Πανεπιστήμιο Κρήτης

ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ-ΜΑΘΗΜΑΤΙΚΟ

ΠΡΟΓΡΑΜΜΑ ΠΡΟΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΑΡΙΘΜΗΤΙΚΗ ΛΥΣΗ ΓΡΑΜΜΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΜΕ ΠΙΝΑΚΑ ΘΕΤΙΚΑ ΟΡΙΣΜΕΝΟ ΚΑΙ ΣΥΜΜΕΤΡΙΚΟ ΜΕ ΤΗ ΜΕΘΟΔΟ CONJUGATE GRADIENT

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΚΟΤΡΩΤΣΙΟΥ ΜΑΡΙΑ

Επιβλέπων καθηγητής: Ζουράρης Γεώργιος
Καθηγητής Πανεπιστημίου Κρήτης

Ηράκλειο, 2019

Πανεπιστήμιο Κρήτης

ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ-ΜΑΘΗΜΑΤΙΚΟ

ΠΡΟΓΡΑΜΜΑ ΠΡΟΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΑΡΙΘΜΗΤΙΚΗ ΛΥΣΗ ΓΡΑΜΜΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΜΕ ΠΙΝΑΚΑ ΘΕΤΙΚΑ ΟΡΙΣΜΕΝΟ ΚΑΙ ΣΥΜΜΕΤΡΙΚΟ ΜΕ ΤΗ ΜΕΘΟΔΟ CONJUGATE GRADIENT

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΚΟΤΡΩΤΣΙΟΥ ΜΑΡΙΑ

Επιβλέπων καθηγητής: Ζουράρης Γεώργιος
Καθηγητής Πανεπιστημίου Κρήτης

Την επιτροπή αξιολόγησης αυτής της πτυχιακής εργασίας αποτέλεσαν οι:

1. Ζουράρης Γεώργιος(Επιβλέπων)
2. Χατζηπαντελίδης Παναγιώτης
3. Χαρμανδάρης Βαγγέλης

Ηράκλειο, 2019

Πανεπιστήμιο Κρήτης, Τμήμα Μαθηματικών και Εφαρμοσμένων Μαθηματικών-Μαθηματικό

Κοτρώτσιου Μαρία

© 2019-Με την επιφύλαξη παντός δικαιώματος. All rights reserved.

ΕΥΧΑΡΙΣΤΙΕΣ

Πρώτα απ' όλα, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου Κύριο Γεώργιο Ζουράρη για την εμπιστοσύνη που μου έδειξε ως προς την ανάθεση της παρούσας πτυχιακής εργασίας , καθώς επίσης και για το αμέριστο ενδιαφέρον του κατά τη διάρκεια της εκπόνησής της.

Επιπλέον, θα ήθελα να ευχαριστήσω την οικογένειά μου, η οποία ήταν δίπλα μου σε κάθε μου βήμα και με βοήθησε με τον δικό της τρόπο σε ότι έχω πετύχει ως σήμερα, μαθαίνοντάς μου να θέτω στόχους και να τους πετυχαίνω.

Στην οικογένειά μου

ΠΕΡΙΛΗΨΗ

Η παρούσα εργασία επικεντρώνεται στην ανάλυση των επαναληπτικών μεθόδων που χρησιμοποιούνται για την επίλυση αραιών γραμμικών συστημάτων, εστιάζοντας στα πλεονεκτήματα και στην αποδοτικότητα της κάθε μεθόδου. Αρχικά, στο εισαγωγικό μέρος, γίνεται μια αναφορά στις δυο μεγάλες κατηγορίες που χωρίζονται οι αριθμητικές μέθοδοι, στις στάσιμες και τις επαναληπτικές. Στη συνέχεια αναφέρονται κάποιες βασικές έννοιες της Γραμμικής Άλγεβρας, όπως ορισμοί νορμών πινάκων και διανυσμάτων, εσωτερικών γινομένων , πινάκων ειδικής μορφής, καθώς και η έννοια της ορθογωνιότητας, οι οποίες χρησιμοποιούνται και στα επόμενα κεφάλαια. Κατόπιν, στο κύριο μέρος της εργασίας, παρουσιάζονται αναλυτικά οι επαναληπτικές μέθοδοι επίλυσης αραιών γραμμικών συστημάτων, με τη χρονολογική σειρά την οποία εισήχθησαν. Ξεκινάμε από την Μέθοδο Απότομης Καθόδου (Steepest Descent Method) και έπειτα, αφού γίνει μια εισαγωγή στους υποχώρους Krylov, εισάγουμε τη Μέθοδο Συζυγών Κλίσεων (Conjugate Gradient Method), η οποία είναι πιο εξελιγμένη ως προς την πολυπλοκότητα, τις απαιτήσεις μνήμης και την ευστάθεια. Προχωρώντας παρακάτω, δίνονται οι τροποποιήσεις της Μεθόδου Συζυγών Κλίσεων, εισάγοντας της τεχνική της προρρύθμισης, χάρη στην οποία το γραμμικό σύστημα συγκλίνει ταχύτερα. Φυσικά, σε κάθε μέθοδο παρουσιάζεται η βασική μαθηματική θεωρία, οι συνθήκες σύγκλισης των νορμών των διανυσμάτων υπολοίπου και σφάλματος, καθώς και τα πλεονεκτήματα και τα μειονεκτήματα της κάθε μεθόδου αντίστοιχα. Έπειτα, δίνονται μερικά παραδείγματα για κάθε μέθοδο που αναλύσαμε, προκειμένου να γίνουν πλήρως κατανοητές από τον αναγνώστη. Τέλος, αφού κάνουμε μια εισαγωγή στο μαθηματικό λογισμικό MATLAB και αφού παρουσιάσουμε ορισμένους κανόνες και ιδιότητές του, υλοποιούμε σε αυτό τη Μέθοδο της Απότομης Καθόδου και τη Μέθοδο Συζυγών Κλίσεων με ένα παράδειγμα για την κάθε μία, αντίστοιχα.

ΛΕΞΕΙΣ-ΚΛΕΙΔΙΑ

Επαναληπτικές Μέθοδοι, Αραιά Γραμμικά Συστήματα, ταχύτητα σύγκλισης, Μέθοδος Συζυγών Κλίσεων, Υπόχωροι Krylov, Προρρυθμιστές

ABSTRACT

This current thesis focuses on the analysis of iterative methods, which are used in order to solve sparse linear systems, focusing on the advantages and the efficiency of each method. Firstly, on the introductory part, there is a reference on the two major categories that numerical methods are divided, the directs, and the repetitive ones. Continually, some basic means of Linear Algebra are mentioned, like definitions of norms of vectors and matrices, of inner products, matrices of special form, as well as the concept of orthogonality, which are used to the next sections. After that, on the main part of this assignment, the iterative methods of solving sparse linear systems are presented, with chronological order in which they were introduced. We begin with the Steepest Descent Method, and then, after an introduction on Krylov subspaces is made, we introduce the Conjugate Gradient Method, which is more developed in terms of complexity, memory requirement and stability. Moving forward the modifications of Conjugate Gradient Method are introduced, by introducing the preconditioning technique, thanks to which the linear system converge faster. Of course, in every method, we present the basic mathematical theory, the conditions of converging the norms, the residuum vector and error, as well as the advantages and disadvantages of every method are presented. After that, some examples of every method we analyzed are given, in order to be more understandable from the reader. Finally, after making an introduction to mathematical software MATLAB and introducing some of its rules and properties, we implement the Steepest Descent Method and the Conjugate Gradient Method in this with an example for each one, respectively.

KEY-WORDS

Iterative Methods, Sparse Linear Systems, Convergence Speed, Conjugate Gradient Method, Krylov Subspaces, Preconditions

ΠΕΡΙΕΧΟΜΕΝΑ

1 Εισαγωγή.....17-21

1.1 Γενικά περί γραμμικών συστημάτων και τρόπων επίλυσής τους.....17-18

1.2 Βασική θεωρία Γραμμικής Άλγεβρας.....18-21

2 Μέθοδος Απότομης Καθόδου (Steepest Descent Method).....22-32

3 Μέθοδος Συζυγών Κλίσεων.....33-43

3.1 Μέθοδος Gram-Schmidt και A-ορθογωνιότητα.....33

3.2 Μέθοδος Συζυγών Κλίσεων (Conjugate Gradient Method)……....33-43

4 Προρρυθμισμένη Μέθοδος Συζυγών Κλίσεων.....44-46

5 Παραδείγματα Επαναληπτικών Μεθόδων.....47-49

6 Υλοποίηση Αριθμητικών Μεθόδων στο Matlab……..50-52

7 Βιβλιογραφία.....53

Παράρτημα.....54-61

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

1.1 Γενικά περί γραμμικών συστημάτων και τρόπων επίλυσής τους

Ένα ιδιαίτερα ενδιαφέρον πρόβλημα στην περιοχή των επιστημονικών υπολογισμών είναι η επίλυση γραμμικών αλγεβρικών συστημάτων που περιγράφονται από την εξίσωση $Ax=b, a\in\mathbb{C}^{n,n}, det(A)\neq 0, b\in\mathbb{C}^{n\setminus\{0\}}$.Το x είναι το διάνυσμα αγνώστων ως προς το οποίο καλούμαι να λύσω το γραμμικό σύστημα.

Οι αριθμητικές μέθοδοι επίλυσης γραμμικών συστημάτων χωρίζονται σε δύο μεγάλες κατηγορίες:

- Άμεσες μέθοδοι (direct methods)
- Έμμεσες μέθοδοι ή αλλιώς επαναληπτικές μέθοδοι (iterative methods)

Στην πρώτη κατηγορία ανήκουν η απαλοιφή Gauss και ο αλγόριθμος Cholesky. Βασική ιδέα για την επίλυση γραμμικών συστημάτων με την χρήση άμεσων μεθόδων είναι η μετατροπή του συστήματος $Ax=b$ σε ένα ισοδύναμο τριγωνικό, το οποίο επιλύεται ευκολότερα εκτελώντας έναν προκαθορισμένο αριθμό βημάτων ο οποίος εξαρτάται από το μέγεθος n του γραμμικού συστήματος. Επίσης, όσον αφορά την τάξη μεγέθους, οι άμεσες μέθοδοι επίλυσης, όπως η απαλοιφή Gauss, κοστίζει $O(n^3)$ σε πράξεις και $O(n^2)$ σε θέσεις μνήμης, διότι ο πίνακας είναι πυκνός και απαιτεί n^2 θέσεις μνήμης για την αποθήκευσή του.

Περνάμε τώρα στην δεύτερη κατηγορία μεθόδων επίλυσης, αυτή των επαναληπτικών μεθόδων. Οι επαναληπτικές μέθοδοι με τη σειρά τους χωρίζονται σε δύο μικρότερες κατηγορίες:

1. Στάσιμες μέθοδοι, οι οποίες είναι της μορφής $x^{n+1}=F(x^n)$
2. Μη στάσιμες μέθοδοι, οι οποίες είναι της μορφής $x^{n+1}=F^n(x^n)$

Στις στάσιμες μεθόδους εντάσσονται η μέθοδος Jaccobi, η Gauss-Seidel, η Μέθοδος Διαδοχικής Υπερχαλάρωσης (SOR-**S**uccessive **O**ver **R**elaxation) κτλ. Οι μέθοδοι αυτές ειναι απλούστερες και παλαιότερες, αλλά δυστυχώς λιγότερο αποτελεσματικές, διότι σε κάθε επανάληψη εκτελούν τις ίδιες λειτουργίες στα διανύσματα επανάληψης, αλλά δεν συγκλίνουν πάντοτε στη λύση.

Οι επαναληπτικές υπολογιστικές μέθοδοι αποτελούν μια πολύ καλή εναλλακτική έναντι των άμεσων μεθόδων, καθώς έχουν μειωμένες απαιτήσεις σε υπολογιστική ισχύ και μνήμη, παρόλο που είναι λίγο πιο πολύπλοκες στην εφαρμογή τους. Οι βασικές πράξεις που εκτελούνται εντός μιας επαναληπτικής μεθόδου είναι πράξεις πολλαπλασιασμού πίνακα με διάνυσμα και εσωτερικά γινόμενα διανυσμάτων. Οι βασικότερες από αυτές είναι η Μέθοδος Συζυγών Κλίσεων (CG-**C**onjugate **G**radient **M**ethod), η Μέθοδος της Απότομης Καθόδου (**S**teepest **D**escent), η Μέθοδος Συζυγών Διευθύνσεων (**C**onjugate **D**irections), η Προρρυθμισμένη Μέθοδος Συζυγών Κλίσεων κτλ.

Γενικά, βασική ιδέα κάθε επαναληπτικής μεθόδου είναι να ξεκινήσουμε από μια αρχική προσέγγιση x^0 (που μπορεί να είναι και αυθαίρετη αν δεν γνωρίζουμε κάτι για την λύση) και

να κατασκευάσουμε διαδοχικές προσεγγίσεις $x^1, x^2, \ldots, x^n$ που να συγκλίνουν στη λύση του συστήματος των γραμμικών εξισώσεων, δηλαδή στο x. Η σύγκλιση αυτών των μεθόδων εξαρτάται από τον πίνακα επανάληψης, δηλαδή από το φάσμα του ή αλλιώς το σύνολο των ιδιοτιμών του. Οι επαναληπτικές μέθοδοι είναι οικονομικότερες από πλευρά μνήμης, γιατί εκμεταλλεύονται την αραιή μορφή του πίνακα A. Με αυτές τις μεθόδους δεν χρειάζεται να φτάσει κανείς μέχρι την πραγματική λύση, αλλά μπορεί να σταματήσει μόλις επιτευχθεί κάποια επιθυμητή ακρίβεια που σχετίζεται με την ακρίβεια των δεδομένων του προβλήματος.

Ας εξηγήσουμε τώρα τι εννοούμε με την έννοια αραιοί πίνακες. Ως αραιός ορίζεται ένας πίνακας με αρκετά μηδενικά στοιχεία τάξης 0(n) έτσι ώστε να αξίζει να τα εκμεταλλευτούμε στους υπολογισμούς μας. Η πρόκληση που εισάγει ένας αραιός πίνακας είναι ότι:

- Ταχύτερος υπολογισμός του γινομένου του με ένα διάνυσμα, αφού εκτελούνται πράξεις μόνο με τα μη-μηδενικά στοιχεία του πίνακα με αποτέλεσμα την βελτίωση του χρόνου εκτέλεσης
- Εξοικονόμηση μνήμης, αφού τα μηδενικά στοιχεία δεν αποθηκεύονται

Όμως, επειδή η αποθήκευση αραιών πινάκων απαιτεί εξειδικευμένες δομές δεδομένων, θα πρέπει τα οφέλη της ταχύτητας να υπερκαλύπτουν το κόστος διαχείρισης (λόγω μη εκτέλεσης πράξεων με τα μηδενικά στοιχεία του πίνακα).

Η εργασία αυτή θα επικεντρωθεί στη μελέτη της Μεθόδου Συζυγών Κλίσεων (Conjugate Gradient Method) για την επίλυση γραμμικών συστημάτων της μορφής $Ax=b$, όπου A είναι ένας τετραγωνικός, συμμετρικός και θετικά ορισμένος πίνακας και $x, b \in \mathbb{R}^n$.

1.2 Βασική θεωρία Γραμμικής Άλγεβρας

Θα χρειαστεί σε αυτό το σημείο, για καλύτερη κατανόηση, να διατυπώσουμε απαραίτητους ορισμούς και βασικά θεωρήματα της γραμμικής άλγεβρας. Θα δοθούν ορισμοί εσωτερικού γινομένου, νόρμας διανύσματος και πίνακα, φασματικής ακτίνας και δείκτη κατάστασης πινάκων, καθώς και κάποιες ιδιότητες των πινάκων που θα χρησιμοποιηθούν στη συνέχεια.

Ορισμός 1.2.1: Το Ευκλείδειο εσωτερικό γινόμενο δύο διανυσμάτων $x, y \in \mathbb{C}^n$ με τη σειρά αυτή ορίζεται ως εξής $(x, y)_2 := \sum_{i=1}^{n} x_i \bar{y}_i$ και συμβολίζεται με $(x,y)_2$ ή απλά $\langle x, y \rangle$.

Το Ευκλείδειο εσωτερικό γινόμενο ορίζει την Ευκλείδεια νόρμα $\|\cdot\|_2$ ως εξής $\|x\|_2 = \sqrt{\langle x, x \rangle}$.

Οι νόρμες που χρησιμοποιούνται στο $\mathbb{C}^n$ περισσότερο είναι οι εξής:

i. $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ (ℓ_1 -νόρμα ή φυσική νόρμα)

ii. $\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}$ (ℓ_2 -νόρμα ή ευκλείδεια νόρμα)

iii. $\|x\|_\infty = \max_{i=1,\dots,n} |x_i|$ (ℓ_∞ -νόρμα)

Παρατήρηση 1.2.1: Μια διανυσματική νόρμα ορίζει προφανώς μια απόσταση $d(x,y)=\|x-y\|$ στον $\mathbb{C}^n$, η οποία έχει τις εξής ιδιότητες:

- $d(x,x)=0, \forall x \in \mathbb{C}^n$ και $d(x,y)=0$ αν και μόνο αν $x=y$
- $d(x,y)=d(y,x), \forall x,y \in \mathbb{C}^n$
- $d(x,y) \le d(x,z)+d(z,y), \forall x,y,z \in \mathbb{C}^n$.

Ορισμός 1.2.2: Δύο νόρμες με $\|\cdot\|_\alpha$ και $\|\cdot\|_\beta$ στο $\mathbb{C}^n$ λέμε ότι είναι ισοδύναμες (συγκρίσιμες), όταν υπάρχουν σταθερές $c_1, c_2 > 0$ τ.ω.:

$$c_1\|x\|_\alpha \le \|x\|_\beta \le c_2\|x\|_\alpha, \forall x \in \mathbb{C}^n \quad .$$

Ορισμός 1.2.3: Η απεικόνιση $\|A\| := \max_{x \in \mathbb{C}^{n\setminus\{0\}}} \frac{\|Ax\|}{\|x\|}, \forall A \in \mathbb{C}^{n,n}$ ορίζει μια φυσική νόρμα πίνακα, όπου $\|\cdot\|$ είναι μια νόρμα στο $\mathbb{C}^n$.

Όπως στις νόρμες διανυσμάτων, έτσι στις νόρμες πινάκων, οι συνηθέστερα χρησιμοποιούμενες είναι οι παρακάτω:

i. $\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |\alpha_{ij}|$

ii. $\|A\|_2 = (\rho(A^H A))^{\frac{1}{2}}$, φασματική ακτίνα του πίνακα $A^H A$

iii. $\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |\alpha_{ij}|$

οι οποίες παράγονται, αντίστοιχα, από τις $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ νόρμες διανύσμάτων.

<u>Σημείωση 1:</u> Ως φασματική ακτίνα ενός πίνακα Α ορίζεται η ποσότητα $\max_{1 \le i \le n} |\lambda_i|$, όπου $(\lambda_i)^n{}_{i=1}$ οι ιδιοτιμές του πίνακα Α.

Πρόταση 1.2.1: $\rho(A) \le \|A\|$ όπου $\|\cdot\|$ μια φυσική νόρμα πίνακα.

Απόδειξη: Έστω $\lambda \in \mathbb{C}$ ιδιοτιμή του πίνακα Α και $x \in \mathbb{C}^{n\setminus\{0\}}$ αντίστοιχο ιδιοδιάνυσμα. Τότε ισχύει $Ax = \lambda x$. Οπότε, παίρνοντας νόρμες και στα δύο μέλη έχουμε $\|Ax\| = \|\lambda x\|$. Ξέρω όμως ότι ισχύει $\|Ax\| = \|\lambda x\| = |\lambda|\|x\|$. Από τον ορισμό της φυσικής νόρμας έχω

$$\frac{\|Ax\|}{\|x\|} \le \|A\| \Rightarrow \frac{|\lambda|\|x\|}{\|x\|} \le \|A\| \Rightarrow |\lambda| \le \|A\| \quad .$$

Άρα $\rho(A) \leq \|A\|$. □

Ορισμός 1.2.4: Δείκτης κατάστασης ενός αντιστρέψιμου πίνακα $A \in \mathbb{C}^{n,n}$ ως προς μια φυσική νόρμα $\|\cdot\|$ καλείται ο αριθμός $\kappa(A) = \|A\| \|A^{-1}\|$.

<u>Σχόλιο:</u> Για τον δείκτη κατάστασης ενός αντιστρέψιμου πίνακα $A \in \mathbb{C}^{n,n}$ ισχύει προφανώς λόγω του ορισμού ότι $\kappa(A) \geq 1$, αφού $1 = \|I\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\| = \kappa(A)$.

Πρόταση 1.2.2: Έστω $A \in \mathbb{C}^{n,n}$ ένας ερμιτιανός και αντιστρέψιμος πίνακας. Ο δείκτης κατάστασης $\kappa_2(A)$ ως προς την Ευκλείδεια νόρμα $\|\cdot\|_2$ δίνεται από τον τύπο:

$$\kappa_2(A) = \frac{\max_i |\lambda_i|}{\min_i |\lambda_i|} \quad ,$$

όπου λ_i είναι οι ιδιοτιμές του πίνακα A.

Απόδειξη:

$$\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = [p(A^H A)]^{\frac{1}{2}} [p(((A^{-1}))^H A^{-1})]^{\frac{1}{2}} = [p(A^2)]^{\frac{1}{2}} [p((A^{-1})^2)]^{\frac{1}{2}} =$$
$$\sqrt{\max_i |\lambda_i|^2} \sqrt{\max_i \frac{1}{|\lambda_i|^2}} = \max_i |\lambda_i| \max_i |\frac{1}{\lambda_i}| = \max_i |\lambda_i| \frac{1}{\min_i |\lambda_i|} = \frac{\max_i |\lambda_i|}{\min_i |\lambda_i|} .$$

□

<u>Σημείωση 2:</u> Έστω $B \in \mathbb{C}^{n,n}$ ο οποίος είναι αυτοσυζυγής (Ερμιτιανός). Τότε υπάρχουν n ιδιοδιανύσματα $(v_i)_{i=1}^n$ τα οποία είναι ορθοκανονικά και επομένως αποτελούν βάση του $\mathbb{C}^n$. Έστω λ_i η ιδιοτιμή που αντιστοιχεί στο v_i για $i = 1, \ldots, n$.

Στην προηγούμενη πρόταση χρησιμοποιήσαμε το ακόλουθο αποτέλεσμα

Πρόταση 1.2.3: Όταν $B = B^H$, οι ιδιοτιμές του B^2 είναι $(\lambda_i^2)_{i=1}^n$, όπου $(\lambda_i)_{i=1}^n$ οι ιδιοτιμές του B.

Απόδειξη:

"$\Rightarrow$" Αν λ είναι ιδιοτιμή και v ιδιοδιάνυσμα που αντιστοιχεί σε αυτή, τότε $B^2 v = B(Bv) = B(\lambda v) = \lambda B v = \lambda^2 v$. Άρα, λ^2 είναι ιδιοτιμή του B.

"$\Leftarrow$" Έστω μ μια ιδιοτιμή του B^2 και $w \in \mathbb{C}^{n \setminus \{0\}}$ ένα αντίστοιχο ιδιοδιάνυσμα.

Τότε $B^2 w = \mu w$. Επειδή υπάρχει βάση $(v_i)_{i=1}^n$ του $\mathbb{C}^n$, έπεται ότι $w = \sum_{i=1}^n w_i v_i$ είναι ιδιοδιανύσματα του B.

Έτσι, $B^2 w = \sum_{i=1}^n w_i B^2 v_i = \sum_{i=1}^n w_i \lambda_i^2 v_i$. Έτσι, $B^2 w = \mu w \Rightarrow \sum_{i=1}^n w_i \lambda_i^2 v_i = \sum_{i=1}^n w_i \mu v_i \Rightarrow$

$\Rightarrow \sum_{i-1}^{n} (\mu - \lambda_i^2) v_i = 0 \Rightarrow w_i (\mu - \lambda_i^2) = 0$, για $i = 1, \dots, n$. Επειδή $w \neq 0$ για κάποιο $i_0 \in \{1, \dots, n\}$, $w_{i_0} \neq 0$. Άρα, $\mu = \lambda_{i_0}^2$. □

ΚΕΦΑΛΑΙΟ 2

Μέθοδος Απότομης Καθόδου (Steepest Descent Method)

Η Μέθοδος των Συζυγών κλίσεων αποτελεί βελτίωση της Μεθόδου Απότομης Καθόδου και γι' αυτό τον λόγο θα παρουσιάσουμε αρχικά την Μέθοδο της Απότομης Καθόδου και έπειτα την Μέθοδο Συζυγών Κλίσεων. Με άλλα λόγια ξεκινάμε από ένα τυχαίο σημείο x^0 , όπως γίνεται άλλωστε και σε κάθε επαναληπτική διαδικασία, και εκτελώντας μια σειρά βημάτων $x^1, x^2, \ldots, x^n$ προσεγγίζουμε τελικά τη λύση του συστήματος $Ax=b$.

Θεωρούμε το $n \times n$ πραγματικό σύστημα $Ax=b$, με A συμμετρικό και θετικά ορισμένο πίνακα και $b \in \mathbb{R}^n$.

Ορίζουμε το συναρτησιακό $f:\mathbb{R}^n \to \mathbb{R}$ με

$$f(x)=\frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle = \frac{1}{2}x^T A x - b^T x \quad , \quad \forall x \in \mathbb{R}^n \quad \textbf{(2.1)}$$

Η συνάρτηση f(x) έχει ελάχιστο στο x_0 αν $f(x) \geq f(x_0), \forall x \in \mathbb{R}$.

Ορισμός 2.1: Έστω συνάρτηση $f:\mathbb{R}^n \to \mathbb{R}$. Η κλίση της συνάρτησης f στο x ορίζεται ως το διάνυσμα των μερικών παραγώγων της f στο x (εφόσον υπάρχουν) και δίνεται από την παρακάτω σχέση:

$$\nabla f(x)=\left(\frac{\partial}{\partial x_1}f(x), \ldots, \frac{\partial}{\partial x_n}f(x)\right)^T \quad .$$

Πρόταση 2.1: Η κλίση της συνάρτησης f(x) που δίνεται στη **(2.1)** είναι

$$\nabla f(x)=Ax-b \quad . \textbf{(2.2)}$$

Επιπλέον:
αν ο A είναι συμμετρικός και θετικά ορισμένος, τότε η f παρουσιάζει μοναδικό ελάχιστο στο x_0 που είναι η λύση του γραμμικού συστήματος $Ax=b$

Απόδειξη: Υπολογίζουμε την μερική παράγωγο της f(x):

$$\frac{\partial}{\partial x_k}f(x)=\alpha_{kk}x_k+\frac{1}{2}\sum_{i\neq k}\alpha_{ik}x_i+\frac{1}{2}\sum_{i\neq k}\alpha_{ki}x_i-b_k=\sum_i \alpha_{ki}x_i-b_k=(Ax-b)_k$$

και ως εκ τούτου εύκολα συμπεραίνουμε ότι $\nabla f(x)=Ax-b$. Επειδή ο A είναι ένας πραγματικός συμμετρικός πίνακας, υπάρχει ορθοκανονική βάση $(\hat{e}_i)^n{}_{i=1}$ ιδιοδιανυσμάτων του A με $(\lambda_i)^n{}_{i=1}$ οι αντίστοιχες ιδιοτιμές. Θέτοντας $x=\sum_{i=1}^{n}\hat{x}_i\hat{e}_i$ και $b=\sum_{i=1}^{n}\hat{b}_i\hat{e}_i$ έχουμε:

$$f(x)=\frac{1}{2}\sum_{i=1}^{n}\lambda_i \hat{x}_i^2-\sum_{i=1}^{n}\hat{b}_i\hat{x}_i=\frac{1}{2}\sum_{i=1}^{n}(\lambda_i(\hat{x}_i-\frac{\hat{b}_i}{\lambda_i})^2-\frac{\hat{b}_i^2}{\lambda_i}) \quad .$$

Επειδή ο A είναι θετικά ορισμένος, τότε $\lambda_i>0, \forall 1\leq i\leq n$. Σε αυτή την περίπτωση ελαχιστοποιούμε την f(x), ελαχιστοποιώντας κάθε τετραγωνικό της όρο. Υπάρχει ένα μοναδικό ελάχιστο x για το οποίο ισχύει $\hat{x}_i=\frac{\hat{b}_i}{\lambda_i}$,και αυτό το διάνυσμα είναι η μοναδική λύση του συστήματος Ax=b. □

Πόρισμα 2.1: Θεωρώ τον A να είναι ένας πραγματικός θετικά ορισμένος συμμετρικός πίνακας. Έστω f(x) να είναι η συνάρτηση που ορίζεται από την σχέση **(2.1)** και F να είναι ένας διανυσματικός υπόχωρος του $\mathbb{R}^n$. Υπάρχει ένα μοναδικό διάνυσμα

$$x_0\in F\, \tau.\omega.\, f(x_0)\leq f(x), \forall x\in F \quad .$$

Επιπλέον, το x_0 είναι το μοναδικό διάνυσμα του F τ.ω. $\langle Ax_0-b, y\rangle=0, \forall y\in F$.

Απόδειξη: Συμβολίζω με P την ορθογώνια προβολή του $\mathbb{R}^n$ στο διανυσματικό υπόχωρο F. Επίσης συμβολίζουμε με P τον πίνακα της κανονικής βάσης αυτής της ορθογώνιας προβολής. Από τον ορισμό της ορθογώνιας προβολής έχουμε ότι: $P^T=P, P^2=P$ και η απεικόνιση $P:\mathbb{R}^n\rightarrow F$ είναι "επί". Συνεπώς $\min_{x\in F} f(x)=\min_{y\in\mathbb{R}^n} f(Py)$.

Υπολογίζουμε $f(Py)=\frac{1}{2}\langle P^T APy, y\rangle-\langle P^T b, y\rangle$. Η πρόταση (2.1) μπορεί να εφαρμοστεί σε αυτή την συνάρτηση. Αφού ο A είναι θετικά ορισμένος, εύκολα βλέπουμε ότι ο $P^T A P$ είναι μη-αρνητικός. Αν $P^T b$ δεν ανήκει στην τάξη του $(P^T A P)$, τότε μπορούμε να συμπεράνουμε ότι το infimum της f(Py) είναι $-\infty$, το οποίο είναι αδύνατο, άρα $\inf_{y\in\mathbb{R}^n} f(Py)=\inf_{x\in F} f(x)\geq \min_{x\in\mathbb{R}^n} f(x)>-\infty$ επειδή ο A είναι θετικά ορισμένος. Έτσι, $P^T b\in Im(P^T AP)$ και το ελάχιστο της f(Py) λαμβάνεται από όλες τις λύσεις της εξίσωσης $P^TAP=P^Tb$. Τώρα θα δείξουμε ότι η εξίσωση αυτή έχει μοναδική λύση στον υπόχωρο F. Έστω x_1 και x_2 να είναι δύο λύσεις της εξίσωσης $P^T A P x=P^T b$. Έχουμε: $P^T A P(x_1-x_2)=0\Rightarrow\langle A(Px_1-Px_2),(Px_1-Px_2)\rangle=0\Rightarrow Px_1=Px_2$. Αφού το x_1 και το x_2 ανήκουν στον F, έχουμε $x_1=Px_1=Px_2=x_2$. Πολλαπλασιάζοντας την εξίσωση $P^T A P x_0=P^T b\, με\, y\in F$ και χρησιμοποιώντας τη σχέση $Px_0=x_0$ και $Py=y$, επειδή $x_0, y\in F$, προκύπτει η σχέση $\langle Ax_0-b, y\rangle=0\, \forall y\in F$. □

Θεώρημα 2.1: Θεωρώ τον A να είναι ένας συμμετρικός θετικά ορισμένος πίνακας και την f να είναι η συνάρτηση στον $\mathbb{R}^n$ από την σχέση **(2.1)**.

Θεωρώ το $x\in\mathbb{R}^n$ να είναι τ.ω. $\nabla f(x)\neq 0$. Τότε $\forall\alpha\in(0,\frac{2}{\rho(A)})$, έχουμε $f(x-\alpha\nabla f(x))<f(x)$.

Απόδειξη:
Θεωρώ το x να είναι ένα σημείο τ.ω. $\nabla f(x)\neq 0$ και θέτω $\delta=-\alpha(Ax-b)$. Αφού ο A είναι συμμετρικός, υπολογίζουμε:

$$f(x+\delta)=\frac{1}{2}\langle A(x+\delta),(x+\delta)\rangle-\langle b,x+\delta\rangle=f(x)+\frac{1}{2}\langle A\delta,\delta\rangle+\langle Ax-b,\delta\rangle \quad .$$

Τώρα έχουμε: $\langle A\delta,\delta\rangle\leq\|A\|_2\|\delta\|_2^2$, και για έναν πραγματικό συμμετρικό πίνακα ισχύει: $\|A\|=\rho(A)$. Άρα: $f(x+\delta)\leq f(x)+\alpha(\alpha\frac{\rho(A)}{2}-1)\|Ax-b\|^2$.

Συνεπώς, $f(x+\delta)<f(x)$ αν $0<\alpha<\frac{2}{\rho(A)}$. □

Μια πρώτη μορφή της Μεθόδου Απότομης Καθόδου είναι η εξής:

$$\left\{\begin{array}{c} x^0\in\mathbb{R} \\ x^{k+1}=x^k+\alpha(b-Ax^k),\forall k\geq 1 \end{array}\right\},$$

με $A=M-N\Rightarrow x=M^{-1}N+M^{-1}b$, όπου $M=\frac{1}{\alpha}I_n$, $N=(\frac{1}{\alpha}I_n-A)$ και α να είναι μια μη-μηδενική παράμετρος.

Οπότε λύνω το εξής σύστημα:

$$(M-N)x=b\Rightarrow Mx=Nx+b\Rightarrow x=M^{-1}Nx+M^{-1}b\Rightarrow x=\alpha Nx+\alpha b\Rightarrow x=(I_n-\alpha Ax)+\alpha b\Rightarrow x=I_n-\alpha(Ax-b).$$

Για τη συγκεκριμένη επαναληπτική μέθοδο επίλυσης του γραμμικού συστήματος $Ax=b$, $A\in\mathbb{C}^{n,n}$ ισχύουν κάποιες σχέσεις οι οποίες δίνονται στη συνέχεια.

Θεώρημα 2.2: Η επαναληπτική μέθοδος που ορίζεται παραπάνω συγκλίνει αν και μόνο αν η φασματική ακτίνα του $M^{-1}N$ ικανοποιεί τη σχέση $\rho(M^{-1}N)<1$.

Θεώρημα 2.3: Έστω $A\in\mathbb{R}$ με ιδιοτιμές $\lambda_{min}=\lambda_1\leq\lambda_2\leq...\leq\lambda_n=\lambda_{max}$. Τότε:

i. Αν $\lambda_1\leq 0\leq\lambda_n$, τότε η Μέθοδος της Απότομης Καθόδου δε συγκλίνει για καμία τιμή του α.
ii. Αν $0<\lambda_1\leq\lambda_2\leq...\leq\lambda_n$, τότε η παραπάνω μέθοδος συγκλίνει αν και μόνο αν $0<\alpha<\frac{2}{\lambda_{max}}$. Σε αυτή την περίπτωση, η βέλτιστη παράμετρος α, η οποία ελαχιστοποιεί τη φασματική ακτίνα $\rho(M^{-1}N)$, είναι $\alpha_{opt}=\frac{2}{\lambda_1+\lambda_n}$, με την φασματική ακτίνα να παίρνει την τιμή $\min_{\alpha}\rho(M^{-1}N)=\min_{a}|1-\alpha A|=\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}=\frac{\kappa(A)-1}{\kappa(A)+1}$.

Απόδειξη: Σύμφωνα με το Θεώρημα 2.2 ξέρουμε ότι η Μέθοδος της Απότομης Καθόδου συγκλίνει αν και μόνο αν $\rho(M^{-1}N)<1$. Εδώ έχουμε $M^{-1}N=(I_n-aA)$. Άρα:

$$\rho(M^{-1}N)<1\Leftrightarrow\rho(I_n-aA)<1\Leftrightarrow|1-\alpha\lambda_i|<1\Leftrightarrow-1<1-\alpha\lambda_i<1,\forall i\Leftrightarrow 2>\alpha\lambda_i>0,\forall 1\leq i\leq n \quad .$$

i. Όταν $\lambda_1 \leq 0 \leq \lambda_n$, τότε χρειαζόμαστε $2>\alpha\lambda_1>0$ και $2>\alpha\lambda_n>0$.
Όταν λ_1 ή λ_n είναι 0 δεν ισχύει. Ας υποθέσουμε ότι $\lambda_1, \lambda_n \neq 0$, δηλαδή $\lambda_1<0<\lambda_n$. Τότε $\alpha\lambda_1>0$, δηλαδή $\alpha<0$, και $\alpha\lambda_n>0$, δηλαδή $\alpha>0$. Άτοπο, δηλαδή δεν μπορούν να ισχύουν οι σχέσεις $2>\lambda_i>0, \forall i \in \{1,\dots,n\}$.

ii. Αν, από την άλλη πλευρά, έχουμε $0<\lambda_1 \leq \dots \leq \lambda_n$, τότε συμπεραίνουμε ότι:

$$2>\alpha\lambda_i>0, \forall 1 \leq i \leq n$$
$$\Leftrightarrow 0<\alpha<\frac{2}{\lambda_i}, i=1,\dots,n$$
$$\Leftrightarrow 0<\alpha<\frac{2}{\lambda_{max}}=\frac{2}{\lambda_n}$$

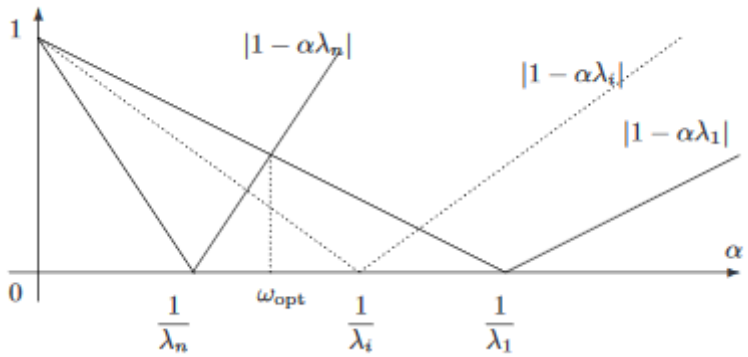
Για να υπολογίσουμε την βέλτιστη παράμετρο α_{opt} , η οποία ελαχιστοποιεί την φασματική ακτίνα, παρατηρούμε ότι η συνάρτηση $\lambda \to |1-\alpha\lambda|$ είναι φθίνουσα στο ($-\infty, \frac{1}{\alpha}$] και αύξουσα στο [$\frac{1}{a}, +\infty$).
Έτσι,

$$\rho(M^{-1}N)=max\{|1-\alpha\lambda_1|, |1-\alpha\lambda_n|\} \quad .$$

Άρα σημειώνει το ελάχιστό της στο σημείο α_{opt}. Αυτό επιτυγχάνεται όταν:

$$1-\alpha_{opt}\lambda_1=\alpha_{opt}\lambda_n-1 \Rightarrow 2=\alpha_{opt}\lambda_1+\alpha_{opt}\lambda_n \Rightarrow 2=\alpha_{opt}(\lambda_1+\lambda_n) \Rightarrow \alpha_{opt}=\frac{2}{\lambda_1+\lambda_n} \quad .$$

Επομένως: $\rho(M^{-1}N)=\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}=\frac{\kappa(A)-1}{\kappa(A)+1}$, όπου $\kappa(A)=\kappa=\frac{\lambda_n}{\lambda_1}$ ο δείκτης κατάστασης του συστήματος.



Εικόνα 1: Ιδιοδιανύσματα του πίνακα $B_\alpha=I-\alpha A$ □

Θεώρημα 2.4: Έστω ένας $n \times n$ συμμετρικός και θετικά ορισμένος πίνακας Α. Για τη νόρμα του διανύσματος υπολοίπου ισχύει

$$\|r^k\| \le \left(\frac{\kappa-1}{\kappa+1}\right)^k \|r^0\| \quad , \quad \kappa = \frac{\lambda_n}{\lambda_1} = \frac{\lambda_{max}}{\lambda_{min}}$$

Απόδειξη: Από τον ορισμό του διανύσματος-υπόλοιπο έχω:

$$r^{k+1} = b - Ax^{k+1} = b - A(x^k + \alpha_{opt} r^k) = b - Ax^k - A\alpha_{opt} r^k = r^k -_K r^k = (I - \alpha_{opt} A) r^k$$

Περνώντας νόρμες και στα δύο μέλη παίρνουμε:

$$\|r^{k+1}\| = \|(I - \alpha_{opt} A) r^k\| \le \|(I - \alpha_{opt} A)\| \|r^k\|$$

Χρησιμοποιώντας την ισότητα $Ax = \lambda x$ η παραπάνω σχέση γίνεται:

$$\|r^{k+1}\| \le \rho(I - \alpha_{opt} A) \|r^k\|$$

Στη συγκεκριμένη περίπτωση ψάχνουμε την ελάχιστη ποσότητα του $|1 - \lambda \alpha_{opt}|$.
Αυτό συμβαίνει μόνο όταν ισχύει

$$|1 - \alpha_{opt} \lambda_1| = |1 - \alpha_{opt} \lambda_n| \Leftrightarrow 1 - \alpha_{opt} \lambda_1 = \alpha_{opt} \lambda_n - 1 \Leftrightarrow \alpha_{opt} = \frac{2}{\lambda_1 + \lambda_n}$$

Επομένως

$$\|r^{k+1}\| \le \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \|r^k\|$$

Στο κλάσμα του δεύτερου μέλους διαιρώ αριθμητή και παρονομαστή με τη μικρότερη ιδιοτιμή, δηλαδή με λ_1 και καταλήγω στη σχέση

$$\|r^{k+1}\| \le \frac{\frac{\lambda_n}{\lambda_1} - 1}{\frac{\lambda_n}{\lambda_1} + 1} \|r^k\| = \left(\frac{\kappa-1}{\kappa+1}\right)^k \|r^0\| \quad .$$

□

Ας ορίσουμε τη μέθοδο Απότομης Καθόδου. Έστω, λοιπόν, x^k η προσεγγιστική λύση του παραπάνω συστήματος στο k-οστό βήμα της μεθόδου. Τότε κάθε επόμενη επανάληψη λαμβάνεται από τον ακόλουθο τύπο:

$$\left\{ \begin{array}{c} x^0 \in \mathbb{R}^n \\ x^{k+1} = x^k + \alpha_k r^k, \forall\, k \ge 1 \end{array} \right\}.$$

Με r^k συμβολίζουμε το διάνυσμα-υπόλοιπο (residual-vector) το οποίο ισούται με $r^k = b - Ax^k$ και είναι ο δείκτης του πόσο απέχουμε από τη σωστή τιμή του διανύσματος b.

Ως α_k επιλέγεται ο αριθμός που ελαχιστοποιεί τη συνάρτηση $g(\alpha) = f(x^k - \alpha \nabla f(x^k))$,

για $\alpha \in \mathbb{R}$

Οπότε έχω: $$g'(\alpha) = -\nabla f(x^k - \alpha \nabla f(x^k)) \nabla f(x^k) = -\langle A(x^k - \alpha(Ax^k - b)) - b, Ax^k - b \rangle =$$
$$= -\langle (Ax^k - b) - \alpha(Ax^k - b), A x^k - b \rangle = \langle -r^k + \alpha A r^k, -r^k \rangle = -\langle r^k, r^k \rangle + \alpha \langle A r^k, r^k \rangle$$

Άρα $g'(\alpha) = 0 \Leftrightarrow \alpha = \dfrac{\|r^k\|_2^2}{\langle A r^k, r^k \rangle}$ και $g''(\alpha) = \langle A r^k, r^k \rangle > 0$.

Παρατήρηση 2.1: Από γεωμετρική σκοπιά στη Μέθοδο Απότομης Καθόδου τα διανύσματα-υπόλοιπο είναι ανά δύο κάθετα μεταξύ τους. Άρα, δύο διαδοχικά διανύσματα-υπόλοιπο θα είναι επίσης κάθετα. Δηλαδή έχω:

$$r^{k+1} = b - Ax^{k+1} = b - A(x^k + \alpha_k r^k) = b - Ax^k - \alpha_k A r^k = r^k - \alpha_k A r^k \quad \textbf{(2.3)}$$

Οπότε:

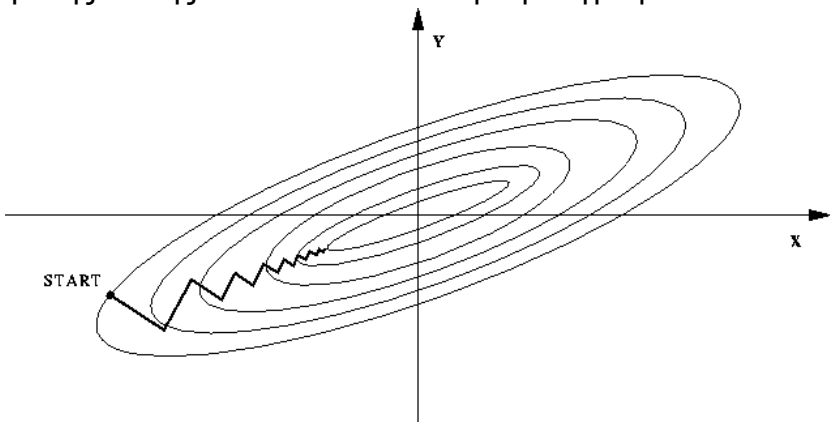
$$\langle r^{k+1}, r^k \rangle = \langle r^k - \alpha_k A r^k, r^k \rangle = \langle r^k, r^k \rangle - \frac{\|r^k\|_2^2}{\langle A r^k, r^k \rangle} \langle A r^k, r^k \rangle = 0 \quad .$$

Άμεση συνέπεια αυτού είναι τα διανύσματα $x^{k+1} - x^k$ και $x^k - x^{k-1}$ είναι ορθογώνια, δηλαδή ισχύει:

$$\langle x^{k+1} - x^k, x^k - x^{k-1} \rangle = 0 \quad .$$

Αυτό αποδεικνύεται εύκολα διότι έχουμε $\langle x^{k+1} - x^k, x^k - x^{k-1} \rangle = \langle \alpha_{k+1} r^k, \alpha_k r^{k-1} \rangle = 0$.

Οι ισοϋψείς επιφάνειες f(x)=σταθ. του συναρτησιακού f είναι ελλειψοειδή στον $\mathbb{R}^n$ με κέντρο το σημείο $A^{-1}b$ και άξονες παράλληλους προς τα ιδιοδιανύσματα $e^i, i = 1,2,\ldots,n$ του πίνακα A, με μήκη αξόνων ανάλογα των αριθμών $\lambda_i^{-1/2}$. Ακόμη, αν $\lambda_{max} \gg \lambda_{min}$ (λ_{max} και λ_{min} η μέγιστη και η ελάχιστη ιδιοτιμή του πίνακα A αντίστοιχα) τα ελλειψοειδή είναι υπερβολικά στενόμακρα και η κάθοδος προς το κέντρο γίνεται με "ζίγκ-ζάγκ" στις διαδοχικές ορθογώνιες κατευθύνσεις $r^k = -\nabla f(x^k)$ μεταξύ των πλαγιών, πράγμα που καθυστερεί πολύ την κάθοδο. Έτσι, η τακτική της ελαχιστοποίησης κατά μήκος των διευθύνσεων μέγιστης κλίσης αποδεικνύεται κακή στρατηγική.



Εικόνα 2: Εδώ η Μέθοδος Απότομης Καθόδου σταματάει τις επαναλήψεις όταν επιτύχουμε σύγκλιση και η γραμμή "ζίγκ-ζάγκ" εμφανίζεται επειδή κάθε κλίση είναι ορθογώνια με την προηγούμενή της.

Παρατήρηση 2.2: Στην περίπτωση που ο $A \in \mathbb{C}^{n,n}$ είναι ένας ερμιτιανός θετικά ορισμένος πίνακας ελαχιστοποιώ το $g(\alpha)=\|r^k-\alpha A r^k\|=\|(I-\alpha A)r^k\|$.

Έστω η νόρμα του διανύσματος-υπόλοιπο στην k+1 επανάληψη:

$$\begin{gathered}(g(\alpha))^2=\langle r^k-\alpha Ar^k, r^k-\alpha A r^k\rangle= \\ \langle r^k, r^k\rangle-\alpha\langle r^k, Ar^k\rangle+|\alpha|^2\langle Ar^k, Ar^k\rangle-\overline{\alpha_k}\langle Ar^k, r^k\rangle= \\ \langle r^k, r^k\rangle-2\Re\langle\alpha(r^k, Ar^k)\rangle+|\alpha|^2\langle Ar^k, Ar^k\rangle.\end{gathered}$$

Γράφω τους αριθμούς α_k και $\langle r^k, Ar^k\rangle$ σε μιγαδική μορφή και έχω $\alpha_k=\alpha+\beta i, \alpha, \beta \in \mathbb{R}$ και $(r^k, Ar^k)=\gamma+\delta i, \gamma, \delta \in \mathbb{R}$.

Επομένως, $F(\alpha,\beta):=\|r^{k+1}\|^2=\|r^k\|^2-2(\gamma\alpha-\beta\delta)+(\alpha^2+\beta^2)\|Ar^k\|^2$.

Το μόνο που απομένει είναι να προσδιορίσουμε τα κρίσιμα σημεία της παραπάνω ποσότητας και έπειτα να βρούμε εκείνο που ελαχιστοποιεί την τιμή της.
Έτσι,

$$\nabla F(\alpha,\beta)=0 \Leftrightarrow \frac{\partial}{\partial\alpha}F(\alpha,\beta)=0 \Leftrightarrow -2\gamma+2\alpha\|Ar_k\|^2=0 \Leftrightarrow \alpha=\frac{\gamma}{\|Ar^k\|^2}$$

και

$$\frac{\partial}{\partial\beta}F(\alpha,\beta)=0 \Leftrightarrow \beta=-\frac{\delta}{\|Ar^k\|^2} \quad .$$

Ορίζουμε τώρα τον Εσσιανό πίνακα της H_F της F, ο οποίος είναι

$$H_F=\begin{bmatrix} 2\|Ar^k\|^2 & 0 \\ 0 & 2\|Ar^k\|^2 \end{bmatrix}.$$

Ο H_F είναι συμμετρικός, αφού προφανώς ισχύει $H_F{}^T=H_F$, και θετικά ορισμένος, οπότε το κρίσιμο σημείο που βρήκαμε είναι το ελάχιστο της F. Κάνοντας αντικατάσταση των α και β προκύπτει ότι:

$$\alpha=\frac{\gamma}{\|Ar_k\|^2}-i\frac{\delta}{\|Ar_k\|^2}=\frac{\gamma-\delta i}{(Ar_k, Ar_k)}=\frac{\langle Ar_k, r_k\rangle}{\langle Ar_k, Ar_k\rangle}=\frac{\|Ax^k-b\|^2}{\langle A(Ax^k-b),(Ax^k-b)\rangle} \quad \textbf{(2.4)}$$

Συμπεραίνουμε, λοιπόν, ότι προκύπτει το ίδιο α με την περίπτωση που $A \in \mathbb{R}^{n,n}$.

Τώρα, αν $\varepsilon^k=x^k-x$ είναι το διάνυσμα-σφάλμα στο k-οστό βήμα της επανάληψης είναι δυνατόν να βρεθούν απλές σχέσεις που να το συνδέουν με το διάνυσμα-υπόλοιπο και οι

οποίες μπορούν να φανούν χρήσιμες στη συνέχεια Για παράδειγμα:

i. $A\varepsilon^k = A(x^k - x) = Ax^k - b = -r^k$

ii. $\langle A\varepsilon^k, r^{k-1}\rangle = 0$

iii. $\varepsilon^k = x^k - x = x^{k-1} + \alpha_{k-1} r^{k-1} - x = \varepsilon^{k-1} + \alpha_{k-1} r^{k-1}$

Λαμβάνοντας υπόψιν όλα τα παραπάνω καταλήγουμε στον ακόλουθο αλγόριθμο για τη Μέθοδο Απότομης Καθόδου.

- <u>Αλγόριθμος της Μεθόδου Απότομης Καθόδου</u>

Δεδομένα/Είσοδος: $A \in \mathbb{C}^{n,n}, A^H = A$, θετικά ορισμένος, $b \in \mathbb{C}^n, \varepsilon \in \mathbb{R}^+$ η μέγιστη επιτρεπτή ανοχή σφάλματος.
Έξοδος: $x = x^k$ (κατά προσέγγιση λύση)
Αρχικοποίηση:
Διαλέγω $x^0 \in \mathbb{R}^n$
Υπολογίζω: $r^0 = b - Ax^0$
$k = 0$
While $\|r^k\|_2 > \varepsilon \|b\|_2$
$\alpha_k = \dfrac{\|r^k\|_2^2}{\langle Ar^k, r^k\rangle}$
$x^{k+1} = x^k + \alpha_k r^k$
$r^{k+1} = r^k - \alpha_k A r^k$
$k = k+1$
End while

<u>Ανάλυση αλγορίθμου:</u> Ο αλγόριθμος τερματίζεται όταν ικανοποιηθούν ένα ή περισσότερα κριτήρια τερματισμού των επαναληπτικών μεθόδων, στην περίπτωσή μας όταν $\|r\|_2 \le \varepsilon \|b\|_2$, με ε να είναι η επιθυμητή ανοχή σφάλματος. Το υπολογιστικό κόστος της Μεθόδου Απότομης Καθόδου εξαρτάται από τον υπολογισμό πίνακα με διάνυσμα και στην περίπτωσή μας είναι $O(n^2)$ πράξεις ανά βήμα, αφού υπάρχουν δύο πολλαπλασιασμοί πίνακα επί διάνυσμα (ο Ax^k και ο Ar^k). Αυτοί οι δύο μπορούν να αναχθούν σε έναν, αφού ισχύει

$$r^{k+1} = b - Ax^{k+1} = b - A(x^k + \alpha_k r^k) = b - Ax^k - \alpha_k A r^k = r^k - \alpha_k A r^k \quad .$$

Έτσι υπολογίζουμε μόνο το Ar^k κατά τον υπολογισμό του α_k και στη συνέχεια το επαναχρησιμοποιούμε για να υπολογίσουμε το r^{k+1} .

Θεώρημα 2.5: Έστω $A \in \mathbb{C}^{n,n}$ ένας ερμιτιανός και θετικά ορισμένος πίνακας, Για τη νόρμα ενέργειας του διανύσματος-σφάλμα που παράγεται από τη Μέθοδο Απότομης Καθόδου ισχύει: $\|\varepsilon^k\|_A \le \left(\dfrac{\kappa - 1}{\kappa + 1}\right)^k \|\varepsilon^0\|_A$, όπου $\kappa = \dfrac{\lambda_n}{\lambda_1}$ ο δείκτης κατάστασης του συστήματος.

Παρατήρηση 2.3: Είναι φανερό ότι η ταχύτητα σύγκλισης της ακολουθίας x^k μπορεί να είναι πολύ μικρή αν ο λόγος $\kappa = \frac{\lambda_n}{\lambda_1}$ είναι μεγάλος, με άλλα λόγια αν ο δείκτης κατάστασης του πίνακα A είναι μεγάλος.

Συμπεράσματα:
Από αυτό το θεώρημα λαμβάνουμε ως συμπέρασμα ότι παράγεται μια επαναληπτική μέθοδος με σκοπό την ελαχιστοποίηση της f. Κατασκευάζουμε, δηλαδή, μια ακολουθία σημείων $(x^k)_k$ τ.ω. η ακολουθία $(f(x^k))_k$ να είναι φθίνουσα. Η x^{k+1} ορίζεται ως το πλησιέστερο στο x^k ελάχιστο της f(x) όταν το x περιορίζεται πάνω στην ακτίνα που διέρχεται από το x^k και έχει τη διεύθυνση του $-\nabla f(x^k)$. Βρίσκουμε, δηλαδή, τον μικρότερο αριθμό που η συνάρτηση $f(x^k - \alpha \nabla f(x^k))$ έχει ελάχιστο, και έτσι τελικά ορίζουμε την $x^{k+1} = x^k - \alpha \nabla f(x^k) = x^k + \alpha(b - Ax^k)$. Με άλλα λόγια, έχουμε δείξει ότι λύνοντας ένα γραμμικό σύστημα του οποίου ο πίνακας είναι συμμετρικός και θετικά ορισμένος είναι ισοδύναμο με την ελαχιστοποίηση μιας τετραγωνικής συνάρτησης.

Ωστόσο, η Μέθοδος Απότομης Καθόδου δεν είναι η αποτελεσματικότερη, διότι παρόλο που η σχέση $r^k = b - Ax^k$ χρειάζεται για τον υπολογισμό του r^0 , η $r^{k+1} = r^k - \alpha_k A r^k$ χρησιμοποιείται για κάθε επανάληψη από εκεί και μετά. Βέβαια, το γινόμενο Ar^k εμφανίζεται στη σχέση **(2.3)** και **(2.4)** και παρόλο που υπολογίζεται μόνο μια φορά, το μειονέκτημα της χρήσης αυτής της αναδρομής είναι ότι η ακολουθία που ορίζεται στη **(2.3)** παράγεται χωρίς καμία ανατροφοδότηση από την τιμή του x^k , ώστε η προσέγγιση του σφάλματος να προκαλεί το x^k να συγκλίνει σε κάποιο σημείο κοντά στο x. Έτσι, προκειμένου να την βελτιώσουμε, πριν περάσουμε στη Μέθοδο Συζυγών Κλίσεων εισάγουμε τη σημαντική έννοια των χώρων Krylov.

Ορισμός 2.2: (Μέθοδος Υποχώρου Krylov). Η μέθοδος υποχώρου Krylov είναι μια επαναληπτική μέθοδος ελαχιστοποίησης με K_k να είναι ο υπόχωρος Krylov

$$K_k(A, r) = span\{r, Ar, \ldots, A^{k-1} r\} \quad .$$

Ο υπόχωρος $K_k(A, r)$ θα συμβολίζεται με K_k όταν δεν υπάρχει αμφισημία.

Επιπλέον,

- Η διάσταση του υποχώρου Krylov αυξάνεται κατά ένα σε κάθε βήμα της επαναληπτικής διαδικασίας. Μια βασική ιδιότητα του υποχώρου Krylov είναι πως ο K_k είναι ο υπόχωρος όλων των διανυσμάτων στον $\mathbb{R}^n$ που μπορούν να γραφούν στη μορφή $x = p(A) r$, όπου με p συμβολίζουμε ένα πολυώνυμο βαθμού το πολύ k-1. Το ελάχιστο πολυώνυμο ενός διανύσματος r είναι το μικρότερου βαθμού μη μηδενικό πολυώνυμο p , τ.ω. $p(A) r = 0$. Ο βαθμός του ελαχίστου πολυωνύμου του διανύσματος r ως προς τον πίνακα A καλείται βαθμός του r ως προς τον A, ή απλά βαθμός του R όταν δεν υπάρχει αμφισημία και συμβολίζεται με $\mu = grade(r)$.

- Μια δεύτερη ιδιότητα των υποχώρων Krylov είναι ότι αν μ είναι ο βαθμός του r, τότε ο υπόχωρος K_μ είναι αναλλοίωτος στην επίδραση του πίνακα A και $K_k = K_\mu$ για

όλα τα $k \geq \mu$.

- Μια τρίτη ιδιότητα των υποχώρων Krylov είναι ότι ο υπόχωρος K_k είναι διάστασης k εαν και μόνο αν ο βαθμός μ του r ως προς τον A δεν είναι μικρότερος του k, δηλαδή

$$dim(K_k) = k \Leftrightarrow \mu \geq k \quad .$$

Επομένως,

$$dim(K_k) = min\{k, \mu\} \quad .$$

Αυτή η ιδιότητα αποδεικνύεται εύκολα.

Έστω ο k-διάστατος υπόχωρος Krylov

$$K_k = \{r, Ar, A^2 r, \ldots, A^{k-1} r\} \quad .$$

Τα διανύσματα $r, Ar, A^2 r, \ldots, A^{k-1} r$ αποτελούν βάση του K_k , άρα είναι γραμμικά ανεξάρτητα. Δηλαδή για κάποιο σύνολο k πραγματικών αριθμών α_i , με τουλάχιστον ένα από τα $\alpha_i \neq 0$, έπεται ότι ο γραμμικός συνδυασμός $\sum_{i=0}^{k-1} \alpha_i A^i b = r_{k-1}(A) r$ είναι διάφορος του μηδενικού διανύσματος. Ισοδύναμα, το μοναδικό πολυώνυμο βαθμού $\leq k-1$ με $p(A) r = 0$ είναι το μηδενικό πολυώνυμο. Επομένως, $\mu \geq k$.

Έστω $\mu < k$ και $p_\mu(A)$ το ελάχιστο πολυώνυμο του r. Τότε, για $y \in K_k$ υπάρχει πολυώνυμο q_{k-1} το πολύ k-1 βαθμού τ.ω. $y = q_{k-1}(A) r$. Συνεπώς,

$$q_{k-1}(A) r = (s_{k-\mu-1}(A) p_\mu(A) + v_{\mu-1}(A)) r =$$

$$= s_{k-\mu-1}(A) p_\mu(A) r + v_{\mu-1}(A) r =$$

$$= v_{\mu-1}(A) r \in K_\mu,$$

όπου s είναι το πολυώνυμο-πηλίκο της διαίρεσης $q_k(A) / p_k(A)$ βαθμού το πολύ k-μ-1 και v το αντίστοιχο υπόλοιπο βαθμού το πολύ μ-1 αντίστοιχα. Αυτό σημαίνει ότι $dim(K_k) = \mu$. Άρα τελικά, $dim(K_k) = min\{k, \mu\}$. □

Οι χώροι Krylov $(K_k)_{k \geq 0}$ σχηματίζουν μια αύξουσα ακολουθία από διανυσματικούς υπόχωρους. Αφού $K_k \subset \mathbb{R}^n$, αυτή η ακολουθία γίνεται σταθερή για κάποιο k.

Η ακολουθία των χώρων Krylov όπως αναφέραμε είναι αύξουσα και ισχύει $K_k \subset K_{k+1}, \forall k \geq 0$. Επιπλέον για όλα τα διανύσματα $r_0 \neq 0, \exists k_0 \in \{0,1,\ldots,n-1\} \tau.\omega.$:

$$\begin{Bmatrix} dim\, K_k = k+1, 0 \leq \kappa \leq \kappa_0 \\ dim\, K_k = k_0 + 1, k_0 \leq k \end{Bmatrix} \quad .$$

Ο ακέραιος k_0 ονομάζεται κρίσιμη διάσταση Krylov.

Τώρα, λοιπόν, αφού ορίσαμε τον χώρο Krylov θεωρούμε τη Μέθοδο της Απότομης Καθόδου (με σταθερό ή μεταβλητό αριθμό βημάτων) ως εξής:

$$\left\{ \begin{array}{c} x^0 \in \mathbb{R}^n, \alpha\rho\chi\iota\kappa\acute{\eta}\ \epsilon\pi\iota\lambda o\gamma\acute{\eta} \\ x^{k+1} = x^k + \alpha_k (b - Ax^k) \end{array} \right\} .$$

Το διάνυσμα $r^k = b - Ax^k$, που ονομάζεται υπόλοιπο ικανοποιεί τις ακόλουθες ιδιότητες:

i. Το r^k ανήκει στον χώρο Krylov K_k που αντιστοιχεί στο αρχικό υπόλοιπο r^0
ii. Το x^{k+1} ανήκει στον αφφινικό χώρο $[\chi_0 + K_k]$ και ορίζεται ως το σύνολο των διανυσμάτων x τ.ω. το $x-x_0$ να ανήκει στον διανυσματικό υπόχωρο K_k.

ΚΕΦΑΛΑΙΟ 3

Μέθοδος Συζυγών Κλισεων

3.1 Μέθοδος Gram-Schmidt και A-ορθογωνιότητα

Πριν περάσουμε στη Μέθοδο των Συζυγών Κλίσεων θα αναφέρουμε κάποιες βασικές έννοιες ως προς την μέθοδο Gram-Schmidt και την A-ορθογωνιότητα.

Θεωρώ τον A να είναι ένας συμμετρικός και θετικά ορισμένος πίνακας. Θεωρώ τα διανύσματα $p^k, k=0,\ldots,k_0$ τα οποία είναι ορθογώνια και τα γραμμικώς ανεξάρτητα διανύσματα $r^k, k=0,\ldots,k_0$. Για να κατασκευάσουμε τα διανύσματα p^k παίρνουμε τα r^k και αφαιρούμε τις συνιστώσες που δεν είναι A-ορθογώνιες με τα προηγούμενα διανύσματα διεύθυνσης.
Με άλλα λόγια, θέτω $p^0=r^0$ και για $0\le k\le k_0$ εφαρμόζουμε την μέθοδο των Gram-Schmidt:

$$p^k=r^k+\sum_{j=0}^{k-1}\beta_{kj}\,p^j \quad ,$$

όπου το β_{kj} ορίζεται για $k>j$. Τώρα, το μόνο που μένει είναι να βρούμε τις τιμές για το β_{ki} .
Απαίτηση για A-ορθογωνιότητα:

$$\langle Ap^i,p^k\rangle=\langle Ap^i,r^k+\sum_{j=0}^{k-1}\kappa_{kj}\,p^j\rangle=\langle Ap^i,r^k\rangle+\sum_{j=0}^{k-1}\beta_{kj}\langle Ap^i,p^j\rangle=\langle Ap^i,r^k\rangle+\beta_{kj}\langle Ap^i,p^i\rangle=0$$

Άρα, $\beta_{ki}=-\dfrac{\langle Ap^i,r^k\rangle}{\langle Ap^i,p^i\rangle}$, k>i.

3.2 Μέθοδος Συζυγών Κλίσεων (Conjugate Gradient Method)

Η μέθοδος ανήκει στην κατηγορία μεθόδων ελαχιστοποίησης και αποτελεί βελτίωση της Μεθόδου Απότομης Καθόδου. Η μέθοδος προτάθηκε από τους Hestenes και Stiefel, οι οποίοι παρουσίασαν τις μεθόδους τους το 1952 και χρησιμοποιείται για την επίλυση γραμμικών συστημάτων της μορφής $Ax=b$ με συμμετρικό και θετικά ορισμένο πίνακα συντελεστών A και b σταθερό διάνυσμα. Η μέθοδος αυτή στοχεύει στην εύρεση της λύσης, έστω x^* , του συστήματος. Όπως και στη Μέθοδο Απότομης Καθόδου, η τιμή x^* θα ταυτίζεται με το σημείο ελαχιστοποίησης της συνάρτησης f με τετραγωνική μορφή

$$f(x)=\frac{1}{2}\langle Ax,x\rangle-\langle b,x\rangle \quad .$$

Δοθέντος, λοιπόν, μιας συνάρτησης f και μιας αρχικής προσέγγισης x^0 , αναζητούμε τη λύση του προβλήματος.

Έστω, λοιπόν, το σύστημα $Ax^*=b$ με $A\in\mathbb{R}^{n,n}$ να είναι ένας συμμετρικός και θετικά ορισμένος πίνακας και $b\in\mathbb{R}^{n\setminus\{0\}}$.

Επιπλέον, ορίζουμε $D_1=K_1(A,b)$, $D_2=K_2(A,b)$ ...,
όπου $K_k(A,b)=\{b,Ab,\dots,A^{k-1}b\}$.

Η ιδέα είναι ότι, αφού ο A είναι ένας συμμετρικός και θετικά ορισμένος πίνακας, το $\langle x,y\rangle_A=\langle Ax,y\rangle_2$ είναι ένα εσωτερικό γινόμενο στο $\mathbb{R}^n$. Έπειτα, ορίζουμε το $x^k\in D_k$ να είναι η βέλτιστη προσέγγιση του x^* από τον D_k . Αυτό ισοδυναμεί με:

$$\langle x^k-x^*,y\rangle_A=0,\forall\, y\in D_k=K_k(A,b)$$

$$\Leftrightarrow\langle A\,x^k-A\,x^*,y\rangle_2=0,\forall\, y\in D_k=K_k(A,b)$$

$$\Leftrightarrow\langle A\,x^k-b,y\rangle_2=0,\forall\, y\in D_k=K_k(A,b)$$

$$\Leftrightarrow\langle r^k,y\rangle_2=0,\forall\, y\in D_k=K_k(A,b) \qquad \textbf{(3.1)}$$

Στόχος μας είναι να υπολογίσουμε τα x^k . Υποθέτουμε ότι έχουμε μια βάση $(p^j)_{j=0}{}^{n-1}$ του $\mathbb{R}^n$ με διανύσματα τα οποία είναι A-ορθογώνια, δηλαδή

$$\langle A\,p^j,p^\ell\rangle=0,\ell\neq j,\ell,j=0,\dots,n-1 \qquad \textbf{(3.2)}$$

και ας υποθέσουμε: $D_k=span\{p^0,\dots,p^{k-1}\}$. Τότε:

$$x^k=\sum_{j=0}^{k-1}\frac{\langle x,p^j\rangle_A}{\langle p^j,p^j\rangle_A}p^j$$

$$\Leftrightarrow x^k=\sum_{j=0}^{k-1}\frac{\langle Ax,p^j\rangle_2}{\langle p^j,p^j\rangle_A}p^j$$

$$\Leftrightarrow x^k=\sum_{j=0}^{k-1}\frac{\langle b,p^j\rangle_2}{\langle p^j,p^j\rangle_A}p^j=\sum_{j=0}^{k-1}\frac{\langle r^j,p^j\rangle_2}{\langle p^j,p^j\rangle_A}p^j$$

Για k=0: $x^0=0$ και $r^0=b$

Για k=1: $x^1=\frac{\langle b,p^0\rangle_2}{\langle p^0,p^0\rangle_A}p^0$

Για k=2: $x^2=\frac{\langle b,p^0\rangle_2}{\langle p^0,p^0\rangle_A}p^0+\frac{\langle b,p^1\rangle_2}{\langle p^1,p^1\rangle_A}p^1$

$$\langle r^j, p^j\rangle = \langle b - A x^j, p^j\rangle$$
$$= \langle b, p^j\rangle - \langle A x^j, p^j\rangle$$

Για $j \geq 1$: $$= \langle b, p^j\rangle - \langle A \sum_{\ell=0}^{j-1} \frac{\langle b, p^\ell\rangle_2}{\langle p^\ell, p^\ell\rangle_A} p^\ell, p^j\rangle$$
$$= \langle b, p^j\rangle - \sum_{\ell=0}^{j-1} \frac{\langle b, p^\ell\rangle_2}{\langle p^\ell, p^\ell\rangle_A} \langle A p^\ell, p^j\rangle$$

Οπότε από τη σχέση **(3.2)** προκύπτει ότι $\langle r^j, p^j\rangle = \langle b, p^j\rangle$.

Το πρόβλημα είναι η επιλογή των $(p^j)_{j=0}^{n-1}$. Θεωρώ τον $span\{p^0, \dots, p^{k-1}\} = K_k(A, b)$, οπότε έχω:

$$x^{k+1} = \sum_{j=0}^{k} \frac{\langle r^j, p^j\rangle}{\langle p^j, p^j\rangle_A} p^j = \frac{\langle r^k, p^k\rangle}{\langle p^k, p^k\rangle_A} + \sum_{j=0}^{k-1} \frac{\langle r^j, p^j\rangle}{\langle p^j, p^j\rangle_A} p^j$$

$$\Rightarrow x^{k+1} = x^k + \sum_{j=0}^{k-1} \frac{\langle r^j, p^j\rangle}{\langle p^j, p^j\rangle_A} p^j$$, με k=0,...

$$\Rightarrow x^{k+1} = x^k + \alpha_\kappa p^k$$, με $$\alpha_k = \frac{\langle r^k, p^k\rangle}{\langle p^k, p^k\rangle_A}$$

Επίσης, για το αντίστοιχο διάνυσμα-υπόλοιπο στην κ+1 επανάληψη ισχύει:

$$r^{k+1} = b - A x^{k+1} = b - A x^k - \frac{\langle r^k, p^k\rangle_A}{\langle p^k, p^k\rangle_A} p^k$$

$$\Rightarrow r^{k+1} = r^k - \frac{\langle r^k, p^k\rangle_2}{\langle p^k, p^k\rangle_A} A p^k$$.

Πρόταση 3.2.1: Ο υπόχωρος $K_k(A, b)$ που ορίζεται ως $K_k = span\{b, Ab, \dots, A^k b\}$, ικανοποιεί τη σχέση:

$$span\{b = r^0, \dots, r^{k-1}\} = span\{b = p^0, \dots, p^{k-1}\} = K_k(A, b)$$

όπου p^0=b και $$p^k = r^k - \frac{\langle A r^k, p^{k-1}\rangle_2}{\langle A p^{k-1}, p^{k-1}\rangle_2} p^{k-1}$$, με κ=1,... .

Απόδειξη: Εφαρμόζω επαγωγή, οπότε έχω:

Για k=1: $span\{p^0 = b\} = span\{r^0 = b\} = K_1(A, b) = span\{b\}$ ✓

Για $k \to k+1$: $span\{p^0, \dots, p^{k-1}\} = span\{r^0, \dots, r^{k-1}\} = K_k(A, b) = span\{b, \dots, A^{k-1} b\}$

$$p^k = r^k - \sum_{j=0}^{k-1} \frac{\langle r^k, p^j \rangle_A}{\langle p^j, p^j \rangle_A} p^j = r^k - \frac{\langle r^k, p^{k-1} \rangle_A}{\langle p^{k-1}, p^{k-1} \rangle_A} p^{k-1} - \sum_{j=0}^{k-2} \frac{\langle r^k, p^j \rangle_A}{\langle p^j, p^j \rangle_A} p^j \quad .$$

Παρατηρούμε μετά ότι τα $\{p^0, \dots, p^{k-1}, r^k\}$ είναι γραμμικώς ανεξάρτητα όταν $r^k \neq 0$,

διότι $\langle r^k, y \rangle_2 = 0, \forall\, y \in K_k(A, b)$

$\Rightarrow \langle r^k, p^j \rangle_2 = 0$, επειδή $\langle r^k, p^j \rangle \in K_k(A, b)$.

Για $0 \leq j \leq k-2$: $\langle r^k, p^j \rangle_A = \langle Ar^k, p^j \rangle_2 = \langle r^k, Ap^j \rangle_2 = 0$, αφού ο A είναι συμμετρικός

και ισχύει $p^j \in span\{p^0, \dots, p^{k-2}\} = K_{k-1}(A, b)$ και $Ap^j \in K_k(A, b)$

και έχω από Gram-Schmidt $span\{p^0, \dots, p^k\} = span\{r^0, \dots, r^k\}$. **(3.3)**

Έτσι, $p^k = r^k - \frac{\langle r^k, p^{k-1} \rangle_A}{\langle p^{k-1}, p^{k-1} \rangle_A} p^{k-1}$.

Το μόνο που μένει να αποδείξουμε είναι ότι η σχέση **(3.3)** ισούται με $K_{k+1}(A, b)$.

Θέλουμε να δείξουμε ότι $r^k \in K_{k+1}(A, b)$.

Ισχύει ότι $r^k = r^{k-1} - \frac{\langle r^{k-1}, p^{k-1} \rangle_2}{\langle p^{k-1}, p^{k-1} \rangle_A} A\, p^{k-1}$, με $p^{k-1} \in K_k(A, b)$ και $Ap^{k-1} \in K_{k+1}(A, b)$.

Επίσης $r^{k-1} \in K_k(A, b)$. Οπότε $r^k \in K_{k+1}(A, b)$. □

Από την παραπάνω πρόταση προκύπτει ότι: $p^k = r^k - \frac{\langle Ar^k, p^{k-1} \rangle_2}{\langle Ap^{k-1}, p^{k-1} \rangle_2} p^{k-1} = r^k + \beta_{k-1} p^{k-1}$.
Ολοκληρώνουμε, λοιπόν, τη Μέθοδο Συζυγών Κλίσεων δείχνοντας ότι $\beta_{k-1} = \frac{\langle r^k, r^k \rangle}{\langle r^k, r^k \rangle} = \frac{\|r^k\|^2}{\|r^{k-1}\|^2}$ και $\alpha_k = \frac{\|r^k\|_2^{\,2}}{\langle Ap^k, p^k \rangle_2}$.

Πρώτα απ΄ όλα ισχύει $\langle r^k, p^k \rangle_2 = \langle r^k, r^k - \frac{\langle Ar^k, p^{k-1} \rangle_2}{\langle A\, p^{k-1}, p^{k-1} \rangle_2} p^{k-1} \rangle = \|r^k\|^2 - \frac{\langle Ar^k, p^{k-1} \rangle_2}{\langle A\, p^{k-1}, p^{k-1} \rangle_2} \langle p^{k-1}, r^k \rangle$
.

Όμως ισχύει $r^k \perp K_k(A, b)$, αφού $p^{k-1} \in K_k(A, b)$. Οπότε, έπεται ότι $\|r^k\|^2 = \langle r^k, p^k \rangle_2$.

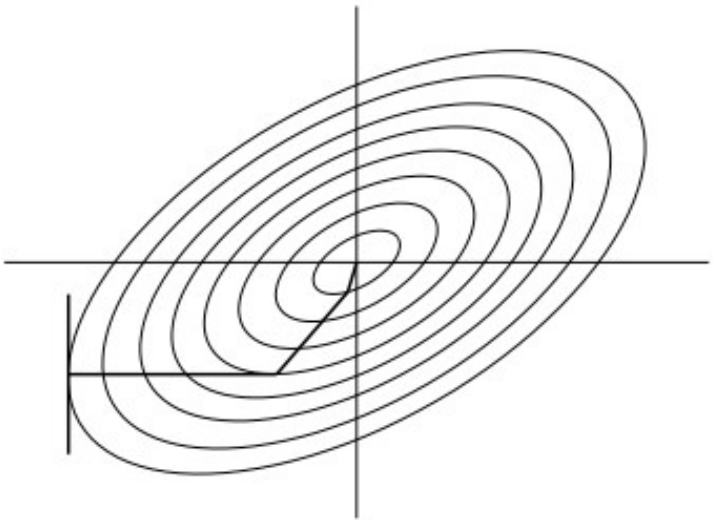
Ακόμη έχουμε ότι $r^k = r^{k-1} - \frac{\langle r^{k-1}, p^{k-1} \rangle_2}{\langle p^{k-1}, p^{k-1} \rangle_A} A\, p^{k-1} = r^{k-1} - \frac{\langle r^{k-1}, r^{k-1} \rangle_2}{\langle p^{k-1}, p^{k-1} \rangle_A} A\, p^{k-1}$.

Οπότε $\|r^k\|^2=\langle r^k,r^{k-1}\rangle-\frac{\|r^{k-1}\|^2}{\|p^{k-1}\|_A{}^2}\langle A\,p^{k-1},r^k\rangle_2=-\frac{\|r^{k-1}\|^2}{\|p^{k-1}\|_A{}^2}\langle A\,p^{k-1},r^k\rangle_2$, διότι έχουμε δείξει ότι $\langle r^k,r^{k-1}\rangle=0$, αφού $r^{k-1}\in K_k(A,b)$. Επομένως, $\beta_{\kappa-1}=-\frac{\langle p^{k-1},A\,r^k\rangle_2}{\|p^{k-1}\|_A{}^2}=\frac{\|r^k\|^2}{\|r^{k-1}\|^2}$.

- <u>Αλγόριθμος της Μεθόδου Συζυγών Κλίσεων</u>

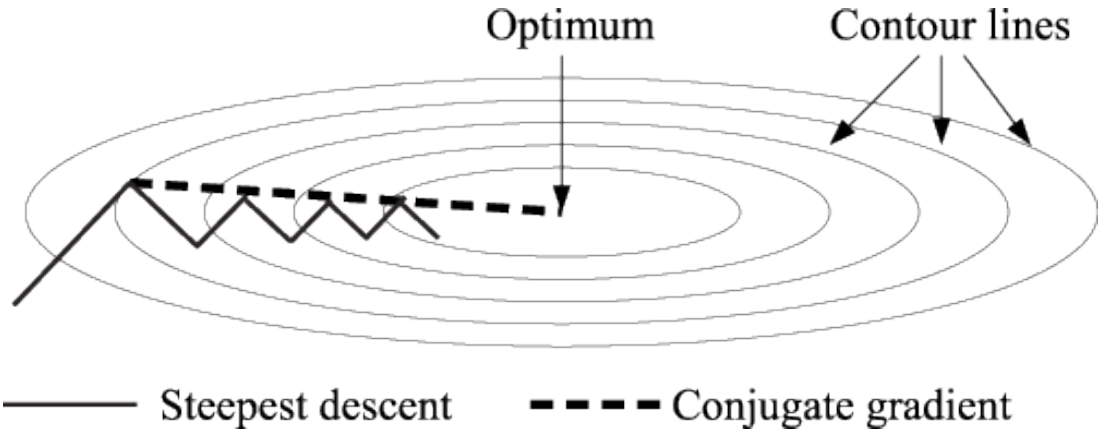
Δεδομένα/Είσοδος: $A\in\mathbb{C}^{n,n},A^H=A$, θετικά ορισμένος, $b\in\mathbb{C}^n,\varepsilon\in\mathbb{R}^+$ η μέγιστη επιτρεπτή ανοχή σφάλματος.
Έξοδος: $x=x^k$ (κατά προσέγγιση λύση)
Αρχικοποίηση:
Διαλέγω $x^0\in\mathbb{R}^n$
Υπολογίζω $r^0=b-Ax^0$
Θέτω $p^0=r^0$
$\alpha_0=\frac{\|r^0\|^2}{\langle Ap^0,p^0\rangle}$
$x^1=x^0+\alpha_0\,p^0$
$r^1=b-Ax^1$
$k=1$
While $\|r^k\|>\varepsilon$ #επαναληπτική διαδικασία
$\beta_k=\frac{\|r^{k+1}\|^2}{\|r^k\|^2}$ #υπολογισμός της παραμέτρου β
$p^k=r^k+\beta_{k-1}\,p^{k-1}$ #ανανέωση του διανύσματος κατεύθυνσης
$\alpha_k=\frac{\|r^k\|^2}{\langle Ap^k,p^k\rangle}$ #υπολογισμός του βήματος προς την κατεύθυνση αναζήτησης
$x^{k+1}=x^k+\alpha_k\,p^k$ #υπολογισμός της νέας προσέγγισης της λύσης
$r^{k+1}=b-Ax^k$ #ανανέωση του διανύσματος υπολοίπου
$k=k+1$
End While

<u>Σχόλια:</u> Αν θεωρήσουμε τη Μέθοδο Συζυγών Κλίσεων ως μια άμεση μέθοδο, μπορούμε να υπολογίσουμε τον αριθμό των απαραίτητων για την επίλυση του γραμμικού συστήματος πράξεων στις περισσότερες δυσμενείς περιπτώσεις, $k_0=n-1$. Σε κάθε επανάληψη, από την μία πλευρά, προσμετράται το γινόμενο ενός πίνακα διάνυσμα ($O(n^2)$), και από την άλλη πλευρά, δύο βαθμωτά γινόμενα και τρις γραμμικώς ανεξάρτητοι συνδυασμοί διανυσμάτων όπως το $x^k+\alpha_k A\,p^k$. Μετά από n επαναλήψεις, βρίσκουμε έναν αριθμό πράξεων ίσο περίπου με n^3 . Ο αριθμός αυτός είναι λιγότερο αποτελεσματικός από τις μεθόδους Gauss και Cholesky. Αλλά πρέπει να θυμηθούμε ότι χρησιμοποιείται σαν μια επαναληπτική μέθοδος και γενικά συγκλίνει σε λιγότερες από n επαναλήψεις. Στην πράξη, όμως, μας ενδιαφέρει το κριτήριο σύγκλισης ($\|r^k\|\le\varepsilon$), που επιτυγχάνεται πολύ πιο πριν το k γίνει ίσο με n σε προβλήματα με καλή κατάσταση. Σε προβλήματα με κακή κατάσταση προτιμάται η τεχνική της προρύθμισης, η οποία μετατρέπει το πρόβλημα σε ένα ισοδύναμο, με καλή κατάσταση. Τη μέθοδο αυτή θα τη δούμε αναλυτικά στη συνέχεια.



Εικόνα 3: Το διάνυσμα κατεύθυνσης στη Μέθοδο Συζυγών Κλίσεων

Στη Μέθοδο Συζυγών Κλίσεων, όπως είδαμε, κάθε διάνυσμα είναι κάθετο σε όλα τα προηγούμενα, άρα είναι και γραμμικώς ανεξάρτητα. Άρα, σε σύγκριση με τη Μέθοδο της Απότομης Καθόδου, η μέθοδος αυτή είναι ταχύτερη και πιο ευσταθής. Αυτό φαίνεται και στο παρακάτω σχήμα (Εικόνα 4).



Εικόνα 4: Η Μέθοδος Συζυγών Κλίσεων (διακεκομμένη γραμμή) συγκλίνει πιο γρήγορα από τη Μέθοδο Απότομης Καθόδου (ευθεία γραμμή).

3.2.1. Σύγκλιση της Μεθόδου Συζυγών Κλίσεων

Υπενθυμίζουμε ότι η Μέθοδος Απότομης Καθόδου μειώνει το σφάλμα σε κάθε επανάληψη με συντελεστή $\frac{\kappa(A)-1}{\kappa(A)+1}$. Τώρα, θα αποδείξουμε ότι η ταχύτητα σύγκλισης της Μεθόδου Συζυγών Κλίσεων είναι γρηγορότερη, αφού μειώνει το σφάλμα σε κάθε επανάληψη με μικρότερο συντελεστή $\frac{\sqrt{(\kappa(A))}-1}{\sqrt{(\kappa(A))}+1}$.

Πρόταση 3.2.1.1: Έστω A ένας συμμετρικός και θετικά ορισμένος πίνακας και x η ακριβής λύση του συστήματος $Ax=b$. Όταν στο σύστημα αυτό εφαρμόσουμε τη Μέθοδο Συζυγών Κλίσεων παρατηρούμε ότι συγκλίνει το πολύ σε n επαναλήψεις, δηλαδή συγκλίνει γραμμικά, οπότε έχουμε:

$$\frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|}\leq\frac{\sqrt{(\kappa(A))}-1}{\sqrt{(\kappa(A))}+1} .$$

Απόδειξη: Αρχικά, σε κάθε βήμα της μεθόδου μας, η τιμή του σφάλματος e^k επιλέγεται από τον υπόχωρο e^0+K_k , όπου

$$span\{r^0, Ar^0, A^2r^0,\dots,A^{k-1}r^0\}=span\{Ae^0, A^2e^0, A^3e^0,\dots,A^ke^0\}=K_k(A,b) .$$

Επιπλέον, ισχύει $x^k=x^0+\sum_{j=0}^{k-1}\alpha_j p^j$, με κάθε p^j να είναι ένα πολυώνυμο του A (βαθμού $\leq j$). Επομένως, υπάρχει ένα πολυώνυμο $q_{k-1}\in P_{k-1}$, τ.ω. $x^k=x^0+q_{k-1}(A)p^0$. Γνωρίζουμε ότι $\sum_{j=0}^{k-1}\alpha_j p^j\in span\{p^0,\dots,p^{k-1}\}=K_k(A,b)$ και αφού $p^0=r^0=b-Ax^0=A(x-x^0)$ το σφάλμα e^k παίρνει τη μορφή

$$e^k=x^k-x=e^0+q_{k-1}(A)p^0=Q_k(A)e^0 ,$$

όπου το Q_k είναι πολυώνυμο βαθμού το πολύ k. Αυτό ισχύει διότι

Συμβολίζουμε με u_j την ορθογώνια βάση των ιδιοδιανυσμάτων του πίνακα A, τ.ω. $Au_j=\lambda_j u_j$, και με $e_{0,j}$ τις καταχωρήσεις του αρχικού σφάλματος e^0 σε αυτή τη βάση. Άρα εφαρμόζοντας το πολυώνυμο Q_k στο e^0 παίρνουμε:

$$e^0=\sum_{j=1}^{n} e_{0,j}u_j$$

$$\|e^0\|^2_A=\langle e^0, Ae^0\rangle=\sum_{j=1}^{n}\lambda_j|e_{0,j}|^2$$

και

$$\|e^k\|^2_A=\|Q_k(A)e^0\|^2_A=\sum_{j=1}^{n}\lambda_j|e_{0,j}Q_k(\lambda_j)|^2 .$$

Αφού η Μέθοδος Συζυγών Κλίσεων βρίσκει το πολυώνυμο που ελαχιστοποιεί την παραπάνω σχέση, έχουμε την απαίτηση $Q_k(0)=1$. Έτσι, το πρόβλημά μας ανάγεται σε ένα διακριτό πρόβλημα της μορφής

$$\frac{\|e^k\|_A^2}{\|e^0\|_A^2} \le \min_{Q_k \in P_k} \max_{1 \le j \le n} |Q_k(\lambda_j)|^2 \le \min_{Q_k \in P_k} \max_{\lambda_1 \le x \le \lambda_n} |Q_k(x)|^2 \quad \textbf{(3.4)}$$

στον χώρο P_k των πολυωνύμων βαθμού το πολύ k. Στόχος μας είναι να ελαχιστοποιήσουμε την **(3.4)** στο διάστημα $[\lambda_1, \lambda_n]$, όπου λ_1 η ελάχιστη και λ_n η μέγιστη ιδιοτιμή του συνόλου ιδιοτιμών του πίνακα A. Απάντηση σε προβλήματα αυτής της μορφής δίνουν τα **πολυώνυμα Chebyshev**.

Τα πολυώνυμα πρώτου είδους, βαθμού k, ορίζονται ως εξής:

$$T_k(t) = \cos(k \arccos t), \forall t \in [-1,1] \quad .$$

Ισχύει προφανώς ότι $|T_k(t)| \le 1$, δηλαδή τα πολυώνυμα φράσσονται κατά απόλυτη τιμή από τη μονάδα και επίσης το $|T_k(t)|$ είναι μέγιστο για $t \notin [-1,1]$.

Η σχέση **(3.4)** ελαχιστοποιείται όταν επιλέγουμε ως πολυώνυμο το

$$Q_k(x) = \frac{T_k\left(\frac{2x-(\lambda_1+\lambda_n)}{\lambda_n-\lambda_1}\right)}{T_k\left(\frac{-(\lambda_1+\lambda_n)}{\lambda_n-\lambda_1}\right)}$$

και η μέγιστη τιμή που λαμβάνεται στο διάστημα $[\lambda_1, \lambda_n]$ ισούται με

$$\frac{1}{\left|T_k\left(\frac{-(\lambda_1+\lambda_n)}{\lambda_n-\lambda_1}\right)\right|} \quad .$$

Οπότε καταλήγουμε στη σχέση

$$\min_{Q_k \in P_k} \max_{\lambda_1 \le x \le \lambda_n} |Q_k(x)| = \frac{1}{\left|T_k\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)\right|} \quad \textbf{(3.5)}.$$

Παρατηρώ ότι

- Καθώς ο A είναι συμμετρικός έχουμε $\frac{\lambda_1+\lambda_n}{\lambda_n-\lambda_1} = \frac{\kappa(A)+1}{\kappa(A)-1}$, όπου $\kappa(A) = \frac{\lambda_n}{\lambda_1}$
- Τα πολυώνυμα Chebyshev ικανοποιούν τη σχέση

$$T_k(x) = \frac{1}{2}\left(\left[x+\sqrt{x^2-1}\right]^k + \left[x-\sqrt{x^2-1}\right]^k\right) \quad .$$

Οπότε, συμπεραίνουμε το ακόλουθο κάτω φράγμα:

$$\left(2T_K\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)\right)^{\frac{1}{k}}\geq\frac{\kappa(A)+1}{\kappa(A)-1}+\sqrt{\left(\frac{\kappa(A)+1}{\kappa(A)-1}\right)^2-1}\geq\frac{1}{\kappa(A)-1}((\kappa(A)+1)+2\sqrt{\kappa(A)})=\frac{\sqrt{\kappa(A)}+1}{\sqrt{\kappa(A)}-1}$$

Συνδυάζοντας τα αποτελέσματα της ισότητας **(3.5)** και του άνω φράγματος **(3.4)** συμπεραίνουμε ότι

$$\frac{\|e^k\|^2}{\|e^0\|^2}\leq\left[T_k\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)\right]^{-2}=\left[T_k\left(\frac{\kappa(A)+1}{\kappa(A)-1}\right)\right]^{-2}=4\left[\left(\frac{\sqrt{\kappa(A)}+1}{\sqrt{\kappa(A)}-1}\right)^k+\left(\frac{\sqrt{\kappa(A)}-1}{\sqrt{\kappa(A)}+1}\right)^k\right]^{-2} .$$

Όμως, ο δεύτερος προσθετέος μέσα στην αγκύλη συγκλίνει στο 0 καθώς το i αυξάνεται, δηλαδή $\left(\frac{\sqrt{\kappa(A)}-1}{\sqrt{\kappa(A)}+1}\right)\to 0$. Οπότε προκύπτει άμεσα ότι

$$\frac{\|e^k\|^2}{\|e^0\|^2}\leq 4\left(\frac{\sqrt{\kappa(A)}+1}{\sqrt{\kappa(A)}-1}\right)^{2k} . \qquad \square$$

Ωστόσο, αν υποθέσουμε ότι αναφερόμαστε σε πραγματικά τόξα τα οποία ορίζονται στο διάστημα $[-1,1]$ και κατ' επέκταση στο $\mathbb{R}$, τότε θέτω

$$arcos(t)=\theta(\in\mathbb{R})\Leftrightarrow t=\cos(\theta) .$$

Οπότε ισχύει:

$$|\cos(\theta)|\leq 1\Leftrightarrow -1\leq\cos(\theta)\leq 1 .$$

Άρα, σε αυτή την περίπτωση τα πολυώνυμα Chebyshev $T_k(k\geq 1)$ δίνονται από την ακόλουθη σχέση

$$T_\kappa(t)=\cos(k\dot{\theta}) \quad , \quad \theta\in[0,\pi] .$$

Εύκολα παρατηρούμε ότι $T_0(t)=1$, $T_1(t)=t$ και $T_{k+1}(t)=2tT_k(t)-T_{k-1}(t),\forall k\geq 1$ **(3.6)**, γεγονός που μας δείχνει ότι το T_k είναι ένα πολυώνυμο βαθμού k και μπορεί να ορισθεί σε ολόκληρο το $\mathbb{R}$.

Η απόδειξη της σχέσης **(3.6)** δίνεται μέσω της τριγωνομετρικής ταυτότητας

$$\cos(k+1)\theta+\cos(k-1)\theta=2\cos(\theta)\cos(k\theta) .$$

Θέτω $\theta=arcos(t)\Leftrightarrow t=\cos(\theta)$ οπότε έχω:

$$\cos(k+1)arcos(t)+\cos(k-1)arcos(t)=2\cos(arcos(t))\cos(k\,arcos(t))\Leftrightarrow$$

$$\Leftrightarrow T_{k+1}(\theta)-2tT_k(\theta)+T_{k-1}(\theta)=0 . \qquad \square$$

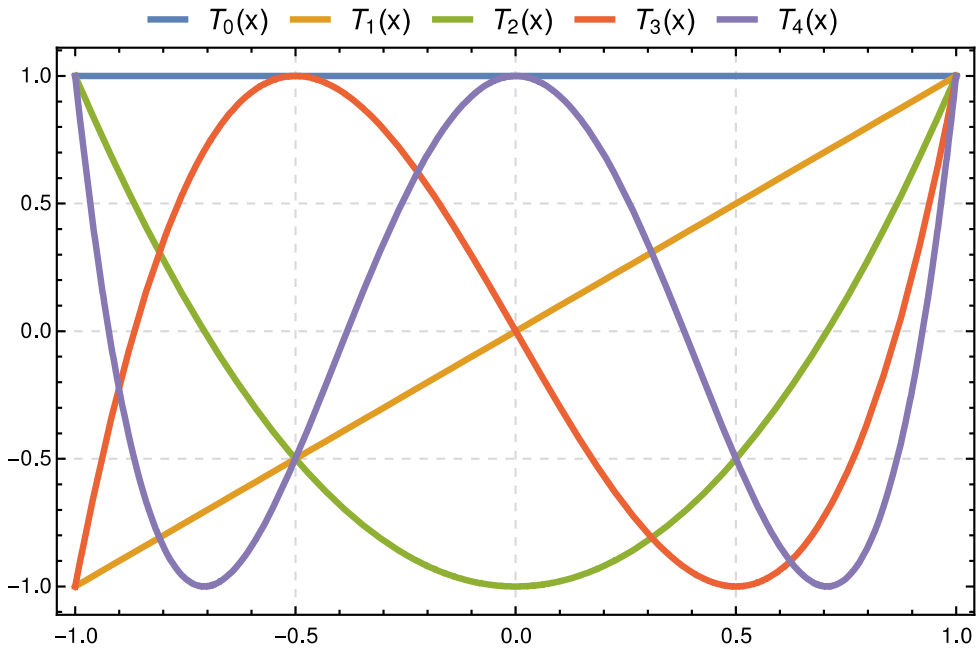
Ελέγχουμε ακόμη εύκολα ότι το T_k έχει:

- k μηδενικά, τα οποία ονομάζονται σημεία Chebyshev και είναι

$$t_k=\cos(\theta_k), \theta_k=\frac{\pi}{2}k+n\frac{\pi}{k}, 0\le n\le k-1$$

- k+1 ακρότατα σημεία όπου το πολυώνυμο παίρνει τις τιμές -1 και 1 και είναι τα

$$\tilde{t}_n=\cos(\eta\pi/k), 0\le n\le k \quad .$$



Εικόνα 5: Τα πέντε πρώτα πολυώνυμα Chebyshev T_k , $k=0,\dots,4$.

Εκφράζοντας το cosine σε εκθετική μορφή έχουμε

$$2T_k(t)=e^{ik\theta}+e^{-ik\theta}=(\cos(\theta)+i\sin(\theta))^k+(\cos(\theta)-i\sin(\theta))^k \quad .$$

Αφού $\theta\in[0,\pi]$, έχουμε $\sin(\theta)=\sqrt{1-t^2}$ και το T_k μπορεί να γραφεί συναρτήσει του t:

$$2T_k(t)=(t+i\sqrt{1-t^2})^n+(t-i\sqrt{1-t^2})^n \quad .$$

Για $\alpha\in\mathbb{R}$ με $|\alpha|>1$, δηλώνουμε με $\bar{P}^\alpha{}_k$ το σύνολο των πολυωνύμων του P_n που παίρνουν την τιμή 1 στο σημείο α:

$$\bar{P}^\alpha{}_k=\{p\in P_k, p(\alpha)=1\} \quad .$$

Πρόταση 3.2.1.2: Επιπλέον τα πολυώνυμα Chebyshev ικανοποιούν την ακόλουθη ιδιότητα:

$$\min_{p\in\bar{P}^{\alpha}_{k}}\ \max_{t\in[-1,1]}|\bar{T}_k(t)|=\frac{1}{|\bar{T}_k(t)|}\quad ,$$

όπου $\bar{T}_k=\frac{T_k}{T_k(\alpha)}$.

Απόδειξη: Ξέρουμε ότι $T_k(\alpha)\neq 0$, επειδή όλα τα μηδενικά, ή οι ρίζες του T_k ανήκουν στο διάστημα $(-1,1)$. Άρα, το $\bar{T}_k$ είναι καλά ορισμένο. Από τη μια πλευρά, αν $\alpha>1$, τότε $T_k(\alpha)>0$, αφού τα πολυώνυμα T_k δεν αλλάζουν πρόσημο στο διάστημα $(1,+\infty)$ και $T_k(1)=1$. Από την άλλη πλευρά, αν $\alpha<-1$, τότε το $T_k(\alpha)$ έχει το ίδιο πρόσημο με το $T_k(-1)=(-1)^k$. Θα μπορούσαμε, ωστόσο, να υποθέσουμε στη συνέχεια χωρίς βλάβη της γενικότητας, ότι υπάρχει $p\in\bar{P}_k$ τ.ω.

$$\max_{[-1,1]}|p(t)|<\frac{1}{T_k(\alpha)}\quad \textbf{(3.7)}.$$

Στη συνέχεια δείχνουμε ότι $p=\bar{T}_k$ με προρρυθμιστή την υπόθεση **(3.7)**. Σε κάθε ένα από τα k+1 ακρότατα σημεία $\tilde{t}_n$ του T_k , το πολυώνυμο $q=p=\bar{T}_k$ αλλάζει πρόσημο, αφού

- $q(\tilde{t}_0)=p(\tilde{t}_0)-\frac{1}{T_k(\alpha)}<0$,
- $q(\tilde{t}_1)=p(\tilde{t}_1)+\frac{1}{T_k(\alpha)}$,
- κτλ.

Επιπλέον, $q(\alpha)=0$, άρα το πολυώνυμο $q\in P_k$ έχει τουλάχιστον κ+1 διακριτά μηδενικά, επειδή τα $\tilde{t}_n$ και $\alpha\notin[-1,1]$. □

Παρατηρήσεις: Είναι δυνατό να βελτιώσουμε ακόμη περισσότερο την ταχύτητα σύγκλισης αν περιορίσουμε το εύρος διασποράς των ιδιοτιμών έτσι ώστε οι περισσότερες εξ αυτών να συγκεντρώνονται σε ένα μόνο διάστημα. Οι υπόλοιπες θα πρέπει να βρίσκονται εκτός του παραπάνω διαστήματος προκειμένου η ταχύτητα σύγκλισης να μην επηρεάζεται σοβαρά. Επιπλέον , όσο πιο κοντά είναι ο δείκτης κατάστασης στην τιμή 1, τόσο μεγαλύτερη είναι και η ταχύτητα σύγκλισης, που σημαίνει ότι ο πίνακας A είναι καλά ορισμένος για μια γρήγορη σύγκλιση της Μεθόδου Συζυγών Κλίσεων. Τέλος, η Μέθοδος Συζυγών Κλίσεων συγκλίνει γρηγορότερα από μια απλή μέθοδο κλίσεων μιας και η σύγκλιση είναι τετραγωνική και όχι γραμμική.

Συμπερασματικά, θεωρώντας τη Μέθοδο Απότομης Καθόδου και τη Μέθοδο Συζυγών Κλίσεων, είναι φανερό ότι η τελευταία υπερτερεί της πρώτης αφού ολοκληρώνεται σε n βήματα, ενώ η πρώτη συγκλίνει μόνο ασυμπτωτικά.

Κεφάλαιο 4

Προρρυθμισμένη Μέθοδος Συζυγών Κλίσεων(Preconditioning Conjugate Gradient)

Ένα πιθανό μειονέκτημα της Μεθόδου Συζυγών Κλίσεων είναι ότι για αρκετά μεγάλο n, φαίνεται να συγκλίνει αργά. Για τη βελτίωση της ταχύτητας σύγκλισης χρησιμοποιείται η μέθοδος της προρρύθμισης. Η ταχύτητα σύγκλισης είναι αύξουσα συνάρτηση του δείκτη κατάστασης $\kappa(A)$ και γι' αυτό προσπαθούμε να τροποποιήσουμε τι αρχικό σύστημα, μετατρέποντάς το σε ένα ισοδύναμο με καλή κατάσταση, δηλαδή με αρκετά μικρότερο δείκτη κατάστασης. Έστω ένας επιπλέον συμμετρικός και θετικά ορισμένος πίνακας $C \in \mathbb{R}^{n,n}$ ο οποίος χρησιμοποιείται ως προρρυθμιστής, τ.ω. αφενός να είναι πιο εύκολα αντιστρέψιμος από τον A και αφετέρου να ισχύει

$$\frac{\lambda_{max}(c^{-1}A)}{\lambda_{min}(c^{-1}A)} \ll \kappa(A)$$

όπου λ_{max} και λ_{min} η μεγαλύτερη και η μικρότερη ιδιοτιμή του πίνακα $C^{-1}A$, αντίστοιχα. Έτσι, το ισοδύναμο σύστημα προς επίλυση στο οποίο καταλήγουμε είναι το εξής:

$$C^{-1}Ax = C^{-1}b \quad .$$

Η ιδέα της προρρύθμισης είναι ότι αν $\kappa(C^{-1}A) \ll \kappa(A)$, η Μέθοδος Συζυγών Κλίσεων με προρρύθμιση συγκλίνει γρηγορότερα από την κλασική Μέθοδο Συζυγών Κλίσεων. Φυσικά, η αιτία αυτής της γρηγορότερης σύγκλισης είναι η απαίτηση αντιστροφής του C . Παρ' όλα αυτά, υπενθυμίζουμε ότι δεν είναι απαραίτητο να δημιουργήσουμε τον πίνακα $C^{-1}A$. Μπορούμε απλώς να πολλαπλασιάσουμε τους πίνακες A και C^{-1} . Το πρόβλημα εδώ είναι να επιλέξουμε έναν πίνακα C ο οποίος να είναι μια εύκολη αντίστροφη προσέγγιση του A.

Παρατηρούμε ότι δεν είναι προφανές εάν μπορούμε να διαλέξουμε έναν πίνακα C με τέτοιο τρόπο ώστε αφενός να ισχύει $\kappa(C^{-1}A) \ll \kappa(A)$ και αφετέρου ο $C^{-1}A$ να παραμένει συμμετρικός και θετικά ορισμένος (απαραίτητη προϋπόθεση για την εφαρμογή της Μεθόδου Συζυγών Κλίσεων). Για κάθε συμμετρικό πίνακα C υπάρχει ένας πίνακας E για τον οποίο ισχύει $C = EE^T$ (Έναν τέτοιον πίνακα μας δίνει και η ανάλυση Cholesky). Οπότε, υποκαθιστούμε το αρχικό σύστημα $Ax = b$, με το ισοδύναμο

$$\tilde{A}\tilde{x} = \tilde{b} \quad ,$$

όπου $\tilde{A} = B^{-1}AB^{-T}, \tilde{b} = B^{-1}b$ και $\tilde{x} = B^T x$.

- <u>Αλγόριθμος της Προρρυθμισμένης Μεθόδου Συζυγών Κλίσεων</u>

Δεδομένα/Είσοδος: $\tilde{A} \in \mathbb{C}^{n,n}, A = A^H$, θετικά ορισμένος, $\tilde{b} \in \mathbb{C}^n$ Έξοδος: $\tilde{x} = \tilde{x}^k$ (κατά προσέγγιση λύση) Αρχικοποίηση:

Διαλέγω $\tilde{x}^0 \in \mathbb{R}^n$
Υπολογίζω: $\tilde{r}^0 = \tilde{p}^0 = \tilde{b} - \tilde{A}\tilde{x}^0$
Επαναλήψεις
For $k \geq 1$ **and** $\tilde{r}^k \neq 0$
$\tilde{\alpha}_k = \frac{\|\tilde{r}^k\|^2}{\langle \tilde{A}\tilde{p}^k, \tilde{p}^k \rangle}$ #υπολογισμός του βήματος προς την κατεύθυνση αναζήτησης
$\tilde{x}^{k+1} \tilde{x}^k + \tilde{\alpha}^k \tilde{p}^k$ #η προσέγγιση της λύσης στη k+1-επανάληψη
$\tilde{r}^{k+1} = \tilde{r}^k - \tilde{\alpha}^k \tilde{A}\tilde{p}^k$ #ανανέωση του διανύσματος υπόλοιπο
$\tilde{\beta}^k = \frac{\|\tilde{r}^{k+1}\|^2}{\|\tilde{r}^k\|^2}$ #υπολογισμός της παραμέτρου β
$\tilde{p}^{k+1} = \tilde{r}^{k+1} + \tilde{\beta}^k \tilde{p}^k$ #ανανέωση του διανύσματος κατεύθυνσης
End

Η επιλογή του κατάλληλου προρρυθμιστή πίνακα C έχει τις ίδιες δυσκολίες που έχει κανείς στην επιλογή προρρυθμιστή στις γενικές επαναληπτικές μεθόδους. Λόγω της φύσης του A, επιλέγονται συμμετρικοί και θετικά ορισμένοι πίνακες. Αυτοί οι προρρυθμιστές εμφανίζονται και στις κλασικές επαναληπτικές μεθόδους, δεδομένου ότι είναι συμμετρικοί και θετικά ορισμένοι πίνακες. Μπορούμε, λοιπόν, να επιλέξουμε τον $C = diag(\alpha_{11,} \alpha_{22,} \alpha_{nn})$ και τότε η μέθοδος που αντιστοιχεί είναι η **Jacobi-Συζυγών Κλίσεων (Jacobi-CG)**.

Η Jacobi-Συζυγών Κλίσεων είναι ο απλούστερος προρρυθμιστής όλων, αφού επιλέγουμε ο C να είναι ένας διαγώνιος πίνακας με στοιχεία τα διαγώνια στοιχεία του πίνακα A. Στην περίπτωση που ο A είναι ερμιτιανός και θετικά ορισμένος πίνακας μπορούμε να παραγοντοποιήσουμε τον διαγώνιο πίνακα του προρρυθμιστή Jacobi σε $C = C^{1/2} C^{1/2}$, με τον πίνακα $C^{1/2}$ να είναι η τετραγωνική ρίζα του πίνακα C , Ο τρόπος αυτός μας επιτρέπει να εφαρμόσουμε τη Μέθοδο Συζυγών Κλίσεων χρησιμοποιώντας αριστερό και δεξιό προρρυθμιστή. Ωστόσο, η εισαγωγή του προρρυθμιστή αυξάνει το κόστος.

Την παραπάνω μέθοδο έρχεται να βελτιώσει η μέθοδος **Block Jacobi-Συζυγών Κλίσεων (Block Jacobi-CG)**. Η βασική ιδέα της μεθόδου αυτής είναι η διάσπαση του πίνακα A σε "μπλοκ" υποπίνακες. Προφανώς η διάσπαση αυτή δεν είναι μοναδική. Η φύση του προβλήματος καθώς και η τεχνική που επιλέγουμε για την επίλυσή του συνήθως μας επιβάλλουν έναν συγκεκριμένο διαχωρισμό. Η μέθοδος Block Jacobi έχει μικρό κόστος και δίνει καλύτερα αποτελέσματα σε σχέση με έναν απλό προρρυθμιστή Jacobi.

Τέλος, μια ακόμη μέθοδος είναι η **SSOR-Συζυγών Κλίσεων (SSOR-CG)**. Ο προρρυθμιστής SSOR είναι ευρέως χρησιμοποιούμενος και προέρχεται από την διάσπαση του A ως εξής:

$$A = C - E = E^T \quad .$$

Για $\omega \in (0,2)$ ο συμμετρικός πίνακας

$$C_\omega = \frac{\omega}{2-\omega}\left(\frac{D}{\omega} - E\right) D^{-1} \left(\frac{D}{\omega} - E^T\right)$$

είναι θετικά ορισμένος. Σημειώνεται ότι η μέθοδος αυτή χρησιμοποιείται συνήθως με $\omega=1$, διότι παρά το γεγονός ότι ο προρρυθμιστής για την βέλτιστη τιμή ω, δηλαδή την ω_{opt} , μπορεί να δώσει πολύ καλά αποτελέσματα, το κόστος όμως για την εύρεση του ω_{opt} είναι απαγορευτικό στη χρήση ενός τέτοιου προρρυθμιστή.

Πράγματι, για $x\neq 0$ έχουμε

$$\langle C_\omega x, x\rangle = \frac{\omega}{2-\omega}\left\langle D^{-1}\left(\frac{D}{\omega}-E\right)^T x, \left(\frac{D}{\omega}-E\right)^T x\right\rangle \quad ,$$

και αφού ο D^{-1} είναι συμμετρικός και θετικά ορισμένος και ο $\frac{D}{\omega}-E$ δεν είναι μοναδικός, έχουμε $\langle C_\omega x, x\rangle > 0$ αν και μόνο αν $\frac{\omega}{2-\omega} > 0$.

ΚΕΦΑΛΑΙΟ 5

Παραδείγματα Επαναληπτικών Μεθόδων

Παράδειγμα 5.1: Έστω ότι ζητείται επίλυση του συστήματος με πίνακα συντελεστών και σταθερό διάνυσμα :

$$A=\begin{pmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix} \quad \text{και} \quad b=\begin{pmatrix} -1 \\ 3 \\ -3 \end{pmatrix}$$

κάνοντας χρήση της Μεθόδου Απότομης Καθόδου με ακρίβεια 3 δεκαδικών ψηφίων και αρχική προσέγγιση $x^0=(0,0,0)^T$.

Γνωρίζουμε ότι η ακριβής λύση του συστήματος είναι $x^*=\left(\frac{1}{5},\frac{4}{5},-\frac{4}{5}\right)^T$.

<u>Λύση</u>

Αρχικά υπολογίζω το διάνυσμα υπόλοιπο r^0 και το αντίστοιχο βήμα α_0 .

$$r^0=b-Ax^0=b=(-1,3,-3)^T$$

$$\alpha_0=\frac{(r^0)^T r^0}{(r^0)^T A r^0}=\frac{\begin{pmatrix} -1 & 3 & -3 \end{pmatrix}\begin{pmatrix} -1 \\ 3 \\ -3 \end{pmatrix}}{\begin{pmatrix} -1 & 3 & -3 \end{pmatrix}\begin{pmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix}\begin{pmatrix} -1 \\ 3 \\ -3 \end{pmatrix}}=\frac{19}{69}=0.275$$

Άρα, η νέα προσέγγιση x^1 θα είναι:

$$x^1=x^0+\alpha_0 r^0=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}+0.275\begin{pmatrix} -1 \\ 3 \\ -3 \end{pmatrix}=\begin{pmatrix} -0.275 \\ 0.825 \\ -0.825 \end{pmatrix}$$

Συνεχίζουμε στην επόμενη επανάληψη ακολουθώντας ακριβώς την ίδια διαδικασία

$$r^1=b-Ax^1=\begin{pmatrix} -1 \\ 3 \\ -3 \end{pmatrix}-\begin{pmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix}\begin{pmatrix} -0.275 \\ 0.825 \\ -0.825 \end{pmatrix}=\begin{pmatrix} -1 \\ 3 \\ -3 \end{pmatrix}-\begin{pmatrix} -2.475 \\ 3.575 \\ -3.575 \end{pmatrix}=\begin{pmatrix} 1.475 \\ 0.575 \\ 0.575 \end{pmatrix}$$

$$\alpha_1=\frac{(r^1)^T r^1}{(r^1)^T A r^1}=\frac{\begin{pmatrix}-2.475 & 3.575 & -3.575\end{pmatrix}\begin{pmatrix}-2.476\\3.575\\-3.575\end{pmatrix}}{\begin{pmatrix}-2.475 & 3.575 & -3.575\end{pmatrix}\begin{pmatrix}3 & -1 & 1\\-1 & 3 & -1\\1 & -1 & 3\end{pmatrix}\begin{pmatrix}-2.475\\3.575\\-3.575\end{pmatrix}}=\frac{0.257}{120.621}=0.002$$

$$x^2=x^1+\alpha_1 r^1=\begin{pmatrix}-0.275\\0.825\\-0.825\end{pmatrix}+0.002\begin{pmatrix}1.475\\0.575\\0.575\end{pmatrix} \quad .$$

Η διαδικασία συνεχίζεται μέχρι να πληρούνται τα κριτήρια σύγκλισης.

Παράδειγμα 5.2: Έστω ότι ζητείται η επίλυση του παρακάτω συστήματος με πίνακα συντελεστών A και σταθερό διάνυσμα b κάνοντας χρήση της Μεθόδου Συζυγών Κλίσεων.

$$A=\begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}, b=\begin{pmatrix}1\\0\end{pmatrix} \quad .$$

Γνωρίζουμε ότι η ακριβής λύση του συστήματος είναι $x^*=\left(\frac{2}{3},\frac{1}{3}\right)^T$.

<u>Λύση</u>

Έστω $x^0=(0,0)^T$. Υπολογίζω το υπόλοιπο $r^0=b-Ax^0=b$. Άρα, το αρχικό διάνυσμα κατεύθυνσης είναι $p^0=r^0=b=(1,0)^T$. Προχωράμε υπολογίζοντας το μέγεθος του βήματος α_0 :

$$\alpha_0=\frac{(r^0)^T r^0}{(p^0)^T A p^0}=\frac{\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}}{\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}}=\frac{1}{2}$$

Έτσι, η προσέγγιση της ρίζας στην πρώτη επανάληψη θα είναι:

$$x^1=x^0+\alpha_0 p^0=\begin{pmatrix}0\\0\end{pmatrix}+\frac{1}{2}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1/2\\0\end{pmatrix}$$

Στη συνέχεια εφαρμόζοντας τα υπόλοιπα βήματα του αλγορίθμου της Μεθόδου Συζυγών Κλίσεων υπολογίζω:

$$r^1=r^0-\alpha_0 A p^0=\begin{pmatrix}1\\0\end{pmatrix}-\frac{1}{2}\begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}0\\1/2\end{pmatrix}$$

$$\beta_0=\frac{(r^1)^T r^1}{(r^0)^T r^0}=\frac{\begin{pmatrix}0 & 1/2\end{pmatrix}\begin{pmatrix}0\\1/2\end{pmatrix}}{\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}}=\frac{1/4}{1}=\frac{1}{4}$$

$$p^1=r^1+\beta_0 p^0=\begin{pmatrix}0\\1/2\end{pmatrix}+\frac{1}{4}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1/4\\1/2\end{pmatrix}$$

$$\alpha_1=\frac{(r^1)^T r^1}{(p^1)^T A p^1}=\frac{\begin{pmatrix}0 & 1/2\end{pmatrix}\begin{pmatrix}0\\1/2\end{pmatrix}}{\begin{pmatrix}1/4 & 1/2\end{pmatrix}\begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}\begin{pmatrix}1/4\\1/2\end{pmatrix}}=\frac{1/4}{3/8}=\frac{2}{3}$$

$$x^2=x^1+\alpha_1 p^1=\begin{pmatrix}1/2\\0\end{pmatrix}+\frac{2}{3}\begin{pmatrix}1/4\\1/2\end{pmatrix}=\begin{pmatrix}2/3\\1/3\end{pmatrix}$$

Επομένως, καταλήξαμε στη ζητούμενη λύση του συστήματος η οποία είναι $x^*=(2/3,1/3)^T$

ΚΕΦΑΛΑΙΟ 6

Υλοποίηση Αριθμητικών Μεθόδων στο Matlab

Συνδυάζοντας όλα όσα αναφέρθηκαν στο παρόν κεφάλαιο, υλοποιούμε τις μεθόδους επίλυσης γραμμικών συστημάτων που αναλύσαμε στα προηγούμενα κεφάλαια στο προγραμματιστικό περιβάλλον MATLAB.

6.1 Μέθοδος Απότομης Καθόδου

Θα εφαρμόσουμε τη Μέθοδο Απότομης Καθόδου για να λύσουμε το σύστημα $Ax=b$, με

```
A =

     4     1     0     0     0     0     0     0     0     0
     1     4     1     0     0     0     0     0     0     0
     0     1     4     1     0     0     0     0     0     0
     0     0     1     4     1     0     0     0     0     0
     0     0     0     1     4     1     0     0     0     0
     0     0     0     0     1     4     1     0     0     0
     0     0     0     0     0     1     4     1     0     0
     0     0     0     0     0     0     1     4     1     0
     0     0     0     0     0     0     0     1     4     1
     0     0     0     0     0     0     0     0     1     4
```

και

```
b =

     1
    -3
     2
     8
    -7
     6
     3
    -2
     1
     4
```

Εφαρμογή:

```
v=[4;4;4;4;4;4;4;4;4;4];
w=[1;1;1;1;1;1;1;1;1];
C=diag(v,0);
B= diag (w,-1);
K= diag (w,1);
A=C+B+K;
x0=[0;0;0;0;0;0;0;0;0;0];
b=[1;-3;2;8;-7;6;3;-2;1;4];
r=b;
rnorm(1)=norm(r,2);
rold=r'*r;
mit=1000;
e=0.00005;
xnew=x0;
k=0;
while (k<mit)
   k=k+1;
 Ak=r'*A;
 ak=rold/(Ak*r);
 xnew=xnew+ak*r;
 betta=ak*A;
 r=r-betta*r;
 rnew=r'*r;
 if norm(r,2)<e
   break;
 end
 rold=rnew;
end
xnew
k
```

```
xnew =

    0.469040
   -0.876158
    0.035581
    2.733834
   -2.970933
    2.149899
    0.371320
   -0.635179
    0.169382
    0.957653

k =  16
```

Οπότε, η υλοποίηση του προγράμματος μας δίνει έξοδο την λύση σε 16 επαναλήψεις.

```
xnew =

   0.469040
  -0.876158
   0.035581
   2.733834
  -2.970933
   2.149899
   0.371320
  -0.635179
   0.169382
   0.957653
```

6.2 Μέθοδος Συζυγών Κλίσεων

Θα εφαρμόσουμε τη Μέθοδο Συζυγών Κλίσεων για να λύσουμε το σύστημα $Ax=b$, με

```
A =

     4     1     0     0     0     0     0     0     0     0
     1     4     1     0     0     0     0     0     0     0
     0     1     4     1     0     0     0     0     0     0
     0     0     1     4     1     0     0     0     0     0
     0     0     0     1     4     1     0     0     0     0
     0     0     0     0     1     4     1     0     0     0
     0     0     0     0     0     1     4     1     0     0
     0     0     0     0     0     0     1     4     1     0
     0     0     0     0     0     0     0     1     4     1
     0     0     0     0     0     0     0     0     1     4
```

και

```
b =

     1
    -3
     2
     8
    -7
     6
     3
    -2
     1
     4
```

Εφαρμογή:

```
v=[4;4;4;4;4;4;4;4;4;4];
w=[1;1;1;1;1;1;1;1;1];
C=diag(v,0);
B= diag (w,-1);
K= diag (w,1);
A=C+B+K;
x0=[0;0;0;0;0;0;0;0;0;0];
b=[1;-3;2;8;-7;6;3;-2;1;4];
r=b;
p=b;
rnorm(1)= norm(r,2);
rold=r'*r;
mit=1000;
e=0.00005;
xnew=x0;
k=0;
while (k<mit)
        k=k+1;
        a= (r'*r)/(p'*A*p);
        y= xnew + a * p;
        xnew=y;
        rold = r;
        r=rold-a*A*p;
        betta=(r'*r)/(rold' * rold);
        if sqrt(r)<e
          break;
        end
        p = r+ betta * p;
end
k
xnew
```

```
xnew =

   0.469039
  -0.876154
   0.035579
   2.733839
  -2.970936
   2.149905
   0.371317
  -0.635174
   0.169380
   0.957655

k =  10
```

Οπότε η υλοποίηση του προγράμματος μας δίνει έξοδο την λύση σε 10 επαναλήψεις.

```
xnew =

    0.469039
   -0.876154
    0.035579
    2.733839
   -2.970936
    2.149905
    0.371317
   -0.635174
    0.169380
    0.957655
```

Συμπεραίνουμε, λοιπόν, ότι η ταχύτητα σύγκλισης της Μεθόδου Συζυγών Κλίσεων είναι ταχύτερη της ταχύτητας σύγκλισης της Μεθόδου Απότομης Καθόδου.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] **Allaire G.**, **Mahmound Kaber S.**, (2008), *Texts in Applied Mathematics 55, Numerical Linear Algebra*, Springer Science+Business Media, LLC.

[2] **Δούγαλης Β.**, **Νούτσος Δ.**, **Χατζηδήμος Α.**, (2016), *Αριθμητική Γραμμική Άλγεβρα*, Πανεπιστήμιο Ιωαννίνων, Ιωάννινα.

[3] **Γεωργίου Γ.**, **Ξενοφώντος Χ.**, (2007), *Σημειώσεις του μαθήματος Μαθηματικά με Υπολογιστές* , Τμήμα Μαθηματικών και Στατιστικής του Πανεπιστημίου Κύπρου, Λευκωσία.

[4] https://en.wikipedia.org/wiki/Conjugate_gradient_method

[5] **Schewchuck J. R.**, (1994), *An introduction to the Conjugate Gradient Method Without the Agonizing Pain*, School of Computer Science, University of Pittsburgh.

[6] **Σπυράκη Μ.**, (2016), *Μέθοδοι Υποχώρων Krylov Για Την Επίλυση Γραμμικών Συστημάτων*, Τμήμα Μαθηματικών του Πανεπιστημίου Ιωαννίνων, Ιωάννινα.

[7] **Pozrikidis C.**, (2006), *Αριθμητικές Υπολογιστικές Μέθοδοι στην Επιστήμη και τη Μηχανική*, Εκδόσεις Τζιόλα, Θεσσαλονίκη

[8] **Ζαφειράκης Γ.**, (2017), *Σύγκριση Αλγορίθμων Αριθμητικής Βελτιστοποίησης και Επίλυσης Γραμμικών και Μη-Γραμμικών Συστημάτων: Χρήση του Matlab*, Πανεπιστήμιο Πατρών, Πάτρα

ΠΑΡΑΡΤΗΜΑ

Το MATLAB είναι ένα σύγχρονο ολοκληρωμένο μαθηματικό λογισμικό πακέτο που χρησιμοποιείται κατά κύριο λόγο για την επίλυση μαθηματικών προβλημάτων, ωστόσο μπορεί να χρησιμοποιηθεί και για προγραμματισμό. Το όνομά του προέρχεται από τα αρχικά γράμματα των λέξεων **MAT**trix **LAB**oratory. Πρόκειται για ένα διαδραστικό (interactive) πρόγραμμα για αριθμητικούς υπολογισμούς και οπτικοποίηση δεδομένων (data visualization) με δυνατότητες προγραμματισμού που το καθιστούν ένα ισχυρό και χρήσιμο εργαλείο στις μαθηματικές και φυσικές επιστήμες.

Όπως υποδηλώνεται και από το όνομα του, το MATLAB είναι ειδικά σχεδιασμένο για υπολογισμούς με πίνακες, όπως η επίλυση γραμμικών συστημάτων, η εύρεση ιδιοτιμών και ιδιοδιανυσμάτων, η αντιστροφή τετραγωνικών πινάκων κλπ. Το MATLAB δε βρίσκει την ακριβή λύση του προβλήματος, αλλά μια προσεγγιστική, δηλαδή είναι σχεδιασμένο φια την αριθμητική επίλυση προβλημάτων σε αριθμητική πεπερασμένης ακρίβειας. Η γλώσσα προγραμματισμού του MATLAB δίνει στον χρήστη την ευχέρεια να το επεκτείνει με δικά του προγράμματα και συχνά θα γράφουμε MATLAB εννοώντας τη γλώσσα προγραμματισμού και όχι το πακέτο MATLAB.

1 Ξεκινώντας με το MATLAB

Αφού εγκαταστήσουμε πρώτα το πρόγραμμα στον υπολογιστή μας, υπάρχουν ορισμένοι κανόνες που πρέπει να ακολουθήσουμε προκειμένου να εξασφαλιστεί η σωστή λειτουργία του προγράμματος. Κάποιοι από αυτούς είναι:

1. Όλες οι ποσότητες θεωρούνται πίνακες και αποθηκεύονται κατά στήλες στη μνήμη του υπολογιστή. Ακόμη και οι αριθμοί θεωρούνται 1×1 πίνακες.
2. Η χρήση του ; στο τέλος μιας εντολής έχει ως αποτέλεσμα να γίνει ο υπολογισμός, αλλά το αποτέλεσμα να κρατηθεί στη μνήμη του υπολογιστή χωρίς να τυπωθεί στην οθόνη.
3. Στα ονόματα των μεταβλητών ή των συναρτήσεων δεν είναι δεκτά τα σύμβολα ή το κενό.
4. Την υποδιαστολή τη δηλώνουμε με το κόμμα (,).
5. Αν θέλουμε να διακόψουμε μια γραμμή και να συνεχίσουμε στην επόμενη βάζουμε στο υέλος της γραμμής 3 τελείες (...).

2 Βασικές πράξεις

Αριθμητικοί τελεστές	Πράξη
+	Πρόσθεση
-	Αφαίρεση
*	Πολλαπλασιασμός
\	Διαίρεση
^	Ύψωση σε δύναμη

Παράδειγμα 1:

```
1  1+4          ans =  5
2  5-2          ans =  3
3  3*6          ans =  18
4  8/4          ans =  2
5  2^3          ans =  8
```

Συγκριτικοί τελεστές	Πράξη
==	Ίσο
<	μικρότερο
>	μεγαλύτερο
<=	μικρότερο ή ίσο
>=	μεγαλύτερο ή ίσο
~=	διάφορο

Προσοχή: Ο τελεστής = χρησιμοποιείται για την εκχώρηση τιμής σε μια μεταβλητή, ενώ ο συγκριτικός τελεστής == χρησιμοποιείται για να ελέγξουμε αν δύο τιμές είναι ίσες,

Παράδειγμα 2:

```
1  2==3         ans = 0
2  2==2         ans = 1
3  2<3          ans = 1
4  2>3          ans = 0
5  2<=3         ans = 1
6  2>=3         ans = 0
7  2~=3         ans = 1
```

Παρατηρώ ότι όταν το αποτέλεσμα της πράξης των συγκριτικών τελεστών είναι αληθές τότε το αποτέλεσμα είναι 1 και όταν είναι ψευδές είναι 0.

Λογικοί τελεστές	Πράξη
&&	Και
||	ή

~	όχι

3 Σταθερές και Μεταβλητές

Όπως όλες οι γλώσσες προγραμματισμού, έτσι και το MATLAB όταν ου ζητάμε να εκτελέσει μια πράξη χρησιμοποιεί **μεταβλητές** και δημιουργεί αυτόματα μια μεταβλητή με το όνομα **ans** (**ans**wer) και αποθηκεύει σε αυτήν το αποτέλεσμα της πράξης. Ωστόσο, στο MATLAB υπάρχουν κάποιες μεταβλητές στις οποίες έχουν καταχωρηθεί σταθερές τιμές. Αυτές είναι οι εξής:

- Το $\pi = 3.14$ το οποίο το καλούμε γράφοντας **pi**.
- Το άπειρο (∞) , το οποίο το καλούμε γράφοντας **Inf**.
- Το **NaN** (Not a Number), το οποίο χρησιμοποιείται για να μας πει ότι το αποτέλεσμα μιας πράξης δεν είναι αριθμός.

Παράδειγμα 3:

```
1  x=1
2  y=3
3  z=(x+y)^2
```

```
x =  1
y =  3
z =  16
```

Για τα ονόματα μεταβλητών χρησιμοποιούνται κυρίως γράμματα του αγγλικού αλφαβήτου. Το MATLAB κάνει διάκριση μεταξύ μικρών και κεφαλαίων γραμμάτων. Για παράδειγμα οι μεταβλητές Y και y δεν είναι ίδιες. Για τα ονόματα των μεταβλητών ισχύουν οι παρακάτω κανόνες:

1. Το όνομα αρχίζει με γράμμα (του αγγλικού αλφαβήτου).
2. Το όνομα περιέχει μόνο γράμματα, αριθμούς και underscore.
3. Δε χρησιμοποιούνται ονόματα που έχουν δεσμευτεί από το MATLAB (π.χ. συναρτήσεις βιβλιοθήκης και εργαλειοθηκών).
4. Προτιμούνται μικρά ονόματα για πρακτικούς λόγους αν και δεν υπάρχει περιορισμός στο μήκος των ονομάτων.

Μπορούμε να γράψουμε περισσότερες από μια εντολές σε μια γραμμή, τις οποίες χωρίζουμε είτε με κόμματα είτε με ερωτηματικά (αν δε θέλουμε να εμφανιστεί στην οθόνη το αποτέλεσμα της πράξης).

Παράδειγμα 4:

```
1  x=3; y=4; sum=x+y,
```

```
sum =  7
```

Παρατηρούμε ότι στο παράθυρο εντολών τυπώθηκε μόνο το sum, αφου μετά τις υπόλοιπες εντολές χρησιμοποιήσαμε ερωτηματικό και όχι κόμμα.

4 Βαθμωτές συναρτήσεις στο MATLAB

Το ΜΑΤΛΑΒ είναι εφοδιασμένο με αρκετές συναρτήσεις βιβλιοθήκης οι οποίες περιλαμβάνουν τους λογαρίθμους, τις εκθετικές συναρτήσεις, τις τριγωνομετρικές, τις υπερβολικές συναρτήσεις, συναρτήσεις στατιστικής ανάλυσης και άλλες. Στον παρακάτω πίνακα δίνουμε τις πιο χρήσιμες από αυτές:

Συνάρτηση	Ερμηνεία
sin(x)	Ημίτονο
cos(x)	συνημίτονο
tan(x)	εφαπτομένη
asin(x)	τόξο ημιτόνου
acos(x)	τόξο συνημιτόνου
atan(x)	τόξο εφαπτομένης
exp(x)	εκθετική συνάρτηση
log(x)	φυσικός λογάριθμος
log10(x)	λογάριθμος με βάση το 10
abs(x)	απόλυτη τιμή
sqrt(x)	τετραγωνική ρίζα
mod(x)	προσημασμένο υπόλοιπο διαίρεσης
round(x)	υπόλοιπο διαίρεσης
ceil(x)	στρογγυλοποίηση στον πλησιέστερο ακέραιο
floor(x)	στρογγυλοποίηση προς το μείον άπειρο
fix(x)	στρογγυλοποίηση προς το μηδέν
rand(x)	δημιουργία τυχαίων αριθμών στο διάστημα (0,1)
real(x)	πραγματικό μέρος μιγαδικού αριθμού
image(x)	φανταστικό μέρος μιγαδικού αριθμού

5 Λίστες στο MATLAB

Οι λίστες είναι ένα από τα βασικότερα στοιχεία του MATLAB και χρησιμοποιούνται αντί για του πίνακες, που έχουμε δει σε άλλες γλώσσες προγραμματισμού, που σημαίνει ότι μας δίνουν τη δυνατότητα να αποθηκεύουμε πολλά δεδομένα στη μορφή μονοδιάστατου ή και πολυδιάστατου πίνακα. Επίσης, μπορούμε να κάνουμε πράξει μεταξύ των πινάκων, όπως ορίζονται στη Γραμμική Άλγεβρα.

Για να δημιουργήσουμε μια λίστα χρησιμοποιούμε τις τετραγωνικές αγκύλες, μέσα στις οποίες βάζουμε τα στοιχεία της.

Παράδειγμα 5:

```
A=[2,3;6,7]
B=[2, 3]
```

```
A =

     2     3
     6     7

B =

     2     3
```

Όπως και στην Γραμμική Άλγεβρα, έτσι και εδώ, όταν θέλουμε να αναφερθούμε στο στοιχείο μιας λίστας αναφερόμαστε στον δείκτη της γραμμής και της στήλης στην οποία βρίσκεται. Για παράδειγμα αν θέλουμε να αναφερθούμε στο στοιχείο της λίστας A που βρίσκεται στην δεύτερη γραμμή και δεύτερη στήλη λέμε A(2,2).

Η ένωση δύο ή περισσότερων λιστών γίνεται ως εξής:

```
A=[2,3;6,7]
B=[2 ,3; 8,9]
C=[A,B]
```

```
A =

     2     3
     6     7

B =

     2     3
     8     9

C =

     2     3     2     3
     6     7     8     9
```

Στοιχειώδεις λίστες

Συνάρτηση	Ερμηνεία
x=a:b:c	Παράγει μονοδιάστατη λίστα από το a εως το c με βημα b
linspace(a,b,c)	παράγει μονοδιάστατη λίστα c στοιχείων με ισοκατανεμημένες τιμές στο διάστημα [a,b]
eye(x)	παράγει λίστα με 1 στην κύρια διαγώνιο και 0 αλλού
zeros(x)	μηδενική λίστα
ones(x)	λίστα με 1 σε όλες τις θέσεις
rand(x)	ομοιόμορφα ψευδο-τυχαία λίστα
diag(x)	λίστα με στοιχεία μόνο στην κύρια διαγώνιο και μήδεν αλλού
randn(x)	κανονικά ψευδο-τυχαία λίστα

pascal(x) hilb(x) invhilb(x)	(τετραγωνική) λίστα Pascal (τετραγωνική) λίστα Hilbert αντίστροφη λίστα Hilbert

6 Βασικές δομές προγραμματισμού

Στον προγραμματισμό υπάρχουν 3 βασικές δομές:

1. Δομή ακολουθίας κατά την οποία οι εντολές ενός προγράμματος εκτελούνται διαδοχικά η μια μετά την άλλη χωρίς καμία εξαίρεση.
2. Δομή επιλογής κατά την οποία ένα πρόγραμμα επιλέγει μόνο του ποιες εντολές θα εκτελέσει. Υλοποιείται ως εξής:

```
1  if (συνθήκη)
2    εντολές 1
3  else
4    εντολές 2
5  end
```

3. Δομή επανάληψης κατά την οποία ένα μέρος του προγράμματος εκτελείται περισσότερες από μια φορές. Η δομή επανάληψης υλοποιείται με την εντολή **for** ή **while**. Η for χρησιμοποιείται όταν γνωρίζουμε πόσες φορές θέλουμε να εκτελεστεί κάποιο κομμάτι του κώδικα, ενώ η while όταν θέλουμε να εκτελείται κάποιο κομμάτι του κώδικα όσο ισχύει μια συνθήκη. Όταν σταματήσει να ισχύει η συνθήκη αυτή, η δομή τερματίζεται.

7 Γραφικές παραστάσεις στο MATLAB

Οι γραφικές παραστάσεις αποτελούν αναμφίβολα έναν τρόπο να κατανοήσουμε καλύτερα το αποτέλεσμα ενός προβλήματος, διότι έχουμε και μια οπτική αναπαράσταση. Το προγραμματιστικό περιβάλλον MATLAB μας δίνει τη δυνατότητα δημιουργίας γραφικών παραστάσεων στο επίπεδο και τον χώρο, με το σύνολο των εντολών που περιέχει.

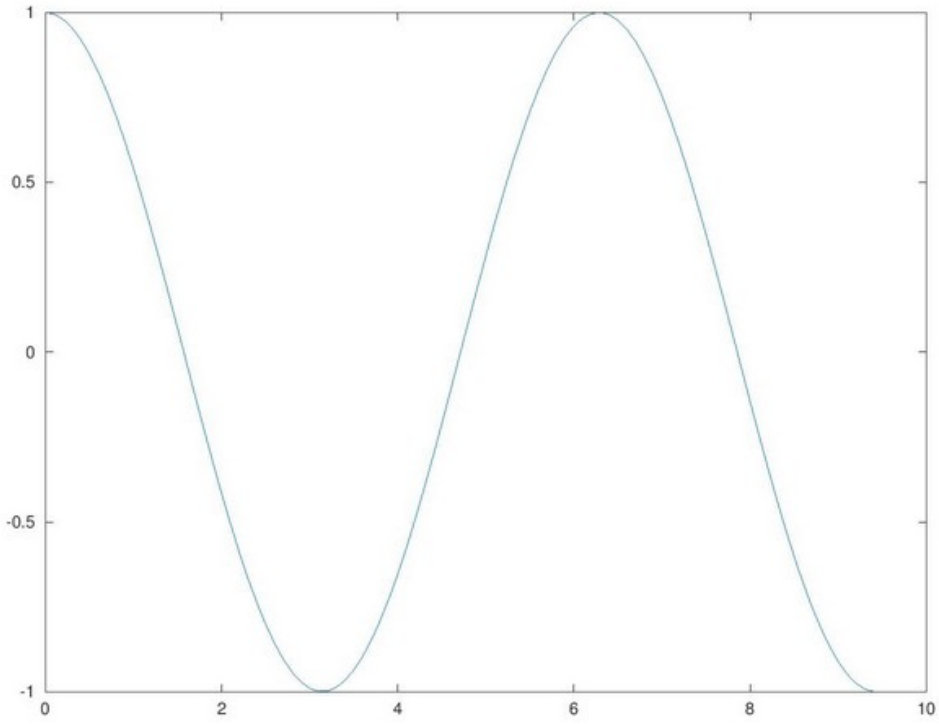
7.1 Γραφικές παραστάσεις στο επίπεδο

Η βασική εντολή δημιουργίας γραφικής παράστασης στο επίπεδο είναι:

plot(x,y),

όπου **x,y** είναι μονοδιάστατες λίστες. Για παράδειγμα, αν θέλουμε να φτιάξουμε τη γραφική παράσταση της συνάρτησης $f(x)=\cos(x)$ δημιουργούμε αρχικά μια λίστα **x** που εκφάζει τον άξονα xx' και περιέχει κοντινές τιμές σε ένα διάστημα, έστω το $(0,3\pi)$. Έπειτα, δημιουργούμε μια λίστα $y=\cos(x)$ και χρησιμοποιούμε την παραπάνω εντολή. Οπότε έχουμε:

```
1  x=0:0.1:3*pi;
2  y=cos(x);
3  plot(x,y)
```



Ωστόσο, η μορφή της γραφικής παράστασης δεν είναι μόνο η προεπιλεγμένη. Αντίθετα, το MATLAB μας δίνει τη δυνατότητα να προσαρμόσουμε τη γραφική παράσταση σύμφωνα με την δική μας προτίμηση. Παρακάτω αναφέρονται ορισμένες εντολές που διαμορφώνουν την γραφική παράσταση.

- Καθορισμός του εύρους αξονων

axis([xmin, xmax, ymin, ymax])

Οι δύο πρώτες μεταβλητές καθορίζουν τον άξονα xx' και οι δύο τελευταίες τον άξονα yy'.

- Μπορούμε να δώσουμε τίτλους στον άξονες xx' και yy' με τις εντολές **xlabel('titlos')** και **ylabel('titlos')** αντίστοιχα.

- Ο τίτλος του γραφήματος δίνεται με την εντολή **title('titlos')**.

- Η εισαγωγή πλέγματος γίνεται με την εντολή **grid on**.

- Η διαμόρφωση της καμπύλης γίνεται με την εντολή **plot(x,y,'ABC')** όπου:

  A: επιλέγει το χρώμα της καμπύλης

  B: επιλέγει το σύμβολο που θα τοποθετηθεί για κάθε σημείο

  C: επιλέγει το είδος της γραμμής

  Οι τιμές των τριών αυτών παραμέτρων δίνονται με βάση τον παρακάτω πίνακα.

A	Χρώμα	B	Σύμβολο	C	Είδος Γραμμής
B	blue	.	point	-	solid
g	green	o	circle	:	dotted
r	red	x	cross	-.	Dach-dot
c	cyan	+	Plus sign	--	dashed
m	magenta	*	asterisk	[κενό]	Χωρίς γραμμη
y	yellow	s	square		
k	black	d	diamond		
w	white	v	triangle(down)		
		^	triangle(up)		
		<	triangle(left)		
		>	triangle(right)		
		p	pentagon		
		h	hexagon		