



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΑΝΤΙΣΤΡΟΦΑ ΠΡΟΒΛΗΜΑΤΑ

ΑΝΑΚΑΤΑΣΚΕΥΗ ΕΙΚΟΝΑΣ ΜΕ ΟΛΟΚΛΗΡΩΤΙΚΕΣ ΕΞΙΣΩΣΕΙΣ FREDHOLM

ΠΡΩΤΟΥ ΕΙΔΟΥΣ

ΠΑΝΤΟΥ ΚΥΡΙΑΚΗ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:

ΔΡΟΣΟΣ ΓΚΙΝΤΙΔΗΣ

ΑΝ.ΚΑΘΗΓΗΤΗΣ Ε.Μ.Π

ΑΘΗΝΑ, ΜΑΙΟΣ 2017



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΑΝΤΙΣΤΡΟΦΑ ΠΡΟΒΛΗΜΑΤΑ

ΑΝΑΚΑΤΑΣΚΕΥΗ ΕΙΚΟΝΑΣ ΜΕ ΟΛΟΚΛΗΡΩΤΙΚΕΣ ΕΞΙΣΩΣΕΙΣ FREDHOLM

ΠΡΩΤΟΥ ΕΙΔΟΥΣ

ΠΑΝΤΟΥ ΚΥΡΙΑΚΗ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:

ΔΡΟΣΟΣ ΓΚΙΝΤΙΔΗΣ

ΑΝ.ΚΑΘΗΓΗΤΗΣ Ε.Μ.Π

Εγκρίθηκε από την τριμελή επιτροπή:

κ. Δρόσος Γκιντίδης

κ. Κυριάκη Κυριακή

κ. Χαραλαμπόπουλος Αντώνιος

Αν. Καθηγητής Ε.Μ.Π

Καθηγήτρια Ε.Μ.Π

Καθηγητής Ε.Μ.Π

ΑΘΗΝΑ, ΜΑΙΟΣ 2017

Ευχαριστίες

Με την ολοκλήρωση της διπλωματικής μου εργασίας, θα ήθελα να εκφράσω την ευγνωμοσύνη μου στην οικογένειά μου, που με στηρίζει σε κάθε μου βήμα. Ιδιαίτερες ευχαριστίες στον επιβλέποντα καθηγητή μου κ. Δρόσο Γκιντίδη, για τη συμπαράστασή του και την άρτια καθοδήγηση που μου προσέφερε, καθ' όλη τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας. Επίσης, θέλω να ευχαριστήσω ξεχωριστά, τον πατέρα μου Αλέξανδρο, τον αδερφό μου Σταμάτη και το φίλο μου Γιώργο, που στάθηκαν δίπλα μου και με βοήθησαν να αντιμετωπίσω κάθε δυσκολία που συνάντησα κατά τη διάρκεια των σπουδών μου στην Αθήνα. Τέλος, με ιδιαίτερη συγκίνηση, θα ήθελα να ευχαριστήσω για τις συμβουλές που μου παρείχε σε επιστημονικό αλλά και σε προσωπικό επίπεδο, τη μητέρα μου, Μαρία Πάντου, εκπαιδευτικό και απόφοιτο του Καποδιστριακού Πανεπιστημίου Αθηνών, η οποία ήταν συνοδοιπόρος σε όλο αυτό το ταξίδι των φοιτητικών μου χρόνων.

Περίληψη

Τα αντίστροφα προβλήματα αποτελούν τα θεμέλια για την επίλυση πολλών προβλημάτων, εκ των οποίων πολλά από αυτά συμπεριλαμβάνονται στην καθημερινότητά μας. Σε αυτή τη μελέτη ασχολούμαστε κυρίως με το αντίστροφο πρόβλημα το οποίο με τη βοήθεια της ολοκληρωτικής εξίσωσης Fredholm πρώτου είδους θα μας οδηγήσει στην ανάλυση και ανακατασκευή μιας εικόνας. Πέραν του θεωρητικού μέρους, θα δούμε μερικές εφαρμογές του αντίστροφου προβλήματος, στην ανάλυση εικόνας με τους κατάλληλους αλγόριθμους.

Στο 1^ο κεφάλαιο της διπλωματικής εργασίας, ορίζουμε το αντίστροφο πρόβλημα και τα μη καλά τοποθετημένα προβλήματα.

Στο 2^ο κεφάλαιο της εργασίας, εστιάζουμε την προσοχή μας στην ολοκληρωτική εξίσωση Fredholm Πρώτου Είδους για τη λύση των αντίστροφων προβλημάτων, καθώς με τη βοήθεια του λήμματος Riemann-Lebesgue αναφέρουμε ένα παράδειγμα από τη Γεωφυσική στο οποίο εφαρμόζουμε την εξίσωση Fredholm.

Στο 3^ο κεφάλαιο της εργασίας, αναλύουμε δύο τρόπους διακριτοποίησης της ολοκληρωτικής εξίσωσης Fredholm, δηλαδή τη μέθοδο αριθμητικής προσέγγισης των ολοκληρωμάτων και τη μέθοδο του αναπτύγματος Galerkin. Ακολούθως, γίνεται σύγκριση μεταξύ των δύο μεθόδων, όπου καταγράφονται τα αρνητικά και τα θετικά τους. Τέλος, αναλύουμε το ρόλο των μεθόδων απόδοσης ιδιαζουσών τιμών SVE (Singular Value Evaluation) και ανάλυσης ιδιαζουσών τιμών SVD (Singular Value Decomposition) στα αντίστροφα προβλήματα, καθώς και τη συμβολή τους στη συνθήκη Picard.

Στο 4^ο κεφάλαιο, ορίζουμε και αναλύουμε τους τρεις τρόπους ομαλοποίησης των διακριτών συστημάτων και δίνουμε περισσότερη έμφαση στον τρόπο επίλυσης με την καμπύλη τύπου L .

Στο 5^ο κεφάλαιο ασχολούμαστε με τις υπολογιστικές μεθόδους της ομαλοποίησης των διακριτών συστημάτων και τις στάσιμες επαναληπτικές διαδικασίες.

Στο 6^ο κεφάλαιο, αναφερόμαστε στο κυρίως μέρος της εργασίας, το οποίο είναι η «αποθόλωση» της εικόνας. Βασικά εξηγούμε τη μέθοδο με την οποία θα οδηγηθούμε στην ανάλυση της εικόνας ,χρησιμοποιώντας τον αλγόριθμο CGLS , καθώς επίσης θα μελετήσουμε και τη τομογραφία σε δύο διαστάσεις.

Τέλος, στο 7^ο κεφάλαιο, εκτελώντας τις κατάλληλες εντολές στο MatLab, θα εφαρμόσουμε τον CGLS αλγόριθμο, ώστε να ανακατασκευάσουμε την εικόνα που θα επεξεργαστούμε.

Λέξεις-Κλειδιά:

Αντίστροφα προβλήματα, ολοκληρωτική εξίσωση Fredholm Πρώτου Είδους, μη καλά τοποθετημένα προβλήματα, λήμμα Riemann-Lebesque, μέθοδοι διακριτοποίησης, μέθοδοι ομαλοποίησης, συνθήκη Picard, απόδοση ιδιάζουσων τιμών(SVE), παραγοντοποίηση ιδιάζουσων τιμών(SVD), καμπύλη τύπου L, στάσιμες επαναληπτικές διαδικασίες, ανακατασκευή εικόνας.

Abstract

Inverse problems are the foundation for solving many other problems, many of which are included in our everyday life. In this study we mainly deal with the inverse problem which, with the help of the Fredholm 1st integral equation, will lead us to the analysis and reconstruction of an image. Beyond the theoretical part, we will see some applications of the inverse problem in image analysis with the appropriate algorithms.

In the 1st part of the diploma thesis, we define the inverse problem and the ill posed problems.

In the 2nd part of the paper, we focus on the complete Fredholm equation which is the first kind to solve the inverse problems, with the help of the Riemann-Lebesgue entry we mention an example from the geophysics applied by the Fredholm equation.

In the 3rd part of the paper, we analyse two ways of distinguishing the integral Fredholm equation, namely the quadrature method and the method of expansion. Finally, we analyse the role of the two most important methods, Singular Value Evaluation (SVE) and Singular Value Decomposition (SVD) on the inverse problems, as well as the contribution of the Picard Treaty.

In the 4th part, we define and analyse the three ways of normalizing discrete systems and put more emphasis on how to solve the L-Curve curve.

In the 5th part, we deal with the computational methods of normalization of discrete systems.

In the 6th part of the study, we refer to the main part of the work, namely the clarification of an image, in which the CGLS algorithm will lead to the implementation of the image analysis, as we will see the example of tomography in two dimensions.

Finally, in the 7th part, by executing the appropriate commands in MatLab, we will apply a specific algorithm image to reconstruct it.

Keywords:

Inverse Problems, Fredholm 1st integral equation, Incorrectly Fixed Problems, Riemann-Lebesgue Entry, Discrete Methods, Normalization Methods, Picard Condition, Singular Value Evaluation (SVE Method), Singular Value Decomposition (SVD Method), L-Curve , image reconstruction.

Πίνακας περιεχομένων

Κεφάλαιο 1 Εισαγωγή	9
1.Η έννοια των αντίστροφων προβλημάτων	9
1.1 Μη καλά τοποθετημένα προβλήματα	11
1.1.1 Πρόβλημα χωρίς λύση	11
1.1.2 Πρόβλημα στο οποίο υπάρχουν περισσότερες από μια λύσεις	12
1.1.3 Πρόβλημα όπου η λύση δεν εξαρτάται με συνεχή τρόπο από τα δεδομένα	12
Κεφάλαιο 2	14
2.Ολοκληρωτικές Εξισώσεις Fredholm	14
2.1 Ολοκληρωτικές Εξισώσεις Fredholm Πρώτου Είδους	14
2.2 Λήμμα Riemann-Lebesgue	15
2.3 Παράδειγμα ενός προβλήματος από τη γεωφυσική	15
2.3.1 Διατύπωση του προβλήματος	15
Κεφάλαιο 3	19
3.Μέθοδοι διακριτοποίησης	19
3.1 Μέθοδος αριθμητικής ολοκλήρωσης	19
3.2 Μέθοδος αναπτύγματος Galerkin(expansion method)	21
3.3 Σύγκριση Μεθόδων και Κριτήρια Επιλογής	24
3.4 Μέθοδος απόδοσης ιδιάζουσων τιμών (SVE)	25
3.4.1 Ο ρόλος της μεθόδου απόδοσης ιδιάζουσων τιμών SVE	26
3.5 Σύγκλιση και απόκλιση της μεθόδου απόδοσης ιδιάζουσων τιμών SVE	27
3.6 Συνθήκη Picard	29
3.6.1 Μη μοναδικότητα της λύσης	31
3.7. Παραγοντοποίηση ιδιάζουσων ιδιοτιμών (SVD)	33
3.8 Ο ρόλος της μεθόδου παραγοντοποίησης ιδιάζουσων ιδιοτιμών (SVD)	35
3.8.1 Συνθήκη Picard στην διακριτή περίπτωση σε συνδυασμό με την SVD	36
Κεφάλαιο 4	37
4. Μέθοδοι Ομαλοποίησης	37
4.1 Μέθοδος Αποκοπής SVD (Truncated SVD)	39
4.2 Επιλεκτική μέθοδος SVD (Selective SVD)	41
4.3 Μέθοδος Tikhonov SVD	41
4.4 Μέθοδος της καμπυλης L (L-curve)	44

Κεφάλαιο 5	48
5.Επαναληπτικές Μέθοδοι Ομαλοποίησης	48
5.1 Στάσιμες επαναληπτικές διαδικασίες ομαλοποίησης	49
5.1.1 Επαναληπτική Μέθοδος Landweber	49
5.1.2 Επαναληπτική Μέθοδος Cimmino	50
5.2 Μέθοδοι προβολής	51
5.3 Μέθοδοι Υπόχωρου Krylov	52
Κεφάλαιο 6	54
6. Αποθόλωση Εικόνας	54
6.1 Ο ρόλος της Συνάρτησης Διασποράς Σημείου	54
6.2 Πίνακες Πρώτου Βαθμού PSF και Γρήγοροι Αλγόριθμοι	56
6.3 Αποσυνέλιξη και Συνέλιξη	58
6.4 Τομογραφία στις 2 Διαστάσεις(2D)	58
6.5 Αλγόριθμος CGLS (Conjugate Gradient Least Squares)	61
Κεφάλαιο 7	62
7. Εφαρμογή του αλγορίθμου CGLS	62
7.1 Συνοπτική ανάλυση αποτελεσμάτων	69
Επίλογος	71
Βιβλιογραφία	72

Κεφάλαιο 1

Εισαγωγή

1. Η έννοια των αντίστροφων προβλημάτων

Στην καθημερινότητά μας, εμφανίζονται πολλά και ποικίλα προβλήματα, που μας απασχολούν είτε στον επιστημονικό χώρο, είτε στον προσωπικό και κοινωνικό μας χώρο. Τόσο η διατύπωση όσο και η αντιμετώπιση ενός προβλήματος αποτελούν διαδικασίες οι οποίες για να μας οδηγήσουν στην επιθυμητή λύση, πρέπει να υποστούν σωστές επεξεργασίες. Με τα κατάλληλα εργαλεία όπως θα δούμε αναλυτικά στην παρούσα διπλωματική εργασία, όσο μη προφανές και αν φαίνεται, θα διαπιστώσουμε ότι υπάρχει μια μεγάλη γκάμα προβλημάτων, η οποία λύνεται με την έννοια του «αντίστροφου προβλήματος».

Στο αντίστροφο πρόβλημα, συνήθως το ζητούμενο είναι η εκτίμηση παραμέτρων ενός φυσικού φαινομένου ή ενός μαθηματικού προβλήματος, όταν οι μετρήσεις μιας συνάρτησης ή κάποιου μεγέθους είναι γνωστές.

Στις Φυσικές Επιστήμες γνωρίζουμε βέβαια, ότι οι παράμετροι ενός φυσικού φαινομένου υπάρχουν. Άρα, μέσω της επίλυσης ενός αντίστροφου προβλήματος πρέπει να βρεθούν οι κατάλληλες εκτιμήσεις τους.

Εντούτοις, η λύση των αντίστροφων προβλημάτων δεν είναι πάντα εύκολο να προσδιοριστεί, διότι ανήκουν στην κατηγορία των μη καλά τοποθετημένων προβλημάτων και επιδέχονται από καμιά έως πολλές λύσεις.

Το αντικείμενο της μελέτης μας, ασχολείται με την ανακατασκευή μιας «θολής» εικόνας, βασισμένης στη θεωρία του αντίστροφου προβλήματος. Εφόσον καταφέρουμε να μοντελοποιήσουμε το πρόβλημά μας και να διαχωρίσουμε δεδομένα και παραμέτρους, μέσω κάποιου πυρήνα ο οποίος καθορίζεται από το κατάλληλο μαθηματικό μοντέλο της εικόνας, τότε θα μπορούμε να διατυπώσουμε μια ολοκληρωτική εξίσωση, αυτή της εξίσωσης Fredholm Πρώτου Είδους. Έπειτα από κατάλληλες διαδικασίες που θα ακολουθήσουμε, όπως διακριτοποίηση και ομαλοποίηση της ολοκληρωτικής εξίσωσης και με τη βοήθεια κατάλληλων αλγορίθμων, θα αναλύσουμε την εικόνα μας και θα την ανακατασκευάσουμε.

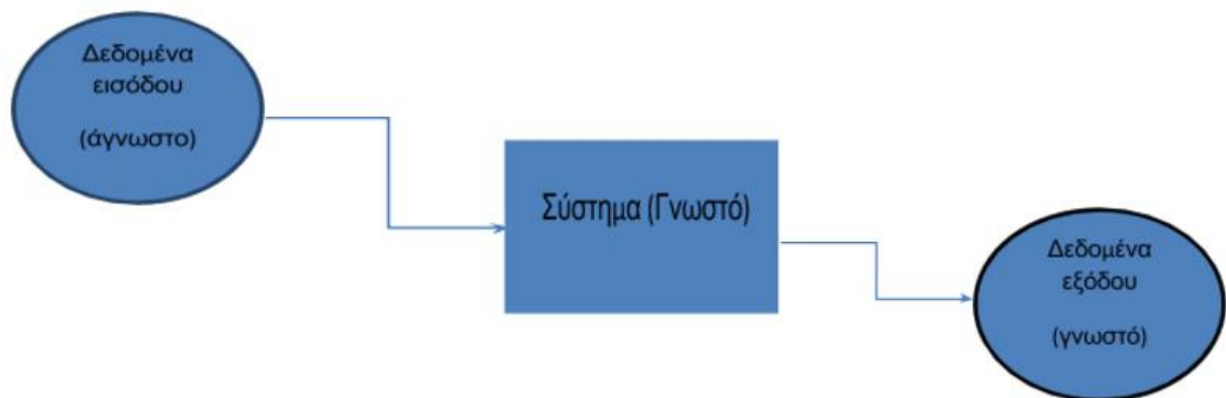
Υπάρχουν πολλά φυσικά προβλήματα τα οποία για να μπορούν να λυθούν, πρέπει να μοντελοποιηθούν μέσω ενός απλού γραμμικού συστήματος. Σε αυτές τις κατηγορίες προβλημάτων έχουμε ως αρχικό πλάνο ένα σύστημα στο οποίο εισέρχονται τα δεδομένα εισόδου και αφού υποστούν την κατάλληλη επεξεργασία εντός του συστήματος, λαμβάνονται τα δεδομένα εξόδου.

Επομένως, τα προβλήματα αυτά μπορούν να οριστούν με δύο τρόπους:

α) ως ένα Ευθύ Πρόβλημα το οποίο προϋποθέτει γνώση των δεδομένων εισόδου, ενώ τα δεδομένα εξόδου θεωρούνται άγνωστα.

β) ως ένα Αντίστροφο Πρόβλημα όπου γνωρίζουμε το σύστημα και τα δεδομένα εξόδου, και πρέπει να ανακατασκευάσουμε τα δεδομένα εισόδου. Μετέπειτα θα δούμε ότι τα αντίστροφα προβλήματα είναι μη καλά τοποθετημένα.

Μερικά παραδείγματα τα οποία ακολουθούν τη φιλοσοφία των αντίστροφων προβλημάτων είναι: α) η ανακατασκευή μιας θολής εικόνας σε μια ευκρινέστερη, χρησιμοποιώντας τις ολοκληρωτικές εξισώσεις Fredholm πρώτου είδους, β) η μέτρηση του μαγνητικού πεδίου στη επιφάνεια της Γης γ) η ανακατασκευή της μαγνήτισης στο εσωτερικό του ηφαιστείου, δ) ο υπολογισμός των μαζών και των σταθερών ελατηρίου από ένα σύστημα ταλάντωσης το οποίο αποτελείται από N μάζες.



Εικόνα 1- Απεικόνιση Αντίστροφου Προβλήματος

1.1 Μη καλά τοποθετημένα προβλήματα

Ένα σημαντικό χαρακτηριστικό των αντίστροφων προβλημάτων, το οποίο δυσκολεύει τη λύση τους είναι ότι δεν είναι καλά τοποθετημένα.

Σύμφωνα με τον Hadamard ο οποίος ασχολείτο με τα μαθηματικά μοντέλα των φυσικών φαινομένων από τις αρχές του 20^{ου} αιώνα, ένα πρόβλημα είναι μη καλά τοποθετημένο αν :

- ✓ Το πρόβλημα δεν έχει λύση.
- ✓ Υπάρχουν περισσότερες από μια λύσεις.
- ✓ Η λύση δεν εξαρτάται με συνεχή τρόπο από τα δεδομένα του προβλήματος.

Στη συνέχεια παρουσιάζονται παραδείγματα μη καλά τοποθετημένων προβλημάτων, τα οποία υπολογίζονται αριθμητικά, με τη βοήθεια του υπολογιστικού μαθηματικού προγράμματος MatLab.

1.1.1 Πρόβλημα χωρίς λύση

Έστω ότι έχουμε δύο πίνακες:

$$A = \begin{pmatrix} 2 \\ 9 \end{pmatrix} \text{ και } b = \begin{pmatrix} 2 \\ 10 \end{pmatrix}$$

και θέλουμε να βρούμε τη λύση του γραμμικού συστήματος $Ax = b$. Παρατηρούμε ότι $2x = 2$ και $9x = 10$, δηλαδή το σύστημα δεν έχει λύση.

Αυτό μπορεί να διορθωθεί με τη μέθοδο των ελαχίστων τετραγώνων έτσι ώστε να μπορέσουμε να βρούμε ένα μοναδικό x , οπότε ελαχιστοποιώντας την παράσταση:

$$\min_x \|(2 - 9)x - (2 - 10)\|_2^2 = \min_x ((2x - 2)^2 + (9x - 10)^2)$$

Βρίσκουμε τη γενικευμένη λύση $x = \mathbf{1.1059}$.

1.1.2 Πρόβλημα στο οποίο υπάρχουν περισσότερες από μια λύσεις

Ένα απλό παράδειγμα, το οποίο εύκολα παρατηρούμε ότι είναι κακώς τοποθετημένο, είναι το σύστημα $x_1 + x_2 = 5$. Το σύστημα αυτό έχει άπειρες λύσεις. Για να έχουμε μοναδική λύση, τροποποιούμε το πρόβλημά μας προσθέτοντας την επιπλέον απαίτηση ελαχιστοποίησης της L_2 νόρμας $\|x\|^2 = (x_1^2 + x_2^2)^{\frac{1}{2}}$, οπότε βρίσκουμε την μοναδική λύση $x_1 = x_2 = 2.5$.

1.1.3 Πρόβλημα όπου η λύση δεν εξαρτάται με συνεχή τρόπο από τα δεδομένα

Η συνθήκη της ευστάθειας, είναι πολύ πιο δύσκολο να αντιμετωπιστεί, επειδή η παραβίαση της, υπονοεί ότι αυθαίρετα μικρές διαταραχές των δεδομένων, μπορούν να προκαλέσουν αυθαίρετα μεγάλες διαταραχές της λύσης. Τουλάχιστον, αυτό ισχύει για τη συνεχή περίπτωση. Βέβαια, για το διακριτό πρόβλημα οι διαταραχές είναι πάντοτε πεπερασμένες, αλλά αυτό εξακολουθεί να είναι προβληματικό, εφόσον αυτές είναι πολύ μεγάλες.

Το κλειδί είναι να αναδιατυπώσουμε το πρόβλημα έτσι ώστε η λύση στο καινούργιο μας πρόβλημα να είναι λιγότερο ευαίσθητη στις διαταραχές. Τότε λέμε ότι σταθεροποιούμε ή ομαλοποιούμε το πρόβλημα, ώστε η λύση του να γίνει πιο σταθερή ή πιο ομαλή.

Έστω το γραμμικό πρόβλημα $Ax = b$, όπου

$$A = \begin{pmatrix} 4 & 5 & 3 & 1 \\ 1 & 6 & 1 & 0 \end{pmatrix} \text{ και } b = \begin{pmatrix} 19.249 \\ 6.843 \end{pmatrix}$$

Ο δείκτης κατάστασης ενός πίνακα A ορίζεται ως:

$$k(A) = \|A^{-1}\| \|A\| \geq \|A^{-1}A\| = 1$$

και παίρνει τιμές μεγαλύτερες από τη μονάδα.

Τιμές κοντά στη μονάδα σημαίνουν πως ο πίνακας A είναι καλώς τοποθετημένος.

Σε αντίθετη περίπτωση, ο πίνακας είναι μη καλώς τοποθετημένος και αυτό σημαίνει πως η επίλυση ενός γραμμικού συστήματος αυτού του πίνακα, ενδέχεται να δώσει μεγάλα σφάλματα.

Ο πίνακας A του παραδείγματός μας έχει δείκτη κατάστασης ίσο με $k(A) = 3363$, δηλαδή αρκετά υψηλό. Αποτέλεσμα αυτού είναι, μικρές διαταραχές στα δεδομένα του αρχικού προβλήματος, να προκαλούν μεγάλες διαταραχές στη λύση του. Πράγματι, παρατηρούμε ότι για $b = \begin{pmatrix} 20 \\ 7 \end{pmatrix}$, $x' = \begin{pmatrix} -30 \\ 50 \end{pmatrix}$.

Δηλαδή για μεταβολή του b μικρότερη από 3% έχουμε αλλαγή στη λύση της τάξης του 750% και 10000% για τις δύο συντεταγμένες του x αντίστοιχα.

Παρ' όλα αυτά, μπορούμε να τροποποιήσουμε το παραπάνω πρόβλημα ώστε η λύση του να είναι πιο σταθερή, δηλαδή λιγότερο ευαίσθητη σε διαταραχές.

Για παράδειγμα, μπορούμε να επιβάλουμε ένα ανώτατο όριο δ για τη νόρμα της λύσης και στη νέα διατύπωση να αντιμετωπίσουμε το τροποποιημένο πρόβλημα ελαχιστοποίησης:

$$\min_x \|Ax - b\|_2 \text{ με } \|x\|_2 < \delta$$

όπου πλέον η λύση εξαρτάται μοναδικά, αλλά όχι γραμμικά από το δ .

Η λύση $x_{1.37}$ ($\delta = 1.37$) είναι αρκετά κοντά στην ακριβή λύση. Με την κατάλληλη πληροφορία, μπορούμε να υπολογίσουμε προσεγγιστικά τη λύση. Βασική δυσκολία είναι με πιο κριτήριο θα επιλέξουμε την παράμετρο δ , όταν δεν γνωρίζουμε την ακριβή λύση.

Έτσι, όταν λύνουμε ένα αντίστροφο πρόβλημα αριθμητικά στον υπολογιστή, αντιμετωπίζουμε δυσκολίες παρόμοιες με τις πιο πάνω, διότι τα προβλήματα που υπολογίζονται αριθμητικά, είναι μη καλά τοποθετημένα.

Οι μέθοδοι ομαλοποίησης είναι η λύση όλων αυτών των προβλημάτων. Στη συνέχεια θα αναπτύξουμε αυτές τις μεθόδους.

Κεφάλαιο 2

2. Ολοκληρωτικές Εξισώσεις Fredholm

Σε αυτή την ενότητα θα ασχοληθούμε με μια σημαντική κατηγορία των αντίστροφων γραμμικών προβλημάτων και συγκεκριμένα εκείνων που λαμβάνουν τη μορφή των ολοκληρωτικών εξισώσεων τύπου Fredholm. Αυτά τα προβλήματα ανακύπτουν σε πολλές εφαρμογές στην επιστήμη και την τεχνολογία, όπως για παράδειγμα στην ιατρική απεικόνιση και την γεωλογική έρευνα. Βασική λειτουργία αυτής της ολοκληρωτικής εξίσωσης, είναι να περιγράψει τη σχέση μεταξύ του πυρήνα, των μετασχηματισμένων δεδομένων και των παρατηρήσεων.

2.1 Ολοκληρωτικές Εξισώσεις Fredholm Πρώτου Είδους

Η ολοκληρωτική εξίσωση Fredholm Πρώτου Είδους διατυπώνεται μαθηματικά ως εξής:

$$\int_{\alpha}^{\beta} K(s,t) f(t) dt = g(s), \quad \alpha \leq s \leq \beta \quad (2.1)$$

όπου η συνάρτηση $k(s,t)$ είναι γνωστή ως πυρήνας. Σε αυτή την περίπτωση τόσο ο πυρήνας k όσο και η $g(s)$ είναι γνωστές συναρτήσεις, ενώ η $f(t)$ είναι άγνωστη. Η ολοκληρωτική εξίσωση Fredholm δημιουργεί μια γραμμική σχέση μεταξύ των συναρτήσεων $f(t)$ και $g(s)$ και ο πυρήνας $k(s,t)$ προσδιορίζει τον βαθμό συσχέτισης μεταξύ αυτών των δύο συναρτήσεων.

Όπως είναι λογικό, εάν οι συναρτήσεις f και k είναι γνωστές, τότε μπορούμε να προσδιορίσουμε την g υπολογίζοντας το παραπάνω ολοκλήρωμα. Το αντίστροφο πρόβλημα συνίσταται στην εύρεση της f δοθέντων των άλλων δύο συναρτήσεων.

Στην περίπτωση που ο πυρήνας έχει τη μορφή $k(s,t) = h(s-t)$, τότε η ολοκληρωτική εξίσωση παίρνει την πιο κάτω μορφή:

$$\int_{\alpha}^{\beta} h(s-t) f(t) dt = g(s), \quad 0 \leq s \leq 1 \quad (2.2)$$

και χαρακτηρίζεται ως πρόβλημα αποσυνέλιξης.

Όταν η εξίσωση Fredholm Πρώτου Είδους προκύπτει σε αντίστροφα προβλήματα, συνήθως ο πυρήνας K και η συνάρτηση $g(s)$ είναι γνωστά, ενώ το f άγνωστο. Το συγκεκριμένο πρόβλημα λύνεται αριθμητικά με μεθόδους ομαλοποίησης.

Οι Riemann-Lebesgue διατύπωσαν μία από τις σημαντικότερες ιδιότητες της ολοκληρωτικής εξίσωσης Fredholm Πρώτου Είδους, η οποία είναι γνωστή και ως Λήμμα Riemann-Lebesgue.

2.2 Λήμμα Riemann-Lebesgue

Σύμφωνα με το Λήμμα Riemann-Lebesgue για συνάρτηση f_p ως: $f_p(t) = \sin(2\pi p t)$, $p=1,2,\dots$ και για οποιαδήποτε συνάρτηση πυρήνα $k(s,t)$ ισχύει ότι:

$$g_p(s) = \int_0^1 k(s,t) f_p(t) dt \rightarrow 0, \text{ για } p \rightarrow \infty \quad (2.3)$$

Αυτό σημαίνει ότι με την αύξηση της συχνότητας της f_p το εύρος ταλάντωσης της g_p ελαττώνεται.

Η ολοκλήρωση «ομαλοποιεί» τις υψηλές συχνότητες καθιστώντας τη συνάρτηση g «ομαλότερη» σε σχέση με την f .

2.3 Παράδειγμα ενός προβλήματος από τη γεωφυσική

Για να γίνει πιο εύκολα κατανοητή η θεωρία της ολοκληρωτικής εξίσωσης Fredholm, θα ασχοληθούμε με ένα απλό μοντέλο αντίστροφου προβλήματος, όπως το πρόβλημα της βαρυτικής χαρτογράφησης σε απλοποιημένη μορφή, που ακολουθεί.

2.3.1 Διατύπωση του προβλήματος

Έστω κατανομή μάζας με πυκνότητα μάζας $f(t)$, σε βάθος d από την επιφάνεια, με $t \in [0,1]$. Υποθέτουμε ότι δεν υπάρχει μάζα έξω από την πηγή, η οποία παράγει βαρυτικό πεδίο γύρω της. Μετρώντας το βαρυτικό πεδίο το οποίο δημιουργείται από την πηγή, προσπαθούμε να βρούμε την άγνωστη κατανομή μάζας. Πιο συγκεκριμένα, όπως φαίνεται στο πιο κάτω σχήμα (βλ. Εικόνα 2.3.1), στον άξονα s μετράμε την κατακόρυφη συνιστώσα του βαρυτικού πεδίου η οποία αναφέρεται ως μια συνάρτηση $g(s)$ ορισμένη από το 0 μέχρι και το 1.

Το πεδίο βαρύτητας που δημιουργείται από απειροελάχιστα τμήματα μάζας $f(t)$ μήκους dt , είναι ισοδύναμο με αυτό που δημιουργείται από σημειακές μάζες $f(t)dt$.

Ως εκ τούτου, το μέγεθος του πεδίου βαρύτητας κατά τον άξονα s , είναι $\frac{f(t)dt}{r^2}$, όπου $r = \sqrt{d^2 + (s - t)^2}$ η απόσταση της πηγής στη θέση t από την απειροστή μάζα στη θέση s (βλ. Εικόνα 2.3.1).

Το βαρυτικό πεδίο κατευθύνεται από την απειροστή μάζα προς την πηγή όπου το μέγεθος $g(s)$ γράφεται ως :

$$dg = \frac{\sin\theta}{r^2} f(t)dt$$

όπου θ είναι η γωνία που σχηματίζεται από τον άξονα t και την ευθεία $f(t) - g(t)$.

Αντικαθιστώντας το $\sin\theta = \frac{d}{r}$ παίρνουμε την εξίσωση:

$$\frac{\sin\theta}{r^2} f(t)dt = \frac{d}{(d^2 + (s - t)^2)^{\frac{3}{2}}} f(t)dt$$

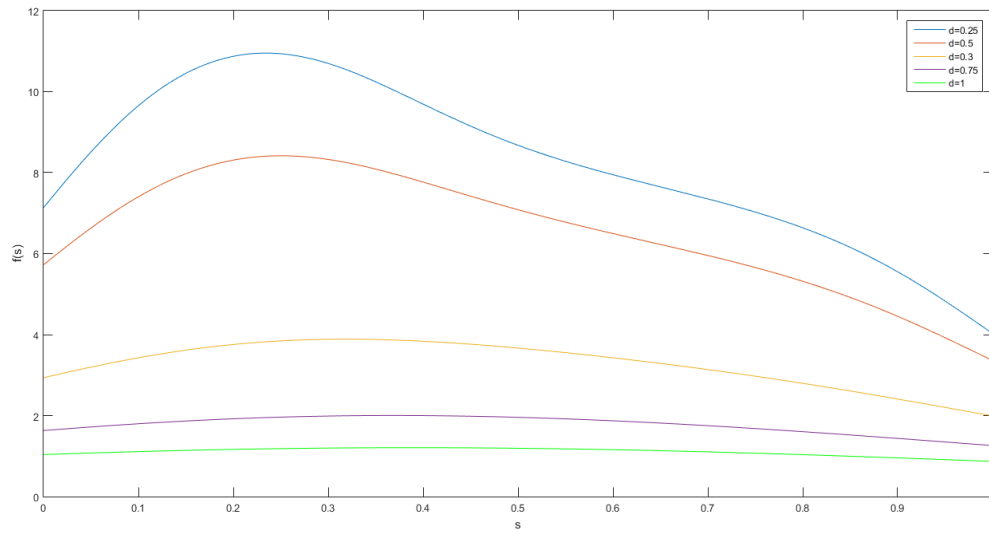
Συνεπώς, η συνολική ποσότητα δίνεται από το ολοκλήρωμα:

$$\int_0^1 k(s, t) f(t) dt = g(s), \quad 0 \leq s \leq 1$$

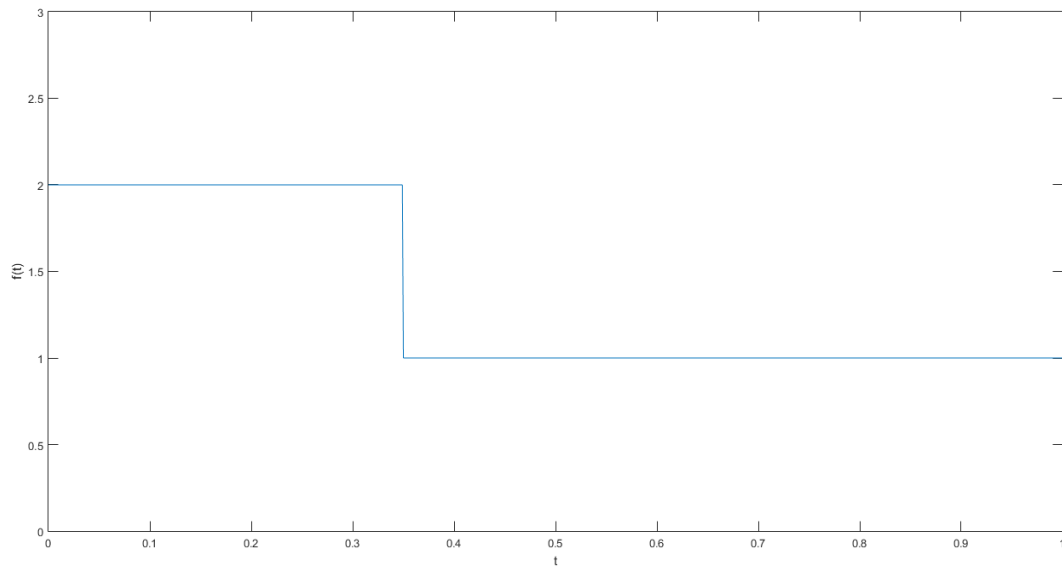
$$\text{όπου } k(s, t) = \frac{d}{(d^2 + (s - t)^2)^{\frac{3}{2}}}$$

Συμπερασματικά, μπορούμε να πούμε ότι, η συνάρτηση $k(s, t)$ είναι η κάθετη συνιστώσα της βαρυτικής δύναμης, μετρημένης στο s από μια απειροστή μάζα στο t και το βαρυτικό πεδίο $g(s)$ είναι μετρήσιμο μέγεθος.

Στην Εικόνα 2.3.1 βλέπουμε τις γραφικές παραστάσεις του $g(s)$ με δεδομένη την κατανομή μάζας $f(t)$ σε πέντε διαφορετικές τιμές της απόστασης d . Προφανώς όταν το d είναι μεγαλύτερο τότε πιο δύσκολα παίρνω πληροφορία για το $g(s)$. Επιπλέον, παρατηρούμε ότι στο $g(s)$ τα δεδομένα κατανέμονται με συνεχή τρόπο, ενώ τα δεδομένα της $f(t)$ είναι εμφανές ότι κατανέμονται με μη συνεχή τρόπο.



Εικόνα 2.3.1- Γραφική παράσταση του βαρυτικού πεδίου $g(s)$ για πέντε διαφορετικές τιμές της απόστασης d .



Εικόνα 2.3.2-Γραφική παράσταση της συνάρτησης $f(t)$

Στο παραπάνω παράδειγμα, παρατηρούμε ότι στην χαρτογράφηση της f με την g , όταν κατευθυνόμαστε από την πηγή f προς το βαρυτικό πεδίο $g(s)$, οι συνιστώσες της f με υψηλές συχνότητες αποσβένουν, σε αντίθεση με τις χαμηλότερες συχνότητες. Συνεπώς, όταν η f ολοκληρώνεται με την συνάρτηση k , η k έχει ομαλή επίδραση στην συνάρτηση f , έτσι ώστε η g να παρουσιάζεται ομαλότερη.

Η παραπάνω ερμηνεία εκφράζεται μαθηματικά με το λήμμα Riemann-Lebesgue (βλ. Σχέση 2.3). Καθώς η συχνότητα του f αυξάνεται, όταν ο ακέραιος αριθμός p παίρνει τιμές, το πλάτος g_{xp} μειώνεται. Δηλαδή, για ψηλές συχνότητες οι συνιστώσες του f αποσβένουν άρα το g θα είναι πιο ομαλό από το f .

Σαφώς, για μια μικρή τυχαία διατάραξη του g μπορεί να προκληθεί μεγάλη διαταραχή στην f , αν δεν έχει μια συνιστώσα ψηλής συχνότητας.

Η ομαλοποίηση που λαμβάνει χώρα στο προηγούμενο παράδειγμα, είναι ένα γενικότερο φαινόμενο στις ολοκληρωτικές εξισώσεις. Παρατηρούμε ότι όσο μεγαλύτερη συνιστώσα έχουμε για το f , τόσο πιο μεγάλη απόσβεση έχουμε, όσο μειώνεται η συχνότητα.

Κεφάλαιο 3

3. Μέθοδοι διακριτοποίησης

Μέχρι τώρα έχουμε μελετήσει την ολοκληρωτική μορφή Fredholm Πρώτου Είδους και αναφέραμε κάποιες θεωρητικές πτυχές. Αυτό όμως δεν είναι αρκετό, διότι για να καταφέρουμε να λύσουμε την εξίσωση Fredholm πρέπει να την τροποποιήσουμε με τέτοιο τρόπο ώστε να μπορεί να υπολογιστεί αριθμητικά. Ο λόγος που καταφεύγουμε σε άλλους τρόπους επίλυσης είναι ότι τα συνεχή προβλήματα είναι συνήθως γνωστά για την πολυπλοκότητά τους και την ακρίβειά τους, έτσι ώστε να μη μπορούν να λυθούν αναλυτικά.

Ως εκ τούτου, σε αυτό το κεφάλαιο θα ασχοληθούμε με δύο βασικές μεθόδους διακριτοποίησης. Η διακριτοποίηση μετατρέπει ένα συνεχές πρόβλημα σε διακριτό, το οποίο είναι εκφρασμένο σε γραμμικές σχέσεις πινάκων. Αυτές τις γραμμικές σχέσεις μπορούμε να τις λύσουμε αριθμητικά.

Θα εξετάσουμε επιπλέον την μέθοδο SVD δηλαδή την Παραγοντοποίηση Ιδιάζουσων Τιμών για την αντιμετώπιση των αντίστροφων διακριτών προβλημάτων που προκύπτουν από τη διακριτοποίηση της ολοκληρωτικής εξίσωσης Fredholm πρώτου είδους.

Για να μετατρέψουμε την ολοκληρωτική εξίσωση Fredholm σε γραμμική σχέση χρησιμοποιούμε μεθόδους διακριτοποίησης. Υπάρχουν αρκετές μέθοδοι οι οποίες διακριτοποιούν το σύστημα, αλλά θα αναλύσουμε τις εξής:

1. τη μέθοδο αριθμητικής ολοκλήρωσης.
2. τη μέθοδο του αναπτύγματος.

3.1 Μέθοδος Αριθμητικής Ολοκλήρωσης

Η διακριτοποίηση βασίζεται στον γενικό κανόνα προσέγγισης του ολοκληρώματος:

$$\int_0^1 \varphi(t) dt = \sum_{j=1}^n \omega_j \varphi(t_j) + e_n \quad (3.1.1)$$

όπου φ η συνάρτηση που έχουμε για ολοκλήρωση, $t_j, j = 1, 2, \dots, n$ επιλεγμένα σημεία από μηδέν έως ένα, ω_j το βάρος της συνεισφοράς του κάθε σημείου στην προσέγγιση του ολοκληρώματος και e_n το σφάλμα της προσέγγισης λόγω διακριτοποίησης. Με τον κανόνα της μέσης τιμής για παράδειγμα έχουμε

$$t_j = \frac{j-\frac{1}{2}}{n}, \quad \omega_j = \frac{1}{n} \quad j = 1, 2, \dots, n$$

Βάσει του παραπάνω κανόνα προσέγγισης του ολοκληρώματος, το ολοκλήρωμα

$$\psi(s) = \int_0^1 k(s, t) f(t) dt$$

προσεγγίζεται από το παρακάτω άθροισμα

$$\psi(s) = \int_0^1 k(s, t) f(t) dt = \sum_{j=1}^n \omega_j k(s, t_j) f(t_j) + e_n(s) \quad (3.1.2)$$

Στη συνέχεια απαιτούμε $\psi(s_i) = g(s_i), \quad i=1, 2, \dots, n$

Όπου $g(s_i)$ οι τιμές της συνάρτησης που έχουν μετρηθεί πειραματικά.

Έτσι οδηγούμαστε στο σύστημα εξισώσεων

$$\sum_{j=1}^n \omega_j k(s, t_j) f(t_j) = g(s_i) - e_n(s_i) \quad (3.1.3)$$

Αντικαθιστώντας τις πραγματικές τιμές της συνάρτησης f στα σημεία $t_j, j = 1, 2, \dots, n$ με τις αντίστοιχες προσεγγιστικές f_j και αμελώντας το σφάλμα $e_n(s)$ καταλήγουμε στο σύστημα:

$$\sum_{j=1}^n \omega_j k(s, t_j) f_j = g(s_i) \quad \text{για } i = 1, \dots, n \quad (3.1.4)$$

$$\begin{bmatrix} \omega_1 k(s_1, t_1) & \cdots & \omega_1 k(s_1, t_n) \\ \vdots & \ddots & \vdots \\ \omega_1 k(s_n, t_1) & \cdots & \omega_1 k(s_n, t_n) \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} g(s_1) \\ \vdots \\ g(s_n) \end{bmatrix} \quad (3.1.5)$$

που είναι ένα γραμμικό σύστημα $Ax = b$ με ,

$$a_{ij} = \omega_j k(s_i, t_j) \quad x_j = f(t_j) \quad b_i = g(s_i) \quad (3.1.6)$$

όπου $i, j = 1, \dots, n$

και άγνωστο το διάνυσμα $\chi^T = (f_1 \ f_2 \ \dots \ f_n)^T$ των προσεγγιστικών τιμών της συνάρτησης f .

3.2 Μέθοδος αναπτύγματος Galerkin (Expansion Method)

Για τη μέθοδο αυτή η διακριτοποίηση στηρίζεται στη χρήση προεπιλεγμένης βάσης συναρτήσεων και την αξιοποίηση της ιδιότητας της ορθογωνιότητας για την επίλυση των διπλών ολοκληρωμάτων που προκύπτουν. Ειδικότερα, η εν λόγω μέθοδος δομείται ως εξής:

Αρχικά γίνεται επιλογή δύο βάσεων συναρτήσεων με τη βοήθεια των οποίων θα επιχειρηθεί η προσέγγιση των f και g :

$$\Phi_n = \{\varphi_1 \dots \varphi_n\},$$

$$\Psi_n = \{\psi_1 \dots \psi_n\}.$$

Στη συνέχεια γίνεται προσέγγιση των f, g καθώς και του ολοκληρώματος που προκύπτει με τη βοήθεια των ανωτέρω συναρτησιακών βάσεων, δηλαδή:

$$f^{(n)}(t) = \sum_{j=1}^n \zeta_j \varphi_j(t) \quad (3.2.1)$$

$$g^{(n)}(s) = \sum_{j=1}^n \xi_j \psi_j(s)$$

ή

$$f(t) = f^{(n)}(t) + e_f(t), \quad f^{(n)} \in \text{span}\{\varphi_1, \dots, \varphi_n\} \quad (3.2.2)$$

$$g(s) = g^n(s) + e_g(s), \quad g^n \in \text{span}\{\psi_1, \dots, \psi_n\} \quad (3.2.3)$$

Δηλαδή, οι $f^n(t)$ και $g^n(s)$ μπορούν να θεωρηθούν προσεγγίσεις των f και g αντίστοιχα, ενώ τα $E_f(t)$ και $E_g(s)$ τα σχετικά σφάλματα τους.

Αν τώρα οριστεί ως $\theta(s)$ η ολοκληρωτική εξίσωση Fredholm πρώτου είδους, τότε

$$\theta(s) = \int_0^1 k(s,t)f^n(t) dt = \sum_{j=1}^n \zeta_j \int_0^1 k(s,t)\varphi_j(t) dt \quad (3.2.4)$$

Δεδομένου ότι η συνάρτηση θ στη γενική περίπτωση δεν ανήκει στο $\text{span}(\Psi)$, μπορεί να περιγραφεί μαθηματικά ως εξής:

$$\theta(s) = \theta^n(s) + e_\theta(s), \quad \text{και η } \theta^n \text{ να ανήκει μέσα στο διάστημα } \{\psi_1 \dots \psi_n\}.$$

Θέλουμε να εξασφαλίσουμε τη μοναδικότητα των προσεγγιστικών λύσεων. Έστω ότι οι θ^n και η g^n είναι οι ορθογώνιες προβολές των θ και g αντίστοιχα, μέσα στο διάστημα $\{\psi_1 \dots \psi_n\}$. Θα εξισώσουμε τις δύο προσεγγιστικές λύσεις $\theta^n(s)$ και η $g^n(s)$ και θα αντικαταστήσουμε την πρώτη με $\theta(s) - e_\theta(s)$, και την δεύτερη με $g(s) - e_g(s)$.

$$\theta^n(s) = g^n(s) \Leftrightarrow \theta(s) - e_\theta(s) = g(s) - e_g(s) \Leftrightarrow \theta(s) - g(s) = e_\theta(s) - e_g(s)$$

Άρα καταλήγουμε σε ένα αποτέλεσμα όπου η διαφορά των συναρτήσεων είναι ίση με την διαφορά των υπολοίπων. Δεδομένου ότι και οι δύο συναρτήσεις των σφαλμάτων είναι ορθογώνιες στο διάστημα $\psi_1 \dots \psi_n$ τότε και το υπόλοιπο $\theta(s) - g(s)$ είναι ορθογώνιο σε κάθε μια από τις συναρτήσεις ψ_1, \dots, ψ_n .

Προκειμένου να καταλήξουμε σε ένα σύστημα εξισώσεων, το οποίο θα υπολογίζει τους άγνωστους συντελεστές ζ_1, \dots, ζ_n , θα επιβάλλουμε την ορθογωνιότητα των υπολοίπων με τη συνάρτηση ψ_i . Δηλαδή απαιτούμε $\langle \psi_i, \theta - g \rangle = 0$, για $i=1, \dots, n$ και θα καταλήξουμε σε ένα $n \times n$ απλό γραμμικό σύστημα $Ax = b$. Με τις πιο κάτω ισοδυναμίες για $i = 1, \dots, n$ έχουμε ότι:

$$\langle \psi_i, \theta - g \rangle = 0 \Leftrightarrow \langle \psi_i, g \rangle - \langle \psi_i, \theta \rangle = \langle \psi_i, g \rangle = \langle \psi_i, \theta \rangle = \langle \psi_i, \int_0^1 k(s,t)f^n(t)dt \rangle$$

$$\Leftrightarrow \langle \Psi_i, g \rangle = \sum_{j=1}^n \zeta_j \langle \Psi_i, \int_0^1 k(s, t) \varphi_j(t) dt \rangle \quad (3.2.5)$$

Η τελευταία σχέση ουσιαστικά αναπαριστά ένα σύστημα της μορφής $Ax = b$ όπου:

$$a_{ij} = \int_0^1 \int_0^1 \Psi_i k(s, t) \varphi_j(t) ds dt$$

$$x_j = \zeta_j, \quad i, j = 1, 2, \dots, n$$

$$b_i = \int_0^1 \Psi_i(s) g(s) ds$$

Ιδιαίτερα σημαντικός είναι ο τρόπος με τον οποίο επιλέγονται οι βάσεις συναρτήσεων που χρησιμοποιούνται στην εν λόγω μέθοδο διακριτοποίησης. Για παράδειγμα, οι συναρτήσεις ψ_j είναι σημαντικό να επιλέγονται έτσι ώστε να συλλέγονται οι πιο σημαντικές πληροφορίες από τα μετρήσιμα δεδομένα $g(s)$.

Επίσης, η χρήση συναρτήσεων με επιθυμητές ιδιότητες, όπως για παράδειγμα ορθοκανονικών συναρτήσεων, απλοποιεί τους αριθμητικούς υπολογισμούς που απαιτούνται.

Τέλος, αν η συνάρτηση πυρήνα είναι συμμετρική, αυτό αντανακλάται και στη διακριτή διατύπωση του προβλήματος, δηλαδή στην περίπτωση αυτή και ο πίνακας του προκύπτοντος γραμμικού συστήματος είναι συμμετρικός.

3.3 Σύγκριση Μεθόδων και Κριτήρια Επιλογής

Οι μέθοδοι διακριτοποίησης που παρουσιάστηκαν στις προηγούμενες παραγράφους, έχουν πλεονεκτήματα και μειονεκτήματα. Τίθεται λοιπόν το ερώτημα ποια είναι η πλέον ενδεδειγμένη προς χρήση. Για να αποφασιστεί ποια είναι η καλύτερη κάθε φορά μέθοδος πρέπει πρώτα να απαντηθούν τα ακόλουθα ερωτήματα:

- Ποιο είναι το πρόβλημα στο οποίο ζητείται απάντηση και τι ακριβώς ζητείται;
- Ποια είναι τα δεδομένα που δίνονται για το πρόβλημα αυτό;
- Τι μπορεί να υπολογιστεί;

Σε κάθε περίπτωση, αν επιχειρηθεί σύγκριση των δύο μεθόδων θα πρέπει να καταγραφούν τα θετικά και αρνητικά τους. Το αποτέλεσμα της σύγκρισης καταγράφεται στον Πίνακα 3.3.1 που ακολουθεί.

Μέθοδος αριθμητικής ολοκλήρωσης	Μέθοδος Αναπτύγματος
(+) Απλή υπολογιστικά	(-) Εμπεριέχει υπολογισμό διπλού ολοκληρώματος, ο οποίος εν γένει είναι δύσκολος υπολογιστικά
(-) Προσέγγιση μόνο πάνω σε συγκεκριμένα σημεία του πλέγματος	(+) Προσέγγιση με τη βοήθεια συναρτήσεων
(+) Δυνατότητα εξειδίκευσης της μεθόδου, ώστε να αντανακλά τις ιδιότητες των συναρτήσεων K και f	(+) Δυνατότητα επιλογής κατάλληλων συναρτήσεων φ_j (+) Αν οι συναρτήσεις φ_j είναι ορθοκανονικές τότε υφίσταται σχέση μεταξύ των διατυπώσεων αναπτύγματος ιδιαζουσών τιμών και παραγοντοποίησης ιδιαζουσών τιμών
(+) Οι τιμές δειγμάτων που χρησιμοποιούνται για το δεξί μέλος του γραμμικού συστήματος είναι διακριτές	(+) Δύο εναλλακτικές για την επιλογή δειγμάτων: 1) $\psi_i(s) = \delta(s - s_i)$ και 2) $\psi_i(s) = \chi_i(s)$

Πίνακας 3.3.1 : Σύγκριση Μεθόδων Αριθμητικής Ολοκλήρωσης και Αναπτύγματος

3.4 Μέθοδος απόδοσης ιδιαζουσών τιμών (Singular Value Evaluation)

Η μέθοδος απόδοσης ιδιαζουσών τιμών SVE, είναι ένα μαθηματικό εργαλείο το οποίο μας βοηθάει να βρούμε μια ομαλή λύση στην ολοκληρωτική εξίσωση Fredholm Πρώτου Είδους. Αρχικά θα πρέπει να εισάγουμε ένα συμβολισμό. Έστω δύο συναρτήσεις φ και ψ που ορίζονται σε ένα διάστημα από το 0 μέχρι το 1. Το εσωτερικό τους γινόμενο ορίζεται ως :

$$\langle \varphi, \psi \rangle = \int_0^1 \varphi(t)\psi(t) dt \quad (3.4.1)$$

Η νόρμα-δύο της φ ορίζεται ως:

$$\|\varphi\|_2 = \langle \varphi, \varphi \rangle - \frac{1}{2} = \left(\int_0^1 \varphi(t)^2 dt \right) - \frac{1}{2} \quad (3.4.2)$$

Ο πυρήνας K της ολοκληρωτικής εξίσωσης είναι τετραγωνικά ολοκληρώσιμος αν το ολοκλήρωμα $\int_0^1 \int_0^1 K(s,t)^2 ds dt$ είναι ορισμένο και για κάθε τετραγωνικό ολοκληρώσιμο του πυρήνα K , η μέθοδος απόδοσης ιδιαζουσών τιμών παίρνει τη μορφή:

$$K(s,t) = \sum_{i=1}^{\infty} \mu_i u_i(s) v_i(t) \quad (3.4.3)$$

Στη μέθοδο απόδοσης ιδιαζουσών τιμών ισχύουν τα ακόλουθα:

- ✓ Οι μ_i είναι οι ιδιάζουσες τιμές και μπορούν να γραφούν σε φθίνουσα σειρά, $\mu_1 \geq \mu_2 \geq \dots \geq 0$.
- ✓ Η συνάρτηση u_i καλείται αριστερά ιδιάζουσα συνάρτηση.
- ✓ Η συνάρτηση v_i καλείται δεξιά ιδιάζουσα συνάρτηση.
- ✓ Όλες οι συναρτήσεις u_i και v_i είναι ορθογώνιες .
- ✓ Επίσης Ισχύει $\langle u_i, u_j \rangle = \langle v_i, v_j \rangle = \delta_{ij}$, για $i = 1, 2, \dots$
- ✓ Στην περίπτωση που βρίσκουμε πεπερασμένο αριθμό ιδιάζουσων ιδιοτιμών οι οποίες να είναι μη μηδενικές, τότε ο πυρήνας K , λέγεται εκφυλισμένος.

Η θεμελιώδης σχέση που ικανοποιεί, τις ιδιάζουσες τιμές και τις ιδιάζουσες συναρτήσεις, για $i=1,2,\dots$ είναι:

$$\int_0^1 K(s, t) v_i(t) dt = \mu_i u_i(s) \quad (3.4.4)$$

3.4.1 Ο ρόλος της μεθόδου απόδοσης ιδιαζουσών τιμών

Ο ρόλος του αναπτύγματος ιδιαζουσών τιμών SVE είναι ιδιαίτερα σημαντικός, διότι οι ιδιάζουσες τιμές μ_i φθίνουν προς το μηδέν. Μάλιστα, όσο πιο ομαλή (smooth) είναι η συνάρτηση πυρήνα, τόσο ταχύτερη είναι η σύγκλιση. Αν επιχειρηθεί η συσχέτιση της ύπαρξη λύσης συνάρτησης f , με την ομαλότητα των ιδιοτιμών μ_i πάνω σε δύο ορθοκανονικές βάσεις του L^2 , θα πρέπει καταρχάς να εκφραστούν οι f, g βάσει των v_i, u_i

$$f(t) = \sum_{i=1}^{\infty} \langle v_i, f \rangle v_i(t) \quad (3.4.1.1)$$

$$g(s) = \sum_{i=1}^{\infty} \langle u_i, g \rangle u_i(s) \quad (3.4.1.2)$$

Αν εισάγουμε το ανάπτυγμα f στην ολοκληρωτική εξίσωση και χρησιμοποιήσουμε την θεμελιώδη σχέση, τότε θα δούμε ότι η συνάρτηση g μπορεί να γραφτεί και ως:

$$g(s) = \sum_{i=1}^{\infty} \mu_i \langle v_i, f \rangle u_i(s) \quad (3.4.1.3)$$

Αν αντικατασταθούν οι συναρτήσεις f και g στην ολοκληρωτική εξίσωση, κάνοντας χρήση της θεμελιώδους σχέσης, τότε λαμβάνεται η ακόλουθη σχέση:

$$\sum_{i=1}^{\infty} \mu_i \langle v_i, f \rangle u_i(s) = \sum_{i=1}^{\infty} \langle u_i, g \rangle u_i(s) \quad (3.4.1.4)$$

Από την ανωτέρω σχέση προκύπτουν οι ακόλουθες δύο συνθήκες:

- Αν υπάρχει i τέτοιο ώστε $\mu_i = 0$ τότε η $f(t)$ έχει λύση αν και μόνο αν:

$$\langle u_i, g \rangle u_i(s) = 0$$

- Αν για κάθε i $\mu_i \neq 0$ τότε η $f(t)$ περιγράφεται από την ακόλουθη σχέση:

$$f(t) = \sum_{i=1}^{\infty} \frac{\langle u_i, g \rangle}{\mu_i} v_i(t) \quad (3.4.1.5)$$

Βάσει της ανωτέρω ανάλυσης προκύπτει και η Συνθήκη Picard, σύμφωνα με την οποία προκειμένου η συνάρτηση f να είναι τετραγωνικά ολοκληρώσιμη συνάρτηση, θα πρέπει η L_2 νόρμα της f να είναι φραγμένη. Θα πρέπει δηλαδή το $\langle u_i, g \rangle$ να συγκλίνει στο μηδέν ταχύτερα απ' ότι οι ιδιάζουσες τιμές μ_i .

3.5 Σύγκλιση και απόκλιση της μεθόδου απόδοσης ιδιαζουσών τιμών

Όπως έχουμε ήδη αναφέρει, για να μπορεί να γίνει η ανάλυση της SVE, πρέπει ο πυρήνας K να είναι τετραγωνικά ολοκληρώσιμος, δηλαδή να έχει την εξής μορφή:

$$\int_0^1 \int_0^1 K(x, t)^2 ds dt < \infty \quad (3.5.1)$$

Χρησιμοποιούμε τη μέθοδο ανάλυσης ιδιαζουσών τιμών SVD (Singular Value Decomposition), για να υπολογίσουμε προσεγγίσεις στη μέθοδο απόδοσης ιδιαζουσών τιμών SVE (Singular Value Evaluation). Αρχικά θα χρησιμοποιήσουμε τον μετασχηματισμό Laplace, ο οποίος είναι πολύ βασικός στον υπολογισμό της συνάρτησης f , όταν δίνεται από μια συνάρτηση g .

Ο μετασχηματισμός Laplace $g(s) = \int_0^{\infty} e^{-st} f(t) dt$ σε συνδυασμό με την

ολοκληρωτική εξίσωση Fredholm Πρώτου Είδους δίνει την εξίσωση:

$$g(s) = \int_0^{\infty} e^{-st} f(t) dt \quad (3.5.2)$$

Παρατηρούμε όμως, ότι δεν μπορεί να αναλυθεί με τη μέθοδο απόδοσης ιδιαιζουσών τιμών SVE, διότι ο πυρήνας δεν είναι τετραγωνικά ολοκληρώσιμος. Αυτό φαίνεται και από το ολοκλήρωμα:

$$\int_0^{\alpha} (e^{-st})^2 ds = \frac{1 - e^{-2ta}}{2t} \rightarrow \frac{1}{2t} \quad (3.5.3)$$

στο οποίο, αν ολοκληρώσουμε το t^{-1} στο διάστημα $[0, \infty)$, τότε τείνει στο άπειρο και αν ολοκληρώσουμε τετραγωνικά το $(e^{-st})^2$ στο ίδιο διάστημα, θα τείνει και αυτό στο άπειρο.

Ωστόσο, αν η $f(t)$ ορίζεται για $t \rightarrow \infty$, τότε το ολοκλήρωμα υπάρχει για όλα τα $s \geq 0$, και προτιμάμε μικρή αποκοπή σφάλματος για μεγάλους αριθμούς του α .

Αυτό ισχύει και όταν η συνάρτηση $g(s)$ φθίνει για $s \rightarrow \infty$, περιορίζοντας την στο διάστημα $[0, \alpha]$. Αυτό θα έχει ως αποτέλεσμα να έχει ένα τετραγωνικά ολοκληρώσιμο πυρήνα, λόγω του ότι το $k(s, t) = e^{-st}$ θα ορίζεται σε ένα πεπερασμένο διάστημα.

Μπορούμε να διακριτοποιήσουμε αυτή την ολοκληρωτική εξίσωση, χρησιμοποιώντας τις προσεγγίσεις $a_{ij} = hK\left(\left(i - \frac{1}{2}\right)h, \left(j - \frac{1}{2}\right)h\right)$, για $h = \frac{\alpha}{n}$.

Αν το α πάρει μικρότερη τιμή από το n και αν ο πυρήνας είναι τετραγωνικά ολοκληρώσιμος σε ένα πεπερασμένο διάστημα και επομένως θα υπάρχει και η ανάλυση SVE, τότε οι προσεγγίσεις $u_j^{(n)}$ συγκλίνουν προς τις ιδιάζουσες συναρτήσεις του πυρήνα. Οι προσεγγίσεις $u_j^{(n)}$ συγκλίνουν γρηγορότερα για μικρές τιμές του s .

Αυτές οι προσεγγίσεις ισχύουν και στην περίπτωση που σταθεροποιούμε το n , και μεγαλώνουμε το διάστημα, με τη διαφορά ότι, όσο αυξάνουμε το α , δεν βλέπουμε κάποια σύγκλιση. Είναι προφανές, διότι η SVE δεν αναλύεται όταν το α τείνει στο άπειρο. Ως εκ τούτου είναι λογικό να μην περιμένουμε σύγκλιση.

3.6 Συνθήκη Picard

Η λύση είναι τετραγωνικά ολοκληρώσιμη, αν:

$$\|f\|_2^2 = \int_0^1 f(t)^2 = \sum_{i=1}^{\infty} \langle v_i, f \rangle^2 = \sum_{i=1}^{\infty} \left(\frac{\langle u_i, g \rangle}{\mu_i} \right)^2 < \infty \quad (3.6.1)$$

Για να συγκλίνει η σειρά της συνθήκης Picard, πρέπει οι συντελεστές του δεξιού μέρους του εσωτερικού γινομένου $\langle u_i, g \rangle$, να φθίνουν γρηγορότερα στο μηδέν από τις ιδιάζουσες τιμές μ_i . Αυτό αναγνωρίζεται και ως συνθήκη ομαλοποίησης του δεξιού μέρους.

Το πιο βασικό πρόβλημα της ολοκληρωτικής εξίσωσης Fredholm πρώτου είδους, είναι ότι ακόμα και αν τα δεδομένα ικανοποιούν την συνθήκη Picard, οι μετρήσεις και ο θόρυβος παραβιάζουν την συνθήκη. Όταν δεν υπάρχει θόρυβος, η συνθήκη ικανοποιείται πλήρως.

Εντούτοις, όταν υπάρχει θόρυβος, η συμπεριφορά των ποσοτήτων του δεξιού μέρους $\langle u_i, g \rangle$ και $\frac{\langle u_i, g \rangle}{\mu_i}$ αλλάζει, διότι η συνθήκη Picard παραβιάζεται.

Συμπεραίνουμε ότι, η παραβίαση της συνθήκης Picard είναι μια γρήγορη εξήγηση της αστάθειας των γραμμικών αντίστροφων προβλημάτων, της ολοκληρωτικής εξίσωσης πρώτου είδους. Έτσι, το όφελος που αποκομίζουμε από την μέθοδο απόδοσης ιδιάζουσών τιμών SVE, είναι το πώς θα φιλτράρουμε το θόρυβο, ώστε να ικανοποιείται η συνθήκη Picard.

Παράδειγμα μη ύπαρξης λύσης:

Σε αυτό το σημείο βλέπουμε ένα παράδειγμα της ολοκληρωτικής εξίσωσης, το οποίο δεν έχει τετραγωνικά ολοκληρώσιμη λύση. Δηλαδή, δεν υπάρχει λύση που η L_2 νόρμα να ορίζεται. Έστω ότι για $0 \leq s \leq 1$ έχουμε την εξής εξίσωση:

$$\int_0^1 \frac{1}{s+t+1} f(t) dt = 1 \quad (3.6.2)$$

Άρα ο πυρήνας παίρνει την μορφή $k(s, t) = (s + t + 1)^{-1}$ και η συνάρτηση του δεξιού μέρους, είναι $g(s) = 1$.

Αν τώρα επεκτείνουμε τη δεξιό μέρος σε όρους της αριστερής ιδιάζουσας συνάρτησης, θα πάρουμε μια προσέγγιση της πιο κάτω μορφής. Καθώς το σφάλμα στην g_k μειώνεται όταν το k αυξάνεται, προκύπτει η πιο κάτω σχέση:

$$g_k(s) = \sum_{i=1}^k \langle u_i, g \rangle u_i(s) \quad (3.6.3)$$

$$\|g - g_k\|_2 \rightarrow 0 \text{ για } k \rightarrow \infty \quad (3.6.4)$$

θεωρούμε το ολοκλήρωμα:

$$\int_0^1 \frac{1}{s+t+1} f_k(t) dt = g_k(s) \quad (3.6.5)$$

με λύση:

$$f_k(t) = \sum_{i=1}^{\infty} \frac{\langle u_i, g \rangle}{\mu_i} v_i(t) \quad (3.6.6)$$

Σαφώς, η L_2 νόρμα της λύσης f_k ορίζεται για όλα τα k , αλλά σε κάποια σημεία φαίνεται ότι το πλάτος της f_k , αυξάνει με το k , δηλαδή $\|f_k\|_2 \rightarrow \infty$ και $k \rightarrow \infty$. Άρα αυτό είναι ένα σημάδι, το οποίο δείχνει ότι αυτές οι συναρτήσεις δεν συγκλίνουν σε μια τετραγωνικά ολοκληρώσιμη λύση.

Επισημάνσεις:

- ✓ Αν ο πυρήνας K είναι συνεχής, τότε για όλες τις εξισώσεις, που περιλαμβάνουν άπειρα αθροίσματα ιδιάζουσων συναρτήσεων του K , το σύμβολο της ισότητας, σημαίνει ότι το άθροισμα συγκλίνει ομοιόμορφα στην αριστερή πλευρά.
- ✓ Αν ο K είναι ασυνεχής τότε θα έχουμε σύγκλιση στο τετραγωνικό μέσο.

Παράδειγμα:

Έστω: $f_k(t) = \sum_{i=1}^{\infty} \frac{\langle u_i, g \rangle}{\mu_i} v_i(t)$, τότε θα λέμε ότι συγκλίνει ομοιόμορφα αν για κάθε ε

θετικό υπάρχει ακέραιος αριθμός $n \in \mathbb{N}: \forall k > n$ έτσι ώστε να έχουμε:

$$\left| f(t) - \sum_{i=1}^{\infty} \frac{\langle u_i, g \rangle}{\mu_i} v_i(t) \right| < \varepsilon \quad (3.6.7)$$

Για τη σύγκλιση στο μέσο του τετραγώνου, η ανισότητα παίρνει την μορφή:

$$\left\| f(t) - \sum_{i=1}^{\infty} \frac{\langle u_i, g \rangle}{\mu_i} v_i(t) \right\|_2 < \varepsilon \quad (3.6.8)$$

Το ανάπτυγμα μιας ορθοκανονικής βάσης παίρνει την μορφή:

$$f(t) = \sum_{i=1}^{\infty} \langle \varphi_i, f \rangle \varphi_i(t) \quad (3.6.9)$$

Μια βασική μέθοδος στην συναρτησιακή ανάλυση είναι ότι το ανάπτυγμα των συντελεστών γίνεται με $\zeta_i = \langle \varphi_i, f \rangle$. Η SVE όμως είναι μοναδική και με αυτό εννοούμε ότι αν οι ιδιάζουσες ιδιοτιμές είναι απομονωμένες, τότε μπορούμε πάντα να αντικαθιστούμε τις συναρτήσεις u_i και v_i , με τις αντίστοιχες $-u_i$ και $-v_i$. Αν όμως οι ιδιάζουσες ιδιοτιμές έχουν πολλαπλότητα μεγαλύτερη από το 1, δηλαδή $\mu_{i-1} > \mu_i = \dots = \mu_{i+v} > \mu_{i+v+1}$ τότε τα αντίστοιχα ιδιάζοντα ιδιοδιανύσματα δεν είναι μοναδικά, αλλά ο υπόχωρος τους είναι μοναδικός.

Πιο αναλυτικά, αν έχουμε δύο ίδιες ιδιοτιμές $\mu_1 = \mu_2$, θα μπορούμε να αντικαθιστούμε τα u_1 και v_1 με δύο γραμμικές σχέσεις $\alpha u_1 + \beta u_2$ και $\alpha v_1 + \beta v_2$, εφόσον ισχύει $\alpha^2 + \beta^2 = 1$, έτσι ώστε να εξασφαλίσουμε ότι ιδιάζουσες συναρτήσεις ικανοποιούν την L_2 νόρμα.

3.6.1 Μη μοναδικότητα της λύσης

Η συνθήκη Hadamard είναι σημαντική διότι εξασφαλίζει την ύπαρξη και τη μοναδικότητα της λύσης. Ο λόγος είναι ότι δεν έχουν όλα τα προβλήματα της ολοκληρωτικής εξίσωσης Fredholm Πρώτου Είδους μοναδική λύση, έτσι το πρόβλημα μας είναι ασταθές. Η μη ύπαρξη μοναδικής λύσης προκαλεί ασάφεια στα προβλήματα. Σε αυτές τις περιπτώσεις, διακρίνουμε δύο είδη παραβίασης της μοναδικότητας (ambiguity):

(α) Μη τετριμμένος πυρήνας (null space ambiguity).

Στην παραβίαση μοναδικότητας αυτού του είδους υπάρχουν συναρτήσεις $f_{null} \neq 0$ τέτοιες ώστε

$$\int_0^1 k(s, t) f_{null}(t) dt = 0 \quad (3.6.1.1)$$

Οι συναρτήσεις f_{null} οι οποίες αναφέρονται και ως μηδενιστές (annihilators), είναι ανεπιθύμητες, και πιθανόν η συμβολή τους στο αποτέλεσμα της λύσης να είναι ορατό.

Το σύνολο των γραμμικών συνδυασμών των μηδενιστών ονομάζεται και μηδενοχώρος (null space). Από την θεμελιώδη εξίσωση:

$$\int_0^1 k(s, t) v_i(t) dt = \mu_i u_i(s) \quad (3.6.1.2)$$

Καταλαβαίνουμε ότι ο μηδενοχώρος αποτελείται από κάθε γραμμικό συνδυασμό των αριστερών ιδιαιζουσών συναρτήσεων $v_i(t)$ που αντιστοιχούν στις ιδιάζουσες τιμές μ_i . Ως εκ τούτου ο πυρήνας $k(s, t)$ είναι εκφυλισμένος, εφόσον περισσότερο από μια των ιδιάζουσων τιμών, κάνει μηδέν. Συνήθως αντιμετωπίζουμε προβλήματα με τον μηδενοχώρο. Αυτό αποτελεί μια ιδανική περίπτωση, όπου τα δεξιά ιδιάζοντα ιδιοδιανύσματα v_i , αντιστοιχούν στις ιδιάζουσες τιμές $\mu_i = 0$. Η διάσταση του μηδενοχώρου είναι πεπερασμένη μόνο αν έχει ένα πεπερασμένο αριθμό ιδιάζουσων τιμών που είναι ίσες με το μηδέν. Αλλιώς η διάσταση είναι άπειρη, όταν υπάρχει ένας άπειρος αριθμός με μη μηδενικές ιδιάζουσες τιμές.

Γενικότερα, οι μηδενιστές χρησιμοποιούνται σε πολλές εφαρμογές διότι μπορούν να ανιχνεύουν τη λύση και να τη λύνουν αριθμητικά. Όμως, λόγω των σφαλμάτων που προκύπτουν από την διακριτοποίηση ενός συνεχούς προβλήματος και των στρογγυλοποιήσεων του υπολογιστή, πολλές φορές είναι δύσκολο να λυθεί αριθμητικά με σωστό αποτέλεσμα.

(β) Πρόβλημα με την διατύπωση (formulation ambiguity)

Το δεύτερο είδος, είναι το πρόβλημα με την διατύπωση (formulation ambiguity) και παρουσιάζεται σε εφαρμογές, στις οποίες δύο διαφορετικές ολοκληρώσιμες μορφές με δύο διαφορετικούς πυρήνες, οδηγούν στην ίδια δεξιό μέρος συνάρτησης g , η οποία συνδέεται με μια φυσική ποσότητα, όπως το παράδειγμα του δυναμικού της θεωρίας του Green.

3.7 Παραγοντοποίηση ιδιάζουσων ιδιοτιμών (SVD)

Μέχρι τώρα, ακολουθήσαμε τις μεθόδους διακριτοποίησης, για να μετατρέψουμε ένα συνεχές πρόβλημα σε διακριτό γραμμικό σύστημα από τις εξισώσεις της γραμμικής άλγεβρας. Έπειτα εξηγήσαμε τη συμβολή της μεθόδου απόδοσης ιδιάζουσων τιμών SVE στην ανάλυση μιας ολοκληρωτικής εξίσωσης Fredholm Πρώτου Είδους. Πολλές από τις αρχές της SVE μας οδηγούν στην παραγοντοποίηση ιδιάζουσων ιδιοτιμών SVD.

Σε αυτό το σημείο, είναι απαραίτητο να εξετάσουμε τον τρόπο που θα υπολογίσουμε το γραμμικό σύστημα, το οποίο θα προκύψει με τη μέθοδο διακριτοποίησης. Τη λύση στο διακριτό πρόβλημα προσφέρει η SVD (Singular Value Decomposition) ή αλλιώς παραγοντοποίηση ιδιάζουσων ιδιοτιμών.

Ορισμός SVD (Παραγοντοποίηση ιδιάζουσων ιδιοτιμών)

Έστω ότι έχω την γραμμική εξίσωση $Ax = b$ με πεπερασμένη τάξη πινάκων με $A \in \mathcal{R}_{n \times m}$ και $m \geq n$. Τότε η παραγοντοποίηση ιδιάζουσων ιδιοτιμών προκύπτει όταν ο πίνακας A λαμβάνει την πιο κάτω μορφή:

$$A = U\Sigma V^T = \sum_{i=1}^n u_i \sigma_i v_i^T \quad (3.7.1)$$

Επίσης αν υπάρχει ο αντίστροφος A^{-1} και ο A είναι ορθογώνιος τότε ισχύει:

$$A^{-1} = V\Sigma^{-1}U^T \quad (3.7.2)$$

Επισημάνσεις:

- ✓ Τα u_i, v_i καλούνται ιδιάζοντα ιδιοδιανύσματα.
- ✓ Ο πίνακας U έχει διαστάσεις $m \times m$ και αποτελείται από τα ιδιάζοντα ιδιοδιανύσματα $\{u_1, u_2, \dots, u_n\}$.
- ✓ Ο πίνακας V έχει διαστάσεις $n \times m$ και αποτελείται από τα ιδιάζοντα ιδιοδιανύσματα $\{v_1, v_2, \dots, v_n\}$.

- ✓ Ο πίνακας Σ είναι ένας **διαγώνιος πίνακας** με διαστάσεις $n \times n$ και στη διαγώνιο του βρίσκονται οι ιδιάζουσες ιδιοτιμές $\{\sigma_1, \sigma_2, \dots, \sigma_n\}$ κατά φθίνουσα σειρά και είναι πάντα θετικές, δηλαδή $\sigma_1 > \sigma_2 > \dots > \sigma_n$
- ✓ Όταν και οι δύο πίνακες είναι ορθογώνιοι $u_i^T u_j = v_i^T v_j = \delta_{ij}$, για $i, j = 1, \dots, n$.
- ✓ Το πιο πάνω σημείο μπορεί να γραφτεί πιο απλά και ως $U^T U = V^T V = I$.

Οι νόρμες πινάκων εκφράζονται μέσω ιδιάζουσων τιμών ως εξής:

$$\begin{aligned}
 1. \quad \|A\|_F &= \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \right)^{1/2} \\
 &= \left(\text{trace}(A^T A) \right)^{1/2} \\
 &= \left(\text{trace}(V \Sigma^2 V^T) \right)^{1/2} \\
 &= \left(\text{trace}(\Sigma^2) \right)^{1/2} \\
 &= (\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2)^{1/2}
 \end{aligned}$$

$$2. \quad \|A\|_2 = \max_{\|x\|_2 < 1} \|Ax\|_2 = \sigma_1$$

$$3. \quad \text{Η } L_2 \text{ νόρμα, του αντιστρόφου δίνεται και από την σχέση } \|A^{-1}\|_2 = \sigma_n^{-1}.$$

Ως εκ τούτου, ο δείκτης κατάστασης του πίνακα A προσδιορίζεται από τον πιο κάτω τύπο:

$$\text{cond}(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_1}{\sigma_n} \quad (3.7.3)$$

3.8 Ο ρόλος της SVD μεθόδου

Η παραγοντοποίηση ιδιάζουσων τιμών ή αλλιώς κοινώς γνωστή και ως SVD ακολουθεί μια σειρά μαθηματικών σχέσεων που είναι παρόμοιες με τις σχέσεις της SVE. Είναι ενδιαφέρον να γνωρίζουμε ότι η SVD μέθοδος αναπτύχθηκε πολύ αργότερα από την SVE. Παρά το γεγονός αυτό, όπως θα δούμε στη συνέχεια, ο ρόλος της SVD είναι εξίσου πολύ σημαντικός για τη λύση των διακριτών προβλημάτων.

Επομένως οι ιδιάζουσες τιμές και τα ιδιάζοντα διανύσματα θα ακολουθούν μια αντίστοιχη σειρά σχέσεων με αυτές της SVE με τις εξής βασικές σχέσεις:

$$1) Av_i = \sigma_i u_i \text{ και } \|Av_i\|_2 = \sigma_i \text{ για } i = 1, \dots, n \quad (3.8.1)$$

$$2) A^{-1}u_i = \sigma_i^{-1}v_i \text{ και } \|A^{-1}u_i\|_2 = \sigma_i^{-1} \text{ για } i = 1, \dots, n \quad (3.8.2)$$

Η πιο σημαντική σχέση στην οποία θα εστιάσουμε την προσοχή μας είναι αυτή που προκύπτει ως λύση της γραμμικής εξίσωσης $Ax = b$, δηλαδή θα προκύψει η εξής σχέση:

$$x = A^{-1}b = \sum_{i=1}^n \frac{u_i^T b}{\sigma_i} v_i \quad (3.8.3)$$

Απόδειξη:

Αρχικά πρέπει να γνωρίζουμε ότι ο πίνακας V είναι ορθογώνιος και ότι ο πίνακας A ορίζεται και είναι τετραγωνικός. Επομένως γράφουμε μαζί με την SVD, τα διανύσματα x και b και προκύπτει η πιο κάτω λύση:

$$\begin{array}{l}
 x = VV^T x = V \begin{bmatrix} V_1^T x \\ \vdots \\ V_n^T x \end{bmatrix} = \sum_{i=1}^n (v_i^T x) v_i \\
 b = \sum_{i=1}^n (u_i^T b) u_i \\
 \xRightarrow{SVD} A = U\Sigma V^T = \sum_{i=1}^n u_i \sigma_i v_i^T
 \end{array}
 \left. \vphantom{\begin{array}{l} x = VV^T x \\ b = \sum_{i=1}^n (u_i^T b) u_i \\ \xRightarrow{SVD} A = U\Sigma V^T \end{array}} \right\} Ax = \sum_{i=1}^n \sigma_i (v_i^T x) u_i$$

$$\downarrow$$

$$x = A^{-1}b = \sum_{i=1}^n \frac{u_i^T b}{\sigma_i} v_i \text{ (λύση)}$$

3.8.1 Συνθήκη Picard στην διακριτή περίπτωση, σε συνδυασμό με την μέθοδο παραγοντοποίησης ιδιάζουσων τιμών SVD

Ένα πλεονέκτημα της SVD είναι ότι οι συντελεστές της $\mathbf{u}_i^T \mathbf{b}$ είναι προσεγγίσεις προς τα αντίστοιχα εσωτερικά γινόμενα $\langle \mathbf{u}_i, \mathbf{g} \rangle$, στην βάση SVE. Αυτό παρατηρείται όταν στο εσωτερικό γινόμενο $\langle \mathbf{u}_j^{(n)}, \mathbf{g}^{(n)} \rangle$, γίνονται καλύτερες προσεγγίσεις από ότι στο εσωτερικό γινόμενο $\langle \mathbf{u}_i, \mathbf{g} \rangle$. Με βάση τον ορισμό του εσωτερικού γινομένου και των εκφράσεων $\mathbf{u}_j^{(n)}, \mathbf{g}^{(n)}$ προκύπτει το $\mathbf{u}_j^T \mathbf{b}$:

$$\langle \mathbf{u}_j^{(n)}, \mathbf{g}^{(n)} \rangle = \int_0^1 \sum_{i=1}^n u_{ij} \psi_i(s) \sum_{k=1}^n b_k \psi_k(s) ds = \sum_{i=1}^n u_{ij} \sum_{k=1}^n b_k \langle \psi_i \psi_k \rangle = \sum_{i=1}^n u_{ij} b_i = \mathbf{u}_j^T \mathbf{b} \quad (3.8.1.1)$$

Η συνθήκη Picard είναι πολύ βασική στην διακριτή περίπτωση και μας βοηθάει να βρούμε την κατάλληλη προσεγγιστική λύση στην περίπτωση που διερευνούμε ένα πρόβλημα με την ανάλυση της SVD. Συμπεραίνουμε ότι μπορούμε να χρησιμοποιήσουμε τις αριθμητικές ποσότητες της ανάλυσης του πίνακα A με την SVD και να μελετάμε κατά πόσο ικανοποιείται η συνθήκη Picard.

Επομένως, η συνθήκη Picard ικανοποιείται όταν, για όλες τις τιμές που είναι μεγαλύτερες από το τ , (εξ.3.8.3) το γινόμενο του παρανομαστή $|\mathbf{u}_i^T \mathbf{b}|$ μειώνεται γρηγορότερα από το σ_i .

Οπότε, η συνθήκη Picard καθορίζει μέχρι πιο σημείο τ θα σταματήσουμε να υπολογίζουμε τις ιδιάζουσες τιμές, αφού ακόμα και για πολύ μικρό θόρυβο η λύση που θα βρούμε θα απέχει πολύ από την ακριβή.

Κεφάλαιο 4

4. Μέθοδοι Ομαλοποίησης

Σε αυτό το κεφάλαιο θα ασχοληθούμε κυρίως με τις δυσκολίες που αντιμετωπίζουμε, όταν θέλουμε να βρούμε την ακριβή λύση σε μη καλά τοποθετημένα προβλήματα. Θα χρησιμοποιήσουμε μεθόδους, οι οποίες βοηθούν στην εύρεση προσεγγιστικής λύσης.

Όταν διακριτοποιούμε ένα συνεχές πρόβλημα, συνήθως είναι μη καλά τοποθετημένο και κύριο χαρακτηριστικό του είναι ο πολύ μεγάλος δείκτης κατάστασης του συστήματος. Ο στόχος των μεθόδων ομαλοποίησης είναι η «ομαλοποίηση» της λύσης, η οποία επιτυγχάνεται μέσω της εξάλειψης του θορύβου που υπεισέρχεται κατά την προσεγγιστική επίλυση ενός διακριτού προβλήματος.

Κατά κανόνα, οι μέθοδοι ομαλοποίησης παράγουν προσεγγιστικές λύσεις, οι οποίες μπορούν να εκφραστούν βάσει ενός φιλτραρισμένου αναπτύγματος παραγοντοποίησης ιδιάζουσων τιμών (SVD) όπως θα δούμε αναλυτικά πιο κάτω.

Ισχύει επίσης ότι:

$$\|x_{\text{ακριβές}} - x\|_2 \|x_{\text{ακριβές}}\|_2 \leq \text{cond}(A) \|\varepsilon\|_2 \|b_{\text{ακριβές}}\|_2 \quad (4.1)$$

Έστω ότι έχουμε δύο λύσεις, την ακριβή λύση $x_{\text{ακριβές}}$, την προσεγγιστική λύση x

και έστω ε μια πολύ μικρή διαταραχή. Ο σκοπός είναι, με τις κατάλληλες μεθόδους να βρούμε μια λύση που να είναι πολύ κοντά με την ακριβή. Όταν όμως αντιστρέψουμε το πρόβλημα και το λύσουμε όπως ένα σύστημα γραμμικής άλγεβρας, δηλαδή $x=A^{-1}b$, τότε η λύση που θα πάρουμε θα αποκλίνει πολύ από την ακριβή ακόμα και για μικρές διαταραχές.

Αυτές οι μέθοδοι που θα βοηθήσουν στην εύρεση της προσεγγιστικής λύσης, καλούνται μέθοδοι ομαλοποίησης. Οι μέθοδοι ομαλοποίησης λειτουργούν με τέτοιο τρόπο ώστε να απαιτούν την ομαλότητα της λύσης. Επομένως, οι μέθοδοι αυτοί συμβάλλουν στην καταστολή των ανεπιθύμητων σφαλμάτων, πιο συγκεκριμένα στο θόρυβο, με αποτέλεσμα να έχουμε μια προσεγγιστική λύση.

Η βάση των μεθόδων είναι η SVD και μπορεί να θεωρηθεί ως μια φασματική βάση φιλτραρίσματος. Σε αυτή την ενότητα θα ασχοληθούμε με τρεις μεθόδους ομαλοποίησης οι οποίες παράγουν λύσεις και είναι οι εξής:

- i. Μέθοδος αποκοπής SVD (Truncated SVD) ή TSVD
- ii. Επιλεκτική μέθοδος SVD (Selective SVD) ή SSVD
- iii. Tikhonov SVD

Η μορφή εξίσωσης που χρησιμοποιούμε για να ομαλοποιήσουμε τη λύση μας είναι ένα φιλτραρισμένο ανάπτυγμα παραγοντοποίησης ιδιάζουσων τιμών (SVD). Οι συντελεστές, φ_i της εξίσωσης αυτής είναι οι παράγοντες που σχετίζονται με τη μέθοδο φιλτραρίσματος. Οπότε έχουμε την πιο κάτω σχέση:

$$X_{reg} = \sum_{i=1}^n \varphi_i \frac{u_i^T b}{\sigma_i} v_i \quad (4.2)$$

Όπως έχουμε δει, όταν διακριτοποιούμε ένα συνεχές πρόβλημα, συνήθως είναι μη καλά τοποθετημένο και κύριο χαρακτηριστικό του είναι ο πολύ μεγάλος δείκτης κατάστασης του συστήματος. Αυτό υποδηλώνει, ότι η λύση είναι πολύ ευαίσθητη σε οποιαδήποτε διαταραχή του δεξιού μέρους και ότι παρουσιάζονται σφάλματα από τα δεδομένα.

Επομένως, συμπεραίνουμε ότι χαρακτηριστικό των διακριτών μη καλά τοποθετημένων προβλημάτων είναι οι πίνακες συντελεστών, οι οποίοι έχουν πολύ μεγάλο δείκτη κατάστασης. Αυτό πρακτικά σημαίνει ότι η ευστάθεια της λύσης επηρεάζεται σημαντικά ακόμη και από μικρά σφάλματα στα δεδομένα b .

Έστω ότι έχουμε δύο λύσεις, την ακριβή λύση $x_{ακριβες}$, την προσεγγιστική λύση x και έστω ϵ μια πολύ μικρή διαταραχή:

$$Ax = b$$

$$Ax_{ακριβες} = b_{ακριβες}$$

Ψάχνουμε λύση

$$x \simeq x_{ακριβες}$$

Ισχύει επίσης :

$$\frac{\|x_{\text{ακριβές}} - x\|_2}{\|x_{\text{ακριβές}}\|_2} \leq \text{cond}(A) \frac{\|e\|_2}{\|b_{\text{ακριβές}}\|_2} \quad (4.3)$$

Όσο πιο μεγάλο δείκτη κατάστασης έχουμε τόσο πιο μεγάλο αναμένεται να είναι το σφάλμα, δηλαδή τόσο περισσότερο απέχει η προσεγγιστική λύση από την ακριβή λύση. Αυτός είναι και ο λόγος για τον οποίο εισάγονται οι μέθοδοι ομαλοποίησης, οι οποίες παρουσιάζουν μεγάλη «ανοχή» στα σφάλματα που υπεισέρχονται στα δεδομένα b .

Η λύση $x = A^{-1}b$, λόγω της ασθενούς κατάστασης του πίνακα A , αναμένεται να βρίσκεται αρκετά μακριά από την ακριβή λύση $x_{\text{ακριβές}}$, ακόμη και εάν η διαταραχή είναι μικρή, δηλαδή όταν $\|e\|_2 \leq \|b\|_2$.

4.1 Μέθοδος Αποκοπής SVD (Truncated SVD ή TSVD)

Η μέθοδος Αποκομμένης Παραγοντοποίησης Ιδιάζουσων Τιμών (Truncated Singular Value Decomposition - TSVD) αξιοποιεί το γεγονός ότι το εξαιρετικά μεγάλο σφάλμα στην αντίστροφη λύση προέρχεται από τον θόρυβο που προκαλείται όταν λύνουμε το σύστημα μας με το ανάπτυγμα της SVD μεθόδου. Ο θόρυβος αυτός προκύπτει από τις μικρότερες ιδιάζουσες ιδιοτιμές που υπάρχουν στη λύση. Άρα παρατηρούμε ότι η πραγματική λύση κυριαρχείται από τους SVD συντελεστές της πιο κάτω μορφής, όπου e είναι η διαταραχή που προκαλείται από το b :

$$\frac{u_i^T b}{\sigma_i} \approx \frac{u_i^T e}{\sigma_i} \quad (4.1.1)$$

Εντούτοις, ορισμένοι από τους συντελεστές SVD είναι αξιόπιστοι και προκύπτει ότι:

$$\frac{u_i^T b}{\sigma_i} \approx \frac{u_i^T b_{\text{ακριβές}}}{\sigma_i} \quad (4.1.2)$$

όπου $b_{\text{ακριβές}} = Ax_{\text{ακριβές}}$

Επίσης, μπορούμε να υποθέσουμε ότι η ακριβής λύση του δεξιού μέρους, ικανοποιεί τη διακριτή κατάσταση Picard. Έτσι βρίσκουμε, ότι οι λύσεις με το μεγαλύτερο μέγεθος είναι ακριβώς εκείνες όπου οι συντελεστές προσεγγίζονται καλύτερα.

Για μικρούς δείκτες i έχουμε ότι:

$$\frac{u_i^T b}{\sigma_i} \approx \frac{u_i^T b^{\text{ακριβές}}}{\sigma_i} = v_i^T x^{\text{ακριβές}} \quad (4.1.3)$$

Αποκόπτοντας τις συνιστώσες της SVD οι οποίες προκαλούν θόρυβο, προκύπτει και η μέθοδος αποκοπής SVD (Truncated SVD ή TSVD). Επομένως, ορίζουμε την ομαλοποιημένη λύση SVD (TSVD) ως x_k , η οποία προκύπτει αν διατηρήσουμε τις πρώτες k συνιστώσες της πραγματικής λύσης, ως εξής:

$$x_k \equiv \sum_{i=1}^k \frac{u_i^T b}{\sigma_i} v_i \quad (4.1.4)$$

Η παράμετρος αποκοπής k θα πρέπει να επιλέγεται, έτσι ώστε να απορρίπτεται ο θόρυβος, ο οποίος προκαλείται από την SVD. Μία κατάλληλη τιμή του k συχνά μπορεί να βρεθεί, από τη γραφική παράσταση της συνθήκης Picard.

Εξίσου σημαντική είναι και μια εναλλακτική διατύπωση της μεθόδου TSVD. Ενώ είδαμε ότι τα x_k υπολογίζονται από την σχέση (4.1.4), μπορούμε να ορίσουμε τη λύση με τον πίνακα A_k (k -τάξης) της μεθόδου TSVD, ο οποίος δίνει καλύτερα αποτελέσματα και πιο αξιόπιστα διότι είναι καλά τοποθετημένος.

$$A_k = \begin{pmatrix} | & & | \\ u_1 & \dots & u_k \\ | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{pmatrix} \begin{pmatrix} | & & | \\ v_1 & \dots & v_k \\ | & & | \end{pmatrix}^T = \sum_{i=1}^k u_i \sigma_i v_i^T \quad (4.1.5)$$

Ο δείκτης κατάστασης του πίνακα A_k είναι $\frac{\sigma_1}{\sigma_k}$, ο οποίος είναι προφανώς μικρότερος από το δείκτη κατάστασης του πίνακα A , ο οποίος είναι $\frac{\sigma_1}{\sigma_n}$.

4.2 Επιλεκτική μέθοδος SVD (Selective SVD ή SSVD)

Η μέθοδος Επιλεκτικής Παραγοντοποίησης Ιδιάζουσων Τιμών (Selective Singular Value Decomposition - SSVD) αξιοποιεί επιλεγμένες συνιστώσες με μεγάλη συνεισφορά στην ομαλοποιημένη λύση. Ειδικότερα, δοθέντος ενός ορίου συντελεστών τ , η λύση x_τ δίνεται από την ακόλουθη σχέση:

$$x_\tau \equiv \sum_{|u_i^T b|} \frac{u_i^T b}{\sigma_i} v_i \quad (4.2.1)$$

Όπως φαίνεται από την πιο πάνω σχέση, οι παράγοντες φίλτρου για τη μέθοδο SSVD είναι:

$$\varphi_i^{[\tau]} = \begin{cases} 1, & |u_i^T b| > \tau \\ 0, & \text{διαφορετικά} \end{cases} \quad (4.2.2)$$

όπου το τ πρέπει να επιλεγεί κατάλληλα έτσι ώστε $u_i^T \cong u_i^T e$.

Οι μέθοδοι TSVD και SSVD δίνουν περίπου το ίδιο αποτέλεσμα, με απόλυτο σφάλμα της ίδιας τάξης μεγέθους.

Επίσης, για τις δύο μεθόδους, μπορούμε να επιλέξουμε κατάλληλο k ή τ έτσι ώστε, κοιτώντας το διάγραμμα Picard της λύσης $x = A^{-1}b$, να καθορίσουμε το σημείο το

οποίο οι ιδιάζουσες τιμές $\frac{u_i^T b}{\sigma_i}$ της λύσης, αρχίζουν να αποκλίνουν.

Ωστόσο σε μεγάλα υπολογιστικά προβλήματα οι μέθοδοι TSVD και SSVD είναι δύσκολο να εφαρμοστούν. Τη λύση του προβλήματος σε αυτές τις περιπτώσεις θα μας την δώσει η μέθοδος Tikhonov SVD.

4.3 Tikhonov SVD

Στις προηγούμενες δύο μεθόδους TSVD και SSVD, το δομικό τους στοιχείο βασίζεται στον υπολογισμό ενός αναπτύγματος SVD, ο οποίος σε προβλήματα μικρής κλίμακας είναι γρήγορος και εύκολος. Ωστόσο, οι μέθοδοι αυτοί σε προβλήματα μεγάλης κλίμακας είναι χρονοβόρες, ως εκ τούτου η Tikhonov SVD καθίσταται πιο αποτελεσματική τις περισσότερες φορές. Η μέθοδος αυτή εξασφαλίζει ακριβή αποτελέσματα και έχει σαν βάση, την επίλυση ενός προβλήματος ελαχίστων τετραγώνων με κάποιον περιορισμό.

Η ομαλοποιημένη λύση x_λ είναι η λύση του ακόλουθου τύπου των ελαχίστων τετραγώνων:

$$\min \left\| \begin{pmatrix} A \\ \lambda L \end{pmatrix} x - \begin{pmatrix} b \\ \lambda Lx^* \end{pmatrix} \right\|_2 \quad (4.3.1)$$

Η λύση αυτή προκύπτει από το πρόβλημα των ελαχίστων τετραγώνων το οποίο είναι:

$$\begin{cases} \min_x \|Ax - b\|_2^2 \\ \|x\|_2^2 \leq \delta \end{cases} \quad (4.3.2)$$

Ο περιορισμός $\|x\|_2^2 \leq \delta$ εισάγεται για την ομαλοποίηση της λύσης (προκύπτει από τον ορισμό της Ευκλείδειας νόρμας).

Έπειτα, θέλουμε να βελτιστοποιήσουμε το πρόβλημα μας και αυτό γίνεται εύκολα, όταν ένας πολλαπλασιαστικός παράγοντας γ , εισαχθεί μέσα στο πρόβλημα ελαχίστων τετραγώνων. Πιο κάτω ισοδύναμα με το (4.3.2) προκύπτει το:

$$\min_x \{ \|Ax - b\|_2^2 + \gamma (\|x\|_2^2 - \delta^2) \} \quad (4.3.3)$$

Αν παραγωγίσουμε την (4.3.3) ως προς x και την εξισώσουμε με το μηδέν, έχουμε:

$$\min_x \{ \|Ax - b\|_2^2 + \lambda^2 \|x\|_2^2 \} \quad (4.3.4)$$

όπου $\gamma = \lambda^2$.

Επεξήγηση της έκφρασης (4.3.4)

- Ο παράγοντας λ είναι μια παράμετρος ομαλοποίησης των δεδομένων και η επιλογή της καθορίζει αν θα δοθεί βαρύτητα στην πρόβλεψη των δεδομένων από τη λύση, ή την πρόβλεψη στην ομαλοποιημένη λύση. Όσο πιο μεγάλο είναι το λ τόσο πιο μεγάλο βάρος δίνεται στην ομαλοποίηση της λύσης.
- Ο παράγοντας $\|Ax - b\|_2^2$ περιγράφει πόσο καλά προβλέπει τα δεδομένα b , η λύση x , τα οποία περιέχουν θόρυβο.
- Ο παράγοντας $\|x\|_2^2$ ομαλοποιεί την λύση.

Για να εκφραστεί η λύση της μεθόδου Tikhonov σε SVD ανάπτυγμα, πρέπει η μορφή $\min_x \{ \|Ax - b\|_2^2 + \lambda^2 \|x\|_2^2 \}$ να αναχθεί με μαθηματικούς υπολογισμούς σε μια απλούστερη μορφή, και να καταλήξει στην :

$$x_\lambda = \sum_{i=1}^n \varphi_i^{[\lambda]} \frac{u_i^T b}{\sigma_i} v_i \quad (4.3.5)$$

Με παράγοντες φιλτραρίσματος:

$$\varphi_i^{[\lambda]} = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} = \begin{cases} 1, & \text{αν } \sigma_i \gg \lambda \\ \frac{\sigma_i^2}{\lambda^2}, & \text{αν } \sigma_i \ll \lambda \end{cases} \quad (4.3.6)$$

Επομένως από τη σχέση (4.3.5), εξάγουμε την πληροφορία ότι για τις ιδιάζουσες τιμές σ_i , που είναι μεγαλύτερες από το λ , οι παράγοντες φιλτραρίσματος παίρνουν τιμές κοντά στη μονάδα.

Ως εκ τούτου, οι σταθερές $u_i^T b$ που αντιστοιχούν σε μεγάλες ιδιάζουσες τιμές σ_i , συνεισφέρουν στη λύση Tikhonov 100%.

Όμως, για ιδιάζουσες τιμές μικρότερες από το λ , οι παράγοντες φιλτραρίσματος τείνουν στο μηδέν καθώς τα σ_i μικραίνουν, με αποτέλεσμα η συνεισφορά τους στη λύση να μειώνεται δραστικά. Ο λόγος είναι ότι οι σταθερές SVD κυριαρχούνται από τον θόρυβο $\frac{u_i^T \varepsilon}{\sigma_i}$.

Όταν βρισκόμαστε σε μια ενδιάμεση κατάσταση όπου οι ιδιάζουσες τιμές είναι της ίδιας τάξης μεγέθους με την σταθερά λ , τότε οι παράγοντες φιλτραρίσματος μειώνονται με ομαλό τρόπο, καθώς μεταβαίνουν από μεγαλύτερα σε μικρότερα σ_i . Άρα, βλέπουμε τη σημαντικότητα της επιλογής του λ , για την ομαλοποίηση της λύσης.

4.4 Μέθοδος της Καμπύλης L (L-curve)

Μέχρι τώρα έχουμε δει ότι, για να μελετήσουμε τη συμπεριφορά των δύο πιο σημαντικών μεθόδων ομαλοποίησης, χρησιμοποιήσαμε τη μέθοδο SVD. Το επόμενο βήμα είναι να αξιολογήσουμε τη μέθοδο αυτή, έτσι ώστε να διαχωρίσει το πραγματικό μέρος της λύσης από τη νόρμα και το αντίστοιχο υπόλοιπο. Οι νόρμες αυτές βοηθάνε στην επίλυση των κακώς τοποθετημένων προβλημάτων, διότι πάντα γίνονται υπολογισμοί της λύσης, ανεξάρτητα με τη μέθοδο ομαλοποίησης που ακολουθούμε. Επίσης, δεν απαιτούν τον υπολογισμό της SVD ή οποιασδήποτε άλλης διακριτοποίησης του συστήματος. Η νόρμα της TSVD λύσης και της αντίστοιχης υπολειπόμενης λύσης αλλάζει μονότονα με το k .

Πιο κάτω βλέπουμε μια σειρά από το ε (δίνεται ως $\varepsilon = \|I - UU^T b\|_2$), το οποίο είναι μια νόρμα του συντελεστή b , η οποία βρίσκεται έξω από τη στήλη του χώρου A .

$$\|x_k\|_2^2 = \sum_{i=1}^k \left(\frac{u_i^T b}{\sigma_i} \right)^2 \leq \|x_{k+1}\|_2^2 \quad (4.4.1)$$

$$\|Ax_k - b\|_2^2 = \sum_{i=1}^k (u_i^T b)^2 + \varepsilon^2 \geq \|Ax_{k+1} - b\|_2^2 \quad (4.4.2)$$

Ομοίως, η νόρμα της λύσης Tikhonov και το συνδεόμενο υπόλοιπο της, δίνεται από τις πιο κάτω σχέσεις (όπου $\varphi_i^{[\lambda]} = \frac{\sigma_i^2}{(\sigma_i^2 + \lambda^2)}$ είναι ο παράγοντας φιλτραρίσματος) :

$$\|x_\lambda\|_2^2 = \sum_{i=1}^n \left(\varphi_i^{[\lambda]} \frac{u_i^T b}{\sigma_i} \right)^2 \quad (4.4.3)$$

$$\|Ax_\lambda - b\|_2^2 = \sum_{i=1}^n [(1 - \varphi_i^{[\lambda]}) u_i^T b]^2 + \varepsilon^2 \quad (4.4.4)$$

Για να δούμε καλύτερα τη μονότονη αλλαγή της νόρμας με το λ χρησιμοποιούμε τις εξής σχέσεις:

Για $\xi' < 0$

$$\xi = \|x_\lambda\|_2^2 \text{ και } \xi' = \frac{d\xi}{d\lambda} = -\frac{4}{\lambda} \sum_{i=1}^n (1 - \varphi_i^{[\lambda]}) (\varphi_i^{[\lambda]})^2 \frac{(u_i^T b)^2}{\sigma_i} \quad (4.4.5)$$

Για $\rho' > 0$

$$\rho = \|Ax_\lambda - b\|_2^2 \text{ και } \rho' = \frac{d\rho}{d\lambda} = \frac{4}{\lambda} \sum_{i=1}^n (1 - \varphi_i^{[\lambda]})^2 \varphi_i^{[\lambda]} (u_i^T b)^2 = -\lambda^2 \xi' \quad (4.4.6)$$

Οπότε παρατηρούμε ότι $\xi' < 0$ και $\rho' > 0$ για όλα τα λ . Έτσι, αποδεχόμαστε τη μονοτονία της νόρμας ως συνάρτηση του λ . Επίσης, από τη σχέση $\rho' = -\lambda^2 \xi'$ παίρνουμε την παράγωγο $\frac{d\xi}{d\rho} = -\lambda^{-2}$ και με αυτό τον τρόπο δείχνουμε ότι η τετραγωνική λύση της νόρμας $\|x_\lambda\|_2^2$ είναι μια φθίνουσα και μονότονη συνάρτηση του τετραγώνου του υπολείπου $\|Ax_\lambda - b\|_2^2$. Αφού η τετραγωνική ρίζα είναι μονότονη συνάρτηση, τότε το ίδιο ισχύει και για τις νόρμες. Έτσι, η καμπύλη που προκύπτει από τη γραφική παράσταση της ξ συναρτήσει της ρ , με παράμετρο λ , είναι η c_λ και προκύπτει ως εξής:

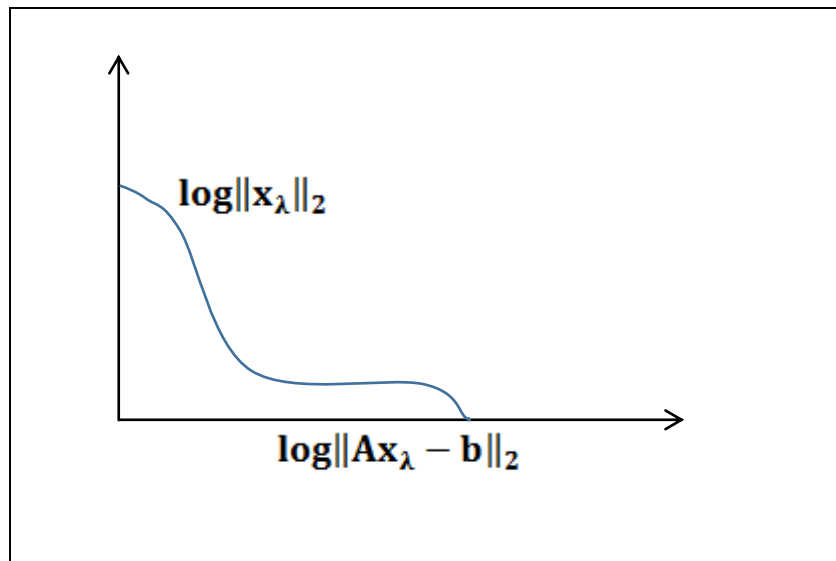
$$\begin{aligned} \rho'' = \frac{d^2\rho}{d\lambda^2} = \frac{d}{d\lambda} (-\lambda^2 \xi') &= -2\lambda \xi' - \lambda^2 \xi'', \quad \xi'' = \frac{d^2\xi}{d\lambda^2} \\ \Rightarrow c_\lambda &= \frac{\rho' \xi'' - \rho'' \xi'}{((\rho')^2 + (\xi')^2)^{3/2}} = \frac{2\lambda (\xi')^2}{((\rho')^2 + (\xi')^2)^{3/2}} \end{aligned} \quad (4.4.7)$$

Άρα, παρατηρούμε ότι $c_\lambda > 0$ για όλα τα λ , και ως εκ τούτου η καμπύλη (ρ, ξ) είναι κυρτή.

Η καμπύλη c_λ είναι πολύ σημαντική διότι δείχνει τον τρόπο με τον οποίο αλλάζει η ομαλοποιημένη παράμετρος λ . Δεδομένου ότι, οποιοδήποτε σημείο (ρ, ξ) στην καμπύλη είναι μια λύση στο πρόβλημα, με $\rho = \min \|Ax - b\|_2^2$, και με τον περιορισμό $\|x\|_2^2 \leq \xi$, έπεται ότι η καμπύλη, ορίζει ένα όριο μεταξύ δύο περιοχών του πρώτου τεταρτημόριου.

Όταν η καμπύλη χαράσσεται σε διπλή λογαριθμική κλίμακα και έχει την εξής μορφή $(\frac{1}{2} \log \rho, \frac{1}{2} \log \xi) = (\log \|Ax_\lambda - b\|_2, \log \|x_\lambda\|_2)$, τότε καλείται ως L-καμπύλη ομαλοποίησης της μεθόδου Τικχονοφ, η οποία διαδραματίζει το βασικότερο ρόλο για την ανάλυση ενός μη καλά τοποθετημένου προβλήματος.

Πιο κάτω βλέπουμε σχηματικά (Εικόνα 4.4.1) την ομαλοποιημένη L-καμπύλη για τη μέθοδο Τικχονοφ, στην οποία παρατηρούμε το διπλό λογαριθμικό γράφημα της λύσης της νόρμας $\|x_\lambda\|_2$, συναρτήσεως της υπολειπόμενης λύσης της νόρμας $\|Ax_\lambda - b\|_2$ με την παράμετρο λ . Παρατηρούμε ότι τα κάθετα και τα οριζόντια μέρη της γραφικής παράστασης αντιστοιχούν στις ομαλοποιημένες λύσεις.



Εικόνα 4.4.1-Διπλό λογαριθμικό γράφημα της λύσης της L-καμπύλης

Η γωνία που σχηματίζεται στο γράφημα (Εικόνα 4.4.1) και χωρίζει τα δύο αυτά μέρη, βρίσκεται περίπου στο σημείο: $(\log \|e\|_2, \log \|x_{\text{ακριβές}}\|_2)$.

Προς τα δεξιά το $\lambda \rightarrow \infty$ και η καμπύλη αρχίζει να λυγίζει προς τα κάτω, όσο αυξάνεται η ποσότητα της ομαλοποίησης. Έτσι αναγκάζει τη λύση να πάει προς το μηδέν.

Όμοια, στο πάνω αριστερό και πάνω από το κάθετο μέρος, η καμπύλη τελικά θα γίνει λιγότερο απότομη καθώς το $\lambda \rightarrow 0$ και η νόρμα $\|x_\lambda\|_2 \rightarrow \|A^{-1}b\|_2$. Επίσης, η L-καμπύλη μπορεί να προσδιοριστεί για κατάλληλες τιμές των παραμέτρων ομαλοποίησης.

Η L - καμπύλη έχει δύο ευδιάκριτα μέρη:

- Το προσεγγιστικό οριζόντιο τμήμα
- Το προσεγγιστικό κατακόρυφο τμήμα

Για μεγάλες τιμές του λ έχουμε ότι:

Η νόρμα $\|x_\lambda\|_2 \approx \|x_{\text{ακριβές}}\|_2$ είναι μια σταθερά, ενώ η νόρμα $\|Ax_\lambda - b_{\text{exact}}\|_2$ αυξάνεται με τις τιμές του λ .

Για μικρές τιμές του λ έχουμε ότι:

Η νόρμα $\|Ax_\lambda - b\|_2 \approx \|e\|_2$ είναι μια σταθερά, ενώ η νόρμα $\|x_\lambda\|_2$ αυξάνεται με τις τιμές του $\frac{1}{\lambda}$.

Για τον υπολογισμό του Κριτηρίου L της λύσης Tikhonov, απαιτείται αρχικά ο προσδιορισμός της καμπυλότητας για τις διάφορες τιμές του λ , και στη συνέχεια η κατάλληλη τιμή εκείνου του λ , το οποίο μεγιστοποιεί την καμπυλότητα.

Για παράδειγμα, όταν το λ είναι μεγάλο, τότε το x_λ προσδιορίζεται από τους συντελεστές της SVD, η οποία συμβάλει στην εύρεση της λύσης χρησιμοποιώντας το ακριβές δεξιό μέρος, $b_{\text{ακριβές}}$, όπου η λύση είναι ομαλοποιημένη. Ενώ για μικρές τιμές η λύση της μεθόδου Tikhonov, κυριαρχείται από τις διαταραχές των σφαλμάτων, που προέρχονται από το θόρυβο, που προκαλείται όταν αντιστρέφουμε το σύστημα.

Κεφάλαιο 5

5. Επαναληπτικές Μέθοδοι Ομαλοποίησης

Οι άμεσες μέθοδοι που μελετήθηκαν έως τώρα είναι αποτελεσματικές για περιπτώσεις προβλημάτων μικρής κλίμακας. Τα περισσότερα πρακτικά προβλήματα οδηγούν σε μεγάλους πίνακες, γεγονός που καθιστά τις μεθόδους αυτές πρακτικά ανεφάρμοστες.

Από την άλλη, η διακριτοποίηση ορισμένων κλάσεων αντίστροφων προβλημάτων μεγάλης κλίμακας οδηγούν σε «αραιούς» (sparse) πίνακες συντελεστών, των οποίων η επίλυση παρουσιάζει σημαντικά πλεονεκτήματα.

Προκειμένου μία μέθοδος ομαλοποίησης να είναι αποτελεσματική για την επίλυση προβλημάτων μεγάλης κλίμακας, θα πρέπει:

- ✓ τα κύρια δομικά της στοιχεία να είναι πολλαπλασιασμοί πίνακα-διανύσματος (προκειμένου να αποφεύγεται οποιαδήποτε παραγοντοποίηση πίνακα) και
- ✓ αφετέρου να είναι εφικτή η επιλογή της παραμέτρου ομαλοποίησης λ , χωρίς να απαιτείται η εκ νέου συνολική επίλυση του προβλήματος για κάθε νέα παράμετρο.

Οι επαναληπτικές (iterative) μέθοδοι ομαλοποίησης κατά κανόνα στηρίζονται σε πολλαπλασιασμούς πίνακα-διανύσματος και ως εκ τούτου ικανοποιούν την πρώτη από τις παραπάνω συνθήκες, χωρίς όμως να ισχύει απαραίτητα κάτι αντίστοιχο και για τη δεύτερη από τις προαναφερθείσες συνθήκη.

Οι επαναληπτικές μέθοδοι εκκινούν με ένα αρχικό διάνυσμα αυθαίρετα ορισμένο, συνήθως το μηδενικό. Ακολουθώς παράγεται μία ακολουθία λύσεων $x[1], x[2], \dots$, η οποία συγκλίνει σε κάποια τελική λύση. Κάποιες επαναληπτικές μέθοδοι, μετά από ένα σχετικά μεγάλο αριθμό επαναλήψεων μπορεί να αρχίζουν να αποκλίνουν από την ακριβή λύση και συγκλίνουν στην αντίστροφη λύση. Πρόκειται για σύγκλιση η οποία είναι γνωστή ως ημισύγκλιση (semiconvergence) και στις περιπτώσεις αυτές καθίσταται σημαντική η χρονική στιγμή κατά την οποία πρέπει να τερματιστούν οι επαναλήψεις .

Κλασσικές Επαναληπτικές Μέθοδοι

Στην παρούσα ενότητα παρουσιάζεται ένας αριθμός κλασσικών επαναληπτικών μεθόδων, οι οποίες παρουσιάζουν το φαινόμενο της ημισύγκλισης. Οι μέθοδοι αυτές παρότι παρουσιάζουν σχετικά αργή σύγκλιση προτιμώνται σε πολύ συγκεκριμένες περιπτώσεις εφαρμογών για τις οποίες κρίνονται αρκετά αποτελεσματικές συγκρινόμενες με άλλες μεθόδους.

5.1 Στάσιμες επαναληπτικές διαδικασίες ομαλοποίησης

Οι μέθοδοι Landweber και Cimmino που παρουσιάζονται στην παρούσα ενότητα, αποτελούν ειδικές περιπτώσεις της προαναφερθείσας ευρύτερης κατηγορίας επαναληπτικών μεθόδων και εν γένει παρουσιάζουν το φαινόμενο της ημισύγκλισης.

5.1.1 Επαναληπτική Μέθοδος Landweber

Μια από τις πιο γνωστές μεθόδους που παρουσιάζει **ημισύγκλιση** είναι η επαναληπτική μέθοδος **Landweber**. Στη βασική της μορφή η μέθοδος εκφράζεται ως εξής:

$$x^{[k+1]} = x^{[k]} + \omega A^T(b - Ax^{[k]}), \quad k = 1, 2, \dots \quad (5.1.1.1)$$

όπου ω πραγματικός αριθμός που ικανοποιεί τη σχέση:

$$0 < \omega < 2\|A^T A\|_2^{-1} = \frac{2}{\sigma_i^2} \quad (5.1.1.2)$$

Σε κάθε επανάληψη της μεθόδου απαιτείται ο υπολογισμός του διανύσματος υπολοίπου:

$$r^k = b - Ax^k \quad (5.1.1.3)$$

Όπως συμβαίνει και με τις λύσεις των μεθόδων TSVD και Tikhonov, η επανάληψη $x^{[k]}$ μπορεί να εκφραστεί με τη βοήθεια της συνάρτησης φίλτρου ως εξής:

$$x^{[k]} = V \Phi^{[k]} \Sigma^{-1} U^T b \quad (5.1.1.4)$$

όπου τα στοιχεία της διαγωνίου $\Phi^{[k]} = \text{diag}(\varphi_1^{[k]}, \dots, \varphi_n^{[k]})$ είναι οι συντελεστές φίλτρου για την επανάληψη $\mathbf{x}^{[k]}$ και υπολογίζονται ως εξής:

$$\varphi_i^{[k]} = 1 - (1 - \sigma_i^2)^k, i = 1, 2, \dots \quad (5.1.1.5)$$

Βασισμένοι στους συντελεστές φίλτρου μπορούμε να ορίσουμε ως σημείο καμπής της συγκεκριμένης μεθόδου, την τιμή $\sigma_{break}^{[k]}$ των σ_i , για τα οποία $\varphi_i^{[k]} = 0.5$, όπου θα έχουμε:

$$\sigma_{break}^{[k]} = \sqrt{\frac{1 - \left(\frac{1}{2}\right)^{\frac{1}{k}}}{\omega}} \quad (5.1.1.6)$$

Κάθε φορά που ο αριθμός των επαναλήψεων k διπλασιάζεται, το σημείο καμπής τείνει να μειώνεται κατά $\sqrt{2} \approx 1.4$ φορές, με αποτέλεσμα την αργή σύγκλιση της επαναληπτικής μεθόδου Landweber.

5.1.2 Επαναληπτική Μέθοδος Cimmino

Η επαναληπτική μέθοδος Cimmino συγκαταλέγεται και αυτή, στις κλασικές επαναληπτικές διαδικασίες που παρουσιάζουν **ημισύγκλιση**. Η βασική της μορφή είναι η εξής:

$$\mathbf{x}^{[k+1]} = \mathbf{x}^{[k]} + \omega \mathbf{A}^T \mathbf{D} (\mathbf{b} - \mathbf{A} \mathbf{x}^{[k]}), \quad k = 1, 2, \dots \quad (5.1.2.1)$$

όπου $\mathbf{D} = \text{diag}(d_i)$, διαγώνιος πίνακας όπου τα στοιχεία του ορίζονται ως εξής:

$$d_i = \begin{cases} \frac{1}{m \|a_i\|_2^2}, & a_i \neq 0, \\ 0, & a_i = 0. \end{cases} \quad (5.1.2.2)$$

5.2 Μέθοδοι προβολής

Όπως έχουμε ήδη δει στην ανάλυση των διακριτών και κακώς ορισμένων προβλημάτων, είναι αδύνατο να υπολογίσουμε με ακρίβεια το $\mathbf{x}_{\text{ακριβές}}$.

Ως εκ τούτου, η SVD ανάλυση μας πληροφορεί ότι το καλύτερο που μπορούμε να κάνουμε, είναι να υπολογίσουμε μία προσέγγιση για το $\mathbf{x}_{\text{ακριβές}}$, η οποία βρίσκεται σε έναν (χαμηλών-διαστάσεων) υπόχωρο του \mathbf{R}^n . Για την TSVD, αυτός ο υπόχωρος φτιάχνεται από τα πρώτα k (με δεξιόστροφη φορά) μοναδιαία διανύσματα $\mathbf{v}_1, \dots, \mathbf{v}_k$ και γνωρίζουμε ότι αυτός ο υπόχωρος, καταγράφει κυρίως χαμηλές συχνότητες του $\mathbf{x}_{\text{ακριβές}}$, αφού αυτά τα μοναδιαία διανύσματα έχουν μικρές αλλαγές στα πρόσημα.

Για τη μέθοδο Tikhonov, καθώς και για μια κατάλληλα επιλεγμένη τιμή του λ , όλες οι σημαντικές συνιστώσες του SVD βρίσκονται περίπου στον ίδιο υπόχωρο χαμηλών διαστάσεων (low-dimensional).

Για προβλήματα μεγάλης κλίμακας είναι μη αποδοτικό ή ανέφικτο να υπολογιστεί ο SVD του A . Αλλά από την SVD ανάλυση γνωρίζουμε ότι η κανονικοποιημένη λύση θα πρέπει να «κυριαρχείται» από στοιχεία χαμηλών συχνοτήτων. Πράγματι, μπορεί να είμαστε σε θέση να βρούμε ένα σύνολο διανυσμάτων βάσης $\mathbf{w}_1, \dots, \mathbf{w}_k$, που έχουν τα ίδια συνολικά χαρακτηριστικά με τα μοναδιαία, δηλαδή που κυριαρχούνται από τα στοιχεία χαμηλών συχνοτήτων.

Ας υποθέσουμε τώρα ότι έχουμε επιλέξει την κατάλληλη ορθοκανονική βάση ενός διακριτού συνημιτονοειδούς μετασχηματισμού (DCT- Discrete Cosine Transformation) ($\mathbf{w}_1, \dots, \mathbf{w}_k$).

Έστω λοιπόν ότι έχει επιλεχθεί μία κατάλληλη βάση διανυσμάτων, η οποία δομεί έναν υπόχωρο χαμηλής διάστασης:

$$\mathbf{W}_k = (\mathbf{w}_1, \dots, \mathbf{w}_k) \in \mathbf{R}^{n \times k}$$

Εάν το προς επίλυση πρόβλημα διατυπωθεί ως πρόβλημα ελαχίστων τετραγώνων τότε μπορεί να περιγραφεί από την ακόλουθη σχέση:

$$\min \|\mathbf{Ax} - \mathbf{b}\|_2 \quad \text{με } \mathbf{x} \in \mathbf{W}_k = \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\} \quad (5.2.1)$$

Επαναδιατυπώνοντας το παραπάνω πρόβλημα εισάγοντας έναν νέο άγνωστο της μορφής $x = W_k y, y \in \mathbb{R}^k$ οδηγούμαστε σε ένα νέο πρόβλημα ελαχίστων τετραγώνων (συναρτήσει του νέου αγνώστου) της μορφής:

$$x^{(k)} = W_k y^{(k)}, \quad y^{(k)} = \arg \min_y \|(A W_k) y - b\|_2 \quad (5.2.2)$$

Με βάση την πιο πάνω διατύπωση, μπορεί κανείς να ισχυριστεί ελαχίστων τετραγώνων ως προς y , αποτελεί πρόβλημα προβολής, γιατί προκύπτει προβάλλοντας το αρχικό πρόβλημα στον υπόχωρο k -διάστασης $\text{span}\{w_1, \dots, w_k\}$. Μάλιστα, αν το k είναι επαρκώς μεγάλο, τότε δύναται να υπολογιστεί αναλυτικά ο πίνακας:

$$A W_k \in \mathbb{R}^{n \times k} \quad (5.2.3)$$

και στη συνέχεια να λυθεί το πρόβλημα προβολής (πρόβλημα ελαχίστων τετραγώνων ως προς y).

ότι το πρόβλημα

Γίνεται εύκολα αντιληπτό πως το σημαντικότερο πλεονέκτημα των μεθόδων προβολής είναι ότι με μία κατάλληλα επιλεγμένη βάση διανυσμάτων οι απαιτούμενοι υπολογισμοί μπορούν να εκτελεστούν ταχύτατα.

Από την άλλη, το κυριότερο μειονέκτημα αυτής της προσέγγισης είναι ότι η εν λόγω βάση διανυσμάτων δεν μπορεί να είναι πάντα προσαρμοσμένη στο προς επίλυση πρόβλημα.

5.3 Μέθοδοι Υπόχωρου Krylov

Η βασική ιδέα των μεθόδων αυτών είναι η αξιοποίηση ενός κατάλληλου υποχώρου Krylov διάστασης k , ο οποίος έχει μια σειρά από επιθυμητές ιδιότητες, εκφράζεται συναρτήσει των A και b και ορίζεται ως εξής:

$$K_k \equiv \text{span} \left\{ A^T b, (A^T A) A^T b, (A^T A)^2 A^T b, \dots, (A^T A)^{k-1} A^T b \right\} \quad (5.3.1)$$

Αν και ο υπόχωρος που ορίζεται από την παραπάνω σχέση εμπεριέχει σημαντικές πληροφορίες για το πρόβλημα και μπορεί να διαδραματίσει το ρόλο του υπόχωρου W_k όπως ορίστηκε στη μέθοδο Προβολής της προηγούμενης παραγράφου, ωστόσο εμπεριέχει διανύσματα τα οποία δε διευκολύνουν τους υπολογισμούς στην πράξη.

Για να επιτευχθεί μία καλύτερη μαθηματική αναπαράσταση ορθοκανονικοποιούνται τα διανύσματα του Kb ως εξής:

$$\left\{ \begin{array}{l} w_1 \leftarrow A^T b; \quad w_1 \leftarrow \frac{w_1}{\|w_1\|_2}; \\ w_2 \leftarrow A^T A w_1; \quad w_2 \leftarrow w_2 - w_1^T w_2 w_1; \quad w_2 \leftarrow \frac{w_2}{\|w_2\|_2}; \\ w_3 \leftarrow A^T A w_2; \quad w_3 \leftarrow w_3 - w_1^T w_3 w_1; \quad w_3 \leftarrow w_3 - w_2^T w_3 w_2; \quad w_3 \leftarrow \frac{w_3}{\|w_3\|_2}; \end{array} \right.$$

Υποθέτοντας ότι ισχύει η διακριτή συνθήκη Picard, η ανωτέρω βάση διανυσμάτων περιγράφει επαρκώς την πληροφορία που εμπεριέχεται στα δεξιά ιδιάζοντα διανύσματα v_1, v_2, \dots

Τα διανύσματα $A^T b, (A^T A) A^T b, (A^T A)^2 A^T b, \dots, (A^T A)^{k-1} A^T b$, τα οποία ορίζουν τον υπόχωρο Krylov, κυριαρχούνται ολοένα και περισσότερο από τις συνιστώσες του v_1 , ενώ τα ορθοκανονικά διανύσματα της βάσης w_1, \dots, w_k έχουν διαφορετική συμπεριφορά, αφού η πληροφορία που εμπεριέχουν κυριαρχείται από μικρό αριθμό ιδιάζοντων διανυσμάτων v_j με δείκτη $j \approx i$.

Συνεπώς, ο υπόχωρος Krylov αποτελεί μία καλή περίπτωση υπόχωρου για να χρησιμοποιηθεί στις μεθόδους προβολής, υπό την προϋπόθεση βέβαια ότι τα προς επίλυση προβλήματα μπορούν να διατυπωθούν ως ασθενώς ορισμένα αντίστροφα προβλήματα και ικανοποιούν τη διακριτή συνθήκη Picard.

Κεφάλαιο 6

6. Αποθόλωση Εικόνας

Η αποθόλωση εικόνας (image deblurring) είναι η διαδικασία της ανακατασκευής μιας επεξεργασίας «καθαρισμού» (blurred) ψηφιακής εικόνας, βασισμένη σ' ένα μαθηματικό μοντέλο. Το πρόβλημα αυτό περιγράφεται ως μια ολοκληρωτική εξίσωση Fredholm πρώτου είδους (Κεφάλαιο 2.1):

$$\int_0^1 \int_0^1 K(s,t) f(t) dt_1 dt_2 = g(s), \quad s \in [0,1] \times [0,1] \quad (6.1)$$

όπου $f(t)$, $g(s)$ οι συναρτήσεις για την ακριβή μορφή της εικόνας (sharp) και της μη αποσαφηνισμένης (blurred) αντίστοιχα. Οι συγκεκριμένες συναρτήσεις είναι διπαραμετρικές και λαμβάνουν τιμές στο διάστημα $[0,1] \times [0,1]$.

Έστω δύο πίνακες X και B μεγέθους $n \times n$ που εκφράζουν την ακριβή εικόνα και την καταγεγραμμένη θολή-θορυβώδη εικόνα αντίστοιχα.

Έστω X_{ij} και B_{ij} τα στοιχεία του πίνακα που εκφράζουν τα εικονοστοιχεία (pixels) των δύο εικόνων.

Επίσης, έστω τα διανύσματα x και b , με μήκος $n = N^2$, που αποτελούνται από τη διαδοχική παράταξη των στηλών των πινάκων X και B :

$$x_l = X_{ij} \text{ και } b_l = B_{ij}, \quad \text{όπου } l = (j - 1)N + 1 \quad (6.2)$$

Τότε ένα σύνηθες καλό μοντέλο για την εισαγωγή θορύβου στην εικόνα είναι της μορφής $Ax = b$, όπου A πίνακας $N^2 \times N^2$.

6.1 Ο ρόλος της Συνάρτησης Διασποράς Σημείου

Για να καταλήξουμε στο μοντέλο $Ax = b$, πρέπει πρώτα να εξετάσουμε τη Συνάρτηση Διασποράς Σημείου (Point Spread Function) για την εισαγωγή θορύβου στην εικόνα, η οποία υποδεικνύει πως η φωτεινότητα ενός εικονοστοιχείου της ακριβούς εικόνας X , εκτείνεται σε έναν αριθμό εικονοστοιχείων στη μη αποσαφηνισμένη εικόνα.

Συγκεκριμένα θα ορίσουμε τον πίνακα P , που ονομάζεται Συνάρτηση Διασποράς Σημείου (PSF), ως την απεικόνιση του μοναδικού άσπρου εικονοστοιχείου της ακριβούς

εικόνας στο σύνολο των εικονοστοιχείων της μη αποσαφηνισμένης εικόνας , όπως φαίνεται στην Εικόνα 6.1.1.



Εικόνα 6.1.1-Η Συνάρτηση Διασποράς Σημείου είναι η εικόνα του μοναδικού φωτεινού εικονοστοιχείου.

Θεωρούμε ότι τα στοιχεία του \mathbf{P} αθροίζουν στη μονάδα, γεγονός που συνεπάγεται τη διατήρηση της ενέργειας κατά τη διαδικασία της διατάραξης-“θόλωσης” της εικόνας.

Η Συνάρτηση Διασποράς Σημείου, μπορεί να προκύψει από ένα μαθηματικό μοντέλο «θόλωσης» διαδικασιών ή από πραγματικές μετρήσεις της συγκεκριμένης συνάρτησης.

Οι διαστάσεις του υπολογιζόμενου ή μετρήσιμου πίνακα PSF είναι συχνά πολύ μικρότερες από N . Αυτό οφείλεται στο γεγονός ότι η εξάπλωση της φωτεινότητας λαμβάνει χώρα μόνο σε μια μικρή περιοχή γύρω από το λευκό εικονοστοιχείο.

Από την υπόθεση για χωρικές αναλλοίωτες και τον ορισμό του \mathbf{P} προκύπτει ότι αν $X_{ij} = 1$ για ένα συγκεκριμένο δείκτη (i, j) και διάφορο του μηδενός, τότε η θολή εικόνα \mathbf{B} δίνεται από τη σχέση:

$$B_{kl} = P_{k-i, l-j}, \quad i, j, k, l = 1, \dots, N \quad (6.1.1)$$

Με αυτή τη σύμβαση των δεικτών, τα στοιχεία του PSF πίνακα \mathbf{P} επιτρέπεται να έχουν μηδέν ή αρνητικούς δείκτες, και το κέντρο του \mathbf{P} είναι το στοιχείο P_{00} .

Πίσω από την ολοκληρωτική διατύπωση της εξίσωσης κρύβεται η παραδοχή ότι το «blurring» είναι μια γραμμική διαδικασία. Αυτό συνεπάγεται ότι για κάθε εικόνα \mathbf{X} , η αντίστοιχη θολωμένη εικόνα \mathbf{B} είναι το άθροισμα των θολωμένων εικόνων όλων των εικονοστοιχείων του \mathbf{X} , όπως απεικονίζεται στην Εικόνα 6.1.1.

Ως εκ τούτου, τα εικονοστοιχεία της \mathbf{X} και της \mathbf{B} συνδέονται μέσω μιας διακριτής συνέλιξης (convolution) δύο διαστάσεων που λαμβάνει τη μορφή:

$$\sum_{i,j=1}^N P_{k-i,l-j} X_{ij} = B_{kl}, \quad k,l = 1, \dots, N \quad (6.1.2)$$

Επομένως, η θολή εικόνα είναι το άθροισμα όλων των θολωμένων εικονοστοιχείων (pixels).

Δεδομένου ότι το μοντέλο (6.1.2) περιγράφει μία γραμμική σχέση μεταξύ των στοιχείων μιας ακριβούς και μιας θολωμένης εικόνας, η γραμμική σχέση διατηρείται και όταν τα στοιχεία των \mathbf{X} και \mathbf{B} αναδιατάσσονται εντός των φορέων \mathbf{x} και \mathbf{b} σύμφωνα με την (6.1).

Ως εκ τούτου, υπάρχει μια μοναδική σχέση μεταξύ των στοιχείων του PSF πίνακα \mathbf{P} και των στοιχείων του πίνακα \mathbf{A} στο γραμμικό σύστημα $\mathbf{Ax} = \mathbf{b}$.

Λόγω του μεγέθους του πίνακα συντελεστών \mathbf{A} , είναι απαραίτητο να αποφύγουμε το σχηματισμό του. Αντί αυτού, πρέπει να είμαστε σε θέση να εκτελέσουμε όλους τους υπολογισμούς με τον \mathbf{A} μέσω της (6.2) γνωρίζοντας τον PSF πίνακα \mathbf{P} .

6.2 PSF Πίνακες Πρώτου Βαθμού και Γρήγοροι Αλγόριθμοι

Για ορισμένες συναρτήσεις παίρνουμε πίνακες \mathbf{A} με περισσότερη «δομή», η οποία μας επιτρέπει να υπολογίσουμε τον SVD του. Εδώ συζητάμε για πίνακες \mathbf{P} πρώτου βαθμού, δηλαδή πίνακες που μπορούν να γραφτούν ως

$$P_{ij} = \mathbf{a} \bar{\mathbf{a}}^T \quad (6.2.1)$$

όπου τα \mathbf{a} και $\bar{\mathbf{a}}$ είναι δύο διανύσματα των οποίων δείκτες, όπως αυτοί των στοιχείων του \mathbf{P} , μπορεί να είναι μηδέν και αρνητικοί. Μπορούμε επίσης να υποθέσουμε ότι και το \mathbf{a} και το $\bar{\mathbf{a}}$ είναι συμμετρικά γύρω από το "μεσαίο" στοιχείο τους, \mathbf{a}_0 και $\bar{\mathbf{a}}_0$ αντίστοιχα, δηλαδή, $\mathbf{a}_i = \mathbf{a} - i$ και $\bar{\mathbf{a}}_i = \bar{\mathbf{a}} - i$ για όλες τις τιμές του i .

Ο PSF πίνακας P που προκύπτει είναι διπλά συμμετρικός, δηλαδή είναι συμμετρικός και σε κάθε γραμμή και σε κάθε στήλη σε σχέση με το κέντρο του.

Αν επανεξετάσουμε την συνέλιξη δύο διαστάσεων στην (6.1.2) και αντικαταστήσουμε τα στοιχεία του P από την έκφραση (6.2.1), τότε για $k, l = 1, \dots, N$ παίρνουμε την σχέση:

$$B_{kl} = \sum_{i,j=1}^N a_{k-i} \bar{a}_{i-j} X_{ij} = (a_{k-1}, \dots, a_{k-N}) X \begin{pmatrix} \bar{a}_{i-j} \\ \vdots \\ \bar{a}_{i-N} \end{pmatrix} \quad (6.2.2)$$

Άρα, άμεσα προκύπτει ότι μπορούμε να γράψουμε τη θολωμένη εικόνα B , ως

$$B = A \times \bar{A}^T \quad (6.2.3)$$

όπου A και είναι \bar{A} δύο συμμετρικοί πίνακες που καθορίζονται ως:

$$A = \begin{pmatrix} a_0 & a_{-1} & a_{-(N-1)} \\ a_1 & a_0 & a_{-(N-2)} \\ \vdots & \vdots & \vdots \\ a_{N-1} & a_{N-2} & a_0 \end{pmatrix}, \quad \bar{A} = \begin{pmatrix} \bar{a}_0 & \bar{a}_{-1} & \bar{a}_{-(N-1)} \\ \bar{a}_1 & \bar{a}_0 & \bar{a}_{-(N-2)} \\ \vdots & \vdots & \vdots \\ \bar{a}_{N-1} & \bar{a}_{N-2} & \bar{a}_0 \end{pmatrix} \quad (6.2.4)$$

Αυτό αντιστοιχεί στην υπόθεση του μοντέλου μας για μηδενικές συντοριακές συνθήκες.

Για να καταλήξουμε σε ένα γρήγορο αλγόριθμο ομαλοποίησης προχωρούμε όπως και πριν, χρησιμοποιώντας ότι A_{BC} και ένα \bar{A}_{BC} είναι και οι δύο διαγώνιοι DCT πίνακες του W_n

$$A_{BC} = W_n D W_n^T, \bar{A}_{BC} = W_n \bar{D} W_n^T \quad (6.2.5)$$

που οδηγεί στις σχέσεις:

$$B = W_n D W_n^T X W_n \bar{D} W_n^T \Leftrightarrow X = W_n (D^{-1} (W_n^T B W_n) \bar{D}^{-1}) W_n^T \quad (6.2.6)$$

Εδώ αναγνωρίζουμε τον πίνακα:

$$\hat{B} = W_n^T B W_n \quad (6.2.7)$$

ως έναν δύο διαστάσεων DCT του B, ενώ τον πίνακα:

$$X = W_n \hat{C} W_n^T \quad (6.2.8)$$

ως έναν DCT αντίστροφο δύο διαστάσεων του

$$\hat{C} = D^{-1} \hat{B} D^{-1} \quad (6.2.9)$$

6.3 Αποσυνέλιξη και Συνέλιξη

Λαμβάνοντας υπόψη ένα πρόβλημα με δεδομένα με θόρυβο, και λαμβάνοντας υπόψη μια συγκεκριμένη μέθοδο κανονικοποίησης, μας ενδιαφέρει να μελετήσουμε τη συνέλιξη, δηλαδή πόση λεπτομέρεια μπορούμε να ανακτήσουμε αξιόπιστα στην κανονικοποιημένη λύση. Η ανάλυση της συνέλιξης είναι στενά συνδεδεμένη με τον αριθμό των συστατικών SVD που μπορούν να ανακτηθούν από τα θορυβώδη δεδομένα, καθώς και με τις φασματικές ιδιότητες των μοναδικών διανυσμάτων.

Γνωρίζουμε ότι οι φασματικές ιδιότητες μιας κανονικοποιημένης λύσης καθορίζονται από τα μοναδικά διανύσματα που λαμβάνουν μέρος στη λύση. Αυτό είναι προφανές για τη μέθοδο TSVD αλλά ισχύει για οποιαδήποτε μέθοδο περιλαμβάνει φιλτραρισμένη επέκταση στην SVD βάση. Ως εκ τούτου για το πρόβλημα της αποσυνέλιξης, εάν συμπεριλάβουμε τα πρώτα k, w_1, \dots, w_k DCT διανύσματα της βάσης τότε το μέγεθος της ακριβούς λεπτομέρειας που μπορούμε να ανακτήσουμε, καθορίζεται από το πλάτος του ταλαντωτή, με τη μεγαλύτερη συχνότητα, του διανύσματος w_k .

6.4 Τομογραφία στις 2 Διαστάσεις

Θεωρούμε ένα δισδιάστατο (2D) πρόβλημα στο διάστημα $\Omega = [0,1] \times [0,1]$, στο οποίο μας δίνεται μια άγνωστη συνάρτηση $f(t) = f(t_1, t_2)$, την οποία θέλουμε να ανακατασκευάσουμε. Υποθέτουμε ότι η συγκεκριμένη συνάρτηση εκφράζει την απόσβεση ενός σήματος το οποίο διαπερνά ένα απειροελάχιστα μικρό μέρος D_t μιας ακτίνας στη θέση t , έτσι ώστε η απόσβεση να είναι ανάλογη με το γινόμενο $f(t)D_t$. Τα

δεδομένα στην Τομογραφία αποτελούνται από μετρήσεις της απόσβεσης των σημάτων, τα οποία διαπερνούν την περιοχή Ω κατά μήκος μιας ευθείας γραμμής.

Στο παραπάνω μοντέλο, η b_i παρατήρηση i για $i = 1, \dots, m$, εκφράζει την απόσβεση ενός σήματος που διαπερνά την περιοχή Ω κατά μήκος μιας ευθείας γραμμής, στην οποία θα αναφερόμαστε ως ακτίνα i . Όλα τα σημεία t_i στην ακτίνα i δίνονται από τη σχέση,

$$t^i(\tau) = t^{i,0} + \tau d^i, \quad \tau \in R \quad (6.4.1)$$

όπου t_i είναι ένα αυθαίρετο σημείο επί της ακτίνας, και d_i μοναδιαίο διάνυσμα που δείχνει την κατεύθυνση της ακτίνας. Λόγω της παραπάνω υπόθεσης, η απόσβεση είναι ανάλογη με το ολοκλήρωμα της συνάρτησης $f(t)$ κατά μήκος της ακτίνας. Συγκεκριμένα, για την παρατήρηση i η απόσβεση που συνδέεται με την ακτίνα i δίνεται από τον εξής τύπο:

$$b_i = \int_{-\infty}^{\infty} f(t^i(\tau)) d\tau, \quad i = 1, \dots, m \quad (6.4.2)$$

όπου το $d\tau$ ορίζει την ολοκλήρωση κατά μήκος της ακτίνας.

Η Διακριτοποίηση του παραπάνω προβλήματος μπορεί να πραγματοποιηθεί με τη διαίρεση της περιοχής Ω σε ένα $N \times N$ πίνακα εικονοστοιχείων, όπου σε κάθε εικονοστοιχείο (k, l) υποθέτουμε ότι η συνάρτηση $f(t)$ είναι μια σταθερά f_{kl} :

$$f_t = f_{kl}, \quad \text{για } t_1 \in L_k \text{ και } t_2 \in L_l \quad (6.4.3)$$

και ορίζουμε το διάστημα $L_k = \left[\frac{k-1}{N}, \frac{k}{N} \right]$ $k = 1, \dots, N$ (όμοια για L_l), με την παραδοχή ότι η $f(t)$ είναι τμηματικά συνεχής. Η έκφραση για τη μέτρηση k λαμβάνει την απλούστερη μορφή, για $i = 1, \dots, m$:

$$b_i = \sum_{(k,l) \in \text{ακτίνα } i} f_{kl} \Delta L_{kl}^i \quad (6.4.5)$$

Όπου ΔL_{kl}^i το μήκος ακτίνας i στο εικονοστοιχείο (k, l) .

Η παραπάνω εξίσωση είναι στην πραγματικότητα, ένα γραμμικό σύστημα εξισώσεων με N^2 αγνώστους f_{kl} . Για να καταλήξουμε σε μια πιο βολική έκφραση για αυτό το σύστημα, εισάγουμε το διάνυσμα x μήκους $n = N^2$ τα στοιχεία του οποίου είναι η οι άγνωστες τιμές f_{kl} , διατεταγμένες ως εξής:

$$x_j = f_{kl}, \quad j = (l-1)N + k \quad (6.4.6)$$

Είναι σαφές ότι υπάρχει γραμμική σχέση μεταξύ των δεδομένων b_k και των αγνώστων του διανύσματος x , πράγμα που σημαίνει ότι μπορούμε να γράψουμε:

$$b_i = \sum_{j=1}^N a_{ij} x_j, \quad i = 1, \dots, m \quad (6.4.7)$$

Με βάση τους παραπάνω ορισμούς προκύπτει ότι θα καταλήξουμε σε ένα γραμμικό σύστημα εξισώσεων $Ax = b$ με έναν $m \times n$ πίνακα του οποίου τα στοιχεία δίνονται από τη σχέση:

$$a_{ij} = \begin{cases} \Delta L_{kl}^i, & (k, l) \in \text{ακτίνα } i \\ 0, & \text{αλλού} \end{cases} \quad (6.4.8)$$

Υπενθυμίζουμε ότι ο δείκτης i δηλώνει την παρατήρηση i (που αντιστοιχεί στην ακτίνα i) και ο δείκτης j δηλώνει τον αριθμό των εικονοστοιχείων στο διατεταγμένο διάνυσμα x όπου $j = (l-1)N + k$. Ο πίνακας A είναι αραιός και ο αριθμός των μη μηδενικών στοιχείων σε οποιαδήποτε σειρά είναι πάνω από $2N$.

Στο πρόβλημα της αποσαφήνισης εικόνας που περιγράφεται στο κεφάλαιο αυτό, δεν μπορεί να γίνει ακριβής αναφορά στον πίνακα συντελεστών, εν αντιθέσει με το πρόβλημα της Τομογραφίας όπου είναι συχνά εφικτό να σχηματιστεί λεπτομερώς ο πίνακας A .

6.5 CGLS Αλγόριθμος

Μία από τις πλέον ευσταθείς παραλλαγές του αλγορίθμου CG είναι ο αλγόριθμος CGLS (Conjugate Gradient Least Squares). Σε αντίθεση με τις διατυπώσεις των μεθόδων προβολής που αναπτύχθηκαν στην προηγούμενη ενότητα, η μέθοδος CGLS δίνει τη δυνατότητα περιγραφής της λύσης $x(k)$ ως φιλτραρισμένο ανάπτυγμα SVD. Η λύση $x(k)$ μπορεί να διατυπωθεί ως γραμμικός συνδυασμός των διανυσμάτων βάσης Krylov. Συνεπώς, υπάρχουν σταθερές c_1, c_2, \dots, c_k τέτοιες ώστε:

$$x^{(k)} = c_1 A^T b + c_2 (A^T A) A^T b + \dots + c_k (A^T A)^{k-1} A^T b \quad (6.5.1)$$

Εισάγοντας το ανάπτυγμα SVD του A υπολογίζουμε την ποσότητα $A^T A$ ως εξής:

$$A^T A = V \Sigma U^T U \Sigma V^T = V \Sigma^2 V^T \quad (6.5.2)$$

Με αποτέλεσμα,

$$x^{(k)} = (c_1 + c_2 V \Sigma^2 V^T + \dots + c_k V \Sigma^{2(k-1)} V^T) V \Sigma U^T b = V \Phi^{(k)} \Sigma^{-1} U^T b \quad (6.5.3)$$

Όπου,

$$\Phi^{(k)} = \text{diag}(\varphi_1^{(k)}, \dots, \varphi_n^{(k)}) \quad (6.5.4)$$

ένας διαγώνιος πίνακας αποτελούμενος από τους φιλτραρισμένους CGLS παράγοντες της μορφής:

$$\varphi_i^{(k)} = c_1 \sigma_i^2 + c_2 \sigma_i^4 + \dots + c_k \sigma_i^{2k}, i = 1, 2, \dots, n \quad (6.5.5)$$

Κεφάλαιο 7

7. Εφαρμογή του αλγορίθμου CGLS

Στην ενότητα αυτή θα ασχοληθούμε με την εφαρμογή των αλγορίθμων που παρουσιάστηκαν στις προηγούμενες ενότητες και συγκεκριμένα των CGLS αλγορίθμων.

Με την υλοποίηση των αλγορίθμων αυτών, θα καταφέρουμε να δούμε και να συγκρίνουμε τις διαφορετικές περιπτώσεις, με τις οποίες μπορούμε να αναλύσουμε μια εικόνα, καθώς και να την ανακατασκευάσουμε.

Η διαδικασία που πρέπει να ακολουθήσουμε θα γίνει με τη βοήθεια του μαθηματικού προγράμματος MatLab και χρησιμοποιώντας κατάλληλους αλγόριθμους, θα εκτελέσουμε τις εντολές μας, έτσι ώστε να δούμε πως μπορούμε να ανακατασκευάσουμε μια εικόνα.

Βήματα Ανακατασκευής και ανάλυση εικόνας:

Στην προκειμένη περίπτωση αυτό που διαθέτουμε είναι μια φωτογραφία της Πάρου το καλοκαίρι, με σκοπό να έχουμε μια καθαρή φωτογραφία έτσι ώστε να μπορεί να αναλυθεί σωστά.

Πιο συγκεκριμένα, ακολουθούμε μια κατασκευαστική διαδικασία η οποία ακολουθεί την εξής σειρά:

- ✓ Θα γνωρίζουμε την αρχική μας λύση, δηλαδή την αρχική μας εικόνα.
- ✓ Έπειτα, θα προσθέσουμε κάποιου είδους αναταραχή έτσι ώστε να αλλάξει η λύση μας από την αρχική, γεγονός το οποίο θα προκαλέσει θόλωμα στην αρχική εικόνα.
- ✓ Τέλος, με τις κατάλληλες εντολές θα προκύψει και πάλι η αρχική μας εικόνα.

Αρχικά, θα «φορτώσουμε» στο MatLab τη φωτογραφία από το λιμανάκι της Πάρου όπως βλέπουμε πιο κάτω (Εικόνα 7.1).



Εικόνα 7.1-Φωτογραφία «το λιμανάκι της Πάρου»

Όμως, η εικόνα τελικά που τέθηκε σε επεξεργασία με το πρόγραμμα, έπρεπε να είναι ασπρόμαυρη και τετράγωνη σε διαστάσεις. Έτσι, η εικόνα που τελικά επεξεργάστηκε το MatLab και η οποία θα θεωρήσουμε ότι είναι η αρχική μας εικόνα, είναι η Εικόνα 7.2.

Sharp Image



Εικόνα 7.2- Αρχική Εικόνα

Για να επεξεργαστούμε όμως την εικόνα 7.1 πρέπει η εικόνα μας να είναι σε μορφή double.

Οι εντολές για την εισαγωγή εικόνας, οι οποίες την μετατρέπουν σε ασπρόμαυρη, τετραγωνική και σε double format είναι οι εξής:

```

RGB=imread('Paros.jpg');
Clear caption map;
ImageToScale=rgb2gray (RGB);
X=double(imageToScale);
X=X(:,1162:4149);

```

Συγκεκριμένα, οι διαστάσεις της καινούριας εικόνας είναι 2988 pixels. Επίσης, ο τετραγωνικός πίνακας που περιγράφει την εικόνα μας θα πρέπει να μετατραπεί σε έναν πίνακα-στήλη, όπως έχει περιγραφεί στη σχέση 6.1. Με κατάλληλη εντολή του αλγορίθμου, φτιάξαμε τη διάσταση της εικόνας και την κάναμε τετραγωνική. Συγκεκριμένα, η ακριβής λύση με $N=2988$, γράφτηκε ως $X_exact=X(:,1:N)$; και έπειτα $x_exact=X_exact(:)$; για να «καθετοποιήσουμε» τον πίνακα μας.

Στη συνέχεια, δημιουργούμε μια συνάρτηση διασποράς σημείου P . Για να λάβουμε τη θολωμένη εικόνα έπρεπε να προσθέσουμε στις εντολές μας το «θόρυβο», δηλαδή $b=b_exact+e$, έτσι ώστε με την εντολή $B=reshape(b,N,N)$ να δημιουργήσουμε μια αλλοιωμένη εικόνα. Το σφάλμα e το ορίσαμε κατάλληλα με τις εντολές: $e=randn(size(b_exact))$, $e=e/norm(e)$ και τέλος με την εντολή $e=0.01*norm(b_exact)*e$ η οποία θα καθορίσει το «θόρυβο» που θα βάλουμε στην εικόνα μας, ανάλογα με το ποσοστό που θα πολλαπλασιάσουμε, δηλαδή με 1%,3% ή 5% κ.ο.κ. Η εικόνα που προκύπτει είναι η εικόνα 7.3.

Ακολούθως, καλέσαμε τον αλγόριθμο *CGLS* για την ανακατασκευή της αρχικής εικόνας από τη θολωμένη. Αποφασίσαμε να καλέσουμε 150 επαναλήψεις για να δούμε τι συμβαίνει κατά τη διάρκειά τους. Η εντολή που καλέσαμε είναι η $z = cgls(P, b, 150)$ η οποία καλεί τον *CGLS* αλγόριθμο για την ανακατασκευή της εικόνας μας. Ως εκ τούτου, το αποτέλεσμα της ανακατασκευής είναι η εικόνα 7.4.

Blurred Image



Εικόνα7.3- "Θολωμένη" Εικόνα

Reconstructed Image



Εικόνα7.4-Ανακατασκευασμένη Εικόνα

Μιας και ο CGLS αλγόριθμος ανήκει στην οικογένεια των αλγορίθμων που ημισυγκλίνουν, αυτό σημαίνει ότι περιμένουμε μέχρι κάποιο βήμα να συγκλίνει και από εκείνο το βήμα και μετά να αποκλίνει από την αρχική μας εικόνα. Έτσι, ξέροντας ποια είναι η αρχική μας εικόνα, υπολογίζουμε το σφάλμα που έχει η λύση που βρίσκει ο αλγόριθμος σε κάθε βήμα από την αρχική. Κάνουμε χρήση L_2 νόρμας.

Άρα οι εντολές που χρησιμοποιούμε για την εκτίμηση του σφάλματος ανακατασκευής είναι οι:

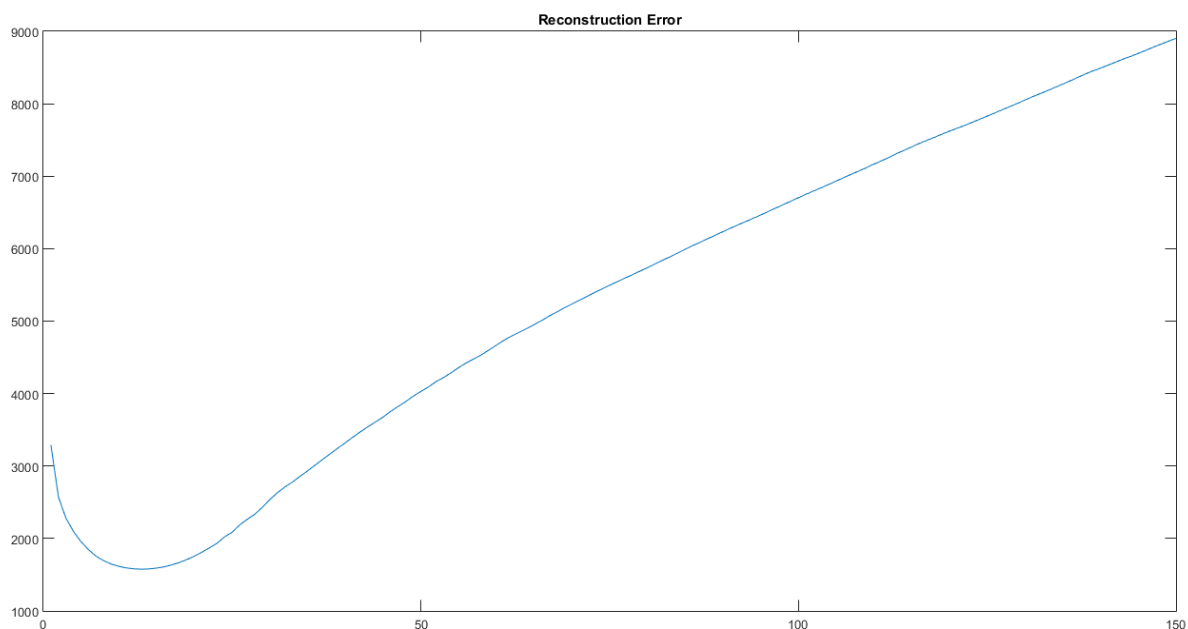
```
error=bsxfun(@minuw,z,x_exact) και norms=zeros(1,150) με βήμα x=1:150.
```

Τέλος, με τις εντολές `plot(x,norms)` και `[min,i]=min(norms)` λαμβάνουμε τη γραφική παράσταση (εικόνα 7.5) για την εκτίμηση σφάλματος του αλγορίθμου, σε σχέση με τα βήματα που έχει κάνει ο αλγόριθμος για εισαγωγή 1% θορύβου.

Στον οριζόντιο άξονα της γραφικής παράστασης είναι τα βήματα που κάνει ο αλγόριθμος και στον κατακόρυφο το σφάλμα που λαμβάνουμε μετά την ανακατασκευή της εικόνας (σε σχέση με την αρχική).

Το ελάχιστο της συνάρτησης αυτής το πετυχαίνουμε για 13 βήματα. Όπως παρατηρούμε και από το γράφημα, το μικρότερο σφάλμα λαμβάνεται στη δέκατη τρίτη επανάληψη του αλγορίθμου.

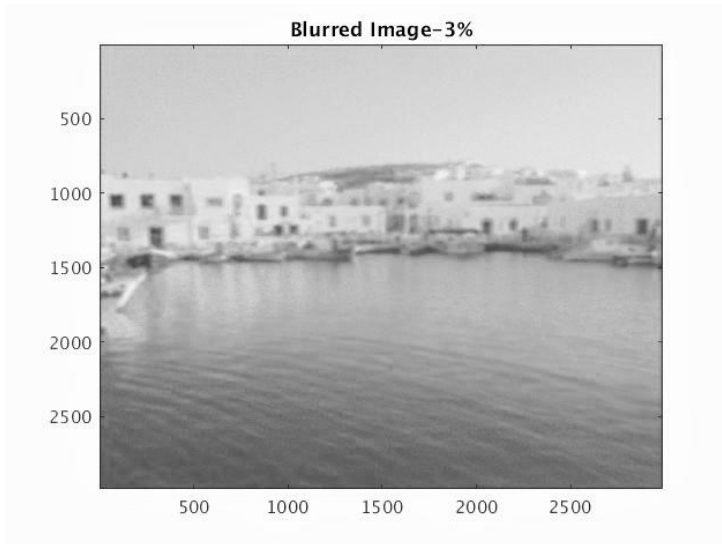
Πιο συγκεκριμένα, η εικόνα που παραθέσαμε αποτελεί τη λύση του αλγορίθμου που παρουσιάζει το μικρότερο δυνατό σφάλμα προτού αρχίσει να αποκλίνει ξανά.



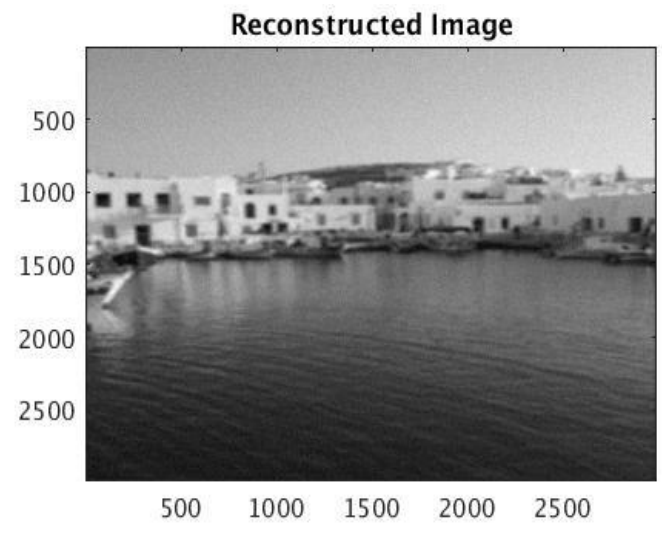
Εικόνα 7.5- Σφάλμα απόκλισης από αρχική Εικόνα

Ακολουθώντας την ίδια διαδικασία που χρησιμοποιήσαμε στο πιο πάνω παράδειγμα, τροποποιούμε στον κώδικα το ποσοστιαίο σφάλμα θορύβου από 1% σε 3% και 5% :

Για θόρυβο 3% χρειαστήκαμε 20 επαναλήψεις του αλγόριθμου για ανακατασκευή της εικόνας (Εικόνα 7.7)

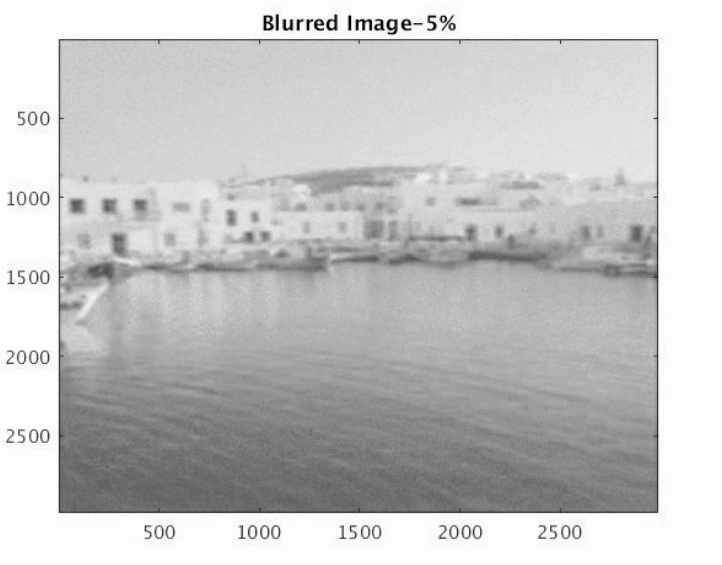


Εικόνα 7.6

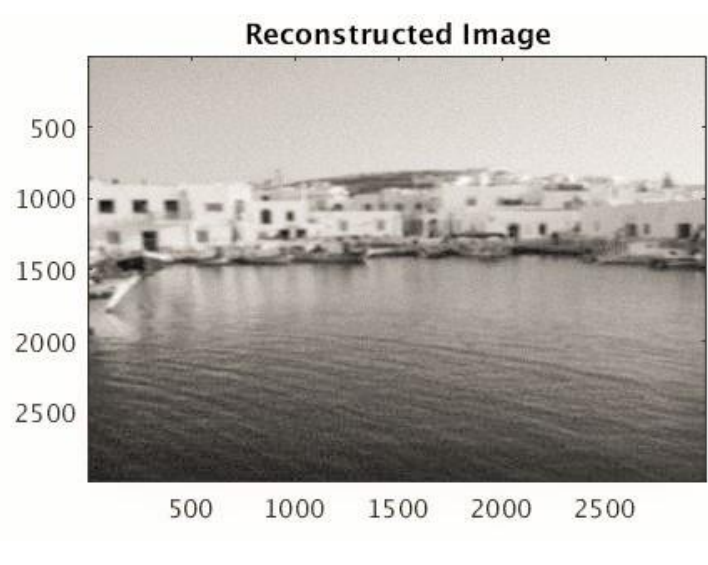


Εικόνα 7.7

Για θόρυβο 5% χρειαστήκαμε 50 επαναλήψεις του αλγόριθμου για την ανακατασκευή της εικόνας (εικόνα 7.9).

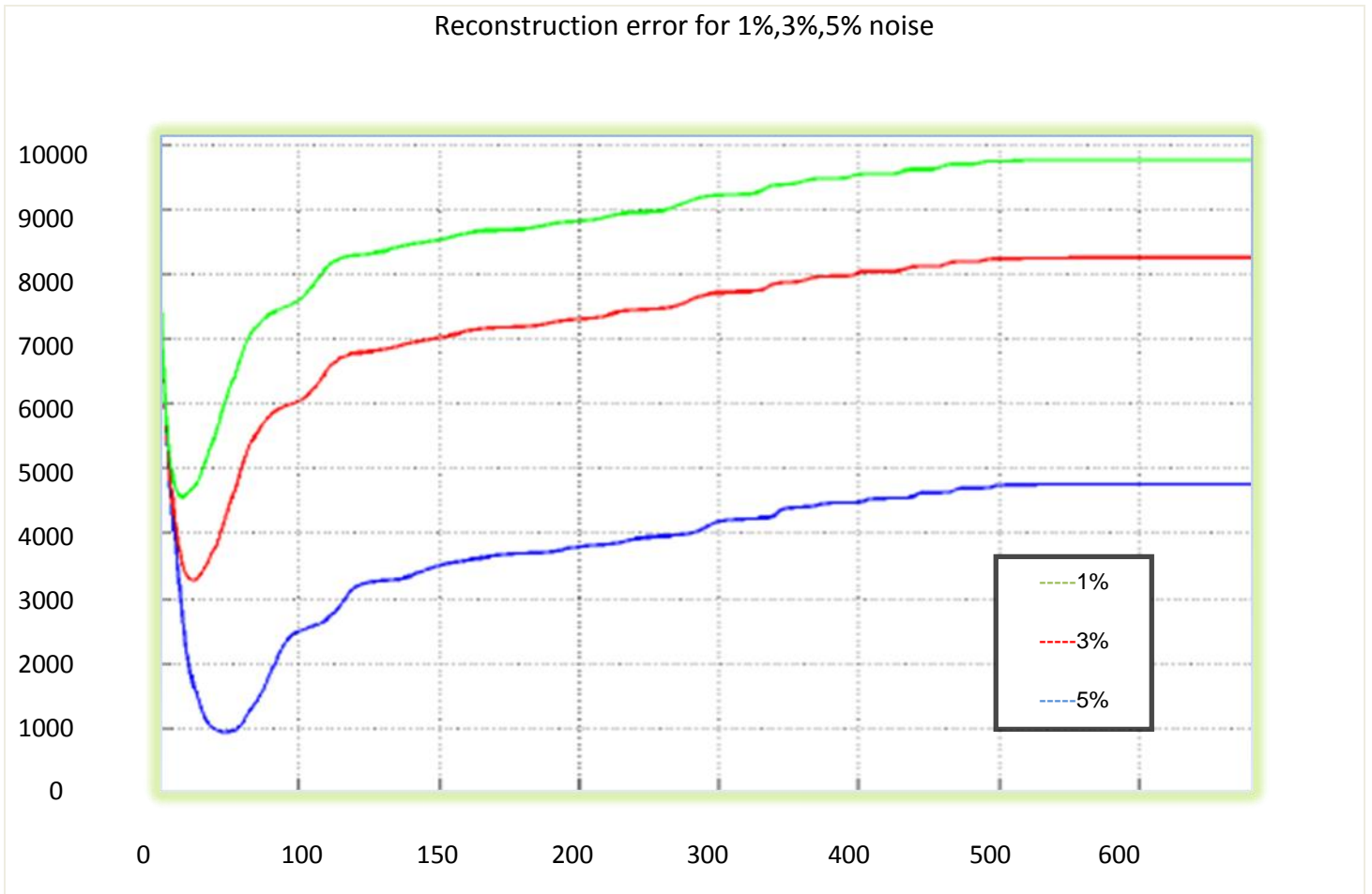


Εικόνα 7.8



Εικόνα 7.9

Πιο κάτω βλέπουμε το γράφημα με τα αποτελέσματα των τριών περιπτώσεων, των πιο πάνω παραδειγμάτων, όσον αφορά τα σφάλματα απόκλισης από την αρχική εικόνα.



Εικόνα 7.10 Σφάλματα απόκλισης από την αρχική Εικόνα

7.1 Ανάλυση αποτελεσμάτων – Σχολιασμός

Το πρόβλημα ανακατασκευής εικόνας (image deblurring) χρησιμοποιεί μία εικόνα – δείγμα, η οποία παράγεται από τη συνάρτηση του Matlab “blurred”.

Στην εικόνα 7.3 απεικονίζεται η εν λόγω εικόνα στη θολή, (blurred) μορφή και στην εικόνα 7.4 στην ακριβή ανακατασκευασμένη(reconstructed) μορφή της.

Στις εικόνες 7.6 και 7.7 απεικονίζονται ανακατασκευασμένα (deblurred) στιγμιότυπα της εικόνας – δείγματος, τα οποία προκύπτουν μετά την εφαρμογή της μεθόδου CGLS για 20 επαναλήψεις , για το Πρόβλημα -Μοντέλο στο οποίο έχει προστεθεί θόρυβος 3%.

Στις εικόνες 7.8 και 7.9 απεικονίζονται ανακατασκευασμένα (deblurred) στιγμιότυπα της εικόνας – δείγματος, τα οποία προκύπτουν μετά την εφαρμογή της μεθόδου CGLS για 50 επαναλήψεις , για το Πρόβλημα – Μοντέλο στο οποίο έχει προστεθεί θόρυβος 5%.

Παρατηρεί εύκολα κανείς ότι αρχικά, όσο αυξάνεται ο αριθμός επαναλήψεων η εικόνα πλησιάζει περισσότερο στην ανακατασκευασμένη(reconstructed) μορφή της, αλλά μετά από ένα αριθμό επαναλήψεων ο θόρυβος κυριαρχεί και αλλοιώνει την εικόνα.

Στην εικόνα 7.10 παρατηρούμε από το γράφημα, τα σφάλματα απόκλισης από την αρχική εικόνα. Το μικρότερο σφάλμα λαμβάνεται στην περίπτωση θορύβου 5%. Πιο συγκεκριμένα, η εικόνα που παραθέσαμε αποτελεί τη λύση του αλγορίθμου που παρουσιάζει το μικρότερο δυνατό σφάλμα προτού αρχίσει να αποκλίνει ξανά.

Επίλογος

Ολοκληρώνοντας τη μελέτη της εργασίας αυτής, με την εφαρμογή των αλγορίθμων CGLS συμπεραίνουμε ότι, τα αντίστροφα προβλήματα είναι από τη φύση τους μη καλά τοποθετημένα προβλήματα και ως εκ τούτου είναι πολύ δύσκολο να βρούμε μια ακριβή λύση .

Όπως βλέπουμε στο παράδειγμα του κεφαλαίου 7, προσπαθούμε να βρούμε τη λύση μας με τυχαίο τρόπο, δηλαδή με δοκιμές, έτσι ώστε να καταλήξουμε σε μια προσεγγιστική λύση. Αυτό αποτελεί απόδειξη ότι τα αντίστροφα προβλήματα είναι μη καλά τοποθετημένα.

Παρά το γεγονός αυτό, η σημαντικότητα των αντίστροφων προβλημάτων είναι αξιοσημείωτη στον επιστημονικό χώρο, διότι μια προσεγγιστική λύση, πιθανόν να οδηγήσει σε μια διεξοδική έρευνα, σαφώς με άλλες πιο εξειδικευμένες μεθόδους.

Ένα παράδειγμα πολύ ενδιαφέρον είναι από το χώρο της Ιατρικής Φυσικής η αξονική τομογραφία, η οποία μπορεί να απεικονίσει σε κάθετες τομές ολόκληρο το ανθρώπινο σώμα, χρησιμοποιώντας την ακτινοβολία Χ. Έτσι, απεικονίζεται το εσωτερικό του σώματος και επιτρέπεται στον εξεταστή να αναζητήσει βλάβες μέσα στα όργανα ή να εντοπίσει ανωμαλίες σε σημεία που είναι δύσκολο να εντοπιστούν με μια απλή ακτινογραφία.

Τέλος, η ελικοειδής αξονική τομογραφία, η οποία είναι μια τελευταία εξέλιξη της αξονικής τομογραφίας, παρέχει τη δυνατότητα ανακατασκευής των εικόνων, έτσι ώστε να έχουμε και τομές σε άλλα επίπεδα.

Βιβλιογραφία

- 1 .Hansen, P. C. (2010). *Discrete Inverse Problems Insight and Algorithms*. Society for Industrial and Applied Mathematics (SIAM).
2. Ν. Καδιανάκης & Σ.Καρανάσιος, *Γραμμική Άλγεβρα και Εφαρμογές*, Τομέας Μαθηματικών ΣΕΜΦΕ
3. A.Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer – Verlag
4. Γ.Παπαγεωργίου ,Χ.Τσίτουρας,*Αριθμητική Ανάλυση με εφαρμογές σε Matlab και Mathematica(4^η έκδοση)*, ΣΥΜΕΩΝ
5. R.Kress, *Linear Integral Equations(2nd edition)*, Springer.
6. D.Wright, *Introduction to Linear Algebra*, McGraw- Hill