Κεφάλαιο 1

Μήκος και ορθογωνιότητα

1.1 Μήκος διανύσματος στον $\mathbb{R}^n$

Θεωρούμε το $x = (x_1, x_2)$ ως το διάνυσμα συντεταγμένων του διανύσματος $\overrightarrow{OX}$ ως προς ένα ορθοκανονικό σύστημα αναφοράς στο E^2. Ορίζουμε το μήκος $||x||$ ενός διανύσματος $x = (x_1, x_2)$ να είναι ίσο με το μήκος του γεωμετρικού διανύσματος $\overrightarrow{OX}$. Αφού οι άξονες του συστήματος αναφοράς είναι κάθετοι μεταξύ τους, από το Πυθαγόρειο θεώρημα έχουμε:

$$||x||^2 = x_1^2 + x_2^2\,.$$

Ανάλογα, για το μήκος ενός διανύσματος στο χώρο E^3, με διάνυσμα συντεταγμένων $x = (x_1, x_2, x_3)$ ως προς ένα ορθοκανονικό σύστημα αναφοράς, εφαρμόζουμε το Πυθαγόρειο θεώρημα 2 φορές: εάν $x = (x_1, x_2, x_3)$ και $u = (x_1, x_2, 0)$

$$\begin{aligned} ||x||^2 &= ||u||^2 + x_3^2 \\ &= x_1^2 + x_2^2 + x_3^2\,. \end{aligned}$$

Χρησιμοποιώντας το συμβολισμό του ανάστροφου, αυτό γράφεται

$$||x||^2 = x^T x\,.$$

Κατ' αναλογία, ορίζουμε το μήκος $||x||$ ενός διανύσματος στο $\mathbb{R}^n$:

$$\begin{aligned} ||x||^2 &= x_1^2 + x_2^2 + \cdots + x_n^2 \\ &= x^T x\,. \end{aligned}$$

Παράδειγμα 1.1 Το μήκος του διανύσματος $x = (1, 2, -3)$ είναι $\sqrt{14}$:

$$\|x\|^2 = x^T x = [1\ 2\ -3] \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} = 1^2 + 2^2 + (-3)^2 = 14\,.$$

1.2 Ορθογώνια διανύσματα στον $\mathbb{R}^n$

Εκτός από τα μήκη, θέλουμε να ορίσουμε και γωνίες μεταξύ διανυσμάτων. Αργότερα θα μιλήσουμε για όλες τις γωνίες, αλλά προς το παρόν μας ενδιαφέρουν οι **ορθές γωνίες**. Πότε είναι δύο διανύσματα $x\,,y$ **ορθογώνια**;

Το Πυθαγόρειο θεώρημα ισχύει και αντίστροφα: ένα τρίγωνο είναι ορθογώνιο **μόνον** όταν το τετράγωνο της υποτείνουσας είναι ίσο με το άθροισμα των τετραγώνων των 2 πλευρών. Μπορούμε να εργαστούμε στο $\mathbb{R}^n$, αλλά στην πραγματικότητα οι μετρήσεις θα είναι μέσα στο επίπεδο που περιέχει το τρίγωνο, δηλαδή μέσα στο διανυσματικό υπόχωρο που παράγεται από τα διανύσματα x και y. Η γωνία $\angle(x,\, y)$ είναι ορθή εάν και μόνον εάν

$$||x||^2 + ||y||^2 = ||x-y||^2\,,$$

ή

$$x_1^2 + \cdots + x_n^2 + y_1^2 + \cdots + y_n^2 = (x_1-y_1)^2 + \cdots + (x_n-y_n)^2\,,$$

δηλαδή εάν και μόνον εάν

$$x_1y_1 + \cdots + x_ny_n = 0$$

ή

$$x^Ty = 0\,.$$

Η ποσότητα x^Ty γενικεύει στους χώρους $\mathbb{R}^n$ το εσωτερικό γινόμενο δύο διανυσμάτων που γνωρίζουμε από την Αναλυτική Γεωμετρία.

Ορισμός 1.1. Δύο διανύσματα $x,\, y$ του $\mathbb{R}^n$ λέγονται **ορθογώνια** εάν το εσωτερικό τους γινόμενο x^Ty είναι 0.

Παράδειγμα 1.2 Το διάνυσμα $x = (2,\, 2,\, -1)$ είναι ορθογώνιο στο $y = (-1,\, 2,\, 2)$:

$$x^Ty = [2\ 2\ -1]\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = 0\,.$$

Ένα διάνυσμα είναι ορθογώνιο στον εαυτό του μόνον εάν έχει μηδενικό μήκος: $x^Tx = 0$. Το μοναδικό τέτοιο διάνυσμα του $\mathbb{R}^n$ είναι το 0.

Πρόταση 1.1 *Εάν τα διανύσματα $v_1, \ldots, v_k$ είναι μη μηδενικά και ορθογώνια μεταξύ τους, τότε είναι γραμμικά ανεξάρτητα.*

Απόδειξη. Έστω ένας γραμμικός συνδυασμός $c_1v_1 + \cdots + c_kv_k = 0$. Παίρνουμε το εσωτερικό γινόμενο με το v_1:

$$v_1^T(c_1v_1 + \cdots + c_kv_k) = v_1^T0 = 0$$

Αλλά $v_1^T v_i = 0$ για κάθε $i \neq 1$, άρα έχουμε

$$v_1^T c_1 v_1 = c_1 ||v_1||^2 = 0$$

και εφόσον $||v_1|| \neq 0$, έχουμε $c_1 = 0$.

Παρόμοια, $c_i = 0$ για κάθε i, και συμπεραίνουμε οτι τα διανύσματα είναι γραμμικά ανεξάρτητα.

□

Είναι προφανές οτι δεν ισχύει το αντίστροφο: δυο γραμμικά ανεξάρτητα διανύσματα δεν είναι υποχρεωτικά ορθογώνια.

Δραστηριότητα 1.1 Βρείτε τα μήκη και το εσωτερικό γινόμενο των $x = (1, 4, 0, 2)$ και $y = (2, -2, 1, 3)$.

Δραστηριότητα 1.2 Ποία ζεύγη από τα διανύσματα u_1, u_2, u_3, u_4 είναι ορθογώνια;

$$u_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}, \ u_2 = \begin{bmatrix} 4 \\ 0 \\ 4 \\ 0 \end{bmatrix}, \ u_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \ u_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

1.3 Ορθογώνιοι υπόχωροι στον $\mathbb{R}^n$

Στον $\mathbb{R}^3$, μία ευθεία είναι κάθετη σε ένα επίπεδο όταν σχηματίζει ορθή γωνία με κάθε ευθεία του επιπέδου που την τέμνει.

Ανάλογα, δύο υπόχωροι V και W του χώρου $\mathbb{R}^n$ είναι **ορθογώνιοι** όταν **κάθε** διάνυσμα του V είναι ορθογώνιο σε **κάθε** διάνυσμα του W.

Παρατηρούμε οτι δύο επίπεδα W_1 και W_2 στο $\mathbb{R}^3$ που σχηματίζουν ορθή δίεδρη γωνία **δεν** ικανοποιούν αυτή τη συνθήκη. Πράγματι, ας θεωρήσουμε μία βάση από δύο ορθογώνια διανύσματα σε κάθε επίπεδο, u_1, v_1 στο W_1, u_2, v_2 στο W_2. Εάν τα W_1 και W_2 ήταν ορθογώνια, τότε θα είχαμε 4 διανύσματα u_1, v_1, u_2, v_2 ορθογώνια μεταξύ τους. Από την Πρόταση 1.1 αυτά θα ήταν γραμμικά ανεξάρτητα. Αλλά στον $\mathbb{R}^3$ δεν υπάρχουν 4 γραμμικά ανεξάρτητα διανύσματα.

Θα συμβολίζουμε την ορθογωνιότητα δύο γραμμικών υπόχωρων U και V του $\mathbb{R}^n$ με $U \perp V$.

Παράδειγμα 1.3 Στο $\mathbb{R}^4$ θεωρούμε το επίπεδο V που παράγεται από τα διανύσματα $v_1 = (1, 0, 0, 0)$ και $v_2 = (1, 1, 0, 0)$. Το διάνυσμα $w = (0, 0, 4, 5)$ είναι ορθογώνιο προς τα v_1 και v_2. Συνεπώς η ευθεία W που παράγεται από το w είναι υπόχωρος του $\mathbb{R}^4$ ορθογώνιος προς τον V. Αλλά μέσα στο $\mathbb{R}^4$ υπάρχει χώρος για ακόμη έναν υπόχωρο ορθογώνιο στους V και W: το διάνυσμα $z = (0, 0, 5, -4)$ είναι ορθογώνιο προς τα v_1, v_2 και w. Η ευθεία U που παράγεται από το z είναι ορθογώνια προς τους υπόχωρους V και W:

$$U \perp V\,,\ U \perp W\,,\ V \perp W\,.$$

Δραστηριότητα 1.3 Δείξτε οτι οι υπόχωροι του $\mathbb{R}^3$, $U = \{(x, y, z, \in \mathbb{R}^3 : x + y = 0\}$ και $V = \langle(1, 1, 0)\rangle$ είναι ορθογώνιοι.

Δραστηριότητα 1.4 Βρείτε a τέτοιο ώστε οι υπόχωροι του $\mathbb{R}^3$, $U = \langle(1, 1, a)\rangle$ και $V = \langle(1, a, 2)\rangle$ είναι ορθογώνιοι.

Πρόταση 1.2 *Δίδεται ένας $m \times n$ πίνακας A. Τότε*

α΄. Στο $\mathbb{R}^n$ ο χώρος γραμμών του A είναι ορθογώνιος στο μηδενόχωρο του A:

$$\mathcal{R}(A^T) \perp \mathcal{N}(A)$$

β΄. Στο $\mathbb{R}^m$ ο χώρος στηλών του A είναι ορθογώνιος στον αριστερό μηδενόχωρο του A:

$$\mathcal{R}(A) \perp \mathcal{N}(A^T)\,.$$

Απόδειξη. Αρκεί να δείξουμε την πρώτη περίπτωση, αφού η δεύτερη προκύπτει εξετάζοντας τον ανάστροφο πίνακα A^T.

Θεωρούμε ένα $x \in \mathcal{N}(A)$ και ένα $v \in \mathcal{R}(A^T)$, και θέλουμε να δείξουμε οτι $v^T x = 0$.

Έχουμε $Ax = 0$. Εφόσον $v \in \mathcal{R}(A^T)$, το v είναι γραμμικός συνδυασμός των γραμμών $r_1, \ldots r_m$ του A,

$$v = z_1 r_1 + \cdots + z_m r_m\,,$$

δηλαδή υπάρχει $z \in \mathbb{R}^m$ τέτοιο ώστε $v^T = z^T A$. Έχουμε

$$v^T x == (z^T A)x = z^T(Ax) = z^T 0 = 0\,.$$

□

1.4 Ορθογώνιο συμπλήρωμα

Υπενθυμίζουμε οτι οι διαστάσεις των θεμελιωδών υποχώρων ενός $m \times n$ πίνακα A ικανοποιούν τις σχέσεις:

$$\dim \mathcal{R}(A^T) + \dim \mathcal{N}(A) = n \tag{1.1}$$

$$\dim \mathcal{R}(A) + \dim \mathcal{N}(A^T) = m \tag{1.2}$$

Αυτή η παρατήρηση υποδεικνύει οτι ο χώρος γραμμών και ο μηδενόχωρος δεν είναι δύο οποιοιδήποτε ορθογώνιοι υπόχωροι του $\mathbb{R}^n$: οι δύο υπόχωροι 'γεμίζουν' τον $\mathbb{R}^n$. Ας εξετάσουμε πιο προσεκτικά την κατάσταση. Αν W είναι το σύνολο όλων των διανυσμάτων που είναι ορθογώνια σε όλα τα διανύσματα του χώρου γραμμών $\mathcal{R}(A^T)$, η Πρόταση 1.2 λέει οτι $\mathcal{N}(A) \subseteq W$.

Εύκολα όμως βλέπουμε οτι ισχύει και ο αντίθετος εγκλεισμός, $W \subseteq \mathcal{N}(A)$, δηλαδή ο μηδενόχωρος περιέχει κάθε διάνυσμα που είναι ορθογώνιο σε όλα τα διανύσματα του χώρου γραμμών. Πράγματι, εάν $x \in W$ τότε το x είναι ορθογώνιο σε κάθε γραμμή του A και $Ax = 0$. Αυτή η κατάσταση παρουσιάζει αρκετό ενδιαφέρον ώστε να της δώσουμε ένα όνομα:

Ορισμός 1.2. Θεωρούμε γραμμικό υπόχωρο V του $\mathbb{R}^n$. Το σύνολο όλων των διανυσμάτων του $\mathbb{R}^n$ που είναι ορθογώνια σε κάθε διάνυσμα του V ονομάζεται **ορθογώνιο συμπλήρωμα** του V στον $\mathbb{R}^n$, και συμβολίζεται $V^\perp$:

$$V^\perp = \{w \in \mathbb{R}^n : w^T v = 0 \text{ για κάθε } v \in V\}.$$

Λήμμα 1.3 *Το ορθογώνιο συμπλήρωμα $V^\perp$ είναι γραμμικός υπόχωρος του $\mathbb{R}^n$.*

Απόδειξη. Πρέπει να δείξουμε οτι $V^\perp \neq \varnothing$ και οτι για κάθε w_1, $w_2 \in V^\perp$ και $a \in \mathbb{R}$, $aw_1 + w_2 \in V^\perp$. Προφανώς $0 \in V^\perp$, άρα $V^\perp \neq \varnothing$. Εάν για κάθε $v \in V$ ισχύουν $w_1^T v = 0$ και $w_2^T v = 0$, τότε $(aw_1 + w_2)^T v = aw_1^T v + w_2^T v = 0$.

□

Δραστηριότητα 1.5 Δείξτε οτι ο υπόχωρος U της Δραστηριότητας 1.3 είναι το ορθογώνιο συμπλήρωμα του υπόχωρου V.

Έχουμε δείξει οτι ο μηδενόχωρος είναι το ορθογώνιο συμπλήρωμα του χώρου γραμμών:

$$\mathcal{N}(A) = \mathcal{R}(A^T)^\perp.$$

Θα δείξουμε οτι και ο χώρος γραμμών είναι το ορθογώνιο συμπλήρωμα του μηδενόχωρου:

$$\mathcal{R}(A^T) = \mathcal{N}(A)^\perp.$$

Η Πρόταση 1.2 λέει οτι $\mathcal{R}(A^T) \subseteq \mathcal{N}(A)^\perp$. Για να δείξουμε τον αντίθετο εγκλεισμό θεωρούμε ένα διάνυσμα z ορθογώνιο στο $\mathcal{N}(A)$. Έστω A' ο πίνακας που προκύπτει από τον A επισυνάπτοντας ως μία επί πλέον γραμμή τη z^T. Ο A' έχει τον ίδιο μηδενόχωρο με τον A, αφού η νέα εξίσωση $z^T x = 0$ ικανοποιείται για κάθε $x \in \mathcal{N}(A)$. Επίσης έχει τον ίδιο αριθμό στηλών, n. Συγκρίνοντας τη σχέση

$$\dim \mathcal{R}(A'^T) + \dim \mathcal{N}(A') = n$$

με την 1.1, και αφού $\mathcal{N}(A') = \mathcal{N}(A)$, συμπεραίνουμε οτι $\dim \mathcal{R}(A'^T) = \dim \mathcal{R}(A^T)$. Αλλά αυτό σημαίνει οτι το διάνυσμα z εξαρτάται γραμμικά από τα διανύσματα μιας βάσης του $\mathcal{R}(A^T)$, δηλαδή οτι ανήκει στο $\mathcal{R}(A^T)$. Δείξαμε οτι κάθε διάνυσμα z του $\mathcal{N}(A)^\perp$ ανήκει στο χώρο γραμμών $\mathcal{R}(A^T)$. Συμπεραίνουμε οτι $\mathcal{R}(A^T) = \mathcal{N}(A)^\perp$. Έχουμε αποδείξει το πρώτο μέρος του ακόλουθου θεωρήματος.

Θεώρημα 1.4 *Δίδεται ένας $m \times n$ πίνακας A. Τότε*

α΄. Ο μηδενόχωρος $\mathcal{N}(A)$ είναι το ορθογώνιο συμπλήρωμα του χώρου γραμμών $\mathcal{R}(A^T)$ στον $\mathbb{R}^n$, και ο χώρος γραμμών είναι το ορθογώνιο συμπλήρωμα του μηδενόχωρου στον $\mathbb{R}^n$,

$$\mathcal{N}(A) = \mathcal{R}(A^T)^{\perp} \text{ και } \mathcal{R}(A^T) = \mathcal{N}(A)^{\perp}\,.$$

β΄. Ο αριστερός μηδενόχωρος $\mathcal{N}(A^T)$ είναι το ορθογώνιο συμπλήρωμα του χώρου στηλών $\mathcal{R}(A)$ στον $\mathbb{R}^m$, και ο χώρος στηλών είναι το ορθογώνιο συμπλήρωμα του αριστερού μηδενόχωρου στον $\mathbb{R}^m$,

$$\mathcal{N}(A^T) = \mathcal{R}(A)^{\perp} \text{ και } \mathcal{R}(A) = \mathcal{N}(A^T)^{\perp}\,.$$

γ΄.

$$\mathbb{R}^n = \mathcal{R}(A^T) \oplus \mathcal{N}(A)\,, \qquad \mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^T)\,.$$

Απόδειξη. Έχουμε αποδείξει το πρώτο μέρος του θεωρήματος. Το δεύτερο μέρος αποδεικνύεται θεωρώντας τον ανάστροφο πίνακα. Για το τρίτο μέρος, αρκεί να δείξουμε οτι κάθε διάνυσμα $x \in \mathbb{R}^n$, γράφεται με μοναδικό τρόπο ως άθροισμα ενός διανύσματος του $\mathcal{R}(A^T)$ και ενός διανύσματος του $\mathcal{N}(A)$.

Πρώτα δείχνουμε οτι $\mathcal{R}(A^T) \cap \mathcal{N}(A) = \{0\}$. Εάν $x \in \mathcal{R}(A^T) \cap \mathcal{N}(A)$ τότε x είναι ορθογώνιο προς τον εαυτό του, συνεπώς $x = 0$.

Εάν $\{v_1, \ldots, v_r\}$ είναι βάση του $\mathcal{R}(A^T)$ και $\{w_1, \ldots, w_{n-r}\}$ βάση του $\mathcal{N}(A)$, τότε το σύνολο $\{v_1, \ldots, v_r, w_1, \ldots, w_{n-r}\}$ έχει n στοιχεία. Εάν δείξουμε οτι είναι γραμμικά ανεξάρτητο, τότε είναι βάση του $\mathbb{R}^n$. Υποθέτουμε οτι $a_1v_1+\cdots+a_rv_r+b_1w_1+\cdots+b_{n-r}w_{n-r}) = 0$. Θέλουμε να δείξουμε οτι τότε όλα τα a_i και b_j είναι 0. Έχουμε

$$a_1v_1 + \cdots + a_rv_r = y = -(b_1w_1 + \cdots + b_{n-r}w_{n-r})\,.$$

Αλλά η αριστερή πλευρά ανήκει στο $\mathcal{R}(A^T)$, η δεξιά πλευρά ανήκει στο $\mathcal{N}(A)$. Άρα το διάνυσμα y ανήκει στην τομή των δύο υποχώρων, και συνεπώς y είναι το μηδενικό διάνυσμα $0 \in \mathbb{R}^n$. Άρα $a_1v_1 + \cdots + a_rv_r = 0$, και αφού $v_1, \ldots, v_r$ είναι γραμμικά ανεξάρτητα, όλα τα a_i είναι 0. Παρόμοια, $b_1w_1 + \cdots + b_{n-r}w_{n-r} = 0$, και αφού το $w_1, \ldots, w_{n-r}$ είναι γραμμικά ανεξάρτητα, όλα τα b_j είναι 0.

Αφού $\{v_1, \ldots, v_r, w_1, \ldots, w_{n-r}\}$ είναι βάση του $\mathbb{R}^n$, κάθε διάνυσμα του $\mathbb{R}^n$ γράφεται ως άθροισμα ενός στοιχείου του $\mathcal{R}(A^T)$ και ενός στοιχείου του $\mathcal{N}(A)$. Αφού η τομή των δύο υποχώρων είναι $\{0\}$,

$$\mathbb{R}^n = \mathcal{R}(A^T) \oplus \mathcal{N}(A)\,.$$

□

Παρατηρήστε οτι η απόδειξη δεν δίδει τις συνιστώσες του x στο $\mathcal{R}(A^T)$ και στο $\mathcal{N}(A)$. Αυτό θα το εξετάσουμε σε επόμενη παράγραφο.

Με αυτό το Θεώρημα ολοκληρώνεται η περιγραφή των τεσσάρων θεμελιωδών υποχώρων ενός πίνακα, οι οποίοι αποτελούν δύο ζεύγη ορθογωνίων συμπληρωμάτων.

Η ακόλουθη Πρόταση δίδει τις βασικές ιδιότητες του ορθογωνίου συμπληρώματος.

Πρόταση 1.5 *Έστω ένας διανυσματικός υπόχωρος V του $\mathbb{R}^n$, και W το ορθογώνιο συμπλήρωμα του V, $W = V^\perp$. Τότε*

α΄. *Η διάσταση του W είναι $\dim W = n - \dim V$, και $V \cap W = \{0\}$.*

β΄. *Το ορθογώνιο συμπλήρωμα του W είναι ο V: εάν $W = V^\perp$ τότε $V = W^\perp$.*

γ΄. *Εάν $\{v_1, \ldots, v_k\}$ είναι βάση του V και $\{w_1, \ldots, w_{n-k}\}$ βάση του W, τότε $\{v_1, \ldots, v_k, w_1, \ldots, w_{n-k}\}$ είναι βάση του $\mathbb{R}^n$.*

δ΄. *Κάθε διάνυσμα $x \in \mathbb{R}^n$, γράφεται με μοναδικό τρόπο ως άθροισμα ενός διανύσματος του V και ενός διανύσματος του W.*

Απόδειξη. 1. Θεωρούμε μια βάση $v_1, \ldots, v_k$ του V και τον πίνακα A που έχει ως γραμμές τα διανύσματα $v_1, \ldots, v_k$. Τότε V είναι ο χώρος γραμμών του A, και ο μηδενοχώρος του A είναι ίσος με το ορθογώνιο συμπλήρωμα του V, $\mathcal{N}(A) = W$. Άρα $\dim W = \dim \mathcal{N}(A) = n - k$.

Έστω τώρα διάνυσμα $x \in V \cap W$, $x = (x_1, x_2, \ldots, x_n)$. Αφού $x \in V$ και $x \in W = V^\perp$, το x είναι ορθογώνιο στον εαυτό του, $xx^T = 0$. Δηλαδή $x_1^2 + x_2^2 + \cdots x_n^2 = 0$ και συνεπώς $x = 0$.

2. Από το Θεώρημα 1.4, το ορθογώνιο συμπλήρωμα του $W = \mathcal{N}(A)$ είναι ο $\mathcal{R}(A^T) = V$.

3. Το σύνολο $\{v_1, \ldots, v_k, w_1, \ldots, w_{n-k}\}$ έχει n στοιχεία. Εάν δείξουμε οτι είναι γραμμικά ανεξάρτητα, τότε θα είναι βάση του $\mathbb{R}^n$. Υποθέτουμε οτι $a_1 v_1 + \cdots + a_k u_k + b_1 w_1 + \cdots + b_{n-k} w_{n-k}) = 0$. Θέλουμε να δείξουμε οτι τότε όλα τα a_i και b_j είναι 0. Έχουμε

$$a_1 v_1 + \cdots + a_k u_k = y = -(b_1 w_1 + \cdots + b_{n-k} w_{n-k}) .$$

Αλλά η αριστερή πλευρά ανήκει στο V, η δεξιά πλευρά ανήκει στο W. Άρα το διάνυσμα y ανήκει στην τομή $V \cap W$, και συνεπώς y είναι το μηδενικό διάνυσμα $0 \in \mathbb{R}^n$. Άρα $a_1 u_1 + \cdots + a_k u_k = 0$, και αφού $u_1, \ldots, u_k$ είναι γραμμικά ανεξάρτητα, όλα τα a_i είναι 0. Παρόμοια, $b_1 w_1 + \cdots + b_{n-k} w_{n-k} = 0$, και αφού το $w_1, \ldots, w_{n-k}$ είναι γραμμικά ανεξάρτητα, όλα τα b_j είναι 0.

4. Αφού $\{u_1, \ldots, u_k, w_1, \ldots, w_{n-k}\}$ είναι βάση του $\mathbb{R}^n$, κάθε διάνυσμα $x \in \mathbb{R}^n$ γράφεται ως γραμμικός συνδυασμός

$$\begin{aligned} x &= a_1 u_1 + \cdots + a_k u_k + b_1 w_1 + \cdots + b_{n-k} w_{n-k} \\ &= x' + x'' \end{aligned}$$

όπου $x' = a_1 u_1 + \cdots + a_k u_k \in V$ και $x'' = b_1 w_1 + \cdots + b_{n-k} w_{n-k} \in W$.

Υποθέτουμε οτι ισχύει επίσης $x = \tilde{x} + \hat{x}$, όπου $\tilde{x} \in V$ και $\hat{x} \in W$. Τότε $x' + x'' = \tilde{x} + \hat{x}$, και συνεπώς $x' - \tilde{x} = \hat{x} - x''$, αλλά η αριστερή πλευρά ανήκει στο V, η δεξιά πλευρά ανήκει στο W, και όπως πιο πάνω, είναι και οι δύο μηδέν. Άρα $x' = \tilde{x}$ και $x'' = \hat{x}$.

$\square$

1.5 Η δράση του πίνακα A.

Τώρα μπορούμε να ολοκληρώσουμε την περιγραφή της γραμμικής απεικόνισης T_A που πολλαπλασιάζει κάθε διάνυσμα του $\mathbb{R}^n$ με τον $m \times n$ πίνακα A.

Εάν $x \in \mathcal{N}(A)$, τότε $T_A(x) = 0$.

Εάν το x είναι ορθογώνιο στο μηδενοχώρο, τότε $x \in \mathcal{R}(A^T)$, και $T_A(x) \in \mathcal{R}(A)$. Αλλά το σημαντικό είναι οτι αυτή η απεικόνιση, από το χώρο γραμμών $\mathcal{R}(A^T)$ στο χώρο στηλών $\mathcal{R}(A)$, είναι αμφιμονοσήμαντη.

Πρόταση 1.6 *Για κάθε διάνυσμα $y \in \mathcal{R}(A)$, υπάρχει ένα, και μόνον ένα, διάνυσμα $x \in \mathcal{R}(A^T)$, τέτοιο ώστε $Ax = y$.*

Απόδειξη. Αφού το y ανήκει στο χώρο στηλών, υπάρχει $u \in \mathbb{R}^n$ τέτοιο ώστε $Au = y$. Από την Πρόταση 1.5, υπάρχουν μοναδικά διανύσματα $x \in \mathcal{R}(A^T)$ και $w \in \mathcal{N}(A)$, τέτοια ώστε $u = x + w$. Αλλά $Aw = 0$, άρα $Ax = Au = y$. Εάν υπάρχει άλλο διάνυσμα $x' \in \mathcal{R}(A^T)$ με $Ax' = y$, τότε $x - x' \in \mathcal{N}(A)$. Αλλά αφού $\mathcal{R}(A^T)$ είναι διανυσματικός υπόχωρος, $x - x' \in \mathcal{R}(A^T)$. Συνεπώς $x - x' = 0$, και έχουμε μοναδικότητα.

$\square$

Εάν $u \in \mathbb{R}^n$ γράφεται ως $u = x + w$, με $x \in \mathcal{R}(A^T)$ και $w \in \mathcal{N}(A)$, ορίζουμε $P : \mathbb{R}^n \longrightarrow \mathcal{R}(A^T)$ με $P(u) = x$ και $E : \mathcal{R}(A) \longrightarrow \mathbb{R}^m$ με $E(y) = y$. Συμβολίζουμε $L : \mathcal{R}(A^T) \longrightarrow \mathcal{R}(A)$ τον ισομορφισμό από το χώρο γραμμών στο χώρο στηλών του A. Τότε η απεικόνιση T_A παραγοντοποιείται ως σύνθεση του επιμορφισμού P, του ισομορφισμού L και του μονομορφισμού E.

$$T_A = E \circ L \circ P\,.$$

Ανακεφαλαιώνουμε την περιγραφή της δράσης του πολλαπλασιασμού με ένα πίνακα.

Εάν A είναι ένας $m \times n$ πίνακας τάξεως r, και $T_A : \mathbb{R}^n \to \mathbb{R}^m$ η γραμμική απεικόνιση $x \mapsto Ax$, τότε

α′. Ο χώρος γραμμών $\mathcal{R}(A^T)$ είναι υπόχωρος του $\mathbb{R}^n$, διάστασης r.

β′. Ο χώρος στηλών $\mathcal{R}(A)$ είναι υπόχωρος του $\mathbb{R}^m$, διάστασης r.

γ′. Ο μηδενοχώρος $\mathcal{N}(A)$ είναι υπόχωρος του $\mathbb{R}^n$ διάστασης $n - r$.

δ′. Ο αριστερός μηδενόχωρος $\mathcal{N}(A^T)$ είναι υπόχωρος του $\mathbb{R}^m$ διάστασης $m - r$.

ε′. Ο μηδενοχώρος είναι το ορθογώνιο συμπλήρωμα του χώρου γραμμών, $\mathcal{N}(A) = \mathcal{R}(A^T)^\perp$ και $\mathbb{R}^n = \mathcal{R}(A^T) \oplus \mathcal{N}(A)$.

ϛ′. Ο αριστερός μηδενοχώρος είναι το ορθογώνιο συμπλήρωμα του χώρου στηλών, $\mathcal{N}(A^T) = \mathcal{R}(A)^\perp$ και $\mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$.

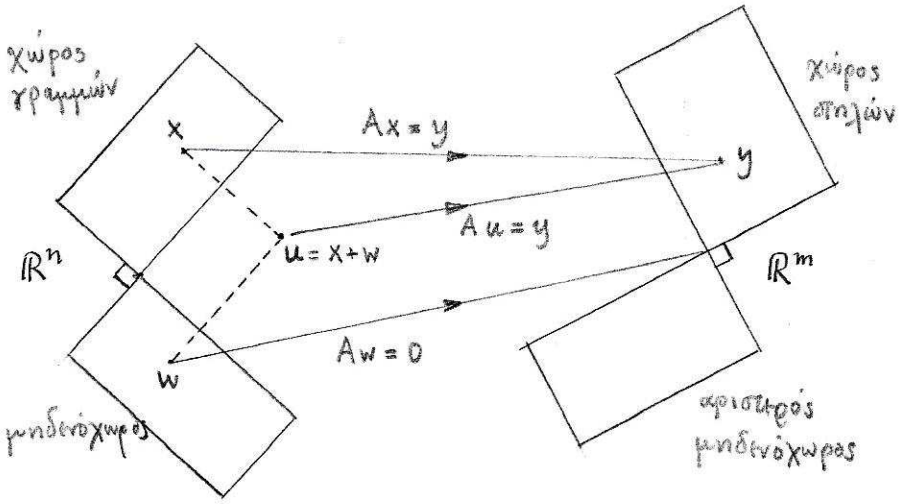
ζ′. Η γραμμική απεικόνιση T_A απεικονίζει το διανυσματικό υπόχωρο $\mathcal{R}(A^T)$ του $\mathbb{R}^n$ αμφιμονοσήμαντα στο διανυσματικό υπόχωρο $\mathcal{R}(A)$ του $\mathbb{R}^m$.

η′. Η γραμμική απεικόνιση T_{A^T} απεικονίζει το διανυσματικό υπόχωρο $\mathcal{R}(A)$ του $\mathbb{R}^m$ αμφιμονοσήμαντα στο διανυσματικό υπόχωρο $\mathcal{R}(A^T)$ του $\mathbb{R}^n$.

ϑ′. Η γραμμική απεικόνιση T_A παραγοντοποιείται ως σύνθεση ενός επιμορφισμού $P : \mathbb{R}^n \longrightarrow \mathcal{R}(A^T)$, ενός ισομορφισμού $L : \mathcal{R}(A^T) \longrightarrow \mathcal{R}(A)$ και ενός μονομορφισμού $E : \mathcal{R}(A) \longrightarrow \mathbb{R}^m$, $T_A = E \circ L \circ P$.

Προσέξτε οτι οι δύο αμφιμονοσήμαντες απεικονίσεις στα ζ′ και η′ δεν είναι υποχρεωτικά αντίστροφες η μία της άλλης.

Αυτή η εικόνα περιγράφεται παραστατικά στο Σχήμα 1.1.



Σχήμα 1.1: Η δράση του $m \times n$ πίνακα A.

1.6 Ασκήσεις

Άσκηση 1.1 Βρείτε ένα διάνυσμα x ορθογώνιο στο χώρο γραμμών του A, ένα διάνυσμα ορθογώνιο στο χώρο στηλών, και ένα διάνυσμα ορθογώνιο στο μηδενοχώρο:

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 3 & 6 & 4 \end{bmatrix}.$$

Άσκηση 1.2 Βρείτε όλα τα διανύσματα του $\mathbb{R}^3$ που είναι ορθογώνια στο $(1,\,1,\,1)$ και στο $(1,\,-1,\,0)$.

Άσκηση 1.3 Δύο ευθείες στο επίπεδο είναι ορθογώνιες όταν το γινόμενο των κλίσεων τους είναι -1. Εφαρμόστε αυτό το κριτήριο στις ευθείες που παράγονται από τα διανύσματα $x = (x_1,\, x_2)$ και $y = (y_1,\, y_2)$, οι οποίες έχουν κλίσεις x_2/x_1 και y_2/y_1, για να βρείτε το κριτήριο ορθογωνιότητας των διανυσμάτων, $x^T y = 0$.

Άσκηση 1.4 Πώς γνωρίζουμε οτι η i γραμμή ενός αντιστρέψιμου πίνακα B είναι ορθογώνια στην j στήλη του B^{-1}, εάν $i \neq j$;

Άσκηση 1.5 Δείξτε οτι το διάνυσμα $x - y$ είναι ορθογώνιο στο $x + y$ εάν και μόνον εάν $||x|| = ||y||$. Ποιά ιδιότητα των ρόμβων εκφράζει αυτό το αποτέλεσμα;

Άσκηση 1.6 Βρείτε μία βάση για το ορθογώνιο συμπλήρωμα του χώρου γραμμών του A:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{bmatrix}.$$

Διαχωρίστε το $x = (3,\, 3,\, 3)$ σε μία συνιστώσα στο χώρο γραμμών, και σε μία συνιστώσα στο μηδενοχώρο του A.

Άσκηση 1.7 Θεωρήστε τον υποχώρο S του $\mathbb{R}^4$ που περιέχει όλα τα διανύσματα που ικανοποιούν την $x_1 + x_2 + x_3 + x_4 = 0$. Βρείτε μία βάση για το χώρο $S^\perp$, που περιέχει όλα τα διανύσματα που είναι ορθογώνια στον S.

Άσκηση 1.8 Για να βρείτε το ορθογώνιο συμπλήρωμα του επιπέδου που παράγεται από τα διανύσματα $(1,\, 1,\, 2)$ και $(1,\, 2,\, 3)$, θεωρήστε αυτά τα διανύσματα ως γραμμές του πίνακα A, και λύστε την εξίσωση $Ax = 0$. Θυμηθείτε οτι το συμπλήρωμα είναι ολόκληρη ευθεία.

Άσκηση 1.9 Εάν V και W είναι ορθογώνιοι υπόχωροι, δείξτε οτι το μόνο κοινό διάνυσμα είναι το μηδενικό: $V \cap W = \{0\}$.

Άσκηση 1.10 Βρείτε έναν πίνακα του οποίου ο χώρος γραμμών περιέχει το $(1,\, 2,\, 1)$ και ο μηδενοχώρος περιέχει το $(1,\, -2,\, 1)$, ή δείξτε οτι δεν υπάρχει τέτοιος πίνακας.

Άσκηση 1.11 Κατασκευάστε μία ομογενή εξίσωση σε τρείς αγνώστους, της οποίας οι λύσεις είναι οι γραμμικοί συνδυασμοί των διανυσμάτων $(1,\, 1,\, 2)$ και $(1,\, 2,\, 3)$. Αυτό είναι το αντίστροφο της προηγούμενης άσκησης, αλλά τα δύο προβλήματα είναι ουσιαστικά τα ίδια.

Άσκηση 1.12 Σχεδιάστε στο επίπεδο τους τέσσερεις θεμελιώδεις υπόχωρους των πινάκων

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \text{ και } B = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}.$$

Άσκηση 1.13 Σχεδιάστε τους τέσσερεις θεμελιώδεις υπόχωρους του A, και βρείτε τις συνιστώσες του x στο χώρο γραμμών και στο μηδενοχώρο του A, όπου

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ και } x = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Άσκηση 1.14 Σε κάθε περίπτωση, κατασκευάστε έναν πίνακα A με τη ζητούμενη ιδιότητα ή εξηγήστε γιατί αυτό δεν είναι δυνατό

α΄. Ο χώρος στηλών περιέχει τα διανύσματα $(1, 2, -3)$ και $(2, -3, 5)$, και ο μηδενοχώρος περιέχει το $(1, 1, 1)$.

β΄. Ο χώρος γραμμών περιέχει τα $(1, 2, -3)$ και $(2, -3, 5)$ και ο μηδενοχώρος περιέχει το $(1, 1, 1)$.

γ΄. Η εξίσωση $Ax = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ έχει λύση, και $A^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

δ΄. Το άθροισμα των στηλών είναι το διάνυσμα $(0, 0, 0)$, και το άθροισμα των γραμμών είναι το διάνυσμα $(1, 1, 1)$.

Άσκηση 1.15 Υποθέστε οτι ο υπόχωρος S παράγεται από τα διανύσματα $(1, 2, 2, 3)$ και $(1, 3, 3, 2)$. Βρείτε δύο διανύσματα που παράγουν τον υπόχωρο $S^{\perp}$. Αυτό ισοδυναμεί με το να λύσετε την εξίσωση $Ax = 0$ για κάποιο πίνακα A. Ποιός είναι ο A;

Άσκηση 1.16 Δείξτε οτι εάν ο υπόχωρος S περιέχεται στον υπόχωρο V, τότε ο $S^{\perp}$ περιέχει τον $V^{\perp}$.

Άσκηση 1.17 Βρείτε το ορθογώνιο συμπλήρωμα $S^{\perp}$ όταν

α΄. S είναι ο μηδενικός υπόχωρος του $\mathbb{R}^3$.

β΄. S είναι ο υπόχωρος που παράγεται από το $(1, 1, 1)$.

γ΄. S είναι ο υπόχωρος που παράγεται από τα $(2, 0, 0)$ και $(0, 0, 3)$.

Άσκηση 1.18 Κατασκευάστε έναν 3×3 πίνακα A, χωρίς μηδενικά στοιχεία, του οποίου οι στήλες είναι ανά δύο κάθετες. Υπολογίστε το γινόμενο $A^T A$. Γιατί είναι το γινόμενο διαγώνιος πίνακας;

Άσκηση 1.19 Βρείτε έναν πίνακα που περιέχει το διάνυσμα $u = (1, 2, 3)$ στο χώρο γραμμών και στο χώρο στηλών. Βρείτε έναν άλλο πίνακα που περιέχει το u στο μηδενοχώρο και στο χώρο στηλών. Σε ποιά ζεύγη υποχώρων ενός πίνακα δεν μπορεί να περιέχεται το u;

Άσκηση 1.20 Δείξτε οτι

$$V^{\perp} = \{w \in \mathbb{R}^n \,:\, \forall v \in V,\, w^T v = 0\}$$

είναι πράγματι γραμμικός υπόχωρος του $\mathbb{R}^n$, δηλαδή οτι είναι ένα μη κενό υποσύνολο του $\mathbb{R}^n$, κλειστό ως προς γραμμικούς συνδυασμούς.

Άσκηση 1.21 Δείξτε οτι εάν V είναι γραμμικός υπόχωρος του $\mathbb{R}^n$ και $W = V^{\perp}$, τότε $W^{\perp} = V$, δηλαδή οτι εάν ο W είναι το ορθογώνιο συμπλήρωμα του V, τότε και ο V είναι το ορθογώνιο συμπλήρωμα του W.

Άσκηση 1.22 Αποδείξτε οτι η εξίσωση $Ax = b$ έχει λύση εάν και μόνον εάν $y^T b = 0$ για κάθε y που ικανοποιεί $y^T A = 0$.

1.7 Βέλτιστες λύσεις και Προβολές

Επιστρέφουμε ακόμη μία φορά στην εξίσωση $Ax = b$. Έχουμε δει οτι η εξίσωση έχει λύσεις μόνον όταν το διάνυσμα b ανήκει στο χώρο στηλών του πίνακα A. Συχνά όμως θέλουμε να βρούμε την *καλύτερη δυνατή* λύση της εξίσωσης, ακόμη και όταν το b δεν ανήκει στον $\mathcal{R}(A)$. Αυτό συμβαίνει συχνά στην ανάλυση πειραματικών δεδομένων, όπου για να περιορίσουμε την πιθανότητα τυχαίου σφάλματος, παίρνουμε περισσότερες μετρήσεις. Το αποτέλεσμα είναι να έχουμε ένα σύστημα με αρκετά περισσότερες εξισώσεις παρά αγνώστους, όπου δεν περιμένουμε να υπάρχει ακριβής λύση.

Εάν αντικαταστήσουμε το b με ένα διάνυσμα b' του χώρου στηλών $\mathcal{R}(A)$ τότε η εξίσωση $Ax = b'$ έχει λύση. Μπορούμε να βρούμε μια *βέλτιστη λύση* της εξίσωσης, εάν αντικαταστήσουμε το διάνυσμα b με το διάνυσμα του χώρου στηλών του A που είναι πλησιέστερο στο b από κάθε άλλο διάνυσμα του χώρου στηλών. Αυτό το διάνυσμα είναι η *ορθογώνια προβολή του b στο χώρο στηλών*.

Εάν συμβολίσουμε p την ορθογώνια προβολή του b στο χώρο στηλών, έχουμε μία νέα εξίσωση

$$A\hat{x} = p.$$

Οι λύσεις αυτής της εξίσωσης ονομάζονται *βέλτιστες λύσεις ελαχίστων τετραγώνων* της αρχικής εξίσωσης $Ax = b$, (δείτε την Άσκηση 1.7).

Παράδειγμα 1.4 Υποθέτουμε οτι μελετάμε την εξάρτηση μίας ποσότητας b από μία ποσότητα a, και αναμένουμε οτι η b είναι ανάλογη προς την a. Θέλουμε να βρούμε τον σταθερό λόγο λ για τον οποίο

$$b = \lambda a\,.$$

Υποθέτουμε οτι οι πειραματικές μετρήσεις δίδουν τις τιμές b_1 για $a = 2$, b_2 για $a = 3$ και b_3 για $a = 4$. Για να βρούμε το λ θεωρούμε τρείς εξισώσεις με ένα άγνωστο.

$$\begin{aligned} 2x &= b_1 \\ 3x &= b_2 \\ 4x &= b_3\,. \end{aligned}$$

Όμως αυτό το σύστημα έχει λύση μόνο όταν το διάνυσμα (b_1, b_2, b_3) είναι ένα πολλαπλάσιο του $(2, 3, 4)$. Η εξίσωση

$$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \tag{1.3}$$

έχει λύση μόνον όταν το (b_1, b_2, b_3) ανήκει στο χώρο στηλών. Για κάθε τιμή του x ορίζουμε το *σφάλμα*

$$\varepsilon = \|ax - b\| = \sqrt{(2x - b_1)^2 + (3x - b_2)^2 + (4x - b_3)^2}\,,$$

το οποίο μηδενίζεται μόνο όταν x αποτελεί λύση της εξίσωσης 1.3. Στην περίπτωση που η εξίσωση 1.3 δεν έχει λύση, θεωρούμε την *βέλτιστη λύση*, την τιμή του x η οποία κάνει το σφάλμα ε όσο το δυνατόν μικρότερο. Αυτό συμβαίνει όταν το διάνυσμα ax είναι ίσο με την ορθογώνια προβολή του διανύσματος b στο χώρο στηλών, δηλαδή όταν $ax - b$ είναι κάθετο στο a.

> **Δραστηριότητα 1.6** Στο Παράδειγμα 1.4, θέτουμε $b = (4, 6, 9)$. Σχεδιάστε στο επίπεδο τα σημεία με συντεταγμένες $(2, 4)$, $(3, 6)$, $(4, 9)$. Μπορείτε να σχεδιάσετε μία ευθεία που να περνάει από τα 4 σημεία $(0, 0)$, $(2, 4)$, $(3, 6)$, $(4, 9)$. Υπολογίστε την παράγωγο $\frac{d}{dx}(\varepsilon^2)$, και δείξτε οτι μηδενίζεται ακριβώς όταν $ax - b$ είναι κάθετο στο a.

> **Δραστηριότητα 1.7** Υπολογίστε την παράγωγο $\frac{d}{dx}(\varepsilon^2)$, και δείξτε οτι μηδενίζεται ακριβώς όταν $(ax - b)^T a = 0$, δηλαδή όταν $ax - b$ είναι κάθετο στο a.

Προβολή σε ευθεία

Ας εξετάσουμε πρώτα την προβολή σε μία ευθεία. Θεωρούμε τα διανύσματα a και b στο επίπεδο. Στην ''Εισαγωγή στη Γραμμική Άλγεβρα'' είδαμε (Κεφάλαιο 4) την προβολή του επιπέδου $\mathbb{R}^2$ στον θ-άξονα, δηλαδή στην ευθεία των διανυσμάτων που είναι συγγραμμικά με το $(\cos\theta, \sin\theta)$. Τώρα θέλουμε να υπολογίσουμε την προβολή ενός σημείου b του $\mathbb{R}^n$ πάνω στην ευθεία των διανυσμάτων που είναι συγγραμμικά με το $a \in \mathbb{R}^n$. Το διάνυσμα προβολής p χαρακτηρίζεται από τις ακόλουθες ιδιότητες:

α'. Το p είναι συγγραμμικό με το a, δηλαδή $p = \hat{x}a$ για κάποιο αριθμό $\hat{x} \in \mathbb{R}$.

β'. Η διαφορά $b - p$ είναι ορθογώνια στο a, δηλαδή $a^T(b - p) = 0$.

Από αυτές τις ιδιότητες λαμβάνουμε την εξίσωση

$$a^T(b - \hat{x}a) = 0$$

την οποία μπορούμε να λύσουμε για να βρούμε το $\hat{x}$:

$$\hat{x} = \frac{a^T b}{a^T a}.$$

Συνεπώς το διάνυσμα προβολής p είναι

$$p = \hat{x}a = \frac{a^T b}{a^T a} a\,,$$

και η βέλτιστη λύση ελαχίστων τετραγώνων της εξίσωσης $Ax = b$ είναι

$$\hat{x} = \frac{a^T b}{a^T a}\,.$$

Θέλουμε να εκφράσουμε την προβολή ως μία γραμμική απεικόνιση από τον $\mathbb{R}^n$ στον $\mathbb{R}^n$, η οποία απεικονίζει κάθε διάνυσμα του $\mathbb{R}^n$ στην ευθεία $V = \{ta : t \in \mathbb{R}\}$, και να βρούμε τον αντίστοιχο πίνακα. Στον προηγούμενο υπολογισμό μπορούμε να αντιστρέψουμε τη διάταξη του διανύσματος a και του αριθμού $\hat{x}$:

$$p = a\hat{x} = a\frac{a^T b}{a^T a} = \frac{1}{a^T a}\, a(a^T b)\,,$$

και να εφαρμόσουμε την προσεταιριστική ιδιότητα:

$$p = \frac{1}{a^T a}\,(aa^T)b\,.$$

Παρατηρήστε οτι $a^T a$ είναι θετικός αριθμός, το τετράγωνο του μήκους του a, ενώ aa^T είναι τετραγωνικός πίνακας.

Τον πίνακα

$$P = \frac{1}{a^T a}\, aa^T$$

ονομάζουμε **πίνακα προβολής**. Για να προβάλουμε το διάνυσμα $b \in \mathbb{R}^n$ στην ευθεία που ορίζει το διάνυσμα a, αρκεί να το πολλαπλασιάσουμε με τον πίνακα P.

Παράδειγμα 1.5 Συνεχίζουμε το Παράδειγμα 1.4, με $b = (4, 6, 9)$, δηλαδή θεωρούμε το σύστημα

$$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} x = \begin{bmatrix} 4 \\ 6 \\ 9 \end{bmatrix}.$$

Αυτό δεν έχει λύση, αφού το διάνυσμα (4, 6, 9) δεν ανήκει στο χώρο που παράγει το (2, 3, 4). Η βέλτιστη λύση είναι $\hat{x}$, τέτοια ώστε

$$\begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \left(\begin{bmatrix} 4 \\ 6 \\ 9 \end{bmatrix} - \hat{x} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \right) = 0\,,$$

δηλαδή

$$\hat{x} = \frac{(2,\,3,\,4)\cdot(4,\,6,\,9)}{2^2+3^2+4^2} = \frac{62}{29}\,.$$

Συμπεραίνουμε οτι η βέλτιστη τιμή για το λ που προκύπτει από τα 3 σημεία (2, 4), (3, 6) και (4, 9) είναι $\lambda = \frac{62}{29}$.

Προβολή σε υπόχωρο

Θεωρούμε τώρα το πρόβλημα σε περισσότερες διαστάσεις. Θέλουμε να προβάλουμε το διάνυσμα b σε ένα υπόχωρο V διάστασης k μέσα στον $\mathbb{R}^m$. Μπορούμε για ευκολία να υποθέσουμε οτι $k=2$ και $m=3$, χωρίς ουσιαστική διαφορά στη διαδικασία. Θεωρούμε λοιπόν δύο διανύσματα a_1 και a_2 του $\mathbb{R}^3$, τα οποία αποτελούν βάση του V, και τον $m \times k$ πίνακα A με στήλες τα διανύσματα a_i, έτσι ώστε $V = \mathcal{R}(A)$. Αφού η προβολή p βρίσκεται στο χώρο στηλών του A, έχουμε

$$p = A\hat{x}$$

για κάποιο $\hat{x} \in \mathbb{R}^k$. Αφού η προβολή είναι ορθογώνια, το διάνυσμα $b - A\hat{x}$ είναι ορθογώνιο στο χώρο στηλών του A, και από το Θεώρημα 1.4 ανήκει στον αριστερό μηδενοχώρο του A:

$$A^T(b - A\hat{x}) = 0\,.$$

Έτσι έχουμε την εξίσωση

$$A^T A\hat{x} = A^T b\,, \tag{1.4}$$

για τη βέλτιστη λύση $\hat{x}$, η οποία έχει μοναδική λύση για κάθε $b \in \mathbb{R}^m$ όταν A^TA είναι μη ιδιόμορφος πίνακας.

Παράδειγμα 1.6 Συνεχίζουμε το Παράδειγμα 1.4, με $b = (4,\,6,\,9)$, αλλά τώρα υποθέτουμε οτι η σχέση μεταξύ των ποσοτήτων a και b είναι

$$b = \lambda a + \mu\,.$$

Με τα ίδια δεδομένα, (2, 4), (3, 6) και (4, 9), έχουμε τρείς εξισώσεις με δύο αγνώστους για να βρούμε τα λ και μ:

$$\begin{array}{rcrcl} 2x_1 & + & x_2 & = & 4 \\ 3x_1 & + & x_2 & = & 6 \\ 4x_1 & + & x_2 & = & 9\,, \end{array}$$

τις οποίες γράφουμε ως ένα σύστημα

$$Ax = \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 9 \end{bmatrix}.$$

Το διάνυσμα (4, 6, 9) δεν ανήκει στο χώρο στηλών του πίνακα A, και το σύστημα δεν έχει λύση. Το σφάλμα $\varepsilon = \|b - Ax\|$ ελαχιστοποιήται για την τιμή $\hat{x}$ του $x = (x_1, x_2)$ για την οποία το διάνυσμα $b - A\hat{x}$ είναι ορθογώνιο στο χώρο στηλών του A. Έτσι έχουμε την εξίσωση *ελαχίστων τετραγώνων*:

$$A^T(b - A\hat{x}) = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 1 & 1 \end{bmatrix} \left(\begin{bmatrix} 4 \\ 6 \\ 9 \end{bmatrix} - \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} \right) = 0\,,$$

δηλαδή

$$\begin{bmatrix} 29 & 9 \\ 9 & 3 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 62 \\ 19 \end{bmatrix},$$

η οποία έχει μοναδική λύση

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 3 & -9 \\ -9 & 29 \end{bmatrix} \begin{bmatrix} 62 \\ 19 \end{bmatrix} = \begin{bmatrix} \frac{15}{6} \\ \frac{7}{6} \end{bmatrix}.$$

Άρα η βέλτιστη ευθεία που καθορίζεται από τα σημεία (2, 4), (3, 6) και (4, 9) έχει εξίσωση

$$6b = 15a + 7\,.$$

Στην γενική περίπτωση, που ο πίνακας A έχει περισσότερες από 2 στήλες, θα ήταν πιο οικονομικό να λύσουμε την εξίσωση 1.4 με απαλοιφή Gauss παρά να υπολογίσουμε τον αντίστροφο πίνακα.

Εάν ο πίνακας A^TA είναι αντιστρέψιμος, μπορούμε να υπολογίσουμε την προβολή p από την 1.4. Η προβολή του b στον υπόχωρο $V = \mathcal{R}(A)$ είναι

$$p = A\hat{x} = A(A^TA)^{-1}A^Tb\,.$$

Ο πίνακας προβολής που όταν τον πολλαπλασιάσουμε με το b δίδει την προβολή p, είναι

$$P = A(A^TA)^{-1}A^T.$$

Σε αυτή την έκφραση, A είναι $m \times k$ πίνακας, οπότε A^TA είναι τετραγωνικός $k \times k$ πίνακας, και P είναι $m \times m$ πίνακας.

Προσέξτε οτι σε αυτή την περίπτωση δεν μπορούμε να αλλάξουμε τη διάταξη των παραγόντων, αφού αυτοί οι πίνακες δεν μετατίθενται. Επίσης, αφού A δεν είναι τετραγωνικός πίνακας, δεν μπορούμε να γράψουμε τον αντίστροφο $(A^TA)^{-1}$ ως γινόμενο αντιστρόφων πινάκων. Αν συγκρίνουμε με την περίπτωση της προβολής σε ευθεία, όπου $k = 1$, βλέπουμε οτι ο $m \times 1$ πίνακας A είναι το διάνυσμα a, και ο αντιστρέψιμος $k \times k$ πίνακας $A^T A$ είναι ο θετικός αριθμός $a^T a$, με αντίστροφο $\frac{1}{a^T a}$. Αυτός μετατίθεται με τον πίνακα A, και συνεπώς μπορούμε να γράψουμε

$$a(a^T a)^{-1} a^T = \frac{1}{a^T a}\, a a^T.$$

Θα δείξουμε οτι η υπόθεση πως $A^T A$ είναι αντιστρέψιμος ικανοποιείται πάντα όταν οι στήλες του A είναι γραμμικά ανεξάρτητες, όπως στην περίπτωση που αποτελούν βάση του υπόχωρου V.

Λήμμα 1.7 *Ο πίνακας $A^T A$ έχει τον ίδιο μηδενόχωρο με τον A.*

Απόδειξη. Είναι προφανές οτι εάν $Ax = 0$ τότε $A^T Ax = 0$, δηλαδή οτι $\mathcal{N}(A) \subseteq \mathcal{N}(A^T A)$.

Για να δείξουμε τον αντίστροφο εγκλεισμό θεωρούμε x τέτοιο ώστε $A^T Ax = 0$, οπότε

$$x^T(A^T Ax) = 0.$$

Αλλά $x^T(A^T Ax) = (x^T A^T)Ax = (Ax)^T Ax = ||Ax||^2$.

Άρα το διάνυσμα Ax έχει μηδενικό μήκος, και συνεπώς $Ax = 0$, δηλαδή $x \in \mathcal{N}(A)$.

□

Συγκεντρώνουμε τα προηγούμενα αποτελέσματα σε μία Πρόταση.

Πρόταση 1.8 *Εάν A είναι $m \times k$ πίνακας με γραμμικά ανεξάρτητες στήλες, τότε*

α΄. Ο πίνακας $A^T A$ είναι αντιστρέψιμος.

β΄. Η ορθογώνια προβολή του $\mathbb{R}^m$ στο χώρο στηλών του πίνακα A είναι

$$P(w) = A(A^T A)^{-1} A^T w\,.$$

γ΄. Η βέλτιστη λύση ελαχίστων τετραγώνων της εξίσωσης $Ax = b$, για κάθε $b \in \mathbb{R}^m$, είναι η λύση της εξίσωσης

$$A^T A\hat{x} = A^T b\,.$$

Κλείνουμε αυτή την παράγραφο με έναν χαρακτηρισμό των πινάκων ορθογώνιας προβολής σε υπόχωρο του $\mathbb{R}^m$.

Πρόταση 1.9 *Ένας $m \times m$ πίνακας P είναι πίνακας προβολής σε ένα υπόχωρο του $\mathbb{R}^m$ εάν και μόνον εάν P είναι συμμετρικός και $P^2 = P$.*

Απόδειξη. Έστω V ένας υπόχωρος του $\mathbb{R}^m$, και A ο πίνακας με στήλες τα διανύσματα μίας βάσης του V. Τότε ο πίνακας προβολής στον υπόχωρο V είναι ο $P = A(A^T A)^{-1} A^T$. Εύκολα ελέγχουμε οτι $P^2 = P$,

$$\begin{aligned} P^2 &= A(A^T A)^{-1} A^T A (A^T A)^{-1} A^T \\ &= A(A^T A)^{-1} A^T \end{aligned}$$

$$= P.$$

Ο ανάστροφος του P είναι ο πίνακας

$$\begin{aligned} P^T &= (A(A^TA)^{-1}A^T)^T \\ &= (A^T)^T\left((A^TA)^{-1}\right)^T A^T \\ &= A\left((A^TA)^T\right)^{-1} A^T \\ &= A(AA^T)^{-1}A^T \\ &= P \end{aligned}$$

Αντιστρόφως, εάν ο $m \times m$ πίνακας P ικανοποιεί τις σχέσεις $P^2 = P$ και $P = P^T$, θα δείξουμε οτι P είναι ο πίνακας προβολής στο χώρο στηλών του. Προφανώς, για κάθε $b \in \mathbb{R}^m$, Pb ανήκει στο χώρο στηλών του P. Για να δείξουμε οτι Pb είναι η προβολή του b στον υπόχωρο $V = \mathcal{R}(P)$ αρκεί να δείξουμε οτι $b - Pb$ είναι ορθογώνιο στον V.

Έστω v διάνυσμα του V. Τότε v είναι γραμμικός συνδυασμός των στηλών του P, δηλαδή υπάρχει $c \in \mathbb{R}^m$ τέτοιο ώστε $v = Pc$, και έχουμε

$$\begin{aligned} (b - Pb)^T v &= (b - Pb)^T Pc \\ &= (b^T - b^TP^T)Pc \\ &= b^T(I - P^T)Pc \\ &= b^T(P - P^TP)c. \end{aligned}$$

Αλλά $P^T = P$ και $P^2 = P$, άρα $P - P^TP = P - P = 0$.

□

1.8 Ασκήσεις

Άσκηση 1.23 Βρείτε την προβολή του διανύσματος (7, 4) πάνω στον υπόχωρο που παράγεται από το διάνυσμα (1, 2).

Άσκηση 1.24 Βρείτε τον πίνακα προβολής που αντιστοιχεί στην προβολή των διανυσμάτων του επιπέδου $\mathbb{R}^2$ πάνω στην ευθεία $3x - 2y = 0$.

Άσκηση 1.25 Βρείτε τον πίνακα προβολής P_1 στην ευθεία με διεύθυνση $a = (1, 3)$, καθώς και τον πίνακα προβολής P_2 στην ευθεία που είναι κάθετη στο a. Υπολογίστε τους πίνακες $P_1 + P_2$ και P_1P_2. Εξηγήστε το αποτέλεσμα.

Άσκηση 1.26 Στον χώρο $\mathbb{R}^n$, ποιά γωνία σχηματίζει το διάνυσμα $(1, 1, \dots, 1)$ με τους άξονες συντεταγμένων; Βρείτε τον πίνακα προβολής σε αυτό το διάνυσμα.

Άσκηση 1.27 Ποιό πολλαπλάσιο του $a = (1, 1, 1)$ είναι πλησιέστερο στο σημείο $b = (2, 4, 4)$; Βρείτε επίσης το σημείο στην ευθεία με διεύθυνση b που είναι πλησιέστερο στο a.

Άσκηση 1.28 Δείξτε οτι ο πίνακας προβολής $P = \frac{1}{a^T a}\, aa^T$ είναι συμμετρικός και ικανοποιεί τη σχέση $P^2 = P$.

Άσκηση 1.29 Ποιος πίνακας P προβάλει κάθε σημείο του $\mathbb{R}^3$ στην ευθεία όπου τέμνονται τα επίπεδα $x + y + t = 0$ και $x - t = 0$;

Άσκηση 1.30 Για τα ακόλουθα διανύσματα, σχεδιάστε στο καρτεσιανό επίπεδο την προβολή του b στο a, και στη συνέχεια υπολογίστε την προβολή, από την έκφραση $p = \hat{x}a$:

α΄. $b = \begin{bmatrix} \cos\vartheta \\ \sin\vartheta \end{bmatrix}$ και $a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

β΄. $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ και $a = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Άσκηση 1.31 Υπολογίστε την προβολή του b στην ευθεία με διεύθυνση a, και ελέγξτε οτι το διανυσματικό σφάλμα $e = b - p$ είναι ορθογώνιο στο a:

α΄. $b = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ και $a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

β΄. $b = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ και $a = \begin{bmatrix} -1 \\ -3 \\ -1 \end{bmatrix}$

Άσκηση 1.32 Έστω a διάνυσμα του $\mathbb{R}^n$ και έστω P ο πίνακας προβολής του a. Δείξτε οτι το άθροισμα των διαγωνίων στοιχείων του P ισούται με 1.

Άσκηση 1.33 Έστω $a = (1, 2, -1, 3)$.

α΄. Βρείτε τον πίνακα προβολής P στο διάνυσμα a.

β΄. Βρείτε μια βάση του μηδενόχωρου $\mathcal{N}(P)$.

γ΄. Βρείτε ένα μη μηδενικό διάνυσμα v του $\mathbb{R}^4$ του οποίου η προβολή στο a να είναι το μηδενικό διάνυσμα.

Άσκηση 1.34 Βρείτε τη βέλτιστη λύση ελαχίστων τετραγώνων της εξίσωσης $Ax = b$, και υπολογίστε την προβολή $p = A\hat{x}$, εάν

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \text{ και } b = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Επαληθεύστε οτι το διανυσματικό σφάλμα $e = b - p$ είναι ορθογώνιο στις στήλες του A.

Άσκηση 1.35 Υπολογίστε το τετράγωνο του σφάλματος $\varepsilon^2 = ||Ax - b||^2$, και βρείτε τις μερικές παραγώγους του ε^2 ως προς u και v, εάν

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \; x = \begin{bmatrix} u \\ v \end{bmatrix} \text{ και } b = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}.$$

Θέσατε τις παραγώγους ίσες με μηδέν, και συγκρίνετε με τις εξισώσεις $A^T A\hat{x} = A^T b$, για να δείξετε οτι ο λογισμός και η γεωμετρία καταλήγουν στις ίδιες εξισώσεις. Υπολογίστε το $\hat{x}$ και την προβολή $p = A\hat{x}$. Γιατί είναι $p = b$;

Άσκηση 1.36 Βρείτε την προβολή του $b = (4, 3, 1, 0)$ πάνω στο χώρο στηλών του

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

Άσκηση 1.37 Βρείτε την βέλτιστη λύση ελαχίστων τετραγώνων $\hat{x}$, του συστήματος εξισώσεων $3x = 10$ και $4x = 5$. Ποιο είναι το τετράγωνο του σφάλματος ε^2 που ελαχιστοποιείται; Επαληθεύστε οτι το διανυσματικό σφάλμα $e = (10 - 3\hat{x}, 5 - 4\hat{x})$ είναι ορθογώνιο στη στήλη $(3, 4)$.

Άσκηση 1.38 Βρείτε την προβολή του b στο χώρο στηλών του A:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -2 & 4 \end{bmatrix} \quad , \quad b = \begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}.$$

Διαχωρήστε το b σε άθροισμα $p + q$, με p στο χώρο στηλών του A και q ορθογώνιο προς αυτόν. Σε ποιό θεμελιώδη υπόχωρο του A βρίσκεται το διάνυσμα q;

Άσκηση 1.39 Δείξτε οτι εάν ο πίνακας P ικανοποιεί τη σχέση $P = P^T P$, τότε P είναι πίνακας προβολής. Είναι ο μηδενικός πίνακας $P = 0$ πίνακας προβολής, και σε ποιο υπόχωρο;

Άσκηση 1.40 Τα διανύσματα $a_1 = (1,\, 1,\, 0)$ και $a_2 = (1,\, 1,\, 1)$ παράγουν ένα επίπεδο στο $\mathbb{R}^3$. Βρείτε τον πίνακα προβολής στο επίπεδο, και ένα μη μηδενικό διάνυσμα b το οποίο προβάλεται στο 0.

Άσκηση 1.41 Εάν V είναι ο υπόχωρος που παράγεται από τα $(1,\, 1,\, 0,\, 1)$ και $(0,\, 0,\, 1,\, 0)$ βρείτε

α΄. μία βάση για το ορθογώνιο συμπλήρωμα $V^\perp$.

β΄. τον πίνακα προβολής P στο V.

γ΄. το διάνυσμα στο V το οποίο είναι πλησιέστερο προς το $(0,\, 1,\, 0,\, -1) \in V^\perp$

Άσκηση 1.42 Εάν P είναι η προβολή στο χώρο στηλών του πίνακα A, ποιά είναι η προβολή στον αριστερό μηδενοχώρο του A;

Άσκηση 1.43 Εάν $P_\sigma = A(A^T A)^{-1} A^T$ είναι ο πίνακας προβολής στο χώρο στηλών του A, ποιός είναι ο πίνακας προβολής P_γ στο χώρο γραμμών του A;

Άσκηση 1.44 Θεωρούμε τον διανυσματικό υπόχωρο V του $\mathbb{R}^4$ που παράγεται από τα διανύσματα

$$(1,\, 2,\, 0,\, 3), \quad (2,\, 1,\, 1,\, 2) \quad (-1,\, 4,\, -2,\, 5)$$

α΄. Βρείτε το ορθογώνιο συμπλήρωμα $V^\perp$ του V.

β΄. Γράψτε το διάνυσμα $x = (-4,\, 15,\, 7,\, 8)$ ως άθροισμα $x = v + w$, όπου $v \in V$ και $w \in V^\perp$.

1.9 Ορθογώνιοι πίνακες

Θεωρούμε ένα διανυσματικό υπόχωρο $V \subseteq \mathbb{R}^n$ διάστασης $\dim V = k$, και μία βάση $v_1, \ldots, v_k$ του V. Τότε κάθε διάνυσμα $u \in V$ μπορεί να εκφραστεί ως γραμμικός συνδυασμός

$$u = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k.$$

Το εσωτερικό γινόμενο των διανυσμάτων u και $w = b_1 v_1 + b_2 v_2 + \cdots + b_k v_k$ είναι

$$\begin{aligned} u^T w &= (a_1 v_1^T + \cdots a_k v_k^T)(b_1 v_1 + \cdots + b_k v_k) \\ &= \sum_{i=1}^{k} \sum_{j=1}^{k} a_i b_j v_i^T v_j. \end{aligned}$$

Εάν υποθέσουμε οτι τα διανύσματα της βάσης είναι ανά δύο ορθογώνια μεταξύ τους, δηλαδή οτι $v_i^T v_j = 0$ εάν $i \neq j$, το εσωτερικό γινόμενο $u^T w$ γίνεται

$$\begin{aligned} u^T w &= \sum_{i=1}^{k} a_i b_i v_i^T v_i \\ &= \sum_{i=1}^{k} a_i b_i ||v_i||^2. \end{aligned}$$

Είναι φανερό οτι μία βάση από ορθογώνια διανύσματα μπορεί να απλοποιήσει σημαντικά τους υπολογισμούς. Σε μία τέτοια βάση, μόνο μία ακόμη βελτίωση μπορούμε να κάνουμε: το μήκος κάθε διανύσματος της βάσης να είναι $||v_i||^2 = 1$. Τότε το εσωτερικό γινόμενο των u και w λαμβάνει την απλούστερη δυνατή μορφή:

$$u^T w = \sum_{i=1}^{k} a_i b_i$$

Μία τέτοια βέλτιστη βάση την ονομάζουμε *ορθοκανονική*.

Ορισμός 1.3. Τα διανύσματα $q_1, q_2, \ldots, q_k$ είναι **ορθοκανονικά** εάν

$$q_i^T q_j = \begin{cases} 0 & \text{εάν} \quad i \neq j \\ \\ 1 & \text{εάν} \quad i = j \end{cases}$$

Μία βάση που αποτελείται από ορθοκανονικά διανύσματα ονομάζεται **ορθοκανονική βάση**. Ένας τετραγωνικός πίνακας του οποίου οι στήλες είναι ορθοκανονικά διανύσματα ονομάζεται **ορθογώνιος**.

Προσέξτε οτι ο όρος ορθογώνιος χρησιμοποιείται μόνο για τετραγωνικούς πίνακες. Ένας μη τετραγωνικός πίνακας με ορθοκανονικές στήλες δεν ονομάζεται ορθογώνιος.

> **Δραστηριότητα 1.8** Δείξτε οτι εάν $q_1, \ldots, q_k$ είναι μία ορθοκανονική βάση του V, τότε οι συντεταγμένες $a_1, \ldots, a_k$ του διανύσματος $u = a_1 q_1 + \cdots + a_k q_k$ είναι $a_i = q_i^T u$.

Το σημαντικότερο παράδειγμα ορθοκανονικής βάσης είναι η κανονική βάση $e_1, \ldots, e_n$ του $\mathbb{R}^n$. Ο ορθογώνιος πίνακας που έχει αυτά τα διανύσματα ως στήλες, με τη διάταξη $e_1, \ldots, e_n$ είναι ο ταυτοτικός $n \times n$ πίνακας $\mathbf{I}$. Οι ίδιες στήλες με διαφορετική διάταξη δίδουν τους πίνακες μετάθεσης, οι οποίοι είναι επίσης ορθογώνιοι πίνακες.

Πρόταση 1.10 *Εάν ο $m \times n$ πίνακας M έχει ορθοκανονικές στήλες τότε $M^T M = \mathbf{I}_n$. Ειδικότερα, εάν Q είναι ορθογώνιος πίνακας, τότε ο ανάστροφος πίνακας είναι και αντίστροφος,*

$$Q^T = Q^{-1}.$$

Απόδειξη. Υποθέτουμε οτι $q_1, \ldots, q_n \in \mathbb{R}^m$ είναι οι στήλες του M. Τότε $M^T M$ είναι ο

$n \times n$ πίνακας με στοιχείο στη θέση (i, j) το $q_i^T q_j$. Αλλά

$$q_i^T q_j = \begin{cases} 0 & \text{εάν} \quad i \neq j \\ 1 & \text{εάν} \quad i = j\,. \end{cases}$$

Συνεπώς $M^T M$ είναι ο ταυτοτικός $n \times n$ πίνακας και M^T είναι αριστερό αντίστροφο του M.

Εάν ο πίνακας είναι τετραγωνικός και έχει αριστερό αντίστροφο, γνωρίζουμε οτι είναι αντιστρέψιμος. Άρα Q^T είναι ο αντίστροφος πίνακας.

□

Παράδειγμα 1.7 Ο πίνακας περιστροφής $Q = \begin{bmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{bmatrix}$ είναι ορθογώνιος. Ο Q περιστρέφει κατά γωνία ϑ, ενώ ο ανάστροφος $Q^T = \begin{bmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{bmatrix}$ περιστρέφει κατά γωνία $-\vartheta$. Οι στήλες είναι ορθογώνιες, και αφού $\sin^2\vartheta + \cos^2\vartheta = 1$, έχουν μήκος 1.

Παράδειγμα 1.8 Όπως αναφέραμε προηγουμένως, κάθε πίνακας μετάθεσης είναι ορθογώνιος. Ειδικότερα ο πίνακας

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

που παριστάνει την ανάκλαση στον άξονα $x = y$. Γεωμετρικά, κάθε ορθογώνιος πίνακας είναι σύνθεση μίας περιστροφής και μίας ανάκλασης.

Οι ορθογώνιοι πίνακες έχουν ακόμα μία σημαντική ιδιότητα:

Πρόταση 1.11 *Ο πολλαπλασιασμός με ένα ορθογώνιο πίνακα Q αφήνει το μήκος αμετάβλητο: για κάθε $x \in \mathbb{R}^n$,*

$$||Qx|| = ||x||.$$

Γενικότερα, πολλαπλασιασμός με ορθογώνιο πίνακα αφήνει το εσωτερικό γινόμενο αμετάβλητο: για κάθε $x, y \in \mathbb{R}^n$,

$$(Qx)^T(Qy) = x^T y.$$

Απόδειξη. Γνωρίζουμε οτι $(Qx)^T = x^T Q^T$. Αλλά $Q^T Q = \mathbf{I}$ και έχουμε:

$$(Qx)^T(Qy) = x^T Q^T Q y = x^T \mathbf{I} y = x^T y.$$

□

Εάν ο πίνακας Q είναι ορθογώνιος, τότε $Q^T = Q^{-1}$, και συνεπώς $QQ^T = \mathbf{I}$. Αυτό σημαίνει οτι *οι γραμμές ενός ορθογώνιου πίνακα είναι επίσης ορθοκανονικά διανύσματα.*

1.10 Ορθοκανονικοποίηση Gram-Schmidt

Θα δείξουμε οτι εάν $v_1, \ldots, v_k$ είναι οποιαδήποτε βάση του υποχώρου $V \subseteq \mathbb{R}^m$, μπορούμε να κατασκευάσουμε από την $v_1, \ldots, v_k$ μία ορθοκανονική βάση $q_1, \ldots, q_k$, τέτοια ώστε για κάθε $j = 1, \ldots, k$, τα διανύσματα $q_1, \ldots, q_j$ παράγουν τον ίδιο υπόχωρο που παράγουν τα διανύσματα $v_1, \ldots, v_j$. Αυτή η διαδικασία ονομάζεται **ορθοκανονικοποίηση Gram-Schmidt**. Θα την περιγράψουμε στην περίπτωση τριών διανυσμάτων v_1, v_2, v_3. Υποθέτουμε οτι τα v_1, v_2, v_3 είναι γραμμικά ανεξάρτητα. Αρχικά θα τα αντικαταστήσουμε με τρία ορθογώνια διανύσματα w_1, w_2, w_3. Στη συνέχεια, διαιρούμε κάθε διάνυσμα w_i με το μήκος του και έχουμε τα ορθοκανονικά διανύσματα q_1, q_2, q_3.

> **Δραστηριότητα 1.9** Εάν v_1 και v_2 είναι διανύσματα του $\mathbb{R}^m$, βρείτε το $x \in \mathbb{R}$ για το οποίο $v_2 - xv_1$ είναι ορθογώνιο στο v_1.

Θέτουμε $w_1 = v_1$. Θέλουμε w_2 ορθογώνιο στο w_1 και τέτοιο ώστε τα w_1 και w_2 να παράγουν τον ίδιο υπόχωρο που παράγουν τα v_1 και v_2. Αφαιρούμε από το v_2 την προβολή του στην ευθεία που παράγεται από το w_1.

$$w_2 = v_2 - \frac{w_1^T v_2}{w_1^T w_1} w_1 \,.$$

Ελέγχουμε οτι w_1 και w_2 είναι ορθογώνια:

$$\begin{aligned} w_1^T w_2 &= w_1^T v_2 - \frac{w_1^T v_2}{w_1^T w_1} w_1^T w_1 \\ &= 0 \end{aligned}$$

Το v_3 δεν περιέχεται στο επίπεδο που παράγουν τα w_1, w_2, αφού υποθέσαμε οτι τα v_1, v_2 και v_3 είναι γραμμικά ανεξάρτητα. Για να βρούμε το w_3 θα αφαιρέσουμε την προβολή του v_3 στο επίπεδο που παράγουν τα v_1 και v_2. Έστω A ο πίνακας με στήλες τα διανύσματα w_1 και w_2. Τότε

$$A^T A = \begin{bmatrix} w_1^T w_1 & 0 \\ 0 & w_2^T w_2 \end{bmatrix}$$

και, χρησιμοποιώντας το συμβολισμό των μπλοκ, $A = [w_1\, w_2]$, η προβολή του v_3 στον υπόχωρο που παράγεται από τα w_1 και w_2 είναι

$$\begin{aligned} A(A^T A)^{-1} A^T v_3 &= [w_1\, w_2] \begin{bmatrix} w_1^T w_1 & 0 \\ & \\ 0 & w_2^T w_2 \end{bmatrix}^{-1} \begin{bmatrix} w_1^T \\ w_2^T \end{bmatrix} v_3 \\ &= \frac{w_1}{w_1^T w_1} w_1^T v_3 + \frac{w_2}{w_2^T w_2} w_2^T v_3 \\ &= \frac{w_1^T v_3}{w_1^T w_1} w_1 + \frac{w_2^T v_3}{w_2^T w_2} w_2 \end{aligned}$$

Παρατηρούμε οτι, επειδή τα διανύσματα w_1 και w_2 είναι ορθογώνια, η προβολή στο επίπεδο που παράγουν τα w_1 και w_2 είναι το άθροισμα των προβολών στις ευθείες των w_1 και w_2. Καταλήγουμε πως

$$w_3 = v_3 - \frac{w_1^T v_3}{||w_1||^2} w_1 - \frac{w_2^T v_3}{||w_2||^2} w_2.$$

Τα διανύσματα w_1, w_2, w_3 είναι τώρα ορθογώνια. Για να βρούμε την ορθοκανονική βάση του V αρκεί να διαιρέσουμε κάθε διάνυσμα με το μήκος του,

$$q_1 = \frac{w_1}{||w_1||}, \quad q_2 = \frac{w_2}{||w_2||}, \quad q_3 = \frac{w_3}{||w_3||}.$$

Παράδειγμα 1.9 Θεωρούμε τα διανύσματα

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Τότε

$$w_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

$$w_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \end{bmatrix},$$

$$w_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{3/2} \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ 1 \\ -\frac{2}{3} \\ -\frac{2}{3} \end{bmatrix}.$$

Άρα

$$q_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad q_2 = \frac{\sqrt{2}}{\sqrt{3}} \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ 0 \\ -\frac{1}{\sqrt{6}} \\ \frac{\sqrt{2}}{\sqrt{3}} \end{bmatrix}, \quad q_3 = \frac{3}{\sqrt{21}} \begin{bmatrix} \frac{2}{3} \\ 1 \\ -\frac{2}{3} \\ -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{21}} \\ \frac{3}{\sqrt{21}} \\ -\frac{2}{\sqrt{21}} \\ -\frac{2}{\sqrt{21}} \end{bmatrix}.$$

Όπως περιγράψαμε τη διαδικασία της απαλοιφής Gauss μέσω της παραγοντοποίησης $A = LU$, μπορούμε να περιγράψουμε και την ορθοκανονικοποίηση Gram-Schmidt μέσω μίας παραγοντοποίησης του πίνακα A ο οποίος έχει ως στήλες τα γραμμικά ανεξάρτητα διανύσματα $v_1, v_2, \ldots, v_k$:

$$A = QR,$$

όπου Q είναι ο πίνακας με ορθοκανονικές στήλες $q_1, q_2, \ldots, q_k$, και R είναι ο πίνακας που αντιστρέφει τη διαδικασία Gram-Schmidt. Αφού τα $q_1, \ldots, q_k$ είναι ορθογώνια και $v_j \in$

$\langle q_1, \ldots, q_j \rangle$,

$$\begin{aligned} v_1 &= c_{11} q_1 = (q_1^T v_1) q_1 \\ v_2 &= c_{12} q_1 + c_{22} q_2 = (q_1^T v_2) q_1 + (q_2^T v_2) q_2 \\ \cdots &= \cdots \\ v_k &= c_{1k} q_1 + \cdots + c_{kk} q_k = (q_1^T v_k) q_1 + \cdots + (q_k^T v_k) q_k \,. \end{aligned}$$

Βλέπουμε οτι εάν A είναι $m \times k$ πίνακας, R είναι ο άνω τριγωνικός $k \times k$ πίνακας με στοιχείο στη θέση i, j

$$R_{ij} = q_i^T v_j \quad \text{για } j \geq i \,.$$

Έχουμε αποδείξει την ακόλουθη Πρόταση.

Πρόταση 1.12 *Κάθε $m \times k$ πίνακας A με γραμμικά ανεξάρτητες στήλες μπορεί να παραγοντοποιηθεί στη μορφή*

$$A = QR \,,$$

όπου ο Q είναι $m \times k$ πίνακας με ορθοκανονικές στήλες, και ο R είναι $k \times k$ άνω τριγωνικός και αντιστρέψιμος. Εάν $m = k$, τότε Q είναι ορθογώνιος πίνακας.

Παράδειγμα 1.10 Στο Παράδειγμα 1.9, εφαρμόσαμε τη διαδικασία ορθοκανονικοποίησης στα γραμμικά ανεξάρτητα διανύσματα v_1, v_2, v_3 για να βρούμε τα ορθοκανονικά διανύσματα q_1, q_2, q_3. Αντιστρέφοντας αυτή τη διαδικασία έχουμε

$$v_1 = \sqrt{2} q_1 \,, \quad v_2 = \frac{1}{sqrt2} q_1 + \frac{3}{\sqrt{6}} q_2 \,, \quad v_3 = \sqrt{2} q_1 + \frac{2}{\sqrt{6}} q_2 + \frac{7}{\sqrt{21}} q_3 \,.$$

Συνεπώς ο πίνακας A που έχει ως στήλες τα διανύσματα v_1, v_2, v_3 παραγοντοποιείται ως $A = QR$, όπου Q είναι ο πίνακας που έχει ως στήλες τα διανύσματα q_1, q_2, q_3.

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{21}} \\ 0 & 0 & \frac{3}{\sqrt{21}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{-2}{\sqrt{21}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{-2}{\sqrt{21}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \sqrt{2} \\ 0 & \frac{3}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ 0 & 0 & \frac{7}{\sqrt{21}} \end{bmatrix} .$$

Η παραγοντοποίηση $A = QR$ απλοποιεί το πρόβλημα εύρεσης βέλτιστης λύσης ελαχίστων τετραγώνων. Η εξίσωση

$$A^T A \hat{x} = A^T b$$

γίνεται

$$R^T Q^T Q R \hat{x} = R^T Q^T b.$$

Αφού $Q^T Q = I$ και R^T είναι αντιστρέψιμος, έχουμε

$$R \hat{x} = Q^T b$$

και αφού R είναι άνω τριγωνικός, το διάνυσμα $\hat{x}$ υπολογίζεται με ανάδρομη αντικατάσταση.

1.11 Ασκήσεις

Άσκηση 1.45 Εάν u είναι μοναδιαίο διάνυσμα, δείξτε οτι $Q = I - 2uu^T$ είναι συμμετρικός ορθογώνιος πίνακας. (Είναι μία ανάκλαση, και ονομάζεται μετασχηματισμός Householder). Υπολογίστε τον Q όταν $u^T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$.

Άσκηση 1.46 Θεωρούμε οτι οι στήλες του $n \times n$ πίνακα A είναι οι $n \times 1$ πίνακες $C_1, C_2, \ldots, C_n$, και οι γραμμές του $n \times n$ πίνακα B είναι οι $1 \times n$ πίνακες $R_1, R_2, \ldots, R_n$. Για κάθε $i = 1, \ldots, n$, το γινόμενο $C_i R_i$ είναι ένας $n \times n$ πίνακας. Εκφράστε το γινόμενο AB ως άθροισμα τέτοιων πινάκων.

Άσκηση 1.47 Δείξτε οτι εάν A είναι $m \times k$ πίνακας με ορθογώνιες στήλες $w_1, \ldots, w_k$, τότε η προβολή P στο διανυσματικό υπόχωρο $\langle w_1, \ldots, w_k \rangle$ είναι το άθροισμα των προβολών P_i στους μονοδιάστατους υπόχωρους $\langle w_i \rangle$, για $i = 1, \ldots, k$.

Άσκηση 1.48 Προβάλετε το διάνυσμα $b = (1, 2)$ σε δύο μή ορθογώνια διανύσματα, $a_1 = (1, 0)$ και $a_2 = (1, 1)$. Επαληθεύστε οτι το άθροισμα των δύο προβολών δεν είναι ίσο προς το b.

Άσκηση 1.49 Δείξτε οτι ένας άνω τριγωνικός ορθογώνιος πίνακας πρέπει να είναι διαγώνιος.

Άσκηση 1.50 Από τα μη ορθογώνια διανύσματα v_1, v_2, v_3, βρείτε ορθογώνια διανύσματα q_1, q_2, q_3.

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Άσκηση 1.51 Ποιά είναι τα δυνατά διανύσματα v_1 και v_2, που δίνουν μετά από ορθοκανονικοποίηση Gram-Schmidt τα διανύσματα q_1 και q_2.

Άσκηση 1.52 Ποιό πολλαπλάσιο του $a_1 = (1, 1)$ πρέπει να αφαιρεθεί από το $a_2 = (4, 0)$, ώστε το αποτέλεσμα να είναι ορθογώνιο προς το a_1. Παραγοντοποιήστε τον πίνακα $\begin{bmatrix} 1 & 4 \\ 1 & 0 \end{bmatrix}$ σε γινόμενο QR όπου Q είναι ορθογώνιος.

Άσκηση 1.53 Εφαρμόστε τη διαδικασία Gram-Schmidt στα διανύσματα

$$v_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

και εκφράστε το αποτέλεσμα στη μορφή $A = QR$.

Άσκηση 1.54 Εάν $A = QR$, όπου οι στήλες του Q είναι ορθογώνια διανύσματα, βρείτε έναν απλό τύπο για τον πίνακα προβολής στο χώρο στηλών του A.

Άσκηση 1.55 Βρείτε τρία ορθοκανονικά διανύσματα q_1, q_2, $q_3 \in \mathbb{R}^3$, τέτοια ώστε τα q_1, q_2 να παράγουν το χώρο στηλών του πίνακα

$$A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix}.$$

Ποιός θεμελιώδης υπόχωρος του A περιέχει το διάνυσμα q_3; Βρείτε τη βέλτιστη λύση ελαχίστων τετραγώνων της εξίσωσης $Ax = b$, όταν $b^T = [1\,2\,7]$.

Άσκηση 1.56 Με τον πίνακα A της Άσκησης 1.55, και το διάνυσμα $b = [1\,1\,1]^T$, χρησιμοποιήστε την παραγοντοποίηση $A = QR$ για να λύσετε το πρόβλημα ελαχίστων τετραγώνων $Ax = b$.

Άσκηση 1.57 Εφαρμόστε τη διαδικασία Gram-Schmidt στα διανύσματα $(1,\,-1,\,0)$, $(0,\,1,\,-1)$ και $(1,\,0,\,-1)$ για να βρείτε ορθοκανονική βάση του επιπέδου $x_1+x_2+x_3 = 0$. Πόσα μή μηδενικά διανύσματα προκύπτουν από τη διαδικασία Gram-Schmidt;

Άσκηση 1.58 Βρείτε ορθογώνια διανύσματα w_1, w_2, w_3 από τα διανύσματα

$$v_1 = (1,\,-1,\,0,\,0),\; v_2 = (0,\,1,\,-1,\,0),\; v_3 = (0,\,0,\,1,\,-1)\,.$$

Τα v_1, v_2, v_3 αποτελούν βάση του υποχώρου που είναι ορθογώνιος στο $(1,\,1,\,1,\,1)$.