ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών
ΙΔΡΥΘΕΝ ΤΟ 1837

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ελάχιστα Τετράγωνα: Αλγόριθμοι και Εφαρμογές

Νικόλαος Ι. Πουλίδης

Επιβλέπουσα: Μαριλένα Μητρούλη, Καθηγήτρια

ΑΘΗΝΑ
ΑΠΡΙΛΙΟΣ 2023

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ελάχιστα Τετράγωνα: Αλγόριθμοι και Εφαρμογές

Νικόλαος Ι. Πουλίδης
Α.Μ.: 192603

ΕΠΙΒΛΕΠΟΥΣΑ: Μαριλένα Μητρούλη, Καθηγήτρια

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Μαριλένα Μητρούλη, Καθηγήτρια

Ευαγγελία Κόττα-Αθανασιάδου, Αναπληρώτρια Καθηγήτρια

Μιχαήλ Δρακόπουλος, Επίκουρος Καθηγητής

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία είναι μια αλγοριθμική εισαγωγή στη μέθοδο των ελαχίστων τετραγώνων και τις εφαρμογές της. Αρχικά παρουσιάζονται οι βασικές έννοιες πλαισιωμένες από αποδείξεις χρήσιμων προτάσεων και εισάγεται ο απαραίτητος συμβολισμός που θα χρειαστεί στη συνέχεια. Ακολουθεί ο ορισμός του γραμμικού προβλήματος ελαχίστων τετραγώνων(ΓΠΕΤ) και η απόδειξη ύπαρξης λύσης. Για την επίλυση του ΓΠΕΤ χρησιμοποιείται η παραγοντοποίηση QR ενός πίνακα μέσω μετασχηματισμών Householder και περιγράφεται αναλυτικά ο αλγόριθμος υλοποίησης της. Στη συνέχεια ορίζονται οι έννοιες του υπερκαθορισμένου ΓΠΕΤ και του υποκαθορισμένου ΓΠΕΤ ενώ παρουσιάζονται τρόποι επίλυσης τους. Τέλος, γίνεται εφαρμογή του ΓΠΕΤ στα πολυώνυμα παρεμβολής και στην παρεμβολή με splines.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Εφαρμοσμένα Μαθηματικά
ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: νόρμα, πίνακας, γραμμικό πρόβλημα ελαχίστων τετραγώνων, μετασχηματισμοί Householder, παραγοντοποίηση QR

ABSTRACT

This thesis is an algorithmic introduction to the least-squares method and its applications. Firstly, the basic concepts are presented along with proofs of useful propositions and the necessary notation that will be needed later is introduced. Following is the definition of the linear least-squares problem(LLSP) and the proof of the existence of a solution. QR factorization of a matrix via Householder transformations is used to solve the LLSP and its implementation algorithm is described in detail. Next, the concepts of the overdetermined LLSP and the underdetermined LLSP are defined while ways of solving them are presented. Finally, the LLSP is applied to polynomial and spline interpolation.

SUBJECT AREA: Applied Mathematics
KEYWORDS: norm, matrix, linear least-squares problem, Householder transformations, QR factorization

ΕΥΧΑΡΙΣΤΙΕΣ

Για την διεκπεραίωση της παρούσας διπλωματικής εργασίας θα ήθελα να ευχαριστήσω την επιβλέπουσα καθηγήτρια Μαριλένα Μητρούλη για την βοήθεια και καθοδηγησή της καθώς επίσης την αναπληρώτρια καθηγήτρια Ευαγγελία Κόττα-Αθανασιάδου και τον επίκουρο καθηγητή Μιχαήλ Δρακόπουλο που δέχθηκαν να συμμετάσχουν ως μέλη της εξεταστικής επιτροπής.

Περιεχόμενα

ΠΡΟΛΟΓΟΣ

Πολλά προβλήματα της επιστήμης και της τεχνολογίας περιγράφονται από γραμμικά μαθηματικά πρότυπα(μοντέλα) στα οποία το πλήθος των εξισώσεων είναι διαφορετικό από το πλήθος των αγνώστων(ζητούμενων παραμέτρων). Τέτοια προβλήματα εμφανίζονται σε στατιστικές και γεωμετρικές εφαρμογές καθώς και σε τεχνολογικές εφαρμογές όπως η επεξεργασία σήματος και εικόνας. Για την επίλυση των προβλημάτων αυτών επιδιώκεται η ελαχιστοποίηση κάποιας ποσότητας η οποία υπολογίζεται με νόρμα που παράγεται από εσωτερικό γινόμενο και είναι γνωστά ως προβλήματα ελαχίστων τετραγώνων. Στην παρούσα διπλωματική εργασία παρουσιάζεται το γραμμικό πρόβλημα ελαχίστων τετραγώνων, αλγόριθμοι επιλυσής του και εφαρμογές πλαισιωμένα από το απαραίτητο θεωρητικό υπόβαθρο.

1 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

1.1 Διανυσματικός Χώρος

Ένα σύνολο V λέγεται διανυσματικός χώρος αν ισχύουν τα ακόλουθα:

1. Για κάθε a, b όπου $a \in V, b \in V$ ορίζεται το άθροισμα τους που είναι στοιχείο του V και συμβολίζεται ως $a + b$.

2. Για κάθε λ, a όπου $\lambda \in \mathbb{R}, a \in V$ ορίζεται το γινομενό τους που είναι στοιχείο του V και συμβολίζεται ως $\lambda \cdot a$ ή λa.

3. Για το άθροισμα και το γινόμενο ισχύουν τα ακόλουθα:

   i. a+b=b+a για κάθε a, b όπου $a \in V, b \in V$.
   ii. a+(b+c)=(a+b)+c για κάθε a, b, c όπου $a \in V, b \in V, c \in V$.
   iii. Υπάρχει στοιχείο του V που συμβολίζεται ως 0_V τέτοιο ώστε $a + 0_V = a$ για κάθε a όπου $a \in V$.
   iv. Για κάθε a όπου $a \in V$ υπάρχει στοιχείο του V που συμβολίζεται ως $-a$ τέτοιο ώστε $a + (-a) = 0_V$.
   v. $1a = a$.
   vi. $\lambda(\mu a) = (\lambda\mu)a$ για κάθε λ, μ όπου $\lambda \in \mathbb{R}, \mu \in \mathbb{R}$.
   vii. $(\lambda + \mu)a = \lambda a + \mu a$ για κάθε λ, μ όπου $\lambda \in \mathbb{R}, \mu \in \mathbb{R}$.
   viii. $\lambda(a + b) = \lambda a + \lambda b$ για κάθε λ όπου $\lambda \in \mathbb{R}$.

Πρόταση 1.1.1. Αν V είναι διανυσματικός χώρος τότε ισχύουν τα ακόλουθα:

1. Το 0_V είναι το μοναδικό στοιχείο του V για το οποίο ισχύει ότι $a + 0_V = a$ για κάθε a όπου $a \in V$.

2. Για κάθε a όπου $a \in V$ το $-a$ είναι το μοναδικό στοιχείο του V για το οποίο ισχύει ότι $a + (-a) = 0_V$.

Απόδειξη.

1. Αν $b \in V$ και $a + b = a$ για κάθε a όπου $a \in V$ τότε $b = b + 0_V = 0_V + b = 0_V$.

2. Αν $b \in V$ και $a + b = 0_V$ τότε $b = b + 0_V = b + (a + (-a)) = (b + a) + (-a) = (a + b) + (-a) = 0_V + (-a) = -a$.

□

Πρόταση 1.1.2. Αν V είναι διανυσματικός χώρος και $a \in V$ τότε ισχύουν τα ακόλουθα:

1. $0a = 0_V$.

2. $(-\lambda)a = -\lambda a$ για κάθε λ όπου $\lambda \in \mathbb{R}$.

3. $-(-a) = a$.

Απόδειξη.

1. $0a = 0a+0_V = 0a+(a+(-a)) = (0a+a)+(-a) = (0a+1a)+(-a) = (0+1)a+(-a) = 1a + (-a) = a + (-a) = 0_V$.
2. $\lambda a + (-\lambda a) = 0_V$ και $\lambda a + (-\lambda)a = (\lambda + (-\lambda))a = 0a = 0_V$ άρα από την Πρόταση 1.1.1(2) προκύπτει ότι $(-\lambda)a = -\lambda a$.
3. $-a + (-(-a)) = 0_V$ και $-a + a = a + (-a) = 0_V$ άρα από την Πρόταση 1.1.1(2) προκύπτει ότι $-(-a) = a$.

□

Πρόταση 1.1.3. Αν V είναι διανυσματικός χώρος και $a_i \in V$ για κάθε i όπου $i \in \{1, \ldots, n\}$ τότε $-\sum_{i=1}^{n} a_i = \sum_{i=1}^{n}(-a_i)$ για κάθε n όπου $n \in \mathbb{N}$.

Απόδειξη. Αν $n = 1$ τότε $-\sum_{i=1}^{n} a_i = -\sum_{i=1}^{1} a_i = -a_1 = \sum_{i=1}^{1}(-a_i) = \sum_{i=1}^{n}(-a_i)$.
$(a_1 + a_2) + (-(a_1 + a_2)) = 0_V$ και $(a_1 + a_2) + (-a_1 + (-a_2)) = ((a_1 + a_2) + (-a_1)) + (-a_2) = ((a_2+a_1)+(-a_1))+(-a_2) = (a_2+(a_1+(-a_1)))+(-a_2) = (a_2+0_V)+(-a_2) = a_2+(-a_2) = 0_V$ άρα από την Πρόταση 1.1.1(2) προκύπτει ότι $-(a_1 + a_2) = -a_1 + (-a_2)$.
Υποθέτουμε ότι αν $n \in \{1, \ldots, m\}$ τότε $-\sum_{i=1}^{n} a_i = \sum_{i=1}^{n}(-a_i)$.
Αν $n = m+1$ τότε $-\sum_{i=1}^{m+1} a_i = -(\sum_{i=1}^{m} a_i+a_{m+1}) = -\sum_{i=1}^{m} a_i+(-a_{m+1}) = \sum_{i=1}^{m}(-a_i)+(-a_{m+1}) = \sum_{i=1}^{m+1}(-a_i)$ άρα από την αρχή της μαθηματικής επαγωγής προκύπτει ότι $-\sum_{i=1}^{n} a_i = \sum_{i=1}^{n}(-a_i)$ για κάθε n όπου $n \in \mathbb{N}$. □

Τα στοιχεία ενός διανυσματικού χώρου λέγονται διανύσματα. Αν a, b είναι διανύσματα τότε θέτουμε $a - b = a + (-b)$ και το διάνυσμα $a - b$ λέγεται η διαφορά του a από το b.

Πρόταση 1.1.4. Αν V είναι διανυσματικός χώρος και $a \in V, b \in V$ τότε η εξίσωση $x+a = b$ έχει ως μοναδική λύση το διάνυσμα $x = b - a$.

Απόδειξη. Το διάνυσμα $b - a$ είναι λύση της εξίσωσης $x + a = b$ διότι $(b - a) + a = (b + (-a)) + a = b + (-a + a) = b + 0_V = b$.
Αν c είναι διάνυσμα τέτοιο ώστε $c+a = b$ τότε $c = c+0_V = c+(a+(-a)) = (c+a)+(-a) = b + (-a) = b - a$. □

Πρόταση 1.1.5. Αν V είναι διανυσματικός χώρος και $a \in V, b \in V$ τέτοια ώστε $a+c = b+c$ τότε $a = b$.

Απόδειξη. $a = a + 0_V = a + (c + (-c)) = (a + c) + (-c) = (b + c) + (-c) = b + (c + (-c)) = b + 0_V = b$. □

Έστω V διανυσματικός χώρος και $\lambda_i \in \mathbb{R}, a_i \in V$ για κάθε i όπου $i \in \{1, \ldots, n\}$. Το διάνυσμα $\sum_{i=1}^{n}(\lambda_i a_i)$ λέγεται γραμμικός συνδυασμός των διανυσμάτων $a_1, \ldots, a_n$. Αν $\sum_{i=1}^{n}(\lambda_i a_i) = 0_V$ συνεπάγεται $\lambda_i = 0$ για κάθε i όπου $i \in \{1, \ldots, n\}$ τότε τα διανύσματα $a_1, \ldots, a_n$ λέγονται γραμμικώς ανεξάρτητα, διαφορετικά λέγονται γραμμικώς εξαρτημένα. Υπόχωρος του V λέγεται κάθε υποσυνολό του που είναι διανυσματικός χώρος. Από τον ορισμό του διανυσματικού χώρου προκύπτει ότι ένα υποσύνολο U ενός διανυσματικού χώρου V είναι διανυσματικός χώρος αν $a \in U$ και $b \in U$ συνεπάγεται $\lambda a + \mu b \in U$ για κάθε λ, μ όπου $\lambda \in \mathbb{R}, \mu \in \mathbb{R}$. Το σύνολο όλων των γραμμικών συνδυασμών των διανυσμάτων $a_1, \ldots, a_n$ είναι υπόχωρος του V και λέγεται διανυσματική θήκη του συνόλου $\{a_1, \ldots, a_n\}$ ενώ συμβολίζεται ως $\text{span}(\{a_1, \ldots, a_n\})$ δηλαδή:

$$\text{span}(\{a_1, \ldots, a_n\}) = \{\sum_{i=1}^{n}(\lambda_i a_i) : \lambda_i \in \mathbb{R} \text{ για κάθε } i \text{ όπου } i \in \{1, \ldots, n\}\}$$

Πρόταση 1.1.6. Αν $a_1, \ldots, a_n$ είναι γραμμικώς ανεξάρτητα διανύσματα του χώρου V και $a \in \text{span}(\{a_1, \ldots, a_n\})$ τότε το a είναι μοναδικός γραμμικός συνδυασμός των $a_1, \ldots, a_n$.

Απόδειξη. Αν $\lambda_i \in \mathbb{R}, \mu_i \in \mathbb{R}$ για κάθε i όπου $i \in \{1, \ldots, n\}$ και $a = \sum_{i=1}^{n}(\lambda_i a_i) = \sum_{i=1}^{n}(\mu_i a_i)$ τότε $\sum_{i=1}^{n}((\lambda_i - \mu_i)a_i) = 0_V$ άρα $\lambda_i - \mu_i = 0$ για κάθε i όπου $i \in \{1, \ldots, n\}$ δηλαδή $\lambda_i = \mu_i$ για κάθε i όπου $i \in \{1, \ldots, n\}$ επομένως το a είναι μοναδικός γραμμικός συνδυασμός των $a_1, \ldots, a_n$. □

Ένας διανυσματικός χώρος V λέγεται πεπερασμένης διάστασης αν υπάρχει φυσικός αριθμός n που λέγεται διάσταση του V και συμβολίζεται ως $\dim(V)$ τέτοιος ώστε ο V περιέχει n γραμμικώς ανεξάρτητα διανύσματα και οποιαδήποτε $n+1$ διανύσματα του V είναι γραμμικώς εξαρτημένα. Κάθε σύνολο n γραμμικώς ανεξάρτητων διανυσμάτων ενός διανυσματικού χώρου V διάστασης n λέγεται βάση του V.

1.2 Εσωτερικό Γινόμενο - Νόρμα

Ένας διανυσματικός χώρος V λέγεται χώρος εσωτερικού γινομένου αν μπορεί να οριστεί συνάρτηση $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ η οποία λέγεται εσωτερικό γινόμενο τέτοια ώστε να ισχύουν τα ακόλουθα:

1. $\langle a, a \rangle \geqslant 0$ για κάθε a όπου $a \in V$ και αν $\langle a, a \rangle = 0$ τότε $a = 0_V$.
2. $\langle a, b \rangle = \langle b, a \rangle$ για κάθε a, b όπου $a \in V, b \in V$.
3. $\langle \lambda a + \mu b, c \rangle = \lambda \langle a, c \rangle + \mu \langle b, c \rangle$ για κάθε a, b, c, λ, μ όπου $a \in V, b \in V, c \in V, \lambda \in \mathbb{R}, \mu \in \mathbb{R}$.

Πρόταση 1.2.1. Αν V είναι διανυσματικός χώρος τότε $\langle 0_V, a \rangle = \langle a, 0_V \rangle = 0$ για κάθε a όπου $a \in V$.

Απόδειξη. $\langle 0_V, a \rangle = \langle 0_V + 0_V, a \rangle = \langle 0_V, a \rangle + \langle 0_V, a \rangle \Leftrightarrow \langle 0_V, a \rangle = \langle a, 0_V \rangle = 0$. □

Πρόταση 1.2.2. Αν V είναι διανυσματικός χώρος και $b \in V, a_i \in V, \lambda_i \in \mathbb{R}$ για κάθε i όπου $i \in \{1, \ldots, n\}$ τότε $\langle \sum_{i=1}^{n} (\lambda_i a_i), b \rangle = \langle b, \sum_{i=1}^{n} (\lambda_i a_i) \rangle = \sum_{i=1}^{n} (\lambda_i \langle a_i, b \rangle)$ για κάθε n όπου $n \in \mathbb{N}$.

Απόδειξη. Αν $n = 1$ τότε $\langle \sum_{i=1}^{n} (\lambda_i a_i), b \rangle = \langle \sum_{i=1}^{1} (\lambda_i a_i), b \rangle = \langle \lambda_1 a_1, b \rangle = \lambda_1 \langle a_1, b \rangle = \sum_{i=1}^{1} (\lambda_i \langle a_i, b \rangle) = \sum_{i=1}^{n} (\lambda_i \langle a_i, b \rangle)$.

Υποθέτουμε ότι αν $n \in \{1, \ldots, m\}$ τότε $\langle \sum_{i=1}^{n} (\lambda_i a_i), b \rangle = \sum_{i=1}^{n} (\lambda_i \langle a_i, b \rangle)$.

Αν $n = m+1$ τότε $\langle \sum_{i=1}^{m+1} (\lambda_i a_i), b \rangle = \langle \sum_{i=1}^{m} (\lambda_i a_i) + \lambda_{m+1} a_{m+1}, b \rangle = \langle \sum_{i=1}^{m} (\lambda_i a_i), b \rangle + \langle \lambda_{m+1} a_{m+1}, b \rangle = \sum_{i=1}^{m} (\lambda_i \langle a_i, b \rangle) + \langle \lambda_{m+1} a_{m+1}, b \rangle = \sum_{i=1}^{m+1} (\lambda_i \langle a_i, b \rangle)$ άρα από την αρχή της μαθηματικής επαγωγής προκύπτει ότι $\langle \sum_{i=1}^{n} (\lambda_i a_i), b \rangle = \langle b, \sum_{i=1}^{n} (\lambda_i a_i) \rangle = \sum_{i=1}^{n} (\lambda_i \langle a_i, b \rangle)$ για κάθε n όπου $n \in \mathbb{N}$. $\square$

Αν για τα διανύσματα a, b ισχύει ότι $\langle a, b \rangle = 0$ τότε τα a, b λέγονται ορθογώνια ή κάθετα και συμβολίζουμε $a \perp b$ ενώ αν υπάρχει λ όπου $\lambda \in \mathbb{R}$ τέτοιο ώστε $a = \lambda b$ ή $b = \lambda a$ τότε τα a, b λέγονται συγγραμμικά.

Πρόταση 1.2.3 (ανισότητα Cauchy-Schwarz). Αν V είναι διανυσματικός χώρος και $a \in V, b \in V$ τότε:

$$|\langle a, b \rangle| \leqslant \langle a, a \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}} \tag{1}$$

Η ισότητα στην έκφραση (1) ισχύει αν και μόνο αν τα a, b είναι συγγραμμικά.

Απόδειξη. Έστω $\lambda \in \mathbb{R}$.
$\langle \lambda a + b, \lambda a + b \rangle \geqslant 0 \Leftrightarrow \lambda \langle a, \lambda a + b \rangle + \langle b, \lambda a + b \rangle \geqslant 0 \Leftrightarrow \lambda^2 \langle a, a \rangle + \lambda \langle a, b \rangle + \lambda \langle b, a \rangle + \langle b, b \rangle \geqslant 0 \Leftrightarrow$

$$\langle a, a \rangle \lambda^2 + 2 \langle a, b \rangle \lambda + \langle b, b \rangle \geqslant 0 \tag{2}$$

Αν $a = 0_V$ τότε $|\langle a, b \rangle| = |\langle 0_V, b \rangle| = 0 = 0 \langle b, b \rangle^{\frac{1}{2}} = \langle 0_V, 0_V \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}} = \langle a, a \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}}$.
Αν $a \neq 0_V$ τότε $\langle a, a \rangle > 0$ και από την έκφραση (2) προκύπτει ότι το τριώνυμο $p(\lambda) = \langle a, a \rangle \lambda^2 + 2 \langle a, b \rangle \lambda + \langle b, b \rangle$ είναι μεγαλύτερο ή ίσο από το 0 άρα η διακρίνουσα του $p(\lambda)$ είναι μικρότερη ή ίση από το 0 δηλαδή $4 \langle a, b \rangle^2 - 4 \langle a, a \rangle \langle b, b \rangle \leqslant 0 \Leftrightarrow 4(|\langle a, b \rangle|^2 - \langle a, a \rangle \langle b, b \rangle) \leqslant 0 \Leftrightarrow (|\langle a, b \rangle| - \langle a, a \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}})(|\langle a, b \rangle| + \langle a, a \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}}) \leqslant 0 \Leftrightarrow |\langle a, b \rangle| - \langle a, a \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}} \leqslant 0 \Leftrightarrow |\langle a, b \rangle| \leqslant \langle a, a \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}}$.
Αν $|\langle a, b \rangle| = \langle a, a \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}}$ και $a \neq 0_V$ τότε η διακρίνουσα του $p(\lambda)$ είναι ίση με 0 άρα το $p(\lambda)$ έχει μια διπλή πραγματική ρίζα λ_0 δηλαδή $p(\lambda_0) = 0 \Leftrightarrow \langle \lambda_0 a + b, \lambda_0 a + b \rangle = 0 \Leftrightarrow \lambda_0 a + b = 0_V \Leftrightarrow b = (-\lambda_0) a$.
Αν τα a, b είναι συγγραμμικά τότε υπάρχει μ όπου $\mu \in \mathbb{R}$ τέτοιο ώστε $a = \mu b$ ή $b = \mu a$ άρα $|\langle a, b \rangle| = |\langle \mu b, b \rangle| = |\mu \langle b, b \rangle| = |\mu| |\langle b, b \rangle| = (|\mu|^2)^{\frac{1}{2}} |\langle b, b \rangle|^{\frac{1}{2}} |\langle b, b \rangle|^{\frac{1}{2}} = (\mu^2 \langle b, b \rangle)^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}} = \langle \mu b, \mu b \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}} = \langle a, a \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}}$ ή $|\langle a, b \rangle| = |\langle a, \mu a \rangle| = |\mu \langle a, a \rangle| = |\mu| |\langle a, a \rangle| = (|\mu|^2)^{\frac{1}{2}} |\langle a, a \rangle|^{\frac{1}{2}} |\langle a, a \rangle|^{\frac{1}{2}} = \langle a, a \rangle^{\frac{1}{2}} (\mu^2 \langle a, a \rangle)^{\frac{1}{2}} = \langle a, a \rangle^{\frac{1}{2}} \langle \mu a, \mu a \rangle^{\frac{1}{2}} = \langle a, a \rangle^{\frac{1}{2}} \langle b, b \rangle^{\frac{1}{2}}$ αντίστοιχα. $\square$

Μια συνάρτηση $\|\cdot\| : V \to \mathbb{R}$ λέγεται νόρμα παραγόμενη από εσωτερικό γινόμενο αν $\|a\| = \langle a, a\rangle^{\frac{1}{2}}$ για κάθε διάνυσμα a του διανυσματικού χώρου V.

Πρόταση 1.2.4. Αν $\|\cdot\|$ είναι νόρμα παραγόμενη από εσωτερικό γινόμενο $\langle \cdot, \cdot\rangle$ σε διανυσματικό χώρο V τότε ισχύουν τα ακόλουθα:

1. $\|a\| = 0 \Leftrightarrow a = 0_V$ για κάθε a όπου $a \in V$.
2. $\|\lambda a\| = |\lambda|\|a\|$ για κάθε λ, a όπου $\lambda \in \mathbb{R}, a \in V$.
3. $\|a + b\| \leqslant \|a\| + \|b\|$ για κάθε a, b όπου $a \in V, b \in V$.

Απόδειξη.

1. $\|a\| = 0 \Leftrightarrow \|a\|^2 = 0 \Leftrightarrow \langle a, a\rangle = 0 \Leftrightarrow a = 0_V$.
2. $\|\lambda a\| = \langle \lambda a, \lambda a\rangle^{\frac{1}{2}} = (\lambda^2\langle a, a\rangle)^{\frac{1}{2}} = (|\lambda|^2)^{\frac{1}{2}}\langle a, a\rangle^{\frac{1}{2}} = |\lambda|\|a\|$.
3. $\|a + b\|^2 = \langle a + b, a + b\rangle = \langle a, a + b\rangle + \langle b, a + b\rangle = \langle a, a\rangle + \langle a, b\rangle + \langle b, a\rangle + \langle b, b\rangle = \|a\|^2 + 2\langle a, b\rangle + \|b\|^2 \leqslant \|a\|^2 + 2|\langle a, b\rangle| + \|b\|^2$ και από την Πρόταση 1.2.3 προκύπτει ότι $\|a\|^2 + 2|\langle a, b\rangle| + \|b\|^2 \leqslant \|a\|^2 + 2\langle a, a\rangle^{\frac{1}{2}}\langle b, b\rangle^{\frac{1}{2}} + \|b\|^2 = \|a\|^2 + 2\|a\|\|b\| + \|b\|^2 = (\|a\| + \|b\|)^2$ άρα $\|a + b\| \leqslant \|a\| + \|b\|$.

□

Η Πρόταση 1.2.4 μπορεί να χρησιμοποιηθεί για τον ορισμό νόρμας σε διανυσματικό χώρο V χωρίς την χρήση εσωτερικού γινομένου.

Πρόταση 1.2.5. Αν $\|\cdot\|$ είναι νόρμα σε διανυσματικό χώρο V τότε ισχύουν τα ακόλουθα:

1. $\|a\| \geqslant 0$ για κάθε a όπου $a \in V$.
2. $\|a + b\| \geqslant |\|a\| - \|b\||$ για κάθε a, b όπου $a \in V, b \in V$.

Απόδειξη.

1. $0 = \|a - a\| \leqslant \|a\| + \|a\| = 2\|a\|$ άρα $\|a\| \geqslant 0$.
2. $\|a\| = \|a + b - b\| \leqslant \|a + b\| + \|b\|$ άρα

$$\|a + b\| \geqslant \|a\| - \|b\| \tag{1}$$

$\|b\| = \|b + a - a\| \leqslant \|a + b\| + \|a\|$ άρα

$$\|a + b\| \geqslant \|b\| - \|a\| \tag{2}$$

Από τις εκφράσεις (1),(2) προκύπτει ότι $\|a + b\| \geqslant |\|a\| - \|b\||$.

□

Πρόταση 1.2.6. Αν $\|\cdot\|$ είναι νόρμα παραγόμενη από εσωτερικό γινόμενο $\langle \cdot, \cdot\rangle$ σε διανυσματικό χώρο V και $a_i \in V$ για κάθε i όπου $i \in \{1, \ldots, n\}$ τότε ισχύουν τα ακόλουθα:

1. $\|a_1 + a_2\|^2 + \|a_1 - a_2\|^2 = 2\|a_1\|^2 + 2\|a_2\|^2$.

2. Αν $\langle a_i, a_j\rangle = 0$ για κάθε i, j όπου $i \in \{1, \dots, n\}, j \in \{1, \dots, n\}, i \neq j$ τότε $\|\sum_{i=1}^{n} a_i\|^2 = \sum_{i=1}^{n}\|a_i\|^2$ για κάθε n όπου $n \in \mathbb{N}$.

Απόδειξη.

1. $\|a_1 + a_2\|^2 = \langle a_1 + a_2, a_1 + a_2\rangle = \langle a_1, a_1\rangle + \langle a_1, a_2\rangle + \langle a_2, a_1\rangle + \langle a_2, a_2\rangle \Leftrightarrow$
$$\|a_1 + a_2\|^2 = \|a_1\|^2 + 2\langle a_1, a_2\rangle + \|a_2\|^2 \tag{1}$$
$\|a_1 - a_2\|^2 = \langle a_1 - a_2, a_1 - a_2\rangle = \langle a_1, a_1\rangle - \langle a_1, a_2\rangle - \langle a_2, a_1\rangle + \langle a_2, a_2\rangle \Leftrightarrow$
$$\|a_1 - a_2\|^2 = \|a_1\|^2 - 2\langle a_1, a_2\rangle + \|a_2\|^2 \tag{2}$$
Αθροίζοντας κατά μέλη τις εκφράσεις (1),(2) προκύπτει ότι $\|a_1 + a_2\|^2 + \|a_1 - a_2\|^2 = 2\|a_1\|^2 + 2\|a_2\|^2$.

2. Αν $n = 1$ τότε $\|\sum_{i=1}^{n} a_i\|^2 = \|\sum_{i=1}^{1} a_i\|^2 = \|a_1\|^2 = \sum_{i=1}^{1}\|a_i\|^2$.
$\|a_1 + a_2\|^2 = \|a_1\|^2 + 2\langle a_1, a_2\rangle + \|a_2\|^2$ και $\langle a_1, a_2\rangle = 0$ άρα $\|a_1 + a_2\|^2 = \|a_1\|^2 + \|a_2\|^2$.
Υποθέτουμε ότι αν $n \in \{1, \dots, m\}$ τότε $\|\sum_{i=1}^{n} a_i\|^2 = \sum_{i=1}^{n}\|a_i\|^2$.
Αν $n = m + 1$ τότε $\|\sum_{i=1}^{m+1} a_i\|^2 = \|\sum_{i=1}^{m} a_i + a_{m+1}\|^2$ και $\langle\sum_{i=1}^{m} a_i, a_{m+1}\rangle = \sum_{i=1}^{m}\langle a_i, a_{m+1}\rangle = \sum_{i=1}^{m} 0 = 0$ άρα $\|\sum_{i=1}^{m} a_i + a_{m+1}\|^2 = \|\sum_{i=1}^{m} a_i\|^2 + \|a_{m+1}\|^2 = \sum_{i=1}^{m}\|a_i\|^2 + \|a_{m+1}\|^2 = \sum_{i=1}^{m+1}\|a_i\|^2$ επομένως από την αρχή της μαθηματικής επαγωγής προκύπτει ότι αν $\langle a_i, a_j\rangle = 0$ για κάθε i, j όπου $i \in \{1, \dots, n\}, j \in \{1, \dots, n\}, i \neq j$ τότε $\|\sum_{i=1}^{n} a_i\|^2 = \sum_{i=1}^{n}\|a_i\|^2$ για κάθε n όπου $n \in \mathbb{N}$.

□

1.3 Χώρος Hilbert

Μια ακολουθία διανυσμάτων (a_n) σε ένα διανυσματικό χώρο V λέγεται συγκλίνουσα αν υπάρχει a όπου $a \in V$ τέτοιο ώστε για κάθε ε όπου $\varepsilon > 0$ υπάρχει n_0 όπου $n_0 \in \mathbb{N}$ τέτοιο ώστε $\|a_n - a\| < \varepsilon$ για κάθε n όπου $n \geqslant n_0$ ενώ το a λέγεται όριο της ακολουθίας (a_n).

Πρόταση 1.3.1. Αν ακολουθία διανυσμάτων (a_n) σε διανυσματικό χώρο V είναι συγκλίνουσα τότε το οριό της είναι μοναδικό.

Απόδειξη. Αν a_1, a_2 είναι όρια της (a_n) όπου $a_1 \neq a_2$ τότε υπάρχουν n_1, n_2 όπου $n_1 \in \mathbb{N}, n_2 \in \mathbb{N}$ τέτοια ώστε $\|a_n - a_1\| < \dfrac{\|a_1 - a_2\|}{2}$ για κάθε n όπου $n \geqslant n_1$ και $\|a_n - a_2\| < \dfrac{\|a_1 - a_2\|}{2}$ για κάθε n όπου $n \geqslant n_2$ άρα αν $n \geqslant \max(\{n_1, n_2\})$ τότε $\|a_1 - a_2\| = \|a_1 - a_n + a_n - a_2\| \leqslant \|a_n - a_1\| + \|a_n - a_2\| < \dfrac{\|a_1 - a_2\|}{2} + \dfrac{\|a_1 - a_2\|}{2} = \|a_1 - a_2\|$ το οποίο είναι άτοπο άρα η (a_n) έχει μοναδικό όριο. □

Αν a είναι το όριο ακολουθίας (a_n) τότε συμβολίζουμε $\lim_{n\to\infty} a_n = a$ ή $\lim a_n = a$ ή $a_n \to a$. Από τον ορισμό του ορίου ακολουθίας διανυσμάτων (a_n) σε διανυσματικό χώρο V προκύπτει ότι $\lim_{n\to\infty} a_n = a \Leftrightarrow \lim_{n\to\infty}(a_n - a) = 0_V \Leftrightarrow \lim_{n\to\infty}\|a_n - a\| = 0$.

Πρόταση 1.3.2. Κάθε συγκλίνουσα ακολουθία διανυσμάτων (a_n) σε διανυσματικό χώρο V είναι φραγμένη δηλαδή υπάρχει M όπου $M > 0$ τέτοιο ώστε $\|a_n\| \leqslant M$ για κάθε n όπου $n \in \mathbb{N}$.

Απόδειξη. Αν $\lim_{n\to\infty} a_n = a$ τότε υπάρχει n_0 όπου $n_0 \in \mathbb{N}$ τέτοιο ώστε $\|a_n - a\| < 1$ για κάθε n όπου $n \geqslant n_0$ και $\|a_n\| - \|a\| \leqslant |\|a_n\| - \|a\|| \leqslant \|a_n - a\|$ για κάθε n όπου $n \in \mathbb{N}$ άρα $\|a_n\| \leqslant \max(\{\|a_1\|, \ldots, \|a_{n_0-1}\|, 1 + \|a\|\})$ για κάθε n όπου $n \in \mathbb{N}$. □

Πρόταση 1.3.3. Αν $(a_n), (b_n)$ είναι ακολουθίες διανυσμάτων στον διανυσματικό χώρο V και $\lim_{n\to\infty} a_n = a$, $\lim_{n\to\infty} b_n = b$ τότε ισχύουν τα ακόλουθα:

1. $\lim_{n\to\infty}(\lambda a_n + \mu b_n) = \lambda a + \mu b$ για κάθε λ, μ όπου $\lambda \in \mathbb{R}, \mu \in \mathbb{R}$.
2. $\lim_{n\to\infty}\langle a_n, b_n\rangle = \langle a, b\rangle$.
3. $\lim_{n\to\infty}\|a_n\| = \|a\|$.
4. $\lim_{n\to\infty}(\lambda_n a_n) = \lambda a$ για κάθε συγκλίνουσα ακολουθία πραγματικών αριθμών (λ_n) όπου $\lim_{n\to\infty}(\lambda_n) = \lambda$.

Απόδειξη. Έστω $\varepsilon > 0$. Η (a_n) είναι φραγμένη ως συγκλίνουσα άρα υπάρχει M όπου $M > 0$ τέτοιο ώστε $\|a_n\| \leqslant M$ για κάθε n όπου $n \in \mathbb{N}$.

1. Για κάθε i όπου $i \in \{1, 2\}$ θέτουμε:

$$\nu_i = \begin{cases} \lambda, & \text{αν } i = 1 \\ \mu, & \text{αν } i = 2 \end{cases} \tag{1}$$

$$\varepsilon_i = \begin{cases} \dfrac{\varepsilon}{4|\nu_i|}, & \text{αν } \nu_i \neq 0 \\ \dfrac{\varepsilon}{4}, & \text{αν } \nu_i = 0 \end{cases} \tag{2}$$

   Από τις εκφράσεις (1),(2) προκύπτει ότι $|\lambda|\varepsilon_1 \leqslant \dfrac{\varepsilon}{4}$ και $|\mu|\varepsilon_2 \leqslant \dfrac{\varepsilon}{4}$. Υπάρχουν n_1, n_2 όπου $n_1 \in \mathbb{N}, n_2 \in \mathbb{N}$ τέτοια ώστε $\|a_n - a\| < \varepsilon_1$ για κάθε n όπου $n \geqslant n_1$ και $\|b_n - b\| < \varepsilon_2$ για κάθε n όπου $n \geqslant n_2$ άρα αν $n \geqslant \max(\{n_1, n_2\})$ τότε $\|(\lambda a_n + \mu b_n) - (\lambda a + \mu b)\| = \|(\lambda a_n - \lambda a) + (\mu b_n - \mu b)\| \leqslant \|\lambda a_n - \lambda a\| + \|\mu b_n - \mu b\| = \|\lambda(a_n - a)\| + \|\mu(b_n - b)\| = |\lambda|\|a_n - a\| + |\mu|\|b_n - b\| \leqslant |\lambda|\varepsilon_1 + |\mu|\varepsilon_2 \leqslant \dfrac{\varepsilon}{4} + \dfrac{\varepsilon}{4} = \dfrac{\varepsilon}{2} < \varepsilon$ επομένως $\lim_{n\to\infty}(\lambda a_n + \mu b_n) = \lambda a + \mu b$.

2. Θέτουμε:

$$\varepsilon_0 = \begin{cases} \dfrac{\varepsilon}{4\|b\|}, & \text{αν } b \neq 0_V \\ \dfrac{\varepsilon}{4}, & \text{αν } b = 0_V \end{cases} \tag{3}$$

   Από την έκφραση (3) προκύπτει ότι $\varepsilon_0\|b\| \leqslant \dfrac{\varepsilon}{4}$. Υπάρχουν n_1, n_2 όπου $n_1 \in \mathbb{N}, n_2 \in \mathbb{N}$ τέτοια ώστε $\|a_n - a\| < \varepsilon_0$ για κάθε n όπου $n \geqslant n_1$ και $\|b_n - b\| < \dfrac{\varepsilon}{4M}$ για κάθε n

όπου $n \geqslant n_2$ άρα αν $n \geqslant \max(\{n_1, n_2\})$ τότε $|\langle a_n, b_n\rangle - \langle a, b\rangle| = |\langle a_n, b_n\rangle - \langle a_n, b\rangle + \langle a_n, b\rangle - \langle a, b\rangle| \leqslant |\langle a_n, b_n\rangle - \langle a_n, b\rangle| + |\langle a_n, b\rangle - \langle a, b\rangle| = |\langle a_n, b_n - b\rangle| + |\langle a_n - a, b\rangle| \leqslant \|a_n\|\|b_n - b\| + \|a_n - a\|\|b\| \leqslant M\frac{\varepsilon}{4M} + \varepsilon_0\|b\| \leqslant \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2} < \varepsilon$ επομένως $\lim\limits_{n\to\infty} \langle a_n, b_n\rangle = \langle a, b\rangle$.

3. $|\|a_n\| - \|a\|| \leqslant \|a_n - a\|$ για κάθε n όπου $n \in \mathbb{N}$ και $\lim\limits_{n\to\infty} \|a_n - a\| = 0$ άρα $\lim\limits_{n\to\infty} (\|a_n\| - \|a\|) = 0 \Leftrightarrow \lim\limits_{n\to\infty} \|a_n\| = \|a\|$.

4. Θέτουμε:
$$\varepsilon_0 = \begin{cases} \dfrac{\varepsilon}{4|\lambda|}, & \text{αν } \lambda \neq 0 \\ \dfrac{\varepsilon}{4}, & \text{αν } \lambda = 0 \end{cases} \tag{4}$$
Από την έκφραση (4) προκύπτει ότι $|\lambda|\varepsilon_0 \leqslant \frac{\varepsilon}{4}$. Υπάρχουν n_1, n_2 όπου $n_1 \in \mathbb{N}, n_2 \in \mathbb{N}$ τέτοια ώστε $\|\lambda_n - \lambda\| < \frac{\varepsilon}{4M}$ για κάθε n όπου $n \geqslant n_1$ και $\|a_n - a\| < \varepsilon_0$ για κάθε n όπου $n \geqslant n_2$ άρα αν $n \geqslant \max(\{n_1, n_2\})$ τότε $\|\lambda_n a_n - \lambda a\| = \|\lambda_n a_n - \lambda a_n + \lambda a_n - \lambda a\| \leqslant \|\lambda_n a_n - \lambda a_n\| + \|\lambda a_n - \lambda a\| = \|(\lambda_n - \lambda) a_n\| + \|\lambda(a_n - a)\| = |\lambda_n - \lambda|\|a_n\| + |\lambda|\|a_n - a\| \leqslant \frac{\varepsilon}{4M} M + |\lambda|\varepsilon_0 \leqslant \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2} < \varepsilon$ επομένως $\lim\limits_{n\to\infty} (\lambda_n a_n) = \lambda a$.

□

Μια ακολουθία διανυσμάτων (a_n) σε ένα διανυσματικό χώρο V λέγεται ακολουθία Cauchy αν για κάθε ε όπου $\varepsilon > 0$ υπάρχει n_0 όπου $n_0 \in \mathbb{N}$ τέτοιο ώστε $\|a_m - a_n\| < \varepsilon$ για κάθε m, n όπου $m \geqslant n_0, n \geqslant n_0$.

Πρόταση 1.3.4. Κάθε συγκλίνουσα ακολουθία διανυσμάτων (a_n) σε διανυσματικό χώρο V είναι ακολουθία Cauchy.

Απόδειξη. Έστω $\varepsilon > 0$ και $\lim\limits_{n\to\infty} a_n = a$. Υπάρχει n_0 όπου $n_0 \in \mathbb{N}$ τέτοιο ώστε $\|a_n - a\| < \frac{\varepsilon}{2}$ για κάθε n όπου $n \geqslant n_0$ άρα αν $m \geqslant n_0, n \geqslant n_0$ τότε $\|a_m - a_n\| = \|a_m - a + a - a_n\| \leqslant \|a_m - a\| + \|a_n - a\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ επομένως η (a_n) είναι ακολουθία Cauchy. □

Ένας διανυσματικός χώρος V λέγεται πλήρης αν κάθε ακολουθία Cauchy στον V είναι συγκλίνουσα. Ένας πλήρης χώρος εσωτερικού γινομένου V λέγεται χώρος Hilbert, δηλαδή χώρος Hilbert είναι ένας χώρος εσωτερικού γινομένου στον οποίο κάθε ακολουθία Cauchy είναι συγκλίνουσα ως προς την νόρμα που παράγεται από το εσωτερικό γινόμενο. Ένας υπόχωρος U του V λέγεται κλειστός υπόχωρος του V αν για κάθε συγκλίνουσα ακολουθία (a_n) τέτοια ώστε $a_n \in U$ για κάθε n όπου $n \in \mathbb{N}$ και $\lim\limits_{n\to\infty} a_n = a$ ισχύει ότι $a \in U$. Κάθε κλειστός υπόχωρος ενος χώρου Hilbert είναι χώρος Hilbert.

2 ΠΙΝΑΚΕΣ

2.1 Ορισμός Πίνακα

Έστω $m \in \mathbb{N}, n \in \mathbb{N}$. Μια διευθέτηση $m \cdot n$ το πλήθος πραγματικών αριθμών $a_{i,j}$ όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}$ σε m γραμμές και n στήλες λέγεται m επί n πραγματικός πίνακας ή απλά m επί n πίνακας και έχει την ακόλουθη μορφή:

$$\begin{bmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \vdots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{bmatrix}$$

Όταν χρησιμοποιούμε ένα κεφαλαίο γράμμα για να συμβολίσουμε ένα πίνακα τότε το αντίστοιχο πεζό γράμμα με δείκτη i, j αναφέρεται στο στοιχείο με συντεταγμένες (i, j) του πίνακα. Το σύνολο όλων των m επί n πραγματικών πινάκων συμβολίζεται ως $\mathbb{R}^{m \times n}$ και αν $A \in \mathbb{R}^{m \times n}$ τότε συμβολίζουμε $A = [a_{i,j}]_{m \times n}$ ενώ λέμε ότι ο A είναι διάστασης m επί n. Επίσης συμβολίζουμε το στοιχείο με συντεταγμένες (i, j) ενός m επί n πίνακα A ως $A(i, j)$ δηλαδή $A = [A(i, j)]_{m \times n}$. Δύο m επί n πίνακες A, B είναι ίσοι αν και μόνο αν $a_{i,j} = b_{i,j}$ για κάθε i, j όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}$. Ο m επί n πίνακας που όλα τα στοιχεία του είναι ίσα με 0 λέγεται μηδενικός και συμβολίζεται ως $0_{m \times n}$ δηλαδή $0_{m \times n} = [0]_{m \times n}$. Η κύρια διαγώνιος ενός m επί n πίνακα αποτελείται από όλα τα στοιχεία της μορφής $a_{i,i}$. Ένας 1 επί n πίνακας p λέγεται πίνακας γραμμή και συμβολίζουμε $p = [p(j)]_{1 \times n}$ ενώ ένας m επί 1 πίνακας q λέγεται πίνακας στήλη και συμβολίζουμε $q = [q(i)]_{m \times 1}$. Ένας n επί n πίνακας δηλαδή ένας πίνακας που έχει το ίδιο πλήθος γραμμών και στηλών λέγεται τετραγωνικός διάστασης n. Ένας τετραγωνικός πίνακας A λέγεται άνω τριγωνικός(αντίστοιχα αυστηρά άνω τριγωνικός) αν $a_{i,j} = 0$ για κάθε i, j όπου $i > j$(αντίστοιχα $i \geqslant j$) και κάτω τριγωνικός(αντίστοιχα αυστηρά κάτω τριγωνικός) αν $a_{i,j} = 0$ για κάθε i, j όπου $i < j$(αντίστοιχα $i \leqslant j$). Ένας n επί n πίνακας A λέγεται διαγώνιος αν είναι άνω και κάτω τριγωνικός ενώ συμβολίζουμε $A = \text{diag}(a_{1,1}, \ldots, a_{n,n})$. Ένας n επί n διαγώνιος πίνακας που έχει όλα τα στοιχεία της κυρίας διαγωνίου ισα με 1 λεγεται μοναδιαίος και συμβολίζεται ως I_n άρα $I_n = [\delta_{i,j}]_{n \times n}$ όπου:

$$\delta_{i,j} = \begin{cases} 1, & \text{αν } i = j \\ 0, & \text{αν } i \neq j \end{cases}$$

Τέλος για κάθε k όπου $k \in \{1, \ldots, n\}$ συμβολίζουμε ως $e_{k,n}$ τον n επί 1 πίνακα όπου:

$$e_{k,n}(i) = \begin{cases} 1, & \text{αν } i = k \\ 0, & \text{αν } i \neq k \end{cases}$$

2.2 Πράξεις Πινάκων

Αν $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{m \times n}$ τότε ορίζουμε ως άθροισμα του A με τον B τον m επί n πίνακα C όπου $c_{i,j} = a_{i,j} + b_{i,j}$ για κάθε i, j όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}$ και συμβολίζουμε $C = A + B$ ενώ ορίζουμε ως αντίθετο του A τον m επί n πίνακα D όπου $d_{i,j} = -a_{i,j}$ για κάθε i, j όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}$ και συμβολίζουμε $D = -A$.

Πρόταση 2.2.1. Αν $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{m \times n}, C \in \mathbb{R}^{m \times n}$ τότε ισχύουν τα ακόλουθα:

1. $A + B = B + A$.
2. $A + (B + C) = (A + B) + C$.

3. $A + 0_{m\times n} = A$.
4. $A + (-A) = 0_{m\times n}$.

Απόδειξη.

1. $A + B = [a_{i,j} + b_{i,j}]_{m\times n} = [b_{i,j} + a_{i,j}]_{m\times n} = B + A$.
2. $A + (B + C) = [a_{i,j} + (b_{i,j} + c_{i,j})]_{m\times n} = [(a_{i,j} + b_{i,j}) + c_{i,j}]_{m\times n} = (A + B) + C$.
3. $A + 0_{m\times n} = [a_{i,j} + 0]_{m\times n} = [a_{i,j}]_{m\times n} = A$.
4. $A + (-A) = [a_{i,j} + (-a_{i,j})]_{m\times n} = [0]_{m\times n} = 0_{m\times n}$.

□

Αν $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times l}$ τότε ορίζουμε ως γινόμενο του A με τον B τον m επί l πίνακα C όπου $c_{i,j} = \sum_{k=1}^{n}(a_{i,k}b_{k,j})$ για κάθε i, j όπου $i \in \{1, \dots, m\}, j \in \{1, \dots, l\}$ και συμβολίζουμε $C = A \cdot B$ ή $C = AB$.

Πρόταση 2.2.2. Αν $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times l}, C \in \mathbb{R}^{l\times r}, D \in \mathbb{R}^{n\times l}, E \in \mathbb{R}^{m\times n}$ τότε ισχύουν τα ακόλουθα:

1. $A(BC) = (AB)C$.
2. $AI_n = I_m A = A$.
3. $A(B + D) = AB + AD$.
4. $(A + E)B = AB + EB$.

Απόδειξη.

1. $A(BC) = [\sum_{k=1}^{n}(a_{i,k}\sum_{s=1}^{l}(b_{k,s}c_{s,j}))]_{m\times r} = [\sum_{k=1}^{n}\sum_{s=1}^{l}(a_{i,k}b_{k,s}c_{s,j})]_{m\times r} =$
   $[\sum_{s=1}^{l}\sum_{k=1}^{n}(a_{i,k}b_{k,s}c_{s,j})]_{m\times r} = [\sum_{s=1}^{l}(\sum_{k=1}^{n}(a_{i,k}b_{k,s})c_{s,j})]_{m\times r} = (AB)C$.
2. $AI_n = [\sum_{k=1}^{n}(a_{i,k}I_n(k, j))]_{m\times n} = [a_{i,j}]_{m\times n} = [\sum_{k=1}^{m}(I_m(i, k)a_{k,j})]_{m\times n} = [a_{i,j}]_{m\times n} = A$.
3. $A(B + D) = [\sum_{k=1}^{n}(a_{i,k}(b_{k,j} + d_{k,j}))]_{m\times l} = [\sum_{k=1}^{n}(a_{i,k}b_{k,j} + a_{i,k}d_{k,j})]_{m\times l} =$
   $[\sum_{k=1}^{n}(a_{i,k}b_{k,j}) + \sum_{k=1}^{n}(a_{i,k}d_{k,j})]_{m\times l} = AB + AD$.
4. $(A + E)B = [\sum_{k=1}^{n}((a_{i,k} + e_{i,k})b_{k,j})]_{m\times l} = [\sum_{k=1}^{n}(a_{i,k}b_{k,j} + e_{i,k}b_{k,j})]_{m\times l} =$
   $[\sum_{k=1}^{n}(a_{i,k}b_{k,j}) + \sum_{k=1}^{n}(e_{i,k}b_{k,j})]_{m\times l} = AB + EB$.

□

Αν $A \in \mathbb{R}^{n\times n}$ τότε ορίζουμε αναδρομικά:

$$A^m = \begin{cases} A, & \text{αν } m = 1 \\ A^{m-1}A, & \text{αν } m \in \mathbb{N} \setminus \{1\} \end{cases}$$

Αν $A \in \mathbb{R}^{n\times n}$ και υπάρχει n επί n πίνακας B τέτοιος ώστε $AB = BA = I_n$ τότε ο A λέγεται αντιστρέψιμος και ο B λέγεται αντίστροφος του A. Αν ο A δεν είναι αντιστρέψιμος τότε λέγεται ιδιάζων.

Πρόταση 2.2.3. Αν $A \in \mathbb{R}^{n\times n}$ και ο A είναι αντιστρέψιμος τότε ο αντιστροφός του είναι μοναδικός.

Απόδειξη. Έστω $B \in \mathbb{R}^{n\times n}, C \in \mathbb{R}^{n\times n}$ τέτοιοι ώστε $AB = BA = I_n$ και $AC = CA = I_n$.
$B = BI_n = B(AC) = (BA)C = I_nC = C$. □

Αν $A \in \mathbb{R}^{n\times n}$ και ο A είναι αντιστρέψιμος τότε ο αντιστροφός του συμβολίζεται ως A^{-1} άρα $AA^{-1} = A^{-1}A = I_n$ επομένως ο A^{-1} είναι αντιστρέψιμος και $(A^{-1})^{-1} = A$.

Πρόταση 2.2.4. Αν $A_i \in \mathbb{R}^{n\times n}$ και ο A_i είναι αντιστρέψιμος για κάθε i όπου $i \in \{1,\ldots,m\}$ τότε ο $A_1 \cdots A_m$ είναι αντιστρέψιμος και $(A_1 \cdots A_m)^{-1} = A_m^{-1} \cdots A_1^{-1}$ για κάθε m όπου $m \in \mathbb{N}$.

Απόδειξη. Αν $m = 1$ τότε $(A_1)^{-1} = A_1^{-1}$.
Υποθέτουμε ότι αν $m \in \{1,\ldots,l\}$ τότε ο $A_1 \cdots A_m$ είναι αντιστρέψιμος και $(A_1 \cdots A_m)^{-1} = A_m^{-1} \cdots A_1^{-1}$.
Αν $m = l+1$ τότε $(A_1 \cdots A_l A_{l+1})(A_{l+1}^{-1} A_l^{-1} \cdots A_1^{-1}) = A_1 \cdots A_l(A_{l+1}(A_{l+1}^{-1} A_l^{-1} \cdots A_1^{-1})) = A_1 \cdots A_l((A_{l+1}A_{l+1}^{-1})A_l^{-1} \cdots A_1^{-1}) = A_1 \cdots A_l(I_n A_l^{-1} \cdots A_1^{-1}) = A_1 \cdots A_l(A_l^{-1} \cdots A_1^{-1}) = A_1 \cdots A_l(A_1 \cdots A_l)^{-1} = I_n \Leftrightarrow$

$$(A_1 \cdots A_l A_{l+1})(A_{l+1}^{-1} A_l^{-1} \cdots A_1^{-1}) = I_n \tag{1}$$

$(A_{l+1}^{-1} \cdots A_2^{-1} A_1^{-1})(A_1 A_2 \cdots A_{l+1}) = A_{l+1}^{-1} \cdots A_2^{-1}(A_1^{-1}(A_1 A_2 \cdots A_{l+1})) = A_{l+1}^{-1} \cdots A_2^{-1}((A_1^{-1}A_1)A_2 \cdots A_{l+1}) = A_{l+1}^{-1} \cdots A_2^{-1}(I_n A_2 \cdots A_{l+1}) = A_{l+1}^{-1} \cdots A_2^{-1}(A_2 \cdots A_{l+1}) = (A_2 \cdots A_{l+1})^{-1}(A_2 \cdots A_{l+1}) = I_n \Leftrightarrow$

$$(A_{l+1}^{-1} \cdots A_2^{-1} A_1^{-1})(A_1 A_2 \cdots A_{l+1}) = I_n \tag{2}$$

Από τις εκφράσεις (1),(2) προκύπτει ότι ο $A_1 \cdots A_{l+1}$ αντιστρέφεται και $(A_1 \cdots A_{l+1})^{-1} = A_{l+1}^{-1} \cdots A_1^{-1}$ άρα από την αρχή της μαθηματικής επαγωγής προκύπτει ότι αν $A_i \in \mathbb{R}^{n\times n}$ και ο A_i είναι αντιστρέψιμος για κάθε i όπου $i \in \{1,\ldots,m\}$ τότε ο $A_1 \cdots A_m$ είναι αντιστρέψιμος και $(A_1 \cdots A_m)^{-1} = A_m^{-1} \cdots A_1^{-1}$ για κάθε m όπου $m \in \mathbb{N}$. □

Πρόταση 2.2.5. Αν $A \in \mathbb{R}^{n\times n}$ και ο A είναι αντιστρέψιμος τότε ο A^m είναι αντιστρέψιμος και $(A^m)^{-1} = (A^{-1})^m$ για κάθε m όπου $m \in \mathbb{N}$.

Απόδειξη. Προκύπτει από την Πρόταση 2.2.4 αν θέσουμε $A_i = A$ για κάθε i όπου $i \in \{1,\ldots,m\}$. □

Αν $\lambda \in \mathbb{R}, A \in \mathbb{R}^{m\times n}$ τότε ορίζουμε ως βαθμωτό γινόμενο του λ με τον A τον m επί n πίνακα C όπου $c_{i,j} = \lambda a_{i,j}$ για κάθε i,j όπου $i \in \{1,\ldots,m\}, j \in \{1,\ldots,n\}$ και συμβολίζουμε $C = \lambda \cdot A$ ή $C = \lambda A$.

Πρόταση 2.2.6. Αν $\lambda \in \mathbb{R}, \mu \in \mathbb{R}, A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{m\times n}, C \in \mathbb{R}^{n\times l}$ τότε ισχύουν τα ακόλουθα:

1. $\lambda(A+B) = \lambda A + \lambda B$.
2. $(\lambda+\mu)A = \lambda A + \mu A$.
3. $(\lambda\mu)A = \lambda(\mu A)$.
4. $\lambda(AC) = (\lambda A)C = A(\lambda C)$.

Απόδειξη.

1. $\lambda(A+B) = [\lambda(a_{i,j} + b_{i,j})]_{m\times n} = [\lambda a_{i,j} + \lambda b_{i,j}]_{m\times n} = \lambda A + \lambda B$.
2. $(\lambda+\mu)A = [(\lambda+\mu)a_{i,j}]_{m\times n} = [\lambda a_{i,j} + \mu a_{i,j}]_{m\times n} = \lambda A + \mu A$.
3. $(\lambda\mu)A = [(\lambda\mu)a_{i,j}]_{m\times n} = [\lambda(\mu a_{i,j})]_{m\times n} = \lambda(\mu A)$.
4. $\lambda(AC) = [\lambda \sum_{k=1}^{n}(a_{i,k}c_{k,j})]_{m\times n} = [\sum_{k=1}^{n}((\lambda a_{i,k})c_{k,j})]_{m\times n} = (\lambda A)C =$
   $[\sum_{k=1}^{n}(a_{i,k}(\lambda c_{k,j}))]_{m\times n} = A(\lambda C)$.

□

Αν $A \in \mathbb{R}^{m\times n}$ τότε ορίζουμε ως ανάστροφο του A τον n επί m πίνακα B όπου $b_{i,j} = a_{j,i}$ για κάθε i, j όπου $i \in \{1, \dots, n\}, j \in \{1, \dots, m\}$ και συμβολίζουμε $B = A^T$. Αν $A = A^T$ τότε ο A λέγεται συμμετρικός και αν $A = -A^T$ τότε ο A λέγεται αντισυμμετρικός.

Πρόταση 2.2.7. Αν $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times l}, \lambda \in \mathbb{R}$ τότε ισχύουν τα ακόλουθα:

1. $(A^T)^T = A$.
2. $(AB)^T = B^T A^T$.
3. $(\lambda A)^T = \lambda A^T$.
4. $I_n^T = I_n$.

Απόδειξη.

1. $(A^T)^T = [A^T(j,i)]_{m\times n} = [A(i,j)]_{m\times n} = A$.
2. $(AB)^T = [\sum_{k=1}^{n}(a_{j,k}b_{k,i})]_{l\times m} = [\sum_{k=1}^{n}(b_{k,i}a_{j,k})]_{l\times m} = [\sum_{k=1}^{n}(B^T(i,k)A^T(k,j))]_{l\times m} = B^T A^T$.
3. $(\lambda A)^T = [\lambda a_{j,i}]_{n\times m} = [\lambda A^T(i,j)]_{n\times m} = \lambda A^T$.
4. $I_n^T = [\delta_{j,i}]_{n\times n} = [\delta_{i,j}]_{n\times n} = I_n$.

□

Πρόταση 2.2.8. Αν $A_i \in \mathbb{R}^{m\times n}$ για κάθε i όπου $i \in \{1,\ldots,l\}$ τότε $(A_1+\cdots+A_l)^T = A_1^T+\cdots+A_l^T$.

Απόδειξη. $(A_1+\cdots+A_l)^T = [\sum_{k=1}^{l} A_k(j,i)]_{n\times m} = [\sum_{k=1}^{l} A_k^T(i,j)]_{n\times m} = A_1^T+\cdots+A_l^T$. □

Πρόταση 2.2.9. Αν $A_i \in \mathbb{R}^{n\times n}$ για κάθε i όπου $i \in \{1,\ldots,m\}$ τότε $(A_1\cdots A_m)^T = A_m^T\cdots A_1^T$ για κάθε m όπου $m \in \mathbb{N}$.

Απόδειξη. Αν $m=1$ τότε $(A_1)^T = A_1^T$.
Υποθέτουμε ότι αν $m \in \{1,\ldots,l\}$ τότε $(A_1\cdots A_m)^T = A_m^T\cdots A_1^T$.
Αν $m=l+1$ τότε $(A_1\cdots A_l A_{l+1})^T = ((A_1\cdots A_l)A_{l+1})^T = A_{l+1}^T(A_1\cdots A_l)^T = A_{l+1}^T A_l^T\cdots A_1$ άρα από την αρχή της μαθηματικής επαγωγής προκύπτει ότι αν $A_i \in \mathbb{R}^{n\times n}$ για κάθε i όπου $i \in \{1,\ldots,m\}$ τότε $(A_1\cdots A_m)^T = A_m^T\cdots A_1^T$ για κάθε m όπου $m \in \mathbb{N}$. □

Πρόταση 2.2.10. Αν $A \in \mathbb{R}^{n\times n}$ και ο A είναι αντιστρέψιμος τότε ο A^T είναι αντιστρέψιμος και $(A^T)^{-1} = (A^{-1})^T$.

Απόδειξη. $AA^{-1} = A^{-1}A = I_n \Leftrightarrow (AA^{-1})^T = (A^{-1}A)^T = I_n^T \Leftrightarrow$

$$(A^{-1})^T A^T = A^T (A^{-1})^T = I_n \tag{1}$$

Από την έκφραση (1) προκύπτει ότι ο πίνακας A^T είναι αντιστρέψιμος και $(A^T)^{-1} = (A^{-1})^T$.
□

2.3 Σύνθετοι Πίνακες

Σύνθετοι λέγονται οι πίνακες που τα στοιχεία τους είναι πίνακες. Για κάθε m επί n πίνακα A υπάρχουν $p\cdot q$ το πλήθος m_i επί n_j πίνακες $A_{i,j}$ όπου $i \in \{1,\ldots,p\}, j \in \{1,\ldots,q\}$ και $\sum_{k=1}^{p} m_k = m, \sum_{k=1}^{q} n_k = n$ τέτοιοι ώστε ο πίνακας A μπορεί να παρασταθεί ως σύνθετος πίνακας με την ακόλουθη μορφή:

$$A = \begin{bmatrix} A_{1,1} & \ldots & A_{1,q} \\ \vdots & \vdots & \vdots \\ A_{p,1} & \ldots & A_{p,q} \end{bmatrix} \tag{1}$$

Αν $B \in \mathbb{R}^{m\times n}$ τότε υπάρχουν $p\cdot q$ το πλήθος m_i επί n_j πίνακες $B_{i,j}$ όπου $i \in \{1,\ldots,p\}, j \in \{1,\ldots,q\}$ και $\sum_{k=1}^{p} m_k = m, \sum_{k=1}^{q} n_k = n$ τέτοιοι ώστε ο πίνακας B μπορεί να παρασταθεί ως σύνθετος πίνακας με την ακόλουθη μορφή:

$$B = \begin{bmatrix} B_{1,1} & \ldots & B_{1,q} \\ \vdots & \vdots & \vdots \\ B_{p,1} & \ldots & B_{p,q} \end{bmatrix} \tag{2}$$

Χρησιμοποιώντας τις εκφράσεις (1),(2) το άθροισμα του m επί n σύνθετου πίνακα A με τον m επί n σύνθετο πίνακα B παριστάνεται ως εξής:

$$A+B = \begin{bmatrix} A_{1,1}+B_{1,1} & \ldots & A_{1,q}+B_{1,q} \\ \vdots & \vdots & \vdots \\ A_{p,1}+B_{p,1} & \ldots & A_{p,q}+B_{p,q} \end{bmatrix}$$

Αν $C \in \mathbb{R}^{n \times l}$ τότε υπάρχουν $q \cdot r$ το πλήθος n_i επί l_j πίνακες $C_{i,j}$ όπου $i \in \{1, \ldots, q\}, j \in \{1, \ldots, r\}$ και $\sum_{k=1}^{q} n_k = n, \sum_{k=1}^{p} l_k = l$ τέτοιοι ώστε ο πίνακας C μπορεί να παρασταθεί ως σύνθετος πίνακας με την ακόλουθη μορφή:

$$C = \begin{bmatrix} C_{1,1} & \ldots & C_{1,r} \\ \vdots & \vdots & \vdots \\ C_{q,1} & \ldots & C_{q,r} \end{bmatrix} \tag{3}$$

Χρησιμοποιώντας τις εκφράσεις (1),(3) το γινόμενο του m επί n σύνθετου πίνακα A με τον n επί l σύνθετο πίνακα C παριστάνεται ως εξής:

$$AC = \begin{bmatrix} A_{1,1}C_{1,1} + \cdots + A_{1,q}C_{q,1} & \ldots & A_{1,1}C_{1,r} + \cdots + A_{1,q}C_{q,r} \\ \vdots & \vdots & \vdots \\ A_{p,1}C_{1,1} + \cdots + A_{p,q}C_{q,1} & \ldots & A_{p,1}C_{1,r} + \cdots + A_{p,q}C_{q,r} \end{bmatrix}$$

Αν $\lambda \in \mathbb{R}$ τότε χρησιμοποιώντας την έκφραση (1) το βαθμωτό γινόμενο του λ με τον m επί n σύνθετο πίνακα A παριστάνεται ως εξής:

$$\lambda A = \begin{bmatrix} \lambda A_{1,1} & \ldots & \lambda A_{1,q} \\ \vdots & \vdots & \vdots \\ \lambda A_{p,1} & \ldots & \lambda A_{p,q} \end{bmatrix}$$

Χρησιμοποιώντας την έκφραση (1) ο n επί m σύνθετος πίνακας A^T παριστάνεται ως εξής:

$$A^T = \begin{bmatrix} A_{1,1}^T & \ldots & A_{p,1}^T \\ \vdots & \vdots & \vdots \\ A_{1,q}^T & \ldots & A_{p,q}^T \end{bmatrix}$$

Ειδική περίπτωση σύνθετου πίνακα είναι ο m επί n πίνακας A που τα στοιχεία του είναι είτε πίνακες γραμμή, είτε πίνακες στήλη δηλαδή $A \in \mathbb{R}^{m \times n}$ και $b_i \in \mathbb{R}^{1 \times n}, c_j \in \mathbb{R}^{m \times 1}$ για κάθε i, j όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}$ τέτοιοι ώστε:

$$A = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = [c_1 \ \ldots \ c_n]$$

2.4 Τάξη Πίνακα

Αν $x = (x_1, \ldots, x_n)$ όπου $x_i \in \mathbb{R}$ για κάθε i όπου $i \in \{1, \ldots, n\}$ τότε $x \in \mathbb{R}^n$ και συμβολίζουμε:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{1}$$

Από την έκφραση (1) προκύπτει ότι ταυτίζουμε το $\mathbb{R}^n$ με το $\mathbb{R}^{n \times 1}$ άρα τα διανύσματα του $\mathbb{R}^n$ είναι πίνακες στήλη και θέτουμε $0_n = 0_{n \times 1}$.
Ορίζουμε ως χώρο στηλών ενός m επί n πίνακα A το σύνολο:

$$R(A) = \{y : y \in \mathbb{R}^m \text{ και υπάρχει } x \text{ όπου } x \in \mathbb{R}^n \text{ τέτοιο ώστε } y = Ax\}$$

Αν $A = [a_1 \ \dots \ a_n]$ όπου $a_j \in \mathbb{R}^m$ για κάθε j όπου $j \in \{1, \dots, n\}$ και $x \in \mathbb{R}^n$ τότε $Ax = [\sum_{j=1}^{n}(a_j(1)x(j)) \ \dots \ \sum_{j=1}^{n}(a_j(m)x(j))]^T = [a_1(1)x(1) \ \dots \ a_1(m)x(1)]^T + \dots + [a_n(1)x(n) \ \dots \ a_n(m)x(n)]^T = x(1)[a_1(1) \ \dots \ a_1(m)]^T + \dots + x(n)[a_n(1) \ \dots \ a_n(m)]^T = x(1)a_1 + \dots + x(n)a_n$ άρα:

$$R(A) = \text{span}(\{a_1, \dots, a_n\}) \tag{2}$$

Η τάξη ενός m επί n πίνακα A ορίζεται ως η διάσταση του χώρου στηλών του και συμβολίζεται ως $\text{rank}(A)$ δηλαδή $\text{rank}(A) = \dim(R(A))$ ενώ από την έκφραση (2) προκύπτει ότι η τάξη ενός πίνακα ισούται με το μέγιστο πλήθος των γραμμικώς ανεξάρτητων στηλών του άρα οι στήλες του A είναι γραμμικώς ανεξάρτητες αν και μόνο αν $\text{rank}(A) = n$.
Ορίζουμε ως πυρήνα ενός m επί n πίνακα A το σύνολο:

$$N(A) = \{x : x \in \mathbb{R}^n \text{ και } Ax = 0_m\}$$

Από την έκφραση (2) προκύπτει ότι αν οι στήλες του A είναι γραμμικώς ανεξάρτητες δηλαδή αν $\text{rank}(A) = n$ τότε $N(A) = \{0_n\}$.
Ένας m επί n πίνακας A λέγεται πλήρους τάξης αν $\text{rank}(A) = \min(\{m, n\})$ και ελλιπούς τάξης αν $\text{rank}(A) < \min(\{m, n\})$.
Από την Γραμμική Άλγεβρα είναι γνωστή η ακόλουθη πρόταση:

Πρόταση 2.4.1. Ένας τετραγωνικός πίνακας είναι αντιστρέψιμος αν και μόνο αν είναι πλήρους τάξης.

Από την Πρόταση 2.4.1 προκύπτει ότι ένας n επί n πίνακας A είναι αντιστρέψιμος αν και μόνο αν $N(A) = \{0_n\}$.
Έστω $A \in \mathbb{R}^{m \times n}$ και $b_i \in \mathbb{R}^{1 \times n}$ για κάθε i όπου $i \in \{1, \dots, m\}$ τέτοιοι ώστε:

$$A = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

Αν $r \in \{1, \dots, m\}$ και $p_k \in \{1, \dots, m\}$ για κάθε k όπου $k \in \{1, \dots, r\}$ τότε λέμε ότι οι γραμμές $b_{p_1}, \dots, b_{p_r}$ του πίνακα A είναι γραμμικώς ανεξάρτητες αν οι πίνακες στήλη $b_{p_1}^T, \dots, b_{p_r}^T$ είναι γραμμικώς ανεξάρτητοι.

Πρόταση 2.4.2. Ένας n επί n πίνακας A έχει n γραμμικώς ανεξάρτητες στήλες αν και μόνο αν έχει n γραμμικώς ανεξάρτητες γραμμές.

Απόδειξη. Έστω $A \in \mathbb{R}^{n \times n}$ και $b_i \in \mathbb{R}^{1 \times n}, c_j \in \mathbb{R}^n$ για κάθε i, j όπου $i \in \{1, \dots, n\}, j \in \{1, \dots, n\}$ τέτοιοι ώστε:

$$A = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = [c_1 \ \dots \ c_n] \tag{1}$$

Από την έκφραση (1) προκύπτει ότι $b_i = [c_1(i) \ \dots \ c_n(i)]$ για κάθε i όπου $i \in \{1, \dots, n\}$ και αν $x_i \in \mathbb{R}$ για κάθε i όπου $i \in \{1, \dots, n\}$ τότε $x_1 b_1^T + \dots + x_n b_n^T = 0_n \Leftrightarrow x_1[c_1(1) \ \dots \ c_n(1)]^T + \dots + x_n[c_1(n) \ \dots \ c_n(n)]^T = 0_n \Leftrightarrow [x_1 c_1(1) \ \dots \ x_1 c_n(1)]^T + \dots + [x_n c_1(n) \ \dots \ x_n c_n(n)]^T = 0_n \Leftrightarrow [x_1 c_1(1) + \dots + x_n c_1(n) \ \dots \ x_1 c_n(1) + \dots + x_n c_n(n)]^T = 0_n \Leftrightarrow$

$$\begin{bmatrix} x_1c_1(1)+\cdots+x_nc_1(n) \\ \vdots \\ x_1c_n(1)+\cdots+x_nc_n(n) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} c_1(1) & \dots & c_1(n) \\ \vdots & \vdots & \vdots \\ c_n(1) & \dots & c_n(n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Leftrightarrow$$
$$\begin{bmatrix} c_1^T \\ \vdots \\ c_n^T \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \xLeftrightarrow{x=[x_1\ \dots\ x_n]^T}$$

$$A^T x = 0_n \tag{2}$$

Αν ο πίνακας A έχει n γραμμικώς ανεξάρτητες στήλες τότε $N(A) = \{0_n\}$ άρα ο πίνακας A είναι αντιστρέψιμος επομένως ο πίνακας A^T είναι αντιστρέψιμος συνεπώς από την έκφραση (2) προκύπτει ότι $(A^T)^{-1}A^Tx = (A^T)^{-1}0_n \Leftrightarrow x = 0_n$ οπότε ο πίνακας A έχει n γραμμικώς ανεξάρτητες γραμμές.
Από την έκφραση (1) προκύπτει ότι:

$$A^T = [b_1^T\ \dots\ b_n^T] \tag{3}$$

$x_1c_1 + \cdots + x_nc_n = 0_n \Leftrightarrow x_1[c_1(1)\ \dots\ c_1(n)]^T + \cdots + x_n[c_n(1)\ \dots\ c_n(n)]^T = 0_n \Leftrightarrow$ $[x_1c_1(1)\ \dots\ x_1c_1(n)]^T + \cdots + [x_nc_n(1)\ \dots\ x_nc_n(n)]^T = 0_n \Leftrightarrow$

$$[x_1c_1(1)+\cdots+x_nc_n(1)\ \dots\ x_1c_1(n)+\cdots+x_nc_n(n)]^T = 0_n \Leftrightarrow \begin{bmatrix} x_1c_1(1)+\cdots+x_nc_n(1) \\ \vdots \\ x_1c_1(n)+\cdots+x_nc_n(n) \end{bmatrix} =$$
$$\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} c_1(1) & \dots & c_n(1) \\ \vdots & \vdots & \vdots \\ c_1(n) & \dots & c_n(n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Leftrightarrow$$

$$Ax = 0_n \tag{4}$$

Αν ο πίνακας A έχει n γραμμικώς ανεξάρτητες γραμμές τότε από την έκφραση (3) προκύπτει ότι $N(A^T) = \{0_n\}$ άρα ο πίνακας A^T είναι αντιστρέψιμος επομένως ο πίνακας $(A^T)^T = A$ είναι αντιστρέψιμος συνεπώς από την έκφραση (4) προκύπτει ότι $A^{-1}Ax = A^{-1}0_n \Leftrightarrow x = 0_n$ οπότε ο πίνακας A έχει n γραμμικώς ανεξάρτητες στήλες. □

Πρόταση 2.4.3. Έστω $A \in \mathbb{R}^{m\times n}$ και $r \in \{1,\dots,\min(\{m,n\})\}$. Ο πίνακας A έχει r γραμμικώς ανεξάρτητες στήλες αν και μόνο αν έχει r γραμμικώς ανεξάρτητες γραμμές.

Απόδειξη. Θέτουμε $A = [a_1\ \dots\ a_n]$ όπου $a_j \in \mathbb{R}^m$ για κάθε j όπου $j \in \{1,\dots,n\}$.
Αν $q_k \in \{1,\dots,\min(\{m,n\})\}$ για κάθε k όπου $k \in \{1,\dots,r\}$ με $q_s < q_{s+1}$ για κάθε s όπου $s \in \{1,\dots,r-1\}$ και οι στήλες $a_{q_1},\dots,a_{q_r}$ είναι γραμμικώς ανεξάρτητες τότε υπάρχει p_k όπου $p_k \in \{1,\dots,\min(\{m,n\})\}$ για κάθε k όπου $k \in \{1,\dots,r\}$ με $p_s < p_{s+1}$ για κάθε s όπου $s \in \{1,\dots,r-1\}$ τέτοιο ώστε οι πίνακες στήλη $[a_{q_1}(p_1)\ \dots\ a_{q_1}(p_r)]^T,\dots,[a_{q_r}(p_1)\ \dots\ a_{q_r}(p_r)]^T$ είναι γραμμικώς ανεξάρτητοι. Ορίζουμε τον r επί r πίνακα:

$$A_r = \begin{bmatrix} a_{q_1}(p_1) & \dots & a_{q_r}(p_1) \\ \vdots & \vdots & \vdots \\ a_{q_1}(p_r) & \dots & a_{q_r}(p_r) \end{bmatrix}$$

Ο r επί r πίνακας A_r έχει r γραμμικώς ανεξάρτητες στήλες άρα έχει r γραμμικώς ανεξάρτητες γραμμές επομένως ο πίνακας A έχει r γραμμικώς ανεξάρτητες γραμμές.
Αν ο πίνακας A έχει r γραμμικώς ανεξάρτητες γραμμές τότε αυτές είναι r γραμμικώς ανεξάρτητες στήλες για τον πίνακα A^T άρα ο πίνακας A^T έχει r γραμμικώς ανεξάρτητες γραμμές οι οποίες είναι r γραμμικώς ανεξάρτητες στήλες για τον πίνακα A. □

Πρόταση 2.4.4. Αν $A \in \mathbb{R}^{m\times n}$ τότε rank$(A) =$ rank(A^T).

Απόδειξη. Η απόδειξη προκύπτει άμεσα από την Πρόταση 2.4.3. □

2.5 Ευκλείδεια Νόρμα

Ορίζουμε $\langle\cdot,\cdot\rangle_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ όπου $\langle x, y\rangle_2 = x^T y$ για κάθε x, y όπου $x \in \mathbb{R}^n, y \in \mathbb{R}^n$.

Πρόταση 2.5.1. Αν $x \in \mathbb{R}^n, y \in \mathbb{R}^n, z \in \mathbb{R}^n, \lambda \in \mathbb{R}, \mu \in \mathbb{R}$ τότε ισχύουν τα ακόλουθα:

1. $\langle x, x\rangle_2 \geqslant 0$ και αν $\langle x, x\rangle_2 = 0$ τότε $x = 0_n$.
2. $\langle x, y\rangle_2 = \langle y, x\rangle_2$.
3. $\langle \lambda x + \mu y, z\rangle_2 = \lambda\langle x, z\rangle_2 + \mu\langle y, z\rangle_2$.

Απόδειξη.

1. $\langle x, x\rangle_2 = x^T x = \sum_{i=1}^{n}(x^2(i)) \geqslant 0$ και αν $\langle x, x\rangle_2 = 0$ τότε $\sum_{i=1}^{n}(x^2(i)) = 0 \Leftrightarrow x(i) = 0$ για κάθε i όπου $i \in \{1, \dots, n\}$ άρα $x = 0_n$.
2. $\langle x, y\rangle_2 = x^T y = \sum_{i=1}^{n}(x(i)y(i)) = \sum_{i=1}^{n}(y(i)x(i)) = y^T x$.
3. $\langle \lambda x + \mu y, z\rangle_2 = (\lambda x + \mu y)^T z = (\lambda x^T + \mu y^T)z = \lambda x^T z + \mu y^T z = \lambda\langle x, z\rangle_2 + \mu\langle y, z\rangle_2$.

□

Από την Πρόταση 2.5.1 προκύπτει ότι η συνάρτηση $\langle\cdot,\cdot\rangle_2$ είναι εσωτερικό γινόμενο και την λέμε Ευκλείδειο εσωτερικό γινόμενο. Συμβολίζουμε ως $\|\cdot\|_2$ την νόρμα που παράγεται από το Ευκλείδειο εσωτερικό γινόμενο και την λέμε Ευκλείδεια νόρμα. Από την Ανάλυση είναι γνωστή η ακόλουθη πρόταση:

Πρόταση 2.5.2. Κάθε ακολουθία Cauchy είναι συγκλίνουσα στον $\mathbb{R}^n$ ως προς την Ευκλείδεια νόρμα.

Από την Πρόταση 2.5.2 προκύπτει ότι ο $\mathbb{R}^n$ είναι χώρος Hilbert.

2.6 Φασματική Νόρμα

Αν $A \in \mathbb{R}^{m\times n}$ τέτοιος ώστε $A = [a_1 \ \dots \ a_n]$ όπου $a_j \in \mathbb{R}^m$ για κάθε j όπου $j \in \{1, \dots, n\}$ και $x \in \mathbb{R}^n \backslash \{0_n\}$ τότε $\frac{\|Ax\|_2}{\|x\|_2} = \frac{\|x(1)a_1 + \dots + x(n)a_n\|_2}{\|x\|_2} \leqslant \frac{\sum_{j=1}^{n}(|x(j)|\|a_j\|_2)}{\|x\|_2} \leqslant \frac{\sum_{j=1}^{n}(\|x\|_2\|a_j\|_2)}{\|x\|_2} = \sum_{j=1}^{n}(\|a_j\|_2)$ άρα υπάρχει το sup$(\{\frac{\|Ax\|_2}{\|x\|_2} : x \in \mathbb{R}^n \setminus \{0_n\}\})$ και ορίζουμε $\|\cdot\|_2 : \mathbb{R}^{m\times n} \to \mathbb{R}$ όπου $\|A\|_2 =$ sup$(\{\frac{\|Ax\|_2}{\|x\|_2} : x \in \mathbb{R}^n \setminus \{0_n\}\})$ για κάθε A όπου $A \in \mathbb{R}^{m\times n}$ επομένως $\|Ax\|_2 \leqslant \|A\|_2\|x\|_2$ για κάθε x όπου $x \in \mathbb{R}^n$.

Πρόταση 2.6.1. Αν $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{m\times n}, C \in \mathbb{R}^{n\times l}, \lambda \in \mathbb{R}$ τότε ισχύουν τα ακόλουθα:

1. $\|A\|_2 = 0 \Leftrightarrow A = 0_{m\times n}$.
2. $\|\lambda A\|_2 = |\lambda|\|A\|_2$.
3. $\|A+B\|_2 \leqslant \|A\|_2 + \|B\|_2$.
4. $\|AC\|_2 \leqslant \|A\|_2\|C\|_2$.

Απόδειξη.

1. Θέτουμε $A = [a_1 \ \dots \ a_n]$ όπου $a_j \in \mathbb{R}^m$ για κάθε j όπου $j \in \{1,\dots,n\}$. Αν $\|A\|_2 = 0$ τότε $\dfrac{\|Ae_{j,n}\|_2}{\|e_{j,n}\|_2} = 0$ για κάθε j όπου $j \in \{1,\dots,n\}$ δηλαδή $\|Ae_{j,n}\|_2 = 0$ για κάθε j όπου $j \in \{1,\dots,n\}$ δηλαδή $Ae_{j,n} = 0_m$ για κάθε j όπου $j \in \{1,\dots,n\}$ δηλαδή $e_{j,n}(1)a_1 + \cdots + e_{j,n}(n)a_n = 0_m$ για κάθε j όπου $j \in \{1,\dots,n\}$ δηλαδή $a_j = 0_m$ για κάθε j όπου $j \in \{1,\dots,n\}$ δηλαδή $A = 0_{m\times n}$.
   Αν $A = 0_{m\times n}$ τότε $\|A\|_2 = \|0_{m\times n}\|_2 = \sup(\{\dfrac{\|0_{m\times n}x\|_2}{\|x\|_2} : x \in \mathbb{R}^n \setminus \{0_n\}\}) = \sup(\{0\}) = 0$.

2. Αν $\lambda = 0$ τότε $\|\lambda A\|_2 = \|0A\|_2 = \|0_{m\times n}\|_2 = 0 = |0|\|A\|_2 = |\lambda|\|A\|_2$.
   Αν $\lambda \neq 0$ τότε $\|\lambda A\|_2 = \sup(\{\dfrac{\|\lambda Ax\|_2}{\|x\|_2} : x \in \mathbb{R}^n \setminus \{0_n\}\})$ άρα $\dfrac{\|\lambda Ax\|_2}{\|x\|_2} \leqslant \|\lambda A\|_2$ για κάθε x όπου $x \in \mathbb{R}^n \setminus \{0_n\}$ δηλαδή $\dfrac{|\lambda|\|Ax\|_2}{\|x\|_2} \leqslant \|\lambda A\|_2$ για κάθε x όπου $x \in \mathbb{R}^n \setminus \{0_n\}$ δηλαδή $\dfrac{\|Ax\|_2}{\|x\|_2} \leqslant \dfrac{\|\lambda A\|_2}{|\lambda|}$ για κάθε x όπου $x \in \mathbb{R}^n \setminus \{0_n\}$ δηλαδή:
   $$\|\lambda A\|_2 \geqslant |\lambda|\|A\|_2 \tag{1}$$
   Αν $x \in \mathbb{R}^n \setminus \{0_n\}$ τότε $\dfrac{\|Ax\|_2}{\|x\|_2} \leqslant \|A\|_2 \Leftrightarrow \dfrac{|\lambda|\|Ax\|_2}{\|x\|_2} \leqslant |\lambda|\|A\|_2 \Leftrightarrow \dfrac{\|\lambda Ax\|_2}{\|x\|_2} \leqslant |\lambda|\|A\|_2 \Leftrightarrow$
   $$\|\lambda A\|_2 \leqslant |\lambda|\|A\|_2 \tag{2}$$
   Από τις εκφράσεις (1),(2) προκύπτει ότι $\|\lambda A\|_2 = |\lambda|\|A\|_2$.

3. Αν $x \in \mathbb{R}^n \setminus \{0_n\}$ τότε $\dfrac{\|(A+B)x\|_2}{\|x\|_2} = \dfrac{\|Ax+Bx\|_2}{\|x\|_2} \leqslant \dfrac{\|Ax\|_2 + \|Bx\|_2}{\|x\|_2} = \dfrac{\|Ax\|_2}{\|x\|_2} + \dfrac{\|Bx\|_2}{\|x\|_2} \leqslant \|A\|_2 + \|B\|_2$ δηλαδή $\|A+B\|_2 \leqslant \|A\|_2 + \|B\|_2$.

4. Αν $x \in \mathbb{R}^l \setminus \{0_l\}$ τότε $\dfrac{\|ACx\|_2}{\|x\|_2} = \dfrac{\|A(Cx)\|_2}{\|x\|_2} \leqslant \|A\|_2\dfrac{\|Cx\|_2}{\|x\|_2} \leqslant \|A\|_2\|C\|_2$ δηλαδή $\|AC\|_2 \leqslant \|A\|_2\|C\|_2$.

□

Από την Πρόταση 2.6.1 προκύπτει ότι η $\|\cdot\|_2$ είναι νόρμα πίνακα και λέγεται φασματική νόρμα.

Πρόταση 2.6.2. $\|I_n\|_2 = 1$.

Απόδειξη. $\|I_n\|_2 = \sup(\{\frac{\|I_n x\|_2}{\|x\|_2} : x \in \mathbb{R}^n \setminus \{0_n\}\}) = \sup(\{\frac{\|x\|_2}{\|x\|_2} : x \in \mathbb{R}^n \setminus \{0_n\}\}) = \sup(\{1\}) = 1$. □

Πρόταση 2.6.3. Αν $A \in \mathbb{R}^{m\times n}$ τότε ο $R(A)$ είναι κλειστός υπόχωρος του $\mathbb{R}^m$.

Απόδειξη. Έστω συγκλίνουσα ακολουθία (y_l) τέτοια ώστε $y_l \in R(A)$ για κάθε l όπου $l \in \mathbb{N}$ και $\lim\limits_{l\to\infty} y_l = y$. Θέτουμε $A = [a_1 \ \dots \ a_n]$ όπου $a_j \in \mathbb{R}^m$ για κάθε j όπου $j \in \{1, \dots, n\}$ και από τον ορισμό της (y_l) προκύπτει ότι υπάρχει ακολουθία (x_l) τέτοια ώστε $y_l = Ax_l$ για κάθε l όπου $l \in \mathbb{N}$ δηλαδή $y_l = x_l(1)a_1 + \dots + x_l(n)a_n$ για κάθε l όπου $l \in \mathbb{N}$.
Αν $A = 0_{m\times n}$ τότε $y_l = 0_{m\times n}x_l = 0_m$ άρα $\lim\limits_{l\to\infty} y_l = \lim\limits_{l\to\infty} 0_m = 0_m$ και $0_m \in R(0_{m\times n})$.
Αν $\operatorname{rank}(A) = r$ όπου $r \in \{1, \dots, \min(\{m, n\})\}$ τότε για κάθε k όπου $k \in \{1, \dots, r\}$ υπάρχει q_k όπου $q_k \in \{1, \dots, n\}$ με $q_s < q_{s+1}$ για κάθε s όπου $s \in \{1, \dots, r-1\}$ τετοιο ώστε τα $a_{q_1}, \dots, a_{q_r}$ είναι γραμμικώς ανεξάρτητα άρα υπάρχουν μοναδικές ακολουθίες πραγματικών αριθμών $(z_{1,l}), \dots, (z_{r,l})$ τέτοιες ώστε $y_l = z_{1,l}a_{q_1} + \dots + z_{r,l}a_{q_r}$ επομένως:

$$y_l(i) = \sum_{j=1}^{r}(z_{j,l}a_{q_j}(i)), i \in \{1, \dots, m\} \tag{1}$$

Ο m επί r πίνακας $[a_{q_1} \ \dots \ a_{q_r}]$ έχει r γραμμικώς ανεξάρτητες στήλες άρα έχει r γραμμικώς ανεξάρτητες γραμμές. Αν οι δείκτες των r γραμμικώς ανεξάρτητων γραμμών του $[a_{q_1} \ \dots \ a_{q_r}]$ είναι $p_1, \dots, p_r$ όπου $p_k \in \{1, \dots, m\}$ για κάθε k όπου $k \in \{1, \dots, r\}$ με $p_s < p_{s+1}$ για κάθε s όπου $s \in \{1, \dots, r-1\}$ τότε ορίζουμε τους r επί 1 πίνακες $h_l = [y_l(p_1) \ \dots \ y_l(p_r)]^T, z_l = [z_{1,l} \ \dots \ z_{r,l}]^T$ και τον αντιστρέψιμο r επί r πίνακα:

$$A_r = \begin{bmatrix} a_{q_1}(p_1) & \dots & a_{q_r}(p_1) \\ \vdots & \vdots & \vdots \\ a_{q_1}(p_r) & \dots & a_{q_r}(p_r) \end{bmatrix} \tag{2}$$

Από την έκφραση (1) προκύπτει ότι $y_l(p_k) = \sum\limits_{j=1}^{r}(z_{j,l}a_{q_j}(p_k))$ για κάθε k όπου $k \in \{1, \dots, r\}$ άρα από την έκφραση (2) προκύπτει ότι $h_l(k) = \sum\limits_{j=1}^{r}(A_r(k,j)z_{j,l})$ για κάθε k όπου $k \in \{1, \dots, r\}$ δηλαδή $h_l = A_r z_l \Leftrightarrow z_l = A_r^{-1}h_l$ δηλαδή $z_l(k) = \sum\limits_{j=1}^{r}(A_r^{-1}(k,j)h_l(j))$ για κάθε k όπου $k \in \{1, \dots, r\}$ δηλαδή $z_l(k) = \sum\limits_{j=1}^{r}(A_r^{-1}(k,j)y_l(p_j))$ για κάθε k όπου $k \in \{1, \dots, r\}$ επομένως $\lim\limits_{l\to\infty} z_l(k) = \lim\limits_{l\to\infty} \sum\limits_{j=1}^{r}(A_r^{-1}(k,j)y_l(p_j)) \Leftrightarrow \lim\limits_{l\to\infty} z_{k,l} = \sum\limits_{j=1}^{r}(A_r^{-1}(k,j)\lim\limits_{l\to\infty} y_l(p_j))$ συνεπώς η ακολουθία $(z_{k,l})$ είναι συγκλίνουσα για κάθε k όπου $k \in \{1, \dots, r\}$ και από την έκφραση (1) προκύπτει ότι $y_l = 0a_1 + \dots + 0a_{q_1-1} + z_{1,l}a_{q_1} + 0a_{q_1+1} + \dots + 0a_{q_2-1} + z_{2,l}a_{q_2} + \dots + z_{r-1,l}a_{q_{r-1}} + 0a_{q_{r-1}+1} + \dots + 0a_{q_r-1} + z_{r,l}a_{q_r} + 0a_{q_r+1} + \dots + 0a_n \Leftrightarrow y_l = A[0 \ \dots \ 0 \ z_{1,l} \ 0 \ \dots \ 0 \ z_{2,l} \ \dots \ z_{r-1,l} \ 0 \ \dots \ 0 \ z_{r,l} \ 0 \ \dots \ 0]^T$.
Θέτουμε $v_l = [0 \ \dots \ 0 \ z_{1,l} \ 0 \ \dots \ 0 \ z_{2,l} \ \dots \ z_{r-1,l} \ 0 \ \dots \ 0 \ z_{r,l} \ 0 \ \dots \ 0]^T$ και $\lim\limits_{l\to\infty} z_{k,l} = t_k$ για κάθε k όπου $k \in \{1, \dots, r\}$ άρα $y_l = Av_l$ και
$\lim\limits_{l\to\infty} v_l = [0 \ \dots \ 0 \ \lim\limits_{l\to\infty} z_{1,l} \ 0 \ \dots \ 0 \ \lim\limits_{l\to\infty} z_{2,l} \ \dots \ \lim\limits_{l\to\infty} z_{r-1,l} \ 0 \ \dots \ 0 \ \lim\limits_{l\to\infty} z_{r,l} \ 0 \ \dots \ 0]^T = [0 \ \dots \ 0 \ t_1 \ 0 \ \dots \ 0 \ t_2 \ \dots \ t_{r-1} \ 0 \ \dots \ 0 \ t_r \ 0 \ \dots \ 0]^T$.

Έστω $\varepsilon > 0$. Θέτουμε $v = [0 \ \dots \ 0 \ t_1 \ 0 \ \dots \ 0 \ t_2 \ \dots \ t_{r-1} \ 0 \ \dots \ 0 \ t_r \ 0 \ \dots \ 0]^T$ και υπάρχει l_0 όπου $l_0 \in \mathbb{N}$ τέτοιο ώστε $\|v_l - v\|_2 < \dfrac{\varepsilon}{\|A\|_2}$ για κάθε l όπου $l \geqslant l_0$ άρα $\|y_l - Av\|_2 = \|Av_l - Av\|_2 = \|A(v_l - v)\|_2 \leqslant \|A\|_2\|v_l - v\|_2 < \|A\|_2\dfrac{\varepsilon}{\|A\|_2} = \varepsilon$ για κάθε l όπου $l \geqslant l_0$ επομένως $\lim\limits_{l\to\infty} y_l = Av = y$ συνεπώς $y \in R(A)$ οπότε ο $R(A)$ είναι κλειστός υπόχωρος του $\mathbb{R}^m$. □

2.7 Ορθογωνιότητα

Ένα σύνολο διανυσμάτων $\{x_1, \dots, x_l\}$ του $\mathbb{R}^n$ λέγεται ορθογώνιο αν $\langle x_i, x_j\rangle_2 = 0$ για κάθε i, j όπου $i \in \{1, \dots, l\}, j \in \{1, \dots, l\}, i \neq j$ και ορθοκανονικό αν $\langle x_i, x_j\rangle_2 = \delta_{i,j}$ για κάθε i, j όπου $i \in \{1, \dots, l\}, j \in \{1, \dots, l\}$. Ορίζουμε ως ορθογώνιο συμπλήρωμα ενός υπόχωρου V του $\mathbb{R}^n$ το σύνολο $V^\perp = \{x : x \in \mathbb{R}^n$ και $\langle x, y\rangle_2 = 0$ για κάθε y όπου $y \in V\}$.

Πρόταση 2.7.1. Αν V είναι υπόχωρος του $\mathbb{R}^n$ τότε $V \cap V^\perp = \{0_V\}$.

Απόδειξη. Αν $x \in V \cap V^\perp$ τότε $\langle x, x\rangle_2 = 0 \Leftrightarrow x = 0_V$ άρα $V \cap V^\perp = \{0_V\}$. □

Πρόταση 2.7.2. Αν $A \in \mathbb{R}^{m\times n}$ τότε $R(A)^\perp = N(A^T)$.

Απόδειξη. Αν $x \in R(A)^\perp$ τότε $\langle x, y\rangle_2 = 0$ για κάθε y όπου $y \in R(A)$ και $\langle A^T x, A^T x\rangle_2 = x^T A A^T x = \langle x, AA^T x\rangle_2 = 0$ διότι $AA^T x \in R(A)$ άρα $A^T x = 0_n$ επομένως $x \in N(A^T)$ συνεπώς

$$R(A)^\perp \subseteq N(A^T) \tag{1}$$

Αν $x \in N(A^T)$ τότε $A^T x = 0_n$ και για κάθε y οπου $y \in R(A)$ υπάρχει z όπου $z \in \mathbb{R}^n$ τέτοιο ώστε $y = Az$ άρα $\langle x, y\rangle_2 = \langle x, Az\rangle_2 = \langle Az, x\rangle_2 = z^T A^T x = z^T 0_n = 0$ επομένως $x \in R(A)^\perp$ συνεπώς

$$N(A^T) \subseteq R(A)^\perp \tag{2}$$

Από τις εκφράσεις (1),(2) προκύπτει ότι $R(A)^\perp = N(A^T)$. □

Πρόταση 2.7.3 (θεώρημα προβολής)**.** Αν V είναι κλειστός υπόχωρος του $\mathbb{R}^n$ τότε για κάθε x όπου $x \in \mathbb{R}^n$ υπάρχει μοναδικό v όπου $v \in V$ τέτοιο ώστε να ισχύουν τα ακόλουθα:

1. $\|x - v\|_2 = \inf(\{\|x - y\|_2 : y \in V\})$.
2. $x - v \in V^\perp$.

Απόδειξη.

1. Για κάθε y όπου $y \in V$ ισχύει ότι $\|x - y\|_2 \geqslant 0$ άρα υπάρχει το $\inf(\{\|x - y\|_2 : y \in V\}) = r \geqslant 0$ επομένως για κάθε m όπου $m \in \mathbb{N}$ υπάρχει y_m όπου $y_m \in V$ τέτοιο ώστε $r \leqslant \|x - y_m\|_2 < r + \dfrac{1}{m}$ συνεπώς ορίζεται ακολουθία (y_m) τέτοια ώστε:

$$\lim_{m\to\infty} \|x - y_m\|_2 = r \tag{1}$$

$$\|x - y_m + (x - y_l)\|_2^2 + \|x - y_m - (x - y_l)\|_2^2 = 2\|x - y_m\|_2^2 + 2\|x - y_l\|_2^2 \Leftrightarrow$$

$$\|y_m - y_l\|_2^2 = 2\|x - y_m\|_2^2 + 2\|x - y_l\|_2^2 - 4\|x - \frac{y_m + y_l}{2}\|_2^2 \tag{2}$$

$\|x - \frac{y_m + y_l}{2}\|_2 = \|\frac{1}{2}(x - y_m + x - y_l)\|_2 \leqslant \frac{1}{2}\|x - y_m\|_2 + \frac{1}{2}\|x - y_l\|_2$ άρα από την έκφραση (1) προκύπτει ότι $\lim\limits_{(m,l)\to\infty}\|x - \frac{y_m + y_l}{2}\|_2 \leqslant \frac{1}{2}r + \frac{1}{2}r \Leftrightarrow$

$$\lim_{(m,l)\to\infty}\|x - \frac{y_m + y_l}{2}\|_2 \leqslant r \tag{3}$$

Από τον ορισμό του r προκύπτει ότι $\|x - \frac{y_m + y_l}{2}\|_2 \geqslant r$ άρα:

$$\lim_{(m,l)\to\infty}\|x - \frac{y_m + y_l}{2}\|_2 \geqslant r \tag{4}$$

Από τις εκφράσεις (3),(4) προκύπτει ότι:

$$\lim_{(m,l)\to\infty}\|x - \frac{y_m + y_l}{2}\|_2 = r \tag{5}$$

Από τις εκφράσεις (1),(2),(5) προκύπτει ότι $\lim\limits_{(m,l)\to\infty}\|y_m - y_l\|_2 = 2r^2 + 2r^2 - 4r^2 = 0$ άρα η (y_m) είναι ακολουθία Cauchy στον κλειστό υπόχωρο V του $\mathbb{R}^n$ επομένως υπάρχει v όπου $v \in V$ τέτοιο ώστε:

$$\lim_{m\to\infty} y_m = v \tag{6}$$

$\|x - v\|_2 = \|x - y_m + y_m - v\|_2 \leqslant \|x - y_m\|_2 + \|y_m - v\|_2$ άρα $\lim\limits_{m\to\infty}\|x - v\|_2 \leqslant \lim\limits_{m\to\infty}\|(\|x - y_m\|_2 + \|y_m - v\|_2)$ και από τις εκφράσεις (1),(6) προκύπτει ότι:

$$\|x - v\|_2 \leqslant r \tag{7}$$

Από τον ορισμό του r προκύπτει ότι:

$$\|x - v\|_2 \geqslant r \tag{8}$$

Από τις εκφράσεις (7),(8) προκύπτει ότι $\|x - v\|_2 = r$.
$r = 0 \Leftrightarrow \|x - v\|_2 = 0 \Leftrightarrow v = x$ άρα αν $x \in V$ τότε το $v = x$ είναι μοναδικό.
Έστω $x \notin V$ και υπάρχουν $v_1 \in V, v_2 \in V$ τέτοια ώστε $r = \|x - v_1\|_2 = \|x - v_2\|_2 > 0$.
Ο V είναι χώρος Hilbert άρα $\frac{1}{2}(v_1 + v_2) \in V$ επομένως:

$$\|x - \frac{1}{2}(v_1 + v_2)\|_2 \geqslant r \tag{9}$$

$\|x - \frac{1}{2}(v_1 + v_2)\|_2 = \|\frac{1}{2}(x - v_1 + x - v_2)\|_2 = \frac{1}{2}\|x - v_1 + x - v_2\|_2 \leqslant \frac{1}{2}\|x - v_1\|_2 + \frac{1}{2}\|x - v_2\|_2 = \frac{1}{2}r + \frac{1}{2}r = r \Leftrightarrow$

$$\|x - \frac{1}{2}(v_1 + v_2)\|_2 \leqslant r \tag{10}$$

Από τις εκφράσεις (9),(10) προκύπτει ότι $\|x - \frac{1}{2}(v_1 + v_2)\|_2 = r$ άρα
$\|\frac{1}{2}(x - v_1 + x - v_2)\|_2 = \frac{1}{2}(r + r) = \frac{1}{2}(\|x - v_1\|_2 + \|x - v_2\|_2) \Leftrightarrow$
$\|x - v_1 + x - v_2\|_2 = \|x - v_1\|_2 + \|x - v_2\|_2 \Leftrightarrow \|x - v_1 + x - v_2\|_2^2 = (\|x - v_1\|_2 + \|x - v_2\|_2)^2 \Leftrightarrow$
$\langle x - v_1 + x - v_2, x - v_1 + x - v_2\rangle_2 = \|x - v_1\|_2^2 + 2\|x - v_1\|_2\|x - v_2\|_2 + \|x - v_2\|_2^2 \Leftrightarrow$
$\|x - v_1\|_2^2 + 2\langle x - v_1, x - v_2\rangle_2 + \|x - v_2\|_2^2 = \|x - v_1\|_2^2 + 2\|x - v_1\|_2\|x - v_2\|_2 + \|x - v_2\|_2^2 \Leftrightarrow$

$\langle x - v_1, x - v_2 \rangle_2 = \|x - v_1\|_2 \|x - v_2\|_2 = \langle x - v_1, x - v_1 \rangle_2^{\frac{1}{2}} \langle x - v_2, x - v_2 \rangle_2^{\frac{1}{2}}$ επομένως υπάρχει λ όπου $\lambda \in \mathbb{R}$ τέτοιο ώστε $x - v_1 = \lambda(x - v_2) \Leftrightarrow$

$$(1 - \lambda)x = v_1 - \lambda v_2 \tag{11}$$

Αν $\lambda = 1$ τότε από την έκφραση (11) προκύπτει ότι $v_1 = v_2$.
Αν $\lambda \neq 1$ τότε από την έκφραση (11) προκύπτει ότι $x = \dfrac{1}{1 - \lambda}(v_1 - v_2)$ άρα $x \in V$ το οποίο είναι άτοπο.
Συμπεραίνουμε λοιπόν ότι το v είναι μοναδικό.

2. Έστω $x - v \notin V^{\perp}$ δηλαδή υπάρχει u όπου $u \in V$ και $u \neq 0_V$ τέτοιο ώστε $\langle x - v, u \rangle_2 \neq 0$. Θέτουμε $w = v + \dfrac{\langle x - v, u \rangle_2}{\langle u, u \rangle_2} u$ άρα $w \in V$ και $\|x - w\|_2^2 = \langle x - w, x - w \rangle_2 =$ $\langle x - v - \dfrac{\langle x - v, u \rangle_2}{\langle u, u \rangle_2} u, x - v - \dfrac{\langle x - v, u \rangle_2}{\langle u, u \rangle_2} u \rangle_2 = \langle x - v, x - v \rangle_2 - 2\dfrac{\langle x - v, u \rangle_2^2}{\langle u, u \rangle_2} +$ $\dfrac{\langle x - v, u \rangle_2^2}{\langle u, u \rangle_2^2} \langle u, u \rangle_2 = \|x - v\|_2^2 - \dfrac{\langle x - v, u \rangle_2^2}{\langle u, u \rangle_2} < \|x - v\|_2^2 = r^2$ επομένως $\|x - w\|_2 < r$ το οποίο είναι άτοπο.
Συμπεραίνουμε λοιπόν ότι $\langle x - v, y \rangle_2 = 0$ για κάθε y όπου $y \in V$ δηλαδή $x - v \in V^{\perp}$.

□

Από την Πρόταση 2.7.3 προκύπτει ότι κάθε x όπου $x \in \mathbb{R}^n$ μπορεί να παρασταθεί μοναδικά ως $x = v + u$ όπου $v \in V, u \in V^{\perp}$ και ο V είναι κλειστός υπόχωρος του $\mathbb{R}^n$.
Αν $A \in \mathbb{R}^{n \times n}$ και $AA^T = A^T A = I_n$ τότε ο πίνακας A λέγεται ορθογώνιος και $A^{-1} = A^T$.

Πρόταση 2.7.4. Αν $A \in \mathbb{R}^{n \times n}, x \in \mathbb{R}^n, y \in \mathbb{R}^n$ και ο πίνακας $A = [a_1 \ \dots \ a_n]$ είναι ορθογώνιος τότε ισχύουν τα ακόλουθα:

1. $\langle a_i, a_j \rangle_2 = \delta_{i,j}$.
2. $\langle Ax, Ay \rangle_2 = \langle x, y \rangle_2$.
3. $\|Ax\|_2 = \|x\|_2$.
4. $\|A\|_2 = 1$.

Απόδειξη.

1. $A^T A = I_n \Leftrightarrow \begin{bmatrix} a_1^T \\ \vdots \\ a_n^T \end{bmatrix} [a_1 \ \dots \ a_n] = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & 1 \end{bmatrix} \Leftrightarrow$

$$\begin{bmatrix} a_1^T a_1 & \dots & a_1^T a_n \\ \vdots & \vdots & \vdots \\ a_n^T a_1 & \dots & a_n^T a_n \end{bmatrix} = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & 1 \end{bmatrix} \tag{1}$$

Από την έκφραση (1) προκύπτει ότι $a_i^T a_j = \delta_{i,j} \Leftrightarrow \langle a_i, a_j \rangle_2 = \delta_{i,j}$.

2. $\langle Ax, Ay \rangle_2 = (Ax)^T Ay = x^T A^T Ay = x^T I_n y = x^T y = \langle x, y \rangle_2$.

3. $\|Ax\|_2^2 = \langle Ax, Ax\rangle_2 = (Ax)^T Ax = x^T A^T Ax = x^T I_n x = x^T x = \langle x, x\rangle_2 = \|x\|_2^2 \Leftrightarrow \|Ax\|_2 = \|x\|_2$.

4. $\|A\|_2 = \sup(\{\frac{\|Ax\|_2}{\|x\|_2} : x \in \mathbb{R}^n \backslash \{0_n\}\}) = \sup(\{\frac{\|x\|_2}{\|x\|_2} : x \in \mathbb{R}^n \backslash \{0_n\}\}) = \sup(\{1\}) = 1$.

□

Πρόταση 2.7.5. Αν $A_i \in \mathbb{R}^{n\times n}$ και ο πίνακας A_i είναι ορθογώνιος για κάθε i όπου $i \in \{1, \ldots, m\}$ τότε ο πίνακας $A_1 \cdots A_m$ είναι ορθογώνιος για κάθε m όπου $m \in \mathbb{N}$.

Απόδειξη. Αν $m = 1$ τότε ο πίνακας A_1 είναι ορθογώνιος.
Υποθέτουμε ότι αν $m \in \{1, \ldots, l\}$ τότε ο πίνακας $A_1 \cdots A_m$ είναι ορθογώνιος δηλαδή $A_1 \cdots A_m(A_1 \cdots A_m)^T = (A_1 \cdots A_m)^T A_1 \cdots A_m = I_n$.
Αν $m = l + 1$ τότε $A_1 \cdots A_l A_{l+1}(A_1 \cdots A_l A_{l+1})^T = A_1 \cdots A_l A_{l+1} A_{l+1}^T A_l^T \cdots A_1^T = A_1 \cdots A_l I_n A_l^T \cdots A_1^T = A_1 \cdots A_l(A_1 \cdots A_l)^T = I_n \Leftrightarrow$

$$A_1 \cdots A_l A_{l+1}(A_1 \cdots A_l A_{l+1})^T = I_n \tag{1}$$

$(A_1A_2 \cdots A_{l+1})^T A_1A_2 \cdots A_{l+1} = A_{l+1}^T \cdots A_2^T A_1^T A_1 A_2 \cdots A_{l+1} = A_{l+1}^T \cdots A_2^T I_n A_2 \cdots A_{l+1} = (A_2 \cdots A_{l+1})^T A_2 \cdots A_{l+1} = I_n \Leftrightarrow$

$$(A_1A_2 \cdots A_{l+1})^T A_1A_2 \cdots A_{l+1} = I_n \tag{2}$$

Από τις εκφράσεις (1),(2) προκύπτει ότι ο πίνακας $A_1 \cdots A_{l+1}$ είναι ορθογώνιος άρα από την αρχή της μαθηματικής επαγωγής προκύπτει ότι αν $A_i \in \mathbb{R}^{n\times n}$ και ο A_i είναι ορθογώνιος για κάθε i όπου $i \in \{1, \ldots, m\}$ τότε ο $A_1 \cdots A_m$ είναι ορθογώνιος για κάθε m όπου $m \in \mathbb{N}$. □

2.8 Μεταθετικοί Πίνακες

Συμβολίζουμε ως $P_{r\leftrightarrow s,n}$ τον πίνακα που προκύπτει από τον I_n αν μεταθέσουμε την r-οστή γραμμή του I_n με την s-οστή γραμμή του I_n όπου $r \in \{1, \ldots, n\}, s \in \{1, \ldots, n\}, r \neq s$ δηλαδή:

$$P_{r\leftrightarrow s,n} = \begin{array}{c} \left[\begin{array}{ccccccc} 1 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & \ldots & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 1 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & \ldots & 0 & \ldots & 1 \end{array}\right] \end{array} \begin{array}{ll} \leftarrow & 1\text{-οστή γραμμή} \\ \vdots & \quad\vdots \\ \leftarrow & r\text{-οστή γραμμή} \\ \vdots & \quad\vdots \\ \leftarrow & s\text{-οστή γραμμή} \\ \vdots & \quad\vdots \\ \leftarrow & n\text{-οστή γραμμή} \end{array}$$

Έστω $A \in \mathbb{R}^{m\times n}$. Αν $B = P_{r\leftrightarrow s,m}A$ τότε $B(i,j) = \sum_{k=1}^{m}(P_{r\leftrightarrow s,m}(i,k)A(k,j))$ για κάθε i, j όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}$ άρα:

$$B(i,j) = \begin{cases} A(s,j), & \text{αν } i = r \\ A(r,j), & \text{αν } i = s \\ A(i,j), & \text{αν } i \in \{1, \ldots, m\} \setminus \{r, s\} \end{cases} \tag{1}$$

Από την έκφραση (1) συμπεραίνουμε ότι ο πίνακας $P_{r\leftrightarrow s,m}A$ προκύπτει από τον πίνακα A αν μεταθέσουμε την r-οστή γραμμή του A με την s-οστή γραμμή του A.

Αν $C = AP_{r\leftrightarrow s,n}$ τότε $C(i,j) = \sum_{k=1}^{n}(A(i,k)P_{r\leftrightarrow s,n}(k,j))$ για κάθε i,j όπου $i \in \{1,\dots,m\}, j \in \{1,\dots,n\}$ άρα:

$$C(i,j) = \begin{cases} A(i,s), & \text{αν } j = r \\ A(i,r), & \text{αν } j = s \\ A(i,j), & \text{αν } j \in \{1,\dots,n\} \setminus \{r,s\} \end{cases} \tag{2}$$

Από την έκφραση (2) συμπεραίνουμε ότι ο πίνακας $AP_{r\leftrightarrow s,n}$ προκύπτει από τον πίνακα A αν μεταθέσουμε την r-οστή στήλη του A με την s-οστή στήλη του A.
Για κάθε l όπου $l \in \mathbb{N}$ το γινόμενο $P_{r_1\leftrightarrow s_1,n}\cdots P_{r_l\leftrightarrow s_l,n}$ λέγεται μεταθετικός πίνακας.

Αν $D = P^2_{r\leftrightarrow s,n}$ τότε $D(i,j) = \sum_{k=1}^{n}(P_{r\leftrightarrow s,n}(i,k)P_{r\leftrightarrow s,n}(k,j))$ για κάθε i,j όπου $i \in \{1,\dots,n\}, j \in \{1,\dots,n\}$ άρα:

$$P^2_{r\leftrightarrow s,n}(i,j) = \delta_{i,j} \tag{3}$$

Από την έκφραση (3) προκύπτει ότι $P^2_{r\leftrightarrow s,n} = I_n$ άρα ο $P_{r\leftrightarrow s,n}$ είναι αντιστρέψιμος και $P^{-1}_{r\leftrightarrow s,n} = P_{r\leftrightarrow s,n} = P^T_{r\leftrightarrow s,n}$ επομένως ο $P_{r\leftrightarrow s,n}$ είναι ορθογώνιος.

2.9 Ορίζουσα

Έστω $A \in \mathbb{R}^{n\times n}$ και $A_{i,j}$ είναι ο $n-1$ επί $n-1$ πίνακας που προκύπτει από τον A αν αγνοήσουμε την i-οστή γραμμή του A και την j-οστή στήλη του A. Η ορίζουσα του A είναι η συνάρτηση $\det(\cdot) : \mathbb{R}^{n\times n} \to \mathbb{R}$ η οποία ορίζεται αναδρομικά ως εξής:

$$\det(A) = \begin{cases} A, & \text{αν } n = 1 \\ \sum_{i=1}^{n}((-1)^{i+1}a_{i,1}\det(A_{i,1})), & \text{αν } n \geqslant 2 \end{cases}$$

Πρόταση 2.9.1. Αν ο n επί n πίνακας A είναι άνω τριγωνικός τότε $\det(A) = \prod_{i=1}^{n} a_{i,i}$.

Απόδειξη. Αν $n = 1$ τότε $\det(A) = A = a_{1,1} = \prod_{i=1}^{1} a_{i,i} = \prod_{i=1}^{n} a_{i,i}$.

Υποθέτουμε ότι αν $n \in \{1,\dots,m\}$ τότε $\det(A) = \prod_{i=1}^{n} a_{i,i}$.

Αν $n = m+1$ τότε $\det(A) = \sum_{i=1}^{m+1}(a_{i,1}\det(A_{i,1})) = a_{1,1}\det(A_{1,1}) = a_{1,1}\prod_{i=2}^{m+1} a_{i,i} = \prod_{i=1}^{m+1} a_{i,i}$ άρα από την αρχή της μαθηματικής επαγωγής προκύπτει ότι αν ο n επί n πίνακας A είναι άνω τριγωνικός τότε $\det(A) = \prod_{i=1}^{n} a_{i,i}$. □

Από την Πρόταση 2.9.1 προκύπτει ότι $\det(I_n) = 1$.

Πρόταση 2.9.2. Αν $A \in \mathbb{R}^{n\times n}$ και $r \in \{1,\dots,n-1\}$ τότε $\det(P_{r\leftrightarrow r+1,n}A) = -\det(A)$.

Απόδειξη. Έστω $B \in \mathbb{R}^{n\times n}$ όπου $B = P_{r\leftrightarrow r+1,n}A$. Αν $n = 2$ τότε $r = 1$ και $\det(B) = \sum_{i=1}^{2}((-1)^{i+1}b_{i,1}\det(B_{i,1})) = b_{1,1}\det(B_{1,1}) - b_{2,1}\det(B_{2,1}) = a_{2,1}a_{1,2} - a_{1,1}a_{2,2} =$
$- (a_{1,1}a_{2,2} - a_{2,1}a_{1,2}) = -(a_{1,1}\det(A_{1,1}) - a_{2,1}\det(A_{2,1})) = -\sum_{i=1}^{2}((-1)^{i+1}a_{i,1}\det(A_{i,1})) =$
$-\det(A)$.
Υποθέτουμε ότι αν $n \in \{2,\ldots,m\}$ τότε $\det(B) = -\det(A)$.
Αν $n = m+1$ τότε $\det(B) = \sum_{i=1}^{m+1}((-1)^{i+1}b_{i,1}\det(B_{i,1}))$

$$= \sum_{i=1}^{r-1}((-1)^{i+1}b_{i,1}\det(B_{i,1})) + (-1)^{r+1}b_{r,1}\det(B_{r,1}) + (-1)^{r+2}b_{r+1,1}\det(B_{r+1,1})$$
$$+ \sum_{i=r+2}^{m+1}((-1)^{i+1}b_{i,1}\det(B_{i,1})) = \sum_{i=1}^{r-1}((-1)^{i+1}a_{i,1}(-\det(A_{i,1}))) + (-1)(-1)^{r+2}a_{r+1,1}\det(A_{r+1,1})$$
$$+ (-1)(-1)^{r+1}a_{r,1}\det(A_{r,1}) + \sum_{i=r+2}^{m+1}((-1)^{i+1}a_{i,1}(-\det(A_{i,1}))) = -\sum_{i=1}^{r-1}((-1)^{i+1}a_{i,1}\det(A_{i,1}))$$
$$- (-1)^{r+2}a_{r+1,1}\det(A_{r+1,1}) - (-1)^{r+1}a_{r,1}\det(A_{r,1}) - \sum_{i=r+2}^{m+1}((-1)^{i+1}a_{i,1}\det(A_{i,1}))$$
$$= -\left(\sum_{i=1}^{r-1}((-1)^{i+1}a_{i,1}\det(A_{i,1})) + (-1)^{r+1}a_{r,1}\det(A_{r,1})\right)$$
$$- \left((-1)^{r+2}a_{r+1,1}\det(A_{r+1,1}) + \sum_{i=r+2}^{m+1}((-1)^{i+1}a_{i,1}\det(A_{i,1}))\right)$$
$$= -\left(\sum_{i=1}^{r}((-1)^{i+1}a_{i,1}\det(A_{i,1})) + \sum_{i=r+1}^{m+1}((-1)^{i+1}a_{i,1}\det(A_{i,1}))\right)$$

$= -\sum_{i=1}^{m+1}((-1)^{i+1}a_{i,1}\det(A_{i,1})) = -\det(A)$ άρα από την αρχή της μαθηματικής επαγωγής προκύπτει ότι αν $A \in \mathbb{R}^{n\times n}$ και $r \in \{1,\ldots,n-1\}$ τότε $\det(P_{r\leftrightarrow r+1,n}A) = -\det(A)$. □

Πρόταση 2.9.3. Αν $A \in \mathbb{R}^{n\times n}, r \in \{1,\ldots,n-1\}, s \in \{1,\ldots,n-1\}, r+s \leqslant n$ τότε $\det(P_{r\leftrightarrow r+s,n}A) = -\det(A)$.

Απόδειξη. Αν $s = 1$ τότε $\det(P_{r\leftrightarrow r+s}A) = \det(P_{r\leftrightarrow r+1}A) = -\det(A)$.
Υποθέτουμε ότι αν $s \in \{1,\ldots,m\}$ και $r+s \leqslant n$ τότε $\det(P_{r\leftrightarrow r+s,n}A) = -\det(A)$.
Αν $s = m+1$ και $r+m+1 \leqslant n$ τότε $P_{r\leftrightarrow r+m+1,n}A = P_{r\leftrightarrow r+1,n}(P_{r+1\leftrightarrow m+1,n}(P_{r\leftrightarrow r+1,n}A))$ άρα $\det(P_{r\leftrightarrow r+m+1,n}A) = \det(P_{r\leftrightarrow r+1,n}(P_{r+1\leftrightarrow m+1,n}(P_{r\leftrightarrow r+1,n}A))) =$
$-\det(P_{r+1\leftrightarrow m+1,n}(P_{r\leftrightarrow r+1,n}A)) = \det(P_{r\leftrightarrow r+1,n}A) = -\det(A)$ επομένως από την αρχή της μαθηματικής επαγωγής προκύπτει ότι αν $A \in \mathbb{R}^{n\times n}, r \in \{1,\ldots,n-1\}$,
$s \in \{1,\ldots,n-1\}, r+s \leqslant n$ τότε $\det(P_{r\leftrightarrow r+s,n}A) = -\det(A)$. □

Πρόταση 2.9.4. Αν $r \in \{1,\ldots,n-1\}, s \in \{1,\ldots,n-1\}, r+s \leqslant n$ τότε $\det(P_{r\leftrightarrow r+s,n}) = -1$.

Απόδειξη. $\det(P_{r\leftrightarrow r+s,n}) = \det(P_{r\leftrightarrow r+s,n}I_n) = -\det(I_n) = -1$. □

Από τις Προτάσεις 2.9.3, 2.9.4 προκύπτει ότι κάθε μεταθετικός πίνακας έχει ορίζουσα ίση με -1 ή με 1.
Από την Γραμμική Άλγεβρα είναι γνωστή η ακόλουθη πρόταση:

Πρόταση 2.9.5. Ένας τετραγωνικός πίνακας A είναι αντιστρέψιμος αν και μόνο αν det$(A) \neq 0$.

2.10 Ιδιοτιμές - Ιδιοδιανύσματα

Οι ιδιοτιμές ενός n επί n πίνακα A είναι οι ρίζες του χαρακτηριστικού πολυωνύμου $p(t) =$ det$(A - tI_n)$ άρα κάθε n επί n πίνακας έχει n ιδιοτιμές οι οποίες μπορεί να είναι πραγματικές ή μιγαδικές.

Πρόταση 2.10.1. Αν λ είναι ιδιοτιμή του n επί n πίνακα A τότε υπάρχει x όπου $x \in \mathbb{R}^n \backslash \{0_n\}$ το οποίο λέγεται ιδιοδιάνυσμα του A που σχετίζεται με την ιδιοτιμή λ τέτοιο ώστε $Ax = \lambda x$.

Απόδειξη. Αν το 0_n είναι το μόνο διάνυσμα που ικανοποιεί την έκφραση $Ax = \lambda x$ δηλαδή $(A - \lambda I_n)x = 0_n$ αν και μόνο αν $x = 0_n$ τότε $N(A) = \{0_n\}$ άρα ο πίνακας $A - \lambda I_n$ είναι αντιστρέψιμος επομένως det$(A - \lambda I_n) \neq 0$ το οποίο είναι άτοπο συνεπώς υπάρχει x όπου $x \in \mathbb{R}^n \setminus \{0_n\}$ τέτοιο ώστε $Ax = \lambda x$. □

Το σύνολο των ιδιοτιμών ενός τετραγωνικού πίνακα A λέγεται φάσμα του A ενώ το μέγιστο των απολύτων τιμών των ιδιοτιμών του A λέγεται φασματική ακτίνα του A και συμβολίζεται ως $\rho(A)$.

Πρόταση 2.10.2. Αν $A \in \mathbb{R}^{m \times n}$ τότε $\|A\|_2 = \rho^{\frac{1}{2}}(A^T A)$.

Η απόδειξη της Πρότασης 2.10.2 παρατίθεται στο Παράρτημα.

3 ΕΛΑΧΙΣΤΑ ΤΕΤΡΑΓΩΝΑ

3.1 Γραμμικό Πρόβλημα Ελαχίστων Τετραγώνων

Έστω $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$. Το πρόβλημα προσδιορισμού ενός n επί 1 πίνακα x τέτοιου ώστε να ελαχιστοποιείται η νόρμα $\|Ax-b\|_2$ λέγεται Γραμμικό Πρόβλημα Ελαχίστων Τετραγώνων(ΓΠΕΤ). Το ΓΠΕΤ μπορεί να θεωρηθεί ως διατύπωση του προβλήματος της προσεγγιστικής επίλυσης του συστήματος $Ax=b$ με την έννοια ότι $x = \min(\{\|Av-b\|_2 : v \in \mathbb{R}^n\})$.

Πρόταση 3.1.1. Αν $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$ τότε ισχύουν τα ακόλουθα:

1. Το ΓΠΕΤ του συστήματος $Ax = b$ έχει πάντα λύση x όπου $x \in \mathbb{R}^n$ η οποία είναι μοναδική αν και μόνο αν $\text{rank}(A) = n$.

2. Η λύση ελαχίστων τετραγώνων x όπου $x \in \mathbb{R}^n$ του συστήματος $Ax = b$ ικανοποιεί το σύστημα των κανονικών εξισώσεων $A^TAx = A^Tb$ ενώ αν είναι μοναδική τότε ο A^TA είναι αντιστρέψιμος και η λύση δίνεται από τον αναλυτικό τύπο $x = (A^TA)^{-1}A^Tb$.

Απόδειξη.

1. Ο $R(A)$ είναι κλειστός υπόχωρος του $\mathbb{R}^m$ άρα υπάρχουν μοναδικά b_1, b_2 όπου $b_1 \in R(A), b_2 \in R(A)^\perp$ τέτοια ώστε $b = b_1 + b_2$ άρα αν $x \in \mathbb{R}^n$ τότε $\|Ax-b\|_2^2 = \|Ax - b_1 - b_2\|_2^2 \Leftrightarrow$
$$\|Ax-b\|_2^2 = \|Ax - b_1 + (-b_2)\|_2^2 \tag{1}$$
Έχουμε ότι $Ax - b_1 \in R(A)$ άρα $\langle Ax - b_1, b_2\rangle = 0$ επομένως από την έκφραση (1) προκύπτει ότι:
$$\|Ax-b\|_2^2 = \|Ax-b_1\|_2^2 + \|-b_2\|_2^2 \tag{2}$$
Από την έκφραση (2) προκύπτει ότι το ΓΠΕΤ του συστήματος $Ax = b$ είναι ισοδύναμο με το πρόβλημα ελαχιστοποίησης της νόρμας $\|Ax - b_1\|_2$ όπου $x \in \mathbb{R}^n$. Έχουμε ότι $b_1 \in R(A)$ άρα το σύστημα $Ax = b_1$ έχει λύση x όπου $x \in \mathbb{R}^n$ η οποία είναι λύση ελαχίστων τετραγώνων του συστήματος $Ax = b$. Η λύση του συστήματος $Ax = b_1$ είναι μοναδική αν και μόνο αν $\text{rank}(A) = n$ άρα το ΓΠΕΤ του συστήματος $Ax = b$ έχει μοναδική λύση αν και μόνο αν $\text{rank}(A) = n$.

2. Αν $x \in \mathbb{R}^n$ τέτοιο ώστε να είναι λύση ελαχίστων τετραγώνων του συστήματος $Ax = b$ τότε υπάρχουν b_1, b_2 όπου $b_1 \in R(A), b_2 \in R(A)^\perp$ τέτοια ώστε $b = b_1 + b_2$ και $b_1 = Ax$ άρα $b - Ax = b_1 + b_2 - Ax = b_1 - Ax + b_2 = 0_m + b_2 = b_2$ επομένως:
$$A^T(b - Ax) = A^Tb_2 \tag{3}$$
$\langle A^Tb_2, A^Tb_2\rangle_2 = (A^Tb_2)^TA^Tb_2 = b_2^T(A^T)^TA^Tb_2 = b_2^TA(A^Tb_2) = \langle b_2, A(A^Tb_2)\rangle_2 \overset{b_2\in R(A)^\perp}{=\!=\!=\!=} 0 \Leftrightarrow$
$$A^Tb_2 = 0_n \tag{4}$$
Από τις εκφράσεις (3),(4) προκύπτει ότι $A^T(b - Ax) = 0_n \Leftrightarrow A^Tb - A^TAx = 0_n \Leftrightarrow A^TAx = A^Tb$. Έστω $x \in \mathbb{R}^n$ τέτοιο ώστε να είναι μοναδική λύση ελαχίστων τετραγώνων του συστήματος $Ax = b$. Αν ο A^TA δεν είναι αντιστρέψιμος τότε $N(A^TA) \neq \{0_n\}$ δηλαδή υπάρχει z όπου $z \in \mathbb{R}^n \setminus \{0_n\}$ τέτοιο ώστε $A^TAz = 0_n$ άρα $\langle Az, Az\rangle_2 = (Az)^TAz = z^TA^TAz = z^T0_n = 0$ επομένως $Az = 0_m$ συνεπώς $N(A) \neq \{0_n\}$ το οποίο είναι άτοπο διότι $\text{rank}(A) = n$ οπότε ο A^TA είναι αντιστρέψιμος και $A^TAx = A^Tb \Leftrightarrow x = (A^TA)^{-1}A^Tb$.

□

Αν $A \in \mathbb{R}^{m\times n}$ και rank$(A) = n$ τότε ο A^TA είναι αντιστρέψιμος και ορίζουμε ως γενικευμένο αντίστροφο του A ή ψευδοαντίστροφο του A ή αντίστροφο Moore-Penrose του A τον n επί m πίνακα $A^+ = (A^TA)^{-1}A^T$. Αν $A \in \mathbb{R}^{n\times n}$ και rank$(A) = n$ τότε $A^+ = (A^TA)^{-1}A^T = A^{-1}(A^T)^{-1}A^T = A^{-1}I_n = A^{-1}$ και η μοναδική λύση ελαχίστων τετραγώνων x όπου $x \in \mathbb{R}^n$ του συστήματος $Ax = b$ όπου $b \in \mathbb{R}^n$ συμπίπτει με την μοναδική κλασική λύση $x = A^{-1}b$ του συστήματος $Ax = b$.

3.2 Μετασχηματισμοί Householder

Ένας m επί m πίνακας H λέγεται πίνακας Householder ή μετασχηματισμός Householder αν και μόνο αν υπάρχει u όπου $u \in \mathbb{R}^m \setminus \{0_m\}$τέτοιο ώστε $H = I_m - \dfrac{2}{u^Tu}uu^T$.

Πρόταση 3.2.1. Αν ο m επί m πίνακας H είναι μετασχηματισμός Householder τότε ισχύουν τα ακόλουθα:

1. Ο H είναι συμμετρικός.
2. Ο H είναι ορθογώνιος.

Απόδειξη.

1. $H^T = (I_m - \dfrac{2}{u^Tu}uu^T)^T = I_m^T - \dfrac{2}{u^Tu}(uu^T)^T = I_m - \dfrac{2}{u^Tu}uu^T = H$.
2. $HH^T = H^TH = H^2 = (I_m - \dfrac{2}{u^Tu}uu^T)(I_m - \dfrac{2}{u^Tu}uu^T) = I_m - \dfrac{2}{u^Tu}uu^T - \dfrac{2}{u^Tu}uu^T + \dfrac{4}{(u^Tu)^2}uu^Tuu^T = I_m - \dfrac{4}{u^Tu}uu^T + \dfrac{4}{u^Tu}uu^T = I_m$.

□

Ορίζουμε συνάρτηση $\text{sgn}_0 : \mathbb{R} \to \mathbb{R}$ τέτοια ώστε:

$$\text{sgn}_0(x) = \begin{cases} \text{sgn}(x), & \text{αν } x \in \mathbb{R} \setminus \{0\} \\ 1, & \text{αν } x = 0 \end{cases}$$

Πρόταση 3.2.2. Αν $x \in \mathbb{R}$ τότε ισχύουν τα ακόλουθα:

1. $x = \text{sgn}_0(x)|x|$.
2. $\text{sgn}_0(-x) = -\text{sgn}_0(x)$ όπου $x \neq 0$.

Απόδειξη.

1. Αν $x > 0$ τότε $x = |x| = 1|x| = \text{sgn}(x)|x| = \text{sgn}_0(x)|x|$.
   Αν $x = 0$ τότε $x = 1 \cdot |0| = \text{sgn}_0(x)|x|$.
   Αν $x < 0$ τότε $x = -|x| = -1|x| = \text{sgn}(x)|x| = \text{sgn}_0(x)|x|$.
2. Αν $x > 0$ τότε $\text{sgn}_0(-x) = \text{sgn}(-x) = -1 = -\text{sgn}(x) = -\text{sgn}_0(x)$.
   Αν $x < 0$ τότε $\text{sgn}_0(-x) = \text{sgn}(-x) = 1 = -(-1) = -\text{sgn}(x) = -\text{sgn}_0(x)$.

□

Η υλοποίηση της συνάρτησης $\text{sign}_0 : \mathbb{R} \to \mathbb{R}$ στην γλώσσα MATLAB είναι η ακόλουθη:

```
function [y]=sgn_0(x)
if x~=0
    y=sign(x);
else
    y=1;
end
end
```

Πρόταση 3.2.3. Αν $x \in \mathbb{R}^m \setminus \{0_m\}, s = \text{sgn}_0(x(1))\|x\|_2$ τότε ισχύουν τα ακόλουθα:

1. $x \neq -se_{1,m}$.

2. Αν $u = x + se_{1,m}$ τότε για τον μετασχηματισμό Householder $H = I_m - \dfrac{2}{u^T u} uu^T$ ισχύει ότι $Hx = -se_{1,m}$.

Απόδειξη.

1. Αν $x = -se_{1,m}$ τότε $\|x\|_2 = \|-se_{1,m}\|_2 = |-s|\|e_{1,m}\|_2 = |s| \neq 0$ και $x(1) = -s$ άρα $\text{sgn}_0(x(1)) = \text{sgn}_0(-s) = -\text{sgn}_0(s)$ επομένως $s = -\text{sgn}_0(s)|s| = -s \Leftrightarrow 2s = 0 \Leftrightarrow s = 0$ το οποίο είναι άτοπο συνεπώς $x \neq -se_{1,m}$.

2. $$Hx = (I_m - \frac{2}{u^T u} uu^T)x = x - \frac{2}{(x + se_{1,m})^T (x + se_{1,m})}(x + se_{1,m})(x + se_{1,m})^T x =$$
$$x - \frac{2}{(x^T + se_{1,m}^T)(x + se_{1,m})}(x + se_{1,m})(x^T + se_{1,m}^T)x =$$
$$x - \frac{2}{x^T x + sx^T e_{1,m} + se_{1,m}^T x + s^2 e_{1,m}^T e_{1,m}}(x + se_{1,m})(x^T x + se_{1,m}^T x) =$$
$$x - \frac{2}{\|x\|_2^2 + sx(1) + sx(1) + \|x\|_2^2}(x + se_{1,m})(\|x\|_2^2 + sx(1)) =$$
$$x - \frac{1}{\|x\|_2^2 + sx(1)}(\|x\|_2^2 + sx(1))(x + se_{1,m}) = x - x - se_{1,m} = -se_{1,m}.$$

□

Οι αλγόριθμοι στο κείμενο παρουσιάζονται μέσω ψευδοκώδικα εμπλουτισμένου με στοιχεία από την γλώσσα MATLAB.

Αλγόριθμος 3.2.1. Ο αλγόριθμος unit που παρουσιάζεται παρακάτω δέχεται ως είσοδο το στοιχείο x όπου $x \in \mathbb{R}^m$ και επιστρέφει ως έξοδο το στοιχείο u όπου $u \in \mathbb{R}^m$ και $u = x + \text{sgn}_0(x(1))\|x\|_2 e_{1,m}$.

algorithm unit$(x) = u$
  $m \leftarrow$ length(x)
  $s \leftarrow \|x\|_2$
  if $s = 0$ **then**
    $u \leftarrow 0_m$
    return

end if
$s \leftarrow$ sgn$_0(x(1))s$
$x(1) \leftarrow x(1) + s$
$u \leftarrow x$

end

Η υλοποίηση του αλγορίθμου unit στην γλώσσα MATLAB είναι η ακόλουθη:

```
function [u]=unit(x)
m=length(x);
s=norm(x,2);
if s==0
    u=zeros(m,1);
    return;
end
s=sgn_0(x(1))*s;
x(1)=x(1)+s;
u=x;
end
```

Αλγόριθμος 3.2.2. Ο αλγόριθμος hous που παρουσιάζεται παρακάτω δέχεται ως είσοδο το στοιχείο (x, u, a) όπου $x \in \mathbb{R}^m, u \in \mathbb{R}^m \setminus \{0_m\}, a = u^T u$ και επιστρέφει ως έξοδο το στοιχείο h όπου $h \in \mathbb{R}^m$ και $h = (I_m - \frac{2}{u^T u} u u^T)x = x - \frac{2u^T x}{u^T u} u$.

algorithm hous$(x, u, a) = h$
$m \leftarrow$ length(x)
$s \leftarrow$ sgn$_0(x(1))\|x\|_2$
$h \leftarrow 0_m$
$h(1) \leftarrow s$
if u=x+h **then**
$h \leftarrow -h$
return
end if
$s \leftarrow u^T$
$s \leftarrow -\frac{2sx}{a}$
$h \leftarrow x + su$

end

Η υλοποίηση του αλγορίθμου hous στην γλώσσα MATLAB είναι η ακόλουθη:

```
function [h]=hous(x,u,a)
m=length(x);
s=sgn_0(x(1))*norm(x,2);
h=zeros(m,1);
h(1)=s;
if u==x+h
    h=-h;
    return;
end
s=u';
```

```
s=-2*(s*x)/a;
h=x+s*u;
end
```

Πρόταση 3.2.4. Έστω $H \in \mathbb{R}^{m \times m}$ και $n+m$ επί $n+m$ πίνακας H_0 τέτοιος ώστε:

$$H_0 = \begin{bmatrix} I_n & 0_{n \times m} \\ 0_{m \times n} & H \end{bmatrix}$$

Αν ο H είναι μετασχηματισμός Householder τότε ο H_0 είναι μετασχηματισμός Householder.

Απόδειξη. Έστω $H = I_m - \frac{2}{u^T u} u u^T$ όπου $u \neq 0_m$ και $u_0 \in \mathbb{R}^{n+m}$ όπου $u_0 = [0_n^T u^T]^T$.

$I_{n+m} - \frac{2}{u_0^T u_0} u_0 u_0^T = I_{n+m} - \frac{2}{[0_n^T u^T][0_n^T u^T]^T}[0_n^T u^T]^T [0_n^T u^T] =$

$\begin{bmatrix} I_n & 0_{n \times m} \\ 0_{m \times n} & I_m \end{bmatrix} - \frac{2}{u^T u} \begin{bmatrix} 0_n \\ u \end{bmatrix} [0_n^T u^T] = \begin{bmatrix} I_n & 0_{n \times m} \\ 0_{m \times n} & I_m \end{bmatrix} - \frac{2}{u^T u} \begin{bmatrix} 0_{n \times n} & 0_{n \times m} \\ 0_{m \times n} & u u^T \end{bmatrix} =$

$\begin{bmatrix} I_n & 0_{n \times m} \\ 0_{m \times n} & I_m - \frac{2}{u^T u} u u^T \end{bmatrix} = \begin{bmatrix} I_n & 0_{n \times m} \\ 0_{m \times n} & H \end{bmatrix} \Leftrightarrow$

$$H_0 = I_{n+m} - \frac{2}{u_0^T u_0} u_0 u_0^T \tag{1}$$

Από την έκφραση (1) προκύπτει ότι ο H_0 είναι μετασχηματισμός Householder. □

3.3 Υπερκαθορισμένο ΓΠΕΤ

Το πρόβλημα προσδιορισμού λύσης ελαχίστων τετραγώνων x όπου $x \in \mathbb{R}^n$ του συστήματος $Ax = b$ όπου $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geqslant n$ λέγεται υπερκαθορισμένο ΓΠΕΤ.

Πρόταση 3.3.1 (παραγοντοποίηση QR)**.** Έστω $A \in \mathbb{R}^{m \times n}$ όπου $m \geqslant n$ και rank$(A) = n$.

1. Αν $m > n$ τότε υπάρχουν μετασχηματισμοί Householder $H_1, \ldots, H_n$ όπου $H_k \in \mathbb{R}^{m \times m}$ για κάθε k όπου $k \in \{1, \ldots, n\}$ και $R \in \mathbb{R}^{n \times n}$ όπου R άνω τριγωνικός πίνακας με μη μηδενικά διαγώνια στοιχεία τέτοιοι ώστε:

   $$H_n \cdots H_1 A = \begin{bmatrix} R \\ 0_{(m-n) \times n} \end{bmatrix}$$

   Θέτουμε $Q = H_1^T \cdots H_n^T$ άρα ο Q είναι ορθογώνιος m επί m πίνακας και $Q^T = (H_1^T \cdots H_n^T)^T = H_n \cdots H_1$ επομένως:

   $$Q^T A = \begin{bmatrix} R \\ 0_{(m-n) \times n} \end{bmatrix}$$

2. Αν $m = n$ τότε υπάρχουν μετασχηματισμοί Householder $H_1, \ldots, H_{n-1}$ όπου $H_k \in \mathbb{R}^{n \times n}$ για κάθε k όπου $k \in \{1, \ldots, n-1\}$ και $R \in \mathbb{R}^{n \times n}$ όπου R άνω τριγωνικός πίνακας με μη μηδενικά διαγώνια στοιχεία τέτοιοι ώστε:

   $$H_{n-1} \cdots H_1 A = R$$

   Θέτουμε $Q = H_1^T \cdots H_{n-1}^T$ άρα ο Q είναι ορθογώνιος n επί n πίνακας και $Q^T = (H_1^T \cdots H_{n-1}^T)^T = H_{n-1} \cdots H_1$ επομένως:

   $$Q^T A = R$$

Απόδειξη.

1. Αν $n = 1$ τότε $A = [a]$ όπου $a \in \mathbb{R}^m$ και $\text{rank}(A) = \text{rank}(a) = 1$ άρα $a \neq 0_m$ επομένως υπάρχει μετασχηματισμός Householder H_1 όπου $H_1 \in \mathbb{R}^{m\times m}$ τέτοιος ώστε $H_1 a = -\text{sgn}_0(a(1))\|a\|_2 e_{1,m}$ συνεπώς:

$$H_1 A = \begin{bmatrix} -\text{sgn}_0(a(1))\|a\|_2 \\ 0_{m-1} \end{bmatrix}$$

Υποθέτουμε ότι αν $n \in \{1, \ldots, l\}$ και $\text{rank}(A) = n$ τότε υπάρχουν μετασχηματισμοί Householder $H_1, \ldots, H_n$ όπου $H_k \in \mathbb{R}^{m\times m}$ για κάθε k όπου $k \in \{1, \ldots, n\}$ και $R \in \mathbb{R}^{n\times n}$ όπου R άνω τριγωνικός πίνακας με μη μηδενικά διαγώνια στοιχεία τέτοιοι ώστε:

$$H_n \cdots H_1 A = \begin{bmatrix} R \\ 0_{(m-n)\times n} \end{bmatrix}$$

Αν $n = l+1$ τότε $A = [A_l\ a]$ όπου $A_l \in \mathbb{R}^{m\times l}, a \in \mathbb{R}^m$ και $\text{rank}(A) = \text{rank}([A_l\ a]) = l+1$ ενώ $\text{rank}(A_l) \leqslant l$ και $\text{rank}(a) \leqslant 1$ άρα αν $\text{rank}(A_l) < l$ τότε $\text{rank}([A_l\ a]) < l+1$ το οποίο είναι άτοπο επομένως $\text{rank}(A_l) = l$ συνεπώς υπάρχουν μετασχηματισμοί Householder $H_1, \ldots, H_l$ όπου $H_k \in \mathbb{R}^{m\times m}$ για κάθε k όπου $k \in \{1, \ldots, l\}$ και $R_l \in \mathbb{R}^{l\times l}$ όπου R_l άνω τριγωνικός πίνακας με μη μηδενικά διαγώνια στοιχεία τέτοιοι ώστε:

$$H_l \cdots H_1 A_l = \begin{bmatrix} R_l \\ 0_{(m-l)\times l} \end{bmatrix} \tag{1}$$

Έχουμε ότι $H_l \cdots H_1 a \in \mathbb{R}^m$ άρα υπάρχουν μοναδικοί πίνακες a_1, a_2 όπου $a_1 \in \mathbb{R}^l, a_2 \in \mathbb{R}^{m-l}$ τέτοιοι ώστε:

$$H_l \cdots H_1 a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{2}$$

$H_l \cdots H_1 A = H_l \cdots H_1 [A_l\ a] \underset{(2)}{\overset{(1)}{\Longleftrightarrow}}$

$$H_l \cdots H_1 A = \begin{bmatrix} R_l & a_1 \\ 0_{(m-l)\times l} & a_2 \end{bmatrix} \tag{3}$$

Αν $a_2 = 0_{m-l}$ τότε $\text{rank}(H_l \cdots H_1 A) = \text{rank}(R_l) = l < l+1 = \text{rank}(A)$ το οποίο είναι άτοπο άρα $a_2 \neq 0_{m-l}$ επομένως υπάρχει μετασχηματισμός Householder H όπου $H \in \mathbb{R}^{(m-l)\times(m-l)}$ τέτοιος ώστε $Ha_2 = -\text{sgn}_0(a_2(1))\|a_2\|_2 e_{1,m-l}$. Έστω $H_0 \in \mathbb{R}^{m\times m}$ τέτοιος ώστε:

$$H_0 = \begin{bmatrix} I_l & 0_{l\times(m-l)} \\ 0_{(m-l)\times l} & H \end{bmatrix} \tag{4}$$

Από την Πρόταση 3.2.4 προκύπτει ότι ο H_0 είναι μετασχηματισμός Householder και:

$H_0 H_l \cdots H_1 A = \begin{bmatrix} I_l & 0_{l\times(m-l)} \\ 0_{(m-l)\times l} & H \end{bmatrix} \begin{bmatrix} R_l & a_1 \\ 0_{(m-l)\times l} & a_2 \end{bmatrix} =$

$\begin{bmatrix} I_l R_l + 0_{l\times(m-l)} 0_{(m-l)\times l} & I_l a_1 + 0_{l\times(m-l)} a_2 \\ 0_{(m-l)\times l} R_l + H 0_{(m-l)\times l} & 0_{(m-l)\times l} a_1 + H a_2 \end{bmatrix} = \begin{bmatrix} R_l & a_1 \\ 0_{(m-l)\times l} & Ha_2 \end{bmatrix} \Leftrightarrow$

$$H_0 H_l \cdots H_1 A = \begin{bmatrix} R_l & a_1 \\ 0 & -\text{sgn}_0(a_2(1))\|a_2\|_2 \\ 0_{(m-l-1)\times l} & 0_{m-l-1} \end{bmatrix} \tag{5}$$

Από την έκφραση (5) και την αρχή της μαθηματικής επαγωγής προκύπτει ότι αν $A \in \mathbb{R}^{m\times n}$ όπου $m > n$ και $\text{rank}(A) = n$ τότε υπάρχουν μετασχηματισμοί Householder

$H_1, \dots, H_n$ όπου $H_k \in \mathbb{R}^{m\times m}$ για κάθε k όπου $k \in \{1, \dots, n\}$ και $R \in \mathbb{R}^{n\times n}$ όπου R άνω τριγωνικός πίνακας με μη μηδενικά διαγώνια τέτοιοι ώστε:

$$H_n \cdots H_1 A = \begin{bmatrix} R \\ 0_{(m-n)\times n} \end{bmatrix}$$

2. Αν $n = 1$ τότε $A = a$ όπου $a \in \mathbb{R}$ και rank$(A) =$ rank$(a) = 1$ άρα $a \neq 0$.
Υποθέτουμε ότι αν $n \in \{1, \dots, l\}$ και rank$(A) = n$ τότε υπάρχουν μετασχηματισμοί Householder $H_1, \dots, H_{n-1}$ όπου $H_k \in \mathbb{R}^{n\times n}$ για κάθε k όπου $k \in \{1, \dots, n-1\}$ και $R \in \mathbb{R}^{n\times n}$ όπου R άνω τριγωνικός πίνακας με μη μηδενικά διαγώνια στοιχεία τέτοιοι ώστε:

$$H_{n-1} \cdots H_1 A = R$$

Αν $n = l+1$ τότε $A = [A_l\ a]$ όπου $A_l \in \mathbb{R}^{(l+1)\times l}, a \in \mathbb{R}^{l+1}$ και rank$(A) =$ rank$([A_l\ a]) = l+1$ ενώ rank$(A_l) \leqslant l$ και rank$(a) \leqslant 1$ άρα αν rank$(A_l) < l$ τότε rank$([A_l\ a]) < l+1$ το οποίο είναι άτοπο επομένως rank$(A_l) = l$ συνεπώς υπάρχουν μετασχηματισμοί Householder $H_1, \dots, H_l$ όπου $H_k \in \mathbb{R}^{(l+1)\times(l+1)}$ για κάθε k όπου $k \in \{1, \dots, l\}$ και $R_l \in \mathbb{R}^{l\times l}$ όπου R_l άνω τριγωνικός πίνακας με μη μηδενικά διαγώνια στοιχεία τέτοιοι ώστε:

$$H_l \cdots H_1 A_l = \begin{bmatrix} R_l \\ 0_l^T \end{bmatrix} \tag{6}$$

Έχουμε ότι $H_l \cdots H_1 a \in \mathbb{R}^{l+1}$ άρα υπάρχουν μοναδικοί πίνακες a_1, a_2 όπου $a_1 \in \mathbb{R}^l, a_2 \in \mathbb{R}$ τέτοιοι ώστε:

$$H_l \cdots H_1 a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{7}$$

$H_l \cdots H_1 A = H_l \cdots H_1 [A_l\ a] \overset{(6)}{\underset{(7)}{\Longleftrightarrow}}$

$$H_l \cdots H_1 A = \begin{bmatrix} R_l & a_1 \\ 0_l^T & a_2 \end{bmatrix} \tag{8}$$

Αν $a_2 = 0$ τότε rank$(H_l \cdots H_1 A) =$ rank$(R_l) = l < l+1 =$ rank(A) το οποίο είναι άτοπο άρα:

$$a_2 \neq 0 \tag{9}$$

Από τις εκφράσεις (8),(9) και την αρχή της μαθηματικής επαγωγής προκύπτει ότι αν $A \in \mathbb{R}^{n\times n}$ και rank$(A) = n$ τότε υπάρχουν μετασχηματισμοί Householder $H_1, \dots, H_{n-1}$ όπου $H_k \in \mathbb{R}^{n\times n}$ για κάθε k όπου $k \in \{1, \dots, n-1\}$ και $R \in \mathbb{R}^{n\times n}$ όπου R άνω τριγωνικός πίνακας με μη μηδενικά διαγώνια στοιχεία τέτοιοι ώστε:

$$H_{n-1} \cdots H_1 A = R$$

□

Πρόταση 3.3.2 (παραγοντοποίηση QR με οδήγηση κατά στήλες). Αν $A \in \mathbb{R}^{m\times n}$, rank$(A) = r$ όπου $1 \leqslant r \leqslant$ min$(\{m, n\}) - 1$ τότε υπάρχει m επί m ορθογώνιος πίνακας Q, n επί n μεταθετικός πίνακας P, r επί r άνω τριγωνικός με μη μηδενικά διαγώνια στοιχεία πίνακας $R_{1,1}$, r επί $n-r$ πίνακας $R_{1,2}$ τέτοιοι ώστε:

$$Q^T A P = \begin{bmatrix} R_{1,1} & R_{1,2} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix}$$

Απόδειξη. Έχουμε ότι rank$(A) = r$ άρα υπάρχει μεταθετικός n επί n πίνακας P τέτοιος ώστε $AP = [A_1\ A_2]$ όπου $A_1 \in \mathbb{R}^{m\times r}, A_2 \in \mathbb{R}^{m\times(n-r)}$ και ο A_1 έχει γραμμικώς ανεξάρτητες στήλες επομένως υπάρχει m επί m ορθογώνιος πίνακας Q και r επί r άνω τριγωνικός πίνακας με μη μηδενικά διαγώνια στοιχεία $R_{1,1}$ τέτοιοι ώστε:

$$Q^T A_1 = \begin{bmatrix} R_{1,1} \\ 0_{(m-r)\times r} \end{bmatrix} \tag{1}$$

Έχουμε ότι $Q^T A_2 \in \mathbb{R}^{m\times(n-r)}$ άρα υπάρχουν μοναδικοί πίνακες $R_{1,2}, R_{2,2}$ όπου $R_{1,2} \in \mathbb{R}^{r\times(n-r)}, R_{2,2} \in \mathbb{R}^{(m-r)\times(n-r)}$ τέτοιοι ώστε:

$$Q^T A_2 = \begin{bmatrix} R_{1,2} \\ R_{2,2} \end{bmatrix} \tag{2}$$

$Q^T AP = Q^T[A_1\ A_2] = [Q^T A_1\ Q^T A_2] \overset{(1)}{\underset{(2)}{\Longleftrightarrow}}$

$$Q^T AP = \begin{bmatrix} R_{1,1} & R_{1,2} \\ 0_{(m-r)\times r} & R_{2,2} \end{bmatrix} \tag{3}$$

Έχουμε ότι rank$(Q^T AP)$ = rank$(A) = r$ και rank$(R_{1,1}) = r$ άρα από την έκφραση (3) προκύπτει ότι $R_{2,2} = 0_{(m-r)\times(n-r)}$. □

Από την Πρόταση 3.3.2 προκύπτει ότι μέσω της παραγοντοποίησης QR με οδήγηση κατά στήλες ενός m επί n πίνακα A αποκαλύπτεται η τάξη του r η οποία είναι ίση με την διάσταση του r επί r άνω τριγωνικού με μη μηδενικά διαγώνια στοιχεία πίνακα $R_{1,1}$. Για τον μεταθετικό πίνακα P ισχύει ότι $P = P_1 \cdots P_k \cdots P_r$ όπου P_k είναι μεταθετικός πίνακας για κάθε k όπου $k \in \{1, \ldots, r\}$ ο οποίος επιδρά στον A πριν επιδράσει στον A πίνακας Householder H_k τέτοιος ώστε $Q^T = H_r \cdots H_k \cdots H_1$. Η περιγραφή του αλγορίθμου παραγοντοποίησης QR με οδήγηση κατά στήλες ενός m επί n πίνακα A με $m \geqslant n$ είναι η ακόλουθη:

Βήμα 1: Θέτουμε $l = \min(\{m, n\})$.

Βήμα 2: Θέτουμε $r = 0$.

Βήμα 3: Βρίσκουμε την στήλη με την μέγιστη νόρμα στον πίνακα A και θέτουμε την μεταβλητή a ίση με την τιμή της μέγιστης νόρμας.

Βήμα 4: **Ενόσω** $a > 0$:

- Βήμα 1: Αυξάνουμε την τιμή της μεταβλητής r κατά 1.
- Βήμα 2: **Αν** $r = 1$ **τότε**:
    - Βήμα 1: Προσδιορίζουμε n επί n μεταθετικό πίνακα P_1 τέτοιο ώστε ο m επί n πίνακας AP_1 να έχει ως πρώτη στήλη αυτή με την μέγιστη νόρμα.
    - Βήμα 2: Προσδιορίζουμε m επί m πίνακα Householder H_1 τέτοιο ώστε ο m επί n πίνακας H_1AP_1 να έχει μηδενικά στην πρώτη στήλη κάτω από το στοιχείο της πρώτης γραμμής.
    - Βήμα 3: Υπολογίζουμε τον m επί n πίνακα $A_1 = H_1AP_1$ ο οποίος έχει μηδενικά στην πρώτη στήλη κάτω από το στοιχείο της πρώτης γραμμής.

    Αλλιώς:

Βήμα 1: Προσδιορίζουμε τον $n-r+1$ επί $n-r+1$ μεταθετικό πίνακα $\bar{P}_r$ τέτοιο ώστε ο $m-r+1$ επί $n-r+1$ πίνακας $\bar{A}_{r-1}\bar{P}_r$ να έχει ως πρώτη στήλη αυτή με την μέγιστη νόρμα.

Βήμα 2: Προσδιορίζουμε $m-r+1$ επί $m-r+1$ πίνακα Householder $\bar{H}_r$ τέτοιο ώστε ο $m-r+1$ επί $n-r+1$ πίνακας $\bar{H}_r\bar{A}_{r-1}\bar{P}_r$ να έχει μηδενικά στην πρώτη στήλη κάτω από το στοιχείο της πρώτης γραμμής.

Βήμα 3: Ορίζουμε n επί n μεταθετικό πίνακα P_r τέτοιο ώστε:

$$P_r = \begin{bmatrix} I_{r-1} & 0_{(r-1)\times(n-r+1)} \\ 0_{(n-r+1)\times(r-1)} & \bar{P}_r \end{bmatrix}$$

Βήμα 4: Ορίζουμε m επί m πίνακα Householder H_r τέτοιο ώστε:

$$H_r = \begin{bmatrix} I_{r-1} & 0_{(r-1)\times(m-r+1)} \\ 0_{(m-r+1)\times(r-1)} & \bar{H}_r \end{bmatrix}$$

Βήμα 5: Υπολογίζουμε τον m επί n πίνακα $A_r = H_rA_{r-1}P_r$ ο οποίος έχει μηδενικά στην r-οστή στήλη κάτω από το στοιχείο της r-οστής γραμμής.

Βήμα 3: **Αν** $r \leqslant l-1$ **τότε**:
Βρίσκουμε την στήλη με την μέγιστη νόρμα στον $m-r$ επί $n-r$ πίνακα $\bar{A}_r$ που προκύπτει από τον A_r αν αφαιρέσουμε τις πρώτες r γραμμές και τις πρώτες r στήλες ενώ θέτουμε την μεταβλητή a ίση με την τιμή της μέγιστης νόρμας.
Αλλιώς:
Θέτουμε $a = 0$.

Βήμα 5: **Αν** $r = 0$ **τότε**:
$A = 0_{m\times n}$.
Αλλιώς:
$HAP = A_r$ όπου H, P είναι ο m επί m ορθογώνιος πίνακας, ο n επί n μεταθετικός πίνακας αντίστοιχα όπως προκύπτουν από την παραγοντοποίηση QR με οδήγηση κατά στήλες του A.

Αλγόριθμος 3.3.1. Ο αλγόριθμος qar που παρουσιάζεται παρακάτω δέχεται ως είσοδο το στοιχείο A όπου $A \in \mathbb{R}^{m\times n}, m \geqslant n$ και επιστρέφει ως έξοδο το στοιχείο (H, B, p, r) όπου B είναι m επί n πίνακας τέτοιος ώστε $HAP = B$ και H, P, p, r είναι ο m επί m ορθογώνιος πίνακας, ο n επί n μεταθετικός πίνακας, το 1 επί n διάνυσμα που αντιστοιχεί στον P και η τάξη του A αντίστοιχα όπως προκύπτουν από την παραγοντοποίηση QR με οδήγηση κατά στήλες του A.

algorithm qar$(A) = (H, B, p, r)$
  $(m, n) \leftarrow$ size(A)
  $l \leftarrow$ min$(\{m, n\})$
  $H \leftarrow I_m$
  $p \leftarrow 0_n^T$
  for $j \leftarrow 1 : n$ **do**
    $p(j) \leftarrow j$
  end for
  $r \leftarrow 0$

$a \leftarrow A(1:m, p(1):p(1))$
$a \leftarrow \|a\|_2$
for $j \leftarrow 2:n$ **do**
  $q \leftarrow A(1:m, p(j):p(j))$
  $q \leftarrow \|q\|_2$
  if $q > a$ **then**
    $a \leftarrow q$
    $q \leftarrow p(1)$
    $p(1) \leftarrow p(j)$
    $p(j) \leftarrow q$
  end if
end for
while a>0 **do**
  $r \leftarrow r+1$
  $q \leftarrow A(r:m, p(r):p(r))$
  $u \leftarrow$ unit(q)
  $a \leftarrow \|u\|_2^2$
  $q \leftarrow$ hous(q,u,a)
  $A(r:m, p(r):p(r)) \leftarrow q$
  for $j \leftarrow r+1:n$ **do**
    $q \leftarrow A(r:m, p(j):p(j))$
    $q \leftarrow$ hous(q,u,a)
    $A(r:m, p(j):p(j)) \leftarrow q$
  end for
  $u_0 \leftarrow [0_{r-1}^T \; u^T]^T$
  $q_0 \leftarrow u_0^T$
  $a \leftarrow -\frac{2}{a}$
  $H \leftarrow H + a u_0 q_0 H$
  if $r \leqslant l-1$ **then**
    $a \leftarrow A(r+1:m, p(r+1):p(r+1))$
    $a \leftarrow \|a\|_2$
    for $j \leftarrow r+2:n$ **do**
      $q \leftarrow A(r+1:m, p(j):p(j))$
      $q \leftarrow \|q\|_2$
      if $q > a$ **then**
        $a \leftarrow q$
        $q \leftarrow p(r+1)$
        $p(r+1) \leftarrow p(j)$
        $p(j) \leftarrow q$
      end if
    end for
  else
    $a \leftarrow 0$
  end if
end while
$B \leftarrow A$

end

Η υλοποίηση του αλγορίθμου qar στην γλώσσα MATLAB είναι η ακόλουθη:

```
function [H,B,p,r]=qar(A)
[m,n]=size(A);
l=min([m,n]);
H=eye(m);
p=zeros(1,n);
for j=1:n
    p(j)=j;
end
r=0;
a=A(1:m,p(1):p(1));
a=norm(a,2);
for j=2:n
    q=A(1:m,p(j):p(j));
    q=norm(q,2);
    if q>a
        a=q;
        q=p(1);
        p(1)=p(j);
        p(j)=q;
    end
end
epsilon=1;
while epsilon+1>1
    epsilon=epsilon/2;
end
epsilon=100*epsilon;
while a>epsilon
    r=r+1;
    q=A(r:m,p(r):p(r));
    u=unit(q);
    a=norm(u,2)^2;
    q=hous(q,u,a);
    A(r:m,p(r):p(r))=q;
    for j=r+1:n
        q=A(r:m,p(j):p(j));
        q=hous(q,u,a);
        A(r:m,p(j):p(j))=q;
    end
    u_0=[zeros(r-1,1);u];
    q_0=u_0';
    a=-2/a;
    H=H+a*u_0*(q_0*H);
    if r<=l-1
        a=A(r+1:m,p(r+1):p(r+1));
        a=norm(a,2);
        for j=r+2:n
            q=A(r+1:m,p(j):p(j));
            q=norm(q,2);
            if q>a
                a=q;
```

```
                    q=p(r+1);
                    p(r+1)=p(j);
                    p(j)=q;
                end
            end
        else
            a=0;
        end
end
if r<=n-1
        for j=r+1:n
            A(r+1:m,p(j):p(j))=zeros(m-r,1);
        end
end
B=A;
end
```

Στο ακόλουθο παράδειγμα χρησιμοποιούμε τα δεδομένα από το Παράδειγμα 5.3.1 του βιβλίου [1].

Παράδειγμα 3.3.1. Έστω 3 επί 2 πίνακας A τέτοιος ώστε:

$$A = \begin{bmatrix} 1 & 1 \\ 0.0001 & 0 \\ 0 & 0.0001 \end{bmatrix}$$

Το πρόγραμμα στην γλώσσα MATLAB που προσδιορίζει την παραγοντοποίηση QR του πίνακα A είναι το ακόλουθο:

```
A=[1 1;0.0001 0;0 0.0001]
[H,B,p,r]=qar(A)
```

Ο 3 επί 2 πίνακας B που επιστρέφει το παραπάνω πρόγραμμα συμφωνεί με τα αποτελέσματα του Παραδείγματος 5.3.1 του βιβλίου [1] και είναι ο ακόλουθος:

$$B = \begin{bmatrix} -1 & -1 \\ 0 & 0.0001 \\ 0 & 0 \end{bmatrix}$$

Για τον προσδιορισμό λύσης ελαχίστων τετραγώνων x όπου $x \in \mathbb{R}^n$ του συστήματος $Ax = b$ όπου $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, m \geqslant n$ ισχύουν τα ακόλουθα:

1. Αν rank$(A) = n$ και $m > n$ τότε υπάρχει m επί m ορθογώνιος πίνακας Q και n επί n άνω τριγωνικός με μη μηδενικά διαγώνια στοιχεία πίνακας R τέτοιοι ώστε:

$$Q^T A = \begin{bmatrix} R \\ 0_{(m-n)\times n} \end{bmatrix} \tag{1}$$

Έχουμε ότι $Q^T b \in \mathbb{R}^m$ άρα ο $Q^T b$ μπορεί να παρασταθεί ως σύνθετος πίνακας στήλη δύο πινάκων b_1, b_2 όπου $b_1 \in \mathbb{R}^n, b_2 \in \mathbb{R}^{m-n}$ τέτοιων ώστε:

$$Q^T b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{2}$$

$\|Ax-b\|_2^2 = \|Q^T(Ax-b)\|_2^2 = \|Q^TAx - Q^Tb\|_2^2 \overset{(1)}{\underset{(2)}{=\!=}} \left\| \begin{bmatrix} Rx \\ 0_{(m-n)\times n}x \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\|_2^2 =$
$\left\| \begin{bmatrix} Rx-b_1 \\ -b_2 \end{bmatrix} \right\|_2^2 \Leftrightarrow$

$$\|Ax-b\|_2^2 = \|Rx-b_1\|_2^2 + \|-b_2\|_2^2 \tag{3}$$

Από την έκφραση (3) συμπεραίνουμε ότι η νόρμα $\|Ax-b\|_2$ ελαχιστοποιείται όταν $Rx - b_1 = 0_n \Leftrightarrow$

$$Rx = b_1 \tag{4}$$

Η λύση x του ΓΠΕΤ προκύπτει από την επίλυση του n επί n άνω τριγωνικού συστήματος (4).

2. Αν rank$(A) = n$ και $m = n$ τότε υπάρχει n επί n ορθογώνιος πίνακας Q και n επί n άνω τριγωνικός με μη μηδενικά διαγώνια στοιχεία πίνακας R τέτοιοι ώστε:

$$Q^TA = R \tag{5}$$

$\|Ax-b\|_2 = \|Q^T(Ax-b)\|_2 = \|Q^TAx - Q^Tb\|_2 \overset{(5)}{\Longleftrightarrow}$

$$\|Ax-b\|_2 = \|Rx - Q^Tb\|_2 \tag{6}$$

Από την έκφραση (6) προκύπτει ότι η νόρμα $\|Ax-b\|_2$ ελαχιστοποιείται όταν $Rx - Q^Tb = 0_n \Leftrightarrow$

$$Rx = Q^Tb \tag{7}$$

Η λύση x του ΓΠΕΤ προκύπτει από την επίλυση του n επί n άνω τριγωνικού συστήματος (7).

3. Αν rank$(A) = r$ όπου $1 \leqslant r \leqslant n-1$ τότε υπάρχει m επί m ορθογώνιος πίνακας Q, n επί n μεταθετικός πίνακας P, r επί r άνω τριγωνικός με μη μηδενικά διαγώνια στοιχεία πίνακας $R_{1,1}$, r επί $n-r$ πίνακας $R_{1,2}$ τέτοιοι ώστε:

$$Q^TAP = \begin{bmatrix} R_{1,1} & R_{1,2} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} \tag{8}$$

Έχουμε ότι $Q^Tb \in \mathbb{R}^m$ άρα ο Q^Tb μπορεί να παρασταθεί ως σύνθετος πίνακας στήλη δύο πινάκων b_1, b_2 όπου $b_1 \in \mathbb{R}^r, b_2 \in \mathbb{R}^{m-r}$ τέτοιων ώστε:

$$Q^Tb = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{9}$$

Έχουμε ότι $P^Tx \in \mathbb{R}^n$ άρα ο P^Tx μπορεί να παρασταθεί ως σύνθετος πίνακας στήλη δύο πινάκων y_1, y_2 όπου $y_1 \in \mathbb{R}^r, y_2 \in \mathbb{R}^{n-r}$ τέτοιων ώστε:

$$P^Tx = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{10}$$

$\|Ax-b\|_2^2 = \|Q^T(AI_nx-b)\|_2^2 = \|Q^TAPP^Tx - Q^Tb\|_2^2 \overset{(8)}{\underset{(9),(10)}{=\!=\!=}}$
$\left\| \begin{bmatrix} R_{1,1} & R_{1,2} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\|_2^2 =$
$\left\| \begin{bmatrix} R_{1,1}y_1 + R_{1,2}y_2 \\ 0_{(m-r)\times r}y_1 + 0_{(m-r)\times(n-r)}y_2 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} R_{1,1}y_1 + R_{1,2}y_2 - b_1 \\ -b_2 \end{bmatrix} \right\|_2^2 \Leftrightarrow$

$$\|Ax-b\|_2^2 = \|R_{1,1}y_1 + R_{1,2}y_2 - b_1\|_2^2 + \|-b_2\|_2^2 \tag{11}$$

Από την έκφραση (11) προκύπτει ότι η νόρμα $\|Ax-b\|_2^2$ ελαχιστοποιείται όταν $R_{1,1}y_1 + R_{1,2}y_2 - b_1 = 0_r \Leftrightarrow$

$$R_{1,1}y_1 = b_1 - R_{1,2}y_2 \tag{12}$$

Από την έκφραση (12) συμπεραίνουμε ότι το ΓΠΕΤ έχει άπειρες λύσεις άρα αν θέσουμε $y_2 = 0_{n-r}$ τότε προκύπτει ότι:

$$R_{1,1}y_1 = b_1 \tag{13}$$

Από την έκφραση (10) προκύπτει ότι $P^T x = [y_1^T\ y_2^T]^T \Leftrightarrow PP^T x = P[y_1^T\ y_2^T]^T \Leftrightarrow I_n x = P[y_1^T\ y_2^T]^T \Leftrightarrow$

$$x = P[y_1^T\ y_2^T]^T \tag{14}$$

Μια λύση x του ΓΠΕΤ προκύπτει από την επίλυση του r επί r άνω τριγωνικού συστήματος (13) και αντικατάσταση της λύσης αυτής στην έκφραση (14) όπου θέτουμε $y_2 = 0_{n-r}$.

Αλγόριθμος 3.3.2. Ο αλγόριθμος over που παρουσιάζεται παρακάτω δέχεται ως είσοδο το στοιχείο (A, b) όπου $A \in \mathbb{R}^{m\times n}, m \geqslant n, b \in \mathbb{R}^m$ και επιστρέφει ως έξοδο το στοιχείο x όπου $x \in \mathbb{R}^n$ τέτοιο ώστε να είναι λύση ελαχίστων τετραγώνων του συστήματος $Ax = b$.

algorithm over$(A, b) = x$
  $(m, n) \leftarrow$ size(A)
  $y \leftarrow 0_n$
  $x \leftarrow 0_n$
  $(H, B, p, r) \leftarrow$ qar(A)
  $b \leftarrow Hb$
  for $i \leftarrow r : -1 : 1$ **do**
    $s \leftarrow 0$
    for $j \leftarrow i+1 : r$ **do**
      $s \leftarrow s + B(i, p(j))y(j)$
    end for
    $y(i) \leftarrow \dfrac{b(i) - s}{B(i, p(i))}$
  end for
  for $i \leftarrow 1 : r$ **do**
    $x(i) \leftarrow y(p(i))$
  end for
end

Η υλοποίηση του αλγορίθμου over στην γλώσσα MATLAB είναι η ακόλουθη:

```
function [x]=over(A,b)
[m,n]=size(A);
y=zeros(n,1);
x=zeros(n,1);
[H,B,p,r]=qar(A);
b=H*b;
for i=r:-1:1
    s=0;
    for j=i+1:r
        s=s+B(i,p(j))*y(j);
    end
```

```
    y(i)=(b(i)-s)/B(i,p(i));
end
for i=1:r
    x(i)=y(p(i));
end
end
```

Στο ακόλουθο παράδειγμα χρησιμοποιούμε τα δεδομένα από το Παράδειγμα 16.1 του βιβλίου [10].

Παράδειγμα 3.3.2. Έστω $A \in \mathbb{R}^{4\times 2}, b \in \mathbb{R}^4$ τέτοιοι ώστε:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 3 & 8 \\ 2 & 9 \end{bmatrix}$$

$$b = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 8 \end{bmatrix}$$

Το πρόγραμμα στην γλώσσα MATLAB που προσδιορίζει την λύση ελαχίστων τετραγώνων του συστήματος $Ax = b$ είναι το ακόλουθο:

```
A=[1 3;2 4;3 8;2 9]
b=[1;3;5;8]
x=over(A,b)
```

Ο 2 επί 1 πίνακας x που επιστρέφει το παραπάνω πρόγραμμα συμφωνεί με τα αποτελέσματα του Παραδείγματος 16.1 του βιβλίου [10] και είναι ο ακόλουθος:

$$x = \begin{bmatrix} -1.0797 \\ 1.0837 \end{bmatrix}$$

Στο ακόλουθο παράδειγμα χρησιμοποιούμε τα δεδομένα από το Παράδειγμα 16.4 του βιβλίου [10].

Παράδειγμα 3.3.3. Έστω $A \in \mathbb{R}^{6\times 3}, b \in \mathbb{R}^3$ τέτοιοι ώστε:

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

$$b = \begin{bmatrix} 2474 \\ 3882 \\ 4834 \\ 1422 \\ 2354 \\ 950 \end{bmatrix}$$

Το πρόγραμμα στην γλώσσα MATLAB που προσδιορίζει την λύση ελαχίστων τετραγώνων του συστήματος $Ax = b$ είναι το ακόλουθο:

```
A=[1 0 0;0 1 0;0 0 1;−1 1 0;−1 0 1;0 −1 1]
b=[2474;3882;4834;1422;2354;950]
x=over(A,b)
```

Ο 3 επί 1 πίνακας x που επιστρέφει το παραπάνω πρόγραμμα συμφωνεί με τα αποτελέσματα του Παραδείγματος 16.4 του βιβλίου [10] και είναι ο ακόλουθος:

$$x = \begin{bmatrix} 24720 \\ 38860 \\ 48320 \end{bmatrix}$$

Στο ακόλουθο παράδειγμα χρησιμοποιούμε τα δεδομένα από το Παράδειγμα 8.5 του βιβλίου [8].

Παράδειγμα 3.3.4. Έστω $A \in \mathbb{R}^{3\times 2}, b \in \mathbb{R}^3$ τέτοιοι ώστε:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 5 \end{bmatrix}$$

$$b = \begin{bmatrix} 3 \\ 5 \\ 9 \end{bmatrix}$$

Το πρόγραμμα στην γλώσσα MATLAB που προσδιορίζει την λύση ελαχίστων τετραγώνων του συστήματος $Ax = b$ είναι το ακόλουθο:

```
A=[1 2;2 3;4 5]
b=[3;5;9]
x=over(A,b)
```

Ο 2 επί 1 πίνακας x που επιστρέφει το παραπάνω πρόγραμμα συμφωνεί με τα αποτελέσματα του Παραδείγματος 8.5 του βιβλίου [8] και είναι ο ακόλουθος:

$$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

3.4 Υποκαθορισμένο ΓΠΕΤ

Το πρόβλημα προσδιορισμού λύσης ελαχίστων τετραγώνων x όπου $x \in \mathbb{R}^n$ του συστήματος $Ax = b$ όπου $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, m < n$ λέγεται υποκαθορισμένο ΓΠΕΤ και ισχύουν τα ακόλουθα:

1. Αν rank$(A) = m$ τότε υπάρχει n επί n ορθογώνιος πίνακας Q και m επί m άνω τριγωνικός με μη μηδενικά διαγώνια στοιχεία πίνακας R τέτοιοι ώστε:
$$Q^T A^T = \begin{bmatrix} R \\ 0_{(n-m)\times m} \end{bmatrix} \Leftrightarrow A^T = Q \begin{bmatrix} R \\ 0_{(n-m)\times m} \end{bmatrix} \Leftrightarrow (A^T)^T = (Q \begin{bmatrix} R \\ 0_{(n-m)\times m} \end{bmatrix})^T \Leftrightarrow$$
$$A = [R^T \; 0_{m\times(n-m)}]Q^T \tag{1}$$
Έχουμε ότι $Q^T x \in \mathbb{R}^n$ άρα ο $Q^T x$ μπορεί να παρασταθεί ως σύνθετος πίνακας στήλη δύο πινάκων y_1, y_2 όπου $y_1 \in \mathbb{R}^m, y_2 \in \mathbb{R}^{n-m}$ τέτοιων ώστε:
$$Q^T x = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{2}$$

$\|Ax-b\|_2 \overset{(1)}{=\!=} \|[R^T\ 0_{m\times(n-m)}]Q^Tx-b\|_2 \overset{(2)}{=\!=} \left\|[R^T\ 0_{m\times(n-m)}]\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - b\right\|_2 =$
$\|R^Ty_1+0_{m\times(n-m)}y_2-b\|_2 \Leftrightarrow$

$$\|Ax-b\|_2 = \|R^Ty_1-b\|_2 \tag{3}$$

Από την έκφραση (3) προκύπτει ότι μπορούμε να θεωρήσουμε το y_2 αυθαίρετο και η νόρμα $\|Ax-b\|_2$ ελαχιστοποιείται όταν $R^Ty_1-b=0_m \Leftrightarrow$

$$R^Ty_1=b \tag{4}$$

Από την έκφραση (2) προκύπτει ότι $Q^Tx=[y_1^T\ y_2^T]^T \Leftrightarrow QQ^Tx=Qy \Leftrightarrow I_nx = Q[y_1^T\ y_2^T]^T \Leftrightarrow$

$$x=Q[y_1^T\ y_2^T]^T \tag{5}$$

Μια λύση x του ΓΠΕΤ προκύπτει από την επίλυση του m επί m κάτω τριγωνικού συστήματος (4) και αντικατάσταση της λύσης αυτής στην έκφραση (5) όπου θέτουμε $y_2=0_{n-m}$.

2. Αν rank$(A)=r$ όπου $1\leqslant r\leqslant m-1$ τότε ισχύουν τα ακόλουθα:

(α΄) Υπάρχει n επί n ορθογώνιος πίνακας Q, m επί m μεταθετικός πίνακας P, r επί r άνω τριγωνικός με μη μηδενικά διαγώνια στοιχεία πίνακας $R_{1,1}$, r επί $m-r$ πίνακας $R_{1,2}$ τέτοιοι ώστε:
$Q^TA^TP=\begin{bmatrix} R_{1,1} & R_{1,2} \\ 0_{(n-r)\times r} & 0_{(n-r)\times(m-r)} \end{bmatrix} \Leftrightarrow A^T=Q\begin{bmatrix} R_{1,1} & R_{1,2} \\ 0_{(n-r)\times r} & 0_{(n-r)\times(m-r)} \end{bmatrix}P^T \Leftrightarrow$
$(A^T)^T=(Q\begin{bmatrix} R_{1,1} & R_{1,2} \\ 0_{(n-r)\times r} & 0_{(n-r)\times(m-r)} \end{bmatrix}P^T)^T \Leftrightarrow$

$$A=P\begin{bmatrix} R_{1,1}^T & 0_{r\times(n-r)} \\ R_{1,2}^T & 0_{(m-r)\times(n-r)} \end{bmatrix}Q^T \tag{6}$$

Έχουμε ότι $Q^Tx\in\mathbb{R}^n$ άρα ο Q^Tx μπορεί να παρασταθεί ως σύνθετος πίνακας στήλη δύο πινάκων y_1,y_2 όπου $y_1\in\mathbb{R}^r, y_2\in\mathbb{R}^{n-r}$ τέτοιων ώστε:

$$Q^Tx=\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{7}$$

Έχουμε ότι $P^Tb\in\mathbb{R}^m$ άρα ο P^Tb μπορεί να παρασταθεί ως σύνθετος πίνακας στήλη δύο πινάκων b_1,b_2 όπου $b_1\in\mathbb{R}^r, b_2\in\mathbb{R}^{m-r}$ τέτοιων ώστε:

$$P^Tb=\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{8}$$

$\|Ax-b\|_2^2 \overset{(6)}{=\!=} \left\|P\begin{bmatrix} R_{1,1}^T & 0_{r\times(n-r)} \\ R_{1,2}^T & 0_{(m-r)\times(n-r)} \end{bmatrix}Q^Tx-b\right\|_2^2 =$
$\left\|P^T(P\begin{bmatrix} R_{1,1}^T & 0_{r\times(n-r)} \\ R_{1,2}^T & 0_{(m-r)\times(n-r)} \end{bmatrix}Q^Tx-b)\right\|_2^2 =$
$\left\|P^TP\begin{bmatrix} R_{1,1}^T & 0_{r\times(n-r)} \\ R_{1,2}^T & 0_{(m-r)\times(n-r)} \end{bmatrix}Q^Tx-P^Tb\right\|_2^2 =$
$\left\|I_m\begin{bmatrix} R_{1,1}^T & 0_{r\times(n-r)} \\ R_{1,2}^T & 0_{(m-r)\times(n-r)} \end{bmatrix}Q^Tx-P^Tb\right\|_2^2 =$

$$\left\| \begin{bmatrix} R_{1,1}^T & 0_{r\times(n-r)} \\ R_{1,2}^T & 0_{(m-r)\times(n-r)} \end{bmatrix} Q^T x - P^T b \right\|_2^2 \overset{(7)}{\underset{(8)}{=\!=}}$$
$$\left\| \begin{bmatrix} R_{1,1}^T & 0_{r\times(n-r)} \\ R_{1,2}^T & 0_{(m-r)\times(n-r)} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\|_2^2 =$$
$$\left\| \begin{bmatrix} R_{1,1}^T y_1 + 0_{r\times(n-r)} y_2 \\ R_{1,2}^T y_1 + 0_{(m-r)\times(n-r)} y_2 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} R_{1,1}^T y_1 - b_1 \\ R_{1,2}^T y_1 - b_2 \end{bmatrix} \right\|_2^2 \Leftrightarrow$$
$$\|Ax - b\|_2^2 = \|R_{1,1}^T y_1 - b_1\|_2^2 + \|R_{1,2}^T y_1 - b_2\|_2^2 \tag{9}$$

Από την έκφραση (9) προκύπτει ότι μπορούμε να θεωρήσουμε το y_2 αυθαίρετο και η νόρμα $\|Ax-b\|_2^2$ ελαχιστοποιείται όταν ελαχιστοποιείται η ποσότητα $\|R_{1,1}^T y_1 - b_1\|_2^2 + \|R_{1,2}^T y_1 - b_2\|_2^2$. Ορίζουμε $f : \mathbb{R}^r \to \mathbb{R}$ όπου $f(v) = \|R_{1,1}^T v - b_1\|_2^2 + \|R_{1,2}^T v - b_2\|_2^2$ για κάθε v όπου $v \in \mathbb{R}^r$ και θέτουμε $v = (v_1, \dots, v_r)$ άρα $f(v) =$
$$f(v_1,\dots,v_r) = \sum_{i=1}^{r}\left(\sum_{j=1}^{r}(R_{1,1}^T(i,j)v_j) - b_1(i)\right)^2 + \sum_{i=1}^{m-r}\left(\sum_{j=1}^{r}(R_{1,2}^T(i,j)v_j) - b_2(i)\right)^2.$$
Το ΓΠΕΤ έχει πάντα λύση άρα η f παρουσιάζει ελάχιστο σε σημείο του $\mathbb{R}^r$ και η f είναι διαφορίσιμη στο $\mathbb{R}^r$ επομένως η f ελαχιστοποιείται όταν $\nabla f(v) = 0_r$ δηλαδή όταν για κάθε k όπου $k \in \{1,\dots,r\}$ ισχύει ότι $\dfrac{\partial f}{\partial v_k}(v) = 0 \Leftrightarrow$
$$2\sum_{i=1}^{r}\left(R_{1,1}^T(i,k)\left(\sum_{j=1}^{r}(R_{1,1}^T(i,j)v_j) - b_1(i)\right)\right) +$$
$$2\sum_{i=1}^{m-r}\left(R_{1,2}^T(i,k)\left(\sum_{j=1}^{r}(R_{1,2}^T(i,j)v_j) - b_2(i)\right)\right) = 0 \Leftrightarrow$$
$$\sum_{i=1}^{r}\left(R_{1,1}^T(i,k)\sum_{j=1}^{r}(R_{1,1}^T(i,j)v_j) - R_{1,1}^T(i,k)b_1(i)\right) +$$
$$\sum_{i=1}^{m-r}\left(R_{1,2}^T(i,k)\sum_{j=1}^{r}(R_{1,2}^T(i,j)v_j) - R_{1,2}^T(i,k)b_2(i)\right) = 0 \Leftrightarrow$$
$$\sum_{i=1}^{r}\left(R_{1,1}^T(i,k)\sum_{j=1}^{r}(R_{1,1}^T(i,j)v_j)\right) - \sum_{i=1}^{r}(R_{1,1}^T(i,k)b_1(i)) +$$
$$\sum_{i=1}^{m-r}\left(R_{1,2}^T(i,k)\sum_{j=1}^{r}(R_{1,2}^T(i,j)v_j)\right) - \sum_{i=1}^{m-r}(R_{1,2}^T(i,k)b_2(i)) = 0 \Leftrightarrow$$
$$\sum_{i=1}^{r}\left(R_{1,1}^T(i,k)\sum_{j=1}^{r}(R_{1,1}^T(i,j)v_j)\right) + \sum_{i=1}^{m-r}\left(R_{1,2}^T(i,k)\sum_{j=1}^{r}(R_{1,2}^T(i,j)v_j)\right) =$$
$$\sum_{i=1}^{r}(R_{1,1}^T(i,k)b_1(i)) + \sum_{i=1}^{m-r}(R_{1,2}^T(i,k)b_2(i)) \Leftrightarrow$$
$$\sum_{i=1}^{r}\left(R_{1,1}(k,i)\sum_{j=1}^{r}(R_{1,1}^T(i,j)v_j)\right) + \sum_{i=1}^{m-r}\left(R_{1,2}(k,i)\sum_{j=1}^{r}(R_{1,2}^T(i,j)v_j)\right) =$$
$\displaystyle\sum_{i=1}^{r}(R_{1,1}(k,i)b_1(i)) + \sum_{i=1}^{m-r}(R_{1,2}(k,i)b_2(i))$ και η τελευταία εξίσωση ισχύει για κάθε k όπου $k \in \{1,\dots,r\}$ συνεπώς $R_{1,1}R_{1,1}^T v + R_{1,2}R_{1,2}^T v = R_{1,1}b_1 + R_{1,2}b_2$ οπότε:
$$(R_{1,1}R_{1,1}^T + R_{1,2}R_{1,2}^T)v = R_{1,1}b_1 + R_{1,2}b_2 \tag{10}$$
Το ΓΠΕΤ έχει πάντα λύση άρα το σύστημα (10) έχει λύση v_0 όπου $v_0 \in \mathbb{R}^r$. Ορίζουμε $g_1 : \mathbb{R}^r \to \mathbb{R}^r, g_2 : \mathbb{R}^r \to \mathbb{R}^{m-r}$ τέτοιες ώστε $g_1(v) = R_{1,1}^T v - b_1, g_2(v) =$

$R_{1,2}^T v - b_2$ για κάθε v όπου $v \in \mathbb{R}^r$ άρα $f(v) = \|g_1(v)\|_2^2 + \|g_2(v)\|_2^2 =$
$\|R_{1,1}^T v - b_1\|_2^2 + \|R_{1,2}^T v - b_2\|_2^2 =$
$\|R_{1,1}^T v - R_{1,1}^T v_0 + R_{1,1}^T v_0 - b_1\|_2^2 + \|R_{1,2}^T v - R_{1,2}^T v_0 + R_{1,2}^T v_0 - b_2\|_2^2 =$
$\|R_{1,1}^T(v - v_0) + g_1(v_0)\|_2^2 + \|R_{1,2}^T(v - v_0) + g_2(v_0)\|_2^2 =$
$\langle R_{1,1}^T(v - v_0) + g_1(v_0), R_{1,1}^T(v - v_0) + g_1(v_0)\rangle_2 +$
$\langle R_{1,2}^T(v - v_0) + g_2(v_0), R_{1,2}^T(v - v_0) + g_2(v_0)\rangle_2 =$
$(R_{1,1}^T(v - v_0) + g_1(v_0))^T(R_{1,1}^T(v - v_0) + g_1(v_0)) +$
$(R_{1,2}^T(v - v_0) + g_2(v_0))^T(R_{1,2}^T(v - v_0) + g_2(v_0)) =$
$(g_1^T(v_0) + (v - v_0)^T R_{1,1})(R_{1,1}^T(v - v_0) + g_1(v_0)) +$
$(g_2^T(v_0) + (v - v_0)^T R_{1,2})(R_{1,2}^T(v - v_0) + g_2(v_0)) =$
$g_1^T(v_0)R_{1,1}^T(v-v_0) + g_1^T(v_0)g_1(v_0) + (v-v_0)^T R_{1,1}R_{1,1}^T(v-v_0) + (v-v_0)^T R_{1,1}g_1(v_0) +$
$g_2^T(v_0)R_{1,2}^T(v-v_0) + g_2^T(v_0)g_2(v_0) + (v-v_0)^T R_{1,2}R_{1,2}^T(v-v_0) + (v-v_0)^T R_{1,2}g_2(v_0) =$
$(R_{1,1}g_1(v_0))^T(v - v_0) + \|g_1(v_0)\|_2^2 + \|R_{1,1}^T(v - v_0)\|_2^2 + (v - v_0)^T R_{1,1}g_1(v_0) +$
$(R_{1,2}g_2(v_0))^T(v - v_0) + \|g_2(v_0)\|_2^2 + \|R_{1,2}^T(v - v_0)\|_2^2 + (v - v_0)^T R_{1,2}g_2(v_0) =$
$(R_{1,1}g_1(v_0) + R_{1,2}g_2(v_0))^T(v - v_0) + (v - v_0)^T(R_{1,1}g_1(v_0) + R_{1,2}g_2(v_0)) +$
$\|g_1(v_0)\|_2^2 + \|g_2(v_0)\|_2^2 + \|R_{1,1}^T(v - v_0)\|_2^2 + \|R_{1,1}^T(v - v_0)\|_2^2 \geqslant$
$(R_{1,1}(R_{1,1}^T v_0 - b_1) + R_{1,2}(R_{1,2}^T v_0 - b_2))^T(v - v_0) +$
$(v - v_0)^T(R_{1,1}(R_{1,1}^T v_0 - b_1) + R_{1,2}(R_{1,2}^T v_0 - b_2)) + f(v_0) =$
$(R_{1,1}R_{1,1}^T v_0 - R_{1,1}b_1 + R_{1,2}R_{1,2}^T v_0 - R_{1,2}b_2)^T(v - v_0) +$
$(v - v_0)^T(R_{1,1}R_{1,1}^T v_0 - R_{1,1}b_1 + R_{1,2}R_{1,2}^T v_0 - R_{1,2}b_2) + f(v_0) =$
$((R_{1,1}R_{1,1}^T + R_{1,2}R_{1,2}^T)v_0 - (R_{1,1}b_1 + R_{1,2}b_2))^T(v - v_0) +$
$(v - v_0)^T((R_{1,1}R_{1,1}^T + R_{1,2}R_{1,2}^T)v_0 - (R_{1,1}b_1 + R_{1,2}b_2)) + f(v_0) \overset{(10)}{=\!=}$
$0_r^T(v - v_0) + (v - v_0)^T 0_r + f(v_0) = f(v_0)$ επομένως:

$$f(v) \geqslant f(v_0) \tag{11}$$

Από την έκφραση (11) προκύπτει ότι κάθε λύση v_0 του συστήματος (10) ελαχιστοποιεί την f ενώ από την έκφραση (7) προκύπτει ότι $Q^T x = [y_1^T\ y_2^T]^T \Leftrightarrow QQ^T x = Q[y_1^T\ y_2^T]^T \Leftrightarrow I_n x = Q[y_1^T\ y_2^T]^T \Leftrightarrow$

$$x = Q[y_1^T\ y_2^T]^T \tag{12}$$

Μια λύση x του ΓΠΕΤ προκύπτει αν θέσουμε $y_1 = v_0, y_2 = 0_{n-r}$ και αντικαταστήσουμε στην έκφραση (12).

(β΄) Υπάρχει m επί m ορθογώνιος πίνακας Q, n επί n μεταθετικός πίνακας P, r επί r άνω τριγωνικός με μη μηδενικά διαγώνια στοιχεία πίνακας $R_{1,1}$, r επί $n - r$ πίνακας $R_{1,2}$ τέτοιοι ώστε:

$$Q^T A P = \begin{bmatrix} R_{1,1} & R_{1,2} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} \tag{13}$$

Έχουμε ότι $Q^T b \in \mathbb{R}^m$ άρα ο $Q^T b$ μπορεί να παρασταθεί ως σύνθετος πίνακας στήλη δύο πινάκων b_1, b_2 όπου $b_1 \in \mathbb{R}^r, b_2 \in \mathbb{R}^{m-r}$ τέτοιων ώστε:

$$Q^T b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{14}$$

Έχουμε ότι $P^T x \in \mathbb{R}^n$ άρα ο $P^T x$ μπορεί να παρασταθεί ως σύνθετος πίνακας στήλη δύο πινάκων y_1, y_2 όπου $y_1 \in \mathbb{R}^r, y_2 \in \mathbb{R}^{n-r}$ τέτοιων ώστε:

$$P^T x = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{15}$$

$$\|Ax-b\|_2^2 = \|Q^T(AI_nx-b)\|_2^2 = \|Q^TAPP^Tx-Q^Tb\|_2^2 \overset{(13)}{\underset{(14),(15)}{=\!=}}$$
$$\left\| \begin{bmatrix} R_{1,1} & R_{1,2} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\|_2^2 =$$
$$\left\| \begin{bmatrix} R_{1,1}y_1 + R_{1,2}y_2 \\ 0_{(m-r)\times r}y_1 + 0_{(m-r)\times(n-r)}y_2 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} R_{1,1}y_1 + R_{1,2}y_2 - b_1 \\ -b_2 \end{bmatrix} \right\|_2^2 \Leftrightarrow$$
$$\|Ax-b\|_2^2 = \|R_{1,1}y_1 + R_{1,2}y_2 - b_1\|_2^2 + \|-b_2\|_2^2 \tag{16}$$

Από την έκφραση (16) προκύπτει ότι η νόρμα $\|Ax-b\|_2^2$ ελαχιστοποιείται όταν $R_{1,1}y_1 + R_{1,2}y_2 - b_1 = 0_r \Leftrightarrow$

$$R_{1,1}y_1 = b_1 - R_{1,2}y_2 \tag{17}$$

Από την έκφραση (17) συμπεραίνουμε ότι το ΓΠΕΤ έχει άπειρες λύσεις άρα αν θέσουμε $y_2 = 0_{n-r}$ τότε προκύπτει ότι:

$$R_{1,1}y_1 = b_1 \tag{18}$$

Από την έκφραση (15) προκύπτει ότι $P^Tx = [y_1^T \ y_2^T]^T \Leftrightarrow PP^Tx = P[y_1^T \ y_2^T]^T \Leftrightarrow I_nx = P[y_1^T \ y_2^T]^T \Leftrightarrow$

$$x = P[y_1^T \ y_2^T]^T \tag{19}$$

Μια λύση x του ΓΠΕΤ προκύπτει από την επίλυση του r επί r άνω τριγωνικού συστήματος (18) και αντικατάσταση της λύσης αυτής στην έκφραση (19) όπου θέτουμε $y_2 = 0_{n-r}$.

Αλγόριθμος 3.4.1. Ο αλγόριθμος under που παρουσιάζεται παρακάτω δέχεται ως είσοδο το στοιχείο (A, b) όπου $A \in \mathbb{R}^{m\times n}, m < n, b \in \mathbb{R}^m$ και επιστρέφει ως έξοδο το στοιχείο x όπου $x \in \mathbb{R}^n$ τέτοιο ώστε να είναι λύση ελαχίστων τετραγώνων του συστήματος $Ax = b$.

algorithm under$(A, b) = x$
  $(m, n) \leftarrow$ size(A)
  $y \leftarrow 0_n$
  $x \leftarrow 0_n$
  $(H, B, p, r) \leftarrow$ qar(A^T)
  if r=m **then**
    $B \leftarrow B^T$
    for $i \leftarrow 1 : m$ **do**
      $s \leftarrow 0$
      for $j \leftarrow 1 : i-1$ **do**
        $s \leftarrow s + B(p(i), j)y(j)$
      end for
      $y(i) \leftarrow \dfrac{b(p(i)) - s}{B(p(i), i)}$
    end for
    $H \leftarrow H^T$
    $x \leftarrow Hy$
  else
    $x \leftarrow$ over(A, b)
  end if
end

Η υλοποίηση του αλγορίθμου under στην γλώσσα MATLAB είναι η ακόλουθη:

```
function [x]=under(A,b)
[m,n]=size(A);
y=zeros(n,1);
x=zeros(n,1);
[H,B,p,r]=qar(A');
if r==m
    B=B';
    for i=1:m
        s=0;
        for j=1:i-1
            s=s+B(p(i),j)*y(j);
        end
        y(i)=(b(p(i))-s)/B(p(i),i);
    end
    H=H';
    x=H*y;
else
    x=over(A,b);
end
end
```

Στο ακόλουθο παράδειγμα χρησιμοποιούμε τα δεδομένα από το Παράδειγμα 16.11 του βιβλίου [10].

Παράδειγμα 3.4.1. Έστω $A \in \mathbb{R}_{3\times 5}, b \in \mathbb{R}_{3\times 1}$ τέτοιοι ώστε:

$$A = \begin{bmatrix} 1 & 3 & 5 & 7 & 9 \\ -1 & -2 & -3 & -4 & -5 \\ 6 & 12 & 8 & 9 & 10 \end{bmatrix}$$

$$b = \begin{bmatrix} 1 \\ 5 \\ 8 \end{bmatrix}$$

Το πρόγραμμα στην γλώσσα MATLAB που προσδιορίζει την λύση ελαχίστων τετραγώνων του συστήματος $Ax = b$ είναι το ακόλουθο:

```
A=[1 3 5 7 9;-1 -2 -3 -4 -5;6 12 8 9 10]
b=[1;5;8]
x=under(A,b)
```

Ο 5 επί 1 πίνακας x που επιστρέφει το παραπάνω πρόγραμμα συμφωνεί με τα αποτελέσματα του Παραδείγματος 16.11 του βιβλίου [10] και είναι ο ακόλουθος:

$$x = \begin{bmatrix} -18.4286 \\ 13.6 \\ -7.5143 \\ -2.0571 \\ 3.4 \end{bmatrix}$$

Αλγόριθμος 3.4.2. Ο αλγόριθμος leasq που παρουσιάζεται παρακάτω δέχεται ως είσοδο το στοιχείο (A, b) όπου $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$ και επιστρέφει ως έξοδο το στοιχείο x όπου $x \in \mathbb{R}^n$ τέτοιο ώστε να είναι λύση ελαχίστων τετραγώνων του συστήματος $Ax = b$.

algorithm leasq$(A, b) = x$

$(m, n) \leftarrow$ size(A)
$l \leftarrow$ length(b)
if $(m = 0) \vee (n = 0) \vee (l = 0) \vee (l \neq m)$ **then**
print("Invalid input.")
return
end if
if $m \geqslant n$ **then**
$x \leftarrow$ over(A, b)
else
$x \leftarrow$ under(A, b)
end if

end

Η υλοποίηση του αλγορίθμου leasq στην γλώσσα MATLAB είναι η ακόλουθη:

```
function [x]=leasq(A,b)
[m,n]=size(A);
l=length(b);
if (m==0)||(n==0)||(l==0)||(l~=m)
     fprintf("Invalid input.\n");
     return;
end
if m>=n
     x=over(A,b);
else
     x=under(A,b);
end
end
```

3.5 Εφαρμογές

Έστω άγνωστη συνάρτηση f της οποίας είναι γνωστές οι τιμές $f_i = f(t_i)$ στα σημεία t_i όπου $i \in \{1, \ldots, m\}$ και την οποία θέλουμε να προσεγγίσουμε από πολυώνυμο $p(t) = \sum_{j=1}^{n+1}(x_{j-1}t^{j-1})$ βαθμού το πολύ n(συμβολίζουμε $p \in \mathbb{P}_n$) όπου $x_j \in \mathbb{R}$ για κάθε j όπου $j \in \{0, \ldots, n\}$ τέτοιο ώστε να παρεμβάλλεται στα σημεία (t_i, f_i) και να ελαχιστοποιείται το άθροισμα $\sum_{i=1}^{m}(p(t_i) - f_i)^2 = \sum_{i=1}^{m}\left(\sum_{j=1}^{n+1}(x_{j-1}t_i^{j-1}) - f_i\right)^2$ δηλαδή θέλουμε να προσδιορίσουμε $x = (x_0, \ldots, x_n)$ τέτοιο ώστε η συνάρτηση $r : \mathbb{R}^{n+1} \to [0, \infty]$ όπου $r(x) = r(x_0, \ldots, x_n) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n+1}(t_i^{j-1}x_{j-1}) - f_i\right)^2$ να ελαχιστοποιείται. Ορίζουμε m επί $n+1$ πίνακα A τέτοιο ώστε $A(i, j) = t_i^{j-1}$ για κάθε i, j όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, n+1\}$ και m επί 1 πίνακα b τέτοιο ώστε $b(i) = f_i$ για κάθε i όπου $i \in \{1, \ldots, m\}$ άρα το πρόβλημα ελαχιστοποίησης της συνάρτησης r είναι ισοδύναμο με το πρόβλημα προσδιορισμού λύσης ελαχίστων τετραγώνων του συστήματος $Ax = b$.

Αλγόριθμος 3.5.1. Ο αλγόριθμος inter που παρουσιάζεται παρακάτω δέχεται ως είσοδο το στοιχείο (n, v, y) όπου $n \in \mathbb{N}, v \in \mathbb{R}^m, y \in \mathbb{R}^m$ τέτοια ώστε n είναι ο μέγιστος βαθμός του πολυωνύμου $p(t) = \sum_{j=1}^{n+1}(x_{j-1}t^{j-1})$ που παρεμβάλλεται στα σημεία $(t_i, f(t_i))$ της συνάρτησης $f : \mathbb{R} \to \mathbb{R}$ όπου $i \in \{1, \dots, m\}$ και $v(i) = t_i, y(i) = f(t_i)$ για κάθε i όπου $i \in \{1, \dots, m\}$ ενώ επιστρέφει ως έξοδο το στοιχείο x όπου $x \in \mathbb{R}^{n+1}$ τέτοιο ώστε $x = (x_0, \dots, x_n)$.

algorithm inter$(n, v, y) = x$
  $m \leftarrow$ length(v)
  if $(m = 0) \vee (n \leqslant 0)$ **then**
    print("Invalid input.")
    return
  end if
  $A \leftarrow 0_{m\times(n+1)}$
  for $i \leftarrow 1 : m$ **do**
    for $j \leftarrow 1 : n + 1$ **do**
      $A(i, j) \leftarrow v(i)^{j-1}$
    end for
  end for
  $x \leftarrow$ leasq(A, y)

end

Η υλοποίηση του αλγορίθμου inter στην γλώσσα MATLAB είναι η ακόλουθη:

```
function [x]=inter(n,v,y)
m=length(v);
if (m==0)||(n<=0)
    fprintf("Invalid input.\n");
    return;
end
A=zeros(m,n+1);
for i=1:m
    for j=1:n+1
        A(i,j)=v(i)^(j-1);
    end
end
x=leasq(A,y);
end
```

Στο ακόλουθο παράδειγμα χρησιμοποιούμε τα δεδομένα από το Παράδειγμα 4.2 του βιβλίου [2].

Παράδειγμα 3.5.1. Έστω συνάρτηση $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ όπου $f(t) = \frac{1}{t}$ για κάθε t όπου $t \in \mathbb{R}\setminus\{0\}$. Το πρόγραμμα στην γλώσσα MATLAB που προσδιορίζει τους συντελεστές του πολυωνύμου p όπου $p \in \mathbb{P}_2$ τέτοιο ώστε να παρεμβάλλεται στα σημεία $(t_1, f(t_1)), (t_2, f(t_2))$, $(t_3, f(t_3))$ όπου $t_1 = 2, t_2 = 2.5, t_3 = 4$ είναι το ακόλουθο:

```
v=[2;2.5;4]
y=v.^(-1)
x=inter(2,v,y)
```

Ο 3 επί 1 πίνακας x των συντελεστών του πολυωνύμου p που επιστρέφει το παραπάνω πρόγραμμα συμφωνεί με τα αποτελέσματα του Παραδείγματος 4.2 του βιβλίου [2] και είναι ο ακόλουθος:

$$x = \begin{bmatrix} 1.15 \\ -0.425 \\ 0.05 \end{bmatrix}$$

Στο ακόλουθο παράδειγμα χρησιμοποιούμε τα δεδομένα από το Παράδειγμα 8.6 του βιβλίου [8].

Παράδειγμα 3.5.2. Έστω άγνωστη συνάρτηση f της οποίας είναι γνωστές οι τιμές στα σημεία t_i όπου $i \in \{1, \ldots, 7\}$ όπως παρουσιάζονται στον ακόλουθο πίνακα:

i	1	2	3	4	5	6	7
t_i	0	2	5	7	9	13	24
$f(t_i)$	0	6	7.9	8.5	12	21.5	35

Το πρόγραμμα στην γλώσσα MATLAB που προσδιορίζει τους συντελεστές των πολυωνύμων p_1, p_2 όπου $p_1 \in \mathbb{P}_1, p_2 \in \mathbb{P}_2$ τέτοια ώστε να παρεμβάλλονται στα σημεία $(t_i, f(t_i))$ για κάθε i όπου $i \in \{1, \ldots, 7\}$ είναι το ακόλουθο:

```
v=[0;2;5;7;9;13;24]
y=[0;6;7.9;8.5;12;21.5;35]
x_1=inter(1,v,y)
x_2=inter(2,v,y)
```

Ο 2 επί 1 πίνακας x_1 των συντελεστών του πολυωνύμου p_1 και ο 3 επί 1 πίνακας x_2 των συντελεστών του πολυωνύμου p_2 που επιστρέφει το παραπάνω πρόγραμμα συμφωνούν με τα αποτελέσματα του Παραδείγματος 8.6 του βιβλίου [8] και είναι οι ακόλουθοι:

$$x_1 = \begin{bmatrix} 0.6831 \\ 1.4353 \end{bmatrix}$$

$$x_2 = \begin{bmatrix} 0.8977 \\ 1.3695 \\ 0.0027 \end{bmatrix}$$

Στο ακόλουθο παράδειγμα χρησιμοποιούμε τα δεδομένα από το Παράδειγμα 5.1.2-1 του βιβλίου [12].

Παράδειγμα 3.5.3. Έστω άγνωστη συνάρτηση f της οποίας είναι γνωστές οι τιμές στα σημεία t_i όπου $i \in \{1, \ldots, 4\}$ όπως παρουσιάζονται στον ακόλουθο πίνακα:

i	1	2	3	4
t_i	-0.5	0.3	0.7	1.5
$f(t_i)$	1.2	2	1	-1

Το πρόγραμμα στην γλώσσα MATLAB που προσδιορίζει τους συντελεστές του πολυωνύμου p όπου $p \in \mathbb{P}_1$ τέτοιο ώστε να παρεμβάλλεται στα σημεία $(t_i, f(t_i))$ για κάθε i όπου $i \in \{1, \ldots, 4\}$ είναι το ακόλουθο:

```
v=[-0.5;0.3;0.7;1.5]
y=[1.2;2;1;-1]
x=inter(1,v,y)
```

Ο 2 επί 1 πίνακας x των συντελεστών του πολυωνύμου p που επιστρέφει το παραπάνω πρόγραμμα συμφωνούν με τα αποτελέσματα του Παραδείγματος 5.1.2-1 του βιβλίου [12] και είναι ο ακόλουθος:

$$x = \begin{bmatrix} 1.3769 \\ -1.1538 \end{bmatrix}$$

Έστω $[a, b] \subset \mathbb{R}, \Delta_{[a,b],1:m} : a = t_1 < \cdots < t_m = b$ ένας διαμερισμός του $[a, b]$ και $n \in \mathbb{N}$. Πολυωνυμικές splines ή απλά splines βαθμού n ως προς τον διαμερισμό $\Delta_{[a,b],1:m}$ λέγονται οι συνεχείς συναρτήσεις $s : [a, b] \to \mathbb{R}$ οι οποίες σε κάθε διάστημα $[t_i, t_{i+1}]$ όπου $i \in \{1, \ldots, m-1\}$ είναι πολυώνυμα βαθμού το πολύ n. Ορίζουμε ως χώρο των splines βαθμού n ως προς τον διαμερισμό $\Delta_{[a,b],1:m}$ το σύνολο:

$$S_n(\Delta_{[a,b],1:m}) = \{s : s \in C([a, b]) \text{ και } s|_{[t_i,t_{i+1}]} \in \mathbb{P}_n \text{ όπου } i \in \{1, \ldots, m-1\}\}$$

Αλγόριθμος 3.5.2. Ο αλγόριθμος splin που παρουσιάζεται παρακάτω δέχεται ως είσοδο το στοιχείο (n, v, y) όπου $n \in \mathbb{N}, v \in \mathbb{R}^m, y \in \mathbb{R}^m$ τέτοια ώστε n είναι ο βαθμός της spline $s|_{[t_i,t_{i+1}]}(t) = \sum_{j=1}^{n+1}(x_{j-1,i}t^{j-1})$ που παρεμβάλλεται στα σημεία $(t_i, f(t_i)), (t_{i+1}, f(t_{i+1}))$ της συνάρτησης $f : \mathbb{R} \to \mathbb{R}$ όπου $i \in \{1, \ldots, m-1\}$ και $v(i) = t_i, y(i) = f(t_i)$ για κάθε i όπου $i \in \{1, \ldots, m\}$ ενώ επιστρέφει ως έξοδο το στοιχείο X όπου $X \in \mathbb{R}^{(n+1)\times(m-1)}$ τέτοιο ώστε $X(j, i) = x_{j-1,i}$ για κάθε j, i όπου $j \in \{1, \ldots, n+1\}, i \in \{1, \ldots, m-1\}$.

algorithm splin$(n, v, y) = X$
  $m \leftarrow$ length(v)
  if $(m = 0) \vee (n \leqslant 0)$ **then**
    print("Invalid input.")
    return
  end if
  $X \leftarrow 0_{(n+1)\times(m-1)}$
  for $i \leftarrow 1 : m-1$ **do**
    $v_0 \leftarrow v(i : i+1)$
    $y_0 \leftarrow y(i : i+1)$
    $X(1 : n+1, i : i) \leftarrow$ inter(n, v_0, y_0)
  end for
end

Η υλοποίηση του αλγορίθμου splin στην γλώσσα MATLAB είναι η ακόλουθη:

```
function [X]=splin(n,v,y)
m=length(v);
if (m==0)||(n<=0)
    fprintf("Invalid input.\n");
    return;
end
X=zeros(n+1,m-1);
for i=1:m-1
    v_0=v(i:i+1);
    y_0=y(i:i+1);
    X(1:n+1,i:i)=inter(n,v_0,y_0);
end
end
```

Στο ακόλουθο παράδειγμα χρησιμοποιούμε τα δεδομένα από το παράδειγμα της Ενότητας 6.2.1 του βιβλίου [11].

Παράδειγμα 3.5.4. Έστω άγνωστη συνάρτηση f της οποίας είναι γνωστές οι τιμές στα σημεία t_i όπου $i \in \{1, \ldots, 4\}$ όπως παρουσιάζονται στον ακόλουθο πίνακα:

i	1	2	3	4
t_i	0	1	4	9
$f(t_i)$	0	1	2	3

Το πρόγραμμα στην γλώσσα MATLAB που προσδιορίζει τους συντελεστές της spline s βαθμού 1 ως προς τον διαμερισμό $\Delta_{[0,9],1:4}$ είναι το ακόλουθο:

```
v=[0;1;4;9]
y=[0;1;2;3]
X=splin(1,v,y)
```

Ο 2 επί 3 πίνακας X των συντελεστών της spline s που επιστρέφει το παραπάνω πρόγραμμα συμφωνεί με τα αποτελέσματα του παραδείγματος της Ενότητας 6.2.1 του βιβλίου [11] και είναι ο ακόλουθος:

$$X = \begin{bmatrix} 0 & 0.6667 & 1.2 \\ 1 & 0.3333 & 0.2 \end{bmatrix}$$

4 ΣΥΜΠΕΡΑΣΜΑΤΑ

Οι μετασχηματισμοί Householder εξασφαλίζουν ευστάθεια στους υπολογισμούς διότι οδηγούν στην εμφάνιση ορθογώνιων πινάκων κατά την ανάλυση ενός πίνακα σε γινόμενο πινάκων. Για παράδειγμα έστω $x \in \mathbb{R}^m$ και $\bar{x} = x + \varepsilon$ είναι μια προσέγγιση του x όπου ο m επί 1 πίνακας ε είναι το σφάλμα που υπεισέρχεται κατά τον υπολογισμό του x. Αν Q είναι m επί m ορθογώνιος πίνακας τότε $Q\bar{x} = Q(x + \varepsilon) = Qx + Q\varepsilon$ άρα το σφάλμα που υπεισέρχεται κατά τον υπολογισμό του $Q\bar{x}$ είναι $Q\varepsilon$ και $\|Q\varepsilon\|_2 = \|\varepsilon\|_2$. Ομοίως αν $A \in \mathbb{R}^{m\times n}$ και $\bar{A} = A + E$ είναι μια προσέγγιση του A όπου ο m επί n πίνακας E είναι το σφάλμα που υπεισέρχεται κατά τον υπολογισμό του A τότε $Q\bar{A} = Q(A+E) = QA+QE$ άρα το σφάλμα που υπεισέρχεται κατά τον υπολογισμό του $Q\bar{A}$ είναι QE και $\|QE\|_2 = \|E\|_2$. Συμπεραίνουμε λοιπόν ότι αν ένας μετασχηματισμός Householder εφαρμόζεται σε ένα διάνυσμα ή σε ένα πίνακα τότε το σφάλμα ως προς την νόρμα $\|\cdot\|_2$ δεν αυξάνεται και αυτή η ιδιοτητά τον καθιστά ένα από τα σημαντικότερα εργαλεία της Αριθμητικής Γραμμικής Άλγεβρας.

ΠΑΡΑΡΤΗΜΑ

Θα αποδείξουμε τέσσερα λήμματα που θα χρειαστούν για την απόδειξη της Πρότασης 2.10.2.

Λήμμα 1. Αν $A \in \mathbb{R}^{m\times m}, X \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times n}$ όπου $m > n$ τέτοιοι ώστε $AX = XB$ και rank$(X) = n$ τότε υπάρχουν πίνακες $C_{1,1}, C_{1,2}, C_{2,2}, Q$ όπου $C_{1,1} \in \mathbb{R}^{n\times n}, C_{1,2} \in \mathbb{R}^{n\times(m-n)}, C_{2,2} \in \mathbb{R}^{(m-n)\times(m-n)}, Q \in \mathbb{R}^{m\times m}$ και ο Q είναι ορθογώνιος τέτοιοι ώστε:

$$Q^T AQ = \begin{bmatrix} C_{1,1} & C_{1,2} \\ 0_{(m-n)\times n} & C_{2,2} \end{bmatrix}$$

Απόδειξη. Αν εφαρμόσουμε παραγοντοποίηση QR στον πίνακα X θα προκύψουν πίνακες Q, R όπου ο Q είναι m επί m ορθογώνιος πίνακας και ο R είναι n επί n άνω τριγωνικός πίνακας με μη μηδενικά διαγώνια στοιχεία τέτοιοι ώστε:

$$X = Q \begin{bmatrix} R \\ 0_{(m-n)\times n} \end{bmatrix} \tag{1}$$

$$AX = XB \overset{(1)}{\Longleftrightarrow} AQ \begin{bmatrix} R \\ 0_{(m-n)\times n} \end{bmatrix} = Q \begin{bmatrix} R \\ 0_{(m-n)\times n} \end{bmatrix} B \Leftrightarrow$$

$$Q^T AQ \begin{bmatrix} R \\ 0_{(m-n)\times n} \end{bmatrix} = \begin{bmatrix} RB \\ 0_{(m-n)\times n} \end{bmatrix} \tag{2}$$

Υπάρχουν πίνακες $C_{1,1}, C_{1,2}, C_{2,1}, C_{2,2}$ όπου $C_{1,1} \in \mathbb{R}^{n\times n}, C_{1,2} \in \mathbb{R}^{n\times(m-n)}, C_{2,1} \in \mathbb{R}^{(m-n)\times n}$, $C_{2,2} \in \mathbb{R}^{(m-n)\times(m-n)}$ τέτοιοι ώστε:

$$Q^T AQ = \begin{bmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \end{bmatrix} \tag{3}$$

Από τις εκφράσεις (2),(3) προκύπτει ότι:

$$\begin{bmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \end{bmatrix} \begin{bmatrix} R \\ 0_{(m-n)\times n} \end{bmatrix} = \begin{bmatrix} RB \\ 0_{(m-n)\times n} \end{bmatrix} \Leftrightarrow \begin{bmatrix} C_{1,1}R + C_{1,2}0_{(m-n)\times n} \\ C_{2,1}R + C_{2,2}0_{(m-n)\times n} \end{bmatrix} = \begin{bmatrix} RB \\ 0_{(m-n)\times n} \end{bmatrix} \Leftrightarrow$$

$$\begin{bmatrix} C_{1,1}R \\ C_{2,1}R \end{bmatrix} = \begin{bmatrix} RB \\ 0_{(m-n)\times n} \end{bmatrix} \tag{4}$$

Ο άνω τριγωνικός πίνακας πίνακας R έχει μη μηδενικά διαγώνια στοιχεία άρα det$(R) = \prod_{i=1}^{n} R(i,i) \neq 0$ επομένως ο R είναι αντιστρέψιμος και από την έκφραση (4) προκύπτει ότι $C_{2,1}R = 0_{(m-n)\times n} \Leftrightarrow C_{2,1}RR^{-1} = 0_{(m-n)\times n}R^{-1} \Leftrightarrow C_{2,1} = 0_{(m-n)\times n}$ συνεπώς:

$$Q^T AQ = \begin{bmatrix} C_{1,1} & C_{1,2} \\ 0_{(m-n)\times n} & C_{2,2} \end{bmatrix}$$

□

Λήμμα 2(διάσπαση Schur). Αν $A \in \mathbb{R}^{m\times m}$ τότε υπάρχουν m επί m πίνακες Q, D, U όπου ο Q είναι ορθογώνιος, ο D είναι διαγώνιος, ο U είναι αυστηρά άνω τριγωνικός τέτοιοι ώστε $Q^T AQ = D + U$.

Απόδειξη. Αν $m = 1$ τότε $1^T \cdot A \cdot 1 = A + 0$.
Υποθέτουμε ότι αν $m \in \{1, \ldots, l\}$ τότε υπάρχουν m επί m πίνακες Q, D, U όπου ο Q είναι ορθογώνιος, ο D είναι διαγώνιος, ο U είναι αυστηρά άνω τριγωνικός τέτοιοι ώστε $Q^T AQ = D + U$.
Έστω $A \in \mathbb{R}^{(m+1)\times(m+1)}$. Αν λ είναι ιδιοτιμή του πίνακα A τότε υπάρχει x όπου $x \in \mathbb{R}^m \backslash \{0_m\}$ τέτοιο ώστε $Ax = \lambda x$ άρα από το Λήμμα 1 προκύπτει ότι υπάρχουν πίνακες c, C, V όπου $c \in \mathbb{R}^{1\times m}, C \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{(m+1)\times(m+1)}$ και ο V είναι ορθογώνιος τέτοιοι ώστε:

$$V^T AV = \begin{bmatrix} \lambda & c \\ 0_m & C \end{bmatrix} \quad (1)$$

Από την επαγωγική υπόθεση υπάρχει m επί m ορθογώνιος πίνακας W τέτοιος ώστε ο πίνακας $W^T CW$ είναι ίσος με το άθροισμα ενός διαγώνιου πίνακα με έναν αυστηρά άνω τριγωνικό πίνακα δηλαδή ο πίνακας $W^T CW$ είναι άνω τριγωνικός και θέτουμε:

$$Q = V \begin{bmatrix} 1 & 0_m^T \\ 0_m & W \end{bmatrix} \quad (2)$$

$$Q^T AQ = \begin{bmatrix} 1 & 0_m^T \\ 0_m & W^T \end{bmatrix} V^T AV \begin{bmatrix} 1 & 0_m^T \\ 0_m & W \end{bmatrix} = \begin{bmatrix} 1 & 0_m^T \\ 0_m & W^T \end{bmatrix} \begin{bmatrix} \lambda & c \\ 0_m & C \end{bmatrix} \begin{bmatrix} 1 & 0_m^T \\ 0_m & W \end{bmatrix} =$$
$$\begin{bmatrix} \lambda & c \\ 0_m & W^T C \end{bmatrix} \begin{bmatrix} 1 & 0_m^T \\ 0_m & W \end{bmatrix} \Leftrightarrow$$

$$Q^T AQ = \begin{bmatrix} \lambda & cW \\ 0_m & W^T CW \end{bmatrix} \quad (3)$$

Ο πίνακας $W^T CW$ είναι άνω τριγωνικός άρα από την έκφραση (3) προκύπτει ότι ο πίνακας $Q^T AQ$ είναι άνω τριγωνικός και κάθε άνω τριγωνικός πίνακας μπορεί να παρασταθεί ως το άθροισμα ενός διαγώνιου πίνακα με έναν αυστηρά άνω τριγωνικό πίνακα επομένως από την αρχή της μαθηματικής επαγωγής προκύπτει ότι για κάθε m όπου $m \in \mathbb{N}$ υπάρχουν m επί m πίνακες Q, D, U όπου ο Q είναι ορθογώνιος, ο D είναι διαγώνιος, ο U είναι αυστηρά άνω τριγωνικός τέτοιοι ώστε $Q^T AQ = D + U$. □

Λήμμα 3. Αν ο A είναι m επί m συμμετρικός πίνακας τότε υπάρχουν m επί m πίνακες Q, D όπου ο Q είναι ορθογώνιος, ο D είναι διαγώνιος τέτοιοι ώστε $Q^T AQ = D$.

Απόδειξη. Από το Λήμμα 2 προκύπτει ότι υπάρχουν m επί m πίνακες Q, D, U όπου ο Q είναι ορθογώνιος, ο D είναι διαγώνιος, ο U είναι αυστηρά άνω τριγωνικός τέτοιοι ώστε:

$$Q^T AQ = D + U \quad (1)$$

$$(Q^T AQ)^T = (D+U)^T \Leftrightarrow Q^T A^T (Q^T)^T = D^T + U^T \Leftrightarrow Q^T AQ = D + U^T \Leftrightarrow D + U = D + U^T \Leftrightarrow$$

$$U = U^T \quad (2)$$

Ο m επί m πίνακας U είναι αυστηρά άνω τριγωνικός άρα $U(i,j) = 0$ για κάθε i, j όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, m\}, i \geqslant j$ και από την έκφραση (1) προκύπτει ότι $U(i,j) = U(j,i)$ για κάθε i, j όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, m\}$ επομένως $U(i,j) = U(j,i) = 0$ για κάθε i, j όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, m\}, j \geqslant i$ συνεπώς $U(i,j) = 0$ για κάθε i, j όπου $i \in \{1, \ldots, m\}, j \in \{1, \ldots, m\}$ δηλαδή:

$$U = 0_{m\times m} \quad (3)$$

Από τις εκφράσεις (1),(3) προκύπτει ότι $Q^T AQ = D$. □

Λήμμα 4. Κάθε συμμετρικός πίνακας έχει πραγματικές ιδιοτιμές στις οποίες αντιστοιχεί μια ορθοκανονική βάση που αποτελείται από ιδιοδιανύσματα.

Απόδειξη. Από το Λήμμα 3 προκύπτει ότι για κάθε m επί m συμμετρικό πίνακα A υπάρχει m επί m ορθογώνιος πίνακας $Q = [q_1 \ \dots \ q_m]$ και m επί m διαγώνιος πίνακας $D = \text{diag}(d_{1,1}, \dots, d_{m,m})$ τέτοιοι ώστε:
$Q^T AQ = D \Leftrightarrow AQ = QD \Leftrightarrow A[q_1 \ \dots \ q_m] = Q\text{diag}(d_{1,1}, \dots, d_{m,m}) \Leftrightarrow$

$$[Aq_1 \ \dots \ Aq_m] = [d_{1,1}q_1 \ \dots \ d_{m,m}q_m] \tag{1}$$

Από την έκφραση (1) προκύπτει ότι $Aq_k = d_{k,k}q_k$ για κάθε k όπου $k \in \{1, \dots, m\}$ άρα κάθε $d_{k,k}$ και κάθε q_k είναι ιδιοτιμή και ιδιοδιάνυσμα αντίστοιχα του πίνακα A. □

Πρόταση 2.10.2. Αν $A \in \mathbb{R}^{m \times n}$ τότε $\|A\|_2 = \rho^{\frac{1}{2}}(A^T A)$.

Απόδειξη. Έχουμε ότι $(A^T A)^T = A^T A$ άρα ο n επί n πίνακας $A^T A$ είναι συμμετρικός επομένως έχει πραγματικές ιδιοτιμές $\lambda_1, \dots, \lambda_n$ στις οποίες αντιστοιχεί μια ορθοκανονική βάση που αποτελείται από ιδιοδιανύσματα $q_1, \dots, q_n$ συνεπώς για κάθε x όπου $x \in \mathbb{R}^n$ ισχύει ότι:

$$x \in \text{span}(\{q_1, \dots, q_n\}) \tag{1}$$

Από την έκφραση (1) προκύπτει ότι για κάθε i όπου $i \in \{1, \dots, n\}$ υπάρχουν s_i όπου $s_i \in \mathbb{R}$ τέτοια ώστε $x = \sum_{i=1}^{n}(s_i q_i)$ άρα $\|x\|_2^2 = \langle x, x \rangle_2 = \langle \sum_{i=1}^{n}(s_i q_i), \sum_{j=1}^{n}(s_j q_j) \rangle_2 =$
$\sum_{i=1}^{n}\sum_{j=1}^{n}(s_i s_j \langle q_i, q_j \rangle_2) = \sum_{i=1}^{n}\sum_{j=1}^{n}(s_i s_j \delta_{i,j}) \Leftrightarrow$

$$\|x\|_2^2 = \sum_{i=1}^{n} s_i^2 \tag{2}$$

$\|Ax\|_2^2 = \langle Ax, Ax \rangle_2 = (Ax)^T Ax = x^T A^T Ax = (A^T Ax)^T x = \langle A^T Ax, x \rangle_2 =$
$\langle \sum_{i=1}^{n}(\lambda_i s_i q_i), \sum_{j=1}^{n}(s_j q_j) \rangle_2 = \sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i s_i s_j \langle q_i, q_j \rangle_2) = \sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i s_i s_j \delta_{i,j}) \Leftrightarrow$

$$\|Ax\|_2^2 = \sum_{i=1}^{n}(\lambda_i s_i^2) \tag{3}$$

Αν $x \in \mathbb{R}^n \setminus \{0_n\}$ τότε από τις εκφράσεις (2),(3) προκύπτει ότι $\dfrac{\|Ax\|_2^2}{\|x\|_2^2} = \dfrac{\sum_{i=1}^{n}(\lambda_i s_i^2)}{\sum_{i=1}^{n} s_i^2} \leqslant$
$\dfrac{\sum_{i=1}^{n}(|\lambda_i| s_i^2)}{\sum_{i=1}^{n} s_i^2} \leqslant \dfrac{\sum_{i=1}^{n}(\max(\{|\lambda_i| : i \in \{1, \dots, n\}\}) s_i^2)}{\sum_{i=1}^{n} s_i^2} = \max(\{|\lambda_i| : i \in \{1, \dots, n\}\}) \dfrac{\sum_{i=1}^{n} s_i^2}{\sum_{i=1}^{n} s_i^2} =$
$\rho(A^T A)$ άρα $\left(\dfrac{\|Ax\|_2}{\|x\|_2}\right)^2 \leqslant \rho(A^T A) \Leftrightarrow \dfrac{\|Ax\|_2}{\|x\|_2} \leqslant \rho^{\frac{1}{2}}(A^T A)$ για κάθε x όπου $x \in \mathbb{R}^n \setminus \{0_n\}$ επομένως:

$$\|A\|_2 \leqslant \rho^{\frac{1}{2}}(A^T A) \tag{4}$$

Υπάρχει k όπου $k \in \{1, \ldots, n\}$ τέτοιο ώστε $|\lambda_k| = \rho(A^T A)$ άρα $0 \leqslant \|Aq_k\|_2^2 = \langle Aq_k, Aq_k \rangle_2 = \langle A^T Aq_k, q_k \rangle_2 = \langle \lambda_k q_k, q_k \rangle_2 = \lambda_k \langle q_k, q_k \rangle_2 = \lambda_k = |\lambda_k| = \rho(A^T A)\|q_k\|_2^2 \Leftrightarrow \dfrac{\|Aq_k\|_2^2}{\|q_k\|_2^2} = \rho(A^T A) \Leftrightarrow \left(\dfrac{\|Aq_k\|_2}{\|q_k\|_2}\right)^2 = \rho(A^T A) \Leftrightarrow \dfrac{\|Aq_k\|_2}{\|q_k\|_2} = \rho^{\frac{1}{2}}(A^T A)$ επομένως:

$$\|A\|_2 \geqslant \rho^{\frac{1}{2}}(A^T A) \tag{5}$$

Από τις εκφράσεις (4),(5) προκύπτει ότι $\|A\|_2 = \rho^{\frac{1}{2}}(A^T A)$. $\square$

ΑΝΑΦΟΡΕΣ

[1] Μαριλένα Μητρούλη, Αριθμητική Γραμμική Άλγεβρα και Επιστημονικοί Υπολογισμοί, Εκδόσεις Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών, 2022.
[2] Γεώργιος Δ. Ακρίβης - Βασίλειος Α. Δουγαλής, Εισαγωγή στην Αριθμητική Ανάλυση, Τέταρτη Έκδοση, Πανεπιστημιακές Εκδόσεις Κρήτης, 2014.
[3] Gene H. Golub - Charles F. Van Loan, Θεωρία και Υπολογισμοί Μητρώων, Τέταρτη Έκδοση, Εκδόσεις Πεδίο, 2015.
[4] Στυλιανός Νεγρεπόντης - Σταύρος Γιωτόπουλος - Ευστάθιος Γιαννακούλιας, Απειροστικός Λογισμός, Τόμος Ι, Εκδόσεις Συμμετρία, 1999.
[5] Τηλέμαχος Ε. Χατζηφράτης, Απειροστικός Λογισμός σε Πολλές Μεταβλητές, Εκδόσεις Συμμετρία, 2009.
[6] Χαρά Χαραλάμπους - Ανέστης Φωτιάδης, Μια Εισαγωγή στη Γραμμική Άλγεβρα για τις Θετικές Επιστήμες, Ηλεκτρονική Έκδοση, Ανοικτές Ακαδημαϊκές Εκδόσεις Κάλλιπος, 2015.
[7] Βασίλειος Δουγαλής - Δημήτριος Νούτσος - Απόστολος Χατζηδήμος, Σημειώσεις Αριθμητικής Γραμμικής Άλγεβρας, Ηλεκτρονική Έκδοση, 2011.
[8] Biswa Nath Datta, Numerical Linear Algebra and Applications, Second Edition, Society for Industrial and Applied Mathematics, 2010.
[9] Vassilios A. Dougalis, Finite Element Methods for the Numerical Solution of Partial Differential Equations, Revised Electronic Edition, 2019.
[10] Willliam Ford, Numerical Linear Algebra with Applications, First Edition, Elsevier, 2015.
[11] Σταύρος Παπαϊωάννου - Χρήστος Βοζίκης, Αριθμητική Ανάλυση, Ηλεκτρονική Έκδοση, Ανοικτές Ακαδημαϊκές Εκδόσεις Κάλλιπος, 2015.
[12] Αθανάσιος Μπράτσος, Μαθήματα Εφαρμοσμένων Μαθηματικών, Ηλεκτρονική Έκδοση, Ανοικτές Ακαδημαϊκές Εκδόσεις Κάλλιπος, 2015.