9] H. J. and L. D. D., “Extended grassmann kernels for subspace-based learning,” 2008.
- [10] M. T. Harandi, M. Salzmann, S. Jayasumana, and R. L. H. Hartley, “Expanding the family of grassmannian kernels: An embedding perspective.,” *ECCV*, 2014.



- [11] M. T. Harandi, C. Sanderson, S. Shirazi, and B. C. Lovell, “Graph embedding discriminant analysis on grassmannian manifolds for improved image set matching,” *CVPR*, 2011.
- [12] R. Vemulapalli, J. Pillai, and R. Chellapa, “Kernel learning for extrinsic classification of manifold features,”
- [13] H. J. L. D.D., “Grassmann discriminant analysis: a unifying view on subspace-based learning,” *ICML, 2008*, 2008.
- [14] Y. Wong, “Differential geometry of grassmann manifolds,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 57(3), p. 589, 1967.
- [15] A. Edelman and A. S. S. Arias, T.A., “The geometry of algorithms with orthogonality constraints,” *SIAM journal on Matrix Analysis and Applications*, vol. 20(2), pp. 303–353, 1998.
- [16] A. Srivastava and E. Klassen, “Bayesian and geometric subspace tracking,” *Advances in Applied Probability*, pp. 43–56, 2004.
- [17] R. Absil, P.A. and Mahony and R. Sepulchre, *Optimization algorithms on matrix manifolds*. 2008.
- [18] D. Pham and S. Venkatesh, “Robust learning of discriminative projection for multicategory classification on the stiefel manifold,” *CVPR*, pp. 43–56, 2008.
- [19] J. H. Manton, “Optimization algorithms exploiting unitary constraints,” *IEEE Transactions on Signal Processing*, pp. 4311–4322, Mar 2002.

## ΣΥΝΤΟΜΟ ΒΙΟΓΡΑΦΙΚΟ

---

Katsigiannis Theophilos was born in Athens in 1971. He was admitted at the Mathematics Department of the University of Ioannina in 1992. He received his BSc degree in mathematics in 1996 with degree 6.72 “very good”. That year he became a postgraduate student at the same institution from where he graduated in 1998. In 2008 he was admitted at the Computer Science Department of the University of Ioannina. He received his BSc degree in computer science in 2015 with degree 7.89 “very good”. That year he became a postgraduate student at the same institution from where he graduated in February 2018. His main research interests lie in the field of Computer Vision and Pattern Recognition. He is working at Secondary School Education since 2005 as a mathematician.