2η Περίπτωση: Αν το Γραμμικό σύστημα είναι underdetermined

Θεωρούμε $Ax = b$, όπου $A \in \mathbb{R}^{m \times n}$, $m < n$ και $b \in \mathbb{R}^m$. Σε αυτό το σύστημα ο αριθμός των αγνώστων είναι μεγαλύτερος από τον αριθμό των εξισώσεων και η λύση x δεν είναι συγκεκριμένη, με την έννοια ότι πολλές επιλογές για το x αντιστοιχούν στο ίδιο δεξί μέλος b.

Αν ο A είναι full rank ($rankA = m$), τότε για κάθε $b \in \mathbb{R}^m$ υπάρχει ένα σύνολο λύσεων, της μορφής: $\{x \in \mathbb{R}^n : Ax = b\} = \{x_p - z,\, z \in N(A)\}$, όπου x_p είναι οποιαδήποτε ειδική λύση, δηλ. $Ax_p = b$, και $\forall z : Az = 0$. Συνεπώς, η λύση έχει $dimN(A) = n - m$ 'βαθμούς ελευθερίας', και μπορεί κανείς να επιλέξει οποιοδήποτε $z \in N(A)$ για να ικανοποιηθούν διαφορετικές ιδιότητες της λύσης ή για βελτιστοποίηση ανάμεσα στις λύσεις. Εμείς ψάχνουμε από αυτές τη λύση x^*, για την οποία η $\|x\|$ ελαχιστοποιείται.

Θεώρημα 1.1. *Έστω A $m \times n$ πίνακας, $m < n$, με $rank(A) = m$ και έστω x^* μια λύση του $Ax = b$. Τότε η x^* είναι λύση ελαχίστης νόρμας αν και μόνο αν $x^* \in R(A^t)$.*

Απόδειξη: Αρχικά, μετασχηματίζουμε το πρόβλημα της εύρεσης ενός διανύσματος x^* που ελαχιστοποιεί τη $\|x\|$ για όλα τα διανύσματα x για τα οποία $Ax = b$, στο πρόβλημα εύρεσης του διανύσματος $z^* \in N(A)$ που ελαχιστοποιεί το $\|x_p - z\|$.

Έστω $x = x_p - z$ μια λύση του $Ax = b$.

Από τη Γραμμική Άλγεβρα γνωρίζουμε ότι αν ένας πίνακας X έχει n γραμμικά ανεξάρτητες στήλες, τότε αυτές σχηματίζουν μια βάση για τον $R(X)$. Επίσης, για δοθέν διάνυσμα $v \in R(X)$, το διάνυσμα y για το οποίο ισχύει $v = Xy$ είναι μοναδικό και δίνεται από τη σχέση $y = X^+ v$.

Από το Θεμελιώδες Θεώρημα των Γραμμικών Απεικονίσεων (βλ. [16]) έχουμε ότι ο πυρήνας του A έχει διάσταση $dimN(A) = n - rank(A) = n - m$. Μπορούμε λοιπόν να θεωρήσουμε μια ορθογώνια βάση $\{b_1, ..., b_{n-m}\}$, $b_i \in \mathbb{R}^n$ για τον $N(A)$, οπότε $N(A) = span\{b_1, ..., b_{n-m}\}$. Αν ορίσουμε πίνακα $B = (b_1, \ldots, b_{n-m})$ έχουμε ότι $R(B) = N(A)$. Θεωρούμε $z \in N(A)$, οπότε $z = By$, $y \in \mathbb{R}^{n-m}$. Το πρόβλημα εύρεσης του z^* ανάγεται τώρα στο πρόβλημα εύρεσης διανύσματος y^* ώστε: $\|x_p - By^*\| \leq \|x_p - By\|$, $\forall y \in \mathbb{R}^{n-m}$. Αυτό αποτελεί στην ουσία ένα overdetermined πρόβλημα ελαχίστων τετραγώνων, με λύση $y^* = B^+ x_p = (B^t B)^{-1} B^t x_p$. Άρα $z^* = By^* = BB^+ x_p$.

Τώρα, για το $x_p - z^*$ έχουμε:

$B^t(x_p - z^*) = B^t x_p - B^t BB^+ x_p = B^t x_p - B^t B(B^t B)^{-1} B^t x_p = B^t x_p - B^t x_p = 0$.

Επομένως το $x_p - z^*$ είναι ορθογώνιο ως προς τις στήλες του B, άρα και ως προς κάθε διάνυσμα του $N(A) = R(B)$. Συνεπώς, για το $x^* = x_p - z^*$ ισχύει $x^* \in (N(A))^\perp$. Ξέρουμε ότι $(N(A))^\perp = R(A^t)$, άρα έχουμε το ζητούμενο.

Αυτό το Θεώρημα προσφέρει μία μέθοδο για την απόκτηση της λύσης ελαχίστης νόρμας του $Ax = b$.

Πράγματι, επειδή $x^* \in R(A^t)$, έχουμε $x^* = A^t w$ για κάποιο $w \in \mathbb{R}^m$. Άρα το w είναι η λύση του συστήματος

$$AA^t w = b. \tag{1.3}$$

Ο πίνακας AA^t είναι $m \times m$, και επειδή ο A είναι full rank, ο AA^t είναι αντιστρέψιμος. Συνεπώς, μπορούμε να λύσουμε αυτό το σύστημα και να αποκτήσουμε τη μοναδική του λύση w, και ύστερα απλά να υπολογίσουμε τη x^* ως εξής: $x^* = A^t w$. Συνοψίζοντας έχουμε ότι η

$$x^* = A^t (AA^t)^{-1} b \tag{1.4}$$

είναι η λύση του underdetermined προβλήματος ελαχίστων τετραγώνων, και οι **κανονικές εξισώσεις** για αυτό το πρόβλημα είναι οι

$$x = A^t w, \quad \textit{όπου} \quad w = (AA^t)^{-1} b. \tag{1.5}$$

1.3 Τρόποι επίλυσης Underdetermined Προβλημάτων Ελαχίστων Τετραγώνων

Είδαμε προηγουμένως πώς βρίσκουμε το διάνυσμα x που προσεγγίζει, κατά το δυνατόν, τη λύση του overdetermined συστήματος εξισώσεων $Ax = b$, όπου $\mathrm{A} \in \mathbb{R}^{m\times n}, m \geq n$ με γραμμικά ανεξάρτητες στήλες. Τώρα θα θεωρήσουμε το underdetermined σύστημα $Ax = b$, όπου $\mathrm{A} \in \mathbb{R}^{m\times n}, m < n$ με γραμμικά ανεξάρτητες γραμμές. Αυτό το underdetermined σύστημα έχει άπειρες λύσεις, εμείς όμως ψάχνουμε από αυτές τη λύση x^*, για την οποία η $\|x\|$ ελαχιστοποιείται.
Επειδή ο άμεσος υπολογισμός της λύσης $x^* = A^t(AA^t)^{-1}b$ είναι ασταθής και απαιτεί πολλές πράξεις, κυρίως λόγω του υπολογισμού του αντιστρόφου πίνακα $(AA^t)^{-1}$, χρησιμοποιούμε διάφορες μεθόδους παραγοντοποίησης, όπως στην περίπτωση του overdetermined προβλήματος. Αυτή τη φορά παραγοντοποιούμε τον A^t και εργαζόμαστε με παρόμοιο τρόπο σε κάθε περίπτωση:

1.3.1 Η μέθοδος των κανονικών εξισώσεων με τη βοήθεια της ανάλυσης Cholesky

Στην περίπτωση που ο **A είναι full-rank**, ο AA^t είναι συμμετρικός και θετικά ορισμένος και άρα μπορούμε να πάρουμε την ανάλυση Cholesky αυτού: $AA^t = GG^t$. Επομένως, το (1.5) γίνεται $GG^tw = b$ και θέτοντας $z = G^tw$, μπορούμε να υπολογίσουμε τη λύση x^* ως εξής:

Αλγόριθμος 1.3.1 Έστω A $m \times n$, $m < n$, $rank(A) = m$.

- Υπολογίζουμε τον παράγοντα Cholesky G του AA^t.
- Επιλύουμε τα τριγωνικά συστήματα: $Gz = b$ (ως προς z) και $G^tw = z$ (ως προς w).
- Υπολογίζουμε το x^* από τη σχέση $x^* = A^tw$.

1.3.2 Υπολογισμός λύσης ελαχίστων τετραγώνων με τη βοήθεια της QR παραγοντοποίησης

χρησιμοποιούμε την QR-παραγοντοποίηση του A^t για την ανάλυση του συστήματός μας, ως εξής:

$$A^t = QR = Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \tag{1.6}$$

όπου ο $Q \in \mathbb{R}^{n\times n}$ είναι ορθογώνιος και ο $R_1 \in \mathbb{R}^{m\times m}$ είναι άνω τριγωνικός.
Έχουμε $b = Ax = \begin{bmatrix} R_1^t & 0 \end{bmatrix} Q^t x = R_1^t y_1$, όπου $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = Q^t x$, $y_1 \in \mathbb{R}^m$.
Αν ο A είναι full rank , τότε το $y_1 = (R_1^t)^{-1} b$ είναι μοναδικά προσδιορισμένο και όλες οι λύσεις του $Ax = b$ δίνονται από:

$$x = Q \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \tag{1.7}$$

όπου $y_2 \in \mathbb{R}^{n-m}$ αυθαίρετο.

Θέτοντας $y_2 = 0$ παίρνουμε τη μοναδική λύση x^* που ελαχιστοποιεί τη $\|x\|$.
Πράγματι, έχουμε

$$x^* = Q \begin{bmatrix} {R_1}^{-t} b \\ 0 \end{bmatrix} = Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} {R_1}^{-1} {R_1}^{-t} b = Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} ({R_1}^t R_1)^{-1} b = A^t (AA^t)^{-1} b = A^+ b, \tag{1.8}$$

όπου ο $A^+ = A^t(AA^t)^{-1}$ είναι ο ψευδοαντίστροφος του A, διότι $AA^t = \begin{bmatrix} R_1^t & 0 \end{bmatrix} Q^t Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = {R_1}^t R_1$. Συνεπώς μπορούμε να γράψουμε $x^* = A^t w$, όπου w είναι η λύση των κανονικών εξισώσεων $AA^t w = b$.
Η εξίσωση $x^* = Q \begin{bmatrix} {R_1}^{-t} b \\ 0 \end{bmatrix}$ μας δίνει έναν τρόπο υπολογισμού της x^*. Εναλλακτικά, προκειμένου να αποφύγουμε την αποθήκευση του Q, μπορούμε να χρησιμοποιήσουμε την QR-παραγοντοποίηση (1.6) και να υπολογίσουμε το $x^* = A^t w$, με w λύση της $R^t R w = R_1^t R_1 w = b$.

Τότε έχουμε τον εξής αλγόριθμο:

Αλγόριθμος 1.3.2 Έστω A $m \times n$, $m < n$, $rank(A) = m$.

- Υπολογίζουμε την QR-παραγοντοποίηση του A^t.

- Για να υπολογίσουμε τη λύση w της $R^tRw = b$, επιλύουμε τα εξής επιμέρους τριγωνικά συστήματα: $R_1^t y = b$ (ως προς y) και $R_1 w = y$ (ως προς w).

- Υπολογίζουμε το x^* από τη σχέση $x^* = A^t w$.

1.3.3 Υπολογισμός λύσης ελαχίστων τετραγώνων με τη βοήθεια της SVD παραγοντοποίησης

Το βασικό πλεονέκτημα αυτής της μεθόδου επίλυσης underdetermined ΓΠΕΤ, είναι ότι δουλεύει και στις περιπτώσεις που ο A είναι rank-deficient , σε αντίθεση με τις άλλες δύο μεθόδους που περιγράψαμε.

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

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

όπου: $U^t x = y$ και $V^t b = b'$.

Έτσι με τη χρήση της SVD το ΓΠΕΤ παίρνει την ακόλουθη μορφή:

"Προσδιορισμός του y ώστε να ελαχιστοποιείται η $\|\Sigma^t y - b'\|$."

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

$$\|\Sigma^t y - b'\| = (\sum_{i=1}^{k} |\sigma_i y_i - b_i'|^2 + \sum_{i=k+1}^{m} |b_i'|^2)^{\frac{1}{2}}.$$

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

$$y_i = \begin{cases} \frac{b_i'}{\sigma_i}, \quad 1 \leq i \leq k \\ \text{αυθαίρετο}, \; k < i \leq m \end{cases}$$

Τέλος, υπολογίζουμε το x ως εξής: $x = Uy$.

Λόγω των αυθαίρετων τιμών στα στοιχεία του y, προκύπτουν άπειρες λύσεις ελαχίστων τετραγώνων του συστήματος $Ax = b$. Από αυτές μας ενδιαφέρει η λύση ελαχίστης νόρμας, η οποία προκύπτει θέτοντας $y_i = 0, i = k+1, ..., m$. Τότε έχουμε:

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

Αλγόριθμος 1.3.3

- Προσδιορίζουμε την SVD του A^t: $A^t = U\Sigma V^t$.

- Δημιουργούμε το διάνυσμα: $b' = V^t b = [b_1', b_2', ..., b_m']^t$.

- Υπολογίζουμε το y:

$$y_i = \begin{cases} \frac{b_i'}{\sigma_i}, \quad 1 \leq i \leq k \\ 0, \; k < i \leq m \end{cases}$$

- Υπολογίζουμε τη λύση ελαχίστων τετραγώνων: $x = Uy$.