Κεφάλαιο 1

Ελάχιστα Τετράγωνα

1.1 Εισαγωγή

Σε προηγούμενο κεφάλαιο αναπτύξαμε διάφορες μεθόδους για την επίλυση του γραμμικού συστήματος

$$A\underline{x} = \underline{b}$$

όπου ο πίνακας υποθέσαμε ότι είναι τετραγωνικός και μη ιδιάζων.
Παρόλα αυτά στις περισσότερες εφαρμογές και κυρίως όσες προέρχονται από τη Στατιστική και την Επεξεργασία σήματος χρειάζεται η επίλυση ενός συστήματος όπου ο πίνακας A δεν είναι τετραγωνικός και ενδεχόμενα ιδιάζων. Σαύτές τις περιπτώσεις μπορεί είτε να μην υπάρχουν λύσεις ή να υπάρχουν άπειρες το πλήθος. Για παράδειγμα, όταν ο A είναι $m \times n$ και $m > n$ έχουμε ένα overdetermined σύστημα το οποίο τυπικά δεν έχει λύση. Αντίθετα ένα υνδερδετερμινεδ σύστημα $(m < n)$ τυπικά έχει άπειρο αριθμό από λύσεις.
Σαύτές τις περιπτώσεις το καλλίτερο που μπορεί να συμβεί είναι να βρούμε ένα διάνυσμα x το οποίο κάνει το Ax να βρίσκεται όσο το δυνατόν πιο κοντά στο b. Με άλλα λόγια ψάχνουμε ένα διάνυσμα x τέτοιο ώστε να ελαχιστοποιεί την ποσότητα $||r(x)|| = ||Ax - b||$
Όταν χρησιμοποιείται η Ευκλείδια νόρμα η λύση αυτή αναφέρεται σα λύση ελαχίστων τετραγώνων για το σύστημα $Ax = b$. Ο όρος λύση ελαχίστων τετραγώνων προκύπτει από το γεγονός ότι η λύση αυτή ελαχιστοποιεί την Ευκλείδια νόρμα του διανύσματος του υπολοίπου, το τετράγωνο της οποίας από τον ορισμό της είναι ακριβώς το άθροισμα των τετραγώνων των συνιστωσών του διανύσματος. Το πρόβλημα του προσδιορισμού λύσεων ελαχίστων τετραγώνων για το γραμμικό σύστημα $Ax = b$ είναι γνωστό σα το **Γραμμικό Πρόβλημα Ελαχίστων Τετραγώνων (ΓΠΕΤ) (Linear Least-squares problem LSP)**.

Το ΓΠΕΤ τυπικά ορίζεται ως εξής:

Εστω $A \in \mathbb{R}^{m \times n}$ ένας δοσμένος πίνακας και $b \in \mathbb{R}^{n}$. Ζητείται να προσδιοριστεί διάνυσμα x έτσι ώστε η συνάρτηση $||r(x)|| = ||Ax - b||_2$ να ελαχιστοποιείται.

Εάν το πρόβλημα έχει περισσότερες από μία λύση αυτή που έχει την ελάχιστη Ευκλείδια νόρμα ονομάζεται λύση ελάχιστου μήκους ή λύση ελάχιστης νόρμας.

Υπαρξη και μοναδικότητα των λύσεων

Οπως και στην περίπτωση επίλυσης ενός γραμμικού συστήματος, προκύπτουν οι ακόλουθες ερωτήσεις.
1. Υπάρχει πάντοτε μία λύση ελαχίστων τετραγώνων για το $Ax = b$ ·
2. Είναι η λύση αυτή μοναδική·
3. Πως μπορούμε να προσδιορίσουμε τέτοιες λύσεις·

1.2 Overdetermined ΓΠΕΤ

Υποθέτουμε ότι $A \in \mathbb{R}^{m\times n}$, $m \geq n$ δηλαδή το σύστημα είναι τλοερδετερμινεδ ή τετραγωνικό.

Θεώρημα 1.2.1 Υπαρξη και Μοναδικότητα Ελαχίστων Τετραγώνων

α) Έστω $A \in \mathbb{R}^{m\times n}$, $m \geq n$ και $b \in \mathbb{R}^m$. Τότε υπάρχει λύση ελαχίστων τετραγώνων $x \in \mathbb{R}^n$ του συστήματος $Ax = b$ αν και μόνο αν ικανοποιεί τις κανονικές εξισώσεις $A^tAx = A^tb$.

β) Η λύση ελαχίστων τετραγώνων όταν υπάρχει είναι μοναδική αν και μόνο αν $rankA = n$, δηλαδή ο A είναι full-rank.

Απόδειξη:

α) Ύπαρξη
Έστω $r(x) = b - Ax$ το υπόλοιπο και $y \in \mathbb{R}^n$, τότε $r(y) = b - Ay = r(x) + Ax - Ay = r(x) + A(x - y)$. Οπότε,

$$\begin{aligned}\|r(y)\|^2 &= (r(x) + A(x-y))^t(r(x) + A(x-y)) \\ &= r(x)^t r(x) + (x-y)^t A^t r(x) + r(x)^t A(x-y) + (x-y)^t A^t A(x-y) \\ &= \|r(x)\|^2 + 2(x-y)^t A^t r(x) + \|A(x-y)\|^2. \qquad \blacksquare\end{aligned}$$

($\Rightarrow$) Αρχικά υποθέτουμε ότι το x ικανοποιεί το σύστημα $A^tAx = A^tb$ (δηλαδή ότι αποτελεί λύση των κανονικών εξισώσεων), έτσι ώστε $A^tr(x) = 0$.
Τότε έχουμε $\|r(y)\|^2 = \|r(x)\|^2 + \|A(x-y)\|^2 \geq \|r(x)\|^2$.
Επομένως, το x είναι λύση ελαχίστων τετραγώνων.
($\Leftarrow$) Υποθέτουμε, τώρα, ότι το x δεν ικανοποιεί τις κανονικές εξισώσεις, οπότε

$A^t r(x) \neq 0$ και θέτουμε $A^t r(x) = z$. Ορίζουμε το διάνυσμα y ώστε $y = x + cz$,
c: βαθμωτό. Τότε $r(y) = r(x) + A(x-y) = r(x) - cAz$.
Έτσι έχουμε: $\|r(y)\|^2 = (r(x) - cAz)^t (r(x) - cAz) = \|r(x)\|^2 + c^2\|Az\|^2 - 2cz^t A^t r(x)$ και αφού $A^t r(x) = z$, τελικά

$$\|r(y)\|^2 = \|r(x)\|^2 + c^2\|Az\|^2 - 2c\|z\|^2. \tag{1.1}$$

Για να είναι η x λύση ελαχίστων τετραγώνων, θα πρέπει για κάθε $y \in \mathbb{R}^n$ να ισχύει:
$\|r(x)\|^2 \leq \|r(y)\|^2$. Όμως, για κάθε $c > 0$ (οπότε $\forall y$) αν $Az = 0$ ή επιλέγοντας $0 < c < \frac{2\|z\|^2}{\|Az\|^2}$ εάν $Az \neq 0$, από την (1.1) έχουμε ότι: $\|r(y)\|^2 \leq \|r(x)\|^2$. Άρα το x δεν είναι λύση ελαχίστων τετραγώνων σ' αυτές τις περιπτώσεις.

β) Μοναδικότητα
Είπαμε ότι η λύση του ΓΠΕΤ ικανοποιεί το $A^t Ax = A^t b$. Όμως η λύση αυτού του συστήματος είναι μοναδική, εάν και μόνο αν ο $A^t A$ είναι μη-ιδιάζων. Θα δείξουμε ότι αν $rank(A) = n$, ο $A^t A$ είναι θετικά ορισμένος, άρα μη-ιδιάζων.
Πράγματι, αν $rank(A) = n$, για $x \neq 0$, $y = Ax \Rightarrow y \neq 0$, συνεπώς $x^t A^t Ax = y^t y = \|y\|^2 > 0$. Άρα $A^t A$ θετικά ορισμένος. Επίσης, αν $rankA \neq n$, υπάρχει $x \neq 0$ τέτοιο ώστε $Ax = 0 \Rightarrow A^t Ax = 0$, άρα ο $A^t A$ όχι θετικά ιδιάζων.

□

Ορισμός: Το σύστημα των εξισώσεων

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

λέγονται κανονικές εξισώσεις (normal equations)

Ορισμός: Ο πίνακας $A^+ = (A^T A)^{-1} A^T$, όταν $A\ \ m \times n\ \ (m \geq n)$ και rank(A)=n ονομάζεται **ψευδοαντίστροφος (pseudoinverse)** του A. Ο ψευδοαντίστροφος αναφέρεται επίσης σαν Moore-Penrose γενικευμένος αντίστροφος του A.

Από τα προηγούμενα φαίνεται ότι η μοναδική λύση ελαχίστων τετραγώνων για το overdetermined σύστημα $Ax = b$ δίνεται από:

$$x = (A^T A)^{-1} A^T b = A^+ b$$

Ο ορισμός του γενικευμένου αντιστρόφου γενικεύει τον συνήθη ορισμό του αντιστρόφου ενός τετραγωνικού πίνακα. Όταν ο πίνακας A είναι τετραγωνικός και αντιστρέψιμος τότε

$$A^+ = (A^T A)^{-1} A^T = A^{-1} (A^T)^{-1} A^T = A^{-1}$$

Ορισμός: Εστω $A\ \ m \times n$ πίνακας και $rank(A) = n$, τότε

$$cond(A) = ||A|| \cdot ||A^+||$$

Παράδειγμα:

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 5 \end{pmatrix}, \quad rank(A) = 2$$

Ετσι ο A έχει full rank

$$A^{+} = (A^T A)^{-1} A^T = \begin{pmatrix} -1.2857 & -0.5714 & 0.8577 \\ 1 & 0.5000 & -0.5000 \end{pmatrix}$$

$$Cond_2(A) = ||A||_2 ||A||_2^{+} = 7.6656 \cdot 2.0487 = 15.7047$$

□

1.2.1 Υπολογιστικές μέθοδοι

Έστω το ΓΠΕΤ $Ax = b$, με A $\in \mathbb{R}^{m\times n}, m \geq n, b \in \mathbb{R}^m$. Η λύση του συστήματος των κανονικών εξισώσεων σε αναλυτική μορφή είναι:

$$x = (A^t A)^{-1} A^t b = A^{+} b.$$

Επειδή η άμεση επίλυση του συστήματος των κανονικών εξισώσεων είναι ασταθής και απαιτεί πολλές πράξεις (άρα είναι μη-υλοποιήσιμη), κυρίως λόγω του υπολογισμού του αντιστρόφου πίνακα $(A^t A)^{-1}$ (διότι το γινόμενο πινάκων δεν είναι ευσταθές, πολλώ δε μάλλον η αντιστροφή αυτού!), χρησιμοποιούμε για απλούστευση των πράξεων και περισσοτερη ευστάθεια υπολογισμών διάφορες μεθόδους παραγοντοποίησης.

Η μέθοδος των κανονικών εξισώσεων

Μία από τις πιο διαδεδομένους μεθόδους (ειδικά στη Στατιστική) για τον υπολογισμό της λύσης ελαχίστων τετραγώνων είναι η μέθοδος των κανονικών εξισώσεων. Η μέθοδος αυτή στηρίζεται στη λύση του συστήματος των κανονικών εξισώσεων.

$$A^T A x = A^T b$$

Εχουμε υποθέσει ότι A είναι $m \times n \quad (m > n)$ και έχει full rank. Επειδή σάυτήν την περίπτωση ο πίνακας $A^T A$ είναι συμμετρικός και θετικά ορισμένος και επομένως μπορούμε να πάρουμε την ανάλυση Cholesky

$$A^T A = H H^T$$

Αλγόριθμος Κανονικών Εξισώσεων

Εστω $A \quad m \times n, \quad m > n$, $rank(A) = n$. Ο παρακάτω αλγόριθμος υπολογίζει τη λύση x ελαχίστων τετραγώνων από τις κανονικές εξισώσεις χρησιμοποιώντας την παραγοντοποίηση Cholesky.

Βήμα 1: Διαμορφώστε $c = A^T b$
Βήμα 2: Υπολογίστε τον παράγονται Cholesky H του $A^T A$.
Βήμα 3: Επιλύστε τα τριγωνικά συστήματα
$Hy = x$ (ως προς y)
$H^T x = y$ (ως προς x)

Πολυπλοκότητα: Για τον υπολογισμό των ποσοτήτων $A^T A$ και $A^T b$ απαιτούνται $mn^2/2$ φλοπς, για τον υπολογισμό της ανάλυσης Cholesky απαιτούνται $n^3/6$ flops. Ετσι συνολικά έχουμε: $mn^2/2 + n^3/6$ flops οπότε η μέθοδος είναι αρκετά αποτελεσματική.

1.2.2 Householder QR παραγοντοποίηση

Στο πρόβλημα ελαχίστων τετραγώνων θέλουμε να βρούμε το x έτσι ώστε η έκφραση $\|b - Ax\|$ να έχει ελάχιστο. Από το Θεώρημα **;;** προκύπτει ότι για κάθε ορθογώνιο πίνακα H θα ισχύει: $\|b-Ax\| = \|H(b-Ax)\| = \|Hb-HAx\|$. Συνεπώς το πρόβλημα μετασχηματίζεται στην εύρεση κατάλληλου πίνακα H για τον οποίο το HA έχει την απλούστερη δυνατή μορφή και συνεπώς η λύση θα προκύπτει εύκολα.

1. Αν ο A είναι full-rank
   Έστω $A = QR \Rightarrow Q^t A = R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ η QR παραγοντοποίηση του A.

   Θεωρούμε ότι $Q^t b = \begin{bmatrix} c \\ d \end{bmatrix}$.
   Επειδή η Ευκλείδεια νόρμα είναι ορθογώνια αναλλοίωτη, πολλαπλασιάζουμε με τον Q^t στη νόρμα του υπολοίπου κι έχουμε:

$$\|r(x)\|^2 = \|Ax - b\|^2 = \|Q^t Ax - Q^t b\|^2 = \left\| \begin{matrix} R_1 x - c \\ d \end{matrix} \right\|^2 .$$

   Συνεπώς, η νόρμα του υπολοίπου θα ελαχιστοποιείται αν επιλέξουμε x τέτοιο ώστε $R_1 x - c = 0$ (και τότε θα ισχύει $\|r\| = \|d\|$), οπότε η λύση του ΓΠΕΤ προκύπτει από την επίλυση του $n \times n$ άνω-τριγωνικού συστήματος $R_1 x = c$.

Αλγόριθμος QR full rank

Εστω $A \;\; m \times n, \;\; m > n$, $rank(A) = n$. Ο παρακάτω αλγόριθμος υπολογίζει τη λύση x ελαχίστων τετραγώνων χρησιμοποιώντας την παραγοντοποίηση QR.

Βήμα 1: Υπολογίζουμε την QR παραγοντοποίηση του A: $A = QR$.
Βήμα 2: Διαμορφώνουμε την ποσότητα: $Q^t b = \begin{bmatrix} c \\ d \end{bmatrix}$.
Βήμα 3: Επιλύουμε το $n \times n$ άνω-τριγωνικό σύστημα:

$R_1 x = c$, όπου $R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$.

Πολυπλοκότητα : Για τον υπολογισμό της QR παραγοντοποίησης απαιτούνται $n^2(m - n/3) flops$, για τον υπολογισμό της ποσότητας $Q^t b$ απαιτούνται $m^2 flops$, και για την επίλυση του άνω-τριγωνικού συστήματος $O(n^2) flops$. Έτσι συνολικά έχουμε: $mn^2/2 - n^3/6 + m^2 + n^2 flops$, οπότε η μέθοδος είναι αρκετά αποτελεσματική.

2. Αν ο A είναι rank-deficient

Έστω $A \in \mathbb{R}^{m \times n}, m \geq n, rank(A) = r < n$. Χρησιμοποιώντας την QR παραγοντοποίηση με μερική οδήγηση έχουμε ότι:

$$Q^t A P = \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}$$

όπου ο P είναι ο μεταθετικός πίνακας και ο R_{11} είναι $r \times r$ άνω-τριγωνικός πίνακας.
Επειδή $PP^t = I_n$ και επειδή η Ευκλείδεια νόρμα είναι ορθογώνια αναλλοίωτη, έχουμε:

$$\|r(x)\|^2 = \|Ax - b\|^2 = \|Q^t(Ax - b)\|^2 = \|Q^t APP^t x - Q^t b\|^2.$$

Θέτουμε $P^t x = y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ και $Q^t b = \begin{bmatrix} c \\ d \end{bmatrix}$, οπότε έχουμε:

$$\|r(x)\|^2 = \left\| \begin{pmatrix} R_{11} & R_{12} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} c \\ d \end{pmatrix} \right\|^2 = \|R_{11} y_1 + R_{12} y_2 - c\|^2 + \|d\|^2.$$

Έτσι η $\|r\|$ ελαχιστοποιείται εάν επιλέξουμε y τέτοιο ώστε: $R_{11} y_1 = c - R_{12} y_2$.

Για την επίλυση αυτού του άνω-τριγωνικού συστήματος, επιλέγουμε αυθαίρετα το διάνυσμα y_2 και υπολογίζουμε το y_1. Τέλος, υπολογίζουμε τη λύση x του ΓΠΕΤ από τη σχέση:

$$x = Py \quad , \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Επειδή η επιλογή του y_2 είναι αυθαίρετη, έχουμε άπειρες λύσεις. Η λύση που προκύπτει αν θέσουμε $y_2 = 0$ λέγεται βασική λύση, και αν επιπλέον R_{12}, έχουμε λύση ελαχίστης νόρμας.

Αλγόριθμος QR rank deficient

Εστω $A \;\; m \times n, \;\; m > n, rank(A) = r < n$. Ο παρακάτω αλγόριθμος υπολογίζει τη λύση x ελαχίστων τετραγώνων χρησιμοποιώντας την παραγοντοποίηση QR.

Βήμα 1: Χρησιμοποιώντας QR με μερική οδήγηση υπολογίζουμε την παραγοντοποίηση:
$$AP = Q \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}.$$
Βήμα 2:Διαμορφώνουμε $Q^t b = \begin{bmatrix} c \\ d \end{bmatrix}$.
Βήμα 3: Επιλέγουμε ένα αυθαίρετο διάνυσμα y_2.
Βήμα 4: Επιλύουμε το $r \times r$ μη-ιδιάζον. άνω-τριγωνικό σύστημα $R_{11}y_1 = c - R_{12}y_2$.
Βήμα 5: Υπολογίζουμε το x από τη σχέση: $x = Py \quad , \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$.

Πολυπλοκότητα : Για τον υπολογισμό της QR με οδήγηση κατά στήλες απαιτούνται $2mnr^2 - r^2(m+n) + 2r^3/3flops$, για τον υπολογισμό της ποσότητας $Q^t b$ απαιτούνται $m^2 flops$, και για την επίλυση του άνω-τριγωνικού συστήματος $O(r^2)flops$. Τέλος, για τον υπολογισμό του x απαιτούνται $O(n)flops$. Έτσι συνολικά έχουμε: $O(mnr^2)flops$, οπότε η μέθοδος είναι λιγότερο αποτελεσματική.

Λύση ελαχίστης νόρμας : Επειδή η επιλογή του είναι αυθαίρετη, η προηγούμενη μέθοδος έχει άπειρες λύσεις. Η λύση που προκύπτει εάν θέσουμε $y_2 = 0$ ονομάζεται βασική λύση. Στην περίπτωση όπου R_{12} η βασική λύση είναι λύση ελαχίστων τετραγώνων ελάχιστης νόρμας (δηλαδή ανάμεσα στις άπειρες λύσεις ελαχίστων τετραγώνων, αυτή έχει ελάχιστη νόρμα).

MGS ορθογωνοποίηση

Η μέθοδος MGS είναι αριθμητικά ισοδύναμη με την εφαρμογή παραγοντοποίησης QR με μετασχηματισμούς Householder στον A. Εφαρμόζουμε τη μέθοδο MGS στον πίνακα $[A, b]$, ώστε το q_{n+1} ορθογώνιο ως προς τον Q:

$[A,b]=[Q,q_{n+1}]\begin{bmatrix} R & z \\ 0 & p \end{bmatrix}$

$\Rightarrow r(x)=Ax-b=[A,b]\begin{bmatrix} x \\ -1 \end{bmatrix}=[Q,q_{n+1}]\begin{bmatrix} R & z \\ 0 & p \end{bmatrix}\begin{bmatrix} x \\ -1 \end{bmatrix}=Q(Rx-z)-pq_{n+1}$

όπου $Q\in\mathbb{R}^{m\times n}, R\in\mathbb{R}^{n\times n}$ οι παράγοντες της QR παραγοντοποίησης του A,

και $q_{n+1}\in\mathbb{R}^m, p\in\mathbb{R}$ και $z\in\mathbb{R}^n, z=\begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$.

Εφόσον το q_{n+1} είναι ορθογώνιο στο Q, τότε η $\|r\|$ ελαχιστοποιείται όταν $Rx=z$. Έτσι το ΓΠΕΤ επιλύεται προσδιορίζοντας τη λύση του $Rx=z$ και το αντίστοιχο υπόλοιπο θα είναι το $r=pq_{n+1}$.

Αλγόριθμος MGS

Εστω A $m\times n$, $m>n$, $rank(A)=r<n$. Ο παρακάτω αλγόριθμος υπολογίζει τη λύση x ελαχίστων τετραγώνων χρησιμοποιώντας την ορθογωνοποίηση MGS.

Βήμα 1: Εφαρμόζουμε MGS στον A και προσδιορίζουμε $Q=[q_1,...,q_n]$ και R.
Βήμα 2:Εφαρμογή MGS μετασχηματισμών στο b:

Για $k=1,...,n$: $\begin{array}{l} \delta_k=q_k^t b \\ b=b-\delta_k q_k \end{array}$ Βήμα 3: Επιλύουμε το σύστημα: $Rx=(\delta_1,...,\delta_n)^t$.

Πολυπλοκότητα : Η μέθοδος αυτή απαιτεί $mn^2 flops$ για τον υπολογισμό των Q,R (περισσότερα από την QR με μετασχηματισμούς Householder), $2mnflops$ για το μετασχηματισμό του b και $O(n^2)flops$ για την επίλυση του άνω-τριγωνικού συστήματος.

SVD παραγοντοποίηση

Έστω $A=U\Sigma V^t$ η SVD του $m\times n$ πίνακα A. Θέλουμε να ελαχιστοποιήσουμε τη νόρμα του υπολοίπου:

$$\begin{aligned}\|r(x)\| &= \|Ax-b\| \\ &= \|U\Sigma V^t x-b\| \\ &= \|U\Sigma V^t x-UU^t b\| \\ &= \|U(\Sigma V^t x-U^t b)\| \\ &= \|\Sigma y-b'\|\end{aligned}$$

όπου: $V^t x=y$ και $U^t b=b'$.

Έτσι με τη χρήση της SVD το ΓΠΕΤ παίρνει την ακόλουθη μορφή:
"Προσδιορισμός του y ώστε να ελαχιστοποιείται η $\|\Sigma y - b'\|$."

Έστω k το πλήθος των μη-μηδενικών ιδιαζουσών τιμών του A. Τότε:

$$\|\Sigma y - b'\| = (\sum_{i=1}^{k} |\sigma_i y_i - b'_i|^2 + \sum_{i=k+1}^{m} |b'_i|^2)^{\frac{1}{2}}$$

Οπότε, το διάνυσμα $y = (y_1, ...y_n)^t$ θα δίνεται από τη σχέση:

$$y_i = \begin{cases} \frac{b'_i}{\sigma_i}, \sigma_i \neq 0 \\ \text{αυθαίρετο}, \sigma_i = 0 \end{cases}$$

- Αν $k = n$, δηλ. ο A είναι full rank τότε η λύση ελαχίστων τετραγώνων είναι μοναδική.
- Αν $k < n$, δηλ. ο A είναι rank-deficient τότε:

- οι $y_{k+1}, ..., y_n$ δεν εμφανίζονται στο $\|r\|$, άρα δεν το επηρεάζουν, οπότε μπορούμε να τους δώσουμε αυθαίρετες τιμές
- έχουμε άπειρες λύσεις ελαχίστων τετραγώνων.

Υπολογίζουμε τελικώς τη λύση ελαχίστων τετραγώνων ως εξής: $x = Vy$.

Αν ο A είναι rank-deficient, από την άπειρη οικογένεια των λύσεων, μας ενδιαφέρει η λύση ελαχίστης νόρμας. Αυτή προκύπτει θέτοντας $y_i = 0, i = k+1, ..., n$ και άρα τότε:

$$x = \sum_{i=1}^{k} \frac{u_i^t b_i}{\sigma_i} v_i.$$

Αλγόριθμος SVD

Εστω $A \ \ m \times n, \ \ m > n$, $rank(A) = r < n$. Ο παρακάτω αλγόριθμος υπολογίζει τη λύση x ελαχίστων τετραγώνων χρησιμοποιώντας την παραγοντοποίηση SVD.

Βήμα 1: Προσδιορίζουμε την SVD του A: $A = U\Sigma V^t$.
Βήμα 2: Δημιουργούμε το διάνυσμα: $b' = U^t b = [b'_1, b'_2, ..., b'_m]^t$.
Βήμα 3: Υπολογίζουμε τη λύση y:
$y_i = \frac{b'_i}{\sigma_i}, \sigma_i \neq 0$, $y_i =$ αυθαίρετο, $\sigma_i = 0$
Βήμα 4: Υπολογίζουμε την οικογένεια των λύσεων ελαχίστων τετραγώνων: $x = Vy$.

1.4 Ασκήσεις

Ασκηση 1

Εστω ένα σύστημα αριθμητικής κινητής υποδιαστολής με $\beta = 10$ και mantissa 8 ψηφίων. Εάν προσπαθήσουμε να λύσουμε το ΓΠΕΤ με τη μέθοδο των κανονικών εξισώσεων θα δούμε ότι θα οδηγηθούμε σε αδιέξοδο.

Έστω $A = \begin{bmatrix} 1 & 1 \\ 10^{-4} & 0 \\ 0 & 10^{-4} \end{bmatrix}$

Παρατηρούμε ότι rank(A)=2. Δημιουργούμε το γινόμενο $A^T A$.

$A^T A = \begin{bmatrix} 1+10^{-8} & 1 \\ 1 & 1+10^{-8} \end{bmatrix}$

Παρατηρούμε ότι και πάλι rank $(A^T A) = 2$. Εάν όμως υπολογίσουμε το παραπάνω γινόμενο στη δοσμένη αριθμητική κινητής υποδιαστολής θα έχουμε. $fl(A^T A) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

Παρατηρούμε ότι τώρα rank (fl $(A^T A)) = 1$, δηλαδή ο πίνακας είναι ιδιάζων και επομένως δεν μπορούμε να προχωρήσουμε σε επίλυση του συστήματος των κανονικών εξισώσεων. Υπενθυμίζουμε ότι αυτό συνέβη λόγω της εκτέλεσης της πράξης :

$x_1 = 1, x_2 = 10^{-8}, \quad fl(x_1 + x_2)$:

$$\begin{array}{rl} 0.10000000 & \times 10 \\ +\ 0.000000001 & \times 10 \\ \hline 0.100000001 & \times 10 \end{array}$$

$\rightarrow$ (στρογγύλευση 0.10000000×10

Άρα $fl(x_1 + x_2) = 1$

□

Ασκηση 2

Να επιλυθεί το ΓΠΕΤ όπου

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} \text{ και } b = \begin{bmatrix} 3 \\ 5 \\ 9 \end{bmatrix}.$$

Βήμα 1: Υπολογίζουμε την παραγοντοποίηση $A = QR$

$$Q = \begin{bmatrix} -0.2673 & 0.8724 & 0.4082 \\ -0.5345 & 0.2182 & -0.8165 \\ -0.8018 & -0.4364 & 0.4082 \end{bmatrix}$$

$$R = \begin{bmatrix} -3.7417 & -5.3452 \\ 0 & 0.6547 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} R_1 \\ \emptyset \end{bmatrix}$$

Βήμα 2: Υπολογίζουμε την ποσότητα $Q^T b$

$Q^T b = \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} -10.6904 \\ -0.2182 \\ 0.8165 \end{bmatrix}$

Βήμα 3: Επιλύουμε το σύστημα $R_1 x = c$

$$\begin{bmatrix} -3.7417 & -5.3452 \\ 0 & 0.6547 \end{bmatrix} \cdot x = \begin{matrix} -10.6904 \\ -0.2182 \\ 0.8165 \end{matrix}$$

Η λύση ελαχίστων τετραγώνων είναι

$$x = \begin{matrix} 3.3332 \\ -0.3333 \end{matrix}$$

Η νόρμα του υπολοίπου είναι $||r||_2 = ||d||_2 = 0.8165$.

□

Ασκηση 3

Να επιλυθεί το ΓΠΕΤ όπου

$A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & \end{bmatrix}$ και $b = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$.

Παρατηρούμε ότι rank(A)=1. Επομένως πρέπει να εφαρμόσουμε QR με οδήγηση κατά στήλες.

Βήμα 1: Υπολογίζουμε την παραγοντοποίηση $AP = QR$

$Q = \begin{bmatrix} -0.4472 & -0.8944 & 0 \\ -0.8944 & -0.4472 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$R = \begin{bmatrix} -2.2361 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Βήμα 2: Υπολογίζουμε την ποσότητα $Q^T b$

$Q^T b = \begin{bmatrix} -6.7082 \\ 0 \\ 0 \end{bmatrix}, c = (-6.7082)$

Βήμα 3: Επιλέγουμε $y_2 = 0$

Βήμα 3: Επιλύουμε το σύστημα $R_{11} y_1 = c - R12 y_2$

$$y_1 = \frac{c}{R_{11}} = \frac{-6.7082}{-2.2361} = 3$$

Η λύση ελαχίστης νόρμας είναι

$$x = y = \begin{matrix} 3 \\ 0 \end{matrix}$$

□

Ασκηση 4
Να επιλυθεί το ΓΠΕΤ όπου

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 2 & 3 \end{pmatrix}, \ \underline{b} = \begin{pmatrix} 6 \\ 9 \\ 6 \end{pmatrix}$$

Βήμα 1: $\sigma_1 = 7.5358,\ \sigma_2 = 0.4957,\ \sigma_3 = 0$
O A είναι rank deficient

$$U = \begin{pmatrix} 0.4956 & 0.5044 & 0.7071 \\ 0.7133 & -0.7008 & -0.0000 \\ 0.4956 & 0.5044 & -0.7071 \end{pmatrix}$$

$$V = \begin{pmatrix} 0.3208 & -0.8546 & 0.4082 \\ 0.5470 & -0.1847 & -0.8165 \\ 0.7732 & 0.4853 & 0.4082 \end{pmatrix}$$

Βήμα 2: $\underline{b}' = U^T\underline{b} = [12.3667, -0.2547, 0]^T$

Βήμα 3: $\underline{y} = [1.6411, -0.5541, 0]^T$

Βήμα 4: Η ελάχιστη ως προς τη 2-νόρμα λύση ελαχίστων τετραγώνων
δίνεται από
$\underline{x} = V\underline{y} = [1, 1, 1]^T$

□