

TEM-401–Θέματα στην Αριθμητική Ανάλυση

Μιχάλης Πλεξουσάκης
Τμήμα Εφαρμοσμένων Μαθηματικών
Πανεπιστήμιο Κρήτης

Χειμερινό Εξάμηνο 2011-12

Πρόλογος

Οι σημειώσεις αυτές γράφτηκαν το φθινόπωρο του 2011 για τις διδακτικές ανάγκες του μαθήματος “Θέματα στην Αριθμητική Ανάλυση”. Σκοπός του συγκεκριμένου μαθήματος είναι η παρουσίαση και η ανάλυση θεμάτων, από διάφορες περιοχές της Αριθμητικής Ανάλυσης, τα οποία καλύπτονται ελάχιστα ή καθόλου σε εισαγωγικά μαθήματα της περιοχής.

1 Ιδιοτιμές και ιδιοδιανύσματα

Από την περιοχή της αριθμητικής γραμμικής άλγεβρας παρουσιάζεται το πρόβλημα του υπολογισμού ιδιοτιμών και ιδιοδιανυσμάτων ενός πίνακα.

1.1 Βασικές έννοιες από την άλγεβρα πινάκων

Με $\mathbb{R}^{m \times n}$ θα συμβολίζουμε το διανυσματικό χώρο των πραγματικών $m \times n$ πινάκων. Το στοιχείο του πίνακα $A \in \mathbb{R}^{m \times n}$ στη γραμμή i , στήλη j θα συμβολίζεται ως a_{ij} .

$$A \in \mathbb{R}^{m \times n} \iff A = [a_{ij}] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

Οι βασικές πράξεις πινάκων είναι η πρόσθεση ($\mathbb{R}^{m \times n} \times \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$)

$$C = A + B \quad c_{ij} = a_{ij} + b_{ij},$$

πολλαπλασιασμός με πραγματικό αριθμό ($\mathbb{R} \times \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$)

$$C = \alpha A \quad c_{ij} = \alpha a_{ij},$$

ο πολλαπλασιασμός πινάκων ($\mathbb{R}^{m \times p} \times \mathbb{R}^{p \times n} \rightarrow \mathbb{R}^{m \times n}$)

$$C = AB \quad c_{ij} = \sum_{k=1}^p a_{ik} b_{kj},$$

η αναστροφή πίνακα ($\mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{n \times m}$)

$$C = A^T \quad c_{ij} = a_{ji},$$

και η παραγωγήιση ($\mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$)

$$C = (c_{ij}(x)) \quad \dot{C} = \frac{d}{dx}C = [\dot{c}_{ij}(x)].$$

Οι $n \times n$ πίνακες θα ονομάζονται *τετραγωνικοί*. Ο $n \times n$ ταυτοτικός πίνακας συμβολίζεται με I_n και η k -στήλη του με $e_k^{(n)}$ (όταν η διάσταση n εννοείται από το περιεχόμενο θα παραλείπεται):

$$I_n = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \quad e_k^{(n)} = (0, \dots, 0, 1, \dots, 0)^T.$$

Αν $A, B \in \mathbb{R}^{n \times n}$ ικανοποιούν $AB = I$, θα λέμε ότι ο B είναι ο *αντίστροφος* του A και θα συμβολίζεται με A^{-1} . Αν ο αντίστροφος ενός πίνακα υπάρχει τότε ο πίνακας ονομάζεται *αντιστρέψιμος*. Με A^{-T} θα συμβολίζουμε τον πίνακα $(A^{-1})^T = (A^T)^{-1}$. Παρατηρήστε ότι $(A^{-1})^{-1} = A$ και ότι αν $A, B \in \mathbb{R}^{n \times n}$ είναι αντιστρέψιμοι τότε ο πίνακας AB είναι επίσης αντιστρέψιμος και μάλιστα $(AB)^{-1} = B^{-1}A^{-1}$.

Αν $A = (a) \in \mathbb{R}^{1 \times 1}$ τότε η ορίζουσα του A είναι $\det(A) = a$. Αν $A \in \mathbb{R}^{n \times n}$ έχουμε

$$\det(A) = \sum_{j=1}^n (-1)^{j+1} a_{1j} \det(A_{1j}),$$

όπου A_{1j} είναι ο $(n-1) \times (n-1)$ πίνακας που προκύπτει από τον A αν παραλείψουμε την πρώτη γραμμή και την j στήλη. Χρήσιμες ιδιότητες της ορίζουσας είναι:

1. $\det(AB) = \det(BA) = \det(A) \det(B)$, $A, B \in \mathbb{R}^{n \times n}$
2. $\det(A^T) = \det(A)$, $A \in \mathbb{R}^{n \times n}$
3. $\det(cA) = c^n \det(A)$, $c \in \mathbb{R}, A \in \mathbb{R}^{n \times n}$
4. $\det(A) \neq 0 \iff A$ είναι αντιστρέψιμος, $A \in \mathbb{R}^{n \times n}$

Το *ίχνος* (trace) ενός τετραγωνικού πίνακα $A \in \mathbb{R}^{n \times n}$ είναι

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}.$$

Είναι εύκολο να δειχθεί ότι $\text{tr}(AB) = \text{tr}(BA)$ και $\text{tr}(A+B) = \text{tr}(A) + \text{tr}(B)$.

Με \mathbb{R}^m θα συμβολίζουμε το χώρο $\mathbb{R}^{m \times 1}$ των διανυσμάτων-στηλών. Αν $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$ και $y = Ax$ τότε

$$y_i = \sum_{j=1}^n a_{ij}x_j, \quad i = 1, \dots, m.$$

Το *εξωτερικό γινόμενο* των $x \in \mathbb{R}^m, y \in \mathbb{R}^n$ ορίζεται ως

$$xy^T = \begin{bmatrix} x_1y_1 & \cdots & x_1y_n \\ \vdots & & \vdots \\ x_my_1 & \cdots & x_my_n \end{bmatrix} \in \mathbb{R}^{m \times n}$$

και το *εσωτερικό γινόμενο* (αν $m = n$) ως

$$x^T y = \sum_{i=1}^n x_i y_i = y^T x.$$

Οι παραπάνω έννοιες γενικεύονται εύκολα στην περίπτωση των μιγαδικών πινάκων. Θα συμβολίζουμε με $\mathbb{C}^{m \times n}$ το διανυσματικό χώρο των $m \times n$ μιγαδικών πινάκων. Αν $A \in \mathbb{C}^{m \times n}$ τότε ο συζυγής ανάστροφος A^H ορίζεται ως $A^H = (\overline{a_{ji}})$. Το εσωτερικό γινόμενο στο χώρο \mathbb{C}^n είναι

$$x^H y = \sum_{i=1}^n \bar{x}_i y_i = \overline{y^H x}$$

1.2 Ειδικοί πίνακες

Υπάρχουν πολλοί σημαντικοί τύποι τετραγωνικών πινάκων. Λέμε ότι ένας πίνακας A είναι

1. *συμμετρικός* αν A είναι πραγματικός και $A^T = A$.
2. *αντισυμμετρικός* αν A είναι πραγματικός και $A^T = -A$.
3. *ερμιτιανός* αν $A^H = A$.
4. *ορθογώνιος* αν A είναι πραγματικός και $AA^T = A^T A = I$.
5. *ορθομοναδιαίος* (unitary) αν $AA^H = A^H A = I$
6. *κανονικός* (normal) αν $AA^H = A^H A$
7. *θετικά ορισμένος* αν $x^T A x > 0, 0 \neq x \in \mathbb{R}^n$

1.3 Ανεξαρτησία, ορθογωνιότητα, υπόχωροι

Το υποσύνολο $\{a_1, \dots, a_n\}$ του \mathbb{R}^m λέγεται *γραμμικά ανεξάρτητο* αν

$$\sum_{j=1}^n \beta_j a_j = 0 \iff \beta_1 = \dots = \beta_n = 0.$$

Ένας *υπόχωρος* του \mathbb{R}^m είναι ένα υποσύνολο το οποίο είναι επίσης διανυσματικός χώρος. Το σύνολο όλων των γραμμικών συνδιασμών των $a_1, \dots, a_n \in \mathbb{R}^m$ είναι υπόχωρος και συμβολίζεται με $\text{span}\{a_1, \dots, a_n\}$.

Δύο σημαντικοί υπόχωροι που σχετίζονται με ένα πίνακα $A \in \mathbb{R}^{m \times n}$ είναι η *εικόνα* (range) του A

$$R(A) = \{y \in \mathbb{R}^m \mid y = Ax \text{ για κάποιο } x \in \mathbb{R}^n\},$$

και ο *μηδενχώρος* (null space)

$$N(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}.$$

Αν $A = [a_1, \dots, a_n]$ τότε $R(A) = \text{span}\{a_1, \dots, a_n\}$. Ο *βαθμός* (rank) ενός πίνακα A είναι

$$\text{rank}(A) = \dim[R(A)]$$

Μια και $\text{rank}(A) = \text{rank}(A^T)$, ο βαθμός ενός πίνακα είναι ο μέγιστος αριθμός γραμμικά ανεξάρτητων γραμμών ή στηλών. Για κάθε $A \in \mathbb{R}^{m \times n}$ έχουμε $\dim[N(A)] + \text{rank}(A) = n$. Επιπλέον, αν $m = n$ τα παρακάτω είναι ισοδύναμα:

1. A είναι αντιστρέψιμος
2. $N(A) = \{0\}$
3. $\text{rank}(A) = n$

Το σύνολο διανυσμάτων $\{x_1, \dots, x_k\}$ του \mathbb{R}^m λέγεται *ορθογώνιο* αν $x_i^T x_j = 0$ όταν $i \neq j$ και *ορθοκανονικό* όταν $x_i^T x_j = \delta_{ij}$. Γενικότερα, μια συλλογή υποχώρων S_1, \dots, S_k του \mathbb{R}^m είναι αμοιβαία ορθογώνιοι αν $x^T y = 0$ όποτε $x \in S_i, y \in S_j$ και $i \neq j$. Το *ορθογώνιο συμπλήρωμα* ενός υπόχωρου $S \subset \mathbb{R}^m$ είναι

$$S^\perp = \{y \in \mathbb{R}^m \mid y^T x = 0 \text{ for all } x \in S\}$$

Μπορεί να δειχθεί ότι $\dim(S) + \dim(S^\perp) = m$ και, για $A \in \mathbb{R}^{m \times n}$, ότι $N(A) = R(A)^\perp$.

Τα διανύσματα v_1, \dots, v_k αποτελούν μια *ορθοκανονική βάση* του υπόχωρου $S \subset \mathbb{R}^m$ αν είναι ορθοκανονικά και παράγουν τον S . Μια τέτοια βάση μπορεί πάντα να επεκταθεί σε ορθοκανονική βάση $\{v_1, \dots, v_m\}$ του \mathbb{R}^m . Σε αυτή την περίπτωση

$$S^\perp = \text{span}\{v_{k+1}, \dots, v_m\}.$$

1.4 Ιδιοτιμές και ιδιοδιανύσματα. Θεωρητικό υπόβαθρο

Οι *ιδιοτιμές* (eigenvalues) $\lambda_i(A)$, $i = 1, \dots, n$, ενός τετραγωνικού πίνακα $A \in \mathbb{R}^{n \times n}$ είναι οι ρίζες του *χαρακτηριστικού πολυωνύμου* του πίνακα A

$$p_A(\lambda) = \det(A - \lambda I).$$

Το *φάσμα* (spectrum) του A είναι το σύνολο

$$\text{sp}(A) = \bigcup_{i=1}^n \{\lambda_i(A)\}$$

και η *φασματική ακτίνα* (spectral radius) $\rho(A) = \{|\lambda_i(A)| : 1 \leq i \leq n\}$. Από τον ορισμό των ιδιοτιμών προκύπτουν οι σχέσεις

$$\text{tr}(A) = \sum_{i=1}^n \lambda_i(A), \quad \det(A) = \prod_{i=1}^n \lambda_i(A). \quad (1.1)$$

Υπενθυμίζουμε ότι αν $p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ με $a_n \neq 0$, τότε οι ρίζες της εξίσωσης $p(x) = 0$ ικανοποιούν

$$x_1 + x_2 + \dots + x_n = -\frac{a_{n-1}}{a_n}, \quad x_1 x_2 \dots x_n = (-1)^n \frac{a_0}{a_n},$$

ή γενικότερα

$$\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} x_{i_1} x_{i_2} \dots x_{i_k} = -\frac{a_{n-k}}{a_n}.$$

Σε κάθε ιδιοτιμή λ αντιστοιχεί τουλάχιστον ένα διάνυσμα p τέτοιο ώστε

$$p \neq 0 \text{ και } Ap = \lambda p. \quad (1.2)$$

Το διάνυσμα p ονομάζεται *ιδιοδιάνυσμα* (eigenvector) που αντιστοιχεί στην ιδιοτιμή λ . Επομένως, αν $\lambda \in \text{sp}(A)$ η διάσταση του *ιδιόχωρου* $\{v \in \mathbb{R}^n : Av = \lambda v\}$ είναι τουλάχιστον ένα.

Παρατηρήστε ότι αν ο βαθμός του πίνακα $A - \lambda I$ είναι μικρότερος του $n - 1$ τότε θα υπάρχουν περισσότερα του ενός γραμμικά ανεξάρτητα διανύσματα p που ικανοποιούν την σχέση (1.2). Είναι επίσης προφανές ότι αν p ικανοποιεί την (1.2) τότε και κp ικανοποιεί την ίδια εξίσωση για οποιαδήποτε τιμή της σταθεράς κ . Ακόμα και όταν $\text{rank}(A - \lambda I) = n - 1$ το ιδιοδιάνυσμα p που αντιστοιχεί στην ιδιοτιμή λ είναι μοναδικό αν εξαιρέσει κανείς πολλαπλασιαστικές σταθερές. Πολύ συχνά διαλέγουμε την πολλαπλασιαστική σταθερά κ έτσι ώστε $\|p\|_2 = 1$.

Μια και η ορίζουσα ενός πίνακα είναι ίση με την ορίζουσα του ανάστροφού του, προκύπτει ότι οι πίνακες A και A^T έχουν τις ίδιες ιδιοτιμές. Τα ιδιοδιανύσματα είναι, γενικά, διαφορετικά. Έστω λ_i οι ιδιοτιμές του πίνακα A και x_i, y_i τα αντίστοιχα ιδιοδιανύσματα των πινάκων A και A^T . Από τις σχέσεις

$$x_i^T A^T = \lambda_i x_i^T, \quad A^T y_j = \lambda_j y_j,$$

έχουμε $(\lambda_i - \lambda_j)x_i^T y_j = 0$, επομένως

$$x_i^T y_j = 0, \quad \lambda_i \neq \lambda_j. \quad (1.3)$$

Σημειώστε ότι επειδή τα x_i και y_j είναι, εν γένει, μιγαδικά, $x_i^T y_j$ δεν είναι εσωτερικό γινόμενο.

Στην περίπτωση που οι ιδιοτιμές ενός πίνακα είναι διακριτές, η δομή του συστήματος ιδιοδιανυσμάτων είναι αρκετά απλή. Ας υποθέσουμε ότι $\lambda_1, \dots, \lambda_n$ είναι οι διακριτές ιδιοτιμές ενός πίνακα A και x_1, \dots, x_n τα αντίστοιχα ιδιοδιανύσματα. Είναι εύκολο να δειχθεί ότι τα ιδιοδιανύσματα είναι γραμμικά ανεξάρτητα. Διαφορετικά, έστω r ο μικρότερος αριθμός γραμμικά εξαρτημένων διανυσμάτων και ας υποθέσουμε ότι αυτά είναι τα x_1, x_2, \dots, x_r . Επομένως

$$\sum_{i=1}^r \alpha_i x_i = 0, \quad \alpha_i \neq 0, \quad i = 1, \dots, r. \quad (1.4)$$

Προφανώς $r \geq 2$ μια και κανένα από τα x_i δεν είναι το μηδενικό διάνυσμα. Πολλαπλασιάζοντας την (1.4) από αριστερά με τον A έχουμε

$$\sum_{i=1}^r \alpha_i \lambda_i x_i = 0.$$

Πολλαπλασιάζοντας τώρα την (1.4) με λ_r και αφαιρώντας την προηγούμενη σχέση έχουμε

$$\sum_{i=1}^{r-1} \alpha_i (\lambda_r - \lambda_i) x_i = 0, \quad \alpha_i \neq 0, \quad i = 1, \dots, r-1.$$

Η σχέση όμως αυτή λέει ότι τα x_1, \dots, x_{r-1} είναι γραμμικά εξαρτημένα το οποίο έρχεται σε αντίθεση με την υπόθεσή μας. Άρα, τα n ιδιοδιανύσματα είναι γραμμικά ανεξάρτητα και παράγουν τον \mathbb{C}^n . Παρόμοια επιχειρήματα δείχνουν και τη μοναδικότητα των ιδιοδιανυσμάτων modulo πολλαπλασιαστικές σταθερές.

Αν y_j είναι τα ιδιοδιανύσματα του A^T , όπως δείξαμε παραπάνω, $x_i^T y_j = 0$, όταν $\lambda_i \neq \lambda_j$. Στην περίπτωση που οι ιδιοτιμές είναι διακριτές, έχουμε επιπλέον

$$x_i^T y_i \neq 0, \quad i = 1, \dots, n. \quad (1.5)$$

Σε διαφορετική περίπτωση το x_i θα ήταν κάθετο προς όλα τα διανύσματα y_1, y_2, \dots, y_n επομένως θα έπρεπε να είναι το μηδενικό διάνυσμα. Αν κανονικοποιήσουμε τα ιδιοδιανύσματα x_i, y_j έτσι ώστε

$$y_i^T x_i = 1, \quad i = 1, \dots, n, \quad (1.6)$$

οι σχέσεις (1.3), (1.6) συνεπάγονται ότι ο πίνακας Y^T όπου η i γραμμή είναι το ιδιοδιάνυσμα y_i^T είναι ο αντίστροφος του πίνακα X , του οποίου η i στήλη είναι το ιδιοδιάνυσμα x_i . Γράφοντας τις σχέσεις $Ax_i = \lambda x_i$ στη μορφή

$$AX = X \text{diag}(\lambda_i)$$

και πολλαπλασιάζοντας από αριστερά με X^{-1} έχουμε

$$X^{-1}AX = Y^T AX = \text{diag}(\lambda_i) \quad (1.7)$$

Μετασχηματισμοί της παραπάνω μορφής ονομάζονται *μετασχηματισμοί ομοιότητας* και έχουν θεμελιώδη θεωρητική αλλά και πρακτική χρησιμότητα στο πρόβλημα των ιδιοτιμών. Η συζήτηση παραπάνω έδειξε την ορθότητα της πρότασης

Πρόταση 1.1. Έστω $A \in \mathbb{C}^{n \times n}$ με διακριτές ιδιοτιμές $\lambda_i, i = 1, \dots, n$. Τότε υπάρχει αντιστρέψιμος πίνακας X τέτοιος ώστε

$$X^{-1}AX = \text{diag}(\lambda_i) \quad (1.8)$$

Επιπλέον, οι στήλες του πίνακα X είναι τα ιδιοδιανύσματα του A .

Αντίστροφα, αν $X^{-1}AX = \text{diag}(\lambda_i)$ έχουμε ισοδύναμα $AX = X \text{diag}(\lambda_i)$ δηλαδή η i στήλη του X είναι ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή λ_i .

Τέλος, παρατηρήστε οι ιδιοτιμές ενός πίνακα είναι αναλοιώτες κάτω από μετασχηματισμούς ομοιότητας, μια και

$$Ax = \lambda x \Rightarrow X^{-1}Ax = \lambda X^{-1}x \Rightarrow X^{-1}AX(X^{-1}x) = \lambda(X^{-1}x),$$

αλλά τα ιδιοδιανύσματα πολλαπλασιάζονται με X^{-1} . Πολλές από τις αριθμητικές μεθόδους για τον υπολογισμό ιδιοτιμών που θα δούμε στην συνέχεια βασίζονται στη κατασκευή ενός μετασχηματισμού ομοιότητας ο οποίος παράγει ένα πίνακα για τον οποίο το πρόβλημα ιδιοτιμών μπορεί να λυθεί απλούστερα.

Αναφέρουμε τέλος ότι στην περίπτωση που ο πίνακας A είναι ερμιτιανός τότε οι ιδιοτιμές του είναι πραγματικές. Πράγματι, από τη σχέση $Ax = \lambda x$ έχουμε ισοδύναμα $x^H Ax = \lambda x^H x$. Προφανώς $x^H x$ είναι πραγματικός αριθμός και επιπλέον $(x^H Ax)^H = x^H A^H x = x^H Ax$, άρα λ πραγματικός. Παρατηρήστε επίσης ότι από την σχέση $A^H x = \lambda x$ έχουμε $A^T \bar{x} = \bar{\lambda} \bar{x}$ άρα τα ιδιοδιανύσματα του A^T είναι τα μιγαδικά συζυγή των ιδιοδιανυσμάτων του A . Σύμφωνα με τον συμβολισμό που εισάγαμε νωρίτερα έχουμε $y_i = \bar{x}_i$ και επομένως

$$x_i^H x_j = 0, \quad i \neq j, \quad x_i^H x_i \neq 0$$

Αν κανονικοποιήσουμε έτσι ώστε $x_i^H x_i = 1$ οι παραπάνω σχέσεις συνεπάγονται ότι ο πίνακας X με στήλες τα ιδιοδιανύσματα x_i ικανοποιεί $X^H X = I$, είναι δηλαδή ορθομοναδιαίος. Αποδείξαμε λοιπόν την πρόταση

Πρόταση 1.2. Έστω ότι ο πίνακας A έχει διακριτές ιδιοτιμές. Τότε

- Αν ο A είναι ερμιτιανός τότε υπάρχει ορθομοναδιαίος πίνακας U τέτοιος ώστε

$$U^H A U = \text{diag}(\lambda_i) \quad (1.9)$$

- Αν ο A είναι συμμετρικός τότε υπάρχει ορθογώνιος πίνακας U τέτοιος ώστε

$$U^T A U = \text{diag}(\lambda_i) \quad (1.10)$$

Μια κατηγορία πινάκων για την οποία μπορούμε να βρούμε τις ιδιοτιμές αναλυτικά είναι τριδιαγώνιοι πίνακες της μορφής

$$A = \begin{bmatrix} a & b & & & \\ c & a & b & & \\ & c & \ddots & \ddots & \\ & & \ddots & a & b \\ & & & c & a \end{bmatrix} \in \mathbb{R}^{m \times m}$$

Το πρόβλημα ιδιοτιμών $Ax = \lambda x$ είναι ισοδύναμο με τις σχέσεις

$$cx_{j-1} + (a - \lambda)x_j + bx_{j+1} = 0, \quad j = 1, \dots, m, \quad (1.11)$$

όπου έχουμε θέσει $x_0 = x_{m+1} = 0$. Ψάχνουμε λύσεις της μορφής $x_j = \alpha r_1^j + \beta r_2^j$, όπου r_1, r_2 , είναι οι ρίζες του χαρακτηριστικού πολυωνύμου $p(r) = br^2 + (a - \lambda)r + c$ της εξίσωσης διαφορών (1.11).

Χρησιμοποιώντας τη σχέση $x_0 = 0$ έχουμε αμέσως $\alpha + \beta = 0$, επομένως $x_j = \alpha(r_1^j - r_2^j)$. Από τη σχέση $x_{m+1} = 0$ τώρα, έχουμε $r_1^{m+1} = r_2^{m+1}$. Αφού r_1, r_2 είναι ρίζες του χαρακτηριστικού πολυωνύμου της εξίσωσης διαφορών (1.11), ικανοποιούν $r_1 r_2 = c/b$, οπότε

$$\left(\frac{r_1}{r_2}\right)^{m+1} = \left(\frac{r_1^2}{r_1 r_2}\right)^{m+1} = \left(\frac{r_1^2}{c/b}\right)^{m+1} = 1 \Rightarrow r_1^2 = \frac{c}{b} e^{2\pi i \left(\frac{s}{m+1}\right)}, \quad s = 1, \dots, m.$$

Έχουμε λοιπόν

$$r_{1,s} = \sqrt{\frac{c}{b}} e^{\pi i \left(\frac{s}{m+1}\right)}, \quad r_{2,s} = \sqrt{\frac{c}{b}} e^{-\pi i \left(\frac{s}{m+1}\right)}, \quad s = 1, \dots, m.$$

Από τη σχέση $r_{1,s} + r_{2,s} = -\frac{a-\lambda}{b}$ λαμβάνουμε

$$\lambda_s = a + 2\sqrt{bc} \cos\left(\frac{\pi s}{m+1}\right), \quad s = 1, \dots, m.$$

Το αντίστοιχο ιδιοδιάνυσμα είναι

$$x_{s,j} = \alpha(r_{1,s}^j - r_{2,s}^j) = 2i\alpha \left(\frac{a}{b}\right)^{j/2} \sin\left(\frac{\pi j s}{m+1}\right), \quad j = 1, \dots, m.$$

Ένα πολύ σημαντικό αποτέλεσμα από την γραμμική άλγεβρα είναι το θεώρημα ανάλυσης ιδιαζουσών τιμών ενός πίνακα.

Θεώρημα 1.3 (Singular Valued Decomposition, SVD). Έστω A ένας πραγματικός $m \times n$ πίνακας. Τότε, υπάρχουν δύο ορθογώνιοι πίνακες

$$U = [u_1, \dots, u_m] \in \mathbb{R}^{m \times m}, \quad V = [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$$

τέτοιοι ώστε

$$U^T A V = \text{diag}(\sigma_1, \dots, \sigma_p), \quad p = \min\{m, n\}$$

όπου $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$ είναι οι ιδιάζουσες τιμές του A .

Απόδειξη. Έστω $x \in \mathbb{R}^n$ και $y \in \mathbb{R}^m$ τέτοια ώστε $\|x\|_2 = \|y\|_2 = 1$ και $Ax = \sigma y$ με $\sigma = \|A\|_2$. Έστω $V = [x, V_1] \in \mathbb{R}^{n \times n}$ και $U = [y, U_1] \in \mathbb{R}^{m \times m}$ ορθογώνιοι πίνακες. Τότε

$$U^T A V = \begin{bmatrix} y^T A x & y^T A V_1 \\ U_1^T A x & U_1^T A V_1 \end{bmatrix} \equiv A_1 = \begin{bmatrix} \sigma & w^T \\ 0 & B \end{bmatrix}$$

Τώρα,

$$\left\| A_1 \begin{pmatrix} \sigma \\ w \end{pmatrix} \right\|_2^2 \geq (\sigma^2 + w^T w)^2,$$

άρα $\|A_1\|_2^2 \geq \sigma^2 + w^T w$. Όμως $\sigma^2 = \|A\|_2^2 = \|A_1\|_2^2$, άρα πρέπει $w = 0$. Ένα επαγωγικό επιχείρημα και το γεγονός ότι $\sigma_2 = \|B\|_2 \leq \|A_1\|_2 = \|A\|_2 = \sigma$ ολοκληρώνει την απόδειξη του θεωρήματος. \square

Παρατηρήσεις. Από το παραπάνω θεώρημα προκύπτει ότι η 2-νόρμα του πίνακα A ισούται με την μεγαλύτερη ιδιάζουσα τιμή του. Στην απόδειξη της ανάλυσης SVD γίνεται χρήση του γεγονότος ότι αν Q, Z είναι ορθογώνιοι πίνακες κατάλληλης διάστασης, τότε

$$\|QAZ\|_2 = \|A\|_2, \quad (1.12)$$

η απόδειξη του οποίου αφήνεται ως άσκηση.

Στα παρακάτω λήμματα συνδέουμε την ανάλυση ιδιαζουσών ιδιοτιμών με το βαθμό ενός πίνακα.

Λήμμα 1.4. Έστω ότι η ανάλυση ιδιαζουσών ιδιοτιμών ενός πίνακα A είναι όπως στο Θεώρημα 1.3 και

$$\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_p = 0.$$

Τότε

- $\text{rank}(A) = r$
- $N(A) = \text{span}\{v_{r+1}, \dots, v_n\}$, $R(A) = \text{span}\{u_1, \dots, u_r\}$
- $A = \sum_{i=1}^r \sigma_i u_i v_i^T = U_r \Sigma_r V_r^T$, όπου

$$U_r = [u_1, \dots, u_r], \quad V_r = [v_1, \dots, v_r], \quad \Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r)$$

- $\|A\|_F^2 = \sigma_1^2 + \dots + \sigma_p^2, \quad \|A\|_2 = \sigma_1$

Λήμμα 1.5. Έστω ότι η ανάλυση ιδιαζουσών ιδιοτιμών ενός πίνακα $A \in \mathbb{R}^{m \times n}$ είναι όπως στο Θεώρημα 1.3. Αν $k < r = \text{rank}(A)$ και

$$A_k = \sum_{i=1}^k \sigma_i u_i v_i^T,$$

τότε

$$\min_{\text{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1}.$$

1.5 Μετασχηματισμοί Householder

Έστω $0 \neq v \in \mathbb{R}^n$. Ένας $n \times n$ πίνακας P της μορφής

$$P = I - 2vv^T/v^T v \tag{1.13}$$

λέγεται πίνακας Householder. Από τον ορισμό προκύπτει εύκολα ότι ο P είναι συμμετρικός και ορθογώνιος. Οι πίνακες Householder είναι χρήσιμοι γιατί μπορούν όταν εφαρμοστούν σε ένα διάνυσμα να μηδενίσουν συγκεκριμένα στοιχεία του. Για παράδειγμα, ας κατασκευάσουμε ένα πίνακα P της μορφής (1.13) τέτοιον ώστε Px να είναι πολλαπλάσιο του e_1 , όπου x μη μηδενικό διάνυσμα στον \mathbb{R}^n . Έχουμε

$$Px = x - \frac{2v^T x}{v^T v} v$$

οπότε $Px \in \text{span}\{e_1\}$ μόνο αν $v \in \text{span}\{x, e_1\}$. Αν θέσουμε $v = x + \alpha e_1$ τότε

$$Px = \left[1 - 2 \frac{x^T x + \alpha x_1}{x^T x + 2\alpha x_1 + \alpha^2} \right] x - 2\alpha \frac{v^T x}{v^T v} e_1$$

Ο συντελεστής του x είναι μηδέν όταν $\alpha = \pm \|x\|_2$, οπότε η επιλογή $v = x \pm \|x\|_2 e_1$ οδηγεί στη σχέση $Px = \pm \|x\|_2 e_1$. Μπορούμε να διαλέξουμε το πρόσημο της παραμέτρου α ως εξής: αν το διάνυσμα x είναι κοντά σε πολλαπλάσιο του e_1 τότε το διάνυσμα $v = x - \text{sign}(x_1) \|x\|_2 e_1$ θα έχει μικρή νόρμα. Επομένως περιμένουμε μεγάλο σχετικό σφάλμα κατά τον υπολογισμό της ποσότητας $2/v^T v$. Μπορούμε να παρακάμψουμε αυτή τη δυσκολία αν διαλέξουμε το πρόσημο της παραμέτρου α να είναι το ίδιο με αυτό της πρώτης συνιστώσας του x , δηλαδή $v = x + \text{sign}(x_1) \|x\|_2 e_1$.

1.6 Η ανάλυση QR

Έστω $A \in \mathbb{R}^{m \times n}$ με $m \geq n$. Θα αποδείξουμε ότι υπάρχει ορθογώνιος πίνακας Q τέτοιος ώστε $Q^T A = R$ είναι άνω τριγωνικός. Η απόδειξη είναι κατασκευαστική και στηρίζεται σε μετασχηματισμούς Householder. Για παράδειγμα, ας υποθέσουμε ότι $A \in \mathbb{R}^{6 \times 5}$ και ότι έχουμε ήδη υπολογίσει πίνακες Householder P_1, P_2 τέτοιους ώστε

$$P_2 P_1 A = \begin{bmatrix} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & a_{3,3} & x & x \\ 0 & 0 & a_{4,3} & x & x \\ 0 & 0 & a_{5,3} & x & x \\ 0 & 0 & a_{6,3} & x & x \end{bmatrix}$$

Κατόπιν, βρίσκουμε ένα πίνακα Householder $\tilde{P}_3 \in \mathbb{R}^{4 \times 4}$ τέτοιον ώστε

$$\tilde{P}_3 \begin{bmatrix} a_{3,3} \\ a_{4,3} \\ a_{5,3} \\ a_{6,3} \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

και θέτουμε $P_3 = \text{diag}(I_2, \tilde{P}_3)$. Έτσι,

$$P_3 P_2 P_1 A = \begin{bmatrix} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & \alpha & x & x \\ 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & x & x \end{bmatrix}$$

Τελικά, έχουμε $P_n P_{n-1} \cdots P_1 A = R$ άνω τριγωνικό, οπότε θέτοντας $Q = P_1 \cdots P_n$ αποδείξαμε την ανάλυση $A = QR$.

Παρατήρηση Τόσο η ανάλυση QR όσο και η ανάλυση ιδιαζουσών ιδιοτιμών (SVD) επεκτείνονται στην περίπτωση μιγαδικών πινάκων. Θα γράφουμε λοιπόν, $A = U \Sigma V^H$ με U, V ορθομοναδιαίους και Σ διαγώνιο ενώ στην ανάλυση QR ο πίνακας Q θα είναι ορθομοναδιαίος.

1.7 Ορθομοναδιαίοι μετασχηματισμοί

Είδαμε σε προηγούμενη παράγραφο ότι οι πίνακες $X^{-1}AX$ και A , όπου X^{-1} αντιστρέψιμος, έχουν τις ίδιες ιδιοτιμές. Χρησιμοποιώντας μετασχηματισμούς ομοιότητας είναι δυνατόν να μετασχηματίσει τον ένα πίνακα σε αρκετές κανονικές μορφές. Σε αυτή την παράγραφο μελετάμε μετασχηματισμούς βασισμένους σε ορθομοναδιαίους πίνακες.

Λήμμα 1.6. Αν $A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{C}^{p \times p}$ και $X \in \mathbb{C}^{n \times p}$ με $\text{rank}(X) = p$, ικανοποιούν

$$AX = XB \tag{1.14}$$

τότε υπάρχει ορθομοναδιαίος πίνακας $Q \in \mathbb{C}^{n \times n}$ τέτοιος ώστε

$$Q^H A Q = T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \begin{matrix} p \\ n-p \\ p & n-p \end{matrix}$$

όπου $\lambda(T_{22}) = \lambda(A) \cap \lambda(B)$.

Απόδειξη. Έστω

$$X = Q \begin{bmatrix} R \\ 0 \end{bmatrix}, \quad Q \in \mathbb{C}^{n \times n}, R \in \mathbb{C}^{p \times p}$$

η ανάλυση QR του πίνακα X . Αντικαθιστώνοντας στη σχέση (1.14) παίρνουμε

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} = \begin{bmatrix} R \\ 0 \end{bmatrix} B$$

όπου

$$Q^H A Q = T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{matrix} p \\ n-p \\ p & n-p \end{matrix}$$

Χρησιμοποιώντας το γεγονός ότι R είναι αντιστρέψιμος και την εξίσωση $T_{21}R = 0$ συμπεραίνουμε ότι $T_{21} = 0$. Από την εξίσωση $T_{11}R = RB$ έχουμε ότι $\lambda(T_{11}) = \lambda(B)$. Ο ισχυρισμός του λήμματος προκύπτει από το γεγονός ότι $\lambda(A) = \lambda(T) = \lambda(T_{11}) \cup \lambda(T_{22})$. \square

Θεώρημα 1.7 (Ανάλυση Schur). Αν $A \in \mathbb{C}^{n \times n}$ τότε υπάρχει ορθομοναδιαίος πίνακας $Q \in \mathbb{C}^{n \times n}$ τέτοιος ώστε

$$Q^H A Q = T = D + N,$$

όπου $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ και $N \in \mathbb{C}^{n \times n}$ είναι αυστηρά άνω τριγωνικός. Επιπλέον, ο πίνακας Q μπορεί να επιλεγεί έτσι ώστε οι ιδιοτιμές λ_i να εμφανίζονται με οποιαδήποτε σειρά στον πίνακα D .

Απόδειξη. Τι θεώρημα ισχύει τετρωμμένα για $n = 1$. Ας υποθέσουμε ότι ισχύει για πίνακες τάξης $n - 1$ ή μικρότερης. Αν $Ax = \lambda x$ όπου $x \neq 0$ τότε από το Λήμμα 1.6 ξέρουμε ότι υπάρχει ορθομοναδιαίος πίνακας U τέτοιος ώστε

$$U^H A U = \begin{bmatrix} \lambda & w^H \\ 0 & C \end{bmatrix}.$$

Από την επαγωγική υπόθεση τώρα, υπάρχει ορθομοναδιαίος πίνακας \tilde{U} τέτοιος ώστε $\tilde{U}^H C \tilde{U}$ είναι άνω τριγωνικός. Άρα, αν $Q = U \text{diag}(1, \tilde{U})$ τότε $Q^H A Q$ είναι άνω τριγωνικός. \square

Λήμμα 1.8. $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν υπάρχει ορθομοναδιαίος $Q \in \mathbb{C}^{n \times n}$ τέτοιος ώστε $Q^H A Q = \text{diag}(\lambda_1, \dots, \lambda_n)$.

Απόδειξη. Είναι εύκολο να διαπιστώσει κανείς ότι αν ο πίνακας A είναι όμοιος, κάτω από ένα μετασχηματισμό με ορθομοναδιαίους πίνακες, με ένα διαγώνιο πίνακα, τότε είναι κανονικός. Αντίστροφα, αν ο A είναι κανονικός και $Q^H A Q = T$ είναι η ανάλυση Schur, τότε και ο T είναι κανονικός. Η απόδειξη του λήμματος προκύπτει από το γεγονός ότι ένας κανονικός πάνω τριγωνικός πίνακας είναι διαγώνιος. \square

1.8 Πρακτική εφαρμογή της ανάλυσης QR

Έστω ότι $A \in \mathbb{C}^{n \times n}$ και $U_0 \in \mathbb{C}^{n \times n}$ είναι δοσμένος ορθομοναδιαίος πίνακας. Θεωρούμε την επαναληπτική διαδικασία

$$\begin{aligned} T_0 &= U_0^H A U_0 \\ \text{Για } k &= 1, 2, \dots \\ \text{Έστω } T_{k-1} &= U_k R_k \text{ η ανάλυση } QR \text{ του } T_{k-1} \\ \text{Θέτουμε } T_k &= R_k U_k \end{aligned} \tag{1.15}$$

Επειδή

$$T_k = (U_0 U_1 \cdots U_k)^H A (U_0 U_1 \cdots U_k). \quad (1.16)$$

είναι προφανές ότι ο πίνακας T_k είναι όμοιος με τον A . Αυτό που μπορεί να δειχθεί είναι ότι συγκλίνει σε άνω τριγωνική μορφή, δηλαδή στην ανάλυση Schur του πίνακα A . Η επαναληπτική διαδικασία (1.15) λέγεται *επανάληψη QR* και αποτελεί τη βάση πολλών αλγορίθμων για τον υπολογισμό της ανάλυσης Schur. Σε αυτή την παράγραφο θα δείξουμε πως μπορούμε να κάνουμε την επανάληψη (1.15) μια αποτελεσματική μέθοδο για τον υπολογισμό της ανάλυσης Schur. Θα περιοριστούμε στην περίπτωση όπου ο πίνακας A είναι πραγματικός, οπότε θα θεωρήσουμε την επαναληπτική διαδικασία

$$\begin{aligned} H_0 &= U_0^T A U_0 \\ \text{Για } k &= 1, 2, \dots \end{aligned} \quad (1.17)$$

Έστω $H_{k-1} = U_k R_k$ η ανάλυση QR του H_{k-1}

Θέτουμε $H_k = R_k U_k$

Εδώ, $A \in \mathbb{R}^{n \times n}$, κάθε $U_i \in \mathbb{R}^{n \times n}$ είναι ορθογώνιος πίνακας και $R_i \in \mathbb{R}^{n \times n}$ είναι άνω τριγωνικός. Μια επιπλέον δυσκολία με την επαναληπτική διαδικασία (1.17) είναι η περίπτωση που ο πίνακας A έχει μιγαδικές ιδιοτιμές. Σε αυτή την περίπτωση δεν μπορούμε να περιμένουμε ότι οι πίνακες H_k θα συγκλίνουν σε τριγωνική μορφή, ισχύει όμως το ακόλουθο ανάλογο της ανάλυσης Schur

Θεώρημα 1.9. Αν $A \in \mathbb{R}^{n \times n}$ τότε υπάρχει ορθογώνιος πίνακας $Q \in \mathbb{R}^{n \times n}$ τέτοιος ώστε

$$Q^T A Q = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ 0 & R_{22} & \cdots & R_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R_{mm} \end{bmatrix} \quad (1.18)$$

όπου κάθε R_{ii} είναι είτε 1×1 είτε ένας 2×2 πίνακας με συζηγείς μιγαδικές ιδιοτιμές.

Το κόστος κάθε επανάληψης QR είναι $O(n^3)$, το οποίο όμως μπορούμε να το μειώσουμε αν διαλέξουμε τον ορθογώνιο πίνακα U_0 έτσι ώστε $U_0^T A U_0 = H_0 = (h_{ij})$ να είναι άνω Hessenberg, να ισχύει δηλαδή $h_{ij} = 0$, $i > j + 1$. Αυτό μπορεί να γίνει εύκολα με χρήση των πινάκων Givens: αν $1 \leq i < j \leq n$ και $\theta \in \mathbb{R}$ θεωρούμε τον πίνακα

$$J(i, j, \theta) = \begin{bmatrix} 1 & \vdots & \vdots & & \\ \cdots & c & \cdots & s & \cdots \\ & \vdots & & \vdots & \\ \cdots & -s & \cdots & c & \cdots \\ & \vdots & & \vdots & 1 \end{bmatrix} \begin{matrix} \\ i \\ \\ j \\ \end{matrix} \quad (1.19)$$

όπου $c = \cos \theta$, $s = \sin \theta$. Προφανώς ο πίνακας $J(i, j, \theta)$ είναι ορθογώνιος. Αν $x, y \in \mathbb{R}^n$ και $y = J(i, j, \theta)x$, τότε

$$y_i = cx_i + sx_j, \quad y_j = -sx_i + cx_j, \quad y_k = x_k, \quad k \neq i, j.$$

Μπορούμε να επιβάλουμε $y_j = 0$ αν διαλέξουμε

$$c = x_i / (x_i^2 + x_j^2)^{1/2}, \quad s = x_j / (x_i^2 + x_j^2)^{1/2}.$$

Ο αλγόριθμος που ακολουθεί, δεδομένου ενός άνω Hessenberg πίνακα H υπολογίζει την ανάλυση QR $H = UR$ και αποθηκεύει στον ίδιο πίνακα γινόμενο RU :

```

for  $k = 1$  to  $n - 1$  do
  Determine  $c_k = \cos \theta_k$ ,  $s_k = \sin \theta_k$  such that
      
$$\begin{bmatrix} c_k & s_k \\ -s_k & c_k \end{bmatrix} \begin{bmatrix} h_{kk} \\ h_{k+1,k} \end{bmatrix} = \begin{bmatrix} * \\ 0 \end{bmatrix}$$

  for  $j = k$  to  $n$  do
    
$$\begin{bmatrix} h_{kj} \\ h_{k+1,j} \end{bmatrix} = \begin{bmatrix} c_k & s_k \\ -s_k & c_k \end{bmatrix} \begin{bmatrix} h_{kj} \\ h_{k+1,j} \end{bmatrix}$$

  end
end
for  $k = 1$  to  $n - 1$  do
  for  $i = 1$  to  $k + 1$  do
    
$$[h_{ik}, h_{i,k+1}] = [h_{ik}, h_{i,k+1}] \begin{bmatrix} c_k & s_k \\ -s_k & c_k \end{bmatrix}$$

  end
end

```

Παρατηρήστε ότι μια και ο πίνακας

$$U^T = J(n-1, n, \theta_{n-1}) \cdots J(1, 2, \theta_1)$$

είναι κάτω Hessenberg, το γινόμενο RU θα είναι άνω Hessenberg. Επομένως οι πίνακες H_k της επαναληπτικής διαδικασίας (1.17) είναι άνω Hessenberg. Μένει να δείξουμε πώς μπορούμε να υπολογίσουμε ένα ορθογώνιο πίνακα U_0 έτσι ώστε $U_0^T A U_0$ να είναι άνω Hessenberg. Ο πίνακας U_0 θα είναι γινόμενο πινάκων Householder P_1, P_2, \dots, P_{n-2} όπου ο ρόλος του πίνακα P_k είναι να μηδενίσει τα στοιχεία του πίνακα κάτω από την πρώτη υποδιαγώνιο, όπως στο παρακάτω παράδειγμα:

$$P_1 A = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \end{bmatrix}, \quad P_2 P_1 A = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{bmatrix}, \quad P_3 P_2 P_1 A = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * \end{bmatrix}$$

Γενικότερα, μετά από $k - 1$ βήματα θα έχουμε υπολογίσει πίνακες Householder P_1, \dots, P_{k-1} τέτοιους ώστε

$$(P_1 \cdots P_{k-1})^T A (P_1 \cdots P_{k-1}) = \left[\begin{array}{ccc|c} B & & & C \\ - & - & - & - \\ 0 & b & & D \end{array} \right]$$

με $B \in \mathbb{R}^{k \times k}$ άνω Hessenberg. Αν \tilde{P}_k είναι ένας $(n - k) \times (n - k)$ πίνακας Householder τέτοιος ώστε $\tilde{P}_k b$ να είναι πολλαπλάσιο του $e_1^{(n-k)}$ και $P_k = \text{diag}(I_k, \tilde{P}_k)$ τότε

$$(P_1 \cdots P_{k-1} P_k)^T A (P_1 \cdots P_{k-1} P_k) = \left[\begin{array}{ccc|c} B & & & C \tilde{P}_k \\ - & - & - & - \\ 0 & \tilde{P}_k b & & \tilde{P}_k D \tilde{P}_k \end{array} \right]$$

είναι άνω Hessenberg.

1.9 Η μέθοδος των δυνάμεων

Ας υποθέσουμε ότι $X^{-1}AX = \text{diag}(\lambda_1, \dots, \lambda_n)$ με $X = [x_1, \dots, x_n]$ και $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$. Δοθέντος ενός διανύσματος $v^{(0)} \in \mathbb{C}^n$, η μέθοδος των δυνάμεων παράγει την ακολουθία διανυσμάτων $V^{(k)}$ ως εξής:

$$\begin{aligned} \text{Για } k = 1, 2, \dots \\ z^{(k)} &= Av^{(k-1)} \\ \lambda^{(k)} &= z_i^{(k)}, \text{ όπου } |z_i^{(k)}| = \|z^{(k)}\|_\infty \\ v^{(k)} &= z^{(k)} / \lambda^{(k)} \end{aligned} \quad (1.20)$$

Παρατηρήστε ότι αν $v^{(0)} = a_1x_1 + a_2x_2 + \dots + a_nx_n$ με $a_1 \neq 0$ τότε

$$A^k v^{(0)} = a_1 \lambda_1^k \left[x_1 + \sum_{j=2}^n \frac{a_j}{a_1} \left(\frac{\lambda_j}{\lambda_1} \right)^k x_j \right].$$

Επειδή $v^{(k)} \in \text{span}\{A^k v^{(0)}\}$ έχουμε ότι

$$\text{dist}(v^{(k)}, \text{span}\{x_1\}) = O \left[\left(\frac{|\lambda_2|}{|\lambda_1|} \right)^k \right]$$

και γενικότερα,

$$|\lambda_1 - \lambda^{(k)}| = O \left[\left(\frac{|\lambda_2|}{|\lambda_1|} \right)^k \right]$$

Όταν $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$ λέμε ότι λ_1 είναι *κυρίαρχη ιδιοτιμή*. Η μέθοδος των δυνάμεων συγκλίνει αν λ_1 είναι κυρίαρχη ιδιοτιμή και το διάνυσμα $v^{(0)}$ έχει μη μηδενική συνιστώσα προς την κατεύθυνση του κυρίαρχου ιδιοδιανύσματος x_1 . Για παράδειγμα, έστω

$$A = \begin{bmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{bmatrix}, \quad (1.21)$$

και $v^{(0)} = (1, 0, 0)^T$. Έχουμε $\lambda(A) = \{10, 4, 3\}$ και η επαναληπτική διαδικασία (1.20) δίνει, διαδοχικά, τις προσεγγίσεις που φαίνονται στον Πίνακα 1.

1.10 Ιδιοτιμές και ιδιοδιανύσματα με το MATLAB

Το περιβάλλον εργασίας MATLAB υπολογίζει ιδιοτιμές και ιδιοδιανύσματα πινάκων με τη συνάρτηση `eig`. Η συγκεκριμένη συνάρτηση όχι μόνο αναγνωρίζει τον τύπο των στοιχείων του πίνακα εισόδου (πραγματικά ή μιγαδικά) αλλά και ιδιότητές τους, όπως για παράδειγμα η συμμετρία. Οι ιδιοτιμές ενός πίνακα A υπολογίζονται καλώντας τη συνάρτηση `eig`: η εντολή `e = eig(A)` αναθέτει στο διάνυσμα e τις ιδιοτιμές του πίνακα A . Για παράδειγμα,

```
>> A = [2 3; 2 1];
>> e = eig(A);
```

αναθέτουν στο διάνυσμα e τις τιμές 4 και -1 οι οποίες είναι οι ιδιοτιμές του πίνακα A , όπως εύκολα μπορεί να διαπιστώσει κανείς. Η εντολή `[V,D] = eig(A)` υπολογίζει αν n -επί- n διαγώνιο πίνακα D και ένα n -επί- n πίνακα V τέτοιους ώστε $A*V = V*D$. Ο πίνακας D περιέχει τις ιδιοτιμές και οι στήλες του V είναι τα αντίστοιχα ιδιοδιανύσματα. Για τον πίνακα A παραπάνω έχουμε

Πίνακας 1: Προσεγγίσεις της ιδιοτιμής λ_1 που παράγει η μέθοδος (1.20) όταν εφαρμοστεί στον πίνακα (1.21).

k	$\lambda^{(k)}$	$ \lambda_1 - \lambda^{(k)} $	Τάξη
1	-800.0000		
2	13.0375		
3	10.7105	0.7105	
4	10.2033	0.2033	0.2861
5	10.0615	0.0615	0.3025
6	10.0190	0.0190	0.3089
7	10.0060	0.0060	0.3157
8	10.0019	0.0019	0.3166
9	10.0006	0.0006	0.3157
10	10.0002	0.0002	0.3333

```
>> [V,D] = eig(A)
```

```
V =
```

```
    0.8321    -0.7071
    0.5547     0.7071
```

```
D =
```

```
    4     0
    0    -1
```

Το MATLAB κανονικοποιεί τις στήλες του πίνακα ιδιοδιανυσμάτων V έτσι ώστε να έχουν μοναδιαία 2-νόρμα. Για ερμιτιανούς πίνακες, το MATLAB επιστρέφει τις ιδιοτιμές ταξινομημένες σε αύξουσα σειρά και ορθομοναδιαίο πίνακα ιδιοδιανυσμάτων:

```
>> A = [2 -1; -1 1];
```

```
>> [V,D] = eig(A)
```

```
V =
```

```
   -0.5257   -0.8507
   -0.8507    0.5257
```

```
D =
```

```
    0.3820     0
    0    2.6180
```

Οι συντελεστές του χαρακτηριστικού πολυωνύμου ενός πίνακα μπορούν να υπολογιστούν με τη συνάρτηση `poly`:

```
>> poly(A)
```

```
ans =
```

```
    1    -3     1
```

δηλαδή το χαρακτηριστικό πολυώνυμο του πίνακα A είναι το $\lambda^2 - 3\lambda + 1$.

Η συνάρτηση `svd` υπολογίζει την ανάλυση ιδιζουσών ιδιοτιμών του ορίσμάτος της. Συγκεκριμένα, η κλήση `[U, S, V] = svd(X)` επιστρέφει ένα διαγώνιο πίνακα S της ίδιας διάστασης με τον πίνακα X , με μη αρνητικά στοιχεία σε φθίνουσα σειρά, και ορθομοναδιαίους πίνακες U , V τέτοιους ώστε

$$X = USV^H.$$

Για παράδειγμα,

```
>> X = [1 2; 3 4; 5 6];
```

```
>> [U, S, V] = svd(X)
```

U =

```
0.2298    0.8835   -0.4082
0.5247    0.2408    0.8165
0.8196   -0.4019   -0.4082
```

S =

```
9.5255    0
0    0.5143
0    0
```

V =

```
0.6196   -0.7849
0.7849    0.6196
```

Η συνάρτηση `qr` υπολογίζει την ανάλυση QR ενός πίνακα. Η κλήση `[Q,R] = qr(X)` όπου X είναι ένας $m \times n$ πίνακας επιστρέφει ένα $m \times n$ άνω τριγωνικό πίνακα R και ένα ορθομοναδιαίο πίνακα Q έτσι ώστε $X = QR$. Για τον πίνακα X παραπάνω, η κλήση της συνάρτησης `qr` επιστρέφει

Q =

```
0.1690    0.8971   -0.4082
0.5071    0.2760    0.8165
0.8452   -0.3450   -0.4082
```

R =

```
5.9161    7.4374
0    0.8281
0    0
```

Ασκήσεις

- 1.1 Αποδείξτε ότι ο αντίστροφος ενός πίνακα, όταν υπάρχει, είναι μοναδικός.
- 1.2 Αποδείξτε ότι αν $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{m \times p}$ τότε $(AB)^T = B^T A^T$.
- 1.3 Έστω ότι $A \in \mathbb{R}^{n \times n}$ είναι αντιστρέψιμος και $u, v \in \mathbb{R}^n$ είναι τέτοια ώστε $v^T A^{-1} u + 1 \neq 0$. Αποδείξτε τον τύπο των Sherman–Morrison

$$(A + uv^T)^{-1} = A^{-1} - \frac{1}{1 + v^T A^{-1} u} A^{-1} uv^T A^{-1}$$

- 1.4 Δείξτε ότι αν τα $v_1, \dots, v_k \in \mathbb{R}^m$ είναι ορθοκανονικά τότε είναι και γραμμικά ανεξάρτητα.
- 1.5 Έστω A, B τετραγωνικοί πίνακες. Δείξτε ότι οι πίνακες AB και BA έχουν το ίδιο χαρακτηριστικό πολυώνυμο.
- 1.6 Αποδείξτε τις σχέσεις (1.1).
- 1.7 Βρείτε τις ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα

$$A = \begin{bmatrix} 4 & -1 & -1 \\ -1 & 4 & -1 \\ -1 & -1 & 4 \end{bmatrix}$$

- 1.8 (Θεώρημα Gershgorin–Hadamard) Δείξτε ότι

$$\text{sp}(A) \subseteq \bigcup_{i=1}^n \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \right\}.$$

Χρησιμοποιήστε το παραπάνω αποτέλεσμα για να δείξετε ότι αν ο πίνακας A έχει αυστηρά κυριαρχική διαγώνιο, δηλαδή, $a_{ii} > \sum_{j \neq i} |a_{ij}|$, $1 \leq i \leq n$, τότε είναι αντιστρέψιμος.

- 1.9 Δείξτε ότι αν ο πίνακας S είναι αντισυμμετρικός τότε ο πίνακας $I - S$ είναι αντιστρέψιμος και ο πίνακας $(I - S)^{-1}(I + S)$ είναι ορθογώνιος.
- 1.10 Αποδείξτε ότι ο $A \in \mathbb{R}^{n \times n}$ είναι θετικά ορισμένος αν και μόνο αν το συμμετρικό του μέρος, $(A + A^T)/2$, είναι θετικά ορισμένο.
- 1.11 Έστω $A \in \mathbb{C}^{n \times n}$ ερμιτιανός πίνακας. Δείξτε ότι $A = P + iQ$, όπου $P^T = P \in \mathbb{R}^{n \times n}$ και $-Q^T = Q \in \mathbb{R}^{n \times n}$.
- 1.12 Δείξτε ότι οι ιδιοτιμές του πίνακα

$$A = \begin{bmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -29 & -9 & -25 \end{bmatrix}$$

είναι $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$. Βρείτε τον πίνακα X της Πρότασης 1.1 και βεβαιωθείτε ότι ισχύει $X^{-1}AX = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$.

- 1.13 Αποδείξτε τη σχέση (1.12). Αποδείξτε ακόμα ότι η ίδια σχέση ισχύει αν η νόρμα $\|\cdot\|_2$ αντικατασταθεί από τη νόρμα $\|\cdot\|_F$, όπου

$$\|A\|_F^2 = \sum_{i,j=1}^n |a_{ij}|^2$$

- 1.14 Χρησιμοποιείστε την ανάλυση ιδιαιζουσών ιδιοτιμών για να δείξετε ότι $\text{rank}(A) = \text{rank}(A^T)$.

- 1.15 Έστω ότι ο πίνακας $A \in \mathbb{R}^{m \times n}$ έχει βαθμό n . Δείξτε ότι $\|A(A^T A)^{-1} A^T\|_2 = 1$.

- 1.16 Αν $x, y \in \mathbb{R}^n$, δείξτε ότι $\det(I_n + xy^T) = 1 + x^T y$.

- 1.17 Έστω ότι $A \in \mathbb{R}^{n \times n}$ είναι συμμετρικός πίνακας και $x \in \mathbb{R}^n$. Δείξτε ότι η ελάχιστη τιμή της συνάρτησης $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(\lambda) = \|(A - \lambda I)x\|_2$ λαμβάνεται για

$$\lambda = r(x) = \frac{x^T A x}{x^T x}.$$

Η παράσταση $r(x)$ ονομάζεται *πηλίκο Rayleigh* του x . Προφανώς, αν x είναι ένα κατά προσέγγιση ιδιοδιάνυσμα τότε $r(x)$ είναι η αντίστοιχη, κατά προσέγγιση, ιδιοτιμή. Αντιστρόφως, αν λ είναι προσέγγιση μιας ιδιοτιμής του πίνακα A τότε η λύση του συστήματος $(A - \lambda I)x = b$, $b \in \mathbb{R}^n$ θα είναι καλή προσέγγιση του αντίστοιχου ιδιοδιανύσματος. Θεωρούμε λοιπόν την ακόλουθη επαναληπτική διαδικασία: Για δεδομένο $x_0 \in \mathbb{R}^n$ με $\|x_0\|_2 = 1$ υπολογίζουμε x_k , $k = 1, 2, \dots$ ως εξής:

$$\mu_k = r(x_k), \quad (A - \mu_k I)z_{k+1} = x_k, \quad x_{k+1} = z_{k+1}/\|z_{k+1}\|.$$

Δείξτε ότι για οποιοδήποτε k ισχύει,

$$(A + E_k)z_{k+1} = \mu_k z_{k+1}, \quad E_k = -x_k z_{k+1}^T / \|z_{k+1}\|_2^2.$$

Από την παραπάνω σχέση προκύπτει ότι (δείτε το βιβλίο του Wilkinson, *The Algebraic Eigenvalue problem*, σελ. 100–102) ότι $|\mu_k - \lambda| \leq 1/\|z_{k+1}\|_2$, για κάποια ιδιοτιμή λ του πίνακα A . Μπορεί να αποδειχθεί ότι η παραπάνω επαληπτική διαδικασία συγκλίνει για οποιοδήποτε αρχικό διάνυσμα x_0 και ότι η ταχύτητα σύγκλισης είναι κυβική. Μια ιδέα της απόδειξης μπορεί κανείς να πάρει στην περίπτωση $n = 2$. Θεωρούμε, χωρίς βλάβη της γενικότητας ότι $A = \text{diag}(\lambda_1, \lambda_2)$ με $\lambda_1 > \lambda_2$. Γράφουμε

$$x_k = \begin{bmatrix} c_k \\ s_k \end{bmatrix}, \quad c_k^2 + s_k^2 = 1.$$

Δείξτε ότι

$$\mu_k = \lambda_1 c_k^2 + \lambda_2 s_k^2, \quad z_{k+1} = \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} c_k/s_k^2 \\ -s_k/c_k^2 \end{bmatrix},$$

και συμπεράνετε ότι

$$c_{k+1} = c_k^3 / (c_k^6 + s_k^6)^{1/2}, \quad s_{k+1} = -s_k^3 / (c_k^6 + s_k^6)^{1/2}.$$

Η παραπάνω σχέση λέει ότι τα x_k συγκλίνουν κυβικά είτε στό $\text{span}\{e_1\}$ είτε στο $\text{span}\{e_2\}$, εφόσον $|c_k| \neq |s_k|$. Γράψτε ένα πρόγραμμα στο MATLAB το οποίο υλοποιεί την παραπάνω διαδικασία. Εφαρμόστε την στον πίνακα Pascal τάξης 6 και επιβεβαιώστε αριθμητικά την ταχύτητα σύγκλισης της μεθόδου. Γράψτε `help pascal` στη γραμμή εντολών του MATLAB για να δείτε ποιός είναι ο πίνακας Pascal και υπολογίστε τις ιδιοτιμές του χρησιμοποιώντας τη συνάρτηση `eig`. Δοκιμάστε με το αρχικό διάνυσμα $x_0 = (1, 1, 1, 1, 1, 1)^T / \sqrt{6}$ και επαναλάβετε με ένα άλλο, τυχαίο διάνυσμα x_0 με 2-νόρμα ίση με ένα. Σε ποιά από τις ιδιοτιμές του πίνακα συγκλίνει η μέθοδός σας;

2 Φασματικές μέθοδοι

Στο κεφάλαιο αυτό παρουσιάζουμε συνοπτικά φασματικές μεθόδους για την αριθμητική επίλυση διαφορικών εξισώσεων. Έμφαση δίνεται τόσο στην θεωρία των μεθόδων αλλά και στην πρακτική εφαρμογή τους για την αριθμητική λύση προβλημάτων συνοριακών και αρχικών τιμών. Η κινητήρια ιδέα είναι η σύνδεση των πινάκων παραγωγίσης με την παράγωγο μιας παρεμβάλλουσας συνάρτησης και τον υπολογισμό της μέσω του μετασχηματισμού Fourier, στην περίπτωση περιοδικών πλεγμάτων. Στην τελευταία παράγραφο αυτού του κεφαλαίου εισάγεται η παρεμβάλλουσα στα σημεία του Chebyshev και εξηγείται πως μπορεί να υπολογιστεί η διακριτή παράγωγος μέσω του μετασχηματισμού Fourier.

2.1 Πίνακες παραγωγίσης

Έστω ένα ομοιόμορφο πλέγμα σημείων $\{x_1, \dots, x_N\}$ με $x_{j+1} - x_j = h$ για κάθε j , και το αντίστοιχο σύνολο τιμών $\{u_1, \dots, u_N\}$ μιας συνάρτησης u . Έστω w_j μια προσέγγιση της $u'(x_j)$. Θα μπορούσαμε, για παράδειγμα, να πάρουμε

$$w_j = \frac{u_{j+1} - u_{j-1}}{2h} \quad (2.1)$$

Αν θεωρήσουμε ότι το πρόβλημα είναι περιοδικό και πάρουμε $u_0 = u_N, u_{N+1} = u_1$ τότε ο υπολογισμός της διακριτής παραγωγού έχει τη μορφή γινομένου πίνακα επί διάνυσμα:

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{N-1} \\ w_N \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & \cdots & & -1/2 \\ -1/2 & 0 & \ddots & & \\ & & \ddots & & \\ & & & \ddots & 0 & 1/2 \\ 1/2 & & & -1/2 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{N-1} \\ u_N \end{bmatrix} \quad (2.2)$$

Ο πίνακας που εμφανίζεται στην παραπάνω σχέση ονομάζεται *πίνακας παραγωγίσης*. Παρατηρήστε ότι είναι πίνακας Toeplitz, δηλαδή έχει σταθερές τιμές κατά μήκος των διαγωνίων του. Είναι επίσης ένας *κυκλικός πίνακας*, δηλαδή η τιμή του στοιχείου στη γραμμή i , στήλη j εξαρτάται μόνο από το $(i - j) \bmod N$.

Ένας διαφορετικός τρόπος να υπολογίσουμε τη διακριτή παράγωγο είναι μέσω του πολυώνυμου παρεμβολής:

Για $j = 1, 2, \dots, N$:

- Έστω p_j το μοναδικό πολυώνυμο βαθμού ≤ 2 με $p_j(x_{j-1}) = u_{j-1}$, $p_j(x_j) = u_j$ και $p_j(x_{j+1}) = u_{j+1}$.
- Θέτουμε $w_j = p'_j(x_j)$.

Είναι εύκολο να δειχθεί ότι για κάθε σταθερό j το πολυώνυμο p_j δίνεται από τον τύπο

$$p_j(x) = u_{j-1}a_{-1}(x) + u_j a_0(x) + u_{j+1}a_1(x),$$

όπου

$$a_{-1}(x) = \frac{1}{2h^2}(x-x_j)(x-x_{j+1}), \quad a_0(x) = -\frac{1}{h^2}(x-x_{j-1})(x-x_{j+1}), \quad a_1(x) = \frac{1}{2h^2}(x-x_{j-1})(x-x_j).$$

Επιπλέον, $p'_j(x_j) = w_j$, όπου τα w_j δίνονται από τη σχέση (2.2). Το πλεονέκτημα αυτής της οπτικής για τον υπολογισμό της διακριτής παραγωγού είναι ότι μπορεί να γενικευτεί. Για παράδειγμα, η παρακάτω

“συνταγή” παράγει προσεγγίσεις στις τιμές $u'(x_j)$ οι οποίες συγκλίνουν με τάξη $O(h^4)$ καθώς $h \rightarrow 0$:

Για $j = 1, 2, \dots, N$:

- Έστω p_j το μοναδικό πολυώνυμο βαθμού ≤ 4 με $p(x_{j\pm 2}) = u_{j\pm 2}$, $p(x_{j\pm 1}) = u_{j\pm 1}$ και $p_j(x_j) = u_j$.
- Θέτουμε $w_j = p'(x_j)$.

Η ιδέα πίσω από τις φασματικές μεθόδους είναι να πάρουμε το όριο αυτής της διαδικασίας όταν $h \rightarrow 0$ και να χρησιμοποιήσουμε, τουλάχιστον στη θεωρία, ένα άπειρο πίνακα παραγωγίσης. Στην πράξη, βασιζόμαστε στον ακόλουθο τρόπο προσέγγισης της παραγωγού:

- Έστω p μια συνάρτηση που ικανοποιεί $p(x_j) = u_j$ για όλα τα j .
- Θέτουμε $w_j = p'(x_j)$.

Στις επόμενες παραγράφους θα διαλέξουμε τη συνάρτηση p κατάλληλα και θα αναπτύξουμε γρήγορες τεχνικές για τον υπολογισμό των διακριτών παραγώγων. Ξεκινάμε με την περίπτωση ενός άπειρου πλέγματος.

2.2 Ο ημιδιακριτός μετασχηματισμός Fourier

Συμβολίζουμε με $h\mathbb{Z}$ το σύνολο των σημείων $x_j = jh$ για $j \in \mathbb{Z}$. Ο μετασχηματισμός Fourier μιας συνάρτησης $u(x)$, $x \in \mathbb{R}$ είναι η συνάρτηση $\hat{u}(k)$ που ορίζεται από τη σχέση

$$\hat{u}(k) = \int_{-\infty}^{\infty} e^{-ikx} u(x) dx, \quad k \in \mathbb{R}. \quad (2.3)$$

Αντίστροφα, μπορούμε να ανακατασκευάσουμε την u από την \hat{u} μέσω του αντίστροφου μετασχηματισμού Fourier

$$u(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} \hat{u}(k) dk, \quad x \in \mathbb{R}. \quad (2.4)$$

Θα ονομάζουμε την μεταβλητή x φυσική μεταβλητή και την μεταβλητή k μεταβλητή Fourier ή κυματαριθμό. Μας ενδιαφέρει να περιορίσουμε τη φυσική μεταβλητή στο σύνολο $h\mathbb{Z}$ οπότε ο κυματαριθμός θα περιοριστεί σε ένα διάστημα μήκους $2\pi/h$ και μια βολική επιλογή για το διάστημα αυτό είναι το διάστημα $[-\pi/h, \pi/h]$. Ο λόγος του περιορισμού του κυματαριθμού σε ένα διάστημα μήκους $2\pi/h$ είναι το φαινόμενο aliasing (ψευδωνυμίας?): Τα μιγαδικά εκθετικά $f(x) = \exp(ik_1x)$ και $g(x) = \exp(ik_2x)$ δεν συμπίπτουν στο \mathbb{R} αν $k_1 \neq k_2$, ενώ, αν περιοριστούν στο σύνολο $h\mathbb{Z}$ τότε παίρνουν τις ίδιες τιμές στα σημεία x_j , αν και μόνο αν, $k_1 - k_2$ είναι πολλαπλάσιο του $2\pi/h$. Αρκεί λοιπόν να υπολογίσει κανείς κυματαριθμούς σε ένα διάστημα αυτού του μήκους.

Αν v είναι μια συνάρτηση ορισμένη στο $h\mathbb{Z}$ με τιμές v_j στα σημεία x_j τότε ο ημιδιακριτός μετασχηματισμός Fourier της v ορίζεται ως

$$\hat{v}(k) = h \sum_{j=-\infty}^{\infty} e^{-ikx_j} v_j, \quad k \in [-\pi/h, \pi/h], \quad (2.5)$$

και ο αντίστροφος ημιδιακριτός μετασχηματισμός Fourier ως

$$v_j = \frac{1}{2\pi} \int_{-\pi/h}^{\pi/h} e^{ikx_j} \hat{v}(k) dk, \quad j \in \mathbb{Z}. \quad (2.6)$$

Η τελευταία σχέση μας επιτρέπει να ορίσουμε μια παρεμβάουσα της συνάρτησης v ως

$$p(x) = \frac{1}{2\pi} \int_{-\pi/h}^{\pi/h} e^{ikx} \hat{v}(k) dk, \quad x \in \mathbb{R}. \quad (2.7)$$

Η συνάρτηση p που ορίζεται από την σχέση (2.7) είναι μια αναλυτική συνάρτηση της μεταβλητής x και, από κατασκευή, ικανοποιεί $p(x_j) = v_j$ για κάθε $j \in \mathbb{Z}$. Επιπλέον, ο μετασχηματισμός Fourier της p είναι

$$\hat{p}(k) = \begin{cases} \hat{v}(k), & k \in [-\pi/h, \pi/h] \\ 0, & \text{διαφορετικά.} \end{cases} \quad (2.8)$$

Επομένως \hat{p} έχει συμπαγή φορέα στο διάστημα $[-\pi/h, \pi/h]$. Λέμε ότι p είναι η παρεμβάουσα περιορισμένου φάσματος της v (band limited interpolant). Το αποτέλεσμα αυτό είναι γνωστό ως το θεώρημα δειγματοληψείας (sampling theorem) των Shannon, Nyquist και Whittaker.

Με τη χρήση της παρεμβάουσας που ορίζει η σχέση (2.7) μπορούμε τώρα να προσεγγίσουμε την παράγωγο μιας συνάρτησης v ορισμένη στο πλέγμα $h\mathbb{Z}$ ως εξής:

- Δεδομένης της συνάρτησης v , υπολογίζουμε την παρεμβάουσα (2.7)
- Θέτουμε $w_j = p'(x_j)$

Μπορούμε επίσης να υπολογίσουμε την διακριτή παράγωγο χρησιμοποιώντας τον ημιδιακριτό μετασχηματισμό Fourier: αν παραγωγίσουμε την (2.4) ως προς x έχουμε την πολύ σημαντική σχέση

$$\widehat{u'}(k) = ik \hat{u}(k) \quad (2.9)$$

Επομένως μπορούμε να υπολογίσουμε τη διακριτή παράγωγο ως εξής:

- Δεδομένης της v , υπολογίζουμε τον ημιδιακριτό μετασχηματισμό Fourier από την (2.5)
- Ορίζουμε $\hat{w}(k) = ik \hat{v}(k)$
- Υπολογίζουμε το w από το \hat{w} μέσω της σχέσης (2.6).

Με τη βοήθεια της παραπάνω διαδικασίας μπορούμε να υπολογίσουμε τον πίνακα παραγωγίσης: συμβολίζουμε με δ τη συνάρτηση δέλτα του Kronecker

$$\delta_j = \begin{cases} 1, & j = 0, \\ 0, & j \neq 0. \end{cases}$$

Ο μετασχηματισμός Fourier της δ είναι, σύμφωνα με τη σχέση (2.5), σταθερή: $\hat{\delta}(k) = h$ για κάθε $k \in [-\pi/h, \pi/h]$. Από τη σχέση (2.7) έχουμε ότι η παρεμβάουσα περιορισμένου φάσματος της συνάρτησης δ είναι

$$p(x) = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} e^{ikx} dk = \frac{\sin(\pi x/h)}{\pi x/h}$$

με τιμή 1 στο $x = 0$. Η συνάρτηση στο δεξί μέλος της παραπάνω σχέσης ονομάζεται συνάρτηση sinc και συμβολίζεται με $S_h(x)$. Από την γραμμικότητα του μετασχηματισμού Fourier και το γεγονός ότι για μια συνάρτηση v ορισμένη στο πλέγμα $h\mathbb{Z}$ ισχύει

$$v_j = \sum_{m=-\infty}^{\infty} v_m \delta_{j-m},$$

έχουμε ότι

$$p(x) = \sum_{m=-\infty}^{\infty} v_m S_h(x - x_m), \quad (2.10)$$

και παραγωγίζοντας ως προς x

$$w_j = p'(x_j) = \sum_{m=-\infty}^{\infty} v_m S'_h(x_j - x_m). \quad (2.11)$$

Τώρα $S'_h(x_j) = 0$ για $j = 0$ ενώ $S'_h(x_j) = (-1)^j/jh$ αν $j \neq 0$, επομένως από την προηγούμενη σχέση προκύπτει ότι η στήλη του πίνακα παραγωγίσισης που αντιστοιχεί στην τιμή $m = 0$ είναι

$$h^{-1}(\dots, 1/3, -1/2, 1, 0, -1, 1/2, -1/3, \dots)^T.$$

Οι υπόλοιπες στήλες του πίνακα παραγωγίσισης προκύπτουν από κατάλληλη μετατόπιση της στήλης που αντιστοιχεί στην τιμή $m = 0$.

2.3 Ο διακριτός μετασχηματισμός Fourier

Εξετάζουμε τώρα το πρόβλημα του υπολογισμού της διακριτής παραγώγου σε ένα φραγμένο περιοδικό πλέγμα. Υποθέτουμε ότι το περιοδικό μας πλέγμα είναι υποσύνολο του $[0, 2\pi]$ και $0 < x_1 < x_2 < \dots < x_{N/2} < \dots < x_{N-1} < x_N = 2\pi$ με N άρτιο. Η απόσταση δύο διαδοχικών σημείων στο πλέγμα είναι $h = 2\pi/N$, με άλλα λόγια,

$$\frac{\pi}{h} = \frac{N}{2} \quad (2.12)$$

Αν τώρα περιορίσουμε τη φυσική μεταβλητή x στα σημεία του περιοδικού πλέγματος τότε οι κυματαριθμοί περιορίζονται στις τιμές $-N/2 + 1, -N/2 + 2, \dots, N/2$ μια και τα κύματα e^{ikx} έχουν περίοδο 2π μόνο αν ο κυματαριθμός k είναι ακέραιος. Ορίζουμε λοιπόν το διακριτό μετασχηματισμό Fourier από τη σχέση

$$\hat{v}_k = h \sum_{j=1}^N e^{-ikx_j} v_j, \quad k = -\frac{N}{2} + 1, \dots, \frac{N}{2}, \quad (2.13)$$

και τον αντίστροφο διακριτό μετασχηματισμό Fourier ως

$$v_j = \frac{1}{2\pi} \sum_{k=-N/2+1}^{N/2} e^{ikx_j} \hat{v}_k, \quad j = 1, \dots, N. \quad (2.14)$$

Για τον υπολογισμό της παραγώγου μιας συνάρτησης v χρησιμοποιούμε την ιδέα της προηγούμενης παραγράφου. Πρώτα ορίζουμε την παρεμβάλλουσα περιορισμένου φάσματος αντικαθιστώντας το x_j στη σχέση (2.14) με x . Σημειώνουμε όμως ότι με αυτή την αντικατάσταση, η παράγωγος του όρου $e^{iNx/2}$ είναι τώρα $(iN/2)e^{iNx/2}$ η οποία δεν είναι μηδέν στα σημεία του πλέγματος! Μπορούμε να διορθώσουμε αυτό το πρόβλημα θέτοντας $\hat{v}_{-N/2} = \hat{v}_{N/2}$ και να ορίσουμε την παρεμβάλλουσα περιορισμένου φάσματος από τη σχέση

$$p(x) = \frac{1}{2\pi} \sum'_{k=-N/2}^{N/2} e^{ikx} \hat{v}_k, \quad x \in [0, 2\pi], \quad (2.15)$$

όπου ο τόνος στο σύμβολο του αθροίσματος δηλώνει ότι ο πρώτος και ο τελευταίος όρος πολλαπλασιάζονται με $1/2$. Η συνάρτηση $p(x)$ που ορίζεται από τη σχέση (2.15) είναι ένα τριγωνομετρικό πολυώνυμο βαθμού το πολύ $N/2$ και μπορεί επομένως να γραφεί ως γραμμικός συνδιασμός των συναρτήσεων $1, \sin(x), \cos(x), \sin(2x), \cos(2x), \dots, \sin(Nx/2), \cos(Nx/2)$.

Συμβολίζουμε και πάλι με δ την περιοδική συνάρτηση που ορίζεται από τις σχέση

$$\delta_j = \begin{cases} 1, & j \equiv 0 \pmod{N} \\ 0, & \text{διαφορετικά} \end{cases}$$

Από τη σχέση (2.13) έχουμε ότι $\hat{\delta}_k = h$ για όλα τα k επομένως από τη σχέση (2.15) έχουμε ότι η παρεμβά-
λουσα περιορισμένου φάσματος $S_N(x)$ της περιοδικής συνάρτησης δ είναι

$$S_N(x) = \frac{\sin(Nx/2)}{(2\pi/h) \tan(x/2)}. \quad (2.16)$$

Παραγωγίζοντας την προηγούμενη σχέση και υπολογίζοντας στο $x = x_j$ έχουμε ότι

$$S'_N(x_j) = \begin{cases} 0, & j \equiv 0 \pmod{N} \\ \frac{1}{2}(-1)^j \cot(jh/2), & \text{διαφορετικά,} \end{cases} \quad (2.17)$$

επομένως ο πίνακας παραγωγίσης είναι

$$D_N = \begin{bmatrix} 0 & & & & -\frac{1}{2} \cot \frac{1h}{2} \\ -\frac{1}{2} \cot \frac{1h}{2} & \ddots & & & \frac{1}{2} \cot \frac{2h}{2} \\ \frac{1}{2} \cot \frac{2h}{2} & & \ddots & & -\frac{1}{2} \cot \frac{3h}{2} \\ -\frac{1}{2} \cot \frac{3h}{2} & & & \ddots & \vdots \\ \vdots & & & \ddots & \frac{1}{2} \cot \frac{1h}{2} \\ \frac{1}{2} \cot \frac{1h}{2} & & & & 0 \end{bmatrix} \quad (2.18)$$

Υψηλότερης τάξης παράγωγοι μπορούν να υπολογιστούν με διαδοχικές παραγωγίσεις της σχέσης (2.16). Όπως και στην περίπτωση του άπειρου πλέγματος μπορούμε να υπολογίσουμε την παράγωγο με τη βο-
ήθεια του διακριτού μετασχηματισμού Fourier:

- Δεδομένης της συνάρτησης v υπολογίζουμε το διακριτό μετασχηματισμό Fourier \hat{v} από τη σχέση (2.13)
- Ορίζουμε $\hat{w}_k = ik\hat{v}_k$ και $\hat{w}_{N/2} = 0$
- Υπολογίζουμε το w από το \hat{w} μέσω της (2.14)

Γενικότερα, για την παράγωγο τάξης ν :

- Δεδομένης της συνάρτησης v υπολογίζουμε το διακριτό μετασχηματισμό Fourier \hat{v} από τη σχέση (2.13)
- Ορίζουμε $\hat{w}_k = (ik)^\nu \hat{v}_k$ και $\hat{w}_{N/2} = 0$ αν το ν είναι περιττός αριθμός
- Υπολογίζουμε το w από το \hat{w} μέσω της (2.14)

Ο υπολογισμός του διακριτού μετασχηματισμού Fourier μπορεί να γίνει μέσω του αλγορίθμου Fast Fourier Transform (FFT) που ανακαλύφθηκε το 1965 από τους Cooley και Tukey. Αν το N είναι γινόμενο πολλών μικρών πρώτων αριθμών, ο αλγόριθμος FFT υπολογίζει τον διακριτό μετασχηματισμό Fourier (2.13) κάνον-
τας $O(N \log N)$ πράξεις. Ο αντίστροφος διακριτός μετασχηματισμός Fourier (2.14) μπορεί, προφανώς, να υπολογιστεί με το ίδιο κόστος πράξεων. Το MATLAB υλοποιεί τον αλγόριθμο FFT και τον αντίστροφο του με τις συναρτήσεις `fft` και `ifft`, αντίστοιχα. Σημειώνουμε όμως ότι οι κυματαριθμοί αποθηκεύονται με τη σειρά $0, 1, \dots, \frac{N}{2}, -\frac{N}{2} + 1, -\frac{N}{2} + 2, \dots, -1$. Επομένως, αν το διάνυσμα v μήκους N περιέχει τις τιμές μιας συνάρτησης στα σημεία του περιοδικού πλέγματος η διαδικασία για τον υπολογισμό των πα-
ραγώγων w είναι:

```
v_hat = fft(v);
w_hat = 1i*[0:N/2-1 0 -N/2+1:-1]' .* v_hat;
w = ifft(w_hat);
```

2.4 Πολυωνυμική παρεμβολή σε μη περιοδικά πλέγματα

Στις προηγούμενες παραγράφους είδαμε πως να προσεγγίσουμε την παράγωγο μιας συνάρτησης σε περιοδικά πλέγματα. Εξετάζουμε τώρα την περίπτωση που το πρόβλημά μας τίθεται σε ένα φραγμένο, μη περιοδικό χωρίο, και για ευκολία θα θεωρήσουμε το διάστημα $[-1, 1]$. Για να χρησιμοποιήσουμε τις μεθόδους που αναπτύξαμε προηγούμενα θα μπορούσαμε να επεκτείνουμε περιοδικά τη συνάρτηση της οποίας την παράγωγο προσεγγίζουμε αλλά η επέκταση θα είναι, εν γένει, μη ομαλή. Οι ασυνέχειες όμως, καταστρέφουν την ακρίβεια των φασματικών μεθόδων (φαινόμενο Gibbs).

Αντί να κάνουμε χρήση τριγωνομετρικών πολυωνύμων, θα χρησιμοποιήσουμε αλγεβρικά πολυώνυμα $p(x) = a_0 + a_1x + \dots + a_Nx^N$. Γνωρίζουμε όμως ότι όταν ομαλές συναρτήσεις παραμβάλονται σε $N + 1$ ισαπέχοντα σημεία από πολυώνυμα, το σφάλμα της παρεμβολής όχι μόνο πάει στο μηδέν όσο το N αυξάνει αλλά μπορεί να μεγαλώσει με ρυθμό ανάλογο του 2^N . Το γεγονός αυτό, γνωστό και ως φαινόμενο Runge, μας επιβάλλει να θεωρήσουμε παρεμβολή με πολυώνυμα σε μη ομοιόμορφα πλέγματα. Ένα σύνολο σημείων για το οποίο μπορεί να αποδειχθεί (δείτε το επόμενο κεφάλαιο) μειώνεται εκθετικά όταν $N \rightarrow \infty$ είναι τα σημεία του Chebyshev

$$x_j = \cos(j\pi/N), \quad j = 0, 1, \dots, N. \quad (2.19)$$

Όπως και σε προηγούμενες παραγράφους, θα προσεγγίσουμε την παράγωγο με το σχήμα:

- Έστω p το μοναδικό πολυώνυμο βαθμού $\leq N$ με $p(x_j) = v_j, j = 0, \dots, N$.
- Θέτουμε $w_j = p'(x_j)$

Η διαδικασία αυτή είναι γραμμική, επομένως μπορεί να αναπαρασταθεί ως τη δράση ενός $(N + 1) \times (N + 1)$ πίνακα πάνω σε ένα διάνυσμα:

$$w = D_N v.$$

Για παράδειγμα, αν $N = 1$ τότε

$$p(x) = \frac{1}{2}(1+x)v_0 + \frac{1}{2}(1-x)v_1, \quad p'(x) = \frac{1}{2}v_0 - \frac{1}{2}v_1,$$

άρα

$$D_1 = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Στην περίπτωση $N = 2$ έχουμε $x_0 = 1, x_1 = 0, x_2 = -1$ οπότε

$$p(x) = \frac{1}{2}x(1+x)v_0 + (1+x)(1-x)v_1 + \frac{1}{2}x(x-1)v_2$$

και

$$p'(x) = (x + \frac{1}{2})v_0 - 2xv_1 + (x - \frac{1}{2})v_2.$$

Ο πίνακας παραγωγίσης D_2 είναι τώρα

$$D_2 = \begin{bmatrix} \frac{3}{2} & -2 & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 2 & -\frac{3}{2} \end{bmatrix}.$$

Το επόμενο θεώρημα δίνει τα στοιχεία του πίνακα παραγωγίσης D_N :

Θεώρημα 2.1. Έστω ότι οι γραμμές και οι στήλες του $(N+1) \times (N+1)$ πίνακα παραγωγίσις D_N αριθμούνται από το 0 ως το N . Τότε

$$(D_N)_{00} = \frac{2N^2+1}{6}, \quad (D_N)_{NN} = -\frac{2N^2+1}{6}, \quad (2.20)$$

$$(D_N)_{jj} = \frac{-x_j}{2(1-x_j^2)}, \quad j = 1, \dots, N-1, \quad (2.21)$$

$$(D_N)_{ij} = \frac{c_i}{c_j} \frac{(-1)^{i+j}}{(x_i-x_j)}, \quad i \neq j, \quad i, j = 0, \dots, N, \quad (2.22)$$

όπου

$$c_i = \begin{cases} 2, & i = 0 \text{ ή } N \\ 1, & \text{διαφορετικά.} \end{cases}$$

Επομένως,

$$D_N = \begin{bmatrix} \frac{2N^2+1}{6} & \dots & 2\frac{(-1)^j}{1-x_j} & \dots & \frac{1}{2}(-1)^N \\ \vdots & \ddots & & \frac{(-1)^{i+j}}{x_i-x_j} & \vdots \\ -\frac{1}{2}\frac{(-1)^i}{1-x_i} & & \frac{-x_j}{2(1-x_j^2)} & & \frac{1}{2}\frac{(-1)^{N+i}}{1+x_i} \\ \vdots & \frac{(-1)^{i+j}}{x_i-x_j} & & \ddots & \vdots \\ -\frac{1}{2}(-1)^N & & -2\frac{(-1)^{N+j}}{1+x_j} & & -\frac{2N^2+1}{6} \end{bmatrix}$$

Η j στήλη του πίνακα D_N περιέχει τις τιμές της παραγώγου του πολυωνύμου βαθμού N , το οποίο παρεμβάλλει τη συνάρτηση δέλτα με τιμή ένα στο σημείο x_j , στα σημεία $\{x_i\}$. Οι παρακάτω εντολές της MATLAB κατασκευάζουν τον πίνακα D_N :

```
function [D,x] = cheb(N)
    if N == 0, D = 0; x = 1; return, end
    x = cos(pi*(0:N)/N)';
    c = [2; ones(N-1,1); 2].*(-1).^(0:N)';
    X = repmat(x, 1, N+1);
    dX = X - X';
    D = (c*(1./c)') ./ (dX+(eye(N+1)));
    D = D - diag(sum(D'));
```

2.5 Προβλήματα συνοριακών τιμών

Σε αυτή την παράγραφο δείχνουμε πως ο πίνακας παραγωγίσις D_N μπορεί να χρησιμοποιηθεί στη λύση προβλημάτων συνοριακών τιμών για μερικές ή συνήθεις διαφορικές εξισώσεις.

Σαν πρώτο παράδειγμα θεωρούμε το πρόβλημα

$$u_{xx} = e^{4x}, \quad -1 < x < 1, \quad u(\pm 1) = 0, \quad (2.23)$$

του οποίου η ακριβής λύση είναι $u(x) = (e^{4x} - x \sinh(4) - \cosh(4))/16$. Προσεγγίζουμε την παράγωγο u_{xx} με τον πίνακα D_N^2 , το τετράγωνο του πίνακα D_N . Τις ομογενείς συνοριακές συνθήκες Dirichlet μπορούμε να τις χειριστούμε ως εξής: έστω x_1, \dots, x_{N-1} τα εσωτερικά σημεία Chebyshev και $v = (v_1, \dots, v_{N-1})^T$ το αντίστοιχο διάνυσμα των τιμών της άγνωστης συνάρτησης u . Θεωρούμε το μοναδικό πολυώνυμο p

βαθμού $\leq N$ το οποίο ικανοποιεί $p(\pm 1) = 0$ και $p(x_j) = v_j$, $1 \leq j \leq N - 1$ και θέτουμε $w_j = p''(x_j)$. Ισοδύναμα, στο γραμμικό σύστημα

$$\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N-1} \\ w_N \end{bmatrix} = D_N^2 \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \\ v_N \end{bmatrix}, \quad (2.24)$$

να θέσουμε $v_0 = v_N = 0$ και να αγνοήσουμε τα w_0 και w_N . Στη γλώσσα του MATLAB, θα μπορούσαμε να γράψουμε:

```
[D, x] = cheb(N);
D2 = D^2;
D2 = D2(2:N, 2:N);
```

Παρατηρήστε ότι η τελευταία εντολή αναθέτει στην μεταβλητή D2 τον πίνακα παραγωγίσις D_N^2 με εξαίρεση την πρώτη και τελευταία γραμμή και στήλη.

Στην περίπτωση ενός μη γραμμικού προβλήματος συνοριακών τιμών, όπως για παράδειγμα το

$$u_{xx} = e^u, \quad -1 < x < 1, \quad u(\pm 1) = 0, \quad (2.25)$$

δεν μπορούμε απλά να αντιστρέψουμε τον πίνακα D_N^2 αλλά μπορούμε να βρούμε τη λύση με την επαναληπτική διαδικασία

$$D_N^2 v^{(i+1)} = e^{v^{(i)}}, \quad i = 1, 2, \dots, \quad (2.26)$$

όπου $(e^{v^{(i)}})_j = e^{v_j^{(i)}}$, $j = 0, 1, \dots, N$. Η επαναληπτική διαδικασία σταματάει όταν η διαφορά δύο διαδοχικών επαναλήψεων στη νόρμα απείρου είναι μικρότερη από ένα προκαθορισμένο, από το χρήστη, όριο.

Για προβλήματα συνοριακών τιμών στις δύο διαστάσεις, όπως για παράδειγμα το πρόβλημα

$$u_{xx} + u_{yy} = 10 \sin(8x(y - 1)), \quad -1 < x, y < 1, \quad u = 0 \text{ στο σύνορο}, \quad (2.27)$$

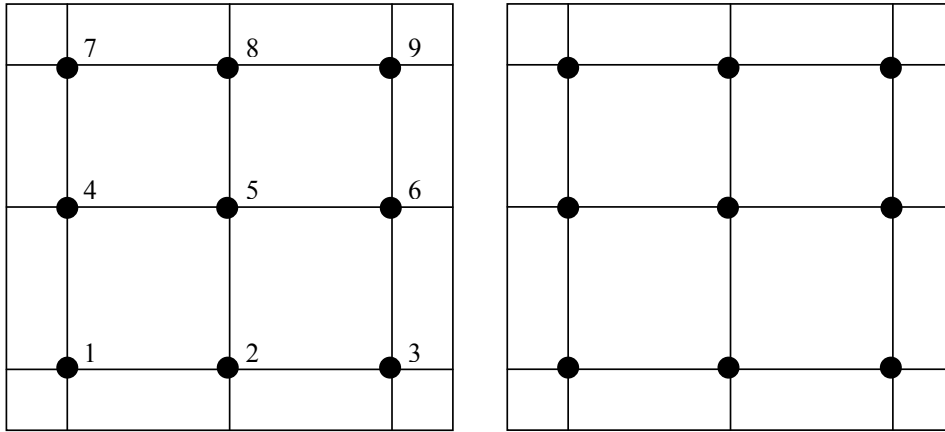
μπορούμε να προχωρήσουμε ως εξής: θεωρούμε το λεγόμενο *τανυστικό πλέγμα* (x_i, x_j) πάνω από το χωρίο $[-1, 1]^2$, όπου $\{x_i\}$ είναι τα σημεία του Chebyshev. Αριθμούμε τα σημεία του πλέγματος στο εσωτερικό του χωρίου με τη λεγόμενη λεξικογραφική διάταξη (από κάτω προς τα επάνω και από αριστερά προς στα δεξιά), όπως στο παράδειγμα για $N = 4$, στο παρακάτω σχήμα. Ο ευκολότερος τρόπος να ορίσουμε τη διακριτή Λαπλασιανή στα σημεία του πλέγματος είναι μέσω του τανυστικού γινομένου ή γινομένου Kronecker δύο πινάκων: αν $A \in \mathbb{R}^{p \times q}$ και $B \in \mathbb{R}^{r \times s}$ τότε το τανυστικό γινόμενο $A \otimes B$ είναι ένας $p \times q$ block-πίνακας διάστασης $pr \times qs$ όπου το ij -block είναι $a_{ij}B$. Για παράδειγμα,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b & 2a & 2b \\ c & d & 2c & 2d \\ 3a & 3b & 4a & 4b \\ 3c & 3d & 4c & 4d \end{bmatrix}$$

Έτσι, αν I είναι ο ταυτοτικός πίνακας διάστασης $(N - 1) \times (N - 1)$ και \tilde{D}_N^2 είναι ο πίνακας που προκύπτει από το τετράγωνο του $(N + 1) \times (N + 1)$ πίνακα παραγωγίσις D_N αν αφαιρέσουμε την πρώτη και την τελευταία γραμμή και στήλη, η διακριτή Λαπλασιανή L_N ορίζεται ως

$$L_N = I \otimes \tilde{D}_N^2 + \tilde{D}_N^2 \otimes I. \quad (2.28)$$

Στο MATLAB, η συνάρτηση `kron` υπολογίζει το τανυστικό γινόμενο δύο πινάκων και οι παρακάτω εντολές κατασκευάζουν την διακριτή Λαπλασιανή:



```
[D, x] = cheb(N);
D2 = D^2; D2 = D2(2:N, 2:N);
I = eye(N-1);
L = kron(I, D2) + kron(D2, I);
```

2.5.1 Μη ομογενείς συνοριακές συνθήκες

Εξετάζουμε τώρα προβλήματα συνοριακών τιμών με μη ομογενείς συνθήκες Dirichlet ή/και συνθήκες Neumann. Θεωρούμε κατ' αρχήν το πρόβλημα συνοριακών τιμών

$$u_{xx} = e^{4x}, \quad -1 < x < 1, \quad u(-1) = 0, u(1) = 1. \quad (2.29)$$

Μπορούμε να υπολογίσουμε τη λύση του παραπάνω προβλήματος ως εξής: συνάρτηση $q(x) = (x+1)/2$ ικανοποιεί $q(-1) = 0$ και $q(1) = 1$. Αν ορίσουμε $v = u - q$ τότε εύκολα προκύπτει ότι η συνάρτηση v ικανοποιεί

$$v_{xx} = e^{4x}, \quad -1 < x < 1, \quad v(-1) = v(1) = 0, \quad (2.30)$$

δηλαδή ένα πρόβλημα συνοριακών τιμών με ομογενείς συνοριακές συνθήκες του οποίου τη λύση μπορούμε να υπολογίσουμε. Η λύση, λοιπόν, αρχικού προβλήματος είναι η $v + q$:

```
[D, x] = cheb(N);
D2 = D^2; D2 = D2(2:N, 2:N);
f = exp(4*x(2:N));
u = D2 \ f;
u = [0; u; 0] + (x+1)/2;
```

Στην περίπτωση ενός προβλήματος συνοριακών τιμών με συνθήκες Neumann, όπως για παράδειγμα το

$$u_{xx} = e^{4x}, \quad -1 < x < 1, \quad u_x(-1) = u(1) = 0, \quad (2.31)$$

μπορούμε να προχωρήσουμε ως εξής: για να επιβάλλουμε την συνοριακή συνθήκη $u(1) = 0$ θα αφαιρέσουμε από τον πίνακα παραγωγισής D_N^2 την πρώτη γραμμή και στήλη. Για να επιβάλλουμε τώρα τη συνθήκη Neumann $u_x(-1) = 0$ θα αντικαταστήσουμε την τελευταία γραμμή του πίνακα D_N^2 από την τελευταία γραμμή του πίνακα D_N :

```

[D, x] = cheb(N);
D2 = D^2;
D2(N+1, :) = D(N+1, :);
D2 = D2(2:N+1, 2:N+1);
f = exp(4*x(2:N));
u = D2\ f;
u = [0; u];

```

2.6 Υπολογισμός της παραγώγου με τη βοήθεια του ταχύ μετασχηματισμού Fourier

Όπως και στην περίπτωση των περιοδικών πλεγμάτων μπορούμε να δείξουμε ότι η διακριτή παράγωγος μπορεί να υπολογιστεί με τη βοήθεια του ταχύ μετασχηματισμού Fourier. Έστω $z \in \mathbb{C}$ με $|z| = 1$ και $x = \operatorname{Re} z = \cos \theta$, όπου θ το όρισμα του z . Για κάθε $x \in [-1, 1]$ υπάρχουν δύο συζυγείς μιγαδικοί αριθμοί τέτοιοι ώστε

$$x = \operatorname{Re} z = \frac{1}{2}(z + z^{-1}) = \cos \theta \in [-1, 1]. \quad (2.32)$$

Το n -οστό πολυώνυμο Chebyshev T_n ορίζεται από την σχέση

$$T_n(x) = \operatorname{Re} z^n = \frac{1}{2}(z^n + z^{-n}) = \cos n\theta. \quad (2.33)$$

Χρησιμοποιώντας την αναδρομική σχέση $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$ είναι εύκολο να δειχθεί ότι το $T_n(x)$ είναι όντως πολυώνυμο βαθμού n και ότι ο συντελεστής του μεγιστοβάθμιου όρου είναι 2^{n-1} . Αν p είναι πολυώνυμο βαθμού N μπορούμε να γράψουμε

$$p(x) = \sum_{n=0}^N a_n T_n(x), \quad x \in [-1, 1], \quad (2.34)$$

και να ορίσουμε, δείτε τη σχέση (2.32),

$$\mathbf{p}(z) = \frac{1}{2} \sum_{n=0}^N a_n (z^n + z^{-n}), \quad |z| = 1, \quad (2.35)$$

και

$$P(\theta) = \sum_{n=0}^N a_n \cos n\theta, \quad \theta \in \mathbb{R}. \quad (2.36)$$

Προφανώς οι συναρτήσεις που ορίζονται από τις σχέσεις (2.34)–(2.36) είναι ισοδύναμες με την έννοια ότι $p(x) = \mathbf{p}(z) = P(\theta)$ όταν τα x, z και θ συνδέονται μέσω της σχέσης (2.32). Όμοια, για μια αυθαίρετη συνάρτηση $f(x)$ ορισμένη για $x \in [-1, 1]$ μπορούμε να ορίσουμε $\mathbf{f}(z)$ και $F(\theta)$ από τις σχέσεις

$$\mathbf{f}(z) = f\left(\frac{z + z^{-1}}{2}\right), \quad F(\theta) = f(\cos \theta).$$

Έχουμε λοιπόν τις ισοδυναμίες

$$\begin{aligned}
P(\theta) \text{ παρεμβάλει την } F(\theta) \text{ στα σημεία } \theta_j = j\pi/N \\
\Updownarrow \\
\mathbf{p}(z) \text{ παρεμβάλει την } \mathbf{f}(z) \text{ στα σημεία } z_j = e^{i\theta_j} \\
\Updownarrow \\
p(x) \text{ παρεμβάλει την } f(x) \text{ στα σημεία } x_j = \cos \theta_j = \operatorname{Re} z_j
\end{aligned}$$

Η κεντρική ιδέα που μας επιτρέπει τώρα τον υπολογισμό της παραγώγου μέσω του ταχύ μετασχηματισμού Fourier είναι ότι μπορούμε να παραγωγίσουμε μια παρεμβάλουσα q της f βρίσκοντας πρώτα το τριγωνομετρικό πολυώνυμο Q που παρεμβάλλει την αντίστοιχη συνάρτηση F , η οποία είναι περιοδική και άρτια (δείτε τη σχέση (2.36)), παραγωγίζοντας στο χώρο Fourier και μετασχηματίζοντας, μέσω της (2.32) πίσω στη φυσική μεταβλητή x . Συνοψίζουμε τη διαδικασία αυτή στον παρακάτω αλγόριθμο:

- Έστω v_0, \dots, v_N οι τιμές μιας συνάρτησης στα σημεία $x_0 = 1, \dots, x_N = -1$. Επεκτείνουμε τα δεδομένα αυτά σε ένα διάνυσμα V μήκους $2N$ μέσω της σχέσης $V_{2N-j} = v_j, j = 1, 2, \dots, N-1$.
- Χρησιμοποιώντας το μετασχηματισμό Fourier υπολογίζουμε

$$\hat{V}_k = \frac{\pi}{N} \sum_{j=1}^{2N} e^{-ik\theta_j} V_j, \quad k = -N+1, \dots, N.$$

- Ορίζουμε $W_k = ik\hat{V}_k$ και $\hat{W}_N = 0$.
- Υπολογίζουμε την παράγωγο του τριγωνομετρικού πολυωνύμου Q μέσω του αντίστροφου μετασχηματισμού Fourier

$$W_j = \frac{1}{2\pi} \sum_{k=-N+1}^N e^{ik\theta_j} \hat{W}_k, \quad j = 1, \dots, 2N.$$

- Υπολογίζουμε την παράγωγο του πολυωνύμου παρεμβολής q στα σημεία Chebyshev από τις σχέσεις

$$w_j = -\frac{W_j}{\sqrt{1-x_j^2}}, \quad j = 1, \dots, N-1, \quad (2.37)$$

και

$$w_0 = \frac{1}{2\pi} \sum_{n=0}^{N-1} n^2 \hat{v}_n, \quad w_N = \frac{1}{2\pi} \sum_{n=0}^{N-1} (-1)^{n+1} n^2 \hat{v}_n \quad (2.38)$$

όπου ο τόνος δηλώνει ότι ο πρώτος και ο τελευταίος όρος του αθροίσματος πολλαπλασιάζονται επί $1/2$.

Η σχέση (2.37) δικαιολογείται από το γεγονός ότι

$$P(\theta) = \frac{1}{2\pi} \sum_{k=-N+1}^N e^{ik\theta} \hat{V}_k = \sum_{n=0}^N a_n \cos n\theta$$

οπότε έχουμε ότι το πολυώνυμο που παίρνει τις τιμές $\{v_j\}$ στα σημεία $\{x_j\}$ είναι το $p(x) = P(\theta)$ για $x = \cos \theta$, και η παράγωγός του είναι

$$q'(x) = \frac{Q'(\theta)}{dx/d\theta} = \frac{\sum_{n=0}^N n a_n \sin n\theta}{\sin \theta} = \frac{\sum_{n=0}^N n a_n \sin n\theta}{\sqrt{1-x^2}}.$$

Οι σχέσεις (2.38) προκύπτουν από τον υπολογισμό της $q'(x)$ στα σημεία $x = \pm 1$ με χρήση του κανόνα l'Hôpital. Η συνάρτηση `chebfft` που φαίνεται παρακάτω υλοποιεί τη διαδικασία αυτή στο MATLAB:

```

function w = chebfft(v)
    N = length(v)-1; if N == 0, w = 0; return, end
    x = cos((0:N)'*pi/N);
    ii = 0:N-1;
    v = v(:); V = [v; flipud(v(2:N))];
    U = real(fft(V));
    W = real(ifft(1i*[ii 0 1-N:-1]' .* U));
    w = zeros(N+1,1);
    w(2:N) = -W(2:N) ./ sqrt(1-x(2:N).^2);
    w(1) = sum(ii' .* 2 .* U(ii+1))/N + 0.5*N*U(N+1);
    w(N+1) = sum((-1).^(ii+1)' .* ii .* 2 .* U(ii+1))/N + ...
        0.5*(-1)^(N+1)*N*U(N+1);

```

Ασκήσεις

2.1 Έστω $\hat{u} = \mathcal{F}\{u\}$, $\hat{v} = \mathcal{F}\{v\}$ ο μετασχηματισμός Fourier, της συνάρτησης u , αντίστοιχα της συνάρτησης v . Αποδείξτε τις παρακάτω ιδιότητες του μετασχηματισμού Fourier.

(α') $\mathcal{F}\{u + v\}(k) = \hat{u}(k) + \hat{v}(k)$, $\mathcal{F}\{cu\} = c\hat{u}(k)$.

(β') Αν $x_0 \in \mathbb{R}$ τότε $\mathcal{F}\{u(x + x_0)\}(k) = e^{ikx_0}\hat{u}(k)$.

(γ') $\mathcal{F}\{\bar{u}\}(k) = \overline{\hat{u}(-k)}$

(δ') $\mathcal{F}\{u_x\}(k) = ik\hat{u}(k)$.

2.2 Έστω $u \in L^2(\mathbb{R})$ και \hat{u} ο μετασχηματισμός Fourier της. Αποδείξτε ότι αν u είναι πραγματική συνάρτηση τότε \hat{u} είναι ερμιτιανή (Σημείωση: f είναι ερμιτιανή αν $f(-x) = \overline{f(x)}$).

2.3 Οι ρουτινές `fft` και `ifft` του MATLAB χρησιμοποιούν μιγαδική αριθμητική ακόμα και αν το διάστημα εισόδου είναι πραγματικό. Εξηγήστε πως μπορούμε να υπολογίσουμε την παράγωγο δύο διακριτών πραγματικών συναρτήσεων χρησιμοποιώντας μία μόνο κλίση της `fft` και της `ifft`. Αλλάξτε κατάλληλα το πρόγραμμα `spdiff2.m` από την ιστοσελίδα του μαθήματος για να υπολογίσετε την παράγωγο των συναρτήσεων $f_1(x) = \exp(\sin(x))$ και $f_2(x) = \exp(\sin^2(x))$.

2.4 Χρησιμοποιήστε τις μεθόδους της παραγράφου 2.5 για να λύσετε το πρόβλημα συνοριακών τιμών

$$u_{xx} + 4u_x + e^x u = \sin(8x), \quad -1 < x < 1, \quad u(\pm 1) = 0,$$

2.5 Υποθέστε ότι $p(x) = \sum_{n=0}^N a_n T_n(x) = \sum_{n=0}^N c_n x^n$ και θεωρήστε τα $(N + 1)$ -διανύσματα $a = (a_0, \dots, a_N)^T$ και $c = (c_0, \dots, c_N)^T$. Έστω A ο $(N + 1) \times (N + 1)$ πίνακας τέτοιος ώστε $c = Aa$. Υπολογίστε τον A στην περίπτωση $N = 3$. Γράψτε μια συνάρτηση στο MATLAB η οποία υπολογίζει τον πίνακα A με όρισμα εισόδου το N . Χρησιμοποιήστε την για να εκφράσετε το πολυώνυμο $1 - 2x + 3x^2 + 2x^3 + x^4 - x^5$ συναρτήσει των πολυωνύμων Chebyshev βαθμού το πολύ πέντε.

2.6 Λύστε την εξίσωση Burgers $u_t + (u^2)_x = \epsilon u_{xx}$ όπου $\epsilon > 0$ είναι σταθερά. Χρησιμοποιήστε μεθόδους Fourier στο διάστημα $[-\pi, \pi]$ και κάντε χρήση ολοκληρωτικού παράγοντα όπως στο πρόγραμμα `kdv.m` στα εργαστήρια του μαθήματος. Χρησιμοποιήστε την αρχική συνθήκη $\sin^2 x$ στο διάστημα $[-\pi, 0]$ και μηδέν στο $[0, \pi]$. Ζωγραφίστε τη λύση στους χρόνους $0, 1/2, 1, \dots, 3$ για $N = 64, 128, 256$ με αρκετά μικρό βήμα σε κάθε περίπτωση έτσι ώστε να μην παρατηρούνται ταλαντώσεις της λύσης. Πάρτε $\epsilon = 0.25$ σε όλες τις περιπτώσεις και χρησιμοποιήστε την άμεση μέθοδο Runge-Kutta για την διακριτοποίηση στο χρόνο, όπως στο πρόγραμμα `kdv.m`.