ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ

ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

# *Συναρτήσεις Πινάκων και Εφαρμογές τους*

Διπλωματική Εργασία

ΠΑΠΑΒΑΣΙΛΕΙΟΥ ΙΩΑΝΝΗΣ

Επιβλέπων Καθηγητής: Ψαρράκος Παναγιώτης
Καθηγητής Ε.Μ.Π.

ΑΘΗΝΑ, Ιούνιος 2020

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ

ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Επιβλέπων Καθηγητής: Ψαρράκος Παναγιώτης
Καθηγητής Ε.Μ.Π.

Τριμελής Εξεταστική Επιτροπή:

Β. Κανελλόπουλος, Αν. Καθηγητής Ε.Μ.Π.

Γ. Σμυρλής, Αν. Καθηγητής Ε.Μ.Π.

Π. Ψαρράκος, Καθηγητής Ε.Μ.Π. (Επιβλέπων)

ΑΘΗΝΑ, Ιούνιος 2020

# Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον κο Παναγιώτη Ψαρράκο, για την πολύτιμη βοήθειά του στην επιλογή του θέματος, αλλά και για την ουσιαστική συμβολή του στην επίβλεψη και υλοποίηση της διπλωματικής μου εργασίας. Οι καίριες παρατηρήσεις και διορθώσεις του, καθώς και η καθοδήγησή του, συνέβαλαν στο να ολοκληρωθεί η εργασία χωρίς άγχος και με ιδιαίτερη ευχαρίστηση. Επίσης, θέλω να ευχαριστήσω για την παρουσία τους τα μέλη της εξεταστικής επιτροπής κο Β. Κανελλόπουλο και κο Γ. Σμυρλή. Τέλος, θέλω να ευχαριστήσω την οικογένεια μου, που μου παρείχε κάθε είδους υποστήριξη κατά τη διάρκεια των σπουδών μου.

# Περίληψη


Σκοπός της παρούσας διπλωματικής εργασίας, είναι η παρουσίαση της θεωρίας των συναρτήσεων πινάκων, με μια κομψή και συνοπτική μεθοδολογία. Υπάρχουν αρκετοί ισοδύναμοι τρόποι για τον ορισμό μιας συνάρτησης πίνακα $\boldsymbol{f(A)}$. Εμείς εστιάζουμε τη μελέτη μας σε τρεις συγκεκριμένους, οι οποίοι είναι ιδιαίτερου ενδιαφέροντος. Στην ανάλυσή μας, δίνουμε έμφαση στον σαφή ορισμό τους, αλλά και στις ιδιότητες που απορρέουν από τον καθένα. Επίσης, ασχολούμαστε και με πεδία, στα οποία η εφαρμογή των αποτελεσμάτων που θα αναπτύξουμε για τις συναρτήσεις πινάκων, είναι εξαιρετικά χρήσιμη και ταυτόχρονα, συμβάλλει στην δημιουργία νέων ιδεών.

Συγκεκριμένα, στο πρώτο κεφάλαιο καταγράφουμε μερικά εισαγωγικά στοιχεία από την ανάλυση πινάκων και τη γραμμική άλγεβρα, τα οποία είναι απαραίτητα για την στέρεη θεμελίωση των εννοιών που θα χρησιμοποιήσουμε στη συνέχεια.

Στο δεύτερο κεφάλαιο, καθορίζουμε τις συνθήκες, για τις οποίες μια συνάρτηση ορίζεται στο φάσμα ενός πίνακα και παρατηρούμε μερικές σημαντικές ιδιότητες των πολυώνυμων παρεμβολής. Συνδυάζοντας τα αποτελέσματα αυτών των δύο παραγράφων, ορίζουμε μια συνάρτηση πίνακα μέσω πολυωνύμων παρεμβολής, κάνοντας χρήση του τύπου $\boldsymbol{f(A) = p(A)}$.

Προχωρώντας στο τρίτο κεφάλαιο, αποδεικνύουμε δύο θεωρημάτα που αφορούν, τον ορισμό μιας συνάρτησης στο φάσμα ενός διαγώνιου υποπίνακα και την διατήρηση των ιδιοτήτων ομοιότητας για τις συναρτήσεις δύο όμοιων πινάκων, αντίστοιχα. Λαμβάνοντας υπόψη αυτά τα δύο θεωρήματα, αλλά και την κανονική μορφή Jordan, επισημαίνουμε ένα τρίτο σχετικό θεώρημα, με το οποίο μπορούμε να ορίσουμε μια συνάρτηση πίνακα, μέσω της κανονικής μορφής Jordan, χρησιμοποιώντας την σχέση $\boldsymbol{f(A) = P\mathbf{diag}[f(J_1), f(J_2), \ldots, f(J_t)]\, P^{-1}}$.

Στα επόμενα κεφάλαια, πραγματοποιούμε μια διάσπαση του πίνακα $\boldsymbol{f(A)}$, ώστε να αποκτήσουμε μια χρήσιμη και σαφή αναπαράστασή του σε θεμελιώδεις όρους του ίδιου του $\boldsymbol{A}$. Συγκεκριμένα, εισάγουμε στην ανάλυσή μας τους πίνακες-συντελεστές, οι οποίοι αποτελούν ένα εξαιρετικό εργαλείο, καθώς συντελούν στον υπολογισμό μιας συνάρτησης πίνακα. Επιπρόσθετα, ορίζουμε την επιλύουσα ενός πίνακα, $\boldsymbol{R_\lambda = (zI - A)^{-1}}$, η οποία αποτελεί μια συνάρτηση που ορίζεται πάνω στο φάσμα του $\boldsymbol{A}$ και μας επιτρέπει να εφαρμόζουμε αποτελέσματα της μιγαδικής ανάλυσης στη θεωρία πινάκων. Ταυτόχρονα, παρατηρούμε μερικές ακόμη ιδιότητες των συναρτήσεων πινάκων, μέσω της παρουσίασης ενός αριθμού κατάλληλων εφαρμογών.

Ο τελευταίος τρόπος ορισμού μιας συνάρτησης πίνακα, παρουσιάζεται στο έβδομο κεφάλαιο. Ειδικότερα, χρησιμοποιούμε το θεώρημα Cauchy για μια συνάρτηση, η οποία είναι συνεχής πάνω σε ένα περίγραμμα $\boldsymbol{L}$ και αναλυτική στο εσωτερικό του και για έναν πίνακα, του οποίου οι διακεκριμένες ιδιοτιμές βρίσκονται στο εσωτερικό του $\boldsymbol{L}$. Λαμβάνουμε με αυτόν τον τρόπο τον τύπο $\boldsymbol{f(A) = \dfrac{1}{2\pi i}\int_L f(\lambda) R_\lambda d\lambda}$.

Ολοκληρώνουμε την ανάλυσή μας με το έννατο κεφάλαιο και την αναφορά μας στο πεδίο των παρατηρήσιμων και ελέγξιμων συστημάτων, στο οποίο η χρήση των συναρτήσεων πινάκων, μας επιτρέπει να παρουσιάσουμε μερικές ενδιαφέρουσες ιδέες και αποτελέσματα από την θεωρία συστημάτων ελέγχου.

# Abstract


The purpose of this diploma thesis, is to present the theory of matrix functions, with an elegant and concise methodology. There are several equivalent ways to define a matrix function $\boldsymbol{f(A)}$. We focus our study on three specific, which are of particular interest. In our analysis, we emphasize in their explicit definition, but also in the properties that derive from each. In addition, we deal with fields, in which the implementation of the results we will develop for matrix functions, is extremely useful and at the same time contributes to the creation of new ideas.

Specifically, in the first chapter we record some introductory elements from matrix analysis and linear algebra, which are necessary for the solid foundation of the concepts we will use next.

In the second chapter, we determine the conditions, for which a function is defined on the spectrum of a matrix and we observe some important properties of polynomial interpolation. Combining the results of these two paragraphs, we define a matrix function through polynomial interpolation with the use of the formula $\boldsymbol{f(A) = p(A)}$.

Advancing to the third chapter, we prove two theorems that refer, to the definition of a function in the spectrum of a block-diagonal matrix and to the conservation of the similarity properties for the functions of two similar matrices, respectively. In view of these two theorems and the Jordan theorem, we highlight a related third theorem, with which we can define a matrix function, through Jordan forms, using the formula
$\boldsymbol{f(A) = P\mathrm{diag}[f(J_1), f(J_2), \ldots, f(J_t)]\, P^{-1}}$.

In the following chapters, we decompose the matrix $\boldsymbol{f(A)}$, in order to get a useful and explicit representation of it, in terms of fundamental properties of $\boldsymbol{A}$. In particular, we introduce in our analysis the component-matrices, which are considered as a great tool, as they contribute to the computation of a matrix function. Moreover, we define the resolvent of a matrix, $\boldsymbol{R_\lambda = (zI - A)^{-1}}$, which is a function defined on the spectrum of $\boldsymbol{A}$, that allows us to apply results of complex analysis in matrix theory. Simultaneously, we observe some further properties of matrix functions, through a number of appropriate applications.

Our ultimate way of defining a matrix function, is illustrated in the seventh chapter. Namely, we use the Cauchy theorem for a function, which is continuous in and on contour $\boldsymbol{L}$ and analytic within $\boldsymbol{L}$ and for a matrix, whose distinct eigenvalues are inside $\boldsymbol{L}$. This way we obtain the formula $\boldsymbol{f(A) = \frac{1}{2\pi i}\int_L f(\lambda) R_\lambda d\lambda}$.

We conclude our analysis with the ninth chapter and our reference to the field of observable and controllable systems, in which the use of matrix functions allows us to present some interesting ideas and results from the theory of systems and control.

# Περιεχόμενα

# Κεφάλαιο 1

# Εισαγωγικά Στοιχεία

Προκειμένου να αναπτύξουμε το κεφάλαιο που αφορά στις συναρτήσεις πινάκων, απαιτείται αρχικά να καταγράψουμε μερικούς βασικούς ορισμούς και θεωρήματα από τη γραμμική άλγεβρα και την ανάλυση πινάκων.

## 1.1 Ομοιότητα, Εικόνα και Πυρήνας Πίνακα

**Ορισμός 1.1.1.** *Δύο τετραγωνικοί πίνακες* $\boldsymbol{A}$ *και* $\boldsymbol{B}$ *λέγονται* ***όμοιοι****, συμβολίζουμε*

$$\boldsymbol{A} \approx \boldsymbol{B},$$

*αν υπάρχει ένας αντιστρέψιμος πίνακας* $\boldsymbol{P}$*, τέτοιος ώστε*

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}.$$

*Για τους όμοιους πίνακες* $\boldsymbol{A}$ *και* $\boldsymbol{B}$ *έχουμε τις εξής ιδιότητες:*

$$\begin{array}{ll}
(\alpha) & \det \boldsymbol{A} = \det \boldsymbol{B}. \\
(\beta) & \boldsymbol{A} - \lambda \boldsymbol{I} \approx \boldsymbol{B} - \lambda \boldsymbol{I}. \\
(\gamma) & \boldsymbol{A}^\top \approx \boldsymbol{B}^\top. \\
(\delta) & p(\boldsymbol{A}) \approx p(\boldsymbol{B}), \quad \text{όπου } p(\lambda) \text{ βαθμωτό πολυώνυμο.} \\
(\varepsilon) & \operatorname{rank} \boldsymbol{A} = \operatorname{rank} \boldsymbol{B}.
\end{array}$$

**Ορισμός 1.1.2.** *Το σύνολο όλων των διανυσμάτων* $\boldsymbol{x}$*, για τα οποία*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0},$$

*ονομάζεται* ***πυρήνας*** *του πίνακα* $\boldsymbol{A}$ *και συμβολίζεται με* $\ker \boldsymbol{A}$.

**Ορισμός 1.1.3.** *Αν* $\boldsymbol{A}$ *είναι ένας* $m \times n$ *πίνακας, τότε το σύνολο*

$$R(\boldsymbol{A}) = \operatorname{Im} \boldsymbol{A} \triangleq \{\boldsymbol{y} \in \mathcal{F}^m : \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}, \quad \text{για κάποια } \boldsymbol{x} \in \mathcal{F}^n\},$$

*λέγεται* ***εικόνα*** *του* $\boldsymbol{A}$.

**Ορισμός 1.1.4.** *Ένας πίνακας* $\boldsymbol{A} \in \mathcal{F}$ *είναι* ***απλός****, αν και μόνο αν είναι όμοιος με ένα διαγώνιο πίνακα.*

## 1.2 Γραμμικές Απεικονίσεις

**Ορισμός 1.2.1.** *Έστω* $\mathcal{L}_1$ *και* $\mathcal{L}_2$ *γραμμικοί χώροι πάνω σε ένα σώμα* $\mathcal{F}$*. Μια απεικόνιση* $\mathcal{T}: \mathcal{L}_1 \to \mathcal{L}_2$*, λέγεται ότι είναι μια **γραμμική απεικόνιση**, αν για κάθε δύο στοιχεία* $x_1$ *και* $x_2$ *από το* $L_1$ *και οποιοδήποτε βαθμωτό* $\alpha \in \mathcal{F}$*,*

$$T(x_1 + x_2) = T(x_1) + T(x_2),$$
$$T(\alpha x_1) = \alpha T(x_1).$$

**Ορισμός 1.2.2.** *Έστω* $T \in \mathcal{L}(\mathcal{S})$*. Ένας υπόχωρος* $\mathcal{S}_0 \subset \mathcal{S}$ *καλείται **αναλλοίωτος** υπό τον* $T$ *ή* $T$*-αναλλοίωτος, αν για κάθε* $x \in \mathcal{S}_0$*, η εικόνα* $T(x)$ *ανήκει επίσης στο* $\mathcal{S}_0$*.*

**Θεώρημα 1.2.1.** *Έστω* $T \in \mathcal{L}(\mathcal{S})$ *και* $\sigma(T) = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ *και έστω ότι* $\lambda_i$ *έχουν δείκτες* $m_i$*, για κάθε* $i$*. Τότε*

$$T = \sum_{i=1}^{k} T|_{\mathcal{G}_i} \quad \text{όπου} \quad \mathcal{G}_i = Ker((T - \lambda_i I)^{m_i}),$$

*είναι ένας* $T$*-αναλλοίωτος υπόχωρος.*

**Πρόταση 1.2.1.** *Έστω* $T \in \mathcal{L}(\mathcal{S})$ *με* $\mathcal{F} = \mathbb{C}$*. Υπάρχει μια βάση στο* $\mathcal{S}$*, στην οποία η αναπαράσταση* $A$ *της* $T$ *είναι ένας διαγώνιος υποπίνακας με μιγαδικά στοιχεία*

$$A = \mathrm{diag}\{A_1, A_2, \ldots, A_k\}$$

*και το φάσμα των* $A_i$ $(1 \leq i \leq k)$ *αποτελείται μόνο από ένα σημείο. Επιπλέον, αυτά τα σημεία είναι διακεκριμένα και αν* $q_i$ *είναι το μέγεθος του* $A_i$ $(1 \leq i \leq k)$ *τότε,* $q_1, q_2, \ldots, q_k$ *ορίζονται μοναδικά από την απεικόνιση* $T$*.*

**Ορισμός 1.2.3.** *Έστω* $T \in \mathcal{L}(\mathcal{S})$ *με* $\mathcal{F} = \mathbb{C}$*,* $\lambda \in \sigma(T)$ *και* $\mathcal{S}_r = \mathrm{Ker}\,(T - \lambda I)^r$ *όπου* $r$ *μεταβάλλεται πάνω στο σύνολο των μη αρνητικών ακεραίων. Τότε έχουμε, ότι*

$$\{0\} = \mathcal{S}_0 \subset \mathcal{S}_1 \subset \mathcal{S}_2 \subset \cdots \subset \mathcal{S}_p = \mathcal{S}_{p+1} = \cdots \subset \mathcal{S},$$

*για κάποιο θετικό ακέραιο* $p$*. Ο υπόχωρος* $\mathcal{S}_r$ $(1 \leq r \leq p)$ *της παραπάνω σχέσης καλείται **γενικευμένος ιδιόχωρος** της απεικόνισης* $T$ *τάξης* $1$ *και αντιστοιχεί στην ίδια ιδιοτιμή.*

**Ορισμός 1.2.4.** *Έστω* $x$ *ένα στοιχείο διάφορο του μηδενός, τέτοιο ώστε* $x \in \mathcal{S}_r$*, αλλά* $x \notin \mathcal{S}_{r-1}$*. Τότε το* $x$ *καλείται **γενικευμένο ιδιοδιάνυσμα** της απεικόνισης* $T$ *τάξης* $r$ *και αντιστοιχεί στην ιδιοτιμή* $\lambda$*.*

## 1.3 Ιδιοτιμές, Ιδιοδιανύσματα και Χαρακτηριστικό Πολυώνυμο Πινάκα

**Ορισμός 1.3.1.** *Έστω $\boldsymbol{A} \in \mathcal{F}^{n\times n}$. Ορίζουμε ένα μη-μηδενικό διάνυσμα $\boldsymbol{\alpha} \in \mathcal{F}^{n}$, να είναι ένα* ***ιδιοδιάνυσμα*** *του $\boldsymbol{A}$, που αντιστοιχεί στην ιδιοτιμή $\boldsymbol{\lambda}$, αν*

$$\boldsymbol{A\alpha} = \boldsymbol{\lambda\alpha}, \quad \boldsymbol{\alpha} \neq \mathbf{0}.$$

**Ορισμός 1.3.2.** *Το* ***φάσμα*** *ενός πίνακα $\boldsymbol{A} \in \mathcal{F}^{n\times n}$, γράφεται $\boldsymbol{\sigma(A)}$ και ορίζεται, ως το σύνολο όλων των διακεκριμένων ιδιοτιμών του $\boldsymbol{A}$.*

**Πρόταση 1.3.1.** *Οι όμοιοι πίνακες έχουν το ίδιο φάσμα.*

**Πρόταση 1.3.2.** *Το φάσμα $\boldsymbol{\sigma(A)}$ ενός πίνακα $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ δεν είναι κενό και αποτελείται από το πολύ $\boldsymbol{n}$ διακεκριμένες ιδιοτιμές. Αν $\boldsymbol{A}$ έχει $\boldsymbol{n}$ διακεκριμένες ιδιοτιμές, τότε είναι απλός.*

**Θεώρημα 1.3.1.** *Έστω $\boldsymbol{A} \in \mathcal{F}^{n\times n}$ είναι απλός πίνακας και έστω $\boldsymbol{D}$ είναι διαγώνιος πίνακας, ώστε*

$$\boldsymbol{A} = \boldsymbol{PDP^{-1}}, \quad \mathbf{det}\boldsymbol{P} \neq \mathbf{0}.$$

*Τότε*

$$\boldsymbol{D} = \mathbf{diag}[\boldsymbol{\lambda}_j]_{j=1}^{n},$$

*με $\boldsymbol{\lambda_j} \in \boldsymbol{\sigma(A)}$, $\quad \boldsymbol{j = 1, 2, \ldots, n}$ $\quad$ και $\quad \boldsymbol{P} = [\boldsymbol{x_1} \ \boldsymbol{x_2} \ \cdots \ \boldsymbol{x_n}]$, το οποίο $\boldsymbol{x_j}$ $(\boldsymbol{1 \leq j \leq n})$ είναι ένα ιδιοδιάνυσμα του $\boldsymbol{A}$, που αντιστοιχεί στην ιδιοτιμή $\boldsymbol{\lambda_j}$ $(\boldsymbol{1 \leq j \leq n})$.*

**Θεώρημα 1.3.2. (Φασματικό Θεώρημα)** *Έστω $\boldsymbol{A} \in \mathcal{F}^{n\times n}$ είναι ένας απλός πίνακας με ιδιοτιμές $\boldsymbol{\lambda_1, \ldots, \lambda_s}$ και σχετιζόμενα ιδιοδιανύσματα $\boldsymbol{x_1, x_2, \ldots, x_n}$. Τότε υπάρχουν αριστερά ιδιοδιανύσματα $\boldsymbol{y_1, y_2, \ldots, y_n}$, για τα οποία*

$$\boldsymbol{y_j^{\top} x_k} = \boldsymbol{\delta_{jk}}, \quad \boldsymbol{1 \leq j}, \quad \boldsymbol{k \leq n}$$

*και*

$$\boldsymbol{A} = \sum_{j=1}^{n} \boldsymbol{\lambda_j x_j y_j^{\top}}.$$

**Πρόταση 1.3.3.** *Μια ιδιοτιμή $\boldsymbol{\lambda} \in \boldsymbol{\sigma(A)}$, αν και μόνο αν ο πίνακας $\boldsymbol{\lambda I - A}$ δεν είναι αντιστρέψιμος, ή ισοδύναμα*

$$\mathbf{det}(\boldsymbol{\lambda I - A}) = \mathbf{0}.$$

**Ορισμός 1.3.3.** *Μπορούμε να γράψουμε*

$$\det(\lambda I - A) = c_0 + c_1\lambda + \cdots + c_{n-1}\lambda^{n-1} + c_n\lambda^n, \quad c_n = 1.$$

*Το πολυώνυμο*

$$c(\lambda) \triangleq \sum_{i=0}^{n} c_i\lambda^i,$$

*καλείται ως το* ***χαρακτηριστικό πολυώνυμο*** *του πίνακα* $A$.

**Πρόταση 1.3.4.** *Ένα βαθμωτό* $\lambda \in \mathcal{F}$ *είναι μια ιδιοτιμή του* $A \in \mathcal{F}^{n\times n}$*, αν και μόνο αν* $\lambda$ *είναι μια ρίζα του χαρακτηριστικού πολυωνύμου του* $A$.

**Πρόταση 1.3.5.** *Το χαρακτηριστικό πολυώνυμο* $c(\lambda)$ *του* $A$ *μπορεί να παραγοντοποιηθεί και να δώσει ένα γινόμενο από* $n$ *γραμμικούς παράγοντες*

$$c(\lambda) = \prod_{i=1}^{n}(\lambda - \lambda_i),$$

*όπου* $\lambda_i \in \sigma(A), \quad i = 1, 2, \ldots, n$.

**Ορισμός 1.3.5.** *Ορίζουμε την* ***αλγεβρική πολλαπλότητα*** *της ιδιοτιμής* $\lambda_i$ *του* $A$*, να είναι ο αριθμός των περιπτώσεων που ο παράγοντας* $(\lambda - \lambda_i)$ *προκύπτει στην παραπάνω σχέση. Με άλλα λόγια η αλγεβρική πολλαπλότητα του* $\lambda_i \in \sigma(A)$*, είναι η πολλαπλότητα του* $\lambda_i$ *σαν ρίζα του χαρακτηριστικού πολυωνύμου.*

**Ορισμός 1.3.6.** *Η* ***γεωμετρική πολλαπλότητα*** *μιας ιδιοτιμής* $\lambda_0$ *του* $A$ *ορίζεται, ως η διάσταση του υποχώρου* $\ker\,(A - \lambda_0 I)$.

**Ορισμός 1.3.7.** *Ορίζουμε ένα μη-μηδενικό βαθμωτό πολυώνυμο* $p(\lambda)$ *πάνω στον χώρο* $\mathcal{F}$*, να είναι ένα* ***πολυώνυμο μηδενισμού*** *για τον πίνακα* $A$*, αν*

$$p(A) = 0.$$

**Θεώρημα 1.3.3.** *Το ελάχιστο πολυώνυμο του* $A$ *διαιρεί όλα τα πολυώνυμα μηδενισμού του* $A$.

**Πρόταση 1.3.6.** *Το ελάχιστο πολυώνυμο του* $A$ *είναι μοναδικό.*

**Θεώρημα 1.3.4.** *Το σύνολο των διακεκριμένων ριζών του ελαχίστου πολυωνύμου του* $\boldsymbol{A}$ *συμπίπτει με το φάσμα του* $\boldsymbol{A}$.

**Πρόταση 1.3.7.** *Δύο όμοιοι πίνακες έχουν το ίδιο ελάχιστο πολυώνυμο.*

**Πρόταση 1.3.8.** *Έστω* $\boldsymbol{\mathcal{F}} = \mathbb{C}$ *και έστω πίνακας* $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n \times n}}$ *με* $\boldsymbol{\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_s\}}$. *Το ελάχιστο πολυώνυμο* $\boldsymbol{m(\lambda)}$ *του* $\boldsymbol{A}$ *δίνεται από την σχέση*

$$\boldsymbol{m(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{m_i}},$$

*για θετικούς ακεραίους* $\boldsymbol{m_i \; (i = 1, 2, \ldots, s)}$.

## 1.4 Κανονική Μορφή Jordan

**Ορισμός 1.4.1.** *Έστω* $\boldsymbol{x_r}$ *γενικευμένο ιδιοδιάνυσμα της απεικόνισης* $\boldsymbol{T}$ *τάξης* $\boldsymbol{r}$. *Τότε υπάρχουν διανύσματα* $\boldsymbol{x_{r-1}, \dots, x_2, x_1}$, *για τα οποία*

$$\begin{aligned} \boldsymbol{T(x_1)} &= \boldsymbol{\lambda x_1}, \\ \boldsymbol{T(x_2)} &= \boldsymbol{\lambda x_2 + x_1}, \\ &\vdots \\ \boldsymbol{T(x_r)} &= \boldsymbol{\lambda x_r + x_{r-1}}, \end{aligned}$$

*όπου* $\boldsymbol{x_j \in \mathcal{S}_j}$ *για* $\boldsymbol{j = 1, 2, \dots, r}$. *Μια τέτοια ακολουθία* $\boldsymbol{x_1, x_2, \dots, x_r}$ *καλείται* ***αλυσίδα Jordan μήκους*** $\boldsymbol{r}$ *και αντιστοιχεί στην ιδιοτιμή* $\boldsymbol{\lambda}$.

**Ορισμός 1.4.2** *Ορίζουμε τον* ***υπόχωρο Jordan,*** *για την απεικόνιση* $\boldsymbol{T}$, *ως τον υπόχωρο που δημιουργείται από τα στοιχεία μιας αλυσίδας Jordan, για την απεικόνιση* $\boldsymbol{T}$ *μήκους* $\boldsymbol{r}$ *και διάστασης* $\boldsymbol{r}$ *και συμβολίζουμε*

$$\boldsymbol{\mathcal{J}} = \mathrm{\mathbf{span}}\boldsymbol{\{x_1, x_2, \dots, x_r\}}.$$

**Ορισμός 1.4.3.** *Έστω* $\boldsymbol{A}$ *τετραγωνικός πίνακας. Ορίζουμε ως* ***Jordan-block*** *τον πίνακα*

$$\boldsymbol{J_i} = \begin{bmatrix} \boldsymbol{\lambda_i} & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\lambda_i} & \mathbf{1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \mathbf{1} \\ \mathbf{0} & \cdots & & \mathbf{0} & \boldsymbol{\lambda_i} \end{bmatrix},$$

*τάξης* $\boldsymbol{r}$ *που αντιστοιχεί στην ιδιοτιμή* $\boldsymbol{\lambda_i}$.

**Θεώρημα 1.4.1.** *Ορίζουμε, ως* ***κανονική μορφή Jordan*** *τον πίνακα*

$$\boldsymbol{J} = \mathrm{\mathbf{diag}}[\boldsymbol{J(\lambda_1), J(\lambda_2), \dots, J(\lambda_s)}],$$

*όπου για* $\boldsymbol{1 \le i \le s}$

$$\boldsymbol{J(\lambda_i)} = \mathrm{\mathbf{diag}}\left[\boldsymbol{J_p^{(i)}, \dots, J_p^{(i)}, J_{p-1}^{(i)}, \dots, J_{p-1}^{(i)}, \dots, J_1^{(i)}, \dots, J_1^{(i)}}\right],$$

*στην οποία ο* $\boldsymbol{j \times j}$ *πίνακας* $\boldsymbol{J_j^{(i)}}$ *εμφανίζεται* $\boldsymbol{t_j^{(i)}}$ *φορές. Οι αριθμοί* $\boldsymbol{t_j^{(i)} (i = 1, 2, \dots, s,\ j = 1, 2, \dots, p)}$ *καθορίζονται μοναδικά από την απεικόνιση* $T$.

**Πρόταση 1.4.1.** *Ο πίνακας* $\boldsymbol{A}$ *είναι όμοιος με μια κανονική μορφή Jordan, αν υπάρχει αντιστρέψιμος πίνακας* $\boldsymbol{P}$, *τέτοιος ώστε*

$$\boldsymbol{J = P^{-1} A P}.$$

## 1.5 Θεώρημα Chrystal

**Ορισμός 1.5.1** *Έστω ένα σύνολο από $\boldsymbol{n}$ ομογενείς βαθμωτές διαφορικές εξισώσεις, με σταθερούς μιγαδικούς συντελεστές σε $\boldsymbol{n}$ βαθμωτές μεταβλητές $\boldsymbol{x_1(t), \ldots, x_n(t)}$. Ορίζουμε $\boldsymbol{x(t)}$, ως*

$$\boldsymbol{x(t) = [x_1(t) \ \cdots \ x_n(t)]^\top}$$

*και έστω $\boldsymbol{l}$ είναι η ανώτατη τάξη παραγώγων στις $\boldsymbol{n}$ εξισώσεις. Τότε οι εξισώσεις μπορούν να γραφούν στη μορφή*

$$\boldsymbol{L_l x^{(l)}(t) + L_{l-1} x^{(l-1)}(t) + \cdots + L_1 x^{(1)}(t) + L_0 x(t) = 0}, \qquad \textbf{(1.11.1)}$$

*για συγκεκριμένους πίνακες $\boldsymbol{L_0, L_1, \ldots, L_l} \in \mathbb{C}^{\boldsymbol{n \times n}}$, όπου $\boldsymbol{L_l \neq 0}$ και οι δείκτες στο διάνυσμα $\boldsymbol{x(t)}$ συμβολίζουν τις παραγώγους των συντελεστών. Το πολυώνυμο πίνακα*

$$\boldsymbol{L(\lambda) = \sum_{j=0}^{l} \lambda^j L_j},$$

*της εξίσωσης $\textbf{(1.11.1)}$, μπορεί να γραφεί στη μορφή*

$$\boldsymbol{L\left(\frac{d}{dt}\right) x(t) = 0},$$

*συνεπώς*

$$\boldsymbol{L\left(\frac{d}{dt}\right) x(t) \triangleq \sum_{j=0}^{l} L_j \frac{\mathrm{d^j x}}{\mathrm{dt^j}}}.$$

**Θεώρημα 1.5.1. (Θεώρημα Chrystal)** *Αν $\mathbf{det}\ \boldsymbol{L(\lambda) \neq 0}$, τότε ο χώρος των λύσεων της εξίσωσης $\textbf{(1.11.1)}$ έχει διάσταση ίση με τον βαθμό του $\mathbf{det}\ \boldsymbol{L(\lambda)}$.*

# Κεφάλαιο 2

# Ορισμός μιας Συνάρτησης Πίνακα μέσω Πολυωνύμων Παρεμβολής

Στην ανάλυσή μας, θεωρούμε πίνακες $\boldsymbol{A}$ από τον χώρο $\mathbb{C}^{n\times n}$, που είναι το σύνολο όλων των τετραγωνικών πινάκων με διαστάσεις $\boldsymbol{n} \times \boldsymbol{n}$, των οποίων τα στοιχεία είναι μιγαδικοί αριθμοί (περιλαμβάνοντας και αυτούς που ανήκουν στον χώρο $\mathbb{R}^{n\times n}$, που συμβολίζει το σύνολο των τετραγωνικών πινάκων με διαστάσεις $\boldsymbol{n} \times \boldsymbol{n}$, των οποίων τα στοιχεία είναι πραγματικοί αριθμοί, σαν μια ειδική περίπτωση) και διερευνούμε το αν υπάρχει η δυνατότητα, να δώσουμε κάποιο νόημα στο $\boldsymbol{f(A)}$, όπου $\boldsymbol{f(\lambda)}$ είναι μια μιγαδική συνάρτηση, μιας μιγαδικής μεταβλητής $\boldsymbol{\lambda}$.

Θα θέλαμε ο ορισμός του $\boldsymbol{f(A)}$, να ισχύει για όσο το δυνατόν περισσότερες κατηγορίες συναρτήσεων $\boldsymbol{f(\lambda)}$. Έχουμε παρατηρήσει, ότι το παραπάνω ερώτημα επιλύεται εύκολα, αν

$$\boldsymbol{f(\lambda) = p(\lambda)},$$

είναι ένα πολυώνυμο πάνω στους μιγαδικούς αριθμούς. Επομένως,

$$\boldsymbol{p(A) \triangleq \sum_{i=1}^{l} p_i A^i},$$

αν

$$\boldsymbol{p(\lambda) = \sum_{i=0}^{l} p_i \lambda^i}.$$

Επιπλέον, αν $\boldsymbol{m(\lambda)}$ είναι το ελάχιστο πολυώνυμο για τον πίνακα $\boldsymbol{A}$, τότε για ένα πολυώνυμο $\boldsymbol{p(\lambda)}$ υπάρχουν πολυώνυμα $\boldsymbol{q(\lambda)}$ και $\boldsymbol{r(\lambda)}$, τέτοια ώστε

$$\boldsymbol{p(\lambda) = m(\lambda)q(\lambda) + r(\lambda)},$$

όπου το $\boldsymbol{r(\lambda)}$ είναι το μηδενικό πολυώνυμο, ή έχει βαθμό μικρότερο από αυτόν του $\boldsymbol{m(\lambda)}$. Συνεπώς, αφού $\boldsymbol{m(A) = 0}$, έχουμε ότι $\boldsymbol{p(A) = r(A)}$.

Οι πιο γενικές συναρτήσεις $\boldsymbol{f(\lambda)}$ που θα θεωρήσουμε, διατηρούν αυτή την ιδιότητα, ήτοι, δοθέντος $\boldsymbol{f(\lambda)}$ και $\boldsymbol{A}$, υπάρχει ένα πολυώνυμο $\boldsymbol{r(\lambda)}$, (με βαθμό μικρότερο από αυτόν του ελαχίστου πολυωνύμου του $\boldsymbol{A}$) τέτοιο ώστε

$$\boldsymbol{f(A) = r(A)}.$$

Είναι αναπάντεχο, αλλά αυτός ο περιορισμός αφήνει σημαντική ελευθερία στον χαρακτηρισμό των συναρτήσεων, που μπορούν να προσαρμοστούν.

## 2.1 Συναρτήσεις Ορισμένες στο Φάσμα ενός Πίνακα

Έστω $A \in \mathbb{C}^{n\times n}$ και υποθέτουμε, ότι $\lambda_1, \lambda_2, \ldots, \lambda_s$ είναι οι διακεκριμένες ιδιοτιμές του πίνακα $A$, έτσι ώστε

$$m(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_s)^{m_s}, \tag{2.1.1}$$

να είναι το ελάχιστο πολυώνυμο του $A$ με βαθμό

$$m = m_1 + m_2 + \cdots + m_s.$$

Γνωρίζουμε ότι η πολλαπλότητα $m_k$ του $\lambda_k$ σαν μια ρίζα του ελαχίστου πολυωνύμου αναφέρεται, ως ο δείκτης της ιδιοτιμής $\lambda_k$ και είναι ίση με τον μέγιστο βαθμό των στοιχειωδών διαιρετών που σχετίζονται με το $\lambda_k$ $(1 \leq k \leq s)$.

Δοθέντος συνάρτησης $f(\lambda)$, λέμε ότι ***αυτή ορίζεται στο φάσμα του πίνακα*** $A$, αν οι αριθμοί

$$f(\lambda_k), f'(\lambda_k), \ldots, f^{(m_k-1)}(\lambda_k), \quad k = 1, 2, \ldots, s, \tag{2.1.2}$$

που καλούνται τιμές της $\lambda_k$ στο φάσμα του $A$, υπάρχουν. Είναι προφανές, πως κάθε πολυώνυμο στον χώρο $\mathbb{C}$, ορίζεται στο φάσμα οποιουδήποτε πίνακα που ανήκει στον χώρο $\mathbb{C}^{n\times n}$.

**Πρόταση 2.1.1.** *Αν $p_1(\lambda)$ και $p_2(\lambda)$ είναι πολυώνυμα στον χώρο $\mathbb{C}$ και $A \in \mathbb{C}^{n\times n}$, τότε*

$$p_1(A) = p_2(A),$$

*αν και μόνο αν τα πολυώνυμα $p_1(\lambda)$ και $p_2(\lambda)$ έχουν τις ίδιες τιμές στο φάσμα του $A$.*

**Απόδειξη.** Αν θεωρήσουμε, ότι

$$p_1(A) = p_2(A) \quad \text{και} \quad p_0(\lambda) = p_1(\lambda) - p_2(\lambda),$$

τότε είναι προφανές, ότι

$$p_0(A) = 0$$

και έτσι $\boldsymbol{p_0(\lambda)}$ είναι ένα πολυώνυμο μηδενισμού για τον $\boldsymbol{A}$. Επομένως, (Θεώρημα ***1.3.3***) το $\boldsymbol{p_0(\lambda)}$ μπορεί να διαιρεθεί από το ελάχιστο πολυώνυμο $\boldsymbol{m(\lambda)}$ του πίνακα $\boldsymbol{A}$ που δίνεται στην εξίσωση **(2.1.1)** και υπάρχει ένα πολυώνυμο $\boldsymbol{q(\lambda)}$, τέτοιο ώστε

$$\boldsymbol{p_0(\lambda) = q(\lambda)m(\lambda)}.$$

Υπολογίζοντας τις τιμές του $\boldsymbol{p_0(\lambda)}$ στο φάσμα του $\boldsymbol{A}$ και γνωρίζοντας ότι οι τιμές του ελάχιστου πολυωνύμου του $\boldsymbol{A}$ στο φάσμα του $\boldsymbol{A}$ είναι όλες ίσες με μηδέν, προκύπτει εύκολα ότι

$$\boldsymbol{p_1^{(j)}(\lambda_k) - p_2^{(j)}(\lambda_k) = p_0^{(j)}(\lambda_k) = 0}, \qquad \textbf{(2.1.3)}$$

για $\boldsymbol{j = 0, 1, \ldots, m_k - 1, \quad 1 \leq k \leq s}$.
Έτσι, τα δύο πολυώνυμα $\boldsymbol{p_1(\lambda)}$ και $\boldsymbol{p_2(\lambda)}$ έχουν τις ίδιες τιμές στο φάσμα του $\boldsymbol{A}$, δεδομένου ότι

$$\boldsymbol{p_1(A) = p_2(A)}.$$

Αντίστροφα, η εξίσωση **(2.1.3)** ισχύει, τότε το $\boldsymbol{p_0(\lambda)}$ έχει μία ρίζα πολλαπλότητας $\boldsymbol{m_k}$ στο σημείο $\boldsymbol{\lambda_k}$, για κάθε $\boldsymbol{k = 1, 2, \ldots, s}$. Συνεπώς, το $\boldsymbol{p_0(\lambda)}$ πρέπει να διαιρείται από το $\boldsymbol{m(\lambda)}$, όπως ορίζεται στην εξίσωση **(2.1.1)** , το οποίο συνεπάγεται ότι

$$\boldsymbol{p_0(A) = 0},$$

(Πρόταση ***1.3.6***), ή ισοδύναμα

$$\boldsymbol{p_1(A) = p_2(A)}. \quad \blacksquare$$

Είναι αυτή η ιδιότητα των πολυωνύμων με ορίσματα πινάκων που χρησιμοποιούμε, για να διευρύνουμε τον ορισμό του $\boldsymbol{f(A)}$ πάνω σε περισσότερες γενικευμένες συναρτήσεις $\boldsymbol{f(\lambda)}$.

Συνεπώς, θα απαιτήσουμε όλες οι συναρτήσεις που ορίζονται στο φάσμα του πίνακα $\boldsymbol{A}$ και που λαμβάνουν τις ίδιες τιμές, να αποδίδουν τον ίδιο πίνακα $\boldsymbol{f(A)}$. Συγκεκριμένα, για οποιαδήποτε συνάρτηση $\boldsymbol{f(\lambda)}$ που ορίζεται στο φάσμα του $\boldsymbol{A}$, θα μπορούμε να γράφουμε

$$\boldsymbol{f(A) = p(A)},$$

όπου $\boldsymbol{p(\lambda)}$ είναι ένα πολυώνυμο με τις ίδιες τιμές στο φάσμα του $\boldsymbol{A}$.

Η ύπαρξη ενός τέτοιου πολυωνύμου $\boldsymbol{p(\lambda)}$ με τις προκαθορισμένες ιδιότητες, προκύπτει από την επίλυση του προβλήματος της γενικευμένης ερμιτιανής παρεμβολής, με την οποία θα ασχοληθούμε στην επόμενη παράγραφο.

## 2.2 Πολυώνυμα Παρεμβολής

**Πρόταση 2.2.1.** *Δοθέντων διακεκριμένων αριθμών* $\lambda_1, \lambda_2, \ldots, \lambda_s$, *θετικών ακεραίων* $m_1, m_2, \ldots, m_s$ *με*

$$m = \sum_{k=1}^{s} m_k,$$

*και μιας σειράς αριθμών*

$$f_{k,0},\ f_{k,1}\ ,\ldots,\ f_{k,m_{k-1}},\quad k = 1, 2, \ldots, s,$$

*υπάρχει ένα πολυώνυμο* $p(\lambda)$ *βαθμού μικρότερου του* $m$, *τέτοιο ώστε*

$$p(\lambda_k) = f_{k,0},\quad p^{(1)}(\lambda_k) = f_{k,1},\quad \ldots,\quad p^{m_k-1}(\lambda_k) = f_{k,m_k-1}, \tag{2.2.1}$$

για $k = 1, 2, \ldots, s$.

**Απόδειξη.** Φαίνεται εύκολα, ότι το πολυώνυμο

$$p_k(\lambda) = \alpha_k(\lambda)\psi_k(\lambda),^{\dagger}\quad 1 \leq k \leq s,$$

$$\alpha_k(\lambda) = \alpha_{k,0} + \alpha_{k,1}(\lambda - \lambda_k) + \cdots + \alpha_{k,m_{k-1}}(\lambda - \lambda_k)^{m_k-1},$$

$$\psi_k(\lambda) = \prod_{j=1,\ j\neq k}^{s} (\lambda - \lambda_j)^{m_j},$$

έχει βαθμό μικρότερου του $m$ και ικανοποιεί τις συνθήκες

$$p_k(\lambda_i) = p_k^{(1)}(\lambda_i) = \cdots = p_k^{(m_i-1)}(\lambda_i) = 0,$$

για $i \neq k$ και αυθαίρετα $\alpha_{k,0},\ \alpha_{k,1},\ \ldots,\ \alpha_{k,m_k-1}$. Συνεπώς, το πολυώνυμο

$$p(\lambda) = p_1(\lambda) + p_2(\lambda) + \cdots + p_s(\lambda), \tag{2.2.2}$$

ικανοποιεί τις συνθήκες στην εξίσωση **(2.2.1)**, αν και μόνο αν

$$p_k(\lambda_k) = f_{k,0},\quad p_k^{(1)}(\lambda_k) = f_{k,1},\quad \ldots,\quad p_k^{(m_k-1)}(\lambda_k) = f_{k,m_k-1}, \tag{2.2.3}$$

για κάθε $1 \leq k \leq s$.

† Αν $s = 1$, από τον ορισμό του, $\psi_1(\lambda) \equiv 1$.

Διαφορίζοντας, παίρνουμε

$$p_k^{(j)}(\lambda_k) = \sum_{i=0}^{j} \binom{j}{i} \alpha_k^{(i)}(\lambda_k) \psi_k^{(j-i)}(\lambda_k),$$

για $1 \leq k \leq s, \quad 0 \leq j \leq m_k - 1$.

Χρησιμοποιώντας την εξίσωση $(2.2.3)$ και ανακαλώντας τον ορισμό του $\alpha_k(\lambda)$, έχουμε για $k = 1, 2, \ldots, s, \quad j = 0, 1, \ldots, m_k - 1$,

$$f_{k,\,j} = \sum_{i=0}^{j} \binom{j}{i} i!\, \alpha_{k,i}\, \psi_k^{(j-i)}(\lambda_k). \tag{2.2.4}$$

Αφού $\psi_k(\lambda_k) \neq 0$ για κάθε σταθερό $k$, η εξίσωση $(2.2.4)$ μπορεί να επιλυθεί διαδοχικά, (ξεκινώντας $j = 0$) για να βρούμε τους σταθερούς συντελεστές $\alpha_{k,0}, \ldots, \alpha_{k,m_k-1}$, για τους οποίους η εξίσωση $(2.2.3)$ ισχύει. Συνεπώς, ένα πολυώνυμο $p(\lambda)$ της μορφής που δίνεται στην εξίσωση $(2.2.2)$, ικανοποιεί τις απαιτούμενες συνθήκες. ■

**Πρόταση 2.2.2.** *Το πολυώνυμο $p(\lambda)$ της Πρότασης **2.2.1** είναι μοναδικό.*

Το μοναδικό πολυώνυμο βαθμού μικρότερου του $m$ που ικανοποιεί τις συνθήκες της εξίσωσης $(2.2.1)$, καλείται ***ερμιτιανό πολυώνυμο παρεμβολής.***

Αυτό είναι ένα σημαντικό βήμα στην επιχειρηματολογία μας, αλλά η απόδειξη είναι πολύπλοκη και για αυτό το λόγο θα παραληφθεί. Μπορεί να βασιστεί στην αντιστρεψιμότητα ενός γενικευμένου πίνακα ***Vandermonde.***

Αν για τις συνθήκες της εξίσωσης $(2.2.1)$, ισχύει ότι

$$m_1 = m_2 = \cdots = m_s = 1,$$

τότε παίρνουμε

$$p(\lambda_k) = f_{k,0}, \quad \text{για} \quad k = 1, 2, \ldots, s.$$

Αυτή τη συγκεκριμένη μορφή της εξίσωσης $(2.2.1)$, επαληθεύει το ***πολυώνυμο παρεμβολής Lagrange***

$$p(\lambda) = \sum_{k=1}^{s} f_{k,0} \frac{(\lambda - \lambda_1) \cdots (\lambda - \lambda_{k-1})(\lambda - \lambda_{k+1}) \cdots (\lambda - \lambda_s)}{(\lambda_k - \lambda_1) \cdots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1}) \cdots (\lambda_k - \lambda_s)}. \tag{2.2.5}$$

Σημαντικό ρόλο παίζουν εκείνα τα πολυώνυμα, για τα οποία οι αριθμοί $f_{k,i}$ της εξίσωσης $(2.2.1)$, είναι όλοι ίσοι με μηδέν εκτός ενός, τον οποίο συμβολίζουμε με $f_{k,j}$ και ισούται με την μονάδα. Συνεπώς, οι Προτάσεις ***2.2.1*** και ***2.2.2*** εξασφαλίζουν την ύπαρξη ενός μοναδικού πολυωνύμου, $\varphi_{kj}(\lambda)$ βαθμού μικρότερου του $\mathbf{m}$, (με $1 \leq k \leq s$ και $0 \leq j \leq m_k - 1$) τέτοιο ώστε

$$\varphi_{kj}^{(r)}(\lambda_k) = \delta_{jr}, \quad r = 0, 1, \ldots, m_k - 1 \tag{2.2.6}$$

και όταν $i \neq k$,

$$\varphi_{kj}^{(r)}(\lambda_i) = 0, \quad r = 0, 1, \ldots, m_i - 1. \tag{2.2.7}$$

Αυτά τα $m$ πολυώνυμα καλούνται ***θεμελιώδη ή απόλυτα*** πολυώνυμα του προβλήματος παρεμβολής.

Αν αυτά τα πολυώνυμα είναι γνωστά, η λύση κάθε προβλήματος που ορίζεται από τις γενικές συνθήκες της εξίσωσης $(2.2.1)$, μπορεί να εκφραστεί ως ένας γραμμικός συνδυασμός αυτών. Αν $p(\lambda)$ ικανοποιεί την εξίσωση $(2.2.1)$, τότε επαληθεύεται εύκολα, ότι

$$p(\lambda) = \sum_{k=1}^{s} \sum_{j=0}^{m_k-1} f_{k,\,j}\, \varphi_{kj}(\lambda). \tag{2.2.8}$$

## 2.3 Μερικές Εφαρμογές στις Συναρτήσεις Πινάκων

Χρησιμοποιώντας τις έννοιες και τα αποτελέσματα των προηγούμενων παραγράφων, ένας γενικός ορισμός του πίνακα $\boldsymbol{f(A)}$ σε όρους του $\boldsymbol{A}$ και της συνάρτησης $\boldsymbol{f}$ μπορεί να δημιουργηθεί, υπό την προϋπόθεση, ότι η συνάρτηση $\boldsymbol{f(\lambda)}$ ορίζεται στο φάσμα του πίνακα $\boldsymbol{A}$. Αν $\boldsymbol{p(\lambda)}$ ένα οποιοδήποτε πολυώνυμο που παίρνει τις ίδιες τιμές με αυτές της $\boldsymbol{f(\lambda)}$ στο φάσμα του πίνακα $\boldsymbol{A}$, ορίζουμε απλά

$$\boldsymbol{f(A)} \triangleq \boldsymbol{p(A)}.$$

Η Πρόταση ***2.2.1*** μας διαβεβαιώνει, ότι ένα τέτοιο πολυώνυμο υπάρχει, ενώ η Πρόταση ***2.2.2*** μας δείχνει, ότι για τον σκοπό του τυπικού ορισμού, η επιλογή του $\boldsymbol{p(\lambda)}$ δεν είναι σημαντική. Επιπλέον, η εξίσωση **(2.2.8)** υποδεικνύει πως το $\boldsymbol{p(\lambda)}$ θα μπορούσε να επιλεχθεί με τον μικρότερο δυνατό βαθμό.

**Παράδειγμα 2.3.1.**

Αν $\boldsymbol{A}$ είναι ένας απλός πίνακας με διακεκριμένες ιδιοτιμές $\boldsymbol{\lambda_1, \dots, \lambda_s}$ και $\boldsymbol{f(\lambda)}$ είναι οποιαδήποτε συνάρτηση που είναι καλά ορισμένη στις ιδιοτιμές του πίνακα $\boldsymbol{A}$, τότε μπορούμε να γράψουμε το $\boldsymbol{f(A)}$ με έναν όμορφο τύπο. Πρώτα, παρατηρούμε ότι το ερμιτιανό πολυώνυμο παρεμβολής $\boldsymbol{p}$, για έναν τυχαίο πίνακα $\boldsymbol{A}$, δίνεται αναλυτικά από την φόρμουλα Lagrange-Hermite.

$$\boldsymbol{p(\lambda) = \sum_{k=1}^{s} \left[ \left( \sum_{j=0}^{m_k-1} \frac{1}{j!} \varphi_i^j(\lambda_i)(\lambda - \lambda_k)^j \right) \prod_{j \neq k} (\lambda - \lambda_j)^{m_j} \right]},$$

όπου

$$\boldsymbol{\varphi_k(\lambda) = \frac{f(\lambda)}{\prod\limits_{j \neq k} (\lambda - \lambda_j)^{m_j}}} \quad .$$

Για έναν απλό πίνακα με διακεκριμένες ιδιοτιμές, ο τύπος μετατρέπεται στην απλή μορφή του Lagrange

$$\boldsymbol{p(\lambda) = \sum_{k=1}^{s} f(\lambda_k) \prod_{j=1,\, j \neq k}^{s} \left( \frac{\lambda - \lambda_j}{(\lambda_k - \lambda_j)} \right)}$$

και έτσι έχοντας ορίσει, ότι

$$\boldsymbol{f(A)} \triangleq \boldsymbol{p(A)},$$

παίρνουμε

$$\boldsymbol{f(A) = \sum_{k=1}^{s} f(\lambda_k) \prod_{j=1,\, j \neq k}^{s} \left( \frac{A - \lambda_j / I}{(\lambda_k - \lambda_j)} \right)}. \quad \square$$

**Παράδειγμα 2.3.2.**

Χρησιμοποιώντας τους ορισμούς και τις προτάσεις που καταγράψαμε στις προηγούμενες παραγράφους, είμαστε σε θέση να υπολογίσουμε το $\boldsymbol{f(A)}$ για συγκεκριμένες γνωστές μας συναρτήσεις.

Έστω ότι $\quad \boldsymbol{f(\lambda) = e^{2\lambda}} \quad$ και $\quad \boldsymbol{A = \begin{bmatrix} 6 & -1 \\ 3 & 2 \end{bmatrix}}$.

Συγκεκριμένα, για τον υπολογισμό θα χρησιμοποιήσουμε, ότι

$$\boldsymbol{f(A) \triangleq p(A)}.$$

Οι ιδιοτιμές του $\boldsymbol{f(A)}$ είναι οι $\boldsymbol{\lambda_1 = 3}$, $\boldsymbol{\lambda_2 = 5}$, άρα το ελάχιστο πολυώνυμο $\boldsymbol{m(\lambda)}$ του $\boldsymbol{A}$ είναι

$$\boldsymbol{m(\lambda) = (\lambda - 3)(\lambda - 5)}.$$

Το πολυώνυμο παρεμβολής του Lagrange είναι

$$\begin{aligned} \boldsymbol{p(\lambda)} \quad &\boldsymbol{= f(\lambda_1)\frac{\lambda - \lambda_2}{\lambda_1 - \lambda_2} + f(\lambda_2)\frac{\lambda - \lambda_1}{\lambda_2 - \lambda_1}} \\ &\boldsymbol{= -\frac{1}{2}e^{6}(\lambda - 5) + \frac{1}{2}e^{10}(\lambda - 3)}. \end{aligned}$$

Συνεπώς, από τον ορισμό,

$$\begin{aligned} \boldsymbol{e^{2A}} &\boldsymbol{= -\frac{1}{2}e^{6}(A - 5I) + \frac{1}{2}e^{10}(A - 3I)} \\ &\boldsymbol{= \frac{1}{2}\begin{bmatrix} 3e^{10} - e^{6} & -e^{10} + e^{6} \\ 3e^{10} - 3e^{6} & -e^{10} + 3e^{6} \end{bmatrix}}. \quad \square \end{aligned}$$

**Παρατήρηση 2.3.1.** Αν $\boldsymbol{f(\lambda)}$ και $\boldsymbol{g(\lambda)}$ ορίζονται στο φάσμα του πίνακα $\boldsymbol{A}$ και

$$\boldsymbol{h(\lambda) = \alpha f(\lambda) + \beta g(\lambda)},$$

όπου $\boldsymbol{\alpha, \beta} \in \mathbb{C}$ και

$$\boldsymbol{k(\lambda) = f(\lambda)g(\lambda)},$$

μπορούμε να δείξουμε ότι $\boldsymbol{h(\lambda)}$ και $\boldsymbol{k(\lambda)}$ ορίζονται στο φάσμα του $\boldsymbol{A}$ και

$$(\boldsymbol{\alpha})\quad \boldsymbol{h(A) = \alpha f(A) + \beta g(A)},$$

$$(\boldsymbol{\beta})\quad \boldsymbol{k(A) = f(A) g(A)}.$$

Πράγματι, το κομμάτι $(\boldsymbol{\alpha})$ προκύπτει άμεσα από τον ορισμό $\boldsymbol{f(A) \triangleq p(A)}$, αφού έχουμε ότι $\boldsymbol{f(\lambda)}$ και $\boldsymbol{g(\lambda)}$ ορίζονται στο φάσμα του πίνακα $\boldsymbol{A}$ και $\boldsymbol{h(\lambda)}$ μπορεί να θεωρηθεί, ως ένα πολυώνυμο αυτών. Συνεπώς, μπορούμε να γράψουμε

$$\boldsymbol{h(A) = \alpha f(A) + \beta g(A)}.$$

Για το κομμάτι $(\boldsymbol{\beta})$, έστω $\boldsymbol{p}$ και $\boldsymbol{q}$ πολυώνυμα παρεμβολής των $\boldsymbol{f}$ και $\boldsymbol{g}$, αντίστοιχα, στο φάσμα του πίνακα $\boldsymbol{A}$, έτσι ώστε

$$\boldsymbol{p(A) = f(A)} \quad \text{και} \quad \boldsymbol{q(A) = g(A)}.$$

Διαφορίζοντας και χρησιμοποιώντας τον κανόνα γινομένου βρίσκουμε, ότι οι συναρτήσεις

$$\boldsymbol{h(\lambda)} \quad \text{και} \quad \boldsymbol{r(\lambda) = p(\lambda) q(\lambda)},$$

έχουν τις ίδιες τιμές στο φάσμα του πίνακα $\boldsymbol{A}$. Συνεπώς,

$$\boldsymbol{h(A) = r(A) = p(A) q(A) = f(A) g(A)}. \quad \square$$

**Παράδειγμα 2.3.3.**

Χρησιμοποιώντας τις ιδέες αυτής της παραγράφου μπορούμε να ορίσουμε μια θετικά ορισμένη τετραγωνική ρίζα, για έναν θετικά ορισμένο πίνακα και να παρατηρήσουμε, ότι υπάρχουν γενικά αρκετές τετραγωνικές ρίζες.

Για παράδειγμα ας υποθέσουμε, ότι θέλουμε να βρούμε τις ρίζες του πίνακα

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

δηλαδή να λύσουμε

$$\boldsymbol{X^2 = A}.$$

Παίρνοντας

$$\boldsymbol{f(\lambda) = \sqrt{\lambda}},$$

οι συνθήκες παρεμβολής της Πρότασης ***2.2.1*** είναι (με $\boldsymbol{s = 1}$, $\boldsymbol{m_1 = 1}$) απλώς

$$\boldsymbol{p(1) = \sqrt{1}}.$$

Το πολυώνυμο παρεμβολής είναι συνεπώς, ή $\boldsymbol{p(\lambda) = 1}$, ή $\boldsymbol{p(\lambda) = -1}$, αντιστοιχώντας στις δύο τετραγωνικές ρίζες του $\boldsymbol{1}$, δίνοντας $\boldsymbol{I}$ και $\boldsymbol{-I}$ σαν τετραγωνικές ρίζες του πίνακα $\boldsymbol{A}$. Αυτές οι δύο τετραγωνικές ρίζες είναι πολυώνυμα του πίνακα $\boldsymbol{A}$. Μπορούμε όμως να πάρουμε και άλλες δύο τετραγωνικές ρίζες, τις

$$\begin{bmatrix} \boldsymbol{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{1} \end{bmatrix}, \quad \begin{bmatrix} \boldsymbol{1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{-1} \end{bmatrix}.$$

Επιπλέον, αφού

$$\boldsymbol{A = ZIZ^{-1}},$$

είναι μια κανονική μορφή Jordan, για οποιοδήποτε αντιστρέψιμο πίνακα $\boldsymbol{Z}$, μπορούμε να πάρουμε ακόμη τους πίνακες

$$\boldsymbol{Z} \begin{bmatrix} \boldsymbol{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{1} \end{bmatrix} \boldsymbol{Z^{-1}}, \quad \boldsymbol{Z} \begin{bmatrix} \boldsymbol{1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{-1} \end{bmatrix} \boldsymbol{Z^{-1}}. \quad \square$$

# Κεφάλαιο 3

## 3.1 Ορισμός μιας Συνάρτησης Πίνακα μέσω της Κανονικής Μορφής Jordan

Σε αυτό το κεφάλαιο αποκτούμε μερικές σημαντικές ιδιότητες του πίνακα $\boldsymbol{f(A)}$, οι οποίες συγκεκριμένα, μας επιτρέπουν να υπολογίσουμε τον πίνακα $\boldsymbol{f(A)}$ εύκολα, υπό την προϋπόθεση, ότι η κανονική μορφή Jordan του πίνακα $\boldsymbol{A}$ είναι γνωστή. Αυτές οι ιδιότητες θα καθορίσουν σημαντικές σχέσεις ανάμεσα στις κανονικές μορφές Jordan του πίνακα $\boldsymbol{A}$ και του πίνακα $\boldsymbol{f(A)}$.

**Θεώρημα 3.1.1.** *Αν $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$ είναι ένας διαγώνιος τετραγωνικός πίνακας, ο οποίος αποτελείται από τετραγωνικούς υποπίνακες,*

$$\boldsymbol{A} = \mathbf{diag}[\boldsymbol{A_1}, \boldsymbol{A_2}, \ldots, \boldsymbol{A_t}],$$

*και η συνάρτηση $\boldsymbol{f(\lambda)}$ ορίζεται στο φάσμα του πίνακα $\boldsymbol{A}$, τότε*

$$\boldsymbol{f(A)} = \mathbf{diag}[\boldsymbol{f(A_1)}, \boldsymbol{f(A_2)}, \ldots, \boldsymbol{f(A_t)}]. \qquad \textbf{(3.1.1)}$$

**Απόδειξη.** Αρχικά, είναι γνωστό, πως για οποιοδήποτε πολυώνυμο $\boldsymbol{q(\lambda)}$,

$$\boldsymbol{q(A)} = \mathbf{diag}[\boldsymbol{q(A_1)}, \boldsymbol{q(A_2)}, \ldots, \boldsymbol{q(A_t)}].$$

Συνεπώς, αν $\boldsymbol{p(\lambda)}$ είναι το πολυώνυμο παρεμβολής για την συνάρτηση $\boldsymbol{f(\lambda)}$ στο φάσμα του πίνακα $\boldsymbol{A}$, έχουμε ότι

$$\boldsymbol{f(A)} = \boldsymbol{p(A)} = \mathbf{diag}[\boldsymbol{p(A_1)}, \boldsymbol{p(A_2)}, \ldots, \boldsymbol{p(A_t)}].$$

Επιπλέον, αφού το φάσμα του υποπίνακα $\boldsymbol{A_j}$ $\boldsymbol{(1 \le j \le t)}$ είναι προφανώς υποσύνολο του φάσματος του πίνακα $\boldsymbol{A}$, η συνάρτηση $\boldsymbol{f(\lambda)}$ ορίζεται στο φάσμα του υποπίνακα $\boldsymbol{A_j}$, για κάθε $\boldsymbol{j = 1, 2, \ldots, t}$. (Σημειώνουμε επίσης ότι ο δείκτης μιας ιδιοτιμής του υποπίνακα $\boldsymbol{A_j}$, δεν μπορεί να υπερβαίνει τον δείκτη της ίδιας ιδιοτιμής του πίνακα $\boldsymbol{A}$.) Επιπροσθέτως, αφού $\boldsymbol{f(\lambda)}$ και $\boldsymbol{p(\lambda)}$ παίρνουν τις ίδιες τίμες στο φάσμα του πίνακα $\boldsymbol{A}$, θα πρέπει να έχουν επίσης τις ίδιες τιμές στο φάσμα του υποπίνακα $\boldsymbol{A_j}$ $\boldsymbol{(j = 1, 2, \ldots, t)}$. Συνεπώς,

$$\boldsymbol{f(A_j)} = \boldsymbol{p(A_j)}$$

και λαμβάνουμε την εξίσωση $\textbf{(3.1.1)}$. ■

Το επόμενο αποτέλεσμα μας προσφέρει μία σχέση μεταξύ του $\boldsymbol{f(A)}$ και του $\boldsymbol{f(B)}$, όταν $\boldsymbol{A}$ και $\boldsymbol{B}$ είναι όμοιοι πίνακες.

**Θεώρημα 3.1.2.** *Αν $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{P} \in \mathbb{C}^{\boldsymbol{n\times n}}$, όπου $\boldsymbol{B} = \boldsymbol{PAP^{-1}}$ και $\boldsymbol{f(\lambda)}$ ορίζεται στο φάσμα του πίνακα $\boldsymbol{A}$, τότε*

$$\boldsymbol{f(B) = Pf(A)P^{-1}}. \tag{3.1.2}$$

**Απόδειξη.** Αφού $\boldsymbol{A}$ και $\boldsymbol{B}$ είναι όμοιοι πίνακες, τότε έχουν το ίδιο ελάχιστο πολυώνυμο, Πρόταση ***1.3.7***. Συνεπώς, αν $\boldsymbol{p(\lambda)}$ είναι το πολυώνυμο παρεμβολής, για το $\boldsymbol{f(\lambda)}$ στο φάσμα του πίνακα $\boldsymbol{A}$, τότε είναι επίσης το πολυώνυμο παρεμβολής για το $\boldsymbol{f(\lambda)}$ στο φάσμα του πίνακα $\boldsymbol{B}$. Συνεπώς, έχουμε ότι

$$\boldsymbol{f(A) = p(A)} \quad \text{και} \quad \boldsymbol{f(B) = p(B)}.$$
$$\boldsymbol{B = PAP^{-1}} \quad \Rightarrow \quad \boldsymbol{p(B) = Pp(A)P^{-1}}$$

από όπου προκύπτει η εξίσωση **(3.1.2)**. ■

Λαμβάνοντας υπόψη την κανονική μορφή Jordan, τα Θεωρήματα ***3.1.1*** και ***3.1.2*** υποδεικνύουν ένα σχετικό θεώρημα, για τις συναρτήσεις πινάκων.

**Θεώρημα 3.1.3.** *Έστω $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$ και έστω $\boldsymbol{J} = \mathbf{diag}[\boldsymbol{J_j}]_{\boldsymbol{j=1}}^{\boldsymbol{t}}$, είναι η κανονική μορφή Jordan του πίνακα $\boldsymbol{A}$, όπου $\boldsymbol{A = PJP^{-1}}$ και $\boldsymbol{J_j}$ είναι ο $\boldsymbol{j}$-οστός υποπίνακας Jordan του πίνακα $\boldsymbol{J}$.* Τότε

$$\boldsymbol{f(A) = P}\mathbf{diag}[\boldsymbol{f(J_1), f(J_2), \ldots, f(J_t)}]\boldsymbol{P^{-1}}. \tag{3.1.3}$$

Το τελευταίο βήμα για τον υπολογισμό του πίνακα $\boldsymbol{f(A)}$ με τη χρήση της κανονικής μορφής Jordan του πίνακα $\boldsymbol{A}$ αποτελείται, επομένως, από τον ακόλουθο τύπο.

**Θεώρημα 3.1.4.** *Έστω $\boldsymbol{J_0}$ είναι ένας υποπίνακας Jordan μεγέθους $\boldsymbol{l}$ σχετιζόμενος μϵ $\boldsymbol{\lambda_0}$*

$$\boldsymbol{J_0} = \begin{bmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{bmatrix}.$$

*Αν $\boldsymbol{f(\lambda)}$ ϵίναι μια $\boldsymbol{(l-1)}$-φορές διαφορίσιμη συνάρτηση σϵ μια πϵριοχή του $\boldsymbol{\lambda_0}$, τότϵ*

$$\boldsymbol{f(J_0)} = \begin{bmatrix} f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \cdots & \frac{1}{(l-1)!}f^{(l-1)}(\lambda_0) \\ 0 & f(\lambda_0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{1}{1!}f'(\lambda_0) \\ 0 & \cdots & 0 & f(\lambda_0) \end{bmatrix}. \qquad (3.1.4)$$

**Απόδειξη.** Το ελάχιστο πολυώνυμο του $\boldsymbol{J_0}$ είναι $\boldsymbol{(\lambda-\lambda_0)^i}$ και οι τιμές της συνάρτησης $\boldsymbol{f(\lambda)}$ στο φάσμα του πίνακα $\boldsymbol{J_0}$ είναι επομένως

$$\boldsymbol{f(\lambda_0), f'(\lambda_0), \ldots, f^{(l-1)}(\lambda_0)}.$$

Το πολυώνυμο παρεμβολής $\boldsymbol{p(\lambda)}$, που ορίζεται από τις τιμές του $\boldsymbol{f(\lambda)}$ στο φάσμα $\boldsymbol{\{\lambda_0\}}$ του πίνακα $\boldsymbol{J_0}$, βρίσκεται θέτοντας

$$\boldsymbol{s=1, \quad m_k=1, \quad \lambda_1=\lambda_0 \quad \text{και} \quad \psi_1(\lambda)\equiv 1},$$

στην εξίσωση $(\mathbf{2.2.2})$ και χρησιμοποιώντας την σχέση $(\mathbf{2.2.4})$. Παίρνουμε

$$\boldsymbol{p(\lambda) = \sum_{i=0}^{l-1} \frac{1}{i!} f^{(i)}(\lambda_0)(\lambda-\lambda_0)^i}.$$

Το γεγονός ότι το πολυώνυμο $\boldsymbol{p(\lambda)}$ επιλύει το πρόβλημα παρεμβολής

$$\boldsymbol{p^{(j)}(\lambda_0) = f^{(j)}(\lambda_0), \quad 1 \le j \le l-1},$$

μπορεί επίσης να ελεγχθεί εύκολα με έναν άμεσο υπολογισμό. Τότε έχουμε $\boldsymbol{f(J_0) = p(J_0)}$ και συνεπώς

$$\boldsymbol{f(J_0) = \sum_{i=0}^{l-1} \frac{1}{i!} f^{(i)}(\lambda_0)(J_0 - \lambda_0 I)^i}.$$

Υπολογίζοντας τις δυνάμεις του $\boldsymbol{(J_0 - \lambda_0 I)}$, παίρνουμε

$$\boldsymbol{(J_0 - \lambda_0 I)^i} = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & & & & & \ddots & \vdots \\ & & & & \ddots & & 0 \\ & & & \ddots & & & 1 \\ & & & & & & 0 \\ & & & & & & \vdots \\ & & & & & & 0 \end{bmatrix}$$

και έτσι προκύπτει η εξίσωση **(3.1.4)** ■

Συνεπώς, αν γνωρίζουμε μια κανονική μορφή Jordan του πίνακα $\boldsymbol{A}$, ο πίνακας $\boldsymbol{f(A)}$ μπορεί να βρεθεί εύκολα, αν συνδυάσουμε τα Θεωρήματα ***3.1.3*** και ***3.1.4***

**Θεώρημα 3.1.5.** *Χρησιμοποιώντας την εξίσωση* **(3.1.3)** *του Θεωρήματος* ***3.1.3***, *μπορούμε να γράψουμε,*

$$\boldsymbol{f(A) = P\mathbf{diag}[f(J_1), f(J_2), \ldots, f(J_t)]P^{-1}},$$

*όπου* $\boldsymbol{f(J_i)}$ $\boldsymbol{(i = 1, 2, \ldots, t)}$ *είναι άνω-τριγωνικοί πίνακες της μορφής, που δίνεται στην εξίσωση* **(3.1.4)**.

Παρατηρούμε, ότι αν ο υποπίνακας $\boldsymbol{J_i}$ σχετίζεται με την ιδιοτιμή $\boldsymbol{\lambda_i}$, τα διαγώνια στοιχεία του υποπίνακα $\boldsymbol{f(J_i)}$ είναι όλα ίσα με $\boldsymbol{(\lambda_i)}$ και αφού οι ιδιοτιμές ενός τριγωνικού πίνακα αποτελούν τα διαγώνια στοιχεία του, προκύπτει ότι οι ιδιοτιμές του πίνακα $\boldsymbol{f(A)}$, πιθανώς πολλαπλές, είναι $\boldsymbol{f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_s)}$.

**Θεώρημα 3.1.6.** *Αν* $\boldsymbol{\lambda_1, \lambda_2, \ldots, \lambda_n}$ *είναι οι ιδιοτιμές του πίνακα* $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n \times n}}$ *και* $\boldsymbol{f(\lambda)}$ *είναι ορισμένη στο φάσμα του πίνακα* $\boldsymbol{A}$, *τότε οι ιδιοτιμές του πίνακα* $\boldsymbol{f(A)}$ *είναι*

$$\boldsymbol{f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)}.$$

# Κεφάλαιο 4

## 4.1 Φασματική Ανάλυση του Πίνακα $f(A)$

Η ακόλουθη διάσπαση ή ***φασματική ανάλυση***, του πίνακα $f(A)$ είναι σημαντική και προφανώς σχετίζεται άμεσα με το φασματικό θεώρημα, για τους απλούς πίνακες, συγκεκριμένα Θεώρημα ***1.3.2*** . Μας προσφέρει μια χρήσιμη και σαφή αναπαράσταση του πίνακα $f(A)$ σε όρους θεμελιωδών ιδιοτήτων του πίνακα $A$.

**Θεώρημα 4.1.1** *Αν ο πίνακας $A \in \mathbb{C}^{n\times n}$, $f(\lambda)$ ορίζεται στο φάσμα $\{\lambda_1, \lambda_2, \ldots, \lambda_s\}$ του πίνακα $A$, $f_{k,j}$ είναι η τιμή της $j$-οστής παραγώγου του $f(\lambda)$ στην ιδιοτιμή $\lambda_k$ $(k = 1, 2, \ldots, s,\quad j = 0, 1, \ldots, m_k - 1)$ και $m_k$ είναι ο δείκτης της ιδιοτιμής $\lambda_k$, τότε υπάρχουν πίνακες $Z_{kj}$, ανεξάρτητοι του $f(\lambda)$, τέτοιοι ώστε*

$$f(A) = \sum_{k=1}^{s} \sum_{j=0}^{m_k-1} f_{k,\,j} Z_{kj}. \tag{4.1.1}$$

*Επιπλέον, οι πίνακες $Z_{kj}$ είναι γραμμικά ανεξάρτητα μέλη του $\mathbb{C}^{n\times n}$ και ισχύει, ότι διατηρούν την αντιμεταθετική ιδιότητα στον πολλαπλασιασμό με τον πίνακα $A$, καθώς και στον μεταξύ τους πολλαπλασιασμό.*

**Απόδειξη.** Ανακαλούμε ότι η συνάρτηση $f(\lambda)$ και το πολυώνυμο $p(\lambda)$, που ορίζεται στην εξίσωση **(2.2.8)**, παίρνουν τις ίδιες τιμές στο φάσμα του πίνακα $A$. Συνεπώς,

$$f(A) = p(A)$$

και η εξίσωση **(4.1.1)** προκύπτει τοποθετώντας

$$Z_{kj} = \varphi_{kj}(A).$$

Συνεπάγεται άμεσα ότι ο πίνακας $A$ και όλοι οι πίνακες $Z_{kj}$, είναι αμοιβαία αντιμεταθετικοί.

Παραμένει μόνο να αποδείξουμε ότι αυτοί οι πίνακες είναι γραμμικά ανεξάρτητοι. Υποθέτουμε, ότι

$$\sum_{k=1}^{s} \sum_{j=0}^{m_k-1} c_{kj} Z_{kj} = 0$$

και ορίζουμε

$$\boldsymbol{h(\lambda) = \sum_{k,\, j} c_{kj}\varphi_{kj}(\lambda)}.$$

Τότε $\boldsymbol{h(\lambda)}$ είναι ένα πολυώνυμο του χώρου $\boldsymbol{\mathcal{P}_{m-1}}$, δηλαδή ο βαθμός του $\boldsymbol{h(\lambda)}$ δεν ξεπερνά το $\boldsymbol{m-1}$. Τότε η εξίσωση

$$\boldsymbol{Z_{kj} = \varphi_{kj}(A)},$$

υποδεικνύει, ότι

$$\boldsymbol{h(A) = 0}.$$

Αλλά το $\boldsymbol{m}$ είναι ο βαθμός του ελαχίστου πολυωνύμου του πίνακα $\boldsymbol{A}$ και συνεπώς, το $\boldsymbol{h}$ μηδενίζει τον πίνακα $\boldsymbol{A}$, άρα συνεπάγεται από το Θεώρημα ***1.3.3***, ότι το $\boldsymbol{h}$ είναι το μηδενικό πολυώνυμο. Επομένως, αφού τα $\boldsymbol{\varphi_{kj}}$ είναι ανεξάρτητα, έχουμε ότι

$$\boldsymbol{c_{kj} = 0}, \quad \text{για κάθε} \quad \boldsymbol{k} \quad \text{και} \quad \boldsymbol{j}.$$

Έτσι, οι πίνακες $\boldsymbol{Z_{kj}}$ είναι γραμμικά ανεξάρτητοι. ■

Θα καλούμε τους πίνακες $\boldsymbol{Z_{kj}}$, ως ***πίνακες-συντελεστές*** του $\boldsymbol{A}$, σημειώνοντας, ότι επειδή είναι γραμμικά ανεξάρτητοι, κανένας από αυτούς δεν μπορεί να είναι ο μηδενικός-πίνακας.

Ας εξετάσουμε την υπόθεση των απλών πινάκων πιο προσεκτικά. Πλέον, έχουμε ότι

$$\boldsymbol{m_1 = m_2 = \cdots = m_s = 1},$$

στην εξίσωση $\boldsymbol{(4.1.1)}$ και έτσι για ένα απλό πίνακα $\boldsymbol{A}$,

$$\boldsymbol{f(A) = \sum_{k=1}^{s} f_k Z_{k0}}, \quad \boldsymbol{f_k = f_{k,0}}, \quad \boldsymbol{k = 1, 2, \ldots, s}.$$

Ανακαλούμε ότι,

$$\boldsymbol{Z_{k0} = \varphi_{k0}(A)}.$$

Παρατηρούμε ότι ο $\boldsymbol{Z_{k0}}$ είναι ταυτοδύναμος πίνακας και το εύρος του $\boldsymbol{Z_{k0}}$ είναι ο (δεξιός) ιδιόχωρος που σχετίζεται με την ιδιοτιμή $\boldsymbol{\lambda_k}$ (ο οποίος χώρος είναι, ο πυρήνας του $\boldsymbol{\lambda_k I - A}$). Συμπεραίνουμε λοιπόν, ότι για ένα απλό πίνακα $\boldsymbol{A}$

$$\boldsymbol{\sum_{k=1}^{s} Z_{k0} = I} \quad \text{και} \quad \boldsymbol{Z_{k0}Z_{j0} = \delta_{kj}Z_{k0}},$$

παρόλο που το πρώτο από τα αποτελέσματα φαίνεται να είναι αληθές, για όλους τους πίνακες $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n}\times\boldsymbol{n}}$. Θα δείξουμε στο Θεώρημα ***4.1.2***, ότι και το δεύτερο αποτέλεσμα είναι γενικώς αληθές.

Τέλος, σημειώνουμε, ότι

$$\boldsymbol{Z_{k0}} = \boldsymbol{\varphi_{k0}(A)} = \prod_{j=1,\ j\neq k}^{s} (\boldsymbol{A} - \boldsymbol{\lambda_j I}) \Big/ \prod_{j=1,\ j\neq k}^{s} (\boldsymbol{\lambda_k} - \boldsymbol{\lambda_j}).$$

Ας εξετάσουμε αναλυτικότερα τις ιδιότητες των πινάκων-συντελεστών ενός πίνακα $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n}\times\boldsymbol{n}}$. Σημειώνουμε, ότι ιδιαίτερα το τμήμα $(\boldsymbol{\gamma})$ του επόμενου θεωρήματος, μας δείχνει ότι οι πίνακες-συντελεστές $\boldsymbol{Z_{10}, \ldots, Z_{s0}}$ οποιουδήποτε πίνακα, είναι ταυτοδύναμοι και από το τμήμα $(\boldsymbol{\alpha})$ βλέπουμε ότι το άθροισμά τους είναι ο ταυτοτικός πίνακας. Ο τύπος του τμήματος $(\boldsymbol{\alpha})$ μερικές φορές καλείται ως ***ανάλυση του ταυτοτικού πίνακα.***

**Θεώρημα 4.1.2.** *Οι πίνακες-συντελεστές* $\boldsymbol{Z_{kj}}$ $(\boldsymbol{j = 0, 1, \ldots, m_k - 1},\ \ \boldsymbol{k = 1, 2, \ldots, s})$ *του πίνακα* $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n}\times\boldsymbol{n}}$, *ικανοποιούν τις συνθήκες*

$(\boldsymbol{\alpha})$ $\displaystyle\sum_{k=1}^{s} \boldsymbol{Z_{k0}} = \boldsymbol{I}.$

$(\boldsymbol{\beta})$ $\boldsymbol{Z_{kp} Z_l = 0}$, *αν* $\boldsymbol{k \neq l}$.

$(\boldsymbol{\gamma})$ $\boldsymbol{Z_{kj}^2 = Z_{kj}}$, *αν και μόνο αν* $\boldsymbol{j = 0}$.

$(\boldsymbol{\delta})$ $\boldsymbol{Z_{k0} Z_{kr} = Z_{kr}}$, $\boldsymbol{r = 0, 1, \ldots, m_k - 1}$.

**Απόδειξη.** Η σχέση $(\boldsymbol{\alpha})$ προκύπτει άμεσα από τον ορισμό

$$\boldsymbol{Z_{k0}} = \boldsymbol{\varphi_{k0}(A)}.$$

Όμως αυτή, αλλά και οι άλλες ιδιότητες προκύπτουν εύκολα από την κατασκευή των πινάκων-συντελεστών $\boldsymbol{Z_{kj}}$.

Πρώτα παρατηρούμε, ότι αν $\boldsymbol{A = PJP^{-1}}$, όπου $\boldsymbol{J = [J_1, J_2, \ldots, J_t]}$ είναι μια κανονική μορφή Jordan για τον πίνακα $\boldsymbol{A}$, τότε από το Θεώρημα ***3.1.5***

$$Z_{kj} = \varphi_{kj}(A) = P\mathrm{diag}[\varphi_{kj}(J_1), \ldots, \varphi_{kj}(J_t)]P^{-1}, \tag{4.1.2}$$

όπου $\varphi_{kj}(J_i)$ $(1 \leq i \leq t, \quad 1 \leq k \leq s, \quad 0 \leq j \leq m_k - 1)$ είναι ένας πίνακας της μορφής που παρουσιάζεται στη σχέση **(3.1.4)**. Η δομή του $\varphi_{kj}(J_i)$ μελετάται λαμβάνοντας υπόψη δύο ενδεχόμενα.

**Περίπτωση 1.** Ο υποπίνακας Jordan $J_i$ (μεγέθους $n_i$) σχετίζεται με την ιδιοτιμή $\lambda_k$. Από τη σχέση **(2.2.6)**, έχουμε ότι

$$\varphi_{kj}^{(r)}(\lambda_k) = \delta_{jr}\,, \quad (r,\, j = 0, 1, \ldots, m_k - 1)$$

και συνεπάγεται από την σχέση **(3.1.4)**, ότι

$$\varphi_{kj}(J_i) = \frac{1}{j!} N_{n_i}^{j}, \tag{4.1.3}$$

όπου $N_{n_i}$ είναι ο $n_i \times n_i$ υποπίνακας Jordan που σχετίζεται με την μηδενική ιδιοτιμή και θεωρείται, ότι για $j = 0$ η έκφραση στο δεξιό μέλος της εξίσωσης **(4.1.3)** είναι ο $n_i \times n_i$ ταυτοτικός πίνακας. Σημειώνουμε επίσης ότι $n_i \leq m_k$.

**Περίπτωση 2.** Ο υποπίνακας Jordan $J_i$ (τάξης $n_i$) σχετίζεται με μια ιδιοτιμή $\lambda_l \neq \lambda_k$. Τότε πάλι, $n_i \leq m_l$ (όπως στην σχέση **(2.2.7)**) και παρατηρούμε ότι το $\varphi_{kj}(\lambda)$, μαζί με τις πρώτες $m_l - 1$ παραγώγους του, είναι μηδέν στο $\lambda = \lambda_l$. Χρησιμοποιώντας ξανά την σχέση **(3.1.4)** βλέπουμε, ότι

$$\varphi_{kj}(J_i) = 0. \tag{4.1.4}$$

Για να απλοποιήσουμε τον συμβολισμό, θεωρούμε την περίπτωση $k = 1$. Αντικαθιστούμε τις εξισώσεις **(4.1.3)** και **(4.1.4)** στην εξίσωση **(4.1.2)** και διαπιστώνουμε, ότι

$$Z_{1j} = \frac{1}{j!} P\mathrm{diag}[N_{n_1}^{j}, N_{n_2}^{j}, \ldots, N_{n_q}^{j}, 0, \ldots, 0]P^{-1}, \tag{4.1.5}$$

για $j = 0, 1, \ldots, m_1 - 1$, υπό την προϋπόθεση, ότι $J_1, J_2, \ldots, J_q$ είναι οι μοναδικοί υποπίνακες Jordan του $J$ που σχετίζονται με την ιδιοτιμή $\lambda_1$. Πιο συγκεκριμένα, η εξίσωση **(4.1.5)** υποδεικνύει, ότι

$$\boldsymbol{Z_{10} = P\mathrm{diag}[I_r, 0]P^{-1}},\tag{4.1.6}$$

όπου $\boldsymbol{r = n_1 + n_2 + \cdots + n_q}$. ■

Τα αποτελέσματα των Θεωρημάτων ***4.1.1*** και ***4.2.2***, μπορούν να βοηθήσουν στον υπολογισμό του $\boldsymbol{f(A)}$ και ιδιαίτερα στην ταυτόχρονη εύρεση αρκετών συναρτήσεων του ίδιου πίνακα. Στην τελευταία περίπτωση κάθε συνάρτηση μπορεί να παρουσιαστεί όπως στην εξίσωση **(4.1.1)**, μόνο που οι συντελεστές $\boldsymbol{f_{k,\,j}}$ θα διαφέρουν από την μια συνάρτηση στην άλλη.

Όταν βρισκόμαστε αντιμέτωποι με το πρόβλημα της εύρεσης των πινάκων $\boldsymbol{Z_{kj}}$, η τεχνική που χρησιμοποιούμε είναι πρώτα να βρίσκουμε τα πολυώνυμα παρεμβολής $\boldsymbol{\varphi_{kj}(\lambda)}$ και έπειτα να τοποθετούμε

$$\boldsymbol{Z_{kj} = \varphi_{kj}(A)}.$$

Όμως, ο υπολογισμός των πολυωνύμων $\boldsymbol{\varphi_{kj}(\lambda)}$ είναι μια αρκετά δύσκολη διαδικασία. Το επόμενο παράδειγμα υποδεικνύει πως αυτό το πρόβλημα μπορεί να αποφευχθεί.

**Παράδειγμα 4.1.1.**

Θέλουμε να βρούμε τους πίνακες-συντελεστές του πίνακα

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

Θέλουμε επίσης να βρούμε τον $\boldsymbol{\ln A}$ χρησιμοποιώντας το Θεώρημα ***4.1.1***. Μπορούμε να δούμε εύκολα, ότι το ελάχιστο πολυώνυμο του πίνακα $\boldsymbol{A}$ είναι $\boldsymbol{(\lambda - 2)(\lambda - 4)^2}$ και έτσι, για οποιαδήποτε $\boldsymbol{f}$ που ορίζεται στο φάσμα του πίνακα $\boldsymbol{A}$, το Θεώρημα ***4.1.1*** δίνει

$$\boldsymbol{f(A) = f(2)Z_{10} + f(4)Z_{20} + f^{(1)}(4)Z_{21}}\ .\tag{4.1.7}$$

Αντικαθιστούμε το $\boldsymbol{f(\lambda)}$ με $\boldsymbol{3}$ συνετά διαλεγμένα γραμμικά ανεξάρτητα πολυώνυμα και επιλύουμε τις συνεπαγόμενες εξισώσεις ταυτόχρονα, για $\boldsymbol{Z_{10}, Z_{20}, Z_{21}}$. Συνεπώς, τοποθετώντας $\boldsymbol{f(\lambda) = 1}$, $\boldsymbol{f(\lambda) = \lambda - 4}$ και $\boldsymbol{f(\lambda) = (\lambda - 4)^2}$ με τη σειρά, παίρνουμε

$$\boldsymbol{Z_{10} + Z_{20} = I},$$
$$\boldsymbol{-2Z_{10} + Z_{21} = A - 4I} = \begin{bmatrix} 2 & 2 & 2 \\ -2 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix},$$

$$\boldsymbol{4Z_{10} = (A - 4I)^2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -4 \\ 0 & 0 & 4 \end{bmatrix}.$$

Αυτές οι εξισώσεις δίνουν

$$\boldsymbol{Z_{10}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{Z_{20}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{Z_{21}} = \begin{bmatrix} 2 & 2 & 2 \\ -2 & -2 & -2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Συγκριτικά, ο ευθύς υπολογισμός ενός πίνακα-συντελεστή, έστω $\boldsymbol{Z_{20}}$, χρησιμοποιώντας τον ορισμό του, $\boldsymbol{Z_{20} = \varphi_{20}(A)}$, είναι περισσότερο περίπλοκος. Ψάχνοντας ένα πολυώνυμο $\boldsymbol{\varphi_{20}(\lambda)}$ βαθμού που δεν ξεπερνά το $\mathbf{2}$ και τέτοιο ώστε

$$\boldsymbol{\varphi_{20}(2) = 0}, \quad \boldsymbol{\varphi_{20}(4) = 1}, \quad \boldsymbol{\varphi'_{20}(4) = 0}, \tag{4.1.8}$$

γράφουμε

$$\boldsymbol{\varphi_{20}(\lambda) = [\alpha_0 + \alpha_1(\lambda - 4)](\lambda - 2)}$$

και παρατηρούμε ότι οι συνθήκες της σχέσης $(\mathbf{4.1.8})$ δίνουν $\boldsymbol{\alpha_0} = \frac{\mathbf{1}}{\mathbf{2}}$, $\boldsymbol{\alpha_1} = -\frac{\mathbf{1}}{\mathbf{4}}$. Επομένως,

$$\boldsymbol{Z_{20} = -\frac{1}{4}A^2 + 2A - 3I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

όπως απαιτείται.

Σημειώνουμε, ότι ο τύπος $(\mathbf{4.1.7})$ για $\boldsymbol{f(\lambda) = \ln\lambda}$ δίνει μια άλλη μέθοδο, για να βρίσκουμε το $\boldsymbol{\ln A}$:

$$\boldsymbol{\ln A = (\ln 2)Z_{10} + (\ln 4)Z_{20} + \frac{1}{4}Z_{21}} = \begin{bmatrix} \ln 4 + \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \ln 4 - \frac{1}{2} & \ln 2 - \frac{1}{2} \\ 0 & 0 & \ln 2 \end{bmatrix}. \quad \square$$

Το επόμενο αποτέλεσμα μας επιτρέπει, να βρούμε όλους τους πίνακες-συντελεστές αρκετά εύκολα, από τη στιγμή που οι πίνακες-προβολές $\boldsymbol{Z_{k0}}$ , $\boldsymbol{k = 1, 2, \ldots, s}$, μας είναι γνωστοί.

**Θεώρημα 4.1.3.** *Για* $\boldsymbol{k = 1, 2, \ldots, s}$ *και* $\boldsymbol{j = 0, 1, \ldots, m_k - 1}$,

$$\boldsymbol{Z_{kj} = \frac{1}{j!}(A - \lambda_k I)^j Z_{k0}} \ . \tag{4.1.9}$$

**Απόδειξη.** Παρατηρούμε ότι το αποτέλεσμα είναι τετριμμένο για $\boldsymbol{j = 0}$. Χωρίς βλάβη της γενικότητας, θα υποθέσουμε πάλι ότι $\boldsymbol{k = 1}$ και θα χρησιμοποιήσουμε τους σχετικούς συμβολισμούς που παρουσιάστηκαν προηγουμένως.

Αφού

$$\boldsymbol{A - \lambda_1 I = P(J - \lambda_1 I)P^{-1}},$$

συνεπάγεται ότι

$$\boldsymbol{J - \lambda_1 I = P^{-1}(A - \lambda_1 I)P}$$

$$\boldsymbol{= \mathrm{diag}[N_{n_1}, \ldots, N_{n_q}, J_{q+1} - \lambda_1 I, \ldots, J_t - \lambda_1 I]}, \tag{4.1.10}$$

όπου $\boldsymbol{J_1, J_2, \ldots, J_q}$ είναι οι μοναδικοί υποπίνακες Jordan του $\boldsymbol{J}$ που σχετίζονται με την ιδιοτιμή $\boldsymbol{\lambda_1}$. Πολλαπλασιάζοντας την εξίσωση **(4.1.10)** στο δεξιό της μέλος με $\boldsymbol{P^{-1} Z_{1j} P}$ και χρησιμοποιώντας την εξίσωση **(4.1.5)**, παίρνουμε

$$\boldsymbol{j! P^{-1}(A - \lambda_1 I) Z_{1j} P = \mathrm{diag}[N_{n_1}^{j+1}, \ldots, N_{n_q}^{j+1}, 0, \ldots, 0]}.$$

Χρησιμοποιώντας την εξίσωση **(4.1.5)** ξανά, έχουμε συνεπώς

$$\boldsymbol{j! P^{-1}(A - \lambda_1 I) Z_{1j} P = (j+1)! P^{-1} Z_{1,\, j+1} P}$$

και

$$\boldsymbol{Z_{1,\, j+1} = \frac{1}{j+1}(A - \lambda_1 I) Z_{1j}},$$

για $\boldsymbol{j = 0, 1, \ldots, m_k - 1}$. Εφαρμόζοντας αυτό το αποτέλεσμα επανειλημμένα, παίρνουμε την εξίσωση **(4.1.9)**, για $\boldsymbol{k = 1}$. ∎

Το Θεώρημα ***4.1.3*** θα συμπληρωθεί τώρα με μερική περιγραφή των πινάκων $\boldsymbol{Z_{k0}}$, $\boldsymbol{k = 1, 2, \ldots, s}$.

**Θεώρημα 4.1.4.** *Με τον προηγούμενο συμβολισμό,*

$$\mathbf{Im}\ \boldsymbol{Z_{k0}} = \mathbf{Ker}(\boldsymbol{A} - \boldsymbol{\lambda_k I})^{\boldsymbol{m_k}}.$$

**Απόδειξη.** Θα ασχοληθούμε ξανά μόνο με την περίπτωση $\boldsymbol{k = 1}$. Αν $\boldsymbol{J}$ είναι η κανονική μορφή Jordan του πίνακα $\boldsymbol{A}$ που χρησιμοποιείται στην παραπάνω συζήτηση, τότε

$$(\boldsymbol{J} - \boldsymbol{\lambda_1 I})^{\boldsymbol{m_1}} = \mathbf{diag}[\mathbf{0}, \ldots, \mathbf{0}, (\boldsymbol{J_{q+1}} - \boldsymbol{\lambda_1 I})^{\boldsymbol{m_1}}, \ldots, (\boldsymbol{J_t} - \boldsymbol{\lambda_1 I})^{\boldsymbol{m_1}}],$$

αφού $\boldsymbol{n_1, n_2, \ldots, n_q \leq m_1}$. Οι υποπίνακες $(\boldsymbol{J_r} - \boldsymbol{\lambda_1 I})^{\boldsymbol{m_1}}$ είναι αντιστρέψιμοι για $\boldsymbol{r = q + 1, \ldots, t}$ και συνεπώς ο πυρήνας του $(\boldsymbol{J} - \boldsymbol{\lambda_1 I})^{\boldsymbol{m_1}}$ αποτελείται από τα μοναδιαία διανύσματα $\boldsymbol{e_1, e_2, \ldots, e_r}$, όπου

$$\boldsymbol{r = n_1 + n_2 + \cdots + n_q}.$$

Ας συμβολίσουμε αυτό τον χώρο με $\boldsymbol{\mathcal{F}}$.

Έχουμε επίσης

$$(\boldsymbol{A} - \boldsymbol{\lambda_1 I})^{\boldsymbol{m_1}} = \boldsymbol{P}(\boldsymbol{J} - \boldsymbol{\lambda_1 I})^{\boldsymbol{m_1}} \boldsymbol{P}^{-\mathbf{1}}$$

και συνεπάγεται άμεσα, ότι

$$\boldsymbol{x} \in \mathbf{Ker}(\boldsymbol{A} - \boldsymbol{\lambda_1 I})^{\boldsymbol{m_1}},$$

αν και μόνο αν

$$\boldsymbol{P}^{-\mathbf{1}}\boldsymbol{x} \in \boldsymbol{\mathcal{F}}.$$

Χρησιμοποιώντας την εξίσωση $\mathbf{(4.1.6)}$, βλέπουμε ότι

$$\boldsymbol{x} \in \mathbf{Im}\, \boldsymbol{Z_{10}},$$

αν και μόνο αν

$$\boldsymbol{P}^{-\mathbf{1}}\boldsymbol{x} \in \mathbf{Im}[\boldsymbol{I_r}\ \ \mathbf{0}]\boldsymbol{P}^{-\mathbf{1}} = \boldsymbol{\mathcal{F}}$$

και συνεπώς

$$\mathbf{Ker}\ (\boldsymbol{A} - \boldsymbol{\lambda_1 I})^{\boldsymbol{m_1}} = \mathbf{Im}\, \boldsymbol{Z_{10}}\ \ . \quad \blacksquare$$

Προκύπτει ότι το αποτέλεσμα του προηγουμένου θεωρήματος δεν μπορεί να διευρυνθεί, για να δώσει μια αντίστοιχα απλή περιγραφή του $\mathbf{Im}\, \boldsymbol{Z_{kj}}$, για $\boldsymbol{j \geq 1}$.

# Κεφάλαιο 5

# Ιδιότητες των Συναρτήσεων Πινάκων

Σε αυτό το κεφάλαιο αποκτούμε μερικές επιπρόσθετες ιδιότητες των συναρτήσεων πινάκων, οι οποίες μας επιτρέπουν να διευρύνουμε μερικές γνωστές μας ταυτότητες, για βαθμωτές συναρτήσεις, στις αντίστοιχες συναρτήσεις που έχουν για τιμές τους πίνακες. Για παράδειγμα, ελέγχουμε αν η ταυτότητα

$$\sin^2 \boldsymbol{\alpha} + \cos^2 \boldsymbol{\alpha} = \mathbf{1},$$

υποδεικνύει την ύπαρξη της ταυτότητας

$$\sin^2 \boldsymbol{A} + \cos^2 \boldsymbol{A} = \boldsymbol{I}$$

και αν η ταυτότητα

$$\boldsymbol{e^{\alpha} e^{-\alpha}} = \mathbf{1},$$

υποδεικνύει την ύπαρξη της ταυτότητας

$$\boldsymbol{e^{A} e^{-A}} = \boldsymbol{I}.$$

Επιπλέον, θα μελετήσουμε αναλυτικά τους πίνακες-συντελεστές $\boldsymbol{Z_{kj}}$ και τις ιδιότητές τους, κάνοντας μια πρώτη αναφορά και στην επιλύουσα

$$(\boldsymbol{zI} - \boldsymbol{A})^{-1},$$

του πίνακα $\boldsymbol{A}$, αλλά και στους αναλλοίωτους υποχώρους που προκύπτουν από τους πίνακες-προβολές $\boldsymbol{Z_{k0}}$.

## 5.1 Μερικές Ιδιότητες Συναρτήσεων για Πίνακες

**Θεώρημα 5.1.1.** *Έστω* $\boldsymbol{P(u_1, u_2, \ldots, u_t)}$ *είναι ένα βαθμωτό πολυώνυμο των* $\boldsymbol{u_1, u_2,} \boldsymbol{\ldots, u_t}$ *και έστω* $\boldsymbol{f_1(\lambda), f_2(\lambda), \ldots, f_t(\lambda)}$ *είναι συναρτήσεις ορισμένες στο φάσμα του πίνακα* $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$. *Αν η συνάρτηση*

$$\boldsymbol{f(\lambda) = P(f_1(\lambda), f_2(\lambda), \ldots, f_t(\lambda))},$$

*παίρνει μηδενικές τιμές στο φάσμα του πίνακα* $\boldsymbol{A}$, *τότε*

$$\boldsymbol{f(A) = P(f_1(A), f_2(A), \ldots, f_t(A)) = 0}.$$

**Απόδειξη.** Έστω $\boldsymbol{p_1(\lambda), p_2(\lambda), \ldots, p_t(\lambda)}$ είναι πολυώνυμα παρεμβολής στο φάσμα του πίνακα $\boldsymbol{A}$, για $\boldsymbol{f_1(\lambda), f_2(\lambda), \ldots, f_t(\lambda)}$, αντίστοιχα. Τότε από τον ορισμό των $\boldsymbol{p_i}$, έχουμε

$$\boldsymbol{f_i(A) = p_i(A)},$$

για $\boldsymbol{i = 1, 2, \ldots, t}$. Συμβολίζουμε το πολυώνυμο $\boldsymbol{P(p_1(\lambda), p_2(\lambda), \ldots, p_t(\lambda))}$ με $\boldsymbol{p(\lambda)}$ και παρατηρούμε, ότι αφού $\boldsymbol{p_i(\lambda)}$ και $\boldsymbol{f_i(\lambda)}$ έχουν τις ίδιες τιμές στο φάσμα του πίνακα $\boldsymbol{A}$ $\boldsymbol{(1 \leq i \leq t)}$, το ίδιο θα ισχύει και για τα $\boldsymbol{p(\lambda)}$ και $\boldsymbol{f(\lambda)}$. Συνεπώς, η υπόθεσή μας, ότι $\boldsymbol{f(\lambda)}$ είναι μηδέν στο φάσμα του πίνακα $\boldsymbol{A}$ υποδεικνύει το ίδιο, για το $\boldsymbol{P(\lambda)}$ και επομένως

$$\boldsymbol{f(A) = p(A) = 0}. \quad \blacksquare$$

Για τις συγκεκριμένες περιπτώσεις που αναφέρθηκαν από πάνω, επιλέγουμε $\boldsymbol{f_1(\lambda) = \sin\lambda,}$ $\boldsymbol{f_2(\lambda) = \cos\lambda}$ και

$$\boldsymbol{P(u_1, u_2) = u_1^2 + u_2^2 - 1},$$

προκειμένου να αποδείξουμε ότι, για κάθε πίνακα $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$, είναι αληθές ότι

$$\boldsymbol{\sin^2 A + \cos^2 A = I}.$$

Έπειτα, παίρνοντας $\boldsymbol{f_1(\lambda) = e^{\lambda}}$, $\boldsymbol{f_2(\lambda) = e^{-\lambda}}$ και

$$\boldsymbol{P(u_1, u_2) = u_1 u_2 - 1},$$

προκύπτει

$$\boldsymbol{e^A e^{-A} = I}.$$

Συνεπώς,

$$\boldsymbol{e^{-A} = (e^{A})^{-1}}.$$

Μπορούμε επίσης να ασχοληθούμε με το ενδεχόμενο, να πάρουμε τις $\boldsymbol{p}$-οστές ρίζες ενός πίνακα με τη βοήθεια του Θεωρήματος ***5.1.1***. Υποθέτουμε, ότι $\boldsymbol{p}$ είναι ένας θετικός ακέραιος και $\boldsymbol{f_1(\lambda) = \lambda^{1/p}}$, $\boldsymbol{f_2(\lambda) = \lambda}$ και

$$\boldsymbol{P(u_1, u_2) = u_1^p - u_2}.$$

Συνήθως, θα διαλέγαμε ένα μοναδικά τιμώμενο τμήμα της συνάρτησης $\boldsymbol{\lambda^{1/p}}$ για $\boldsymbol{f_1(\lambda)}$, αλλά αυτό δεν είναι απαραίτητο. Αν όμως ο πίνακας $\boldsymbol{A}$ έχει μια μηδενική ιδιοτιμή με δείκτη μεγαλύτερο του $\mathbf{1}$, τότε αφού οι παράγωγοι του $\boldsymbol{f_1(\lambda)}$ χρειάζονται, $\boldsymbol{f_1(\lambda)}$ δεν θα μπορούσε να οριστεί στο φάσμα του πίνακα $\boldsymbol{A}$. Συνεπώς, αν $\boldsymbol{A}$ αντιστρέψιμος ή μη αντιστρέψιμος με μια μηδενική ιδιοτιμή με δείκτη $\mathbf{1}$, ορίζουμε $\boldsymbol{f_1(\lambda) = \lambda^{1/p}}$ και

$$\boldsymbol{A^{1/p} = f_1(A) = \sum_{k=1}^{s} \sum_{j=0}^{m_k-1} f_1^{(j)}(\lambda_k) Z_{kj}}.$$

Έπειτα, αν εφαρμόσουμε το Θεώρημα ***5.1.1*** στις συναρτήσεις $\boldsymbol{f_1(\lambda), f_2(\lambda)}$ και $\boldsymbol{P(u_1, u_2)}$ είναι το πολυώνυμο που ορίσαμε πιο πάνω,

$$\boldsymbol{(A^{1/p})^p = A}.$$

Να σημειωθεί, ότι ο ορισμός που δίνουμε στο $\boldsymbol{A^{1/2}}$, για παράδειγμα, δεν περιλαμβάνει όλους τους πιθανούς πίνακες $\boldsymbol{A}$ για τους οποίους $\boldsymbol{B^2 = A}$. Για να αποτυπώσουμε καλύτερα αυτό το γεγονός, κάθε πραγματικός ορθογώνιος πίνακας $\boldsymbol{B}$ έχει την ιδιότητα $\boldsymbol{B^2 = I}$, αλλά ο πίνακας $\boldsymbol{B}$, δεν θα παραχθεί αναγκαστικά με το να εφαρμόσουμε τους ορισμούς μας για να βρούμε τον πίνακα $\boldsymbol{I^{1/2}}$. Ένα συγκεκριμένο παράδειγμα είναι το σύνολο πινάκων που ορίζεται από

$$\boldsymbol{B(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}}.$$

**Θεώρημα 5.1.2.** *Έστω $h(\lambda) = g(f(\lambda))$, όπου $f(\lambda)$ ορίζεται στο φάσμα του πίνακα $A \in \mathbb{C}^{n\times n}$ και $g(\mu)$ ορίζεται στο φάσμα του πίνακα $B = f(A)$. Τότε $h(\lambda)$ ορίζεται στο φάσμα του πίνακα $A$ και*

$$h(A) = g(f(A)).$$

**Απόδειξη.** Έστω $\lambda_1, \lambda_2, \ldots, \lambda_s$ είναι οι διακεκριμένες ιδιοτιμές του πίνακα $A$ με δείκτες $m_1, m_2, \ldots, m_s$, αντίστοιχα. Συμβολίζουμε $f(\lambda) = \mu_k$, για $k = 1, 2, \ldots, s$. Αφού

$$\begin{aligned} h(\lambda_k) &= g(\mu_k), \\ h'(\lambda_k) &= g'(\mu_k) f'(\lambda_k), \\ &\vdots \\ h^{(m_k-1)}(\lambda_k) &= g^{(m_k-1)}(\mu_k)[f'(\lambda_k)]^{m_k-1} + \cdots + g'(\mu_k) f^{(m_k-1)}(\lambda_k), \end{aligned} \tag{5.1.1}$$

συνεπάγεται, ότι $h(\lambda)$ ορίζεται στο φάσμα του πίνακα $A$ και εναποθέτει τιμές που δίνονται από τους τύπους της σχέσης (**5.1.1**). Έστω ότι το $r(\mu)$ συμβολίζει οποιοδήποτε πολυώνυμο που έχει τις ίδιες τιμές με το $g(\mu)$ στο φάσμα του πίνακα $B$ και θεωρούμε τη συνάρτηση $r(f(\lambda))$. Αν οι τιμές των $h(\lambda)$ και $r(f(\lambda))$ συμπίπτουν στο φάσμα του πίνακα $A$, τότε $h(\lambda) - r(f(\lambda))$ εναποθέτει μηδενικές τιμές στο φάσμα του πίνακα $A$ και από το Θεώρημα ***5.1.1***,

$$h(A) = r(f(A)) = r(B) = g(B) = g(f(A))$$

και το αποτέλεσμα ισχύει.

Παρατηρούμε τώρα έχοντας υπόψη την σχέση (**5.1.1**) ότι οι τιμές των $h(\lambda)$ και $r(f(\lambda))$ συμπίπτουν στο φάσμα του πίνακα $A$, αν

$$r^i(\mu_k) = g^i(\mu_k), \tag{5.1.2}$$

για $i = 0, 1, \ldots, m_k - 1, \quad k = 1, 2, \ldots, s.$

Τώρα θα δείξουμε, ότι

$$\{g(\mu_k), g'(\mu_k), \ldots, g^{(m_k-1)}(\mu_k)\},$$

περιέχει όλες τις τιμές του $g(\mu)$ στο φάσμα του πίνακα $B$. Αυτό το συμπέρασμα είναι ισοδύναμο με το να πούμε, ότι ο δείκτης $n_k$ της ιδιοτιμής $\mu_k$ του πίνακα $B$ δεν ξεπερνά το $m_k$ ή ότι το πολυώνυμο

$$p(\mu) = (\mu - \mu_1)^{m_1}(\mu - \mu_2)^{m_2} \cdots (\mu - \mu_s)^{m_s},$$

μηδενίζει τον πίνακα $B$ $(1 \leq k \leq s)$. Για να το δούμε αυτό, παρατηρούμε ότι

$$q(\lambda) = p(f(\lambda)) = (f(\lambda) - f(\lambda_1))^{m_1} \cdots (f(\lambda) - f(\lambda_s))^{m_s},$$

έχει ρίζες $\lambda_k$ με πολλαπλότητα τουλάχιστον $m_k$ $(k = 1, 2, \ldots, s)$ και επομένως $q(\lambda)$ εναποθέτει μηδενικές τιμές στο φάσμα του πίνακα $A$. Συνεπώς, από το Θεώρημα ***5.1.1*** με $t = 1$,

$$p(B) = p(f(A)) = 0$$

και $p(\lambda)$ είναι ένα πολυώνυμο μηδενισμού για τον πίνακα $B$. Τώρα ορίζουμε το πολυώνυμο $r(\mu)$, έτσι ώστε να ικανοποιεί τις συνθήκες της σχέσης $(5.1.2)$ και παίρνουμε ένα πολυώνυμο με τιμές ίδιες με αυτές του $g(\mu)$ στο φάσμα του πίνακα $B$ και η απόδειξη ολοκληρώθηκε. Σημειώνουμε ότι $r(\mu)$ δεν είναι αναγκαστικά το πολυώνυμο παρεμβολής, για το $g(\mu)$ στο φάσμα του πίνακα $B$ του ελαχίστου δυνατού βαθμού. $\blacksquare$

## 5.2 Πίνακες-Συντελεστές και Αναλλοίωτοι Υπόχωροι

Γνωρίζουμε από το Θεώρημα ***1.1.1***, ότι δοθέντος ενός αυθαίρετου πίνακα $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n}\times\boldsymbol{n}}$, ο χώρος $\mathbb{C}^{\boldsymbol{n}}$ είναι ένα άμεσο άθροισμα των γενικευμένων ιδιόχωρων του πίνακα $\boldsymbol{A}$. Τώρα, το Θεώρημα ***4.1.4*** αναφέρει, ότι ο πίνακας-συντελεστής $\boldsymbol{Z_{k0}}$ είναι μια προβολή πάνω στο γενικευμένο ιδιόχωρο που σχετίζεται με την ιδιοτιμή $\boldsymbol{\lambda_k}$, $\boldsymbol{k=1,2,\dots,s}$. Επιπρόσθετα, το αποτέλεσμα του Θεωρήματος ***4.1.2*** $\boldsymbol{(\alpha)}$, η ανάλυση του ταυτοτικού πίνακα, δείχνει ότι η συμπληρωματική προβολή $\boldsymbol{I-Z_{k0}}$, μπορεί να εκφραστεί στη μορφή

$$\boldsymbol{I-Z_{k0}=\sum_{j=0,\, j\neq k}^{s} Z_{j0}}$$

και αφού

$$\mathbf{Im}\ \boldsymbol{Z_{i0}} \cap \mathbf{Im}\ \boldsymbol{Z_{j0}} = \{\mathbf{0}\}, \quad \text{όταν} \quad \boldsymbol{i \neq j},$$

$$\mathbf{Ker}\ \boldsymbol{Z_{k0}} = \mathbf{Im}(\boldsymbol{I-Z_{k0}}) = \sum_{\boldsymbol{j=0,\, j\neq k}}^{\boldsymbol{s}} \mathbf{Im}\, \boldsymbol{Z_{j0}}. \tag{5.2.1}$$

Συνεπώς, $\mathbf{Im}\,\boldsymbol{Z_{k0}}$, $\mathbf{Ker}\boldsymbol{Z_{k0}}$ είναι $\boldsymbol{A}$-αναλλοίωτοι και το επόμενο αποτέλεσμα ισχύει.

**Θεώρημα 5.2.1.** *Ο πίνακας-συντελεστής $\boldsymbol{Z_{k0}}$ που σχετίζεται με μια ιδιοτιμή $\boldsymbol{\lambda_k}$ του πίνακα $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n}\times\boldsymbol{n}}$, είναι μια προβολή πάνω στο γενικευμένο ιδιόχωρο του $\boldsymbol{\lambda_k}$ και κατά μήκος του αθροίσματος των γενικευμένων ιδιόχωρων που σχετίζονται με όλες τις ιδιοτιμές του πίνακα $\boldsymbol{A}$, διάφορες από $\boldsymbol{\lambda_k}$.*

Αν διαλέξουμε βάσεις στους χώρους $\mathbf{Im}\ \boldsymbol{Z_{k0}}$ και $\mathbf{Ker}\ \boldsymbol{Z_{k0}}$, τότε αυτοί οι χώροι θα προσδιορίσουν μια βάση για τον χώρο $\mathbb{C}^{\boldsymbol{n}}$, στην οποία ο πίνακας $\boldsymbol{A}$ έχει μια αναπαράσταση $\mathbf{diag}[\boldsymbol{A_1, A_2}]$, όπου τα μεγέθη των πινάκων $\boldsymbol{A_1}$ και $\boldsymbol{A_2}$ είναι απλώς οι διαστάσεις του $\mathbf{Im}\ \boldsymbol{Z_{k0}}$ και $\mathbf{Ker}\ \boldsymbol{Z_{k0}}$ , αντίστοιχα. Συγκεκριμένα, αν αυτές οι βάσεις αποτελούνται από αλυσίδες Jordan για τον πίνακα $\boldsymbol{A}$, τότε οι πίνακες $\boldsymbol{A_1}$ και $\boldsymbol{A_2}$ έχουν κανονική μορφή Jordan.

Για οποιονδήποτε πίνακα $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n}\times\boldsymbol{n}}$ και οποιοδήποτε $\boldsymbol{z} \notin \boldsymbol{\sigma(A)}$, η συνάρτηση

$$\boldsymbol{f(\lambda) = (z-\lambda)^{-1}},$$

ορίζεται στο $\boldsymbol{\sigma}(\mathrm{A})$, συνεπώς ***η επιλύουσα του*** $\boldsymbol{A}$, $\boldsymbol{(zI-A)^{-1}}$, υπάρχει. Διατηρώντας τον πίνακα $\boldsymbol{A}$ σταθερό, η επιλύουσα μπορεί να θεωρηθεί, ως μια συνάρτηση του $\boldsymbol{A}$ σε οποιαδήποτε περιοχή του $\mathbb{C}$ που δεν περιέχει σημεία του $\boldsymbol{\sigma(A)}$ και αυτός ο τρόπος θεώρησης παίζει ένα σημαντικό ρόλο στην ανάλυση πινάκων.

Μια αναπαράσταση της επιλύουσας είναι

$$\boldsymbol{(zI - A)^{-1} = \sum_{k=1}^{s} \sum_{j=0}^{m_k-1} \frac{j!}{(z-\lambda_k)^{j+1}} Z_{kj}} \ .$$

Χρησιμοποιώντας τις ιδιότητες καθετότητας των πινάκων $\boldsymbol{Z_{kj}}$, προκύπτει ότι

$$\boldsymbol{(zI - A)^{-1}(I - Z_{10}) = \sum_{k=2}^{s} \sum_{j=0}^{m_k-1} \frac{j!}{(z-\lambda_k)^{j+1}} Z_{kj}}$$

και χρησιμοποιώντας την εξίσωση $\mathbf{(5.2.1)}$, η επιλύουσα έχει ως εύρος $\mathbf{Im}\,\boldsymbol{(I - Z_{10})}$.

Καθώς το $\boldsymbol{z \to \lambda_1}$, τα στοιχεία του $\boldsymbol{(zI - A)^{-1}}$ δεν θα είναι φραγμένα. Αλλά, αφού το όριο στο δεξιό μέλος της εξίσωσης υπάρχει, παρατηρούμε ότι ο πολλαπλασιασμός της επιλύουσας με $\boldsymbol{I - Z_{10}}$ έχει ως αποτέλεσμα την εξουδετέρωση της απροσδιοριστίας του $\boldsymbol{(zI - A)^{-1}}$ στο $\boldsymbol{z = \lambda_1}$. Με άλλα λόγια, ο περιορισμός της επιλύουσας στον υπόχωρο

$$\mathbf{Im}\,\boldsymbol{(I - Z_{10}) = \sum_{j=2}^{s}}\mathbf{Im}\,\boldsymbol{Z_{j0}},$$

δεν παρουσιάζει απροσδιοριστία στο $\boldsymbol{z = \lambda_1}$.

Ορίζουμε τον πίνακα

$$\boldsymbol{E_1 = \lim_{z\to\lambda_1}(zI - A)^{-1}(I - Z_{10}) = \sum_{k=2}^{s} \sum_{j=0}^{m_k-1} \frac{j!}{(\lambda_1-\lambda_k)^{j+1}} Z_{kj}}. \qquad \mathbf{(5.2.2)}$$

**Θεώρημα 5.2.2.** *Χρησιμοποιώντας τους συμβολισμούς των προηγουμένων παραγράφων,*

$$\boldsymbol{I - Z_{10} = E_1(\lambda_1 I - A)}.$$

**Απόδειξη.** Ο απλός τύπος

$$\boldsymbol{\lambda_1 I - A = (\lambda_1 - z)I + (zI - A)},$$

υποδεικνύει

$$(\boldsymbol{zI} - \boldsymbol{A})^{-1}(\boldsymbol{\lambda_1 I} - \boldsymbol{A}) = (\boldsymbol{\lambda_1} - \boldsymbol{z})(\boldsymbol{zI} - \boldsymbol{A})^{-1} + \boldsymbol{I}.$$

Πολλαπλασιάζοντας στο αριστερό μέλος με $\boldsymbol{I} - \boldsymbol{Z_{10}}$ και παίρνοντας το όριο καθώς $\boldsymbol{z} \to \boldsymbol{\lambda_1}$ και τέλος χρησιμοποιώντας τον ορισμό του $\boldsymbol{E_1}$ που δίνεται στην σχέση $\mathbf{(5.2.2)}$, παίρνουμε το ζητούμενο αποτέλεσμα. ■

Αυτή η θεώρηση που εκφράζει το $\boldsymbol{I} - \boldsymbol{Z_{10}} = \sum_{\boldsymbol{k=2}}^{\boldsymbol{s}} \boldsymbol{Z_{k0}}$ σε όρους του $\boldsymbol{E_1}$ χρησιμοποιείται στην Θεωρία Διαταραχής.

# Κεφάλαιο 6

## 6.1 Ακολουθίες, Σειρές και ένας Αυστηρότερος Ορισμός Συνάρτησης Πίνακα

Σε αυτή την παράγραφο θα δούμε πως άλλες ιδέες της ανάλυσης μπορούν να εφαρμοστούν στους πίνακες και ιδιαίτερα στις συναρτήσεις πινάκων όπως τις έχουμε ορίσει. Θεωρούμε πρώτα ένα σύνολο από πίνακες που ορίζονται στους μη αρνητικούς ακέραιους αριθμούς, δηλαδή ακολουθίες πινάκων.

Έστω $\boldsymbol{A_1, A_2, \ldots}$ είναι μια ακολουθία από πίνακες που ανήκουν στο $\mathbb{C}^{\boldsymbol{m\times n}}$ και έστω $\boldsymbol{\alpha_{ij}^{(p)}}$ είναι το $\boldsymbol{j}$-οστό στοιχείο του πίνακα $\boldsymbol{A_p}$ , $\boldsymbol{p = 1, 2, \ldots}$

Η ακολουθία $\boldsymbol{A_1, A_2, \ldots}$ λέγεται ότι συγκλίνει στον πίνακα $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{m\times n}}$, αν υπάρχουν αριθμοί $\boldsymbol{\alpha_{ij}}$, (τα στοιχεία του πίνακα $\boldsymbol{A}$) τέτοιοι ώστε

$$\boldsymbol{\alpha_{ij}^{(p)} \to \alpha_{ij}},$$

καθώς $\boldsymbol{p \to \infty}$, για κάθε ζεύγος δεικτών $\boldsymbol{i}$, $\boldsymbol{j}$.

Μια ακολουθία που δεν συγκλίνει λέγεται ότι αποκλίνει. Συνεπώς, η σύγκλιση των ακολουθιών των πινάκων ορίζεται ως σημειακή σύγκλιση. Ο ορισμός περιλαμβάνει επίσης την σύγκλιση των διανυσματικών στηλών και σειρών, ως ιδιαίτερες περιπτώσεις.

**Πρόταση 6.1.1.** *Έστω*

$$\boldsymbol{\{A_p\}_{p=1}^{\infty}}, \quad \boldsymbol{\{B_p\}_{p=1}^{\infty}},$$

*είναι συγκλίνουσες ακολουθίες πινάκων του* $\mathbb{C}^{\boldsymbol{m\times n}}$ *και υποθέτουμε, ότι*

$$\boldsymbol{A_p \to A}, \quad \boldsymbol{B_p \to B},$$

*καθώς* $\boldsymbol{p \to \infty}$. *Τότε αποδεικνύεται, ότι*

$(\boldsymbol{\alpha})$ $\qquad \boldsymbol{A_p + B_p \to A + B}$.

$(\boldsymbol{\beta})$ $\qquad \boldsymbol{A_pB_p \to AB}$ *και* συγκεκριμένα,

$$\boldsymbol{\alpha_p B_p \to \alpha B}, \quad \textit{για οποιαδήποτε ακολουθία} \quad \boldsymbol{\alpha_p \to \alpha} \textit{ στο } \mathbb{C}.$$

(γ) $$\boldsymbol{Q A_p Q^{-1} \to Q A Q^{-1}}.$$

**(δ)** $$\mathbf{diag}\,[\boldsymbol{A_p, B_p}] \to \mathbf{diag}[\boldsymbol{A, B}].$$

Θεωρούμε τώρα την περίπτωση, στην οποία τα μέλη μιας ακολουθίας πινάκων ορίζονται, ως συναρτήσεις του πίνακα $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$. Συνεπώς, έστω ότι η ακολουθία συναρτήσεων $\boldsymbol{f_1(\lambda), f_2(\lambda), \ldots}$ ορίζεται στο φάσμα του πίνακα $\boldsymbol{A}$ και ορίζουμε

$$\boldsymbol{A_p \triangleq f_p(A)}\,, \quad \boldsymbol{p = 1, 2, \ldots}$$

Για κάθε όρο της ακολουθίας $\boldsymbol{A_1, A_2, \ldots}$ , το Θεώρημα ***4.1.1***, μας δίνει

$$\boldsymbol{f_p(A) = \sum_{k=1}^{s} \sum_{j=0}^{m_k-1} f_p^{(j)}(\lambda_k) Z_{kj}} \ . \tag{6.1.1}$$

Τώρα οι πίνακες-συντελεστές $\boldsymbol{Z_{kj}}$ εξαρτώνται μόνο από τον πίνακα $\boldsymbol{A}$ και όχι από το $\boldsymbol{p}$, έτσι ώστε κάθε όρος της ακολουθίας να ορίζεται από τους $\boldsymbol{m}$ αριθμούς $\boldsymbol{f_p^{(j)}(\lambda_k)}$, για $\boldsymbol{k = 1, 2, \ldots, s}$ και $\boldsymbol{j = 0, 1, \ldots, m_k - 1}$. Θα μπορούσαμε επομένως να περιμένουμε, ότι η σύγκλιση των $\boldsymbol{A_1, A_2, \ldots}$ μπορεί να δημιουργηθεί, έτσι ώστε να εξαρτάται από την σύγκλιση των $\boldsymbol{m}$ βαθμωτών ακολουθιών

$$\boldsymbol{f_1^{(j)}(\lambda_k), f_2^{(j)}(\lambda_k), \ldots}\ ,$$

παρά στην σύγκλιση των $\boldsymbol{n^2}$ βαθμωτών ακολουθιών των στοιχείων των πινάκων $\boldsymbol{A_1, A_2, \ldots}$

**Θεώρημα 6.1.1.** *Έστω οι συναρτήσεις* $\boldsymbol{f_1, f_2, \ldots}$ *ορίζονται στο φάσμα του πίνακα* $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$ *και έστω*

$$\boldsymbol{A_p = f_p(A)}, \quad \boldsymbol{p = 1, 2, \ldots}$$

*Τότε, η ακολουθία* $\boldsymbol{A_1, A_2, \ldots, A_p}$ *συγκλίνει, καθώς* $\boldsymbol{p \to \infty}$*, αν και μόνο αν οι* $\boldsymbol{m}$ *βαθμωτές ακολουθίες*

$$\boldsymbol{f_1^{(j)}(\lambda_k), f_2^{(j)}(\lambda_k), \ldots, f_p^{(j)}(\lambda_k), \ldots}\ ,$$

$k = 1, 2, \ldots, s, \quad j = 0, 1, \ldots, m_k - 1, \quad m = m_1 + m_2 + \ldots + m_s,$ *συγκλίνουν, καθώς* $p \to \infty$.

*Επιπλέον, αν*

$$f_p^{(j)}(\lambda_k) \to f^{(j)}(\lambda_k), \quad \textit{καθώς} \quad p \to \infty,$$

*για κάθε* $j$ *και* $k$ *και για μια συναρτήση* $f(\lambda)$, *τότε*

$$A_p \to f(A), \quad \textit{καθώς} \quad p \to \infty.$$

*Αντίστροφα, αν* $\lim A_p$ *υπάρχει, καθώς* $p \to \infty$, *τότε υπάρχει μια συνάρτηση* $f(\lambda)$ *που ορίζεται στο φάσμα του πίνακα* $A$, *τέτοια ώστε*

$$\lim A_p = f(A), \quad \textit{καθώς} \quad p \to \infty.$$

**Απόδειξη.** Αν θεωρήσουμε, ότι

$$f_p^{(j)}(\lambda_k) \to f^{(j)}(\lambda_k), \quad \textit{καθώς} \quad p \to \infty,$$

για κάθε $j$ και $k$ και για μια συνάρτηση $f(\lambda)$, τότε από την εξίσωση $(6.1.1)$ και την Πρόταση ***6.1.1*** έχουμε

$$\lim_{p\to\infty} A_p = \lim_{p\to\infty} f_p(A) = \sum_{k=1}^{s} \sum_{j=0}^{m_k-1} \left( \lim_{p\to\infty} f_p^{(j)}(\lambda_k) \right) Z_{kj} \qquad (6.1.2)$$

και

$$\lim_{p\to\infty} A_p = \sum_{k=1}^{s} \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{kj} = f(A). \qquad (6.1.3)$$

Αντίστροφα, υποθέτουμε ότι $\lim A_p$ υπάρχει, καθώς $p \to \infty$ και συμβολίζουμε τα στοιχεία του

$$A_p = f_p(A) \quad \text{με} \quad \alpha_{ir}^{(p)}, \quad i,\, r = 1, 2, \ldots, n, \quad p = 1, 2, \ldots$$

Βλέποντας την εξίσωση $(6.1.1)$ σαν ένα σύνολο από $n^2$ εξισώσεις για τις

$$m = m_1 + m_2 + \ldots + m_s,$$

άγνωστες συναρτήσεις $f_p^{(j)}(\lambda_k)$, όπου $p$ είναι σταθερό και ανακαλώντας την γραμμική ανεξαρτησία (Θεώρημα ***4.1.1***) των πινάκων-συντελεστών $Z_{kj}$, ως μέλοι του χώρου $\mathbb{C}^{n\times n}$, συμπεραίνουμε ότι ο $\boldsymbol{n^2 \times m}$ πίνακας-συντελεστής αυτού του συστήματος έχει βαθμό $\boldsymbol{m}$. Αφού ο βαθμός $\boldsymbol{m}$ του ελαχίστου πολυωνύμου του πίνακα $\boldsymbol{A}$ δεν ξεπερνά το βαθμό του χαρακτηριστικού του πολυώνυμου, συνεπάγεται ότι $\boldsymbol{m \le n \le n^2}$ και επομένως το σύστημα της εξίσωσης $\mathbf{(6.1.1)}$ έχει μια μοναδική λύση της μορφής

$$\boldsymbol{f_p^{(j)}(\lambda_k) = \sum_{i,\, r=1}^{n} \gamma_{ir} \alpha_{ir}^{(p)}}, \tag{6.1.4}$$

όπου τα $\boldsymbol{\gamma_{ir}}$ $\boldsymbol{(i,\, r = 1, 2, \ldots, n)}$ εξαρτώνται από το $\boldsymbol{j}$ και $\boldsymbol{k}$ αλλά όχι από το $\boldsymbol{p}$. Τώρα αν το όριο $\boldsymbol{\lim A_p}$ υπάρχει, καθώς $\boldsymbol{p \to \infty}$, τότε υπάρχει και το όριο $\boldsymbol{\lim \alpha_{ir}^{(p)}}$ και από την εξίσωση $\mathbf{(6.1.4)}$ προκύπτει, ότι το όριο $\boldsymbol{f_p^{(j)}(\lambda_k)}$ υπάρχει επίσης, καθώς $\boldsymbol{p \to \infty}$, για κάθε $\boldsymbol{j}$ και $\boldsymbol{k}$. Έτσι ολοκλήρωνεται το πρώτο κομμάτι του θεωρήματος.

Αν θεωρήσουμε, ότι

$$\boldsymbol{f_p^{(j)}(\lambda_k) \to f^{(j)}(\lambda_k)}, \quad \text{καθώς} \quad \boldsymbol{p \to \infty},$$

για κάθε $\boldsymbol{j}$ και $\boldsymbol{k}$ και για μια συνάρτηση $\boldsymbol{f(\lambda)}$, τότε η εξίσωση $\mathbf{(6.1.3)}$ μας δείχνει, ότι

$$\boldsymbol{A_p \to f(A)}, \quad \text{καθώς} \quad \boldsymbol{p \to \infty}.$$

Αντίστροφα, η ύπαρξη του ορίου $\boldsymbol{\lim A_p}$, καθώς $\boldsymbol{p \to \infty}$, εξασφαλίζει την ύπαρξη του ορίου $\boldsymbol{\lim f_p^{(j)}(\lambda_k)}$, καθώς $\boldsymbol{p \to \infty}$ και επομένως η εξίσωση $\mathbf{(6.1.2)}$ ισχύει. Συνεπώς, θα μπορούσαμε να διαλέξουμε για $\boldsymbol{f(\lambda)}$, το πολυώνυμο παρεμβολής, του οποίου οι τιμές στο φάσμα του πίνακα $\boldsymbol{A}$ είναι οι αριθμοί

$$\boldsymbol{\lim_{p\to\infty} f_p^{(j)}(\lambda_k)}, \quad \boldsymbol{k = 1, 2, \ldots, s}, \quad \boldsymbol{j = 0, 1, \ldots, m_k - 1}. \quad \blacksquare$$

Τώρα θα θέλαμε να αναπτύξουμε ένα παρόμοιο θεώρημα για τις σειρές πινάκων. Αρχικά, η ***σειρά*** των πινάκων $\boldsymbol{\sum_{p=0}^{\infty} A_p}$ λέγεται ότι ***συγκλίνει*** στον πίνακα $\boldsymbol{A}$, (το ***άθροισμα*** της σειράς) αν η ακολουθία των μερικών αθροισμάτων $\boldsymbol{B_\nu = \sum_{p=1}^{\nu} A_p}$ συγκλίνει στον πίνακα $\boldsymbol{A}$, καθώς $\boldsymbol{\nu \to \infty}$. Μια σειρά που δεν συγκλίνει λέγεται ότι αποκλίνει.

Το επόμενο θεώρημα προκύπτει άμεσα, αν εφαρμόσουμε το Θεώρημα ***6.1.1*** σε μια ακολουθία μερικών αθροισμάτων.

**Θεώρημα 6.1.2.** *Αν οι συναρτήσεις $\boldsymbol{u_1, u_2, \dots}$ ορίζονται στο φάσμα του πίνακα $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$, τότε η σειρά*

$$\sum_{p=1}^{\infty} \boldsymbol{u_p(A)},$$

*συγκλίνει, αν και μόνο αν η σειρά*

$$\sum_{p=1}^{\infty} \boldsymbol{u_p^{(j)}(\lambda_k)},$$

*συγκλίνει, για* $\boldsymbol{k = 1, 2, \dots, s}$ *και* $\boldsymbol{j = 0, 1, \dots, m_k - 1}$.

*Επιπλέον, αν*

$$\sum_{p=1}^{\infty} \boldsymbol{u_p^{(j)}(\lambda_k) = f^{(j)}(\lambda_k)},$$

*για κάθε* $\boldsymbol{j}$ *και* $\boldsymbol{k}$, *τότε*

$$\sum_{p=1}^{\infty} \boldsymbol{u_p(A) = f(A)},$$

*και αντίστροφα.*

Μπορούμε τώρα να εφαρμόσουμε αυτό το θεώρημα για να λάβουμε ένα σημαντικό αποτέλεσμα το οποίο, μας βοηθά στην αιτιολόγηση της ανάπτυξης της θεωρίας των συναρτήσεων των πινάκων. Το επόμενο θεώρημα δείχνει την ασάφεια του ορισμού που έγινε στην Παράγραφο **2.3**, ο οποίος ίσως φάνηκε αρκετά αυθαίρετος στην πρώτη ανάγνωση. Πραγματικά, οι ιδιότητες που περιγράφονται στο επόμενο θεώρημα χρησιμοποιούνται συχνά σαν (περισσότερο αυστηρός) ορισμός μιας συνάρτησης ενός πίνακα.

**Θεώρημα 6.1.3.** *Έστω ότι η συνάρτηση $\boldsymbol{f(\lambda)}$ αναπαριστάται από μια σειρά Taylor στο σημείο $\boldsymbol{\lambda_0} \in \mathbb{C}$ :*

$$\boldsymbol{f(\lambda)} = \sum_{p=0}^{\infty} \boldsymbol{\alpha_p(\lambda - \lambda_0)^p}, \tag{6.1.5}$$

*με ακτίνα σύγκλισης $\boldsymbol{r}$. Αν $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$, τότε $\boldsymbol{f(A)}$ ορίζεται και δίνεται από τον τύπο*

$$\boldsymbol{f(A)} = \sum_{p=0}^{\infty} \boldsymbol{\alpha_p(A - \lambda_0 I)^p}, \tag{6.1.6}$$

*αν και μόνο αν κάθε μια από τις διακεκριμένες ιδιοτιμές* $\boldsymbol{\lambda_1, \lambda_2, \ldots, \lambda_s}$ *του πίνακα* $\boldsymbol{A}$ *ικανοποιεί μια από τις ακόλουθες συνθήκες :*

$(\boldsymbol{\alpha})$ $\quad |\boldsymbol{\lambda_k - \lambda_0}| < \boldsymbol{r}.$

$(\boldsymbol{\beta})$ $\quad |\boldsymbol{\lambda_k - \lambda_0}| = \boldsymbol{r}$

*και η σειρά για* $\boldsymbol{f^{(m_k-1)}(\lambda)}$ *(όπου* $\boldsymbol{m_k}$ *είναι ο δείκτης της ιδιοτιμής* $\boldsymbol{\lambda_k}$*) είναι συγκλίνουσα στο σημείο* $\boldsymbol{\lambda = \lambda_k}$ , $\boldsymbol{1 \leq k \leq s}$.

**Απόδειξη.** Πρώτα ανακαλούμε, ότι $\boldsymbol{f(A)}$ υπάρχει, αν και μόνο αν $\boldsymbol{f^{(j)}(\lambda_k)}$ υπάρχει, για $\boldsymbol{j = 0, 1, \ldots, m_k - 1}$ και $\boldsymbol{k = 1, 2, \ldots, s}$. Επιπλέον, η συνάρτηση $\boldsymbol{f(\lambda)}$ που δίνεται στην εξίσωση $\mathbf{(6.1.5)}$ έχει παραγώγους όλων των τάξεων σε οποιοδήποτε σημείο μέσα στον κύκλο σύγκλισης και αυτές οι παράγωγοι μπορούν να αποκτηθούν παραγωγίζοντας τη σειρά για τη συνάρτηση $\boldsymbol{f(\lambda)}$ όρο προς όρο. Συνεπώς, $\boldsymbol{f(\lambda)}$ είναι προφανώς ορισμένη στο φάσμα του πίνακα $\boldsymbol{A}$, αν

$$|\boldsymbol{\lambda_k - \lambda_0}| < \boldsymbol{r}, \quad \text{για} \quad \boldsymbol{k = 1, 2, \ldots, s}.$$

Επίσης, αν

$$|\boldsymbol{\lambda_k - \lambda_0}| = \boldsymbol{r},$$

για μερικά $\boldsymbol{k}$ και η σειρά που προκύπτει από $\boldsymbol{m_k - 1}$ διαφορίσεις της σειράς στην εξίσωση $\mathbf{(6.1.5)}$, είναι συγκλίνουσα στο σημείο $\boldsymbol{\lambda = \lambda_k}$, τότε

$$\boldsymbol{f(\lambda_k), f'(\lambda_k), \ldots, f^{(m_k-1)}(\lambda_k)},$$

υπάρχουν και συνεπώς $\boldsymbol{f(A)}$ ορίζεται. Σημειώνουμε, ότι ξεκινώντας με την ύπαρξη του $\boldsymbol{f(A)}$, το προηγούμενο επιχείρημα μπορεί να αντιστραφεί.

Αν τώρα γράψουμε

$$\boldsymbol{u_p(\lambda) = \alpha_p(\lambda - \lambda_0)^p}, \quad \text{για} \quad \boldsymbol{p = 0, 1, 2, \ldots} \ ,$$

τότε $\boldsymbol{u_p(\lambda)}$ είναι σίγουρα ορισμένο στο φάσμα του πίνακα $\boldsymbol{A}$ και υπό την υπόθεση $(\boldsymbol{\alpha})$ ή $(\boldsymbol{\beta})$, η σειρά

$$\sum_{\boldsymbol{p=1}}^{\infty} \boldsymbol{u_p(\lambda_k)}, \quad \boldsymbol{k = 1, 2, \ldots, s}, \quad \boldsymbol{j = 0, 1, \ldots m_k - 1},$$

συγκλίνει. Προκύπτει τώρα από το Θεώρημα $\mathbf{(6.1.2)}$, ότι $\displaystyle\sum_{\boldsymbol{p=0}}^{\infty} \boldsymbol{u_p(A)}$ συγκλίνει και έχει άθροισμα $\boldsymbol{f(A)}$ και αντίστροφα. ■

Σημειώνουμε ότι για τις συναρτήσεις της μορφής $\mathbf{(6.1.5)}$, μια συνάρτηση $\boldsymbol{f(A)}$ του πίνακα $\boldsymbol{A}$ μπορεί να οριστεί από τον τύπο $\mathbf{(6.1.6)}$. Όμως είναι ξεκάθαρο, ότι αυτός ο ορισμός δεν μπορεί να εφαρμοστεί για συναρτήσεις που έχουν ένα πεπερασμένο αριθμό παραγώγων στα σημεία $\boldsymbol{\lambda_k} \in \boldsymbol{\sigma(A)}$, όπως υποθέσαμε και στην Παράγραφο $\mathbf{2.3}$.

Αν η σειρά Taylor για τη συνάρτηση $\boldsymbol{f(\lambda)}$ στην αρχή των αξόνων συγκλίνει για όλα τα σημεία στη μιγαδική περιοχή, (η ακτίνα της σύγκλισης είναι άπειρη) τότε $\boldsymbol{f(\lambda)}$ λέγεται ***ολόμορφη*** συνάρτηση και η αναπαράσταση δυναμοσειράς για $\boldsymbol{f(A)}$ θα συγκλίνει για όλους τους πίνακες $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$. Ανάμεσα στις πιο σημαντικές συναρτήσεις αυτού του είδους, βρίσκονται οι τριγωνομετρικές και οι εκθετικές συναρτήσεις. Συνεπώς, για τα αντίστοιχα αποτελέσματα για τις μιγαδικές βαθμωτές συναρτήσεις, συμπεραίνουμε ότι για όλους τους πίνακες $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$,

$$\sin A = \sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+1)!} A^{2p+1},$$

$$\cos A = \sum_{p=0}^{\infty} \frac{(-1)^p}{(2p)!} A^{2p},$$

$$e^A = \sum_{p=0}^{\infty} \frac{A^p}{p!},$$

$$\sinh A = \sum_{p=0}^{\infty} \frac{A^{2p+1}}{(2p+1)!},$$

$$\cosh A = \sum_{p=0}^{\infty} \frac{A^{2p}}{(2p)!}.$$

Αν οι ιδιοτιμές $\boldsymbol{\lambda_1, \lambda_2, \dots, \lambda_n}$ του $\boldsymbol{A}$ ικανοποιούν τις συνθήκες $|\boldsymbol{\lambda_j}| < \mathbf{1}$, για $\boldsymbol{j = 1, 2, \dots, n}$, τότε συμπεραίνουμε από το διωνυμικό θεώρημα και τη λογαριθμική σειρά, ότι

$$(I - A)^{-1} = \sum_{p=0}^{\infty} A^p,$$

$$\log(I + A) = \sum_{p=1}^{\infty} \frac{(-1)^{p-1}}{p} A^p. \tag{6.1.7}$$

# Κεφάλαιο 7

## 7.1 Η Επιλύουσα και Ορισμός μιας Συνάρτησης Πίνακα μέσω του Θεωρήματος Cauchy

Η θεωρία των (βαθμωτών) συναρτήσεων μιας μιγαδικής μεταβλητής παρέχει μια τρίτη προσέγγιση, για τον ορισμό του $f(A)$ που μπορεί να εφαρμοστεί όταν $f(\lambda)$ είναι μια αναλυτική συνάρτηση. Αυτή η προσέγγιση είναι επίσης συνεπής, ως προς τον γενικό ορισμό της Παραγράφου **2.3**.

Θεωρούμε ένα πίνακα $A(\lambda)$, του οποίου τα στοιχεία είναι συναρτήσεις μιας μιγαδικής μεταβλητής $\lambda$ και ορίζουμε την ***παράγωγο*** (ή ***ολοκλήρωμα***) του πίνακα, να είναι ο πίνακας που παίρνουμε διαφορίζοντας (ή ολοκληρώνοντας) κάθε στοιχείο του $A(\lambda)$. Για $r = 0, 1, 2, \ldots$ γράφουμε

$$\frac{d^r}{d\lambda^r} A(\lambda),$$

ή

$$A^{(r)}(\lambda),$$

για την παράγωγο του $A(\lambda)$ τάξης $r$.

Ο συμβολισμός

$$\int_L A(\lambda) d\lambda,$$

θα χρησιμοποιηθεί για το ολοκλήρωμα του $A(\lambda)$ πάνω στη θετική διεύθυνση κατά μήκος ενός μονοπατιού $L$ στην μιγαδική περιοχή, το οποίο για τους σκοπούς μας θα είναι πάντα ένα απλό κατά σημείο λείο κλειστό περίγραμμα, ή ένα πεπερασμένο σύστημα τέτοιων περιγραμμάτων, χωρίς σημεία τομής.

Έστω $A \in \mathbb{C}^{n \times n}$. Ο πίνακας $(\lambda I - A)^{-1}$ είναι γνωστός ως ***η επιλύουσα του*** πίνακα $A$ και έχει παίξει ήδη ένα σημαντικό ρόλο στην ανάλυσή μας. Αν την δούμε, ως συνάρτηση μιας μιγαδικής μεταβλητής $\lambda$, η επιλύουσα ορίζεται μόνο σε εκείνα τα σημεία που δεν ανήκουν στο φάσμα του πίνακα $A$. Αφού ο πίνακας $A$ είναι σταθερός σε αυτή την παράγραφο, χρησιμοποιούμε τον συμβολισμό

$$R_\lambda = (\lambda I - A)^{-1},$$

για την επιλύουσα του πίνακα $A$.

**Θεώρημα 7.1.1.** *Η επιλύουσα $\boldsymbol{R_\lambda}$ του πίνακα $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$ είναι μια ρητή συνάρτηση του $\boldsymbol{\lambda}$ με πόλους, τα σημεία του φάσματος του πίνακα $\boldsymbol{A}$ και $\boldsymbol{R_\infty} = \boldsymbol{0}$. Επιπλέον, κάθε $\boldsymbol{\lambda_k} \in \boldsymbol{\sigma(A)}$ είναι ένας πόλος του $\boldsymbol{R_\lambda}$ τάξης $\boldsymbol{m_k}$, όπου $\boldsymbol{m_k}$ είναι ο δείκτης της ιδιοτιμής $\boldsymbol{\lambda_k}$.*

**Απόδειξη.** Έστω ότι

$$\boldsymbol{m(\lambda)} = \sum_{\boldsymbol{i=0}}^{\boldsymbol{m}} \boldsymbol{\alpha_i \lambda^i},$$

είναι το ελάχιστο πολυώνυμο του πίνακα $\boldsymbol{A}$. Εύκολα φαίνεται ότι

$$\frac{\boldsymbol{m(\lambda) - m(\mu)}}{\boldsymbol{\lambda - \mu}} = \sum_{\boldsymbol{i,\, j=0}}^{\boldsymbol{m-1}} \gamma_{ij} \boldsymbol{\lambda^i \mu^j},$$

$\boldsymbol{\lambda \neq \mu}$, για κάποια $\gamma_{ij}$ $\boldsymbol{(1 \leq i, \quad j \leq m-1)}$
και επομένως, αντικαθιστώντας το $\boldsymbol{\mu}$ με τον πίνακα $\boldsymbol{A}$, έχουμε για $\boldsymbol{\lambda \notin \sigma(A)}$

$$\boldsymbol{(m(\lambda)I - m(A))R_\lambda} = \sum_{\boldsymbol{j=0}}^{\boldsymbol{m-1}} \left( \sum_{\boldsymbol{i=0}}^{\boldsymbol{m-1}} \gamma_{ij} \boldsymbol{\lambda^i} \right) \boldsymbol{A^j}.$$

Αφού $\boldsymbol{m(A) = 0}$, προκύπτει, ότι για τις τιμές του $\boldsymbol{\lambda}$, για τις οποίες $\boldsymbol{m(\lambda) \neq 0}$,

$$\boldsymbol{R_\lambda} = \frac{\boldsymbol{1}}{\boldsymbol{m(\lambda)}} \sum_{\boldsymbol{j=0}}^{\boldsymbol{m-1}} \boldsymbol{m_j(\lambda) A^j}, \qquad \textbf{(7.1.1)}$$

όπου $\boldsymbol{m_j(\lambda)}$ $\boldsymbol{(0 \leq j \leq m-1)}$, είναι ένα πολυώνυμο βαθμού, ο οποίος δεν ξεπερνά το $\boldsymbol{m-1}$. Τα απαιτούμενα αποτελέσματα τώρα ακολουθούν από την αναπαράσταση **(7.1.1)**. ■

Συνεπώς, χρησιμοποιώντας την ορολογία της μιγαδικής ανάλυσης, το φάσμα ενός πίνακα μπορεί να περιγραφεί σε όρους της επιλύουσάς του. Αυτό το γενονός, μας παροτρύνει να εφαρμόσουμε αποτελέσματα της μιγαδικής ανάλυσης στην θεωρία πινάκων και αντίστροφα. Στο επόμενο θεώρημα χρησιμοποιούμε το κλασικό θεώρημα ολοκλήρωσης του Cauchy, για να πάρουμε την γενίκευσή του, για την περίπτωση των πινάκων και αυτό αποδεικνύεται να είναι ένα σημαντικό και χρήσιμο αποτέλεσμα στην θεωρία των πινάκων.

Αρχικά, ας συμφωνήσουμε στο να καλούμε μια συνάρτηση $\boldsymbol{f(\lambda)}$ της μιγαδικής μεταβλητής $\boldsymbol{\lambda}$, ***αναλυτική*** πάνω σε ένα ανοικτό σύνολο σημείων του $\boldsymbol{D}$ της μιγαδικής γειτονιάς, αν $\boldsymbol{f(\lambda)}$ είναι συνεχώς διαφορίσιμη σε κάθε σημείο του $\boldsymbol{D}$. Ή, ισοδύναμα, $\boldsymbol{f(\lambda)}$ είναι αναλυτική αν έχει ως αναπαράσταση μια συγκλίνουσα σειρά Taylor, για κάθε σημείο του $\boldsymbol{D}$. Αν $\boldsymbol{f(\lambda)}$ είναι αναλυτική μέσα σε μια ανοικτή περιοχή $\boldsymbol{D}$ που φράσσεται από ένα περίγραμμα $\boldsymbol{L}$ και είναι συνεχής πάνω στην κλειστότητα $\overline{\boldsymbol{D}}$, τότε το θεώρημα Cauchy μας διαβεβαιώνει, ότι

$$\boldsymbol{f^{(j)}(\lambda_0) = \frac{j!}{2\pi i}\int_L \frac{f(\lambda_0)}{(\lambda-\lambda_0)^{j+1}} d\lambda}, \qquad \mathbf{(7.1.2)}$$

για οποιοδήποτε $\boldsymbol{\lambda_0 \in D}$ και $\boldsymbol{j = 0, 1, \dots}$

**Θεώρημα 7.1.2.** *Αν $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$ έχει διακεκριμένες ιδιοτιμές $\boldsymbol{\lambda_1, \dots, \lambda_s}$, το μονοπάτι $\boldsymbol{L}$ είναι ένα κλειστό περίγραμμα το οποίο έχει $\boldsymbol{\lambda_1, \dots, \lambda_s}$, στο εσωτερικό του και $\boldsymbol{f(\lambda)}$ είναι συνεχής στο εσωτερικό και στο σύνορο του $\boldsymbol{L}$, τότε*

$$\boldsymbol{f(A) = \frac{1}{2\pi i}\int_L f(\lambda)(\lambda I - A)^{-1} d\lambda = \frac{1}{2\pi i}\int_L f(\lambda) R_\lambda d\lambda}. \qquad \mathbf{(7.1.3)}$$

**Απόδειξη.** Αν $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$ και $\boldsymbol{\lambda \notin \sigma(A)}$, έχουμε ότι

$$\boldsymbol{R_\lambda(A) \triangleq (\lambda I - A)^{-1} = \sum_{k=1}^{s}\sum_{j=0}^{m_k-1} \frac{j!}{(\lambda-\lambda_k)^{j+1}} Z_{kj}}.$$

Αν πολλαπλασιάσουμε και τα δύο μέλη με $\boldsymbol{f(\lambda)(2\pi i)^{-1}}$ και εφαρμόσουμε τον τύπο της σχέσης $\mathbf{(7.1.2)}$ για $\boldsymbol{\lambda_0 = \lambda_k}$ $\boldsymbol{(1 \le k \le s)}$ και ολοκληρώσουμε κατά μήκος του $\boldsymbol{L}$, παίρνουμε

$$\boldsymbol{\frac{1}{2\pi i}\int_L f(\lambda)(\lambda I - A)^{-1} d\lambda = \sum_{k=1}^{s}\sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{kj}} \ ,$$

και το αποτέλεσμα προκύπτει από τη φασματική ανάλυση του Θεωρήματος ***4.1.1*** ■

Σημειώνουμε ότι η σχέση $\mathbf{(7.1.3)}$ είναι ακριβώς ανάλογη της $\mathbf{(7.1.2)}$ στην περίπτωση που $\boldsymbol{j = 0}$. Το $\boldsymbol{(\lambda - \lambda_0)^{-1}}$ έχει αντικατασταθεί από $\boldsymbol{(\lambda I - A)^{-1}}$. Θα επισημάνουμε επίσης, ότι η σχέση $\mathbf{(7.1.3)}$ χρησιμοποιείται συχνά στον ορισμό του $\boldsymbol{f(A)}$ για αναλυτικές συναρτήσεις $\boldsymbol{f(\lambda)}$.

**Συμπέρασμα 7.1.1.** *Με τον συμβολισμό του Θεωρήματος **7.1.2**,*

$$\boldsymbol{I} = \frac{1}{2\pi i}\int_L \boldsymbol{R_\lambda d\lambda}\,, \quad \boldsymbol{A} = \frac{1}{2\pi i}\int_L \boldsymbol{\lambda R_\lambda d\lambda}.$$

*Οι πίνακες-συντελεστές $\boldsymbol{Z_{kj}}$ του πίνακα $\boldsymbol{A}$ μπορούν επίσης να περιγραφούν σε όρους της επιλύουσας του $\boldsymbol{A}$.*

**Θεώρημα 7.1.3.** *Διατηρώντας τον προηγούμενο συμβολισμό, έστω $\boldsymbol{L_k}$ ένα απλό κατά σημείο λείο κλειστό περίγραμμα, που περιέχει $\boldsymbol{\lambda_k \in \sigma(A)}$ στο εσωτερικό του και υποθέτουμε, ότι όχι άλλα σημεία του φάσματος του πίνακα $\boldsymbol{A}$ βρίσκονται στο εσωτερικό, ή πάνω στο σύνορο του $\boldsymbol{L_k}$. Τότε*

$$\boldsymbol{Z_{kj}} = \frac{1}{j!2\pi i}\int_{L_k} \boldsymbol{(\lambda - \lambda_k)^j R_\lambda d\lambda}, \tag{7.1.4}$$

*για* $\boldsymbol{j = 0, 1, \dots, m_k - 1}, \quad \boldsymbol{k = 1, 2, \dots, s}.$

**Απόδειξη.** Πολλαπλασιάζοντας και τα δύο μέλη της εξίσωσης **(4.1.2)** με

$$\boldsymbol{(\lambda - \lambda_p)^r, \quad 1 \le p \le s, \quad 0 \le r \le m_p - 1}$$

και ολοκληρώνοντας γύρω από $\boldsymbol{L_p}$, παίρνουμε ότι

$$\int_{L_p} \boldsymbol{(\lambda - \lambda_p)^r R_\lambda d\lambda} = \sum_{k=1}^{s}\sum_{j=0}^{m_k-1} \boldsymbol{j! Z_{kj}} \int_{L_p} \frac{\boldsymbol{(\lambda - \lambda_p)^r}}{\boldsymbol{(\lambda - \lambda_k)^{j+1}}} \boldsymbol{d\lambda}. \tag{7.1.5}$$

Οι συναρτήσεις $\boldsymbol{(\lambda - \lambda_p)^r(\lambda - \lambda_k)^{-j-1}}$ με $\boldsymbol{k \neq p}$ ή $\boldsymbol{k = p}$ και $\boldsymbol{j \le r - 1}$ είναι αναλυτικές στην περιοχή που φράσσεται από $\boldsymbol{L_p}$. Επομένως, τα ολοκληρώματα αυτών των συναρτήσεων γύρω από $\boldsymbol{L_p}$ είναι ίσα με μηδέν. Τώρα παρατηρούμε ότι για τις εναπομείναντες περιπτώσεις, έχοντας υπόψη την εξίσωση **(7.1.2)** για $\boldsymbol{f(\lambda) \equiv 1}$,

$$\int_{L_p} (\lambda - \lambda_p)^{r-j-1} d\lambda = \begin{cases} 2\pi i & \text{αν} \quad j = r, \\ 0 & \text{αν} \quad j > r. \end{cases}$$

Συνεπώς, η εξίσωση **(7.1.5)** υποδεικνύει

$$\int_{L_p} (\lambda - \lambda_p)^r R_\lambda d\lambda = r! Z_{pr}(2\pi i),$$

το οποίο είναι ισοδύναμο με την σχέση **(7.1.4)**. ■

Τώρα αν θυμηθούμε τις επεκτάσεις Laurent, η εξίσωση **(7.1.4)** εκφράζει το $Z_{kj}$ σαν ένα συντελεστή της επέκτασης Laurent για την επίλυση $R_\lambda$ στο σημείο $\lambda_k$.

Χρησιμοποιώντας τα αποτελέσματα των Θεωρημάτων ***4.1.2*** και ***4.1.4*** και του αποτελέσματος του Θεωρήματος ***7.1.3*** για $j = 0$, το ακόλουθο χρήσιμο αποτέλεσμα προκύπτει.

**Θεώρημα 7.1.4.** *Με τον προηγούμενο συμβολισμό, το ολοκλήρωμα*

$$P_k = Z_{k0} = \frac{1}{2\pi i} \int_{L_k} R_\lambda d\lambda, \quad 1 \leq k \leq s,$$

*που καλείται προβολή Riesz, είναι μια προβολή. Το σύστημα* $\{P_k\}_{k=1}^{s}$ *είναι ένα ορθογώνιο σύστημα από προβολές υπό την έννοια ότι*

$$P_k P_j = 0, \quad \text{για} \quad k \neq j\ ,$$

*δίνοντας μια φασματική ανάλυση του ταυτοτικού πίνακα (ως προς τον πίνακα* $A$*):*

$$I = \sum_{k=1}^{s} P_k \quad \text{και} \quad \operatorname{Im} P_k = \operatorname{Ker}(A - \lambda_k I)^{m_k},$$

*όπου* $m_k$ *είναι ο δείκτης της ιδιοτιμής* $\lambda_k$ *του πίνακα* $A$.

# Κεφάλαιο 8

## 8.1 Οι Συναρτήσεις Πινάκων στις Διαφορικές Εξισώσεις

Θεωρούμε τα συστήματα διαφορικών εξισώσεων με σταθερούς συντελεστές. Θα θέλαμε να βρούμε συναρτήσεις του χρόνου, $\boldsymbol{x(t)}$, με τιμές στον χώρο $\mathbb{C}^{\boldsymbol{n}}$, που ικανοποιούν την ομογενή εξίσωση

$$\dot{\boldsymbol{x}}(\boldsymbol{t}) = \boldsymbol{A}\boldsymbol{x}(\boldsymbol{t}), \tag{8.1.1}$$

ή την μη ομογενή εξίσωση

$$\dot{\boldsymbol{x}}(\boldsymbol{t}) = \boldsymbol{A}\boldsymbol{x}(\boldsymbol{t}) + \boldsymbol{f}(\boldsymbol{t}), \tag{8.1.2}$$

όπου $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n}\times\boldsymbol{n}}$ και $\boldsymbol{f(t)}$ είναι μια συνάρτηση που καθορίζεται από τον χρόνο. Θα αναλύσουμε τώρα αυτά τα προβλήματα χωρίς να κάνουμε υποθέσεις πάνω στον $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n}\times\boldsymbol{n}}$.

Οι λύσεις της μορφής $\boldsymbol{x_0 e^{\lambda_0 t}}$, όπου $\boldsymbol{\lambda_0}$ είναι μια ιδιοτιμή και $\boldsymbol{x_0}$ είναι ένα σχετιζόμενο ιδιοδιάνυσμα, παίζουν έναν σημαντικό ρόλο. Γενικότερα, ψάχνουμε για μια λύση της εξίσωσης **(8.1.1)** στη μορφή

$$\boldsymbol{x}(\boldsymbol{t}) = \boldsymbol{x_0}(\boldsymbol{t})\boldsymbol{e}^{\boldsymbol{\lambda_0 t}}, \quad \boldsymbol{\lambda_0} \in \mathbb{C}, \tag{8.1.3}$$

όπου $\boldsymbol{x_0(t)}$ είναι ένα διάνυσμα τάξης $\boldsymbol{n}$, του οποίου οι συντεταγμένες είναι πολυώνυμα του $\boldsymbol{t}$. Για ευκολία γράφουμε $\boldsymbol{x_0(t)}$ στη μορφή

$$\boldsymbol{x_0}(\boldsymbol{t}) = \frac{\boldsymbol{t^{k-1}}}{\boldsymbol{(k-1)!}}\boldsymbol{x_1} + \cdots + \frac{\boldsymbol{t}}{\boldsymbol{1!}}\boldsymbol{x_{k-1}} + \boldsymbol{x_k}\ , \tag{8.1.4}$$

όπου $\boldsymbol{x_i} \in \mathbb{C}^{\boldsymbol{n}}\ ,\ \boldsymbol{i = 1, 2, \ldots, k}$ και $\boldsymbol{x_1}$ θεωρείται να είναι μη μηδενικό.

Φαίνεται εύκολα, ότι το $\boldsymbol{x(t)}$ της μορφής **(8.1.3)** είναι μια λύση της **(8.1.1)**, αν και μόνο αν

$$\dot{\boldsymbol{x}}_{\boldsymbol{0}}(\boldsymbol{t}) = (\boldsymbol{A} - \boldsymbol{\lambda_0 I})\boldsymbol{x_0}(\boldsymbol{t}),$$

όπου $\boldsymbol{x_0(t)}$ δίνεται από την εξίσωση **(8.1.4)**. Διαφορίζοντας, παίρνουμε

$$\boldsymbol{x_0^{(i)}(t) = (A - \lambda_0 I)^i x_0(t), \quad i = 1, 2, \dots} \tag{8.1.5}$$

και αφού $\boldsymbol{x_0(t)}$ είναι ένα διανυσματικά τιμώμενο πολυώνυμο βαθμού $\boldsymbol{k-1}$, προκύπτει ότι

$$\boldsymbol{0 \ = (A - \lambda_0 I)^k x_0(t)}.$$

Τώρα βρίσκουμε από την εξίσωση $\mathbf{(8.1.4)}$, ότι

$$\boldsymbol{x_0^{(k-1)}(t) = x_1},$$

και συγκρίνοντάς το με την εξίσωση $\mathbf{(8.1.5)}$,

$$\boldsymbol{x_1 = (A - \lambda_0 I)^{k-1} x_0(t)}.$$

Πολλαπλασιάζοντας στο αριστερό μέλος με $\boldsymbol{(A - \lambda_0 I)}$, παίρνουμε

$$\boldsymbol{(A - \lambda_0 I) x_1 = 0}$$

και αφού $\boldsymbol{x_1 \neq 0}$, από υπόθεση, $\boldsymbol{x_1}$ είναι ένα ιδιοδιάνυσμα του πίνακα $\boldsymbol{A}$ που αντιστοιχεί στην ιδιοτιμή $\boldsymbol{\lambda_0}$.

Τώρα συγκρίνουμε τις $\boldsymbol{(k-2)}$-στες παραγώγους του $\boldsymbol{x_0(t)}$ υπολογισμένες από την εξίσωση $\mathbf{(8.1.4)}$ και την εξίσωση $\mathbf{(8.1.5)}$, για να πάρουμε

$$\boldsymbol{t x_1 + x_2 = (A - \lambda_0 I)^{k-2} x_0(t)}.$$

Πολλαπλασιάζοντας στο αριστερό μέλος με $\boldsymbol{(A - \lambda_0 I)}$ και χρησιμοποιώντας τα αποτελέσματα της προηγούμενης παραγράφου, παίρνουμε

$$\boldsymbol{(A - \lambda_0 I) x_2 = x_1}.$$

Συνεπώς, $\boldsymbol{x_1, x_2}$ σχηματίζουν μια αλυσίδα Jordan μήκους $\mathbf{2}$.

Συνεχίζοντας με τον ίδιο τρόπο, βρίσκουμε ότι $\boldsymbol{x_1, x_2, \dots, x_k}$ δημιουργούν μια αλυσίδα Jordan για τον πίνακα $\boldsymbol{A}$ σχετιζόμενη με την ιδιοτιμή $\boldsymbol{\lambda_0}$. Ένα παρόμοιο επιχείρημα δείχνει, ότι αν $\boldsymbol{x_1, x_2, \dots, x_k}$ δημιουργούν μια αλυσίδα Jordan για τον πίνακα $\boldsymbol{A}$ που αντιστοιχεί στην ιδιοτιμή $\boldsymbol{\lambda_0}$, τότε η εξίσωση $\mathbf{(8.1.3)}$ μαζί με την εξίσωση $\mathbf{(8.1.4)}$ δίνουν μια λύση της εξίσωσης $\mathbf{(8.1.1)}$. Συνεπώς, έχουμε αποδείξει το ακόλουθο.

**Πρόταση 8.1.1.** *Η συνάρτηση $\boldsymbol{x(t)}$ που ορίζεται από τις εξισώσεις $\mathbf{(8.1.3)}$ και $\mathbf{(8.1.4)}$ είναι μια λύση της εξίσωσης $\mathbf{(8.1.1)}$, αν και μόνο αν $\boldsymbol{x_1, x_2, \ldots, x_k}$ συνιστούν μια αλυσίδα Jordan για τον πίνακα $\boldsymbol{A}$, σχετιζόμενη με την ιδιοτιμή $\boldsymbol{\lambda_0}$ του πίνακα $\boldsymbol{A}$.*

Τώρα ανακαλούμε, ότι από το Θεώρημα ***1.5.1***, η διάσταση του χώρου των λύσεων $\boldsymbol{\mathcal{F}}$ της εξίσωσης $\mathbf{(8.1.1)}$ είναι $\boldsymbol{n}$ και έχοντας υπόψη την κανονική μορφή Jordan, υπάρχει μια βάση στον $\mathbb{C}^{\boldsymbol{n}}$ που αποτελείται αποκλειστικά από ιδιοδιανύσματα και γενικευμένα ιδιοδιανύσματα του πίνακα $\boldsymbol{A}$. Συνεπώς, το ακόλουθο αποτέλεσμα είναι λογικό.

**Θεώρημα 8.1.1.** *Έστω $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$ και έστω $\boldsymbol{\lambda_1, \lambda_2, \ldots, \lambda_s}$ οι διακεκριμένες ιδιοτιμές του πίνακα $\boldsymbol{A}$ αλγεβρικής πολλαπλότητας $\boldsymbol{\alpha_1, \alpha_2, \ldots, \alpha_s}$, αντίστοιχα. Αν οι αλυσίδες Jordan για $\boldsymbol{\lambda_k}$ $\boldsymbol{(1 \leq k \leq s)}$ είναι*

$$\begin{array}{llll}
\boldsymbol{x_{11}^{(k)}}, & \boldsymbol{x_{12}^{(k)}}, & \ldots, & \boldsymbol{x_{1\nu_{k1}}^{(k)}}, \\
\boldsymbol{x_{21}^{(k)}}, & \boldsymbol{x_{22}^{(k)}}, & \ldots, & \boldsymbol{x_{2\nu_{k2}}^{(k)}}, \\
\vdots & & & \\
\boldsymbol{x_{r_k,1}^{(k)}}, & \boldsymbol{x_{r_k,2}^{(k)}}, & \ldots, & \boldsymbol{x_{r_k\nu_{k,r_k}}^{(k)}},
\end{array}$$

*όπου $\displaystyle\sum_{\boldsymbol{i=1}}^{\boldsymbol{r_k}} \boldsymbol{\nu_{ki} = \alpha_k}$, τότε οποιαδήποτε λύση $\boldsymbol{x(t)}$ της σχέσης ***(8.1.1)*** είναι ένας γραμμικός συνδυασμός των $\boldsymbol{n}$ διανυσματικά-τιμώμενων συναρτήσεων*

$$\left[\frac{\boldsymbol{t^{j-1}}}{\boldsymbol{(j-1)!}}\boldsymbol{x_{i1}^{(k)}} + \cdots + \frac{\boldsymbol{t}}{\mathbf{1!}}\boldsymbol{x_{i,\,j-1}^{(k)}} + \boldsymbol{x_{ij}^{(k)}}\right] \boldsymbol{e^{\lambda_k t}}, \qquad \mathbf{(8.1.6)}$$

όπου $\quad \boldsymbol{j = 1, 2, \ldots, \nu_{ki}}, \quad \boldsymbol{i = 1, 2, \ldots, r_k}, \quad \boldsymbol{k = 1, 2, \ldots, s}$.

Με άλλα λόγια, οι εκφράσεις στην σχέση $\mathbf{(8.1.6)}$ σχηματίζουν μια βάση στον χώρο επίλυσης $\boldsymbol{\mathcal{F}}$ της σχέσης $\mathbf{(8.1.1)}$ και επομένως αναφέρονται, ως οι ***θεμελιώδεις λύσεις*** της σχέσης $\mathbf{(8.1.1)}$. Σημειώνουμε επίσης, ότι η ιδιοτιμή $\boldsymbol{\lambda_k}$ έχει γεωμετρική πολλαπλότητα $\boldsymbol{r_k}$ και

$$\boldsymbol{x_{11}^{(k)}, x_{21}^{(k)}, \ldots, x_{r_k1}^{(j)}},$$

αποτελούν το σύνολο $\quad \mathbf{Ker}\boldsymbol{(A - \lambda_k I)}$.

**Παράδειγμα 8.1.1.**

Εφαρμόζοντας όσα καταγράψαμε έως τώρα, θα βρούμε τη γενική μορφή των λύσεων (*τη γενική μορφή της λύσης*) του συστήματος

$$\dot{\boldsymbol{x}}(\boldsymbol{t}) = \begin{bmatrix} \dot{\boldsymbol{x}}_1 \\ \dot{\boldsymbol{x}}_2 \\ \dot{\boldsymbol{x}}_3 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \end{bmatrix} = \boldsymbol{A}\boldsymbol{x}(\boldsymbol{t}).$$

Ο παραπάνω πίνακας $\boldsymbol{A}$ έχει μια ιδιοτιμή $\boldsymbol{\lambda_1 = 2}$ με ένα σχετιζόμενο ιδιοδιάνυσμα $[\mathbf{0} \; -\mathbf{1} \quad \mathbf{1}]^{\boldsymbol{T}}$ και μια ιδιοτιμή $\boldsymbol{\lambda_2 = 4}$ σχετιζόμενη με το ιδιοδιάνυσμα $[\mathbf{2} \; -\mathbf{2} \quad \mathbf{0}]^{\boldsymbol{T}}$ και με το γενικευμένο ιδιοδιάνυσμα $[\mathbf{1} \quad \mathbf{0} \quad \mathbf{0}]^{\boldsymbol{T}}$. Συνεπώς, η γενική λύση, σύμφωνα με την σχέση **(8.1.6)** δίνεται από την

$$\boldsymbol{x}(\boldsymbol{t}) = \boldsymbol{\beta}_1 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \boldsymbol{e}^{2t} + \boldsymbol{\beta}_2 \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} \boldsymbol{e}^{4t} + \boldsymbol{\beta}_3 \left( \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} \boldsymbol{t} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) \boldsymbol{e}^{4t}. \qquad \textbf{(8.1.7)}$$

□

Είναι αρκετά συχνή η περίπτωση στην οποία απαιτείται μια συγκεκριμένη λύση της σχέσης **(8.1.1)** η οποία προσδιορίζεται από επιπλέον συνθήκες. Τα ***προβλήματα αρχικής τιμής*** είναι αυτού του είδους, στα οποία ένα διάνυσμα $\boldsymbol{x_0} \in \mathbb{C}^{\boldsymbol{n}}$ δίνεται και μια λύση $\boldsymbol{x}(\boldsymbol{t})$ της σχέσης **(8.1.1)** είναι να βρεθεί, όπου $\boldsymbol{x}(\boldsymbol{t})$ παίρνει την τιμή $\boldsymbol{x_0}$ στο $\boldsymbol{t = 0}$.

**Θεώρημα 8.1.2.** *Έστω $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n \times n}}$. Υπάρχει ένα μοναδικό $\boldsymbol{x}(\boldsymbol{t})$, τέτοιο ώστε*

$$\dot{\boldsymbol{x}}(\boldsymbol{t}) = \boldsymbol{A}\boldsymbol{x}(\boldsymbol{t}) \quad \text{και} \quad \boldsymbol{x}(0) = \boldsymbol{x_0}, \qquad \textbf{(8.1.8)}$$

*για οποιοδήποτε δοσμένο διάνυσμα $\boldsymbol{x_0} \in \mathbb{C}^{\boldsymbol{n}}$.*

**Απόδειξη.** Από το Θεώρημα ***8.1.1***, οποιαδήποτε λύση $\boldsymbol{x(t)}$ της σχέσης $\mathbf{(8.1.1)}$ είναι ένας γραμμικός συνδυασμός των $\boldsymbol{n}$ εκφράσεων $\boldsymbol{x_p(t)}$ $\boldsymbol{(p = 1, 2, \ldots, n)}$ της μορφής $\mathbf{(8.1.6)}$

$$\boldsymbol{x(t) = \sum_{p=1}^{n} \beta_p x_p(t)}.$$

Παρατηρούμε, ότι $\boldsymbol{x_p(0)}$ , $\boldsymbol{p = 1, 2, \ldots, n}$, δίνει τα διανύσματα της βάσης Jordan του πίνακα $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n}}$ και επομένως ο πίνακας

$$\boldsymbol{B} = [\boldsymbol{x_1(0)} \;\; \boldsymbol{x_2(0)} \;\; \cdots \;\; \boldsymbol{x_n(0)}],$$

είναι αντιστρέψιμος. Συνεπώς, το σύστημα

$$\boldsymbol{\sum_{p=1}^{n} \beta_p x_p(0) = x_0},$$

ή σε μορφή πινάκων

$$\boldsymbol{B\beta = x_0}, \quad \text{με} \quad \boldsymbol{\beta = [b_1 \;\; b_2 \;\; \cdots \;\; b_n]^T},$$

έχει μια μοναδική λύση για κάθε $\boldsymbol{x_0} \in \mathbb{C}^{\boldsymbol{n}}$. ■

**Παράδειγμα 8.1.2.**

Θα βρούμε τη λύση του συστήματος του Παραδείγματος ***8.1.1*** που ικανοποιεί τη συνθήκη $\boldsymbol{x(0) = [0 \;\; 1 \;\; 1]^T}$. Φαίνεται εύκολα, ότι αν βάλουμε $\boldsymbol{t = 0}$ στην εξίσωση $\mathbf{(8.1.7)}$, τότε παίρνουμε ένα γραμμικό σύστημα

$$\begin{bmatrix} \mathbf{0} & \mathbf{2} & \mathbf{1} \\ \mathbf{-1} & \mathbf{-2} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta_1} \\ \boldsymbol{\beta_2} \\ \boldsymbol{\beta_3} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix},$$

με έναν αντιστρέψιμο πίνακα-συντελεστή. Το διάνυσμα $[\boldsymbol{\beta_1} \;\; \boldsymbol{\beta_2} \;\; \boldsymbol{\beta_3}]^{\boldsymbol{T}}$ βρίσκεται ότι είναι $\mathbf{[1 - 1 \;\;\; 2]}^{\boldsymbol{T}}$ και επομένως η ζητούμενη λύση είναι

$$\boldsymbol{x(t)} = \begin{bmatrix} \mathbf{0} \\ \mathbf{-1} \\ \mathbf{1} \end{bmatrix} \boldsymbol{e^{2t}} + \left( \begin{bmatrix} \mathbf{0} \\ \mathbf{2} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{4} \\ \mathbf{-4} \\ \mathbf{0} \end{bmatrix} \boldsymbol{t} \right) \boldsymbol{e^{4t}}. \quad \square$$

Τώρα θα δείξουμε πως η λύση $(\mathbf{8.1.8})$ μπορεί να βρεθεί με τη χρήση της έννοιας μιας συνάρτησης του πίνακα $\boldsymbol{A}$. Ξεκινάμε, χρησιμοποιώντας την ιδέα με την οποία αρχίσαμε την παράγραφο, αλλά την εφαρμόζουμε άμεσα στην διαφορική εξίσωση πίνακα $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}$. Ψάχνουμε για λύσεις της μορφής $\boldsymbol{e^{At}z}$ με $\boldsymbol{z} \in \mathbb{C}^{\boldsymbol{n}}$. Γνωρίζουμε ότι η διανυσματικά-τιμώμενη συνάρτηση $\boldsymbol{x(t)} = \boldsymbol{e^{At}z}$ είναι μια λύση της σχέσης $(\mathbf{8.1.1})$ για οποιοδήποτε $\boldsymbol{z} \in \mathbb{C}^{\boldsymbol{n}}$,

$$\dot{\boldsymbol{x}}(\boldsymbol{t}) = \frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{e^{At}z} = \boldsymbol{Ae^{At}z} = \boldsymbol{Ax(t)}.$$

Προφανώς $\boldsymbol{x(t)} = \boldsymbol{e^{At}z}$ ικανοποιεί την αρχική συνθήκη $\boldsymbol{x}(\mathbf{0}) = \boldsymbol{x_0}$, αν και μόνο αν $\boldsymbol{z} = \boldsymbol{x_0}$. Συνεπώς, χρησιμοποιώντας το Θεώρημα ***8.1.2***, βρίσκουμε το ακόλουθο.

**Θεώρημα 8.1.3.** *Η μοναδική λύση του προβλήματος* $(\mathbf{8.1.8})$ *μπορεί να γραφεί στη μορφή*

$$\boldsymbol{x(t)} = \boldsymbol{e^{At}x_0}. \tag{8.1.9}$$

Έστω $\boldsymbol{x(t)}$ λύση της σχέσης $(\mathbf{8.1.1})$. Υπολογίζοντας το $\boldsymbol{x_0}$, βλέπουμε από το Θεώρημα ***8.1.3*** ότι

$$\boldsymbol{x(t)} = \boldsymbol{e^{At}x}(\mathbf{0}).$$

Συνεπώς έχουμε μια περαιτέρω βελτίωση.

**Θεώρημα 8.1.4.** *Κάθε λύση* $\boldsymbol{x(t)}$ *της σχέσης* $(\mathbf{8.1.1})$ *μπορεί να αναπαρασταθεί στη μορφή*

$$\boldsymbol{x(t)} = \boldsymbol{e^{At}z},$$

*για κάποια* $\boldsymbol{z} \in \mathbb{C}^{\boldsymbol{n}}$.

Είναι αξιοσημείωτο γεγονός, το ότι η πολυπλοκότητα του Θεωρήματος ***8.1.1*** μπορεί να βελτιστοποιηθεί στην απλή δήλωση του Θεωρήματος ***8.1.4***. Αυτό δείχνει την δύναμη του συναρτησιακού λογισμού.

Μια άλλη αναπαράσταση για τις λύσεις της σχέσης **(8.1.1)** προκύπτει, αν βάλουμε

$$\boldsymbol{f(\lambda) = e^{\lambda}},$$

στο Θεώρημα ***4.1.1*** και χρησιμοποιήσουμε την φασματική ανάλυση του $\boldsymbol{e^{At}}$. Συνεπώς, συμβολίζοντας $\boldsymbol{x(0) = x_0}$, χρησιμοποιούμε το Θεώρημα ***8.1.3***, για να πάρουμε

$$\boldsymbol{x(t) = \sum_{k=1}^{s} \sum_{j=0}^{m_k-1} Z_{kj} t^j e^{\lambda_k t} x_0 = \sum_{k=1}^{s} (z_{k0} + z_{k1} t + \cdots + z_{k,m_k-1} t^{m_k-1}) e^{\lambda_k t}},$$

όπου $\boldsymbol{z_{kj} = Z_{kj} x_0}$ και είναι ανεξάρτητο του $\boldsymbol{t}$. Συνεπώς, παρατηρούμε ότι πάλι η λύση είναι ένας συνδυασμός πολυωνύμων του $\boldsymbol{t}$ με εκθετικές συναρτήσεις. Αν $\boldsymbol{A}$ είναι ένας απλός πίνακας, η λύση μειώνεται σε

$$\boldsymbol{x(t) = \sum_{k=1}^{s} z_{k0} e^{\lambda_k t}},$$

και βλέπουμε ότι κάθε στοιχείο του $\boldsymbol{x(t)}$ είναι ένας γραμμικός συνδυασμός από εκθετικές συναρτήσεις χωρίς την παρουσία πολυωνύμων.

Τέλος, θεωρούμε την ομογενή εξίσωση **(8.1.2)**, στην οποία τα στοιχεία του $\boldsymbol{f}$ είναι καθορισμένες συνεχείς κατά σημείο συναρτήσεις $\boldsymbol{f_1(t), \ldots, f_n(t)}$ του $\boldsymbol{t}$, οι οποίες δεν είναι όλες ταυτοτικά μηδέν. Έστω $\boldsymbol{\mathcal{F}}$ είναι ο χώρος όλων των λύσεων που σχετίζονται με το ομογενές πρόβλημα

$$\boldsymbol{\dot{x} = Ax}$$

και έστω $\boldsymbol{\mathcal{J}}$ είναι το σύνολο όλων των λύσεων του

$$\boldsymbol{\dot{x} = Ax + f}.$$

Αν μας δίνεται μόνο μια λύση $\boldsymbol{y} \in \boldsymbol{\mathcal{J}}$, τότε $\boldsymbol{z} \in \boldsymbol{\mathcal{J}}$, αν και μόνο αν

$$\boldsymbol{z = y + w},$$

για κάποιο $\boldsymbol{w} \in \boldsymbol{\mathcal{F}}$. Συνεπώς, για να πάρουμε τον χώρο $\boldsymbol{\mathcal{J}}$ χρειάζεται να γνωρίζουμε μια λύση της μη ομογενής εξίσωσης, γνωστή ως ένα ***συγκεκριμένο ολοκλήρωμα*** και έπειτα να το προσθέσουμε σε όλες τις λύσεις της ομογενής εξίσωσης και να εκφράσουμε την κάθε μια στην μορφή $\boldsymbol{e^{At}x}$, για κάποιο $\boldsymbol{x} \in \mathbb{C}^{\boldsymbol{n}}$.

Ψάχνουμε για μια λύση του $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}$ στη μορφή

$$\boldsymbol{x}(\boldsymbol{t}) = \boldsymbol{e}^{\boldsymbol{A}\boldsymbol{t}}\boldsymbol{z}(\boldsymbol{t}).$$

Αν υπάρχει μια τέτοια λύση, τότε

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e}^{\boldsymbol{A}\boldsymbol{t}}\dot{\boldsymbol{z}} \quad \text{και έτσι} \quad \boldsymbol{e}^{\boldsymbol{A}\boldsymbol{t}}\dot{\boldsymbol{z}} = \boldsymbol{f},$$

συνεπώς

$$\boldsymbol{z}(\boldsymbol{t}) = \boldsymbol{e}^{-\boldsymbol{A}\boldsymbol{t}_0}\boldsymbol{x}_0 + \int_{\boldsymbol{t}_o}^{\boldsymbol{t}} \boldsymbol{e}^{-\boldsymbol{A}\boldsymbol{\tau}}\boldsymbol{f}(\boldsymbol{\tau})\boldsymbol{d}\boldsymbol{\tau},$$

όπου $\boldsymbol{x}_0 = \boldsymbol{e}^{\boldsymbol{A}\boldsymbol{t}_0}\boldsymbol{z}(\boldsymbol{t}_0) = \boldsymbol{x}(\boldsymbol{t}_0)$. Αυτό υποδεικνύει, ότι

$$\boldsymbol{x}(\boldsymbol{t}) = \boldsymbol{e}^{\boldsymbol{A}\boldsymbol{t}}\left(\boldsymbol{e}^{-\boldsymbol{A}\boldsymbol{t}_0}\boldsymbol{x}_0 + \int_{\boldsymbol{t}_0}^{\boldsymbol{t}} \boldsymbol{e}^{-\boldsymbol{A}\boldsymbol{\tau}}\boldsymbol{f}(\boldsymbol{\tau})\boldsymbol{d}\boldsymbol{\tau}\right)$$

$$= \boldsymbol{e}^{\boldsymbol{A}(\boldsymbol{t}-\boldsymbol{t}_0)}\boldsymbol{x}_0 + \int_{\boldsymbol{t}_0}^{\boldsymbol{t}} \boldsymbol{e}^{\boldsymbol{A}(\boldsymbol{t}-\boldsymbol{\tau})}\boldsymbol{f}(\boldsymbol{\tau})\boldsymbol{d}\boldsymbol{\tau},$$

είναι η λύση του $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{f}$, όπου $\boldsymbol{x}(\boldsymbol{t}_0) = \boldsymbol{x}_0$.

**Θεώρημα 8.1.5.** *Αν $\boldsymbol{f}(\boldsymbol{t})$ είναι μια κατά σημείο συνεχής συνάρτηση του $\boldsymbol{t}$ με τιμές στον χώρο $\mathbb{C}^{\boldsymbol{n}}$, $\boldsymbol{t}_0 \in \mathbb{R}$ και $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n}\times\boldsymbol{n}}$, τότε κάθε λύση της διαφορικής εξίσωσης*

$$\dot{\boldsymbol{x}}(\boldsymbol{t}) = \boldsymbol{A}\boldsymbol{x}(\boldsymbol{t}) + \boldsymbol{f}(\boldsymbol{t}),$$

*έχει τη μορφή*

$$\boldsymbol{x}(\boldsymbol{t}) = \boldsymbol{e}^{\boldsymbol{A}(\boldsymbol{t}-\boldsymbol{t}_0)}\boldsymbol{x}_0 + \int_{\boldsymbol{t}_0}^{\boldsymbol{t}} \boldsymbol{e}^{\boldsymbol{A}(\boldsymbol{t}-\boldsymbol{\tau})}\boldsymbol{f}(\boldsymbol{\tau})\boldsymbol{d}\boldsymbol{\tau} \qquad \textbf{(8.1.10)}$$

*και*

$$\boldsymbol{x}_0 = \boldsymbol{x}(\boldsymbol{t}_0).$$

**Παράδειγμα 8.1.3.**

Θα λύσουμε το σύστημα $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{f}$, όπου $\boldsymbol{x}(0) = \boldsymbol{x}_0$, αν $\boldsymbol{A}$ και $\boldsymbol{x}_0$ ορίζονται όπως στα Παραδείγματα ***8.1.1*** και ***8.1.2*** και

$$\boldsymbol{f} = \begin{bmatrix} \alpha e^{4t} \\ 0 \\ 0 \end{bmatrix} = \alpha e^{4t} \boldsymbol{e}_1.$$

Το διάνυσμα $\boldsymbol{e}^{\boldsymbol{A}t}\boldsymbol{x}_0$ είναι γνωστό από το Παράδειγμα ***8.1.1***. Χρησιμοποιώντας την εξίσωση **(4.1.1)**, για να βρούμε το συγκεκριμένο ολοκλήρωμα, έχουμε

$$\boldsymbol{y}(t) = \int_0^t e^{\boldsymbol{A}(t-\tau)} \boldsymbol{f}(\tau) d\tau = \sum_{k=1}^{s} \sum_{j=0}^{m_k - 1} \boldsymbol{Z}_{kj} \int_0^t (t-\tau)^{j-1} e^{\lambda_k (t-\tau)} \boldsymbol{f}(\tau) d\tau.$$

Για τα ολοκληρώματα ισχύει

$$\alpha \int_0^t e^{\boldsymbol{A}(t-\tau)} e^{4\tau} = \alpha t e^{4t},$$

$$\alpha \int_0^t (t-\tau) e^{4(t-\tau)} e^{4\tau} = \frac{1}{2} \alpha t^2 e^{4t},$$

$$\alpha \int_0^t e^{2(t-\tau)} e^{4\tau} d\tau = \frac{1}{2} \alpha e^{2t} (e^{2t} - 1).$$

Συνεπώς,

$$\boldsymbol{y}(t) = \alpha t e^{4t} \boldsymbol{Z}_{10} \boldsymbol{e}_1 + \frac{1}{2} \alpha t^2 e^{4t} \boldsymbol{Z}_{11} \boldsymbol{e}_1 + \frac{1}{2} \alpha e^{2t} (e^{2t} - 1) \boldsymbol{Z}_{20} \boldsymbol{e}_1$$

$$= \frac{1}{2} \alpha t e^{4t} \left( 2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} \alpha(1+t) t e^{4t} \\ -\alpha t^2 e^{4t} \\ 0 \end{bmatrix}.$$

Η λύση δίνεται από το $\boldsymbol{x}(t) = \boldsymbol{e}^{\boldsymbol{A}t} \boldsymbol{x}_0 + \boldsymbol{y}(t)$. $\square$

# Κεφάλαιο 9

## 9.1 Παρατηρήσιμα και Ελέγξιμα Συστήματα

Σε αυτό το κεφάλαιο συγκεντρώνουμε διάφορα σημαντικά αποτελέσματα που αναπτύξαμε έως τώρα. Αυτό θα μας επιτρέψει να παρουσιάσουμε μερικές ενδιαφέρουσες ιδέες και αποτελέσματα από την θεωρία συστημάτων και ελέγχου.

Έστω $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ και $\boldsymbol{C} \in \mathbb{C}^{r\times n}$ και θεωρούμε το σύστημα εξισώσεων

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0,$$

$$\boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t). \tag{9.1.1}$$

Συνεπώς $\boldsymbol{x}(t)$ έχει τιμές στο $\mathbb{C}^n$ που ορίζονται από το πρόβλημα αρχικών τιμών (υποθέτοντας ότι $\boldsymbol{x}_0 \in \mathbb{C}^n$ δίνεται) και καλείται ***διάνυσμα κατάστασης***. Τότε $\boldsymbol{y}(t) \in \mathbb{C}^r$, για κάθε $t \geq 0$, είναι το ***παρατηρήσιμο διάνυσμα*** ή ***διάνυσμα απόδοσης***. Παριστάνει τις πληροφορίες για την κατάσταση του συστήματος που μπορεί να καταγραφεί, αφού φιλτραριστεί μέσω του πίνακα $\boldsymbol{C}$. Το σύστημα **(9.1.1)** λέγεται ότι είναι ***παρατηρήσιμο***, αν

$$\boldsymbol{y}(t) \equiv \boldsymbol{0} \quad \text{και προκύπτει μόνο όταν} \quad \boldsymbol{x}_0 = \boldsymbol{0}.$$

Συνεπώς, για να είναι το σύστημα παρατηρήσιμο, ***όλα*** τα μη μηδενικά αρχικά διανύσματα πρέπει να δίνουν λύσεις που μπορούν να παρατηρηθούν, με την έννοια ότι το διάνυσμα απόδοσης δεν είναι ταυτοτικά μηδέν.

Η σύνδεση με τα παρατηρήσιμα ζευγάρια, πραγματοποιείται με το πρώτο θεώρημα.

Για διευκόλυνση, γράφουμε

$$\boldsymbol{Q} = \boldsymbol{Q}(\boldsymbol{C}, \boldsymbol{A}) = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{C}\boldsymbol{A} \\ \vdots \\ \boldsymbol{C}\boldsymbol{A}^{n-1} \end{bmatrix},$$

ένας πίνακας μεγέθους $nr \times n$.

**Θεώρημα 9.1.1.** *Το σύστημα* $(\mathbf{9.1.1})$ *είναι παρατηρήσιμο, αν και μόνο αν*

$$\operatorname{Ker} Q(C, A) = \{0\},$$

*το οποίο ισχύει, αν και μόνο αν το ζεύγος* $(C, A)$ *είναι παρατηρήσιμο.*

**Λήμμα 9.1.1.** *Για κάθε πίνακα* $A \in \mathbb{C}^{n\times n}$ *υπάρχουν ολόμορφες συναρτήσεις* $\psi_1(t), \ldots, \psi_n(t)$, *τέτοιες ώστε*

$$e^{At} = \sum_{j=1}^{n} \psi_j(t) A^{j-1}. \tag{9.1.2}$$

**Απόδειξη.** Έστω

$$f(\lambda) = e^{\lambda t},$$

τότε $f$ είναι μια ολόμορφη συνάρτηση του $\lambda$ που ορίζεται στο φάσμα οποιουδήποτε $n \times n$ πίνακα $A$. Αφού

$$f^{(j)}(\lambda) = t^j e^{\lambda t},$$

προκύπτει από το Θεώρημα ***4.1.1***, ότι

$$f(A) = e^{At} = \sum_{k=1}^{s} \sum_{j=0}^{m_k-1} t^j e^{\lambda_k t} Z_{kj} \ . \tag{9.1.3}$$

Αφού

$$Z_{kj} = \varphi_{kj}(A),$$

για κάθε $k$, $j$ και $\varphi_{kj}$ είναι ένα πολυώνυμο ο βαθμός, του οποίου δεν ξεπερνά το $n$ και επειδή τα $\varphi_{kj}$ εξαρτώνται μόνο από τον πίνακα $A$, ο δεξιός όρος της εξίσωσης $(\mathbf{9.1.3})$ μπορεί να αναδιαταχθεί όπως στην εξίσωση $(\mathbf{9.1.2})$. Επιπρόσθετα, κάθε συντελεστής $\psi_j(t)$ είναι απλά ένας γραμμικός συνδυασμός από συναρτήσεις της μορφής $t^j e^{\lambda_k t}$ και είναι συνεπώς ολόμορφος. ■

**Απόδειξη Θεωρήματος 9.1.1.** Αν $\mathbf{Ker}\boldsymbol{Q(C,A)} \neq \{\mathbf{0}\}$, τότε υπάρχει ένα μη μηδενικό $\boldsymbol{x_0}$, τέτοιο ώστε

$$\boldsymbol{CA^j x_0 = 0}, \quad \text{για} \quad \boldsymbol{j = 0, 1, \ldots, n-1}.$$

Συνεπάγεται από το Λήμμα ***9.1.1***, ότι

$$\boldsymbol{Ce^{At}x_0 = 0},$$

για κάθε πραγματικό $\boldsymbol{t}$. Αλλά, τότε το Θεώρημα ***8.1.3*** και η εξίσωση **(9.1.1)** δίνουν

$$\boldsymbol{y(t) = Ce^{At}x_0 \equiv 0}$$

και το σύστημα δεν είναι παρατηρήσιμο. Συνεπώς, αν το σύστημα είναι παρατηρήσιμο, το ζεύγος $\boldsymbol{(C,A)}$ πρέπει να είναι παρατηρήσιμο.

Αντίστροφα, αν το σύστημα δεν είναι παρατηρήσιμο, υπάρχει ένα $\boldsymbol{x_0 \neq 0}$, τέτοιο ώστε

$$\boldsymbol{y(t) = Ce^{At}x_0 \equiv 0}.$$

Διαφορίζοντας επανειλημμένα ως προς $\boldsymbol{t}$, προκύπτει ότι

$$\boldsymbol{Q(C,A)e^{At}x_0 \equiv 0} \quad \text{και} \quad \boldsymbol{Q(C,A) \neq \{0\}}.$$

Επομένως, η παρατηρησιμότητα ενός ζεύγους $\boldsymbol{(C,A)}$ πρέπει να υποδεικνύει την παρατηρησιμότητα του συστήματος. ■

Τώρα θα αναπτύξουμε τον όρο ελεγξιμότητα της θεωρίας συστημάτων. Έστω $\boldsymbol{A} \in \mathbb{C}^{\boldsymbol{n\times n}}$ και $\boldsymbol{B} \in \mathbb{C}^{\boldsymbol{n\times m}}$ και θεωρούμε μια διαφορική εξίσωση με αρχικές συνθήκες

$$\boldsymbol{\dot{x}(t) = Ax(t) + Bu(t)}, \quad \boldsymbol{x(0) = x_0}. \qquad \mathbf{(9.1.4)}$$

Για να είμαστε συνεπείς, απαιτείται $\boldsymbol{x(t)} \in \mathbb{C}^{\boldsymbol{n}}$ και $\boldsymbol{u(t)} \in \mathbb{C}^{\boldsymbol{m}}$, για κάθε $\boldsymbol{t}$. Θα βλέπουμε το $\boldsymbol{u(t)}$, ως μια διανυσματική συνάρτηση, η οποία καλείται ***έλεγχος*** και μπορεί να χρησιμοποιηθεί, για να οδηγήσει ή να ελέγξει, τη λύση $\boldsymbol{x(t)}$ του προβλήματος αρχικών τιμών, ώστε να πληρεί κάποιο σκοπό.

Ο σκοπός που θεωρούμε είναι αυτός του ελέγχου της λύσης $\boldsymbol{x(t)}$ κατά τέτοιο τρόπο, ώστε μετά από κάποιο χρόνο $\boldsymbol{t}$, η συνάρτηση $\boldsymbol{x(t)}$ να παίρνει κάποιες προκαθορισμένες διανυσματικές τιμές στον χώρο $\mathbb{C}^{\boldsymbol{n}}$.

Για να κάνουμε αυτές τις ιδέες περισσότερο ακριβείς, πρώτα ορίζουμε μια κλάση συναρτήσεων αποδοχής ελέγχου $\boldsymbol{u(t)}$. Ορίζουμε $\boldsymbol{\mathcal{U}_m}$ να είναι ο γραμμικός χώρος των συναρτήσεων που παίρνουν τιμές στον χώρο $\mathbb{C}^{\boldsymbol{m}}$, που ορίζονται και είναι κατά σημείο συνεχείς στην περιοχή $\boldsymbol{t \geq 0}$ και θεωρούμε ότι $\boldsymbol{u(t)} \in \boldsymbol{\mathcal{U}_m}$.

Τότε παρατηρούμε, ότι χρησιμοποιώντας την εξίσωση $\mathbf{(8.1.10)}$ και διαλέγοντας $\boldsymbol{u(t)} \in \boldsymbol{\mathcal{U}_m}$, η λύση του προβλήματος αρχικών τιμών μπορεί να γραφεί αναλυτικά στην μορφή

$$\boldsymbol{x(t; x_0, u) = e^{At}x_0 + \int_0^t e^{(t-s)A}Bu(s)ds}. \qquad \mathbf{(9.1.5)}$$

Ο σκοπός του πολύπλοκου συμβολισμού στα αριστερά είναι να δώσει έμφαση στην εξάρτηση της λύσης $\boldsymbol{x}$ από τη συνάρτηση ελέγχου $\boldsymbol{u}$ και την αρχική συνθήκη $\boldsymbol{x_0}$, καθώς και από την ανεξάρτητη μεταβλητή $\boldsymbol{t}$.

Ένα διάνυσμα $\boldsymbol{y} \in \mathbb{C}^{\boldsymbol{n}}$ λέγεται, ότι είναι ***προσβάσιμο από το αρχικό διάνυσμα*** $\boldsymbol{x_0}$, αν υπάρχει ένα $\boldsymbol{u(t)} \in \mathcal{U}_{\boldsymbol{m}}$, τέτοιο ώστε

$$\boldsymbol{x(t; x_0, u) = y},$$

για κάποια θετικά $\boldsymbol{t}$. Κάτω από την ίδια συνθήκη μπορούμε να πούμε, ότι το $\boldsymbol{y}$ είναι ***προσβάσιμο από το*** $\boldsymbol{x_0}$ ***σε χρόνο*** $\boldsymbol{t}$.

Το ***σύστημα*** της εξίσωσης $\mathbf{(9.1.4)}$ λέγεται ότι είναι ***ελέγξιμο***, αν κάθε διάνυσμα στον χώρο $\mathbb{C}^{\boldsymbol{n}}$ είναι προσβάσιμο από κάθε άλλο διάνυσμα στον χώρο $\mathbb{C}^{\boldsymbol{n}}$. Ο επόμενός μας στόχος είναι είναι να δείξουμε ότι το σύστημα της σχέσης $\mathbf{(9.1.4)}$ είναι ελέγξιμο, αν και μόνο αν το ζεύγος $\boldsymbol{(A, B)}$ είναι ελέγξιμο και αυτό θα προκύψει εύκολα από το επόμενο αποτέλεσμα.

Ανακαλούμε πρώτα ότι ο ελέγξιμος υπόχωρος του $\boldsymbol{(A, B)}$, τον οποίο θα συμβολίζουμε με $\boldsymbol{\mathcal{L}}$, δίνεται από

$$\boldsymbol{\mathcal{L}_{A,B} = \sum_{r=0}^{n-1} \mathrm{\mathbf{Im}}(A^r B)}. \qquad \mathbf{(9.1.6)}$$

**Θεώρημα 9.1.2.** *Το διάνυσμα* $\boldsymbol{z} \in \mathbb{C}^{\boldsymbol{n}}$ *είναι προσβάσιμο από* $\boldsymbol{y} \in \mathbb{C}^{\boldsymbol{n}}$ *σε χρόνο* $\boldsymbol{t}$*, αν και μόνο αν*

$$\boldsymbol{z} - \boldsymbol{e}^{\boldsymbol{At}}\boldsymbol{y} \in \boldsymbol{\mathcal{L}_{A,B}}.$$

Η απόδειξη θα αναλυθεί με τη βοήθεια ενός άλλου λήμματος, το οποίο προσφέρει μια χρήσιμη περιγραφή του ελέγξιμου υποχώρου του $(\boldsymbol{A}, \boldsymbol{B})$.

**Λήμμα 9.1.2.** *Η συνάρτηση που έχει για τιμές πίνακες* $\boldsymbol{W(t)}$ *που ορίζεται, ως*

$$\boldsymbol{W(t)} = \int_{\boldsymbol{0}}^{\boldsymbol{t}} \boldsymbol{e^{As}BB^{*}e^{A^{*}s}ds}, \tag{9.1.7}$$

έχει την ιδιότητα, ότι για οποιοδήποτε $\boldsymbol{t > 0}$,

$$\mathbf{Im}\,\boldsymbol{W(t)} = \boldsymbol{\mathcal{L}_{A,B}}.$$

**Απόδειξη.** Είναι ξεκάθαρο ότι, $\boldsymbol{W(t)}$ είναι ερμιτιανός, για κάθε $\boldsymbol{t > 0}$ και συνεπώς

$$\mathbf{Im}\,\boldsymbol{W(t)} = (\mathbf{Ker}\,\boldsymbol{W(t)})^{\perp}.$$

Επίσης, αν $\boldsymbol{\mathcal{K}_{A,B}}$ συμβολίζει τον μη παρατηρήσιμο υπόχωρο του $(\boldsymbol{B^{*}, A^{*}})$, τότε

$$\boldsymbol{\mathcal{L}_{A,B}} = (\boldsymbol{\mathcal{K}_{A,B}})^{\perp}.$$

Συνεπώς, το συμπέρασμα του λήμματος είναι ισοδύναμο με τη δήλωση, ότι

$$\mathbf{Ker}\,\boldsymbol{W(t)} = \boldsymbol{\mathcal{K}_{A,B}}$$

και θα το αποδείξουμε σε αυτή τη μορφή.

Αν $\boldsymbol{x} \in \mathbf{Ker}\,\boldsymbol{W(t)}$, τότε

$$\boldsymbol{x^{*}W(t)x} = \int_{\boldsymbol{0}}^{\boldsymbol{t}} \|\boldsymbol{B^{*}e^{A^{*}s}x}\|^{\boldsymbol{2}}\boldsymbol{ds} = \boldsymbol{0},$$

όπου εδώ χρησιμοποιείται η ευκλείδεια διανυσματική νόρμα. Συνεπώς, για όλα τα $\boldsymbol{s} \in [\boldsymbol{0}, \boldsymbol{t}]$,

$$\boldsymbol{B^{*}e^{A^{*}s}x} = \boldsymbol{0}.$$

Διαφορίζουμε αυτή την σχέση, ως προς $s$ επανειλημμένα και παίρνουμε το όριο, καθώς $s \to 0_+$, για να πάρουμε

$$B^*(A^*)^{j-1}x = 0, \quad \text{για} \quad j = 1, 2, \ldots, n.$$

Με άλλα λόγια,

$$x \in \bigcap_{j=1}^{n} \mathrm{Ker}\,(B^*(A^*)^{j-1}) = \mathcal{K}_{A,B}.$$

Συνεπώς,

$$\mathrm{Ker}\, W(t) \subset \mathcal{K}_{A,B}.$$

Είναι εύκολο να επαλήθευσουμε το, ότι αντιστρέφοντας το επιχείρημα και χρησιμοποιώντας το Λήμμα ***9.1.1***, το αντίθετο συμπέρασμα προκύπτει. Έτσι

$$\mathrm{Ker}\, W(t) = \mathcal{K}_{A,B}$$

και το λήμμα αποδείχθηκε. ■

Αυτό το λήμμα δημιουργεί ένα ακόμη κριτήριο για την ελεγξιμότητα του ζεύγους $(A, B)$.

**Συμπέρασμα 9.1.1.** *Το ζεύγος $(A, B)$ είναι ελέγξιμο, αν και μόνο αν ο πίνακας $W(t)$ της εξίσωσης $(9.1.7)$ είναι θετικά ορισμένος, για όλα τα $t > 0$.*

**Απόδειξη.** Παρατηρώντας τον ορισμό $(9.1.7)$ φαίνεται ότι ο $W(t)$ είναι τουλάχιστον θετικά ημιορισμένος. Όμως, για $(A, B)$ να είναι ελέγξιμο, τότε

$$\mathcal{L}_{A,B} = \mathbb{C}^n,$$

και έτσι το Λήμμα $(9.1.2)$ υποδεικνύει, ότι

$$\mathrm{Im}\, W(t) = \mathbb{C}^n.$$

Συνεπώς, $W(t)$ έχει βαθμό $n$ και είναι επομένως θετικά ορισμένος. ■

**Απόδειξη Θεωρήματος 9.1.2.** Αν $\boldsymbol{z}$ είναι προσβάσιμο από $\boldsymbol{y}$ σε χρόνο $\boldsymbol{t}$, τότε χρησιμοποιώντας την εξίσωση **(9.1.5)**,

$$\boldsymbol{z - e^{At}y = \int_0^t e^{(t-s)A} Bu(s)ds},$$

για κάποια $\boldsymbol{u} \in \boldsymbol{\mathcal{U}_m}$. Χρησιμοποιώντας το Λήμμα ***9.1.1*** γράφουμε

$$\boldsymbol{z - e^{At}y = \sum_{j=1}^{n} A^{j-1}B \int_0^t \psi_j(t-s)u(s)ds},$$

και χρησιμοποιώντας τον ορισμό **(9.1.6)** φαίνεται άμεσα, ότι

$$\boldsymbol{z - e^{At}y \in \mathcal{L}_{A,B}}.$$

Για το αντίστροφο, έστω

$$\boldsymbol{z - e^{At}y \in \mathcal{L}_{A,B}}.$$

Θα κατασκευάσουμε μια είσοδο ή έλεγχο, $\boldsymbol{u(t)}$ που θα οδηγεί το σύστημα από το $\boldsymbol{y}$ στο $\boldsymbol{z}$. Χρησιμοποιώντας το Λήμμα ***9.1.2*** γράφουμε

$$\boldsymbol{z - e^{At}y = W(t)w = \int_0^t e^{As} BB^* e^{A^* s} w ds},$$

για κάποια $\boldsymbol{w} \in \mathbb{C}^{\boldsymbol{n}}$. Βάζοντας $\boldsymbol{s = t - \tau}$, παίρνουμε

$$\boldsymbol{z - e^{At}y = \int_0^t e^{A(t-\tau)} BB^* e^{A^*(t-\tau)} w d\tau = \int_0^t e^{A(t-\tau)} Bu(\tau) d\tau},$$

όπου ορίζουμε

$$\boldsymbol{u(\tau) = B^* e^{A^*(t-\tau)} w \in \mathcal{U}_m}.$$

Συγκρίνοντας το αποτέλεσμα που πήραμε με την εξίσωση **(9.1.5)**, βλέπουμε ότι το $\boldsymbol{z}$ είναι προσβάσιμο από το $\boldsymbol{y}$ σε χρόνο $\boldsymbol{t}$. ■

**Συμπέρασμα 9.1.2.** *Το ζεύγος* $(\boldsymbol{A}, \boldsymbol{B})$ *είναι ελέγξιμο, αν και μόνο αν το σύστημα*

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u}, \ \boldsymbol{x}(0) = \boldsymbol{x_0},$$

*είναι ελέγξιμο.*

**Απόδειξη.** Αν το ζεύγος $(\boldsymbol{A}, \boldsymbol{B})$ είναι ελέγξιμο, τότε

$$\boldsymbol{\mathcal{L}_{A,B}} = \mathbb{C}^{\boldsymbol{n}}.$$

Συνεπώς, για οποιαδήποτε $\boldsymbol{y}, \boldsymbol{z} \in \mathbb{C}^{\boldsymbol{n}}$, προκύπτει τετριμμένα ότι

$$\boldsymbol{z} - \boldsymbol{e^{At}y} \in \boldsymbol{\mathcal{L}_{A,B}}.$$

Τότε, το Θεώρημα ***9.1.2*** μας διαβεβαιώνει ότι το σύστημα είναι ελέγξιμο.

Αντίστροφα, αν το σύστημα είναι ελέγξιμο, τότε συγκεκριμένα, κάθε διάνυσμα $\boldsymbol{z} \in \boldsymbol{\mathcal{L}^n}$ είναι προσβάσιμο από το μηδενικό διάνυσμα και από το Θεώρημα ***9.1.2***,

$$\boldsymbol{\mathcal{L}_{A,B}} = \mathbb{C}^{\boldsymbol{n}}.$$

Συνεπώς το ζεύγος $(\boldsymbol{A}, \boldsymbol{B})$ είναι ελέγξιμο. ■

Όταν το σύστημα είναι ελέγξιμο, προκύπτει ότι οποιοδήποτε $\boldsymbol{z} \in \mathbb{C}^{\boldsymbol{n}}$ είναι προσβάσιμο από οποιοδήποτε $\boldsymbol{y} \in \mathbb{C}^{\boldsymbol{n}}$ σε οποιοδήποτε χρόνο $\boldsymbol{t} > \boldsymbol{0}$. Για να το δούμε αυτό, κατασκευάζουμε μια σαφή συνάρτηση ελέγχου $\boldsymbol{u(s)}$ που θα παίρνει την λύση από το $\boldsymbol{y}$ στο $\boldsymbol{z}$ σε οποιοδήποτε προκαθορισμένο χρόνο $\boldsymbol{t}$. Συνεπώς, στην ελέγξιμη περίπτωση, ο $\boldsymbol{W(t)}$ του Λήμματος ***9.1.2*** είναι θετικά ορισμένος και μπορούμε να ορίσουμε μια συνάρτηση

$$\boldsymbol{u(s)} = \boldsymbol{B^* e^{A^*(t-s)} W(t)^{-1} (z - e^{At} y)}. \tag{9.1.8}$$

Τότε η εξίσωση **(9.1.5)** δίνει

$$\boldsymbol{x(t; y, u)} = \boldsymbol{e^{At} y} + \left( \int_{\boldsymbol{0}}^{\boldsymbol{t}} \boldsymbol{e^{A(t-s)} B B^* e^{A^*(t-s)} ds} \right) \boldsymbol{W(t)^{-1} (z - e^{At} y)}.$$

Χρησιμοποιώντας τον ορισμό **(9.1.7)** του $\boldsymbol{W(t)}$, αυτή η εξίσωση μειώνεται σε

$$\boldsymbol{x(t; y, u)} = \boldsymbol{e^{At} y} + \boldsymbol{z} - \boldsymbol{e^{At} y} = \boldsymbol{z},$$

όπως απαιτείται.

# Βιβλιογραφία


Lancaster, P., and Tismenetsky, M. *The Theory of Matrices* New York: Academic Press, 1969.

Bellman, R. *Introduction to Matrix Analysis.* New York: McGraw-Hill, 1960.

Dunford, N., and Schwartz, J. T. *Linear Operators, Part I: General Theory.* New York: Wiley, 1958.

Gantmacher, F. R. *The Theory of Matrices,* vols. 1 and 2. New York: Chelsea, 1959.

Glazman, I. M., and Liubich, J. I. *Finite-Dimensional Linear Analysis.* Cambridge, Mass.: M.I.T. Press, 1974.

Marcus, M., and Minc, H. *A Survey of Matrix Theory and Matrix Inequalities.* Boston: Allyn and Bacon, 1964.

Russell, D. L. *Mathematics of Finite-Dimensional Control Systems.* New York: Marcel Dekker, 1979.

Wonham, W. M.*Linear Multivariable Control: A Geometric Approach.* Berlin: Springer-Verlag, 1979.