VI

Κανονικές Μορφές

VI.1 Ιδιοτιμές - Ιδιοδιανύσματα

Αν έχουμε ένα διανυσματικό χώρο V, με μία γραμμική συνάρτηση $L : V \to V$ και ένα υπόχωρο $W \subset V$, τότε ο περιορισμός της L στον W δεν δίνει πάντα μια συνάρτηση $W \to W$. Πράγματι, θα μπορούσε να υπάρχει $w \in W$ με $L(w) \in L \backslash W$. Για παράδειγμα η γραμμική συνάρτηση $L : \mathbb{R}^2 \to \mathbb{R}^2$ η οποία ως προς την κανονική βάση $= \{e_1, e_2\}$ έχει πίνακα

$$(L, B, B) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

στέλνει το e_1 στο $L(e_1) = e_1 + e_2$, συνεπώς η L αν περιοριστεί στον $W = \langle e_1 \rangle$ δεν δίνει συνάρτηση $W \to W$.

Ορισμός VI.1.1. Έστω V ένας $\mathbb{F}$-διανυσματικός χώρος και $L : V \to V$ μια γραμμική αποικόνιση. Ένας υπόχωρος $W \subset V$ θα λέγεται L-αναλλοίωτος αν και μόνο αν $L(W) \subset W$.

Παράδειγμα VI.1.2. Οι τετριμμένοι υπόχωροι $V \subset V$ και $\{0\} \subset V$ είναι L-αναλλοίωτοι.

Λήμμα VI.1.3. *Ένας μονοδιάστατος υπόχωρος $\langle v \rangle \subset V$ είναι L-αναλλοίωτος αν και μόνο αν $Lv = \lambda v$, για κάποιο $\lambda \in \mathbb{F}$.*

Απόδειξη. Αν $Lv = \lambda v$ για κάποιο $\lambda \in \mathbb{F}$ τότε $L(\langle v \rangle) \subset \langle v \rangle$. Αντιστρόφως αν $L(\langle v \rangle) \subset \lambda v$, τότε $L(v) \in \langle v \rangle$, άρα $L(v) = \lambda v$ για κάποιο $\lambda \in \mathbb{F}$. □

Ορισμός VI.1.4. Έστω V ένας $\mathbb{F}$-διανυσματικός χώρος και $L : V \to V$ μία γραμμική απεικόνιση. Ένα $\lambda \in \mathbb{F}$ θα λέγεται ιδιοτιμή αν υπάρχει μή μηδενικό $v \in V$, ώστε $L(v) = \lambda v$. Το διάνυσμα v θα λέγεται ιδιοδιάνυσμα ως προς την ιδιοτιμή λ. Ο χώρος όλων των ιδιοδιανυσμάτων μιας ιδιοτιμής λ, θα λέγεται ιδιόχωρος της λ. Θα συμβολίζουμε τον ιδιόχωρο της ιδιοτιμής λ με E_λ.

Παρατήρηση VI.1.5. Ένα στοιχείο $v \in E_\lambda$ αν και μόνο αν $Lv - \lambda \mathrm{Id}_V v = 0$ αν και μόνο αν $v \in \mathrm{Ker}(L - \lambda \mathrm{Id}_V)$. Συνεπώς $E_\lambda = \mathrm{Ker}(L - \lambda \mathrm{Id}_V)$ είναι υπόχωρος του V. Θέλουμε ο υπόχωρος αυτός να μην είναι ο μηδενικός και αυτό συμβαίνει αν και μόνο αν $\det(L - \lambda \mathrm{Id}_V) = 0$.

Λήμμα VI.1.6. *Έστω* $L : V \to V$ *γραμμική συνάρτηση,* λ_1, λ_2 *ιδιοτιμές και* $E_{\lambda_1}, E_{\lambda_2}$ *οι ιδιόχωροι. Αν* $\lambda_1 \neq \lambda_2$ *έχουμε* $E_{\lambda_1} \cap E_{\lambda_2} = \{\mathbf{0}_V\}$.

Απόδειξη. Αν $v \in E_{\lambda_1}$ τότε $Lv = \lambda_1 v$ και αν $v \in E_{\lambda_2}$ τότε $Lv = \lambda_2 v$. Συνεπώς $(\lambda_1 - \lambda_2)v = \not{\mathcal{V}}_V$ και αφού $\lambda_1 - \lambda_2 \neq 0$ καταλήγουμε στο ότι $v = \mathbf{0}_V$. □

Ορισμός VI.1.7. Αν ο χώρος V έχει πεπερασμένη διάσταση τότε το πολυώνυμο

$$\mathrm{Ch}_L(x) = \det(L - x\mathrm{Id}_V)$$

θα λέγεται χαρακτηριστικό πολυώνυμο της L.

Παρατήρηση VI.1.8. Αν B είναι μια βάση του V, και $A = [L, B, B]$ είναι ο πίνακας της L ως προς την βάση B, τότε

$$\mathrm{Ch}_L(x) = \det(A - x\mathbb{I}_n). \tag{VI.1}$$

Παρατηρούμε ότι η έκφραση στην εξίσωση (VI.1) είναι ανεξάρτητη της επιλογής της βάσης B. Πράγματι, για μια διαφορετική επιλογή βάσης B', αντί για τον πίνακα A θα είχαμε τον πίνακα QAQ^{-1}, Q αντιστρέψιμος. Όμως

$$\det\left(QAQ^{-1} - x\mathbb{I}_n\right) = \det\left(QAQ^{-1} - xQ\mathbb{I}_nQ^{-1}\right) = \det(Q)\det(A - x\mathbb{I}_n)\det\left(Q^{-1}\right) = \det(A - x\mathbb{I}_n).$$

Η ιδιότητα αυτή εκφράζεται συχνά και ως: «όμοιοι πίνακες έχουν το ίδιο χαρακτηριστικό πολυώνυμο».

Λήμμα VI.1.9. *Αν* $A \in \mathbb{F}^{n,n}$, *η ποσότητα*

$$\det(A - x\mathbb{I}_n)$$

είναι ένα πολυώνυμο βαθμού n.

Απόδειξη. Υπολογίζουμε το χαρακτηριστικό πολυώνυμο με την βοήθεια της πρότασης III.6.4

$$\det(A - x\mathbb{I}_n) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)},$$

όπου $b_{ij} = a_{ij} - x\delta_{ij}$. Είναι σαφές ότι ο μεγαλύτερος βαθμός του πολυωνύμου $b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n,\sigma(n)}$ είναι n, οπότε το ζητούμενο έπεται. □

Παρατήρηση VI.1.10. Ο σταθερός όρος του χαρακτηριστικού πολυωνύμου είναι η ορίζουσα του L, αφού $\mathrm{Ch}_L(0) = \det(L)$.

Παρατήρηση VI.1.11. Οι ρίζες λ του χαρακτηριστικού πολυωνύμου είναι ακριβώς οι τιμές $\lambda \in \mathbb{F}$, για τις οποίες το ομογενές σύστημα

$$Lv - \lambda v = 0$$

έχει, εκτός από την μηδενική λύση και άλλες λύσεις, δηλαδή οι ιδιοτιμές.

Λήμμα VI.1.12. *Έστω* $\lambda_1, \dots, \lambda_n$ *ανά δύο διαφορετικές ιδιοτιμές μιας γραμμικής συνάρτησης* $L : V \to V$ *με αντίστοιχα ιδιοδιανύσματα* $v_1, \dots, v_n$. *Τότε τα* $v_1, \dots, v_n$ *είναι γραμμικά ανεξάρτητα.*

Απόδειξη. Θα δόσουμε μια επαγωγική απόδειξη στο πλήθος n των ιδιοδιανυσμάτων $v_1, \dots, v_n$. Αν $n = 1$, τότε το $v_1 \neq 0$ είναι γραμμικά ανεξάρτητο. Υποθέτουμε ότι ισχυρισμός είναι αληθής για $n-1$ το πλήθος ιδιοδιανύσματα. Έστω $v_1, \dots, v_n$ ιδιοδιανύσματα τα οποία αντιστοιχούν στις διαφορετικές ανά δύο ιδιοτιμές $\lambda_1, \dots, \lambda_n$. Έστω ότι αυτά ικανοποιούν μια σχέση

$$a_1 v_1 + a_2 v_2 + \dots + a_n v_n = 0. \tag{VI.2}$$

Εφαρμόζουμε την L για να πάρουμε

$$a_1 L(v_1) + a_2 L(v_2) + \dots + a_n L(v_n) = a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 + \dots + a_n \lambda_n v_n = L(\mathbf{0}_V) = \mathbf{0}_V. \tag{VI.3}$$

Πολλαπλασιάζουμε την σχέση (VI.2) με λ_n και αφαιρούμε την σχέση (VI.3) για να πάρουμε:

$$a_1(\lambda_n - \lambda_1) v_1 + \dots + (\lambda_n - \lambda_{n-1}) v_{n-1} = 0.$$

Από την επαγωγική υπόθεση έχουμε ότι τα $v_1, \dots, v_{n-1}$ είναι γραμμικά ανεξάρτητα, συνεπώς $a_\nu(\lambda_n - \lambda_\nu) = 0$ για κάθε $1 \leq \nu \leq n-1$. Επειδή $\lambda_n - \lambda_\nu \neq 0$ θα πρέπει $a_1 = \dots = a_{n-1} = 0$. Τότε από την σχέση (VI.2) έχουμε ότι και $a_n = 0$, άρα τα $v_1, \dots, v_n$ είναι γραμμικά ανεξάρτητα. □

Πόρισμα VI.1.13. *Έστω ότι το χαρακτηριστικό πολυώνυμο* $\mathrm{Ch}_L(x)$ *έχει* n*-το πλήθος διαφορετικές ρίζες, οι οποίες είναι ιδιοτιμές. Τότε υπάρχει μια βάση* $B = \{v_1, \dots, v_n\}$ *αποτελούμενη από ιδιοδιανύσματα ώστε*

$$(L, B, B) = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

Απόδειξη. Τα n το πλήθος ιδιοδιανύσματα $v_1, \dots, v_n$ που αντιστοιχούν στις διαφορετικές ανά δύο ιδιοτιμές είναι γραμμικά ανεξάρτητα και είναι τόσα όσα η διάσταση, άρα αποτελούν μια βάση B του V, σύμφωνα με το θεώρημα IV.3.12. Η διαγώνια μορφή του πίνακα (L, B, B) προκύπτει από το ότι $Lv_i = \lambda_i v_i$ για κάθε $1 \leq i \leq n$. □

Παρατήρηση VI.1.14. Αν δουλέψουμε σε ένα αλγεβρικά κλειστό σώμα τότε το χαρακτηριστικό πολυώνυμο της $L : V \to V$ έχει $\dim V = n$ το πλήθος ρίζες.

Ορισμός VI.1.15. Ένας πίνακας $A \in \mathbb{F}^{n,n}$ λέγεται διαγωνοποιήσιμος αν υπάρχει αντιστρέψιμος πίνακας $Q \in \mathbb{F}^{n,n}$ ώστε

$$Q^{-1} A Q = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

Παρατήρηση VI.1.16. Υπάρχουν πίνακες που δεν είναι διαγωνοποιήσιμοι. Για παράδειγμα ο πίνακας

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

δεν είναι διαγωνοποιήσιμος. Παρατηρούμε ότι το χαρακτηριστικό πολυώνυμο του A είναι $(x-1)^2$. Πράγματι αν υπήρχε αντιστρέψιμος πίνακας $Q \in \mathbb{F}^{2,2}$ ώστε

$$Q^{-1} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} Q = \mathrm{diag}(\lambda_1, \lambda_2), \tag{VI.4}$$

τότε $\mathrm{Ch}_{\mathrm{diag}(\lambda_1,\lambda_2)}(x) = (x-\lambda_1)(x-\lambda_2)$ το οποίο όμως θα πρέπει να είναι ίσο με το $(x-1)^2$ αφού όμοιοι πίνακες έχουν το ίδιο χαρακτηριστικό πολυώνυμο. Συνεπώς $\lambda_1 = \lambda_2 = 1$ και η εξίσωση (VI.4) γίνεται

$$Q^{-1}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}Q = \mathbb{I}_2 \Rightarrow \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} = \mathbb{I}_2,$$

το οποίο είναι άτοπο.

Ορισμός VI.1.17. Θεωρούμε μια ιδιοτιμή λ της γραμμικής συνάρτησης $L: V \to V$.

- Θα ονομάζουμε *αλγεβρική πολλαπλότητα* της ιδιοτιμής λ τον φυσικό αριθμό $\mu(\lambda)$ για τον οποίο ισχύει
$$\mathrm{Ch}_L(x) = (x-\lambda)^{\mu(\lambda)}q(x),$$
όπου $q(x) \in \mathbb{F}[x]$ με $q(\lambda) \neq 0$, δηλαδή την πολλαπλότητα της ρίζας λ στο χαρακτηριστικό πολυώνυμο.

- θα ονομάζουμε *γεωμετρική πολλαπλότητα* $\rho(\lambda)$ της ιδιοτιμής λ την διάσταση $\dim E_\lambda$ του ιδιόχωρου της ιδιοτιμής λ.

Πρόταση VI.1.18. *Για μια ιδιοτιμή λ της γραμμικής συνάρτησης L η αλγεβρική πολλαπλότητα $\mu(\lambda)$ και η γεωμετρική πολλαπλότητα $\rho(\lambda)$ ικανοποιούν την σχέση*

$$1 \leq \rho(\lambda) \leq \mu(\lambda) \leq \dim(V) = n.$$

Απόδειξη. Η πρώτη και η τελευταία ανισότητα είναι προφανής. Θα αποδείξουμε ότι $\rho(\lambda) \leq \mu(\lambda)$. Θεωρούμε μια βάση $\{v_1,\dots,v_{\rho(\lambda)}\}$ του ιδιόχωρου E_λ την οποία επεκτείνουμε σε μια βάση $B = \{v_1,\dots,v_{\rho(\lambda)},w_1,\dots,w_k\}$ του V. Γράφουμε τον πίνακα της L ως προς την βάση B

$$(L,B,B) = \begin{pmatrix}
\lambda & 0 & \cdots & 0 & a_{1,\rho(\lambda)+1} & \cdots & a_{1n}\\
0 & \lambda & \ddots & \vdots & a_{2,\rho(\lambda)+1} & \cdots & a_{2n}\\
\vdots & \ddots & \ddots & 0 & \vdots & & \vdots\\
0 & \cdots & 0 & \lambda & a_{\rho(\lambda),\rho(\lambda)+1} & \cdots & a_{\rho(\lambda),n}\\
0 & \cdots & \cdots & 0 & a_{\rho(\lambda)+1,\rho(\lambda)+1} & \cdots & a_{\rho(\lambda)+1,n}\\
\vdots & & & \vdots & \vdots & & \vdots\\
0 & \cdots & \cdots & 0 & a_{n,\rho(\lambda)+1} & \cdots & a_{n,n}
\end{pmatrix} = \left(\begin{array}{c|c}\lambda\mathbb{I}_{\rho(\lambda)} & A\\ \hline \mathbf{0}_{n-\rho(\lambda),\rho(\lambda)} & A'\end{array}\right).$$

Παρατηρούμε ότι το χαρακτηριστικό πολυώνυμο του παραπάνω πίνακα

$$\det\left(\begin{array}{c|c}(\lambda-x)\mathbb{I}_{\rho(\lambda)} & A\\ \hline \mathbf{0}_{n-\rho(\lambda),\rho(\lambda)} & A'-x\mathbb{I}_{n-\rho(\lambda)}\end{array}\right) = (x-\lambda)^{\rho(\lambda)}(-1)^{\rho(\lambda)}g(x).$$

διαιρείται από το $(x-\lambda)^{\rho(\lambda)}$, από όπου προκύπτει η ζητούμενη ανισότητα, η οποία μπορεί να είναι γνήσια αν $g(\lambda) = 0$. □

Πρόταση VI.1.19. *Ένας πίνακας είναι διαγωνοποιήσιμος αν και μόνο αν για κάθε ιδιοτιμή* $\rho(\lambda) = \mu(\lambda)$.

Απόδειξη. Ας υποθέσουμε ότι πίνακας $A \in \mathbb{F}^{n,n}$ είναι διαγωνοποιήσιμος. Υπάρχει μια βάση $v_1, \dots, v_n$ ώστε $Av_i = \lambda_i v_i$, δηλαδή μια βάση αποτελούμενη από ιδιοδιανύσματα. Έστω $\lambda_{i_1}, \dots, \lambda_{i_r}$ οι ανά δύο διαφορετικές ιδιοτιμές. Το χαρακτηριστικό πολυώνυμο υπολογίζεται ως

$$\mathrm{Ch}_A(x) = \prod_{\nu=1}^{s} (\lambda_{i_\nu} - x)^{\mu(\lambda_{i_\nu})}.$$

Είναι σαφές ότι

$$n = \dim \mathbb{F}^n = \sum_{\nu=1}^{s} \rho(\lambda_{i_\nu}) \leq \sum_{\nu=1}^{s} \mu(\lambda_{i_\nu}) = \deg \mathrm{Ch}_A(x) = n$$

από όπου προκύπτει ότι $\rho(\lambda) = \mu(\lambda)$ για κάθε ιδιοτιμή λ.

Αντιστρόφως αν για κάθε ιδιοτιμή λ έχουμε $\rho(\lambda) = \mu(\lambda)$ τότε μπορούμε να βρούμε

$$n = \deg \mathrm{Ch}_A(x) = \sum_{\nu=1}^{s} \mu(\lambda_{i_\nu}) = \sum_{\nu=1}^{s} \rho(\lambda_{i_\nu})$$

το πλήθος γραμμικά ανεξάρτητα ιδιοδιανύσματα τα οποία αποτελούν μια βάση του χώρου $\mathbb{F}^n$ και συνεπώς ο πίνακας A είναι διαγωνοποιήσιμος. □

Παρατήρηση VI.1.20. Υποθέτουμε ότι ο πίνακας $A \in \mathbb{F}^{n,n}$ είναι διαγωνοποιήσιμος, δηλαδή υπάρχει ένας αντιστρέψιμος πίνακας $Q \in \mathbb{F}^{n,n}$ και διαγώνιος πίνακας $\Delta = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ ώστε

$$Q^{-1}AQ = \Delta \Leftrightarrow AQ = Q\Delta.$$

Αν γράψουμε $Q = (q^1, q^2, \dots, q^n)$, όπου $q^1, \dots, q^n$ είναι οι γραμμικά ανεξάρτητες στήλες του Q, τότε έχουμε $Aq^\nu = \lambda_\nu q^\nu$, για $1 \leq \nu \leq n$, δηλαδή ο πίνακας Q έχει ως στήλες ιδιοδιανύσματα.

Μέθοδος 2 (Διαγωνοποίηση). *Για να διαγωνοποιήσουμε ένα $n \times n$ πίνακα εκτελούμε τα παρακάτω βήματα:*

1. *Υπολογισμός χαρακτηριστικού πολυωνύμου*
2. *Εύρεση ιδιοτιμών $\lambda_1, \dots, \lambda_n$*
3. *Υπολογισμός βάσεων B_i των ιδιοχώρων E_{λ_i}.*
4. *Ο πίνακας Q είναι ένας πίνακας που έχει ως στήλες τα στοιχεία των βάσεων B_i.*

Παράδειγμα VI.1.21. Θέλουμε να διαγωνοποιήσουμε τον πίνακα

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Υπολογίζουμε το χαρακτηριστικό πολυώνυμο

$$\mathrm{Ch}_A(x) = \det(A - x\mathbb{I}_2) = \det \begin{pmatrix} -x & 1 \\ 1 & 1-x \end{pmatrix} = x^2 - x - 1 = \left(x - \frac{1-\sqrt{5}}{2}\right)\left(x + \frac{1-\sqrt{5}}{2}\right)$$

Συνεπώς οι ιδιοτιμές είναι οι $\frac{1-\sqrt{5}}{2}, \frac{1+\sqrt{5}}{2}$. Στην συνέχεια υπολογίζουμε τους ιδιόχωρους

Για την ιδιοτιμή $\frac{1-\sqrt{5}}{2}$ ο $E_{\frac{1-\sqrt{5}}{2}}$ είναι ο χώρος λύσεων της

$$\left(A - \frac{1-\sqrt{5}}{2}\mathbb{I}_2\right)x = 0 \text{ ισοδύναμα } \begin{pmatrix} \frac{1}{2}(\sqrt{5}-1) & 1 \\ 1 & 1 + \frac{1}{2}(\sqrt{5}-1) \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

το οποίο λύνουμε με την μέθοδο απαλοιφής του Gauss και παίρνουμε ως βάση το χώρου λύσεων το $((-1-\sqrt{5})/2, 1)^t$.
Ομοίως για την ιδιοτιμή $\frac{1+\sqrt{5}}{2}$ ο $E_{\frac{1+\sqrt{5}}{2}}$ είναι ο χώρος λύσεων της

$$\left(A - \frac{1+\sqrt{5}}{2}\mathbb{I}_2\right)x = 0 \text{ ισοδύναμα } \begin{pmatrix} \frac{1}{2}\left(-1-\sqrt{5}\right) & 1 \\ 1 & 1+\frac{1}{2}\left(-1-\sqrt{5}\right) \end{pmatrix}$$

ο οποίος έχει βάση του χώρου λύσεων το $((-1+\sqrt{5})/2, 1)^t$

Σχηματίζουμε τον πίνακα Q ο οποίος έχει ως πρώτη στήλη το μοναδικό στοιχείο της βάσης του ιδιόχωρου της ιδιοτιμής $(1-\sqrt{5})/2$ και ως δεύτερη στήλη το μοναδικό στοιχείο της βάσης του ιδιόχωρου της ιδιοτιμής $(1+\sqrt{5})/2$, δηλαδή

$$Q = \begin{pmatrix} \frac{-1-\sqrt{5}}{2} & \frac{-1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}$$

Υπολογίζουμε ότι

$$Q^{-1} = \begin{pmatrix} -\frac{1}{\sqrt{5}} & -\frac{1-\sqrt{5}}{2\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{-1-\sqrt{5}}{2\sqrt{5}} \end{pmatrix}$$

και τέλος επαληθεύουμε ότι

$$Q^{-1}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} Q = \begin{pmatrix} \frac{(1-\sqrt{5})}{2} & 0 \\ 0 & \frac{(1+\sqrt{5})}{2} \end{pmatrix}. \tag{VI.5}$$

Παράδειγμα VI.1.22. Ας προσπαθήσουμε να διαγωνοποιήσουμε τον πίνακα με την παραπάνω μέθοδο.

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

Υπολογίζουμε το χαρακτηριστικό πολυώνυμο

$$\mathrm{Ch}_A(x) = \det(A - x\mathbb{I}_3) = -(x-2)^3.$$

Η μοναδική ιδιοτιμή είναι το 2. Στην συνέχεια υπολογίζουμε τον ιδιόχωρο E_2 ο οποίος δίνεται από τις λύσεις του ομογενούς συστήματος

$$(A - 2\mathbb{I}_3)x = 0 \text{ ισοδύναμα } \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

το οποίο λύνουμε με την μέθοδο απαλοιφής του Gauss, ο πίνακας είναι ήδη σε κλιμακωτή μορφή και ο χώρος λύσεων είναι μονοδιάστατος με βάση $(1, 0, 0)^t$. Δεν μπορούμε να βρούμε μια βάση του $\mathbb{R}^3$ αποτελούμενη από ιδιοδιανύσματα, οπότε ο πίνακας A, δεν είναι διαγωνοποιήσιμος.

Παράδειγμα VI.1.23. Θέλουμε να διαγωνοποιήσουμε τον πίνακα

$$A = \begin{pmatrix} -3 & 6 & -2 & 6 \\ 16 & -49 & 22 & -46 \\ 7 & -30 & 15 & -27 \\ -15 & 42 & -18 & 40 \end{pmatrix}$$

Υπολογίζουμε το χαρακτηριστικό πολυώνυμο

$$\mathrm{Ch}_A(x) = \det(A - x\mathbb{I}_4) = -2 + 3x + x^2 - 3x^3 + x^4$$

Αυτός γενικά είναι ένας δύσκολος και απαιτητικός υπολογισμός αν ο πίνακας A δεν είναι κάποιας ειδικής μορφής.

Στην συνέχεια υπολογίζουμε τις ιδιοτιμές. Αυτό για ένα πίνακα $n \times n$ προϋποθέτει ότι μπορούμε να υπολογίσουμε τις ρίζες ενός πολυωνύμου n-στου βαθμού. Μια τέτοια λύση εν γένει δεν μπορεί να γίνει με ακριβή τρόπο αν το $n \geq 5$ και στην πράξη χρησιμοποιούμε μεθόδους αριθμητικής ανάλυσης για να υπολογίσουμε ρητές προσεγγίσεις των ριζών. Στο συγκεκριμένο παράδειγμα έχουμε ότι

$$\mathrm{Ch}_A(x) = (x-2)(x-1)^2(x+1).$$

Οι ιδιοτιμές είναι οι $2, 1, -1$ και το χαρακτηριστικό πολυώνυμο έχει την ρίζα -1 με πολλαπλότητα 2, οπότε δεν μπορούμε να χρησιμοποιήσουμε το πόρισμα VI.1.13 για να εξασφαλίσουμε ότι ο πίνακας A είναι διαγωνοποιήσιμος. Ο πίνακας A μπορεί να είναι διαγωνοποιήσιμος μπορεί και όχι και πρέπει να το ερευνήσουμε περισσότερο.

Υπολογίζουμε βάσεις για τους ιδιόχωρους των ιδιοτιμών:

Για την ιδιοτιμή 2 Ο ιδιόχωρος E_2 είναι η λύση του ομογενούς συστήματος

$$(A - 2\mathbb{I}_4)x = 0 \text{ ισοδύναμα } \begin{pmatrix} -5 & 6 & -2 & 6 \\ 16 & -51 & 22 & -46 \\ 7 & -30 & 13 & -27 \\ -15 & 42 & -18 & 38 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Το παραπάνω σύστημα το λύνουμε με την μέθοδο της απαλοιφής του Gauss για να βρούμε μια βάση του ιδιόχωρου E_2 η οποία δίνεται από το διάνυσμα $(0, -4, -3, 3)^t$

Για την ιδιοτιμή 1 Ο ιδιόχωρος E_1 είναι η λύση του ομογενούς συστήματος

$$(A - \mathbb{I}_4)x = 0 \text{ ισοδύναμα } \begin{pmatrix} -4 & 6 & -2 & 6 \\ 16 & -50 & 22 & -46 \\ 7 & -30 & 14 & -27 \\ -15 & 42 & -18 & 39 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Και αυτό το σύστημα το λύνουμε με την μέθοδο της απαλοιφής του Gauss για να βρούμε μια βάση του ιδιόχωρου $_1$ η οποία δίνεται από τα διανύσματα $(3, -11, 0, 13)^t, (4, 7, 13, 0)^t$.

Για την ιδιοτιμή -1 Ο ιδιόχωρος E_{-1} είναι η λύση του ομογενούς συστήματος

$$(A + \mathbb{I}_4)x = 0 \text{ ισοδύναμα } \begin{pmatrix} -2 & 6 & -2 & 6 \\ 16 & -48 & 22 & -46 \\ 7 & -30 & 16 & -27 \\ -15 & 42 & -18 & 41 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Το παραπάνω σύστημα το λύνουμε με την μέθοδο της απαλοιφής του Gauss για να βρούμε μια βάση του ιδιόχωρου E_{-1} η οποία δίνεται από το διάνυσμα $(1, -3, -1, 3)^t$.

Σχηματίζουμε τον πίνακα

$$Q = \begin{pmatrix} 0 & 3 & 4 & 1 \\ -4 & -11 & 7 & -3 \\ -3 & 0 & 13 & -1 \\ 3 & 13 & 0 & 3 \end{pmatrix}$$

Η πρώτη στήλη είναι το ιδιοδιάνυσμα της ιδιοτιμής 2, οι δεύτερη και η τρίτη στήλη είναι βάση του ιδιόχωρου της ιδιοτιμής 1 και η τέταρτη στήλη είναι ιδιοδιάνυσμα της ιδιοτιμής -1.

Μπορούμε τώρα να επαληθεύσουμε ότι

$$Q^{-1} = \begin{pmatrix} -1 & 8 & -4 & 7 \\ -\frac{3}{13} & -\frac{15}{13} & \frac{9}{13} & -\frac{11}{13} \\ -\frac{1}{13} & \frac{21}{13} & -\frac{10}{13} & \frac{18}{13} \\ 2 & -3 & 1 & -3 \end{pmatrix}$$

και επαληθεύουμε ότι

$$Q^{-1}AQ = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Άσκηση VI.1.24. Να δειχθεί ότι κάθε άνω τριγωνικός πίνακας $A = (a_{ij})$ έχει χαρακτηριστικό πολυώνυμο το

$$\mathrm{Ch}_A(x) = \prod_{i=1}^{n} (a_{ii} - x).$$

Λύση VI.1.25. Υπολογίζουμε

$$\mathrm{Ch}_A(x) = \det(A - x\mathbb{I}_n) = \begin{pmatrix} a_{11} - x & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & a_{22} - x & \cdots & \cdots & a_{2n} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & a_{nn} - x \end{pmatrix} = \prod_{i=1}^{n} (a_{ii} - x).$$

Άσκηση VI.1.26. Να αποδειχθεί ότι οι πίνακες $A, A^t \in \mathbb{F}^{n,n}$ έχουν το ίδιο χαρακτηριστικό πολυώνυμο.

Λύση VI.1.27. Υπολογίζουμε ότι

$$\mathrm{Ch}_A(x) = \det(A - x\mathbb{I}_n) = \det(A - x\mathbb{I}_n)^t = \det\left(A^t - x\mathbb{I}_n\right) = \mathrm{Ch}_{A^t}(x).$$

Άσκηση VI.1.28. Δίνεται ο πίνακας $A = (a_{ij}) \in \mathbb{F}^{n,n}$ με $a_{ij} = 1$ για κάθε $1 \leq i, j \leq n$. Να βρεθεί αντιστρέψιμος πίνακας Q ώστε $Q^{-1}AQ = \Delta$, όπου Δ διαγώνιος πίνακας.

Λύση VI.1.29. Ξεκινάμε υπολογίζοντας το χαρακτηριστικό πολυώνυμο του A,

$$\begin{aligned} \mathrm{Ch}_A(x) = \det(A - x\mathbb{I}_n) &= \det \begin{pmatrix} 1-x & 1 & \cdots & 1 \\ 1 & 1-x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 1-x \end{pmatrix} = (n-x) \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1-x & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 1-x \end{pmatrix} \\ &= \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & -x & 0 & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -x \end{pmatrix} = (-1)^n (x-n) x^{n-1}. \end{aligned}$$

Στον παραπάνω υπολογισμό αρχικά προσθέσαμε όλες τις στήλες στην πρώτη, ενώ στην στην συνέχεια αφαιρέσαμε την πρώτη γραμμή από τις υπόλοιπες. Παρατηρούμε ότι έχουμε το n ως ιδιοτιμή με αλγεβρική πολλαπλότητα $\mu(n) = 1$ και συνεπώς με γεωμετρική πολλαπλότητα $\rho(n) = 1$, αφού $1 \leq \rho(n) \leq \mu(n) = 1$. Από την άλλη έχουμε την ιδιοτιμή 0 με αλγεβρική πολλαπλότητα $\mu(0) = n - 1$. Η γεωμετρική πολλαπλότητα $\rho(0) = \dim \mathrm{Ker} A = n - 1$ αφού η τάξη του πίνακα A είναι $r(A) = 1$. Άρα ο πίνακας είναι διαγωνοποιήσιμος αφού η αλγεβρική πολλαπλότητα κάθε ιδιοτιμής ταυτίζεται με την γεωμετρική της πολλαπλότητα.

Υπολογίζουμε τώρα τον ιδιόχωρο της ιδιοτιμής n. Θα πρέπει να λύσουμε το ομογενές σύστημα

$$\begin{pmatrix} 1-n & 1 & \cdots & 1 \\ 1 & 1-n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 1-n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

το οποίο έχει μονοδιάστατο χώρο λύσεων που παράγεται από το $(1, \ldots, 1)^t$.

Ο ιδιόχωρος της ιδιοτιμής 0 είναι ίσος με τον πυρήνα της A ο οποίος παράγεται από τα διανύσματα

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, v_{n-1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix},$$

Υπολογίζουμε τον πίνακα

$$Q = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & -1 \end{pmatrix}$$

και έχουμε

$$Q^{-1}AQ = \operatorname{diag}(n, 0, \dots, 0).$$

Άσκηση VI.1.30. Θεωρούμε τον πίνακα

$$A = \begin{pmatrix} a & b & \cdots & b \\ b & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & a \end{pmatrix}$$

Να βρεθεί αντιστρέψιμος πίνακας Q ώστε $Q^{-1}AQ = \Delta$, όπου Δ διαγώνιος πίνακας.

Λύση VI.1.31. Ξεκινάμε υπολογίζοντας το χαρακτηριστικό πολυώνυμο του πίνακα A

$$\begin{aligned}
\mathrm{Ch}_A(x) = \det \begin{pmatrix} a-x & b & \cdots & b \\ b & a-x & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & a-x \end{pmatrix} &= \det \begin{pmatrix} a-x+(n-1)b & b & \cdots & b \\ a-x+(n-1b & a-x & \ddots & \vdots \\ \vdots & & \ddots & b \\ a-x+(n-1)b & \cdots & b & a-x \end{pmatrix} \\
&= (a+(n-1)b-x) \det \begin{pmatrix} 1 & b & \cdots & b \\ 1 & a-x & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ 1 & \cdots & b & a-x \end{pmatrix} \\
&= (a+(n-1)b-x) \det \begin{pmatrix} 1 & b & \cdots & \cdots & b \\ 0 & a-x-b & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & a-x-b \end{pmatrix} \\
&= (a+(n-1)b-x)(a-b-x)^{n-1}.
\end{aligned}$$

Η ιδιοτιμή $(n-1)b + a$ έχει αλγεβρική πολλαπλότητα και γεωμετρική πολλαπλότητα ίση με 1. Υπολογίζουμε τον ιδιόχωρο $\mathrm{Ker}(A - (a+(n-1)b)\mathbb{I}_n)$ ως λύση του ομογενούς συστήματος,

$$\begin{pmatrix} -(n-1)b & b & \cdots & b \\ b & -(n-1)b & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & -(n-1)b \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

ο οποίος έχει μονοδιάστατη λύση $(1,\dots,1)^t$.

Στην συνέχεια υπολογίζουμε τον ιδιόχωρο $\mathrm{Ker}(A-(b-a)\mathbb{I}_n)$ της ιδιοτιμής $a-b$ ως λύση του ομογενούς συστήματος

$$\begin{pmatrix} b & b & \cdots & b \\ b & b & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & b \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

το οποίο έχει λύσεις:

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, v_{n-1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix},$$

Υπολογίζουμε τον πίνακα

$$Q = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & -1 \end{pmatrix}$$

και έχουμε

$$Q^{-1}AQ = \mathrm{diag}(n,0,\dots,0).$$

Άσκηση VI.1.32. Να υπολογιστεί το χαρακτηριστικό πολυώνυμο του πίνακα

$$A = \begin{pmatrix} 0 & a & \cdots & a \\ b & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a \\ b & \cdots & b & 0 \end{pmatrix}.$$

Λύση VI.1.33. Αν $a=b$ το χαρακτηριστικό πολυώνυμο έχει υπολογιστεί στην άσκηση VI.1.30 και είναι ίσο με $((n-1)b-x)(-b-x)^{n-1}$. Υποθέτουμε ότι $a \neq b$ και υπολογίζουμε

$$\mathrm{Ch}_A(x) = \det\begin{pmatrix} -x & a & \cdots & a \\ b & -x & \ddots & \vdots \\ \vdots & \ddots & \ddots & a \\ b & \cdots & b & -x \end{pmatrix} = \frac{a(-x-b)^n - b(-x-a)^n}{a-b}$$

σύμφωνα με την άσκηση III.8.7.

Άσκηση VI.1.34. Δίνονται οι πίνακες $A \in \mathbb{F}^{m,n}$, $B \in \mathbb{F}^{n,m}$. Τι σχέση έχουν τα χαρακτηριστικά πολυώνυμα των πίνακων AB και BA;

Λύση VI.1.35. Θεωρούμε τους μπλοκ πίνακες

$$\mathbb{F}^{n+m,n+m} \ni C = \begin{pmatrix} x\mathbb{I}_m & A \\ B & \mathbb{I}_n \end{pmatrix} \quad D = \begin{pmatrix} \mathbb{I}_m & \mathbf{0}_{m,n} \\ -B & x\mathbb{I}_n \end{pmatrix} \in \mathbb{F}^{n+m,n+m}.$$

και υπολογίζουμε ότι

$$CD = \begin{pmatrix} x\mathbb{I}_m - AB & xA \\ \mathbf{0}_{n,m} & x\mathbb{I}_n \end{pmatrix} \qquad DC = \begin{pmatrix} x\mathbb{I}_m & A \\ \mathbf{0}_{n,m} & -BA + x\mathbb{I}_n \end{pmatrix}$$

Κάνοντας χρήση του $\det(CD) = \det(CD)$ και της άσκησης III.8.5 έχουμε ότι

$$x^n(-1)^n\mathrm{Ch}_{AB}(x) = x^n\det(\mathbb{I}_m - AB) = x^m\det(\mathbb{I}_m - BA) = x^m(-1)^m\mathrm{Ch}_{AB}(x).$$

Άσκηση VI.1.36. Στην άσκηση VI.1.34 αν $n = m$ να δειχθεί ότι οι πίνακες AB και BA έχουν το ίδιο χαρακτηριστικό πολυώνυμο.

Λύση VI.1.37. Το ζητούμενο προκύπτει άμεσα από την άσκηση VI.1.34.

Εναλλακτικά, θα μπορούσαμε να επιχειρηματολογήσουμε ως εξής: Αν A είναι αντιστρέψιμος τότε

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

δηλαδή οι πίνακες AB και BA είναι όμοιοι και έχουν το ίδιο χαρακτηριστικό πολυώνυμο. Αν ο A δεν είναι αντιστρέψιμος τότε υπάρχει μια ακολουθία αντιστρέψιμων πινάκων A_n ώστε $A_n \to A$, δείτε το θεώρημα III.7.3. Ισχύει ότι

$$\begin{aligned}
\mathrm{Ch}_{AB}(x) &= \det\left(\lim_{n\to\infty} A_n B - x\mathbb{I}_n\right) = \lim_{n\to\infty} \det(A_n B - x\mathbb{I}_n) = \lim_{n\to\infty} \det(BA_n - x\mathbb{I}_n) \\
&= \det\left(B \lim_{n\to\infty} A_n - x\mathbb{I}_n\right) = \det(BA - x\mathbb{I}_n) = \mathrm{Ch}_{BA}(x).
\end{aligned}$$

Άσκηση VI.1.38. Να δείξετε ότι η γραμμική συνάρτηση $T : V \to V$, όπου V είναι ο διανυσματικός χώρος των συνεχών συναρτήσεων $f : \mathbb{R} \to \mathbb{R}$ και δίνεται από $Tf(x) = \int_0^x f(t)dt$ δεν έχει ιδιοτιμές.

Λύση VI.1.39. Έστω λ μια ιδιοτιμή του T, οπότε θα πρέπει να υπάρχει μη-μηδενική συνάρτηση f με

$$\int_0^x f(t)dt = \lambda f(t). \tag{VI.6}$$

Η παραπάνω ισότητα εξασφαλίζει ότι το ιδιοδιάνυσμα f είναι παραγωγίσιμο οπότε έχουμε

$$f(x) = \lambda f'(x) \to f(x) = ce^{\lambda x}.$$

Όμως από την (VI.6) έχουμε ότι $c = 0$.

Άσκηση VI.1.40. Θεωρούμε την γραμμική συνάρτηση $D : V \to V$, όπου V είναι ο διανυσματικός χώρος των διαφορίσιμων συναρτήσεων $f : \mathbb{R} \to \mathbb{R}$ και $Df = f'$. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα της f.

Λύση VI.1.41. Ψάχνουμε για συναρτήσεις τέτοιες ώστε $f'(x) = \lambda f(x)$. Αυτές είναι οι συναρτήσεις της μορφής $ce^{\lambda x}$. Δηλαδή κάθε πραγματικός αριθμός λ είναι ιδιοτιμή της D με μονοδιάστατο ιδιόχωρο $\langle e^{\lambda x}\rangle$.

Άσκηση VI.1.42. Να αποδειχθεί ότι $(2+\sqrt{3})^n + (2-\sqrt{3})^n \in \mathbb{Z}$ για κάθε $n \in \mathbb{N}$.

Λύση VI.1.43. Παρατηρούμε ότι ο πίνακας

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

έχει ιδιοτιμές $2 \pm \sqrt{3}$ και συνεπώς είναι διαγωνοποιήσιμος. Η ζητούμενη έκφραση είναι το $\mathrm{tr}(A^n)$ υπολογισμένη είτε ως δύναμη του πίνακα A συνεπώς είναι μια ακέραια τιμή είτε ως δύναμη του διαγώνιου πίνακα.

VI.2 Το θεώρημα Caley-Hamilton

Ορισμός VI.2.1. Έστω $f(x) \in \mathbb{F}[x]$,

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

για ένα τετραγωνικό πίνακα $A \in \mathbb{F}^{m,m}$ μπορούμε να θεωρήσουμε τον πίνακα

$$f(A) = a_0\mathbb{I}_m + a_1A + a_2A^2 + \cdots + a_nA^n.$$

Παρατηρούμε ότι ο διανυσματικός χώρος $\mathbb{F}^{m,m}$ έχει διάσταση m^2, οπότε οποιαδήποτε m^2+1 στοιχεία του είναι γραμμικά εξαρτημένα. Ειδικότερα, τα m^2+1 στοιχεία

$$\mathbb{I}_m, A, A^2, A^3, \ldots, A^{m^2}$$

είναι γραμμικά εξαρτημένα, δηλαδή υπάρχουν στοιχεία $\lambda_0, \lambda_1, \ldots, \lambda_{m^2+1} \in \mathbb{F}$ όχι όλα 0 ώστε

$$\lambda_0\mathbb{I}_m + \lambda_1 A + \ldots + \lambda_{m^2}A^{m^2} = \mathbf{0}_{\mathbb{F}^{m,m}}.$$

Άρα για το μη-μηδενικό πολυώνυμο $f(x) = \sum_{\nu=0}^{m^2} \lambda_\nu x^\nu$ έχουμε ότι $f(A) = \mathbf{0}_{\mathbb{F}^{m,m}}$. Θεωρούμε το σύνολο των πολυωνύμων τα οποία μηδενίζουν τον πίνακα A.

$$I_A = \{f \in \mathbb{F}[x] : f(A) = \mathbf{0}_{\mathbb{F}^{m,m}}\}$$

Πρόταση VI.2.2. *Το σύνολο I_A είναι ένα ιδεώδες του δακτυλίου $\mathbb{F}[x]$, δηλαδή*

1. *Περιέχει μη-μηδενικά πολυώνυμα,*
2. *Αν $f, g \in I_A$, τότε $f + g \in I_A$*
3. *Αν $f \in I_A$, $g \in \mathbb{F}[x]$, τότε $fg \in I_A$.*

Απόδειξη. Το ότι το σύνολο I_A περιέχει και μη-μηδενικά πολυώνυμα το έχουμε ήδη δείξει. Είναι σαφές ότι αν $f, g \in I_A$ τότε $f(A) = g(A) = \mathbf{0}_{\mathbb{F}^{m,m}}$ οπότε και $(f+g)(A) = f(A) + g(A) = \mathbf{0}_{\mathbb{F}^{m,m}}$. Ομοίως αν $f \in I_A$ και $g \in \mathbb{F}[x]$, τότε $(fg)(A) = f(A)g(A) = \mathbf{0}_{\mathbb{F}^{m,m}} g(A) = \mathbf{0}_{\mathbb{F}^{m,m}}$. □

Πόρισμα VI.2.3. *Υπάρχει ένα μονικό πολυώνυμο $p_A(x)$, το οποίο διαιρεί κάθε πολυώνυμο που μηδενίζει τον A. Το πολυώνυμο αυτό θα το λέμε το ελάχιστο πολυώνυμο του πίνακα A.*

Απόδειξη. Γνωρίζουμε ότι τα ιδεώδη του $\mathbb{F}[x]$ είναι κύρια (δείτε το θεώρημα I.9.3, οπότε υπάρχει πολυώνυμο $p_A(x)$, ώστε $p_A(x)\mathbb{F}[x] = I_A$. □

Παρατήρηση VI.2.4. Όμοιοι πίνακες έχουν το ίδιο ελάχιστο πολυώνυμο. Πράγματι, αφού $QA^kQ^{-1} = (QAQ^{-1})^k$ παρατηρούμε ότι $f(QAQ^{-1}) = Qf(A)Q^{-1}$, συνεπώς $I_A = I_{QAQ^{-1}}$.

Παραδείγματα VI.2.5.

- Θεωρούμε ένα διαγωνοποιήσιμο πίνακα, ο οποίος εξ ορισμού είναι όμοιος με ένα διαγώνιο. Οπότε σύμφωνα με την παρατήρηση VI.2.4 αρκεί να υπολογίσουμε το ελάχιστο πολυώνυμο του $\mathrm{diag}(\lambda_1, \ldots, \lambda_s)$. Παρατηρούμε ότι για κάθε πολυώνυμο $f(x) \in \mathbb{F}[x]$ έχουμε

  $$f(\mathrm{diag}(\lambda_1, \ldots, \lambda_n)) = \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_r)),$$

  οπότε το ελάχιστο πολυώνυμο είναι το

  $$p_{\mathrm{diag}(\lambda_1,\ldots,\lambda_n)} = \prod_{\nu=1}^{s} (x - \lambda_{i_\nu}),$$

  όπου $\{\lambda_{i_1}, \ldots, \lambda_{i_s}\}$ είναι το σύνολο των διαφορετικών $\lambda_1, \ldots, \lambda_n$.

- Ένας πίνακας θα λέγεται μηδενοδύναμος βαθμού k αν το ελάχιστο πολυώνυμο του είναι το x^k για κάποιο φυσικό αριθμό $k \geq 1$. Για παράδειγμα ο $n \times n$ πίνακας

$$N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

έχει ελάχιστο πολυώνυμο x^n. Πράγματι μπορούμε να γράψουμε τον πίνακα $N = (n_{ij})$ με $n_{ij} = \delta_{i,j+1}$. Για κάθε $0 < r \leq n$ συμβολίζουμε τον πίνακα $N^r = (n_{ij}^{(r)})$ ο οποίος σύμφωνα με την άσκηση II.3.3 είναι

$$n_{ij}^{(r)} = \sum_{\nu_1=1}^{\ell_1} \sum_{\nu_2=1}^{\ell_2} \cdots \sum_{\nu_r=1}^{\ell_r} n_{i,\ell_1} n_{\ell_1,\ell_2} \cdots n_{\ell_{r-1},j},$$

το οποίο είναι μη μηδενικό αν και μόνο αν $\ell_1 = i+1, \ell_2 = \ell_1 + 1 = i+2, \ldots, j = \ell_{r-1} + 1 = i + r$. Συνεπώς $n_{ij}^{(r)} = \delta_{i,i+r}$. Άρα για $r = n-1$ έχουμε ότι $n_{1,n-1}^{(n-1)} = 1$, οπότε $N^{n-1} \neq \mathbf{0}_{\mathbb{F}^{n,n}}$. Ενώ $N^n = \mathbf{0}_{\mathbb{F}^{n,n}}$. Μπορούμε να σκεφτόμαστε τον πίνακα N ως ένα πίνακα με μονάδες ακριβώς στην διαγώνιο που είναι μια θέση πάνω από την κύρια διαγώνιο, ενώ σε κάθε δύναμη N^k η μη μηδενική διαγώνιος με τις μονάδες μετακινείται μια θέση προς τα πάνω.

Θεώρημα VI.2.6 (Caley-Hamilton). *Το χαρακτηριστικό πολυώνυμο* $\mathrm{Ch}_A(x)$ *ενός* $n \times n$ *πίνακα μηδενίζει τον πίνακα* A*, δηλαδή* $\mathrm{Ch}_A(A) = \mathbf{0}_{\mathbb{F}^{n,n}}$.

Απόδειξη. Υποθέτουμε ότι

$$\mathrm{Ch}_A(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Παρατηρούμε ότι ισχύει

$$(A - x\mathbb{I}_n)\mathrm{adj}(A - x\mathbb{I}_n) = \det(A - x\mathbb{I}_n)\mathbb{I}_n.$$

Ο πίνακας $\mathrm{adj}(A - x\mathbb{I}_n)$ έχει ως στοιχεία πολυώνυμα βαθμού το πολύ $n-1$ συνεπώς μπορεί να γραφεί ως:

$$\mathrm{adj}(A - x\mathbb{I}_n) = B_0 + B_1 x + \cdots + B_{n-1}x^{n-1}, 0, 1, \ldots, B_{n-1} \in \mathbb{F}^{n,n}.$$

Συνεπώς έχουμε

$$(A - x\mathbb{I}_n)(B_0 + B_1 x + \cdots + B_{n-1}x^{n-1}) = (x^n + a_{n-1}x^{n-1} + \cdots + a_0)\mathbb{I}_n.$$

Συγκρίνοντας τους συντελεστές έχουμε

$$\begin{aligned} AB_0 &= a_0\mathbb{I}_n \\ -B_0 + AB_1 &= a_1\mathbb{I}_n \\ &\cdots \\ -B_{n-2} + AB_{n-1} &= a_{n-1}\mathbb{I}_n \\ -B_{n-1} &= \mathbb{I}_n \end{aligned}$$

Πολλαπλασιάζουμε τις παραπάνω σχέσεις με $\mathbb{I}_n, A, \ldots, A^n$ και προσθέτουμε κατά μέλη για να πάρουμε

$$0 = a_0\mathbb{I}_n + a_1 A + \cdots + a_{n-1}A^{n-1} + A^n,$$

δηλαδή το ζητούμενο. □

Πόρισμα VI.2.7. *Το ελάχιστο πολυώνυμο* $p_A(x)$ *ενός πίνακα* A *διαιρεί το χαρακτηριστικό πολυώνυμο* $\mathrm{Ch}_A(x)$.

Πόρισμα VI.2.8. *Κάθε αντιστρέψιμος πίνακας μπορεί να γραφεί ως μια πολυωνυμική έκραση βαθμού* $n-1$ *στον* A.

Απόδειξη. Έχουμε ότι

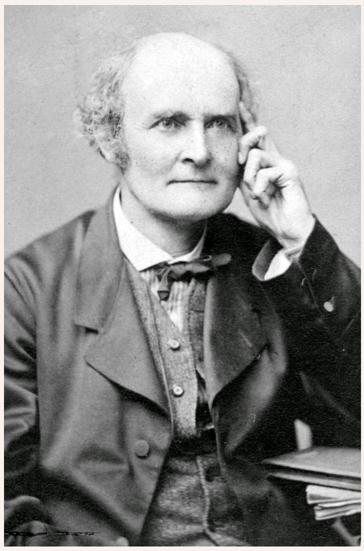
$$\mathrm{Ch}_A(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + (-1)^n \det(A)\mathbb{I}_n = 0$$

άρα

$$A^{-1} = \frac{(-1)^{n+1}}{\det(A)}\left(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1\mathbb{I}_n\right).$$

□

Ο Arthur Cayley (16 Αυγούστου 1821 - 26 Ιανουαρίου 1895) ήταν Άγγλος πολυγραφότατος Μαθηματικός. Σπούδασε στο King's College, και στο Trinity College στο Πανεπιστήμιο του Cambridge. Εργάστηκε ως δικηγόρος για 14 έτη. Σε ηλικία 42 ετών έγινε καθηγητής στο Πανεπιστήμιο του Cambridge. Εργάστηκε κυρίως στην Άλγεβρα και ήταν ένας από τους πρώτους Μαθηματικούς που όρισαν την έννοια της ομάδας με τον σύγχρονο τρόπο, ως ένα σύνολο εφοδιασμένο με μια πράξη που ικανοποιεί μια σειρά από ιδιότητες. Οι Μαθηματικοί της εποχής του όταν αναφέρονταν στην έννοια της ομάδας εννοούσαν ομάδες μεταθέσεων. Διάφορα θεωρήματα στην θεωρία ομάδων φέρουν το όνομα του όπως για παράδειγμα το θεώρημα ότι κάθε πεπερασμένη ομάδα είναι υποομάδα μιας ομάδας μεταθέσεων. Ασχολήθηκε και με την Αλγεβρική γεωμετρία, την θεωρία των επιφανειών, τις πολλαπλότητες Chow, τις ruled surfaces αλλά και με την θεωρία των ελλειπτικών, αβελιανών και θήτα συναρτήσεων.



Ο William Rowan Hamilton (3/4 Αυγούστου 1805 – 2 Σεπτεμβρίου 1865) ήταν ένας Ιρλανδός Μαθηματικός, Αστρονόμος και Φυσικός. Έδειξε το ταλέντο του από μικρή ηλικία τόσο στα Μαθηματικά όσο και στην εκμάθηση ξένων γλωσσών. Σπούδασε στο κολέγιο Trinity του Δουβλίνου. Στην ηλικία των 21 ετών και ενώ ήταν ακόμα προπτυχιακός φοιτητής του δόθηκε θέση καθηγητή Αστρονομίας στο Trinity college που φοιτούσε. Εργάστηκε σε θέματα Άλγεβρας και Γραμμικής Άλγεβρας, εισήγαγε την έννοια των quaterions και επίσης μελέτησε την πεμπτοβάθμια εξίσωση και τα νέα αποτελέσματα του Niels Henrik Abel. Επίσης είχε μέγιστη συνεισφορά στην οπτική και στην κλασική μηχανική εισάγοντας την έννοια της Χαμιλτονιανής μια θεμελιώδη έννοια της θεωρητικής φυσικής με εφαρμογές στον ηλεκτρομαγνητισμό και στην κβαντομηχανική.

VI.2.1 Εφαρμογές

Πρόταση VI.2.9. *Για ένα πίνακα A θεωρούμε το χαρακτηριστικό του πολυώνυμο*

$$\mathrm{Ch}_A(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 A + a_0$$

Παρατηρούμε ότι αν $z_1, \dots, z_n$ είναι οι ιδιοτιμές του A τότε οι εκφράσεις

$$z_1^k + \cdots + z_n^k = \mathrm{tr}(A^k).$$

VI.2.2 Συνοδός πίνακας

Ορισμός VI.2.10. Για ένα πολυώνυμο

$$f(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + x^n \in \mathbb{F}[x]$$

ο παρακάτω $n \times n$ πίνακας

$$C(f) := \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

θα λέγεται συνοδός πίνακας (companion matrix) του $f(x)$.

Πρόταση VI.2.11. *Το χαρακτηριστικό πολυώνυμο του $C(f)$ είναι ίσο με $(-1)^n f(x)$ ενώ το ελάχιστο πολυώνυμο του $C(f)$ είναι ίσο με το $f(x)$.*

Απόδειξη. Το χαρακτηριστικό πολυώνυμο είναι ίσο με

$$\mathrm{Ch}_{C(f)}(x) = \det \begin{pmatrix} -x & 0 & \cdots & 0 & -a_0 \\ 1 & -x & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & -x & \vdots \\ 0 & \cdots & 0 & 1 & -(x+a_{n-1}) \end{pmatrix}$$

Αν πολλαπλασιάσουμε την τελευταία γραμμή με x και την προσθέσουμε στην προτελευταία στην συνέχεια την προτελευταία στην προηγούμενη της και συνεχίσουμε με τον ίδιο τρόπο μέχρι την δεύτερη γραμμή η ορίζουσα δεν αλλάζει και καταλήγουμε στην ορίζουσα

$$\mathrm{Ch}_{C(f)}(x) = \det \begin{pmatrix} 0 & 0 & \cdots & 0 & f(x) \\ 1 & 0 & \cdots & 0 & * \\ 0 & 1 & \ddots & \vdots & * \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & -(x+a_{n-1}) \end{pmatrix}$$

όπου στην πάνω δεξιά γωνία έχουμε πάρει το

$$f(x) = -(a_0 + x(a_1 + x(\cdots(a_{n-2} + x(a_{n-1} + x)))))$$

μία γραφή του πολυωνύμου x γνωστή και ως σχήμα Horner. Το αποτέλεσμα προκύπτει αναπτύσοντας κατά Laplace την ορίζουσα κατά μήκος της πρώτης γράμμης.

Για το ελάχιστο πολυώνυμο τώρα: Παρατηρούμε ότι για $x = (x_1, x_2, \ldots, x_n)^t$

$$C(f)x = x_n \begin{pmatrix} -a_0 \\ -a_1 \\ \vdots \\ -a_{n-1} \end{pmatrix} + \begin{pmatrix} 0 \\ x_1 \\ \vdots \\ x_{n-2} \end{pmatrix}.$$

Συνεπώς, για $v = (1, 0, \ldots, 0)^t$, έχουμε

$$C(f)v = (0,1,\ldots,0)^t, C(f)^2 v = (0,0,1,\ldots,0)^t, \ldots, C(f)^{n-1} v = (0,\ldots,0,1)^t.$$

Άρα για μια επιλογή $b_0, b_1, \ldots, b_{n-1}$

$$\sum_{j=0}^{n-1} b_j C(f)^j v = (b_0, b_1, \ldots, b_{n-1})^t.$$

Το τελευταίο όμως σημαίνει ότι δεν υπάρχει μη-μηδενικό πολυώνυμο $g(x) = \sum_{j=0}^{n-1} b_j x^j$, ώστε $g(C(f)) = 0$, άρα το ελάχιστο πολυώνυμο έχει βαθμό n και ταυτίζεται με το $f(x)$. □

Παρατήρηση VI.2.12. Θα δώσουμε άλλη μια κομψή απόδειξη της πρότασης VI.2.11 η οποία είναι λιγότερο στοιχειώδης από αυτή που ήδη γράψαμε.

Θεωρούμε τον διανυσματικό χώρο $V = \mathbb{F}[x]/\langle f \rangle$. Η συνάρτηση

$$\begin{aligned} \phi : F[x]/\langle f \rangle &\longrightarrow F[x]/\langle f \rangle \\ g + \langle f \rangle &\longmapsto x \cdot g + \langle f \rangle \end{aligned}$$

είναι καλά ορισμένη και γραμμική. Ο πίνακας της γραμμικής συνάρτησης της ϕ ως προς την βάση $\{1, x, x^2, \ldots, x^{n-1}\}$ είναι ο συνοδός πίνακας του πολυωνύμου $f(x)$. Έστω

$$p_\phi(x) = \sum_{i=1}^{d} b_i x^i \in \mathbb{F}[x]$$

το ελάχιστο πολυώνυμο της ϕ. Και $\mathrm{Ch}_\phi(x)$ το χαρακτηριστικό. Τότε το p_ϕ είναι η μηδενική συνάρτηση $V \to V$, δηλαδή

$$\begin{aligned} 0 + \langle f\rangle = p_\phi(\phi)(1+\langle\phi\rangle) &= \sum_{i=0}^{d} b_i \phi^i (1+\langle f\rangle) \\ &= \left(\sum_{i=0}^{d} a_i x^i\right) + \langle f\rangle = p_\phi + \langle f\rangle \end{aligned}$$

Συνεπώς $f \mid p_\phi \mid \mathrm{Ch}_\phi$. Αφού όλα τα πολυώνυμα είναι μονικά βαθμού n η ισότητα $f(x) = p_f(x) = \mathrm{Ch}_\phi(x)$ έπεται.

Παρατήρηση VI.2.13. Παρατηρούμε ότι για μια ρίζα λ του $f(x)$ έχουμε

$$(1,\lambda,\lambda^2,\dots,\lambda^{n-1})\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} = \lambda(1,\lambda,\lambda^2,\dots,\lambda^{n-1})$$

Δηλαδή, το $(1,\lambda,\lambda^2,\dots,\lambda^{n-1})^t$ είναι ένα ιδιοδιάνυσμα του $C(f)^t$. Γνωρίζουμε ότι αν το $f(x)$ έχει διαφορετικές ρίζες τότε ο πίνακας $C(f)$ άρα και ο $C(f)^t$ είναι διαγωνοποιήσιμοι. Παρατηρούμε ότι σε αυτή την περίπτωση, αν $\lambda_1,\dots,\lambda_n$ είναι οι διαφορετικές ρίζες του $f(x)$ τότε οι στήλες του πίνακα

$$Q = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}$$

αποτελούν μία βάση από ιδιοδιανύσματα του $C(f)^t$. Παρατηρήστε ότι

$$\det Q = \prod_{1\le i<j\le n} (\lambda_i - \lambda_j),$$

παράγραφος III.8 συμβατό με το ότι ο Q είναι αντιστρέψιμος.

Εφαρμογή 1. *θα δώσουμε τώρα μια απόδειξη των ταυτοτήτων του Newton (πρόταση I.9.32) με την βοήθεια της γραμμικής άλγεβρας. Ακολουθούμε τον Dan Kalman A Matrix proof of Newton Identities*

Θεωρούμε το πολυώνυμο

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

με ρίζες $z_1,\dots,z_n$. Σύμφωνα με την πρόταση I.9.32 οι συντελεστές του πολυωνύμου a_i δίνονται από

$$a_i = (-1)^{n-i} e_i(z_1,\dots,z_n)$$

Οπότε αν $p_i(z_1,\dots,z_n) = z_1^i + \cdots + z_n^i$, τότε οι ταυτότητες Newton γράφονται ως

$$\begin{aligned} p_k + a_{n-1}p_{k-1} + \cdots + a_0 p_{k-n} &= 0 \text{ αν } k > n \\ p_k + a_{n-1}p_{k-1} + \cdots + a_{n-k+1}p_1 &= -k a_{n-k} \text{ αν } 1 \le k \le n. \end{aligned}$$

θεωρούμε τώρα τον $n \times n$ *συνοδό πίνακα* C *του* $f(x)$. *Αυτός έχει ιδιοτιμές τις ρίζες* $z_1, \dots, z_n$ *του* $f(x)$ *ενώ τα αθροίσματα*

$$p_k(z_1, \dots, z_n) = \mathrm{tr}(C^k).$$

Συνεπώς, για $k > n$ *η πρώτη ταυτότητα Newton γράφεται*

$$\mathrm{tr}(C^k) + a_{n-1}\mathrm{tr}(C^{k-1}) + \cdots + a_0\mathrm{tr}(C^{k-n}) = 0$$

και επειδή η συνάρτηση ίχνους είναι γραμμική έχουμε ότι η παραπάνω εξίσωση είναι ισοδύναμη με την

$$\mathrm{tr}\left(C^k + a_{n-1}C^{k-1}) + \cdots + a_0C^{k-n}\right) = 0 \Leftrightarrow \mathrm{tr}(C^{k-n}f(C)) = 0$$

η οποία είναι άμεση συνέπεια του θεωρήματος Caley-Hamilton.

Για την περίτωση $1 \le k \le n$ *η ταυτότητα Newton γράφεται ως*

$$\mathrm{tr}(C^k) + a_{n-1}\mathrm{tr}(C^{k-1}) + \cdots + a_{n-k+1}\mathrm{tr}(C) = -ka_{n-k}$$

την οποία την γράφουμε στην ισοδύναμη μορφή:

$$\mathrm{tr}\left(C^k + a_{n-1}C^{k-1} + \cdots + a_{n-k+1}C + a_{n-k}\mathbb{I}_n\right) = (n-k)a_{n-k}. \tag{VI.7}$$

Γράφουμε $X = x\mathbb{I}_n$ *και υπολογίζουμε ότι*

$$\begin{aligned} f(X) = (X - C)\big(&X^{n-1} + (C + a_{n-1}\mathbb{I}_n)X^{n-2} + (C^2 + a_{n-1}C + a_{n-2}\mathbb{I}_n)X^{n-3} \\ &+ \cdots + (C^{n-1} + a_{n-1}C^{n-2} + \cdots + a_1\mathbb{I}_n)\mathbb{I}_n\big). \end{aligned}$$

Θα θέλαμε να χρησιμοποιήσουμε την συνάρτηση ίχνους, η οποία όμως δεν συμπεριφέρεται με κάποιο εύκολο τρόπο στα γινόμενα. Παρατηρούμε ότι αν το x *δεν είναι ιδιοτιμή του* C *ο πίνακας* $X - C$ *είναι αντιστρέψιμος οπότε με αυτή την παραδοχή γράφουμε*

$$\begin{aligned} (X - C)^{-1}f(X) = X^{n-1} &+ (C + a_{n-1}\mathbb{I}_n)X^{n-2} + (C^2 + a_{n-1}C + a_{n-2}\mathbb{I}_n)X^{n-3} \\ &+ \cdots + (C^{n-1} + a_{n-1}C^{n-2} + \cdots + a_1\mathbb{I}_n)\mathbb{I}_n. \end{aligned} \tag{VI.8}$$

Στην συνέχεια εφαρμόζουμε τον τύπο του ίχνους για να πάρουμε

$$\begin{aligned} \mathrm{tr}\left((X - C)^{-1}f(X)\right) = nx^{n-1} &+ \mathrm{tr}(C + a_{n-1}\mathbb{I}_n)x^{n-2} + \mathrm{tr}(C^2 + a_{n-1}C + a_{n-2}\mathbb{I}_n)x^{n-3} \\ &+ \cdots + \mathrm{tr}(C^{n-1} + a_{n-1}C^{n-2} + \cdots + a_1\mathbb{I}_n). \end{aligned}$$

Θα δείξουμε ότι το αριστερό μέρος της παραπάνω εξίσωσης είναι το $f'(x)$. *Με αυτή την παραδοχή, εξισώνοντας τους συντελεστές από το αριστερό μέρος έχουμε το ζητούμενο. Πράγματι, εξισώνοντας τον συντελεστή του* x^{n-k-1} *στο* $p'(x)$ *με τον συντελεστή του αριστερού μέρους της εξίσωσης (VI.12) έχουμε*

$$(n-k)a_{n-k} = \mathrm{tr}\left(C^k + a_{n-1}C^{k-1} + \cdots + a_{n-k+1}C + a_{n-k}\mathbb{I}_n\right)$$

η οποία είναι ακριβώς η (VI.7).

θα υπολογίσουμε τώρα το $A = (X - C)^{-1}f(X)$. *Παρατηρούμε ότι* $f(X) = f(x\mathbb{I}_n) = f(x)\mathbb{I}_n$, *συνεπώς* $A = f(x)(x\mathbb{I}_n) - C)^{-1}$. *Συνεπώς*

$$\mathrm{tr}(A) = f(x)\mathrm{tr}(x\mathbb{I}_n - C)^{-1}.$$

Είναι σαφές ότι το ίχνος ενός πίνακα είναι το άθροισμα των ιδιοτιμών του και οι ιδιοτιμές του $(x\mathbb{I}_n - C)^{-1}$ *είναι οι* $1/(x-z_1), 1/(x-z_2), \ldots, 1/(x-z_n)$*. Συνεπώς*

$$\operatorname{tr}(A) = f(x)\left(\frac{1}{x-z_1} + \frac{1}{x-z_2} + \cdots + \frac{1}{x-z_n}\right)$$

το οποίο είναι η παράγωγος του $f(x) = (x-z_1)(x-z_2)\cdots(x-z_n)$.

VI.3 Κανονική μορφή Jordan

VI.3.1 Ανάλυση σε άθροισμα αναλλοίωτων υπόχωρων

Παρατηρούμε ότι αν ένας διανυσματικός χώρος V είναι ευθύ άθροισμα L-αναλλοίωτων υπόχωρων

$$V = W_1 \bigoplus W_2 \bigoplus \cdots \bigoplus W_s,$$

δηλαδή $L(W_i) \subset W_i$ για κάθε $1 \le i \le s$ τότε μπορούμε να διαλέξουμε βάσεις B_i των χώρων V_i για $1 \le i \le s$ ώστε

$$(L, B, B) = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_s \end{pmatrix} \tag{VI.9}$$

όπου $A_i = (L|_{W_i}, B_i, B_i)$. Μια τέτοια ανάλυση της γραμμικής συνάρτησης L θα την λέμε ανάλυση κατά μπλόκ της L. Παρατηρούμε ότι μια γραμμική συνάρτηση L (ή ο αντίστοιχος πίνακας) είναι διαγωνοποίσιμη αν και μόνο αν υπάρχει μια ανάλυση του χώρου V ως ευθύ άθροισμα μονοδιάστατων L-αναλλοίωτων υπόχωρων.

Πρόταση VI.3.1. *Το χαρακτηριστικό πολυώνυμο μιας γραμμικής συνάρτησης η οποία δέχεται μια κατά μπλοκ διαγώνια ανάλυση όπως στην εξίσωση (VI.9) είναι το γινόμενο των χαρακτηριστικών πολυωνύμων των περιορισμών δηλαδή*

$$\operatorname{Ch}_L(x) = \prod_{i=1}^{s} \operatorname{Ch}_{L|_{W_i}}(x).$$

Απόδειξη. Το χαρακτηριστικό πολυώνυμο της L είναι το ίδιο με το το χαρακτηριστικό πολυώνυμο του πίνακα της L ως προς οποιαδήποτε βάση B. Υπολογίζουμε ότι

$$\operatorname{Ch}_L(x) = \det\begin{pmatrix} A_1 - x\mathbb{I}_{\dim(W_1)} & 0 & \cdots & 0 \\ 0 & A_2 - x\mathbb{I}_{\dim(W_2)} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & A_s - x\mathbb{I}_{\dim(W_s)} \end{pmatrix}$$

και το ζητούμενο προκύπτει από την άσκηση III.8.5. □

Παρατήρηση VI.3.2. Το ελάχιστο πολυώνυμο της L που έχει μια κατά μπλοκ διαγώνια ανάλυση όπως στην εξίσωση (VI.9) δεν είναι κατ ανάγκη το γινόμενο των ελαχίστων πολυώνυμων των $L|_{W_i}$. Πράγματι $m_{\operatorname{diag}(\lambda,\ldots,\lambda)}(x) = x - \lambda$ και όχι $(x-\lambda)^n$.

Λήμμα VI.3.3. *Έστω πίνακας*

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_s \end{pmatrix}.$$

Το ελάχιστο πολυώνυμο του A *είναι το ελάχιστο κοινό πολαπλάσιο των ελαχίστων πολυωνύμων των πινάκων* A_i.

Απόδειξη. Παρατηρούμε ότι για ένα πολυώνυμο $f(x) \in \mathbb{F}[x]$ έχουμε

$$f(A) = \begin{pmatrix} f(A_1) & 0 & \cdots & 0 \\ 0 & f(A_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & f(A_s) \end{pmatrix}.$$

Άρα το $f(A) = \mathbf{0}$ αν και μόνο αν $f(A_i) = 0$ για κάθε $1 \le i \le s$. Συνεπώς, έχουμε ότι $m_{A_i}(x) \mid m_A(x)$ για κάθε $1 \le i \le s$ και κάθε κοινό πολλαπλάσιο των $m_{A_i}(x)$ μηδενίζει τον A. Συνεπώς

$$m_A(x) = \text{Ε.Κ.Π}(m_{A_1}(x), \dots, m_{A_s}(x)).$$

□

Λήμμα VI.3.4. *Αν* $L_i : V \to V$, $i = 1, 2$ *είναι γραμμικές συναρτήσεις και* $L_1L_2 = L_2L_1$ *τότε ο* $\mathrm{Ker}L_1$ *και ο* $\mathrm{Im}L_1$ *είναι* L_2*-αναλλοίωτοι.*

Απόδειξη. Αν $v \in \mathrm{Ker}L_1$ τότε

$$L_1L_2(v) = L_2L_1(v) = L_2\mathbf{0}_V = \mathbf{0}_V$$

συνεπώς $L_2(v) \in \mathrm{Ker}L_1$.

Αν $v \in \mathrm{Im}L_1$ τότε υπάρχει $w \in V$, με $L_1w = v$. Έχουμε ότι

$$L_2v = L_2L_1w = L_1(L_2w),$$

άρα και $L_2v \in \mathrm{Im}L_1$. □

Πρόταση VI.3.5. *Αν το ελάχιστο πολυώνυμο* $m_L(x)$ *της* L *διασπάται ως γινόμενο* $m_L(x) = m_1(x)m_2(x)$ *δύο μεταξύ τους πρώτων πολυωνύμων* $m_1(x), m_2(x)$ *τότε ο* V *διασπάται σε ένα ευθύ άθροισμα* L*-αναλλοίωτων υπόχωρων* W_1, W_2*, όπου*

1. $W_i = \mathrm{Ker}m_i(L) = \{v \in V : m_i(L)v = \mathbf{0}_V\}$
2. *Το ελάχιστο πολυώνυμο της* $L|_{W_i}$ *είναι το* $m_i(x)$.

Απόδειξη. Από την εξίσωση (I.10) στην παρατήρηση I.9.7 έχουμε ότι υπάρχουν πολυώνυμα $g_1(x), g_2(x)$ ώστε

$$1 = g_1(x)m_1(x) + p_1(x)m_2(x),$$

συνεπώς

$$\mathrm{Id}_V = g_1(L)m_1(L) + p_1(L)m_2(L) \tag{VI.10}$$

Για κάθε $v \in V$ έχουμε $m_1(L)g_1(L)v \in W_2$ και $m_2(L)g_2(L)v \in W_1$. Επιπλέον, $W_1 \cap W_2 = \{\mathbf{0}_V\}$. Πράγματι αν $v \in W_1 \cap W_2$ τότε η σχέση $m_1(L)v = m_2(L)v = \mathbf{0}_V$ μαζί με την εξίσωση (VI.10) επιβάλουν ότι $v = \mathbf{0}_V$.

Είναι σαφές ότι οι χώροι W_1, W_2 είναι L-αναλλοίωτοι σύμφωνα το λήμμα VI.3.4 αφού η L αντιμετατίθεται με κάθε πολυώνυμο του L, άρα $Lm_i(L) = m_i(L)L$ για $i = 1, 2$.

Τέλος παρατηρούμε ότι το $m_i(L)$ μηδενίζει εξ ορισμού κάθε διάνυσμα του $W_i = \mathrm{Ker} m_i(L)$, άρα το ελάχιστο πολυώνυμο $m_i'(x)$ του $L|_{W_i}$ διαιρεί το $m_i(x)$. Επίσης, αφού κάθε $v \in V$ γράφεται ως $v = v_1 + v_2$ με $v_i \in W_i$ έχουμε ότι το $m'(x) = m_1'(x)m_2'(x)$ μηδενίζει τον L, αφού

$$m'(L)(v) = m_1'(L)m_2'(L)(v) = m_1'(L)m_2'(L)(v_1) + m_1'(L)m_2'(L)(v_2) = \mathbf{0}_V.$$

Συνεπώς,

$$m_1(x)m_2(x) = m(x) \mid m_1'(x)m_2'(x) \mid m_1(x)m_2(x)$$

άρα $m_i'(x) = m_i(x)$, $i = 1, 2$. □

Επαγωγικά η παραπάνω πρόταση γενικεύεται στην

Πρόταση VI.3.6. *Αν το ελάχιστο πολυώνυμο $m_L(x)$ της L διασπάται ως γινόμενο $m_L(x) = m_1(x)m_2(x)\cdots m_s(x)$ s το πλήθος μεταξύ τους πρώτων πολυωνύμων $m_1(x), m_2(x), \ldots, m_s(x)$ τότε ο V διασπάται σε ένα ευθύ άθροισμα L-αναλλοίωτων υπόχωρων $W_1, W_2, \ldots, W_s$, όπου*

1. $W_i = \mathrm{Ker} m_i(L) = \{v \in V : m_i(L)v = \mathbf{0}_V\}$
2. *Το ελάχιστο πολυώνυμο της $L|_{W_i}$ είναι το $m_i(x)$.*

Θεώρημα VI.3.7. *Η γραμμική συνάρτηση $L : V \to V$, όπου V είναι διανυσματικός χώρος επί του αλγεβρικά κλειστού σώματος $\mathbb{F}$ είναι διαγωνοποίσιμη αν και μόνο αν όταν το ελάχιστο πολυώνυμο της είναι γινόμενο διαφορετικών μεταξύ τους γραμμικών παραγόντων.*

Απόδειξη. Αν η L είναι διαγωνοποίσιμη και $\dim V = n$, τότε υπάρχει βάση B ώστε $(L, B, B) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Αν $\lambda_{i_1}, \ldots, \lambda_{i_s}$ είναι οι διαφορετικές ανά δύο ιδιοτιμές τότε το ελάχιστο πολυώνυμο είναι το

$$m_L(x) = \prod_{\nu=1}^{s}(x - \lambda_{i_\nu}).$$

□

Αντιστρόφως, αν το ελάχιστο πολυώνυμο είναι γινόμενο γραμμικών παραγόντων τότε η πρόταση VI.3.6 μας δίνει μια διάσπαση του V σε L-αναλλοίωτους υπόχωρους

$$V = W_1 \bigoplus \cdots \bigoplus W_s$$

όπου ο περιορισμοί $L|_{W_i}$ έχουν ελάχιστο πολυώνυμο ίσο με $x - \lambda_i$ δηλαδή $L|_{W_i}$ είναι διαγώνιος πίνακας με ίδια ιδιοτιμή λ_i.

Λήμμα VI.3.8. *Αν $L : V \to V$ είναι διαγωνοποιήσιμη γραμμική συνάρτηση και $W \subset V$ είναι L-αναλλοίωτος υπόχωρος τότε $L|_W$ είναι διαγωνοποιήσιμη.*

Απόδειξη. Το ελάχιστο πολυώνυμο $m_{L|_W}(x)$ της $L|_W$ διαιρεί το ελάχιστο πολυώνυμο $m_L(x)$ της L. Πράγματι, αφού $m_L(L)(v) = \mathbf{0}_V$ για κάθε $v \in V$, το ίδιο θα ισχύει και για κάθε $w \in W$, δηλαδή $m_L(L|W) = \mathbf{0}_W$. Όμως η L είναι διαγωνοποιήσιμη και συνεπώς το ελάχιστο πολυώνυμο της διασπάται σε γραμμικούς, διαφορετικούς ανά δύο γραμμικούς παράγοντας. Το ίδιο θα συμβαίνει και με το $m_{L|_W}(x)$ και συνεπώς και η $m_{L|_W}$ είναι διαγωνοποιήσιμη. □

Πρόταση VI.3.9. *Θεωρούμε δύο διαγωνοποιήσιμες γραμμικές συναρτήσεις $L_1, L_2 : V \to V$. Υπάρχει κοινή βάση B ώστε $(L_1, B, B), (L_2, B, B)$ διαγώνιοι πίνακες αν και μόνο αν $L_1L_2 = L_2L_1$.*

Ισοδύναμα, για δύο διαγωνοποιήσιμους πίνακες $A_1, A_2 \in \mathbb{F}^{n,n}$ υπάρχει κοινός αντιστρέψιμος πίνακας $Q \in \mathbb{F}^{n,n}$ αν και μόνο αν $A_1A_2 = A_2A_1$.

Απόδειξη. Αν υπάρχει τέτοια βάση B, ώστε $(L_1, B, B), (L_2, B, B)$ διαγώνιοι, τότε $(L_1, B, B)(L_2, B, B) = (L_1, B, B), L_2, B, B)$ γιατί οι διαγώνιοι πίνακες αντιμετατίθενται. Συνεπώς $L_1L_2 = L_2L_1$.

Αντιστρόφως, έστω $L_1L_2 = L_2L_1$. Αφού η L_1 είναι διαγωνοποιήσιμη, έχει μια βάση αποτελούμενη από ιδιοδιανύσματα, δηλαδή

$$V = \mathrm{Ker}(L_1 - \lambda_1 I_V) \bigoplus \cdots \bigoplus \mathrm{Ker}(L_1 - \lambda_s I_V)$$

όπου $\lambda_1, \ldots, \lambda_s$ είναι οι ανά δύο διαφορετικές ιδιοτιμές της L_1. Σύμφωνα με το λήμμα VI.3.4 οι ιδιόχωροι $\mathrm{Ker}(L_1 - \lambda_i \mathrm{Id}_V)$ είναι L_2-αναλλοίωτοι, και από το λήμμα VI.3.8 οι $L_2|_{\mathrm{Ker}(L_1 - \lambda_i \mathrm{Id}_V)}$ είναι διαγωνοποιήσιμη. Αυτό σημαίνει ότι ο ιδιόχωρος $\mathrm{Ker}(L_1 - \lambda_i \mathrm{Id}_V)$ της ιδιοτιμής L_1, ο οποίος αποτελείται από ιδιοδιανύσματα της της L_1 ως προς την ιδιοτιμή λ_i έχει μια βάση από ιδιοδιανύσματα της L_2. Δηλαδή, υπάρχει μια βάση που αποτελείται ταυτόχρονα από ιδιοδιανύσματα και για την L_1, L_2, δηλαδή το ζητούμενο.

Η ταυτόχρονη διαγονοποίηση των πινάκων είναι ισοδύναμη με την ταυτόχρονη διαγωνοποίηση των αντίστοιχων γραμμικών συναρτήσεων. □

Ορισμός VI.3.10. Μια γραμμική συνάρτηση $L : V \to V$ θα λέγεται ημιαπλή αν και μόνο αν για κάθε L-αναλλοίωτο υπόχωρο $W \subset V$ υπάρχει L-αναλλοίωτος υπόχωρος W' ώστε $V = W \oplus W'$.

Πρόταση VI.3.11. *Μια γραμμική συνάρτηση $L : V \to V$, όπου V είναι διανυσματικός χώρος επί του αλγεβρικά κλειστού σώματος $\mathbb{F}$, είναι ημιαπλή αν και μόνο αν είναι διαγωνοποιήσιμη.*

Απόδειξη. Αν η L είναι διαγωνοποιήσιμη τότε σύμφωνα με το λήμμα VI.3.8 και ο περιορισμός της στο W είναι διαγωνοποιήσιμος. Διαλέγουμε μια βάση του W την $w_1, \ldots, w_s$ αποτελούμενη από ιδιοδιανύσματα την οποία την συμπληρώνουμε σε μία βάση από ιδιοδιανύσματα $w'_1, \ldots, w'_r$ του V. Αυτό μπορεί να γίνει με την βοήθεια της πρότασης IV.3.23. Το ζητούμενο αναλλοίωτο συμπλήρωμα είναι το $W' = \langle w'_1, \ldots, w'_r \rangle$.

Για το αντίστροφο θα χρησιμοποιήσουμε επαγωγή. Έστω ότι κάθε ημιαπλή γραμμική συνάρτηση $L' : W \to W$ είναι διαγωνοποιήσιμη για κάθε διανυσματικό χώρο διάστασης $n-1$. Θεωρούμε την $L : V \to V$, όπου $\dim V = n$. Αφού το $\mathbb{F}$ είναι αλγεβρικά κλειστό η L έχει μια ιδιοτιμή και έναν L-αναλλοίωτο μονοδιάστατο χώρο $\langle v \rangle$ που αντιστοιχεί σε ένα ιδιοδιάνυσμα $v \neq 0$ της ιδιοτιμής. Η ημιαπλότητα της L μας δίνει την ύπαρξη ενός L-αναλλοίωτου W', ώστε $V = \langle v \rangle \oplus W'$. Το ζητούμενο έπεται από την επαγωγική υπόθεση αφού $\dim W' = n - 1$. □

Παρατήρηση VI.3.12. Έστω $L : V \to V$ και έστω W ένας L-αναλλοίωτος υπόχωρος. Μπορούμε να ορίζουμε μια γραμμική συνάρτηση $L_{V/W} : V/W \to V/W$, ως εξής:

$$V/W \ni v/w \longmapsto L_{V/W}(v+W) = L(v) + W \in V/W.$$

Η παραπάνω είναι καλά ορισμένη αφού αν $v = v' + w$, έχουμε ότι $L(v) = L(v') + L(w)$, με $L(w) \in W$. Αν διαλέξουμε μία βάση $B_W = \{w_1, \ldots, w_n\}$ του W την οποία την επεκτείνουμε σε βάση $B_V = \{w_1, \ldots, w_n, v_1, \ldots, v_s\}$ του V τότε ο πίνακας

$$(L, B_V, B_V) = \left(\begin{array}{c|c} (L|_W, B_W, B_W) & * \\ \hline \mathbf{0}_{s,n} & (L_{V/W}, B_{V/W}, B_{V/W}) \end{array}\right), \tag{VI.11}$$

όπου $B_{V/W} = \{v_1 + W, \ldots, v_s + W\}$ είναι μια βάση του V/W.

Πρόταση VI.3.13. *Τα χαρακτηριστικά πολυώνυμα* $\mathrm{Ch}_L(x)$, $\mathrm{Ch}_{L|_W}(x)$ *και* $\mathrm{Ch}_{L_{V/W}}(x)$ *συνδέονται με την σχέση*

$$\mathrm{Ch}_L(x) = \mathrm{Ch}_{L|_W}(x) \cdot \mathrm{Ch}_{L_{V/W}}(x).$$

Απόδειξη. Από την εξίσωση (VI.11) έχουμε ότι

$$\mathrm{Ch}_L(x) = \det\left(\begin{array}{c|c} (L|_W, B_W, B_W) - x\mathbb{I}_n & * \\ \hline \mathbf{0}_{s,n} & (L_{V/W}, B_{V/W}, B_{V/W}) - x\mathbb{I}_s \end{array}\right) = \mathrm{Ch}_{L|_W}(x) \cdot \mathrm{Ch}_{L_{V/W}}(x).$$

□

VI.3.2 Τριγωνίσιμοι Πίνακες

Θα αποδείξουμε μια

Θεώρημα VI.3.14. *Θεωρούμε μια γραμμική συνάρτηση* $L : V \to V$, *όπου ο* V *έχει πεπερασμένη διάσταση* V, *πάνω από ένα αλγεβρικά κλειστό σώμα* $\mathbb{F}$. *Υπάρχει μια βάση* $B = \{v_1, \ldots, v_n\}$ *του* V *στην οποία ο πίνακας* (L, B, B) *είναι άνω τριγωνικός δηλαδή της μορφής*

$$(L, B, B) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

Απόδειξη. Αν $\dim V = 1$ το αποτέλεσμα είναι σαφές. Θα αποδείξουμε το ζητούμενο με επαγωγή. Υποθέτουμε ότι το θεώρημα ισχύει για κάθε διανυσματικό χώρο διάστασης $n-1$.

Το χαρακτηριστικό πολυώνυμο του έχει βαθμό n και συνεπώς έχει μια ρίζα που είναι ιδιοτιμή και έχει ένα μη-μηδενικό ιδιοδιάνυσμα v_1. Συμπληρώνουμε σε μία βάση $B_1 = \{v_1, b_2, \ldots, b_n\}$ του V ως προς την οποία ο πίνακας της L γράφεται

$$(L, B_1, B_1) = \left(\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline 0 & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & a_{n-1,n} \\ 0 & a_{n2} & \cdots & a_{nn} \end{array}\right) = \left(\begin{array}{c|c} a_{11} & a_{12} \ \cdots \ a_{1n} \\ \hline 0 & \\ \vdots & (L_{V/\langle v_1 \rangle}, B_1', B_1') \\ 0 & \end{array}\right),$$

όπου $L_{V/\langle v_1 \rangle}$ είναι η επαγώμενη συνάρτηση στον χώρο πηλίκο $V/\langle v_1 \rangle$ και $B_1 = \{b_2 + \langle v_1 \rangle, \ldots, b_n + \langle v_1 \rangle\}$ είναι μια βάση του χώρου πηλίκου. Παρατηρούμε ότι $\dim V/\langle v_1 \rangle = n-1$ οπότε από την επαγωγική υπόθεση έχουμε ότι υπάρχει μια βάση $_1 = \{v_2 + \langle v_1 \rangle, \ldots, v_n + \langle v_1 \rangle\}$ ως προς την οποία ο $(L_{V/\langle v_1 \rangle}, B_1, B_1)$ να είναι άνω τριγωνικός. Ο πίνακας της L ως προς την βάση $B = \{v_1, v_2, \ldots, v_n\}$, έχει την ζητούμενη μορφή. □

Παρατηρούμε ότι τα στοιχεία της διαγωνίου είναι ιδιοτιμές, δηλαδή $a_{ii} = \lambda_i$. Έχουμε το χαρακτηριστικό πολυώνυμο

$$\mathrm{Ch}_A(x) = (-1)^n \prod_{i=1}^{n}(x-\lambda_i) = (-1)^n\left(x^n - \sum_{i=1}^{n} x^{n-1}\lambda_i + \cdots + (-1)^n\lambda_1\lambda_2\cdots\lambda_n\right).$$

Δηλαδή ο σταθερός όρος του χαρακτηριστικού πολυωνύμου είναι το γινόμενο των ιδιοτιμών δηλαδή η ορίζουσα, κάτι το οποίο γνωρίζαμε και από τον ίδιο τον ορισμό του χαρακτηριστικού πολυωνύμου. Αλλά το άθροισμα των ιδιοτιμών είναι το ίχνος του πίνακα και σχετίζεται με τον συντελεστή a_{n-1} του x^{n-1} στο χαρακτηριστικό πολυώνυμο από την σχέση

$$\mathrm{tr}(A) = (-1)^{n+1}a_{n-1}.$$

Παρατηρήστε ότι όμοιοι πίνακες έχουν το ίδιο ίχνος, άρα το ίχνος μπορεί να υπολογιστεί από το άθροισμα των ιδιοτιμών, δείτε και την άσκηση II.3.7.

Άσκηση VI.3.15. Δίνεται ένας 3×3 πίνακας με συντελεστές πραγματικούς. Δείξτε ότι αν ο πίνακας δεν είναι τριγωνοποιήσιμος στο $\mathbb{R}$ τότε είναι διαγωνοποιήσιμος στο $\mathbb{C}$.

Λύση VI.3.16. Ένας τριγωνοποιήσιμος πίνακας στο $\mathbb{R}$ έχει όλα τα στοιχεία του στο $\mathbb{R}$ άρα και τα διαγώνια στοιχεία που είναι πραγματικοί αριθμοί. Ένας πίνακας αποτυγχάνει να είναι τριγωνοποιήσιμος στο $\mathbb{R}$ αν το χαρακτηριστικό του πολυώνυμο δεν έχει όλες τις ρίζες πραγματικές. Συνεπώς έχει μια πραγματική ρίζα και δύο συζυγείς μιγαδικές στο $\mathbb{C}$. Άρα στους μιγαδικούς αριθμούς έχει τρεις διαφορετικές ρίζες και είναι διαγωνοποιήσιμος.

VI.3.3 Ανάλυση Jordan

Ορισμός VI.3.17. Θα ονομάζουμε *Jordan block* $J(\lambda, k)$ τον $k \times k$ πίνακα

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

Λήμμα VI.3.18. *Το χαρακτηριστικό και ελάχιστο πολυώνυμο του $J(\lambda, k)$ υπολογίζονται*

$$\mathrm{Ch}_{J(\lambda,k)}(x) = (-1)^k(x-\lambda)^k$$
$$p_{J(\lambda,k)}(x) = (x-\lambda)^k$$

Απόδειξη. Το χαρακτηριστικό πολυώνυμο προκύπτει από τον υπολογισμό της ορίζουσας

$$\mathrm{Ch}_{J(\lambda,k)}(x) = \det(J(\lambda,k) - x\mathbb{I}_k) = \det\begin{pmatrix} \lambda-x & 1 & 0 & \cdots & 0 \\ 0 & \lambda-x & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda-x & 1 \\ 0 & \cdots & \cdots & 0 & \lambda-x \end{pmatrix} = (\lambda-x)^k.$$

Το ελάχιστο πολυώνυμο $p_{J(\lambda,k)}$ διαιρεί το χαρακτηριστικό πολυώνυμο συνεπώς είναι της μορφής $(x-\lambda)^\nu$ για κάποιο $0 < \nu \leq k$. Από το παράδειγμα VI.2.5 έχουμε ότι $\nu = k$. □

Παρατήρηση VI.3.19. Ο πίνακας $J(\lambda, k)$ για $k \geq 2$ δεν είναι διαγωνοποιήσιμος. Πράγματι η αλγεβρική πολλαπλότητα της ιδιοτιμής λ είναι ίση με k ενώ η γεωμετρική ίση με 1.

Θεώρημα VI.3.20 (Κανονική μορφή Jordan). *Για ένα $\mathbb{F}$-διανυσματικό χώρο όπου το $\mathbb{F}$ είναι αλγεβρικά κλειστό σώμα, οποιοσδήποτε πίνακας $A \in \mathbb{F}^{n,n}$ είναι όμοιος με τον πίνακα*

$$\begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_s \end{pmatrix} \tag{VI.12}$$

όπου για κάθε $1 \le i \le s$ οι πίνακες $J_i = J(\lambda_i, k_i)$ για $\lambda_i \in \mathbb{F}$ και $0 < k \in \mathbb{N}$.

Παρατήρηση VI.3.21. Στην παραπάνω ανάλυση οι τιμές $\lambda_1, \ldots, \lambda_s$ δεν είναι ανάγκη να είναι διαφορετικές.

Η απόδειξη του θεωρήματος της κανονικής μορφής Jordan βασίζεται στα παρακάτω λήμματα:

Λήμμα VI.3.22. *Έστω V ένας $\mathbb{F}$-διανυσματικός χώρος, όπου το σώμα $\mathbb{F}$ είναι αλγεβρικά κλειστό και $L : V \to V$ γραμμική συνάρτηση. Έστω $\lambda_1, \ldots, \lambda_r$ διαφορετικές ανά δύο ιδιοτιμές της L. Τότε υπάρχουν ακέραιοι $s_1, \ldots, s_r$, ώστε*

$$V = \mathrm{Ker}(L - \lambda_1 \mathrm{Id}_V)^{s_1} \bigoplus \mathrm{Ker}(L - \lambda_2 \mathrm{Id}_V)^{s_2} \bigoplus \cdots \bigoplus \mathrm{Ker}(L - \lambda_1 \mathrm{Id}_V)^{s_r},$$

όπου με $(L - \lambda_\nu \mathrm{Id}_V)^{s_\nu}$ στον παραπάνω τύπο εννοούμε την s_ν φορές διαδοχική σύνθεση της γραμμικής συνάρτησης $L - \lambda_\nu \mathrm{Id}_V$.

Απόδειξη. Διαλέγουμε μία από τις ιδιοτιμές του L, έστω λ. Η απόδειξη του λήμματος θα γίνει σε 5 βήματα.
Βήμα 1. Ορίζουμε τον χώρο $W_i = \mathrm{Ker}(L - \lambda \mathrm{Id}_V)^i$ για κάθε $i \ge 1$ οι όποιοι έχουν την παρακάτω σχέση εγκλεισμού

$$W_1 \subset W_2 \subset W_3 \subset \cdots \subset W_i \subset \cdots$$

Αφού ο χώρος V είναι πεπερασμένης διάστασης η παραπάνω ακολουθία δεν μπορεί να είναι γνήσια αύξουσα για κάθε i, άρα υπάρχει ένα t ώστε $W_t = W_{t+1}$. Υποθέτουμε ότι αυτό είναι το μικρότερο δυνατό t. Είναι σαφές, ότι αν $W_t = W_{t+1}$ τότε $W_t = W_{t+1} = W_{t+2} = \cdots$ δηλαδή η ακολουθία των χώρων γίνεται σταθερή. Πράγματι $W_{t+1} \subset W_{t+2}$ και αν $v \in W_{t+2} = \mathrm{Ker}(L - \lambda \mathrm{Id}_V)^{t+2}$ τότε $(L - \lambda \mathrm{Id}_V)v \in \mathrm{Ker}(L - \lambda \mathrm{Id}_V)^{t+1} = \mathrm{Ker}(L - \lambda \mathrm{Id}_V)^t$, άρα $(L - \lambda \mathrm{Id}_V)^t (L - \lambda \mathrm{Id}_V)v = \mathbf{0}_V$ οπότε $v \in \mathrm{Ker}(L - \lambda \mathrm{Id}_V)^{t+1}$.
Βήμα 2. Τώρα θα αποδείξουμε ότι $\mathrm{Ker}(L - \lambda \mathrm{Id}_V)^t \cap \mathrm{Im}(L - \lambda \mathrm{Id}_V)^t = \mathbf{0}_V$. Έστω ότι για ένα $v \in V$ ανήκει στην τομή. Συνεπώς υπάρχει ένα $w \in V$ με $v = (L - \lambda \mathrm{Id}_V)^t w$. Από την άλλη $(L - \lambda \mathrm{Id}_V)^t v = \mathbf{0}_V$, οπότε $(L - \lambda \mathrm{Id}_V)^{2t} w = (L - \lambda \mathrm{Id}_V)^t v = \mathbf{0}_V$, άρα $w \in \mathrm{Ker}(L - \lambda \mathrm{Id}_V)^{2t} = W_{2t} = W_t$, το οποίο δίνει ότι $v = (L - \lambda \mathrm{Id}_V)^t w = 0$.
Βήμα 3. Γνωρίζουμε ότι $\dim V = \dim\mathrm{Ker}(L - \lambda \mathrm{Id}_V)^t + \dim \mathrm{Im}(L - \lambda \mathrm{Id}_V)^t$. Επιπλέον γνωρίζουμε ότι

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Οπότε για $V_1 = \mathrm{Ker}(L - \lambda \mathbb{I}_n)^t$, $V_2 = \mathrm{Im}(L - \lambda \mathrm{Id}_V)^t$ που έχουν μηδενική τομή έχουμε

$$\dim(\mathrm{Ker}(L - \lambda \mathrm{Id}_V)^t + \mathrm{Im}(L - \lambda \mathrm{Id}_V)^t) = \dim V$$

άρα

$$V = \mathrm{Ker}(L - \lambda \mathrm{Id}_V)^t \bigoplus \mathrm{Im}(L - \lambda \mathrm{Id}_V)^t.$$

Βήμα 4. Παρατηρούμε ότι οι χώροι $\mathrm{Ker}(L - \lambda \mathrm{Id}_V)^t$ και $\mathrm{Im}(L - \lambda \mathrm{Id}_V)^t$ είναι και οι δύο L-αναλλοίωτοι. Πράγματι, παρατηρούμε ότι

$$L(L - \lambda \mathbb{I}_n) = (L - \lambda \mathrm{Id}_V)L$$

άρα αν $v \in \mathrm{Ker}(L - \lambda \mathrm{Id}_V)^t$ τότε

$$(L - \lambda \mathrm{Id}_V)^t L(v) = L(L - \lambda \mathrm{Id}_V)^t v = \mathbf{0}_V,$$

οπότε $L\mathrm{Ker}(L - \lambda \mathrm{Id}_V)^t \subset \mathrm{Ker}(L - \lambda \mathrm{Id}_V)^t$. Ομοίως αν $v \in \mathrm{Im}(L - \lambda \mathrm{Id}_V)^t$, τότε $v = (L - \lambda \mathrm{Id}_V)^t w$, οπότε

$$Lv = L(L - \lambda \mathrm{Id}_V)^t w = (L - \lambda \mathrm{Id}_V)^t L w,$$

άρα $L(\mathrm{Im}(L - \lambda \mathrm{Id}_V)^t) \subset \mathrm{Im}(L - \lambda \mathrm{Id}_V)^t$.

Βήμα 5. Το λήμμα μπορεί να αποδειχθεί με επαγωγή στο πλήθος των διαφορετικών ιδιοτιμών της συνάρτησης L. Αν λοιπόν $\lambda_1, \ldots, \lambda_s$ είναι οι διαφορετικές ανά δύο ιδιοτιμές της L, εφαρμόζουμε τα βήματα 1-4 για κάθε $\lambda = \lambda_1$ και στην συνέχεια έχουμε τον περιορισμό της L στον $\mathrm{Im}(L - \lambda \mathrm{Id}_V)^t$ ο οποίος έχει ιδιοτιμές $\lambda_2, \ldots, \lambda_s$, οπότε από την επαγωγική υπόθεση έχουμε ότι

$$\mathrm{Im}(L - \lambda_1 \mathrm{Id}_V)^t = \mathrm{Ker}(L - \lambda_2 \mathrm{Id}_V)^{s_2} \bigoplus \cdots + \bigoplus \mathrm{Ker}(L - \lambda \mathrm{Id}_V)^{s_r}.$$

□

Λήμμα VI.3.23 (Mark Wildon). *Έστω V είναι ένας $\mathbb{F}$-διανυσματικός χώρος διάστασης n, και $T : V \to V$ είναι γραμμική συνάρτηση με $T^s = \mathbf{0}_{\mathrm{Env}(V)}$ για κάποιο φυσικό αριθμό s, τότε υπάρχουν διανύσματα $u_1, \ldots, u_k$ και φυσικοί αριθμοί $a_1, \ldots, a_k$ ώστε*

$$T^{a_i}(u_i) = \mathbf{0}_V, \ \textit{για}\ i = 1, \ldots, k,$$

και τα διανύσματα

$$u_1, T(u_1), \ldots, T^{a_1-1}u_1, \ldots, u_k, T(u_k), \ldots, T^{a_k-1}(u_k)$$

σχηματίζουν μια βάση του χώρου V.

Απόδειξη. Αν η T είναι η μηδενική συνάρτηση, τότε μπορούμε να διαλέξουμε τα $u_1, \ldots, u_k$ να είναι μια βάση του V και $a_1 = \cdots = a_k = 1$.

Θα δώσουμε μια επαγωγική απόδειξη στην διάσταση του V. Αν το $n = 1$ τότε T^s είναι η μηδενική συνάρτηση αν και μόνο αν η T είναι η μηδενική συνάρτηση οπότε επαναλαμβάνουμε το επιχείρημα για μηδενική συνάρτηση.

Υποθέτουμε ότι το λήμμα ισχύει για διαστάσεις μικρότερες του n και θα το αποδείξουμε για την διάσταση n. Θεωρούμε τον χώρο ImT. Αν $\dim \mathrm{Im}T = 0$ τότε η T είναι η μηδενική συνάρτηση και το λήμμα έπεται. Αν $\dim \mathrm{Im}T = n$ τότε η T είναι 1-1 και δεν μπορεί T^s να είναι η μηδενική συνάρτηση για κάποιο s, άτοπο. Άρα μπορούμε να υποθέσουμε ότι $0 < \dim \mathrm{Im}T < n$, και από την επαγωγική υπόθεση έχουμε ότι υπάρχουν διανύσματα $v_1, \ldots, v_\ell$ και φυσικοί αριθμοί $b_1, \ldots, b_\ell$ ώστε

$$T^{b_i}(v_i) = \mathbf{0}_V \text{ για } i = 1, \ldots, \ell$$

και

$$v_1, T(v_1), \ldots, T^{b_1-1}(v_1), \ldots, v_\ell, T(v_\ell), \ldots, T^{b_\ell-1}(v_\ell) \tag{VI.13}$$

να αποτελούν μια βάση του ImT.

Αφού τα $v_1, \ldots, v_\ell \in \mathrm{Im}(T)$ υπάρχουν διανύσματα $w_1, \ldots, w_\ell$ ώστε $T(w_i) = v_i$ για $i = 1, \ldots, \ell$. Από την άλλη τα διανύσματα $T^{b_1-1}v_1, \ldots, T^{b_\ell-1}v_\ell$ είναι γραμμικά ανεξάρτητα διανύσματα του KerT, τα οποία τα επεκτείνουμε σε μία βάση

$$T^{b_1-1}v_1, \ldots, T^{b_\ell-1}v_\ell, z_1, \ldots, z_m \tag{VI.14}$$

του $\mathrm{Ker}T$. Παρατηρούμε ότι $T^j(w_i) = T^{j-1}(v_i)$ για όλες τις επιλογές των i, j. Αρκεί να αποδείξουμε ότι

$$w_1, T(w_1), \dots, T^{b_1}(w_1), \dots, w_\ell, T(w_\ell), \dots, T^{b_\ell}(w_\ell), z_1, \dots, z_m \tag{VI.15}$$

αποτελούν μια βάση του V.

Θα δείξουμε πρώτα ότι είναι γραμμικά ανεξάρτητα. Υποθέτουμε την ύπαρξη ενός γραμμικού συνδυασμού

$$\sum_{i=1}^{\ell}\sum_{j=0}^{b_i} \alpha_{i,j} T^j(w_i) + \sum_{\nu=1}^{m} \beta_\nu z_\nu = \mathbf{0}_V. \tag{VI.16}$$

Εφαρμόζουμε τον T στην παραπάνω σχέση για να πάρουμε

$$\sum_{i=1}^{\ell}\sum_{j=0}^{b_i} \alpha_{i,j} T^{j+1}(w_i) = \sum_{i=1}^{\ell}\sum_{j=0}^{b_i-1} \alpha_{i,j} T^j(v_i)$$

συνεπώς τα

$$\alpha_{1,0} = \alpha_{1,1} = \alpha_{1,b_1-1} = \cdots = \alpha_{\ell,0} = \alpha_{\ell,1} = \cdots = \alpha_{\ell,b_\ell-1} = 0.$$

Επιστρέφουμε με αυτή την πληροφορία στην εξίσωση (VI.16) για να πάρουμε

$$\alpha_{1,b_1} T^{b_1}(w_1) + \cdots + \alpha_{\ell,b_\ell} T^{b_\ell} + \beta_1 z_1 + \cdots + \beta_m z_m = \not{0}_V,$$

όμως τα διανύσματα που εμφανίζονται στο παραπάνω άθροισμα είναι βάση του $\mathrm{Ker}T$, οπότε

$$\alpha_{1,b_1} = \cdots = \alpha_{\ell,b_\ell} = \beta_1 = \cdots = \beta_m = 0.$$

Για να δείξουμε ότι το σύνολο των διανυσμάτων της εξίσωσης (VI.15) είναι βάση αρκεί, σύμφωνα με το θεώρημα IV.3.12, να δείξουμε ότι το πλήθος τους είναι όση και η διάσταση του χώρου V. Τα στοιχεία της εξίσωσης (VI.13) είναι μια βάση του $\mathrm{Im}T$, οπότε $\dim \mathrm{Im}T = b_1 + \cdots + b_\ell$. Τα διανύσματα της εξίσωσης (VI.14) είναι μια βάση του πυρήνα οπότε $\dim \mathrm{Ker}T = l + m$. Συνεπώς

$$\begin{aligned}\dim V = \dim \mathrm{Ker}T + \dim \mathrm{Im}T &= b_1 + \cdots + b_\ell + \ell + m\\ &= (1+b_1) + \cdots + (1+b_\ell) + m,\end{aligned}$$

δηλαδή ακριβώς το πλήθος των διανυσμάτων της εξίσωσης (VI.15). □

Παράδειγμα VI.3.24. Θεωρούμε τον πίνακα

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

Έχουμε ότι $\mathrm{Im}(A) = \langle (1,0,0)^t \rangle$. Διαλέγουμε $v_1 = (1,0,0)^t$. Τα διανύσματα $(0,1,0)^t$ και $(0,0,1)^t$ απεικονίζονται μέσω της A στο v_1, όπως και κάθε διάνυσμα της μορφής $(x_1, x_2, 1-x_2)^t$. Διαλέγουμε ένα από αυτά, για παράδειγμα το $u_1 = (0,0,1)^t$. Τέλος, συμπληρώνουμε την βάση με ένα διάνυσμα του πυρήνα για παράδειγμα το $u_2 = (0,1,-1)^t$. Έτσι έχουμε την βάση $\{u_1, T(u_1) = v_1, u_2\}$, οπότε $a_1 = 2, a_2 = 1$ και παρατηρούμε ότι $T^2(u_1) = \mathbf{0} = T(u_2)$.

Προχωράμε τώρα στην απόδειξη του θεωρήματος VI.3.20. Από το λήμμα VI.3.22 έχουμε ότι υπάρχουν ακέραιοι $s_1, \dots, s_r$ ώστε

$$V = \mathrm{Ker}(L - \lambda_1 \mathrm{Id}_V)^{s_1} \bigoplus \cdots \bigoplus \mathrm{Ker}(L - \lambda_1 \mathrm{Id}_V)^{s_r},$$

όπου $\lambda_1, \dots, \lambda_r$ είναι οι διαφορετικές ιδιοτιμές του L. Από το λήμμα VI.3.23 διαλέγουμε βάσεις για τους υπόχωρους $\mathrm{Ker}(L - \lambda_1 \mathrm{Id}_V), \dots, \mathrm{Ker}(L - \lambda_s \mathrm{Id}_V)$ της οποίες τις γράφουμε την μία μετά την άλλη και παρατηρούμε ότι ο πίνακας της γραμμικής συνάρτησης ως προς αυτή την βάση είναι πράγματι ο πίνακας της εξίσωσης (VI.12).

Μέθοδος 3 (Υπολογισμός κανονικής μορφής Jordan). *Θεωρούμε μια γραμμική συνάρτηση* $L : V \to V$, *η οποία εκφράζεται με την βοήθεια πίνακα* A. *Για να υπολογίσουμε την κανοική μορφή Jordan, εκτελούμε τα παρακάτω βήματα:*

1. *Υπολογίζουμε τις ιδιοτιμές του πίνακα* A *και τα ιδιοδιανύσματα που αντιστοιχούν σε κάθε ιδιοτιμή.*

2. *Το πλήθος των γραμμικά ανεξάρτητων ιδιοδιανυσμάτων που υπολογίσαμε δίνει το πλήθος των διαφορετικών Jordan blocks.*

3. *Για κάθε ιδιοδιάνυσμα* v *που υπολογίσαμε, και που αντιστοιχεί στην ιδιοτιμή* λ, *λύνουμε το σύστημα* $(A - \lambda\mathbb{I}_n)w_1 = v$. *Στην συνέχεια υπολογίζουμε το* $(A - \lambda\mathbb{I}_n)w_2 = w_1$ *και συνεχίζουμε με τον ίδιο τρόπο υπολογίζοντας*
$$(A - \lambda\mathbb{I}_n)w_3 = w_2, (A - \lambda\mathbb{I}_n)w_4 = w_3, \dots, (A - \lambda\mathbb{I}_n)w_n = w_{n-1},$$
και η διαδικασία σταματάει όταν το σύστημα $(A - \lambda\mathbb{I}_n)w_{n+1} = w_n$ *είναι αδύνατο.*

4. *Σχηματίζουμε τον πίνακα* P *ο οποίος έχει ως στήλες τα διανύσματα* $v, w_1, \dots, w_n$ *τα οποία υπολογίσαμε για κάθε ιδιοτιμή. Ο πίνακας* $P^{-1}AP$ *είναι ο πίνακας της κανονικής μορφής Jordan.*

Παραδείγματα VI.3.25. 1. Θεωρούμε τον πίνακα
$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$
ο οποίος έχει χαρακτηριστικό πολυώνυμο $(x-2)^3$ και ιδιοτιμή $\lambda = 2$. Το σύστημα $(A-2\mathbb{I}_n)x = 0$ έχει μονοδιάστατο πυρήνα ο οποίος παράγεται από το $v = (1, 0, 0)^t$. Στην συνέχεια λύνουμε το σύστημα
$$(A - 2\mathbb{I}_n)w_1 = v \Leftrightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$
το οποίο έχει χώρο λύσεων $x(1, 0, 0)^t + (0, 1, 0)^t, t \in \mathbb{R}$. Διαλέγουμε την απλούστερη λύση $w_1 = (0, 1, 0)^t$. Επαναλαμβάνουμε την διαδικασία
$$(A - 2\mathbb{I}_n)w_2 = w_1 \Leftrightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$
το οποίο έχει χώρο λύσεων $x(1, 0, 0)^t + (0, 0, 1)^t, t \in \mathbb{R}$. Διαλέγουμε ως λύση το $w_2 = (0, 0, 1)^t$. Τα διανύσματα v, w_1, w_2 είναι γραμμικά ανεξάρτητα και αποτελούν μια βάση του $\mathbb{R}^3$ οπότε μπορούμε να σταματήσουμε εδώ. Παρατηρούμε ότι στην επόμενη επανάληψη
$$(A - 2\mathbb{I}_n)w_3 = w_2 \Leftrightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$
το σύστημα είναι αδύνατο οπότε δεν υπάρχει άλλο w_3. Ο πίνακας P που αποτελείται από τα διανύσματα v, w_1, w_2 που υπολογίσαμε ως στήλες είναι ο ταυτικός. Αυτό είναι προφανές, αφού ξεκινήσαμε να εκτελούμε την μέθοδο εύρεσης του πίνακα Jordan σε ένα πίνακα που ήταν ήδη σε αυτή την μορφή.

2. **Τρία γραμμικά ανεξάρτητα ιδιοδιανύσματα** Ας προσπαθήσουμε να υπολογίζουμε την κανονική μορφή Jordan του πίνακα

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

Υπολογίζουμε το χαρακτηριστικό πολυώνυμο $-(x-3)^2(x-5)$ το οποίο έχει τις ιδιοτιμές $\lambda = 3$ και $\lambda = 5$. Υπολογίζουμε τους ιδιόχωρους E_3, E_5 λύνοντας τα ομογενή συστήματα:

$$(A - 3\mathbb{I}_3)x = \mathbf{0}, \qquad (A - 5\mathbb{I}_3)x = \mathbf{0}$$

και έχουμε ότι

$$E_3 = \langle (0,1,0)^t, (-1,0,1)^t \rangle, E_5 = \langle (1,2,1)^t \rangle.$$

Παρατηρούμε ότι έχουμε βρει $3 = \dim \mathbb{R}^3$ γραμμικά ανεξάρτητα διανύσματα. Ο πίνακας είναι διαγωνοποιήσιμος. Θέτουμε

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

και υπολογίζουμε

$$P^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ -1 & 1 & -1 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \qquad P^{-1}AP = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Ο πίνακας είναι διαγωνοποιήσιμος και τα Jordan blocks είναι οι 1×1 πίνακες $[5], [3], [3]$.

3. **Δύο γραμμικά ανεξάρτητα ιδιοδιανύσματα** Θεωρούμε τον πίνακα

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

το οποίο έχει χακτηριστικό πολυωνυμο $-(x-1)^3$. Υπολογίζουμε μια βάση του ιδιόχωρου E_1 της ιδιοτιμής $\lambda = 1$ ως την λύση του συστήματος

$$(A - \mathbb{I}_3)x = \mathbf{0} \Leftrightarrow \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

$$E_1 = \langle (1,0,0)^t, (0,1,-1)^t \rangle.$$

Προσπαθούμε να βρούμε ένα διάνυσμα w_1 που να ικανοποιεί το σύστημα

$$(A - \mathbb{I}_3)w_1 = (0,1,-1)^t \Leftrightarrow \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Το σύστημα αυτό όμως είναι αδύνατο και η διαδικασία στο βήμα 3 της μεθόδου 4 σταματάει. Προσπαθούμε τώρα να βρούμε ένα διάνυσμα w_1 που να ικανοποιεί το σύστημα

$$(A - \mathbb{I}_3)w_1 = (1,0,0)^t \Leftrightarrow \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Το σύστημα αυτό έχει λύσεις (παρατηρήστε ότι το αντίστοιχο ομογενές το έχουμε ήδη λύσει)

$$x\begin{pmatrix}1\\0\\0\end{pmatrix}+y\begin{pmatrix}0\\1\\-1\end{pmatrix}+\begin{pmatrix}0\\0\\1\end{pmatrix}$$

Διαλέγουμε στον χώρο λύσεων την απλούστερη λύση, δηλαδή την $w_1 = (0,0,1)^t$. Έχουμε ήδη βρεί τρία γραμμικά ανεξάρτητα διανύσματα, τα $(1,0,0)^t,(0,1,-1)^t,(0,0,1)^t$ οπότε μπορούμε να σταματήσουμε εδώ. Αν προσπαθούσαμε να συνεχίσουμε θα ψάχναμε για μια λύση στο σύστημα

$$A-\mathbb{I}_3)w_2=(0,0,1)^t \Leftrightarrow \begin{pmatrix}0&1&1\\0&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\1\end{pmatrix}$$

θα βλέπαμε ότι το σύστημα είναι αδύνατο. Μπορούμε τώρα να θεωρήσουμε τον πίνακα

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

και να επαληθεύσουμε ότι

$$P^{-1}=\begin{pmatrix}0&1&0\\1&0&0\\0&1&1\end{pmatrix}\qquad P^{-1}AP=\left(\begin{array}{c|cc}1&0&0\\\hline 0&1&1\\0&0&1\end{array}\right)$$

4. **Ένα γραμμικά ανεξάρτητο ιδιοδιάνυσμα** Σε αυτό το παράδειγμα θεωρούμε τον πίνακα

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

του οποίου θέλουμε να υπολογίσουμε την κανονική μορφή Jordan του. Υπολογίζουμε ότι το χαρακτηριστικό του πολυώνυμο είναι το $-(x+1)^3$, συνεπώς έχουμε μοναδική ιδιοτιμή $\lambda=-1$ και τον ιδιόχωρο $E_{-1}=\{x:\mathbb{R}^3:(A+\mathbb{I}_3)x=\mathbf{0}\}$, ο οποίος έχει διάσταση ένα και παράγεται από το $v=(1,0,0)^t$. Αναζητούμε ένα διάνυσμα

$$(A+\mathbb{I}_3)w_1=v\Leftrightarrow\begin{pmatrix}0&-1&0\\0&0&-2\\0&0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1\\0\\0\end{pmatrix}$$

Παρατηρούμε ότι ο χώρος λύσεων είναι το $x(1,0,0)^t+(0,-1,0)^t, x\in\mathbb{R}$ και διαλέγουμε την απλούστερη λύση $w_1=(0,-1,0)^t$. Συνεχίζουμε, αναζητώντας τις λύσεις του συστήματος

$$(A+\mathbb{I}_3)w_1=v\Leftrightarrow\begin{pmatrix}0&-1&0\\0&0&-2\\0&0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\-1\\0\end{pmatrix}$$

το οποίο έχει λύσεις $x(1,0,0)^t+(0,0,1/2)^t, x\in\mathbb{R}$. Και πάλι διαλέγουμε την λύση $w_2=(0,0,1/2)^t$. Τα v,w_1,w_2 είναι γραμμικά εξαρτημένα και έχουμε τελειώσει. Αν επιχειρούσαμε να βρούμε μια λύση $(A+\mathbb{I}_3)w_3=w_2$, θα διαπιστώναμε ότι το τελευταίο σύστημα είναι αδύνατο. Διαλέγουμε τον πίνακα P

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

και επαληθεύουμε

$$P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \qquad P^{-1}AP = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Ο Marie Ennemond Camille Jordan (5 Ιανουαρίου 1838 - 22 Ιανουαρίου 1922) ήταν Γάλλος μαθηματικός με σημαντική συνεισφορά στην θεωρία ομάδων και ιδιαίτερα γνωστός για το σύγγραμμά του Cours d'analyse.

Άφησε το όνομα του σε μια σειρά από αποτελέσματα, όπως το θεώρημα της κλειστής καμπύλης του Jordan στην τοπολογία ή το θεώρημα Jordan–Hölder στις συνθετικές σειρές στην θεωρία ομάδων την έννοια των οποίων εισήγαγε όπως και έθεσε το πρόβλημα της ταξινόμησης των επιλύσιμων ομάδων. Την εργασία του στην θεωρία των ομάδων δημοσίευσε το 1870 στο πρώτο ιστορικά βιβλίο θεωρίας ομάδων: Traité des substitutions et des équations algebraique, το οποίο περιείχε και μια περιεκτική μελέτη της θεωρίας Galois. Επίσης είχε συνεισφορά στην μελέτη των ομάδων πινάκων τόσο πάνω από σώματα χαρακτηριστικής μηδέν όσο και πάνω από σώματα θετικής χαρακτηριστικής.

Θεώρημα VI.3.26. *Για κάθε πίνακα $A \in \mathbb{C}^{n,n}$ υπάρχει μια ακολουθία διαγωνοποιήσιμων πινάκων $A_n \to A$.*

Απόδειξη. Κάθε πίνακας A είναι όμοιος με ένα άνω τριγωνικό πίνακα B ο οποίος έχει ιδιοτιμές $\lambda_1, \dots, \lambda_k$, δηλαδή υπάρχει αντιστρέψιμος Q ώστε $Q^{-1}AQ = B$. Αν ο πίνακας έχει διαφορετικές ιδιοτιμές είναι διαγωνοποιήσιμος. Αν ο πίνακας έχει κάποιες ιδιοτιμές ίδιες, τότε μπορούμε να βρούμε μια ακολουθία από διαγώνιους πίνακες D_n με $\lim_{n\to\infty} D_n = \mathbf{0}_{n,n}$ ώστε οι πίνακες $B + D_n$ να έχουν διαφορετικές ιδιοτιμές για κάθε $n \in \mathbb{N}$ και συνεπώς είναι διαγωνοποιήσιμοι. Ισχύει ότι

$$\lim_{n\to\infty} Q(B + D_n)Q^{-1} = A,$$

δηλαδή το ζητούμενο. □

Παράδειγμα VI.3.27. Οι πίνακες

$$A_n = \begin{pmatrix} 1+\frac{1}{n} & 1 & 0 \\ 0 & 1+\frac{2}{n} & 0 \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = A.$$

Οι πίνακες A_n έχουν δύο διαφορετικές ιδιοτιμές και είναι διαγωνοποιήσιμοι, ενώ ο πίνακας A δεν είναι. Ο πίνακας D_n στην απόδειξη του θεωρήματος VI.3.26 είναι ο $\mathrm{diag}(1/n, 2/n, 0)$. Είναι ενδιαφέρον ότι μια βάση από ιδιοδιανύσματα είναι οι στήλες του πίνακα

$$Q = \begin{pmatrix} \frac{n^2}{2} & 1 & n \\ -\frac{n}{2} & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

οι οποίες δεν έχουν νόημα στο όριο $n \to \infty$ εκτός από το ιδιοδιάνυσμα $(1,0,0)^t$.

Παρατήρηση VI.3.28 (Αναλυτική απόδειξη του θεωρήματος Caley-Hamilton)**.** Παρατηρούμε ότι το Caley-Hamilton ισχύει για διαγωνοποιήσιμους πίνακες. Πράγματι, το χαρακτηριστικό πολυώνυμο ενός διαγωνοποιήσιμου πίνακα $A = Q\mathrm{diag}(\lambda_1, \dots, \lambda_n)Q^{-1}$ είναι

$$\mathrm{Ch}_A(x) = \prod_{i=1}^{n}(x - \lambda_i)$$

$$\mathrm{Ch}_A(A) = \prod_{i=1}^{n}(A - \lambda_i \mathbb{I}_n) = Q\prod_{i=1}^{n}\mathrm{diag}(\lambda_1 - \lambda_i, \dots, \lambda_n - \lambda_i)Q^{-1} = \mathbf{0}_{\mathbb{F}^{n,n}}.$$

Για την τελευταία ισότητα παρατηρήστε ότι έχουμε ένα γινόμενο $n \times n$ διαγώνιων πινάκων, όπου ο κάθε ένας έχει τουλάχιστον στην i θέση της διαγωνίου ένα μηδενικό στοιχείο.

Για ένα μη διαγωνοποιήσιμο πίνακα A, αν το $\mathbb{F} = \mathbb{R}$ ή $\mathbb{F} = \mathbb{C}$ μπορούμε να βρούμε μια ακολουθία διαγωνοποιήσιμων πινάκων $A_n \to A$. Το χαρακτηριστικό πολυώνυμο είναι μια συνεχής συνάρτηση $\mathbb{F}^{n,n} \to \mathbb{F}[x]$ οπότε

$$\mathbf{0}_{n,n} = \lim_{n\to\infty} \mathrm{Ch}_{A_n}(A_n) = \mathrm{Ch}_A(\lim_{n\to\infty} A_n) = \mathrm{Ch}_A(A),$$

από όπου έχουμε το ζητούμενο.

Άσκηση VI.3.29. Θεωρούμε ένα 6×6 πίνακα με ελάχιστο πολυώνυμο

$$m(x) = (x-1)^2(x-2)$$

και χαρακτηριστικό πολυώνυμο

$$f(x) = (x-1)^4(x-2)^2.$$

Να βρεθεί η πιθανή μορφή Jordan του πίνακα.

Λύση VI.3.30. Από την μορφή του ελαχίστου πολυωνύμου έχουμε ότι ο χώρος $\mathbb{F}^6$ διασπάτε σε ένα ευθύ άθροισμα δύο αναλλοίωτων υποχώρων $\mathbb{F}^6 = W_1 \oplus W_2$, όπου ο L_{W_1} έχει ελάχιστο πολυώνυμο $(x-2)$ αναγκαστικά χαρακτηριστικό $(x-2)^2$, ενώ ο L_{W_2} έχει ελάχιστο πολυώνυμο $(x-1)^2$ και αναγκαστικά χαρακτηριστικό $(x-1)^4$.

Η πιθανή μορφή Jordan είναι ένας μποκ διαγώνιος πίνακας της μορφής

$$\left(\begin{array}{c|c} A_1 & \mathbf{0}_{2,4} \\ \hline \mathbf{0}_{4,2} & A_2 \end{array}\right)$$

όπου οι πίνακες $A_1 \in \mathbb{F}^{2,2}$ και $A_2 \in \mathbb{F}^{4,4}$. Το ελάχιστο πολυώνυμο του A_1 είναι $(x-2)$ άρα ο πίνακας A_1 είναι διαγωνοποιήσιμος της μορφής $\mathrm{diag}(2,2)$. Από την άλλη το ελάχιστο πολυώνυμο του πίνακα A_2 είναι $(x-1)^2$ άρα ο πίνακας A_2 δεν μπορεί να είναι διαγωνοποίσιμος. Επίσης δεν μπορεί να έχει ως μπλοκ το $J(1,3)$ ούτε το $J(1,4)$ γιατί αυτά έχουν ελάχιστα πολυώνυμα $(x-1)^3$ και $(x-1)^4$ αντίστοιχα. Συνεπώς οι δυνατές μορφές για τον πίνακα είναι οι παρακάτω:

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{ή} \quad \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Άσκηση VI.3.31. Να βρεθεί η μορφή Jordan του πίνακα

$$\begin{pmatrix} 4 & -4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 2 & -3 \\ 0 & 0 & 1 & 3 & 1 \\ 0 & 0 & 8 & -3 & 8 \end{pmatrix}$$

Λύση VI.3.32. Ο παραπάνω πίνακας είναι μπλοκ διαγώνιος. Θα μελετήσουμε τα κομμάτια ξεχωριστά. Ο πίνακας

$$A_1 = \begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix}$$

έχει χαρακτηριστικό πολυώνυμο $(x-2)^2$. Αν δεν μας ενδιαφέρει να υπολογίσουμε τον πίνακα μετατροπής Q μπορούμε να επιχειρηματολογήσουμε ως εξής: Είτε ο πίνακας είναι διαγώνιος με ιδιοτιμή 2, οπότε το ελάχιστο του είναι το $x-2$. Όμως $A_1 - 2\mathbb{I}_2 \neq \mathbf{0}$ άρα ο πίνακας όμοιος με τον

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

Ο πίνακας

$$A_2 = \begin{pmatrix} -2 & 2 & -3 \\ 1 & 3 & 1 \\ 8 & -3 & 8 \end{pmatrix}$$

έχει χαρακτηριστικό πολυώνυμο $(3-x)^3$. Αν δεν μας ενδιαφέρει και πάλι ο υπολογισμός του πίνακα μετατροπής υπολογίζουμε:

$$B - 3\mathbb{I}_3 = \begin{pmatrix} -5 & 2 & -3 \\ 1 & 0 & 1 \\ 8 & -3 & 5 \end{pmatrix}, \quad (B - 3\mathbb{I}_2)^2 = \begin{pmatrix} 3 & -1 & 2 \\ 3 & -1 & 2 \\ -3 & 1 & -2 \end{pmatrix}, \quad (B - 3\mathbb{I}_3)^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Το συμπέρασμα είναι ότι το ελάχιστο πολυώνυμο είναι το $(x-3)^3$ και ότι ο αρχικός πίνακας είναι όμοιος με τον

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

VI.4 Εφαρμογές στις αναδρομικές ακολουθίες

Ας υποθέσουμε ότι έχουμε την αναδρομική ακολουθία

$$u_0 = 1, u_n = Au_0, A \in \mathbb{F}.$$

Είναι σαφές ότι ο γενικός όρος της ακολουθίας δίνεται από τον τύπο

$$u_n = A^n u_0 = A^n.$$

Την διαδικασία αυτή θέλουμε να την γενικεύσουμε. θα ξεκινήσουμε από ένα κλασικό πρόβλημα.

Η ακολουθία Fibonacci είναι μια αναδρομική ακολουθία η οποία ορίζεται ως εξής: θέτουμε $F_0 = F_1 = 1$ και

$$F_n = F_{n-1} + F_{n-2}$$

δηλαδή κάθε όρος της είναι το άθροισμα των δύο προηγούμενων. Παρατηρούμε ότι η παραπάνω αναδρομική σχέση μπορεί να εφραστεί με την βοήθεια του πινάκων:

$$\begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_{n-2} \\ F_{n-1} \end{pmatrix} \tag{VI.17}$$

Αν λοιπόν θέσουμε

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad u_n = \begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix}$$

τότε η ακολουθία Fibonacci δίνεται από τον τύπο

$$u_n = Au_{n-1}, \quad u_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

και συνεπώς ο γενικός όρος της δίνεται από τον τύπο

$$u_{n+1} = \begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = A^n u_1.$$

Συνεπώς χρειαζόμαστε έναν αποτελεσματικό τρόπο να υπολογίζουμε δυνάμεις του πίνακα A. Στο παράδειγμα VI.1.21 δείξαμε ότι

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = Q \begin{pmatrix} \frac{1-\sqrt{5}}{2} & 0 \\ 0 & \frac{1+\sqrt{5}}{2} \end{pmatrix} Q^{-1}, \text{ όπου } Q = \begin{pmatrix} \frac{-1-\sqrt{5}}{2} & \frac{-1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}.$$

Συνεπώς,

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = A^n u_0 = Q \begin{pmatrix} \left(\frac{1-\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1+\sqrt{5}}{2}\right)^n \end{pmatrix} Q^{-1} u_0 = \begin{pmatrix} \frac{2^{-n}\left((1+\sqrt{5})^n - (1-\sqrt{5})^n\right)}{\sqrt{5}} \\ \frac{2^{-2n-1}\left((\sqrt{5}-1)(2-2\sqrt{5})^n + 2^n(1+\sqrt{5})^{n+1}\right)}{\sqrt{5}} \end{pmatrix}$$

Από όπου προκύπτει ότι

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right).$$

Ο παραπάνω τύπος ανακατώνει ρητές δυνάμεις του άρρητου αριθμού $\sqrt{5}$ οι οποίες με κάποιο μαγικό τρόπο αλληλοαναιρούνται για να δώσουν ως τελικό αποτέλεσμα έναν ακέραιο.

Ασκήσεις με ορίζουσες, Binet κτλ

Άσκηση VI.4.1. 1. Είναι διαγονοποιήσιμος ο πίνακας

$$A = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix};$$

2. Θεωρούμε δύο ακολουθίες (u_n) και (v_n) που ορίζονται ως εξής: $u_0 = v_0 = 1$, $u_{n+1} = 6u_n - 2v_n$, $v_{n+1} = -2u_n + 9v_n$. Να υπολογιστούν τα u_n, v_n συναρτήσει του n.

Λύση VI.4.2. Υπολογίζουμε το χαρακτηριστικό πολυώνυμο του πίνακα το οποίο είναι $x^2 - 15x + 50 = (x-5)(x-10)$. Αφού το χαρακτηριστικό πολυώνυμο έχει δύο διαφορετικές ρίζες ο πίνακας A είναι διαγωνοποιήσιμος.

Παρατηρούμε ότι

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix} \begin{pmatrix} u_n \\ v_n \end{pmatrix}$$

από όπου προκύπτει ότι

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = A^n \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Θέλουμε να βρούμε ένα πίνακα Q ώστε $Q^{-1}AQ = \mathrm{diag}(5, 10)$. Για αυτό λύνουμε το ισοδύναμο σύστημα

$$AQ = Q\mathrm{diag}(5,10) \text{ ισοδύναμα } \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix},$$

από όπου προκύπτει

$$Q = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}, \text{ και στην συνέχεια } Q^{-1} = \frac{1}{5} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$

Το ζητούμενο αποτέλεσμα είναι το

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = Q \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} Q^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -5^{n+1}(2^n - 6) \\ 5^{n-1}(2^{n+1} + 3) \end{pmatrix},$$

VI.4.1 Ρητή Κανονική μορφή

Στην περίπτωση που ο διανυσματικός χώρος είναι ορισμένος υπέρ ενός αλγεβρικά κλειστού σώματος $\mathbb{F}$ η κανονική μορφή Jordan είναι το καλύτερο που μπορούμε να κάνουμε για να περιγράψουμε τις κλάσεις ομοιότητας πινάκων. Στην παράγραφο αυτή θα εξετάσουμε την ρητή κανονική μορφή η οποία δεν προϋποθέτει να είναι το σώμα $\mathbb{F}$ αλγεβρικά κλειστό. Στην πορεία θα επανεξετάσουμε από μια διαφορετική οπτική την θεωρία των πολυωνύμων γραμμικών συναρτήσεων.

Modules

Δίνεται μια γραμμική συνάρτηση $L : V \to V$. Μπορούμε να ορίσουμε μια συνάρτηση

$$\begin{aligned} \mathbb{F}[x] \times V &\longrightarrow V \\ (f, v) &\longmapsto f(L)v =: fv \end{aligned} \tag{VI.18}$$

Την παραπάνω συνάρτηση την θεωρούμε ως ένα «πολλαπλασιασμό» πολυωνύμων με στοιχεία του διανυσματικού χώρου V. Ο πολλαπλασιασμός αυτός φυσικά εξαρτάται από την γραμμική συνάρτηση V και ικανοποιεί τις ιδιότητες:

1. $(f + g)v = (f + g)(L)v = f(L)v + g(L)V = fv + gv$ για κάθε $f, g \in \mathbb{F}[x]$ και κάθε $v \in V$.
2. $f(v + w) = f(L)(v + w) = f(L)v + f(L)w = fv + fw$, για κάθε $f \in \mathbb{F}[x]$ και κάθε $v, w \in V$
3. $(fg)v = fg(L)v = f(L)g(L)v = f(gv)$, για κάθε $f, g \in \mathbb{F}[x]$ και κάθε $v \in V$.
4. $1v = \mathrm{Id}_V v = v$.

Η παρακάτω κατασκευή μας δίνει την ανάγκη να ορίσουμε την έννοια του module (το οποίο στην Ελληνική βιβλιογραφία αναφέρεται συχνά ως «πρότυπο»)

Ορισμός VI.4.3. Μία αβελιανή ομάδα $(V, +)$ θα είναι ένα R-module πάνω από τον αντιμεταθετικό δακτύλιο R, ο οποίος έχει μονάδα 1_R, αν υπάρχει μια συνάρτηση

$$\begin{aligned} R \times V &\longrightarrow V \\ (f, v) &\longmapsto f(L)v =: fv \end{aligned}$$

για την οποία να ισχύουν οι ιδιότητες:

1. $(f + g)v = fv + gv$ για κάθε $f, g \in R$ και κάθε $v \in V$.
2. $f(v + w) = fv + fw$, για κάθε $f \in R$ και κάθε $v, w \in V$
3. $(fg)v = f(gv)$, για κάθε $f, g \in R$ και κάθε $v \in V$.
4. $1_R v = v$.

Παρατήρηση VI.4.4. Κάθε αβελιανή ομάδα V αποκτά με φυσιολογικό τρόπο δομή $\mathbb{Z}$-module θέτοντας

$$nv = \begin{cases} \underbrace{v + \cdots + v}_{n-\text{φορες}} & \text{αν } n > 0 \\ 0 & \text{αν } n = 0 \\ \underbrace{(-v) + \cdots + (-v)}_{-n-\text{φορες}} & \text{αν } n < 0 \end{cases}$$

Στην πραγματικότητα η έννοια της αβελιανής ομάδας και του $\mathbb{Z}$-module ταυτίζονται.

Παρατήρηση VI.4.5. Η έννοια του R-module γενικεύει την έννοια του διανυσματικού χώρου. Πράγματι ένας διανυσματικός χώρος πάνω από το σώμα $\mathbb{F}$ είναι το ίδιο με ένα $\mathbb{F}$-module, όπου έχουμε απαιτήσει ότι ο δακτύλιος R να είναι το σώμα $\mathbb{F}$.

Όμως υπάρχουν και σοβαρές διαφορές. Πράγματι, σε ένα R-module μπορεί να έχουμε ότι $fv = 0$ με $f \in R$ $f \neq 0$ και $v \in V, v \neq 0$. Σκεφτείτε το $5(1 \mod 5) = 0$ στην ομάδα $V = \mathbb{Z}/5\mathbb{Z}$. Για τις ανάγκες της γραμμικής άλγεβρας έχουμε ότι αν $m \neq 0$ είναι το ελάχιστο πολυώνυμο του L, τότε $mv = m(L)v = \mathbf{0}_V$ για κάθε $v \in V$.

Ορισμός VI.4.6. Ένα υποσύνολο W ενός R-module θα λέγεται υποmodule αν και μόνο αν οι περιορισμοί των πράξεων της πρόσθεσης και του και του εξωτερικού πολλαπλασιασμού $R \times V \to V$ στο W, το εφοδιάζουν με την δομή ενός R-module.

Παρατήρηση VI.4.7. Αυτό που χρειάζεται να ελεγχθεί είναι κατά πόσο οι πράξεις είναι κλειστές στο W, δηλαδή

- $x - y \in W$ για κάθε $x, y \in W$
- $rx \in W$ για κάθε $r \in R$ και $x \in W$.

Παρατήρηση VI.4.8. Κάθε υποσύνολο I του R είναι R-module αν και μόνο αν είναι ιδεώδες του R.

Παρατήρηση VI.4.9. Στην περίπτωση που έχουμε την δράση του $\mathbb{F}[x]$ στον διανυσματικό χώρο V, μέσω της γραμμικής συνάρτησης $L : V \to V$, έχουμε ότι ένας υπόχωρος $W \subset V$ είναι R-υποmodule αν και μόνο αν είναι L-αναλλοίωτος.

Ορισμός VI.4.10. Μία συνάρτηση $f : V_1 \to V_2$ ανάμεσα σε R-modules θα λέγεται μορφισμός από R-modules αν διατηρεί τις πράξεις, δηλαδή αν

- $f(v_1 + v_2) = f(v_1) + f(v_2)$ για κάθε $v_1, v_2 \in V_1$
- $f(rv) = rf(v)$ για κάθε $v \in V_1$ και κάθε $r \in R$.

Ένας μορφισμός από R-modules που είναι 1-1 θα λέγεται μονομοριφμός από R-modules, ενώ ένας μορφισμός από R-modules που είναι επί θα λέγεται επιμορφισμός από R-modules. Τέλος ένας μορφισμός από R-modules που είναι και 1-1 και επί θα λέγεται ισομορφισμός από R-modules.

Παρατήρηση VI.4.11. Οι παραπάνω ορισμοί γενικεύουν την έννοια της γραμμικής συνάρτησης που είναι ένας μορφισμός από $\mathbb{F}$-modules.

Ορισμός VI.4.12. Ένα R-module V θα λέγεται ελεύθερο τάξης n, αν υπάρχουν στοιχεία $v_1, .., v_n \in V$, ώστε κάθε στοιχείο $v \in V$ να γράφεται με μοναδικό τρόπο ως

$$v = r_1v_1 + r_2v_2 + \cdots + r_nv_n, \text{ για } r_1, \ldots, r_n \in R.$$

Παρατήρηση VI.4.13. Ένα ελεύθερο R-module τάξης n είναι ισόμορφο με το R-module, R^n το οποίο αποτελείται από το σύνολο των n-άδων

$$R^n = (r_1, r_2, \ldots, r_n)$$

πρόσθεση κατά συντεταγμένες και εξωτερικό πολλαπλασιασμό $r(r_1, r_2, \ldots, r_n) = (rr_1, rr_2, \ldots, rr_n)$. Πράγματι, αρκεί να θεωρήσουμε την συνάρτηση

$$
\begin{aligned}
R^n &\longrightarrow V \\
(r_1, \ldots, r_n) &\longmapsto v = \sum_{i=1}^{n} r_i e_i
\end{aligned}
$$

και να τσεκάρουμε ότι είναι ισομορφισμός.

Παρατήρηση VI.4.14. Δύο ισόμορφα ελεύθερα R-modules, με R-αντιμεταθετικός δακτύλιος έχουν την ίδια τάξη. Δηλαδή το R^n δεν μπορεί να είναι ισόμορφο με το R^m για $n \neq m$. Ισχυριζόμαστε ότι κάθε αντιμεταθετικός δακτύλιος έχει ένα τουλάχιστον μέγιστο ιδεώδες, έστω $\mathfrak{m}$. Τότε έχουμε ότι $R^n/\mathfrak{m}R^n \cong k^n$ ενώ $R^m/\mathfrak{m}R^m = k^m$, όπου k είναι το σώμα $R/\mathfrak{m}$. Ένας ισομορφισμός ανάμεσα στο R^n και το R^m θα επάγει ένα ισομορφισμό ανάμεσα στους διανυσματικούς χώρους k^n και k^m για $n \neq m$, άτοπο.

Για να αποδείξουμε τώρα ότι ένας αντιμεταθετικός δακτύλιος R έχει ένα μέγιστο ιδεώδες, θα χρησιμοποιήσουμε το λήμμα του Zorn, I.4.19. Θα αποδείξουμε μάλιστα ότι κάθε ιδεώδες $\neq R$ περιέχεται σε μέγιστο ιδεώδες του R. Θεωρούμε το σύνολο Σ όλων των ιδεωδών του R τα οποία είναι διάφορα του R και περιέχουν το R. Το σύνολο R είναι επαγωγικό ως προς την διάταξη του εγκλεισμού, δηλαδή κάθε ολικά διαταγμένο υποσύνολο του έχει άνω φράγμα που ανήκει στο Σ. Πράγματι, αν $\{I_\nu\}_{\nu \in A}$ είναι ένα ολικά διατεταγμένο υποσύνολο του Σ τότε το $\cup_{\nu \in A} I_\nu$ είναι ένα ιδεώδες του R, το οποίο είναι διάφορο του R και αποτελεί ένα άνω φράγμα του $\{I_\nu\}_{\nu \in A}$. Έστω ένα μέγιστο στοιχείο $\mathfrak{m}$ του Σ, η ύπαρξη του οποίου εξασφαλίζεται από το λήμμα του Zorn. Αυτό είναι μέγιστο ιδεώδες, γιατί διαφορετικά θα υπήρχε $\mathfrak{m} \subset J \subsetneq R$, και είναι σαφές ότι $J \in \Sigma$, το οποίο είναι άτοπο με βάση τον ορισμό του $\mathfrak{m}$.

Παρατήρηση VI.4.15. Κάθε πεπερασμένα παραγόμενος διανυσματικός χώρος είναι ελεύθερο $\mathbb{F}$-module. Αυτό σημαίνει ότι έχει μια βάση. Στην περίπτωση των R-modules το αντίστοιχο θεώρημα δεν είναι σωστό. Έτσι το πεπερασμένο $\mathbb{Z}$-module $\mathbb{Z}/5\mathbb{Z}$ δεν είναι ελεύθερο, αφού δεν είναι ισόμορφο με το $\mathbb{Z}^n$ που έχει άπειρα το πλήθος στοιχεία.

Ορισμός VI.4.16. Ένα R-module V θα λέγεται torsion module αν για κάθε $v \in V$ υπάρχει $r \in R$, ώστε $rv = 0$.

Παράδειγμα VI.4.17. Κάθε πεπερασμένης διάστασης $\mathbb{F}$-διανυσματικός χώρος, εφοδιασμένος με την δομή ενός $\mathbb{F}[x]$-module μέσω μιας γραμμικής συνάρτησης $L : V \to V$ είναι ένα torsion module, αφού για το ελάχιστο πολυώνυμο m_L του L, έχουμε $m_L v = \mathbf{0}_V$.

Από εδώ και στο εξής ο δακτύλιος R θα είναι μια περιοχή κυρίων ιδεωδών. Το βασικό παράδειγμα το οποίο έχουμε στο μυαλό μας για τις ανάγκες της γραμμικής άλγεβρας είναι ο $R = \mathbb{F}[x]$.

Θεωρούμε το R-module V, και έστω $v \in V$. Θεωρούμε την συνάρτηση $\phi : R \to Rv$ η οποία στέλνει το $r \in R$ στο $rv \in Rv$. Θεωρούμε τον πυρήνα της συνάρτησης ϕ ο οποίος είναι ένα κύριο ιδεώδες το οποίο παράγεται από ένα στοιχείο $m_v \in R$. Το στοιχείο m_v θα το λέμε περίοδο του $v\ inV$. Ένα στοιχείο m θα λέγεται ένας εκθέτης για το V (αντίστοιχα για το $v \in V$) αν $mV = 0$ (αντίστοιχα $mv = 0$). Για παράδειγμα στην δράση του $\mathbb{F}[x]$ στον διανυσματικό χώρο V, μέσω της γραμμικής συνάρτησης $L : V \to V$, το ελάχιστο πολυώνυμο είναι ένας εκθέτης για το V. Για ένα $m \in R$ ορίζουμε το σύνολο

$$V_m = \{v \in V : mv = 0\},$$

δηλαδή τα στοιχεία του V που έχουν εκθέτη m. Ένα R-module Θα λέγεται κυκλικό αν είναι ισόμορφο με το $R/\langle a \rangle$, για κάποιο $a \in R$, $a \neq 0$. Το a θα το λέμε τάξη του κυκλικού module. Για ένα ανάγωγο στοιχείο $p \in R$, ορίζουμε

$$V(p) = \{v \in V : pv = 0\},$$

δηλαδή το υποmodule του V των στοιχείων του V που έχουν εκθέτη p.

Θεώρημα VI.4.18. *Έστω V ένα μη-μηδενικό πεπερασμένο torsion R-module. Τότε*

$$V = \bigoplus_p V(p), \tag{VI.19}$$

όπου το άθροισμα λαμβάνεται πάνω σε όλα τα ανάγωγα p με $V(p) \neq 0$. Κάθε $V(p)$ με την σειρά του γράφεται ως

$$V(p) = R/\langle p^{\nu_1}\rangle \bigoplus \cdots \bigoplus R/\langle p^{\nu_s}\rangle, \tag{VI.20}$$

με $1 \le \nu_1 \le \cdots \le \nu_s$, όπου οι ακολουθία $\nu_1, \dots, \nu_s$ είναι μονοσήμαντα ορισμένη.

Απόδειξη. Έστω m ένας εκθέτης του V τον οποίο αναλύουμε σε γινόμενο πρώτων μεταξύ τους ανάγωγων πολυωνύμων. Όπως ακριβώς στην απόδειξη της πρότασης VI.3.6 αποδεικνύουμε την αλήθεια της ανάλυσης στην εξίσωση (VI.19).

Θα προχωρήσουμε τώρα στην απόδειξη της (VI.20). Τα στοιχεία $v_1, \dots, v_n \in V$ θα λέγονται εξαρτημένα αν οποτεδήποτε έχουμε μια σχέση

$$r_1 v_1 + \cdots + r_n v_n = 0$$

με $r_1, \dots, r_n \in R$ ισχύει ότι $r_i v_i = 0$, για κάθε $1 \le i \le n$. Η συνθήκη αυτή είναι ασθενέστερη της γραμμικής ανεξαρτησίας. Είναι σαφές ότι τα στοιχεία $v_1, \dots, v_n$ είναι ανεξάρτητα αν και μόνο αν έχουμε την παρακάτω ανάλυση από R-modules:

$$\langle v_1, \dots, v_n \rangle = \langle v_1 \rangle \bigoplus \cdots \bigoplus \langle v_n \rangle,$$

όπου τα $\langle v_i \rangle = Rv_i \cong R/m_{v_i}R$, είναι κυκλικά modules.

Θεωρούμε ένα torsion module εκθέτη p^r, $r \ge 1$ για κάποιο ανάγωγο p. Έστω ένα στοιχείο v με περίοδο p^r. Θεωρούμε το πηλίκο $\overline{V} = V/\langle v \rangle$. Έστω $\overline{w}_1, \dots, \overline{w}_m$ ανεξάρτητα στοιχεία του $\overline{V}$. Θα δείξουμε ότι για κάθε ένα $\overline{w}_i$ υπάρχει ένας αντιπρόσωπος w_i της κλάσης modulo $\langle v \rangle$ ώστε ο εκθέτης του $\overline{w}_i$ να είναι ίδιος με τον εκθέτη του w_i.

Πράγματι, έστω $\overline{w}$ να έχει περίοδο p^n για κάποιο $n \ge 1$. Έστω w ένας αντιπρόσωπος της κλάσης το $\overline{w}$. Ισχύει ότι $p^n w \in \langle v \rangle$, συνεπώς $p^n w = p^s c v$, με $c \in R$, $(c, p) = 1$, $0 \le s \le r$. Αν $s = r$ τότε η $p^n w = 0$ και το w, έχει τον ίδιο εκθέτη όπως και το $\overline{w}$. Αν το $s < r$, τότε το $p^s c v$ έχει περίοδο p^{r-s} και συνεπώς το w έχει περίοδο p^{n+r-s}. Είναι σαφές ότι $n + r - s \le r$, αφού το p^r είναι εκθέτης του V. Συνεπώς $n \le s$ και το στοιχείο $w - p^{s-n} c v$ είναι ένας αντιπρόσωπος του $\overline{w}$ με εκθέτη n.

Διαλέγουμε αντιπροσώπους w_i των κλάσεων $\overline{w}_i$ με την ίδια περίοδο. Θα αποδείξουμε ότι τα $v, w_1, \dots, w_m$ είναι ανεξάρτητα. Έστω $r, r_1, \dots, r_m \in R$ ώστε

$$r v + r_1 w_1 + \cdots + r_m w_m = 0. \tag{VI.21}$$

Είναι σαφές ότι

$$r_1 \overline{w}_1 + \cdots r_m \overline{w}_m = 0$$

Από την υπόθεση της ανεξαρτησίας των $\overline{w}_1, \dots, \overline{w}_m$ έχουμε ότι $r_i \overline{w}_i = 0$ για όλα τα i. Αν p^{ν_i} είναι η περίοδος του $\overline{w}_i$ τότε p^{ν_i} διαιρεί το r_i. Τότε όμως $r_i w_i = 0$ για κάθε i. Επιστρέφοντας στην εξίσωση (VI.21) έχουμε ότι και $rv = 0$.

Διαλέγουμε το $v \in V(p)$ με τέτοιο τρόπο ώστε η περίοδος του p^r να έχει μέγιστο δυνατό r. Παρατηρούμε ότι το V_p, το οποίο είναι το σύνολο των στοιχείων με περίοδο p, γίνεται ένας $k = R/pR$-διανυσματικός χώρος. Έχουμε ότι $\dim_k \overline{V}_p < \dim_k V_p$, αφού μπορούμε στο $\langle v \rangle$ να βρούμε στοιχείο περιόδου p το οποίο να είναι γραμμικά ανεξάρτητο από μια βάση του $\overline{V}_p$.

Θα ολοκληρώσουμε την απόδειξη του θεωρήματος με επαγωγή. Αν το $\overline{V}$ δεν είναι μηδενικό, τότε υπάρχουν στοιχεία $\overline{v}_2, \dots, \overline{v}_s$ τα οποία έχουν αντίστοιχες περιόδους $p^{r_2}, \dots, p^{r_s}$ ώστε $r_2 \ge \cdots \ge$

r_s. Διαλέγουμε αντιπροσώπους $v_2, \dots, v_s$ στο V με τις ίδιες περιόδους ώστε τα $v_1, v_2, \dots, v_s$ να είναι ανεξάρτητα. Αφού το p^{r_1} έχει διαλεχτεί με τέτοιο τρόπο ώστε r_1 να είναι μέγιστο, η ανισότητα $r_1 \geq r_2 \geq \cdots$, είναι σαφής.

Η μοναδικότητα θα αποδειχθεί ως συνέπεια του γενικότερου θεωρήματος VI.4.19. □

Θεώρημα VI.4.19. *Έστω V ένα πεπερασμένα παραγόμενο μη-μηδενικό torsion R-module. Τότε το V είναι ισόμορφο με ένα ευθύ άθροισμα*

$$V = R/\langle q_1 \rangle \bigoplus \cdots \bigoplus R/\langle q_r \rangle, \tag{VI.22}$$

όπου $q_1, \dots, q_r$ είναι μη-μηδενικά στοιχεία του R, και $q_1 \mid q_2 \mid \cdots \mid q_r$. Η ακολουθία των ιδεωδών $\langle q_1 \rangle, \dots, \langle q_r \rangle$, είναι μονοσήμαντα ορισμένη από τις παραπάνω συνθήκες.

Απόδειξη. Με χρήση του θεωρήματος VI.4.18 διασπάμε το V ως

$$V = V(p_1) \bigoplus \cdots \bigoplus V(p_l),$$

όπου κάθε $V(p_i)$ διασπάται με την σειρά του ως ευθύ άθροισμα κυκλικών υποmodules περιόδων $p_i^{r_{ij}}$. Κρατάμε την πληροφορία των εκθετών στον παρακάτω πίνακα:

$$\begin{array}{c|c} V(p_1) & r_{11} \leq r_{12} \leq \cdots \\ V(p_2) & r_{21} \leq r_{22} \leq \cdots \\ \vdots & \vdots \\ V(p_l) & r_{l1} \leq r_{l2} \leq \cdots \end{array}$$

Θέτουμε τα $q_1, q_2, \dots, q_r$ να είναι τα γινόμενα κατά μήκος των παραπάνω στηλών, δηλαδή

$$q_1 = p_1^{r_{11}} p_2^{r_{21}} \cdots p_l^{r_{l1}}, \quad q_2 = p_1^{r_{12}} p_2^{r_{22}} \cdots, p_l^{r_{l2}}, \dots$$

Από το θεώρημα του Κινέζου να γραφεί στους δακτυλίους και να δοθεί αναφορά το $R/\langle q_1 \rangle$ είναι ισόμορφο με το ευθύ άθροισμα των modules κατά μήκος της πρώτης στήλης δηλαδή

$$R/\langle q_1 \rangle = R/\langle p_1^{r_{11}} \rangle \bigoplus R/\langle p_2^{r_{21}} \rangle \bigoplus \cdots \bigoplus R/\langle p_l^{r_{l1}} \rangle$$

ομοίως και για τα $R/\langle q_2 \rangle, \dots R/\langle q_r \rangle$. Με αυτό τον τρόπο αποδεικνύεται η εξίσωση (VI.22), ενώ η συνθήκη $q_1 \mid q_2 \mid \cdots \mid q_r$ είναι προφανής από τις ανισότητες των εκθετών.

Για την μοναδικότητα τώρα. Θεωρούμε ένα ανάγωγο στοιχείο p και θεωρούμε την περίπτωση που $V' = R/\langle pb \rangle$ για κάποιο $b \in R$, $b \neq 0$. Το V'_p είναι το υποmodule $bR/\langle pb \rangle$, όπως προκύπτει από την μοναδική παραγοντοποίηση του R. Θεωρούμε την συνάρτηση $V' \to pV'$ η οποία είναι επί και έχει πυρήνα V'_p. Συνεπώς

$$pV' = V'/V'_p = \frac{R/\langle pb \rangle}{bR/\langle pb \rangle} \cong R/bR.$$

Θεωρούμε τώρα το V να εκφράζεται όπως το θεώρημα VI.4.19 ως ένα ευθύ άθροισμα από r προσθετέους. Το τυχαίο στοιχείο $v \in V$ γράφεται

$$v = v_1 + v_2 + \cdots + v_r$$

και είναι στοιχείο του V_p αν και μόνο αν $pv_i = 0$ για όλα τα $i, 1 \leq i \leq r$. Δηλαδή το V_p είναι το ευθύ άθροισμα των πυρήνων του πολλαπλασιασμού με p σε κάθε προσθετέο. Όμως το V_p, αφού ο πολλαπλασιασμός με p είναι μηδενικός, είναι ένας διανυσματικός χώρος πάνω από το $k = R/\langle p \rangle$ και η διάστασή του είναι ίση με το πλήθος των $R/\langle q_i \rangle$ που το $p \mid q_i$.

Αν το p είναι ανάγωγο που διαιρεί το q_1 και συνεπώς όλα τα επόμενα q_i για $1 \le i \le r$. Υποθέτουμε ότι V έχει και μια δεύτερη ανάλυση

$$V = R/\langle q'_1 \rangle \bigoplus \cdots \bigoplus R/\langle q'_s \rangle,$$

η οποία ικανοποιεί τις συνθήκες του θεωρήματος. Τότε θα πρέπει το p να διαιρεί τουλάχιστον r από τα στοιχεία q'_j και συνεπώς $r \le s$. Για λόγους συμμετρίας έχουμε ότι $r = s$ και ότι το p διαιρεί όλα τα q'_j.

Θεωρούμε το module pV. Αν γράψουμε $q_i = pb_i$ τότε

$$pV \cong R/\langle b_1 \rangle \bigoplus \cdots \bigoplus R/\langle b_r \rangle,$$

και $b_1 \mid \cdots \mid b_r$. Κάποια από τα b_i μπορεί να είναι μονάδες. Αν όχι συνεχίζουμε να βγάζουμε έξω και άλλους πρώτους διαιρέτες του b_1 και συνεπώς όλως των b_i. Κάποια στιγμή θα έχουμε ότι

$$p_1 p_2 \cdots p_t V \cong R/\langle b'_1 \rangle \bigoplus \cdots R/\langle b'_r \rangle,$$

όπου $\langle b'_1 \rangle = \cdots = \langle b'_j \rangle = R$. Αν υπάρχει $\langle b_{j+1} \rangle \ne R$ τότε η μοναδικότητα προκύπτει με επαγωγή στο πλήθος των προσθεταίων r, αφού έχουμε καταλήξει σε ένα module με μικρότερο πλήθος προσθετέων. Διαφορετικά έχουμε και πάλι το μονοσήμαντο της ανάλυσης αφού τα $b_1, .., b_r$ είναι όλα ίσα με $p_1 p_2 \cdots p_t$ μέχρι πολλαπλασιασμό με αντιστρέψιμο στοιχείο, συνεπώς τα ιδεώδη που παράγουν είναι ίσα. □

Ρητή κανονική μορφή

Επιστρέφουμε και πάλι στην δράση του $\mathbb{F}[x]$ στο V όπως ορίστηκε στην (VI.18).

Ορισμός VI.4.20. Αν υπάρχει ένα στοιχείο $v \in V$ ώστε $V = \mathbb{F}[x]v$, τότε θα λέμε ότι το V είναι ένα κύριο $\mathbb{F}[x]$-module.

Ένα κύριο $\mathbb{F}[x]$-module V παράγεται ως $\mathbb{F}$-διανυσματικός χώρος από τα στοιχεία $v, xv = Lv, x^2 v = L^2 v, \ldots$. Θεωρούμε το ελάχιστο πολυώνυμο $m_L(x)$ του L. Είναι σαφές ότι τα

$$v, xv, x^2 v, \ldots, x^{n-1} v \tag{VI.23}$$

αποτελούν μια βάση του $V = \mathbb{F}[x]v$. Πράγματι, μια σχέση γραμμικής εξάρτησης θα έδινε ένα πολυώνυμο $g(x)$ μικρότερου βαθμού από το ελάχιστο πολυώνυμο ώστε $g(L)v = 0$. Αυτό θα έδινε ότι το $g(L) = \mathbf{0}$, αφού κάθε στοιχείο $w \in V$, είναι $w = h(x)v = h(L)v$ συνεπώς $g(L)w = g(L)h(L)v = h(L)g(L)v = \mathbf{0}$. Επίσης τα στοιχεία της εξίσωσης παράγουν τον V, αφού κάθε πολυώνυμο $f(x) \in \mathbb{F}[x]$, γράφεται ως $f(x) = q(x)m_L(x) + u(x)$, με $u(x) = 0$ ή $\deg u(x) \le n$, άρα

$$V \ni w = f(x)v = q(x)m_L(x)v + u(x)v = u(x)v.$$

Ο πίνακας της L ως προς την παραπάνω βάση είναι σαφές ότι είναι ο

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

όπου $m_L(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{F}[x]$.

Παρατήρηση VI.4.21. Το κύριο module V είναι ισόμορφο με το $\mathbb{F}[x]/\langle m_L(x) \rangle$, όπως παρατηρεί κανείς από τον μορφισμό $\mathbb{F}[x] \to \mathbb{F}[x]v = V$ με πυρήνα $\langle m_L(x) \rangle$.

Άμεση συνέπεια του θεωρήματος VI.4.19 είναι το

Θεώρημα VI.4.22. *Έστω V ένας μη-μηδενικός πεπερασμένης διάστασης διανυσματικός χώρος πάνω από το σώμα $\mathbb{F}$ και έστω $L : V \to V$ γραμμική συνάρτηση. Τότε ο V διασπάται ως ευθύ άθροισμα L-αναλλοίωτων υπόχωρων*

$$V = V_1 \bigoplus \cdots \bigoplus V_r,$$

όπου κάθε V_i είναι ένα κύριο $\mathbb{F}[x] = \mathbb{F}[L]$-module με ελάχιστο πολυώνυμο $q_i \neq 0$ ώστε $q_1 \mid q_2 \mid \cdots \mid q_r$. Η ακολουθία $\{q_1, \ldots, q_r\}$ είναι μονοσήμαντα ορισμένη.

Απόδειξη. Αυτή είναι μια άμεση μεταφορά του θεωρήματος VI.4.19 στο $\mathbb{F}[x]$-module V. □

Παρατήρηση VI.4.23. Η παραπάνω διάσπαση ονομάζεται ρητή κανονική μορφή του $L : V \to V$. Έχει το πλεονέκτημα ότι δεν απαιτεί από το σώμα $\mathbb{F}$ να είναι αλγεβρικά κλειστό.

Η κανονική μορφή Jordan μπορεί να προκύψει από την ρητή κανονική μορφή ως εξής:

Θεώρημα VI.4.24. *Αν $\mu_L(x) \in \mathbb{F}[x]$ είναι της μορφής $m_L(x) = (x - \lambda)^k$ τότε το κύριο module $V = \mathbb{F}[x]v$ έχει μια βάση B ως $\mathbb{F}$-διανυσματικός χώρος ώστε ο πίνακας*

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

Απόδειξη. Το V είναι ένα κύριο $\mathbb{F}[x]$-module με ελάχιστο πολυώνυμο $(x - \lambda)^k$ και ταυτίζεται με το $\mathbb{F}[z]v$ για κάποιο $v \in V$. Τα στοιχεία

$$v, (x - \lambda)v, (x - \lambda)^2 v, \ldots, (x - \lambda)^{k-1} v \tag{VI.24}$$

αποτελούν μια βάση του V. Είναι γραμμικά ανεξάρτητα, διότι κάθε γραμμική σχέση εξάρτησης οδηγεί σε ένα μη-μηδενικό πολυώνυμο που μηδενίζει τον L και είναι βαθμού μικρότερου του k.

Παρατηρούμε ότι αν θεωρήσουμε την διατεταγμένη βάση με την ανάποδη διάταξη από αυτή της (VI.24) δηλαδή

$$e_1 = (x - \lambda)^{k-1} v, e_2 = (x - \lambda)^{k-2} v, \ldots, e_j = (x - \lambda)^{k-j} v, \ldots, e_k = v$$

τότε ο πίνακας έχει ακριβώς την μορφή που θέλουμε. Πράγματι

$$x e_j = (x - \lambda) e_j + \lambda e_j = (x - \lambda)(x - \lambda)^{k-j} v + \lambda e_j = \lambda e_j + e_{j-1}.$$

Ας σημειωθεί ότι στον παραπάνω τύπο για λόγους ομοιομορφίας θέσαμε $e_0 = 0$. □