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 $(\lambda_1 - \lambda_2)v = \nvdash_V$ 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

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

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

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

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 $\{v_1,\ldots,v_{\rho(\lambda)}\}$ του ιδιόχωρου E_λ την οποία επεκτείνουμε σε μια βάση $B = \{v_1,\ldots,v_{\rho(\lambda)},w_1,\ldots,w_k\}$ 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 $v_1, \ldots, v_n$ ώστε $Av_i = \lambda_i v_i$, δηλαδή μια βάση αποτελούμενη από ιδιοδιανύσματα. Έστω $\lambda_{i_1}, \ldots, \lambda_{i_r}$ 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 $\Delta = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ ώστε

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

Αν γράψουμε $Q = (q^1, q^2, \ldots, q^n)$, όπου $q^1, \ldots, q^n$ 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 3 (Διαγωνοποίηση). *Για να διαγωνοποιήσουμε ένα $n \times n$ πίνακα εκτελούμε τα παρακάτω βήματα:*

1. *Υπολογισμός χαρακτηριστικού πολυωνύμου*
2. *Εύρεση ιδιοτιμών $\lambda_1, \ldots, \lambda_n$*
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 $1 \le i, j \le n$. 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 $1 \le \rho(n) \le \mu(n) = 1$. 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

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

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.

Άσκηση VI.1.44. Δείξτε ότι αν ο πίνακας A είναι αντιστρέψιμος και λ είναι μια ιδιοτιμή του A τότε λ^{-1} είναι μια ιδιοτιμή του A^{-1}.

Λύση VI.1.45. Αφού ο είναι αντιστρέψιμος καμία ιδιοτιμή του δεν είναι μηδενική. Πράγματι αν υπήρχε ιδιοτιμή $\lambda = 0$ τότε ο ιδιόχωρος της E_0 θα ήταν μη-τετριμμένος. Όμως ο ιδιόχωρος της μηδενικής ιδιοτιμής ταυτίζεται με τον πυρήνα του A, άτοπο. Έστω λ μια ιδιοτιμή του πίνακα A. Έχουμε ότι υπάρχει ιδιοδιάνυσμα $\nu \neq \mathbf{0}$, ώστε

$$A\nu = \lambda\nu \text{ συνεπώς } \nu = \lambda A^{-1}\nu \text{ δηλαδή } A^{-1}\nu = \lambda^{-1}\nu.$$

Άσκηση VI.1.46. Αν ο πίνακας $A \in \mathbb{F}^{n,n}$ είναι αντιστρέψιμος, είναι όμοιος με τον A^{-1} και ο n είναι περιτός τότε έχει μια ιδιοτιμή $\lambda = \pm 1$.

Λύση VI.1.47. Αφού οι πίνακες A, A^{-1} είναι όμοιοι έχουν το ίδιο χαρακτηριστικό πολυώνυμο. Συνεπώς αν $\lambda_1, \ldots, \lambda_n$ είναι οι ιδιοτιμές του A τότε κάθε ιδιοτιμή λ_i είναι ίση με μια ιδιοτιμή $\lambda_{j(i)}^{-1}$. Δεν είναι δυνατόν ένα πεπερασμένο σύνολο με περιττό αριθμό στοιχείων να το χωρίσουμε σε ζευγάρια οπότε για τουλάχιστον μία ιδιοτιμή ισχύει $\lambda = \lambda^{-1}$, συνεπώς $\lambda = \pm 1$.

Άσκηση VI.1.48. Δείξτε ότι κάθε πίνακας A με τάξη 1 έχει χαρακτηριστικό πολυώνυμο

$$\mathrm{Ch}_A(x) = (-x)^{n-1}(\mathrm{tr}(A) - x)$$

Στην συνέχεια υπολογίστε το χαρακτηριστικό πολυώνυμο του πίνακα

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

Λύση VI.1.49. Ένας πίνακας τάξης 1 έχει πυρήνα τάξης $n-1$ και αυτός είναι ο ιδιόχωρος της ιδιοτιμής 0. Επίσης κάθε πίνακας είναι όμοιος με ένα άνω τριγωνικό πίνακα ο οποίος έχει στην διαγώνιο $n-1$ μηδενικά στοιχεία και ένα στοιχείο λ. Αφού όμοιοι πίνακες έχουν το ίδιο ίχνος, έχουμε ότι $\mathrm{tr}(A) = \lambda$ και προκύπτει η πρώτη ισότητα σχετικά με το χαρακτηριστικό πολυώνυμο.

Παρατηρούμε ότι το χαρακτηριστικό πολυώνυμο του πίνακα $B = (b_{ij})$ με $b_{ij} = b$ είναι ίσο με $\mathrm{Ch}_B(x) = (-x)^{n-1}(nb - x)$. Επίσης παρατηρούμε ότι $A = B - (b-a)\mathbb{I}_n$. Συνεπώς:

$$\det(A - x\mathbb{I}_n) = \det(- (b-a)\mathbb{I}_n - x\mathbb{I}_n).$$

Αν θέσουμε $y = b - a + x$ η παραπάνω παράσταση είναι ίση με

$$\mathrm{Ch}_A(x) = (-y)^{n-1}(nb - y) = (a - b - x)^{n-1}(nb - b + a - x).$$

Άσκηση VI.1.50.

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

$$f(A) = a_0\mathbb{I}_m + a_1 A + a_2 A^2 + \cdots + a_n A^n.$$

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

$$\mathrm{Ch}_A(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0.$$

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:

$$\mathrm{adj}(A - x\mathbb{I}_n) = B_0 + B_1x + \cdots + B_{n-1}x^{n-1}, 0, 1, \ldots, B_{n-1} \in \mathbb{F}^{n,n}.$$

Συνεπώς έχουμε

$$(A - x\mathbb{I}_n)(B_0 + B_1x + \cdots + B_{n-1}x^{n-1}) = (x^n + a_{n-1}x^{n-1} + \cdots + a_0)\mathbb{I}_n.$$

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

$$0 = a_0\mathbb{I}_n + a_1A + \cdots + a_{n-1}A^{n-1} + A^n,$$

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 έκφραση βαθμού* $n-1$ *στον* A.

Απόδειξη. Έχουμε ότι

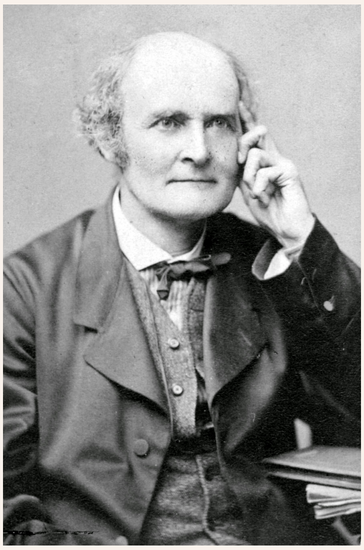
$$\mathrm{Ch}_A(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1A + (-1)^n\det(A)\mathbb{I}_n = 0$$

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 $z_1, \ldots, z_n$ 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

$$f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n \in \mathbb{F}[x]$$

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

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

Δηλαδή, το $(1, \lambda, \lambda^2, \ldots, \lambda^{n-1})^t$ 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 $\lambda_1, \ldots, \lambda_n$ 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 $z_1, \ldots, z_n$. *Σύμφωνα με την πρόταση I.9.32 οι συντελεστές του πολυωνύμου* a_i *δίνονται από*

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

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

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

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

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

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

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

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

$$\operatorname{tr}\left(C^k + a_{n-1}C^{k-1}) + \cdots + a_0 C^{k-n}\right) = 0 \Leftrightarrow \operatorname{tr}(C^{k-n} f(C)) = 0$$

η οποία είναι άμεση συνέπεια του θεωρήματος Caley-Hamilton.

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

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

την οποία την γράφουμε στην ισοδύναμη μορφή:

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

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

$$f(X) = (X - C)\left(X^{n-1} + (C + a_{n-1}\mathbb{I}_n)X^{n-2} + (C^2 + a_{n-1}C + a_{n-2}\mathbb{I}_n)X^{n-3} + \cdots + (C^{n-1} + a_{n-1}C^{n-2} + \cdots + a_1\mathbb{I}_n)\mathbb{I}_n\right).$$

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*

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

Στην συνέχεια εφαρμόζουμε τον τύπο του ίχνους για να πάρουμε

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

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

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

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

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

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 $f(x) = (x-z_1)(x-z_2)\cdots(x-z_n)$.

Άσκηση VI.2.14. Δίνονται δύο γραμμικές συναρτήσεις $L, G : V \to V$ με ελάχιστα πολυώνυμα $p_L(x), p_G(x)$. Αν ο μέγιστος κοινός διαιρέτης $f(x)p_L(x) + g(x)p_G(x) = 1$ δείξτε ότι η $p_L(G) : V \to V$ είναι ισομορφισμός. Επίσης δείξτε ότι αν $\ker G \neq \{\mathbf{0}\}$ τότε η L είναι ισομορφισμός.

Λύση VI.2.15. Αφού $(p_L(x), p_G(x)) = 1$ υπάρχουν πολυώνυμα $f(x), g(x) \in \mathbb{F}[x]$ ώστε

$$f(x)p_L(x) + g(x)p_G(x) = 1 \text{ συνεπώς } f(G)p_L(G) = \mathrm{Id}_V,$$

από όπου προκύπτει ότι ο $f(G)$ είναι ο αντίστροφος της $p_L(G)$.

Υποθέτουμε τώρα ότι υπάρχει μη μηδενικό $v \in V$ ώστε $Gv = \mathbf{0}$. Γράφουμε

$$p_L(x) = a_0 + a_1x + \cdots + a_nx^n$$

και έχουμε ότι αφού $p_L(G)$ είναι ισομορφισμός

$$\mathbf{0} \neq p_L(G)v = a_0v = a_0v + a_1Gv + a_2G^2v + \cdots + a_nG^nv.$$

Συνεπώς $a_0 \neq 0$. Αν $w \in \operatorname{Ker} L$, τότε αφού $p_L(L) = \mathbf{0}$ έχουμε

$$a_0w = -(a_1L + a_2L^2 + \cdots a_nL^n)w = \mathbf{0}$$

και αφού $a_0 \neq 0$ έχουμε ότι $w = \mathbf{0}$. Συνεπώς η L είναι 1-1 $L : V \to V$, άρα είναι και επί.

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 $1 \leq i \leq s$ τότε μπορούμε να διαλέξουμε βάσεις B_i των χώρων V_i για $1 \leq i \leq s$ 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*

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

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

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

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 $m_{\mathrm{diag}(\lambda,\dots,\lambda)}(x) = x - \lambda$ 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 $1 \leq i \leq s$. Συνεπώς, έχουμε ότι $m_{A_i}(x) \mid m_A(x)$ για κάθε $1 \leq i \leq s$ 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 $(L, B, B) = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Αν $\lambda_{i_1}, \dots, \lambda_{i_s}$ 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

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

όπου $\lambda_1, \dots, \lambda_s$ είναι οι ανά δύο διαφορετικές ιδιοτιμές της L_1. Σύμφωνα με το λήμμα VI.3.4 οι ιδιόχωροι $\operatorname{Ker}(L_1 - \lambda_i \mathrm{Id}_V)$ είναι L_2-αναλλοίωτοι, και από το λήμμα VI.3.8 οι $L_2|_{\operatorname{Ker}(L_1 - \lambda_i \mathrm{Id}_V)}$ είναι διαγωνοποιήσιμη. Αυτό σημαίνει ότι ο ιδιόχωρος $\operatorname{Ker}(L_1 - \lambda_i \mathrm{Id}_V)$ 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 $w_1, \dots, w_s$ αποτελούμενη από ιδιοδιανύσματα την οποία την συμπληρώνουμε σε μία βάση από ιδιοδιανύσματα $w'_1, \dots, w'_r$ του V. Αυτό μπορεί να γίνει με την βοήθεια της πρότασης IV.3.23. Το ζητούμενο αναλλοίωτο συμπλήρωμα είναι το $W' = \langle w'_1, \dots, w'_r \rangle$.

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 $B_W = \{w_1, \dots, w_n\}$ του W την οποία την επεκτείνουμε σε βάση $B_V = \{w_1, \dots, w_n, v_1, \dots, v_s\}$ 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 $B_{V/W} = \{v_1 + W, \dots, v_s + W\}$ 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

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

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*

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

Απόδειξη. Το χαρακτηριστικό πολυώνυμο προκύπτει από τον υπολογισμό της ορίζουσας

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

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 $1 \leq i \leq s$ *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 $i \geq 1$ 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 $\mathrm{Im}T$. 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 $\mathrm{Im}T$.

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 $\mathrm{Ker}T$, 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

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

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

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

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 $s_1, \ldots, s_r$ 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 $\lambda_1, \ldots, \lambda_r$ είναι οι διαφορετικές ιδιοτιμές του L. Από το λήμμα VI.3.23 διαλέγουμε βάσεις για τους υπόχωρους $\mathrm{Ker}(L - \lambda_1 \mathrm{Id}_V), \ldots, \mathrm{Ker}(L - \lambda_s \mathrm{Id}_V)$ 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 4 (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*

$$(A - \lambda\mathbb{I}_n)w_3 = w_2, (A - \lambda\mathbb{I}_n)w_4 = w_3, \ldots, (A - \lambda\mathbb{I}_n)w_n = w_{n-1},$$

   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* $v, w_1, \ldots, w_n$ *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 5 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,

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

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 υπomodule 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-υπomodule 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

$$v = r_1 v_1 + r_2 v_2 + \cdots + r_n v_n, \text{ για } r_1, \ldots, r_n \in R.$$

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 υπomodule 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 $1 \leq \nu_1 \leq \cdots \leq \nu_s$, *όπου οι ακολουθία* $\nu_1, \ldots, \nu_s$ *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 $v_1, \ldots, v_n \in V$ θα λέγονται εξαρτημένα αν οποτεδήποτε έχουμε μια σχέση

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

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

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

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

Θεωρούμε ένα torsion module εκθέτη p^r, $r \geq 1$ για κάποιο ανάγωγο p. Έστω ένα στοιχείο v με περίοδο p^r. Θεωρούμε το πηλίκο $\overline{V} = V/\langle v\rangle$. Έστω $\overline{w}_1, \ldots, \overline{w}_m$ 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 $n \geq 1$. 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 $c \in R$, $(c, p) = 1$, $0 \leq s \leq r$. 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 $n + r - s \leq r$, αφού το p^r είναι εκθέτης του V. Συνεπώς $n \leq s$ 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 $v, w_1, \ldots, w_m$ είναι ανεξάρτητα. Έστω $r, r_1, \ldots, r_m \in R$ ώστε

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

Είναι σαφές ότι

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

Από την υπόθεση της ανεξαρτησίας των $\overline{w}_1, \ldots, \overline{w}_m$ 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 $\overline{v}_2, \ldots, \overline{v}_s$ τα οποία έχουν αντίστοιχες περιόδους $p^{r_2}, \ldots, p^{r_s}$ ώστε $r_2 \geq \cdots \geq r_s$. Διαλέγουμε αντιπροσώπους $v_2, \ldots, v_s$ στο V με τις ίδιες περιόδους ώστε τα $v_1, v_2, \ldots, v_s$ 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 υπομodules περιόδων $p_i^{r_{ij}}$. Κρατάμε την πληροφορία των εκθετών στον παρακάτω πίνακα:

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

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 υπομodule $bR/\langle pb\rangle$, 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 $i, 1 \le i \le r$. 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 $\langle b_{j+1} \rangle \neq R$ 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 $\deg u(x) \leq n$, 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 $e_0 = 0$. □

VII

Χώροι με εσωτερικό γινόμενο

VII.1 Διανυσματικοί χώροι με νόρμα

Ορισμός VII.1.1. Ένας διανυσματικός χώρος με νόρμα είναι ένας διανυσματικός χώρος V πάνω από το σώμα $\mathbb{F}$ των πραγματικών ή μιγαδικών αριθμών, συνοδευόμενος με μια νόρμα ή απόλυτη τιμή, συμβολίζεται με $\|\cdot\| : V \to \mathbb{F}$ και ικανοποιεί τις παρακάτω ιδιότητες για κάθε διάνυσμα $u, v \in V$ και κάθε $c \in \mathbb{F}$:

1. Θετικά ορισμένη: $\|u\| \geq 0$, και $\|u\| = 0$ αν και μόνο αν $u = \mathbf{0}$.
2. Ομογένεια: $\|cu\| = |c|\|u\|$
3. Τριγωνική Ανισότητα: $\|u + v\| \leq \|u\| + \|v\|$

Παρατήρηση VII.1.2. Στην δεύτερη σχέση το $|c|$ είναι η απόλυτη τιμή ενός στοιχείου του $\mathbb{R}$, $c \in \mathbb{R}$ ή είναι το μέτρο ενός μιγαδικού αριθμού $|c| = \sqrt{\mathrm{Re}^2 c + \mathrm{Im}^c}$ αν $c \in \mathbb{C}$.

Παρατήρηση VII.1.3. Είναι σαφές από τα αξιώματα της νόρμας ότι $\|v\| \geq 0$ για κάθε $v \in V$. Πράγματι,

$$0 = \|v - v\| \leq \|v\| + \|-v\| = 2\|v\|.$$

Παρατήρηση VII.1.4. Ισχύει

$$\big|\|v\| - \|u\|\big| \leq \|v - u\| \text{ για κάθε } v, u \in V.$$

Πράγματι,

$$\|u\| = \|(u - v) + v\| \leq \|u - v\| + \|v\|,$$

δηλαδή $\|u\| - \|v\| \leq \|u - v\|$ και αλλάζοντας τους ρόλους των u, v έχουμε $\|v\| - \|u\| \leq \|v - u\| = \|u - v\|$ δηλαδή

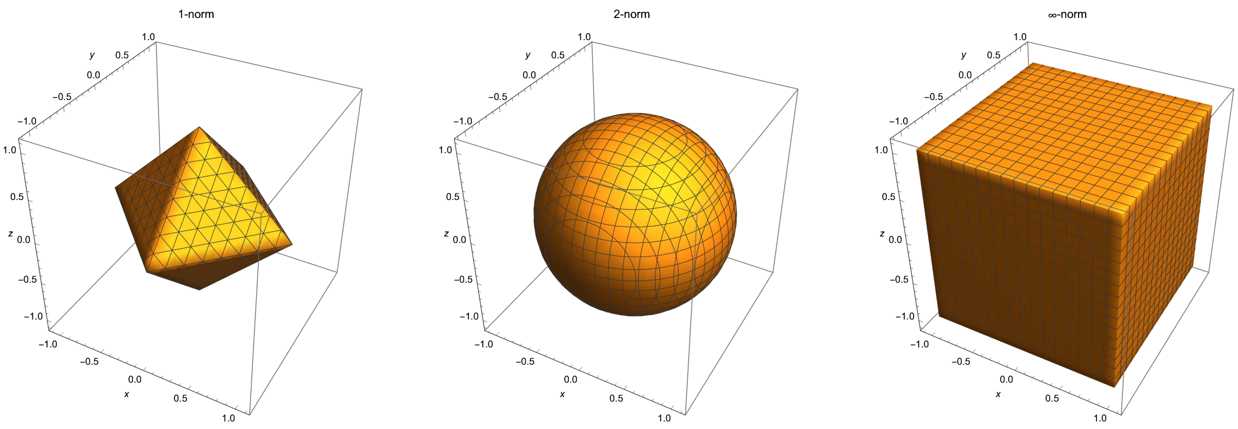
$$-\|u\| + \|v\| \leq \|u - v\| \leq \|u\| - \|v\|,$$

το οποίο είναι και το ζητούμενο.

Παραδείγματα VII.1.5.

1. Ο χώρος $(\mathbb{R}, \|\cdot\|)$ με $\|x\| = |x|$ για κάθε $x \in \mathbb{R}$.
2. Ο χώρος $(\mathbb{C}, \|\cdot\|)$ με $\|x\| = |x|$ για κάθε $x \in \mathbb{C}$.
3. Ο χώρος $(\mathbb{R}^n, \|\cdot\|_1)$ με

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \text{ όπου } x = (x_1, \ldots, x_n).$$



Σχήμα VII.1: Οι μοναδιαίες μπάλες $B = x : \|x\| \leq 1$ στον χώρο $\mathbb{R}^3$ για $\|\cdot\| = \|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$.

4. Ο χώρος $(\mathbb{R}^n, \|\cdot\|_2)$ με

$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}, \text{ όπου } x = (x_1, \ldots, x_n).$$

5. Ο χώρος $(\mathbb{R}^n, \|\cdot\|_\infty)$ με

$$\|x\|_\infty = \max_{1\leq i\leq n} |x_i|, \text{ όπου } x = (x_1, \ldots, x_n).$$

6. Ο χώρος $(C[a,b], \|\cdot\|_\infty)$ των συνεχών συναρτήσεων με νόρμα

$$\|f\| = \max_{a\leq x\leq b} |f(x)|.$$

Στο παραπάνω είναι σημαντικό η f να είναι συνεχής σε κλειστό διάστημα ώστε να έχει μέγιστο.

Παρατήρηση VII.1.6. Σε ένα διανυσματικό χώρο με εσωτερικό γινόμενο μπορούμε να ορίσουμε την έννοια της σύγκλισης μιας ακολουθίας (v_n) στοιχείων $v_n \in V$ στο $v \in V$ ως εξής:
Το όριο $\lim_{n->\infty} v_n = v$ αν και μόνο αν

$$\text{Για κάθε } \epsilon > 0 \text{ υπάρχει } n_0 \in \mathbb{N}, \text{ ώστε για } n > n_0 \text{ ισχύει } \|v_n - v\| < \epsilon.$$

Οι ιδιότητες του ορίου τέτοιων ακολουθιών είναι παρόμοιες με αυτές που ισχύουν για τα όρια ακολουθιών πραγματικών αριθμών γιατί ισχύουν οι ιδιότητες της νόρμας που είναι παρόμοιες με αυτές της απόλυτης τιμής, την οποία και γενικεύουν.

Έστω ένας διανυσματικός χώρος V. Επιλέγοντας μια βάση B έχουμε ένα ισομορφισμό $[\cdot]_B : V \to \mathbb{R}^n$, οπότε μπορούμε να ορίσουμε μια νόρμα ως εξής:

$$\|v\|_V = \|[v]_B\|_{\mathbb{R}^n}.$$

Δηλαδή μετράμε με μια νόρμα του $\mathbb{R}^n$ τις συντεταγμένες του διανύσματος. Η νόρμα αυτή εξαρτάται από την επιλογή της βάσης. Θα ορίσουμε μια έννοια ισοδυναμίας νορμών και θα δείξουμε στην πρόταση VII.1.11 ότι η εξάρτηση αυτή από την βάση δεν πειράζει και πολύ αφού όλες οι νόρμες σε έναν πεπερασμένης διάστασης διανυσματικό χώρο είναι ισοδύναμες.

Ορισμός VII.1.7. Θα λέμε ότι οι νόρμες $\|\cdot\| : V \to \mathbb{R}$ και $\|\cdot\|' : V \to \mathbb{R}$ είναι ισοδύναμες αν υπάρχουν σταθερές $A > 0, B > 0$ ώστε

$$A\|v\|' \leq \|v\| \leq B\|v\|' \text{ για κάθε } v \in V.$$

Παρατήρηση VII.1.8. Η παραπάνω σχέση είναι σχέση ισοδυναμίας, δηλαδή

- Κάθε νόρμα είναι ισοδύναμη με τον εαυτό της. Πράγματι, αρκεί να διαλέξουμε $A, B = 1$.
- Αν η νόρμα $\|\cdot\|$ είναι ισοδύναμη με την $\|\cdot\|'$ δηλαδή υπάρχουν A, B ώστε να ισχύουν
$$A\|v\|' \leq \|v\| \leq B\|v\|' \Rightarrow B^{-1}\|v\| \leq \|v\|' \leq A^{-1}\|v\|$$
- Αν η $\|\cdot\|$ είναι ισοδύναμη με την $\|\cdot\|'$ είναι ισοδύναμη με την $\|\cdot\|''$ δηλαδή υπάρχουν A_1, A_2 και B_1, B_2 ώστε
$$A_1\|v\|' \leq \|v\| \leq B_1\|v\|'$$
$$A_2\|v\|'' \leq \|v\|' \leq B_2\|v\|''$$
τότε
$$A_1A_2\|v\|'' \leq A_1\|v\|' \leq \|v\| \leq B_1\|v\|' \leq B_1B_2\|v\|$$
δηλαδή και $\|\cdot\|$ είναι ισοδύναμη με την $\|\cdot\|''$.

Η έννοια της νόρμας σε ένα χώρο μας δίνει την δυνατότητα να ορίσουμε όρια.

Ορισμός VII.1.9. Θα λέμε ότι η ακολουθία (a_m) στοιχείων του V συγκλίνει στο $a \in V$ αν για κάθε $\epsilon > 0$ υπάρχει $n_0 \in \mathbb{N}$ ώστε για κάθε $n \geq n_0$ να ισχύει $\|a_n - a\| < \epsilon$.

Θα λέμε ότι η ακολουθία (a_m) στοιχείων του V είναι Cauchy αν για κάθε $\epsilon > 0$ υπάρχει $n_0 \in \mathbb{N}$ ώστε για κάθε $m, \ell \geq n_0$ να ισχύει $\|a_m - a_\ell\| < \epsilon$.

Τέλος ο χώρος V θα λέγεται πλήρης αν και μόνο αν κάθε ακολουθία Cauchy σε αυτόν έχει όριο στον V.

Παρατήρηση VII.1.10. Αν οι μετρικές $\|\cdot\|$ και $\|\cdot\|'$ είναι ισοδύναμες τότε αν το όριο της (a_n) ως προς την $\|\cdot\|$ υπάρχει και είναι $a \in V$ τότε το όριο υπάρχει και ως προς την $\|\cdot\|'$ και είναι επίσης a.

Πρόταση VII.1.11. *Όλες οι νόρμες σε ένα πεπερασμένης διάστασης διανυσματικό χώρο πάνω από ένα πλήρες σώμα είναι ισοδύναμες.*

Απόδειξη. Θεωρούμε ένα πλήρες σώμα $\mathbb{F}$ ως προς την νόρμα $|\cdot|$. Θα κάνουμε επαγωγή ως προς την διάσταση n του V. Η περίπτωση $\{\mathbf{0}\}$ είναι τετριμμένη. Ας υποθέσουμε ότι $\dim V = 1$ και έστω $v_0 \neq \mathbf{0}$ ένα στοιχείο του V, οπότε $V = \mathbb{F}v_0$, οπότε $\|cv_0\| = |c|\|v_0\|$. Συνεπώς δύο νόρμες $\|\cdot\|$, $\|\cdot\|'$ στον μονοδιάστατο χώρο V είναι η μία πολλαπλάσιο της άλλης.

Υποθέτουμε ότι $n \geq 2$ και ότι η πρόταση ισχύει για όλους τους διανυσματικούς χώρους διάστασης $n-1$. Έστω V διανυσματικός χώρος διάστασης n. Διαλέγουμε μια $\mathbb{F}$ βάση $\{v_1, \ldots, v_n\}$ του V. Για μια οποιαδήποτε νόρμα $\|\cdot\|$ του V θα δείξουμε ότι είναι ισοδύναμη με την νόρμα $\|\cdot\|_\infty$ η οποία ορίζεται ως

$$\left\|\sum_{i=1}^{n} c_i v_i\right\| = \max_{1 \leq i \leq n} |c_i|.$$

Παρατηρούμε ότι ο V ως προς την $\|\cdot\|_\infty$ είναι πλήρης διότι αν έχουμε μια ακολουθία Cauchy (w_n) στον V ως προς την $\|\cdot\|_\infty$ αυτό σημαίνει ότι για κάθε $\epsilon > 0$ υπάρχει $n_0 \in \mathbb{N}$, ώστε για $\ell, m > n_0$ έχουμε $\|a_\ell - a_m\| < \epsilon$. Γράφοντας τα a_m ως προς την βάση $v_1, \ldots, v_n$ έχουμε ότι

$$a_m = c_{1m}v_1 + \cdots + c_{nm}v_n$$

οπότε ο ορισμός της $\|\cdot\|_\infty$ δίνει ότι οι οι ακολουθίες $(c_{im})_{m\in\mathbb{N}}$ είναι Cauchy ως προς την $|\cdot|$, και λόγω πληρότητας υπάρχει το όριο $c_i = \lim_{m\to\infty} c_{im}$. Τότε το $a = \sum c_i v_i$ είναι το όριο της (a_m) στον V:

$$\|a - a_m\|_\infty = \max_{1\le i\le n} |c_i - c_{im}| \to 0.$$

Θα πρέπει να βρούμε σταθερές ώστε

$$A\|v\|_\infty \le \|v\| \le B\|v\|_\infty.$$

Το τυχαίο $v \in V$ το γράφουμε $v = c_1v_1 + \cdots + c_nv_n$. Έχουμε

$$\|v\| \le \sum_{i=1}^n \|c_iv_i\| = \sum_{i=1}^n |c_i|\|v_i\| \le B \max |c_i| = B\|v\|_\infty,$$

όπου $B = \sum_{\|v\|} > 0$.

Θα δείξουμε τώρα ότι υπάρχει $A > 0$ ώστε $A\|v\|_\infty \le \|v\|$. Ας υποθέσουμε ότι αυτό δεν είναι σωστό για κανένα A, οπότε για κάθε $A = 1/k$, $k = 1, 2, \ldots$ υπάρχει ένα $u_k \in V$, ώστε

$$\|u_k\| < (1/k)\|u_k\|_\infty.$$

Από τον ορισμό του $\|\cdot\|_\infty$ σε όρους της βάσης $v_1, \ldots, v_n$ έχουμε ότι το $\|u_k\|_\infty$ είναι η απόλυτη τιμή κάποιας συντεταγμένης του u_k. Αφού το σύνολο u_k είναι άπειρο και υπάρχουν πεπερασμένα το πλήθος στοιχεία της βάσης $v_1, \ldots, v_n$, υπάρχει κάποιος δείκτης i, $1 \le i \le n$ για τον οποίο ισχύει ότι άπειρα u_k έχουν $\|u_k\|$ ίσο με την συντεταγμένη που αντιστοιχεί στο v_i στοιχείο της βάσης. Χωρίς περιορισμό της γενικότητας μπορούμε να υποθέσουμε ότι $i = n$. Παρατηρούμε ότι αν πολλαπλασιάσουμε το u_k με κάποιο μη μηδενικό στοιχείο του $\mathbb{F}$ η ανισότητα $\|u_k\| < (1/k)\|u_k\|_\infty$ παραμένη σε ισχύ, αφού και τα δύο μέλη της πολλαπλασιάζονται με το ίδιο στοιχείο.

Έτσι μπορούμε, αφού περάσουμε σε κάποια υποακολουθία $\{u_{k_j}\}$ να υποθέσουμε ότι

1. $\|u'_{k_j}\|_\infty = 1$,
2. Η συντεταγμένη του v_n του u'_{k_j} να είναι ίση με 1.
3. $\|u'_{k_j}\| < (1/k_j)\|v'_{k_k}\|_\infty = 1/k_j$.

Αφού το $k_j \to \infty$ όταν $j \to \infty$ έχουμε ότι $\|u'_{k_j}\| \to 0$, όταν $j \to \infty$. Θέτουμε $w_j = u'_{k_j} - v_n$, οπότε w_j ανήκει στον χώρο $W = \langle v_1, \ldots, v_n\rangle$ και έχουμε

$$\|w_j + v_n\| = \|u'_{k_j}\| \to 0 \tag{VII.1}$$

Θεωρούμε τις διαφορές

$$\|w_j - w_{j'}\| = \|(w_j + v_n) - (w_{j'} + v_n)\| \le \|(w_j + v_n)\| - \|(w_{j'} + v_n)\| \to 0. \tag{VII.2}$$

Συνεπώς η ακολουθία w_j είναι ακολουθία Cauchy στον W ως προς την $\|\cdot\|$ περιορισμένη στον W.

Αφού ο χώρος W έχει διάσταση $n-1$ από την επαγωγική υπόθεση όλες οι νόρμες είναι ισοδύναμες σε αυτόν, συνεπώς $\|\cdot\|_\infty \le C\|\cdot\|$ και $\|\cdot\| \le C'\|\cdot\|_\infty$ στον W για κάποια $C > 0, C' > 0$. Από την (VII.2) έχουμε ότι

$$\|w_j - w_{j'}\|_\infty \le C\|w_j - w_{j'}\| \to 0.$$

Συνεπώς η ακολουθία w_j είναι ακολουθία Cauchy στον W ως προς την $\|\cdot\|_\infty$ περιορισμένη στον W. Ο χώρος $(W, \|\cdot\|_\infty)$ είναι πλήρης οπότε υπάρχει ένα όριο w για τα w_j ως προς την $\|\cdot\|_\infty$, δηλαδή

$$\|w - w_j\|_\infty \to 0.$$

Τότε $\|w - w_j\| \leq C'\|w - w_j\|_\infty - > 0$ συνεπώς $\|w - w_j\| \to 0$. Σε συνδυασμό με την (VII.1) έχουμε ότι

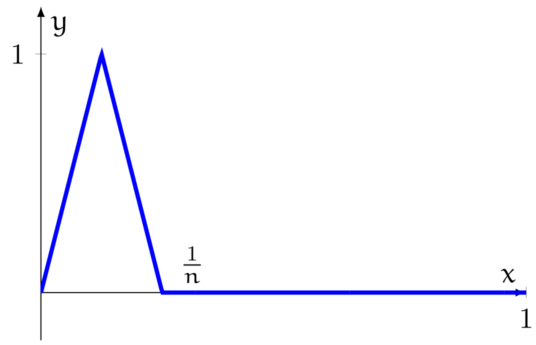
$$\|w + v_n\| = \|(w - w_j) + (w_j + v_n)\| \leq \|w - w_j\| + \|w_j + v_n\| \to 0$$

οπότε $w = -v_n$, όμως $w \in W$ και $-v_n \notin W$, από όπου έχουμε μια αντίφαση. □

Παρατήρηση VII.1.12. Η πρόταση VII.1.11 δεν ισχύει αν η διάσταση του V δεν είναι πεπερασμένη. Για παράδειγμα στον χώρο $C([0,1])$ των συνεχών συναρτήσεων $f : [0,1] \to \mathbb{R}$ ορίζονται οι νόρμες

$$\|f\|_\infty = \max_{0 \leq x \leq 1} |f(x)| \qquad \|f\|_1 = \int_0^1 |f(x)| dx$$

Είναι σαφές ότι $\|f\|_1 \leq \|f\|_\infty$ αλλά δεν υπάρχει C ώστε $\|f\| \leq C\|f\|_1$. Μπορούμε να διαλέξουμε τις συναρτήσεις $f_n(x)$ οι οποίες να είναι 0 στο διάστημα $(1/n, 1)$ και να έχουν το τρίγωνο σχήμα με κορυφή 1 στο $1/2n$ όπως στο παρακάτω σχήμα:



Η ακολουθία f_n συγκλίνει στην μηδενική συνάρτηση με την $\|\cdot\|_1$ νόρμα αφού το εμβαδόν της διαφοράς τείνει στο 0 αλλά όχι με την $\|\cdot\|_\infty$ αφού η νόρμα $\|\cdot\|_\infty$ βλέπει την μεγαλύτερη τιμή της $f_n(x)$ που είναι 1.

VII.2 Χώροι με εσωτερικό γινόμενο

Στο κεφάλαιο αυτό θα περιορίσουμε το σώμα $\mathbb{F}$ στην περίπτωση που $\mathbb{F} = \mathbb{R}$ ή $\mathbb{F} = \mathbb{C}$.

Ορισμός VII.2.1. Ένας χώρος με εσωτερικό γινόμενο είναι ένας διανυσματικός χώρος V πάνω από το σώμα $\mathbb{F}$, με ένα εσωτερικό γινόμενο που συμβολίζεται ως

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$$

για κάθε όλα τα $u, v, w \in V$ και κάθε $c \in \mathbb{F}$

1. Γραμμικότητα ως προς το πρώτο όρισμα: $\langle cu, v\rangle = c\langle u, v\rangle$
2. Συζυγής Συμμετρία: $\langle u, v\rangle = \overline{\langle v, u\rangle}$
3. Γραμμικότητα ως προς το δεύτερο όρισμα: $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$
4. Θετικά ορισμένη:$\langle u, u\rangle \geq 0$, και $\langle u, u\rangle = 0$ αν και μόνο αν $u = \mathbf{0}_V$.

Παραδείγματα VII.2.2. 1. Θεωρούμε τα στοιχεία $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$. Ορίζουμε την συνάρτηση

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i, \end{aligned}$$

το οποίο είναι ένα εσωτερικό γινόμενο στον $\mathbb{R}^n$.

2. Θεωρούμε τα στοιχεία $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{C}^n$. Ορίζουμε την συνάρτηση

$$\begin{aligned} \mathbb{C}^n \times \mathbb{C}^n &\longrightarrow \mathbb{C} \\ (x, y) &\longmapsto \langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y}_i, \end{aligned}$$

το οποίο είναι ένα εσωτερικό γινόμενο στον $\mathbb{R}^n$. Αυτό στην βιβλιογραφία είναι γνωστό και ως ερμητιανό εσωτερικό γινόμενο.

Γιατί δεν ορίσαμε και στους μιγαδικούς αριθμούς το εσωτερικό γινόμενο όπως και στους πραγματικούς; Δηλαδή

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

και βάλαμε μια συζυγία στα y_i; Από τον ορισμό του εσωτερικού γινομένου χρειαζόμαστε

$$\langle x, x \rangle = \sum_{i=1}^{n} x_i^2 \geq 0$$

το οποίο δεν μπορεί να εξασφαλιστεί αν $x_i \in \mathbb{C}$ αφού τα τετράγωνα μπορεί να είναι αρνητικά. Ούτε μπορούμε να εξασφαλίσουμε με αυτόν τον ορισμό ότι

$$\langle x, x \rangle = 0 \Leftrightarrow x = 0.$$

Αντιθέτως για το ερμητιανό εσωτερικό γινόμενο έχουμε

$$\langle x, x \rangle = \sum_{i=1}^{n} x_i \overline{x_i} = \sum_{i=1}^{n} |x_i|^2$$

από το οποίο προκύπτουν οι ζητούμενες ιδιότητες του γινομένου.

3. Για τους πραγματικούς αριθμούς $a < b$ θεωρούμε τον χώρο $C([a, b])$ των συνεχών συναρτήσεων $f : [a, b] \to \mathbb{R}$. Ορίζουμε την συνάρτηση

$$\begin{aligned} C([a, b]) \times C([a, b]) &\longrightarrow \mathbb{R} \\ (f, g) &\longmapsto \int_a^b f(t) g(t) dt, \end{aligned}$$

το οποίο είναι ένα εσωτερικό γινόμενο στο $C([a, b])$.

4. Θεωρούμε τον χώρο $X := \{f : [a, b] \to \mathbb{C}, f \text{ συνεχής}\}$. Ορίζουμε την συνάρτηση

$$\begin{aligned} X \times X &\longrightarrow \mathbb{C} \\ (f, g) &\longmapsto \int_a^b f(t) \overline{g(t)} dt \end{aligned}$$

το οποίο είναι ένα εσωτερικό γινόμενο στο X.

Ο **Charles Hermite** (24 Δεκεμβρίου 1822 - 14 Ιανουαρίου 1901) ήταν Γάλλος Μαθηματικός που εργάστηκε στην Θεωρία Αριθμών, τις τετραγωνικές μορφές, την θεωρία των Ελλειπτικών συναρτήσεων και γενικότερα στην Άλγεβρα. Μια αναπηρία στο δεξί του πόδι τον εμπόδισε να συνεχίσει τις σπουδές του στην περίφημη École Polytechnique. Ανέπτυξε έντονη ερευνητική δραστηριότητα και το 1856 έγινε μέλος της Γαλλικής Ακαδημίας επιστημών. Απλοποίησε την απόδειξη του Abel σχετικά με την αδύναμια επίλυσης της πεπτοβάθμιας εξίσωσης με ριζικά και κατάφερε να επιλύσει την πεπτοβάθμια εξίσωση, όχι με ριζικά αλλά με την βοήθεια ελλειπτικών συναρτήσεων. Το 1862 εντάχθηκε ξανά στην École Polytechnique ως μέλος του διδακτικού προσωπικού ενώ του 1869 έγινε καθηγητής. Είναι επίσης γνωστός για την απόδειξη του ότι ο αριθμός e είναι υπερβατικός, ενώ η μέθοδός του ακολουθήθηκε από τον Lindemann στην απόδειξη του ότι και ο π είναι υπερβατικός.

Άσκηση VII.2.3. Έστω $V = \mathbb{C}^{n,n}$. Η συνάρτηση $V \times V \to V$, $(A,B) \mapsto \langle A,B\rangle = \mathrm{tr}(AB^*)$, όπου $B^* = \overline{B}^t$ είναι ένα εσωτερικό γινόμενο. Αν $B = (b_{ij})$ τότε $\overline{B}$ είναι ο πίνακας $(\overline{b}_{ij})$, δηλαδή έχουμε εφαρμόσει την μιγαδική συζυγία σε κάθε στοιχείο του πίνακα B.

Λύση VII.2.4. Ο χώρος των $n \times n$ πινάκων μπορεί να ταυτιστεί με το $\mathbb{C}^{n^2}$. Παρατηρούμε ότι

$$\mathrm{tr}(AB^*) = \sum_{i=1}^{n}\sum_{\nu=1}^{n} a_{i\nu}\overline{b}_{i\nu},$$

δηλαδή ταυτίζεται με το ερμητιανό εσωτερικό γινόμενο στον $\mathbb{C}^{n^2}$.

Άσκηση VII.2.5. Θεωρούμε το σύνολο $V = \mathbb{C}^n$ το οποίο το ταυτίζουμε με πίνακες στήλες. Για κάθε αντιστρέψιμο πίνακα $A \in \mathbb{C}^{n,n}$ θεωρούμε την συνάρτηση $V \times V \to V$, $(x,y) \mapsto \langle x,y\rangle_A = x^*A^*Ay$. Δείξτε ότι είναι ένα εσωτερικό γινόμενο.

Λύση VII.2.6. Αν $\langle\cdot,\cdot\rangle$ είναι το κανονικό ερμητιανό εσωτερικό γινόμενο τότε παρατηρούμε ότι

$$\langle x,y\rangle_A = \langle Ax, Ay\rangle_H,$$

οπότε οι ζητούμενες ιδιότητες του εσωτερικού γινομένου ελέγχονται εύκολα με την βοήθεια του ερμητιανού εσωτερικού γινομένου.

Άσκηση VII.2.7. Δίνεται ένα εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ σε ένα διανυσματικό χώρο με βάση $B = \{v_1,\dots,v_n\}$. Να αποδειχθεί ότι οι τιμές $\langle v_i, v_j\rangle$ καθορίζουν πλήρως το εσωτερικό γινόμενο. Επιπλέον δείξτε ότι αν $[x]_B = (x_1,\dots,x_n)^t$, $[y]_B = (y_1,\dots,y_n)^t$ τότε

$$\langle x,y\rangle = \overline{y}^t Qx = y^*Qx, \tag{VII.3}$$

όπου ο Q είναι ο πίνακας $Q = (q_{ij})$ με $q_{ij} = \langle v_j, v_i\rangle$.

Λύση VII.2.8. Θωρούμε δύο στοιχεία x, y τα οποία τα εκφράζουμε με μοναδικό τρόπο ως γραμμικούς συνδυασμούς

$$x = \sum_{i=1}^{n} x_i v_i, \quad y = \sum_{j=1}^{n} y_j v_j$$

και υπολογίζουμε

$$\langle x,y\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\overline{y_j}\langle v_i, v_j\rangle$$

από όπου προκύπτει το ζητούμενο.

Άσκηση VII.2.9. Δείξτε ότι ικανή και αναγκαία συνθήκη για να ορίζεται ένα εσωτερικό γινόμενο από μια σχέση στις συντεταγμένες όπως δίνεται στην εξίσωση (VII.3) είναι να ισχύει $Q = Q^*$ και $x^*Qx > 0$ όταν $x \neq 0$.

Λύση VII.2.10. Ορίζουμε ένα εσωτερικό γινόμενο από την εξίσωση $\langle x, y\rangle = y^*Qx$. Η απαίτηση $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ προκύπτει από την επιμεριστική ιδιότητα του πολλαπλασιασμού ως προς την πρόσθεση. Το ίδιο και η ιδιότητα $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$.

Υπολογίζουμε ότι

$$\langle x, y\rangle = y^*Qx, \quad \langle y, x\rangle = x^*Qy,$$

και

$$\overline{\langle y, x\rangle} = \langle y, x\rangle^* = (x^*Qy)^* = y^*Q^*x.$$

Παρατηρούμε ότι η τελευταία ποσότητα είναι ίση με $\langle x, y\rangle = y^*Qx$ αν και μόνο αν $Q^* = Q$. Πράγματι η ποσότητα $e_i^*Qe_j = q_{ij}$.

Τέλος η συνθήκη $x^*Qx > 0$ προκύπτει από την συνθήκη $\langle x, x\rangle > 0$ και $\langle x, x\rangle = 0$ αν και μόνο αν $x = 0$.

Άσκηση VII.2.11. Αποδείξτε ότι ο τύπος

$$\left\langle \sum_{i=0}^{n} a_i x^i, \sum_{j=0}^{m} b_j x^j \right\rangle = \sum_{i=0}^{n}\sum_{j=0}^{m} \frac{a_i b_j}{i+j+1},$$

ορίζει ένα εσωτερικό γινόμενο στον χώρο των πολυωνύμων $\mathbb{R}[x]$.

Λύση VII.2.12. Γνωρίζουμε ότι στον χώρο $C[0,1]$ των συνεχών συναρτήσεων στο διάστημα $[0,1]$ το

$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt$$

είναι ένα εσωτερικό γινόμενο. Αν περιοριστούμε στον υπόχωρο των πολυώνυμων και ιδιαίτερα στην βάση $\{1, x, x^2, \ldots\}$ έχουμε ότι

$$\langle x^i, x^j\rangle = \int_0^1 t^{i+j}dt = \frac{1}{i+j+1},$$

από όπου προκύπτει το ζητούμενο.

VII.3 Ανισότητα Cauchy-Schwartz

Ορισμός VII.3.1. Σε ένα χώρο με εσωτερικό γινόμενο $(X, \langle\cdot,\cdot\rangle)$ ορίζεται μια συνάρτηση

$$\begin{aligned} \|\cdot\| : X &\longrightarrow \mathbb{R} \\ x &\longmapsto \langle x, x\rangle^{1/2} \end{aligned}$$

Πρόταση VII.3.2 (Ανισότητα Cauchy-Schwarz)**.** *Σε ένα χώρο με εσωτερικό γινόμενο* $(X, \langle\cdot,\cdot\rangle)$ *ισχύει η ανισότητα των Cauchy-Schwartz*

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|, \ \text{για κάθε } x, y \in X. \tag{VII.4}$$

Απόδειξη. Αν $y = 0$ τότε η ανισότητα (VII.4) ισχύει προφανώς. Υποθέτουμε ότι $y \neq 0$. Για ένα $\lambda \in \mathbb{F}$ ισχύει $\langle x - \lambda y, x - \lambda y \rangle \geq 0$. Υπολογίζουμε

$$\begin{aligned}\langle x - \lambda y, x - \lambda y \rangle &= \langle x, x - \lambda y \rangle - \lambda \langle y, x - \lambda y \rangle \\ &= \langle x, x \rangle - \bar{\lambda}\langle x, y \rangle - \lambda \langle y, x \rangle + \lambda\bar{\lambda}\langle y, y \rangle \\ &= \langle x, x \rangle - \bar{\lambda}\langle x, y \rangle - \lambda \overline{\langle x, y \rangle} + \lambda\bar{\lambda}\langle y, y \rangle,\end{aligned}$$

συνεπώς $\langle x, x \rangle - \bar{\lambda}\langle x, y \rangle - \lambda \overline{\langle x, y \rangle} + \lambda\bar{\lambda}\langle y, y \rangle \geq 0$. Επιλέγουμε $\lambda = \frac{\langle x,y \rangle}{\langle y,y \rangle}$ οπότε καταλήγουμε στο

$$\langle x, x \rangle - \frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle}\langle x, y \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle}\overline{\langle x, y \rangle} + \frac{\langle x, y \rangle}{\langle y, y \rangle}\frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle}\langle y, y \rangle$$

και συνεπώς

$$\langle x, x \rangle - \frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle}\langle x, y \rangle \geq 0 \Rightarrow \langle x, y \rangle \overline{\langle x, y \rangle} \leq \langle x, x \rangle \langle y, y \rangle$$

ή ισοδύναμα

$$|\langle x, y \rangle| \leq \langle x, x \rangle^{1/2} \langle y, y \rangle^{1/2} = \|x\| \cdot \|y\|.$$

□

Παρατήρηση VII.3.3. Αν είχαμε ένα πραγματικό εσωτερικό γινόμενο θα μπορούσαμε απλούστερα να γράψουμε την σχέση

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - 2\lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle.$$

Η παραπάνω σχέση είναι θετική για κάθε $\lambda \in \mathbb{R}$ οπότε η τετραγωνική εξίσωση στο λ έχει αρνητική διακρίνουσα δηλαδή

$$\langle x, y \rangle^2 - \langle x, x \rangle \langle y, y \rangle \leq 0$$

και η ανισότητα (VII.4) έπεται.

Παρατήρηση VII.3.4. Οι ανισότητες

$$\left| \sum_{i=1}^{n} x_i \bar{y}_i \right| \leq \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^{1/2} |y_i|^2 \right)^{1/2}$$

για μιγαδικούς αριθμούς $x_i, y_i, i = 1, 2, \ldots, n$

$$\left| \int_a^b f(t)g(t)dt \right| \leq \left(\int_a^b |f(t)|^2 \right)^{1/2} \left(\int_a^b |g(t)|^2 \right)^{1/2}$$

για συνεχείς πραγματικές συναρτήσεις είναι ειδικές περιπτώσεις της ανισότητας Cauchy-Schwartz στους χώρους $(\mathbb{C}^n, \langle \cdot, \cdot \rangle)$ και $\big(C([a, b]), \langle \cdot, \cdot \rangle\big)$.

Πρόταση VII.3.5. *Για ένα χώρο $(X, \langle \cdot, \cdot \rangle)$ με εσωτερικό γινόμενο η συνάρτηση $\| \cdot \| : X \to \mathbb{R}$ που ορίστηκε στον ορισμό VII.3.1 είναι μια νόρμα.*

Απόδειξη. Είναι σαφές ότι $\|x\| = 0$ αν και μόνο αν $\langle x, x \rangle = 0$ αν και μόνο αν $x = 0$. Επίσης είναι σαφές ότι

$$\|\lambda x\| = \langle \lambda x, \lambda x \rangle^{1/2} = (\lambda \bar{\lambda})^{1/2} \langle x, x \rangle^{1/2} = |\lambda| \|x\|.$$

Για να αποδείξουμε την τριγωνική ανισότητα έχουμε

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2\\ &= \|x\|^2 + 2\mathrm{Re}\langle x,y\rangle + \|y\|^2\\ &\overset{(\text{VII.4})}{\leq} \|x\|^2 + 2\|x\|\cdot\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2\end{aligned}$$

από την οποία έπεται η τριγωνική ανισότητα. □

Σε ένα χώρο με εσωτερικό γινόμενο ισχύει το Πυθαγόρειο θεώρημα

Θεώρημα VII.3.6.

$$\langle x,y\rangle = 0 \Rightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2 \text{ για κάθε } x,y \in X \tag{VII.5}$$

ή γενικότερα

$$\text{Αν για κάθε } i \neq j \text{ έχουμε } \langle x_i, x_j\rangle = 0 \text{ τότε } \|x_1 + x_2 + \cdots + x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_n\|^2. \tag{VII.6}$$

Απόδειξη. Έχουμε

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2$$

η οποία για $\langle x,y\rangle = 0$ δίνει την εξίσωση (VII.5). Θα αποδείξουμε την γενίκευση της με επαγωγή, την έχουμε ήδη αποδείξει στην περίπτωση $n = 2$. Θεωρούμε ότι ισχύει στην περίπτωση n η

$$\|x_1 + x_2 + \cdots + x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_n\|^2$$

και θα υπολογίσουμε το άθροισμα $\|x_1+x_2+\cdots+x_n+x_{n+1}\|^2$. Είναι σαφές ότι $\langle x_1+x_2+\cdots+x_n, x_{n+1}\rangle = 0$ συνεπώς έχουμε

$$\|x_1 + x_2 + \cdots + x_n + x_{n+1}\|^2 = \|x_1 + x_2 + \cdots + x_n\|^2 + \|x_{n+1}\|^2 = \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_n\|^2 + \|x_{n+1}\|^2.$$

□

Πρόταση VII.3.7. *Σε ένα χώρο με εσωτερικό γινόμενο $(X, \langle\cdot\rangle)$ ισχύει η ισότητα του παραλληλογράμμου*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \text{ για κάθε } x,y \in X.$$

Αντιστρόφως αν σε ένα χώρο με νόρμα $(X, \|\cdot\|)$, ισχύει η παραπάνω ισότητα τότε η νόρμα προέρχεται από ένα εσωτερικό γινόμενο.

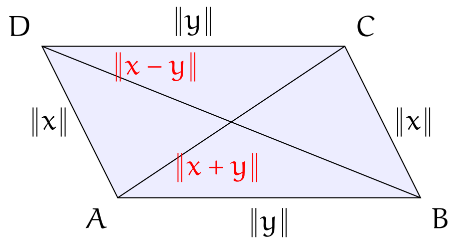
Απόδειξη. Προσθέτουμε κατά μέλη τις σχέσεις

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2\\ \|x-y\|^2 &= \langle x-y, x-y\rangle = \|x\|^2 - \langle x,y\rangle - \langle y,x\rangle + \|y\|^2\end{aligned}$$

οπότε προκύπτει η ισότητα του παραλληλογράμμου.

Αντιστρόφως, αν ισχύει η ισότητα του παραλληλογράμμου τότε το

$$\langle x,y\rangle := \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) \tag{VII.7}$$



Σχήμα VII.2: Γεωμετρική αναπαράσταση της ισότητας του παραλληλογράμμου. Τα $\|x\|, \|y\|$ είναι τα μήκη των πλευρών και $\|x+y\|, \|x-y\|$ είναι τα μήκη των διαγωνίων.

στην περίπτωση που ο διανυσματικός χώρος είναι πραγματικός και το

$$\langle x, y\rangle := \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right)$$

στην περίπτωση που ο διανυσματικός χώρος είναι μιγαδικός, ορίζουν εσωτερικό γινόμενο από το οποίο παράγεται η νόρμα $\|\cdot\|$.

Θα αποδείξουμε πρώτα την πραγματική περίπτωση. Υποθέτουμε ότι η νόρμα ικανοποιεί τον νόμο του παραλληλογράμμου, και θεωρούμε το εσωτερικό γινόμενο που ορίζεται από την εξίσωση (VII.7). Είναι σαφές ότι ικανοποιούνται οι ταυτότητας $\langle x, y\rangle = \langle y, x\rangle$ και $\|x\| = \sqrt{\langle x, y\rangle}$.

Παρατηρούμε ότι $\langle x+y+z\rangle = \langle x, z\rangle + \langle y, z\rangle$. Πράγματι, από τον κανόνα του παραλληλογράμμου έχουμε ότι

$$\|x+y+z\|^2 + \|x-y+z\|^2 = 2\|x+z\|^2 + 2\|y\|^2.$$

Αυτό μας δίνει ότι

$$\begin{aligned}\|x+y+z\|^2 &= 2\|x+z\|^2 + 2\|y\|^2 - \|x-y+z\|^2\\ &= 2\|y+z\|^2 + 2\|x^2\| - \|y-x+z\|^2,\end{aligned}$$

όπου η δεύτερη γραμμή προκύπτει εναλλάσσοντας τους ρόλους των x, y. Συνεπώς

$$\|x+y+z\|^2 = \|x\|^2 + \|y\|^2 + \|x+z\|^2 + \|y+z\|^2 - \frac{1}{2}\|x-y+z\|^2 - \frac{1}{2}\|y-x+z\|^2.$$

Με αντικατάσταση του z με $-z$ στην τελευταία εξίσωση παίρνουμε

$$\|x+y-z\|^2 = \|x\|^2 + \|y\|^2 + \|x-z\|^2 + \|y-z\|^2 - \frac{1}{2}\|x-y-z\|^2 - \frac{1}{2}\|y-x-z\|^2.$$

Τέλος χρησιμοποιώντας τα παραπάνω και τον ορισμό έχουμε

$$\begin{aligned}\langle x+y, z\rangle &= \frac{1}{4}\left(\|x+y+z\|^2 - \|x+y-z\|^2\right)\\ &= \frac{1}{4}\left(\|x+z\|^2 - \|x-z\|^2\right) + \frac{1}{4}\left(\|y+z\|^2 - \|y-z\|^2\right)\\ &= \langle x, z\rangle + \langle y, z\rangle\end{aligned}$$

Τώρα θα αποδείξουμε ότι $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$. Παρατηρούμε ότι η ζητούμενη ιδιότητα ισχύει για $\lambda = -1$, ενώ με βάση την προσθετική ιδιότητα και επαγωγή μπορούμε να αποδείξουμε ότι ισχύει για κάθε $\lambda \in \mathbb{N}$ και αφού ισχύει και για $\lambda = -1$ ισχύει για κάθε $\lambda \in \mathbb{Z}$. Αν τώρα $\lambda = p/q$ με $p, q \in \mathbb{Z}, q \neq 0$ τότε για $x' = x/q$ έχουμε ότι

$$q\langle \lambda x, y\rangle = q\langle px', y\rangle = p\langle qx', y\rangle = p\langle x, y\rangle,$$

οπότε διαιρώντας με q έχουμε την ζητούμενη ιδιότητα:

$$\langle \lambda x, y\rangle = \lambda\langle x, y\rangle \text{ για κάθε } x, y \in V, \lambda \in \mathbb{Q}.$$

Για να μπορέσουμε να περάσουμε σε $\lambda \in \mathbb{R}$, παρατηρούμε ότι η συνάρτηση

$$\begin{aligned} f(t) : \mathbb{R}^* &\longrightarrow \mathbb{R} \\ t &\longmapsto \frac{1}{t}\langle tx, y\rangle \end{aligned}$$

είναι συνεχής στο πεδίο ορισμού της, δηλαδή για κάθε ακολουθία (a_n) με $a_n \to l$ έχουμε $f(a_n) \to f(l)$. Επίσης παρατηρούμε ότι η $f(t) = \langle x, y\rangle$ για κάθε $t \in \mathbb{Q}$. Οπότε αν $\lambda \in \mathbb{R}^*$, διαλέγουμε μια ακολουθία ρητών αριθμών (a_n) με $a_n \to \lambda$ και έχουμε ότι $f(\lambda) = \frac{1}{\lambda}\langle \lambda x, y\rangle = \langle x, y\rangle$, δηλαδή το ζητούμενο. Η περίπτωση $\lambda = 0$ είναι προφανής.

Τέλος για την περίπτωση του μιγαδικού εσωτερικού γινομένου παρατηρούμε ότι $i\langle x, y\rangle = i\langle x, y\rangle$ και ότι $\langle x, y\rangle = \overline{\langle y, x\rangle}$. Για την απόδειξη της γραμμικότητας επαναλαμβάνουμε τα επιχειρήματα που εφραμόσαμε στο πραγματικό εσωτερικό γινόμενο για το πραγματικό $\mathrm{Re}\langle\cdot,\cdot\rangle$ και το φαντασικό $\mathrm{Im}\langle\cdot,\cdot\rangle$ μέρος του $\langle\cdot,\cdot\rangle$. □

VII.3.1 Γωνία διανυσμάτων

Η ανισότητα Cauchy-Schwartz (VII.4) εξασφαλίζει ότι στην περίπτωση πραγματικού εσωτερικού γινομένου

$$-1 \le \frac{\langle x, y\rangle}{\|x\| \cdot \|y\|} \le 1$$

και συνεπώς μπορούμε να ορίσουμε την γωνία θ, $0 \le \theta < 2\pi$ των διανυσμάτων x, y ώστε

$$\langle x, y\rangle = \|x\| \cdot \|y\| \cos(\theta).$$

Στην περίπτωση που $\langle x, y\rangle = 0$ θα λέμε ότι τα x, y είναι κάθετα.

Ορισμός VII.3.8. Ένα σύνολο στοιχείων $S = \{e_\mu \in V, \mu \in I\}$ θα λέγεται ορθογώνιο αν και μόνο αν $\langle e_\mu, e_\nu\rangle = \delta_{i,j}\|e_\mu\|^2$. Θα λέγεται ορθοκανονικό αν επιπλέον $\|e_\mu\| = 1$ για κάθε $\mu \in I$.

Παρατήρηση VII.3.9. Ένα ορθογώνιο σύνολο S μπορεί να μετασχηματιστεί σε ορθοκανονικό $S' = \{e_\mu/\|e_\mu\| : \mu \in I\}$.

Πρόταση VII.3.10. *Ένα ορθογώνιο σύνολο είναι γραμμικά ανεξάρτητο.*

Απόδειξη. Πράγματι για οποιαδήποτε πεπερασμένη επιλογή $e_{i_1}, \dots, e_{i_r}$ από το ορθογώνιο σύνολο και σχέση

$$\mathbf{0}_V = \sum_{\nu=1}^{r} \lambda_\nu e_{i_\nu}$$

έχουμε ότι

$$\mathbf{0}_V = \langle \mathbf{0}_V, e_{i_j}\rangle = \left\langle \sum_{\nu=1}^{r} \lambda_\nu e_{i_\nu}, e_{i_j} \right\rangle = \sum \nu = 1^r \lambda_\nu \langle e_{i_\nu}, e_{i_j} = \lambda_j,$$

□

δηλαδή όλοι οι συντελεστές που εμφανίζονται είναι μηδενικοί.

Πρόταση VII.3.11 (Ανάλυση Fourier). *Αν σε ένα διανυσματικό χώρο έχουμε μια ορθοκανονική βάση* $e_1, \dots, e_n$, *τότε για κάθε διάνυσμα* $v \in V$ *γράφεται ως*

$$\sum_{i=1}^{n} \langle v, e_i \rangle e_i.$$

Απόδειξη. Αφού τα $e_1, \dots, e_n$ είναι βάση, το τυχαίο στοιχείο $v \in V$ γράφεται ως

$$v = \sum_{\nu=1}^{n} \lambda_\nu \nu.$$

Υπολογίζουμε ότι

$$\langle v, e_i \rangle = \langle \sum_{\nu=1}^{n} \lambda_\nu e_\nu, e_i \rangle = \sum_{\nu=1}^{n} \lambda_\nu \langle e_\nu, e_i \rangle = \lambda_i.$$

□

Θεωρία αναπαραστάσεων...

Παράδειγμα VII.3.12. Θεωρούμε τον χώρο

$$X := \{f : [0, 2\pi] \to \mathbb{R} : f(0) = f(2\pi), f \text{ συνεχής } \}.$$

Ο χώρος αυτός μπορεί να ταυτιστεί με τον χώρο των περιοδικών συναρτήσεων $f : \mathbb{R} \to \mathbb{R}$ οι οποίες είναι περιοδικές με περίοδο 2π.

Στον χώρο αυτό εισάγουμε το εσωτερικό γινόμενο:

$$\langle f, g \rangle = \int_0^{2\pi} f(t)g(t)dt.$$

Παρατηρούμε ότι οι συναρτήσεις

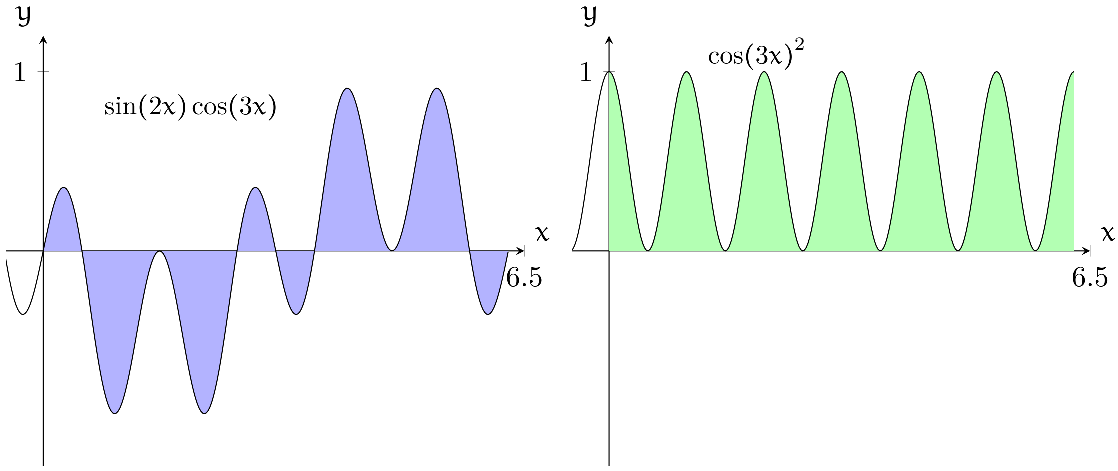
$$e_0(s) = \frac{1}{\sqrt{2\pi}}, e_1(s) = \frac{1}{\sqrt{\pi}}\sin(s), e_2(s) = \frac{1}{\sqrt{\pi}}\cos(s), \dots, e_{2\nu-1}(s) = \frac{1}{\sqrt{\pi}}\sin(\nu s), e_{2\nu}(s) = \frac{1}{\sqrt{\pi}}\cos(\nu s), \dots$$

είναι στοιχεία του χώρου X, και μάλιστα ότι

$$\langle e_\nu, e_\mu \rangle = \delta_{\nu,\mu} \text{ για } \nu, \mu \in \mathbb{N}.$$

Πράγματι

$$\begin{aligned}
\int_0^{2\pi} \sin(\nu t)\sin(\mu t)dt &= \frac{1}{2}\int_0^{2\pi} \left(\cos\left((\nu-\mu)t\right) - \cos\left((\nu+\mu)t\right)\right) dt = \pi \cdot \delta_{\nu,\mu} \\
\int_0^{2\pi} \cos(\nu t)\cos(\mu t)dt &= \frac{1}{2}\int_0^{2\pi} \left(\cos\left((\nu-\mu)t\right) + \cos\left((\nu+\mu)t\right)\right) dt = \pi \cdot \delta_{\nu,\mu} \\
\int_0^{2\pi} \sin(\nu t)\cos(\mu t)dt &= \frac{1}{2}\int_0^{2\pi} \left(\sin\left((\nu-\mu)t\right) + \sin\left((\nu+\mu)t\right)\right) dt = 0 \\
\int_0^{2\pi} e_0(t)e_\mu(t)dt &= \delta_{0,\mu}.
\end{aligned}$$



Σχήμα VII.3: Το αριστερό εμβαδόν είναι 0, τα εμβαδά κάτω από τον άξονα των x είναι αρνητικά. Το δεξί εμβαδόν είναι π

VII.4 Μια εφαρμογή στις διαφορικές εξισώσεις

VII.4.1 Συστήματα διαφορικών εξισώσεων

Ας ξεκινήσουμε για προθέρμανση με το εξής πρόβλημα διαφορικών εξισώσεων: Να βρεθεί μια συνάρτηση $f(x)$ ώστε

$$f'(x) = \lambda f(x). \tag{VII.8}$$

Αν υποθέσουμε ότι έχει αναλυτική λύση

$$f(x) = \sum_{i=0}^{\infty} a_i x^i$$

τότε η εξίσωση (VII.8) μας δίνει ότι

$$\sum_{i=0}^{\infty} a_i i x^{i-1} = \lambda \sum_{i=0}^{\infty} a_i x^i$$

από όπου προκύπτει ότι

$$a_{i+1} = a_i \lambda / i, \text{ συνεπώς } a_i = \frac{\lambda^i}{i!} a_0.$$

και καταλήγουμε στην

$$f(x) = a_0 \sum_{i=0}^{\infty} \frac{(\lambda x)^i}{i!} = a_0 e^{\lambda x}.$$

VII.4.2 Ο εκθετικός πίνακας

Έχουμε ήδη δει ότι έχει νόημα να ορίσουμε για $A \in \mathbb{F}^{n,n}$ και $f(x) = \sum_{i=0}^{n} a_i x^i \in \mathbb{F}[x]$ τον πίνακα

$$f(A) = \sum_{i=0}^{n} a_i A^i \in \mathbb{F}^{n,n}.$$

Θέλουμε να επεκτείνουμε τον παραπάνω ορισμό για συναρτήσεις οι οποίες εκφράζονται μέσω δυναμοσειρών όπως τις ορίσαμε στην παράγραφο I.9.7. Για παράδειγμα ιδιαίτερα ενδιαφέρον είναι ο εκθετικός πίνακας

$$\exp(A) := \sum_{i=0}^{n} \frac{A^i}{i!}.$$

Για ένα τέτοιο πίνακα χρειαζόμαστε την έννοια της σύγκλισης. Θα χρειαστεί να εισάγουμε μερικές έννοιες για νόρμες πινάκων. Ο χώρος των $n \times n$ πινάκων είναι ένας διανυσματικός χώρος, διάστασης n^2 οπότε μπορούμε να ορίσουμε την έννοια της νόρμας όπως κάναμε στην παράγραφο VII.1. Χρειαζόμαστε νόρμες οι οποίες να σέβονται την πράξη του πολλαπλασιασμού πινάκων

Ορισμός VII.4.1. Μια νόρμα στον χώρο των $n \times n$ πινάκων θα λέγεται υποπολλαπλασιαστική αν και μόνο αν ισχύει

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

Επίσης, θα λέμε ότι η νόρμα πινάκων είναι συμβατή με την νόρμα του διανυσματικού χώρου αν ισχύει

$$\|Av\| \leq \|A\|\|v\|.$$

Ορισμός VII.4.2. Έστω $\mathbb{F} = \mathbb{C}$ ή $\mathbb{F} = \mathbb{R}$. Από κάθε νόρμα διανυσματικών χώρων $\|\cdot\| : \mathbb{F}^n \to \mathbb{R}$ μπορούμε να ορίσουμε την *επαγόμενη νόρμα*

$$\|A\| = \max_{v \in \mathbb{F}^n \setminus \{\mathbf{0}\}} \frac{\|Av\|}{\|v\|} = \max_{v \in \mathbb{F}^n : \|v\|=1} \|Av\|. \tag{VII.9}$$

Λήμμα VII.4.3. *Η εξίσωση (VII.9) ορίζει νόρμα συμβατή και υποπολλαπλασιαστική. Επιπλέον* $\|\mathbb{I}_n\| = 1$.

Απόδειξη. Από τον ορισμό της νόρμας έχουμε για κάθε $v \in \mathbb{R}^n$, $v \neq \mathbf{0}$

$$\|Av\|/\|v\| \leq \|A\|$$

από όπου προκύπτει ότι η νόρμα είναι συμβατή. Η περίπτωση $v = \mathbf{0}$ είναι προφανής.

Παρατηρούμε, χρησιμοποιώντας την συμβατότητα της νόρμας, ότι για κάθε $v \in \mathbb{R}^n$

$$\|ABv\| \leq \|A\|\|Bv\| \leq \|A\|\|B\|\|v\|$$

από όπου έχουμε ότι

$$\|AB\| \leq \|A\|\|B\|$$

από όπου έχουμε ότι η νόρμα είναι υποπολλαπλασιαστική.

Τέλος η νόρμα του $\|\mathbb{I}_n\| = 1$ από τον ορισμό αφού το μέγιστο υπολογίζεται πάνω στην σταθερή ποσότητα $\|v\|/\|v\| = 1$. □

Ασκήσεις από το 2020_03_25.pdf

Λήμμα VII.4.4. *Αν μια σειρά πινάκων ικανοποιεί*

$$\left\|\sum_{n=0}^{\infty} a_i A^i\right\| \leq \sum_{n=0}^{\infty} M_i \leq \infty$$

τότε αυτή συγκλίνει.

Απόδειξη. Ο χώρος των $n \times n$ είναι πεπερασμένης διάστασης πάνω από το πλήρες σώμα $\mathbb{F}$ συνεπώς είναι πλήρης. Αυτό σημαίνει ότι μια ακολουθία είναι Cauchy αν και μόνο αν συγκλίνει. Η συνθήκη Cauchy για μια νόρμα πινάκων είναι η

$$\text{Για κάθε } \epsilon > 0 \text{ υπάρχει } n_0 \in \mathbb{N}, \text{ ώστε } n \geq m > n_0 \text{ ισχύει } \left\| \sum_{i=m}^{n} a_n A^n \right\| \leq \epsilon$$

□

Ορισμός VII.4.5. Για σταθερό $A \in \mathbb{F}^{n,n}$ ορίζουμε την εκθετική συνάρτηση $\mathbb{F} \to \mathbb{F}^{n,n}$ με

$$\exp(Ax) = \sum_{i=0}^{n} \frac{A^n x^n}{n!}.$$

Για να δείξουμε ότι η παραπάνω συνάρτηση είναι καλά ορισμένη αρκεί να χρησιμοποιήσουμε το λήμμα VII.4.4 και να παρατηρήσουμε ότι για μια υποπολλαπλασιαστική νόρμα έχουμε

$$\left\| \sum_{i=0}^{n} \frac{A^n x^n}{n!} \right\| \leq \sum_{i=0}^{n} \frac{\|A\|^n |x|^n}{n!}$$

το οποίο συγκλίνει αφού αναπαριστά το $e^{\|A\||x|}$.

Πρόταση VII.4.6. *Στην περίπτωση που ο πίνακας A είναι διαγωνοποιήσιμος, δηλαδή υπάρχει αντιστρέψιμος πίνακας Q ώστε $A = Q\mathrm{diag}(\lambda_1, \dots, \lambda_n)Q^{-1}$ έχουμε*

$$\exp(At) = Q\mathrm{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})Q^{-1}.$$

Απόδειξη. Από τον ορισμό έχουμε

$$\mathrm{ext}(At) = \sum_{i=0}^{\infty} \frac{A^i t^i}{i!} = Q\mathrm{diag}\left(\sum_{i=0}^{\infty} \frac{\lambda_1^i t^i}{i!}, \dots, \sum_{i=0}^{\infty} \frac{\lambda_n^i t^i}{i!} \right) Q^{-1} = Q\mathrm{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})Q^{-1}.$$

□

Πρόταση VII.4.7. *Έχουμε ότι η παράγωγος*

$$\frac{d}{dt}\exp(At) = A \cdot \exp(At).$$

Απόδειξη.

$$\frac{d}{dt}\exp(At) = \frac{d}{dt}\left(\sum_{i=0}^{\infty} \frac{A^i t^i}{i!} \right) = \sum_{i=1}^{\infty} \frac{A^i i t^{i-1}}{i!} = A \sum_{i=1}^{\infty} \frac{A^i t^{i-1}}{(i-1)!} = A \sum_{i=0}^{\infty} \frac{A^i t^i}{i!} = A \cdot \exp(At).$$

□

Πρόταση VII.4.8. *Για κάθε πίνακα $A \in \mathbb{C}^{n,n}$ ο εκθετικός πίνακας $\exp(A)$ είναι αντιστρέψιμος.*

Το $\exp(At)$ *είναι μια οικογένεια αντιστρέψιμων πινάκων η οποία για* $t = 0$ *δίνει:*

$$\exp(At)|_{t=0} = \mathbb{I}_n, \quad \frac{d}{dt}\exp(At)|_{t=0} = A.$$

Απόδειξη. Το ότι $\exp(At)|_{t=0} = \exp(\mathbf{0}_{\mathbb{C}^{n,n}}) = \mathbb{I}_n$ προκύπτει από τον ορισμό του εκθετικού πίνακα VII.4.5. Η αντιστρεψιμότητα προκύπτει από το ότι

$$\mathbb{I}_n = \mathbf{0}_{\mathbb{C}^{n,n}} = \exp(At - At) = \exp(At)\exp(-At).$$

Ειδικότερα έχουμε ότι

$$\exp(At)^{-1} = \exp(-At).$$

□

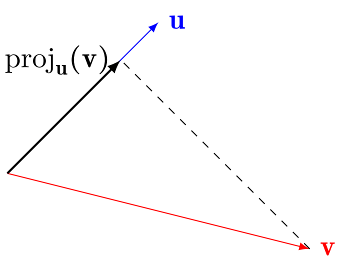
Επιστροφή στις Διαφορικές εξισώσεις

VII.5 Ορθοκανονικοποίηση

Στην παράγραφο αυτή θα περιγράψουμε μία μέθοδο η οποία δοσμένου ενός πεπερασμένου συνόλου γραμμικά ανεξάρτητων διανυσμάτων $v_1, \ldots, v_n$ σε ένα χώρο με εσωτερικό γινόμενο δίνει μια ορθοκανονική βάση του χώρου $\langle v_1, \ldots, v_n\rangle$ που αυτά παράγουν.

Ορισμός VII.5.1. Η προβολή ενός στοιχείου $v \in V$ επί ενός μη μηδενικού στοιχείου $u \in V$ ορίζεται ως

$$\mathrm{proj}_u(v) = \frac{\langle v, u\rangle}{\langle u, u\rangle}u.$$



Σχήμα VII.4: Γεωμετρική αναπαράσταση της προβολής του διανύσματος v στο διάνυσμα u.

Παρατήρηση VII.5.2. Είναι σαφές ότι το διάνυσμα $\mathrm{proj}_u(v)$ είναι παράλληλο στο u και ότι το $v - \mathrm{proj}_u(v)$ είναι κάθετο με το u αφού

$$\langle v - \mathrm{proj}_u(v), u\rangle = \langle v, u\rangle - \frac{\langle v, u\rangle}{\langle u, u\rangle}\langle u, u\rangle = 0.$$

Η διαδικασία ορθοκανονικοποίησης κατά Gram-Schmidt είναι η παρακάτω:

Μέθοδος 5 (Gram-Schmidt ορθοκανονικοποίηση)**.** *Δίνεται ένα πεπερασμένο σύνολο διανυσμάτων* $v_1, \ldots, v_n$ *για να βρούμε μια ορθοκανονική βάση* $e_, \ldots, e_n$ *του χώρου που παράγουν εκτελούμε τα παρακάτω βήματα:*

1 *Θέτουμε* $u_1 = v_1$

2 *Θέτουμε* $u_2 = v_2 - \mathrm{proj}_{u_1}(v_2)$

3 *Θέτουμε* $u_3 = v_3 - \mathrm{proj}_{u_1}(v_3) - \mathrm{proj}_{u_2}(v_3)$

4 *Θέτουμε* $u_4 = v_4 - \mathrm{proj}_{u_1}(v_4) - \mathrm{proj}_{u_2}(v_4) - \mathrm{proj}_{u_3}(v_4)$

5 $\vdots$

6 *Θέτουμε* $u_n = v_n - \sum_{j=1}^{n-1} \mathrm{proj}_{u_j}(v_n)$.

7 *Θέτουμε* $e_i = \frac{u_i}{\|u_i\|}$.

Πρόταση VII.5.3. *Η μέθοδος Gram-Schmidt δίνει πράγματι μια ορθοκανονική βάση του χώρου που παράγουν τα* $v_1, \ldots, v_n$.

Απόδειξη. Πράγματι, παρατηρούμε ότι το σύνολο των διανυσμάτων $u_1, \ldots, u_n$ είναι ανά δύο κάθετα. Αυτό για $n = 1$ δεν έχει νόημα και για $n = 2$ το είδαμε στην παρατήρηση VII.5.2. Υποθέτουμε επαγωγικά ότι τα $u_1, \ldots, u_k$ είναι ανά δύο κάθετα και θα δείξουμε ότι και το u_{k+1} είναι κάθετο με όλα τα προηγούμενα. Έχουμε ότι για $1 \le i \le k$

$$\begin{aligned}
\langle u_{k+1}, u_i \rangle &= \left\langle v_{k+1} - \sum_{j=1}^{k} \mathrm{proj}_{u_j}(v_{k+1}), u_i \right\rangle \\
&= \langle v_{k+1}, u_i \rangle - \langle \mathrm{proj}_{u_i}(v_{k+1}), u_i \rangle \\
&= \langle v_{k+1}, u_i \rangle - \frac{\langle v_{k+1}, u_i \rangle}{\langle u_i, u_i \rangle} \langle u_i, u_i \rangle = 0.
\end{aligned}$$

Παρατηρούμε ότι τα $e_i \in \langle v_1, \ldots, v_n$ για $1 \le i \le n$, συνεπώς $\langle e_1, \ldots, e_n \rangle \subset \langle v_1, \ldots, v_n \rangle$. Τέλος τα $\{e_1, \ldots, e_n\}$ είναι ως ορθοκανονικό σύνολο γραμμικά ανεξάρτητο άρα $\langle e_1, \ldots, e_n \rangle = \langle v_1, \ldots, v_n \rangle$. □

Παράδειγμα VII.5.4. Έχουμε τα διανύσματα

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

από τα οποία θα κατασκευάσουμε μια ορθοκανονική βάση με την μέθοδο Gram-Schmidt.

Ξεκινάμε με

$$u_1 = v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Υπολογίζουμε

$$\mathrm{proj}_{u_1}(v_2) = \frac{\langle v_2, u_1 \rangle}{\|u_1\|^2} u_1 = \frac{2}{7} u_1$$

Στη συνέχεια, υπολογίζουμε το u_2 ως εξής:

$$u_2 = v_2 - \mathrm{proj}_{u_1}(v_2) = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - \frac{2}{7} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 12 \\ 3 \\ -6 \end{pmatrix}$$

Συνεχίζουμε με

$$\mathrm{proj}_{u_1}(v_3) = \frac{\langle v_3, u_1\rangle}{\|u_1\|^2} u_1 = \frac{7}{14} u_1 = \frac{1}{2} u_1$$
$$\mathrm{proj}_{u_2}(v_3) = \frac{\langle v_3, u_2\rangle}{\|u_2\|^2} u_2 = \frac{-21}{27} u_2$$

και

$$\begin{aligned} u_3 &= v_3 - \mathbf{proj}_{u_1}(v_3) - \mathbf{proj}_{u_2}(v_3) \\ &= \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \frac{3}{27}\begin{pmatrix} 12 \\ 3 \\ -6 \end{pmatrix} = \begin{pmatrix} -1/6 \\ 1/3 \\ -1/6 \end{pmatrix}. \end{aligned}$$

Τέλος υπολογίζουμε ότι

$$\|u_1\| = \sqrt{14}, \qquad \|u_2\| = \sqrt{27/7}, \qquad \|u_3\| = \|\sqrt{1/6}\|$$

και έχουμε την ορθοκανονική βάση

$$e_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{14}} u_1, \qquad e_2 = \frac{u_2}{\|u_2\|} = \sqrt{7/27} u_2, \qquad e_3 = \frac{u_3}{\|u_3\|} = \sqrt{6} u_3.$$

VII.5.1 Εφαρμογή σε προβλήματα βελτιστοποίησης

Ορισμός VII.5.5. Θεωρούμε ένα διανυσματικό χώρο $(V, \|\cdot\|)$ με νόρμα και ένα υποσύνολο $\varnothing \neq \Sigma \subset V$. Ένα $x \in \Sigma$ θα λέγεται βέλτιστη προσέγγιση του $z \in V$ αν ισχύει

$$\|x - z\| \leq \|x - y\| \text{ για κάθε } y \in V.$$

Παρατήρηση VII.5.6. Ένα τέτοιο στοιχείο μπορεί να μην υπάρχει, θεωρήστε ένα ανοιχτό διάστημα $\Sigma = (0,1) \in \mathbb{R}$, ή να μην είναι μοναδικό.

Έστω $(V, \langle\cdot,\cdot\rangle)$ ένας χώρος με εσωτερικό γινόμενο και $W \subset V$ ένας υπόχωρος του. Θεωρούμε μια ορθοκανονική βάση $\{w_1, \dots, w_n\}$ του υπόχωρου W και ορίζουμε το στοιχείο

$$x = \langle z, w_1\rangle w_1 + \cdots + \langle z, w_n\rangle w_n. \tag{VII.10}$$

Για κάθε $1 \leq j \leq n$ υπολογίζουμε

$$\langle z - x, w_j\rangle = \langle z, w_j\rangle - \langle x, w_j\rangle = \langle z, w_j\rangle - \langle z, w_j\rangle = 0,$$

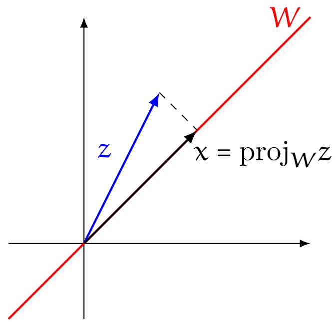
λαμβάνοντας υπόψιν ότι $\langle w_i, w_j\rangle = \delta_{ij}$. Συνεπώς το διάνυσμα $z - x$ είναι ορθογώνιο με κάθε γραμμικό συνδυασμό των $w_1, \dots, w_n$ δηλαδή με κάθε στοιχείο του υποχώρου W:

$$\langle z - x, w\rangle = 0 \text{ για κάθε } w \in W. \tag{VII.11}$$

Το σημείο x είναι η ορθογώνια προβολή του z στον χώρο W. Θα αποδείξουμε ότι το x είναι πραγματικά μια βέλτιστη προσέγγιση του z από στοιχεία του W. Θεωρούμε ένα στοιχείο $y \in W$. Ισχύει ότι $x - y \in W$ και η (VII.11) δίνει $\langle z - x, x - y\rangle = 0$. Σύμφωνα με το Πυθαγόρειο θεώρημα έχουμε ότι

$$\|z - y\|^2 = \|(z - x) + (x - y)\|^2 = \|z - x\|^2 + \|x - y\|^2 \geq \|z - x\|^2,$$

δηλαδή το x είναι μια βέλτιστη προσέγγιση. Αν $x \neq y$ τότε $\|z - y\| > \|z - x\|$, δηλαδή η βέλτιστη προσέγγιση είναι μοναδική.



Σχήμα VII.5: Προβολή του z στον υπόχωρο W, η διακεκομμένη γραμμή είναι η ελάχιστη απόσταση του z από τον υπόχωρο W.

Παράδειγμα VII.5.7. Θεωρούμε τον χώρο

$$X := \{f : [0, 2\pi] \to \mathbb{R} : f(0) = f(2\pi), f \text{ συνεχής } \}.$$

του παραδείγματος VII.3.12 και τον πεπερασμένο υπόχωρο

$$T_n = \left\{ x \in X : x(t) = \alpha_0 + \sum_{\nu=1}^{n} (\alpha_\nu \cos(\nu t) + \beta_\nu \sin(\nu s)) \right\}$$

των τριγωνομετρικών πολυώνυμων βαθμού το πολύ n. Οι συναρτήσεις

$$e_0(t) = \frac{1}{\sqrt{2\pi}}, e_{2\nu-1}(t) = \frac{1}{\sqrt{\pi}} \sin(\nu t), e_{2\nu} = \frac{1}{\sqrt{\pi}} \cos(\nu t), \nu = 1, \dots, n$$

αποτελούν μια ορθοκανονική βάση του T_n. Η βέλτιστη προσέγγιση μιας τυχαίας περιοδικής συνάρτησης από τον T_n είναι το μερικό άθροισμα της λεγόμενης σειράς Fourier της x:

$$S_n := \frac{1}{2\pi} \int_0^{2\pi} x(t) dt + \frac{1}{\sqrt{\pi}} \sum_{\nu=1}^{2n} \left(\int_0^{2\pi} x(t) e_\nu(t) dt \right) e_\nu.$$

Η μελέτη των σειρών Fourier είναι κλάδος της Ανάλυσης γνωστός ως Αρμονική Ανάλυση και έχει πολλές εφαρμογές στα Θεωρητικά και Εφαρμοσμένα Μαθηματικά, στην Φυσική και στην τεχνολογία.

VII.5.2 Ορθογώνιες προβολές

Έστω ένας πεπερασμένης διάστασης διανυσματικός χώρος V με εσωτερικό γινόμενο και έστω W ένας υπόχωρος. Η μέθοδος Gram-Schmidt μας δίνει μία μέθοδο να κατασκευάσουμε μια ορθοκανονική βάση $B_W = \{e_1, \dots, e_n\}$ του W την οποία μπορούμε να την επεκτείνουμε σε μία ορθοκανονική βάση $B_V = \{e_1, \dots, e_n, \epsilon_1, \dots, \epsilon_m\}$ του V. Είναι σαφές ότι αν θέσουμε

$$W^\perp = \langle \epsilon_1, \dots, \epsilon_m \rangle$$

τότε

$$V = W \bigoplus W^\perp.$$

Μπορούμε να εφοδιάσουμε τον χώρο των γραμμικών συναρτήσεων $\mathrm{End}(V)$ από τον $V \to V$ με δύο στοιχεία $p_W, p_{W^\perp}$, τις ορθογώνιες προβολές, όπου

$$p_W : \left\{ \begin{array}{rl} V & \longrightarrow W \subset V \\ w + w' = v & \longmapsto w \end{array} \right. \qquad p_{W^\perp} : \left\{ \begin{array}{rl} V & \longrightarrow W^\perp \subset V \\ w + w' = v & \longmapsto w' \end{array} \right.$$

όπου $v = w + w'$ είναι η μοναδική γραφή του $v \in V$ ως άθροισμα δύο στοιχείων $w \in W, w' \in W^{\perp}$.

Είναι σαφές ότι $\mathrm{Id}_V = p_W + p_{W^{\perp}}$, $p_W^2 = p_W$, $p_{W^{\perp}}^2 = p_{W^{\perp}}$ και $p_W p_{W^{\perp}} = p_{W^{\perp}} p_W$. Τέλος για τους πίνακες των γραμμικών συναρτήσεων $p_W, p^{W^{\perp}}$ ως προς την βάση B_V έχουμε

$$(p_W, B_V, B_V) = \left(\begin{array}{c|c} \mathbb{I}_n & \mathbf{0}_{n,m} \\ \hline \mathbf{0}_{m,n} & \mathbf{0}_{m,m} \end{array}\right) \qquad (p_{W^{\perp}}, B_V, B_V) = \left(\begin{array}{c|c} \mathbf{0}_{n,n} & \mathbf{0}_{n,m} \\ \hline \mathbf{0}_{m,n} & \mathbb{I}_n \end{array}\right)$$

Παρατήρηση VII.5.8. Για ένα πίνακα $A = (a_{ij}) \in \mathbb{R}^{n,m}$ θεωρούμε το γραμμικό σύστημα

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,m} \\ a_{21} & a_{22} & \cdots & a_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Έστω $a_1, \ldots, a_n$ οι γραμμές του A και έστω $W \subset \mathbb{R}^m$ ο χώρος γραμμών του πίνακα. Είναι σαφές ότι ο χώρος λύσεων του συστήματος ταυτίζεται με τον

$$W^{\perp} = \{x = (x_1, \ldots, x_m)^t \in \mathbb{R}^m : \langle w, x \rangle = 0 \text{ για κάθε } w \in W\}.$$

οπότε ισχύει ότι $\dim W + \dim W^{\perp} = m$ και επίσης $\dim \mathrm{Im} A + \dim W^{\perp} = m$, και αφού $\dim \mathrm{Im}(A)$ είναι η διάσταση του χώρου στηλών, έχουμε μια διαφορετική απόδειξη ότι η διάσταση του χώρου γραμμών ταυτίζεται με αυτή του χώρου στηλών.

Άσκηση VII.5.9. Δίνεται ένα σύνολο $u_1, \ldots, u_n$ ορθογώνιων διανυσμάτων σε διανυσματικό χώρο V με εσωτερικό γινόμενο. Για κάθε διάνυσμα $v \in V$ ισχύει

$$\sum_{\nu=1}^{n} \frac{|\langle v, u_\nu \rangle|^2}{\|u_\nu\|} \leq \|v\|^2$$

με την ισότητα να ισχύει μόνο αν

$$v = \sum_{\nu=1}^{n} \frac{\langle v, u_\nu \rangle}{\|u_\nu\|} u_\nu.$$

Λύση VII.5.10. Θεωρούμε τον χώρο $W = \langle u_1, \ldots, u_n \rangle$. Γράφουμε το $v \in V$ ως

$$v = \sum_{\nu=1}^{n} \frac{\langle v, u_\nu \rangle}{\|u_\nu\|} u_\nu + w'$$

με $w' \in W^{\perp}$. Αφού για $w = \sum_{\nu=1}^{n} \frac{\langle v, u_\nu \rangle}{\|u_\nu\|} u_\nu$ έχουμε ότι $\langle w, w' \rangle = 0$ από το πυθαγόρειο θεώρημα έχουμε

$$\|v\|^2 = \|w\|^2 + \|w'\|^2 \geq \sum_{\nu=1}^{n} \frac{|\langle v, u_\nu \rangle|^2}{\|u_\nu\|}.$$

Η ισότητα ισχύει αν και μόνο αν $w' = 0$.

Άσκηση VII.5.11. Με το εσωτερικό γινόμενο στον χώρο των συνεχών συναρτήσεων $C[0,1] = \{f : [0,1] \to \mathbb{R}, f \text{ συνεχής}\}$ να βρεθεί μια ορθοκανονική βάση του χώρου που παράγουν τα $\{1, x, x^2, x^3\}$.

Λύση VII.5.12. Ξεκινάμε την διαδικασία Gram-Schmidt. Θέτουμε $u_1 = 1$ και στην συνέχεια υπολογίζουμε

$$u_2 = x - \frac{\langle x^2, 1 \rangle}{\langle 1, 1 \rangle} 1 = x - \frac{1}{2}.$$

Στην συνέχεια θέτουμε

$$u_3 = x^2 - \frac{\langle x^2, 1 \rangle}{\langle 1, 1 \rangle} - \frac{\langle x^2, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 = x^3 - 1/3 - u_2$$

και

$$\begin{aligned} u_4 &= x^3 - \frac{\langle x^3, 1\rangle}{\langle 1, 1\rangle} - \frac{\langle x^3, u_2\rangle}{\langle u_2, u_2\rangle}u_2 - \frac{\langle x^3, u_3\rangle}{\langle u_3, u_3\rangle}u_3 \\ &= x^3 - 1/4 - 9/10u_2 - 3/2u_3. \end{aligned}$$

Στην συνέχεια υπολογίζουμε ότι

$$\|u_1\| = 1, \|u_2\| = \frac{1}{2\sqrt{3}}, \|u_3\| = \frac{1}{6\sqrt{5}}, \|u_4\| = \frac{1}{20\sqrt{7}},$$

οπότε έχουμε την ζητούμενη ορθοκανονική βάση

$$e_1 = 1, e_2 = 2\sqrt{3}u_2, e_3 = 6\sqrt{5}u_3, e_4 = 20\sqrt{7}u_4.$$

Άσκηση VII.5.13. Θεωρούμε τον διανυσματικό χώρο $\mathbb{C}^{n,n}$, με το εσωτερικό γινόμενο $\langle A, B\rangle = \mathrm{tr}(AB^*)$, όπου για πίνακα $B = (b_{ij})$ ο πίνακας $B^* = \overline{B}^t = (\overline{b}_{ji})$. Να βρεθεί το ορθογώνιο συμπλήρωμα του χώρου των διαγωνίων πινάκων.

Λύση VII.5.14. Παρατηρούμε ότι οι πίνακες E_{ij} οι οποίοι έχουν μια μονάδα στην i-γραμμή και στην j-στήλη και σε όλες τις άλλες θέσεις του πίνακα έχουν μηδενικά, αποτελούν μια ορθοκανονική βάση ως προς το εσωτερικό γινόμενο που μόλις ορίσαμε. Από την άλλη οι διαγώνιοι πίνακες είναι ο χώρος που παράγουν οι πίνακες $E_{11}, E_{22}, \ldots, E_{nn}$. Ένας πίνακας $A = (a_{ij})$ γράφεται ως

$$A = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}E_{ij}$$

ενώ

$$\langle A, E_{ii}\rangle = a_{ii}$$

Συνεπώς το ορθογώνιο συμπλήρωμα του χώρου των διαγωνίων πινάκων αποτελείται από τον διανυσματικό χώρο των πινάκων που έχουν $a_{11} = a_{22} = \cdots = a_{nn} = 0$.

VII.5.3 Η μέθοδος των ελαχίστων τετραγώνων

Υποθέτουμε ότι μια ποσότητα εξαρτάται με γραμμικό τρόπο από τις μεταβλητές $x_1, \ldots, x_n$ δηλαδή εκφράζεται ως συνάρτηση

$$f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n.$$

Θέλουμε να υπολογίσουμε τα $a_1, \ldots, a_n$. Σε ένα κόσμο που μπορούμε να κάνουμε μετρήσεις με όσο μεγάλη ακρίβεια θέλουμε θα αρκούσε να εκτελέσουμε n το πλήθος πειράματα μετρώντας κάθε φορά τις τιμές των $x_1, \ldots, x_n, f(x_1, x_2, \ldots, x_n)$ και στην συνέχεια να λύσουμε το σύστημα με αγνώστους $a_1, \ldots, a_n$. Το πρόβλημα σε αυτή την προσέγγιση είναι ότι δεν είναι δυνατόν ούτε να μετρήσουμε με απόλυτη ακρίβεια τα $x_1, \ldots, x_n, f(x_1, \ldots, x_n)$. Για αυτό τον λόγο κάνουμε μια σειρά πειραμάτων, πολύ περισσότερα από n και στην συνέχεια προσπαθούμε να προσδιορίσουμε όσο γίνεται καλύτερα τις τιμές $a_1, \ldots, a_n$ με τον ακόλουθο τρόπο.

Υποθέτουμε ότι στο i πείραμα έχουμε τις τιμές $x_{1i}, \ldots, x_{ni}$ και $y_i = f(x_{1i}, \ldots, x_{ni})$ για $1 \le i \le m$, $m > n$ οπότε έχουμε το σύστημα

$$\begin{pmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & & \vdots \\ x_{1m} & x_{2m} & \cdots & x_{nm} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Το παραπάνω σύστημα είναι αδύνατο και αυτό οφείλεται στα λάθη τα οποία υπεισέρχονται στις πειραματικές μετρήσεις. Θέλουμε να προσδιορίσουμε τα $a_1, \ldots, a_n$ ελαχιστοποιώντας την ποσότητα

$$\sum_{i=1}^{m}(a_1x_{1i} + a_2x_{2i} + \cdots a_nx_{ni} - y_i)^2.$$

Αν τον πίνακα

$$X = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & & \vdots \\ x_{1m} & x_{2m} & \cdots & x_{nm} \end{pmatrix}$$

τότε ζητάμε να διαλέξουμε διάνυσμα $a = (a_1, \ldots, a_n)^t$ ώστε να ελαχιστοποιείται η απόσταση $\|Xa - y\|$, όπου $y = (y_1, \ldots, y_m)^t$. Δηλαδή ψάχνουμε να βρούμε την ελάχιστη απόσταση του y από τον χώρο που παράγουν οι στήλες του πίνακα . Με άλλα λόγια το σύστημα είναι αδύνατο δηλαδή το y δεν ανήκει στην εικόνα της γραμμικής συνάρτησης που ορίζει ο πίνακας X, οπότε ψάχνουμε να βρούμε το πλησιέστερο διάνυσμα της εικόνας στο y. Συνεπώς αν $x^1, \ldots, x^{n'}$, $n \leq n'$ είναι μια βάση του χώρου στηλών του πίνακα X ψάχνουμε για $v = X'a$ στον χώρο στηλών ώστε

$$\langle v - y, x^i \rangle = 0 \text{ για κάθε } i = 1, \ldots, n'.$$

Δηλαδή έχουμε το σύστημα

$$a_1\langle x^1, x^i \rangle + a_2\langle x^2, x^i \rangle + \cdots + a_{n'}\langle x^{n'}, x^i \rangle = \langle y, x^i \rangle \text{ για } i = 1, \ldots, n', \tag{VII.12}$$

με

$$\langle y, x^i \rangle = \sum_{\nu=1}^{m} x_{i\nu}y_\nu \quad \langle x^j, x^i \rangle = \sum_{\nu=1}^{m} x_{j\nu}x_{i,\nu}.$$

αυτό είναι ένα σύστημα n'-εξισώσεων με n'-αγνώστους το οποίο έχει μοναδική λύση αρκεί ο πίνακας

$$A = (a_{ij}) \text{ με } a_{ij} = \langle x^j, x^i \rangle$$

να είναι αντιστρέψιμος. Παρατηρούμε ότι ο πίνακας $A = (X')^tX'$, όπου X' είναι ο πίνακας που έχει σαν στήλες τα $x^1, \ldots, x^{n'}$. Στο επόμενο λήμμα θα δείξουμε ότι ο πίνακας αυτός είναι αντιστρέψιμος και συνεπώς έχουμε ότι οι λύσεις το συστήματος (VII.12) δίνονται από

$$(a_1, \ldots, a_{n'})^t = ((X')^tX')^{-1}(X')^t(y_1, \ldots, y_m)^t.$$

Λήμμα VII.5.15. *Έστω πίνακας X' πίνακας $m \times n'$ με στήλες τις $x^1, \ldots, x^{n'}$. Το σύνολο στηλών $x^1, \ldots, x^{n'}$ είναι γραμμικά ανεξάρτητο αν και μόνο αν ο πίνακας $(X')^t(X')$ είναι αντιστρέψιμος.*

Απόδειξη. Ας υποθέσουμε ότι ο $n' \times n'$ πίνακας $(X')^tX'$ δεν είναι αντιστρέψιμος. Τότε υπάρχει μη-μηδενικό $v \in \mathbb{R}^{n',1}$ με $(X')^tX'v = \mathbf{0}_{n',1}$ και συνεπώς και το $v^t(X')^tX'v = 0$ άρα $\langle X'v, X'v \rangle = \|X'v\| = 0$ το οποίο μας δίνει ότι $X'v = \mathbf{0}_{\mathbb{R}^m}$ συνεπώς και ο πίνακας X' έχει μη τετριμμένο πυρήνα, άρα οι στήλες του είναι γραμμικά εξαρτημένες.

Αντιστρόφως, κάθε σχέση εξάρτησης μεταξύ των στηλών δίνει ένα μη-μηδενικό $w \in \mathbb{R}^n$, ώστε $'w = 0$. Άρα $(X')^tXw = 0$ συνεπώς ο $n' \times n'$ πίνακας $(X')^tX$ δεν μπορεί να είναι αντιστρέψιμος. $\square$

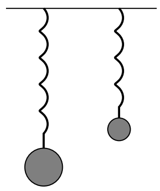
Πόρισμα VII.5.16. *Στην περίπτωση που έχουμε μια εξίσωση της μορφής* $f(x) = ax$ *και ζητούμε να προσδιορίσουμε το* a*, δηλαδή θέλουμε να βρούμε την «καλύτερη» ευθεία της μορφής* $y = ax$ *που να προσεγγίζει τα σημεία* $(x_1, y_1), \dots, (x_n, y_n)$ *τότε θέτουμε* $x = (x_1, \dots, x_n)^t$ $y = (y_1, \dots, y_n)^t$ *και το* a *δίνεται από το*

$$a = (x^t x)^{-1}(x^t y) = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}.$$

Παράδειγμα VII.5.17. Είναι γνωστό ότι η δύναμη που ασκεί ένα ελατήριο όταν τεντωθεί είναι ανάλογη της μετατόπισης, δηλαδή

$$F = -kx$$

για κάποια σταθερά k. Θέλουμε να την μετρήσουμε την σταθερά k για κάποιο συγκεκριμένο ελατήριο. Υποβάλουμε το ελατήριο σε μια σειρά από παραμορφώσεις και μετράμε την δύναμη που ασκεί αυτό και πειραματικά λογαριάζουμε τον πίνακα: Ας θεωρήσουμε ότι έχουμε τις εξής μετρήσεις:



Αποκλίση (x)	Δύναμη (F)
0.0	0.0
0.1	5.2
0.2	10.5
0.3	15.7
0.4	20.9

από όπου υπολογίζουμε ότι

$$\sum_{i=1}^{5} x_i^2 = 0.01 + 0.04 + 0.09 + 0.16 = 0.3$$

$$\sum_{i=1}^{5} F_i x_i = 0.0 + 0.52 + 2.1 + 4.71 + 8.36 = 15.69$$

από όπου έχουμε ότι $k = \frac{15.69}{0.3} \approx 52.3$

Παράδειγμα VII.5.18. Η πλησιέστερη έλλειψη σε μία σειρά σημείων. Αν μας δοθούν μια σειρά σημείων στον επίπεδο μπορούμε να υπολογίσουμε την πλησιέστερη έλλειψη σε αυτά;

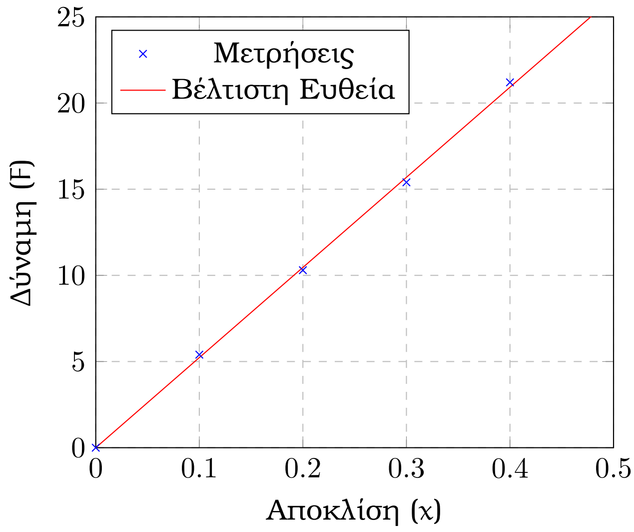
Αυτό είναι ένα πρόβλημα το οποίο έλυσε ο Gauss σε ηλικία 24 ετών με την μέθοδο των ελαχίστων τετραγώνων την οποία είχε επινοήσει ο ίδιος. Οι αστρονόμοι είχαν παρατηρήσει ένα νέο πλανήτη τον αστεροειδή «Δήμητρα» και είχαν δεδομένα για την θέση του μια χρονική στιγμή αλλά δεν ήταν δυνατόν να τον ξανά παρατηρήσουν. Ο Gauss υπολόγισε την τροχιά της Δήμητρας και είπε στους αστρονόμους πότε και που να στρέψουν τα τηλεσκόπια τους ώστε να ξαναδούν τον χαμένο αστεροειδή. Η ανακάλυψη της Δήμητρας ήταν ένας θρίαμβος των Μαθηματικών και του νεαρού τότε Gauss. Ο Gauss έλυσε το πρόβλημα στον χώρο, εμείς για χάρη ευκολίας θα περιοριστούμε σε ένα πρόβλημα στο επίπεδο.

Μια έλλειψη δίνεται ως ένα σύνολο σημείων $(x, y) \in \mathbb{R}^2$ που ικανοποιούν μια εξίσωση της μορφής

$$x^2 + By^2 + Cxy + Dx + Ey + F = 0.$$

Η παραπάνω εξίσωση μπορεί να αντιστοιχεί και σε άλλες κωνικές τομές, κύκλους, παραβολές και υπερβολές. Αν μας δοθούν τα σημεία $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ τότε υπάρχει μια έλλειψη η οποία περνάει από αυτά αν και μόνο αν

$$x_i^2 + By_i^2 + Cx_i y_i + Dx_i + Ey_i + F = 0 \text{ για κάθε } 1 \le i \le n.$$



Σχήμα VII.6: Μετρήσεις και βέλτιστη ευθεία

Μπορεί να υπάρχουν λάθη στις συντεταγμένες των σημείων και να μην ανήκουν όλα σε μια εξίσωση της παραπάνω μορφής. Ο στόχος μας είναι να υπολογίσουμε την πλησιέστερη κωνική τομή στα δεδομένα σημεία. Το πρόβλημα μπορεί να γραφεί σε μορφή πινάκων ως εξής

$$\begin{pmatrix} y_1^2 & x_1y_1 & x_1 & y_1 & 1 \\ y_2^2 & x_2y_2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_n^2 & x_ny_n & x_n & y_n & 1 \end{pmatrix} \begin{pmatrix} B \\ C \\ D \\ E \\ F \end{pmatrix} = - \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{pmatrix}$$

Στο παραπάνω πρόβλημα τα x_i, y_i είναι γνωστά και ζητούμε να υπολογίσουμε τα B, C, D, E, F. Έστω

$$A = \begin{pmatrix} y_1^2 & x_1y_1 & x_1 & y_1 & 1 \\ y_2^2 & x_2y_2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_n^2 & x_ny_n & x_n & y_n & 1 \end{pmatrix} \text{ και } b = - \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{pmatrix}$$

Έστω το σύνολο των σημείων του επιπέδου

$$\{(8, 2), (-6.07, -3.18), (-1.53, -6.82), (4.39, -4.74), (5.27, 0.66), (3, 5), (-5.31, 6.38)\}.$$

Υπολογίζουμε τους πίνακες A, b

$$A = \begin{pmatrix} 4 & -16 & -8 & 2 & 1 \\ 10.1124 & 19.3026 & -6.07 & -3.18 & 1 \\ 46.5124 & 10.4346 & -1.53 & -6.82 & 1 \\ 22.4676 & -20.8086 & 4.39 & -4.74 & 1 \\ 0.4356 & 3.4782 & 5.27 & 0.66 & 1 \\ 25 & 15 & 3 & 5 & 1 \\ 40.7044 & -33.8778 & -5.31 & 6.38 & 1 \end{pmatrix} \text{ και } b = \begin{pmatrix} -64 \\ -36.8449 \\ -2.3409 \\ -19.2721 \\ -27.7729 \\ -9 \\ -28.1961 \end{pmatrix}$$

Στην συνέχεια υπολογίζουμε

$$A^tA = \begin{pmatrix} 5068.49 & -853.446 & -204.758 & -62.8869 & 149.232 \\ -853.446 & 2555.27 & 146.74 & -204.758 & -22.471 \\ -204.758 & 146.74 & 187.427 & -22.471 & -8.25 \\ -62.8869 & -204.758 & -22.471 & 149.232 & -0.7 \\ 149.232 & -22.471 & -8.25 & -0.7 & 7. \end{pmatrix}$$

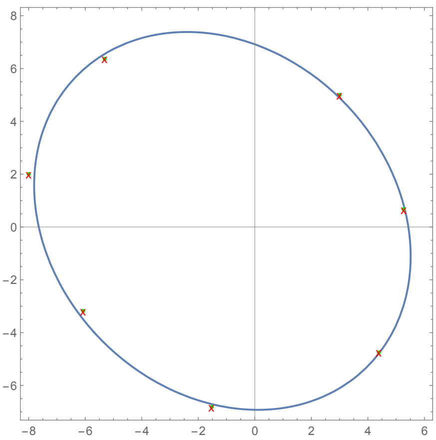
Τέλος έχουμε ότι

$$\begin{pmatrix} B \\ C \\ D \\ E \\ F \end{pmatrix} = (A^t A)^{-1} A^t b = \begin{pmatrix} 0.866743 \\ 0.350403 \\ 2.21155 \\ 0.000974531 \\ -41.5219 \end{pmatrix}$$

όποτε καταλήγουμε στην εξίσωση της έλλειψης:

$$x^2 + 0.866743y^2 + 0.350403xy + 2.21155x + 0.000974531y - 41.5219 = 0.$$

Στο σχήμα VII.7 έχουμε σχεδιάσει την παραπάνω εξίσωση μαζί με τα αρχικά σημεία.



Σχήμα VII.7: Η βέλτιστη έλλειψη από μία επιλογή σημείων

VII.6 Η συζυγής γραμμική συνάρτηση

Ορισμός VII.6.1. Ο δυικός χώρος ενός $\mathbb{F}$-διανυσματικού χώρου V είναι ο χώρος των γραμμικών συναρτήσεων

$$\mathrm{Hom}(V, \mathbb{F}) := \{f : V \to \mathbb{F}, f \text{ είναι γραμμική}\}.$$

Τα στοιχεία του δυικού χώρου λέγονται και γραμμικά συναρτησοειδή.

Είναι σαφές ότι είναι και αυτός ένας διανυσματικός χώρος με πράξεις

$$\begin{aligned} + : V^* \times V^* &\longrightarrow V^* \\ (f, g) &\longmapsto f + g : v \to (f+g)(v) = f(v) + g(v) \end{aligned}$$

$$\begin{aligned} \cdot : \mathbb{F} \times V^* &\longrightarrow V^* \\ (\lambda, g) &\longmapsto \lambda g : v \to (\lambda f)(v) = \lambda f(v) \end{aligned}$$

Παρατήρηση VII.6.2. Κάθε στοιχείο $a \in V$ ορίζει ένα στοιχείο στο V^*, το $v \mapsto \langle v, a \rangle$. Θα δείξουμε ότι σε πεπερασμένης διάστασης χώρους ισχύει και το αντίθετο.

Πρόταση VII.6.3. *Θεωρούμε ένα διανυσματικό χώρο πεπερασμένης διάστασης* V *και* $f \in V^*$. *Τότε υπάρχει μοναδικό* $a \in V$ *ώστε*

$$f(v) = \langle v, a\rangle, \text{ για κάθε } v \in V.$$

Απόδειξη. Θεωρούμε μια ορθοκανονική βάση $\{e_1, \ldots, e_n\}$ του V. Θέτουμε

$$a = \sum_{j=1}^{n} \overline{f(e_j)} e_j,$$

και υπολογίζουμε

$$\langle e_i, a\rangle = \left\langle e_i, \sum_{j=1}^{n} \overline{f(e_j)} e_j \right\rangle = f(e_i)$$

από όπου προκύπτει το ζητούμενο. □

Παρατήρηση VII.6.4. Σε άπειρης διάστασης διανυσματικούς χώρους η παραπάνω πρόταση δεν είναι αληθείς. Θεωρούμε τον διανυσματικό χώρο $V = \mathbb{C}[x]$, με εσωτερικό γινόμενο

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)} dt.$$

Όπως είδαμε στην άσκηση VII.2.11 για $f = \sum_\nu a_\nu x^\nu$, $g = \sum_\mu b_\mu x^\mu$

$$\langle f, g\rangle = \sum_{\nu,\mu} a_\nu \overline{b_\mu} \frac{1}{\nu + \mu + 1}.$$

Για ένα σταθερό $z \in \mathbb{C}$ Θεωρούμε το στοιχείο του V^*

$$\begin{aligned} \mathbb{C}[x] &\longrightarrow \mathbb{C} \\ f &\longmapsto \mathrm{eval}(f) = f(z). \end{aligned}$$

Αυτή είναι μια γραμμική συνάρτηση δηλαδή

$$\mathrm{eval}(f+g) = (f+g)(z) = f(z) + g(z) = \mathrm{eval}(f) + \mathrm{eval}(g) \text{ για κάθε } f, g \in \mathbb{C}[x]$$

και

$$\mathrm{eval}(\lambda f) = \lambda f(z) = \lambda \mathrm{eval}(f) \text{ για κάθε } \lambda \in \mathbb{C}, f \in \mathbb{C}[x].$$

Υπάρχει πολυώνυμο $g \in \mathbb{C}[x]$ ώστε

$$\mathrm{eval}(f) = \langle f, g\rangle \text{ για κάθε } f \in \mathbb{C}[x];$$

Η απάντηση είναι όχι. Πράγματι, υποθέτουμε ότι το $f(x) = (x - z)f_1(x)$. Τότε έχουμε

$$0 = f(z) = \mathrm{eval}(f) = \int_0^1 f(t)\overline{g(t)} dt = \int_0^1 (x-z) f_1(t)\overline{g(t)} dt,$$

και η τελευταία σχέση είναι αληθής για κάθε $f_1(x)$ άρα και για $f_1(x) = \overline{x - z}\overline{g(x)}$, το οποίο δίνει ότι

$$0 = \int_0^1 |t - z|^2 |g(t)|^2 dt$$

το οποίο μπορεί να συμβεί μόνο αν $g = 0$. Όμως το eval δεν είναι το μηδενικό συναρτησοειδές, άτοπο.

Το αποτέλεσμα είναι αναμενόμενο, δεν είναι δυνατόν ένα πολυώνυμο g, το οποίο έχει πεπερασμένη πληροφορία διάστασης $\deg g + 1$ να περιγράψει μια γραμμική συνάρτηση που έχει πληροφορία στο πως δρα σε ένα άπειρης διάστασης χώρο. Στο μάθημα της συναρτησιακής ανάλυσης θα δείτε ότι με επιπλέον συνθήκες, για παράδειγμα την πληρότητα της μετρικής, δηλαδή εξασφαλίζοντας ότι κάθε ακολουθία Cauchy ως προς την μετρική $\|v\| = \langle v, v\rangle^{1/2}$ συγκλίνει, έχουμε ένα θεώρημα αναπαράστασης των "συνεχών" συναρτησοειδών f στην μορφή $f(\cdot) = \langle \cdot, v\rangle$ και για απειροδιάστατους χώρους[1].

[1] https://en.wikipedia.org/wiki/Riesz_representation_theorem

Παρατήρηση VII.6.5. Έστω $f : V \to \mathbb{F}$ ένα στοιχείο του δυικού χώρου. Θέτουμε $W = \ker(f)$. Αν το $f \neq 0$, τότε $\mathrm{rk}(f) = 1$ και $\dim W = \dim V - 1$. Θεωρούμε την ανάλυση $V = W \oplus W^{\perp}$. Η τιμή του f προσδιορίζεται από την τιμή στο $W^{\perp}$. Αν $p_{W^{\perp}} : V \to W^{\perp}$ είναι η ορθογώνια προβολή, τότε $f(v) = f(p_{W^{\perp}}(v))$ για κάθε $v \in V$. Αν $f \neq 0$, και w ένα μη μηδενικό διάνυσμα του $W^{\perp}$, τότε

$$p_{W^{\perp}}(v) = \frac{\langle v, w\rangle}{\|w\|^2} w \text{ για κάθε } v \in V.$$

Συνεπώς,

$$f(v) = \langle v, w\rangle \frac{f(w)}{\|w\|^2} = \left\langle v, \overline{\frac{f(w)}{\|w\|^2}} w \right\rangle \text{ για κάθε } v \in W.$$

Λήμμα VII.6.6. *Αν για κάθε $v \in V$ ισχύει*

$$\langle v, w_1\rangle = \langle v, w_2\rangle$$

τότε $w_1 = w_2$.

Απόδειξη. Έχουμε ότι για κάθε $v \in V$ $\langle v, w_1 - w_2\rangle = 0$, οπότε αρκεί να πάρουμε $v = w_1 - w_2$. □

Θεώρημα VII.6.7. *Έστω V πεπερασμένης διάστασης διανυσματικός χώρος με εσωτερικό γινόμενο. Για κάθε γραμμική συνάρτηση $T : V \to V$, υπάρχει μοναδική γραμμική συνάρτηση T^* ώστε*

$$\langle Tv, w\rangle = \langle v, T^* w\rangle \text{ για κάθε } v, w \in V.$$

Απόδειξη. Σταθεροποιούμε ένα $w \in V$. Τότε η συνάρτηση

$$\begin{aligned} V &\longrightarrow \mathbb{F} \\ v &\longmapsto \langle Tv, w\rangle \end{aligned}$$

είναι ένα στοιχείο του δυικού χώρου και συνεπώς υπάρχει μοναδικό $w' \in V$, το οποίο θα το συμβολίζουμε με $T^* w$, ώστε

$$\langle Tv, w\rangle = \langle v, T^* w\rangle.$$

Θα πρέπει να δείξουμε ότι το $T^* w$ είναι μια γραμμική συνάρτηση του w. Πράγματι, για $\lambda_1, \lambda_2 \in \mathbb{F}$ και $w_1, w_2 \in V$ έχουμε

$$\begin{aligned} \langle v, T^*(\lambda_1 w_1 + \lambda_2 w_2)\rangle &= \langle Tv, \lambda_1 w_1 + \lambda_2 w_2\rangle \\ &= \overline{\lambda_1}\langle Tv, w_1\rangle + \overline{\lambda_2}\langle Tv, w_2\rangle \\ &= \overline{\lambda_1}\langle v, T^* w_1\rangle + \overline{\lambda_2}\langle v, T^* w_2\rangle \\ &= \langle v, \lambda_1 T^*(w_1) + \lambda_2 T^*(w_2)\rangle, \end{aligned}$$

από όπου έχουμε, σύμφωνα με το λήμμα VII.6.6, ότι

$$T^*(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 T^*(w_1) + \lambda_2 T^*(w_2).$$

□

Θεώρημα VII.6.8. *Σε ένα πεπερασμένης διάστασης διανυσματικό χώρο* V *θεωρούμε τον χώρο*

$$\mathrm{End}(V) = \{T : V \to V : T \text{ γραμμική}\}.$$

Η συνάρτηση $T \mapsto T^*$ *είναι μια συνάρτηση* $\mathrm{End}(V) \to \mathrm{End}(V)$ *η οποία ικανοποιεί*

1. $(T_1 + T_2)^* = T_1^* + T_2^*$, *για κάθε* $T_1, T_2 \in \mathrm{End}(V)$
2. $(\lambda T)^* = \overline{\lambda} T^*$, *για κάθε* $\lambda \in \mathbb{C}$ *και κάθε* $T \in \mathrm{End}(V)$
3. $(T_1 T_2)^* = T_2^* T_1^*$ *για κάθε* $T_1, T_2 \in \mathrm{End}(V)$
4. $(T^*)^* = T$ *για κάθε* $T \in \mathrm{End}(V)$

Απόδειξη. Για το 1. παρατηρούμε ότι για κάθε $v, w \in V$ έχουμε

$$\langle v, (T_1 + T_2)^* w\rangle = \langle (T_1 + T_2)v, w\rangle = \langle T_1 v, w\rangle + \langle T_2 v, w\rangle = \langle v, T_1^* w\rangle + \langle v, T_2^* w\rangle = \langle v, (T_1^* + T_2^*)w\rangle$$

οπότε η ισότητα προκύπτει λόγω του μονοσημάντου του $(T_1 + T_2)^*$.

Το 2. προκύπτει με παρόμοιο τρόπο

$$\langle v, (\lambda T)^* w\rangle = \langle \lambda T v, w\rangle = \lambda \langle T v, w\rangle = \lambda \langle v, T^* w\rangle = \langle v, \overline{\lambda} T^* w\rangle.$$

Για το 3. έχουμε

$$\langle v, (T_1 T_2)^* w\rangle = \langle T_1 T_2 v, w\rangle = \langle T_2 v, T_1^* w\rangle = \langle v, T_2^* T_1^* w\rangle.$$

Για το 4. υπολογίζουμε

$$\langle v, (T^*)^* w\rangle = \langle T^* v, w\rangle = \overline{\langle w, T^* v\rangle} = \overline{\langle T w, v\rangle} = \langle v, T w\rangle.$$

□

Πρόταση VII.6.9. *Σε ένα πεπερασμένης διάστασης διανυσματικό χώρο με εσωτερικό γινόμενο και* $B = \{e_1, \ldots, e_n\}$ *μια διατεταγμένη ορθοκανονική βάση, αν* $(T, B, B) = (a_{ij})$ *είναι ο πίνακας μιας γραμμικής συνάρτησης* $T : V \to V$, *έχουμε ότι* $a_{ij} = \langle T e_j, e_i\rangle$.

Απόδειξη. Εκφράζουμε κάθε $v \in V$ ως γραμμικό συνδυασμό των στοιχείων της ορθοκανονικής βάσης:

$$v = \sum_{i=1}^{n} \langle v, e_i\rangle e_i.$$

Ο πίνακας του T υπολογίζεται από την σχέση

$$T e_j = \sum_{i=1}^{n} a_{i,j} e_i$$

άρα

$$\langle T e_j, e_i\rangle = \sum_{i=1}^{n} \langle a_{ij} e_j, e_i\rangle = a_{ij}.$$

□

Πόρισμα VII.6.10. *Σε κάθε ορθοκανονική βάση* B *του πεπερασμένης διάστασης χώρου* V*, και κάθε γραμμική* $T : V \to V$ *ισχύει ότι*

$$(T^*, B, B) = (T, B, B)^*,$$

όπου για ένα πίνακα A *θέτουμε* $A^* = \overline{A}^t$*, δηλαδή τον ανάστροφο στον οποίο έχουμε εφαρμόσει την μιγαδική συζυγία σε κάθε στοιχείο του.*

Απόδειξη. Έστω $B = \{e_1, \ldots, e_n\}$ μία ορθοκανονική βάση του V και $A = (a_{ij}) = (T, B, B)$, $B = (b_{ij}) = (T^*, B, B)$. Έχουμε ότι

$$a_{ij} = \langle Te_j, e_i \rangle, \qquad b_{ij} = \langle T^* e_j, e_i \rangle.$$

Συνεπώς

$$b_{ij} = \langle T^* e_j, e_i \rangle = \langle e_j, Te_i \rangle = \overline{\langle Te_i, e_j \rangle} = \overline{a_{ji}},$$

από όπου προκύπτει το ζητούμενο. □

Παράδειγμα VII.6.11. Αν θεωρήσουμε τον χώρο $\mathbb{C}^n$ των διανυσμάτων στηλών, τότε η συνάρτηση

$$\begin{aligned} \mathbb{C}^n \times \mathbb{C}^n &\longrightarrow \mathbb{C} \\ (x, y) &\longmapsto y^* x = \langle x, y \rangle \end{aligned}$$

ορίζει το ερμητιανό εσωτερικό γινόμενο. Μια γραμμική συνάρτηση δίνεται μέσω ενός $n \times n$ πίνακα A, δηλαδή $x \mapsto Ax$. Έχουμε ότι

$$\langle Ax, y \rangle = y^* Ax = (A^* y)^* x = \langle x, A^* y \rangle.$$

Άσκηση VII.6.12. Θεωρούμε τον χώρο $\mathbb{C}^{n,n}$ με το εσωτερικό γινόμενο

$$\langle A, B \rangle = \mathrm{tr}(B^* A).$$

Για ένα σταθερό πίνακα M θεωρούμε την γραμμική συνάρτηση

$$\begin{aligned} L_M : \mathbb{C}^{n,n} &\longrightarrow \mathbb{C}^{n,n} \\ A &\longmapsto MA \end{aligned}$$

Δείξτε ότι $L_M^* = L_{M^*}$.

Λύση VII.6.13. Υπολογίζουμε ότι

$$\begin{aligned} \langle L_M(A), B \rangle &= \mathrm{tr} B^* MA) = \mathrm{tr} MAB^* = \mathrm{tr}(AB^* M) \\ &= \mathrm{tr}(A(M^* B)^*) = \mathrm{tr}(M^* B)^* A = \langle A, M^* B \rangle = \langle A, L_{M^*} B \rangle. \end{aligned}$$

Άσκηση VII.6.14. Θεωρούμε τον διανυσματικό χώρο $\mathbb{C}[x]$ με το εσωτερικό γινόμενο

$$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)} dt.$$

- Δείξτε ότι για σταθερό $f \in \mathbb{C}[x]$ για την γραμμική συνάρτηση

$$\begin{aligned} M_f : \mathbb{C}[x] &\longrightarrow \mathbb{C}[x] \\ g &\longmapsto fg \end{aligned}$$

  ισχύει ότι $M_f^* = M_{\overline{f}}$.

- Για την γραμμική συνάρτηση

$$\begin{aligned} D : \mathbb{C}[x] &\longrightarrow \mathbb{C}[x] \\ g &\longmapsto Dg = g' \end{aligned}$$

δεν υπάρχει γραμμική συνάρτηση D^*, ώστε $\langle Df, g\rangle = \langle f, D^*g\rangle$ για κάθε $f, g \in \mathbb{C}[x]$.

Λύση VII.6.15. Υπολογίζουμε ότι

$$\begin{aligned} \langle M_f g, h\rangle = \langle fg, h\rangle &= \int_0^1 f(t)g(t)\overline{h(t)}dt \\ &= \int_0^1 g(t)\overline{\overline{f(t)}h(t)}dt = \langle g, \overline{f}h\rangle = \langle g, M_{\overline{f}}(h)\rangle. \end{aligned}$$

Για την συνάρτηση παραγώγισης D παρατηρούμε ότι

$$\langle Df, g\rangle = f(1)g(1) - f(0)g(0) - \langle f, Dg\rangle.$$

Για g σταθερό υποθέτουμε ότι υπάρχει πολυώνυμο D^* ώστε $\langle Df, g\rangle = \langle f, D^*g\rangle$. Τότε

$$\langle f, D^*g\rangle = f(1)g(1) - f(0)g(0) - \langle f, Dg\rangle$$

ισοδύναμα

$$\langle f, D^*g + Dg\rangle = f(1)g(1) - f(0)g(0).$$

Παρατηρούμε ότι για g σταθερό το $L(f) = f(1)g(1) - f(0)g(0)$ είναι ένα στοιχείο του δυικού χώρου όπως στην παρατήρηση VII.6.4 και δεν υπάρχει $h \in \mathbb{C}[x]$ ώστε $L(f) = \langle f, h\rangle$ εκτός αν $L = 0$. Συνεπώς η ύπαρξη του D^*g μας δίνει ότι το $h = D^*g + Dg$ πρέπει να είναι μηδενικό και $g(0) = g(1) = 0$. Αν διαλέξουμε g ώστε $g(0) \neq 0$ ή $g(1) \neq 0$ το D^*g δεν ορίζεται και δεν υπάρχει συζυγής.

VII.7 Μοναδιαίες γραμμικές συναρτήσεις

Ορισμός VII.7.1. Θα λέμε ότι μία γραμμική συνάρτηση $: V \to W$ ανάμεσα σε διανυσματικούς χώρους V, W διατηρεί το εσωτερικό γινόμενο όταν

$$\langle Tv, Tv\rangle = \langle v, v\rangle, \text{ για κάθε } v \in V.$$

Ένας ισομορφισμός ανάμεσα σε χώρους με εσωτερικό γινόμενο θα είναι ένας ισομορφισμός που επιπλέον διατηρεί το εσωτερικό γινόμενο.

Παρατήρηση VII.7.2. Μια γραμμική συνάρτηση $V \to W$ που διατηρεί το εσωτερικό γινόμενο δεν μπορεί να έχει μη τετριμμένο πυρήνα, αφού $\|Tv\| = \|v\|$ και αν $v \in \ker$ τότε $v = \mathbf{0}$.

Συνεπώς αν $\dim V = \dim W$ κάθε γραμμική συνάρτηση $V \to W$ που διατηρεί το εσωτερικό γινόμενο είναι ισομορφισμός.

Παρατήρηση VII.7.3. Κάθε γραμμική συνάρτηση $T : V \to W$ που διατηρεί το εσωτερικό γινόμενο μεταφέρει ορθοκανονικές βάσεις σε ορθοκανονικές βάσεις. Αντιστρόφως αν $\{e_1, \dots, e_n\}$ είναι μια ορθοκανονική βάση του V, ώστε $\{Te_1, \dots, Te_n\}$ να είναι μια ορθοκανονική βάση του W, τότε για κάθε $v, w \in V$, έχουμε

$$v = \sum_{i=1}^{n} \lambda_i e_i, w = \sum_{j=1}^{n} \mu_j e_j$$

και

$$\langle v, w\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \overline{\mu}_j \langle e_i, e_j\rangle = \sum_{i=1}^{n} \lambda_i \overline{\mu}_j.$$

Ομοίως

$$\langle Tv, Tw\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\overline{\mu}_j\langle Te_i, Te_j\rangle = \sum_{i=1}^{n}\lambda_i\overline{\mu}_j,$$

δηλαδή

$$\langle v, w\rangle = \langle Tv, Tw\rangle.$$

Ορισμός VII.7.4. Μία μοναδιαία (unitary) γραμμική συνάρτηση είναι ένας ισομορφισμός που διατηρεί το εσωτερικό γινόμενο.

Παρατήρηση VII.7.5. Η σύνθεση δύο μοναδιαίων γραμμικών συναρτήσεων είναι μοναδιαία γραμμική συνάρτηση, η ταυτοτική συνάρτηση είναι μοναδιαία γραμμική συνάρτηση. Συνεπώς οι μοναδιαίες γραμμικές συναρτήσεις αποτελούν υποομάδα της ομάδας των αντιστρέψιμων γραμμικών συναρτήσεων.

Πρόταση VII.7.6. *Μία γραμμική συνάρτηση είναι μοναδιαία αν και μόνο αν είναι αντιστρέψιμη και* $U^{-1} = U^*$.

Απόδειξη. Υποθέτουμε ότι U είναι μοναδιαία, δηλαδή διατηρεί το εσωτερικό γινόμενο. Παρατηρούμε ότι για κάθε $v, w \in W$ και $U : V \to U$ μοναδιαία έχουμε

$$\langle Uv, w\rangle = \langle Uv, \mathrm{Id}_V w\rangle \overset{\mathrm{Id}_V = UU^{-1}}{=} \langle Uv, UU^{-1}w\rangle = \langle v, U^{-1}w\rangle$$

από όπου προκύπτει ότι $U^{-1} = U^*$.

Αντιστρόφως, αν $U^{-1} = U^*$, τότε

$$\langle Uv, Uw\rangle = \langle v, U^*Uw\rangle = \langle v, w\rangle \text{ για κάθε } v, w \in V.$$

□

Ορισμός VII.7.7. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ θα λέγεται μοναδιαίος (unitary) αν και μόνο αν για $A^* = \overline{A}^t$ ισχύει $AA^* = \mathbb{I}_n$.

Πρόταση VII.7.8. *Σε ένα διανυσματικό χώρο* V *πεπερασμένης διάστασης η γραμμική συνάρτηση* $T : V \to V$ *είναι μοναδιαία αν και μόνο αν ο πίνακας της ως προς μια ορθοκανονική βάση είναι είναι μοναδιαίος.*

Απόδειξη. Παρατηρούμε ότι αν $A = (T, B, B)$ είναι ο πίνακας της T ως προς μια ορθοκανονική βάση B τότε ο πίνακας (T^*, B, B) της T^* είναι ο $A^* = \overline{A}^t$. □

Ορισμός VII.7.9. Ένας πραγματικός ή μιγαδικός πίνακας θα λέγεται ορθογώνιος αν $AA^t = \mathbb{I}_n$.

Παρατήρηση VII.7.10. Ένας πραγματικός πίνακας είναι ορθογώνιος αν και μόνο αν είναι μοναδιαίος.

Πρόταση VII.7.11. *Ένας πίνακας είναι μοναδιαίος αν και μόνο αν οι γραμμές ή οι στήλες του αποτελούν ορθοκανονική βάση του $\mathbb{C}^n$ με το ερμητιανό εσωτερικό γινόμενο $\langle x, y\rangle = y^* x$.*

Απόδειξη. Γράφουμε τον πίνακα $A = (a^1, \dots, a^n)$, όπου a^i είναι η i-στήλη Παρατηρούμε ότι ο πίνακας $A^* = (c_{ij})$ έχει

$$c_{ij} = (a^i)^* a^j = \langle a^j, a^i\rangle.$$

Η συνθήκη $^* = \mathbb{I}_n$ είναι ισοδύναμη με το ότι $c_{ij} = \delta_{ij} = \langle a^j, a^i\rangle$ δηλαδή οι στήλες αποτελούν μια ορθοκανονική βάση.

Για να δείξουμε ότι και οι στήλες αποτελούν μια ορθοκανονική βάση παρατηρούμε ότι

$$A^* A = \mathbb{I}_n \Leftrightarrow (A^t)(A^t)^* = \mathbb{I}_n,$$

δηλαδή ο A είναι μοναδιαίος αν και μόνο αν ο A^t είναι μοναδιαίος. □

Η μέθοδος Gram-Schmidt έχει την παρακάτω ενδιαφέρουσα συνέπεια:

Πρόταση VII.7.12. *Για κάθε αντιστρέψιμο $n \times n$ πίνακα A υπάρχει μοναδικός κάτω τριγωνικός πίνακας B με θετικά στοιχεία στην διαγώνιο ώστε BA να είναι μοναδιαίος.*

Απόδειξη. Οι γραμμές $a_1, \dots, a_n$ του πίνακα είναι γραμμικά ανεξάρτητες αφού ο πίνακας A είναι αντιστρέψιμος. Εκτελούμε την μέθοδο Gram-Schmidt στις γραμμές $a_1, \dots, a_n$ και καταλήγουμε σε μία ορθογώνια βάση $u_1, \dots, u_n$ του $\mathbb{C}^n$. Κατά την διαδικασία παρατηρούμε ότι σε κάθε βήμα της αναδρομικής ορθοκανονικοποίησης τα $\{u_1, \dots, u_i\}$ αποτελούν μια ορθογώνια βάση του διανυσματικού χώρου $\langle a_1, \dots, a_i\rangle$ και ότι

$$u_k = a_k - \sum_{j=1}^{k-1} \frac{\langle a_k, u_j\rangle}{\|u_j\|^2} u_j$$

Δηλαδή, υπάρχουν μοναδικοί συντελεστές γ_{kj} ώστε

$$u_k = a_k - \sum_{j=1}^{k-1} \gamma_{kj} a_j.$$

Θεωρούμε τον unitary πίνακα U με γραμμές $u_1/\|u_1\|, \dots, u_n/\|u_n\|$ και B τον πίνακα που ορίζεται από

$$b_{kj} = \begin{cases} -\frac{\gamma_{kj}}{\|u_k\|} & \text{αν } j < k \\ \frac{1}{\|u_k\|} & \text{αν } j = k \\ 0 & \text{αν } j > k \end{cases}$$

Παρατηρούμε ότι ο B είναι κάτω τριγωνικός, έχει θετικές τιμές στην διαγώνιο και επίσης

$$\frac{u_k}{\|u_k\|} = \sum_{j=1}^{n} b_{kj} a_j, \text{ για κάθε } 1 \le k \le n,$$

δηλαδή $U = BA$.

Θα αποδείξουμε τώρα ότι η γραφή είναι μοναδική. Θεωρούμε το σύνολο $T^+(n)$ των $n \times n$ κάτω τριγωνικών πινάκων με θετικά στοιχεία στην διαγώνιο και με $U(n)$ το σύνολο των μοναδιαίων $n \times n$ πινάκων.

Υποθέτουμε ότι $B_1, B_2 \in T^+(n)$, ώστε $B_1 A, B_2 A \in U(n)$. Έχουμε ότι ο αντίστροφος κάθε μοναδιαίου πίνακα είναι μοναδιαίος και ότι το γινόμενο δύο μοναδιαίων πινάκων είναι μοναδιαίος. Συνεπώς,

$$(B_A)(B_2 A)^{-1} = B_1 B_2^{-1} \in U(n).$$

Επίσης παρατηρούμε ότι ο αντίστροφος κάθε πίνακα στο $T^+(n)$ ανήκει στο $T^+(n)$ όπως και το γινόμενο δύο πινάκων στο $T^+(n)$ ανήκει στο $T^+(n)$, δείτε την άσκηση VII.7.14.

Αφού $B_1B_2^{-1} \in U(n)$, έχουμε ότι $(B_1B_2^{-1})^{-1} = (B_1B_2^{-1})^*$. Δηλαδή ο πίνακας $({}_12^{-1})^{-1}$ είναι κάτω τριγωνικός, ως αντίστροφος στοιχείου στο $T^+(n)$ αλλά και άνω τριγωνικός ως ανάστροφος κάτω τριγωνικού. Άρα είναι διαγώνιος. Και αφού είναι μοναδιαίος τα στοιχεία της διαγωνίου του έχουν μέτρο 1 και είναι και ταυτόχρονα θετικά. Άρα $B_1B_2^{-1} = \mathbb{I}_n$. □

Παράδειγμα VII.7.13. Να βρεθεί κάτω τριγωνικός πίνακας B με θετικά στοιχεία στην διαγώνιο για τον πίνακα A

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

ώστε ο πίνακας BA να είναι μοναδιαίος.

Στο παράδειγμα VII.5.4 κάναμε μια ορθοκανονικοποίηση των γραμμών του πίνακα υπολογίζοντας ότι

$$\begin{aligned} u_1 = v_1, u_2 = v_2 - \frac{2}{7}v_1, u_3 &= v_3 - \frac{1}{2}v_1 + \frac{21}{27}u_2 \\ &= v_3 - \frac{1}{2}v_1 + \frac{21}{27}(v_2 - \frac{2}{7}v_1) \\ &= v_3 + \frac{21}{27}v_2 - \frac{13}{18}v_1 \end{aligned}$$

οπότε έχουμε τον πίνακα

$$B' = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{2}{7} & 1 & 0 \\ \frac{13}{18} & \frac{7}{9} & 1 \end{pmatrix}$$

ώστε

$$B'A = \begin{pmatrix} 1 & 2 & 3 \\ \frac{12}{7} & \frac{3}{7} & -\frac{6}{7} \\ -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} \end{pmatrix}$$

Ο τελευταίος πίνακας δεν είναι μοναδιαίος γιατί οι γραμμές είναι ορθογώνιες αλλά δεν έχουν μέτρο 1. Αντικαθιστώντας τον πίνακα B' με τον πίνακα

$$B = \begin{pmatrix} \frac{1}{\sqrt{14}} & 0 & 0 \\ -\frac{2\sqrt{7}}{7\sqrt{27}} & \sqrt{7/27} & 0 \\ \frac{13}{18}\sqrt{6} & \frac{7}{9}\sqrt{6} & \sqrt{6} \end{pmatrix}$$

έχουμε την ζητούμενη ανάλυση.

Άσκηση VII.7.14. Έστω $T^+(n)$ το σύνολο των $n \times n$ κάτω τριγωνικών πινάκων με θετικά στοιχεία στην διαγώνιο. Δείξτε ότι για κάθε $B_1, B_2 \in T^+(n)$ το γινόμενο $B_1B_2 \in T^+(n)$ και ότι για κάθε $B \in T^+(n)$ ισχύει ότι $B^{-1} \in T^+(n)$.

Λύση VII.7.15. Είναι σαφές ότι το γινόμενο δύο κάτω τριγωνικών πινάκων είναι κάτω τριγωνικός πίνακας αλλά και ότι το γινόμενο δύο πινάκων που ο καθένας έχει θετικά στοιχεία στην διαγώνιο θα έχει και αυτό θετικά στοιχεία στην διαγώνιο.

Θεωρούμε ένα κάτω τριγωνικό αντιστρέψιμο πίνακα B τον οποίο τον γράφουμε στην μορφή

$$B = \Delta(\mathbb{I}_n + L)$$

όπου ο Δ είναι διαγώνιος πίνακας με μή-μηδενικά στοιχεία ενώ ο L είναι κάτω τριγωνικός με μηδενικά στοιχεία στην διαγώνιο. Ισχύει ότι $L^n = 0$ (το οποίο το βλέπουμε είτε από την διαγώνια μορφή του πίνακα είτε εξυπνότερα από το Θεώρημα Caley-Hamilton). Τώρα υπολογίζουμε ότι

$$(\mathbb{I}_n + L)(\mathbb{I}_n - L + L^2 - L^3 + \cdots + (-1)^{n-1}L^{n-1}) = \mathbb{I}_n.$$

Άρα

$$B^{-1} = (\mathbb{I}_n + L)^{-1}\Delta^{-1} = (\mathbb{I}_n - L + L^2 - L^3 + \cdots + (-1)^{n-1}L^{n-1})\Delta^{-1}$$

ο οποίος είναι κάτω τριγωνικός, ως άθροισμα γινομένων κάτω τριγωνικών πινάκων. Η απαίτηση για θετικά στοιχεία στην διαγώνιο είναι εμφανής από τον υπολογισμό αντιστρόφου διαγώνιου πίνακα.

Άσκηση VII.7.16. Θεωρούμε τον χώρο $V = C[0,1]$ των συνεχών συναρτήσεων $[0,1] \to \mathbb{R}$, με το εσωτερικό γινόμενο

$$\langle f, g\rangle_1 = \int_0^1 f(t)g(t)dt.$$

και W τον ίδιο χώρο με το εσωτερικό γινόμενο

$$\langle f, g\rangle_2 = \int_0^1 f(t)g(t)dt.$$

Αφού δείξετε ότι το $\langle\cdot,\cdot\rangle_2$ είναι εσωτερικό γινόμενο δείξτε ότι η

$$\begin{aligned} T : W &\longrightarrow V \\ f(x) &\longmapsto xf(x) \end{aligned}$$

διατηρεί το εσωτερικό γινόμενο αλλά δεν είναι ισομορφισμός.

Λύση VII.7.17. Το ότι το $\langle\cdot,\cdot\rangle_2$ είναι εσωτερικό γινόμενο είναι μια εύκολη επαλήθευση.

Παρατηρούμε ότι

$$\langle f, g\rangle_2 = \int_0^1 f(t)g(t)t^2dt = \langle Tf, Tg\rangle_1.$$

Η συνάρτηση $: W \to V$ είναι 1-1 αλλά επειδή η διάστασεις των V, W δεν είναι πεπερασμένες δεν είναι επί. Πράγματι δεν υπάρχει τρόπος να πάρουμε την σταθερή συνάρτηση $1 : [0,1] \to \mathbb{R}$ ως γινόμενο $xg(x) = 1$, η συνάρτηση $g(x) = 1/x$ δεν ορίζεται στο 0.

Άσκηση VII.7.18. Να βρεθούν όλοι οι $n \times n$ μοναδιαίοι και ορθογώνιοι πίνακες για $n = 1, 2$.

Λύση VII.7.19. Ένας 1×1 πίνακας (a) είναι ορθογώνιος αν και μόνο αν $a^2 = 1$ δηλαδή αν και μόνο αν $a = \pm 1$. Ένας 1×1 πίνακας (a) είναι μοναδιαίος αν και μόνο αν $a\bar{a} = \|a\|^2 = 1$ δηλαδή το a είναι ένας μιγαδικός πάνω στον μοναδιαίο κύκλο, $a = e^{i\theta}$, $\theta \in \mathbb{R}$.

Έστω τώρα ένας 2×2 πίνακας

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ με } A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Αν ο πίνακας είναι ορθογώνιος, δηλαδή αν $AA^t = \mathbb{I}_n$, τότε $\det(A) = \pm 1$, αφού $\det A^t = \det A$. Συνεπώς ο πίνακας είναι ορθογώνιος αν

$$\det A = -1 \text{ και } \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -d & c \\ b & -a \end{pmatrix} \Rightarrow A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

$$\det A = 1 \text{ και } \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -c \\ b & a \end{pmatrix} \Rightarrow A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

και στις δύο περιπτώσεις ο υπολογισμός της ορίζουσας δίνει την σχέση $a^2 + b^2 = 1$.

Στην περίπτωση ενός πραγματικού πίνακα έχουμε ότι αυτός είναι ένας πίνακας της μορφής

$$A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Στην μιγαδική περίπτωση τώρα, παρατηρούμε ότι για μοναδιαίους πίνακες η σχέση $AA^* = I_2$, δίνει ότι $\det A\overline{\det A} = 1$ συνεπώς η ορίζουσα έχει μέτρο 1, δηλαδή $\det A = e^{i\theta}$, $\theta \in \mathbb{R}$. Επιπλέον ο πίνακας είναι μοναδιαίος αν και μόνο αν

$$\begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} -d & -b \\ -c & -a \end{pmatrix} \Rightarrow = \begin{pmatrix} a & b \\ -e^{i\theta}\overline{b} & e^{i\theta}\overline{a} \end{pmatrix}$$

και $|a|^2 + |b|^2 = 1$.

VII.8 Κανονικές γραμμικές συναρτήσεις

Σε αυτή την παράγραφο θα απαντήσουμε στο πότε μια γραμμική συνάρτηση $L : V \to V$, όπου ο V είναι ένας πεπερασμένης διάστασης διανυσματικός χώρος έχει μια ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα. Ας υποθέσουμε ότι $B = \{e_1, \dots, e_n\}$ είναι μια τέτοια βάση. Από την μία έχουμε ότι

$$Le_j = \lambda_j e_j$$

δηλαδή ο (L, B, B) είναι ένας διαγώνιος πίνακας, συνεπώς ο $(L^*, B, B) = (L, B, B)^*$ δηλαδή είναι και αυτός διαγωνοποιήσιμος ως προς την ίδια βάση. Αυτό σημαίνει ότι $L^*L = LL^*$, δείτε και την πρόταση VI.3.9.

Ορισμός VII.8.1. Έστω $L : V \to V$ γραμμική συνάρτηση, όπου ο V είναι πεπερασμένης διάστασης διανυσματικός χώρος. Η L θα λέγεται κανονική (normal) αν και μόνο αν $L^*L = LL^*$.

Ορισμός VII.8.2. Έστω $L : V \to V$ γραμμική συνάρτηση, όπου ο V είναι πεπερασμένης διάστασης διανυσματικός χώρος. Η L θα λέγεται αυτοσυζυγής ή ερμητιανή αν και μόνο αν $L^* = L$.

Παρατήρηση VII.8.3. Είναι σαφές ότι μια ερμητιανή γραμμική συνάρτηση όπως και μια μοναδιαία γραμμική συνάρτηση είναι κανονική.

Πρόταση VII.8.4. *Οι ιδιοτιμές μιας ερμητιανής γραμμικής συνάρτησης είναι όλες πραγματικές. Επίσης τα ιδιοδιανύσματα που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι κάθετα μεταξύ τους.*

Απόδειξη. Έστω v ένα ιδιοδιάνυσμα της $L \to L$ το οποίο αντιστοιχεί στην ιδιοτιμή $\lambda \in \mathbb{C}$. Έχουμε ότι

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Lv, v\rangle = \langle v, L^*v\rangle = \langle v, Lv\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$$

και αφού $\langle v, v\rangle \neq 0$ έχουμε ότι $\lambda = \overline{\lambda}$, δηλαδή $\lambda \in \mathbb{R}$.

Από την άλλη αν $Lv = \lambda v$ και $Lw = \mu w$ με $\mu \neq \lambda$ τότε

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle Lv, w\rangle = \langle v, L^*w\rangle = \langle v, \mu w\rangle = \overline{\mu}\langle v, w\rangle = \mu\langle v, w\rangle,$$

συνεπώς $\langle v, w\rangle = 0$. □

Λήμμα VII.8.5. *Έστω V ένας πεπερασμένης διάστασης διανυσματικός χώρος και $L : V \to V$ γραμμική συνάρτηση. Θεωρούμε ένα L-αναλλοίωτο υπόχωρο W. Το ορθογώνιο συμπλήρωμα $W^\perp$ είναι L^* αναλλοίωτο.*

Απόδειξη. Έχουμε ότι $Lw \in W$ για κάθε $w \in W$. Έστω $w' \in W^\perp$ δηλαδή $\langle w, w' \rangle = 0$ για κάθε $w \in W$. Θέλουμε να δείξουμε ότι $L^*w' \in W^\perp$ δηλαδή $\langle w, L^*w' \rangle = 0$ για κάθε $w \in W$. Όμως $\langle w, L^*w' \rangle = \langle Lw, w' \rangle = 0$ αφού $Lw \in W$. □

Θεώρημα VII.8.6. *Για κάθε αυτοσυζυγή γραμμική συνάρτηση $L : V \to V$, όπου ο V είναι πεπερασμένης διάστασης διανυσματικός χώρος υπάρχει ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα.*

Απόδειξη. Υποθέτουμε ότι $\dim V \geq 1$. Πάνω από το $\mathbb{C}$, έχουμε ότι ο L έχει μια ιδιοτιμή λ η οποία είναι μάλιστα πραγματική. Το αντίστοιχο ιδιοδιάνυσμα v μπορούμε να το κανονικοποιήσουμε στο $e = v/\|v\|$ ώστε να έχει μέτρο 1. Επιπλέον αν δουλεύουμε σε πραγματικούς χώρους μπορούμε να υποθέσουμε ότι το διάνυσμα v έχει πραγματικούς συντελεστές σε οποιαδήποτε βάση, διότι αν $A = (L, B, B) \in \mathbb{R}^{n,n}$ το ιδιοδιάνυσμα έχει συντεταγμένες $[v]_B \in \mathbb{R}^n$ που είναι λύση του συστήματος $(A - \lambda \mathbb{I}_n)x = \mathbf{0}$, το οποίο έχει μη τετριμμένες λύσεις από την επιλογή του λ ως ρίζα του χαρακτηριστικού πολυωνύμου.

Είναι προφανές ότι αν $\dim V = 1$ υπάρχει ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα το $e = v/\|v\|$ που μόλις κατασκευάσαμε. Θα αποδείξουμε το ζητούμενο με επαγωγή. Υποθέτουμε ότι η πρόταση είναι αληθής για όλες τις αυτοσυζυγείς γραμμικές συναρτήσεις σε διανυσματικούς χώρους διάστασης $n-1$. Έστω ότι έχουμε ένα διανυσματικό χώρο διάστασης n. Θεωρούμε την ανάλυση

$$V = \langle e \rangle \bigoplus \langle e \rangle^\perp$$

Από το λήμμα Είναι σαφές από το λήμμα VII.8.5 ότι ο

$$L_{\langle e \rangle^\perp} = L^*|_{\langle e \rangle^\perp} : \langle e \rangle^\perp \longrightarrow \langle e \rangle^\perp$$

περιορίζεται στον $\langle e \rangle^\perp$ σε μια αυτοσυζυγή συνάρτηση η οποία λόγω της επαγωγικής υπόθεσης έχει μια ορθοκανονική βάση $\{e_2, ..., e_n\}$ αποτελούμενη από ιδιοδιανύσματα. Η $\{e, e_2, \ldots, e_n\}$ είναι η ζητούμενη ορθοκανονική βάση αποτελούμενη από ιδιοδιανύσματα για όλο τον χώρο. □

Λήμμα VII.8.7. *Για μία κανονική γραμμική συνάρτηση $L : V \to V$ για $v \in V$ έχουμε $\|Lv\| = \|L^*v\|$.*

Απόδειξη. Έχουμε ότι

$$\|Lv\|^2 = \langle Lv, Lv \rangle = \langle v, L^*Lv \rangle = \langle v, LL^*v \rangle = \langle L^*v, L^*v \rangle = \|L^*v\|^2.$$

□

Πρόταση VII.8.8. *Έστω V πεπερασμένης διάστασης διανυσματικός χώρος και $L : V \to V$ κανονική γραμμική συνάρτηση. Τότε το $v \in V$ είναι ιδιοδιάνυσμα για την L με ιδιοτιμή λ αν και μόνο αν v είναι ιδιοδιάνυσμα για την T^* με ιδιοτιμή $\overline{\lambda}$.*

Απόδειξη. Για μια τιμή $\lambda \in \mathbb{C}$ έχουμε ότι η συνάρτηση $T = L - \lambda \mathrm{Id}_V$ είναι κανονική αφού $T^* = L^* - \bar{\lambda}\mathrm{Id}_V$. Υπολογίζουμε ότι

$$\begin{aligned} TT^* &= (L - \lambda\mathrm{Id}_V)(L^* - \bar{\lambda}\mathrm{Id}_V) = LL^* - \bar{\lambda}L - \lambda L^* - \lambda\bar{\lambda}\mathrm{Id}_V \\ &= L^*L - \bar{\lambda}L - \lambda L^* - \lambda\bar{\lambda}\mathrm{Id}_V = (L^* - \bar{\lambda}\mathrm{Id}_V)(L - \lambda\mathrm{Id}_V) = T^*T. \end{aligned}$$

Συνεπώς

$$\|(L - \lambda\mathrm{Id}_V)v\| = \|(L^* - \bar{\lambda}\mathrm{Id}_V)v)\|$$

από όπου προκύπτει ότι $(L - \lambda\mathrm{Id}_V)v = 0$ αν και μόνο αν $(L^* - \bar{\lambda}\mathrm{Id}_V)v = 0$, δηλαδή το ζητούμενο. □

Ορισμός VII.8.9. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ θα λέγεται κανονικός αν και μόνο αν $AA^* = A^*A$.

Πρόταση VII.8.10. *Σε έναν πεπερασμένης διάστασης διανυσματικό χώρο* V, *θεωρούμε μια γραμμική συνάρτηση* $L : V \to V$ *ώστε ο πίνακας της* (L, B, B) *ως προς μία ορθοκανονική βάση* B *να είναι άνω τριγωνικός. Η* L *είναι κανονική αν και μόνο αν* (L, B, B) *είναι διαγώνιος πίνακας.*

Απόδειξη. Έστω $A = (L, B, B)$. Γνωρίζουμε ότι $A^* = (L^*, B, B)$. Αν $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ τότε $A^* = \mathrm{diag}(\bar{\lambda}_1, \ldots, \bar{\lambda}_n)$. Είναι σαφές ότι $A^*A = AA^*$ δηλαδή ότι ο A και συνεπώς και η L είναι κανονικοί.

Αντιστρόφως, έστω ότι L είναι κανονική και έστω $B = \{e_1, \ldots, e_n\}$ μια ορθοκανονική βάση του V. Έχουμε ότι e_1 είναι ιδιοδιάνυσμα αφού ο A είναι άνω τριγωνικός και μάλιστα $Le_1 = a_{11}e_1$. Από την πρόταση VII.8.8 έχουμε ότι $L^*e_1 = \bar{a}_{11}e_1$. Σύμφωνα με τον ορισμό όμως της $A^* = \bar{A}^t$ αυτό σημαίνει ότι $a_{1j} = 0$ για $j > 1$. Συνεπώς $Le_2 = a_{22}e_2$, άρα και το e_2 είναι ιδιοδιάνυσμα της L. Όπως και πριν έχουμε ότι $L^*e_2 = \bar{a}_{22}e_2$, και αυτό μας δίνει όπως και πριν ότι $a_{2j} = 0$ για $j > 2$. Συνεχίζουμε με τον ίδιο τρόπο για να δείξουμε ότι ο πίνακας A είναι διαγώνιος. □

Πρόταση VII.8.11 (Διάσπαση Schur)**.** *Σε ένα πεπερασμένης διάστασης διανυσματικό χώρο* V *και* $L : V \to V$ *υπάρχει μια ορθοκανονική βάση* B, *ώστε ο πίνακας* (L, B, B) *να είναι άνω τριγωνικός.*

Απόδειξη. Αν $\dim V = 1$ τότε το αποτέλεσμα είναι σαφές. Θα αποδείξουμε το ζητούμενο με επαγωγή στην διάσταση του V. Υποθέτουμε ότι η πρόταση αληθεύει για κάθε γραμμική συνάρτηση από ένα διανυσματικό χώρο διάστασης $n-1$ στον εαυτό του. Θεωρούμε ένα διανυσματικό χώρο V διάστασης n. Θεωρούμε ένα ιδιοδιάνυσμα v της L^* το οποίο το κανονικοποιούμε στο $v/\|v\|$ ώστε να έχει μέτρο 1. Ο χώρος $\langle v/\|v\|\rangle$ είναι L^*-αναλλοίωτος συνεπώς ο $\langle v/\|v\|\rangle^\perp$ είναι L-αναλλοίωτος και έχει διάσταση $n - 1$, συνεπώς μπορούμε να εφαρμόσουμε την επαγωγική υπόθεση και να κατασκευάσουμε μια ορθοκανονική βάση $\{e_1, \ldots, e_{n-1}\}$ στην οποία ο L να είναι άνω τριγωνικός. Η ζητούμενη βάση είναι η $B = \{e_1, \ldots, e_{n-1}, v/\|v\|\}$. □

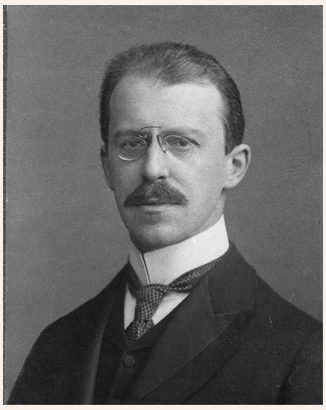
Πόρισμα VII.8.12. *Για κάθε πίνακα* $A \in \mathbb{C}^{n,n}$ *υπάρχει μοναδιαίος πίνακας* U, *ώστε* U^*AU *να είναι άνω τριγωνικός.*

Τέλος ο συνδυασμός των προτάσεων VII.8.10 και VII.8.11 δίνει το

Θεώρημα VII.8.13. *Έστω ένας πεπερασμένης διάστασης διανυσματικός χώρος και* $L : V \to V$ *μια κανονική γραμμική συνάρτηση. Υπάρχει μια ορθοκανονική βάση του* V *η οποία να αποτελείται από ιδιοδιανύσματα.*

Πόρισμα VII.8.14. *Για κάθε κανονικό πίνακα* A *υπάρχει μοναδιαίος πίνακας* U *ώστε* $U^{-1}AU$ *να είναι διαγώνιος πίνακας.*

Ο **Issai Schur** (10 Ιανουαρίου 1875 - 10 Ιανουαρίου 1941) ήταν Ρωσικής καταγωγής Μαθηματικός που έζησε και εργάστηκε στην Γερμανία. Ασχολήθηκε με την θεωρία ομάδων και ιδιαίτερα την θεωρία αναπαραστάσεων αλλά και με την συδυαστική, την Θεωρία Αριθμών και την Φυσική. Ήταν μαθητής του Frobenius και στην διδακτορική διατριβή του μελέτησε τις ρητές αναπαραστάσεις της γενικής γραμμικής ομάδας. Έχει αφήσει το όνομα του σε μια σειρά από έννοιες της άλγεβρας όπως η συνθήκη του Schur, ο πολλαπλασιαστής του Schur, οι συναρτήσεις του Schur, που σχετίζονται με τις αναπαραστάσεις της συμμετικής ομάδας και με την συνδυαστική.



Για να υπολογίσουμε ένα μοναδιαίο πίνακα U για ένα ερμητιανό πίνακα A, ώστε ο UAU^{-1} να είναι διαγώνιος παρατηρούμε ότι οι ιδιόχωροι που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι ανά δύο κάθετοι μεταξύ τους.

Μέθοδος 6 (Διαγωνοποίηση Ερμητιανών Πινάκων)**.**

1. *Υπολογισμός χαρακτηριστικού πολυωνύμου*
2. *Εύρεση ιδιοτιμών* $\lambda_1, \dots, \lambda_n$
3. *Υπολογισμός βάσεων* B_i *των ιδιοχώρων* E_{λ_i}.
4. *Ορθοκανονικοποίηση κατά Gram-Schmidt κάθε ιδιόχωρου.*

Παράδειγμα VII.8.15. Θεωρούμε τον ερμητιανό πίνακα

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

και θέλουμε να βρούμε μοναδιαίο πίνακα P, ώστε ο P^*AP να είναι διαγώνιος πίνακας.

Το χαρακτηριστικό πολυώνυμο του είναι το $\mathrm{Ch}_A[x] = (1+x)^2(-5+x)$, το οποίο μπορούμε να το υπολογίσουμε είτε με τον ορισμό είτε ως ειδική περίπτωση της άσκησης VI.1.30. Υπολογίζουμε ότι ο ιδιόχωρος της ιδιοτιμής $\lambda = 5$ είναι το διάνυσμα $v_1 = (1,1,1)^t$ ενώ ο ιδιόχωρος της ιδιοτιμής $\lambda = -1$ παράγεται από τα διανύσματα $v_2 = (1,-1,0)^t, v_3 = (1,0,-1)^t$. Θα μπορούσαμε να πάρουμε

$$Q = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \text{ με αντίστροφο } Q^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

και να είχαμε ότι $Q^{-1}AQ = \mathrm{diag}(5,-1,-1)$. Όμως ο πίνακας Q δεν είναι μοναδιαίος. Για να κατασκευάσουμε τον μοναδιαίο πίνακα P ορθοκανονικοποιούμε σε $e_1 = v_1/\|v_1\| = \frac{1}{\sqrt{3}}(1,1,1)^t$, $e_2 = v_2/\|v_2\| = 1/\sqrt{2}(1,-1,0)$, ενώ $e_3 = u_3/\|u_3\|$, όπου

$$u_3 = v_3 - \frac{\langle v_2, v_3\rangle}{\langle v_2, 2\rangle} v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \\ -1 \end{pmatrix} \text{ συνεπώς } e_3 = \sqrt{2/3}\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -1 \end{pmatrix}$$

Σχηματίζουμε τον πίνακα

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\sqrt{\frac{2}{3}} \end{pmatrix} \text{ με } P^{-1} = P^* = P^t = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\sqrt{\frac{2}{3}} \end{pmatrix}$$

και έχουμε ότι $P^*AP = \mathrm{diag}(5,-1,-1)$.

Άσκηση VII.8.16. Έστω $A \in \mathbb{C}^{n,n}$. Για ένα πολυώνυμο $f(x) = \sum_{\nu=0}^{n} a_\nu x^\nu$ θα συμβολίζουμε με $\overline{f}(x) = \sum_{\nu=0}^{n} \overline{a}_\nu x^\nu$. Αποδείξτε ότι

1. $\mathrm{ch}_{A^*}(x) = \overline{\mathrm{ch}_A}(x)$
2. $m_{A^*}(x) = \overline{m}_A(x)$
3. Αν λ είναι ιδιοτιμή του A τότε $\overline{\lambda}$ είναι ιδιοτιμή του A^*.

Λύση VII.8.17. 1. Παρατηρούμε ότι

$$\mathrm{ch}_{A^*}(x) = \det(A^* - x\mathbb{I}_n) = \overline{\det(A^t - \overline{x}\mathbb{I}_n)} = \overline{\det(A - \overline{x}\mathbb{I}_n)} = \overline{\mathrm{ch}_A(\overline{x})} = \overline{\mathrm{ch}}_A(x).$$

2. Έχουμε ότι αν $f(x) = \sum_{\nu=0}^{n} a_\nu x^\nu \in \mathbb{C}[x]$ πολυώνυμο ώστε $f(A) = 0$ τότε

$$0 = 0^* = \overline{\sum_{\nu=0}^{n} a_\nu (A^t)^\nu} = \sum_{\nu=0}^{n} \overline{a}_\nu (A^*)^\nu.$$

   Συνεπώς αν $m_A(x)$ είναι το ελάχιστου βαθμού πολυώνυμο που μηδενίζει τον A, το $\overline{m}_A(x)$ είναι το ελαχίστου βαθμού πολυώνυμο που μηδενίζει το A^*.

3. Σαφές από το 1.

Άσκηση VII.8.18. Έστω ένας μη-μηδενικός πίνακας γραμμή $x = (x_1, \ldots, x_n) \in \mathbb{F}^{n,1}$ με $n \geq 2$. Θεωρούμε την ευκλείδεια νόρμα $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$. Θέτουμε

$$A = \mathbb{I}_n + B, \text{ όπου } B = x^t x.$$

Αποδείξτε ότι

1. $B^2 = \|x\|^2 B$ και βρείτε το ελάχιστο πολυώνυμό του.
2. ο A είναι διαγωνοποιήσιμος και βρείτε τις διαστάσεις των ιδιοχώρων του
3. $\det A = 1 + \|x\|^2$.

Λύση VII.8.19. Παρατηρούμε ότι αν $B = (b_{ij})$ τότε $b_{ij} = x_i x_j$. Υπολογίζουμε ότι αν $B^2 = (c_{ij})$ τότε

$$c_{ij} = \sum_{\nu=1}^{n} b_{i\nu} b_{\nu j} = \sum_{\nu=1}^{n} x_i x_\nu^2 x_j = b_{ij}\|x\|^2.$$

Αφού $B^2 = \|x\|B$ έχουμε ότι για το πολυώνυμο

$$f(x) = x^2 - (2 + \|x\|^2)x + 1 + \|x\| = (x-1)(x-(1+\|x\|^2))$$

ισχύει ότι $f(A) = \mathbf{0}$. Επειδή δε κανένας από τους δύο γραμμικούς παράγοντες δεν μηδενίζουν τον A αυτό είναι το ελάχιστο.

Για τον ιδιόχωρο της ιδιοτιμής 1 έχουμε ότι αποτελείται από τα διανύσματα v ώστε

$$(A - \mathbb{I}_n)v = x^t x v = 0$$

και ότι κάθε διάνυσμα v που είναι κάθετο στο x ανήκει στον ιδιόχωρο αυτό. Άρα ο ιδιοχώρος της ιδιοτιμής 1 έχει διάσταση $\geq n-1 = \dim\langle x\rangle^\perp$. Αφού ο ιδιόχωρος της ιδιοτιμής $1+\|x\|^2$ έχει διάσταση ≥ 1 έχουμε ότι οι ιδιαστάσεις των ιδιόχωρων είναι $n-1$ και 1 αντίστοιχα. Τέλος η ορίζουσα είναι το γινόμενο των ιδιοτιμών άρα ίση με $1 + \|x\|^2$.

Άσκηση VII.8.20. Έστω $A \in \mathbb{C}^{n,n}$ και $a_1, \ldots, a_n$ οι γραμμές του A. Ισχύει ότι

$$\|a_1\| \cdots \|a_n\| \geq |\det A|$$

Η ισότητα ισχύει μόνο αν τα $a_1, \ldots, a_n$ είναι ορθογώνια ή κάποιο $a_i = \mathbf{0}$.

Λύση VII.8.21. Αν A δεν είναι αντιστρέψιμος η ανισότητα ισχύει με προφανή τρόπο και γίνεται ισότητα στην περίπτωση που κάποιο $a_i = \mathbf{0}$.

Αν ο A είναι αντιστρέψιμος τότε από την πρόταση VII.7.12 υπάρχει μοναδικός κάτω τριγωνικός πίνακας B με θετικά στοιχεία στην διαγώνιο ώστε ο πίνακας $BA = U$ να είναι μοναδιαίος. Έχουμε ότι

$$|\det U|^2 = \det U \det U^* = \det UU^* = \det \mathbb{I}_n = 1.$$

Γράφουμε

$$|\det(A)| = |\det\left(B^{-1}U\right)| = |\det\left(B^{-1}\right)| = b_{11}^{-1} \cdots b_{nn}^{-1}$$

από την άλλη έχουμε ότι πίνακας $B^{-1} = (c_{ij})$ είναι κάτω τριγωνικός με $c_{ii} = b_{ii}^{-1}$.

$$\begin{aligned} a_i &= (c_{i1}, c_{i2}, \ldots, b_{ii}^{-1}, 0, \ldots, 0)U \\ &= c_{i1}U^1 + c_{i2}U^2 + \cdots + b_{ii}^{-1}U^i \text{ όπου } U^1, \ldots, U^n \text{ οι στήλες του } U. \end{aligned}$$

Άρα

$$\|a_i\|^2 = |c_{i1}|^2\|U^1\|^2 + \cdots + |b_{ii}^{-1}|^2\|U^i\|^2 = |c_{i1}|^2 + \cdots + |b_{ii}^{-1}|^2 \geq |b_{ii}^{-1}|^2. \tag{VII.13}$$

Από την τελευταία σχέση προκύπτει ότι

$$|a_1| \cdots |a_n| \geq b_{11}^{-1} \cdots b_{nn}^{-1} = |\det(A)|.$$

Τέλος, παρατηρούμε ότι η ανισότητα (VII.13) γίνεται ισότητα αν και μόνο αν $c_{i1} = \cdots = c_{i,i-1} = 0$, ισοδύναμα αν και μόνο αν $a_i = b_{ii}^{-1}U^i$, αν και μόνο αν οι γραμμές a_i είναι ανά δύο ορθογώνιες.

VIII
Γραμμικές μορφές

VIII.1 Τετραγωνικές μορφές

VIII.1.1 Τετραγωνικές εξισώσεις στο επίπεδο

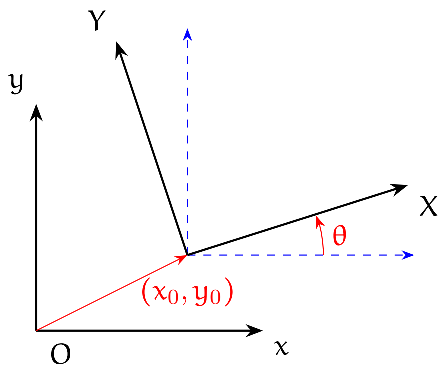
Η γενική δευτεροβάθμια εξίσωση στο επίπεδο έχει την μορφή

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0. \tag{VIII.1}$$

Σε μαθήματα αναλυτικής γεωμετρίας η παραπάνω εξίσωση μελετάται με την βοήθεια αλλαγής συστήματος συντεταγμένων. Δηλαδή θεωρούμε την αλλαγή μεταβλητών

$$\begin{aligned} x &= x_0 + X\cos\theta - Y\sin\theta \\ y &= y_0 + X\sin\theta + Y\cos\theta \end{aligned} \tag{VIII.2}$$

η οποία οδηγεί σε ένα σύστημα συντεταγμένων μετατοπισμένο κατά (x_0, y_0) και περιστραμένο κατά γωνία θ όπως στο σχήμα Στο νέο σύστημα συντεταγμένων η τετραγωνική εξίσωση έχει



Σχήμα VIII.1: Μετακίνηση και περιστροφή συστήματος συντεταγμένων

μετατραπεί στην

$$A'X^2 + 2B'XY + C'Y^2 + 2D'X + 2E'Y + F = 0$$

όπου

$$\begin{aligned}
A' &= A\cos^{\theta} + 2B\sin\theta\cos\theta + C\sin^2\theta \\
B' &= -A\sin\theta\cos\theta + B(\cos^2\theta - \sin^2\theta) + C\sin\theta\cos\theta \\
C' &= A\sin^2\theta - 2B\sin\theta\cos\theta + C\cos^2\theta \\
D' &= (Ax_0 + By_0 + D)\cos\theta + (Bx_0 + Cy_0E)\sin\theta \\
E' &= -(Ax_0 + By_0 + D)\sin\theta + (Bx_0 + Cy_0 + E)\cos\theta \\
F' &= (Ax_0 + By_0 + D)x_0 + B(x_0 + Cy_0 + E)y_0 + Dx_0 + Ey_0 + F
\end{aligned}$$

Στην συνέχεια επιλέγουμε το θ ώστε $B' = 0$. Αν $A = C$ τότε αυτό γίνεται αν $\cos 2\theta = 0$ δηλαδή για $\theta = \pi/4$ εκτός αν $B = 0$ οπότε δεν χρειάζεται στροφή. Αν $A \neq B$, τότε θα πρέπει

$$\frac{2B}{A - C} = \frac{2\sin\theta\cos\theta}{\cos^2\theta - \sin^2\theta} = \tan 2\theta$$

από όπου προσδιορίζεται η γωνία θ ώστε $B' = 0$. Στην συνέχεια επιλέγουμε τα (x_0, y_0) ώστε $D', F' = 0$ το οποίο γίνεται λύνοντας το σύστημα

$$\begin{pmatrix} A & B \\ B & C \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} D \\ E \end{pmatrix}$$

αν αυτό έχει λύση. Τελικός σκοπός μας είναι να φέρουμε την εξίσωση σε μια μορφή την οποία μπορούμε να αναγνωρίσουμε όπως

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1 \text{ Έλλειψη} \tag{VIII.3}$$

$$\begin{aligned}
\frac{X^2}{a^2} + \frac{Y^2}{b^2} &= -1 \text{ Κενό σύνολο} \\
\frac{X^2}{a^2} - \frac{Y^2}{b^2} &= 1 \text{ Υπερβολή} \\
-\frac{X^2}{a^2} + \frac{Y^2}{b^2} &= 1 \text{ Υπερβολή} \\
AX^2 + BY^2 &= 0 \text{ Ζευγάρι πραγματικών ευθείων ή σημείο (Μιγαδικές ευθείες)} \\
A'X^2 + 2D'X + 2E'Y + F' &= 0 \text{ Παραβολή ή ζεύγος ευθειών πραγματικών ή μιγαδικών} \\
C'Y^2 + 2D'X + 2E'Y + F' &= 0 \text{ Παραβολή ή ζεύγος ευθειών πραγματικών ή μιγαδικών}
\end{aligned} \tag{VIII.4}$$

Η μέθοδος αυτή έχει το πλεονέκτημα της γεωμετρικής εποπτίας αλλά απαιτεί αρκετές πράξεις και γενικεύται δύσκολα σε μεγαλύτερες διαστάσεις. Επιπλέον η διερεύνηση είναι αρκετά πολύπλοκη.

Πρωτού προχωρήσουμε σε μία μέθοδο βασισμένη στις τεχνικές του μαθήματος αυτού ας παρατηρήσουμε ότι η αλλαγή συστήματος συντεταγμένων στην εξίσωση (VIII.2) μπορεί να γραφεί σε μορφή πινάκων ως

$$\begin{pmatrix} x \\ y \end{pmatrix} = P\begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \text{ με } P = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{VIII.5}$$

όπου ο πίνακας P είναι μοναδιαίος. Επίσης παρατηρούμε ότι η εξίσωση (VIII.1) γράφεται στην μορφή

$$(x, y)\begin{pmatrix} A & B \\ B & C \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + 2\begin{pmatrix} E & D \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + F = 0 \tag{VIII.6}$$

Δηλαδή, ο υπολογισμός της γωνίας θ αντιστοιχεί στον υπολογισμό ενός μοναδιαίου πίνακα P ώστε

$$P^t \begin{pmatrix} A & B \\ B & C \end{pmatrix} P = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

με $\lambda_1, \lambda_2 \in \mathbb{R}$ ώστε αντικαθιστώντας την (VIII.5) στην (VIII.6) έχουμε ότι

$$\begin{aligned}
(x, y) \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} X & Y \end{pmatrix} P^t \begin{pmatrix} A & B \\ B & C \end{pmatrix} P \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} x_0 & y_0 \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\
&= \begin{pmatrix} X & Y \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} x_0 & y_0 \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\
&= \lambda_1 X^2 + \lambda_2 Y^2 + \begin{pmatrix} x_0 & y_0 \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}
\end{aligned}$$

Στην παραπάνω εξίσωση ο πίνακας P είναι αντιστρέψιμος οπότε θέτουμε

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = P \begin{pmatrix} x_0' \\ y_0' \end{pmatrix}$$

και έχουμε ότι

$$\begin{aligned}
(x, y) \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \lambda_1 X^2 + \lambda_2 Y^2 + \begin{pmatrix} x_0 & y_0 \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\
&= \lambda_1 X^2 + \lambda_2 Y^2 + \lambda_1 (x_0')^2 + \lambda_2 (y_0')^2 \\
&= \lambda_1 (X - x_0')^2 + \lambda_2 (Y - y_0')^2 + 2 \begin{pmatrix} x_0' & y_0' \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}
\end{aligned}$$

Από την άλλη έχουμε

$$2 \begin{pmatrix} E & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} E & D \end{pmatrix} P \begin{pmatrix} X \\ Y \end{pmatrix} + 2 \begin{pmatrix} E & D \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

Οπότε αν μπορέσουμε να διαλέξουμε το (x_0, y_0) με τέτοιο τρόπο ώστε

$$\begin{pmatrix} x_0' & y_0' \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} + \begin{pmatrix} E & D \end{pmatrix} P = 0$$

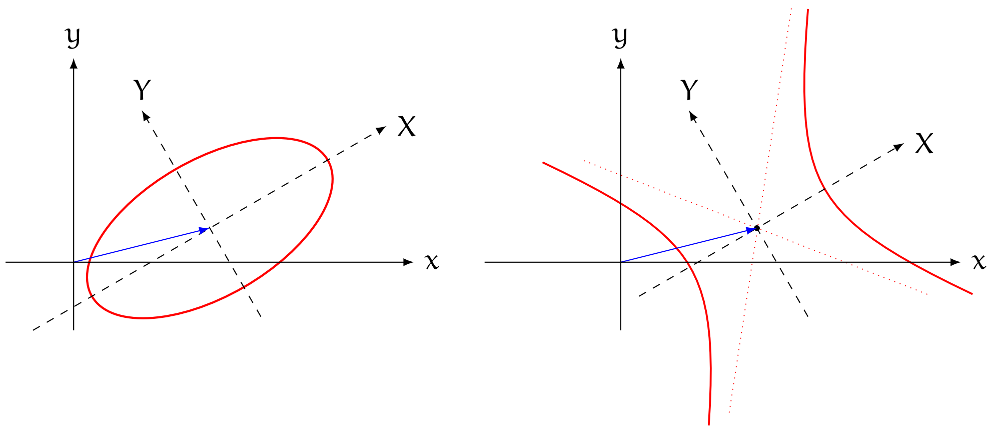
ισοδοδύναμα αν το σύστημα

$$A^t \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = - \begin{pmatrix} E \\ D \end{pmatrix} \tag{VIII.7}$$

έχει λύση $(x_0, y_0)^t$, τότε η αρχική εξίσωση γράφεται στην μορφή

$$\lambda_1 (X - x_0')^2 + \lambda_2 (Y - y_0')^2 + 2 \begin{pmatrix} E & D \end{pmatrix} P \begin{pmatrix} x_0' \\ y_0' \end{pmatrix} + F = 0.$$

Αυτή την περίπτωση την λέμε καμπύλη με κέντρο συμμετρίας το (x_0, y_0). Το παραπάνω σύστημα έχει μοναδική λύση αν και μόνο αν η ποσότητα $AC - B^2 = \lambda_1 \lambda_2 \neq 0$. Οι ιδιοτιμές είναι πραγματικές, ως ιδιοτιμές συμμετρικού πίνακα. Καταλλήγουμε με αυτό τον τρόπο στις πέντε πρώτες εξισώσεις της ταξινόμισης στην εξίσωση (VIII.3). Στο παρακάτω σχήμα σχεδιάζουμε μία περιστραμένη έλλειψη και μία περιστραμένη υπερβολή.



Αν υποθέσουμε τώρα ότι το σύστημα (VIII.7) δεν έχει λύση. Αυτό σημαίνει ότι μία τουλάχιστον ιδιοτιμή είναι μηδενική. Υποθέτουμε ότι $\lambda_1 \neq 0$ και $\lambda_2 = 0$. Η εξίσωση της καμπύλης γίνεται σε αυτή την περίπτωση

$$\lambda_1(X - x_0')^2 + 2\begin{pmatrix} x_0' & y_0' \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix} + 2\begin{pmatrix} E & D \end{pmatrix} P \begin{pmatrix} X \\ Y \end{pmatrix} + 2\begin{pmatrix} E & D \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + F = 0.$$

Σε αυτή την περίπτωση διαλέγουμε x_0' ώστε

$$x_0'\lambda_1 + Ep_{11} + Dp_{22} = 0$$

οπότε η εξίσωση γίνεται

$$\lambda_1(X - x_0')^2 + 2(Ep_{12} + Dp_{22})Y + 2\begin{pmatrix} E & D \end{pmatrix} P \begin{pmatrix} x_0' \\ y_0' \end{pmatrix} + F = 0$$

Αν $Ep_{12} + Dp_{22} \neq 0$ η παραπάνω εξίσωση παριστάνει μία παραβολή ενώ αν $Ep_{12} + Dp_{22} = 0$ η εξίσωση είναι δύο ευθείες πραγματικές ή μιγαδικές.

Αν και οι δύο ιδιοτιμές $\lambda_1 = \lambda_2 = 0$ είναι μηδενικές τότε η καμπύλη δεν είναι τετραγωνική αλλά εκφυλίζεται στην τόπο μηδενισμού της

$$2Dx + 2Ey + F = 0$$

η οποία είναι ευθεία ή αδύνατη (αν $D = E = 0$ και $F \neq 0$) ή όλο το $\mathbb{R}^2$ αν $D = E = F = 0$.

VIII.1.2 Τετραγωνικές εξισώσεις στον χώρο

Η γενική μορφή μιας τετραγωνικής εξίσωσης στον χώρο είναι η

$$(x_1, \dots, x_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + 2\begin{pmatrix} E_1 & E_2 & \cdots & E_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + F = 0.$$

Θέτουμε $x = (x_1, \dots, x_n)^t$, $E = (E_1, \dots, E_n)$ και $A = (a_{ij})$ είναι ένας συμμετρικός πίνακας. Συνεπώς υπάρχει μοναδιαίος πίνακας $P \in \mathbb{R}^{n,n}$ ώστε

$$P^t AP = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

Θέτουμε $X = (X_1, \dots, X_n)$ από την εξίσωση $X = P^t x$ δηλαδή $x = PX$. Έχουμε την εξίσωση

$$X^t P^t APX + 2EPX + F = X^t \mathrm{diag}(\lambda_1, \dots, \lambda_n)X + 2EPX + F = 0.$$

Η παραπάνω εξίσωση δίνει

$$\sum_{i=1}^{n} \lambda_i X_i^2 + 2\sum_{i=1}^{n}\sum_{\nu=1}^{n} E_\nu P_{\nu,i} X_i + F = 0.$$

Υποθέτουμε ότι οι ιδιοτιμές $\lambda_1, \ldots, \lambda_\kappa$ είναι διαφορετικές του μηδενός ενώ $\lambda_{\kappa+1} = \cdots = \lambda_n = 0$. Για κάθε ιδιοτιμή λ_i, $1 \le i \le \kappa$ γράφουμε $Y_i = X_i + a_i$ και έχουμε ότι

$$\begin{aligned} \lambda_i (X_i + a_i - a_i)^2 + 2\sum_{\nu=1}^{n} E_\nu P_{\nu,i}(X_i + a_i - a_i) &= \lambda_i Y_i^2 - 2a_i\lambda_i Y_i + \lambda_i a_i^2 + 2\sum_{\nu=1}^{n} E_\nu P_{\nu,i} Y_i - 2\sum_{\nu=1}^{n} E_\nu P_{\nu,i} a_i \\ &= \lambda_i Y_i^2 - \frac{1}{\lambda_i}\left(\sum_{\nu=1}^{n} E_\nu P_{\nu,i}\right)^2 \end{aligned} \tag{VIII.8}$$

επιλέγοντας

$$a_i = \frac{1}{\lambda_i}\sum_{\nu=1}^{n} E_\nu P_{\nu,i},$$

ώστε να μηδενιστεί ο συντελεστής του Y_i.

Με αυτό τον τρόπο καταλλήγουμε στην εξίσωση

$$\sum_{i=1}^{\kappa} \lambda_i Y_i^2 + \sum_{i=\kappa+1}^{n} Q_i X_i - \sum_{i=1}^{\kappa} \frac{1}{\lambda_i} Q_i^2 + F = 0 \tag{VIII.9}$$

όπου

$$Q_i = \sum_{\nu=1}^{n} E_\nu P_{\nu,i}.$$

Αν $\kappa = n$ ή αν $Q_{\kappa+1} = \cdots = Q_n = 0$ τότε η εξίσωση γράφεται ως

$$\sum_{i=1}^{\kappa} \lambda_i Y_i^2 + F = 0$$

η οποία ονομάζεται κανονική εξίσωση τετραγωνικής υπερεπιφάνειας με κέντρο. Αν $\kappa = n$ τότε ονομάζεται μη εκφυλισμένη διαφορετικά ονομάζεται εκφυλισμένη.

Η εξίσωση (VIII.9) μπορεί να απλοποιηθεί περαιτέρω αν θέσουμε ως νέα μεταβλητή την

$$Y = \sum_{i=\kappa+1}^{n} Q_i X_i.$$

Προκειμένου να έχουμε διατήρηση του εσωτερικού γινομένου θέλουμε οι αλλαγές μεταβλητών να γίνονται μέσω μοναδιαίων πινάκων. Οπότε θεωρούμε τον πίνακα

$$P_1 = \left(\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \vdots & & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \hline 0 & \cdots & \cdots & 0 & -\frac{1}{q}Q_{\kappa+1} & \cdots & -\frac{1}{q}Q_n \\ q_{k+2,1} & q_{k+2,2} & \cdots & \cdots & \cdots & \cdots & q_{k+1,n} \\ \vdots & \vdots & & & & & \vdots \\ q_{k+2,1} & q_{k+2,2} & \cdots & \cdots & \cdots & \cdots & q_{k+1,n} \end{array}\right)$$

στον οποίο το πάνω δεξιά $\kappa \times \kappa$ κομμάτι είναι ο ταυτοτικός πίνακας, η $k+1$ γραμμή έχει διαιρεθεί με το μέτρο της

$$q = \sqrt{\sum_{\nu=\kappa+1}^{n} Q_\nu^2}$$

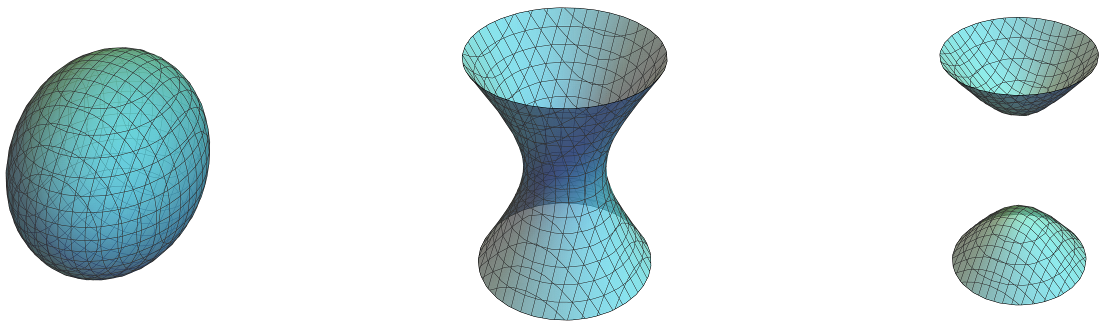
ώστε οι πρώτες $\kappa + 1$ γραμμές να είναι ορθοκανονικές ενώ έχουμε συμπληρώσει τις υπόλοιπες γραμμές σε μία ορθοκανονική βάση του $\mathbb{R}^n$ και ο πίνακας P_1 να είναι μοναδιαίος. Εκτελούμε μια αλλαγή βάσης $Y = P_1^t Z$ οπότε καταλήγουμε στην εξίσωση

$$\sum_{i=1}^{\kappa} \lambda_i Z_i^2 = 2qZ_{\kappa+1} - D.$$

Αν $D \neq 0$ τότε με μια παράλληλη μετατόπιση $T = Z_{\kappa+1} - \frac{F}{2q}$ έχουμε την εξίσωση

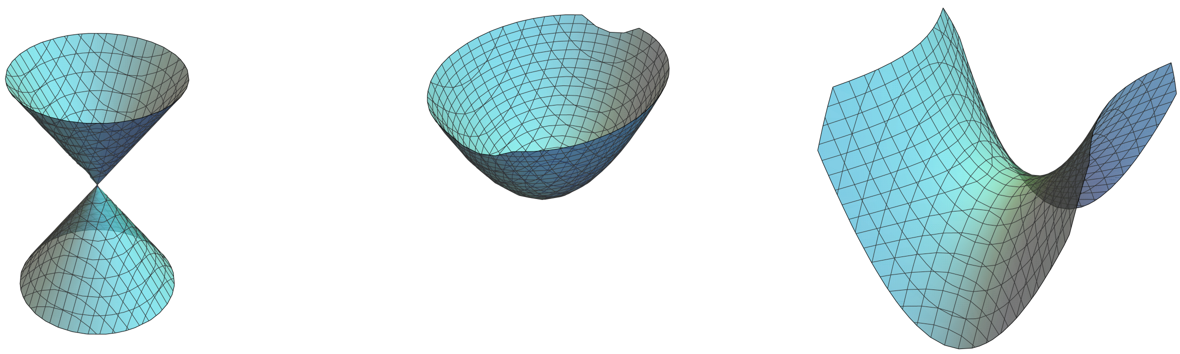
$$\sum_{i=1}^{\kappa} \lambda_i Z_i^2 = 2qT,$$

η οποία ονομάζεται η κανονική εξίσωση μιας υπερεπιφάνειας χωρίς κέντρο.



(α΄) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^3} = 1$ (β΄) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ (γ΄) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$

Σχήμα VIII.2: $\kappa = 3$ μη-ιδιόμορφες



(α΄) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^3} = 0$ (β΄) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = z$ (γ΄) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = z$

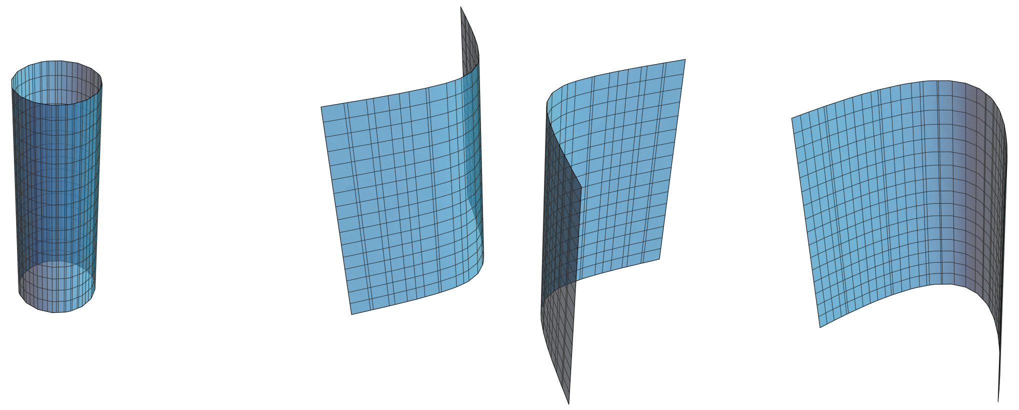
Σχήμα VIII.3: $\kappa = 3$ μη-ιδιόμορφες

Παράδειγμα VIII.1.1. Για να μελετήσουμε την επιφάνεια

$$x^2 - 2xy + 2y^2 = 1,$$

την γράφουμε στην μορφή

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.$$



(α΄) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ (β΄) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ (γ΄) $\frac{x^2}{a^2} = 2y$

Σχήμα VIII.4: $\kappa = 1$ κύλινδροι

Υπολογίζουμε τις ιδιοτιμές $\lambda_1 = \frac{3-\sqrt{5}}{2}$, $\lambda_2 = \frac{3+\sqrt{5}}{2}$ και βρίσκουμε τους αντίστοιχους ιδιόχωρους $E_{\lambda_1} = \langle (\frac{1+\sqrt{5}}{2}, 1)^t \rangle$, $E_{\lambda_2} = \langle (\frac{1-\sqrt{5}}{2}, 1)^t \rangle$. Τα ιδιοδιανύσματα είναι κάθετα μεταξύ τους γιατί αντιστοιχούν σε ιδιοτιμές συμμετρικού πίνακα. Για να καταλήξουμε σε μία ορθοκανονική βάση τα διαιρούμε με το μέτρο τους και έχουμε τα ορθοκανονικά διανύσματα

$$e_1 = \frac{(\frac{1+\sqrt{5}}{2}, 1)^t}{\|(\frac{1+\sqrt{5}}{2}, 1)^t\|} = \begin{pmatrix} \frac{1+\sqrt{5}}{\sqrt{2(5+\sqrt{5})}} & \sqrt{\frac{2}{5+\sqrt{5}}} \end{pmatrix}^t \qquad e_2 = \frac{(\frac{1-\sqrt{5}}{2}, 1)^t}{\|(\frac{1-\sqrt{5}}{2}, 1)^t\|} = \begin{pmatrix} -\frac{\sqrt{5}-1}{\sqrt{2(5-\sqrt{5})}} & \sqrt{\frac{2}{5-\sqrt{5}}} \end{pmatrix}^t$$

τα οποία τα τοποθετούμε ως στήλες του μοναδιαίου πίνακα P

$$P = \begin{pmatrix} \frac{1+\sqrt{5}}{\sqrt{2(5+\sqrt{5})}} & -\frac{\sqrt{5}-1}{\sqrt{2(5-\sqrt{5})}} \\ \sqrt{\frac{2}{5+\sqrt{5}}} & \sqrt{\frac{2}{5-\sqrt{5}}} \end{pmatrix}.$$

Θέτοντας ως , από τον τύπο

$$\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} X, Y \end{pmatrix}$$

έχουμε την εξίσωση στην μορφή

$$\lambda_1 X^2 + \lambda_2 Y^2 = 1$$

η οποία παριστάνει μια έλλειψη αφού και οι δύο ιδιοτιμές λ_1, λ_2 είναι θετικές, όπως φαίνεται στο σχήμα VIII.5. Στην συνέχεια θέλουμε να μελετήσουμε την μορφή της καμπύλης

$$x^2 - 2xy + 2y^2 + x + y + F = 0$$

για διάφορες τιμές της παραμέτρου $F \in \mathbb{R}$. Υπολογίζουμε ότι

$$\begin{pmatrix} Q_1 & Q_2 \end{pmatrix} = \begin{pmatrix} E_1 & E_2 \end{pmatrix} P = \begin{pmatrix} \frac{3+\sqrt{5}}{\sqrt{2(5+\sqrt{5})}} & \frac{3-\sqrt{5}}{\sqrt{2(5-\sqrt{5})}} \end{pmatrix}.$$

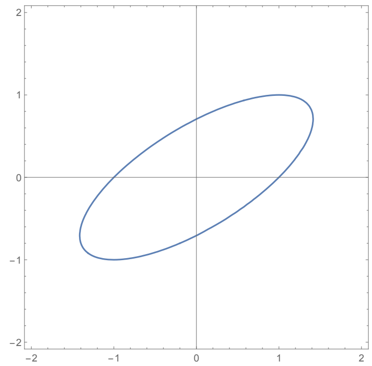
Υπολογίζουμε ότι

$$a_1 = Q_1/\lambda_1 = -\frac{(3+\sqrt{5})\sqrt{\frac{2}{5+\sqrt{5}}}}{\sqrt{5}-3} \qquad a_2 = Q_2/\lambda_2 = -\frac{(\sqrt{5}-3)\sqrt{\frac{1}{10}(5+\sqrt{5})}}{3+\sqrt{5}},$$

οπότε για τις μεταβλητές $Y_i = X_i + a_i$, $i = 1, 2$, σύμφωνα με την εξίσωση (VIII.8) έχουμε την εξίσωση

$$\lambda_1 Y_1^2 + \lambda_2 Y_2^2 = -F + \frac{Q_1^2}{\lambda_1} + \frac{Q_2^2}{\lambda_2} = 5 - F.$$

Οπότε η εξίσωση παριστάνει όπως και πριν έλλειψη αρκεί $5 - F > 0$ ή σημείο αν $F = 5$ ή το κενό σύνολο αν $5 - F < 0$.



Σχήμα VIII.5: Η έλλειψη $x^2 - 2xy + 2y^2 = 0$

VIII.2 Μιάμιση και διγραμμικές μορφές

Ορισμός VIII.2.1. Έστω V ένας διανυσματικός χώρος πάνω από το $\mathbb{F}$, όπου $\mathbb{F} = \mathbb{R}$ ή $\mathbb{F} = \mathbb{C}$. Μία συνάρτηση

$$f : V \times V \longrightarrow \mathbb{F}$$

θα λέγεται μιάμιση (sesqui) μορφή αν ικανοποιεί

1. $f(\lambda_1 v_1 + \lambda_2 v_2, v_3) = \lambda_1 f(v_1, v_3) + \lambda_2 f(v_2, v_3)$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{F}$ και κάθε $v_1, v_2, v_3 \in V$.
2. $f(v_1, \lambda_1 v_2 + \lambda_2 v_3) = \overline{\lambda}_1 f(v_1, v_2) + \overline{\lambda}_2 f(v_1, v_3)$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{F}$ και κάθε $v_1, v_2, v_3 \in V$.

Θα λέγεται διγραμμική αν ικανοποιεί τις σχέσεις

1. $f(\lambda_1 v_1 + \lambda_2 v_2, v_3) = \lambda_1 f(v_1, v_3) + \lambda_2 f(v_2, v_3)$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{F}$ και κάθε $v_1, v_2, v_3 \in V$.
2. $f(v_1, \lambda_1 v_2 + \lambda_2 v_3) = \lambda_1 f(v_1, v_2) + \lambda_2 f(v_1, v_3)$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{F}$ και κάθε $v_1, v_2, v_3 \in V$.

Παρατήρηση VIII.2.2. Αν το $\mathbb{F} = \mathbb{R}$ τότε οι έννοιες της διγραμμικής και μιάμιση μορφής ταυτίζονται.

Παράδειγμα VIII.2.3. Σε ένα διανυσματικό χώρο V υπέρ του σώματος $\mathbb{C}$ ένα ερμητιανό εσωτερικό γινόμενο είναι μια μιαμιση μορφή. Επίσης αν $L : V \to V$ είναι μια γραμμική συνάρτηση τότε η

$$f(v, w) = \langle Lv, w \rangle$$

είναι μια μιάμιση μορφή.

Σχόλιο: Μία διγραμμική συνάρτηση είναι διπλά γραμμική. Αντίθετα μια μιάμιση μορφή είναι γραμμική στο πρώτο όρισμα αλλά κάτι λιγότερο από γραμμική στο δεύτερο αφού ισχύει μόνο η προσθετική ιδιότητα της γραμμικής συνάρτησης. Η λατινική λέξη “sesqui” σημαίνει ένα και μισό.

VIII.2.1 Μιάμιση μορφές

Θεώρημα VIII.2.4. *Για ένα διανυσματικό χώρο V πεπερασμένης διάστασης με εσωτερικό γινόμενο και μια μιάμιση μορφή f υπάρχει μοναδική γραμμική συνάρτηση $L_f : V \to V$ ώστε*

$$f(v, w) = \langle L_f v, w \rangle \text{ για κάθε } v, w \in W.$$

Απόδειξη. Για ένα σταθερό διάνυσμα w έχουμε ότι η συνάρτηση

$$\begin{aligned} f(v, w) : V &\longrightarrow \mathbb{F} \\ v &\longmapsto f(v, w) \end{aligned}$$

είναι μια γραμμική συνάρτηση και συνεπώς από την πρόταση VII.6.3 έχουμε ότι υπάρχει ένα $u(w) \in V$ ώστε

$$f(v, w) = \langle v, u(w) \rangle.$$

Στην παραπάνω εξίσωση συμβολίσαμε το στοιχείο με $u(w)$ για να δείξουμε την εξάρτησή του από το w. Θα δείξουμε τώρα ότι

$$u(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 u(w_1) + \lambda_2 u(w_2) \tag{VIII.10}$$

δηλαδή το στοιχείο $u(w)$ είναι γραμμική συνάρτηση του w. Πράγματι έχουμε για κάθε $v, w_1, w_2 \in V$ και κάθε $\lambda_1, \lambda_2 \in \mathbb{F}$ ότι

$$\begin{aligned} \langle v, u(\lambda_1 w_1 + \lambda_2 w_2) \rangle &= f(v, \lambda_1 w_1 + \lambda_2 w_2) \\ &= \overline{\lambda_1} f(v, w_1) + \overline{\lambda_2} f(v, w_2) \\ &= \langle v, \lambda_1 u(w_1) \rangle + \langle v, \lambda_2 u(w_2) \rangle \\ &= \langle v, \lambda_1 u(w_1) + \lambda_2 u(w_2) \rangle \end{aligned}$$

από όπου προκύπτει ότι ισχύει η (VIII.10). Έτσι αρκεί να πάρουμε $L_f = u^*$. Τέλος παρατηρούμε ότι αν υπήρχαν δύο συναρτήσεις L_1, L_2 οι οποίες να είχαν την ζητούμενη ιδιότητα τότε θα είχαμε ότι για κάθε $v, w \in V$

$$\langle L_1(v), w \rangle = \langle L_2(v), w \rangle \Rightarrow \langle L_1(v) - L_2(v), w \rangle = 0$$

συνεπώς $L_1 = L_2$. □

Έστω F ο χώρος των μιαμιση μορφών είναι ένας διανυσματικός χώρος πάνω από το $\mathbb{F}$. Πράγματι αν προσθέσουμε δύο μιάμιση μορφές παίρνουμε μια μιάμιση μορφή και αν πολλαπλασιάσουμε μια μιάμιση μορφή με ένα στοιχείο $\lambda \in \mathbb{F}$ παίρνουμε μια μιάμιση μορφή.

Πρόταση VIII.2.5. *Υπάρχει ένας ισομορφισμός διανυσματικών χώρων*

$$\begin{aligned} \psi : F &\longrightarrow \mathrm{End}(V) \\ f &\longmapsto L_f \end{aligned}$$

Απόδειξη. Παρατηρούμε ότι για οποιαδήποτε στοιχεία $f_1, f_2 \in F$ και οποιαδήποτε $\lambda_1, \lambda_2 \in \mathbb{F}$ έχουμε ότι

$$\begin{aligned} \langle L_{\lambda_1 f_1 + \lambda_2 f_2}(v), w \rangle &= (\lambda_1 f_1 + \lambda_2 f_2)(v, w) = \lambda_1 f_1(v, w) + \lambda_2 f_2(v, w) \\ &= \lambda_1 \langle L_{f_1}(v), w \rangle + \lambda_2 \langle L_{f_2}(v), w \rangle = \langle (\lambda_1 L_{f_1} + \lambda_2 L_{f_2})(v), w \rangle \end{aligned}$$

από όπου προκύπτει ότι

$$L_{\lambda_1 f_1 + \lambda_2 f_2} = \lambda_1 L_{f_1} + \lambda_2 L_{f_2}$$

δηλαδή η συνάρτηση ψ είναι γραμμική. Επίσης κάθε $L : V \to V$ ορίζει μια $f(v, w) = \langle Lv, w \rangle$ άρα η ψ είναι επί. Τέλος $\ker\psi$ είναι οι $f \in F$ ώστε $L_f = 0$ αλλά τότε $f(v, w) = \langle L_f v, w \rangle = \langle 0, w \rangle = 0$, δηλαδή η ψ είναι και 1-1. □

Ορισμός VIII.2.6. Για μια μιάμιση μορφή και για μία διατεταγμένη βάση $B = \{v_1, \ldots, v_n\}$ του V, ορίζουμε τον πίνακα $A = (a_{ij})$ με

$$a_{ij} = f(v_j, f_i),$$

τον οποίο ονομάζουμε πίνακα της μιάμιση μορφής f.

Ο πίνακας A της μιάμιση μορφής γνωρίζει ότι την πληροφορία της f, όπως δείχνει η παρακάτω πρόταση.

Πρόταση VIII.2.7. *Για v, w συμβολίζουμε με $[v]_B, [w]_B \in \mathbb{R}^n$ τα διανύσματα στήλες των συντεταγμένων των v, w. Ισχύει*

$$f(v, w) = [w]_B^* A [v]_B.$$

Απόδειξη. Γράφουμε $[v]_B = (\lambda_1, \ldots, \lambda_n)^t$ και $[w]_W = (\mu_1, \ldots, \mu_n)^t$. Έχουμε ότι

$$\begin{aligned} f(v, w) &= f(\sum_{i=1}^n \lambda_i v_i, \sum_{j=1}^n \mu_j v_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n \overline{\mu}_j \lambda_i f(v_i, v_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{j,i} \overline{\mu}_j \lambda_i \end{aligned}$$

από όπου προκύπτει το ζητούμενο. □

Παρατήρηση VIII.2.8. Αν επιλέξουμε την βάση $v_1, \ldots, v_n$ να είναι ορθοκανονική τότε έχουμε ότι

$$f(v_j, v_i) = \langle T_f e_j, e_i \rangle$$

ο οποίος είναι ο πίνακας της γραμμικής συνάρτησης T_f ως προς την βάση B, δείτε την πρόταση VII.6.9.

Παρατήρηση VIII.2.9. Αν κάνουμε την αλλαγή βάσης $v_j' = \sum_{i=1}^n p_{ij} v_i$ από την βάση B στην βάση $B' = \{v_1', \ldots, v_n'\}$ τότε έχουμε ότι ο πίνακας A' της μιάμιση μορφής ως προς την νέα βάση δίνεται από την εξίσωση

$$A' = P^* A P.$$

Πράγματι,

$$\begin{aligned} a_{ij}' = f(v_j', v_i') &= f(\sum_{\nu=1}^n p_{\nu,j} v_\nu, \sum_{\mu=1}^n p_{\mu,i} v_i) \\ &= \sum_{\nu=1}^n \sum_{\mu=1}^n p_{\nu,j} \overline{p}_{\mu,i} f(v_\nu, v_\mu) \\ & \sum_{\nu=1}^n \sum_{\mu=1}^n \overline{p}_{\mu,i} a_{\mu,\nu} p_{\nu,j} \end{aligned}$$

Ορισμός VIII.2.10. Μια μιάμιση μορφή f θα λέγεται ερμητιανή αν και μόνο αν

$$f(v,w) = \overline{f(w,v)} \text{ για κάθε } v,w \in V.$$

Παρατήρηση VIII.2.11. Αν $L_f : V \to V$ είναι η γραμμική συνάρτηση που αντιστοιχεί στην μιάμιση μορφή f, έχουμε ότι

$$\begin{aligned} f(v,w) &= \langle L_f(v), w\rangle \\ \overline{f(w,v)} &= \langle L_f(w), v\rangle = \overline{\langle w, L_f^*(v)\rangle} \end{aligned}$$

από όπου προκύπτει ότι η f είναι ερμητιανή αν και μόνο αν η L_f είναι ερμητιανή, δηλαδή $L_f^* = L_f$.

Θεώρημα VIII.2.12. *Για κάθε ερμητιανή μιάμιση μορφή f σε ένα πεπερασμένης διάστασης διανυσματικό χώρο V με εσωτερικό γινόμενο υπάρχει μια ορθοκανονική βάση στην οποία ο f αναπαρίσταται από ένα διαγώνιο πίνακα με πραγματικά στοιχεία στην διαγώνιο.*

Απόδειξη. Αφού η f είναι ερμητιανή, σύμφωνα με την παρατήρηση VIII.2.11 η γραμμική της συνάρτηση L_f είναι ερμητιανή, δηλαδή $L_f = L_f^*$. Σύμφωνα με το θεώρημα VII.8.6 υπάρχει ορθοκανονική βάση $B = \{e_1, \ldots, e_n\}$ του V η οποία αποτελείται από ιδιοδιανύσματα της L_f, δηλαδή $L_f(e_i) = \lambda_i e_i$. Έχουμε ότι

$$f(e_i, e_j) = \langle L_f(e_i), e_j\rangle = \lambda_i \delta_{ij}.$$

Τα $\lambda_i \in \mathbb{R}$ σύμφωνα με την πρόταση VII.8.4. □

Ορισμός VIII.2.13. Μία μιάμιση μορφή f θα λέγεται μη-αρνητική αν και μόνο αν είναι ερμητιανή και επιπλέον $f(v,v) \geq 0$ για κάθε $v \in V$. Η μιάμιση μορφή θα λέγεται θετική αν είναι ερμητιανή και επιπλέον $f(v,v) > 0$ για κάθε $v \in V, v \neq 0$.

Παρατήρηση VIII.2.14. Απαιτούμε η f να είναι ερμητιανή γιατί τότε $f(v,v) \in \mathbb{R}$, δείτε την άσκηση VIII.2.25, οπότε η ανισότητα έχει νόημα.

Παρατήρηση VIII.2.15. Μια θετική μιάμιση μορφή είναι ένα εσωτερικό γινόμενο στον V. Μια μη-αρνητική έχει όλες τις ιδιότητες εκτός από το ότι μπορεί να υπάρχουν και μη-μηδενικά $v \in V$ με $f(v,v) = 0$.

Παρατήρηση VIII.2.16. Αν $A = (a_{ij})$ με $a_{ij} = f(v_j, v_i)$ είναι ο πίνακας της f ως προς μία διατεταγμένη βάση $B = \{v_1, \ldots, v_n\}$ του V και $v = \lambda_1 v_1 + \cdots + \lambda_n v_n \in V$ τότε

$$f(v,v) = [v]_B^* A [v]_B = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}\lambda_j \overline{\lambda}_i. \tag{VIII.11}$$

Συνεπώς για να είναι ο A μη-αρνητικός θα πρέπει $A = A^*$ και επιπλέον θα πρέπει η έκφραση της εξίσωσης (VIII.11) να είναι θετική για κάθε $(\lambda_1, \ldots, \lambda_n)^t \in \mathbb{R}^n$.

Πρόταση VIII.2.17. *Έστω $A \in \mathbb{F}^{n,n}$. Η συνάρτηση*

$$\begin{aligned} \mathbb{F}^n \times \mathbb{F}^n &\longrightarrow \mathbb{F} \\ (x,y) &\longmapsto f(x,y) = y^* A x, \end{aligned}$$

δίνει μια θετική μορφή αν και μόνο αν υπάρχει αντιστρέψιμος $P \in \mathbb{F}^{n,n}$ ώστε $A = P^*P$.

Απόδειξη. Παρατηρούμε ότι αν $A = P^*P$ τότε

$$f(x,x) = x^*Ax = x^*P^*Px = (Px)^*Px \geq 0.$$

Επιπλέον αν $(Px)^*Px = 0$ τότε $Px = 0$ και αν ο P είναι αντιστρέψιμος τότε $x = 0$.

Αντιστρόφως, αν η $f(x,y)$ είναι θετική τότε είναι εσωτερικό γινόμενο και μπορούμε να διαλέξουμε μία ορθοκανονική βάση $Q_1, \dots, Q_n \in \mathbb{F}^n$, ώστε

$$f(Q_i, Q_j) = Q_j^* A Q_i = \delta_{ij}.$$

Το τελευταίο έχει ως συνέπεια ότι ο πίνακας P που έχει ως στήλες τα $Q_1, \dots, Q_n$ ικανοποιεί

$$Q^*AQ = \mathbb{I}_n \Rightarrow A = (Q^{-1})^*Q,$$

δηλαδή το ζητούμενο, αφού ο P είναι αντιστρέψιμος. □

Ορισμός VIII.2.18. Για ένα πίνακα $A = (a_{ij})$ οι κύριες ελάσσονες ορίζουσες $\Delta_k(A)$ είναι οι ορίζουσες

$$\Delta_k(A) = \det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} \quad 1 \leq k \leq n.$$

Λήμμα VIII.2.19. *Τα παρακάτω είναι ισοδύναμα για $A \in \mathbb{F}^{n,n}$*

1. *Υπάρχει άνω τριγωνικός πίνακας P με μονάδες στην διαγώνιο ώστε ο πίνακας $B = AP$ να είναι κάτω τριγωνικός.*
2. *Τα $\Delta_k(A) \neq 0$ για κάθε $1 \leq k \leq n$.*

Απόδειξη. Για ένα πίνακα $P = (p_{ij})$ και $A = (a_{ij})$ έχουμε

$$b_{ij} = \sum_{\nu=1}^{n} a_{i\nu} p_{\nu j}.$$

Αν ο πίνακας P είναι άνω τριγωνικός, δηλαδή $p_{\nu,j} = 0$ για $\nu > j$ και $p_{jj} = 1$ τότε η παραπάνω σχέση δίνει:

$$\sum_{\nu=1}^{j-1} a_{i\nu} p_{\nu,j} = b_{ij} - a_{jj} \text{ για κάθε } 1 \leq j \leq n.$$

Για να είναι ο πίνακας $B = (b_{ij})$ κάτω τριγωνικός θα πρέπει $b_{ij} = 0$ για $i < j$ το οποίο συμβαίνει αν και μόνο αν

$$\sum_{\nu=1}^{j-1} a_{i\nu} p_{\nu,j} = -a_{jj}$$

για όλα τα $1 \leq i \leq j-1$ και $2 \leq j \leq n$. Αυτό μας δίνει ένα σύστημα $j-1$ εξισώσεων με $j-1$ αγνώστους $p_{1,j}, \dots, p_{j-1,j}$ με πίνακα συντελεστών τον

$$\begin{pmatrix} a_{11} & \cdots & a_{1,j-1} \\ \vdots & & \vdots \\ a_{j-1} & \cdots & a_{j-1,j-1} \end{pmatrix}$$

ο οποίος έχει ορίζουσα $\Delta_{j-1}(A)$. Αν όλα τα $\Delta_{j-1}(A) \neq 0$ τότε τα παραπάνω συστήματα έχουν μοναδικές λύσεις από τις οποίες υπολογίζεται με μοναδικό τρόπο ο πίνακας P. Δηλαδή αν ισχύει το 2. τότε ισχύει το 1.

Αντιστρόφως, αν ισχύει το 1. και $a^1, \dots, a^n$ είναι οι στήλες του πίνακα A και $b^1, \dots, b^n$ είναι οι στήλες του πίνακα B. Έχουμε δείξει ότι ισχύει

$$b^1 = a^1$$
$$b^r = \sum_{j=1}^{r-1} p_{jr} a^j + a^r \text{ για } r > 1$$

Για σταθερό k, $1 \leq k \leq n$ έχουμε ότι η r-στήλη του πίνακα

$$\begin{pmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & & \vdots \\ b_{k1} & \cdots & b_{kk} \end{pmatrix}$$

δίνεται προσθέτοντας στην r στήλη του πίνακα

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

ένα γραμμικό συνδυασμό των άλλων στηλών, ένας μετασχηματισμός που διατηρεί την ορίζουσα. Αυτό, μαζί με το γεγονός ότι ο πίνακας B είναι τριγωνικός δίνει ότι

$$\Delta_k(A) = \Delta_k(B) = b_{11}b_{22}\cdots b_{kk} \text{ για κάθε } 1 \leq k \leq n.$$

Αφού ο A και ο P είναι αντιστρέψιμοι και ο B είναι αντιστρέψιμος οπότε έχουμε

$$\det(B) = b_{11}b_{22}\cdots b_{nn} \neq 0$$

άρα και όλοι οι $\Delta_k(B) = \Delta_k(A) \neq 0$. □

Θεώρημα VIII.2.20. *Μια μιάμιση μορφή* f *σε ένα πεπερασμένης διάστασης διανυσματικό χώρο* V *με πίνακα* $A = (a_{ij})$, $a_{ij} = f(v_j, v_i)$ *ως προς την διατεταγμένη βάση* $B = \{v_1, \dots, v_n\}$ *είναι θετική αν και μόνο αν* $A = A^*$ *και όλες οι ελάσσονες ορίζουσες* $\Delta_k(A) > 0$.

Απόδειξη. Υποθέτουμε ότι $A = A^*$ και ότι $\Delta_k(A) > 0$ για κάθε $1 \leq k \leq n$. Σύμφωνα με το λήμμα VIII.2.19 υπάρχει μοναδικός άνω τριγωνικός πίνακας P με μονάδες στην διαγώνιο ώστε ο $B = AP$ να είναι κάτω τριγωνικός. Ο πίνακας P^* είναι κάτω τριγωνικός συνεπώς ο πίνακας $P^*B = P^*AP$ είναι επίσης κάτω τριγωνικός. Αφού ο A^* είναι αυτοσυζυγής και ο πίνακας $D = P^*AP$ είναι αυτοσυζυγής και είναι και τριγωνικός, συνεπώς είναι διαγώνιος σύμφωνα με την πρόταση VII.8.10. Έχουμε

$$\Delta_k(D) = \Delta_k(P^*A) = \Delta_k(B) = \Delta_k(A),$$

και ο πίνακας $D = \mathrm{diag}(d_{11}, \dots, d_{nn})$ έχει $\Delta_k(D) = d_{11}d_{22}\cdots d_{kk}$. Η σχέση $\Delta_k(D) > 0$ για $1 \leq k \leq n$ μας δίνει διαδοχικά ότι $d_{kk} > 0$ για κάθε k.

Έχουμε τώρα ότι για τον πίνακα A της f ως προς την βάση B ότι $D = P^*AP$ είναι ο πίνακας της f ως προς την βάση $B' = \{v'_1, \dots, v'_n\}$ η οποία ορίζεται από

$$v'_k = \sum_{i=1}^{n} p_{ij} v_i$$

και αφού ο D είναι διαγώνιος με θετικά στοιχεία στην διαγώνιο έχουμε ότι

$$X^*DX > 0 \text{ για κάθε } X \neq 0.$$

Αντιστρόφως, αν f είναι θετική έχουμε ότι $A = A^*$. Θεωρούμε τον υπόχωρο V_k που παράγεται από τα k πρώτα διανύσματα $B_k := \{v_1, \ldots, v_k\}$ της βάσης B και f_k τον περιορισμό της f στον $V_k \times V_k$. Είναι σαφές ότι η f_k είναι θετική μορφή στο V_k και στην βάση B_k αναπαρίσταται από τον πίνακα

$$A_k := \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}.$$

Σύμφωνα με την πρόταση VIII.2.17 υπάρχει αντιστρέψιμος πίνακας P_k, ώστε $A_k = P_k^* P_k$. Τότε όμως έχουμε

$$\Delta_k(A) = \det(A_k) = \det(P^*P) = \det(P^*)\det(P) = \overline{\det(P)}\det(P) = |\det(P)|^2 > 0.$$

□

Έστω V πεπερασμένης διάστασης διανυσματικός χώρος με εσωτερικό γινόμενο. Αν f είναι μια μη-αρνητική μιάμιση μορφή, υπάρχει μοναδική αυτοσυζυγής γραμμική συνάρτηση $L_f : V \to V$, ώστε

$$f(v, w) = \langle L_f(v), w\rangle \text{ για κάθε } v, w \in V,$$

ώστε $\langle T_f(v), v\rangle \geq 0$.

Ορισμός VIII.2.21. Μία γραμμική συνάρτηση $L : V \to V$, στον πεπερασμένης διάστασης διανυσματικό χώρο V με εσωτερικό γινόμενο θα λέγεται μη-αρνητική αν $L = L^*$ και $\langle Lv, v\rangle \geq 0$ για κάθε $v \in V$. Θα λέγεται θετική αν είναι μη-αρνητική και επιπρόσθετα $\langle Lv, v\rangle = 0$ αν και μόνο αν $v = 0$.

Άσκηση VIII.2.22. Δείξτε ότι για δύο μιάμιση μορφές f, g η εξίσωση

$$\langle f, g\rangle = \operatorname{tr}(L_g^* L_f)$$

ορίζει ένα εσωτερικό γινόμενο στον διανυσματικό χώρο των μιάμιση μορφών.

Άσκηση VIII.2.23. Έστω f μια μιάμιση μορφή σε ένα πεπερασμένης διάστασης διανυσματικό χώρο με εσωτερικό γινόμενο. Δείξτε ότι υπάρχει μια ορθοκανονική βάση του V στην οποία ο πίνακας της f είναι άνω τριγωνικός.

Λύση VIII.2.24. Έστω $L : V \to V$ η γραμμική συνάρτηση ώστε $f(v, w) = \langle Lv, w\rangle$. Σύμφωνα με την πρόταση VII.8.11 υπάρχει ορθοκανονική βάση $B = \{e_1, \ldots, e_n\}$ στην οποία ο (L, B, B) να είναι άνω τριγωνικός. Έχουμε ότι

$$f(e_i, e_j) = \langle Le_i, a_j\rangle = 0,$$

για $j > i$.

Άσκηση VIII.2.25. Δείξτε ότι μια μιάμιση μορφή είναι ερμητιανή αν και μόνο αν $f(v, v) \in \mathbb{R}$ για κάθε $v\ inV$.

Λύση VIII.2.26. Αν η f είναι ερμητιανή τότε $f(v, v) = \overline{f(v, v)}$ για κάθε $v \in V$, συνεπώς $f(v, v) \in \mathbb{R}$.

Αντιστρόφως, έστω $v, w \in V$. Θα πρέπει να δείξουμε ότι $f(v, w) = \overline{f(w, v)}$ για κάθε $v, w \in V$. Έχουμε ότι

$$f(v, w) + f(w, v) = f(v + w, v + w) - f(v, v) - f(w, w) \in \mathbb{R}.$$

Επίσης

$$-if(v,w)+if(w,v)=f(v+iw,v+iw)-f(v,v)-f(w,w)\in\mathbb{R}.$$

Συνεπώς

$$f(v,w)+f(w,v)=\overline{f(v,w)}+\overline{f(w,v)},\qquad -if(v,w)+if(v,w)=i\overline{f(v,w)}-i\overline{f(w,v)}.$$

Αν πολλαπλασιάσουμε την δεύτερη εξίσωση με i και την προσθέσουμε στην πρώτη έχουμε ότι

$$2f(v,w)=2f(w,v).$$

VIII.2.2 Διγραμμικές μορφές

Θα μελετήσουμε τώρα την θεωρία των διγραμμικών μορφών όπως αυτές ορίστηκαν στο VIII.2.1. Είναι σαφές ότι στην περίπτωση ενός πραγματικού διανυσματικού χώρου οι μιάμιση μορφές ταυτίζονται με τις διγραμμικές. Ας δούμε μερικά ακόμα παραδείγματα διγραμμικών μορφών που δουλεύουν εξίσου καλά και στην περίπτωση οποιουδήποτε σώματος $\mathbb{F}$.

Παρατήρηση VIII.2.27. Παρατηρούμε ότι το σύνολο $L(V,V,\mathbb{F})$ των διγραμμικών μορφών $V\times V\to\mathbb{F}$, όπου V διανυσματικός χώρος πάνω από το σώμα $\mathbb{F}$ είναι διανυσματικός χώρος.

Παράδειγμα VIII.2.28. Έστω ένας διανυσματικός χώρος V και έστω $L_1,L_2:V\to\mathbb{F}$. Η συνάρτηση $V\times V\to F$ $f(v,w)=L_1(v)L_2(w)$ είναι διγραμμική.

Παράδειγμα VIII.2.29. Θεωρούμε τον διανυσματικό χώρο $V=\mathbb{F}^{n,m}$ των $n\times m$ πινάκων πάνω από το σώμα $\mathbb{F}$. Για ένα σταθερό πίνακα $A\in\mathbb{F}^{m,m}$ η συνάρτηση

$$\begin{aligned} V\times V &\longrightarrow \mathbb{F}\\ (X,Y)&\longmapsto f_A(X,Y)=\mathrm{tr}(X^tAY)\end{aligned}$$

είναι μια διγραμμική μορφή στον χώρο V.

Μπορούμε στην πραγματικότητα να περιγράψουμε όλες τις τετραγωνικές μορφές σε ένα διανυσματικό χώρο ως εξής: Θεωρούμε μια διατεταγμένη βάση $B=\{v_1,\dots,v_n\}$ του V και $f(v,w)$ μια διγραμμική μορφή. Αν $v=x_1v_1+\dots+x_nv_n$, $w=y_1v_1+\dots+y_nv_n$ τότε έχουμε

$$f(v,w)=\sum_{i=1}^{n}\sum_{j=1}x_iy_jf(v_i,v_j).$$

Οπότε θέτοντας $A=(a_{ij})$ με $a_{ij}=f(v_i,v_j)$ έχουμε ότι

$$f(v,w)=[v]_B^tA[w]_B.$$

Αντιστρόφως, κάθε πίνακας A ορίζει μια διγραμμική μορφή με τον παραπάνω τρόπο.

Ορισμός VIII.2.30. Ο πίνακας A θα λέγεται ο πίνακας της διγραμμικής μορφής f ως προς την διατεταγμένη βάση B και θα τον συμβολίζουμε με $[f]_B$.

Θεώρημα VIII.2.31. *Για ένα πεπερασμένης διάστασης διανυσματικό χώρο V και μια διατεταγμένη βάση B, η συνάρτηση*

$$\begin{aligned} L(V,V,\mathbb{F}) &\longrightarrow \mathbb{F}^{n,n}\\ f&\longmapsto [f]_B\end{aligned}$$

είναι ισομορφισμός διανυσματικών χώρων.

Απόδειξη. Δεδομένης μιας διγραμμικής μορφής f ο πίνακας της $[f]_B$ είναι μονοσήμαντα ορισμένος, ενώ για κάθε πίνακα $A \in \mathbb{F}^{n,n}$ υπάρχει μοναδική διγραμμική συνάρτηση που αντιστοιχεί σε αυτόν. Τέλος, η συνάρτηση αντιστοίχισης είναι γραμμική δηλαδή

$$[\lambda_1 f_1 + \lambda_2 f_2]_B = \lambda_1[f_1]_B + \lambda_2[f_2]_B \text{ για κάθε } \lambda_1, \lambda_2 \in \mathbb{F}, f_1, f_2 \in L(V, V, \mathbb{F}).$$

□

Πόρισμα VIII.2.32. *Αν $B = \{v_1, \ldots, v_n\}$ είναι μια διατεταγμένη βάση του V και $B^* = \{v_1^*, \ldots, v_n^*\}$ είναι η δυική βάση του $V^* = \mathrm{Hom}(V, \mathbb{F})$, δηλαδή $v_i^*(v_j) = \delta_{ij}$ τότε οι n^2 το πλήθος διγραμμικές μορφές*

$$f_{ij}(v, w) = v_i^*(v)v_j^*(w), 1 \le i, j \le n$$

αποτελούν μια βάση του χώρου $L(V, V, \mathbb{F})$. Η διάσταση του χώρου $L(V, V, \mathbb{F})$ είναι n^2.

Απόδειξη. Παρατηρούμε ότι οι f_{ij} είναι πράγματι διγραμμικές μορφές, όπως αυτές ορίστηκαν στο παράδειγμα VIII.2.28 και μάλιστα

$$f_{ij}(v, w) = x_i y_j$$

όπου $[v]_B = (x_1, \ldots, x_n)^t$ $[w]_B = (y_1, \ldots, y_n)^t$. Η γενική διγραμμική μορφή που αντιστοιχεί στον πίνακα $A = (a_{ij})$ γράφεται ως

$$f_A(v, w) = [v]_B^t A[w]_B = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} f_{ij}.$$

Δηλαδή, οι f_{ij} είναι n^2 το πλήθος μορφές, γραμμικά ανεξάρτητες που παράγουν τον χώρο $L(V, V, \mathbb{F})$. □

Παρατήρηση VIII.2.33. Αν P είναι ο πίνακας αλλαγής βάσης ώστε $[v]_B = P[v]_{B'}$ από την διατεταγμένη βάση B στην διατεταγμένη βάση B' τότε

$$f(v, w) = [v]_B^t [f]_B [w]_B = (P[v]_{B'})^t [f]_B P[v]_{B'}$$

και από το μονοσήμαντο του πίνακα της διγραμμικής συνάρτησης έχουμε ότι

$$[f]_{B'} = P^t [f]_B P.$$

Παρατήρηση VIII.2.34. Αν σταθεροποιήσουμε ένα $v \in V$ τότε η διγραμμική μορφή f δίνει μια συνάρτηση

$$\begin{aligned} L_f(v) : V &\longrightarrow \mathbb{F} \\ w &\longmapsto f(v, w) \end{aligned}$$

δηλαδή δίνει μια συνάρτηση

$$\begin{aligned} V &\longrightarrow V^* \\ v &\longmapsto L_f(\cdot) = f(v, \cdot) \end{aligned}$$

Ομοίως, αν σταθεροποιήσουμε ένα $w \in V$ έχουμε

$$\begin{aligned} R_f(w) : V &\longrightarrow \mathbb{F} \\ v &\longmapsto f(v, w) \end{aligned}$$

δηλαδή δίνει μια συνάρτηση

$$\begin{aligned} V &\longrightarrow V^* \\ w &\longmapsto R_f(\cdot) = f(\cdot, w) \end{aligned}$$

Θεώρημα VIII.2.35. *Έστω μία διγραμμική μορφή σε ένα πεπερασμένης διάστασης διανυσματικό χώρο* V *και έστω* $L_f, R_f : V \to V^*$. *Έχουμε ότι* $r(L_f) = r(R_f)$.

Απόδειξη. Αρκεί να αποδείξουμε ότι $\dim \ker R_f = \dim \ker L_f$. Γράφουμε την $f(v,w)$ ως

$$f(v,w) = [v]_B^t A[w]_B.$$

Αν $w \in \ker R_f$ τότε έχουμε ότι $f(v,w) = 0$ για κάθε $v \in V$, δηλαδή $A[w]_B = \mathbf{0} \in \mathbb{F}^{n,n}$. Με τον ίδιο τρόπο αν $v \in \mathrm{Ker} L_f$ έχουμε ότι $f(v,w) = 0$ για κάθε $w \in V$, δηλαδή $A^t[v]_B = 0$. Όμως $\dim \ker A = \dim \ker A^t$ και το ζητούμενο έπεται. □

Ορισμός VIII.2.36. Η τάξη μιας διγραμικής μορφής $r(f)$ ορίζεται να είναι ο $r(L_f) = r(R_f) = r([f]_B)$.

Πόρισμα VIII.2.37. *Για μια διγραμμική μορφή* $f : V \times V \to \mathbb{F}$ *τα παρακάτω είναι ισοδύναμα*

1. $r(f) = \dim V$.
2. *Για κάθε* $v \neq 0, v \in V$ *υπάρχει μη-μηδενικό* $w \in V$ *ώστε* $f(v,w) \neq 0$.
3. *Για κάθε* $w \neq 0, w \in V$ *υπάρχει μη-μηδενικό* $v \in V$ *ώστε* $f(v,w) \neq 0$.

Απόδειξη. Το 2. λέει ότι $\ker L_f = \{0\}$ ενώ το 3. ότι $\ker R_f = \{0\}$. Οι γραμμικές συναρτήσεις L_f και R_f έχουν μηδενικό πυρήνα αν και μόνο αν η τάξη τους, άρα και η τάξη της f είναι n. □

Ορισμός VIII.2.38. Μια διγραμμική μορφή θα λέγεται μη-ιδιόμορφη αν ισχύει μία από τις ισοδύναμες προτάσεις του πορίσματος VIII.2.37.

Ορισμός VIII.2.39. Μια διγραμμική μορφή θα λέγεται συμμετρική αν μόνο αν

$$f(v,w) = f(w,v) \text{ για κάθε } v,w \in V.$$

Παρατήρηση VIII.2.40. Παρατηρούμε ότι η $f(v,w) = [v]_B^t[f]_B[w]_B$ είναι συμμετρική αν και μόνο αν $[f]_B = [f]_B^t$.

Σε κάθε συμμμετρική διγραμμική μορφή αντιστοιχεί μια τετραγωνική μορφή

$$q(v) = f(v,v) = [v]_B^t[f]_B[v]_B. \tag{VIII.12}$$

Η τετραγωνική μορφή καθορίζει την διγραμμική μορφή σύμφωνα με την ταυτότητα

$$f(v,w) = \frac{1}{4}q(v+w) - \frac{1}{4}q(v-w).$$

Θεώρημα VIII.2.41. *Σε ένα πεπερασμένης διάστασης διανυσματικό χώρο πάνω από το σώμα $\mathbb{F}$, χαρακτηριστικής 0 υπάρχει μια διατεταγμένη βάση του V ώστε ο πίνακας $[f]_B$ της συμμετρικής διγραμμικής μορφής f να είναι διαγώνιος.*

Απόδειξη. Ψάχνουμε να βρούμε διατεταγμένη βάση $B = \{v_1, \ldots, v_n\}$ ώστε $f(v_i, v_j) = 0$ αν $i \neq j$. Αν $f = 0$ ή $n = \dim V = 1$ τότε το θεώρημα είναι προφανές. Αν $q(v) = f(v, v) = 0$ για κάθε $v \in V$ τότε η σχέση (VIII.12) μας δίνει ότι η $f = 0$. Υποθέτουμε ότι υπάρχει $v \in V$ ώστε $q(v) \neq 0$. Θεωρούμε τον χώρο $W = \langle v \rangle$ και ορίζουμε τον χώρο $W^{\perp} = \{w \in V : f(v, w) = 0\}$. Θα δείξουμε ότι $V = W \oplus W^{\perp}$. Είναι σαφές ότι $W \cap W^{\perp} = \{0\}$. Πράγματι, για ένα στοιχείο στον W της μορφής λv έχουμε ότι $0 = f(\lambda v, \lambda v) = \lambda^2 f(v, v)$ και αφού $f(v, v) \neq 0$ έχουμε ότι $\lambda = 0$. Κάθε διάνυσμα στο V είναι άθροισμα ενός στοιχείου του W και ενός στοιχείου του $W^{\perp}$. Πράγματι, για $w \in V$, θέτουμε

$$w_1 = w - \frac{f(w, v)}{f(v, v)} v$$

και παρατηρούμε ότι

$$f(w_1, v) = f(w, v) - \frac{f(w, v)}{f(v, v)} f(v, v) = 0.$$

Συνεπώς $w_1 \in W^{\perp}$ και $w = w_1 + \frac{f(w,v)}{f(v,v)} v$. Ο περιορισμός της f στο $W^{\perp}$ είναι μια συμμετρική διγραμμική μορφή και αφού $\dim W^{\perp} = n - 1$ το αποτέλεσμα προκύπτει επαγωγικά. Μπορούμε δηλαδή να υποθέσουμε ότι ο $W^{\perp}$ έχει μια διατεταγμένη βάση $v_1, \ldots, v_{n-1}$, ώστε $f(v_i, v_j) = 0$ για $i \neq j$, $1 \leq i, j \leq n - 1$. Η ζητούμενη βάση είναι η $\{v_1, \ldots, v_{n-1}, v_n\}$. □

Πόρισμα VIII.2.42. *Αν $\mathbb{F}$ είναι ένα υπόσωμα του $\mathbb{C}$ και $A \in \mathbb{F}^{n,n}$, $A^t = A$, τότε υπάρχει αντιστρέψιμος πίνακας P, ώστε $P^t A P$ να είναι διαγώνιος.*

Παρατήρηση VIII.2.43. Στην περίπτωση που $\mathbb{F} = \mathbb{R}$ τον πίνακα P μπορούμε να τον διαλέξουμε ώστε $P^t = P^{-1}$.

Θεώρημα VIII.2.44. *Αν V είναι διανυσματικός χώρος πάνω από το $\mathbb{C}$ και f είναι μια συμμετρική διγραμμική μορφή τότε υπάρχει διατεταγμένη βάση $B = \{v_1, \ldots, v_n\}$ ώστε $f(v_i, v_j) = 0$ αν $i \neq j$,*

$$f(v_i, v_i) = \begin{cases} 1 & \text{αν } 1 \leq i \leq r \\ 0 & \text{αν } i > r \end{cases}$$

Απόδειξη. Υπάρχει μια βάση $B = \{v_1, \ldots, v_n\}$ στην οποία ο πίνακας της f να είναι διαγώνιος. Αν ο $[f]_B$ έχει τάξη $0 \leq r \leq n$ τότε στην διαγώνιο έχουμε ακριβώς r το πλήθος μη-μηδενικά στοιχεία. Αναδιατάσσοντας την βάση μπορούμε να υποθέσουμε ότι αυτά είναι τα r πρώτα. Στην συνέχεια θέτουμε

$$v_j' = \frac{1}{\sqrt{f(v_i, v_i)}} v_i \text{ για κάθε } 1 \leq j \leq r.$$

Η ζητούμενη βάση είναι η $v_1', \ldots, v_r', v_{r+1}, \ldots, v_n$. □

Στους πραγματικούς αριθμούς οι αρνητικοί αριθμοί δεν έχουν τετραγωνικές ρίζες. Μπορούμε όμως να αποδείξουμε ότι

Θεώρημα VIII.2.45. *Έστω V ένας πραγματικός διανυσματικός χώρος διάστασης n και έστω f διγραμμική συμμετρική μορφή. Υπάρχει μια βάση B, ώστε ο πίνακας $[f]_B$ να είναι διαγώνιος και ώστε*

$$f(v_i, v_i) = \begin{cases} \pm 1 & \text{αν } 1 \le i \le r \\ 0 & \text{αν } i > r \end{cases}$$

Επιπλέον το πλήθος των διανυσμάτων της βάσης ώστε $f(v_i, v_i) = 1$ είναι ανεξάρτητο της επιλογής της βάσης.

Απόδειξη. Όπως και πριν υπάρχει μια βάση $B = \{v_1, \ldots, v_n\}$ ώστε

$$\begin{aligned} f(v_i, v_j) &= 0 \text{ αν } i \neq j \\ f(v_i, v_i) &\neq 0 \text{ αν } 1 \le i \le r \\ f(v_i, v_i) &= 0 \text{ αν } i > r \end{aligned}$$

Θέτουμε

$$v_i' = |f(v_i, v_i)|^{-1/2} v_i \text{ αν } 1 \le i \le r.$$

Η ζητούμενη βάση είναι η $\{v_1', \ldots, v_r', v_{r+1}, \ldots, v_n\}$.

Έστω p το πλήθος των $v_1', \ldots, v_r'$ ώστε $f(v_i', v_i') = 1$. Θα δείξουμε ότι το p είναι ανεξάρτητο από την επιλογή της βάσης που κάναμε αν έχουμε την ζητούμενη μορφή. Θέτουμε V^+ τον χώρο που παράγεται από τα διανύσματα $v_1', \ldots, v_r'$ και $f(v_i', v_i') = 1$ και V^- τον χώρο που παράγεται από τα $v_1', \ldots, v_r'$ ώστε $f(v_i', v_i') = -1$. Ισχύει ότι $p = \dim V^+$. Έστω V_0 ο υπόχωρος που παράγεται από τα $v_{r+1}, \ldots, v_n$.

Για κάθε μη-μηδενικό διάνυσμα $v \in V^+$ έχουμε $f(v, v) > 0$ ενώ για κάθε μη-μηδενικό διάνυσμα $v \in V^-$ έχουμε $f(v, v) < 0$. Ισχύει ότι

$$V = V^+ \oplus V^- \oplus V^0. \tag{VIII.13}$$

Έστω ένας υπόχωρος W του V ώστε $f(v, v) > 0$ για κάθε $v \in W$. Ισχύει ότι αν έχουμε ένα άθροισμα

$$w + v^- + v_0 = 0, w \in W, v^- \in V^-, v_0 \in V^0,$$

τότε $w = v^- = v_0 = 0$. Πράγματι,

$$\begin{aligned} 0 &= f(w, w + v^- + v_0) = f(w, w) + f(w, v^-) + f(w, v_0) \\ 0 &= f(v^-, w + v^- + v_0) = f(v^-, w) + f(v^-, v^-) + f(v^-, v_0). \end{aligned}$$

Παρατηρούμε ότι αν $v_0 \in V^0$ τότε $f(v, v_0) = 0$ για κάθε $v \in V$. Συνεπώς οι παραπάνω σχέσεις γίνονται

$$\begin{aligned} 0 &= f(w, w) + f(w, v^-) \\ 0 &= f(v^-, w) + f(v^-, v^-). \end{aligned}$$

Επειδή η διγραμμική είναι συμμετρική η παραπάνω σχέση δίνει ότι $f(w, w) = f(v^-, v^-)$ και επειδή $f(w, w) \ge 0$ και $f(v^-, v^-) \le 0$ έχουμε ότι $f(w, w) = f(v^-, v^-) = 0$ και συνεπώς $w = v^- = 0$ το οποίο δίνει και ότι $v_0 = 0$. Η σχέση (VIII.13) μαζί με την ανεξαρτησία των W, V^-, V^0 επιβάλει ότι $\dim W \le \dim V^+$. Άρα αν σχηματίσουμε τους ίδιους χώρους V_1^+, V_1^-, V_1^0 ως προς μια άλλη βάση θα έχουμε ότι $\dim V_1^+ \le \dim V^+$ και για συμμετρικούς λόγους καταλήγουμε στο ότι $\dim V^+ = \dim V_1^+$. □

Παρατήρηση VIII.2.46.

$$\dim V^+ + \dim V^- = r(f).$$

Η ποσότητα $\dim V^+ - \dim V^-$ λέγεται υπογραφή της διγραμμικής μορφής και μαζί με το n και $r(f)$ μπορούμε να ανακτήσουμε τις διαστάσεις των χώρων V^+, V^-, V^0.

VIII.3 Φασματική Θεωρία

Θεώρημα VIII.3.1. *Έστω $L : V \to V$ μία κανονική γραμμική συνάρτηση, όπου V είναι πεπερασμένης διάστασης μιγαδικός διανυσματικός χώρος ή είναι μια αυτοσυζυγής γραμμική συνάρτηση σε ένα πεπερασμένης διάστασης πραγματικό χώρο. Θεωρούμε το σύνολο των διαφορετικών ανά δύο ιδιοτιμών της L, $\lambda_1, \dots, \lambda_k$, $k \leq \dim V$. Έστω E_i ο ιδιόχωρος της ιδιοτιμής λ_i. Τότε οι ιδιόχωροι E_i είναι ανά δύο κάθετοι. Επίσης, αν $\pi_i : V \to E_i$ είναι η προβολή του V στον χώρο E_i, τότε*

$$L = \lambda_1\pi_1 + \cdots + \lambda_k\pi_k.$$

Απόδειξη. Στην περίπτωση που η L είναι κανονική απεικόνιση επί μιγαδικού διανυσματικού χώρου για να δείξουμε ότι οι χώροι E_i, E_j είναι κάθετοι αν $i \neq j$, διαλέγουμε $v_i \in E_i, v_j \in E_j$ και υπολογίζουμε

$$\lambda_i\langle e_i, e_j\rangle = \langle Le_i, e_j\rangle = \langle e_i, L^* e_j\rangle \overset{\text{VII.8.8}}{=} \langle e_i, \overline{\lambda_j} e_j\rangle = \lambda_j\langle e_i, e_j\rangle$$

από όπου προκύπτει το ζητούμενο, αφού $\lambda_i \neq \lambda_j$.

Στην περίπτωση που η L είναι αυτοσυζυγής έχουμε δείξει στην πρόταση VII.8.4 με παρόμοιο τρόπο ότι οι ιδιόχωροι είναι κάθετοι μεταξύ τους.

Αφού οι κανονικές γραμμικές συναρτήσεις είναι διαγωνοποιήσιμες ο χώρος V έχει μια βάση αποτελούμενη από ιδιοδιανύσματα οπότε

$$V = E_1 + \cdots + E_k.$$

Έχουμε ότι αν $v_i \in E_i$ και $v_1 + \cdots + v_k = 0$, τότε

$$0 = \langle v_i, v_1 + \cdots + v_k\rangle = \langle v_i, v_i\rangle = \|v_i\|^2,$$

άρα κάθε $v_i = 0$ και το άθροισμα είναι ευθύ. Συνεπώς

$$\mathrm{Id}_V = \pi_1 + \cdots + \pi_k$$

από όπου έχουμε

$$L = L \circ \pi_1 + \cdots + L \circ \pi_k = \lambda_1\pi_1 + \cdots + \lambda_k\pi_k.$$

□

Πρόταση VIII.3.2. *Οι προβολές στους ιδιόχωρους είναι πολυωνιμικές συναρτήσεις του L. Για*

$$p_i(x) = \prod_{\substack{j=1\\ j\neq i}}^{n} \left(\frac{x - \lambda_i}{\lambda_j - \lambda_i}\right)$$

έχουμε $p_i(L) = \pi_i$.

Απόδειξη. Αφού $\pi_i\pi_j = \delta_{ij}$ έχουμε ότι

$$L^2 = \lambda_1^2\pi_1 + \cdots\lambda_k^2\pi_k,$$

από όπου επαγωγικά προκύπτει ότι

$$L^n = \lambda_1^n\pi_1 + \cdots + \lambda_k\pi_k^n \text{ για κάθε } n \geq 0.$$

Συνεπώς, για ένα οποιοδήποτε πολυώνυμο

$$f(x) = \sum_{n=0}^{d} a_n x^n,$$

έχουμε ότι

$$f(L) = \sum_{n=0}^{d} a_n L^n = \sum_{n=0}^{d} a_n \sum_{j=1}^{k} \lambda_j^n \pi_j = \sum_{j=1}^{k} \left(\sum_{n=0}^{d} a_n \lambda_j^n \right) = \sum_{j=1}^{k} f(\lambda_j \pi_j).$$

Αφού $p_i(\lambda_j) = \delta_{ij}$, έχουμε ότι $\pi_j(L) = \pi_j$. □

Παρατήρηση VIII.3.3. Οι προβολές $\pi_1, \dots, \pi_k$ εξαρτώνται από τον L. Την γραφή

$$\mathrm{Id}_V = \pi_1 + \cdots + \pi_k$$

θα την λέμε ανάλυση της ταυτότητας ορισμένη από τον L.

Ορισμός VIII.3.4. Θεωρούμε μια διαγωνοποιήσιμη κανονική γραμμική συνάρτηση $L : V \to V$ σε ένα πεπεπερασμένης διάστασης διανυσματικό χώρο V με εσωτερικό γινόμενο και έστω

$$L = \sum_{j=1}^{j} \lambda_j \pi_j$$

η φασματική της ανάλυση. Αν $f : \mathbb{C} \supseteq U \to \mathbb{C}$ είναι μια συνάρτηση που το πεδίο ορισμού της περιέχει τα $\{\lambda_1, \dots, \lambda_k\}$, τότε μπορούμε να ορίσουμε την γραμμική συνάρτηση $f(L) : V \to V$ ως εξής:

$$f(L) = \sum_{j=1}^{k} f(\lambda_j) \pi_j.$$

Θεώρημα VIII.3.5. *Έστω $L : V \to V$ διαγωνοποιήσιμη κανονική γραμμική συνάρτηση με σύνολο ιδιοτιμών $S = \{\lambda_1, \dots, \lambda_k\}$, όπου ο V είναι πεπερασμένης διάστασης διανυσματικός χώρος με εσωτερικό γινόμενο. Αν $f : \mathbb{C} \supset U \to \mathbb{C}$ συνάρτηση με το πεδίο ορισμού της U να περιέχει το S. Τότε ο $f(T)$ είναι διαγωνοποιήσιμη κανονική γραμμική συνάρτηση με σύνολο ιδιοτιμών $f(S)$. Αν $Q : V \to V'$ είναι μοναδιαία γραμμική συνάρτηση και $L' = QLQ^{-1}$, τότε το σύνολο ιδιοτιμών της L' είναι επίσης το S και*

$$f(L') = Qf(L)Q^{-1}.$$

Απόδειξη. Για να αποδείξουμε ότι η $f(L)$ είναι κανονική υπολογίζουμε ότι

$$f(L)^* = \sum_{j=1}^{k} \overline{f(\lambda_i)} \pi_j,$$

συνεπώς

$$f(L)f(L)^* = \sum_{j=1}^{k} \sum_{i=1}^{k} f(\lambda_i) \overline{f(\lambda_i)} \pi_j \pi_i \overset{\pi_j \pi_i = \delta_{ij} \pi_i}{=} \sum_{i=1}^{k} |f(\lambda_i)|^2 \pi_i = f(L)^* f(L).$$

Επιπλέον είναι σαφές ότι για κάθε $v \in \pi_j(V)$ έχουμε ότι

$$f(L)v = f(\lambda_j)v,$$

συνεπώς το σύνολο $f(S)$ περιέχεται στο σύνολο των ιδιοτιμών του $f(L)$. Αντιστρόφως, αν $v \neq 0$ και

$$f(L)v = \lambda v,$$

τότε αφού $v = \sum_{j=1}^{k} \pi_j v$ έχουμε ότι

$$f(L)v = \sum_{j=1}^{k} f(L)\pi_j v = \sum_{j=1}^{k} f(\lambda_j)\pi_j v = \sum_{j=1}^{k} \lambda \pi_j v$$

Από εδώ έχουμε ότι

$$\left\| \sum_{j=1}^{k} f(\lambda_j) - \lambda \pi_j v \right\|^2 = \sum_{j=1}^{k} |f(\lambda_j) - \lambda|^2 \|\pi_j v\|^2 = 0$$

δηλαδή $f(\lambda_j) - \lambda = 0$ ή $\pi_j v = 0$. Συνεπώς $f(\lambda_i) = \lambda$ για κάποιο i και συνεπώς το $f(S)$ είναι το σύνολο των ιδιοτιμών του $f(L)$.

Υποθέτουμε $Q : V \to V'$ είναι μοναδιαία γραμμική συνάρτηση και ότι $L' = QLQ^{-1}$. Τότε η εξίσωση

$$Lv = \lambda v$$

ισχύει αν και μόνο αν

$$L'Qv = \lambda Qv.$$

Συνεπώς, το S είναι επίσης το σύνολο των ιδιοτιμών του L' και η Q μεταφέρει τον ιδιόχωρο της ιδιοτιμής λ για τον L στον ιδιόχωρο της ιδιοτιμής λ για τον L'. Έτσι γράφουμε την φασματική ανάλυση για τον L

$$L' = \sum_{j=1}^{j} \lambda_j \pi'_j, \text{ όπου } \pi'_j = Q\pi_j Q^{-1}$$

και έχουμε

$$f(L') = \sum_{j=1}^{k} f(\lambda_j)\pi'_j = \sum_{j=1}^{k} f(\lambda_j) Q\pi_j Q^{-1} = Q\left(\sum_{j=1}^{k} f(\lambda_j)\pi_j\right) Q^{-1} = Qf(L)Q^{-1}.$$

□

Άσκηση VIII.3.6. Έστω L μια διαγωνοποιήσιμη γραμμική συνάρτηση $L : V \to V$ σε ένα πεπερασμένης διάστασης χώρο V με εσωτερικό γινόμενο. Δείξτε ότι

1. L είναι αυτοσυζυγής αν και και μόνο αν οι ιδιοτιμές της είναι πραγματικές.
2. L είναι μη-αρνητική αν και μόνο αν οι ιδιοτιμές της είναι μη-αρνητικές.
3. L είναι μοναδιαία αν και μόνο αν οι ιδιοτιμές της έχουν μέτρο 1.

Λύση VIII.3.7. Υποθέτουμε ότι η L έχει την φασματική ανάλυση

$$L = \lambda_1\pi_1 + \cdots + \lambda_k\pi_k.$$

Η L είναι αυτοσυζυγής αν και μόνο αν $L = L^*$ οπότε η παραπάνω εξίσωση μας δίνει:

$$(\lambda_1 - \overline{\lambda}_1)\pi_1 + \cdots + (\lambda_k - \overline{\lambda}_k)\pi_k = 0.$$

Η παραπάνω σχέση είναι αληθής αν και μόνο αν $\lambda_i = \overline{\lambda}_i$ για κάθε i. Πράγματι αν $\lambda_i = \overline{\lambda}_i$ τότε η σχέση είναι αληθής και αν η σχέση είναι αληθής τότε συνθέτοντας με την π_j και χρησιμοποιώντας ότι καμία π_j δεν είναι η μηδενική συνάρτηση έχουμε $\lambda_i = \overline{\lambda}_i$.

Για να εξετάσουμε τις μη-αρνητικές κανονικές γραμμικές συναρτήσεις υπολογίζουμε

$$\langle Lv, v\rangle = \langle \sum_{j=1}^{k} \lambda_k \pi_k v, \sum_{j=1}^{k} \pi_k v = \sum_{j=1}^{k}\sum_{i=1}^{k} \lambda_j \langle \pi_j v, \pi_i v\rangle = \sum_{j=1}^{k} \lambda_j \|\pi_j v\|^2.$$

Στην παραπάνω σχέση χρησιμοποιήσαμε το ότι $\langle \pi_j v, \pi_i v \rangle = 0$ για $i \neq j$. Συνεπώς η σχέση $\langle Lv, v \rangle \geq 0$ ισχύει αν και μόνοα $\lambda_j \geq 0$ για κάθε j.

Τέλος για τις μοναδιαίες γραμμικές συναρτήσεις έχουμε ότι

$$LL^* = |\lambda_1|^2 \pi_1 + \cdots + |\lambda_k|^2 \pi_k,$$

οπότε αν $LL^* = Id_V$ τότε

$$Id_V = |\lambda_1|^2 \pi_1 + \cdots + |\lambda_k|^2 \pi_k,$$

οπότε συνθέτοντας με π_i έχουμε

$$\pi_i = |\lambda_i|^2 p_i,$$

και αφού $\pi_i \neq 0$ καταλλήγουμε στο $|\lambda_j| = 1$ για κάθε j.