Γραμμική Άλγεβρα

Ενότητα **4** : Ορθογωνιότητα

Ευστράτιος Γαλλόπουλος
Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Άδειες Χρήσης

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.

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

2 Ορθογωνιότητα (μέρος *III*): Γραμμικό πρόβλημα ελαχίστων τετραγώνων

Υπενθύμιση και πρόγραμμα διάλεξης

Στην προηγούμενη διάλεξη μιλήσαμε για ορισμένες χρήσεις μητρώων και διανυσμάτων.

- Προβολές και ορθογώνιες προβολές (συνέχεια)
- Η μέθοδος ορθοκανονικοποίησης Gram-Schmidt
- Παραγοντοποίηση QR μέσω Gram-Schmidt

Σήμερα θα συζητήσουμε:

1. Γενική θεώρηση προβλημάτων ελαχιστοποίησης
2. Προβλήματα και προσεγγίσεις ελαχίστων τετραγώνων

Υπό συζήτηση ενότητες

1 Εισαγωγή στα Διανύσματα 1
- 1.1 Διανύσματα και Γραμμικοί Συνδυασμοί . . . 2
- 1.2 Μήκη και Στικτά Γινόμενα . . . 13

2 Επίλυση Γραμμικών Εξισώσεων 27
- 2.1 Διανύσματα και Γραμμικές Εξισώσεις . . . 27
- 2.2 Η Έννοια της Απαλοιφής . . . 44
- 2.3 Απαλοιφή Χρησιμοποιώντας Πίνακες . . . 58
- 2.4 Κανόνες για τις Πράξεις Πινάκων . . . 71
- 2.5 Αντίστροφοι Πίνακες . . . 89
- 2.6 Απαλοιφή = Παραγοντοποίηση: $A = LU$ . . . 105
- 2.7 Ανάστροφοι και Μεταθέσεις . . . 122

3 Διανυσματικοί Χώροι και Υπόχωροι 141
- 3.1 Χώροι Διανυσμάτων . . . 141
- 3.2 Ο Μηδενόχωρος του A: Επίλυση της $Ax = \mathbf{0}$ . . . 156
- 3.3 Η Τάξη και η Μορφή Ανηγμένων Γραμμών . . . 171
- 3.4 Η Πλήρης Λύση της $Ax = b$ . . . 184
- 3.5 Ανεξαρτησία, Βάση και Διάσταση . . . 199
- 3.6 Διαστάσεις των Τεσσάρων Υποχώρων . . . 219

4 Ορθογωνιότητα 233
- 4.1 Ορθογωνιότητα των Τεσσάρων Υποχώρων . . . 233
- 4.2 Προβολές . . . 246
- 4.3 Προσεγγίσεις Ελάχιστων Τετραγώνων . . . 261
- 4.4 Ορθογώνιες Βάσεις και **Gram – Schmidt** . . . 277

5 Ορίζουσες 295
- 5.1 Οι Ιδιότητες των Οριζουσών . . . 295
- 5.2 Μεταθέσεις και Αλγεβρικά Συμπληρώματα . . . 309
- 5.3 Κανόνας **Cramer**, Αντίστροφοι και Όγκοι . . . 327

6 Ιδιοτιμές και Ιδιοδιανύσματα 347
- 6.1 Εισαγωγή στις Ιδιοτιμές . . . 347
- 6.2 Διαγωνοποιώντας έναν Πίνακα . . . 365
- 6.3 Εφαρμογές στις Διαφορικές Εξισώσεις . . . 383
- 6.4 Συμμετρικοί Πίνακες . . . 401
- 6.5 Θετικά Ορισμένοι Πίνακες . . . 416
- 6.6 Όμοιοι Πίνακες . . . 432
- 6.7 Ανάλυση Ιδιαζουσών Τιμών (**SVD**) . . . 443

7 Γραμμικοί Μετασχηματισμοί 457
- 7.1 Η Έννοια του Γραμμικού Μετασχηματισμού . . . 457
- 7.2 Ο Πίνακας ενός Γραμμικού Μετασχηματισμού . . . 468
- 7.3 Αλλαγή Βάσης . . . 485
- 7.4 Η Διαγωνοποίηση και ο Ψευδοαντίστροφος . . . 494

8 Εφαρμογές 507
- 8.1 Πίνακες στη Μηχανική . . . 507
- 8.2 Γραφήματα και Δίκτυα . . . 521
- 8.3 Πίνακες **Markov** και Οικονομικά Μοντέλα . . . 535
- 8.4 Γραμμικός Προγραμματισμός . . . 545
- 8.5 Σειρές **Fourier**: Γραμμική Άλγεβρα για Συναρτήσεις . . . 553
- 8.6 Γραφικά με Ηλεκτρονικό Υπολογιστή . . . 561

9 Αριθμητική Γραμμική Άλγεβρα 569
- 9.1 Η Απαλοιφή **Gauss** στην Πράξη . . . 569
- Στάθμες και Δείκτες Κατάστασης . . . 581
- Επαναληπτικές Μέθοδοι για τη Γραμμική Άλγεβρα . . . 589

10 Μιγαδικά Διανύσματα και Πίνακες 603
- 10.1 Μιγαδικοί Αριθμοί . . . 603
- 10.2 Ερμιτιανοί και Μοναδιαίοι Πίνακες . . . 614
- 10.3 Ο Ταχύς Μετασχηματισμός **Fourier** . . . 625

Λύσεις σε Επιλεγμένες Ασκήσεις 635

Ένα Τελικό Διαγώνισμα 689

Παραγοντοποιήσεις Πινάκων 693

Ερωτήσεις Ανασκόπησης επί των Εννοιών 697

Γλωσσάριο 705

Κώδικες Διδασκαλίας MATLAB 717

Η Γραμμική Άλγεβρα με Δύο Λόγια 719

Ευρετήριο 721

Απόσταση από υπόχωρο και το γενικό πρόβλημα της προσέγγισης

Απόσταση από υπόχωρο

Έστω $\mathcal{S}$ ένα μη κενό υποσύνολο ενός διανυσματικού χώρου $\mathcal{V}$, που διαθέτει τη νόρμα $\|\cdot\|$.
Δοθέντος $v \in \mathcal{V}$, η απόσταση του v από το $\mathcal{S}$ ορίζεται ως

$$d(v, \mathcal{S}) = \inf_{s \in \mathcal{S}} \|v - s\|.$$

Ένα στοιχείο $\hat{s} \in \mathcal{S}$ αποκαλείται βέλτιστη προσέγγιση (ή πλησιέστερο στοιχείο) στο v από τον $\mathcal{S}$ αν

$$\|v - \hat{s}\| = d(v, \mathcal{S}).$$

Απόσταση από υπόχωρο και το γενικό πρόβλημα της προσέγγισης

Απόσταση από υπόχωρο

Έστω $\mathcal{S}$ ένα μη κενό υποσύνολο ενός διανυσματικού χώρου $\mathcal{V}$, που διαθέτει τη νόρμα $\|\cdot\|$.
Δοθέντος $v \in \mathcal{V}$, η απόσταση του v από το $\mathcal{S}$ ορίζεται ως

$$d(v, \mathcal{S}) = \inf_{s \in \mathcal{S}} \|v - s\|.$$

Ένα στοιχείο $\hat{s} \in \mathcal{S}$ αποκαλείται βέλτιστη προσέγγιση (ή πλησιέστερο στοιχείο) στο v από τον $\mathcal{S}$ αν

$$\|v - \hat{s}\| = d(v, \mathcal{S}).$$

Η αναζήτηση βέλτιστης προσέγγισης και τα συναφή προβλήματα (ύπαρξη, μοναδικότητα, κατασκευή) είναι κεντρικό ζήτημα των Μαθηματικών με πάρα πολλές εφαρμογές. Η αποτελεσματική επίλυσή του απαιτεί συνδυασμό μαθηματικών τεχνικών και υπολογιστικών τεχνολογιών και γνώσεις της εκάστοτε εφαρμογής.

Παράδειγμα

Πρόβλημα: Λαμβάνουμε μέτρησεις $\psi_1, \psi_2, \ldots, \psi_m$ (π.χ. ένα σήμα) στις χρονικές στιγμές $t = t_1, t_2, \ldots, t_m$. Παρατηρούμε ότι το σήμα έχει εκθετική και ταλαντούμενη συμπεριφορά. Αποφασίζουμε να το μοντελοποιήσουμε με μια συνάρτηση

$$f(t;x) \quad = \quad \xi_1 + \xi_2 e^{-(\xi_3 - t)^2/\xi_4} + \xi_5 \sin(\xi_6 t)$$

όπου τα στοιχεία του $x = (\xi_1, \xi_2, \ldots, \xi_6)^\top$ είναι οι παράμετροι του μοντέλου. Για να ορίσουμε τη συνάρτηση, θέλουμε να επιλέξουμε το x ώστε να επιτυγχάνεται η βέλτιστη προσαρμογή τιμών $f(t_j;x)$ στις μετρήσεις.

Διατύπωση ως πρόβλημα ελαχίστων τετραγώνων Να υπολογιστεί $x \in \mathbb{R}^6$ τέτοιο ώστε να ελαχιστοποιείται το άθροισμα των τετραγώνων

$$\min_{x \in \mathbb{R}^6} \sum_{j=1}^{m} r_j^2(x)$$

όπου $r_j(x) = \psi_j - f(t_j;x)$, $j = 1, \ldots, m$ είναι τα κατάλοιπα.

Ορολογία Πρόκειται για μη γραμμικό πρόβλημα ελαχίστων τετραγώνων. Είναι μη γραμμικό γιατί η συνάρτηση μοντέλο f εξαρτάται μη γραμμικά από τις παραμέτρους.

☺ Δεν θα ασχοληθούμε με μη γραμμικά προβλήματα!

Γραμμικό πρόβλημα ελαχίστων τετραγώνων[1]

- Στο γραμμικό πρόβλημα ελαχίστων τετραγώνων, η συνάρτηση μοντέλο $f(t;x)$ εξαρτάται γραμμικά από τις παραμέτρους:

$$f(t;x) \quad = \quad \xi_1 f_1(t) + \cdots + \xi_n f_n(t)$$

- Στο γραμμικό πρόβλημα ελαχίστων τετραγώνων, οι συναρτήσεις $f_j(t)$ μπορεί να είναι μη γραμμικές!
- π.χ.

$$f(t;x) \quad = \quad \xi_1 + \xi_2 e^{-t^2} + \xi_3 \sin^2(t\pi)$$

[1] Ονομάζεται και «πρόβλημα γραμμικών ελαχίστων τετραγώνων»

Υπενθύμιση

Γνωρίζουμε ότι για κάθε διανυσματικό υπόχωρο $\mathcal{S}$ ενός δ.χ. $\mathcal{V}$, ισχύουν τα εξής:

1. (θεώρημα προβολής) $\mathcal{V} = \mathcal{S} \oplus \mathcal{S}^{\perp}$ όπου
$$\mathcal{S}^{\perp} = \{y | y \perp x, \forall x \in \mathcal{S}\}$$
2. Κάθε $x \in \mathcal{V}$ γράφεται με μοναδικό τρόπο ως $x = x_{\mathcal{S}} + x_{\mathcal{S}^{\perp}}$, όπου $x_{\mathcal{S}} \in \mathcal{S}$ και $x_{\mathcal{S}^{\perp}} \perp \mathcal{S}$.
3. Ισχύουν ότι $x_{\mathcal{S}} = Px$ και $x_{\mathcal{S}^{\perp}} = (I - P)x$, όπου P είναι ο τελεστής ορθογώνιας προβολής επί του $\mathcal{S}$.
4. $\|x\|^2 = \|x_{\mathcal{S}} + x_{\mathcal{S}^{\perp}}\|^2 = \|x_{\mathcal{S}}\|^2 + \|x_{\mathcal{S}^{\perp}}\|^2$

Προσέγγιση ελαχίστων τετραγώνων

Γραμμικό πρόβλημα ελαχίστων τετραγώνων

Περιγραφή

Δίνονται $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $m \geq n$. Το **γραμμικό πρόβλημα ελαχίστων τετραγώνων** συνίσταται στην εύρεση ενός διανύσματος από το σύνολο

$$\mathcal{X} = \{x \in \mathbb{R}^n : \text{ ελαχιστοποιεί το } \rho(x) = \|Ax - b\|_2\}.$$

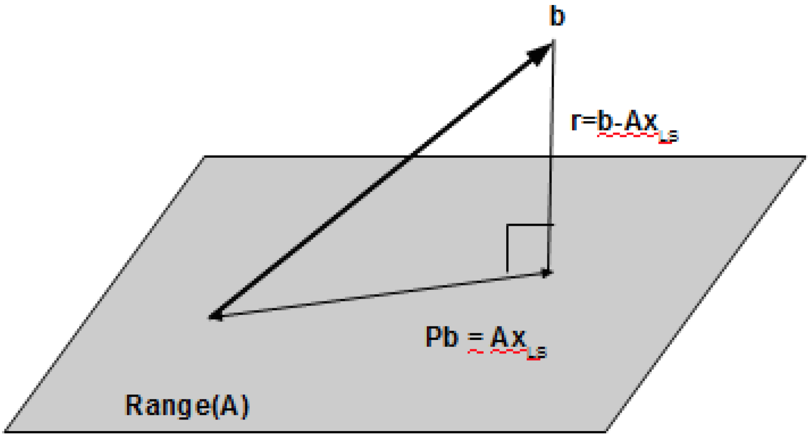
- Η ελαχιστοποίηση του $\|Ax - b\|_2$ ισοδυναμεί με την εύρεση διανύσματος $x_{LS} \in \mathbb{R}^n$ τέτοιου ώστε, μεταξύ όλων των διανυσμάτων που παράγονται από γραμμικό συνδυασμό των στηλών του A, το $p = Ax_{LS}$ να είναι το πλησιέστερο στο b (ως προς την ευκλείδεια νόρμα).
- Για να συμβαίνει αυτό, το $r = b - Ax_{\mathrm{LS}}$ είναι κάθετο στον υπόχωρο $\mathrm{range}(A)$ άρα $r \in \mathrm{null}(A^\top)$. Επομένως, για οποιοδήποτε $y \in \mathbb{R}^m$

$$0 = (Ay)^\top (b - Ax_{LS}) = y^\top (A^\top b - A^\top A x_{LS}).$$

Επομένως το x_{LS} είναι η λύση του συστήματος $\boxed{A^\top A x_{\mathrm{LS}} = A^\top b}$ που αποκαλούνται κανονικές εξισώσεις του προβλήματος των ελαχίστων τετραγώνων[2]

[2] Το μητρώο $A^\top A$ αποκαλείται συχνά μητρώο Gram ή Γραμμιανό.

Οπτικοποίηση



Επιλυσιμότητα των κανονικών εξισώσεων $A^\top Ax = A^\top b$

Υπάρχει λύση; Είναι μοναδική;

- Επιλυσιμότητα Είναι εξασφαλισμένο ότι υπάρχει τουλάχιστον μία λύση: Γιατί; Η τάξη $r = \mathrm{rank}(A^\top A) \leq n$. Αν $r = n$ τότε υπάρχει μοναδική λύση. Αν $r < n$, τότε ο μηδενόχωρος του $A^\top A$ έχει διάσταση $n - r$ και υπάρχουν άπειρες λύσεις. Στην περίπτωση αυτή, επιλέγουμε από τις λύσεις με κάποιο επιπλέον κριτήριο, π.χ. το διάνυσμα που έχει το μικρότερο μήκος (διπλή βελτιστοποίηση).
- Παρατήρηση: Ικανή και αναγκαία συνθήκη για να μην είναι αντιστρέψιμο το $A^\top A$ είναι να είναι γραμμικά εξαρτημένες οι στήλες του A (αποδείξτε το!)

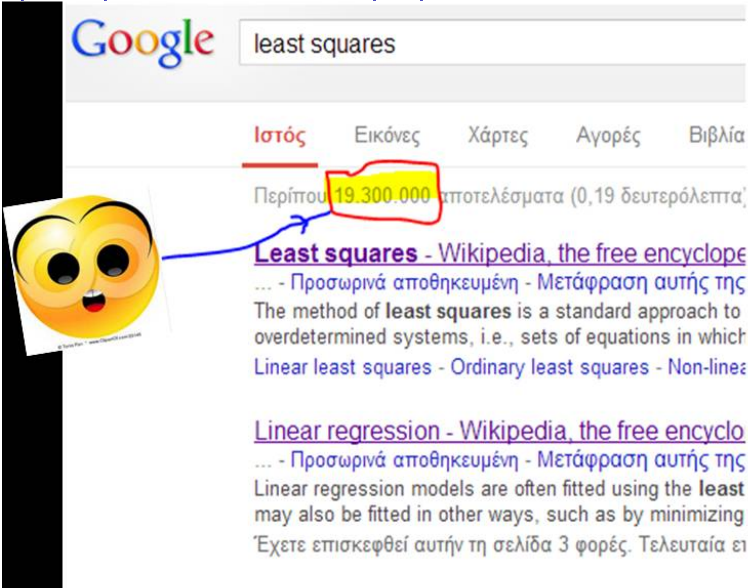
Γενική μορφή γραμμικού προβλήματος ελαχίστων τετραγώνων

$$f(\xi; c) = \gamma_1\phi_1(\xi) + \cdots + \gamma_n\phi_n(\xi).$$

Υπολογίζουμε τις τιμές $[\phi_1(\xi_j), \phi_2(\xi_j), \ldots, \phi_n(\xi_j)]$ για κάθε ξ_j, όπου $\phi_1, \ldots, \phi_n$ είναι γνωστές (συνήθως απλές) συναρτήσεις, ενώ οι συντελεστές $\gamma_1, \ldots, \gamma_n$ είναι σταθερές που πρέπει να καθοριστούν έτσι ώστε να ελαχιστοποιείται το άθροισμα των τετραγώνων του σφάλματος σε κάθε σημείο. Αν χρησιμοποιήσουμε μητρώα, το πρόβλημα γράφεται ισοδύναμα ως $\arg\min_{c\in\mathbb{R}^n} \|y - Ac\|$ όπου

$$A = \begin{pmatrix} \phi_1(\xi_1) & \cdots & \phi_n(\xi_1) \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \phi_1(\xi_m) & \cdots & \phi_n(\xi_m) \end{pmatrix}, c = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}. \text{ και } \quad y = \begin{pmatrix} \psi_1 \\ \vdots \\ \vdots \\ \psi_m \end{pmatrix}$$

Τι λέγεται για τα ελάχιστα τετράγωνα



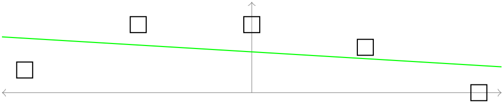
Τι λέγεται για τα ελάχιστα τετράγωνα

Προσαρμογή ελαχίστων τετραγώνων με ευθεία γραμμή

Πρόβλημα: Να υπολογίσουμε την ευθεία γραμμή $\gamma_1 + \gamma_2\xi$ που προσαρμόζεται σε ένα σύνολο γνωστών τιμών $\{(\xi_j, \psi_j)\}_{j=1}^m$ με το ελάχιστο δυνατό σφάλμα.

Ορολογία: Στη Στατιστική λέγεται «γραμμική παλινδρόμηση».

Πώς μετράμε το σφάλμα; Το ορίζουμε ως το άθροισμα των τετραγώνων των αποκλίσεων των τιμών της (γραμμικής) συνάρτησης, $\{\gamma_1 + \gamma_2\xi_j\}_{j=1}^m$ από τις τιμές των μετρήσεων ψ_j.

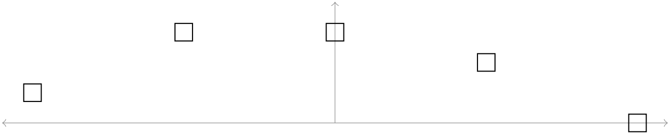


Με γραμμοαλγεβρική γραφή: $\gamma_1 + \gamma_2\xi = \psi$ γράφεται και ως $\begin{pmatrix} 1 & \xi \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = \psi$

Άλλες μετρικές σφάλματος Ανάλογα με την εφαρμογή, μπορεί οι απαιτήσεις να είναι διαφορετικές (και πιο απαιτητικές). Για παράδειγμα στα ολικά ελάχιστα τετράγωνα (total least squares) ελαχιστοποιούμε το άθροισμα τετραγώνων της απόστασης κάθε σημείου (ξ_j, ψ_j) από την ευθεία $\gamma_0 + \gamma_1\xi = \psi$ (δηλ. το μήκος της καθέτου από το κάθε σημείο στην ευθεία).

Παράδειγμα

Σημεία:	$\xi_1 = -1$	$\xi_2 = -1/2$	$\xi_3 = 0$	$\xi_4 = 1/2$	$\xi_5 = 1$
Τιμές:	$\psi_1 = 0.1$	$\psi_2 = 0.3$	$\psi_3 = 0.3$	$\psi_4 = 0.2$	$\psi_5 = 0.0$



$\phi_0(\xi) = 1, \phi_1(\xi) = \xi$ τότε $\phi(\xi) = \gamma_0\phi_0(\xi) + \gamma_1\phi_1(\xi)$ οπότε

$$A = \begin{pmatrix} 1.0 & -1.0 \\ 1.0 & -0.5 \\ 1.0 & 0.0 \\ 1.0 & 0.5 \\ 1.0 & 1.0 \end{pmatrix}$$

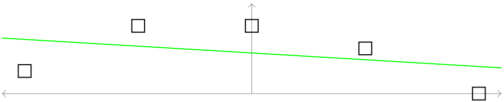
$$A^\top A = \begin{pmatrix} 5.0 & 0.0 \\ 0.0 & 2.5 \end{pmatrix}, A^\top b = \begin{pmatrix} 0.9 \\ -0.15 \end{pmatrix}$$

Αν κάνουμε τους υπολογισμούς

$$(A^\top A)^{-1} A^\top b = \begin{pmatrix} 0.18 \\ -0.06 \end{pmatrix}$$

$$\Rightarrow \boxed{\phi(\xi) = 0.18 - 0.06\xi}$$

σφάλμα $\|b - Ax_{\mathrm{LS}}\|_2 = 0.2429$



Προσαρμογή με καμπύλη γραμμή 2ου βαθμού

$\phi_0(x) = 1, \phi_1(\xi) = \xi, \phi_2(\xi) = \xi^2$ και $\phi(\xi) = \gamma_0\phi_0(\xi) + \gamma_1\phi_1(\xi) + \gamma_2\phi_2(\xi)$

$$A = \begin{pmatrix} 1.0 & -1.0 & 1.0 \\ 1.0 & -0.5 & 0.25 \\ 1.0 & 0.0 & 0.0 \\ 1.0 & 0.5 & 0.25 \\ 1.0 & 1.0 & 1.0 \end{pmatrix}$$

Αν κάνουμε τους υπολογισμούς

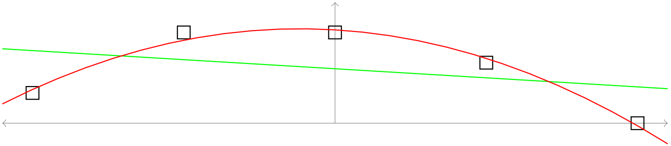
$$x_{\mathrm{LS}} = (A^\top A)^{-1} A^\top b = \begin{pmatrix} 0.3086 \\ -0.0600 \\ -0.2571 \end{pmatrix}$$

$$A^\top A = \begin{pmatrix} 5.0 & 0.0 & 2.5 \\ 0.0 & 2.5 & 0.0 \\ 2.5 & 0.0 & 2.125 \end{pmatrix},$$

επομένως $\boxed{\phi(\xi) = 0.3086 - 0.0600\xi - 0.2571\xi^2}$

σφάλμα $\|b - Ax_{\mathrm{LS}}\|_2 = 0.0338$

$$A^\top b = \begin{pmatrix} 0.9 \\ -0.15 \\ 0.225 \end{pmatrix}$$



Βιβλιογραφία I

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

Χρήση Έργου Τρίτων I

1. http://www.google.gr/ (βλ. σελ 15-16)

Σημείωμα Αναφοράς

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

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

Τέλος Ενότητας