Γραμμικη Αλγεβρα II

Θεωρητικα Θεματα

Ακαδημαϊκο Ετος

2011 - 2012

Διδασκοντες: Ν. Μαρμαρίδης - Α. Μπεληγιάννης

Βοηθος Ασκησεων: Χ. Ψαρουδάκης

Ιστοσελιδα Μαθηματος:
http://www.math.uoi.gr/∼abeligia/LinearAlgebraII/LAII.html

30 Ιουνιου 2012

"If you can reduce a mathematical problem to a problem in Linear Algebra, you can most likely solve it, provided that you know enough Linear Algebra."

Peter Lax
Βραβείο Abel 2005

Μέρος 1. Η Δομή Ενός Ενδομορφισμού

1. Χαρακτηριστικό Πολυώνυμο Γινομένου Πινάκων

1.1. **Ιδιοτιμές Σύνθεσης Γραμμικών Απεικονίσεων και Γινομένου Πινάκων.** Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης υπεράνω ενός σώματος $\mathbb{K}$ και

$$f,\, g:\ \mathcal{E} \longrightarrow \mathcal{E}$$

δύο γραμμικές απεικονίσεις. Θεωρούμε τις γραμμικές απεικονίσεις

$$f \circ g,\, g \circ f:\ \mathcal{E} \longrightarrow \mathcal{E}$$

Πρόταση 1.1. *Οι γραμμικές απεικονίσεις $f \circ g$ και $g \circ f$ έχουν τις ίδιες ιδιοτιμές.*

Απόδειξη. **1.** Δείχνουμε ότι κάθε ιδιοτιμή της $f \circ g$ είναι και ιδιοτιμή της $g \circ f$.

Έστω λ μια ιδιοτιμή της $f \circ g$ με αντίστοιχο ιδιοδιάνυμσα $\vec{e}$, δηλαδή:

$$(f \circ g)(\vec{e}) = \lambda \vec{e}, \qquad \vec{e} \neq \vec{0}$$

(α΄) Αν $\lambda = 0$, τότε σύμφωνα με όσα γνωρίζουμε από την θεωρία, η γραμμική απεικόνιση $f \circ g$ δεν είναι ισομορφισμός. Τότε όμως και η γραμμική απεικόνιση $g \circ f$ δεν είναι ισομορφισμός. Πραγματικά: αν η $g \circ f$ είναι ισομορφισμός, τότε η f θα είναι μονομορφισμός διότι αν $f(\vec{x}) = 0$, τότε $g(f(\vec{x})) = \vec{0} \implies \vec{x} = 0$ διότι από την υπόθεση η $g \circ f$ είναι ισομορφισμός. Γνωρίζουμε όμως ότι ένας μονομορφισμός $f \colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι πάντα ισομορφισμός. Άρα η f είναι ισομορφισμός και παρόμοια δείχνουμε η g είναι ισομορφισμός. Επειδή η σύνθεση ισομορφσιμών είναι ισομορφισμός έπεται ότι και η $f \circ g$ είναι ισομορφισμός το οποίο είναι άτοπο. Καταλήγουμε ότι η γραμμική απεικόνιση $g \circ f$ δεν είναι ισομορφισμός. Αυτό είναι ισοδύναμο με το ότι η γραμμική απεικόνιση $g \circ f$ έχει το $\lambda = 0$ ως ιδιοτιμή.
Επομένως θα έχουμε ότι: η γραμμική απεικόνιση $f \circ g$ έχει το $\lambda = 0$ ως ιδιοτιμή $\Longrightarrow$ η γραμμική απεικόνιση $g \circ f$ έχει το $\lambda = 0$ ως ιδιοτιμή.

(β΄) Υποθέτουμε ότι $\lambda \neq 0$. Θεωρούμε το διάνυσμα $\vec{\varepsilon} := g(\vec{e})$. Τότε $\vec{\varepsilon} \neq 0$ διότι διαφορετικά αν $\vec{\varepsilon} = g(\vec{e}) = \vec{0}$, τότε $f(g(\vec{e})) = \vec{0}$. Επειδή $(f \circ g)(\vec{e}) = \lambda \vec{e}$ θα έχουμε $\lambda \vec{e} = \vec{0}$. Όμως $\vec{e} \neq \vec{0}$ και επομένως $\lambda = 0$ το οποίο είναι άτοπο. Άρα θα έχουμε $\vec{\varepsilon} = g(\vec{e}) \neq 0$.
Τότε $f(\vec{\varepsilon}) = f(g(\vec{e})) = (f \circ g)(\vec{e}) = \lambda \vec{e} \neq \vec{0}$. Επιπρόσθετα:

$$(g \circ f)(\vec{\varepsilon}) = g(f(\vec{\varepsilon}) = g(\lambda \vec{e}) = \lambda g(\vec{e}) = \lambda \vec{\varepsilon}, \qquad \vec{\varepsilon} \neq \vec{0}$$

Η παραπάνω σχέση δείχνει ότι το λ είναι ιδιοτιμή της $g \circ f$ με αντίστοιχο ιδιοδιάνυσμα $\vec{\varepsilon}$.

2. Παρόμοια δείχνουμε ότι κάθε ιδιοτιμή της $g \circ f$ είναι και ιδιοτιμή της $f \circ g$. □

Πόρισμα 1.2. *Αν $A, B \in \mathsf{M}_{n \times n}(\mathbb{K})$, τότε οι πίνακες AB και BA έχουν τις ίδιες ιδιοτιμές.*

Απόδειξη. Θεωρούμε τις γραμμικές απεικονίσεις

$$f_A:\ \mathbb{K}_n \longrightarrow \mathbb{K}_n, \quad f_A(X) = AX$$

$$f_B:\ \mathbb{K}_n \longrightarrow \mathbb{K}_n, \quad f_B(X) = BX$$

Τότε $f_{AB}(X) = (AB)X = A(BX) = Af_B(X) = f_A(f_B(X)) = (f_A \circ f_B)(X)$, $\forall X \in \mathbb{K}_n$. Αυτό σημαίνει ότι $f_{AB} = f_A \circ f_B$. Παρόμοια $f_{BA} = f_B \circ f_A$.

Από την Πρόταση 1.1, έπεται ότι οι γραμμικές απεικονίσεις f_{AB} και f_{BA} έχουν τις ίδιες ιδιοτιμές. Αυτό όμως είναι ισοδύναμο με το ότι οι πίνακες AB και BA έχουν τις ίδιες ιδιοτιμές. □

Παρατήρηση 1.3. *Από την Πόρισμα 1.2, έπεται αν $A, B \in \mathsf{M}_{n\times n}(\mathbb{K})$, τότε οι πίνακες AB και BA έχουν τις ίδιες ιδιοτιμές. Αυτό δεν σημαίνει ότι οι πίνακες είναι όμοιοι. Για παράδειγμα οι πίνακες*

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

έχουν το ίδιο χαρακτηριστικό πολυώνυμο και άρα τις ίδιες ιδιοτιμές. Όμως $A \cdot B = \mathbb{O} \neq B \cdot A$ και άρα οι πίνακες $A \cdot B$ και $B \cdot A$ δεν μπορεί να είναι όμοιοι.

1.2. **Χαρακτηριστικό Πολυώνυμο Γινομένου Πινάκων.** Ισχύει κάτι ισχυρότερο από το συμπέρασμα της Πρότασης 1.1, ή ισοδύναμα του Πορίσματος 1.2:

Θεώρημα 1.4. *Αν $A, B \in \mathsf{M}_{n\times n}(\mathbb{K})$. Τότε οι πίνακες AB και BA έχουν το ίδιο χαρακτηρισικό πολυώνυμο:*

$$P_{AB}(t) = P_{BA}(t)$$

Απόδειξη. *Πρώτη Περίπτωση*: Ένας εκ των πινάκων A, B είναι αντιστρέψιμος. Έστω ότι ο A είναι αντιστρέψιμος. Τότε:

$$BA = A^{-1}ABA = A^{-1}(AB)A$$

και επομένως οι πίνακες AB και BA είναι όμοιοι. Επειδή όμοιοι πίνακες έχουν το ίδιο χαρακτηρσιτικό πολυώνυμο, έπεται ότι: $P_{AB}(t) = P_{BA}(t)$. Παρόμοια αν ο B είναι αντιστρέψιμος, τότε $AB = B^{-1}BAB = B^{-1}(BA)B$, δηλαδή οι πίνακες AB και BA είναι όμοιοι και άρα $P_{AB}(t) = P_{BA}(t)$.

Δεύτερη Περίπτωση: Υποθέτουμε ότι ο A δεν είναι αντιστρέψιμος. Τότε η βαθμίδα του είναι $\mathbf{r}(A) := r < n$. Από την Γραμμική Άλγεβρα Ι γνωρίζουμε ότι ο A είναι ισοδύναμος με τον πίνακα

$$\widetilde{I}_r := \begin{pmatrix} I_r & \mathbb{O}_{r\times n-r} \\ \mathbb{O}_{n-r\times r} & \mathbb{O}_{n-r\times n-r} \end{pmatrix}$$

Δηλαδή υπάρχουν αντιστρέψιμοι πίνακες $Q_1, P_1 \in \mathsf{M}_{n\times n}(\mathbb{K})$ έτσι ώστε:

$$Q_1AP_1 = \widetilde{I}_r$$

Θέτοντας $Q = Q_1^{-1}$ και $P := P_1^{-1}$ θα έχουμε τότε:

$$A = Q\widetilde{I}_rP \tag{1.1}$$

Θεωρούμε τον πίνακα $C := PBQ$ τον οποίο τον χωρίζουμε σε υποπίνακες:

$$C = PBQ = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

όπου

$$C_{11} \in \mathsf{M}_{r\times r}(\mathbb{K}), \quad C_{12} \in \mathsf{M}_{r\times n-r}(\mathbb{K}), \quad C_{21} \in \mathsf{M}_{n-r\times r}(\mathbb{K}), \quad C_{22} \in \mathsf{M}_{n-r\times n-r}(\mathbb{K})$$

και τότε

$$B = P^{-1}CQ^{-1} = P^{-1} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} Q^{-1}$$

Επομένως:

$$AB = AP^{-1} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} Q^{-1} = Q\widetilde{I}_rPP^{-1} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} Q^{-1} = Q\widetilde{I}_r \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} Q^{-1}$$

Εύκολα βλέπουμε ότι:

$$\widetilde{I}_r \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ \mathbb{O}_{r\times r} & \mathbb{O}_{n-r\times n-r} \end{pmatrix}$$

και επομένως θα έχουμε:

$$AB = Q \begin{pmatrix} C_{11} & C_{12} \\ \mathbb{O}_{r\times r} & \mathbb{O}_{n-r\times n-r} \end{pmatrix} Q^{-1}$$

Δηλαδή ο πίνακας AB είναι όμοιος με τον πίνακα

$$\begin{pmatrix} C_{11} & C_{12} \\ \mathbb{O}_{r\times r} & \mathbb{O}_{n-r\times n-r} \end{pmatrix}$$

Επίσης:

$$BA = P^{-1}CQ^{-1}A = P^{-1} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} Q^{-1}\, Q\widetilde{I}_r P = P^{-1} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \widetilde{I}_r P$$

Εύκολα βλέπουμε ότι:

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \widetilde{I}_r = \begin{pmatrix} C_{11} & \mathbb{O}_{r\times n-r} \\ C_{21} & \mathbb{O}_{n-r\times n-r} \end{pmatrix}$$

Άρα θα έχουμε

$$BA = P^{-1} \begin{pmatrix} C_{11} & \mathbb{O}_{r\times n-r} \\ C_{21} & \mathbb{O}_{n-r\times n-r} \end{pmatrix} P$$

Δηλαδή ο πίνακας BA είναι όμοιος με τον πίνακα

$$\begin{pmatrix} C_{11} & \mathbb{O}_{r\times n-r} \\ C_{21} & \mathbb{O}_{n-r\times n-r} \end{pmatrix}$$

Επειδή όμοιοι πίνακες έχουν το ίδιο χαρακτηριστικό πολυώνυμο θα έχουμε

$$P_{AB}(t) = P_M(t), \qquad \text{όπου} \qquad M := \begin{pmatrix} C_{11} & C_{12} \\ \mathbb{O}_{r\times r} & \mathbb{O}_{n-r\times n-r} \end{pmatrix}$$

$$P_{BA}(t) = P_N(t), \qquad \text{όπου} \qquad N := \begin{pmatrix} C_{11} & \mathbb{O}_{r\times n-r} \\ C_{21} & \mathbb{O}_{n-r\times n-r} \end{pmatrix}$$

Όμως:

$$P_M(t) = \begin{vmatrix} C_{11} - tI_r & C_{12} \\ \mathbb{O}_{r\times r} & \mathbb{O}_{n-r\times n-r} - tI_{n-r\times n-r} \end{vmatrix} = (-1)^{n-r} P_{C_{11}}(t)$$

$$P_N(t) = \begin{vmatrix} C_{11} - tI_r & \mathbb{O}_{r\times n-r} \\ C_{21} & \mathbb{O}_{n-r\times n-r} - tI_{n-r\times n-r} \end{vmatrix} = (-1)^{n-r} P_{C_{11}}(t)$$

Άρα $P_M(t) = P_N(t)$ και επομένως

$$P_{AB}(t) = P_{BA}(t) \qquad \square$$

Θεώρημα 1.5. *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης, και έστω $f,\ g : \mathcal{E} \longrightarrow \mathcal{E}$ δύο γραμμικές απεικονίσεις. Τότε οι $f\circ g$ και $g \circ f$ έχουν το ίδιο χαρακτηριστικό πολυώνυμο:*

$$P_{f\circ g}(t) = P_{g\circ f}(t)$$

Απόδειξη. Έστω $\mathcal{B}$ μια βάση του $\mathcal{E}$ και $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$ και $B = M^{\mathcal{B}}_{\mathcal{B}}(g)$. Τότε από το Θεώρημα 1.4 έχουμε: $P_{AB}(t) = P_{BA}(t)$. Επειδή το χαρακτηριστικό πολυώνυμο μιας γραμμικής απεικόνισης συμπίπτει με το χαρακτηριστικό πολυώνυμο του πίνακά της σε τυχούσα βάση, έπεται ότι:

$$P_f(t) = P_A(t) \qquad \text{και} \qquad P_g(t) = P_B(t)$$

Επειδή ο πίνακας της $f\circ g$ στην βάση $\mathcal{B}$ είναι ο AB και ο πίνακας της $g\circ f$ στην βάση $\mathcal{B}$ είναι ο BA, θα έχουμε:

$$P_{f\circ g}(t) = P_{AB}(t) = P_{BA}(t) = P_{g\circ f}(t) \qquad \square$$

2. Τριγωνοποίηση

2.1. **Άνω Τριγωνικοί Πίνακες και Κάτω Τριγωνικοί Πίνακες.** Όπως γνωρίζουμε ένας τετραγωνικός πίνακας $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ καλείται *τριγωνοποιήσιμος* αν και μόνον αν ο A είναι όμοιος με έναν άνω τριγωνικό πίνακα.

Το παρακάτω αποτέλεσμα δείχνει ότι ο A είναι τριγωνοποιήσιμος αν και μόνον αν ο A είναι όμοιος με έναν *κάτω* τριγωνικό πίνακα.

Πρόταση 2.1. *Για έναν πίνακα $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ τα ακόλουθα είναι ισοδύναμα:*

(1) *Ο A είναι όμοιος με έναν άνω τριγωνικό πίνακα.*
(2) *Ο A είναι όμοιος με έναν κάτω τριγωνικό πίνακα.*

Απόδειξη. Θεωρούμε τον πίνακα:

$$T = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \iddots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

Εύκολα υπολογίζουμε ότι $T^2 = I_n$ και επομένως ο πίνακας T είναι αντιστρέψιμος και $T^{-1} = T$.

(1) $\Longrightarrow$ (2) Έστω ότι ο A είναι όμοιος με έναν άνω τριγωνικό πίνακα B. Τότε υπάρχει αντιστρέψιμος πίνακας P έτσι ώστε:

$$P^{-1} \cdot A \cdot P = B$$

Τότε:

$$TP^{-1} \cdot A \cdot PT = TBT \quad \Longrightarrow \quad T^{-1}P^{-1} \cdot A \cdot PT = TBT \quad \Longrightarrow \quad (PT)^{-1} \cdot A \cdot PT = TBT$$

Υπολογίζουμε:

$$TBT =$$

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \iddots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} k_{11} & k_{12} & k_{13} & \cdots & k_{1\,n-2} & k_{1\,n-1} & k_{1n} \\ 0 & k_{22} & k_{23} & \cdots & k_{2\,n-2} & k_{2,n-1} & k_{2n} \\ 0 & 0 & k_{33} & \cdots & k_{3\,n-2} & k_{3\,n-1} & k_{3\,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & k_{n-2\,n-2} & k_{n-2\,n-1} & k_{n-2\,n} \\ 0 & 0 & 0 & \cdots & 0 & k_{n-1\,n-1} & k_{n-1\,n} \\ 0 & 0 & 0 & \cdots & 0 & 0 & k_{nn} \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \iddots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} k_{nn} & 0 & 0 & \cdots & 0 & 0 & 0 \\ k_{n-1\,n} & k_{n-1\,n-1} & 0 & \cdots & 0 & 0 & 0 \\ k_{n-2\,n} & k_{n-2\,n-1} & k_{n-2\,n-2} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ k_{2\,n} & k_{2\,n-1} & k_{2\,n-2} & \cdots & k_{23} & k_{22} & 0 \\ k_{1n} & k_{1\,n-1} & k_{1\,n-2} & \cdots & k_{13} & k_{12} & k_{11} \end{pmatrix}$$

Ο τελευταίος πίνακας είναι κάτω τριγωνικός. Άρα ο A είναι όμοιος με έναν κάτω τριγωνικό πίνακα.

(2) $\Longrightarrow$ (1) Η απόδειξη είναι παρόμοια (χρησιμοποιώντας πάλι τον πίνακα T). $\square$

2.2. **Πότε είναι ένας πίνακας όμοιος με έναν άνω τριγωνικό πίνακα;**

Θεώρημα 2.2. *Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας $n \times n$ πίνακας και $P_A(t)$ το χαρακτηριστικό πολυώνυμο του A. Ο πίνακας A είναι όμοιος με έναν άνω τριγωνικό πίνακα, αν και μόνο αν, όλες οι ρίζες τού $P_A(t)$ ανήκουν στο σώμα $\mathbb{K}$.*

Απόδειξη. «$\Longrightarrow$» Υπενθυμίζουμε ότι όμοιοι πίνακες έχουν το ίδιο χαρακτηριστικό πολυώνυμο. Αν λοιπόν, ο A είναι όμοιος με έναν άνω τριγωνικό πίνακα $D \in \mathsf{M}_{n\times n}(\mathbb{K})$, τότε $P_A(t) = P_D(t)$. Αλλά το $P_D(t)$ ισούται με[1] $(t-\lambda_1)(t-\lambda_2)\dots(t-\lambda_n)$, όπου τα $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{K}$ είναι τα στοιχεία τής κύριας διαγωνίου του D. Επιπλέον, το $P_A(t) = P_D(t)$ είναι ένα πολυώνυμο τού $\mathbb{K}[t]$ βαθμού n και αφού εκφράζεται ως γινόμενο γραμμμικών παραγόντων, τα $\lambda_1, \lambda_2, \dots, \lambda_n$ είναι ακριβώς όλες οι ρίζες του, οι οποίες ανήκουν στο $\mathbb{K}$.

«$\Longleftarrow$» Η απόδειξη θα εκτελεστεί με επαγωγή ως προς το μέγεθος n τού πίνακα A.

(*Επαγωγική Θεμελίωση*) Η αλήθεια τού θεωρήματος είναι φανερή για $n = 1$, αφού κάθε 1×1 πίνακας είναι άνω τριγωνικός.

(*Επαγωγική Υπόθεση*) Έστω ότι κάθε $(n-1)\times(n-1)$ πίνακας, τού οποίου όλες οι ρίζες τού χαρακτηριστικού του πολυωνύμου ανήκουν στο $\mathbb{K}$, είναι όμοιος με έναν άνω τριγωνικό πίνακα.

(*Επαγωγική Απόδειξη*) Θα δείξουμε ότι το θεώρημα ισχύει, για κάθε $n \times n$ πίνακα A, τού οποίου όλες οι ρίζες τού χαρακτηριστικού του πολυωνύμου ανήκουν στο $\mathbb{K}$.

Ας είναι A ένας τέτοιος πίνακας και λ_1 μια ρίζα τού $P_A(t)$. Υπεθυμίζουμε ότι ο A ορίζει έναν $\mathbb{K}$-γραμμικό ενδομορφισμό

$$f_A : \mathsf{M}_{n\times 1}(\mathbb{K}) = \mathbb{K}_n \longrightarrow \mathsf{M}_{n\times 1}(\mathbb{K}) = \mathbb{K}_n, \qquad \begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \vdots \\ \kappa_n \end{pmatrix} \longmapsto A \cdot \begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \vdots \\ \kappa_n \end{pmatrix}$$

τού χώρου $\mathsf{M}_{n\times 1}(\mathbb{K})$ των $n \times 1$ πινάκων, δηλαδή τού χώρου $\mathbb{K}_n$ των στηλών με n συνιστώσες από το $\mathbb{K}$, και επιπλέον ο A είναι ο πίνακας που παριστάνει τον f_A ως προς την κανονική βάση $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ τού $\mathsf{M}_{n\times 1}(\mathbb{K})$, δηλαδή $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$.

Ας είναι $\vec{v}_1 \in \mathsf{M}_{n\times 1}(\mathbb{K})$ ένα ιδιοδιάνυσμα τού f_A που αντιστοιχεί στην ιδιοτιμή λ_1, δηλαδή $f_A(\vec{v}_1) = A\vec{v}_1 = \lambda_1\vec{v}_1$. Επειδή το $\vec{v}_1 \neq \vec{0}$, μπορούμε να θεωρήσουμε μια νέα βάση $\mathcal{B}'$ τού $\mathsf{M}_{n\times 1}(\mathbb{K})$ με πρώτο διάνυσμά της το $\vec{v}_1$, δηλαδή $\mathcal{B}' = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$. Ως προς τη νέα βάση $\mathcal{B}'$ ο πίνακας που παριστάνει τον f_A είναι ο

$$M_{\mathcal{B}'}^{\mathcal{B}'}(f) = \left(\begin{array}{c|ccccc} \lambda_1 & * & * & * & \dots & * \\ \hline 0 & & & & & \\ \vdots & & & & & \\ 0 & & & B & & \\ \vdots & & & & & \\ 0 & & & & & \end{array}\right),$$

όπου B είναι ένας $(n-1)\times(n-1)$ πίνακας.

Οι πίνακες $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ και $M_{\mathcal{B}'}^{\mathcal{B}'}(f)$ είναι όμοιοι, αφού παριστάνουν τον ίδιο γραμμικό ενδομορφισμό, ως προς τις βάσεις $\mathcal{B}$ και $\mathcal{B}'$ αντιστοίχως και γι' αυτό το χαρακτηριστικό πολυώνυμο $P_A(t)$ τού A ισούται με το χαρακτηριστικό πολυώνυμο $q(t)$ τού πίνακα $M_{\mathcal{B}'}^{\mathcal{B}'}(f)$. Αλλά το $q(t) = (t-\lambda_1)P_B(t)$, όπου $P_B(t)$ είναι το χαρακτηριστικό πολυώνυμο τού $(n-1)\times(n-1)$ υποπίνακα B, ο οποίος εμφανίζεται στην προηγούμενη μορφή τού $M_{\mathcal{B}'}^{\mathcal{B}'}(f)$, αφού

[1] Αυτό ισχύει όταν το χαρακτηριστικό πολυώνυμο ορίζεται ως $P_A(t) = \mathrm{Det}(XI_n - A)$. Όταν ορίζεται ως $P_A(t) = \mathrm{Det}(A - XI_n)$, τότε $P_D(t) = (-1)^n(t-\lambda_1)(t-\lambda_2)\dots(t-\lambda_n)$

$$\mathrm{Det}(tI_n - M_{\mathcal{B}'}^{\mathcal{B}'}(f)) = \mathrm{Det}\begin{pmatrix} t-\lambda_1 & * \quad * \quad * \quad \cdots \quad * \\ 0 & \\ \vdots & \\ 0 & \boxed{tI_{n-1}-B} \\ \vdots & \\ 0 & \end{pmatrix} = (t-\lambda_1)\mathrm{Det}(tI_{n-1}-B),$$

λόγω τής ειδικής μορφής που έχει ο $\mathrm{Det}M_{\mathcal{B}'}^{\mathcal{B}'}(f)$. Επειδή λοιπόν, $P_A(t) = q(t) = (t-\lambda_1)P_B(t)$, όλες οι ρίζες τού χαρακτηριστικού πολυωνύμου $P_B(t)$ τού B ανήκουν στο $\mathbb{K}$ και γι' αυτό, λόγω τής επαγωγικής υπόθεσης, ο B είναι όμοιος με έναν άνω τριγωνικό πίνακα C, δηλαδή υπάρχει ένας αντιστρέψιμος $(n-1)\times(n-1)$ πίνακας S, έτσι ώστε ο $S^{-1}BS = C$ να είναι ένας άνω τριγωνικός πίνακας.

Τώρα έχουμε:

$$\begin{pmatrix} 1 & 0 \quad 0 \quad 0 \quad \cdots \quad 0 \\ 0 & \\ \vdots & \\ 0 & \boxed{S^{-1}} \\ \vdots & \\ 0 & \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & * \quad * \quad * \quad \cdots \quad * \\ 0 & \\ \vdots & \\ 0 & \boxed{B} \\ \vdots & \\ 0 & \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \quad 0 \quad 0 \quad \cdots \quad 0 \\ 0 & \\ \vdots & \\ 0 & \boxed{S} \\ \vdots & \\ 0 & \end{pmatrix} =$$

$$\begin{pmatrix} \lambda_1 & * \quad * \quad * \quad \cdots \quad * \\ 0 & \\ \vdots & \\ 0 & \boxed{S^{-1}BS} \\ \vdots & \\ 0 & \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \quad * \quad * \quad \cdots \quad * \\ 0 & \\ \vdots & \\ 0 & \boxed{C} \\ \vdots & \\ 0 & \end{pmatrix}$$

Επομένως, ο πίνακας

$$M_{\mathcal{B}'}^{\mathcal{B}'}(f) = \begin{pmatrix} \lambda_1 & * \quad * \quad * \quad \cdots \quad * \\ 0 & \\ \vdots & \\ 0 & \boxed{B} \\ \vdots & \\ 0 & \end{pmatrix}.$$

είναι όμοιος με τον πίνακα

$$\begin{pmatrix} \lambda_1 & * \quad * \quad * \quad \cdots \quad * \\ 0 & \\ \vdots & \\ 0 & \boxed{C} \\ \vdots & \\ 0 & \end{pmatrix},$$

που είναι άνω τριγωνικός επειδή ο C είναι άνω τριγωνικός. Έτσι ο αρχικός πίνακας A είναι όμοιος με τον συγκεκριμένο άνω τριγωνικό πίνακα, αφού είναι όμοιος με τον $M^{\mathcal{B}'}_{\mathcal{B}'}(f)$. $\square$

2.3. **Αλγόριθμος Τριγωνοποίησης Πίνακα.** Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας πίνακας και υποθέτουμε ότι όλες οι ιδιοτιμές του $\lambda_1, \lambda_2, \cdots, \lambda_n$ ανήκουν στο σώμα $\mathbb{K}$, έτσι ώστε το χαρακτηριστικό πολυώνυμο του A να αναλύεται ως εξής:

$$P_A(t) = (-1)^n (t-\lambda_1)^{a_1})(t-\lambda_2)^{a_2} \ \cdots \ (t-\lambda_k)^{a_k}$$

Έτσι $\lambda_1, \lambda_2, \cdots, \lambda_k \in \mathbb{K}$ είναι οι διακεκριμμένες ιδιοτιμές του A. Τότε όπως γνωρίζουμε ο πίνακας A είναι τριγωνοποιήσιμος. Επομένως υπάρχει αντιστρέψιμος πίνακας P έτσι ώστε:

$$P^{-1}\ A\ P = \begin{pmatrix} \lambda_1 & \epsilon_{12} & \cdots & \epsilon_{1n} \\ 0 & \lambda_2 & \cdots & \epsilon_{2n} \\ \vdots & \vdots & \ddots & \epsilon_{n-1\,n} \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Αλγόριθμος Τριγωνοποίησης του A

1. Επιλέγουμε μια ιδιοτιμή λ_1 του A με αντίστοιχο ιδιοδιάνυσμα-στήλη E_1:

$$A \cdot E_1 = \lambda_1 E_1$$

και συμπλήρώνουμε το ιδιοδιάνυσμα E_1 σε μια βάση

$$\mathcal{B}_1 = \big\{E_1, E_2, \cdots, E_n\big\}$$

του χώρου των στηλών $\mathbb{K}_n$.

Θεωρούμε τον αντιστρέψιμο πίνακα

$$P_1 = \big(E_1 E_2 \cdots E_n\big)$$

ο οποίος σχηματίζεται από τις στήλες της βάσης $\mathcal{B}_1$. Τότε ο πίνακας $P_1^{-1} \cdot A \cdot P_1$ θα έχει την ακόλουθη μορφή:

$$P_1^{-1} \cdot A \cdot P_1 = \left(\begin{array}{c|c} \lambda_1 & * \quad * \quad * \quad \dots \quad * \\ \hline 0 & \\ \vdots & \\ 0 & B_1 \\ \vdots & \\ 0 & \end{array}\right)$$

Ο $(n-1)\times(n-1)$ πίνακας B_1 έχει όλες τις ιδιοτιμές του στο $\mathbb{K}$ διότι

$$P_A(t) = P_{P_1^{-1}\cdot A \cdot P_1}(t) = (\lambda_1 - t) P_{B_1}(t)$$

(α΄) Αν ο $(n-1)\times(n-1)$ πίνακας B_1 είναι άνω τριγωνικός, ο αλγόριθμος σταματάει και έχουμε βρει την άνω τριγωνική μορφή του A η οποία είναι ο πίνακας $P^{-1}\cdot A \cdot P_1$.

(β΄) Αν ο $(n-1)\times(n-1)$ πίνακας B_1 δεν είναι άνω τριγωνικός, τότε προχωρούμε στο επόμενο βήμα:

2. Επιλέγουμε μια ιδιοτιμή λ_2 του B_1 (η οποία εκ΄ κατασκευής είναι και ιδιοτιμή του A με αντίστοιχο ιδιοδιάνυσμα E_2':

$$B_1 \cdot E_2' = \lambda_2 E_2'$$

και συμπλήρώνουμε το ιδιοδιάνυσμα E_2' σε μια βάση

$$\mathcal{B}_2 = \left\{E_2', E_3', \cdots, E_n'\right\}$$

του χώρου των στηλών $\mathbb{K}_{n-1}$.

Θεωρούμε τον αντιστρέψιμο $(n-1) \times (n-1)$ πίνακα

$$P_2 = \left(E_2' E_3' \cdots E_n'\right)$$

ο οποίος σχηματίζεται από τις στήλες της βάσης $\mathcal{B}_2$. Τότε ο πίνακας $P_2^{-1} \cdot B_1 \cdot P_2$ θα έχει την ακόλουθη μορφή:

$$P_2^{-1} \cdot B_1 \cdot P_2 = \begin{pmatrix} \lambda_2 & * & * & * & \ldots & * \\ 0 & & & & & \\ \vdots & & & & & \\ 0 & & & B_2 & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix}$$

Ο $(n-2) \times (n-2)$ πίνακας B_2 έχει όλες τις ιδιοτιμές του στο $\mathbb{K}$ διότι

$$P_{B_1}(t) = P_{P_2^{-1} \cdot B_1 \cdot P_2}(t) = (\lambda_2 - t) P_{B_2}(t)$$

(α΄) Αν ο $(n-1) \times (n-1)$ πίνακας B_2 είναι άνω τριγωνικός, ο αλγόριθμος σταματάει και έχουμε βρει την άνω τριγωνική μορφή του A ως εξής:

Θέτουμε

$$Q_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & & & & & \\ \vdots & & & & & \\ 0 & & & P_2 & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix}$$

Τότε ο πίνακας Q_2 είναι αντιστρέψιμος, διότι $|Q_2| = 1|P_2| = |P_2| \neq 0$, και

$$Q_2^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & & & & & \\ \vdots & & & & & \\ 0 & & & P_2^{-1} & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix}$$

Επιπλέον έχουμε:

$$(P_1 Q_2)^{-1} \cdot A (P_1 Q_2) = Q_2^{-1} \cdot (P_1^{-1} \cdot A \cdot P_1) \cdot Q_2 =$$

$$Q_2^{-1} \cdot \begin{pmatrix} \lambda_1 & * & * & * & \dots & * \\ 0 & & & & & \\ \vdots & & & & & \\ 0 & & & B_1 & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix} \cdot Q_2 =$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & & & & & \\ \vdots & & & & & \\ 0 & & & P_2^{-1} & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & * & * & * & \dots & * \\ 0 & & & & & \\ \vdots & & & & & \\ 0 & & & B_1 & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & & & & & \\ \vdots & & & & & \\ 0 & & & P_2 & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix} =$$

$$\begin{pmatrix} \lambda_1 & * & * & * & \dots & * \\ 0 & & & & & \\ \vdots & & & & & \\ 0 & & & P_2^{-1} \cdot B_1 \cdot P_2 & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix} = \begin{pmatrix} \lambda_1 & * & * & * & \dots & * \\ 0 & \lambda_2 & * & * & \dots & * \\ 0 & 0 & & & & \\ \vdots & & & & & \\ 0 & 0 & & B_2 & & \\ \vdots & \vdots & & & & \\ 0 & 0 & & & & \end{pmatrix}$$

Επειδή ο B_2 είναι άνω τριγωνικός, έπεται ότι και ο A είναι άνω τριγωνικός και ο τελευταίος πίνακας, δηλαδή ο πίνακας $(P_1Q_2)^{-1} \cdot A(P_1Q_2)$ είναι μια άνω τριγωνική μορφή του A.

(β΄) Αν ο $(n-2) \times (n-2)$ πίνακας B_2 δεν είναι άνω τριγωνικός, τότε προχωρούμε στο επόμενο βήμα:

3. Το βήμα αυτό είναι η επανάληψη του βήματος **2.** για τον $(n-2) \times (n-2)$ πίνακα B_2.

4. Συνεχίζουμε την παραπάνω διαδικασία η οποία μετά από πεπερασμένο πλήθος βημάτων θα μας οδηγήσει στην άνω τριγωνική μορφή του πίνακα A.

Παράδειγμα 2.3. *Θεωρούμε τον πίνακα*

$$A = \begin{pmatrix} 8 & 14 & 11 \\ -3 & -5 & -5 \\ -1 & -2 & 0 \end{pmatrix}$$

Υπολογίζουμε το χαρακτηριστικό πολυώνυμο του πίνακα A:

$$P_A(t) = -(t-1)^3$$

και άρα η μόνη ιδιοτιμή του A είναι η $\lambda = 1$ με αλγεβρική πολλαπλότητα 3.

Επομένως ο πίνακας A είναι τριγωνοποιήσιμος. Εφαρμόζουμε το Αλγόριθμο Τριγωνοποίησης για να υπολογίσουμε μια άνω τριγωνική μορφή του A:

1. *Για την ιδιοτιμής $\lambda = 1$, υπολογίζουμε τον ιδιοχώρο*

$$\mathcal{V}(1) = \Big\{ \begin{pmatrix} k \\ -k/2 \\ 0 \end{pmatrix} \in \mathbb{R}_3 \mid k \in \mathbb{R} \Big\} = \Big\langle \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \Big\rangle$$

και άρα έβα ιδιοδιάνυσμα του A το οποίο αντιστοιχεί στην ιδιοτιμή $\lambda = 1$ είναι το

$$E_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$$

Συμπληρώνουμε το E_1 σε μια βάση $\mathcal{B}_1$ του $\mathbb{R}_3$, για παράδειγμα:

$$\mathcal{B}_1 = \Big\{ E_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \ \ E_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \ \ E_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \Big\}$$

Θεωρούμε τον αντιστρέψιμο πίνακα $P_1 = (E_1 E_2 E_3)$:

$$P_1 = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Υπολογίζουμε τον αντίστροφο του P_1:

$$P_1^{-1} = \begin{pmatrix} 1/2 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

και ακολούθως τον πίνακα $P_1^{-1} \cdot A \cdot P_1$:

$$P_1^{-1} \cdot A \cdot P_1 = \begin{pmatrix} 1 & 7 & 1/2 \\ 0 & 2 & -2 \\ 0 & -2 & 0 \end{pmatrix}$$

Ο 2×2 πίνακας

$$B_1 := \begin{pmatrix} 2 & -2 \\ -2 & 0 \end{pmatrix}$$

δεν είναι άνω τριγωνικός, επομένως προχωρούμε στο επόμενο βήμα:

2. *Οι ιδιοτιμές τοπυ B_1 είναι οι ιδιοτιμές του A, και άρα η μόνη ιδιοτιμή του B_1 είναι η $\lambda = 1$ με αλγεβρική πολλαπλότητα 2.*

Για την ιδιοτιμής $\lambda = 1$ του 2×2 πίνακα B_1, υπολογίζουμε τον ιδιοχώρο

$$\mathcal{V}(1) = \Big\{ \begin{pmatrix} k \\ -2k \end{pmatrix} \in \mathbb{R}_2 \mid k \in \mathbb{R} \Big\} = \Big\langle \begin{pmatrix} 1 \\ -2 \end{pmatrix} \Big\rangle$$

και άρα ένα ιδιοδιάνυσμα του B_1 το οποίο αντιστοιχεί στην ιδιοτιμή $\lambda = 1$ είναι το

$$E_2' = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Συμπληρώνουμε το E_2' σε μια βάση $\mathcal{B}_2$ του $\mathbb{R}_2$, για παράδειγμα:

$$\mathcal{B}_2 = \Big\{ E_2' = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \ \ E_3' = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Big\}$$

Θεωρούμε τον αντιστρέψιμο πίνακα $P_1 = (E_2' E_3')$:

$$P_2 = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$$

Υπολογίζουμε τον αντίστροφο του P_2:

$$P_1^{-1} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

και ακολούθως τον πίνακα $P_1^{-1} \cdot A \cdot P_1$:

$$P_1^{-2} \cdot B_1 \cdot P_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} := B_2$$

ο οποίος είναι άνω τριγωνικός. Άρα ο αλγόριθμος τριγωνοποίησης σταματάει και μπορούμε να υπολογίσουμε την άνω τριγωνική μορφή ως εξής:

Θέτουμε

$$Q_2 = \begin{pmatrix} 1 & \begin{matrix} 0 & 0 \end{matrix} \\ \begin{matrix} 0 \\ 0 \end{matrix} & \boxed{P_1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$$

Τότε μπορούμε να υπολογίσουμε την άνω τριγωνική μορφή του πίνακα A ως εξής:

$$(P_1 Q_2)^{-1} \cdot A \cdot (P_1 Q_2) = \begin{pmatrix} 1 & -4 & 1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

3. **Το Θεώρημα των Cayley-Hamilton**

3.1. **Πολυωνυμικές Γραμμικές Απεικονίσεις και Πολυωνυμικοί Πίνακες.**

1. Έστω $\mathcal{E}$ ένας διανυσματικός χώρος πεπερασμένης διάστασης υπεράνω ενός σώματος $\mathbb{K}$.

Σταθεροποιούμε μια γραμμική απεικόνιση $f\colon \mathcal{E} \longrightarrow \mathcal{E}$.

Για κάθε πολυώνυμο $P(t) = a_0 + a_1 t + \cdots + a_{n-1}t^{n-1} + a_n t^n$ ορίζουμε μια απεικόνιση

$$P(f)\ :\ \mathcal{E} \longrightarrow \mathcal{E}, \quad \vec{x} \longmapsto P(f)(\vec{x})$$

όπου

$$\begin{aligned} P(f)(\vec{x}) &:= a_0 \mathrm{Id}_{\mathcal{E}}(\vec{x}) + a_1 f(\vec{x}) + \cdots + a_{n-1} f^{n-1}(\vec{x}) + a_n f^n(\vec{x}) \\ &= a_0 \vec{x} + a_1 f(\vec{x}) + \cdots + a_{n-1} f^{n-1}(\vec{x}) + a_n f^n(\vec{x}) \end{aligned}$$

Αν $\mathsf{Hom}_{\mathbb{K}}(\mathcal{E},\mathcal{E}) = \{f\colon \mathcal{E} \longrightarrow \mathcal{E} \mid f \text{ είναι γραμμική}\}$ είναι ο $\mathbb{K}$-διανυσματικός χώρος των γραμμικών απεικονίσεων από τον $\mathcal{E}$ στον εαυτό του, τότε η παραπάνω διαδικασία ορίζει μια απεικόνιση

$$\mathbb{K}[t] \times \mathsf{Hom}_{\mathbb{K}}(\mathcal{E},\mathcal{E}) \longrightarrow \mathsf{Hom}_{\mathbb{K}}(\mathcal{E},\mathcal{E}), \quad \big(P(t), f\big) \longmapsto P(f)$$

Γραμμικές απεικονίσεις της μορφής $P(f)$, όπου $P(t) \in \mathbb{K}[t]$, καλούνται *πολυωνυμικές γραμμικές απεικονίσεις*.

2. Σταθεροποιούμε έναν $n \times n$ πίνακα $A \in \mathsf{M}_{n\times n}(\mathbb{K})$.

Για κάθε πολυώνυμο $P(t) = a_0 + a_1 t + \cdots + a_{n-1}t^{n-1} + a_n t^n$ ορίζουμε έναν $n \times n$ πίνακα

$$P(A) := a_0 I_n + a_1 A + \cdots + a_{n-1} A^{n-1} + a_n A^n \in \mathsf{M}_{n\times n}(\mathbb{K})$$

Η παραπάνω διαδικασία ορίζει μια απεικόνιση

$$\mathbb{K}[t] \times \mathsf{M}_{n\times n}(\mathbb{K}) \longrightarrow \mathsf{M}_{n\times n}(\mathbb{K}), \quad \big(P(t), A\big) \longmapsto P(A)$$

Τετραγωνικοί πίνακες της μορφής $P(A)$, όπου $P(t) \in \mathbb{K}[t]$, καλούνται *πολυωνυμικοί πίνακες*.

3.2. **Το Θεώρημα των Cayley-Hamilton.**

Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος και $P(t) \in \mathbb{K}[t]$ ένα πολυώνυμο με συντελεστές από το σώμα $\mathbb{K}$. Έστω επίσης $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση, και $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας τετραγωνικός πίνακας.

Ορισμός 3.1. (1) *Το $P(t)$ μηδενίζει την γραμμική απεικόνιση $f\colon \mathcal{E} \longrightarrow \mathcal{E}$, αν: $P(f) = 0$.*
(2) *Το $P(t)$ μηδενίζει τον τετραγωνικό πίνακα $A \in \mathsf{M}_{n\times n}(\mathbb{K})$, αν: $P(A) = \mathbb{O}$.*

Το ακόλουθο σημαντικό Θεώρημα των Cayley-Hamilton πιστοποιεί ότι το χαρακτηριστικό πολυώνυμο ενός πίνακα μηδενίζει τον πίνακα. Στην απόδειξη του Θεωρήματος εμπλέκονται πίνακες τα στοιχεία των οποίων είναι πολυώνυμα. Οι βασικές ιδιότητες πινάκων με στοιχεία αριθμούς από το σώμα που θα χρησιμοποιηθούν στην απόδειξη, είναι εύκολο να δει κανείς ότι ισχύουν και για πίνακες με στοιχεία πολυώνυμα.

Θεώρημα 3.2. *(Cayley-Hamilton) Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας $n \times n$ πίνακας. και $P_A(t)$ το χαρακτηριστικό πολυώνυμο του A. Τότε ο πολυωνυμικός πίνακας $P_A(A)$ είναι ο μηδενικός:*

$$P_A(A) = \mathbb{O}$$

Απόδειξη. Γνωρίζουμε ότι $P_A(t) = |A - tI_n| = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$. Θα δείξουμε ότι

$$P_A(A) = (-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n = \mathbb{O} \tag{3.1}$$

Υπενθυμιζουμε ότι για κάθε τετραγωνικό πίνακα $M \in \mathsf{M}_{n\times}(\mathbb{K})$ ισχύει:

$$M \cdot \mathsf{adj}(M) = |M| \cdot I_n$$

Επομένως θα έχουμε την ακόλουθη ισότητα πινάκων με στοιχεία πολυώνυμα:

$$(A - tI_n) \cdot \mathsf{adj}(A - tI_n) = |A - tI_n| \cdot I_n = P_A(t) \cdot I_n \tag{3.2}$$

Τα στοιχεία του πίνακα πολυωνύμων $\mathsf{adj}(A - tI_n)$ είναι πολυώνυμα τα οποία προκύπτουν ως ελάσσονες ορίζουσες τάξης $n-1$ του πίνακα $A-tI_n$, και άρα θα είναι πολυώνυμα βαθμού το πολύ $n-1$. Επομένως υπάρχουν πίνακες $B_0, B_1, \cdots, B_{n-1} \in \mathsf{M}_{n\times n}(\mathbb{K})$ έτσι ώστε:

$$\mathsf{adj}(A - tI_n) = B_{n-1}t^{n-1} + B_{n-2}t^{n-2} + \cdots + B_1 t + B_0 \tag{3.3}$$

Τότε συνδυάζοντας τις (2.2) και (2.3) θα έχουμε:

$$\begin{aligned}
(A - tI_n) \cdot \mathsf{adj}(A - tI_n) = P_A(t) \cdot I_n = ((-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0) \cdot I_n \quad &\Longrightarrow \\
A \cdot B_{n-1}t^{n-1} + A \cdot B_{n-2}t^{n-2} + \cdots + A \cdot B_1 t + A \cdot B_0 - B_{n-1}t^n - B_{n-2}t^{n-1} - \cdots - B_1 t^2 - B_0 t \quad &= \\
(-1)^n t^n I_n + a_{n-1}t^{n-1} I_n + \cdots + a_1 t I_n + a_0 I_n \quad &\Longrightarrow \\
-B_{n-1}t^n - (A \cdot B_{n-1} + B_{n-2})t^{n-1} + \cdots + (A \cdot B_2 - B_1)t^2 + (A \cdot B_1 - B_0)t + A \cdot B_0 \quad &= \\
(-1)^n t^n I_n + a_{n-1}t^{n-1} I_n + \cdots + a_1 t I_n + a_0 I_n &
\end{aligned}$$

Η τελευταία ισότητα είναι μια ισότητα πινάκων με στοιχεία πολυώνυμα, και επομένως θα έχουμε:

$$\begin{gathered}
-B_{n-1} = (-1)^n I_n \\
A \cdot B_{n-1} + B_{n-2} = a_{n-1} I_n \\
\vdots \\
A \cdot B_1 - B_0 = a_1 I_n \\
A \cdot B_0 = a_0 I_n
\end{gathered}$$

Πολλαπλασιάζουμε την πρώτη ισότητα πινάκων με τον πίνακα A^n, την δεύτερη με τον πίνακα A^{n-1}, $\cdots$, την προτελευταία με τον πίνακα A, και την τελευταία με τον πίνακα I_n. Τότε θα έχουμε τις ακόλουθες ισότητες πινάκων:

$$\begin{gathered}
-A^n \cdot B_{n-1} = (-1)^n A^n \cdot I_n = (-1)^n A^n \\
A^n \cdot B_{n-1} + A^{n-1} \cdot B_{n-2} = a_{n-1}A^{n-1} \cdot I_n = a_{n-1}A^{n-1} \\
\vdots \\
A^2 \cdot B_1 - A \cdot B_0 = a_1 A \cdot I_n = a_1 A \\
A \cdot B_0 = a_0 I_n
\end{gathered}$$

Προσθέτοντας τις παραπάνω ισότητες πινάκων, βλέπουμε ότι το πρώτο μέλος της προκύπτουσας ισότητας είναι ο μηδενικός πίνακας $\mathbb{O}$ και το δεύτερο μέλος είναι ο πίνακας $(-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n = P_A(A)$. Επομένως δείξαμε ότι:

$$P_A(A) = \mathbb{O}$$

□

Θα δείξουμε στη συνέχεια ότι το Θεώρημα των Cayley-Hamilton ισχύει και για γραμμικές απεικονίσεις. Πριν περάσουμε στην απόδειξη, θα χρειασθούμε το ακόλουθο Λήμμα:

Λήμμα 3.3. *Έστω $\mathcal{E}$ ένας διανυσματικός χώρος πεπερασμένης διάστασης υπεράνω ενός σώματος $\mathbb{K}$ και $\mathcal{B}$ μια βάση του $\mathcal{E}$. Έστω $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση, και $P(t) \in \mathbb{K}[t]$ ένα πολυώνυμο. Τότε:*

$$M^{\mathcal{B}}_{\mathcal{B}}\big(P(f)\big) = P\big(M^{\mathcal{B}}_{\mathcal{B}}(f)\big)$$

Απόδειξη. Έστω $P(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_{m-1}t^{m-1} + a_m t^m$. Τότε:

$$P(M^{\mathcal{B}}_{\mathcal{B}}(f)) = a_0 I_n + a_1 M^{\mathcal{B}}_{\mathcal{B}}(f) + a_2 M^{\mathcal{B}}_{\mathcal{B}}(f)^2 + \cdots + a_{m-1} M^{\mathcal{B}}_{\mathcal{B}}(f)^{m-1} + a_m M^{\mathcal{B}}_{\mathcal{B}}(f)^m \tag{3.4}$$

Υπενθυμίζουμε από την Γραμμική Άλγεβρα Ι, ότι αν $\mathcal{B}$ είναι μια βάση του $\mathcal{E}$, τότε η απεικόνιση

$$\Phi : \ \mathsf{Hom}_{\mathbb{K}}(\mathcal{E}, \mathcal{E}) \ \longrightarrow \ \mathsf{M}_{n\times n}(\mathbb{K}), \quad \Phi(f) = M^{\mathcal{B}}_{\mathcal{B}}(f)$$

είναι ισομορφισμός $\mathbb{K}$-διανυσματικών χώρων, όπου $n = \dim_{\mathbb{K}}\mathcal{E}$. Επιπλέον η Φ στέλνει την σύνθεση γραμμικών απεικονίσεων στο γινόμενο των αντίστοιων πινάκων, δηλαδή ισχύει:

$$\Phi(f \circ g) = \Phi(f) \cdot \Phi(g) = M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot M^{\mathcal{B}}_{\mathcal{B}}(g)$$

Ιδιαίτερα θα έχουμε:

$$\Phi(f^k) = \Phi(f \circ f \circ \cdots \circ f) = \Phi(f) \cdot \Phi(f) \cdot \ \cdots \ \cdot \Phi(f) = \Phi(f)^k = M^{\mathcal{B}}_{\mathcal{B}}(f)^k$$

Επίσης η Φ και στέλνει την ταυτοτική γραμμική απεικόνιση $\mathrm{Id}_{\mathcal{E}}$ στον ταυτοτικό $n \times n$ πίνακα I_n:

$$\Phi(\mathrm{Id}_{\mathcal{E}}) = I_n$$

Χρησιμοποιώντας τις παραπάνω ιδιότητες του ισομορφισμού Φ, η σχέση (3.4) δίνει:

$$P(M^{\mathcal{B}}_{\mathcal{B}}(f)) = a_0\Phi(\mathrm{Id}_{\mathcal{E}}) + a_1\Phi(f) + a_2\Phi(f)^2 + \cdots + a_n\Phi(f)^m =$$
$$\Phi(a_0\mathrm{Id}_{\mathcal{E}} + a_1 f + a_2 f^2 + \cdots + a_m f^m) = \Phi(P(f)) = M^{\mathcal{B}}_{\mathcal{B}}(P(f))$$

□

Πόρισμα 3.4. *Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας $n \times n$ πίνακας και και $P(t) \in \mathbb{K}[t]$ ένα πολυώνυμο. Αν $\mathcal{B}$ είναι η κανονική βάση του $\mathbb{K}_n$, και $f_A \colon \mathbb{K}_n \longrightarrow \mathbb{K}_n$, $f_A(X) = AX$, είναι η επαγόμενη γραμμική απεικόνιση, τότε:*

$$P(A) = M^{\mathcal{B}}_{\mathcal{B}}\big(P(f_A)\big)$$

Απόδειξη. Γνωρίζουμε ότι ο πίνακας της γραμμικής απεικόνισης f_A στην κανονική βάση $\mathcal{B}$ του $\mathbb{K}_n$ είναι ο A: $M^{\mathcal{B}}_{\mathcal{B}}(f_A) = A$. Τότε το συμπέρασμα προκύπτει άμεσα από το Λήμμα 3.3. □

Θεώρημα 3.5. *(Cayley-Hamilton) Έστω $\mathcal{E}$ ένας διανυσματικός χώρος πεπερασμένης διάστασης υπεράνω ενός σώματος $\mathbb{K}$. Έστω $f : \mathcal{E} \longrightarrow \mathcal{E}$ είναι μια γραμμική απεικόνιση, και $P_f(t)$ το χαρακτηριστικό πολυώνυμο της f. Τότε η γραμμική απεικόνσιση $P_f(f)$ είναι η μηδενική:*

$$P_f(f) = 0$$

Απόδειξη. Θεωρώντας τυχούσα βάση $\mathcal{B}$ του $\mathcal{E}$, και θέτοντας $P(t) = P_f(t)$ στο Λήμμα 3.3, έπεται ότι:

$$M^{\mathcal{B}}_{\mathcal{B}}\big(P_f(f)\big) = P_f\big(M^{\mathcal{B}}_{\mathcal{B}}(f)\big) \tag{3.5}$$

Όπως γνωρίζουμε, το χαρακτηριστικό πολυώνυμο μιας γραμμκής απεικόνισης συμπίπτει με το χαρακτηριστικό πολυώνυμο του πίνακά της σε τυχούσα βάση του $\mathcal{E}$:

$$P_f(t) = P_{M^{\mathcal{B}}_{\mathcal{B}}(f)}(t)$$

και επομένως:

$$P_f(M^{\mathcal{B}}_{\mathcal{B}}(f)) = P_{M^{\mathcal{B}}_{\mathcal{B}}(f)}(M^{\mathcal{B}}_{\mathcal{B}}(f)) \tag{3.6}$$

Απο το Θεώρημα των Cayley-Hamilton 3.2 για πίνακες έπεται ότι $P_{M^{\mathcal{B}}_{\mathcal{B}}(f)}(M^{\mathcal{B}}_{\mathcal{B}}(f)) = \mathbb{O}$. Επομένως $P_f(M^{\mathcal{B}}_{\mathcal{B}}(f)) = \mathbb{O}$, και τότε η σχέση (3.6) δείχνει ότι

$$M^{\mathcal{B}}_{\mathcal{B}}\big(P_f(f)\big) = \mathbb{O}$$

δηλαδή ο πίνακας της πολυωνυμικής γραμμικής απεικόνισης $P_f(f)$ στην βάση $\mathcal{B}$ είναι ο μηδενικός πίνακας. Επειδή η μόνη γραμμική απεικόνιση της οποίας ο πίνακας σε τυχούσα βάση του $\mathcal{E}$ είναι ο μηδενικός, έπεται ότι η γραμμική απεικόνιση $P_f(f)$ είναι η μηδενική:

$$P_f(f) = 0$$

□

3.3. **Μια άλλη απόδειξη τού Θεωρήματος Cayley-Hamilton.**

Πρόταση 3.6. *Έστω $A = (a_{ij}) \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας άνω τριγωνικός πίνακας. Τότε ο πίνακας $B = (A - a_{11}I_n)(A - a_{22}I_n)\dots(A - a_{nn}I_n)$ ισούται με τον μηδενικό $n \times n$ πίνακα, δηλαδή τον $\mathbb{O}_{n\times n}$.*

Απόδειξη. Για να δείξουμε ότι ο B είναι ο $\mathbb{O}_{n\times n}$, αρκεί να δείξουμε ότι

$$B \cdot \vec{e}_j = \vec{0}, \forall j, 1 \leq j \leq n,$$

όπου $\vec{e}_j$ είναι ο $n \times 1$ πίνακας, τού οποίου όλα τα στοιχεία είναι ίσα με μηδέν, εκτός από το στοιχείο τής j–οστής γραμμής το οποίο ισούται με 1, αφού το γινόμενο $B \cdot \vec{e}_j$ ισούται με την j–οστή στήλη τού B.

Πριν προχωρήσουμε παρατηρούμε ότι οποιοδήποτε γινόμενο

$$(A - a_{i_1i_1}I_n)(A - a_{i_2i_2}I_n)\dots(A - a_{i_ni_n}I_n),$$

όπου τα $a_{i_1i_1}, a_{i_2i_2}, \dots, a_{i_ni_n}$, είναι μια μετάθεση των $a_{11}, a_{22}, \dots a_{nn}$, δηλαδή είναι οι αριθμοί $a_{11}, a_{22}, \dots a_{nn}$, ενδεχόμενα με διαφορετική σειρά, ισούται και πάλι τον πίνακα B, αφού ο B ισούται με την τιμή $p(A)$ τού πολυωνύμου $p(t) = (t - a_{11})(t - a_{22})\dots(t - a_{nn})$, οι παράγοντες τού οποίου μετατίθενται, δηλαδή $p(t) = (t - a_{i_1i_1})(t - a_{i_2i_2})\dots(t - a_{i_ni_n})$.

Θα δείξουμε επαγωγικώς ότι

$$(A - a_{11}I_n)(A - a_{22}I_n)\dots(A - a_{ii}I_n) \cdot \vec{e}_j = \vec{0}, \forall j, 1 \leq j \leq i, \tag{*}$$

Για $i = 1$ έχουμε:

$$(A - a_{11}I_n)(\vec{e}_1 = A \cdot \vec{e}_1 - a_{11}\vec{e}_1) = a_{11}\vec{e}_1 - a_{11}\vec{e}_1 = \vec{0},$$

επειδή ο A είναι άνω τριγωνικός.
Έστω ότι η (∗) είναι αληθής για $i = \kappa$. Θα δείξουμε την (∗) για $i = \kappa + 1$.
Για κάθε $\vec{e}_j, 1 \leq j \leq \kappa$, έχουμε:

$$\begin{aligned}&(A - a_{11}I_n)(A - a_{22}I_n)\dots(A - a_{\kappa\kappa}I_n)(A - a_{\kappa+1,\kappa+1}I_n) \cdot \vec{e}_j = \\ &(A - a_{\kappa+1,\kappa+1}I_n)(A - a_{11}I_n)(A - a_{22}I_n)\dots(A - a_{\kappa\kappa}I_n) \cdot \vec{e}_j = \vec{0},\end{aligned}$$

αφού λόγω τής επαγωγικής υπόθεσης $(A - a_{11}I_n)(A - a_{22}I_n)\dots(A - a_{\kappa\kappa}I_n) \cdot \vec{e}_j = \vec{0}$, όταν $1 \leq j \leq \kappa$.
Υπολείπεται η απόδειξη για $j = \kappa + 1$.
Παρατηρούμε ότι

$$\begin{aligned}&(A - a_{\kappa+1,\kappa+1}I_n) \cdot \vec{e}_{\kappa+1,\kappa+1} = A \cdot \vec{e}_{\kappa+1,\kappa+1} - a_{\kappa+1,\kappa+1}\vec{e}_{\kappa+1} = \\ &a_{1,\kappa+1}\vec{e}_1 + a_{2,\kappa+1}\vec{e}_2 + \dots + a_{\kappa,\kappa+1}\vec{e}_\kappa + a_{\kappa+1,\kappa+1}\vec{e}_{\kappa+1} - a_{\kappa+1,\kappa+1}\vec{e}_{\kappa+1} = \\ &a_{1,\kappa+1}\vec{e}_1 + a_{2,\kappa+1}\vec{e}_2 + \dots + a_{\kappa,\kappa+1}\vec{e}_\kappa.\end{aligned}$$

Συνεπώς, το $(A - a_{\kappa+1,\kappa+1}I_n) \cdot \vec{e}_{\kappa+1,\kappa+1}$ είναι ένας $\mathbb{K}$–γραμμικός συνδυασμός των $\vec{e}_j, 1 \leq j \leq \kappa$. Αφού ο πίνακας $(A - a_{11}I_n)(A - a_{22}I_n)\dots(A - a_{\kappa\kappa}I_n)$, μηδενίζει τα $\vec{e}_j$, για κάθε $j, 1 \leq j \leq \kappa$, μηδενίζει και κάθε γραμμικό συνδυασμό τους. Έτσι, ο συγκεκριμένος πίνακας μηδενίζει και το $(A - a_{\kappa+1,\kappa+1}I_n) \cdot \vec{e}_{\kappa+1,\kappa+1}$. Ώστε,

$$(A - a_{11}I_n)(A - a_{22}I_n)\dots(A - a_{\kappa\kappa}I_n)(A - a_{\kappa+1,\kappa+1}I_n) \cdot \vec{e}_{\kappa+1,\kappa+1} = \vec{0}.$$

Επομένως, $B \cdot \vec{e}_j = \vec{0}, \forall j, 1 \leq j \leq n$. □

Παρατήρηση 3.7. *Παρατηρούμε ότι το πολυώνυμο $(t - a_{11})(t - a_{22})\dots(t - a_{nn})$ ισούται με το χαρακτηριστικό πολυώνυμο $P_A(t) = \mathrm{Det}(tI_n - A)$ τού άνω τριγωνικού πίνακα A. Συνεπώς, η ανωτέρω Πρόταση δηλώνει ότι $P_A(A) = (A - a_{11}I_n)(A - a_{22}I_n)\dots(A - a_{nn}I_n) = \mathbb{O}_{n\times n}$, που είναι το Θεώρημα Cayley-Hamilton, όταν ο A είναι ένας άνω τριγωνικός πίνακας. Ας δούμε πώς αποδεικνύεται η γενική περίπτωση:*

Θεώρημα 3.8. *(Cayley-Hamilton) Έστω A ένας $n \times n$ πίνακας με συνιστώσες από το σώμα $\mathbb{K}$, όπου $\mathbb{K} = \mathbb{C}$ ή $\mathbb{R}$. Αν $P_A(t)$ είναι το χαρακτηριστικό πολυώνυμο τού A, τότε $P_A(A) = \mathbb{O}_{n\times n}$.*

Απόδειξη. Το χαρακτηριστικό πολυώνυμο $P_A(t)$ τού A ισούται με το χαρακτηριστικό πολυώνυμο $P_B(t)$ οποιουδήποτε πίνακα που είναι όμοιος με τον A. Όλες οι ρίζες τού $P_A(t)$ ανήκουν στο σώμα $\mathbb{C}$ και γι' αυτό ο A είναι όμοιος με έναν άνω τριγωνικό πίνακα B με πιθανόν μιγαδικές συνιστώσες, σύμφωνα με το Θεώρημα 2.2. Ας πούμε λοιπόν ότι $B = S^{-1}AS$, όπου ο $S \in \mathsf{M}_{n\times n}(\mathbb{C})$ είναι ένας αντιστρέψιμος πίνακας. Αλλά για τον άνω τριγωνικό πίνακα B γνωρίζουμε ότι $P_B(B) = \mathbb{O}_{n\times n}$, λόγω τής Πρότασης 3.6, και αφού $P_A(t) = P_B(t)$, έχουμε $P_A(B) = \mathbb{O}_{n\times n}$. Επιπλέον είναι:

$$P_A(A) = P_A(SBS^{-1}) = SP_A(B)S^{-1} = S\mathbb{O}_{n\times n}S^{-1} = \mathbb{O}_{n\times n}, \tag{**}$$

□

Παρατήρηση 3.9. *Σχετικά με τον υπολογισμό που εκτελέσαμε στην (**) ισχύει το γενικότερο:*

Αν, $P(t) = a_0 + a_1t + a_2t \cdots + a_mt^m$ είναι οποιοδήποτε πολυώνυμο και $A, S^{-1}AS$ είναι δύο όμοιοι τετραγωνικοί $n \times n$ πίνακες, τότε η τιμή $P(S^{-1}AS)$ τού $P(t)$ στον πίνακα $S^{-1}AS$ ισούται με $S^{-1}P(A)S$:

$$P(S^{-1}AS) = S^{-1}P(A)S$$

Πραγματικά:

$$\begin{aligned} P(S^{-1}AS) = a_0I_n + a_1(S^{-1}AS) + a_2(S^{-1}AS)^2 + \cdots + a_m(S^{-1}AS)^m &= \\ a_0S^{-1}S + a_1(S^{-1}AS) + a_2S^{-1}A^2S + \cdots + a_mS^{-1}A^mS &= \\ S^{-1}(a_0I_n + a_1A + a_2A^2 + \cdots + a_mA^m)S = S^{-1}P(A)S. \end{aligned}$$

4. **Ελάχιστο Πολυώνυμο**

Στην παρούσα παράγραφο θε δείξουμε ένα πολύ χρήσιμο κριτήριο διαγωνοποιησιμότητας για γραμμικές απεικονίσεις και πίνακες.

4.1. **Πυρήνες Πολυωνυμικών Γραμμικών Απεικονίσεων.** Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης και $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση.

Παρατήρηση 4.1. *Όπως γνωρίζουμε, αν $f, g\colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι δύο γραμμικές απεικονίσεις, τότε γενικά ισχύει ότι: $f \circ g \neq g \circ f$. Αν όμως η g προκύπτει ως πολυωνυμική απεικόνιση από την f, δηλαδή $g = P(f)$ για κάποιο πολυώνυμο $P(t) \in \mathbb{K}[t]$, είναι εύκολο να δει κανείς ότι οι f και g μετατίθενται: $f \circ g = g \circ f$.*

Γενικότερα ισχύει ότι (η απόδειξη είναι εύκολη και αφήνεται ως άσκηση):

$$F(f) \circ G(f) = G(f) \circ F(f), \qquad \forall\, F(t),\ G(t)\ \in\ \mathbb{K}[t]$$

Για κάθε πολυώνυμο $P(t) \in \mathbb{K}[t]$, θεωρούμε την πολυωνυμική γραμμική απεικόνιση

$$P(f):\ \mathcal{E}\ \longrightarrow\ \mathcal{E},\ \ \vec{x}\ \longmapsto\ P(f)(\vec{x})$$

και ορίζουμε έναν υπόχωρο του $\mathcal{E}$, τον *πυρήνα της πολυωνυμικής γραμμικής απεικόνισης* $P(t)$, ως εξής:

$$\mathsf{N}_{P(t)} := \mathsf{Ker}P(f) = \big\{\vec{x} \in \mathcal{E} \mid P(f)(\vec{x}) = \vec{0}\big\}$$

Στην απόδειξη του κριτηρίου διαγωοποιησιμότητας θα χρησιμοποιήσουμε την ακόλουθη χρήσιμη πρόταση.

Πρόταση 4.2. *Έστω $P(t) \in \mathbb{K}[t]$ ένα πολυώνυμο έτσι ώστε $P(f) = 0$, και υποθέτουμε ότι:*

$$P(t) = (t-\kappa_1)(t-\kappa_2)\cdots(t-\kappa_m), \qquad \kappa_i \neq \kappa_j,\ \ 1 \leq i \neq j \leq m$$

1. $\mathsf{N}_{t-\kappa_i} = \{\vec{x} \in \mathcal{E} \mid f(\vec{x}) = \kappa_i\vec{x}\},\ 1 \leq i \leq m.$
2. $\mathcal{E} = \mathsf{N}_{t-\kappa_1} \oplus \mathsf{N}_{t-\kappa_2} \oplus \cdots \oplus \mathsf{N}_{t-\kappa_m}.$

Απόδειξη. **1.** Για κάθε $1 \leq i \leq m$:

$$\mathsf{N}_{t-\kappa_i} = \{\vec{x} \in \mathcal{E} \mid (f-\kappa_i\mathrm{Id}_{\mathcal{E}})(\vec{x}) = \vec{0}\} = \{\vec{x} \in \mathcal{E} \mid f(\vec{x}) - \kappa_i\mathrm{Id}_{\mathcal{E}}(\vec{x}) = \vec{0}\} =$$
$$\{\vec{x} \in \mathcal{E} \mid f(\vec{x}) - \kappa_i\vec{x} = \vec{0}\} = \{\vec{x} \in \mathcal{E} \mid f(\vec{x}) = \kappa_i\vec{x}\}$$

2. Επειδή $\kappa_i \neq \kappa_j,\ \ 1 \leq i \neq j \leq m$, έπεται ότι τα πολυώνυμα $t-\kappa_1, t-\kappa_2, \cdots, t-\kappa_m$ είναι πρώτα μεταξύ τους, δηλαδή $\mathsf{MK\Delta}\big(t-\kappa_i, t-\kappa_j\big) = 1$, για $1 \leq i \neq j \leq m$. Επομένως θέτοντας

$$F_i(t) = \frac{P(t)}{t-\kappa_i} = (t-\kappa_1)(t-\kappa_2)\cdots(t-\kappa_{i-1})(t-\kappa_{i+1})\cdots(t-\kappa_m)$$

έπεται ότι ο μέγιστος κοινός διαιρέτης των πολυωνύμων $F_1(t), F_2(t), \cdots, F_m(t)$ είναι το σταθερό πολυώνυμο 1: $\mathsf{MK\Delta}\big(F_1(t), \cdots, F_m(t)\big) = 1$. Από την Θεωρία Πολυωνύμων γνωρίζουμε τότε ότι υπάρχουν πολυώνυμα $G_1(t), \cdots G_m(t)$, έτσι ώστε:

$$G_1(t)F_1(t) + G_2(t)F_2(t) + \cdots + G_m(t)F_m(t) = 1$$

Τότε όμως για τις αντίστοιχες πολυωνυμικές απεικονίσεις θα έχουμε:

$$G_1(f)\circ F_1(f) + G_2(f)\circ F_2(f) + \ \cdots\ + G_m(f)\circ F_m(f) = \mathrm{Id}_{\mathcal{E}}$$

και επομένως για κάθε $\vec{x} \in \mathcal{E}$ θα έχουμε:

$$G_1(f)(F_1(f)(\vec{x})) + G_2(f)(F_2(f)(\vec{x})) + \ \cdots\ + G_m(f)(F_m(f)(\vec{x})) = \mathrm{Id}_{\mathcal{E}}(\vec{x}) = \vec{x} \tag{4.1}$$

Θέτουμε:

$$\vec{x}_i = G_i(f)(F_i(f)(\vec{x})), \quad \forall \vec{x} \in \mathcal{E}, \qquad 1 \leq i \leq m$$

και άρα η σχέση (4.1) γράφεται:

$$\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_m = \vec{x}, \quad \forall \vec{x} \in \mathcal{E} \tag{4.2}$$

Ισχυρισμός: $\vec{x}_i \in \mathsf{N}_{t-\kappa_i}$.

Πραγματικά: χρησιμοποιώντας την Παρατήρηση 4.1 θα έχουμε $(f-\kappa_i \mathrm{Id}_{\mathcal{E}}) \circ G_i(f) \circ F_i(f) = G_i(f) \circ (f-\kappa_i \mathrm{Id}_{\mathcal{E}}) \circ F_i(f) = G_i(f) \circ F_i(f) \circ (f-\kappa_i \mathrm{Id}_{\mathcal{E}})$. Όμως

$$F_i(f) \circ (f-\kappa_i \mathrm{Id}_{\mathcal{E}}) =$$
$$= (f-\kappa_1 \mathrm{Id}_{\mathcal{E}})(f-\kappa_2 \mathrm{Id}_{\mathcal{E}}) \cdots (f-\kappa_{i-1} \mathrm{Id}_{\mathcal{E}})(f-\kappa_{i+1} \mathrm{Id}_{\mathcal{E}}) \cdots (f-\kappa_m \mathrm{Id}_{\mathcal{E}}) \circ (f-\kappa_i \mathrm{Id}_{\mathcal{E}}) =$$
$$= (f-\kappa_1 \mathrm{Id}_{\mathcal{E}})(f-\kappa_2 \mathrm{Id}_{\mathcal{E}}) \cdots (f-\kappa_{i-1} \mathrm{Id}_{\mathcal{E}}) \circ (f-\kappa_i \mathrm{Id}_{\mathcal{E}}) \circ (f-\kappa_{i+1} \mathrm{Id}_{\mathcal{E}}) \cdots (f-\kappa_m \mathrm{Id}_{\mathcal{E}}) = P(f)$$

Επομένως

$$(f-\kappa_i \mathrm{Id}_{\mathcal{E}})(\vec{x}_i) = (f-\kappa_i \mathrm{Id}_{\mathcal{E}})\big(G_i(f)(F_i(f)\big)(\vec{x})) = (f-\kappa_i \mathrm{Id}_{\mathcal{E}}) \circ G_i(f) \circ F_i(f))(\vec{x}) =$$
$$= (G_i(f) \circ (f-\kappa_i \mathrm{Id}_E) \circ F_i(f))(\vec{x}) = (G_i(f) \circ P(f))(\vec{x}) =$$
$$= G_i(f)(P(f)(\vec{x})) = \vec{0}$$

διότι $P(f)=0$. Άρα $(f-\kappa_1 \mathrm{Id}_{\mathcal{E}})(\vec{x}_i) = 0$ το οποίο σημαίνει ότι

$$f(\vec{x}_i) = \kappa_i(\vec{x}_i) \quad \Longrightarrow \quad \vec{x}_i \in \mathsf{N}_{t-\kappa_i}$$

Επομένως συνδυάζοντας την παραπάνω σχέση με την (4.2) θα έχουμε ότι:

$$\mathcal{E} = \mathsf{N}_{t-\kappa_1} + \mathsf{N}_{t-\kappa_2} + \cdots + \mathsf{N}_{t-\kappa_m}$$

Μένει να δείξουμε ότι το παραπάνω άθροισμα είναι ευθύ. Ως γνωστόν αρκεί να δείξουμε ότι:

$$\vec{x}_1 + \cdots + \vec{x}_m = \vec{0} \quad \text{και} \quad \vec{x}_i \in \mathsf{N}_{t-\kappa_i} \quad 1 \le i \le m \quad \Longrightarrow \quad \vec{x}_i = \vec{0}, \ \ 1 \le i \le m$$

Υποθέτουμε ότι κάποια από τα $\vec{x}_1, \cdots, \vec{x}_m$ είναι μη-μηδενικά. Χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι τα διανύσματα $\vec{x}_1, \cdots, \vec{x}_\rho$ είναι μη-μηδενικά και $\vec{x}_{\rho+1} = \cdots = \vec{x}_m = \vec{0}$. Τότε επειδή $\vec{x}_i \in \mathsf{N}_{t-\kappa_i})$, $1 \le i \le \rho$, από το **1.** έπεται ότι τα $\vec{x}_1, \cdots, \vec{x}_\rho$ είναι ιδιοδιανύσματα της f τα οποία αντιστοιχούν στις ιδιοτιμές $\kappa_1, \cdots, \kappa_\rho$. Επειδή οι ιδιοτιμές $\kappa_1, \cdots, \kappa_\rho$ είναι ανα δύο διαφορετικές, γνωρίζουμε ότι τα διανύσματα $\vec{x}_1, \cdots, \vec{x}_\rho$ είναι γραμμικά ανεξάρτητα το οποίο είναι άτοπο διότι $\vec{x}_1 + \cdots + \vec{x}_\rho = \vec{0}$. Στο άτοπο καταλήξαμε διότι υποθέσαμε ότι κάποια από τα $\vec{x}_1, \cdots, \vec{x}_m$ είναι μη-μηδενικά. Άρα θα έχουμε $\vec{x}_1 = \cdots = \vec{x}_m = \vec{0}$. Επομένως το άθροισμα $\mathsf{N}_{t-\kappa_1} + \mathsf{N}_{t-\kappa_2} + \cdots + \mathsf{N}_{t-\kappa_m}$ είναι ευθύ και άρα:

$$\mathcal{E} = \mathsf{N}_{t-\kappa_1} \oplus \mathsf{N}_{t-\kappa_2} \oplus \ \cdots \ \oplus \mathsf{N}_{t-\kappa_m} \qquad \square$$

Παρατήρηση 4.3. *Η παραπάνω Πρόταση είναι ειδική περίπτωση του ακόλουθου Θεωρήματος Πρωταρχικής Ανάλυσης. Πρώτα υπενθυμίζουμε ότι ένα πολυώνυμο $P(t)$ καλείται ανάγωγο αν $P(t) = P_1(t)P_2(t)$ έπεται ότι είτε το $P_1(t)$ ή το $P_2(t)$ είναι σταθερό πολυώνυμο. Από την Θεωρία Πολυωνύμων γνωρίζουμε ότι ισχύει το ακόλουθο σημαντικό Θεώρημα:*

Θεώρημα: *Κάθε μη-σταθερό πολυώνυμο $P(t)$ γράφεται κατά μοναδικό τρόπο ως γινόμενο αναγώγων πολυωνύμων κατά μοναδικό τρόπο. Δηλαδή υπάρχουν μοναδικά κανονικά ανάγωγα πολυώνυμα $P_1(t), \cdots,$ $P_k(t)$, και μοναδικό $\alpha \in \mathbb{K}$, έτσι ώστε:*

$$P(t) = \alpha P_1(t) P_2(t) \cdots P_k(t)$$

και η παραπάνω γραφή είναι μοναδική (μη-λαμβάνοντας υπ' όψιν τη σειρά των παραγόντων).

Θεώρημα Πρωταρχικής Ανάλυσης: *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης και $f : \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμικής απεικόνιση. Υποθέτουμε ότι*

$$Q_f(t) = Q_1(t)^{\alpha_1} Q_2(t)^{\alpha_2} \cdots Q_m(t)^{\alpha_m}$$

είναι η ανάλυση του ελαχίστου πολυωνύμου $Q_f(t)$ της f σε διακεκριμένα ανάγωγα κανονικά πολυώνυμα. Τότε:

$$\mathcal{E} = \mathsf{N}_{Q_1(t)^{\alpha_1}} \oplus \mathsf{N}_{Q_2(t)^{\alpha_2}} \oplus \cdots \oplus \mathsf{N}_{Q_m(t)^{\alpha_m}}$$

Η απόδειξη είναι παρόμοια με την απόδειξη της Πρότασης 4.2 και αφήνεται ως Άσκηση.

4.2. **Κριτήριο Διαγωνοποίησης.** Μπορούμε να αποδείξουμε τώρα το ακόλουθο κριτήριο διαγωνοποιησιμότητας:

Θεώρημα 4.4. *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης και $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση. Τότε τα ακόλυθα είναι ισοδύναμα:*

(1) *Η f είναι διαγωνοποιήσιμη.*

(2) *Το ελάχιστο πολυώνυμο $Q_f(t)$ της f αναλύεται σε γινόμενο διακεκριμένων πρωτοβαθμίων παραγόντων:*

$$Q_f(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_m), \qquad \lambda_i \neq \lambda_j,\ \ 1 \leq i \neq j \leq m$$

Απόδειξη. (1) $\Longrightarrow$ (2) Θεωρούμε το πολυώνυμο $Q(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_m)$, όπου λ_i, $1 \leq i \leq m$ είναι οι διακεκριμένες ιδιοτιμές της f. Θα δείξουμε ότι $Q(t) = Q_f(t)$. Για αυτό το σκοπό, δείχνουμε πρώτα ότι το πολυώνυμο $Q(t)$ μηδενίζει την f: $Q(f) = 0$.

Επειδή η f είναι διαγωνοποιήσιμη, ο χώρος $\mathcal{E}$ θα είναι το ευθύ άθροισμα των ιδιοχώρων της f:

$$\mathcal{E} = \mathcal{V}(\lambda_1) \oplus \mathcal{V}(\lambda_2) \oplus \cdots \oplus \mathcal{V}(\lambda_m) \tag{4.3}$$

Ισχυρισμός: $Q(f)(\vec{x}_i) = \vec{0},\ \forall \vec{x}_i \in \mathcal{V}(\lambda_i),\ 1 \leq i \leq m.$

Πραγματικά επειδή $\vec{x}_i \in \mathcal{V}(\lambda_i) \implies f(\vec{x}_i) = \lambda_i \vec{x}_i$, δηλαδή $(f - \lambda_i \mathrm{Id}_{\mathcal{E}})(\vec{x}_i) = \vec{0}$, και επειδή, λόγω της Παρατήρησης 4.1, οι γραμμικές απεικονίσεις $(f - \lambda_i \mathrm{Id}_{\mathcal{E}})$ και $(f - \lambda_j \mathrm{Id}_{\mathcal{E}})$ μετατίθενται, θα έχουμε:

$$Q(f)(\vec{x}_i) = [(f-\lambda_1\mathrm{Id}_{\mathcal{E}}) \circ (f-\lambda_2\mathrm{Id}_{\mathcal{E}}) \circ \ \cdots \ \circ (f-\lambda_m\mathrm{Id}_{\mathcal{E}})](\vec{x}_i) =$$
$$= [(f-\lambda_1\mathrm{Id}_{\mathcal{E}}) \circ \cdots \circ (f-\lambda_{i-1}\mathrm{Id}_{\mathcal{E}}) \circ (f-\lambda_i\mathrm{Id}_{\mathcal{E}}) \circ (f-\lambda_{i+1}\mathrm{Id}_{\mathcal{E}}) \circ \ \cdots \ \circ (f-\lambda_m\mathrm{Id}_{\mathcal{E}})](\vec{x}_i) =$$
$$= [(f-\lambda_1\mathrm{Id}_{\mathcal{E}}) \circ \cdots \circ (f-\lambda_{i-1}\mathrm{Id}_{\mathcal{E}}) \circ (f-\lambda_{i+1}\mathrm{Id}_{\mathcal{E}}) \circ \ \cdots \ \circ (f-\lambda_m\mathrm{Id}_{\mathcal{E}}) \circ (f-\lambda_i\mathrm{Id}_{\mathcal{E}})](\vec{x}_i) =$$
$$= [(f-\lambda_1\mathrm{Id}_{\mathcal{E}}) \circ \cdots \circ (f-\lambda_{i-1}\mathrm{Id}_{\mathcal{E}}) \circ (f-\lambda_{i+1}\mathrm{Id}_{\mathcal{E}}) \circ \ \cdots \ \circ (f-\lambda_m\mathrm{Id}_{\mathcal{E}})](f-\lambda_i\mathrm{Id}_{\mathcal{E}})(\vec{x}_i) =$$
$$= [(f-\lambda_1\mathrm{Id}_{\mathcal{E}}) \circ \cdots \circ (f-\lambda_{i-1}\mathrm{Id}_{\mathcal{E}}) \circ (f-\lambda_{i+1}\mathrm{Id}_{\mathcal{E}}) \circ \ \cdots \ \circ (f-\lambda_m\mathrm{Id}_{\mathcal{E}})](\vec{0}) = \vec{0}$$

Επομένως δείξαμε ότι $Q(f)(\vec{x}_i) = \vec{0}$, $\forall \vec{x}_i \in \mathcal{V}(\lambda_i)$, $1 \leq i \leq m$.

Επειδή το άθροισμα (4.3) είναι ευθύ, κάθε διάνυσμα $\vec{x} \in \mathcal{E}$ έχει μοναδική γραφή:

$$\vec{x} = \vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_m, \qquad \vec{x}_i \in \mathcal{V}(\lambda_i),\ \ 1 \leq i \leq m$$

Τότε:

$$Q(f)(\vec{x}) = Q(f)(\vec{x}_1) + Q(f)(\vec{x}_2) + \cdots + Q(f)(\vec{x}_m) = \vec{0}$$

και άρα η γραμμική απεικόνιση $Q(f)$ είναι η μηδενική:

$$Q(f) = 0$$

διότι μηδενίζει κάθε διάνυσμα του $\mathcal{E}$. Τότε όμως $Q_f(t)/Q(t)$. Επειδή τα πολυώνυμα $Q_f(t)$ και $Q(t)$ είναι κανονικά και κάθε ρίζα του $Q(t)$ (δηλαδή μια ιδιοτιμή της f) είναι και ρίζα του $Q_f(t)$, έπεται ότι τα πολυώνυμα $Q_f(t)$ και $Q(t)$ συμπίπτουν: $Q_f(t) = Q(t)$.

(2) $\Longrightarrow$ (1) Επειδή τα $\lambda_1, \cdots, \lambda_m$ είναι ανα δύο διαφορετικά, σύμφωνα με την Πρόταση 4.1 θα έχουμε:

$$\mathcal{E} = \mathsf{N}_{t-\lambda_1} \oplus \mathsf{N}_{t-\lambda_2} \oplus \cdots \oplus \mathsf{N}_{t-\lambda_m}$$

Επειδή οι ρίζες του ελαχίστου πολυωνύμου συμπίπτουν με τις ρίζες του χαρακτηριστικού πολυωνύμου της f, δηλαδή με τις ιδιοτιμές της f, έπεται ότι τα $\lambda_1, \cdots, \lambda_m$ είναι οι ιδιοτιμές της f με αντίστοιχους ιδιοχώρους $\mathsf{N}_{t-\lambda_1}, \mathsf{N}_{t-\lambda_2}, \cdots, \mathsf{N}_{t-\lambda_m}$. Δηλαδή:

$$\mathcal{E} = \mathcal{V}(\lambda_1) \oplus \mathcal{V}(\lambda_2) \oplus \cdots \oplus \mathcal{V}(\lambda_m)$$

Επομένως ο χώρος $\mathcal{E}$ είναι το ευθύ άθροισμα των ιδιοχώρων της f και τότε όπως γνωρίζουμε η f είναι διαγωνοποιήσιμη. $\square$

Πριν περάσουμε να δούμε το αντίστοιχο του παραπάνω κριτηρίου για πίνακες, χρειαζόμαστε το ακόλουθο Λήμμα:

Λήμμα 4.5. *Έστω $\mathcal{E}$ ένας διανυσματικός χώρος πεπερασμένης διάστασης υπεράνω ενός σώματος $\mathbb{K}$ και $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση. Τότε το ελάχιστο πολυώνυμο της f συμπίπτει με το ελάχιστο πολυώνυμο του πίνακά της σε μια τυχούσα βάση $\mathcal{B}$ του $\mathcal{E}$:*

$$Q_f(t) = Q_{M^{\mathcal{B}}_{\mathcal{B}}(f)}(t)$$

Απόδειξη. Έστω $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$. Τότε θέτοντας $P(t) = Q_f(t)$ στο Λήμμα 3.3, θα έχουμε:

$$M^{\mathcal{B}}_{\mathcal{B}}(Q_f(f)) = Q_f(M^{\mathcal{B}}_{\mathcal{B}}(f)) = Q_f(A)$$

Επειδή $Q_f(f) = 0$ και ο πίνακας της μηδενικής γραμμικής απεικόνισης ως προς τυχούσα βάση είναι ο μηδενικός, έπεται ότι $Q_f(A) = \mathbb{O}$, δηλαδή το ελάχιστο πολυώνυμο της f μηδενίζει τον πίνακα A. Τότε όμως θα έχουμε:

$$Q_A(t)/Q_f(t) \tag{4.4}$$

Από την άλλη πλευρά, θέτοντας $P(t) = Q_A(t)$ στο Λήμμα 3.3, θα έχουμε:

$$M^{\mathcal{B}}_{\mathcal{B}}(Q_A(f)) = Q_A(M^{\mathcal{B}}_{\mathcal{B}}(f)) = Q_A(A)$$

Επειδή $Q_A(A) = \mathbb{O}$ και επειδή η μοναδική γραμμική απεικόνιση με μηδενικό πίνακα σε τυχούσα βάση είναι η μηδενική, έπεται ότι η γραμμική απεικόνιση $Q_A(f)$ είναι η μηδενική, δηλαδή το ελάχιστο πολυώνυμο του A μηδενίζει την γραμμική απεικόνιση f. Τότε όμως θα έχουμε:

$$Q_f(t)/Q_A(t) \tag{4.5}$$

Επειδή τα πολυώνυμα $Q_A(t)$ και $Q_f(t)$ είναι κανονικά, από τις (4.4) και (4.5) έπεται ότι $Q_A(t) = Q_f(t)$, δηλαδή: $Q_{M^{\mathcal{B}}_{\mathcal{B}}(f)}(t) = Q_f(t)$. □

Θεώρημα 4.6. *Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας $n \times n$ πίνακας. Τότε τα ακόλουθα είναι ισοδύναμα:*

(1) *Ο A είναι διαγωνοποιήσιμος.*
(2) *Το ελάχιστο πολυώνυμο $Q_A(t)$ του A αναλύεται σε γινόμενο διακεκριμένων πρωτοβαθμίων παραγόντων:*

$$Q_A(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_m), \qquad \lambda_i \neq \lambda_j,\ \ 1 \leq i \neq j \leq m$$

Απόδειξη. Θεωρούμε την γραμμική απεικόνιση $f_A\colon \mathbb{K}_n \longrightarrow \mathbb{K}_n$, $f_A(X) = AX$. Έστω $\mathcal{B}$ η κανονική βάση του $\mathbb{K}_n$. Τότε γνωρίζουμε ότι $M^{\mathcal{B}}_{\mathcal{B}}(f_A) = A$. Από το Λήμμα 3.3 θα έχουμε τότε:

$$Q_A(t) = Q_{f_A}(t)$$

Επειδή, όπως γνωρίζουμε, ο πίνακας A είναι διαγωνοποιήσιμος αν και μονον αν η γραμμική απεικόνιση f_A είναι διαγωνοποιήσιμη, το ζητούμενο προκύπτει από την παραπάνω ισότητα ελαχίστων πολυωνύμων σε συνδυασμό με το Θεώρημα 4.4. □

Παράδειγμα 4.7. *Θεωρούμε τον 4×4 πίνακα παργματικών αριθμών*

$$\begin{pmatrix} 0 & 4 & 1 & -2 \\ -1 & 4 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ -1 & 3 & 0 & 0 \end{pmatrix}$$

Υπολογίζουμε το ελάχιστο πολυώνυμο του πίνακα A:

$$Q_A(t) = t^3 - 4t^2 + 5t - 2$$

Επειδή όπως βλέπουμε εύκολα

$$Q_A(t) = t^3 - 4t^2 + 5t - 2 = (t-1)^2(t-2)$$

το ελάχιστο πολυώνυμο του A δεν αναλύεται σε γινόμενο διακεκριμένων πρωτοβαθμίων παραγόντων. Επομένως ο πίνακας A δεν είναι διαγωνοποιήσιμος.

Γνωρίζουμε ότι μια γραμμική απεικόνιση, αντίστοιχα ένας πίνακας, είναι τριγωνοποιήσιμη, αντίστοιχα τριγωνοποιήσιμος, αν και μόνον αν το χαρακτηριστικό της, αντίστοιχα το χαρακτηριστικό του, το πολυώνυμο αναλύεται σε γινόμενο πρωτοβαθμίων παραγόντων υπεράνω του $\mathbb{K}$. Όπως γνωρίζουμε το χαρακτηριστικό και το ελάχιστο πολυώνυμο έχουν τις ίδιες ρίζες. Σημειώνουμε εδώ ότι μπορούμε να περιορισθούμε στο ελάχιστο πολυώνυμο, όπως δείχνει η παρακάτω εφαρμογή.

Πόρισμα 4.8. (1) *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος και $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση. Τότε τα ακόλουθα είναι ισοδύναμα:*
- (α') *Η γραμμική απεικόνιση f είναι τριγωνοποιήσιμη.*
- (β') *Το ελάχιστο πολυώνυμο $Q_f(t)$ της f έχει όλες τις ρίζες του στο σώμα $\mathbb{K}$.*

(2) *$A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας $n \times n$ πίνακας. Τότε τα ακόλουθα είναι ισοδύναμα:*
- (α') *Ο πίνακας A είναι τριγωνοποιήσιμος.*
- (β') *Το ελάχιστο πολυώνυμο $Q_A(t)$ του A έχει όλες τις ρίζες του στο σώμα $\mathbb{K}$.*

Απόδειξη. Προκύπτει άμεσα από το γεγονός ότι το ελάχιστο πολυώνυμο διαιρεί, και έχει τις ιδιες ρίζες με, το χαρακτηριστικό πολυώνυμο. □

Άσκηση 4.9. *Επειδή όλες οι ιδιοτιμές του πίνακα A του Παραδείγματος 4.7 ανήκουν στο $\mathbb{R}$, ο πίνακας A είναι τριγωνοποιήσιμος. Να βρεθεί η άνω τριγωνική μορφή του πίνακα A.*

4.3. **Μηδενοδύναμοι Ενδομορφισμοί και Πίνακες.** Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος και $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση. Έστω επίσης $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας τετραγωνικός πίνακας.

Ορισμός 4.10. (1) *Η γραμμική απεικόνιση καλείται μηδενοδύναμη αν: $f^m = 0$, για κάποιο $m \geq 1$.*
(2) *Ο πίνακας A καλείται μηδενοδύναμος αν: $A^m = 0$, για κάποιο $m \geq 1$.*

Πρόταση 4.11. *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος διάστασης $\dim_{\mathbb{K}}\mathcal{E} = n$. και $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση. Τότε τα ακόλουθα είναι ισοδύναμα:*
1. *Η f είναι μηδενοδύναμη: $f^m = 0$, για κάποιο $m \geq 1$.*
2. *Η μόνη ιδιοτιμή της f είναι η μηδενική.*
3. *$f^n = 0$.*
4. *Το ελάχιστο πολυώνυμο $Q_f(t)$ είναι της μορφής: $Q_f(t) = t^m$, για κάποιο $m \geq 1$.*

Απόδειξη. (1) $\Longrightarrow$ (2) Αν $f^m = 0$, τότε το πολυώνυμο $Q(t) = t^m$ μηδενίζει την f και άρα διαιρείται από το ελάχιστο πολυώνυμο $Q_f(t)$. Προφανώς τότε το ελάχιστο πολυώνυμο της f είναι της μορφής $Q_f(t) = t^k$ του οποίου η μοναδική ρίζα είναι η $\lambda = 0$. Επειδή οι ρίζες του ελαχίστου πολυωνύμου είναι οι ιδιοτιμές της f έπεται ότι η μόνη ιδιοτιμή της f είναι η μηδενική.

(2) $\Longrightarrow$ (3) Αν η μόνη ιδιοτιμή της f είναι η μηδενική, τότε προφανώς το ελάχιστο πολυώνυμο της f θα είναι της μορφής $Q_f(t) = t^m$ για κάποιο $1 \leq m \leq n$. Επειδή το χαρακτηριστικό και το ελάχιστο πολυώνυμο έχουν τις ίδιες ρίζες, έπεται ότι $P_f(t) = (-1)^n t^n$. Από το Θεώρημα των Cayley-Hamilton θα έχουμε: $0 = P_f(f) = (-1)^n f^n$ και άρα $f^n = 0$.

(3) $\Longrightarrow$ (4) Αν $f^n = 0$, τότε το πολυώνυμο $Q(t) = t^n$ μηδενίζει την f και άρα διαιρείται το από το ελάχιστο πολυώνυμο $Q_f(t)$ της f. Προφανώς τότε $Q_f(t) = t^m$ για κάποιο $1 \leq m \leq n$.

(4) $\Longrightarrow$ (1) Αν το ελάχιστο πολυώνυμο $Q_f(t)$ είναι της μορφής $Q_f(t) = t^m$, για κάποιο $m \geq 1$, τότε $0 = Q_f(f) = f^m$. □

Χρησιμοποιώντας την γραμμική απεικόνιση $f_A\colon \mathbb{K}_n \longrightarrow \mathbb{K}_n,\ f_A(X) = AX$ και την Πρόταση 4.11 έχουμε το ακόλουθο πόρισμα:

Πόρισμα 4.12. *Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας τετραγωνικός πίνακας. Τότε τα ακόλουθα είναι ισοδύναμα:*

(1) *Ο A είναι μηδενοδύναμος: $A^m = 0$, για κάποιο $m \geq 1$.*
(2) *Η μόνη ιδιοτιμή του A είναι η μηδενική.*
(3) $A^n = 0$.
(4) *Το ελάχιστο πολυώνυμο $Q_A(t)$ είναι της μορφής: $Q_A(t) = t^m$, για κάποιο $m \geq 1$.*

5. **Κανονική Μορφή Fitting**

Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας πίνακας. Σκοπός μας στην παρούσα παράγραφο είναι η απόδειξη του ακόλουθου Θεωρήματος:

Θεώρημα 5.1. *Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας πίνακας. Τότε ο A είναι όμοιος με έναν πίνακα της μορφής*

$$B = \begin{pmatrix} \boxed{N} & \mathbb{O} \\ \mathbb{O} & \boxed{P} \end{pmatrix}$$

όπου ο P είναι αντιστρέψιμος και ο N είναι μηδενοδύναμος.

Η παραπάνω μορφή καλείται *μορφή Fitting του πίνακα A.*

Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης υπεράνω του σώματος $\mathbb{K}$, και

$$f : \mathcal{E} \longrightarrow \mathcal{E}$$

μια γραμμική απεικόνιση.

Συμβολίζουμε με $f^k = f \circ f \circ \cdots \circ f$ (k-φορές) την σύνθεση της f με τον εαυτό της k-φορές, $\forall k \geq 1$. Υπενθυμίζουμε ότι η f καλείται *μηδενοδύναμη* αν $f^k = 0$.

5.1. **Αποσύνθεση Fitting.** Για να αποδείξουμε το Θεώρημα 5.1 χρειαζόμαστε πρώτα κάποια προεργασία.

Πρόταση 5.2. (1) *Υπάρχει μια (αύξουσα) ακολουθία υπόχωρων του $\mathcal{E}$:*

$$\{\vec{0}\} \subseteq \mathsf{Ker}(f) \subseteq \mathsf{Ker}(f^2) \subseteq \cdots \subseteq \mathsf{Ker}(f^k) \subseteq \mathsf{Ker}(f^{k+1}) \subseteq \cdots \subseteq \mathcal{E} \tag{5.1}$$

και φυσικός αριθμός $\mu \geq 0$ έτσι ώστε: $\mathsf{Ker}(f^\mu) = \mathsf{Ker}(f^{\mu+1}) = \mathsf{Ker}(f^{\mu+2}) = \cdots$.

(2) *Υπάρχει μια (φθίνουσα) ακολουθία υποχώρων του $\mathcal{E}$:*

$$\{\vec{0}\} \subseteq \cdots \subseteq \mathsf{Im}(f^{k+1}) \subseteq \mathsf{Im}(f^k) \subseteq \cdots \subseteq \mathsf{Im}(f^2) \subseteq \mathsf{Im}(f) \subseteq \mathcal{E} \tag{5.2}$$

και φυσικός αριθμός $\lambda \geq 0$ έτσι ώστε: $\mathsf{Im}(f^\lambda) = \mathsf{Im}(f^{\lambda+1}) = \mathsf{Im}(f^{\lambda+2}) = \cdots$.

Απόδειξη. (1) Δείχνουμε πρώτα ότι

$$\mathsf{Ker}(f^k) \subseteq \mathsf{Ker}(f^{k+1}), \quad \forall k \geq 0$$

Έστω $\vec{x} \in \mathsf{Ker}(f^k)$, δηλαδή $f^k(\vec{x}) = \vec{0}$. Τότε $f(f^k(\vec{x})) = \vec{0}$ και άρα $(f \circ f^k)(\vec{x}) = f^{k+1}(\vec{x}) = \vec{0}$. Άρα $\vec{x} \in \mathsf{Ker}(f^{k+1}(\vec{x}) = \vec{0}$ και επομένως $\mathsf{Ker}(f^k) \subseteq \mathsf{Ker}(f^{k+1})$. Έτσι έχουμε την ακολουθία υπόχωρων (5.1). Τότε όμως θα έχουμε και

$$\dim_{\mathbb{K}} \mathsf{Ker}(f^k) \leq \dim_{\mathbb{K}} \mathsf{Ker}(f^{k+1}), \quad \forall k \geq 0$$

Τότε η ακολουθία υπόχωρων (5.1) επάγει την αύξουσα ακολουθία φυσικών αριθμών

$$0 \leq \dim_{\mathbb{K}} \mathsf{Ker}(f) \leq \dim_{\mathbb{K}} \mathsf{Ker}(f^2) \leq \cdots \leq \dim_{\mathbb{K}} \mathsf{Ker}(f^k) \leq \dim_{\mathbb{K}} \mathsf{Ker}(f^{k+1}) \leq \cdots \leq \dim_{\mathbb{K}} \mathcal{E}$$

Επειδή $\dim_{\mathbb{K}} \mathcal{E} < \infty$, έπεται το σύνολο φυσικών αριθμών $\big\{\dim_{\mathbb{K}} \mathsf{Ker}(f^k),\ k \geq 0\big\}$ είναι πεπερασμένο. Άρα υπάρχει φυσικός $\mu \geq 0$ έτσι ώστε: $\dim_{\mathbb{K}} \mathsf{Ker}(f^\mu) = \dim_{\mathbb{K}} \mathsf{Ker}(f^{\mu+1}) = \dim_{\mathbb{K}} \mathsf{Ker}(f^{\mu+2}) = \cdots$. Επειδή από την ακολουθία (5.1) έχουμε $\mathsf{Ker}(f^\mu) \subseteq \mathsf{Ker}(f^{\mu+1}) \subseteq \mathsf{Ker}(f^{\mu+2}) \subseteq \cdots$, έπεται ότι:

$$\mathsf{Ker}(f^\mu) = \mathsf{Ker}(f^{\mu+1}) = \mathsf{Ker}(f^{\mu+2}) = \cdots$$

(2) Δείχνουμε πρώτα ότι

$$\mathsf{Im}(f^{k+1}) \subseteq \mathsf{Im}(f^k), \quad \forall k \geq 0$$

Έστω $\vec{x} \in \mathsf{Im}(f^{k+1})$, δηλαδή υπάρχει $\vec{y} \in \mathcal{E}$ έτσι ώστε: $f^{k+1}(\vec{y}) = \vec{x}$. Τότε $f^{k+1}(\vec{y}) = (f^k \circ f)(\vec{y}) = f^k(f(\vec{y})) = \vec{x}$, και άρα $\vec{x} \in \mathsf{Im}(f^k)$. Έτσι $\mathsf{Im}(f^{k+1}) \subseteq \mathsf{Im}(f^k)$. Έτσι έχουμε την ακολουθία υπόχωρων (5.2). Τότε όμως θα έχουμε και

$$\mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f^{k+1}) \;\leq\; \mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f^k), \quad \forall k \geq 0$$

Τότε η ακολουθία υπόχωρων (5.2) επάγει την αύξουσα ακολουθία φυσικών αριθμών

$$0 \;\leq\; \cdots \;\leq\; \mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f^{k+1}) \;\leq\; \mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f^k) \;\leq\; \cdots \;\leq\; \mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f^2) \;\leq\; \mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f) \;\leq\; \mathsf{dim}_{\mathbb{K}} \mathcal{E}$$

Επειδή $\mathsf{dim}_{\mathbb{K}}\, \mathcal{E} < \infty$, έπεται το σύνολο φυσικών αριθμών $\big\{\, \mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f^k),\; k \geq 0 \big\}$ είναι πεπερασμένο. Άρα υπάρχει φυσικός $\lambda \geq 0$ έτσι ώστε: $\cdots = \mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f^{\lambda+2}) = \mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f^{\lambda+1}) = \mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f^{\lambda})$. Επειδή από τη ακολουθία (5.2) έχουμε $\cdots \subseteq \mathsf{Im}(f^{\lambda+2}) \subseteq \mathsf{Im}(f^{\lambda+1}) \subseteq \mathsf{Im}(f^{\lambda})$, έπεται ότι:

$$\cdots = \mathsf{Im}(f^{\lambda+2}) = \mathsf{Im}(f^{\lambda+1}) = \mathsf{Im}(f^{\lambda})$$

□

Λήμμα 5.3. *Με τους συμβολισμούς της Πρότασης* 5.2*, έστω* $m := \mathsf{max}\{\mu, \lambda\}$.

1. $f(\mathsf{Ker}(f^m) \;\subseteq\; \mathsf{Ker}(f^m)$ *και άρα η* $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ *επάγει μια γραμμική απεικόνιση:*

$$f_1 :\; \mathsf{Ker}(f^m) \;\longrightarrow\; \mathsf{Ker}(f^m), \quad f_1(\vec{x}) = f(\vec{x})$$

Επιπλέον η f_1 *είναι μηδενοδύναμη, δηλαδή* $f_1^r = 0$*, για κάποιον* $r \geq 1$.

2. $f(\mathsf{Im}(f^m) \;\subseteq\; \mathsf{Im}(f^m)$ *και άρα η* $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ *επάγει μια γραμμική απεικόνιση:*

$$f_2 :\; \mathsf{Im}(f^m) \;\longrightarrow\; \mathsf{Im}(f^m), \quad f_2(\vec{x}) = f(\vec{x})$$

Επιπλέον η f_2 *είναι ισομορφισμός.*

Απόδειξη. (1) Για το **1.** θα έχουμε:

(α) Έστω $\vec{x} \in \mathsf{Ker}(f^m)$, δηλαδή $f^m(\vec{x}) = \vec{0}$. Τότε $f(f^m(\vec{x})) = f(\vec{0}) = \vec{0}$ και άρα $f^{m+1}(\vec{x}) = \vec{0}$, δηλαδή $\vec{x} \in \mathsf{Ker}(f^{m+1})$. Όμως επειδή $m \geq \mu$, έχουμε $\mathsf{Ker}(f^{m+1}) = \mathsf{Ker}(f^m)$ και άρα $\vec{x} \in \mathsf{Ker}(f^m$. Επομένως $f(\mathsf{Ker}(f^m)) \subseteq \mathsf{Ker}(f^m)$. Προφανώς τότε η f επάγει μια γραμμική απεικόνιση $f_1\colon\, \mathsf{Ker}(f^m) \;\longrightarrow\; \mathsf{Ker}(f^m)$, ορίζοντας $f_1(\vec{x}) = f(\vec{x})$, $\forall \vec{x} \in \mathsf{Ker}(f^m)$.

(β) Έστω $\vec{x} \in \mathsf{Ker}(f^m)$, δηλαδή $f^m(\vec{x}) = \vec{0}$. Τότε $f_1^{m+1}(\vec{x} = f^{m+1}(\vec{x}) = f(f^m(\vec{0}) = f(\vec{0}) = \vec{0}$. Άρα θέτοντας $r = m+1$ έχουμε $f_1^r(\vec{x}) = 0$, $\forall \vec{x} \in \mathsf{Ker}(f^m)$ και επομένως $f_1^r = 0$, δηλαδή η f_1 είναι μηδενοδύναμη.

(2) Για το **2.** θα έχουμε:

(α) Έστω $\vec{x} \in f(\mathsf{Im}(f^m))$, δηλαδή υπάρχει $\vec{y} \in \mathsf{Im}(f^m)$ έτσι ώστε: $f(\vec{y}) = \vec{x}$. Επειδή $\vec{y} \in \mathsf{Im}(f^m)$ έπεται ότι υπάρχει $\vec{z} \in \mathcal{E}$ έτσι ώστε: $f^m(\vec{z}) = \vec{y}$. Τότε $\vec{x} = f(\vec{y}) = f(f^m(\vec{z})) = f^{m+1}(\vec{z}$ και άρα $\vec{x} \in \mathsf{Im}(f^{m+1})$. Επειδή $m \geq \lambda$, έπεται ότι θα έχουμε $\mathsf{Im}(f^{m+1}) = \mathsf{Im}(f^m)$ και άρα $\vec{x} \in \mathsf{Im}(f^m$. Επομένως $f(\mathsf{Im}(f^m)) \subseteq \mathsf{Im}(f^m)$. Προφανώς τότε η f επάγει μια γραμμική απεικόνιση $f_1\colon\, \mathsf{Im}(f^m) \;\longrightarrow\; \mathsf{Im}(f^m)$, ορίζοντας $f_2(\vec{x}) = f(\vec{x})$, $\forall \vec{x} \in \mathsf{Im}(f^m)$.

(β) Έστω $\vec{y} \in \mathsf{Im}(f^m)$. Επειδή $\mathsf{Im}(f^m) = \mathsf{Im}(f^{m+1})$, έπεται ότι $\vec{y} \in \mathsf{Im}(f^{m+1})$ και άρα υπάρχει $\vec{x} \in \mathcal{E}$ έτσι ώστε: $f^{m+1}(\vec{x}) = \vec{y}$. Τότε: $\vec{y} = f^{m+1}(\vec{x}) = f(f^m(\vec{x})$ και θέτοντας $\vec{z} := f^m(\vec{x}) \in \mathsf{Im}(f^m)$ θα έχουμε $\vec{y} = f(\vec{z}) = f_2(\vec{z})$. Αυτό σημαίνει ότι η f_2 είναι επιμορφισμός. Επειδή $\mathsf{dim}_{\mathbb{K}}\, \mathsf{Im}(f^m) \leq \mathsf{dim}_{\mathbb{K}}\, \mathcal{E} < \infty$, έπεται ότι η f_2 είναι ισομορφισμός.

□

Θεώρημα 5.4. *Έστω* $\mathcal{E}$ *ένας* $\mathbb{K}$*-διανυσματικός χώρος πεπερασμένης διάστασης υπεράνω του σώματος* $\mathbb{K}$*. Αν* $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ *είναι μια γραμμική απεικόνιση, τότε υπάρχει* $m \geq 1$ *έτσι ώστε:*

$$\mathcal{E} = \mathsf{Ker}(f^m) \oplus \mathsf{Im}(f^m)$$

Απόδειξη. Έστω $m = \mathsf{max}\{\mu, \lambda\}$ όπως στην Πρόταση 5.2 ή στο Λήμμα 5.3.

(1) Δείχνουμε πρώτα ότι: $\mathsf{Ker}(f^m) \cap \mathsf{Im}(f^m) = \{\vec{0}\}$.

Έστω $\vec{x} \in \mathsf{Ker}(f^m) \cap \mathsf{Im}(f^m)$. Τότε $f^m(\vec{x}) = \vec{0}$ και υπάρχει $\vec{y} \in \mathcal{E}$ έτσι ώστε $f^m(\vec{y}) = \vec{x}$. Τότε $f^{2m}(\vec{y}) = f^m(f^m(\vec{y})) = f^m(\vec{x}) = \vec{0}$ και άρα $\vec{y} \in \mathsf{Ker}(f^{2m})$. Επειδή $\mathsf{Ker}(f^{2m}) = \mathsf{Ker}(f^m)$, έπεται ότι $\vec{y} \in \mathsf{Ker}(f^m)$ και άρα $\vec{x} = f^m(\vec{y}) = \vec{0}$. Επομένως $\mathsf{Ker}(f^m) \cap \mathsf{Im}(f^m) = \{\vec{0}\}$.

(2) Δείχνουμε ότι: $\mathcal{E} = \mathsf{Ker}(f^m) + \mathsf{Im}(f^m)$.

Έστω $\vec{x} \in \mathcal{E}$. Τότε $f^m(\vec{x}) \in \mathsf{Im}(f^m)$. Επειδή $\mathsf{Im}(f^m) = \mathsf{Im}(f^{2m})$ έπεται ότι $f^m(\vec{x}) \in \mathsf{Im}(f^{2m})$ και άρα υπάρχει $\vec{y} \in \mathcal{E}$ έτσι ώστε: $f^m(\vec{x}) = f^{2m}(\vec{y})$. Τότε: $f^m(\vec{x}) = f^{2m}(\vec{y}) = f^m(f^m(\vec{y}))$ και άρα $f^m(\vec{x} - f^m(\vec{y})) = \vec{0}$. Θέτοντας $\vec{z} := \vec{x} - f^m(\vec{y})$ έπεται ότι $\vec{z} \in \mathsf{Ker}(f^m)$ και

$$\vec{x} = \vec{z} + f^m(\vec{y}, \qquad \vec{z} \in \mathsf{Ker}(f^m), \quad f^m(\vec{y}) \in \mathsf{Im}(f^m)$$

Η τελευταία σχέση δείχνει ότι: $\mathcal{E} = \mathsf{Ker}(f^m) + \mathsf{Im}(f^m)$.

Από τα (1) και (2) έχουμε ότι $\mathcal{E} = \mathsf{Ker}(f^m) \oplus \mathsf{Im}(f^m)$. □

Θεώρημα 5.5. *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης υπεράνω του σώματος $\mathbb{K}$. Αν $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι μια γραμμική απεικόνιση, τότε υπάρχει βάση $\mathcal{B}$ του $\mathcal{E}$ έτσι ώστε ο πίνακας της f στην βάση $\mathcal{B}$ να είναι της μορφής:*

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} \boxed{N} & \mathbb{O} \\ \mathbb{O} & \boxed{P} \end{pmatrix}$$

όπου ο P είναι αντιστρέψιμος και ο N είναι μηδενοδύναμος.

Απόδειξη. Έστω ο φυσικός αριθμός m όπως στο παραπάνω Θεώρημα 5.3. Τότε θα έχουμε $\mathcal{E} = \mathsf{Ker}(f^m) \oplus \mathsf{Im}(f^m)$. Από την άλλη πλευρά από την Πρόταση 5.2 η f επάγει έναν ισομορφισμό

$$f_2\colon\ \mathsf{Im}(f^m) \longrightarrow \mathsf{Im}(f^m), \quad f_2 = f|_{\mathsf{Im}(f^m)}$$

και μια μηδενοδύναμη γραμμική απεικόνιση

$$f_1\colon\ \mathsf{Ker}(f^m) \longrightarrow \mathsf{Ker}(f^m, \quad f_1 = f|_{\mathsf{Ker}(f^m)}$$

Έστω $\mathcal{B}_1$ μια βάση του υπόχωρου $\mathsf{Ker}(f^m)$ και $\mathcal{B}_2$ μια βάση του υπόχωρου $\mathsf{Im}(f^m)$. Τότε ο πίνακας $P := M_{\mathcal{B}_2}^{\mathcal{B}_2}(f_2)$ της f_2 στην βάση $\mathcal{B}_2$ θα είναι αντιστρέψιμος (επειδή η f_2 είναι ισομορφσιμός) και ο πίνακας $N := M_{\mathcal{B}_1}^{\mathcal{B}_1}(f_1)$ της f_1 στην βάση $\mathcal{B}_1$ θα είναι μηδενοδύναμος (επειδή η f_1 είναι μηδενοδύναμη.

Επειδή το άθροισμα $\mathsf{Ker}(f^m) + \mathsf{Im}(f^m)$ είναι ευθύ και μας δίνει τον χώρο $\mathcal{E}$, έπεται ότι το σύνολο $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ είναι μια βάση του $\mathcal{E}$. Τότε προφανώς $f(\mathcal{B}_2) \subseteq \mathsf{Im}(f^m)$ και $f(\mathcal{B}_1) \subseteq \mathsf{Ker}(f^m)$ και τότε ο πίνακας της f στην βάση $\mathcal{B}$ θα είναι της μορφής

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} \boxed{N} & \mathbb{O} \\ \mathbb{O} & \boxed{P} \end{pmatrix}$$

όπου ο P είναι αντιστρέψιμος και ο N είναι μηδενοδύναμος. □

5.2. **Κανονική Μορφή Fitting.** Είμαστε τώρα σε θέση να αποδείξουμε το Θεώρημα 5.1.

Απόδειξη του Θεωρήματος 5.1**:** Θεωρούμε την γραμμική αεπικόνιση

$$f_A\ :\ \mathbb{K}_n\ \longrightarrow\ \mathbb{K}_n, \qquad f_A(X) = A \cdot X$$

Από το Θεώρημα 5.4 έπεται ότι υπάρχει βάση $\mathcal{B}$ του $\mathbb{K}_n$ έτσι ώστε ο πίνακας της f_A στην βάση $\mathcal{B}$ να είναι της μορφής:

$$M^{\mathcal{B}}_{\mathcal{B}}(f_A) = \begin{pmatrix} \boxed{N} & \mathbb{O} \\ \mathbb{O} & \boxed{P} \end{pmatrix}$$

Επειδή ο πίνακας της f_A στην κανονική βάση του $\mathbb{K}_n$ είναι ο A έπεται ότι ο πίνακας A είναι όμοιος με τον παραπάνω πίνακα. $\square$

5.3. **Ευθύ Άθροισμα Γραμμικών Απεικονίσεων και Πινάκων.** Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος υπεράνω του σώματος $\mathbb{K}$.

Υποθέτουμε ότι $\mathcal{E} = \mathcal{V} \oplus \mathcal{W}$ και έστω $f : \mathcal{V} \longrightarrow \mathcal{V}$ και $g : \mathcal{W} \longrightarrow \mathcal{W}$ δύο γραμμικές απεικονίσεις.

Ορισμός 5.6. *Το* **ευθύ άθροισμα** $f \oplus g$ *των γραμμικών απεικονίσεων* f *και* g *ορίζεται να είναι η απεικόνιση*

$$f \oplus g : \ \mathcal{E} = \mathcal{V} \oplus \mathcal{W} \ \longrightarrow \ \mathcal{E} = \mathcal{V} \oplus \mathcal{W}, \quad (f \oplus g)(\vec{v} + \vec{w}) = f(\vec{v}) + g(\vec{w})$$

Είναι εύκολο να δει κανείς ότι η $f \oplus g$ είναι μια γραμμική απεικόνιση.

Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ και $B \in \mathsf{M}_{m\times m}(\mathbb{K})$ δύο τετραγωνικοί πίνακες, ενδεχομένως διαφορετικού μεγέθους.

Ορισμός 5.7. *Το* **ευθύ άθροισμα** $A \oplus B$ *των* $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ *και* $B \in \mathsf{M}_{m\times m}(\mathbb{K})$ *ορίζεται να είναι ο* $(n+m) \times (n+m)$ *πίνακας*

$$A \oplus B = \begin{pmatrix} \boxed{A} & \mathbb{O} \\ \mathbb{O} & \boxed{B} \end{pmatrix}$$

Παράδειγμα 5.8. *Έστω* $\mathcal{E}$ *ένας διανυσματικός χώρος πεπερασμένης διάστασης και υποθέτουμε ότι* $\mathcal{E} = \mathcal{V} \oplus \mathcal{W}$, *για κάποιους υπόχωρους* $\mathcal{V}, \mathcal{W}$ *του* $\mathcal{E}$. *Έστω* $f : \mathcal{V} \longrightarrow \mathcal{V}$ *και* $g : \mathcal{W} \longrightarrow \mathcal{W}$ *δύο γραμμικές απεικονίσεις. Αν* $\mathcal{B}_{\mathcal{V}}$ *είναι μαι βάση του* $\mathcal{V}$ *και* $\mathcal{B}_{\mathcal{W}}$ *είναι μια βάση του* $\mathcal{W}$, *τότε όπως γνωρίζουμε το σύνολο* $\mathcal{B} = \mathcal{B}_{\mathcal{V}} \cup \mathcal{B}_{\mathcal{W}}$ *είναι μια βάση του* $\mathcal{E}$.

Εύκολα βλέπουμε ότι θέτοντας $B = M^{\mathcal{B}_{\mathcal{V}}}_{\mathcal{B}_{\mathcal{V}}}(f)$ *και* $C = M^{\mathcal{B}_{\mathcal{W}}}_{\mathcal{B}_{\mathcal{W}}}(g)$, *έχουμε:*

$$M^{\mathcal{B}}_{\mathcal{B}}(f \oplus g) = B \oplus C = \begin{pmatrix} \boxed{B} & \mathbb{O} \\ \mathbb{O} & \boxed{C} \end{pmatrix}$$

Παράδειγμα 5.9. *Έστω* $\mathcal{E}$ *ένας διανυσματικός χώρος πεπερασμένης διάστασης και υποθέτουμε ότι* $\mathcal{E} = \mathcal{V} \oplus \mathcal{W}$, *για κάποιους υπόχωρους* $\mathcal{V}, \mathcal{W}$ *του* $\mathcal{E}$. *Έστω* $f : \mathcal{E} \longrightarrow \mathcal{E}$ *μια γραμμική απεικόνιση.*

Υποθέτουμε ότι: $f(\mathcal{V}) \subseteq \mathcal{V}$ *και* $f(\mathcal{W}) \subseteq \mathcal{W}$.

Τότε ορίζονται οι περιορισμοί της f *στους υπόχωρους* $\mathcal{V}$ *και* $\mathcal{W}$:

$$f_{\mathcal{V}} : \ \mathcal{V} \ \longrightarrow \ \mathcal{V}, \quad f_{\mathcal{V}}(\vec{v}) = f(\vec{v})$$

$$f_{\mathcal{W}} : \ \mathcal{W} \ \longrightarrow \ \mathcal{W}, \quad f_{\mathcal{W}}(\vec{w}) = f(\vec{w})$$

Προφανώς θα έχουμε:

$$f = f_{\mathcal{V}} \oplus f_{\mathcal{W}}$$

Θεώρημα 5.10. *Έστω $f : \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση, όπου $\mathsf{dim}_{\mathbb{K}}\mathcal{E} < \infty$. Τότε η f μπορεί να γραφεί ώς ευθύ άθροισμα*

$$f = g \oplus h$$

κατάλληλων γραμμικών απεικονίσεων $g\colon \mathcal{V} \longrightarrow \mathcal{V}$ και $h\colon \mathcal{W} \longrightarrow \mathcal{W}$, όπου $\mathcal{V}, \mathcal{W}$ είναι υπόχωροι του $\mathcal{E}$, και όπου:

(1) *η γραμμική απεικόνιση $g : \mathcal{V} \longrightarrow \mathcal{V}$ είναι ισομορφισμός.*
(2) *η γραμμική απεικόνιση $h : \mathcal{W} \longrightarrow \mathcal{W}$ είναι μηδενοδύναμη, δηλ. $h^m = 0$, για κάποιο $m \geq 1$.*

Απόδειξη. Θεωρούμε τον φυσικό αριθμό m του Θεωρήματος 5.4, και θέτουμε:

$$\mathcal{V} := \mathsf{Im}(f^m) \qquad \text{και} \qquad \mathcal{W} := \mathsf{Im}(f^m)$$

Τότε από το Θεώρημα 5.4 θα έχουμε:

$$\mathcal{E} = \mathcal{V} \oplus \mathcal{W}$$

Από το Λήμμα 5.3 έπεται ότι $f(\mathcal{V}) \subseteq \mathcal{V}$ και η επαγόμενη γραμμική απεικόνιση $f_{\mathcal{V}}\colon \mathcal{V} \longrightarrow \mathcal{V}$ είναι ισομορφισμός, και $f(\mathcal{W}) \subseteq \mathcal{W}$ και η επαγόμενη γραμμική απεικόνιση $f_{\mathcal{W}}\colon \mathcal{W} \longrightarrow \mathcal{W}$ είναι μηδενοδύναμη. Θέτοντας $g = f_{\mathcal{V}}$ και $h = f_{\mathcal{W}}$, με χρήση του Παραδείγματος 5.9 θα έχουμε το ζητούμενο. □

Θεώρημα 5.11. *Έστω $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ ένας τετραγωνικός πίνακας. Τότε ο A είναι όμοιος με έναν πίνακα D ο οποίος ώς ευθύ άθροισμα πινάκων:*

$$D = B \oplus C$$

όπου $B \in \mathsf{M}_{k\times k}(\mathbb{K})$, $C \in \mathsf{M}_{r\times r}(\mathbb{K})$, όπου $k + r = n$, και:

(1) *ο πίνακας B είναι αντιστρέψιμος.*
(2) *ο πίνακας C είναι μηδενοδύναμος, δηλ. $C^m = 0$, για κάποιο $m \geq 1$.*

Απόδειξη. Εφαρμόζουμε το Θεώρημα 5.10 για την γραμμική απεικόνιση

$$f_A : \mathbb{K}_n \longrightarrow \mathbb{K}_n, \qquad f_A(X) = A \cdot X$$

Υπάρχουν υπόχωροι $\mathcal{V}$ και $\mathcal{W}$ του $\mathbb{K}_n$ έτσι ώστε:

$$\mathbb{K}_n = \mathcal{V} \oplus \mathcal{V}$$

και $f_A(\mathcal{V}) \subseteq \mathcal{V}$ και $f_A(\mathcal{W}) \subseteq \mathcal{W}$. Με χρήση του Παραδείγματος 5.8 και του Θεωρήματος 5.11, αν $\mathcal{B}_1$ είναι μια βάση του $\mathcal{V}$ και $\mathcal{B}_2$ είναι μια βάση του $\mathcal{W}$, τότε θέτοντας $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$, θα έχουμε μια βάση του $\mathbb{K}_n$ στην οποία ο πίνακας D της f_A θα είναι το ευθύ άθροισμα $B \oplus C$ πινάκων, όπου B είναι ένας αντιστρέψιμος $k \times k$ πίνακας, ο C είναι ένας μηδενοδύναμος $r \times r$ πίνακας, και $k + r = n$. Επομένως ο πίνακας A είναι όμοιος με τον πίνακα $D = B \oplus C$. □

6. Ταυτόχρονη Διαγωνοποίηση

6.1. **Ταυτόχρονη Διαγωνοποίηση Πινάκων.** Έστω $A, B \in \mathsf{M}_{n\times n}(\mathbb{K})$ δύο $n \times n$ πίνακες με στοιχεία από το σώμα $\mathbb{K}$.

Ορισμός 6.1. *Οι πίνακες A, B καλούνται* **ταυτόχρονα διαγωνοποιήσιμοι** *αν υπάρχει αντιστρέψιμος πίνακας P έτσι ώστε:*

$$P^{-1} \cdot A \cdot P = \Delta_1 \qquad \text{και} \qquad P^{-1} \cdot B \cdot P = \Delta_2$$

όπου οι πίνακες Δ_1 και Δ_2 είναι διαγώνιοι. Δηλαδή αν υπάρχει αντιστρέψιμος πίνακας P ο οποίος διαγωνοποιεί ταυτόχρονα τους A, B.

Σκοπός μας στην παρούσα παράγραφο είναι να αποδείξουμε το ακόλουθο κριτήριο ταυτόχρονης διαγωνοποίησης:

Θεώρημα 6.2. *Έστω $A, B \in \mathsf{M}_{n\times n}(\mathbb{K})$ δύο $n \times n$ πίνακες με στοιχερία από το σώμα $\mathbb{K}$. Τότε τα ακόλυθα είναι ισοδύναμα:*

1. *Οι πίνακες A, B είναι ταυτόχρονα διαγωνοποιήσιμοι.*
2. *Οι πίνακες A, B είναι διαγωνοποιήσιμοι και: $A \cdot B = B \cdot A$.*

Απόδειξη. **1.** $\Longrightarrow$ **2.** Υποθέτουμε ότι οι πίνακες A, B είναι ταυτόχρονα διαγωνοποιήσιμοι. Τότε οι πίνακες είναι διαγωνοποιήσιμοι και υπάρχει αντιστρέψιμος πίνακας P έτις ώστε:

$$P^{-1} \cdot A \cdot P = \Delta_1 \qquad \text{και} \qquad P^{-1} \cdot B \cdot P = \Delta_2$$

όπου οι πίνακες Δ_1 και Δ_2 είναι διαγώνιοι. Τότε θα έχουμε:

$$P^{-1} \cdot (A \cdot B) \cdot P = P^{-1} \cdot A \cdot B \cdot P = P^{-1} \cdot A \cdot P \cdot P^{-1} \cdot B \cdot P = \Delta_1 \cdot \Delta_2$$

$$P^{-1} \cdot (B \cdot A) \cdot P = P^{-1} \cdot B \cdot A \cdot P = P^{-1} \cdot B \cdot P \cdot P^{-1} \cdot A \cdot P = \Delta_2 \cdot \Delta_1$$

Επειδή προφαβώς διαγώνιοι πίνακες μετατίθενται, θα έχουμε $\Delta_1 \cdot \Delta_2 = \Delta_2 \cdot \Delta_1$. Τότε

$$P^{-1} \cdot (A \cdot B) \cdot P = P^{-1} \cdot (B \cdot A) \cdot P \implies (A \cdot B) \cdot P = (B \cdot A) \cdot P \implies A \cdot B = B \cdot A$$

2. $\Longrightarrow$ **1.** Για τη απόδειξη αυτής της κατεύθυνσης χρειαζόμαστε κάποια προεργασία. Η απόδειξη θα ολοκληρωθεί στο Πόρισμα 6.8 στο τέλος της παραγράφου. □

Λήμμα 6.3. *Έστω $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση, όπου $\dim_{\mathbb{K}} \mathcal{E} < \infty$. Υποθέτουμε ότι*

$$\mathcal{E} = \mathcal{V} \oplus \mathcal{W} \qquad \text{και} \qquad f(\mathcal{V}) \subseteq \mathcal{V}$$

Θεωρούμε την επαγόμενη γραμμική απεικόνιση $f_{\mathcal{V}} = f|_{\mathcal{V}}\colon \mathcal{V} \longrightarrow \mathcal{V}$, όπου $f|_{\mathcal{V}}$ είναι ο περιορισμός της f στον υπόχωρο $\mathcal{V}$. Τότε το ελάχιστο πολυώνυμο $Q_{f_{\mathcal{V}}}(t)$ της $f_{\mathcal{V}}$ διαιρεί το ελάχιστο πολυώνυμο $Q_f(t)$ της f:

$$Q_{f_{\mathcal{V}}}(t) / Q_f(t)$$

Απόδειξη. Έστω $Q_f(t) = a_0 + a_1 t + \cdots a_{m-1} t^{m-1} + t^m$ το ελάχιστο πολυώνυμο της f, και θεωρούμε την γραμμική απεικόνιση

$$Q_f(f_{\mathcal{V}}) = a_0 \mathsf{Id}_{\mathcal{V}} + a_1 f_{\mathcal{V}} + \cdots + a_{m-1}(f_{\mathcal{V}})^{m-1} + (f_{\mathcal{V}})^m : \mathcal{V} \longrightarrow V$$

Τότε, $\forall \vec{v} \in \mathcal{V}$: $Q_f(f_{\mathcal{V}})(\vec{v}) = a_0(\vec{v}) + a_1 f_{\mathcal{V}}(\vec{v}) + \cdots + a_{m-1}(f_{\mathcal{V}})^{m-1}(\vec{v}) + (f_{\mathcal{V}})^m(\vec{v})$. Επειδή $f_{\mathcal{V}}(\vec{v}) = f(\vec{v})$, $\forall \vec{v} \in \mathcal{V}$, έπεται ότι θα έχουμε: $Q_f(f_{\mathcal{V}})(\vec{v}) = a_0(\vec{v}) + a_1 f(\vec{v}) + \cdots + a_{m-1} f^{m-1}(\vec{v}) + f^m(\vec{v}) = (a_0 \mathsf{Id}_{\mathcal{E}} + a_1 f + \cdots a_{m-1} f^{m-1} + f^m)(\vec{v}) = Q_f(f)(\vec{v}) = 0$. Αυτό σημαίνει ότι το πολυώνυμο $Q_f(t)$ μηδενίζει την $f_{\mathcal{V}}$. Τότε όμως το $Q_f(t)$ θα διαιρείται από το ελάχιστο πολυώνυμο της $f_{\mathcal{V}}$: $Q_{f_{\mathcal{V}}}(t) / Q_f(t)$. □

Πόρισμα 6.4. *Έστω $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση, όπου* $\dim_{\mathbb{K}} \mathcal{E} < \infty$. *Υποθέτουμε ότι*

$$\mathcal{E} = \mathcal{V} \oplus \mathcal{W} \qquad \textit{και} \qquad f(\mathcal{V}) \subseteq \mathcal{V}$$

Θεωρούμε την επαγόμενη γραμμική απεικόνιση $f_{\mathcal{V}}\colon \mathcal{V} \longrightarrow \mathcal{V}$, όπου $f_{\mathcal{V}} = f|_{\mathcal{V}}$ είναι ο περιορισμός της f στον υπόχωρο $\mathcal{V}$. Τότε:

$$f \ \textit{είναι διαγωνοποιήσιμη} \quad \Longrightarrow \quad f_{\mathcal{V}} \ \textit{είναι διαγωνοποιήσιμη}$$

Απόδειξη. Από το Λήμμα 6.3, έχουμε $Q_{f_{\mathcal{V}}}(t)/Q_f(t)$. Επειδή η f είναι διαγωνοπιήσιμη, έπεται ότι το ελάχιστο πολυώνυμο $Q_f(t)$ της f αναλύεται σε γινόμενο διακεκριμμένων πρωτοβαθμίων παραγόντων. Επειδή το ελάχιστο πολυώνυμο της $f_{\mathcal{V}}$ είναι διαιρέτης του $Q_f(t)$, έπεται ότι το ελάχιστο πολυώνυμο $Q_{f_{\mathcal{V}}}(t)$ της $f_{\mathcal{V}}$ αναλύεται σε γινόμενο διακεκριμμένων πρωτοβαθμίων παραγόντων. Επομένως η $f_{\mathcal{V}}$ είναι διαγωνοποιήσιμη. ☐

Πόρισμα 6.5. *Έστω $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια γραμμική απεικόνιση, όπου* $\dim_{\mathbb{K}} \mathcal{E} < \infty$. *Υποθέτουμε ότι*

$$\mathcal{E} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_k \qquad \textit{και} \qquad f(\mathcal{V}_i) \subseteq \mathcal{V}_i, \ \ 1 \leq i \leq k$$

Θεωρούμε τις επαγόμενες γραμμικές απεικόνισεις $f_{\mathcal{V}_i}\colon \mathcal{V}_i \longrightarrow \mathcal{V}_i$, όπου $f_{\mathcal{V}_i} = f|_{\mathcal{V}_i}$ είναι ο περιορισμός της f στον υπόχωρο $\mathcal{V}_i$. Τότε:

$$f \ \textit{είναι διαγωνοποιήσιμη} \quad \Longleftrightarrow \quad f_{\mathcal{V}_i} \ \textit{είναι διαγωνοποιήσιμη}, \quad 1 \leq i \leq k$$

Απόδειξη. "$\Longrightarrow$" Έπεται άμεσα από το Πόρισμα 6.4.

"$\Longleftarrow$" Έστω $\mathcal{B}_i$ μια βάση του $\mathcal{V}_i$, $1 \leq i \leq k$, η οποία αποτελείται από ιδιοδιανύσματα της f_i. Από το ορισμό των απεικονίσεων f_i, έπεται προφανώς ότι τα διανύσματα κάθε βάσης $\mathcal{B}_i$ είναι και ιδιοδιανύσματα της f. Επειδή το άθροισμα $\mathcal{V}_1 + \mathcal{V}_2 + \cdots + \mathcal{V}_k$ είναι ευθύ και μας δίνει τον χώρο $\mathcal{E}$, έπεται ότι το σύνολο

$$\mathcal{C} := \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$$

είναι μια βάση του $\mathcal{E}$ η οποία αποτελείται από ιδιοδιανύσματα της f. Άρα η f είναι διαγωνοποιήσιμη. ☐

Θεώρημα 6.6. *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης και $f, g\colon \mathcal{E} \longrightarrow \mathcal{E}$ δύο γραμμικές απεικονίσεις. Υποθέτουμε ότι οι f, g είναι διαγωνοποιήσιμες και $f \circ g = g \circ f$. Τότε υπάρχει βάση $\mathcal{C}$ του $\mathcal{E}$ η οποία αποτελείται από ιδιοδιανύσματα της f και της g.*

Απόδειξη. Επειδή η g είναι διαγωνοποιήσιμη, όπως γνωρίζουμε θα έχουμε:

$$\mathcal{E} = \mathcal{V}_g(\lambda_1) \oplus \mathcal{V}_g(\lambda_2) \oplus \cdots \oplus \mathcal{V}_g(\lambda_k)$$

όπου $\lambda_1, \cdots \lambda_k$ είναι οι διακεκριμμένες ιδιοτιμές της g και $\mathcal{V}_g(\lambda_i)$ είναι οι αντίστοιχοι ιδιοχώροι:

$$\mathcal{V}_g(\lambda_i) = \{\vec{x} \in \mathcal{E} \mid g(\vec{x}) = \lambda_i \vec{x}\}, \qquad 1 \leq i \leq k$$

Έστω $\vec{x} \in \mathcal{V}_g(\lambda_i)$. Τότε $g(\vec{x}) = \lambda_i \vec{x}$ και επομένως, $\forall i = 1, 2, \cdots, k$:

$$f(g(\vec{x})) = f(\lambda_i \vec{x}) \implies (f \circ g)(\vec{x}) = \lambda_i f(\vec{x}) \implies (g \circ f)(\vec{x}) = \lambda_i f(\vec{x}) \implies g(f(\vec{x})) = \lambda_i f(\vec{x})$$

Επομένως το διάνυσμα $f(\vec{x}) \in \mathcal{V}_g(\lambda_i)$, και άρα: $f(\mathcal{V}_g(\lambda_i) \subseteq \mathcal{V}_g(\lambda_i)$. Αυτό σημαίναι ότι μπορούμε να περιορίσουμε την f σε μια γραμμική απεικόνιση

$$f_i := f|_{\mathcal{V}_g(\lambda_i)} \, : \, \mathcal{V}_g(\lambda_i) \longrightarrow \mathcal{V}_g(\lambda_i), \quad f_i(\vec{x}) = f(\vec{x})$$

Επειδή η f είναι διαγωνοποιήσιμη, από το Πόρισμα 6.5 έπεται ότι η απεικόνιση f_i είναι διαγωνοποιήσιμη, για κάθε $i = 1, 2, \cdots, k$. Έστω $\mathcal{B}_i$ μια βάση του $\mathcal{V}_g(\lambda_i)$ η οποία αποτελείται από ιδιοδιανύσματα της f_i. Από το ορισμό των απεικονίσεων f_i, έπεται προφανώς ότι τα διανύσματα κάθε βάσης $\mathcal{B}_i$ είναι και ιδιοδιανύσματα της f. Επειδή το άθροισμα $\mathcal{V}_g(\lambda_1) + \mathcal{V}_g(\lambda_2) + \cdots + \mathcal{V}_g(\lambda_k)$ είναι ευθύ και μας δίνει τον χώρο $\mathcal{E}$, έπεται ότι το σύνολο

$$\mathcal{C} := \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$$

είναι μια βάση του $\mathcal{E}$ η οποία αποτελείται από ιδιοδιανύσματα της f. Επειδή $\mathcal{B}_i \subseteq \mathcal{V}_g(\lambda_i)$ έπεται ότι τα διανύσματα κάθε βάσης $\mathcal{B}_i$ είναι και ιδιοδιανύσματα της g. Καταλήγουμε ότι η βάση $\mathcal{C}$ του $\mathcal{E}$ αποτελείται από ιδιοδιανύσματα της f και της g. □

6.2. **Ταυτόχρονη διαγωνοποίηση Γραμμικών Απεικονίσεων.**

Ορισμός 6.7. *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης και $f, g \colon \mathcal{E} \longrightarrow \mathcal{E}$ δύο γραμμικές απεικονίσεις. Τότε οι f, g καλούνται* **ταυτόχρονα διαγωνοποιήσιμες**, *αν υπάρχει βάση $\mathcal{C}$ του $\mathcal{E}$ η οποία αποτελείται από ιδιοδιανύσματα της f και της g.*

Θεώρημα 6.8. *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης και $f, g \colon \mathcal{E} \longrightarrow \mathcal{E}$ δύο γραμμικές απεικονίσεις. Τότε τα ακόλουθα είναι ισοδύναμα:*

(1) *Οι f, g είναι ταυτόχρονα διαγωνοποιήσιμες.*
(2) *Οι f, g είναι διαγωνοποιήσιμες και $f \circ g = g \circ f$.*

Απόδειξη. (2) $\Longrightarrow$ (1) Η κατεύθυνση αυτή αποδείχθηκε στο Θεώρημα 6.6.

(1) $\Longrightarrow$ (2) Έστω ότι οι f, g είναι ταυτόχρονα διαγωνοποιήσιμες, δηλαδή υπάρχει βάση

$$\mathcal{C} = \big\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\big\}$$

του $\mathcal{E}$ η οποία αποτελείται από ιδιοδιανύσματα της f και της g, τα οποία αντιστοιχούν στις ιδιοτιμές $\lambda_1, \lambda_2, \cdots, \lambda_n$ της f και στις ιδιοτιμές $\mu_1, \mu_2, \cdots, \mu_n$ της f Τότε:

$$(f \circ g)(\vec{e}_i) = f(g(\vec{e}_i)) = f(\mu_i \vec{e}_i) = \mu_i f(\vec{e}_i) = \mu_i \lambda_i \vec{e}_i$$
$$(g \circ f)(\vec{e}_i) = g(f(\vec{e}_i)) = g(\lambda_i \vec{e}_i) = \lambda_i g(\vec{e}_i) = \lambda_i \mu_i \vec{e}_i$$

Επομένως $(f \circ g)(\vec{e}_i) = (g \circ f)(\vec{e}_i)$, για κάθε διάνυσμα $\vec{e}_i$ της βάσης $\mathcal{C}$. Τότε όμως $f \circ g = g \circ f$. □

Το ακόλουθο Πόρισμα ολοκληρώνει την απόδειξη της κατεύθυνσης **2.** $\Longrightarrow$ **1.** του Θεωρήματος 6.2.

Πόρισμα 6.9. *Έστω A και B δύο διαγωνοποιήσιμοι $n \times n$ πίνακες υπεράνω του σώματος $\mathbb{K}$. Αν $A \cdot B = B \cdot A$, τότε υπάρχει αντιστρέψιμος πίνακας P έτσι ώστε:*

$$P^{-1} \cdot A \cdot P = \Delta_1 \qquad \text{και} \qquad P^{-1} \cdot B \cdot P = \Delta_2$$

όπου οι πίνακες Δ_1 και Δ_2 είναι διαγώνιοι.

Απόδειξη. Θεωρούμε τις γραμμικές απεικονίσεις

$$f_A,\ f_B : \mathbb{K}_n \longrightarrow \mathbb{K}_n, \qquad f_A(X) = A \cdot X \quad \text{και} \quad f_B(X) = B \cdot X$$

Επειδή οι πίνακες A και B είναι διαγωνοποιήσιμοι, έπεται ότι και οι γραμμικές απεικονλίσεις f_A και f_B είναι διαγωνοποιήσιμες. Επειδή $A \cdot B = B \cdot A$, έπεται άμεσα ότι $f_A \circ f_B = f_B \circ f_A$. Τότε από το Θεώρημα 6.5 έπεται ότι υπάρχει βάση $\mathcal{C}$ του $\mathbb{K}_n$ η οποία αποτελείται από ιδιοδιανύσματα της f_A και της f_B. Άρα ο πίνακας της f_A στην $\mathcal{C}$ είναι ένας διαγώνιος πίνακας Δ_1 και ο πίνακας της f_B στην $\mathcal{C}$ είναι ένας διαγώνιος πίνακας Δ_2. Τότε όμως υπάρχει αντιστρέψιμος πίνακας P έτσι ώστε οι πίνακας $P^{-1} \cdot A \cdot P = \Delta_1$ και $P^{-1} \cdot B \cdot P = \Delta_2$. □

Τα παραπάνω γενικεύονται και για παραπάνω από δύο πίνακες.

Ορισμός 6.10. *Έστω $A_1, A_2, \cdots, A_r$ $n \times n$ πίνακες υπεράνω του σώματος $\mathbb{K}$. Οι πίνακες $A_1, A_2, \cdots, A_r$ καλούνται* **ταυτόχρονα διαγωνοποιήσιμοι** *αν υπάρχει αντιστρέψιμος πίνακας P έτσι ώστε:*

$$P^{-1} \cdot A_i \cdot P = \Delta_i \qquad 1 \leq i \leq r$$

όπου οι πίνακες Δ_i διαγώνιοι. Δηλαδή αν υπάρχει αντιστρέψιμος πίνακας P ο οποίος διαγωνοποιεί ταυτόχρονα τους $A_1, A_2, \cdots, A_r$.

Ορισμός 6.11. *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης υπεράνω του σώματος $\mathbb{K}$, και $f_1, f_2, \cdots, f_r : \mathcal{E} \longrightarrow \mathcal{E}$ γραμμικές απεικονίσεις. Οι $f_1, f_2, \cdots, f_r$ καλούνται* **ταυτόχρονα διαγωνοποιήσιμες** *αν υπάρχει βάση $\mathcal{B}$ του $\mathcal{E}$ η οποία αποτελείται από ιδιοδιανύσματα των $f_1, f_2, \cdots, f_r$.*

Με βάση τα Θεωρήματα 6.2 και 6.8 και με χρήση επαγωγής αποδεικνύεται εύκολα και το ακόλουθο Θεώρημα:

Θεώρημα 6.12. (1) *Έστω $A_1, A_2, \cdots, A_r$ $n \times n$ πίνακες υπεράνω του σώματος $\mathbb{K}$. Τότε τα ακόλουθα είναι ισοδύναμα:*
- (α') *Οι πίνακες $A_1, A_2, \cdots, A_r$ είναι ταυτόχρονα διαγωνοποιήσιμοι.*
- (β') *Οι πίνακες $A_1, A_2, \cdots, A_r$ είναι διαγωνοποιήσιμοι και $A_i \cdot A_j = A_j \cdot A_i$, $1 \leq i \neq j \leq r$.*

(2) *Έστω $f_1, f_2, \cdots, f_r : \mathcal{E} \longrightarrow \mathcal{E}$ γραμμικές απεικονίσεις, όπου $\dim_{\mathbb{K}} \mathcal{E} < \infty$. Τότε τα ακόλουθα είναι ισοδύναμα:*
- (α') *Οι $f_1, f_2, \cdots, f_r$ είναι ταυτόχρονα διαγωνοποιήσιμες.*
- (β') *Οι $f_1, f_2, \cdots, f_r$ είναι διαγωνοποιήσιμες και $f_i \circ f_j = f_j \circ f_i$, $1 \leq i \neq j \leq r$.*

Άσκηση 6.13. 1. *Να ορισθούν οι έννοιες "ταυτόχρονα τριγωνοποιήσιμοι πίνακες" και "ταυτόχρονα τριγωνοποιήσιμες γραμμικές απεικονίσεις"*

2. *Να εξετασθεί αν το Θεώρημα 6.12 ισχύει αν στην διατύπωσή του έχουμε παντού τριγωνοποιήσιμους πίνακες ή τριγωνοποιήσιμες γραμμικές απεικονίσεις αντί διαγωνοποιήσιμους πίνακες ή διαγωνοποιήσιμες γραμμικές απεικονίσεις.*

7. **Η Κανονική Μορφή Jordan - I**

Στην παρούσα παράγραφο θα μελετήσουμε μια σημαντική κανονική μορφή πινάκων οι οποίοι έχουν την ιδιότητα ότι το χαρακτηριστικό τους πολυώνυμο αναλύεται σε γινόμενο πρωτοβαθμίων, όχι κατ' ανάγκη διακεκριμμένων, παραγόντων.

Έστω $\mathbb{K}$ ένα σώμα.

Ορισμός 7.1. *Έστω $n \geq 1$ και $\lambda \in \mathbb{K}$. Ο τετραγωνικός $n \times n$ πίνακας*

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \lambda \end{pmatrix}$$

καλείται ο **στοιχειώδης** *$n \times n$* **πίνακας Jordan** *ο οποίος αντιστοιχεί στο λ.*

Ένας πίνακας $n \times n$ πινακας J καλείται **πίνακας Jordan** *αν είναι το ευθύ άθροισμα στοιχειωδών πινάκων Jordan:*

$$J = J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \cdots \oplus J_{n_k}(\lambda_k)$$

όπου $\lambda_i \in \mathbb{K}$, $1 \leq i \leq k$, και $n_1 + n_2 + \cdots + n_k = n$.

Στην παρούσα παράγραφο θα αποδείξουμε ότι κάθε $n \times n$ πίνακας A με στοιχεία από το σώμα $\mathbb{K}$, ο οποίος έχει όλες τις ιδιοτιμές του στο $\mathbb{K}$ είναι όμοιος με έναν πίνακα Jordan.

7.1. **Μηδενοδύναμες Γραμμικές Απεικονίσεις.** Στην παρούσα παράγραφο συμβολίζουμε με $\mathbb{K}$ ένα σώμα.

Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος πεπερασμένης διάστασης και

$$f : \mathcal{E} \longrightarrow \mathcal{E}$$

μια γραμμική απεικόνιση.

Θεώρημα 7.2. *Υποθέτουμε ότι η f είναι μηδενοδύναμη: $f^m = 0$. Τότε υπάρχουν διανύσματα*

$$\vec{z}_1,\ \vec{z}_2,\ \cdots,\ \vec{z}_k \ \in \ \mathcal{E}$$

και φυσικοί αριθμοί

$$\rho_1,\ \rho_2,\ \cdots,\ \rho_k \ \geq \ 1$$

έτσι ώστε: $f^{\rho_1}(\vec{z}_1) = f^{\rho_2}(\vec{z}_2) = \cdots = f^{\rho_k}(\vec{z}_k) = \vec{0}$, δηλαδή:

$$f^{\rho_i}(\vec{z}_i) = \vec{0}, \qquad 1 \leq i \leq k$$

και το σύνολο

$$\mathcal{B} = \big\{\vec{z}_1, f(\vec{z}_1), \cdots, f^{\rho_1 - 1}(\vec{z}_1),\ \ \vec{z}_2, f(\vec{z}_2), \cdots, f^{\rho_2 - 1}(\vec{z}_2),\ \ \cdots,\ \ \vec{z}_k, f(\vec{z}_k), \cdots, f^{\rho_k - 1}(\vec{z}_k)\big\}$$

να είναι βάση του $\mathcal{E}$.

Απόδειξη. Η απόδειξη θα γίνει με επαγωγή στην διάσταση $\dim_{\mathbb{K}} \mathcal{E} := n$.

Αν $\dim_{\mathbb{K}} \mathcal{E} = 0$, τότε $\mathcal{E} = \{\vec{0}\}$ και $f = 0$. Το συμπέρασμα τότε ισχύει τετριμμένα.

1. Υποθέτουμε ότι $\boxed{n=1}$. Θεωρούμε μιά βάση $\{\vec{e}\}$ του $\mathcal{E}$, και τότε $\forall \vec{x} \in \mathcal{E}$, θα έχουμε $\vec{x} = k\vec{e}$ και άρα: $f(\vec{x}) = f(k\vec{e}) = kf(\vec{e})$. Έτσι θέτοντας $f(\vec{e}) = \vec{\varepsilon} \in \mathbb{K}$, έπεται ότι $\vec{\varepsilon} = \lambda\vec{e}$, για κάποιο $\lambda \in \mathbb{K}$, και άρα θα έχουμε: $f(\vec{x}) = k\vec{\varepsilon} = k\lambda\vec{e} = \lambda\vec{x}$, $\forall \vec{x} \in \mathcal{E}$. Επειδή $f^m = 0$, έπεται ότι: $f^m(\vec{x}) = \lambda^m\vec{x} = \vec{0}$. Ιδιαίτερα $f^m(\vec{e}) = \lambda^m\vec{e} = \vec{0}$ και άρα $\lambda^m = 0$ διότι $\vec{e} \neq \vec{0}$. Επομένως $\lambda = 0$ και άρα $f = 0$. Τότε θέτοντας $k = 1$, $\rho_1 = 1$, και $\vec{z}_1 = \vec{e}$, έπεται ότι το σύνολο $\mathcal{B} = \{\vec{e}\}$ είναι μια βάση του $\mathcal{E}$ με την επιθυμιτή ιδιότητα.

2. Επαγωγική Υπόθεση: Υποθέτουμε ότι το συμπέρασμα ισχύει για όλους τους $\mathbb{K}$-διανυσματικούς χώρους $\mathcal{F}$ με $\dim_{\mathbb{K}} \mathcal{F} < n$ και κάθε μηδενοδύναμη γραμμική απεικόνιση $g : \mathcal{F} \longrightarrow \mathcal{F}$.

3. Γενική Περίπτωση: Υποθέτουμε, όπως στην εκφώνηση του Θεωρήματος, ότι $\dim_{\mathbb{K}} \mathcal{E} = n \geq 1$ και η $f : \mathcal{E} \longrightarrow \mathcal{E}$ είναι μηδενοδύναμη: $f^m = 0$, για κάποιον φυσικό $m \geq 1$. Επιπλέον υποθέτουμε ότι $f \neq 0$, διότι αν $f = 0$, τότε το συμπέρασμα ισχύει κατά τετριμμένο τρόπο.

Θέτουμε $\mathcal{F} := \mathsf{Im}(f) = f(\mathcal{E})$.

<u>Ισχυρισμός 1:</u> $0 < \dim_{\mathbb{K}} \mathcal{F} < n = \dim_{\mathbb{K}} \mathcal{E}$.

Πραγματικά: Αν $0 = \dim_{\mathbb{K}} \mathcal{F}$, τότε $\mathcal{F} = f(\mathcal{E}) = \{\vec{0}\}$, δηλαδή: $f(\vec{x}) = \vec{0}$, και άρα $f = 0$ και το συμπέρασμα ισχύει.

Αν $\dim_{\mathbb{K}} \mathcal{F} < n = \dim_{\mathbb{K}} \mathcal{E}$, τότε προφανώς θα έχουμε $\mathcal{F} = \mathsf{Im}(f) = \mathcal{E}$ και επομένως επειδή $f^m = 0$, θα καταλήξουμε στο άτοπο:

$$f^2(\mathcal{E}) = f(f(\mathcal{E})) = f(\mathcal{E}) = \mathcal{E}, \quad \cdots, \quad \{\vec{0}\} = f^m(\mathcal{E}) = f^{m-1}(f(\mathcal{E})) = f^{m-1}(\mathcal{E}) = \cdots = f(\mathcal{E}) = \mathcal{E}$$

Συμφωνα με τον παραπάνω Ισχυρισμό 1, θα έχουμε $\{\vec{0}\} \neq \mathcal{F} \neq \mathcal{E}$.

Από την άλλη πλευρά είναι προφανές ότι η μηδενοδύναμη γραμμική απεικόνιση $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ επάγει μια γραμμική απεικόνιση

$$g : \mathcal{F} \longrightarrow \mathcal{F}, \qquad g = f|_{\mathcal{F}}, \qquad \text{δηλαδή} \qquad g(\vec{x}) = f(\vec{x}), \quad \forall \vec{x} \in \mathcal{F}$$

Προφανώς η g είναι μηδενοδύναμη, διότι η f είναι μηδενοδύναμη, και άρα από την επαγωγική υπόθεση, υπάρχουν διανύσματα

$$\vec{w}_1,\ \vec{w}_2,\ \cdots,\ \vec{w}_l\ \in\ \mathcal{F}$$

και φυσικοί αριθμοί

$$\sigma_1,\ \sigma_2,\ \cdots,\ \sigma_l\ \geq\ 1$$

έτσι ώστε: $g^{\sigma_1}(\vec{w}_1) = g^{\sigma_2}(\vec{w}_2) = \cdots = g^{\sigma_l}(\vec{w}_l) = \vec{0}$, δηλαδή:

$$g^{\sigma_i}(\vec{w}_i) = \vec{0}, \qquad 1 \leq i \leq l$$

και το σύνολο

$$\mathcal{C} = \left\{\vec{w}_1, g(\vec{w}_1), \cdots, g^{\sigma_1-1}(\vec{w}_1),\ \ \vec{w}_2, g(\vec{w}_2), \cdots, g^{\sigma_2-1}(\vec{w}_2),\ \ \cdots,\ \ \vec{w}_l, g(\vec{w}_l), \cdots, g^{\sigma_l-1}(\vec{w}_l)\right\}$$

να είναι βάση του $\mathcal{F}$.

Επειδή $f(\mathcal{E}) = \mathcal{F} \neq \mathcal{E}$, έπεται ότι για τα διανύσματα $\vec{w}_i \in \mathcal{F}$, $1 \leq i \leq l$, υπάρχουν διανύσματα $\vec{z}_i \in \mathcal{E}$ έτσι ώστε:

$$f(\vec{z}_i) = \vec{w}_i, \qquad 1 \leq i \leq l$$

Παρατηρούμε ότι επειδή

$$g^{\sigma_i}(\vec{w}_i) = f^{\sigma_i}(\vec{w}_i) = \vec{0}, \qquad 1 \leq i \leq l$$

έπεται ότι τα διανύσματα

$$f^{\sigma_i-1}(\vec{w}_i)\ \in\ \mathsf{Ker}(f), \qquad 1 \leq i \leq l$$

<u>Ισχυρισμός 2:</u> Το σύνολο

$$\{f^{\sigma_1-1}(\vec{w}_1),\ \cdots,\ f^{\sigma_l-1}(\vec{w}_l)\} \qquad \text{είναι γραμμικά ανεξάρτητο}$$

Πραγματικά: αυτό προκύπτει διότι τα παραπάνω διανύσματα είναι στοιχεία της βάσης $\mathcal{C}$.

Συμπληρώνουμε το σύνολο $\{f^{\sigma_1-1}(\vec{w}_1), \cdots, f^{\sigma_l-1}(\vec{w}_l)\}$ σε μια βάση του $\mathsf{Ker}(f)$:

$$\mathcal{D} = \left\{f^{\sigma_1-1}(\vec{w}_1), \cdots, f^{\sigma_l-1}(\vec{w}_l), \ \vec{y}_1, \vec{y}_2, \cdots, \vec{y}_t\right\}$$

Επομένως

$$\mathsf{dim}_{\mathbb{K}} \mathsf{Ker}(f) = l + t$$

Από την άλλη πλευρά θα έχουμε

$$\mathsf{dim}_{\mathbb{K}}\mathcal{F} = \mathsf{dim}_{\mathbb{K}} \mathsf{Im}(f) = \sigma_1 + \sigma_2 + \cdots \sigma_l$$

διότι το σύνολο $\mathcal{D}$ είναι βάση του $\mathcal{F}$ και $|\mathcal{D}| = \sigma_1 + \sigma_2 + \cdots \sigma_l$.

Από την Θεμελιώδη εξίσωση διαστάσεων για την γραμμική απεικόνιση $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ θα έχουμε

$$\mathsf{dim}_{\mathbb{K}}\mathcal{E} = \mathsf{dim}_{\mathbb{K}} \mathsf{Ker}(f) + \mathsf{dim}_{\mathbb{K}} \mathsf{Im}(f) = \sigma_1 + \sigma_2 + \cdots \sigma_l + l + t$$

Ισχυρισμός 3: Το σύνολο

$$\mathcal{B} = \left\{\vec{z}_1, f(\vec{z}_1), \cdots, f^{\sigma_1}(\vec{z}_1), \ \vec{z}_2, f(\vec{z}_2), \cdots, f^{\sigma_2}(\vec{z}_2), \ \cdots, \ \vec{z}_l, f(\vec{z}_l), \cdots, f^{\sigma_l}(\vec{z}_l), \ y_1, y_2, \cdots, y_t\right\}$$

είναι μια βάση του $\mathcal{E}$.

Απόδειξη του Ισχυρισμού 3: Πρώτα παρατηρούμε ότι:

$$|\mathcal{B}| = (\sigma_1 + 1) + (\sigma_2 + 1) + \cdots (\sigma_l + 1) + t = \sigma_1 + \sigma_2 + \cdots \sigma_l + l + t = \mathsf{dim}_{\mathbb{K}}\mathcal{E}$$

Άρα για να είναι το σύνολο $\mathcal{B}$ βάση του $\mathcal{E}$, αρκεί να είναι γραμμικά ανεξάρτητο.

Υποθέτουμε ότι:

$$\begin{aligned}
\vec{0} &= \lambda_0\vec{z}_1 + \lambda_1 f(\vec{z}_1) + \cdots + \lambda_{\sigma_1} f^{\sigma_1}(\vec{z}_1) \\
&= \mu_0\vec{z}_2 + \mu_1 f(\vec{z}_2) + \cdots + \mu_{\sigma_2} f^{\sigma_2}(\vec{z}_2) \\
&\vdots \\
&= \nu_0\vec{z}_l + \nu_1 f(\vec{z}_l) + \cdots + \nu_{\sigma_l} f^{\sigma_l}(\vec{z}_l) \\
&= \xi_1 y_1 + \xi_2 y_2 + \cdots \xi_t y_t
\end{aligned}$$

Εφαρμόζοντας την f στην παραπάνω σχέση και χρησιμοποιώντας ότι $y_i \in \mathsf{Ker}(f)$, $1 \leq i \leq t$, θα έχουμε:

$$\begin{aligned}
\vec{0} &= \lambda_0 f(\vec{z}_1) + \lambda_1 f^2(\vec{z}_1) + \cdots + \lambda_{\sigma_1} f^{\sigma_1+1}(\vec{z}_1) \\
&= \mu_0 f(\vec{z}_2) + \mu_1 f^2(\vec{z}_2) + \cdots + \mu_{\sigma_2} f^{\sigma_2+1}(\vec{z}_2) \\
&\vdots \\
&= \nu_0 f(\vec{z}_l) + \nu_1 f^2(\vec{z}_l) + \cdots + \nu_{\sigma_l} f^{\sigma_l+1}(\vec{z}_l)
\end{aligned} \tag{7.1}$$

Όμως

$$f(\vec{z}_i) = \vec{w}_i, \qquad 1 \leq i \leq l$$

και επομένως η τελευταία σχέση γράφεται:

$$
\begin{aligned}
\vec{0} &= \lambda_0\vec{w}_1 + \lambda_1 f(\vec{w}_1) + \cdots + \lambda_{\sigma_1} f^{\sigma_1}(\vec{w}_1) \\
&= \mu_0\vec{w}_2 + \mu_1 f(\vec{w}_2) + \cdots + \mu_{\sigma_2} f^{\sigma_2}(\vec{w}_2) \\
&\vdots \\
&= \nu_0\vec{w}_l + \nu_1 f(\vec{w}_l) + \cdots + \nu_{\sigma_l} f^{\sigma_l}(\vec{w}_l)
\end{aligned}
\tag{7.2}
$$

Επειδή τα διανύσματα $\vec{w}_i$ ανήκουν στον υπόχωρο $\mathcal{F} = \mathsf{Im}(f)$ και $g(\vec{x}) = f(\vec{x})$, $\forall \vec{x} \in \mathcal{F}$, η τελευταία σχέση είναι ένας γραμμικός συνδυασμός των διανυσμάτων της βάσης $\mathcal{C}$ του $\mathcal{F}$. Επομένως θα έχουμε:

$$\lambda_0 = \cdots = \lambda_{\sigma_1} = \ \cdots \ = \mu_0 = \cdots = \mu_{\sigma_2} = \ \cdots \ = \nu_0 = \cdots = \nu_{\sigma_l}$$

και άρα το σύνολο $\mathcal{B}$ είναι γραμμικά ανεξάρτητο.

Καταλήγουμε ότι το σύνολο $\mathcal{B}$ είναι μια βάση του $\mathcal{E}$.

Η βάση $\mathcal{B}$ έχει την επιθυμιτή ιδιότητα της εκφώνησης, θέτοντας:

$$k = l + t, \quad \rho_1 = \sigma_1 + 1, \cdots, \rho_l = \sigma_l + 1,\ \rho_{l+1} = 1, \cdots, \rho_k = 1$$

και

$$\vec{z}_{l+1} = \vec{y}_1,\ \cdots,\ \vec{z}_k = \vec{y}_k$$

□

Θεώρημα 7.3. *Έστω $\mathcal{E}$ ένας $\mathbb{K}$-διαννσματικός χώρος πεπερασμένης διάστασης και $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια μηδενοδύναμη γραμμική απεικόνιση. Τότε υπάρχει μια βάση του $\mathcal{B}$ του $\mathcal{E}$ στην οποία ο πίνακας της f είναι ένας πίνακας Jordan της μορφής:*

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = J_{\rho_1}(0) \oplus J_{\rho_2}(0) \oplus \cdots \oplus J_{\rho_k}(0)$$

όπου $\rho_1 + \rho_2 + \cdots + \rho_k = n$ και:

$$J_{n_i}(0) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

Απόδειξη. Σύμφωνα με το Θεώρημα 7.2 υπάρχει βάση $\mathcal{B}$ του $\mathcal{E}$ της μορφής

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$$

όπου

$$\mathcal{B}_i = \left\{ f^{\rho_i - 1}(\vec{z}_i),\ f^{\rho_i - 2}(\vec{z}_i),\ \cdots,\ f(\vec{z}_i),\ \vec{z}_i \right\}, \qquad 1 \le i \le k$$

και

$$f^{\rho_i}(\vec{z}_i) = \vec{0}, \quad 1 \le i \le k$$

Είναι προφανές από τις παραπάνω σχέσεις ότι ο πίνακας της f στην παραπάνω βάση $\mathcal{B}$ είναι ο ζητούμενος.

Αναλυτικότερα: θέτουμε $\mathcal{V}_i$ να είναι ο υπόχωρος του $\mathcal{E}$ ο οποίος παράγεται από το σύνολο $\mathcal{V}_i$:

$$\mathcal{V}_i = \langle \mathcal{B}_i \rangle = \left\langle f^{\rho_i - 1}(\vec{z}_i),\ f^{\rho_i - 2}(\vec{z}_i),\ \cdots,\ f(\vec{z}_i),\ \vec{z}_i \right\rangle, \quad 1 \le i \le k$$

Παρατηρούμε ότι το σύνολο $\mathcal{B}_i$ είναι βάση του $\mathcal{V}_i$ και επιπλέον $f(\mathcal{V}_i) \subseteq \mathcal{V}_i$. Επομένως η $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ επάγει για κάθε $i = 1, 2, \cdots, k$, γραμμικές απεικονίσεις

$$f_i := f|_{\mathcal{V}_i} : \mathcal{V}_i \longrightarrow \mathcal{V}_i, \quad f_i(\vec{x}) = f(\vec{x}), \quad \forall\, \vec{x} \in \mathcal{V}_i$$

Επειδή $f^{\rho_i}(\vec{z}_i) = f_i^{\rho_i}(\vec{z}_i) = \vec{0}$, έπεται ότι ο πίνακας της f_i στην βάση $\mathcal{B}_i$ είναι ο στοιχειώδης πίνακας Jordan $J_{\rho_i}(0)$:

$$M_{\mathcal{B}_i}^{\mathcal{B}_i}(f_i) = J_{\rho_i}(0)$$

Επειδή το σύνολο $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$ είναι βάση του $\mathcal{E}$, έπεται ότι το άθροισμα υπόχωρων $\mathcal{V}_1 + \mathcal{V}_2 + \cdots + \mathcal{V}_k$ είναι ευθύ και μας δίνει τον $\mathcal{E}$:

$$\mathcal{E} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_k$$

και η γαρμμική απεικόνιση f είναι το ευθύ άθροισμα των f_i:

$$f = f_1 \oplus f_2 \oplus \cdots \oplus f_k$$

Τότε όμως ο πίνακας $M_{\mathcal{B}}^{\mathcal{B}}(f)$ της f στην βάση $\mathcal{B}$ είναι το ευθύ άθροισμα των πινάκων $M_{\mathcal{B}_i}^{\mathcal{B}_i}(f_i) = J_{\rho_i}(0)$ της f_i στην βάση $\mathcal{B}_i$ του $\mathcal{V}_i$:

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = J_{\rho_1}(0) \oplus J_{\rho_2}(0) \oplus \cdots \oplus J_{\rho_k}(0)$$ □

7.2. **Μηδενοδύναμοι Πίνακες.** Μια άμεση συνέπεια του Πορίσματος 7.3 είναι το ακόλουθο σημαντικό αποτέλεσμα:

Θεώρημα 7.4. *Κάθε μηδενοδύναμος $n \times n$ πίνακας A είναι όμοιος με έναν πίνακα Jordan της μορφής*

$$J = J_{\rho_1}(0) \oplus J_{\rho_2}(0) \oplus \cdots \oplus J_{\rho_k}(0), \qquad \rho_1 + \rho_2 + \cdots + \rho_k = n$$

Απόδειξη. Προκύπτει άμεσα από το Πόρισμα 7.3 αν το τελευταίο εφαρμοσθεί στην γραμμική απεικόνιση

$$f_A : \mathbb{K}_n \longrightarrow \mathbb{K}_n, \quad f_A(X) = A \cdot X$$

η οποία είναι προφανώς μηδενοδύναμη. □

7.3. **Βάσεις Jordan.** Θεωρούμε την γραμμική απεικόνιση $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ και υποθέτουμε ότι όλες οι ιδιοτιμές της ανήκουν στο $\mathbb{K}$. Επομένως το χαρακτηριστικό πολυώνυμο της f γράφεται ως εξής:

$$P_f(t) = (-1)^n (t - \lambda_1)^{k_1} \cdot (t - \lambda_2)^{k_2} \cdot \quad \cdots \quad (t - \lambda_r)^{k_r}$$

όπου $\lambda_1, \cdots, \lambda_r$ είναι οι διακεκριμμένες ιδιοτιμές της f με αντίστοιχες πολλαπλότητες $k_1, \cdots, k_r$, όπου: $k_1 + \cdots + k_r = n = \mathsf{dim}_{\mathbb{K}}\, \mathcal{E}$.

Θεώρημα 7.5. *Υπάρχει μια βάση $\mathcal{J}$ του $\mathcal{E}$ στην οποία ο πίνακας της f είναι ένας πίνακας Jordan της μορφής:*

$$M_{\mathcal{J}}^{\mathcal{J}}(f) = J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \cdots \oplus J_{n_r}(\lambda_r) \qquad \dim_{\mathbb{K}} \mathcal{E} = n_1 + n_2 + \cdots + n_r$$

όπου $\lambda_1, \cdots, \lambda_r$ είναι οι διακεκριμμένες ιδιοτιμές της f, και κάθε πίνακας Jordan $J_{n_i}(\lambda_i)$ είναι ευθύ άθροισμα στοιχειωδών πινάκων Jordan κατάλληλων μεγεθών ως προς την ιδιοτιμή λ_i:

$$J_{n_i}(\lambda_i) = J_{n_{i_1}}(\lambda_i) \oplus \cdots \oplus J_{n_{i_k}}(\lambda_i), \qquad n_i = n_{i_1} + \cdots n_{i_k}, \qquad 1 \leq i \leq r$$

Η βάση $\mathcal{J}$ καλείται **βάση Jordan** *του $\mathcal{E}$ για την f.*

Απόδειξη. Επειδή η f έχει όλες τις ιδιοτιμές της στο $\mathbb{K}$, έπεται ότι το χαρακτηριστικό πολυώνυμο της αναλύεται όπως παραπάνω σε γινόμενο (δυνάμεων) πρωτοβαθμίων παταγόντων:

$$P_f(t) = (-1)^n (t-\lambda_1)^{k_1} \cdot (t-\lambda_2)^{k_2} \cdot \quad \cdots \quad (t-\lambda_r)^{k_r}$$

Θέτουμε:

$$\mathcal{N}_i := \big\{ \vec{x} \in \mathcal{E} \mid (f - \lambda_i \mathrm{Id}_{\mathcal{E}})^{k_i}(\vec{x}) = \vec{0} \big\}, \qquad 1 \ \leq \ i \ \leq \ r$$

Αν $\vec{x} \in \mathcal{N}_i$, δηλαδή $(f - \lambda_i \mathrm{Id}_{\mathcal{E}})^{k_i}(\vec{x}) = \vec{0}$, έστω $(f - \lambda_i \mathrm{Id}_{\mathcal{E}})(\vec{x}) = \vec{y}$. Τότε $(f - \lambda_i \mathrm{Id}_{\mathcal{E}})^{k_i}(\vec{y}) = (f - \lambda_i \mathrm{Id}_{\mathcal{E}})^{k_i+1}(\vec{x}) = (f - \lambda_i \mathrm{Id}_{\mathcal{E}})((f - \lambda_i \mathrm{Id}_{\mathcal{E}})^{k_i}(\vec{x})) = \vec{0}$. Επομένως $\vec{y} \in \mathcal{N}_i$ και επομένως η γραμμική απεικόνιση $f - \lambda_i \mathrm{Id}_{\mathcal{E}} \colon \mathcal{E} \longrightarrow \mathcal{E}$ επάγει μια γραμμική απεικόνιση

$$f_i := f - \lambda_i \mathrm{Id}_{\mathcal{E}} \ : \ \mathcal{N}_i \ \longrightarrow \ \mathcal{N}_i, \quad f_i(\vec{x}) = (f - \lambda_i \mathrm{Id}_{\mathcal{E}})(\vec{x}) = f(\vec{x}) - \lambda_i \vec{x}$$

<u>Ισχυρισμός:</u> Η γραμμική απεικόνιση $f_i \ : \ \mathcal{N}_i \ \longrightarrow \ \mathcal{N}_i$ είναι μηδενοδύναμη και:

$$\mathcal{E} = \mathcal{N}_1 \oplus \mathcal{N}_2 \oplus \cdots \oplus \mathcal{N}_r$$

Πράγματικά: το ότι η f_i είναι μηδενοδύναμη προκύπτει άμεσα από τον ορισμό του υπόχωρου $\mathcal{N}_i$. Το ότι ο $\mathcal{E}$ είναι το ευθύ άθροισμα των υπόχωρων $\mathcal{N}_i$ προκύπτει ακριβώς όπως στην Πρόταση 4.2, βλέπε και Παρατήρηση 4.3, διότι επειδή οι ιδιοτιμές $\lambda_1, \cdots, \lambda_r$ είναι διακεκριμμένες, ο μέγιστος κοινός διαιρέτης των πολυωνύμων $(t-\lambda_1)^{k_1}, \cdots, (t-\lambda_r)^{k_r}$ είναι το σταθερό πολυώνυμο 1.

Επειδή η γραμμική επεικόνιση $f_i \colon \mathcal{N}_i \longrightarrow \mathcal{N}_i$ είναι μηδενοδύναμη, από το Θεώρημα 7.3 έπεται ότι υπάρχει βάση $\mathcal{B}_i$ του $\mathcal{N}_i$ στην οποία ο πίνακας της f_i είναι ένας πίνακας Jordan της μορφής

$$M^{\mathcal{B}_i}_{\mathcal{B}_i}(f_i) = M^{\mathcal{B}_1}_{\mathcal{B}_1}(f_i + \lambda_i \mathrm{Id}_{\mathcal{N}_i}) := J_i = J_{\rho_{i1}}(0) \oplus J_{\rho_{i2}}(0) \oplus \cdots \oplus J_{\rho_{ik}}(0)$$

όπου $\rho_{i1} + \rho_{i2} + \cdots + \rho_{ik} = \dim_{\mathbb{K}} \mathcal{N}_i := n_1$ και:

$$J_{ij}(0) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}, \qquad 1 \leq j \leq k$$

Επειδή $\mathcal{E} = \mathcal{N}_1 \oplus \mathcal{N}_2 \oplus \cdots \oplus \mathcal{N}_r$, έπεται ότι το σύνολο

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$$

θα είναι μια βάση του $\mathcal{E}$.

Επειδή $f_i = f - \lambda_i \mathrm{Id}_{\mathcal{E}}$, έπεται ότι ο περιορισμός $f|_{\mathcal{N}_i}$ της f στον υπόχωρο $\mathcal{N}_i$ θα είναι της μορφής $f|_{\mathcal{N}_i} = f_i + \lambda_i \mathrm{Id}_{\mathcal{N}_i}$.

Παρατηρούμε ότι ο πίνακας της $f_i + \lambda_i \mathrm{Id}_{\mathcal{N}_i}$ στην βάση $\mathcal{B}_i$ του $\mathcal{N}_i$ είναι της μορφής

$$M^{\mathcal{B}_i}_{\mathcal{B}_i}(f_i + \lambda_i \mathrm{Id}_{\mathcal{N}_i}) := J_i(\lambda_i) = J_{\rho_{i1}}(\lambda_i) \oplus J_{\rho_{i2}}(\lambda_i) \oplus \cdots \oplus J_{\rho_{ik}}(\lambda_i)$$

όπου $\rho_{i1} + \rho_{i2} + \cdots + \rho_{ik} = \dim_{\mathbb{K}} \mathcal{N}_i = n_i$ και:

$$J_{ij}(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \lambda_i \end{pmatrix}, \qquad 1 \leq j \leq k$$

Λαμβάνοντας υπ' όψιν τα παραπάνω, θα έχουμε ότι η f είναι το ευθύ άθροισμα

$$f = f|_{\mathcal{N}_1} \oplus f|_{\mathcal{N}_2} \oplus \cdots \oplus f|_{\mathcal{N}_r} = (f_1 + \lambda_1 \mathrm{Id}_{\mathcal{N}_1}) \oplus (f_2 + \lambda_2 \mathrm{Id}_{\mathcal{N}_2}) \oplus \cdots \oplus (f_r + \lambda_r \mathrm{Id}_{\mathcal{N}_r})$$

και άρα ο πίνακας της f στην βάση $\mathcal{B}$ θα είναι της μορφής:

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = M_{\mathcal{B}_1}^{\mathcal{B}_1}(f_1 + \lambda_1 \mathrm{Id}_{\mathcal{N}_1}) \oplus M_{\mathcal{B}_1}^{\mathcal{B}_1}(f_2 + \lambda_2 \mathrm{Id}_{\mathcal{N}_2}) \oplus \cdots \oplus M_{\mathcal{B}_1}^{\mathcal{B}_1}(f_r + \lambda_r \mathrm{Id}_{\mathcal{N}_r})$$

δηλαδή

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \cdots \oplus J_{n_r}(\lambda_r)$$

ο οποίος είναι ένας πίνακας Jordan. Αυτό σημαίνει ότι η βάση $\mathcal{B}$ του $\mathcal{E}$ είναι η επιθυμητή βάση Jordan $\mathcal{J}$ του $\mathcal{E}$ για την f. □

7.4. **Η Κανονική Μορφή Jordan ενός πίνακα.** Μια άμεση συνέπεια του Πορίσματος 7.5 είναι το ακόλουθο σημαντικό αποτέλεσμα:

Θεώρημα 7.6. *Κάθε $n \times n$ πίνακας $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ο οποίος έχει όλες τις διακεκριμμένες ιδιοτιμές του $\lambda_1, \cdots, \lambda_k$ στο $\mathbb{K}$ είναι όμοιος με έναν πίνακα Jordan της μορφής*

$$J = J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda) \oplus \cdots \oplus J_{n_k}(\lambda_k), \qquad n_1 + n_2 + \cdots + n_k = n$$

όπου κάθε πίνακας Jordan $J_{n_i}(\lambda_i)$, $1 \leq i \leq k$, είναι ευθύ άθροισμα στοιχειωδών πινάκων Jordan κατάλληλων μεγεθών που αντιστοιχούν στην ιδιοτιμή λ_i:

$$J_{n_i}(\lambda_i) = J_{n_{i_1}}(\lambda_i) \oplus \cdots \oplus J_{n_{i_r}}(\lambda_i), \qquad n_i = n_{i_1} + \cdots n_{i_r}$$

Απόδειξη. Προκύπτει άμεσα από το Πόρισμα 7.3 αν το τελευταίο εφαρμοσθεί στην γραμμική απεικόνιση

$$f_A : \mathbb{K}_n \longrightarrow \mathbb{K}_n, \quad f_A(X) = A \cdot X$$

η οποία είναι προφανώς έχει όλες τις ιδιοτιμές της (= ιδιοτιμές του A) στο σώμα $\mathbb{K}$. □

7.5. **Αλγόριθμος Εύρεσης Κανονικής Μορφής Jordan.**

- Έστω $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ ένας $n \times n$-πίνακας με στοιχεία απο το σώμα $\mathbb{K}$.
- Υποθέτουμε ότι όλες οι ιδιοτιμές του A ανήκουν στο σώμα $\mathbb{K}$.

Τότε υπάρχει αντιστρέψιμος πίνακας P έτσι ώστε ο πίνακας

$$P^{-1} \cdot A \cdot P = J$$

να είναι η κανονική μορφή Jordan J του A.

- Περιγράφουμε έναν αλγόριθμο εύρεσης της κανονικής μορφής Jordan J του πίνακα A:

Βήμα 1. *Υπολογίζουμε το χαρακτηριστικό πολυώνυμο $P_A(t)$ και βρίσκουμε το σύνολο*

$$\mathsf{Spec}(A) := \{\lambda_1, \lambda_2, \cdots, \lambda_k\}$$

των διακεκριμένων ιδιοτιμών του A.

Βήμα 2. *Για κάθε ιδιοτιμή $\lambda \in \mathsf{Spec}(A)$, υπολογίζουμε τις βαθμίδες των πινάκων:*

$$(A - \lambda \cdot I_n), \quad (A - \lambda \cdot I_n)^2, \ \cdots, \quad (A - \lambda \cdot I_n)^n$$

Βήμα 3. *Για κάθε ιδιοτιμή* $\lambda \in \mathsf{Spec}(A)$ *υπολογίζουμε τον ελάχιστο φυσικό αριθμό* $s := s(\lambda)$ *για τον οποίο ισχύει:*

$$\mathbf{r}(A - \lambda \cdot I_n)^s = \mathbf{r}(A - \lambda \cdot I_n)^{s+1}$$

και θέτουμε:

$$q_m := \mathbf{r}(A - \lambda \cdot I_n)^{m-1} - \mathbf{r}(A - \lambda \cdot I_n)^m, \quad \forall m = 1, 2, \cdots, s+1.$$

Σημειώνουμε ότι:

$$0 \ = \ q_{s+1} \ < \ q_s \ \leq \ q_{s-1} \ \leq \ \cdots \ \leq q_1 \ = \ n - \mathbf{r}(A - \lambda \cdot I_n)$$

Βήμα 4. *Για κάθε ιδιοτιμή* $\lambda \in \mathsf{Spec}(A)$ *και για κάθε* $m = 1, 2, \cdots, s(\lambda)$, *θεωρούμε:*

$q_m - q_{m+1}$ *στοιχειώδεις πίνακες Jordan οι οποίοι είναι* $m \times m$ *και έχουν διαγώνιο στοιχείο τον αριθμό* λ.

Βήμα 5. *Για κάθε ιδιοτιμή* $\lambda \in \mathsf{Spec}\, A$, *θεωρούμε το ευθύ άθροισμα των στοιχειωδών πινάκων Jordan του Βήματος* 4. *Ο πίνακας που προκύπτει είναι η κανονική μορφή Jordan του* A. *(Ως συνήθως διατάσσουμε Jordan blocks έτσι ώστε τα blocks με μεγαλύτερη τάξη να εμφανίζονται πρώτα).*

Παράδειγμα 7.7. *Θεωρούμε τον* 3×3 *πίνακα πραγματικών αριθμών*

$$A = \begin{pmatrix} 0 & 4 & 2 \\ -3 & 8 & 3 \\ 4 & -8 & -2 \end{pmatrix}$$

Το χαρακτηριστικό πολυώνυμο του A *είναι*

$$P_A(t) = (t-2)^3$$

και άρα η μόνη ιδιοτιμή του A *είναι η* $\lambda = 2$ *πολλαπλότητας* 3. *Επομένως*

$$\mathsf{Spec}(A) = \{2\}$$

(1) *Θα έχουμε:*

$$A - 2 \cdot I_3 = \begin{pmatrix} -2 & 4 & 2 \\ -3 & 6 & 3 \\ 4 & -8 & -4 \end{pmatrix}$$

και άρα όπως μπορούμε να δούμε εύκολα

$$\mathbf{r}(A - 2 \cdot I_3) = 1$$

(2) *Θα έχουμε*

$$(A - 2 \cdot I_3)^2 = 0$$

και άρα

$$\mathbf{r}(A - 2 \cdot I_3)^2 = \mathbf{r}(A - 2 \cdot I_3)^3 = 0$$

(3) *Επομένως*

$$s = 2, \qquad q_1 = 3 - 1 = 2$$

και

$$q_2 = \mathbf{r}(A - 2 \cdot I_3) - \mathbf{r}(A - 2 \cdot I_3)^2 = 1 - 0 = 1 \qquad \text{και} \qquad q_3 = 0$$

(4) *Για τον πίνακα* A *θα έχουμε:*
- $q_1 - q_2 = 1 - 0 = 1$ *στοιχειώδη πίνακα Jordan ο οποίος θα είναι* 1×1 *και θα έχει το* $\lambda = 2$ *στην διαγώνιο.*
- $q_2 - q_3 = 1 - 0 = 1$ *στοιχειώδη πίνακα Jordan ο οποίος θα είναι* 2×2 *και θα έχει το* $\lambda = 2$ *στην διαγώνιο.*

- $q_3 - q_4 = 0$ *στοιχειώδη πίνακα Jordan ο οποίος θα είναι 3×3 και θα έχει το $\lambda = 2$ στην διαγώνιο.*

Επομένως η κανονική μορφή Jordan του A θα είναι ο πίνακας

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

ο οποίος είναι το ευθύ άθροισμα των στοιχειωδών πινάκων Jordan:

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \oplus \begin{pmatrix} 2 \end{pmatrix}$$

7.6. **Αλγόριθμος Εύρεσης Αντιστρέψιμου Πίνακα P έτσι ώστε ο πίνακας $P^{-1}AP$ να είναι η Κανονική Μορφή Jordan του A.**

- Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας $n \times n$-πίνακας με στοιχεία απο το σώμα $\mathbb{K}$.

- Υποθέτουμε ότι όλες οι ιδιοτιμές του A ανήκουν στο σώμα $\mathbb{K}$.

Τότε υπάρχει αντιστρέψιμος πίνακας P έτσι ώστε ο πίνακας

$$P^{-1} \cdot A \cdot P$$

να είναι η κανονική μορφή Jordan του A.

- Περιγράφουμε έναν αλγόριθμο εύρεσης του αντιστρέψιμου πίνακα P:

Βήμα 6. *Για κάθε ιδιοτιμή $\lambda \in \mathsf{Spec}\, A$ $(s = s(\lambda))$:*

(1) *Βρίσκουμε q_s γραμμικά ανεξαρτητες λύσεις του συστήματος εξισώσεων:*

$$(A - \lambda \cdot I_n)^s \cdot X = 0$$

έτσι ώστε να ισχύει:

$$(A - \lambda \cdot I_n)^{s-1} \cdot X \neq 0$$

(2) *Βρίσκουμε $q_{s-1} - q_s$ γραμμικά ανεξαρτητες λύσεις του συστήματος εξισώσεων:*

$$(A - \lambda \cdot I_n)^{s-1} \cdot X = 0$$

έτσι ώστε

$$(A - \lambda \cdot I_n)^{s-2} \cdot X \neq 0$$

έτσι ώστε αυτές οι λύσεις μαζί με τις q_s λυσεις του (1) να αποτελούν γραμμικά ανεξάρτητο σύνολο.

(3) *Συνεχίζουμε αυτή την διαδικασία και τελικά βρίσκουμε $q_1 - q_2$ γραμμικά ανεξάρτητες λύσεις του συστήματος*

$$(A - \lambda \cdot I_n) \cdot X = 0$$

έτσι ώστε αυτές οι λύσεις μαζί με τις q_2 λύσεις που βρέθηκαν προηγούμενα να αποτελούν γραμμικά ανεξάρτητο σύνολο.

Βήμα 7. *Έστω $\{X_1, X_2, \cdots, X_k\}$ οι λύσεις που βρέθηκαν στο Βήμα 6. Εκ' κατασκευής οι λύσεις αυτές είναι γενικευμένα ιδιοδιανύσματα του A που αντιστοιχούν σε κάποια ιδιοτιμή λ, δηλαδή ικανοποιούν την σχέση*

$$(A - \lambda \cdot I_n)^t \cdot X = 0 \qquad \text{για κάποιο} \quad t \geq 0$$

Για κάθε $i = 1, 2, \cdots k$, Βρίσκουμε τον φυσικό αριθμό

$$m_i := \min\{t \geq 0 \mid (A - \lambda \cdot I_n)^t \cdot X = 0\}$$

και θεωρούμε τον πίνακα

$$P_i := \big((A-\lambda\cdot I_n)^{m_i-1}\cdot X_i,\ \cdots\ (A-\lambda\cdot I_n)\cdot X_i,\ X_i\big)$$

δηλαδή ο πίνακας P_i έχει σαν στήλες τα διανύσματα

$$(A-\lambda\cdot I_n)^r\cdot X_i,\quad 0\le r\le m_i-1$$

Βήμα 8. *Ο αντιστρέψιμος πίνακας P έτσι ώστε ο πίνακας $P^{-1}AP$ να ειναι η Κανονικη Μορφη Jordan του A, είναι τότε ο πίνακας:*

$$P = \big(P_1\ P_2\ \cdots\ P_k\big).$$

Παράδειγμα 7.8. *Θεωρούμε τον 4×4 πίνακα πραγματικών αριθμών*

$$A=\begin{pmatrix}5 & 1 & -2 & 4\\ 0 & 5 & 2 & 2\\ 0 & 0 & 5 & 3\\ 0 & 0 & 0 & 4\end{pmatrix}$$

(1) *Υπολογίζουμε το χαρακτηριστικό πολυώνυμο του A:*

$$P_A(t) = (t-5)^3(t-4)$$

(2) *Άρα θα έχουμε:*

$$\mathsf{Spec}(A)=\{5,4\}$$

(α') *Για την ιδιοτιμή $\lambda=5$, έχουμε:*

$$s=3,\ \ q_3=1,\ \ q_2=1,\ \ q_1=1$$

Άρα υπάρχει ένα Jordan block τάξης 3.

(β') *Για την $\lambda=4$, έχουμε:*

$$s=1,\ \ q_1=1$$

Άρα υπάρχει ένα Jordan block τάξης 1.

(3) *Η κανονική μορφή Jordan είναι η*

$$\begin{pmatrix}5 & 1 & 0 & 0\\ 0 & 5 & 1 & 0\\ 0 & 0 & 5 & 0\\ 0 & 0 & 0 & 4\end{pmatrix} = \begin{pmatrix}5 & 1 & 0\\ 0 & 5 & 1\\ 0 & 0 & 5\end{pmatrix}\oplus\big(4\big)$$

(4) *Ακολουθώντας την διαδικασία των Βημάτων $6,7,8$ βρίσκουμε ότι:*

$$P=\begin{pmatrix}2 & -1 & 0 & -14\\ 0 & 2 & 0 & 4\\ 0 & 0 & 1 & -3\\ 0 & 0 & 0 & 1\end{pmatrix}.$$

Άσκηση 7.9. *Να βρεθεί η κανονική μορφή Jordan των ακόλουθων πινάκων πραγματικών αριθμών:*

$$A=\begin{pmatrix}1 & -3 & 0 & 3\\ -2 & -6 & 0 & 13\\ 0 & -3 & 1 & 3\\ -1 & -4 & 0 & 8\end{pmatrix}\qquad \textit{και} \qquad B=\begin{pmatrix}3 & -1 & 0 & 0\\ 1 & 1 & 0 & 0\\ 3 & 0 & 5 & -3\\ 4 & -1 & 3 & -1\end{pmatrix}$$

Άσκηση 7.10. *Να δείξετε ότι η κανονική μορφή Jordan του πίνακα πραγματικών αριθμών:*

$$A = \begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$$

είναι (ως ευθύ άθροισμα στοιχειωδών πινάκων Jordan):

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

Άσκηση 7.11. *Θεωρούμε τον ακόλουθο πίνακα πραγματικών αριθμών:*

$$A = \begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}$$

Να βρεθεί αντιστρέψιμος πίνακας P έτσι ώστε ο πίνακας $P^{-1} \cdot A \cdot P$ να είναι η κανονική μορφή Jordan του A.

8. **Εφαρμογές της Κανονικής Μορφής Jordan**

Στην παρούσα παράγραφο θα δούμε κάποιες εφαρμογές της κανονικής μορφής Jordan. Οι πίνακες $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ με τους οποίους θα ασχοληθούμε θα έχουν όλες οι ιδιοτιμές τους στο σώμα $\mathbb{K}$. Αν $\mathbb{K} = \mathbb{C}$, τότε ως γνωστόν όλοι οι πίνακες $A \in \mathsf{M}_{n\times n}(\mathbb{C})$ έχουν αυτή την ιδιότητα.

8.1. **Κάθε τετραγωνικός πίνακας είναι όμοιος με τον ανάστροφό του.**

Ώς γνωστόν κάθε τετραγωνικός πίνακας A έχει το ίδιο χαρακτηριστικό πολυώνυμο με τον ανάστροφό του ${}^tA$, και άρα οι πίνακες A και ${}^tA$ έχουν τις ίδιες ιδιοτιμές.

Το ακόλουθο αποτέλεσμα δείχνει ότι οι πίνακες A και ${}^tA$ είναι όμοιοι, με την προυπόθεση ότι οι ιδιοτιμές τους ανήκουν στο $\mathbb{K}$.

Θεώρημα 8.1. *Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας τετραγωνικός πίνακας, και υποθέτουμε ότι όλες οι ιδιοτιμές του A ανήκουν στο $\mathbb{K}$. Τότε ο A είναι όμοιος με το ανάστροφό του ${}^tA$.*

Απόδειξη. Ώς γνωστόν οι πίνακες A και ${}^tA$ έχουν τις ίδιες ιδιοτιμές. Άρα όλες οι ιδιοτιμές του ${}^tA$ ανήκουν στο σώμα $\mathbb{K}$. Τότε οι πίνακες A και ${}^tA$ είναι όμοιοι με πίνακες Jordan J_1 και J_2 αντίστοιχα:

$$P^{-1} \cdot A \cdot P = J_1 \qquad \text{και} \qquad Q^{-1} \cdot {}^tA \cdot Q = J_2$$

για κατάλληλους αντιστρέψιμους πίνακες P και Q. Τότε:

$$P^{-1} \cdot A \cdot P = J_1 \quad \Longrightarrow \quad A = P \cdot J_1 \cdot P^{-1} \quad \Longrightarrow \quad {}^tA = {}^t(P^{-1}) \cdot {}^tJ_1 \cdot {}^tP = ({}^tP)^{-1} \cdot {}^tJ_1 \cdot {}^tP$$

$$Q^{-1} \cdot {}^tA \cdot Q = J_2 \quad \Longrightarrow \quad {}^tA = Q \cdot J_2 \cdot Q^{-1}$$

Άρα:

$$({}^tP)^{-1} \cdot {}^tJ_1 \cdot {}^tP = Q \cdot J_2 \cdot Q^{-1} \quad \Longrightarrow \quad J_2 = Q^{-1} \cdot ({}^tP)^{-1} \cdot {}^tJ_1 \cdot {}^tP \cdot Q$$

και επομένως οι πίνακες ${}^tJ_1$ και J_2 είναι όμοιοι:

$$J_2 = ({}^tP \cdot Q)^{-1} \cdot {}^tJ_1 \cdot ({}^tP \cdot Q) \tag{8.1}$$

Για κάθε στοιχειώδη πίνακα Jordan

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \lambda \end{pmatrix}$$

ο οποίος εμφανίζεται στον πίνακα Jordan J_1 και αντιστοιχεί σε μια ιδιοτιμή λ του A, άρα και του ${}^tA$, θεωρούμε τον $m \times m$ πίνακα

$$T_m = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & ⋰ & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

Εύκολα υπολογίζουμε ότι $T_m^2 = T_m$ και άρα $T_m^{-1} = T_m$. Επιπλέον:

$$T_m^{-1} \cdot J_m(\lambda) \cdot T^m = {}^t J_m(\lambda) = \begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & \lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & \lambda & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & \lambda \end{pmatrix}$$

Επομένως ο στοιχειώδης πίνακας Jordan $J_m((\lambda)$ είναι όμοιος με τον ανάστροφό του. Θεωρώντας το ευθύ άθροισμα T πινάκων της μορφής T_m για κατάλληλα μεγέθη $m \geq 1$, έπεται ότι ο πίνακας Jordan J_1 είναι όμοιος με τον ανάστροφό του ${}^t J_1$: υπαρχει αντιστρέψιμος πίνακας R έτσι ώστε:

$$T^{-1} \cdot {}^t J_1 \cdot T = J_1 \qquad \Longrightarrow \qquad {}^t J_1 = T \cdot J_1 \cdot T^{-1} \tag{8.2}$$

Επειδή η σχέση ομοιότητας πινάκων είναι μεταβατική, από τις (8.1) και (8.2) έπεται ότι οι πίνακες Jordan J_1 και J_2 είναι όμοιοι. Ακλριβέστερα θα έχουμε:

$$J_2 = (T^{-1} \cdot {}^t P \cdot Q)^{-1} \cdot J_1 \cdot (T^{-1} \cdot {}^t P \cdot Q)$$

Έτσι:

$${}^t A \ \sim \ J_2 \quad \text{και} \quad J_2 \ \sim \ J_1 \quad \text{και} \quad J_1 \ \sim \ A$$

όπου η σχέση $\sim$ υποδηλώνει σχέση ομοιότητας. Λόγω μεταβατικότητας της σχέσης ομοιότητας, έπεται τελικά ότι:

$$A \ \sim \ {}^t A$$

και άρα ο A είναι όμοιος με τον A. □

8.2. **Ανάλυση τετραγωνικού πίνακα σε γινόμενο δύο συμμετρικών πινάκων ένας εκ των οποίων είναι αντιστρέψιμος.**

Το βασικό αποτέλεσμα της παρούσης παραγράφου είναι το ακόλουθο:

Θεώρημα 8.2. *Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας τετραγωνικός πίνακας, και υποθέτουμε ότι όλες οι ιδιοτιμές του A ανήκουν στο $\mathbb{K}$. Τότε υπάρχουν συμμετρικοί πίνακες $B, C \in \mathsf{M}_{n\times n}(\mathbb{K})$ έτσι ώστε:*

$$A = B \cdot C$$

και ο πίνακας B είναι αντιστρέψιμος.

Απόδειξη. Επειδή όλες οι ιδιοτιμές του A ανήκουν στο σώμα $\mathbb{K}$, έπεται ότι ο A είναι όμοιος με έναν πίνακα Jordan J:

$$P^{-1} \cdot A \cdot P = J \qquad \Longrightarrow \qquad A = P \cdot J \cdot P^{-1}$$

Τότε θα έχουμε:

$${}^t A = {}^t P^{-1} \cdot {}^t J \cdot {}^t P$$

Αν $\lambda_1, \cdots, \lambda_k$ είναι οι διακεκριμμένες ιδιοτιμές του A, θα έχουμε:

$$J = J_{n_1}(\lambda_1) \oplus \cdots \oplus J_{n_k}(\lambda_k)$$

και κάθε πίνακας Jordan $J_{n_i}(\lambda_i)$, $1 \leq i \leq k$, είναι ευθύ άθροισμα στοιχειωδών πινάκων Jordan κατάλληλων μεγεθών που αντιστοιχούν στην ιδιοτιμή λ_i:

$$J_{n_i}(\lambda_i) = J_{n_{i_1}}(\lambda_i) \oplus \cdots \oplus J_{n_{i_r}}(\lambda_i), \qquad n_i = n_{i_1} + \cdots n_{i_r}$$

Όπως και στην απόδειξη του Θεωρήματος 8.1, θέτοντας

$$T = T_{n_1} \oplus \cdots \oplus T_{n_k}, \qquad n = n_1 + n_2 + \cdots + n_k$$

όπου

$$T_{n_i} = T_{n_{i_1}} \oplus \cdots \oplus T_{n_{i_r}}, \qquad n_i = n_{i_1} + \cdots n_{i_r}$$

και

$$T_m = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \iddots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}, \qquad \forall m \geq 1$$

θα έχουμε:

$$J^t = T^{-1} \cdot J \cdot T$$

και άρα:

$${}^tA = {}^tP^{-1} \cdot {}^tJ \cdot {}^tP = {}^tP^{-1} \cdot T^{-1} \cdot J \cdot T \cdot {}^tP = {}^tP^{-1} \cdot T^{-1}P^{-1} \cdot A \cdot P \cdot T \cdot {}^tP$$

και άρα:

$${}^tA = B^{-1} \cdot A \cdot B, \qquad \text{όπου} \qquad B := P \cdot T \cdot {}^tP$$

Παρατηρούμε ότι ο πίνακας B είναι συμμετρικός και αντιστρέψιμος. Πραγματικά επειδή ο P και ο T είναι αντιστρέψιμοι πίνακες και ο T είναι προφανώς συμμετρικός, θα έχουμε:

$${}^tB = {}^t(P \cdot T \cdot {}^tP) = {}^t({}^tP) \cdot {}^tT \cdot {}^tP = P \cdot T \cdot {}^tP$$

Επειδή ο B είναι συμμετρικός, έπεται ότι και ο B είναι συμμετρικός: ${}^t(B^{-1}) = B^{-1}$. Θέτοντας

$$C := B^{-1} \cdot A$$

θα έχουμε:

$${}^tC = {}^t(B^{-1} \cdot A) = {}^tA \cdot {}^t(B^{-1}) = B^{-1} \cdot A \cdot B \cdot B^{-1} = B^{-1} \cdot A = C$$

Επομένως θα έχουμε:

$$A = B \cdot C, \qquad B: \text{ συμμετρικός και αντιστρέψιμος} \quad \text{και} \qquad C: \text{ συμμετρικός}$$

□

8.3. **Ανάλυση τετραγωνικού πίνακα σε άθροισμα διαγωνοποιήσιμου και μηδενοδύναμου πίνακα.**

Θεώρημα 8.3. *(Ανάλυση Jordan) Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{K})$ ένας τετραγωνικός πίνακας, και υποθέτουμε ότι όλες οι ιδιοτιμές του A ανήκουν στο $\mathbb{K}$. Τότε ο A μπορεί να γραφεί ως εξής:*

$$A = \widetilde{\Delta} + \widetilde{N}$$

όπου:

(1) *ο $\widetilde{\Delta}$ είναι διαγωνοποιήσιμος.*
(2) *ο $\widetilde{N}$ είναι μηδενοδύναμος.*
(3) $\widetilde{\Delta} \cdot \widetilde{N} = \widetilde{N} \cdot \widetilde{\Delta}$, $\;A \cdot \widetilde{\Delta} = \widetilde{\Delta} \cdot A$, $\;A \cdot \widetilde{N} = \widetilde{N} \cdot A$.
(4) *Υπάρχουν πολυώνυμα $R(t), S(t) \in \mathbb{K}[t]$ έτσι ώστε:*

$$\widetilde{\Delta} = R(A) \qquad \text{και} \qquad \widetilde{N} = S(A)$$

Απόδειξη. Υποθέτουμε πτώτα ότι ο πίνακας A είναι ένας στοιχειώδης πίνακας Jordan $J_m(\lambda)$ μεγέθους $m \geq 1$ που αντιστοιχεί σε έναν αριθμό $\lambda \in \mathbb{K}$:

$$A = J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \lambda \end{pmatrix}$$

Τότε θέτοντας

$$\Delta_m(\lambda) = \lambda \cdot I_m = \begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \lambda \end{pmatrix} \qquad \text{και} \qquad N_m = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

θα έχουμε:

$$J_m(\lambda) = \Delta_m(\lambda) + N_m$$

και προφανώς:

(1) ο $\Delta_m(\lambda)$ είναι διαγώνιος (ακριβέστερα είναι βαθμωτός, δηλαδή βαθμωτό πολλαπλάσιο του I_m).
(2) ο είναι μηδενοδύναμος: $N_m^m = \mathbb{O}$.

Επιπλέον:

(1)

$$\Delta_m(\lambda) \cdot N_m = \lambda \cdot I_m \cdot N_m = \lambda \cdot N_m = \lambda \cdot N_m \cdot I_m = N_m \cdot \lambda \cdot I_m = N_m \cdot \Delta_m(\lambda)$$

(2)

$$J_m(\lambda) \cdot \Delta_m(\lambda) = (\Delta_m(\lambda) + N_m) \cdot \Delta_m = \Delta_m(\lambda)\Delta_m(\lambda) + N_m \cdot \Delta_m(\lambda) =$$
$$\Delta_m(\lambda)\Delta_m(\lambda) + \Delta_m(\lambda) \cdot N_m = \Delta_m(\lambda) \cdot (\Delta_m(\lambda) + N_m) = \Delta_m(\lambda) \cdot J_m(\lambda)$$

(3)

$$J_m(\lambda) \cdot N_m = (\Delta_m(\lambda) + N_m) \cdot N_m = \Delta_m(\lambda) \cdot N_m + N_m \cdot N_m =$$
$$N_m \cdot \Delta_m(\lambda) + N_m \cdot N_m = N_m \cdot (\Delta_m(\lambda) + N_m) = N_m \cdot J_m(\lambda)$$

Άρα οι πίνακες $\Delta_m(\lambda)$ και N_m μετατίθενται μεταξύ τους και με τον $A = J_m(\lambda)$.

Στην γενική περίπτωση, επειδή ο A έχει όλες τις ιδιοτιμές του στο $\mathbb{K}$, έπεται ότι ο A είναι όμοιος με ένας πίνακα Jordan J, δηλαδή υπάρχει αντιστρέψιμος πίνακας P έτσι ώστε:

$$P^{-1} \cdot A \cdot P = J$$

όπου, αν $\lambda_1, \cdots, \lambda_k$ είναι οι διακεκριμμένες ιδιοτιμές του A, θα έχουμε:

$$J = J_{n_1}(\lambda_1) \oplus \cdots \oplus J_{n_k}(\lambda_k)$$

και κάθε πίνακας Jordan $J_{n_i}(\lambda_i)$, $1 \leq i \leq k$, είναι ευθύ άθροισμα στοιχειωδών πινάκων Jordan κατάλληλων μεγεθών που αντιστοιχούν στην ιδιοτιμή λ_i:

$$J_{n_i}(\lambda_i) = J_{n_{i_1}}(\lambda_i) \oplus \cdots \oplus J_{n_{i_r}}(\lambda_i), \qquad n_i = n_{i_1} + \cdots n_{i_r}$$

Τότε θα έχουμε

$$A = P \cdot J \cdot P^{-1}$$

και κάθε στοιχειώδης πίνακα Jordan $J_{n_{i_s}}(\lambda_i)$, σύμφωνα με την παραπάνω ανάλυση, μπορεί να γραφεί ως άθροισμα:

$$J_{n_{i_s}}(\lambda_i) = \Delta_{n_{i_s}}(\lambda_i) + N_{n_{i_s}}$$

Θέτουμε:

$$\Delta_{n_i}(\lambda_i) = \Delta_{n_{i_1}}(\lambda_i) \oplus \cdots \oplus \Delta_{n_{i_r}}(\lambda_i) \qquad \text{και} \qquad \Delta = \Delta_{n_1}(\lambda_1) \oplus \cdots \oplus \Delta_{n_k}(\lambda_k)$$

$$N_{n_i} = N_{n_{i_1}} \oplus \cdots \oplus N_{n_{i_r}} \qquad \text{και} \qquad N = N_{n_1} \oplus \cdots \oplus N_{n_k}$$

Τότε, με χρήση των παραπάνω σχέσεων, εύκολα βλέπουμε ότι θα έχουμε:

$$J = \Delta + N, \quad \Delta \cdot N = N \cdot \Delta, \quad J \cdot N = N \cdot J, \qquad J \cdot \Delta = \Delta \cdot J$$

και ο πίνακας Δ είναι διαγωνοπιήσιμος (ως ευθύ άθροισμα βαθμωτών πινάκων) και ο πίνακας N είναι μηδενοδύναμος (ως ευθύ άθροισμα μηδενοδύναμων πινάκων).

Τότε όμως θα έχουμε:

$$A = P^{-1} \cdot J \cdot P = P^{-1} \cdot (\Delta + N) \cdot P = P^{-1} \cdot \Delta \cdot P + P^{-1} \cdot N \cdot P$$

Τέλος θέτοντας $\widetilde{\Delta} = P^{-1} \cdot \Delta \cdot P$ και $\widetilde{N} = P^{-1} \cdot N \cdot P$, έχουμε έναν διαγωνοποιήσιμο πίνακα $\widetilde{\Delta}$ και έναν μηδενοδύναμο πίνακα $\widetilde{N}$ έτσι ώστε:

$$A = \widetilde{\Delta} + \widetilde{N}$$

και

$$A \cdot \widetilde{\Delta} = \widetilde{\Delta} \cdot A, \qquad A \cdot \widetilde{N} = \widetilde{N} \cdot A, \qquad \widetilde{N} \cdot \widetilde{\Delta} = \widetilde{\Delta} \cdot \widetilde{N}$$

Τέλος έστω ένα πολυώνυμο $R(t) \in \mathbb{K}[t]$ με την ιδιότητα:

$$R(\lambda_i) = \lambda_i, \quad 1 \leq i \leq k, \qquad \text{και} \qquad R^{(m)}(\lambda_i) = 0, \qquad 1 \leq m \leq n_i$$

όπου n_i είναι η πολλαπλότητα της ιδιοτιμής λ_i (προφανώς τέτοια πολυώνυμα υπάρχουν), και εκ' κατασκευής το $R(t)$ έχει την ιδιότητα:

$$\widetilde{\Delta} = R(A)$$

Θέτοντας $S(t) = t - R(t)$ έχουμε ένα πολυώνυμο $S(t) \in \mathbb{K}[t]$ έτσι ώστε:

$$\widetilde{N} = S(A)$$

□

Ορισμός 8.4. *Η ανάλυση του πίνακα A:*

$$A = \widetilde{\Delta} + \widetilde{N}$$

στο Θεώρημα 8.3 καλείται **ανάλυση Jordan** *του A.*

Άσκηση 8.5. *Να δείξετε ότι η ανάλυση Jordan του A είναι μοναδική: Αν*

$$A = \Delta' + N'$$

όπου ο πίνακας Δ' είναι διαγωνοποιήσιμος και ο πίνακας N' είναι μηδενοδύναμος, και οι πίνακες Δ' και N' μετατίθενται μεταξύ τους (άρα και με τον A), τότε:

$$\widetilde{\Delta} = \Delta' \qquad \text{και} \qquad \widetilde{N} = N'$$

8.4. **Κριτήριο Ομοιότητας Πινάκων.** Το ακόλουθο σημαντικό Θεώρημα μας δίνει ένα χρήσιμο κριτήριο για το πότε δύο πίνακες είναι όμοιοι.

Θεώρημα 8.6. *Έστω A, B δύο τετραγωνικοί πίνακες με στοιχεία από το σώμα $\mathbb{K}$, και υποθέτουμε ότι οι A, B έχουν όλες τις ιδιοτιμές του στο $\mathbb{K}$. Τότε τα ακόλουθα είναι ισοδύναμα:*

(1) *Οι πίνακες A και B είναι όμοιοι.*

(2) (α') *Οι πίνακες A και B έχουν τις ίδιες ιδιοτιμές $\lambda_1, \lambda_2, \cdots, \lambda_k$.*

(β') *Για κάθε $1 \leq j \leq k$ και για κάθε $m \geq 1$:*

$$\mathbf{r}\big(A - \lambda_j I_n\big)^m \;=\; \mathbf{r}\big(B - \lambda_j I_n\big)^m$$

Απόδειξη. □

8.5. **Κριτήριο Διαγωνοποίησης Πινάκων.** Το ακόλουθο σημαντικό Θεώρημα μας δίνει ένα χρήσιμο κριτήριο για το πότε ένας πίνακας είναι διαγωνοποίησιμος.

Θεώρημα 8.7. *Έστω A ένας τετραγωνικός πίνακας με στοιχεία από το σώμα $\mathbb{K}$, και έστω $s = \deg Q_A(t)$ ο βαθμός του ελαχίστου πολυωνύμου $Q_A(t)$ του A. Τότε τα ακόλουθα είναι ισοδύναμα:*

(1) *Ο πίνακας A είναι διαγωνοποίησιμος.*

(2) *Όλες οι ιδιοτιμές του A ανήκουν στο σώμα $\mathbb{K}$, και:*

$$\mathsf{Det}\big(\mathsf{Tr}(A^{i+j})\big) \;\neq\; 0, \qquad\qquad 0 \leq i, j \leq s-1$$

9. **Η Κανονική Μορφή Jordan - II**

Στην παρούσα παράγραφό θα δούμε μια διαφορετική προσέγγιση στην Κανονική Μορφή Jordan.

9.1. **Αναλλοίωτοι και κυκλικοί υπόχωροι.**

Έστω ότι $\mathcal{E}$ είναι ένας $\mathbb{K}$-διανυσματικός χώρος υπεράνω ενός σώματος $\mathbb{K}$ και ότι $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι ένας $\mathbb{K}$-γραμμικός ενδομορφισμός του.

Ορισμός 9.1. *Ένας υπόχωρος $\mathcal{V}$ του $\mathcal{E}$ ονομάζεται f-αναλλοίωτος, αν $f(\mathcal{V}) \subseteq \mathcal{V}$.*

Αν $\mathcal{V}$ είναι ένας f-αναλλοίωτος υπόχωρος τού $\mathcal{E}$, τότε ο περιορισμός $f_{|\mathcal{V}} : \mathcal{V} \longrightarrow \mathcal{E}$ του f στον υπόχωρο $\mathcal{V}$ αποτελεί έναν ενδομορφισμό του $\mathcal{V}$, αφού $f(\mathcal{V}) \subseteq \mathcal{V}$, δηλαδή $f_{|\mathcal{V}} : \mathcal{V} \longrightarrow \mathcal{V}$.

Παράδειγμα 9.2. *Δοθέντος ενός ενδομορφισμού $f\colon \mathcal{E} \longrightarrow \mathcal{E}$, οι προφανείς υπόχωροι $\{\vec{0}\}$ και $\mathcal{E}$ είναι f-αναλλοίωτοι. Επιπλέον, για κάθε $n \in \mathbb{N}$, οι υπόχωροι $\mathsf{Ker}(f^n)$ και $\mathsf{Im}(f^n)$ είναι επίσης f-αναλλοίωτοι υπόχωροι του $\mathcal{E}$. (Γιατί;)*

Αν $\vec{v}$ είναι οποιοδήποτε διάνυσμα του $\mathcal{E}$, τότε ενδιαφερόμαστε για τον «μικρότερο»[2] f-αναλλοίωτο υπόχωρο του $\mathcal{E}$ που περιέχει το $\vec{v}$.

Λήμμα 9.3. *Ο υπόχωρος τού $\mathcal{E}$ που παράγεται από το σύνολο των διανυσμάτων*

$$\left\{f^i(\vec{v}) \mid i \in \mathbb{N} \cup \{0\}\right\} = \left\{\vec{v} = f^0(\vec{v}), f(\vec{v}), f^2(\vec{v}), \dots, f^n(\vec{v}), f^{(n+1)}(\vec{v}), \dots\right\}$$

είναι ο μικρότερος f-αναλλοίωτος υπόχωρος του $\mathcal{E}$ που περιέχει το διάνυσμα $\vec{v}$.

Απόδειξη. Αν $\vec{w}$ είναι ένα στοιχείο του υπόχωρου $\langle\{f^i(\vec{v}) \mid i \in \mathbb{N} \cup \{0\}\}\rangle$, τότε είναι ένας πεπερασμένος $\mathbb{K}$-γραμμικός συνδυασμός κάποιων δυνάμεων $f^i(\vec{v}), i \in \mathbb{N} \cup \{0\}$. Έστω ότι $\vec{w} = a_1 f^{i_1}(\vec{v}) + a_2 f^{i_2}(\vec{v}) + \cdots + a_s f^{i_s}(\vec{v}), a_j \in \mathbb{K}, \forall j, 1 \leq j \leq s$.

Εφαρμόζοντας στο $\vec{w}$ τον f προκύπτει το διάνυσμα:

$$f(\vec{w}) = f(a_1 f^{i_1}(\vec{v}) + a_2 f^{i_2}(\vec{v}) + \cdots + a_s f^{i_s}(\vec{v})) = a_1 f^{i_1+1}(\vec{v}) + a_2 f^{i_2+1}(\vec{v}) + \cdots + a_s f^{i_s+1}(\vec{v}),$$

το οποίο είναι γραμμικός συδυασμός των $f^{i_1+1}(\vec{v}), f^{i_2+1}(\vec{v}), \dots, f^{i_s+1}(\vec{v})$ και συνεπώς ανήκει στον $\langle\{f^i(\vec{v}) \mid i \in \mathbb{N} \cup \{0\}\}\rangle$. Επομένως, ο $\langle\{f^i(\vec{v}) \mid i \in \mathbb{N} \cup \{0\}\}\rangle$ είναι ένας f-αναλλοίωτος υπόχωρος του $\mathcal{E}$.

Αν τώρα, $\mathcal{V}$ είναι ένας f-αναλλοίωτος υπόχωρος που περιέχει το συγκεκριμένο διάνυσμα $\vec{v}$, τότε οφείλει να περιέχει και κάθε δύναμη $f^n(\vec{v}), n \in \mathbb{N} \cup \{0\}$, αφού είναι f-αναλλοίωτος. Συνεπώς, περιέχει και το σύνολο $\left\{f^i(\vec{v}) \mid i \in \mathbb{N} \cup \{0\}\right\}$, που είναι ακριβώς το σύνολο των γεννητόρων τού $\langle\{f^i(\vec{v}) \mid i \in \mathbb{N} \cup \{0\}\}\rangle$. Επομένως, $\langle\{f^i(\vec{v}) \mid i \in \mathbb{N} \cup \{0\}\}\rangle \subseteq \mathcal{V}$. Ώστε, ο $\langle\{f^i(\vec{v}) \mid i \in \mathbb{N} \cup \{0\}\}\rangle$ είναι ο μικρότερος f-αναλλοίωτος υπόχωρος του $\mathcal{E}$ που περιέχει το διάνυσμα $\vec{v}$. □

Ορισμός 9.4. *Έστω ότι $\mathcal{E}$ είναι ένας $\mathbb{K}$-διανυσματικός χώρος, ότι $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι ένας $\mathbb{K}$-γραμμικός ενδομορφισμός και ότι $\vec{v}$ είναι ένα διάνυσμα τού $\mathcal{E}$. Ο υπόχωρος $\langle\{f^i(\vec{v}) \mid i \in \mathbb{N} \cup \{0\}\}\rangle$ καλείται ο f-κυκλικός υπόχωρος που παράγεται από το ζεύγος $(f, \vec{v})$ και συμβολίζεται με $C(f; \vec{v})$.*

Λήμμα 9.5. *Έστω ότι $\mathcal{E}$ είναι ένας $\mathbb{K}$-διανυσματικός χώρος, ότι $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι ένας $\mathbb{K}$-γραμμικός ενδομορφισμός και ότι $\vec{v}$ είναι ένα μη μηδενικό διάνυσμα τού $\mathcal{E}$.*

Το $\vec{v}$ είναι ιδιοδιάνυσμα τού f, αν και μόνο αν, η $\mathbb{K}$-διάσταση τού f-κυκλικού υπόχωρου $C(f; \vec{v})$ ισούται με 1.

---

[2] ως προς τη σχέση του υποσυνόλου "$\subseteq$"

Απόδειξη. «$\Longrightarrow$» Αν το $\vec{v}$ είναι ένα ιδιοδιάνυσμα τού f που αντιστοιχεί στην ιδιοτιμή λ, τότε κάθε ένας από τους γεννήτορες $f^i(\vec{v}), i > 0$ του $C(f;\vec{v})$ ισούται με $f^i(\vec{v}) = \lambda^i\vec{v}$, αφού $f(\vec{v}) = \lambda\vec{v}$. Σύμφωνα με την υπόθεση, το $\vec{v} \neq \vec{0}$, επομένως, $\dim C(f;\vec{v}) = 1$.

«$\Longleftarrow$» Αν $\dim C(f;\vec{v}) = 1$, τότε το σύνολο που αποτελείται από οποιονδήποτε μη μηδενικό γεννήτορα τού $C(f;\vec{v})$ είναι και μια βάση του. Γι' αυτό το $\{\vec{v}\}$ είναι μια βάση τού $C(f;\vec{v})$. Το $f(\vec{v}) \in C(f;\vec{v})$ είναι γραμμικά εξαρτημένο από το $\vec{v}$. Συνεπώς, υπάρχει κάποιο $\lambda \in \mathbb{K}$ με $f(\vec{v}) = \lambda\vec{v}$ και γι'αυτό το $\vec{v}$ είναι ιδιοδιάνυσμα. □

Παρατήρηση 9.6. *Η διάσταση τού f-κυκλικού υπόχωρου $C(f;\vec{v})$ «μετρά» το πόσο «απέχει» ένα μη μηδενικό διάνυσμα $\vec{v}$ από το να είναι ιδιοδιάνυσμα.*

Ας στρέψουμε τώρα την προσοχή μας σε πεπερασμένης διάστασης διανυσματικούς χώρους.

Ας είναι $\mathcal{E}$ ένας $\mathbb{K}$–διανυσματικός χώρος με $\dim_{\mathbb{K}} \mathcal{E} = n \in \mathbb{N} \cup \{0\}$, ας είναι $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ ένας $\mathbb{K}$–γραμμικός ενδομορφισμός, ας είναι $\vec{v}$ ένα μη μηδενικό διάνυσμα τού $\mathcal{E}$ και τέλος ας είναι $C(f;\vec{v})$ ο αντίστοιχος f–κυκλικός υπόχωρος. Επειδή, $\dim_{\mathbb{K}} \mathcal{E} = n$, θα είναι επίσης $\dim_{\mathbb{K}} C(f;\vec{v}) \leq n$.

Υπενθυμίζουμε, ότι δοθέντος οποιουδήποτε πολυωνύμου $P(t) = a_0 + a_1 t + \cdots + a_s t^s \in \mathbb{K}[t]$ μπορούμε να σχηματίσουμε τον γραμμικό ενδομορφισμό $P(f) := a_0 \mathrm{Id}_{\mathcal{E}} + a_1 f + \cdots + a_s f^s$, όπου $\mathrm{Id}_{\mathcal{E}}$ είναι ο ταυτοτικός ενδομορφισμός τού $\mathcal{E}$. Εφαρμόζοντας τον $P(f)$ στο διάνυσμα $\vec{v}$, παίρνουμε $P(f)(\vec{v}) := a_0\vec{v} + a_1 f(\vec{v}) + \cdots + a_s f^s(\vec{v})$.

Παρατηρούμε ότι το σύνολο των πολυωνύμων $S_{(f;v)} = \{P(t) \in \mathbb{K}[t] \mid P(t) \neq 0, P(f)(\vec{v}) = \vec{0}\}$ είναι μη κενό. Πράγματι, το $S_{(f;v)}$ περιέχει το ελάχιστο πολυώνυμο $Q_f(t)$ τού f, επειδή το $Q_f(t) \neq 0$, ενώ ο ενδομορφισμός $Q_f(f)$ ισούται με τον μηδενικό ενδομορφισμό $\zeta_{\mathcal{E}}$ του $\mathcal{E}$ και γι'αυτό $Q_f(f)(\vec{v}) = \vec{0}$.

Όπως και στην περίπτωση τού ελαχίστου πολυωνύμου ενός ενδομορφισμού, μπορούμε να ορίσουμε το ελάχιστο πολυώνυμο τού f–κυκλικού υπόχωρου ως εξής:

(1) Μεταξύ των πολυωνύμων του $S_{(f;v)}$ θεωρούμε κάποιο πολυώνυμο με τον μικρότερο βαθμό.
(2) Πολλαπλασιάζοντας αυτό το πολυώνυμο με τον αντίστροφο τού συντελεστή τού μεγιστοβαθμίου όρου προκύπτει ένα πολυώνυμο που ανήκει στο $S_{(f;v)}$ που έχει τον μικρότερο βαθμό μεταξύ των πολυωνύμων τού $S_{(f;v)}$.
(3) Αποδεικνύεται ότι αυτό το πολυώνυμο είναι μοναδικό, το ονομάζουμε το *ελάχιστο πολυώνυμο* του f–κυκλικού υπόχωρου $C(f;\vec{v})$ και το συμβολίζουμε με $Q_{(f;v)}(t)$.
(4) Τέλος, αποδεικνύεται ακριβώς όπως και στην περίπτωση τού ελαχίστου πολυωνύμου ενός ενδομορφισμού, ότι κάθε πολυώνυμο τού $S_{(f;v)}$ είναι πολλαπλάσιο τού $Q_{(f;v)}(t)$.

Θεώρημα 9.7. *Έστω ότι $\mathcal{E}$ είναι ένας πεπερασμένης διάστασης $\mathbb{K}$-διανυσματικός χώρος, ότι $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι ένας $\mathbb{K}$-γραμμικός ενδομορφισμός, ότι $\vec{v}$ είναι ένα μη μηδενικό διάνυσμα τού $\mathcal{E}$ και ότι $C(f;\vec{v})$ είναι ο αντίστοιχος f-κυκλικός υπόχωρος με ελάχιστο πολυώνυμο το $Q_{(f;v)}(t)$.*

(1) *Ο βαθμός τού $Q_{(f;v)}(t)$ ισούται με $m \in \mathbb{N}$, αν και μόνο αν, το σύνολο $\{\vec{v} = f^0(\vec{v}), f(\vec{v}), f^2(\vec{v}), \ldots, f^{(m-1)}(\vec{v})\}$ είναι βάση τού f-κυκλικού υπόχωρου $C(f;\vec{v})$.*
(2) *Το ελάχιστο πολυώνυμο τού περιορισμού $f_{|C(f;\vec{v})}$ τού f στον f-κυκλικό υπόχωρο $C(f;\vec{v})$ συμπίπτει με το $Q_{(f;v)}(t)$.*

Απόδειξη. Παρατηρούμε ότι ένας πεπερασμένος $\mathbb{K}$–γραμμικός συνδυασμός

$$\vec{w} = \sum_{j=1}^{s} a_{i_j} f^{i_j}(\vec{v}), i_j \in \mathbb{N} \cup \{0\}, \quad \forall j, 1 \leq j \leq s \tag{*}$$

διανυσμάτων τού $C(f;\vec{v})$ προκύπτει ως εικόνα τού ενδομορφισμού $P(f)$ εφαρμοσμένη πάνω στο διάνυσμα $\vec{v}$, όπου $P(t)$ είναι το πολυώνυμο $\sum_{j=1}^{s} a_{i_j} t^{i_j}$, δηλαδή $\vec{w} = P(f)(\vec{v})$. (Υπενθυμίζουμε ότι $t^0 = 1_{\mathbb{K}}$.)

(1) «$\Longrightarrow$» Αν ήταν τα διανύσματα $\vec{v} = f^0(\vec{v}), f(\vec{v}), f^2(\vec{v}), \ldots, f^{(m-1)}(\vec{v})$, γραμμικά εξαρτημένα, τότε θα υπήρχαν $a_i \in \mathbb{K}, 0 \leq i \leq m-1$, όχι όλα ίσα με 0, τέτοια ώστε, $\sum_{i=0}^{m-1} a_i f^i(\vec{v}) = \vec{0}$. Τότε το μη μηδενικό πολυώνυμο $R(t) = \sum_{i=0}^{m-1} a_i t^i$, που έχει βαθμό γνήσια μικρότερο από τον βαθμό m του $Q_{(f;v)}(t)$, ικανοποιεί την $R(f)(\vec{v}) = \vec{0}$. Συνεπώς, το $R(t)$ ανήκει στο σύνολο $S_{(f;v)} = \{P(t) \in \mathbb{K}[t] \mid P(t) \neq 0, P(f)(\vec{v}) = \vec{0}\}$. Αυτό είναι άτοπο, αφού το $Q_{(f;v)}(t)$ είναι ένα πολυώνυμο που έχει τον μικρότερο βαθμό m μεταξύ των πολυωνύμων τού $S_{(f;v)}$. Ώστε το σύνολο $\mathcal{B} = \{\vec{v} = f^0(\vec{v}), f(\vec{v}), f^2(\vec{v}), \ldots, f^{(m-1)}(\vec{v})\}$ αποτελείται από γραμμικά ανεξάρτητα διανύσματα.

Θα δείξουμε ότι το $\mathcal{B}$ είναι επίσης σύνολο γεννητόρων του $C(f;\vec{v})$ και συνεπώς μια βάση του. Αν $\vec{w} \in C(f;\vec{v})$, τότε $\vec{w} = \sum_{j=1}^{s} a_{i_j} f^{i_j}(\vec{v})$, βλ. (*). Επομένως, $\vec{w} = P(f)(\vec{v})$ με $P(t) = \sum_{j=1}^{s} a_{i_j} t^{i_j}$. Εκτελώντας την ευκλείδεια διαίρεση $P(t)$ δια $Q_{(f;v)}(t)$ προκύπτει

$$P(t) = L(t)Q_{(f;v)}(t) + R(t), \text{ όπου ή } R(t) = 0 \text{ ή βαθμός } R(t) < m = \text{ βαθμός } Q_{(f;v)}(t).$$

Συνεπώς, το πολυώνυμο $R(t)$ έχει τη γενική μορφή $R(t) = a_0 + a_1 t + \cdots + a_{m-1} t^{m-1}$ με $a_i \in \mathbb{K}, \forall i, 0 \leq i \leq m-1$, όπου κάποια ή και όλα τα a_i είναι ίσα με το $0 \in \mathbb{K}$. Επιπλέον,

$$\begin{aligned}\vec{w} &= P(f)(\vec{v}) = L(f) \circ Q_{(f;v)}(f)(\vec{v}) + R(f)(\vec{v}) = L(f)(Q_{(f;v)}(f)(\vec{v})) + R(f)(\vec{v}) = L(f)(\vec{0}) + R(f)(\vec{v}) \\ &= a_0\vec{v} + a_1 f(\vec{v}) + \cdots + a_{m-1} f^{m-1}(\vec{v}).\end{aligned}$$

Επομένως, το $\vec{w}$ είναι ένας $\mathbb{K}$–γραμμικός συνδυασμός των στοιχείων τού συνόλου $\mathcal{B}$. Ώστε το $\mathcal{B}$ αποτελεί μια βάση τού $C(f;\vec{v})$.

(1) «$\Longleftarrow$» Επειδή το σύνολο $\mathcal{B} = \{\vec{v} = f^0(\vec{v}), f(\vec{v}), f^2(\vec{v}), \ldots, f^{(m-1)}(\vec{v})\}$ είναι βάση τού f–κυκλικού υπόχωρου $C(f;\vec{v})$, το σύνολο $\mathcal{B} \cup \{f^m(\vec{v})\} = \{\vec{v} = f^0(\vec{v}), f(\vec{v}), f^2(\vec{v}), \ldots, f^{(m-1)}(\vec{v}), f^m(\vec{v})\}$ αποτελείται από γραμμικώς εξαρτημένα διανύσματα. Επομένως, υπάρχουν $a_0, a_1, \ldots, a_m \in \mathbb{K}$, όχι όλα ίσα με μηδέν, τέτοια ώστε

$$a_0\vec{v} + a_1 f(\vec{v}) + a_2 f^2(\vec{v}) + \cdots + a_{m-1} f^{(m-1)}(\vec{v}) + a_m f^m(\vec{v}) = \vec{0}. \quad (**)$$

Επιπλέον, ο συντελεστής a_m οφείλει να μην ισούται με μηδέν, αφού αν $a_m = 0$, τότε από την (**) έπεται ότι το $\mathcal{B}$ αποτελείται απο γραμμικά εξαρτημένα διανύσματα, πράγμα άτοπο. Ώστε το πολυώνυμο $P(f) = a_0 + a_1 t + \cdots + a_m t^m$ έχει βαθμό m και ανήκει στο σύνολο $S_{(f;v)}$. Επομένως, το ελάχιστο πολυώνυμο $Q_{(f;v)}(t)$ έχει βαθμό $\leq m$. Αν ο βαθμός s τού $Q_{(f;v)}(t) = b_0 + b_1 t + \cdots + t^s$ ήταν γνήσια μικρότερος από m, τότε από $\vec{0} = Q_{(f;v)}(f)(\vec{v}) = b_0\vec{v} + b_1 f(\vec{v}) + \cdots + f^s(\vec{v})$ έπεται ότι τα διανύσματα $\vec{v}, f(\vec{v}), \ldots, f^s(\vec{v})$ είναι γραμμικώς εξαρτημένα, αφού τουλάχιστον ένας από τους συντελεστές b_i (του μη μηδενικού πολυωνύμου $Q_{(f;v)}(t)$) είναι διάφορος τού μηδενός. Αλλά αυτό είναι άτοπο, αφού τα $\vec{v}, f(\vec{v}), \ldots, f^s(\vec{v})$ με $s \leq m$, ανήκουν στη βάση $\mathcal{B}$. Ώστε ο βαθμός τού $Q_{(f;v)}(t)$ ισούται με m.

(2) Θεωρούμε τον περιορισμό τού f στον f–αναλλοίωτο υπόχωρο $C(f;\vec{v})$, δηλαδή τον ενδομορφισμό $f_{|C(f;\vec{v})} : C(f;\vec{v}) \longrightarrow C(f;\vec{v})$ και ας είναι $R(t)$ το ελάχιστο πολυώνυμο τού $f_{|C(f;\vec{v})}$. Επειδή το $R(t)$ είναι το ελάχιστο πολυώνυμο τού $f_{|C(f;\vec{v})}$, ο ενδομορφισμός $R(f_{|C(f;\vec{v})})$ είναι ο μηδενικός ενδομορφισμός τού $C(f;\vec{v})$ και γι'αυτό $R(f_{|C(f;\vec{v})})(\vec{v}) = \vec{0}$. Επομένως, το $Q_{(f;v)}(t)$ διαιρεί το $R(t)$.

Έστω $P(t)$ οποιοδήποτε πολυώνυμο και $P(f)$ ο αντίστοιχος ενδομορφισμός τού $\mathcal{E}$. Επειδή ο $C(f;\vec{v})$ είναι ένας f–αναλλοίωτος υπόχωρος, ο ενδομορφισμός $P(f)$ περιορισμένος στον $C(f;\vec{v})$ αποτελεί έναν ενδομορφισμό τού $C(f;\vec{v})$.

Έστω ότι το ελάχιστο πολυώνυμο τού f–κυκλικού υπόχωρου $C(f;\vec{v})$ είναι το $Q_{(f;v)}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_{m-1} t^{(m-1)} + t^m$.

Παρατηρούμε ότι για κάθε διάνυσμα $f^i(\vec{v}), 1 \leq i \leq m-1$, της βάσης $\mathcal{B}$ του $C(f;\vec{v})$, έχουμε:

$$\begin{aligned} Q_{(f;v)}(f)(f^i(\vec{v})) &= a_0 f^i(\vec{v}) + a_1 f(f^i(\vec{v})) + a_2 f^2(f^i(\vec{v})) + \cdots + a_{m-1} f^{(m-1)}(f^i(\vec{v})) + f^m(f^i(\vec{v})) \\ &= f^i(a_0\vec{v} + a_1 f(\vec{v}) + a_2 f^2(\vec{v}) + \cdots + a_{m-1} f^{(m-1)}(\vec{v}) + f^m(\vec{v})) \\ &= f^i(Q_{(f;v)}(f)(\vec{v})) = f^i(\vec{0}) = \vec{0}. \end{aligned}$$

Επομένως, ο ενδομορφισμός $Q_{(f;v)}(f)$ περιορισμένος στον υπόχωρο $C(f;\vec{v})$ ισούται με τον μηδενικό ενδομορφισμό τού $C(f;\vec{v})$, αφού η τιμή τού $Q_{(f;v)}(f)$ ισούται με το μηδενικό διάνυσμα πάνω σε κάθε στοιχείο τής βάσης $\mathcal{B}$. Γι'αυτό το ελάχιστο πολυώνυμο $R(t)$ τού $f_{|C(f;\vec{v})}$ διαιρεί το $Q_{(f;v)}(t)$.

Συνεπώς το $Q_{(f;v)}(t)$ διαιρεί το $R(t)$ και το $R(t)$ διαιρεί το $Q_{(f;v)}(t)$. Τέλος, επειδή ο συντελεστής των μεγιστοβαθμίων όρων και των δύο πολυωνύμων ισούται με 1, έπεται $Q_{(f;v)}(t) = R(t)$. $\square$

Παρατήρηση 9.8. *Ας είναι $\mathcal{E}$ ένας $\mathbb{K}$-διανυσματικός χώρος με $\dim_{\mathbb{K}} \mathcal{E} = n \in \mathbb{N} \cup \{0\}$, ας είναι $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ ένας $\mathbb{K}$-γραμμικός ενδομορφισμός, ας είναι $\vec{v}$ ένα μη μηδενικό διάνυσμα τού $\mathcal{E}$, ας είναι $C(f;\vec{v})$ ο αντίστοιχος f-κυκλικός υπόχωρος και τέλος ας είναι $Q_{(f;v)}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_{m-1} t^{(m-1)} + t^m$ το ελάχιστο πολυώνυμο τού $C(f;\vec{v})$. Χρησιμοποιώντας το προηγούμενο θεώρημα, βλέπουμε ότι ο $m \times m$ πίνακας $M_{\mathcal{B}}^{\mathcal{B}}(f_{|C(f;\vec{v})})$ που παριστάνει τον ενδομορφισμό*
$f_{|C(f;\vec{v})} : C(f;\vec{v}) \longrightarrow C(f;\vec{v})$ ως προς τη βάση $\mathcal{B} = \{\vec{v}, f(\vec{v}), f^2(\vec{v}), \ldots, f^{(m-1)}(\vec{v})\}$ είναι ο

$$M_{\mathcal{B}}^{\mathcal{B}}(f_{|C(f;\vec{v})}) = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 & -a_{m-2} \\ 0 & 0 & 0 & \ldots & 0 & 1 & -a_{m-1} \end{pmatrix}.$$

Πράγματι, για κάθε $f^i(\vec{v}), 0 \leq i \leq m-2$, τής βάσης $\mathcal{B}$ έχουμε $f(f^i(\vec{v})) = f^{i+1}(\vec{v})$. Για το διάνυσμα $f^{m-1}(\vec{v})$ έχουμε

$$f(f^{m-1}(\vec{v})) = f^m(\vec{v}) = (-a_0)\vec{v} + (-a_1)f(\vec{v}) + (-a_2)f^2(\vec{v}) + \cdots + (-a_{m-1})f^{(m-1)}(\vec{v}),$$

αφού $Q_{(f;v)}(\vec{v}) = a_0\vec{v} + a_1 f(\vec{v}) + a_2 f^2(\vec{v}) + \cdots + a_{m-1} f^{(m-1)}(\vec{v}) + f^m(\vec{v}) = \vec{0}$.

9.2. **Μηδενοδύναμοι ενδομορφισμοί–κυκλικοί υπόχωροι.**

Υπενθυμίζουμε ότι ένας ενδομορφισμός $f : \mathcal{E} \longrightarrow \mathcal{E}$, ονομάζεται μηδενοδύναμος, αν υπάρχει $n \in \mathbb{N}$ με f^n ίσο με τον μηδενικό ενδομορφισμό $\zeta_{\mathcal{E}}$ τού $\mathcal{E}$. Ονομάζουμε *δείκτη μηδενοδυναμικότητας* τον μικρότερο $k \in \mathbb{N}$ με $f^k = \zeta_{\mathcal{E}}$.

Πρόταση 9.9. *Έστω ότι $\mathcal{E}$ είναι ένας πεπερασμένης διάστασης $\mathbb{K}$-διανυσματικός χώρος με $\dim_{\mathbb{K}} \mathcal{E} \geq 1$ και ότι $f : \mathcal{E} \longrightarrow \mathcal{E}$ είναι ένας μηδενοδύναμος ενδομορφισμός με δείκτη μηδενοδυναμικότητας k. Τότε υπάρχει ένας f-κυκλικός υπόχωρος $C(f;\vec{v})$ με $\dim_{\mathbb{K}} C(f;\vec{v}) = k$.*

Απόδειξη. Αν $k = 1$, τότε ο $f = f^1$ είναι ο μηδενικός ενδομορφισμός $\zeta_{\mathcal{E}}$ και για οποιοδήποτε διάνυσμα $\vec{v} \neq \vec{0}$ έχουμε $f(\vec{v}) = \vec{0}$. Γι' αυτό, ο υπόχωρος $C(f;\vec{v}) = \langle \vec{v}, f(\vec{v}) \ldots, f^i(\vec{v}), f^{i+1}(\vec{v}), \ldots \rangle = \langle \vec{v} \rangle$ είναι διάστασης 1.
Ας δούμε τώρα την περίπτωση $k \geq 2$. Αφού ο f^k είναι ο μηδενικός ενδομορφισμός $\zeta_{\mathcal{E}}$ του $\mathcal{E}$, ενώ $f^{k-1} \neq \zeta_{\mathcal{E}}$, υπάρχει κάποιο $\vec{v} \in \mathcal{E}$ με $f^{k-1}(\vec{v}) \neq \vec{0}$. Θεωρούμε τον υπόχωρο $C(f;\vec{v}) = \langle \vec{v}, f(\vec{v}) \ldots, f^{k-1}(\vec{v}) \rangle$. Το ελάχιστο πολυώνυμο $Q_{(f;v)}(t)$ τού $C(f;\vec{v})$ διαιρεί το πολυώνυμο $P(t) = t^k$, αφού $P(f)(v) = f^k(\vec{v})$. Συνεπώς, $Q_{(f;v)}(t) = t^s$, όπου $s \geq k$. Αλλά, αν $s < k$, δηλαδή αν $k - s \geq 1$, τότε $Q_{(f;v)}(f)(\vec{v}) =$

$f^s(\vec{v}) = \vec{0}$ και γι'αυτό και $f^{k-1}(\vec{v}) = f^{k-s-1}(f^s(\vec{v})) = \vec{0}$. Το τελευταίο είναι άτοπο. Επομένως, το ελάχιστο πολυώνυμο $Q_{(f;v)}(t)$ τού $C(f;\vec{v})$ ισούται με t^k. Σύμφωνα με το προηγούμενο Θεώρημα 9.7, $\dim_{\mathbb{K}} C(f;\vec{v}) = k$. $\square$

Παρατήρηση 9.10. *Ο πίνακας $M_{\mathcal{B}}^{\mathcal{B}}(f_{|C(f;\vec{v})})$ που παριστάνει τον περιορισμό $f_{|C(f;\vec{v})}$ του μηδενοδύναμου ενδομορφισμού $f : \mathcal{E} \longrightarrow \mathcal{E}$ στον f-κυκλικό υπόχωρο $C(f;\vec{v})$ τής προηγούμενης πρότασης, ως προς τη βάση $\mathcal{B} = \{\vec{v}, f(\vec{v}) \dots, f^{k-1}(\vec{v})\}$ είναι ο*

$$M_{\mathcal{B}}^{\mathcal{B}}(f_{|C(f;\vec{v})}) = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

Πρόταση 9.11. *Έστω ότι $\mathcal{E}$ είναι ένας πεπερασμένης διάστασης $\mathbb{K}$-διανυσματικός χώρος με $\dim_{\mathbb{K}} \mathcal{E} \geq 1$, ότι $f : \mathcal{E} \longrightarrow \mathcal{E}$ είναι ένας μηδενοδύναμος ενδομορφισμός με δείκτη μηδενοδυναμικότητας k και ότι $\vec{v} \in \mathcal{E}$ είναι ένα μη μηδενικό διάνυσμα, όπου ο αντίστοιχος f-κυκλικός υπόχωρος $C(f;\vec{v})$ έχει $\mathbb{K}$-διάσταση k. Τότε υπάρχει ένας f-αναλλοίωτος υπόχωρος W του $\mathcal{E}$ με $\mathcal{E} = C(f;\vec{v}) \oplus W$.*

Απόδειξη. Επαγωγή ως προς τον δείκτη μηδενοδυναμικότητας k.
Αν $k = 1$, τότε ο $f = f^1$ ισούται με τον μηδενικό ενδομορφισμό $\zeta_{\mathcal{E}}$ του χώρου $\mathcal{E}$. Ας είναι $\vec{v}$ οποιοδήποτε μη μηδενικό διάνυσμα τού $\mathcal{E}$. Η $\mathbb{K}$–διάσταση τού χώρου $C(f;\vec{v})$ ισούται με 1 και οποιοδήποτε ευθύ συμπλήρωμα W τού $C(f;\vec{v})$, δηλαδή οποιοσδήποτε υπόχωρος W τού $\mathcal{E}$ με $C(f;\vec{v}) \oplus W = \mathcal{E}$, είναι f–αναλλοίωτος, αφού ο f είναι ο μηδενικός ενδομορφισμός $\zeta_{\mathcal{E}}$.

Ας υποθέσουμε ότι η πρόταση είναι αληθής για κάθε φυσικό αριθμό $< k$, θα την αποδείξουμε για τον φυσικό k.
Από την Πρόταση 9.9 γνωρίζουμε ότι υπάρχει $\vec{v} \in \mathcal{E}, \vec{v} \neq \vec{0}$ τέτοιο, ώστε ο f–κυκλικός υπόχωρος $C(f;\vec{v}) = \langle \vec{v}, f(\vec{v}) \dots, f^{k-1}(\vec{v}) \rangle$ να έχει διάσταση τον δείκτη μηδενοδυναμικότητας k. Θεωρούμε τον περιορισμό τού f στον f–αναλλοίωτο υπόχωρο $\mathsf{Im}(f) = f(\mathcal{E})$, δηλαδή τον $f_{|f(\mathcal{E})} : f(\mathcal{E}) \longrightarrow f(\mathcal{E})$. Ο δείκτης μηδενοδυναμικότητας τού $f_{|f(\mathcal{E})}$ ισούται με $k-1$, αφού $f^{k-1}(f(\mathcal{E})) = f^k(\mathcal{E}) = \{\vec{0}\}$, ενώ $f^{k-2}(f(\mathcal{E})) = f^{k-1}(\mathcal{E}) \neq \{\vec{0}\}$. Επιπλέον, η εικόνα

$$f(C(f;\vec{v})) = f(\langle \vec{v}, f(\vec{v}) \dots, f^{k-1}(\vec{v}) \rangle) = \langle f(\vec{v}), f(f(\vec{v})) \dots, f(f^{k-2}(\vec{v})) \rangle = C(f_{|f(\mathcal{E})}; f(\vec{v}))$$

τού $C(f;\vec{v})$ είναι ένας $f_{|f(\mathcal{E})}$–κυκλικός υπόχωρος με $\dim_{\mathbb{K}} C(f_{|f(\mathcal{E})}; f(\vec{v}))$ ίσο με τον δείκτη μηδενοδυναμικότητας $k-1$ τού $f_{|f(\mathcal{E})}$, αφού τα διανύσματα $f(\vec{v}), f(f(\vec{v})) = f^2(\vec{v}) \dots, f(f^{k-2}(\vec{v})) = f^{k-1}(\vec{v})$ είναι $\mathbb{K}$–γραμμικώς ανεξάρτητα.
Συνεπώς, λόγω της επαγωγικής υπόθεσης, υπάρχει ένας $f_{|f(\mathcal{E})}$–αναλλοίωτος υπόχωρος W' τού $f(\mathcal{E})$, δηλαδή $f_{|f(\mathcal{E})}(W') = f(W') \subseteq W'$, με $f(\mathcal{E}) = C(f_{|f(\mathcal{E})}; f(\vec{v})) \oplus W'$.
Ας είναι $W_0 = \{\vec{w} \in \mathcal{E} \mid f(\vec{w}) \in W'\}$. Προφανώς ο W_0 είναι ένας υπόχωρος τού $\mathcal{E}$. Επιπλέον, επειδή ο W' είναι f–αναλλοίωτος, έχουμε $W' \subseteq W_0$, (**) και αφού $f(W_0) \subseteq W'$, έπεται ότι ο W_0 είναι επίσης f–αναλλοίωτος.

Ισχυριζόμαστε ότι $\mathcal{E} = C(f;\vec{v}) + W_0$. Πράγματι, αν $\vec{x} \in \mathcal{E}$, τότε $f(\vec{x}) \in f(\mathcal{E}) = C(f_{|f(\mathcal{E})}; f(\vec{v})) \oplus W'$ (*). Επομένως, $f(\vec{x}) = \vec{z} + \vec{w}'$, όπου $\vec{z} \in C(f_{|f(\mathcal{E})}; f(\vec{v})) = f(C(f;\vec{v}))$ και $\vec{w}' \in W'$. Συνεπώς, $f(\vec{x}) = f(\vec{s}) + \vec{w}'$, όπου $\vec{z} = f(\vec{s}), \vec{s} \in C(f;\vec{v})$. Τώρα, $\vec{x} = \vec{s} + (\vec{x} - \vec{s})$ με $s \in C(f;\vec{v})$. Επιπλέον, το $\vec{x} - \vec{s}$ ανήκει στο W_0, αφού $f(\vec{x} - \vec{s}) = f(\vec{x}) - f(\vec{s}) = \vec{w}' \in W'$. Συνεπώς, $\mathcal{E} = C(f;\vec{v}) + W_0$.

Ισχυριζόμαστε ότι η τομή $C(f;\vec{v}) \cap W_0$ περιέχεται στον υπόχωρο $C(f_{|f(\mathcal{E})}; f(\vec{v})) = f(C(f;\vec{v}))$. Αν $\vec{x} \in C(f;\vec{v}) \cap W_0$, τότε $f(\vec{v}) \in f(C(f;\vec{v})) \cap f(W_0)$. Αλλά $f(C(f;\vec{v})) = C(f_{|f(\mathcal{E})}; f(\vec{v}))$, $f(W_0) \subseteq W_0$ και $C(f_{|f(\mathcal{E})}; f(\vec{v})) \cap W_0 = \{\vec{0}\}$. Επομένως, $f(\vec{x}) = \vec{0}$. Τώρα, το $\vec{x}$ ως στοιχείο τού $C(f;\vec{v})$, ισούται με

$\vec{x} = a_0\vec{x}+a_1 f(\vec{x})+\cdots+a_{k-1}f^{k-1}(\vec{v})$ και γι' αυτό $\vec{0} = f(\vec{x}) = a_0 f(\vec{x})+a_1 f^2(\vec{x})+\cdots+a_{k-2}f^{k-1}(\vec{v})$. Αλλά, αφού τα $f(\vec{x}), f^2(\vec{x}), \ldots, f^{k-1}(\vec{v})$ είναι $\mathbb{K}$–γραμμικώς ανεξάρτητα, έχουμε $a_0 = a_1 = \cdots = a_{k-2} = 0$ και επομένως το $\vec{x} = a_{k-1}f^{k-1}(\vec{v})$ ανήκει στον υπόχωρο $C(f_{|f(\mathcal{E})}; f(\vec{v})) = f(C(f;\vec{v}))$.

Αφού $C(f;\vec{v}) \cap W_0 \subseteq C(f_{|f(\mathcal{E})}; f(\vec{v}))$ και $W' \cap C(f_{|f(\mathcal{E})}; f(\vec{v})) = \{\vec{0}\}$, λόγω της (*), έχουμε $W' \cap (C(f;\vec{v}) \cap W_0) = \{\vec{0}\}$. Επομένως, το άθροισμα $W' \oplus (C(f;\vec{v}) \cap W_0)$ είναι ευθύ και περιέχεται στο W_0, επειδή το $W' \subseteq W_0$, λόγω τής (**).

Ας είναι W'' ένας υπόχωρος τού $\mathcal{E}$ με $W_0 = W'' \oplus [W' \oplus (C(f;\vec{v}) \cap W_0)]$, (***). Θεωρούμε τον υπόχωρο $W = W'' \oplus W'$. Παρατηρούμε ότι ο W'' είναι f–αναλλοίωτος, ως υπόχωρος τού f–αναλλοίωτου υπόχωρου W_0. Αφού και ο W' είναι f–αναλλοίωτος, έπεται ότι ο W είναι επίσης f–αναλλοίωτος.

Για να ολοκληρωθεί η απόδειξη τής πρότασης, αρκεί να δείξουμε ότι $\mathcal{E} = C(f;\vec{v}) \oplus W$. Παρατηρούμε ότι η τομή $C(f;\vec{v}) \cap W = C(f;\vec{v}) \cap (W'' \oplus W') \subseteq C(f;\vec{v}) \cap (W'' \oplus W') \cap W_0$, αφού $W'' \oplus W' \subseteq W_0$. Αλλά λόγω τής (***), η τομή $C(f;\vec{v}) \cap (W'' \oplus W') \cap W_0$ ισούται με $\{\vec{0}\}$. Επομένως, $C(f;\vec{v}) \cap W = \{\vec{0}\}$.

Τέλος,

$$\mathcal{E} = C(f;\vec{v}) + W_0 = C(f;\vec{v}) + W'' + W' + (C(f;\vec{v}) \cap W_0) = C(f;\vec{v}) + W'' + W' = C(f;\vec{v}) + W.$$

Ώστε, $\mathcal{E} = C(f;\vec{v}) \oplus W$, όπου ο W είναι ένας f–αναλλοίωτος υπόχωρος τού $\mathcal{E}$. $\square$

Η προηγούμενη πρόταση αποτελεί το επαγωγικό βήμα τού επόμενου ισχυρισμού:

Πρόταση 9.12. *Έστω ότι $\mathcal{E}$ είναι ένας πεπερασμένης διάστασης $\mathbb{K}$-διανυσματικός χώρος με $\dim_{\mathbb{K}} \mathcal{E} \geq 1$, ότι $f : \mathcal{E} \longrightarrow \mathcal{E}$ είναι ένας μηδενοδύναμος ενδομορφισμός με δείκτη μηδενοδυναμικότητας k.*
Τότε υπάρχει μια πεπερασμένη ακολουθία $k = k_1 \geq k_2 \geq \cdots \geq k_s$ φυσικών με $k_1 + k_2 + \ldots k_s = \dim_{\mathbb{K}} \mathcal{E}$ και μη μηδενικά διανύσματα $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_s$ τέτοια, ώστε

$$\mathcal{E} = C(f;\vec{v}_1) \oplus C(f;\vec{v}_2) \oplus \cdots \oplus C(f;\vec{v}_s),$$

όπου $\dim_{\mathbb{K}} C(f;\vec{v}_i) = k_i, \forall i, 1 \leq i \leq s$.

Απόδειξη. Λόγω των Προτάσεων 9.9 και 9.11, γνωρίζουμε ότι υπάρχει ένας f–κυκλικός υπόχωρος $C(f;\vec{v}_1) \subseteq \mathcal{E}$ διάστασης $k = k_1$ τέτοιος, ώστε $\mathcal{E} = C(f;\vec{v}_1) \oplus W_1$, όπου ο W_1 είναι ένας f–αναλλοίωτος υπόχωρος διάστασης $\dim_{\mathbb{K}} \mathcal{E} - k_1$. Τώρα, υπάρχει ένας f–κυκλικός υπόχωρος $C(f;\vec{v}_2) \subseteq W_1$ διάστασης k_2 με $W_1 = C(f;\vec{v}_2) \oplus W_2$, όπου ο W_2 είναι ένας f–αναλλοίωτος υπόχωρος διάστασης $\dim_{\mathbb{K}} \mathcal{E} - k_1 - k_2$. Επιπλέον,

$$\mathcal{E} = C(f;\vec{v}_1) \oplus W_1 = C(f;\vec{v}_1) \oplus C(f;\vec{v}_2) \oplus W_2$$

Σχηματίζοντας, για $i = 1, 2, \ldots, j, j+1, \ldots$, τον f–κυκλικό υπόχωρο $C(f;\vec{v}_i) \subseteq W_{i-1}$ διάστασης k_i και κατόπιν τον f–αναλλοίωτο υπόχωρο $W_i \subseteq W_{i-1}$ έτσι, ώστε $W_{i-1} = C(f;\vec{v}_i) \oplus W_i$, όπου $W_0 = \mathcal{E}$, με $\dim_{\mathbb{K}} W_i = \dim_{\mathbb{K}} \mathcal{E} - k_1 - k_2 - \cdots - k_i, i \geq 1$, διαπιστώνουμε ότι οι διαστάσεις των W_i σχηματίζουν μια φθίνουσα ακολουθία φυσικών αριθμών που οφείλει να σταματήσει μετά από πεπερασμένα, ας πούμε, s βήματα. Τότε $\dim_{\mathbb{K}} W_s = 0$ και

$$\mathcal{E} = C(f;\vec{v}_1) \oplus C(f;\vec{v}_2) \oplus \cdots \oplus C(f;\vec{v}_s).$$

$\square$

Παρατήρηση 9.13. *Αν $C(f;\vec{v}_i)$ είναι ένας f-κυκλικός υπόχωροςδιάστασης k_i, τότε γνωρίζουμε ότι το σύνολο $\{\vec{v}_i, f(\vec{v}_i), \ldots, f^{k-1}(\vec{v}_i)\}$ αποτελεί μια βάση του. Συνεπώς, αν $f : \mathcal{E} \longrightarrow \mathcal{E}$ είναι ένας μηδενοδύναμος ενδομορφισμός με δείκτη μηδενοδυναμικότητας ίσο με k, τότε, λόγω τής Πρότασης 9.12, ο χώρος $\mathcal{E}$ είναι το ευθύ άθροισμα των $C(f;\vec{v}_i), i = 1, 2, \ldots, s$ και η ένωση των αντίστοιχων βάσεων, δηλαδή το σύνολο*

$$\mathcal{B} = \{\vec{v}_1, f(\vec{v}_1), \ldots, f^{k_1-1}(\vec{v}_1)\} \cup \{\vec{v}_2, f(\vec{v}_2), \ldots, f^{k_2-1}(\vec{v}_2)\} \cup \cdots \cup \{\vec{v}_s, f(\vec{v}_s), \ldots, f^{k_s-1}(\vec{v}_s)\}$$

αποτελεί μια βάση τού $\mathcal{E}$. *Ο πίνακας που παριστάνει τον ενδομορφισμό* f *ως προς τη συγκεκριμένη βάση, είναι ο*

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} \boxed{B_1} & \\ & \boxed{B_2} \end{pmatrix}$$

Μέρος 2. **Ευκλείδειοι Χώροι**

10. **Σταθμητοί Χώροι και Ευκλείδειοι Χώροι**

10.1. **Σταθμητοί Χώροι.** Έστω $(\mathcal{E}, \langle , \rangle)$ ένας Ευκλείδειος χώρος. Τότε όπως γνωρίζουμε η απεικόνιση "μήκος διανύσματος"

$$\| \cdot \| : \mathcal{E} \longrightarrow \mathbb{R}, \quad \vec{x} \longmapsto \|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle}$$

ικανοποιεί, μεταξύ άλλων, τις ακόλουθες ιδιότητες:

1. $\|r\vec{x}\| = |r| \, \|\vec{x}\|, \quad \forall r \in \mathbb{R}, \ \forall \vec{x} \in \mathcal{E}$.
2. (Τριγωνική ανισότητα): $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|, \ \forall \vec{x}, \vec{y} \in \mathcal{E}$.
3. $\|\vec{x}\| \geq 0, \ \forall \vec{x} \in \mathcal{E}$, και αν $\|\vec{x}\| = 0$, τότε $\vec{x} = \vec{0}$.

Ορισμός 10.1. *Μια απεικόνιση* $\| \cdot \| : \mathcal{E} \longrightarrow \mathbb{R}$ *ορισμένη υπεράνω του* $\mathbb{R}$*-διανυσματικού χώρου* $\mathcal{E}$ *η οποία ικανοποιεί τις τρείς παραπάνω ιδιότητες καλείται* **στάθμη** *επί του* $\mathcal{E}$.

Ένας $\mathbb{R}$-**σταθμητός χώρος** *είναι ένα ζεύγος* $(\mathcal{E}, \| \cdot \|)$, *όπου* $\mathcal{E}$ *είναι ένας* $\mathbb{R}$*-διανυσματικός χώρος και* $\| \cdot \|$ *είναι μια στάθμη επί του* $\mathcal{E}$.

Υπενθυμίζουμε επίσης ότι αν $(\mathcal{E}, \langle , \rangle)$ είναι ένας Ευκλείδειος χώρος, τότε ικανοποιείται ο ακόλουθος

$$2\|\vec{x}\|^2 + 2\|\vec{y}\|^2 = \|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2 \qquad \text{(Νόμος Παραλληλογράμμου)}$$

Άρα αν $(\mathcal{E}, \langle , \rangle)$ είναι ένας Ευκλείδειος χώρος, τότε ορίζεται επί του $\mathcal{E}$ μια στάθμη $\| \cdot \| : \mathcal{E} \longrightarrow \mathbb{R}$, $\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle}$, η οποία ικανοποιεί τον νόμο του Παραλληλογράμμου.

10.2. **Το Θεώρημα των Jordan-Von Neumann.** Αντίστροφα ένα σημαντικό Θεώρημα, το οποίο οφείλεται στους Jordan και Von Neumann, βλέπε P. Jordan and J. Von Neumann: "*On Inner product spaces in linear metric spaces*", Annals of Mathematics, **36**, (1935), δείχνει ότι οι Ευκλείδειοι χώροι χαρακτηρίζονται ώς οι σταθμητοί χώροι των οποίων η στάθμη ικανοποιεί τον Νόμο του Παραλληλογράμμου. Για την απόδειξη του, η οποία χρησιμοποιεί βασικά στοιχεία τοπολογίας, χρειαζόμαστε κάποια προεργασία.

Ορισμός 10.2. *Ένας* **μετρικός χώρος** *είναι ένα ζεύγος* $(\mathcal{E}, \rho)$, *όπου* $\mathcal{E}$ *είναι ένα μη-κενό σύνολο και*

$$\rho : \mathcal{E} \times \mathcal{E} \longrightarrow \mathbb{R}, \ (x, y) \longmapsto \rho(x, y)$$

είναι μια απεικόνιση, η οποία καλείται **μετρική**, *έτσι ώστε να ικανοποιούνται οι ακόλουθες ιδιότητες:*

(1) $\forall x, y \in \mathcal{E}$: $\rho(x, y) \geq 0$
(2) $\rho(x, y) = 0$ *αν και μόνον αν* $x = y$.
(3) $\forall x, y \in \mathcal{E}$: $\rho(x, y) = \rho(y, x)$.
(4) $\forall x, y, z \in \mathcal{E}$: $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$.

Ένας μετρικός χώρος ο οποίος είναι και $\mathbb{R}$*-διανυσματικός χώρος, καλείται* **διανυσματικός μετρικός χώρος**.

Λήμμα 10.3. *Έστω $(\mathcal{E}, \|\cdot\|)$ ένας σταθμητός χώρος. Τότε η απεικόνιση*

$$\rho : \mathcal{E} \times \mathcal{E} \longrightarrow \mathbb{R}, \qquad \rho(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$$

είναι μια μετρική επί του $\mathcal{E}$ και το ζεύγος $(\mathcal{E}, \rho)$ είναι ένας μετρικός διανυσματικός χώρος.

Απόδειξη. Το ότι η απεικόνιση ρ είναι μετρική είναι άμεση συνέπεια των ιδιοτήτων της στάθμης $\|\cdot\|$. $\square$

Σύμφωνα με τα παραπάνω έχουμε τις ακόλουθες συνεπαγωγές

$$(\mathcal{E}, \langle , \rangle) : \text{ Ευκλείδειος Χώρος } \implies (\mathcal{E}, \|\cdot\|) : \text{ Σταθμητός Χώρος } \implies (\mathcal{E}, \rho) : \text{ Μετρικός Διανυσματικός Χώρος}$$

Σε έναν μετρικό (διανυσματικό) χώρο $(\mathcal{E}, \rho)$ μπορούμε να ορίσουμε την έννοια της σύγκλισης και συνέχειας ως εξής:

Έστω $(x_n)_{n\in\mathbb{N}}$ μια ακολουθία στοιχείων του $\mathcal{E}$.

(1) Η ακολουθία $(x_n)_{n\in\mathbb{N}}$ **συγκλίνει ισχυρά** στο στοιχείο $y \in \mathcal{E}$, και θα γράφουμε:

$$\lim_{n\to\infty} x_n = y$$

αν:

$$\forall \varepsilon > 0, \;\; \exists n_0 \in \mathbb{N} : \;\; \|x_n - y\| < \varepsilon, \quad \forall n > n_0$$

Προφανώς $\lim_{n\to\infty} x_n = y$ αν και μόνον αν η ακολουθία πραγματικών αριθμών $\|x_n - y\|$ συγκλίνει στο 0: $\lim_{n\to\infty} \|x_n - y\| = 0$.

(2) Η ακολουθία $(x_n)_{n\in\mathbb{N}}$ καλείται **ακολουθία Cauchy** αν:

$$\forall \varepsilon > 0, \;\; \exists n_0 \in \mathbb{N} : \;\; \|x_n - x_m\| < \varepsilon, \quad \forall n, m > n_0$$

Είναι εύκολο να δεί κανείς ότι κάθε ισχυρά συγκλίνουσα ακολουθία στοιχείων του $\mathcal{E}$ είναι ακολουθία Cauchy, και το όριο μιας ισχυρά συγκλίνουσας ακολουθίας (αν υπάρχει) είναι μοναδικό.

(3) Έστω $f : \mathcal{E} \longrightarrow \mathbb{R}$, μια συνάρτηση. Τότε η f καλείται **συνεχής** αν για κάθε στοιχείο y του $\mathcal{E}$ και κάθε ισχυρά συγκλίνουσα ακολουθία στοιχείων $(x_n)_{n\in\mathbb{N}}$ έτσι ώστε $\lim_{n\to\infty} x_n = y$, ισχύει ότι η ακολουθία πραγματικών αριθμών $(f(x_n))_{n\in\mathbb{N}}$ συγκλίνει στο $f(y)$:

$$\lim_{n\to\infty} x_n = y \qquad \implies \qquad \lim_{n\to\infty} f(x_n) = f(y)$$

Στην απόδειξη του κύριου Θεωρήματος της παρούσης παραγράφου θα χρειαστούμε την ακόλουθη:

Πρόταση 10.4. *Έστω $(\mathcal{E}, \|\cdot\|)$ ένας σταθμητός χώρος.*

(1) $\forall \vec{x}, \vec{y} \in \mathcal{E}$: $\big|\|\vec{x}\| - \|\vec{y}\|\big| \leq \|\vec{x} - \vec{y}\|$.

(2) *Η απεικόνιση $\|\cdot\| : \mathcal{E} \longrightarrow \mathbb{R}$ είναι συνεχής.*

Απόδειξη. (1) Θα έχουμε:

$$\|\vec{x}\| = \|\vec{y} + (\vec{x} - \vec{y})\| \;\leq\; \|\vec{y}\| + \|\vec{x} - \vec{y}\| \quad \text{και άρα} \quad \|\vec{x}\| - \|\vec{y}\| \;\leq\; \|\vec{x} - \vec{y}\|$$

Παρόμοια θα έχουμε: $\|\vec{y}\| - \|\vec{x}\| \leq \|\vec{x} - \vec{y}\|$ απ' όπου προκύπτει το ζητούμενο.

(2) Έστω $(\vec{x}_n)_{n\in\mathbb{N}}$ μια ισχυρά συγκλίνουσα ακολουθία στοιχείων του $\mathcal{E}$ έτσι ώστε $\lim_{n\to\infty} \vec{x}_n = \vec{y}$. Τότε από το (1) θα έχουμε:

$$\big|\|\vec{x}_n\| - \|\vec{y}\|\big| \;\leq\; \|\vec{x}_n - \vec{y}\|$$

και επομένως:

$$0 \;\leq\; \lim_{n\to\infty} \big|\|\vec{x}_n\| - \|\vec{y}\|\big| \;\leq\; \lim_{n\to\infty} \|\vec{x}_n - \vec{y}\| \;=\; 0$$

Αυτό όμως σημαίνει ότι η ακολουθία πραγματικών αριθμών $\|\vec{x}_n\|_{n\in\mathbb{N}}$ συγκλίνει στον $\|\vec{y}\|$:

$$\lim_{n\to\infty} \|\vec{x}_n\| = \|y\|$$

και άρα η στάθμη $\|\cdot\|$ είναι συνεχής. $\square$

Θεώρημα 10.5 (Jordan - Von Neumann (1935)). *Έστω $\mathcal{E}$ ένας $\mathbb{R}$-διανυσματικός χώρος και $\|\cdot\| : \mathcal{E} \longrightarrow \mathbb{R}$ μια στάθμη επί του $\mathcal{E}$ η οποία ικανοποιεί τον Νόμο του Παραλληλογράμμου. Τότε η απεικόνιση*

$$\langle \, , \rangle : \mathcal{E} \times \mathcal{E} \longrightarrow \mathbb{R}, \quad \langle \vec{x}, \vec{y} \rangle = \frac{\|\vec{x}+\vec{y}\|^2 - \|\vec{x}-\vec{y}\|^2}{4} \tag{$\dagger$}$$

ορίζει ένα μοναδικό εσωτερικό γινόμενο επί του $\mathcal{E}$ έτσι ώστε $\forall \vec{x} \in \mathcal{E}$:

$$\langle \vec{x}, \vec{x} \rangle = \|\vec{x}\|^2$$

Επομένως αν $\mathcal{E}$ είναι ένας $\mathbb{R}$-διανυσματικός χώρος, τότε τα ακόλουθα είναι ισοδύναμα:

(1) *Το ζεύγος $(\mathcal{E}, \langle , \rangle)$ είναι ένας Ευκλείδειος χώρος*

(2) *Το ζεύγος $(\mathcal{E}, \|\cdot\|)$ είναι σταθμητός χώρος και η στάθμη του $\|\cdot\|$ ικανοποιεί τον Νόμο του Παραλληλογράμμου.*

Απόδειξη. (**a**) Έστω $\vec{x}, \vec{y} \in \mathcal{E}$. Τότε χρησιμοποιώντας τις ιδιότητες της στάθμης $\|\cdot\|$ θα έχουμε:

$$\langle \vec{x}, \vec{y} \rangle = \frac{\|\vec{x}+\vec{y}\|^2 - \|\vec{x}-\vec{y}\|^2}{4} = \frac{\|\vec{y}+\vec{x}\|^2 - \|(-1)(\vec{y}-\vec{x}\|^2}{4} = \frac{\|\vec{y}+\vec{x}\|^2 - (-1)^2\|\vec{y}-\vec{x}\|^2}{4} =$$

$$= \frac{\|\vec{y}+\vec{x}\|^2 - \|\vec{y}-\vec{x}\|^2}{4} = \langle \vec{y}, \vec{x} \rangle$$

Άρα $\forall \vec{x}, \vec{y} \in \mathcal{E}$: $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$.

(**c**) Έστω $\vec{x} \in \mathcal{E}$. Τότε:

$$\langle \vec{x}, \vec{x} \rangle = \frac{\|\vec{x}+\vec{x}\|^2 - \|\vec{x}-\vec{x}\|^2}{4} = \frac{\|2\vec{x}\|^2}{4} = \frac{|2|^2\|\vec{x}\|^2}{4} = \frac{4\|\vec{x}\|^2}{4} = \|\vec{x}\|^2 \geq 0$$

Αν $\langle \vec{x}, \vec{x} \rangle = 0$, τότε από την παραπάνω σχέση θα έχουμε: $\|\vec{x}\|^2$ και επειδή η απεικόνιση $\|\cdot\|$ είναι στάθμη θα έχουμε $\vec{x} = \vec{0}$.

(**b**) Έστω $\vec{x}, \vec{y}, \vec{z} \in \mathcal{E}$ και $\lambda \in \mathbb{R}$.

(i) Από τον ορισμό της απεικόνισης $\langle \, , \rangle : \mathcal{E} \times \mathcal{E} \longrightarrow \mathbb{R}$, θα έχουμε:

$$\langle \vec{x}, \vec{z} \rangle = \frac{\|\vec{x}+\vec{z}\|^2 - \|\vec{x}-\vec{z}\|^2}{4} \quad \text{και} \quad \langle \vec{y}, \vec{z} \rangle = \frac{\|\vec{y}+\vec{z}\|^2 - \|\vec{y}-\vec{z}\|^2}{4}$$

και επομένως:

$$\langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle = \frac{\|\vec{x}+\vec{z}\|^2 + \|\vec{y}+\vec{z}\|^2 - \|\vec{x}-\vec{z}\|^2 - \|\vec{y}-\vec{z}\|^2}{4} \tag{$*$}$$

Εφαρμώζοντας το Νόμο του Παραλληλογράμμου για τα διανύσματα $\vec{x}+\vec{z}$ και $\vec{y}+\vec{z}$ θα έχουμε:

$$2\|\vec{x}+\vec{z}\|^2 + 2\|\vec{y}+\vec{z}\|^2 = \|\vec{x}+\vec{y}+2\vec{z}\|^2 + \|\vec{x}-\vec{y}\|^2$$

η οποία γράφεται ισοδύναμα και ως εξής:

$$2\|\vec{x}+\vec{z}\|^2 + 2\|\vec{y}+\vec{z}\|^2 = \|\vec{x}+\vec{y}+2\vec{z}\|^2 + \|\vec{x}+\vec{y}\|^2 + \|\vec{x}-\vec{y}\|^2 - \|\vec{x}+\vec{y}\|^2$$

Εφαρμόζοντας στην τελευταία σχέση πάλι το Νόμο του Παραλληλογράμμου για τα διανύσματα $\vec{x}+\vec{y}+\vec{z}$ και $\vec{z}$ θα έχουμε:

$$2\|\vec{x}+\vec{z}\|^2 + 2\|\vec{y}+\vec{z}\|^2 = 2\|\vec{x}+\vec{y}+\vec{z}\|^2 + 2\|\vec{z}\|^2$$

και επομένως η προτελευταία σχέση γράφεται:

$$2\|\vec{x}+\vec{z}\|^2+2\|\vec{y}+\vec{z}\|^2 \;=\; 2\|\vec{x}+\vec{y}+\vec{z}\|^2+2\|\vec{z}\|^2-\|\vec{x}+\vec{y}\|^2+\|\vec{x}-\vec{y}\|^2 \tag{10.1}$$

Θέτοντας $-\vec{z}$ στην θέση του $\vec{z}$ η παραπάνω σχέση δίνει:

$$2\|\vec{x}-\vec{z}\|^2+2\|\vec{y}-\vec{z}\|^2 \;=\; 2\|\vec{x}+\vec{y}-\vec{z}\|^2+2\|\vec{z}\|^2-\|\vec{x}+\vec{y}\|^2+\|\vec{x}-\vec{y}\|^2 \tag{10.2}$$

Αφαιρώντας την (10.2) από την (10.1) τότε θα έχουμε την ταυτότητα:

$$\|\vec{x}+\vec{z}\|^2+\|\vec{y}+\vec{z}\|^2-\|\vec{x}-\vec{z}\|^2-\|\vec{y}-\vec{z}\|^2 \;=\; \|\vec{x}+\vec{y}+\vec{z}\|^2+\|\vec{x}+\vec{y}-\vec{z}\|^2$$

Τότε η $(*)$ γράφεται ώς εξής:

$$\langle\vec{x},\vec{z}\rangle+\langle\vec{y},\vec{z}\rangle \;=\; \frac{\|\vec{x}+\vec{y}+\vec{z}\|^2+\|\vec{x}+\vec{y}-\vec{z}\|^2}{4} \;=\; \langle\vec{x}+\vec{y},\vec{z}\rangle$$

Άρα θα έχουμε:

$$\langle\vec{x}+\vec{y},\vec{z}\rangle \;=\; \langle\vec{x},\vec{z}\rangle+\langle\vec{y},\vec{z}\rangle \tag{10.3}$$

(ii) Μένει να δείξουμε ότι: $\langle\lambda\vec{x},\vec{y}\rangle=\lambda\langle\vec{x},\vec{y}\rangle$.

• Από την σχέση (10.3) έπεται άμεσα ότι: $\langle 2\vec{x},\vec{y}\rangle=2\langle\vec{x},\vec{y}\rangle$. Επαγωγικά προκύπτει εύκολα ότι:

$$\langle n\vec{x},\vec{y}\rangle=n\langle\vec{x},\vec{y}\rangle, \qquad \forall n\in\mathbb{N}$$

Από την άλλη πλευρά, θα έχουμε:

$$\langle 0\vec{x},\vec{y}\rangle \;=\; \langle\vec{0},\vec{y}\rangle \;=\; \frac{\|\vec{0}+\vec{y}\|^2-\|\vec{0}-\vec{y}\|^2}{4} \;=\; \frac{\|\vec{y}\|^2-\|-\vec{y}\|^2}{4} \;=\; \frac{\|\vec{y}\|^2-(-1)^2\|\vec{y}\|^2}{4} \;=$$

$$=\; \frac{\|\vec{y}\|^2-\|\vec{y}\|^2}{4} \;=\; 0$$

Επίσης αν $n\in\mathbb{N}$, τότε:

$$n\langle\vec{x},\vec{y}\rangle+\langle(-n)\vec{x},\vec{y}\rangle \;=\; \langle(n+(-n))\vec{x},\vec{y}\rangle \;=\; \langle 0\vec{x},\vec{y}\rangle \;=\; \langle\vec{0},\vec{y}\rangle=0$$

και επομένως

$$\langle(-n)\vec{x},\vec{y}\rangle \;=\; (-n)\langle\vec{x},\vec{y}\rangle$$

Συνοψίζοντας θα έχουμε:

$$\langle n\vec{x},\vec{y}\rangle=n\langle\vec{x},\vec{y}\rangle, \qquad \forall n\in\mathbb{Z} \tag{10.4}$$

• Έστω τώρα $r=\frac{m}{n}\in\mathbb{Q}$. Τότε:

$$n\langle\frac{m}{n}\vec{x},\vec{y}\rangle \;=\; \langle n\frac{m}{n}\vec{x},\vec{y}\rangle \;=\; \langle m\vec{x},\vec{y}\rangle \;=\; m\langle\vec{x},\vec{y}\rangle$$

και άρα θα έχουμε:

$$\langle\frac{m}{n}\vec{x},\vec{y}\rangle \;=\; \frac{m}{n}\langle\vec{x},\vec{y}\rangle$$

Δηλαδή

$$\langle r\vec{x},\vec{y}\rangle=r\langle\vec{x},\vec{y}\rangle, \qquad \forall r\in\mathbb{Q} \tag{10.5}$$

• Τέλος υποθέτουμε ότι $\lambda \in \mathbb{R}$. Τότε, όπως είναι γνωστό από τον Απειροστικό Λογισμό, υπάρχει μια ακολουθία $(r_n)_{n\in\mathbb{N}}$ ρητών αριθμών έτσι ώστε:

$$\lim_{n\to\infty} r_n = \lambda$$

Επειδή από την Πρόταση 10.4 η απεικόνιση $\mathcal{E} \longrightarrow \mathbb{R}$, $\vec{x} \longmapsto \|\vec{x}\|$ είναι συνεχής, έπεται ότι και η απεικόνιση, για σταθερό $\vec{y} \in \mathcal{E}$:

$$\langle -, \vec{y}\rangle : \mathcal{E} \longrightarrow \mathbb{R}, \quad \vec{x} \longmapsto \langle \vec{x}, \vec{y}\rangle = \frac{\|\vec{x}+\vec{y}\|^2 - \|\vec{x}-\vec{y}\|^2}{4}$$

είναι συνεχής συνάρτηση του $\vec{x}$. Τότε θα έχουμε:

$$\langle \lambda\vec{x}, \vec{y}\rangle = \langle \lim_{n\to\infty} r_n\vec{x}, \vec{y}\rangle = \frac{\|\lim_{n\to\infty} r_n\vec{x}+\vec{y}\|^2 - \|\lim_{n\to\infty} r_n\vec{x}-\vec{y}\|^2}{4} =$$

$$= \frac{\|\lim_{n\to\infty}(r_n\vec{x}+\vec{y})\|^2 - \|\lim_{n\to\infty}(r_n\vec{x}-\vec{y})\|^2}{4} =$$

$$= \frac{\lim_{n\to\infty}\|r_n\vec{x}+\vec{y}\|^2 - \lim_{n\to\infty}\|r_n\vec{x}-\vec{y}\|^2}{4} =$$

$$= \lim_{n\to\infty} \frac{\|r_n\vec{x}+\vec{y}\|^2 - \|r_n\vec{x}-\vec{y}\|^2}{4} = \lim_{n\to\infty}\langle r_n\vec{x}, \vec{y}\rangle \stackrel{(*)}{=} \lim_{n\to\infty} r_n\langle \vec{x}, \vec{y}\,\rangle = \lambda\langle \vec{x}, \vec{y}\rangle$$

όπου στην ισότητα $(*)$ χρησιμοποιήσαμε την σχέση (10.5).

Επομένως δείξαμε ότι η απεικόνιση $\langle , \rangle$ είναι ένα εσωτερικό γινόμενο στον $\mathcal{E}$, το οποίο λόγω της (**c**) ικανοποιεί τη σχέση $\langle \vec{x}, \vec{x}\rangle = \|\vec{x}\|^2$, $\forall \vec{x} \in \mathcal{E}$.

Αν $\langle\!\langle , \rangle\!\rangle : \mathcal{E} \times \mathcal{E} \longrightarrow \mathbb{R}$, είναι ένα άλλο εσωτερικό γινόμενο στον $\mathcal{E}$ το οποίο ικανοποιεί την σχέση $\langle\!\langle \vec{x}, \vec{x}\rangle\!\rangle = \|\vec{x}\|^2$, $\forall \vec{x} \in \mathcal{E}$, τότε η σχέση (†) δείχνει ότι: $\langle\!\langle \vec{x}, \vec{x}\rangle\!\rangle = \langle \vec{x}, \vec{y}\rangle$, $\forall \vec{x}, \vec{y} \in \mathcal{E}$. □

Το ακόλουθο είναι ένα παράδειγμα σταθμητού χώρου του οποίου η στάθμη δεν ικανοποιεί τον Νόμο του Παραλληλογράμμου και άρα ο χώρος δεν είναι Ευκλείδειος.

Παράδειγμα 10.6. *Έστω $p \geq 1$ και έστω το ακόλουθο σύνολο ακολουθιών πραγματικών αριθμών:*

$$\ell^p(\mathbb{R}) := \Big\{(x_n)_{n\in\mathbb{N}} \mid x_n \in \mathbb{R} \quad \textit{και} \quad \sum_{n=0}^{\infty} |x_n|^p < \infty\Big\}$$

Τότε αποδεικνύεται εύκολα ότι το σύνολο $\ell^p(\mathbb{R})$ είναι ένας διανυσματικός χώρος άπειρης διάστασης υπεράνω του (R). Ορίζουμε μια απεικόνιση:

$$\|\cdot\|_p : \ell^p(\mathbb{R}) \longrightarrow \mathbb{R}, \qquad \|(x_n)_{n\in\mathbb{N}}\|_p = \Big(\sum_{n=0}^{\infty} |x_n|^p\Big)^{\frac{1}{p}}$$

Αποδεικνύεται εύκολα ότι η απεικόνιση $\|\cdot\|_p$ είναι μια στάθμη επί του $\ell^p(\mathbb{R})$ και άρα το ζεύγος $(\ell^p(\mathbb{R}), \|\cdot\|_p)$ είναι ένας σταθμητός χώρος.

(1) *Αν $p = 2$, τότε βλέπουμε εύκολα ότι η απεικόνιση*

$$\langle , \rangle \; : \; \ell^2(\mathbb{R}) \times \ell^2(\mathbb{R}) \longrightarrow \mathbb{R}, \quad \big\langle (x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \big\rangle = \sum_{n=0}^{\infty} x_n y_n$$

είναι ένα εσωτερικό γινόμενο επί του $\ell^2(\mathbb{R})$ έτσι ώστε: $\|(x_n)_{n\in\mathbb{N}}\|_p^2 = \langle (x_n)_{n\in\mathbb{N}}, (x_n)_{n\in\mathbb{N}} \rangle$. Άρα το ζεύγος $(\ell^2(\mathbb{R}), \langle , \rangle)$ είναι ένας Ευκλείδειος χώρος.

(2) *Θα δείξουμε ότι αν $p > 2$, τότε δεν ισχύει ο Νόμος του Παραλληλογράμμου. Πραγματικά: όπως έχουμε δει αν ισχύει ο Νόμος του Παραλληλογράμμου, τότε, $\forall \vec{x}, \vec{y} \in \ell^2(\mathbb{R})$:*

$$\|\vec{x}+\vec{y}\|^2 \; + \; \|\vec{x}-\vec{y}\|^2 \; = \; 2\|\vec{x}\|^2 \; + \; 2\|\vec{y}\|^2 \tag{+}$$

Θεωρούμε τις ακολουθίες:

$$(x_n)_{n\in\mathbb{N}} = (1, 1, 0, 0, \cdots) \qquad \textit{και} \qquad (y_n)_{n\in\mathbb{N}} = (1, -1, 0, 0, \cdots)$$

οι οποίες ποροφανώς ανήκουν στον σταθμητό χώρο $\ell^p(\mathbb{R})$. Υπολογίζουμε εύκολα:

$$\|(x_n)_{n\in\mathbb{N}}\|_p = \|(x_n)_{n\in\mathbb{N}}\|_p = 2^{\frac{1}{p}} \qquad \textit{και} \qquad \|(x_n)_{n\in\mathbb{N}} + (y_n)_{n\in\mathbb{N}}\|_p = \|(x_n)_{n\in\mathbb{N}} - (y_n)_{n\in\mathbb{N}}\|_p = 2$$

Άρα λόγω της σχέσης $(+)$ θα πρέπει

$$2(2^{\frac{1}{p}})^2 + 2(2^{\frac{1}{p}})^2 = 2^2 + 2^2 \quad \Longrightarrow \quad 2^2 = 2^p$$

το οποίο ισχύει αν και μόνον αν $p = 2$. Άρα αν $p > 2$, δεν ισχύει η σχέση $(+)$ και άρα και ο Νόμος του Παραλληλογράμμου. Συμπεραίνουμε ότι αν $p > 2$, τότε δεν υπάρχει εσωτερικό γινόμενο $\langle , \rangle$ στον σταθμητό χώρο $\ell^p(\mathbb{R})$, έτσι ώστε: $\langle (x_n)_{n\in\mathbb{N}}, (x_n)_{n\in\mathbb{N}} \rangle = \|(x_n)_{n\in\mathbb{N}}\|_p^2$.

Παρατήρηση 10.7. *Υπάρχουν περισσότεροι από 350 χαρακτηρισμοί του πότε ένας σταθμητός χώρος είναι Ευκλείδειος, βλέπε το βιβλίο: D. Amir: Characterizations of Inner Product Spaces, Birkhäuser Verlag, Basel, 1986. Ο χαρακτηρισμός του Θεωρήματος 10.5, ο οποίος οφείλεται στους Jordan και Von Neumann και αποδείχθηκε το 1935, είναι ίσως ο πρώτος ιστορικά χαρακτηρισμός.*

Αναφέρουμε μόνο τους ακόλουθους ενδιαφέροντες χαρακτηρισμούς:

Θεώρημα 10.8. *Για έναν σταθμητό χώρο $(\mathcal{E}, \|\cdot\|)$ οι ακόλουθες συνθήκες είναι ισοδυναμες:*

(1) *Υπάρχει ένα (μοναδικό) εσωτερικό γινόμενο $\langle , \rangle : \mathcal{E} \times \mathcal{E} \longrightarrow \mathbb{R}$, έτσι ωστε:*

$$\langle \vec{x}, \vec{x} \rangle = \|\vec{x}\|^2, \qquad \forall \vec{x} \in \mathcal{E}$$

(2) *(Lorch (1947)) $\forall \vec{x}, \vec{y} \in \mathcal{E}$:*

$$\Big(\|\vec{x}\| + \|\vec{y}\|\Big)\Big\| \frac{\vec{x}}{\|\vec{x}\|} - \frac{\vec{y}}{\|\vec{y}\|} \Big\| \; \leq \; 2\|\vec{x}-\vec{y}\|$$

(3) *(Lorch (1947)) $\forall \vec{x}, \vec{y} \in \mathcal{E}$, $\forall \lambda \in \mathbb{R}$:*

$$\|\vec{x}\| = \|\vec{y}\| \qquad \Longrightarrow \qquad \|\vec{x} + \lambda\vec{y}\| = \|\lambda\vec{x} + \vec{y}\|$$

(4) *(Lorch (1947)) $\forall \vec{x}, \vec{y}, \vec{z} \in \mathcal{E}$:*

$$\|\vec{x}\| = \|\vec{y}\| \quad \textit{και} \quad \vec{x} + \vec{y} + \vec{z} = \vec{0} \qquad \Longrightarrow \qquad \|\vec{x} - \vec{z}\| = \|\vec{y} - \vec{z}\|$$

(Σε ένα ισοσκελές τρίγωνο οι διχοτόμοι των ίσων γωνιών είναι ίσες).

(5) *(Day (1947)) $\forall \vec{x}, \vec{y} \in \mathcal{E}$:*

$$\|\vec{x}\| = \|\vec{y}\| = 1 \qquad \Longrightarrow \qquad \|\vec{x}+\vec{y}\|^2 + \|\vec{x}-\vec{y}\|^2 = 4$$

11. **Η Ορίζουσα Gram και οι Εφαρμογές της**

Έστω $(\mathcal{E}, \langle , \rangle)$ ένας Ευκλείδειος χώρος. Στην παρούσα παράγραφο θα δούμε, μεταξύ άλλων, μια γεωμετρική ερμηνεία της ορίζουσας ενός τετραγωνικού πίνακα, και ένα χρήσιμο κριτήριο για το πότε ένα πεπερασμένο σύνολο διανυσμάτων

$$\mathcal{S} = \big\{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m\big\}$$

είναι γραμμικά ανεξάρτητο.

11.1. **Πίνακας και Ορίζουσα Gram.**

Ορισμός 11.1. *Ο* **πίνακας Gram** *των διανυσμάτων $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m$ ορίζεται να είναι ο πίνακας:*

$$G(\vec{x}_1, \cdots, \vec{x}_m) := \begin{pmatrix} \langle \vec{x}_1, \vec{x}_1\rangle & \langle \vec{x}_1, \vec{x}_2\rangle & \cdots & \langle \vec{x}_1, \vec{x}_m\rangle \\ \langle \vec{x}_2, \vec{x}_1\rangle & \langle \vec{x}_2, \vec{x}_2\rangle & \cdots & \langle \vec{x}_2, \vec{x}_m\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{x}_m, \vec{x}_1\rangle & \langle \vec{x}_m, \vec{x}_2\rangle & \cdots & \langle \vec{x}_m, \vec{x}_m\rangle \end{pmatrix}$$

Η **ορίζουσα Gram** *των διανυσμάτων $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m$ ορίζεται να είναι η ορίζουσα του πίνακα Gram:*

$$\big|G(\vec{x}_1, \cdots, \vec{x}_m)\big|$$

Παρατηρούμε ότι η ορίζουσα Gram είναι η ορίζουσα ενός συμμετρικού πίνακα διότι:

$$\langle \vec{x}_i, \vec{x}_j\rangle = \langle \vec{x}_j, \vec{x}_i\rangle, \qquad 1 \leq i, j \leq m$$

- Αν $n = 1$, τότε η ορίζουσα Gram του διανύσματος $\vec{x}_1$ είναι:

$$\big|G(\vec{x}_1)\big| = \langle \vec{x}_1, \vec{x}_1\rangle = \|\vec{x}_1\|^2 \ \geq 0$$

και η ισότητα ισχύει αν και μόνον αν $\vec{x}_1 = \vec{0}$.

- Αν $n = 2$, τότε η ορίζουσα Gram των διανυσμάτων $\vec{x}_1, \vec{x}_2$ είναι:

$$\big|G(\vec{x}_1, \vec{x}_2)\big| = \begin{vmatrix} \langle \vec{x}_1, \vec{x}_1\rangle & \langle \vec{x}_1, \vec{x}_2\rangle \\ \langle \vec{x}_2, \vec{x}_1\rangle & \langle \vec{x}_2, \vec{x}_2\rangle \end{vmatrix} = \begin{vmatrix} \|\vec{x}_1\|^2 & \langle \vec{x}_1, \vec{x}_2\rangle \\ \langle \vec{x}_2, \vec{x}_1\rangle & \|\vec{x}_2\|^2 \end{vmatrix} = \|\vec{x}_1\|^2 \, \|\vec{x}_2\|^2 - \langle \vec{x}_1, \vec{x}_2\rangle^2 \ \geq \ 0$$

όπου η τελευταία ανισότητα ισχύει λόγω της ανισότητας των Cauchy-Schwarz. Επιπλέον

$$\big|G(\vec{x}_1, \vec{x}_2)\big| = 0 \quad \text{αν και μόνον αν} \quad \text{τα} \quad \vec{x}_1, \vec{x}_2 \quad \text{είναι γραμμικά εξαρτημένα}$$

Θα δούμε ότι ισχύει κάτι ανάλογο και για $n \geq 3$.

Πρόταση 11.2. *Τα ακόλουθα είναι ισοδύναμα:*

(1) *Το σύνολο διανυσμάτων $\mathcal{S} = \big\{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m\big\}$ του $\mathcal{E}$ είναι γραμμικά ανεξάρτητο.*
(2) $\big|G(\vec{x}_1, \cdots, \vec{x}_m)\big| \neq 0$.

Απόδειξη. (1) $\Longrightarrow$ (2) Έστω ότι το $\mathcal{S}$ είναι γραμμικά εξαρτημένο. Τότε υπάρχουν αριθμοί $\lambda_1, \cdots, \lambda_m \in \mathbb{R}$, έτσι ώστε: $(\lambda_1, \cdots, \lambda_m) \neq (0, \cdots, 0)$ και:

$$\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \cdots + \lambda_m\vec{x}_m = \vec{0} \tag{11.1}$$

Τότε θεωρώντας το εσωτερικό γινόμενο των δύο μελών της (11.1) με το διάνυσμα $\vec{x}_i$, $1 \leq i \leq m$, θα έχουμε:

$$\langle \vec{x}_i, \lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \cdots + \lambda_m\vec{x}_m\rangle = \langle \vec{x}_i, \vec{0}\rangle = 0 \qquad 1 \leq i \leq m$$

Η παραπάνω σχέση γράφεται:

$$\lambda_1\langle\vec{x}_i,\vec{x}_1\rangle + \lambda_2\langle\vec{x}_i,\vec{x}_2\rangle + \cdots + \lambda_m\langle\vec{x}_i,\vec{x}_m\rangle = 0 \qquad 1 \le i \le m$$

Οι τελευταίες σχέσεις δείχνουν ότι θα έχουμε μια σχέση γραμμικής εξάρτησης

$$\lambda_1\Gamma_1 + \lambda_2\Gamma_2 + \cdots \lambda_m\Gamma_m = 0$$

μεταξύ των γραμμών

$$\Gamma_i = \Big(\langle\vec{x}_i,\vec{x}_1\rangle,\ \ \langle\vec{x}_i,\vec{x}_2\rangle,\ \ \cdots,\ \ \langle\vec{x}_i,\vec{x}_m\rangle\Big)$$

του πίνακα Gramm $G(\vec{x}_1,\cdots,\vec{x}_m)$. Τότε όμως η ορίζουσα Gram είναι μηδέν. Άρα αν το σύνολο $\mathcal{S}$ είναι γραμμικά ανεξάρτητο, έπεται ότι: $\big|G(\vec{x}_1,\cdots,\vec{x}_m)\big| \neq 0$.

(2) $\Longrightarrow$ (1) Έστω ότι: $\big|G(\vec{x}_1,\cdots,\vec{x}_m)\big| = 0$. Τότε οι γραμμές του πίνακα $G(\vec{x}_1,\cdots,\vec{x}_m)$ είναι γραμμικά εξαρτημένες ή ισοδύναμα οι στήλες του $G(\vec{x}_1,\cdots,\vec{x}_m)$ είναι γραμμικά εξαρτημένες. Επομένως υπάρχουν αριθμοί $\lambda_1,\cdots,\lambda_m \in \mathbb{R}$, όπου: $(\lambda_1,\cdots,\lambda_m) \neq (0,\cdots,0)$, έτσι ώστε:

$$\lambda_1\Sigma_1 + \lambda_2\Sigma_2 + \cdots \lambda_m\Sigma_m = 0$$

Δηλαδή:

$$\lambda_1\begin{pmatrix}\langle\vec{x}_1,\vec{x}_1\rangle\\ \langle\vec{x}_2,\vec{x}_1\rangle\\ \vdots\\ \langle\vec{x}_m,\vec{x}_1\rangle\end{pmatrix} + \lambda_2\begin{pmatrix}\langle\vec{x}_1,\vec{x}_2\rangle\\ \langle\vec{x}_2,\vec{x}_2\rangle\\ \vdots\\ \langle\vec{x}_m,\vec{x}_2\rangle\end{pmatrix} + \cdots + \lambda_m\begin{pmatrix}\langle\vec{x}_1,\vec{x}_m\rangle\\ \langle\vec{x}_2,\vec{x}_m\rangle\\ \vdots\\ \langle\vec{x}_m,\vec{x}_m\rangle\end{pmatrix} = \begin{pmatrix}0\\0\\ \vdots\\0\end{pmatrix}$$

Αυτό σημαίνει ότι:

$$\lambda_1\langle\vec{x}_i,\vec{x}_1\rangle + \lambda_2\langle\vec{x}_i,\vec{x}_2\rangle + \cdots + \lambda_m\langle\vec{x}_i,\vec{x}_m\rangle = 0 \qquad 1 \le i \le m$$

και επομένως

$$\langle\vec{x}_i,\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \cdots + \lambda_m\vec{x}_m\rangle = 0 \qquad 1 \le i \le m$$

Πολλαπλασιάζοντας διαδοχικά κάθε σχέση από τις παραπάνω με λ_1, $\cdots$, λ_m και ακολούθως προσθέτοντας τις σχέσεις που προκύπτουν καταλήγουμε στην σχέση:

$$\|\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \cdots + \lambda_m\vec{x}_m\|^2 = 0$$

από την οποία προκύπτει ότι:

$$\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \cdots + \lambda_m\vec{x}_m = \vec{0}$$

και επομένως το σύνολο διανυσμάτων $\mathcal{S}$ είναι γραμμικά εξαρτημένο. Καταλήγουμε: αν η ορίζουσα Gram $\big|G(\vec{x}_1,\cdots,\vec{x}_m)\big| \neq 0$, τότε το σύνολο $\mathcal{S}$ είναι γραμμικά ανεξάρτητο. □

Έστω $(\mathcal{E},\langle,\rangle)$ ένας Ευκλείδειος χώρος πεπερασμένης διάστασης και $\mathcal{S} = \big\{\vec{x}_1,\vec{x}_2,\cdots,\vec{x}_m\big\}$ ένα σύνολο διανυσμάτων του $\mathcal{E}$.

Έστω $\mathcal{B} = \{\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n\}$ μια ορθοκανονική βάση του $\mathcal{E}$. Τότε θα έχουμε μοναδική γραφή των διανυσμάτων του συνόλου $\mathcal{S}$ ως γραμμικό συνδυασμό των διανυσμάτων της βάσης $\mathcal{B}$:

$$\vec{x}_1 = x_{11}\vec{e}_1 + x_{12}\vec{e}_2 + \cdots + x_{1n}\vec{e}_n$$

$$\vdots$$

$$\vec{x}_i = x_{i1}\vec{e}_1 + x_{i2}\vec{e}_2 + \cdots + x_{in}\vec{e}_n$$

$$\vdots$$

$$\vec{x}_m = x_{m1}\vec{e}_1 + x_{m2}\vec{e}_2 + \cdots + x_{mn}\vec{e}_n$$

Συμβολίζουμε με H τον $m \times n$ πίνακα του οποίου οι γραμμές είναι οι συνιστώσες των διανυσμάτων $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m$ στην βάση $\mathcal{B}$:

$$H = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix}$$

Υπολογίζουμε εύκολα:

$$H \cdot {}^tH = G(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$$

Αυτό προκύπτει από το γεγονός ότι:

$$\langle \vec{x}_i, \vec{x}_j \rangle = \sum_{k=1}^{n} x_{ik}x_{jk} = (H \cdot {}^tH)_{ij}$$

Πρόταση 11.3. *Έστω $(\mathcal{E}, \langle , \rangle)$ ένας Ευκλείδειος χώρος πεπερασμένης διάστασης και $\mathcal{S} = \{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m\}$ ένα σύνολο διανυσμάτων του $\mathcal{E}$. Τότε:*

$$\left|G(\vec{x}_1, \cdots, \vec{x}_m)\right| \ \geq \ 0$$

Επιπλέον:

(1) $\left|G(\vec{x}_1, \cdots, \vec{x}_m)\right| = 0$ *αν και μόνον αν το σύνολο διανυσμάτων $\mathcal{S}$ είναι γραμμικά εξαρτημένο.*
(2) $\left|G(\vec{x}_1, \cdots, \vec{x}_m)\right| > 0$ *αν και μόνον αν το σύνολο διανυσμάτων $\mathcal{S}$ είναι γραμμικά ανεξάρτητο.*

Απόδειξη. Δείχνουμε πρώτα ότι η ορίζουσα Gram είναι μη-αρνητική. Από την προηγούμενη Πρόταση αρκεί να δείξουμε ότι αν το σύνολο $\mathcal{S}$ είναι γραμμικά ανεξάρτητο, τότε η ορίζουσα Gram είναι θετική.

Έστω $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$ μια ορθοκανονική βάση του $\mathcal{E}$. Από την ανάλυση που προηγήθηκε έπεται ότι:

$$H \cdot {}^tH = G(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$$

Από το Φασματικό Θεώρημα έπεται ότι ο συμμετρικός πίνακας $G(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ είναι διαγωνοποιήσιμος και άρα έχει όλες τις ιδιοτιμές του στο $\mathbb{R}$. Έστω λ μια ιδιοτιμή του $G := G(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ με αντίστοιχο ιδιοδιάνυσμα-στήλη Y:

$$G \cdot Y = \lambda Y, \qquad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Τότε εργαζόμενοι στον Ευκλείδειο χώρο $\mathbb{R}_m$ θα έχουμε:

$$\|Y\|^2 = \langle Y, Y \rangle = {}^tY \cdot Y$$

Άρα:

$$\lambda \|Y\|^2 = \lambda({}^tY \cdot Y) = ({}^tY) \cdot (\lambda \cdot Y) = {}^tY \cdot G \cdot Y = {}^tY \cdot H \cdot {}^tH \cdot Y = {}^t({}^tH \cdot Y) \cdot ({}^tH \cdot Y) =$$

$$= \langle {}^tH \cdot Y, {}^tH \cdot Y \rangle = \|{}^tH \cdot Y\|^2 \ \geq \ 0$$

Η τελευταία σχέση δείχνει ότι $\lambda \geq 0$. Επειδή το σύνολο $\mathcal{S}$ είναι γραμμικά ανεξάρτητο, έπεται ότι η ορίζουσα Gram είναι $\neq 0$ και επομένως ο πίνακας Gram δεν έχει το 0 ως ιδιοτιμή. Καταλήγουμε ότι όλες ιδιοτιμές του πίνακα Gram ανήκουν στο $\mathbb{R}$ και είναι όλες > 0. Επειδή η ορίζουσα ενός πίνακα είναι το γινόμενο των ιδιοτιμών του, αυτό σημαίνει ότι η ορίζουσα Gram είναι θετική. $\square$

11.2. **Διαδικασία Gram-Schmidt.** Υπενθυμίζουμε την Διαδικασία Gram-Schmidt για το *γραμμικά ανεξάρτητο* σύνολο διανυσμάτων

$$\mathcal{S} = \left\{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m\right\}$$

του $\mathcal{E}$:

Υπενθυμίζουμε πρώτα ότι η *ορθογώνια προβολή* του $\vec{y} \neq \vec{0}$ στο $\vec{x}$ ορίζεται να είναι το διάνυσμα:

$$\Pi_{\vec{y}}(\vec{x}) := \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \, \|\vec{y}\|} \cdot \vec{y}$$

Για την απλοποίηση των υπολογισμών που θα ακολουθήσουν, ορίζουμε την *αριθμητική προβολή* του $\vec{y}$ στο $\vec{x}$ να είναι ο αριθμός:

$$\pi_{\vec{y}}(\vec{x}) := \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \, \|\vec{y}\|}$$

Τότε θα έχουμε:

$$\Pi_{\vec{y}}(\vec{x}) = \pi_{\vec{y}}(\vec{x}) \cdot \vec{y}$$

<u>*Κατασκευή Ορθογώνιου Συνόλου*</u> $\mathcal{T} = \left\{\vec{y}_1, \vec{y}_2, \cdots, \vec{y}_m\right\}$

$(\mathbf{GS_1})$ Θέτουμε:

$$\vec{y}_1 = \vec{x}_1$$

$(\mathbf{GS_2})$ Θέτουμε:

$$\vec{y}_2 = \vec{x}_2 - \pi_{y_1}(\vec{x}_1) \cdot \vec{y}_1$$

$$\vdots$$

$(\mathbf{GS_k})$ Επαγωγικά:

$$\vec{y}_k = \vec{x}_k - \pi_{y_1}(\vec{x}_k) \cdot \vec{y}_1 - \pi_{y_2}(\vec{x}_k) \cdot \vec{y}_2 - \cdots - \pi_{y_{k-1}}(\vec{x}_k) \cdot \vec{y}_{k-1} \; = \; \vec{x}_k - \sum_{i=1}^{k-1} \pi_{y_i}(\vec{x}_k) \cdot \vec{y}_i$$

Τότε το σύνολο $\mathcal{T} = \{\vec{y}_1, \vec{y}_2, \cdots, \vec{y}_m\}$ έχει τις ακόλουθες ιδιότητες:

1. Το $\mathcal{T}$ είναι ορθογώνιο: $\langle \vec{y}_i, \vec{y}_j \rangle = 0, \ \forall 1 \leq i \neq j \leq m$.

2. Για κάθε $1 \leq k \leq m$: ο υπόχωρος $\mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_k)$ ο οποίος παράγεται από το σύνολο $\mathcal{S}$ συμπίπτει με τον υπόχωρο $\mathcal{L}(\vec{y}_1, \vec{y}_2, \cdots, \vec{y}_k)$ ο οποίος παράγεται από το σύνολο $\mathcal{T}$:

$$\mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_k) \; = \; \mathcal{L}(\vec{y}_1, \vec{y}_2, \cdots, \vec{y}_k), \qquad 1 \leq k \leq m \tag{11.2}$$

3. Επομένως, για κάθε $2 \leq k \leq m$:

$$\vec{y}_k \; \in \; \mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{k-1})^{\perp} \tag{11.3}$$

4. Αν $\mathcal{L}(\vec{y}_k)$ είναι ο υπόχωρος του $\mathcal{E}$ ο οποίος παράγεται από το $\vec{y}_k$, τότε:

$$\mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_k) = \mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{k-1}) \oplus \mathcal{L}(\vec{y}_k) \qquad \text{(Ορθογώνιο Ευθύ Άθροισμα)}$$

5.

$$0 \; \leq \; \|\vec{y}_k\| \; \leq \; \|\vec{x}_k\| \qquad 1 \leq k \leq m \tag{11.4}$$

Πράγματικά: χρησιμοποιώντας την $(\mathbf{GS_k})$ και τα **2**. και **4**. Θα έχουμε

$$\vec{x}_k = \vec{z}_k + \vec{y}_k$$

$$\vec{z}_k = \pi_{y_1}(\vec{x}_k) \cdot \vec{y}_1 + \pi_{y_2}(\vec{x}_k) \cdot \vec{y}_2 + \cdots + \pi_{y_{k-1}}(\vec{x}_k) \cdot \vec{y}_{k-1} \quad \in \quad \mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{k-1})$$

Επειδή $\langle \vec{z}_k, \vec{y}_k \rangle = 0$, από το Πυθαγόρειο Θεώρημα, θα έχουμε:

$$0 \ \leq \ \|\vec{y}\|^2 \ \leq \ \|\vec{z}_k\|^2 + \|\vec{y}\|^2 = \|\vec{x}_k\|^2 \qquad \Longrightarrow \qquad 0 \ \leq \ \|\vec{y}_k\| \ \leq \ \|\vec{x}_k\|$$

6. Επειδή το σύνολο $\mathcal{T} = \{\vec{y}_1, \vec{y}_2, \cdots, \vec{y}_m\}$ είναι ορθογώνιο, θα έχουμε:

$$G(\vec{y}_1, \cdots, \vec{y}_m) := \begin{pmatrix} \langle \vec{y}_1, \vec{y}_1 \rangle & 0 & \cdots & 0 \\ 0 & \langle \vec{y}_2, \vec{y}_2 \rangle & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \langle \vec{y}_m, \vec{y}_m \rangle \end{pmatrix}$$

και άρα:

$$\left| G(\vec{y}_1, \cdots, \vec{y}_m) \right| = \|\vec{y}_1\|^2 \cdot \|\vec{y}_2\|^2 \cdot \ \cdots \ \cdot \|\vec{y}_m\|^2 \tag{11.5}$$

7. Το σύνολο διανυσμάτων

$$\left\{ \frac{\vec{y}_1}{\|\vec{y}_1\|}, \frac{\vec{y}_2}{\|\vec{y}_2\|}, \cdots, \frac{\vec{y}_m}{\|\vec{y}_m\|} \right\}$$

είναι ορθοκανονικό και άρα αποτελεί μια ορθοκανονική βάση του υπόχωρου $\mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$.

Δείχνουμε τώρα ότι η ορίζουσα Gram των διανυσμάτων $\mathcal{S} = \{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m\}$ είναι ίση με την ορίζουσα Gram των διανυσμάτων $\mathcal{T} = \{\vec{y}_1, \vec{y}_2, \cdots, \vec{y}_m\}$.

Πρόταση 11.4. *Έστω $\mathcal{S} = \{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m\}$ ένα σύνολο διανυσμάτων του Ευκλείδειου χώρου $(\mathcal{E}, \langle , \rangle)$, και έστω $\mathcal{T} = \{\vec{y}_1, \vec{y}_2, \cdots, \vec{y}_m\}$ το ορθογώνιο σύνολο το οποίο προκύπτει από το $\mathcal{S}$ με την διαδικασία Gram-Schmidt. Τότε:*

$$\left| G(\vec{x}_1, \cdots, \vec{x}_m) \right| \ = \ \left| G(\vec{y}_1, \cdots, \vec{y}_m) \right| \ = \ \langle \vec{y}_1, \vec{y}_1 \rangle \cdot \langle \vec{y}_2, \vec{y}_2 \rangle \cdot \ \cdots \ \cdot \langle \vec{y}_m, \vec{y}_m \rangle$$

Απόδειξη. Θα αποδείξουμε τον ισχυρισμό εκτελώντας στοιχειώδεις πράξεις, οι οποίες δεν αλλάζουν την ορίζουσά, στις γραμμές και στις στήλες του πίνακα Gram:

$$G(\vec{x}_1, \cdots, \vec{x}_m) = \begin{pmatrix} \langle \vec{x}_1, \vec{x}_1 \rangle & \langle \vec{x}_1, \vec{x}_2 \rangle & \cdots & \langle \vec{x}_1, \vec{x}_m \rangle \\ \langle \vec{x}_2, \vec{x}_1 \rangle & \langle \vec{x}_2, \vec{x}_2 \rangle & \cdots & \langle \vec{x}_2, \vec{x}_m \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{x}_m, \vec{x}_1 \rangle & \langle \vec{x}_m, \vec{x}_2 \rangle & \cdots & \langle \vec{x}_m, \vec{x}_m \rangle \end{pmatrix}$$

Κατ' αρχήν αν το σύνολο $\mathcal{S}$ είναι γραμμικά εξαρτημένο, τότε προφανώς και το σύνολο $\mathcal{T}$ είναι γραμμικά εξαρτημένο, και τότε $\left| G(\vec{x}_1, \cdots, \vec{x}_m) \right| = 0 = \left| G(\vec{y}_1, \cdots, \vec{y}_m) \right|$.

Υποθέτουμε ότι το $\mathcal{S}$, άρα και το $\mathcal{T}$, είναι γραμμικά ανεξάρτητο.

(1) Θέτουμε $\vec{y}_1 = \vec{x}_1$ παντού στον πίνακα $G(\vec{x}_1, \cdots, \vec{x}_m)$.

(2) Το διάνυσμα $\vec{y}_2$ είναι της μορφής: $\vec{y}_2 = \vec{x}_2 + \kappa \vec{y}_1$. Πολλαπλασιάζουμε κάθε στοιχείο της πρώτης στήλης του πίνακα $G(\vec{x}_1, \cdots, \vec{x}_m)$ με το κ, και ακολούθως προσθέτουμε την προκύπτουσα στήλη στην δεύτερη στήλη του πίνακα $G(\vec{x}_1, \cdots, \vec{x}_m)$. Έπειτα πολλαπλασιάζουμε την πρώτη γραμμή του πίνακα $G(\vec{x}_1, \cdots, \vec{x}_m)$ με το κ και ακολούθως προσθέτουμε την προκύπτουσα γραμμή στην δεύτερη γραμμή του πίνακα $G(\vec{x}_1, \cdots, \vec{x}_m)$.

Έτσι, με τις παραπάνω πράξεις οι οποίες δεν αλλάζουν την ορίζουσα, τα διανύσματα $\vec{y}_1$ και $\vec{y}_2$ εμφανίζονται σε κάθε στοιχείο της ορίζουσας όπου εμφανίζονταν τα $\vec{y}_1$ και $\vec{x}_2$, και επομένως θα έχουμε:

$$\left|G(\vec{x}_1,\cdots,\vec{x}_m)\right| = \mathsf{Det}\begin{pmatrix} \langle\vec{y}_1,\vec{y}_1\rangle & \langle\vec{y}_1,\vec{y}_2\rangle & \langle\vec{y}_1,\vec{x}_3\rangle & \cdots & \langle\vec{y}_1,\vec{x}_m\rangle \\ \langle\vec{y}_2,\vec{y}_1\rangle & \langle\vec{y}_2,\vec{y}_2\rangle & \langle\vec{y}_2,\vec{x}_3\rangle & \cdots & \langle\vec{y}_2,\vec{x}_m\rangle \\ \langle\vec{x}_3,\vec{y}_1\rangle & \langle\vec{x}_3,\vec{y}_2\rangle & \langle\vec{x}_3,\vec{x}_3\rangle & \cdots & \langle\vec{x}_3,\vec{x}_m\rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \langle\vec{x}_m,\vec{y}_1\rangle & \langle\vec{x}_m,\vec{y}_2\rangle & \langle\vec{y}_m,\vec{x}_3\rangle & \cdots & \langle\vec{x}_m,\vec{x}_m\rangle \end{pmatrix}$$

(3) Ακολούθως το διάνυσμα $\vec{y}_3$ είναι της μορφής $\vec{y}_3 = \vec{x}_3 + \lambda_1\vec{y}_1 + \lambda_2\vec{y}_2$. Τότε πολλαπλασιάζουμε κάθε στοιχείο της πρώτης στήλης του πίνακα $G(\vec{y}_1,\cdots,\vec{x}_m)$ με το λ_1, και κάθε στοιχείο της δεύτερης στήλης με λ_2 και προσθέτουμε τις προκύπτουσες στήλες στην τρίτη στήλη του πίνακα $G(\vec{y}_1,\cdots,\vec{x}_m)$. Εκτελούμε τις ίδιες πράξεις με τις γραμμές του πίνακα $G(\vec{y}_1,\cdots,\vec{x}_m)$.

Έτσι, με τις παραπάνω πράξεις οι οποίες δεν αλλάζουν την ορίζουσα, τα διανύσματα $\vec{y}_3$ εμφανίζεται σε κάθε στοιχείο της ορίζουσας όπου εμφανιζόταν τα $\vec{q}_3$, και επομένως θα έχουμε:

$$\left|G(\vec{x}_1,\cdots,\vec{x}_m)\right| \;=\; \mathsf{Det}\begin{pmatrix} \langle\vec{y}_1,\vec{y}_1\rangle & \langle\vec{y}_1,\vec{y}_2\rangle & \langle\vec{y}_1,\vec{y}_3\rangle & \cdots & \langle\vec{y}_1,\vec{x}_m\rangle \\ \langle\vec{y}_2,\vec{y}_1\rangle & \langle\vec{y}_2,\vec{y}_2\rangle & \langle\vec{y}_2,\vec{y}_3\rangle & \cdots & \langle\vec{y}_2,\vec{x}_m\rangle \\ \langle\vec{y}_3,\vec{y}_1\rangle & \langle\vec{y}_3,\vec{y}_2\rangle & \langle\vec{y}_3,\vec{y}_3\rangle & \cdots & \langle\vec{y}_3,\vec{x}_m\rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \langle\vec{x}_m,\vec{y}_1\rangle & \langle\vec{x}_m,\vec{y}_2\rangle & \langle\vec{x}_m,\vec{y}_3\rangle & \cdots & \langle\vec{x}_m,\vec{x}_m\rangle \end{pmatrix}$$

(4) Συνεχίζοντας αυτή τη διαδικασία, αντικαθιστούμε, με στοιχειώδεις πράξεις οι οποίες δεν αλλάζουν την ορίζουσα, όλα τα διάνυσματα $\vec{x}_i$, $1 \leq i \leq m$, τα οποία εμφανίζονται ως (i,j)-στοιχεία $\langle\vec{x}_i,\vec{x}_j\rangle$ της ορίζουσας $\left|G(\vec{x}_1,\cdots,\vec{x}_m)\right|$, με τα διανύσματα $\vec{y}_i$, $1 \leq i \leq m$, τα οποία προκύπτουν από τα $\vec{x}_i$, $1 \leq i \leq m$, με την διαδικασία Gram-Schmidt. Επομένως θα έχουμε:

$$\left|G(\vec{x}_1,\cdots,\vec{x}_m)\right| \;=\; \mathsf{Det}\begin{pmatrix} \langle\vec{y}_1,\vec{y}_1\rangle & \langle\vec{y}_1,\vec{y}_2\rangle & \langle\vec{y}_1,\vec{y}_3\rangle & \cdots & \langle\vec{y}_1,\vec{y}_m\rangle \\ \langle\vec{y}_2,\vec{y}_1\rangle & \langle\vec{y}_2,\vec{y}_2\rangle & \langle\vec{y}_2,\vec{y}_3\rangle & \cdots & \langle\vec{y}_2,\vec{y}_m\rangle \\ \langle\vec{y}_3,\vec{y}_1\rangle & \langle\vec{y}_3,\vec{y}_2\rangle & \langle\vec{y}_3,\vec{y}_3\rangle & \cdots & \langle\vec{y}_3,\vec{y}_m\rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \langle\vec{y}_m,\vec{y}_1\rangle & \langle\vec{y}_m,\vec{y}_2\rangle & \langle\vec{y}_m,\vec{y}_3\rangle & \cdots & \langle\vec{y}_m,\vec{y}_m\rangle \end{pmatrix} \;=\; \left|G(\vec{y}_1,\cdots,\vec{y}_m)\right|$$

Επειδή το σύνολο $\{\vec{y}_1,\cdots,\vec{y}_m\}$ είναι ορθογώνιο, έπεται ότι:

$$\left|G(\vec{x}_1,\cdots,\vec{x}_m)\right| \;=\; \left|G(\vec{y}_1,\cdots,\vec{y}_m)\right| \;=\; \prod_{i=1}^{m}\langle\vec{y}_i,\vec{y}_i\rangle \;=\; \prod_{i=1}^{m}\|\vec{y}_i\|^2 \qquad \square$$

Πρόταση 11.5. *Έστω $\mathcal{S} = \{\vec{x}_1,\vec{x}_2,\cdots,\vec{x}_m\}$ ένα σύνολο διανυσμάτων του Ευκλείδειου χώρου $(\mathcal{E},\langle,\rangle)$, και έστω $\mathcal{T} = \{\vec{y}_1,\vec{y}_2,\cdots,\vec{y}_m\}$ το ορθογώνιο σύνολο το οποίο προκύπτει από το $\mathcal{S}$ με την διαδικασία Gram-Schmidt. Τότε:*

$$0 \;\leq\; \left|G(\vec{x}_1,\cdots,\vec{x}_m)\right| \;\leq\; \|\vec{x}_1\|^2\,\|\vec{x}_2\|^2\cdots\|\vec{x}_m\|^2$$

Επιπλέον:

(1) $\left|G(\vec{x}_1,\cdots,\vec{x}_m)\right| = 0$ *αν και μόνον αν το σύνολο διανυσμάτων $\mathcal{S}$ είναι γραμμικά εξαρτημένο.*
(2) $\left|G(\vec{x}_1,\cdots,\vec{x}_m)\right| > 0$ *αν και μόνον αν το σύνολο διανυσμάτων $\mathcal{S}$ είναι γραμμικά ανεξάρτητο.*
(3) $\left|G(\vec{x}_1,\cdots,\vec{x}_m)\right| \;=\; \|\vec{x}_1\|^2\,\|\vec{x}_2\|^2\cdots\|\vec{x}_m\|^2$ *αν και μόνον αν το σύνολο $\mathcal{S}$ είναι ορθογώνιο.*

Απόδειξη. Από την σχέση (12.4) και την Πρόταση 12.4 έπεται ότι

$$\left|G(\vec{x}_1,\cdots,\vec{x}_m)\right| = \left|G(\vec{y}_1,\cdots,\vec{y}_m)\right| = \|\vec{y}_1\|^2\cdot\|\vec{y}_2\|^2\cdot\;\cdots\;\cdot\|\vec{y}_m\|^2 \;\leq\; \|\vec{x}_1\|^2\cdot\|\vec{x}_2\|^2\cdot\;\cdots\;\cdot\|\vec{x}_m\|^2$$

Άρα $0 \ \leq \ \left|G(\vec{x}_1, \cdots, \vec{x}_m)\right| \ \leq \ \|\vec{x}_1\|^2 \, \|\vec{x}_2\|^2 \cdots \|\vec{x}_m\|^2$. Οι ισχυρισμοί (1) και (2) αποδείχθηκαν στην Πρόταση 12.3.

Προφανώς αν το σύνολο $\mathcal{S}$ είναι ορθογώνιο, τότε η παραπάνω ανισότητα είναι ισότητα. Αντίστροφα εύκολα βλέπουμε με επαγωγή στο πλήθος m των διανυσμάτων και με χρήση της σχέσης (11.4), ότι αν η παραπάνω ανισότητα είναι ισότητα, τότε το σύνολο διανυσμάτων $\mathcal{S}$ είναι ορθογώνιο, □

11.3. **Όγκος Παραλληλεπιπέδου σε Ευκλείδειους Χώρους.** Έστω $\mathcal{S} = \{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m\}$ ένα σύνολο διανυσμάτων του Ευκλείδειου χώρου $(\mathcal{E}, \langle , \rangle)$.

Υπενθυμίζουμε ότι αν $\mathcal{V}$ είναι ένας υπόχωρος του $\mathcal{E}$, τότε:

$$\mathcal{E} = \mathcal{V} \oplus \mathcal{V}^\perp \qquad \text{(Ορθογώνιο Ευθύ Άθροισμα)}$$

και άρα κάθε διάνυσμα $\vec{x}$ του $\mathcal{E}$ γράφεται μοναδικά ως εξής:

$$\vec{x} = \Pi_{\mathcal{V}}(\vec{x}) + \mathsf{K}_{\mathcal{V}}(\vec{x})$$

όπου: $\Pi_{\mathcal{V}}(\vec{x}) \ \in \ \mathcal{V}$ και $\mathsf{K}_{\mathcal{V}}(\vec{x}) \ \in \ \mathcal{V}^\perp$.

Ορισμός 11.6. *Έστω $\mathcal{S} = \{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m\}$ ένα σύνολο διανυσμάτων του Ευκλείδειου χώρου $(\mathcal{E}, \langle , \rangle)$.*

(1) *Το m-**διάστατο παραλληλεπίπεδο** του $\mathcal{E}$ το οποίο ορίζουν τα διανύσματα $\mathcal{S} = \{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m\}$ είναι το σύνολο*

$$\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m) \ = \ \Big\{ \sum_{i=1}^m r_i \vec{x}_i \ | \ 0 \leq r_i \leq 1 \Big\}$$

(2) *Η **βάση** του παραλληλεπιπέδου $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ ορίζεται ως το $(n-1)$-διάστατο παραλληλεπίπεδο $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{m-1})$ το οποίο ορίζουν τα διανύσματα $\{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{m-1}\}$.*

(3) *Το **ύψος** του m-διάστατου παραλληλεπιπέδου $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ ορίζεται να είναι η κάθετη προβολή του διανύσματος $\vec{x}_m$ στον υπόχωρο $\mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{m-1})$ ο οποίος παράγεται απο τα διανύσματα $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{m-1}$:*

$$\vec{\mathsf{h}}_m := \mathsf{K}_{\mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{m-1})}(\vec{x}_m)$$

Παρατήρηση 11.7. *Έστω $\mathcal{E} = \mathbb{R}^3$ εφοδιασμένος με το συνηθισμένο εσωτερικό γινόμενο, ο οποίος είναι ο χώρος που εξετάζει η Αναλυτική Γεωμετρία. Τότε το 1-διάστατο παραλληλεπίπεδο το οποίο ορίζεται από το σύνολο $\mathcal{S} = \{\vec{x}_1\}$ είναι η ευθεία η οποία διέρχεται από το $\vec{0}$ και από το τέλος του διανύσματος $\vec{x}_1$. Το 2-διάστατο παραλληλεπίπεδο το οποίο ορίζεται από το σύνολο $\mathcal{S} = \{\vec{x}_1, \vec{x}_2\}$ είναι το παραλληλόγραμμο με πλευρές τα διανύσματα $\vec{x}_1$ και $\vec{x}_2$. Το 3-διάστατο παραλληλεπίπεδο το οποίο ορίζεται από το σύνολο $\mathcal{S} = \{\vec{x}_1, \vec{x}_2\}$ είναι το παραλληλόγραμμο με πλευρές τα διανύσματα $\vec{x}_1$, $\vec{x}_2$, και $\vec{x}_3$.*

Παρατήρηση 11.8. *Από την Αναλυτική Γεωμετρία το εμβαδόν ενός παραλληλογράμου το οποίο ορίζουν δύο διανύσματα $\vec{x}_1$ και $\vec{x}_2$, είναι το γινόμενο του μήκους του ενός από αυτά, π.χ. το $\vec{x}_1$ το οποίο θεωρούμε ως βάση, με το μήκος της καθέτου από το τέλος του $\vec{x}_2$ στην ευθεία στην οποία κείται το $\vec{x}_1$.*

Παρόμοια ο όγκος ενός παραλληλεπιπέδου το οποίο ορίζουν τρία διανύσματα $\vec{x}_1$, $\vec{x}_2$, και $\vec{x}_3$, είναι το γινόμενο του εμβαδού του παραλληλογράμμου το οποίο σχηματίζεται από δύο εκ των διανυσμάτων, π.χ. $\vec{x}_1$ και $\vec{x}_2$ τα οποία θεωρούνται ως βάση, με το μήκος της καθέτου από το τρίτο διάνυσμα $\vec{x}_3$ στο επίπεδο το οποίο ορίζουν τα $\vec{x}_1$ και $\vec{x}_2$.

Με βάση τις παραπάνω παρατηρήσεις, μπορούμε να ορίσουμε τον όγκο ενός m-διάστατου παραλληλεπίπεδου ως εής:

Ορισμός 11.9. *Ο* **όγκος** $\mathsf{V}_m := \mathsf{V}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ *του* m*-διάστατου παραλληλεπίπεδου του* $\mathcal{E}$ *το οποίο ορίζουν τα διανύσματα* $\mathcal{S} = \{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m\}$ *του* $\mathcal{E}$ *ορίζεται επαγωγικά ως εξής:*

1. *Αν* $m = 1$, *τότε:*

$$\mathsf{V}_1 := \|\vec{x}_1\|$$

2. *Αν* $m = 2$, *τότε:*

$$\mathsf{V}_2 := \mathsf{V}_1 \cdot \|\vec{\mathsf{h}}_1\|$$

όπου:

$$\vec{\mathsf{h}}_1 = \mathsf{K}_{\mathcal{L}(\vec{x}_1)}(\vec{x}_2) \quad \text{και} \quad \mathsf{V}_1 = \text{ μήκος του} \vec{x}_1$$

3. *Αν* $m = 3$, *τότε:*

$$\mathsf{V}_3 := \mathsf{V}_2 \cdot \|\vec{\mathsf{h}}_2\|$$

όπου:

$$\vec{\mathsf{h}}_2 = \mathsf{K}_{\mathcal{L}(\vec{x}_1, \vec{x}_2)}(\vec{x}_3) \quad \text{είναι η κάθετη προβολή του } \vec{x}_3 \text{ στο επίπεδο } \mathcal{L}(\vec{x}_1, \vec{x}_2)$$

και

$$\mathsf{V}_2 = \text{ εμβαδόν παραλληλογράμμου } \mathcal{P}(\vec{x}_1, \vec{x}_2) \text{ το οποίο ορίζουν τα } \vec{x}_1, \vec{x}_2$$

$\vdots$

k. *Για* $1 \leq k \leq m$:

$$\mathsf{V}_k := \mathsf{V}_{k-1} \cdot \|\vec{\mathsf{h}}_{k-1}\|$$

όπου:

$$\vec{\mathsf{h}}_{k-1} = \mathsf{K}_{\mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{k-1})}(\vec{x}_k) \quad \text{είναι η κάθετη προβολή του } \vec{x}_k \text{ στον υπόχωρο } \mathcal{L}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{k-1})$$

και

$$\mathsf{V}_{k-1} = \text{ όγκος παραλληλεπιπέδου } \mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{k-1}) \text{ το οποίο ορίζουν τα } \vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{k-1}.$$

Παρατήρηση 11.10. *Όπως είναι φανερό από τον παραπάνω επαγωγικό ορισμό, ο όγκος* V_m *του* m*-διάστατου παραλληλεπιπέδου* $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ *είναι:*

$$\mathsf{V}_m = \|\vec{x}_1\| \cdot \|\vec{\mathsf{h}}_1\| \cdot \|\vec{\mathsf{h}}_2\| \cdot \cdots \cdot \|\vec{\mathsf{h}}_{m-1}\|$$

Λαμβάνοντας υπ' όψιν τη σχέση (11.4), *την Πρόταση* 11.5, *και το γεγονός ότι τα διανύσματα* $\vec{\mathsf{h}}_i$, $1 \leq i \leq m$ *προκύπτουν με βάση την διαδικασία Gram-Schmidt από τα διανύσματα* $\vec{x}_i$, $1 \leq i \leq m$, *θα έχουμε:*

$$(\mathsf{V}_m)^2 = \big|G(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)\big|$$

Ορισμός 11.11. *Έστω* $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ *το* m*-διάστατο παραλληλεπίπεδο το οποίο ορίζεται από τα διανύσματα* $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m$ *του Ευκλείδειου χώρου* $(\mathcal{E}, \langle , \rangle)$.

Το $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ *καλείται* **ορθογώνιο παραλληλεπίπεδο**, *αντίστοιχα* **ορθοκανονικό παραλληλεπίπεδο**, *αν το σύνολο διανυσμάτων* $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m$ *είναι ορθογώνιο, αντίστοιχα ορθοκανονικό.*

Θεώρημα 11.12. *Έστω* $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ *το* m*-διάστατο παραλληλεπίπεδο το οποίο ορίζεται από τα διανύσματα* $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m$ *του Ευκλείδειου χώρου* $(\mathcal{E}, \langle , \rangle)$. *Τότε το τετράγωνο του όγκου του* $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ *είναι ίσο με την ορίζουσα Gram των διανυσμάτων* $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m$. *Άρα:*

$$\mathsf{V}_m = \pm\sqrt{\big|G(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)\big|}$$

Επιπλέον ο όγκος ενός m*-διάστατου παραλληλεπιπέδου είναι μικρότερος ή ίσος από το γινόμενο των μηκών των πλευρών του και είναι ίσος με αυτό αν και μόνον αν το παραλληλεπίπεδο είναι ορθογώνιο.*

Πόρισμα 11.13. *Έστω $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ το m-διάστατο παραλληλεπίπεδο το οποίο ορίζεται από τα διανύσματα $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m$ του Ευκλείδειου χώρου $(\mathcal{E}, \langle, \rangle)$.*

(1) *Αν το παραλληλεπίπεδο $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ είναι ορθογώνιο, τότε:*

$$\mathsf{V}_m = \|\vec{x}_1\| \cdot \|\vec{x}_2\| \cdot \ \cdots \ \cdot \|\vec{x}_m\|$$

(2) *Αν το παραλληλεπίπεδο $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m)$ είναι ορθοκανονικό, τότε:*

$$\mathsf{V}_m = 1$$

11.4. **Η Γεωμετρική Ερμηνεία της Ορίζουσας και η Ανισότητα του Hadamard.**

Έστω A ένας $n \times n$ πίνακας πραγματικών αριθμών:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Τότε χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι οι στήλες του A είναι οι συνιστώσες διανυσμάτων $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n$ ως προς μια ορθοκανονική βάση $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$ σε έναν Ευκλείδειο χώρο $(\mathcal{E}, \langle, \rangle)$ διάστασης n, για παράδειγμα στον $\mathbb{R}^n$. Έτσι θα έχουμε τα διανύσματα:

$$\vec{x}_1 := a_{11}\vec{e}_1 + a_{21}\vec{e}_2 + \cdots + a_{n1}\vec{e}_n$$

$$\vec{x}_2 := a_{12}\vec{e}_1 + a_{22}\vec{e}_2 + \cdots + a_{n2}\vec{e}_n$$

$$\vdots$$

$$\vec{x}_n := a_{1n}\vec{e}_1 + a_{2n}\vec{e}_2 + \cdots + a_{nn}\vec{e}_n$$

Θεώρημα 11.14. *Έστω A ένας $n \times n$ πίνακας πραγματικών αριθμών. Τότε:*

$$\big|\mathsf{Det}(A)\big| = \mathsf{V}_n$$

όπου V_n είναι ο όγκος του n-διάστατου παραλληλεπιπέδου $\mathcal{P}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n)$ το οποίο σχηματίζεται από τα διανύσματα $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n$ με συνιστώσες σε μια ορθοκανονκή βάση του $\mathbb{R}^n$ τις στήλες του A.

Απόδειξη. Υπολογίζουμε εύκολα ότι:

$$G(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n) = {}^tA \cdot A$$

και άρα χρησιμοποιώντας το Θεώρημα 11.12, θα έχουμε:

$$\mathsf{V}_n^2 = \mathsf{Det}\big(G(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n)\big) = \mathsf{Det}({}^tA \cdot A) = \mathsf{Det}({}^tA) \cdot \mathsf{Det}(A) = \mathsf{Det}(A)^2$$

απ' όπου προκύπτει το ζητούμενο. □

Παρατήρηση 11.15. *Το παραπάνω Θεώρημα δίνει μια γεωμετρική ερμηνεία της ορίζουσας, ακριβέστερα της απόλυτης τιμής της ορίζουσας, ενός τετραγωνικού $n \times n$ πίνακα A ως ο όγκος του n-διάστατου παραλληλεπιπέδου το οποίο σχηματίζεται από τα διανύσματα-στήλες του πίνακα, θεωρούμενες ως συνιστώσες διανυσμάτων $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n$ σε μια ορθοκανονική βάση ενός n-διάστατου Ευκλείδειου χώρου $\mathcal{E}$. Όσον αφορά το πρόσημο αυτό ερμηνεύεται ως ο "προσανατολισμός" των διανυσμάτων αναφορικά με την ορθοκανονική βάση $\mathcal{B}$ του $\mathcal{E}$.*

Προχωρούμε με σκοπό να αποδείξουμε μια σημαντική ανισότητα του Hadamard η οποία δίνει μια εκτίμηση για την ορίζουσα ενός τετραγωνικού πίνακα πραγματικών αριθμών.

Έστω $(\mathbb{R}_n, \langle , \rangle)$ ο Ευκλείδειος χώρος των στηλών πραγματικών αριθμών με n στοιχεία εφοδιασμένος με το συνηθισμένο εσωτερικό γινόμενο.

Υπενθυμίζουμε ότι αν $A \in \mathsf{M}_{n\times n}(\mathbb{R})$ είναι ένας συμμετρικός πίνακας, τότε ο A καλείται **θετικός** αν:

$$\langle A \cdot X, X\rangle \; > \; 0, \qquad \forall X \in \mathbb{R}_n, \quad X \neq 0$$

Ο πίνακας A καλείται **μη-αρνητικός** αν:

$$\langle A \cdot X, X\rangle \; \geq \; 0, \qquad \forall X \in \mathbb{R}_n$$

Λήμμα 11.16. *Έστω $A = (a_{ij})$ ένας συμμετρικός πίνακας πραγματικών αριθμών. Αν ο A είναι θετικός, αντίστοιχα μη-αρνητικός, τότε:*

$$a_{ii} > 0, \qquad \textit{αντίστοιχα} \quad a_{ii} \geq 0, \qquad 1 \leq i \leq n$$

Απόδειξη. Θεωρούμε τα διανύσματα-στήλες της κανονικής βάσης του $\mathbb{R}_n$:

$$E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots, \quad E_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

Τότε εύκολα βλέπουμε ότι:

$$\langle A \cdot E_i, E_i\rangle = a_{ii}$$

από όπου προκύπτει το ζητούμενο. □

Υπενθυμίζουμε το ακόλουθο θεμελιώδες

Θεώρημα 11.17 (Φασματικό Θεώρημα). *Αν $A \in \mathsf{M}_{n\times n}(\mathbb{R})$ είναι ένας συμμετρικός πίνακας, τότε όλες οι ιδιοτιμές του $\lambda_1, \cdots, \lambda_n$ ανήκουν στο $\mathbb{R}$, και υπάρχει ένας ορθογώνιος πίνακα P έτσι ώστε*

$${}^tP \cdot A \cdot P = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Επιπλέον:

(1) *Ο A είναι θετικός $\iff$ ο A είναι αντιστρέψιμος $\iff$ $\lambda_i > 0$, $\forall i = 1, 2, \cdots, n$.*
(2) *Ο A είναι μη-αρνητικός $\iff$ $\lambda_i \geq 0$, $\forall i = 1, 2, \cdots, n$.*

Θεώρημα 11.18. *Αν $A \in \mathsf{M}_{n\times n}(\mathbb{R})$ είναι ένας θετικός συμμετρικός πίνακας, τότε:*

$$\mathsf{Det}(A) \; \leq \; \prod_{i=1}^{n} a_{ii}$$

Απόδειξη. Από το Λήμμα 11.16 έπεται ότι τα διαγώνια στοιχεία a_{ii} του πίνακα A είναι θετικοί αριθμοί. Θέτουμε:

$$\kappa_i = \frac{1}{\sqrt{a_{ii}}}, \qquad 1 \leq i \leq n$$

και έστω ο διαγώνιος πίνακας

$$D = \begin{pmatrix} \kappa_1 & 0 & \cdots & 0 \\ 0 & \kappa_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \kappa_n \end{pmatrix}$$

Θεωρούμε τον πίνακα

$$B := D \cdot A \cdot D$$

και παρατηρούμε ότι ο B είναι επίσης συμμετρικός και τα στοιχεία της διαγωνίου του είναι όλα ίσα με 1:

$$b_{ii} = 1, \qquad 1 \leq i \leq n$$

Επιπλέον ο πίνακας B είναι θετικός, διότι αν X είναι ένα μη-μηδενικό διάνυσμα-στήλη, τότε:

$$\langle B \cdot X, X \rangle = \langle D \cdot A \cdot D \cdot X, X \rangle = \langle A \cdot (D \cdot X), {}^tD \cdot X \rangle = \langle A \cdot (D \cdot X), D \cdot X \rangle \ > \ 0$$

διότι ο A είναι θετικός και προφανώς $D \cdot X \neq 0$.

Για την ορίζουσα του B θα έχουμε:

$$\mathsf{Det}(B) = \mathsf{Det}(D \cdot A \cdot D) = \mathsf{Det}(D)^2 \cdot \mathsf{Det}(A) = \mathsf{Det}(D^2) \cdot \mathsf{Det}(A) = (\frac{1}{\sqrt{a_{ii}}})^2 \cdot \mathsf{Det}(A) = \frac{\mathsf{Det}(A)}{\prod_{i=1}^n a_{ii}}$$

Άρα για να αποδείξουμε το Θεώρημα, αρκεί να δείξουμε ότι: $\mathsf{Det}(B) \leq 1$. Έστω

$$F = \begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n \end{pmatrix}$$

η διαγώνια μορφή του συμμετρικού πίνακα B, όπου $\mu_1, \cdots, \mu_n$ είναι οι ιδιοτιμές του B, οι οποίες είναι θετικοί αριθμοί, όπως προκύπτει από το Φασματικό Θεώρημα, διότι ο B είναι θετικός.

Από την, εύκολα αποδεικνυόμενη σχέση, θετικών πραγματικών αριθμών:

$$\prod_{i=1}^n \mu_i \ \leq \ \sum_{i=1}^n (\frac{\mu_i}{n})^n$$

θα έχουμε:

$$\mathsf{Det}(B) \ = \ \mathsf{Det}(F) \ = \ \prod_{i=1}^n \mu_i \ \leq \ \sum_{i=1}^n (\frac{\mu_i}{n})^n \ \leq \ \left(\frac{\mathsf{Tr}(B)}{n}\right)^n = (\frac{n}{n})^n = 1$$

διότι, καθώς όλα τα διαγώνια στοιχεία του B είναι ίσα με 1, έπεται ότι $\mathsf{Tr}(B) = n$. Άρα $\mathsf{Det}(B) \ \leq \ 1$ και επομένως:

$$\mathsf{Det}(A) \ \leq \ \prod_{i=1}^n a_{ii} \qquad \square$$

Μπορούμε τώρα να αποδείξουμε την ανισότητα του Hadamard:

Θεώρημα 11.19 (Ανισότητα Hadamard)**.** *Έστω* $A = (a_{ij}) \in \mathsf{M}_{n\times n}(\mathbb{R})$*. Τότε:*

$$\mathsf{Det}(A) \;\leq\; \sqrt{\prod_{j=1}^{n} \cdot \sum_{i=1}^{n} a_{ij}^2}$$

Απόδειξη. Αν $\mathsf{Det}(A) = 0$, τότε η ανισότητα ισχύει τετριμμένα.

Υποθέτουμε ότι $\mathsf{Det}(A) \neq 0$, και θεωρούμε τον πίνακα $T = (t_{ij})$, όπου $T := {}^tA \cdot A$. Τότε προφανώς ο T είναι συμμετρικός, και αντιστρέψιμος διότι ο A είναι αντιστρέψιμος. Επομένως από το Φασματικό Θεώρημα έπεται ότι T είναι θετικός. Τότε από το Θεώρημα 11.18, θα έχουμε:

$$\mathsf{Det}(T) \;\leq\; \prod_{j=1}^{n} t_{jj}$$

Όμως προφανώς:

$$t_{jj} = (T)_{jj} = ({}^tA \cdot A)_{jj} = \sum_{i=1}^{n} a_{ij}a_{ij} = \sum_{i=1}^{n} a_{ij}^2$$

και επομένως:

$$\mathsf{Det}(T) \;\leq\; \prod_{j=1}^{n} t_{jj} = \prod_{j=1}^{n}\sum_{i=1}^{n} a_{ij}^2$$

Επειδή $\mathsf{D}(T) = \mathsf{Det}({}^tA \cdot A) = \mathsf{Det}(A^2) = \mathsf{Det}(A)^2$, θα έχουμε:

$$\mathsf{Det}(A)^2 \;=\; \mathsf{Det}(T) \;\leq\; \prod_{j=1}^{n} t_{jj} = \prod_{j=1}^{n}\sum_{i=1}^{n} a_{ij}^2$$

απ' όπου προκύπτει το ζητούμενο. □

Προφανώς η ισότητα στην Ανισότητα Hadamard ισχύει όταν ο πίνακας είναι διαγώνιος.

Ώς άμεση συνέπεια της Ανισότητας Hadamard, έχουμε το ακόλουθο

Πόρισμα 11.20. *Έστω* $A = (a_{ij}) \in \mathsf{M}_{n\times n}(\mathbb{R})$*. Αν* $|a_{ij}| \leq \varepsilon,\ 1 \leq i, i \leq n$*, τότε:*

$$\mathsf{Det}(A) \;\leq\; \varepsilon^n \cdot n^{\frac{n}{2}}$$

Πόρισμα 11.21. *Από όλα τα* n*-διάστατα παραλληλεπίπεδα* $\mathcal{P}(\vec{x}_1, \cdots, \vec{x}_n)$ *με δεδομένο μήκος* $\|\vec{x}_i\|$ *διανυσμάτων πλευρών, εκείνο με τον μεγαλύτερο όγκο είναι το ορθογώνιο.*

Παράδειγμα 11.22. *Θωρούμε τον* 4×4 *πίνακα*

$$A = \begin{pmatrix} 2 & -2 & 1 & 2 \\ -1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & -2 \end{pmatrix}$$

Τότε θέτοντας $\varepsilon = 2$*, από το Πόρισμα* 11.20 *θα έχουμε* $\mathsf{Det}(A) \leq 2^4 \cdot 4^{\frac{4}{2}} = 16 \cdot 16 = 256$*. Πραγματικά:* $\mathsf{Det}(A) = 12$.

Παράδειγμα 11.23. *Θωρούμε τον* 3×3 *πίνακα*

$$A = \begin{pmatrix} 5 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 3 \end{pmatrix}$$

Τότε θέτοντας $\varepsilon = 5$, *από το Πόρισμα* 11.20 *θα έχουμε* $\mathsf{Det}(A) \leq 5^3 \cdot 3^{\frac{3}{2}}$ *που είναι περίπου* 156. *Πραγματικά:* $\mathsf{Det}(A) = 50$.

12. **Ισομετρίες**

12.1. **Χαρακτηρισμός Ισομετριών.** Έστω $(\mathcal{E}, \langle , \rangle)$ ένας Ευκλείδειος χώρος πεπερασμένης διάστασης.

Ορισμός 12.1. *Μια, όχι κατ΄ ανάγκην γραμμική, απεικόνιση $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ καλείται* **ισομετρία** *αν η f διατηρεί αποστάσεις, δηλαδή:*

$$\|f(\vec{x}) - f(\vec{y})\| \;=\; \|\vec{x} - \vec{y}\|, \qquad \forall \vec{x}, \vec{y} \in \mathcal{E}$$

Μια απεικόνιση $g\colon \mathcal{E} \longrightarrow \mathcal{E}$ καλείται **μεταφορά**, κατά το διάνυσμα $\vec{a}$, αν:

$$g(\vec{x}) = \vec{x} + \vec{a}, \qquad \forall \vec{x} \in \mathcal{E}$$

Την μεταφορά κατά το διάνυσμα $\vec{a}$ θα την συμβολίζουμε με $g_{\vec{a}}$.

Προφανώς κάθε μεταφορά είναι ισομετρία. Στην παρούσα παράγραφο με τον όρο ισομετρία θα ενοούμε ισομετρία με την έννοια του ορισμού 12.1.

Θεώρημα 12.2. (1) *Κάθε ισομετρία $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μπορεί να γραφεί μοναδικά ως σύνθεση μιας μεταφοράς $g_{\vec{a}}$ και μιας ισομετρίας h η οποία στέλνει το $\vec{0}$ στο $\vec{0}$:*

$$f = g_{\vec{a}} \circ h, \qquad \text{η} \qquad h\colon \mathcal{E} \longrightarrow \mathcal{E} \quad \text{είναι ισομετρία} \qquad \text{και} \qquad h(\vec{0}) = \vec{0}$$

δηλαδή: $f(\vec{x}) = h(\vec{x}) + \vec{a}$, $\forall \vec{x} \in \mathcal{E}$.

(2) *Τα ακόλουθα είναι ισοδύναμα:*
   (α΄) *Η $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι ισομετρία και: $f(\vec{0}) = \vec{0}$.*
   (β΄)

$$\langle\, f(\vec{x}), f(\vec{y})\,\rangle \;=\; \langle\, \vec{x}, \vec{y}\rangle, \qquad \forall \vec{x}, \vec{y} \in \mathcal{E}$$

Ιδιαίτερα: $\|f(\vec{x})\| \;=\; \|\vec{x}\|$, $\forall \vec{x} \in \mathcal{E}$.

(3) *Έστω $\mathcal{B} = \{\vec{e}_1, \cdots, \vec{e}_n\}$ μια ορθοκανονική βάση του $\mathcal{E}$ και $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια ισομετρία. Αν $f(\vec{0}) = \vec{0}$ και $f(\vec{e}_i) = \vec{e}_i$, $1 \leq i \leq n$, τότε: $f = \mathrm{Id}_{\mathcal{E}}$.*

Απόδειξη. □

Θεώρημα 12.3. *Κάθε ισομετρία $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ για την οποία ισχύει ότι $f(\vec{0}) = \vec{0}$, είναι γραμμική απεικόνιση, και επομένως είναι ένας ισομορφισμός.*

Απόδειξη. □

Πόρισμα 12.4. *Για μια απεικόνιση $f\colon \mathcal{E} \longrightarrow \mathcal{E}$, τα ακόλουθα είναι ισοδύναμα:*

(1) *Η f είναι ισομετρία και $f(\vec{0}) = \vec{0}$.*
(2) *Η f είναι γραμμική ισομετρία.*

Θεώρημα 12.5. (1) *Έστω $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια ισομετρία, τότε υπάρχει μοναδικό διάνυσμα $\vec{a} \in \mathcal{E}$ και μια γραμμική ισομετρία $h\colon \mathcal{E} \longrightarrow \mathcal{E}$ έτσι ώστε:*

$$f = g_{\vec{a}} \circ h$$

(2) *Αντίστροφα κάθε απεικόνιση της παραπάνω μορφής είναι ισομετρία.*

Πόρισμα 12.6. *Έστω $f\colon \mathbb{R}_n \longrightarrow \mathbb{R}_n$ μια απεικόνιση. Τότε τα ακόλουθα είναι ισοδύναμα:*

(1) *Η f είναι ισομετρία.*

(2) *Υπάρχει μοναδικό διάνυσμα-στήλη $Y \in \mathbb{R}_n$ και μοναδικός ορθογώνιος πίνακας A έτσι ώστε:*

$$f(X) = A \cdot X + Y, \qquad \forall X \in \mathbb{R}_n$$

12.2. **Κανονική μορφή Ορθογωνίων Πινάκων.** Έστω $(\mathcal{E}, \langle , \rangle)$ ένας Ευκλείδειος χώρος πεπερασμένης διάστασης $\mathsf{dim}_{\mathbb{R}} \mathcal{E} = n \geq 2$.

Συμβολισμός: Υπενθυμίζουμε ότι ένα ευθύ άθροισμα $\mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_k$ υπόχωρων καλείται **ορθογώνιο ευθύ άθροισμα** αν οι υπόχωροι $\mathcal{V}_i$, $1 \leq i \leq k$ είναι ανά δύο ορθογώνιοι, δηλαδή:

$$\mathcal{V}_i \perp \mathcal{V}_j, \qquad 1 \leq i \neq j \leq k$$

με άλλα λόγια: $\langle \vec{x}_i, \vec{x}_j \rangle = 0,\ \forall \vec{x}_i \in \mathcal{V}_i,\ \forall \vec{x}_j \in \mathcal{V}_j$.

Σ' αυτή την περίπτωση θα γράφουμε το ορθογώνιο ευθύ άθροισμα ως εξής: $\mathcal{V}_1 \boxplus \mathcal{V}_2 \boxplus \cdots \boxplus \mathcal{V}_k$.

Από τώρα και στο εξής: με τον όρο ισομετρία θα εννούμε γραμμική ισομετρία.

Θεώρημα 12.7. *Αν $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι μια ισομετρία, τότε:*

$$\mathcal{E} = \mathcal{V}_1 \boxplus \mathcal{V}_2 \boxplus \cdots \boxplus \mathcal{V}_k$$

όπου κάθε υπόχωρος $\mathcal{V}_i$ είναι f-αναλλοίωτος, δηλαδή: $f(\mathcal{V}_i) \subseteq \mathcal{V}_i$, και $\mathsf{dim}_{\mathbb{R}} \mathcal{V} = 1$ ή 2.

Απόδειξη. Η απόδειξη θα γίνει με χρήση επαγωγής σε κάποια βήματα.

Βήμα 1: Ο Ευκλείδειος χώρος $\mathcal{E}$ έχει έναν f-αναλλοίωτο υπόχωρο $\mathcal{V}$ με $\mathsf{dim}_{\mathbb{R}} \mathcal{V} = 1$ ή 2.

Απόδειξη: Έστω $P_f(t)$ το χαρακτηρσιτικό πολυώνυμο της f. Επειδή τα μόνα ανάγωγα πολυώνυμα με συντελεστές από το $\mathbb{R}$ είναι βαθμού 1 ή 2, έπεται ότι μπορούμε να γράψουμε το $P_f(t)$ ως γινόμενο αναγώγων πολυωνύμων

$$P_f(t) = P_1(t) \cdot P_2(t) \cdot \ \cdots \ \cdot P_k(t) \qquad\qquad \deg P_i(t) = 1 \text{ ή } 2$$

Από το Θεώρημα των Cayley-Hamilton, θα έχουμε $P_f(f) = 0$, και άρα $P_f(f)(\vec{x}) = \vec{0}$, $\forall \vec{x} \in \mathcal{E}$.

□

Θεώρημα 12.8. *Έστω $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ μια ισομετρία. Τότε υπάρχει ορθοκανονική βάση $\mathcal{B}$ του $\mathcal{E}$ έτσι ώστε ο πίνακας της f στην $\mathcal{B}$ να είναι της μορφής:*

$$A = \mathsf{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{pmatrix} A_1 & \mathbb{O} & \cdots & \mathbb{O} \\ \mathbb{O} & A_2 & \cdots & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{O} & \mathbb{O} & \cdots & A_k \end{pmatrix} \qquad (*)$$

όπου:

$$A_i = \begin{pmatrix} \cos(\varphi_i) & \sin(\varphi_i) \\ -\sin(\varphi_i) & \cos(\varphi_i) \end{pmatrix}, \qquad 1 \leq i \leq k-2$$

$$A_{k-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \quad \in \quad \mathsf{M}_{r \times r}(\mathbb{R})$$

$$A_k = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix} \quad \in \quad \mathsf{M}_{s\times s}(\mathbb{R})$$

$$2(k-2)+r+s = \dim_{\mathbb{R}}\mathcal{E}$$

Ιδιαίτερα κάθε ορθογώνιος πίνακας είναι όμοιος με έναν πίνακα της μορφής $(*)$.

12.3. **Ανακλάσεις.** Στο παρόν εδάφιο θα ασχοληθούμε με ανακλάσεις οι οποίες είναι ειδικού τύπου ισομετρίες.

Ορισμός 12.9. *Μια γραμμική απεικόνιση* $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ *καλείται* **ανάκλαση** *αν υπάρχει ένα υπερεπίπεδο* $\mathcal{H}$ *του* $\mathcal{E}$, *το* **υπερεπίπεδο ανάκλασης**, *δηλαδή ένας υπόχωρος* $\mathcal{H}$ *του* $\mathcal{E}$ *διάστασης* $\dim_{\mathbb{R}}\mathcal{H} = n-1$, *έτσι ώστε:*

$$f(\vec{x}) = \vec{x}, \quad \forall \vec{x} \in \mathcal{H} \qquad \text{και} \qquad f(\vec{x}) = -\vec{x}, \quad \forall \vec{x} \in \mathcal{H}^{\perp}$$

Πρόταση 12.10. *Έστω* $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ *μια ανάκλαση ως προς ένα υπερεπίπεδο* $\mathcal{H}$ *του* $\mathcal{E}$. *Τότε υπάρχει ένα μη-μηδενικό διάνυσμα* $\vec{v} \in \mathcal{H}^{\perp}$, *έτσι ώστε:*

$$f(\vec{x}) = \vec{x} - 2\,\frac{\langle\, \vec{x}, \vec{v}\,\rangle}{\langle\, \vec{v}, \vec{v}\,\rangle}\cdot \vec{v}, \qquad \forall\, \vec{x} \,\in\, \mathcal{E}$$

Αν $\vec{w} \in \mathcal{H}^{\perp}$ *είναι ένα άλλο διάνυσμα έτσι ώστε να ικανοποιείται η παραπάνω σχέση, τότε* $\vec{w} = \lambda\vec{v}$, *για κάποιο* $\lambda \in \mathbb{R}$.

Απόδειξη. □

Σύμφωνα με την παραπάνω Πρόταση, μια ανάκλαση $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ προσδιορίζεται από ένα μη-μηδενικό διάνυσμα $\vec{v} \in \mathcal{E}$, δηλαδή είναι της μορφής $f = f_{\vec{v}}$, όπου

$$f_{\vec{v}}\,:\; \mathcal{E} \longrightarrow \mathcal{E}, \qquad f_{\vec{v}}(\vec{x}) = \vec{x} - 2\,\frac{\langle\, \vec{x}, \vec{v}\,\rangle}{\langle\, \vec{v}, \vec{v}\,\rangle}\cdot \vec{v}, \qquad \forall\, \vec{x} \,\in\, \mathcal{E}$$

και τότε το υπερεπίπεδο ανάκλασης είναι: $\mathcal{H} = \vec{v}^{\perp}$.

Πρόταση 12.11. *Κάθε ανάκλαση* $f\colon \mathcal{E} \longrightarrow \mathcal{E}$ *είναι ισομετρία και αυτοπροσαρτημένη. Ιδιαίτερα* $f^2 = \mathrm{Id}_{\mathcal{E}}$ *και υπάρχει ορθοκανονική βάση* $\mathcal{B}$ *του* $\mathcal{E}$ *στην οποία ο πίνακας της* f *είναι της μορφής:*

$$\begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Απόδειξη. □

Λήμμα 12.12. *Έστω* $\vec{v}$, $\vec{w} \in \mathcal{E}$ *δύο διανύσματα του* $\mathcal{E}$. *Τότε τα ακόλουθα είναι ισοδύναμα:*

(1) $\|\vec{v}\| = \|\vec{w}\|$.
(2) *Υπάρχει μια ανάκλαση* $f\colon \mathcal{E} \longrightarrow \mathcal{E}$, *έστι ώστε:* $f(\vec{v}) = \vec{w}$.

Επιπλέον $f(\vec{w}) = \vec{v}$, *και αν* $\vec{v} \neq \vec{0} \neq \vec{w}$, *τότε η ανάκλαση* f *για την οποία ισχύει* $f(\vec{v}) = \vec{w}$ *είναι μοναδική και ίση με την ανάκλαση:*

$$f = f_{\vec{v}-\vec{w}}$$

Απόδειξη. □

Θεώρημα 12.13 (Cartan-Dieudonné)**.** *Αν* $f\colon \mathcal{E} \longrightarrow \mathcal{E}$, *είναι μια γραμμική απεικόνιση, τότε τα ακόλυθα είναι ισοδύναμα:*

(1) *Η* f *είναι ισομετρία.*
(2) *Η* f *είναι σύνθεση* k *το πλήθος ανακλάσεων, δηλαδή υπάρχουν μη-μηδενικά διανύσματα* $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k$, $1 \leq k \leq n = \mathsf{dim}_{\mathbb{R}}\mathcal{E}$, *έτσι ώστε:*

$$f = f_{\vec{v}_1} \circ f_{\vec{v}_1} \circ \cdots \circ f_{\vec{v}_k}$$

Απόδειξη. □

13. Παραγοντοποιήσεις Πινάκων και Γραμμικών Απεικονίσεων

13.1. **Η Παραγοντοποίηση QR ενός πίνακα.** Έστω A ένας $n \times n$ πίνακας πραγματικών αριθμών:

$$A = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix}$$

Στην παρούσα παράγραφο, με χρήση της διαδικασίας Gram-Schmidt θα δείξουμε ότι αν ο A είναι αντιστρέψιμος, τότε ο A μπορεί να γραφεί με μοναδικό τρόπο ως γινόμενο ενός ορθογώνιου πίνακα και ενός άνω τριγωνικού πίνακα ο οποίος έχει θετικούς αριθμούς στην διαγώνιό του.

Θεώρημα 13.1. *Έστω A ένας αντιστρέψιμος $n \times n$ πίνακας πραγματικών αριθμών. Τότε υπάρχουν:*

(1) *ένας ορθογώνιος πίνακας Q, και*
(2) *ένας άνω τριγωνικός πίνακας R όλα τα στοιχεία της διαγωνίου του οποίου είναι θετικοί αριθμοί,*

έτσι ώστε:

$$A = Q \cdot R$$

και η παραπάνω παραγοντοποίηση είναι μοναδική. Δηλαδή αν $A = Q_1 \cdot R_1$, όπου Q_1 είναι ένας ορθογώνιος πίνακας, και R_1 είναι ένας άνω τριγωνικός πίνακας R όλα τα στοιχεία της διαγωνίου του οποίου είναι θετικοί αριθμοί, τότε: $Q = Q_1$ και $R = R_1$.

Απόδειξη. Έστω

$$\Sigma_i = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}, \qquad 1 \leq j \leq n$$

η j-στήλη του πίνακα A. Τότε ο πίνακας A ορίζει το σύνολο διανυσμάτων

$$\Big\{ \Sigma_1, \ \Sigma_2, \ \cdots, \ \Sigma_n \Big\}$$

του Ευκλείδειου χώρου $(\mathbb{R}_n, \langle , \rangle)$. Επειδή ο πίνακας A είναι αντιστρέψιμος, έπεται ότι οι στήλες του $\Sigma_1, \Sigma_2, \cdots, \Sigma_n$ αποτελούν ένα γραμμικά ανεξάρτητο σύνολο διανυσμάτων του $\mathbb{R}_n$. Έστω

$$Q_j = \begin{pmatrix} q_{1j} \\ q_{2j} \\ \vdots \\ q_{nj} \end{pmatrix}, \qquad 1 \leq j \leq n$$

το σύνολο διανυσμάτων το οποίο προκύπτει από το σύνολο των στηλών του A με την διαδικασία Gram-Schmidt. Ως γνωστόν τότε το διάνυσμα Q_j είναι γραμμικός συνδυασμός των $\Sigma_1, \cdots, \Sigma_j$, και έτσι θα έχουμε, βλέπε εδάφιο 11.2, ότι το σύνολο :

$$Q_1 = c_{11}\Sigma_1$$

$$Q_2 = c_{12}\Sigma_1 + c_{22}\Sigma_2$$

$$\vdots$$

$$Q_n = c_{1n}\Sigma_1 + c_{2n}\Sigma_2 + \cdots + c_{nn}\Sigma_n$$

είναι ένα ορθοκανονικό σύνολο διανυσμάτων του $\mathbb{R}_n$, και $c_{jj} \neq 0,\ 1 \leq j \leq n$. Οι παραπάνω σχέσεις μπορούν να γραφούν με μορφή πινάκων ως εξής:

$$\Big(Q_1,\ Q_2,\ \cdots,\ Q_n\Big) \ = \ \Big(\Sigma_1,\ \Sigma_2,\ \cdots,\ \Sigma_n\Big) \ \cdot \ \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ 0 & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_{nn} \end{pmatrix}$$

Επειδή $c_{jj} \neq 0$, $1 \leq j \leq n$, ο άνω τριγωνικός πίνακας

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ 0 & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_{nn} \end{pmatrix}$$

είναι αντιστρέψιμος και τότε ο αντίστροφός του

$$C^{-1} = \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ 0 & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix}$$

θα είναι άνω τριγωνικός. Έτσι θα έχουμε:

$$\Big(\Sigma_1,\ \Sigma_2,\ \cdots,\ \Sigma_n\Big) \ = \ \Big(Q_1,\ Q_2,\ \cdots,\ Q_n\Big) \ \cdot \ \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ 0 & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix} \qquad (*)$$

και $d_{jj} \neq 0$, $1 \leq j \leq n$. Τότε θα έχουμε τις σχέσεις:

$$\begin{aligned} \Sigma_1 &= d_{11}Q_1 \\ \Sigma_2 &= d_{12}Q_1 + d_{22}Q_2 \\ &\vdots \\ \Sigma_n &= d_{1n}Q_1 + d_{2n}Q_2 + \cdots + d_{nn}Q_n \end{aligned}$$

Αν κάποιο από τα d_{jj} είναι < 0, τότε πολλαπλασιάζουμε κάθε στήλη Q_j με το -1 και έτσι χωρίς να επηρεάζεται η ορθοκανονικότητα του συνόλου $\Big\{Q_1, Q_2, \cdots, Q_n\Big\}$, μπορούμε να γράψουμε τις παραπάνω σχέσεις ως εξής:

$$\begin{aligned} \Sigma_1 &= r_{11}Q_1' \\ \Sigma_2 &= r_{12}Q_1' + r_{22}Q_2' \\ &\vdots \\ \Sigma_n &= r_{1n}Q_1' + r_{2n}Q_2' + \cdots + r_{nn}Q_n' \end{aligned}$$

όπου

$$r_{11},\ r_{22},\ \cdots,\ r_{nn}\ >\ 0$$

Οι παραπάνω σχέσεις γράφονται με μορφή πινάκων:

$$\Big(\Sigma_1,\ \Sigma_2,\ \cdots,\ \Sigma_n\Big) \ = \ \Big(Q_1',\ Q_2',\ \cdots,\ Q_n'\Big) \ \cdot \ \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix} \qquad (*)$$

ή ισοδύναμα:

$$A = Q \cdot R$$

όπου

$$Q = \Big(Q'_1,\ Q'_2,\ \cdots,\ Q'_n\Big) \qquad \text{και} \qquad R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix}$$

η οποία είναι η ζητούμενη παραγοντοποίηση, διότι ο πίνακας Q είναι ορθογώνιος επειδή οι στήλες του αποτελούν ορθοκανονική βάση του $\mathbb{R}_n$, και ο πίνακας R είναι άνω τριγωνικός με θετικά στοιχεία στην διαγώνιο.

Έστω $P = (p_{ij})$ ένας ορθογώνιος πίνακας και $S = (s_{ij})$ ένας άνω τριγωνικός πίνακας με θετικά στοιχεία στην διαγώνιο, έτσι ώστε: $A = P \cdot S$. Τότε επειδή οι πίνακας R είναι αντιστρέψιμος, θα έχουμε:

$$Q \cdot R = P \cdot S \qquad \Longrightarrow \qquad P^{-1} \cdot Q = S \cdot R^{-1}$$

Επειδή οι πίνακες P^{-1} και Q είναι ορθογώνιοι, έπεται ότι και ο πίνακας $P^{-1} \cdot Q$ θα είναι ορθογώνιος. Άρα ο πίνακας $S \cdot R^{-1}$ ορθογώνιος. Όμως ο πίνακας $S \cdot R^{-1}$ είναι άνω τριγωνικός (ως γινόμενο άνω τριγωνικών πινάκων). Επειδή, όπως μπορούμε να δούμε εύκολα ο μόνος ορθογώνιος και άνω τριγωνικός πίνακας με θετικά στοιχεία στην κύρια διαγώνιο είναι ο μοναδιαίος, έπεται ότι $S \cdot R^{-1} = I_n$ και άρα $S = R$. Τότε προφανώς $Q = P$. □

Παρατήρηση 13.2. *Η παραγοντοποίηση QR ενός αντιστρέψιμου πίνακα χρησιμοποιείται στην επίλυση γραμμικών συστημάτων: Έστω $A \in \mathsf{M}_{n\times n}(\mathbb{R}$ και έχτω το γραμμικό σύστημα*

$$A \cdot X = B$$

Έστω $A = Q \cdot R$, όπου Q είναι ένας ορθογώνιος πίνακας, δηλαδή $Q^{-1} = {}^tQ$, και R είναι ένας άνω τριγωνικός πίνακας με θετικά στοιχεία στην κύρια διαγώνιο. Τότε θα έχουμε το σύστημα

$$Q \cdot R \cdot X = B \qquad \Longrightarrow \qquad R \cdot X = {}^tQ \cdot B$$

το οποίο είναι της μορφής

$$\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} b'_1 \\ b'_2 \\ \cdot \\ b'_n \end{pmatrix}$$

όπου στα δεξιά είναι ο πίνακας-στήλη ${}^tQ \cdot B$, και το οποίο επιλύεται εύκολα διότι ο πίνακας R είναι άνω τριγωνικός και $r_{jj} > 0$, $1 \leq j \leq n$. Για παράδειγμα

$$x_n = \frac{r_{nn}}{b'_n}, \qquad x_{n-1} = \frac{1}{r_{n-1,n-1}}\big(b'_{n-1} - r_{n-1,n}\frac{r_{nn}}{b'_n}\big), \qquad \cdots$$

Παρατήρηση 13.3. *Η παραγοντοποίηση QR, χωρίς την μοναδικότητα, ισχύει και για μη-τετραγωνικούς (και ιδιαίτερα και για μη-αντιστρέψιμους) πίνακες:*

Κάθε πίνακας $A \in \mathsf{M}_{m\times n}(\mathbb{R})$ μπορεί να γραφεί ως γινόμενο

$$A = Q \cdot R$$

όπου $Q \in \mathsf{M}_{m\times n}(\mathbb{R})$ είναι ένας πίνακας του οποίου οι στήλες αποτελούν ένα ορθοκανονικό σύνολο, και $R \in \mathsf{M}_{n\times n}(\mathbb{R})$ είναι ένας άνω τριγωνικός πίνακας με μη-αρνητικά στοιχεία στην κύρια διαγώνιο.

Η απόδειξη είναι παρόμοια με την απόδειξη του Θεωρήματος 13.1.

13.2. **Πολική Ανάλυση.** Έστω $z \in \mathbb{C}$ ένας μη-μηδενικός μιγαδικός αριθμός. Τότε ως γνωστόν ο z μπορεί να γραφεί ως:

$$z = r \cdot e^{i\theta}, \qquad r > 0, \qquad |e^{i\theta}| = 1$$

Διαφορετικά: $z = r \cdot w$, όπου $w = \frac{z}{|z|}$ και $r = |z|$.

Στην παρούσα παράγραφο θα δούμε μια ανάλογη παραγοντοποίηση για γραμμικές απεικονίσεις σε Ευκλείδειους χώρους.

Θεώρημα 13.4. *Έστω $(\mathcal{E}, \langle , \rangle)$ ένας Ευκλείδειος χώρος πεπερασμένης διάστασης και $f : \mathcal{E} \longrightarrow \mathcal{E}$ μια αντιστρέψιμη γραμμική απεικόνιση. Τότε η f μπορεί να γραφεί μοναδικά ως:*

$$f = g \circ h$$

όπου:

(1) *Η γραμμική απεικόνιση $g\colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι αυτοπροσαρτημένη και θετική.*
(2) *Η γραμμική απεικόνιση $h\colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι ισομετρία.*

Απόδειξη. Υποθέτουμε ότι η f είναι ισομορφισμός. Τότε όπως γνωρίζουμε η αυτοπροσαρτημένη της f^* είναι επίσης ισομορφισμός και ισχύει $(f^*)^{-1} = (f^{-1})^*$. Θεωρούμε την γραμμική απεικόνιση

$$f \circ f^* : \mathcal{E} \longrightarrow \mathcal{E}$$

η οποία είναι θετική διότι, $\forall \vec{x} \neq \vec{0}$:

$$\langle f \circ f^*)(\vec{x}), \vec{x}\rangle = \langle f(f^*(\vec{x})), \vec{x}\rangle = \langle f^*(\vec{x}), f^*(\vec{x})\rangle = \|f^*(\vec{x})\|^2 > 0$$

και η τελευταία ανισότητα ισχύει διότι $\|f^*(\vec{x})\|^2 \geq 0$ και $\|f^*(\vec{x})\|^2 = 0$ αν και μόνον αν $f^*(\vec{x}) = 0$ και άρα αν και μόνον αν $\vec{x} = \vec{0}$ το οποίο είναι άτοπο διότι $\vec{x} \neq 0$. Άρα:

$$f \circ f^* > 0$$

Επειδή όπως γνωρίζουμε κάθε θετική γραμμική απεικόνιση έχει μια μοναδική θετική τετραγωνική ρίζα, έπεται ότι υπάρχει μοναδική αυτοπροσαρτημένη θετική γραμμική απεικόνιση $g\colon \mathcal{E} \longrightarrow \mathcal{E}$ έτσι ώστε:

$$f \circ f^* = g^2$$

Επειδή κάθε θετική γραμμική απεικόνιση είναι ισομορφισμός, μπορούμε να Θεωρήσουμε τότε την γραμμική απεικόνιση

$$h := g^{-1} \circ f$$

και άρα

$$h^* = (g^{-1} \circ f)^* = f^* \circ (g^{-1})^* = f^* \circ (g^*)^{-1} = f^* \circ g^{-1}$$

και επομένως

$$h \circ h^* = h \circ f^* \circ g^{-1} = g^{-1} \circ f \circ f^* \circ g^{-1} = g^{-1} \circ g^2 \circ g^{-1} = \mathrm{Id}_{\mathcal{E}}$$

το οποίο δείχνει ότι η h είναι ισομετρία. Επειδή $h := g^{-1} \circ f$ θα έχουμε $f = g \circ h$, όπου η g είναι αυτοπροσαρτημένη θετική γραμμική απεικόνιση και η h είναι ισομετρία.

Για την μοναδικότητα: έστω $f = g_1 \circ h_1$, όπου η g_1 είναι θετική αυτοπροσαρτημένη και η h_1 είναι ισομετρία. Τότε χρησιμοποιώντας ότι η h_1 είναι ισομετρία, θα έχουμε:

$$f^* = (g_1 \circ h_1)^* = h_1^* \circ g_1^* = h_1^* \circ g_1 \implies g^2 = f \circ f^* = f \circ h_1^* \circ g_1 = g_1 \circ h_1 \circ h_1^* \circ g_1 = g_1^2$$

Επομένως $g^2 = g_1^2$ και άρα $g = \sqrt{g^2} = \sqrt{g_1^2} = g_1$, λόγω της μοναδικότητας της θετικής τετραγωνικής ρίζας θετικής αυτοπροσαρτημένης γραμμικής απεικόνισης. Τέλος θα έχουμε:

$$h = g^{-1} \circ f = g_1^{-1} \circ f = h_1 \qquad \square$$

Πόρισμα 13.5. *Κάθε αντιστρέψιμος $n \times n$ πίνακας πραγματικών αριθμών A γράφεται μοναδικά ως:*

$$A = B \cdot C$$

όπου:

(1) *Ο πίνακας B είναι συμμετρικός και θετικός.*
(2) *Ο πίνακας C είναι ορθογώνιος.*

Απόδειξη. Θεωρούμε την γραμμική απεικόνιση

$$f_A : \mathbb{R}_n \longrightarrow \mathbb{R}_n, \quad f_A(X) = A \cdot X$$

Από το παραπάνω Θεώρημα η f_A γράφεται μοναδικά ως:

$$f_A = g \circ h$$

όπου $g \colon \mathbb{R}_n \longrightarrow \mathbb{R}_n$ είναι είναι συμμετρικός και θετικός, και η $h \colon \mathbb{R}_n \longrightarrow \mathbb{R}_n$ είναι ισομετρία. Τότε θεωρώντας πίνακες στην κανονική βάση $\mathcal{B}$ του $\mathbb{R}_n$ η οποία είναι ορθοκανονική, θα έχουμε $A = B \cdot C$, διότι $M^{\mathcal{B}}_{\mathcal{B}}(f_A = A$, και ο πίνακας B της αυτοπροσαρτημένης θετικής γραμμικής απεικόνισης g είναι συμμετρικός και θετικός, και ο πίνακας C της ισομετρίας h είναι ορθογώνιος. □

Πρόταση 13.6. *Έστω $f, g : \mathcal{E} \longrightarrow \mathcal{E}$ δύο αυτοπροσαρτημένες θετικές γραμμικές απεικονίσεις, και έστω $h : \mathcal{E} \longrightarrow \mathcal{E}$ μια ισομετρία. Υποθέτουμε ότι:*

$$f = g \circ h \qquad \text{ή} \qquad f = h \circ g$$

Τότε $h = \mathrm{Id}_{\mathcal{E}}$ και $f = g$.

Απόδειξη. Υποθέτουμε ότι: $f = g \circ h$. Επειδή οι f, g είναι θετικές, έπεται ότι f, g είναι αντιστρέψιμες. Τότε:

$$f = g \circ h \qquad \Longrightarrow \qquad h = g^{-1} \circ f$$

και επομένως χρησιμοποιώντας οτι οι f, g είναι αυτοπροσαρτημένες, θα έχουμε:

$$h^* = (g^{-1} \circ f)^* = f^* \circ (g^{-1})^* = f \circ (g^*)^{-1} = f \circ g^{-1}$$

Επίσης χρησιμοποιώντας ότι η h είναι ισομετρία, θα έχουμε $h^* \circ h = \mathrm{Id}_{\mathcal{E}}$ και άρα

$$h^* = h^{-1} \qquad \Longrightarrow \qquad f \circ g^{-1} = (g^{-1} \circ f)^{-1} = f^{-1} \circ g \qquad \Longrightarrow \qquad f^2 = g^2$$

Λόγω της μοναδικότητας της θετικής τετραγωνικής ρίζας αυτοπροσαρτημένων θετικών γραμμικών απεικονίσεων, θα έχουμε $f = g$ και άρα $h = \mathrm{Id}_{\mathcal{E}}$.

Παρόμοια εργαζόμαστε αν $f = h \circ g$. □

Θα δείξουμε τώρα ότι το Θεώρημα 13.4 ισχύει για κάθε γραμμική απεικόνιση f, όχι κατ' ανάγκην αντιστρέψιμη.

Θεώρημα 13.7. *Έστω $(\mathcal{E}, \langle , \rangle)$ ένας Ευκλείδειος χώρος πεπερασμένης διάστασης και $f : \mathcal{E} \longrightarrow \mathcal{E}$ μια αντιστρέψιμη γραμμική απεικόνιση. Τότε η f μπορεί να γραφεί ως:*

$$f = g \circ h$$

όπου:

(1) *Η γραμμική απεικόνιση $g \colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι αυτοπροσαρτημένη και μη-αρνητική.*
(2) *Η γραμμική απεικόνιση $h \colon \mathcal{E} \longrightarrow \mathcal{E}$ είναι ισομετρία.*

Απόδειξη. Υποθέτουμε πρώτα ότι η f είναι δεν είναι ισομορφισμός. Τότε η γραμμική απεικόνιση $f \circ f^* : \mathcal{E} \longrightarrow \mathcal{E}$ είναι αυτοπροσαρτημένη και μη-αρνητική, και άρα έχει μια μοναδική μη-αρνητική τετραγωνική ρίζα $g \colon \mathcal{E} \longrightarrow \mathcal{E}$, δηλαδή έτσι ώστε:

$$f \circ f^* = g^2$$

Τότε

$$\|g(\vec{x})\|^2 = \langle g(\vec{x}, g(\vec{x})\rangle = \langle \vec{x}, g^*(g(\vec{x}))\rangle = \langle \vec{x}, g^2(\vec{x}\rangle = \langle \vec{x}, (f \circ f^*)(\vec{x})\rangle = \langle (f \circ f^*)(\vec{x}), \vec{x}\rangle =$$
$$= \langle f^*(\vec{x}), f^*(\vec{x})\rangle = \|f^*(\vec{x})\|^2$$

Ορίζουμε μια γραμμική απεικόνιση

$$\widetilde{h} : \mathsf{Im}(g) \longrightarrow \mathcal{E}, \quad \widetilde{h}(g(\vec{x})) := f^*(\vec{x})$$

Η $\widetilde{h}$ είναι καλά ορισμένη διότι αν $g(\vec{x}), g(\vec{y}) \in \mathsf{Im}(g)$ και $g(\vec{x}) = g(\vec{y})$, τότε θα έχουμε:

$$\|g(\vec{x})\|^2 = \|g(\vec{y})\|^2 \implies \|f^*(\vec{x})\|^2 = \|f^*(\vec{y})\|^2 \implies \|f^*(\vec{x} - \vec{y})\| = 0 \implies f^*(\vec{x}) = f^*(\vec{y})$$

Επομένως έχουμε μια ισομετρία $\widetilde{h} \colon \mathsf{Im}(g) \longrightarrow \mathcal{E}$. □

13.3. **Ορθογώνια Τριγωνοποίηση.** Όπως γνωρίζουμε, για κάθε τετραγωνικό πίνακα A ο οποίος έχει όλες τις ιδιοτιμές του στο $\mathbb{K}$, είναι όμοιος με έναν άνω τριγωνικό πίνακα: υπάρχει αντιστρέψιμος πίνακας P έτσι ώστε ο πίνακας $P^{-1} \cdot A \cdot P$ να είναι άνω τριγωνικός.

Αν $\mathbb{K} = \mathbb{R}$, τότε το ακόλουθο Θεώρημα δείχνει ότι ο πίνακας A είναι ορθογώνια όμοιος με έναν άνω τριγωνικό πίνακα.

Θεώρημα 13.8 (Θεώρημα του Schur)**.** *Αν A είναι ένας τετραγωνικός πίνακας πραγματικών αριθμών ο οποίος έχει όλες τις ιδιοτιμές του στο $\mathbb{R}$, τότε υπάρχει ορθογώνιος πίνακας P έτσι ώστε ο πίνακας:*

$$P^{-1} \cdot A \cdot P \qquad \text{είναι άνω τριγωνικός}$$

Απόδειξη. □

14. **Αντισυμμετρικοί Πίνακες**

15. **Η Μέθοδος των Ελαχίστων Τετραγώνων**

16. **Ο Δυϊκός Χώρος**

17. **Κανονικοί Ενδομορφισμοί και Κανονικοί Πίνακες**

18. **Τετραγωνικές Μορφές**

19. **Μέτρο Πίνακα**

20. Βιβλιογραφία

1. PETER LAX, *Linear Algebra and its applications*, John Wiley & Sons, (2007).
2. SHELDOM AXLER, *Linear Algebra done right*, Springer-Verlag, (1996).
3. S. TREIL, *Linear Algebra done wrong*, http://www.math.brown.edu/ treil/papers/LADW/LADW.html
4. PAUL HALMOS, *Finite dimensional vector spaces*, Springer-Verlag, (1974).
5. PAUL HALMOS, *Linear Algebra Problem Book*, American Mathematical Society, (1995).
6. K. HOFFMAN AND R. KUNZE, *Linear Algebra*, Prentice Hall, (1971).
7. SERGE LANG, *Linear Algebra*, Springer-Verlag, (1987).
8. STEVEN ROMAN, *Advanced Linear Algebra*, Springer-Verlag, (1992).
9. E. VINBERG, *A Course in Algebra*, American Mathematical Society, (2003).
10. J. GOLAN, *The Linear Algebra a Beginning Graduate Student Ought to Know*, Kluwer, (2004).
11. SETH WARNER, *Modern Algebra*, Dover, (1990).
12. S. FRIEDBERG, A. INSEL, AND L. SPENCE, *Linear Algebra*, Prentice Hall, (1989).
13. I. M. GELFAND, *Lectures on Linear Algebra*, Interscience, (1961).
14. G. SHILOV, *Linear Algebra*, Dover, (1971).
15. I.V. PROSKURYAKOV, *Problems in Linear Algebra*, Mir, (1978).
16. V. PRASOLOV, *Problems and Theorems in Linear Algebra*, American Mathematical Society, (1994).
17. W. GREUB, *Linear Algebra*, Springer-Verlag, (1967).
18. I. N. HERSTEIN, *Topics in Algebra*, Wiley, (2006).
19. M. POSTNIKOV, *Lectures in Geometry, Semesters I-V*, Mir, (1986).