Διάλεξη 16

(βίντεο)

ΜΕΜ255 ΘΕΩΡΙΑ ΠΡΟΣΕΓΓΙΣΗΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

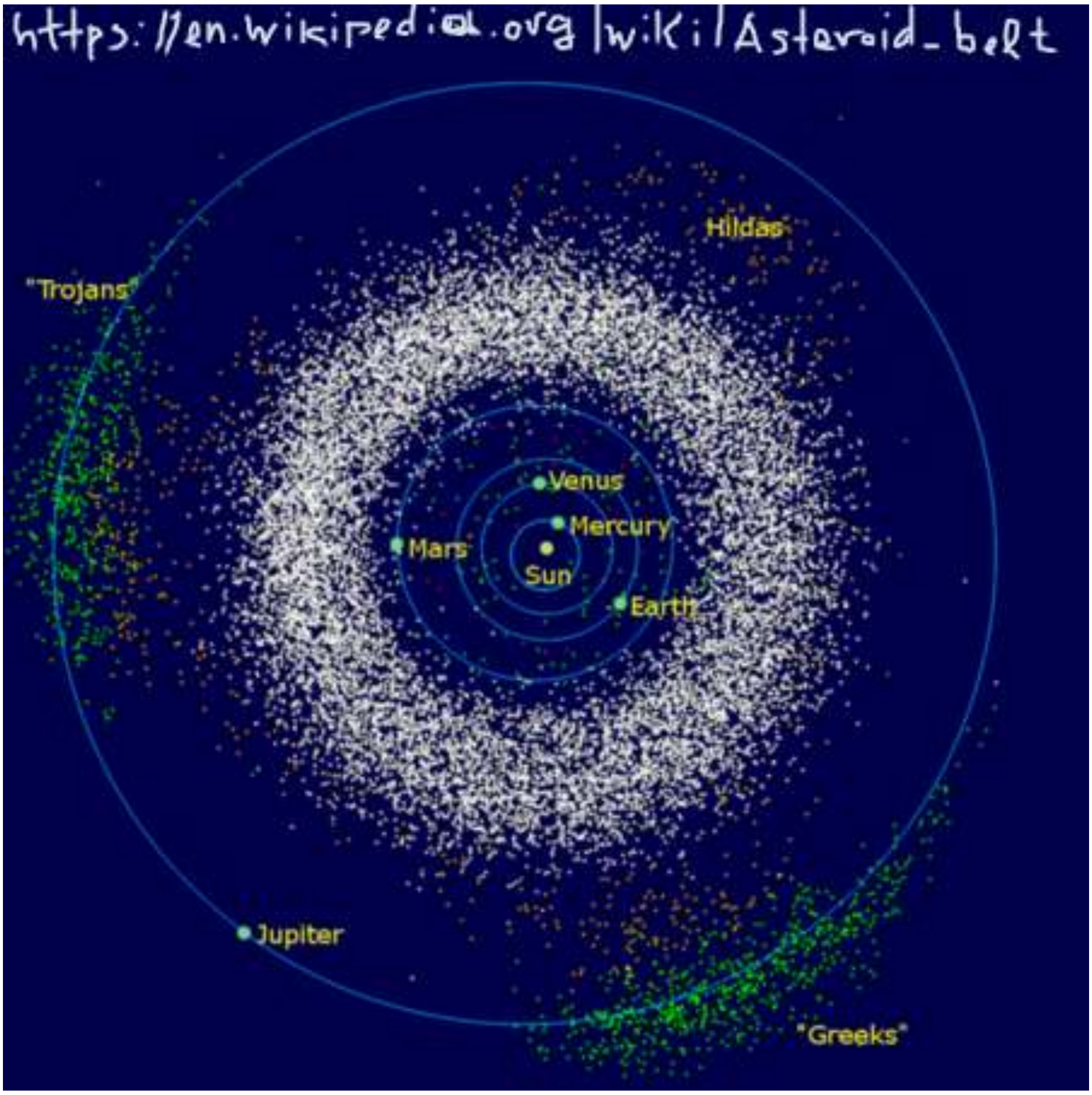
ΧΕ 2020

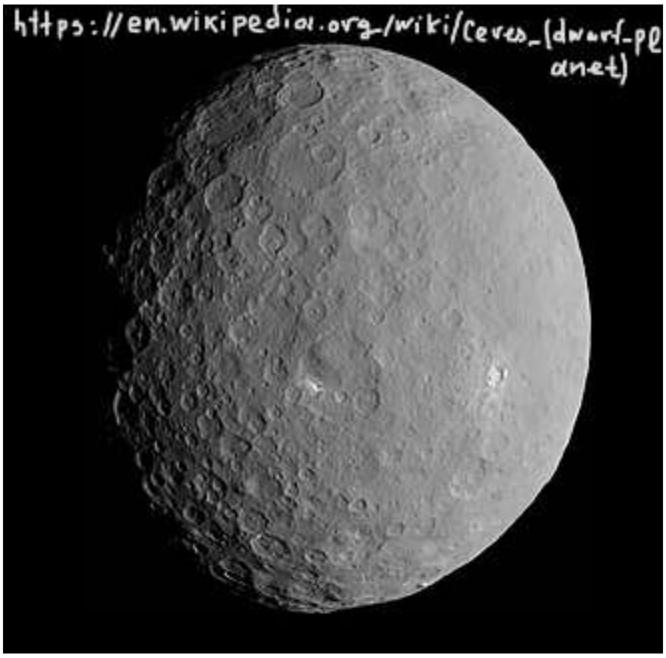
UoC

Η μέθοδος των ελαχίστων τετραγώνων

Δήμητρα

Ανάμεσα στον 'Αρη και το Δία υπάρχει ζώνη αστεροειδών της οποίας το μεγαλύτερο αντικείμενο είναι ο Ceres με διάμετρο 940 km περίπου και είναι ο μοναδικός πλανήτης νάνος στο ηλιακό σύστημα εντός της τροχιάς του Ποσειδώνα. Ανακαλύφθηκε από τον Giuseppe Piazzi το 1801 στο Αστεροσκοπείο του Παλέρμο και αρχικά θεωρήθηκε ότι ήταν ένας πλανήτης του ηλιακού συστήματος μέχρι που ανακαλύφθηκαν και άλλα παρόμοια σώματα όπως επανονομάστηκε ως αστεροειδή το 1850 περίπου. Μετά τη δημοσίευση της ανακάλυψης από τον Piazzi, λόγω της κίνησης της Γης ο Ceres μετατοπίστηκε πιο κοντά στον Ήλιο και ήταν αδύνατη η παρατήρηση του για αρκετούς μήνες από τους αστρονόμους που ήθελαν να επιβεβαιώσουν τις παρατηρήσεις του Piazzi. Το πρόβλημα που προέκυψε ήταν ότι δεν ήξεραν σε πιό


https://en.wikipedia.org/wiki/Asteroid_belt
Hildas
"Trojans"
Venus
Mercury
Mars
Sun
Earth
Jupiter
"Greeks"



σημείο θα γινόταν ορατός όταν απομακρυνόταν από τη λάμψη του Ήλιου. Τότε ο Gauss χρησιμοποιώντας τις 3 τελευταίες καταγεγραμμένες θέσεις του Ceres έκανε πρόβλεψη της τροχιάς του και αργότερα παρατηρήθηκε κοντά στο προβλεπόμενο σημείο. Το σκεπτικό πίσω από τη μέθοδο του Gauss οδήγησε σε αυτό που σήμερα λέμε μέθοδο ελαχίστων τετραγώνων η οποία έχει μεγάλη χρήση και εφαρμογή σε πολλά επιστημονικά πεδία. 16.2

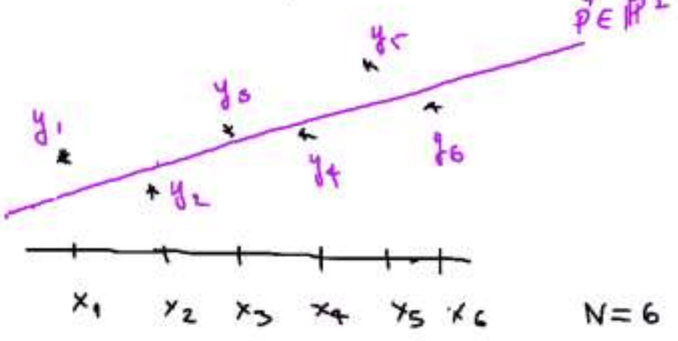
Στη συνέχεια θα παρουσιάσουμε αναλυτικά την πιο συνηθισμένη εκδοχή της μεθόδου ελαχίστων τετραγώνων.

<u>Πρόβλημα:</u> Έστω $N\in\mathbb{N}$ και δεδομένα $x\in\mathbb{R}^N$ και $y\in\mathbb{R}^N$, όπου: $x_i\neq x_j$ όταν $i\neq j$ και $1\le i,j\le N$. Έστω $m\in\mathbb{N}_0$ με: $m<N$. Αναζητούμε $\overset{*}{p}\in\mathbb{P}^m$ τ.ω.

$$\left[\sum_{i=1}^{N}\left|\overset{*}{p}(x_i)-y_i\right|^2\right]^{\frac{1}{2}}=\min_{q\in\mathbb{P}^m}\left[\sum_{i=1}^{N}\left|q(x_i)-y_i\right|^2\right]^{\frac{1}{2}}.$$

π.χ.

$m=1$



<u>Διατύπωση του προβλήματος ως πρόβλημα βέλτιστης προσέγγισης από υπόχωρο ενός γραμμικού χώρου</u>

Έστω: $X=(\mathbb{R},\mathbb{R}^N,+,\cdot,(\cdot,\cdot)_2)$ όπου: $(z,w)_2:=\sum_{i=1}^{N}z_iw_i$ για κάθε $z,w\in\mathbb{R}^N$. Στη συνέχεια ορίζουμε το σύνολο:

$$A^m=\{v\in\mathbb{R}^N:\ \text{υπάρχει } q\in\mathbb{P}^m \text{ τ.ω. } v_i=q(x_i) \text{ για } i=1,\dots,N\}$$

<u>Σημείωση:</u> Το A^m δεν είναι κενό διότι κάθε σταθερό διάνυσμα ανήκει στο A^m, δηλ. για κάθε $c\in\mathbb{R}$ έχουμε: $(c,\dots,c)^T\in A^m$, για κάθε $m\in\mathbb{N}_0$.

<u>Πρόταση:</u> Ο A^m είναι υπόχωρος του $\mathbb{R}^N$.

<u>Απόδειξη:</u> Έστω $\lambda\in\mathbb{R}$ και $v,w\in A^m$. Τότε υπάρχουν $q_1,q_2\in\mathbb{P}^m$ τ.ω. $v_i=q_1(x_i)$ και $w_i=q_2(x_i)$ για $i=1,\dots,N$. Έστω $p(t)=\lambda q_1+q_2\in\mathbb{P}^m$. Τότε $\lambda v+w\in A^m$ καθώς $(\lambda v+w)_i=p(x_i)$ για $i=1,\dots,N$. ∎

<u>Πρόταση</u>: Έστω $(q_\ell)_{\ell=0}^{m}$ μία βάση του $\mathbb{P}^m$ και:
$(z^\ell)_{\ell=1}^{m+1} \subset A^m$ με: $z_i^\ell := q_{\ell-1}(x_i)$ για $i=1,\dots,N$, $\ell=1,\dots,m+1$.
Τότε τα $(z^\ell)_{\ell=1}^{m+1}$ είναι γραμμικώς ανεξάρτητα και συνιστούν μία βάση του A^m.

<u>Απόδειξη</u>:

A. <u>Γραμμική ανεξαρτησία</u>:
Έστω $(\lambda_k)_{k=1}^{m+1} \subset \mathbb{R}$ τ.ω. $\sum_{k=1}^{m+1} \lambda_k \cdot z^k = 0$ ή ισοδύναμα
$\sum_{k=1}^{m+1} \lambda_k \, q_k(x_i) = 0$ για: $i=1,\dots,N$. Ας ορίσουμε:
$p \in \mathbb{P}^m$ με $p = \sum_{k=1}^{m+1} \lambda_k q_k$. Τότε $p(x_i)=0$ για $i=1,\dots,N$, δηλ. το p έχει $N>m$ ρίζες! Έτσι $p \equiv 0 \Leftrightarrow \lambda_k = 0$ για $k=1,\dots,m+1$.

B. Έστω $z \in A^m$. Τότε $z_i = q(x_i)$ για $i=1,\dots,N$, για κάποιο $q \in \mathbb{P}^m$. Προφανώς υπάρχουν $(\mu_i)_{i=0}^{m} \subset \mathbb{R}$ τ.ω. $q = \sum_{\ell=0}^{m} \mu_\ell q_\ell$.
Άρα: $z_i = \sum_{\ell=0}^{m} \mu_\ell \, q_\ell(x_i) = \sum_{\ell=0}^{m} \mu_\ell \, z_i^{\ell+1} = \sum_{\ell=1}^{m+1} \mu_{\ell-1} z_i^\ell$ για $i=1,\dots,N$, που σημαίνει ότι $z \in \operatorname{span}\{z^\ell : \ell=1,\dots,m+1\}$. ∎

<u>Σημ.</u>: $\dim A^m = m+1 = \dim \mathbb{P}^m$.

Συνεπώς το Πρόβλημα διατυπώνεται ισοδύναμα ως εξής:
βρες $z^* \in A^m$ τ.ω. $\|y - z^*\|_2 = \min_{z \in A^m} \|y - z\|_2$, όπου $\|\cdot\|_2$ η Ευκλείδια νόρμα η οποία παράγεται από το εσωτερικό γινόμενο $(\cdot,\cdot)_2$. Επειδή το σύνολο A^m είναι υπόχωρος του $\mathbb{R}^N$ με πεπερασμένη διάσταση, έχουμε ότι η βέλτιστη προσέγγιση $z^* \in A^m_{m+1}$ του y από το A^m υπάρχει και είναι μοναδική. Επομένως $z^* = \sum_{k=1}^{m+1} \lambda_k^* z^k$ και

και χρησιμοποιώντας τη συνθήκη ορθογωνιότητας:

$$(z^* - y, v)_2 = 0 \quad \forall v \in A^m$$

καταλήγουμε στο γραμμικό σύστημα:

$$\sum_{k=1}^{m+1} \lambda_k^* \, (z^k, z^\ell)_2 = (y, z^\ell)_2 \,, \quad \ell = 1, \dots, m+1 .$$

Εφαρμογή: Ας δούμε την περίπτωση όπου $m=1$. ($N > m=1$) Τότε μια βάση του P^1 μπορεί να είναι: $q_1(t)=1$ και $q_2(t)=t$. Άρα: $z^1 = (1,\dots,1)^T \in \mathbb{R}^N$ και $z^2 = (x_1, \dots, x_N)^T \in \mathbb{R}^N$ $\underset{=x}{}$. Το σύστημα των κανονικών εξισώσεων παίρνει τη μορφή:

$$\begin{bmatrix} (z^1,z^1)_2 & (z^1,z^2)_2 \\ (z^2,z^1)_2 & (z^2,z^2)_2 \end{bmatrix} \begin{bmatrix} \lambda_1^* \\ \lambda_2^* \end{bmatrix} = \begin{bmatrix} (y,z^1)_2 \\ (y,z^2)_2 \end{bmatrix}$$

$$\Leftrightarrow \quad \begin{bmatrix} N & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 \end{bmatrix} \begin{bmatrix} \lambda_1^* \\ \lambda_2^* \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N y_i x_i \end{bmatrix} .$$

Θέτοντας $c = (1,\dots,1)^T \in \mathbb{R}^N$, έπεται:

$$\lambda_1^* = \frac{(y,c)_2 \|x\|_2^2 - (x,c)_2 (x,y)_2}{N \|x\|_2^2 - (x,c)_2^2}$$

$$\lambda_2^* = \frac{N (x,y)_2 - (x,c)_2 (y,c)_2}{N \|x\|_2^2 - (x,c)_2^2}$$

$\alpha^2 - b^2 = (\alpha - b)(\alpha + b)$

ΠΡΟΣΟΧΗ στον παρονομαστή

Είναι προτιμότερο μερικές φορές να χρησιμοποιήσει κανείς απαλοιφή Gauss ή Cholesky.

και επομένως: $P^*(t) = \lambda_1^* + \lambda_2^* t$.

16.9

Σύνδεση με το γραμμικό πρόβλημα ελ. τετραγώνων.

Το Γ.Π.Ε.Τ. έχει τη μορφή: βρες $x^* \in \mathbb{R}^{n_2}$ τ.ω.

$$\|Ax^* - b\|_2 = \min_{x \in \mathbb{R}^{n_2}} \|Ax - b\|_2$$

όπου: $A \in \mathbb{R}^{n_1 \times n_2}$, $b \in \mathbb{R}^{n_1}$, $\|.\|_2$ η Ευκλείδια νόρμα στον $\mathbb{R}^{n_1}$ και n_1, n_2.

Το Πρόβλημα μπορούμε να το γράψουμε ως εξής:

$$\min_{q \in \mathbb{P}^m} \left\{ \sum_{i=1}^{N} (y_i - q(x_i))^2 \right\}^{1/2} = \min_{\lambda \in \mathbb{R}^{m+1}} \left\{ \sum_{i=1}^{N} \left(y_i - \sum_{j=1}^{m+1} \lambda_j q_j(x_i)\right)^2 \right\}^{1/2}$$

$$= \min_{\lambda \in \mathbb{R}^{m+1}} \|A\lambda - y\|_2 \quad \text{όπου: } A_{ij} = q_j(x_i), \text{ για: } \begin{array}{l} i = 1, \dots, N \\ j = 1, \dots, m+1 \end{array}$$

16.10

$\to (z^k)_{k=1}^{m+1}$ βάση του $\mathbb{P}^m$

Προφανώς: $A \in \mathbb{R}^{N \times (m+1)}$ και $A = [z^1 | \cdots | z^{m+1}]$, δηλ. ο A έχει γραμμικώς ανεξάρτητες στήλες (full rank). Το Γ.Π.Ε.Τ. έχει πάντα λύση η οποία δεν είναι μοναδική όταν ο A δεν έχει γραμμ. ανεξάρτητες στήλες. Η γραμμική ανεξαρτησία των στηλών του A εξασφαλίζεται από την υπόθεση $m < N$ και τις ιδιότητες των πολυωνύμων.

<u>ΕΡΩΤΗΣΗ</u>: Μπορώ να έχω κάποιο άλλο χώρο συναρτήσεων στη θέση του $\mathbb{P}^m$?

<u>ΑΠΑΝΤΗΣΗ</u>: ΝΑΙ. Όμως για να πάρουμε ανάλογα αποτελέσματα χρειαζόμαστε την ισχύ της συνθήκης του Haar, η οποία έχει ως εξής:

Συνθήκη Haar.

16.11

Έστω $(g_i)_{i=1}^{m} \subset C([a,b];\mathbb{R})$ οι οποίες είναι γραμμικώς ανεξάρτητες. Λέμε ότι ο χώρος $H=\operatorname{span}\{g_1,\dots,g_m\}$ ικανοποιεί τη συνθήκη του Haar όταν κάθε μηδενικό $g\in H$ έχει το πολύ $m-1$ διαφορετικές ρίζες.

↳ $\dim H$

<u>Παράδειγμα 1:</u> ο $\mathbb{P}^m[a,b]$ ικανοποιεί τη συνθήκη του Haar, καθώς κάθε μη μηδενικό $p\in\mathbb{P}^m[a,b]$ έχει το πολύ m ρίζες και $m=\dim\mathbb{P}^m-1$

<u>Παράδειγμα 2:</u> Έστω $m\in\mathbb{N}$ και $g_j(t)=e^{(j-1)t}$ για $j=1,\dots,m$.
Οι $(g_j)_{j=1}^{m}$ ικανοποιούν τη συνθήκη Haar. Πράγματι, έστω ότι η $g(t)=\sum_{j=1}^{m}\mu_j g_j(t)\in\operatorname{span}\{g_1,\dots,g_m\}$ έχει m διαφορετικές ρίζες $(t_k)_{k=1}^{m}$.
Τότε το πολυώνυμο $p(t)=\sum_{j=1}^{m}\mu_j t^{j-1}\in\mathbb{P}^{m-1}$ έχει m διαφορετικές ρίζες $(e^{t_k})_{k=1}^{m}$.
Άρα: $p\equiv 0$, δηλαδή $\mu_j=0$ για $j=1,\dots,m$. ∎