ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΑΝΟΙΚΤΑ ΑΚΑΔΗΜΑΪΚΑ ΜΑΘΗΜΑΤΑ

Τίτλος Μαθήματος: Γραμμική Άλγεβρα ΙΙ

Ενότητα: Ελάχιστο Πολυώνυμο

Διδάσκων: Καθηγητής Νικόλαος Μαρμαρίδης, Καθηγητής Ιωάννης Μπεληγιάννης

Τμήμα: Μαθηματικών

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

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 άμεσδα 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 $\mathsf{dim}_{\mathbb{K}}\mathcal{E} = n$. 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 $m \geq 1$.*

Ανοικτά Ακαδημαϊκά Μαθήματα

Πανεπιστήμιο Ιωαννίνων

Τέλος Ενότητας

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Ιωαννίνων**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Σημειώματα

Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Ιωαννίνων, Διδάσκων: Καθηγητής Νικόλαος Μαρμαρίδης, Καθηγητής Ιωάννης Μπεληγιάννης «Γραμμική Άλγεβρα ΙΙ». Έκδοση: 1.0. Ιωάννινα 2014. Διαθέσιμο από τη δικτυακή διεύθυνση: http://ecourse.uoi.gr/course/view.php?id=1249.