



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Πολιτικών Μηχανικών

Εργαστήριο Στατικής και Αντισεισμικών Ερευνών

Διπλωματική Εργασία

**Επιτάχυνση ανάλυσης Monte Carlo σε
δυναμικό πρόβλημα με χρήση μεθόδων
μείωσης διαστάσεων**

Όνοματεπώνυμο Φοιτητή:
Κωνσταντίνος Ατζαράκης

Επιβλέπων καθηγητής:
Βησσαρίων Παπαδόπουλος



10 Νοεμβρίου 2020

«Truth is ever to be found in simplicity, and not in the multiplicity and confusion of things.»

Sir Isaac Newton

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Περίληψη

Σχολή Πολιτικών Μηχανικών

Εργαστήριο Στατικής και Αντισεισμικών Ερευνών

Διπλωματική

Επιτάχυνση ανάλυσης Monte Carlo σε δυναμικό πρόβλημα με χρήση μεθόδων μείωσης διαστάσεων

by Κωνσταντίνος Ατζαράκης

Στην σύγχρονη εποχή, παρά την ύπαρξη εξαιρετικών μεθόδων (FEM, BEM etc.) και παρά την αύξηση της υπολογιστικής ισχύος, αποτελεί ακόμα ζητούμενο ή χρήση ταχύτερων αλγορίθμων για την επίλυση φυσικών προβλημάτων μηχανικού. Η αριθμητική ανάλυση *Monte Carlo*, αν και αξιόπιστη για την επίλυση του στοχαστικού δυναμικού προβλήματος, συνήθως απομειώνει εξαιρετικά σύντομα τους όποιους διαθέσιμους υπολογιστικούς πόρους και ιδιαίτερα σε συστήματα με πολλούς βαθμούς ελευθερίας. Η παρούσα εργασία εξετάζει την χρήση δύο μεθόδων μείωσης διαστάσεων (*Dimensionality Reduction methods*) προς επιτάχυνση της ανάλυσης Monte Carlo. Μετά από μία σύντομη εισαγωγή στις έννοιες της υποκατάστατης μοντελοποίησης (*surrogate modeling*) και των μοντέλων μειωμένης τάξης (*Reduced Order Modeling*), δίνονται συνοπτικά, κάποιες χρήσιμες μαθηματικές έννοιες για τον μη εξοικειωμένο αναγνώστη. Στην συνέχεια, σε ξεχωριστά κεφάλαια, παρουσιάζονται οι δύο μέθοδοι μείωσης διαστάσεων που θα χρησιμοποιηθούν, ή κλασική μέθοδος της ανάλυσης σε Κύριες Συνιστώσες (*Principal Component Analysis*) και η σχετικά πρόσφατη μέθοδος των *Απεικονίσεων (ή Χαρτών) Διάχυσης (Diffusion Maps)*. Οι μέθοδοι εφαρμόζονται σε αριθμητικό παράδειγμα για την δημιουργία δύο υποκατάστατων φυσικών μοντέλων τα αποτελέσματα των οποίων συγκρίνονται, εν τέλει, με αυτά της ακριβούς Monte Carlo ανάλυσης.

Εὐχαριστίες

Εὐχαριστῶ ἰδιαίτερος τὸν ἐπιβλέποντα καθηγητὴ μου κ. Βησσαρίωνα Παπαδόπουλο γιὰ τὴν καθοδήγησή του καθ' ὅλη την διάρκεια ἐκπόνησης τῆς ἐργασίας, γιὰ τὶς ἔγκαιρες ἀποκρίσεις του σὲ ὅλα μου τὰ ἐρωτήματα καὶ τὶς παροτρύνσεις του νὰ παρευρίσκομαι, ἀκαδημαϊκὰ νήπιος ὢν, σὲ δραστηριότητες τῆς ἐπιστημονικῆς του ὁμάδας.

Εὐχαριστῶ θερμὰ τὸν ὑποψήφιο διδάκτορα κ. Ιωάννη Καλογερό για τὸν πολύτιμο χρόνο ποὺ ἀφιέρωσε ἀφειδῶς μοιραζόμενος τὴν ἐμπειρία του, δίνοντας χρήσιμες συμβουλές καὶ προλαμβάνοντας τυχούσες παρατυπίες. Παρόμοιο ζήλο καὶ προθυμία γιὰ βοήθεια ἐπέδειξαν ὅλα τὰ μέλη τῆς ἐπιστημονικῆς ὁμάδας στὰ ὁποῖα ὀφείλω εὐγνωμοσύνη καὶ ποὺ συγχαίρω γιὰ τὸ κλίμα ἐνότητας καὶ συνεργασίας ποὺ ἔχουν δημιουργήσει.

Τέλος, εὐχαριστῶ ὅσους συγγενεῖς καὶ φίλους, ἐμφανῶς ἢ ἀφανῶς, μὲ παρότρυναν, μὲ στήριξαν καὶ μὲ βοήθησαν νὰ ὀλοκληρώσω τὶς προπτυχιακὲς σπουδές μου, δίνοντάς μου τὴν πολυτέλεια νὰ ἀσχολούμαι ἀμέριμνα μὲ τὴν ἐκμάθηση τῆς ἐπιστήμης τοῦ Πολιτικοῦ Μηχανικοῦ.

Περιεχόμενα

Περίληψη	v
Ευχαριστίες	vii
Κατάλογος Σχημάτων	xi
Κατάλογος Πινάκων	xiii
1 Εισαγωγή	1
1.1 Μεταμοντελοποίηση	1
1.2 Προσομιώματα Μειωμένης Τάξης	2
1.3 Μαθηματική διατύπωση του προβλήματος	2
1.4 Δυναμικό πρόβλημα	3
1.5 Σύνοψη μεθοδολογίας	5
2 Στοιχεία από την Θεωρία Στοχαστικῶν Ἀνελιξέων	7
2.1 Ὅρισμός	7
2.2 Συναρτήσεις μάζας και πυκνότητας	8
2.3 Συναρτήσεις αὐτοσυσχέτισης και αὐτοδιακύμανσης	9
2.4 Γνωστές στοχαστιές διαδικασίες	10
2.4.1 Διαδικασίες Gauss	10
2.4.2 Διαδικασίες Markov	10
Ἄλυσίδες Markov	10
2.5 Ἀναπαράσταση στοχαστικῆς διαδικασίας	12
2.6 Ἡ μέθοδος Karhunen-Loève	13
3 Στοιχεία από την Θεωρία Γραφημάτων	15
3.1 Στόχος	15
3.2 Ὅρισμός	15
3.2.1 Διαφορικοί τελεστές σὲ γράφους	16
3.2.2 Σύνδεση με την Θεωρία Πινάκων	16
3.2.3 Ἰδιομορφική Ἀνάλυση τῆς Λαπλασιανῆς	18
4 Ἡ μέθοδος τῆς Ἀνάλυσης σὲ Κύριες Συνιστώσες (PCA)	19
4.1 Επισκόπηση	19
4.2 Μεγιστοποίηση τῆς διασπορᾶς	20
4.3 Ἰδιομορφική ἀνάλυση	21
4.4 Ὄρθη και ἀντίστροφη ἀπεικόνιση – Μείωση Διαστάσεων	22
4.5 Σύνοψη τῆς μεθόδου	23
5 Ἡ μέθοδος τῶν Ἀπεικονίσεων Διάχυσης (Diffusion Maps)	25
5.1 Επισκόπηση	25
5.2 Πυρήνας και ἐγγύτητα δεδομένων	25

5.3	Κατασκευή οικογένειας διαχύσεων	26
5.4	Κατασκευή τυχαίου περιπάτου στα δεδομένα	28
5.5	Διαδικασία Διάχυσης	28
5.6	Διαχυτική Απόσταση	29
5.7	Διαχυτική Απεικόνιση - Diffusion Map	30
5.8	Αντίστροφη Απεικόνιση	31
5.9	Σύνοψη τής μεθόδου	32
6	Αριθμητικές Έφαρμογές	33
6.1	Έποπτικό Παράδειγμα μείωσης διαστάσεων	33
6.2	Δυναμικό πρόβλημα	36
6.2.1	Όρισμός	36
6.2.2	Καθορισμός μοντέλου	36
6.2.3	Διαλογή δειγμάτων	37
6.2.4	Αποτελέσματα	39
6.2.5	Συμπεράσματα	44
6.2.6	Σημεία πρὸς ἐρευνητική ἐμβάθυνση	44
	Βιβλιογραφία	47

Κατάλογος Σχημάτων

1.1	Διάγραμμα κατασκευής του υποκατάστατου μοντέλου. Με κόκκινο χρώμα σημειώνεται η συνήθης πορεία υπολογισμού των αποτελεσμάτων, ενώ με πράσινο ή πορεία μέσω της χρήσης υποκατάστατου μοντέλου.	2
1.2	Μετασχηματισμοί μεταξύ του $\mathcal{E} \in \mathbb{R}^p$ και του χώρου μειωμένων διαστάσεων και μεταξύ του χώρου μειωμένων διαστάσεων και του $\mathcal{U} \in \mathbb{R}^d$ [1].	4
2.1	Μία στοχαστική διαδικασία αποτελεί απεικόνιση από τον χώρο των τυχαίων γεγονότων στον χώρο των συναρτήσεων.	7
2.2	Πραγματοποίηση μιᾶς στοχαστικής διαδικασίας Wiener, γνωστής και ως κίνησης Brown. Το σύνολο δεικτών της διαδικασίας είναι οι μη αρνητικοί πραγματικοί αριθμοί, ενώ ο χώρος κατάστασης είναι ο τριδιάστατος εὐκλείδιος χώρος.	8
2.3	Γράφημα ή διάγραμμα κατάστασης μαρκοβιανής αλυσίδας	10
2.4	Φθίνουσες ιδιοτιμές από την λύση του ολοκληρώματος Fredholm δευτέρου εΐδους για $M = 10$ [3].	14
4.1	Το σημεία ενός συνόλου δεδομένων και τὰ δύο συστήματα συντεταγμένων, τὸ ἀρχικό σύστημα (μαῦρο) και τὸ κύριο σύστημα (κόκκινο).	19
5.1	Ἡ παράμετρος ϵ ἐπιλέγεται μακριά ἀπὸ τὰ ἄκρα ποὺ ἀποτελοῦν ἐκφυλισμένες περιπτώσεις με ὑπερβολικά μεγάλα ἢ μικρά βάρη στὸ ἀντίστοιχο μητρώο.	27
5.2	Μονοπάτια κατὰ τὴν γεωμετρική δομή τῶν δεδομένων ἐμφανίζονται με μεγαλύτερη πιθανότητα.	29
6.1	Τυχαῖες ὑλοποιήσεις τῶν t και ϕ	33
6.2	Σύνολο δεδομένων "swiss roll" σὲ 3Δ σύμφωνα με τὴν ἐξίσωση $(t, \phi \sin(\phi), \phi \cos(\phi))$	34
6.3	Ἰδιοτιμές: κανονικοποιημένη γιὰ τὴν PCA, ἐνῶ γιὰ τὰ DMAPs ἡ πρώτη δεν δίνει πληροφορία ὄντας πάντα ἴση με τὴν μονάδα.	35
6.4	Τὸ σύνολο δεδομένων στὶς 2Δ, ὅπως μειώθηκε ἀπὸ κάθε τεχνική – παράμετροι DMAPs ($\epsilon = 1, \alpha = 0, t = 1$).	35
6.5	Οἱ διαστάσεις, οἱ συνοριακές συνθήκες και τὸ πλέγμα τοῦ προβλήματος.	36
6.6	Στοχαστικά πεδία τοῦ μέτρου ἐλαστικότητας (E) και τῆς σεισμικῆς διέγερσης ($\vec{\alpha}_g$).	37
6.7	Φόρτιση με σταθερή ἐπιτάχυνση ἐδάφους α_g και ἡ ἀντίστοιχη ἀπόκριση κόμβου ὅπως προκύπτει ἀπὸ τὸ πλήρες μοντέλο (Full Order Model).	38
6.8	Φόρτιση με ἡμιτονοειδή ἐπιτάχυνση ἐδάφους α_g και ἡ ἀντίστοιχη ἀπόκριση κόμβου ὅπως προκύπτει ἀπὸ τὸ πλήρες μοντέλο FOM.	38
6.9	Ἀπὸ κάθε χρονοσειρά ἐπιλέγονται $N_S = \frac{N_{\text{timesteps}}}{S}$ λύσεις, ὅπου S τὸ χρονικό βήμα.	39

6.10	Σύγκριση αποκρίσεων για τόν βαθμό ελευθερίας 101 – στην βάση του προβόλου – με χρήση $k = 3$ ιδιοτιμών στα υποκατάστατα μοντέλα (Reduced Order Models).	40
6.11	Σύγκριση αποκρίσεων για τόν βαθμό ελευθερίας 2498 – στο ελεύθερο άκρο του προβόλου – με χρήση $k = 3$ ιδιοτιμών στα υποκατάστατα μοντέλα (ROM)	40
6.12	Σύγκριση αποκρίσεων για τόν βαθμό ελευθερίας 100 με χρήση $k = 3$ ιδιοτιμών στα υποκατάστατα μοντέλα.	41
6.13	Σύγκριση αποκρίσεων για τόν βαθμό ελευθερίας 100 με χρήση $k = 5$ ιδιοτιμών στα υποκατάστατα μοντέλα.	41
6.14	Συναρτήσεις πυκνότητας πιθανότητας (pdf) για τόν βαθμό ελευθερίας 100 όπως προκύπτει από την πλήρη Monte Carlo ανάλυση (FOM) και τα υποκατάστατα μοντέλα (ROM-DMAPs, ROM-PCA).	42
6.15	Συναρτήσεις πυκνότητας πιθανότητας (pdf) για τόν βαθμό ελευθερίας 2499 όπως προκύπτει από την πλήρη Monte Carlo ανάλυση (FOM) και τα υποκατάστατα μοντέλα (ROM-DMAPs, ROM-PCA).	43

Κατάλογος Πινάκων

6.1	Παράμετροι τῶν τυχαίων μεταβλητῶν.	34
6.2	Παράμετροι τῶν τυχαίων μεταβλητῶν.	37
6.3	Χρόνοι ἐκτέλεσης γιὰ τὸ πλήρες καὶ τὰ ὑποκατάστατα μοντέλα.	44

Κεφάλαιο 1

Είσαγωγή

1.1 Μεταμοντελοποίηση

Πολλές από τις επιστημονικές εφαρμογές ή τις εφαρμογές μηχανικού απαιτούν την προσομοίωση φυσικών συστημάτων με ακριβή μοντέλα υψηλού κόστους. Έξ όρισμού, καθώς ο φυσικός κόσμος, καίτοι περιγραφτός, δεν δεσμεύεται να είναι και απλός, τὰ προκύπτοντα φυσικά προβλήματα κληρονομούν συχνά την πραγματική περιπλοκότητα, παρὰ τις ανθρώπινες απλοποιητικές παραδοχές. Σὲ αὐτήν τὴν ὑπάρχουσα περιπλοκότητα ἔρχεται ἐπιπλέον νὰ προστεθεῖ καὶ αὐτὴ τῶν μεθόδων ἐπίλυσης τῶν προβλημάτων. Ἐφαρμογές, ὅπως γιὰ παράδειγμα ὁ σχεδιασμός, ἡ βελτιστοποίηση καὶ ἡ στοχαστικὴ μοντελοποίηση, περιλαμβάνουν συχνά μεγάλο ἀριθμὸ μεταβλητῶν καὶ ἐπαναληπτικότητα διαδικασιῶν. Ἔτσι, λοιπόν, δὲν εἶναι σπάνιο τὸ ἐνδεχόμενο ἡ ἀριθμητικὴ προσομοίωση νὰ χρειάζεται ικανότατο χρόνο γιὰ νὰ ὁλοκληρωθεῖ – πολλές ὥρες ἢ καὶ μέρες – ἀκόμα καὶ μὲ τὴν χρῆση τῆς ἐπιτομῆς τοῦ σύγχρονου τεχνολογικοῦ ὑλικοῦ. Αὐτὸ καθιστᾷ τις προαναφερθεῖσες ἐφαρμογές δυσπροσάρμοστες γιὰ πρακτικὴ χρῆση, παρόλο πὺ εἶναι καθόλα χρήσιμες καὶ ὠφέλιμες.

Ἡ *Μεταμοντελοποίηση (Meta-modeling)* ἢ *Ἐποκατάστατη Μοντελοποίηση (Surrogate Modeling)* εἶναι τεχνικὴ πὺ ἔχει σκοπὸ νὰ μειώσῃ τὸ ὑπολογιστικὸ κόστος τῶν ἀκριβῶν προσομοιωμάτων ἀντικαθιστώντας τα μὲ ἀποδοτικότερα προσεγγιστικὰ μοντέλα. Ἡ πρόκληση εἶναι τὸ ὑποκατάστατο μοντέλο νὰ παρέχει ικανοποιητικὴ ἀκρίβεια, χρησιμοποιώντας τις ελάχιστες δυνατές πληροφορίες ἀπὸ τὸ πρωτότυπο.

Τὰ *ἀμιγῆ μεταμοντέλα* κατασκευάζονται μὲ καθαρὰ δεδομενοήλατη (data-driven) λογικὴ καὶ δὲν στηρίζονται στὴν γνώση τῆς ἐσωτερικῆς συμπεριφορᾶς τοῦ ἀκριβοῦς μοντέλου, ἀλλὰ μόνο στὴν σχέση τῶν δεδομένων εἰσόδου-ἐξόδου. Δὲν ἀπαιτεῖται, δηλαδή, καμία γνώση τῆς δομῆς τοῦ πρωτοτύπου κατὰ τὴν χρῆση τοῦ μοντέλου¹. Σὲ αὐτὴν τὴν κατηγορίᾳ ἀνήκουν, γιὰ παράδειγμα, τὰ εὐρέως διαδεδομένα νευρωνικὰ δίκτυα.

Τὰ *μεικτὰ μεταμοντέλα*, ἀπὸ τὴν ἄλλη, παρεμβάλλονται μεταξὺ τοῦ πρωτοτύπου καὶ τῶν ἀποτελεσμάτων, δηλαδή χρησιμοποιοῦν καὶ τὴν ἐσωτερικὴ δομὴ τοῦ ἀρχικοῦ προσομοιωματος. Μεικτὰ μοντέλα εἶναι, παραδείγματος χάριν, ὅλα τὰ *μοντέλα μειωμένης τάξης*, τὰ ὁποῖα θὰ ἀναλυθοῦν στὴν συνέχεια.

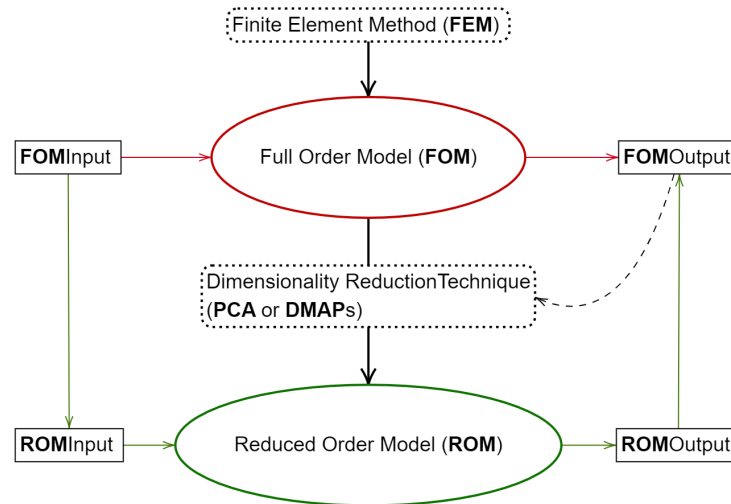
Ἄξιζει νὰ σημειωθεῖ, ὅτι δὲν ὑπάρχει περιορισμὸς στὴν σύνθεση μεταμοντέλων. Εἶναι δυνατὸν, ἐπὶ παραδείγματι, νὰ γίνῃ συνδυασμὸς νευρωνικῶν δικτύων καὶ μείωσης τῆς διάστασης τῶν δεδομένων.

¹Φυσικά, γίνεται χρῆση τοῦ ἀρχικοῦ μοντέλου ἅπαξ, προκειμένου νὰ ἐξαχθοῦν τὰ δεδομένα ἐξόδου μὲ τὰ ὁποῖα ἐκπαιδεύεται τὸ μεταμοντέλο.

1.2 Προσομοιώματα Μειωμένης Τάξης

Κινούμενο στο προηγούμενο πλαίσιο, τὸ ενδιαφέρον τῆς ἐπιστημονικῆς κοινότητας ἐσχάτως στρέφεται, μεταξύ ἄλλων, σὲ μεθόδους μείωσης τῶν διαστάσεων τῶν μεταβλητῶν τοῦ συστήματος. Τὸ σύνολο τῶν μεθόδων αὐτῶν στὴν βιβλιογραφία αναφέρεται ὡς *Προσομοίωση Μειωμένης Τάξης* ἢ συντομότερα **PMT** (*Reduced Order Modeling* – **ROM**).

Ἡ παρούσα ἐργασία ἐπικεντρώνεται στὴν ἐπιστήμη τῆς ὑπολογιστικῆς μηχανικῆς καὶ γιὰ τὴν κατασκευή μεταμοντέλων θὰ γίνεταί χρήση τῆς μεθόδου τῶν *Πεπερασμένων Στοιχείων*. Τὸ ἕκαστοτε **ROM** θὰ ὑλοποιηθεῖ μὲ συνδυασμὸ τῆς **FEM**, ἀφενός, καὶ μιᾶς τεχνικῆς μείωσης διαστάσεων ἀφετέρου. Οἱ συνδυαζόμενες μὲ τὴν **FEM** τεχνικὲς ἀνήκουν στὴν εὐρύτατη οἰκογένεια τῆς *μηχανικῆς μάθησης* καὶ συγκεκριμένα στὸν κλάδο τῆς *πολυπτυγματικῆς μάθησης* (*manifold learning*). Ἡ πρώτη εἶναι ἡ γνωστὴ **PCA** (βλ. σελ. 19), ἐνῶ ἡ δευτέρη ἡ ἀνερχόμενη μέθοδος τῶν **Diffusion Maps** (βλ. σελ. 25). Οἱ τεχνικὲς θὰ συνδυαστοῦν ἀμφότερες ἀνὰ μία μὲ τὴν **FEM** παράγοντας δύο μεταμοντέλα.



Σχῆμα 1.1: Διάγραμμα κατασκευῆς τοῦ ὑποκατάστατου μοντέλου. Μὲ κόκκινο χρῶμα σημειώνεται ἡ συνήθης πορεία ὑπολογισμοῦ τῶν ἀποτελεσμάτων, ἐνῶ μὲ πράσινο ἡ πορεία μέσω τῆς χρήσης ὑποκατάστατου μοντέλου.

Δέον νὰ σημειωθεῖ, ὅτι παρόλη τὴν ἐξειδίκευση τῆς παρουσίης ἐργασίας στὸν κλάδο τῆς μηχανικῆς, οἱ ἐφαρμοζόμενες τεχνικὲς μείωσης διαστάσεων, λόγω τοῦ γενικοῦ τους χαρακτήρα δεν περιορίζονται στὶς ἐπιστῆμες μηχανικοῦ, ἀλλὰ ἐφαρμόζονται σὲ πλῆθος ἐπιστημῶν ὅπως στὴν Ἐπιστῆμη Δεδομένων, στὴν Στατιστικὴ, στὴν Μηχανικὴ Μάθηση κ.α.

1.3 Μαθηματικὴ διατύπωση τοῦ προβλήματος

Σὲ αὐτὸ τὸ σημεῖο εἶναι πρέπον νὰ ξεκινήσει ἡ μαθηματικὴ περιγραφή τοῦ προβλήματος, ἔτσι ὥστε νὰ γίνεῖ κατανοητὸ τὸ ὅλο πλαίσιο ἀπὸ τὸν ἀναγνώστη.

Ἐστω, λοιπόν, ὅτι δίδεται ἀκριβὲς μοντέλο $\mathcal{M} : \mathbb{R}^p \rightarrow \mathbb{R}^d$ ποὺ χρησιμοποιεῖ d μεταβλητές. Μὲ \mathbb{R}^p σημειώνεται ὁ χῶρος τῶν παραμέτρων τοῦ μοντέλου, δηλαδή τῶν δεδομένων εἰσόδου καὶ μὲ \mathbb{R}^d ὁ χῶρος τῶν ἀποτελεσμάτων, δηλαδή τῶν δεδομένων ἐξόδου. Συχνότατα στὶς ἐφαρμογὲς ἰσχύει $d \gg p$, ἀπαιτοῦνται πολλοὶ ὑπολογισμοὶ γιὰ διάφορες τιμὲς τῶν παραμέτρων καὶ ὁ ὑπολογισμὸς τῶν ἀποτελεσμάτων ἀπευθείας μέσω τοῦ ἀκριβοῦς

μοντέλου είναι, όπως προειπώθηκε, κοστοβόρος. Προκύπτει, λοιπόν, ένα εύλογο ερώτημα: είναι απαραίτητο, ενώ κάθε φορά αλλάζουν μόνο ν μεταβλητές, να γίνονται εκ νέου υπολογισμοί που απαιτούν d μεταβλητές;

Η απάντηση είναι αρνητική, αν συλλογιστεί κανείς τα επόμενα. Είναι γνωστό, ότι ο μετασχηματισμός \mathcal{M} είναι για τα προβλήματα της μηχανικής εν γένει συνεχής, όπως και σε πολλά φυσικά συστήματα. Τότε, λοιπόν, είναι ασφαλής ή υπόθεση ότι τα αποτελέσματα $\mathcal{M}(\mathbf{e}) \in \mathbb{R}^d$ που προκύπτουν από τις διάφορες παραμέτρους $\mathbf{e} \in \mathbb{R}^\nu$ δεν είναι ασυνάρτητα, αλλά κείνται κατά προσέγγιση κι αυτά σε κάποιο υπόχωρο του \mathbb{R}^d . Ο υπόχωρος αυτός είναι εν γένει μη γραμμικός και αποτελεί ένα ν -διάστατο πολύπτυγμα έμβαπτισμένο στον \mathbb{R}^d . Αν με κάποιο τρόπο βρεθεί ή δομηθεί αυτό το πολύπτυγμα, τότε δεν θα είναι απαραίτητη η χρήση του πλήρους μοντέλου, αλλά θα χρησιμοποιείται η πληροφορία του πολύπτυγματος για την κατασκευή ενός μεταμοντέλου.

1.4 Δυναμικό πρόβλημα

Ειδικεύοντας, η δυναμική εξίσωση ισορροπίας ενός γραμμικού συστήματος γράφεται ως στοχαστικό πρόβλημα d μεταβλητών

$$\mathbf{M}(\mathbf{e})\ddot{\mathbf{u}}(\mathbf{e}) + \mathbf{C}(\mathbf{e})\dot{\mathbf{u}}(\mathbf{e}) + \mathbf{K}(\mathbf{e})\mathbf{u}(\mathbf{e}) = \mathbf{f}(\mathbf{e}), \quad (1.1)$$

όπου

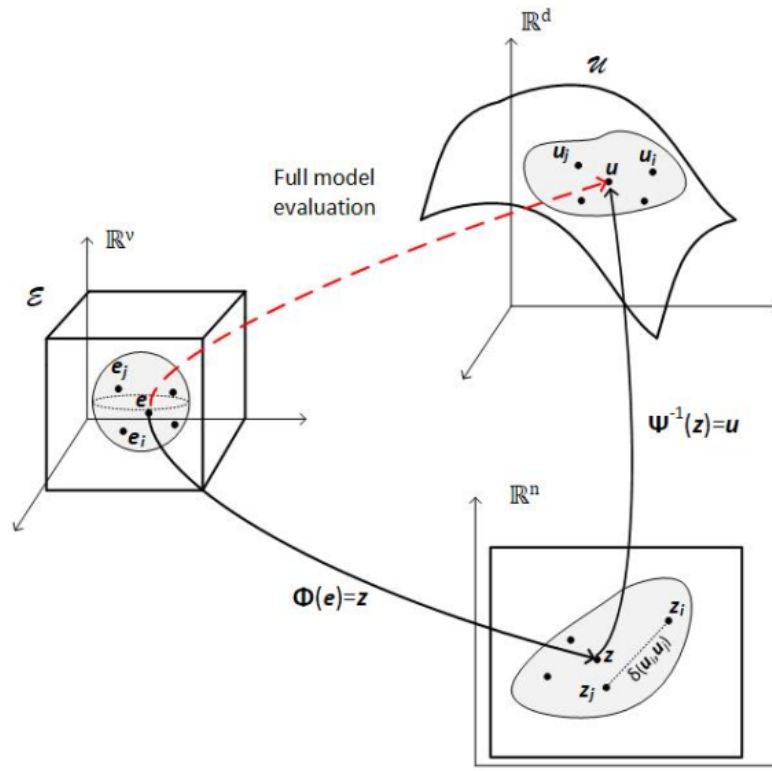
- $\mathbf{e} \in \mathbb{R}^\nu$ διάνυσμα που περιγράφει κάθε τυχαία παράμετρο του συστήματος,
- $\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}} \in \mathbb{R}^d$ τα διανύσματα της μετακίνησης, ταχύτητας και επιταχύνσεως αντίστοιχα,
- $\mathbf{M}, \mathbf{C}, \mathbf{K} \in \mathbb{R}^{d \times d}$ το μητρώο μάζας, το μητρώο απόσβεσεως και το μητρώο στιβαρότητας,
- $\mathbf{f} \in \mathbb{R}^d$ το διάνυσμα των εξωτερικών δράσεων.

Αναζητείται, καταρχήν, η γεωμετρία του πολύπτυγματος \mathcal{M} στο οποίο κείνται οι λύσεις \mathbf{u} του συστήματος. Στην συνέχεια ή γνώση αυτή θα βοηθήσει στην δημιουργία του υποκατάστατου μοντέλου για το φυσικό σύστημα. Το μοντέλο αυτό θα πρέπει να είναι ικανό να προσεγγίσει με ακρίβεια και αποδοτικότητα την πραγματική λύση \mathbf{u} , δεδομένων παραμέτρων \mathbf{e} , χωρίς την χρήση της εξίσωσης 1.1.

Προκειμένου, λοιπόν, να βρεθεί ή απεικόνιση \mathbf{u} ενός παραμετρικού διανύσματος \mathbf{e} , βρίσκεται, αντί της επίλυσης του πλήρους συστήματος, ή απεικόνιση \mathbf{z} του \mathbf{e} σε χώρο μειωμένων διαστάσεων μέσω του μετασχηματισμού $\Phi(\mathbf{e}) = \mathbf{z}$. Αν δεν εγκαταλειφθεί πλήρως το φυσικό σύστημα, όπως στην παρούσα εργασία, το \mathbf{z} βρίσκεται έμμεσα από την επίλυση της εξίσωσης της δυναμικής ισορροπίας στον χώρο μειωμένων διαστάσεων \mathbb{R}^n . Στην συνέχεια μέσω της σύνθεσης $(\Psi^{-1} \circ \Phi)(\mathbf{e})$ βρίσκεται ή ζητούμενη απεικόνιση του \mathbf{e} στο πολύπτυγμα \mathcal{M} . Τον όρθο $\Psi : \mathbb{R}^d \rightarrow \mathbb{R}^n$ τόν δίνει ή εκάστοτε μέθοδος μείωσης διαστάσεων. Μια σχηματική αναπαράσταση των μετασχηματισμών μπορεί να δεί κανείς στην εικόνα 1.2.

Πιο συγκεκριμένα, για γραμμικό μετασχηματισμό Ψ^{-1} , δηλαδή $\mathbf{u} = \Psi^{-1}\mathbf{z}$ και παραλείποντας την εξάρτηση από το \mathbf{e} για συντομία ή 1.1 γίνεται

$$\begin{aligned} & \mathbf{M}\Psi^{-1}\ddot{\mathbf{z}} + \mathbf{C}\Psi^{-1}\dot{\mathbf{z}} + \mathbf{K}\Psi^{-1}\mathbf{z} = \mathbf{f} \\ \implies & (\Psi^{-1})^T \mathbf{M}\Psi^{-1}\ddot{\mathbf{z}} + (\Psi^{-1})^T \mathbf{C}\Psi^{-1}\dot{\mathbf{z}} + (\Psi^{-1})^T \mathbf{K}\Psi^{-1}\mathbf{z} = (\Psi^{-1})^T \mathbf{f} \end{aligned}$$



Σχήμα 1.2: Μετασχηματισμοί μεταξύ του $\mathcal{E} \in \mathbb{R}^v$ και του χώρου μειωμένων διαστάσεων και μεταξύ του χώρου μειωμένων διαστάσεων και του $\mathcal{U} \in \mathbb{R}^d$ [1].

$$\implies \mathbf{M}_r \ddot{\mathbf{z}} + \mathbf{C}_r \dot{\mathbf{z}} + \mathbf{K}_r \mathbf{z} = \mathbf{f}_r, \quad (1.2)$$

όπου

- $\mathbf{z}, \dot{\mathbf{z}}, \ddot{\mathbf{z}} \in \mathbb{R}^n$ τὰ διανύσματα μειωμένων διαστάσεων τῆς μετακίνησης, ταχύτητας και ἐπιταχύνσεως ἀντίστοιχα,
- $\mathbf{M}_r, \mathbf{C}_r, \mathbf{K}_r \in \mathbb{R}^{n \times n}$ τὸ μητρώο μάζας, τὸ μητρώο ἀποσβέσεως και τὸ μητρώο στιβαρότητας στὶς μειωμένες διαστάσεις,
- $\mathbf{f}_r \in \mathbb{R}^n$ τὸ ἀντίστοιχο διάνυσμα τῶν ἐξωτερικῶν δράσεων.

Ἐτσι, επιτυγχάνεται μείωση τοῦ ὑπολογιστικοῦ κόστους, ἀφοῦ, ἀντὶ νὰ λυθεῖ ἡ διαφορική ἐξίσωση 1.1 μὲ πλῆθος ἀγνώστων ὅσο και οἱ βαθμοὶ ἐλευθερίας τοῦ πλήρους μοντέλου, λύνεται ἡ 1.2 ποὺ θὰ δώσει, ὅπως προειπώθηκε, προσεγγιστικὰ τὸ ἴδιο ἀποτέλεσμα, ἀλλὰ μὲ ἀρκετὰ λιγότερο ἀριθμὸ μεταβλητῶν.

Ἀφοῦ ἔχει βρεθεῖ τὸ \mathbf{z} , τότε $\mathbf{u} = \Psi^{-1} \mathbf{z}$ χρησιμοποιώντας τὸν γραμμικὸ μετασχηματισμὸ Ψ^{-1} , ποὺ μέχρι στιγμῆς εἶναι ἀγνώστος. Οἱ μέθοδοι εὑρεσῆς του θὰ παρουσιαστοῦν σὲ ἐπόμενα κεφάλαια. Δέον νὰ σημειωθεῖ ὅτι ἀπὸ τὴν στιγμή ποὺ τὸ μητρώο Ψ^{-1} εἶναι σταθερό, ὡς γραμμικὸς μετασχηματισμὸς, ἀρκεῖ νὰ βρεθεῖ μόνο μία φορά και δὲν χρειάζεται ἀνανυπολογισμὸς του κάθε φορά ποὺ ἀλλάζουν οἱ ἀρχικοὶ παράμετροι τοῦ μοντέλου.

1.5 Σύνοψη μεθοδολογίας

Άρχικά πρέπει να σχηματιστεί το σύνολο δεδομένων χρησιμοποιώντας τον αλγόριθμο 1.

Algorithm 1: Full Order Model

Data: FOM : $\mathbb{R}^{\nu} \rightarrow \mathbb{R}^d$

Result: Form dataset \mathbf{U} from solutions $\mathbf{u}_t \in \mathbb{R}^d$

```

1 Initialize empty matrix  $\mathbf{U}$ ;
2 COMPUTE  $\mathbf{M}, \mathbf{C}, \mathbf{K}$  matrices of FOM;
3 for  $i \leftarrow 1$  to  $N_{train}$  do
4   for  $t \leftarrow 0$  to  $N_{timesteps}$  do
5     SOLVE equilibrium equation with FOM;
6     GET solution  $\mathbf{u}_t$ ;
7     STORE  $\mathbf{u}_t$  in memory;
8   end
9   SAMPLE a subset of the solutions  $\mathbf{u}_t$  and store it in  $\mathbf{U}$ ;
10 end
```

Έπειτα, όπως φαίνεται στον αλγόριθμο 2, εκπαιδεύεται η μέθοδος στο σύνολο δεδομένων των λύσεων και οι υπόλοιπες λύσεις δίνονται από το μοντέλο μειωμένης τάξης.

Algorithm 2: Reduced Order Model

Data: Solutions dataset $\mathbf{U} \in \mathbb{R}^{d \times N}$

Result: Compute solutions $\mathbf{u}_t^{new} \in \mathbb{R}^d$ with ROM

```

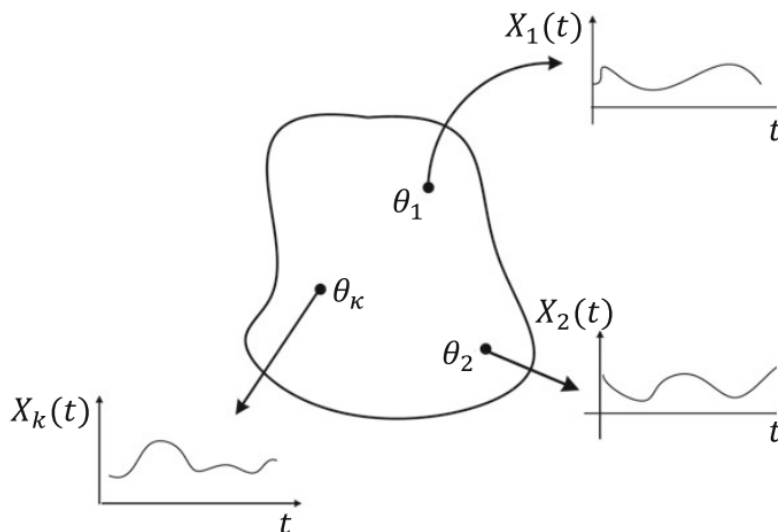
1 SELECT Dimensionality Reduction Method  $R$ ;
2 TRAIN the method  $R$  using solutions dataset  $\mathbf{U}$ ;
3 COMPUTE  $\Psi^{-1}$  matrix from  $R$ ;
4 COMPUTE reduced matrices  $\mathbf{M}_r, \mathbf{C}_r, \mathbf{K}_r$  of ROM using  $\Psi^{-1}$ ;
5 for  $i \leftarrow 1$  to  $N_{test}$  do
6   for  $t \leftarrow 0$  to  $N_{timesteps}$  do
7     SOLVE equilibrium equation with ROM;
8     GET solution  $\mathbf{u}_t^{new}$ ;
9   end
10 end
```

Κεφάλαιο 2

Στοιχεία από την Θεωρία Στοχαστικῶν Ἀνελιξέων

2.1 Ὅρισμός

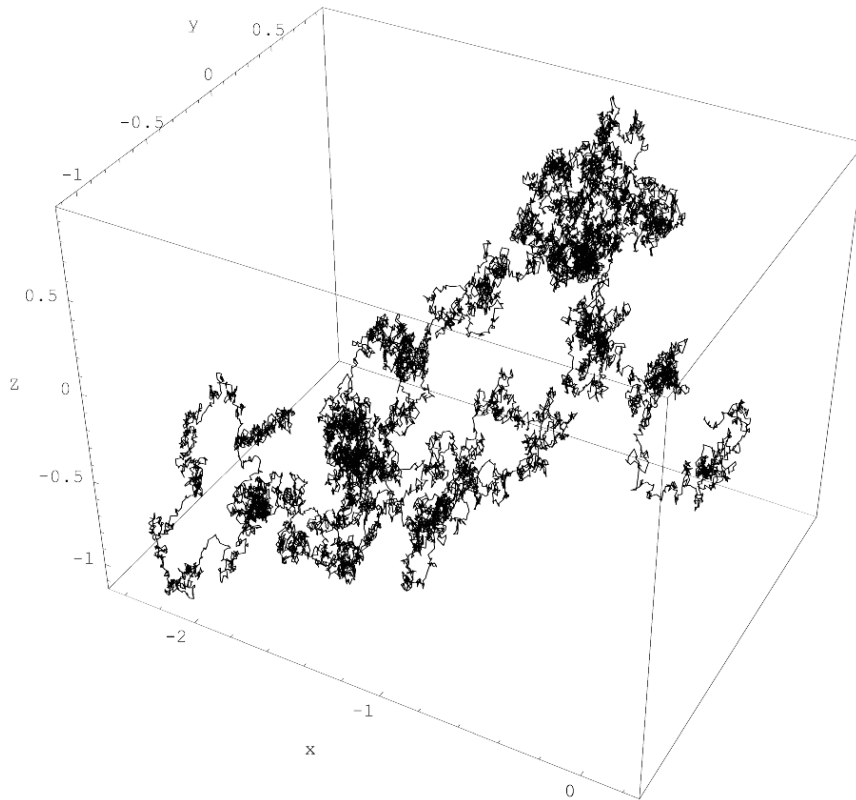
Οἱ *στοχαστικές διαδικασίες* ἢ *στοχαστικές ἀνελιξίεις* γενικεύουν τὸν ὅρισμό τῶν τυχαιῶν μεταβλητῶν. μιὰ στοχαστική διαδικασία ὀρίζεται ὡς ἓνα σύνολο τυχαιῶν μεταβλητῶν ἢ διανυσμάτων ποὺ βρίσκονται σὲ 1-1 ἀντιστοιχία μὲ ἓνα ἄλλο μαθηματικό σύνολο, ἔστω \mathbb{I} . Τὸ σύνολο \mathbb{I} ὀνομάζεται σύνολο δεικτῶν καὶ ἱστορικά ξεκίνησε σὰν ἓνα ὑποσύνολο τῶν πραγματικῶν ἀριθμῶν – π.χ. φυσικοὶ ἀριθμοὶ – γιὰ νὰ ἀποδοθεῖ ἡ ἔννοια τοῦ χρόνου. Τὸ σύνολο τιμῶν τῶν τυχαιῶν μεταβλητῶν \mathcal{S} ὀνομάζεται *χῶρος κατάστασης*. Ὅπως εἶναι κατανοητό, ἐν γένει, δὲν ὑπάρχει κανένας περιορισμὸς οὔτε στὸ σύνολο δεικτῶν, οὔτε στὸ χῶρο κατάστασης τῶν τυχαιῶν μεταβλητῶν. Ἔτσι, μιὰ στοχαστική διαδικασία μπορεῖ νὰ παράγει πολλές συναρτήσεις λόγω τῆς τυχαιότητάς. Κάθε συνάρτηση ποὺ παράγεται ὀνομάζεται *δείγμα* ἢ *πραγματοποίηση* τῆς στοχαστικῆς διαδικασίας.



Σχῆμα 2.1: Μία στοχαστική διαδικασία ἀποτελεῖ ἀπεικόνιση ἀπὸ τὸν χῶρο τῶν τυχαιῶν γεγονότων στὸν χῶρο τῶν συναρτήσεων.

Πιο συγκεκριμένα, μιὰ στοχαστική διαδικασία $X(t, \theta)$ μπορεῖ νὰ ὀριστεῖ ὡς ἡ ἀπεικόνιση ἀπὸ τὸν χῶρο πιθανοτήτων $\{\Theta, \mathcal{F}, P\}$ σὲ μιὰ συνάρτηση τοῦ χρόνου $X_i(t)$ (*στοχαστική διαδικασία*) ἢ καὶ τοῦ χῶρου (*στοχαστικό πεδίο*), ὅπου Θ ὁ δειγματικός χῶρος, \mathcal{F} μιὰ σ -ἄλγεβρα ἐπὶ τοῦ Θ καὶ P τὸ μέτρο πιθανότητας. Σὲ μιὰ γενική περίπτωση ἓνα στοχαστικό

πεδίο $\mathbf{X}(\mathbf{t}, \theta)$ μπορεί να υλοποιείται από πολυμεταβλητές συναρτήσεις $\mathbf{X}_i(\mathbf{t}, \theta) \in \mathbb{R}^n$ με $\mathbf{t} \in \mathbb{R}^m$, όπως άλλωστε ισχύει για όλες τις συναρτήσεις. Για συντομία μπορεί να γραφτεί ότι το προηγούμενο πεδίο μπορεί να κατηγοριοποιηθεί ως mD-nV (m-Dimensional, n-Variate). Σύμφωνα, λοιπόν, με τα παραπάνω μία 1D-1V διαδικασία $X(t)$ ερμηνεύεται ως ακολουθία από κοινού κατανεμημένων τυχαίων μεταβλητών $X(t_i)$ με $i \in \mathbb{I}$ για τις οποίες πρέπει να όριστεί από κοινού κατανομή.



Σχήμα 2.2: Πραγματοποίηση μίας στοχαστικής διαδικασίας Wiener, γνωστής και ως κίνησης Brown. Το σύνολο δεικτών της διαδικασίας είναι οι μη αρνητικοί πραγματικοί αριθμοί, ενώ ο χώρος κατάστασης είναι ο τριδιάστατος εὐκλείδειος χώρος.

2.2 Συναρτήσεις μάζας και πυκνότητας

Ἄν τὸ \mathbb{I} εἶναι μετρήσιμο, πεπερασμένο σύνολο ἢ ἀπὸ κοινῆ συνάρτηση κατανομῆς μάζας ὀρίζεται ὡς

$$F_{X(t_1), \dots, X(t_n)}(x_1, \dots, x_n) = P[X(t_1) \leq x_1 \wedge \dots \wedge X(t_n) \leq x_n] \quad (2.1)$$

καὶ ἡ ἀντίστοιχη ἀπὸ κοινῆ κατανομῆ πυκνότητας ὡς

$$f_{X(t_1), \dots, X(t_n)}(x_1, \dots, x_n) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} F_{X(t_1), \dots, X(t_n)}(x_1, \dots, x_n) \quad (2.2)$$

section Ροπές Στοχαστικῶν Διαδικασιῶν

Έστω $f_X(x)$ ή σ.π.π. της τυχαίας μεταβλητής $X(t)$ κατά την χρονική στιγμή t . Η n -στή ροπή ορίζεται ως

$$m_n(t) = \mathbb{E}[X^n(t)] = \int_{-\infty}^{+\infty} x^n f_X(x) dx \quad (2.3)$$

Για $n = 1$ λαμβάνεται η μέση τιμή της διαδικασίας $m_1(t) = \mathbb{E}[X(t)] = \mu_X(t)$.

Οι κεντρικές ροπές βρίσκονται αντίστοιχα από τον τύπο

$$\mathbb{E}[(X(t) - \mu_X(t))^n] = \int_{-\infty}^{+\infty} (x - \mu_X(t))^n f_X(x) dx \quad (2.4)$$

Για $n = 2$ λαμβάνεται η διακύμανση της διαδικασίας $\mathbb{E}[(X(t) - \mu_X(t))^2] = \text{Var}[X(t)] = s_X^2(t)$.

2.3 Συναρτήσεις αυτοσυσχέτισης και αυτοδιακύμανσης

Η συνάρτηση αυτοσυσχέτισης $R_X(t_i, t_j)$ μιᾶς τυχαίας 1D-1V διαδικασίας ποσοτικοποιεί την σχέση μεταξύ τῶν τιμῶν τῆς $X(t)$ σὲ δύο χρονικὲς στιγμὲς καὶ . Σὲ ὀρισμένες περιπτώσεις μπορεῖ νὰ ὀριστεῖ ἀπὸ τὴν διαφορὰ μεταξύ τῶν δύο χρονικῶν στιγμῶν $\tau = t_i - t_j$ ἢ τὴν ἀπόσταση, ἂν γίνεταί λόγος γιὰ στοχαστικὸ πεδίο. Ἡ συνάρτηση αυτοσυσχέτισης μπορεῖ νὰ ὀριστεῖ ὡς

$$R_X(t_i, t_j) = \mathbb{E}[X(t_i)X(t_j)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_i x_j f_{X(t_i), X(t_j)}(x_i, x_j) dx_i dx_j, \quad (2.5)$$

ὅπου $X(t_i), X(t_j)$ οἱ τιμές τῆς στοχαστικῆς διαδικασίας κατὰ τὶς στιγμὲς t_i καὶ t_j μὲ ἀπὸ κοινῶ συνάρτηση τὴν $f_{X(t_i), X(t_j)}(x_i, x_j)$. Στὴν οὐσία ἡ συνάρτηση αυτοσυσχέτισης προκύπτει ἀπὸ τὸν ἄνα δύο συνδυασμὸ κάθε τυχαίας μεταβλητῆς τῆς διαδικασίας μὲ τὶς ὑπόλοιπες. Ἡ συνάρτηση αυτοσυσχέτισης ικανοποιεῖ τὶς παρακάτω ιδιότητες:

- Συμμετρία: $R_X(t_i, t_j) = R_X(t_j, t_i)$
- Ἀνισότητα Cauchy-Schwarz: $R_X(t_i, t_j)^2 \leq R_X(t_i, t_i)R_X(t_j, t_j)$
- Μὴ ἀρνητικὰ ὀρισμένη: $\sum_{-\infty}^{+\infty} \sum_{-\infty}^{+\infty} R_X(t_i - t_j)g(t_i)g(t_j) \geq 0 \forall$ συνάρτηση g γιὰ τὴν ὁποία τὸ ἄθροισμα συγκλίνει,

Γιὰ $t = t_i = t_j$ προκύπτει

$$R_X(t, t) = \mathbb{E}[X^2(t)] \quad (2.6)$$

ποὺ εἶναι τὸ μέσο τετραγωνικὸ σφάλμα τὴν χρονικὴ στιγμή t .

Μὲ ὅμοιο τρόπο ὀρίζεται ἡ συνάρτηση αυτοδιακύμανσης

$$C_X(t_i, t_j) = \mathbb{E}[(X(t_i) - \mu_X(t_i))(X(t_j) - \mu_X(t_j))] = R_X(t_i, t_j) - \mu_X(t_i)\mu_X(t_j). \quad (2.7)$$

Ἡ τιμὴ $C_X(t, t)$ στὴν διαγώνιο ἀποτελεῖ τὴν διακύμανση τῆς στοχαστικῆς διαδικασίας

$$C_X(t, t) = \mathbb{E}[(X(t) - \mu_X(t))^2] = \text{Var}[X(t)]. \quad (2.8)$$

Γιὰ μιὰ διαδικασία με μηδενικὴ μέση τιμὴ $\mu_X(t_i) = \mu_X(t_j) = 0$ ἡ συσχέτιση εἶναι ἴση μὲ τὴν συνδιακύμανση

$$C_X(t_i, t_j) = R_X(t_i, t_j) \quad (2.9)$$

ή, διαφορετικά, κάτι αρκετά βολικό στις πρακτικές εφαρμογές, από κάθε συνάρτηση συσχέτισης μπορεί να προκύψει συνάρτηση συνδιακύμανσης, αν αφαιρεθεί η μέση τιμή της διαδικασίας.

Σύμφωνα με όσα ορίστηκαν πριν, αν κανονικοποιηθεί η συνάρτηση συνδιακύμανσης, όπως και στην κλασική θεωρία πιθανοτήτων, προκύπτει ο συντελεστής αυτοσυσχέτισης

$$\rho_X(t_i, t_j) = \frac{C_X(t_i, t_j)}{\sqrt{C_X(t_i, t_i)C_X(t_j, t_j)}} \quad (2.10)$$

2.4 Γνωστές στοχαστιές διαδικασίες

2.4.1 Διαδικασίες Gauss

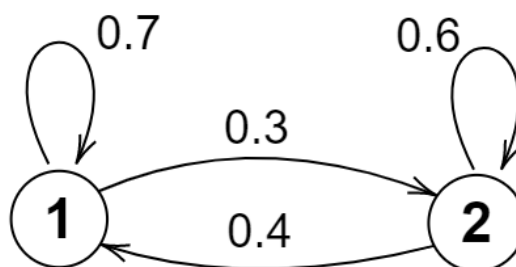
Όπως προαναφέρθηκε μια στοχαστική διαδικασία μπορεί να θεωρηθεί ως ακολουθία τυχαίων μεταβλητών. Αν οι τυχαίες αυτές μεταβλητές ακολουθούν την κατανομή Gauss, τότε αντίστοιχα και η διαδικασία ονομάζεται γκαουσιανή. Όταν ύποστούν γραμμικούς μετασχηματισμούς, οι διαδικασίες αυτές παραμένουν γκαουσιανές.

2.4.2 Διαδικασίες Markov

Σε μια διαδικασία Markov ή μελλοντική συμπεριφορά της μεταβλητής εξαρτάται αποκλειστικώς από το παρόν και όχι από το παρελθόν της μεταβλητής. Για τον λόγο αυτό οι διαδικασίες αυτές ανήκουν στην κατηγορία των άμνημων – χωρίς μνήμη – διαδικασιών. Αν για παράδειγμα οι διακριτές μεταβλητές X_i αντιστοιχούν στους χρόνους t_i , τότε έχουμε την εξής ιδιότητα

$$P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}, \dots, X_{t_1} = x_1) = P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}) \quad (2.11)$$

Άλυσίδες Markov



Σχήμα 2.3: Γράφημα ή διάγραμμα κατάστασης μαρκοβιανής αλυσίδας

Αν, λοιπόν, ο χρόνος παίρνει διακριτές τιμές, η διαδικασία ονομάζεται αλυσίδα Markov και χαρακτηρίζεται από τις πιθανότητες μετάβασης P_{ij} μεταξύ των καταστάσεων, με P_{ij} να εκφράζει την πιθανότητα μετάβασης από την κατάσταση i στην κατάσταση j . Συχνά συναντώνται συστήματα που οι μεταβατικές πιθανότητες είναι ανεξάρτητες του χρόνου,

δηλαδή έχουμε στάσιμη κατανομή

$$P(X_n = j \mid X_{n-1} = i) = P_{ij} \quad (2.12)$$

Ο πίνακας με στοιχεία P_{ij} , εκφράζει την πιθανότητα μετάβασης από την κορυφή i του διαγράμματος κατάστασης στην κορυφή j σε ένα χρονικό βήμα.

$$\mathbf{P} = \begin{pmatrix} P_{11} & \dots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \dots & P_{nn} \end{pmatrix}$$

και έχει την προφανή ιδιότητα

$$\forall i = 1, \dots, n, \sum_{j=1}^n P_{ij} = 1. \quad (2.13)$$

Συχνά οι αλυσίδες Markov είναι χρήσιμο να αναπαρασταθούν με την μορφή περιπάτου πάνω σε γράφημα. Το γράφημα αυτό ονομάζεται διάγραμμα κατάστασης της αλυσίδας. Οι κορυφές του αντιστοιχούν σε μιὰ κατάσταση i , ενώ τὰ βάρη στις κατευθυνόμενες ακμές αναπαριστούν την πιθανότητα P_{ij} . Επί παραδείγματι, στον πίνακα μετάβασης

$$\mathbf{P} = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \begin{pmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{pmatrix}.$$

αντιστοιχεί το γράφημα της εικόνας 2.3.

Το στοιχείο P_{12} αφορά στο ένδεχομένο μετάβασης από τον κόμβο 1 στον κόμβο 2, ενώ το στοιχείο P_{11} στο ένδεχομένο παραμονής στον κόμβο 1. Όμοίως και για τὰ στοιχεία της 2ης γραμμής. Όταν το \mathbf{P} τετραγωνίζεται, γίνεται

$$\begin{aligned} \mathbf{P}^2 &= \begin{pmatrix} P_{11}^2 & P_{12}^2 \\ P_{21}^2 & P_{22}^2 \end{pmatrix} \\ &= \begin{pmatrix} P_{11}P_{11} + P_{12}P_{21} & P_{11}P_{12} + P_{12}P_{22} \\ P_{21}P_{11} + P_{22}P_{21} & P_{21}P_{12} + P_{22}P_{22} \end{pmatrix} \\ &= \begin{pmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{pmatrix} \end{aligned}$$

Κάθε στοιχείο του \mathbf{P}^2 άθροίζει τις πιθανότητες δύο ένδεχομένων, π.χ. το $\mathbf{P}_{11}^2 = P_{11}P_{11} + P_{12}P_{21}$ άθροίζει την πιθανότητα της παραμονής στον κόμβο 1 και την πιθανότητα μετάβασης στον κόμβο 2 και ξανά πίσω στον κόμβο 1. Έν γενεί, κάθε στοιχείο P_{ij}^t άθροίζει τις πιθανότητες όλων τών μονοπατιών με μήκος t από τον κόμβο i στον κόμβο j .

Δεδομένης, λοιπόν, μιās κατανομής π_t πάνω στις κορυφές του αντίστοιχου γραφήματος, ή κατανομή στην επόμενη χρονική στιγμή δίνεται από την εξίσωση $\pi_{t+1} = \pi_t \mathbf{P}$. Εύκολα γίνεται αντιληπτό λόγω της άναδρομικότητας της σχέσης, ότι $\pi_{t+1} = \pi_{t-1} \mathbf{P}^2 = \pi_0 \mathbf{P}^{t+1}$, όπου π_0 ή αρχική κατανομή την χρονική στιγμή $t = 0$.

Σύμφωνα με το θεώρημα του Markov για μιὰ πεπερασμένη, άπεριοδική, μη ύποβιβάσιμη αλυσίδα Markov με καταστάσεις $0, 1, \dots, n$ ύπάρχουν όριακές πιθανότητες π_i , οι όποιες

είναι λύσεις του συστήματος εξισώσεων

$$\pi_j = \sum_{i=0}^n \pi_i P_{ij} \quad (2.14)$$

υπό την συνθήκη

$$\sum_{i=0}^n \pi_i = 1. \quad (2.15)$$

Η εξίσωση 2.14 αποτελεί πρόβλημα ιδιοτιμών με ιδιοτιμή $\lambda = 1$.

2.5 Αναπαράσταση στοχαστικής διαδικασίας

Παρά τα όσα ειπώθηκαν μέχρι στιγμής, η αριθμητική ανάλυση που χρησιμοποιείται σε προβλήματα μηχανικού απαιτεί απτά αριθμητικά δεδομένα. Αυτό σημαίνει ότι για να μπορούν να λυθούν τα προβλήματα αυτά θα πρέπει τα στοχαστικά πεδία να μπορούν να αναπαρασταθούν στην γλώσσα της αριθμητικής ανάλυσης. Έτσι, μιὰ συνεχής διαδικασία $\mathbf{X}(t, \theta)$ χρειάζεται να παρασταθεί από διακριτές τιμές \mathbf{X}_i σε κάποια σημεία $i = 1, \dots, N$. Το έρώτημα, λοιπόν, που προκύπτει είναι: πώς μπορεί να καθορίσει κανείς το βέλτιστο προσεγγιστικό πεδίο $\hat{\mathbf{X}}(\cdot)$ που θα περιγράφει το αρχικό πεδίο $\mathbf{X}(\cdot)$ με τον ελάχιστο αριθμό τυχαίων μεταβλητών $\{\mathbf{X}_i\}$; Ζητείται δηλαδή

$$\mathbf{X}(t, \theta) \approx \hat{\mathbf{X}}(t, \theta) = \{\mathbf{X}_i\} \quad (2.16)$$

Χωρίς βλάβη τις γενικότητας, για ένα 1D-1V πεδίο θα παρουσιαστούν συνοπτικά οι κατηγορίες των μεθόδων διακριτοποίησης:

1. Μέθοδοι **διακριτοποίησης σημείων** που οι τυχαίες μεταβλητές X_i είναι τιμές της $X(t, \theta)$ σε κάποια δεδομένα σημεία t_i .

$$\hat{X}(t, \theta) = \{X_i\} = \{X(t_i, \theta)\} \quad (2.17)$$

2. Μέθοδοι **διακριτοποίησης μέσου** που οι τυχαίες μεταβλητές X_i είναι σταθμισμένα ολοκλήρωματα της $X(t, \theta)$ στο χωρίο Ω_i .

$$\hat{X}(t, \theta) = \{X_i\} = \left\{ \int_{\Omega_i} X(t, \theta) c(t) d\Omega_i, t \in \Omega_i \right\} \quad (2.18)$$

3. Μέθοδοι **παρεμβολής** κατά τις οποίες η στοχαστική διαδικασία παρεμβάλεται σε κάποια σημεία. Συνήθως συνδυάζεται με κάποια μέθοδο διακριτοποίησης με σημεία.
4. Μέθοδοι **επέκτασης σειρών** στις οποίες η στοχαστική διαδικασία εκφράζεται προσεγγιστικά ως άθροισμα πεπερασμένων όρων άπειροσειρών. Από τις πλέον διαδεδομένες μεθόδους είναι η μέθοδος Karhunen-Loève (KL) και η μέθοδος της φασματικής απεικόνισης (SR). Και οι δύο ανήκουν στις φασματικές μεθόδους στις οποίες η $\hat{X}(t, \theta)$ εκφράζεται ως άπειροσειρά

$$\hat{X}(t, \theta) = \sum_{j=1}^{\infty} g_j(t) \xi_j(\theta) \quad (2.19)$$

όπου $\{\xi_j(\theta)\}$ τυχαίες μεταβλητές που χρησιμοποιούνται ως συντελεστές της σειράς και $\{g_j(t)\}$ αίτιοκρατικές συναρτήσεις στις οποίες προβάλλεται το πεδίο.

2.6 Η μέθοδος Karhunen-Loève

Η επέκταση Karhunen-Loève¹ (KL) μιὰς μηδενικού μέσου διαδικασίας $X(t, \theta)$ βασίζεται στην φασματική ανάλυση της συνάρτησης συνδιακύμανσης που ορίζεται ως

$$C_X(t_i, t_j) = \sigma_X(t_i)\sigma_X(t_j)\rho(t_i, t_j) \quad (2.20)$$

όπου ρ ο συντελεστής συσχέτισης. Έξ' όρισμοῦ ἡ $C_X(t_i, t_j)$ εἶναι φραγμένη, συμμετρική και ἔχει τὴν ἐξῆς φασματική ἀναπαράσταση

$$C_X(t_i, t_j) = \sum_{n=1}^{\infty} \lambda_n \phi_n(t_i) \phi_n(t_j) \quad (2.21)$$

Οἱ ϕ_n εἶναι οἱ ἰδιοσυναρτήσεις τῆς συνάρτησης συνδιακύμανσης καὶ λ_n οἱ ἀντίστοιχες ἰδιοτιμῆς. Ἐπιβάλλεται ὀρθογωνιότητα στὴν βάση, δηλαδή

$$\int_D \phi_n(t) \phi_m(t) dt = \delta_{nm} \quad (2.22)$$

όπου δ_{nm} τὸ δέλτα τοῦ Kronecker.

Ἡ οὐσία τῆς μεθόδου εἶναι ἡ εὑρεση τῶν ἰδιοσυναρτήσεων καὶ τῶν ἰδιοτιμῶν τῆς ἐξίσωσης 2.23 καὶ βρίσκονται ἀπὸ τὴν ἐπίλυση τῆς ὁμογενοῦς ἐξίσωσης τοῦ ὀλοκληρώματος Fredholm δευτέρου εἴδους μὲ πυρήνα συνδιακύμανσης

$$\int_D C_X(t_i, t_j) \phi_n(t_j) dt_j = \lambda_n \phi_n(t_i) \quad (2.23)$$

όπου D τὸ πεδίο στὸ ὁποῖο ορίζεται ἡ διαδικασία. Ἐπειδὴ ἀναλυτικὴ λύση τῆς 2.23 προσφέρεται μόνον σὲ εἰδικές περιπτώσεις ἡ ἐπίλυσή της γίνεται γενικὰ ἀριθμητικά.

Ὅποιαδήποτε πραγματοποίηση τῆς $X(t, \theta)$ μπορεῖ νὰ προβληθεῖ στὴν βάση ὡς ἐξῆς

$$X(t, \theta) = \sum_{n=1}^{\infty} \sqrt{\lambda_n} \phi_n(t) \xi_n(\theta), \quad t \in D \quad (2.24)$$

όπου $\{\xi_n(\theta)\}$ εἶναι ἓνα σύνολο ἀπὸ ἀσυσχέτιστες τυχαῖες μεταβλητές μὲ μέσο $\mathbb{E}[\xi_n(\theta)] = 0$ καὶ συνάρτηση συνδιακύμανσης $\mathbb{E}[\xi_n(\theta) \xi_m(\theta)] = \delta_{nm}$. Κάθε τυχαία μεταβλητὴ μπορεῖ νὰ ἐκφρασθεῖ ὡς

$$\xi_n(\theta) = \frac{1}{\sqrt{\lambda_n}} \int_D X(t, \theta) \phi_n(t) dt. \quad (2.25)$$

Ἡ ἐξίσωση 2.24 εἶναι γνωστὸ ὅτι συγκλίνει ὡς πρὸς τὸ μέσο τετραγωνικὸ σφάλμα γιὰ ὁποιαδήποτε κατανομή. Ἡ KL σειρά μιὰς διαδικασίας Gauss ἔχει τὴν ιδιότητα ὅτι οἱ $\xi_n(\theta)$ εἶναι ἀνεξάρτητες κανονικὲς μεταβλητές.

¹Τὸ ὄνομά της τὸ ὀφείλει στοὺς μαθηματικούς Kari Karhunen (1907-1979) καὶ Michel Loève (1915-1992)

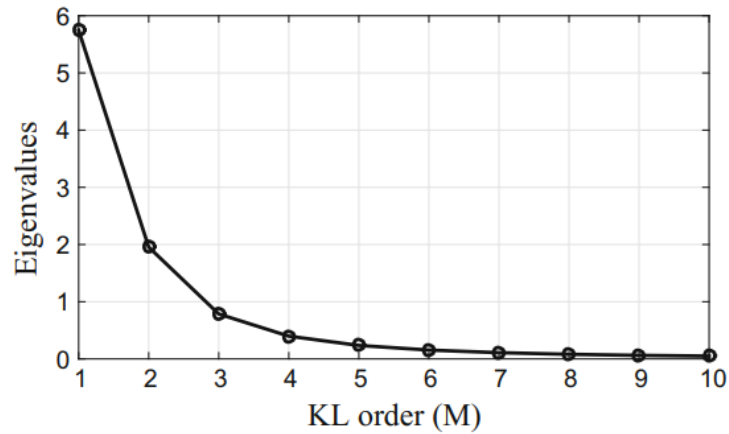
Για πρακτικές εφαρμογές ή σειρά προσεγγίζεται από πεπερασμένο αριθμό όρων M , δηλαδή

$$X(t, \theta) = \hat{X}(t, \theta) = \sum_{n=1}^M \sqrt{\lambda_n} \phi_n(t) \xi_n(\theta). \quad (2.26)$$

Η αντίστοιχη συνάρτηση συνδιακύμανσης προσεγγίζεται από την

$$\hat{C}_X(t_i, t_j) = \sum_{n=1}^M \lambda_n \phi_n(t_i) \phi_n(t_j). \quad (2.27)$$

Οι Ghanem και Spanos[2] απέδειξαν ότι αυτές οι πεπερασμένες σειρές είναι βέλτιστες ως προς το μέσο τετραγωνικό σφάλμα, καθώς οι ιδιοτιμές λ_n της 2.23 συγκλίνουν γρήγορα προς το μηδέν 2.4.



Σχήμα 2.4: Φθίνουσες ιδιοτιμές από την λύση του ολοκληρώματος Fredholm δευτέρου είδους για $M = 10$ [3].

Κεφάλαιο 3

Στοιχεία από την Θεωρία Γραφημάτων

3.1 Στόχος

Ο στόχος του κεφαλαίου αυτού, καθώς και του έπομένου είναι η συνοπτική παρουσίαση των μαθηματικῶν εργαλείων που θα χρησιμοποιηθοῦν στην εργασία. Δεδομένου ὅτι ὁ ἀναγνώστης ἔχει ἕνα ἰκανοποιητικό μαθηματικό υπόβαθρο καὶ προκειμένου ἡ εργασία νὰ παραμείνει ἀνεπαχθῆς χωρὶς νὰ ξεφύγει ἀπὸ τὸν σκοπὸ της, θὰ παρουσιαστοῦν μόνο οἱ ἔννοιες ἐκεῖνες ποὺ ἐν γένει δὲν διδάσκονται σὲ προπτυχιακά μαθήματα πολιτικῶν μηχανικῶν. Ἄν, λοιπόν, ὁ ἀναγνώστης ἀνακαλύψει ἄγνωστες μαθηματικές ὁρολογίες, αὐτὲς κατὰ πάσα πιθανότητα καλύπτονται στὰ ἀντίστοιχα μαθήματα προπτυχιακοῦ ἐπιπέδου στὰ ὁποῖα μπορεῖ καὶ νὰ ἀνατρέξει.

3.2 Ὅρισμός

Ἐστω μὴ κατευθυνόμενος γράφος $\mathcal{G} = (\mathbf{V}, \mathbf{E}, w)$, ὅπου:

- $\mathbf{V} := \{v_1, \dots, v_n\}$ τὸ πεπερασμένο σύνολο τῶν κορυφῶν τοῦ γράφου,
- $\mathbf{E} \subseteq \mathbf{V} \times \mathbf{V}$ τὸ σύνολο τῶν ἀκμῶν τοῦ γράφου,
- $\mathcal{V} := \{f : \mathbf{V} \rightarrow \mathbb{R}\}$ ὁ χῶρος Hilbert τῶν συναρτήσεων ὀριζομένων στὶς κορυφές,
- $\mathcal{E} := \{F : \mathbf{V} \times \mathbf{V} \rightarrow \mathbb{R} \mid (v_i, v_j) \notin \mathbf{E} \implies F(v_i, v_j) = 0\}$ ὁ χῶρος Hilbert τῶν συναρτήσεων ποὺ ὀρίζονται στὶς ἀκμές, ἐπεκτεταμένος στὸν $\mathbf{V} \times \mathbf{V}$,
- $w \in \mathcal{E}$ ἡ συνάρτηση βάρους τοῦ γράφου.

Σὲ ἕνα μὴ κατευθυνόμενο γράφο γιὰ κάθε ἀκμὴ (v_i, v_j) ὑπάρχει ἡ ἀκμὴ (v_j, v_i) , ἐνῶ ἡ συνάρτηση βάρους πρέπει νὰ εἶναι συμμετρικὴ $w(v_i, v_j) = w(v_j, v_i)$. Γιὰ τὰ συνήθη προβλήματα $w : \mathbf{V} \times \mathbf{V} \rightarrow \mathbb{R}_0^+$ ἢ ἀκόμα καὶ $w : \mathbf{V} \times \mathbf{V} \rightarrow [0, 1]$.

Ἡ κορυφή v_i ὀνομάζεται γείτονας τῆς v_j καὶ συμβολίζεται $v_i \sim v_j$ ἂν ὑπάρχει ἀκμὴ $(v_i, v_j) \in \mathbf{E}$. Γειτονιά $\mathcal{N}(v_i) := \{v_j \in \mathbf{V} \mid v_j \sim v_i\}$ τῆς κορυφῆς v_i εἶναι ἀπλὰ τὸ σύνολο τῶν γειτόνων της. Βαθμὸς μιᾶς κορυφῆς εἶναι τὸ σταθμικὸ ἄθροισμα τῶν γειτόνων της:

$$d_i = \deg(v_i) = \sum_{j: v_j \sim v_i} w(v_i, v_j)$$

3.2.1 Διαφορικοί τελεστές σε γράφους

Έστω γράφος $\mathcal{G} = (\mathbf{V}, \mathbf{E}, w)$. Κατ' αναλογία με τούς συνεχείς τελεστές της διανυσματικής ανάλυσης, ορίζονται οι έξης διακριτοί τελεστές γράφου:

- Ο τελεστής κλίσεως γράφου grad ή $\nabla : \mathbf{V} \rightarrow \mathbf{E}$ πού δρᾶ ὡς:

$$(\nabla f)(v_i, v_j) := \partial_{v_j} f(v_i) = f(v_j) - f(v_i) \quad (3.1)$$

ὅπου $f \in \mathbf{V}$ συνάρτηση ὀρισμένη στις κορυφές. Μερικές προφανείς ιδιότητες τῆς κλίσεως:

1. $\nabla f(v_i, v_j) = -\nabla f(v_j, v_i)$,
2. $\nabla f(v_i, v_i) = 0$,
3. $f(v_i) = f(v_j) \implies \nabla f(v_i, v_j) = 0$.

- Ο τελεστής ἀποκλίσεως γράφου div ή $\nabla \cdot : \mathbf{E} \rightarrow \mathbf{V}$ ὡς:

$$(\nabla \cdot F)(v_i) := \sum_{v_j \sim v_i} w(v_i, v_j) F(v_i, v_j) \quad (3.2)$$

ὅπου $F \in \mathbf{E}$ συνάρτηση ὀρισμένη στις ἄκμές.

- Ἡ Λαπλασιανή γράφου Δ ή $\nabla^2 := \nabla \cdot \nabla : \mathbf{V} \rightarrow \mathbf{V}$ πού ὀρίζεται ὅπως καὶ τὸ συνεχές ἀνάλογο τῆς ἀπὸ τὴν κλασικὴ ἀνάλυση:

$$\begin{aligned} (\nabla^2 f)(v_i) &:= (\nabla \cdot (\nabla f))(v_i) = \sum_{v_j \sim v_i} w(v_i, v_j) \nabla f(v_i, v_j) \\ &\implies (\Delta f)(v_i) = \sum_{v_j \sim v_i} w(v_i, v_j) (f(v_j) - f(v_i)). \end{aligned} \quad (3.3)$$

Οἱ διακριτοὶ τελεστές ἐξ' ὀρισμοῦ μιμοῦνται τὴν συμπεριφορὰ τῶν συνεχῶν τελεστῶν τῆς κλασικῆς διανυσματικῆς ἀνάλυσης. Ὅπως καὶ στὴν διανυσματικὴ ἀνάλυση, ἡ κλίση μετράει τὸ ρυθμὸ μεταβολῆς, ἡ ἀπόκλιση μετρά τὴν συνολικὴ σταθμισμένη ἐκροή ἀπὸ τὸν κόμβο v_i καὶ, τέλος, ἡ Λαπλασιανὴ ποσοτικοποιεῖ τὴν ἔννοια τῆς διαφορᾶς τῆς τιμῆς κορυφῆς $f(v_i)$ ἀπὸ τὶς ἀντίστοιχες γειτονικὲς τῆς. Ἐξακολουθεῖ π.χ. νὰ ἰσχύει, ὡς γνωστὸν, ὅτι $\Delta f_i > 0$, ἂν $f(v_i)$ εἶναι τοπικὸ ἐλάχιστο, ἔνω $\Delta f_i < 0$, ἂν $f(v_i)$ τοπικὸ μέγιστο.

Ἐκτὸς τῆς ὁμοιότητος τῶν ἴδιων τῶν τελεστῶν στις δύο θεωρίες, ἀξίζει νὰ παρατηρήσει κανεὶς τὴν ἀναλογία τῶν πεδίων ὀρισμοῦ τῶν συναρτήσεων. Ἡ κλίση γράφου ἔχει πεδίο ὀρισμοῦ τὸ σύνολο τῶν κορυφῶν \mathbf{V} , ἐνῶ ἡ συνεχῆς κλίση ὀρίζεται σὲ πολύπτυγμα \mathcal{M} . Τοῦτο παρακινεῖ στὸ νὰ νοεῖται τὸ \mathbf{V} ὡς τὸ πολύπτυγμα στὸ ὁποῖο ἐκτελεῖται ἡ πραγματικὴ ἀνάλυση. Κατ' ἀντιστοιχία ἡ ἀπόκλιση γράφου ἔχει πεδίο ὀρισμοῦ τὸ ἐπεκτεταμένον σύνολο τῶν ἄκμῶν $\mathbf{V} \times \mathbf{V}$, ἐνῶ ἡ συνεχῆς ἀπόκλιση ὀρίζεται πάνω σὲ διανυσματικὸ πεδίο. Ἐτσι, τὸ $\mathbf{V} \times \mathbf{V}$ δύναται νὰ λογιστεῖ ὡς διανυσματικὸ πεδίο. Γιὰ παράδειγμα, ἡ ἔννοια τῆς φορᾶς ὑπάρχει ἤδη ἀπὸ τὸν ὀρισμὸ τῆς ἄκμῆς $(i, j) \neq (j, i)$, ἐνῶ ἡ ιδιότητα $\nabla f(v_i, v_j) = -\nabla f(v_j, v_i)$ θυμίζει τὸ διανυσματικὸ πρόσημο $\vec{r}_{AB} = -\vec{r}_{BA}$.

3.2.2 Σύνδεση με τὴν Θεωρία Πινάκων

Ὅπως διαπιστώθηκε προηγουμένως, μπορεῖ κατ' ἀναλογία νὰ θεωρηθεῖ ὅτι ἡ $f \in \mathbf{V}$ πεδίο ἀνάλυσης καὶ ἡ $F \in \mathbf{E}$ διανυσματικὸ πεδίο. Μὲ τὴν ἴδια λογικὴ, καὶ γιὰ νὰ συνδεθεῖ ἡ

θεωρία γραφημάτων με την θεωρία πινάκων διαπιστώνεται το ἑξῆς: Ἐφόσον μιὰ συνάρτηση $f \in \mathcal{V}$ ὀρίζεται σὲ πεπερασμένο σύνολο κορυφῶν \mathbf{V} , τότε δύναται νὰ παρασταθεῖ ὡς ἓνα n -διάστατο διάνυσμα $\mathbf{f} \in \mathbb{R}^n$, ὅπου $n = |\mathbf{V}|$. Ἐξ' αὐτοῦ μπορεῖ κανεὶς νὰ ἐξάγει τοὺς καθοριστικῆς σημασίας ἰσομορφισμοὺς $\mathcal{V} \cong \mathbb{R}^n$ καὶ $\mathcal{E} \cong \mathbb{R}^{n \times n}$ ποὺ θὰ ἐπιτρέψουν τὴν μετάβαση.

Ἐχοντας, λοιπὸν, κάνει αὐτὴν τὴν διαπίστωση, θὰ γραφεῖ ἡ Λαπλασιανὴ μετὴν μορφή πίνακα, παρατηρώντας τὸν τρόπο που δρᾷ στὸ διάνυσμα $\mathbf{f} \in \mathbb{R}^n$. Ἡ ἐξίσωση 3.3, θέτοντας χάριν συντομίας $f_i = f(v_i)$ καὶ $w_{ij} = w(v_i, v_j)$ ¹ γράφεται:

$$\begin{aligned} (\Delta f)_i &= \sum_{j=1}^n w_{ij}(f_j - f_i) \\ &= \sum_{j=1}^n w_{ij}f_j - \sum_{j=1}^n w_{ij}f_i \\ \implies \Delta f_i &= \sum_{j=1}^n w_{ij}f_j - f_i \sum_{j=1}^n w_{ij}, \end{aligned} \quad (3.4)$$

Ἡ Λαπλασιανὴ γράφου μετὴν μητρικὴ μορφή ὀρίζεται ὡς ἡ διαφορά:

$$\mathbf{L} = \mathbf{D} - \mathbf{W} \quad (3.5)$$

ὅπου:

- \mathbf{f} τὸ διάνυσμα μετὴν τιμὲς τῆς $f(v_i)$ σὲ κάθε κόμβο τοῦ γράφου:

$$f_i = f(v_i) \quad (3.6)$$

- \mathbf{W} τὸ συμμετρικὸ μητρῶο βαρῶν μετὴν στοιχεία:

$$W_{ij} = w(v_i, v_j) \quad (3.7)$$

- \mathbf{D} τὸ διαγώνιο μητρῶο ποὺ περιέχει τὸν βαθμὸ κάθε κορυφῆς:

$$D_{ii} = d_i = \deg(v_i) = \sum_{j=1}^n W_{ij} \quad (3.8)$$

Σημειώνεται ὅτι μετὴν μητρικὴ γραφή ἡ 3.4 γράφεται:

$$\Delta \mathbf{f} = \mathbf{W}\mathbf{f} - \mathbf{D}\mathbf{f} = -(\mathbf{D} - \mathbf{W})\mathbf{f} = -\mathbf{L}\mathbf{f} \quad (3.9)$$

Συχνὰ ἐφαρμόζεται μιὰ μέθοδος γνωστὴ ὡς «κανονικοποίηση Λαπλασιανῆς γράφου» (*graph Laplacian normalization*) [4] ἡ ὅποια μπορεῖ νὰ γίνῃ μετὴν δύο τρόπους:

1. Συμμετρικῶς κανονικοποιημένη:

$$\mathbf{L}_n = \mathbf{D}^{-\frac{1}{2}} \mathbf{L} \mathbf{D}^{-\frac{1}{2}} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}} \quad (3.10)$$

¹Στὸ σημεῖο αὐτὸ φαίνεται ὁ λόγος γιὰ τὸν ὅποιο ἐπιλέχθηκε τὸ πεδίο ὀρισμοῦ τῶν $F \in \mathcal{E}$ νὰ εἶναι τὸ $\mathbf{V} \times \mathbf{V}$, ἀντὶ τοῦ $E \subseteq \mathbf{V} \times \mathbf{V}$, καθὼς δὲν θα ἦταν δυνατὸν μετὴν ἄλλο τρόπο νὰ δημιουργηθοῦν κανονικοὶ πίνακες.

2. Τυχαιού περιπάτου:

$$\mathbf{L}_s = \mathbf{D}^{-1}\mathbf{L} = \mathbf{I} - \mathbf{P} = \mathbf{I} - \mathbf{D}^{-1}\mathbf{W} \quad (3.11)$$

Όπου

$$\mathbf{P} = \mathbf{D}^{-1}\mathbf{W} \quad (3.12)$$

στοχαστικός πίνακας ή μητρώο μετάβασης το οποίο επιτρέπει τον όρισμό μαρκοβιανής άλυσίδας πάνω στους κόμβους του γράφου. Ο στοχαστικός πίνακας θα χρησιμοποιηθεί αργότερα από την μέθοδο των DMAPs.

3.2.3 Ιδιομορφική Ανάλυση τής Λαπλασιανής

Θα εξεταστεί ή σχέση τών φασμαστικῶν χαρακτηριστικῶν τής συμμετρικά κανονικοποιημένης Λαπλασιανής με αὐτῶν τοῦ στοχαστικοῦ πίνακα. Ἐστω τὸ ιδιοπρόβλημα

$$\mathbf{L}_n \mathbf{f} = \lambda_i \mathbf{f} \quad (3.13)$$

ὅπου λ_i ή i -στη ιδιοτιμή τοῦ Λαπλασιανοῦ πίνακα. Ἄν αντικατασταθεῖ ὁ πίνακας αὐτὸς, ἀπὸ τήν σχέση 3.10

$$\begin{aligned} \left(\mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}} \right) \mathbf{f} &= \lambda_i \mathbf{f} \\ \iff \mathbf{D}^{-\frac{1}{2}} \mathbf{f} - \mathbf{D}^{-1} \mathbf{W} \mathbf{D}^{-\frac{1}{2}} \mathbf{f} &= \lambda_i \mathbf{D}^{-\frac{1}{2}} \mathbf{f}. \end{aligned}$$

Χρησιμοποιώντας τήν 3.12 καί θέτοντας

$$\mathbf{g} = \mathbf{D}^{-\frac{1}{2}} \mathbf{f} \quad (3.14)$$

ή προηγούμενη σχέση γίνεται

$$\begin{aligned} \mathbf{g} - \mathbf{P} \mathbf{g} &= \lambda_i \mathbf{g}. \iff \mathbf{P} \mathbf{g} = (1 - \lambda_i) \mathbf{g}. \\ \iff \mathbf{P} \mathbf{g} &= \omega_i \mathbf{g} \end{aligned} \quad (3.15)$$

ή ὅποια ἐξίσωση δεν εἶναι παρὰ ἓνα ἀκόμη πρόβλημα ιδιοτιμῶν, αὐτήν τήν φορά τοῦ στοχαστικοῦ πίνακα με ιδιοτιμές

$$\omega_i = 1 - \lambda_i. \quad (3.16)$$

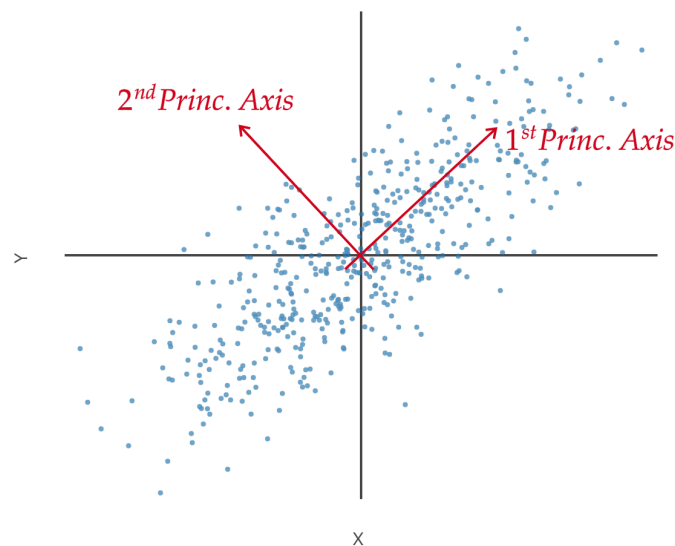
Εἶναι φανερό λοιπὸν ὅτι τὰ φασματικά χαρακτηριστικά τῶν \mathbf{P} καί \mathbf{L}_n συνδέονται ἄμεσα ἀπὸ τῆς ἐξισώσεις 3.14 καί 3.16. Ἐπιπλέον, συμπεραίνεται ὅτι οἱ ἐφόσον τὰ φασματικά χαρακτηριστικά συνδέονται με ἓνα πραγματικό, γραμμικό μετασχηματισμό καί ἐφόσον ὁ συμμετρικός \mathbf{L}_n ἔχει πραγματικές ιδιοτιμές, τὸ ἴδιο ἰσχύει καί γιὰ τὸν \mathbf{P} . Δεδομένου μάλιστα ὅτι οἱ ιδιοτιμές ἑνὸς στοχαστικοῦ πίνακα $\lambda_i \in [0, 1]$ δύναται νὰ παρατηρήσει κανεῖς ὅτι τὰ ω_i καί λ_i εἶναι συμπληρωματικά ὡς πρὸς τήν μονάδα $\omega_i + \lambda_i = 1$. Τὰ παραπάνω συμπεράσματα εἶναι ἐκμεταλλεύσιμα κατὰ τήν ἀριθμητική φασματική ἀνάλυση τῶν πινάκων, καθὼς ή εὔρεση τῶν φασματικῶν χαρακτηριστικῶν τοῦ μὴ συμμετρικοῦ στοχαστικοῦ μητρώου, μπορεῖ νὰ ἀναχθεῖ στὸ εὐκολότερο ὑπολογιστικά πρόβλημα με συμμετρικό πίνακα.

Κεφάλαιο 4

Ἡ μέθοδος τῆς Ἀνάλυσης σὲ Κύριες Συνιστώσες (PCA)

4.1 Επισκόπηση

Ἡ Ἀνάλυση Κυρίων Συνιστωσῶν (*Principal Component Analysis* ἢ ἐν συντομίᾳ *PCA*) εἶναι μία ἀπὸ τὶς πλέον γνωστὲς γραμμικὲς μεθόδους μείωσης διαστάσεων. Ἀνάλογα μὲ τὸ πλαίσιο στὸ ὁποῖο ἐφαρμόζεται, πολλὲς φορές ἀναφέρεται καὶ ὡς *Proper Orthogonal Decomposition (POD)*, ἐνῶ ἔχει ἄμεση σχέση μὲ ἄλλες μεθόδους ἰδιομορφικῆς ἀνάλυσης, ὅπως ἡ *SVD*¹. Χρησιμοποιεῖται ἐπιτυχῶς, παραδοσιακὰ στὴν Στατιστικὴ, ἐξ ἧς καὶ ἡ προέλευση, ἀλλὰ ἐσχάτως καὶ στὴν Ἐπιστῆμη Δεδομένων, στὴν Μηχανικὴ Μάθηση, στὴν ἐπεξεργασία σήματος, καθὼς καὶ σὲ πληθώρα ἄλλων ἐπιστημονικῶν κλάδων².



Σχῆμα 4.1: Τὸ σημεία ἐνὸς συνόλου δεδομένων καὶ τὰ δύο συστήματα συντεταγμένων, τὸ ἀρχικὸ σύστημα (μαῦρο) καὶ τὸ κύριο σύστημα (κόκκινο).

Ἡ PCA χρησιμοποιεῖ ἓναν ὀρθογώνιο μετασχηματισμὸ γιὰ νὰ ἀπεικονίσει ἓνα σύνολο συσχετισμένων δεδομένων σὲ νέες γραμμικὰ ἀσυσχέτιστες συντεταγμένες οἱ ὁποῖες ὀνομάζονται *κύριες συνιστώσες*. Ὁ ὀρθογώνιος μετασχηματισμὸς προκύπτει ἀπὸ τὴν φασματικὴ

¹Singular Value Decomposition

²Ἀξίζει νὰ σημειωθεῖ, ὅτι ἡ μέθοδος δημιουργήθηκε ἀπὸ τὸν K. Pearson κατ' ἀντιστοιχία μὲ τὸ θέωρημα τῶν Κυρίων Ἀξόνων τῆς μηχανικῆς καὶ τῆς μαθηματικῆς ἀνάλυσης, γι'αὐτὸ ἡ μεθοδολογία τῆς πιθανόν νὰ φανεῖ οἰκεία στὸν ἀναγνώστη.

άνάλυση του μητρώου εμπειρικής συσχέτισης, ενώ η διασπορά των δεδομένων κατά τους κύριους άξονες είναι φθίνουσα άκολουθία. Αυτό τελικώς επιτρέπει την διατήρηση μόνο εκείνων των πρώτων σημαντικων διαστάσεων που εμφανίζουν σχετικά μεγάλη διασπορά μειώνοντας έτσι τις διαστάσεις του αρχικού συνόλου.

4.2 Μεγιστοποίηση τής διασποράς

Έστω ένα σύνολο δεδομένων (dataset) $\mathbb{U} = \{\mathbf{u}_i \in \mathbb{R}^d, i = 1, 2, \dots, N | E(\mathbf{u}) = 0\}$, δηλαδή ένα σύνολο N ύλοποιήσεων ενός τυχαίου διανύσματος στον \mathbb{R}^d με μέση τιμή το μηδενικό διάνυσμα. Όπως αναφέρθηκε, αναζητείται αρχικά ένας όρθογώνιος μετασχηματισμός $\Psi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ό οποίος απεικονίζει τα δεδομένα σε ένα νέο σύστημα συντεταγμένων, τέτοιο ώστε ή μεγαλύτερη διασπορά να κείται κατά την πρώτη συντεταγμένη, τουτέστιν την πρώτη κύρια συνιστώσα, ή δεύτερη μεγαλύτερη διασπορά κατά την δεύτερη συνιστώσα κ.ο.κ.

Έστω το σύνολο d -διάστατων μοναδιαίων διανυσμάτων $\{\psi_i \in \mathbb{R}^d, i = 1, 2, \dots, n, n \leq d\}$ τα όποια θα άποτελέσουν την κύρια βάση απεικονίζοντας το σύνολο δεδομένων από το αρχικό στο κύριο σύστημα

$$P_{ij} = \psi_i \cdot \mathbf{u}_j \quad (4.1)$$

όπου P_{ij} είναι ή i κύρια συνιστώσα τής j παρατήρησης. Άς σημειωθεί, ότι το γινόμενο $\psi_i \cdot \mathbf{u}_j$ δέν είναι παρά ή κάθετη προβολή του \mathbf{u}_j στην διεύθυνση ψ_i .

Προκειμένου να μεγιστοποιείται ή διασπορά, το πρώτο διάνυσμα πρέπει ίκανοποιεί

$$\psi_1 = \operatorname{argmax}_{\|\psi\|=1} \left\{ \sum_j^N (P_{1j})^2 \right\} = \operatorname{argmax}_{\|\psi\|=1} \left\{ \sum_j^N (\psi \cdot \mathbf{u}_j)^2 \right\}$$

ή με μητρική γραφή

$$\psi_1 = \operatorname{argmax}_{\|\psi\|=1} \left\{ \|\mathbf{U}^T \psi\|^2 \right\} = \operatorname{argmax}_{\|\psi\|=1} \left\{ \psi^T \mathbf{U} \mathbf{U}^T \psi \right\}$$

Άφοϋ, όμως, $\|\psi\| = 1$, τελικώς προκύπτει ότι

$$\psi_1 = \operatorname{argmax}_{\|\psi\|=1} \left\{ \psi^T \mathbf{C} \psi \right\} = \operatorname{argmax}_{\|\psi\|=1} \left\{ \frac{\psi^T \mathbf{C} \psi}{\psi^T \psi} \right\}, \quad (4.2)$$

όπου

- $\mathbf{U} = (\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_N) \in \mathbb{R}^{d \times N}$ ό πίνακας που προκύπτει από το αρχικό σύνολο δεδομένων,
- $\mathbf{C} = \mathbf{U} \mathbf{U}^T \in \mathbb{R}^{d \times d}$. Άς σημειωθεί ότι $\mathbf{C} \propto \operatorname{Cov}(i, j) \frac{\mathbf{U} \mathbf{U}^T}{d-1}$, όπου $\frac{\mathbf{U} \mathbf{U}^T}{d-1}$ ό εμπειρικός πίνακας συνδιακύμανσης των δεδομένων με $C_{ij} = \sum_m U_{im} U_{jm}$

Δεδομένου του ψ_1 , ή πρώτη κύρια συνιστώσα του \mathbf{u}_j στο κύριο σύστημα είναι $P_{1j} = \mathbf{u}_j \cdot \psi_1$, ενώ στο αρχικό σύστημα συντεταγμένων είναι $(\mathbf{u}_j \cdot \psi_1) \psi_1$. Για κάθε επόμενο ψ_{k+1} άρκει να αφαιρέσουμε από τα δεδομένα τις προηγούμενες $k + 1$ συνιστώσες και να επαναλάβουμε την διαδικασία

$$\mathbf{u}_j^{k+1} = \mathbf{u}_j - \sum_{i=1}^k (\psi_i \cdot \mathbf{u}_j) \psi_i \quad (4.3)$$

ή μητρικιά

$$\mathbf{U}_{k+1} = \mathbf{U} - \sum_{i=1}^k \psi_k (\psi_k^T \mathbf{U}). \quad (4.4)$$

Το ψ_k βρίσκεται ξανά ως τὸ διάνυσμα ἐκεῖνο πὸν μεγιστοποιεῖ τὴν διασπορά

$$\psi_k = \operatorname{argmax}_{\|\psi\|=1} \{ \psi^T \mathbf{C}_k \psi \} = \operatorname{argmax}_{\|\psi\|=1} \left\{ \frac{\psi^T \mathbf{C}_k \psi}{\psi^T \psi} \right\}, \quad (4.5)$$

ὅπου $\mathbf{C}_k = \mathbf{U}_k \mathbf{U}_k^T$

4.3 Ίδιομορφική ανάλυση

Στὴν ἐξίσωση 4.5 ἐμφανίζεται ἡ ποσότητα

$$R(\mathbf{C}, \psi) = \frac{\psi^T \mathbf{C} \psi}{\psi^T \psi} \quad (4.6)$$

πὸν εἶναι γνωστή ὡς *πηλίκο Rayleigh* καὶ ἐξαρτάται μόνο ἀπὸ τὴν διεύθυνση τοῦ ψ καὶ ὄχι ἀπὸ τὸ μέτρο του: $R(\mathbf{C}, \psi) = R(\mathbf{C}, c\psi)$. Ὅπως εἶναι ἀποδεδειγμένο, γιὰ ἕνα θετικὰ ἡμιορισμένο πίνακα σὰν τὸν \mathbf{C} ἡ μεγαλύτερη τιμὴ πὸν μπορεῖ νὰ λάβει εἶναι ἴση μὲ τὴν μέγιστη ἰδιοτιμὴ τοῦ \mathbf{C} μὲ ἀντίστοιχο ἰδιοδιάνυσμα τὸ ψ . Ἀκόμη, ἂν θυμηθεῖ κανεὶς ὅτι τὰ ἰδιοδιανύσματα ἑνὸς συμμετρικοῦ πίνακα ὅπως ὁ \mathbf{C} συνιστοῦν μίαν ὀρθοκανονικὴν βάση, ὅπως δηλαδὴ καὶ τὰ τα ζητούμενα ψ_k , τότε ἀναζητᾶ τὴν πιθανὴ σχέση μετὰξὺ τους. Ὅπως, θὰ ἀποδειχθεῖ τὰ ψ_k ταυτίζονται μὲ τὰ ἰδιοδιανύσματα τοῦ ψ_k .

Πιὸ ἀναλυτικὰ, ἐξετάζεται ἡ ἐπίλυση τῆς 4.5 γιὰ ὅλες τὶς κύριες συνιστώσες. Ἀναζητοῦνται μὲ τὴν μέθοδο τῶν πολλαπλασιαστῶν Lagrange τὰ στάσιμα σημεῖα τῆς

$$R(\mathbf{C}, \psi) = \psi^T \mathbf{C} \psi$$

ὑπὸ τὸν περιορισμὸ $\|\psi\|^2 = \psi^T \psi = 1$, δηλαδὴ τὰ στάσιμα σημεῖα τῆς

$$\mathcal{L}(\psi) = \psi^T \mathbf{C} \psi - \lambda(\psi^T \psi - 1) \quad (4.7)$$

Ἡ παραγώγιση γίνεται ἀναλυτικὰ μὲ χρῆση γραφῆς Einstein ³

$$\begin{aligned} \mathcal{L} &= \psi_i C_{ij} \psi_j - \lambda(\psi_i \psi_i - 1) \\ &\implies \frac{\partial \mathcal{L}}{\partial \psi_k} = 0 \\ &\implies \frac{\partial}{\partial \psi_k} (\psi_i C_{ij} \psi_j) - \lambda \frac{\partial}{\partial \psi_k} (\psi_i \psi_i) = 0 \end{aligned}$$

³Ἡ σύμβαση Einstein εἶναι κατὰ τὸν γράφοντα ὁ πιὸ πρόσφορος τρόπος γραφῆς κατὰ τὴν χρῆση τῆς ἀνάλυσης σὲ διανυσματικούς χώρους. Στὴν πιὸ ἀπλή μορφή της καὶ ἂν δὲν δηλώνεται κάτι διαφορετικὸ ὑπονοεῖ τὴν ἄθροιση κάθε μεταβλητῆς κατὰ τὴν διεύθυνση τοῦ δείκτη πὸν ἐπαναλαμβάνεται στὸν ἴδιο ὄρο. Αὐτὸ συμβαίνει χωρὶς τὴν χρῆση τοῦ συμβόλου \sum , γιὰ εὐχέρεια τόσο γραφῆς ὅσο καὶ ἀνάγνωσης π.χ. $a \cdot b = a_i b_i = a_1 b_1 + a_2 b_2 + a_3 b_3$.

καί χρησιμοποιώντας τον κανόνα του γινομένου

$$\begin{aligned} \frac{\partial \psi_i}{\partial \psi_k} C_{ij} \psi_j + \psi_i C_{ij} \frac{\partial \psi_j}{\partial \psi_k} - \lambda \left(\frac{\partial \psi_i}{\partial \psi_k} \psi_i + \psi_i \frac{\partial \psi_i}{\partial \psi_k} \right) &= 0 \\ \implies \delta_{ik} C_{ij} \psi_j + \psi_i C_{ij} \delta_{jk} - \lambda (\delta_{ik} \psi_i + \psi_i \delta_{ik}) &= 0 \\ \implies C_{kj} \psi_j + \psi_i C_{ik} - \lambda (\psi_k + \psi_k) &= 0 \end{aligned}$$

Όπου δ_{ik} τὸ δέλτα τοῦ Kronecker. Ἐπειδὴ τὰ i, j εἶναι ψευδοδείκτες ἄθροισης ποὺ δὲν εἶναι καθοριστικοί, θέτοντας ὅπου $j = i$ λαμβάνεται

$$C_{ki} \psi_i + \psi_i C_{ik} = 2\lambda \psi_k$$

Λόγω συμμετρίας $C_{ki} = C_{ik}$, συνεπῶς

$$2C_{ki} \psi_i = 2\lambda \psi_k \implies C_{ki} \psi_i = \lambda \psi_k$$

ἢ ἐπιστρέφοντας στὴν γραφὴ μὲ πίνακες

$$\mathbf{C}\boldsymbol{\psi} = \lambda\boldsymbol{\psi}. \quad (4.8)$$

Ἡ ἐξίσωση 4.8 δὲν εἶναι τίποτα ἄλλο, παρά πρόβλημα ιδιοτιμῶν. Ἐτσι, συμπεραίνεται ὅτι τὰ ζητούμενα ψ_k εἶναι τὰ ιδιοδιανύσματα ($\psi_1, \psi_2, \dots, \psi_n$) τοῦ μητρώου \mathbf{C} καὶ ἀποτελοῦν τὰ στάσιμα σημεῖα τοῦ πηλίκου Rayleigh, ἐνῶ οἱ ἀντίστοιχες ιδιοτιμὲς ($\lambda_1, \lambda_2, \dots, \lambda_n$) τὶς στάσιμες τιμὲς τοῦ R

$$R(\mathbf{C}, \psi_k) = \frac{\psi_k^T \mathbf{C} \psi_k}{\psi_k^T \psi_k} = \lambda_k \frac{\psi_k^T \psi_k}{\psi_k^T \psi_k} = \lambda_k \quad (4.9)$$

Ἀκόμη, ἀπὸ τὴν 4.5 διαπιστώνεται ὅτι

$$\lambda_k = \sum_j^N (P_{kj})^2 = \sum_j^N (\psi_k \cdot \mathbf{u}_j)^2, \quad (4.10)$$

δηλαδή οἱ ιδιοτιμὲς ταυτίζονται μὲ τὴν —μεγιστοποιημένη— διακύμανση κατὰ τοὺς ἀντιστοίχους κυρίου ἀξονες.

4.4 Ὄρθη καὶ ἀντίστροφη ἀπεικόνιση – Μείωση Διαστάσεων

Μὲ βάση τὰ προηγούμενα, λοιπόν, ὁ πίνακας τῶν κυρίων συνιστωσῶν P_{ij} εἶναι

$$\mathbf{P} = \boldsymbol{\Psi}\mathbf{U} \in \mathbb{R}^{n \times d}, \quad (4.11)$$

ὅπου

$$\boldsymbol{\Psi} = \begin{pmatrix} \psi_1^T \\ \psi_2^T \\ \vdots \\ \psi_n^T \end{pmatrix} \quad (4.12)$$

τὸ μητρῶο με γραμμὲς ὅλα τὰ ιδιοδιανύσματα τῶν ὁποίων οἱ ἀντίστοιχες ιδιοτιμὲς ($\lambda_1, \lambda_2, \dots, \lambda_n$) εἶναι φθίνουσα ἀκολουθία — ἄρα τὰ ιδιοδιανύσματα λαμβάνονται κατὰ φθίνουσα σημαντικότητα. Αὐτὸ γίνεται, γιὰ νὰ μειωθεῖ στὴν συνέχεια ἡ διάσταση τῶν δεδομένων. Ἡ

μείωση τών διαστάσεων επιτυγχάνεται εύκολα αν ληφθούν υπόψιν μόνο κάποια πρώτα σημαντικά ιδιοδιανύσματα, δηλαδή στην περίπτωση που $n < d$.

Γενικότερα, έστω $\Psi : \mathbb{R}^d \rightarrow \mathbb{R}^n$ ή όρθή απεικόνιση, από τόν άρχικό χώρο στόν κύριο χώρο. Τότε, κάθε διάνυσμα εκφράζεται στο κύριο σύστημα με τήν εξίσωση

$$\psi = \Psi \mathbf{u}. \quad (4.13)$$

Άντίστοιχα, ή αντίστροφη απεικόνιση $\Psi^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^d$ βρίσκεται αν θέσουμε άρχικώς $n = d$. Λόγω όρθογωνιότητας, ισχύει

$$\Psi_d^T = \Psi_d^{-1}.$$

Τελικώς,

$$\mathbf{u} = \Psi_d^T \psi$$

καί κρατώντας $n < d$ ιδιοδιανύσματα

$$\mathbf{u} \approx \Psi^T \psi, \quad (4.14)$$

4.5 Σύνοψη τής μεθόδου

Algorithm 3: PCA algorithm

Data: Dataset matrix $\mathbf{U} \in \mathbb{R}^{d \times N}$

Result: COMPUTE the matrix Ψ^{-1}

- 1 Initialize;
 - 2 COMPUTE the covariance matrix $C = \mathbf{U}\mathbf{U}^T$;
 - 3 COMPUTE the n largest eigenvalues λ_i and corresponding eigenvectors ψ_i ;
 - 4 COMPUTE the inverse $\Psi^{-1} = (\psi_1, \dots, \psi_n)$
-

Κεφάλαιο 5

Ἡ μέθοδος τῶν Ἀπεικονίσεων Διάχυσης (Diffusion Maps)

5.1 Επισκόπηση

Οἱ Ἀπεικονίσεις Διάχυσης (*Diffusion Maps* ἢ ἐν συντομίᾳ *DMAPs*) ἀνήκουν στὶς μὴ γραμμικὲς μεθόδους μείωσης διαστάσεων, καθὼς προσπαθοῦν νὰ προσεγγίσουν τὴν τοπικὴ γεωμετρία τοῦ συνόλου δεδομένων. Εἶναι μὴ ἀρκετὰ πρόσφατη πού φέρει ἀρκετὲς προσδοκίες καὶ τῆς ὁποίας ἔχουν ἀρχίσει τελευταία νὰ διαφαίνονται ἡ χρησιμότητα καὶ τὰ πλεονεκτήματα.

Ἡ τοπικὴ συνδεσιμότητα τῶν δεδομένων ὑπολογίζεται ἀπὸ ἓνα μέτρο ἐγγύτητας μεταξύ κάθε ζεύγους δεδομένων. Μὲ βάση αὐτό, ὀρίζεται ἓνας μαρκοβιανὸς γράφος ἐπάνω στὰ δεδομένα πού συνδυάζεται μὲ μὴ μὴ διαδικασία διάχυσης πού ἐξελίσσεται στὸν χρόνο. Καθὼς ἐξελίσσεται ἡ διαδικασία, ἀποκαλύπτει γεωμετρικὲς δομὲς σὲ ὄλοένα καὶ μεγαλύτερες κλίμακες. Γιὰ κάθε τέτοια κλίμακα μπορεῖ νὰ ὀριστεῖ ἓνα μέτρο ἀπόστασης τὸ ὁποῖο βασίζεται στὴν ἀποκεκαλυμμένη γεωμετρία.

Ἡ ἀπεικόνιση (*diffusion map*) μετασχηματίζει τὰ δεδομένα ἀπὸ τὸν ἀρχικὸ χῶρο στὸν μικρότερον διαστάσεων χῶρο τῆς Διάχυσης. Οἱ διαστάσεις τοῦ χῶρου αὐτοῦ ἐξαρτῶνται ἀπὸ τὴν γεωμετρικὴ δομὴ τῶν δεδομένων καὶ ἀπὸ τὴν ἐπιθυμητὴ ἰσορροπία μεταξύ συμπύκνωσης τῆς πληροφορίας καὶ ἀκριβείας.

Ἐνα χαρακτηριστικὸ τῆς μεθόδου εἶναι ὅτι τὸ κόστος τῆς ἐξαρτάται ὡς ἐπὶ τῶν πλείστων ἀπὸ τὸν ἀριθμὸ τῶν δεδομένων καὶ ὄχι ἀπὸ τὴν διάστασή τους. Αὐτὸ ἀποτελεῖ ἐξαιρετικὸ πλεονέκτημα ὅταν οἱ διαστάσεις τοῦ ἀρχικοῦ χῶρου εἶναι μεγαλύτερες ἀπὸ τὸ πλῆθος τῶν δεδομένων καὶ τὴν καθιστᾷ ἀνταγωνιστικὴ ὡς πρὸς τὸ χρόνο ἐκπαίδευσης σὲ αὐτὴν τὴν περίπτωση.

5.2 Πυρήνας καὶ ἐγγύτητα δεδομένων

Ἐστω ἓνα σύνολο δεδομένων (*dataset*) $U = \{\mathbf{u}_i \in \mathbb{R}^d, i = 1, 2, \dots, N\}$ τὸ ὁποῖο μπορεῖ νὰ νοηθεῖ ὡς ἓνα σύνολο N ὑλοποιήσεων ἑνὸς τυχαίου διανύσματος στὸν \mathbb{R}^d . Ἐπίσης, ἔστω (U, A, μ) ἓνας μετρικὸς χῶρος ὅπου U τὸ προαναφερθὲν σύνολο, A ἡ σ -ἄλγεβρα καὶ μ ἡ κατανομή τῶν σημείων πάνω στὸν U . Δίδεται πυρήνας (*kernel*):

$$k : U \times U \rightarrow \mathbb{R}$$

μὲ τὶς ἐξῆς ιδιότητες:

- Συμμετρία: $k(\mathbf{x}, \mathbf{y}) = k(\mathbf{y}, \mathbf{x})$

- Μη αρνητικότητα: $k(\mathbf{x}, \mathbf{y}) \geq 0, \forall \mathbf{x}, \mathbf{y} \in \mathbb{U}$

Η επιλογή του εξαρτάται από το είδος της εφαρμογής, καθώς κάθε πυρήνας εξάγει διαφορετικά χαρακτηριστικά από τα δεδομένα. Σύνηθες παράδειγμα είναι ο Γκαουσιανός πυρήνας (*gaussian kernel*):

$$k(\mathbf{x}, \mathbf{y}) = e^{-\left(\frac{\|\mathbf{x}-\mathbf{y}\|}{\epsilon}\right)^2} \quad (5.1)$$

ή ο πυρήνας Epanechnikov:

$$k(\mathbf{x}, \mathbf{y}) = \begin{cases} 1 - \left(\frac{\|\mathbf{x}-\mathbf{y}\|}{\epsilon}\right)^2, & \text{αν } \|\mathbf{x} - \mathbf{y}\| \leq \epsilon \\ 0, & \text{αν } \|\mathbf{x} - \mathbf{y}\| > \epsilon \end{cases} \quad (5.2)$$

Ο πυρήνας προσδίδει μία έννοια εγγύτητας μεταξύ των δεδομένων, καθώς εκφράζει την σχέση μεταξύ δύο σημείων του \mathbb{U} . Στην ουσία ορίζει την *τοπική γεωμετρία* που θα αναλάβει να αποκαλύψει η μέθοδος και αποτελεί την κύρια διαφοροποίηση από μία γραμμική, μη τοπική μέθοδο όπως η PCA.

Στο σημείο αυτό είναι σημαντικό να τονιστεί η επίδραση της παραμέτρου ϵ ή οποία εξαρτάται από το πρόβλημα και πρέπει να επιλεγθεί καταλλήλως. Μία απλή μέθοδος που έχει προταθεί [5] για τον προσδιορισμό του ϵ χρησιμοποιεί την συνάρτηση:

$$M(\epsilon) = \sum_{i=1}^N \sum_{j=1}^N k(\mathbf{u}_i, \mathbf{u}_j)(\epsilon) \quad (5.3)$$

την οποία υπολογίζει και απεικονίζει σε διπλό λογαριθμικό διάγραμμα. Το ϵ επιλέγεται τυχαία από το εύρος που η γραφική παράσταση της $M(\epsilon)$ είναι γραμμική ή τουλάχιστον στο εύρος που η $M(\epsilon)$ δεν αποτελεί έκφυλισμένη περίπτωση.

5.3 Κατασκευή οικογένειας διαχύσεων

Άρχικά, ορίζεται το συμμετρικό μητρώο $\mathbf{W} \in \mathbb{R}^{N \times N}$ μέσω των στοιχείων του:

$$W_{ij} = k(\mathbf{u}_i, \mathbf{u}_j), \quad i, j \in \{1, \dots, N\} \quad (5.4)$$

το οποίο εν συνεχεία μετασχηματίζεται ως εξής:

$$\mathbf{W}^{(\alpha)} = \mathbf{D}^{-\alpha} \mathbf{W} \mathbf{D}^{-\alpha} \quad (5.5)$$

ή με χρήση δεικτών:

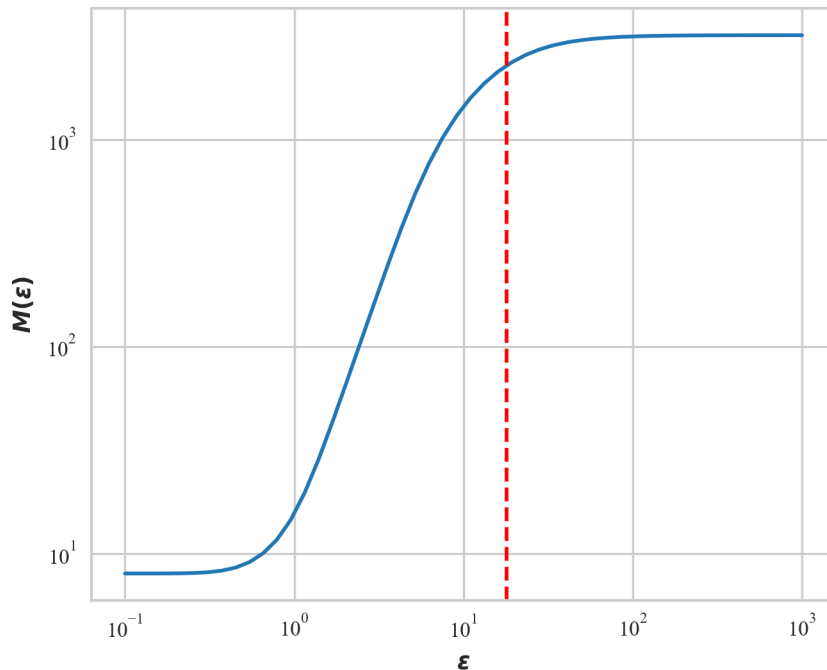
$$W_{ij}^{(\alpha)} = \frac{W_{ij}}{d_i^\alpha d_j^\alpha} \quad (5.6)$$

όπου \mathbf{D} το διαγώνιο μητρώο που προκύπτει από το άθροισμα των γραμμών του \mathbf{W} :

$$D_{ij} = \begin{cases} d_i, & i = j \\ 0, & i \neq j \end{cases} \quad (5.7)$$

και

$$d_i = \sum_{r=1}^N W_{ir} \quad (5.8)$$



Σχήμα 5.1: Η παράμετρος ϵ επιλέγεται μακριά από τα άκρα που αποτελούν έκφυλισμένες περιπτώσεις με υπερβολικά μεγάλα ή μικρά βάρη στο αντίστοιχο μητρώο.

Η παράμετρος α δημιουργεί μία οικογένεια από διαδικασίες διάχυσης. Είναι ουσιαστικά αντισταθμιστική και για το φάσμα της διάχυσης καθορίζει την έπιρροή της πυκνότητας των σημείων έναντι της γεωμετρίας του πολύπτυγματος στο οποίο αυτά ανήκουν. Όπως δημοσιεύτηκε στο [6], τρεις τιμές της παραμέτρου παρουσιάζουν ιδιαίτερο ενδιαφέρον:

- Όταν $\alpha = 0$, τότε η διάχυση συμπίπτει με την κλασική κανονικοποίηση της Λαπλασιανής του γράφου με ισότροπα βάρη π.χ. $e^{-(\|\mathbf{x}-\mathbf{y}\|/\epsilon)^2}$. Στην περίπτωση αυτή η έπιρροή της πυκνότητας των σημείων είναι μέγιστη.
- Στην ενδιάμεση περίπτωση όπου $\alpha = \frac{1}{2}$ ή μαρκοβιανή αλυσίδα που προκύπτει αποτελεί προσέγγιση της διάχυσης μίας εξίσωσης Focke-Planck έπιτρέπουσα τον προσδιορισμό της μακροχρόνιας συμπεριφοράς και της κατανομής των σημείων ενός συστήματος που περιγράφεται από συγκεκριμένη στοχαστική διαφορική εξίσωση.
- Όταν $\alpha = 1$ και υποθέτοντας ότι τα σημεία προσεγγιστικά κείνται σε κάποιο υποπολύπτυγμα του \mathbb{R}^d , τότε προκύπτει ο γνωστός *τελεστής Laplace-Beltrami*¹ Σε αυτήν την περίπτωση εξάγεται μόνο η Ριμμάνεια γεωμετρία του συνόλου δεδομένων που είναι ανεξάρτητη της κατανομής των σημείων.

¹Ο τελεστής αυτός αποτελεί γενίκευση του γνωστού τελεστού $\Delta = \nabla^2$ όριζόμενος και σε μη Ευκλείδειες Γεωμετρίες.

5.4 Κατασκευή τυχαίου περιπάτου στα δεδομένα

Είναι γνωστό ότι σε κάθε αντιστρέψιμη μαρκοβιανή αλυσίδα μπορεί κανείς να αντιστοιχίσει ένα συμμετρικό γράφο. Όμως, ισχύει και το αντίστροφο, δηλαδή δεδομένου συμμετρικού γράφου ορισμένου από τα (\mathbb{U}, k) μπορεί να κατασκευασθεί μία αντιστρέψιμη μαρκοβιανή αλυσίδα πάνω στον \mathbb{U} μέσω μίας μεθόδου γνωστής ως «κανονικοποίηση Λαπλασιανής γράφου» (*graph Laplacian normalization*) [4]. Κανονικοποιείται, λοιπόν, το $\mathbf{W}^{(\alpha)}$ ως προς τις γραμμές του:

$$\mathbf{P} = \left(\mathbf{D}^{(a)}\right)^{-1} \mathbf{W}^{(\alpha)} \quad (5.9)$$

ή με χρήση δεικτών:

$$P_{ij} = \frac{W_{ij}^{(\alpha)}}{d_i^{(\alpha)}} \quad (5.10)$$

όπου $\mathbf{D}^{(\alpha)}$ το διαγώνιο μητρώο που προκύπτει από το άθροισμα των γραμμών του $\mathbf{W}^{(\alpha)}$:

$$D_{ij}^{(\alpha)} = \begin{cases} d_i^{(\alpha)}, & i = j \\ 0, & i \neq j \end{cases} \quad (5.11)$$

και

$$d_i^{(\alpha)} = \sum_{r=1}^N W_{ir}^{(\alpha)} \quad (5.12)$$

Είναι προφανές ότι το \mathbf{P} παρόλο που δεν διατηρεί την συμμετρία, εντούτοις ικανοποιεί δύο βασικές ιδιότητες:

1. $P_{ij} \geq 0, \forall i, j \in \{1, \dots, N\}$
2. $\sum_{j=1}^N P_{ij} = 1 \implies \mathbf{P}\mathbf{1} = \mathbf{1}$

Οι δύο παραπάνω ιδιότητες επιτρέπουν στο $\mathbf{P} \in \mathbb{R}^{N \times N}$ να θεωρηθεί στοχαστικός (*stochastic*) ή μεταβατικός (*transition*) πίνακας.

5.5 Διαδικασία Διάχυσης

Οι δυνάμεις του στοχαστικού μητρώου συνεπώς, ορίζουν την διαδικασία διάχυσης (*diffusion process*), καθώς αυξανόμενου του t παρρατηρούνται δομές των δεδομένων σε διαφορετικές κλίμακες. Όταν το t είναι μικρό, τότε αποκαλύπτεται η τοπική γεωμετρία του συνόλου δεδομένων, ενώ όταν μεγαλώνει αποκαλύπτεται σταδιακά ή καθολική γεωμετρία των σημείων. Αυτό είναι λογικό, διότι, όπως προαναφέρθηκε το \mathbf{P}^t άθροίζει πιθανότητες όλων των μονοπατιών μήκους t .

Καθώς η διαδικασία εξελίσσεται, η πιθανότητα να ακολουθηθεί ένα μονοπάτι μήκους t αυξάνεται και συνεπώς αυξάνεται η πιθανότητα το μονοπάτι αυτό να είναι μέρος της υποκείμενης γεωμετρικής δομής των δεδομένων. Τοῦτο συμβαίνει γιατί η πυκνότητα των σημείων - άρα και η πιθανότητα αλληλομετάβασης - είναι μεγάλη εκεί που ακολουθείται η γεωμετρία των δεδομένων. Μονοπάτια που ακολουθούν την γεωμετρική δομή των δεδομένων, αποτελούνται από μικρά άλματα με μεγάλη πιθανότητα. Αντίθετα, όπου δεν υπάρχει κάποια υποκείμενη δομή, ή σύνδεση μεταξύ των σημείων είναι λιγότερο πιθανή, καθώς τα μονοπάτια περιλαμβάνουν μεγάλα άλματα με μικρή πιθανότητα.

Στην εικόνα 5.2, το κόκκινο μονοπάτι είναι πιθανότερο να διανυθεί από ότι το πράσινο, καθώς το t αυξάνει. Το πρώτο αποτελείται από μικρά άλματα με μεγάλη πιθανότητα, ενώ



Σχήμα 5.2: Μονοπάτια κατά την γεωμετρική δομή τών δεδομένων εμφανίζονται με μεγαλύτερη πιθανότητα.

το δεύτερο αποτελείται από μεγάλα άλματα που συνεχίζουν να αντιστοιχίζονται με μικρές πιθανότητες όσο κι αν αυξάνει η χρονική μεταβλητή.

5.6 Διαχυτική Απόσταση

Εν συνεχεία θα οριστεί το διαχυτικό μέτρο ή διαχυτική απόσταση που θα υπολογίζει αποστάσεις κατά την γεωμετρική δομή τών δεδομένων που προαναφέρθηκε. Η απόσταση αυτή μετρά την εγγύτητα δύο σημείων στον αρχικό χώρο τών δεδομένων ως την πιθανότητα μετάβασης από το ένα στο άλλο. Μπορεί να βρεθεί για κάθε τιμή της χρονικής παραμέτρου και χρησιμοποιεί κάθε δυνατό μονοπάτι αντιστοίχου μήκους που συνδέει τα δύο σημεία. Σχετίζεται άμεσα με το μητρώο μετάβασης και ορίζεται ως

$$\delta_t^2(\mathbf{u}_i, \mathbf{u}_j) = \sum_{\mathbf{x} \in \mathcal{U}} |P_t(\mathbf{u}_i, \mathbf{x}) - P_t(\mathbf{u}_j, \mathbf{x})|^2 = \sum_k |P_{ik}^t - P_{jk}^t|^2. \quad (5.13)$$

Ο όρος $P_t(\mathbf{u}_i, \mathbf{x})$ είναι πιθανότητα μετάβασης από το \mathbf{u}_i στο τυχόν \mathbf{x} σε t χρονικά βήματα. Επομένως, αφού η εξίσωση άθροίζει διαφορές, η διαχυτική απόσταση είναι μικρή όταν $P_t(\mathbf{u}_i, \mathbf{x}) \approx P_t(\mathbf{u}_j, \mathbf{x})$, άρα όταν τόσο το \mathbf{u}_i όσο και το \mathbf{u}_j είναι καλά συνδεδεμένα μέσω του \mathbf{x} , δηλαδή όταν υπάρχουν πολλά ισοπίθανα μονοπάτια μήκους t μεταξύ δύο σημείων. Το μέτρο, λοιπόν, αυτό δεν συγκρίνει άπευθείας τα δύο σημεία, αλλά έμμεσα μέσω τών γειτόνων τους.

Μπορεί, όμως, κανείς να θεωρήσει την διαχυτική απόσταση από μια διαφορετική έποψη. Η γραμμή \mathbf{P}_{i*}^t του στοχαστικού πίνακα, σύμφωνα με προηγούμενη διαπίστωση, μπορεί να νοηθεί ως μία διακριτή κατανομή στις κορυφές του γράφου όταν άφετηρία είναι ο κόμβος i — άς μὴν λησμονείται ότι $\sum_{j=1}^N P_{ij} = 1$. Άρα, η διαχυτική απόσταση $\sum_k |P_{ik}^t - P_{jk}^t|^2$ δὲν είναι παρά διαφορά δύο κατανομῶν με άφετηρία τα σημεία i και j . Εν όλίγοις, η διαχυτική απόσταση σε άντιδιαστολή με την εύκλείδια, δὲν έλέγχει το πού βρίσκονται τα σημεία, αλλά το πρὸς τα πού κατευθύνουν τὸν περίπατο.

Τέλος, σημαντικό πλεονέκτημα είναι ότι λόγω της συνεισφοράς όλων τών σημείων κατά τὸν υπολογισμό της, η διαχυτική απόσταση είναι άρκετά άκλινης σε σχέση με άλλες μεθόδους στὸν θόρυβο που συχνά παρουσιάζεται στις πρακτικές εφαρμογές.

5.7 Διαχυτική Άπεικόνιση - Diffusion Map

Στή συνέχεια ορίζεται ο μετασχηματισμός ο οποίος απεικονίζει τα αρχικά δεδομένα στον χώρο της διάχυσης

$$\mathbf{Y}^t(\mathbf{u}_i) = \begin{pmatrix} P_t(\mathbf{u}_i, \mathbf{u}_1) \\ P_t(\mathbf{u}_i, \mathbf{u}_2) \\ \vdots \\ P_t(\mathbf{u}_i, \mathbf{u}_N) \end{pmatrix} = \begin{pmatrix} P_{i1}^t \\ P_{i2}^t \\ \vdots \\ P_{iN}^t \end{pmatrix} = \mathbf{P}_{i*}^t = \mathbf{Y}_i^t, \quad (5.14)$$

και έτσι η διαχυτική απόσταση στον αρχικό χώρο μπορεί να γραφτεί ως εὐκλείδια απόσταση στον χώρο της διάχυσης

$$\|\mathbf{Y}_i^t - \mathbf{Y}_j^t\|^2 = \sum_{\mathbf{x} \in \mathbb{U}} |P_t(\mathbf{u}_i, \mathbf{x}) - P_t(\mathbf{u}_j, \mathbf{x})|^2 = \sum_k |P_{ik}^t - P_{jk}^t|^2 = \delta_t^2(\mathbf{u}_i, \mathbf{u}_j). \quad (5.15)$$

Η απεικόνιση αυτή διατηρεί την εσωτερική γεωμετρία των δεδομένων τα οποία υποθέτουμε ότι βρίσκονται σε κάποιο υποπολύπτυγμα. Ακόμη, όμως, δεν έχει επιτευχθεί ο τελικός στόχος που είναι η μείωση των διαστάσεων, καθώς μένει να βρεθούν ποιές διαστάσεις είναι κυρίαρχες και ποιές μπορούν να άμεληθούν. Το πρόβλημα αυτό, ως εἴθισται, τὸ λύνει ἡ φασματική ἀνάλυση.

Ἐπειδὴ ὁ πίνακας \mathbf{P} ἔχει ἄμεση σχέση μὲ τὴν κανονικοποιημένη Λαπλασιανή του γράφου που εἶναι συμμετρική, ἀποδεικνύεται ὅτι ἔχει πραγματικές ιδιοτιμές και ιδιοδιανύσματα. Ἐστω $\{\lambda_0, \lambda_1, \dots, \lambda_n\}$ οἱ πρῶτες $n < d$ ιδιοτιμές τοῦ \mathbf{P} κατὰ φθίνουσα σειρά και $\{\phi^0, \phi^1, \dots, \phi^n\}$ τὰ ἀντίστοιχα ιδιοδιανύσματα. Ὀρίζουμε τὴν συνάρτηση $\Psi^t : \mathbb{U} \rightarrow \mathbb{R}^n$ ὡς

$$\Psi^t(\mathbf{u}_i) = \begin{pmatrix} \lambda_1^t \phi^1(\mathbf{u}_i) \\ \lambda_2^t \phi^2(\mathbf{u}_i) \\ \vdots \\ \lambda_n^t \phi^n(\mathbf{u}_i) \end{pmatrix} = \begin{pmatrix} \lambda_1^t \phi_i^1 \\ \lambda_2^t \phi_i^2 \\ \vdots \\ \lambda_n^t \phi_i^n \end{pmatrix} = \psi_i^t. \quad (5.16)$$

Ὁ τελευταῖος μετασχηματισμός απεικονίζει τὰ δεδομένα σὲ ἕναν n -διάστατο εὐκλείδιο χώρο, ἐνῶ κάθε στοιχείο τοῦ διανύσματος $\Psi^t(\mathbf{u}_i)$ ἀποτελεῖ τὴν νέα συντεταγμένη διάχυσης σὸν μειωμένο χώρο. Ἐπειδὴ $\phi^0 = 1$, ἡ πρώτη ιδιομορφή μορφή μπορεί νὰ παραλείπεται καθὼς δὲν προσφέρει πληροφορία. Ὁ χώρος \mathbb{R}^n , ἐπειδὴ ἐδῶ ὑλοποιεῖται ἡ μείωση τῶν διαστάσεων, κατέχει ἀντὶ τοῦ προηγούμενου τὴν ὀνομασία «χώρος διάχυσης» (diffusion space ἢ DMAP space). Γιὰ νὰ εἶναι ἀποτελεσματική ἡ μείωση πρέπει $n \ll d$ ὅπου d ἡ διάσταση τῶν δεδομένων σὸν ἀρχικό χώρο.

Μὲ τὴν ὀργάνωση σὲ ἕνα ἐνιαῖο μητρώο ὄλων τῶν συντεταγμένων, ὄλων τῶν δεδομένων κατακευάζεται ἡ βάση τοῦ χώρου διάχυσης $\Psi^t \in \mathbb{R}^{n \times N}$

$$\Psi^t(\mathbb{U}) = \Psi_U^t = (\psi_1^t \quad \psi_2^t \quad \dots \quad \psi_N^t) = \begin{pmatrix} \lambda_1^t \phi^1 \\ \lambda_2^t \phi^2 \\ \vdots \\ \lambda_n^t \phi^n \end{pmatrix}. \quad (5.17)$$

Το σύνολο τών ὀρθογωνίων ιδιοδιανυσμάτων σχηματίζουν τὴν βάση τοῦ μειωμένου χώρου διάχυσης, ἐνῶ τὸ μέγεθος κάθε ιδιοτιμῆς δείχνει τὴν βαρύτητα τῆς ἀντίστοιχης διάστασης. Ὅπως καὶ πρὶν, ἡ εὐκλείδεια ἀπόσταση στὸν ἄρχικο χώρο τῆς διάχυσης ἀποδεικνύεται ὅτι προσεγγίζει τὴν διαχυτική ἀπόσταση στὸν ἀρχικό χώρο, δηλαδή

$$\delta_t^2(\mathbf{u}_i, \mathbf{u}_j) \approx \|\Psi^t(\mathbf{u}_i) - \Psi^t(\mathbf{u}_j)\|^2 = \|\psi_i^t - \psi_j^t\|^2 = \sum_k \left(\lambda_k \left| \phi_i^k - \phi_j^k \right| \right)^2. \quad (5.18)$$

Νὰ τονιστεῖ στὸ σημεῖο αὐτὸ ὅτι τὸ κόστος τῆς φασματικῆς ἀνάλυσης ἐξαρτάται μόνο ἀπὸ τὸν ἀριθμὸ τῶν δεδομένων N καὶ ὄχι ἀπὸ τὴν διάστασή τους d , καθὼς ὁ $\mathbf{P} \in \mathbb{R}^{N \times N}$. Αὐτὸ εἶναι ἐξαιρετικὸ πλεονέκτημα ὅταν οἱ διαστάσεις τοῦ ἀρχικοῦ χώρου εἶναι μεγαλύτερες ἀπὸ τὸ πλῆθος τῶν δεδομένων καὶ τὴν καθιστᾷ ἀνταγωνιστικὴ ὡς πρὸς τὴν PCA σὲ ἀνάλογες περιπτώσεις.

5.8 Αντίστροφη Άπεικόνιση

Μέχρι στιγμῆς ἔχει ὀλοκληρωθεῖ ἡ μείωση τῶν διαστάσεων τοῦ προβλήματος, δηλαδή ἔχει κατασκευασθεῖ ἡ ὀρθή ἀπεικόνιση ἀπὸ τὸν περιβάλλοντα εὐκλείδιο χώρο τῶν δεδομένων στὸν μειωμένων διαστάσεων χώρο τῆς διάχυσης. Ἡ ἀντίστροφη διαδικασία δὲν ἀποτελεῖ εὐκόλο πρόβλημα. Ἐδῶ τὸ πλεονέκτημα τῆς μεθόδου, ποὺ εἶναι ἡ μὴ γραμμικότητα τῆς ὀρθῆς ἀπεικόνισης, γίνεται μειονέκτημα, καθὼς πρέπει νὰ ἐπινοηθεῖ μιὰ ἐπίσης μὴ γραμμικὴ ἀντίστροφη ἀπεικόνιση. Στὴν βιβλιογραφία ὑπάρχουν ἀρκετές τέτοιες μέθοδοι. Στὸ πλαίσιο τῆς παρούσης ἐργασίας, ὅμως, θὰ χρησιμοποιηθεῖ ἕνας κλασικὸς γραμμικὸς μετασχηματισμὸς ποὺ, ναί μὲν δὲν διαθέτει τὴν μαθηματικὴ λεπτομερῆ πρὸς τὴν σύνθετων μεθόδων, εἶναι, ὅμως, ἀσυγκρίτως ταχύτερος ἀπὸ αὐτές.

Ζητεῖται ἡ εὕρεση μητρῶου $\Psi^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^d$, $n < d$, τέτοιου ὥστε κατὰ τὸ δυνατόν νὰ ἰσχύει

$$\mathbf{u}_i \approx \Psi^{-1} \psi_i, \quad i = 1, \dots, N \quad (5.19)$$

Ἀμέσως γίνεται ἀντιληπτὸ ὅτι τὸ παραπάνω πρόβλημα δὲν μπορεῖ νὰ ἰσχύει χωρὶς σφάλματα, καθότι $\mathbf{u}_i \in \mathbb{R}^d$, $\psi_i \in \mathbb{R}^n$ καὶ $\Psi^{-1} \in \mathbb{R}^{d \times n}$, $n < d$. Ἄρα, τὸ πρόβλημα μετατρέπεται σε κλασικὸ πρόβλημα βελτιστοποίησης καὶ λύνεται, ὡς γνωστὸν, λόγῳ τῆς γραμμικότητάς, μὲ τὴν μέθοδο ἐλαχίστων τετραγώνων

$$\begin{aligned} \mathbf{U} &\approx \Psi^{-1} \Psi(\mathbf{U}) = \Psi^{-1} \Psi_U \\ &\implies (\mathbf{U})^T \approx (\Psi^{-1} \Psi_U)^T \\ &\implies (\Psi_U)^T (\Psi^{-1})^T \approx \mathbf{U}^T \end{aligned} \quad (5.20)$$

ὅπου

- $\mathbf{U} = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_N) \in \mathbb{R}^{d \times N}$ ὁ πίνακας ποὺ προκύπτει ἀπὸ τὸ ἀρχικὸ σύνολο δεδομένων
- $\Psi_U^2 \in \mathbb{R}^{n \times N}$ ὁ μὴ ἀντιστρέψιμος πίνακας τῆς 5.17

Παρατηρῶντας τὴν 5.20 φαίνεται ἐμφανῶς ὅτι πρόκειται γιὰ ὑπερορισμένο σύστημα N ἐξισώσεων. Ἐτσι, πολλαπλασιάζοντάς ἐκατέρωθεν τὴν ἴδια ἐξίσωση μὲ Ψ_U καὶ χρησιμοποιώντας τὴν συμμετρία τοῦ πίνακα $\Psi_U \Psi_U^T$ προκύπτει

²Γιὰ νὰ εἶναι εὐανάγνωστες οἱ ἐξισώσεις, παραλείπεται τὸ t στὸν ἐκθέτη.

$$\Psi^{-1} \approx \mathbf{U} \Psi_U^T (\Psi_U \Psi_U^T)^{-1} \quad (5.21)$$

Έτσι, κάθε νέο διάνυσμα στον αρχικό χώρο, μπορεί να γραφτεί

$$\mathbf{u} \approx \Psi^{-1} \psi \quad (5.22)$$

Άς ξανατονιστεί στο σημείο αυτό, ότι για την μέθοδο των Diffusion Maps ο γραμμικός μητρικός μετασχηματισμός Ψ^{-1} μειώνει την ακρίβεια και αναιρεί πολλά από τα όφελη της μη γραμμικότητας προς όφελος της ταχύτητας.

5.9 Σύνοψη τής μεθόδου

Algorithm 4: DMAPs algorithm

Data: Dataset matrix $\mathbf{U} \in \mathbb{R}^{d \times N}$

Result: Compute the matrix Ψ^{-1}

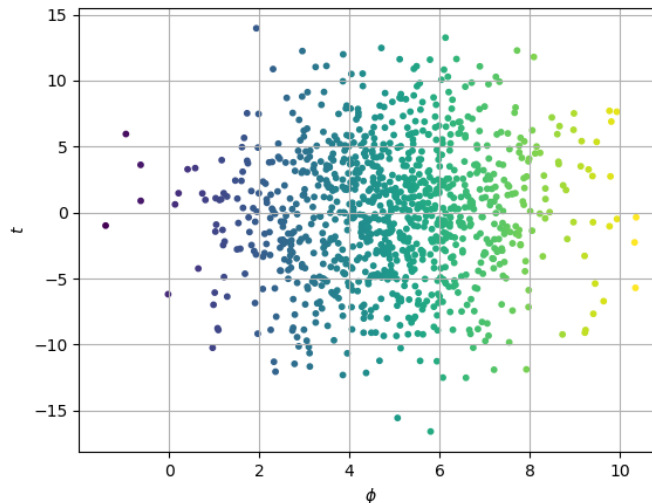
- 1 Initialize;
 - 2 SELECT a suitable kernel function $k : \mathbf{U} \times \mathbf{U} \rightarrow \mathbb{R}$;
 - 3 SELECT a valid hyperparameter ϵ ;
 - 4 COMPUTE the matrix with elements $W_{ij} = k(\mathbf{u}_i, \mathbf{u}_j)$;
 - 5 SELECT the parameter α based on your problem;
 - 6 COMPUTE the degrees of each point $d_i = \sum_k W_{ik}$;
 - 7 NORMALIZE by degree $W_{ij}^\alpha = \frac{W_{ij}}{d_i^\alpha d_j^\alpha}$;
 - 8 NORMALIZE by row sum $P_{ij} = \frac{W_{ij}^\alpha}{\sum_k W_{ik}^\alpha}$;
 - 9 COMPUTE n largest eigenvalues λ_i and eigenvectors ϕ_i of \mathbf{P} ;
 - 10 FORM the matrix $\Psi = \begin{pmatrix} \lambda_1 \phi_1 \\ \vdots \\ \lambda_n \phi_n \end{pmatrix}$;
 - 11 COMPUTE the pseudo-inverse $\Psi^{-1} = \mathbf{U} \Psi_U^T (\Psi_U \Psi_U^T)^{-1}$
-

Κεφάλαιο 6

Άριθμητικές Έφαρμογές

6.1 Έποπτικό Παράδειγμα μείωσης διαστάσεων

Πριν αναλυθεί το κύριο πρόβλημα της δυναμικής ανάλυσης, θα παρουσιαστεί ένα πολύ σύντομο παράδειγμα χρήσης των μεθόδων μείωσης διαστάσεων για έποπτικούς κυρίως λόγους. Συγκεκριμένα, θα εξακριβωθεί το πώς επιδρά κάθε μέθοδος μείωσης διαστάσεων στο ίδιο σύνολο δεδομένων.



Σχήμα 6.1: Τυχαίες υλοποιήσεις των t και ϕ .

Έστω λοιπόν η παραμετρική επιφάνεια $S : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ με εξίσωση

$$S(t, \phi) = \begin{pmatrix} x(t, \phi) \\ y(t, \phi) \\ z(t, \phi) \end{pmatrix} = \begin{pmatrix} t \\ \phi \sin(\phi) \\ \phi \cos(\phi) \end{pmatrix}. \quad (6.1)$$

όπου t και ϕ τυχαίες μεταβλητές με κατανομή όπως στον πίνακα 6.1. Στην βιβλιογραφία η S είναι γνωστή και ως "swiss roll" για προφανείς λόγους.

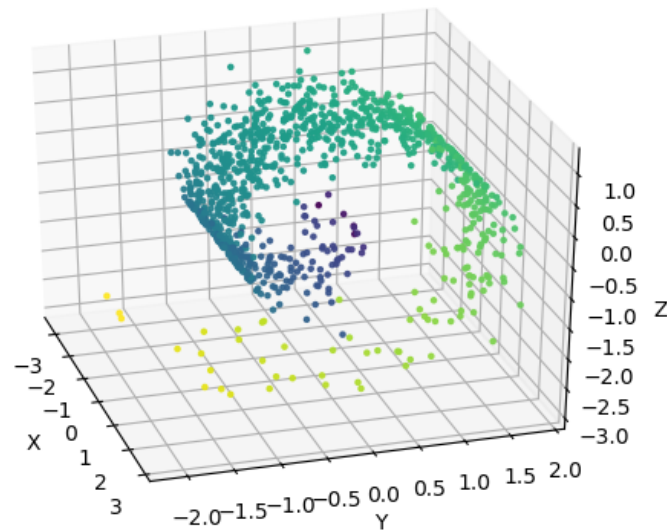
Πίνακας 6.1: Παράμετροι τών τυχαίων μεταβλητών.

Μεταβλητή	Κατανομή
t	$\mathcal{N}(0, 5^2)$
ϕ	$\mathcal{N}(5, 2^2)$

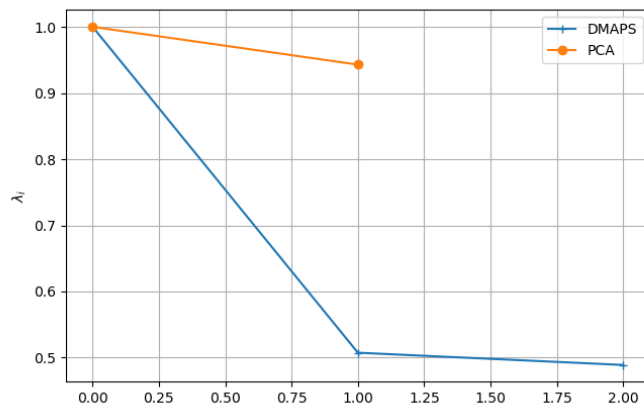
Έστω N υλοποιήσεις τών t και ϕ και ο αντίστοιχος μετασχηματισμός τους από την συνάρτηση S . Αν συγκεντρωθούν σε ένα σύνολο δεδομένων

$$U = \{\mathbf{u}_i \mid \mathbf{u}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} x(t_i, \phi_i) \\ y(t_i, \phi_i) \\ z(t_i, \phi_i) \end{pmatrix}, i = 1, 2, \dots, N\}, \quad (6.2)$$

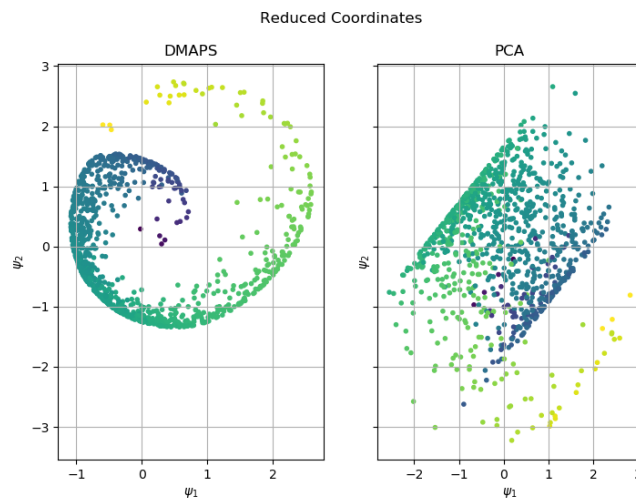
τότε αυτό μπορεί να τροφοδοτηθεί στις μεθόδους μείωσης διαστάσεων. Αυτές με την σειρά τους θα δώσουν τις αναπαραστάσεις τών δεδομένων ξανά πίσω στον \mathbb{R}^2 .



Σχήμα 6.2: Σύνολο δεδομένων "swiss roll" σε 3Δ σύμφωνα με την εξίσωση $(t, \phi \sin(\phi), \phi \cos(\phi))$.



Σχήμα 6.3: Ίδιοτιμές: κανονικοποιημένη για την PCA, ενώ για τα DMAPs η πρώτη δεν δίνει πληροφορία όντας πάντα ίση με την μονάδα.



Σχήμα 6.4: Το σύνολο δεδομένων στις 2Δ, όπως μειώθηκε από κάθε τεχνική – παράμετροι DMAPs ($\epsilon = 1, \alpha = 0, t = 1$).

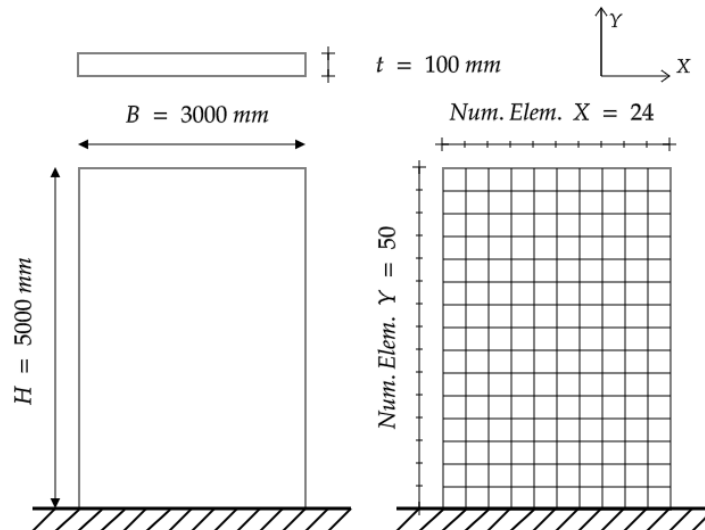
6.2 Δυναμικό πρόβλημα

6.2.1 Όρισμός

Όπως έχει ήδη προαναφερθεί, ζητούνται λύσεις της στοχαστικής, γραμμικής διαφορικής εξίσωσης ενός δυναμικού μηχανικού συστήματος. Η επίλυση του στοχαστικού προβλήματος γίνεται με την μέθοδο Monte Carlo που απαιτεί πολλά δεδομένα εισόδου και εξόδου, συνεπώς και πολλές λύσεις της δυναμικής εξίσωσης. Αφού, λοιπόν ληφθούν κάποιες αρχικές λύσεις με το πλήρες μοντέλο, αυτές χρησιμοποιούνται για την εκπαίδευση του υποκατάστατου μοντέλου. Το υποκατάστατο μοντέλο θα δώσει λύσεις προσεγγιστικές, αλλά θα μειώσει το υπολογιστικό κόστος.

6.2.2 Καθορισμός μοντέλου

Για την αριθμητική εφαρμογή επιλέχτηκε ο πρόβλος του σχήματος 6.5 στο οποίο αναγράφονται οι διαστάσεις, οι συνοριακές συνθήκες και η διακριτοποίηση του πλέγματος των πεπερασμένων στοιχείων.



Σχήμα 6.5: Οι διαστάσεις, οι συνοριακές συνθήκες και το πλέγμα του προβλήματος.

Ο φορέας προσομοιώνεται από διδιάστατα, τετρακομβικά πεπ. στ. επίπεδης έντασης. Συνολικά, ο φορέας αποτελείται από $N_{elX} \cdot N_{elY} = 24 \cdot 50 = 1200$ στοιχεία και, συνεπώς οι συνολικοί βαθμοί ελευθερίας του μοντέλου είναι $N_{dof_s} = 2 \cdot 50 \cdot (24 + 1) = 2500$, δηλαδή οι λύσεις του συστήματος $\mathbf{u}_i \in \mathbb{R}^{2500}$. Προφανώς για ένα τέτοιου μεγέθους μοντέλο δεν θα επιχειρούσε κάποιος να εφαρμόσει τεχνικές μείωσης διαστάσεων, παρά μόνο σε έρευνητικά πλαίσια.

Το μέτρο ελαστικότητας είναι 1D-1V στοχαστικό πεδίο επάνω σε διδιάστατο πεδίο όρισμού και μεταβάλλεται κατά την διεύθυνση Y

$$E = \tilde{\mathbf{E}}(y, \vartheta) = \tilde{\mathbf{E}}(y)$$

το οποίο ελήφθη από ανάπτυγμα Karhunen-Loève κανονικής κατανομής. Η εκπαιδευτική φόρτιση αποτελείται από 3D-2V στοχαστικό ήμιτονικό επιταχυνσιογράφημα στην βάση

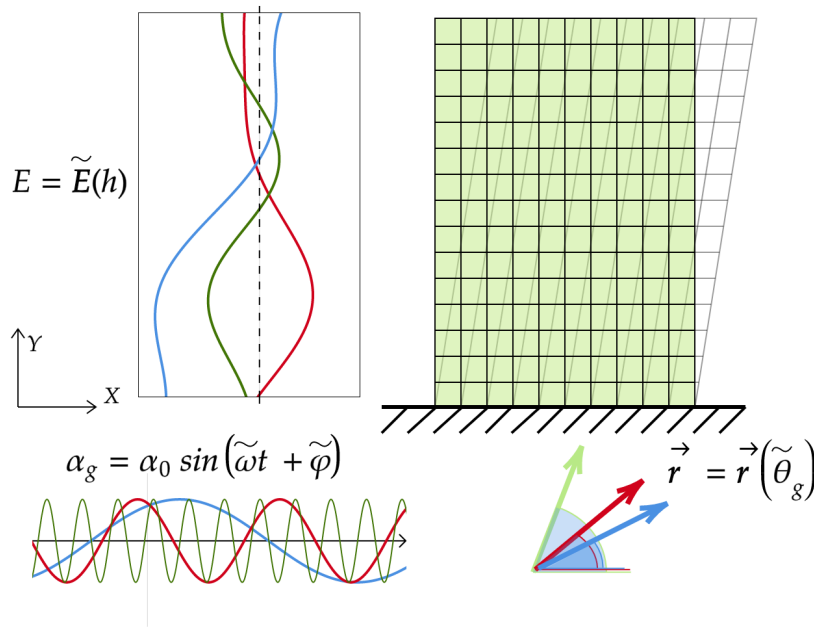
Πίνακας 6.2: Παράμετροι τών τυχαίων μεταβλητών.

Μεταβλητή	Κατανομή	Μονάδες
$\tilde{E}(h)$	$\mathcal{N}(30, 6^2)$	GPa
$\tilde{\omega}$	$\mathcal{U}(0, 50)$	Hz
$\tilde{\varphi}$	$\mathcal{U}(0, 2\pi)$	rad
$\tilde{\theta}$	$\mathcal{U}(0, \pi/2)$	rad

του φορέα.

$$\vec{\alpha}_g = \alpha_g(\tilde{\omega}, \tilde{\varphi}) \vec{r}(\tilde{\theta}) = \alpha_0 \sin(\tilde{\omega}t + \tilde{\varphi}) \begin{pmatrix} \cos(\tilde{\theta}) \\ \sin(\tilde{\theta}) \end{pmatrix}.$$

Κυρίως λόγω του στοχαστικού μέτρου ελαστικότητας πρέπει να σημειωθεί ότι οι λύσεις αναμένεται να κινούνται σε έναν μη γραμμικό υπόχωρο. Οι παράμετροι τών τυχαίων μεταβλητών αναγράφονται στον πίνακα 6.2, ενώ, τέλος, ο συντελεστής απόσβεσης είναι $\zeta = 5\%$.

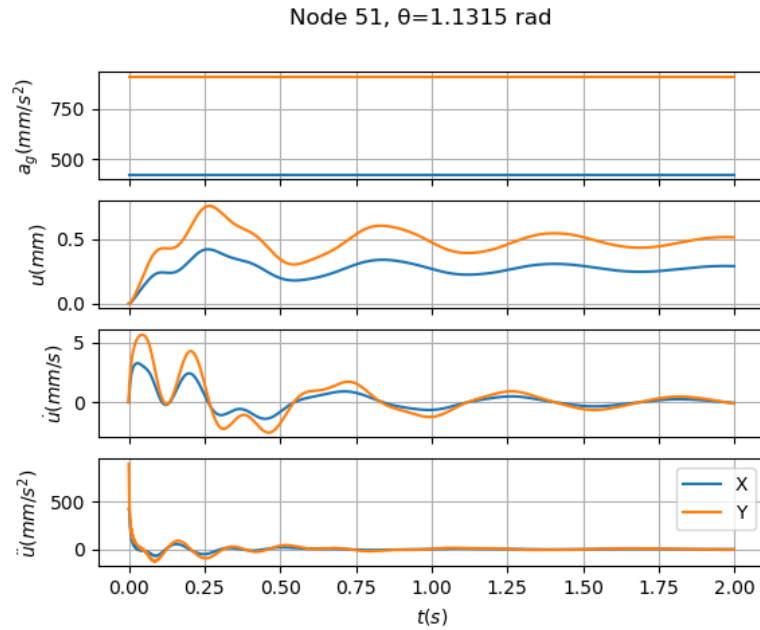


Σχήμα 6.6: Στοχαστικά πεδία του μέτρου ελαστικότητας (E) και της σεισμικής διέγερσης ($\vec{\alpha}_g$).

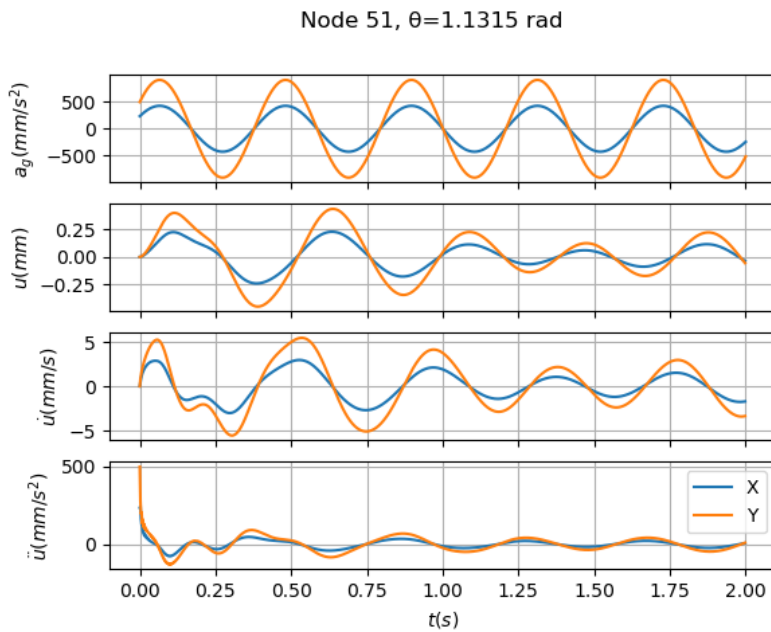
Άρχικώς, παρατίθενται τὰ επιταχυνσιογραφήματα δύο φορτίσεων και οι αντίστοιχες αποκρίσεις ενός κόμβου για να επιβεβαιωθεί – ποιοτικά τουλάχιστον – ότι το πλήρες μοντέλο λειτουργεί σωστά.

6.2.3 Διαλογή δειγμάτων

Προκειμένου να ελεγχθεί η ακρίβεια, παράχθηκαν με την μέθοδο πεπερασμένων στοιχείων 3000 λύσεις από το πλήρες προσομοίωμα οι οποίες θα συγκριθούν με τις αντίστοιχες λύσεις τών υποκατάστατων μοντέλων. Έτσι, όπως προειπώθηκε, σε πρώτη φάση γίνεται προσπάθεια να βρεθεί η δομή ενός υποπολυτύγματος – αλλιώς πολλαπλότητας – του \mathbb{R}^{2500} επάνω στο οποίο κινούνται κατά προσέγγιση οι λύσεις του συστήματος. Για τόν

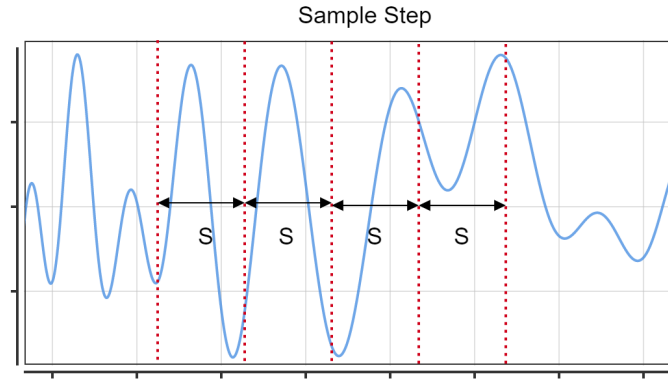


Σχήμα 6.7: Φόρτιση με σταθερή επιτάχυνση εδάφους a_g και η αντίστοιχη απόκριση κόμβου όπως προκύπτει από το πλήρες μοντέλο (Full Order Model).



Σχήμα 6.8: Φόρτιση με ήμιτονοειδή επιτάχυνση εδάφους a_g και η αντίστοιχη απόκριση κόμβου όπως προκύπτει από το πλήρες μοντέλο FOM.

λόγο αυτό χρησιμοποιείται ένα σύνολο από N εκπαιδευτικές φορτίσεις, δηλαδή σημεία στον \mathbb{R}^{2500} , των οποίων οι λύσεις βρίσκονται από το ακριβές, πλήρες μοντέλο του φορέα.



Σχήμα 6.9: Από κάθε χρονοσειρά επιλέγονται $N_S = \frac{N_{\text{timesteps}}}{S}$ λύσεις, όπου S το χρονικό βήμα.

Ός γνωστόν η υπολογιστική ισχύς τόσο σε μνήμη όσο και σε ταχύτητα είναι πεπερασμένη. Επειδή χρειάζεται ένας ικανός αριθμός N_{train} από εκπαιδευτικές προσομοιώσεις και καθώς κάθε πλήρης χρονοσειρά ενδέχεται να έχει πολλά χρονικά βήματα, εγείρεται το θέμα της επιλογής των δειγμάτων. Προφανώς δεν μπορούν να επιλεγούν όλα τα $N_{\text{timesteps}}$ διανύσματα κάθε προσομοίωσης, γιατί αυτό θα άυξανε πολύ γρήγορα τον συνολικό αριθμό εκπαιδευτικών διανυσμάτων.

Χρειάζεται προσοχή, λοιπόν, κατά την διαλογή των δειγμάτων, να επιλεγούν έτσι ώστε να είναι αντιπροσωπευτικά τόσο διαφορετικών φορτίσεων όσο και διαφορετικών χρονικών στιγμών. Για τους παραπάνω λόγους και εφαρμόζοντας το απλούστερο δυνατό σχήμα, από κάθε χρονοσειρά με $N_{\text{timesteps}} = 100$ επιλέγονται με βήμα $S = \text{Sample step} = 10$ ορισμένα μόνο, $N_S = \frac{N_{\text{timesteps}}}{S}$ το πλήθος, στιγμιότυπα της απόκρισης. Στην συνέχεια αυτά συλλέγονται σε ένα ενιαίο σύνολο εκπαίδευσης \mathbb{U} που αποτελείται από τα επιλεχθέντα στιγμιότυπα των λύσεων όλων των προσομοιώσεων. Έτσι, συνολικά τα σημεία του \mathbb{R}^{2500} που χρησιμοποιούνται στα υποκατάστατα μοντέλα είναι

$$N = N_{\text{train}} \cdot N_S = N_{\text{train}} \frac{N_{\text{timesteps}}}{S} \quad (6.3)$$

και το σύνολο

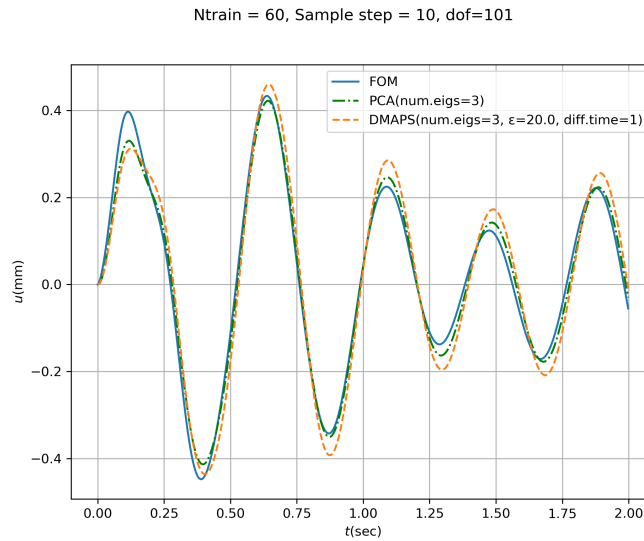
$$\begin{aligned} \mathbb{U}_{\text{training}} &= \{\mathbf{u}_{11}, \dots, \mathbf{u}_{1N_S}, \dots, \mathbf{u}_{N_{\text{train}}1}, \dots, \mathbf{u}_{N_{\text{train}}N_S}\} \\ &= \{\mathbf{u}_{ij} : i = 1, \dots, N_{\text{train}}, j = 1, \dots, N_S\} \end{aligned} \quad (6.4)$$

Έπειτα, οι δύο μέθοδοι μείωσης διαστάσεων – DMAPs και PCA – εφαρμόζονται στα προηγούμενα διανύσματα των λύσεων προκειμένου να βρεθεί ο μετασχηματισμός από τον αρχικό χώρο στον χώρο μικρότερης διάστασης που αντιστοιχεί στο υποπολύπτυγμα.

6.2.4 Αποτελέσματα

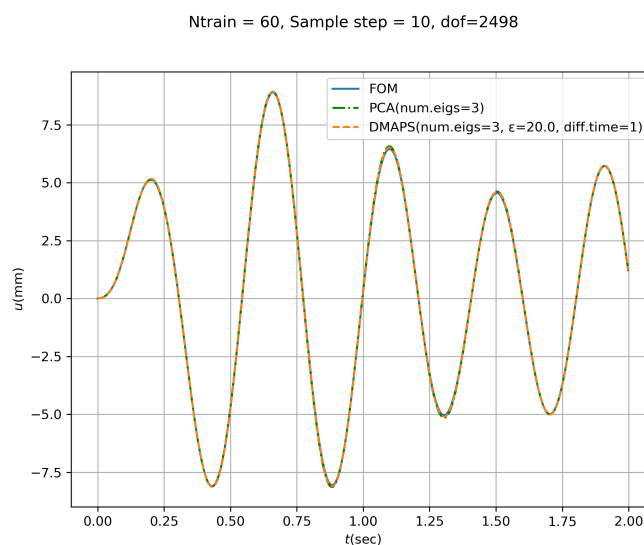
Για να υπάρχει μέτρο σύγκρισης επιλέχτηκε ο χώρος αυτός να έχει την ίδια διάσταση και για τις δύο μεθόδους. Η σύγκριση των μεθόδων και της ακριβούς λύσης γίνεται σε ορισμένους βαθμούς ελευθερίας με σημείο αναφοράς τις χρονοσειρές των αποκρίσεων και τις συναρτήσεις πυκνότητας πιθανότητας όλης της απόκρισης ($t = [0, 2] \text{sec}$). Οι συγκρίσεις της χρονοσειράς της απόκρισης φαίνεται στο σχήμα 6.10 και 6.11.

Άξιζει να ερευνηθεί η επίρροή του χρησιμοποιούμενου αριθμού ιδιοτιμών k στην ακρίβεια των υποκατάστατων μοντέλων που φαίνεται στα σχήματα 6.14 και 6.15.

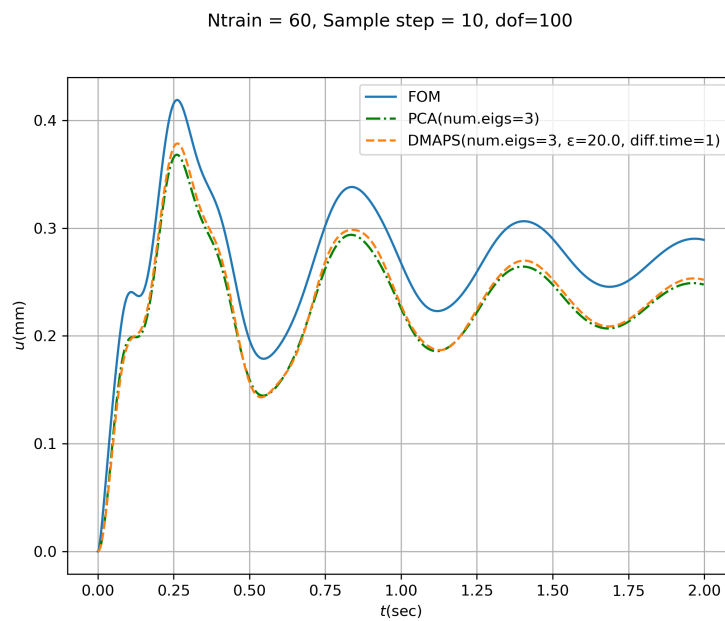


Σχήμα 6.10: Σύγκριση αποκρίσεων για τόν βαθμό ελευθερίας 101 – στην βάση του προβόλου – με χρήση $k = 3$ ιδιοτιμών στα υποκατάστατα μοντέλα (Reduced Order Models).

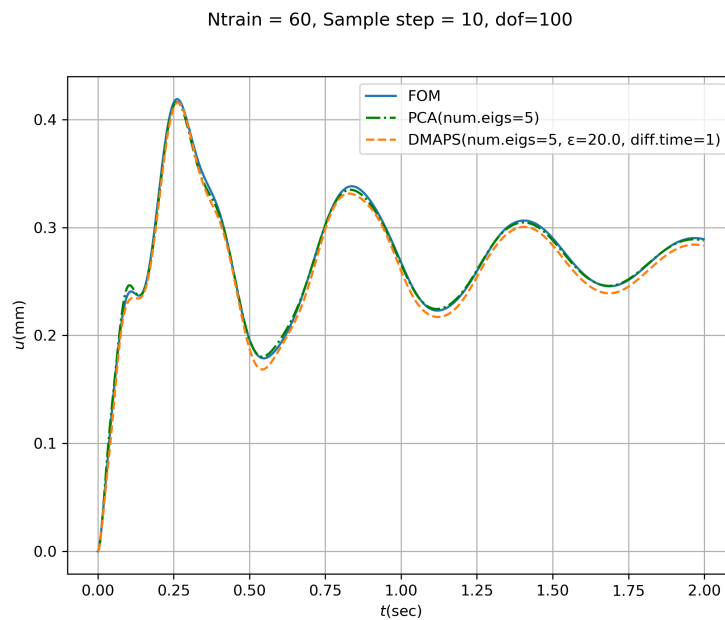
Έν συνεχεία παρουσιάζονται οι κατανομές τών μετακινήσεων για όρισμένους βαθμούς ελευθερίας κατά τις χρονικές στιγμές $t = 1sec$ και $t = 2sec$.



Σχήμα 6.11: Σύγκριση αποκρίσεων για τόν βαθμό ελευθερίας 2498 – στο ελεύθερο άκρο του προβόλου – με χρήση $k = 3$ ιδιοτιμών στα υποκατάστατα μοντέλα (ROM)

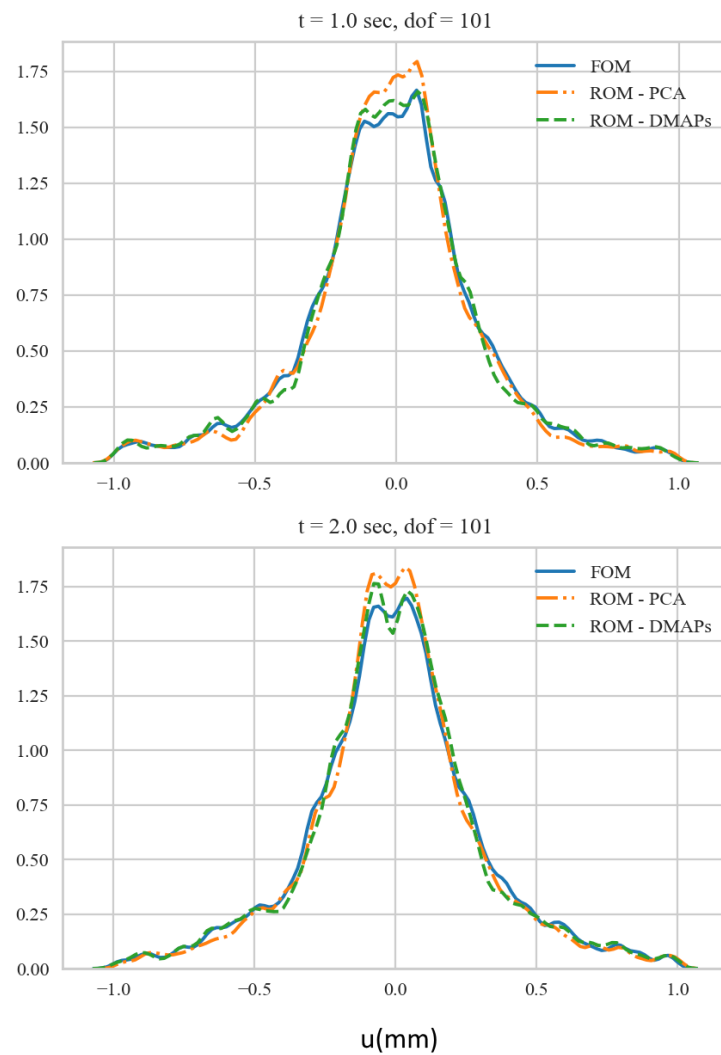


Σχήμα 6.12: Σύγκριση αποκρίσεων για τον βαθμό ελευθερίας 100 με χρήση $k = 3$ ιδιοτιμών στα υποκατάστατα μοντέλα.

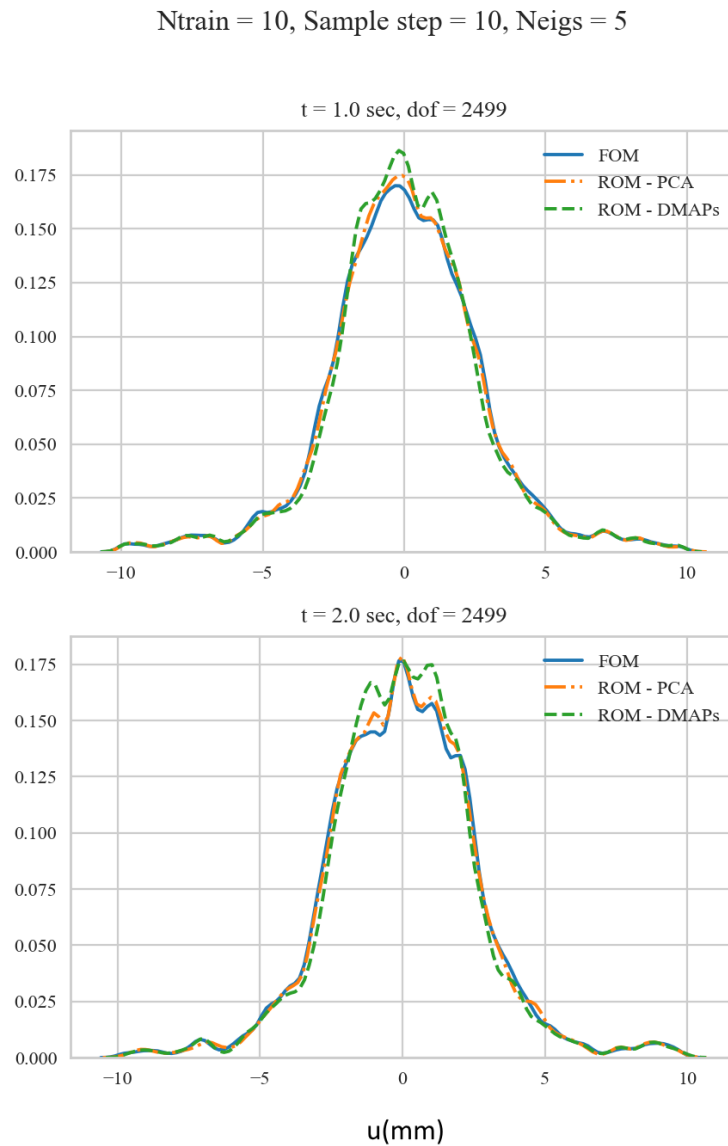


Σχήμα 6.13: Σύγκριση αποκρίσεων για τον βαθμό ελευθερίας 100 με χρήση $k = 5$ ιδιοτιμών στα υποκατάστατα μοντέλα.

Ntrain = 10, Sample step = 10, Neigs = 5



Σχήμα 6.14: Συναρτήσεις πυκνότητας πιθανότητας (pdf) για τον βαθμό ελευθερίας 100 όπως προκύπτει από την πλήρη Monte Carlo ανάλυση (FOM) και τα υποκατάστατα μοντέλα (ROM-DMAPs, ROM-PCA).



Σχήμα 6.15: Συναρτήσεις πυκνότητας πιθανότητας (pdf) για τον βαθμό έλευθερίας 2499 όπως προκύπτει από την πλήρη Monte Carlo ανάλυση (FOM) και τα υποκατάστατα μοντέλα (ROM-DMAPs, ROM-PCA).

Τέλος, δέν θα πρέπει νά λησμονηθεῖ κι ὁ χρόνος ἐπίλυσης τοῦ συστήματος, ἡ ἱκανή μείωση τοῦ ὁποῖου εἶναι ἐξίσου σημαντική μέ τήν ἐπίτευξη τῆς ακρίβειας. Προκειμένου νά ὑπάρχει σύγκριση στήν ἀκρίβεια τῶν ὑποκατάστατων, χρησιμοποιήθηκε γιά ὅλες τίς ἀναλύσεις ὁ ἴδιος ἀριθμός ἰδιοτιμῶν, καί δεδομένου ὅτι καί οἱ δύο μέθοδοι ἐφαρμόζονται γραμμικά στήν ἐξίσωση ἰσορροπίας, οἱ δύο χρόνοι εἶναι κοινοί. Τά ἀποτελέσματα φαίνονται στόν πίνακα 6.3.

Πίνακας 6.3: Χρόνοι ἐκτέλεσης γιά τὸ πλήρες καί τὰ ὑποκατάστατα μοντέλα.

Μοντέλο	Διάσταση πινάκων	Ἀριθμός ἀναλύσεων	Συνολικός Χρόνος	Μέσος χρόνος
FOM	2500	3000	103.29 min	2066 ms
ROM	5	2800	30.78 min	659 ms

Ἐπίσης, σημειώνεται ὅτι οἱ χρόνοι ἐκπαίδευσης τῶν ὑποκατάστατων μοντέλων εἶναι τῆς τάξης τῶν μερικών ἑκατοντάδων ms, τουλάχιστον γιά τὸ παρόν πρόβλημα, καί ἀποτελοῦν ἀμελητέο ποσοστὸ μιάς πλήρους ἀνάλυσης. Παρόλα αὐτὰ σὲ προβλήματα μέ πολύ μεγαλύτερες διαστάσεις, ὁ χρόνος ἐκπαίδευσης τῶν μεθόδων ἀναμένεται νά εἶναι καθοριστικός παράγοντας ὁ ὁποῖος θά κρίνει τὴν ἀποτελεσματικότητά τους.

6.2.5 Συμπεράσματα

Ἐν ὀλίγοις, ἡ διαδικασία ποὺ ἀκολουθήθηκε μπορεῖ νά συνοψισθεῖ στὰ ἑξῆς. Τὸ ἀρχικὸ πρόβλημα στόν ἀρχικὸ χῶρο ἀνάχθηκε σὲ δύο ὑποπροβλήματα.

- Τὸ πρῶτο ὑποπρόβλημα εἶναι ἡ ἐκπαίδευση μιᾶς μεθόδου manifold learning που μαθηματικά ἰσοδυναμεῖ μέ τὴν φασματικὴ ἀνάλυση ἑνός πίνακα πολλῶν διαστάσεων.
- Τὸ δεύτερο ὑποπρόβλημα ἀποτελεῖ ἡ ἐπίλυση τοῦ δυναμικοῦ προβλήματος σὲ χῶρους πολύ μικρότερων διαστάσεων ἀπὸ τὸν ἀρχικὸ.

Παρατηρώντας τὰ γραφήματα, δύναται κανεὶς νά συμπεράνει τὰ ἑξῆς:

- Παρὰ τὴν γραμμικότητα τῶν τεχνικῶν (PCA - πλήρως γραμμική, DMAPs - μόνο ὁ μετασχηματισμός) καί τὰ δύο ὑποκατάστατα μοντέλα κατάφεραν νά προσεγγίσουν ἱκανοποιητικά τίς λύσεις τοῦ πλήρους μοντέλου σὲ ἓνα μὴ γραμμικὸ πρόβλημα.
- Οἱ δύο μέθοδοι παρήγαγαν μοντέλα μέ παραπλήσια ἀκρίβεια. Λίγο καλύτερα ἀποτελέσματα φαίνεται νά δίνει ἡ PCA.
- Ἡ μείωση τοῦ χρόνου ἐπίλυσης τῆς ἀνάλυσης βρίσκεται στὰ ἐπιθυμητὰ ὅρια γιά ὑποκατάστατα μοντέλα. Ἐν προκειμένῳ τὰ ὑποκατάστατα μοντέλα εἶναι 3.14 φορές ταχύτερα.
- Σὲ βαθμούς ἐλευθερίας κοντὰ στὸ ἐλεύθερο ἄκρο (π.χ. 2098), ὅπου οἱ μετακινήσεις ἦταν μεγαλύτερες κατὰ ἀπόλυτη τιμή, τὰ ὑποκατάστατα μοντέλα προσαρμόζονταν γρηγορότερα — ἤδη ἀπὸ $k = 3$ —, ἐνῶ σὲ βαθμούς ἐλευθερίας κοντὰ στήν βάση (π.χ. 101) χρειάζονταν περισσότερες ἰδιοτιμές.

6.2.6 Σημεῖα πρὸς ἐρευνητικὴ ἐμβάθυνση

Στὴν παρούσα ἐργασία χρησιμοποιήθηκαν μεικτὰ μοντέλα, δηλαδή μοντέλα που συνδυάζουν τὴν φυσικὴ τοῦ προβλήματος (FEM) μέ μεθόδους μείωσης διαστάσεων (DMAPs, PCA).

Ἐπίσης, ἡ μέθοδος τῶν DMAPs ὅπως ἔχει ἤδη προαναφερθεῖ εἶναι μὴ γραμμικὴ ἀπὸ τὴν φύση τῆς. Στὴν παρούσα ἐργασία συνδυάστηκε μέ γραμμικὸ μετασχηματισμὸ ἀπὸ τὸν

χῶρο τῶν λύσεων στὸν χῶρο μειωμένων διαστάσεων πράγμα ποὺ ἀναιρεῖ πολλὰ ὠφέλη τῆς μὴ γραμμικότητας πρὸς ὄφελος τῆς ταχύτητας.

Σύμφωνα μὲ τὰ παραπάνω, ἐρευνητικό ἐνδιαφέρον ἐνδεχομένως νὰ ἔχουν οἱ ἐξῆς προτάσεις, οἱ ὁποῖες κατὰ τὴν γνώμη τοῦ γράφοντα θὰ μπορούσαν νὰ ἀποτελέσουν τὸ ἔναυσμα γιὰ περαιτέρω μελέτη:

- Χρήση πλήρους μὴ γραμμικῆς τεχνικῆς γιὰ τὴν μέθοδο τῶν DMAPs, προκειμένου νὰ γίνει ἐκμετάλλευση τῆς μεθόδου στὸν μέγιστο βαθμὸ.
- Πλήρης ἀπεξάρτηση ἀπὸ τὸ φυσικὸ μοντέλο μὲ χρήση ἐναλλακτικῶν ὑποκατάστατων μοντέλων μηχανικῆς μάθησης (π.χ. νευρωνικῶν δικτύων).
- Συνδυασμὸς μηχανικῆς μάθησης καὶ μεθόδων μείωσης διαστάσεων γιὰ ταχύτερη ἐκπαίδευση καὶ μεγαλύτερη ἀπόδοση τοῦ μοντέλου.

Γενικά, τὸ πεδίο τῆς μεταμοντελοποίησης εἶναι εὐρὺ, ἀλλὰ καὶ ταχέως ἀναπτυσσόμενο, καθὼς ὄλο καὶ περισσότεροι ἐπιστήμονες, εἴτε ἐκ τῶν πραγμάτων εἴτε ἐξ ἐρευνητικοῦ ἐνδιαφέροντος στρέφονται πρὸς ὑποκατάστατα μοντέλα γιὰ τὴν λύση τῶν φυσικῶν προβλημάτων. Σὲ μιὰ ἐποχὴ ποὺ στηρίζεται στὶς στέρεες θεωρητικὲς ἐπιστημονικὲς βάσεις παλαιότερων ἐτῶν, ἡ πρόκληση πλέον εἶναι ἡ εὕρεση τοῦ βέλτιστου συνδυασμοῦ γνωσιακῶν ἐργαλείων ποὺ θὰ παρέχουν τὸ ἐκάστοτε ἐπιθυμητὸ ἀποτέλεσμα.

Βιβλιογραφία

- ¹V. P. I. Kalogeris, «Diffusion maps-based surrogate modeling - An alternative machine learning approach», *International Journal for numerical methods in Engineering* (2019).
- ²S. Ghanem, *Stochastic Finite Elements, A Spectral Approach* (Springer, 1991).
- ³D. G. G. Vissarion Papadopoulos, *Stochastic Finite Element Methods, An Introduction* (2018).
- ⁴F. R. K. Chung, *Spectral Graph Theory* (1997).
- ⁵I. G. K. R. R. C. A. Singer R. Erban, «Detecting intrinsic slow variables in stochastic dynamical systems by anisotropic diffusion maps», *Proceedings of the National Academy of Sciences* (2009).
- ⁶S. L. Ronald R. Coifman, «Diffusion Maps», *Applied and Computational Harmonic Analysis* **21**, 5–30 (2006).
- ⁷W. H. S. J. v. d. W. J. de la Porte B. M. Herbst, «An Introduction to Diffusion Maps»,
- ⁸. Λουλάκης, *Στοχαστικές Διαδικασίες* (2015).
- ⁹K. J. Bathe, *Finite Element Methods, A Spectral Approach* (Springer, 2016).