Γραμμική Άλγεβρα

Ενότητα **4** : Ορθογωνιότητα

Ευστράτιος Γαλλόπουλος
Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης
ΕΣΠΑ 2007-2013
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Πατρών**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σκοπός Ενότητας

- Ορθογωνιότητα των 4ων υποχώρων
- Προβολές
- Ορθογώνιες βάσεις και μέθοδος Gram-Schmidt
- Προβλήματα γραμμικών ελαχίστων τετραγώνων

Περιεχόμενα

1 Υπενθύμιση

Υπενθύμιση και πρόγραμμα διάλεξης

Στην προηγούμενη διάλεξη μιλήσαμε για ορισμένες χρήσεις μητρώων και διανυσμάτων.

- Γενική θεώρηση προβλημάτων ελαχιστοποίησης
- Προβλήματα και προσεγγίσεις ελαχίστων τετραγώνων

Σήμερα θα συζητήσουμε:

1. Παραδείγματα προσεγγίσεων ελαχίστων τετραγώνων.
2. Ελάχιστα τετράγωνα μέσω διαφορικού λογισμού.
3. Ελάχιστα τετράγωνα σε διανυσματικούς χώρους συναρτήσεων.

Επιχειρηματικό Κίνητρο: Βραβείο Netflix $1,000,000 I

With Netflix you can rent as many DVDs as you want and watch movies instantly on your PC for one low price (...) no late fees (...) no due dates, and DVD shipping free both ways. Plans start from only \$4.99 (...) With our most popular plan, you can rent as many DVDs as you want (3 at-a-time) (...) all for just \$17.99 a month +tax. ...

Επιχειρηματικό Κίνητρο: Βραβείο Netflix $1,000,000 II

Περιγραφή (Wikipedia)

- The Netflix Prize was an open competition for the best collaborative filtering algorithm to predict user ratings for films, based on previous ratings without any other information about the users or films, i.e. without the users or the films being identified ...
- Netflix provided a training data set of 100,480,507 ratings that 480,189 users gave to 17,770 movies. Each training rating is a quadruplet of the form <user, movie, date of grade, grade>. The user and movie fields are integer IDs, while grades are from 1 to 5 (stars).
- Prizes were based on improvement over Netflix's own algorithm, called Cinematch, or the previous year's score if a team has made improvement beyond a certain threshold. A trivial algorithm that predicts for each movie in the quiz set its average grade from the training data produces an RMSE of 1.0540. Cinematch uses "straightforward statistical linear models with a lot of data conditioning".

Επιχειρηματικό Κίνητρο: Βραβείο Netflix $1,000,000 III

The BigChaos Solution to the Netflix (

Andreas Töscher and Michael Jahre

commendo research & consulting
Neuer Weg 23, A-8580 Köflach, Austr
{andreas.toescher,michael.jahrer}@comme

Robert M. Bell*

AT&T Labs - Research
Florham Park, NJ

September 24, 2009

Input: A matrix $\mathbf{P}$ with all previous probe predictions. $\mathbf{P}$ always includes a constant predictor (a column with ones).

Exclude the Probe ratings $\mathbf{r}$ from the training set.
Initialize the weights.
$RMSE_{best} = \infty$
$RMSE_{epoch} = 1000$
$epochs = 0$
while $RMSE_{epoch} \le RMSE_{best}$ **do**
  Train one epoch.
  if $RMSE_{epoch} \le RMSE_{best}$ **then**
    $RMSE_{best} = RMSE_{epoch}$
    Save the current weights.
  end
  Predict the probe set $\hat{\mathbf{p}}$.
  Merge current probe prediction $\hat{\mathbf{p}}$ and previous predictions: $\mathbf{X} = [\mathbf{P}\ \hat{\mathbf{p}}]$
  Calculate blending weights: $\mathbf{w} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{r}$
  Calculate prediction of the current blend: $\mathbf{p} = \mathbf{X} \cdot \mathbf{w}$
  ~~Calculate the RMSE of the blend: $RMSE_{epoch} = \sqrt{\frac{1}{R}\sum_{i=1}^{R}(p_i - r_i)^2}$; r_i is~~ probe rating i, R is #ratings in the probe set
  $epochs = epochs + 1$
end
Load the weights.
Generate predictions for the probe set.
Insert the probe set into the training set.
Initialize the weights (use the same random seed as for the first initialization).
while $epochs > 0$ **do**
  Train one epoch

Συμμετρικά θετικά (ημι)ορισμένα μητρώα

Ενδιαφέρον χαρακτηριστικό π.χ. στο μητρώο των κανονικών εξισώσεων

$$(A^\top A)x = A^\top b.$$

Συμμετρικά θετικά (ημι)ορισμένα μητρώα

Ενδιαφέρον χαρακτηριστικό π.χ. στο μητρώο των κανονικών εξισώσεων

$$(A^\top A)x = A^\top b.$$

Το $B = A^\top A$:

- είναι συμμετρικό
- ισχύει πάντα ότι

$$x^\top A^\top A x \geq 0$$

- όταν οι στήλες του A είναι γραμμικά ανεξάρτητες, $x^\top A^\top A x > 0$, είναι αυστηρά θετικό.

Συμμετρικά θετικά (ημι)ορισμένα μητρώα

Ενδιαφέρον χαρακτηριστικό π.χ. στο μητρώο των κανονικών εξισώσεων

$$(A^\top A)x = A^\top b.$$

Το $B = A^\top A$:

- είναι συμμετρικό
- ισχύει πάντα ότι

$$x^\top A^\top Ax \geq 0$$

- όταν οι στήλες του A είναι γραμμικά ανεξάρτητες, $x^\top A^\top Ax > 0$, είναι αυστηρά θετικό.

Ορισμός

Κάθε $A \in \mathbb{R}^{n\times n}$ για το οποίο ισχύει ότι $A^\top = A$ και $f(x) := x^\top Ax \geq 0$ για κάθε $x \in \mathbb{R}^n$ και $f(x) = 0$ ανν $x = 0$ αποκαλείται **συμμετρικό θετικά ορισμένο μητρώο (ΣΘΟ)**. Αν υπάρχει $x \neq 0$ ώστε $f(x) = 0$ τότε αποκαλείται **συμμετρικό θετικά ημιορισμένο μητρώο**. Αν λαμβάνει και αρνητικές τιμές, αποκαλείται **αόριστο**.

Συμμετρικά θετικά (ημι)ορισμένα μητρώα

Ενδιαφέρον χαρακτηριστικό π.χ. στο μητρώο των κανονικών εξισώσεων

$$(A^\top A)x = A^\top b.$$

Το $B = A^\top A$:

- είναι συμμετρικό
- ισχύει πάντα ότι

$$x^\top A^\top Ax \geq 0$$

- όταν οι στήλες του A είναι γραμμικά ανεξάρτητες, $x^\top A^\top Ax > 0$, είναι αυστηρά θετικό.

Ορισμός

Κάθε $A \in \mathbb{R}^{n\times n}$ για το οποίο ισχύει ότι $A^\top = A$ και $f(x) := x^\top Ax \geq 0$ για κάθε $x \in \mathbb{R}^n$ και $f(x) = 0$ ανν $x = 0$ αποκαλείται **συμμετρικό θετικά ορισμένο μητρώο (ΣΘΟ)**. Αν υπάρχει $x \neq 0$ ώστε $f(x) = 0$ τότε αποκαλείται **συμμετρικό θετικά ημιορισμένο μητρώο**. Αν λαμβάνει και αρνητικές τιμές, αποκαλείται **αόριστο**.

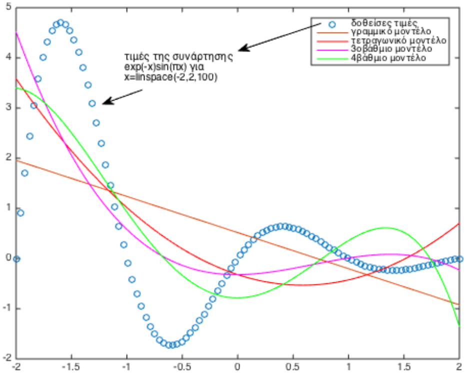
Επομένως, το $A^\top A$ είναι ΣΘΟ ανν το A έχει γραμμικά ανεξάρτητες στήλες, διαφορετικά είναι θετικά ημιορισμένο. Δεν μπορεί όμως να είναι αόριστο.

Παραδείγματα προσέγγισης ελαχίστων τετραγώνων

Χρησιμοποιούμε τον παρακάτω κώδικα MATLAB για διάφορες επιλογές `x, y`

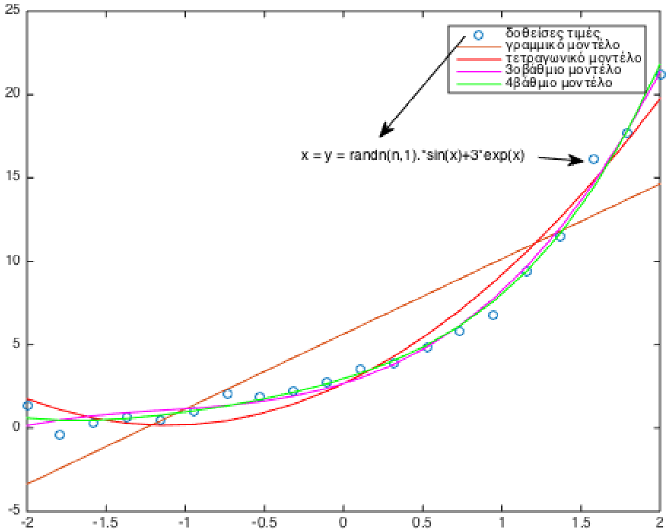
```
% kataskeu'h twn mhtr'wwn
A1 = (ones(n,1),x); % grammik'o mont'elo
A2 = (A1,x.^2); % tetragwnik'o mont'elo
A3 = (A2,x.^3); % 3b'ajmio mont'elo
A4 = (A3,x.^4); % 4b'ajmio mont'elo
% ep'ilush kanonik'wn exis'wsewn
z1 = (A1'*A1)\(A1'*y); z2 = (A2'*A2)\(A2'*y);
z3 = (A3'*A3)\(A3'*y); z4 = (A4'*A4)\(A4'*y);
% optikopo'ihsh
plot(x,y,'o',x,A1*z1,x,A2*z2,'r',x,A3*z3,'m',x,A4*z4,'g')
% sf'almata
nerr(1) = norm(y-A1*z1); nerr(2) = norm(y-A2*z2);
nerr(3) = norm(y - A3*z3); nerr(4) = norm(y - A4*z4);
```

`n=100; x=linspace(-2,2,n)'; y = exp(-x).*sin(pi*x)`



Σφάλματα: `nerr(1:4)= 14.4664 12.3761 11.7957 11.0530`

```
n=20;  x = linspace(-2,2,n)';
y = randn(n,1).*sin(x)+3*exp(x)
```



Σφάλματα: `nerr(1:4)= 12.7322 4.5668 2.7300 2.4798`

Βέλτιστη προσέγγιση συναρτήσεων: Εισαγωγή I

- Γνωρίζουμε ότι πολλές κατηγορίες συναρτήσεων αποτελούν δ.χ.
- Παραδείγματα: α) $\mathbb{P}_n$: τα πολυώνυμα βαθμού $\leq n$. (ο δ.χ. έχει διάσταση $n+1$). β) $C[-1,1]$: χώρος συνεχών συναρτήσεων στο κλειστό διάστημα $[-1,1]$. γ) Τα πολυώνυμα οποιουδήποτε βαθμού (οι δ.χ. είναι απειροδιάστατοι).
- Για ολοκληρώσιμες συναρτήσεις f, g στο διάστημα $[a, b]$ και θετική συνάρτηση w, το ορισμένο ολοκλήρωμα

$$\langle f, g \rangle := \int_a^b f(t)g(t)w(t)dt$$

  ορίζει εσωτερικό γινόμενο.
- Επομένως μπορούμε να ορίσουμε και καθετότητα συναρτήσεων. Για παράδειγμα, αν $w(x) = 1$, οι συναρτήσεις $\{1, x, \frac{1}{2}(3x^2 - 1)\}$ είναι κάθετες μεταξύ τους. Επίσης είναι γραμμικά ανεξάρτητες και αποτελούν βάση για τον $\mathbb{P}_2[-1,1]$ που είναι υπόχωρος του $C[-1,1]$.

Βέλτιστη προσέγγιση συναρτήσεων: Εισαγωγή II

- Μπορούμε να εφαρμόσουμε τη διαδικασία Gram-Schmidt σε πολυώνυμα και να κατασκευάσουμε Ορθογώνια Πολυώνυμα (σε κάποιο διάστημα και ως προς δεδομένη συνάρτηση w.)

Ερώτημα Δίνεται συνάρτηση f από κάποιο χώρο συναρτήσεων και θέλουμε να βρούμε βέλτιστη προσέγγισή της από κάποιον επιλεγμένο γραμμικό υπόχωρο, π.χ. τα πολυώνυμα, με βάση το παραπάνω εσωτερικό γινόμενο.

Επίλυση όπως πριν, θα αξιοποιήσουμε το ότι η διαφορά της υπό προσέγγιση συνάρτησης από τη βέλτιστη προσέγγιση θα είναι κάθετη σε όλες τις συναρτήσεις του υποχώρου απ' όπου «χτίζουμε» την προσέγγιση.

Ιδέα

Από το θεώρημα προβολής, αν αφαιρέσουμε από τη συνάρτηση το γραμμικό συνδυασμό που παράγει τη συνάρτηση βέλτιστης προσέγγισης, το κατάλοιπο θα είναι κάθετο σε όλες τις συναρτήσεις ϕ_j, επομένως $\langle f - \sum_{j=1}^{n} \gamma_j \phi_j, \phi_i \rangle = 0$ για $i = 1, \ldots, n$. Επομένως

$$\begin{pmatrix} \langle f, \phi_1 \rangle \\ \vdots \\ \langle f, \phi_n \rangle \end{pmatrix} = \underbrace{\begin{pmatrix} \langle \phi_1, \phi_1 \rangle & \cdots & \langle \phi_1, \phi_n \rangle \\ \vdots & & \vdots \\ \langle \phi_n, \phi_1 \rangle & \cdots & \langle \phi_n, \phi_n \rangle \end{pmatrix}}_{\text{Μητρώο Gram}} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}.$$

Παράδειγμα

Να υπολογιστεί η βέλτιστη προσέγγιση της συνάρτησης x^2 από τον υπόχωρο $\mathrm{span}\langle 1, x\rangle$ με εσωτερικό γινόμενο $\langle f, g\rangle := \int_{-1}^{1} f(x)g(x)dx$. Έστω $\phi(x) = \gamma_1 + \gamma_2 x$ η ζητούμενη προσέγγιση, τότε αν $r(x) = x^2 - \phi(x)$ θα πρέπει από το θεώρημα ορθογωνιότητας $\langle r, 1\rangle = \langle r, x\rangle = 0$. Επομένως

$$\begin{aligned}
\langle x^2, 1\rangle &= \int_{-1}^{1} x^2 dx = \frac{2}{3} \\
\langle x^2, x\rangle &= \int_{-1}^{1} x^3 dx = 0 \\
\langle 1, 1\rangle &= 2, \langle 1, x\rangle = 0 \\
\langle x, 1\rangle &= 0, \langle x, x\rangle = \frac{2}{3}
\end{aligned}$$

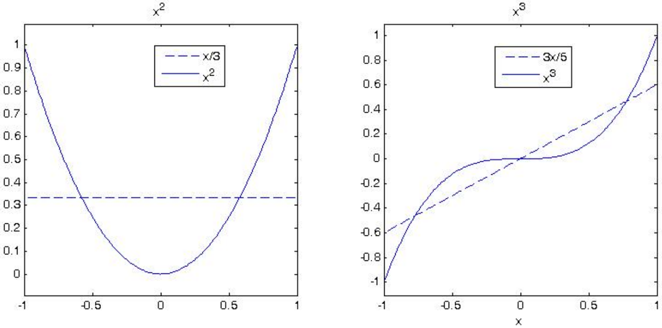
$$\begin{pmatrix} \frac{2}{3} \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}. \Rightarrow \gamma_1 = \frac{1}{3}, \gamma_2 = 0.$$

Άρα η βέλτιστη προσέγγιση είναι $\phi(x) = \frac{1}{3}$.
Προσέξτε επίσης ότι $\langle 1, x \rangle = 0$ δηλ, οι συναρτήσεις $\phi_1(x) = 1$ και $\phi_2(x) = x$ είναι κάθετες ως προς το παραπάνω εσωτερικό γινόμενο.
Προσοχή: Με τον ίδιο τρόπο αλλά λύνοντας το σύστημα

$$\begin{pmatrix} 0 \\ \frac{2}{5} \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}. \Rightarrow \gamma_1 = 0, \gamma_1 = \frac{3}{5}, \gamma_3 = 0.$$

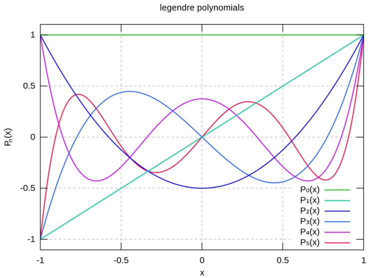
προκύπτει ότι η βέλτιστη προσέγγιση του x^3 από το χώρο πολυωνύμων $\langle 1, x, x^2 \rangle$ ως προς παραπάνω εσωτερικό γινόμενο είναι $\phi(x) = \frac{3}{5}x$.

Οπτικοποίηση των βέλτιστων προσεγγίσεων



Ορθογώνια πολυώνυμα Legendre

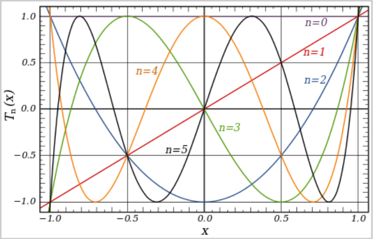
n	$P_n(x)$
0	1
1	x
2	$\frac{1}{2}(3x^2-1)$
3	$\frac{1}{2}(5x^3-3x)$
4	$\frac{1}{8}(35x^4-30x^2+3)$
5	$\frac{1}{8}(63x^5-70x^3+15x)$
6	$\frac{1}{16}(231x^6-315x^4+105x^2-5)$
7	$\frac{1}{16}(429x^7-693x^5+315x^3-35x)$
8	$\frac{1}{128}(6435x^8-12012x^6+6930x^4-1260x^2+35)$
9	$\frac{1}{128}(12155x^9-25740x^7+18018x^5-4620x^3+315x)$
10	$\frac{1}{256}(46189x^{10}-109395x^8+90090x^6-30030x^4+3465x^2-63)$



- Παράγονται με διαδικασία Gram-Schmidt επί των $\{1, x, x^2, \ldots\}$ με το εσωτερικό γινόμενο $\int_{-1}^{1} f(x)g(x)\mathrm{d}x$
- Ισχύει ότι $\int_{-1}^{1} P_n(x)P_m(x)dx = \frac{2}{2n+1}\delta_{nm}$ (δέλτα Kronecker)
- Υπάρχουν πολλές οικογένειες ορθογωνίων πολυωνύμων (π.χ. Chebyshev polynomials, Gegenbauer

Ορθογώνια πολυώνυμα Chebyshev

$$
\begin{aligned}
T_0(x) &= 1\\
T_1(x) &= x\\
T_2(x) &= 2x^2 - 1\\
T_3(x) &= 4x^3 - 3x\\
T_4(x) &= 8x^4 - 8x^2 + 1\\
T_5(x) &= 16x^5 - 20x^3 + 5x\\
T_6(x) &= 32x^6 - 48x^4 + 18x^2 - 1\\
T_7(x) &= 64x^7 - 112x^5 + 56x^3 - 7x\\
T_8(x) &= 128x^8 - 256x^6 + 160x^4 - 32x^2 + 1\\
T_9(x) &= 256x^9 - 576x^7 + 432x^5 - 120x^3 + 9x.\\
T_{10}(x) &= 512x^{10} - 1280x^8 + 1120x^6 - 400x^4 + 50x^2 - 1.\\
T_{11}(x) &= 1024x^{11} - 2816x^9 + 2816x^7 - 1232x^5 + 220x^3 - 11x.
\end{aligned}
$$



The first few Chebyshev polynomials of the first kind in the domain $-1 < x < 1$: The flat T_0, T_1, T_2, T_3, T_4 and T_5.

- Παράγονται με διαδικασία Gram-Schmidt επί των $\{1, x, x^2, \ldots\}$ με το εσωτερικό γινόμενο $\int_{-1}^{1} f(x)g(x)\frac{1}{\sqrt{1-x^2}}\mathrm{d}x$
- Ισχύει ότι $\int_{-1}^{1} T_n(x)T_m(x)dx = 0$ αν $m \neq n$.
- Υπάρχουν πολλές οικογένειες ορθογωνίων πολυωνύμων (π.χ. Chebyshev polynomials, Gegenbauer

Προσέγγιση ελαχίστων τετραγώνων μέσω διαφορικού λογισμού

Για το γραμμικό μοντέλο

Δοθέντων m ζευγών τιμών $\{(\xi_j,\psi_j)\}_{j=1}^m$ (θεωρούμε ότι τα ξ δεν είναι όλα ίσα μεταξύ τους) ενδιαφερόμαστε να υπολογίσουμε τις τιμές των παραμέτρων γ_1,γ_2 ώστε να ελαχιστοποιείται η συνάρτηση

$$E(\gamma_1,\gamma_2)=\sum_{j=1}^{m}(\psi_j-\gamma_2\xi_j-\gamma_1)^2$$

Από βασική θεωρία, για να υπάρχει τοπικό ακρότατο, πρέπει να ικανοποιούνται οι ακόλουθες σχέσεις (κριτήριο 1ης παραγώγου):

$$\begin{aligned}0 &= \frac{\partial E}{\partial \gamma_1}=2\sum_{j=1}^{m}(\psi_j-\gamma 2\xi_j-\gamma_1)(-1)\\ 0 &= \frac{\partial E}{\partial \gamma_2}=2\sum_{j=1}^{m}(\psi_j-\gamma_2\xi_j-\gamma_1)(-\xi_j)\end{aligned}$$

Τα τοπικά ακρότατα είναι στις τιμές $\tilde{\gamma}_1,\tilde{\gamma}_2$ που ικανοποιούν το σύστημα

$$\sum_{j=1}^{m}\psi_j = m\gamma_1+(\sum_{j=1}^{m}\xi_j)\gamma_2$$

Αν η λύση είναι $(\tilde{\gamma}_1, \tilde{\gamma}_2)$, για να ελαχιστοποιεί το $E(\tilde{\gamma}_1, \tilde{\gamma}_2)$, θα πρέπει (κριτήριο 2ης παραγώγου)

$$\frac{\partial^2 E}{\partial \gamma_1^2} > 0 \text{ και } \Delta = \frac{\partial^2 E}{\partial \gamma_1^2}\frac{\partial^2 E}{\partial \gamma_2^2} - \frac{\partial^2 E}{\partial \gamma_1 \partial \gamma_2} > 0 \frac{\partial^2 E}{\partial \gamma_1^2} > 0.$$

Η πρώτη ανισότητα προφανώς ικανοποιείται:

$$\frac{\partial^2 E}{\partial \gamma_1^2} = 2m > 0$$

Σχετικά με τη δεύτερη: Υπολογίζουμε

$$\frac{\partial^2 E}{\partial \gamma_2^2} = 2\sum_{j=1}^{m} \xi_j^2, \quad \frac{\partial^2 E}{\partial \gamma_1 \partial \gamma_2} = 2\sum_{j=1}^{m} \xi_j.$$

άρα

$$\Delta > 0 \Leftrightarrow m\sum_{j=1}^{m} \xi_j^2 > (\sum_{j=1}^{m} \xi_j)^2.$$

Αλλά $m = e^\top e, x^\top x = \sum_{j=1}^m \xi_j^2$ και $e^\top x = \sum_{j=1}^m \xi$. Επομένως η ανισότητα

$$\sum_{j=1}^m \xi_j^2 > (\sum_{j=1}^m \xi_j)^2$$

μπορεί να γραφτεί ως

$$\|e\|^2 \|x\|^2 > \langle e, x\rangle|^2.$$

Θυμηθείτε τώρα την ανισότητα Cauchy-Schwarz. Για κάθε $z, x, \in \mathbb{R}^m$

$$\langle z, x\rangle| \leq \|z\| \|x\|$$

όπου ισότητα ισχύει μόνον αν τα z, x είναι συγγραμμικά, δηλ. $z = \alpha x$ για κάποιο $\alpha \in \mathbb{R}$. Όμως στην περιπτωσή μας αυτό δεν μπορεί να συμβεί γιατί τότε θα ίσχυε ότι $x = \alpha e$, άρα $\xi_1 = \xi_2 = \cdots = \xi_m = \alpha$. Αυτό αποκλείεται αφού θεωρούμε ότι δεν είναι όλα τα ξ_j ίδια.

Είδαμε ότι η λύση του παρακάτω συστήματος ελαχιστοποιεί το $E(\gamma_1, \gamma_2)$:

$$\begin{pmatrix} m & \sum_{j=1}^m \xi_j \\ \sum_{j=1}^m \xi_j & \sum_{j=1}^m \xi_j^2 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m \psi_j \\ \sum_{j=1}^m \xi_j \psi_j \end{pmatrix}$$

Ενδιαφέρον: Αν θέσουμε

$$A = \begin{pmatrix} 1 & \xi_1 \\ \vdots & \vdots \\ 1 & \xi_m \end{pmatrix}, \quad y = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_m \end{pmatrix}$$

το παραπάνω είναι ίδιο με το σύστημα των κανονικών εξισώσεων

$$A^\top Ac = A^\top y$$

που θα προέκυπτε απευθείας αν επιχειρούσαμε να λύσουμε το $Ax = y$ πολλαπλασιάζοντας πρώτα από τα αριστερά με το $A^\top$ ή αν χρησιμοποιούσαμε την ιδέα της προβολής του y στον υπόχωρο στηλών του A.

Προσοχή: Η τάξη $\mathrm{rank}(A^\top A) \leq 2 = \min(m, n)$ επομένως το σύστημα θα έχει μοναδική λύση ανν υπάρχουν τουλάχιστον δύο i, j τ.ώ. $\xi_i \neq \xi_j$.

Βιβλιογραφία I

G. Strang.
Εισαγωγή στη Γραμμική Άλγεβρα.
Εκδόσεις Πανεπιστημιου Πατρών, 2006.

Χρήση Έργου Τρίτων I

1. http://www.netflixprize.com/ (βλ. σελ 10)
2. http://www.netflixprize.com/assets/GrandPrize2009_BPC_BigChaos.pdf (βλ. σελ 10)

Σημείωμα Αναφοράς

``Γραμμική Άλγεβρα'', Έκδοση: 1.0, Πάτρα 2014-2015.
Διαθέσιμο από τη δικτυακή διεύθυνση:

https://eclass.upatras.gr/courses/CEID1097/

Τέλος Ενότητας

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης