10.0

Διάλεξη 10
(βίντεο)

ΜΕΜ 255 Θεωρία Προσέγγισης και Εφαρμογές
χε 2020
UoC

10.1

Αναδρομή Διάλεξης 9

1. $F=\mathbb{R}$ ή $\mathbb{C}$, $(F,V,+,\cdot)$ ορίσαμε εσωτερικό γινόμενο στο $(F,V,+,\cdot)$ $b: V\times V \to F$

$$b(v,w)=\overline{b(w,v)}, \quad b(v+w,z)=b(v,z)+b(w,z)$$

$$b(\lambda v,w)=\lambda\, b(v,w), \quad \forall v\in V \text{ με: } v\neq 0 \text{ ισχύει } b(v,v)>0$$

2. Συνέπειες ορισμού.

3. Θεώρημα (ανισότητα Cauchy-Schwarz)

$$|(x,y)| \leq \sqrt{(x,x)}\ \sqrt{(y,y)} \quad \forall x,y\in V$$

η ισότητα ισχύει ανν τα x,y γραμμικώς εξαρτημένα
$(\cdot,\cdot)$ είναι ένα εσωτερικό γινόμενο στο $(F,V,+,\cdot)$

Έστω:

Θεώρημα: $F = \mathbb{R}$ ή $\mathbb{C}$, $(F, V, +, \cdot, (\cdot,\cdot))$ ένα χώρος με εσωτερικό γινόμενο και $\|\cdot\| : V \to [0,+\infty)$ με $\|v\| = \sqrt{(v,v)}$ για κάθε $v \in V$. Τότε η $\|\cdot\|$ είναι μία νόρμα στον $(F, V, +, \cdot)$ η οποία επιπλέον είναι αυστηρά κυρτή. Σ' αυτή την περίπτωση λέμε ότι η $\|\cdot\|$ είναι η νόρμα που παράγεται από το εσωτερικό γινόμενο $(\cdot,\cdot)$.

Απόδειξη:

1. Αν $v=0$ τότε: $(v,v) = (0,0) = 0$, το οποίο συνεπάγεται ότι: $\|v\| = 0$. Αν για κάποιο $v \in V$ ισχύει $\|v\| = 0$, τότε $(v,v) = 0$, το οποίο σημαίνει ότι $v = 0$ (σημ. όταν $v \neq 0$ τότε από τον ορισμό του εσωτερικού γινομένου έπεται $(v,v) > 0$).

2. Έστω $\lambda \in F$ και $v \in V$. Τότε: $\|\lambda v\| = [(\lambda v, \lambda v)]^{1/2} = [\lambda (v, \lambda v)]^{1/2}$
$= [\lambda \bar{\lambda} (v,v)]^{1/2} = [|\lambda|^2 (v,v)]^{1/2} = |\lambda| \sqrt{(v,v)} = |\lambda| \|v\|$.

3. Έστω $v, w \in V$. Θα εξασφαλίσουμε στη συνέχεια την τριγωνική ανισότητα.

$$\|v+w\|^2 = (v+w, v+w) = (v,v) + (v,w) + (w,v) + (w,w)$$
$$= \|v\|^2 + \|w\|^2 + (v,w) + \overline{(v,w)}$$
$$= \|v\|^2 + \|w\|^2 + 2\,\mathrm{Re}(v,w).$$

3.1. $F = \mathbb{R}$

Έτσι: $\|v+w\|^2 = \|v\|^2 + \|w\|^2 + 2(v,w)$
$$\underset{C\text{-}S}{\leq} \|v\|^2 + \|w\|^2 + 2\|v\|\|w\| = (\|v\| + \|w\|)^2.$$

Άρα: $\|v+w\| \leq \|v\| + \|w\|$.

Ισότητα: $\|v+w\| = \|v\| + \|w\| \Leftrightarrow \|v+w\|^2 = (\|v\| + \|w\|)^2$
$\Leftrightarrow 2(v,w) = 2\|v\|\|w\| \Leftrightarrow (v,w) = \|v\|\|w\|$
$\Leftrightarrow |(v,w)| = \|v\|\|w\| \underset{\text{ισότητα στην CS}}{\Leftrightarrow} v, w$ γραμμ. εξαρτημένα

Αυτό σημαίνει ότι η $\|\cdot\|$ είναι αυστηρά κυρτή

3.2. $F=\mathbb{C}$

Τότε: $\|v+w\|^2 = \|v\|^2 + \|w\|^2 + 2\,\mathrm{Re}\,(v,w)$ $\qquad \mathrm{Re}\,z \le |\mathrm{Re}\,z| \le |z|$

$$\le \|v\|^2 + \|w\|^2 + 2\,|(v,w)|$$

$$\underset{CS}{\le} \|v\|^2 + \|w\|^2 + 2\,\|v\|\,\|w\|$$

$$\le (\|v\| + \|w\|)^2.$$ Άρα: $\|v+w\| \le \|v\| + \|w\|$.

Ισότητα: $\|v+w\| = \|v\| + \|w\|$

$\iff \|v+w\|^2 = (\|v\| + \|w\|)^2 \iff |(v,w)| = \|v\|\cdot\|w\|$

$\underset{\text{ισότητα στην CS}}{\iff}$ v, w είναι γραμμ. εξαρτημένα.

Έτσι η $\|\cdot\|$ είναι αυστηρά κυρτή. ∎

Παραδείγματα.

1. $F=\mathbb{R}$, $(\mathbb{R}, \mathbb{R}^m, +, \cdot)$, $m \in \mathbb{N}$, $(\cdot,\cdot)_2 : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$

$$(x,y)_2 = \sum_{i=1}^{m} x_i y_i \quad \forall x,y \in \mathbb{R}^m$$

Ευκλείδιο εσωτερικό γινόμενο

Η αντίστοιχη νόρμα είναι η Ευκλείδια νόρμα:

$$\|x\|_2 := \sqrt{(x,x)_2} = \left\{\sum_{i=1}^{m} x_i^2\right\}^{1/2} \quad \forall x \in \mathbb{R}^m$$

2. $F = \mathbb{R}$ ή $\mathbb{C}$, $(F, \mathbb{C}^m, +, \cdot)$ $\quad m \in \mathbb{N}$

$$(z,w)_2 = \sum_{i=1}^{m} z_i \bar{w}_i \quad \forall z,w \in \mathbb{C}^m$$

$$\|z\|_2 = \left(\sum_{i=1}^{m} |z_i|^2\right)^{1/2} \quad \forall z \in \mathbb{C}^m$$

3. $F=\mathbb{R}$ ή $\mathbb{C}$ $\quad (F, C([0,1],\mathbb{C}),+,\cdot)$

L2 εσ. γινόμενο $\quad (f,g)_2 := \int_0^1 f(t)\overline{g(t)}\,dt \quad \forall f,g\in C([0,1];\mathbb{C})$

L2 νόρμα $\quad \|f\|_2 = \left(\int_0^1 |f(t)|^2\, dt\right)^{\frac{1}{2}} \quad \forall f\in C([0,1];\mathbb{C})$

(4) Θεώρημα: (χαρακτηρισμός της βέλτιστης προσέγγισης)

Έστω: $F=\mathbb{R}$ ή $\mathbb{C}$, $X=(F,V,+,\cdot,(\cdot,\cdot))$ γραμμικός χώρος με εσωτερικό γινόμενο και $\|\cdot\|$ η νόρμα που παράγεται από το εσωτερικό γινόμενο.
Έστω ακόμα Y ένας υπόχωρος του V και $x\in V$. Τότε ισχύει το ακόλουθο:
ένα στοιχείο $y_*\in Y$ είναι βέλτιστη προσέγγιση του x από τον Y ανν
$(x-y_*,y)=0 \ \forall y\in Y$ ή $x-y_*\perp Y$.

Απόδειξη:

"$\Leftarrow$": Υποθέτουμε ότι υπάρχει $y_*\in Y$ τ.ω. $(x-y_*,y)=0 \ \forall y\in Y$.
Ο στόχος μας είναι να αποδείξουμε ότι: $\|x-y_*\|\le\|x-y\| \ \forall y\in Y$.
δηλ. $\|x-y_*\|=dist(x,Y)$.

Έστω $y\in Y$. Τότε:

$$\|x-y\| = \|(x-y_*)+(y_*-y)\| = \left\{\left((x-y_*)+(y_*-y),(x-y_*)+(y_*-y)\right)\right\}^{\frac{1}{2}}$$

$$= \left\{\|x-y_*\|^2+\|y_*-y\|^2+(y_*-y,x-y_*)+(x-y_*,y_*-y)\right\}^{\frac{1}{2}}$$

$$= \left\{\|x-y_*\|^2+\|y_*-y\|^2+2\,\mathrm{Re}\,(x-y_*,y_*-y)\right\}^{\frac{1}{2}}$$

$$= \left\{\|x-y_*\|^2+\|y_*-y\|^2+2\,\mathrm{Re}\left[\underbrace{(x-y_*,\overset{\in Y}{y_*})}_{=0}-\underbrace{(x-y_*,\overset{\in Y}{y})}_{=0}\right]\right\}^{\frac{1}{2}}$$

$$= \left(\|x-y_*\|^2+\|y_*-y\|^2\right)^{\frac{1}{2}} \ge \sqrt{\|x-y_*\|^2} = \|x-y_*\|.$$

10.8

"$\Rightarrow$" Υποθέτουμε ότι υπάρχει $y_* \in Y$ τ.ω. $\|x-y_*\| = \operatorname{dist}(x,Y)$.
Ο σκοπός μας είναι να αποδείξουμε ότι: $(x-y_*, y) = 0 \ \forall y \in Y$.
Η απόδειξη θα γίνει με απαγωγή σε άτοπο. Υποθέτουμε ότι υπάρχει $z_0 \in Y$ τ.ω. $(x-y_*, z_0) \neq 0$. (Φυσικά: $z_0 \neq 0$). Στη συνέχεια εισάγουμε μια βοηθητική συνάρτηση: $\varphi : F \to \mathbb{R}$ η οποία ορίζεται ως εξής: $\varphi(\lambda) = \|x-y_*-\lambda z_0\|^2 \ \forall \lambda \in F$.

Έστω $\lambda \in F$. Τότε:

$$\begin{aligned}\varphi(\lambda) &= (x-y_*-\lambda z_0, x-y_*-\lambda z_0)\\ &= \|x-y_*\|^2 + |\lambda|^2 \|z_0\|^2 - 2\operatorname{Re}[(x-y_*, \lambda z_0)]\\ &= \|x-y_*\|^2 + |\lambda|^2\|z_0\|^2 - 2\operatorname{Re}[\bar{\lambda}(x-y_*, z_0)] \geq 0.\end{aligned}$$

<u>Περίπτωση 1</u>: $F = \mathbb{R}$

Έτσι: $\varphi(\lambda) = \|x-y_*\|^2 + \lambda^2\|z_0\|^2 - 2\lambda(x-y_*, z_0) \quad \forall \lambda \in \mathbb{R}$.
Επιπλέον: $\varphi'(\lambda) = 2\lambda\|z_0\|^2 - 2(x-y_*, z_0) \quad \forall \lambda \in \mathbb{R}$.

10.9

Άρα: $\varphi'(\lambda) = 0 \Leftrightarrow \lambda\|z_0\|^2 = (x-y_*, z_0)$

$$\Leftrightarrow \lambda = \frac{(x-y_*, z_0)}{\|z_0\|^2}.$$

Επιπλέον: $\varphi''(\lambda) = 2\|z_0\|^2 > 0$. Το $\lambda_0 := \frac{(x-y_*, z_0)}{\|z_0\|^2}$ είναι σημείο ολικού ελαχίστου της φ στο $\mathbb{R}$, δηλ. $\varphi(\lambda) \geq \varphi(\lambda_0) \ \forall \lambda \in \mathbb{R}$. Επιπλέον το λ_0 είναι το μοναδικό σημείο ολικού ελαχίστου. Έτσι:

$$\varphi(\lambda) > \varphi(\lambda_0) \quad \forall \lambda \in \mathbb{R} \text{ με } \lambda \neq \lambda_0.$$

$\Rightarrow \quad \varphi(0) > \varphi(\lambda_0)$ καθώς: $0 \neq \lambda_0 \Leftarrow 0 \neq (x-y_*, z_0)$

$\Rightarrow \quad \|x-y_*\|^2 > \|x-(y_*+\lambda_0 z_0)\|^2$

$\Rightarrow \quad \underbrace{\|x-y_*\|}_{= \operatorname{dist}(x,Y)} > \|x - \underbrace{(y_*+\lambda_0 z_0)}_{\in Y}\| \longleftarrow$ Άτοπο

Επειδή Y υπόχωρος έπεται: $y_*+\lambda_0 \cdot z_0 \in Y$

Περίπτωση 2 : $F = \mathbb{C}$

$$\varphi(\lambda) = \|x-y_*\|^2 + |\lambda|^2 \|z_0\|^2 - 2\,\mathrm{Re}\left[\bar{\lambda}\,(x-y_*, z_0)\right] \quad \forall \lambda \in \mathbb{C}.$$

Υποπερίπτωση 1: $\mathrm{Re}\,(x-y_*, z_0) \neq 0$.

Παρατηρούμε ότι:

$$\varphi(t) = \|x-y_*\|^2 + t^2 \|z_0\|^2 - 2t\,\mathrm{Re}(x-y_*, z_0) \quad \forall t \in \mathbb{R}.$$

Όπως στην περίπτωση 1 στο $t_0 = \dfrac{\mathrm{Re}\,(x-y_*, z_0)}{\|z_0\|^2}$ η φ λαμβάνει ολικό ελάχιστο και μόνο εκεί. Άρα: $\varphi(t) > \varphi(t_0) \quad \forall t \in \mathbb{R},\ t \neq t_0$

$\Rightarrow \varphi(0) > \varphi(t_0)$ (καθώς $t_0 \neq 0 (\Leftrightarrow \mathrm{Re}(x-y_*, z_0) \neq 0)$).

$\Rightarrow \|x-y_*\| > \|x-y_*-t_0 z_0\|$ Άτοπο.

Υποπερίπτωση 2: $\mathrm{Re}\,(x-y_*, z_0) = 0$.

Επειδή έχουμε υποθέσει ότι: $(x-y_*, z_0) \neq 0$, αναγκαστικά: $\mathrm{Im}\,(x-y_*, z_0) \neq 0$. Στη συνέχεια ορίζουμε: $\psi: \mathbb{R} \to \mathbb{R}$ ως εξής: $\psi(t) = \varphi(\underline{i}t) \quad \forall t \in \mathbb{R}$. Έτσι:

$$\begin{aligned}\psi(t) &= \|x-y_*\|^2 + |it|^2 \|z_0\|^2 - 2\,\mathrm{Re}\left[\overline{t\,i}\,(x-y_*, z_0)\right] \\ &= \|x-y_*\|^2 + t^2 \|z_0\|^2 - 2t\,\mathrm{Re}\left[-i\,(x-y_*, z_0)\right] \\ &= \|x-y_*\|^2 + t^2 \|z_0\|^2 - 2t\,\mathrm{Im}\,(x-y_*, z_0) \quad \forall t \in \mathbb{R}.\end{aligned}$$

Όπως και στις προηγούμενες περιπτώσεις στο $t_0 = \dfrac{\mathrm{Im}\,(x-y_*, z_0)}{\|z_0\|^2}$, ώστε:

$\psi(t) > \psi(t_0) \quad \forall t \in \mathbb{R}$ με $t \neq t_0$

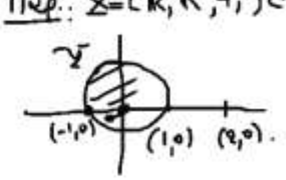
$\Rightarrow \psi(0) > \psi(t_0)$ (καθώς $t_0 \neq 0 (\Leftrightarrow \mathrm{Im}(x-y_*, z_0) \neq 0)$)

$\Rightarrow \|x-y_*\| > \|x-\underbrace{(y_*+t_0 z_0)}_{\in Y}\|$ Άτοπο. ∎

10.12

Παρατήρηση: Όταν αποδείξαμε το παράδοξο, δηλ. "⇐", δεν χρησιμοποιήσαμε την υπόθεση ότι ο Y είναι υπόχωρος.
Δηλ. αποδείξαμε ότι αν $Y \subset V$ και υπάρχει $y_* \in Y$ τ.ω. $(x-y_*, y) = 0$ $\forall y \in Y$ τότε: $\|x-y_*\| = dist(x,Y)$ δηλ. το y_* είναι η βέλτιστη προσέγγιση του x από το Y. Για το ευθύ "⇒" χρειαστήκαμε την υπόθεση ότι το Y υπόχωρος. Ένα λογικό ερώτημα είναι αν ο περιορισμός είναι αναγκαίος δηλ. αν γίνεται να έχουμε Y υποσύνολο του V, $x \in V$, y_* βέλτιστη προσέγγιση του x από το Y και $x - y_* \not\perp Y$.

Παρ.: $\mathbb{X} = (\mathbb{R}, \mathbb{R}^2, +, \cdot, (\cdot,\cdot)_2)$ $Y = \{z \in \mathbb{R}^2 : \|z\|_2 \leq 1\}$, $x = (2,0)$



Το Y είναι κλειστό και φραγμένο, το $\mathbb{R}^2$ πεπερασμένης διάστασης, άρα το Y συμπαγές. Άρα υπάρχει βέλτιστη προσέγγιση $y_* \in Y$ του x. Αυτή είναι μοναδική διότι $\|\cdot\|_2$ είναι αυστηρά κυρτή και το Y είναι κυρτό. Εδώ: $y_* = (1,0)$. Έτσι: $e = x - y_* = (1,0)$

10.13

Το ερώτημα είναι αν $e = (1,0)$ είναι κάθετο στο Y?

Απ. Όχι. Αν πάρουμε το $y = (-1,0)$. Τότε: $y \in Y$ και
$(e,y)_2 = 1 \cdot (-1) + 0 \cdot 0 = -1 \neq 0$.

Συμπέρασμα: $F = \mathbb{R}$ ή $\mathbb{C}$, $\mathbb{X} = (F, V, +, \cdot, (\cdot,\cdot))$ γραμμ. χώρος με εσωτερικό γινόμενο.
$x \in V$, Y υποχ. πεπ. διάστασης 1)

1. Επειδή ο Y έχει πεπερασμένη διάσταση υπάρχει βέλτιστη προσέγγιση $x_* \in Y$ του x από τον Y.
2. Επειδή Y κυρτό σύνολο και η νόρμα που προέρχεται από το εσωτερικό γινόμενο αυστηρά κυρτή, έπεται η βέλτιστη προσέγγιση $x_* \in Y$ του x από τον Y είναι μοναδική.
3. Μπορούμε να βρούμε το x_* χρησιμοποιώντας την (2) $(x - x_*, y) = 0$ $\forall y \in Y$ η οποία ισχύει επειδή ο Y είναι υπόχωρος.