



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΣΤΕΡΕΑΣ ΕΛΛΑΔΑΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΪΑΤΡΙΚΗ**

Θεωρία Πινάκων και Εφαρμογές

Πτυχιακή εργασία

Ιωάννη Γ. Καρακίτσιου

**Υπεύθυνος
Κόκκινος Χριστόδουλος
Αναπληρωτής Καθηγητής Ε.Μ.Π**

Λαμία, 2008

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΣΤΕΡΕΑΣ ΕΛΛΑΔΑΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΪΑΤΡΙΚΗ**

Θεωρία Πινάκων και Εφαρμογές

Πτυχιακή εργασία

Ιωάννη Γ. Καρακίτσιου

**Υπεύθυνος
Κόκκινος Χριστόδουλος
Αναπληρωτής Καθηγητής Ε.Μ.Π**

Λαμία, 2008

ΕΥΧΑΡΙΣΤΙΕΣ

Έχοντας ολοκληρώσει την παρούσα πτυχιακή εργασία, θα ήθελα να ευχαριστήσω:

Τον επιβλέποντα καθηγητή μου, κ. Κόκκινο Χριστόδουλο, Αναπληρωτή Καθηγητή Ε.Μ.Π. για την υπόδειξη του θέματος, τη συνεχή παρακολούθηση και ενθάρρυνση καθόλη τη διάρκεια περάτωσης της εργασίας.

Την κα. Αδάμ Μαρία, Δρ. του Ε.Μ.Π., για την πολύτιμη βοήθεια, την υποδομή που μου έδωσε από τα αρχικά εξάμηνα και τις κρίσιμες υποδείξεις καθόλη τη διάρκεια της συγγραφής αυτής.

Τα μέλη της Τριμελούς Εξεταστικής Επιτροπής, την Αναπληρώτρια Καθηγήτρια κα. Καραγκούνη-Κύρτσου Αμαλία και τον Επίκουρο Καθηγητή κ. Πλαγιανάκο Βασίλειο, οι οποίοι δέχτηκαν να συμμετέχουν στην Εξεταστική Επιτροπή. Έγινε κάθε δυνατή προσπάθεια, ώστε οι παρατηρήσεις τους να ενσωματωθούν στο τελικό κείμενο.

Το Τμήμα Πληροφορικής με Εφαρμογές στη Βιοϊατρική του Πανεπιστημίου Στερεάς Ελλάδας για το θετικό ακαδημαϊκό κλίμα και ιδιαίτερα τους διδάσκοντες και συμφοιτητές μου για την εν γένει συνεργασία τους.

Το Ίδρυμα Κρατικών Υποτροφιών για την οικονομική υποστήριξη που μου παρείχε κατά το πρώτο έτος των σπουδών μου καθώς ήμουν υπότροφός του.

Τέλος, οφείλω ευχαριστίες στην οικογένειά μου για την καθολική στήριξή τους σε όλα τα χρόνια των σπουδών μου.

Καρακίτσος Γ. Ιωάννης

Σεπτέμβριος 2008

ABSTRACT

Matrix Theory is a sector of Linear Algebra that constitutes an essential part of the mathematical background of many scientific branches with applications in physical, technological and economical sciences and its evolution was astonishing in Mathematical Analysis, Probability Theory and Statistics. This project does not contain any unconventional conclusion, while that presupposes postgraduate-level knowledge. The contribution of this project in the preexistent bibliography is the clarification of many deficient acceptations. Finally, through the studying of some chapters of Matrix Theory, we approach acceptable solutions through more general spaces, such as Vector Spaces, Norm Spaces and Metric Spaces.

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΙΣΑΓΩΓΗ	1
ΚΕΦΑΛΑΙΟ 1 – ΠΡΟΚΑΤΑΡΚΤΙΚΕΣ ΕΝΝΟΙΕΣ	
1.1. Βασικοί ορισμοί πινάκων.....	3
1.2. Βασικές πράξεις και ιδιότητες.....	5
1.2.1. Πρόσθεση πινάκων.....	5
1.2.2. Γινόμενο πίνακα επί έναν αριθμό.....	5
1.2.3. Γινόμενο διανυσμάτων.....	6
1.2.4. Πολλαπλασιασμός πινάκων.....	6
1.2.5. Ιδιότητες πράξεων.....	6
1.3. Διανυσματικοί χώροι.....	7
1.4. Γραμμικοί μετασχηματισμοί.....	8
1.5. Γραμμική ανεξαρτησία διανυσμάτων τύπου $n \times 1$	8
1.6. Εσωτερικό γινόμενο.....	9
1.7. Ορίζουσες.....	10
1.7.1. Βασικοί ορισμοί.....	10
1.7.2. Ιδιότητες οριζουσών.....	11
1.8. Ειδικές κατηγορίες πινάκων.....	12
1.8.1. Συμπληρωματικός πίνακας.....	12
1.8.2. Αντίστροφος πίνακας.....	13
1.8.3. Ορθομοναδιαίος πίνακας.....	13
1.8.4. Ρητές συναρτήσεις πινάκων.....	14
ΚΕΦΑΛΑΙΟ 2 – ΓΡΑΜΜΙΚΑ ΣΥΣΤΗΜΑΤΑ	
2.1. Εισαγωγικές έννοιες.....	17
2.2. Βαθμός πίνακα.....	18
2.3. Στοιχειώδεις μετασχηματισμοί.....	20
2.4. Επίλυση γραμμικών συστημάτων.....	25

ΚΕΦΑΛΑΙΟ 3 – ΙΔΙΟΤΙΜΕΣ ΚΑΙ ΙΔΙΟΔΙΑΝΥΣΜΑΤΑ

3.1. Εισαγωγικοί ορισμοί.....	31
3.2. Υπολογισμός ιδιοτιμών και ιδιοδιανυσμάτων.....	31
3.3. Βασικές ιδιότητες χαρακτηριστικών μεγεθών.....	35
3.4. Χαρακτηριστικά μεγέθη πινάκων ειδικής μορφής.....	36
3.5. Χαρακτηριστικά μεγέθη ρητών συναρτήσεων.....	38
3.6. Αναγωγή πίνακα σε απλούστερη μορφή.....	40
3.7. Χρήση της διαγωνοποίησης στη λύση προβλημάτων.....	42
3.8. Θεώρημα Cayley-Hamilton.....	47

ΚΕΦΑΛΑΙΟ 4 – ΝΟΡΜΕΣ ΚΑΙ ΘΕΩΡΙΑ ΔΙΑΤΑΡΑΧΩΝ

4.1. Χώροι νορμέ.....	53
4.2. Συνεχείς γραμμικές απεικονίσεις.....	55
4.3. Απόλυτος τιμή και Νόρμα πίνακα.....	56
4.3.1. Απόλυτος τιμή.....	56
4.3.2. Νόρμα πίνακα $m \times n$	57
4.3.3. Νόρμα τετραγωνικού πίνακα.....	58
4.4. Φράγμα της φασματικής ακτίνας.....	64
4.5. Θεωρία διαταραχών (perturbations) για γραμμικές εξισώσεις.....	68
4.6. Θεωρία διαταραχών για το γραμμικό σύστημα.....	68
4.6.1. Διαταραχή του δευτέρου μέλους.....	68
4.6.2. Διαταραχή του πίνακα.....	71
4.6.3. Διαταραχή του πίνακα και του δευτέρου μέλους.....	71
4.7. Υπολογίσιμα φράγματα σφάλματος για γραμμικές εξισώσεις.....	72

ΚΕΦΑΛΑΙΟ 5 – ΜΕΤΡΙΚΟΙ ΧΩΡΟΙ - ΑΚΟΛΟΥΘΙΕΣ ΠΙΝΑΚΩΝ

5.1. Μετρικοί χώροι.....	75
5.2. Ακολουθίες πινάκων.....	78

ΚΕΦΑΛΑΙΟ 6 –ΣΥΣΤΟΛΙΚΗ ΑΠΕΙΚΟΝΙΣΗ ΚΑΙ ΕΠΑΝΑΛΗΠΤΙΚΕΣ ΜΕΘΟΔΟΙ

6.1. Συστολική απεικόνιση.....	83
6.2. Εφαρμογή του θεωρήματος Banach στις γραμμικές εξισώσεις.....	85
6.3. Επαναληπτικές μέθοδοι.....	87
6.3.1. Μέθοδος Jacobi (E.M.J.).....	87
6.3.2. Μέθοδος Gauss-Seidel (E.M.G-S.).....	91

ΚΕΦΑΛΑΙΟ 7 – ΣΥΝΑΡΤΗΣΕΙΣ ΠΙΝΑΚΩΝ

7.1. Σειρά πινάκων.....	95
7.2. Πράξεις με σειρές πινάκων.....	97
7.3. Επέκταση των βαθμωτών συναρτήσεων σε συναρτήσεις πινάκων.....	101
7.4. Επέκταση των βαθμωτών ταυτοτήτων.....	105
7.5. Υπολογισμός συναρτήσεων πινάκων.....	106
7.6. Η συνάρτηση e^{At}	110

ΚΕΦΑΛΑΙΟ 8 – ΔΙΑΦΟΡΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

8.1. Συναρτήσεις πινάκων μίας μεταβλητής.....	113
8.2. Γραμμικές διαφορικές εξισώσεις.....	116
8.3. Θεμελιώδεις λύσεις.....	118
8.4. Μη ομογενές σύστημα διαφορικών εξισώσεων.....	124

ΣΥΜΠΕΡΑΣΜΑ	127
------------------	-----

ΠΑΡΑΡΤΗΜΑ

Χρήση της MATLAB για τη λύση προβλημάτων με την E.M.J.	129
---	-----

ΒΙΒΛΙΟΓΡΑΦΙΑ	131
--------------------	-----

ΠΕΡΙΛΗΨΗ ΣΤΑ ΑΓΓΛΙΚΑ (ABSTRACT).....	133
--------------------------------------	-----

ΕΙΣΑΓΩΓΗ

Η Θεωρία Πινάκων είναι ένας κλάδος της Γραμμικής Άλγεβρας, ο οποίος αποτελεί ουσιώδες μέρος του μαθηματικού υπόβαθρου πολλών επιστημονικών κλάδων με εφαρμογές στις φυσικές, τεχνολογικές και οικονομικές επιστήμες και η εξέλιξή της υπήρξε ραγδαία στη Μαθηματική Ανάλυση, στη Θεωρία Πιθανοτήτων και Στατιστικής κ.α. Η εργασία μου αυτή δεν περιέχει ουσιαστικά πρωτότυπα αποτελέσματα από μαθηματική άποψη, αφού αυτό προϋποθέτει και γνώσεις μεταπτυχιακού επιπέδου. Η συμβολή της παρούσης πτυχιακής εργασίας στην προϋπάρχουσα βιβλιογραφία είναι η αποσαφήνιση και συμπλήρωση πολλών εννοιών που ήταν ελλειμματικές. Στην εργασία αυτή, παράλληλα με τη μελέτη ορισμένων κεφαλαίων της Θεωρίας Πινάκων, προσεγγίζουμε διάφορα γνωστά αποτελέσματα με τη βοήθεια γενικότερων χώρων, όπως είναι οι Διανυσματικοί χώροι, οι Χώροι Νορμέ και οι Μετρικοί Χώροι.

Η μελέτη της Θεωρίας Πινάκων στην παρούσα πτυχιακή εργασία κατανέμεται σε οκτώ κεφάλαια. Στο πρώτο κεφάλαιο, για την αυτοδύναμη ανάγνωσή της, δίνουμε κάποιες εισαγωγικές έννοιες για τους πίνακες. Πιο συγκεκριμένα, αναφέρουμε τους βασικούς ορισμούς, τις πράξεις μεταξύ πινάκων και τις ιδιότητές τους, τον ορισμό και τις ιδιότητες της ορίζουσας ενός τετραγωνικού πίνακα, καθώς και τους απαραίτητους ορισμούς και ιδιότητες των πινάκων που ανήκουν σε ειδικές κατηγορίες. Παράλληλα αναφερόμαστε στους διανυσματικούς χώρους και σε ορισμένες έννοιες των χώρων αυτών, όπως, γραμμικοί μετασχηματισμοί, γραμμική ανεξαρτησία και εσωτερικό γινόμενο διανυσμάτων.

Στο δεύτερο κεφάλαιο της διπλωματικής αυτής εισάγουμε την έννοια του βαθμού ενός πίνακα και παρουσιάζουμε τον αναλυτικό τρόπο επίλυσης των γραμμικών συστημάτων.

Το τρίτο κεφάλαιο επικεντρώνεται στη μελέτη των χαρακτηριστικών μεγεθών των τετραγωνικών πινάκων, δηλαδή των ιδιοτιμών και ιδιοδιανυσμάτων τους. Αφού παρουσιάσουμε τις βασικές ιδιότητες των χαρακτηριστικών μεγεθών για έναν τυχαίο

πίνακα, στη συνέχεια, δίνουμε τις ιδιότητες των ιδιοτιμών και ιδιοδιανυσμάτων πινάκων ειδικής μορφής και των ρητών συναρτήσεων. Έπειτα, αναλύουμε τη μέθοδο της αναγωγής πίνακα σε απλούστερη μορφή, καθώς και τη χρήση της διαγωνοποίησης στην επίλυση προβλημάτων. Τέλος, περιγράφουμε το θεώρημα Cayley-Hamilton και τη σημασία του.

Στο επόμενο κεφάλαιο παρουσιάζουμε, αρχικά, τους χώρους νορμέ, στη συνέχεια ορίζουμε την απόλυτη τιμή πίνακα και δίνουμε την έννοια καθώς και τις ιδιότητες της νόρμας ενός πίνακα. Η μελέτη μας επικεντρώνεται στις εφαρμογές που βρίσκει η παραπάνω θεωρία τόσο στη θεωρία διαταραχών στις γραμμικές εξισώσεις και στα γραμμικά συστήματα όσο και στον υπολογισμό φραγμάτων σφάλματος των γραμμικών εξισώσεων.

Στο πέμπτο εισάγουμε την έννοια της μετρικής στο σύνολο των πινάκων του ιδίου τύπου, με συνέπεια όλα τα γνωστά αποτελέσματα της σύγκλισης των ακολουθιών των πραγματικών αριθμών να μεταφέρονται στις ακολουθίες πινάκων.

Στο έκτο κεφάλαιο αποδεικνύουμε το βασικό θεώρημα Banach (Θεώρημα Σταθερού Σημείου) για τις γραμμικές εξισώσεις και παρουσιάζουμε την εφαρμογή του στις επαναληπτικές μεθόδους Jacobi και Gauss-Seidel για την αριθμητική επίλυση των γραμμικών συστημάτων. Επιλύοντας τη μέθοδο Jacobi μέσω Matlab, ο κώδικας της οποίας βρίσκεται στο Παράρτημα του βιβλίου, για προκαθορισμένη ακρίβεια, διαπιστώνουμε τη διαφορά μεταξύ θεωρητικών και πρακτικών αποτελεσμάτων.

Στόχος του εβδομού κεφαλαίου είναι ο υπολογισμός συναρτήσεων πινάκων. Πιο συγκεκριμένα, δίνουμε τον ορισμό της σειράς πινάκων και επεκτείνουμε τις γνωστές ιδιότητες και φράγματα των βαθμωτών συναρτήσεων και ταυτοτήτων σε συναρτήσεις πινάκων. Γίνεται ειδική αναφορά στη συνάρτηση e^{At} .

Στο όγδοο κεφάλαιο, παρουσιάζουμε τρόπους επίλυσης γραμμικών διαφορικών εξισώσεων χρησιμοποιώντας τη θεωρία πινάκων.

Στο τέλος της πτυχιακής εργασίας υπάρχουν τα συμπεράσματα και ένας κατάλογος συγγραμμάτων, που πολλά από αυτά αναφέρονται στο κείμενο και είναι χρήσιμα για περαιτέρω μελέτη και εμβάθυνση των προαναφερόμενων εννοιών.

ΚΕΦΑΛΑΙΟ 1

ΠΡΟΚΑΤΑΡΚΤΙΚΕΣ ΕΝΝΟΙΕΣ

1.1. Βασικοί ορισμοί πινάκων.

Ονομάζουμε **πίνακα** τύπου $m \times n$ ένα σύνολο από πραγματικούς ή μιγαδικούς αριθμούς, διατεταγμένους σε ένα ορθογώνιο σχήμα της μορφής:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix} \quad (1.1)$$

όπου a_{ij} ($i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$) παριστάνει το στοιχείο της i γραμμής και της j στήλης.

Το **διάνυσμα στήλης** $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ είναι ένας πίνακας τύπου $n \times 1$.

Το **διάνυσμα γραμμής** $\mathbf{u} = [u_1 \ u_2 \ \dots \ u_n]$ είναι ένας πίνακας τύπου $1 \times n$. Τον πίνακα (1.1) θα τον παριστάνουμε συντόμως με το σύμβολο

$$A = [a_{ij}].$$

Ομοτάξια στοιχεία δυο πινάκων A και B του ίδιου τύπου $m \times n$ λέγονται εκείνα τα στοιχεία που έχουν τον ίδιο δείκτη γραμμής i και τον ίδιο δείκτη στήλης j .

Δύο πίνακες A και B του ίδιου τύπου είναι **ίσοι** μεταξύ τους και γράφουμε $A = B$ αν και μόνο αν τα ομοτάξια στοιχεία τους είναι ίσα. Δηλαδή,

$$A = B \Leftrightarrow a_{ij} = b_{ij}.$$

Αν $a_{ij} = -b_{ij}$, τότε ο πίνακας A λέγεται **αντίθετος** του B και γράφουμε $A = -B$.

Ένας πίνακας τύπου $n \times n$:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{bmatrix} \quad (1.2)$$

Ένας συμμετρικός ή αντισυμμετρικός πίνακας είναι πάντοτε τετραγωνικός.

Συζυγής του πίνακα $A = [a_{ij}]$ είναι ο πίνακας

$$\bar{A} = [\bar{a}_{ij}]$$

του ίδιου τύπου με τον A , με στοιχεία τα ομοτάξια συζυγή των στοιχείων του A .

Από τα παραπάνω προκύπτει άμεσα ότι από τις δυο συνθήκες

$$A = \bar{A}, \quad A = -\bar{A},$$

η πρώτη είναι αναγκαία και ικανή για να είναι ο πίνακας πραγματικός και η δεύτερη για να είναι καθαρά φανταστικός.

Αναστροφosuζυγής του πίνακα $A = [a_{ij}]$ τύπου $m \times n$ είναι ο πίνακας

$$A^* = \bar{A}^T = \overline{A^T} = [\bar{a}_{ji}]$$

τύπου $n \times m$, που είναι ο ανάστροφος του \bar{A} .

Ένας πίνακας A λέγεται **ερμιτιανός** όταν ικανοποιεί τη σχέση

$$A^* = A,$$

λέγεται δε **αντιερμιτιανός** όταν ικανοποιεί τη σχέση

$$A^* = -A.$$

Ένας ερμιτιανός ή αντιερμιτιανός πίνακας είναι πάντοτε τετραγωνικός. Κάθε πραγματικός ερμιτιανός πίνακας είναι συμμετρικός και κάθε πραγματικός αντιερμιτιανός πίνακας είναι αντισυμμετρικός.

1.2. Βασικές πράξεις και ιδιότητες.

1.2.1. Πρόσθεση πινάκων.

Αν $A = [a_{ij}]$ και $B = [b_{ij}]$ είναι δύο πίνακες τύπου $m \times n$, τότε το άθροισμα $C = A + B$ είναι ένας πίνακας του ίδιου τύπου του οποίου το στοιχείο c_{ij} είναι το άθροισμα των αντίστοιχων ομοτάξιων στοιχείων των πινάκων A και B . Άρα

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}].$$

1.2.2. Γινόμενο πίνακα επί έναν αριθμό.

Το **γινόμενο** ενός πίνακα A επί έναν **οποιοδήποτε αριθμό** λ ορίζεται από τον τύπο

$$\lambda A = [\lambda a_{ij}].$$

Ορίζουμε τη **διαφορά** $A - B$ ίση με τον πίνακα $A + (-B)$.

1.2.3. Γινόμενο διανυσμάτων.

$$\text{Αν } \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \text{ και } \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \text{ τότε ορίζουμε το γινόμενο του διανύσματος γραμμής } \mathbf{u}^T$$

επί το διάνυσμα στήλης \mathbf{v} από τον τύπο:

$$\mathbf{u}^T \mathbf{v} = [u_1 \quad u_2 \quad \cdots \quad u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{k=1}^n u_k v_k .$$

1.2.4. Πολλαπλασιασμός πινάκων.

Αν $A = [a_{ij}]$ είναι ένας πίνακας τύπου $m \times l$ και $B = [b_{ij}]$ είναι ένας πίνακας τύπου $l \times n$, τότε το γινόμενο $C = AB$ είναι ένας πίνακας τύπου $m \times n$, του οποίου το στοιχείο c_{ij} είναι το γινόμενο των στοιχείων της i γραμμής του πίνακα A επί των στοιχείων της j γραμμής του πίνακα B . Δηλαδή

$$AB = C = [c_{ij}] ,$$

$$\text{όπου } c_{ij} = [a_{i1} \quad a_{i2} \quad \cdots \quad a_{il}] \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{lj} \end{bmatrix} = \sum_{k=1}^l a_{ik} b_{kj} .$$

1.2.5. Ιδιότητες πράξεων.

Έστω οι πίνακες A, B, C , έτσι ώστε οι ακόλουθες σημειούμενες πράξεις να έχουν έννοια, τότε:

- □ $A + B = B + A$.
- $A + (B + C) = (A + B) + C$.
- $\lambda(A + B) = \lambda A + \lambda B = (A + B)\lambda$.
- υπάρχει πίνακας D , τέτοιος ώστε $A + D = B$.
- $(A + B)^* = A^* + B^*$.
- $tr(A + B) = trA + trB$.
- □ $A(B + C) = AB + AC$.
- $(A + B)C = AC + BC$.
- $(AB)^* = B^* A^*$.

- $tr(AB) = tr(BA)$.
- Όμως :
 - $AB \neq BA$.
 - $AB = \mathbf{O}$ δεν έπεται αναγκαστικά $A = \mathbf{O}$ ή $B = \mathbf{O}$.
 - $AB = CB$ δεν έπεται αναγκαστικά $A = C$.
- Αν οι πίνακες A, B είναι τετραγωνικοί του ίδιου τύπου, αντιμεταθετικοί και ο πρώτος από αυτούς διαγώνιος με τα διαγώνια στοιχεία διάφορα μεταξύ τους τότε και ο δεύτερος είναι διαγώνιος.

1.3. Διανυσματικοί χώροι.

Ένα μη κενό σύνολο V με στοιχεία \mathbf{x}, \mathbf{y} λέγεται **διανυσματικός χώρος** (ή **γραμμικός χώρος**) και τα στοιχεία αυτού λέγονται **διανύσματα**, αν στο V είναι ορισμένες δύο αλγεβρικές πράξεις, η πρόσθεση διανυσμάτων και ο βαθμωτός πολλαπλασιασμός, ως εξής:

(A₁) Σε κάθε ζεύγος διανυσμάτων \mathbf{x} και \mathbf{y} του V αντιστοιχεί ένα μοναδικό διάνυσμα του V , που λέγεται **άθροισμα** των \mathbf{x} και \mathbf{y} , και συμβολίζεται με $\mathbf{x} + \mathbf{y}$ και ικανοποιούνται οι ιδιότητες:

(A₂) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ (η πρόσθεση είναι αντιμεταθετική),

(A₃) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$ (η πρόσθεση είναι προσεταιριστική),

(A₄) Υπάρχει ένα μηδενικό στοιχείο $\mathbf{0}$, τέτοιο ώστε : $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$, για κάθε \mathbf{x} .

(A₅) Για κάθε \mathbf{x} , υπάρχει ένα αρνητικό διάνυσμα $-\mathbf{x}$, τέτοιο ώστε : $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

Επιπλέον, υπάρχει ένα σώμα αριθμών K , του οποίου τα στοιχεία συμβολίζονται με a, b, \dots που συνήθως είναι, είτε το σώμα των πραγματικών αριθμών \mathbb{R} , είτε το σώμα των μιγαδικών αριθμών \mathbb{C} , τα στοιχεία του σώματος K χρησιμοποιούνται στο βαθμωτό πολλαπλασιασμό για να ικανοποιούνται τα παρακάτω αξιώματα:

(S₁) Σε κάθε αριθμό $a \in K$ και κάθε διάνυσμα $\mathbf{x} \in V$ αντιστοιχεί ένα μοναδικό βαθμωτό γινόμενο $a\mathbf{x} \in V$.

(S₂) $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$.

(S₃) $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$.

(S₄) $a(b\mathbf{x}) = (ab)\mathbf{x}$.

(S₅) $1 \cdot \mathbf{x} = \mathbf{x}$.

Εάν είναι $K = \mathbb{R}$, ο διανυσματικός χώρος λέγεται **πραγματικός**, ενώ αν είναι $K = \mathbb{C}$ λέγεται **μιγαδικός**.

Εύκολα, αποδεικνύεται ότι για $\lambda \in K$, ισχύουν οι ιδιότητες:

- (i) $0 \cdot \mathbf{x} = \mathbf{0}$.
- (ii) $\lambda \cdot \mathbf{0} = \mathbf{0}$.
- (iii) $-(\lambda \mathbf{x}) = (-\lambda) \mathbf{x} = \lambda(-\mathbf{x})$.

Για μία πλήρη ανάπτυξη των διανυσματικών χώρων, βλέπε οποιοδήποτε βιβλίο Γραμμικής Άλγεβρας. (Για παράδειγμα Δ. Δασκαλόπουλος, [9], Κεφ. 1, Γ. Δονάτος-Μ. Αδάμ Κεφ.4, Ι. Μαρουλά Κεφ.4, Ν. Καδιανάκη- Σ. Καρανάσκιος Κεφ. 6).

Παράδειγμα 1.1 Το σύνολο $V = \mathbb{R}^n$ ($V = \mathbb{C}^n$) όλων των πραγματικών (μιγαδικών) $n \times 1$ διανυσμάτων στήλης, με τις πράξεις της πρόσθεσης και του βαθμωτού πολλαπλασιασμού όπως ορίστηκαν στις πράξεις πινάκων, αποτελεί ένα διανυσματικό χώρο.

Παράδειγμα 1.2 Το σύνολο V όλων των πολυωνύμων, βαθμού $\leq n$, αποτελεί ένα διανυσματικό χώρο.

1.4. Γραμμικοί μετασχηματισμοί.

Έστω V και U είναι δύο διανυσματικοί χώροι ορισμένοι πάνω στο ίδιο σώμα K . Μία απεικόνιση $T: V \rightarrow U$ λέγεται **γραμμική**, αν για κάθε $\mathbf{x}_1, \mathbf{x}_2 \in V$ και για κάθε $a, b \in K$ ισχύει:

$$T(a\mathbf{x}_1 + b\mathbf{x}_2) = aT\mathbf{x}_1 + bT\mathbf{x}_2.$$

Μία γραμμική απεικόνιση $T: V \rightarrow V$ λέγεται **γραμμικός μετασχηματισμός**.

Παράδειγμα 1.3 Ένας πραγματικός τετραγωνικός πίνακας $A = [a_{ij}]$ ορίζει ένα γραμμικό μετασχηματισμό $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$, με

$$\mathbf{y} = A\mathbf{x}.$$

1.5. Γραμμική ανεξαρτησία διανυσμάτων τύπου $n \times 1$.

Ένα σύστημα k διανυσμάτων στηλών τύπου $n \times 1$, $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$, είναι **γραμμικά ανεξάρτητα**, αν από τη σχέση

$$\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \dots + \lambda_k \mathbf{u}_k = \mathbf{0} \tag{1.3}$$

συνεπάγεται ότι $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$. Διαφορετικά, αν υπάρχουν k συντελεστές $\lambda_1, \lambda_2, \dots, \lambda_k$, όχι όλοι μηδέν, τέτοιοι ώστε να ισχύει η (1.3) τότε τα διανύσματα είναι **γραμμικά εξαρτημένα**.

Σημείωση : (i) Αν $k > n$, τότε τα διανύσματα, τύπου $n \times 1$, \mathbf{u}_i είναι γραμμικά εξαρτημένα.

(ii) Τα n γραμμικά ανεξάρτητα διανύσματα αποτελούν μία βάση του χώρου των διανυσμάτων στήλης τύπου $n \times 1$, δηλαδή, κάθε διάνυσμα στήλης μπορεί να γραφεί ως γραμμικός συνδυασμός των γραμμικά ανεξάρτητων διανυσμάτων.

(iii) Τα διανύσματα:

$$\mathbf{e}_1^T = [1 \ 0 \ 0 \ \dots \ 0], \mathbf{e}_2^T = [0 \ 1 \ 0 \ \dots \ 0], \dots, \mathbf{e}_n^T = [0 \ 0 \ 0 \ \dots \ 1]$$

είναι μία βάση του χώρου των διανυσμάτων στήλης $n \times 1$ και λέγεται **κανονική**.

Αν $\mathbf{u} = [u_1 \ u_2 \ \dots \ u_n]^T$, τότε $\mathbf{u} = u_1 \mathbf{e}_1 + u_2 \mathbf{e}_2 + \dots + u_n \mathbf{e}_n$.

1.6. Εσωτερικό γινόμενο.

Ορισμός 1.1 Το **εσωτερικό γινόμενο** δυο διανυσμάτων $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ορίζεται από τη σχέση:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i = \mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x}$$

αν $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$, $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^T$.

Για $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, το εσωτερικό γινόμενο ορίζεται από τη σχέση:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i \bar{y}_i = \mathbf{y}^* \mathbf{x}.$$

Ορίζουμε για κάθε $\mathbf{x} \in \mathbb{R}^n$ ή $\mathbf{x} \in \mathbb{C}^n$ το μήκος του \mathbf{x} με

$$\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2} \quad (1.4)$$

και λέγεται **Ευκλείδεια νόρμα**.

Τα επόμενα αποτελέσματα προκύπτουν άμεσα από τους παραπάνω ορισμούς. Έτσι, αν συμβολίσουμε με V το διανυσματικό χώρο \mathbb{R}^n ή \mathbb{C}^n , ισχύουν τα παρακάτω:

i) Για κάθε $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$,

$$\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle,$$

$$\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle.$$

ii) Για κάθε $\mathbf{x}, \mathbf{y} \in V$,

$$\langle a\mathbf{x}, \mathbf{y} \rangle = a \langle \mathbf{x}, \mathbf{y} \rangle, \text{ με } a \in \mathbb{R} \text{ ή } a \in \mathbb{C}$$

$$\text{και } \square \text{ για } V = \mathbb{R}^n, a \in \mathbb{R}: \langle \mathbf{x}, a\mathbf{y} \rangle = a \langle \mathbf{x}, \mathbf{y} \rangle,$$

$$\square \text{ για } V = \mathbb{C}^n, a \in \mathbb{C}: \langle \mathbf{x}, a\mathbf{y} \rangle = \bar{a} \langle \mathbf{x}, \mathbf{y} \rangle.$$

$$\text{iii)} \square \text{ Αν } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \text{ τότε } \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle,$$

$$\square \text{ αν } \mathbf{x}, \mathbf{y} \in \mathbb{C}^n, \text{ τότε } \langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}.$$

$$\text{iv)} \text{ Για κάθε } \mathbf{x} \in V, \text{ ισχύει}$$

$$\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$$

και

$$\langle \mathbf{x}, \mathbf{x} \rangle = 0, \text{ αν και μόνο αν } \mathbf{x} = \mathbf{0}.$$

$$\text{v)} \text{ Ανισότητα Cauchy-Schwartz.}$$

Για κάθε $\mathbf{x}, \mathbf{y} \in V$ ισχύει :

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle,$$

όπου η ισότητα ισχύει, αν και μόνο αν τα διανύσματα \mathbf{x} και \mathbf{y} είναι γραμμικά εξαρτημένα.

Χρησιμοποιώντας την Ευκλείδεια νόρμα η ανισότητα μπορεί να γραφεί:

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2.$$

Ορισμός 1.2 Δύο διανύσματα $\mathbf{x}, \mathbf{y} \in V$ λέγονται **ορθογώνια**, αν $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. Αν $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ είναι μία βάση του \mathbb{R}^n ή \mathbb{C}^n και αν $\langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0$, για κάθε $i \neq j, 1 \leq i, j \leq n$, τότε η βάση $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ λέγεται **ορθογώνια**.

Αν επιπλέον $\|\mathbf{x}_i\|_2 = 1, 1 \leq i \leq n$, τότε η βάση λέγεται **ορθοκανονική**.

1.7. Ορίζουσες.

1.7.1. Βασικοί ορισμοί.

Η **ορίζουσα** ενός τετραγωνικού πίνακα $A = [a_{ij}]$, συμβολίζεται με $\det A$ ή $|a_{ij}|$ και είναι ένας αριθμός που υπολογίζεται από τα στοιχεία του πίνακα A .

- Για τον πίνακα τάξης 1 έχουμε :

$$\det A = |a_{11}| = a_{11}.$$

- Για τον πίνακα τάξης 2 έχουμε:

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

- Οι ορίζουσες των πινάκων τάξης $n > 2$ υπολογίζονται από μία επαναληπτική διαδικασία, χρησιμοποιώντας τα *αλγεβρικά συμπληρώματα* του πίνακα A .

Ορισμός 1.3 Το *αλγεβρικό συμπλήρωμα* A_{ij} του στοιχείου a_{ij} του πίνακα A τάξης n ορίζεται από τον τύπο:

$$A_{ij} = (-1)^{i+j} M_{ij},$$

όπου M_{ij} είναι η ορίζουσα του πίνακα τάξης $n-1$, που προκύπτει από τον A , αν

παραλείψουμε τη γραμμή και τη στήλη στις οποίες ανήκει το a_{ij} , δηλαδή, την i -γραμμή και την j -στήλη.

Τώρα, για κάθε i ή j ($i, j = 1, 2, \dots, n$),

$$\det A = \sum_{k=1}^n a_{ik} A_{ik} = \sum_{k=1}^n a_{kj} A_{kj} \quad (1.5)$$

Για κάθε i , το πρώτο άθροισμα στην (1.5) παριστάνει το ανάπτυγμα της ορίζουσας κατά τα στοιχεία της i -γραμμής του πίνακα A . Για κάθε j , το δεύτερο άθροισμα στην (1.5) παριστάνει το ανάπτυγμα της ορίζουσας κατά τα στοιχεία της j -στήλης του πίνακα A .

Αν η ορίζουσα του πίνακα A είναι διάφορη του μηδενός, ο πίνακας A λέγεται **αντιστρέψιμος (ομαλός)**, ενώ αν είναι μηδενική, ο πίνακας A λέγεται **μη αντιστρέψιμος (ανώμαλος)**.

1.7.2. Ιδιότητες οριζουσών.

i) $\det(A^*) = \det(\overline{A}) = \overline{\det A}$.

ii) Αν A και B είναι δυο τετραγωνικοί πίνακες της ίδιας τάξης, τότε

$$\det(AB) = \det A \det B.$$

iii) Η ορίζουσα ενός τετραγωνικού πίνακα, άνω ή κάτω τριγωνικού, ισούται με το γινόμενο των διαγώνιων στοιχείων.

iv) Αν B είναι ο πίνακας που προκύπτει από έναν τετραγωνικό πίνακα A , αν εναλλάξουμε δύο γραμμές ή δύο στήλες, τότε $\det B = -\det A$.

v) Αν B είναι ο πίνακας που προκύπτει από έναν τετραγωνικό πίνακα A , πολλαπλασιάζοντας τα στοιχεία μίας γραμμής ή μίας στήλης του επί έναν αριθμό λ , τότε $\det B = \lambda \det A$.

- vi)** Αν μία γραμμή (ή μία στήλη) ενός τετραγωνικού πίνακα A είναι άθροισμα δύο γραμμών (ή δύο στηλών), τότε η ορίζουσα του πίνακα A είναι το άθροισμα των ορίζουσών δύο πινάκων, οι οποίοι αντιστοιχούν στις δύο προσθετέες γραμμές (ή στήλες).
- vii)** Αν δύο γραμμές (ή στήλες) ενός τετραγωνικού πίνακα A είναι ίσες, τότε $\det A = 0$. Πιο γενικά, αν δύο γραμμές (ή στήλες) του A είναι γραμμικά εξαρτημένες, τότε $\det A = 0$.
- viii)** Αν B είναι ο πίνακας που προκύπτει από έναν τετραγωνικό πίνακα A , αν σε μία γραμμή (ή στήλη) προσθέσουμε το γινόμενο μίας άλλης γραμμής (ή στήλης) αυτού επί έναν αριθμό, τότε

$$\det A = \det B.$$

Σημείωση : **(i)** Αν τα στοιχεία της i -γραμμής ή j -στήλης είναι όλα μηδέν, τότε από τον τύπο (1.5) έπεται ότι $\det A = 0$.

(ii) Χρησιμοποιώντας το σύμβολο δέλτα του Kronecker το οποίο ορίζεται

$$\delta_{ik} = \begin{cases} 1, & \text{αν } i = k \\ 0, & \text{αν } i \neq k \end{cases} \quad (1.6)$$

τα δύο αθροίσματα στην (1.5) γράφονται :

$$\begin{aligned} a_{k1}A_{11} + a_{k2}A_{12} + \dots + a_{kn}A_{1n} &= \det A \cdot \delta_{ik} \\ a_{1k}A_{1i} + a_{2k}A_{2i} + \dots + a_{nk}A_{ni} &= \det A \cdot \delta_{ik} \end{aligned} \quad (1.7)$$

1.8. Ειδικές κατηγορίες πινάκων.

1.8.1. Συμπληρωματικός πίνακας.

Ονομάζουμε συμπληρωματικό πίνακα του τετραγωνικού πίνακα τάξης $n \geq 2$ και τον συμβολίζουμε με $adjA$, τον πίνακα:

$$adjA = \begin{bmatrix} A_{11} & \dots & A_{j1} & \dots & A_{n1} \\ \vdots & & \vdots & & \vdots \\ A_{1j} & \dots & A_{ji} & \dots & A_{ni} \\ \vdots & & \vdots & & \vdots \\ A_{1n} & \dots & A_{jn} & \dots & A_{nn} \end{bmatrix} = [A_{ij}]^T$$

όπου A_{ij} είναι το αλγεβρικό συμπλήρωμα του στοιχείου a_{ij} .

Από τους τύπους (1.5) και τους κανόνες πολλαπλασιασμού πινάκων προκύπτουν οι τύποι:

$$A \cdot (\text{adj}A) = (\text{adj}A) \cdot A = \det A \cdot I \quad (1.8)$$

1.8.2. Αντίστροφος πίνακας.

Αν ένας πίνακας A , τάξης n είναι αντιστρέψιμος, τότε ο αντίστροφός του ορίζεται μονοσήμαντα από τη σχέση:

$$A^{-1} = \frac{1}{\det A} \text{adj}A$$

και ισχύει ο τύπος

$$AA^{-1} = A^{-1}A = I.$$

Ιδιότητες αντίστροφου πίνακα.

- i) $(AB)^{-1} = B^{-1}A^{-1}$.
- ii) $\det(A^{-1}) = (\det A)^{-1}$.
- iii) $(\overline{A})^{-1} = \overline{(A^{-1})}$.
- iv) $(A^T)^{-1} = (A^{-1})^T$.
- v) $(A^*)^{-1} = (A^{-1})^*$.

με την προϋπόθεση ότι οι παραπάνω σημειούμενοι αντίστροφοι πίνακες υπάρχουν.

Από τον αντίστροφο πίνακα A ορίζουμε τις αρνητικές και ακέραιες δυνάμεις του, σύμφωνα με τον τύπο:

$$A^{-k} = (A^{-1})^k = (A^k)^{-1}, \quad (k = 2, 3, \dots).$$

1.8.3. Ορθομοναδιαίος πίνακας.

Ένας τετραγωνικός πίνακας A λέγεται **ορθομοναδιαίος**, όταν ισχύει μία από τις δύο επόμενες ισοδύναμες σχέσεις:

$$AA^* = I, \quad A^*A = I.$$

Άρα

$$AA^* = A^*A = I.$$

Ιδιότητες.

- i) Ο αναστροφosuζυγής κάθε ορθομοναδιαίου πίνακα A ισούται με τον αντίστροφο του πίνακα αυτού. Δηλαδή, $A^{-1} = A^*$.
- ii) Αν ο A είναι ορθομοναδιαίος πίνακας, τότε $|\det A| = 1$.
- iii) Αν οι πίνακες A, B είναι ορθομοναδιαίοι, τότε ορθομοναδιαίοι είναι και οι πίνακες: \overline{A} , A^T , A^{-1} , AB .

Ένας πραγματικός ορθομοναδιαίος πίνακας λέγεται **ορθογώνιος** και ισχύουν οι σχέσεις:

$$AA^T = A^T A = I \quad \text{και} \quad \det A = \pm 1.$$

Ορισμός 1.4 Ένας πίνακας A λέγεται **κανονικός**, όταν ισχύει η σχέση:

$$AA^* = A^* A.$$

Ορισμός 1.5 Λέμε ότι ο πίνακας A είναι **ισοδύναμος** προς τον πίνακα B , όταν υπάρχουν δυο ομαλοί πίνακες R, S , τέτοιοι ώστε :

$$A = RBS.$$

Από αυτή προκύπτει $B = R^{-1}AS^{-1}$, δηλαδή, ο πίνακας B είναι ισοδύναμος προς τον A . Κατόπιν αυτού έχει νόημα η έκφραση «οι πίνακες A, B είναι **ισοδύναμοι**».

Ορισμός 1.6 Λέμε ότι ο πίνακας A είναι **όμοιος** προς τον πίνακα B , όταν υπάρχει αντιστρέψιμος πίνακας P , τέτοιος ώστε:

$$A = P^{-1}BP.$$

Από αυτήν προκύπτει ότι $B = P^{-1}AP$, δηλαδή ο πίνακας B είναι όμοιος προς τον A . Άρα έχει νόημα η έκφραση «οι πίνακες A, B είναι **όμοιοι**».

Ορισμός 1.7 Λέμε ότι ο πίνακας A είναι **ισότιμος** προς τον B , όταν υπάρχει αντιστρέψιμος πίνακας P , τέτοιος ώστε:

$$A = P^T BP.$$

Από αυτήν προκύπτει ότι $B = (P^{-1})^T AP^{-1}$.

Άρα έχει νόημα η έκφραση «οι πίνακες A, B είναι **ισότιμοι**».

1.8.4. Ρητές συναρτήσεις πινάκων.

Θεωρούμε το βαθμωτό πολυώνυμο:

$$f(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n$$

και τον τετραγωνικό πίνακα A . Ορίζουμε τον πίνακα $f(A)$, με τη σχέση:

$$f(A) = a_0 A^n + a_1 A^{n-1} + \dots + a_{n-1} A + a_n I$$

και την ονομάζουμε **πολυωνυμική συνάρτηση** του πίνακα A , αντίστοιχη του βαθμωτού πολυωνύμου $f(x)$. Λέμε επίσης ότι και το $f(x)$ είναι αντίστοιχο της πολυωνυμικής συνάρτησης $f(A)$.

Αν $f(x), g(x)$ είναι δύο οποιαδήποτε βαθμωτά πολυώνυμα, τότε για κάθε πίνακα A ισχύουν οι σχέσεις:

$$f(x) + g(x) = \sigma(x) \Rightarrow f(A) + g(A) = \sigma(A) \quad (1.9)$$

$$f(x)g(x) = \sigma(x) \Rightarrow f(A)g(A) = \sigma(A)$$

Αν, επιπλέον, ο πίνακας $g(A)$ είναι αντιστρέψιμος, ορίζουμε το πηλίκο αυτών με τον τύπο

$$\frac{f(A)}{g(A)} = [g(A)]^{-1} f(A) = f(A)[g(A)]^{-1}. \quad (1.10)$$

ΚΕΦΑΛΑΙΟ 2

ΓΡΑΜΜΙΚΑ ΣΥΣΤΗΜΑΤΑ

2.1. Εισαγωγικές έννοιες.

Βασικός σκοπός του κεφαλαίου αυτού είναι η λύση της εξίσωσης $Ax = b$.

Προηγουμένως θα δώσουμε απαραίτητους ορισμούς, που θα χρησιμοποιήσουμε στη συνέχεια:

Ονομάζουμε **υποπίνακα** ενός δοθέντος πίνακα, ο οποίος προκύπτει από αυτόν, αν διαγράψουμε όλα τα στοιχεία που ανήκουν σ' ένα ορισμένο πλήθος γραμμών είτε σε ένα ορισμένο πλήθος στηλών. Την ορίζουσα κάθε τετραγωνικού υποπίνακα τάξης p ενός πίνακα $A = [a_{ij}]$ θα την ονομάσουμε **ελάσσονα ορίζουσα** τάξης p του πίνακα.

Αν οι γραμμές και οι στήλες που απομένουν μετά τη διαγραφή είναι τάξης i_1, i_2, \dots, i_p και k_1, k_2, \dots, k_p αντίστοιχα, τότε η ελάσσονα ορίζουσα είναι:

$$\begin{vmatrix} a_{i_1 k_1} & \dots & a_{i_1 k_p} \\ \vdots & \dots & \vdots \\ a_{i_p k_1} & \dots & a_{i_p k_p} \end{vmatrix}.$$

Ειδικότερα, αν ο πίνακας A είναι τετραγωνικός και $i_1 = k_1, i_2 = k_2, \dots, i_p = k_p$, τότε η παραπάνω ορίζουσα λέγεται **πρωτεύουσα ελάσσονα** ορίζουσα τάξης p . Προφανώς τα στοιχεία της πρωτεύουσας αυτής διαγωνίου είναι συγχρόνως και στοιχεία της πρωτεύουσας διαγωνίου του πίνακα A . Αν, επί πλέον, είναι $1 = i_1 = k_1, 2 = i_2 = k_2, \dots, p = i_p = k_p$, τότε λέγεται **κύρια πρωτεύουσα ελάσσονα** τάξης p του πίνακα. Η ορίζουσα ενός τετραγωνικού πίνακα είναι προφανώς η μέγιστης τάξης κύρια πρωτεύουσα ελάσσονα ορίζουσα του πίνακα.

Παράδειγμα 2.1. Έστω ο πίνακας

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

□ Τότε η ορίζουσα:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

είναι κύρια πρωτεύουσα ελάσσονα δεύτερης τάξης.

□ Η ορίζουσα:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

είναι κύρια πρωτεύουσα ελάσσονα τρίτης τάξης.

□ Η ορίζουσα:

$$\begin{vmatrix} \alpha_{33} & \alpha_{34} \\ \alpha_{43} & \alpha_{44} \end{vmatrix}$$

είναι πρωτεύουσα ελάσσονα ορίζουσα δεύτερης τάξης.

□ Η ορίζουσα

$$\begin{vmatrix} a_{11} & a_{13} & a_{14} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix}$$

είναι μία ελάσσονα ορίζουσα τρίτης τάξης κ.ο.κ. ■

2.2. Βαθμός πίνακα.

Λέμε ότι ένας μη μηδενικός πίνακας A έχει **διάσταση** ή **βαθμό** p και γράφουμε $rank A = p$, όταν υπάρχει ελάσσονα ορίζουσα αυτού τάξης p , διάφορος του μηδενός, ενώ κάθε ελάσσονα του μεγαλύτερης τάξης είναι μηδέν. Ο βαθμός κάθε μηδενικού πίνακα ορίζεται ίσος με μηδέν.

Από τον ορισμό αυτό προκύπτει ότι ο βαθμός ενός πίνακα $m \times n$ είναι ένας μη αρνητικός ακέραιος, ο οποίος δεν υπερβαίνει το $\min\{m, n\}$. Από τον ορισμό του βαθμού πίνακα και τις γνωστές ιδιότητες των οριζουσών προκύπτουν αμέσως οι ιδιότητες:

i) $rank A = rank A^T$.

ii) Για κάθε $\lambda \neq 0$ είναι $rank(\lambda A) = rank A$.

iii) Για κάθε υποπίνακα B του A είναι $rank B \leq rank A$.

iv) Αν εναλλάξουμε τις γραμμές ή τις στήλες ενός πίνακα, προκύπτει πίνακας του ίδιου βαθμού με τον αρχικό.

Θεώρημα 2.1 Ο βαθμός πίνακα ισούται με το μέγιστο πλήθος των γραμμικά ανεξάρτητων γραμμών του πίνακα, καθώς και με το μέγιστο πλήθος των γραμμικά ανεξάρτητων στηλών του.

Απόδειξη. Βλέπε για παράδειγμα Δ. Δασκαλόπουλος, [9], σελ. 141. ■

Θεώρημα 2.2 Ένας πίνακας έχει βαθμό p , αν και μόνον αν ο διανυσματικός χώρος που προκύπτει από τις γραμμές (ή στήλες) του έχει διάσταση ίση με p .

Απόδειξη. Βλέπε για παράδειγμα Δ. Δασκαλόπουλος, [9], σελ. 142. ■

Θεώρημα 2.3 Για έναν πίνακα, αν υπάρχει ελάχιστο ορίζουσα τάξης p διάφορη του μηδενός, ενώ όλες οι προκύπτουσες ελάχιστονες ορίζουσες τάξης $p+1$ είναι μηδέν, τότε και κάθε άλλη ελάχιστο ορίζουσα τάξης $p+1$ είναι μηδέν και ο βαθμός του πίνακα είναι p .

Απόδειξη. Βλέπε για παράδειγμα Δ. Δασκαλόπουλος, [9], σελ. 143. ■

Παράδειγμα 2.2. Να υπολογιστεί ο βαθμός του πίνακα

$$A = \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 2 & 1 & 1 \\ 1 & -2 & -2 & -1 \end{bmatrix}.$$

Λύση. Η ελάχιστο ορίζουσα δεύτερης τάξης $\begin{vmatrix} 1 & 4 \\ 2 & 2 \end{vmatrix}$ είναι διάφορη του μηδενός.

Υπάρχουν τέσσερις ελάχιστονες τρίτης τάξης, αλλά κατά το θεώρημα 2.3, αρκεί να εξετάσουμε μόνο τις ορίζουσες:

$$\begin{vmatrix} 1 & 4 & 3 \\ 2 & 2 & 1 \\ 1 & -2 & -2 \end{vmatrix}, \begin{vmatrix} 1 & 4 & 2 \\ 2 & 2 & 1 \\ 1 & -2 & -1 \end{vmatrix},$$

οι οποίες προκύπτουν από την πλαισίωση της μη μηδενικής ελάχιστο ορίζουσα δεύτερης τάξης.

Εύκολα διαπιστώνουμε ότι και οι δύο αυτές ορίζουσες είναι μηδέν.

Πράγματι, η πρώτη ορίζουσα έχει την 1^η με την 3^η γραμμή γραμμικά εξαρτημένες, επειδή ισχύει

$$\begin{vmatrix} 1 & 4 & 3 \\ 2 & 2 & 1 \\ 1 & -2 & -2 \end{vmatrix} = \begin{vmatrix} -1 & 2 & 2 \\ 2 & 2 & 1 \\ 1 & -2 & -2 \end{vmatrix} = 0,$$

και η δεύτερη ορίζουσα έχει τη 2^η με την 3^η στήλη γραμμικά εξαρτημένες.

Επομένως $\text{rank}A = 2$. ■

2.3. Στοιχειώδεις μετασχηματισμοί.

Έστω ο πίνακας $A = [a_{ij}]$, τύπου $m \times n$. Θα συμβολίζουμε τις γραμμές και τις στήλες αυτού ως εξής:

$$\mathbf{a}^i = [a_{i1} \quad a_{i2} \quad \dots \quad a_{in}] \quad \text{και} \quad \mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Χρησιμοποιώντας τους συμβολισμούς αυτούς, θα γράφουμε τον πίνακα A καταχρηστικά και με τη μορφή:

$$A = [a_{ij}] = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n] = \begin{bmatrix} \mathbf{a}^1 \\ \mathbf{a}^2 \\ \vdots \\ \mathbf{a}^m \end{bmatrix}.$$

Ονομάζουμε **στοιχειώδη μετασχηματισμό** ενός πίνακα $A = [a_{ij}]$ μία από τις επόμενες πράξεις ως προς τις γραμμές ή στήλες του:

- Η εναλλαγή δύο γραμμών \mathbf{a}^i , \mathbf{a}^j συμβολίζεται με H_{ij} .

Η εναλλαγή δύο στηλών \mathbf{a}_i , \mathbf{a}_j συμβολίζεται με Θ_{ij} .

- Ο πολλαπλασιασμός της γραμμής \mathbf{a}^i επί έναν αριθμό $\lambda \neq 0$ συμβολίζεται με $H_i(\lambda)$.

Ο πολλαπλασιασμός της στήλης \mathbf{a}_i επί έναν αριθμό $\lambda \neq 0$ συμβολίζεται με $\Theta_i(\lambda)$.

- Πρόσθεση του γινομένου μίας γραμμής \mathbf{a}^i επί έναν αριθμού λ στη γραμμή \mathbf{a}^j συμβολίζεται με $H_{ij}(\lambda)$.

- Πρόσθεση του γινομένου μίας στήλης \mathbf{a}_i επί έναν αριθμού λ στη στήλη \mathbf{a}_j συμβολίζεται με $\Theta_{ij}(\lambda)$.

Τους στοιχειώδεις μετασχηματισμούς γραμμών θα συμβολίζουμε **H-μετασχηματισμούς** και των στηλών **Θ-μετασχηματισμούς**.

Παράδειγμα 2.3. Να εφαρμοστούν ορισμένοι μετασχηματισμοί στον πίνακα

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ -1 & 1 & 2 \\ 1 & 1 & 0 \end{bmatrix}.$$

Λύση. Ο παραπάνω πίνακας μέσω της πράξης Θ_{23} γίνεται:

$$\begin{bmatrix} 1 & 3 & 2 \\ 1 & 1 & 0 \\ -1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} = B.$$

Με $H_1(2)$ έχουμε:

$$\begin{bmatrix} 2 & 6 & 4 \\ 1 & 1 & 0 \\ -1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} = C,$$

και εκτελώντας την $H_{23}(3)$:

$$\begin{bmatrix} 2 & 6 & 4 \\ 1 & 1 & 0 \\ 2 & 5 & 1 \\ 1 & 0 & 1 \end{bmatrix} = D. \quad \blacksquare$$

Ονομάζουμε **αντίστροφο** ενός στοιχειώδους μετασχηματισμού την πράξη σε έναν πίνακα, η οποία αναιρεί το αποτέλεσμα του εν λόγω μετασχηματισμού.

Οι αντίστροφοι στοιχειώδεις μετασχηματισμοί είναι:

- i) $H_{ij}^{-1} = H_{ij}, \quad \Theta_{ij}^{-1} = \Theta_{ij}.$
- ii) $H_i^{-1}(\lambda) = H_i\left(\frac{1}{\lambda}\right), \quad \Theta_i^{-1}(\lambda) = \Theta_i\left(\frac{1}{\lambda}\right).$
- iii) $H_{ij}^{-1}(\lambda) = H_{ij}(-\lambda), \quad \Theta_{ij}^{-1}(\lambda) = \Theta_{ij}(-\lambda).$

Κατόπιν αυτών, για τους πίνακες A, B, C, D του παραδείγματος 2.3 ισχύει επίσης και το ακόλουθο σχήμα:

$$A \xleftarrow{\Theta_{23}} B \xleftarrow{H_1\left(\frac{1}{2}\right)} C \xleftarrow{H_{23}(-3)} D.$$

Ονομάζουμε **στοιχειώδη πίνακα** n τάξης τον πίνακα που προκύπτει από το μοναδιαίο πίνακα n τάξης, μ' ένα στοιχειώδη μετασχηματισμό. Χάριν απλότητας, θα παριστάνουμε τους στοιχειώδεις πίνακες με τα αντίστοιχα σύμβολα των στοιχειωδών μετασχηματισμών. Είναι προφανές ότι για τους στοιχειώδεις πίνακες της ίδιας τάξης ισχύουν οι σχέσεις:

$$H_{ij} = \Theta_{ij}, \quad H_i(\lambda) = \Theta_i(\lambda), \quad H_{ij}(\lambda) = \Theta_{ji}(\lambda). \quad (2.1)$$

Παράδειγμα για τους στοιχειώδεις πίνακες τρίτης τάξης, οι οποίοι αντιστοιχούν στους στοιχειώδεις μετασχηματισμούς H_{23} , $H_2(2)$, $H_{23}(-2)$, είναι :

$$H_{23} = \Theta_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad H_2(2) = \Theta_2(2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$H_{23}(-2) = \Theta_{32}(-2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}.$$

Επειδή κάθε μοναδιαίος πίνακας είναι αντιστρέψιμος, από τις ιδιότητες των οριζουσών συμπεραίνεται άμεσα ότι κάθε στοιχειώδης πίνακας είναι επίσης αντιστρέψιμος. Ο αντίστροφος ενός στοιχειώδους πίνακα αντιστοιχεί προφανώς στον αντίστροφο στοιχειώδη μετασχηματισμό, επομένως ισχύουν οι τύποι :

$$H_{ij}^{-1} = H_{ij}, \quad H_i^{-1}(\lambda) = H_i\left(\frac{1}{\lambda}\right), \quad H_{ij}^{-1}(\lambda) = H_{ij}(-\lambda)$$

$$\Theta_{ij}^{-1} = \Theta_{ij}, \quad \Theta_i^{-1}(\lambda) = \Theta_i\left(\frac{1}{\lambda}\right), \quad \Theta_{ij}^{-1}(\lambda) = \Theta_{ij}(-\lambda).$$

Θεώρημα 2.4 Κάθε στοιχειώδης μετασχηματισμός στηλών (γραμμών) ενός πίνακα $A = [a_{ij}]$, τύπου $m \times n$, είναι ισοδύναμος με πολλαπλασιασμό του A προς τα δεξιά (αριστερά) επί τον αντίστοιχο στοιχειώδη πίνακα, ο οποίος προκύπτει από το I_n (I_m).

Απόδειξη. Θέτουμε $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n]$ και $I = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_n]$. Σύμφωνα με τον κανόνα πολλαπλασιασμού σύνθετων πινάκων ισχύει για κάθε στήλη $\mathbf{e}_i = [\delta_{i1} \quad \delta_{i2} \quad \dots \quad \delta_{in}]^T$ του πίνακα I_n , η σχέση:

$$A\mathbf{e}_i = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} \delta_{1i} \\ \vdots \\ \delta_{ii} \\ \vdots \\ \delta_{ni} \end{bmatrix} = \mathbf{a}_1\delta_{1i} + \dots + \mathbf{a}_i\delta_{ii} + \dots + \mathbf{a}_n\delta_{ni} = \mathbf{a}_i,$$

όπου με δ_{ik} συμβολίζεται το δέλτα του Kronecker (βλέπε σχέση (1.6)).

Οπότε έχουμε:

$$\begin{aligned} A\Theta_{ij} &= A \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_j & \dots & \mathbf{e}_i & \dots & \mathbf{e}_n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{e}_1 & A\mathbf{e}_2 & \dots & A\mathbf{e}_j & \dots & A\mathbf{e}_i & \dots & A\mathbf{e}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_j & \dots & \mathbf{a}_i & \dots & \mathbf{a}_n \end{bmatrix} \end{aligned}$$

$$\begin{aligned} A\Theta_i(\lambda) &= A \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \lambda\mathbf{e}_i & \dots & \mathbf{e}_n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{e}_1 & A\mathbf{e}_2 & \dots & \lambda A\mathbf{e}_i & \dots & A\mathbf{e}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \lambda\mathbf{a}_i & \dots & \mathbf{a}_n \end{bmatrix} \end{aligned}$$

$$\begin{aligned} A\Theta_{ij}(\lambda) &= A \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \lambda\mathbf{e}_i + \mathbf{e}_j & \dots & \mathbf{e}_n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{e}_1 & A\mathbf{e}_2 & \dots & \lambda A\mathbf{e}_i + A\mathbf{e}_j & \dots & A\mathbf{e}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \lambda\mathbf{a}_i + \mathbf{a}_j & \dots & \mathbf{a}_n \end{bmatrix} \end{aligned}$$

Από τις σχέσεις αυτές προκύπτει η αλήθεια του θεωρήματος για τους στοιχειώδεις μετασχηματισμούς στηλών.

Με ανάλογο τρόπο αποδεικνύεται και για τους στοιχειώδεις μετασχηματισμούς γραμμών. ■

Θεώρημα 2.5 Ο τύπος και ο βαθμός ενός πίνακα δε μεταβάλλονται μετά από κάθε στοιχειώδη μετασχηματισμό του πίνακα.

Απόδειξη. Έστω ο πίνακας A , τύπου $m \times n$, τότε κάθε H -μετασχηματισμός αυτού είναι ισοδύναμος με τον πίνακα HA , ο οποίος είναι πάλι τύπου $m \times n$, επειδή εξ' ορισμού ο πίνακας H είναι τύπου $m \times m$. Αντίστοιχα, κάθε Θ -μετασχηματισμός του πίνακα A είναι ισοδύναμος με το γινόμενο $A\Theta$, ο οποίος είναι πάλι τύπου $m \times n$, επειδή εξ' ορισμού ο πίνακας Θ είναι τύπου $n \times n$. Έτσι αποδείξαμε το πρώτο μισό του θεωρήματος.

Για την απόδειξη του δεύτερου μέρους εργαζόμαστε ως εξής:

Έστω $\text{rank}A = \lambda$ και P μία οποιαδήποτε ελάσσονα ορίζουσα του, τάξης $\lambda + 1$, τότε, από το θεώρημα 2.3 ισχύει ότι $P = 0$.

Μετά από ένα μετασχηματισμό H_{ij} του πίνακα A προκύπτει ο πίνακας $B = H_{ij}A$.

Οπότε η ορίζουσα P μετασχηματίζεται σε μία άλλη, την οποία ονομάζουμε Q . Ο μετασχηματισμός H_{ij} σημαίνει εναλλαγή των γραμμών \mathbf{a}^i και \mathbf{a}^j . Τότε:

- Η ορίζουσα P δεν έχει στοιχεία επί των δύο εναλλασσόμενων γραμμών του πίνακα A , δηλαδή, η P μένει αμετάβλητη, οπότε $P = Q = 0$.
- Η ορίζουσα P έχει μόνο στοιχεία επί μίας των εναλλασσόμενων γραμμών, π.χ. επί της α^i . Οπότε η Q είναι μία ελάσσονα ορίζουσα τάξης $\lambda + 1$, επομένως $Q = 0$.
- Η ορίζουσα P έχει στοιχεία και επί των δύο εναλλασσόμενων γραμμών α^i και α^j . Οπότε, από την ιδιότητα (iv) των οριζουσών, $Q = -P = 0$. Άρα κάθε ελάσσονα ορίζουσα του πίνακα B , τάξης $\lambda + 1$ είναι πάντοτε μηδέν.

Έτσι καταλήγουμε ότι ισχύει

$$\text{rank}B = \text{rank}(H_{ij}A) \leq \lambda = \text{rank}A.$$

Επειδή $H_{ij}^{-1} = H_{ij}$ από την ισότητα $B = H_{ij}A$ προκύπτει ότι ισχύει $H_{ij}B = A$, οπότε

$$\text{rank}A = \text{rank}(H_{ij}B) \leq \text{rank}B.$$

Άρα $\text{rank}A = \text{rank}B = \lambda$.

Με τον ίδιο τρόπο ακριβώς γίνεται η απόδειξη και για τους δυο άλλους μετασχηματισμούς. ■

Θεώρημα 2.6 Κάθε πίνακας βαθμού λ ανάγεται με ένα πεπερασμένο πλήθος στοιχειωδών μετασχηματισμών στην κανονική μορφή K , δηλαδή, σε έναν πίνακα με μία από τις παρακάτω τέσσερις μορφές:

$$I_\lambda, \quad [I_\lambda \quad \mathbf{0}], \quad \begin{bmatrix} I_\lambda \\ \mathbf{0} \end{bmatrix}, \quad \begin{bmatrix} I_\lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Ειδικότερα, αν ο πίνακας είναι αντιστρέψιμος τάξης n , τότε η κανονική μορφή του είναι ο μοναδιαίος πίνακας I_n .

Απόδειξη. Βλέπε για παράδειγμα Δ. Δασκαλόπουλος, [9], σελ.151 και R. Bronson, σελ. 41. ■

Θεώρημα 2.7 Αν K είναι η κανονική μορφή του πίνακα A , τότε υπάρχουν δύο ομαλοί πίνακες P, Q , τέτοιοι ώστε να ισχύει η σχέση $PAQ = K$.

Απόδειξη. Έστω ότι ο πίνακας A ανάγεται στην κανονική μορφή K με εφαρμογή σ' αυτόν διαδοχικά των H -μετασχηματισμών $P_1, P_2, \dots, P_\sigma$ και των Θ -μετασχηματισμών Q_1, Q_2, \dots, Q_p . Σύμφωνα με το προηγούμενο θεώρημα, θα έχουμε τη σχέση:

$$(P_\sigma \cdots P_1)A(Q_1 \cdots Q_p) = K.$$

Επειδή όμως οι στοιχειώδεις πίνακες είναι ομαλοί, θα είναι ομαλοί και οι πίνακες $P = P_\sigma \cdots P_1$ και $Q = Q_1 \cdots Q_p$, δηλαδή,

$$PAQ = K,$$

όπου P και Q ομαλοί πίνακες. ■

Θεώρημα 2.8 Κάθε αντιστρέψιμος πίνακας μπορεί να γραφεί ως γινόμενο στοιχειωδών πινάκων.

Απόδειξη. Χρησιμοποιώντας τα σύμβολα του θεωρήματος 2.7, προκύπτει ότι

$$PAQ = I_n,$$

όπου P, Q ομαλοί. Άρα

$$A = P^{-1}I_nQ^{-1} = P^{-1}Q^{-1} = (P_\sigma \cdots P_1)^{-1}(Q_1 \cdots Q_p)^{-1} = P_1^{-1} \cdots P_\sigma^{-1}Q_p^{-1} \cdots Q_1^{-1},$$

όπου P_i^{-1} , ($i=1,2,\dots,\sigma$), Q_j^{-1} , ($j=1,2,\dots,p$) είναι αντίστροφοι στοιχειώδεις πίνακες, οι οποίοι, ως γνωστό, είναι επίσης στοιχειώδεις πίνακες. ■

Από τα παραπάνω θεωρήματα προκύπτει η ισοδυναμία των παρακάτω ισχυρισμών:

- i) Οι πίνακες A, B είναι ισοδύναμοι.
- ii) Οι πίνακες A, B είναι του ίδιου τύπου και βαθμού.
- iii) Οι πίνακες A, B έχουν την ίδια κανονική μορφή.

Επίσης, για το βαθμό πίνακα ισχύουν και οι επόμενες προτάσεις:

Θεώρημα 2.9 Αν ο πίνακας A είναι αντιστρέψιμος, τότε $rank(AB) = rankB$. Το ίδιο ισχύει και για το γινόμενο BA .

Θεώρημα 2.10 Ο βαθμός του γινομένου AB δεν υπερβαίνει το βαθμό του κάθε παράγοντα.

Θεώρημα 2.11 Αν $AB = \mathbb{O}$ και ένας από τους δύο πίνακες A, B έχει βαθμό λ , τότε ο βαθμός του άλλου δεν υπερβαίνει τον αριθμό $m - \lambda$, όπου m το πλήθος των στηλών του A .

2.4. Επίλυση γραμμικών συστημάτων.

Ορισμός 2.1 Ονομάζουμε **γραμμικό σύστημα** m εξισώσεων με n αγνώστους κάθε σύστημα της μορφής:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_n
\end{aligned}
\tag{2.2}$$

Αν τα δεύτερα μέλη των εξισώσεων είναι μηδέν, το σύστημα λέγεται **ομογενές**.

Το σύστημα (2.2) με $m = n$ και πίνακα A αντιστρέψιμο λέγεται **σύστημα Cramer**.

Ορισμός 2.2 Ονομάζουμε **λύση** του συστήματος (2.2) κάθε διατεταγμένη n -άδα αριθμών x_1, x_2, \dots, x_n , οι οποίοι επαληθεύουν το σύστημα.

Αν υπάρχει λύση, τότε το σύστημα (2.2) λέγεται **συμβιβαστό**.

Αν θέσουμε $A = [a_{ij}]$, $\mathbf{b} = [b_1 \ b_2 \ \dots \ b_m]^T$ και $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_m]^T$, το σύστημα (2.2) γράφεται με την ισοδύναμη μορφή:

$$A\mathbf{x} = \mathbf{b} \tag{2.3}$$

Ορισμός 2.3 Ο πίνακας

$$[A | \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right]$$

λέγεται **επαυξημένος** πίνακας του συστήματος (2.2).

Οι πίνακες A και $[A | \mathbf{b}]$ παίζουν σπουδαίο ρόλο στην επίλυση των γραμμικών συστημάτων.

Θεώρημα 2.12 Κάθε σύστημα Cramer έχει τη μοναδική λύση:

$$\mathbf{x} = A^{-1}\mathbf{b}$$

ή

$$[x_1 \ x_2 \ \dots \ x_n]^T = \frac{1}{|A|} [|A_1| \ |A_2| \ \dots \ |A_n|]^T \tag{2.4}$$

όπου $A_i, i=1, 2, \dots, n$ παριστάνει τον πίνακα που προκύπτει από τον A , με αντικατάσταση της i στήλης από τη στήλη \mathbf{b} .

Αν $\mathbf{b} = \mathbf{0}$, δηλαδή, $A\mathbf{x} = \mathbf{0}$, τότε η μοναδική λύση είναι το μηδενικό διάνυσμα (τετριμμένη λύση).

Σημείωση : Ο συμπληρωματικός πίνακας και ο κανόνας Cramer (τύποι (2.4)) παρουσιάζουν μόνο θεωρητικό ενδιαφέρον. Ο υπολογισμός της ορίζουσας με το ανάπτυγμα (1.5), του αντίστροφου πίνακα με το συμπληρωματικό και η λύση του συστήματος Cramer χρησιμοποιώντας τους τύπους (2.4) δεν έχουν πρακτικό ενδιαφέρον. Αν ένας υπολογιστής εκτελεί περισσότερους από 10^6 πολλαπλασιασμούς και διαιρέσεις 13-ψηφίων αριθμών το δευτερόλεπτο, τότε ο χρόνος που απαιτείται για τη λύση ενός συστήματος Cramer 100 εξισώσεων με 100 αγνώστους είναι 10^{141} αιώνες(!).

Στην πράξη, για τη λύση του συστήματος (2.2), χρησιμοποιούμε μία από τις αριθμητικές μεθόδους Gauss, Cholesky κ.λπ., ανάλογα με τη μορφή του πίνακα A , με τις οποίες το σύστημα αυτό λύνεται σε κλάσματα δευτερολέπτων!

Βλέπε για παράδειγμα Α. Μπακόπουλος- Ι. Χρυσοβέργης, σελ. 1-22 και Γ. Σοφριανός- Ε. Τυχόπουλος σελ.29-62.

Θεώρημα 2.13 Το σύστημα (2.2) είναι συμβιβαστό, αν και μόνο αν

$$\text{rank}A = \text{rank}[A | \mathbf{b}].$$

Σημείωση : Όταν το σύστημα (2.2) είναι ομογενές, είναι πάντοτε συμβιβαστό, επειδή έχει τη μηδενική λύση και είναι πάντοτε $\text{rank}A = \text{rank}[A | \mathbf{0}]$.

Ορισμός 2.4 Λέμε ότι το σύστημα (2.2) έχει k -**παραμετρική απειρία λύσεων**, την οποία συμβολίζουμε ∞^k , όταν όλες οι λύσεις του συστήματος προκύπτουν δίνοντας αυθαίρετες τιμές σε κατάλληλα επιλεγμένους αγνώστους από τους x_1, x_2, \dots, x_n , ενώ οι υπόλοιποι $n-k$ υπολογίζονται, λύνοντας ισάριθμες, κατάλληλα επιλεγμένες εξισώσεις του συστήματος.

Θεώρημα 2.14 Αν για το σύστημα (2.2) έχουμε $\text{rank}A = \text{rank}[A | \mathbf{b}] = k$, τότε το σύστημα αυτό έχει $n-k$ παραμετρική απειρία λύσεων.

Σημείωση : Για να έχει ένα ομογενές σύστημα και μη μηδενικές λύσεις, σύμφωνα με το θεώρημα 2.14, πρέπει και αρκεί να ισχύει $\text{rank}A < n$. Η συνθήκη αυτή πραγματοποιείται πάντοτε για $m < n$, ενώ για $m = n$ πραγματοποιείται αν και μόνο αν $\det A = 0$. Για $m = n-1 = \text{rank}A$, το ομογενές σύστημα $A\mathbf{x} = \mathbf{0}$ έχει μονοπαραμετρική απειρία λύσεων.

Θεώρημα 2.15 Αν το ομογενές σύστημα $A\mathbf{x} = \mathbf{0}$ έχει τις λύσεις $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$, τότε έχει και τη λύση $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_k\mathbf{x}_k$, όπου c_1, c_2, \dots, c_n αυθαίρετες σταθερές.

Ορισμός 2.5 Λέμε ότι οι λύσεις $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ ενός ομογενούς συστήματος $A\mathbf{x} = \mathbf{0}$ αποτελούν ένα k -τάξης θεμελιώδες σύστημα λύσεων του, όταν είναι γραμμικά ανεξάρτητες και κάθε άλλη λύση αυτού μπορεί να γραφεί με τη μορφή

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_k\mathbf{x}_k,$$

για κατάλληλες τιμές των αυθαίρετων σταθερών c_1, c_2, \dots, c_k .

Η λύση

$$\mathbf{x}_0 = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_k\mathbf{x}_k$$

ονομάζεται γενική λύση του ομογενούς συστήματος.

Θεώρημα 2.16 Έστω το ομογενές σύστημα $A\mathbf{x} = \mathbf{0}$, όπου ο πίνακας $A = [a_{ij}]$ είναι τύπου $m \times n$. Αν $\text{rank}A = \tau < n$, τότε το σύστημα έχει ένα $(n - \tau)$ -τάξης θεμελιώδες σύστημα λύσεων.

Ισοδύναμα : Το σύνολο όλων των λύσεων $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$ ενός ομογενούς συστήματος $A\mathbf{x} = \mathbf{0}$ αποτελεί ένα διανυσματικό χώρο διάστασης $d = n - \text{rank}A$.

Θεώρημα 2.17 Αν ξ είναι μία λύση του συστήματος $A\mathbf{x} = \mathbf{b}$ και \mathbf{x}_0 είναι η γενική λύση του αντίστοιχου ομογενούς $A\mathbf{x} = \mathbf{0}$, τότε κάθε λύση \mathbf{x} του συστήματος μπορεί να γραφεί με τη μορφή $\mathbf{x} = \xi + \mathbf{x}_0$, όπου οι αυθαίρετες σταθερές της \mathbf{x}_0 είναι κατάλληλες.

Για τις αποδείξεις των παραπάνω θεωρημάτων βλέπε για παράδειγμα Δ. Δασκαλόπουλος, [9], σελ. 168-172.

Δίνουμε το παρακάτω παράδειγμα για να επαληθεύσουμε τις προτάσεις που αφορούν τη λύση του συστήματος $A\mathbf{x} = \mathbf{b}$.

Παράδειγμα 2.4. Να λυθεί το σύστημα:

$$2x_1 - 2x_2 + x_3 - x_4 + x_5 = 1$$

$$x_1 + 2x_2 - x_3 + x_4 - 2x_5 = 1$$

$$4x_1 - 10x_2 + 5x_3 - 5x_4 + 7x_5 = 1$$

$$2x_1 - 14x_2 + 7x_3 - 7x_4 + 11x_5 = -1$$

Λύση. Ο πίνακας του συστήματος είναι:

$$A = \begin{bmatrix} 2 & -2 & 1 & -1 & 1 \\ 1 & 2 & -1 & 1 & -2 \\ 4 & -10 & 5 & -5 & 7 \\ 2 & -14 & 7 & -7 & 11 \end{bmatrix}$$

Επειδή η 2^η, η 3^η και η 4^η στήλη του πίνακα A είναι γραμμικά εξαρτημένες, συμπεραίνουμε ότι $\text{rank}A = 2$ (θεώρημα 2.1).

Για τον επαυξημένο πίνακα, που είναι

$$[A|\mathbf{b}] = \left[\begin{array}{ccccc|c} 2 & -2 & 1 & -1 & 1 & 1 \\ 1 & 2 & -1 & 1 & -2 & 1 \\ 4 & -10 & 5 & -5 & 7 & 1 \\ 2 & -14 & 7 & -7 & 11 & -1 \end{array} \right]$$

παρατηρούμε ότι ισχύει

$$B = H_{14}(1)H_{13}(-1)H_{12}(-1)[A|\mathbf{b}] = \left[\begin{array}{ccccc|c} 2 & -2 & 1 & -1 & 1 & 1 \\ -1 & 4 & -2 & 2 & -3 & 0 \\ 2 & -8 & 4 & -4 & 6 & 0 \\ 4 & -16 & 8 & -8 & 12 & 0 \end{array} \right],$$

όπου οι τρεις τελευταίες γραμμές του είναι γραμμικά εξαρτημένες.

Επομένως σύμφωνα με το θεώρημα 2.5 έχουμε

$$\text{rank}B = \text{rank}[A|\mathbf{b}] = 2.$$

Άρα $\text{rank}A = \text{rank}[A|\mathbf{b}] = 2$, δηλαδή, το σύστημα είναι συμβιβαστό (θεώρημα 2.13).

Επομένως, το σύστημα έχει $5 - 2 = 3$ παραμετρική απειρία λύσεων (θεώρημα 2.14).

Επειδή $\begin{vmatrix} 2 & -2 \\ 1 & 2 \end{vmatrix} \neq 0$, θεωρούμε το σύστημα

$$\begin{aligned} 2x_1 - 2x_2 &= 1 - x_3 + x_4 - x_5 \\ x_1 + 2x_2 &= 1 + x_3 - x_4 + 2x_5 \end{aligned}$$

όπου οι άγνωστοι x_3, x_4, x_5 παίρνουν αυθαίρετες τιμές.

Λύνοντας το παραπάνω σύστημα βρίσκουμε

$$\begin{aligned} x_1 &= \frac{2}{3} + 0x_3 + 0x_4 + \frac{1}{3}x_5 \\ x_2 &= \frac{1}{6} + \frac{1}{2}x_3 - \frac{1}{2}x_4 + \frac{5}{6}x_5 \end{aligned}$$

Η γενική λύση του συστήματος γράφεται :

$$\begin{aligned}
\mathbf{x} &= [x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5]^T \\
&= \left[\frac{2}{3} + 0x_3 + 0x_4 + \frac{1}{3}x_5 \quad \frac{1}{6} + \frac{1}{2}x_3 - \frac{1}{2}x_4 + \frac{5}{6}x_5 \quad x_3 \quad x_4 \quad x_5 \right]^T \\
&= \left[\frac{2}{3} \quad \frac{1}{6} \quad 0 \quad 0 \quad 0 \right]^T \\
&\quad + c_1 \left[0 \quad \frac{1}{2} \quad 1 \quad 0 \quad 0 \right]^T + c_2 \left[0 \quad -\frac{1}{2} \quad 0 \quad 1 \quad 0 \right]^T + c_3 \left[\frac{1}{3} \quad \frac{5}{6} \quad 0 \quad 0 \quad 1 \right]^T \\
&= \boldsymbol{\xi} + \mathbf{x}_0,
\end{aligned}$$

όπου $\boldsymbol{\xi} = \left[\frac{2}{3} \quad \frac{1}{6} \quad 0 \quad 0 \quad 0 \right]^T$ είναι μία λύση του δοθέντος συστήματος και

$$\mathbf{x}_0 = c_1 \left[0 \quad \frac{1}{2} \quad 1 \quad 0 \quad 0 \right]^T + c_2 \left[0 \quad -\frac{1}{2} \quad 0 \quad 1 \quad 0 \right]^T + c_3 \left[\frac{1}{3} \quad \frac{5}{6} \quad 0 \quad 0 \quad 1 \right]^T$$

είναι η γενική λύση του αντίστοιχου ομογενούς συστήματος.

Τα παραπάνω εξαγόμενα επαληθεύουν το θεώρημα 2.17.

ΚΕΦΑΛΑΙΟ 3

ΙΔΙΟΤΙΜΕΣ ΚΑΙ ΙΔΙΟΔΙΑΝΥΣΜΑΤΑ

3.1. Εισαγωγικοί ορισμοί.

Πολλά προβλήματα στις πρακτικές εφαρμογές ανάγονται στη λύση του παρακάτω μαθηματικού προβλήματος:

Έστω ο τετραγωνικός πίνακας $A = [a_{ij}]$. Αναζητούνται οι τιμές του λ και τα αντίστοιχα \mathbf{x} , έτσι ώστε να ισχύει η εξίσωση:

$$A\mathbf{x} = \lambda\mathbf{x} \quad (3.1)$$

Αν θεωρήσουμε τον αντίστοιχο γραμμικό μετασχηματισμό, προς τον πίνακα A , $\mathbf{y} = A\mathbf{x}$, τότε τα παραπάνω σημαίνουν την εύρεση όλων των \mathbf{x} και λ , για τα οποία οι εικόνες $A\mathbf{x}$ είναι διανύσματα, παράλληλα προς το αρχέτυπο \mathbf{x} . Κάθε τέτοιο **διάνυσμα** λέγεται **αναλλοίωτο διάνυσμα**.

Προφανώς, το μηδενικό διάνυσμα $\mathbf{x} = \mathbf{0}$ είναι ένα αναλλοίωτο διάνυσμα.

Η τιμή αυτή δεν έχει ιδιαίτερο ενδιαφέρον. Μια τιμή του λ , για την οποία η (3.1) έχει λύση $\mathbf{x} \neq \mathbf{0}$ λέγεται **ιδιοτιμή** (ή **χαρακτηριστική τιμή**) του A . Το σύνολο των ιδιοτιμών λέγεται **φάσμα**. Η μεγαλύτερη, κατά μέτρο, ιδιοτιμή λέγεται **φασματική ακτίνα** και συμβολίζεται με $\rho(A)$. Δηλαδή

$$\rho(A) = \max_i \{|\lambda_i| : \lambda_i \text{ είναι ιδιοτιμή του πίνακα } A\}.$$

Οι λύσεις $\mathbf{x} \neq \mathbf{0}$ της (3.1) λέγονται **ιδιοδιανύσματα** (ή **χαρακτηριστικά διανύσματα**) του A , αντίστοιχα της ιδιοτιμής λ .

Το σύνολο των ιδιοδιανυσμάτων που αντιστοιχούν σε μία ιδιοτιμή του A , μαζί με το $\mathbf{0}$ σχηματίζουν ένα διανυσματικό χώρο, που λέγεται **ιδιόχωρος** του A , αντίστοιχος της ιδιοτιμής αυτής.

Οι ιδιοτιμές και τα ιδιοδιανύσματα λέγονται **χαρακτηριστικά μεγέθη** του πίνακα A .

Το πρόβλημα υπολογισμού των ιδιοτιμών και ιδιοδιανυσμάτων του πίνακα A λέγεται **πρόβλημα ιδιοτιμών**.

3.2. Υπολογισμός ιδιοτιμών και ιδιοδιανυσμάτων.

Η εξίσωση (3.1) γράφεται:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= \lambda x_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= \lambda x_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= \lambda x_n
\end{aligned}$$

ή ισοδύναμα

$$\begin{aligned}
(a_{11} - \lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + (a_{22} - \lambda)x_2 + \cdots + a_{2n}x_n &= 0 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + (a_{mn} - \lambda)x_n &= 0
\end{aligned} \tag{3.2}$$

Σε μορφή πίνακα:

$$(A - \lambda I)\mathbf{x} = \mathbf{0}. \tag{3.2a}$$

Από το θεώρημα Cramer, (θεώρημα 2.12), το ομογενές γραμμικό σύστημα (3.2a) έχει μία μηδενική λύση $\mathbf{x} \neq \mathbf{0}$, αν και μόνο αν η αντίστοιχη ορίζουσα, **χαρακτηριστική ορίζουσα** $D(\lambda)$, των συντελεστών είναι μηδέν:

$$D(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} - \lambda \end{vmatrix} = 0. \tag{3.3}$$

Η εξίσωση (3.3) λέγεται **χαρακτηριστική εξίσωση** του πίνακα A .

Αναπτύσσοντας την $D(\lambda)$, παίρνουμε ένα πολυώνυμο n -οστού βαθμού ως προς λ .

Το πολυώνυμο αυτό λέγεται **χαρακτηριστικό πολυώνυμο** του A .

Οι ιδιοτιμές του πίνακα A είναι οι ρίζες της χαρακτηριστικής εξίσωσης (3.3) του A .

Άρα, το πλήθος των διαφορετικών ιδιοτιμών ενός πίνακα $n \times n$ είναι από μία έως n .

Θεώρημα 3.1 Αν \mathbf{x} είναι ένα ιδιοδιάνυσμα του πίνακα A , αντίστοιχο της ιδιοτιμής λ , τότε και το $k\mathbf{x}$, όπου $k \neq 0$, είναι ένα ιδιοδιάνυσμα του A , αντίστοιχο της ιδιοτιμής λ .

Θεωρητικά, δεν υπάρχει δυσκολία για τον καθορισμό των ιδιοτιμών, αλλά στην πράξη συμβαίνει το αντίθετο, επειδή ο υπολογισμός της ορίζουσας (3.3) απαιτεί $n!$ πολλαπλασιασμούς, για μεγάλη τιμή του n το πλήθος των πράξεων του προβλήματος είναι αριθμός απαγορευτικός.

Άλλωστε ο υπολογισμός των ριζών του χαρακτηριστικού πολυωνύμου είναι ένα δύσκολο αλγεβρικό πρόβλημα, γι' αυτό στην πράξη χρησιμοποιούμε αριθμητικές μεθόδους. Για περισσότερες λεπτομέρειες βλέπε για παράδειγμα Α. Μπακόπουλος- Ι. Χρυσοβέργης, σελ. 61, ή Κ. Atkinson chap. 9, ή R. Bronson, chap. 9 και 20.

Δίνουμε το παρακάτω παράδειγμα ως εφαρμογή του προβλήματος υπολογισμού των χαρακτηριστικών μεγεθών ενός πίνακα A .

Το ίδιο παράδειγμα θα βοηθήσει στην κατανόηση των ορισμών που θα δώσουμε αμέσως μετά.

Παράδειγμα 3.1. Να βρεθούν τα χαρακτηριστικά μεγέθη του πίνακα A .

$$A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}.$$

Λύση. Σχηματίζουμε τη χαρακτηριστική εξίσωση:

$$\det(A - \lambda I) = \begin{vmatrix} -2 - \lambda & 2 & -3 \\ 2 & 1 - \lambda & -6 \\ -1 & -2 & -\lambda \end{vmatrix} = 0,$$

δηλαδή, $\lambda^3 + \lambda^2 - 21\lambda - 45 = 0$.

Οι ρίζες (ιδιοτιμές του A) είναι $\lambda_1 = 5$, $\lambda_2 = \lambda_3 = -3$ (διπλή ρίζα).

Άρα το πλήθος των διαφορετικών ριζών είναι 2.

□ Για $\lambda = 5$, προκύπτει το ομογενές σύστημα:

$$\begin{aligned} -7x_1 + 2x_2 - 3x_3 &= 0 \\ 2x_1 - 4x_2 - 6x_3 &= 0 \\ -x_1 - 2x_2 - 5x_3 &= 0 \end{aligned}$$

Ο βαθμός του πίνακα του συστήματος είναι 2. Άρα το σύστημα έχει $(3 - 2 = 1)$ μονοπαραμετρική απειρία λύσεων (θεώρημα 2.14).

Από τις δύο πρώτες εξισώσεις παίρνουμε:

$$\begin{aligned} -7x_1 + 2x_2 &= 3x_3 \\ 2x_1 - 4x_2 &= 6x_3 \end{aligned}$$

όπου η μεταβλητή x_3 παίρνει αυθαίρετες τιμές.

Η ορίζουσα του πίνακα του συστήματος είναι $\begin{vmatrix} -7 & 2 \\ 2 & -4 \end{vmatrix} = 24$. Από το θεώρημα

Cramer, (θεώρημα 2.12), προκύπτουν οι λύσεις:

$$x_1 = \frac{\begin{vmatrix} 3x_3 & 2 \\ 6x_3 & -4 \end{vmatrix}}{24} = \frac{-24x_3}{24} = -x_3, \quad x_2 = \frac{\begin{vmatrix} -7 & 3x_3 \\ 2 & 6x_3 \end{vmatrix}}{24} = \frac{-48x_3}{24} = -2x_3.$$

Άρα η γενική λύση του συστήματος είναι $\mathbf{x} = [-x_3 \ -2x_3 \ x_3]^T$, δηλαδή, $\mathbf{x} = c[-1 \ -2 \ 1]^T$, οπότε ένα ιδιοδιάνυσμα αντίστοιχο της ιδιοτιμής $\lambda = 5$ είναι το $\mathbf{x} = [-1 \ -2 \ 1]^T$ (θεώρημα 3.1).

□ Για $\lambda = -3$, το αντίστοιχο ομογενές σύστημα είναι:

$$\begin{aligned}x_1 + 2x_2 - 3x_3 &= 0 \\2x_1 + 4x_2 - 6x_3 &= 0 \\-x_1 - 2x_2 + 3x_3 &= 0\end{aligned}$$

Ο βαθμός του πίνακα του συστήματος είναι 1. Άρα το σύστημα έχει $(3-1=2)$ δυπαμετρική απειρία λύσεων. Από την πρώτη παίρνουμε:

$$x_1 = -2x_2 + 3x_3,$$

οπότε η γενική λύση αυτού είναι:

$$\mathbf{x} = [-2x_2 + 3x_3 \ x_2 \ x_3]^T = x_2[-2 \ 1 \ 0]^T + x_3[3 \ 0 \ 1]^T$$

και τα δύο γραμμικά ανεξάρτητα ιδιοδιανύσματα αντίστοιχα της $\lambda = -3$ είναι:

$$\mathbf{x}_2 = [-2 \ 1 \ 0]^T, \quad \mathbf{x}_3 = [3 \ 0 \ 1]^T. \quad \blacksquare$$

Σημείωση: (i) Η πολλαπλότητα M_λ μίας ιδιοτιμής ως ρίζας του χαρακτηριστικού πολυωνύμου λέγεται **αλγεβρική πολλαπλότητα**.

Ο αριθμός m_λ των γραμμικά ανεξάρτητων ιδιοδιανυσμάτων, που αντιστοιχούν στην ιδιοτιμή του λ , λέγεται **γεωμετρική πολλαπλότητα** και, κατά το θεώρημα 2.16, ισχύει $m_\lambda = n - \text{rank}(A - \lambda I)$.

Επειδή το χαρακτηριστικό πολυώνυμο είναι βαθμού n το άθροισμα όλων των αλγεβρικών πολλαπλοτήτων είναι ίσο με n .

Η διαφορά $\Delta_\lambda = M_\lambda - m_\lambda$ λέγεται **έλλειμμα** της ιδιοτιμής λ .

Στο παράδειγμα 3.1 έχουμε $M_5 = m_5 = 1$, $M_{-3} = m_{-3} = 2$, $\Delta_5 = 0$, $\Delta_{-3} = 0$.

(ii) α) Γενικά ισχύει

$$1 \leq m_\lambda \leq M_\lambda.$$

β) Αν ο πίνακας A έχει k διαφορετικές μεταξύ τους ιδιοτιμές: $\lambda_1, \lambda_2, \dots, \lambda_k$, με αντίστοιχες γεωμετρικές πολλαπλότητες m_1, m_2, \dots, m_k , τότε ο πίνακας A έχει ακριβώς

$$m = m_1 + m_2 + \dots + m_k \leq n$$

γραμμικά ανεξάρτητα ιδιοδιανύσματα. Η ισότητα ισχύει αν και μόνο αν $m_1 = M_1$, $m_2 = M_2, \dots, m_k = M_k$, όπως συμβαίνει στο παράδειγμα.

(iii) Αν $\text{rank}A = p$, τότε, τουλάχιστον $n - p$ ιδιοτιμές είναι μηδέν.

(iv) Ένας πίνακας είναι μη αντιστρέψιμος, αν το πλήθος των γραμμικά ανεξάρτητων διανυσμάτων είναι μικρότερο του n .

3.3. Βασικές ιδιότητες χαρακτηριστικών μεγεθών.

Δίνουμε τις κυριότερες από τις ιδιότητες των χαρακτηριστικών μεγεθών ενός τετραγωνικού πίνακα τάξης n .

Θεώρημα 3.2 Αν $\lambda_1, \lambda_2, \dots, \lambda_n$ είναι οι ιδιοτιμές του πίνακα A , αριθμώντας και τις πολλαπλές, τότε

i) $\text{tr}A = \lambda_1 + \lambda_2 + \dots + \lambda_n$.

ii) $\det A = \lambda_1 \lambda_2 \dots \lambda_n$.

iii) Οι πίνακες A, A^T έχουν τις ίδιες ιδιοτιμές.

iv) Οι ιδιοτιμές των πινάκων \overline{A}, A^* είναι οι συζυγείς των ιδιοτιμών του A .

Θεώρημα 3.3 Οι ιδιοτιμές ενός τριγωνικού πίνακα είναι ακριβώς τα διαγώνια στοιχεία του.

Θεώρημα 3.4 Όμοιοι πίνακες A και $B = P^{-1}AP$ έχουν το ίδιο χαρακτηριστικό πολυώνυμο και τα ιδιοδιανύσματα \mathbf{x}, \mathbf{y} , που αντιστοιχούν στην ίδια ιδιοτιμή, συνδέονται με τη σχέση $\mathbf{x} = P\mathbf{y}$.

Απόδειξη. Χρησιμοποιώντας την ιδιότητα (ii) των οριζουσών έχουμε:

$$\begin{aligned}\det(A - \lambda I) &= \det(P^{-1}) \det(A - \lambda I) \det(P) \\ &= \det(P^{-1}(A - \lambda I)P) = \det(P^{-1}AP - \lambda I) = \det(B - \lambda I).\end{aligned}$$

Επίσης,

$$A\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow P^{-1}APP^{-1}\mathbf{x} = \lambda P^{-1}\mathbf{x} \Leftrightarrow P^{-1}AP(P^{-1}\mathbf{x}) = \lambda P^{-1}\mathbf{x} \Leftrightarrow$$

$$B\mathbf{y} = \lambda\mathbf{y}, \text{ όπου } \mathbf{y} = P^{-1}\mathbf{x} \Rightarrow \mathbf{x} = P\mathbf{y}. \quad \blacksquare$$

Θεώρημα 3.5 Αν οι τετραγωνικοί πίνακες A, B είναι της ίδιας τάξης, οι πίνακες AB και BA έχουν την ίδια χαρακτηριστική εξίσωση.

Θεώρημα 3.6 Ιδιοδιανύσματα $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ αντίστοιχα σε διαφορετικές μεταξύ τους ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_k$ του πίνακα A είναι γραμμικά ανεξάρτητα.

Για την απόδειξη των παραπάνω θεωρημάτων, βλέπε για παράδειγμα Δ.Δασκαλόπουλος, [9], σελ. 206 και Γ. Δονάτος-Μ. Αδάμ, σελ. 337.

3.4. Χαρακτηριστικά μεγέθη πινάκων ειδικής μορφής.

Θεώρημα 3.7 **i)** Οι ιδιοτιμές ενός ερμιτιανού (συμμετρικού) πίνακα είναι πραγματικοί αριθμοί.

ii) Οι ιδιοτιμές ενός αντιερμιτιανού (αντισυμμετρικού) πίνακα είναι καθαρά φανταστικοί αριθμοί ή μηδέν.

iii) Οι ιδιοτιμές ενός ορθομοναδιαίου (ορθογώνιου) πίνακα είναι αριθμοί μέτρου ένα.

Απόδειξη. Έστω λ, \mathbf{x} δύο αντίστοιχα χαρακτηριστικά μεγέθη. Τότε $A\mathbf{x} = \lambda\mathbf{x}$.

i), ii) Αν ο A είναι ερμιτιανός ή αντιερμιτιανός, τότε έχουμε:

$$\begin{aligned} A\mathbf{x} = \lambda\mathbf{x} &\Rightarrow (A\mathbf{x})^* = (\lambda\mathbf{x})^* \Rightarrow \mathbf{x}^* A^* = \bar{\lambda}\mathbf{x}^* \Rightarrow \pm \mathbf{x}^* A = \bar{\lambda}\mathbf{x}^* \\ &\Rightarrow \pm \mathbf{x}^* A\mathbf{x} = \bar{\lambda}\mathbf{x}^* \mathbf{x} \Rightarrow \pm \mathbf{x}^* \lambda\mathbf{x} = \bar{\lambda}\mathbf{x}^* \mathbf{x} \Rightarrow (\lambda \mp \bar{\lambda})\mathbf{x}^* \mathbf{x} = 0. \end{aligned}$$

Επειδή $\mathbf{x} \neq \mathbf{0}$, θα έχουμε $\mathbf{x}^* \mathbf{x} = \|\mathbf{x}\|_2^2 > 0$ και, επομένως, $\lambda \mp \bar{\lambda} = 0$.

Άρα, οι προτάσεις (i), (ii) είναι αληθείς.

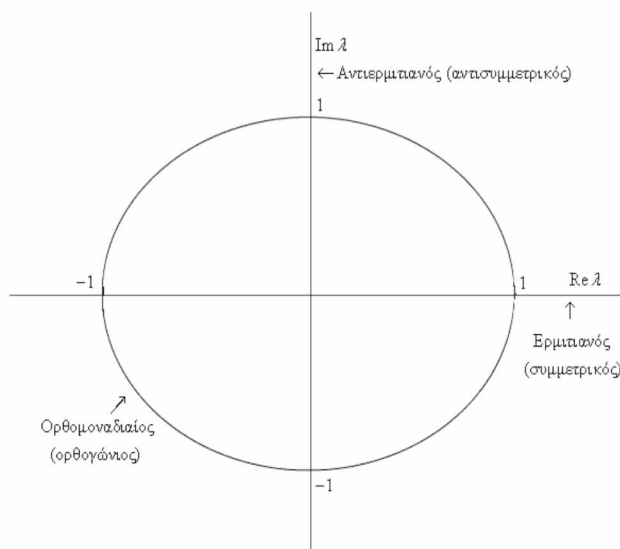
iii) Αν A είναι ορθομοναδιαίος πίνακας (βλέπε παράγραφο 1.7.3), τότε

$$A\mathbf{x} = \lambda\mathbf{x} \Rightarrow \mathbf{x}^* A^* = \bar{\lambda}\mathbf{x}^* \Rightarrow \mathbf{x}^* A^* A\mathbf{x} = \bar{\lambda}\mathbf{x}^* A\mathbf{x} \Rightarrow \mathbf{x}^* \mathbf{x} = \bar{\lambda}\lambda\mathbf{x}^* \mathbf{x} \Rightarrow (1 - \bar{\lambda}\lambda)\mathbf{x}^* \mathbf{x} = 0,$$

και επειδή $\mathbf{x}^* \mathbf{x} \neq 0$, από την προηγούμενη ισότητα συμπεραίνεται

$$\bar{\lambda}\lambda = 1 \Rightarrow |\lambda|^2 = 1. \quad \blacksquare$$

Η γεωμετρική ερμηνεία των ιδιοτήτων που αποδείχθηκαν στο θεώρημα 3.7 απεικονίζεται στο σχήμα που ακολουθεί.



Θεώρημα 3.8 Αν λ_1, λ_2 είναι δύο διαφορετικές χαρακτηριστικές τιμές ενός ερμιτιανού (συμμετρικού), αντιερμιτιανού (αντισυμμετρικού) ή ορθομοναδιαίου (ορθογώνιου) πίνακα A , τότε δύο αντίστοιχα προς αυτές ιδιοδιανύσματα $\mathbf{x}_1, \mathbf{x}_2$ είναι ορθογώνια.

Απόδειξη. Από την υπόθεση, έχουμε $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$ και $A\mathbf{x}_2 = \lambda_2\mathbf{x}_2$.

Από την πρώτη ισότητα, παίρνουμε τη σχέση:

$$\mathbf{x}_1^* A^* = \bar{\lambda}_1 \mathbf{x}_1^* \quad (1)$$

□ Αν ο πίνακας A είναι ερμιτιανός, τότε $A^* = A$ και $\bar{\lambda}_1 = \lambda_1$, οπότε από την (1) παίρνουμε $\mathbf{x}_1^* A = \lambda_1 \mathbf{x}_1^*$. Πολλαπλασιάζοντας αυτήν από δεξιά επί \mathbf{x}_2 , βρίσκουμε:

$$\mathbf{x}_1^* A \mathbf{x}_2 = \lambda_1 \mathbf{x}_1^* \mathbf{x}_2 \Rightarrow \mathbf{x}_1^* \lambda_2 \mathbf{x}_2 = \lambda_1 \mathbf{x}_1^* \mathbf{x}_2 \Rightarrow (\lambda_2 - \lambda_1) \mathbf{x}_1^* \mathbf{x}_2 = 0. \quad (2)$$

Επειδή $\lambda_2 \neq \lambda_1$, από την τελευταία ισότητα στην (2) προκύπτει ότι $\mathbf{x}_1^* \mathbf{x}_2 = 0$, δηλαδή, τα ιδιοδιανύσματα $\mathbf{x}_1, \mathbf{x}_2$ είναι ορθογώνια.

□ Αν ο πίνακας A είναι αντιερμιτιανός, τότε $A^* = -A$, $\bar{\lambda}_1 = -\lambda_1$, οπότε, ακολουθώντας την ίδια διαδικασία φτάνουμε πάλι στην (2).

□ Αν ο πίνακας A είναι ορθομοναδιαίος, τότε $A^* A = I$ και $\bar{\lambda}_1 \lambda_1 = 1 \Rightarrow \bar{\lambda}_1 = \frac{1}{\lambda_1}$.

Πολλαπλασιάζοντας την (1) κατά μέλη προς τα δεξιά με την $A\mathbf{x}_2 = \lambda_2\mathbf{x}_2$, βρίσκουμε

$$\mathbf{x}_1^* A^* A \mathbf{x}_2 = \bar{\lambda}_1 \mathbf{x}_1^* \lambda_2 \mathbf{x}_2.$$

Από την παραπάνω προκύπτει:

$$\mathbf{x}_1^* \mathbf{x}_2 = \frac{\lambda_2}{\lambda_1} \mathbf{x}_1^* \mathbf{x}_2 \Rightarrow (\lambda_2 - \lambda_1) \mathbf{x}_1^* \mathbf{x}_2 = 0,$$

από την οποία προκύπτει πάλι ότι $\mathbf{x}_1^* \mathbf{x}_2 = 0$. ■

Θεώρημα 3.9 Κάθε ερμιτιανός (συμμετρικός) πίνακας n -τάξης έχει ακριβώς n γραμμικά ανεξάρτητα ιδιοδιανύσματα.

Θεώρημα 3.10 Για κάθε ερμιτιανό (συμμετρικό) πίνακα n -τάξης, υπάρχει ένα σύνολο, n το πλήθος, ορθοκανονικών ιδιοδιανυσμάτων.

3.5. Χαρακτηριστικά μεγέθη ρητών συναρτήσεων.

Θεώρημα 3.11 Αν είναι λ, \mathbf{x} δύο χαρακτηριστικά μεγέθη του πίνακα A , τότε είναι λ^k, \mathbf{x} δύο επίσης αντίστοιχα χαρακτηριστικά μεγέθη του πίνακα A , για k θετικό ακέραιο. Τα παραπάνω ισχύουν και για αρνητικό ακέραιο k , όταν ο πίνακας A^k είναι αντιστρέψιμος.

Απόδειξη. Από την υπόθεση, $A\mathbf{x} = \lambda\mathbf{x}$. Άρα η σχέση $A^k\mathbf{x} = \lambda^k\mathbf{x}$ ισχύει για $k=1$.

Υποθέτουμε ότι ισχύει και για το φυσικό αριθμό k . Οπότε

$$A(A^k\mathbf{x}) = \lambda^k A\mathbf{x} \Rightarrow A^{k+1}\mathbf{x} = \lambda^k (\lambda\mathbf{x}) \Rightarrow A^{k+1}\mathbf{x} = \lambda^{k+1}\mathbf{x}.$$

Δηλαδή, η σχέση ισχύει για κάθε φυσικό αριθμό k .

Αν ο πίνακας A είναι αντιστρέψιμος, τότε $\lambda \neq 0$ και $A^{-1}A\mathbf{x} = \lambda A^{-1}\mathbf{x} \Rightarrow A^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}$.

Αν υποθέσουμε ότι ισχύει για έναν αρνητικό ακέραιο k και πολλαπλασιάσουμε αυτή προς τα αριστερά επί A^{-1} , αποδεικνύουμε, όπως και προηγούμενα, ότι ισχύει για κάθε αρνητικό ακέραιο. ■

Θεώρημα 3.12 Αν είναι λ, \mathbf{x} δύο χαρακτηριστικά μεγέθη του πίνακα A και $f(A)$ μία πολυωνυμική συνάρτησή του, τότε είναι $f(\lambda), \mathbf{x}$ δύο επίσης αντίστοιχα χαρακτηριστικά μεγέθη του πίνακα $f(A)$.

Απόδειξη. Αν είναι $f(A) = a_0 A^k + a_1 A^{k-1} + \dots + a_k I$, τότε, από το προηγούμενο θεώρημα θα έχουμε:

$$\begin{aligned} f(A)\mathbf{x} &= a_0 A^k \mathbf{x} + a_1 A^{k-1} \mathbf{x} + \dots + a_k \mathbf{x} \\ &= a_0 \lambda^k \mathbf{x} + a_1 \lambda^{k-1} \mathbf{x} + \dots + a_k \mathbf{x} \\ &= (a_0 \lambda^k + a_1 \lambda^{k-1} + \dots + a_k) \mathbf{x} \\ &= f(\lambda) \mathbf{x} \end{aligned}$$

$$\Rightarrow f(A)\mathbf{x} = f(\lambda)\mathbf{x}. \quad \blacksquare$$

Θεώρημα 3.13 Αν είναι $\lambda_1, \lambda_2, \dots, \lambda_n$ οι ιδιοτιμές του πίνακα A και $f(A)$ μία πολυωνυμική συνάρτησή του, τότε είναι:

$$\det(f(A)) = f(\lambda_1)f(\lambda_2)\cdots f(\lambda_n).$$

Θεώρημα 3.14 Αν είναι $\lambda_1, \lambda_2, \dots, \lambda_n$ οι ιδιοτιμές του πίνακα A και $f(A) = \frac{g(A)}{h(A)}$ είναι μία ρητή συνάρτησή του, τότε είναι:

$$\det(f(A)) = f(\lambda_1)f(\lambda_2)\cdots f(\lambda_n).$$

Απόδειξη. Από τον ορισμό (βλέπε (1.10)), έχουμε $f(A) = g(A)h^{-1}(A)$. Επειδή είναι

$$\det(h^{-1}(A)) = \frac{1}{\det(h(A))},$$

θα έχουμε

$$\det(f(A)) = \det(g(A)) \cdot \det(h^{-1}(A)) = \frac{\det(g(A))}{\det(h(A))},$$

οπότε, από το προηγούμενο θεώρημα θα έχουμε:

$$\det(f(A)) = \frac{g(\lambda_1)g(\lambda_2)\cdots g(\lambda_n)}{h(\lambda_1)h(\lambda_2)\cdots h(\lambda_n)} = f(\lambda_1)f(\lambda_2)\cdots f(\lambda_n). \quad \blacksquare$$

Θεώρημα 3.15 Αν είναι $\lambda_1, \lambda_2, \dots, \lambda_n$ οι ιδιοτιμές του πίνακα A και $f(A) = \frac{g(A)}{h(A)}$ είναι μία ρητή συνάρτησή του, τότε οι ιδιοτιμές του πίνακα $f(A)$ είναι $f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)$.

Θεώρημα 3.16 Αν είναι λ, \mathbf{x} δύο χαρακτηριστικά μεγέθη του πίνακα A , τότε $f(\lambda), \mathbf{x}$ είναι δύο επίσης αντίστοιχα χαρακτηριστικά μεγέθη της ρητής συνάρτησης

$$f(A) = \frac{g(A)}{h(A)}.$$

Απόδειξη. Σύμφωνα με το θεώρημα 3.12 θα έχουμε:

$$g(A)\mathbf{x} = g(\lambda)\mathbf{x}, \quad h(A)\mathbf{x} = h(\lambda)\mathbf{x}.$$

Επειδή ο πίνακας $h(A)$ είναι αντιστρέψιμος, από τη δεύτερη σχέση και από το θεώρημα 3.11 προκύπτει:

$$h^{-1}(A)\mathbf{x} = h^{-1}(\lambda)\mathbf{x}.$$

Άρα

$$\begin{aligned}
f(A)\mathbf{x} &= \frac{g(A)}{h(A)}\mathbf{x} = g(A)h^{-1}(A)\mathbf{x} \\
&= g(A)h^{-1}(\lambda)\mathbf{x} \\
&= h^{-1}(\lambda)g(A)\mathbf{x} \\
&= h^{-1}(\lambda)g(\lambda)\mathbf{x} \\
&= \frac{g(\lambda)}{h(\lambda)}\mathbf{x} = f(\lambda)\mathbf{x}. \quad \blacksquare
\end{aligned}$$

3.6. Αναγωγή πίνακα σε απλούστερη μορφή.

Αν ο τετραγωνικός πίνακας A είναι όμοιος με τον πίνακα B , τότε, σύμφωνα με τον ορισμό 1.6, υπάρχει αντιστρέψιμος πίνακας X , τέτοιος ώστε να είναι:

$$A = XBX^{-1}$$

και λέμε ότι ο πίνακας A προκύπτει από τον B από ένα **μετασχηματισμό ομοιότητας**. Ειδικά αν ο πίνακας X είναι ορθομοναδιαίος, τότε $X^* = X^{-1}$, οπότε η παραπάνω σχέση γράφεται:

$$A = XBX^*$$

και λέμε ότι ο πίνακας A προκύπτει από τον B από έναν **ορθομοναδιαίο μετασχηματισμό**.

Επιπλέον, σε ένα μετασχηματισμό ομοιότητας ή σε έναν ορθομοναδιαίο μετασχηματισμό, επειδή ο πίνακας X είναι αντιστρέψιμος, σύμφωνα με το θεώρημα 2.9 ο βαθμός του πίνακα A ισούται με το βαθμό του πίνακα B . Αν, λοιπόν, ο πίνακας B είναι διαγώνιος ή τριγωνικός, ο βαθμός του υπολογίζεται ευκολότερα, λόγω της απλούστερης μορφής του, οπότε και ο βαθμός του A . Από αυτό διαφαίνεται η θεωρητική και πρακτική σημασία της αναγωγής ενός πίνακα A , από ένα μετασχηματισμό ομοιότητας, σε έναν πίνακα B , διαγώνιο ή τριγωνικό, για τη μελέτη των συμφυών ιδιοτήτων του.

Για το μετασχηματισμό ενός τετραγωνικού πίνακα n -τάξης σε διαγώνιο ή τριγωνική μορφή ισχύουν οι παρακάτω χρήσιμες προτάσεις:

Θεώρημα 3.17 Ένας πίνακας A , τάξης n είναι όμοιος με διαγώνιο πίνακα, αν και μόνο αν έχει n το πλήθος γραμμικά ανεξάρτητα ιδιοδιανύσματα.

Η σχέση ομοιότητας είναι της μορφής:

$$A = X \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) X^{-1}, \quad (3.4)$$

όπου $\lambda_1, \lambda_2, \dots, \lambda_n$ οι ιδιοτιμές του πίνακα A , οι οποίες δεν είναι αναγκαστικά όλες διαφορετικές μεταξύ τους και $X = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]$ ο πίνακας με στήλες τα αντίστοιχα προς αυτές γραμμικά ανεξάρτητα ιδιοδιανύσματα.

Για κάθε πολυωνυμική ή και ρητή συνάρτηση $f(A)$ του πίνακα A , για τον οποίον ισχύει ο προηγούμενος τύπος, ισχύει επίσης και ο τύπος

$$f(A) = X \text{diag}(f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)) X^{-1}. \quad (3.5)$$

Αν έχουμε μία σχέση της μορφής (3.4), τότε λέμε ότι ο πίνακας A **διαγωνοποιείται** με τον πίνακα X ή ακόμα ότι ο πίνακας X **διαγωνοποιεί** τον A .

Θεώρημα 3.18 Ένας πίνακας A είναι ερμιτιανός (συμμετρικός), αν και μόνο αν είναι ορθομοναδιαία (ορθογώνια) όμοιος με διαγώνιο πραγματικό πίνακα. Η σχέση ομοιότητας είναι της μορφής:

$$A = X \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) X^*$$

με τους ίδιους συμβολισμούς, όπως προηγούμενα, όπου τώρα τα $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ αποτελούν ένα ορθομοναδιαίο σύστημα ιδιοδιανυσμάτων.

Θεώρημα 3.19 Ένας πίνακας είναι συμμετρικός, αν και μόνο αν είναι ισότιμος με διαγώνιο πραγματικό πίνακα. Για ένα συμμετρικό πίνακα A , βαθμού p , η σχέση ισοτιμίας είναι της μορφής:

$$S^T A S = \text{diag}(d_1, \dots, d_p, 0, \dots, 0).$$

Ο παραπάνω τύπος, με κατάλληλο μετασχηματισμό μπορεί να πάρει τη μορφή:

$$Z^T A Z = \begin{bmatrix} I_p & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}.$$

Αν τα $k > 0$ διαγώνια στοιχεία από τα d_1, \dots, d_p είναι θετικά και τα υπόλοιπα αρνητικά, τότε έχουμε ακόμη τον τύπο

$$N^T A N = \begin{bmatrix} I_k & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & -I_{p-k} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} \end{bmatrix}.$$

Το θεώρημα 3.19 και οι παραπάνω τύποι ισχύουν και για ερμιτιανούς πίνακες, όπου το σύμβολο T αντικαθίσταται από το $*$.

Θεώρημα 3.20 Για κάθε τετραγωνικό πίνακα A υπάρχει ορθομοναδιαίος πίνακας U , τέτοιος ώστε ο πίνακας U^*AU να είναι άνω τριγωνικός, με τις ιδιοτιμές του A ως στοιχεία της διαγωνίου του.

Θεώρημα 3.21 Για κάθε τετραγωνικό πίνακα A υπάρχει αντιστρέψιμος πίνακας P , τέτοιος ώστε ο πίνακας $P^{-1}AP$ να είναι άνω τριγωνικός, με τις ιδιοτιμές του A ως διαγώνια στοιχεία του.

Θεώρημα 3.22 Ένας πίνακας είναι κανονικός, αν και μόνο αν είναι ορθομοναδιαία όμοιος με διαγώνιο πίνακα.

Για την απόδειξη των παραπάνω θεωρημάτων 3.20-3.22, βλέπε για παράδειγμα Δ. Δασκαλόπουλος, [9], σελ. 225-234 και Γ. Δονάτος-Μ. Αδάμ, σελ. 388-398.

3.7. Χρήση της διαγωνοποίησης στη λύση προβλημάτων.

Υπάρχει μία στρατηγική για τη λύση προβλημάτων, τα οποία συναντάμε συχνά, συνεπώς αξίζει να δώσουμε προσοχή στην τεχνική της μεθόδου της διαγωνοποίησης.

Εφαρμογή 3.1 Από τη σχέση (3.5) αν $f(A) = A^k$, με k θετικό ακέραιο, τότε έχουμε

$$A^k = P \text{diag}(\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k) P^{-1}. \quad (3.6)$$

Αν ο πίνακας A είναι αντιστρέψιμος, τότε ο παραπάνω τύπος ισχύει και για k αρνητικό ακέραιο.

Παράδειγμα 3.2. Να υπολογιστεί ο πίνακας A^{1010} , αν $A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$.

Λύση. Βρίσκουμε τα χαρακτηριστικά μεγέθη του πίνακα A , που είναι:

$$\lambda_1 = 5, \mathbf{x}_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T \quad \text{και} \quad \lambda_2 = -2, \mathbf{x}_2 = \begin{bmatrix} 4 & -3 \end{bmatrix}^T.$$

Άρα

$$A = X \text{diag}(5, -2) X^{-1},$$

όπου

$$X = \begin{bmatrix} 1 & 4 \\ 1 & -3 \end{bmatrix}$$

με $X^{-1} = -\frac{1}{7} \begin{bmatrix} -3 & -4 \\ -1 & 1 \end{bmatrix}$ και έτσι

$$\begin{aligned} A^{1010} &= \frac{1}{7} \begin{bmatrix} 1 & 4 \\ 1 & -3 \end{bmatrix} \text{diag}(5^{1010}, 2^{1010}) \begin{bmatrix} 3 & 4 \\ 1 & -1 \end{bmatrix} \\ \Rightarrow A^{1010} &= \frac{1}{7} \begin{bmatrix} 5^{1010} & 2^{1012} \\ 5^{1010} & -3(2^{1012}) \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{7} \begin{bmatrix} 3 \cdot 5^{1010} + 2^{1012} & 4 \cdot 5^{1010} - 2^{1012} \\ 3 \cdot 5^{1010} - 3 \cdot 2^{1012} & 4 \cdot 5^{1010} + 3 \cdot 2^{1012} \end{bmatrix}. \quad \blacksquare \end{aligned}$$

Εφαρμογή 3.2 Να λυθεί η εξίσωση

$$f(X) = A, \quad (3.7)$$

όπου το πρώτο μέλος είναι δοθείσα πολυωνυμική συνάρτηση του αγνώστου X και το δεύτερο μέλος γνωστός τετραγωνικός πίνακας τάξης n .

Στην περίπτωση, κατά την οποία ο πίνακας A είναι όμοιος με διαγώνιο πίνακα, η εξίσωση αυτή επιλύεται ως εξής:

Έστω ότι είναι

$$A = P \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) P^{-1} \quad (3.8)$$

και έστω t_i μία ρίζα της εξίσωσης $f(t_i) = \lambda_i$. Για $i = 1, 2, \dots, n$ θα έχουμε

$$f(t_1) = \lambda_1, f(t_2) = \lambda_2, \dots, f(t_n) = \lambda_n. \quad (3.9)$$

Ο πίνακας

$$X = P \text{diag}(t_1, t_2, \dots, t_n) P^{-1} \quad (3.10)$$

είναι λύση της εξίσωσης (3.7).

Πράγματι, λαμβάνοντας υπ' όψη τον τύπο (3.5), έχουμε

$$f(X) = P \text{diag}(f(t_1), f(t_2), \dots, f(t_n)) P^{-1} = P \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) P^{-1} = A.$$

Αν ο βαθμός του πολυωνύμου $f(t)$ είναι k , τότε κάθε μία από τις ρίζες t_i μπορεί να εκλεγεί, εν γένει, κατά k διαφορετικούς τρόπους, οπότε μπορεί να κατασκευάσουμε, γενικά, k^n λύσεις της μορφής (3.10).

Ειδικά για την εξίσωση

$$X^2 = A \quad (3.11)$$

θα δείξουμε ότι για κάθε πίνακα A , με n διαφορετικές μεταξύ τους ιδιοτιμές, έχει ακριβώς 2^n ή 2^{n-1} διαφορετικές μεταξύ τους λύσεις, αντίστοιχα όταν ο πίνακας A είναι αντιστρέψιμος ή μη αντιστρέψιμος.

Πράγματι, αν είναι $\lambda_1, \lambda_2, \dots, \lambda_n$ οι διαφορετικές μεταξύ τους ιδιοτιμές του πίνακα A , συνδυάζοντας το θεώρημα 3.6 με το θεώρημα 3.17, προκύπτει ότι ισχύει μία σχέση της μορφής (3.8). Επομένως, η εξίσωση (3.11) έχει τις λύσεις

$$X = P \operatorname{diag}(\pm\sqrt{\lambda_1}, \pm\sqrt{\lambda_2}, \dots, \pm\sqrt{\lambda_n}) P^{-1} \quad (3.12)$$

για κάθε επιλογή προσήμων.

▫ Εάν ο πίνακας A είναι αντιστρέψιμος, όλες οι ιδιοτιμές του είναι διάφορες του μηδενός, οπότε οι 2^n το πλήθος δυνατές διατάξεις προσήμων οδηγούν σε ισάριθμες διαφορετικές μεταξύ τους λύσεις της εξίσωσης (3.11).

▫ Εάν ο πίνακας A είναι μη αντιστρέψιμος, τότε μία και μάλιστα ακριβώς μία από τις ιδιοτιμές του είναι μηδέν, επειδή από την υπόθεση αυτές είναι διαφορετικές μεταξύ τους. Άρα, ένα μόνο από τα διαγώνια στοιχεία $\pm\sqrt{\lambda_i}$ θα είναι μηδέν, επομένως αντί για δύο θα έχουμε μόνο μία επιλογή. Οπότε από την (3.12) δεν υπάρχουν άλλες λύσεις της εξίσωσης (3.11).

Πράγματι, αν X είναι μία οποιαδήποτε λύση της εξίσωσης αυτής, θα έχουμε

$$(P^{-1}XP)^2 = P^{-1}X^2P = P^{-1}AP = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Αν θέσουμε $B = P^{-1}XP$ και $D = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, τότε η προηγούμενη σχέση γράφεται

$$B^2 = D.$$

Κάνοντας αντικατάσταση τον πίνακα B^2 στις επόμενες ισότητες $B^3 = B^2B = BB^2$, προκύπτει ότι $DB = BD$, δηλαδή, οι πίνακες D, B αντιμετατίθενται. Επειδή όμως ο πίνακας D είναι διαγώνιος, (βλέπε ιδιότητες πράξεων 1.2.5), τότε και ο πίνακας B είναι διαγώνιος, οπότε από την $B^2 = D$, θα έχουμε

$$B = \operatorname{diag}(\pm\sqrt{\lambda_1}, \pm\sqrt{\lambda_2}, \dots, \pm\sqrt{\lambda_n}).$$

Από τη σχέση αυτή και τη σχέση $B = P^{-1}XP$, προκύπτει ότι ο πίνακας X είναι της μορφής (3.12). ■

Σημείωση : Όταν οι ιδιοτιμές του πίνακα A δεν είναι όλες διαφορετικές μεταξύ τους, τότε μία ποικιλία περιπτώσεων προκύπτει ως προς τις λύσεις της εξίσωσης (3.11), για παράδειγμα από τις εξισώσεις

$$X^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad X^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Εύκολα διαπιστώνουμε ότι η πρώτη δεν έχει λύση, ενώ η δεύτερη επαληθεύεται από τον πίνακα

$$X = \begin{bmatrix} \lambda & 1 \\ 1-\lambda^2 & -\lambda \end{bmatrix},$$

όπου λ οποιοσδήποτε αριθμός.

Επεκτείνοντας τη σχετική ορολογία των αριθμών, κάθε λύση της εξίσωσης (3.11) λέγεται **τετραγωνική ρίζα** του πίνακα A .

Παράδειγμα 3.3. Να λυθεί η εξίσωση

$$X^2 + 6X + 9I = \begin{bmatrix} -5 & 9 \\ -6 & 10 \end{bmatrix}.$$

Λύση. Τα χαρακτηριστικά μεγέθη του πίνακα του δευτέρου μέλους είναι:

$$\lambda_1 = 1, \mathbf{x}_1 = [3 \ 2]^T \text{ και } \lambda_2 = 4, \mathbf{x}_2 = [1 \ 1]^T,$$

οπότε

$$P = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \text{ και } P^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}.$$

Οι ρίζες της εξίσωσης $t^2 + 6t + 9 = 1$ είναι $t_1 = -2, t_2 = -4$ και της εξίσωσης $t^2 + 6t + 9 = 4$ είναι $t_1 = -1, t_2 = -5$. Συνδυάζοντας αυτές ανά δύο, από τον τύπο (3.10) βρίσκουμε τις λύσεις:

$$X_1 = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} -4 & 3 \\ -2 & 1 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 4 & -9 \\ 6 & -11 \end{bmatrix}$$

$$X_3 = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -4 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} -10 & 9 \\ -6 & 5 \end{bmatrix}$$

$$X_4 = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} -2 & -3 \\ 2 & -7 \end{bmatrix}$$

Άλλες λύσεις εκτός των παραπάνω δεν υπάρχουν, γιατί αν θέσουμε $X + 3I = Y$, η δοθείσα γράφεται:

$$Y^2 = \begin{bmatrix} -5 & 9 \\ -6 & 10 \end{bmatrix},$$

η οποία, όπως αποδείξαμε, έχει ακριβώς τέσσερις λύσεις. ■

Παράδειγμα 3.4. Να λυθεί η εξίσωση:

$$X^2 = \begin{bmatrix} 11 & -4 & -7 \\ 7 & -2 & -5 \\ 10 & -4 & -6 \end{bmatrix}.$$

Λύση. Τα χαρακτηριστικά μεγέθη του πίνακα του δευτέρου μέλους είναι:

$$\lambda_1 = 0, \quad \mathbf{x}_1 = [1 \ 1 \ 1]^T$$

$$\lambda_2 = 1, \quad \mathbf{x}_2 = [1 \ -1 \ 2]^T$$

$$\lambda_3 = 2, \quad \mathbf{x}_3 = [2 \ 1 \ 2]^T$$

Από τις τιμές αυτές υπολογίζουμε τους πίνακες P και P^{-1} :

$$P = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \\ 1 & 2 & 2 \end{bmatrix},$$

$$\text{με } \det P = \begin{vmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \\ 1 & 2 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 \\ 0 & -2 & -1 \\ 0 & 1 & 0 \end{vmatrix} = 1, \text{ και}$$

$$\begin{aligned} P^{-1} &= \frac{1}{\det P} \text{adj}P = \begin{bmatrix} P_{11} & P_{21} & P_{31} \\ P_{12} & P_{22} & P_{32} \\ P_{13} & P_{23} & P_{33} \end{bmatrix} \\ &= \begin{bmatrix} \begin{vmatrix} -1 & 1 \\ 2 & 2 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} \\ -\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} \\ \begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} \end{bmatrix} \\ &= \begin{bmatrix} -4 & 2 & 3 \\ -1 & 0 & 1 \\ 3 & -1 & -2 \end{bmatrix}. \end{aligned}$$

Λαμβάνοντας υπ' όψη την (3.12) βρίσκουμε όλες τις δυνατές λύσεις ($2^{3-1} = 2^2 = 4$) της δοθείσας εξίσωσης,

$$X = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm\sqrt{2} \end{bmatrix} \begin{bmatrix} -4 & 2 & 3 \\ -1 & 0 & 1 \\ 3 & -1 & -2 \end{bmatrix}.$$

Συνδυάζοντας κατά τους τέσσερις δυνατούς τρόπους τα πρόσημα των διαγώνιων στοιχείων, βρίσκουμε τελικά τις εξισώσεις:

$$X_1 = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} -4 & 2 & 3 \\ -1 & 0 & 1 \\ 3 & -1 & -2 \end{bmatrix} = \begin{bmatrix} -1+6\sqrt{2} & -2\sqrt{2} & 1-4\sqrt{2} \\ 1+3\sqrt{2} & -\sqrt{2} & -1-2\sqrt{2} \\ -2+6\sqrt{2} & -2\sqrt{2} & 2-4\sqrt{2} \end{bmatrix}$$

$$X_2 = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} -4 & 2 & 3 \\ -1 & 0 & 1 \\ 3 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 1+6\sqrt{2} & -2\sqrt{2} & -1-4\sqrt{2} \\ -1+3\sqrt{2} & -\sqrt{2} & 1-2\sqrt{2} \\ 2+6\sqrt{2} & -2\sqrt{2} & -2-4\sqrt{2} \end{bmatrix}$$

$$X_3 = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} -4 & 2 & 3 \\ -1 & 0 & 1 \\ 3 & -1 & -2 \end{bmatrix} = \begin{bmatrix} -1-6\sqrt{2} & 2\sqrt{2} & 1+4\sqrt{2} \\ 1-3\sqrt{2} & \sqrt{2} & -1+2\sqrt{2} \\ -2-6\sqrt{2} & 2\sqrt{2} & 2+4\sqrt{2} \end{bmatrix}$$

$$X_4 = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} -4 & 2 & 3 \\ -1 & 0 & 1 \\ 3 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 1-6\sqrt{2} & 2\sqrt{2} & -1+4\sqrt{2} \\ -1-3\sqrt{2} & \sqrt{2} & 1+2\sqrt{2} \\ 2-6\sqrt{2} & 2\sqrt{2} & -2+4\sqrt{2} \end{bmatrix} \quad \blacksquare$$

3.8. Θεώρημα Cayley-Hamilton.

Το σύνολο όλων των τετραγωνικών πινάκων n -τάξης αποτελεί ένα διανυσματικό χώρο διάστασης $k = n^2$. Κατόπιν αυτού, για κάθε τετραγωνικό πίνακα A , n -τάξης οι $k+1$ πλήθος πίνακες I, A, A^2, \dots, A^k είναι γραμμικά εξαρτημένοι. Επομένως, υπάρχουν συντελεστές c_0, c_1, \dots, c_k , όχι όλοι μηδέν, τέτοιοι ώστε να ισχύει η σχέση:

$$c_0 I + c_1 A + c_2 A^2 + \dots + c_k A^k = \mathbb{O}.$$

Αυτό σημαίνει ότι για κάθε τετραγωνικό πίνακα A , n -τάξης, υπάρχει βαθμωτό πολυώνυμο

$$c(\lambda) = c_0 + c_1 \lambda + c_2 \lambda^2 + \dots + c_k \lambda^k,$$

για το οποίο ισχύει

$$c(A) = \mathbb{O}.$$

Η επόμενη πρόταση, γνωστή στη βιβλιογραφία ως Θεώρημα Cayley-Hamilton, μας λέει ότι αυτό ισχύει και για το πολυώνυμο

$$\delta(\lambda) = \det(\lambda I - A) = \lambda^n + a_1 \lambda^{n-1} + \dots + a_n \quad (3.13)$$

επομένως, και για το χαρακτηριστικό πολυώνυμο του πίνακα A , το οποίο, ως γνωστό, διαφέρει από αυτό μόνο κατά τον παράγοντα $(-1)^n$.

Θεώρημα 3.23 Για κάθε πίνακα A , με χαρακτηριστική εξίσωση $\delta(\lambda) = 0$, ισχύει η σχέση $\delta(A) = \mathbf{O}$, δηλαδή,

$$A^n + a_1 A^{n-1} + \dots + a_n I = \mathbf{O}. \quad (3.14)$$

Απόδειξη. Βλέπε για παράδειγμα Δ. Δασκαλόπουλος, [9], σελ. 246 και Γ. Δονάτος-Μ. Αδάμ, σελ. 347. ■

Εφαρμογή 3.3 Για κάθε αντιστρέψιμο πίνακα A ισχύει:

$$A^{-1} = -\frac{1}{a_n} \{A^{n-1} + a_1 A^{n-2} + \dots + a_{n-1} I\} \quad (3.15)$$

αν

$$\delta(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \dots + a_n.$$

Απόδειξη. Επειδή ο πίνακας A είναι αντιστρέψιμος, τότε ο σταθερός όρος $a_n \neq 0$, γιατί, διαφορετικά, ο πίνακας θα είχε μία ιδιοτιμή $\lambda = 0$, δηλαδή, αντίθετο προς την υπόθεση ότι ο A είναι αντιστρέψιμος.

Πολλαπλασιάζοντας την ισότητα (3.14) επί A^{-1} και επιλύοντας την προκύπτουσα ως προς A^{-1} , παίρνουμε τη σχέση (3.15). ■

Σημείωση : (i) Λύνοντας την (3.14) ως προς A^n , βρίσκουμε:

$$A^n = -\{a_1 A^{n-1} + a_2 A^{n-2} + \dots + a_{n-1} A + a_n I\}.$$

Από αυτή εύκολα συμπεραίνουμε ότι μπορούμε να εκφράσουμε τη δύναμη A^m ($m \geq n$) κάθε τετραγωνικού πίνακα ως συνάρτηση των μη αρνητικών δυνάμεων του A , που είναι μικρότερες του n .

Όμοια, από την (3.15) βρίσκουμε ότι κάθε αρνητική ακέραια δύναμη του αντιστρέψιμου πίνακα A μπορεί να εκφραστεί ως γραμμικός συνδυασμός των $I, A, A^2, \dots, A^{n-1}$.

(ii) Προκειμένου να πετύχουμε κάτι ανάλογο για μία οποιαδήποτε πολυωνυμική συνάρτηση $f(A)$, διαιρούμε το βαθμωτό πολυώνυμο $f(\lambda)$ δια του χαρακτηριστικού πολυωνύμου ή δια του πολυωνύμου $\delta(\lambda)$ του πίνακα A , οπότε βρίσκουμε ένα πηλίκο $\pi(\lambda)$ και ένα υπόλοιπο $\nu(\lambda)$ βαθμού το πολύ $n-1$, για το οποίο, ως γνωστό, ισχύει η ταυτότητα:

$$f(\lambda) = \delta(\lambda)\pi(\lambda) + \nu(\lambda).$$

Από αυτή και τους τύπους (1.9), προκύπτει η ζητούμενη απλοποιημένη μορφή της πολυωνυμικής συνάρτησης, δηλαδή

$$f(A) = \delta(A)\pi(A) + \nu(A) = \nu(A).$$

(iii) Μία ακόμη εφαρμογή του θεωρήματος Cayley-Hamilton αποτελεί η εύρεση του ίδιου του χαρακτηριστικού πολυωνύμου $\delta(\lambda)$.

Για το σκοπό αυτό, πολλαπλασιάζουμε την ισότητα (3.14) επί ένα αυθαίρετο γνωστό διάνυσμα \mathbf{x} , οπότε προκύπτει η εξίσωση:

$$A^n \mathbf{x} + a_1 (A^{n-1} \mathbf{x}) + \dots + a_{n-1} (A \mathbf{x}) + a_n \mathbf{x} = \mathbf{0} \quad (3.6)$$

Από την εξίσωση αυτή προκύπτουν ισοδύναμα n βαθμωτές εξισώσεις με αγνώστους τους συντελεστές a_1, a_2, \dots, a_n .

Λύνοντας τις εξισώσεις αυτές, προσδιορίζονται οι άγνωστοι συντελεστές, οπότε έχουμε το χαρακτηριστικό πολυώνυμο. Η μέθοδος αυτή για μεγάλες τιμές του n είναι σχεδόν απαγορευτική, εξαιτίας των πράξεων.

Παράδειγμα 3.5. Δίνεται ο αντιστρέψιμος πίνακας

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 1 \\ 2 & 3 & 1 \end{bmatrix}.$$

Να υπολογιστούν οι δυνάμεις A^3, A^4 , ο αντίστροφος A^{-1} και ο A^{-2} .

Λύση. Το χαρακτηριστικό πολυώνυμο του A είναι:

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 & 2 \\ 3 & 1-\lambda & 1 \\ 2 & 3 & 1-\lambda \end{vmatrix} = -\lambda^3 + 3\lambda^2 + 7\lambda + 11.$$

Σύμφωνα με το θεώρημα Cayley-Hamilton, έχουμε

$$-A^3 + 3A^2 + 7A + 11I = \mathbf{O},$$

από όπου προκύπτουν οι ζητούμενοι πίνακες:

$$\begin{aligned} \square \quad A^3 &= 3A^2 + 7A + 11I = 3 \begin{bmatrix} 8 & 8 & 5 \\ 8 & 7 & 8 \\ 13 & 8 & 8 \end{bmatrix} + 7 \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 1 \\ 2 & 3 & 1 \end{bmatrix} + 11 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 42 & 31 & 29 \\ 45 & 39 & 31 \\ 53 & 45 & 42 \end{bmatrix}. \end{aligned}$$

$$\square \quad A^4 = 3A^3 + 7A^2 + 11A$$

$$= 3 \begin{bmatrix} 42 & 31 & 29 \\ 45 & 39 & 31 \\ 53 & 45 & 42 \end{bmatrix} + 7 \begin{bmatrix} 8 & 8 & 5 \\ 8 & 7 & 8 \\ 13 & 8 & 8 \end{bmatrix} + 11 \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 1 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 193 & 160 & 144 \\ 224 & 177 & 160 \\ 272 & 224 & 193 \end{bmatrix}.$$

$$\square \quad \text{Επειδή } -A^2 + 3A + 7I + 11A^{-1} = \mathbf{O} \Rightarrow$$

$$A^{-1} = \frac{1}{11} \{-7I - 3A + A^2\} = \frac{1}{11} \left\{ -7 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - 3 \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 1 \\ 2 & 3 & 1 \end{bmatrix} + \begin{bmatrix} 8 & 8 & 5 \\ 8 & 7 & 8 \\ 13 & 8 & 8 \end{bmatrix} \right\}$$

$$= \frac{1}{11} \begin{bmatrix} -2 & 5 & -1 \\ -1 & -3 & 5 \\ 7 & -1 & -2 \end{bmatrix}.$$

$$\square \quad A^{-2} = \frac{1}{11} \{-7A^{-1} - 3I + A\} = \frac{1}{121} \left\{ -7 \begin{bmatrix} -2 & 5 & -1 \\ -1 & -3 & 5 \\ 7 & -1 & -2 \end{bmatrix} - 33 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + 11 \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 1 \\ 2 & 3 & 1 \end{bmatrix} \right\}$$

$$= \frac{1}{121} \begin{bmatrix} -8 & -24 & 29 \\ 40 & -1 & -24 \\ -27 & 40 & -8 \end{bmatrix}. \quad \blacksquare$$

Παράδειγμα 3.6. Ναδειχτεί ότι ο πίνακας

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

επαληθεύει τη σχέση $A^p = A^{p-2} + A^2 - I$, για $p = 3, 4, \dots$ και κατόπιν να υπολογιστεί ο πίνακας A^{100} .

Λύση. Το χαρακτηριστικό πολυώνυμο του πίνακα A είναι:

$$\begin{vmatrix} 1-\lambda & 0 & 0 \\ 1 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = (1-\lambda)(\lambda^2-1) = -\lambda^3 + \lambda^2 + \lambda - 1.$$

Σύμφωνα με το θεώρημα Cayley-Hamilton, έχουμε:

$$-A^3 + A^2 + A - I = \mathbf{O} \Rightarrow A^3 = A^2 + A - I.$$

Δηλαδή, η αποδεικτέα σχέση ισχύει για $p = 3$.

Έστω ότι ισχύει για το φυσικό αριθμό $p \geq 3$. Πολλαπλασιάζοντας και τα δύο μέλη επί A , βρίσκουμε:

$$A^{p+1} = A^{p-1} + A^3 - A$$

$$\Rightarrow A^{p+1} = A^{p-1} + (A + A^2 - I) - A = A^{p-1} + A^2 - I,$$

δηλαδή, η σχέση επαληθεύεται και για το φυσικό αριθμό $p+1$, επομένως ισχύει για κάθε φυσικό αριθμό $p \geq 3$.

Ο υπολογισμός του πίνακα A^{100} γίνεται ως εξής:

$$\begin{aligned}
A^{100} &= A^{98} + A^2 - I \\
A^{98} &= A^{96} + A^2 - I \\
A^{96} &= A^{94} + A^2 - I \\
&\vdots \\
A^4 &= A^2 + A^2 - I
\end{aligned}$$

Το πλήθος των σχέσεων αυτών είναι $100 : 2 - 1 = 49$. Προσθέτοντας τις σχέσεις αυτές κατά μέλη, βρίσκουμε

$$A^{100} = A^2 + 49(A^2 - I) = 50A^2 - 49I,$$

από όπου υπολογίζουμε

$$A^{100} = 50 \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}^2 - 49 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 50 & 1 & 0 \\ 50 & 0 & 1 \end{bmatrix}. \quad \blacksquare$$

Παράδειγμα 3.7. Να εκφραστεί ο αντίστροφος του πίνακα

$$B = A^4 + 5A^3 - 48A^2 - I$$

ως γραμμική συνάρτηση του πίνακα

$$A = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}.$$

Λύση. Το χαρακτηριστικό πολυώνυμο του A είναι:

$$\begin{vmatrix} 2-\lambda & -5 \\ -1 & 3-\lambda \end{vmatrix} = (2-\lambda)(3-\lambda) - 5 = \lambda^2 - 5\lambda + 1.$$

Διαιρώντας το πολυώνυμο

$$f(\lambda) = \lambda^4 + 5\lambda^3 - 48\lambda^2 - 1$$

με το χαρακτηριστικό πολυώνυμο βρίσκουμε

$$f(\lambda) = \lambda^4 + 5\lambda^3 - 48\lambda^2 - 1 = (\lambda^2 - 5\lambda + 1)(\lambda^2 + 10\lambda + 1) - 5\lambda - 2.$$

Άρα

$$B = f(A) = -5A - 2I = -5 \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -12 & 25 \\ 5 & -17 \end{bmatrix}.$$

Το χαρακτηριστικό πολυώνυμο του πίνακα B είναι:

$$\begin{vmatrix} -12-\lambda & 25 \\ 5 & -17-\lambda \end{vmatrix} = (12+\lambda)(17+\lambda) - 125 = \lambda^2 + 29\lambda + 79.$$

Σύμφωνα με το θεώρημα Cayley-Hamilton, έχουμε:

$$B^2 + 29B + 79I = \mathbf{O}.$$

Πολλαπλασιάζοντας και τα δύο μέλη της τελευταίας ισότητας επί τον πίνακα B^{-1} , έχουμε

$$B + 29I + 79B^{-1} = \mathbf{O},$$

από όπου υπολογίζεται ο αντίστροφος του B να είναι:

$$B^{-1} = -\frac{1}{79}(B + 29I) = -\frac{1}{79}(-5A - 2I + 29I).$$

$$\text{Άρα, } B^{-1} = \frac{5}{79}A - \frac{27}{79}I.$$

■

ΚΕΦΑΛΑΙΟ 4

ΝΟΡΜΕΣ ΚΑΙ ΘΕΩΡΙΑ ΔΙΑΤΑΡΑΧΩΝ

4.1. Χώροι Νορμέ.

Η έννοια της Ευκλείδειας νόρμας $\|\mathbf{x}\|_2$ έχει ήδη δοθεί ως το μέτρο ενός διανύσματος \mathbf{x} . Υπάρχουν όμως πολλές περιπτώσεις στις οποίες είναι σκόπιμο να υπολογίσουμε το μέτρο ενός διανύσματος με άλλους τρόπους. Έτσι εισάγουμε τη γενική έννοια της **νόρμας** ενός διανύσματος.

Ορισμός 4.1 Μία **νόρμα** (**norm**) επί ενός (πραγματικού ή μιγαδικού) διανυσματικού χώρου V είναι μία πραγματική συνάρτηση της οποίας η τιμή για κάθε $\mathbf{x} \in V$ συμβολίζεται με $N(\mathbf{x}) = \|\mathbf{x}\|$, και ικανοποιεί τα αξιώματα:

$$(A_1) \quad \|\mathbf{x}\| \geq 0.$$

$$(A_2) \quad \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}.$$

$$(A_3) \quad \|a\mathbf{x}\| = |a|\|\mathbf{x}\|, \quad a \in \mathbb{R} \text{ ή } a \in \mathbb{C}.$$

$$(A_4) \quad \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \text{ (τριγωνική ανισότητα).}$$

Κάθε διανυσματικός χώρος V μαζί με μία νόρμα επ' αυτού λέγεται **διανυσματικός χώρος νορμέ** ή απλούστερα **χώρος νορμέ** και συμβολίζεται με $(V, \|\cdot\|)$ ή απλούστερα με V εφόσον δεν υπάρχει κίνδυνος σύγχυσης. Ισχύουν μεταξύ άλλων οι ακόλουθες ιδιότητες:

Ιδιότητες νόρμας.

$$\text{i)} \quad \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}.$$

$$\text{ii)} \quad \|\mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_n\| \leq \|\mathbf{x}_1\| + \|\mathbf{x}_2\| + \dots + \|\mathbf{x}_n\|.$$

$$\text{iii)} \quad \left| \|\mathbf{x}\| - \|\mathbf{y}\| \right| \leq \|\mathbf{x} - \mathbf{y}\|.$$

Απόδειξη.

Η (i) είναι το αξίωμα (A₂) και το (ii) προκύπτει επαγωγικά από την τριγωνική ανισότητα (A₄).

iii) Για ένα διάνυσμα $\mathbf{x} \in V$ μπορούμε να γράψουμε

$$\|\mathbf{x}\| = \|\mathbf{x} + \mathbf{y} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\| \Rightarrow \|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|.$$

Όμοια δείχνουμε ότι ισχύει

$$\|\mathbf{y}\| - \|\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{y}\|.$$

Οπότε συνδυάζοντας τις δύο ανισώσεις έχουμε

$$\|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|. \quad \blacksquare$$

Για περισσότερες λεπτομέρειες σχετικά με τους χώρους νορμέ βλέπε οποιοδήποτε εισαγωγικό βιβλίο Συναρτησιακής Ανάλυσης. Βλέπε για παράδειγμα Ε. Kreyszig, σελ.58 και C.Groetsch, σελ. 49.

Οι πιο συχνά χρησιμοποιούμενες νόρμες στο χώρο $V = \mathbb{C}^n$ είναι:

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|, \quad 1\text{-νόρμα}$$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}, \quad 2\text{-νόρμα (Ευκλείδεια νόρμα)}$$

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|, \quad \infty\text{-νόρμα}$$

Οι παραπάνω νόρμες είναι ειδικές περιπτώσεις της

$$\|\mathbf{x}\|_p = \sqrt[p]{\sum_{i=1}^n |x_i|^p}, \quad p\text{-νόρμα με } p \geq 1$$

Η ∞ -νόρμα προκύπτει από την p -νόρμα, για $p \rightarrow \infty$.

Σημείωση : (i) (Ισοδύναμες νόρμες). Δύο νόρμες $N_1(\mathbf{x})$ και $N_2(\mathbf{x})$ επί ενός διανυσματικού χώρου V λέγονται **ισοδύναμες** αν υπάρχουν θετικοί αριθμοί α και β τέτοιοι ώστε

$$\alpha N_1(\mathbf{x}) \leq N_2(\mathbf{x}) \leq \beta N_1(\mathbf{x}).$$

(ii) Δύο νόρμες επί ενός διανυσματικού χώρου πεπερασμένης διαστάσεως είναι ισοδύναμες. Έτσι οι νόρμες $\|\mathbf{x}\|_1, \|\mathbf{x}\|_2, \|\mathbf{x}\|_\infty$ επί του διανυσματικού χώρου $V = \mathbb{C}^n$ είναι ισοδύναμες, βλέπε C. DeVito, σελ. 6.

Ανισότητα Hölder.

Έστω $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$ και $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^T$, τότε

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \sum_{i=1}^n |x_i| |y_i| \leq \left\{ \sum_{i=1}^n |x_i|^p \right\}^{1/p} \left\{ \sum_{i=1}^n |y_i|^q \right\}^{1/q}, \quad (4.1)$$

όπου $p, q \geq 1$ και $\frac{1}{p} + \frac{1}{q} = 1$.

Ανισότητα Minkowski.

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \sum_{i=1}^n |x_i| |y_i| \leq \sqrt[p]{\sum_{i=1}^n |x_i + y_i|^p} \leq \sqrt[p]{\sum_{i=1}^n |x_i|^p} + \sqrt[p]{\sum_{i=1}^n |y_i|^p}, \quad p \geq 1 \quad (4.2)$$

Σημείωση : (i) Η (4.2) είναι η τριγωνική ανισότητα για τη νόρμα p , δηλαδή

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p.$$

(ii) Με αφετηρία τη 2-νόρμα, ορίζουμε το **εσωτερικό** (βαθμωτό) **γινόμενο**:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x} = \sum_{i=1}^n x_i \bar{y}_i.$$

Οπότε για $\mathbf{x} \in V$ έχουμε

$$\|\mathbf{x}\|_2 = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2} = \sqrt{\mathbf{x}^* \mathbf{x}} = \sqrt{\sum_{i=1}^n |x_i|^2}.$$

(iii) Η **ανισότητα Schwartz** είναι η ανισότητα του Hölder, για $p = q = 2$, δηλαδή,

$$|\langle \mathbf{x}, \mathbf{y} \rangle| = |\mathbf{y}^* \mathbf{x}| = \left| \sum_{i=1}^n x_i \bar{y}_i \right| \leq \sum_{i=1}^n |x_i| |y_i| \leq \sqrt{\sum_{i=1}^n |x_i|^2} \cdot \sqrt{\sum_{i=1}^n |y_i|^2} = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2. \quad (4.3)$$

(βλέπε σελ.10)

4.2. Συνεχείς γραμμικές απεικονίσεις.

Μία γραμμική απεικόνιση $T: V \rightarrow U$ του χώρου νορμέ V στο χώρο νορμέ U λέγεται **φραγμένη** αν για κάθε $\mathbf{x} \in V$ υπάρχει σταθερός αριθμός $\alpha > 0$ τέτοιος ώστε να είναι

$$\|T(\mathbf{x})\| < \alpha \|\mathbf{x}\|.$$

Πρόταση 4.1 Έστω $T: V \rightarrow U$ γραμμική απεικόνιση του χώρου νορμέ V στο χώρο νορμέ U . Οι παρακάτω προτάσεις είναι ισοδύναμες:

- (i) Η απεικόνιση T είναι φραγμένη.
- (ii) Η απεικόνιση T είναι συνεχής στο σημείο $\mathbf{x}_0 \in V$.
- (iii) Η απεικόνιση T είναι συνεχής στο V .

Απόδειξη. Για την απόδειξη βλέπε Δ. Δασκαλόπουλος, [10], σελ. 60 και C. Groetsch, σελ. 49. ■

Ο χώρος όλων των φραγμένων γραμμικών απεικονίσεων, που συμβολίζεται με $B(V, U)$, είναι ένας διανυσματικός χώρος νορμέ με νόρμα που ορίζεται από τον τύπο:

$$\|T\| = \sup_{\mathbf{x} \neq 0} \frac{\|T(\mathbf{x})\|}{\|\mathbf{x}\|} = \sup_{\|\mathbf{x}\|=1} \|T(\mathbf{x})\|.$$

4.3. Απόλυτος τιμή και Νόρμα πίνακα.

4.3.1. Απόλυτος τιμή.

Αν $A = [\alpha_{ij}]$ και $B = [\beta_{ij}]$ είναι δύο πίνακες του ίδιου τύπου, τότε η σχέση

$$A \leq B \Leftrightarrow \alpha_{ij} \leq \beta_{ij}$$

ορίζει μία σχέση **διάταξης** στο σύνολο όλων των πινάκων του ίδιου τύπου.

Πράγματι,

- Για κάθε πίνακα A ισχύει: $\alpha_{ij} \leq \alpha_{ij} \Rightarrow A \leq A$.
- Εάν $\alpha_{ij} \leq \beta_{ij}$ και $\beta_{ij} \leq \alpha_{ij} \Rightarrow A = B$.
- Εάν $\alpha_{ij} \leq \beta_{ij}$ και $\beta_{ij} \leq \gamma_{ij} \Rightarrow \alpha_{ij} \leq \gamma_{ij}$, οπότε εάν $A \leq B$ και $B \leq \Gamma \Rightarrow A \leq \Gamma$.

Προφανώς το σύνολο όλων των πινάκων του ίδιου τύπου δεν είναι **ολικά διατεταγμένο**. Δηλαδή, δύο τυχαίοι πίνακες του ίδιου τύπου δεν είναι συγκρίσιμοι μεταξύ τους.

Με τον όρο **απόλυτη τιμή** ενός πίνακα A εννοούμε τον πίνακα με στοιχεία τις απόλυτες τιμές των στοιχείων του A , δηλαδή,

$$|A| = [|\alpha_{ij}|].$$

Αν A και B είναι δύο πίνακες για τους οποίους οι πράξεις $A+B$ και AB έχουν έννοια, τότε

- i) $|A+B| \leq |A| + |B|$
- ii) $|AB| \leq |A| \cdot |B|$
- iii) $|aA| = |a| |A|$, όπου a αριθμός

Ειδικά, για έναν τετραγωνικό πίνακα A ισχύει:

$$|A^p| \leq |A|^p.$$

Βλέπε Β. Demidovich, I. Maron, σελ. 242.

4.3.2. Νόρμα πίνακα $m \times n$.

Το σύνολο όλων των πινάκων τύπου $m \times n$ αποτελούν ένα διανυσματικό χώρο διαστάσεως $m \times n$, βλέπε παράδειγμα 1.4.

Έτσι, μία νόρμα ενός πίνακα A , τύπου $m \times n$, είναι ένας πραγματικός αριθμός, που συμβολίζεται με $\|A\|$ και ικανοποιεί τα αξιώματα $A_1 - A_4$.

Επειδή τα διανύσματα στήλης είναι ένας πίνακας τύπου $n \times 1$ μπορούμε να επεκτείνουμε τις διανυσματικές νόρμες $\|\mathbf{x}\|_1, \|\mathbf{x}\|_2, \|\mathbf{x}\|_\infty$ σε οποιοδήποτε διανυσματικό χώρο πινάκων $\Delta_{m \times n}$ ως εξής:

$$\|A\|_1 = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}| \quad (\text{επέκταση της νόρμας } \|\mathbf{x}\|_1) \quad (\mathbf{N}_1)$$

$$\|A\|_E = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} \quad (\text{Ευκλείδεια νόρμα πίνακα}) \quad (\mathbf{N}_2)$$

$$\|A\|_\infty = \max_{i,j} |a_{ij}| \quad (\text{επέκταση της νόρμας } \|\mathbf{x}\|_\infty) \quad (\mathbf{N}_3)$$

Επειδή ο διανυσματικός χώρος $\Delta_{m \times n}$ είναι πεπερασμένης διαστάσεως όλες οι νόρμες που ορίζονται επ' αυτού είναι ισοδύναμες.

Σημείωση : Αν A, B είναι δύο πίνακες του ίδιου τύπου, τότε από την ιδιότητα (iii) (σελ. 53) συμπεραίνουμε

$$\| \|A\| - \|B\| \| \leq \|A - B\|.$$

Ορισμός 4.2 Μία νόρμα πίνακα λέγεται **κανονική**, αν ισχύουν οι παρακάτω δύο συνθήκες:

(i) Έστω ο πίνακας $A = [\alpha_{ij}]$, τότε

$$|\alpha_{ij}| \leq \|A\|$$

και για τον πίνακα αριθμό $A = [\alpha_{11}]$ έχουμε $\|A\| = |\alpha_{11}|$.

(ii) Από την ανισότητα $|A| \leq |B|$ (όπου A και B είναι πίνακες) συμπεραίνεται η ανισότητα

$$\|A\| \leq \|B\|.$$

Ειδικά, $\|A\| = \|A\|$.

Στη συνέχεια της παραγράφου αυτής περιοριζόμαστε στο διανυσματικό χώρο $\Delta_{n \times n}$.

Σημείωση : Από τον ορισμό των \mathbf{N}_1 , \mathbf{N}_2 , \mathbf{N}_3 άμεσα συμπεραίνεται ότι οι νόρμες αυτές είναι κανονικές. Για παράδειγμα,

$$|\alpha_{ij}| \leq \sum_{i=1}^m \sum_{j=1}^n |\alpha_{ij}| = \|A\|_I.$$

Ειδικά, αν $A = [\alpha_{11}]$, τότε $\|A\|_I = |\alpha_{11}|$.

Εάν $|A| \leq B$, τότε $|\alpha_{ij}| \leq |\beta_{ij}|$ και επομένως

$$\|A\|_I = \sum_{i=1}^m \sum_{j=1}^n |\alpha_{ij}| \leq \sum_{i=1}^m \sum_{j=1}^n |\beta_{ij}| = \|B\|_I.$$

Φυσική νόρμα πίνακα.

Σε πολλά προβλήματα ο πίνακας θεωρείται ως μία γραμμική απεικόνιση $A: V \rightarrow U$, όπου V, U διανυσματικοί χώροι πεπερασμένης διάστασης επί του αυτού σώματος K των πραγματικών ή μιγαδικών αριθμών. Επειδή η απεικόνιση αυτή είναι φραγμένη, η νόρμα αυτής μπορεί να οριστεί σύμφωνα με την παράγραφο 4.2 και μάλιστα από τον τύπο

$$\|A\| = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_U}{\|\mathbf{x}\|_V} = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|_U.$$

Προς διάκριση η νόρμα αυτή λέγεται **φυσική νόρμα του A** που επάγεται από τις νόρμες των χώρων V, U .

4.3.3. Νόρμα τετραγωνικού πίνακα.

Το σύνολο όλων των τετραγωνικών πινάκων τάξης $n \times n$ αποτελούν ένα διανυσματικό χώρο διάστασης $n \times n$. Μία νόρμα ενός τετραγωνικού πίνακα A είναι ένας πραγματικός αριθμός που ικανοποιεί τα αξιώματα $A_1 - A_4$ με μία επιπλέον συνθήκη:

$$\|AB\| \leq \|A\| \cdot \|B\|.$$

Έτσι, καταλήγουμε στον ακόλουθο ορισμό.

Ορισμός 4.3 Μία νόρμα ενός τετραγωνικού πίνακα A , τάξης n , είναι ένας πραγματικός αριθμός, που συμβολίζεται με $\|A\|$ και ικανοποιεί τα αξιώματα:

$$(A_1) \quad \|A\| \geq 0.$$

$$(A_2) \quad \|A\| = 0 \Leftrightarrow A = \mathbf{O}.$$

$$(A_3) \quad \|aA\| = |a|\|A\|, \quad a \in \mathbb{R} \text{ ή } a \in \mathbb{C}. \quad (\text{Ειδικά, } \|-A\| = \|A\|)$$

$$(A_4) \quad \|A+B\| \leq \|A\| + \|B\|.$$

$$(A_5) \quad \|AB\| \leq \|A\| \cdot \|B\|.$$

Σημείωση : (i) Από το αξίωμα (A_5) , συμπεραίνουμε ότι $\|I\| \geq 1$.

Πράγματι,

$$\|A\| = \|AI\| \leq \|A\| \cdot \|I\| \Rightarrow \|I\| \geq 1.$$

(ii) Για έναν τετραγωνικό πίνακα A ισχύει:

$$\|A^p\| \leq \|A\|^p, \quad p \in \mathbb{N}.$$

(iii) Μπορούμε να αποδείξουμε ότι οι νόρμες $N_1 = \|A\|_I$ και $N_2 = \|A\|_E$ είναι νόρμες του διανυσματικού χώρου $\Delta_{n \times n}$.

(iv) Η νόρμα $\|\mathbf{x}\|_\infty$ δεν επεκτείνεται στο διανυσματικό χώρο $\Delta_{n \times n}$, επειδή δεν ισχύει το αξίωμα (A_5) .

Για παράδειγμα, θεωρούμε τους πίνακες $A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, οπότε

$$AB = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix},$$

και

$$\|A\|_I = \|B\|_I = 1,$$

ενώ

$$\|AB\|_I = 2,$$

δηλαδή, το αξίωμα (A_5) δεν ισχύει.

Οπότε η $N_3 = \|A\|_I$ δεν είναι νόρμα του διανυσματικού χώρου $\Delta_{n \times n}$.

Ορισμός 4.4 Μία νόρμα διανύσματος και μία νόρμα πίνακα ονομάζονται **συμβιβαστές**, αν ισχύει

$$\|A\mathbf{x}\| \leq \|A\| \cdot \|\mathbf{x}\|.$$

Σημείωση : Οι νόρμες N_1, N_2 (Ευκλείδεια νόρμα) είναι συμβιβαστές με τις νόρμες $\|\mathbf{x}\|_1$ και $\|\mathbf{x}\|_2$ αντίστοιχα.

Ορισμός 4.5 Έστω $\|\cdot\|_p$ νόρμα διανύσματος, τότε, η **φυσική νόρμα** ενός τετραγωνικού πίνακα, ως προς τη νόρμα του διανύσματος, ορίζεται από την έκφραση:

$$\|A\|_p = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p=1} \|A\mathbf{x}\|_p \quad (4.4)$$

Ιδιότητες φυσικής νόρμας τετραγωνικού πίνακα

i) Κάθε φυσική νόρμα πίνακα είναι συμβιβαστή με την αντίστοιχη νόρμα του διανύσματος. Πράγματι,

$$\|A\mathbf{x}\|_p = \|\mathbf{x}\|_p \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} \leq \|\mathbf{x}\|_p \max_{\|\mathbf{x}\|_p=1} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \|\mathbf{x}\|_p \|A\|_p. \quad (4.5)$$

ii) Για κάθε φυσική νόρμα, έχουμε $\|I\| = 1$.

Πράγματι,

$$\|I\| = \max_{\|\mathbf{x}\|=1} \|I\mathbf{x}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{x}\| = 1. \quad \blacksquare$$

Λήμμα 4.1 Μία φυσική νόρμα, ορισμός 4.5, ικανοποιεί τα αξιώματα της νόρμας πίνακα.

Απόδειξη. Αρκεί να αποδειχθούν τα αξιώματα του ορισμού 4.3.

(**A**₁) Αν $A \neq \mathbf{0}$, υπάρχει $\mathbf{x} \neq \mathbf{0}$, τέτοιο ώστε $A\mathbf{x} \neq \mathbf{0}$. Μπορούμε να κανονικοποιήσουμε το \mathbf{x} , δηλαδή $\|\mathbf{x}\|_p = 1$, και να εξακολουθεί να ισχύει η σχέση $A\mathbf{x} \neq \mathbf{0}$. Οπότε, από την (4.4), συμπεραίνεται $\|A\|_p > 0$.

(**A**₂) Αν $A = \mathbf{0}$, τότε $A\mathbf{x} = \mathbf{0}$, για κάθε \mathbf{x} και, επομένως, $\|A\|_p = 0$.

(**A**₃) $\|aA\|_p = \max_{\|\mathbf{x}\|_p=1} \|aA\mathbf{x}\|_p \leq |a| \cdot \max_{\|\mathbf{x}\|_p=1} \|A\mathbf{x}\|_p = |a| \cdot \|A\|_p$.

(**A**₄) Χρησιμοποιώντας την τριγωνική ανισότητα (βλέπε ορισμό 4.1), έχουμε:

$$\begin{aligned} \|A+B\|_p &= \max_{\|\mathbf{x}\|_p=1} \|(A+B)\mathbf{x}\|_p \leq \max_{\|\mathbf{x}\|_p=1} \{ \|A\mathbf{x}\|_p + \|B\mathbf{x}\|_p \} \\ &\leq \max_{\|\mathbf{x}\|_p=1} \|A\mathbf{x}\|_p + \max_{\|\mathbf{x}\|_p=1} \|B\mathbf{x}\|_p = \|A\|_p + \|B\|_p. \end{aligned}$$

(**A**₅) Συνδυάζοντας την ιδιότητα (i) της φυσικής νόρμας τετραγωνικού πίνακα με τον ορισμό 4.4 παίρνουμε :

$$\|AB\|_p = \max_{\|\mathbf{x}\|_p=1} \|AB\mathbf{x}\|_p \leq \max_{\|\mathbf{x}\|_p=1} \{ \|A\|_p \|B\mathbf{x}\|_p \} = \|A\|_p \cdot \max_{\|\mathbf{x}\|_p=1} \|B\mathbf{x}\|_p = \|A\|_p \|B\|_p. \quad \blacksquare$$

Για να αποδείξουμε τους τύπους που χρησιμοποιούμε σε βασικές νόρμες πίνακα, χρειάζεται να υπενθυμίσουμε τον ορισμό που δόθηκε στην παράγραφο 3.1.

Η **φασματική ακτίνα** $\rho(A)$ ενός τετραγωνικού πίνακα A τάξης n ορίζεται από τη σχέση

$$\rho(A) = \max_{1 \leq i \leq n} \{|\lambda_i| : \lambda_i \text{ είναι ιδιοτιμή του } A\}. \quad (4.6)$$

Θεώρημα 4.1 Έστω ο τετραγωνικός πίνακας A τάξης n . Τότε για $p = 1$, $p = \infty$ και $p = 2$ προκύπτουν οι αντίστοιχες φυσικές νόρμες οι οποίες υπολογίζονται εύκολα.

$$\text{i) } \|A\|_1 = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \max_{\|\mathbf{x}\|_1=1} \|A\mathbf{x}\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|.$$

Δηλαδή, $\|A\|_1$ είναι το μεγαλύτερο άθροισμα των μέτρων των στοιχείων στήλης.

$$\text{ii) } \|A\|_\infty = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \max_{\|\mathbf{x}\|_\infty=1} \|A\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

Δηλαδή, $\|A\|_\infty$ είναι το μεγαλύτερο άθροισμα των μέτρων των στοιχείων γραμμής.

iii) $\|A\|_2 = \sqrt{\rho(AA^*)}$, όπου $\rho(AA^*)$ είναι η φασματική ακτίνα του AA^* , (βλέπε σχέση (4.6)). Δηλαδή, $\|A\|_2$ είναι η **φασματική νόρμα**.

Απόδειξη.

i) Χρησιμοποιώντας τη διανυσματική νόρμα

$$\|\mathbf{x}\|_1 = \sum_{j=1}^n |x_j|,$$

έχουμε

$$\|A\mathbf{x}\|_1 = \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}| |x_j| = \sum_{j=1}^n |x_j| \cdot \sum_{i=1}^n |a_{ij}| \leq \|\mathbf{x}\|_1 \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|.$$

Οπότε

$$\frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} \leq \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$$

και

$$\|A\|_1 \leq \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|. \quad (1)$$

Έστω ότι, για $j = k$, έχουμε $\sum_{i=1}^n |a_{ik}| = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$.

Άρα, η (1) γίνεται ισότητα, αν πάρουμε $\mathbf{x} = \mathbf{e}_k$, επειδή τότε θα έχουμε

$$\|\mathbf{x}\|_1 = \|\mathbf{e}_k\|_1 = 1$$

και

$$\frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \|A\mathbf{x}\|_1 = \|A\mathbf{e}_k\|_1 = \sum_{i=1}^n |a_{ik}| = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|.$$

Οπότε, από τη διανυσματική νόρμα $\|\cdot\|_p$, η φυσική νόρμα πίνακα είναι:

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

ii) Χρησιμοποιώντας τη διανυσματική νόρμα

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|,$$

έχουμε

$$\|A\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^n |a_{ij}| \cdot \max_{1 \leq i \leq n} |x_i| \right\} = \|\mathbf{x}\|_\infty \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Οπότε,

$$\frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

και

$$\|A\|_\infty \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|. \quad (2)$$

Έστω ότι, για $i = k$, έχουμε $\sum_{j=1}^n |a_{kj}| = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$.

Άρα, η (2) γίνεται ισότητα, όταν

$$a_{kj} x_j = |a_{kj}|, \quad j = 1, 2, 3, \dots \quad (3)$$

Επομένως,

□ αν ο πίνακας A είναι πραγματικός, τότε, από την (3) συμπεραίνεται ότι πρέπει να διαλέξουμε ένα διάνυσμα \mathbf{x} , τέτοιο ώστε:

$$x_j = \begin{cases} 1, & \text{αν } a_{kj} > 0 \\ 0, & \text{αν } a_{kj} = 0 \\ -1, & \text{αν } a_{kj} < 0 \end{cases}$$

□ Αν ο A είναι μιγαδικός πίνακας, δηλαδή, $a_{kj} = |a_{kj}|e^{-i\theta_j}$, τότε, από την (3) συμπεραίνεται ότι πρέπει να πάρουμε

$$x_j = e^{-i\theta_j}, \quad j=1,2,3,\dots,n.$$

Οπότε, $\|\mathbf{x}\|_\infty = 1$ και

$$\frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \|A\mathbf{x}\|_\infty = \sum_{j=1}^n |a_{kj}| = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|,$$

από την οποία συμπεραίνεται ότι, από τη διανυσματική νόρμα $\|\cdot\|_\infty$, προκύπτει η φυσική νόρμα πίνακα

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

iii)
$$\|A\|_2 = \max_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \max_{\mathbf{x} \neq 0} \frac{\sqrt{\mathbf{x}^* A^* A \mathbf{x}}}{\sqrt{\mathbf{x}^* \mathbf{x}}}.$$

Επειδή $(A^* A)^* = A^* (A^*)^* = A^* A$, ο πίνακας $A^* A$ είναι ερμιτιανός και κανονικός (βλέπε σελ. 3 και ορισμό 1.4, σελ. 12). Επομένως έχει πραγματικές ιδιοτιμές. Αν λ είναι μία ιδιοτιμή του, τότε πρέπει να υπάρχει ένα μη μηδενικό ιδιοδιάνυσμα \mathbf{x} αντίστοιχο της ιδιοτιμής λ , τέτοιο ώστε $A^* A \mathbf{x} = \lambda \mathbf{x}$. Από τον ορισμό του εσωτερικού γινομένου έχουμε:

$$\begin{aligned} 0 \leq \|A\mathbf{x}\|_2^2 &= \langle A\mathbf{x}, A\mathbf{x} \rangle \\ &= (A\mathbf{x})^* A\mathbf{x} = \mathbf{x}^* A^* A \mathbf{x} = \langle A^* A \mathbf{x}, \mathbf{x} \rangle = \langle \lambda \mathbf{x}, \mathbf{x} \rangle = \lambda \langle \mathbf{x}, \mathbf{x} \rangle = \lambda \|\mathbf{x}\|_2^2 \end{aligned}$$

και επειδή $\|\mathbf{x}\|_2^2 > 0$ συμπεραίνουμε ότι $\lambda \geq 0$. Επομένως, ο πίνακας $A^* A$, έχει μη αρνητικές, πραγματικές ιδιοτιμές $\lambda_i \geq 0$, για κάθε $i=1,2,3,\dots,n$, στις οποίες αντιστοιχεί ένα ορθοκανονικό σύστημα ιδιοδιανυσμάτων \mathbf{u}_i , ($i=1,2,3,\dots,n$), δηλαδή, $A^* A \mathbf{u}_i = \lambda_i \mathbf{u}_i$ και $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \mathbf{u}_j^* \mathbf{u}_i = \delta_{ij}$, όπου δ_{ij} είναι το δέλτα του Kronecker, (βλέπε ορισμό 1.2 και σχέση (1.6)).

Άρα, κάθε \mathbf{x} μπορεί να γραφεί:

$$\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i.$$

Οπότε

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^* \mathbf{x}} = \sqrt{\left(\sum_{i=1}^n a_i \mathbf{u}_i \right)^* \cdot \left(\sum_{i=1}^n a_i \mathbf{u}_i \right)} = \sqrt{\left(\sum_{i=1}^n \bar{a}_i \mathbf{u}_i^* \right) \cdot \left(\sum_{i=1}^n a_i \mathbf{u}_i \right)} = \sqrt{\sum_{i=1}^n |a_i|^2}.$$

Χρησιμοποιώντας την παραπάνω σχέση, παίρνουμε

$$\begin{aligned}\|A\mathbf{x}\|_2 &= \sqrt{\mathbf{x}^* A^* A \mathbf{x}} = \sqrt{\left(\sum_{i=1}^n a_i \mathbf{u}_i\right)^* \left(\sum_{i=1}^n \lambda_i a_i \mathbf{u}_i\right)} = \sqrt{\sum_{i=1}^n \lambda_i |a_i|^2} \leq \sqrt{\max_i |\lambda_i|} \cdot \sqrt{\sum_{i=1}^n |a_i|^2} \\ &= \sqrt{\max_i |\lambda_i|} \cdot \|\mathbf{x}\|_2,\end{aligned}$$

από όπου έχουμε :

$$\|A\mathbf{x}\|_2 \leq \sqrt{\max_i |\lambda_i|} \cdot \|\mathbf{x}\|_2 \Rightarrow \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \sqrt{\max_i |\lambda_i|} \Rightarrow \|A\|_2 \leq \sqrt{\max_i |\lambda_i|}.$$

Οπότε, η ισότητα προκύπτει, για $\mathbf{x} = \mathbf{u}_k$, με $\lambda_k = \max_{1 \leq i \leq n} \{|\lambda_i|\}$. ■

Σημείωση : (i) Για τις νόρμες N_1 και N_2 (Ευκλείδεια νόρμα) πίνακα έχουμε αντίστοιχα:

$$\|I\|_1 = n \neq 1 \quad \text{και} \quad \|I\|_2 = \sqrt{n} \neq 1.$$

δηλαδή, οι νόρμες αυτές **δεν** είναι φυσικές.

(ii) Οι νόρμες πινάκων $\|\cdot\|_1$, $\|\cdot\|_\infty$ και $\|\cdot\|_2$ επεκτείνονται και στους διανυσματικούς χώρους πινάκων $\Delta_{m \times n}$.

(iii) Οι νόρμες πινάκων $\|\cdot\|_1$ και $\|\cdot\|_\infty$ είναι κανονικές.

4.4. Φράγμα της φασματικής ακτίνας.

Για κάθε συμβιβαστή νόρμα, αν $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$, έχουμε:

$$|\lambda_i| \cdot \|\mathbf{u}_i\| = \|A\mathbf{u}_i\| \leq \|A\| \cdot \|\mathbf{u}_i\| \Rightarrow |\lambda_i| \leq \|A\|$$

και, επομένως,

$$\rho(A) = \max_{1 \leq i \leq n} |\lambda_i| \leq \|A\|. \quad (4.7)$$

Παράδειγμα 4.1. Έστω $A = \begin{pmatrix} 2 & -4 & 3 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix}$, τότε

$$\|A\|_1 = \max \{ |2| + |-4| + |3|, |1| + |-2| + |1|, |0| + |1| + |-1| \} = 7$$

$$\|A\|_\infty = \max \{ |2| + |-4| + |3|, |1| + |-2| + |1|, |0| + |1| + |-1| \} = 9$$

$$\|A\|_2 = \sqrt{4 + 16 + 9 + 1 + 4 + 1 + 1 + 1} = \sqrt{37}$$

$$\|A\|_1 = \{|2| + |-4| + |3| + |1| + |-2| + |1| + |1| + |-1|\} = 15$$

Από την εξίσωση $\det(A - \lambda I) = 0$, βρίσκουμε τις ιδιοτιμές $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = -1$.

Άρα, η φασματική ακτίνα είναι:

$$\rho(A) = \max\{|1|, |-1|, |-1|\} = 1. \quad \blacksquare$$

Λήμμα 4.2 Αν για μία φυσική νόρμα ισχύει $\|A\| < 1$, τότε

i) ο πίνακας $I \pm A$ είναι αντιστρέψιμος.

$$\text{ii)} \quad \frac{1}{1 + \|A\|} \leq \|(I \pm A)^{-1}\| \leq \frac{1}{1 - \|A\|} \quad (4.8)$$

Απόδειξη. Θεωρούμε μόνο την περίπτωση $I + A$, επειδή $I - A = I + (-A)$.

i) Υποθέτουμε το αντίθετο. Έστω ότι ο πίνακας $I + A$ δεν είναι αντιστρέψιμος. Τότε, υπάρχει ένα διάνυσμα $\mathbf{x} \neq \mathbf{0}$, τέτοιο ώστε:

$$(I + A)\mathbf{x} = \mathbf{0}. \quad (4.9)$$

Αν κανονικοποιήσουμε το \mathbf{x} , έτσι ώστε $\|\mathbf{x}\| = 1$, τότε, από την (4.9) παίρνουμε

$$A\mathbf{x} = -I\mathbf{x} = -\mathbf{x},$$

δηλαδή,

$$\|A\mathbf{x}\| = \|-\mathbf{x}\| = \|\mathbf{x}\| = 1.$$

Επειδή $\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$ και υπάρχει \mathbf{x} , για το οποίο $\|A\mathbf{x}\| = 1$, συμπεραίνουμε ότι

$$\|A\| \geq 1,$$

το οποίο είναι αντίθετο με την υπόθεση $\|A\| < 1$.

(ii) Επειδή ο $I + A$ είναι αντιστρέψιμος, μπορούμε να γράψουμε

$$I = (I + A)^{-1}(I + A) \quad (4.10)$$

Από την ιδιότητα (ii) κάθε φυσικής νόρμας τετραγωνικού πίνακα, τα αξιώματα $(A_4) - (A_5)$ του ορισμού 4.3 και την (4.10) έχουμε:

$$\begin{aligned} 1 = \|I\| &= \|(I + A)^{-1}(I + A)\| \leq \|(I + A)^{-1}\| \cdot \|(I + A)\| \leq \|(I + A)^{-1}\| (\|I\| + \|A\|) \\ &= \|(I + A)^{-1}\| (1 + \|A\|) \end{aligned}$$

Έτσι καταλήγουμε

$$\frac{1}{1 + \|A\|} \leq \|(I + A)^{-1}\|. \quad (4.11)$$

Η σχέση (4.10) μπορεί να γραφεί:

$$I = (I + A)^{-1} + (I + A)^{-1}A \Rightarrow (I + A)^{-1} = I - (I + A)^{-1}A.$$

Από αυτή και με τη βοήθεια των αξιωμάτων της νόρμας, έχουμε:

$$\|(I+A)^{-1}\| = \|I - (I+A)^{-1}A\| \leq 1 + \|(I+A)^{-1}\| \cdot \|A\| \Rightarrow \|(I+A)^{-1}\|(1-\|A\|) \leq 1.$$

Επειδή $\|A\| < 1$, από την τελευταία ανίσωση συμπεραίνουμε

$$\|(I+A)^{-1}\| \leq \frac{1}{1-\|A\|}. \quad (4.12)$$

Από τις σχέσεις (4.11) και (4.12), προκύπτει η (ii). ■

Θεώρημα 4.2 Έστω ότι οι πίνακες A και B είναι δύο τετραγωνικοί πίνακες της ίδιας τάξης. Αν ο πίνακας A είναι αντιστρέψιμος και, για κάθε φυσική νόρμα, ισχύει:

$$\|A-B\| < \frac{1}{\|A^{-1}\|}, \quad (4.13)$$

τότε, ο B είναι αντιστρέψιμος και ισχύει:

$$\|B^{-1}\| \leq \frac{\|A^{-1}\|}{1-\|A^{-1}\| \cdot \|A-B\|}, \quad (4.14)$$

καθώς επίσης και

$$\|A^{-1} - B^{-1}\| \leq \frac{\|A^{-1}\|^2 \|A-B\|}{1-\|A^{-1}\| \cdot \|A-B\|}. \quad (4.15)$$

Απόδειξη. Ισχύει η ισότητα:

$$B = A - (A-B) = A(I - A^{-1}(A-B)).$$

Από το αξίωμα (A_5) του ορισμού 4.3 και την (4.13) μπορούμε να γράψουμε

$$\|A^{-1}(A-B)\| \leq \|A^{-1}\| \cdot \|A-B\| \leq \|A^{-1}\| \frac{1}{\|A^{-1}\|} = 1,$$

οπότε σύμφωνα με το Λήμμα 4.2 (i), συμπεραίνουμε ότι ο πίνακας $I - A^{-1}(A-B)$ είναι αντιστρέψιμος.

□ Ο πίνακας B , ως γινόμενο αντιστρέψιμων πινάκων, είναι αντιστρέψιμος πίνακας, για τον οποίο ισχύει :

$$B^{-1} = (I - A^{-1}(A-B))^{-1} A^{-1}.$$

Συνδυάζοντας την τελευταία ισότητα με το αξίωμα (A_5) του ορισμού 4.3 προκύπτει η (4.14) επειδή :

$$\|B^{-1}\| = \|(I - A^{-1}(A-B))^{-1} A^{-1}\| \leq \|(I - A^{-1}(A-B))^{-1}\| \cdot \|A^{-1}\|$$

$$\begin{aligned} &\leq \frac{\|A^{-1}\|}{1 - \|A^{-1}(A-B)\|}, && \text{(Λήμμα 4.2 (ii))} \\ &\leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|A-B\|}. \end{aligned}$$

Επιπλέον ισχύει η ισότητα,

$$A^{-1} - B^{-1} = A^{-1}(B-A)B^{-1},$$

από την οποία προκύπτει:

$$\begin{aligned} \|A^{-1} - B^{-1}\| &\leq \|A^{-1}(B-A)B^{-1}\| \leq \|A^{-1}\| \cdot \|A-B\| \cdot \|B^{-1}\| \\ &\leq \frac{\|A^{-1}\|^2 \|A-B\|}{1 - \|A^{-1}\| \|A-B\|}, && \text{από την (4.14).} \quad \blacksquare \end{aligned}$$

Το θεώρημα 4.2 έχει εφαρμογές σε ένα πλήθος περιπτώσεων, βλέπε για παράδειγμα Liusternik και Sobolev, σελ. 87-93.

Στην παρούσα φάση, αν θέσουμε $B = A + \Delta A$, όπου ΔA είναι κατάλληλα **μικρές διαταραχές** του αντιστρέψιμου πίνακα A , τότε, από το παραπάνω θεώρημα, προκύπτει ότι ο πίνακας $B = A + \Delta A$ είναι επίσης αντιστρέψιμος.

Παράδειγμα 4.2. Έστω ο πίνακας

$$A = \begin{bmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \mathbb{O} \\ & 1 & 4 & 1 & \\ & & \ddots & \ddots & \ddots \\ \mathbb{O} & & & 1 & 4 & 1 \\ & & & & 1 & 4 \end{bmatrix}.$$

Τότε, $A = 4(I + B)$, όπου

$$B = \begin{bmatrix} 0 & 1/4 & & & \\ 1/4 & 0 & 1/4 & & \mathbb{O} \\ & 1/4 & 0 & 1/4 & \\ & & \ddots & \ddots & \ddots \\ \mathbb{O} & & & 1/4 & 0 & 1/4 \\ & & & & 1/4 & 0 \end{bmatrix}.$$

Χρησιμοποιώντας τη νόρμα γραμμής, $\|B\|_\infty = \frac{1}{2}$, από το Λήμμα 4.2, συμπεραίνουμε ότι ο πίνακας $(I+B)^{-1}$ υπάρχει και

$$\|(I+B)^{-1}\| \leq \frac{1}{1-1/2} = 2.$$

Έτσι, ο πίνακας A^{-1} υπάρχει,

$$A^{-1} = \frac{1}{4}(I+B)^{-1} \quad \text{και} \quad \|A^{-1}\|_\infty < \frac{1}{2}.$$

Χρησιμοποιώντας τη νόρμα γραμμής και την ανισότητα (4.7), παίρνουμε τα φράγματα:

$$\rho(A) \leq \|A\|_\infty = 6 \quad \text{και} \quad \rho(A^{-1}) \leq \|A^{-1}\|_\infty \leq \frac{1}{2}. \quad \blacksquare$$

4.5. Θεωρία διαταραχών (perturbations) για γραμμικές εξισώσεις.

Ένα υπολογιστικό πρόβλημα είναι **ασταθές** (ill-conditioned), αν οι τιμές που πρόκειται να υπολογίσουμε είναι πολύ ευαίσθητες σε μικρές διαταραχές-μεταβολές (perturbations) των αρχικών δεδομένων (πρόβλημα που συναντάμε στις αριθμητικές μεθόδους εξ αιτίας των σφαλμάτων στρογγύλευσης).

4.6. Θεωρία διαταραχών για ένα γραμμικό σύστημα.

Έστω ένας αντιστρέψιμος πίνακας A , που αντιστοιχεί στο γραμμικό σύστημα

$$A\mathbf{x} = \mathbf{b}. \quad (4.16)$$

Ορισμός 4.6 Ο αριθμός

$$C(A) = \|A\| \cdot \|A^{-1}\| \quad (4.17)$$

λέγεται **δείκτης κατάστασης** του πίνακα A .

4.6.1. Διαταραχή του δευτέρου μέλους.

Έστω $\mathbf{x} + \Delta\mathbf{x}$ είναι η λύση του παραλλαγμένου συστήματος:

$$A(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{b} + \Delta\mathbf{b}.$$

Λύνοντας το σύστημα αυτό ως προς $\Delta\mathbf{x}$, κάνοντας χρήση της $A\mathbf{x} = \mathbf{b}$, έχουμε:

$$\Delta\mathbf{x} = A^{-1}\Delta\mathbf{b}.$$

Άρα, για κάθε συμβιβαστή νόρμα έχουμε:

$$\|\Delta\mathbf{x}\| \leq \|A^{-1}\| \cdot \|\Delta\mathbf{b}\|.$$

Διαιρώντας με $\|\mathbf{x}\|$ την ανισότητα αυτή, παίρνουμε

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\|A\|^{-1} \cdot \|\Delta\mathbf{b}\|}{\|\mathbf{x}\|}.$$

Επειδή $\|\mathbf{b}\| = \|A\mathbf{x}\| \leq \|A\| \cdot \|\mathbf{x}\| \Rightarrow \frac{1}{\|\mathbf{x}\|} \leq \|A\| \frac{1}{\|\mathbf{b}\|}$, προκύπτει:

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|A\|^{-1} \|A\| \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} \quad (4.18)$$

και από την (4.17) έχουμε:

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq C(A) \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}. \quad (4.19)$$

Ανεξάρτητα από τη νόρμα που χρησιμοποιούμε για τον υπολογισμό του $C(A)$, οι τιμές αυτές φράσσονται κάτω από το ένα, επειδή

$$1 \leq \|I\| = \|A^{-1}A\| \leq \|A\| \cdot \|A^{-1}\| = C(A).$$

- Αν η τιμή του δείκτη κατάστασης είναι κοντά στο 1, τότε, από την (4.19), βλέπουμε ότι μικρές σχετικές μεταβολές στο \mathbf{b} , οδηγούν σε μικρές σχετικές μεταβολές στη λύση \mathbf{x} .
- Αλλά αν η τιμή του $C(A)$ είναι μεγάλη, τότε, από την (4.19), προκύπτει ότι μικρές σχετικές μεταβολές στο \mathbf{b} , μπορεί να οδηγήσουν σε μεγάλες σχετικές μεταβολές στη λύση \mathbf{x} .

Επειδή η τιμή του $C(A)$ εξαρτάται από την επιλογή της νόρμας, πολλές φορές χρησιμοποιούμε έναν άλλο ορισμό για το δείκτη κατάστασης, που είναι ανεξάρτητος της νόρμας.

Από την (4.7) ισχύει

$$\rho(A) \leq \|A\| \quad \text{και} \quad \rho(A^{-1}) \leq \|A^{-1}\|,$$

προκύπτει ότι

$$C(A) = \|A\| \cdot \|A^{-1}\| \geq \rho(A)\rho(A^{-1}).$$

Επιπλέον, σύμφωνα με το θεώρημα 3.11 οι ιδιοτιμές του πίνακα A^{-1} είναι οι αντίστροφες των ιδιοτιμών του A , οπότε η προηγούμενη ανισότητα γράφεται

$$C(A) \geq \frac{\max_{\lambda \in \sigma(A)} |\lambda|}{\min_{\lambda \in \sigma(A)} |\lambda|} = C(A)_*, \quad (4.20)$$

όπου $\sigma(A)$ είναι το φάσμα του πίνακα A .

Παράδειγμα 4.3. Θεωρούμε το σύστημα

$$\begin{aligned}7x_1 + 10x_2 &= b_1 \\5x_1 + 7x_2 &= b_2\end{aligned}\tag{4.21}$$

Για τον πίνακα των συντελεστών:

$$A = \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix} \text{ είναι } A^{-1} = \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix}.$$

▫ Έστω ότι στην (4.17) ο δείκτης κατάστασης του πίνακα A συμβολίζεται με $C(A)_p$, όταν αυτός υπολογίζεται, χρησιμοποιώντας τη νόρμα πίνακα $\|\cdot\|_p$. Για τον παραπάνω πίνακα, έχουμε:

$$C(A)_1 = \|A\|_1 \|A^{-1}\|_1 = 17 \cdot 17 = 289$$

$$C(A)_\infty = \|A\|_\infty \|A^{-1}\|_\infty = 289$$

$$C(A)_2 = \sqrt{\frac{\max\{\text{ιδιοτιμή του } A^*A\}}{\min\{\text{ιδιοτιμή του } A^*A\}}} = 223$$

$$C(A)_* = 198.$$

Όλοι αυτοί οι δείκτες κατάστασης δηλώνουν ότι το σύστημα (4.21) μπορεί να είναι ευαίσθητο στις διαταραχές του δευτέρου μέλους.

▫ Για να δείξουμε τη δυνατότητα αυτή, θεωρούμε την ειδική περίπτωση:

$$\begin{aligned}7x_1 + 10x_2 &= 1 \\5x_1 + 7x_2 &= 0.7\end{aligned}$$

το οποίο έχει τη λύση $x_1 = 0$, $x_2 = 0.1$.

Το παραλλαγμένο σύστημα

$$\begin{aligned}7\tilde{x}_1 + 10\tilde{x}_2 &= 1.01 \\5\tilde{x}_1 + 7\tilde{x}_2 &= 0.69\end{aligned}$$

έχει τη λύση $\tilde{x}_1 = -0.17$, $\tilde{x}_2 = 0.22$. ■

Οι σχετικές αλλαγές στο \mathbf{x} είναι αρκετά μεγάλες, σε σύγκριση με το μέγεθος των σχετικών αλλαγών στο δεύτερο μέλος \mathbf{b} .

Το προηγούμενο σύστημα (4.21) είναι ασταθές και, ειδικά, χρησιμοποιούμε την αριθμητική της κινούμενης υποδιαστολής, με τρία ή τέσσερα δεκαδικά ψηφία για τη λύση του.

Οι δείκτες κατάστασης $C(A)$ και $C(A)^*$ είναι ενδεικτικοί της αστάθειας. Καθώς αυτοί αυξάνονται με παράγοντα το 10, τότε χάνεται ένα δεκαδικό ψηφίο από την ακρίβεια της λύσης.

4.6.2. Διαταραχή του πίνακα.

Έστω ΔA η διαταραχή του πίνακα, με:

$$\|\Delta A\| < \frac{1}{\|A^{-1}\|},$$

τότε, από το θεώρημα 4.2, προκύπτει ότι ο πίνακας $A + \Delta A$ είναι αντιστρέψιμος και ότι

$$\|(A + \Delta A)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|}, \quad (4.22)$$

όπου $\|A^{-1}\| \cdot \|\Delta A\| < 1$.

Αν $\mathbf{x} + \Delta \mathbf{x}$ είναι η λύση του παραλλαγμένου συστήματος:

$$(A + \Delta A)(\mathbf{x} + \Delta \mathbf{x}) = \mathbf{b},$$

τότε, από τη λύση της εξίσωσης αυτής, ως προς $\Delta \mathbf{x}$, λαμβάνοντας υπόψη ότι $A\mathbf{x} = \mathbf{b}$, παίρνουμε:

$$\Delta \mathbf{x} = -(A + \Delta A)^{-1} \Delta A \mathbf{x}$$

Άρα, από τον ορισμό 4.4 (συμβιβαστικότητα νόρμας) και την (4.22) έχουμε:

$$\|\Delta \mathbf{x}\| \leq \|(A + \Delta A)^{-1}\| \cdot \|\Delta A \mathbf{x}\| \leq \frac{\|A^{-1}\| \cdot \|\Delta A \mathbf{x}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \leq \frac{\|A^{-1}\| \cdot \|\Delta A\| \cdot \|\mathbf{x}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|}.$$

Διαιρώντας με $\|\mathbf{x}\|$ την τελευταία ανίσωση παίρνουμε:

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\|A^{-1}\| \cdot \|\Delta A\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|} = \frac{C(A)}{1 - C(A) \frac{\|\Delta A\|}{\|A\|}} \cdot \frac{\|\Delta A\|}{\|A\|}. \quad (4.23)$$

4.6.3. Διαταραχή του πίνακα και του δευτέρου μέλους.

Έστω ΔA και $\Delta \mathbf{b}$ είναι οι διαταραχές του πίνακα A και του διανύσματος \mathbf{b} .

Υποθέτουμε ότι

$$\|\Delta A\| < \frac{1}{\|A^{-1}\|}.$$

Λύνοντας το παραλλαγμένο σύστημα

$$(A + \Delta A)(\mathbf{x} + \Delta \mathbf{x}) = \mathbf{b} + \Delta \mathbf{b}$$

ως προς $\Delta \mathbf{x}$, κάνοντας χρήση της (4.16), δηλαδή $A\mathbf{x} = \mathbf{b}$, βρίσκουμε

$$\Delta \mathbf{x} = (A + \Delta A)^{-1}(\Delta \mathbf{b} - \Delta A\mathbf{x}).$$

Συνδυάζοντας τη συμβιβαστικότητα κάθε νόρμας με την (4.22) και την τριγωνική ανισότητα προκύπτει

$$\begin{aligned} \|\Delta \mathbf{x}\| &\leq \|(A + \Delta A)^{-1}\| \cdot (\|\Delta \mathbf{b}\| + \|\Delta A\| \cdot \|\mathbf{x}\|) \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|\Delta A\|} (\|\Delta \mathbf{b}\| + \|\Delta A\| \cdot \|\mathbf{x}\|) \\ &= \frac{C(A)}{1 - C(A) \frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta \mathbf{b}\|}{\|A\|} + \frac{\|\Delta A\| \cdot \|\mathbf{x}\|}{\|A\|} \right) \quad (4.24) \end{aligned}$$

Τα φράγματα (4.23) και (4.24) δείχνουν τη σπουδαιότητα του δείκτη κατάστασης για τον καθορισμό του σφάλματος, κατά την αριθμητική επίλυση του συστήματος $A\mathbf{x} = \mathbf{b}$.

Γενικά, μπορούμε να ισχυριστούμε ότι μία μεγάλη τιμή για το $C(A)$ αποτελεί ένδειξη για ένα **ασταθές** πρόβλημα, ενώ μία μικρή τιμή για το $C(A)$ αποτελεί ένδειξη για ένα **ευσταθές** (well-conditioned) σύστημα.

4.7. Υπολογίσιμα φράγματα σφάλματος για γραμμικές εξισώσεις.

➤ Έστω C ο αριθμητικά υπολογισμένος αντίστροφος πίνακας του A . Ορίζουμε τον υπόλοιπο πίνακα με

$$R = I - CA \quad (\text{ή } R = I - AC).$$

Θεώρημα 4.3 Αν $\|R\| < 1$, για κάθε φυσική νόρμα, τότε

i) Οι πίνακες A και C είναι αντιστρέψιμοι.

ii) $\|A^{-1}\| \leq \frac{\|C\|}{1 - \|R\|}.$

iii) $\|A^{-1} - C\| \leq \frac{\|C\| \cdot \|R\|}{1 - \|R\|}.$

Απόδειξη. i) Ο πίνακας $CA = I - R$ είναι αντιστρέψιμος, επειδή $\|R\| < 1$, Λήμμα 4.2, και από την (4.8) ισχύει

$$\|(I - R)^{-1}\| \leq \frac{1}{1 - \|R\|}. \quad (4.25)$$

Εξάλλου,

$$0 \neq \det(I - R) = \det(CA) = \det C \cdot \det A.$$

Επομένως, $\det C \neq 0$ και $\det A \neq 0$, οπότε οι πίνακες A και C είναι αντιστρέψιμοι.

ii) Επειδή ο πίνακας $I - R$ είναι αντιστρέψιμος μπορούμε να γράψουμε

$$C = (I - R)A^{-1} \Rightarrow A^{-1} = (I - R)^{-1}C,$$

οπότε χρησιμοποιώντας το αξίωμα (A_5) του ορισμού 4.3 και την (4.25) έχουμε :

$$\|A^{-1}\| = \|(I - R)^{-1}C\| \leq \|C\| \cdot \|(I - R)^{-1}\| \leq \frac{\|C\|}{1 - \|R\|}.$$

iii) Επειδή ισχύει

$$R = I - CA = (A^{-1} - C)A \Rightarrow A^{-1} - C = RA^{-1},$$

χρησιμοποιώντας το αξίωμα (A_5) του ορισμού 4.3 και την ανισότητα στο (ii) έχουμε:

$$\|A^{-1} - C\| \leq \|R\| \cdot \|A^{-1}\| \leq \frac{\|C\| \cdot \|R\|}{1 - \|R\|}. \quad \blacksquare$$

➤ Έστω ότι $\hat{\mathbf{x}}$ είναι η αριθμητική λύση του συστήματος $A\mathbf{x} = \mathbf{b}$. Ορίζουμε ως υπόλοιπο διάνυσμα το

$$\mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}}.$$

Θεώρημα 4.4 Αν C είναι ο αριθμητικά υπολογισμένος πίνακας και $\|R\| < 1$, για κάθε φυσική νόρμα, όπου $R = I - CA$ (ή $R = I - AC$), τότε

$$\|\mathbf{x} - \hat{\mathbf{x}}\| \leq \frac{\|C\| \cdot \|\mathbf{r}\|}{1 - \|R\|}.$$

Απόδειξη. Επειδή $\|R\| < 1$, τότε, από το θεώρημα 4.3 (i), ο πίνακας A είναι αντιστρέψιμος και, επομένως, υπάρχει μία μοναδική λύση, \mathbf{x} .

Επομένως,

$$\mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}} = A\mathbf{x} - A\hat{\mathbf{x}} = A(\mathbf{x} - \hat{\mathbf{x}})$$

και

$$\mathbf{x} - \hat{\mathbf{x}} = A^{-1}\mathbf{r}.$$

Από την παραπάνω ισότητα και το θεώρημα 4.3 (ii) έχουμε:

$$\|\mathbf{x} - \hat{\mathbf{x}}\| \leq \|A^{-1}\| \cdot \|\mathbf{r}\| \leq \frac{\|C\| \cdot \|\mathbf{r}\|}{1 - \|R\|}. \quad \blacksquare$$

Το αποτέλεσμα του παραπάνω θεωρήματος δεν έχει πρακτική σημασία, επειδή, για τον υπολογισμό του φράγματος, απαιτείται ο υπολογισμός του αντίστροφου πίνακα.

Να σημειωθεί ότι μία μικρή τιμή του υπόλοιπου διανύσματος $\mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}}$ δε σημαίνει αναγκαστικά και μία ακριβής λύση του προβλήματος.

ΚΕΦΑΛΑΙΟ 5

ΜΕΤΡΙΚΟΙ ΧΩΡΟΙ - ΑΚΟΛΟΥΘΙΕΣ ΠΙΝΑΚΩΝ

5.1. Μετρικοί χώροι.

Ορισμός 5.1 Έστω ότι X είναι ένα μη κενό σύνολο. Αν υπάρχει μια συνάρτηση $d: X \times X \rightarrow \mathbb{R}$ για την οποία ισχύουν, για κάθε $x, y, z \in X$, τα αξιώματα:

$$(A_1) \quad d(x, y) \geq 0,$$

$$(A_2) \quad d(x, y) = 0 \Leftrightarrow x = y,$$

$$(A_3) \quad d(x, y) = d(y, x) \text{ (συμμετρική ισότητα),}$$

$$(A_4) \quad d(x, z) \leq d(x, y) + d(y, z) \text{ (τριγωνική ανισότητα),}$$

τότε το ζεύγος (X, d) λέγεται **μετρικός χώρος** και η συνάρτηση $d(x, y)$ λέγεται **μετρική**. Ο μη αρνητικός αριθμός $d(x, y)$ λέγεται **απόσταση** των σημείων x, y και ισχύουν για αυτή οι ιδιότητες:

$$\text{i) } |d(x, z) - d(y, z)| \leq d(x, y).$$

$$\text{ii) } |d(x, y) - d(z, \omega)| \leq d(x, z) + d(y, \omega).$$

Το πιο απλό γνωστό παράδειγμα μετρικού χώρου είναι το σύνολο των πραγματικών αριθμών $X = \mathbb{R}$, εφοδιασμένο με τη μετρική

$$d(x, y) = |x - y|.$$

Ένα πιο αξιόλογο παράδειγμα μετρικού χώρου είναι το σύνολο $X = \mathbb{R}^n$, $n \in \mathbb{N}$ όλων των n -διάστατων διανυσμάτων $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$, εφοδιασμένο με μία από τις μετρικές:

$$d_1(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n |x_i - y_i| \quad (1\text{-μετρική}),$$

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (\text{Ευκλείδεια μετρική}),$$

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max_{1 \leq i \leq n} |x_i - y_i| \quad (\infty\text{-μετρική}).$$

Από τα μαθήματα της Μαθηματικής Ανάλυσης γνωρίζουμε ότι η έννοια της σύγκλισης είναι μία έννοια, που εξαρτάται αποκλειστικά και μόνο από την έννοια της

«απόστασης». Επομένως, είναι φυσικό να περιμένουμε ότι η έννοια της σύγκλισης μπορεί να επεκταθεί σε οποιοδήποτε μετρικό χώρο.

Για να δούμε πόσο απλά γίνεται αυτή η επέκταση υπενθυμίζουμε τη σύγκλιση μίας ακολουθίας στο μετρικό χώρο $(\mathbb{R}, |x - y|)$.

Μία ακολουθία (x_n) σημείων του \mathbb{R} συγκλίνει στο σημείο $x \in \mathbb{R}$ και γράφουμε

$$\lim_{n \rightarrow \infty} x_n = x \quad \text{ή} \quad x_n \rightarrow x, \quad (n \rightarrow \infty),$$

αν και μόνο αν, για κάθε αριθμό $\varepsilon > 0$, υπάρχει φυσικός αριθμός n_0 , που εξαρτάται γενικώς από το ε , τέτοιος ώστε,

$$|x_n - x| < \varepsilon, \quad \text{για κάθε } n > n_0,$$

ή, ισοδύναμα, αν και μόνο αν

$$\lim_{n \rightarrow \infty} |x_n - x| = 0.$$

Η επέκταση της έννοιας της σύγκλισης μίας ακολουθίας σε ένα οποιοδήποτε μετρικό χώρο (X, d) γίνεται πολύ απλά ως εξής:

Ορισμός 5.2 Μία ακολουθία (x_n) σημείων του μετρικού χώρου (X, d) **συγκλίνει** στο $x \in X$ και γράφουμε,

$$\lim_{n \rightarrow \infty} x_n = x \quad \text{ή} \quad x_n \rightarrow x, \quad (n \rightarrow \infty),$$

αν και μόνο αν, για κάθε αριθμό $\varepsilon > 0$, υπάρχει φυσικός αριθμός $n_0(\varepsilon)$, που εξαρτάται γενικά από το ε , τέτοιος ώστε,

$$d(x_n, x) < \varepsilon \quad \text{για κάθε } n > n_0,$$

ή ισοδύναμα αν και μόνο αν

$$\lim_{n \rightarrow \infty} d(x_n, x) = 0.$$

Το σημείο x λέγεται **όριο** της ακολουθίας (x_n) , για το οποίο μπορούμε να αποδείξουμε πολύ απλά ότι είναι μοναδικό.

Οι έννοιες της *περιοχής*, του *οριακού σημείου* (*σημείου συσσωρεύσεως*), του *ανοικτού* και *κλειστού συνόλου*, επεκτείνονται με όμοιο τρόπο σε οποιοδήποτε μετρικό χώρο.

Οι έννοιες αυτές λέγονται *τοπολογικές* και έχουν βασική σημασία σε μία ειδική κατηγορία χώρων που λέγονται *τοπολογικοί χώροι* και είναι γενικότεροι των μετρικών χώρων.

Ορισμός 5.3 Ονομάζουμε ε -σφαιρική περιοχή ενός σημείου $x_0 \in X$ του μετρικού χώρου (X, d) το σύνολο των σημείων:

$$\varpi_d(x_0, \varepsilon) = \{x \in X : d(x_0, x) < \varepsilon\}.$$

Ορισμός 5.4 Δύο μετρικές d και a επί ενός συνόλου λέγονται **ισοδύναμες** αν, για κάθε $x \in X$ και $\varepsilon > 0$, υπάρχουν δύο θετικοί αριθμοί ε_1 και ε_2 (οι οποίοι μπορεί να εξαρτώνται από τα x και ε), τέτοιοι ώστε :

$$\varpi_d(x, \varepsilon_1) \subseteq \varpi_a(x, \varepsilon) \quad \text{και} \quad \varpi_a(x, \varepsilon_2) \subseteq \varpi_d(x, \varepsilon).$$

Σημείωση : Από την παραπάνω σχέση έπεται ότι αν μία ακολουθία συγκλίνει ως προς τη μία μετρική συγκλίνει και ως προς την άλλη, δηλαδή, ισοδύναμες μετρικές δίνουν το ίδιο σύνολο συγκλινοσών ακολουθιών.

Ορισμός 5.5 Μία ακολουθία (x_n) του μετρικού χώρου (X, d) λέγεται **ακολουθία Cauchy**, αν για κάθε $\varepsilon > 0$ υπάρχει $n_0(\varepsilon)$, τέτοιο ώστε

$$d(x_n, x_m) < \varepsilon, \quad \text{για κάθε } n, m > n_0$$

ή ισοδύναμα

$$d(x_{n+p}, x_n) < \varepsilon, \quad \text{για κάθε } n > n_0 \text{ και οποιοδήποτε } p \in \mathbb{N}.$$

Ορισμός 5.6 Ένας μετρικός χώρος λέγεται **πλήρης**, αν κάθε ακολουθία Cauchy αυτού συγκλίνει σε κάποιο σημείο του ίδιου του χώρου.

Για παράδειγμα, ο χώρος \mathbb{C}^n , $n \in \mathbb{N}$, είναι πλήρης.

Εάν X είναι ένας χώρος νορμέ, τότε η συνάρτηση $d : X \times X \rightarrow \mathbb{R}$, που ορίζεται για κάθε $x, y \in X$ με τον τύπο :

$$d(x, y) = \|x - y\|, \tag{5.1}$$

είναι μία μετρική επί του χώρου X . Άρα, κάθε χώρος νορμέ είναι ένας μετρικός χώρος με την παραπάνω μετρική και επομένως ισχύουν όλα γενικά τα περί μετρικών χώρων και ιδιαίτερα όσα σχετίζονται με τη θεωρία του ορίου μίας ακολουθίας.

Από την ιδιότητα **(iii)** του ορισμού της νόρμας,

$$\| \|x\| - \|y\| \| \leq \|x - y\| \tag{5.2}$$

προκύπτει ότι η νόρμα είναι μία συνεχής συνάρτηση ως προς τη μετρική (5.1) .
 Δηλαδή,

$$\lim_{x_n \rightarrow x} \|x_n - a\| = \|x - a\|.$$

Ορισμός 5.7 Ένας χώρος νορμέ που είναι πλήρης ως προς τη μετρική (5.1) λέγεται χώρος **Banach**.

Επειδή κάθε διανυσματικός χώρος νορμέ πεπερασμένης διάστασης είναι πλήρης, τότε αυτός είναι και χώρος Banach. Οι διανυσματικοί χώροι \mathbb{R}^n και $\Delta_{m \times n}$ είναι πλήρεις ως οποιαδήποτε νόρμα.

Για περισσότερες λεπτομέρειες για μετρικούς χώρους και χώρους Banach βλέπε οποιοδήποτε βιβλίο Συναρτησιακής Ανάλυσης. Για παράδειγμα βλέπε E. Kreyszig σελ.3 και σελ.55, και C. Groetsch, σελ.1 και σελ.29.

Ειδικά για μετρικούς χώρους βλέπε το βιβλίο του C.G.C. Pittis με τίτλο «Introduction to metric spaces».

5.2. Ακολουθίες πινάκων.

Ορισμός 5.8 Έστω η ακολουθία πινάκων (A_n) , όπου

$$A_n = [a_{ij}^{(n)}], \quad n = 1, 2, 3, \dots$$

είναι πίνακες του ίδιου τύπου. Ο πίνακας $A = [a_{ij}]$ λέγεται **όριο** της ακολουθίας (A_n) και γράφουμε $\lim_{n \rightarrow \infty} A_n = A$, αν και μόνο αν, για κάθε $\varepsilon > 0$, υπάρχει φυσικός αριθμός $n_0(\varepsilon)$, που εξαρτάται γενικά από το ε , τέτοιος ώστε:

$$\|A_n - A\| < \varepsilon, \quad \text{για κάθε } n > n_0,$$

ή ισοδύναμα αν και μόνο αν

$$\lim_{n \rightarrow \infty} \|A_n - A\| = 0.$$

Θεώρημα 5.1 Αν η νόρμα πίνακα είναι κανονική, τότε οι παρακάτω προτάσεις είναι ισοδύναμες:

- i) $\lim_{n \rightarrow \infty} \|A_n - A\| = 0$
- ii) $\lim_{n \rightarrow \infty} [a_{ij}^{(n)}] = [a_{ij}]$, για κάθε i, j .

Απόδειξη. Αν ισχύει η συνθήκη **i)**, τότε, για κάθε $n > n_0(\varepsilon)$, έχουμε

$$|a_{ij} - a_{ij}^{(n)}| \leq \|A - A_n\| < \varepsilon,$$

επειδή η νόρμα είναι κανονική. Οπότε,

$$\lim_{n \rightarrow \infty} a_{ij}^{(n)} = a_{ij}.$$

Αν ισχύει η **ii)**, τότε

$$|a_{ij}^{(n)} - a_{ij}| < \varepsilon, \text{ για } n > n_0(\varepsilon) \text{ και κάθε } i, j$$

και, επομένως,

$$|A_n - A| = \left[|a_{ij}^{(n)} - a_{ij}| \right] < \varepsilon E,$$

όπου E είναι ένας πίνακας του ίδιου τύπου με τους A, A_n , με όλα τα στοιχεία του ίσα με τη μονάδα. .

Από τις ιδιότητες της νόρμας, έχουμε

$$\|A_n - A\| \leq \varepsilon \|E\|, \text{ για } n > n_0(\varepsilon)$$

οπότε

$$\lim_{n \rightarrow \infty} \|A_n - A\| = 0. \quad \blacksquare$$

Λήμμα 5.1 Μία ακολουθία πινάκων (A_n) συγκλίνει στον πίνακα A , αν και μόνο αν

$$\|A - A_n\| \rightarrow 0, \text{ για } n \rightarrow \infty \quad (5.3)$$

όπου $\|A\|$ είναι μία κανονική νόρμα.

Απόδειξη. Αν $A_n \rightarrow A = [a_{ij}]$, τότε

$$|a_{ij} - a_{ij}^{(n)}| < \varepsilon, \text{ για } n > N(\varepsilon) \text{ και κάθε } i, j$$

και, επομένως,

$$\|A - A_n\| < \varepsilon E,$$

όπου E είναι ένας πίνακας του ίδιου τύπου με τους A, A_n και όλα τα στοιχεία του είναι ίσα με τη μονάδα.

Από τις ιδιότητες της νόρμας, έχουμε

$$\|A - A_n\| < \varepsilon \|E\|, \text{ για } n > N(\varepsilon)$$

οπότε

$$\lim_{n \rightarrow \infty} \|A - A_n\| = 0.$$

Αντίστροφα, αν ισχύει η συνθήκη (5.3), τότε, για κάθε $n > N(\varepsilon)$, έχουμε

$$|a_{ij} - a_{ij}^{(n)}| \leq \|A - A_n\| < \varepsilon,$$

οπότε

$$\lim_{n \rightarrow \infty} a_{ij}^{(n)} = a_{ij},$$

δηλαδή,

$$\lim_{n \rightarrow \infty} A_n = A. \quad \blacksquare$$

Πόρισμα 5.1 Για κάθε φυσική νόρμα, αν $A_n \rightarrow A$, τότε

$$\lim_{n \rightarrow \infty} \|A_n\| = \|A\|.$$

Απόδειξη. Αν $A_n \rightarrow A$, τότε έχουμε:

$$\left| \|A\| - \|A_n\| \right| \leq \|A - A_n\| \rightarrow 0, \quad \text{για } n \rightarrow \infty.$$

Άρα

$$\lim_{n \rightarrow \infty} \|A_n\| = \|A\|. \quad \blacksquare$$

Σημείωση : Αν $A = \mathbf{O}$, τότε αν $\lim_{n \rightarrow \infty} A_n = \mathbf{O} \Rightarrow \lim_{n \rightarrow \infty} \|A_n\| = 0$.

Είναι φανερό ότι, αν

$$\lim_{n \rightarrow \infty} A_n = A \quad \text{και} \quad \lim_{n \rightarrow \infty} B_n = B,$$

τότε

i) $\lim_{n \rightarrow \infty} (A_n \pm B_n) = A \pm B.$

ii) $\lim_{n \rightarrow \infty} (A_n B_n) = AB.$

iii) $\lim_{n \rightarrow \infty} A_n^{-1} = A^{-1}$, (αν ισχύει $\det A \neq 0$).

iv) $\lim_{n \rightarrow \infty} cA_n = cA$ και $\lim_{n \rightarrow \infty} A_n c = Ac$, (c σταθερά).

Όλες οι σημειούμενες παραπάνω πράξεις έχουν έννοια.

Λήμμα 5.2 (Γενίκευση του κριτηρίου Cauchy).

Μία ακολουθία πινάκων (A_n) , $n = 1, 2, 3, \dots$ συγκλίνει αν και μόνο αν για κάθε $\varepsilon > 0$, υπάρχει φυσικός αριθμός $N = N(\varepsilon)$, τέτοιος ώστε

$$\|A_{n+p} - A_n\| < \varepsilon, \quad \text{για } n > N(\varepsilon) \text{ και οποιοδήποτε } p \in \mathbb{N}. \quad (5.4)$$

Απόδειξη. Αν η ανισότητα (5.4) ισχύει, τότε για κάθε στοιχείο $a_{ij}^{(n)}$ του πίνακα A_n εφαρμόζεται το κριτήριο Cauchy και επομένως υπάρχει το

$$\lim_{n \rightarrow \infty} A_n = \left[\lim_{n \rightarrow \infty} a_{ij}^{(n)} \right]. \quad (5.5)$$

Αντίστροφα, αν υπάρχει το όριο (5.5), τότε, από το Λήμμα 5.1, συμπεραίνουμε ότι ισχύει

$$\|A - A_n\| \rightarrow 0, \quad \text{για } n \rightarrow \infty$$

και, επομένως, η ανισότητα (5.4) ισχύει. ■

Τη γνωστή έννοια της συστολικής ακολουθίας πραγματικών αριθμών, βλέπε Ε. Γαλανή, σελ. 60, την επεκτείνουμε για ακολουθίες πινάκων δίνοντας τον ακόλουθο ορισμό.

Ορισμός 5.9 Μία ακολουθία τετραγωνικών πινάκων τάξης m , (A_n) , λέγεται **συστολική**, αν και μόνο αν για κάθε φυσική νόρμα $\|\cdot\|$ ισχύει:

$$\|A_{n+2} - A_{n+1}\| \leq c \|A_{n+1} - A_n\|, \quad n \in \mathbb{N},$$

όπου $0 < c < 1$.

Θεώρημα 5.2 Κάθε συστολική ακολουθία πινάκων είναι ακολουθία Cauchy και επομένως συγκλίνουσα.

Απόδειξη. Έστω (A_n) μία συστολική ακολουθία. Από τον ορισμό 5.9 της συστολικής ακολουθίας πινάκων, έχουμε

$$\|A_{n+2} - A_{n+1}\| \leq c \|A_{n+1} - A_n\| \leq c^2 \|A_n - A_{n-1}\| \leq \dots \leq c^n \|A_2 - A_1\|.$$

Χρησιμοποιώντας την τριγωνική ιδιότητα (βλέπε ορισμό 4.3, αξίωμα (A_4)) και την προηγούμενη ανισότητα παίρνουμε:

$$\begin{aligned} \|A_{n+p} - A_n\| &\leq \|A_{n+p} - A_{n+1}\| + \|A_{n+1} - A_n\| \\ &\leq \|A_{n+p} - A_{n+2}\| + \|A_{n+2} - A_{n+1}\| + \|A_{n+1} - A_n\| \\ &\vdots \\ &\leq \|A_{n+p} - A_{n+p-1}\| + \|A_{n+p-1} - A_{n+p-2}\| + \dots + \|A_{n+1} - A_n\| \\ &\leq c^{n+p-2} \|A_2 - A_1\| + c^{n+p-3} \|A_2 - A_1\| + \dots + c^{n-1} \|A_2 - A_1\| \\ &= c^{n-1} (1 + c + \dots + c^{p-2} + c^{p-1}) \|A_2 - A_1\| \\ &= c^{n-1} \frac{1 - c^p}{1 - c} \|A_2 - A_1\| \\ &\leq \frac{c^{n-1}}{1 - c} \|A_2 - A_1\| \end{aligned}$$

Αφού $0 < c < 1$, τότε $c^{n-1} \rightarrow 0$, για $n \rightarrow \infty$. Άρα, η (A_n) είναι ακολουθία Cauchy και, επομένως, συγκλίνουσα. ■

Πόρισμα 5.2 Αν (A_n) είναι μία συστολική ακολουθία με $0 < c < 1$ και $\lim_{n \rightarrow \infty} A_n = A$, τότε, για κάθε φυσική νόρμα, έχουμε:

(i) $\|A - A_n\| \leq \frac{c^{n-1}}{1-c} \|A_2 - A_1\|.$

(ii) $\|A - A_n\| \leq \frac{c}{1-c} \|A_n - A_{n-1}\|.$

ΚΕΦΑΛΑΙΟ 6

ΣΥΣΤΟΛΙΚΗ ΑΠΕΙΚΟΝΙΣΗ ΚΑΙ ΕΠΑΝΑΛΗΠΤΙΚΕΣ ΜΕΘΟΔΟΙ

6.1. Συστολική απεικόνιση.

Ορισμός 6.1 Έστω $X = (X, d)$ ότι είναι ένας μετρικός χώρος. Μία απεικόνιση $T: X \rightarrow X$ λέγεται **συστολική** επί του X , αν υπάρχει ένας αριθμός c , $0 < c < 1$, τέτοιος ώστε, για κάθε $x, y \in X$ να ισχύει:

$$d(Tx, Ty) \leq cd(x, y) \quad (6.1)$$

Θεώρημα 6.1 (Σταθερό σημείο Banach)

Θεωρούμε ένα μετρικό χώρο $X = (X, d)$. Υποθέτουμε ότι ο χώρος είναι πλήρης και ότι η απεικόνιση $T: X \rightarrow X$ είναι συστολική, τότε η T έχει ένα μοναδικό **σταθερό σημείο**. Δηλαδή υπάρχει ακριβώς ένα σημείο $\xi \in X$, τέτοιο ώστε:

$$T\xi = \xi.$$

Απόδειξη. Ξεκινάμε με ένα $x_0 \in X$ και κατασκευάζουμε την επαναληπτική ακολουθία:

$$x_{n+1} = Tx_n, \quad n = 0, 1, 2, \dots \quad (6.2)$$

Θα δείξουμε ότι η ακολουθία (x_n) είναι Cauchy στο X .

Συνδυάζοντας την (6.2) με τον ορισμό 6.1 παίρνουμε διαδοχικά:

$$\begin{aligned} d(x_{k+1}, x_k) &= d(Tx_k, Tx_{k-1}) \leq cd(x_k, x_{k-1}) \\ &\leq c^2 d(x_{k-1}, x_{k-2}) \\ &\leq c^3 d(x_{k-2}, x_{k-3}) \\ &\vdots \\ &\leq c^k d(x_1, x_0) \end{aligned} \quad (6.3)$$

Επιπλέον, από την τριγωνική ιδιότητα (ορισμός 5.1, αξίωμα (\mathbf{A}_4)) και την (6.3) παίρνουμε:

$$\begin{aligned} d(x_{n+p}, x_n) &\leq d(x_{n+p}, x_{n+1}) + d(x_{n+1}, x_n) \\ &\leq d(x_{n+p}, x_{n+2}) + d(x_{n+2}, x_{n+1}) + d(x_{n+1}, x_n) \\ &\vdots \\ &\leq d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \dots + d(x_{n+2}, x_{n+1}) + d(x_{n+1}, x_n) \\ &\leq c^{n+p-1} d(x_1, x_0) + c^{n+p-2} d(x_1, x_0) + \dots + c^{n+1} d(x_1, x_0) + c^n d(x_1, x_0) \end{aligned}$$

$$\begin{aligned}
&= c^n (c^{p-1} + c^{p-2} + \dots + c + 1) d(x_1, x_0) \\
&= c^n \frac{1-c^p}{1-c} d(x_1, x_0)
\end{aligned}$$

Άρα,

$$d(x_{n+p}, x_n) \leq c^n \frac{1-c^p}{1-c} d(x_1, x_0) \leq \frac{c^n}{1-c} d(x_1, x_0) \quad (6.4)$$

Επειδή $0 < c < 1$ και $d(x_1, x_0)$ είναι σταθερό, συμπεραίνουμε από την (6.4) ότι

$$\frac{c^n}{1-c} d(x_1, x_0) \rightarrow 0. \text{ Οπότε } d(x_{n+p}, x_n) \rightarrow 0, \text{ δηλαδή, η } (x_n) \text{ είναι Cauchy.}$$

Επειδή ο X είναι πλήρης μετρικός χώρος, συμπεραίνεται ότι υπάρχει $\xi \in X$, τέτοιο ώστε:

$$\lim_{n \rightarrow \infty} x_n = \xi$$

Από την (6.2), δηλαδή $x_{n+1} = Tx_n$, και τη συνέχεια της T , επειδή η T είναι φραγμένη, (βλέπε τις ισοδυναμίες της πρότασης 4.1), έχουμε :

$$\xi = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} Tx_n = T \left(\lim_{n \rightarrow \infty} x_n \right) = T\xi \quad (6.5)$$

Άρα η (6.2) ισχύει.

Το ξ είναι μοναδικό σταθερό σημείο. Γιατί αν υπάρχει και ένα άλλο σημείο η , τέτοιο ώστε:

$$\eta = T\eta, \text{ με } \eta \in X$$

τότε:

$$d(\xi, \eta) = d(T\xi, T\eta) \leq cd(\xi, \eta),$$

η οποία ισχύει αν και μόνο αν $d(\xi, \eta) = 0$, δηλαδή, $\xi \equiv \eta$. ■

Πρόταση 6.1 [Εκτιμήσεις Σφαλμάτων]

Όταν οι συνθήκες του θεωρήματος 6.1 ισχύουν, τότε η επαναληπτική ακολουθία (6.2), με $x_0 \in X$ να είναι ένα αυθαίρετο σημείο, συγκλίνει στο μοναδικό σημείο ξ της T . Τότε οι εκτιμήσεις σφαλμάτων είναι:

(i) η προγενέστερη εκτίμηση:

$$d(x_n, \xi) \leq \frac{c^n}{1-c} d(x_1, x_0) \quad (6.6)$$

(ii) η μεταγενέστερη εκτίμηση:

$$d(x_n, \xi) \leq \frac{c}{1-c} d(x_{n-1}, x_n). \quad (6.7)$$

Απόδειξη. (i) Από την (6.4) για $p \rightarrow \infty$, παίρνουμε:

$$d(\xi, x_n) = d(x_n, \xi) \leq \frac{c^n}{1-c} d(x_1, x_0).$$

(ii) Ακολουθώντας την ίδια διαδικασία απόδειξης όπως στην (6.3) προκύπτει:

$$d(x_{n+p-1}, x_{n+p}) \leq c^p d(x_{n-1}, x_n).$$

Επομένως, συνδυάζοντας την προηγούμενη ανισότητα με την τριγωνική ιδιότητα έχουμε:

$$\begin{aligned} d(x_n, x_{n+p}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{n+p-1}, x_{n+p}) \\ &\leq (c + c^2 + \dots + c^p) d(x_{n-1}, x_n) \\ &= c \frac{1-c^p}{1-c} d(x_{n-1}, x_n). \end{aligned}$$

Η τελευταία ανισότητα με $p \rightarrow \infty$ δίνει :

$$d(x_n, \xi) \leq \frac{c}{1-c} d(x_{n-1}, x_n). \quad \blacksquare$$

Σημείωση : Η προγενέστερη εκτίμηση του σφάλματος (6.6) μπορεί να χρησιμοποιηθεί στην αρχή για τον υπολογισμό του αριθμού των επαναληπτικών βημάτων, τα οποία απαιτούνται για να πάρουμε μία προκαθορισμένη ακρίβεια. Η εκτίμηση (6.7) μπορεί να χρησιμοποιηθεί στο τέλος της διαδικασίας και πολλές φορές μας δίνει μεγαλύτερη ακρίβεια από την (6.6).

6.2. Εφαρμογή του θεωρήματος Banach στις γραμμικές εξισώσεις.

Το πρόβλημα προσδιορισμού σταθερού σημείου το συναντάμε σε πολλούς κλάδους των μαθηματικών.

Ειδικά, αυτό συνδέεται με τη λύση της γραμμικής αλγεβρικής εξίσωσης

$$A\mathbf{x} = \mathbf{b}, \quad (6.8)$$

όπου $A = [a_{ij}]$ είναι ένας **αντιστρέψιμος** τετραγωνικός πίνακας τάξης n ,

$\mathbf{b} = [b_1 \ b_2 \ \dots \ b_n]^T$ και \mathbf{x} το άγνωστο διάνυσμα $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$.

Έστω ο τετραγωνικός πίνακας A , τάξης n . Τότε, για κάθε φυσική νόρμα, ισχύει

$$\|A\mathbf{x} - A\mathbf{y}\| \leq \|A\| \cdot \|\mathbf{x} - \mathbf{y}\|.$$

Άρα, ο πίνακας A είναι **συστολικός**, αν και μόνο αν ισχύει $\|A\| \leq 1$.

Για να ορίσουμε μία επαναληπτική μέθοδο επίλυσης της εξίσωσης (6.8), γράφουμε τον πίνακα A ως διαφορά δύο πινάκων:

$$A = B - G,$$

όπου ο πίνακας B είναι αντιστρέψιμος. Οπότε, το σύστημα (6.8) γράφεται ισοδύναμα:

$$\begin{aligned} (B - G)\mathbf{x} &= \mathbf{b} \\ \text{ή} \quad \mathbf{x} &= B^{-1}G\mathbf{x} + B^{-1}\mathbf{b} \\ \text{ή} \quad \mathbf{x} &= U\mathbf{x} + \mathbf{c} \end{aligned} \tag{6.9}$$

όπου $U = B^{-1}G$ και $\mathbf{c} = B^{-1}\mathbf{b}$.

Από την (6.9), βλέπουμε ότι η λύση της (6.8) συμπίπτει με το σταθερό σημείο του τελεστή $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$, που ορίζεται από τον τύπο

$$T\mathbf{x} = U\mathbf{x} + \mathbf{c}.$$

Για κάθε φυσική νόρμα $\|\cdot\|$, έχουμε:

$$\|T\mathbf{x} - T\mathbf{y}\| = \|U\mathbf{x} - U\mathbf{y}\| \leq \|U\| \|\mathbf{x} - \mathbf{y}\|.$$

Από το θεώρημα σταθερού σημείου Banach, θεώρημα 6.1, προκύπτει ότι η επαναληπτική ακολουθία

$$\mathbf{x}_{k+1} = U\mathbf{x}_k + \mathbf{c} \tag{6.10}$$

ή αλλάζοντας τη θέση του δείκτη k , για λόγους ευκολίας, η ακολουθία

$$\mathbf{x}^{(k+1)} = U\mathbf{x}^{(k)} + \mathbf{c}$$

συγκλίνει, αν και μόνο αν ισχύει $\|U\| \leq 1$.

Αν $\xi \in \mathbb{R}^n$ είναι το σταθερό σημείο της (6.9), τότε, από τις (6.6) και (6.7), προκύπτουν αντίστοιχα οι εκτιμήσεις σφαλμάτων:

$$\|\mathbf{x}^{(k)} - \xi\| \leq \frac{\|U\|^k}{1 - \|U\|} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| \tag{6.11}$$

$$\|\mathbf{x}^{(k)} - \xi\| \leq \frac{\|U\|}{1 - \|U\|} \|\mathbf{x}^{(k-1)} - \mathbf{x}^{(k)}\|, \tag{6.12}$$

όπου $\mathbf{x}^{(0)} \in \mathbb{R}^n$ είναι ένα οποιοδήποτε αρχικό διάνυσμα.

6.3. Επαναληπτικές μέθοδοι.

6.3.1. Μέθοδος Jacobi (E.M.J.).

Για τον πίνακα A της εξίσωσης (6.8) υποθέτουμε ότι $a_{ii} \neq 0$, $i = 1, 2, 3, \dots, n$ και ότι ο A έχει **διαγώνια υπεροχή** (κατά γραμμή), δηλαδή

$$\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < |a_{ii}| ; \quad i = 1, 2, 3, \dots, n \quad (6.13)$$

Τότε το σύστημα (6.8) είναι ισοδύναμο με το σύστημα

$$(I - U)\mathbf{x} = \mathbf{c}, \quad (6.14)$$

όπου ο πίνακας $U = [u_{ij}]$ είναι ένας τετραγωνικός πίνακας τάξης n με στοιχεία

$$u_{ij} = \begin{cases} -\frac{a_{ij}}{a_{ii}}, & \text{αν } i \neq j \\ 0, & \text{αν } i = j \end{cases}$$

Το \mathbf{c} είναι το διάνυσμα στήλης $\mathbf{c} = \begin{bmatrix} b_1 & b_2 & \dots & b_n \\ a_{11} & a_{22} & \dots & a_{nn} \end{bmatrix}^T$.

Από τις σχέσεις (6.13) προκύπτει ότι

$$\|U\|_{\infty} = \max_{1 \leq i \leq n} \sum_{\substack{j=1 \\ j \neq i}}^n u_{ij} = \max_{1 \leq i \leq n} \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1.$$

Άρα, ο πίνακας $I - U$ είναι αντιστρέψιμος (Λήμμα 4.2) και, επομένως, το σύστημα (6.8) έχει μία μοναδική λύση.

Η εξίσωση (6.14) γράφεται

$$\mathbf{x} = U\mathbf{x} + \mathbf{c}, \quad \text{με } \|U\|_{\infty} < 1.$$

Άρα, ανεξάρτητα από το διάνυσμα \mathbf{b} , η απεικόνιση

$$T\mathbf{x} = U\mathbf{x} + \mathbf{c}$$

έχει ένα σταθερό σημείο, δηλαδή η εξίσωση (6.9) έχει μία μοναδική λύση, για κάθε $\mathbf{b} \in \mathbb{R}^n$.

Με άλλα λόγια,

οι πίνακες που έχουν διαγώνια υπεροχή (κατά γραμμή) είναι αντιστρέψιμοι και η αρχή της συστολικής απεικόνισης εξασφαλίζει ότι η επαναληπτική ακολουθία

$$\mathbf{x}^{(k+1)} = U\mathbf{x}^{(k)} + \mathbf{c} \quad (6.15)$$

συγκλίνει στη μοναδική λύση της εξίσωσης (6.9), για κάθε $\mathbf{x}^{(0)} \in \mathbb{R}^n$.

Οπότε, η επαναληπτική μέθοδος ορίζεται από τις σχέσεις:

$$\begin{aligned}
 x_1^{(k+1)} &= \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} x_2^{(k)} - \frac{a_{13}}{a_{11}} x_3^{(k)} - \dots - \frac{a_{1j}}{a_{11}} x_j^{(k)} - \dots - \frac{a_{1n}}{a_{11}} x_n^{(k)} \\
 x_2^{(k+1)} &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x_1^{(k)} - \frac{a_{23}}{a_{22}} x_3^{(k)} - \dots - \frac{a_{2j}}{a_{22}} x_j^{(k)} - \dots - \frac{a_{2n}}{a_{22}} x_n^{(k)} \\
 &\vdots \\
 x_i^{(k+1)} &= \frac{b_i}{a_{ii}} - \frac{a_{i1}}{a_{ii}} x_1^{(k)} - \frac{a_{i2}}{a_{ii}} x_2^{(k)} - \frac{a_{i3}}{a_{ii}} x_3^{(k)} - \dots - \frac{a_{in}}{a_{ii}} x_n^{(k)} \\
 &\vdots \\
 x_n^{(k+1)} &= \frac{b_n}{a_{nn}} - \frac{a_{n1}}{a_{nn}} x_1^{(k)} - \frac{a_{n2}}{a_{nn}} x_2^{(k)} - \frac{a_{n3}}{a_{nn}} x_3^{(k)} - \dots - \frac{a_{n(n-1)}}{a_{nn}} x_{n-1}^{(k)}
 \end{aligned}$$

Δηλαδή,

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left\{ b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} x_j^{(k)} \right\}, \quad \text{για κάθε } i = 1, 2, 3, \dots, n.$$

Θεώρημα 6.2 (Δεύτερη ικανή συνθήκη σύγκλισης της μεθόδου Jacobi)

Αν τα στοιχεία του πίνακα A του συστήματος

$$A\mathbf{x} = \mathbf{b} \tag{6.8}$$

ικανοποιούν τις ανισότητες:

$$\sum_{i=1}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1, \quad \text{για κάθε } j = 1, 2, 3, \dots, n,$$

τότε η επαναληπτική μέθοδος Jacobi (E.M.J) συγκλίνει για οποιοδήποτε $\mathbf{x}^{(0)} \in \mathbb{R}^n$.

Απόδειξη. Οι εξισώσεις του συστήματος (6.8) είναι

$$\sum_{j=1}^n a_{ij} x_j = b_i, \quad i = 1, 2, 3, \dots, n$$

όπου θέτουμε

$$x_i = \frac{z_i}{a_{ii}},$$

οπότε τα z_i είναι οι νέοι άγνωστοι του συστήματος

$$\sum_{j=1}^n \frac{a_{ij}}{a_{ii}} z_j = b_i, \quad \text{για κάθε } i = 1, 2, 3, \dots, n \tag{6.8a}$$

για το οποίο η επαναληπτική διαδικασία συγκλίνει ή αποκλίνει συγχρόνως με την επαναληπτική διαδικασία του αρχικού συστήματος (6.8). Ανάγοντας το σύστημα

(6.8α) στην ειδική μορφή (6.9) με το συνηθισμένο τρόπο, η επαναληπτική διαδικασία θα συγκλίνει, όταν

$$\|U\|_1 = \max_j \sum_{i=1}^n |u_{ij}| < 1,$$

από την οποία προκύπτει μία ικανή συνθήκη σύγκλισης για την επαναληπτική μέθοδο Jacobi για το σύστημα (6.8):

$$\sum_{\substack{i=1 \\ i \neq j}}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1 \quad \text{ή} \quad \sum_{\substack{i=1 \\ i \neq j}}^n |a_{ij}| < |a_{ii}|, \quad \text{για κάθε } j = 1, 2, 3, \dots, n.$$

Ειδικά, η επαναληπτική διαδικασία συγκλίνει οπωσδήποτε, αν τα στοιχεία του πίνακα U ικανοποιούν την ανισότητα

$$|u_{ij}| < \frac{1}{n},$$

όπου n είναι ο αριθμός των αγνώστων του συστήματος (6.8).

Παράδειγμα 6.1. Δίνεται το σύστημα:

$$\begin{aligned} 10x_1 - x_2 + 2x_3 - 3x_4 &= 0 \\ x_1 + 10x_2 - x_3 + 2x_4 &= 5 \\ 2x_1 + 3x_2 + 20x_3 - x_4 &= -10 \\ 3x_1 + 2x_2 + x_3 + 20x_4 &= 15 \end{aligned} \tag{1}$$

Ναδειχθεί ότι η επαναληπτική μέθοδος Jacobi (E.M.J) συγκλίνει και να βρεθεί ο αριθμός των επαναλήψεων που χρειάζονται για να πάρουμε τη λύση του συστήματος (1), με σφάλμα μικρότερο του 10^{-4} .

Λύση. ▫ Επειδή τα στοιχεία του πίνακα A του συστήματος έχουν διαγώνια υπεροχή κατά γραμμή, συμπεραίνουμε ότι το σύστημα (1) έχει μία μοναδική λύση, στην οποία συγκλίνει η E.M.J.

▫ Γράφουμε το σύστημα με τη μορφή (6.9):

$$\begin{aligned} x_1 &= 0.1x_2 - 0.2x_3 + 0.3x_4 \\ x_2 &= -0.1x_1 + 0.1x_3 - 0.2x_4 + 0.5 \\ x_3 &= -0.1x_1 - 0.15x_2 + 0.05x_4 - 0.5 \\ x_4 &= -0.15x_1 - 0.1x_2 - 0.05x_3 + 0.75 \end{aligned}$$

Τότε ο πίνακας U είναι:

$$U = \begin{bmatrix} 0 & 0.1 & -0.2 & 0.3 \\ -0.1 & 0 & 0.1 & -0.2 \\ -0.1 & -0.15 & 0 & 0.05 \\ -0.15 & 0 & -0.05 & 0 \end{bmatrix}.$$

Χρησιμοποιώντας τη νόρμα $\|U\|_\infty$, βρίσκουμε:

$$\|U\|_\infty = \max\{0.6, 0.4, 0.3, 0.2\} = 0.6.$$

Για

$$\mathbf{x}^{(0)} = \mathbf{0},$$

τότε

$$\mathbf{x}^{(1)} = [0 \quad 0.5 \quad -0.5 \quad 0.75]^T$$

και

$$\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_\infty = 0.75.$$

Αν k είναι ο αριθμός των απαιτούμενων επαναλήψεων για να επιτύχουμε τη ζητούμενη ακρίβεια, τότε από τον τύπο (6.11) έχουμε:

$$\|\mathbf{x}^{(k)} - \boldsymbol{\xi}\| \leq \frac{\|U\|_\infty^k}{1 - \|U\|_\infty} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_\infty = \frac{(0.6)^k}{1 - 0.6} 0.75 < 10^{-4}.$$

Από την τελευταία ανισότητα παίρνουμε

$$(0.6)^k < \frac{4}{7.5} 10^{-4}$$

και

$$\begin{aligned} k \cdot \log_{10} 0.6 &< \log_{10} 4 - \log_{10} 7.5 - 4 \Rightarrow \\ -k \cdot 0.2218 &< 0.6021 - 0.8751 - 4 = -4.273 \end{aligned}$$

από όπου έχουμε

$$k > \frac{4.273}{0.2218} \approx 19.27.$$

Άρα, μπορούμε να πάρουμε $k = 20$.

Εδώ χρειάζεται να τονίσουμε, ότι αυτή είναι η θεωρητική εκτίμηση του αριθμού των επαναλήψεων που χρειαζόμαστε για να πετύχουμε τη ζητούμενη ακρίβεια. Υλοποιώντας τα δεδομένα του παραδείγματος στο πρόγραμμα της MATLAB, (βλέπε στο παράρτημα) διαπιστώνουμε ότι αρκούν 7 επαναλήψεις, και τότε η λύση του συστήματος είναι :

$$\mathbf{x}^{(7)} = [0.3447 \quad 0.2719 \quad -0.5403 \quad 0.6981]^T. \quad \blacksquare$$

Σημείωση : Ακολουθώντας το γενικό πλαίσιο της κατασκευής επαναληπτικών μεθόδων για τη λύση του συστήματος $A\mathbf{x} = \mathbf{b}$, όπου ο πίνακας A έχει διαγώνια υπεροχή (κατά γραμμή) φθάνουμε στην E.M.J., αν στη διάσπαση του A ,

$$A = B - G$$

πάρουμε

$$B = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}) \quad \text{και} \quad G = B - A = [g_{ij}],$$

όπου ο $G = [g_{ij}]$ είναι τετραγωνικός πίνακας n -τάξης, με στοιχεία

$$g_{ij} = \begin{cases} -a_{ij}, & \text{αν } i \neq j \\ 0, & \text{αν } i = j \end{cases} \quad \text{για κάθε } i, j = 1, 2, 3, \dots, n.$$

Οπότε,

$$B^{-1} = \text{diag}\left(\frac{1}{a_{11}}, \frac{1}{a_{22}}, \dots, \frac{1}{a_{nn}}\right) \quad \text{και} \quad U = B^{-1}G = [u_{ij}],$$

με στοιχεία

$$u_{ij} = \begin{cases} -\frac{a_{ij}}{a_{ii}}, & \text{αν } i \neq j \\ 0, & \text{αν } i = j \end{cases} \quad \text{για κάθε } i, j = 1, 2, 3, \dots, n$$

$$\text{και } \mathbf{c} = \begin{bmatrix} \frac{b_1}{a_{11}} & \frac{b_2}{a_{22}} & \dots & \frac{b_n}{a_{nn}} \end{bmatrix}^T.$$

6.3.2. Μέθοδος Gauss-Seidel (E.M.G-S).

Είναι μία επαναληπτική μέθοδος με **διαδοχικές διορθώσεις**, στην οποία, σε κάθε βήμα, χρησιμοποιούμε όλες τις γνωστές συνιστώσες. Η μέθοδος αυτή ορίζεται από τον επαναληπτικό τύπο:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left\{ b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right\}, \quad (6.16)$$

όπου $i = 1, 2, 3, \dots, n$ και υποθέτουμε ότι για κάθε i ισχύει $a_{ij} \neq 0$.

Θεώρημα 6.3 Αν ο πίνακας A του συστήματος

$$A\mathbf{x} = \mathbf{b}$$

έχει διαγώνια υπεροχή (κατά γραμμή), τότε η E.M.G-S. συγκλίνει στη λύση του συστήματος, ανεξάρτητα από την επιλογή του αρχικού διανύσματος $\mathbf{x}^{(0)}$.

Απόδειξη. Γράφουμε το σύστημα με την ειδική μορφή (6.9)

$$\mathbf{x} = U\mathbf{x} + \mathbf{c},$$

με

$$\|U\|_{\infty} < 1. \quad (6.17)$$

Έστω $\mathbf{x}^{(k+1)} = [x_1^{(k+1)} \quad x_2^{(k+1)} \quad \dots \quad x_n^{(k+1)}]^T$ η $(k+1)$ -τάξης προσέγγιση της E.M.G-S.

Τότε έχουμε:

$$x_i^{(k+1)} = c_i + \sum_{j=1}^{i-1} u_{ij} x_j^{(k+1)} + \sum_{j=i+1}^n u_{ij} x_j^{(k)}, \quad i = 1, 2, 3, \dots, n. \quad (6.18)$$

Από την (6.17), προκύπτει ότι το σύστημα (6.8) έχει μία μοναδική λύση, έστω $\xi = [\xi_1 \quad \xi_2 \quad \dots \quad \xi_n]^T$, την οποία υπολογίζουμε με την Ε.Μ.Ι. Από την (6.9), παίρνουμε

$$\xi_i = \sum_{j=1}^n u_{ij} \xi_j + c_i, \quad i = 1, 2, 3, \dots, n. \quad (6.19)$$

Αφαιρώντας την (6.19) από την (6.18), έχουμε:

$$\xi_i - x_i^{(k+1)} = \sum_{j=1}^{i-1} u_{ij} (\xi_j - x_j^{(k+1)}) + \sum_{j=i+1}^n u_{ij} (\xi_j - x_j^{(k)}),$$

οπότε,

$$\left| \xi_i - x_i^{(k+1)} \right| \leq \sum_{j=1}^{i-1} |u_{ij}| \cdot \left| \xi_j - x_j^{(k+1)} \right| + \sum_{j=i+1}^n |u_{ij}| \cdot \left| \xi_j - x_j^{(k)} \right|, \quad i = 1, 2, 3, \dots, n \quad (6.20)$$

Από τον ορισμό της $\|\cdot\|_\infty$, έχουμε:

$$\|\xi - \mathbf{x}^{(k+1)}\|_\infty = \max_i \left| \xi_i - x_i^{(k+1)} \right|$$

και επομένως

$$\left| \xi_i - x_i^{(k+1)} \right| \leq \|\xi - \mathbf{x}^{(k+1)}\|_\infty, \quad \text{για κάθε } i = 1, 2, 3, \dots, n.$$

Οπότε από την ανισότητα (6.20) παίρνουμε

$$\left| \xi_i - x_i^{(k+1)} \right| \leq a_i \|\xi - \mathbf{x}^{(k+1)}\|_\infty + b_i \|\xi - \mathbf{x}^{(k)}\|_\infty, \quad (6.21)$$

όπου

$$a_i = \sum_{j=1}^{i-1} |u_{ij}|, \quad b_i = \sum_{j=i+1}^n |u_{ij}|, \quad i = 1, 2, 3, \dots, n,$$

με $a_1 = b_n = 0$.

Έστω $s = s(k+1)$ είναι η τιμή του δείκτη i , για την οποία ισχύει

$$\left| \xi_s - x_s^{(k+1)} \right| = \max_i \left| \xi_i - x_i^{(k+1)} \right| = \|\xi - \mathbf{x}^{(k+1)}\|_\infty.$$

Θέτοντας $i = s$ στην ανισότητα (6.21), παίρνουμε:

$$\|\xi - \mathbf{x}^{(k+1)}\|_\infty \leq a_s \|\xi - \mathbf{x}^{(k+1)}\|_\infty + b_s \|\xi - \mathbf{x}^{(k)}\|_\infty$$

ή

$$\|\xi - \mathbf{x}^{(k+1)}\|_\infty \leq \frac{b_s}{1 - a_s} \|\xi - \mathbf{x}^{(k)}\|_\infty.$$

Δηλαδή,

$$\|\xi - \mathbf{x}^{(k+1)}\|_\infty \leq \mu \|\xi - \mathbf{x}^{(k)}\|_\infty, \quad (6.22)$$

$$\text{όπου } \mu = \max_i \frac{b_s}{1-a_s}.$$

Θα αποδείξουμε ότι ισχύει

$$\mu \leq \|U\|_\infty < 1.$$

Πράγματι, επειδή

$$a_i + b_i = \sum_{j=1}^n |u_{ij}| \leq \|U\|_\infty < 1,$$

είναι φανερό ότι ισχύει

$$b_i \leq \|U\|_\infty - a_i.$$

Επομένως, συνδυάζοντας την τελευταία ανίσωση με την (6.17) γράφουμε

$$\frac{b_i}{1-a_i} \leq \frac{\|U\|_\infty - a_i}{1-a_i} \leq \frac{\|U\|_\infty - a_i \|U\|_\infty}{1-a_i} = \|U\|_\infty,$$

$$\text{δηλαδή } \mu = \max_i \frac{b_s}{1-a_s} \leq \|U\|_\infty < 1.$$

Από την ανισότητα (6.22), προκύπτει ότι

$$\|\xi - \mathbf{x}^{(k+1)}\|_\infty \leq \mu^{k+1} \|\xi - \mathbf{x}^{(0)}\|_\infty$$

και επομένως

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \xi. \quad \blacksquare$$

Σημείωση : (i) Επειδή στην E.M.J. έχουμε

$$\|\xi - \mathbf{x}^{(k+1)}\|_\infty \leq \|U\|_\infty \|\xi - \mathbf{x}^{(k)}\|_\infty,$$

ενώ στην E.M.G-S. έχουμε

$$\|\xi - \mathbf{x}^{(k+1)}\|_\infty \leq \mu \|\xi - \mathbf{x}^{(k)}\|_\infty,$$

όπου $\mu \leq \|U\|_\infty$, συμπεραίνουμε ότι η σύγκλιση της E.M.G-S. είναι κάπως καλύτερη από τη σύγκλιση της E.M.J.

(ii) Ακολουθώντας το γενικό πλαίσιο της κατασκευής επαναληπτικών μεθόδων για τη λύση του συστήματος $A\mathbf{x} = \mathbf{b}$, όπου ο A έχει διαγώνια υπεροχή (κατά γραμμή), φθάνουμε στην E.M.G-S., αν στη διάσπαση του πίνακα A :

$$A = B - G$$

Πάρουμε

$$B = \begin{bmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \mathbf{O} \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}, \quad G = B - A = \begin{bmatrix} 0 & -a_{12} & \dots & -a_{1n} \\ 0 & \dots & \dots & -a_{2n} \\ & & \ddots & \vdots \\ \mathbf{O} & & & 0 \end{bmatrix}.$$

Η επαναληπτική εξίσωση εδώ είναι

$$\mathbf{x}^{(k+1)} = U\mathbf{x}^{(k)} + \mathbf{c}, \quad k = 0, 1, 2, \dots, \text{ για κάποιο αρχικό } \mathbf{x}^{(0)},$$

όπου

$$U = B^{-1}G \quad \text{και} \quad \mathbf{c} = B^{-1}\mathbf{b}.$$

Η εξίσωση αυτή μπορεί να γραφεί

$$B\mathbf{x}^{(k+1)} = G\mathbf{x}^{(k)} + \mathbf{b}.$$

Παίρνουμε, δηλαδή, ένα γραμμικό σύστημα ως προς $\mathbf{x}^{(k+1)}$, το οποίο μπορεί να λυθεί με **εμπρός-αντικατάσταση**, για κάθε k

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left\{ b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right\}, \quad i = 1, 2, 3, \dots, n.$$

ΚΕΦΑΛΑΙΟ 7

ΣΥΝΑΡΤΗΣΕΙΣ ΠΙΝΑΚΩΝ

7.1. Σειρά πινάκων.

Έστω ότι (A_n) είναι μία ακολουθία πινάκων του ίδιου τύπου $k \times m$:

$$A_1, A_2, A_3, \dots, A_n, \dots \quad (7.1)$$

Τότε, το συμβολικό άθροισμα

$$\sum_{n=1}^{\infty} A_n = A_1 + A_2 + A_3 + \dots + A_n + \dots \quad (7.2)$$

λέγεται **σειρά πινάκων**.

Οι πίνακες $A_1, A_2, A_3, \dots, A_n, \dots$ λέγονται **όροι** της σειράς και οι αριθμοί

$$S_n = \sum_{k=1}^n A_k \quad (7.3)$$

μερικά αθροίσματα της σειράς.

Χρησιμοποιώντας την έννοια του ορίου της ακολουθίας πινάκων, εισάγουμε την έννοια της **συγκλίνουσας** σειράς πινάκων:

Λέμε ότι η σειρά (7.2) **συγκλίνει**, αν και μόνο αν η ακολουθία (S_n) των μερικών αθροισμάτων συγκλίνει, ενώ λέμε ότι **αποκλίνει**, αν και μόνο αν η ακολουθία αυτή αποκλίνει.

Αν $\lim_{n \rightarrow \infty} S_n = S$, όπου S πίνακας τύπου $k \times m$, τότε ο πίνακας S λέγεται **άθροισμα** της σειράς και γράφουμε:

$$A_1 + A_2 + A_3 + \dots + A_n + \dots = \sum_{n=1}^{\infty} A_n = S = [s_{ij}], \quad 1 \leq i \leq k, \quad 1 \leq j \leq m.$$

Ισοδύναμα, η σειρά πινάκων (7.2) συγκλίνει αν και μόνο αν

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n a_{ij}^{(k)} = s_{ij},$$

δηλαδή, αν και μόνο αν

$$\sum_{n=1}^{\infty} a_{ij}^{(n)} = s_{ij}. \quad (7.4)$$

Θεώρημα 7.1 Αν η σειρά πινάκων (7.2) συγκλίνει, τότε

$$\lim_{n \rightarrow \infty} A_n = \mathbf{O}.$$

Απόδειξη. Έστω

$$S_n = \sum_{k=1}^n A_k.$$

Αν η σειρά (7.2) συγκλίνει, τότε

$$\lim_{n \rightarrow \infty} S_n = S.$$

Επειδή $A_n = S_n - S_{n-1}$, τότε

$$\lim_{n \rightarrow \infty} A_n = \lim_{n \rightarrow \infty} S_n - \lim_{n \rightarrow \infty} S_{n-1} = S - S = \mathbf{O}.$$

■

Παράδειγμα 7.1. Η σειρά πινάκων

$$\sum_{n=1}^{\infty} \begin{bmatrix} \frac{1}{n!} & \frac{1}{n(n+1)} \\ \left(\frac{1}{2}\right)^n & \frac{n}{4^n} \end{bmatrix}$$

συγκλίνει.

Λύση. Πράγματι, επειδή

$$\sum_{n=1}^{\infty} \frac{1}{n!} = e - 1, \quad \sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1, \quad \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n = \frac{1}{1 - \frac{1}{2}} - 1 = 1$$

και

$$\sum_{n=1}^{\infty} \frac{n}{4^n} = \frac{1}{4} + 2 \frac{1}{4^2} + 3 \frac{1}{4^3} + \dots = \frac{1}{4} \{1 + \frac{2}{4} + \frac{3}{4^2} + \dots\} = \frac{1}{4} \frac{1}{\left(1 - \frac{1}{4}\right)^2} = \frac{4}{9},$$

έχουμε

$$\sum_{n=1}^{\infty} \begin{bmatrix} \frac{1}{n!} & \frac{1}{n(n+1)} \\ \left(\frac{1}{2}\right)^n & \frac{n}{4^n} \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^{\infty} \frac{1}{n!} & \sum_{n=1}^{\infty} \frac{1}{n(n+1)} \\ \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n & \sum_{n=1}^{\infty} \frac{n}{4^n} \end{bmatrix} = \begin{bmatrix} e-1 & 1 \\ 1 & \frac{4}{9} \end{bmatrix}.$$

Ενώ η σειρά, $\sum_{n=1}^{\infty} A_n$, που παράγεται από την ακολουθία

$$A_n = \begin{bmatrix} \frac{1}{n^2} & \frac{n+2}{2n+3} \\ 5 & \left(1 + \frac{2}{n}\right)^n \end{bmatrix}$$

αποκλίνει, επειδή

$$\lim_{n \rightarrow \infty} A_n = \begin{bmatrix} 0 & \frac{1}{2} \\ 5 & e^2 \end{bmatrix} \neq \mathbf{O}.$$

■

7.2. Πράξεις με σειρές πινάκων.

Ιδιότητες πράξεων με σειρές πινάκων. Αν $\sum_{n=1}^{\infty} A_n = A$ και $\sum_{n=1}^{\infty} B_n = B$, τότε

i) Η $\sum_{n=1}^{\infty} (A_n + B_n)$ συγκλίνει και $\sum_{n=1}^{\infty} (A_n + B_n) = A + B$.

ii) Η $\sum_{n=1}^{\infty} aA_n = aA$, για κάθε βαθμωτό a .

iii) Η $\sum_{n=1}^{\infty} A_n \mathbf{x} = A\mathbf{x}$, για κάθε διάνυσμα \mathbf{x} .

Όλες οι σημειούμενες παραπάνω πράξεις έχουν έννοια.

Ορισμός 7.1 Λέμε ότι η σειρά πινάκων (7.2) συγκλίνει απολύτως, όταν η σειρά

$$\sum_{n=1}^{\infty} |A_n| \tag{7.5}$$

συγκλίνει.

Θεώρημα 7.2 Μία απολύτως συγκλίνουσα σειρά πινάκων είναι μία συγκλίνουσα σειρά.

Απόδειξη. Έστω $A_n = [a_{ij}^{(n)}]$, $n = 1, 2, 3, \dots$, τότε

$$\sum_{n=1}^{\infty} |A_n| = \left[\sum_{n=1}^{\infty} |a_{ij}^{(n)}| \right],$$

επειδή η σειρά πινάκων (7.5) συγκλίνει, τότε, εξ ορισμού, κάθε μία από τις αριθμητικές σειρές

$$\sum_{n=1}^{\infty} |a_{ij}^{(n)}|, \quad 1 \leq i \leq k, \quad 1 \leq j \leq m,$$

συγκλίνει. Δηλαδή, υπάρχει ένας πίνακας S , τέτοιος ώστε

$$S = \lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \sum_{k=1}^n A_k$$

και επομένως η σειρά πινάκων (7.2) συγκλίνει.

■

Θεώρημα 7.3 Αν $\|A\|$ είναι μία φυσική νόρμα και η αριθμητική σειρά

$$\sum_{n=1}^{\infty} \|A_n\|$$

συγκλίνει, τότε η σειρά πινάκων (7.2) συγκλίνει και μάλιστα συγκλίνει απολύτως.

Απόδειξη. Έστω

$$A_n = [a_{ij}^{(n)}], \quad n = 1, 2, 3, \dots$$

Θεωρούμε τις αριθμητικές σειρές

$$\sum_{n=1}^{\infty} a_{ij}^{(n)}, \quad 1 \leq i \leq k, \quad 1 \leq j \leq m. \quad (7.6)$$

Επειδή

$$|a_{ij}^{(n)}| \leq \|A_n\|,$$

συμπεραίνουμε ότι κάθε μία από τις αριθμητικές σειρές (7.6) συγκλίνει και μάλιστα απολύτως. Επομένως, η σειρά πινάκων:

$$\sum_{n=1}^{\infty} A_n = \left[\sum_{n=1}^{\infty} a_{ij}^{(n)} \right]$$

συγκλίνει και μάλιστα απολύτως. ■

Στις εφαρμογές ιδιαίτερο ενδιαφέρον παρουσιάζουν οι **δυναμοσειρές πινάκων**:

$$\sum_{n=0}^{\infty} A_n X^n \quad (\text{εκ δεξιών}) \quad (7.7)$$

και

$$\sum_{n=0}^{\infty} X^n A_n \quad (\text{εξ αριστερών}) \quad (7.7\alpha)$$

όπου X είναι ένας τετραγωνικός πίνακας m -τάξης. Στην πρώτη περίπτωση, οι A_n είναι $k \times m$ πίνακες ή αριθμοί (για παράδειγμα, οι A_n μπορεί να είναι διανύσματα-γραμμή). Στη δεύτερη περίπτωση, οι A_n είναι $m \times k$ πίνακες ή αριθμοί (για παράδειγμα οι A_n μπορεί να είναι διανύσματα-στήλη).

Θεώρημα 7.4 Αν R είναι η ακτίνα σύγκλισης της δυναμοσειράς

$$\sum_{n=0}^{\infty} \|A_n\| x^n,$$

όπου $\|A_n\|$, ($n = 0, 1, 2, \dots$) είναι κάποια φυσική νόρμα, τότε οι σειρές πινάκων (7.7)

και (7.7α) συγκλίνουν για

$$\|X\| < R. \quad (7.8)$$

Ειδικά, η δυναμοσειρά πινάκων

$$\sum_{n=0}^{\infty} a_n X^n,$$

με αριθμητικούς συντελεστές a_n , ($n = 0, 1, 2, \dots$) συγκλίνει για

$$\|X\| < R,$$

όπου R είναι η ακτίνα σύγκλισης της δυναμοσειράς

$$\sum_{n=0}^{\infty} |a_n| x^n.$$

Απόδειξη. Επειδή

$$\|A_n X^n\| \leq \|A_n\| \cdot \|X\|^n,$$

τότε, όταν η ανισότητα (7.8) ισχύει, η σειρά

$$\sum_{n=0}^{\infty} \|A_n X^n\|$$

συγκλίνει. Επομένως, από το θεώρημα 7.3, συγκλίνει και η δυναμοσειρά (7.7).

Με όμοιο τρόπο, αποδεικνύεται ότι η δυναμοσειρά (7.7α) συγκλίνει.

Ο δεύτερος ισχυρισμός του θεωρήματος συμπεραίνεται από το γεγονός ότι, αν a_n είναι ένας αριθμός, τότε

$$\|a_n\| = |a_n|. \quad \blacksquare$$

Θεώρημα 7.5 Οι γεωμετρικές σειρές

$$A + AX + AX^2 + \dots + AX^n + \dots \quad (7.9)$$

και

$$A + XA + X^2 A + \dots + X^n A + \dots \quad (7.9\alpha)$$

όπου X είναι ένας τετραγωνικός πίνακας, συγκλίνουν αν

$$\|X\| < 1 \quad (7.10)$$

και ισχύει

$$\sum_{n=0}^{\infty} AX^n = A(I - X)^{-1}$$

και

$$\sum_{n=0}^{\infty} X^n A = (I - X)^{-1} A.$$

Απόδειξη. Αν ισχύει η συνθήκη (7.10), τότε, από το θεώρημα 7.4, η γεωμετρική σειρά (7.9) συγκλίνει, δηλαδή, υπάρχει ένας πίνακας

$$S = \sum_{n=0}^{\infty} AX^n.$$

Θεωρούμε την ταυτότητα

$$A(I + X + X^2 + \dots + X^n)(I - X) = A(I - X^{n+1}) \quad (7.11)$$

Παίρνοντας το όριο, για $n \rightarrow \infty$, επειδή $\|X\| < 1$, τότε

$$X^{n+1} \rightarrow \mathbf{O}, \quad \text{για } n \rightarrow \infty$$

και επομένως

$$S(I - X) = A.$$

Σύμφωνα με το Λήμμα 4.2 ο πίνακας $I - X$ είναι αντιστρέψιμος, οπότε

$$S = A(I - X)^{-1},$$

δηλαδή,

$$S = \sum_{n=0}^{\infty} AX^n = A(I - X)^{-1}.$$

Όμοια, αποδεικνύεται ότι

$$\sum_{n=0}^{\infty} X^n A = (I - X)^{-1} A, \quad \text{για } \|X\| < 1. \quad \blacksquare$$

Πόρισμα 7.1 Αν $\|X\| < 1$, τότε

$$(I - X)^{-1} = \sum_{n=0}^{\infty} X^n$$

και

$$\|(I - X)^{-1}\| \leq \sum_{n=0}^{\infty} \|X\|^n.$$

Απόδειξη. Η δυναμοσειρά $\sum_{n=0}^{\infty} X^n = (I - X)^{-1}$ προκύπτει από το θεώρημα 7.5, αν

θεωρήσουμε ως πίνακα A το μοναδιαίο.

Επειδή $\|\cdot\|$ είναι μία φυσική νόρμα, από το Λήμμα 4.2 προκύπτει ότι ισχύει

$$\|(I - X)^{-1}\| \leq \frac{1}{1 - \|X\|} = \sum_{n=0}^{\infty} \|X\|^n. \quad \blacksquare$$

Σημείωση : Αν $\|X\| < 1$, τότε είναι εύκολο να εκτιμήσουμε τη νόρμα του υπόλοιπου της σειράς πινάκων (7.9):

$$R_n = \|A(I-X)^{-1} - A(I+X+X^2+\dots+X^n)\| \leq \|A\| \|X^{n+1} + X^{n+2} + \dots\| = \frac{\|A\| \cdot \|X^{n+1}\|}{1 - \|X\|}.$$

Όμοια, για τη σειρά (7.9α), έχουμε:

$$R_n = \|(I-X)^{-1}A - (I+X+X^2+\dots+X^n)A\| \leq \frac{\|X^{n+1}\| \cdot \|A\|}{1 - \|X\|}.$$

7.3. Επέκταση των βαθμωτών συναρτήσεων σε συναρτήσεις πινάκων.

Θεώρημα 7.6 Υποθέτουμε ότι είναι $R > 0$ η ακτίνα σύγκλισης της δυναμοσειράς

$$\sum_{n=0}^{\infty} a_n x^n \text{ και}$$

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \text{ για } |x| < R \quad (7.12)$$

τότε, η δυναμοσειρά πινάκων $\sum_{n=0}^{\infty} a_n X^n$, όπου X είναι ένας τετραγωνικός πίνακας

τάξης m , συγκλίνει, αν όλες οι ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_m$ του πίνακα X βρίσκονται στο εσωτερικό του κύκλου σύγκλισης ακτίνας R . Στην περίπτωση αυτή, η **συνάρτηση πίνακα** $f(X)$ ορίζεται από τη σχέση

$$f(X) = \sum_{n=0}^{\infty} a_n X^n. \quad (7.13)$$

Εξ ορισμού $X^0 = I$.

Απόδειξη. Έστω ο πίνακας X , τέτοιος ώστε

$$|\lambda_1| < R, \quad |\lambda_2| < R, \dots, \quad |\lambda_m| < R,$$

Χάριν απλότητας, υποθέτουμε ότι οι ιδιοτιμές λ_j του πίνακα X είναι διαφορετικές μεταξύ τους, τότε ο πίνακας X είναι όμοιος με ένα διαγώνιο πίνακα:

$$X = P^{-1} \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m) P.$$

Εισάγοντας τα σύμβολα

$$f_n(X) = \sum_{k=0}^n a_k X^k, \quad f_n(x) = \sum_{k=0}^n a_k x^k, \quad f(x) = \lim_{n \rightarrow \infty} f_n(x) = \sum_{k=0}^{\infty} a_k x^k$$

έχουμε

$$\begin{aligned}
f_n(X) &= \sum_{k=0}^n a_k (P^{-1} \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m) P)^k \\
&= P^{-1} \left(\sum_{k=0}^n a_k \cdot \text{diag}(\lambda_1^k, \lambda_2^k, \dots, \lambda_m^k) \right) P \\
&= P^{-1} \text{diag} \left(\sum_{k=0}^n a_k \lambda_1^k, \sum_{k=0}^n a_k \lambda_2^k, \dots, \sum_{k=0}^n a_k \lambda_m^k \right) P \\
&= P^{-1} \text{diag}(f_n(\lambda_1), f_n(\lambda_2), \dots, f_n(\lambda_m)) P
\end{aligned} \tag{7.14}$$

Επειδή οι ιδιοτιμές λ_j βρίσκονται εντός του κύκλου σύγκλισης της δυναμοσειράς (7.12), συμπεραίνουμε ότι

$$f(\lambda_j) = \lim_{n \rightarrow \infty} f_n(\lambda_j), \quad j = 1, 2, \dots, m.$$

Επομένως, παίρνοντας το όριο της (7.14) για να $n \rightarrow \infty$, παίρνουμε

$$f(X) = \lim_{n \rightarrow \infty} f_n(X) = P^{-1} \text{diag}(f(\lambda_1), f(\lambda_2), \dots, f(\lambda_m)) P. \tag{7.15}$$

Άρα, η σειρά (7.13) συγκλίνει στο X . ■

Σημείωση : (i) Το θεώρημα 7.6 γενικεύεται, ισχύει και στην περίπτωση της πολλαπλότητας των ιδιοτιμών λ_j . Για την απόδειξη του ισχυρισμού, βλέπε για παράδειγμα. N.J. Pullman, σελ. 41.

(ii) Έστω ότι μία τουλάχιστον ιδιοτιμή, για παράδειγμα η λ_1 , βρίσκεται στο εξωτερικό του κύκλου ακτίνας R , δηλαδή,

$$|\lambda_1| > R.$$

Τότε, η $f_n(\lambda_1)$ δεν έχει όριο για $n \rightarrow \infty$ και, επομένως, από την (7.14), προκύπτει ότι η $f_n(X)$ δεν έχει όριο για $n \rightarrow \infty$, δηλαδή, η σειρά (7.13) αποκλίνει.

(iii) Αν υπάρχει μία ιδιοτιμή επί της περιφέρειας του κύκλου ακτίνας R , δηλαδή, αν υπάρχει μία ιδιοτιμή, τέτοια ώστε:

$$|\lambda_j| = R,$$

τότε η σειρά (7.13) συγκλίνει, αν η λ_j είναι απλή ιδιοτιμή και η (7.12) συγκλίνει.

(iv) Αν R είναι ακτίνα σύγκλισης της δυναμοσειράς $\sum_{n=0}^{\infty} a_n x^n$ και X οποιοσδήποτε

τετραγωνικός πίνακας m -τάξης, τότε η δυναμοσειρά πινάκων $\sum_{n=0}^{\infty} a_n X^n$ συγκλίνει, αν

$\rho(X) < R$, όπου $\rho(X)$ είναι η φασματική ακτίνα του πίνακα X (βλέπε ορισμό και σχέση (4.6)), δηλαδή,

$$\rho(X) = \max_{1 \leq j \leq m} \{ |\lambda_j| : \lambda_j \text{ είναι ιδιοτιμή του } X \}.$$

(v) Από τον τύπο (7.15), προκύπτει ότι, αν $\lambda_1, \lambda_2, \dots, \lambda_m$ είναι απλές ιδιοτιμές του πίνακα X , τότε $f(\lambda_1), f(\lambda_2), \dots, f(\lambda_m)$, όπου

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

είναι ιδιοτιμές της συνάρτησης

$$f(X) = \sum_{k=0}^{\infty} a_k X^k.$$

Ειδικά, οι αριθμοί $\lambda_1^n, \lambda_2^n, \dots, \lambda_m^n$ είναι ιδιοτιμές του πίνακα X^n .

Παράδειγμα 7.2. Είναι γνωστό ότι

$$e^x = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \dots + \frac{1}{n!}x^n + \dots = \sum_{n=0}^{\infty} \frac{1}{n!}x^n,$$

για κάθε x , επειδή $R = \infty$. Επειδή οποιαδήποτε ιδιοτιμή λ κάθε τετραγωνικού πίνακα X ικανοποιεί τη συνθήκη $|\lambda| < \infty$, ορίζεται η επέκταση της e^x στη συνάρτηση πίνακα e^X να είναι το ανάπτυγμα

$$e^X = 1 + \frac{1}{1!}X + \frac{1}{2!}X^2 + \dots + \frac{1}{n!}X^n + \dots = \sum_{n=0}^{\infty} \frac{1}{n!}X^n,$$

για οποιοδήποτε τετραγωνικό πίνακα X .

Αν γράψουμε $\exp(x)$ για την e^x , τότε μπορούμε να γράψουμε $\exp(X)$, αντί για e^X .

Μπορούμε να αποδείξουμε τις ακόλουθες ιδιότητες:

- i) $\det[\exp(X)] = \exp(\text{tr}X)$, όπου $\text{tr}X$ είναι το ίχνος του X .
- ii) Ο πίνακας $\exp(X)$ είναι αντιστρέψιμος, για οποιοδήποτε τετραγωνικό πίνακα X .
- iii) $f(X^T) = [f(X)]^T$, όπου η $f(X)$ είναι η επέκταση της συνάρτησης $f(x)$.
- iv) $e^a B = e^{aI} B$, για κάθε a και κάθε πίνακα B .
- v) $e^I = eI$.
- vi) $e^A e^B = e^{A+B}$. ■

Παράδειγμα 7.3. Είναι γνωστό ότι

$$\sum_{n=0}^{\infty} x^n = (1-x)^{-1}, \quad \text{για κάθε } |x| < 1.$$

Άρα, η δυναμοσειρά πινάκων $\sum_{n=0}^{\infty} X^n$ συγκλίνει, για οποιοδήποτε τετραγωνικό πίνακα X , με $\rho(X) < 1$. Επειδή $\rho(X) \leq \|X\|$, για οποιαδήποτε φυσική νόρμα (βλέπε σχέση 4.7), συμπεραίνουμε ότι η δυναμοσειρά πινάκων συγκλίνει, για $\|X\| < 1$. Στην περίπτωση αυτή ισχύει

$$\sum_{n=0}^{\infty} X^n = (I - X)^{-1} \quad (7.16)$$

και

$$(I - X) \sum_{n=0}^{\infty} X^n = I.$$

Σύγκρισε το αποτέλεσμα με το Πόρισμα 7.1.

Αν η σειρά αποκλίνει, τότε ο $X \not\rightarrow \mathbf{O}$, για $n \rightarrow \infty$. ■

Εφαρμογή 7.1 Να υπολογιστεί το άθροισμα

$$\sum_{n=0}^{\infty} \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}^n.$$

Απόδειξη. Επειδή $\|X\|_{\infty} = \max\{0.8, 0.9\} = 0.9 < 1$, από την (7.16) έχουμε

$$\begin{aligned} \sum_{n=0}^{\infty} \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}^n &= (I - X)^{-1} \\ \sum_{n=0}^{\infty} \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}^n &= \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix} \right\}^{-1} \\ &= \begin{bmatrix} 0.9 & -0.7 \\ -0.3 & 0.4 \end{bmatrix}^{-1} \\ &= \frac{1}{0.4 \times 0.9 - 0.3 \times 0.7} \begin{bmatrix} 0.4 & 0.7 \\ 0.3 & 0.9 \end{bmatrix} \\ &= \frac{1}{0.15} \begin{bmatrix} 0.4 & 0.7 \\ 0.3 & 0.9 \end{bmatrix} = \begin{bmatrix} \frac{8}{3} & \frac{14}{3} \\ 2 & 6 \end{bmatrix}. \end{aligned}$$

Εξ άλλου,

$$\sum_{n=1}^{\infty} nx^n = x \sum_{n=1}^{\infty} nx^{n-1} = x(1-x)^{-2}, \quad |x| < 1.$$

Επομένως,

$$\sum_{n=1}^{\infty} nX^n = X(I - X)^{-2}, \quad \rho(X) < 1. \quad \blacksquare$$

7.4. Επέκταση των βαθμωτών ταυτοτήτων.

Λήμμα 7.1 Αν

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \quad g(x) = \sum_{n=0}^{\infty} b_n x^n, \quad s(x) = f(x) + g(x) \quad \text{και} \quad p(x) = f(x)g(x),$$

για κάθε $|x| < R$, τότε, για κάθε τετραγωνικό πίνακα X , με $\rho(X) < R$, έχουμε:

- i) $s(X) = f(X) + g(X)$.
- ii) $p(X) = f(X)g(X)$.

Εφαρμογή 7.2 Ναδειχτεί ότι για κάθε τετραγωνικό πίνακα X ισχύει

- i) $\sin^2 X + \cos^2 X = I$.
- ii) $\sin 2X = 2 \sin X \cos X$.

Απόδειξη. (i) Έστω X ένας τετραγωνικός πίνακας. Υπάρχουν a_n και b_n , τέτοιοι ώστε:

$$\sin^2 x = \sum_{n=0}^{\infty} a_n x^n \quad \text{και} \quad \cos^2 x = \sum_{n=0}^{\infty} b_n x^n.$$

Από το παραπάνω Λήμμα 7.1, συμπεραίνουμε ότι

$$\sin^2 X = \sum_{n=0}^{\infty} a_n X^n \quad \text{και} \quad \cos^2 X = \sum_{n=0}^{\infty} b_n X^n.$$

Από την παράγραφο 7.2 και την ιδιότητα (i) των πράξεων των δυναμοσειρών πινάκων έχουμε

$$\sin^2 X + \cos^2 X = \sum_{n=0}^{\infty} (a_n + b_n) X^n \quad (7.17\alpha)$$

$$\sin^2 x + \cos^2 x = \sum_{n=0}^{\infty} (a_n + b_n) x^n \quad (7.17\beta)$$

Αλλά $\sin^2 x + \cos^2 x = 1$, επομένως $a_0 + b_0 = 1$, ενώ $a_n + b_n = 0$, για κάθε $n > 0$.

Επομένως, από την εξίσωση (7.17α), συμπεραίνεται ότι ισχύει $\sin^2 X + \cos^2 X = I$.

(ii) Υπάρχουν c_n , τέτοια ώστε:

$$\sin x \cos x = \sum_{n=0}^{\infty} c_n x^n,$$

οπότε, από το Λήμμα 7.1, προκύπτει ότι

$$\sin X \cos X = \sum_{n=0}^{\infty} c_n X^n. \quad (7.18)$$

Επειδή $2x \sin x \cos x = \sin 2x$, για κάθε x , έχουμε

$$\sin 2x = \sum_{n=0}^{\infty} 2c_n x^n$$

και επομένως

$$\sin 2X = \sum_{n=0}^{\infty} 2c_n X^n .$$

Εξ άλλου, από την εξίσωση (7.18), έχουμε

$$2 \sin X \cos X = \sum_{n=0}^{\infty} 2c_n X^n .$$

Επομένως,

$$\sin 2X = 2 \sin X \cos X . \quad \blacksquare$$

7.5. Υπολογισμός συναρτήσεων πινάκων.

Γενικά, ένα ανάπτυγμα της μορφής (7.13) δεν είναι χρήσιμο για τον υπολογισμό των στοιχείων του $f(X)$. Από το θεώρημα Cayley-Hamilton, συμπεραίνεται ότι κάθε συνάρτηση, καλά ορισμένη, ενός τετραγωνικού πίνακα n -τάξης, μπορεί να εκφραστεί ως ένα πολυώνυμο $n-1$ βαθμού ως προς τον πίνακα A .

Έτσι

$$f(A) = a_{n-1}A^{n-1} + a_{n-2}A^{n-2} + \dots + a_1A + a_0I, \quad (7.19)$$

όπου $a_{n-1}, a_{n-2}, \dots, a_1, a_0$ είναι προσδιοριστέοι συντελεστές, οι οποίοι υπολογίζονται ως εξής:

Βήμα 7.1 Από την (7.19) σχηματίζουμε το πολυώνυμο

$$r(\lambda) = a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \dots + a_1\lambda + a_0,$$

όπου A^j αντικαθίσταται με λ^j , ($j = 0, 1, 2, \dots, n-1$).

Βήμα 7.2 Για κάθε διακεκριμένη ιδιοτιμή λ_i του A , κατασκευάζουμε την εξίσωση:

$$f(\lambda_i) = r(\lambda_i) \quad (7.20)$$

Βήμα 7.3 Αν λ_i είναι μία ιδιοτιμή με αλγεβρική πολλαπλότητα k , τότε σχηματίζουμε και τις παρακάτω εξισώσεις, που περιέχουν τις παραγώγους του $f(\lambda)$ και $r(\lambda)$ ως προς λ :

$$\begin{aligned}
f'(\lambda)_{\lambda=\lambda_i} &= r'(\lambda)_{\lambda=\lambda_i} \\
f''(\lambda)_{\lambda=\lambda_i} &= r''(\lambda)_{\lambda=\lambda_i} \\
&\vdots \\
f^{(k-1)}(\lambda)_{\lambda=\lambda_i} &= r^{(k-1)}(\lambda)_{\lambda=\lambda_i}
\end{aligned}
\tag{7.21}$$

Βήμα 7.4 Λύνουμε το σύστημα των εξισώσεων (7.20) και (7.21) ως προς $a_{n-1}, a_{n-2}, \dots, a_1, a_0$.

Αντικαθιστώντας τις τιμές αυτές στην (7.19), μπορούμε να υπολογίσουμε την $f(A)$.

Παράδειγμα 7.4. Να βρεθούν οι πίνακες X , για τους οποίους ορίζεται η συνάρτηση $\cos X$ και, κατόπιν, να υπολογιστεί η τιμή της συνάρτησης $\cos A$, με $A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$.

Λύση. ▫ Η σειρά Maclaurin της $\cos x$ είναι

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \dots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}x^{2n},$$

η οποία συγκλίνει, για κάθε x , επειδή $R = \infty$. Κάθε ιδιοτιμή ενός οποιουδήποτε τετραγωνικού πίνακα X ικανοποιεί τη σχέση $|\lambda| < \infty$.

Οπότε,

$$\cos X = I - \frac{1}{2!}X^2 + \frac{1}{4!}X^4 - \frac{1}{6!}X^6 + \dots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}X^{2n},$$

ορίζεται για κάθε τετραγωνικό πίνακα X .

▫ Ο πίνακας $\cos A$ υπολογίζεται από το πολυώνυμο

$$\cos A = a_1 A + a_0 I,$$

όπου a_1, a_0 είναι προσδιοριστέοι συντελεστές.

Οι ιδιοτιμές του πίνακα A είναι:

$$\begin{vmatrix} 2-\lambda & 4 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 4 = 0 \Rightarrow 2-\lambda = \pm 2 \Rightarrow \lambda_1 = 0, \lambda_2 = 4$$

$$f(\lambda) = \cos \lambda \quad \text{και} \quad r(\lambda) = a_1 \lambda + a_0.$$

Για $\lambda = 0$, $f(0) = 1$ και $r(0) = a_0 \Rightarrow a_0 = 1$.

Για $\lambda = 4$, $f(4) = \cos 4$ και $r(4) = a_1 4 + a_0 \Rightarrow 4a_1 + 1 = \cos 4 \Rightarrow a_1 = \frac{\cos 4 - 1}{4}$.

Άρα,

$$\cos A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\cos 4 - 1}{4} \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2(1 + \cos 4) & 4(\cos 4 - 1) \\ \cos 4 - 1 & 2(1 + \cos 4) \end{bmatrix}.$$

Στη συγκεκριμένη περίπτωση μπορούμε να υπολογίσουμε τον πίνακα $\cos A$ από τον τύπο

$$\cos A = P \begin{bmatrix} 1 & 0 \\ 0 & \cos 4 \end{bmatrix} P^{-1},$$

όπου $P = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$ είναι ο πίνακας με στήλες τα ιδιοδιανύσματα που αντιστοιχούν στις ιδιοτιμές λ_1, λ_2 .

Άρα,

$$\cos A = \frac{1}{4} \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \cos 4 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & 2 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2(1 + \cos 4) & 4(\cos 4 - 1) \\ \cos 4 - 1 & 2(1 + \cos 4) \end{bmatrix}. \quad \blacksquare$$

Παράδειγμα 7.5. Να υπολογιστεί ο πίνακας $\cos \pi A$, για $A = \begin{bmatrix} -2 & 2 & -2 & 4 \\ -1 & 2 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ -2 & 1 & -1 & 4 \end{bmatrix}$.

Λύση. Ο πίνακας $\cos \pi A$ υπολογίζεται από το πολυώνυμο:

$$\cos \pi A = a_3 \pi^3 A^3 + a_2 \pi^2 A^2 + a_1 \pi A + a_0 I,$$

όπου a_3, a_2, a_1, a_0 προσδιοριστέοι συντελεστές.

Θέτουμε $B = \pi A$.

Οι ιδιοτιμές του πίνακα A είναι:

$$\begin{vmatrix} -2 - \lambda & 2 & -2 & 4 \\ -1 & 2 - \lambda & -1 & 1 \\ 0 & 0 & 1 - \lambda & 0 \\ -2 & 1 & -1 & 4 - \lambda \end{vmatrix} = \begin{vmatrix} 2 - \lambda & 2 & -2 & 4 \\ 0 & 2 - \lambda & -1 & 1 \\ 0 & 0 & 1 - \lambda & 0 \\ 2 - \lambda & 1 & -1 & 4 - \lambda \end{vmatrix} \\ = (2 - \lambda) \begin{vmatrix} 1 & 2 & -2 & 4 \\ 0 & 2 - \lambda & -1 & 1 \\ 0 & 0 & 1 - \lambda & 0 \\ 1 & 1 & -1 & 4 - \lambda \end{vmatrix}$$

$$\begin{aligned}
&= (2-\lambda) \begin{vmatrix} 1 & 2 & -2 & 4 \\ 0 & 2-\lambda & -1 & 1 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & -1 & 1 & -\lambda \end{vmatrix} \\
&= (2-\lambda) \begin{vmatrix} 2-\lambda & -1 & 1 \\ 0 & 1-\lambda & 0 \\ -1 & 1 & -\lambda \end{vmatrix} \\
&= (2-\lambda)(1-\lambda) \begin{vmatrix} 2-\lambda & 1 \\ -1 & -\lambda \end{vmatrix} \\
&= (2-\lambda)(1-\lambda)(-2\lambda + \lambda^2 + 1) = (2-\lambda)(1-\lambda)^3.
\end{aligned}$$

Άρα, οι ιδιοτιμές του A είναι $\lambda_1 = 2$ και $\lambda_2 = 1$ (πολλαπλότητα 3).

Οι ιδιοτιμές του πίνακα B είναι $\lambda_1 = 2\pi$, $\lambda_2 = \pi$.

$$\begin{aligned}
f(\lambda) &= \cos \lambda, & r(\lambda) &= a_3 \lambda^3 + a_2 \lambda^2 + a_1 \lambda + a_0. \\
f(2\pi) &= \cos 2\pi = 1, & r(2\pi) &= 8a_3 \pi^3 + 4a_2 \pi^2 + 2a_1 \pi + a_0 = 1. \\
f(\pi) &= \cos \pi = -1, & r(\pi) &= a_3 \pi^3 + a_2 \pi^2 + a_1 \pi + a_0 = -1. \\
f'(\lambda) &= -\sin \lambda, & r'(\lambda) &= 3a_3 \lambda^2 + 2a_2 \lambda + a_1. \\
f'(\pi) &= -\sin \pi = 0, & r'(\pi) &= 3a_3 \pi^2 + 2a_2 \pi + a_1 = 0. \\
f''(\lambda) &= -\cos \lambda, & r''(\lambda) &= 6a_3 \lambda + 2a_2. \\
f''(\pi) &= -\cos \pi = 1, & r''(\pi) &= 6a_3 \pi + 2a_2 = 1.
\end{aligned}$$

Από το σύστημα

$$8a_3 \pi^3 + 4a_2 \pi^2 + 2a_1 \pi + a_0 = 1$$

$$a_3 \pi^3 + a_2 \pi^2 + a_1 \pi + a_0 = -1$$

$$3a_3 \pi^2 + 2a_2 \pi + a_1 = 0$$

$$6a_3 \pi + 2a_2 = 1$$

βρίσκουμε $a_0 = \pi^2 - 3$, $a_1 = \frac{12 - 5\pi^2}{2\pi}$, $a_2 = \frac{2\pi^2 - 6}{\pi^2}$, $a_3 = \frac{4 - \pi^2}{2\pi^3}$, οπότε

$$\cos \pi A =$$

$$\begin{aligned}
 &= (\pi^2 - 3)I + \frac{12 - 5\pi^2}{2}A + (2\pi^2 - 6)A^2 + \frac{4 - \pi^2}{2}A^3 \\
 &= (\pi^2 - 3)I + \frac{12 - 5\pi^2}{2} \begin{bmatrix} -2 & 2 & -2 & 4 \\ -1 & 2 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ -2 & 1 & -1 & 4 \end{bmatrix} \\
 &\quad + (2\pi^2 - 6) \begin{bmatrix} -6 & 4 & -4 & 10 \\ -2 & 3 & -2 & 2 \\ 0 & 0 & 1 & 0 \\ -5 & 2 & -2 & 9 \end{bmatrix} + \frac{4 - \pi^2}{2} \begin{bmatrix} -12 & 6 & -6 & 20 \\ -3 & 4 & -3 & 3 \\ 0 & 0 & 1 & 0 \\ -10 & 3 & -3 & 18 \end{bmatrix} \\
 &= \begin{bmatrix} -3 & 0 & 0 & -4 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -2 & 0 & 0 & 3 \end{bmatrix}.
 \end{aligned}$$

■

7.6. Η συνάρτηση e^{At} .

Για κάθε σταθερό πίνακα A και πραγματική μεταβλητή t , η συνάρτηση πίνακα e^{At} υπολογίζεται, θέτοντας $B = At$ και, κατόπιν, υπολογίζουμε τον πίνακα e^B , όπως περιγράφεται στην προηγούμενη παράγραφο.

Οι ιδιοτιμές του $B = At$ είναι οι ιδιοτιμές του A , πολλαπλασιασμένες επί t . Να σημειωθεί ότι οι (7.21) περιέχουν παραγώγους ως προς λ και όχι ως προς t . Η σωστή σειρά των βημάτων στην περίπτωση αυτή είναι να πάρουμε αναγκαστικά πρώτα τις παραγώγους της $f(\lambda)$ και του $r(\lambda)$ και, κατόπιν, να θέσουμε $\lambda = \lambda_j t$. Η αντίστροφη διαδικασία μπορεί να οδηγήσει σε λάθος αποτελέσματα.

Παράδειγμα 7.6. Να υπολογιστεί ο πίνακας e^{At} , για $A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$.

Λύση. Ως γνωστό η συνάρτηση

$$e^X = \sum_{n=0}^{\infty} \frac{1}{n!} X^n$$

ορίζεται για κάθε τετραγωνικό πίνακα X .

Ο πίνακας e^{At} υπολογίζεται από το πολυώνυμο

$$e^{At} = a_1 At + a_0 I,$$

όπου a_1, a_0 είναι προσδιοριστέοι συντελεστές.

Θέτουμε $B = At$. Οι ιδιοτιμές του πίνακα A είναι

$$\begin{vmatrix} 1-\lambda & 4 \\ 3 & 2-\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - 3\lambda - 10 = 0 \Rightarrow \lambda_1 = 5, \lambda_2 = -2.$$

Άρα, οι ιδιοτιμές του πίνακα B είναι $\lambda_1 = 5t$, $\lambda_2 = -2t$.

$$f(\lambda) = e^\lambda, \quad r(\lambda) = a_1\lambda + a_0.$$

$$f(5t) = e^{5t}, \quad r(5t) = 5ta_1 + a_0 \Rightarrow 5ta_1 + a_0 = e^{5t} \quad (1)$$

$$f(-2t) = e^{-2t}, \quad r(-2t) = -2ta_1 + a_0 \Rightarrow -2ta_1 + a_0 = e^{-2t}$$

Από το σύστημα (1) βρίσκουμε

$$a_1 = \frac{1}{7t}(e^{5t} - e^{-2t}), \quad a_0 = \frac{1}{7}(5e^{-2t} + 2e^{5t}),$$

$$e^{At} = \frac{5e^{-2t} + 2e^{5t}}{7} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{e^{5t} - e^{-2t}}{7} \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} 3e^{5t} + 4e^{-2t} & 4e^{5t} - 4e^{-2t} \\ 3e^{5t} - 3e^{-2t} & 4e^{5t} + 3e^{-2t} \end{bmatrix}. \quad \blacksquare$$

ΚΕΦΑΛΑΙΟ 8

ΔΙΑΦΟΡΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

8.1. Συναρτήσεις πινάκων μίας μεταβλητής.

Υποθέτουμε ότι οι συναρτήσεις μίας μεταβλητής:

$$f_{11}(t), f_{12}(t), \dots, f_{ij}(t), \dots, f_{mn}(t)$$

έχουν κοινό πεδίο ορισμού, Δ , τότε, ο πίνακας

$$F(t) = \begin{bmatrix} f_{11}(t) & \dots & f_{1j}(t) & \dots & f_{1n}(t) \\ \vdots & & \vdots & & \vdots \\ f_{i1}(t) & \dots & f_{ij}(t) & \dots & f_{in}(t) \\ \vdots & & \vdots & & \vdots \\ f_{m1}(t) & \dots & f_{mj}(t) & \dots & f_{mn}(t) \end{bmatrix}, t \in \Delta$$

είναι μία **συνάρτηση πίνακα**.

Παράδειγμα 8.1.

$$F(t) = \begin{bmatrix} \sin t & \cos t & t \\ \frac{\sin t}{t} & e^t & t^2 \\ 1 & 0 & t^3 \end{bmatrix}, t \neq 0$$

Παράδειγμα 8.2.

$$\mathbf{x}(t) = \begin{bmatrix} e^t \\ \frac{1}{t^2} \\ 1 \end{bmatrix}, t \neq 0$$

Παράδειγμα 8.3.

$$\mathbf{y}(t) = \begin{bmatrix} \sin \frac{\pi}{t} \\ \sin \pi t \end{bmatrix}, t \neq 0$$

Ορισμοί 8.1 i) Για μία συνάρτηση πίνακα $F(t)$ λέμε ότι $\lim_{t \rightarrow t_0} F(t) = A = [a_{ij}]$, αν και

μόνο αν $\lim_{t \rightarrow t_0} f_{ij}(t) = a_{ij}$, για κάθε i, j .

ii) Λέμε ότι μία συνάρτηση $F(t)$ είναι **συνεχής** στο σημείο $t = t_0$, αν και μόνο αν

$$\lim_{t \rightarrow t_0} F(t) = F(t_0),$$

δηλαδή, αν και μόνο αν $\lim_{t \rightarrow t_0} f_{ij}(t) = f_{ij}(t_0)$.

Ισοδύναμα, μία συνάρτηση πίνακα είναι συνεχής στο σημείο $t = t_0$, αν και μόνο αν οι συναρτήσεις $f_{ij}(t)$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$ είναι συνεχείς στο $t = t_0$.

iii) Λέμε ότι μία συνάρτηση πίνακα, $F(t)$, είναι **παραγωγίσιμη** στο σημείο $t = t_0$, αν και μόνο αν κάθε συνάρτηση $f_{ij}(t)$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$ είναι παραγωγίσιμη στο σημείο $t = t_0$ και γράφουμε

$$\left. \frac{dF(t)}{dt} \right|_{t=t_0} = \lim_{h \rightarrow 0} \frac{F(t_0 + h) - F(t_0)}{h}.$$

iv) Μία συνάρτηση πίνακα, $F(t)$, $t \in [a, b]$ είναι **ολοκληρώσιμη** στο σημείο $[a, b]$, αν και μόνο αν κάθε συνάρτηση $f_{ij}(t)$ είναι ολοκληρώσιμη στο σημείο $[a, b]$.

Έτσι, στο παράδειγμα 8.1, έχουμε

$$\lim_{t \rightarrow 0} F(t) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Στο παράδειγμα 8.2 βρίσκουμε

$$\lim_{t \rightarrow -\infty} \mathbf{x}(t) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Ενώ η συνάρτηση $\mathbf{y}(t)$ του παραδείγματος 8.3 δεν έχει όριο, για $t \rightarrow 0$.

Θεώρημα 8.1 Έστω ότι η συνάρτηση πίνακα $F(t)$ είναι ολοκληρώσιμη σε ένα διάστημα $[a, b]$. Θεωρούμε τον πίνακα

$$\Phi(t) = [\phi_{ij}(t)] = \left[\int_a^t f_{ij}(s) ds \right] = \int_a^t F(s) ds, \quad a \leq t \leq b.$$

Τότε

i) Η $\Phi(t)$ είναι συνεχής στο $[a, b]$.

$$\text{ii) } \frac{d\Phi(t)}{dt} = \left[\frac{d\phi_{ij}(t)}{dt} \right] = [f_{ij}(t)] = F(t),$$

για κάθε t , όπου η $F(t)$ είναι συνεχής.

Άμεσες συνέπειες.

i) Αν οι συναρτήσεις πινάκων $F(t)$ και $G(t)$ είναι παραγωγίσιμες στο Δ , τότε

$$\frac{d[aF(t) + bG(t)]}{dt} = a \frac{dF(t)}{dt} + b \frac{dG(t)}{dt}, \quad t \in \Delta \text{ και } a, b \text{ αριθμοί.}$$

ii) Αν οι συναρτήσεις πινάκων $F(t)$ και $G(t)$ είναι παραγωγίσιμες για $t \in \Delta$, τότε

$$\frac{d[F(t)G(t)]}{dt} = \frac{dF(t)}{dt}G(t) + F(t)\frac{dG(t)}{dt}, \quad t \in \Delta.$$

iii) Αν ο πίνακας $F(t)$ είναι παραγωγίσιμος και αντιστρέψιμος, για κάθε $t \in \Delta$, τότε

$$\frac{dF^{-1}(t)}{dt} = -F^{-1}(t) \frac{dF(t)}{dt} F^{-1}(t).$$

Απόδειξη. Επειδή $F(t)F^{-1}(t) = I$, από την (ii), παίρνουμε:

$$\frac{dF(t)}{dt}F^{-1}(t) + F(t)\frac{dF^{-1}(t)}{dt} = \mathbf{O}.$$

Πολλαπλασιάζοντας από αριστερά επί τον πίνακα $F^{-1}(t)$, παίρνουμε:

$$\frac{dF^{-1}(t)}{dt} = -F^{-1}(t) \frac{dF(t)}{dt} F^{-1}(t).$$

iv) Από το θεώρημα 8.1, προκύπτει ότι

$$\int_a^b F(t) dt = \left[\int_a^b f_{ij}(t) dt \right] \left[g_{ij}(t) \Big|_a^b \right],$$

όπου $g_{ij}(t)$ είναι ένα αόριστο ολοκλήρωμα της συνάρτησης $f_{ij}(t)$, $1 \leq i \leq m, 1 \leq j \leq n$.

Παράδειγμα 8.4. Έστω ο πίνακας

$$F(t) = \begin{bmatrix} t & \sin \pi t \\ 1 & -t \end{bmatrix}.$$

Τότε

$$\int_0^1 F(t) dt = \begin{bmatrix} \int_0^1 t dt & \int_0^1 \sin \pi t dt \\ \int_0^1 dt & \int_0^1 -t dt \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{2}{\pi} \\ 1 & -\frac{1}{2} \end{bmatrix}. \quad \blacksquare$$

Λήμμα 8.1 Αν η συνάρτηση πίνακα $F(t) = [f_{ij}(t)]$ είναι ολοκληρώσιμη στο διάστημα $[a, b]$, τότε

$$\left\| \int_a^b F(t) dt \right\|_I \leq \int_a^b \|F(t)\|_I dt.$$

Απόδειξη. Από τον ορισμό (N_1) της νόρμας $\|\cdot\|_I$ (βλέπε παράγραφο 4.3.2), έχουμε

$$\begin{aligned} \left\| \int_a^b F(t) dt \right\|_I &= \left| \sum_{i=1}^m \sum_{j=1}^n \int_a^b f_{ij}(t) dt \right| \\ &\leq \sum_{i=1}^m \sum_{j=1}^n \int_a^b |f_{ij}(t)| dt = \int_a^b \sum_{i=1}^m \sum_{j=1}^n |f_{ij}(t)| dt = \int_a^b \|F(t)\|_I dt. \quad \blacksquare \end{aligned}$$

8.2. Γραμμικές Διαφορικές Εξισώσεις.

Ονομάζουμε γραμμικό σύστημα διαφορικών εξισώσεων πρώτης τάξης κάθε σύστημα της μορφής:

$$\begin{aligned} x_1'(t) &= a_{11}(t)x_1(t) + \dots + a_{1j}(t)x_j(t) + \dots + a_{1n}(t)x_n(t) + g_1(t) \\ &\quad \vdots \\ x_i'(t) &= a_{i1}(t)x_1(t) + \dots + a_{ij}(t)x_j(t) + \dots + a_{in}(t)x_n(t) + g_i(t) \\ &\quad \vdots \\ x_n'(t) &= a_{n1}(t)x_1(t) + \dots + a_{nj}(t)x_j(t) + \dots + a_{nn}(t)x_n(t) + g_n(t) \end{aligned} \quad (8.1)$$

όπου οι συναρτήσεις $a_{ij}(t)$ και $g_i(t)$ είναι συνεχείς γνωστές συναρτήσεις, ορισμένες στο διάστημα $\Delta \subset \mathbb{R}$, όπου Δ είναι ένα διάστημα της μορφής $[a, b]$, (a, b) , $[a, b)$, με $a, b \in \overline{\mathbb{R}}$ και $x_1(t), x_2(t), \dots, x_n(t)$ είναι άγνωστες συναρτήσεις, τις οποίες θέλουμε να προσδιορίσουμε.

Αν οι συναρτήσεις $g_i(t)$, $i=1, 2, \dots, n$ δεν είναι ταυτοτικά μηδέν, τότε το σύστημα λέγεται **μη ομογενές**, ενώ αν είναι μηδενικές συναρτήσεις, τότε το σύστημα

$$\begin{aligned} x_1'(t) &= a_{11}(t)x_1(t) + \dots + a_{1n}(t)x_n(t) \\ &\quad \vdots \\ x_n'(t) &= a_{n1}(t)x_1(t) + \dots + a_{nn}(t)x_n(t) \end{aligned} \quad (8.2)$$

λέγεται **ομογενές**.

Τα συστήματα (8.1) και (8.2) σε διανυσματική μορφή γράφονται, αντίστοιχα:

$$\dot{\mathbf{x}}(t) = A(t)\mathbf{x}(t) + \mathbf{g}(t) \quad (8.1\alpha)$$

και

$$\dot{\mathbf{x}}(t) = A(t)\mathbf{x}(t) \quad (8.2\alpha)$$

όπου $A(t) = [a_{ij}(t)]$ είναι ένας τετραγωνικός πίνακας τάξης n ,

$$\mathbf{g}(t) = [g_1(t) \quad g_2(t) \quad \dots \quad g_n(t)]^T \quad \text{και} \quad \mathbf{x}(t) = [x_1(t) \quad x_2(t) \quad \dots \quad x_n(t)]^T$$

τα διανύσματα των γνωστών και άγνωστων συναρτήσεων, αντίστοιχα.

Η n -τάξης διαφορική εξίσωση

$$y^{(n)} + a_1(t)y^{(n-1)} + \dots + a_n(t)y = b(t), \quad (8.3)$$

όπου $y = y(t)$ είναι μία άγνωστη βαθμωτή συνάρτηση και $b(t)$, $a_i(t)$, $i=1,2,\dots,n$ είναι συνεχείς και γνωστές συναρτήσεις ορισμένες στο Δ . Η (8.3) είναι μία εξίσωση εξαιρετικού ενδιαφέροντος στις μαθηματικές και φυσικές εφαρμογές.

Θέτοντας

$$\begin{aligned} x_1(t) &= y(t) \\ x_2(t) &= y'(t) \\ &\vdots \\ x_n(t) &= y^{(n-1)}(t) \end{aligned}$$

η εξίσωση (8.3) μπορεί να γραφεί με τη μορφή (8.1α), όπου

$$A(t) = \begin{bmatrix} 0 & 1 & & & & & \\ & 0 & 1 & & & & \\ & & \ddots & \ddots & & & \\ \mathbb{O} & & & 0 & 1 & & \\ & & & & \ddots & \ddots & \\ & & & & & 0 & 1 \\ -a_n(t) & -a_{n-1}(t) & \dots & \dots & \dots & -a_2(t) & -a_1(t) \end{bmatrix}, \quad \mathbf{g}(t) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ b(t) \end{bmatrix}.$$

Αντίστροφα, κάθε σύστημα της μορφής (8.1α) μπορεί να μετατραπεί σε μία γραμμική διαφορική εξίσωση n -τάξης. Άρα, ένα σύστημα της μορφής (8.1) είναι ισοδύναμο με μία γραμμική διαφορική εξίσωση n -τάξης. Η πρόταση αυτή αποτελεί τη βάση της λεγόμενης «μεθόδου απαλοιφής» για τη λύση του συστήματος (8.1).

Επομένως, ένα σύστημα της μορφής (8.1) μπορεί να θεωρηθεί ως μία γενίκευση μίας γραμμικής διαφορικής εξίσωσης n -τάξης. Γι' αυτόν το λόγο, η θεωρία επίλυσης των συστημάτων αυτών αναπτύσσεται εντελώς ανάλογα με εκείνη των γραμμικών διαφορικών εξισώσεων n -τάξης, βλέπε Δ. Δασκαλόπουλος, [8], σελ. 315-325.

Παράδειγμα 8.5. Η διαφορική εξίσωση

$$y''(t) - t^2 y'(t) + e^t y(t) = t^3,$$

είναι ισοδύναμη με το σύστημα

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -e^t & t^2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ t^3 \end{bmatrix}. \quad \blacksquare$$

8.3. Θεμελιώδεις λύσεις.

Θα αρχίσουμε με το ομογενές γραμμικό σύστημα

$$\dot{\mathbf{x}}(t) = A(t)\mathbf{x}(t), \quad \mathbf{x}(t_0) = \mathbf{c} \quad (8.4)$$

όπου \mathbf{c} είναι ένα οποιοδήποτε γνωστό διάνυσμα.

Το σύστημα (8.4) λέγεται **γραμμικό**, επειδή ισχύει το παρακάτω θεώρημα.

Θεώρημα 8.2 Αν $\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_n(t)$ είναι λύσεις του ομογενούς συστήματος (8.4) στο διάστημα Δ , τότε, κάθε γραμμικός συνδυασμός αυτών

$$\mathbf{x}(t) = c_1 \mathbf{x}_1(t) + c_2 \mathbf{x}_2(t) + \dots + c_n \mathbf{x}_n(t)$$

είναι μία λύση του συστήματος (8.4), στο διάστημα Δ .

Απόδειξη.

$$\begin{aligned} A\mathbf{x}(t) &= A(c_1 \mathbf{x}_1(t) + \dots + c_n \mathbf{x}_n(t)) \\ &= c_1 A\mathbf{x}_1(t) + \dots + c_n A\mathbf{x}_n(t) \\ &= c_1 \dot{\mathbf{x}}_1(t) + \dots + c_n \dot{\mathbf{x}}_n(t) = (c_1 \mathbf{x}_1(t) + \dots + c_n \mathbf{x}_n(t))' = \dot{\mathbf{x}}(t). \quad \blacksquare \end{aligned}$$

Ορισμός 8.2 Λέμε ότι οι n λύσεις $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ του ομογενούς συστήματος (8.4) αποτελούν ένα **θεμελιώδες σύστημα λύσεων** του σε ένα διάστημα Δ , αν και μόνο αν οι λύσεις είναι γραμμικά ανεξάρτητες.

Αν μπορέσουμε να βρούμε μία συνάρτηση πίνακα n -τάξης, $X(t) = [x_{ij}(t)]$, που ικανοποιεί τη διαφορική εξίσωση

$$\frac{dX(t)}{dt} = A(t)X(t) \quad (8.5a)$$

και

$$X(t_0) = I \quad (8.5b)$$

τότε η $\mathbf{x}(t) = X(t)\mathbf{c}$ θα είναι μία λύση της εξίσωσης (8.4).

Πράγματι, θα έχουμε:

$$\dot{\mathbf{x}}(t) = \dot{X}(t)\mathbf{c} = A(t)X(t)\mathbf{c} = A(t)\mathbf{x}(t)$$

και

$$\mathbf{x}(t_0) = X(t_0)\mathbf{c} = I\mathbf{c} = \mathbf{c} . \quad \blacksquare$$

Μία συνάρτηση πίνακα $X(t)$, που ικανοποιεί τις εξισώσεις (8.5a) και (8.5b), λέγεται **θεμελιώδης πίνακας της διαφορικής εξίσωσης**

$$\frac{dX(t)}{dt} = A(t)X(t)$$

ή **θεμελιώδης πίνακας του $A(t)$** .

Λήμμα 8.2 (Λήμμα ύπαρξης)

Αν ο $n \times n$ πίνακας $A(t)$ είναι μία συνάρτηση πίνακα, συνεχής για κάθε $t \in \Delta$, τότε υπάρχει ένας θεμελιώδης πίνακας του $A(t)$, ορισμένος για κάθε $t \in \Delta$, δηλαδή, υπάρχει μία συνάρτηση πίνακα $X(t)$, που ικανοποιεί τις σχέσεις:

i) $\dot{X}(t) = A(t)X(t)$, για $t \in \Delta$.

ii) $X(t_0) = I$.

Απόδειξη. Θεωρούμε την επαναληπτική ακολουθία συναρτήσεων πινάκων:

$$X_{n+1}(t) = I + \int_{t_0}^t A(s)X_n(s)ds, \quad X(t_0) = I \quad (8.6)$$

και χρησιμοποιώντας το Λήμμα 8.1 αποδεικνύουμε ότι αυτή είναι μία ακολουθία Cauchy, για κάθε $t \in \Delta$.

Επομένως, η ακολουθία αυτή συγκλίνει σε μία συνάρτηση πίνακα n -τάξης $X(t)$.

Μπορούμε επιπλέον να αποδείξουμε ότι ο πίνακας $X(t)$ είναι συνεχής για κάθε $t \in \Delta$ και ότι ικανοποιεί τη σχέση

$$X(t) = I + \int_{t_0}^t A(s)X(s)ds .$$

Επομένως, $X(t_0) = I$ και $\frac{dX(t)}{dt} = AX(t)$ για κάθε $t \in \Delta$.

Άρα, η συνάρτηση πίνακα $X(t)$ είναι ο ζητούμενος πίνακας. ■

Λήμμα 8.3 (Λήμμα μοναδικότητας)

Αν $A(t)$ είναι η προηγούμενη συνάρτηση πίνακα

$$\frac{dZ}{dt} = A(t)Z(t), \text{ για κάθε } t \in \Delta$$

και

$$Z(t_0) = \mathbf{0},$$

τότε

$$Z(t) = \mathbf{0}, \text{ για κάθε } t \in \Omega.$$

Για περισσότερες λεπτομέρειες, βλέπε για παράδειγμα Ν. J. Pullman, σελ. 127-132.

Τα παραπάνω αποτελέσματα μας οδηγούν στη διατύπωση του επόμενου θεωρήματος.

Θεώρημα 8.3 Αν $A(t)$ είναι μία συνεχής συνάρτηση πίνακα n -τάξης στο διάστημα Δ , τότε υπάρχει μία και μοναδική συνάρτηση πίνακα n -τάξης, τέτοια ώστε:

$$\text{i) } \frac{dX(t)}{dt} = A(t)X(t), \text{ για κάθε } t \in \Delta. \quad (8.7)$$

$$\text{ii) } X(t_0) = I.$$

Ισοδύναμα, ο A έχει ένα μοναδικό θεμελιώδη πίνακα, αν ο A είναι συνεχής στο Δ .

Απόδειξη. ▫ Από το Λήμμα ύπαρξης 8.2 συμπεραίνεται ότι ο πίνακας $A(t)$ έχει ένα θεμελιώδη πίνακα.

▫ Για τη μοναδικότητα: αν $\frac{dY(t)}{dt} = A(t)Y(t)$, για κάθε $t \in \Delta$ και $Y(0) = I$, τότε ορίζουμε

$$Z(t) = X(t) - Y(t), \text{ για κάθε } t \in \Delta.$$

Ο $Z(t)$ ικανοποιεί τις υποθέσεις του Λήμματος μοναδικότητας (λήμμα 8.3).

Επομένως, $Z(t) = \mathbf{0} = X(t) - Y(t)$, για κάθε $t \in \Delta$, οπότε $Y(t) \equiv X(t)$. ■

Πόρισμα 8.1 Αν ισχύουν οι υποθέσεις του θεωρήματος 8.3, για κάθε δοσμένο σταθερό διάνυσμα \mathbf{c} , υπάρχει ένα μοναδικό διάνυσμα στήλης $\mathbf{x}(t)$, τέτοιο ώστε:

$$\frac{d\mathbf{x}(t)}{dt} = A(t)\mathbf{x}(t), \text{ για κάθε } t \in \Delta$$

$$\mathbf{x}(t_0) = \mathbf{c}.$$

Απόδειξη. ▫ Για την ύπαρξη χρησιμοποιούμε τη σχέση

$$\mathbf{x}(t) = X(t)\mathbf{c},$$

όπου $X(t)$ είναι ο θεμελιώδης πίνακας του $A(t)$.

▫ Για τη μοναδικότητα: αν $\mathbf{y}(t)$ είναι μία λύση, έστω

$$\mathbf{z}(t) = \mathbf{x}(t) - \mathbf{y}(t)$$

και $Z(t)$ είναι ο τετραγωνικός, του οποίου κάθε στήλη είναι μία λύση $\mathbf{z}(t)$, τότε

$$\frac{dZ(t)}{dt} = AZ(t) \text{ και } Z(t_0) = \mathbf{O}. \text{ Επομένως, } Z(t) = \mathbf{O}, \text{ για κάθε } t \in \Delta, \text{ οπότε ισχύει και}$$

$\mathbf{z}(t) = \mathbf{0}$, για κάθε $t \in \Delta$. Άρα, $\mathbf{y}(t) = \mathbf{x}(t)$, δηλαδή, η λύση είναι μοναδική. ■

Παράδειγμα 8.6. Αν ο A είναι σταθερός, τότε

$$X_0 = I$$

$$X_1 = I + \int_{t_0}^t A ds = I + A(t - t_0)$$

$$X_2 = I + \int_{t_0}^t A(I + A(s - t_0)) ds = I + A(t - t_0) + \frac{A^2(t - t_0)^2}{2}$$

⋮

$$X_n = I + A(t - t_0) + \frac{A^2(t - t_0)^2}{2} + \dots + \frac{A^n(t - t_0)^n}{n!} + \dots$$

Επομένως, αν ο πίνακας A είναι **σταθερός**, τότε $X = e^{A(t-t_0)}$. Οπότε, η λύση της εξίσωσης (8.4) είναι:

$$\mathbf{x}(t) = e^{A(t-t_0)} \mathbf{c}. \quad \blacksquare$$

Εφαρμογή 8.1 Θεωρούμε το σύστημα:

$$x' = 2x + y$$

$$y' = 3x + 4y$$

με $t_0 = 0$ και $-\infty < t < +\infty$.

Ο θεμελιώδης πίνακας του A είναι $X(t) = e^{At}$.

Ισοδύναμα μπορούμε να γράψουμε $X(t) = e^B$, όπου $B = At$.

Απόδειξη. Οι ιδιοτιμές του πίνακα A είναι:

$$\begin{vmatrix} 2-\lambda & 1 \\ 3 & 4-\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - 6\lambda + 5 = 0 \Rightarrow \lambda_1 = 1, \lambda_2 = 5.$$

Οπότε, οι ιδιοτιμές του πίνακα $B = At$ είναι

$$\lambda_1 = t \text{ και } \lambda_2 = 5t.$$

Θεωρούμε

$$f(\lambda) = e^\lambda, \quad r(\lambda) = a_1\lambda + a_0$$

□ Για $\lambda = t \Rightarrow f(t) = e^t, \quad r(t) = a_1t + a_0$

□ Για $\lambda = 5t \Rightarrow f(t) = e^{5t}$, $r(t) = 5a_1t + a_0$.

Οι συντελεστές a_1 και a_0 , υπολογίζονται από το σύστημα:

$$\begin{aligned} a_1t + a_0 &= e^t \\ 5a_1t + a_0 &= e^{5t} \end{aligned}$$

από όπου βρίσκουμε:

$$a_1 = \frac{e^t(e^{4t} - 1)}{4t}, \quad a_0 = \frac{e^t(5 - e^{4t})}{4}$$

Συνεπώς,

$$\begin{aligned} X(t) &= e^{At} = \frac{e^t(5 - e^{4t})}{4} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{e^t(e^{4t} - 1)}{4t} \begin{bmatrix} 2t & t \\ 3t & 4t \end{bmatrix} \\ \Rightarrow X(t) &= e^{At} = \begin{bmatrix} \frac{3}{4}e^t + \frac{e^{5t}}{4} & -\frac{1}{4}e^t + \frac{1}{4}e^{5t} \\ -\frac{3}{4}e^t + \frac{3}{4}e^{5t} & \frac{1}{4}e^t + \frac{3}{4}e^{5t} \end{bmatrix}, \quad X(0) = I. \end{aligned}$$

Επομένως, κάθε λύση $\varphi(t)$ δίνεται από τον τύπο

$$\varphi(t) = X(t)\varphi(0). \quad \blacksquare$$

Εφαρμογή 8.2 Θεωρούμε το σύστημα:

$$\begin{aligned} x' &= x - y \\ y' &= x + 3y \end{aligned}$$

με $t_0 = 0$ και $-\infty < t < +\infty$.

Ο θεμελιώδης πίνακας του A είναι $X(t) = e^{At}$.

Ισοδύναμα μπορούμε να γράψουμε $X(t) = e^B$, όπου $B = At$.

Απόδειξη. Οι ιδιοτιμές του πίνακα A είναι:

$$\begin{vmatrix} 1 - \lambda & -1 \\ 1 & 3 - \lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - 4\lambda + 4 = 0 \Rightarrow \lambda = 2, \text{ πολλαπλότητας } 2.$$

Οπότε, οι ιδιοτιμές του πίνακα $B = At$ είναι $\lambda = 2t$, με πολλαπλότητα δύο.

Θεωρούμε

$$\begin{aligned} f(\lambda) &= e^\lambda, & r(\lambda) &= a_1\lambda + a_0 \\ f'(\lambda) &= e^\lambda, & r'(\lambda) &= a_1. \end{aligned}$$

Άρα, για $\lambda = 2t$ προκύπτει το σύστημα :

$$\begin{aligned} 2ta_1 + a_0 &= e^{2t} \\ a_1 &= e^{2t} \end{aligned}$$

από το οποίο υπολογίζεται

$$a_1 = e^{2t} \quad \text{και} \quad a_0 = e^{2t} - 2te^{2t}.$$

Συνεπώς,

$$\begin{aligned} X(t) &= e^{At} = (e^{2t} - 2te^{2t}) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + e^{2t} \begin{bmatrix} t & -t \\ t & 3t \end{bmatrix} \\ \Rightarrow X(t) &= e^{At} = \begin{bmatrix} e^{2t} - te^{2t} & -te^{2t} \\ te^{2t} & e^{2t} + te^{2t} \end{bmatrix}, \quad X(0) = I. \end{aligned}$$

Επομένως, κάθε λύση $\boldsymbol{\varphi}(t)$ δίνεται από τον τύπο

$$\boldsymbol{\varphi}(t) = X(t)\boldsymbol{\varphi}(0). \quad \blacksquare$$

Σημείωση : (i) Κάθε στήλη $\mathbf{x}_j = [x_{1j} \quad x_{2j} \quad \dots \quad x_{nj}]^T$, $j = 1, 2, \dots, n$ είναι λύση του συστήματος

$$\dot{\mathbf{x}}(t) = A(t)\mathbf{x}(t),$$

η οποία ικανοποιεί την αρχική συνθήκη, $\mathbf{x}_j(t_0) = \mathbf{e}_j$. Οι λύσεις αυτές αποτελούν ένα θεμελιώδες σύστημα λύσεων του συστήματος αυτού.

Πράγματι, αν $\mathbf{x}(t) = \sum_{j=1}^n c_j \mathbf{x}_j = \mathbf{0}$, για κάθε $t \in \Delta$, με τους συντελεστές c_j όχι όλους μηδέν, τότε

$$\mathbf{x}(t_0) = \sum_{j=1}^n c_j \mathbf{x}_j(t_0) = \sum_{j=1}^n c_j \mathbf{e}_j = [c_1 \quad c_2 \quad \dots \quad c_n]^T = \mathbf{0},$$

Δηλαδή, $c_1 = c_2 = \dots = c_n = 0$, που είναι αντίθετο προς την υπόθεση ότι οι συντελεστές δεν είναι όλοι μηδέν.

Γι' αυτόν το λόγο, ο πίνακας $X(t)$ λέγεται θεμελιώδης πίνακας του συστήματος (8.7).

(ii) Για την ορίζουσα του πίνακα αυτού ισχύει η παρακάτω ταυτότητα του Jacobi:

$$W(t) = \det X(t) = \exp \left\{ \int_{t_0}^t \text{tr}[A(s)] ds \right\} \quad (8.8)$$

Για την απόδειξη, βλέπε για παράδειγμα N. J. Pullman, σελ. 140. ■

Επειδή η εκθετική συνάρτηση είναι πάντοτε διάφορη του μηδενός, συμπεραίνουμε το ακόλουθο πόρισμα ως άμεση συνέπεια της σχέσης (8.8).

Πόρισμα 8.2 Αν $A(t)$ είναι μία συνάρτηση πίνακα, συνεχής στο Δ , τότε ο θεμελιώδης πίνακας του $\frac{dX}{dt} = AX$ είναι αντιστρέψιμος, για κάθε $t \in \Delta$.

8.4. Μη Ομογενές Σύστημα Διαφορικών Εξισώσεων.

Για τη λύση του προβλήματος:

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= A\mathbf{x} + \mathbf{b} \\ \mathbf{x}(t) &= \mathbf{c} \end{aligned} \quad (8.9)$$

μπορούμε να χρησιμοποιήσουμε τη γνωστή μέθοδο της «μεταβολής των παραμέτρων».

Βρίσκουμε το θεμελιώδη πίνακα του συστήματος $\dot{X}(t) = A(t)X(t)$, $X(t_0) = I$, και στη συνέχεια, αναζητούμε μία λύση, $\mathbf{y}(t)$, έτσι ώστε το διάνυσμα $X(t)\mathbf{y}(t)$ να είναι μία λύση του συστήματος (8.9).

Υπενθυμίζουμε ότι, αν $\mathbf{b}(t) = \mathbf{0}$, τότε το διάνυσμα $X(t)\mathbf{c}$ είναι λύση του συστήματος (8.9). Γι' αυτόν το λόγο, προκύπτει η ονομασία της μεθόδου.

Έστω $\mathbf{x}(t) = X(t)\mathbf{y}(t)$, όπου $\mathbf{y}(t)$ είναι η προσδιοριστέα λύση,

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \frac{dX(t)}{dt}\mathbf{y}(t) + X(t)\frac{d\mathbf{y}(t)}{dt} = A(t)X(t)\mathbf{y}(t) + X(t)\frac{d\mathbf{y}(t)}{dt} \\ \dot{\mathbf{x}}(t) &= A(t)\mathbf{x}(t) + X(t)\dot{\mathbf{y}}(t). \end{aligned}$$

Άρα

$$\begin{aligned} A(t)\mathbf{x}(t) + X(t)\dot{\mathbf{y}}(t) &= A(t)\mathbf{x}(t) + \mathbf{b}(t) \Rightarrow \\ X(t)\dot{\mathbf{y}}(t) &= \mathbf{b}(t) \Rightarrow \\ \dot{\mathbf{y}}(t) &= X^{-1}(t)\mathbf{b}(t), \quad \text{πόρισμα 8.2,} \end{aligned}$$

από την οποία παίρνουμε

$$\mathbf{y}(t) = \int_{t_0}^t X^{-1}(s)\mathbf{b}(s)ds$$

και, επομένως, η λύση του προβλήματος (8.9) δίνεται από τον τύπο

$$\mathbf{x}(t) = X(t)\mathbf{c} + X(t) \int_{t_0}^t X^{-1}(s)\mathbf{b}(s)ds \quad (8.10)$$

Αν ο πίνακας A είναι σταθερός, τότε η λύση του συστήματος (8.9) δίνεται από τον τύπο:

$$\mathbf{x}(t) = e^{A(t-t_0)}\mathbf{c} + e^{At} \int_{t_0}^t e^{-As}\mathbf{b}(s)ds \quad (8.11)$$

Παράδειγμα 8.7. Να λυθεί το σύστημα:

$$\frac{d\mathbf{x}(t)}{dt} = \begin{bmatrix} 3 & -2 \\ 4 & -3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \text{ για } \mathbf{x}(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Λύση. Η λύση του συστήματος αυτού δίνεται από τον τύπο (8.11), με $t_0 = 0$,

$$\mathbf{c} = [0 \ 0]^T \text{ και } \mathbf{b} = [1 \ 1]^T,$$

$$\mathbf{x}(t) = e^{At} \int_0^t e^{-As} \begin{bmatrix} 1 \\ 1 \end{bmatrix} ds.$$

□ Υπολογισμός του e^{At} :

$$\begin{vmatrix} 3-\lambda & -2 \\ 4 & -3-\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 = 1 \Rightarrow \lambda_1 = -1, \lambda_2 = 1.$$

Άρα, οι ιδιοτιμές του πίνακα $B = At$ είναι

$$\lambda_1 = -t, \quad \lambda_2 = t.$$

Θεωρούμε

$$f(\lambda) = e^\lambda, \quad r(\lambda) = a_1\lambda + a_0.$$

Από τις σχέσεις

$$\begin{aligned} -a_1t + a_0 &= e^{-t} \\ a_1t + a_0 &= e^t \end{aligned}$$

υπολογίζουμε τους συντελεστές a_1 και a_0 :

$$a_0 = \frac{e^t + e^{-t}}{2} \quad \text{και} \quad a_1 = \frac{e^t - e^{-t}}{2t}.$$

Συνεπώς,

$$e^B = e^{At} = \frac{e^t + e^{-t}}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{e^t - e^{-t}}{2t} \begin{bmatrix} 3t & -2t \\ 4t & -3t \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2}e^t + \frac{1}{2}e^{-t} + \frac{3}{2}e^t - \frac{3}{2}e^{-t} & -e^t + e^{-t} \\ 2e^t - 2e^{-t} & \frac{1}{2}e^t + \frac{1}{2}e^{-t} - \frac{3}{2}e^t + \frac{3}{2}e^{-t} \end{bmatrix},$$

από όπου έχουμε:

$$e^{At} = \begin{bmatrix} 2e^t - e^{-t} & e^{-t} - e^t \\ 2e^t - 2e^{-t} & 2e^{-t} - e^t \end{bmatrix},$$

$$e^{-As} = \begin{bmatrix} 2e^{-s} - e^s & e^s - e^{-s} \\ 2e^{-s} - 2e^s & 2e^s - e^{-s} \end{bmatrix}.$$

Οπότε,

$$e^{-As} \mathbf{b} = \begin{bmatrix} 2e^{-s} - e^s & e^s - e^{-s} \\ 2e^{-s} - 2e^s & 2e^s - e^{-s} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2e^{-s} - e^s + e^s - e^{-s} \\ 2e^{-s} - 2e^s + 2e^s - e^{-s} \end{bmatrix} = \begin{bmatrix} e^{-s} \\ e^{-s} \end{bmatrix}$$

και

$$\int_0^t e^{-As} \mathbf{b} ds = \int_0^t \begin{bmatrix} e^{-s} \\ e^{-s} \end{bmatrix} ds = \begin{bmatrix} \int_0^t e^{-s} ds \\ \int_0^t e^{-s} ds \end{bmatrix} = \begin{bmatrix} 1 - e^{-t} \\ 1 - e^{-t} \end{bmatrix}$$

$$e^{At} \int_0^t e^{-As} \mathbf{b} ds = \begin{bmatrix} 2e^t - e^{-t} & e^{-t} - e^t \\ 2e^t - 2e^{-t} & 2e^{-t} - e^t \end{bmatrix} \begin{bmatrix} 1 - e^{-t} \\ 1 - e^{-t} \end{bmatrix} = \begin{bmatrix} e^t - 1 \\ e^t - 1 \end{bmatrix}$$

Άρα,

$$\mathbf{x}(t) = \begin{bmatrix} e^t - 1 \\ e^t - 1 \end{bmatrix} \quad \blacksquare$$

ΣΥΜΠΕΡΑΣΜΑ

Στη Διπλωματική αυτή εργασία έχει καταβληθεί προσπάθεια μετά τις βασικές έννοιες της Γραμμικής Άλγεβρας να αναπτυχθούν μερικά κεφάλαια της θεωρίας πινάκων. Τα θέματα που αναπτύσσονται έχουν εφαρμογές στην Αριθμητική Ανάλυση, στις Διαφορικές Εξισώσεις, στη θεωρία Πιθανοτήτων και Στατιστικής κ.α..

Στο τέλος της εργασίας αυτής δίνονται προγράμματα MATLAB τα οποία χρησιμοποιήθηκαν για την επαλήθευση κάποιων θεωρητικών αποτελεσμάτων.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Atkinson, K., An Introduction to Numerical Analysis. John Wiley and Sons.
- [2] Ayres, F., Matrices, Schaum's. Outlines Series.
- [3] Berberian, S.K., Lectures in Functional Analysis and Operator Theory. Springer-Verlag.
- [4] Boyce, W.E. and DiPrima, R.C., Στοιχειώδεις Διαφορικές Εξισώσεις και Προβλήματα Συνοριακών Τιμών. Πανεπιστημιακές Εκδόσεις. Ε.Μ.Π.
- [5] Brand, L., Μαθηματική Ανάλυση. (Advanced Calculus), Ε.Μ.Ε. 1984.
- [6] Brickel, I., Matrices and Vector Spaces. George Allen and Unwin. LTD.
- [7] Bronson, R., Matrix Operations. Schaum's. Outlines Series.
- [8] Δασκαλόπουλος, Δ. , Ανώτερα Μαθηματικά V. Διαφορικές Εξισώσεις. Αθήνα 1990.
- [9] Δασκαλόπουλος, Δ. , Εφαρμοσμένη Γραμμική Άλγεβρα Τεύχος Πρώτο. Αθήνα.
- [10] Δασκαλόπουλος, Δ., Τυπολόγιο Ανωτέρων Μαθηματικών. Αθήνα.
- [11] Demidovich, P. B. and Maron, A.I, Computational Mathematics. English Translation. Mir Publishers 1976.
- [12] DeVito, C.I., Function Analysis. Academic Press.
- [13] Δονάτος, Γ., και Αδάμ, Μ., Γραμμική Άλγεβρα. Θεωρία και Εφαρμογές. Εκδόσεις Gutenberg 2008.
- [14] Φελλούρης, Α., Γραμμική Άλγεβρα και Αναλυτική Γεωμετρία. Αθήνα 1988.
- [15] Fulks, W., Advanced Calculus. John Wiley and Sons 1961.
- [16] Fullerton, G.H., Mathematical Analysis. Editors/Alan Jeffery Iain T. Adamson.
- [17] Γαλανής, Ε., Εισαγωγή στην Πραγματική Ανάλυση. Συναρτήσεις μίας Μεταβλητής. Αθήνα.
- [18] Groetsch, C. W., Elements of Applicable Functional Analysis. Marcel Dekker, INC.
- [19] Hille, E., Methods in Classical and Functional Analysis Copyright 1972 by Addison-Wesley. Publishing Company, INC.

- [20] Καδιανάκης Ν. και Καρανάσιος, Σ., Γραμμική Άλγεβρα και Στοιχεία Αναλυτικής Γεωμετρίας. Αθήνα 1989.
- [21] Καδιανάκης Ν., Καρανάσιος Σ. και Φελλούρης, Α., Ανάλυση ΙΙ , Συναρτήσεις Πολλών Μεταβλητών. Αθήνα 1993.
- [22] Knopp, C., Elementw of the Theory of Functions. Dover Publ. Co. New York,1952.
- [23] Kreyszing E., Advanced Enginneering Mathematic 8th Edition . John Wiley and Sons.
- [24] Kreyszing E., Introductory Functional Analysis with Applications. John Wiley and Sons.
- [25] Κυριαζής , Α. Σ., Στοιχεία Απειροστικού Λογισμού, Συναρτήσεις μίας Μεταβλητής. Interbooks.
- [26] Liusternik and Sobolev, Elements of Functions Analysis. UNGAR.
- [27] Μαρουλάς, Ι., Γραμμική Άλγεβρα. Αθήνα.
- [28] Μπακόπουλος, Α και Χρυσοβέργης Ι., Εισαγωγή στην Αριθμητική Ανάλυση. Αθήνα.
- [29] Pitts, C. G. C., Itronduction to Metric Spaces. Editors/Alan Jeffery Iain T. Adamson.
- [30] Pryce, J. D., Basic Methods of Linear Functional Analysis Hutchinson University Library.
- [31] Pullman , N., J., Matrix Theory and its Applications. Marcel Dekker, INC.
- [32] Riesz, F. and Nagy, B. Sz., Function Analysis. UNGAR.
- [33] Σοφιανός Γ. και Τυχόπουλος, Ε., Αριθμητική Ανάλυση. Εκδόσεις: Α.Σταμούλης.
- [34] Strang, G., Γραμμική Άλγεβρα και Εφαρμογές, Πανεπιστημιακές Εκδόσεις Κρήτης.

ΠΑΡΑΡΤΗΜΑ

Χρήση της MATLAB για τη λύση προβλημάτων με την E.M.J.

Χρησιμοποιώντας τον κώδικα που ακολουθεί στη MATLAB, για έναν τετραγωνικό αντιστρέψιμο πίνακα A , αντίστοιχο ενός συστήματος της μορφής

$$Ax = b ,$$

από την (6.14) πρώτα κατασκευάζεται ο πίνακας U και εξετάζεται αν ο πίνακας A έχει διαγώνια υπεροχή. Στη συνέχεια, για προκαθορισμένη ακρίβεια, υπολογίζεται και τυπώνεται η λύση του συστήματος καθώς και ο αριθμός επαναλήψεων που χρειάστηκαν κατά τον υπολογισμό της λύσης.

```
function EMJ
A          (δίνεται ο αρχικός πίνακας)
b          (δίνεται το αρχικό διάνυσμα-στήλη)
er = 0.0001 (σφάλμα)
d=det(A);  (υπολογίζεται η ορίζουσα του πίνακα A)
[n,m]=size(A);
U=0;      (αρχικοποιείται ο πίνακας U)
for i=1:n
    for j=1:m
        U(i,j)=-A(i,j)/A(i,i); (υπολογίζονται τα στοιχεία του U)
    end
end
for i=1:n
    U(i,i)=0;
end;
k=norm(U,inf) (υπολογίζεται η νόρμα του U)
if k>=1
    E
end
```

```

for i=1:n
    c(i)=b(i)/A(i,i);      (υπολογισμός του διανύσματος στήλης c)
end
p=[];
x0=zeros(1,n);
for f=1:18      (ικανός αριθμός επαναλήψεων σχετικά μικρός)
    xm=x0;
    for i=1:n
        sr=U(i,:);
        zr= sr*xm';
        x(i)=zr+c(1,i) ;
    end
    dX=x'-xm';
    N=norm(dX,inf)
    if N<er
        Sx=x;
        f;
        p(f)=f;
        break;
    end
    x0=x;
end

Sx      ( τυπώνεται η λύση του συστήματος)
p      ( τυπώνεται το πλήθος επαναλήψεων που απαιτούνται)

```