1η ΔΙΑΛΕΞΗ
ΕΣΩΤΕΡΙΚΟ ΓΙΝΟΜΕΝΟ

- **Βασικοί ορισμοί και συμβολισμοί**
- **Ανισότητα Cauchy - Schwarz**
- **Norm διανυσμάτων**
- **Γωνία διανυσμάτων**

Διανυσματικοί χώροι με εσωτερικό γινόμενο

Έστω $\mathcal{V} = (V, +, \cdot_{\mathbb{R}})$ ένας πραγματικός διανυσματικός χώρος και $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$.
Η απεικόνιση $\langle \rangle$ ονομάζεται **εσωτερικό γινομένο** επί του $\mathcal{V}$ αν και μόνο αν για κάθε $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ και για κάθε $a \in \mathbb{R}$ ισχύει ότι

1. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$,
2. $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{0} \Leftrightarrow \mathbf{x} = \mathbf{0}$,
3. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$,
4. $\langle a\mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = a \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$.

Στην περίπτωση αυτή έχουμε ένα **πραγματικό διανυσματικό χώρο με εσωτερικό γινόμενο** και γράφουμε $(V, +, \cdot_{\mathbb{R}}, \langle \rangle)$.
Αν επιπλέον, ο $\mathcal{V}$ είναι πεπερασμένης διάστασης, ο διανυσματικός χώρος λέγεται **Ευκλείδειος χώρος**.
Μερικές φορές αντί για $\langle \mathbf{x}, \mathbf{y} \rangle$ χρησιμοποιείται και ο συμβολισμός $\mathbf{x} \cdot \mathbf{y}$.

Διανυσματικοί χώροι με εσωτερικό γινόμενο

Παράδειγμα 1

Για τον διανυσματικό χώρο $(\mathbb{R}^n, +, \cdot_{\mathbb{R}})$ η απεικόνιση $\langle\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ όπου για κάθε $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ με $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ και $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ ορίζουμε

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

είναι ένα εσωτερικό γινόμενο που λέγεται **σύνηθες εσωτερικό γινόμενο επί του** $\mathbb{R}$.

Διανυσματικοί χώροι με εσωτερικό γινόμενο

Πράγματι, έστω $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$, $\mathbf{z} = (z_1, \ldots, z_n)$ και $a \in \mathbb{R}$. Τότε,

1. $$\langle \mathbf{x}, \mathbf{x} \rangle = x_1 x_1 + \cdots x_n x_n = x_1^2 + \cdots + x_n^2 \geq 0.$$

2. $$\begin{aligned} \langle \mathbf{x}, \mathbf{x} \rangle = 0 &\Leftrightarrow x_1^2 + \cdots + x_n^2 = 0 \\ &\Leftrightarrow x_1 = \cdots = x_n = 0 \\ &\Leftrightarrow \mathbf{x} = (0, 0, \ldots, 0). \end{aligned}$$

Διανυσματικοί χώροι με εσωτερικό γινόμενο

3.

$$\begin{aligned}\langle \mathbf{x}, \mathbf{y} \rangle &= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \\ &= y_1 x_1 + y_2 x_2 + \cdots + y_n x_n \\ &= \langle \mathbf{x}, \mathbf{y} \rangle .\end{aligned}$$

4.

$$\begin{aligned}\langle a\mathbf{x} + \mathbf{y}, \mathbf{z} \rangle &= \langle a(x_1, \ldots, x_n) + (y_1, \ldots, y_n), (z_1, \ldots, z_n) \rangle \\ &= \langle (ax_1, \ldots, ax_n) + (y_1, \ldots, y_n), (z_1, \ldots, z_n) \rangle \\ &= \langle (ax_1 + y_1, \ldots, ax_n + y_n), (z_1, \ldots, z_n) \rangle \\ &= (ax_1 + y_1) z_1 + \cdots + (ax_n + y_n) z_n \\ &= ax_1 z_1 + \cdots + ax_n z_n + y_1 z_1 + \cdots + y_n z_n \\ &= a(x_1 z_1 + \cdots + x_n z_n) + (y_1 z_1 + \cdots + y_n z_n) \\ &= a \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle .\end{aligned}$$

Παρατηρήσεις:

1. Ο παραπάνω διανυσματικός χώρος είναι Ευκλείδειος, αφού $\dim \mathbb{R}^n = n$ και η απεικόνιση $\langle\rangle$ ορίζει εσωτερικό γινόμενο επί του $\mathbb{R}^n$.
2. Έστω $A = [a_{ij}] \in \mathcal{M}_{k\times n}$, $B = [b_{ij}] \in \mathcal{M}_{n\times m}$ και $C = [c_{ij}] = AB$.
   Όπως είναι γνωστό, $c_{ij} = \sum\limits_{\lambda=1}^{n} a_{i\lambda} \cdot b_{\lambda j}$.
   Οι γραμμές της A: $R_1, R_2, \ldots, R_k$ και οι στήλες της B: $C_1, C_2, \ldots, C_m$ μπορούν να θεωρηθούν στοιχεία του $\mathbb{R}^n$,
   οπότε $c_{ij} = \langle R_i, C_j\rangle$, όπου $\langle\ \rangle$ το σύνηθες εσωτερικό γινόμενο του $\mathbb{R}^n$.

Διανυσματικοί χώροι με εσωτερικό γινόμενο

Στον διανυσματικό χώρο $(\mathbb{R}^3, +, \cdot)$ με το συνήθες εσωτερικό γινόμενο $\langle\rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ έχουμε

$$\langle (x_1, x_2, x_3), (y_1, y_2, y_3) \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

Έτσι,

- $\langle (2, 1, -2), (-3, 1, 0) \rangle = -6 + 1 + 0 = -5.$
- $\langle (1, 2, 3), (4, 5, 6) \rangle = 4 + 10 + 18 = 32.$
- Έστω $\mathbf{x} = (-1, 2, -1)$, $\mathbf{y} = (2, 2, 0)$, $\mathbf{z} = (-5, -1, 2)$. Τότε

$$\begin{aligned} \langle 3\mathbf{x} + \mathbf{y}, \mathbf{z} \rangle &= \langle 3(-1, 2, -1) + (2, 2, 0), (-5, -1, 2) \rangle \\ &= \langle (-3, 6, -3) + (2, 2, 0), (-5, -1, 2) \rangle \\ &= \langle (-1, 8, -3), (-5, -1, 2) \rangle = 5 - 8 - 6 = -9 \end{aligned}$$

και

$$\begin{aligned} 3\langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle &= 3\langle (-1, 2, -1), (-5, -1, 2) \rangle + \langle (2, 2, 0), (-5, -1, 2) \rangle \\ &= 3(5 - 2 - 2) + (-10 - 2 + 0) = 3 \cdot 1 - 12 = -9, \end{aligned}$$

(επαληθεύοντας την ιδιότητα 4 του ορισμού).

Παράδειγμα 2

Έστω $(V, +, \cdot)$ ο πραγματικός διανυσματικός χώρος των συνεχών συναρτήσεων που ορίζονται στο $[a, b]$, $a < b$. Η απεικόνιση $<>: V^2 \to \mathbb{R}$ για την οποία

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx, \text{ για κάθε } f, g \in V$$

είναι ένα εσωτερικό γινόμενο επί του V.

Διανυσματικοί χώροι με εσωτερικό γινόμενο

Πράγματι, έστω $f, g, h \in V$ και $\lambda \in \mathbb{R}$. Τότε

1. $\langle f, f \rangle = \int_a^b f^2(x)dx \geq \int_a^b 0dx = 0.$
2. $\langle f, f \rangle = 0 \Leftrightarrow \int_a^b f^2(x)dx = 0 \overset{f \text{ συνεχής}}{\Leftrightarrow} f = \mathbf{0}.$
3. $\langle f, g \rangle = \int_a^b f(x)g(x)dx = \int_a^b g(x)f(x)dx = \langle g, f \rangle.$
4. $$\begin{aligned}\langle \lambda f + g, h \rangle &= \int_a^b (\lambda f(x) + g(x))h(x)dx \\ &= \lambda \int_a^b f(x)h(x) + \int_a^b g(x)h(x)dx \\ &= \lambda < f, h > + < g, h >\end{aligned}$$

Ο διανυσματικός χώρος V δεν είναι Ευκλείδειος.

Διανυσματικοί χώροι με εσωτερικό γινόμενο

Παράδειγμα 3

Έστω $(\mathcal{M}_n, +, \cdot)$ ο πραγματικός διανυσματικός χώρος των $n \times n$ μητρών. Η απεικόνιση $\langle\rangle : \mathcal{M}_n \times \mathcal{M}_n \to \mathbb{R}$ για την οποία

$$\langle A, B\rangle = \mathrm{tr}(\mathrm{AB}^{\mathrm{t}})$$

είναι ένα εσωτερικό γινόμενο επί του $\mathcal{M}_n$.

Διανυσματικοί χώροι με εσωτερικό γινόμενο

Πράγματι, έστω $A, B, C \in \mathcal{M}_n$. Αν $R_1, R_2, \ldots, R_n$ είναι οι γραμμές της μήτρας A και $R'_1, R'_2, \ldots, R'_n$ οι γραμμές της μήτρας B τότε, όπως αναφέρθηκε σε προηγούμενη παρατήρηση, αν θεωρήσουμε ότι $R_i, R'_i \in \mathbb{R}^n$ για κάθε $i \in [n]$ τότε

1. $$\langle A, A \rangle = \operatorname{tr}(AA^t) = \langle R_1, R_1 \rangle + \langle R_2, R_2 \rangle + \cdots + \langle R_n, R_n \rangle \geq 0,$$

   όπου $\langle R_i, R_i \rangle$ είναι το σύνηθες εσωτερικό γινόμενο του $(\mathbb{R}^n, +, \cdot)$, για το οποίο γνωρίζουμε ότι για κάθε $\mathbf{x} \in \mathbb{R}^n$ είναι $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$.

2. $$\begin{aligned} \langle A, A \rangle = 0 &\Leftrightarrow \langle R_1, R_1 \rangle + \langle R_2, R_2 \rangle + \cdots + \langle R_n, R_n \rangle = 0 \\ &\Leftrightarrow \langle R_1, R_1 \rangle = \langle R_2, R_2 \rangle = \cdots = \langle R_n, R_n \rangle = 0 \\ &\Leftrightarrow R_1 = R_2 = \cdots = R_n = (0, 0, \ldots, 0) \in \mathbb{R}^n \\ &\Leftrightarrow A = O_{n \times n}. \end{aligned}$$

Διανυσματικοί χώροι με εσωτερικό γινόμενο

3.

$$\begin{aligned}\langle A, B\rangle &= \operatorname{tr}(AB^t)\\ &= \left\langle R_1, R_1'\right\rangle + \left\langle R_2, R_2'\right\rangle + \cdots + \left\langle R_n, R_n'\right\rangle\\ &= \left\langle R_1', R_1\right\rangle + \left\langle R_2', R_2\right\rangle + \cdots + \left\langle R_n', R_n\right\rangle\\ &= \operatorname{tr}(BA^t) = \langle B, A\rangle .\end{aligned}$$

4. Έστω $\lambda \in \mathbb{R}$. Τότε,

$$\begin{aligned}\langle \lambda A + B, C\rangle &= \operatorname{tr}((\lambda A + B)C^t)\\ &= \operatorname{tr}(\lambda AC^t + BC^t)\\ &= \operatorname{tr}(\lambda AC^t) + \operatorname{tr}(BC^t)\\ &= \lambda\operatorname{tr}(AC^t) + \operatorname{tr}(BC^t)\\ &= \lambda\langle A, C\rangle + \langle B, C\rangle .\end{aligned}$$

Ο διανυσματικός χώρος $\mathcal{M}_n$ είναι Ευκλείδειος, αφού $\dim(\mathcal{M}_n) = n$.

Διανυσματικοί χώροι με εσωτερικό γινόμενο

Πρόταση 1

Έστω $\mathcal{V} = (V, +, \cdot)$ ένας πραγματικός διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\rangle$. Τότε,

1. $\langle a\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, a\mathbf{y}\rangle = a\langle \mathbf{x}, \mathbf{y}\rangle$.
2. $\langle a\mathbf{x}, a\mathbf{y}\rangle = a^2\langle \mathbf{x}, \mathbf{y}\rangle$.
3. $\langle \mathbf{0}, \mathbf{x}\rangle = \langle \mathbf{x}, \mathbf{0}\rangle = 0$.
4. $\mathbf{x} \neq \mathbf{0} \Leftrightarrow \langle \mathbf{x}, \mathbf{x}\rangle > 0$.
5. $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z}\rangle = a\langle \mathbf{x}, \mathbf{z}\rangle + b\langle \mathbf{y}, \mathbf{z}\rangle$.
6. $\langle a_1\mathbf{x}_1 + \cdots a_n\mathbf{x}_n, \mathbf{y}\rangle = a_1\langle \mathbf{x}_1, \mathbf{y}\rangle + \cdots + a_n\langle \mathbf{x}_n, \mathbf{y}\rangle$.
7. $\langle \mathbf{x}, a\mathbf{y} + b\mathbf{z}\rangle = a\langle \mathbf{x}, \mathbf{y}\rangle + b\langle \mathbf{x}, \mathbf{z}\rangle$.
8. $\langle a\mathbf{x} + b\mathbf{y}, a\mathbf{x} + b\mathbf{y}\rangle = a^2\langle \mathbf{x}, \mathbf{x}\rangle + 2ab\langle \mathbf{x}, \mathbf{y}\rangle + b^2\langle \mathbf{y}, \mathbf{y}\rangle$.

Διανυσματικοί χώροι με εσωτερικό γινόμενο

Πρόταση 2 (συνέχεια)

9. $$\left\langle \sum_{p\in[n]} a_p \mathbf{x}_p, \sum_{q\in[m]} b_q \mathbf{y}_q \right\rangle = \sum_{(p,q)\in[n]\times[m]} a_p b_q \langle \mathbf{x}_p, \mathbf{y}_q \rangle.$$
10. $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \Leftrightarrow \langle a\mathbf{x}, b\mathbf{y} \rangle = 0.$
11. $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \Leftrightarrow \langle a\mathbf{x}, \mathbf{y} \rangle = 0,$
    $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \Leftrightarrow \langle \mathbf{x}, a\mathbf{y} \rangle = 0.$
12. $\langle \mathbf{x}, \mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{z} \rangle \Leftrightarrow \langle \mathbf{x} - \mathbf{y}, \mathbf{z} \rangle = 0.$

Διανυσματικοί χώροι με εσωτερικό γινόμενο

(Για το σύνηθες εσωτερικό γινόμενο στο $\mathbb{R}^3$.)

1. Ισχύει ότι

   $$\langle 2(1,2,3),(-1,-1,5)\rangle = \langle (2,4,6),(-1,-1,5)\rangle = (-2)+(-4)+30 =$$

   και

   $$2\langle (1,2,3),(-1,-1,5)\rangle = 2((-1)+(-2)+15) = 2\cdot 12 = 24,$$

   (επαληθεύοντας την ιδιότητα 1).
2. Ισχύει ότι

   $$\langle 2(1,3,5), 2(1,3,5)\rangle = 4\langle (1,3,5),(1,3,5)\rangle,$$

   (επαληθεύοντας την ιδιότητα 2).
   Πράγματι,

   $$2(1,3,5), 2(1,3,5) = \langle (2,6,10),(2,6,10)\rangle = 4+36+100 = 140,$$

   και

   $$4\langle (1,3,5),(1,3,5)\rangle = 4(4+9+25) = 4\cdot 35 = 140.$$
3. Ισχύει ότι

Πρόταση 3 (Ανισότητα Cauchy - Schwarz)

Έστω $(V, +, \cdot)$ ένας πραγματικός διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\rangle$. Τότε,

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \cdot \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}$$

όπου $|a|$ είναι η απόλυτη τιμή του a. Η ισότητα ισχύει μόνο αν τα $\mathbf{x}, \mathbf{y}$ είναι γραμμικώς εξαρτημένα.

Παρατήρηση: Η ανισότητα Cauchy - Schwarz θεωρείται από τις πιο σημαντικές ανισότητες. Με τη βοήθειά της, μπορούν να αποδειχθούν πολλές άλλες ανισότητες.

Παράδειγμα

Για κάθε $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n \in \mathbb{R}$ ισχύει ότι

$$(x_1 y_1 + \cdots + x_n y_n)^2 \leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2).$$

Απόδειξη. Στον $\mathbb{R}^n$ με το σύνηθες εσωτερικό γινόμενο, αν $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ και $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, τότε έχουμε ότι

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n = \sum_{k=1}^{n} x_k y_k,$$

$$\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{x_1^2 + \cdots + x_n^2} = \left(\sum_{k=1}^{n} x_k^2 \right)^{\frac{1}{2}}$$

$$\sqrt{\langle \mathbf{y}, \mathbf{y} \rangle} = \sqrt{y_1^2 + \cdots + y_n^2} = \left(\sum_{k=1}^{n} y_k^2 \right)^{\frac{1}{2}}.$$

Επομένως,

$$\left|\sum_{k=1}^{n} x_k y_k\right| \leq \left(\sum_{k=1}^{n} x_k^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} y_k^2\right)^{\frac{1}{2}},$$

ή, ισοδύναμα, υψώνοντας στο τετράγωνο,

$$(x_1 y_1 + \cdots + x_n y_n)^2 \leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2).$$

Εφαρμογή της ανισότητας C-S

Να δειχθεί ότι για κάθε $a, b \in \mathbb{R}$ ισχύει ότι $a^2 + b^2 + (1 - a - b)^2 \geq \frac{1}{3}$.

Λύση Από την προηγούμενη ανισότητα για $x_1 = a$, $x_2 = b$, $x_3 = 1 - a - b$ και $y_1 = 1$, $y_2 = 1$ και $y_3 = 1$ ισχύει ότι

$$\begin{aligned}(a \cdot 1 + b \cdot 1 + (1 - a - b) \cdot 1)^2 &\leq (a^2 + b^2 + (1 - a - b)^2) \cdot (1^2 + 1^2 + 1^2)\\ 1^2 &\leq (a^2 + b^2 + (1 - a - b)^2) \cdot 3\\ \frac{1}{3} &\leq a^2 + b^2 + (1 - a - b)^2\end{aligned}$$

Η ισότητα ισχύει αν και μόνο αν τα διανύσματα $(a, b, 1 - a - b)$ και $(1, 1, 1)$ είναι γραμμικώς εξαρτημένα, δηλαδή όταν $a = b$ $= 1 - a - b = \lambda$ για κάποιο $\lambda \in \mathbb{R}$, επομένως

$$\lambda = 1 - \lambda - \lambda \Leftrightarrow \lambda = \frac{1}{3}$$

άρα, η ισότητα ισχύει όταν $a = b = \frac{1}{3}$. Πράγματι, ισχύει ότι

$$\left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(1 - \frac{1}{3} - \frac{1}{3}\right)^2 = \frac{1}{3}.$$

Εφαρμογή της ανισότητας C-S

Να βρεθεί η ελάχιστη τιμή της παράστασης

$$(x+1)^2+(y-2)^2+(2x+3y-8)^2$$

Λύση Αν τεθεί $\mathbf{x}=(x+1, y-2, 2x+3y-8)$ τότε

$$\langle \mathbf{x}, \mathbf{x}\rangle=(x+1)^2+(y-2)^2+(2x+3y-8)^2.$$

Προκειμένου να εφαρμόσουμε την προηγούμενη ιδέα αναζητούμε διάνυσμα $\mathbf{y}=(c_1, c_2, c_3)\in\mathbb{R}^3\setminus\{\mathbf{0}\}$ έτσι ώστε η παράσταση Π όπου:

$$\begin{aligned}\Pi &= \langle \mathbf{x}, \mathbf{y}\rangle=\langle (x+1, y-2, 2x+3y-8), (c_1, c_2, c_3)\rangle\\ &= c_1(x+1)+c_2(y-2)+c_3(2x+3y-8)\end{aligned}$$

να μην περιέχει x και y. Ισχύει ότι

$$\Pi=(c_1+2c_3)x+(c_2+3c_3)y+c_1-2c_2-8c_3.$$

Πρέπει

$$\begin{cases} c_1+2c_3=0\\ c_2+3c_3=0\end{cases} \Leftrightarrow \begin{cases} c_1=-2c_3\\ c_2=-3c_3\end{cases}$$

Άρα, αν επιλέξουμε για παράδειγμα $c_3 = -1$ έχουμε ότι $c_1 = 2$, $c_2 = 3$, $\mathbf{y} = (2, 3, -1)$ και $\Pi = 2 - 2 \cdot 3 - 8 \cdot (-1) = 4$.
Για τα διανύσματα $\mathbf{x} = (x+1, y-2, 2x+3y-8)$ και $\mathbf{y} = (2, 3, -1)$ από την ανισότητα Cauchy - Schwarz προκύπτει ότι

$$\begin{aligned}
&\langle \mathbf{x}, \mathbf{y} \rangle^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle \Leftrightarrow \\
&(2(x+1) + 3(y-2) + (-1)(2x+3y-8))^2 \leq \\
&\qquad ((x+1)^2 + (y-2)^2 + (2x+3y-8)^2)(2^2 + 3^2 + (-1)^2)
\end{aligned}$$

ή, ισοδύναμα

$$4^2 \leq ((x+1)^2 + (y-2)^2 + (2x+3y-8)^2) \cdot 14$$

άρα,

$$(x+1)^2 + (y-2)^2 + (2x+3y-8)^2 \geq \frac{16}{14}$$

Η ισότητα ισχύει αν και μόνο αν τα διανύσματα $\mathbf{x} = (x+1, y-2, 2x+3y-8)$ και $\mathbf{y} = (2, 3, -1)$ είναι γραμμικώς εξαρτημένα, δηλαδή όταν $\mathbf{x} = \lambda \mathbf{y}$, για κάποιο $\lambda \in \mathbb{R}$, δηλαδή

$$\begin{cases} x+1 = 2\lambda \\ y-2 = 3\lambda \\ 2x+3y-8 = (-1)\lambda \end{cases}$$

οπότε

$$4 = \langle \mathbf{x}, \mathbf{y} \rangle = \langle \lambda \mathbf{y}, \mathbf{y} \rangle = \lambda \langle \mathbf{y}, \mathbf{y} \rangle = 14\lambda$$

επομένως $\lambda = \frac{4}{14}$ άρα

$$x = -\frac{6}{14}, \quad y = \frac{40}{14}.$$

Τότε $(x+1)^2 = \left(2 \cdot \frac{4}{14}\right)^2$, $(y-2)^2 = \left(3 \cdot \frac{4}{14}\right)^2$, $(2x+3y-8)^2 = \left(-\frac{4}{14}\right)^2$ και

$$(x+1)^2 + (y-2)^2 + (2x+3y-8)^2 = \frac{8^2+12^2+4^2}{14^2} = \frac{64+144+16}{14^2} = \frac{16}{14}.$$

Διανυσματικοί χώροι με norm

Έστω $\mathcal{V} = (V, +, \cdot_{\mathbb{R}})$ ένας πραγματικός διανυσματικός χώρος και $\| \; \| : V \to \mathbb{R}$.

Η απεικόνιση $\| \; \|$ λέγεται **norm** (ή **μήκος**) επί του $\mathcal{V}$ αν και μόνο αν για κάθε $\mathbf{x}, \mathbf{y} \in V$ και για κάθε $a \in \mathbb{R}$ ισχύει ότι

1. $\|\mathbf{x}\| \geq 0$,
2. $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$,
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (τριγωνική ανισότητα),
4. $\|a\mathbf{x}\| = |a| \, \|\mathbf{x}\|$.

Ο μη αρνητικός, πραγματικός αριθμός $\|\mathbf{x}\|$ ονομάζεται **norm** (ή **μήκος**) **του διανύσματος** $\mathbf{x}$.

Το $(\mathcal{V}, \| \; \|)$ λέγεται **πραγματικός διανυσματικός χώρος με norm**.

Διανυσματικοί χώροι με norm

Παράδειγμα 1

Η απεικόνιση $\| \ \| : \mathbb{R}^n \to \mathbb{R}$ όπου, αν $\mathbf{x} = (x_1, \ldots, x_n)$, τότε

$$\|\mathbf{x}\| = |x_1| + \cdots + |x_n|$$

είναι μια norm επί του $(\mathbb{R}^n, +, \cdot)$.

Διανυσματικοί χώροι με norm

Πράγματι, αν $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$, τότε

1. $$\|\mathbf{x}\| = |x_1| + \cdots + |x_n| \geq 0.$$

2. $$\begin{aligned} \|\mathbf{x}\| = 0 &\Leftrightarrow |x_1| + \cdots + |x_n| = 0 \\ &\Leftrightarrow x_1 = \cdots = x_n = 0 \\ &\Leftrightarrow \mathbf{x} = (0, \ldots, 0) = \mathbf{0}. \end{aligned}$$

3. $$\begin{aligned} \|\mathbf{x} + \mathbf{y}\| &= \|(x_1 + y_1, \ldots, x_n + y_n)\| \\ &= |x_1 + y_1| + \cdots + |x_n + y_n| \\ &\leq |x_1| + |y_1| + \cdots + |x_n| + |y_n| \\ &= (|x_1| + \cdots + |x_n|) + (|y_1| + \cdots + |y_n|) \\ &= \|\mathbf{x}\| + \|\mathbf{y}\|. \end{aligned}$$

Διανυσματικοί χώροι με norm

4.

$$
\begin{aligned}
\|a\mathbf{x}\| &= \|a(x_1, \ldots, x_n)\| = \|(ax_1, \ldots, ax_n)\| \\
&= |ax_1| + \cdots + |ax_n| = |a|\,|x_1| + \cdots + |a|\,|x_n| \\
&= |a|\,(|x_1| + \cdots + |x_n|) = |a|\,\|\mathbf{x}\|\,.
\end{aligned}
$$

Παρατήρηση: Η norm αυτή ονομάζεται **1-norm** και συνήθως συμβολίζεται με $\| \ \|_1$. Προφανώς, αυτό το παράδειγμα για $n = 1$ αποδεικνύει ότι η απόλυτη τιμή είναι norm στο $\mathbb{R}$.

Διανυσματικοί χώροι με norm

Παράδειγμα 2

Η απεικόνιση $\| \; \| : \mathbb{R}^n \to \mathbb{R}$ όπου, αν $\mathbf{x} = (x_1, \dots, x_n)$, τότε

$$\|\mathbf{x}\| = \max\{|x_1|, \dots, |x_n|\}$$

είναι μια norm επί του $(\mathbb{R}^n, +, \cdot)$.

Διανυσματικοί χώροι με norm

Πράγματι, αν $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$, τότε

1. $$\|\mathbf{x}\| = \max\{|x_1|, \ldots, |x_n|\} \geq 0.$$

2. $$\begin{aligned} \|\mathbf{x}\| = 0 &\Leftrightarrow \max\{|x_1|, \ldots, |x_n|\} = 0 \\ &\Leftrightarrow |x_1| = \cdots = |x_n| = 0 \\ &\Leftrightarrow \mathbf{x} = (0, \ldots, 0) = \mathbf{0}. \end{aligned}$$

3. $$\begin{aligned} \|\mathbf{x} + \mathbf{y}\| &= \|(x_1 + y_1, \ldots, x_n + y_n)\| \\ &= \max\{|x_1 + y_1|, \cdots, |x_n + y_n|\} \\ &\leq \max\{|x_1| + |y_1|, \cdots, |x_n| + |y_n|\} \\ &\leq \max\{|x_1|, \ldots, |x_n|\} + \max\{|y_1|, \ldots, |y_n|\} \\ &= \|\mathbf{x}\| + \|\mathbf{y}\|. \end{aligned}$$

Διανυσματικοί χώροι με norm

4.

$$\begin{aligned}
\|a\mathbf{x}\| &= \|a(x_1, \ldots, x_n)\| \\
&= \|(ax_1, \ldots, ax_n)\| \\
&= \max\{|ax_1|, \cdots, |ax_n|\} \\
&= \max\{|a|\,|x_1|, \cdots + |a|\,|x_n|\} \\
&= |a| \max\{|x_1|, \cdots, |x_n|\} \\
&= |a|\,\|\mathbf{x}\|.
\end{aligned}$$

Παρατήρηση: Η norm αυτή ονομάζεται **max norm** και συνήθως συμβολίζεται με $\| \|_\infty$.

Διανυσματικοί χώροι με norm

Παράδειγμα 3

Αποδεικνύεται (άσκηση) ότι η απεικόνιση $\| \ \| : \mathbb{R}^n \to \mathbb{R}$ όπου, αν $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, τότε

$$\|\mathbf{x}\| = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}$$

όπου $p \in \mathbb{N}^*$ είναι μια norm επί του $(\mathbb{R}^n, +, \cdot)$, η οποία ονομάζεται *p*-**norm** και συμβολίζεται με $\| \ \|_p$.

Παράδειγμα 4

Η απεικόνιση $\| \; \| : V \to \mathbb{R}$ όπου $(V, +, \cdot)$ το σύνολο των συνεχών συναρτήσεων που ορίζονται[α΄] στο $[a, b]$ και αν $f \in V$ τότε

$$\|f\| = \max\{|f(x)| : x \in [a, b]\},$$

είναι μια norm επί του V.

[α΄] Κάθε συνεχής συνάρτηση ορισμένη σε κλειστό διάστημα είναι φραγμένη.

Πράγματι, αν $f, g \in V$ τότε

1. $$\|f\| = \max\{|f(x)| : x \in [a,b]\} \geq 0.$$

2. $$\begin{aligned} \|f\| = 0 &\Leftrightarrow \max\{|f(x)| : x \in [a,b]\} = 0 \\ &\Leftrightarrow |f(x)| = 0 \text{ για κάθε } x \in [a,b] \\ &\Leftrightarrow f(x) = 0 \text{ για κάθε } x \in [a,b]. \end{aligned}$$

3. $$\begin{aligned} \|f+g\| &= \max\{|f(x)+g(x)| : x \in [a,b]\} \\ &\leq \max\{|f(x)| + |g(x)| : x \in [a,b]\} \\ &\leq \max\{|f(x)| : x \in [a,b]\} + \max\{|g(x)| : x \in [a,b]\} \\ &= \|f\| + \|g\|. \end{aligned}$$

Διανυσματικοί χώροι με norm

4.

$$\begin{aligned} \|af\| &= \max\{|af(x)| : x \in [a,b]\} \\ &= \max\{|a|\,|f(x)| : x \in [a,b]\} \\ &= |a| \max\{|f(x)| : x \in [a,b]\} \\ &= |a|\,\|f\| . \end{aligned}$$

Διανυσματικοί χώροι με norm

Για τις norm εύκολα αποδεικνύονται οι παρακάτω ιδιότητες:

Πρόταση 4

Έστω $\mathcal{V} = (V, +, \cdot_{\mathbb{F}})$ ένας πραγματικός διανυσματικός χώρος με norm. Τότε,

1. $\|\mathbf{0}\| = 0$.
2. $\mathbf{x} \neq \mathbf{0} \Leftrightarrow \|\mathbf{x}\| > 0$.
3. $\|-\mathbf{x}\| = \|\mathbf{x}\|$.
4. $\|\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n\| \leq \|\mathbf{x}_1\| + \|\mathbf{x}_2\| + \cdots + \|\mathbf{x}_n\|$.
5. $|\ \|\mathbf{x}\| - \|\mathbf{y}\|\ | \leq \|\mathbf{x} + \mathbf{y}\|$.
6. $|\ \|\mathbf{x}\| - \|\mathbf{y}\|\ | \leq \|\mathbf{x} - \mathbf{y}\|$.

Διανυσματικοί χώροι με norm

Παρατήρηση: Επειδή ορίζονται τόσες πολλές norm σε ένα διανυσματικό χώρο V (για παράδειγμα στον $\mathbb{R}^n$ ορίζονται άπειρες το πλήθος norm) τίθεται το ερώτημα κατά πόσον οι διαφορετικές norm είναι συμβατές μεταξύ τους. Το επόμενο αποτέλεσμα μας εξασφαλίζει ότι σε ένα διανυσματικό χώρο πεπερασμένης διάστασης όλες οι norm είναι κατά κάποιο τρόπο ισοδύναμες.

Πρόταση 5

Αν $\| \, \|_a$ και $\| \, \|_b$ είναι δύο norm σε ένα διανυσματικό χώρο πεπερασμένης διάστασης, τότε υπάρχουν σταθερές $m, M \in \mathbb{R}$ τέτοιες ώστε για κάθε $\mathbf{x} \in \mathbb{R}^n$ να ισχύει ότι

$$m \|\mathbf{x}\|_a \leq \|\mathbf{x}\|_b \leq M \|\mathbf{x}\|_a .$$

Εσωτερικό γινόμενο και norm

Στα προηγούμενα, το εσωτερικό γινόμενο και η norm ορίσθηκαν ανεξάρτητα μεταξύ τους. Παρακάτω θα δούμε ένα ‘‘φυσικό’’ τρόπο να ορίσουμε μια norm σε ένα διανυσματικό χώρο με εσωτερικό γινόμενο.

Έστω $\mathcal{V} = (V, +, \cdot)$ ένας πραγματικός δ.χ. με εσωτερικό γινόμενο $\langle\rangle$. Για κάθε $\mathbf{x} \in V$ ορίζουμε

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

Η απεικόνιση $\| \ \| : V \to \mathbb{R}$ είναι μια norm επί του V.

Εσωτερικό γινόμενο και norm

Πράγματι, για κάθε $\mathbf{x}, \mathbf{y} \in V$ ισχύει ότι

1. $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \geq 0$.
2. $\|\mathbf{x}\| = 0 \Leftrightarrow \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = 0 \Leftrightarrow \langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$.
3. $\|\mathbf{x} + \mathbf{y}\| = \sqrt{\langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle + 2 \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle} =$
   $\sqrt{\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}^2 + 2 \langle \mathbf{x}, \mathbf{y} \rangle + \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}^2}$
   $\leq \sqrt{\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}^2 + 2 |\langle \mathbf{x}, \mathbf{y} \rangle| + \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}^2} \leq$
   $\sqrt{\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}^2 + 2\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}\sqrt{\langle \mathbf{y}, \mathbf{y} \rangle} + \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}^2}$
   $= \sqrt{\left(\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} + \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}\right)^2} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} + \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle} = \|\mathbf{x}\| + \|\mathbf{y}\|$.
4. $\|a\mathbf{x}\| = \sqrt{\langle a\mathbf{x}, a\mathbf{x} \rangle} = \sqrt{a^2 \langle \mathbf{x}, \mathbf{x} \rangle} = |a| \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = |a| \, \|\mathbf{x}\|$.

(Προφανώς, ισχύει ότι $\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle$.)

Παρατήρηση: Στα επόμενα, αν δεν αναφέρεται διαφορετικά, θα χρησιμοποιείται ο παραπάνω ορισμός της norm για τα στοιχεία ενός διανυσματικού χώρου $(\mathcal{V}, \langle\rangle)$.

Εσωτερικό γινόμενο και norm

Από την Πρόταση 3 προκύπτει μια μορφή της ανισότητας Cauchy - Schwarz, η οποία συνδέει εσωτερικό γινόμενο και norm διανυσμάτων.

Πρόταση 6 (Ανισότητα Cauchy - Schwarz (2η εκδοχή))

Για κάθε $\mathbf{x}, \mathbf{y} \in V$ *ισχύει ότι*

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\| .$$

Η ισότητα ισχύει μόνο αν τα διανύσματα $\mathbf{x}, \mathbf{y}$ *είναι γραμμικώς εξαρτημένα.*

Γωνία διανυσμάτων

Έστω ο Ευκλείδειος χώρος $(\mathcal{V}, \langle\rangle)$ και $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$. Τότε, αφού

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \, \|\mathbf{y}\| \Leftrightarrow -1 \leq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \, \|\mathbf{y}\|} \leq 1,$$

υπάρχει $\theta \in [0, \pi]$, τέτοιο ώστε

$$\cos\theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \, \|\mathbf{y}\|}$$

Ο αριθμός θ ονομάζεται **γωνία** των διανυσμάτων $\mathbf{x}, \mathbf{y}$ και το πηλίκο αυτό ονομάζεται **συνημίτονο της γωνίας των** $\mathbf{x}, \mathbf{y}$.

Γωνία διανυσμάτων

Παράδειγμα 1

Για $\mathbf{x} = (1,1,4)$ και $\mathbf{y} = (2,7,7)$ είναι

$$
\begin{aligned}
\cos\theta &= \frac{\langle (1,1,4),(2,7,7)\rangle}{\sqrt{\langle (1,1,4),(1,1,4)\rangle}\sqrt{\langle (2,7,7),(2,7,7)\rangle}} \\
&= \frac{1\cdot 2 + 1\cdot 7 + 4\cdot 7}{\sqrt{1^2+1^2+4^2}\sqrt{2^2+7^2+7^2}} \\
&= \frac{37}{6\sqrt{51}}
\end{aligned}
$$

οπότε $\theta = \arccos\dfrac{37}{6\sqrt{51}} \simeq \dfrac{\pi}{5.943} \simeq \dfrac{\pi}{6}$.

Γωνία διανυσμάτων

Παράδειγμα 2

Για $\mathbf{x} = (1, -1, 2)$ και $\mathbf{y} = (2, 6, 2)$ είναι

$$\begin{aligned}
\cos\theta &= \frac{\langle (1,-1,2),(2,6,2)\rangle}{\sqrt{\langle (1,-1,2),(1,-1,2)\rangle}\sqrt{\langle (2,6,2),(2,6,2)\rangle}} \\
&= \frac{1\cdot 2 + (-1)\cdot 6 + 2\cdot 2}{\sqrt{1^2+(-1)^2+2^2}\sqrt{2^2+6^2+2^2}} \\
&= \frac{0}{\sqrt{6}\sqrt{44}}
\end{aligned}$$

οπότε $\theta = \arccos 0 = \dfrac{\pi}{2}$.

Γωνία διανυσμάτων

Μήτρα όρων – εγγράφων

Έστω ότι έχουμε μια συλλογή από n έγγραφα κειμένου $D_1, D_2, \ldots, D_n$ στην οποία πρόκειται να κάνουμε αρκετές αναζητήσεις με λέξεις κλειδιά (keywords) ή όρους (terms). Ένας κλασσικός τρόπος για να αυξήσουμε την ταχύτητα των αναζητήσεων είναι να κάνουμε προεπεξεργασία των εγγράφων και να φτιάξουμε μια μήτρα όρων – εγγράφων (term-by-document) D η οποία έχει διαστάσεις $m \times n$, όπου m είναι ο αριθμός των διαφορετικών λέξεων κλειδιών ή όρων, στην οποία το στοιχείο της i-γραμμής και j-στήλης είναι ο αριθμός των εμφανίσεων του όρου t_i στο έγγραφο D_j.

Γωνία διανυσμάτων

Έτσι, για παράδειγμα σε μια συλλογή με 9 έγγραφα $D_1, D_2, \ldots, D_9$ για την οποία μας ενδιαφέρει η αναζήτηση με λέξεις κλειδιά που μπορούν να επιλεγούν μεταξύ 5 όρων t_1, t_2, t_3, t_4, t_5 δημιουργήθηκε η παρακάτω μήτρα όρων - εγγράφων:

$$D = \begin{array}{c|ccccccccc} & D_1 & D_2 & D_3 & D_4 & D_5 & D_6 & D_7 & D_8 & D_9 \\ \hline t_1 & 1 & 1 & 2 & 0 & 1 & 0 & 1 & 0 & 1 \\ t_2 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ t_3 & 0 & 2 & 0 & 2 & 0 & 1 & 0 & 1 & 1 \\ t_4 & 1 & 0 & 1 & 0 & 1 & 0 & 2 & 1 & 0 \\ t_5 & 1 & 2 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{array}$$

(Σύμφωνα με την μήτρα D, το έγγραφο D_3 περιέχει δύο εμφανίσεις του όρου t_1, καμία εμφάνιση των όρων t_2, t_3 και από μια εμφάνιση των όρων t_4, t_5.)
Αν υποθέσουμε ότι αναζητούμε τους όρους t_2 και t_4 και t_5 ποιο έγγραφο από τα 9 ταιριάζει καλύτερα στην αναζήτησή μας;

Γωνία διανυσμάτων

Λύση. Εξετάζοντας μια-μία όλες τις στήλες της μήτρας D παρατηρούμε ότι κανένα έγγραφο δεν περιέχει ακριβώς αυτούς τους όρους.
Σύμφωνα με την αναπαράσταση αυτή καθένα από τα έγγραφα D_1, D_2, …, D_9 αντιστοιχεί σε ένα διάνυσμα που ανήκει σε ένα χώρο 5 διαστάσεων με άξονες τα t_1, t_2, …, t_5, ή ισοδύναμα ένα διάνυσμα του $\mathbb{R}^5$. Προκειμένου να επιλέξουμε ποιό από τα D_1, D_2, …, D_9 ταιριάζει καλύτερα στην αναζήτησή μας, θεωρουμε επίσης ότι η αναζήτησή μας αντιστοιχεί στο διάνυσμα

$$\mathbf{q} = (0, 1, 0, 1, 1)$$

(Στην θέση i του $\mathbf{q}$ εμφανίζεται 1 αν η αναζήτησή μας περιέχει τον όρο t_i, αλλιώς εμφανίζεται 0.)
Υπολογίζουμε την ομοιότητα ανάμεσα στο q και τα έγγραφα D_1, D_2, …, D_9 με βάση το συνημίτονο της γωνίας που σχηματίζει διάνυσμα q με καθένα από τα διανύσματα-στήλες που αντιστοιχούν σε κάθε έγγραφο D_i, $i = 1, 2, \ldots, 9$.

Γωνία διανυσμάτων

Έτσι, αν θ_1, θ_2, …, θ_9 είναι οι γωνίες μεταξύ του q και των D_1, D_2, …, D_9 αντίστοιχα, τότε

$$\cos\theta_1 = \frac{\langle \mathbf{q}, D_1\rangle}{\|\mathbf{q}\|\,\|D_1\|} = \frac{\langle (0,1,0,1,1),(1,0,0,1,1)\rangle}{\|0,1,0,1,1\|\,\|1,0,0,1,1\|} = \frac{2}{3} \simeq 0.666666$$

$$\cos\theta_2 = \frac{\langle \mathbf{q}, D_2\rangle}{\|\mathbf{q}\|\,\|D_2\|} = \frac{\langle (0,1,0,1,1),(1,1,2,0,2)\rangle}{\|0,1,0,1,1\|\,\|1,1,2,0,2\|} = \frac{3}{\sqrt{30}} \simeq 0.547723$$

$$\cos\theta_3 = \frac{\langle \mathbf{q}, D_3\rangle}{\|\mathbf{q}\|\,\|D_3\|} = \frac{\langle (0,1,0,1,1),(2,0,0,1,1)\rangle}{\|0,1,0,1,1\|\,\|2,0,0,1,1\|} = \frac{2}{\sqrt{18}} \simeq 0.471405$$

$$\cos\theta_4 = \frac{\langle \mathbf{q}, D_4\rangle}{\|\mathbf{q}\|\,\|D_4\|} = \frac{\langle (0,1,0,1,1),(0,1,2,0,0)\rangle}{\|0,1,0,1,1\|\,\|0,1,2,0,0\|} = \frac{1}{\sqrt{15}} \simeq 0.258199$$

$$\cos\theta_5 = \frac{\langle \mathbf{q}, D_5\rangle}{\|\mathbf{q}\|\,\|D_5\|} = \frac{\langle (0,1,0,1,1),(1,0,0,1,0)\rangle}{\|0,1,0,1,1\|\,\|1,0,0,1,0\|} = \frac{1}{\sqrt{6}} \simeq 0.408248$$

$$\cos\theta_6 = \frac{\langle \mathbf{q}, D_6\rangle}{\|\mathbf{q}\|\,\|D_6\|} = \frac{\langle (0,1,0,1,1),(0,1,1,0,1)\rangle}{\|0,1,0,1,1\|\,\|0,1,1,0,1\|} = \frac{2}{3} \simeq 0.666666$$

Γωνία διανυσμάτων

$$\cos\theta_7 = \frac{\langle \mathbf{q}, D_7 \rangle}{\|\mathbf{q}\| \, \|D_7\|} = \frac{\langle (0,1,0,1,1), (1,1,0,2,0) \rangle}{\|0,1,0,1,1\| \, \|1,1,0,2,0\|} = \frac{1}{\sqrt{2}} \simeq 0.707107$$

$$\cos\theta_8 = \frac{\langle \mathbf{q}, D_8 \rangle}{\|\mathbf{q}\| \, \|D_8\|} = \frac{\langle (0,1,0,1,1), (0,0,1,1,0) \rangle}{\|0,1,0,1,1\| \, \|0,0,1,1,0\|} = \frac{1}{\sqrt{6}} \simeq 0.408248$$

$$\cos\theta_9 = \frac{\langle \mathbf{q}, D_9 \rangle}{\|\mathbf{q}\| \, \|D_9\|} = \frac{\langle (0,1,0,1,1), (1,0,1,0,1) \rangle}{\|0,1,0,1,1\| \, \|1,0,1,0,1\|} = \frac{1}{3} \simeq 0.333333$$

Με βάση τους υπολογισμούς μας, η μικρότερη γωνία εμφανίζεται μεταξύ του $\mathbf{q}$ και του D_7 (διότι όσο μεγαλύτερο το συνημίτονο της γωνίας θ, τόσο μικρότερη η γωνία θ). Επομένως, το έγγραφο D_7 ταιριάζει καλύτερα στην αναζήτηση των όρων t_2, t_4 και t_5 σε σχέση με τα υπόλοιπα έγγραφα της συλλογής μας.

Επιπλέον, με βάση τα συνημίτονα που υπολογίσαμε μπορούμε να διάταξουμε σε φθίνουσα σειρά (ως προς το ταίριασμα με την αναζήτησή μας) τα έγγραφα της συλλογής ως εξής:

$$D_7, D_1, D_6, D_2, D_3, D_5, D_8, D_9, D_4.$$

2η ΔΙΑΛΕΞΗ
ΕΣΩΤΕΡΙΚΟ ΓΙΝΟΜΕΝΟ

- **Μετρικοί χώροι**

Διανυσματικοί μετρικοί χώροι

Έστω $E \neq \emptyset$. Η απεικόνιση $d : E \times E \to \mathbb{R}$ ονομάζεται **μετρική** επί του E αν και μόνο αν, για κάθε $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E$ ισχύουν:

1. $d(\mathbf{x}, \mathbf{y}) \geq 0$,
2. $d(\mathbf{x}, \mathbf{y}) = 0 \Rightarrow \mathbf{x} = \mathbf{y}$,
3. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$,
4. $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ (τριγωνική ανισότητα).

Ο αριθμός $d(\mathbf{x}, \mathbf{y})$ ονομάζεται **απόσταση** των $\mathbf{x}, \mathbf{y}$.

Το ζεύγος (E, d) ονομάζεται **μετρικός χώρος** αν και μόνο αν η d είναι μια μετρική επί του E.
Επίσης, αν $\emptyset \neq E_1 \subseteq E$, d μια μετρική επί του E και d_1 ο περιορισμός της d στο $E_1 \times E_1$, τότε αποδεικνύεται ότι η d_1 είναι μια μετρική επί του E_1. Ο μετρικός χώρος (E_1, d_1) ονομάζεται **μετρικός υπόχωρος** του (E, d).

Αν $\mathcal{V} = (V, +, \cdot_F)$, είναι ένας διανυσματικός χώρος και d είναι μια μετρική επί του V, το ζεύγος $(\mathcal{V}, d)$ ονομάζεται **διανυσματικός μετρικός χώρος**.

Διανυσματικοί μετρικοί χώροι

Παράδειγμα 1

Ένας διανυσματικός μετρικός χώρος, για τον $(\mathbb{R}^2, +, \cdot_{\mathbb{R}})$ δημιουργείται από τη μετρική

$$d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

η οποία ορίζεται ως εξής:
Για $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$,

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \in \mathbb{R}$$

Διανυσματικοί μετρικοί χώροι

Πράγματι, ισχύει ότι

1. $d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \geq 0$
2. $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} = 0 \Leftrightarrow$ $(x_1 - y_1)^2 + (x_2 - y_2)^2 = 0 \Leftrightarrow$ $x_1 - y_1 = 0$ και $x_2 - y_2 = 0 \Leftrightarrow x_1 = y_1$ και $x_2 = y_2 \Leftrightarrow (x_1, y_1) = (x_2, y_2) \Leftrightarrow \mathbf{x} = \mathbf{y}$
3. $d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2} = d(\mathbf{y}, \mathbf{x})$
4. Επίσης αποδεικνύεται (άσκηση) ότι

$$d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$$

Η επόμενη πρόταση συνδέει το εσωτερικό γινόμενο και την norm διανυσμάτων με τις μετρικές.

Πρόταση 7

Εστω $\mathcal{V} = (V, +, \cdot)$ ένας πραγματικός διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\ \rangle$. Η απεικόνιση $d : V \times V \to \mathbb{R}$, με

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|, \qquad \text{για κάθε } \mathbf{x}, \mathbf{y} \in V,$$

είναι μια μετρική επί του V.

Διανυσματικοί μετρικοί χώροι

Απόδειξη Έστω $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$. Ισχύει ότι

1. $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \geq 0$.
2. $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \|\mathbf{x} - \mathbf{y}\| = 0 \Leftrightarrow \mathbf{x} - \mathbf{y} = \mathbf{0} \Leftrightarrow \mathbf{x} = \mathbf{y}$.
3. $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \|(-1)(\mathbf{y} - \mathbf{x})\| = |-1| \, \|\mathbf{y} - \mathbf{x}\| = d(\mathbf{y}, \mathbf{x})$.
4. $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x} - \mathbf{z} + \mathbf{z} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{z}\| + \|\mathbf{z} - \mathbf{y}\| = d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$.

Παρατήρηση: Ένας πραγματικός διανυσματικός χώρος με εσωτερικό γινόμενο είναι και μετρικός χώρος, αφού μπορούμε να ορίσουμε $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle$

Για 25 αυτοκίνητα γνωρίζουμε τις επιδόσεις τους με βάση διάφορους δείκτες. Να χωρισθούν τα αυτοκίνητα σε ομάδες όμοιων αυτοκινήτων (με βάση την απόσταση).

$\mathbf{v}_1$	1.5	2.1
$\mathbf{v}_2$	1.3	3.15
$\mathbf{v}_3$	1.4	4.1
$\mathbf{v}_4$	2.1	2.2
$\mathbf{v}_5$	2.2	5.1
$\mathbf{v}_6$	3.15	8.1
$\mathbf{v}_7$	3.2	9.2
$\mathbf{v}_8$	3.3	10.25
$\mathbf{v}_9$	3.5	11.15
$\mathbf{v}_{10}$	3.1	12.1
$\mathbf{v}_{11}$	3.4	13.2
$\mathbf{v}_{12}$	4	10.2
$\mathbf{v}_{13}$	4.1	11.1
$\mathbf{v}_{14}$	4.15	12.5
$\mathbf{v}_{15}$	4.2	13.6
$\mathbf{v}_{16}$	5.1	9.7
$\mathbf{v}_{17}$	5.3	11.1
$\mathbf{v}_{18}$	6.1	1.4
$\mathbf{v}_{19}$	6.4	2.3
$\mathbf{v}_{20}$	6.2	3.8
$\mathbf{v}_{21}$	7.1	1.4
$\mathbf{v}_{22}$	7.2	2.5
$\mathbf{v}_{23}$	7.15	3.3
$\mathbf{v}_{24}$	8.1	3.1
$\mathbf{v}_{25}$	8.4	4.5

(Εδώ φαίνονται οι δύο πρώτοι δείκτες)

Έστω $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ ένα σύνολο n σημείων του $\mathbb{R}^m$. Το **κέντρο βάρους των** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ ορίζεται ως

$$\mathbf{C} = \frac{1}{n}\left(\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n\right)$$

Για παράδειγμα, για τα σημεία
$\mathbf{v}_1 = (1, 2, 1), \mathbf{v}_2 = (1, 1, 1), \mathbf{v}_3 = (-1, 2, 2), \mathbf{v}_4 = (2, 1, -1)$ του $\mathbb{R}^3$ το κέντρο βάρους είναι το σημείο

$$\mathbf{C} = \frac{1}{4}(\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 + \mathbf{v}_4) = \left(\frac{3}{4}, \frac{6}{4}, \frac{3}{4}\right)$$

Να χωρισθούν τα 25 αυτοκίνητα $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_{25}$ σε 3 ομάδες (συστάδες) έτσι ώστε τα αθροίσματα των αποστάσεων κάθε σημείου από το κέντρο βάρους της ομάδας στην οποία ανήκει να είναι όσο το δυνατόν μικρότερο.

Ένας απλός επαναληπτικός αλγόριθμος για την διαμέριση σημείων του $\mathbb{R}^n$ σε k ομάδες είναι ο λεγόμενος **αλγόριθμος** K-**means**.
Για την εφαρμογή του αλγορίθμου πρέπει να επιλέξουμε το αριθμό των ομάδων k καθώς επίσης και την μετρική που θα χρησιμοποιήσουμε για τον υπολογισμό των αποστάσεων.
Η διατύπωση του αλγορίθμου K-means είναι πολύ απλή:

Είσοδος: Ένα σύνολο σημείων δεδομένων $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ του $\mathbb{R}^m$, ένας φυσικός αριθμός k
Έξοδος: Ένα σύνολο k σημείων κέντρων βαρών $\mathbf{C}_1$, $\mathbf{C}_2$, ..., $\mathbf{C}_k$.

1. Αρχικά επιλέγονται τυχαία k σημεία $\mathbf{C}_1$, $\mathbf{C}_2$, ..., $\mathbf{C}_k$ του $\mathbb{R}^m$, τα οποία αποτελούν τα προσωρινά κέντρα βάρους των ζητούμενων k ομάδων. Για κάθε σημείο $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_n$ υπολογίζονται οι αποστάσεις του από τα k κέντρων βαρών $\mathbf{C}_1$, $\mathbf{C}_2$, ..., $\mathbf{C}_k$ και τοποθετείται στην ομάδα με το πλησιέστερο κέντρο βάρους.
2. Στην συνέχεια υπολογίζονται τα κέντρα βάρους $\mathbf{C}_1$, $\mathbf{C}_2$, ..., $\mathbf{C}_k$ κάθε ομάδας με βάση τα σημεία $\mathbf{v}_i$ που τοποθετήθηκαν σε αυτή.
3. Έπειτα υπολογίζονται εκ νέου οι αποστάσεις των σημείων $\mathbf{v}_i$, $i = 1, \ldots, n$ από τα κέντρα βάρους $\mathbf{C}_j$, $j = 1, \ldots, k$ και το καθένα τοποθετείται στην ομάδα με το πλησιέστερο κέντρο βάρους.
4. Επαναλαμβάνουμε τα βήματα 2 και 3 μέχρις ότου να μην υπάρχει καμία αλλαγή στις τοποθετήσεις των σημείων $\mathbf{v}_i$, $i = 1, \ldots, n$, ή μέχρις ότου τα κέντρα βάρους C_j, $j = 1, \ldots, k$ να αλλάζουν πολύ λίγο από την προηγούμενη επανάληψη.

Για παράδειγμα, έστω ότι επιλέγουμε ως προσωρινά κέντρα των 3 ζητούμενων ομάδων σημείων τα σημεία

$$\mathbf{C}_1 = (1, 1), \mathbf{C}_2 = (5, 5), \mathbf{C}_3 = (8, 8)$$

και έστω ότι επιλέγουμε ως μετρική (απόσταση) στο $\mathbb{R}^2$ την συνάρτηση $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ με

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

για κάθε $\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$.

Τότε οι αποστάσεις κάθε σημείου $\mathbf{v}_i$ από κάθε ένα από τα κέντρα $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$ είναι

$\mathbf{v}_i$	$d(\mathbf{v}_i, \mathbf{C}_1)$	$d(\mathbf{v}_i, \mathbf{C}_2)$	$d(\mathbf{v}_i, \mathbf{C}_3)$	Ομάδα
$\mathbf{v}_1$	1.2083	4.54533	8.77838	1
$\mathbf{v}_2$	2.17083	4.13673	8.27118	1
$\mathbf{v}_3$	3.1257	3.7108	7.66616	1
$\mathbf{v}_4$	1.62788	4.03113	8.27345	1
$\mathbf{v}_5$	4.272	2.80179	6.4846	2
$\mathbf{v}_6$	7.41839	3.61006	4.85103	2
$\mathbf{v}_7$	8.48999	4.56946	4.94773	2
$\mathbf{v}_8$	9.53166	5.51838	5.21081	3
$\mathbf{v}_9$	10.4533	6.33028	5.49295	3
$\mathbf{v}_{10}$	11.2969	7.34983	6.38905	3
$\mathbf{v}_{11}$	12.4338	8.35464	6.94262	3
$\mathbf{v}_{12}$	9.67678	5.29528	4.56508	3
$\mathbf{v}_{13}$	10.565	6.16604	4.98197	3
$\mathbf{v}_{14}$	11.9236	7.54801	5.9222	3
$\mathbf{v}_{15}$	13.	8.63713	6.76757	3
$\mathbf{v}_{16}$	9.61769	4.70106	3.36155	3
$\mathbf{v}_{17}$	10.9772	6.10737	4.11096	3
$\mathbf{v}_{18}$	5.11566	3.76431	6.86804	2
$\mathbf{v}_{19}$	5.55428	3.04138	5.9203	2
$\mathbf{v}_{20}$	5.90593	1.69706	4.56946	2
$\mathbf{v}_{21}$	6.1131	4.16773	6.66108	2
$\mathbf{v}_{22}$	6.37887	3.33017	5.55788	2
$\mathbf{v}_{23}$	6.56601	2.74089	4.77624	2
$\mathbf{v}_{24}$	7.40405	3.63593	4.90102	2
$\mathbf{v}_{25}$	8.18596	3.43657	3.52278	2

Υπολογίζουμε ξανά τα κέντρα βαρών των 3 ομάδων με βάση την τοποθέτηση που κάναμε:

$$\mathbf{C}_1 = \frac{1}{4}\left(\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 + \mathbf{v}_4\right) = (1.575, 2.8875)$$

$$\begin{aligned}\mathbf{C}_2 &= \frac{1}{11}\left(\mathbf{v}_5 + \mathbf{v}_6 + \mathbf{v}_7 + \mathbf{v}_{18} + \mathbf{v}_{19} + \mathbf{v}_{20} + \mathbf{v}_{21} + \mathbf{v}_{22} + \mathbf{v}_{23} + \mathbf{v}_{24} + \mathbf{v}_{25}\right)\\ &= (5.92727, 4.06364)\end{aligned}$$

$$\begin{aligned}\mathbf{C}_3 &= \frac{1}{10}\left(\mathbf{v}_8 + \mathbf{v}_9 + \mathbf{v}_{10} + \mathbf{v}_{11} + \mathbf{v}_{12} + \mathbf{v}_{13} + \mathbf{v}_{14} + \mathbf{v}_{15} + \mathbf{v}_{16} + \mathbf{v}_{17}\right)\\ &= (4.015, 11.49)\end{aligned}$$

Υπολογίζουμε ξανά τις αποστάσεις κάθε σημείου $\mathbf{v}_i$ από κάθε ένα από τα (νέα) κέντρα $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$:

$\mathbf{v}_i$	$d(\mathbf{v}_i, \mathbf{C}_1)$	$d(\mathbf{v}_i, \mathbf{C}_2)$	$d(\mathbf{v}_i, \mathbf{C}_3)$	Ομάδα
$\mathbf{v}_1$	0.791063	4.8432	9.72097	1
$\mathbf{v}_2$	0.380173	4.71661	8.77079	1
$\mathbf{v}_3$	1.22506	4.52742	7.83903	1
$\mathbf{v}_4$	0.865033	4.25689	9.48532	1
$\mathbf{v}_5$	2.29908	3.86867	6.64276	1
$\mathbf{v}_6$	5.44525	4.89953	3.49862	3
$\mathbf{v}_7$	6.5183	5.81551	2.4307	3
$\mathbf{v}_8$	7.56188	6.72113	1.43137	3
$\mathbf{v}_9$	8.48378	7.49054	0.61711	3
$\mathbf{v}_{10}$	9.33787	8.51919	1.09969	3
$\mathbf{v}_{11}$	10.4727	9.47946	1.81723	3
$\mathbf{v}_{12}$	7.70411	6.4319	1.29009	3
$\mathbf{v}_{13}$	8.5919	7.26975	0.399155	3
$\mathbf{v}_{14}$	9.95142	8.62153	1.01898	3
$\mathbf{v}_{15}$	11.0294	9.69152	2.11809	3
$\mathbf{v}_{16}$	7.67045	5.69675	2.09316	3
$\mathbf{v}_{17}$	9.0178	7.06426	1.34288	3
$\mathbf{v}_{18}$	4.76322	2.66923	10.3032	2
$\mathbf{v}_{19}$	4.86064	1.8259	9.49444	2
$\mathbf{v}_{20}$	4.71416	0.379325	7.99439	2
$\mathbf{v}_{21}$	5.72174	2.91037	10.5511	2
$\mathbf{v}_{22}$	5.63833	2.01614	9.53752	2
$\mathbf{v}_{23}$	5.59024	1.4416	8.76951	2
$\mathbf{v}_{24}$	6.52846	2.37684	9.33163	2
$\mathbf{v}_{25}$	7.0129	2.51094	8.25157	2

Παρατηρήστε ότι στην επανάληψη αυτή έγιναν ορισμένες αλλαγές: Το σημείο $\mathbf{v}_5$ από την ομάδα 2 τοποθετήθηκε τώρα στην ομάδα 1, τα σημεία $\mathbf{v}_6$, $\mathbf{v}_7$ από την ομάδα 2 μετακινήθηκαν στην ομάδα 3. Υπολογίζουμε ξανά τα κέντρα βαρών των 3 ομάδων με βάση την (νέα) τοποθέτηση που κάναμε:

$$\mathbf{C}_1 = \frac{1}{5}\left(\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 + \mathbf{v}_4 + \mathbf{v}_5\right) = (1.7, 3.33)$$

$$\begin{aligned}\mathbf{C}_2 &= \frac{1}{8}\left(\mathbf{v}_{18} + \mathbf{v}_{19} + \mathbf{v}_{20} + \mathbf{v}_{21} + \mathbf{v}_{22} + \mathbf{v}_{23} + \mathbf{v}_{24} + \mathbf{v}_{25}\right) \\ &= (7.08125, 2.7875)\end{aligned}$$

$$\begin{aligned}\mathbf{C}_3 &= \frac{1}{12}\left(\mathbf{v}_6 + \mathbf{v}_7 + \mathbf{v}_8 + \mathbf{v}_9 + \mathbf{v}_{10} + \mathbf{v}_{11} + \mathbf{v}_{12} + \mathbf{v}_{13} + \mathbf{v}_{14} + \mathbf{v}_{15} + \mathbf{v}_{16} + \mathbf{v}_{17}\right) \\ &= (3.875, 11.0167)\end{aligned}$$

Υπολογίζουμε ξανά τις αποστάσεις κάθε σημείου $\mathbf{v}_i$ από κάθε ένα από τα (νέα) κέντρα $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$:

$\mathbf{v}_i$	$d(\mathbf{v}_i, \mathbf{C}_1)$	$d(\mathbf{v}_i, \mathbf{C}_2)$	$d(\mathbf{v}_i, \mathbf{C}_3)$	Ομάδα
$\mathbf{v}_1$	1.24615	5.62343	9.22758	1
$\mathbf{v}_2$	0.438634	5.7926	8.27741	1
$\mathbf{v}_3$	0.826378	5.83089	7.34618	1
$\mathbf{v}_4$	1.19871	5.01578	8.9936	1
$\mathbf{v}_5$	1.83927	5.40132	6.14922	1
$\mathbf{v}_6$	4.98552	6.60889	3.00546	3
$\mathbf{v}_7$	6.05862	7.49562	1.93805	3
$\mathbf{v}_8$	7.10256	8.36581	0.95836	3
$\mathbf{v}_9$	8.02449	9.09707	0.397987	3
$\mathbf{v}_{10}$	8.88104	10.1278	1.33198	3
$\mathbf{v}_{11}$	10.0153	11.0441	2.23437	3
$\mathbf{v}_{12}$	7.24478	8.02741	0.826211	3
$\mathbf{v}_{13}$	8.13221	8.83094	0.239925	3
$\mathbf{v}_{14}$	9.49165	10.1452	1.50858	3
$\mathbf{v}_{15}$	10.5699	11.1898	2.60366	3
$\mathbf{v}_{16}$	7.22059	7.19083	1.79842	3
$\mathbf{v}_{17}$	8.56346	8.50121	1.42743	3
$\mathbf{v}_{18}$	4.80467	1.69941	9.87074	2
$\mathbf{v}_{19}$	4.81154	0.83771	9.07505	2
$\mathbf{v}_{20}$	4.52448	1.3423	7.58198	2
$\mathbf{v}_{21}$	5.73454	1.38763	10.1431	2
$\mathbf{v}_{22}$	5.56227	0.311059	9.14275	2
$\mathbf{v}_{23}$	5.45008	0.517091	8.3829	2
$\mathbf{v}_{24}$	6.40413	1.0656	8.97356	2
$\mathbf{v}_{25}$	6.80139	2.16142	7.93366	2

Παρατηρήστε ότι στην επανάληψη αυτή δεν έγινε καμία αλλαγή στις ομάδες. Επομένως, ο αλγόριθμος ολοκληρώθηκε. Τα ζητούμενα κέντρα είναι τα σημεία

$$\mathbf{C}_1 = (1.7, 3.33)\,, \mathbf{C}_2 = (7.08125, 2.7875)\,, \mathbf{C}_3 = (3.875, 11.0167)$$

και οι 3 ομάδες είναι τα παρακάτω σύνολα:

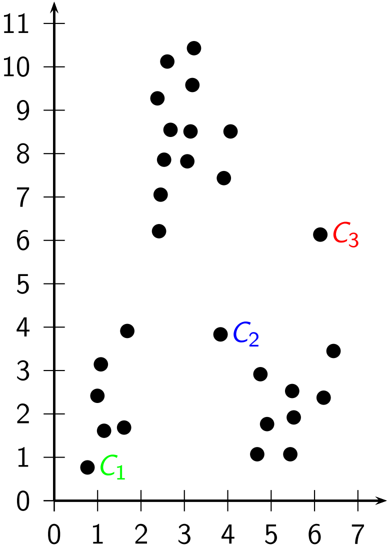
$$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\}$$

$$\{\mathbf{v}_8, \mathbf{v}_9, \mathbf{v}_{10}, \mathbf{v}_{11}, \mathbf{v}_{12}, \mathbf{v}_{13}, \mathbf{v}_{14}, \mathbf{v}_{15}, \mathbf{v}_{16}, \mathbf{v}_{17}\}$$

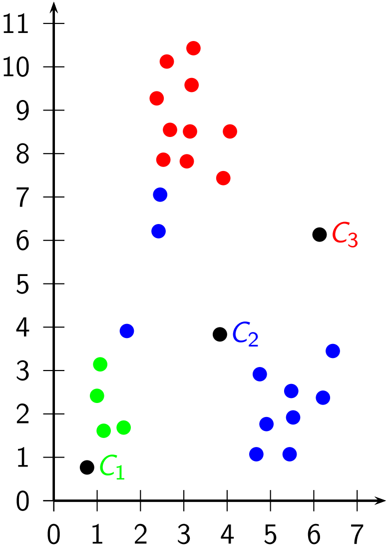
$$\{\mathbf{v}_{18}, \mathbf{v}_{19}, \mathbf{v}_{20}, \mathbf{v}_{21}, \mathbf{v}_{22}, \mathbf{v}_{23}, \mathbf{v}_{24}, \mathbf{v}_{25}\}$$

Επειδή τα σημεία $\mathbf{v}_i$ ανήκουν στο επίπεδο $\mathbb{R}^2$ μπορούμε να δούμε και οπτικά την διαδικασία εκτέλεσης του αλγορίθμου K-means

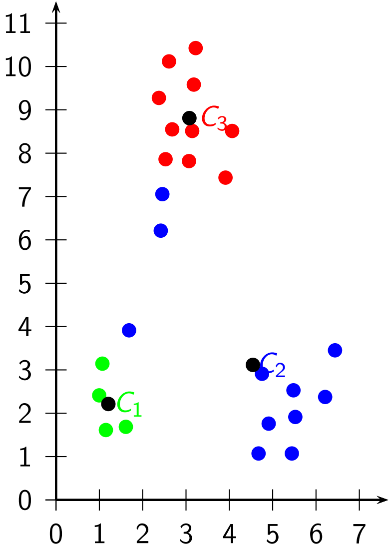
Αλγόριθμος K-Means στον $\mathbb{R}^2$ για 3 ομάδες



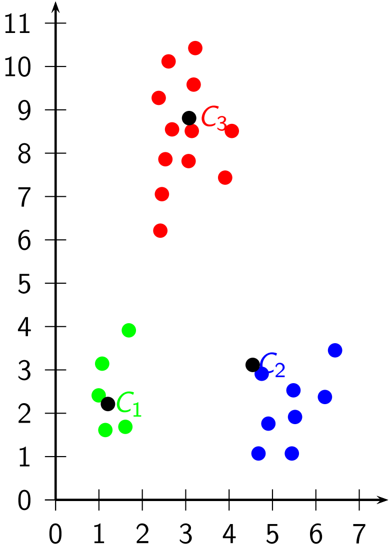
1. Επιλογή των $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$



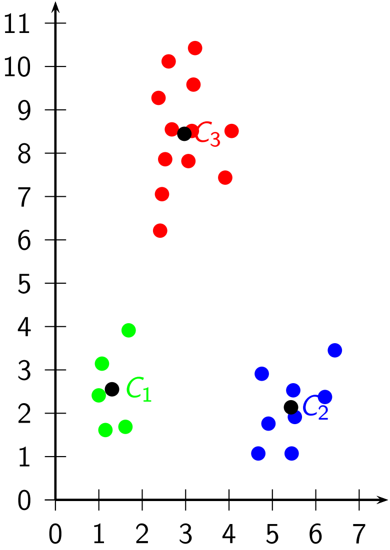
2. Δημιουργία ομάδων



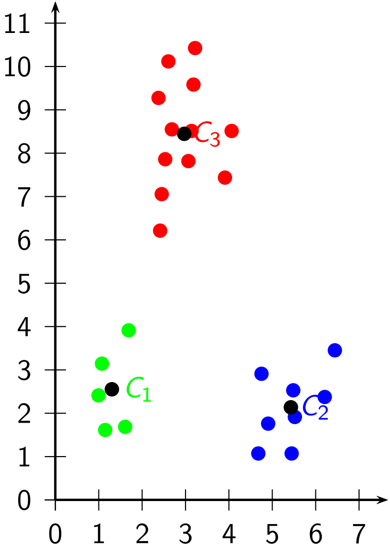
3. Νέος υπολογισμός των $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$



4. Δημιουργία νέων ομάδων



5. Νέος υπολογισμός των $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$



6. Δημιουργία τελικών ομάδων

Παρατηρήσεις:

- Ο αλγόριθμος K-means διαθέτει πολλές παραλλαγές.
- Η απόδοσή του εξαρτάται από την μορφή των δεδομένων, από την μετρική που επιλέγουμε και από την αρχική επιλογή των k κέντρων βαρών. Συνήθως, γίνονται πολλές εκτελέσεις του αλγορίθμου με διαφορετικές αρχικές επιλογές κέντρων βαρών και επιλέγεται η διαμέριση που έχει το μικρότερο συνολικό άθροισμα αποστάσεων κάθε σημείου από το κέντρο βάρους της ομάδας στην οποία ανήκει.
- Το οπτικό πλεονέκτημα που είχαμε στο παράδειγμα μας δεν υφίσταται σε χώρους με διάσταση 4 ή παραπάνω.

3η ΔΙΑΛΕΞΗ
ΕΣΩΤΕΡΙΚΟ ΓΙΝΟΜΕΝΟ

- **Ορθογωνιότητα**
- **Πυθαγόρειο θεώρημα**
- **Ορθοκανονικές βάσεις**
- **Αλγόριθμος Gram - Schmidt**

Ορθογωνιότητα

Έστω V διανυσματικός χώρος με εσωτερικό γινόμενο $\langle \rangle$ και $\mathbf{v}, \mathbf{u} \in V$.
Το $\mathbf{v}$ ονομάζεται **ορθογώνιο** (ή **κάθετο**) στο $\mathbf{u}$ αν και μόνο αν

$$\langle \mathbf{v}, \mathbf{u} \rangle = 0$$

(άρα και $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, δηλαδή και το $\mathbf{u}$ είναι ορθογώνιο στο $\mathbf{v}$).
Γι' αυτό λέμε ότι τα $\mathbf{v}, \mathbf{u}$ είναι **ορθογώνια** ή **κάθετα**, και γράφουμε $\mathbf{v} \perp \mathbf{u}$.

Παράδειγμα: Τα διανύσματα $\mathbf{v} = (1, -1, 1)$, $\mathbf{u} = (-2, -1, 1)$ του $\mathbb{R}^3$, με το σύνηθες εσωτερικό γινόμενο, είναι ορθογώνια. Πράγματι, ισχύει ότι $\langle \mathbf{v}, \mathbf{u} \rangle = \langle (1, -1, 1), (-2, -1, 1) \rangle = 1 \cdot (-2) + (-1) \cdot (-1) + 1 \cdot 1 = 0$.

Παρατηρήσεις:

- Προφανώς, αν $\mathbf{u}, \mathbf{v}$ είναι ορθογώνια τότε για κάθε $a, b \in \mathbb{R}$ και τα $a\mathbf{v}$ και $b\mathbf{u}$ είναι επίσης ορθογώνια.
- $\cos \theta = \dfrac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{v}\| \, \|\mathbf{u}\|} = 0 \Leftrightarrow \theta = \dfrac{\pi}{2}$.

Πρόταση 8 (Πυθαγόρειο Θεώρημα)

Έστω $(\mathcal{V}, \langle\rangle)$. Αν τα **u**, **v** *είναι ορθογώνια, τότε*

$$\|\mathbf{v}+\mathbf{u}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 .$$

Απόδειξη.

Ισχύει ότι

$$\begin{aligned} \|\mathbf{v}+\mathbf{u}\|^2 &= \langle \mathbf{v}+\mathbf{u}, \mathbf{v}+\mathbf{u}\rangle = \langle \mathbf{v}, \mathbf{v}+\mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{v}+\mathbf{u}\rangle \\ &= \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{u}\rangle \\ &\overset{\langle \mathbf{v},\mathbf{u}\rangle=0}{=} \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{u}\rangle = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 . \end{aligned}$$

□

Παρατήρηση: Γενικότερα, ισχύει η σχέση

$$\|\mathbf{v}+\mathbf{u}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 + 2\|\mathbf{v}\|\,\|\mathbf{u}\|\cos\theta$$

Έστω μη κενό σύνολο $A \subseteq V$.
Το σύνολο A ονομάζεται **ορθογώνιο** αν και μόνο αν για κάθε $\mathbf{v}, \mathbf{u} \in A$ ισχύει ότι

$$\mathbf{v} \neq \mathbf{u} \Rightarrow \langle \mathbf{v}, \mathbf{u} \rangle = 0.$$

(Ένα μονοσύνολο θεωρείται πάντα ορθογώνιο.)

Παράδειγμα:
Το σύνολο $\{(2,0,0),(0,3,0),(0,0,1)\}$ είναι ορθογώνιο.

Ορθογωνιότητα

Παρατήρηση: Η ορθογωνιότητα συνεπάγεται την γραμμική ανεξαρτησία.

Πρόταση 9

Έστω $(\mathcal{V}, \langle\rangle)$ *και* $\emptyset \neq A \subseteq V \setminus \{\mathbf{0}\}$. *Αν το σύνολο* $A = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *είναι ορθογώνιο, τότε τα διανύσματα* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *είναι γραμμικώς ανεξάρτητα.*

Απόδειξη. Θεωρούμε την εξίσωση

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_k \mathbf{v}_k + \cdots + \lambda_n \mathbf{v}_n = \mathbf{0}$$

Θα δείξουμε ότι $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.
Για κάθε $k \in [n]$ ισχύει ότι

$$\begin{aligned}
&\langle \mathbf{v}_k, \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_k \mathbf{v}_k + \cdots + \lambda_n \mathbf{v}_n \rangle = \langle \mathbf{v}_k, \mathbf{0} \rangle \\
\Leftrightarrow &\langle \mathbf{v}_k, \lambda_1 \mathbf{v}_1 \rangle + \langle \mathbf{v}_k, \lambda_2 \mathbf{v}_2 \rangle + \cdots + \langle \mathbf{v}_k, \lambda_k \mathbf{v}_k \rangle + \cdots + \langle \mathbf{v}_k, \lambda_n \mathbf{v}_n \rangle = 0 \\
\Leftrightarrow &\lambda_1 \langle \mathbf{v}_k, \mathbf{v}_1 \rangle + \lambda_2 \langle \mathbf{v}_k, \mathbf{v}_2 \rangle + \cdots + \lambda_k \langle \mathbf{v}_k, \mathbf{v}_k \rangle + \cdots + \lambda_n \langle \mathbf{v}_k, \mathbf{v}_n \rangle = 0 \\
\Rightarrow &\lambda_k \langle \mathbf{v}_k, \mathbf{v}_k \rangle = 0 \text{ (διότι } \langle \mathbf{v}_k, \mathbf{v}_j \rangle = 0 \text{, όταν } k \neq j) \\
\Leftrightarrow &\lambda_k \|\mathbf{v}_k\|^2 = 0 \\
\overset{\mathbf{v}_k \neq \mathbf{0}}{\Rightarrow} &\lambda_k = 0
\end{aligned}$$

Επομένως, $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.
Άρα, τα $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ είναι γραμμικώς ανεξάρτητα. □

Παρατηρήσεις:

- Δεν ισχύει το αντίστροφο της προηγούμενης πρότασης.
- Από την προηγούμενη πρόταση προκύπτει ότι σε χώρους πεπερασμένης διάστασης υπάρχει περιορισμός στον αριθμό των ορθογώνιων διανυσμάτων.

Πόρισμα 1

Έστω $(\mathcal{V}, \langle\rangle)$ *με* $\dim V = n$. *Για κάθε ορθογώνιο σύνολο* $A \subseteq V$ *ισχύει ότι* $|A| \leq n$.

Απόδειξη.

Έστω ότι το A περιέχει k διανύσματα. Από την προηγούμενη πρόταση έπεται ότι τα διανύσματα του A είναι γραμμικώς ανεξάρτητα και επειδή σε κάθε χώρο διάστασης n ο μέγιστος αριθμός γραμμικώς ανεξάρτητων διανυσμάτων είναι n, έπεται ότι $k \leq n$. $\square$

Ορθογωνιότητα

Κάθε διάνυσμα $\mathbf{v} \in V$ με $\|\mathbf{v}\| = 1$ ονομάζεται **μοναδιαίο** ή **κανονικό** (normal) διάνυσμα.

Αν $\mathbf{v}$ είναι οποιοδήποτε μη μηδενικό διάνυσμα τότε το διάνυσμα $\frac{1}{\|\mathbf{v}\|}\mathbf{v}$ είναι μοναδιαίο.

Αν $\mathbf{v} \neq 0$ και $\|\mathbf{v}\| \neq 1$, τότε η εύρεση του διανύσματος $\frac{1}{\|\mathbf{v}\|}\mathbf{v}$ ονομάζεται **κανονικοποίηση του $\mathbf{v}$**.

Παράδειγμα: Αν $\mathbf{v} = (2,4,4)$ τότε $\|\mathbf{v}\| = \sqrt{2^2+4^2+4^2} = \sqrt{36} = 6$.
Για να κανονικοποιήσουμε το $\mathbf{v}$ πολλαπλασιάζουμε με $\frac{1}{6}$ οπότε προκύπτει το διάνυσμα $\frac{1}{6}(2,4,4) = (\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$ για το οποίο

$$\left\|(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})\right\| = \sqrt{\frac{1}{9}+\frac{4}{9}+\frac{4}{9}} = \sqrt{\frac{9}{9}} = 1$$

Ορθογωνιότητα

Ένα μη κενό σύνολο διανυσμάτων A ονομάζεται **ορθοκανονικό** αν και μόνο αν

- A ορθογώνιο
- Τα διανύσματα του A είναι μοναδιαία.

Παράδειγμα: Το σύνολο $\{(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0), (\frac{1}{\sqrt{3}}, 0, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}})\}$ είναι ορθοκανονικό.

Παράδειγμα: Στο χώρο των ολοκληρώσιμων συναρτήσεων στο $[-1, 1]$ το σύνολο των συναρτήσεων

$$\frac{1}{\sqrt{2}}, \cos \pi x, \cos 2\pi x, \dots, \cos n\pi x, \dots, \sin \pi x, \sin 2\pi x, \dots, \sin m\pi x, \dots,$$

είναι ορθοκανονικό.
(Εδώ ορίζουμε $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$.)

Έστω V διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\rangle$.
Μια βάση B του V ονομάζεται **ορθοκανονική** αν και μόνο αν το σύνολο B είναι ορθοκανονικό δηλαδή

- Κάθε διάνυσμά του είναι μοναδιαίο
- Κάθε ζεύγος διανυσμάτων της βάσης είναι ορθογώνια.

Παράδειγμα: Στον $\mathbb{R}^3$ τα διανύσματα

$$\mathbf{e}_1 = (1, 0, 0), \quad \mathbf{e}_2 = (0, 1, 0), \quad \mathbf{e}_3 = (0, 0, 1)$$

είναι μια ορθοκανονική βάση του $\mathbb{R}^3$.
Παράδειγμα: Στον $\mathbb{R}^4$ τα διανύσματα

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1, 1, 0, 0) \qquad \mathbf{v}_2 = \frac{1}{\sqrt{2}}(1, -1, 0, 0)$$

$$\mathbf{v}_3 = \frac{1}{\sqrt{2}}(0, 0, 1, 1) \qquad \mathbf{v}_2 = \frac{1}{\sqrt{2}}(0, 0, 1, -1)$$

είναι μια ορθοκανονική βάση του $\mathbb{R}^4$ (βάση Haar).

Ορθοκανονικές βάσεις

Παράδειγμα: Η *κανονική* βάση $\{e_1, e_2, \ldots, e_n\}$ του $(\mathbb{R}^n, \langle\rangle)$, όπου $\langle\rangle$ είναι το σύνηθες εσωτερικό γινόμενο είναι ορθοκανονική.
(Υπενθύμιση: $\mathbf{e}_1 = (1, 0, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \ldots, 0)$, $\ldots$,
$\mathbf{e}_n = (0, 0, 0, \ldots, 1)$ και $\langle(x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle = x_1y_1 + \cdots + x_ny_n$.)
Πράγματι, για κάθε $i \in [n]$ ισχύει ότι

$$\|e_i\| = \sqrt{\langle \mathbf{e}_i, \mathbf{e}_i \rangle} = \sqrt{\left\langle (\underbrace{0, 0, \ldots, 1}_{i \text{ θέσεις}}, \ldots, 0), (\underbrace{0, 0, \ldots, 1}_{i \text{ ϑέσεις}}, \ldots, 0) \right\rangle}$$

$$= \sqrt{0 + 0 + \cdots + 1 + \cdots + 0} = \sqrt{1} = 1,$$

ενώ για $i \neq j$ (με $i < j$)

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \left\langle (\overbrace{\underbrace{0, \ldots, 1}_{i \text{ ϑέσεις}}, \ldots, 0}^{j \text{ θέσεις}}, \ldots, 0), (\overbrace{\underbrace{0, \ldots, 0}_{i \text{ θέσεις}}, \ldots, 1}^{j \text{ θέσεις}}, \ldots, 0) \right\rangle$$

$$= 0{\cdot}0 + 0{\cdot}0 + \cdots 1 \cdot 0 + \cdots + 0{\cdot}1 + \cdots + 0{\cdot}0 = 0.$$

Επίσης, $\langle \mathbf{e}_i, \mathbf{e}_i \rangle = 1$ και άρα

$$\|\mathbf{e}_i\| = \sqrt{\langle \mathbf{e}_i, \mathbf{e}_i \rangle} = \sqrt{1} = 1$$

(αφού $\langle \mathbf{e}_i, \mathbf{e}_i \rangle = \delta_{1,i}^2 + \delta_{2,i}^2 + \cdots + \delta_{n,i}^2$ και μόνο το $\delta_{i,i}^2 = 1$ ενώ όλα τα άλλα είναι ίσα με 0).
Παρατήρηση: Έστω $(\mathcal{V}, \langle\rangle)$ και $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subseteq V$. Το σύνολο $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ είναι ορθοκανονικό αν και μόνο αν για κάθε i, j είναι

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{i,j}.$$

όπου δ_{ij} είναι το σύμβολο δέλτα του Kronecker

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

Δηλαδή, είναι $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$, ενώ $i \neq j \Rightarrow \langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$.

Παρατήρηση: Είναι πάρα πολύ εύκολο να βρούμε την αναπαράσταση ενός διανύσματος σε μια ορθοκανονική βάση.

Πρόταση 10

Αν $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_n$ *είναι μια ορθοκανονική βάση του διανυσματικού χώρου* V, *τότε κάθε διάνυσμα* $\mathbf{v} \in V$ *γράφεται κατά μοναδικό τρόπο στην μορφή*

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{v}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{v}, \mathbf{v}_n \rangle \mathbf{v}_n$$

Οι συντελεστές $\lambda_i = \langle \mathbf{v}, \mathbf{v}_i \rangle$ *ονομάζονται* ***συντελεστές* Fourier *του*** $\mathbf{v}$ ***ως προς την ορθοκανονική βάση*** $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_k$.

Ορθοκανονικές βάσεις

Απόδειξη.

Αφού $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_n$ βάση του V υπάρχουν μοναδικά λ_1, λ_2, …, λ_n ώστε

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n$$

Για κάθε $i \in [n]$ θεωρούμε τα εσωτερικά γινόμενα

$$\begin{aligned}
\langle \mathbf{v}, \mathbf{v}_i \rangle &= \langle \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_i \mathbf{v}_i + \cdots + \lambda_n \mathbf{v}_n, \mathbf{v}_i \rangle \\
&= \lambda_1 \langle \mathbf{v}_1, \mathbf{v}_i \rangle + \lambda_2 \langle \mathbf{v}_2, \mathbf{v}_i \rangle + \cdots + \lambda_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle + \cdots + \lambda_n \langle \mathbf{v}_n, \mathbf{v}_i \rangle \\
&= \lambda_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle \ \ (\text{διότι } \langle \mathbf{v}_j, \mathbf{v}_i \rangle = 0 \text{ αν } j \neq i) \\
&= \lambda_i \|\mathbf{v}_i\|^2 = \lambda_i \ (\text{διότι } \|\mathbf{v}_i\| = 1)
\end{aligned}$$

Άρα,

$$\mathbf{v} = \underbrace{\langle \mathbf{v}, \mathbf{v}_1 \rangle}_{\lambda_1} \mathbf{v}_1 + \underbrace{\langle \mathbf{v}, \mathbf{v}_2 \rangle}_{\lambda_2} \mathbf{v}_2 + \cdots + \underbrace{\langle \mathbf{v}, \mathbf{v}_n \rangle}_{\lambda_n} \mathbf{v}_n$$

□

Παράδειγμα: Στο διανυσματικό χώρο $(\mathbb{R}^3, +, \cdot)$, το σύνολο $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, όπου $\mathbf{x}_1 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$, $\mathbf{x}_2 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$, $\mathbf{x}_3 = \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$ είναι ορθοκανονικό (Άσκηση). Άρα, αποτελεί μια ορθοκανονική βάση του $(\mathbb{R}^3, +, \cdot)$. Το διάνυσμα $\mathbf{x} = (1, -1, 1) \in \mathbb{R}^3$ γράφεται λοιπόν, ως προς την παραπάνω βάση, ως εξής:

$$\mathbf{x} = \langle \mathbf{x}, \mathbf{x}_1 \rangle \, \mathbf{x}_1 + \langle \mathbf{x}, \mathbf{x}_2 \rangle \, \mathbf{x}_2 + \langle \mathbf{x}, \mathbf{x}_3 \rangle \, \mathbf{x}_3,$$

όπου

$$\langle \mathbf{x}, \mathbf{x}_1 \rangle = \left\langle (1, -1, 1), \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right) \right\rangle = 0,$$
$$\langle \mathbf{x}, \mathbf{x}_2 \rangle = \left\langle (1, -1, 1), \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \right\rangle = \frac{1}{\sqrt{3}},$$
$$\langle \mathbf{x}, \mathbf{x}_3 \rangle = \left\langle (1, -1, 1), \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) \right\rangle = \frac{4}{\sqrt{6}},$$

δηλαδή

$$(1, -1, 1) = 0 \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right) + \frac{1}{\sqrt{3}} \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) + \frac{4}{\sqrt{6}} \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right).$$

Ορθοκανονικές βάσεις

Παράδειγμα: Η αναπαράσταση του διανύσματος $\mathbf{v} = (1, -1, 2, 3) \in \mathbb{R}^4$ στην ορθοκανονική βάση Haar του παραδείγματος είναι

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{v}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \langle \mathbf{v}, \mathbf{v}_3 \rangle \mathbf{v}_3 + \langle \mathbf{v}, \mathbf{v}_4 \rangle \mathbf{v}_4$$

όπου

$$\langle \mathbf{v}, \mathbf{v}_1 \rangle = \left\langle (1, -1, 2, 3), \frac{1}{\sqrt{2}}(1, 1, 0, 0) \right\rangle = \frac{0}{\sqrt{2}} = 0$$

$$\langle \mathbf{v}, \mathbf{v}_2 \rangle = \left\langle (1, -1, 2, 3), \frac{1}{\sqrt{2}}(1, -1, 0, 0) \right\rangle = \frac{2}{\sqrt{2}}$$

$$\langle \mathbf{v}, \mathbf{v}_3 \rangle = \left\langle (1, -1, 2, 3), \frac{1}{\sqrt{2}}(0, 0, 1, 1) \right\rangle = \frac{5}{\sqrt{2}}$$

$$\langle \mathbf{v}, \mathbf{v}_4 \rangle = \left\langle (1, -1, 2, 3), \frac{1}{\sqrt{2}}(0, 0, 1, -1) \right\rangle = \frac{-1}{\sqrt{2}}$$

Άρα,

$$\mathbf{v} = 0 \cdot \mathbf{v}_1 + \frac{2}{\sqrt{2}}\mathbf{v}_2 + \frac{5}{\sqrt{2}}\mathbf{v}_3 + \frac{(-1)}{\sqrt{2}}\mathbf{v}_4$$

Ορθοκανονικές βάσεις

Ομοίως αποδεικνύεται και η παρακάτω γενικότερη πρόταση.

Πρόταση 11

Αν $A = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *είναι ένα ορθογώνιο σύνολο διανυσμάτων και*

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n$$

Τότε ισχύει ότι

$$\lambda_i = \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \text{ για κάθε } i \in [n].$$

Οι συντελεστές λ_i *ονομάζονται (και πάλι)* ***συντελεστές* Fourier *του* v *ως προς τα ορθογώνια διανύσματα*** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

Αλγόριθμος Gram - Schmidt

Δεδομένου ότι σε μια ορθοακανονική βάση το πρόβλημα της αναπαράστασης διανυσμάτων είναι υπολογιστικά πολύ απλό μας ενδιαφέρει το επόμενο ερώτημα.

Έστω ότι για κάποιο διανυσματικό χώρο V γνωρίζουμε μια βάση B η οποία δεν είναι ορθοκανονική.
Μπορούμε να βρούμε μια άλλη βάση B' η οποία είναι ορθοκανονική;

Η απάντηση είναι καταφατική.
Ένας τρόπος να γίνει αυτό είναι να χρησιμοποιήσουμε τον αλγόριθμο των Gram - Schmidt.
Συγκεκριμένα, Θα δείξουμε τώρα ότι κάθε πεπερασμένης διάστασης, μη μηδενικός, πραγματικός ή μιγαδικός διανυσματικός χώρος με εσωτερικό γινόμενο, έχει μια ορθοκανονική βάση.

Ορθοκανονικές βάσεις

Ο αλγόριθμος Gram - Schmidt βασίζεται στην επόμενη ιδιότητα:

Πρόταση 12

Αν $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_k$ *είναι ορθοκανονικό σύνολο διανυσμάτων, τότε για κάθε* $\mathbf{v} \in V$ *το διάνυσμα*

$$\mathbf{y} = \mathbf{v} - \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 - \langle \mathbf{v}, \mathbf{v}_2 \rangle \mathbf{v}_2 - \cdots - \langle \mathbf{v}, \mathbf{v}_k \rangle \mathbf{v}_k$$

είναι ορθογώνιο σε κάθε ένα από τα $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_k$.

Ειδικότερα,

- αν $\mathbf{v} \in \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ τότε $\mathbf{y} = \mathbf{0}$.
- αν $\mathbf{v} \notin \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ τότε $\mathbf{y} \neq \mathbf{0}$
  και το $\mathbf{y}$ είναι γραμμικώς ανεξάρτητο από τα $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_k$, δηλαδή το σύνολο $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{y}$ είναι ορθογώνιο, οπότε
  το σύνολο $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}$ είναι ορθοκανονικό.

Γενικότερα, αν τα $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ είναι ορθογώνια ισχύει η επόμενη πρόταση.

Πρόταση 13

Αν $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_k$ είναι ορθογώνιο σύνολο διανυσμάτων, τότε για κάθε $\mathbf{v} \in V$ το διάνυσμα

$$\mathbf{y} = \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{v}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \cdots - \frac{\langle \mathbf{v}, \mathbf{v}_k \rangle}{\|\mathbf{v}_k\|^2}\mathbf{v}_k$$

είναι ορθογώνιο σε κάθε ένα από τα $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_k$.

Ορθοκανονικές βάσεις

Απόδειξη.

Πράγματι, για κάθε $k \in [n]$ ισχύει ότι

$$\begin{aligned}
&\langle \mathbf{y}, \mathbf{v}_k \rangle \\
&= \left\langle \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{v}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \cdots - \frac{\langle \mathbf{v}, \mathbf{v}_k \rangle}{\|\mathbf{v}_k\|^2}\mathbf{v}_k - \cdots - \frac{\langle \mathbf{v}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n, \mathbf{v}_k \right\rangle \\
&= \langle \mathbf{v}, \mathbf{v}_k \rangle - \frac{\langle \mathbf{v}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \langle \mathbf{v}_1, \mathbf{v}_k \rangle - \frac{\langle \mathbf{v}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \langle \mathbf{v}_2, \mathbf{v}_k \rangle - \cdots - \frac{\langle \mathbf{v}, \mathbf{v}_k \rangle}{\|\mathbf{v}_k\|^2} \langle \mathbf{v}_k, \mathbf{v}_k \rangle \\
&\qquad - \cdots - \frac{\langle \mathbf{v}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2} \langle \mathbf{v}_n, \mathbf{v}_k \rangle \\
&= \langle \mathbf{v}, \mathbf{v}_k \rangle - \frac{\langle \mathbf{v}, \mathbf{v}_k \rangle}{\|\mathbf{v}_k\|^2} \langle \mathbf{v}_k, \mathbf{v}_k \rangle \\
&= \langle \mathbf{v}, \mathbf{v}_k \rangle - \frac{\langle \mathbf{v}, \mathbf{v}_k \rangle}{\|\mathbf{v}_k\|^2} \|\mathbf{v}_k\|^2 = 0.
\end{aligned}$$

□

Πρόταση 14 (Διαδικασία ορθοκανονικοποίησης των Gram-Schmidt)

Έστω $(V, \langle\rangle)$ διανυσματικός χώρος με εσωτερικό γινόμενο και $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ ελεύθερο υποσύνολο του V. Αν

$$
\begin{aligned}
\mathbf{v}_1 &= \frac{1}{\|\mathbf{y}_1\|}\mathbf{y}_1, \ \acute{o}\pi o\upsilon\ \mathbf{y}_1 = \mathbf{x}_1, \\
\mathbf{v}_2 &= \frac{1}{\|\mathbf{y}_2\|}\mathbf{y}_2, \ \acute{o}\pi o\upsilon\ \mathbf{y}_2 = \mathbf{x}_2 - \langle \mathbf{x}_2, \mathbf{v}_1 \rangle \mathbf{v}_1, \\
\mathbf{v}_3 &= \frac{1}{\|\mathbf{y}_3\|}\mathbf{y}_3, \ \acute{o}\pi o\upsilon\ \mathbf{y}_3 = \mathbf{x}_3 - \langle \mathbf{x}_3, \mathbf{v}_2 \rangle \mathbf{v}_2 - \langle \mathbf{x}_3, \mathbf{v}_1 \rangle \mathbf{v}_1, \\
\mathbf{v}_4 &= \frac{1}{\|\mathbf{y}_4\|}\mathbf{y}_4, \ \acute{o}\pi o\upsilon\ \mathbf{y}_4 = \mathbf{x}_4 - \langle \mathbf{x}_4, \mathbf{v}_3 \rangle \mathbf{v}_3 - \langle \mathbf{x}_4, \mathbf{v}_2 \rangle \mathbf{v}_2 - \langle \mathbf{x}_4, \mathbf{v}_1 \rangle \mathbf{v}_1, \\
&\vdots \\
\mathbf{v}_n &= \frac{1}{\|\mathbf{y}_n\|}\mathbf{y}_n, \acute{o}\pi o\upsilon\ \mathbf{y}_n = \mathbf{x}_n - \langle \mathbf{x}_n, \mathbf{v}_{n-1} \rangle \mathbf{v}_{n-1} - \cdots - \langle \mathbf{x}_n, \mathbf{v}_1 \rangle \mathbf{v}_1,
\end{aligned}
$$

Απόδειξη.

Επειδή τα διανύσματα $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ είναι γραμμικώς ανεξάρτητα, έπεται ότι και τα διανύσματα $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ είναι μη μηδενικά, άρα και τα διανύσματα $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ είναι επίσης μη μηδενικά και ειδικότερα μοναδιαία. Επίσης, από την Πρόταση 13 έπεται ότι το διάνυσμα $\mathbf{y}_k$ είναι ορθογώνιο με τα διανύσματα $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1}$ και άρα και το διάνυσμα $\mathbf{v}_k$ είναι επίσης ορθογώνιο με τα διανύσματα $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1}$, για κάθε $k \in [n]$. Επομένως, $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ είναι ορθοκανονικά διανύσματα.
Τέλος, αφού ο χώρος $< \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} >$ που παράγουν τα $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ έχει διάσταση n και τα $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ ανήκουν στο χώρο αυτό και είναι γραμμικώς ανεξάρτητα, έπεται ότι τα $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ αποτελούν ορθοκανονική βάση του χώρου $< \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} >$. $\square$

Αλγόριθμος Gram - Schmidt

Πρόταση 15

Έστω $(\mathcal{V}, \langle\rangle)$*, με* $\dim \mathcal{V} = n$*. Τότε υπάρχει ορθοκανονικό υποσύνολο* B *του* V *το οποίο είναι βάση του* $\mathcal{V}$*.*

Απόδειξη.

Πράγματι, από την προηγούμενη πρόταση, έχουμε ότι αν $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ είναι μια βάση του V, τότε το σύνολο $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ που κατασκευάζεται με τη διαδικασία Gram-Schmidt θα είναι μια ορθοκανονική βάση του V. $\square$

Παράδειγμα: Δίδεται το σύνολο των γραμμικώς ανεξάρτητων διανυσμάτων

$$\mathbf{x}_1 = (1,1,0,0), \mathbf{x}_2 = (0,1,1,0), \mathbf{x}_3 = (1,0,1,1)$$

Να βρεθεί μια ορθοκανονική βάση του χώρου που παράγουν τα $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$.

Λύση.

- $\mathbf{y}_1 = \mathbf{x}_1 = (1,1,0,0)$, $\|\mathbf{y}_1\| = \sqrt{1^2+1^2+0^2+0^2} = \sqrt{2}$.
  Άρα,
  $$\mathbf{v}_1 = \frac{1}{\|\mathbf{y}_1\|}\mathbf{y}_1 = \frac{1}{\sqrt{2}}(1,1,0,0)$$

Αλγόριθμος Gram - Schmidt

- $\mathbf{y}_2 = \mathbf{x}_2 - \langle \mathbf{x}_2, \mathbf{v}_1 \rangle \, \mathbf{v}_1$
  $\langle \mathbf{x}_2, \mathbf{v}_1 \rangle = \dfrac{1}{2}(0 \cdot 1 + 1 \cdot 1 + 1 \cdot 0 + 0 \cdot 0) = \dfrac{1}{\sqrt{2}}$
  οπότε
  $\mathbf{y}_2 = (0, 1, 1, 0) - \dfrac{1}{\sqrt{2}} \dfrac{1}{\sqrt{2}}(1, 1, 0, 0) = \dfrac{1}{2}(-1, 1, 2, 0)$
  $\|\mathbf{y}_2\| = \dfrac{1}{2}\sqrt{(-1)^2 + 1^2 + 2^2 + 0^2} = \dfrac{\sqrt{6}}{2}$
  Άρα,

$$\mathbf{v}_2 = \frac{1}{\|\mathbf{y}_2\|}\mathbf{y}_2 = \frac{1}{(\sqrt{6}/2)} \frac{1}{2}(-1, 1, 2, 0) = \frac{1}{\sqrt{6}}(-1, 1, 2, 0)$$

Αλγόριθμος Gram - Schmidt

- $\mathbf{y}_3 = \mathbf{x}_3 - \langle \mathbf{x}_3, \mathbf{v}_1 \rangle \, \mathbf{v}_1 - \langle \mathbf{x}_3, \mathbf{v}_2 \rangle \, \mathbf{v}_2$
  $\langle \mathbf{x}_3, \mathbf{v}_1 \rangle = \dfrac{1}{\sqrt{2}}$
  $\langle \mathbf{x}_3, \mathbf{v}_2 \rangle = \dfrac{1}{\sqrt{6}}$
  οπότε
  $\mathbf{y}_3 = (1, 0, 1, 1) - \dfrac{1}{\sqrt{2}} \dfrac{1}{\sqrt{2}} (1, 1, 0, 0) - \dfrac{1}{\sqrt{6}} \dfrac{1}{\sqrt{6}} (-1, 1, 2, 0) =$
  $\dfrac{1}{6}(4, -4, 4, 6) = \dfrac{1}{3}(2, -2, 2, 3)$
  $\|\mathbf{y}_3\| = \dfrac{1}{3}\sqrt{2^2 + (-2)^2 + 2^2 + 3^2} = \dfrac{\sqrt{21}}{3}$
  Άρα,

$$\mathbf{v}_3 = \frac{1}{\|\mathbf{y}_3\|}\mathbf{y}_3 = \frac{1}{(\sqrt{21}/3)} \frac{1}{3}(2, -2, 2, 3) = \frac{1}{\sqrt{21}}(2, -2, 2, 3)$$

Αλγόριθμος Gram - Schmidt

Παράδειγμα: Έστω $(\mathbb{R}^3, \langle\rangle)$ όπου $\langle\rangle$ το σύνηθες εσωτερικό γινόμενο.
Το σύνολο $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, με $\mathbf{x}_1 = (1,1,0)$, $\mathbf{x}_2 = (0,1,1)$, $\mathbf{x}_3 = (1,0,1)$ είναι προφανώς μια βάση του $\mathbb{R}^3$
Η βάση αυτή **δεν** είναι ορθογώνια, αφού
$\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 1 \cdot 0 + 1 \cdot 1 + 0 \cdot 1 = 1 \neq 0$.
Η βάση αυτή **δεν** είναι κανονική, αφού
$\|\mathbf{x}_1\| = \sqrt{\langle \mathbf{x}_1, \mathbf{x}_1 \rangle} = \sqrt{1^2 + 1^2 + 0^2} = \sqrt{2} \neq 1$.
Από τη βάση αυτή, προκύπτει (χρησιμοποιώντας τη διαδικασία ορθοκανονικοποίησης των Gram-Schmidt) μια **ορθοκανονική βαση** $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ του $\mathbb{R}^3$ ως εξής:

- $\mathbf{v}_1 = \frac{1}{\|\mathbf{y}_1\|}\mathbf{y}_1$, όπου $\mathbf{y}_1 = \mathbf{x}_1$.
  Αλλά $\mathbf{y}_1 = (1,1,0)$. Επομένως,
  $$\|\mathbf{y}_1\| = \sqrt{\langle \mathbf{y}_1, \mathbf{y}_1 \rangle} = \sqrt{1^2 + 1^2 + 0^2} = \sqrt{2}.$$
  Άρα,
  $$\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1,1,0)$$

- Στην συνέχεια,

$$\mathbf{v}_2 = \frac{1}{\|\mathbf{y}_2\|}\mathbf{y}_2, \text{ όπου } \mathbf{y}_2 = \mathbf{x}_2 - \langle \mathbf{x}_2, \mathbf{v}_1 \rangle \mathbf{v}_1.$$

Αλλά,
$\langle \mathbf{x}_2, \mathbf{v}_1 \rangle = \left\langle (0,1,1), \frac{1}{\sqrt{2}}(1,1,0) \right\rangle = \frac{1}{\sqrt{2}}(0 \cdot 1 + 1 \cdot 1 + 1 \cdot 0) = \frac{1}{\sqrt{2}}$,
οπότε
$\mathbf{y}_2 = (0,1,1) - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}(1,1,0) = (0,1,1) - \left(\frac{1}{2}, \frac{1}{2}, 0\right) = \frac{1}{2}\left(-1,1,2\right)$.
Επομένως,

$$\|\mathbf{y}_2\| = \frac{1}{2}\sqrt{(-1)^2 + 1^2 + 2^2} = \frac{\sqrt{6}}{2}.$$

Άρα,

$$\mathbf{v}_2 = \frac{1}{\frac{\sqrt{6}}{2}}\frac{1}{2}\left(-1,1,2\right) = \frac{1}{\sqrt{6}}\left(-1,1,2\right).$$

Αλγόριθμος Gram - Schmidt

- Τέλος,

$$\mathbf{v}_3 = \frac{1}{\|\mathbf{y}_3\|}\mathbf{y}_3, \ \text{όπου } \mathbf{y}_3 = \mathbf{x}_3 - \langle \mathbf{x}_3, \mathbf{v}_2 \rangle \mathbf{v}_2 - \langle \mathbf{x}_3, \mathbf{v}_1 \rangle \mathbf{v}_1$$

Αλλά

$$\langle \mathbf{x}_3, \mathbf{v}_2 \rangle = \left\langle (1,0,1), \frac{1}{\sqrt{6}}(-1,1,2) \right\rangle = \frac{1}{\sqrt{6}}(1\cdot(-1)+0\cdot1+1\cdot2) = \frac{1}{\sqrt{6}},$$

οπότε

$$\langle \mathbf{x}_3, \mathbf{v}_2 \rangle \mathbf{v}_2 = \frac{1}{\sqrt{6}}\frac{1}{\sqrt{6}}(-1,1,2) = \frac{1}{6}(-1,1,2).$$

Επίσης

$$\langle \mathbf{x}_3, \mathbf{v}_1 \rangle = \left\langle (1,0,1), \frac{1}{\sqrt{2}}(1,1,0) \right\rangle = \frac{1}{\sqrt{2}}(1\cdot1+0\cdot1+1\cdot0) = \frac{1}{\sqrt{2}},$$

οπότε

$$\langle \mathbf{x}_3, \mathbf{v}_1 \rangle \mathbf{v}_1 = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}(1,1,0) = \frac{1}{2}(1,1,0).$$

Αλγόριθμος Gram - Schmidt

Επομένως,

$$\mathbf{y}_3 = (1,0,1) - \frac{1}{6}(-1,1,2) - \frac{1}{2}(1,1,0) = \frac{1}{6}(4,-4,4) = \frac{1}{3}(2,-2,2)\,.$$

οπότε

$$\|\mathbf{y}_3\| = \frac{1}{3}\sqrt{2^2+(-2)^2+2^2} = \frac{\sqrt{12}}{3} = \frac{2\sqrt{3}}{3}.$$

Τελικά

$$\mathbf{v}_3 = \frac{1}{\frac{2\sqrt{3}}{3}}\frac{1}{3}(2,-2,2) == \frac{1}{\sqrt{3}}(1,-1,1)\,.$$

Άρα, η ζητούμενη ορθοκανονική βάση είναι το σύνολο

$$\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\} = \{\frac{1}{\sqrt{2}}(1,1,0), \frac{1}{\sqrt{6}}(-1,1,2), \frac{1}{\sqrt{3}}(1,-1,1)\}$$

Αλγόριθμος Gram - Schmidt

(Πράγματι, $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = \|\mathbf{v}_3\| = 1$, δηλαδή η βάση αυτή είναι κανονική. Επίσης,

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \left\langle \frac{1}{\sqrt{2}}(1,1,0), \frac{1}{\sqrt{6}}(-1,1,2) \right\rangle = \frac{1}{\sqrt{12}}(-1+1+0) = 0.$$

Ομοίως, $\langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0$, δηλαδή η βάση αυτή είναι ορθογώνια.)

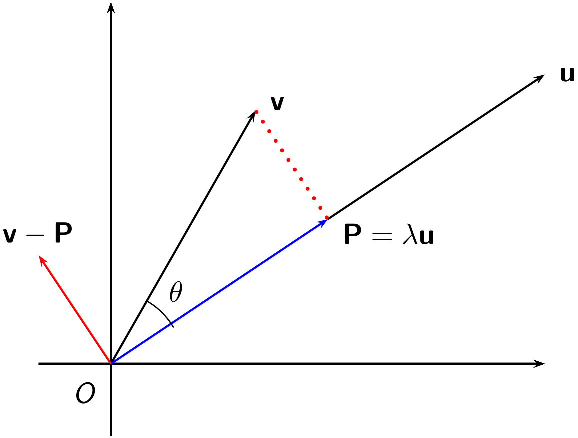
4η ΔΙΑΛΕΞΗ
ΕΣΩΤΕΡΙΚΟ ΓΙΝΟΜΕΝΟ

- **Προβολές**
- **Επίλυση αδύνατων γραμμικών συστημάτων**

Προβολές

Υπενθυμίζεται ότι αν **v**, **u** είναι δύο μη μηδενικά διανύσματα στον $\mathbb{R}^2$ τότε η **προβολή του v πάνω στο u** είναι ένα διάνυσμα **P** το οποίο

- είναι πολλαπλάσιο του **u**, δηλαδή $P = \lambda\mathbf{u}$, $\lambda \in \mathbb{R}$.
- τα διανύσματα $\mathbf{v} - P$ και **u** είναι ορθογώνια.



Προβολές

Έχοντας ως υπόδειγμα τον ορισμό αυτό μπορούμε να ορίσουμε την έννοια της προβολής και για διανύσματα με περισσότερες συντεταγμένες.

Έστω $\mathbf{v}$, $\mathbf{u}$ είναι δύο μη μηδενικά διανύσματα στον $\mathbb{R}^m$ τότε η **προβολή του $\mathbf{v}$ πάνω στο $\mathbf{u}$** είναι ένα διάνυσμα $\mathbf{P}$ το οποίο

- είναι πολλαπλάσιο του $\mathbf{u}$, δηλαδή $\mathbf{P} = \lambda\mathbf{u}$, $\lambda \in \mathbb{R}$.
- τα διανύσματα $\mathbf{v} - \mathbf{P}$ και $\mathbf{u}$ είναι ορθογώνια, δηλαδή $\langle \mathbf{v} - \mathbf{P}, \mathbf{u} \rangle = 0$.

Από τον ορισμό προκύπτει ότι

$$\langle \mathbf{v} - \mathbf{P}, \mathbf{u} \rangle = 0 \Leftrightarrow \langle \mathbf{v} - \lambda\mathbf{u}, \mathbf{u} \rangle = 0 \Leftrightarrow \langle \mathbf{v}, \mathbf{u} \rangle - \lambda \langle \mathbf{u}, \mathbf{u} \rangle = 0 \Leftrightarrow \lambda = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}$$

Άρα,

$$\mathbf{P} = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u}.$$

Προβολές

Παράδειγμα: Η προβολή του $\mathbf{v} = (1, 1, 4)$ στο $\mathbf{u} = (2, 7, 7)$ είναι το διάνυσμα

$$\mathbf{P} = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u} = \frac{1 \cdot 2 + 1 \cdot 7 + 4 \cdot 7}{2^2 + 7^2 + 7^2}(2, 7, 7) = \frac{37}{102}(2, 7, 7) = \left(\frac{74}{102}, \frac{259}{102}, \frac{259}{102}\right)$$

Τα διανύσματα $(2, 7, 7)$ και $(1, 1, 4) - \left(\frac{74}{102}, \frac{259}{102}, \frac{259}{102}\right) = \left(\frac{28}{102}, -\frac{157}{102}, -\frac{149}{102}\right)$ είναι ορθογώνια. Πράγματι

$$\left\langle (2, 7, 7), \left(\frac{28}{102}, -\frac{157}{102}, \frac{149}{102}\right) \right\rangle = 2 \cdot \frac{28}{102} + 7 \cdot \left(-\frac{157}{102}\right) + 7 \cdot \frac{149}{102}$$
$$= \frac{56 - 1099 + 1043}{102} = 0.$$

- Αν $\mathbf{v} = \lambda \mathbf{u}$, $\lambda \in \mathbb{R}$, τότε η προβολή του $\mathbf{v}$ πάνω στο $\mathbf{u}$ είναι το $\mathbf{v}$.
- Το μήκος του $\mathbf{P}$ ισούται με $\dfrac{\langle \mathbf{v}, \mathbf{u} \rangle}{\sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}} = \dfrac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|}$.
- Το διάνυσμα $\mathbf{v} - \mathbf{P}$ είναι ορθογώνιο σε κάθε πολλαπλάσιο του $\mathbf{u}$.

Βασικές ιδιότητες προβολής

- Για κάθε $\lambda \in \mathbb{R}$ ισχύει ότι

$$\|\mathbf{v} - \mathbf{P}\| \leq \|\mathbf{v} - \lambda\mathbf{u}\|$$

Η ισότητα ισχύει μόνο όταν $\mathbf{P} = \lambda\mathbf{u}$.

Απόδειξη. Πράγματι, επειδή τα διανύσματα $\mathbf{v} - \mathbf{P}$ και $\underbrace{\mathbf{P} - \lambda\mathbf{u}}_{\text{πολλαπλάσιο του } \mathbf{u}}$ είναι ορθογώνια, από το Πυθαγόρειο Θεώρημα ισχύει ότι

$$\|\mathbf{v} - \mathbf{P}\|^2 + \|\mathbf{P} - \lambda\mathbf{u}\|^2 = \|(\mathbf{v} - \mathbf{P}) + (\mathbf{P} - \lambda\mathbf{u})\|^2 = \|v - \lambda\mathbf{u}\|^2$$

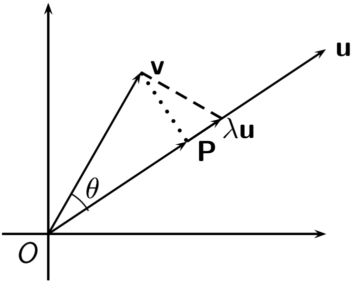
Επειδή $\|\mathbf{P} - \lambda\mathbf{u}\|^2 \geq 0$ έπεται ότι

$$\|\mathbf{v} - \mathbf{P}\|^2 \leq \|\mathbf{v} - \lambda\mathbf{u}\|^2 \Leftrightarrow \|\mathbf{v} - \mathbf{P}\| \leq \|\mathbf{v} - \lambda\mathbf{u}\|$$

Η ισότητα ισχύει μόνο όταν $\|\mathbf{P} - \lambda\mathbf{u}\|^2 = 0 \Leftrightarrow \|\mathbf{P} - \lambda\mathbf{u}\| = 0 \Leftrightarrow \mathbf{P} - \lambda\mathbf{u} = \mathbf{0} \Leftrightarrow \mathbf{P} = \lambda\mathbf{u}$.

$\square$

Η ιδιότητα αυτή εκφράζει το ενδιαφέρον γεγονός ότι η προβολή **P** του **v** πάνω στο **u** έχει την μικρότερη δυνατή απόσταση από το **v** σε σύγκριση με όλα τα υπόλοιπα πολλαπλάσια του **u**.



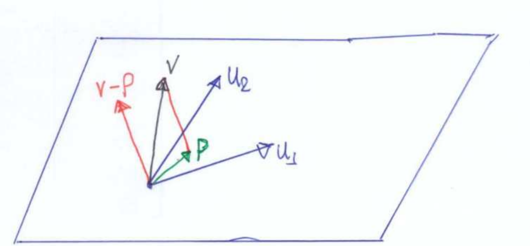
Η ανισότητα $\|\mathbf{v} - P\| \leq \|\mathbf{v} - \lambda\mathbf{u}\|$ γίνεται "φανερή" αν παρατηρήσουμε ότι το διάνυσμα $\mathbf{v} - \lambda\mathbf{u}$ αποτελεί την υποτείνουσα του ορθογωνίου τριγώνου με κάθετες πλευρές $\mathbf{v} - P$ και $P - \lambda\mathbf{u}$.

- Συχνά για λόγους σαφήνειας θα συμβολίζουμε την προβολή του **v** πάνω στο **u** με $P_{\mathbf{u}}\mathbf{v}$.

Προβολές

Υπενθυμίζεται ότι $\mathbf{v}, \mathbf{u}_1, \mathbf{u}_2$ είναι τρία μη μηδενικά διανύσματα του $\mathbb{R}^3$ τότε η **προβολή του v πάνω στο επίπεδο που ορίζουν τα $\mathbf{u}_1$, $\mathbf{u}_2$** είναι ένα διάνυσμα **P** το οποίο:

- ανήκει στο επίπεδο που ορίζουν τα $\mathbf{u}_1$, $\mathbf{u}_2$ δηλαδή
$$\mathbf{P} = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2, \lambda_1, \lambda_2 \in \mathbb{R}.$$
- τα διανύσματα $\mathbf{v} - \mathbf{P}$ και $\mathbf{u}_1$ είναι ορθογώνια, ομοίως τα διανύσματα $\mathbf{v} - \mathbf{P}$ και $\mathbf{u}_2$ είναι ορθογώνια και γενικότερα το διάνυσμα $\mathbf{v} - \mathbf{P}$ είναι ορθογώνιο με κάθε διάνυσμα που ανήκει στο επίπεδο.



Προβολές

Έχοντας ως υπόδειγμα τον ορισμό αυτό μπορούμε να ορίσουμε την έννοια της προβολής και για περισσότερα διανύσματα με περισσότερες συντεταγμένες.

Έστω $\mathbf{v} \in \mathbb{R}^n$ και $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \in \mathbb{R}^n$. Η προβολή του V πάνω στην θήκη που παράγουν τα $\mathbf{u}_1$, $\mathbf{u}_2$, …, $\mathbf{u}_k$ είναι το διάνυσμα $\mathbf{P}$ το οποίο:

- ανήκει στην θήκη που παράγουν τα $\mathbf{u}_1$, $\mathbf{u}_2$, …, $\mathbf{u}_k$, δηλαδή

$$\mathbf{P} = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \cdots + \lambda_k\mathbf{u}_k,\ \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}.$$

- το διάνυσμα $\mathbf{v} - \mathbf{P}$ είναι ορθογώνιο στα διανύσματα $\mathbf{u}_i$, $i = 1, 2, \ldots, k$ και άρα και σε κάθε διάνυσμα $\mathbf{u}$ που ανήκει στην θήκη που παράγουν τα $\mathbf{u}_1$, $\mathbf{u}_2$, …, $\mathbf{u}_k$, δηλαδή

$$\langle \mathbf{v} - \mathbf{P}, \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \cdots + \lambda_k\mathbf{u}_k \rangle = 0$$

για κάθε $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$.

Προβολές

Από τον ορισμό προκύπτει ότι η προβολή **P** ικανοποιεί το σύστημα

$$\begin{cases} \langle \mathbf{u}_1, \mathbf{v} - P \rangle = 0 \\ \langle \mathbf{u}_2, \mathbf{v} - P \rangle = 0 \\ \vdots \\ \langle \mathbf{u}_k, \mathbf{v} - P \rangle = 0 \end{cases} \Leftrightarrow \begin{cases} \langle \mathbf{u}_1, P \rangle = \langle \mathbf{u}_1, \mathbf{v} \rangle \\ \langle \mathbf{u}_2, P \rangle = \langle \mathbf{u}_2, \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{u}_k, P \rangle = \langle \mathbf{u}_k, \mathbf{v} \rangle \end{cases} \Leftrightarrow \begin{cases} \langle \mathbf{u}_1, \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_k \mathbf{u}_k \rangle = \langle \mathbf{u}_1 \\ \langle \mathbf{u}_2, \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_k \mathbf{u}_k \rangle = \langle \mathbf{u}_2 \\ \vdots \\ \langle \mathbf{u}_k, \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_k \mathbf{u}_k \rangle = \langle \mathbf{u}_k \end{cases}$$

$$(\Sigma) \begin{cases} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle \lambda_1 + \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \lambda_2 + \cdots + \langle \mathbf{u}_1, \mathbf{u}_k \rangle \lambda_k = \langle \mathbf{u}_1, \mathbf{v} \rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1 \rangle \lambda_1 + \langle \mathbf{u}_2, \mathbf{u}_2 \rangle \lambda_2 + \cdots + \langle \mathbf{u}_2, \mathbf{u}_k \rangle \lambda_k = \langle \mathbf{u}_2, \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{u}_k, \mathbf{u}_1 \rangle \lambda_1 + \langle \mathbf{u}_k, \mathbf{u}_2 \rangle \lambda_2 + \cdots + \langle \mathbf{u}_k, \mathbf{u}_k \rangle \lambda_k = \langle \mathbf{u}_k, \mathbf{v} \rangle \end{cases}$$

Άρα, οι συντελεστές λ_1, λ_2, ..., λ_k που εκφράζουν την προβολή P του $\mathbf{v}$ πάνω στην θήκη που παράγουν τα $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_k$ προκύπτουν από την λύση του συστήματος (Σ).

Προβολές

Παρατηρήστε ότι αν η βάση $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_k$ είναι ορθογώνια ή ορθοκανονική τότε $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ ανν $i \neq j$, οπότε το σύστημα (Σ') γράφεται ισοδύναμα

$$(\Sigma) \begin{cases} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle \, \lambda_1 = \langle \mathbf{u}_1, \mathbf{v} \rangle \\ \langle \mathbf{u}_2, \mathbf{u}_2 \rangle \, \lambda_2 = \langle \mathbf{u}_2, \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{u}_k, \mathbf{u}_k \rangle \, \lambda_k = \langle \mathbf{u}_k, \mathbf{v} \rangle \end{cases} \Leftrightarrow \begin{cases} \lambda_1 = \dfrac{\langle \mathbf{u}_1, \mathbf{v} \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \\ \lambda_2 = \dfrac{\langle \mathbf{u}_2, \mathbf{v} \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} \\ \vdots \\ \lambda_k = \dfrac{\langle \mathbf{u}_k, \mathbf{v} \rangle}{\langle \mathbf{u}_k, \mathbf{u}_k \rangle} \end{cases}$$

οπότε αν $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_k$ είναι ορθογώνια ή ορθοκανονική βάση του W, τότε η προβολή του $\mathbf{v}$ στο W ισούται με

$$P = \frac{\langle \mathbf{u}_1, \mathbf{v} \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \frac{\langle \mathbf{u}_2, \mathbf{v} \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 + \cdots + \frac{\langle \mathbf{u}_k, \mathbf{v} \rangle}{\langle \mathbf{u}_k, \mathbf{u}_k \rangle}\mathbf{u}_n$$

Προβολές

Παράδειγμα

Να βρεθεί η προβολή του $(1,2,3)$ στον υπόχωρο W του $\mathbb{R}^3$ όπου

$$W = \{(x,y,z) \in \mathbb{R}^3 : x + 2y - z = 0\}.$$

Λύση. Μια βάση του W είναι τα διανύσματα

$$\mathbf{x}_1 = (1,0,1) \text{ και } \mathbf{x}_2 = (0,1,2)$$

Από τη βάση αυτή, προκύπτει (εφαρμόζοντας τη διαδικασία Gram - Schmidt) μια ορθοκανονική βάση $\{\mathbf{v}_1, \mathbf{v}_2\}$ του W ως εξής:

$$\mathbf{v}_1 = \frac{1}{\|\mathbf{y}_1\|}\mathbf{y}_1, \text{ όπου } \mathbf{y}_1 = \mathbf{x}_1.$$

Αλλά $\mathbf{y}_1 = (1,0,1)$. Επομένως,

$$\|\mathbf{y}_1\| = \sqrt{\langle \mathbf{y}_1, \mathbf{y}_1 \rangle} = \sqrt{1^2 + 0^2 + 1^2} = \sqrt{2}.$$

Άρα, $\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1,0,1) = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$.

Προβολές

Επίσης,

$$\mathbf{v}_2 = \frac{1}{\|\mathbf{y}_2\|}\mathbf{y}_2,\ \text{όπου } \mathbf{y}_2 = \mathbf{x}_2 - \langle \mathbf{x}_2, \mathbf{v}_1 \rangle \mathbf{v}_1.$$

Αλλά

$$\langle \mathbf{x}_2, \mathbf{v}_1 \rangle = \left\langle (0,1,2), \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) \right\rangle = \frac{2}{\sqrt{2}} = \sqrt{2}.$$

Επομένως,

$$\mathbf{y}_2 = (0,1,2) - \sqrt{2}\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) = (0,1,2) - (1,0,1) = (-1,1,1),$$

οπότε

$$\|\mathbf{y}_2\| = \sqrt{(-1)^2 + 1^2 + 1^2} = \sqrt{3}.$$

Άρα, $\mathbf{v}_2 = \frac{1}{\sqrt{3}}(-1,1,1) = \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$.

Προβολές

Επομένως, μια ορθοκανονική βάση του W είναι το σύνολο

$$\{\mathbf{v}_1, \mathbf{v}_2\} = \left\{\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)\right\}.$$

Η προβολή του $\mathbf{x} = (1, 2, 3)$ στον υπόχωρο W είναι το διάνυσμα

$$P = \frac{\langle \mathbf{x}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 + \frac{\langle \mathbf{x}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_2$$

όπου

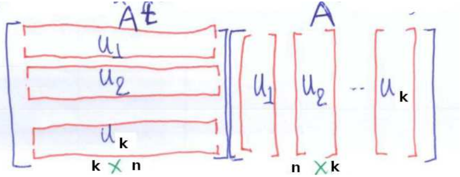
$$\langle \mathbf{x}, \mathbf{v}_1 \rangle = \left\langle (1, 2, 3), \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) \right\rangle = \frac{4}{\sqrt{2}}$$

$$\langle \mathbf{x}, \mathbf{v}_2 \rangle = \left\langle (1, 2, 3), \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \right\rangle = \frac{4}{\sqrt{3}}.$$

Άρα, τελικά

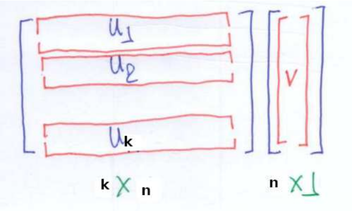
$$\begin{aligned} P &= \frac{4}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) + \frac{4}{\sqrt{3}}\left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \\ &= (2, 0, 2) + \left(-\frac{4}{3}, \frac{4}{3}, \frac{4}{3}\right) = \left(\frac{2}{3}, \frac{4}{3}, \frac{10}{3}\right). \end{aligned}$$

Παρατηρήστε ότι η μήτρα των συντελεστών του συστήματος (Σ) ισούται με A^tA όπου A είναι η μήτρα με στήλες τα $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$.



$$= \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_2 \rangle & \cdots & \langle \mathbf{u}_1, \mathbf{u}_k \rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1 \rangle & \langle \mathbf{u}_2, \mathbf{u}_2 \rangle & \cdots & \langle \mathbf{u}_2, \mathbf{u}_k \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{u}_k, \mathbf{u}_1 \rangle & \langle \mathbf{u}_k, \mathbf{u}_2 \rangle & \cdots & \langle \mathbf{u}_k, \mathbf{u}_k \rangle \end{bmatrix}$$

Επίσης, η μήτρα των σταθερών του συστήματος (Σ) ισούται με A^tV όπου V είναι η μήτρα στήλη που ισούται με το $\mathbf{v}$.



$$= \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{v} \rangle \\ \langle \mathbf{u}_2, \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{u}_k, \mathbf{v} \rangle \end{bmatrix}$$

Προβολές

Άρα, το σύστημα (Σ) γράφεται ισοδύναμα

$$A^t A \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{bmatrix} = A^t V.$$

Επομένως,

Προβολές

Η προβολή $\mathbf{P}$ του διανύσματος $\mathbf{v}$ στην γραμμική θήκη που παράγουν τα διανύσματα $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_k \in \mathbb{R}^n$ δίδεται από τον τύπο:

$$\mathbf{P} = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \cdots + \lambda_k\mathbf{u}_k$$

όπου λ_1, λ_2, ..., λ_k είναι λύση του (συμβιβαστού) συστήματος

$$A^t AX = A^t V$$

όπου

- A είναι η $n \times k$ μήτρα με στήλες τα $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_k$
- V είναι η $n \times 1$ μήτρα στήλη ίση με $\mathbf{v}$
- X είναι η $k \times 1$ μήτρα στήλη ίση με $(\lambda_1, \lambda_2, \ldots, \lambda_k)$.

Παρατήρηση: Αν τα $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_k$ δεν είναι γραμμικώς ανεξάρτητα, τότε υπάρχουν άπειρες λύσεις για τα λ_1, λ_2, ..., λ_k.

Βασικές ιδιότητες προβολής

- Αν $\mathbf{u} \in < \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k >$, τότε η προβολή του $\mathbf{u}$ πάνω στην γραμμική θήκη που παράγουν τα διανύσματα $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_k$ είναι το $\mathbf{u}$.
- Κάποιες φορές, για λόγους σαφήνειας, θα συμβολίζουμε την προβολή του $\mathbf{v}$ πάνω στο W με $P_W\mathbf{v}$.

- Για κάθε διάνυσμα $\mathbf{u}$ που ανήκει στην γραμμική θήκη $< \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k >$ ισχύει ότι

$$\|\mathbf{v} - \mathbf{P}\| \leq \|\mathbf{v} - \mathbf{u}\|$$

Η ισότητα ισχύει μόνο όταν $P = \mathbf{u}$.

Απόδειξη. Πράγματι, επειδή τα διανύσματα $\mathbf{v} - \mathbf{P}$ και $\underbrace{\mathbf{P} - \mathbf{u}}_{\text{στοιχείο της θήκης}}$ είναι ορθογώνια, από το Πυθαγόρειο Θεώρημα ισχύει ότι

$$\|\mathbf{v} - \mathbf{P}\|^2 + \|\mathbf{P} - \mathbf{u}\|^2 = \|(\mathbf{v} - \mathbf{P}) + (\mathbf{P} - \mathbf{u})\|^2 = \|\mathbf{v} - \mathbf{u}\|^2$$

Επειδή $\|\mathbf{P} - \mathbf{u}\|^2 \geq 0$ έπεται ότι

$$\|\mathbf{v} - \mathbf{P}\|^2 \leq \|\mathbf{v} - \mathbf{u}\|^2 \Leftrightarrow \|\mathbf{v} - \mathbf{P}\| \leq \|\mathbf{v} - \mathbf{u}\|$$

Η ισότητα ισχύει μόνο όταν $\|\mathbf{P} - \mathbf{u}\|^2 = 0 \Leftrightarrow \|\mathbf{P} - \mathbf{u}\| = 0 \Leftrightarrow \mathbf{P} - \mathbf{u} = \mathbf{0} \Leftrightarrow \mathbf{P} = \mathbf{u}$. □

Η ιδιότητα αυτή εκφράζει το σημαντικό γεγονός ότι η προβολή $\mathbf{P}$ του $\mathbf{v}$ στην γραμμική θήκη $< \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k >$ έχει την μικρότερη δυνατή απόσταση από το $\mathbf{v}$ σε σύγκριση με όλα τα υπόλοιπα $\mathbf{u} \in < \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k >$.

Επίλυση αδύνατου συστήματος

Επίλυση αδύνατου συστήματος

Να λυθεί το γραμμικό σύστημα

$$(\Sigma)\begin{cases} x+y=5 \\ x-y=3 \\ 2x+y=12 \end{cases}$$

Από τις 2 πρώτες εξισώσεις προκύπτει ότι $x=4$, $y=1$ τα οποία δεν επαληθεύουν την 3η εξίσωση. Άρα το σύστημα είναι αδύνατο.
Το σύστημα γράφεται ισοδύναμα στην μορφή

$$x\underbrace{\begin{bmatrix}1\\1\\2\end{bmatrix}}_{\mathbf{u}_1}+y\underbrace{\begin{bmatrix}1\\-1\\1\end{bmatrix}}_{\mathbf{u}_2}=\underbrace{\begin{bmatrix}5\\3\\12\end{bmatrix}}_{\mathbf{v}}$$

$$x(1,1,2)+y(1,-1,1)=(5,3,12)$$

Επίλυση αδύνατου συστήματος

(1η γραφή) Πότε το σύστημα θα είχε λύση; Θα είχε λύση αν το $\mathbf{v} = (5, 3, 12)$ ήταν γραμμικός συνδυασμός των $\mathbf{u}_1 = (1, 1, 2)$ και $\mathbf{u}_2 = (1, -1, 1)$. Στην περίπτωση μας δεν είναι, άρα δεν μπορούμε να εκφράσουμε το $\mathbf{v} = (5, 3, 12)$ με τα $\mathbf{u}_1 = (1, 1, 2)$ και $\mathbf{u}_2 = (1, -1, 1)$.
Το κοντινότερο διάνυσμα στο $\mathbf{v} = (5, 3, 12)$ που μπορούμε να εκφράσουμε με τα $\mathbf{u}_1 = (1, 1, 2)$ και $\mathbf{u}_2 = (1, -1, 1)$ είναι η προβολή του στην γραμμική θήκη των $\mathbf{u}_1 = (1, 1, 2)$ και $\mathbf{u}_2 = (1, -1, 1)$.
Η ‘‘λύση’’ (x, y) του συστήματος θα είναι οι συντελεστές (λ_1, λ_2) της προβολής $P = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2$.

Επίλυση αδύνατου συστήματος

(**2η γραφή**) Το σύστημα θα είχε λύση αν το $\mathbf{v}$ περιέχονταν στην θήκη που παράγουν τα $\mathbf{u}_1$, $\mathbf{u}_2$, τότε θα υπήρχαν $\lambda_1, \lambda_2 \in \mathbb{R}$ ώστε

$$\mathbf{v} = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 \Leftrightarrow \mathbf{v} - \lambda_1 \mathbf{u}_1 - \lambda_2 \mathbf{u}_2 = \mathbf{0}$$

Στο σύστημα (Σ) δεν υπάρχουν τέτοια x, y και το διάνυσμα

$$\mathbf{e} = \mathbf{v} - x\mathbf{u}_1 - y\mathbf{u}_2$$

δεν γίνεται ποτέ μηδενικό.
Η ιδέα είναι ότι το $\mathbf{e}$ μπορεί να θεωρηθεί ως σφάλμα το οποίο θέλουμε να ελαχιστοποιήσουμε.
Αν χρησιμοποιήσουμε ως κριτήριο ελαχιστοποίησης το μήκος του $\mathbf{e}$, τότε η παράσταση

$$\|\mathbf{e}\| = \|\mathbf{v} - x\mathbf{u}_1 - y\mathbf{u}_2\|$$

γίνεται ελάχιστη αν και μόνο αν τα $x\mathbf{u}_1 + y\mathbf{u}_2$ ισούνται με την προβολή P του $\mathbf{v}$ στην θήκη που παράγουν τα $\mathbf{u}_1$, $\mathbf{u}_2$.

Επίλυση αδύνατου συστήματος

(**Λύση**) Έστω $A = \begin{bmatrix} \begin{bmatrix} \mathbf{u}_1 \end{bmatrix} & \begin{bmatrix} \mathbf{u}_2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix}$, $A^t = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \end{bmatrix}$,
$V = \begin{bmatrix} 5 \\ 3 \\ 12 \end{bmatrix}$ και $X = \begin{bmatrix} x \\ y \end{bmatrix}$.
Τότε η ζητούμενη λύση είναι η λύση του (συμβιβαστού) συστήματος

$$A^t A X = A^t V$$

Έχουμε ότι

$$A^t A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}$$

$$A^t V = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 12 \end{bmatrix} = \begin{bmatrix} 32 \\ 14 \end{bmatrix}$$

Άρα,

$$A^t AX = A^t V \Leftrightarrow$$

$$\begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 32 \\ 14 \end{bmatrix} \Leftrightarrow \begin{cases} 6x + 2y = 32 \\ 2x + 3y = 14 \end{cases} \Leftrightarrow \begin{cases} x = \dfrac{34}{7} \\ y = \dfrac{10}{7} \end{cases}$$

Άρα, ως λύση του αδύνατου συστήματος θα δώσουμε το ζεύγος $(x = 34/7, y = 10/7) \simeq (4.85714, 1.42857)$.
Τότε, το κοντινότερο διάνυσμα P στο $\mathbf{v} = (5, 3, 12)$ είναι το $\mathbf{P} = 34/7(1, 1, 2) + 10/7(1, -1, 1) = (44/7, 24/7, 78/7) = (6.28, 3.48, 11.14)$ και το σφάλμα είναι

$$\|e\| = \|\mathbf{v} - x\mathbf{u}_1 - y\mathbf{u}_2\| = \left\|(5, 3, 12) - \frac{34}{7}(1, 1, 2) - \frac{10}{7}(1, -1, 1)\right\|$$
$$= \|(5, 3, 12) - (44/7, 24/7, 78/7)\| = \|(-9/7, -3/7, -6/7\| = \sqrt{\frac{18}{7}}$$
$$\simeq 1.60357$$

Επίλυση αδύνατου συστήματος

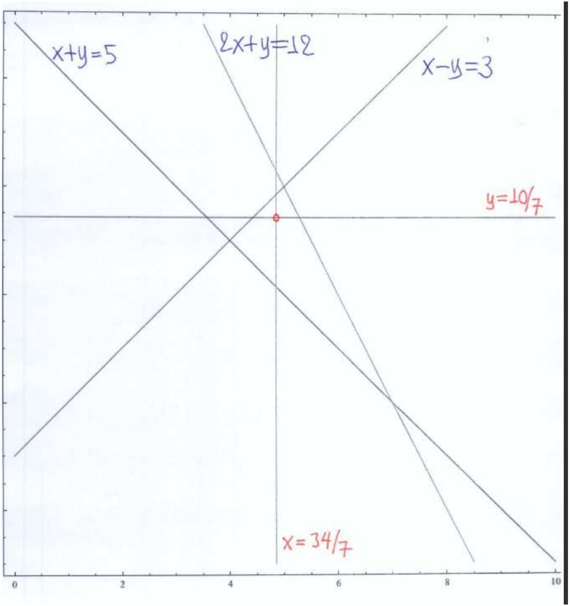
```python
import numpy as np

b = np.array([5,3,12])
A = np.transpose(np.array([[1,1,2],[1,-1,1]]))
print("b =", b)
print("A =\n",A)
#First method
x, residuals, rank, s = np.linalg.lstsq(A, b, rcond=None)
print("Least squares solution of system A x = b =", x)
print("Projection = ", x[0]*np.transpose(A)[0] + x[1]*np.
    transpose(A)[1])
print("Error", np.sqrt(residuals[0]))

#Second method
At = np.transpose(A)
newx = np.linalg.solve(At@A,At@b)
print("Solution of system At*A*x = At*b = ", newx)
```

Επίλυση αδύνατου συστήματος

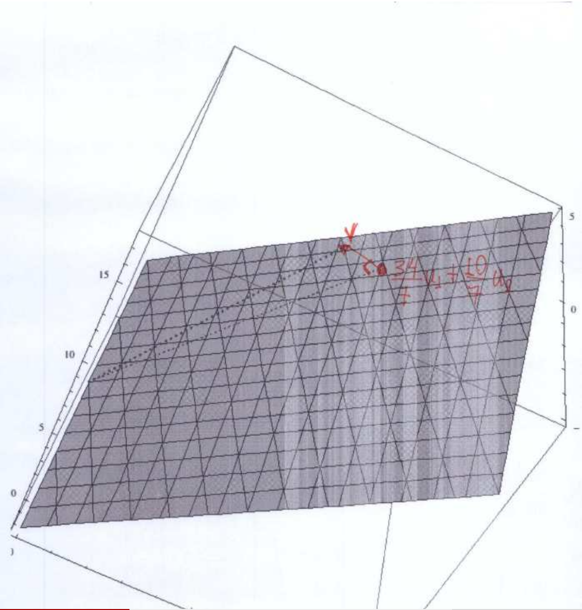
Output:

```
b = [ 5  3 12]
A =
 [[ 1  1]
 [ 1 -1]
 [ 2  1]]
Least squares solution of system A x = b = [4.85714286
   1.42857143]
Projection =  [ 6.28571429  3.42857143 11.14285714]
Error 1.6035674514745457
Solution of system At*A*x = At*b =  [4.85714286 1.42857143]
```

Γεωμετρικές ερμηνείες του συστήματος (Σ)



Γεωμετρικές ερμηνείες του συστήματος (Σ)



Το διάγραμμα ενός κτηρίου ορίζεται από τα παρακάτω 14 σημεία του $\mathbb{R}^3$:

$A(0,0,0)$	$B(12,0,0)$	$C(12,10,0)$	$D(6,17,0)$
$E(0,10,0)$	$F(4.8,2.8,0)$	$G(6.4,0,0)$	$H(12,0,20)$
$I(12,10,20)$	$J(6,17,20)$	$K(0,10,20)$	$L(0,0,20)$
$M(6.4,2.8,0)$	$N(4.8,0,0)$		

Εκ των οποίων τα επόμενα ζεύγη σημείων συνδέονται με 20 ευθύγραμμα τμήματα: $\{A,B\}$, $\{A,E\}$, $\{A,L\}$, $\{B,C\}$, $\{B,H\}$, $\{C,D\}$, $\{C,E\}$, $\{C,I\}$, $\{D,E\}$, $\{D,J\}$, $\{E,K\}$, $\{F,M\}$, $\{F,N\}$, $\{G,M\}$, $\{H,I\}$, $\{H,L\}$, $\{I,J\}$, $\{I,K\}$, $\{J,K\}$, $\{K,L\}$.
Να βρεθεί η προβολή του διαγράμματος του κτηρίου στο επίπεδο $x+y-z=0$ του $\mathbb{R}^3$.

Λύση. Κάθε $\mathbf{v} = (x, y, z)$ που ανήκει στο επίπεδο $x + y - z = 0$ ισχύει ότι

$$\mathbf{v} = (x, y, z) = (x, y, x + y) = x(1, 0, 1) + y(0, 1, 1)$$

οπότε τα $\mathbf{v}_1 = (1, 0, -1)$ και $\mathbf{v}_2 = (0, 1, -1)$ αποτελούν μια βάση των διανυσμάτων του επιπέδου $x + y - z = 0$.
Εφαρμόζοντας την διαδικασία ορθοκανονικοποίησης των Gram - Schmidt προκύπτει ότι μια ορθοκανονική βάση του επιπέδου $x + y - z = 0$ αποτελείται από τα διανύσματα $\mathbf{x}_1 = \frac{1}{\sqrt{2}}(1, 0, 1)$ και $\mathbf{x}_2 = \frac{1}{\sqrt{6}}(-1, 2, 1)$

Για την αναπαράσταση των προβολών των 14 σημείων στο επίπεδο $x + y - z = 0$ μας ενδιαφέρουν μόνο οι συντεταγμένες της προβολής κάθε σημείου (δηλαδή οι συντελεστές που εκφράζουν την προβολή κάθε σημείου ως γραμμικό συνδυασμό των $\mathbf{x}_1$, $\mathbf{x}_2$).
Επειδή η βάση $\mathbf{x}_1$, $\mathbf{x}_2$ είναι ορθοκανονική γνωρίζουμε ότι η προβολή στο επίπεδο κάθε διανύσματος $\mathbf{v} \in \mathbb{R}^3$ δίνεται από την σχέση

$$\langle \mathbf{v}, \mathbf{x}_1 \rangle \mathbf{x}_1 + \langle \mathbf{v}, \mathbf{x}_2 \rangle \mathbf{x}_2$$

επομένως, οι συντεταγμένες της προβολής κάθε διανύσματος $\mathbf{v}$ στο επίπεδο $x + y - z = 0$ είναι:

$$\left(\langle \mathbf{v}, \mathbf{x}_1 \rangle, \langle \mathbf{v}, \mathbf{x}_2 \rangle \right)$$

Οι συντεταγμένες κάθε σημείου είναι:

Για το A: $(\langle (0,0,0), \mathbf{x}_1 \rangle, \langle (0,0,0), \mathbf{x}_2 \rangle) = (0,0)$

Για το B: $(\langle (12,0,0), \mathbf{x}_1 \rangle, \langle (12,0,0), \mathbf{x}_2 \rangle) = (\frac{12}{\sqrt{2}}, -\frac{12}{\sqrt{6}})$.

Για το C: $(\langle (12,10,0), \mathbf{x}_1 \rangle, \langle (12,10,0), \mathbf{x}_2 \rangle) = (\frac{12}{\sqrt{2}}, \frac{8}{\sqrt{6}})$.

Για το D: $(\langle (6,18,0), \mathbf{x}_1 \rangle, \langle (6,18,0), \mathbf{x}_2 \rangle) = (\frac{6}{\sqrt{2}}, \frac{30}{\sqrt{6}})$.

Για το E: $(\langle (0,10,0), \mathbf{x}_1 \rangle, \langle (0,10,0), \mathbf{x}_2 \rangle) = (0, \frac{20}{\sqrt{6}})$.

Για το F: $(\langle (4.8,2.8,0), \mathbf{x}_1 \rangle, \langle (4.8,2.8,0), \mathbf{x}_2 \rangle) = (\frac{24}{5\sqrt{2}}, \frac{4}{5\sqrt{6}})$.

Για το G: $(\langle (6.4,0,0), \mathbf{x}_1 \rangle, \langle (6.4,0,0), \mathbf{x}_2 \rangle) = (\frac{32}{5\sqrt{2}}, -\frac{32}{5\sqrt{6}})$.

Για το H: $(\langle (12,0,20), \mathbf{x}_1 \rangle, \langle (12,0,20), \mathbf{x}_2 \rangle) = (\frac{32}{\sqrt{2}}, \frac{18}{\sqrt{6}})$.

Για το I: $(\langle (12,10,20), \mathbf{x}_1 \rangle, \langle (12,10,20), \mathbf{x}_2 \rangle) = (\frac{32}{\sqrt{2}}, \frac{28}{\sqrt{6}})$.

Για το J: $(\langle (6,18,20), \mathbf{x}_1 \rangle, \langle (6,18,20), \mathbf{x}_2 \rangle) = (\frac{26}{\sqrt{2}}, \frac{50}{\sqrt{6}})$.

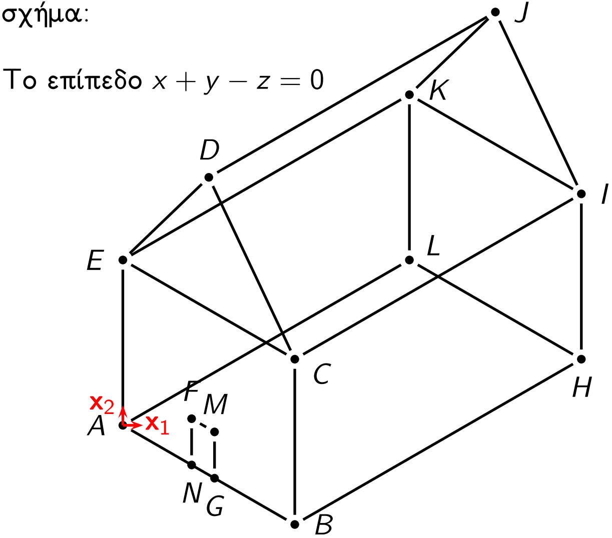
Για το K: $(\langle (0,10,20), \mathbf{x}_1 \rangle, \langle (0,10,20), \mathbf{x}_2 \rangle) = (\frac{20}{\sqrt{2}}, \frac{40}{\sqrt{6}})$.

Για το L: $(\langle (0,0,20), \mathbf{x}_1 \rangle, \langle (0,0,20), \mathbf{x}_2 \rangle) = (\frac{20}{\sqrt{2}}, \frac{20}{\sqrt{6}})$.

Για το M: $(\langle (6.4,2.8,0), \mathbf{x}_1 \rangle, \langle (6.4,2.8,0), \mathbf{x}_2 \rangle) = (\frac{32}{5\sqrt{2}}, -\frac{4}{5\sqrt{6}})$.

Για το N: $(\langle (4.8,0,0), \mathbf{x}_1 \rangle, \langle (4.8,0,0), \mathbf{x}_2 \rangle) = (\frac{24}{5\sqrt{2}}, -\frac{24}{5\sqrt{6}})$.

Σχηματικά, ενώνοντας τα αντίστοιχα σημεία, η προβολή του διαγράμματος του κτηρίου στο επίπεδο $x + y - z = 0$ απεικονίζεται στο επόμενο σχήμα:



Οι ετικέτες στο σχήμα δείχνουν τις προβολές των αντίστοιχων σημείων στο επίπεδο $x + y - z = 0$.

5η ΔΙΑΛΕΞΗ
ΕΣΩΤΕΡΙΚΟ ΓΙΝΟΜΕΝΟ

- **Διάσπαση ιδιαζουσών τιμών**

Διάσπαση ιδιαζουσών τιμών (SVD)

Μια από τις πιο χρήσιμες διασπάσεις μητρών που **ορίζονται και για μη τετραγωνικές μήτρες** είναι η **διάσπαση ιδιαζουσών τιμών** (singular value decomposition (**SVD**)).

Έστω $A \in \mathcal{M}_{m \times n}$. Η **διάσπαση ιδιαζουσών τιμών** της A είναι το γινόμενο

$$A = U \Sigma V^t$$

όπου

- $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p) \in \mathcal{M}_{m \times n}$, $p = \min\{m, n\}$, είναι μια διαγώνια $m \times n$ μήτρα με διαγώνιες τιμές $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$,
- $U \in \mathcal{M}_{m \times m}$ και $V \in \mathcal{M}_{n \times n}$ είναι ορθογώνιες μήτρες.
- Τα σ_i ονομάζονται **ιδιάζουσες τιμές** (singular values) της A.
- Οι στήλες των U, V ονομάζονται **αριστερά** και **δεξιά ιδιάζοντα διανύσματα** (singular vectors) της A αντίστοιχα.

Παρατήρηση: Για την διάσπαση ιδιαζουσών τιμών έχουν καθιερωθεί τα σύμβολα U, Σ και V.

Διάσπαση ιδιαζουσών τιμών (SVD)

Παρατήρηση: Στις μη τετραγωνικές διαγώνιες $m \times n$ μήτρες τα μη μηδενικά στοιχεία βρίσκονται στην i-οστή γραμμή και i-οστή στήλη όπου $i \leq \min\{m, n\}$ και ανάλογα με τα m, n έχουν τις μορφές

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \cdots & a_{nn} \\ 0 & 0 & \cdots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ όταν } n < m \text{ και}$$

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 & 0 & \cdots & 0 \\ \dots & \dots & \ddots & \dots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & a_{mm} & 0 & \cdots & 0 \end{bmatrix} \text{ όταν } m < n.$$

Διάσπαση ιδιαζουσών τιμών (SVD)

Αποδεικνύεται ότι κάθε πραγματική μήτρα A έχει διάσπαση ιδιαζουσών τιμών και ισχύει η παρακάτω πρόταση:

Πρόταση 16

Έστω $A \in \mathcal{M}_{m\times n}$ με διάσπαση ιδιαζουσών τιμών $A = U\Sigma V^t$, όπου $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_p)$ και $p = \min\{m, n\}$. Τότε

1. *Οι ιδιάζουσες τιμές της A ισούνται με τις (μη αρνητικές) τετραγωνικές ρίζες των p μεγαλύτερων ιδιοτιμών της $n \times n$ μήτρας A^tA, (ή της $m \times m$ μήτρας AA^t). Ανάλογα με το p, οι επιπλέον ιδιοτιμές της A^tA ή της A^tA είναι 0.*
2. *Οι στήλες v_i της ορθογώνιας μήτρας V είναι τα ιδιοδιανύσματα της $n \times n$ μήτρας A^tA.*
3. *Οι στήλες u_i της ορθογώνιας μήτρας U είναι τα ιδιοδιανύσματα της $m \times m$ μήτρας AA^t.*
4. *Για $i = 1, 2, \dots, p$, οι i-οστές στήλες u_i και v_i των U και V αντίστοιχα συνδέονται με τις σχέσεις: $Av_i = \sigma_i u_i$ και $A^tu_i = \sigma_i v_i$*

Απόδειξη. Χωρίς βλάβη της γενικότητας έστω $m \geq n$, οπότε $p = n$. Υπενθυμίζεται ότι μια τετραγωνική μήτρα B ονομάζεται ορθογωνίως διαγωνιοποιήσιμη αν $B = PDP^t$ όπου P είναι ορθογώνια μήτρα (δηλαδή $P^{-1} = P^t$). Μια τετραγωνική μήτρα είναι ορθογωνίως διαγωνιοποιήσιμη ανν είναι συμμετρική. Για κάθε μήτρα $A \in \mathcal{M}_{m \times n}$ οι μήτρες A^tA, AA^t είναι συμμετρικές, επομένως είναι ορθογώνιως διαγωνοποιήσιμες.
Ισχύει ότι $A^tA \in \mathcal{M}_n$ με

$$A^tA = (U\Sigma V^t)^t(U\Sigma V^t) = V\Sigma^t U^t U\Sigma V^t = V(\Sigma^t\Sigma)V^t$$

όπου $\Sigma^t\Sigma \in \mathcal{M}_{n \times n}$ με $\Sigma^t\Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$, δηλαδή οι στήλες της V είναι ιδιοδιανύσματα της μήτρας A^tA με αντίστοιχες ιδιοτιμές τα τετράγωνα των ιδιαζουσών τιμών της A.

Διάσπαση ιδιαζουσών τιμών (SVD)

Ομοίως, ισχύει ότι $AA^t \in \mathcal{M}_{m\times m}$ με

$$AA^t = (U\Sigma V^t)(U\Sigma V^t)^t = U\Sigma V^t V\Sigma^t U^t = U(\Sigma\Sigma^t)U^t$$

όπου $\Sigma\Sigma^t \in \mathcal{M}_{m\times m}$ με $\Sigma\Sigma^t = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2, 0, \ldots, 0)$, δηλαδή οι στήλες της U είναι ιδιοδιανύσματα της μήτρας AA^t με αντίστοιχες ιδιοτιμές τα τετράγωνα των ιδιαζουσών τιμών της A και $m-n$ επιπλεον μηδενικές ιδιοτιμές της AA^t.

Επιπλέον, έχουμε ότι

$$AV = U\Sigma V^t V \Leftrightarrow AV = U\Sigma$$

και

$$U^t A = U^t U\Sigma V^t \Leftrightarrow U^t A = \Sigma V^t \Leftrightarrow (U^t A)^t = (\Sigma V^t)^t \Leftrightarrow A^t U = V\Sigma^t.$$

Άρα, οι i-οστές στήλες u_i και v_i των U και V αντίστοιχα συνδέονται με τις σχέσεις:

$$Av_i = \sigma_i u_i \text{ για κάθε } i = 1, \ldots, p$$

και

$$A^t u_i = \sigma_i v_i \text{ για κάθε } i = 1, \ldots, p.$$

Διάσπαση ιδιαζουσών τιμών (SVD)

Από την προηγούμενη πρόταση προκύπτει το επόμενο πόρισμα:

Πόρισμα 2

Η διάσπαση ιδιαζουσών τιμών μιας συμμετρικής τετραγωνικής $n \times n$ μήτρας A με μη αρνητικές ιδιοτιμές ταυτίζεται με την ορθογώνια διαγωνιοποίηση της μήτρας A, όπου οι ιδιοτιμές της A εμφανίζονται στην D σε φθίνουσα σειρά.

Διάσπαση ιδιαζουσών τιμών (SVD)

Παράδειγμα υπολογισμού της SVD

Να βρεθεί η διάσπαση ιδιαζουσών τιμών της μήτρας

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

Εδώ $p = \min\{3, 2\} = 2$. Άρα, η A έχει δυο ιδιάζουσες τιμές. Μας συμφέρει να τις υπολογίσουμε από τις ιδιοτιμές της 2×2 μήτρας AA^t. Έχουμε ότι

$$AA^t = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 14 & 6 \\ 6 & 9 \end{bmatrix}.$$

Από την χαρακτηριστική εξίσωση της AA^t:

$$\begin{vmatrix} 14-\lambda & 6 \\ 6 & 9-\lambda \end{vmatrix} = 0 \Leftrightarrow (14-\lambda)(9-\lambda) - 6^2 = 0 \Leftrightarrow \lambda^2 - 23\lambda + 90 = 0$$

Διάσπαση ιδιαζουσών τιμών (SVD)

προκύπτει ότι οι ιδιοτιμές της AA^t σε φθίνουσα σειρά είναι

$$\lambda_1 = 18 \text{ και } \lambda_2 = 5$$

Επομένως, οι ιδιάζουσες τιμές της A είναι

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{18} \text{ και } \sigma_2 = \sqrt{\lambda_2} = \sqrt{5}.$$

Στην ιδιοτιμή $\lambda_1 = 18$ αντιστοιχεί το ιδιοδιάνυσμα $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$, ενώ στην ιδιοτιμή $\lambda_2 = 5$ αντιστοιχεί το ιδιοδιάνυσμα $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Κανονικοποιώντας τα δύο ιδιοδιανύσματα προκύπτούν αντίστοιχα τα διανύσματα $u_1 = \begin{bmatrix} \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \end{bmatrix}$ και $u_2 = \begin{bmatrix} -\frac{2}{\sqrt{3}} \\ \frac{3}{\sqrt{13}} \end{bmatrix}$, τα οποία αποτελούν τις στήλες της ορθογώνιας μήτρας U, δηλαδή

$$U = \begin{bmatrix} \frac{3}{\sqrt{13}} & -\frac{2}{\sqrt{3}} \\ \frac{2}{\sqrt{13}} & \frac{3}{\sqrt{13}} \end{bmatrix}.$$

Διάσπαση ιδιαζουσών τιμών (SVD)

Απομένει να βρούμε τις στήλες της ορθογώνιας μήτρας V. Οι στήλες της V είναι τα ιδιοδιανύσματα της 3×3 μήτρας $A^t A$.
Τις δύο πρώτες στήλες της V μπορούμε να τις υπολογίσουμε από την σχέση

$$A^t u_i = \sigma_i v_i$$

Πράγματι,

$$v_1 = \frac{1}{\sigma_1} A^t u_1 = \frac{1}{\sqrt{18}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \end{bmatrix} = \frac{1}{\sqrt{18}} \begin{bmatrix} \frac{7}{\sqrt{13}} \\ \frac{4}{\sqrt{13}} \\ \frac{13}{\sqrt{13}} \end{bmatrix} = \begin{bmatrix} \frac{7}{\sqrt{234}} \\ \frac{4}{\sqrt{234}} \\ \frac{13}{\sqrt{234}} \end{bmatrix}$$

και

$$v_2 = \frac{1}{\sigma_2} A^t u_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{13}} \\ \frac{3}{\sqrt{13}} \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \frac{4}{\sqrt{13}} \\ -\frac{7}{\sqrt{13}} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{4}{\sqrt{65}} \\ -\frac{7}{\sqrt{65}} \\ 0 \end{bmatrix}$$

Διάσπαση ιδιαζουσών τιμών (SVD)

Την τρίτη στήλη της V θα την υπολογίσουμε βρίσκοντας το ιδιοδιάνυσμα που αντιστοιχεί την ιδιοτιμή 0 της μήτρας A^tA:
Έχουμε ότι

$$A^tA = \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 7 \\ 0 & 5 & 4 \\ 7 & 4 & 13 \end{bmatrix}.$$

Στην ιδιοτιμή 0 αντιστοιχεί το ιδιοδιάνυσμα $\begin{bmatrix} 7 \\ 4 \\ -5 \end{bmatrix}$. Κανονικοποιώντας, προκύπτει το διάνυσμα $v_3 = \begin{bmatrix} \frac{7}{3\sqrt{10}} \\ \frac{4}{3\sqrt{10}} \\ -\frac{5}{3\sqrt{10}} \end{bmatrix}$. Επομένως, η ορθογώνια 3×3 μήτρα V ισούται με $V = \begin{bmatrix} \frac{7}{\sqrt{234}} & \frac{4}{\sqrt{65}} & \frac{7}{3\sqrt{10}} \\ \frac{4}{\sqrt{234}} & -\frac{7}{\sqrt{65}} & \frac{4}{3\sqrt{10}} \\ \frac{13}{\sqrt{234}} & 0 & -\frac{5}{3\sqrt{10}} \end{bmatrix}$.

Διάσπαση ιδιαζουσών τιμών (SVD)

Συνοψίζοντας, η διάσπαση ιδιαζουσών τιμών της A είναι

$$A = U\Sigma V^t = \begin{bmatrix} \frac{3}{\sqrt{13}} & -\frac{2}{\sqrt{3}} \\ \frac{2}{\sqrt{13}} & \frac{3}{\sqrt{13}} \end{bmatrix} \begin{bmatrix} \sqrt{18} & 0 & 0 \\ 0 & \sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} \frac{7}{\sqrt{234}} & \frac{4}{\sqrt{234}} & \frac{13}{\sqrt{234}} \\ \frac{4}{\sqrt{65}} & -\frac{7}{\sqrt{65}} & 0 \\ \frac{7}{3\sqrt{10}} & \frac{4}{3\sqrt{10}} & -\frac{5}{3\sqrt{10}} \end{bmatrix}$$

Παρατήρηση: Η διάσπαση ιδιαζουσών τιμών ανακαλύφθηκε το 1873 από τον Beltrami αλλά ο πρώτος πρακτικός τρόπος υπολογισμού της για μεγάλες μήτρες δόθηκε το 1965 από τους Golub και Kahan: Η μέθοδός τους υπολογίζει την διάσπαση ιδιαζουσών τιμών χωρίς να χρειαστεί η χρονοβόρα διαδικασία υπολογισμού της ορθογώνιας διαγωνιοποίησης της μήτρας AA^t ή A^tA.

Διάσπαση ιδιαζουσών τιμών (SVD)

Η διάσπαση ιδιαζουσών τιμών δίνει πολλές πληροφορίες για την μήτρα A και τους βασικούς υποχώρους που σχετίζονται με αυτή.

Πρόταση 17

Έστω $A \in \mathcal{M}_{m\times n}$ και $A = U\Sigma V^t$ η διάσπαση ιδιαζουσών τιμών της A. Τότε

1. *Το rank της A ισούται με τον αριθμό των μη μηδενικών ιδιαζουσών τιμών της.*
2. *Η εικόνα της A παράγεται από τις πρώτες r στήλες της U.*
3. *Ο πυρήνας της A παράγεται από τις τελευταίες $n-r$ στήλες της V.*

Διάσπαση ιδιαζουσών τιμών (SVD)

Πρόταση 18 (Διάσπαση μήτρας ως άθροισμα μητρών με rank 1)

Έστω $A \in \mathcal{M}_{m\times n}$ και $A = U\Sigma V^t$ η διάσπαση ιδιαζουσών τιμών της A, με $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p)$, όπου $p = \min\{m, n\}$, U ορθογώνια $m \times m$ μήτρα με στήλες $\mathbf{u}_1$, $\mathbf{u}_2$, $\ldots$, $\mathbf{u}_m$, V ορθογώνια $n \times n$ μήτρα με στήλες $\mathbf{v}_1$, $\mathbf{v}_2$, $\ldots$, $\mathbf{v}_n$. Τότε

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^t + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^t + \cdots + \sigma_p \mathbf{u}_p \mathbf{v}_p^t \qquad (1)$$

Με άλλα λόγια, η μήτρα A εκφράζεται ως άθροισμα μητρών που έχουν rank 1 (οι οποίες προκύπτουν από τις στήλες των μητρών U και V. Κάποιες φορές αντί της γραφής $A = U\Sigma V^t$ η διάσπαση ιδιαζουσών τιμών της A ορίζεται ως το (ισοδύναμο) άθροισμα (1)

Παράδειγμα διάσπασης σε μήτρες με rank 1

Δίδεται η μήτρα

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

Να εκφραστεί ως άθροισμα μητρών που έχουν rank 1.

Λύση. Η διάσπαση ιδιαζουσών τιμών της A είναι

$$A = U\Sigma V^t = \begin{bmatrix} \frac{3}{\sqrt{13}} & \frac{-2}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} & \frac{3}{\sqrt{13}} \end{bmatrix} \begin{bmatrix} \sqrt{18} & 0 & 0 \\ 0 & \sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} \frac{7}{\sqrt{234}} & \frac{4}{\sqrt{234}} & \frac{13}{\sqrt{234}} \\ \frac{4}{\sqrt{65}} & \frac{-7}{\sqrt{65}} & 0 \\ \frac{7}{3\sqrt{10}} & \frac{4}{3\sqrt{10}} & \frac{-5}{3\sqrt{10}} \end{bmatrix}$$

Επομένως, η A εκφράζεται ως άθροισμα μητρών που έχουν rank 1 ως εξής:

$$
\begin{aligned}
A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 2 \end{bmatrix} &= \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t \\
&= \sqrt{18} \begin{bmatrix} \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \end{bmatrix} \begin{bmatrix} \frac{7}{\sqrt{234}} & \frac{4}{\sqrt{234}} & \frac{13}{\sqrt{234}} \end{bmatrix} + \sqrt{5} \begin{bmatrix} \frac{-2}{\sqrt{13}} \\ \frac{3}{\sqrt{13}} \end{bmatrix} \begin{bmatrix} \frac{4}{\sqrt{65}} & \frac{-7}{\sqrt{65}} & 0 \end{bmatrix} \\
&= \frac{\sqrt{18}}{\sqrt{13 \cdot 234}} \begin{bmatrix} 3 \\ 2 \end{bmatrix} \begin{bmatrix} 7 & 4 & 13 \end{bmatrix} + \frac{\sqrt{5}}{\sqrt{13 \cdot 65}} \begin{bmatrix} -2 \\ 3 \end{bmatrix} \begin{bmatrix} 4 & -7 & 0 \end{bmatrix} \\
&= \frac{1}{13} \begin{bmatrix} 21 & 12 & 39 \\ 14 & 8 & 26 \end{bmatrix} + \frac{1}{13} \begin{bmatrix} -8 & 14 & 0 \\ 12 & -21 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{21}{13} & \frac{12}{13} & 3 \\ \frac{14}{13} & \frac{8}{13} & 2 \end{bmatrix} + \begin{bmatrix} \frac{-8}{13} & \frac{14}{13} & 0 \\ \frac{12}{13} & \frac{-21}{13} & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1.61538 & 0.92308 & 3 \\ 1.07692 & 0.61538 & 2 \end{bmatrix} + \begin{bmatrix} -0.61538 & 1.07692 & 0 \\ 0.92308 & -1.61538 & 0 \end{bmatrix}
\end{aligned}
$$

Παρατήρηση: Έστω μια μήτρα A με $\mathrm{rank}(A) = r$, δηλαδή η διάσταση του χώρου των στηλών (ή των γραμμών) της A είναι r. Σε πολλές εφαρμογές όπου η διάσταση r είναι πολύ μεγάλη, ψάχνουμε μια άλλη μήτρα B η οποία διαφέρει όσο το δυνατόν ''λιγότερο'' από την A, αλλά έχει μικρότερο rank. Η επόμενη πρόταση μας δείχνει πως μπορούμε να χρησιμοποιήσουμε την διάσπαση ιδιαζουσών τιμών της A για να βρούμε την ''πλησιέστερη'' μήτρα B η οποία έχει rank k, όπου $k < r$.

Πρόταση 19 (Θεώρημα Eckart - Young)

Έστω $A \in \mathcal{M}_{m\times n}$ με rank r και διάσπαση ιδιαζουσών τιμών

$$A = U\Sigma V^t = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^t + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^t + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^t$$

Για κάθε $k < r = \mathrm{rank(A)}$ η μήτρα

$$A_k = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^t + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^t + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^t$$

έχει rank k και για οποιαδήποτε άλλη μήτρα B με rank k ισχύει ότι

$$\|A - A_k\| \leq \|A - B\|$$

ή ισοδύναμα $\displaystyle\min_{\mathrm{rank(B)=k}} \|A - B\| = \|A - A_k\|$ όπου $\|P\|$ είναι οποιαδήποτε άλλη norm μητρών.
Με άλλα λόγια, η μήτρα A_k είναι η πλησιέστερη μήτρα με rank k στην A από όλες τις μήτρες B που έχουν rank k.

Πρόταση 11 (**Θεώρημα Eckart - Young (συνέχεια)**)

Ειδικότερα, για τις παρακάτω συνηθισμένες norm η απόσταση αυτή είναι:

Αν $\|A\| = \sqrt{\mathrm{tr}(AA^t)} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2}$ *(Frobenius norm) τότε*

$$\|A - A_k\| = \min_{\mathrm{rank(B)=k}} \|A - B\| = \sqrt{\sum_{i=k+1}^{r} \sigma_i^2}.$$

Αν $\|A\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$ *όπου* $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots x_n^2}$ *(2-norm) τότε*

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

Διάσπαση ιδιαζουσών τιμών (SVD)

Παρατηρήσεις:

1. Η μήτρα A_k μπορεί να υπολογισθεί και από το γινόμενο
$$A_k = UD_kV^t$$
όπου
$$D_k = \mathrm{diag}(\sigma_1, \ldots, \sigma_k, 0, \ldots, 0).$$
2. Η μήτρα A_k του θεωρήματος Eckart - Young ονομάζεται **βέλτιστη rank κ προσέγγιση** της A
3. Σε πολλές περιπτώσεις, ενώ οι μήτρες που αντιστοιχούν στα δεδομένα έπρεπε να έχουν υποχρεωτικά μικρό rank, όμως εξαιτίας σφαλμάτων στα παρατηρούμενα δεδομένα τελικά καταλήγουμε με μήτρες που έχουν μεγάλο rank. Το θεώρημα Eckart - Young μας εξασφαλίζει ότι για να λάβουμε μια μήτρα μικροτέρου rank μπορούμε δικαιολογημένα να αγνοήσουμε τις ιδιάζουσες τιμές που είναι της ίδιας τάξης με τα σφάλματα στα δεδομένα μας (δηλαδή να τις θέσουμε ίσες με μηδέν) και να υπολογίσουμε εκ νέου το γινόμενο $U\Sigma V$.

Παράδειγμα εύρεσης βέλτιστης προσέγγισης

Να βρεθεί η βέλτιστη rank 1 προσέγγιση της μήτρας $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 2 \end{bmatrix}$ του προηγούμενου παραδείγματος.

Λύση Προηγουμένως βρήκαμε ότι η μήτρα A έχει διάσπαση ιδιαζουσών τιμών

$$\begin{aligned}
A = U\Sigma V^t &= \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t \\
&= \begin{bmatrix} \frac{3}{\sqrt{13}} & -\frac{2}{\sqrt{3}} \\ \frac{2}{\sqrt{13}} & \frac{3}{\sqrt{13}} \end{bmatrix} \begin{bmatrix} \sqrt{18} & 0 & 0 \\ 0 & \sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} \frac{7}{\sqrt{234}} & \frac{4}{\sqrt{234}} & \frac{13}{\sqrt{234}} \\ \frac{4}{\sqrt{65}} & -\frac{7}{\sqrt{65}} & 0 \\ \frac{7}{3\sqrt{10}} & \frac{4}{3\sqrt{10}} & -\frac{5}{3\sqrt{10}} \end{bmatrix} \\
&= \begin{bmatrix} \frac{21}{13} & \frac{12}{13} & 3 \\ \frac{14}{13} & \frac{8}{13} & 2 \end{bmatrix} + \begin{bmatrix} \frac{-8}{13} & \frac{14}{13} & 0 \\ \frac{12}{13} & \frac{-21}{13} & 0 \end{bmatrix}
\end{aligned}$$

Με βάση το Θεώρημα Eckart - Young, η βέλτιστη rank 1 προσέγγιση A_1 της A προκύπτει κρατώντας μόνο την μεγαλύτερη ιδιάζουσα τιμή της, και θέτωντας ίσες με 0 τις υπόλοιπες. Οπότε

$$A_1 = \sigma_1 u_1 v_1^t = \sqrt{18}\begin{bmatrix} \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \end{bmatrix}\begin{bmatrix} \frac{7}{\sqrt{234}} & \frac{4}{\sqrt{234}} & \frac{13}{\sqrt{234}} \end{bmatrix} = \begin{bmatrix} \frac{21}{13} & \frac{12}{13} & 3 \\ \frac{14}{13} & \frac{8}{13} & 2 \end{bmatrix}$$

Ή, ισοδύναμα

$$A_1 = UD_1V^t = \begin{bmatrix} \frac{3}{\sqrt{13}} & -\frac{2}{\sqrt{3}} \\ \frac{2}{\sqrt{13}} & \frac{3}{\sqrt{13}} \end{bmatrix}\begin{bmatrix} \sqrt{18} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{7}{\sqrt{234}} & \frac{4}{\sqrt{234}} & \frac{13}{\sqrt{234}} \\ \frac{4}{\sqrt{65}} & -\frac{7}{\sqrt{65}} & 0 \\ \frac{7}{3\sqrt{10}} & \frac{4}{3\sqrt{10}} & -\frac{5}{3\sqrt{10}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{21}{13} & \frac{12}{13} & 3 \\ \frac{14}{13} & \frac{8}{13} & 2 \end{bmatrix}$$

Διάσπαση ιδιαζουσών τιμών (SVD)

Παράδειγμα εύρεσης βέλτιστης προσέγγισης

Να βρεθούν οι βέλτιστες rank 1,2 και 3 προσεγγίσεις της 17×4 μήτρας A με

$$A = \begin{bmatrix} 5.1 & 3.5 & 1.4 & 0.2 \\ 4.9 & 3.0 & 1.4 & 0.2 \\ 4.7 & 3.2 & 1.3 & 0.2 \\ 4.6 & 3.1 & 1.5 & 0.2 \\ 5.0 & 3.6 & 1.4 & 0.2 \\ 5.4 & 3.9 & 1.7 & 0.4 \\ 4.6 & 3.4 & 1.4 & 0.3 \\ 5.0 & 3.4 & 1.5 & 0.2 \\ 4.4 & 2.9 & 1.4 & 0.2 \\ 4.9 & 3.1 & 1.5 & 0.1 \\ 5.4 & 3.7 & 1.5 & 0.2 \\ 4.8 & 3.4 & 1.6 & 0.2 \\ 4.8 & 3.0 & 1.4 & 0.1 \\ 4.3 & 3.0 & 1.1 & 0.1 \\ 5.8 & 4.0 & 1.2 & 0.2 \\ 5.7 & 4.4 & 1.5 & 0.4 \\ 5.4 & 3.9 & 1.3 & 0.4 \end{bmatrix}$$

Λύση. Αρχικά, υπολογίζουμε την διάσπαση ιδιαζουσών τιμών της A από την οποία προκύπτει

$$A = U\Sigma V^t$$

όπου U έχει διαστάσεις 17×17, Σ έχει διαστάσεις 17×4 και V έχει διαστάσεις 4×4. Οι ιδιάζουσες τιμές της A είναι $\sigma_1 = 25.7804$, $\sigma_2 = 0.823796$, $\sigma_3 = 0.588009$ και $\sigma_4 = 0.218905$. Άρα, η A έχει rank 4.

Διάσπαση ιδιαζουσών τιμών (SVD)

Μηδενίζοντας την σ_4 και υπολογίζοντας εκ νέου το γινόμενο $U\Sigma V^t$ προκύπτει ότι η βέλτιστη rank 3 προσέγγιση της A είναι η μήτρα

$$A_3 = \begin{bmatrix} 5.10347 & 3.49462 & 1.39846 & 0.215215 \\ 4.87826 & 3.03376 & 1.40968 & 0.104607 \\ 4.69864 & 3.2021 & 1.3006 & 0.194054 \\ 4.60012 & 3.09981 & 1.49995 & 0.200537 \\ 5.01464 & 3.57728 & 1.39349 & 0.26421 \\ 5.38496 & 3.92334 & 1.70669 & 0.334034 \\ 4.59936 & 3.401 & 1.40029 & 0.29718 \\ 5.003 & 3.39534 & 1.49867 & 0.213155 \\ 4.39334 & 2.91034 & 1.40296 & 0.170785 \\ 4.90619 & 3.09039 & 1.49724 & 0.127164 \\ 5.40587 & 3.69088 & 1.49739 & 0.225773 \\ 4.8137 & 3.37874 & 1.5939 & 0.26009 \\ 4.80183 & 2.99716 & 1.39919 & 0.108015 \\ 4.31786 & 2.97227 & 1.09205 & 0.178361 \\ 5.80291 & 3.99549 & 1.19871 & 0.212758 \\ 5.70191 & 4.39704 & 1.49915 & 0.408359 \\ 5.37717 & 3.93544 & 1.31016 & 0.299859 \end{bmatrix}$$

η οποία απέχει απόσταση $\sigma_4 = 0.218905$ από την A αν χρησιμοποιήσουμε την 2-norm

Διάσπαση ιδιαζουσών τιμών (SVD)

Μηδενίζοντας τις σ_3, σ_4 και υπολογίζοντας πάλι το γινόμενο $U\Sigma V^t$ προκύπτει ότι η βέλτιστη rank 2 προσέγγιση της A είναι η μήτρα

$$A_2 = \begin{bmatrix} 5.07927 & 3.51279 & 1.43699 & 0.231072 \\ 4.81528 & 3.08107 & 1.50996 & 0.145878 \\ 4.67784 & 3.21774 & 1.33374 & 0.207691 \\ 4.65511 & 3.0585 & 1.41237 & 0.164499 \\ 5.0291 & 3.56641 & 1.37045 & 0.254731 \\ 5.49528 & 3.84046 & 1.531 & 0.261733 \\ 4.68341 & 3.33784 & 1.26642 & 0.242091 \\ 5.01945 & 3.38299 & 1.47247 & 0.202375 \\ 4.42415 & 2.88719 & 1.3539 & 0.150591 \\ 4.88049 & 3.1097 & 1.53817 & 0.144007 \\ 5.38378 & 3.70748 & 1.53257 & 0.240251 \\ 4.90944 & 3.3068 & 1.44142 & 0.197337 \\ 4.74646 & 3.03876 & 1.48736 & 0.144301 \\ 4.2674 & 3.01018 & 1.17241 & 0.21143 \\ 5.63682 & 4.12028 & 1.46322 & 0.321613 \\ 5.75862 & 4.35443 & 1.40883 & 0.371189 \\ 5.34688 & 3.9582 & 1.35841 & 0.319714 \end{bmatrix}$$

η οποία απέχει απόσταση $\sigma_3 = 0.588009$ από την A αν χρησιμοποιήσουμε την 2-norm.

Διάσπαση ιδιαζουσών τιμών (SVD)

Τέλος, μηδενίζοντας τις σ_2, σ_3 και σ_4 και υπολογίζοντας ξανά το γινόμενο $U\Sigma V^t$ προκύπτει ότι η βέλτιστη rank 1 προσέγγιση της A είναι η μήτρα

$$A_1 = \begin{bmatrix} 5.07871 & 3.51403 & 1.43586 & 0.231524 \\ 4.72891 & 3.272 & 1.33697 & 0.215578 \\ 4.67132 & 3.23215 & 1.32068 & 0.212952 \\ 4.59914 & 3.18221 & 1.30028 & 0.209662 \\ 5.05897 & 3.50038 & 1.43028 & 0.230625 \\ 5.50844 & 3.81137 & 1.55736 & 0.251114 \\ 4.71694 & 3.26372 & 1.33358 & 0.215032 \\ 4.98842 & 3.45156 & 1.41034 & 0.227408 \\ 4.36422 & 3.01967 & 1.23386 & 0.198953 \\ 4.78844 & 3.31319 & 1.3538 & 0.218291 \\ 5.37771 & 3.72091 & 1.5204 & 0.245155 \\ 4.8784 & 3.37543 & 1.37923 & 0.222392 \\ 4.66192 & 3.22565 & 1.31803 & 0.212524 \\ 4.28722 & 2.96639 & 1.21209 & 0.195442 \\ 5.71265 & 3.95266 & 1.61509 & 0.260424 \\ 5.88609 & 4.07267 & 1.66413 & 0.26833 \\ 5.43598 & 3.76123 & 1.53687 & 0.247811 \end{bmatrix}$$

η οποία απέχει απόσταση $\sigma_2 = 0.823796$ από την A αν χρησιμοποιήσουμε την 2-norm.

Διάσπαση ιδιαζουσών τιμών (SVD)

Οι εικόνες με γκρι παλέττα και διαστάσεις $m \times n$ μπορούν να θεωρηθούν ως $m \times n$ μήτρες στις οποίες κάθε στοιχείο αντιστοιχεί σε ένα pixel της εικόνας και τα στοιχεία λαμβάνουν τιμές από 0 εως 255 ανάλογα με την αντίστοιχη τιμή του γκρι που έχει το pixel.
Μια εποπτική εφαρμογή του θεωρήματος Eckart - Young είναι η προσέγγιση της μήτρας μιας γκρι εικόνας με μεγάλο rank από μήτρες (και άρα αντίστοιχες εικόνες) με μικρότερο rank.
Να βρεθούν προσεγγίσεις μικροτέρου rank για την εικόνα (unipi_rank310.eps):

Διάσπαση ιδιαζουσών τιμών (SVD)

Λύση. Η εικόνα αυτή έχει διαστάσεις 361×310. Για τους υπολογισμούς θα χρησιμοποιήσουμε το **Octave**. (Το ερωτηματικό στο τέλος των εντολών που ακολουθούν έχει ως συνέπεια να εκτελεστεί η εντολή αλλά να μην εμφανισθεί το αποτέλεσμα της στην οθόνη).
Πρώτα φορτώνουμε την εικόνα με την εντολή

```
A = imread("unipi_rank310.eps");
```

Επειδή η εικόνα είναι ήδη γκρι δεν χρειάζεται να κάνουμε κάποια μετατροπή, αν ήταν έγχρωμη θα την μετατρέπαμε σε γκρι με την εντολή

```
Agray = rgb2gray(A);
```

Πρώτα θα μετατρέψουμε τα στοιχεία της σε double με την εντολή

```
Adouble = im2double(A);
```

Μπορούμε να δούμε το rank της Adouble με την εντολή

```
rank(Adouble)
ans = 310
```

Η απάντηση μας λέει ότι η μήτρα Adouble έχει rank 310.

Διάσπαση ιδιαζουσών τιμών (SVD)

Επίσης, μπορούμε να δούμε την εικόνα που αντιστοιχεί την μήτρα Adouble με την εντολή

```
imshow(Adouble);
```

Έπειτα θα υπολογίσουμε την διάσπαση ιδιαζουσών τιμών (SVD) της Adouble με την εντολή

```
[U, S, V] = svd(Adouble);
```

Αν θέλουμε να υπολογίσουμε την βέλτιστη rank k προσέγγιση της Adouble, μπορούμε να χρησιμοποιήσουμε τις εντολές

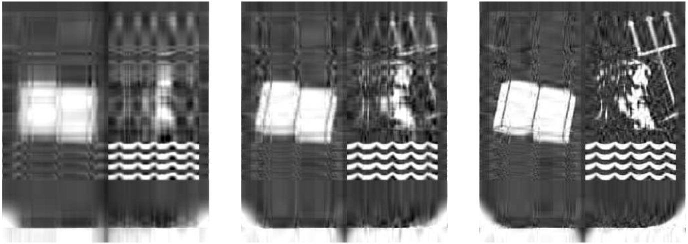
```
C=S; N=k; C(N+1:end,:)=0; C(:,N+1:end)=0; D=U*C*V'; imshow(D)
```

Η D είναι η βέλτιστη rank k προσέγγιση της Adouble.
Για να αποθηκεύσουμε την D σε αρχείο εικόνας μπορούμε να χρησιμοποιήσουμε την εντολή

```
imwrite(D,"mypic.jpg");
```

Ακολουθούν ορισμένες rank k προσεγγίσεις της αρχικής εικόνας:

$k = 5$ $\qquad$ $k = 10$ $\qquad$ $k = 20$



Διάσπαση ιδιαζουσών τιμών (SVD)

$k = 30$

$k = 40$

$k = 50$

$k = 100$

$k = 200$

$k = 310$

Διάσπαση ιδιαζουσών τιμών (SVD)

Παρατήρηση: Μπορούμε να βρούμε προσεγγίσεις μικροτέρου rank και για έγχρωμες εικόνες. Οι έγχρωμες εικόνες με παλέττα rgb αντιστοιχούν σε 3 διαφορετικές μήτρες (μια για κάθε ένα από τα χρώματα της παλέττας red, green και blue).
Για να βρεθούν προσεγγίσεις μικροτέρου rank για την εικόνα (unipic_rank310.eps)

όπως προηγουμένως φορτώνουμε την εικόνα σε μια μήτρα με την εντολή

```
A = imread("unipi_rank310.eps");
```

Διάσπαση ιδιαζουσών τιμών (SVD)

Και πάλι, πρώτα θα μετατρέψουμε τα στοιχεία της σε double

```
Adouble = im2double(A);
```

Στην συνέχεια, θα υπολογίσουμε ξεχωριστά την διάσπαση ιδιαζουσών τιμών (SVD) κάθε χρώματος της Adouble με τις εντολές

```
[U1, S1, V1] = svd(Adouble(:,:,1);
[U2, S2, V2] = svd(Adouble(:,:,2);
[U3, S3, V3] = svd(Adouble(:,:,3);
```

Για να υπολογίσουμε την βέλτιστη rank k προσέγγιση της Adouble, θα υπολογίσουμε την βέλτιστη rank k προσέγγιση κάθε χρώματος με τις εντολές

```
C=S1; N=k; C(N+1:end,:)=0; C(:,N+1:end)=0; D1=U1*C*V1';
C=S2; N=k; C(N+1:end,:)=0; C(:,N+1:end)=0; D2=U2*C*V2';
C=S3; N=k; C(N+1:end,:)=0; C(:,N+1:end)=0; D3=U3*C*V3';
```

Οι $D1$, $D2$, $D3$ είναι οι βέλτιστες rank k προσεγγίσεις κάθε χρώματος της Adouble.

Τέλος, θα ανασυνθέσουμε τις 3 προσεγγίσεις σε μια μήτρα ώστε να προκύψει και πάλι έγχρωμη εικόνα χρησιμοποιώντας τις εντολές:

```
Anew = zeros(size(Adouble));
Anew(:,:,1) = D1;
Anew(:,:,2) = D2;
Anew(:,:,3) = D3;
```

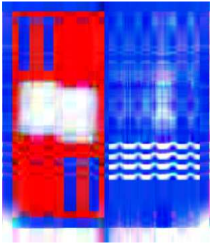
Για να δούμε την εικόνα της προσέγγισης και πάλι χρησιμοποιούμε την εντολή:

```
imshow(Anew);
```

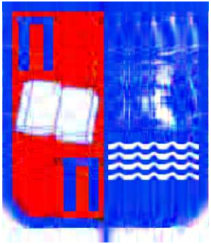
Διάσπαση ιδιαζουσών τιμών (SVD)

Ακολουθούν ορισμένες rank k προσεγγίσεις της αρχικής εικόνας:

$k = 5$



$k = 10$



$k = 20$

Διάσπαση ιδιαζουσών τιμών (SVD)

$k = 30$

$k = 40$

$k = 50$

$k = 100$

$k = 200$

$k = 310$

Παρατήρηση: Μια καλή ιδιότητα της διάσπασης ιδιαζουσών τιμών είναι ότι οι ιδιάζουσες τιμές της A δεν είναι ιδιαίτερα ευαίσθητες σε μικρές αλλαγές των στοιχείων της A. Συγκεριμένα, οι ιδιάζουσες τιμές των A και $A+E$ διαφέρουν το πολύ κατά $\|E\|$ (όπου $\|E\|$ είναι η 2-norm).

6η ΔΙΑΛΕΞΗ
ΕΣΩΤΕΡΙΚΟ ΓΙΝΟΜΕΝΟ

- **Ορθογώνια συμπληρώματα**

Ορθογώνια συμπληρώματα

Έστω $\emptyset \neq A \subseteq V$. Το σύνολο

$$A^{\perp} = \{\mathbf{x} \in V : \langle \mathbf{x}, \mathbf{y} \rangle = 0 \text{ για κάθε } \mathbf{y} \in A\}$$

ονομάζεται **ορθογώνιο συμπλήρωμα του** A.
Στην επόμενη πρόταση συνοψίζονται οι βασικές ιδιότητες των ορθογωνίων συμπληρωμάτων.

Πρόταση 12

Έστω $(\mathcal{V}, \langle\rangle)$ και W υπόχωρος του V. Ισχύουν τα παρακάτω:

1. *$W^{\perp}$ είναι υπόχωρος του V.*
2. *$W \cap W^{\perp} = \{\mathbf{0}\}$.*
3. *Έστω S μια βάση του W. Τότε*

$$W^{\perp} = \{\mathbf{x} \in V : \langle \mathbf{x}, \mathbf{y} \rangle = 0 \text{ για κάθε } \mathbf{y} \in S\}.$$

Πρόταση 12 (συνέχεια)

4. $V = W + W^{\perp}$ *και για κάθε* $\mathbf{v} \in V$ *υπάρχουν μοναδικά* $\mathbf{w} \in W$ *και* $\mathbf{x} \in W^{\perp}$ *ώστε*
$$\mathbf{v} = \mathbf{w} + \mathbf{x}.$$
5. *Έστω* S *μια βάση του* W *και* S' *μια βάση του* $W^{\perp}$. *Το σύνολο* $S \cup S'$ *αποτελεί βάση του* V.
6. *Αν* $\dim V = n$ *και* $\dim W = k$ *τότε* $\dim W^{\perp} = n - k$.
7. $(W^{\perp})^{\perp} = W$.
8. *Έστω* S' *μια βάση του* $W^{\perp}$. *Τότε*
$$W = \{\mathbf{x} \in V : \langle \mathbf{x}, \mathbf{y}' \rangle = 0 \text{ για κάθε } \mathbf{y}' \in S'\}.$$

Απόδειξη:

1. Επειδή $\langle \mathbf{0}, \mathbf{w} \rangle = 0$ για κάθε $\mathbf{w} \in W$ έπεται ότι $W^{\perp} \neq \emptyset$.
   Έστω $\mathbf{x}, \mathbf{y} \in W^{\perp}$ και $\lambda, \mu \in \mathbb{R}$. Ισχύει ότι $\langle \mathbf{x}, \mathbf{w} \rangle = \langle \mathbf{y}, \mathbf{w} \rangle = 0$ για κάθε $\mathbf{w} \in W$.
   Επομένως, για κάθε $\mathbf{w} \in W$ ισχύει ότι

   $$\langle \lambda \mathbf{x} + \mu \mathbf{y}, w \rangle = \lambda \langle \mathbf{x}, \mathbf{w} \rangle + \mu \langle \mathbf{y}, \mathbf{w} \rangle = 0$$

   άρα, $\lambda \mathbf{x} + \mu \mathbf{y} \in W^{\perp}$, δηλαδή το $W^{\perp}$ είναι υπόχωρος του V.
2. Προφανώς $\mathbf{0} \in W \cap W^{\perp}$. Έστω $\mathbf{x} \in W \cap W^{\perp}$. Επειδή $\mathbf{x} \in W^{\perp}$ έπεται ότι $\langle \mathbf{x}, \mathbf{w} \rangle = 0$ για κάθε $\mathbf{w} \in W$. Όμως, $\mathbf{x} \in W$, οπότε για $\mathbf{w} = \mathbf{x}$ ισχύει ότι

   $$\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}.$$

Ορθογώνια συμπληρώματα

3. Έστω S μια βάση του W. Αν $\mathbf{x} \in W^{\perp}$ τότε για κάθε $\mathbf{y} \in S \subseteq W$ έπεται ότι $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. Αντίστροφα, αν $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ για κάθε $\mathbf{y} \in S$, τότε $\mathbf{x} \in W^{\perp}$. Πράγματι, επειδή κάθε στοιχείο $\mathbf{w} \in W$ εκφράζεται ως γραμμικός συνδυασμός των στοιχείων $\mathbf{y}$ της S έπεται ότι
$$\mathbf{w} = \sum_{\mathbf{y} \in S} \lambda_{\mathbf{y}} \mathbf{y} \text{ όπου } \lambda_{\mathbf{y}} \in \mathbb{R}$$
επομένως
$$\langle \mathbf{x}, \mathbf{w} \rangle = \left\langle \mathbf{x}, \sum_{\mathbf{y} \in S} \lambda_{\mathbf{y}} \mathbf{y} \right\rangle = \sum_{\mathbf{y} \in S} \lambda_{\mathbf{y}} \langle \mathbf{x}, \mathbf{y} \rangle = 0,$$
άρα, $\mathbf{w} \in W^{\perp}$.

Ορθογώνια συμπληρώματα

4. Έστω $\mathbf{v} \in V$ και S μια *ορθοκανονική βάση* του W. Θεωρούμε τα διανύσματα

$$\mathbf{w} = \sum_{\mathbf{y} \in S} \langle \mathbf{v}, \mathbf{y} \rangle \, \mathbf{y} \text{ και } \mathbf{x} = \mathbf{v} - \mathbf{w}.$$

Προφανώς $\mathbf{v} = \mathbf{w} + \mathbf{x}$ και $\mathbf{w} \in W$ αφού το $\mathbf{w}$ είναι γραμμικός συνδυασμός των στοιχείων $\mathbf{y}$ της βάσης του W. Επίσης $\mathbf{x} \in W^{\perp}$. Πράγματι για κάθε $\mathbf{y}_0 \in S$ ισχύει ότι

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y}_0 \rangle = \langle \mathbf{v} - \mathbf{w}, \mathbf{y}_0 \rangle &= \left\langle \mathbf{v} - \sum_{\mathbf{y} \in S} \langle \mathbf{v}, \mathbf{y} \rangle \, \mathbf{y}, \mathbf{y}_0 \right\rangle \\
&= \langle \mathbf{v}, \mathbf{y}_0 \rangle - \left\langle \sum_{\mathbf{y} \in S} \langle \mathbf{v}, \mathbf{y} \rangle \, \mathbf{y}, \mathbf{y}_0 \right\rangle \\
&= \langle \mathbf{v}, \mathbf{y}_0 \rangle - \sum_{\mathbf{y} \in S} \langle \langle \mathbf{v}, \mathbf{y} \rangle \, \mathbf{y}, \mathbf{y}_0 \rangle = \langle \mathbf{v}, \mathbf{y}_0 \rangle - \sum_{\mathbf{y} \in S} \langle \mathbf{v}, \mathbf{y} \rangle \langle \mathbf{y}, \mathbf{y}_0 \rangle \\
&= \langle \mathbf{v}, \mathbf{y}_0 \rangle - \langle \mathbf{v}, \mathbf{y}_0 \rangle \langle \mathbf{y}_0, \mathbf{y}_0 \rangle = 0.
\end{aligned}$$

αφού $\langle \mathbf{y}, \mathbf{y}_0 \rangle = 0$ για κάθε $\mathbf{y} \in S$ με $\mathbf{y} \neq \mathbf{y}_0$ και $\langle \mathbf{y}_0, \mathbf{y}_0 \rangle = 1$. Επομένως, από την προηγούμενη ιδιότητα προκύπτει ότι $\mathbf{x} \in W^{\perp}$.
Έστω ότι $\mathbf{v} = \mathbf{w}' + \mathbf{x}'$ με $\mathbf{w}' \in W$ και $\mathbf{x}' \in W^{\perp}$, τότε $\mathbf{w} + \mathbf{x} = \mathbf{w}' + \mathbf{x}'$ επομένως, $\mathbf{w} - \mathbf{w}' = \mathbf{x}' - \mathbf{x}$, όμως $\mathbf{w} - \mathbf{w}' \in W$ και $\mathbf{x}' - \mathbf{x} \in W^{\perp}$ επομένως, $\mathbf{w} - \mathbf{w}', \mathbf{x}' - \mathbf{x} \in W \cap W^{\perp}$ επομένως, $\mathbf{w} - \mathbf{w}' = \mathbf{0}$ και $\mathbf{x}' - \mathbf{x} = \mathbf{0}$, δηλαδή $\mathbf{w} = \mathbf{w}'$ και $\mathbf{x} = \mathbf{x}'$.

Ορθογώνια συμπληρώματα

5. Άσκηση.
6. Άσκηση.
7. Έστω $\mathbf{w} \in W$ τότε $\langle \mathbf{w}, \mathbf{x} \rangle = 0$ για κάθε $\mathbf{x} \in W^\perp$, άρα $\mathbf{w} \in (W^\perp)^\perp$, οπότε $W \subseteq (W^\perp)^\perp$. Έστω $\mathbf{v} \in (W^\perp)^\perp$ τότε από την ιδιότητα 4 υπάρχουν μοναδικά $\mathbf{w} \in W$ και $\mathbf{x} \in W^\perp$ ώστε
$$\mathbf{v} = \mathbf{w} + \mathbf{x}.$$
Ισχύει ότι
$$0 = \langle \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{w} + \mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{w}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle \Leftrightarrow \mathbf{x} = \mathbf{0}.$$
Επομένως,
$$\mathbf{v} = \mathbf{w} + \mathbf{0} = \mathbf{w} \in W$$
άρα, $(W^\perp)^\perp \subseteq W$, δηλαδή $(W^\perp)^\perp = W$.
8. Άσκηση. $\square$

Ορθογώνια συμπληρώματα

Παράδειγμα: Έστω

$$W = \{(x, y, z) \in \mathbb{R}^3 : x + 2y - z = 0\}.$$

Εύκολα προκύπτει ότι το W είναι υπόχωρος του $\mathbb{R}^3$.
Για κάθε $(x, y, z) \in W$ ισχύει ότι

$$(x, y, z) = (x, y, x + 2y) = x(1, 0, 1) + y(0, 1, 2).$$

Άρα, το σύνολο $\{(1, 0, 1), (0, 1, 2)\}$ αποτελεί βάση του W. Επομένως, $\dim W = 2$.
Το ορθογώνιο συμπλήρωμα του W είναι το σύνολο των διανυσμάτων $(x, y, z) \in \mathbb{R}^3$ για τα οποία

$$\langle (x, y, z), (1, 0, 1) \rangle = 0 \text{ και } \langle (x, y, z), (0, 1, 2) \rangle = 0.$$

ή, ισοδύναμα,

$$\begin{cases} x + z = 0 \\ y + 2z = 0 \end{cases} \Leftrightarrow \begin{cases} x = -z \\ y = -2z. \end{cases}$$

Ορθογώνια συμπληρώματα

Επομένως, για κάθε $(x, y, z) \in W^{\perp}$ ισχύει ότι

$$(x, y, z) = (-z, -2z, z) = z(-1, -2, 1).$$

Άρα, το σύνολο $\{-1, -2, 1\}$ είναι βάση του $W^{\perp}$. Επομένως $\dim W^{\perp} = 1 = 3 - 2 = \dim \mathbb{R}^3 - \dim W$, επαληθεύοντας την Πρόταση 12 6.
Επιπλέον, τα διανύσματα $(1, 0, 1)$, $(0, 1, 2)$, $(-1, -2, 1)$ είναι γραμμικώς ανεξάρτητα αφού

$$\begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ -1 & -2 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 2 \\ -2 & 1 \end{vmatrix} + 1 \cdot \begin{vmatrix} 0 & 1 \\ -1 & -2 \end{vmatrix} = 6 \neq 0.$$

Άρα, αποτελούν βάση του $\mathbb{R}^3$.

Επομένως κάθε διάνυσμα $\mathbf{x} \in \mathbb{R}^3$ εκφράζεται ως

$$\mathbf{x} = \lambda_1(1,0,1) + \lambda_2(0,1,2) + \lambda_3(-1,-2,1).$$

Αν τεθούν

$$\mathbf{y} = \lambda_1(1,0,1) + \lambda_2(0,1,2) \in W,$$
$$\mathbf{y}' = \lambda_3(-1,-2,1) \in W^{\perp},$$

τότε

$$\mathbf{x} = \mathbf{y} + \mathbf{y}',$$

όπου $\mathbf{y} \in W$ και $\mathbf{y} \in W^{\perp}$, επαληθεύοντας την Πρόταση 12 4.
Βάσει της Πρότασης 12 8, το διάνυσμα $(3,3,3)$ δεν ανήκει στον W, διότι

$$\langle (3,3,3), (-1,-2,1) \rangle = 3 \cdot (-1) + 3 \cdot (-2) + 3 \cdot 1 = 6 \neq 0,$$

ενώ, το διάνυσμα $(4,-3,-2)$ ανήκει στον W, διότι

$$\langle (4,-3,-2), (-1,-2,1) \rangle = 4 \cdot (-1) + (-3) \cdot (-2) + (-2) \cdot 1 = 0.$$

Έστω W ο χώρος που παράγουν τα διανύσματα

$$\mathbf{x}_1 = (0, 1, 1, 1, 1), \quad \mathbf{x}_2 = (1, 1, 2, 0, 1), \quad \mathbf{x}_3 = (2, 1, 3, 0, 0).$$

- Δίνεται το διάνυσμα $\mathbf{x} = (3, 0, 1, -2, -1)$. Ανήκει το $\mathbf{x}$ στον W;

1η απάντηση. Το $\mathbf{x}$ ανήκει στο W ανν υπάρχουν λ_1, λ_2, λ_3 ώστε $\mathbf{x} = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \lambda_3\mathbf{x}_3$ (Οπότε για κάθε διάνυσμα $\mathbf{x}$ λύνουμε ένα γραμμικό σύστημα.)
Κάθε φορά που αλλάζει το $\mathbf{x}$ έχουμε ένα νέο σύστημα να λύσουμε.
2η απάντηση. Υπάρχει άλλος τρόπος για να κάνουμε τον έλεγχο αυτό;
Ναι, υπάρχει και άλλος τρόπος, μέσα από το ορθόγωνιο συμπλήρωμα του W.
Ιδέα:

- Θα βρούμε μια βάση του ορθογωνίου συμπληρώματος του W.
- Το $\mathbf{x}$ θα ανήκει στο W ανν είναι ορθογώνιο σε κάθε στοιχείο της βάσης του ορθογωνίου συμπληρώματος του W.

Αρχικά θα βρούμε μια βάση του ορθογωνίου συμπληρώματος του W.
Ένα διάνυσμα $\mathbf{y} = (y_1, y_2, y_3, y_4, y_5)$ ανήκει στο ορθογώνιο συμπλήρωμα του W ανν

$$\begin{cases} \langle \mathbf{y}, \mathbf{x}_1 \rangle = 0 \\ \langle \mathbf{y}, \mathbf{x}_2 \rangle = 0 \\ \langle \mathbf{y}, \mathbf{x}_3 \rangle = 0 \end{cases} \Leftrightarrow \begin{cases} y_2 + y_3 + y_4 + y_5 = 0 \\ y_1 + y_2 + 2y_3 + y_5 = 0 \\ 2y_1 + y_2 + 3y_3 = 0 \end{cases}$$

(Αφού τα $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ παράγουν χώρο διάστασης 3 και βρισκόμαστε σε ένα χώρο διάστασης 5 περιμένουμε ότι το ορθογώνιο συμπλήρωμα θα έχει διάσταση $5 - 3 = 2$, δηλαδή το σύστημα θα έχει 2 ελεύθερους αγνώστους και όλοι οι υπόλοιποι θα εκφράζονται μέσω αυτών.)

Παρατηρώντας τις 3 εξισώσεις βλέπουμε ότι μπορούμε να εκφράσουμε τα y_1, y_4, y_5 μέσω των y_2, y_3.

$$\Leftrightarrow \begin{cases} y_2 + y_3 + y_4 + y_5 = 0 \\ 2y_1 + 2y_2 + 4y_3 + 2y_5 = 0 \Leftrightarrow 2y_5 = -y_2 - y_3 \\ 2y_1 = -y_2 - 3y_3 \end{cases}$$

$$\Leftrightarrow \begin{cases} 2y_2 + 2y_3 + 2y_4 + 2y_5 = 0 \Leftrightarrow 2y_4 = -y_2 - y_3 \\ 2y_5 = -y_2 - y_3 \\ 2y_1 = -y_2 - 3y_3 \end{cases}$$

Άρα, το $\mathbf{y}$ ανήκει στο ορθογώνιο συμπλήρωμα του W ανν έχει την μορφή

$$\begin{aligned} \mathbf{y} &= ((-1/2)(y_2 + 3y_3), y_2, y_3, (-1/2)(y_2 + y_3), (-1/2)(y_2 + y_3)) \\ &= ((-1/2)y_2, y_2, 0, (-1/2)y_2, (-1/2)y_2) \\ &\qquad + ((-3/2)y_3, 0, y_3, (-1/2)y_3, (-1/2)y_3) \\ &= (-1/2)y_2(1, -2, 0, 1, 1) + (-1/2)y_3(3, 0, -2, 1, 1) \end{aligned}$$

Άρα, το $\mathbf{y}$ είναι γραμμικός συνδυασμός των γραμμικώς ανεξάρτητων διανυσμάτων

$$\mathbf{y}_1 = (1, -2, 0, 1, 1) \text{ και } \mathbf{y}_2 = (3, 0, -2, 1, 1)$$

άρα, αποτελούν βάση του ορθογωνίου συμπληρώματος του W.
Είναι το $\mathbf{x} = (3, 0, 1, -2, -1)$ γραμμικός συνδυασμός των $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$;
Για να είναι πρέπει $\langle \mathbf{x}, \mathbf{y}_1 \rangle = 0$ και $\langle \mathbf{x}, \mathbf{y}_2 \rangle = 0$.
$\langle \mathbf{x}, \mathbf{y}_1 \rangle = 1 \cdot 3 + (-2) \cdot 0 + 0 \cdot 1 + 1 \cdot (-2) + 1 \cdot (-1) = 0$
$\langle \mathbf{x}, \mathbf{y}_2 \rangle = 3 \cdot 3 + 0 \cdot 0 + (-2) \cdot 1 + 1 \cdot (-2) + 1 \cdot (-1) = 4 \neq 0$.
Άρα, το $\mathbf{x}$ δεν είναι γραμμικός συνδυασμός των $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$.

- Τι ιδιότητες πρέπει να ικανοποιούν οι συντεταγμένες ενός διανύσματος $\mathbf{z} = (z_1, z_2, z_3, z_4, z_5)$ που είναι γραμμικός συνδυασμός των $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$;

Για να είναι το $\mathbf{z}$ γραμμικός συνδυασμός των $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ πρέπει

$$\begin{cases} \langle \mathbf{z}, \mathbf{y}_1 \rangle = 0 \\ \langle \mathbf{z}, \mathbf{y}_2 \rangle = 0 \end{cases} \Leftrightarrow \begin{cases} z_1 - 2z_2 + z_4 + z_5 = 0 \\ 3z_1 - 2z_3 + z_4 + z_5 = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} z_2 = (1/2)(z_1 + z_4 + z_5) \\ z_3 = (1/2)(3z_1 + z_4 + z_5) \end{cases}$$

Παράδειγμα Αν $\mathbf{z} = (5, 5, 10, 2, 3)$ τότε
$(5 + 2 + 3)/2 = 5 = z_2$
$(3 \cdot 5 + 2 + 3)/2 = 10 = z_3$
Άρα, το $\mathbf{z} = (5, 5, 10, 2, 3)$ είναι γραμμικός συνδυασμός των $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$

Παρατήρηση: Τα $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{y}_1$, $\mathbf{y}_2$ αποτελούν βάση του $\mathbb{R}^5$.