ΗΥ213. ΑΡΙΘΜΗΤΙΚΗ ΑΝΑΛΥΣΗ

ΕΛΑΧΙΣΤΑ ΤΕΤΡΑΓΩΝΑ
ΑΝΑΛΥΣΗ ΙΔΙΑΖΟΥΣΩΝ ΤΙΜΩΝ

Π. ΤΣΟΜΠΑΝΟΠΟΥΛΟΥ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Μέθοδος ελαχίστων τετραγώνων

- Τα σφάλματα στις πειραματικές μετρήσεις (από παρατηρήσεις και πειράματα) είναι αναπόφευκτα.
- Για να εξομαλύνουμε τα σφάλματα, λαμβάνουμε *μέσο όρο* από πολλές περιπτώσεις(π.χ., λαμβάνοντας περισσότερες μετρήσεις από ότι χρειαζόμαστε για να υπολογίσουμε κάποιες παραμέτρους).
- Το σύστημα στο οποίο καταλήγουμε είναι *υπερπροσδιορισμένο*.
- *Προβολή* των δεδομένων μεγαλύτερης διάστασης σε χώρους μικρότερης διάστασης, για να αφαιρέσουμε τις άσχετες λεπτομέρειες.
- Η προβολή στο χώρο μικρότερης διάστασης επιτυγχάνεται με τη μέθοδο των ελαχίστων τετραγώνων.

Παράδειγμα: Data Fitting

- Με δοσμένα τα m σημεία (x_i, y_i), υπολογίστε το n-διάστατο διάνυσμα c που καθορίζει τη συνάρτηση $f(x,c)$ η οποία προσεγγίζει τα δεδομένα με τον καλύτερο δυνατό τρόπο:

$$\min_c \sum_{i=1}^{m} \left(y_i - f(x_i, c)\right)$$

- Αν $\varphi_1, \varphi_2, \ldots, \varphi_n$, η βάση του χώρου που ανήκει η f τότε:

$$f(x, \mathbf{c}) = c_1\varphi_1(x) + c_2\varphi_2(x) + \ldots + c_n\varphi_n(x)$$

- Σε μορφή πινάκων το σύστημα γράφεται $Ac \approx b$, με $a_{ij} = \varphi_j(x_i)$ και $b_i = y_i$

Παράδειγμα: Data Fitting

- Αν $\varphi_1, \varphi_2, \ldots, \varphi_n$, είναι πολυώνυμα τότε:

$$f(x, \boldsymbol{c}) = c_1 + c_2 x + \ldots + c_n x^{n-1}$$

οπότε το **c** περιέχει τους συντελεστές του πολυωνύμου και το πρόβλημα της προσέγγισης είναι γραμμικό.

- Αν $\varphi_1, \varphi_2, \ldots, \varphi_n$, είναι εκθετικές συναρτήσεις τότε:

$$f(x,c) = c_1 e^{c_2 x} + c_3 e^{c_4 x} + \ldots + c_{n\text{-}1} e^{c_n x}$$

όπου το πρόβλημα είναι μη-γραμμικό (γιατί;;;).

- Θα μελετήσουμε μόνο γραμμικά προβλήματα ελαχίστων τετραγώνων.

Παράδειγμα: Data Fitting

- Προσέγγιση (με τη μέθοδο ελαχίστων τετραγώνων) 5 σημείων με πολυώνυμο 2ου βαθμού:

$$Ac = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ 1 & x_5 & x_5^2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \cong \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = b$$

όπου ο πίνακας A είναι τύπου *Vandermonde*.

Παράδειγμα: Data Fitting

- Με σημεία τα (-1, 1) , (-0.5, 0.5) , (0, 0) , (0.5, 0.5) , (1, 2), τότε:

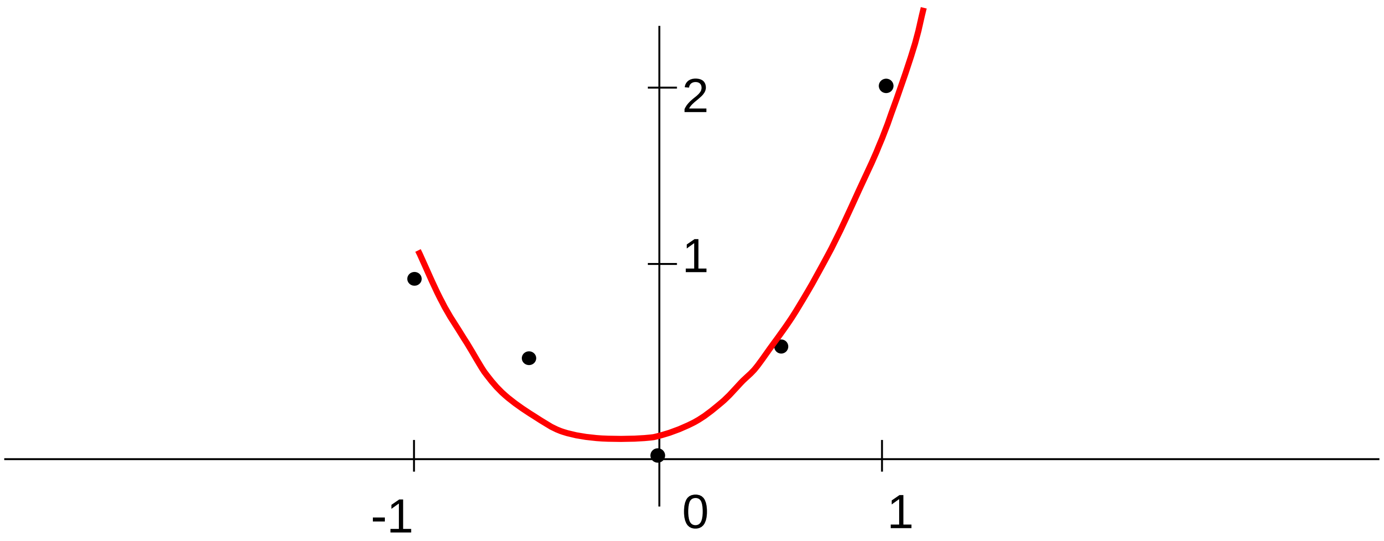
$$Ac = \begin{bmatrix} 1 & -1.0 & 1.0 \\ 1 & -0.5 & 0.25 \\ 1 & 0.0 & 0.0 \\ 1 & 0.5 & 0.25 \\ 1 & 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \cong \begin{bmatrix} 1.0 \\ 0.5 \\ 0.0 \\ 0.5 \\ 2.0 \end{bmatrix} = b$$

η μέθοδος ελαχίστων τετραγώνων δίνει λύση την $[0.086 \ \ 0.4 \ \ 1.4]^T$, οπότε το πολυώνυμο είναι:

$$p(x) = 0.086 + 0.4x + 1.4x^2$$

Παράδειγμα: Data Fitting

Τα 5 σημεία και το πολυώνυμο που τα προσεγγίζει.



Μέθοδος ελαχίστων τετραγώνων

- Για γραμμικά προβλήματα που καταλήγουν σε υπερπροσδιορισμένα γραμμικά συστήματα $Ax=b$, με τον A να είναι διάστασης $m \times n$ και $m>n$.
- Επειδή το σύστημα δεν λύνεται ακριβώς όταν $m>n$, συνήθως γράφουμε $Ax \approx b$.
- Η μέθοδος των ελαχίστων τετραγώνων υπολογίζει το x έτσι ώστε να ελαχιστοποιεί την Ευκλείδεια νόρμα του υπολοίπου:

$$\min_x \|r\|_2^2 = \min_x \|b - Ax\|_2^2$$

yota@inf.uth.gr

Ύπαρξη και μοναδικότητα της λύσης.

- Το πρόβλημα των ελαχίστων τετραγώνων $Ax \approx b$, έχει πάντα λύση.
- Η λύση είναι μοναδική αν και μόνο αν οι στήλες του A είναι γραμμικά ανεξάρτητες, δηλ. $\mathrm{rank}(A) = n$, αφού $m > n$.
- Αν $\mathrm{rank}(A) < n$, τότε η λύση δεν είναι μοναδική.
- Στα προβλήματα που θα λύσουμε θα ισχύει $\mathrm{rank}(A) = n$.

Αναγκαίες μαθηματικές έννοιες.

- Γραμμικοί χώροι.
- Εσωτερικό γινόμενο.
- Νόρμα.
- Βέλτιστη προσέγγιση.
- Βάσεις.
- Ορθογώνιες βάσεις.

Γραμμικοί χώροι - Εσωτερικό γινόμενο.

<u>Ορισμός:</u> *Γραμμικός χώρος* X, είναι ένας χώρος εφοδιασμένος με πρόσθεση και με πολλαπλασιασμό με πραγματικό αριθμό.

<u>Ορισμός:</u> Μία απεικόνιση (.,.): $X^x X \rightarrow R$ λέγεται *εσωτερικό γινόμενο*, αν ικανοποιεί τις ιδιότητες:

$$(x+y,z)=(x,z)+(y,z),\ \forall x,y,z\in X$$

$$(\lambda x,y)=\lambda(x,y),\ \forall x,y\in X$$

$$(x,y)=(y,x),\ \forall x,y\in X$$

$$(x,x)>0,\ \forall x\in X-\{0\}$$

Παραδείγματα.

- Εσωτερικό γινόμενο στον R^2 :

$$x=(a,b),\, y=(c,d)\in R^2, (x,y)=ac+bd$$

- Εσωτερικά γινόμενα στον $C([a,b])$:

$$f,g\in C([a,b]),\ (f,g)=\int_a^b f(x)g(x)dx,$$

ή

$$f,g\in C([a,b]),\ (f,g)=\int_a^b w(x)f(x)g(x)dx.$$

Νόρμα (ή στάθμη ή απόσταση).

<u>Ορισμός:</u> Μία απεικόνιση (.,.): X×X→R λέγεται νόρμα, αν ικανοποιεί τις ιδιότητες:

$$x \in X \quad \|x\| = 0 \Leftrightarrow x = 0$$

$$\forall \lambda \in R, \forall x \in X \quad \|\lambda \ x\| = |\lambda| \ \|x\|$$

$$\forall x, y \in X \quad \|x + y\| \leq \|x\| + \|y\|$$

Παραδείγματα.

- Νόρμα στον R^2 :

$$x = (a,b) \in R^2, \|x\| = \sqrt{a^2 + b^2}$$

- Νόρμες στον $C([a,b])$:

$$f \in C([a,b]),\ \|f\|_2 = \sqrt{\int_a^b f^2(x)dx}\ ,\qquad f \in C([a,b]),\ \|f\|_1 = \int_a^b |f(x)|dx\ ,$$

$$f \in C([a,b]),\ \|f\|_{w,2} = \sqrt{\int_a^b w(x) f^2(x)dx}\ ,\qquad f \in C([a,b]),\ \|f\|_\infty = \max_{a \le x \le b} |f(x)|\ .$$

Βάσεις – Ορθογώνιες Βάσεις.

Ορισμός: Έστω (X, (.,.)) ένας γραμμικός χώρος με εσωτερικό γινόμενο, ένα υποσύνολο σύνολο του $\{x_1,x_2,\ldots,x_n\}$, λέγεται *βάση*, ανν :

$$\forall x \in X, \exists a_1,a_2,\ldots,a_n \in \mathrm{R},\ \ x = \sum_{i=1}^{n} a_i x_i$$

Ορισμός: Έστω (X, (.,.)) ένας γραμμικός χώρος με εσωτερικό γινόμενο, ένα υποσύνολο σύνολο του $\{x_1,x_2,\ldots,x_n\}$, λέγεται *ορθογώνια βάση*, ανν είναι βάση και ισχύει:

$$(x_i,x_j) = 0,\quad i \neq j$$

Βάσεις (συνέχεια).

Έστω (X, (.,.)) γραμμικός χώρος με εσωτερικό γινόμενο, τότε από το εσωτερικό γινόμενο ορίζεται η νόρμα:

$$\forall x \in X, \quad \|x\| = \sqrt{(x,x)}$$

Έστω (X, (.,.)) γραμμικός χώρος με εσωτερικό γινόμενο, Y υποσύνολο του X, x στοιχείο του X, y η βέλτιστη προσέγγιση του x μέσα στο Y (με βάση τη νόρμα που παράγεται από το εσ. γιν.) και $\{z_i\}$, $i=1,\ldots,n$ βάση του Y, τότε ισχύει:

$$(x, z_i) = (y, z_i), \quad i = 1, \ldots, n$$

Ορθογωνιότητα.

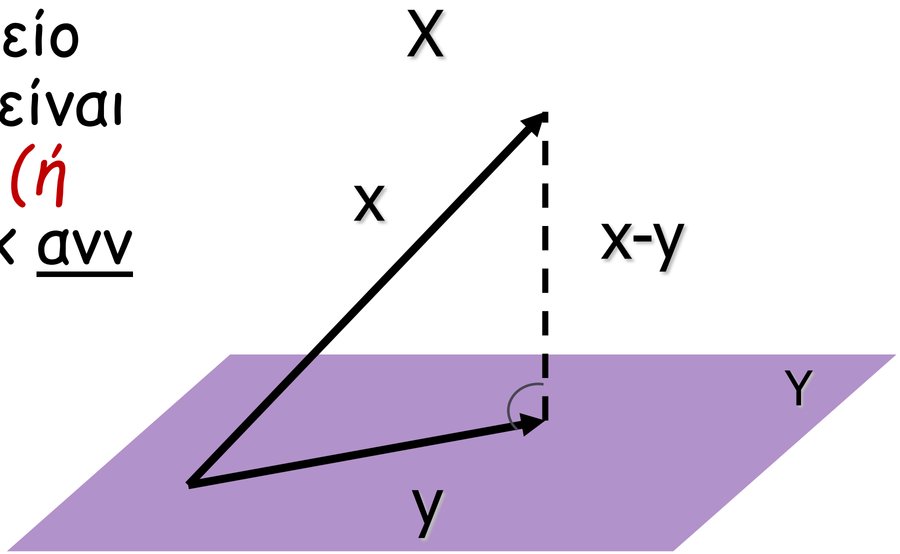
- Δύο διανύσματα v_1, v_2 λέγονται ορθογώνια ανν το εσωτερικό τους γινόμενο είναι 0, $v_1^T v_2 = 0$.
- Ο χώρος που παράγεται από τον A, $\mathrm{span}(A) = \{Ax: x \in \mathbf{R}^n\}$ είναι διάστασης το πολύ n.
- Αν $m>n$ τότε το b δεν ανήκει γενικά στον span(A) → το σύστημα $Ax=b$ δεν έχει λύση.
- Το διάνυσμα $y=Ax$ στον span(A), είναι η βέλτιστη προσέγγιση του b(ως προς την $||.||_2$ νόρμα), όταν το υπόλοιπο είναι κάθετο στο χώρο span(A):

$$0 = A^T r = A^T(b - Ax) \Rightarrow \quad A^T A x = A^T b$$

Βέλτιστη προσέγγιση.

<u>Ορισμός:</u> Έστω (X, ||.||) ένας γραμμικός χώρος με νόρμα, Y ένα υποσύνολο του X και x στοιχείο του X. Ένα στοιχείο y του Y είναι λέγεται *βέλτιστη προσέγγιση (ή προβολή του x στον Y)* του x <u>ανν</u> ισχύει:

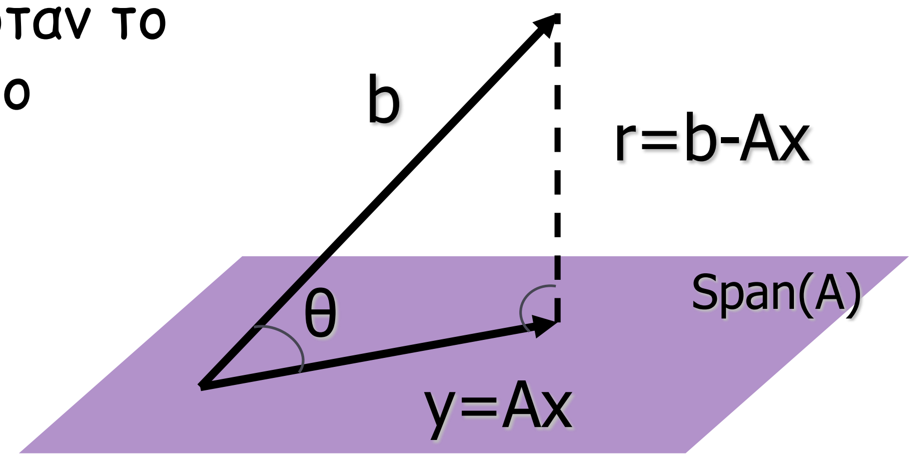
$$\forall z \in Y \ \ \|x-y\| \le \|x-z\|$$



ο Y λέγεται *χώρος προσέγγισης*.

Ορθογωνιότητα – Ιδιότητα της Βέλτιστης προσέγγισης.

Το διάνυσμα $y=Ax$ στον span(A), είναι η βέλτιστη προσέγγιση του b(ώς προς την $||.||_2$ νόρμα), εμφανίζεται όταν το υπόλοιπο είναι κάθετο στο χώρο span(A):



$$0 = A^T r = A^T(b - Ax) \Rightarrow \quad A^T A x = A^T b$$

Ψευδοαντίστροφος και δείκτης κατάστασης.

Για ένα πίνακα A, mxn, δεν υπάρχει αντίστροφος με τη γνωστή έννοια.

Αν rank(A)=n, ο ψευδοαντίστροφος ορίζεται ως:

$$A^{+} = (A^{T}A)^{-1}A^{T} \quad \text{και} \quad \mathrm{cond(A)} = \|\mathrm{A}\|_2 \|\mathrm{A}^{+}\|_2$$

Αν rank(A)<n τότε $\mathrm{cond(A)} = \infty$

Η λύση των ελαχίστων τετραγώνων για το σύστημα $Ax \approx b$, είναι

$$x = A^{+}b$$

Ευαισθησία και συνθήκες

Η μέθοδος των ελαχίστων τετραγώνων επηρεάζεται τόσο από τον πίνακα A όσο και από το διάνυσμα b.

Η γωνία θ μεταξύ του y=Ax και του b σαν:

$$\cos(\theta) = \frac{\|y\|_2}{\|b\|_2} = \frac{\|Ax\|_2}{\|b\|_2}$$

Το σχετικό σφάλμα φράσσεται:

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leq \mathrm{cond}(A) \frac{1}{\cos(\theta)} \frac{\|\Delta b\|_2}{\|b\|_2}$$

ή

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leq \left(\mathrm{cond}^2(A)\tan(\theta) + \mathrm{cond}(A)\right)\frac{\|\Delta A\|_2}{\|A\|_2}$$

Κανονικές εξισώσεις.

Τα ελάχιστα τετράγωνα ελαχιστοποιούν το τετράγωνο της Ευκλείδειας νόρμας του υπολοίπου:

$$\|r\|_2^2 = r^T r, \qquad r = b - Ax$$

Για να ελαχιστοποιήσουμε το υπόλοιπο:

$$\|r\|_2^2 = r^T r = (b - Ax)^T (b - Ax)$$

$$= b^T b - 2x^T A^T b + x^T A^T Ax,$$

παραγωγίζουμε ως προς x, την θέτουμε ίση με 0 και λύνουμε ως προς x:

$$2A^T Ax - 2A^T b = 0 \Leftrightarrow A^T Ax = A^T b$$

όπου το τελευταίο σύστημα είναι nxn και ονομάζεται σύστημα κανονικών εξισώσεων.

Μέθοδος Κανονικών Εξισώσεων

Αν ο mxn πίνακας A έχει τάξη n, τότε ο συμμετρικός nxn πίνακας $A^T A$ είναι θετικά ορισμένος, οπότε η παραγοντοποίηση Cholesky:

$$A^T A = LL^T$$

μπορεί να χρησιμοποιηθεί για τη λύση του συστήματος:

$$A^T Ax = A^T b$$

που έχει την ίδια λύση με αυτή του προβλήματος των γραμμικών ελαχίστων τετραγώνων $Ax \cong b$.

Παράδειγμα: Μέθοδος Κανονικών Εξισώσεων

Με σημεία τα (-1, 1), (-0.5, 0.5) , (0, 0), (0.5, 0.5), (1, 2), τότε:

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1.0 & -0.5 & 0.0 & 0.5 & 1.0 \\ 1.0 & 0.25 & 0.0 & 0.25 & 1.0 \end{bmatrix} \begin{bmatrix} 1 & -1.0 & 1.0 \\ 1 & -0.5 & 0.25 \\ 1 & 0.0 & 0.0 \\ 1 & 0.5 & 0.25 \\ 1 & 1.0 & 1.0 \end{bmatrix} = \begin{bmatrix} 5.0 & 0.0 & 2.5 \\ 0.0 & 2.5 & 0.0 \\ 2.5 & 0.0 & 2.125 \end{bmatrix},$$

$$A^T b = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1.0 & -0.5 & 0.0 & 0.5 & 1.0 \\ 1.0 & 0.25 & 0.0 & 0.25 & 1.0 \end{bmatrix} \begin{bmatrix} 1.0 \\ 0.5 \\ 0.0 \\ 0.5 \\ 2.0 \end{bmatrix} = \begin{bmatrix} 4.0 \\ 1.0 \\ 3.25 \end{bmatrix}$$

Παράδειγμα: Μέθοδος Κανονικών Εξισώσεων

Η παραγοντοποίηση Cholesky για τον $A^T A$ δίνει:

$$A^T A = \begin{bmatrix} 5.0 & 0.0 & 2.5 \\ 0.0 & 2.5 & 0.0 \\ 2.5 & 0.0 & 2.125 \end{bmatrix} =$$

$$\begin{bmatrix} 2.236 & 0 & 0 \\ 0 & 1.581 & 0 \\ 1.118 & 0 & 0.935 \end{bmatrix} \begin{bmatrix} 2.236 & 0 & 1.118 \\ 0 & 1.581 & 0 \\ 0 & 0 & 0.935 \end{bmatrix}$$

$$= LL^T$$

Παράδειγμα: Μέθοδος Κανονικών Εξισώσεων

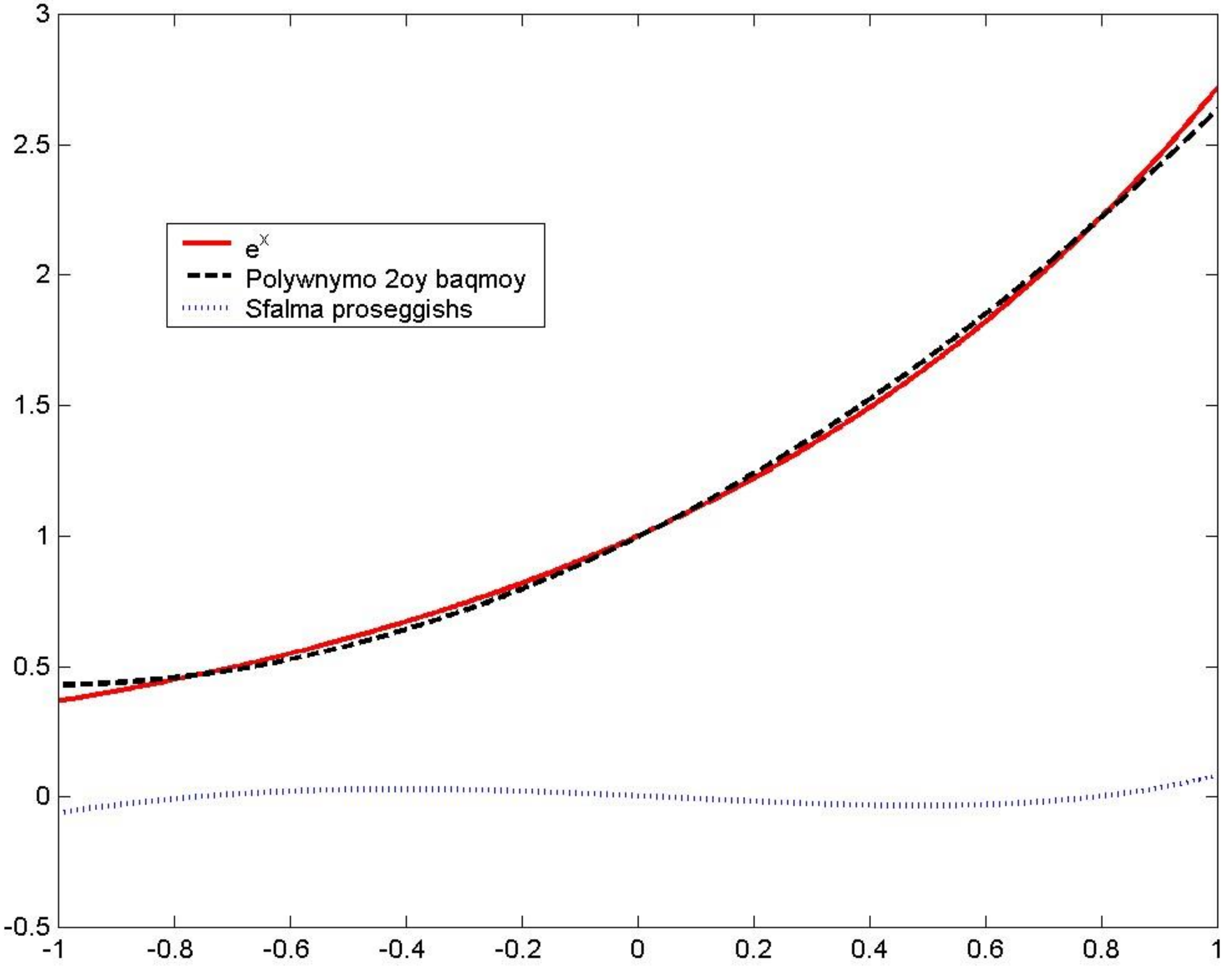
Λύνοντας το κάτω τριγωνικό σύστημα $Lz=A^Tb$ με προς τα εμπρός αντικατάσταση έχουμε:

$$z = \begin{bmatrix} 1.789 \\ 0.632 \\ 1.336 \end{bmatrix}$$

Λύνοντας το άνω τριγωνικό σύστημα $L^Tx=z$, με προς τα πίσω αντικατάσταση έχουμε την λύση των ελαχίστων τετραγώνων:

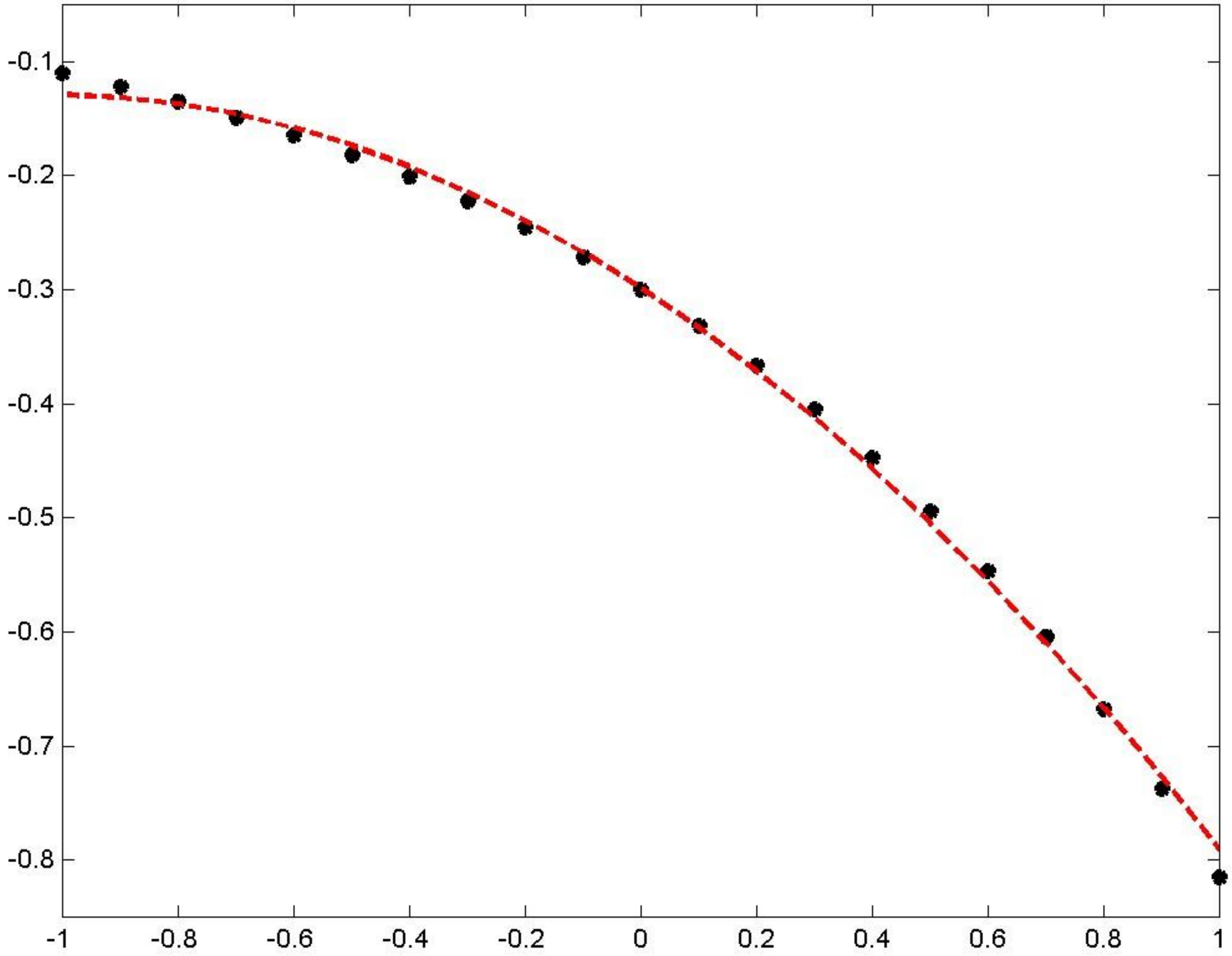
$$x = \begin{bmatrix} 0.086 \\ 0.400 \\ 1.429 \end{bmatrix}$$

Προσέγγιση συναρτήσεων.



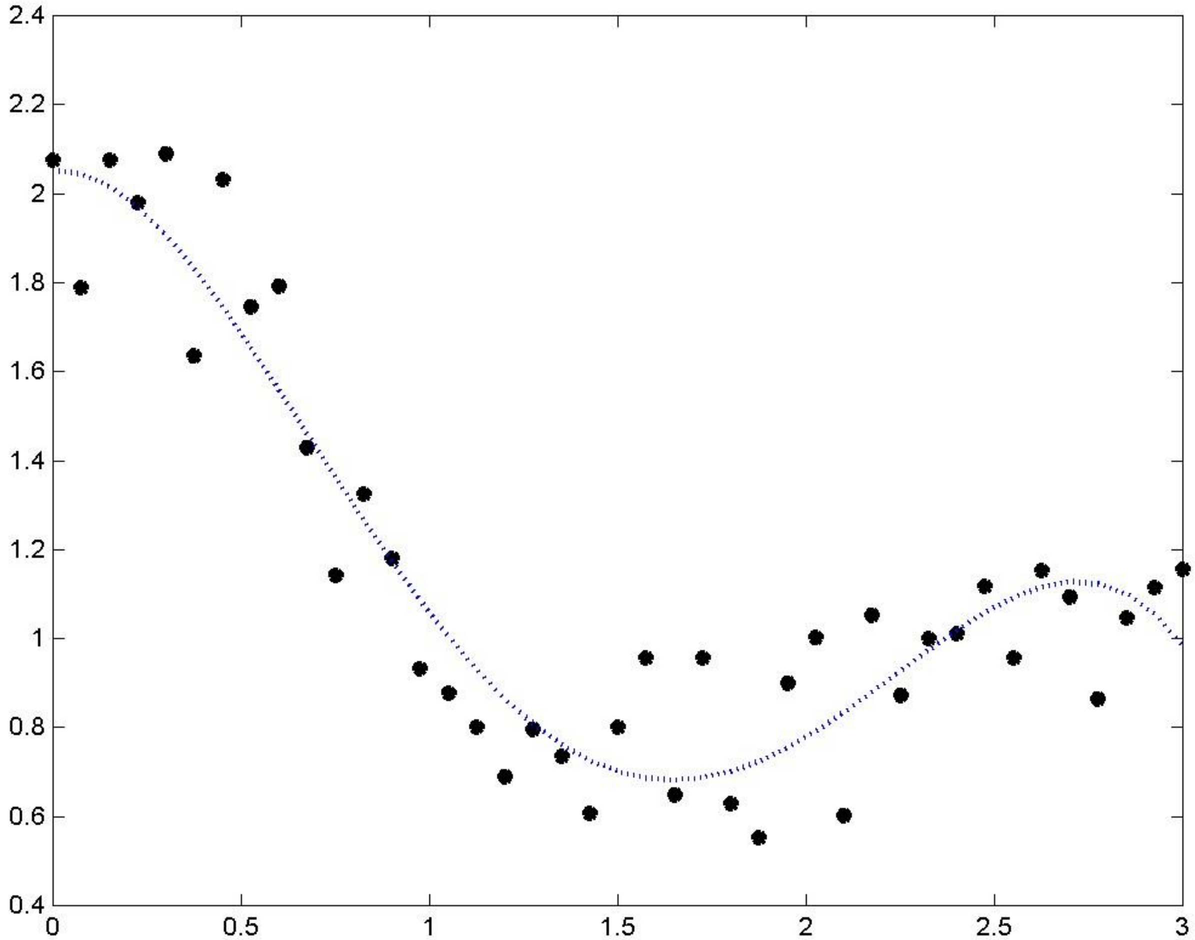
- Υψηλή ακρίβεια (6-30 δεκαδ. ψηφία).
- Ανυπαρξία αβεβαιότητας στη συνάρτηση.
- Χρήση όλων των ιδιοτήτων της συνάρτησης για μείωση πολυπλοκότητας.

Αναπαράσταση δεδομένων.



- Μεσαία ακρίβεια (2-5 δεκαδ. ψηφία).
- Μικρή αβεβαιότητα στα δεδομένα.
- Μικρή αλλά ικανή πληροφορία για τα δεδομένα (π.χ. μεταβάλλονται εκθετικά).
- Δύσκολο να εμπλουτίσουμε το δείγμα μας.

Ανάλυση δεδομένων.



- Χαμηλή ακρίβεια (0.5-3 δεκαδ. ψηφία).
- Μεγάλη αβεβαιότητα στα δεδομένα.
- Καμία πληροφορία για τα δεδομένα.
- Δύσκολο να εμπλουτίσουμε το δείγμα μας.

Διακριτά ελάχιστα τετράγωνα - Προσέγγιση δεδομένων.

Έστω $\mathbf{x}=(x_1,x_2,\ldots,x_m)$, $\mathbf{y}=(y_1,y_2,\ldots,y_m)$ διανύσματα του $\mathbf{R}^m$ που περιέχουν τις τετμημένες και τεταγμένες των δεδομένων.

Για να προσεγγίσουμε τα σημεία με μια συνάρτηση με το βέλτιστο τρόπο στον $\mathbf{R}^n$, ως προς την Ευκλείδεια νόρμα $||.||_2$ η οποία ορίζεται ως:

$$z \in R^m, \quad \|z\|_2 = \sqrt{\sum_{i=1}^{m} z_i^2}$$

αρκεί να υπολογίσουμε τους συντελεστές της f ως προς το σύστημα βάσης συναρτήσεων που επιθυμούμε να χρησιμοποιήσουμε.

Αν $\{\varphi_i,\ i=1,\ldots,n\}$ είναι οι συναρτήσεις βάσης, τότε $$f = \sum_{i=1}^{n} c_i \varphi_i,$$

Επιθυμούμε να ελαχιστοποιήσουμε:

$$\tilde{y} = \left(f(x_1) \quad \ldots \quad f(x_m)\right), \quad \|y-\tilde{y}\|_2 = \left\|y - \sum_{i=1}^{n} c_i\varphi_i(x)\right\|_2$$

Ελαχιστοποιούμε την ποσότητα:

$$r = \|y - \tilde{y}\|_2^2 = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} \mathrm{c}_j \varphi_j(\mathrm{x}_i) - y_i\right)^2$$

$$\frac{\partial r}{\partial c_k} = 0, \quad k = 1, \ldots, n$$

σε μορφή πίνακα (κανονικές εξισώσεις):

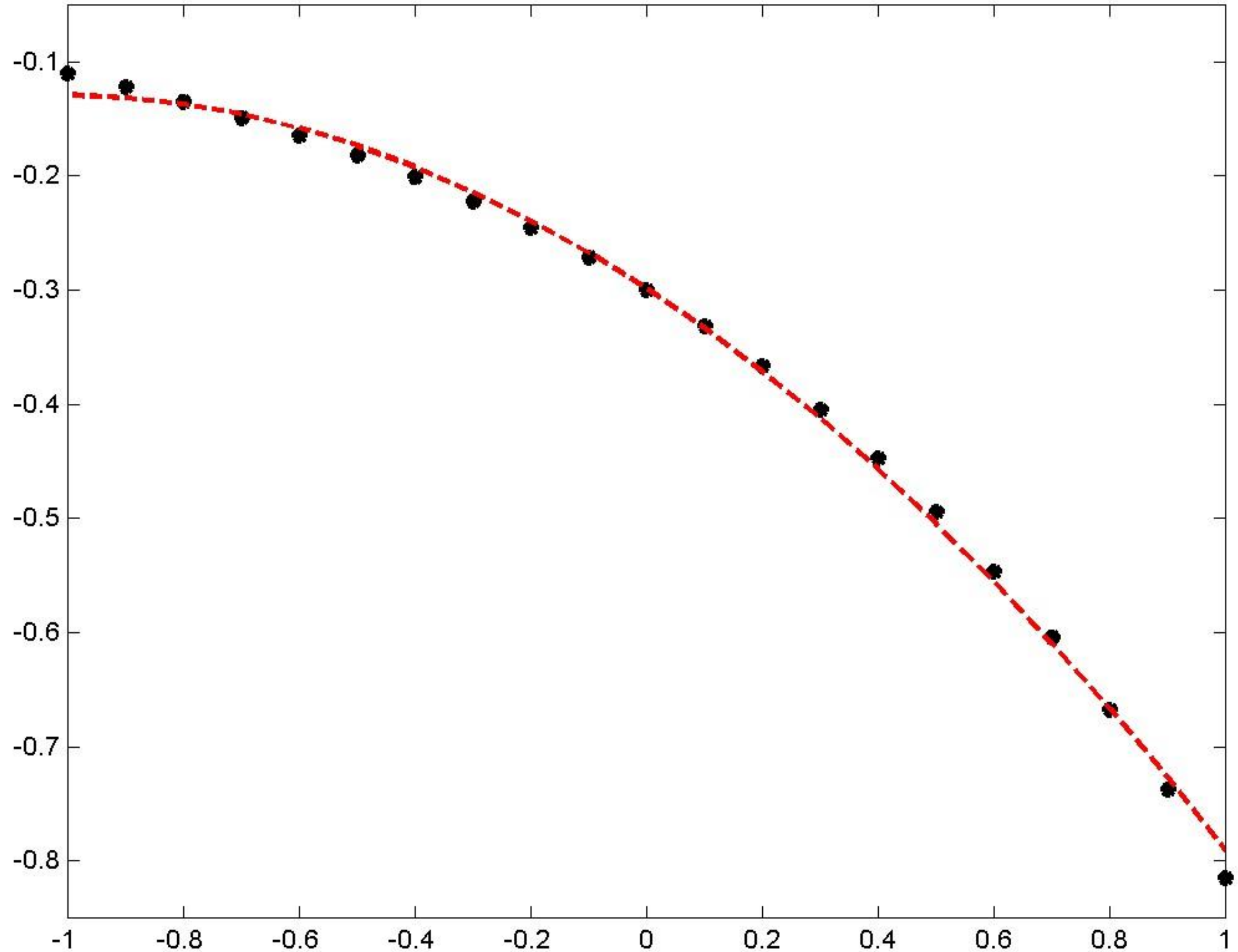
$$\mathrm{Pc} = \mathrm{b},$$

$$\mathrm{p_{ij}} = \sum_{\mathrm{k}=1}^{\mathrm{m}} \varphi_{\mathrm{i}}(\mathrm{x_k})\varphi_{\mathrm{j}}(\mathrm{x_k}), \qquad \mathrm{i, j} = 1, \ldots, \mathrm{n}$$

$$\mathrm{b_i} = \sum_{\mathrm{k}=1}^{\mathrm{m}} \varphi_{\mathrm{i}}(\mathrm{x_k})\mathrm{y_k}, \qquad \mathrm{i} = 1, \ldots, \mathrm{n}$$

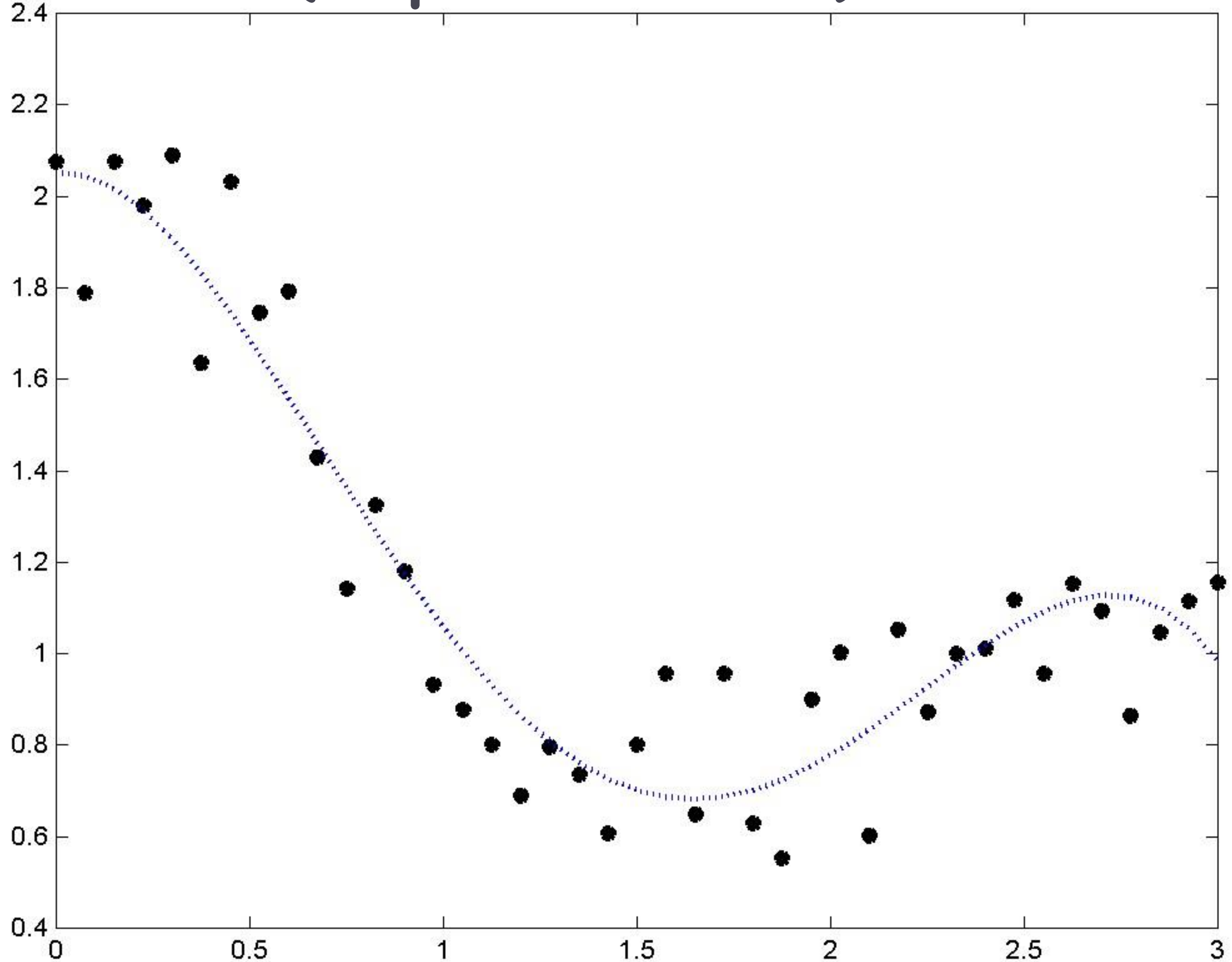
Ερώτηση: Ποια η διαφορά μεταξύ των πινάκων από τις κανονικές εξισώσεις στην προσέγγιση συναρτήσεων και διακριτών δεδομένων;

Εκτίμηση σφάλματος προσέγγισης δεδομένων (a-posteriori).



$$\left\| \mathbf{y} - \tilde{\mathbf{y}} \right\|_2$$

Εκτίμηση σφάλματος προσέγγισης δεδομένων (a-posteriori).



$$\left\| \mathbf{y} - \tilde{\mathbf{y}} \right\|_2$$

Διακριτά ελάχιστα τετράγωνα - Προσέγγιση δεδομένων με ευθεία.

Έστω $\mathbf{x}=(x_1,x_2,\ldots,x_m)$, $\mathbf{y}=(y_1,y_2,\ldots,y_m)$ διανύσματα του $\mathbf{R}^m$ που περιέχουν τις τετμημένες και τεταγμένες των δεδομένων.

Για να προσεγγίσουμε τα σημεία με μια *ευθεία* με το βέλτιστο τρόπο στον $\mathbf{R}^m$, ως προς την Ευκλείδεια νόρμα $||.||_2$ η οποία ορίζεται ως:

$$\mathrm{z} \in R^m, \quad \|\mathrm{z}\|_2 = \sqrt{\sum_{i=1}^{m} z_i^2}$$

αρκεί να υπολογίσουμε τα a_0 και a_1, (αν η ευθεία είναι η $y=a_0-a_1x$)

ώστε η ποσότητα να είναι η ελάχιστη δυνατή:

$$\|\mathbf{y}-\tilde{\mathbf{y}}\|_2 = \|\mathbf{y}-(a_0+a_1\mathbf{x})\|_2$$

- *Ανάλογα δουλεύουμε αν προσεγγίζουμε με παραβολή ή με πολυώνυμο μεγαλύτερου βαθμού.*

Παράδειγμα:

$$n = 2,\ \varphi_{\mathrm{i}}(\mathrm{x}) = x^{i-1},\ i = 1,2$$

σε μορφή πίνακα (κανονικές εξισώσεις):

$$Pc = b$$

$$P = \begin{bmatrix} m & \sum_{i=1}^{m} x_i \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 \end{bmatrix}, \qquad b = \begin{bmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} x_i y_i \end{bmatrix}$$

Προσέγγιση συναρτήσεων.

Αν X είναι ο χώρος των ομαλών συναρτήσεων με πεδίο ορισμού το [a, b] και Y ο χώρος προσέγγισης με βάση $\{B_i\}$, i=1,…,n, τότε αν $f \in X$ τότε η βέλτιστη προσέγγιση $\varphi \in Y$ έχει τη μορφή

$$\varphi(x) = \sum_{i=1}^{n} a_i B_i(x)$$

και ισχύει $(f, B_j) = (\varphi, B_j) \Leftrightarrow (f, B_j) = \sum_{i=1}^{n} a_i (B_i, B_j), \quad \forall j = 1,\ldots,n$

(κανονικές εξισώσεις).

Προσέγγιση συναρτήσεων (συνέχεια).

Το σύστημα που προκύπτει έχει τη μορφή:

$$\begin{bmatrix} (B_1,B_1) & \cdots & (B_n,B_1) \\ \vdots & \ddots & \vdots \\ (B_n,B_n) & \cdots & (B_n,B_n) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} (f,B_1) \\ \vdots \\ (f,B_n) \end{bmatrix}$$

όπου αν η βάση είναι ορθογώνια ισχύει:

$$(B_i,B_j)=0, i \neq j \qquad \Rightarrow$$

$$\begin{bmatrix} (B_1,B_1) & 0 & \cdots & 0 \\ 0 & (B_2,B_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (B_n,B_n) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} (f,B_1) \\ (f,B_2) \\ \vdots \\ (f,B_n) \end{bmatrix} \qquad \Rightarrow \qquad a_j = \frac{(f,B_j)}{(B_j,B_j)}, \quad j=1,\ldots,n$$

Προσέγγιση συναρτήσεων (συνέχεια).

Αν X είναι ο χώρος των ομαλών συναρτήσεων με πεδίο ορισμού το [a, b] και χώρος προσέγγισης είναι ο χώρος των πολυωνύμων βαθμού n, $\mathbf{P}_n([a,b])$, και αν $f \in X$ τότε η βέλτιστη προσέγγιση $p_n^* \in \mathbf{P_n}([a,b])$ έχει τη μορφή

$$p_n^*(x) = a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0$$

και ισχύει $\left\| f - p_n^* \right\| \le \left\| f - p \right\|, \ \forall p \in \mathbf{P_n}([a,b])$

ως προς τη συγκεκριμένη νόρμα.

Προσέγγιση συναρτήσεων (συνέχεια).

Τα μονώνυμα δεν αποτελούν ορθογώνια βάση.

Ορθοκανονικοποίηση Gram-Schmidt: Αν $x_1,\ldots,x_n$ βάση του χώρου X, τότε μπορούμε να κατασκευάσουμε μια ορθοκανονική βάση $e_1,\ldots,e_n$ του χώρου X.

Απόδειξη: Ορίζουμε

$$e_1' = x_1, \qquad e_2' = x_2 - \frac{(x_2, e_1')}{(e_1', e_1')} e_1', \quad \ldots$$

$$e_n' = x_n - \frac{(x_n, e_{n-1}')}{(e_{n-1}', e_{n-1}')} e_{n-1}' - \ldots - \frac{(x_n, e_1')}{(e_1', e_1')} e_1'$$

Θέτουμε:

$$e_i = \frac{e_i'}{\|e_i'\|}, \qquad i = 1,\ldots,n$$

yota@inf.uth.gr

Προσέγγιση συναρτήσεων (συνέχεια).

<u>Ανισότητα Bessel:</u> Αν $x_n \in X_n$ βέλτιστη προσέγγιση του $x \in X$, τότε

$\|x_n\| \le \|x\|$, και αν $e_1,\ldots,e_n$ ορθοκανονική βάση του X_n τότε ισχύει:

$$\sum_{i=1}^{n} |(x, e_i)|^2 \le \|x\|^2$$

<u>Απόδειξη:</u> Γνωρίζουμε ότι $(x-x_n, x_n)=0$, άρα

$$\|x\|^2 = \|(x-x_n)+x_n\|^2 = \|x-x_n\|^2 + \|x_n\|^2 \Rightarrow \|x\|^2 \ge \|x_n\|^2$$

Επίσης:

$$x_n = \sum_{i=1}^{n} (x, e_i) e_i \quad \Rightarrow \quad \|x_n\|^2 = \left\| \sum_{i=1}^{n} (x, e_i) e_i \right\|^2 = \sum_{i=1}^{n} \|(x, e_i) e_i\|^2 = \sum_{i=1}^{n} |(x, e_i)|^2 \cdot \|e_i\|^2 \Rightarrow$$

$$\left. \begin{aligned} \|x_n\|^2 &= \sum_{i=1}^{n} |(x, e_i)|^2 \\ \|x\|^2 &\ge \|x_n\|^2 \end{aligned} \right\} \Rightarrow \|x\|^2 \ge \sum_{i=1}^{n} |(x, e_i)|^2$$

Ορθογώνια πολυώνυμα.

<u>Θεώρημα:</u> Έστω w μια συνάρτηση βάρους(εσωτερικό γινόμενο), τότε υπάρχουν μοναδικά πολυώνυμα $p_n \in P_n$ τα οποία είναι ορθογώνια ως προς το $(.,.)_w$ και ισχύει:

$$\begin{cases} p_{-1}(x) = 0, \\ p_0(x) = 1, \\ p_{n+1}(x) = (x - a_{n+1})p_n(x) - \beta_n p_{n-1}(x), \quad n \geq 0 \end{cases}$$

και
$$\begin{cases} a_{n+1} = \dfrac{(\hat{x}p_n, p_n)_w}{(p_n, p_n)_w} & , \quad n \geq 0 \\ \beta_{n+1}(x) = \begin{cases} 0 & , \quad n = 0 \\ \dfrac{(p_n, p_n)_w}{(p_{n-1}, p_{n-1})_w}, & n \geq 1 \end{cases} \end{cases}$$
, όπου $\hat{x}$ η ταυτοτική συνάρτηση.

Ορθογώνια πολυώνυμα (συνέχεια).

- Πολυώνυμα Legendre: Αν $w(x)=1$ και το πεδίο ορισμού των συναρτήσεων είναι το $[-1,1]$ τότε τα ορθογώνια πολυώνυμα p_n ονομάζονται πολυώνυμα Legendre και ισχύει :

$$\gamma_n p_n(1)=1, \qquad P_n=\gamma_n p_n$$

και

$$P_0(x)=1, \quad P_1(x)=x,$$

$$(n+1)P_{n+1}(x)=(2n+1)xP_n(x)-nP_{n-1}(x), \quad n>1$$

Ορθογώνια πολυώνυμα (συνέχεια).

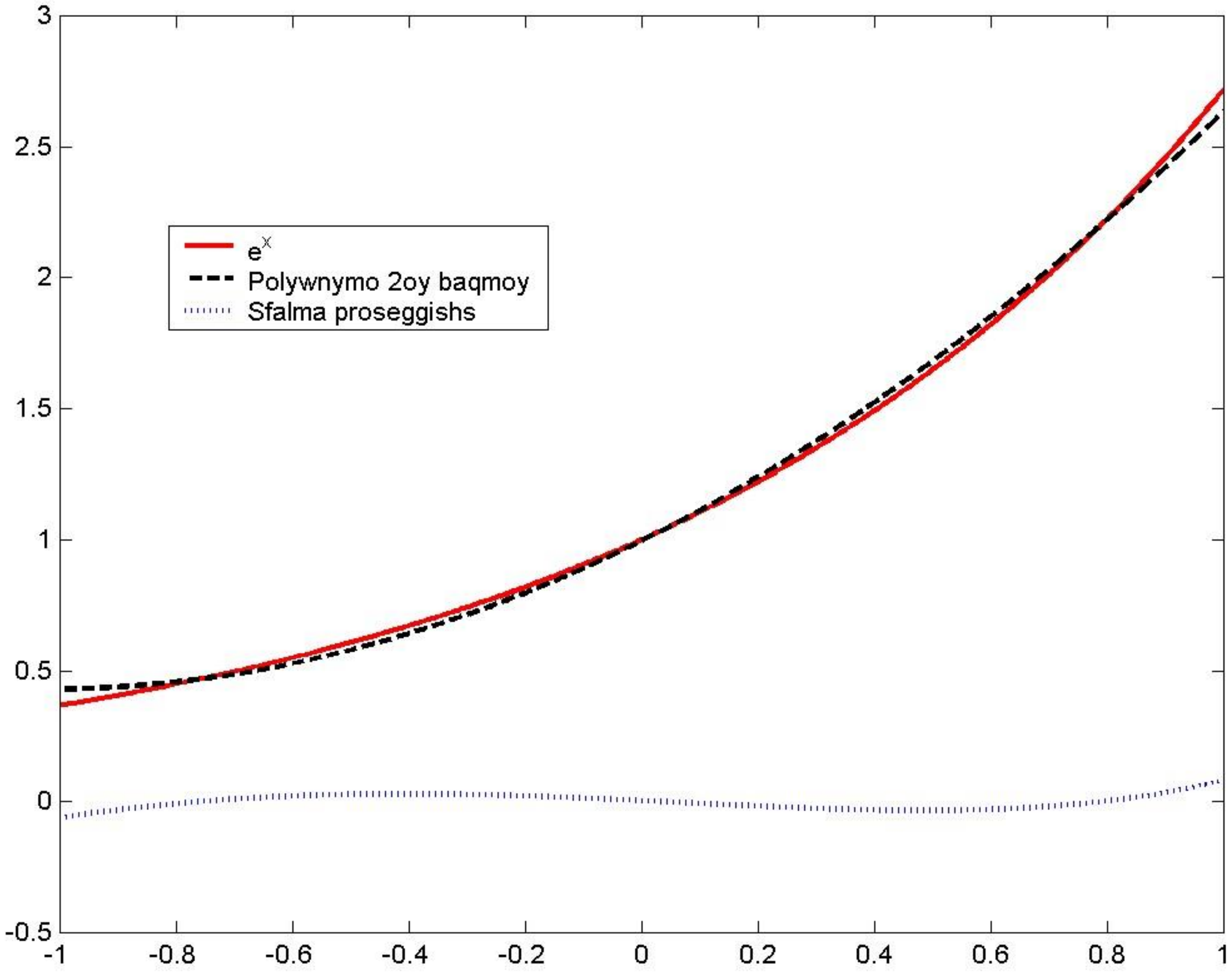
- **Πολυώνυμα Chebychev:** Αν $w(x)=(1-x)^{\alpha}(1+x)^{\beta}$, $\alpha=\beta=1/2$ και το πεδίο ορισμού των συναρτήσεων είναι το $[-1,1]$ τότε τα ορθογώνια πολυώνυμα T_n ονομάζονται πολυώνυμα Chebychev και ισχύει :

$$T_n(x)=\cos(n\arccos(x)), \quad x\in[-1,1], \quad n=0,1,\ldots,$$

ή

$$T_0(x)=1, \ T_1(x)=x, \ T_{n+1}(x)=2xT_n(x)-T_{n-1}(x), \quad x\in[-1,1], \quad n=1,\ldots,$$

Εκτίμηση σφάλματος προσέγγισης συνάρτησης (a-posteriori).



Κανονικές Εξισώσεις.

☹Σοβαρά μειονεκτήματα:

- Ο δείκτης κατάστασης του P είναι μεγάλος.
- Αν οι συναρτήσεις βάσεις δεν είναι ορθογώνιες, τότε ο P μπορεί είναι «σχεδόν μη αντιστρέψιμος».

☺Αντιμετώπιση με μετασχηματισμούς Householder.

- Ορισμός: Ένας τετραγωνικός πίνακας U λέγεται ορθοκανονικός ανν $U^TU=I$, δηλαδή ο ανάστροφος του είναι και ο αντίστροφος του.

Πέμπτη, 23 Απριλίου 2015 10:26 πμ

SVD:

$$U^T A V = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}$$

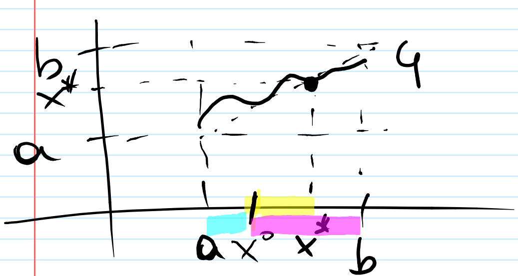
$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^T$$

Συστολή

$\varphi : [a,b] \to [a,b]$, συνεχής συστολή

$$x_0 \rightsquigarrow x_{n+1} = \varphi(x_n)$$

$x_n \longrightarrow x^*$ σταθερό σημείο της φ



$$x_n - x^* = \varphi(x_{n-1}) - x^* = \varphi(x_{n-1}) - \varphi(x^*)$$

όριο της ακολουθίας

x^* σταθερό σημείο της φ.

$$|x_n - x^*| = |\varphi(x_{n-1}) - \varphi(x^*)| \underset{\text{συστολή}}{\leq} L\,|x_{n-1} - x^*| \leq$$

$$\leq L \cdot L \cdot |x_{n-2} - x^*| = L^2 |x_{n-2} - x^*| \leq$$

$$\leq \dots \leq L^n |x_0 - x^*| \leq$$

$$\leq L^n \max\{x_0 - a,\ b - x_0\}$$

$$|x_n - x^*| \leq L^n \max\{x_0 - a,\ b - x_0\}$$

Π.χ.

$a = 0,\ b = 2,\ x_0 = 1$

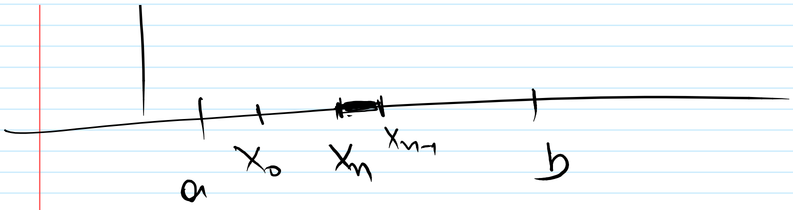
$L = 1/10$

Πόσες επαναλήψεις χρειάζεται ώστε το ~~x_n~~ να απέχει από το σταθερό σημείο το πολύ 10^{-6}

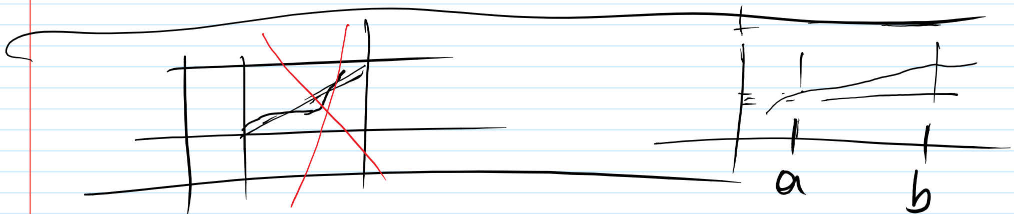
$$|x_n - x^*| \leq 10^{-6} \leq L^n \underbrace{\max\{1-0,\ 2-1\}}_{\to 1}$$

$$= 10^{-n}$$

$$\Rightarrow \boxed{n \geq 6}$$



$$|x_n - x^*| < \frac{L}{1-L}|x_n - x_{n-1}|$$

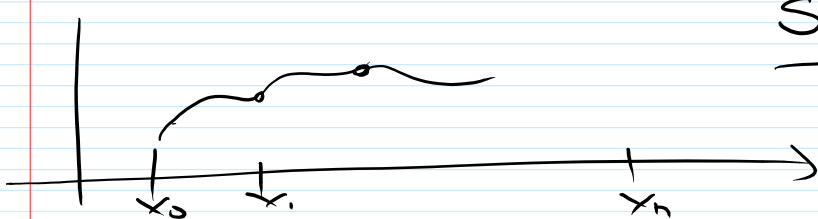


Splines:

α) κυβικά splines s

- Παρεμβολή
- κατά τμήματα κυβικά πολυώνυμα

$$S(x_i) = y_i$$



- $S \in C^2([x_0, x_n])$, συνεχής μέχρι 2η παράγωγο.

$$\Downarrow$$

$$S_-(x_i) = S_+(x_i)$$

$$S'_-(x_i) = S'_+(x_i)$$

$$S''_-(x_i) = S''_+(x_i)$$

β) Hermite splines s:

- παρεμβολή

- κατά τμήματα κυβικά πολυώνυμα.
- συνεχή μέχρι 1η παράγωγο
- $s(x_i) = y_i$
- $s'(x_i) = z_i$

- Νόρμες διανυσμάτων

$x \in \mathbb{R}^n \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

Νόρμα : $\|\cdot\| : X \longrightarrow \mathbb{R}_+$

$\|x\| = 0 \iff x = 0$

$\|\lambda x\| = |\lambda| \cdot \|x\|$

$\|x+y\| \leq \|x\| + \|y\|$

Π.χ.:

$\|\cdot\|_a : X \to \mathbb{R} \quad , \quad X = \mathbb{R}^n \quad a \in \mathbb{R}$

$\|x\|_a = |a x_1 + a x_2 + \dots a x_n|$

$\|\cdot\|_a$ νόρμα ??

ΟΧΙ γιατί για το $x = \begin{bmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \neq 0$

$$\|x\|_a = |a - a + 0 \cdots 0| = 0$$

Π.χ.:

$$\|\cdot\|_\beta : \mathbb{R}^n \to \mathbb{R} \qquad \beta \in \mathbb{R}_+$$

$$\|x\|_\beta = \beta(|x_1| + \dots + |x_n|)$$

νόρμα ??

- $\|x\|_\beta \geq 0 \quad \forall x \in \mathbb{R}^n$
- $\|x\|_\beta = 0 \overset{?}{\iff} x = 0$

$\cdot\ x = 0 \Rightarrow x_i = 0 \ \forall i = 1 \dots n \Rightarrow \|x\|_\beta = 0$

$\cdot\ \|x\|_\beta = 0 \Rightarrow \beta(|x_1| + |x_2| + \dots |x_n|) = 0 \Rightarrow$

$$\xRightarrow{\beta \neq 0} \sum |x_i| = 0 \Rightarrow \forall x_i = 0 \Rightarrow \underline{\underline{x = 0}}$$

- $\|\lambda x\|_\beta \overset{?}{=} |\lambda| \cdot \|x\|_\beta$

$$\|\lambda x\|_\beta = \beta(|\lambda x_1| + |\lambda x_2| + \dots + |\lambda x_n|) =$$

$$= \beta(|\lambda| \cdot (\underbrace{|x_1| + |x_2| + \dots |x_n|})) =$$

$$= |\lambda| (\beta(\qquad)) =$$

$$= |\lambda| \cdot \|x\|_\beta$$

$$- \|x+y\|_\beta \overset{?}{\leq} \|x\|_\beta + \|y\|_\beta$$

$$\|x+y\|_\beta = \beta\left(|x_1+y_1| + |x_2+y_2| + \dots |x_n+y_n|\right)$$

$$\leq \beta\left(|x_1|+|y_1| + |x_2|+|y_2| + \dots + |x_n|+|y_n|\right)$$

$$= \beta\left(\sum |x_i| + \sum |y_i|\right) =$$

$$= \beta \sum |x_i| + \beta \sum |y_i| = \|x\|_\beta + \|y\|_\beta$$

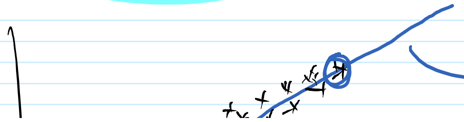
Νόρμες συναρτήσεων:

$$f \in C([a,b])$$

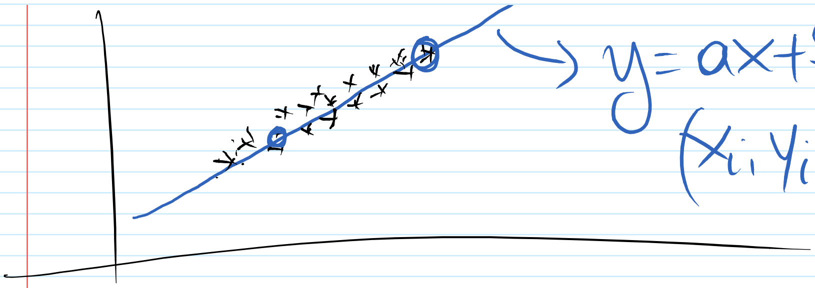
$$\|f\|_1 = \int_a^b |f(x)|\, dx$$

$$\|f\|_2 = \sqrt{\int_a^b f^2(x)\, dx}$$

$$\|f\|_\infty = \max_{a \leq x \leq b} |f(x)|$$



$\rightarrow y = ax + B$



$$y = ax + \beta$$

$$(x_i, y_i) \quad i = 0, \ldots, n$$

$$ax_0 + \beta = y_0$$
$$ax_1 + \beta = y_1$$
$$\vdots$$
$$ax_n + \beta = y_n$$

$$M \begin{bmatrix} a \\ \beta \end{bmatrix} = Y \qquad M = \begin{bmatrix} x_0 & 1 \\ x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}$$

$$r = y - M \begin{bmatrix} a \\ \beta \end{bmatrix} \neq 0 \Longrightarrow r_i = y_i - ax_i - \beta, \; i = 0, \ldots, n$$

Θέλω να υπολογίσω $\begin{bmatrix} a \\ \beta \end{bmatrix}$ τ.ω.

το r να είναι το ελάχιστο δυνατό.

$$\min_{a,\beta} \|r\| = g(a,\beta)$$

$$\xrightarrow{\|\cdot\|_1} \min_{a,\beta} \left(\sum_{i=0}^{n} |y_i - ax_i - \beta| \right)$$

$$\xrightarrow{\|\cdot\|_\infty} \min_{a,\beta} \left(\max_{i=0,\ldots,n} |y_i - ax_i - \beta| \right)$$

$$\xrightarrow{\|\cdot\|_2} \min_{a,\beta} \left(\sqrt{\sum_{i=0}^{n} (y_i - ax_i - \beta)^2} \right)$$

Άρα καταλήγω να δουλέψω το

$$\min_{a,\beta} \left(\sum_{i=0}^{n} (y_i - ax_i - \beta)^2 \right)$$

$$\min_{a,\beta} \Big(\underbrace{\sum_{i=0} (y_i - a x_i - \beta)^2}_{g(a,\beta)} \Big)$$

$$\frac{\partial g}{\partial a} = 0 \Leftrightarrow \frac{\partial}{\partial a}\Big(\sum (\ )^2\Big) = \sum_{i=0}^{n} \frac{\partial}{\partial a}(y_i - a x_i - \beta)^2 = \sum_{i=0}^{n} 2(y_i - a x_i - \beta)(-x_i)$$

$$= \sum(-2 y_i x_i + 2 a x_i^2 + 2\beta x_i)$$

$$= -2\sum y_i x_i + 2a\sum x_i^2 + 2\beta \sum x_i = 0$$

θέτω

Άρα

$$a\sum_{i=0}^{n} x_i^2 + \beta \sum_{i=0}^{n} x_i = \sum_{i=0}^{n} x_i y_i$$

Ομοίως

$$\frac{\partial g}{\partial \beta} = 0 \Rightarrow a\sum_{i=0}^{n} x_i + (n+1)\beta = \sum_{i=0}^{n} y_i$$

$$\Downarrow$$

$$\begin{bmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n+1 \end{bmatrix} \begin{bmatrix} a \\ \beta \end{bmatrix} = \begin{bmatrix} \sum x_i y_i \\ \sum y_i \end{bmatrix}$$