Κατάσταση γραμμικών συστημάτων

Ας θεωρήσουμε το ακόλουθο παράδειγμα γραμμικού συστήματος

$$Ax = b \quad \text{με} \quad A = \begin{pmatrix} .913 & .659 \\ .780 & .563 \end{pmatrix}, \quad b = \begin{pmatrix} .254 \\ .217 \end{pmatrix}$$

Μοναδική λύση του είναι η $x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Παρατηρούμε ότι αν εφαρμόσουμε τις παραλλαγές της μεθόδου της απαλοιφής Gauss που θεωρήσαμε (μερική οδήγηση, ολική οδήγηση) δεν θα προκύψει εναλλαγή γραμμών. Άρα στον Η/Υ το παραπάνω γραμμικό σύστημα θα γίνει με την απλή απαλοιφή Gauss

Θεωρούμε τώρα έναν Η/Υ με ακρίβεια 3 δεκαδικών ψηφίων ώστε μπορούμε να δούμε ότι θα προκύψει η λύση

$$\tilde{x}_2 = -1 \quad \text{και} \quad \tilde{x}_1 = -.443$$

Αν χρησιμοποιήσουμε έναν Η/Υ με ακρίβεια 6 δεκαδικών ψηφίων ώστε θα πάρουμε τη λύση

$$\tilde{x}_2 = -1 \quad \text{και} \quad \tilde{x}_1 = 1$$

Σε αυτή την περίπτωση το σφάλμα μπορεί να διορθωθεί σε ένα Η/Υ με μεγαλύτερη ακρίβεια.

Ας θεωρήσουμε όμως τώρα το γραμμικό σύστημα με τον ίδιο πίνακα και λίγο διαφορετικό δεξιό μέλος b,

$$b = \begin{pmatrix} .253 \\ .218 \end{pmatrix}$$

Η ακριβή του προβλήματος γίνεται τώρα $(Ax = b)$

$$x = \begin{pmatrix} 1223 \\ -1694 \end{pmatrix}$$

Διαταράξαμε λίγο το δεξιό μέλος b κατά $\begin{pmatrix} -0.001 \\ +0.001 \end{pmatrix}$

και αυτό είχε ως αποτέλεσμα να αλλάξει η ακριβής λύση

κατά $\begin{pmatrix} 1222 \\ -1693 \end{pmatrix}$

Τα προβλήματα που αν τα διαταράξουμε "λίγο" οδηγούν σε "μεγάλη" αλλαγή της ακριβούς λύσης, λέμε ότι έχουν "κακή κατάσταση"

Παρατηρούμε ότι στο παράδειγμα μας $\det(A) = -10^{-6}$

Επειδή $\det(A) \neq 0$, προφανώς ο A είναι αντιστρέψιμος και υπάρχει μοναδική λύση του γραμμικού συστήματος $Ax = b$

<u>Ερώτημα</u>: Η μικρή τιμή της $\det(A)$ μήπως είναι ο λόγος της που η απαλοιφή Gauss αποτυγχάνει στον πρώτο υπολογιστή να δώσει καλά αποτελέσματα;

Ας πολ/με την 1η εξίσωση του $Ax=b$ με έναν μεγάλο αριθμό το 10^8 έτσι παίρνουμε το σύστημα.

$$\underbrace{\begin{pmatrix} .913\times 10^8 & .659\times 10^8 \\ .780 & .563 \end{pmatrix}}_{\tilde{A}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} .254\times 10^8 \\ .217 \end{pmatrix}$$

Τώρα $\det(\tilde{A}) = -100$, όμως η λύση που δίνει ο υπολογιστής με ακρίβεια 3 δεκαδικών ψηφίων παραμένει

$$\tilde{x}_1 = -.443 \quad , \quad \tilde{x}_2 = -1$$

Προφανώς <u>δεν</u> είναι η μικρή τιμή της ορίζουσας που δημιουργεί το <u>πρόβλημα</u>

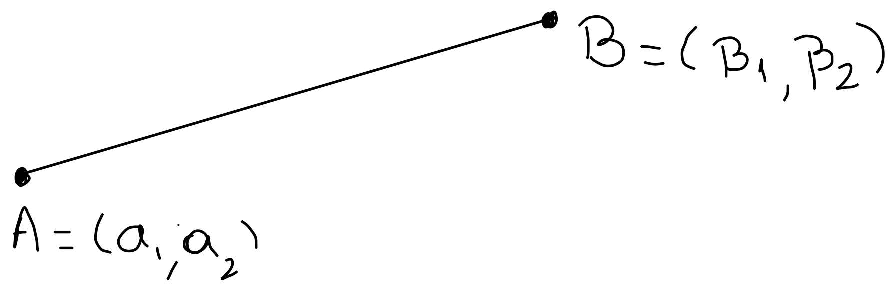
Για να μελετήσουμε το φαινόμενο θα χρειαστούμε μια νέα απεικόνιση, που ονομάζουμε νόρμα.

Με τη βοήθεια της νόρμας θα μπορέσουμε να ορίσουμε απόσταση ανάμεσα σε διανύσματα και πίνακες, καθώς και ένα "μέτρο" κατά πόσο ένα γραμμικό σύστημα είναι σε "κακή κατάσταση"

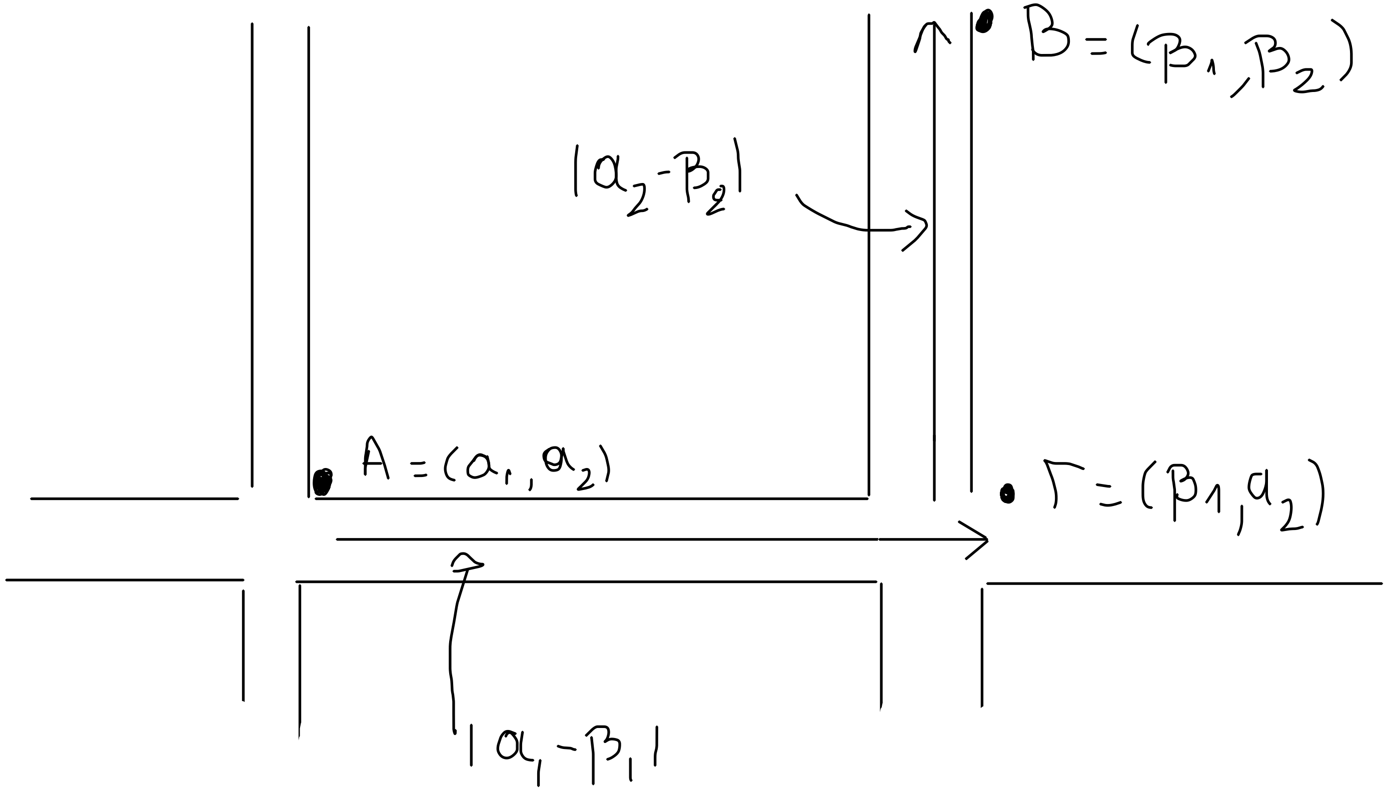
Νόρμες διανυσμάτων

Θα ξεκινήσουμε ορίζοντας την απεικόνιση νόρμα για διανύσματα και στη συνέχεια θα επεκτείνουμε τον ορισμό για πίνακες.

Γνωρίζουμε ήδη στις 2-διαστάσεις πως μπορούμε να μετρήσουμε την απόσταση ανάμεσα σε δύο σημεία A, B με συντεταγμένες $A=(\alpha_1,\alpha_2)$, $B=(\beta_1,\beta_2)$ χρησιμοποιώντας την Ευκλείδια απόσταση $\|A-B\| = \sqrt{(\alpha_1-\beta_1)^2+(\alpha_2-\beta_2)^2}$



Υπάρχουν όμως και άλλα παραδείγματα που δεν μπορούμε να χρησιμοποιήσουμε την ευκλείδεια απόσταση, π.χ. ο χάρτης με οδούς μιας πόλης

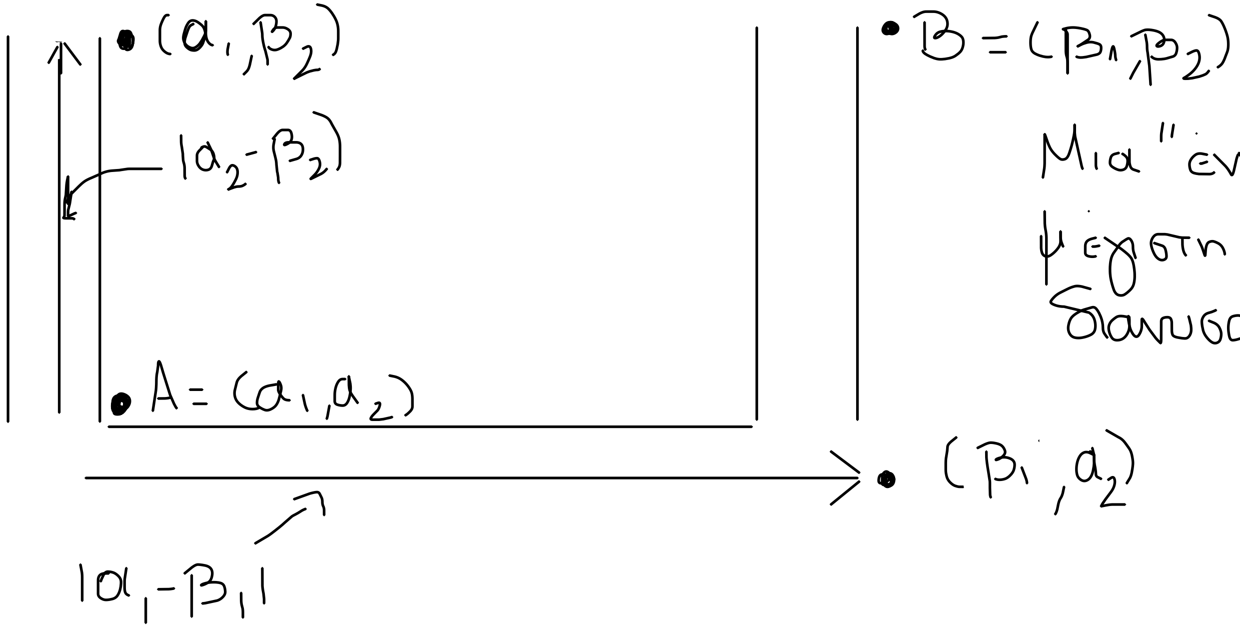


Για να μεταβούμε από το Α στο Β πρέπει να ακολουθούμε συγκεκριμένους δρόμους. Τότε η απόσταση θα είναι το άθροισμα

$$|\alpha_1-\beta_1|+|\alpha_2-\beta_2|$$

Ένα ακόμα παράδειγμα μπορεί να προκύψει αν δεν θέλουμε να μεταβούμε από το ένα σημείο A στο άλλο σημείο B, αλλά να παρατηρήσουμε από μακριά αυτό που βρίσκεται στο σημείο B.

Βρισκόμαστε σε μια πόλη και υπάρχουν ψηλά κτήρια και εμποδίζουν να δούμε αυτό που βρίσκεται στο B από το A, οπότε πρέπει να μετακινηθούμε από το A σε κάποιο άλλο σημείο, ώστε να μπορούμε να παρατηρήσουμε το B



Μια "έννοια απόστασης" μπορεί να είναι η μέγιστη διαδρομή που θα μπορούσαμε να διανύσουμε

$$\max(|a_1-\beta_1|, |a_2-\beta_2|)$$

Ορισμός: Μια απεικόνιση από ένα γραμμικό χώρο X στους πραγματικούς αριθμούς, $\|\cdot\| : X \to \mathbb{R}$, λέγεται νόρμα αν ικανοποιούνται οι ακόλουθες τρεις ιδιότητες

(i) $x \in X$, $\|x\| = 0$ αν και μόνο αν $x = 0$ (0 το μηδενικό στοιχείο του X)

(ii) $\forall x \in X$, $\forall \lambda \in \mathbb{R}$ (ή $\mathbb{C}$), $\|\lambda x\| = |\lambda| \, \|x\|$ ($\lambda \in \mathbb{C}$, $|\lambda|$ μέτρο του μιγαδικού)

(iii) $\forall x, y \in X$, $\|x+y\| \leq \|x\| + \|y\|$

Παρατήρηση: Μια απεικόνιση που ικανοποιεί τις παραπάνω 3 ιδιότητες θα λαμβάνει θετικές τιμές, διότι

$$0 = \|x - x\| \leq \|x\| + \|(-1)x\| = 2\|x\|$$

Παρατήρηση: Για μια απεικόνιση, η οποία ικανοποιεί τις παραπάνω 3 ιδιότητες της νόρμας μπορούμε να δείξουμε ότι ισχύει

$$\forall\, x,y \in X, \quad \|x-y\| \geq \big|\, \|x\| - \|y\| \,\big|$$

Πράγματι,

$$\|x\| = \|(x-y)+y\| \leq \|x-y\| + \|y\| \Rightarrow \|x\| - \|y\| \leq \|x-y\|$$

και αντίστοιχα

$$\|y\| = \|(y-x)+x\| \leq \|y-x\| + \|x\| \Rightarrow \|y\| - \|x\| \leq \|y-x\| = \|x-y\|$$

Παραδείγματα: 1.) $(\mathbb{R}, \|\cdot\|)$, με $\|x\| = |x|$.

Είναι προφανές ότι η απόλυτη τιμή για τους πραγματικούς αριθμούς ικανοποιεί τις ιδιότητες της νόρμας.

2) $(\mathbb{R}^n, \|\cdot\|_1)$ με $\|x\|_1 = \sum_{i=1}^{n} |x_i|$, $x = (x_1, x_2, \ldots, x_n)^T$

α) $\|x\|_1 = 0 \iff \sum_{i=1}^{n} |x_i| = 0 \iff |x_i| = 0$, $i = 1, \ldots, n \iff x = 0$

β) $\|\lambda x\|_1 = \sum_{i=1}^{n} |\lambda x_i| = |\lambda| \sum_{i=1}^{n} |x_i| = |\lambda| \, \|x\|_1$, $x \in \mathbb{R}^n$

γ) $\|x+y\|_1 = \sum_{i=1}^{n} |x_i + y_i| \leq \sum_{i=1}^{n} (|x_i| + |y_i|) = \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i| = \|x\|_1 + \|y\|_1$

3) $(\mathbb{R}^n, \|\cdot\|_2)$ με $\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$

Εύκολα βλέπουμε ότι ισχύουν οι δύο πρώτες ιδιότητες της νόρμας
$\|x\|_2 = 0 \iff x = 0$ και $\|\lambda x\|_2 = |\lambda| \, \|x\|_2$

Για να δείξουμε την 3η ιδιότητα θα δείξουμε ένα ενδιάμεσο αποτέλεσμα.

Καταρχήν ορίζουμε την απεικόνιση $(\cdot,\cdot)_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$
ως $(x,y)_2 = \sum_{i=1}^{n} x_i y_i = x^T y$

Εύκολα βλέπουμε ότι $\|x\|_2 = \sqrt{(x,x)_2}$

Ανισότητα Cauchy-Schwarz

$$|(x,y)_2| \leq \|x\|_2 \|y\|_2 \quad , \forall x,y \in \mathbb{R}^n$$

Απόδειξη: για $\lambda \in \mathbb{R}$, $0 \leq \|x+\lambda y\|_2^2 = (x+\lambda y, x+\lambda y)_2$

Επίσης κάνοντας πράξεις εύκολα βλέπουμε ότι

$$(x+\lambda y, x+\lambda y)_2 = \sum_{i=1}^{n} (x_i + \lambda y_i)^2 = \underbrace{\|x\|_2^2 + 2\lambda (x,y)_2 + \lambda^2 \|y\|_2^2}_{p(\lambda)}$$

Το $p(\lambda) \geq 0$ για κάθε λ, άρα πρέπει η διακρίνουσα του

$\Delta \leq 0$ (διαφορετικά θα υπήρχαν δύο ρίζες λ_1, λ_2 και ανάμεσα τους η τιμή του $p(\lambda)$ θα ήταν αρνητική)

Η $\Delta = 4(x,y)_2^2 - 4\|x\|_2^2\|y\|_2^2 \leq 0$

οπότε $(x,y)_2^2 \leq \|x\|_2^2\,\|y\|_2^2$ και άρα ισχύει το ζητούμενο.

Επιστρέφουμε τώρα να δείξουμε ότι η $\|\cdot\|_2$ είναι μια νόρμα και μένει να δείξουμε την τρίτη ιδιότητα ότι ισχύει.

$$\|x+y\|_2^2 = (x+y, x+y)_2^2 = \sum_{i=1}^{n}(x_i+y_i)^2 = \|x\|_2^2 + 2(x,y)_2 + \|y\|_2^2$$

$$\leq \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2 = (\|x\|_2 + \|y\|_2)^2$$

Συνεπώς δείξαμε και την $3^{η}$ ιδιότητα και άρα η $\|\cdot\|_2$ είναι μια νόρμα στον $\mathbb{R}^n$

4) $(\mathbb{R}^n, \|\cdot\|_\infty)$ με $\|x\|_\infty = \max\limits_{1\leq i\leq n} |x_i|$, $x = (x_1, \dots, x_n)^T$

Εύκολα μπορούμε να δούμε ότι η $\|\cdot\|_\infty$ είναι μία νόρμα.

Στη συνέχεια θα δείξουμε ορισμένα βοηθητικά αποτελέσματα τα οποία θα χρησιμοποιήσουμε παρακάτω

<u>Ορισμός</u>: Λέμε ότι η ακολουθία $(x^{(k)})_{k\in\mathbb{N}} \subset X$ συγκλίνει ως προς τη νόρμα $\|\cdot\|$ του X, αν υπάρχει $x\in X$ τέτοιο ώστε

$$\|x^{(k)} - x\| \to 0 \ , \ k\to\infty$$

και γράφουμε $x^{(k)} \to x$, $k\to\infty$ ή $\lim\limits_{k\to\infty} x^{(k)} = x$

<u>Ορισμός</u>: Έστω X γραμμικός χώρος. Δύο νόρμες $\|\cdot\|$ και $\|\cdot\|'$ στον X λέγονται ισοδύναμες (ή συγκρίσιμες) αν υπάρχουν θετικές σταθερές m και M τέτοιες ώστε

$$\forall x \in X \quad m\|x\| \leq \|x\|' \leq M\|x\|$$

(οπότε φυσικά θα ισχύει και

$$\forall x \in X \quad \frac{1}{M}\|x\|' \leq \|x\| \leq \frac{1}{m}\|x'\| \quad)$$

<u>Λήμμα</u>: Κάθε νόρμα στον $\mathbb{R}^n$ είναι ισοδύναμη με τη νόρμα μεγίστου $\|\cdot\|_\infty$ του $\mathbb{R}^n$

<u>Απόδειξη</u>: Έστω $\{e^1, \dots, e^n\} \subset \mathbb{R}^n$, η κανονική βάση του $\mathbb{R}^n$,

$$e^1 = (1, 0, \dots, 0)^T, \; e^2 = (0, 1, 0, \dots, 0)^T, \; \dots, \; e^n = (0, \dots, 0, 1)^T$$

Τότε $\forall x \in \mathbb{R}^n$, $x = (x_1, \dots, x_n)^T$, $\quad x = \sum_{i=1}^{n} x_i e^i$

$$\|x\| = \Big\| \sum_{i=1}^{n} x_i e^i \Big\| \leq \sum_{i=1}^{n} |x_i| \, \|e^i\| \leq \|x\|_\infty \underbrace{\left(\sum_{i=1}^{n} \|e^i\| \right)}_{M} = M \, \|x\|_\infty$$

Όλες οι νόρμες είναι συνεχείς συναρτήσεις. Αυτό προκύπτει από την τριγωνική ανισότητα, διότι

$$\|x\| \leq \|x-y\| + \|y\| \Rightarrow \big|\, \|x\| - \|y\| \,\big| \leq \|x-y\| \leq M \|x-y\|_\infty .$$

Συνεπώς, αν $x \to y$, $\|x-y\|_\infty \to 0$, θα έχουμε ότι και $\|x\| \to \|y\|$

Λόγω της συνέχειας, έχουμε ότι αν $x \in S$, με S κλειστό και φραγμένο σύνολο από τον απειροστικό λογισμό γνωρίζουμε ότι μια συνεχής συνάρτηση θα λαμβάνει την ελάχιστη τιμή της.

Έστω $S = \{ x \in \mathbb{R}^n , \|x\|_\infty = 1 \}$ και $m = \min\limits_{x \in S} \|x\|$. Είναι προφανές ότι

$m > 0$ γιατί αν $m = 0$ τότε θα υπήρχε $x \in S$ τ.ω. $\|x\| = 0$, όμως $x = 0 \notin S$

$\forall x \in \mathbb{R}^n$, έχουμε ότι $\frac{1}{\|x\|_\infty} x \in S$, διότι $\left\| \frac{x}{\|x\|_\infty} \right\|_\infty = 1$.

άρα $m \leq \left\| \frac{x}{\|x\|_\infty} \right\| = \frac{1}{\|x\|_\infty} \|x\|$ ή $\underline{\underline{m\|x\|_\infty \leq \|x\|}}$

Πρόταση: Όλες οι νόρμες στον $\mathbb{R}^n$ είναι ισοδύναμες μεταξύ τους.

Απόδειξη: Έστω $\|.\|$, $\|.\|'$ δύο νόρμες στον $\mathbb{R}^n$. Υπάρχουν δύο ζευγάρια θετικών σταθερών τ.ω. κάθε μία από τις 2 νόρμες να είναι ισοδύναμη με τη νόρμα μεγίστου $\|.\|_\infty$.

Εύκολα τώρα βλέπουμε ότι και μεταξύ τους θα είναι ισοδύναμες.

$$\forall x \in \mathbb{R}^n, \quad \& \quad \left.\begin{array}{l} m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty \\ m'\|x\|_\infty \leq \|x\|' \leq M'\|x\|_\infty \end{array}\right\} \Rightarrow$$

$$\|x\| \leq M\|x\|_\infty \leq \frac{M}{m'}\|x\|'$$

και $\quad \|x\| \geq m\|x\|_\infty \geq \frac{m}{M'}\|x\|'$

<u>Παρατήρηση</u>: Συμπεραίνουμε λοιπόν ότι αν μια ακολουθία διανυσμάτων $(x^{(k)})_{k\in\mathbb{N}} \subset \mathbb{R}^n$ συγκλίνει ως προς μια νόρμα $\|\cdot\|$ του $\mathbb{R}^n$, θα συγκλίνει και ως προς μια οποιαδήποτε άλλη.

Συνεπώς και ως προς τη $\|\cdot\|_\infty$. Άρα αν $\lim\limits_{k\to\infty} x^{(k)} = x$ ως προς $\|\cdot\|$ τότε $\lim\limits_{k\to\infty} x_i^{(k)} = x_i$, $i=1,\ldots,n$

Νόρμες Πινάκων

Θα συμβολίζουμε και τις νόρμες πινάκων $\|\cdot\|$ και θα είναι καθαρό αν αναφερόμαστε σε νόρμα πίνακα ή διανύσματος από το όρισμα που θα χρησιμοποιούμε.

Μια νόρμα πινάκων $\|\cdot\| : \mathbb{R}^{n\times n} \to \mathbb{R}$ πληροί τις 3 ιδιότητες που θεωρήσαμε για τις νόρμες διανυσμάτων και επιπλέον και μια τέταρτη ιδιότητα

$$\|AB\| \le \|A\| \cdot \|B\|$$

<u>Έτσι</u>
(i) $\|A\| = 0$ αν και μόνο αν $A = 0$ (0 μηδενικός πίνακας)
(ii) $\|\lambda A\| = |\lambda| \, \|A\|$
(iii) $\|A+B\| \le \|A\| + \|B\|$
(iv) $\|AB\| \le \|A\| \, \|B\|$

Για κάθε νόρμα $\|.\|$ στο $\mathbb{R}^n$, μπορούμε να θεωρήσουμε μια νόρμα πινάκων, η οποία όπως θα λέγονται παράγεται από την νόρμα για διανύσματα.

Όλες οι νόρμες πινάκων οι οποίες παράγονται από νόρμα διανύσματος, λέγονται φυσικές νόρμες πινάκων

Ορισμός: Έστω $\|.\|$ μια νόρμα στον $\mathbb{R}^n$. Η απεικόνιση

$$\|.\| : \mathbb{R}^{n\times n} \to \mathbb{R}, \qquad \|A\| = \sup_{\substack{x\in\mathbb{R}^n \\ x\neq 0}} \frac{\|Ax\|}{\|x\|}$$

λέγεται φυσική νόρμα ή νόρμα πινάκων η οποία παράγεται από τη νόρμα $\|.\|$ του $\mathbb{R}^n$