Γραμμική Άλγεβρα Ι

Σημειώσεις

Περιληπτική απόδοση στα ελληνικά αποσπασμάτων του βιβλίου του G. Strang με κάποιες προσθήκες...

Ε. Ιωαννίδης

Οκτώβριος 2022

Πρόλογος

Οι σημειώσεις αυτές **δεν είναι** συγγραφικό πόνημα του «συγγραφέα».

Έκανα απλά μια **περίληψη** --για να βοηθηθούν οι φοιτητές-- των κεφαλαίων του βιβλίου του Strang που μας αφορούν σε αυτό το μάθημα. (Στην αρχή κάθε κεφαλαίου υπάρχουν αναφορές στα αντίστοιχα κεφάλαια του Strang.)

Οι ασκήσεις των Φυλλαδίων είναι και αυτές στο μεγαλύτερό τους μέρος από το βιβλίο του Strang.

Ευχαριστώ τον Τάσο Πλατανιώτη και τη Μαρία Χονδροκούκη για την δακτυλογράφηση των σημειώσεων και των ασκήσεων.

Ε. Ιωαννίδης

Περιγραφή

Το μάθημα κατ' αρχάς γενικεύει τα διανύσματα του επιπέδου σε διανύσματα n διαστάσεων. Αυτά αποτελούν τον $\mathbb{R}^n$. Μελετάμε υποσύνολα αυτού του χώρου (υπόχωρους) και γραμμικές απεικονίσεις μεταξύ υποχώρων. Αυτές εκφράζονται με πίνακες, οι οποίοι επίσης αποτελούν αντικείμενο εκτενούς μελέτης στο μάθημα. Τέλος επικεντρωνόμαστε σε κάποια από τα παραπάνω ζητήματα που έχουν σημαντικές εφαρμογές στη στατιστική: προβολές και ελάχιστα τετράγωνα, καθώς και ιδιοτιμές/ιδιοδιανύσματα πινάκων και τετραγωνικές μορφές.

Σκοπός

Σκοπός του μαθήματος είναι να γίνουν σε βάθος κατανοητές οι παραπάνω έννοιες, να αποκτήσουν οι φοιτητές στοιχεία μιας γεωμετρικής εποπτείας εννοιών όπως η προβολή και να είναι σε θέση να τις εφαρμόζουν πρακτικά.

Συνοπτικά Περιεχόμενα

- Ο $\mathbb{R}^n$ και οι εξισώσεις της ευθείας και του επιπέδου, γραμμικά συστήματα, πίνακες και ο αλγόριθμος Gauss (4 βδομάδες).
- Υπόχωροι, θεμελιώσεις υπόχωροι ενός πίνακα, γραμμικοί μετασχηματισμοί (3 βδομάδες).
- Ορθογωνιότητα, προβολές και ελάχιστα τετράγωνα (2 βδομάδες).

Γραμμική Άλγεβρα: Τι είναι

- Το μάθημα γενικεύει τα διανύσματα του επιπέδου σε **διανύσματα n διαστάσεων**. Αυτά αποτελούν τον $\mathbb{R}^n$.
- Μελετάμε **υποσύνολα** αυτού του χώρου (υπόχωρους) και **γραμμικές απεικονίσεις** μεταξύ υποχώρων.
- Αυτές εκφράζονται με **πίνακες**, οι οποίοι επίσης αποτελούν αντικείμενο εκτενούς μελέτης στο μάθημα.
- Τέλος επικεντρωνόμαστε σε **προβολές και ελάχιστα τετράγωνα** που έχουν σημαντικές εφαρμογές στη στατιστική.

Γραμμική Αλγεβρα: που χρειάζεται στη Στατιστική;

- Η Στατιστική μελετάει **παρατηρήσεις** από μια ή περισσότερες ποσότητες (μεταβλητές) που παρήχθησαν από ένα υποκείμενο μηχανισμό που ενέχει **τυχαιότητα** (τυχαίες μεταβλητές)
- Οι **σχέσεις** μεταξύ τέτοιων μεταβλητών δίνονται από **μοντέλα** που στη περιγραφή και ανάλυσή τους η Γραμμική Αλγεβρα αποδεικνύεται ιδιαίτερα χρήσιμη (παλιδρόμηση, πολυμεταβλητή ανάλυση)

Ιδιοδιανύσματα
ιδιοτιμές
(Γρ.Α. ΙΙ).

- Στη προσπάθεια του Στατιστικού να ανακτήσει από τα δεδομένα πληροφορία για τις υποκείμενες σχέσεις μεταξύ μεταβλητών (παλινδρόμηση, γραμμικό ,μοντέλο) σημαντικό ρόλο παίζουν οι **εκτιμήτριες ελαχίστων τετραγώνων** που προκύπτουν από μια **προβολή σε έναν υπόχωρο**.

Διδασκαλία μαθήματος

- Διδασκαλία 2 δίωρα τη βδομάδα (**δια ζώσης**)
- Φροντιστήριο ένα δίωρο τη βδομάδα (???) (επίλυση ασκήσεων). Πρόσθετες ελάχιστες ασκήσεις στο μάθημα
- Προαπαιτούμενες γνώσεις: Διανύσματα στο επίπεδο
- e-class **εγγραφή στο** (προσοχή: υπάρχουν περισσότερα μαθήματα με τίτλο Γραμμική Άλγεβρα
    - «Γραμμική Άλγεβρα Ι (STAT207) , Εκπαιδευτής Evagelos Ioannidis»).
- Σημειώσεις και διαδικτυακό υλικό
- Επικοινωνία

    Γραφείο: Κοδριγκτώνος 12, 3ος όροφος, Τηλ. 210 8203545, email: eioannid@aueb.gr

    Ώρες γραφείου: Πέμπτη και Παρασκευη 11.00-12.00

- Ερωτήσεις
- Δυσκολίες μαθήματος:
    - **Ταχύτητα** σε σχέση με λύκειο.
    - Τα μαθηματικά **αργούν** να γίνουν κτήμα. Θέλουν παρακολούθηση και ενασχόληση. Δεν αρκούν λίγες μέρες διάβασμα στο τέλος!
    - Στο δεύτερο μισό που έρχονται **τα πιο σημαντικά** για τη στατιστική υπαχει αυξημένη κούραση (μαζί με τη θεωρητική δυσκολία)
- Δυσκολίες σπουδών γενικότερα:
    - **Ο καθένας έχει πια μόνος του την ευθύνη της καθοδήγησης της διαδικασίας μάθησης**

Αναλυτικά Περιεχόμενα μαθήματος

Στοιχεία και πράξεις στον $\mathbb{R}^n$, ευθείες και επίπεδα στον $\mathbb{R}^n$. 2 εβδ

Πίνακες και πολλαπλασιασμός πινάκων, στοιχειώδεις πίνακες.

Γραμμικά συστήματα: απαλοιφή Gauss και η παραγοντοποίηση PA=LDU.

Αντίστροφοι και ανάστροφοι πίνακες, αλγόριθμος Gauss-Jordan. 3-3,5 εβδ

Συμμετρικοί πίνακες και η παραγοντοποίηση Cholesky.

Διανυσματικοί χώροι και υπόχωροι.

Γραμμικά συστήματα: λύση m εξισώσεων με n αγνώστους και τάξη πίνακα.

Γραμμική ανεξαρτησία, βάσεις και διάσταση. 3,5-4

Οι 4 θεμελιώδεις υπόχωροι ενός πίνακα.Θεμελιώδες Θεώρημα της Γραμμικής Άλγεβρας. εβδ

Γραμμικοί μετασχηματισμοί του $\mathbb{R}^n$ και πίνακες.

Ορθογώνιοι υπόχωροι, ορθογώνιο συμπλήρωμα υπόχωρου. 3 εβδ

Προβολές και προσεγγίσεις ελάχιστων τετραγώνων.

Προτεινόμενα συγγράμματα

1) STRANG G., ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ ΚΑΙ ΕΦΑΡΜΟΓΕΣ STRANG G.

2) Lipschutz Seymour, Lipson Marc Lars, Γραμμική Άλγεβρα

3) Δονάτος Γεώργιος Σ., Αδάμ Μαρία Χ., Γραμμική Άλγεβρα

Βιβλιογραφία

- Graybill, F. A. (1969), *Introduction to Matrices with Applications in Statistics*, Wadsworth, Belmont, CA. προχωρημένα

- Healy, M.J.R. (1995), *Matrices for Statistics*, Oxford University Press.

- Harville D.A. (1997), Matrix Algebra from a Statistician's Perspective

- Lay, D. (2011), Linear Algebra and its Applications εισαγωγικό.

- Puntanen S., Styan G.P.H., Isotalo J. (2011), Matrix Tricks for Linear Statistical Models

- Searle, S. R. (1982), *Matrix Algebra Useful for Statistics,* Wiley.

- Ε. Ξεκαλάκη & Ι. Πανάρετος (1993), *Γραμμική Άλγεβρα για Στατιστικές Εφαρμογές*, Αθήνα.

- Η. Φλυτζάνης, Γραμμική Αλγεβρα και Εφαρμογές

Μέθοδοι αξιολόγησης

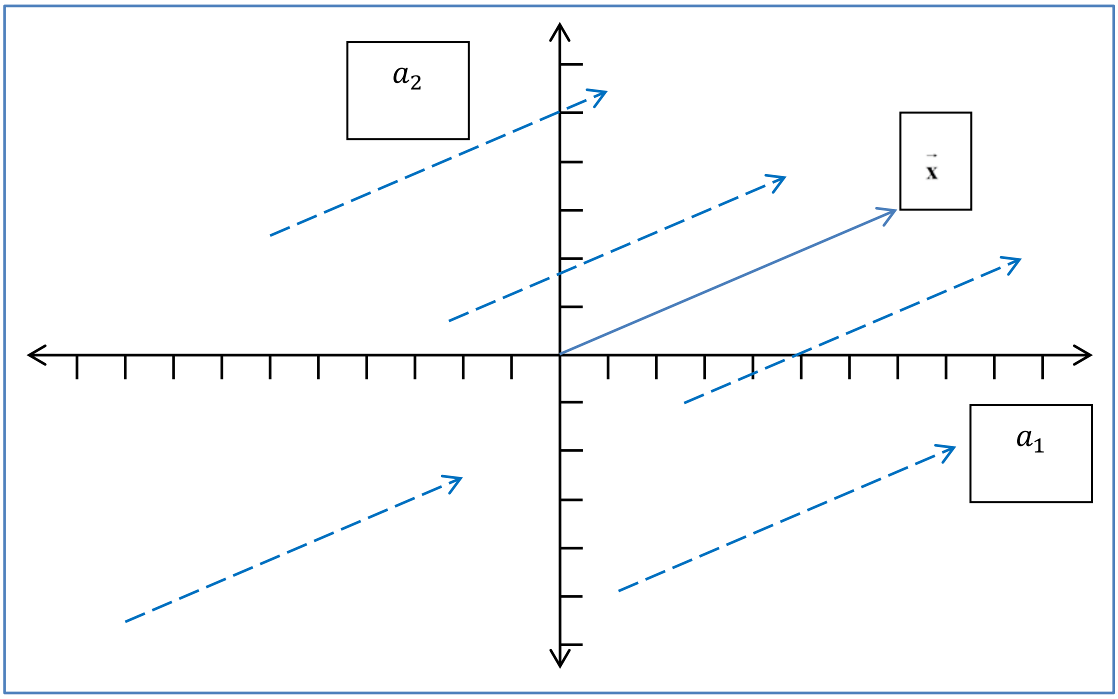
Στο τέλος του εξάμηνου γίνεται Γραπτή εξέταση εφ' όλης της ύλης.

Αξιολογούνται η κατανόηση των εννοιών και η δυνατότητα εφαρμογής και χειρισμού τους σε ασκήσεις.

1. Στοιχεία και πράξεις στον R^n

Διανύσματα στο επίπεδο και ο R^2

- Ας θεωρήσουμε **το επίπεδο** και ένα **σύστημα συντεταγμένων** με άξονες a_1, a_2

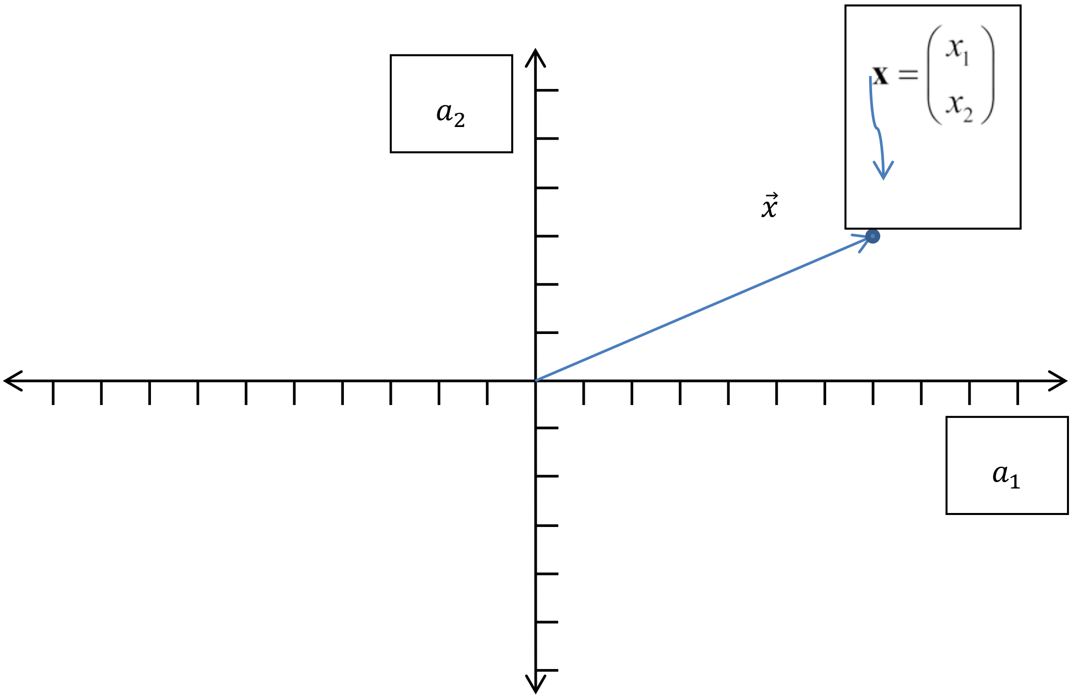


- καθώς και τα διανύσματα $\vec{\mathbf{x}}$ στο επίπεδο.

- **διανύσματα που προκύπτουν το ένα από το άλλο με παράλληλη μετατόπιση**, έχουν δηλαδή ίδια φορά και μήκος, δε διακρίνονται αλλά **θεωρούνται ένα και το αυτό διάνυσμα**.

- **διαλέγουμε ένα διάνυσμα ως «κανονικό εκπρόσωπο»** όλων όσων ταυτίζονται με αυτό, και αυτό θα είναι το διάνυσμα που έχει **ως σημείο εκκίνησης την αρχή των αξόνων**.

- Για να προσδιορίσουμε το διάνυσμα $\vec{\mathbf{x}}$ **αρκεί λοιπόν να γνωρίζουμε τις συντεταγμένες του σημείου στο οποίο καταλήγει** το διάνυσμα, σημείο το οποίο ονομάζουμε $\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.



- Το σύνολο των σημείων $\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ του επιπέδου, ταυτίζεται με το σύνολο των διατεταγμένων δυάδων πραγματικών αριθμών, και συμβολίζεται με $\mathbb{R}^2$.

Σε κάθε διάνυσμα (που ξεκινάει από το 0) λοιπόν αντιστοιχεί ένα σημείο του επιπέδου και σε κάθε σημείο του επιπέδου ένα διάνυσμα (το διάνυσμα που πάει από την αρχή των αξόνων στο σημείο αυτό).

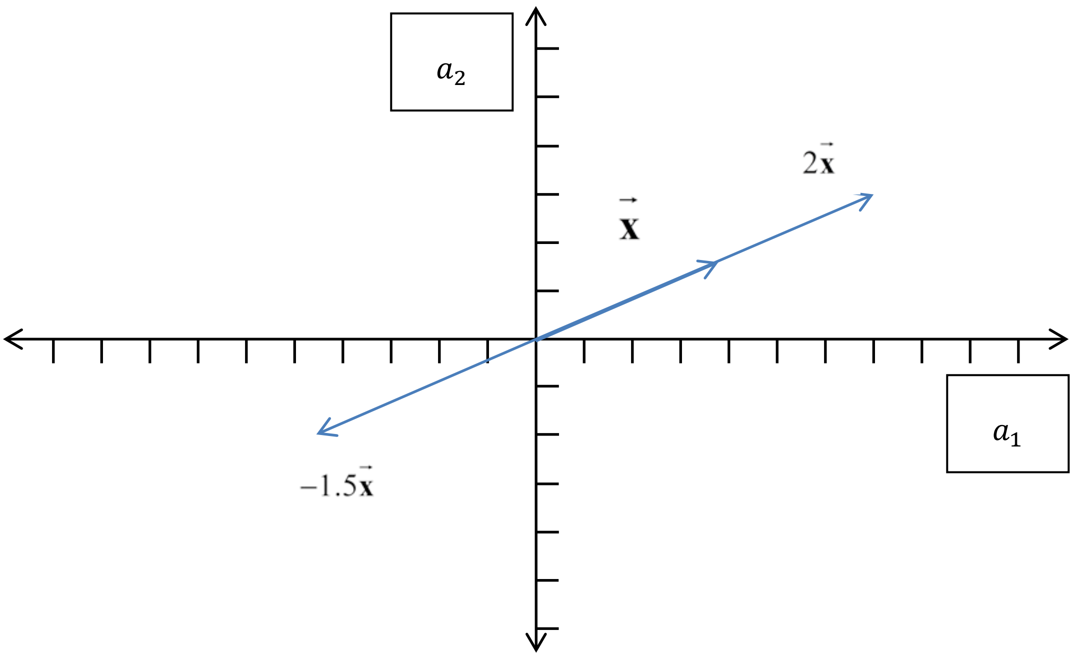
Επίσης κάθε **σημείο του επιπέδου** ορίζεται με μια **δυάδα πραγματικών αριθμών**, τις **δυο συντεταγμένες** του σημείου.

Πράξεις μεταξύ διανυσμάτων και μεταξύ σημείων

Διανύσματα

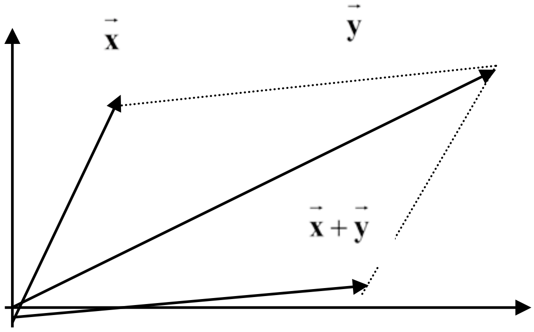
Στα **διανύσματα** έχω δύο πράξεις:

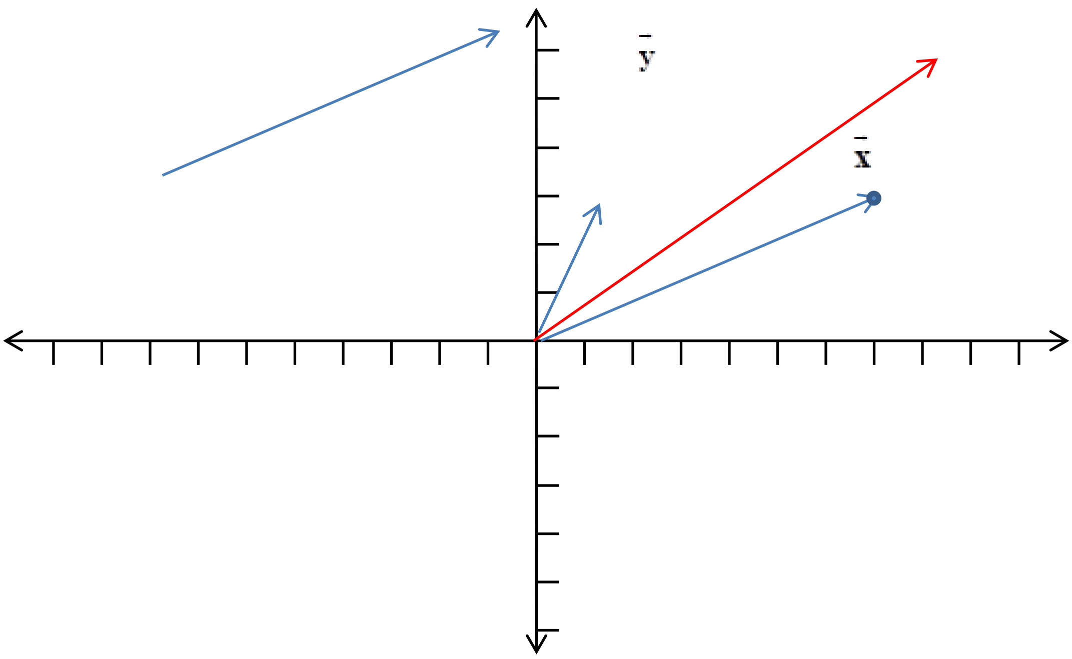
- **πολλαπλασιασμό διανύσματος με πραγματικό αριθμό**, όπου το $\lambda\vec{\mathbf{x}}$ ορίζεται ως το διάνυσμα με (λ)-πλάσιο μήκος και ίδια φορά με το $\vec{\mathbf{x}}$ αν $\lambda > 0$ και αντίθετη αν $\lambda < 0$.



https://www.geogebra.org/m/HYZXHadK

- **πρόσθεση μεταξύ διανυσμάτων** όπου το $\vec{x}+\vec{y}$ προκύπτει από τον κανόνα του παραλληλογράμμου:





Vector Addition https://www.geogebra.org/m/Cy8bxaKS

Vector Addition https://www.geogebra.org/m/dfUmKFZ7

Vector Subtraction https://www.geogebra.org/m/vUAFWvmk

Σημεία του επίπεδου

Στα **σημεία του επιπέδου ή καλύτερα στα στοιχεία του** $\mathbb{R}^2$ μπορώ να ορίσω δύο αντίστοιχες πράξεις

και επιπλέον να ορίσω ότι τα σημεία του επιπέδου εξοπλισμένα με αυτές τις πράξεις αποτελούν τον $\mathbb{R}^2$:

Ορισμός: $\mathbb{R}^2$ είναι το σύνολο των διαταγμένες δυάδων $\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ πραγματικών αριθμών $x_1, x_2 \in \mathbb{R}$,

[με σύμβολα $R^2 := \left\{ \mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| x_1 \in \mathbb{R}, x_2 \in \mathbb{R} \right\}$]

εφοδιασμένων με τις παρακάτω πράξεις:

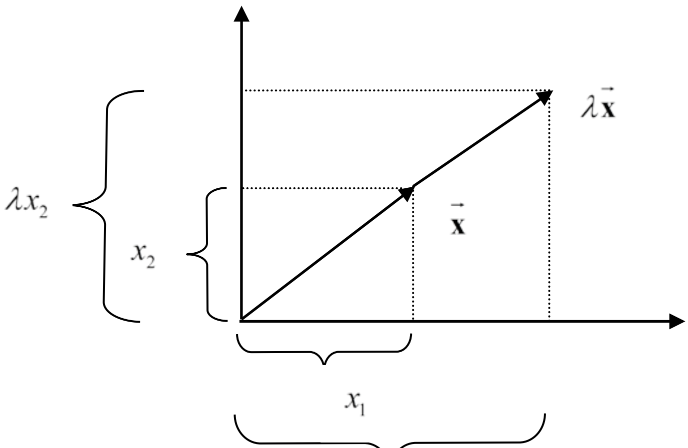
- ***πρόσθεση***: Αν $\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ και $\mathbf{y}=\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ ορίζω το άθροισμα $\mathbf{x}+\mathbf{y}$ ως $\boxed{\mathbf{x}+\mathbf{y} := \begin{pmatrix} x_1+y_1 \\ x_2+y_2 \end{pmatrix}} \in \mathbb{R}^2$.

- ***πολλαπλασιασμό με πραγματικό***: Αν $\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ και $\lambda \in \mathbb{R}$ τότε ορίζω το γινόμενο $\lambda\mathbf{x}$ ως $\boxed{\lambda\mathbf{x} := \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}} \in \mathbb{R}^2$, δηλ. πολλαπλασιάζω και προσθέτω <u>κατά συντεταγμένη</u>.

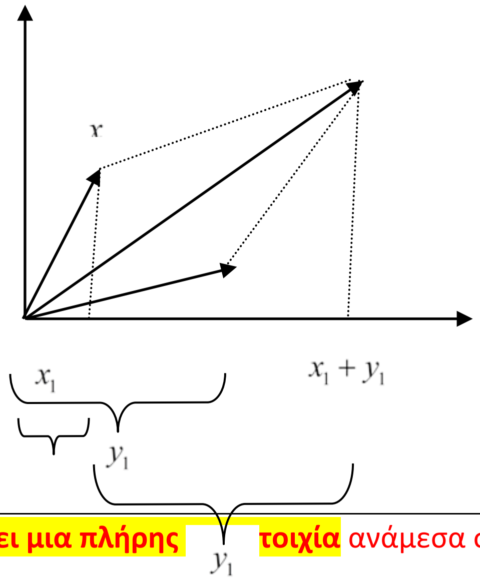
Οι πράξεις στον $\mathbb{R}^2$ έχουν οριστεί έτσι ώστε **η αντιστοιχία που είδαμε μεταξύ διανυσμάτων και στοιχείων** του $\mathbb{R}^2$ (σημείων του επιπέδου) να **διατηρείται και στα αποτελέσματα των εκάστοτε μεταξύ τους πράξεων**.

Δηλαδή:

- αν $\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ το σημείο που αντιστοιχεί στο διάνυσμα $\vec{\mathbf{x}}$, τότε το $\lambda\mathbf{x}=\begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}$ είναι το σημείο που αντιστοιχεί στο διάνυσμα $\lambda\vec{\mathbf{x}}$:



- και επίσης: το σημείο που αντιστοιχεί στο διάνυσμα $\vec{\mathbf{x}}+\vec{\mathbf{y}}$ (που προκύπτει από τον κανόνα του παραλληλογράμμου) είναι ακριβώς το $\begin{pmatrix} x_1+y_1 \\ x_2+y_2 \end{pmatrix}$.



Αφού λοιπόν **υπάρχει μια πλήρης αντιστοιχία** ανάμεσα στα διανύσματα και στα σημεία του επιπέδου (στοιχεία του $\mathbb{R}^2$) καθώς και στο τι μπορώ να κάνω με αυτά (πράξεις), **μπορώ να τα ταυτίσω, δηλ.** να τα βλέπω σαν δύο όψεις του ίδιου πράγματος, **να μη διακρίνω πια μεταξύ διανυσμάτων και R^2.**

- Οι δύο όψεις είναι η <u>γεωμετρική</u> (διανύσματα) και η <u>αλγεβρική</u> (στοιχεία του R^2), δηλ. αυτή στην οποία κάνω πράξεις με σύμβολα και αριθμούς.
    - Η **<u>γεωμετρική θεώρηση</u>** των πραγμάτων με **βοηθάει να φανταστώ καλύτερα ορισμένες ιδιότητες** ή καταστάσεις και έτσι να τις κατανοώ καλύτερα.
    - Η **<u>αλγεβρική θεώρηση</u>** όμως έχει το πλεονέκτημα ότι **μπορώ <u>πιο εύκολα να γενικεύσω</u>** και να επεκτείνω αυτά που κάνουμε με διανύσματα σε χώρους περισσότερων διαστάσεων, τον R^n, στους οποίους δεν έχω γεωμετρική εποπτεία (δεν μπορώ να φανταστώ πάνω από 3 διαστάσεις).

Ο $\mathbb{R}^n$

Η αλγεβρική θεώρηση μας επιτρέπει άμεσα να ορίσουμε τέτοιους χώρους κατ' αναλογία του ορισμού του R^2:

Ορισμός: Για οποιοδήποτε φυσικό αριθμό $n \in N$ ορίζουμε τον $\mathbb{R}^n$ ως το σύνολο των (διατεταγμένων) *n*-αδων πραγματικών αριθμών

$$\mathbb{R}^n := \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \middle| \, x_1, \cdots, x_n \in \mathbb{R} \right\}$$

συντεταγμένες του $\mathbf{x}$

(σύμβαση: γράφουμε αυτές τις *n*-αδες ως στήλες)

- πρόσθεση: Για $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ και $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ ορίζω το άθροισμα $\mathbf{x} + \mathbf{y}$ ως εκείνο το στοιχείο του $\mathbb{R}^n$ που έχει ως συντεταγμένες το άθροισμα των συντεταγμένων των $\mathbf{x}$ και $\mathbf{y}$:

$$\mathbf{x} + \mathbf{y} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

- πολλαπλασιασμός με πραγματικό: Αν $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ και $\lambda \in R$

τότε $\lambda \cdot \mathbf{x} := \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}$.

Προσθέτουμε και πολλαπλασιάζουμε κατά συντεταγμένες.

Ποιος είναι ο $\mathbb{R}^3$?

3D Vector visualization

https://www.geogebra.org/m/bDFcHaUt

sum of 2 vectors in 3d

https://www.geogebra.org/m/uMWZCwzs

Κανόνες πράξεων στον $\mathbb{R}^n$

Έστω $\mathbf{x},\mathbf{y},\mathbf{z} \in \mathbb{R}^n$ και $\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n$. Επίσης για $\mathbf{x} \in \mathbb{R}^n$ ορίζουμε $\mathbf{-x} := \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix}$.

Έχουμε:

- $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$ (αντιμεταθετική)
- $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$ (προσεταιριστική)
- $\mathbf{x}+\mathbf{0} = \mathbf{x}$ (υπάρχει μηδέν της πρόσθεσης)
- $\mathbf{x}+(\mathbf{-x}) = \mathbf{0}$ (υπάρχει αντίθετος)
- $\lambda\cdot(\mu\cdot\mathbf{x}) = (\lambda\cdot\mu)\cdot\mathbf{x}$
- $\lambda(\mathbf{x}+\mathbf{y}) = \lambda\mathbf{x}+\lambda\mathbf{y}$ (επιμεριστική για + διανυσμ.)
- $(\lambda+\mu)\mathbf{x} = \lambda\mathbf{x}+\mu\mathbf{x}$ (επιμεριστική για + βαθμωτών)

Οι αποδείξεις είναι εύκολες: περνάμε σε συντεταγμένες και εκεί χρησιμοποιούμε αντίστοιχες ιδιότητες πραγματικών αριθμών.
π.χ. τελευταία:

$$(\lambda+\mu)\mathbf{x} = \begin{pmatrix} (\lambda+\mu)x_1 \\ \vdots \\ (\lambda+\mu)x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1+\mu x_1 \\ \vdots \\ \lambda x_n+\mu x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix} + \begin{pmatrix} \mu x_1 \\ \vdots \\ \mu x_n \end{pmatrix} = \lambda\mathbf{x}+\mu\mathbf{x}$$

2. Ευθείες και επίπεδα στον R^n

Ευθεία στον R^2

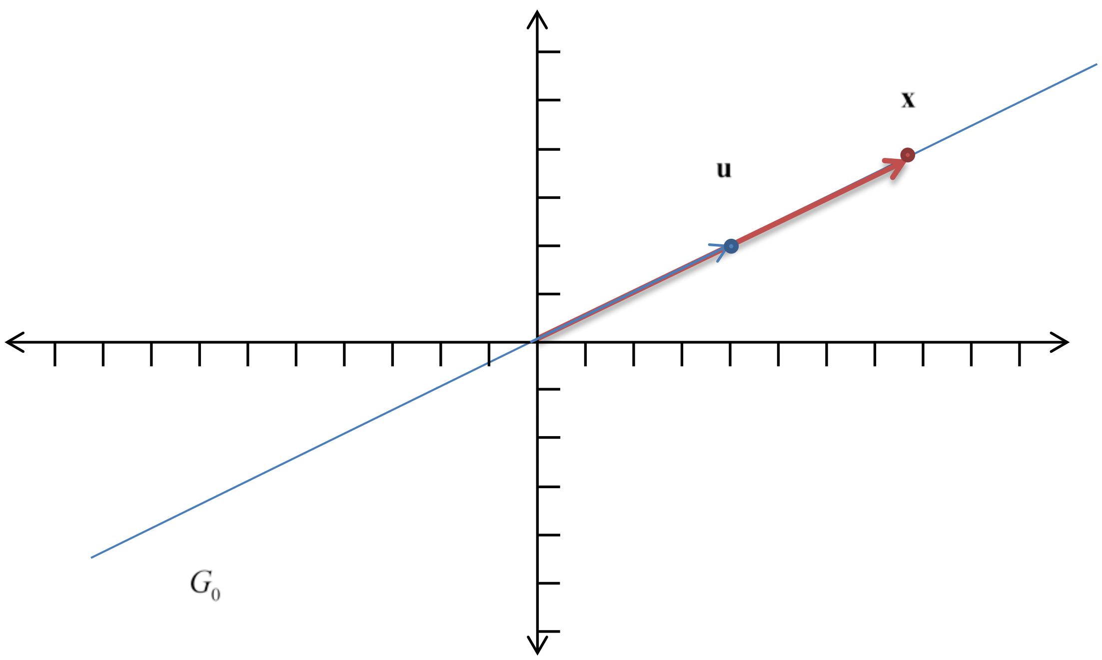
- Η ευθεία είναι ένα σύνολο σημείων.

Θέλουμε να περιγράψουμε αυτό το σύνολο με τη μορφή μιας εξίσωσης σημείων/διανυσμάτων. Να πούμε δηλαδή ότι η **ευθεία είναι όλα εκείνα τα σημεία του επιπέδου που ικανοποιούν την τάδε εξίσωση** (σημείων/διανυσμάτων)

Ευθεία που διέρχεται από το μηδέν.

- Για να ορίσω μια ευθεία G_0 που διέρχεται από την αρχή των αξόνων χρειάζεται άλλο ένα προκαθορισμένο σημείο $\mathbf{u} \in R^2$ από το οποίο να διέρχεται η ευθεία.

- Όλα τα υπόλοιπα σημεία x της ευθείας θα είναι ακριβώς **εκείνα τα $\mathbf{x}$ που ???**



x

- Άρα για να είναι ευθεία που διέρχεται από το **0** και το $\mathbf{u}$ η G_0 πρέπει να είναι το σύνολο των σημείων $\mathbf{x} \in R^2$ που είναι πολλαπλάσια του u, δηλαδή

$$G_0 = \left\{ \mathbf{x} \in R^2 \middle| \exists \lambda : \mathbf{x} = \lambda \mathbf{u} \right\}$$

https://www.geogebra.org/m/KnFAZ8fa

https://www.geogebra.org/m/skj88vSN

- Σε αυτή την εξίσωση το ρόλο του $\mathbf{u}$ μπορεί να τον παίξει και οποιοδήποτε άλλο (αυθαίρετα επιλεγμένο) σημείο της ευθείας:

  Δηλαδή, αν $\mathbf{u}' \in G_0$ και άρα $\mathbf{u}' = \mu \mathbf{u}$ για κατάλληλο μ, τότε

$$G_0 = \left\{ \mathbf{x} \in R^2 \middle| \exists \kappa : \mathbf{x} = \kappa \mathbf{u}' \right\}$$

- Το σύνολο των πολλαπλασίων ενός στοιχείου / διανύσματος / σημείου $\mathbf{u} \in R^2$ λέγεται γραμμική θήκη του $\mathbf{u}$ και το γράφουμε και ως $\mathrm{Span}(\mathbf{u})$. Δηλαδή

  $$G_0 = \left\{ \mathbf{x} \in R^2 \middle| \exists \lambda : \mathbf{x} = \lambda \mathbf{u} \right\} = Span(\mathbf{u})$$

  Δέστε ότι για κάθε συγκεκριμένη τιμή του λ παίρνω και ένα άλλο συγκεκριμένο σημείο $\mathbf{x}$ της G_0. Όταν αφήσω το λ να «τρέχει» στους πραγματικούς αριθμούς τότε το $\mathbf{x}$ τρέχει πάνω στη G_0.

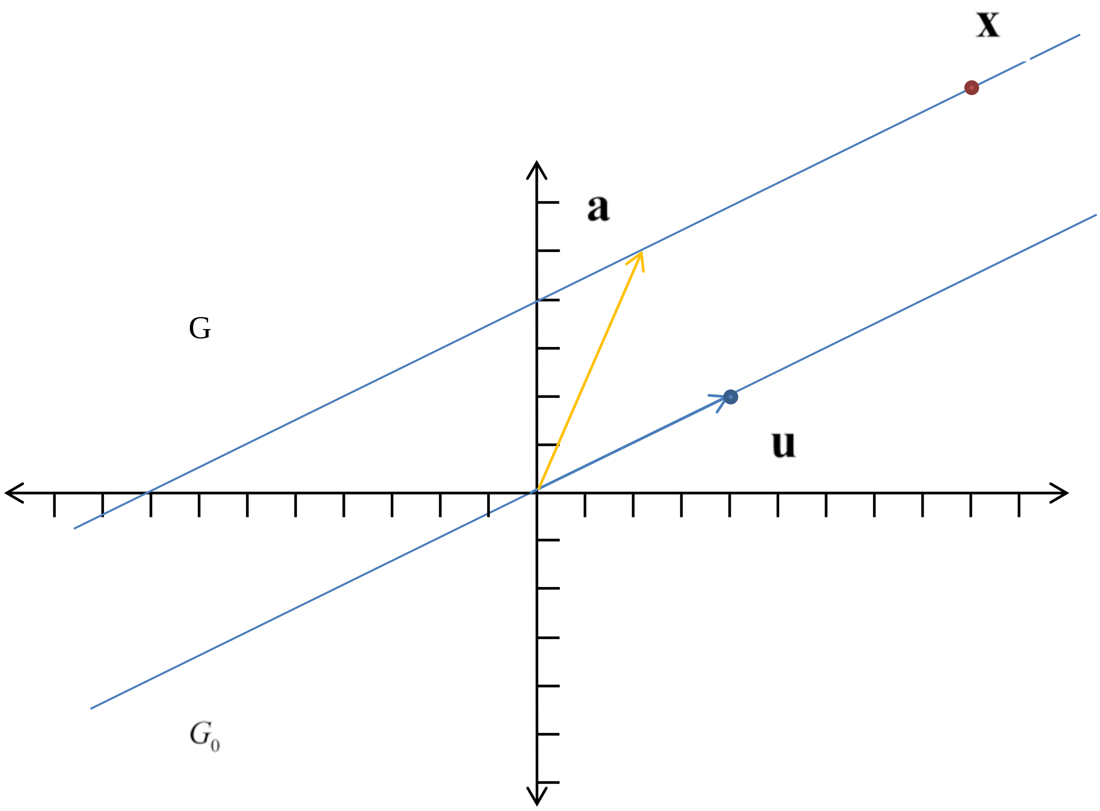
https://www.geogebra.org/m/skj88vSN

- Σημειώστε ότι, όπως και προηγουμένως, αν $\mathbf{u}' \in Span(\mathbf{u})$ και άρα $\mathbf{u}' = \mu \mathbf{u}$ για κατάλληλο μ, τότε

  $$Span(\mathbf{u}) = Span(\mathbf{u}')$$

Ευθεία που δε διέρχεται από μηδέν .

- Ψάχνω εξίσωση που ορίζει την ευθεία G που διέρχεται από κάποιο προκαθορισμένο σημείο $\mathbf{a} \in R^2$ και είναι παράλληλη κάποιας G_0 που διέρχεται από την αρχή των αξόνων.



- Τότε όλα τα υπόλοιπα σημεία x της ευθείας G θα είναι ακριβώς **εκείνα τα $\mathbf{x}$ που γράφονται ως άθροισμα κάποιο σημείου $\mathbf{x}_0 \in G_0$ με το $\mathbf{a}$**. Άρα

$$G = \left\{ \mathbf{x} \in R^2 \middle| \exists \mathbf{x_0} \in G_0 : \mathbf{x} = \mathbf{a} + \mathbf{x_0} \right\}$$

- Καθώς τα σημεία της G_0 είναι τα πολλαπλάσια του $\mathbf{u} \in R^2$ μπορούμε να αντικαταστήσουμε αυτή την ιδιότητα στη παραπάνω εξίσωση:

$$G = \left\{ \mathbf{x} \in R^2 \middle| \exists \lambda \in R : \mathbf{x} = \mathbf{a} + \lambda \mathbf{u} \right\}$$

https://www.geogebra.org/m/KnFAZ8fa

https://www.geogebra.org/m/skj88vSN

- Τα παραπάνω μπορούν να γραφτούν και ως εξίσωση συνόλων

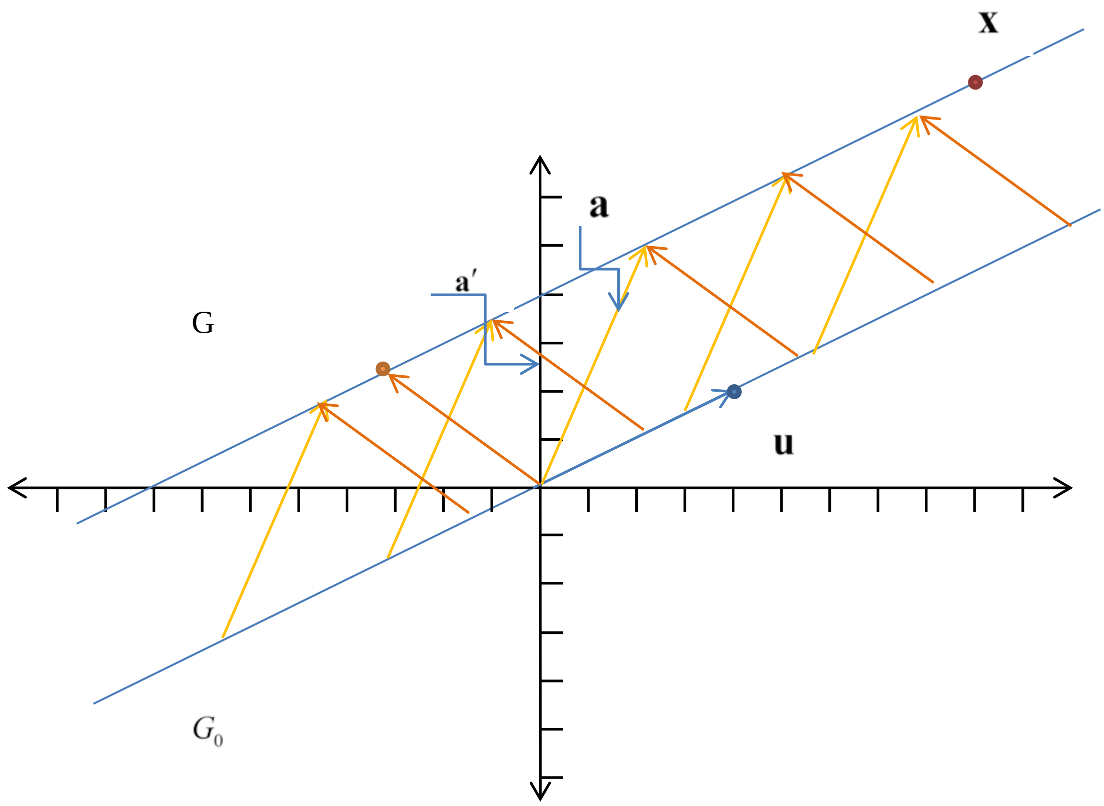
$$G = \mathbf{a} + G_0 = \mathbf{a} + Span(\mathbf{u})$$

Δηλαδή: κάθε σημείο της G προκύπτει ως άθροισμα κάποιου σημείου της G_0 με το $\mathbf{a}$. Αυτό σημαίνει ότι **η G προκύπτει από μετατόπιση της G_0.**

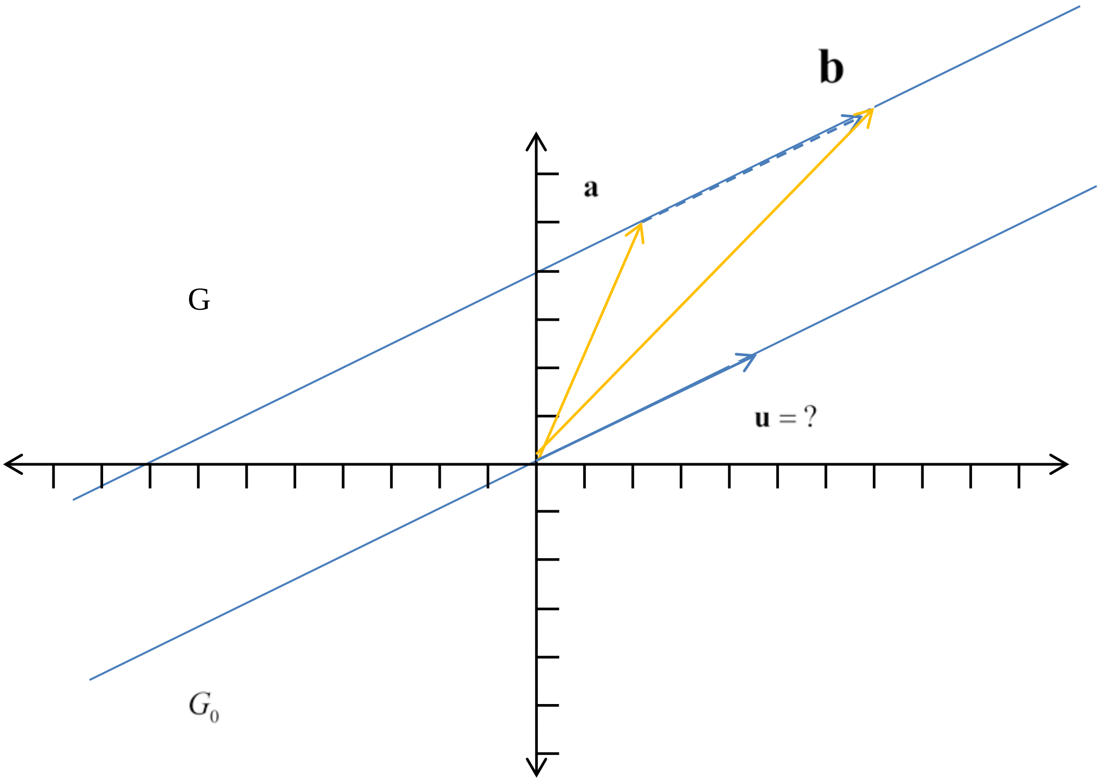
https://www.geogebra.org/m/ZkzF7UmX

- Το ρόλο του $\mathbf{a}$ σε αυτές τις εξισώσεις μπορεί να τον παίξει οποιοδήποτε άλλο στοιχείο $\mathbf{a}' \in G$. Δηλαδή έχουμε

$$G = \mathbf{a}' + G_0 \text{ για οποιοδήποτε } \mathbf{a}' \in G.$$



- Εάν προτιμάμε τώρα να προδιορίσουμε εναλλακτικά την ευθεία G ως εκείνη που διέρχεται από δύο σημεία $\mathbf{a}, \mathbf{b} \in R^2$.
- τότε αυτή θα πρέπει να είναι παράλληλη της ευθείας που διέρχεται από το 0 και το $\mathbf{u} = ??$



- Άρα:

$$G = \left\{ x \in R^2 \middle| \exists \lambda : \mathbf{x} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) \right\}$$

- Αν προτιμώ μπορώ να πάρω σε αυτό τον ορισμό οποιαδήποτε δύο (μη ταυτόσημα) σημεία της G $\mathbf{a}', \mathbf{b}'$ στη θέση των $\mathbf{a}, \mathbf{b}$

Ευθεία στον R^n

Το υποσύνολο $G \subset R^n$ ονομάζεται ευθεία αν υπάρχουν $\mathbf{a},\mathbf{b} \in R^n$ τέτοια ώστε όλα τα σημεία $\mathbf{x}$ της G να γράφονται ως $\mathbf{x} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a})$, για κάποιο λ.

$$G = \left\{ \mathbf{x} \in R^n \middle| \exists \lambda \in R : \mathbf{x} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) \right\}$$

- γράφω και

$$\begin{aligned} G &= \{\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}), \lambda \in R\} \\ &= \{\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a})\} \\ &= \mathbf{a} + \{\lambda(\mathbf{b} - \mathbf{a}), \lambda \in R\} \\ &= \mathbf{a} + Span(\mathbf{b} - \mathbf{a}) \end{aligned}$$

Π.χ. στον R^3

https://www.geogebra.org/m/sb2cedpp

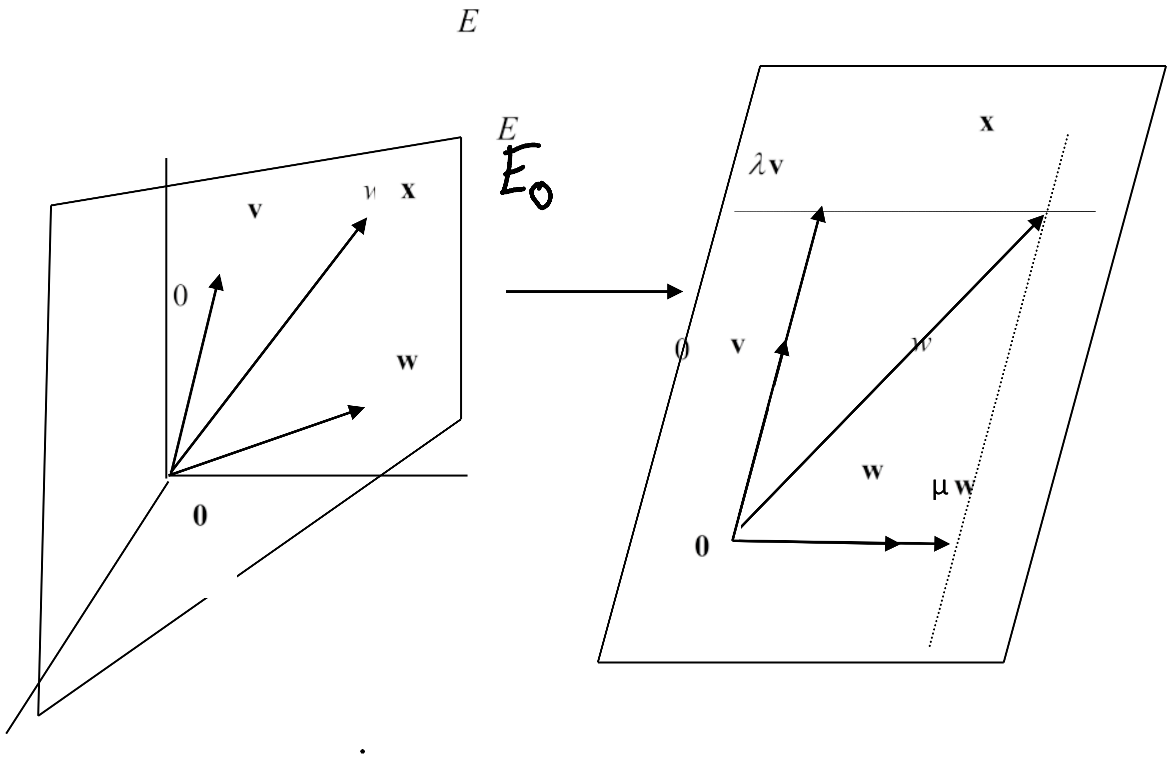
v

Επίπεδο στον R³

Ας θεωρήσουμε διανύσματα στο χώρο, που ακριβώς όπως τα διανύσματα του επιπέδου μπορούν να ταυτιστούν με τα σημεία του επιπέδου, του R^2, έτσι και τα διανύσματα στο χώρο ταυτίζονται με τα σημεία του χώρου, που περιγράφονται από τον R^3.

Επίπεδο που διέρχεται από 0

- Ας προσπαθήσουμε πάλι να περιγράψουμε με μια εξίσωση τα σημεία του επιπέδου που διέρχονται από το 0 και κάποια σημεία $\mathbf{v}$ και $\mathbf{w}$ του χώρου.



- Ένα σημείο x αυτού του επιπέδου διακρίνεται από την ιδιότητα ότι μπορώ να το γράψω σαν

$\mathbf{x}$ = πολ/σιο του $\mathbf{v}$ + πολ/σιο του $\mathbf{w}$

ή υπάρχουν κατάλληλα λ και μ τέτοια ώστε

$$\mathbf{x} = \lambda\mathbf{v} + \mu\mathbf{w}$$

Λέμε: το $\mathbf{x}$ γράφεται ως γραμμικός συνδυασμός των $\mathbf{v}$ & $\mathbf{w}$ με συντελεστές τα λ, μ.

https://www.geogebra.org/m/YHmGKuJN

https://www.geogebra.org/m/HTdWt2YD

Για να βεβαιωθώ για την παραπάνω ιδιότητα μπορώ να φέρω παράλληλο του $\mathbf{w}$ μέσα από το $\mathbf{x}$· αυτή θα τμήσει κάπου την ευθεία που διέρχεται από το **0** και το $\mathbf{v}$. Εκεί θα βρίσκεται το ζητούμενο πολλαπλάσιο του $\mathbf{v}$. Παρομοίως μπορώ να βρω και το πολλαπλάσιο του $\mathbf{w}$.

Το επίπεδο E_0 δίνεται από το σύνολο των σημείων που γράφονται ως γραμμικός συνδυασμός των $\mathbf{v}$ & $\mathbf{w}$:

$$E_0 = \left\{ \mathbf{x} \in R^3 \middle| \exists \lambda, \mu \in R : \mathbf{x} = \lambda \mathbf{v} + \mu \mathbf{w} \right\}$$

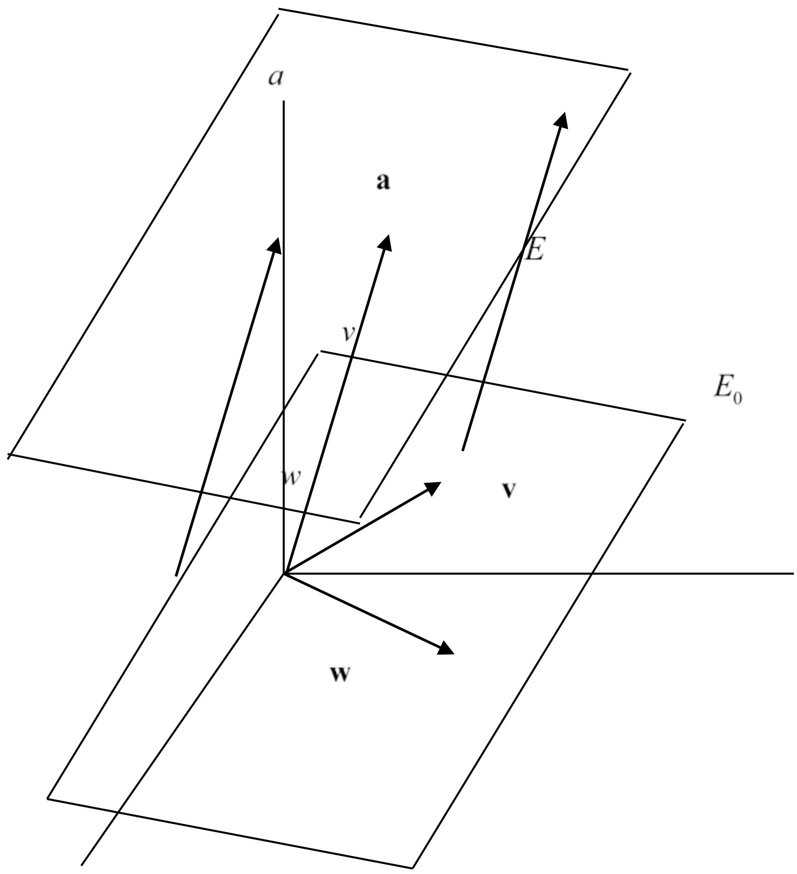
- Αυτό το σύνολο λέγεται γραμμική θήκη των $\mathbf{v},\mathbf{w}$ και γράφεται και ως $Span(\mathbf{v},\mathbf{w})$.
- Δηλαδή $E_0 = Span(\mathbf{v},\mathbf{w})$

Αντίστοιχα με την ευθεία μπορούμε τώρα να πούμε ότι **αφήνοντας δύο πραγματικούς αριθμούς** λ,μ **να «τρέχουν» πάνω στον *R*** (και χρησιμοποιώντας αυτούς τους δύο ως συντελεστές σε γραμμικούς συνδυασμούς των $\mathbf{v},\mathbf{w}$) τα σημεία που παίρνουμε $\mathbf{x} = \lambda \mathbf{v} + \mu \mathbf{w}$ «τρέχουν» πάνω σε ένα επίπεδο (ή απλώνουν ένα επίπεδο, εξού και "Span").

Και πάλι η γραφή του επιπέδου E_0 ως $Span(\mathbf{v},\mathbf{w})$ δεν είναι μοναδική: θα μπορούσα στη θέση των $\mathbf{v}$ & $\mathbf{w}$ να χρησιμοποιήσω δύο οποιαδήποτε άλλα μη συγγραμικά στοιχεία του E_0.

Επίπεδο που δε διέρχεται από το 0

Έστω E το επίπεδο που διέρχεται από κάποιο $\mathbf{a}$ και είναι παράλληλο του $E_0 = Span(\mathbf{v},\mathbf{w})$.



Τα σημεία του επιπέδου αυτού προκύπτουν αν προσθέσω $\mathbf{a}$ στα σημεία του E_0.

Αν δηλαδή «μετατοπίσω» το E_0 κατά a.

$$\boxed{\begin{aligned} &E = \left\{\mathbf{x} \in R^3 \middle| \mathbf{x} = \mathbf{x_0} + \mathbf{a}, \mathbf{x_0} \in E_0\right\} \\ &= \mathbf{a} + E_0 \\ &= \left\{\mathbf{x'} \in R^3 \middle| \exists \lambda, \mu \in R : \mathbf{x} = \mathbf{a} + \lambda\mathbf{v} + \mu\mathbf{w}\right\} \end{aligned}}$$

https://www.geogebra.org/m/wbBUwENF

https://www.geogebra.org/m/HTdWt2YD

- Αν θέλω το επίπεδο που περνά από τα σημεία $\mathbf{a},\mathbf{b},\mathbf{c} \in R^3$, αυτό θα είναι παράλληλο με το επίπεδο $Span(\mathbf{b}-\mathbf{a},\mathbf{c}-\mathbf{a})$.
  Περιγράφεται λοιπόν εναλλακτικά ως ως:

$$\begin{aligned} E &= \mathbf{a} + Span(\mathbf{b}-\mathbf{a},\mathbf{c}-\mathbf{a}) \\ &= \left\{ \mathbf{x} \in R^3 \middle| \exists \lambda, \mu \in R : \mathbf{x} = \mathbf{a} + \lambda(\mathbf{b}-\mathbf{a}) + \mu(\mathbf{c}-\mathbf{a}) \right\} \end{aligned}$$

- Οποιαδήποτε 3 άλλα σημεία του επιπέδου $\mathbf{a}',\mathbf{b}',\mathbf{c}' \in \mathrm{E}_0$ που δεν ανήκουν στην ίδια ευθεία, θα έδιναν το ίδιο επίπεδο:

$$E = \mathbf{a}' + Span(\mathbf{b}'-\mathbf{a}',\mathbf{c}'-\mathbf{a}')$$

Επίπεδο στον R^n

Άμεση είναι η γενίκευση στον ορισμό του επιπέδου στον R^n:

Ορισμός: Ένα υποσύνολο $E \subset R^n$ ονομάζεται επίπεδο αν υπάρχουν $\mathbf{a},\mathbf{v},\mathbf{w} \in R^n$ τέτοια που να ισχύει

$$E = \left\{ \mathbf{x} \in R^n \middle| \exists \lambda, \mu : \mathbf{x} = \mathbf{a} + \lambda \mathbf{v} + \mu \mathbf{w} \right\}$$

δηλ. το E ταυτίζεται με τη μετατόπιση κατά $\mathbf{a}$ του $Span(\mathbf{v},\mathbf{w})$: $E = \mathbf{a} + Span(\mathbf{v},\mathbf{w})$.

3. Μέτρο, εσωτερικό γινόμενο και ορθογωνιότητα στον R^n

Εσωτερικό γινόμενο

Έστω $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in R^n$ και $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in R^n$.

Ορίζουμε το εσωτερικό γινόμενο των $\mathbf{x}$ και $\mathbf{y}$ ως:

$$< \mathbf{x}, \mathbf{y} > := x_1 y_1 + x_2 y_2 + \ldots + x_n y_n \in R \text{ !}$$

https://www.geogebra.org/m/N9pvSPf4

Κανόνες πράξεων

- $< \mathbf{x}, \mathbf{y} > = < \mathbf{y}, \mathbf{x} >$
- $< \mathbf{x} + \mathbf{x}', \mathbf{y} > = < \mathbf{x}, \mathbf{y} > + < \mathbf{x}', \mathbf{y} >$
- $< \lambda \mathbf{x}, \mathbf{y} > = \lambda < \mathbf{x}, \mathbf{y} >$ για $\lambda \in R$

Αποδείξεις…

Συνέπεια:

$$\langle (\lambda \mathbf{x} + \kappa \mathbf{y}), (\xi \mathbf{u} + \psi \mathbf{w}) \rangle = \lambda \xi \langle \mathbf{x}, \mathbf{u} \rangle + \lambda \psi \langle \mathbf{x}, \mathbf{w} \rangle + \kappa \xi \langle \mathbf{y}, \mathbf{u} \rangle + \kappa \psi \langle \mathbf{y}, \mathbf{w} \rangle$$

Δηλ., κάνω πράξεις όπως με γινόμενα παρενθέσεων.

Μέτρο

Στον R^2 έχω ότι το μέτρο ενός διανύσματος $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ δίνεται από $\|\mathbf{x}\| = \sqrt{x_1{}^2 + x_2{}^2}$.

Αυτό όμως ισούται με $\sqrt{<\mathbf{x},\mathbf{x}>}$. Αυτό χρησιμοποιώ ως ορισμό του μέτρου και στον R^n.

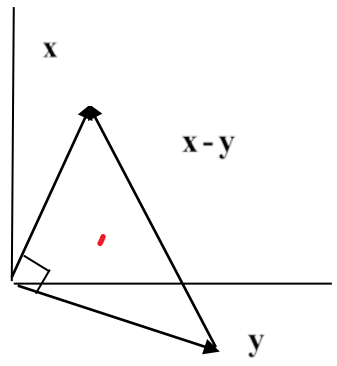
Αν $\mathbf{x} \in R^n$ τότε το μέτρο του $\mathbf{x}$ δίνεται από

$$\|\mathbf{x}\| := \sqrt{<\mathbf{x},\mathbf{x}>} = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Ιδιότητες

- $\|\mathbf{x}\| \geq 0$
- $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = 0$
- $\|\lambda\mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\| \quad \forall \lambda \in R$

Ορθογωνιότητα



Στον R^2 έχουμε $\|\mathbf{x}-\mathbf{y}\|^2=<\mathbf{x}$ [illegible]

$=<\mathbf{x},\mathbf{x}>-<\mathbf{x},\mathbf{y}>-<\mathbf{y},\mathbf{x}>+<\mathbf{y},\mathbf{y}>$

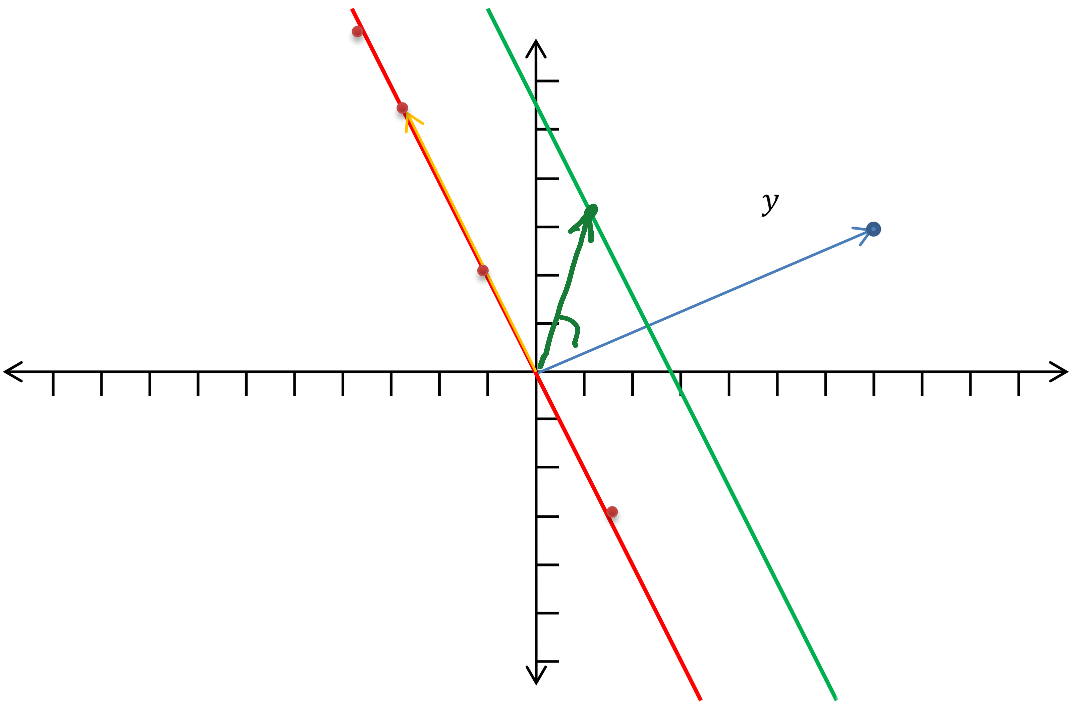
$=\|\mathbf{x}\|^2+\|\mathbf{y}\|^2-2<\mathbf{x},\mathbf{y}>$

Από το πυθαγόρειο γνωρίζουμε ότι αν $\varphi=90^o$ τότε $\|\mathbf{x}-\mathbf{y}\|^2=\|\mathbf{x}\|^2+\|\mathbf{y}\|^2$ και επομένως πρέπει $\underline{<\mathbf{x},\mathbf{y}>=0}$. Αυτό χρησιμοποιούμε και ως ορισμό της ορθογωνιότητας στον R^n.

<u>Ορισμός:</u> Για $\mathbf{x},\mathbf{y}\in R^n$ λέμε ότι $\mathbf{x}$ & $\mathbf{y}$ ορθογώνια και γράφουμε $\mathbf{x}\perp\mathbf{y}$ όταν $<\mathbf{x},\mathbf{y}>=0$.

Προσοχή: $\mathbf{x}\perp\mathbf{y}$ όταν έχουν μεταξύ τους γωνία 90 μοίρες όταν τα φέρω στο κοινό σημείο εκκίνησης.

Π.χ. είναι ορθογώνια στο **y** τα **x** της κόκκινης ευθείας? Της πράσινης?



Εφαρμογή: προβολή σε ευθεία

- Έστω η ευθεία $L_{\mathbf{a}} = Span(\mathbf{a})$ και ένα σημείο $\mathbf{x} \notin L_{\mathbf{a}}$. Θέλω να υπολογίσω την προβολή του $\mathbf{x}$ στην $L_{\mathbf{a}}$, την $P_{\mathbf{a}}\mathbf{x}$.

- Κατ' αρχάς: αφού $P_{\mathbf{a}}\mathbf{x} \in L_{\mathbf{a}} \Rightarrow \exists \lambda_0 : P_{\mathbf{a}}\mathbf{x} = \lambda_0 \mathbf{a}$

  Επίσης: $< \mathbf{a}, \mathbf{x} - P_{\mathbf{a}}\mathbf{x} >= 0$

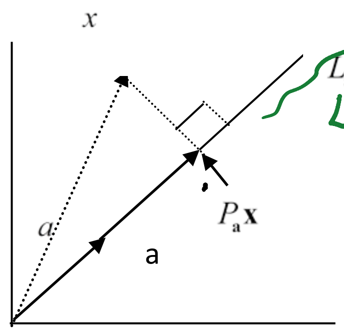
  $\Leftrightarrow < \mathbf{a}, \mathbf{x} - \lambda_0 \boldsymbol{\alpha} >= 0$

  $\Leftrightarrow < \mathbf{a}, \mathbf{x} > - \lambda_0 < \mathbf{a}, \mathbf{a} >= 0$

  $$\Leftrightarrow \lambda_0 = \frac{< \mathbf{a}, \mathbf{x} >}{< \mathbf{a}, \mathbf{a} >}$$

$\langle a,x\rangle + \langle a, -\lambda_0 a\rangle = 0$

$\langle a,x\rangle - \lambda_0 \langle a,a\rangle = 0$



$\langle a,x\rangle - \lambda_0 \langle a,a\rangle$

$\lambda_0 = \frac{\langle a,x\rangle}{\langle a,a\rangle}$

Άρα $\quad P_{\mathbf{a}}\mathbf{x} = \lambda_0 \mathbf{a} = \dfrac{< \mathbf{a}, \mathbf{x} >}{< \mathbf{a}, \mathbf{a} >} \cdot \mathbf{a}$

Παρατηρείστε ότι:

$$\mathrm{P}_{\mathbf{a}}\mathbf{x} \ \textit{και} \ \mathbf{a} \ \text{ομόροπα} \Leftrightarrow \lambda_0 \geq 0 \Leftrightarrow \langle \mathbf{a}, \mathbf{x} \rangle \geq 0$$

https://www.geogebra.org/m/TZvtkHKz

https://www.geogebra.org/m/mkV7F8Jf

Εφαρμογή: Η ανισότητα Cauchy-Schwartz

Έστω $\mathbf{x},\mathbf{a} \in R^n$ και η προβολή του $\mathbf{x}$ στην $L_{\mathbf{a}} : P_{\mathbf{a}}\mathbf{x}$.

Έχουμε: $\mathbf{x} = (\mathbf{x} - P_{\mathbf{a}}\mathbf{x}) + P_{\mathbf{a}}\mathbf{x}$ και άρα

$$\|\mathbf{x}\|^2 = <\mathbf{x},\mathbf{x}> = <(\mathbf{x} - P_{\mathbf{a}}\mathbf{x}) + P_{\mathbf{a}}\mathbf{x}, (\mathbf{x} - P_{\mathbf{a}}\mathbf{x}) + P_{\mathbf{a}}\mathbf{x}>$$

$$= \|\mathbf{x} - P_{\mathbf{a}}\mathbf{x}\|^2 + \|P_{\mathbf{a}}\mathbf{x}\|^2 + \underbrace{2<\mathbf{x} - P_{\mathbf{a}}\mathbf{x}, P_{\mathbf{a}}\mathbf{x}>}_{0}$$

$$\Rightarrow \quad \|P_{\mathbf{a}}\mathbf{x}\| \le \|\mathbf{x}\|$$

$$\Rightarrow \quad \left\|\frac{<\mathbf{a},\mathbf{x}>}{<\mathbf{a},\mathbf{a}>}\cdot\mathbf{a}\right\| \le \|\mathbf{x}\| \quad \Rightarrow \quad |<\mathbf{a},\mathbf{x}>|\cdot\frac{\|\mathbf{a}\|}{\|\mathbf{a}\|^2} \le \|\mathbf{x}\|$$

$$\Rightarrow \boxed{|<\mathbf{a},\mathbf{x}>| \le \|\mathbf{x}\|\cdot\|\mathbf{a}\|}$$

με ισότητα αν $\mathbf{x} = P_{\mathbf{a}}\mathbf{x} \Leftrightarrow \mathbf{x},\mathbf{a}$ συγγραμικά.

Γωνία δύο διανυσμάτων στον R^n

Έστω $\mathbf{x},\mathbf{y}\in R^n$ και η προβολή του $\mathbf{x}$ στην $L_{\mathbf{y}}: P_{\mathbf{y}}\mathbf{x}$.

- Ορίζουμε ως γωνία φ των διανυσμάτων $\mathbf{x},\mathbf{y}$ εκείνο τον αριθμό $\in[0,\pi]$ για τον οποίο:

$$\cos(\varphi)=\begin{cases}\dfrac{\|P_{\mathbf{y}}\mathbf{x}\|}{\|\mathbf{x}\|}=\dfrac{\left\|\dfrac{\langle\mathbf{x},\mathbf{y}\rangle}{\langle\mathbf{y},\mathbf{y}\rangle}\mathbf{y}\right\|}{\|x\|}=\dfrac{|\langle\mathbf{x},\mathbf{y}\rangle|}{\|\mathbf{x}\|\|\mathbf{y}\|}, \alpha\nu\ P_{\mathbf{y}}\mathbf{x}\ \kappa\alpha\iota\ \mathbf{y}\ \text{ομόροπα} \\ -\dfrac{\|P_{\mathbf{y}}\mathbf{x}\|}{\|\mathbf{x}\|}=-\dfrac{\left\|\dfrac{\langle\mathbf{x},\mathbf{y}\rangle}{\langle\mathbf{y},\mathbf{y}\rangle}\mathbf{y}\right\|}{\|\mathbf{x}\|}=-\dfrac{|\langle\mathbf{x},\mathbf{y}\rangle|}{\|\mathbf{x}\|\|\mathbf{y}\|}, \alpha\nu\ P_{\mathbf{y}}\mathbf{x}\ \kappa\alpha\iota\ \mathbf{y}\ \text{αντίροπα}\end{cases} = \frac{\langle\mathbf{x},\mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}$$

- Η ανισότητα Cauchy-Schwartz εγγυάται ότι αυτή η τιμή $\in[-1,1]$

Τριγωνική ανισότητα:

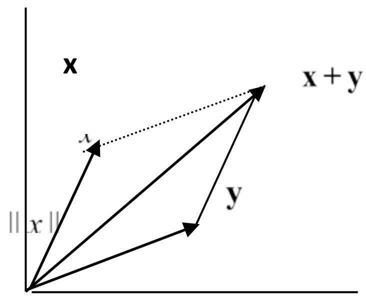
Για $\mathbf{x},\mathbf{y}\in R^n$ έχουμε:

$$\|\mathbf{x}+\mathbf{y}\|^2=<\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}>=<\mathbf{x},\mathbf{x}>+2<\mathbf{x},\mathbf{y}>+<\mathbf{y},\mathbf{y}>$$

$$=\|\mathbf{x}\|^2+\|\mathbf{y}\|^2+2<\mathbf{x},\mathbf{y}>$$

$$\le\|\mathbf{x}\|^2+\|\mathbf{y}\|^2+2\|\mathbf{x}\|\cdot\|\mathbf{y}\|$$

Cauchy-Schwartz



$$=[\|\mathbf{x}\|+\|\mathbf{y}\|]^2$$

$$\Rightarrow \boxed{\|\mathbf{x}+\mathbf{y}\| \quad\le\quad \|\mathbf{x}\|+\|\mathbf{y}\|}$$

Εφαρμογή: οι λύσεις μιας εξίσωσης σε πολλούς αγνώστους

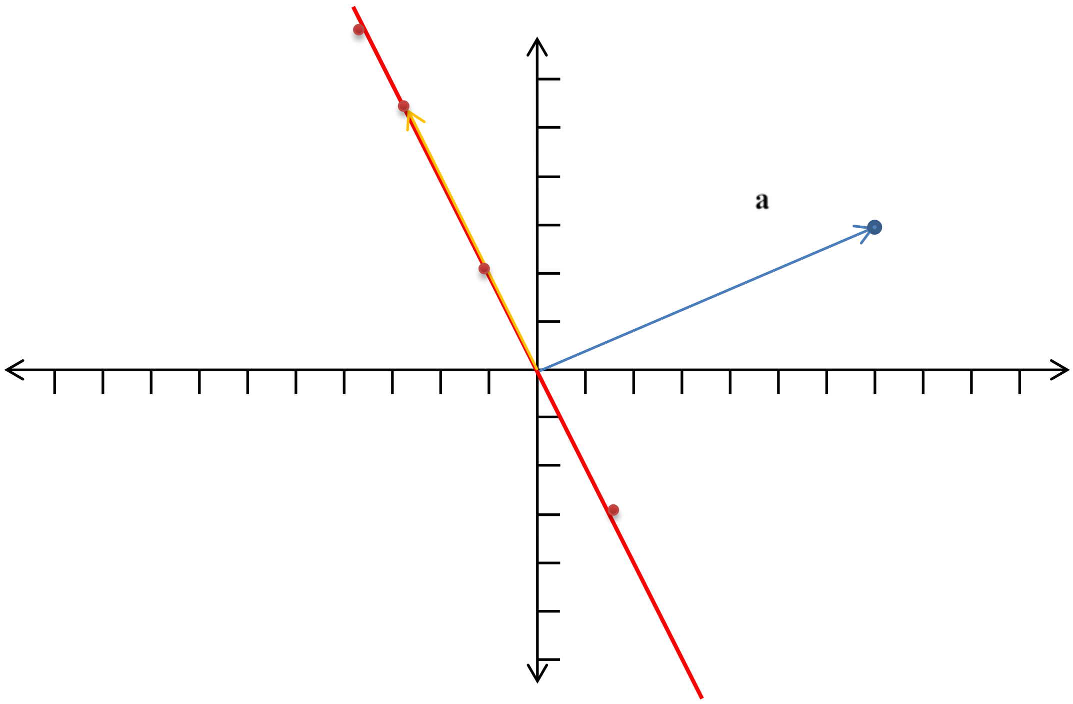
- Ενδιαφέρομαι για εξισώσεις της μορφής $a_1x_1+...+a_nx_n=b$, όπου $a_1,...,a_n$, b γνωστές σταθερές και $x_1,...,x_n$ άγνωστοι: Ποιό είναι το σύνολο των $\mathbf{x}=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ που ικανοποιεί αυτή την εξίσωση?
- Π.χ. $2x_1+3x_2+4x_3+x_4=6$?
- Τέτοιες εξισώσεις γράφονται και ως $\langle \mathbf{a},\mathbf{x}\rangle = b$ με $\mathbf{a}=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$

Πρώτα στον $\mathbb{R}^2$ με b=0:

- $a_1x_1 + a_2x_2 = 0$, δηλ. $\langle \mathbf{a}, \mathbf{x} \rangle = 0$ ή και $\mathbf{x} \perp \mathbf{a}$

  π.χ. $2x_1 + 1x_2 = 0$, δηλ. $x \perp \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

- *<u>Γεωμετρική λύση</u>*: Όλα τα $\mathbf{x}$ που ανήκουν στην ευθεία που περνά από το 0



και είναι κάθετη στο $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$.

- *<u>Αλγεβρική λύση</u>*: Όλα τα x που ικανοποιούν την $x_1 = -\frac{a_2}{a_1}x_2$. Άρα οι λύσεις θα είναι όλα τα $\mathbf{x}$ της μορφής $\mathbf{x} = \begin{pmatrix} -\frac{a_2}{a_1}x_2 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \end{pmatrix}$ για οποιοδήποτε (δηλαδή για όλα τα) $x_2 \in \mathbb{R}$.

  Άρα το σύνολο των λύσεων είναι η ευθεία $Span\left\{\begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \end{pmatrix}\right\}$ που διέρχεται από το 0.

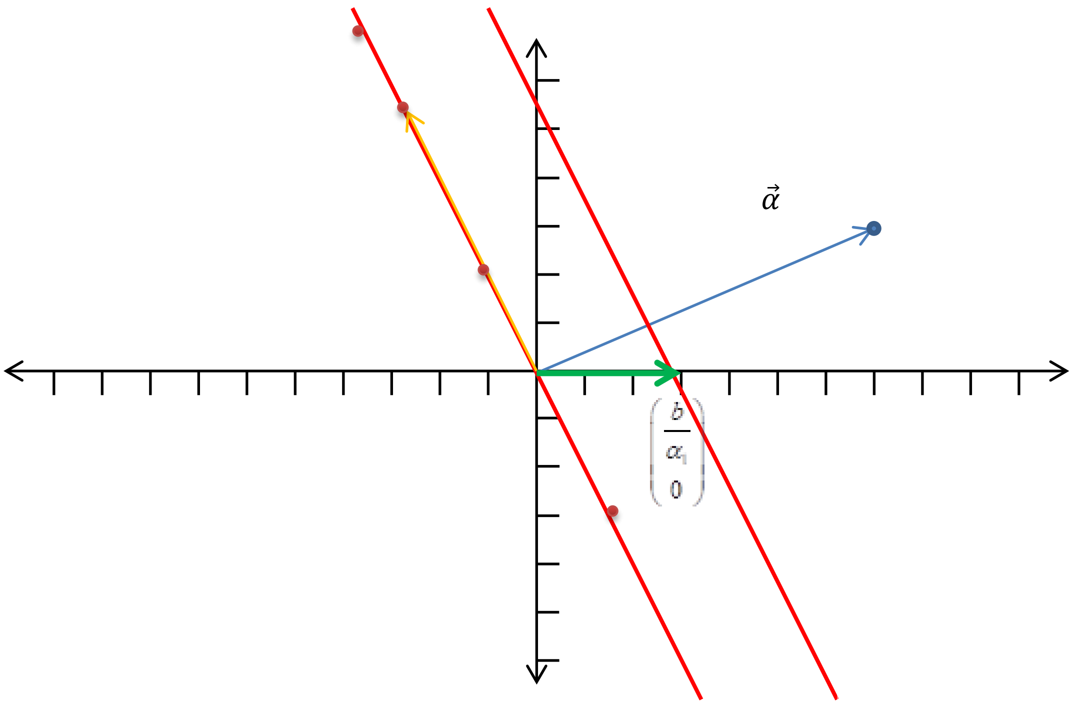
Στον $\mathbb{R}^2$ ***με*** $b \neq 0$***:***

❖ <u>*Αλγεβρική λύση*</u>: Όλα τα $\mathbf{x}$ που ικανοποιούν την $x_1 = \frac{b}{a_1} - \frac{a_2}{a_1}x_2$. Άρα οι λύσεις θα είναι όλα τα x της μορφής

$$\mathbf{x} = \begin{pmatrix} \frac{b}{a_1} - \frac{a_2}{a_1}x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{b}{a_1} \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \end{pmatrix}$$

για οποιοδήποτε (δηλαδή για όλα τα) $x_2 \in \mathbb{R}$.

Άρα το σύνολο των λύσεων είναι η ευθεία $\begin{pmatrix} \frac{b}{a_1} \\ 0 \end{pmatrix} + Span\left\{ \begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \end{pmatrix} \right\}$ που δε διέρχεται από το 0 διότι έχει μετατοπιστεί κατά $\begin{pmatrix} \frac{b}{a_1} \\ 0 \end{pmatrix}$.



Στον $\mathbb{R}^3$ με b=0:

- $a_1x_1+a_2x_2+a_3x_3=0$, δηλ. $\langle \mathbf{a},\mathbf{x}\rangle=0$ ή και $\mathbf{x}\perp\mathbf{a}$

  π.χ. $2x_1+1x_2+2x_3=0$, δηλ. $\mathbf{x}\perp\begin{pmatrix}2\\1\\2\end{pmatrix}$.

- _Γεωμετρική λύση_: Όλα τα $\mathbf{x}$ που ανήκουν στο επίπεδο που περνά από το 0 και είναι κάθετη στο $\mathbf{a}=\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}$.

  https://www.geogebra.org/m/myVDsRRt

- _Αλγεβρική λύση_: Όλα τα $\mathbf{x}$ που ικανοποιούν την $x_1=-\frac{a_2}{a_1}x_2-\frac{a_3}{a_1}x_3$. Άρα οι λύσεις θα είναι όλα τα $\mathbf{x}$ της μορφής $\mathbf{x}=\begin{pmatrix}-\frac{a_2}{a_1}x_2-\frac{a_3}{a_1}x_3\\x_2\\x_3\end{pmatrix}=x_2\begin{pmatrix}-\frac{a_2}{a_1}\\1\\0\end{pmatrix}+x_3\begin{pmatrix}-\frac{a_3}{a_1}\\0\\1\end{pmatrix}$ για οποιοδήποτε (δηλαδή για όλα τα) $x_2,x_3\in\mathbb{R}$.

  Άρα το σύνολο των λύσεων είναι το επίπεδο $Span\left\{\begin{pmatrix}-\frac{a_2}{a_1}\\1\\0\end{pmatrix},\begin{pmatrix}-\frac{a_3}{a_1}\\0\\1\end{pmatrix}\right\}$ που διέρχεται από το 0.

Στον $\mathbb{R}^3$ ***με*** $b \neq 0$***:***

❖ *<u>Αλγεβρική λύση</u>*: Όλα τα x που ικανοποιούν την $x_1 = \frac{b}{a_1} - \frac{a_2}{a_1}x_2 - \frac{a_3}{a_1}x_3$.

Άρα οι λύσεις θα είναι όλα τα $\mathbf{x}$ της μορφής

$$\mathbf{x} = \begin{pmatrix} \frac{b}{a_1} - \frac{a_2}{a_1}x_2 - \frac{a_3}{a_1}x_3 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{b}{a_1} \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -\frac{a_3}{a_1} \\ 0 \\ 1 \end{pmatrix}$$

για οποιοδήποτε (δηλαδή για όλα τα) $x_2, x_3 \in \mathbb{R}$.

Άρα το σύνολο των λύσεων είναι το επίπεδο

$$\begin{pmatrix} \frac{b}{a_1} \\ 0 \\ 0 \end{pmatrix} + Span\left\{ \begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{a_3}{a_1} \\ 0 \\ 1 \end{pmatrix} \right\}$$

που δε διέρχεται από το 0 διότι έχει μετατοπιστεί κατά $\begin{pmatrix} \frac{b}{a_1} \\ 0 \\ 0 \end{pmatrix}$.

https://www.geogebra.org/m/myVDsRRt

4. Γεωμετρική ερμηνεία συστημάτων γραμμικών εξισώσεων στον R3

(Πηγή: Strang, Κεφάλαιο 1.2)

Υπενθύμιση: μια εξίσωση στον R3

- εξισώσεις της μορφής $a_1x_1 + a_2x_2 + a_3x_3 = b$, όπου $\alpha_1, \alpha_2, \alpha_3$, b γνωστές σταθερές και x_1, x_2, x_3 άγνωστοι. Π.χ. $2x_1 + x_2 + x_3 = 5$?

- Τέτοιες εξισώσεις γράφονται και ως $\langle \mathbf{a}, \mathbf{x} \rangle = b$ με $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$

- Μια τέτοια εξίσωση ικανοποιείται από πολλά $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Το σύνολο των $\mathbf{x}$ που ικανοποιούν την εξίσωση σχηματίζουν ένα επίπεδο: Κάθε σημείο $\mathbf{x}$ του επιπέδου είναι λύση της εξίσωσης.

- Πρόκειται για το επίπεδο αυτό είναι κάθετο στο $\mathbf{a}$.

- Αν $b \neq 0$ το επίπεδο δεν περνά από το 0,

  αλλά είναι μετατοπισμένο με μετατόπιση που αυξάνεται όσο αυξάνεται το b.

- Οι λύσεις είναι οι

$$\begin{pmatrix} \frac{b}{a_1} \\ 0 \\ 0 \end{pmatrix} + Span\left\{ \begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{a_3}{a_1} \\ 0 \\ 1 \end{pmatrix} \right\}$$

https://www.geogebra.org/m/myVDsRRt

- Στο παράδειγμά μας

$$\begin{pmatrix} 5/2 \\ 0 \\ 0 \end{pmatrix} + Span\left\{ \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1/2 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Συστήματα γραμμικών εξισώσεων στον R3

- Όταν έχω πολλές τέτοιες εξισώσεις (π.χ. τρεις εξισώσεις) και ψάχνω τα x που λύνουν ταυτόχρονα όλες τις εξισώσεις.

Π.χ.
$$\begin{aligned} 2x_1 + x_2 + x_3 &= 5 \\ 4x_1 - 6x_2 + 0x_3 &= 5 \\ -2x_1 + 7x_2 + 2x_3 &= 9 \end{aligned}$$

- Η γενική μορφή ενός τέτοιου συστήματος είναι λοιπόν η

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned},$$

Όπου συμβολίζουμε τους συντελεστές a_{ij} με διπλό δείκτη που ο πρώτος δίνει τη γραμμή (αριθμό εξίσωσης) και ο δεύτερος τη στήλη (αριθμό αγνώστου).

Στο παράδειγμα…

Άλλη γραφή των συστημάτων γραμμικών εξισώσεων:

- Πριν γράψαμε την εξίσωση $a_1x_1 + a_2x_2 + a_3x_3 = b$ και ως $\langle \mathbf{a}, \mathbf{x} \rangle = b$ με $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, όπου a το διάνυσμα των συντελεστών.

- Τώρα έχουμε τρία διανύσματα συντελεστών ένα για κάθε εξίσωση. Για να τα διακρίνουμε τα συμβολίζουμε με $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3$.

Παρένθεση: Αν θέλουμε να γράψουμε ένα διάνυσμα ως γραμμή αντί για στήλη (να το ξαπλώσουμε) του προθέτουμε ένα δείκτη «Τ» πάνω δεξιά: π.χ. $\mathbf{x}^{\mathrm{T}} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}$

- Έτσι:

$$\begin{aligned} (\mathbf{a}^1)^T &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \end{pmatrix} \\ (\mathbf{a}^2)^T &= \begin{pmatrix} a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} 4 & -6 & 0 \end{pmatrix} \\ (\mathbf{a}^3)^T &= \begin{pmatrix} a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} -2 & 7 & 2 \end{pmatrix} \end{aligned}$$

Και με αυτά το σύστημα γράφεται ως:

$$\begin{aligned} \langle \mathbf{a}^1, x \rangle &= b_1 \\ \langle \mathbf{a}^2, x \rangle &= b_2 \\ \langle \mathbf{a}^3, x \rangle &= b_3 \end{aligned}$$

Οι λύσεις του συστήματος με δυο οπτικές γωνίες:

Α. Οπτική γωνία των γραμμών

- Κάθε μια από τις εξισώσεις $\begin{array}{l}\langle \mathbf{a}^1, x\rangle = b_1\\ \langle \mathbf{a}^2, x\rangle = b_2\\ \langle \mathbf{a}^3, x\rangle = b_3\end{array}$ έχει λύσεις ένα επίπεδο (κάθετο στο $\mathbf{a}^{\mathbf{i}}$ μετατοπισμένο κατά διάνυσμα που αυξάνεται με το b_i). Αυτό δίνει τρία επίπεδα.

- Οι λύσεις του συστήματος πρέπει να ικανοποιούν και τις τρείς εξισώσεις, άρα να ανήκουν ταυτόχρονα και στα τρία παραπάνω επίπεδα.

- Άρα κάθε λύση πρέπει να ανήκει στη τομή των τριών επιπέδων.

- «**Κανονικά**», αν δεν έχουμε μια «εξαιρετική» περίπτωση (που θα τη λέμε **ιδιόμορφη**) δύο επίπεδα τέμνονται σε μια ευθεία και με τη τρίτο επίπεδο αυτή τέμνεται σε ένα σημείο, το $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ που είναι σε αυτή τη περίπτωση και η μοναδική λύση του συστήματος. Αυτή η περίπτωση ονομάζεται **μη-ιδιόμορφη**

2 επίπεδα
https://www.geogebra.org/m/QeMWg5Q8

3 επίπεδα
https://www.geogebra.org/m/pjczxakm

- Ποια είναι η ιδιόμορφη περίπτωση? Να έχουμε είτε άπειρες λύσεις είτε καμία λύση

Υπό ποιές προϋποθέσεις στα $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3, \mathrm{b}_1, \mathrm{b}_2, \mathrm{b}_3$ συμβαίνουν αυτά;

- Αν $\mathbf{a}^2 = \lambda \mathbf{a}^1$ τότε τα αντίστοιχα δύο επίπεδα είναι παράλληλα

  $\Rightarrow \left(\mathbf{a}^1\right)^T = \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \end{pmatrix}$

  $\Rightarrow \left(\mathbf{a}^2\right)^T = \begin{pmatrix} a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 \end{pmatrix}$

  - Αν έχουμε και $b_2 = \lambda b_1$ τότε τα δύο επίπεδα ταυτίζονται (ίσως άπειρες λύσεις, και δη αν $\mathbf{a}^3 \neq \kappa \mathbf{a}^1$)
  - Αν έχουμε και $b_2 \neq \lambda b_1$ τότε τα δύο επίπεδα είναι παράλληλα με κενή τομή (καμία λύση –ασύμβατες εξισώσεις)

- Ή αν $\mathbf{a}^2 = \lambda \mathbf{a}^1$ και $\mathbf{a}^3 = \kappa \mathbf{a}^1$ τότε τα αντίστοιχα τρία επίπεδα είναι παράλληλα
  - Αν έχουμε $b_2 = \lambda b_1$ και $b_3 = \kappa b_1$ τότε τα τρία επίπεδα ταυτίζονται και έχουμε άπειρες λύσεις
  - Αν έχουμε είτε $b_2 \neq \lambda b_1$ είτε $b_3 = \kappa b_1$ τότε τα τρία επίπεδα έχουν κενή τομή (καμία λύση –ασύμβατες εξισώσεις)

- Ή αν $\mathbf{a}^3 = \mu \mathbf{a}^1 + \nu \mathbf{a}^2$ τότε
  - Αν έχουμε και $b_3 = \mu b_1 + \nu b_2$ τότε τα τρία επίπεδα τέμνονται σε μια ευθεία και έχουμε άπειρες λύσεις
  - Αν έχουμε $b_3 \neq \mu b_1 + \nu b_2$ τότε η ευθεία που είναι τομή των δύο πρώτων επιπέδων είναι παράλληλη στο τρίτο επίπεδο και άρα τα τρία επίπεδα έχουν κενή τομή (καμία λύση –ασύμβατες εξισώσεις)

https://www.geogebra.org/m/xu2t99AD

- **Γενικό συμπέρασμα:** Με την οπτική γωνία των γραμμών έχουμε την **ιδιόμορφη περίπτωση** όταν οι αριστερές πλευρές κάποιας γραμμή γράφεται ως γραμμικός συνδυασμός των άλλων.
- Τότε αν ακριβώς το ίδιο συμβαίνει και με τις δεξιές πλευρές έχουμε άπειρες λύσεις. Αν όχι, οι εξισώσεις είναι ασύμβατες και δεν έχουμε καμία λύση.

Β. Οπτική γωνία των στηλών

- Ας δώσουμε ονόματα στις στήλες του συστήματος:

$$\mathbf{a_1} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \mathbf{a_2} = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \mathbf{a_3} = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$$

- Το σύστημα γράφεται ως ισότητα διανυσμάτων

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

➢ Που γίνεται με λίγες πράξεις και:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

➢ Και με τα ονόματα των στηλών που εισάγαμε παίρνουμε:

$x_1\mathbf{a_1} + x_2\mathbf{a_2} + x_3\mathbf{a_3} = b$ θέτοντας και $b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$

- Το σύστημα έχει λύση αν η δεξιά πλευρά (b) γράφεται ως γραμμικός συνδυασμός των στηλών, δηλαδή όταν $b \in Span\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$. Η λύση $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ δίνεται από τους συντελεστές αυτού του γραμμικού συνδυασμού.

- Υπό ποιές προϋποθέσεις στα $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathrm{b}$ έχουμε την **ιδιόμορφη περίπτωση** (καμία ή άπειρες λύσεις);

  - Π.χ. καμία λύση: πότε γίνεται να υπάρχει b τέτοιο ώστε:

$$b \notin Span\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$$

$\Rightarrow$ Για να συμβαίνει αυτό πρέπει

$Span\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\} \neq \mathbb{R}^3$ (οπότε θα υπάρχουν $b \notin Span\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$)

$\Rightarrow$ Αυτό γίνεται μόνο αν κάποια στήλη ανήκει στο επίπεδο που σχηματίζουν οι άλλες δύο, π.χ. αν $\mathbf{a}_3 \in Span\{\mathbf{a}_1, \mathbf{a}_2\}$.

$\Rightarrow$ Διότι τότε $Span\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\} = Span\{\mathbf{a}_1, \mathbf{a}_2\} \neq \mathbb{R}^3$

  - Αποδεικνύεται ότι τότε ακριβώς έχουμε και τις άπειρες λύσεις: αν κάποια στήλη ανήκει στο επίπεδο που σχηματίζουν οι άλλες δύο και $b \in Span\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\} = Span\{\mathbf{a}_1, \mathbf{a}_2\}$ τότε θα γράφεται με πολλούς τρόπους ως γ.σ. των $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$.

Παράδειγμα των παραπάνω:

Έστω δύο συστήματα με ίδιες αριστερές πλευρές

Α. $\begin{aligned} 1x_1+1x_2+x_3&=2\\ 2x_1+0x_2+3x_3&=5\\ 3x_1+1x_2+4x_3&=7 \end{aligned}$ και το Β. $\begin{aligned} 1x_1+1x_2+x_3&=2\\ 2x_1+0x_2+3x_3&=5\\ 3x_1+1x_2+4x_3&=6 \end{aligned}$

Για τις αριστερές πλευρές έχουμε

$\mathbf{a}_1=\frac{1}{3}\mathbf{a}_2+\frac{2}{3}\mathbf{a}_3$ και επομένως $Span\{\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3\}=Span\{\mathbf{a}_1,\mathbf{a}_2\}$

- Όμως: $\frac{1}{3}\begin{pmatrix}1\\0\\1\end{pmatrix}+\frac{5}{3}\begin{pmatrix}1\\3\\4\end{pmatrix}=\begin{pmatrix}2\\5\\7\end{pmatrix}$ και άρα το σύστημα Α έχει λύση.
  - Μάλιστα η παραπάνω λύση δεν είναι μοναδική. Μια άλλη λύση είναι η: $1\begin{pmatrix}1\\2\\3\end{pmatrix}+0\begin{pmatrix}1\\0\\1\end{pmatrix}+1\begin{pmatrix}1\\3\\4\end{pmatrix}=\begin{pmatrix}2\\5\\7\end{pmatrix}$
  - Άρα **βρισκόμαστε στην ιδιόμορφη περίπτωση** διότι το $\begin{pmatrix}2\\5\\7\end{pmatrix}$ γράφεται με πολλούς τρόπους ως γ.σ των $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$.
- Όμως το Β δεν έχει καμία λύση: η τελευταία εξίσωση είναι ασύμβατη με τις δύο πρώτες, διότι το $\begin{pmatrix}2\\5\\6\end{pmatrix}\notin Span\{\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3\}=Span\{\mathbf{a}_1,\mathbf{a}_2\}$. Άρα πάλι **βρισκόμαστε στην ιδιόμορφη περίπτωση.**

- ΣΥΜΠΕΡΑΣΜΑ: **Με την *οπτική γωνία των στηλών* βρισκόμαστε στη ιδιόμορφη περίπτωση** όταν μια στήλη των συντελεστών του συστήματος γράφεται ως γραμμικός συνδυασμός των άλλων.

Συμπέρασμα:

- Με την *οπτική γωνία των* ***γραμμών*** βρισκόμαστε στη **ιδιόμορφη περίπτωση** όταν μια γραμμή των συντελεστών του συστήματος γράφεται ως γραμμικός συνδυασμός των άλλων.
- Με την *οπτική γωνία των* ***στηλών*** βρισκόμαστε στη **ιδιόμορφη περίπτωση** όταν μια στήλη των συντελεστών του συστήματος γράφεται ως γραμμικός συνδυασμός των άλλων.

Άρα: μια **γραμμή** των συντελεστών του συστήματος γράφεται ως γραμμικός συνδυασμός των άλλων **τότε και μόνο τότε όταν** μια **στήλη** των συντελεστών του συστήματος γράφεται ως γραμμικός συνδυασμός των άλλων.

5. Συμβολισμός πινάκων και πράξεις με πίνακες

(Πηγή: Strang, Κεφάλαιο 1.4)

- Ας θεωρήσουμε ένα σύστημα γραμμικών εξισώσεων σε «πολλούς αγνώστους», όπως για παράδειγμα το παρακάτω:

$$\begin{aligned} 2x_1 + 1x_2 + x_3 &= 5 \\ 4x_1 - 6x_2 + 0x_3 &= -2 \end{aligned}$$

- Συνήθως αναρωτιόμαστε ποια $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ λύνουν ταυτόχρονα και τις δύο εξισώσεις. Αλλά με την επίλυση γραμμικών συστημάτων θα ασχοληθούμε στο επόμενο κεφάλαιο.
- Εδώ θα χρησιμοποιήσουμε τα συστήματα ως κίνητρο για να εισάγουμε τους «πίνακες».

- **Η γενική μορφή ενός συστήματος** με m εξισώσεις και n αγνώστους είναι

$$(*)\ \left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_n \end{array}\right\}$$

όπου τα a_{ij} είναι οι συντελεστές των x_j.

δείκτης γραμμής   δείκτης στήλης

Στο παράδειγμά μας: a_{11} a_{12} a_{13} b_1

$$\begin{array}{l} 2x_1 + 1x_2 + 1x_3 = 5 \\ 4x_1 + (-6)x_2 + 0x_3 = -2 \end{array}$$

a_{21} a_{22} a_{23} b_2

Πιο σύντομη γραφή του συστήματος $(*)$:

- Συμβολίζω με $\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in R^n$ και με

- $\mathbf{a}^{\mathbf{i}^T}$ την «ξαπλωμένη» (εξ ου και το T που διαβάζεται «ανάστροφο») *i*-γραμμή των συντελεστών $\mathbf{a}^{\mathbf{i}^T}=\left(a_{i1},a_{i2},\dots,a_{in}\right)$ με $\mathbf{a}^{\mathbf{i}} \in R^n$

  π.χ. $\mathbf{a}^{\mathbf{1}^T}=\left(2,1,1\right)$ και $\mathbf{a}^{\mathbf{2}^T}=\left(4,-6,0\right)$

- <u>τότε</u> η *i*-εξίσωση γράφεται $\boxed{\left\langle \mathbf{a}^{\mathbf{i}},\mathbf{x}\right\rangle = b_i}$ και το σύστημα γράφεται

$$(*)\left\{\begin{matrix} \left\langle \mathbf{a}^{\mathbf{1}},\mathbf{x}\right\rangle = b_1 \\ \left\langle \mathbf{a}^{2},\mathbf{x}\right\rangle = b_2 \\ \vdots \\ \left\langle \mathbf{a}^{\mathbf{m}},\mathbf{x}\right\rangle = b_m \end{matrix}\right\}$$

Ψάχνω **ακόμα «πιο σύντομη» γραφή του συστήματος**:

- Πιάνω όλους τους συντελεστές a_{ij} σε ένα «τετραγωνικό αντικείμενο» $\mathbf{A}$ και
- Θέλω να ορίσω ένα «πολλαπλασιασμό» μεταξύ $\mathbf{A}$ και $\mathbf{x}$ έτσι ώστε να γράφω το σύστημα ως $\boxed{\mathbf{A}\cdot\mathbf{x}=\mathbf{b}}$ με

$$\mathbf{b}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \text{ και } \mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

π.χ. $\mathbf{A}=\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \end{bmatrix}$

Αντικείμενα σαν τον $\mathbf{A}$ τα ονομάζω «πίνακες» και τα ορίζω παρακάτω:

<u>Ορισμός (Πίνακας):</u> Πίνακας $m \times n$ είναι ένα σύνολο $m \cdot n$ πραγματικών αριθμών a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ διατεταγμένων σε m γραμμές και n στήλες.

Ο $m \times n$ πίνακας $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ γράφεται και ως $\mathbf{A} = \left[a_{ij} \right]_{\substack{i=1,\ldots,\mathrm{m} \\ j=1,\ldots,n}}$

- **Πρόσθεση δύο $m \times n$ πινάκων**: Έστω $\mathbf{A} = \left[a_{ij} \right]$ και $\mathbf{B} = \left[b_{ij} \right]$ δύο $m \times n$ πίνακες.

Τότε ορίζω $\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} = \left[a_{ij} + b_{ij} \right]$ (αθροίζω κατά συντεταγμένες).

- **Πολλαπλασιασμός με $\lambda \in R$**. Ορίζω τον πίνακα $\lambda \cdot \mathbf{A} = \begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{bmatrix}$.

Κανόνες πράξεων

Αν $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ $m\times n$ πίνακες, και $\lambda,\mu\in R$,

- $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$
- $(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C})$
- $\exists\ m\times n\ \mathbf{0}:\ \mathbf{A}+\mathbf{0}=\mathbf{A}\ \ \forall\ m\times n\ A$ $\qquad \mathbf{0}=\underbrace{\begin{bmatrix}0 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 0\end{bmatrix}}_{m}\Bigg\}n$
-
- $\forall\mathbf{A}\ \ \exists(\text{-}\mathbf{A}):\ \mathbf{A}+(\text{-}\mathbf{A})=\mathbf{0}$
- $\lambda\cdot(\mu\cdot\mathbf{A})=(\lambda\cdot\mu)\cdot\mathbf{A}$
- $(\lambda+\mu)\cdot\mathbf{A}=\lambda\cdot\mathbf{A}+\mu\cdot\mathbf{A}$
- $\lambda\cdot(\mathbf{A}+\mathbf{B})=\lambda\cdot\mathbf{A}+\lambda\cdot\mathbf{B}$

Σύμβαση: Τα **διανύσματα του R^n τα γράφουμε ως πίνακες με μια στήλη**, δηλαδή **ως $n\times 1$ πίνακες**, όχι ως γραμμές. Αν θέλουμε να γράψουμε το $x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$ ως γραμμή το συμβολίζουμε με x^T. $x^T=(x_1,\ldots,x_n)$ (διαβάζουμε x ανάστροφο)

- Θέλω η αριστερή πλευρά γραμμικού συστήματος να γράφεται ως $\mathbf{A}\cdot\mathbf{x}$.
- Άρα πρέπει να ορίσω πολλαπλασιασμό πίνακα επί διάνυσμα, τέτοιο που το αποτέλεσμα να δίνει την αριστερή πλευρά του συστήματος.

<u>Ορισμός (Πολλαπλασιασμός πίνακα με διάνυσμα):</u> Έστω $\mathbf{A}=\left[a_{ij}\right]$ ένας $m\times n$ πίνακας με *i*-γραμμή την $\mathbf{a}^{\mathbf{i}^{\mathrm{T}}}=\left(a_{i1},\ldots a_{in}\right)$. Και έστω $\mathbf{x}\in R^n$ $\mathbf{x}=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

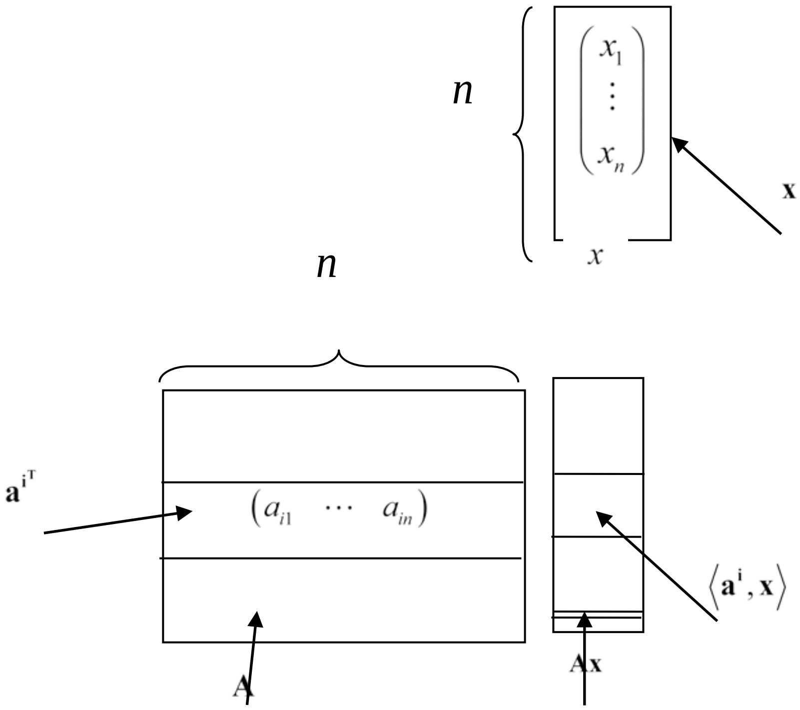
Ορίζουμε $\mathbf{A}\cdot\mathbf{x}\in R^m$ ως

$$\underbrace{\mathbf{A}\cdot\mathbf{x}=\begin{pmatrix} \left\langle \mathbf{a}^1,\mathbf{x}\right\rangle \\ \left\langle \mathbf{a}^2,\mathbf{x}\right\rangle \\ \vdots \\ \left\langle \mathbf{a}^{\mathbf{m}},\mathbf{x}\right\rangle \end{pmatrix}}_{m\ \text{συντεταγμένες}}=\underbrace{\begin{pmatrix} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\ldots \\ \vdots \\ a_{m1}x_1+a_{m2}x_2+\ldots \end{pmatrix}}_{\in R^m}$$

<u>Σημείωση:</u> Για να μπορώ να πολλαπλασιάσω $\mathbf{A}$ και $\mathbf{x}$ πρέπει:

 # στηλών του $\mathbf{A}$ = # συντεταγμένων του $\mathbf{x}$

Μνημονικός κανόνας



π.χ.

$$\begin{pmatrix}1\\3\\5\end{pmatrix}$$

$$A \quad \begin{pmatrix}2 & 1 & 1\\ 4 & -6 & 0\end{pmatrix} \quad \begin{pmatrix}2\cdot 1+1\cdot 3+1\cdot 5\\ 4\cdot 1-6\cdot 3+0\cdot 5\end{pmatrix} = \begin{pmatrix}10\\ -14\end{pmatrix}$$

Τί σχέση έχει το $\mathbf{Ax}$ με τις στήλες του $\mathbf{A}$;

Αν ονομάσουμε $\mathbf{a_i} = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} \in R^m$ την *i*-στήλη του $\mathbf{A}$ τότε

$$\mathbf{Ax} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots \\ a_{21}x_1 + a_{22}x_2 + \dots \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots \end{pmatrix} = \begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{pmatrix} + \begin{pmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{m2}x_2 \end{pmatrix} + \dots = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots$$

$$= x_1\mathbf{a_1} + x_2\mathbf{a_2} + \dots + x_n\mathbf{a}_n$$

Δηλ. **Το $\mathbf{Ax}$ είναι γραμμικός συνδυασμός των στηλών του $\mathbf{A}$ με συντελεστές τις συντεταγμένες του $\mathbf{x}$.**

Κανόνες πράξεων

- $\mathbf{A}\cdot(\mathbf{x}+\mathbf{y}) = \mathbf{A}\mathbf{x}+\mathbf{A}\mathbf{y}$
- $\mathbf{A}\cdot(\lambda\mathbf{x}) = \lambda\cdot(\mathbf{A}\mathbf{x})$
- $(\mathbf{A}+\mathbf{B})\cdot\mathbf{x} = \mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{x}$
- $(\lambda\mathbf{A})\cdot\mathbf{x} = \lambda\cdot(\mathbf{A}\mathbf{x})$

Παραδείγματα:

Ταυτοτικός, Μοναδιαίος

$$\underbrace{\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}}_{I_n} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \mathbf{I_n} \cdot \mathbf{x} = \mathbf{x} \quad \forall x$$

Διαγώνιος

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \\ \vdots \\ \lambda_n x_n \end{pmatrix}$$

Στοιχειώδης $\mathbf{E_{ij}}\left(\lambda\right)$

$$\begin{array}{ll} i- & \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots\lambda & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \\ & \qquad\qquad\qquad\quad | \\ & \qquad\qquad\qquad\quad j \end{array}$$

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots\lambda & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ x_i + \lambda x_j \\ x_{i+1} \\ \vdots \\ x_n \end{pmatrix}$$

1. Έστω $\mathbf{K} = m\left\{\left[\underbrace{\mathbf{A}_{m\times n}}_{n}\middle|\underbrace{\mathbf{B}_{m\times q}}_{q}\right]\right.$ και $\mathbf{x} = \begin{matrix} n \\ q \end{matrix}\left\{\left[\begin{array}{c} \mathbf{z}_{n\times 1} \\ \hline \underbrace{\mathbf{w}_{q\times 1}}_{1} \end{array}\right]\right.$. Να υπολογισθεί το $\mathbf{K}\cdot\mathbf{x}$ ώς συνάρτηση γινομένων των πινάκων $\mathbf{A},\mathbf{B}$ με τα $\mathbf{z},\mathbf{w}$.

(πρώτα m=3,n=2,q=1 μετά γενικά…)

Π.χ. $\mathbf{K} = \begin{pmatrix} a_{11} & a_{12} & b_{11} \\ a_{21} & a_{22} & b_{21} \\ a_{31} & a_{32} & b_{31} \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} z_1 \\ z_2 \\ w_1 \end{pmatrix}$

$$\mathbf{Kx} = \begin{pmatrix} \langle \mathbf{a}^1, \mathbf{z}\rangle + \langle \mathbf{b}^1, \mathbf{w}\rangle \\ \langle \mathbf{a}^2, \mathbf{z}\rangle + \langle \mathbf{b}^2, \mathbf{w}\rangle \\ \langle \mathbf{a}^3, \mathbf{z}\rangle + \langle \mathbf{b}^3, \mathbf{w}\rangle \end{pmatrix} = \mathbf{Az} + \mathbf{Bw}$$

2. Έστω $\mathbf{L} = \begin{array}{c} r \\ s \end{array}\left\{ \underbrace{\begin{bmatrix} \mathbf{C}_{r\times p} \\ \hline \mathbf{D}_{s\times p} \end{bmatrix}}_{p} \right.$ και $\mathbf{y} = p\left\{ \underbrace{\begin{bmatrix} * \end{bmatrix}}_{1} \right.,$. Να υπολογισθεί το $\mathbf{L}\cdot\mathbf{y}$ ώς συνάρτηση γινομένων των πινάκων $\mathbf{C},\mathbf{D}$ με το $\mathbf{y}$.

Π.χ. $\mathbf{L} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$$\mathbf{Ly} = \begin{pmatrix} \left(\mathbf{c}^1\right)^T \\ \vdots \\ \left(\mathbf{c}^r\right)^T \\ \left(\mathbf{d}^1\right)^T \\ \vdots \\ \left(\mathbf{d}^s\right)^T \end{pmatrix}\mathbf{y} = \begin{pmatrix} \left\langle \mathbf{c}^1, \mathbf{y} \right\rangle \\ \vdots \\ \left\langle \mathbf{c}^r, \mathbf{y} \right\rangle \\ \left\langle \mathbf{d}^1, \mathbf{y} \right\rangle \\ \vdots \\ \left\langle \mathbf{d}^s, \mathbf{y} \right\rangle \end{pmatrix} = \begin{pmatrix} \mathbf{Cy} \\ \mathbf{Dy} \end{pmatrix}$$

Επίσης **ορίζουμε πολλαπλασιασμό πινάκων μεταξύ τους**.

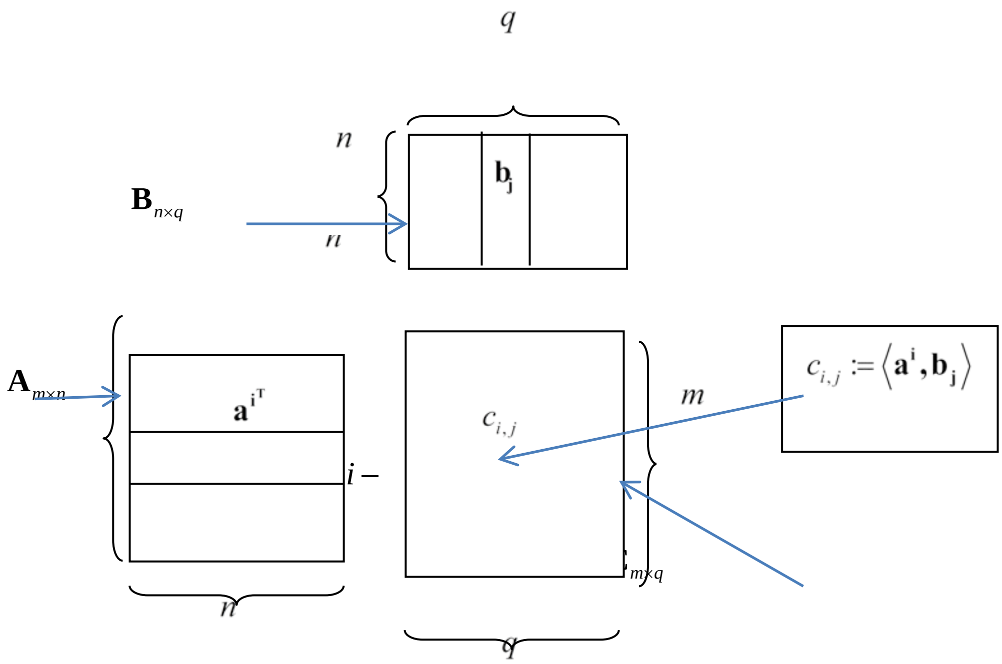
Ορισμός (Πολλαπλασιασμός πίνακα με πίνακα): Έστω $\mathbf{A}$ $m \times n$ πίνακας και $\mathbf{B}$ ένας $n \times q$ πίνακας. Ο πίνακας $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ θα είναι $m \times q$ και ορίζεται ως $\mathbf{C}_{ij} = \left\langle \mathbf{a}^{\mathbf{i}}, \mathbf{b}_{\mathbf{j}} \right\rangle$ (εσωτερικό γινόμενο *i*-γραμμής του $\mathbf{A}$ με *j*-στήλη του $\mathbf{B}$.

Σημείωση: Για να πολλαπλασιάσω $\mathbf{A}$ με $\mathbf{B}$ πρέπει

 # στηλών $\mathbf{A}$ = # γραμμών $\mathbf{B}$

Μνημονικός κανόνας

Για να πολλαπλασιάσω $A \cdot B$:



Παράδειγμα:

$$\begin{matrix} 0 & 1 \\ 2 & 0 \\ 1 & 1 \end{matrix}$$

$$\mathbf{A} \quad \begin{matrix} 1 & 0 & 2 \\ 3 & 1 & 1 \end{matrix} \qquad \begin{matrix} 2 & 3 \\ 3 & 4 \end{matrix} = \mathbf{A}\cdot\mathbf{B}$$

$$1\cdot 0+0\cdot 2+2\cdot 1 \qquad 3\cdot 1+1\cdot 0+1\cdot 1$$

Παράδειγμα:

- $\underline{x^T y = ?}$

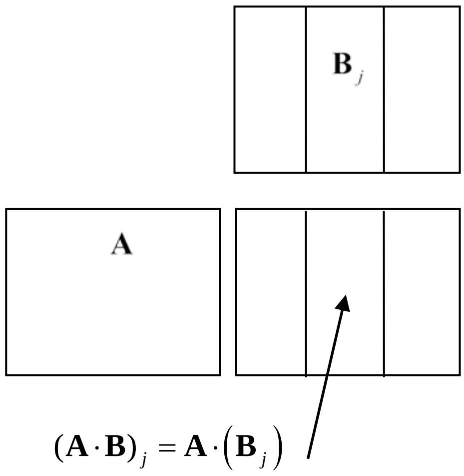
- $\underline{x y^T = ?}$

- $\underline{x^T y = \langle x, y \rangle \in R}$

$$\underline{xy^T = \begin{pmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ x_n y_1 & \cdots & \cdots & x_n y_n \end{pmatrix}}$$

• <u>Παρατήρηση:</u> Η «j-στήλη του $\mathbf{A}\cdot\mathbf{B}$» ισούται με $A\cdot$ «j-στήλη του B».

$$(\mathbf{A}\cdot\mathbf{B})_j = \mathbf{A}\cdot\left(\mathbf{B}_j\right)$$



Κανόνες πράξεων

- Έστω $\mathbf{A}_{m\times n}, \mathbf{E}_{m\times n}, \quad \mathbf{B}_{n\times q}, \mathbf{D}_{n\times q}, \quad \mathbf{C}_{q\times s}$
- $(\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C} = \mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})$
- $\mathbf{A}\cdot(\mathbf{B}+\mathbf{D}) = \mathbf{A}\cdot\mathbf{B}+\mathbf{A}\cdot\mathbf{D}$
- & $(\mathbf{A}+\mathbf{E})\cdot\mathbf{B} = \mathbf{A}\cdot\mathbf{B}+\mathbf{E}\cdot\mathbf{B}$
- $(\boldsymbol{\lambda}\cdot\mathbf{A})\cdot(\boldsymbol{\mu}\cdot\mathbf{B}) = \boldsymbol{\lambda}\cdot\boldsymbol{\mu}\cdot\left(\mathbf{A}\cdot\mathbf{B}\right)$

<u>Προσοχή:</u> Γενικώς $\mathbf{A}\cdot\mathbf{B} \neq \mathbf{B}\cdot\mathbf{A}$ $n\times n$!!!

π.χ.

$$\begin{array}{cc} & \boxed{\begin{matrix}0 & 0\\ 1 & 0\end{matrix}} \\ \boxed{\begin{matrix}1 & 2\\ 0 & 3\end{matrix}} & \boxed{\begin{matrix}2 & 0\\ 3 & 0\end{matrix}} \end{array} \qquad \varepsilon \qquad \begin{array}{cc} & \boxed{\begin{matrix}1 & 2\\ 0 & 3\end{matrix}} \\ \boxed{\begin{matrix}0 & 0\\ 1 & 0\end{matrix}} & \boxed{\begin{matrix}0 & 0\\ 1 & 2\end{matrix}} \end{array}$$

Παραδείγματα:

Μοναδιαίος

$$\mathbf{I}_n = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix} \qquad \boxed{\mathbf{I}_n \cdot \mathbf{A} = ?} \qquad \boxed{\mathbf{A} \cdot \mathbf{I}_n = ?}$$

$$I_n = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix} \Rightarrow (I_n \cdot A)_j = I_n \cdot a_j = a_j \Rightarrow \boxed{\mathbf{I}_n \cdot \mathbf{A} = \mathbf{A}}$$

και $(A \cdot I_n)_j = A \cdot e_j = a_j \Rightarrow \boxed{\mathbf{A} \cdot \mathbf{I}_n = \mathbf{A}}$

Διαγώνιος

Από αριστερά

$$\begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \vdots \\ \vdots & \cdots & \ddots \end{bmatrix}$$

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ \vdots & \lambda_2 & \vdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{bmatrix}$$

Από δεξιά

$$\begin{bmatrix} \lambda_1 & \cdots & \cdots \\ \vdots & \lambda_2 & \vdots \\ \vdots & \cdots & \ddots \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots \\ a_{21} & a_{22} & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \cdots & \cdots & \ddots \end{bmatrix}$$

Από αριστερά

$$\begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \vdots \\ \vdots & \cdots & \ddots \end{bmatrix}$$

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ \vdots & \lambda_2 & \vdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \lambda_1 a_{11} & \lambda_1 a_{12} & \cdots \\ \lambda_2 a_{21} & \lambda_2 a_{22} & \vdots \\ \vdots & \cdots & \ddots \end{bmatrix}$$

Από δεξιά

$$\begin{bmatrix} \lambda_1 & \cdots & \cdots \\ \vdots & \lambda_2 & \vdots \\ \vdots & \cdots & \ddots \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots \\ a_{21} & a_{22} & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \cdots & \cdots & \ddots \end{bmatrix} \begin{bmatrix} \lambda_1 a_{11} & \lambda_2 a_{12} & \cdots \\ \lambda_1 a_{21} & \lambda_2 a_{22} & \vdots \\ \vdots & \cdots & \ddots \end{bmatrix}$$

Το λ_i συντελεστής της *i-γραμμής*   Το λ_i συντελεστής της *i-στήλης*

Στοιχειώδης $\mathbf{E_{ij}}(\lambda)$

<u>*Από αριστερά*</u>

$$\begin{bmatrix} a_{11} & & \\ a_{21} & a_2 & \cdots \\ \vdots & & \\ a_{m1} & & \end{bmatrix}$$

$$\begin{array}{c c} & \begin{bmatrix} 1 & \dots & \dots & \dots \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \lambda & \ddots & \vdots \\ \vdots & \dots & \dots & 1 \end{bmatrix} \end{array}$$

$$\begin{array}{c} i- \end{array}$$

$$\begin{bmatrix} a_{11} & & \\ a_{21} & a_2 & \cdots \\ \vdots & & \\ a_{m1} & & \end{bmatrix}$$

$$i-\begin{bmatrix} 1 & \dots & \dots & \dots \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \lambda & \ddots & \vdots \\ \vdots & \dots & \dots & 1 \end{bmatrix}\begin{bmatrix} a_{11} & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots \\ a_{i-1,1} & \ddots & \ddots & \vdots \\ a_{i,1}+\lambda\alpha_{j1} & \ddots & \ddots & \vdots \\ a_{i+1,1} & \vdots & \vdots & \vdots \\ \vdots & \dots & \dots & \dots \end{bmatrix}=\begin{bmatrix} \vdots \\ a^{i-1^T} \\ a^{i^T}+\lambda a^{j^T} \\ a^{i+1^T} \\ \vdots \end{bmatrix}$$

*Προσθέτει λ-πλάσιο της <u>j-**γραμμής**</u> στην <u>i-**γραμμή**</u>.*

<u>Από δεξιά</u>

$$i - \begin{bmatrix} 1 & \cdots & \cdots & \cdots \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \lambda & \ddots & \vdots \\ \vdots & \cdots & \cdots & 1 \end{bmatrix}$$

a

$$\begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \vdots \\ \vdots & \cdots & \ddots \end{bmatrix}$$

j

$\uparrow$

$a_j + \lambda a_i$

*Προσθέτει λ-πλάσιο της <u>i-**στήλης**</u> στην <u>j-**στήλη**</u>.*

Πίνακες μεταθέσεων

Έστω $\mathbf{P}_{\mathrm{ij}}$ προκύπτει από μοναδιαίο με εναλλαγή *i*- και *j*-γραμμής.

Τότε

- $\mathbf{P}_{\mathrm{ij}} \cdot \mathbf{A}$ εναλλάσσει *i*- & *j*-γραμμή του A
- $\mathbf{A} \cdot \mathbf{P}_{\mathrm{ij}}$ εναλλάσσει *i*- & *j*-στήλη του A.

Έστω $\mathbf{K} = \left(\mathbf{A} \quad \middle| \mathbf{B}\right)$, $\mathbf{A}$ $m \times r$, $\mathbf{B}$ $m \times s$,

$\mathbf{H} = \begin{pmatrix} \mathbf{F} \\ \hline \mathbf{G} \end{pmatrix}$, $\mathbf{F}$ $r \times p$, $\mathbf{G}$ $s \times p$

Τότε	$\mathbf{KH} = \mathbf{AF} + \mathbf{BG}$

Υπενθυμ. Αν $\mathbf{K} = \left[\mathbf{A} \middle| \mathbf{B}\right]$ και $\mathbf{x} = \begin{pmatrix} \mathbf{z} \\ \hline \mathbf{w} \end{pmatrix}$, τότε $\mathbf{Kx} = \mathbf{Az} + \mathbf{Bw}$

Παρατήρηση: $\left(\mathbf{H}\right)_i = \mathbf{h}_i = \begin{pmatrix} \left(\mathbf{F}\right)_i \\ \hline \left(\mathbf{G}\right)_i \end{pmatrix} = \begin{pmatrix} \mathbf{f}_i \\ \hline \mathbf{g}_i \end{pmatrix}$

$$\left(\mathbf{KH}\right)_i = \mathbf{K}\left(\mathbf{H}\right)_i = \mathbf{Kh}_i = \mathbf{K}\begin{pmatrix} \mathbf{f}_i \\ \hline \mathbf{g}_i \end{pmatrix} = \mathbf{Af}_i + \mathbf{Bg}_i = \left(\mathbf{AF}\right)_i + \left(\mathbf{BG}\right)_i$$

Έστω $\mathbf{L} = \begin{pmatrix} \mathbf{C} \\ \hline \mathbf{D} \end{pmatrix}$, $\mathbf{C}$ $r \times p$, $\mathbf{D}$ $s \times p$ και $\mathbf{F}$ $r \times p$.

Τότε $\mathbf{LF} = \begin{pmatrix} \mathbf{CF} \\ \hline \mathbf{DF} \end{pmatrix}$

Υπενθυμ.: Έστω $\mathbf{L} = \begin{pmatrix} \mathbf{C} \\ \hline \mathbf{D} \end{pmatrix}$, $\mathbf{C}$ $r \times p$, $\mathbf{D}$ $s \times p$ και $\mathbf{y} \in \mathbb{R}^p$. Τότε $\mathbf{Ly} = \begin{pmatrix} \mathbf{Cy} \\ \mathbf{Dy} \end{pmatrix}$

$$(\mathbf{LF})_i = \mathbf{L}(\mathbf{F})_i = \mathbf{Lf}_i = \begin{pmatrix} \mathbf{Cf}_i \\ \hline \mathbf{Df}_i \end{pmatrix} = \begin{pmatrix} \mathbf{C}(\mathbf{F})_i \\ \hline \mathbf{D}(\mathbf{F})_i \end{pmatrix} = \begin{pmatrix} (\mathbf{CF})_i \\ \hline (\mathbf{DF})_i \end{pmatrix} = \begin{pmatrix} \mathbf{CF} \\ \hline \mathbf{DF} \end{pmatrix}_i$$

Έστω $\mathbf{A}\ m \times r\ ,$

$\mathbf{P} = \left(\mathbf{M} \quad \middle| \mathbf{N}\right)\ ,\ \mathbf{M}\ r \times p\ ,\ \mathbf{N}\ r \times q\ ,$

Τότε $\mathbf{AP} = \left(\mathbf{AM} \quad \middle| \mathbf{AN}\right)$

$$i \le p\ :\ \left(\mathbf{AP}\right)_i = \mathbf{A}\left(\mathbf{P}\right)_i = \mathbf{A}\left(\mathbf{M}\right)_i = \left(\mathbf{AM} \quad \middle| \mathbf{AN}\right)_i$$

$$i > p\ :\ \left(\mathbf{AP}\right)_i = \mathbf{A}\left(\mathbf{P}\right)_i = \mathbf{A}\left(\mathbf{N}\right)_{i-p} = \left(\mathbf{AM} \quad \middle| \mathbf{AN}\right)_i$$

Συνδυάζοντας τα προηγούμενα:

Έστω

$$\mathbf{A}=\left(\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \hline \mathbf{A}_{21} & \mathbf{A}_{22}\end{array}\right),\ \mathbf{A}_{11}\ \mathrm{r}\times\mathrm{p},\ \mathbf{A}_{12}\ \mathrm{r}\times\mathrm{q},\ \mathbf{A}_{21}\ \mathrm{s}\times\mathrm{p},\ \mathbf{A}_{22}\ \mathrm{s}\times\mathrm{q}\ \text{ και}$$

$$\mathbf{B}=\left(\begin{array}{c|c}\mathbf{B}_{11} & \mathbf{B}_{12}\\ \hline \mathbf{B}_{21} & \mathbf{B}_{22}\end{array}\right),\ \mathbf{B}_{11}\ \mathrm{p}\times\mathrm{u},\ \mathbf{B}_{12}\ \mathrm{p}\times w,\ \mathbf{B}_{21}\ \mathrm{q}\times u,\ \mathbf{B}_{22}\ \mathrm{q}\times w$$

Τότε $$\mathbf{AB}=\left(\begin{array}{c|c}\mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12}+\mathbf{A}_{12}\mathbf{B}_{22}\\ \hline \mathbf{A}_{21}\mathbf{B}_{11}+\mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12}+\mathbf{A}_{22}\mathbf{B}_{22}\end{array}\right),$$

$$\mathbf{AB}=\left(\begin{array}{c}\left(\mathbf{A}_{11}\middle|\ \mathbf{A}_{12}\right)\left(\begin{array}{c}\left(\mathbf{B}_{11}\middle|\ \mathbf{B}_{12}\right)\\ \left(\mathbf{B}_{21}\middle|\ \mathbf{B}_{22}\right)\end{array}\right)\\ \hline \left(\mathbf{A}_{21}\middle|\ \mathbf{A}_{22}\right)\left(\begin{array}{c}\left(\mathbf{B}_{11}\middle|\ \mathbf{B}_{12}\right)\\ \left(\mathbf{B}_{21}\middle|\ \mathbf{B}_{22}\right)\end{array}\right)\end{array}\right)=\left(\begin{array}{c}\mathbf{A}_{11}\left(\mathbf{B}_{11}\middle|\ \mathbf{B}_{12}\right)+\mathbf{A}_{12}\left(\mathbf{B}_{21}\middle|\ \mathbf{B}_{22}\right)\\ \hline \mathbf{A}_{21}\left(\mathbf{B}_{11}\middle|\ \mathbf{B}_{12}\right)+\mathbf{A}_{22}\left(\mathbf{B}_{21}\middle|\ \mathbf{B}_{22}\right)\end{array}\right)$$

$$=\left(\begin{array}{c}\left(\mathbf{A}_{11}\mathbf{B}_{11}\middle|\ \mathbf{A}_{11}\mathbf{B}_{12}\right)+\left(\mathbf{A}_{12}\mathbf{B}_{21}\middle|\ \mathbf{A}_{12}\mathbf{B}_{22}\right)\\ \hline \left(\mathbf{A}_{21}\mathbf{B}_{11}\middle|\ \mathbf{A}_{21}\mathbf{B}_{12}\right)+\left(\mathbf{A}_{22}\mathbf{B}_{21}\middle|\ \mathbf{A}_{22}\mathbf{B}_{22}\right)\end{array}\right)$$

$$=\left(\begin{array}{c}\left(\mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{21}\middle|\ \mathbf{A}_{11}\mathbf{B}_{12}+\mathbf{A}_{12}\mathbf{B}_{22}\right)\\ \hline \left(\mathbf{A}_{21}\mathbf{B}_{11}+\mathbf{A}_{22}\mathbf{B}_{21}\middle|\ \mathbf{A}_{21}\mathbf{B}_{12}+\mathbf{A}_{22}\mathbf{B}_{22}\right)\end{array}\right)$$

6. Γραμμικά Συστήματα: Ο Αλγόριθμος Gauss και η παραγοντοποίηση PA=LDU

(Πηγή: Strang, Κεφάλαιο 1.3 και 1.5)

Έστω το ακόλουθο γραμμικό σύστημα:

$$\begin{aligned}
&2x_1 + 1x_2 + 1x_3 = 5 \\
&4x_1 - 6x_2 = -2 \\
&-2x_1 + 7x_2 + 2x_3 = 9
\end{aligned}$$

<u>Στόχος:</u> Να βρω τις λύσεις, δηλαδή τα $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ που ικανοποιούν και τις τρεις εξισώσεις.

Η διαδικασία επίλυσης **βασίζεται στην αρχή ότι**:

Οι **λύσεις του συστήματος δεν αλλάζουν** αν

- αλλάξω τη σειρά των εξισώσεων

- πολλαπλασιάσω μία εξίσωση με λ

$$\langle \mathbf{a}, \mathbf{x} \rangle = b \Leftrightarrow \langle \lambda \mathbf{a}, \mathbf{x} \rangle = \lambda b$$

Π.χ. $4x_1 - 6x_2 = -2 \Leftrightarrow 8x_1 - 12x_2 = -4$

- προσθέσω σε μία εξίσωση πολλαπλάσιο μίας άλλης εξίσωσης

$$\left.\begin{matrix} \langle \mathbf{a}^1, \mathbf{x} \rangle = b_1 \\ \langle \mathbf{a}^2, \mathbf{x} \rangle = b_2 \end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix} \langle \mathbf{a}^1, x \rangle = b_1 \\ \langle \mathbf{a}^1 + \lambda \mathbf{a}^2, x \rangle = b_1 + \lambda b_2 \end{matrix}\right.$$

Δηλ.: αν $\mathbf{x}$ ικανοποιεί τις δύο εξισώσεις αριστερά θα ικανοποιεί και τις δύο δεξιά και αντίστροφα.

Π.χ.

$$\left\{\begin{matrix} 2x_1 + 1x_2 + 1x_3 = 5 \\ 4x_1 - 6x_2 = -2 \end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix} 2x_1 + 1x_2 + 1x_3 = 5 \\ 8x_1 - 12x_2 + 1x_3 = -4 \end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix} 2x_1 + 1x_2 + 1x_3 = 5 \\ 10x_1 - 11x_2 = 1 \end{matrix}\right\}$$

Ο αλγόριθμος Gauss προτείνει ένα συγκεκριμένο τρόπο να κάνω τέτοιες «γραμμοπράξεις» ώστε να καταλήξω σε λύση.

Κατ' αρχάς το σύστημα γράφεται ως

$$\mathbf{Ax} = \mathbf{b} \text{ με}$$

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \text{ και } \mathbf{b} = \begin{pmatrix} 5 \\ -2 \\ 9 \end{pmatrix}.$$

Αλγόριθμος Gauss

Πρώτα παίρνουμε τον $\mathbf{A}$ και κολλάμε από δεξιά το $\mathbf{b}$.

$\rightarrow [\mathbf{A} \mid \mathbf{b}]$. Στο παράδειγμά μας σχηματίζεται ο $\left[\begin{array}{ccc:c} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{array}\right]$.

Ακολούθως **κάνουμε γραμμοπράξεις στον «διευρυμένο» $\mathbf{A}$ με την εξής σειρά**:

<u>1^ο βήμα:</u>

Αν $a_{11} \neq 0$ το ονομάζουμε **πρώτο οδηγό**. Προσθέτουμε κατάλληλα **<u>πολλαπλάσια της πρώτης γραμμής</u>** (την οποία κρατάμε αναλλοίωτη) στις επόμενες, έτσι ώστε <u>να μηδενιστούν τα στοιχεία κάτω από τον πρώτο οδηγό</u>.

<u>Δηλαδή:</u> στην i γραμμή προσθέτω $-\dfrac{a_{i1}}{a_{11}} \cdot$ της πρώτης γραμμής.

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{array}\right] \xrightarrow[\substack{\gamma_2' = \gamma_2 - 2\gamma_1 \\ \gamma_3' = \gamma_3 + \gamma_1}]{} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{array}\right]$$

<u>2^ο βήμα:</u>

Αν το δεύτερο στοιχείο της διαγωνίου (που προέκυψε μετά το 1^ο βήμα) $a_{22}' \neq 0$ το ονομάζω **δεύτερο οδηγό**. Κρατάω τις δύο πρώτες γραμμές και προσθέτω κατάλληλα **<u>πολλαπλάσια της</u>** <u>δεύτερης γραμμής</u> (της **<u>γραμμής του οδηγού</u>**!!) στις επόμενες ώστε <u>να μηδενιστούν τα στοιχεία κάτω από το δεύτερο οδηγό</u>.

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{array}\right] \xrightarrow[\gamma_3'' = \gamma_3' + \gamma_2']{} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

Η διαδικασία αυτή συνεχίζεται (υπό την προϋπόθεση ότι τα διαγώνια στοιχεία που συναντώ είναι $\neq 0$) **μέχρι να φθάσω δεξιά σε άνω τριγωνικό πίνακα** (τα στοιχεία κάτω από τη διαγώνιο είναι 0) και πάνω στη διαγώνιο έχει τους οδηγούς. Αυτόν τον ονομάζω $\mathbf{U}$.

Στο παράδειγμά μας, όπου ο A είναι 3×3 άρκεσαν δύο βήματα. Αν $\mathbf{A}_{\mathbf{n}\times\mathbf{n}}$ χρειάζονται $n-1$ βήματα του Αλγορίθμου Gauss.

Παράλληλα το $\mathbf{b}$ έχει μετασχηματιστεί σε κάποιο $\mathbf{c}$:

$$\begin{bmatrix}\mathbf{A} & \mathbf{b}\end{bmatrix} \to \begin{bmatrix}\mathbf{U} & \mathbf{c}\end{bmatrix}.$$

Οι λύσεις του $\boxed{\mathbf{Ax}=\mathbf{b}}$ είναι ακριβώς οι λύσεις του $\boxed{\mathbf{Ux}=\mathbf{c}}$. Το πλεονέκτημα είναι τώρα ότι μετατρέποντας το σύστημα σε τριγωνικό, μπορούμε να το λύσουμε εύκολα με «**ανάδρομη αντικατάσταση**»:

$\boxed{Ux=c}$: $2x_1+1x_2+1x_3=5$

$$-8x_2-2x_3=-12$$

$$x_3=2$$

Αντικαθιστώντας την τελευταία στην προτελευταία παίρνουμε:

$$-8x_2-4=-12 \Leftrightarrow x_2=1$$

και αντικαθιστώντας αυτά στην πρώτη:

$$2x_1+1+2=5 \Leftrightarrow x_1=1$$

Συμπέρασμα: Αν για κάποιο $n\times n$ πίνακα A κατά την εφαρμογή του Αλγορίθμου Gauss δε συναντήσω μηδενικό και βρω, λοιπόν, **ένα πλήρες σύστημα οδηγών (δηλ.** n **οδηγούς** $\neq 0$**),** τότε **το σύστημα** $Ax=b$ **θα έχει** **ακριβώς μία λύση για οποιοδήποτε** b**. Ονομάζω αυτή την περίπτωση μη – ιδιόμορφη.**

Η διάσπαση A=L·U

Στο 1ο βήμα του Gauss μετατρέψαμε με γραμμοπράξεις τον $\mathbf{A}$ σε $\mathbf{A}'$.

$$\left[\begin{array}{ccc:c} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{array}\right] \xrightarrow[\gamma_3' = \gamma_3 + \gamma_1]{\gamma_2' = \gamma_2 - 2\gamma_1} \left[\begin{array}{cccc} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{array}\right]$$

Οι γραμμοπράξεις αυτές μπορούν να εκφραστούν με πολλαπλασιασμούς του A με στοιχειώδεις πίνακες από αριστερά:

$$\mathbf{A}' = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}}_{\gamma_3' = \gamma_3 + \gamma_1} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\gamma_2' = \gamma_2 - 2\gamma_1} \mathbf{A}$$

Στο 2ο βήμα πήραμε από τον $\mathbf{A}'$ τον $\mathbf{U}$ πάλι με γραμμοπράξεις:

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{array}\right] \xrightarrow{\gamma_3''=\gamma_3'+\gamma_2'} \left[\begin{array}{cccc|} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

Οι γραμμοπράξεις αυτές μπορούν να εκφραστούν με πολλαπλασιασμούς του A με στοιχειώδεις πίνακες από αριστερά

$$\mathbf{U}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\mathbf{A}'$$

$$\gamma_3''=\gamma_3'+\gamma_2'$$

Συνολικά λοιπόν παίρνουμε:

$$\mathbf{U}=\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}}\cdot\mathbf{A}=\mathbf{L}'\mathbf{A}$$

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

Στο γινόμενο $\mathbf{L}'=\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$ **οι συντελεστές <u>«χάνονται» γενικώς</u>.** (Εδώ είναι σύμπτωση ότι κάποιοι διατηρήθηκαν).

> Για να **ακυρώσω ένα μετασχηματισμό** $\mathbf{a}^{\mathbf{i}} \to \mathbf{a}^{\mathbf{i}}+\lambda\mathbf{a}^{\mathbf{j}}$ αρκεί να τον «ξανακάνω» με <u>αντίθετο πρόσημο του λ</u>: $\mathbf{a}^{\mathbf{i}}+\lambda\mathbf{a}^{\mathbf{j}}-\lambda\mathbf{a}^{\mathbf{j}} \to \mathbf{a}^{\mathbf{i}}$.

Αν **ακυρώσω τον τελευταίο** μετασχηματισμό που έκανα από

$$\mathbf{U} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{A}$$

παίρνω

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

και **αν ακυρώσω έναν – έναν και τους υπόλοιπους**

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}} \mathbf{U} = \mathbf{A}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} =: \mathbf{L}$$

Οι στοιχειώδεις πίνακες εμφανίζονται με **ανάποδη σειρά** και περιέχουν τα **αντίθετα των συντελεστών** του αλγορίθμου Gauss.

Αν σχηματίσω τώρα το γινόμενο $\mathbf{L}=\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ -1 & -1 & 1\end{bmatrix}$ θα δω ότι:

τα αντίθετα των συντελεστών διατηρούνται στον L και βρίσκονται στην ίδια θέση με το στοιχείο που «μηδένισαν» στον $\mathbf{A}$ (για να προκύψει ο U).

$$\mathbf{U}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\cdot\mathbf{A}$$

Αυτό δεν είναι σύμπτωση: αποδεικνύεται ότι ισχύει γενικά. Επομένως:

Αν για κάποιο $n\times n$ πίνακα $\mathbf{A}$ κατά τον Αλγόριθμο Gauss **δε συναντήσω μηδενικά**, αν έχω δηλαδή πλήρες σύστημα οδηγών, τότε θα υπάρχουν

- **κάτω τριγωνικός πίνακας** $\mathbf{L}$ $n\times n$ με «1» στη διαγώνιο και τα αντίθετα των συντελεστών κάτω από τη διαγώνιο και

- **άνω τριγωνικός** $\mathbf{U}$ $n\times n$ με τους οδηγούς στη διαγώνιο,

τέτοιοι ώστε

$$\mathbf{A}=\mathbf{LU}.$$

Κατασκευή του L:

Τον φτιάχνω παράλληλα με τον αλγόριθμο Gauss θέτοντας το αντίθετο του συντελεστή που χρησιμοποιώ για να μηδενίσω ένα στοιχείο του $\mathbf{A}$ στην αντίστοιχη θέση του $\mathbf{L}$.

Παράδειγμα:

$$\begin{bmatrix} 2 & 0 & 2 & 4 \\ 4 & 2 & 2 & 0 \\ 4 & 4 & 2 & 0 \\ 6 & 4 & 2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 0 & 2 & 4 \\ 0 & 2 & -2 & -8 \\ 0 & 4 & -2 & -8 \\ 0 & 4 & -4 & -10 \end{bmatrix} \qquad \left(\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{bmatrix} \right)$$

$$\rightarrow \begin{bmatrix} 2 & 0 & 2 & 4 \\ 0 & 2 & -2 & -8 \\ 0 & 0 & 2 & 8 \\ 0 & 0 & 0 & -6 \end{bmatrix} \qquad \left(\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 \\ 3 & 2 & 0 & 1 \end{bmatrix} \right)$$

Άσκηση. Να βρεθεί η παραγοντοποίηση $\mathbf{A} = \mathbf{LU}$ για το πίνακα

$$\mathbf{A} = \begin{bmatrix} 2 & -2 & 2 & 1 \\ -3 & 6 & 0 & -1 \\ 1 & -7 & 10 & 2 \\ 0 & 12 & 15 & 3 \end{bmatrix}$$

Επίλυση του Ax=b όταν γνωρίζω τη διάσπαση A=LU

Αν αφού έχω κάνει τη διάσπαση $\mathbf{A}=\mathbf{LU}$ μου δοθεί η δεξιά πλευρά b του συστήματος, δεν αξίζει να «ξανακάνω» αλγόριθμο Gauss.

Λύνω το σύστημα $\mathbf{Ax}=\mathbf{b}$ στα εξής δύο βήματα:

$$\mathbf{Ax}=\mathbf{b} \Leftrightarrow \underbrace{\mathbf{LU}}_{A}\mathbf{x}=\mathbf{b} \Leftrightarrow \mathbf{L}\underbrace{\mathbf{Ux}}_{c}=\mathbf{b}$$

1) Βρες $\mathbf{c}$: $\mathbf{Lc}=\mathbf{b}$

2) Βρες $\mathbf{x}$: $\mathbf{Ux}=\mathbf{c}$

Και τα δύο συστήματα είναι τριγωνικά και λύνονται εύκολα.

π.χ. Να λυθεί το $\mathbf{Ax}=\begin{pmatrix}1\\0\\1\end{pmatrix}$.

Υπενθύμ. $\mathbf{A}=\mathbf{LU}$, με $\mathbf{L}=\begin{bmatrix}1&0&0\\2&1&0\\-1&-1&1\end{bmatrix}$ και $\mathbf{U}=\begin{bmatrix}2&1&1\\0&-8&-2\\0&0&1\end{bmatrix}$

1) Βρες c: $Lc=b$

2) Βρες x: $Ux=c$

1) Βρες c: $Lc = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

$\boxed{c_1 = 1}$

$2c_1 + c_2 = 0 \Rightarrow \boxed{c_2 = -2}$

$-c_1 - c_2 + c_3 = 0 \Rightarrow \boxed{c_3 = 0}$

2) Βρες x: $Ux = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$

$\boxed{x_3 = 0}$

$-8x_2 - 2x_3 = -2 \Rightarrow \boxed{x_2 = \frac{1}{4}}$

$2x_1 + x_2 + x_3 = 1 \Rightarrow \boxed{x_1 = \frac{3}{8}}$

Άσκηση για το σπίτι (για αύριο). Να βρεθεί η παραγοντοποίηση $\mathbf{A} = \mathbf{LU}$ για το πίνακα

$\mathbf{A} = \begin{bmatrix} 1 & -3 & 5 \\ 2 & -4 & 7 \\ -1 & -2 & 1 \end{bmatrix}$ και να λυθεί το σύστημα $\mathbf{Ax} = \mathbf{b}$, με $\mathbf{b} = \begin{pmatrix} 10 \\ 22 \\ -12 \end{pmatrix}$, λύνοντας τα δύο τριγωνικά $\mathbf{Lc} = \mathbf{b}$ και $\mathbf{Ux} = \mathbf{c}$.

Παρατήρηση: ο $\mathbf{U}$ γράφεται ως $\mathbf{U}=\mathbf{D}\tilde{\mathbf{U}}$ με $\mathbf{D}$ διαγώνιο με τους οδηγούς στη διαγώνιο και $\tilde{\mathbf{U}}$ άνω τριγωνικός με 1 στη διαγώνιο.

Αν $\mathbf{U}=\begin{bmatrix} u_{11} & u_{12} & \cdots & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & \cdots & u_{2n} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & u_{nn} \end{bmatrix}$ θέτουμε

$\mathbf{D}=\begin{bmatrix} u_{11} & 0 & \cdots & \cdots & 0 \\ 0 & u_{22} & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & u_{nn} \end{bmatrix}$ και $\tilde{\mathbf{U}}=\begin{bmatrix} 1 & \frac{u_{12}}{u_{11}} & \cdots & \cdots & \frac{u_{1n}}{u_{11}} \\ 0 & 1 & \frac{u_{23}}{u_{22}} & \cdots & \frac{u_{2n}}{u_{22}} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}$ δηλ

διαιρούμε κάθε γραμμή του $\mathbf{U}$ με το διαγώνιο στοιχείο του.

Παρατήρηση: Η παραγοντοποίηση $\mathbf{A}=\mathbf{L}\mathbf{D}\tilde{\mathbf{U}}$ είναι μοναδική

Δηλαδή αν υπήρχαν δύο τέτοιες διασπάσεις $\mathbf{A}=\mathbf{L}_1\mathbf{D}_1\tilde{\mathbf{U}}_1$ και $\mathbf{A}=\mathbf{L}_2\mathbf{D}_2\tilde{\mathbf{U}}_2$ με $\mathbf{L}_1,\mathbf{L}_2$ κάτω τριγωνικούς με «1» στη διαγώνιο, $\mathbf{D}_1,\mathbf{D}_2$ διαγώνιους και $\tilde{\mathbf{U}}_1,\tilde{\mathbf{U}}_2$ πάνω τριγωνικούς με «1» στη διαγώνιο, τότε αυτές θα πρέπει να είναι ταυτόσημες:

$$\mathbf{L}_1=\mathbf{L}_2,\ \mathbf{D}_1=\mathbf{D}_2 \text{ και } \tilde{\mathbf{U}}_1=\tilde{\mathbf{U}}_2$$

Αν εμφανιστεί μηδενικό

Υπάρχουν δύο περιπτώσεις, η **«θεραπεύσιμη»** και η **«μη – θεραπεύσιμη»**.

Θεραπεύσιμη π.χ. $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 5 \\ 4 & 6 & 8 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 3 \\ 0 & 2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}$

Μπορώ με εναλλαγή γραμμών να φέρω στην θέση που μ' ενδιαφέρει μη μηδενικό στοιχείο και να το κάνω οδηγό.

Αν μπορώ **με εναλλαγή γραμμών** να έχω ένα πλήρες σύστημα οδηγών (n οδηγούς $\neq 0$) βρίσκομαι στην μη – ιδιόμορφη περίπτωση: **υπάρχει ακριβώς μία λύση για οποιαδήποτε δεξιά πλευρά**.

Σημείωση: Αν γνώριζα από πριν τις αναγκαίες εναλλαγές γραμμών θα μπορούσα να τις είχα κάνει εξ' αρχής και να κάνω μετά τον αλγόριθμο Gauss χωρίς να συναντήσω 0. $\Rightarrow$ **Υπάρχει πίνακας μεταθέσεων $\mathbf{P}$ τέτοιος που $\mathbf{PA} = \mathbf{LU}$.**

Μη – θεραπεύσιμη Όταν όλη η στήλη δεξιά κάτω από τον προηγούμενο οδηγό είναι γεμάτη 0. Άρα δεν μπορώ με εναλλαγή γραμμών να φέρω μη – μηδενικό στοιχείο στη θέση του οδηγού.

π.χ. $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 5 \\ 4 & 4 & 8 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 4 \end{bmatrix}$

Τότε βρίσκομαι στην **ιδιόμορφη** περίπτωση: **Καμμία αναδιάταξη γραμμών δεν παράγει πλήρες σύστημα οδηγών**.

Στην ιδιόμορφη περίπτωση το αν θα υπάρχουν λύσεις και πόσες εξαρτάται από τη δεξιά πλευρά.

- Αν η μετασχηματισμένη δεξιά πλευρά στο παραπάνω παράδειγμα ήταν $\begin{pmatrix} \cdots \\ 6 \\ 8 \end{pmatrix}$ οι δύο τελευταίες εξισώσεις θα ήταν συμβατές ($x_3 = 2$) αλλά θα είχα άπειρες λύσεις γιατί το x_2 είναι ελεύθερο.
- Αν η δεξιά πλευρά ήταν $\begin{pmatrix} \cdots \\ 6 \\ 7 \end{pmatrix}$ οι δύο τελευταίες εξισώσεις θα ήταν μη συμβατές και δεν θα είχα καμμία λύση.

Στην ιδιόμορφη περίπτωση έχω καμμία ή άπειρες λύσεις.

Επιπτώσεις εναλλαγής γραμμών στη κατασκευή του L

	A	**L**	**P**
Αρχ.	$\begin{bmatrix} 3 & 2 & 0 & 0 & 0 \\ 12 & 8 & 0 & 9 & 5 \\ 3 & 22 & 0 & 8 & 1 \\ 6 & 4 & 2 & 1 & 0 \\ 3 & 6 & 0 & 1 & 0 \end{bmatrix}$		
1^o β.	$\begin{bmatrix} 3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 & 5 \\ 0 & 20 & 0 & 8 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 4 & 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 4 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$
2^o β.	$\begin{bmatrix} 3 & 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 1 & 0 \\ 0 & 20 & 0 & 8 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 9 & 5 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$
3^o β.	$\begin{bmatrix} 3 & 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 9 & 5 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 5 & 1 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

4° β. $\begin{bmatrix} 3 & 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 9 & 5 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 \\ 1 & 5 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 1 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

$\mathbf{A}$ $\mathbf{L}$ $\mathbf{P}$

5° β. $\begin{bmatrix} 3 & 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 \\ 1 & 5 & 0 & 1 & 0 \\ 4 & 0 & 0 & 3 & 1 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

Όταν εναλλάσω γραμμές στον $\mathbf{A}$

- εναλλάσω τις ίδιες γραμμές στον $\mathbf{P}$
- εναλλάσω στον $\mathbf{L}$ τις αντίστοιχες γραμμές **μόνο στους συντελεστές πού έχω συμπληρώσει εγώ** (όχι τα 1)

Άσκηση για αύριο για το σπίτι. Να βρεθεί η παραγοντοποίηση $\mathbf{PA} = \mathbf{LD\tilde{U}}$ για το πίνακα

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 1 & 3 \\ 2 & 8 & 5 & 10 \\ 1 & 7 & 7 & 12 \\ 3 & 11 & 7 & 15 \end{bmatrix}$$

7. Αντίστροφοι και Ανάστροφοι Πίνακες

(Πηγή: Strang, Κεφάλαιο 1.6)

Αντίστοφοι πίνακες

- Έστω ο στοιχειώδης $n \times n$ πίνακας $\mathbf{E_{ij}}(\lambda) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & 0 \\ \vdots\lambda & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{bmatrix}$

- Γνωρίζουμε ότι $\mathbf{E_{ij}}(\lambda) \cdot x$ προσθέτει την *λ*-πλάσια της *j*-συντεταγμένης στην *i*-συντεταγμένη του x: $\mathbf{E_{ij}}(\lambda) \cdot \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ x_i + \lambda x_j \\ x_{i+1} \\ \vdots \\ x_n \end{pmatrix}$.

- Αν τώρα πολλαπλασιάσουμε το αποτέλεσμα με $\mathbf{E_{ij}}(-\lambda)$ ξανα-ακυρώνουμε τον μετασχηματισμό που διενήργησε η $\mathbf{E_{ij}}(\lambda)$ και γυρνάμε στο αρχικό x.

  Δηλαδή

$$\mathbf{E_{ij}}(-\lambda) \cdot \left(\mathbf{E_{ij}}(\lambda) \cdot \mathbf{x}\right) = \mathbf{x}$$

Αυτή είναι και η ιδιότητα μέσω της οποίας θα ορίσουμε τον αντίστροφο, $\mathbf{A}^{-1}$, ενός πίνακα $\mathbf{A}$. Εάν υπάρχει αντίστροφος αυτός θα έχει την ιδιότητα ότι

$$\boxed{\mathbf{A^{-1}(Ax) = x}} \quad \forall \mathbf{x} \in R^n \text{ και επομένως } \mathbf{A^{-1}A = I}.$$

Σημείωση: **Αντίστροφος δεν υπάρχει για κάθε πίνακα $\mathbf{A}$:** Έστω ένας πίνακας $\mathbf{A}$ ιδιόμορφος. Τότε θα υπάρχει $\mathbf{x} \neq 0$ με $\mathbf{Ax} = \mathbf{0}$. Εδώ όμως δεν υπάρχει πίνακας που πολλαπλασιάζει το Ax και ξαναγυρνάμε στο x διότι: $\mathbf{B(Ax) = B0 = 0}$

Ορισμός (αντίστροφος πίνακας): Ένας $n \times n$ πίνακας $\mathbf{A}$ λέγεται **αντιστρέψιμος** εάν υπάρχει πίνακας $\mathbf{B}$ τέτοιος που

$\mathbf{AB} = \mathbf{I}$ ($\mathbf{B}$ δεξιός αντίστροφος) και

$\mathbf{BA} = \mathbf{I}$ ($\mathbf{B}$ αριστερός αντίστροφος).

Εάν υπάρχει είναι μοναδικός και συμβολίζεται με $\mathbf{A}^{-1}$.

Παραδείγματα:

- $(a)^{-1} = \left(\dfrac{1}{a}\right)$

- $\left(\mathbf{E_{ij}}(\lambda)\right)^{-1} = \mathbf{E_{ij}}(-\lambda)$

- $\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}^{-1} = \begin{bmatrix} \lambda_1^{-1} & 0 & \dots & 0 \\ 0 & \lambda_2^{-1} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n^{-1} \end{bmatrix}$

Σημείωση:

1) Από τον ορισμό προκύπτει ότι $\mathbf{(A^{-1})^{-1} = A}$

2) Αν ο $\mathbf{A}$ έχει δεξιό αντίστροφο $\mathbf{C}$ και αριστερό $\mathbf{B}$ τότε πρέπει να ταυτίζονται και ο $\mathbf{A}$ είναι αντιστρέψιμος!

Διότι:

$$\mathbf{B = BI = B(AC) = (BA)C = IC = C}$$

Κανόνας: Αν $\mathbf{A}$, $\mathbf{B}$ αντιστρέψιμοι (<u>υποθέτω</u> **ότι υπάρχουν** $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$) τότε και $\mathbf{AB}$ αντιστρέψιμος και

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

<u>Απόδειξη:</u> $(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) \overset{!}{=} \mathbf{I}$ και $\mathbf{AB}(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{I}$

$\Leftrightarrow$ παρομοίως

$$\mathbf{B}^{-1}\underbrace{(\mathbf{A}^{-1}\mathbf{A})}_{\mathbf{I}}\mathbf{B} = \mathbf{I}$$

$\Leftrightarrow$

$$\mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$$

ΟΚ

<u>Προσοχή:</u> Μπορεί να υπάρχουν μη-τετραγωνικοί πίνακες $\mathbf{A}_{n\times m}$, $\mathbf{B}_{m\times n}$ (, οι οποίοι δεν έχουν αντίστροφο), τέτοιοι που ο $n\times n$ πίνακας $\mathbf{AB}$ να έχει αντίστροφο. Τότε προφανώς ο παραπάνω τύπος δεν ισχύει.

Π.χ. $\mathbf{A} = \begin{pmatrix} 1 & 0 \end{pmatrix}$ και $\mathbf{B} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Τότε $\mathbf{AB} = \begin{pmatrix} 1 \end{pmatrix}$ που είναι αντιστρέψιμος

Ή π.χ. $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ και $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. Τότε $\mathbf{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ που είναι αντιστρέψιμος.

<u>Γενικότερα</u>: $(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$ αν $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ αντιστρέψιμοι.

Μέθοδος υπολογισμού του αντιστρόφου (Αλγόριθμος Gauss Jordan)

Έστω $\mathbf{A}$ έχει n οδηγούς. Θα δείξουμε πώς μπορούμε να υπολογίσουμε τον A^{-1} σε αυτή την περίπτωση.

Ας συμβολίσουμε την *i*-στήλη του $\mathbf{A}^{-1}$ με $\mathbf{x_i}$ (άγνωστη) δηλαδή:

$\mathbf{A}^{-1} = (\mathbf{x_1} \,|\, \mathbf{x_2} \,|\, \ldots \,|\, \mathbf{x_n})$, $\mathbf{x_i} \in R^n$! (διάνυσμα, όχι αριθμός!)

Πιο αναλυτικά: $\mathbf{A}^{-1} = (\mathbf{x_1} \,|\, \mathbf{x_2} \,|\, \ldots \,|\, \mathbf{x_n}) = \left(\begin{array}{c|c|c|c} (\mathbf{x_1})_1 & (\mathbf{x}_2)_1 & & (\mathbf{x}_n)_1 \\ (\mathbf{x_1})_2 & (\mathbf{x}_2)_2 & \cdots & (\mathbf{x}_n)_2 \\ \vdots & \vdots & & \vdots \\ (\mathbf{x_1})_n & (\mathbf{x}_2)_n & & (\mathbf{x}_n)_n \end{array}\right)$

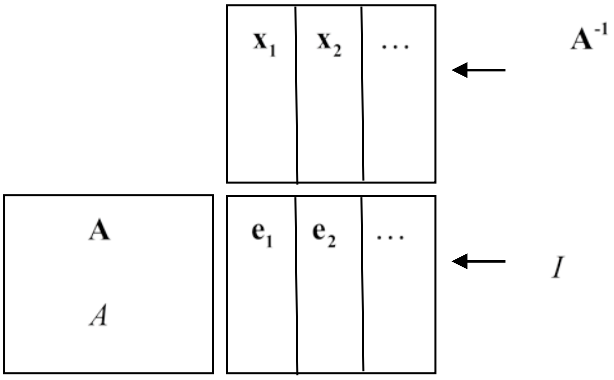
Ας συμβολίσουμε επιπλέον με $\mathbf{e_i} = i - \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}$. Δηλαδή $\mathbf{I} = (\mathbf{e_1} \,|\, \mathbf{e_2} \,|\, \ldots \,|\, \mathbf{e_n})$.

Πιο αναλυτικά:

$$\mathbf{I} = (\mathbf{e}_1 \,|\, \mathbf{e}_2 \,|\, \ldots \,|\, \mathbf{e_n}) = \left(\begin{array}{c|c|c|c} (\mathbf{e}_1)_1 & (\mathbf{e}_2)_1 & & (\mathbf{e}_n)_1 \\ (\mathbf{e}_1)_2 & (\mathbf{e}_2)_2 & \cdots & (\mathbf{e}_n)_2 \\ \vdots & \vdots & & \vdots \\ (\mathbf{e}_1)_n & (\mathbf{e}_2)_n & & (\mathbf{e}_n)_n \end{array}\right) = \left(\begin{array}{c|c|c|c} 1 & 0 & & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & 1 \end{array}\right)$$

- Η εξίσωση που ορίζει τον $\mathbf{A}^{-1}$ είναι η $\boxed{\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}}$

  Δηλαδή:



- **Ισοδύναμα** $\boxed{\begin{array}{l}\mathbf{A}\mathbf{x}_i = \mathbf{e}_i \\ i = 1, \ldots, n\end{array}}$

  <u>Άρα:</u> έχω n γραμμικά συστήματα. Από το πρώτο μπορώ να βρω το $\mathbf{x}_1$, από το δεύτερο το $\mathbf{x}_2$, κλπ.

<u>Πώς:</u>

- πρώτα για το $\mathbf{x}_1$: θα έφτιαχνα τον $\begin{bmatrix}\mathbf{A} & |\mathbf{e}_1\end{bmatrix}$ και θα έκανα αλγόριθμο Gauss μέχρι να φτάσω σε

$$\begin{bmatrix}\mathbf{A} & |\mathbf{e}_1\end{bmatrix} \to \begin{bmatrix}\mathbf{U} & |\mathbf{l}'_1\end{bmatrix} = \begin{bmatrix}\mathbf{L}'\mathbf{A} & |\mathbf{L}'\mathbf{e}_1\end{bmatrix}.$$

Τότε το $\mathbf{A}\mathbf{x}_1 = \mathbf{e}_1$ ισοδυναμεί με $\mathbf{U}\mathbf{x}_1 = \mathbf{l}'_1$

(όπου $\mathbf{l}'_1$ η *1*-στήλη του $\mathbf{L}'$)

- μετά για το $\mathbf{x}_2$: θα έφτειαχνα τον $\begin{bmatrix}\mathbf{A} & |\mathbf{e}_2\end{bmatrix}$ και θα έκανα αλγόριθμο Gauss μέχρι να φτάσω σε

$$\begin{bmatrix}\mathbf{A} & |\mathbf{e}_2\end{bmatrix} \to \begin{bmatrix}\mathbf{U} & |\mathbf{l}'_2\end{bmatrix} = \begin{bmatrix}\mathbf{L}'\mathbf{A} & |\mathbf{L}'\mathbf{e}_2\end{bmatrix}.$$

Τότε το $\mathbf{A}\mathbf{x}_2 = \mathbf{e}_2$ ισοδυναμεί με $\mathbf{U}\mathbf{x}_2 = \mathbf{l}'_2$

(όπου $\mathbf{l}'_2$ η *2*-στήλη του $\mathbf{L}'$)

…

Αντί να κάνω Gauss για κάθε σύστημα ξεχωριστά, κολλάω όλες τις δεξιές πλευρές δεξιά από τον A και κάνω Gauss σε όλες μαζί:

$$\begin{bmatrix}\mathbf{A} \,|\, \mathbf{e}_1 \,|\, \mathbf{e}_2 \,|\, \ldots |\mathbf{e}_n\end{bmatrix} = \begin{bmatrix}\mathbf{A} \,|\, \mathbf{I}\end{bmatrix} \longrightarrow \begin{bmatrix}\mathbf{U} \,|\, \mathbf{L}'\end{bmatrix}$$

Τότε το $\mathbf{A}\mathbf{x}_i = \mathbf{e}_i$ ισοδυναμεί με $\mathbf{U}\mathbf{x}_i = \mathbf{l}'_i$

π.χ. Βρες τον αντίστροφο του $\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}$

$$\left[\begin{array}{ccc:ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 4 & -6 & 0 & 0 & 1 & 0 \\ -2 & 7 & 2 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc:ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 8 & 3 & 1 & 0 & 1 \end{array}\right]$$

A **I**

$$\rightarrow \left[\begin{array}{cccc:cc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right]$$

U **L'**

$\mathbf{A}\mathbf{x}_i = \mathbf{e}_i$ ισοδυναμεί με $\mathbf{U}\mathbf{x}_i = \mathbf{l}'_i$

- Εδώ θα μπορούσα να πάρω τα συστήματα $\mathbf{U}\mathbf{x}_i = \mathbf{l}'_i$ (όπου $\mathbf{l}'_i$ η *i*-στήλη του $\mathbf{L}'$) και να λύσω το καθένα από αυτά με ανάδρομη αντικατάσταση.

  Π.χ. για να βρώ το $\mathbf{x}_1$ (τη πρώτη στήλη του $\mathbf{A}^{-1}$) θα έλυνα με ανάδρομη αντικατάσταση το $\mathbf{U}\mathbf{x}_1 = \mathbf{l}'_1$, δηλαδή το

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_1 = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$$

Ωστόσο, τώρα που έχω πολλές δεξιές πλευρές για τον ίδιο U συμφέρει να κάνω κάτι άλλο:

Να συνεχίσω τους γραμμομετασχηματισμούς μέχρι αριστερά να εμφανιστεί ο ταυτοτικός (πώς; βλέπε παρακάτω) *

$$\left[\mathbf{A}\,|\,\mathbf{I}\right] \rightarrow \left[\mathbf{U}\,|\,\mathbf{L}'\right] \rightarrow \left[\mathbf{I}\,|\,\mathbf{K}\right]$$

Τότε το $\mathbf{Ax}_i = \mathbf{e}_i$ ισοδυναμεί με $\mathbf{Ux}_i = \mathbf{l}'_i$ που ισοδυναμεί με $\mathbf{Ix}_i = \mathbf{k}_i$

και **επομένως η *i*-στήλη του** $\mathbf{K}$ **είναι η *i*-στήλη του** $\mathbf{A}^{-1}$ **δηλαδή** $\boxed{\mathbf{K} = \mathbf{A}^{-1}}$**.** Τελείωσα!

Πώς * : **με ανάποδη απαλοιφή Gauss**.

Βήμα 1^{o}: Προσθέτω πολλαπλάσια της τελευταίας γραμμής στις προηγούμενες έτσι ώστε να μηδενιστούν τα στοιχεία πάνω από τον τελευταίο οδηγό.

Βήμα 2^{o}: Προσθέτω πολλαπλάσια της προτελευταίας γραμμής στις προηγούμενες έτσι ώστε να μηδενιστούν τα στοιχεία πάνω από τον προτελευταίο οδηγό.

⋮

Μέχρι να φθάσω αριστερά σε διαγώνιο πίνακα.

Τέλος: Διαιρώ κάθε γραμμή με στοιχείο της διαγωνίου του αριστερού πίνακα, οπότε αριστερά παίρνω τον ταυτοτικό.

$$\left[\begin{array}{ccc:ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right] \to \left[\begin{array}{ccc:ccc} 2 & 1 & 0 & 2 & -1 & -1 \\ 0 & -8 & 0 & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{ccc:ccc} 2 & 0 & 0 & \frac{12}{8} & -\frac{5}{8} & -\frac{6}{8} \\ 0 & -8 & 0 & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right] \to \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & \frac{12}{16} & -\frac{5}{16} & -\frac{6}{16} \\ 0 & 1 & 0 & \frac{4}{8} & -\frac{3}{8} & -\frac{2}{8} \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right] \longleftarrow A^{-1}$$

<u>Σημείωση:</u> Δείξαμε ότι αν $\mathbf{A}$ μη – ιδιόμορφος $\Rightarrow$ υπάρχει δεξιός αντίστροφος.

Όμως η κατασκευή εξασφαλίζει ότι θα υπάρχει και αριστερός και άρα ο $\mathbf{A}$ είναι αντιστρέψιμος και ο $\mathbf{A}^{-1}$ αυτός που φτιάξαμε.

Μπορεί κανείς να δείξει και το ανάποδο:

Αν $\mathbf{A}$ αντιστρέψιμος $\Rightarrow$ μη – ιδιόμορφος

και άρα:

Αντιστρέψιμοι πίνακες είναι ακριβώς οι μη – ιδιόμορφοι.

Άσκηση για το σπιτι. Να βρεθεί ο αντίστροφος του $\mathbf{A}=\begin{bmatrix} 1 & 2 & -4 \\ -1 & -1 & 5 \\ 2 & 7 & 3 \end{bmatrix}$

Άσκηση. Να βρεθεί ο αντίστροφος του $\mathbf{A}=\begin{bmatrix}1&0&1&2\\2&1&1&0\\3&2&1&0\\2&3&1&0\end{bmatrix}$.

$$\left[\begin{array}{cccc|cccc}1&0&1&2&1&0&0&0\\2&1&1&0&0&1&0&0\\3&2&1&0&0&0&1&0\\2&3&1&0&0&0&0&1\end{array}\right]\to\left[\begin{array}{cccc|cccc}1&0&1&2&1&0&0&0\\0&1&-1&-4&-2&1&0&0\\0&2&-2&-6&-3&0&1&0\\0&3&-1&-4&-2&0&0&1\end{array}\right]$$

$$\to\left[\begin{array}{cccc|cccc}1&0&1&2&1&0&0&0\\0&1&-1&-4&-2&1&0&0\\0&0&0&2&1&-2&1&0\\0&0&2&8&-4&-3&0&1\end{array}\right]\to\left[\begin{array}{cccc|cccc}1&0&1&2&1&0&0&0\\0&1&-1&-4&-2&1&0&0\\0&0&2&8&-4&-3&0&1\\0&0&0&2&1&-2&1&0\end{array}\right]$$

$$\to\left[\begin{array}{cccc|cccc}1&0&1&0&0&2&-1&0\\0&1&-1&0&0&-3&2&0\\0&0&2&0&0&5&-4&1\\0&0&0&2&1&-2&1&0\end{array}\right]\to\left[\begin{array}{cccc|cccc}1&0&0&0&0&-1/2&1&-1/2\\0&1&0&0&0&-1/2&0&1/2\\0&0&2&0&0&5&-4&1\\0&0&0&2&1&-2&1&0\end{array}\right]$$

$$\to\left[\begin{array}{cccc|cccc}1&0&0&0&0&-1/2&1&-1/2\\0&1&0&0&0&-1/2&0&1/2\\0&0&1&0&0&5/2&-2&1/2\\0&0&0&1&-1/2&1&1/2&0\end{array}\right]$$

Άρα $\mathbf{A}^{-1}=\begin{bmatrix}0&-1/2&1&-1/2\\0&-1/2&0&1/2\\0&5/2&-2&1/2\\-1/2&1&1/2&0\end{bmatrix}$

Άσκηση. Αν $\mathbf{A}$ $n\times n$ αντιστρέψιμος και $\mathbf{c},\mathbf{d}\in\mathbb{R}^n$, να δειχτεί ότι:

$$\left(\mathbf{A}+\mathbf{c}\mathbf{d}^{\mathbf{T}}\right)^{-1}=\mathbf{A}^{-1}-\frac{\mathbf{A}^{-1}\mathbf{c}\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}}{1+\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{c}}.$$

Απόδ. Θα δείξουμε πρώτα: $\left(\mathbf{A}+\mathbf{c}\mathbf{d}^{\mathbf{T}}\right)\left[\mathbf{A}^{-1}-\dfrac{\mathbf{A}^{-1}\mathbf{c}\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}}{1+\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{c}}\right]=\mathbf{I}$.

Θέτοντας $\lambda=1+\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{c}$, η αριστερή πλευρά ισούται με

$$\mathbf{A}\mathbf{A}^{-1}-\lambda^{-1}\mathbf{A}\mathbf{A}^{-1}\mathbf{c}\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}+\mathbf{c}\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}-\lambda^{-1}\mathbf{c}\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{c}\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}$$

$$=\mathbf{I}+\left(1-\lambda^{-1}\right)\mathbf{c}\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}-\lambda^{-1}\mathbf{c}\left(\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{c}\right)\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}$$

$$=\mathbf{I}+\left(1-\lambda^{-1}-\lambda^{-1}\left(\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{c}\right)\right)\mathbf{c}\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}=\mathbf{I},$$

καθώς

$$1-\lambda^{-1}-\lambda^{-1}\left(\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{c}\right)=1-\lambda^{-1}\left(1+\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{c}\right)=1-\lambda^{-1}\lambda=0.$$

Η $\left[\mathbf{A}^{-1}-\dfrac{\mathbf{A}^{-1}\mathbf{c}\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}}{1+\mathbf{d}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{c}}\right]\left(\mathbf{A}+\mathbf{c}\mathbf{d}^{\mathbf{T}}\right)=\mathbf{I}$ αποδεικνύεται παρομοίως.

Πρόταση. Έστω $\mathbf{A}=\begin{bmatrix}\mathbf{W} & \mathbf{V}\\ \mathbf{U} & \mathbf{T}\end{bmatrix}$, όπου $\mathbf{W}_{n\times n},\mathbf{T}_{m\times m},\mathbf{U}_{m\times n},\mathbf{V}_{n\times m}$ και $\mathbf{W},\mathbf{T}$ αντιστρέψιμοι. Τότε ο $\mathbf{A}$ θα είναι αντιστρέψιμος αν $\mathbf{Q}:=\mathbf{W}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U}$ είναι αντιστρέψιμος και $\mathbf{A}^{-1}=\mathbf{F}$, με

$$\mathbf{F}:=\begin{bmatrix}\mathbf{Q}^{-1} & -\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}\\ -\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1} & \mathbf{T}^{-1}+\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}\end{bmatrix}.$$

Απόδ.

$$\mathbf{AF}=\begin{bmatrix}\mathbf{Q}^{-1} & -\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}\\ -\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1} & \mathbf{T}^{-1}+\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}\end{bmatrix}$$

$$\begin{bmatrix}\mathbf{W} & \mathbf{V}\\ \mathbf{U} & \mathbf{T}\end{bmatrix}\begin{bmatrix}\mathbf{W}\mathbf{Q}^{-1}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1} & -\mathbf{W}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}+\mathbf{V}\mathbf{T}^{-1}+\mathbf{V}\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}\\ \mathbf{U}\mathbf{Q}^{-1}-\mathbf{T}\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1} & -\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}+\mathbf{T}\mathbf{T}^{-1}+\mathbf{T}\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}\end{bmatrix}$$

$\mathbf{AF}$ Πάνω αριστερά: $\left(\mathbf{W}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U}\right)\mathbf{Q}^{-1}=\mathbf{Q}\mathbf{Q}^{-1}=\mathbf{I}$

$\mathbf{AF}$ Κάτω αριστερά: $\mathbf{U}\mathbf{Q}^{-1}-\mathbf{U}\mathbf{Q}^{-1}=\mathbf{0}$

$\mathbf{AF}$ Πάνω δεξιά: $\left(-\mathbf{W}+\mathbf{Q}+\mathbf{V}\mathbf{T}^{-1}\mathbf{U}\right)\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}=$

$$\left(-\mathbf{W}+\mathbf{W}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U}+\mathbf{V}\mathbf{T}^{-1}\mathbf{U}\right)\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}=\mathbf{0}$$

$\mathbf{AF}$ Κάτω δεξιά: $-\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}+\mathbf{I}+\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1}=\mathbf{I}$

$\mathbf{FA}=\mathbf{I}$ παρομοίως.

Ανάστροφοι πίνακες

Αν $\mathbf{A}$ ένας $m \times n$ πίνακας, τότε ως $\mathbf{A}^{\mathrm{T}}$ ορίζεται ο $n \times m$ πίνακας που προκύπτει **αν στον A εναλλάξω γραμμές και στήλες**:

Η πρώτη γραμμή του $\mathbf{A}$ γίνεται πρώτη στήλη του $\mathbf{A}^{\mathrm{T}}$ κλπ.

Δηλαδή $\left[\mathbf{A}^{\mathrm{T}}\right]_{ij} = \mathbf{A}_{ji}$

π.χ. $A = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 4 & 3 \end{bmatrix}$

<u>Επίσης:</u> Ο $\mathbf{A}^{\mathrm{T}}$ **προκύπτει αν στον $\mathbf{A}$ καθρεφτίσω τα στοιχεία στην κύρια διαγώνιο** $(a_{11}, a_{22}, \ldots)$

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 4 & 3 \end{bmatrix}$$

Ισχύουν οι εξής **κανόνες**:

- $(\mathbf{A}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{A}$ (προφανές)
- $(\mathbf{A}+\mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}}$ (εύκολο...)
- $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ (θέλει λίγη δουλίτσα, αλλά μόνο ορισμοί...)
- $(\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1}$

Απόδειξη του τελευταίου:

Έχουμε $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$

παίρνουμε ανάστροφα $(\mathbf{A}^{-1}\mathbf{A})^{\mathrm{T}} = \mathbf{I} = (\mathbf{A}\mathbf{A}^{-1})^{\mathrm{T}}$

κανόνας $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$: $\mathbf{A}^{\mathrm{T}}(\mathbf{A}^{-1})^{\mathrm{T}} = \mathbf{I} = (\mathbf{A}^{-1})^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$

Αυτός είναι ορισμός του αντιστρόφου του $\mathbf{A}^{\mathrm{T}}$: ένας πίνακας $\mathbf{B}$ με $\mathbf{A}^{\mathrm{T}}\mathbf{B} = \mathbf{I} = \mathbf{B}\mathbf{A}^{\mathrm{T}}$.

Άρα ο $(\mathbf{A}^{-1})^{\mathrm{T}}$ είναι ο $\mathbf{B}$ δηλαδή ο $(\mathbf{A}^{\mathrm{T}})^{-1}$.

<u>**Ορισμός:**</u> Ένας $n \times n$ πίνακας $\mathbf{A}$ λέγεται <u>συμμετρικός</u>.

$$\Leftrightarrow \boxed{\mathbf{A}^{\mathrm{T}} = \mathbf{A}}$$

<u>δηλ.</u> τα συμμετρικά ως προς την κύρια διαγώνιο στοιχεία ταυτίζονται.

<u>**Πρόταση:**</u> Αν $\mathbf{A}$ συμμετρικός και αντιστρέψιμος τότε $\mathbf{A}^{-1}$ συμμετρικός.

!

<u>**Απόδειξη:**</u> Δείξτε $(\mathbf{A}^{-1})^{\mathrm{T}} = \mathbf{A}^{-1}$

Αλλά έχουμε $(\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1}$ και αφού $\mathbf{A}$ συμμετρικός (άρα $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$) το τελευταίο είναι ίσο με $\mathbf{A}^{-1}$. OK

Διάσπαση Cholesky

Αν $\mathbf{A}$ συμμετρικός υπάρχουν κάτω τριγωνικός $\mathbf{L}$ με 1 στη διαγώνιο και διαγώνιος $\mathbf{D}$ τέτοιοι ώστε $\mathbf{A} = \mathbf{LDL}^T$.

Απόδειξη. Έχουμε $\mathbf{A} = \mathbf{LDU}$ και επομένως $\mathbf{A}^{\mathrm{T}} = \mathbf{U}^{\mathrm{T}}\mathbf{DL}^{\mathrm{T}}$ και αφού $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$ και $\mathbf{A} = \mathbf{U}^{\mathrm{T}}\mathbf{DL}^{\mathrm{T}}$. Αλλά αφου η διάσπαση $\mathbf{A} = \mathbf{LDU}$ είναι μοναδική, έπεται $\mathbf{L} = \mathbf{U}^{\mathrm{T}}$ και άρα $\mathbf{A} = \mathbf{LDL}^T$.

<u>**Προσοχή**</u>: Αν $\mathbf{A}$, $\mathbf{B}$ συμμετρικοί μπορεί ο $\mathbf{AB}$ να <u>μην</u> είναι συμμετρικός, καθώς $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} = \mathbf{BA}$ και όχι $\mathbf{AB}$.

π.χ. $\begin{bmatrix} 1 & b \\ b & 0 \end{bmatrix} \longleftarrow$ συμμετρικός

$\begin{bmatrix} 1 & a \\ a & 0 \end{bmatrix} \begin{bmatrix} & b \\ a & \end{bmatrix} \longleftarrow$ όχι συμμετρικός

συμμετρικός

8. Διανυσματικοί χώροι και υπόχωροι

(Πηγή: Strang, Κεφάλαιο 2.1)

Οι $\mathbb{R}^n$ που γνωρίσαμε είναι σύνολα εφοδιασμένα με πρόσθεση $(\mathbf{x},\mathbf{y}\in\mathbb{R}^n \Rightarrow \mathbf{x}+\mathbf{y}\in\mathbb{R}^n)$ και πολλαπλασιασμό με πραγματικό αριθμό $(\mathbf{x}\in\mathbb{R}^n \Rightarrow \lambda\mathbf{x}\in\mathbb{R}^n)$.

Γι' αυτές τις πράξεις ισχύει ένα σύνολο <u>κανόνων:</u>

1) $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$

2) $\mathbf{x}+(\mathbf{y}+\mathbf{z})=(\mathbf{x}+\mathbf{y})+\mathbf{z}$

3) $\exists\mathbf{0}$ δηλ. $\mathbf{x}+\mathbf{0}=\mathbf{0}+\mathbf{x}=\mathbf{x}$

4) $\forall\mathbf{x}\ \exists(-\mathbf{x}):\ \mathbf{x}+(\text{-}\mathbf{x})=\mathbf{0}$

5) $1\cdot\mathbf{x}=\mathbf{x}$

6) $\lambda\cdot(\mu\cdot\mathbf{x})=(\lambda\cdot\mu)\cdot\mathbf{x}$

7) $\lambda(\mathbf{x}+\mathbf{y})=\lambda\mathbf{x}+\lambda\mathbf{y}$

8) $(\lambda+\mu)\mathbf{x}=\lambda\mathbf{x}+\mu\mathbf{x}$

<u>Ορισμός:</u> Ένα σύνολο V εφοδιασμένο με μια πράξη + (πρόσθεση) και ένα πολλαπλασιασμό με πραγματικό αριθμό, τέτοιους ώστε να ισχύουν οι παραπάνω κανόνες λέγεται **<u>διανυσματικός χώρος</u>**.

Η έννοια του διανυσματικού χώρου είναι τόσο γενική που συμπεριλαμβάνει χώρους όπως:

- $\{$ το σύνολο των ακολουθιών $\{a_\nu\}_{\nu\in\mathrm{N}}$ $\}$,

- $\{$ το σύνολο των πραγματικών συναρτήσεων $f:\mathbb{R}\to\mathbb{R}\}$

- $\{$ το σύνολο των συνεχών συναρτήσεων $\}$

- $\{$ το σύνολο των πολυωνύμων βαθμού $\leq k$ $\}$.

Εδώ όμως θα περιοριστούμε στους $\mathbb{R}^n$ ως διανυσματικούς χώρους.

Υπόχωροι

Ορισμός: Έστω V ένας διανυσματικός χώρος και $U\subseteq V$ με $U\neq\varnothing$ και U έχει τις εξής ιδιότητες:

- αν $x,y\in U\Rightarrow x+y\in U$
- αν $x\in U\Rightarrow \lambda\cdot x\in U \quad \forall\lambda$

Τότε U ονομάζεται **υπόχωρος** του V.

Σημείωση:

- Το $\mathbf{0} \in U$ διότι $0 \cdot \mathbf{x} = \mathbf{0} \in U$.

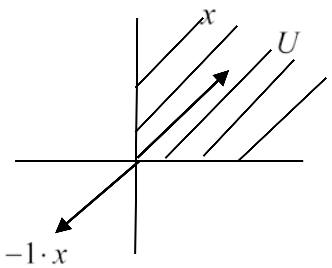
- Ο πιο μικρός υπόχωρος του V είναι ο $U = \{\mathbf{0}\}$.

- Ο πιο μεγάλος υπόχωρος του V είναι ο V.

Παραδείγματα:

Είναι υπόχωροι οι παρακάτω?

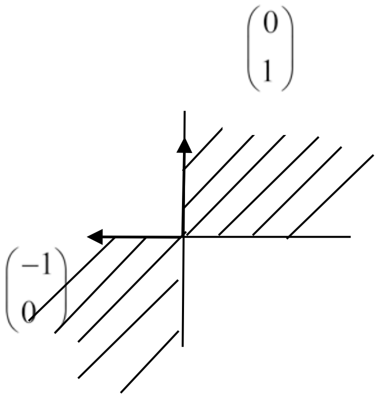
❖ στον $\mathbb{R}^2$



$U = \{x : x_1 \geq 0 \text{ και } x_2 \geq 0\}$

<u>ΟΧΙ</u> διότι αν $x \in U$ τότε $(-1) \cdot x \notin U$

❖ στον $\mathbb{R}^2$



$$U = \left\{ x \middle| \begin{array}{l} x_1 \geq 0 \text{ ή } x_2 \geq 0 \\ x_1 \leq 0 \text{ ή } x_2 \leq 0 \end{array} \right\}$$

<u>ΟΧΙ</u> διότι $\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \notin U$

❖ $\mathbb{R}^2$ ή $\mathbb{R}^n$ ευθεία που διέρχεται από 0.

NAI διότι: $L = Span(v)$ | αν $\left.\begin{array}{l}\mathbf{x} \in L \Rightarrow \mathbf{x} = \lambda \mathbf{v} \\ \mathbf{x}' \in L \Rightarrow \mathbf{x}' = \lambda' \mathbf{v}\end{array}\right\} \Rightarrow \mathbf{x} + \mathbf{x}' = (\lambda + \lambda')\mathbf{v}$

και $\mu \cdot \mathbf{x} = \mu \cdot \lambda \mathbf{v} = (\mu\lambda) \cdot \mathbf{v} \in L$

❖ $\mathbb{R}^2$ ή $\mathbb{R}^n$ ευθεία που δε διέρχεται από 0.

OXI διότι δεν περιέχει το 0.

❖ στον $\mathbb{R}^3$ ή $\mathbb{R}^n$: επίπεδο που διέρχεται από 0.

$E = Span(\mathbf{v}, \mathbf{w})$ **NAI** (βλ. παρακάτω)

και γενικότερα στον $\mathbb{R}^n$:

Αν $U = Span(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ **τότε** U **υπόχωρος**

Απόδειξη: Αν $\mathbf{x} \in U \Rightarrow \exists \lambda_1, \ldots \lambda_k : \ \mathbf{x} = \lambda_1 \mathbf{v_1} + \ldots + \lambda_k \mathbf{v_k}$

$x' \in U \Rightarrow \lambda'_1, \ldots \lambda'_k : \ \mathbf{x}' = \lambda'_1 \mathbf{v_1} + \ldots + \lambda'_k \mathbf{v_k}$

$\Rightarrow \mathbf{x} + \mathbf{x}' = (\lambda_1 + \lambda'_1)\mathbf{v_1} + \ldots + (\lambda_k + \lambda'_k)\mathbf{v_k} \in U$

και $\mu \cdot \mathbf{x} = \lambda_1 \mu \cdot \mathbf{v_1} + \ldots + \lambda_k \mu \cdot \mathbf{v_k} \in U$

Χώρος στηλών ενός mxn πίνακα A

π.χ. $\mathbf{A}=\begin{bmatrix}1 & 0\\ 5 & 4\\ 2 & 4\end{bmatrix}$

- Ψάχνουμε όλα τα $\mathbf{b}=\begin{pmatrix}b_1\\ b_2\\ b_3\end{pmatrix}$ για τα οποία ισχύει: το σύστημα $A\mathbf{x}=\mathbf{b}$ έχει λύση.

- Αλλά τότε: $\exists \mathbf{x}=\begin{pmatrix}x_1\\ x_2\end{pmatrix}$: $\mathbf{b}=A\mathbf{x}=x_1\mathbf{a_1}+x_2\mathbf{a_2}$, δηλαδή $\mathbf{b}\in Span(\mathbf{a_1},\mathbf{a_2})$.

Στο παράδειγμά μας, λοιπόν, τα $\mathbf{b}$ με αυτή την ιδιότητα είναι τα σημεία του επιπέδου που περνά από το 0, το $a_1=\begin{pmatrix}1\\ 5\\ 2\end{pmatrix}$ και $a_2=\begin{pmatrix}0\\ 4\\ 4\end{pmatrix}$.

<u>Γενικώς:</u> Αν A ένας $m\times n$ πίνακας τότε **χώρος στηλών** του είναι το σύνολο των $\mathbf{b}\in R^m$ (των δεξιών πλεύρων) για τα οποία υπάρχει $\mathbf{x}\in R^n$ με $\mathbf{Ax}=\mathbf{b}$.

Ο χώρος αυτός ταυτίζεται με το $Span$ των στηλών του πίνακα. Συμβολίζεται με $R(A)$:

$$\boxed{R(A):=Span(\mathbf{a}_1,\ldots,\mathbf{a}_n)\subset R^m}$$

Είναι υπόχωρος, καθώς γράφεται ως $Span(\mathbf{a}_1,\ldots,\mathbf{a}_n)$.

Άλλη απόδειξη:

- Έστω $\mathbf{b},\mathbf{b}'\in R(\mathbf{A})$.
  - Τότε υπάρχουν $\mathbf{x},\mathbf{x}':\mathbf{Ax}=\mathbf{b}$ και $\mathbf{Ax}'=\mathbf{b}'$.

- Προθέτοντας παίρνουμε $\mathbf{A}(\mathbf{x}+\mathbf{x}')=\mathbf{b}+\mathbf{b}'$ και άρα $\mathbf{b}+\mathbf{b}'\in R(\mathbf{A})$.

- Παρόμοια από $\mathbf{Ax}=\mathbf{b}$ πολλαπλασιάζονατς με λ παίρνουμε
  - $\mathbf{A}(\lambda\mathbf{x})=\lambda\mathbf{b}$ και άρα $\lambda\mathbf{b}\in R(\mathbf{A})$.

Άλλα παραδείγματα:

- $\mathbf{A}=\left[\mathbf{a}|\lambda\mathbf{a}\right]$. Τότε $R(\mathbf{A})=Span\{\mathbf{a}\}$
- $\mathbf{A}=[\mathbf{0}]$. Τότε $R(\mathbf{A})=\{\mathbf{0}\}$
- $\mathbf{A}$ $\mathrm{n}\times\mathrm{n}$ μη ιδιόμορφος. Τότε το $\mathbf{Ax}=\mathbf{b}$ έχει λύση για κάθε $\mathbf{b}\in\mathbb{R}^n$. Άρα $R(\mathbf{A})=\mathbb{R}^n$.

Ο Μηδενόχωρος ενός mxn πίνακα A

<u>Ορισμός:</u> Το σύνολο των $\mathbf{x} \in \mathbb{R}^n$ με $\mathbf{Ax} = \mathbf{0}$ λέγεται **μηδενόχωρος** του $\mathbf{A}$. **Είναι υπόχωρος** του $\mathbb{R}^n$.

<u>Απόδειξη:</u> (Είναι υπόχωρος)

- Αν $\left.\begin{array}{l}\mathbf{x} \in N(\mathbf{A}) \Rightarrow \mathbf{Ax} = \mathbf{0} \\ \mathbf{x}' \in N(\mathbf{A}) \Rightarrow \mathbf{Ax}' = \mathbf{0}\end{array}\right\} \quad \mathbf{A}(\mathbf{x}+\mathbf{x}') = \mathbf{0} \Rightarrow \mathbf{x}+\mathbf{x}' \in N(\mathbf{A})$

- $\mathbf{x} \in N(\mathbf{A}) \Rightarrow \mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{A}\cdot(\lambda\mathbf{x}) = \lambda \cdot \mathbf{Ax} = \mathbf{0} \Rightarrow \lambda\mathbf{x} \in N(\mathbf{A})$

Παραδείγματα

- αν $A = [5 \quad 4]$ το $Ax = 0 \Leftrightarrow 5x_1 + 4x_2 = 0 \Leftrightarrow x_1 = -\frac{4}{5}x_2$.
  - Άρα ο $N(A)$ αποτελείται από όλα τα $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ της μορφής $\begin{pmatrix} -\frac{4}{5}x_2 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} -\frac{4}{5} \\ 1 \end{pmatrix}$. Είναι δηλαδή ο $N(\mathbf{A})$ μια ευθεία που περνά από το 0 και το $\begin{pmatrix} -\frac{4}{5} \\ 1 \end{pmatrix}$.
- Αν $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{pmatrix}$ τότε $N(\mathbf{A}) = \{\mathbf{0}\}$
- Αν $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1+0 \\ 5 & 4 & 5+4 \\ 2 & 4 & 2+4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 5 & 4 & 9 \\ 2 & 4 & 6 \end{pmatrix}$ τότε (από Gauss) $\mathbf{U} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 4 & 4 \\ 0 & 0 & 0 \end{pmatrix}$ και το $\mathbf{Ux} = \mathbf{0}$ έχει λύσεις τα $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ της μορφής $x = \begin{pmatrix} -x_3 \\ -x_3 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$. Δηλαδή

ο $N(\mathbf{A})$ είναι δηλαδή μια ευθεία που περνά από το 0 και το $\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$. Ή

$$N(\mathbf{A}) = Span\left\{\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}\right\}.$$

- $\mathbf{A}$ $\mathrm{n} \times \mathrm{n}$ μη ιδιόμορφος. Τότε το $\mathbf{Ax} = \mathbf{0}$ έχει μία και μοναδική λύση, την $\mathbf{x} = 0$. Άρα $\mathrm{N}(\mathbf{A}) = \{\mathbf{0}\}$.

9. Η λύση m εξισώσεων με n αγνώστους

(Πηγή: Strang, Κεφάλαιο 2.2)

- Ψάχνουμε **<u>το σύνολο</u> των λύσεων** ενός συστήματος $\mathbf{Ax}=\mathbf{b}$ όπου $\mathbf{A}_{\mathbf{m}\times\mathbf{n}}$, $\mathbf{x}\in\mathbb{R}^n$, $\mathbf{b}\in\mathbb{R}^m$.

- Όπως και στην $n\times n$ περίπτωση: οι λύσεις δεν αλλάζουν αν κάνω γραμμομετασχηματισμούς στον $\mathbf{A}$ και στο $\mathbf{b}$ ταυτόχρονα.

- **Εφαρμόζοντας αλγόριθμο Gauss** φθάνουμε τώρα όχι σε άνω τριγωνικό U, αλλά σε «**κλιμακωτό**» U, όπου η μη-θεραπεύσιμη περίπτωση (συναντάω 0 από τα οποία δεν μπορώ να απαλλαγώ) είναι ο κανόνας.

Παράδειγμα:

$$A=\begin{bmatrix}1 & 3 & 3 & 2\\ 2 & 6 & 9 & 5\\ -1 & -3 & 3 & 0\end{bmatrix}\rightarrow\begin{bmatrix}1 & 3 & 3 & 2\\ 0 & 0 & 3 & 1\\ 0 & 0 & 6 & 2\end{bmatrix}\rightarrow\begin{bmatrix}1 & 3 & 3 & 2\\ 0 & 0 & 3 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}$$

$$\gamma'_2=\gamma_2-2\gamma_1 \quad \gamma''_3=\gamma'_3-2\gamma'_2$$
$$\gamma'_3=\gamma_2+\gamma_1$$

Οι **οδηγοί**

- δεν είναι πια πάνω στη διαγώνιο, αλλά
- **είναι το πλησιέστερο μη-μηδενικό στοιχείο κάθε γραμμής**,
- ο οδηγός, αλλά και τα μηδενικά που προηγούνται να έχουν **από κάτω τους στήλη μηδενικών**.

Κατά τα άλλα λειτουργώ όπως στην $n \times n$ περίπτωση: σε κάθε βήμα προσθέτω πολλαπλάσια της γραμμής του οδηγού στις από κάτω της, ώστε να μηδενιστούν τα στοιχεία κάτω από τον οδηγό.

Έτσι φθάνω σε **κλιμακωτό πίνακα** του οποίου η γενική μορφή έχει ως εξής:

$$\begin{bmatrix} 0 & \cdots & 0 & * & * & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & 0 & 0 & \cdots & 0 & * & * & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 0 & \cdots & 0 & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ \vdots & \vdots & 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

$$\underbrace{\qquad\qquad}_{?} \qquad \underbrace{\qquad\qquad}_{?} \qquad \underbrace{\qquad\qquad}_{?}$$

Με γενικά χαρακτηριστικά:

- Πρώτες έρχονται οι μη-μηδενικές γραμμές και τελευταίες οι γραμμές που είναι γεμάτες μηδέν.
- Πρώτο μη-μηδενικό στοιχείο κάθε μη-μηδενικής γραμμής ονομάζεται **οδηγός**.
- Κάτω από κάθε οδηγό έχει **στήλη μηδενικών**.
- Κάθε οδηγός στα δεξιά του οδηγού προηγούμενων γραμμών.

Όπως στη nxn περίπτωση μπορεί να δείξει λοιπόν κανείς ότι:

Αν $\mathbf{A}$ mxn τότε θα υπάρχουν mxm κάτω τριγωνικός $\mathbf{L}$ με "1" στη διαγώνιο και mxm πίνακας μεταθέσεων $\mathbf{P}$, καθώς και mxn κλιμακωτός $\mathbf{U}$ τέτοιοι ώστε $\mathbf{PA} = \mathbf{LU}$.

Προσοχή στις διαστάσεις:

- $\mathbf{U}$ έχει τις ίδιες με τον $\mathbf{A}$,
- ενώ ο $\mathbf{P}$ και ο $\mathbf{L}$ είναι και οι δύο τετραγωνικοί με διαστάσεις τέτοιες που να μπορώ να πολ/σω τους $\mathbf{A}$ και $\mathbf{U}$ από αριστερά: γραμμές του $\mathbf{A}$ επί γραμμές του $\mathbf{A}$

Προσοχή στη κατασκευήτου $\mathbf{L}$:

- $\mathbf{L}$ περιέχει πάλι τα αντίθετα των συντελεστών,
- **αλλά όχι στη θέση του στοιχείου που μηδενίζανε**!
  - Τώρα είναι μεν **στη «σωστή» γραμμή ο καθένας**,
  - αλλά **στην κ-στήλη (κάτω από τον κ-«1») μπαίνουν οι συντελεστές που χρησιμοποιήθηκαν με τον κ-οδηγό** (ακόμα και αν αυτός δεν ήταν στη κ-στήλη),

Στο παράδειγμά μας :
$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$$

Δίνεται ο ακόλουθος πίνακας Α:

$$A = \begin{bmatrix} 6 & 0 & 12 & 6 & 6 \\ 2 & 0 & 4 & 5 & 2 \\ -1 & -1 & -2 & -3 & -3 \\ 3 & 1 & 6 & 11 & 5 \end{bmatrix},$$

(0.5 μον.) Να βρεθεί η παραγοντοποίηση PA=LU, όπου P πίνακας μεταθέσεων, L κάτω τριγωνικός (με 1 στη διαγώνιο) και U κλιμακωτός.

Έστω $\mathbf{b} = \begin{pmatrix} 6 & 8 & -6 & 20 \end{pmatrix}^T$ Περιγράψτε το σύνολο των λύσεων του συστήματος $\mathbf{Ax = b}$ ως $\mathbf{x}_{\text{ειδική}} + \mathrm{Span}(\mathbf{v_1,...,v_k})$ για κατάλληλα **ανεξάρτητα** $\mathrm{v}_1,...,\mathrm{v}_k$. (πρώτα εφαρμόζετε τον αλγόριθμο του <u>Gauss</u> και ακολούθως λύστε τις εξισώσεις ώστε να εκφράσετε τις βασικές μεταβλητές συναρτήσει των ελεύθερων. Αν ο αλγόριθμος Gauss υποδείξει ότι το σύστημα είναι αδύνατο, **πρώτα επισημάνετε το**. Ακολούθως, αγνοείστε τους περιορισμούς που το καθιστούν αδύνατο και προχωρήστε στην επίλυσή του.)

*Λύσεις του **Ax=0***

(δηλ. $b=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$. Το σύστημα λέγεται τότε «ομογενές».)

Κατ’ αρχάς $\mathbf{Ax}=\mathbf{0} \Leftrightarrow \mathbf{Ux}=\mathbf{0}$, όπου $\mathbf{U}$ ο κλιμακωτός στον οποίο φθάνουμε με αλγόριθμο Gauss.

$$Ux=0:\begin{bmatrix}1&3&3&2\\0&0&3&1\\0&0&0&0\end{bmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

- **Κάθε μεταβλητή αντιστοιχεί σε μία στήλη** του U. Οι μεταβλητές που αντιστοιχούν σε **<u>στήλη που δεν έχει οδηγό</u> ονομάζονται <u>ελεύθερες</u>, οι άλλες <u>βασικές</u>**. <u>Εδώ:</u> βασικές: x_1, x_3, ελεύθερες: x_2, x_4.

- Κάθε μη-μηδενική γραμμή του U αντιστοιχεί σε μία εξίσωση. Λύσε τις έτσι ώστε σε όλες **να εκφράζονται <u>οι βασικές μεταβλητές συναρτήσει των ελεύθερων</u>.** (Δηλ.: αριστερά του “=” <u>μόνο</u> βασικές, δεξιά του “=” <u>μόνο</u> ελεύθερες).

  Εδώ: $3x_3+x_4=0 \Rightarrow x_3=-\frac{1}{3}x_4$

  $$x_1+3x_2+3x_3+2x_4=0 \Rightarrow x_1=-3x_2-x_4$$

- Λύσεις: τα x της μορφής: **οι ελεύθερες ως έχουν, οι βασικές συναρτήσει των ελεύθερων.**

Εδώ: Όλα τα x που γράφονται ως

$$x=\begin{pmatrix}-3x_2-x_4\\ x_2\\ -1/3\,x_4\\ x_4\end{pmatrix}=x_2\begin{pmatrix}-3\\ 1\\ 0\\ 0\end{pmatrix}+x_4\begin{pmatrix}-1\\ 0\\ -1/3\\ 1\end{pmatrix}$$

Θέσεις ελεύθερων

Άρα $\{$ λύσεις της $Ax=0$ $\}=Span\left(\begin{pmatrix}-3\\ 1\\ 0\\ 0\end{pmatrix},\begin{pmatrix}-1\\ 0\\ -1/3\\ 1\end{pmatrix}\right).$

Γενικώς:

- ✓ Φτιάχνω **τόσα διανύσματα** $\mathbf{v}_1,\ldots\mathbf{v}_k$ **όσα έχω ελεύθερες μεταβλητές**.
- ✓ Σε καθένα από αυτά έχω «θέσεις» ελεύθερων και βασικών μεταβλητών. **Τις «θέσεις» των ελεύθερων τις γεμίζω με 0/1** έτσι ώστε σε κάθε $\mathbf{v}_j$ να έχω ένα “1” και κάθε «θέση» ελεύθερης να έχει κάπου ένα «1».
- ✓ Για κάθε $\mathbf{v}_j$ έχω δώσει τιμές στις ελεύθερες. Με αυτές τις τιμές **μπαίνω στις εξισώσεις και βρίσκω τις τιμές των βασικών**.
- ✓ Τέλος: $\{$ λύσεις της $Ax=0\ \} = Span\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}$.

π.χ. 1 εξίσωση: $x_2+2x_3+x_4=0 \Rightarrow x_2=-2x_3-x_4$

βασική: x_2, ελεύθερες: x_1,x_3,x_4

$$\Rightarrow \{\text{ λύσεις της } Ax=0\ \} = Span\left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\-2\\1\\0\end{pmatrix},\begin{pmatrix}0\\-1\\0\\1\end{pmatrix}\right\}$$

Θέσεις ελεύθερων

Ax=b

Με τους ίδιους γραμμομετασχηματισμούς στο $\mathbf{b}$ όπως στο $\mathbf{A}$, το b μεταβάλλεται σε κάποιο c.

$$\left[\mathbf{A}|\mathbf{b}\right] \to \left[\mathbf{U}|\mathbf{c}\right]$$

Στην περίπτωσή μας: $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \to \begin{pmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 + b_1 \end{pmatrix} \to \begin{pmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - 2b_2 + 5b_1 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$

Κατ' αρχάς: Για να υπάρχουν λύσεις πρέπει το b να είναι τέτοιο ώστε για το c που προκύπτει **οι συντεταγμένες που αντιστοιχούν στις μηδενικές γραμμές του U να είναι ίσες με 0**.

$$\boxed{\text{Εδώ:}}\, c_3 = b_3 - 2b_2 + 5b_1 = 0$$

Δηλαδή: το σύνολο των εφικτών b (, που εξ ορισμού είναι το span των στηλών,) δεν είναι όλος ο $\mathbb{R}^3$, αλλά μόνο εκείνα τα b που ικανοποιούν τον παραπάνω περιορισμό. Αυτός μας λέει ότι πρόκειται για τα b που είναι κάθετα στο $(1 \quad -2 \quad 5)^T$. Πρόκειται λοιπόν για ένα επίπεδο στον $\mathbb{R}^3$ που περνά από το 0. Αυτό το ξέραμε και λόγω του ότι δυο από τις στήλες του A είναι γραμμικός συνδυασμός των υπολοίπων. Άρα των span των στηλών στην ουσία είναι span δύο στηλών, άρα επίπεδο.

Γενικώς: Αν r οδηγοί $\Rightarrow m - r$ μηδενικές γραμμές. Άρα χρειάζομαι

$$\boxed{c_{r+1} = c_{r+2} = c_m = 0}$$

Αν αυτοί οι περιορισμοί πληρούνται έχω λύσεις. Ποιες είναι αυτές;

π.χ. $b=\begin{pmatrix}1\\5\\5\end{pmatrix} \Rightarrow c=\begin{pmatrix}1\\3\\0\end{pmatrix}$

- Πάλι λύνω τις εξισώσεις που αντιστοιχούν στις μη-μηδενικές γραμμές του U ώστε να εκφράσω τις βασικές συναρτήσει των ελεύθερων.

$$\begin{aligned} x_1+3x_2+3x_3+2x_4&=1 \\ 3x_3+1x_4&=3 \end{aligned} \Rightarrow \begin{aligned} x_3&=1-\frac{1}{3}x_4 \\ x_1&=-2-3x_2-x_4 \end{aligned}$$

- Άρα λύσεις τα x της μορφής

$$x=\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}-2-3x_2-x_4\\x_2\\1-\frac{1}{3}x_4\\x_4\end{pmatrix}=\underset{x_{ειδική}}{\begin{pmatrix}-2\\0\\1\\0\end{pmatrix}}\underbrace{+x_2\begin{pmatrix}-3\\1\\0\\0\end{pmatrix}+x_4\begin{pmatrix}-1\\0\\-\frac{1}{3}\\1\end{pmatrix}}_{+\text{λύσεις της } Ax=0}$$

<u>Άρα:</u> { οι λύσεις της $Ax=b$ } = $x_{ειδική}$ + { λύσεις της $Ax=0$ }

<u>Σημείωση:</u> Το ρόλο της $x_{ειδικής}$ εδώ τον έπαιξε η λύση για την οποία ελεύθερες μεταβλητές = 0. Ωστόσο **<u>οποιαδήποτε</u> άλλη λύση** της $Ax=b$ μπορεί να παίξει το ρόλο της $x_{ειδικής}$.

<u>Σημείωση:</u> Οι λύσεις της $Ax=b$ είναι **ο κατά** $x_{ειδική}$ **«μετατοπισμένος» μηδενόχωρος του** A.

$\Rightarrow$ Οι λύσεις της $Ax=b$ <u>δεν είναι</u> υπόχωρος.

Συμπεράσματα:

Έστω $r =$ **αριθμός οδηγών = τάξη του πίνακα**

$\Rightarrow n-r$ **ελεύθερες μεταβλητές** και $m-r$ **μηδενικές γραμμές του** U.

Αν $m-r =$

- 0 $\Rightarrow$ δεν υπάρχουν μηδενικές γραμμές του U.
  Άρα: δεν υπάρχουν περιορισμοί στην ύπαρξη λύσης.
  $\Rightarrow$ $\boxed{\text{Έχω λύση για κάθε } b \in R^m.}$
- >0
  - Αν κάποιο c_j με $j>r$ είναι $c_j \neq 0$ $\Rightarrow$ <u>δεν έχω λύση</u>.
  - Αν $c_{r+1} = \ldots = c_m = 0$ $\Rightarrow$ <u>υπάρχουν λύσεις</u>.

<u>Αν υπάρχουν λύσεις:</u>

και $n-r =$

- $=0$ c_j δεν υπάρχουν ελεύθερες μεταβλητές.
  $\Rightarrow N(A) = \{0\} \Rightarrow$ μία ακριβώς λύση
  $\{$ λύσεις του $Ax=b$ $\} = \{x_{ειδική}\}$
- >0
  $\{$ λύσεις του $Ax=b$ $\} = \{x_{ειδική}\} + \underbrace{Span(v_1,\ldots,v_{n-r})}_{N(A)}$

Άσκηση: $\mathbf{A}=\begin{bmatrix}1 & -1 & 2 & 1 & 1 & 2\\ 2 & -2 & 4 & 2 & 2 & 4\\ 3 & -3 & 6 & 2 & 3 & 5\end{bmatrix},\mathbf{b}=\begin{bmatrix}1\\2\\5\end{bmatrix}$

$$\left[\mathbf{A}|\mathbf{b}\right]=\left[\begin{array}{cccccc|c}1 & -1 & 2 & 1 & 1 & 2 & 1\\ 2 & -2 & 4 & 2 & 2 & 4 & 2\\ 3 & -3 & 6 & 2 & 3 & 5 & 5\end{array}\right]\to\left[\begin{array}{cccccc|c}1 & -1 & 2 & 1 & 1 & 2 & 1\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & -1 & 0 & -1 & 2\end{array}\right]$$

$$\to\left[\begin{array}{cccccc|c}1 & -1 & 2 & 1 & 1 & 2 & 1\\ 0 & 0 & 0 & -1 & 0 & -1 & 2\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\end{array}\right],\quad \mathbf{P}=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix},\mathbf{L}=\begin{bmatrix}1 & 0 & 0\\ 3 & 1 & 0\\ 2 & 0 & 1\end{bmatrix},$$

Βασικές: x_1, x_4, Ελεύθερες: x_2, x_3, x_5, x_6

$x_4=-2-x_6$,

$$x_1=1+x_2-2x_3-x_4-x_5-2x_6=1+x_2-2x_3-\left(-2-x_6\right)-x_5-2x_6$$

$$=3+x_2-2x_3-x_5-x_6$$

Λύσεις τα $\mathbf{x}$ της μορφής

$$\mathbf{x}=\begin{bmatrix}3+x_2-2x_3-x_5-x_6\\ x_2\\ x_3\\ -2+x_6\\ x_5\\ x_6\end{bmatrix}=\begin{bmatrix}3\\0\\0\\-2\\0\\0\end{bmatrix}+x_2\begin{bmatrix}1\\1\\0\\0\\0\\0\end{bmatrix}+x_3\begin{bmatrix}-2\\0\\1\\0\\0\\0\end{bmatrix}+x_5\begin{bmatrix}-1\\0\\0\\0\\1\\0\end{bmatrix}+x_6\begin{bmatrix}-1\\0\\0\\1\\0\\1\end{bmatrix}$$

Λύσεις = $\mathbf{x}_{\varepsilon\iota\delta\iota\kappa\acute{\eta}}+\mathrm{N}\left(\mathbf{A}\right)$, όπου:

$$\mathbf{x}_{\varepsilon\iota\delta\iota\kappa\acute{\eta}} = \begin{bmatrix} 3 \\ 0 \\ 0 \\ -2 \\ 0 \\ 0 \end{bmatrix}, \mathrm{N}\left(\mathbf{A}\right) = Span\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Άσκηση: $\mathbf{A}=\begin{bmatrix}2&0&0&2&1&0\\2&3&0&2&2&0\\2&-6&0&3&0&2\\6&9&0&9&11&7\\0&0&0&4&4&8\end{bmatrix},\mathbf{b}=\begin{bmatrix}2\\3\\4\\24\\16\end{bmatrix}$

$$\left[\mathbf{A}\middle|\mathbf{b}\right]=\left[\begin{array}{cccccc|c}2&0&0&2&1&0&2\\2&3&0&2&2&0&3\\2&-6&0&3&0&2&4\\6&9&0&9&11&7&24\\0&0&0&4&4&8&16\end{array}\right]\to\left[\begin{array}{cccccc|c}2&0&0&2&1&0&2\\0&3&0&0&1&0&1\\0&-6&0&1&-1&2&2\\0&9&0&3&8&7&18\\0&0&0&4&4&8&16\end{array}\right]$$

$$\to\left[\begin{array}{cccccc|c}2&0&0&2&1&0&2\\0&3&0&0&1&0&1\\0&0&0&1&1&2&4\\0&0&0&3&5&7&15\\0&0&0&4&4&8&16\end{array}\right]\to\left[\begin{array}{cccccc|c}2&0&0&2&1&0&2\\0&3&0&0&1&0&1\\0&0&0&1&1&2&4\\0&0&0&0&2&1&3\\0&0&0&0&0&0&0\end{array}\right]$$

$$\mathbf{P}=\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix},\ \mathbf{L}=\begin{bmatrix}1&0&0&0&0\\1&1&0&0&0\\1&-2&1&0&0\\3&3&3&1&0\\0&0&4&0&1\end{bmatrix},$$

Βασικές: x_1, x_2, x_4, x_5, Ελεύθερες: x_3, x_6

- $2x_5 + x_6 = 3 \Rightarrow x_5 = 3/2 - 1/2\,x_6$
- $x_4 + x_5 + 2x_6 = 4 \Rightarrow x_4 = 4 - x_5 - 2x_6 \Rightarrow x_4 = 4 - \left(3/2 - 1/2\,x_6\right) - 2x_6$

  $\Rightarrow x_4 = 5/2 - 3/2\,x_6$
- $3x_2 + x_5 = 1 \Rightarrow 3x_2 = 1/3 - 1/3\,x_5 \Rightarrow x_2 = 1/3 - 1/3\left(3/2 - 1/2\,x_6\right)$

  $\Rightarrow x_2 = -1/6 + 1/6\,x_6$.
- $2x_1 + 2x_4 + x_5 = 2 \Rightarrow x_1 = 1 - x_4 - 1/2\,x_5$

  $\Rightarrow x_1 = 1 - 5/2 - 3/4 + 3/2\,x_6 + 1/4\,x_6$

  $\Rightarrow x_1 = -9/4 + 7/4\,x_6$

Λύσεις τα $\mathbf{x}$ της μορφής

$$\mathbf{x} = \begin{bmatrix} -9/4 + 7/4\,x_6 \\ -1/6 + 1/6\,x_6 \\ x_3 \\ 5/2 - 3/2\,x_6 \\ 3/2 - 1/2\,x_6 \\ x_6 \end{bmatrix} = \begin{bmatrix} -9/4 \\ -1/6 \\ 0 \\ 5/2 \\ 3/2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} 7/4 \\ 1/6 \\ 0 \\ -3/2 \\ -1/2 \\ 1 \end{bmatrix}$$

Λύσεις = $\mathbf{x}_{\varepsilon\iota\delta\iota\kappa\acute{\eta}} + \mathrm{N}(\mathbf{A})$, όπου:

$$\mathbf{x}_{\varepsilon\iota\delta\iota\kappa\acute{\eta}} = \begin{bmatrix} -9/4 \\ -1/6 \\ 0 \\ 5/2 \\ 3/2 \\ 0 \end{bmatrix}, \mathrm{N}(\mathbf{A}) = Span\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 7/4 \\ 1/6 \\ 0 \\ -3/2 \\ -1/2 \\ 1 \end{bmatrix} \right\}.$$

10. Γραμμική ανεξαρτησία, βάσεις και διάσταση.

(Πηγή: Strang, Κεφάλαιο 2.32)

Γραμμική Ανεξαρτησία

Έστω k στοιχεία ενός διανυσματικού χώρου V: $v_1,\ldots,v_k \in V$. Τότε για οποιαδήποτε $\lambda_1,\ldots,\lambda_k \in R$ έχουμε $\lambda_1 v_1 + \ldots + \lambda_\kappa v_k \in V$.

Επίσης:

Αν $\lambda_1 = \ldots = \lambda_k = 0 \Rightarrow \lambda_1 v_1 + \ldots + \lambda_k v_k = 0$.

Ανεξαρτησία των $\{v_1,\ldots,v_k\}$ έχουμε αν αυτός είναι ο μόνος γραμμικός συνδυασμός τους που ισούται με 0.

Ορισμός: $\{v_1,\ldots,v_k\} \in V$ **(γραμμικά) ανεξάρτητα**

Τότε και μόνο τότε

αν από $\lambda_1 v_1 + \ldots + \lambda_\kappa v_k = 0$ **έπεται** ότι: $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$.

<u>Παρατήρηση:</u> $\{v_1,\ldots,v_k\}$ ανεξάρτητα **ισοδυναμεί** με:

Δε μπορώ να γράψω ένα από αυτά ως γραμμικό συνδυασμό των υπολοίπων.

π.χ. αν $v_1 = \mu_2 v_2 + \ldots + \mu_k v_k \Rightarrow v_1 - \mu_2 v_2 - \ldots - \mu_k v_k = 0$ ενώ $\mu_j \neq 0$.

Παραδείγματα:

- $x, y \in R^n$ εξαρτημένα $\Leftrightarrow x, y$ συγγραμικά.
- $x, y, z \in R^n$ εξαρτημένα $\Leftrightarrow$ κάποιο από αυτά ανήκει στο επίπεδο που δίνουν τα άλλα δύο και το 0: π.χ. $z \in span(x, y)$.
- Αν κάποιο από τα $v_i = 0$, τότε εξαρτημένα.
- Οι στήλες τριγωνικού πίνακα με μη μηδενικά στοιχεία στη διαγώνιο είναι ανεξάρτητες:

$$\lambda_1 \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} * \\ * \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \ldots + \lambda_n \begin{bmatrix} * \\ \vdots \\ * \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Τελευταία εξίσωση $\Rightarrow \lambda_n = 0$.

Αντικαθιστώ αυτό στην προτελευταία $\Rightarrow \lambda_{n-1} = 0$

$\vdots$

- Παρομοίως για τις γραμμές τριγωνικού πίνακα,

- καθώς και για τις γραμμές (ή τις στήλες) κλιμακωτού που <u>περιέχουν οδηγούς</u>.

Οι r γραμμές με οδηγούς κλιμακωτού είναι ανεξάρτητες.

Οι r r στήλες με οδηγούς κλιμακωτού είναι ανεξάρτητες.

Μέθοδος για να ελέγξω αν $v_1,\ldots,v_k \in R^n$ ***είναι ανεξάρτητα***

Σχηματίζω τον $n \times n$ πίνακα $V = \left[v_1 | \ldots | v_k\right]$ και βρίσκω τον μηδενόχωρο $N(V)$.
Αν αυτός αποτελείται μόνο από $\{0\}$, τα $v_1,\ldots,v_k$ είναι ανεξάρτητα. (Δηλαδή όταν δεν έχω ελεύθερες μεταβλητές).

Διότι:

$$Vx = 0 \Rightarrow x = 0$$

ισοδυναμεί με:

$$x_1 v_1 + \ldots + x_n v_n = 0 \Rightarrow x_1 = \ldots = x_n = 0$$

που είναι ακριβώς ορισμός της ανεξαρτησίας.

Παράγοντας ένα υπόχωρο

Ορισμός. Ένα σύνολο διανυσμάτων $\{v_1,\ldots,v_k\} \in U$ **παράγει ένα υπόχωρο** U όταν οποιοδήποτε στοιχείο του U γράφεται ως γραμμικός συνδυασμός των $v_1,\ldots,v_k$.

Τότε θα έχουμε :

$$U = Span(v_1,\ldots,v_k)$$

<u>Παρατήρηση:</u> Ένα στοιχείο του U μπορεί να γράφεται **με (ενδεχομένως) διαφορετικούς τρόπους** ως γραμμικός συνδυασμός των $v_1, \ldots, v_k$.

π.χ.

- Έστω $v \,\&\, w$ δύο συγγραμικά διανύσματα και $U = Span(v, w)$. Τότε οποιοδήποτε $x \in U$ γράφεται είτε ως πολλαπλάσιο του v, είτε ως πολλαπλάσιο του w (είτε και με άλλους άπειρους τρόπους).
- Παρομοίως αν v, w, z ανήκουν στο ίδιο επίπεδο που περνά από 0 και $U = Span(v, w, z)$.

Ορισμός: **Βάση** ενός υπόχωρου U ονομάζεται **ένα σύνολο διανυσμάτων** $\{v_1,...,v_k\}\in U$ που έχουν τις εξής δύο ιδιότητες:

1) Είναι **ανεξάρτητα**

2) **Παράγουν** τον U .

Σημείωση:

Ένα $v\in U$ γράφεται με μοναδικό τρόπο ως γραμμικός συνδυασμός της βάσης.

Για κάθε υπόχωρο υπάρχουν άπειρες βάσεις.

π.χ.

- Αν U ένα **επίπεδο**: τότε οποιαδήποτε δύο μη – συγγραμικά διανύσματα είναι βάση του.
- Οι στήλες ενός μη – ιδιόμορφου $n\times n$ πίνακα είναι βάση του R^n :

Διότι: Το $Ax=b$ έχει πάντα λύση (άρα παράγουν) και

Το $Ax=0$ μόνο τη μηδενική (άρα είναι ανεξάρτητες)

Διάσταση

<u>Πρόταση και ορισμός</u>: Δύο βάσεις ενός υπόχωρου έχουν τον **<u>ίδιο αριθμό διανυσμάτων</u>**, τη **<u>διάσταση</u>** του U (που επομένως χαρακτηρίζει τον U). Γράφουμε $\dim(U)$.

Απόδειξη.

Έστω 2 βάσεις του ίδιου υπόχωρου $\mathbf{v_1,...,v_m}$ και $\mathbf{w_1,...,w_n}$ με διαφορετικό αριθμό στοιχείων. Έστω π.χ. $m<n$.

Αφου $\mathbf{v_1,...,v_m}$ βάση μπορώ να γραψω

$$\mathbf{w}_j = a_{1j}\mathbf{v_1} + ... + a_{mj}\mathbf{v}_m,\ j=1,...,n$$

με κατάλληλα $a_{1j},...,a_{mj},\ j=1,...,n$.

Θέτοντας $\mathbf{V}=\left[\mathbf{v_1}|...|\mathbf{v}_m\right]$, $\mathbf{W}=\left[\mathbf{w_1}|...|\mathbf{w}_n\right]$ και $\mathbf{A}_{m\times n}=\left[a_{ij}\right]$ οι παραπάνω σχέσεις γράφονται ως

$$\mathbf{W = VA}$$

Αφού όμως $m<n$ θα υπάρχει $\mathbf{c}\neq\mathbf{0}$ με $\mathbf{Ac=0}\Rightarrow\mathbf{VAc=0}\Rightarrow\mathbf{Wc=0}$

Αυτό όμως αντιβαίνει στην ανεξαρτησία των $\mathbf{w_1,...,w_n}$. **Αντίφαση!!!**

Αποδεικνύεται ότι :

Αν η διάσταση του $V = k$, τότε

- αν έχω περισσότερα από k διανύσματα: θα είναι εξαρτημένα

π.χ. 3 διανύσματα σε επίπεδο

- αν έχω περισσότερα από k, και αυτά παράγουν τον V, τότε αποκλείοντας μερικά (κατάλληλα) φθάνω σε βάση

π.χ. 3 διανύσματα σε επίπεδο που δεν είναι όλα συγγραμμικά

- αν έχω λιγότερα από k : δεν θα παράγουν τον V

π.χ. 1 διάνυσμα σε επίπεδο θα παράγει μόνο μια ευθεία

- αν έχω λιγότερα από k, και αυτά είναι ανεξάρτητα, τότε μπορώ να συμπληρώσω σε βάση

π.χ. 1 διάνυσμα σε επίπεδο, συμπληρώνω με ένα μη-συγγραμμικό

- αν έχω ακριβώς k, τότε ανεξαρτησία $\Rightarrow$ παραγωγή και παραγωγή $\Rightarrow$ ανεξαρτησία: αν έχω τη μία ιδιότητα θα έχω και την άλλη.

π.χ. 2 διάνυσμα σε επίπεδο, αν είναι μη-συγγραμμικά φθάνουν για να παράξουν τον V, ενώ για να φθάνουν να παράξουν τον V θα πρέπει να είναι μη-συγγραμμικά.

11. Οι τέσσερις θεμελιώδεις υπόχωροι

(Πηγή: Strang, Κεφάλαιο 2.4)

Έστω $\mathbf{A}$ ένας $m \times n$ πίνακας. Έχουμε ήδη γνωρίσει δυο **υπόχωρους που συνδέονται με το $\mathbf{A}$:**

- Το **χώρο στηλών** του $\mathbf{A}$, $R(\mathbf{A})$:=Span(των στηλών του $\mathbf{A}$)$\subseteq \mathbb{R}^m$.
- Το **μηδενόχωρο** του $\mathbf{A}$, $\mathrm{N}(\mathbf{A}) := \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{Ax} = \mathbf{0}\} \subseteq \mathbb{R}^n$.

Προσθέτουμε άλλους δύο:

- Το **χώρο γραμμών** του $\mathbf{A}$, $R(\mathbf{A}^\mathrm{T})$:=Span(των γραμμών του $\mathbf{A}$)= Span(των στηλών του $\mathbf{A}^\mathrm{T}$ $\mathbf{A}$)$\subseteq \mathbb{R}^n$.
- Τον **αριστερό μηδενόχωρο** του $\mathbf{A}$:= το μηδενόχωρο του $\mathbf{A}^\mathrm{T}$, $\mathrm{N}(\mathbf{A}^\mathrm{T}) \subseteq \mathbb{R}^m$.

Στόχος αυτού του κεφαλαίου: **κατασκευή μιας βάσης** για καθένα εξ αυτών.

Χρήσιμη σε αυτή τη κατεύθυνση: η παραγοντοποίηση $\mathbf{PA} = \mathbf{LU}$.

Παράδειγμα:

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix} \longrightarrow \mathbf{U} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Βάση του χώρου γραμμών του $\mathbf{A}$, $R\left(\mathbf{A}^{\mathrm{T}}\right)$

Παρατηρούμε ότι:

- Εκ κατασκευής οι γραμμές του $\mathbf{U}$ είναι γραμμικοί συνδυσασμοί των γραμμών του $\mathbf{A}$. Έπεται:
  Span(των γραμμών του $\mathbf{U}$) $\subseteq$ Span(των γραμμών του $\mathbf{A}$)

- Όμως οι γραμμοπράξεις μπορούν να αντιστραφούν. Άρα οι γραμμές του $\mathbf{U}$ είναι γραμμικοί συνδυασμοί των γραμμών του $\mathbf{A}$. Έπεται:
  Span(των γραμμών του $\mathbf{U}$) $\subseteq$ Span(των γραμμών του $\mathbf{A}$)

Άρα:

Span(των γραμμών του $\mathbf{U}$ **)** $=$ **Span(των γραμμών του** $\mathbf{A}$ **)**

$$R\left(\mathbf{U}^{\mathrm{T}}\right)=R\left(\mathbf{A}^{\mathrm{T}}\right)$$

Άρα: αρκεί να βρώ βάση του $R\left(\mathbf{U}^{\mathrm{T}}\right)$.

- Πάρε ως βάση του $R\left(\mathbf{A}^{\mathrm{T}}\right)$ τις **μη-μηδενικές γραμμές του** $\mathbf{U}$
- Αυτές θα είναι **ανεξάρτητες** (μη μηδενικές κλιμακωτου)
- Άρα θα είναι βάση του $R\left(\mathbf{U}^{\mathrm{T}}\right)$.
- Άρα θα είναι βάση του $R\left(\mathbf{A}^{\mathrm{T}}\right)$.

Στο παράδειγμα: $\left\{\begin{pmatrix}1\\3\\3\\2\end{pmatrix},\begin{pmatrix}0\\0\\3\\1\end{pmatrix}\right\}$ είναι βάση του $R\left(\mathbf{A}^{\mathrm{T}}\right)$ (και του $R\left(\mathbf{U}^{\mathrm{T}}\right)$).

Διάσταση του $R\left(\mathbf{A}^{\mathrm{T}}\right)=$
=πλήθος μη-μηδενικών γραμμων του $\mathbf{U}$ **=πλήθος οδηγών=** r

Βάση Μηδενόχωρου του $\mathbf{A}$, $\mathrm{N}(\mathbf{A})$

- $\mathrm{N}(\mathbf{A}) := \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{A}\mathbf{x} = \mathbf{0}\} = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{U}\mathbf{x} = \mathbf{0}\}$
- Άρα $\mathrm{N}(\mathbf{A}) = \mathrm{N}(\mathbf{U})$

Λύσε $Ax = 0 \Leftrightarrow Ux = 0$, r οδηγοί.

Βάση του $\mathrm{N}(\mathbf{A})$: τα $n-r$ διανύσματα $\{\mathbf{v_1}, \dots, \mathbf{v_{n-r}}\}$ που προκύπτουν αν θέσω τις ελεύθερες μεταβλητές διαδοχικά 0 και 1.

Διότι: Προφανώς παράγουν τον $\mathrm{N}(\mathbf{A}) = \mathrm{N}(\mathbf{U})$ (Έτσι κατασκευάστηκαν). Είχαμε δει $\mathrm{N}(\mathbf{A}) = \mathrm{N}(\mathbf{U}) = Span\{\mathbf{v_1}, \dots, \mathbf{v_{n-r}}\}$.

Και: Η ανεξαρτησία των $\{\mathbf{v_1}, \dots, \mathbf{v_{n-r}}\}$ εξασφαλίζεται από το ότι για κάθε ελεύθερη μεταβλητή το “1” εμφανίζεται μόνο σε ένα από τα $\mathbf{v_k}$, τα άλλα έχουν “0”.

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} c_1 = 0 \\ c_2 = 0 \\ c_3 = 0 \end{matrix}.$$

$$\mathbf{v_1} \qquad \mathbf{v_2} \qquad \mathbf{v_3}$$

Έχουμε:

Διάσταση του $\mathrm{N}(\mathbf{A})$ = πλήθος ελεύθερων μεταβλητών = $n-r$

Βάση του χώρου στηλών του $\mathbf{A}$, $R(\mathbf{A})$

Παράδειγμα:

π.χ. $A=\begin{bmatrix}1 & 3 & 3 & 2\\ 2 & 6 & 9 & 5\\ -1 & -3 & 3 & 0\end{bmatrix} \longrightarrow U=\begin{bmatrix}1 & 3 & 3 & 2\\ 0 & 0 & 3 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}$

=3*πρώτη

=πρώτη + (1/3)*τρίτη

=3*πρώτη

=πρώτη + (1/3)*τρίτη

$R(\mathbf{A}) \neq R(\mathbf{U})$, κ[illegible]αία συντεταγμένη στον $R(\mathbf{U})$ =0, αλλά όχι στον $R(\mathbf{A})$. Όμως οι εξαρτήσεις στις στήλες τους ίδιες:

Ανεξάρτητες είναι στήλες του A αν και οι αντίστοιχες του U είναι ανεξάρτητες.

Αν κάποιες στήλες του A εξαρτημένες

οι ίδιες στήλες του U είναι εξαρτημένες.

$\Updownarrow$

$\exists \mathbf{x} \neq \mathbf{0}$, **με** $x_r = 0$ **για τις στήλες** που δε με ενδιαφέρουν τέτοιο που $\mathbf{Ax}=\mathbf{0}$. $\Leftrightarrow$ $\mathbf{Ux}=\mathbf{0}$.

Πάρε ως <u>βάση του</u> $R(\mathbf{A})$ **τις στήλες του** A που αντιστοιχούν στις στήλες του U που έχουν οδηγούς.

Διότι: με βάση τα παραπάνω, καθως οι στήλες αυτές του $\mathbf{U}$ είναι ανεξάρτητες θα είναι ανεξάρτητες και οι αντίστοιχες στηλες του $\mathbf{A}$.

Εδώ: 1η & 3η $\Rightarrow$ βάση $R(A):\left\{\begin{pmatrix}1\\2\\-1\end{pmatrix},\begin{pmatrix}3\\9\\3\end{pmatrix}\right\}$.

Διάσταση του $R(\mathbf{A})$ = **πλήθος οδηγών=** r

Βάση Αριστερού Μηδενόχωρου του $\mathbf{A}$, $\mathrm{N}\left(\mathbf{A}^{\mathrm{T}}\right)$

- $\mathrm{N}\left(\mathbf{A}^{\mathrm{T}}\right)$ είναι ο Μηδενόχωρος του $\mathbf{A}^{\mathrm{T}}$, δηλαδή τα $:=\left\{\mathbf{y}\in\mathbb{R}^m \middle| \mathbf{A}^T\mathbf{y}=\mathbf{0}\right\}=\left\{\mathbf{y}\in\mathbb{R}^m \middle| \mathbf{y}^T\mathbf{A}=\mathbf{0}^T\right\}$, εξού και αριστερος μηδενόχωρος.

Εδώ θα βρούμε πρώτα την $\dim\left[\mathrm{N}\left(\mathbf{A}^{\mathrm{T}}\right)\right]$:

Για κάθε πίνακα $\mathbf{A}$ γνωρίζουμε:

Πλήθος ελεύθερων μεταβλ. + πλήθος βασικών μεταβλ.= πληθος στηλών

$$n-r \quad + \quad r \quad = \quad n$$

$$\dim\left[\mathrm{N}(\mathbf{A})\right] \quad + \quad \dim\left[R(\mathbf{A})\right] \quad = \text{πληθος στηλών } \mathbf{A}$$

Εφαρμόζοντας το τελευταίο για τον $\mathbf{A}^{\mathrm{T}}$ παιρνουμε:

$$\dim\left[\mathrm{N}\left(\mathbf{A}^{\mathrm{T}}\right)\right] \quad + \quad \dim\left[R\left(\mathbf{A}^{\mathrm{T}}\right)\right] \quad = \text{πληθος στηλών } \mathbf{A}^{\mathrm{T}}$$

Άρα $\quad \dim\left[\mathrm{N}\left(\mathbf{A}^{\mathrm{T}}\right)\right]=m-r$

Άρα ψάχνω $m-r$ ανεξάρτητα $\left\{\mathbf{y}_1,\ldots,\mathbf{y}_{\mathbf{m-r}}\right\}$ τέτοια που $\mathbf{y}_i^T\mathbf{A}=\mathbf{0}$

Όμως από $\mathbf{PA}=\mathbf{LU}$ παίρνω $\mathbf{L}^{-1}\mathbf{PA}=\mathbf{U}$ που σημαίνει ότι οι $m-r$ τελευταίες γραμμές του $\mathbf{L}^{-1}\mathbf{P}$ έχουν ακριβώς αυτή την ιδιότητα, καθως αντιστοιχουν στις μηδενικές γραμμές του $\mathbf{U}$ (σκεφτείτε το με το μνημονικό κανόνα).

Πάρε ως **βάση του** $\mathrm{N}\left(\mathbf{A}^{\mathrm{T}}\right)$ τις $m-r$ τελευταίες γραμμές του $\mathbf{L}^{-1}\mathbf{P}$.

Αρκεί να δείξουμε ότι είναι ανεξάρτητα!

Όμως

- οι γραμμές τριγωνικού πίνακα (L) με 1 στη διαγώνιο είναι ανεξάρτητες.

- Παρομοίως και οι γραμμές και του αντιστρόφου του που είναι παλι τριγωνικός.
- Ο πολ/μος με P από δεξιά εναλλάσσει στήλες και δε χαλάει την ανεξατησία των γραμμών.

Θεμελιώδες Θεωρημα της Γραμμικής Αλγεβρας I.

Αν $\mathbf{A}_{m\times n}$ με r οδηγούς τότε:

1. $\dim\left[R(\mathbf{A})\right]=r, \quad R(\mathbf{A})\subseteq\mathbb{R}^m$
2. $\dim\left[N(\mathbf{A})\right]=n-r, \quad N(\mathbf{A})\subseteq\mathbb{R}^n$
3. $\dim\left[R(\mathbf{A}^T)\right]=r, \quad R(\mathbf{A}^T)\subseteq\mathbb{R}^n$
4. $\dim\left[N(\mathbf{A}^T)\right]=m-r, \quad N(\mathbf{A}^T)\subseteq\mathbb{R}^m$

Άλλο παράδειγμα: Να βρεθούν βάσεις των θεμελιωδών υποχώρων του

$$A=\begin{bmatrix} 6 & 0 & 12 & 6 & 6 \\ 2 & 0 & 4 & 5 & 2 \\ -1 & -1 & -2 & -3 & -3 \\ 3 & 1 & 6 & 11 & 5 \end{bmatrix},$$ όπου $PA = LU$ με

$$U=\begin{bmatrix} 6 & 0 & 12 & 6 & 6 \\ 0 & 1 & 0 & 8 & 2 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad P=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad L=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ -1/6 & -1 & 1 & 0 \\ 1/3 & 0 & 1/2 & 1 \end{pmatrix}$$

βάση $R(A) \to$ 1η, 2η & 4η στήλη A: $\left\{ \begin{pmatrix} 6 \\ 2 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 6 \\ 5 \\ -3 \\ 11 \end{pmatrix} \right\}$

" $R(A^T) \to$ 3 πρώτες γραμ. U: $\left\{ \begin{pmatrix} 6 \\ 0 \\ 12 \\ 6 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 8 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 6 \\ 0 \end{pmatrix} \right\}$

" βάση $N(A)$: έχουμε βρει σ.6 του '09.τεύχ': $\left\{ \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

βάση $N(A^T)$: τελευταία γραμμή του $L^{-1}P$

πρώτα βρες L^{-1}

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1/6 & -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1/3 & 0 & 1/2 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/2 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1/6 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 1 & -1/3 & 0 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1/3 & 1 & 1 & 0 \\ 0 & 0 & 1/2 & 1 & -1/3 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1/3 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1/6 & -1/2 & -1/2 & 1 \end{array}\right]$$

$$L^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1/2 & 1 & 0 & 0 \\ -1/3 & 1 & 1 & 0 \\ -1/6 & -1/2 & -1/2 & 1 \end{pmatrix} \qquad L^{-1} \cdot P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1/2 & 0 & 0 & 1 \\ -1/3 & 0 & 1 & 1 \\ -1/6 & 1 & -1/2 & -1/2 \end{pmatrix}$$

$$\Rightarrow \text{bim } N(A^T) = \left\{ \begin{pmatrix} -1/6 \\ 1 \\ -1/2 \\ -1/2 \end{pmatrix} \right\}$$

Άσκηση για το σπίτι: παρομοιως για τον $\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 & 1 & 1 & 2 \\ 2 & -2 & 4 & 2 & 2 & 4 \\ 3 & -3 & 6 & 2 & 3 & 5 \end{bmatrix}$

12. Γραμμικοί Μετασχηματισμοί

(Πηγή: Strang, Κεφάλαιο 2.6)

Βλέπε και
https://www.khanacademy.org/math/linear-algebra/matrix-transformations#lin-trans-examples

Έστω δύο σύνολα X και Y.

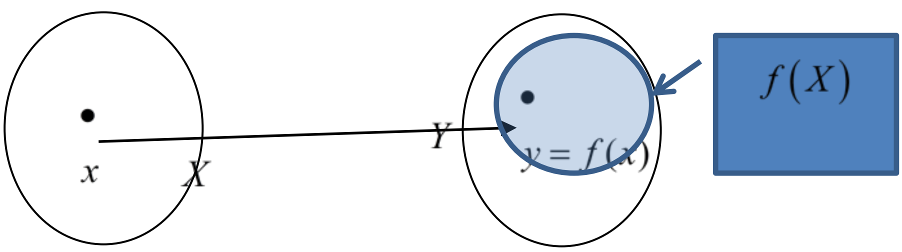
Μια **συνάρτηση ή απεικόνιση ή μετασχηματισμός** f του X στο Y (Γράφουμε $f : X \to Y$) είναι **μια αντιστοίχηση ενός στοιχείου** $y \in Y$ **για κάθε στοιχείο** $x \in X$.

(Σε κάθε $x \in X$ αντιστοιχούμε μόνο ένα $y \in Y$).

Γράφουμε $f : x \mapsto y$ ή $y = f(x)$.

- Το y ονομάζεται **εικόνα** του x.
- Το X ονομάζεται **πεδίο ορισμού** της f και το Y **πεδίο τιμών** της f.
- Τα $y \in Y$ για τα οποία υπάρχει $x \in X$ με $y = f(x)$ συμβολίζονται με $f(X)$ και ονομάζονται **εικόνα του** X **ή και εικόνα της** f.

Καμμια φορά ζωγραφίζουμε την αντιστοίχηση με βελάκια.



($f(X)$ είναι τα $y \in Y$ στα οποία καταλήγει ένα τουλάχιστον βελάκι).

Ο μετασχηματισμός $x \to Ax$

Έστω $\mathbf{A}$ ένας $m \times n$ πίνακας. Τότε **μπορώ να ορίσω ένα μετασχηματισμό** $f_{\mathbf{A}}$.

$$f_{\mathbf{A}} : R^n \to R^m$$

ως εξής: στο $\mathbf{x} \in R^n$ αντιστοιχώ το $\mathbf{y} = \mathbf{Ax} \in R^m$

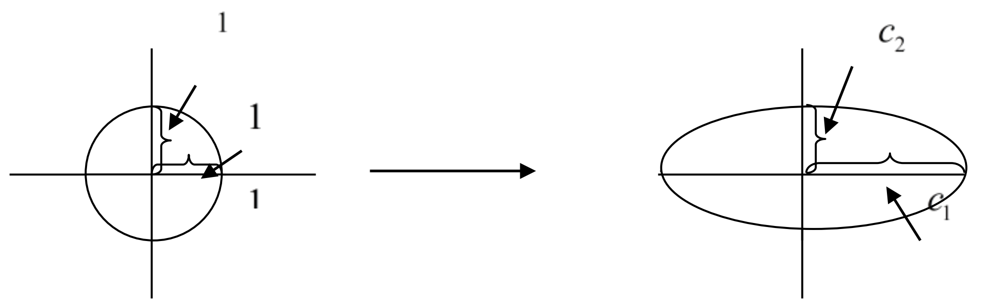
$$f_{\mathbf{A}} : \mathbf{x} \mapsto \mathbf{y} = f_{\mathbf{A}}(\mathbf{x}) := \mathbf{Ax}$$

<u>Δηλ.:</u> η εικόνα του x είναι το γινόμενό του με τον A.

Παραδείγματα:

$$\mathbf{A} = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} \qquad \mathbf{A}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} c_1 x_1 \\ c_2 x_2 \end{pmatrix}$$

Διαστέλλει κατά c_1 στην κατεύθυνση του πρώτου άξονα και κατά c_2 στου δεύτερου.



https://www.geogebra.org/m/E74ww6gn

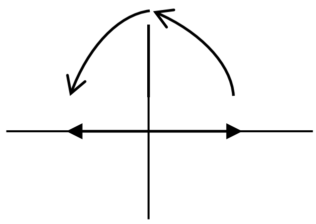
https://www.geogebra.org/m/TSQk4yCY

https://www.geogebra.org/m/VKQmRQg6

Περιστρέφει τον R^2 κατά 90°

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

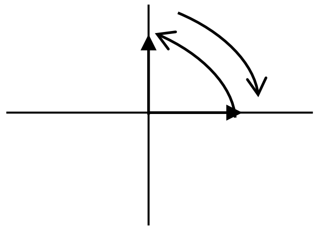
$$\mathbf{A}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix} \text{ και } \mathbf{A}\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}-1\\0\end{pmatrix}$$



Αντανάκλαση στην ευθεία 45°

$$\mathbf{A}=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$ τότε:

$$\mathbf{A}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix} \text{ και } \mathbf{A}\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}$$



$$\mathbf{A}=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} \qquad \mathbf{A}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}x_1\\0\end{pmatrix}$$

Προβολή στον άξονα των x_1

https://www.geogebra.org/m/E74ww6gn

https://www.geogebra.org/m/VKQmRQg6

Μια απεικόνιση $f: X \to Y$ λέγεται «επί»

αν για κάθε $y \in Y$ υπάρχει $x \in X$ τέτοιο που $f(x) = y$, αν δηλαδή η εικόνα της f, $f(X)$ ταυτίζεται με το Y. (Δηλ.: **Καταλήγει βελάκι σε κάθε στοιχείο του** Y).

Πότε είναι επί η $f_{\mathbf{A}}$?

Η $f_{\mathbf{A}}: R^n \to R^m$ είναι «επί»

$\Leftrightarrow$ το $\mathbf{Ax} = \mathbf{y}$ έχει λύση για κάθε $\mathbf{y} \in R^m$

$\Leftrightarrow$ ο $R(\mathbf{A}) = R^m \Leftrightarrow \dim\left(R(\mathbf{A})\right) = \dim\left(R^m\right)$

$\Leftrightarrow r = m$ όπου r η τάξη του πίνακα

Άρα: η $f_{\mathbf{A}}$ **είναι «επί»** $\Leftrightarrow r = m$

Η παραπάνω επιχειρηματολογία δείχνει και ότι έχουμε:

$f_{\mathbf{A}}(X) = R(\mathbf{A})$**, δηλαδή ο** $R(\mathbf{A})$ **είναι η εικόνα της** $f_{\mathbf{A}}$**.**

https://www.geogebra.org/m/ebJbXhXu

Μια απεικόνιση $f: X \to Y$ λέγεται «μονοσήμαντη»

αν **κάθε** $x \in X$ **απεικονίζεται σε διαφορετικό** $y \in Y$, δηλαδή¨

αν x και $x' \in X$ έχουν την ίδια εικόνα, πρέπει να ταυτίζονται:

$$\boxed{f(x) = f(x') \Rightarrow x = x'}$$

(Δηλ.: **Αν σε κάποιο** $y \in Y$ **καταλήγει βελάκι, θα είναι μόνο ένα**).

Πότε είναι μονοσήμαντη η $f_{\mathbf{A}}$?

$x \mapsto Ax$ (από $\mathbf{Ax} = \mathbf{Ax'} \Rightarrow \mathbf{x} = \mathbf{x'}$) $\Leftrightarrow$

(από $\mathbf{A(x - x') = 0} \Rightarrow \mathbf{x - x' = 0}$) $\Leftrightarrow$

δεν υπάρχουν ελεύθερες μεταβλητές $\Leftrightarrow$

$N(\mathbf{A}) = \{0\} \Leftrightarrow \dim(N(\mathbf{A})) = 0 \Leftrightarrow$

$r = n$

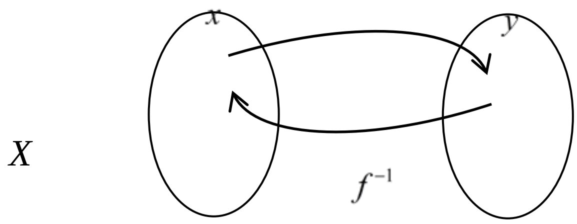
$$\boxed{\text{Άρα: η } f_A \text{ είναι «μονοσήμαντη»} \Leftrightarrow r = n}$$

Μια απεικόνιση $f : X \to Y$ λέγεται «αμφιμονοσήμαντη»

αν είναι «επί» και «μονοσήμαντη».

Τότε μπορώ να **ορίσω την αντίστροφη συνάρτηση** $f^{-1} : Y \to X$**,** ως εξής:

αν $y = f(x)$ τότε $f^{-1}(y) = x$ ή και $f^{-1}\big(f(x)\big) = x$



Πότε είναι αμφι-μονοσήμαντη και επομένως αντιστρέψιμη η $f_{\mathbf{A}}$?

Για να αντιστρέφεται η $f_{\mathbf{A}}$ χρειάζομαι «επί» ($r = m$) και «μονοσήμαντη» ($r = n$).

Άρα $f_{\mathbf{A}}$ αντιστρέψιμη όταν $r = m = n$ και επομένως **ο** $\mathbf{A}$ **είναι αναγκαστικά μη – ιδιόμορφος**.

<u>Ερώτηση:</u> Υπάρχει $B_{n \times n}$ τέτοιος που «να δίνει» την f_A^{-1};

<u>Δηλαδή:</u> $f_A^{-1}(y) = f_B y = B \cdot y$;

Χρειάζομαι $B: B\cdot Ax = x$.

NAI : $B = A^{-1}$

Η αντίστροφη συνάρτηση f_A^{-1} **δίνεται από τον αντίστροφο πίνακα** $f_{A^{-1}}$

$$f_A^{-1}(y) = f_{A^{-1}}y := A^{-1}\cdot y$$

και όντως αν $y = Ax$ τότε $A^{-1}\cdot y = A^{-1}\cdot Ax = x$

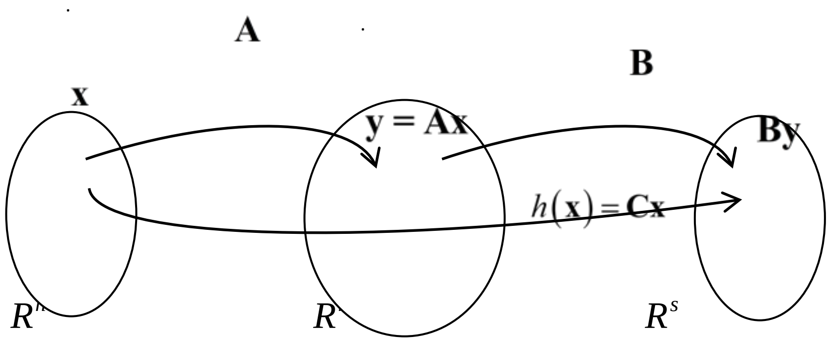
Σύνθεση συναρτήσεων

Έστω $f_{\mathbf{A}} : R^n \to R^m$ και $f_{\mathbf{B}} : R^m \to R^s$

$x \mapsto Ax$ $\qquad\qquad$ $y \mapsto By$

με $\mathbf{A}_{\mathbf{m\times n}}$ και $\mathbf{B}_{\mathbf{s\times m}}$

Ορίζουμε την σύνθεση των συναρτήσεων f_A και f_B ως μια καινούρια συνάρτηση $h = f_B \circ f_A$ με $h : R^n \to R^s$ και $h(x) := f_B\left(f_A(x)\right)$.



Ποιος πίνακας «δίνει» την h;

Δηλαδή: Υπάρχει $\mathbf{C}$ τέτοιος που $h(\mathbf{x}) = f_{\mathbf{C}}(\mathbf{x}) = \mathbf{C}\cdot\mathbf{x}$

ΝΑΙ: $h(\mathbf{x}) = f_{\mathbf{B}}\left(f_{\mathbf{A}}(\mathbf{x})\right) = f_{\mathbf{B}}(\mathbf{A}\cdot\mathbf{x}) = \mathbf{B}\cdot(\mathbf{A}\cdot\mathbf{x})$

Άρα $\mathbf{C} = \mathbf{B}\cdot\mathbf{A}$

Η σύνθεση συναρτήσεων δίνεται από τον πολλαπλασιασμό πινάκων.

Γραμμικοί Μετασχηματισμοί

<u>Ορισμός:</u> **Ένας μετασχηματισμός** f μεταξύ δύο διανυσματικών χώρων V, W $f:V \to W$ **λέγεται γραμμικός αν**

$$f(0)=0$$

$$f(x+y)=f(x)+f(y) \quad \forall x,y \in V$$

$$f(\lambda x)=\lambda f(x) \quad \forall x \in V, \ \lambda \in R$$

<u>Παρατήρηση</u>: Για ένα $m \times n$ πίνακα $\mathbf{A}$ **η απεικόνιση** $f_{\mathbf{A}}$ **είναι γραμμική**:

$$A \cdot 0 = 0$$

$$A(x+y)=Ax+Ay$$

και $A(\lambda x)=\lambda Ax$

Βοηθητική Πρόταση.

Αν γνωρίζω εικόνες της βάσης τότε γνωρίζω όλο το μετασχηματισμό:

‘Εστω f κάποιος γραμμικός μετασχηματισμός μεταξύ δύο διανυσματικών χώρων V, W: $f:V \to W$ και $v_1,\ldots,v_k$ μία βάση του V. **Αν γνωρίζω τις εικόνες** $f(v_1),\ldots,f(v_k)$ **τότε μπορώ να βρώ την εικόνα οποιουδήποτε** $v \in V$**, και επομένως γνωρίζω όλο τον f.**

https://www.geogebra.org/m/TSQk4yCY

https://www.geogebra.org/m/kGJZH7AU

Απόδειξη:

‘Εστω $v \in V$.

Τότε θα υπάρχουν $\lambda_1,\ldots,\lambda_k$ τέτοια που $v = \lambda_1 v_1 + \ldots + \lambda_k v_k$.

Τότε όμως, λόγω της γραμμικότητας της f, παίρνουμε:

$$f(v) = f(\lambda_1 v_1 + \ldots + \lambda_k v_k) = \lambda_1 f(v_1) + \ldots + \lambda_k f(v_k)$$

με γνωστά $\lambda_1,\ldots,\lambda_k$ και $f(v_1),\ldots,f(v_k)$.

Πρόταση : **Κάθε γραμμικό μετασχηματισμό «τον δίνει» ένας πίνακας**

Έστω **οποιοσδήποτε** γραμμικός μετασχηματισμός $f : R^n \to R^m$. Τότε θα υπάρχει $\mathbf{A}_{\mathbf{m}\times\mathbf{n}}$ τέτοιος που $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$, δηλαδή $f \equiv f_{\mathbf{A}}$.

Αν γνωρίζουμε τις εικόνες $\mathbf{y}_{\mathbf{i}} = f(\mathbf{e}_{\mathbf{i}})$ των στοιχείων της κανονικής βάσης $e_i = (0,\ldots,0,1,0,\ldots,0)^T$ (με το 1 στην i-συντεταγμένη) τότε ο πίνακας ***A*** κατασκευάζεται ως $\mathbf{A} = \left[y_1 | y_2 | \ldots | y_n \right]$.

Απόδειξη: Για τα στοιχεία της βάσης $e_1,\ldots,e_n$ όντως έχουμεQ

$$f_{\mathbf{A}}(\mathbf{e}_{\mathbf{i}}) = \mathbf{A}\mathbf{e}_{\mathbf{i}} = \mathbf{y}_{\mathbf{i}} = f(\mathbf{e}_{\mathbf{i}}).$$

Και αφού οι δύο μετασχηματισμοί συμπίπτουν για τα στοιχεία της βάσης, σύμφωνα με τη προγηγούμενη πρόταση συμπίπτουν και σε όλο το V.

https://www.geogebra.org/m/kGJZH7AU

Αν ο f δίνεται με τις εικόνες άλλης βάσης πώς κατασκευάζω τον A;

Πρόταση. Αν για κάποιο γραμμικό μετασχηματισμός $f: R^n \to R^m$ γνωρίζουμε τις εικόνες $\mathbf{y_i} = f(\mathbf{v_i})$ των στοιχείων κάποιας βάσης $\mathbf{v_1,...,v_n}$ του R^n τότε ο πίνακας $\mathbf{A}$ τέτοιος που $f(\mathbf{x}) = \mathbf{Ax}$ κατασκευάζεται ως $\mathbf{A} = \mathbf{YV}^{-1}$, όπου $\mathbf{Y} = \left[\mathbf{y_1}\middle|\mathbf{y}_2\middle|\ldots\middle|\mathbf{y_n}\right]$ και $\mathbf{V} = \left[\mathbf{v_1}\middle|\mathbf{v}_2\middle|\ldots\middle|\mathbf{v_n}\right]$.

Απόδειξη

Θέλω $\mathbf{A}$ *τέτοιο που* $\mathbf{Av_i} = \mathbf{y_i}$.

Αυτό ισοδυναμεί με (μνημονικός κανόνας)

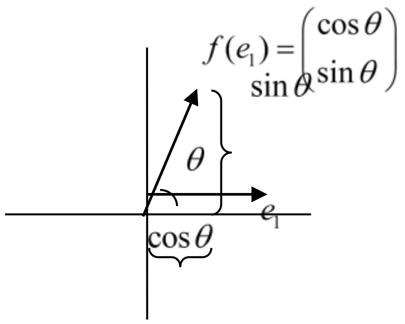
$$\mathbf{AV} = \mathbf{Y},$$

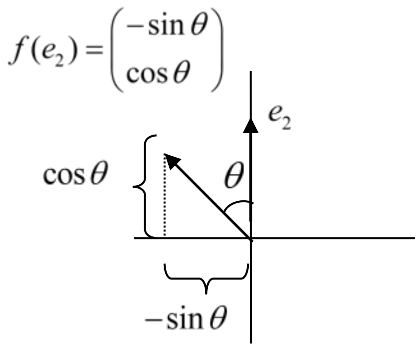
που συνεπάγεται

$$\mathbf{A} = \mathbf{YV}^{-1}.$$

π.χ.

Η απεικόνιση f **«στροφή κατά γωνία θ»**





άρα $A=\left[\, f(e_1) \middle| f(e_2)\,\right]=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$

και **η απεικόνιση προβολή στην** L_a $\;P_{\mathbf{a}}(\mathbf{x})=\dfrac{\langle \mathbf{a},\mathbf{x}\rangle}{\langle \mathbf{a},\mathbf{a}\rangle}\cdot\mathbf{a}$

$$f(\mathbf{e}_1)=\frac{\langle \mathbf{a},\mathbf{e}_1\rangle}{\langle \mathbf{a},\mathbf{a}\rangle}\cdot\mathbf{a}=\begin{pmatrix}\dfrac{a_1^{\,2}}{a_1^{\,2}+a_2^{\,2}}\\[2ex] \dfrac{a_1a_2}{a_1^{\,2}+a_2^{\,2}}\end{pmatrix} \quad \& \quad f(\mathbf{e}_2)=\frac{\langle \mathbf{a},\mathbf{e}_2\rangle}{\langle \mathbf{a},\mathbf{a}\rangle}\cdot\mathbf{a}=\begin{pmatrix}\dfrac{a_1a_2}{a_1^{\,2}+a_2^{\,2}}\\[2ex] \dfrac{a_2^{\,2}}{a_1^{\,2}+a_2^{\,2}}\end{pmatrix}$$

άρα $A=\dfrac{1}{a_1^{\,2}+a_2^{\,2}}\begin{bmatrix}a_1^{\,2} & a_1a_2\\ a_1a_2 & a_2^{\,2}\end{bmatrix}$

Συγκρίνετε με $P_{\mathbf{a}}(x)=\dfrac{\langle \mathbf{a},\mathbf{x}\rangle}{\langle \mathbf{a},\mathbf{a}\rangle}\cdot\mathbf{a}=\dfrac{\mathbf{a}^{\mathrm{T}}\mathbf{x}}{\mathbf{a}^{\mathrm{T}}\mathbf{a}}\cdot\mathbf{a}=\dfrac{\mathbf{a}\mathbf{a}^{\mathrm{T}}}{\mathbf{a}^{\mathrm{T}}\mathbf{a}}\cdot\mathbf{x}$

Στην **αντανάκλαση** $H_{\mathbf{a}}(\mathbf{x})$

έχουμε

$$H_{\mathbf{a}}(\mathbf{x}) = x + 2\left(P_{\mathbf{a}}(\mathbf{x}) - \mathbf{x}\right) = 2P_{\mathbf{a}}(\mathbf{x}) - \mathbf{x} = \left[2\frac{\mathbf{a}\mathbf{a}^{\mathbf{T}}}{\mathbf{a}^{\mathbf{T}}\mathbf{a}} - I\right]\mathbf{x}$$

άρα $$\mathbf{A} = \frac{1}{a_1^{\,2} + a_2^{\,2}}\begin{bmatrix} 2a_1^{\,2} - 1 & 2a_1a_2 \\ 2a_1a_2 & 2a_2^{\,2} - 1 \end{bmatrix}$$

5.52. Find a 2×2 matrix A that maps

(a) $(1,3)^T$ and $(1,4)^T$ into $(-2,5)^T$ and $(3,-1)^T$, respectively.

(b) $(2,-4)^T$ and $(-1,2)^T$ into $(1,1)^T$ and $(1,3)^T$, respectively.

5.53. Find a 2×2 singular matrix B that maps $(1,1)^T$ into $(1,3)^T$.

5.62. For each linear map G, find a basis and the dimension of the kernel and the image of G:

(a) $G: \mathbf{R}^3 \to \mathbf{R}^2$ defined by $G(x,y,z) = (x+y+z,\ 2x+2y+2z)$,

(b) $G: \mathbf{R}^3 \to \mathbf{R}^2$ defined by $G(x,y,z) = (x+y,\ y+z)$,

(c) $G: \mathbf{R}^5 \to \mathbf{R}^3$ defined by

$$G(x,y,z,s,t) = (x+2y+2z+s+t, \quad x+2y+3z+2s-t, \quad 3x+6y+8z+5s-t).$$

5.63. Each of the following matrices determines a linear map from $\mathbf{R}^4$ into $\mathbf{R}^3$:

$$\text{(a)} \quad A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 2 & -1 & 2 & -1 \\ 1 & -3 & 2 & -2 \end{bmatrix}, \text{(b)} \quad B = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 2 & 3 & -1 & 1 \\ -2 & 0 & -5 & 3 \end{bmatrix}.$$

Find a basis as well as the dimension of the kernel and the image of each linear map.

5.64. Find a linear mapping $F: \mathbf{R}^3 \to \mathbf{R}^3$ whose image is spanned by $(1,2,3)$ and $(4,5,6)$.

5.65. Find a linear mapping $G: \mathbf{R}^4 \to \mathbf{R}^3$ whose kernel is spanned by $(1,2,3,4)$ and $(0,1,1,1)$.

13. Ορθογώνιοι υπόχωροι

(Πηγή: Strang, Κεφάλαιο 3.1)

Υπενθύμιση: για $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ λέμε $\mathbf{x} \perp \mathbf{y} \Leftrightarrow \langle \mathbf{x},\mathbf{y} \rangle = 0$.

Πρόταση. Έστω $\mathbf{v_1},\ldots,\mathbf{v_k} \in \mathbb{R}^n$ μη-μηδενικά με $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ για $\mathrm{i} \neq \mathrm{j}$. Τότε τα $\{\mathbf{v_1},\ldots,\mathbf{v_k}\}$ **θα είναι ανεξάρτητα**.

Απόδειξη.

Έστω $\lambda_1\mathbf{v_1} + \ldots + \lambda_k\mathbf{v_k} = \mathbf{0}$.

Πολ/ζω από αριστερά με $\mathbf{v}_j^T$ για κάποιο j. Παίρνουμε

$$\lambda_1\mathbf{v}_j^T\mathbf{v_1} + \ldots + \lambda_k\mathbf{v}_j^T\mathbf{v_k} = \mathbf{0}$$

Αριστερά μόνο $\lambda_j\mathbf{v}_j^T\mathbf{v}_j$ είναι διάφορο του μηδενός εξ αιτίας της καθετότητας των $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ για $\mathrm{i} \neq \mathrm{j}$. Έτσι παίρνουμε:

$$\lambda_j\mathbf{v}_j^T\mathbf{v}_j = \lambda_j\left\|\mathbf{v}_j\right\|^2 = \mathbf{0}$$

Και καθως $\left\|\mathbf{v}_j\right\|^2 \neq 0$ πρέπει $\lambda_j = 0$. Παρομοίως παίρνουμε

$$\lambda_1 = \ldots = \lambda_k = 0.$$

Θέλουμε να ορίσουμε πότε **δύο υπόχωροι θα είναι ο οθογώνιοι** ο ένας στον άλλον.

Απάντηση. Όταν κάθε στοιχείο του ενός είναι ορθογώνιο σε κάθε στοιχείο του άλλου.

Ορισμός. Έστω $V, W \subseteq \mathbb{R}^n$ υπόχωροι. Θα λέμε V ορθογώνιος (ή κάθετος) στον W και συμβολίζουμε $V \perp W$ τότε και μόνο τότε όταν:

$$\mathbf{v} \perp \mathbf{w}, \quad \forall \mathbf{v} \in V, \mathbf{w} \in W.$$

Παραδείγματα

- Στον $\mathbb{R}^2$ δύο ευθείες που περνάνε από το 0 κάθετες μεταξύ τους.
- Στον $\mathbb{R}^3$ ένα επίπεδο και μια ευθεία που περνάνε από το 0 κάθετα μεταξύ τους.
- Στον $\mathbb{R}^3$ δύο επίπεδα **δεν** μπορούν να είναι ορθογώνια καθως θα έχουν ως τομή μία ευθεία και για να είναι ορθογώνια θε έπρεπε όλα τα στοιχεία της ευθείας να είναι κάθετα στον εαυτό τους.

Πρόταση. Έστω $V,W \subseteq \mathbb{R}^n$ υπόχωροι και $\{\mathbf{v_1},...,\mathbf{v_k}\}$ παράγει V και $\{\mathbf{w_1},...,\mathbf{w_r}\}$ παράγει W. Τότε:

- αν $\mathbf{v}_i \perp \mathbf{w}_j,\ \forall i, j$ **θα έπεται** $V \perp W$
- αν για κάθε $\mathbf{v} \in V$ έχουμε $\mathbf{v} \perp \mathbf{w}_j,\ \forall j$ **θα έπεται** $V \perp W$

Απόδειξη.

Αν $\mathbf{v} \in V, \mathbf{w} \in W$ θα έχουμε

$$\mathbf{v} = \lambda_1 \mathbf{v_1} + ... + \lambda_k \mathbf{v_k} \text{ και } \mathbf{w} = \mu_1 \mathbf{w_1} + ... + \mu_r \mathbf{w_r}.$$

και $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{k} \sum_{j=1}^{r} \lambda_i \mu_j \langle \mathbf{v}_i, \mathbf{w}_j \rangle = 0$ καθώς $\langle \mathbf{v}_i, \mathbf{w}_j \rangle = 0$.

Παρόμοια αν $\mathbf{v} \perp \mathbf{w}_j,\ \forall j$ **για κάθε** $\mathbf{v} \in V$.

Παραδείγμα

Στον $\mathbb{R}^4$: Έστω $P = span\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}, L = span\left\{ \begin{pmatrix} 0 \\ 0 \\ 4 \\ 5 \end{pmatrix} \right\}, L' = span\left\{ \begin{pmatrix} 0 \\ 0 \\ -5 \\ 4 \end{pmatrix} \right\}$.

Εύκολα ελέγχουμε

- $P \perp L$,
- $P \perp L'$,
- $L \perp L'$.

Πρόταση. Έστω $\mathbf{A}_{m\times n}$. Τότε:

α) χώρος γραμμών $\mathbf{A}$ $\perp$ Μηδενόχωρο $\mathbf{A}$ $\subseteq \mathbb{R}^n$
α) χώρος στηλών $\mathbf{A}$ $\perp$ Αριστερό μηδενόχωρο $\mathbf{A}$ $\subseteq \mathbb{R}^m$

Απόδειξη.

α) Έστω $\mathbf{x} \in N(\mathbf{A})$. Τότε $\mathbf{Ax} = \mathbf{0}$.

Που ισοδυναμεί με $\left\langle \mathbf{a}^{\mathbf{i}}, \mathbf{x} \right\rangle = 0, i = 1,\ldots,m$.

Και άρα $R\left(\mathbf{A}^{\mathrm{T}}\right) \perp N(\mathbf{A})$.

α) Έστω $\mathbf{y} \in N\left(\mathbf{A}^{T}\right)$. Τότε $\mathbf{y}^{\mathbf{T}}\mathbf{A} = \mathbf{0}^{T}$.

Που ισοδυναμεί με $\left\langle \mathbf{a}_{\mathbf{i}}, \mathbf{y} \right\rangle = 0, i = 1,\ldots,n$.

Και άρα $R(\mathbf{A}) \perp N\left(\mathbf{A}^{T}\right)$.

Ερώτημα: Έστω $V, W \subseteq \mathbb{R}^n$ υπόχωροι με $V \perp W$.

- Μήπως εκτός από τα στοιχεία του W υπάρχουν και άλλα στοιχεία του $\mathbb{R}^n$ που να ειναι ορθογώνια στον V, ή
- μήπως, αντίθετα, **ο** W **περιέχει όλα τα** στοιχεία του $\mathbb{R}^n$ που είναι **ορθογώνια στον** V **;**

Π.χ. στο παράδειγμά μας έχουμε $P \perp L$, αλλά ο L δεν περιέχει όλα τα στοιχεία του $\mathbb{R}^n$ που είναι ορθογώνια στον P: υπάρχουν και τα στοιχεί της L' που είναι ορθογώνια στο P!

Ορισμός: Έστω $V \subseteq \mathbb{R}^n$ υπόχωρος. Ορίζουμε **το ορθογώνιο συμπλήρωμα του** V**,** που το συμβολίζουμε με $V^{\perp}$ ως εκείνο τον υπόχωρο του $\mathbb{R}^n$ που περιέχει όλα τα στοιχεία του $\mathbb{R}^n$ που είναι ορθογώνια στον V, δηλ.

$$V^{\perp} := \left\{ \mathbf{x} \in \mathbb{R}^n \middle| \mathbf{x} \perp \mathbf{v}, \forall \mathbf{v} \in V \right\}$$

Ερώτημα: Έστω $V, W \subseteq \mathbb{R}^n$ υπόχωροι με $V \perp W$. Πως μπορώ να ξέρω **αν είναι ορθογώνια συμπληρώματα** ο ένας του άλλου?

Απάντηση: Η επόμενη πρόταση λέει ότι **είναι ορθογώνια συμπληρώμτα** **αν οι διαστάσεις τους αθροίζονται σε** n.

Πρόταση. Έστω $V, W \subseteq \mathbb{R}^n$ υπόχωροι με $V \perp W$ και με

$$\dim(V) + \dim(W) = n.$$

Τότε:

α) αν $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ βάση του V και $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ βάση του W, τότε θα έπεται πως $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ βάση του $\mathbb{R}^n$.

β) Για κάθε $\mathbf{x} \in \mathbb{R}^n$ υπάρχουν $\mathbf{v} \in V$ και $\mathbf{w} \in W$ τέτοια ώστε $\mathbf{x} = \mathbf{v} + \mathbf{w}$.

γ) $W = V^{\perp}$ και $V = W^{\perp}$.

Σημείωση. Θα δούμε ότι $\mathbf{v} = P_V(\mathbf{x})$ η προβολή του $\mathbf{x}$ στο V, ενώ $\mathbf{w} = \mathbf{x} - \mathbf{v} = P_W(\mathbf{x})$.

Απόδειξη.

α) Καθώς το πλήθος των $\{\mathbf{v_1},\ldots,\mathbf{v_n}\}$ συμπίπτει με τη διάσταση του $\mathbb{R}^n$ αρκεί να δείξουμε ότι είναι ανεξάρτητα. Έστω

$$\lambda_1\mathbf{v_1}+\ldots+\lambda_r\mathbf{v}_r+\lambda_{r+1}\mathbf{v}_{r+1}+\ldots+\lambda_n\mathbf{v}_n=\mathbf{0}$$

Θέτοντας $\mathbf{v}:=\lambda_1\mathbf{v_1}+\ldots+\lambda_r\mathbf{v}_r\in V$ και $\mathbf{w}:=\lambda_{r+1}\mathbf{v}_{r+1}+\ldots+\lambda_n\mathbf{v}_n\in W$ έχουμε:

$$\mathbf{v}+\mathbf{w}=\mathbf{0}\Rightarrow\|\mathbf{v}+\mathbf{w}\|^2=0\Rightarrow\|\mathbf{v}\|^2+\|\mathbf{w}\|^2+2\langle\mathbf{v},\mathbf{w}\rangle=0$$

Αλλά καθως $\mathbf{v}\in V$ και $\mathbf{w}\in W$ έπεται $\langle\mathbf{v},\mathbf{w}\rangle=0$ και με αυτό από τη παραπάνω ισότητα παίρνουμε:

$$\|\mathbf{v}\|^2+\|\mathbf{w}\|^2=0\Rightarrow\|\mathbf{v}\|^2=0 \text{ και } \|\mathbf{w}\|^2=0\Rightarrow\mathbf{v}=\mathbf{0} \text{ και } \mathbf{w}=\mathbf{0}$$

Αλλά από $\mathbf{v}=\lambda_1\mathbf{v_1}+\ldots+\lambda_r\mathbf{v}_r=\mathbf{0}$ με την ανεξάρτησία των $\{\mathbf{v_1},\ldots,\mathbf{v}_r\}$ παίρνουμε $\lambda_1=\ldots=\lambda_r=0$ και παρομοίως $\lambda_{r+1}=\ldots=\lambda_n=0$ από $\mathbf{w}=\mathbf{0}$.

β) Αφού $\{\mathbf{v_1},\ldots,\mathbf{v_n}\}$ είναι βάση του $\mathbb{R}^n$ οποιοδήποτε $\mathbf{x}\in\mathbb{R}^n$ γράφεται ως

$$\mathbf{x}=\mu_1\mathbf{v_1}+\ldots+\mu_r\mathbf{v}_r+\mu_{r+1}\mathbf{v}_{r+1}+\ldots+\mu_n\mathbf{v}_n.$$

Θέτουμε $\mathbf{v}:=\mu_1\mathbf{v_1}+\ldots+\mu_r\mathbf{v}_r\in V$ και $\mathbf{w}:=\mu_{r+1}\mathbf{v}_{r+1}+\ldots+\mu_n\mathbf{v}_n\in W$.

Τότε προφανώς $\mathbf{x}=\mathbf{v}+\mathbf{w}$.

γ) Έστω $\mathbf{x} \in \mathbb{R}^n$ με $\mathbf{x} \perp V$. Θα δείξουμε $\mathbf{x} \in W$.

Από β) γράφουμε το $\mathbf{x}$ ως $\mathbf{x} = \mathbf{v} + \mathbf{w}$ με $\mathbf{v} \in V$ και $\mathbf{w} \in W$.

Από $\mathbf{x} \perp V$ έπεται όμως $\mathbf{v} = \mathbf{0}$.

Διότι $0 = \langle \mathbf{x}, \mathbf{v} \rangle = \langle \mathbf{v} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle = \|\mathbf{v}\|^2 + 0$, καθως $\langle \mathbf{w}, \mathbf{v} \rangle = 0$ και άρα $\mathbf{v} = \mathbf{0}$.

Άρα $\mathbf{x} = \mathbf{w} \in W$.

Θεμελιώδες θεώρημα της Γραμμικής Άλγεβρας, Μέρος ΙΙ.

Έστω $\mathbf{A}_{m\times n}$. Τότε οι θεμελιώδεις υπόχωροι του $\mathbf{A}$ είναι ανα ζεύγη ορθογώνια συμπληρώματα ο ένας του άλλου:

α) χώρος γραμμών $\mathbf{A}$, $R(\mathbf{A}^T) = \mathrm{N}(\mathbf{A})^{\perp} \subseteq \mathbb{R}^n$

α) χώρος στηλών $\mathbf{A}$, $R(\mathbf{A}) = \mathrm{N}(\mathbf{A}^{\mathrm{T}})^{\perp} \subseteq \mathbb{R}^m$

Απόδειξη. Οι διαστάσεις αθροίζονται:

$$\dim\left(R\left(\mathbf{A}^T\right)\right) + \dim\left(\mathrm{N}\left(\mathbf{A}\right)\right) = n$$

$$\dim\left(R\left(\mathbf{A}\right)\right) + \dim\left(\mathrm{N}\left(\mathbf{A}^{\mathrm{T}}\right)\right) = m$$

14. Προβολές και προσεγγίσεις ελαχίστων τετραγώνων, Μέρος Α.

(Πηγή: Strang, Κεφάλαιο 3.3)

Έστω ένα σύστημα με «πολλές» εξισώσεις και λίγους αγνώστους:

$$\mathbf{Ax} = \mathbf{b} \text{ με } \mathbf{A}_{m\times n} \text{ και } m \gg n.$$

Γνωρίζουμε ότι ένα σύστημα έχει λύση αν $\mathbf{b} \in R(\mathbf{A})$.

Μπορώ να κάνω «κάτι» αν $\mathbf{b} \notin R(\mathbf{A})$;

Το πρόβλημα για n=1

Έτσι π.χ. αν $n = 1$, δηλ. ο $\mathbf{A}$ είναι μια στήλη $\mathbf{a}$ και $x \in R$ το σύστημα έχει λύση αν $\mathbf{b} \in Span(\mathbf{a})$.

Μπορώ να κάνω «κάτι» αν $\mathbf{b} \notin Span(\mathbf{a})$;

$$\mathbf{A} = \mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$$

Ιδέα: Αφού $\mathbf{Ax} - \mathbf{b} = \mathbf{0}$ αδύνατο, **βρες τουλάχιστον $\overline{\mathbf{x}}$ τέτοιο ώστε το $\mathbf{A}\overline{\mathbf{x}} - \mathbf{b}$ να είναι όσο πιο «κοντά στο 0» γίνεται, δηλαδή να ελαχιστοποιείται το $\| \mathbf{A}\overline{\mathbf{x}} - \mathbf{b} \|$.**

Βρες $\overline{x} : \| \mathbf{A}\overline{x} - \mathbf{b} \| \le \| \mathbf{A}x - \mathbf{b} \| \ \forall x \in R$

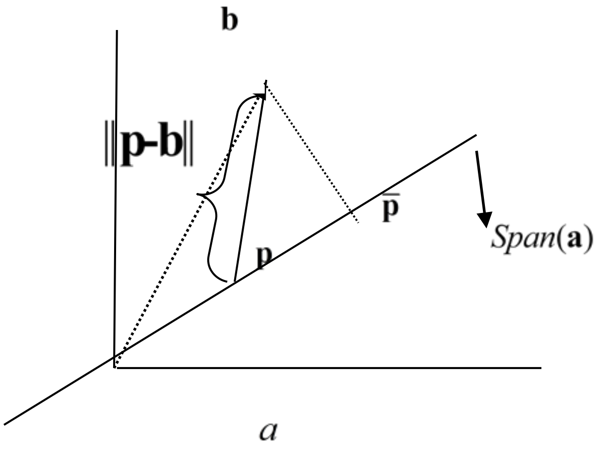
Βρες $\overline{x} : \| \mathbf{A}\overline{x} - \mathbf{b} \| \le \| \mathbf{A}x - \mathbf{b} \| \ \forall x \in R$

Αν θέσω $p = Ax$ τότε $p \in Span(a)$ και αν $\overline{p} = A\overline{x}$ τότε και $\overline{p} \in Span(a)$, το πρόβλημά μου, λοιπόν, μπορεί ισοδύναμα να γραφτεί:

$$\boxed{\text{Βρες } \overline{\mathbf{p}} \in Span(\mathbf{a}) : \| \overline{\mathbf{p}} - \mathbf{b} \| \leq \| \mathbf{p} - \mathbf{b} \| \ \forall \mathbf{p} \in Span(\mathbf{a})}$$

ή με λόγια: βρες εκείνο το σημείο $\overline{\mathbf{p}}$ της ευθείας $Span(\mathbf{a})$ που **να ελαχιστοποιεί την απόστασή του από το** $\mathbf{b}$ (συγκρινόμενο με την απόσταση του $\mathbf{b}$ με άλλα σημεία $\mathbf{p}$ της $Span(\mathbf{a})$).

Σχηματικά:



Γεωμετρική λύση: $\overline{p}$ **η προβολή του** b **στο** $Span(a)$.

Αλγεβρική λύση: Θέτω $E(x) = \| \mathbf{A}x - \mathbf{b} \|^2 = \sum_{i=1}^{n} (a_i x - b_i)^2$

$$\frac{dE(x)}{dx} = 0 \Leftrightarrow \sum_{i=1}^{n} 2a_i(a_i x - b_i) = 0 \Leftrightarrow x\sum a_i^2 - \sum a_i b_i = 0 \Rightarrow \overline{x} = \frac{\sum a_i b_i}{\sum a_i^2} = \frac{<\mathbf{a},\mathbf{b}>}{<\mathbf{a},\mathbf{a}>}$$

και $\overline{\mathbf{p}} = \mathbf{A}\overline{x} = \dfrac{<\mathbf{a},\mathbf{b}>}{<\mathbf{a},\mathbf{a}>} \cdot \mathbf{a} = P_{\mathbf{a}}\mathbf{x}$

Ίδιο πρόβλημα για n=2

$$\mathbf{A} = [\mathbf{a}_1 \,|\, \mathbf{a}_2], \ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Βρες $\overline{\mathbf{x}} = \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \end{pmatrix} : \| \mathbf{A}\overline{\mathbf{x}} - \mathbf{b} \| \leq \| \mathbf{A}\mathbf{x} - \mathbf{b} \| \ \forall \mathbf{x} \in R^2$

Καθώς

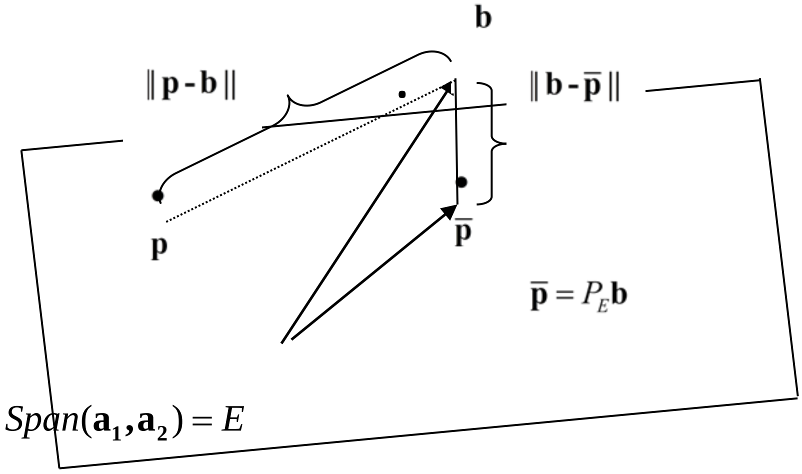
$$\mathbf{A}\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 \in Span(\mathbf{a}_1, \mathbf{a}_2) =: E$$

και θέτοντας $\overline{\mathbf{p}} = \mathbf{A}\overline{\mathbf{x}} \in E$ και $\mathbf{p} = \mathbf{A}\mathbf{x} \in E$ το πρόβλημα γράφεται ισοδύναμα:

Βρες $\overline{\mathbf{p}} \in E = Span(\mathbf{a}_1, \mathbf{a}_2) : \| \overline{\mathbf{p}} - \mathbf{b} \| \leq \| \mathbf{p} - \mathbf{b} \| \ \forall \mathbf{p} \in E$

Ή: **Βρες εκείνο το σημείο** $\overline{\mathbf{p}}$ **του** E **που να έχει ελάχιστη απόσταση από το** $\mathbf{b}$ (συγκρινόμενο με την απόσταση του $\mathbf{b}$ από άλλα σημεία $\mathbf{p}$ του E).

Σχηματικά:



Ίδιο πρόβλημα για οποιοδήποτε n

Γενικώς:

$$\mathbf{A} = [\mathbf{a}_1 | \ldots | \mathbf{a}_n], \ \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Βρες $\overline{\mathbf{x}} = \begin{pmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_n \end{pmatrix} : \| \mathbf{A}\overline{\mathbf{x}} - \mathbf{b} \| \leq \| \mathbf{A}\mathbf{x} - \mathbf{b} \| \ \forall \mathbf{x} \in R^n$

Θέτοντας $\mathbf{p} = \mathbf{A}\mathbf{x} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n \in R(\mathbf{A}) = Span(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ και

$$\overline{\mathbf{p}} = \mathbf{A}\overline{\mathbf{x}} = \overline{x}_1\mathbf{a}_1 + \ldots + \overline{x}_n\mathbf{a}_n \in R(\mathbf{A}) = Span(\mathbf{a}_1, \ldots, \mathbf{a}_n)$$

το πρόβλημα γράφεται ισοδύναμα ως:

Βρες $\overline{\mathbf{p}} \in R(\mathbf{A}) : \| \overline{\mathbf{p}} - \mathbf{b} \| \leq \| \mathbf{p} - \mathbf{b} \| \ \forall \mathbf{p} \in R(\mathbf{A})$

με $\overline{\mathbf{p}} = \mathbf{A}\overline{\mathbf{x}}$ και $\mathbf{p} = \mathbf{A}\mathbf{x}$.

Βρες εκείνο το σημείο $\overline{\mathbf{p}} \in R(\mathbf{A})$ **που να ελαχιστοποιεί την απόσταση από το** b.

Απάντηση: **Θα είναι η προβολή του** $\mathbf{b}$ **στο** $R(\mathbf{A})$**, δηλαδή εκείνο το σημείο** $\overline{\mathbf{p}} \in R(\mathbf{A})$ **με** $\mathbf{b} - \overline{\mathbf{p}} \perp$ **σε κάθε στοιχείο του** $R(\mathbf{A})$**.**

Διότι: Έστω $\mathbf{p} \in R(\mathbf{A})$. Θα δείξουμε $\|\mathbf{b} - \mathbf{p}\|^2 \geq \|\mathbf{b} - \overline{\mathbf{p}}\|^2$

$\|\mathbf{b} - \mathbf{p}\|^2 =$

$= \|\mathbf{b} - \overline{\mathbf{p}} + \overline{\mathbf{p}} - \mathbf{p}\|^2 =$

$= \left\langle (\mathbf{b} - \overline{\mathbf{p}}) + (\overline{\mathbf{p}} - \mathbf{p}), (\mathbf{b} - \overline{\mathbf{p}}) + (\overline{\mathbf{p}} - \mathbf{p}) \right\rangle =$

$= \|\mathbf{b} - \overline{\mathbf{p}}\|^2 + \|\overline{\mathbf{p}} - \mathbf{p}\|^2 + 2\left\langle \mathbf{b} - \overline{\mathbf{p}}, \overline{\mathbf{p}} - \mathbf{p} \right\rangle$

$\left\langle \mathbf{b} - \overline{\mathbf{p}}, \overline{\mathbf{p}} - \mathbf{p} \right\rangle = 0$ εκ κατασκευής διότι $\overline{p} - p \in R(A)$

$\geq \|\mathbf{b} - \overline{\mathbf{p}}\|^2$

Υπολογισμός των $\overline{\mathbf{p}}$, $\overline{\mathbf{x}}$:

$\mathbf{b} - \mathbf{A}\overline{\mathbf{x}} \perp$ σε κάθε στοιχείο του $R(\mathbf{A})$

$$\left.\begin{array}{l} \Leftrightarrow \mathbf{b} - \mathbf{A}\overline{\mathbf{x}} \perp \mathbf{a}_1 \Rightarrow \mathbf{a}_1^{\mathbf{T}}(\mathbf{b} - \mathbf{A}\overline{\mathbf{x}}) = 0 \\ \qquad\qquad \perp \mathbf{a}_2 \Rightarrow \mathbf{a}_2^{\mathbf{T}}(\mathbf{b} - \mathbf{A}\overline{\mathbf{x}}) = 0 \\ \qquad\qquad \vdots \\ \qquad\qquad \perp \mathbf{a}_\mathbf{n} \Rightarrow \mathbf{a}_\mathbf{n}^{\mathbf{T}}(\mathbf{b} - \mathbf{A}\overline{\mathbf{x}}) = 0 \end{array}\right\}$$

$$\Leftrightarrow \mathbf{A}^{\mathbf{T}} \cdot (\mathbf{b} - \mathbf{A}\overline{\mathbf{x}}) = 0$$

$$\Leftrightarrow \mathbf{A}^{\mathbf{T}}\mathbf{A}\overline{\mathbf{x}} = \mathbf{A}^{\mathbf{T}}\mathbf{b}$$

- Η λύση $\overline{x}$ ικανοποιεί την εξίσωση $\mathbf{A}^{\mathbf{T}}\mathbf{A}\overline{\mathbf{x}} = \mathbf{A}^{\mathbf{T}}\mathbf{b}$
- Εάν οι στήλες του $\mathbf{A}$ είναι ανεξάρτητες $\Rightarrow \mathbf{A}^{\mathbf{T}}\mathbf{A}$ αντιστρέψιμος και $\overline{\mathbf{x}} = (\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{b}$
- Τότε και η προβολή του $\mathbf{b}$ στο $R(\mathbf{A})$ $P_{R(\mathbf{A})}\mathbf{b}$ υπολογίζεται ως $\overline{p} = P_{R(\mathbf{A})}\mathbf{b} = \mathbf{A}\overline{\mathbf{x}} = \mathbf{A}(\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{b}$.

Ο πίνακας $\mathbf{P} = \mathbf{A}(\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}$ ονομάζεται **πίνακας προβολής** στο $R(\mathbf{A})$. Έχει την ιδιότητα ότι οποιοδήποτε $\mathbf{b} \in R^m$ πολλαπλασιάσω με τον $\mathbf{P}$ θα πάρω την προβολή του $\mathbf{b}$ στο $R(\mathbf{A})$.

Παράδειγμα 1: $\mathbf{A}=\begin{bmatrix}1 & 2\\ 1 & 3\\ 0 & 0\end{bmatrix}$, $\mathbf{b}=\begin{pmatrix}4\\5\\6\end{pmatrix}$, $R(\mathbf{A})=Span\left(\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}2\\3\\0\end{pmatrix}\right)$

$P=?$, $P_{R(\mathbf{A})}\mathbf{b}=?$, $\overline{\mathbf{x}}=?$

$$\begin{bmatrix}1 & 2\\ 1 & 3\\ 0 & 0\end{bmatrix}$$

Τύπος στο 2×2:

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}^{-1}=\frac{1}{ad-cb}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$$

$$\mathbf{A^TA}=\begin{bmatrix}1 & 1 & 0\\ 2 & 3 & 0\end{bmatrix}\begin{bmatrix}2 & 5\\ 5 & 13\end{bmatrix}$$

άρα $\mathbf{(A^TA)^{-1}}=\dfrac{1}{26-25}\begin{bmatrix}13 & -5\\ -5 & 2\end{bmatrix}$

$$\begin{bmatrix}1 & 1 & 0\\ 2 & 3 & 0\end{bmatrix}$$

$$\mathbf{(A^TA)^{-1}A^T}=\begin{bmatrix}13 & -5\\ -5 & 2\end{bmatrix}\begin{bmatrix}3 & -2 & 0\\ -1 & 1 & 0\end{bmatrix}$$

$$\begin{bmatrix} 3 & -2 & 0 \\ -1 & 1 & 0 \end{bmatrix}$$

και $\mathbf{P = A\Big((A^T A)^{-1} A\Big) =} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \longleftarrow \cancel{P}$

$$P_{R(\mathbf{A})}\mathbf{b} = \mathbf{Pb} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$$

$$\overline{\mathbf{x}} = \left[(A^T A)^{-1} A^T\right] b = \begin{bmatrix} 3 & -2 & 0 \\ -1 & 1 & 0 \end{bmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} \longleftarrow \overline{\mathbf{x}}$$

Δηλ. $P_{R(\mathbf{A})}\mathbf{b} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \mathbf{A}\overline{\mathbf{x}} = \overline{x}_1\mathbf{a_1} + \overline{x}_2\mathbf{a_2} = 2\cdot\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 1\cdot\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$

<u>Παράδειγμα 2:</u> Έστω $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix}$. Να υπολογιστεί ο πίνακας προβολής $\mathbf{\Gamma}$ στο $Span(\mathbf{v})$.

Με $\mathbf{V} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix}$, $\mathbf{\Gamma = V(V^T V)^{-1} V^T = v(v^T v)^{-1} v^T}$

$$= \frac{\mathbf{1}}{\mathbf{v^T v}} \mathbf{v v^T} = \frac{1}{1+1+4+4} \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 2 & 2 \end{pmatrix}$$

$$= \frac{1}{10} \cdot \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 2 & 2 & 4 & 4 \\ 2 & 2 & 4 & 4 \end{bmatrix}$$

Ερώτηση – Άσκηση: Έστω U ένας υπόχωρος και έστω $\{\mathbf{v}_1,\ldots,\mathbf{v}_r\}$ μια βάση του U. Να βρεθεί πίνακας $\boldsymbol{\Gamma}$ τέτοιος που $\boldsymbol{\Gamma}\mathbf{x}$ να είναι για κάθε $\mathbf{x}$ το πλησιέστερο στο $\mathbf{x}$ σημείο του U.

Απάντηση: $\boldsymbol{\Gamma}$ πρέπει να **είναι ο πίνακας προβολής στο** U. Για να τον βρω πρέπει να σχηματίσω πίνακα $\mathbf{A}$, τέτοιο που $U = R(\mathbf{A})$. Ο $\mathbf{A} = [\mathbf{v}_1 | \ldots | \mathbf{v}_r]$ έχει αυτή την ιδιότητα. Άρα έχουμε $\boldsymbol{\Gamma} = \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}$.

Άσκηση για το σπίτι

Σεπτ 2004/ Θέμα 5$^{\text{ο}}$

Δίνονται τα ακόλουθα 4 στοιχεία του R^5:

$$a = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \mathrm{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, c = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} \text{ και } v = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}$$

Έστω E ο υπόχωρος που παράγουν τα a, b και c μαζί και έστω F το ορθογώνιο συμπλήρωμα του E.

5.1 Να βρεθεί η προβολή του v στο E.

5.2 Να βρεθεί πίνακας $\mathbf{A}$ τέτοιος ώστε --για οποιοδήποτε $\mathbf{x}$-- το $\mathbf{Ax}$ να είναι εκείνο το πλησιέστερο στο $\mathbf{x}$ στοιχείο του E.

5.3 Να βρεθεί πίνακας $\mathbf{B}$ τέτοιος ώστε --για οποιοδήποτε $\mathbf{x}$-- το $\mathbf{Bx}$ να είναι το πλησιέστερο στο $\mathbf{x}$ πολλαπλάσιο του a.

5.4 Να βρεθούν $\mathbf{v}_0 \in E$ και $\mathbf{v}_1 \in F$, τέτοια ώστε $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_1$.

5.5 Να βρεθεί ο πίνακας προβολής στον F.

5.2 Έστω $\mathbf{\Delta} = \left[\mathbf{a}|\mathbf{b}|\mathbf{c}\right] = \left[\begin{array}{c|c|c} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 1 & 0 \end{array}\right]$. Τότε $\mathbf{A}$ είναι ο πίνακας προβολής στο E

και υπολογίζεται ως $\mathbf{A} = \mathbf{\Delta}\left(\mathbf{\Delta}^{\mathbf{T}}\mathbf{\Delta}\right)^{\mathbf{-1}}\mathbf{\Delta}^{\mathbf{T}}$.

$$\mathbf{\Delta}^{\mathbf{T}}\mathbf{\Delta} = \begin{bmatrix} 1 & 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 0 \end{bmatrix} \left[\begin{array}{c|c|c} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 1 & 0 \end{array}\right] = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$\left(\mathbf{\Delta}^{\mathbf{T}}\mathbf{\Delta}\right)^{\mathbf{-1}}$=?

$$\left[\begin{array}{ccc|ccc} 3 & 1 & 0 & 1 & 0 & 0 \\ 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 3 & 1 & 0 & 1 & 0 & 0 \\ 0 & 8/3 & 0 & -1/3 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 3 & 0 & 0 & 9/8 & -3/8 & 0 \\ 0 & 8/3 & 0 & -1/3 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3/8 & -1/8 & 0 \\ 0 & 1 & 0 & -1/8 & 3/8 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1/2 \end{array}\right]$$. Άρα $\left(\mathbf{\Delta}^{\mathbf{T}}\mathbf{\Delta}\right)^{\mathbf{-1}} = \begin{pmatrix} 3/8 & -1/8 & 0 \\ -1/8 & 3/8 & 0 \\ 0 & 0 & 1/2 \end{pmatrix}$

$$\left(\mathbf{\Delta}^{\mathbf{T}}\mathbf{\Delta}\right)^{\mathbf{-1}}\mathbf{\Delta}^{\mathbf{T}} = ?$$

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

$$\left(\mathbf{\Delta}^{\mathbf{T}}\mathbf{\Delta}\right)^{\mathbf{-1}}\mathbf{\Delta}^{\mathbf{T}} = \frac{1}{8}\begin{pmatrix}3 & -1 & 0\\ -1 & 3 & 0\\ 0 & 0 & 4\end{pmatrix} \qquad \frac{1}{8}\begin{bmatrix}2 & 0 & 2 & 0 & -4\\ 2 & 0 & 2 & 0 & 4\\ 0 & 4 & 0 & -4 & 0\end{bmatrix}$$

$$\mathbf{A} = \mathbf{\Delta}\left(\mathbf{\Delta}^{\mathbf{T}}\mathbf{\Delta}\right)^{\mathbf{-1}}\mathbf{\Delta}^{\mathbf{T}} = ?$$

$$\frac{1}{8}\begin{bmatrix}2 & 0 & 2 & 0 & -4\\ 2 & 0 & 2 & 0 & 4\\ 0 & 4 & 0 & -4 & 0\end{bmatrix}$$

$$\mathbf{A} = \mathbf{\Delta}\left(\mathbf{\Delta}^{\mathbf{T}}\mathbf{\Delta}\right)^{\mathbf{-1}}\mathbf{\Delta}^{\mathbf{T}} = \begin{bmatrix}1 & 1 & 0\\ 0 & 0 & 1\\ 1 & 1 & 0\\ 0 & 0 & -1\\ -1 & 1 & 0\end{bmatrix} \qquad \frac{1}{8}\begin{bmatrix}4 & 0 & 4 & 0 & 0\\ 0 & 4 & 0 & -4 & 0\\ 4 & 0 & 4 & 0 & 0\\ 0 & -4 & 0 & 4 & 0\\ 0 & 0 & 0 & 0 & 8\end{bmatrix}$$

Άρα $\mathbf{A} = \frac{1}{2}\begin{bmatrix}1 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & -1 & 0\\ 1 & 0 & 1 & 0 & 0\\ 0 & -1 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 2\end{bmatrix}$

5.1

$$P_E\mathbf{v} = \mathbf{A}\mathbf{v} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 \\ -1 \\ 2 \\ 1 \\ 2 \end{pmatrix}.$$

Άρα $P_E\mathbf{v} = \frac{1}{2}\begin{pmatrix} 2 \\ -1 \\ 2 \\ 1 \\ 2 \end{pmatrix}$

$$\overline{\mathbf{x}} = \left(\boldsymbol{\Delta}^{\mathbf{T}}\boldsymbol{\Delta}\right)^{\mathbf{-1}}\boldsymbol{\Delta}^{\mathbf{T}}\mathbf{v} = \frac{1}{8}\begin{bmatrix} 2 & 0 & 2 & 0 & -4 \\ 2 & 0 & 2 & 0 & 4 \\ 0 & 4 & 0 & -4 & 0 \end{bmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{8}\begin{pmatrix} 0 \\ 8 \\ -4 \end{pmatrix}.$$

Άρα $\overline{\mathbf{x}} = \begin{pmatrix} 0 \\ 1 \\ -1/2 \end{pmatrix}$.

$$P_E\mathbf{v} = \boldsymbol{\Delta}\overline{\mathbf{x}} = \overline{x}_1\mathbf{a} + \overline{x}_2\mathbf{b} + \overline{x}_3\mathbf{c} = 0\mathbf{a} + 1\mathbf{b} - \frac{1}{2}\mathbf{c} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1/2 \\ 1 \\ 1/2 \\ 1 \end{pmatrix}$$

5.3 Έστω $\mathbf{a}=\begin{pmatrix}1\\0\\1\\0\\-1\end{pmatrix}$. Να υπολογιστεί ο πίνακας προβολής $\mathbf{B}$ στο $Span(\mathbf{a})$.

Με $\mathbf{Z}=\mathbf{a}=\begin{pmatrix}1\\0\\1\\0\\-1\end{pmatrix}$, $\mathbf{B}=\mathbf{Z}(\mathbf{Z}^{\mathbf{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathbf{T}}=\mathbf{a}(\mathbf{a}^{\mathbf{T}}\mathbf{a})^{-1}\mathbf{a}^{\mathbf{T}}=\dfrac{\mathbf{1}}{\mathbf{a}^{\mathbf{T}}\mathbf{a}}\mathbf{a}\mathbf{a}^{\mathbf{T}}$

$$=\frac{1}{1+1+1}\cdot\begin{pmatrix}1\\0\\1\\0\\-1\end{pmatrix}\cdot\begin{pmatrix}1 & 0 & 1 & 0 & -1\end{pmatrix}=\frac{1}{3}\cdot\begin{bmatrix}1 & 0 & 1 & 0 & -1\\0 & 0 & 0 & 0 & 0\\1 & 0 & 1 & 0 & -1\\0 & 0 & 0 & 0 & 0\\-1 & 0 & -1 & 0 & 1\end{bmatrix}$$

5.4 $\mathbf{v_0}=P_E\mathbf{v}=\dfrac{1}{2}\begin{pmatrix}2\\-1\\2\\1\\2\end{pmatrix}$ και $\mathbf{v_1}=\mathbf{v}-\mathbf{v_0}=\begin{pmatrix}1\\1\\1\\2\\1\end{pmatrix}-\dfrac{1}{2}\begin{pmatrix}2\\-1\\2\\1\\2\end{pmatrix}=\begin{pmatrix}0\\3/2\\0\\3/2\\0\end{pmatrix}$

5.5

$$=\mathbf{I}-\mathbf{A}=\begin{bmatrix}1 & 0 & 0 & 0 & 0\\0 & 1 & 0 & 0 & 0\\0 & 0 & 1 & 0 & 0\\0 & 0 & 0 & 1 & 0\\0 & 0 & 0 & 0 & 1\end{bmatrix}-\frac{1}{2}\begin{bmatrix}1 & 0 & 1 & 0 & 0\\0 & 1 & 0 & -1 & 0\\1 & 0 & 1 & 0 & 0\\0 & -1 & 0 & 1 & 0\\0 & 0 & 0 & 0 & 2\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1 & 0 & -1 & 0 & 0\\0 & 1 & 0 & 1 & 0\\-1 & 0 & 1 & 0 & 0\\0 & 1 & 0 & 1 & 0\\0 & 0 & 0 & 0 & 0\end{bmatrix}$$

14. Προβολές και προσεγγίσεις ελαχίστων τετραγώνων, Μέρος Β.

(Πηγή: Strang, Κεφάλαιο 3.3)

Σημείωση 1.

Αν $\mathbf{b} \in R(\mathbf{A})$ τότε $\exists \mathbf{x} : \mathbf{Ax} = \mathbf{b}$ και $\bar{\mathbf{p}} = \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b} = \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{Ax} = \mathbf{Ax} = \mathbf{b}$

Σημείωση 2.

Αν $\mathbf{b} \perp R(\mathbf{A})$ τότε $\mathbf{A}^{\mathrm{T}}\mathbf{b} = \mathbf{0}$ και $\bar{\mathbf{p}} = \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b} = \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{0} = \mathbf{0}$.

Σημείωση 3.

Αν $\mathbf{A}$ $n \times n$ αντιστρέψιμος τότε $R(\mathbf{A}) = \mathbb{R}^n$ και

$$\bar{\mathbf{p}} = \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b} = \mathbf{A}\mathbf{A}^{-1}\left(\mathbf{A}^{\mathrm{T}}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b} = \mathbf{b}$$

Σημείωση 4. Για $b \in \mathbb{R}^n$ *έχουμε*

$$\mathbf{b} = P_{R(\mathbf{A})}(\mathbf{b}) + \left[\mathbf{b} - P_{R(\mathbf{A})}(\mathbf{b})\right].$$

- *Εδώ η* $\mathbf{b} - P_{R(\mathbf{A})}(\mathbf{b})$ *είναι εκ κατασκευής* $\perp R(\mathbf{A})$*. Άρα* $\mathbf{b} - P_{R(\mathbf{A})}(\mathbf{b}) \in R(\mathbf{A})^{\perp}$*.*
- *Από την άλλη προφανώς* $P_{R(\mathbf{A})}(\mathbf{b}) \in R(\mathbf{A})$*.*
- *Επομένως αυτή είναι η διάσπαση του b σε μία σθνιστώσα στον* $R(\mathbf{A})$ *και μία στο ορθογώνιο συμπλήρωμά του, το* $R(\mathbf{A})^{\perp}$*.*
- *Η δεύτερη ισούται με* $P_{R(\mathbf{A})^{\perp}}(\mathbf{b}) = \left[\mathbf{b} - P_{R(\mathbf{A})}(\mathbf{b})\right] = [\mathbf{I} - \mathbf{P}]\mathbf{b}$*. Άρα ο*

πίνακας προβολής στο $R(\mathbf{A})^{\perp}$ *δίνεται από I-πίνακα προβολής στο* $R(A)$*.*

Απόδειξη του: «**Εάν οι στήλες του** $\mathbf{A}$ **είναι ανεξάρτητες** $\Rightarrow \mathbf{A}^{\mathrm{T}}\mathbf{A}$ **αντιστρέψιμος**»

Ο πίνακας εσωτερικού γινομένου $\mathbf{A}^{\mathrm{T}}\mathbf{A}$

- Είναι συμμετρικός
- Έχει τον **ίδιο μηδενόχωρο** με τον A, διότι

Αν $\mathbf{x} \in N(\mathbf{A}) \Rightarrow \mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{A}^{\mathrm{T}}\mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{x} \in N(\mathbf{A}^{\mathrm{T}}\mathbf{A})$ και

Αν $\mathbf{x} \in N(\mathbf{A}^{\mathrm{T}}\mathbf{A}) \Rightarrow \mathbf{A}^{\mathrm{T}}\mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{Ax} = \mathbf{0}$

$$\Rightarrow \|\mathbf{Ax}\|^2 = \mathbf{0} \Rightarrow \mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{x} \in N(\mathbf{A}).$$

Άρα όντως:

- αν οι στήλες του $\mathbf{A}$ είναι ανεξάρτητες
- τότε ο $\mathbf{A}$ έχει τάξη n
- και τότε ο $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ έχει την ίδια τάξη και άρα είναι αντιστρέψιμος.

Ιδιότητες: Αν $\mathbf{P}$ πίνακας προβολής τότε:

$\mathbf{P}^2 = \mathbf{P}$ (Αν προβάλλω δύο φορές είναι σαν να πρόβαλα μία)

$\mathbf{P}^T = \mathbf{P}$ ($\mathbf{P}$ συμμετρικός)

Ισχύει και το ανάποδο: Ένας συμμετρικός πίνακας με $\mathbf{P}^2 = \mathbf{P}$ θα είναι πίνακας προβολής σε κάποιο υπόχωρο.

Απόδειξη:

$$\mathbf{P}^2 = \mathbf{P} \cdot \mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{P}$$

και $\mathbf{P}^T = \left[\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\right]^T = \mathbf{A}^{T^T}\left((\mathbf{A}^T\mathbf{A})^{-1}\right)^T\mathbf{A}^T$

$$= \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{P}$$

διότι $\mathbf{A}^T\mathbf{A}$ συμμετρικός

Η απεικόνιση $\mathbf{x} \mapsto P_U \mathbf{x}$.

Έστω $U \subseteq \mathbb{R}^m$ ένας υπόχωρος και έστω $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ μια βάση του U.

Είδαμε ότι αν $\mathbf{\Gamma} = \mathbf{\Gamma}_U$ **είναι ο πίνακας προβολής στο** U τότε $\mathbf{\Gamma} = \mathbf{A}(\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}$ με $\mathbf{A} = [\mathbf{v}_1 | \ldots | \mathbf{v}_r]$.

Επίσης $U = R(\mathbf{A})$.

Θεωρούμε την απεικόνιση $\mathbf{x} \mapsto f(\mathbf{x}) = P_U \mathbf{x}$.

- Προφανώς $f : \mathbb{R}^m \to \mathbb{R}^m$.
- Η απεικόνιση αυτή είναι γραμμική και δίνεται από τον $\mathbf{\Gamma}$:
  $f(\mathbf{x}) = P_U \mathbf{x} = \mathbf{\Gamma x}$. Άρα $f = f_{\mathbf{\Gamma}}$
- Η εικόνα της απεικόνισης, το σύνολο δηλαδή των εικόνων $f(\mathbf{x}) = P_U \mathbf{x}$ για $\mathbf{x} \in \mathbb{R}^m$ είναι ο U. (Τα πάντα προβάλλονται στον U και σε κάθε $\mathbf{x} \in U$ προβάλλεται τουλάχιστον το ίδο το $\mathbf{x}$: αν $\mathbf{x} \in U$ τότε $P_U \mathbf{x} = \mathbf{x}$).
- Άρα $R(\mathbf{\Gamma}) = U$ και τάξη($\mathbf{\Gamma}$) = dim(U) = r
- Επίσης αν $\mathbf{x} \in U^{\perp}$ τότε $P_U \mathbf{x} = \mathbf{\Gamma x} = \mathbf{0}$ και άρα $\mathrm{N}(\mathbf{\Gamma}) = U^{\perp}$ και $\dim[\mathrm{N}(\mathbf{\Gamma})] = m - r$.

Πρόταση. Έστω $W \subseteq \mathbb{R}^m$ ένας υπόχωρος και $\{\mathbf{w_1},\ldots,\mathbf{w_r}\}$ μια βάση του και $V \subseteq \mathbb{R}^m$ ένας άλλος υπόχωρος και $\{\mathbf{v_1},\ldots,\mathbf{v_s}\}$ μια βάση του.

Ορίζουμε ένα καινούργιο υπόχωρο $U := Span\{\mathbf{w_1},\ldots,\mathbf{w_r},\mathbf{v_1},\ldots,\mathbf{v_s}\}$.

Τότε ισχύει πως αν $W \perp V$ θα έχουμε $P_U\mathbf{x} = P_V\mathbf{x} + P_W\mathbf{x}$, $\forall \mathbf{x} \in \mathbb{R}^m$.

Απόδειξη. Ορίζω τους πίνακες $\mathbf{\Gamma} = [\mathbf{w_1}\,|\ldots|\,\mathbf{w_r}]$, $\mathbf{\Delta} = [\mathbf{v_1}\,|\ldots|\,\mathbf{v_s}]$ και $\mathbf{E} = [\mathbf{\Gamma}\,|\,\mathbf{\Delta}] = [\mathbf{w_1}\,|\ldots|\,\mathbf{w_r}\,|\,\mathbf{v_1}\,|\ldots|\,\mathbf{v_s}]$.

Τότε $P_U\mathbf{x} = \mathbf{E(E^TE)^{-1}E^Tx}$.

Έχουμε:

$$\mathbf{E}^{\mathrm{T}}\mathbf{E} = \begin{bmatrix} \mathbf{\Gamma}^{\mathrm{T}} \\ \mathbf{\Delta}^{\mathrm{T}} \end{bmatrix} [\mathbf{\Gamma}\,|\,\mathbf{\Delta}] = \begin{bmatrix} \mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma} & \mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Delta} \\ \mathbf{\Delta}^{\mathrm{T}}\mathbf{\Gamma} & \mathbf{\Delta}^{\mathrm{T}}\mathbf{\Delta} \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Delta}^{\mathrm{T}}\mathbf{\Delta} \end{bmatrix}$$

$$\left(\mathbf{E}^{\mathrm{T}}\mathbf{E}\right)^{-1} = \begin{bmatrix} \mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Delta}^{\mathrm{T}}\mathbf{\Delta} \end{bmatrix}^{-1} = \begin{bmatrix} \left(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma}\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{\Delta}^{\mathrm{T}}\mathbf{\Delta}\right)^{-1} \end{bmatrix}$$

$$\mathbf{E}\left(\mathbf{E}^{\mathrm{T}}\mathbf{E}\right)^{-1} = [\mathbf{\Gamma}\,|\,\mathbf{\Delta}] \begin{bmatrix} \left(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma}\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{\Delta}^{\mathrm{T}}\mathbf{\Delta}\right)^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma}\left(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma}\right)^{-1} & \mathbf{\Delta}\left(\mathbf{\Delta}^{\mathrm{T}}\mathbf{\Delta}\right)^{-1} \end{bmatrix}$$

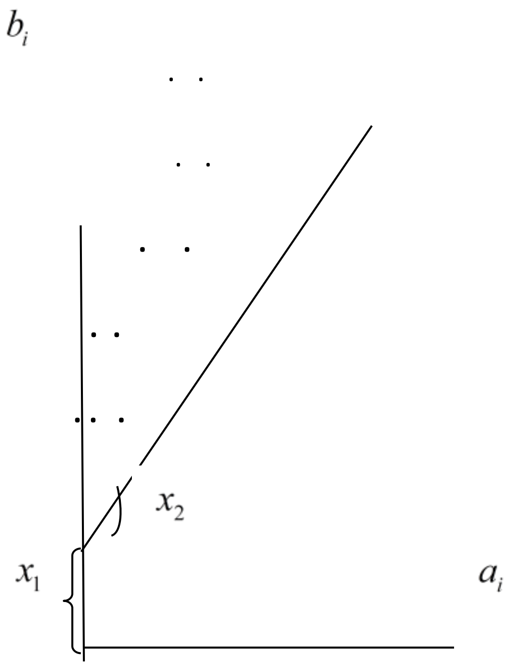
$$\mathbf{E}\left(\mathbf{E}^{\mathrm{T}}\mathbf{E}\right)^{-1}\mathbf{E}^{\mathrm{T}} = \begin{bmatrix} \mathbf{\Gamma}\left(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma}\right)^{-1} & \mathbf{\Delta}\left(\mathbf{\Delta}^{\mathrm{T}}\mathbf{\Delta}\right)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{\Gamma}^{\mathrm{T}} \\ \mathbf{\Delta}^{\mathrm{T}} \end{bmatrix} = \mathbf{\Gamma}\left(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma}\right)^{-1}\mathbf{\Gamma}^{\mathrm{T}} + \mathbf{\Delta}\left(\mathbf{\Delta}^{\mathrm{T}}\mathbf{\Delta}\right)^{-1}\mathbf{\Delta}^{\mathrm{T}}$$

Άρα $P_U\mathbf{x} = \mathbf{E(E^TE)^{-1}E^Tx} = \mathbf{\Gamma}\left(\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma}\right)^{-1}\mathbf{\Gamma}^{\mathrm{T}}\mathbf{x} + \mathbf{\Delta}\left(\mathbf{\Delta}^{\mathrm{T}}\mathbf{\Delta}\right)^{-1}\mathbf{\Delta}^{\mathrm{T}}\mathbf{x} = P_W\mathbf{x} + P_V\mathbf{x}$

Ευθεία ελαχίστων τετραγώνων

Έστω έχω n μετρήσεις a_i, b_i π.χ. b_i πίεση και a_i θερμοκρασία.

Οι μετρήσεις αυτές σε διάγραμμα διασποράς βρίσκονται κοντά σε ευθεία (άγνωστη) και **υποψιάζομαι ότι οι δύο ποσότητες συνδέονται με μια γραμμική σχέση την οποία θέλω να εκτιμήσω.**



$$b_i = x_1 + x_2 a_i + \varepsilon_i$$

Πώς; Ιδέα: Βρες **εκείνη την ευθεία (εκείνα τα x_1, x_2)** που να είναι **«κοντά στα σημεία», δηλ. όλα τα** $b_i - x_1 - x_2 a_i$ **μικρά** ή καλύτερα «ελάχιστα».

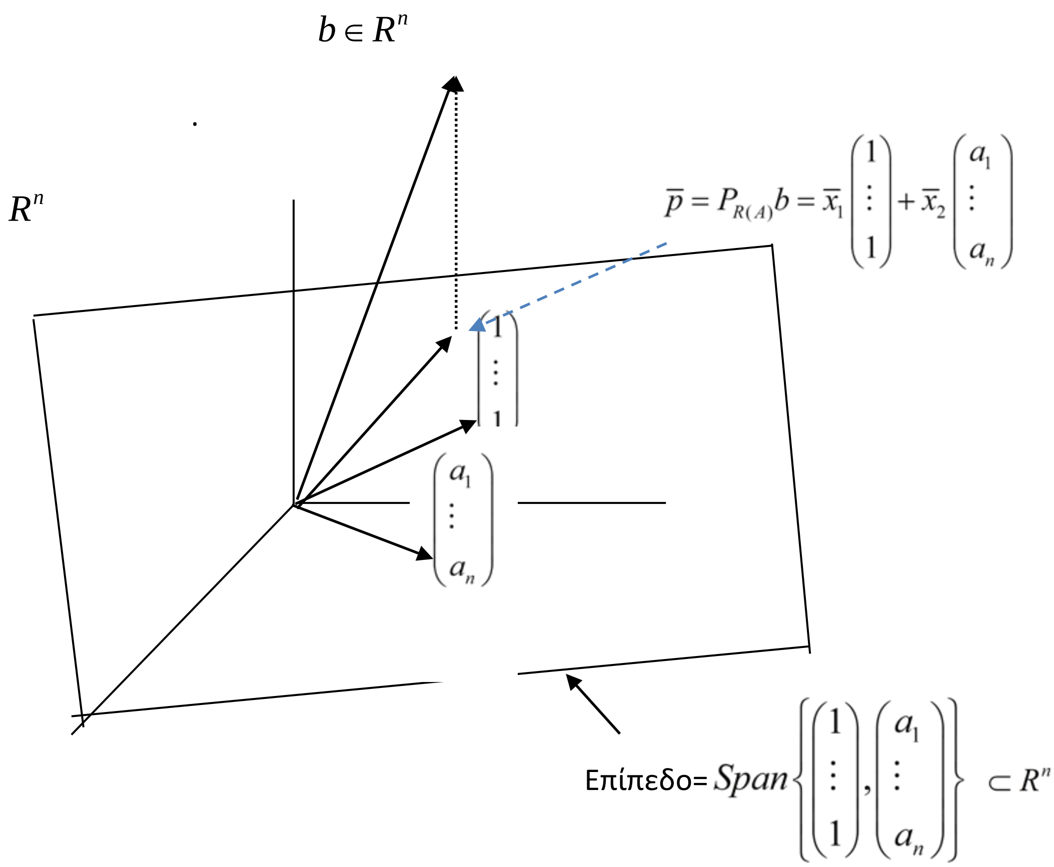
$$\text{π.χ. } \sum_{i=1}^{n} (b_i - x_1 - x_2 a_i)^2 \text{ **ελάχιστο**.}$$

Αλλά

$$\sum_{i=1}^{n}(b_i - x_1 - x_2 a_i)^2 = \left\| \begin{pmatrix} b_1 - x_1 - x_2 a_1 \\ \vdots \\ b_n - x_1 - x_2 a_n \end{pmatrix} \right\|^2 = \left\| \underbrace{\begin{pmatrix} 1 & a_1 \\ \vdots & \vdots \\ 1 & a_n \end{pmatrix}}_{\substack{\mathbf{A} \\ \text{γνωστός πίνακας}}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\substack{\mathbf{x} \\ \text{άγνωστοι}}} - \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}}_{\substack{\mathbf{b} \\ \text{γνωστό διάνυσμα}}} \right\|^2$$

Που ελαχιστοποιείται για $\begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \end{pmatrix} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$ με $\mathbf{A} = \begin{pmatrix} 1 & a_1 \\ \vdots & \vdots \\ 1 & a_n \end{pmatrix}$.

Σημειώστε ότι η λύση της ευθείας ελαχίστων τετραγώνων **συνδέεται με την προβολή του** $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in R^n$ στο $Span\left\{ \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \right\}$



Θα έχουμε $\begin{pmatrix}\overline{x}_1\\ \overline{x}_2\end{pmatrix} = \mathbf{(A^T A)^{-1} A^T b} = \begin{bmatrix} n & \Sigma a_i \\ \Sigma a_i & \Sigma a_i^{\ 2}\end{bmatrix}^{-1} \begin{pmatrix}\Sigma b_i \\ \Sigma a_i b_i\end{pmatrix}$

$$= \frac{1}{n\Sigma a_i^{\ 2} - (\Sigma a_i)^2} \begin{bmatrix} \Sigma a_i^{\ 2} & -\Sigma a_i \\ -\Sigma a_i & n \end{bmatrix} \begin{pmatrix}\Sigma b_i \\ \Sigma a_i b_i\end{pmatrix} = \dots$$

Θέτουμε

$$\overline{a} = \frac{1}{n}\sum a_i, \overline{b} = \frac{1}{n}\sum b_i \, ,$$

$$\overline{AB} := \frac{1}{n}\sum\left(a_i - \overline{a}\right)\left(b_i - \overline{b}\right), \overline{A^2} := \frac{1}{n}\sum\left(a_i - \overline{a}\right)^2$$

Τότε

$$n\overline{A^2} = \sum a_i^2 - n\overline{a}^2 \qquad \text{και} \qquad n\overline{AB} = \sum a_i b_i - n\overline{a}\overline{b}$$

Έτσι

$$\overline{x}_2 = \frac{n\Sigma a_i b_i - \Sigma a_i \Sigma b_i}{n\Sigma a_i^{\ 2} - (\Sigma a_i)^2} = \frac{n^2\overline{AB} + n^2\overline{a}\overline{b} - n^2\overline{a}\overline{b}}{n^2\overline{A^2} + n^2\overline{a}^2 - n^2\overline{a}^2}$$

$$= \frac{\overline{AB}}{\overline{A^2}} = \frac{\sum\left(a_i - \overline{a}\right)\left(b_i - \overline{b}\right)}{\sum\left(a_i - \overline{a}\right)^2}$$

$$\overline{x}_1 = \frac{\Sigma a_i{}^2 \Sigma b_i - \Sigma a_i \Sigma a_i b_i}{n^2\, \overline{A^2}} = \frac{n^2\, \overline{A^2}\overline{b} + n^2 \overline{a}^2 \overline{b} - n^2 \overline{a}\left(\overline{AB} + \overline{a}\overline{b}\right)}{n^2\, \overline{A^2}}$$

$$= \frac{\overline{A^2}\overline{b} - \overline{a}\left(\overline{AB}\right)}{\overline{A^2}} = \overline{b} - \overline{a}\,\frac{\overline{AB}}{\overline{A^2}} = \overline{b} - \overline{a}\ \overline{x}_2$$

Αριθμητικό παράδειγμα:

$$a_1 = -1, b_1 = 1$$

$$a_2 = 1, b_2 = 1$$

$$a_3 = 3, b_3 = 3$$

Άρα $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$, και $\begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \end{pmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}^{-1} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 9/7 \\ 4/7 \end{pmatrix}$

$$\overline{\mathbf{p}} = P_{R(\mathbf{A})}\mathbf{b} = \mathbf{A}\overline{\mathbf{x}} = \begin{pmatrix} \overline{p}_1 \\ \overline{p}_2 \\ \overline{p}_3 \end{pmatrix} = \begin{pmatrix} \overline{x}_1 + \overline{x}_2 a_1 \\ \overline{x}_1 + \overline{x}_2 a_2 \\ \overline{x}_1 + \overline{x}_2 a_3 \end{pmatrix} = \begin{pmatrix} 5/7 \\ 13/7 \\ 17/7 \end{pmatrix}$$

Τα σημεία $(a_i, b_i) \notin$ στην ευθεία $\Leftrightarrow \mathbf{b} \notin R(\mathbf{A}) = Span(\mathbf{1}, \mathbf{a})$

Τα σημεία $(a_i, \overline{p}_i) \notin$ στην ευθεία $\Leftrightarrow \overline{\mathbf{p}} \in R(\mathbf{A}) = Span(\mathbf{1}, \mathbf{a})$

Αποκλίσεις $\varepsilon_i := b_i - \overline{x}_1 - \overline{x}_2 a_i$, $\boldsymbol{\varepsilon} = \mathbf{b} - \overline{\mathbf{p}}$

Άλλο παράδειγμα

Έστω έχω n μετρήσεις y_i, t_i, i=1,…,n

Ψάχνω να βρώ μια παραβολή στο t που να προσεγγίζει τα y:

$$y_i = x_1 + x_2 t_i + x_3 t_i^2 + \varepsilon_i$$

Δηλαδή:

ψάχνω $\overline{\mathbf{x}} = \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix}$ που να ελαχιστοποιεί την $\sum_{i=1}^{n} \left[y_i - \left(x_1 + x_2 t_i + x_3 t_i^2 \right) \right]^2$

Τώρα

$$\sum_{i=1}^{n} \left[y_i - \left(x_1 + x_2 t_i + x_3 t_i^2 \right) \right]^2 = \left\| \begin{pmatrix} y_1 - x_1 - x_2 t_1 - x_3 t_1^2 \\ \vdots \\ y_n - x_1 - x_2 t_n - x_3 t_n^2 \end{pmatrix} \right\|^2 = \left\| \begin{pmatrix} x_1 + x_2 t_1 + x_3 t_1^2 \\ \vdots \\ x_1 + x_2 t_n + x_3 t_n^2 \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\|^2$$

$$= \left\| \begin{pmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\|^2$$

Λύση το $\begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix} = \mathbf{(A^T A)^{-1} A^T y}$ με $\mathbf{A} = \begin{pmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

Σημειώστε ότι η λύση

συνδέεται με την προβολή του $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in R^n$ στο $Span\left\{ \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}, \begin{pmatrix} t_1^2 \\ \vdots \\ t_n^2 \end{pmatrix} \right\}$

Άλλο παράδειγμα

Έστω έχω n μετρήσεις y_i, w_i, z_i , i=1,…,n

Ψάχνω να βρώ μια παραβολή στο t που να προσεγγίζει τα y:

$$y_i = x_1 + x_2 w_i + x_3 z_i + \varepsilon_i$$

Δηλαδή:

ψάχνω $\overline{\mathbf{x}} = \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix}$ που να ελαχιστοποιεί την $\sum_{i=1}^{n} \left[y_i - \left(x_1 + x_2 w_i + x_3 z_i \right) \right]^2$

Τώρα

$$\sum_{i=1}^{n} \left[y_i - \left(x_1 + x_2 w_i + x_3 z_i \right) \right]^2 = \left\| \begin{pmatrix} y_1 - x_1 - x_2 w_1 - x_3 z_1 \\ \vdots \\ y_n - x_1 - x_2 w_n - x_3 z_n \end{pmatrix} \right\|^2 = \left\| \begin{pmatrix} x_1 + x_2 w_1 + x_3 z_1 \\ \vdots \\ x_1 + x_2 w_n + x_3 z_n \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\|^2$$

$$= \left\| \begin{pmatrix} 1 & w_1 & z_1 \\ \vdots & \vdots & \vdots \\ 1 & w_n & z_n \end{pmatrix} \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\|^2$$

Λύση το $\begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix} = (\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{y}$ με $\mathbf{A} = \begin{pmatrix} 1 & w_1 & z_1 \\ \vdots & \vdots & \vdots \\ 1 & w_n & z_n \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

Σημειώστε ότι η λύση **συνδέεται με την προβολή του**

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in R^n \text{ στο } Span\left\{ \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}, \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \right\}$$

Άλλο παράδειγμα

Έστω έχω n μετρήσεις y_i, w_i, i=1,…,n και μια τρίτη δίτιμη μεταβλητή $d_i \in \{0,1\}$ που ορίζει δυο ομάδες π.χ. $d_i = 0$ για άνδρες και $d_i = 1$ για γυναίκες.

Ψάχνω να βρώ μια γραμμιή σχέση στο w που να προσεγγίζει τα y η οποία όμως μπορεί ναέχει διαφορετική σταθερά για άνδρες και γυναίκες.:

$$y_i = \begin{cases} x_1 + x_2 w_i + \varepsilon_i, \text{ για άνδρες} \\ x_1 + x_2 w_i + x_3 + \varepsilon_i, \text{ για γυναίκες} \end{cases} = x_1 + x_2 w_i + x_3 d_i + \varepsilon_i$$

Δηλαδή:

ψάχνω $\overline{\mathbf{x}} = \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix}$ που να ελαχιστοποιεί την $\sum_{i=1}^{n} \left[y_i - (x_1 + x_2 w_i + x_3 d_i) \right]^2$

Τώρα

$$\sum_{i=1}^{n} \left[y_i - (x_1 + x_2 w_i + x_3 d_i) \right]^2 = \left\| \begin{pmatrix} y_1 - x_1 - x_2 w_1 - x_3 d_1 \\ \vdots \\ y_n - x_1 - x_2 w_n - x_3 d_n \end{pmatrix} \right\|^2 = \left\| \begin{pmatrix} x_1 + x_2 w_1 + x_3 d_1 \\ \vdots \\ x_1 + x_2 w_n + x_3 d_n \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\|^2$$

$$= \left\| \begin{pmatrix} 1 & w_1 & d_1 \\ \vdots & \vdots & \vdots \\ 1 & w_n & d_n \end{pmatrix} \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\|^2$$

Λύση το $\begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix} = \mathbf{(A^T A)^{-1} A^T y}$ με $\mathbf{A} = \begin{pmatrix} 1 & w_1 & d_1 \\ \vdots & \vdots & \vdots \\ 1 & w_n & d_n \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

Η λύση **συνδέεται με την προβολή του** $\mathbf{y} \in R^n$ στο $Span\left\{ \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}, \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} \right\}$.

Άλλο παράδειγμα

Έστω έχω n μετρήσεις y_i, w_i, i=1,…,n.

Ψάχνω να βρώ μια γραμμική σχέση στο w που να προσεγγίζει τα y η οποία όμως μπορεί να αλλάζει κλίση για $w \geq w_0$ γιακάποιο γνωστο w_0.

$$y_i = x_1 + x_2 w_i + x_3 (w_i - w_0)^+ + \varepsilon_i$$

με $(w - w_0)^+ := \begin{cases} w - w_0, \text{για } w \geq w_0 \\ 0, \text{για } w < w_0 \end{cases}$.

Δηλαδή: ψάχνω $\overline{\mathbf{x}} = \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix}$ που

να ελαχιστοποιεί την $\sum_{i=1}^{n} \left[y_i - \left(x_1 + x_2 w_i + x_3 (w_i - w_0)^+ \right) \right]^2$

Τώρα

$$\sum_{i=1}^{n} \left[y_i - \left(x_1 + x_2 w_i + x_3 (w_i - w_0)^+ \right) \right]^2 = \left\| \begin{pmatrix} y_1 - x_1 - x_2 w_1 - x_3 (w_1 - w_0)^+ \\ \vdots \\ y_n - x_1 - x_2 w_n - x_3 (w_n - w_0)^+ \end{pmatrix} \right\|^2$$

$$= \left\| \begin{pmatrix} x_1 + x_2 w_1 + x_3 (w_1 - w_0)^+ \\ \vdots \\ x_1 + x_2 w_n + x_3 (w_n - w_0)^+ \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\|^2$$

$$= \left\| \begin{pmatrix} 1 & w_1 & (w_1 - w_0)^+ \\ \vdots & \vdots & \vdots \\ 1 & w_n & (w_n - w_0)^+ \end{pmatrix} \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\|^2$$

Λύση το $\begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix} = (\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{y}$ με $\mathbf{A} = \begin{pmatrix} 1 & w_1 & (w_1 - w_0)^+ \\ \vdots & \vdots & \vdots \\ 1 & w_n & (w_n - w_0)^+ \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

Η λύση **συνδέεται με την προβολή του** $\mathbf{y} \in R^n$ στο $Span\left\{ \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}, \begin{pmatrix} (w_1 - w_0)^+ \\ \vdots \\ (w_n - w_0)^+ \end{pmatrix} \right\}$.