Κεφάλαιο 4

Μήκη και ορθές γωνίες

Μήκος διανύσματος

Στο επίπεδο, $\mathbb{R}^2$, βρίσκουμε το μήκος ενός διανύσματος $x = (x_1,\, x_2)$ χρησιμοποιώντας το Πυθαγόρειο θεώρημα:

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

Στο χώρο $\mathbb{R}^3$, εφαρμόζουμε το Πυθαγόρειο θεώρημα 2 φορές: εάν $x = (x_1,\, x_2,\, x_3)$ και $u = (x_1,\, x_2)$

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

Χρησιμοποιώντας το συμβολισμό του ανάστροφου, αυτό γράφεται

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

Εφαρμόζοντας το Πυθαγόρειο θεώρημα $n-1$ φορές, βρίσκουμε το μήκος ενός διανύσματος στο $\mathbb{R}^n$:

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

Ορθογώνια διανύσματα

Εκτός από τα μήκη, θέλουμε να μετράμε 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. Η γωνία $\angle(x,\, y)$ είναι ορθή εάν και μόνον εάν

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

δηλαδή εάν και μόνον εάν

$$x_1 y_1 + \cdots + x_n y_n = 0$$

ή

$$x^T y = 0$$

Πρόταση 4.1 *Δύο διανύσματα x, y του $\mathbb{R}^n$ είναι **ορθογώνια** εάν και μόνον εάν το εσωτερικό τους γινόμενο $x^T y$ είναι 0.*

□

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

Πρόταση 4.2 *Εάν τα διανύσματα $v_1, \ldots, v_n$ είναι μη μηδενικά και ορθογώνια μεταξύ τους, τότε είναι γραμμικώς ανεξάρτητα.*

Απόδειξη. Έστω ένας γραμμικός συνδυασμός $c_1 v_1 + \cdots + c_n v_n = 0$. Παίρνουμε το εσωτερικό γινόμενο με το v_1:

$$v_1^T(c_1 v_1 + \cdots + c_n v_n) = v_1^T 0 = 0$$

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

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

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

Παρόμοια, $c_i = 0$ για κάθε i.

□

Είναι προφανές ότι δεν ισχύει το αντίστροφο: δυο γραμμικώς ανεξάρτητα διανύσματα δεν είναι υποχρεωτικά ορθογώνια.

Ορθογώνιοι υπόχωροι

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

Ανάλογα, ορίζουμε την ορθογωνιότητα δύο υποχώρων.

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

Προσέξτε, όμως, ότι στο $\mathbb{R}^3$, δύο κάθετα επίπεδα W_1 και W_2 (δηλαδή, τα κάθετα σε αυτά διανύσματα είναι κάθετα μεταξύ τους) **δεν** ικανοποιούν αυτή τη συνθήκη, δηλαδή, ως υπόχωροι του $\mathbb{R}^n$, οι W_1, W_2 **δεν** είναι ορθογώνιοι. 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 4.2 αυτά θα ήταν γραμμικώς ανεξάρτητα. Αλλά στον $\mathbb{R}^3$ δεν υπάρχουν 4 γραμμικώς ανεξάρτητα διανύσματα.

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

Πρόταση 4.3 *Αν $V \perp W$ τότε $V \cap W = \{\mathbf{0}\}$.*

Απόδειξη. Αν $x \in V \cap W$, τότε $x \perp x$, άρα $||x|| = 0$, οπότε $x = \mathbf{0}$.

□

Πρόταση 4.4 *Αν $V \perp W$, $v_1, \ldots, v_k \in V$ είναι γραμμικώς ανεξάρτητα και $w_1, \ldots, w_l \in W$ είναι, επίσης, γραμμικώς ανεξάρτητα, τότε τα $v_1, \ldots, v_k, w_1, \ldots, w_l$ είναι γραμμικώς ανεξάρτητα.*

Απόδειξη. Έστω $x_1v_1+\cdots+x_kv_k+y_1w_1+\cdots+y_lw_l=\mathbf{0}$, όπου $x_1,\ldots,x_k,y_1,\ldots,y_l\in\mathbb{R}$. Θα δείξομε ότι όλοι οι συντελεστές x_i και όλοι οι συντελεστές y_j είναι μηδενικοί. Για τον σκοπό αυτό θέτομε $v=x_1v_1+\cdots+x_kv_k$ και $w=y_1w_1+\cdots+y_lw_l$, οπότε $v+w=\mathbf{0}$. Καθώς $v\in V$ και $w\in W$, η τελευταία ισότητα συνεπάγεται ότι $v\in V\cap W$, άρα, από την Πρόταση 4.3, $v=\mathbf{0}$ και $w=\mathbf{0}$. Συνεπώς, $x_1v_1+\cdots+x_kv_k=\mathbf{0}$ και $y_1w_1+\cdots+y_lw_l=\mathbf{0}$. Λόγω της γραμμικής ανεξαρτησίας των $v_1,\ldots,v_k$ έπεται ότι $x_1=\ldots=x_k=0$ και, ανάλογα, $w_1=\ldots=w_l=0$.

□

Πόρισμα 4.5 *Αν οι V,W είναι ορθογώνιοι υπόχωροι του $\mathbb{R}^n$ και $\dim V+\dim W=n$, τότε η ένωση μιας οποιασδήποτε βάσης του V με οποιαδήποτε βάση του W δίνει μια βάση του $\mathbb{R}^n$*

□

Πόρισμα 4.6 *Αν οι V,W είναι ορθογώνιοι υπόχωροι του $\mathbb{R}^n$ και $\dim V+\dim W=n$, τότε κάθε $x\in\mathbb{R}^n$ γράφεται με ένα και μοναδικό τρόπο ως $x=v+w$, όπου $v\in V$ και $w\in W$.*

Απόδειξη. Θεωρούμε βάσεις $v_1,\ldots,v_k$ του V και $w_1,\ldots,w_l$ του W. Από το πόρισμα 4.5 έπεται ότι υπάρχουν $x_1,\ldots,x_k,y_1,\ldots,y_l\in\mathbb{R}$ τέτοιοι ώστε $x_1v_1+\cdots+x_kv_k+y_1w_1+\cdots+y_lw_l=x$. Θέτομε $v=x_1v_1+\cdots+x_kv_k$ και $w=y_1w_1+\cdots+y_lw_l$, οπότε $v+w=x$ με το $v\in V$ και το $w\in W$. Αν τώρα ισχύει και $x=v'+w'$, όπου $v'\in V$ και $w'\in W$, τότε $v+w=v'+w'$, απ' όπου $v-v'=w'-w$. Αυτό σημαίνει ότι το τελευταίο διάνυσμα ανήκει στην τομή των V και W άρα, από την πρόταση 4.3, είναι το μηδενικό διάνυσμα. Άρα $v'=v$ και $w'=w$.

□

Άσκηση 4.1 Έστω ότι V,W είναι μη μηδενικοί υπόχωροι του $\mathbb{R}^n$, $v_1,\ldots,v_k$ μια οποιαδήποτε βάση του V και $w_1,\ldots,w_l$ μια οποιαδήποτε βάση του W. Αποδείξτε ότι $V\perp W$ αν και μόνο αν $v_i\perp w_j$ για κάθε $i=1,\ldots,k$ και $j=1,\ldots,l$.

Άσκηση 4.2 Έστω ότι V,W είναι μη μηδενικοί υπόχωροι του $\mathbb{R}^n$, και $V=\langle v_1,\ldots,v_k\rangle$. Αποδείξτε ότι $V\perp W$ αν και μόνο αν $w\perp v_i$ για κάθε $i=1,\ldots,k$.

Ορισμός. Έστω V υπόχωρος του $\mathbb{R}^n$. Το σύνολο

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

λέγεται ορθογώνιο συμπλήρωμα του V.

Άσκηση 4.3 Αποδείξτε ότι το ορθογώνιο συμπλήρωμα ενός υποχώρου V του $\mathbb{R}^n$ είναι *υπόχωρος* του $\mathbb{R}^n$.

Θεώρημα 4.7 *Το ορθογώνιο συμπλήρωμα ενός υποχώρου V του $\mathbb{R}^n$ είναι υπόχωρος του $\mathbb{R}^n$ διαστάσεως $n-\dim V$. Ακόμη, $(V^\perp)^\perp=V$.*

Απόδειξη. Αν $V=\mathbf{0}$ τότε, προφανώς, $V^\perp=\mathbb{R}^n$ και ο ισχυρισμός αποδείχθηκε.
Έστω τώρα $\dim V=m\geq 1$, $v_1,\ldots,v_m$ μια βάση του V και A ο πίνακας με γραμμές $v_1^T,\ldots,v_m^T$. Ο A είναι πίνακας $m\times n$ τάξεως m, αφού οι m γραμμές του είναι γραμμικώς ανεξάρτητες. Σύμφωνα με την άσκηση 4.1, ένα $w\in\mathbb{R}^n$ ανήκει στο $V^\perp$ αν και μόνο αν $w\perp v_i$ για κάθε $i=1,\ldots,m$. Αυτή η συνθήκη, όμως, ισοδυναμεί με την $Aw=\mathbf{0}$. Άρα

$w \in V^{\perp} \Leftrightarrow w \in \mathcal{N}(A)$, δηλαδή, $V^{\perp} = \mathcal{N}(A)$. Άρα $\dim V^{\perp} = \dim \mathcal{N}(A) = n - \text{τάξη}(A) = n - m$.

Θα απδείξομε τώρα ότι $(V^{\perp})^{\perp} = V$. Πράγματι, εξ ορισμού του $V^{\perp}$, κάθε $v \in V$ είναι ορθογώνιο πρός κάθε $w \in V^{\perp}$, άρα $V \subseteq (V^{\perp})^{\perp}$. Επιπλέον, οι διαστάσεις των δύο υποχώρων είναι ίσες, αφού $\dim(V^{\perp})^{\perp} = n - \dim V^{\perp} = n - (n-m) = m$. Συνεπώς, οι δύο υπόχωροι ταυτίζονται.

□

Βάσει του θεωρήματος αυτού και του πορίσματος 4.6, αν V είναι μη μηδενικός υπόχωρος του $\mathbb{R}^n$, τότε κάθε $x \in \mathbb{R}^n$ 'αναλύεται' με ένα και μοναδικό τρόπο σε άθροισμα $x = v + w$, όπου $v \in v$ και $w \in V^{\perp}$. Το v λέγεται *προβολή του v στον υπόχωρο V*.

Θεώρημα 4.8 *Σε κάθε πίνακα A ισχύουν τα εξής:*

(α') $\mathcal{R}(A^T)^{\perp} = \mathcal{N}(A)$ και $\mathcal{N}(A)^{\perp} = \mathcal{R}(A^T)$, δηλαδή, για τον χώρο γραμμών και τον μηδενόχωρο ισχύει ότι ο ένας είναι ορθογώνιο συμπλήρωμα του άλλου.

(β') $\mathcal{R}(A)^{\perp} = \mathcal{N}(A^T)$ και $\mathcal{N}(A^T)^{\perp} = \mathcal{R}(A)$, δηλαδή, για τον χώρο στηλών και τον αριστερό μηδενόχωρο ισχύει ότι ο ένας είναι ορθογώνιο συμπλήρωμα του άλλου.

Απόδειξη. (α') Πρώτα αποδεικνύομε ότι $\mathcal{N}(A) = \mathcal{R}(A^T)^{\perp}$. Ο $\mathcal{R}(A^T)$, εξ ορισμού, παράγεται από τις γραμμές του A, άρα, από την άσκηση 4.2, ένα διάνυσμα (στήλη) του $\mathbb{R}^n$ ανήκει στον $\mathcal{R}(A^T)^{\perp}$ αν, και μόνο αν, το x είναι ορθογώνιο προς κάθε γραμμή του A. Προφανώς, αυτό ισοδυναμεί με τη σχέση $Ax = \mathbf{0}$, δηλαδή, με τη σχέση $x \in \mathcal{N}(A)$. Άρα, $\mathcal{N}(A) = \mathcal{R}(A^T)^{\perp}$.

Ο δεύτερος ισχυρισμός του (α') είναι άμεση συνέπεια του δευτέρου ισχυρισμού του θεωρήματος 4.7.

(β') Προκύπτει αμέσως από το (α') αν βάλομε στη θέση του A τον A^T.

□

Έστω A ένας $m \times n$ πίνακας. Σύμφωνα με το θεώρημα 4.8, ο χώρος γραμμών $\mathcal{R}(A^T)$ και ο μηδενόχωρος $\mathcal{N}(A)$ είναι υπόχωροι του $\mathbb{R}^n$, που ο ένας είναι ορθογώνιο συμπλήρωμα του άλλου. Άρα, από την παρατήρηση μετά το θεώρημα 4.7, κάθε $x \in \mathbb{R}^n$ γράφεται με ένα και μοναδικό τρόπο ως άθροισμα $x = x_\gamma + x_\mu$, όπου το x_γ ανήκει στον χώρο γραμμών του πίνακα A και το x_μ στον μηδενόχωρο του A.

Επίσης, ο χώρος στηλών $\mathcal{R}(A)$ και ο αριστερός μηδενόχωρος $\mathcal{N}(A^T)$ είναι υπόχωροι του $\mathbb{R}^m$, που ο ένας είναι ορθογώνιο συμπλήρωμα του άλλου. Το επόμενο θεώρημα μας περιγράφει τη 'δράση' της απεικόνισης $L_A : \mathbb{R}^n \to \mathbb{R}^m$.

Θεώρημα 4.9 *(Θεμελιώδες Θεώρημα της Γραμμικής Άλγεβρας). Έστω A ένας $m \times n$ πίνακας και $L_A : \mathbb{R}^n \to \mathbb{R}^m$ η αντίστοιχη γραμμική απεικόνιση.*

(α') Αν περιορίσομε την L_A στον χώρο γραμμών, τότε παίρνομε αμφιμονοσήμαντη απεικόνιση επί του χώρου στηλών. Δηλαδή, η γραμμική απεικόνιση $L_A : \mathcal{R}(A^T) \to \mathcal{R}(A)$ είναι αμφιμονοσήμαντη (1-1 και επί).

(β') Αν περιορίσομε την L_A στον μηδενόχωρο, τότε παίρνομε τη μηδενική απεικόνιση, δηλαδή, έχομε την τετριμμένη γραμμική απεικόνιση $L_A : \mathcal{N}(A) \to \{\mathbf{0}\} \subseteq \mathbb{R}^m$.

(γ') Για κάθε $x \in \mathbb{R}^n$, αν θεωρήσομε την ανάλυσή του $x = x_\gamma + x_\mu$, όπου το x_γ ανήκει στον χώρο γραμμών του πίνακα A και το x_μ στον μηδενόχωρο του A, τότε $L_A(x) = L_A(x_\gamma) + L_A(x_\mu) = L_A(x_\gamma) + \mathbf{0} = L_A(x_\gamma)$.

(δ') Όλα τα μη μηδενικά διανύσματα του αριστερού μηδενόχωρου του A είναι 'απρόσιτα' από την L_A, δηλαδή, αν κάποιο μη μηδενικό $y \in \mathbb{R}^m$ ανήκει στον αριστερό μηδενόχωρο του A, τότε δεν υπάρχει $x \in \mathbb{R}^n$ που ν' απεικονίζεται μέσω της L_A στο y. Ισοδύναμη διατύπωση: $L_A^{-1}(\mathcal{N}(A)) = \{\mathbf{0}\}$.

Απόδειξη. (α') Έστω r η τάξη του A, άρα r από τις γραμμές του, ας πούμε οι $\gamma_1, \ldots, \gamma_r$ αποτελούν βάση για τον χώρο γραμμών $\mathcal{R}(A^T)$. Αν θέσομε $L_A(\gamma_i) = \sigma_i$, $(i = 1, \ldots, r)$, τότε τα διανύσματα (στήλες) $\sigma_1, \ldots, \sigma_r$ ανήκουν στον χώρο στηλών $\mathcal{R}(A)$, διότι κάθε στήλη της μορφής Ax, για οποιοδήποτε $x \in \mathbb{R}^n$, είναι γραμμικός συνδυασμός των στηλών του A. Επιπλέον, τα $\sigma_1, \ldots, \sigma_r$ είναι γραμμικώς ανεξάρτητα. Πράγματι, έστω ότι για κάποιους $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ ισχύει $\lambda_1\sigma_1 + \cdots + \lambda_r\sigma_r = \mathbf{0}$. Τότε,

$$\mathbf{0} = \lambda_1 L_A(\gamma_1) + \cdots + \lambda_r L_A(\gamma_r) = L_A(\lambda_1\gamma_1 + \cdots + \lambda_r\gamma_r) = A(\lambda_1\gamma_1 + \cdots + \lambda_r\gamma_r).$$

Επομένως, $\lambda_1\gamma_1 + \cdots + \lambda_r\gamma_r \in \mathcal{N}(A)$. Όμως το διάνυσμα στο αριστερό μέλος ανήκει και στον υπόχωρο $\mathcal{R}(A^T)$, ο οποίος είναι ορθογώνιος προς τον $\mathcal{N}(A)$ (θεώρημα 4.8), συνεπώς, λόγω της πρότασης 4.3, το διάνυσμα αυτό είναι το μηδενικό. Άρα $\lambda_1\gamma_1 + \cdots + \lambda_r\gamma_r = \mathbf{0}$ και λόγω της ανεξαρτησίας των $\gamma_1, \ldots, \gamma_r$, οι συντελεστές $\lambda_1, \ldots, \lambda_r$ πρέπει να είναι όλοι μηδέν. Το συμπέρασμά μας, λοιπόν, είναι ότι τα $\sigma_1, \ldots, \sigma_r$ είναι r γραμμικώς ανεξάρτητα διανύσματα του r-διάστατου υποχώρου $\mathcal{R}(A)$, άρα αποτελούν βάση του $\mathcal{R}(A)$.
Τώρα είναι πολύ εύκολο να δείξομε ότι η L_A είναι 'επί'. Κάθε $y \in \mathcal{R}(A)$ γράφεται $y = \mu_1\sigma_1 + \cdots + \mu_r\sigma_r$, οπότε, αν θέσομε $x = \mu_1\gamma_1 + \cdots + \mu_r\gamma_r$, τότε $x \in \mathcal{R}(A^T)$ και $L_A(x) = \mu_1 L_A(\gamma_1) + \cdots \mu_r L_A(\gamma_r) = \mu_1\sigma_1 + \cdots + \mu_r\sigma_r = y$.
Επίσης, η L_A περιορισμένη στον $\mathcal{R}(A^T)$ είναι 1-1. Πράγματι, αν $x_1, x_2 \in \mathcal{R}(A^T)$ και $L_A(x_1) = L_A(x_2)$, τότε $A \cdot (x_1 - x_2) = \mathbf{0}$, άρα $x_1 - x_2 \in \mathcal{N}(A)$. Καθώς, όμως, $x_1 - x_2 \in \mathcal{R}(A^T)$, το θεώρημα 4.8-α' σε συνδυασμό με την πρόταση 4.3 μας οδηγούν στο συμπέρασμα $x_1 - x_2 = \mathbf{0}$.

(β') Το $x \in \mathbb{R}^n$ ανήκει στον μηδενόχωρο $\mathcal{N}(A)$ εξ ορισμού αν $Ax = \mathbf{0}$, δηλαδή, αν και μόνο αν $L_A(x) = \mathbf{0}$.

(γ') Άμεση συνέπεια των (α') και (β').

(δ') Έστω ότι κάποιο $y \in \mathcal{N}(A^T)$ είναι εικόνα μέσω της L_A κάποιου $x \in \mathbb{R}^n$. Από το (γ') έχομε ότι $L_A(x) = L_A(x_\gamma)$ για κάποιο $x_\gamma \in \mathcal{R}(A^T)$ και από το (α'), $L_A(x_\gamma) \in \mathcal{R}(A)$. Άρα $y \in \mathcal{R}(A) \cap \mathcal{N}(A^T)$. Η τελευταία τομή, όμως, είναι το μηδενικό διάνυσμα, λόγω του θεωρήματος 4.8-β' και της πρότασης 4.3.

□