5. Επίλυση Γραμμικών Συστημάτων

Στο κεφάλαιο αυτό θα παρουσιάσουμε μεθόδους επίλυσης γραμμικών συστημάτων. Η επίλυση γραμμικών συστημάτων κατέχει ιδιαίτερη θέση στην υπολογιστική μηχανική διότι στην πλειονότητα των προβλημάτων καταλήγουμε σε ένα μεγάλου μεγέθους γραμμικό σύστημα του οποίου αναζητούμε την ακριβή ή προσεγγιστική λύση. Η λύση των γραμμικών συστημάτων δεν απαιτείται μόνο για γραμμικά προβλήματα, όπως π.χ. η αριθμητική λύση της εξίσωσης Laplace για την εύρεση της κατανομής της θερμότητας ή του ηλεκτροστατικού πεδίου αλλά χρησιμοποιείται και κατά την αριθμητική επίλυση πιο περίπλοκων γραμμικών προβλημάτων ή μη γραμμικών προβλημάτων όπως είναι η εύρεση κατανομής τάσεων και παραμορφώσεων (επίλυση των εξισώσεων γραμμικής ελαστικότητας) και ο υπολογισμός του πεδίου ροής (επίλυση των μη-γραμμικών εξισώσεις ροής Euler και Navier–Stokes).

Η γενική μορφή του γραμμικού συστήματος είναι

$$\begin{aligned}
&a_{11}\ x_1 + a_{12}\ x_2 + ... + a_{1n}x_n = b_1 \\
&a_{21}\ x_1 + a_{22}\ x_2 + ... + a_{2n}\ x_n = b_2 \\
&.. \\
&a_{n1}\ x_1\ a_{n2}\ x_2\ + ... + a_{nn}\ x_n = b_n
\end{aligned} \tag{5.1}$$

ή

$$\begin{bmatrix} a_{11} & a_{12} & .... & a_{1n} \\ a_{21} & a_{22} & .... & a_{2n} \\ & & & \\ a_{n1} & a_{n2} & .... & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ . \\ b_n \end{bmatrix}$$

που γράφεται σε συνεπτυγμένη μορφή ως

$$\left[\mathrm{A}\right]\vec{\mathbf{x}} = \vec{\mathbf{b}} \tag{5.2}$$

όπου $\vec{\mathbf{x}} = \left[x_1, x_2, ..., x_n\right]^T$ είναι το διάνυσμα των αγνώστων ή η λύση, $\vec{\mathbf{b}} = \left[b_1, b_2, ..., b_n\right]^T$ είναι το γνωστό διάνυσμα του δεξιού μέλους, και $A = \left(a_{ii}\right)$ είναι το τετραγωνικό μητρώο συντελεστών.

Υποθέτοντας ότι το σφάλμα στρογγυλοποίησης κατά την διαδικασία επίλυσης είναι αμελητέο υπάρχουν μέθοδοι για την εύρεση της "ακριβούς" λύσης του συστήματος (5.1) που απαιτούν ένα πεπερασμένο αριθμό υπολογισμών. Οι μέθοδοι που μας δίνουν την "ακριβή" αριθμητική λύση του συστήματος ονομάζονται ευθείς μέθοδοι (direct methods) και έχουν πρακτική σημασία όταν ο αριθμός των εξισώσεων δεν είναι μεγάλος. Τυπικό μέγεθος των συστημάτων που επιλύονται με ευθείς μεθόδους είναι συστήματα 100 εξισώσεων. Αρχικά θα παρουσιάσουμε τις ευθείς μεθόδους. Συστήματα μεγαλύτερου μεγέθους, με 1000, 10000 ή ακόμη με100000 εξισώσεις ή περισσότερες λύονται προσεγγιστικά με την εφαρμογή επαναληπτικών μεθόδων που θα παρουσιασθούν στο τέλος του κεφαλαίου.

Η επίλυση γραμμικών συστημάτων απαιτεί ουσιαστικά μετασχηματισμούς μητρώων και προϋποθέτει βασικές γνώσεις γραμμικής άλγεβρας. Η θεωρία μητρώων που αποτελεί απαραίτητο υπόβαθρο για την κατανόηση των μεθόδων επίλυσης γραμμικών συστημάτων ανασκοπείται παρακάτω.

5.1 Στοιχεία θεωρίας μητρώων

Τα στοιχεία ενός μητρώου A με n στήλες και m γραμμές συμβολίζονται ως a_{ij}. Το άθροισμα δυο μητρώων $\mathrm{A} = \left(a_{ij}\right)$ $i = 1, \ldots\ldots, n$ και $j = 1, \ldots, m$ ή διαστάσεως $m \times n$ και $\mathrm{B} = b_{ij}$ επίσης διαστάσεως $m \times n$ είναι ένα μητρώο C διαστάσεως $m \times n$

$$A + B = C = \left(C_{ij}\right) = \left(a_{ij} + b_{ij}\right) \tag{5.1.3}$$

Το γινόμενο δυο μητρώων A διαστάσεων $m \times n$ και B διαστάσεων $n \times p$ είναι ένα μητρώο C διαστάσεων $m \times p$ που έχει στοιχεία $c_{ij} = \sum_{k=1}^{n} a_{ik} \, b_{kj}$

$$\underset{n \times m}{\mathrm{A}} \; \underset{m \times p}{\mathrm{B}} = \underset{m \times p}{C} = \left(c_{ij}\right) \tag{5.1.4}$$

Η διαδοχική άθροιση και πολλαπλασιασμού μητρώων δεν έχει πάντοτε την αντιμεταθετική ιδιότητα έστω και αν οι διαστάσεις είναι συμβατές για την εκτέλεση του πολλαπλασιασμού, δηλαδή

$$\begin{aligned} \mathrm{A}\left(\mathrm{B} + C\right) &= \mathrm{AB} + \mathrm{A}C \\ \left(\mathrm{B} + C\right)\mathrm{A} &= \mathrm{B}A + C\mathrm{A} \end{aligned} \tag{5.5}$$

διότι γενικά $\mathrm{AB} \neq \mathrm{BA}$ έστω και όταν $m = n = p$, όταν δε $\mathrm{AB} = \mathrm{BA}$ τοτε τα μητρώα A και B έχουν την αντιμεταθετική ιδιότητα..
Θεωρούμε το μητρώο $\mathrm{A} = \left(a_{ij}\right)$ με στοιχεία μιγαδικούς αριθμούς και συμβολίζουμε με $\overline{a}_{ij}$ τους συζυγείς μιγαδικούς των στοιχείων a_{ij} το μητρώο $\overline{\mathrm{A}} = \left(\overline{\alpha}_{ij}\right)$ είναι το συζυγές του μητρώου Α και ισχύει $\overline{\mathrm{AB}} = \overline{\mathrm{A}}\,\overline{\mathrm{B}}$.

Το μητρώο που προκύπτει από την αντιμετάθεση γραμμών και στηλών του μητρώου Α διαστάσεων $n \times m$ ονομάζεται ανάστροφο (transpose) και συμβολίζεται με A^{T}. Το ανάστροφο μητρώο έχει στοιχεία $b_{ij} = a_{ji}$ και διαστάσεις $m \times n$.
Γενικά ισχύει

$$\left(\mathrm{A} + \mathrm{B}\right)^{T} = A^{T} + B^{T}$$

$$\left(\mathrm{A}B\right)^{T} = B^{T} A^{T} \tag{5.6}$$

$$\left(A\mathrm{B}C\right)^{T} = C^{T} B^{T} A^{T} \quad \kappa.\lambda.\pi.$$

Συζυγής ανάστροφο μητρώο Α είναι το μητρώο $\left(\overline{\mathrm{A}}\right)^{\mathrm{T}} = \mathrm{A}^{*}$ και ισχύει $\left(\overline{\mathrm{A}^{\mathrm{T}}}\right) = \left(\overline{\mathrm{A}}\right)^{\mathrm{T}}$ και $\left(\mathrm{AB}\right)^{*} = \mathrm{B}^{*}\mathrm{A}^{*}$.

Ένα τετραγωνικό μητρώο $\left(n = m\right)$ είναι τύπου Hermite όταν $\mathrm{A}^{*} = A$. Ένα μητρώο είναι συμμετρικό όταν και μόνο όταν $A = A^{T}$ και ένα μητρώο με στοιχεία πραγματικούς αριθμούς είναι τύπου Hermite όταν και μόνον όταν είναι συμμετρικό.

Ένα τετραγωνικό μητρώο $n \times n$ είναι διαγώνιο όταν έχει μη μηδενικά στοιχεία μόνο στην διαγώνιο $d_{ij} = 0,\ i \neq j$. Στα διαγώνια μητρώα συνήθως συμβολίζουμε τα διαγώνια στοιχεία d_{ii} ως d_{i} και το διαγώνιο μητρώο ως $D = \left[d_{1}, d_{2} ... d_{n}\right]$. Το γινόμενο ενός μητρώου $A\left(m \times n\right)$ με το διαγώνιο μητρώο $D\left(n \times n\right)$ έχει διαστάσεις $m \times n$ και στοιχεία $d_{j}\ a_{ij}$. Όταν όλα τα στοιχεία ενός διαγώνιου μητρώου ίδια, *d*, το μητρώο ονομάζεται βαθμωτό μητρώο και ισχύει $AD = dA = Ad$. Όταν η τιμή των στοιχείων της διαγώνιου είναι μονάδα το βαθμωτό μητρώο ονομάζεται μοναδιαίο και ισχύει $AI = IA = A$.
Θεωρούμε τετραγωνικό μητρώο *A* $n \times n$ και ένα μητρώο *B* τέτοιο ώστε *AB = I*. Το μητρώο *B* είναι μοναδικό και ονομάζεται αντίστροφο του *A*

συμβολίζεται ως A^{-1} και ικανοποιεί $\mathrm{A}\mathrm{A}^{-1} = A^{-1}\mathrm{A} = I$. Το τετραγωνικό μητρώο A που έχει αντίστροφο ονομάζεται nonsingular.

Θεωρούμε δυο τετραγωνικά μητρώα A $n \times n$ και Β $n \times n$ που έχουν αντίστροφα A^{-1} και B^{-1} τότε $\left(AB\right)^{-1} = B^{-1}\ A^{-1}$.Το τετραγωνικό μητρώο A είναι μοναδιαίο όταν το αντίστροφό του είναι το ίδιο με το ανάστροφο συζυγές $A^{-1} = A^{*}$ και ορθογώνιο όταν και μόνον όταν $A^{*} = A^{-1} = A^{T}$.

Ένα τετραγωνικό μητρώο στο οποίο όλα τα στοιχεία κάτω από την διαγώνιο είναι μηδενικά ονομάζεται ανω–τριγωνικό (upper–triagonal). Επί πλέον, όταν και τα στοιχεία της διαγωνίου είναι μηδενικά το μητρώο είναι αυστηρά–άνω–τριγωνικό. Οι ορισμοί του κάτω–τριγωνικού και του αυστηρά–κάτω–τριγωνικού είναι αντίστοιχοι.

Στην υπολογιστική μηχανική είναι συχνή η χρήση μητρώων που έχουν στοιχεία μητρώα π.χ.

$$A = \left[\begin{array}{cc|cc|cc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{55} & a_{46} \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix},$$

$$B = \left[\begin{array}{ccc|cc} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} \\ \hline b_{51} & b_{52} & b_{53} & b_{54} & b_{55} \\ b_{61} & b_{62} & b_{63} & b_{64} & b_{65} \end{array}\right] = \begin{bmatrix} B_{11} & B_{22} \\ B_{12} & B_{22} \end{bmatrix} \qquad (5.7)$$

Για κάθε τετραγωνικό μητρώο A που έχει αντίστροφο υπάρχει ένας αριθμός που το χαρακτηρίζει και ονομάζεται ορίζουσα που έχει τις ιδιότητες

$$\det(A) = \det(A^T) \qquad \det(\mathrm{AB}) = \det(\mathrm{A})\det(\mathrm{B})$$

$$\det(\mathrm{A}^{-1}) = 1/\det(\mathrm{A}) \quad \det(\overline{\mathrm{A}}) = \overline{\det(\mathrm{A})} \tag{5.8}$$

$$\det(\mathrm{A}^{*}) = \overline{\det(\mathrm{A})}$$

Η τιμή της ορίζουσας υπολογίζεται εύκολα μόνο για τα διαγώνια μητρώα

$$\det(D) = d_1\, d_2 ....d_n$$

Διαφορετικά ο υπολογισμός της ορίζουσας απαιτεί πλήθος $n!$ πολλαπλασιασμών και προσθέσεων και δίνεται από

$$\det(A) = \sum_{j=1}^{n} \alpha_{pj}\, A_{pj} = \sum_{j=1}^{n} \alpha_{iq}\, A_{i\phi} \tag{5.9}$$

όπου A_{pj} και A_{iq} είναι τα υπομητρώα που αντιστοιχούν στο κάθε στοιχείο a_{ij} του μητρώου *A*.

5.1.1 Διανυσματικοί χώροι ιδιοτιμές και ιδιοδιανύαματα.

Οι στήλες και οι γραμμές ενός μητρώου είναι διανύσματα που ανήκουν σε ένα διανυσματικό χώρο V_n που είναι ένας χώρος που ορίζεται το εσωτερικό γινόμενο $(\vec{u}_1, \vec{u}_2) = u_1^*\, u_2$. Η διάσταση ενός διανυσματικού χώρου V_q, είναι q τότε και μόνο τότε εάν q γραμμικά ανεξάρτητα διανύσματα ανήκουν στον διανυσματικό χώρο και $(q+1)$ ή περισσότερα διανύσματα από αυτόν τον χώρο είναι γραμμικά εξαρτημένα. Σε ένα διανυσματικό χώρο διάστασης q κάθε διάνυσμα μπορεί να εκφρασθεί ως γραμμικός συνδυασμός q διανυσμάτων.

Θεωρούμε ένα σύνολο διανυσμάτων *W* που ανήκουν σε ένα διανυσματικό χώρο *V*. Το σύνολο διανυσμάτων *W* αποτελεί ένα διανυσματικό υπόχωρο του διανυσματικού χώρου *V* τότε και μόνο τότε για κάθε σταθερά *c* και δυάδα διανυσμάτων $\vec{\mathbf{u}}, \vec{\mathbf{v}}$ που ανήκουν στο σύνολο W τα διανύσματα $\vec{\mathbf{u}} + \vec{\mathbf{v}}$ και $c\,\vec{\mathbf{u}}$ είναι στοιχεία του συνόλου *W*.

Θεωρούμε ένα τετραγωνικό μητρώο A $n \times n$. Τότε όλα τα διανύσματα $A\,\vec{\mathbf{u}}$, όπου $\vec{\mathbf{u}}$ είναι διάνυσμα που ανήκει στον χώρο V_n, συνιστούν ένα διανυσματικό υπόχωρο του V_n, διότι $A\vec{\mathbf{u}}$ είναι ένα μητρώο $n\times 1$ που είναι στοιχείο του V_n. Επί πλέον ισχύει

$$\vec{\mathbf{u}}+\vec{\mathbf{v}} = A\vec{\mathbf{u}} \ + A\vec{\mathbf{v}} \ = A\left(\vec{\mathbf{u}}+\vec{\mathbf{v}}\right)$$

$$c\vec{\mathbf{u}} = cA\vec{\mathbf{u}} = Ac\vec{\mathbf{u}} \tag{5.10}$$

Επειδή όμως όλα τα διανύσματα $A\vec{\mathbf{u}}$ συνιστούν έναν διανυσματικό υπόχωρο του V_n συνεπάγεται ότι ένα σύνολο γραμμικά ανεξάρτητων διανυσμάτων της μορφής $A\vec{\mathbf{u}}$ μπορεί να χρησιμοποιηθεί σαν βάση. Ειδικά επειδή $\mathrm{A}\vec{\boldsymbol{\varepsilon}}_1$ είναι η πρώτη στήλη του μητρώου Α (όπου $\vec{\boldsymbol{\varepsilon}}_1 = \left[1,0,0........0\right]^T$), $A\vec{\boldsymbol{\varepsilon}}_2$ $\left(\text{όπου } \vec{\boldsymbol{\varepsilon}}_2 = \left[0,1,0,......0\right]^{\mathrm{T}}\right)$είναι η δεύτερη κ. ο. κ. και επειδή το διάνυσμα $\vec{\mathbf{u}}$ μπορεί να εκφρασθεί σαν γραμμικός συνδυασμός των διανυσμάτων $\vec{\boldsymbol{\varepsilon}}_{\mathrm{i}}$ που ονομάζονται διανύσματα βάσης

$$\vec{u} = \sum_{j=1}^{n} u_j \ \vec{\varepsilon}_j \tag{5.11}$$

κατ' επέκταση

$$\mathrm{A}\vec{u} = \sum_{j=1}^{n} u_j \ \mathrm{A} \ \vec{\varepsilon}_j \tag{5.12}$$

που σημαίνει ότι μπορούμε να εκλέξουμε ως βάση τον μέγιστο αριθμό των γραμμικά ανεξάρτητων στηλών του μητρώου Α. Η διάσταση του διανυσματικού υπόχωρου W που απαρτίζεται από διανύσματα $A\vec{u}$ είναι η τάξη του μητρώου Α.

Ο μετασχηματισμός ενός διανυσματικού χώρου V στον διανυσματικό του υπόχωρο W μπορεί να θεωρηθεί ως η απεικόνιση κάθε στοιχείου του V σε μια μοναδικά ορισμένη εικόνα στον διανυσματικό υπόχωρο W, δηλαδή $L\,\vec{\mathbf{u}} = \vec{\mathbf{v}}$που σημαίνει ότι ο μετασχηματισμός ή η εικόνα του $\vec{u}$ είναι το διάνυσμα $\vec{\mathbf{v}}$. Ο μετασχηματισμός είναι γραμμικός τότε και μόνο τότε εάν

$$L\left(\vec{u}+\vec{v}\right) = L\vec{u} + L\vec{v}$$

$$L\left(c\vec{u}\right) = c\ L\,\vec{u} \tag{5.13}$$

Έστω ότι Α είναι ένα $n \times n$ τετραγωνικό μητρώο και $\vec{u}$ είναι ένα διάνυσμα του διανυσματικού χώρου V_n τέτοιο ώστε

$$A\,\vec{u} = 0 \tag{5.14}$$

Εάν Α είναι μη ιδιάζουσας μορφής (non-singular) όταν η τιμή της ορίζουσας του μητρώου A, είναι διάφορη του μηδενός, $\det A \neq 0$. Τότε το γραμμικό σύστημα $A\vec{\mathbf{u}} = 0$ έχει την μοναδική τετριμμένη λύση (trivial solution) $\vec{\mathbf{u}} = 0$. Εάν όμως το μητρώο A είναι ιδιάζουσας μορφής (singular), δηλαδή $\det A = 0$, τότε υπάρχει ένα διάνυσμα $\vec{\mathbf{u}} \neq 0$ τέτοιο ώστε $A\vec{\mathbf{u}} = 0$.

Το σύνολο όλων των διανυσμάτων που ικανοποιούν το σύστημα $A\vec{u} = 0$ συνιστούν ένα υπόχωρο του V_n που ονομάζεται ο μηδενικός χώρος του μητρώου Α (null-space of A). Η τάξη του μηδενικού υπόχωρου ονομάζεται μηδενικότητα (nullity) του μητρώου Α. Η μηδενικότητα του Α συν την τάξη του Α είναι η διάσταση του χώρου V_n.

Για κάθε τετραγωνικό μητρώο υπάρχει ένα ειδικό σύνολο διανυσμάτων που ονομάζονται ιδιοδιανύσματα και ένα σύνολο βαθμωτών μεγεθών που συνδέεται με τα ιδιοδιανύσματα και ονομάζονται ιδιοτομές. Ένα διάνυσμα $\vec{u}$ είναι ιδιοδιάνυσμα του μητρώου Α όταν είναι μη – μηδενικό και λ είναι ένας γενικά μιγαδικός αριθμός τέτοιος ώστε

$$A\,\vec{u} = l\,\vec{u} \quad ή \quad (\mathrm{A} - \lambda\, I)\vec{u} = 0 \tag{5.15}$$

ο μιγαδικός αριθμός λ είναι η ιδιοτιμή του μητρώου Α που συνδέεται με το ιδιοδιάνυσμα $\vec{u}$. Κάθε πολλαπλάσιο ενός διανύσματος είναι ιδιοδιάνυσμα.

Η μέγιστη απόλυτη τιμή των ιδιοτιμών ενός μητρώου Α ονομάζεται φασματική ακτίνα και συμβολίζεται ως $\rho(A) = \max\left|\lambda\,(A)\right|$.
Η συνάρτηση

$$\phi(\lambda) = \det(A - \lambda\, I) \tag{5.16}$$

ονομάζεται χαρακτηριστική συνάρτηση του μητρώου Α. Οι ιδιοτιμές λ του μητρώου Α βρίσκονται από την λύση της χαρακτηριστικής εξίσωσης

$$\det\left(A-\lambda I\right)=0$$

(5.17)

$$\begin{bmatrix} \\ a_{11}-\lambda & a_{12}\ldots\ldots\ldots & a_{1n} \\ a_{21} & \alpha_{22}-\lambda\ldots\ldots & a_{2n} \\ \ldots & \ldots\ldots\ldots\ldots\ldots & \ldots \\ a_{n1} & a_{n2}\ldots\ldots\ldots & a_{nn}-\lambda \end{bmatrix}=0$$

Ένα πραγματικό μητρώο $n\ x\ m$ μπορεί να θεωρηθεί ως μια γραμμική απεικόνιση $\mathbb{R}^n \to \mathbb{R}^m$, δηλαδή για κάποιο διάνυσμα $\vec{x}\in\mathbb{R}^n \quad \vec{y}=A\vec{x} \quad y\in\mathbb{R}^m$. Η αναπαράσταση ενός διανύσματος $\vec{x}$ στον n–διαστατό Ευκλείδειο χώρο είναι $\vec{x}=\sum_{j=1}^{n} x_j\ \vec{e}_j^{(n)}$ όπου $\vec{e}_j^{(n)}$ είναι το διάνυσμα βάσης του Ευκλείδειου χώρου $\left(\vec{e}_j^{\,n}\right)_i=\delta_{ji}$. Η αναπαράσταση των διανυσμάτων $\mathbf{y}_j$ στον Ευκλείδειο χώρο $\mathbb{R}^m$ είναι $\vec{y}_j=\sum_{j=1}^{m} y_{ij}e_j^{(m)}$, τότε

$$a_{ij}=\left(\vec{y}\right)_i=\left(Ax\right)_i=\left(Ae_i^{(n)}\right)_{ij}$$

Θεωρώντας ένα τετραγωνικό μητρώο τότε ιδιο-διανύσματα είναι το σύνολο των διανυσμάτων που απεικονίζονται στην κατεύθυνσή τους επί μια σταθερά, δηλαδή διανύσματα $\vec{x}$ για τα οποία υπάρχει κάποιος αριθμός λ (πραγματικός ή μιγαδικός) τέτοιος ώστε $\mathrm{A}\vec{x}=\lambda\vec{x}, \quad \vec{x}\neq 0$ είναι τα ιδιο-διανύσματα του μητρώου Α, και οι αριθμοί λ είναι οι ιδιο-τιμές του μητρώου Α. Το σύνολο όλων των ιδιο-τιμών του μητρώου Α ονομάζεται φάσμα του μητρώου Α και συμβολίζεται ως S(A)

Η χαρακτηριστική συνάρτηση του ανάστροφου A^T είναι η ίδια με την χαρακτηριστική συνάρτηση του Α διότι $\left(A-\lambda I\right)^T=A^T-\lambda I$ αλλά επειδή $\det\left(A^T\right)=\det\left(\mathrm{A}^T\right)\ \det\left(\mathrm{A}-\lambda I\right)^T=\det\left(A^T-\lambda I\right)$.

Θεωρούμε δυο διακριτές ιδιο-τιμές του μητρώου Α $\lambda_1\neq\lambda_2$ και τα αντίστοιχα ιδιο-διανυσμάτων u και υ $Au=\lambda_1\vec{u}$ και $A\vec{v}=\lambda_2\vec{v}$. Τα ιδιο-διανύσματα $\vec{u}$ και $\vec{v}$ είναι κάθετα διότι $\left(\vec{v},A\vec{u}\right)=\left(\vec{v},\lambda_1\vec{u}\right)=\lambda_1\left(\vec{u},\vec{v}\right)$ και $\left(\vec{v},A\vec{u}\right)=\left(A\vec{v},\vec{u}\right)=\left(\lambda_2\vec{v},\vec{u}\right)=\lambda_2\left(\vec{v},\vec{u}\right)$ δηλαδή $\left(\vec{v},\vec{u}\right)=0$.

Η χαρακτηριστική συνάρτηση του μητρώου Α γράφεται ως

$$\phi(\lambda) = (-1)^n \left(\lambda^n - \sigma_1 l^{n-1} + ... + (-1)^n \sigma_n\right) \tag{5.18}$$

όπου σ_i είναι συνδυασμοί των υπο-μητρώων του Α . Μπορεί να αποδειχθεί ότι το άθροισμα των διαγώνιων στοιχείων του μητρώου Α είναι ίσο με το άθροισμα των ιδιο-τιμών του μητρώου Α. Το άθροισμα των ιδιο-τιμών ονομάζεται ίχνος (trace) του μητρώου Α.

$$\mathrm{T}r(\mathrm{A}) = \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii} \tag{5.19}$$

Το γινόμενο όλων των ιδιο-τιμών είναι ίσο με την τιμή της διακρίνουσας

$$\prod_{i=1}^{n} \lambda_i = \det(\mathrm{A})$$

Οι ιδιο-τιμές, τα ιδιο-διανύσματα ή ακόμα ο συνδυασμός ιδιο-τιμών με τα αντίστοιχα ιδιο-διανύσματα δεν χαρακτηρίζουν ένα μητρώο. Θεωρούμε ένα τετραγωνικό μητρώο *T* τάξης *n* μη ιδιάζουσας μορφής nonsingular. Τότε για οποιαδήποτε ιδιο-διανύσματα $\vec{u}$, η σχέση

$$\vec{u} = T\,\vec{v}$$

μπορεί να θεωρηθεί ως ένας μετασχηματισμός συντεταγμένων. Από την σχέση ορισμού ιδιο-τιμών/ιδιο-διανυσμάτων $A\vec{u}_i = \lambda_i\,\vec{u}_i$ και τον μετασχηματισμό συντεταγμένων *T* ορίζουμε το μητρώο $B = T^{-1}AT$ και σχηματίζουμε το γινόμενο

$$\begin{aligned} \mathrm{B}\vec{v}_i &= T^{-1}\Lambda\, T\,\vec{v}_i = \lambda_i\,\vec{v}_i \\ \vec{v}_i &= T^{-1}\,u_1 \end{aligned} \tag{5.20}$$

Το μητρώο $B = T^{-1}A\,T$ έχει τις ίδιες ιδιο-τιμές με το μητρώο Α και ιδιο-διανύσματα που συνδέονται με τα ιδιο-διανύσματα του μητρώου *T* το μητρώο *B* ονομάζεται όμοιο με το μητρώο *A* όταν και μόνο όταν υπάρχει ένα μητρώο *T* τέτοιο ώστε $\mathrm{B} = T^{-1}\,A\,T$.

Υποθέτοντας ότι όλες οι ιδιο-τιμές του μητρώου *A* είναι διαφορετικές $\lambda_1 \neq \lambda_2 \neq ...... \neq \lambda_n$ τότε τα αντίστοιχα ιδιο-διανύσματα

$\vec{u}_1, \vec{u}_2, ...\vec{u}_n$ είναι γραμμικά ανεξάρτητα. Θεωρούμε το μητρώο U του οποίου η i^{th} στήλη είναι το ιδιο-διάνυσμα $\vec{u}_i$ τότε

$\mathrm{A}u = u \; diag(\lambda_1, \lambda_2, ...\lambda_n)$ ή

$$u^{-1}\mathrm{A}u = diag(\lambda_1, \lambda_2, ...\lambda_n) \tag{5.21}$$

που σημαίνει ότι το μητρώο A είναι όμοιο με ένα διαγώνιο μητρώο του οποίου οι τιμές στην διαγώνιο είναι οι ιδιο-τιμές του A.

Θεωρούμε σταθερές $\alpha_0, a_1, ...a_m$ και ένα τετραγωνικό μητρώο A τότε μπορούμε να ορίσουμε ένα πολυωνυμικό μητρώο

$$g(A) = \sum_{j=0}^{m} a_j \mathrm{A}^{m-j} \tag{5.22}$$

όπου $\mathrm{A}^0 = \mathrm{I}$

Σύμφωνα με τον παραπάνω ορισμό για κάθε nonsingular μητρώο B (δηλαδή ένα μητρώο B για το οποίο μπορούμε να βρούμε το αντίστροφό του B^{-1}) παρατηρούμε ότι $B^{-1}A^2B = B^{-1}A\;B\;B^{-1}A\;B$. Συνεπώς εξ επαγωγής συμπεραίνουμε ότι για το μητρώο B ισχύει

$$g(\mathrm{B}^{-1}AB) = B^{-1}g(A)B \tag{5.23}$$

Επί πλέον εάν οι ιδιο-τιμές του μητρώου A είναι $\lambda_1, \lambda_2, ...\lambda_n$ τότε οι ιδιο-τιμές του πολυωνυμικού μητρώου $g(\mathrm{A})$ είναι $g(\lambda_1), g(\lambda_2)..., g(\lambda_n)$.

Έστω ότι U είναι το μητρώο που έχει ως στήλες τα ιδιο-διανύσματα του μητρώου A που υποτίθεται ότι είναι διαφορετικά τότε $AD = D \; diag(\lambda_1, ...\lambda_n)$ και $u^{-1}A\;u = diag(\lambda_1, \lambda_2, ...\lambda_n)$ σύμφωνα με την Εξ. (5.21)

$$g(u^{-1}A\;u) = u^{-1}g(\mathrm{A})\;u \tag{5.24}$$

αλλά το μητρώο ιδιο-τιμών είναι $\Lambda = diag(\lambda_1, \lambda_2, ...\lambda_n)$ και $\Lambda^k = diag(\lambda_1^k, \lambda_2^k, ..., \lambda_n^k)$. Επίσης για οποιουσδήποτε διαγώνια μητρώα $D_1 = diag(a_1, a_2, ...a_n)$, $D_2 = diag(b_1, b_2, ..., b_n)$ και σταθερές β, γ ισχύει

$aD_1 + \beta D_2 = (\beta a_1 + \gamma b_1,\; \beta a_2 + \gamma b_2, ...\beta a_n + \gamma b_n)$ συνεπώς

$$u^{-1}g(\mathrm{A})u = diag\left(g(\lambda_1), g(\lambda_2), ...g(\lambda_n)\right)$$

ή (5.25)

$$g(A)u = u\ diag\left(g(\lambda_1), ...g(\lambda_n)\right)$$

δηλαδή οι ιδιο-τιμές του πολυωνυμικού μητρώου $g(H)$ είναι $g(\lambda_i)$ και επειδή υποθέσαμε ότι τα ιδιο-διανύσματα του Α είναι ξεχωριστά τα ιδιο-διανύσματα του $g(A)$ είναι τα ίδια με τα ιδιο-διανύσματα του Α.

Βασιζόμενοι στα παραπάνω μπορούμε να αποδείξουμε ότι η ικανή και αναγκαία συνθήκη ώστε $\lim_{k\to\infty} A^k = 0$, όπου Α είναι ένα τετραγωνικό μητρώο, είναι το μητρώο Α να έχει ιδιο-τιμές $|\lambda_i| < 1$. Η συνθήκη αυτή ισχύει διότι $u^{-1}\mathrm{A}u = \Lambda \quad ή \quad A = u^{-1}\Lambda u$ οπότε $\mathrm{A}^k = u^{-1}\Lambda^k u$. Όταν το μέτρο όλων των ιδιο-τιμών είναι μικρότερο της μονάδας η σειρά των μητρώων

$$I + A + A^2 + ...A^k + ... \tag{5.26}$$

συγκλίνει οπότε

$$(I - A)^{-1} = I + A + A^2 + ...$$

Συμμετρικά και τύπου Hermite μητρώα

Αναφέρθηκε ότι ένα μητρώο Α είναι τύπου Hermite τότε και μόνο τότε εάν $A = A* = \overline{A}^T$ ενώ το συμμετρικό όταν $\mathrm{A} = \mathrm{A}^T$. Δηλαδή είναι μητρώο είναι συμμετρικό και τύπου Hermite τότε και μόνο τότε εάν Α είναι πραγματικό και $\mathrm{A} = \mathrm{A}^T$.

Το άθροισμα $\sum_{i=1}^{n}\ \sum_{j=1}^{n} \alpha_{ij}\ \overline{u}_i\ u_j$ των στοιχείων α_{ij} ενός μητρώου για τα οποία ισχύει $\alpha_{ij} = \overline{a}_{ij}$ για οποιουσδήποτε μιγαδικούς αριθμούς u_i ονομάζεται Hermitian μορφή και γράφεται ως

$$\left(\overline{u}, A\vec{u}\right) \qquad για \qquad \mathrm{A} = \mathrm{A}^* \tag{5.27}$$

Κατ' αναλογία τετραγωνική μορφή είναι το άθροισμα $\sum_{i=1}^{n}\ \sum_{j=1}^{n} \alpha_{ij}\ u_i\ u_j$ των στοιχείων $a_{ij} = a_{ji}$ ενός συμμετρικού μητρώου. Η τετραγωνική μορφή για $\alpha_{ij}, u_j \in \mathbb{R}$ είναι

$$\left(u, Au\right) \quad για \quad \mathrm{A} = \overline{\mathrm{A}} = \mathrm{A}^{\mathrm{T}} \quad \vec{u} \in \mathbb{R} \tag{5.28}$$

Όταν το διάνυσμα $\vec{u}$ και το μητρώο A ικανοποιούν συνθήκη την (5.27) ή (5.28) τότε

(1) $\forall \vec{u}\ , \quad \left(\vec{u}, A\vec{u}\right) \quad \in \mathbb{R}$

(2) όλες οι ιδιο-τιμές του μητρώου A είναι πραγματικές

(3) Τα ιδιο-διανύσματα που αντιστοιχούν σε ξεχωριστές ιδιο-τιμές είναι κάθετα μεταξύ τους

Ένα μητρώο A είναι θετικά ορισμένο (positive definite) στον χώρο $\mathbb{C}^n$ όταν η τετραγωνική του μορφή είναι πραγματική και θετική $\left(\mathrm{A}\vec{x}, \vec{x}\right) > 0 \quad \forall\, \vec{x} \neq 0,\ \vec{x} \in \mathbb{C}^n$. Ανάλογος είναι ο ορισμός του θετικά ορισμένου μητρώου στον χώρο $\mathbb{R}^n \left(\mathrm{A}\vec{x}, x\right) > 0 \quad \forall\, \vec{x} \neq 0,\ \vec{x} \in \mathbb{R}^n$. Τα συμμετρικά και θετικά ορισμένα μητρώα (symmetric positive definite SPD matrices) έχουν ιδιαίτερη σημασία στην επίλυση γραμμικών συστημάτων όμως τα μητρώα που προκύπτουν στις εφαρμογές αριθμητικών μεθόδων δεν είναι πάντα SPD.

Ένα μητρώο A είναι unitary τότε και μόνο τότε όταν $\mathrm{A}^* = A^{-1}$ ορθογώνιο τότε και μόνο τότε όταν $\mathrm{A}^{\mathrm{T}} = A^{-1}$ (προφανώς τα unitary μητρώα είναι ορθογώνια), και ισομετρικά τότε και μόνον τότε όταν για όλα τα n διαστάσεων διανύσματα $\mathbf{\vec{u}}, \mathbf{\vec{v}} \in \mathbb{R}^n$ ισχύει η σχέση $\left(\vec{u}, \vec{v}\right) = \left(A\vec{u}, A\vec{v}\right)$. Στην συνθήκη ισομετρίας το γινόμενο μητρώου διανύσματος Au μπορεί να θεωρηθεί όπως αναφέραμε σαν ένας μετασχηματισμός μεταβλητής $\vec{w} = \mathrm{A}\vec{u}$ τότε $\left(\vec{u}, \vec{u}\right) = \left(A\vec{u}, A\vec{u}\right) = \left(\vec{w}, \vec{w}\right)$ που σημαίνει ότι μήκος διατηρείται. Ένα μητρώο είναι ισομετρικό μόνο και μόνο όταν είναι unitary.

Θεωρούμε το Hermitian μητρώο A τότε υπάρχει ένα unitary μητρώο τέτοιο ώστε

$$\mathrm{A} = P\Lambda P^*, \ \ P^* = P^{-1} \tag{5.29}$$

όπου $\lambda_1, \lambda_2, \ldots\ldots, \lambda_n$ είναι οι πραγματικές ιδιο-τιμές του μητρώου A και $\Lambda = diag\left(\lambda_1, \lambda_2, \ldots. \lambda_n\right)$. Όταν όλες οι ιδιο-τιμές είναι διαφορετικές μεταξύ τους υπάρχουν n ιδιο-διανύσματα ορθογώνια και γραμμικά ανεξάρτητα θεωρούμε ένα μητρώο A που είναι Hermitian τότε όταν επαληθεύεται μία από τις παρακάτω συνθήκες το μητρώο A είναι θετικά ορισμένο.

(1) Όλες οι ιδιο-τιμές του μητρώου A είναι θετικές

(2) Οι συντελεστές της χαρακτηριστικής εξίσωσης του μητρώου A έχουν εναλλασσόμενο πρόσημο.

Όταν το μητρώο A $m \times n$ είναι θετικά ορισμένο και Hermitioan τότε ικανοποιείται η παρακάτω συνθήκη

$$\lambda_n \le \frac{\left(\vec{u}, A\vec{u}\right)}{\left(\vec{u},\vec{u}\right)} \le \frac{\mathrm{A}\vec{u}, \mathrm{A}\vec{u}}{\left(\vec{u}, \mathrm{A}\vec{u}\right)} \le \lambda_1 \tag{5.30}$$

όπου $\lambda_1 > \lambda_2 > .... > \lambda_n$ και $\vec{u}$ είναι ένα διάνυσμα διάστασης n.

5.1.2 Στοιχειώδεις μετασχηματισμοί μητρώων

Υπάρχουν τρία είδη στοιχειωδών μητρώων που συνδέονται με το μοναδιαίο μητρώο I τα οποία θα ορίσουμε παρακάτω. Πολλαπλασιασμός ενός μητρώου A με αυτά τα στοιχειώδη μητρώα έχει σαν αποτέλεσμα την δημιουργία στοιχειωδών μετασχηματισμών που συχνά ονομάζονται στοιχειώδεις μετάθεση γραμμών (elementary row operations). Τα στοιχειώδη μητρώα είναι

1. Το στοιχειώδες μητρώο Q το οποίο είναι το $n\ x\ n$ διαγώνιο μητρώο που δημιουργείται όταν αντικαταστήσουμε το i^{th} διαγώνιο στοιχείο του μοναδιαίου μητρώου I με μια σταθερά q.

$$Q = \begin{array}{cc} \begin{bmatrix} 1 & 0 & .... & 0 \\ 0 & 1 & .... & 0 \\ .. & .. & .. & .. \\ 0 & 0 & q & 0 \\ .. & .. & .. & .. \\ 0 & 0 & 0 & 1 \end{bmatrix} & \begin{array}{c} \\ \\ \\ i^{th} \\ \\ \\ \end{array} \end{array} \tag{5.31}$$

Τα στοιχειώδες μητρώα Q έχει τις παρακάτω ιδιότητες

$$\begin{aligned} &\det \quad Q = q \\ &Q^{-1} = diag(1,1....1,1/q,1...,1 \end{aligned} \tag{5.32}$$

2. Το στοιχειώδες μητρώο P είναι ένα $n \ x \ n$ μητρώο που προκύπτει όταν εναλλαχθούν δυο γραμμές του μοναδιαίου μητρώου I

$$P = \begin{bmatrix} 1 & .. & 0 & .... & & 0 \\ 0 & & 1 & .... & & 0 \\ 0 & & 0 & .. & 1 & 0 \\ .. & .. & .. & .. & .. & .. \\ 0 & 0 & 1 & 0 & .. & 0 \\ .. & .. & .. & .. & .. & .. \\ 0 & 0 & .. & .. & 0 & 1 \end{bmatrix} \tag{5.33}$$

Το στοιχειώδες μητρώο P έχει τις ιδιότητες

$$\det \quad P = -1$$

$$P = I$$

π. χ.

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

3. Το στοιχειώδες μητρώο s που δημιουργείται από το μοναδιαίο μητρώο I όταν εισαχθεί ένα στοιχείο $s \neq 0$ στην θέση i, j του μοναδιαίου μητρώου

$$s = \begin{array}{cc} \begin{bmatrix} 1 & 0 & .. & 0 & .. & 0 \\ 0 & 1 & .. & 0 & .. & 0 \\ .. & .. & .. & .. & .. & .. \\ 0 & 0 & 1.. & 5 & .. & 0 \\ .. & .. & .. & .. & .. & .. \\ 0 & 0 & .. & 0.. & 0.. & 01 \end{bmatrix} & \begin{array}{c} \\ \\ \\ j \\ \\ \\ \end{array} \end{array} \tag{5.34}$$

Το μητρώο s έχει την ιδιότητα $\det s = 1$, π.χ.

$$4x4 \quad s = \begin{bmatrix} 1 & s & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \det S = 1\,x\,1 - s\,x\,0 + 0\,x\,0 = 1$$

Παραδείγματα στοιχειωδών μετασχηματισμών που προκύπτουν με τον πολλαπλασιασμό μητρώου A με τα στοιχειώδη μητρώα Q, R, S είναι

$$Q\,A \begin{bmatrix} 1 & 0 & 0 \\ 1 & 9 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{ii} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ qa_{12} & q\alpha 22 & qa_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

1. Q A αντιστοιχεί σε πολλαπλασιασμό όλων των στοιχείων μιας γραμμής με q.
2. AQ αντιστοιχεί ανάλογα σε πολλαπλασιασμό όλων των στοιχείων μιας στήλης με q.
3. όταν το μητρώο A είναι τετραγωνικό όπως στο παράδειγμα τότε

$$\det(QA) = \det Q\,x \det A = q \det A$$

$$PA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Το μητρώο P αναφέρεται συχνά ως μητρώο εναλλαγής (permutation matrix).

1. Ο πολλαπλασιασμός PA εναλλάσσει δυο γραμμές
2. Ο πολλαπλασιασμός AP εναλλάσσει δυο στήλες
3. Όταν το μητρώο A είναι τετραγωνικό τότε

$$\det(PA) = \det P\ x \det = -\det A$$

$$SA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & S \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} + Sa_{31} & a_{22} + Sa_{32} & a_{23} + Sa_{33} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

1. Ο πολλαπλασιασμός SA αντιστοιχεί με πρόσθεση των πολλαπλασίων επί 5 της γραμμής j στην γραμμή i
2. Ο πολλαπλασιασμός As αντιστοιχεί ανάλογα με πρόσθεση των πολλαπλασίων στήλης σε στήλη
3. Όταν το μητρώο A είναι τετραγωνικό

$$\det(SA) = \det S \; den \; A = \det \mathrm{A}$$

Η σημαντική ιδιότητα των στοιχειωδών μητρώων είναι ότι ο πολλαπλασιασμός ενός μητρώου με αυτά διατηρεί την τιμή της ορίζουσας. Ένα τετραγωνικό μητρώο του οποίου η ορίζουσα είναι διαφορετική του μηδενός είναι θεωρητικά αντιστρέψιμα. Η αντιστροφή ενός n x n μητρώου απαιτεί όπως θα δείξουμε παρακάτω n^3 υπολογισμούς που σημαίνει ότι τα λάθη στρογγυλοποίησης που εισάγονται κατά την διαδικασία αντιστροφής είναι σημαντικά. Η επίλυση συστημάτων μεγάλου μεγέθους $n > 20$ δεν πραγματοποιείται ποτέ με την αντιστροφή του πίνακα αλλά με την διαδικασία απαλοιφής Gauss που απαιτεί $n^3/3$ υπολογισμούς. Λόγω του μεγάλου αριθμού πράξεων και στις δυο περιπτώσεις τα λάθη στρογγυλοποίησης μπορεί να καταστήσουν ένα μητρώο close to singular . Απαιτείται συνεπώς να γνωρίζουμε εκ των προτέρων πόσο κοντά βρίσκεται κάποιο μητρώο στην κατάσταση του να μην είναι αντιστρέψιμο. Το μέτρο της συμπεριφοράς των μητρώων στις διαδικασίες αντιστροφής και απαλοιφής Gauss είναι ο αριθμός συμβατότητας (condition number) που θα μας απασχολήσει παρακάτω.

5.1.3 Αριθμός συμβατότητας μητρώου

Ο αριθμός συμβατότητας ορίζεται με την βοήθεια μέτρων διανυσμάτων (norms of a vector) που είναι αριθμοί οι οποίοι εκτιμούν το μέγεθος των στοιχείων διανυσμάτων. Το μέτρο διανύσματος **x** που συμβολίζεται l_p για ένα εύρος της παραμέτρου p, $1 \le p \le \infty$ ορίζεται ως

$$l_p = \|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p} \qquad (5.35)$$

Συνήθως χρησιμοποιούμε τα μέτρα l_1, l_2, και l_∞ που για το διάνυσμα $\vec{x} = \left(x_1, x_2, ...x_n \right)^T$ ορίζονται ως

$$\|\vec{x}\|_1 = \sum_{j=1}^{n} |x_j|$$

$$\|\vec{x}\|_2 = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2} \qquad (5.36)$$

$$\|\vec{x}\|_\infty = \max_j |x_j|$$

Η εκλογή της παραμέτρου p δεν έχει συνήθως μεγάλη σημασία και μπορούμε να υποδηλώσουμε το μέτρο ενός διανύσματος ως $\|x\|$. Το μέτρο διανύσματος εξ ορισμού έχει τις παρακάτω βασικές ιδιότητες που σχετίζονται με την απόσταση

$$\|\vec{x}\| > 0 \quad \forall \vec{x} \neq 0$$

$$\|0\| = 0$$

$$(5.37)$$

$$\|c\vec{x}\| = |c| \; \|x\|$$

$\|x + y\| \leq \|x\| + \|y\|$ (ανισότητα τριγώνου)

Ο πολλαπλασιασμός διανύσματος $\vec{x}$ με μητρώο Α δημιουργεί ένα νέο διάνυσμα $A\vec{x}$ του οποίου το μέτρο εξαρτάται από την μορφή του μητρώου Α. Το εύρος των μεταβολών που εισάγονται λόγω του πολλαπλασιασμού με το μητρώο Α εκφράζονται από τους αριθμούς

$$\mathrm{M} = \max \frac{\|Ax\|}{\|x\|}$$

$$(5.38)$$

$$m = \min \frac{\|Ax\|}{\| x \|}$$

Όταν δε $m = 0$, τότε το μητρώο Α είναι ιδιάζουσας μορφής (singular). Ο αριθμός κατάστασης (condition number) του μητρώου Α ορίζεται ώς

$$\kappa(\mathrm{A}) = \frac{\mathrm{M}}{m} = \frac{\max \dfrac{\|Ax\|}{\|x\|}}{\min \dfrac{\|Ax\|}{x}} \tag{5.39}$$

Η αριθμητική τιμή του αριθμού κατάστασης κ(Α) εξαρτάται από το μέτρο που χρησιμοποιούμε για τον ορισμό του. Συνήθως όμως ενδιαφέρει η τάξη μεγέθους του κ(Α) οπότε η εκλογή του μέτρου δεν είναι τόσο σημαντική.

Ο αριθμός κατάστασης αποτελεί μέτρο του πόσο απέχει κάποιο μητρώο από το να γίνει μη αντιστρεπτό και όταν είναι μεγάλος τότε το μητρώο Α είναι σχεδόν ιδιάζουσας μορφής (almost singular). Ο αριθμός κατάστασης μητρώου έχει τιμή μεγαλύτερη της μονάδας επειδή $M \geq m$

$$\kappa(A) \geq 1 \tag{5.40}$$

Πολλαπλασιασμός του διανύσματος $\vec{x}$ με το μητρώο εναλλαγής έχει σαν αποτέλεσμα την αναδιάταξη των στοιχείων του διανύσματος συνεπώς επειδή το μέτρο δεν μεταβάλλεται $\|\vec{x}\| = \|P\vec{x}\|$ και έχουμε

$$\kappa(P) = 1 \tag{5.41}$$

το μητρώο I είναι ειδική περίπτωση του μητρώου P οπότε $\kappa(I) = 1$. Πολλαπλασιασμός του μητρώου A με το μητρώο Q ή με μια σταθερά q πολλαπλασιάζει την τιμή των μέτρων M και m, όπως ορίζονται από την Εξ. (5.38), με q συνεπώς

$$\kappa(qA) = \kappa(A) \tag{5.42}$$

Όταν το μητρώο A είναι το διαγώνιο μητρώο D τότε ο αριθμός κατάστασης βρίσκεται εύκολα από την σχέση

$$\kappa(D) = \frac{\max|d_{ii}|}{\min|d_{ii}|} \tag{5.43}$$

Το παρακάτω παράδειγμα δείχνει ότι ο αριθμός κατάστασης είναι το καλύτερο κριτήριο της αντιστρεψιμότητας ενός πίνακα. Θεωρούμε το διαγώνιο μητρώο 100×100 με στοιχεία 0.01 στην διαγώνιο. Η ορίζουσα αυτού του μητρώου είναι 10^{-200} και το μητρώο μπορεί να θεωρηθεί ως ιδιάζουσας μορφής (singular) διότι η ορίζουσα του είναι πολύ κοντά στο μηδέν. Εφαρμόζοντας όμως κριτήριο αντιστρεψιμότητας τον αριθμό κατάστασης της Εξ. (5.43) συμπεραίνουμε ότι το μητρώο δεν είναι singular. Πράγματι το διαγώνιο μητρώο του παραδείγματος συμπεριφέρεται σαν το μοναδιαίο μητρώο I και όχι σαν ένα singular μητρώο. Η παρατήρηση αυτή προκύπτει και από την ιδιότητα της Εξ. (5.42).

Στις εφαρμογές των αριθμητικών μεθόδων και στις επαναληπτικές μεθόδους επίλυσης γραμμικών συστημάτων ενδιαφερόμαστε να γνωρίζουμε τις μεταβολές που υφίσταται η λύση του συστήματος $Ax = b$ από μικρές μεταβολές $\delta\vec{b}$ στο διάνυσμα $\vec{b}$.
Δηλαδή δεδομένου ότι :

$$A\,\vec{x} = \vec{b}$$

αναζητούμε την μεταβολή $\delta\,\vec{x}$ που προέρχεται από την μεταβολή ή το σφάλμα $\delta\,\vec{b}$

$$A\left(\vec{x} + \delta\vec{x}\right) = \vec{b} + \delta\vec{b}$$

αλλά επειδή $\mathrm{A}\vec{x} = \vec{b}$ προκύπτει ότι $A\delta\vec{x} = \delta\vec{b}$ και με εφαρμογή των ορισμών της Εξ. (5.38) έχουμε :

$$\|b\| \le M\,\|x\|$$

$$n\|\delta\vec{x}\| \le \|\delta\vec{b}\|$$

και επειδή η σταθερά m είναι διάφορη του μηδενός βρίσκουμε ότι:

$$\frac{\|\delta\vec{x}\|}{\|x\|} \le \frac{\mathrm{M}}{m}\frac{\|\delta\vec{b}\|}{\|\vec{b}\|} = \kappa\left(\mathrm{A}\right)\frac{\|\delta\vec{b}\|}{\|\vec{b}\|}$$

Ο λόγος $\|\delta\vec{b}\|/\|\vec{b}\|$ της παραπάνω σχέσης εκφράζει όμως την σχετική μεταβολή του δεξιού μέλους στο σύστημα $\mathrm{A}\vec{x} = \vec{b}$ και ο λόγος $\|\delta\vec{x}\|/\|\vec{x}\|$ εκφράζει την σχετική μεταβολή στην λύση που προκαλείται από την μεταβολή του δεξιού μέλους.

Συνεπώς ο αριθμός κατάστασης είναι ο *παράγοντας μεγέθυνσης* της μεταβολής $\delta\vec{b}$ και είναι προφανές ότι μικρές μεταβολές στο διάνυσμα $\vec{b}$ έχουν σαν αποτέλεσμα μεγάλες μεταβολές στην λύση όταν ο παράγοντας κατάστασης είναι μεγάλος. Μπορεί να αποδειχθεί ότι το ίδιο ισχύει και για μικρές μεταβολές συντελεστών του μητρώου συνεπώς το σφάλμα στρογγυλοποίησης μπορεί να προκαλέσει μεγάλο σφάλμα στην λύση όταν ο αριθμός κατάστασης είναι μεγάλος. Τα παραπάνω συμπεράσματα έχουν μεγάλη σημασία στην εφαρμογή επαναληπτικών μεθόδων προσεγγιστικής επίλυσης συστημάτων αλλά και την ανάλυση και τον ρόλο των σφαλμάτων στρογγυλοποίησης που εισάγονται στην διαδικασία επίλυσης συστημάτων με την μέθοδο απαλοιφής Gauss που θα εξετάσουμε παρακάτω.

Οι παρακάτω ανισότητες είναι χρήσιμες για την ανάλυση σφάλματος που προκύπτει από την επίλυση συστημάτων στον υπολογιστή

$$\frac{\left\| b - \mathrm{A}x^{*} \right\|}{\left\| A \right\| \left\| x^{*} \right\|} \leq \rho\, \varepsilon \tag{5.44}$$

$$\frac{\left\| x - x^{*} \right\|}{\left\| x^{*} \right\|} < \rho\, \kappa\left(\mathrm{A}\right) \varepsilon \tag{5.45}$$

όπου $\vec{x}$ είναι η ακριβής λύση του συστήματος και $\vec{x}^{*}$ είναι η λύση που βρίσκουμε με την επαναληπτική μέθοδο ή την απαλοιφή Gauss, ε (σφάλμα στρογκυλοποίησης machine epsilon) είναι η ακρίβεια της μηχανής και ρ είναι μια σταθερά με τιμή $\rho < 10$. Ο αριθμητής στην ανισότητα (5.44) είναι το υπόλοιπο της προσεγγιστικής λύσης και ο λόγος το σχετικό υπόλοιπο που αναμένεται να είναι της τάξης του σφάλματος στρογγυλοποίησης ανεξάρτητα από τον αριθμό κατάστασης του μητρώου *A*. Ο αριθμητής στην ανισότητα (5.45) είναι το σφάλμα της προσεγγιστικής λύσης και ο λόγος είναι το σχετικό σφάλμα που αναμένεται να είναι μικρό όταν ο αριθμός κατάστασης είναι μικρός ενώ γίνεται πολύ μεγάλο όταν το μητρώο *A* είναι σχεδόν singular.

Η ποσότητα *M* που ορίσαμε προηγουμένως είναι το μέτρο μητρώου και συμβολίζεται ως $\left\|\mathrm{A}\right\|$.

$$\left\| \mathrm{A} \right\| = \max \frac{\left\| Ax \right\|}{\left\| x \right\|} = M \tag{5.46}$$

κατ’ αναλογία η ποσότητα m που ορίσθηκε παραπάνω είναι το μέτρο του αντίστροφου του μητρώου A

$$\left\|A^{-1}\right\| = \frac{1}{m} \qquad (5.47)$$

Από τους ορισμούς των Εξ. (5.46) και (5.47) προκύπτει ο ορισμός του αριθμού κατάστασης μητρώου

$$\kappa(\mathrm{A}) = \|A\| \; \left\|A^{-1}\right\| \qquad (5.48)$$

που είναι ισοδύναμος με τον ορισμό της Εξ. (5.39)

Αναφέρθηκε προηγουμένως ότι η τιμή του μέτρου ενός μητρώου, και συνεπώς ο αριθμός κατάστασης, εξαρτώνται από την εκλογή της μορφής του μέτρου (δες Εξ. (5.35)). Το μέτρο μητρώου που βασίζεται στον l_1 και l_∞ ορισμό μέτρου είναι

$$\|\mathrm{A}\|_1 = \max_j \sum_{i=1}^{n} \left|a_{ij}\right| \qquad (5.49)$$

$$\|A\|_\infty = \max_j \sum_{j=1}^{n} \left|\alpha_{ij}\right| \qquad (5.50)$$

ο υπολογισμός της τιμής του μέτρου που αντιστοιχεί στον ορισμό του l_2 μέτρου διανύσματος απαιτεί την χρήση της τεχνικής singular value decomposition που θα παρουσιασθεί σε επόμενο κεφάλαιο. Ο προσδιορισμός της τιμής του μέτρου ενός μητρώου σύμφωνα με τους ορισμούς των Εξ. (5.49) και (5.50) είναι εύκολος και μπορεί να γίνει σε οποιαδήποτε γλώσσα προγραμματισμό $C++$, C, *ή FORTRAN*. Το λογισμικό *MATLAB* έχει ειδική εντολή για τον προσδιορισμό του μέτρου μητρώου. Η εντολή αυτή είναι

$$norm(A, p) \qquad p = 1, 2, \ \inf$$

Ο προσδιορισμός του αριθμού κατάστασης μητρώου σύμφωνα με τον ορισμό της Εξ. (5.48) απαιτεί τον υπολογισμό του αντιστρόφου A^{-1} του μητρώου Α. Ο υπολογισμός του αντιστρόφου όμως απαιτεί όπως θα δείξουμε στο επόμενο κεφάλαιο περίπου τρεις φορές μεγαλύτερη υπολογιστική προσπάθεια από εκείνη που απαιτείται για την επίλυση του αντίστοιχου συστήματος (πρέπει να σημειωθεί ότι το μέτρο $\left\|A^{-1}\right\|_2$ απαιτεί χρήση της τεχνικής *SVD* που συνεπάγεται περαιτέρω αύξηση της υπολογιστικής προσπάθειας). Η ακριβής τιμή του παράγοντα κατάστασης μητρώου δεν είναι όμως απαραίτητη στις πρακτικές εφαρμογές. Το λογισμικό *MATLAB* έχει τις παρακάτω εντολές για τον προσδιορισμό του αριθμού κατάστασης.

1. $cond(\mathrm{A})$ *or* $cond(\mathrm{A},2)$ προσδιορίζει τον αριθμό κατάστασης $k_2(\mathrm{A})$ χρησιμοποιώντας την συνάρτηση $svd(\mathrm{A})$. Ο προσδιορισμός του αριθμού κατάστασης με την εντολή $cond(\mathrm{A})$ συνίσταται για μητρώα μικρού μεγέθους όπου οι γεωμετρικές ιδιότητες του μέτρου l_2 έχουν μεγαλύτερη σημασία.

2. $cond(\mathrm{A},1)$ προσδιορίζει τον αριθμό κατάστασης $\kappa_1(\mathrm{A})$ και απαιτεί σημαντικά μικρότερο χρόνο υπολογισμού από την εντολή $cond(\mathrm{A},2)$.

3. $cond(\mathrm{A},\inf)$ προσδιορίζει τον αριθμό κατάστασης $\kappa_\infty(\mathrm{A})$ και απαιτεί τον ίδιο υπολογιστικό χρόνο με την εντολή $cond(\mathrm{A},1)$.

4. $cord\ est(\mathrm{A})$ προσδιορίζει τον αριθμό κατάστασης $\kappa_1(\mathrm{A})$ και συνίσταται για μεγάλου μεγέθους αραιά $(sparse)$ μητρώα.

5.2 Απαλοιφή Gauss

Η επίλυση τετραγωνικών γραμμικών συστημάτων $\mathrm{A}\vec{x}=\vec{b}$ δηλαδή συστημάτων τα οποία έχουν τον ίδιο αριθμό εξισώσεων όσο και αγνώστους και το μητρώο Α είναι τετραγωνικό $n\ x\ n$, επιτυγχάνεται με την συστηματική διαδικασία απαλοιφής που επινοήθηκε από τον *Gauss*. Η απαλοιφή *Gauss* είναι μια από τις παλαιότερες αριθμητικές μεθόδους και αποτελείται από δυο βασικά στάδια απαλοιφής (forward elimination) και αντικατάσταση (back substitution). Η απαλοιφή περιλαμβάνει $n-1$ βήματα, όπου στο k^{th} βήμα κατάλληλα πολλαπλάσια της k^{th} εξίσωσης αφαιρούνται από τις εναπομένουσες $n-k$ εξισώσεις με σκοπό την απαλοιφή της k^{th} μεταβλητής. Σαν αποτέλεσμα μετά το πέρας του *n – 1* βήματος απομένει η τελευταία εξίσωση ως προς x_n η οποία αφού λύνεται και η διαδικασία συνεχίζεται στην προ- τελευταία εξίσωση ως προς x_{n-1} κ.ο.κ. μέχρι να υπολογισθεί η μεταβλητή x_1 η διαδικασία αυτή ονομάζεται ανάδρομη αντικατάσταση (back substitution).

Η διαδικασία που περιγράψαμε παραπάνω υλοποιείται ως ακολούθως. Έστω ότι η αρχική μορφή του συστήματος είναι :

$$\mathrm{A}^{(1)}\vec{x}=b^{(1)}$$

όπου $\alpha_{ij}^{(1)}$ είναι τα στοιχεία του τετραγωνικού μητρώου των συντελεστών και $b_i^{(1)}$ είναι τα στοιχεία του διανύσματος στο δεξί μέλος. Απαλοιφή της πρώτης στήλης επιτυγχάνεται όταν πολλαπλασιάσουμε διαδοχικά την πρώτη γραμμή με $-a_{i1}^{(1)}/a_{11}^{(1)}$ και την προσθέσουμε στην i^{th} γραμμή οπότε έχουμε το ισοδύναμο σύστημα

$$
\begin{aligned}
a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \dots\dots\dots + a_{1n}^{(1)}x_n &= b_1^{(1)} \\
a_{22}^{(2)}x_2 + \dots\dots\dots + a_{2n}^{(2)}x_n &= b_2^{(2)} \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\
a_{n2}^{(2)}x_2 + \dots\dots\dots + a_{nn}^{(2)}x_n &= b_n^{(2)}
\end{aligned}
\tag{5.2.1}
$$

όπου

$$\alpha_{ij}^{(2)} = a_{ij}^{(1)} - \left(a_{i1}^{(1)}/a_{11}^{(1)}\right)a_{1j}^{(1)}$$

$$b_i^{(2)} = b_i - \left(a_{i1}^{(1)}/a_{i1}^{(1)}\right)b_1^{(1)}$$

Έστω ότι $a_{22}^{(2)} \neq 0$ τότε η διαδικασία απαλοιφής μπορεί να επαναληφθεί για τις εναπομένουσες $n-2$ εξισώσεις. Όταν $a_{22}^{(2)} = 0$ ή μπορούμε να κάνουμε αντιμετάθεση στηλών ή γραμμών και να συνεχίσουμε την διαδικασία απαλοιφής. Η αντιμετάθεση γραμμών και στηλών είναι αναγκαία διότι πάντα επιθυμούμε να αποφύγουμε την διαίρεση με μικρούς αριθμούς εάν λοιπόν το στοιχείο $a_{22}^{(2)}$ έχει μικρή τιμή κάνουμε αντιμετάθεση γραμμών ώστε να αποφύγουμε την διαίρεση $-a_{i2}^{(2)}/a_{22}^{(2)}$. Συνεχίζοντας την απαλοιφή έστω ότι στο k στάδιο έχουμε το μητρώο συντελεστών.

$$
\mathrm{A}^{(k)} = \begin{bmatrix}
a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \dots & & a_{1n}^{(1)} \\
0 & \alpha_{22}^{(2)} & a_{23}^{(2)} & \dots & & a_{2n}^{(2)} \\
. & 0 & a_{33}^{(3)} & \dots & & \alpha_{3n}^{(3)} \\
 & & & a_{rr}^{(r)} & \dots & a_{rn}^{(r)} \\
 & & & 0 & \dots & 0 \\
 & & & 0 & \dots & 0
\end{bmatrix}
\tag{5.2.2}
$$

Τότε σημαίνει ότι η διαδικασία απαλοιφής δεν μπορεί να συνεχιστεί διότι $a_{ss}^{(s)} \neq 0 \quad 1 \leq s \leq r \leq n, \ k = r \ \ ή \ \ k = r+1$ ενώ για $s > r$ όλοι οι συντελεστές μηδενίζονται. Ο αριθμός $r \leq n$ που ορίζεται από τον παραπάνω περιορισμό ονομάζεται διάσταση ή τάξη $(rank)$ του μητρώου Α. Το εύρος της απεικόνισης $\mathrm{A}: \mathbb{R}^n \to \mathbb{R}^m$ μεταξύ των διανυσματικών χώρων είναι $R(A) = \left\{ \vec{y} = A\vec{x} \ ; \vec{x} \in \mathbb{R}^n \right\}$ και έχει διάσταση που ισούται με την τάξη

$$\dim R(A) = r \quad rank \ of \ A \tag{5.2.3}$$

και αντίστοιχα το εύρος της απεικόνισης του μηδενικού χώρου, $\mathrm{N}(\mathrm{A})$, είναι $\mathrm{N}(\mathrm{A}) = \left\{ \vec{x} \in \mathbb{R}^n ; \mathrm{A}\vec{x} = 0 \right\}$ και έχει διάσταση

$$\dim \mathrm{N}(A) = n - r \quad ή \quad \mathrm{rank}(\mathrm{A}) + \dim \mathrm{N}(\mathrm{A}) = n \tag{5.2.4}$$

Όταν το διάνυσμα $\vec{b} \in R(\mathrm{A})$ τότε το σύστημα $\mathrm{A}\vec{x} = \vec{b}$ είναι ένα συμβατό σύστημα γραμμικών εξισώσεων διαφορετικά είναι μη συμβατό. Εάν το σύστημα $\mathrm{A}\vec{x} = \vec{b}$ είναι συμβατό και η τάξη του μητρώου *A* είναι $rank(\mathrm{A}) = r = n$ τότε υπάρχει μόνο μια λύση. Στην περίπτωση αυτή $\dim N(A) = 0$ και το μητρώο Α είναι nonsingular.

Όταν το μητρώο *A* είναι μη-ιδιάζουσας μορφής, δηλαδή όταν $rank(A) = r = n$, η διαδικασία απαλοιφής μπορεί να συνεχισθεί μέχρι την τελευταία εξίσωση και η διαδικασία επίλυσης του τριδιαγωνίου συστήματος που σχηματίζεται είναι εύκολη. Η διαδικασία απαλοιφής Gauss μπορεί να εκφρασθεί και με την βοήθεια των μετασχηματισμών γραμμών (elementary row operations) που υλοποιείται με τον πολλαπλασιασμό με τα στοιχειώδη μητρώα *Q, P, S* που ορίσαμε προηγουμένως. Η διαδικασία απαλοιφής Gauss με την χρήση στοιχειωδών μετασχηματισμών περιγράφεται για το 3×3 σύστημα γραμμικών εξισώσεων.

Ορίζουμε επαυξημένο μητρώο που αποτελείται από το μητρώο συντελεστών Α και το διάνυσμα του δεξιού μέλους $\vec{b}$.

$$C = \left[A \middle| \vec{u} \right] = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & .. & b_1^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} & .. & b_2^{(1)} \\ a_{31}^{(1)} & a_{32}^{(1)} & a_{33}^{(1)} & .. & b_3^{(1)} \end{bmatrix}$$

και θεωρούμε τα στοιχειώδη μητρώα τύπου –S

$$S_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\dfrac{a_{22}}{a_{11}} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ S_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ -\dfrac{b_{31}}{b_{11}} & 1 & 1 \end{bmatrix},\ S_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\dfrac{b_{32}}{b_{22}} & 1 \end{bmatrix}$$

τότε η διαδικασία απαλοιφής είναι

$$S_3\ S_2\ S_1\ C = \begin{bmatrix} 0_{11}^{(1)} & 0_{12}^{(1)} & 0_{13}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & b_3^{(3)} \end{bmatrix}$$

Η διαδικασία αντικατάστασης υλοποιείται με πολλαπλασιασμό με τα στοιχειώδη μητρώα Q_1, Q_2, Q_3

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \dfrac{1}{b_{33}^{(3)}} \end{bmatrix}, Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \dfrac{1}{b_{22}^{(2)}} & 0 \\ 0 & 0 & 1 \end{bmatrix},\ Q_3 = \begin{bmatrix} \dfrac{1}{b_{11}^{(1)}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Θεωρώντας ακόμη τρία μητρώα S_4, S_5, S_6 η διαδικασία επίλυσης είναι

$$Q_3\ S_6\ S_5\ Q_2\ S_4\ Q_1\ S_3\ S_2\ S_1\ CEC = \begin{bmatrix} 0 & 0 & 0 & x_1 \\ 0 & 1 & 0 & x_2 \\ 0 & 0 & 1 & x_3 \end{bmatrix}$$

ή $\qquad EC = E\left[A \middle| \vec{b} \right] = \left[I \middle| \vec{x} \right]$

(5.2.5)

ή $\qquad \mathrm{EA} = \mathrm{I} \Rightarrow \mathrm{E} = \mathrm{A}^{-1}$

Δηλαδή με την διαδικασία απαλοιφής που περιγράψαμε μπορούμε μα βρούμε και τον αντίστροφο του μητρώου A. Η εύρεση του αντίστροφου μητρώου A δεν είναι όμως απαραίτητη για την λύση του συστήματος και δεν υπολογίζεται διότι όμως θα δείξουμε παρακάτω απαιτεί σημαντικά μεγαλύτερη υπολογιστική προσπάθεια από ότι η λύση του συστήματος.

Η διαδικασία απαλοιφής Gauss χωρίς την χρήση του επαυξημένου μητρώου μπορεί να περιγραφή συστηματικά με την χρήση του μητρώου εναλλαγής $P_k,\ k=1,....,n-1$ και των κάτω τριγωνικών μητρώων τύπου S που περιέχουν τα αρνητικά των πολλαπλασιαστών που χρησιμοποιούνται στο k^{th} στάδιο απαλοιφής. Το τελικό άνω τριγωνικό μητρώο που προκύπτει μετά από τα *n-1* στάδια απαλοιφής περιγράφεται από την εξίσωση μητρώων

$$U = S_{n-1}\ P_{n-1}.......M_2\ P_2\ M_1\ P_1\ \ A \tag{5.2.6}$$

Η παραπάνω εξίσωση μπορεί να γραφεί ως

$$L_1\ L_2.......L_{n-1}\ U = P_{n-1}.....P_2\ P_1\ A \tag{5.2.7}$$

όπου τα κάτω τριγωνικά μητρώα έχουν την μορφή

$$L_1 = \begin{bmatrix} 1 & 0 & .. & 0 \\ -a_{21}^{(1)}/a_{11}^{(1)} & 1 & & \\ . & & & \\ -a_{n1}^{(1)} & 0 & & 1 \end{bmatrix} ....L_k = \begin{bmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ 0 & .. & 1 & & & \\ 0 & .. & -p_{k+1,k}^{(k)}/a_{kk}^{(k)} & 1 & & \\ 0 & .. & -a_{mk}^{(k)}/a_{kk}^{(k)} & 0 & .. & 1 \end{bmatrix}$$

$$l_{i1} = -a_{i1}^{(1)}/a_{11}^{(1)}\ \ 2 \le i \le n \qquad\qquad l_{ik} = -a_{ik}^{(k)}/a_{kk}^{(k)}\quad k+1 < i < n$$

Συμβολίζοντας

$$\begin{aligned} L &= L_1\ L_2\ ....L_{n-1} \\ P &= P_{n-1}.........P_2\ P_1 \end{aligned} \tag{5.2.8}$$

Η Εξ. 5.2.7 γράφεται ως

$$L\,U = P\ A \tag{5.2.9}$$

Με την παραπάνω παραγοντοποίηση το γραμμικό σύστημα $A\vec{x}=\vec{b}$ καθίσταται ισοδύναμο με δυο τριγωνικά συστήματα

$$L\vec{y}=P\vec{b}$$

$$U\vec{x}=\vec{y} \qquad (5.2.10)$$

Η απαλοιφή Gauss απαιτεί κατάλληλη εκλογή του στοιχείου $a_{kk}^{(k)}$ (pivot) που ελαχιστοποιεί τα σφάλματα στρογγυλοποίησης. Η παραλλαγή της απαλοιφής Gauss γνωστή και ως *Gauss – Jordan* απαλοιφή προσπαθεί να αποφύγει την εκλογή του στοιχείου εκτελώντας κανονικοποίηση.

5.2.1 Υπολογιστική περιπλοκότητα

Η υπολογιστική περιπλοκότητα αλγορίθμων είναι ένα θέμα ιδιαίτερου ενδιαφέροντος ιδιαίτερα στους υπολογισμούς μεγάλης κλίμακας. Η περιπλοκότητα αλγορίθμου της απαλοιφής Gauss αφορά το πλήθος αριθμητικών διεργασιών που απαιτείται για την παραγοντοποίηση του μητρώου *A* και την λύση του συστήματος $A\vec{x}=\vec{b}$. Υποθέτοντας ότι η διαίρεση απαιτεί τον ίδιο χρόνο με ένα πολλαπλασιασμό και μια πρόσθεση βρίσκουμε ότι το πλήθος των προσθέσεων και πολλαπλασιασμών κατά την διάρκεια της απαλοιφής εκτιμάται ότι είναι

$$\sum_{k=n}^{1} (k-1)k = \frac{1}{3}n\left(n^2-1\right) \sim \frac{1}{3}n^3 \qquad (5.2.11)$$

Στην παραπάνω εκτίμηση αριθμητικών διεργασιών ο παράγων $k-1$ είναι ο αριθμός των εναπομενουσών γραμμών που πρέπει να απαλειφθούν και *k* είναι ο αριθμός των στοιχείων που απομένουν στην γραμμή. Ο αριθμός των διαιρέσεων για τον υπολογισμό των αντίστροφων είναι (*n – 1*) και η εκτίμηση των αριθμητικών πράξεων αφορά μητρώα τα οποία έχουν παντού μη-μηδενικά στοιχεία. Ο αριθμός πράξεων (operation count) στον υπολογιστή μετράτε σε flop που αναφέρεται σε (floating point) πολλαπλασιασμό συνοδευόμενο από μια πρόσθεση. Δηλαδή η παραγοντοποίηση ενός πλήρους μητρώου A απαιτεί περίπου $n^3/3$ flops. Όταν το μητρώο είναι συμμετρικό τότε ο αριθμός υπολογισμών ελαττώνεται σε μισό και είναι $n^3/6$ flops.

Εκείνο που ενδιαφέρει στις πρακτικές εφαρμογές είναι ο χρόνος που απαιτείται για την επίλυση ενός συστήματος με την μέθοδο απαλοιφής Gauss. Υποθέτοντας ότι ο χρόνος της κεντρικής μονάδας του υπολογιστή central processing unit (CPU) είναι $1\mu s = 10^{-6}$ sec/ *flop* έχουμε την παρακάτω εκτίμηση χρόνου για την επίλυση συστημάτων

n	χρόνος cpu (sec)
100	0.33
1000	330
10000	$33x10^4 \simeq 3.8$ μέρες

Πέραν του κόστους υπολογισμών πρέπει να λάβουμε υπ’ όψη και τον χώρο μνήμης π.χ. όταν $n = 1000$ πρέπει να δεσμεύσουμε $n^2 = 10^6$ θέσης μνήμης. Η διαθέσιμη μνήμη για την αποθήκευση μπορεί να μην είναι αρκετή όταν το σύστημα είναι μεγάλο, και τότε η χρήση της δευτερεύουσας μνήμης μπορεί να επιβραδύνει ακόμη περισσότερο τους υπολογισμούς.

Πρέπει να σημειωθεί ότι στην παραπάνω εκτίμηση χρόνου εκτέλεσης συμπεριλάβαμε μόνο την παραγοντοποίηση. Η πλήρης επίλυση του συστήματος απαιτεί την λύση $u\vec{x} = \vec{y}$ που απαιτεί $\sum_{k=1}^{n} k = n(n+1)/2 \sim n^2/2$ flops οπότε η εκτίμηση του συνολικού χρόνου είναι $n^3/3 + n^2/2 \sim n^3/3$ όταν ο αριθμός εξισώσεων n είναι μεγάλος. Η αντιστροφή ενός μητρώου μπορεί να επιτευχθεί με την λύση του συστήματος $A\vec{x}_i = \vec{e}_i \quad 1 \le i \le n$ όπου $\vec{e}_i$ είναι οι στήλες του μοναδιαίου πίνακα, οπότε επειδή $AA^{-1} = I, \vec{x}_i$ είναι η i στήλη του μητρώου A^{-1}. Συνεπώς ο συνολικός αριθμός υπολογισμών είναι περίπου $n^3/3 + n \cdot n^2 = n^3 4/3$. Ο αριθμός πράξεων μπορεί να σημειωθεί με κατάλληλη τροποποίηση του αλγόριθμου αντιστροφής διότι τα διανύσματα $\vec{e}_i$ περιέχουν μόνο ένα μη – μηδενικό στοιχείο, και τελικά ο αριθμός πράξεων για την αντιστροφή γίνεται ανάλογος του n^3. Πρέπει να σημειωθεί ότι ο αριθμός πράξεων που απαιτείται για τον υπολογισμό δυο μητρώων είναι επίσης ανάλογος του n^3.

Σε πολλές πρακτικές εφαρμογές απαιτείται η λύση ενός συστήματος με το ίδιο μητρώο A και M διαφορετικά διανύσματα δεξιού μέλους. Στην περίπτωση αυτή θα αναμένονταν να επιτύχουμε επιτάχυνση της διαδικασίας εάν πρώτα υπολογίσουμε τον αντίστροφο του πίνακα A εκτελώντας n^3 αριθμητικές πράξεις και μετά να εκτελέσουμε M πολλαπλασιασμούς μητρώου με διάνυσμα που κάθε ένας από αυτούς απαιτεί n^2 flops. Πραγματικά επιτάχυνση των υπολογισμών επιτυγχάνεται όταν $n^3 + Mn^2 < M\ n^3/3$ δηλαδή όταν ο αριθμός των διαφορετικών συστημάτων είναι αρκετά μεγάλος $\mathrm{M} > (n-3)/3\pi$.Ο υπολογισμός του αντίστροφου συνίσταται στην παράλληλη επεξεργασία.

5.2.2 Σφάλμα απαλοιφής Gauss

Μπορεί να αποδειχθεί ότι η λύση x^* που υπολογίζεται με την απαλοιφή Gauss ικανοποιεί το σύστημα

$$(A+E)x^* = b \tag{5.2.12}$$

όπου Ε είναι ένα μητρώο του οποίου τα στοιχεία έχουν περίπου το μέγεθος του σφάλματος στρογγυλοποίησης των στοιχείων του Α. Το μέτρο του μητρώου Ε σχετίζεται με το μέτρο του μητρώου Α με την παρακάτω σχέση.

$$\frac{\|\mathrm{E}\|}{\|\mathrm{A}\|} = \rho\varepsilon \tag{5.2.13}$$

όπου ε είναι η ακρίβεια μηχανής και ρ ένας αριθμός περίπου 10. Η λύση x^* που υπολογίζουμε με απαλοιφή Gauss διαφέρει από την πραγματική λύση x λόγω της συσσώρευσης του σφάλματος στρογγυλοποίησης που όπως αναφέραμε μπορεί να είναι πολύ σημαντικό όταν δεν καταβληθεί προσπάθεια αποφυγής διαίρεσης με μικρούς αριθμούς. Η διαφορά $b - Ax^*$ ονομάζεται υπόλοιπο και σύμφωνα με τα παραπάνω είναι

$$\begin{aligned} &b - Ax^* = Ex^* \\ &\|b - Ax^*\| = \|Ex^*\| \le \|\mathrm{E}\|\ \|x^*\| \end{aligned} \tag{5.2.4}$$

Από την παραπάνω σχέση βρίσκουμε ότι το σχετικό υπόλοιπο είναι

$$\frac{\|b - \mathrm{A}x^*\|}{\|\mathrm{A}\|\ \|x^*\|} \le \frac{\|E\|}{\|A\|} = \rho\varepsilon \tag{5.2.15}$$

Το σφάλμα υπολογισμού είναι $x - x^*$ και επειδή το μητρώο Α είναι αντιστρέψιμο, διότι υποθέσαμε ότι η λύση του συστήματος υπάρχει, βρίσκουμε

$$x - x^* = \mathrm{A}^{-1}\left(b - \mathrm{A}x^*\right)$$

$$ή \tag{5.2.16}$$

$$\left\|x - x^*\right\| \le \left\|\mathrm{A}^{-1}\right\| \left\|\mathrm{E}\right\| \left\|x^*\right\|$$

και το σχετικό σφάλμα της λύσης που υπολογίζουμε με την απαλοιφή Gauss ικανοποιεί την σχέση

$$\frac{\left\|x - x^*\right\|}{\left\|x^*\right\|} \le \rho \left\|\mathrm{A}\right\| \left\|\mathrm{A}^{-1}\right\| \varepsilon$$

$$\tag{5.2.17}$$

$$\frac{\left\|x - x^*\right\|}{\left\|x^*\right\|} < \rho k\left(\mathrm{A}\right)\varepsilon$$

Συνεπώς το σχετικό σφάλμα εξαρτάται από την ακρίβεια μηχανής (single or double precision) αλλά και από τον αριθμό κατάστασης. Η διαδικασία απαλοιφής Gauss γενικά δίνει μικρό σφάλμα υπόλοιπου το οποίο εξαρτάται από την ακρίβεια μηχανής. Το πραγματικό σφάλμα , δηλαδή η απόσταση της λύσης x^* που υπολογίζουμε από την πραγματική λύση, εξαρτάται από τον αριθμό κατάστασης του μητρώου συντελεστών. Το σφάλμα αριθμητικής επίλυσης μπορεί να είναι σημαντικό όταν το μητρώο είναι near singular και ο αριθμός κατάστασης είναι μεγάλος.

Ο υπολογισμός του αριθμού κατάστασης $\kappa\left(\mathrm{A}\right) = \left\|\mathrm{A}\right\| \left\|\mathrm{A}^{-1}\right\|$ του μητρώου Α απαιτεί τον υπολογισμό του αντιστρόφου A^{-1}. Ο υπολογισμός του αντιστρόφου συνεπάγεται όμως τριπλάσιο υπολογιστικό κόστος συγκριτικά με την λύση του συστήματος που σημαίνει ότι ο υπολογισμός του αριθμού κατάστασης είναι δαπανηρός. Η ακριβής τιμή του αριθμού κατάστασης δεν είναι όμως αναγκαία και εκτιμήσεις του αριθμού κατάστασης μπορεί να βρεθούν με την χρήση του λογισμικού *MATLAB* που διαθέτει τις παρακάτω συναρτήσεις για την εκτίμηση του αριθμού κατάστασης.

cond *(A, 2)*	υπολογίζει τον αριθμό κατάστασης	$\kappa_2(A)$
cond *(A, 1)*	υπολογίζει τον αριθμό κατάστασης	$\kappa_1(A)$
cond *(A, inf)*	υπολογίζει τον αριθμό κατάστασης	$\kappa_\infty(A)$

5.2.3 Επίλυση με χρήση παραγοντοποιήσης (LU decomposition)

Η μέθοδος επίλυσης του συστήματος $[\mathbf{A}]\mathbf{x}=\mathbf{b}$ με την απαλοιφή Gauss εφαρμόζεται στο επαυξημένο μητρώο $[\mathbf{A}]|\mathbf{b}$, δηλαδή στο μητρώο $[\mathbf{A}]$ και στο διάνυσμα $\mathbf{b}$ με σκοπό την μετατροπή του μητρώου $[\mathbf{A}]$ σε ένα άνω–διαγώνιο μητρώο. Όταν όμως είναι αναγκαίο να επιλύσουμε ένα σύστημα γραμμικών εξισώσεων με το ίδιο μητρώο συντελεστών $[\mathbf{A}]$ και διαφορετικά διανύσματα δεξιού μέλους $\mathbf{b}_k$, $k=1,...,K$, τότε η μετατροπή του μητρώου $[\mathbf{A}]$ σε άνω–διαγώνιο μητρώο πρέπει να εκτελεσθεί K φορές. Η επανάληψη της διαδικασίας διαγωνιοποίησης μπορεί να αποφευχθεί με την παραγοντοποιήση του μητρώου $[\mathbf{A}]$ σε τριγωνικά μητρώα.

Μπορεί να αποδειχθεί ότι το μητρώο $[\mathbf{A}]$ είναι δυνατόν να γραφεί ως ένα γινόμενο μητρώων

$$\mathbf{A}=\mathbf{L}\ \mathbf{U}$$

όπου $[\mathbf{L}]$ είναι ένα κάτω–τριγωνικό μητρώο και $[\mathbf{U}]$ είναι ένα άνω–τριγωνικό μητρώο της μορφής

$$\mathbf{L}=\begin{bmatrix} l_{11} & 0 & 0 & ... & 0 \\ l_{21} & l_{22} & 0 & ... & 0 \\ l_{31} & l_{32} & l_{33} & & 0 \\ . & . & . & . & . \\ l_{n1} & l_{n2} & l_{n3} & ... & l_{nn} \end{bmatrix}, \qquad \mathbf{U}=\begin{bmatrix} u_{11} & u_{12} & u_{13} & ... & u_{1n} \\ 0 & u_{22} & 0 & ... & u_{2n} \\ 0 & 0 & u_{33} & ... & u_{3n} \\ . & . & . & . & . \\ 0 & 0 & 0 & ... & u_{nn} \end{bmatrix}$$

Το γινόμενο των διαγώνιων στοιχείων l_{kk} και u_{kk} είναι γνωστό $l_{kk}u_{kk}=a_{kk}$ και επειδή η τιμή των στοιχείων l_{kk} και u_{kk} είναι αυθαίρετη μπορεί να υποθέσουμε ότι τα στοιχεία l_{kk} ή u_{kk} είναι μονάδα, δηλαδή

$$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & a_{13} & ... & a_{1n} \\ a_{21} & a_{22} & a_{23} & ... & a_{2n} \\ a_{31} & a_{32} & a_{33} & & a_{3n} \\ . & . & . & . & . \\ a_{n1} & a_{n2} & a_{n3} & ... & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & ... & 0 \\ l_{21} & 1 & 0 & ... & 0 \\ l_{31} & l_{32} & 1 & & 0 \\ . & . & . & . & . \\ l_{n1} & l_{n2} & l_{n3} & ... & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} & ... & u_{1n} \\ 0 & u_{22} & 0 & ... & u_{2n} \\ 0 & 0 & u_{33} & ... & u_{3n} \\ . & . & . & . & . \\ 0 & 0 & 0 & ... & u_{nn} \end{bmatrix} = \mathbf{LU}$$

Δεδομένου ότι τα διαγώνια μητρώα $[\mathbf{L}]$ και $[\mathbf{U}]$ είναι γνωστά η διαδικασία επίλυσης για πολλαπλά διανύσματα $\mathbf{b}$ είναι ως ακολούθως

$[\mathbf{A}]\{\mathbf{x}\}=\{\mathbf{b}\}$

ή

$[\mathbf{L}][\mathbf{U}]\{\mathbf{x}\}=\{\mathbf{b}\}$

Η παραπάνω σχέση μετά από αντικατάσταση $[\mathbf{U}]\{\mathbf{x}\}=\{\mathbf{y}\}$ γίνεται

$[\mathbf{L}]\{\mathbf{y}\}=\{\mathbf{b}\}$

Αλλά το γνωστό μητρώο $\mathbf{L}$ δεν εξαρτάται από το διανύσματα $\mathbf{b}$ και η διαδικασία της επίλυσης εξίσωσης $[\mathbf{L}]\{\mathbf{y}\}=\{\mathbf{b}\}$ είναι η πρόσω αντικατάσταση (forward substitution). Στην συνέχεια από λύση $\{\mathbf{y}\}$ βρίσκουμε την λύση του συστήματος $\{\mathbf{x}\}$ με ανάδρομη αντικατάσταση (back substitution) του συστήματος $[\mathbf{U}]\{\mathbf{x}\}=\{\mathbf{y}\}$ όπου το μητρώο $[\mathbf{U}]$ είναι πάνω διαγώνιο. Σύμφωνα με τα παραπάνω ο αλγόριθμος επίλυσης γραμμικού συστήματος με την μέθοδο παραγοντοποιήσης περιλαμβάνει τρεις φάσεις:

(1) Παραγοντοποιήση $\mathbf{A}=\mathbf{L}\ \mathbf{U}$
(2) Πρόσω αντικατάσταση (forward substitution) $[\mathbf{L}]\{\mathbf{y}\}=\{\mathbf{b}\}$
(3) Ανάδρομη αντικατάσταση (back substitution) $[\mathbf{U}]\{\mathbf{x}\}=\{\mathbf{y}\}$

Η διαδικασία επίλυσης γραμμικών συστημάτων με κάτω-άνω διαγώνια παραγοντοποιήση ($\mathbf{L}-\mathbf{U}$ decomposition) είναι η πλέον συνήθης "ακριβής" μέθοδος επίλυσης γραμμικών συστημάτων γενικής μορφής και έχει το πλεονέκτημα ότι μπορεί να χρησιμοποιηθεί για πολλαπλές επιλύσεις συστημάτων με διαφορετικό διανύσματα $\mathbf{b}$ όταν τα μητρώα $\mathbf{L}$ και $\mathbf{U}$ αποθηκευτούν. Η υλοποίηση της μέθοδου επίλυσης με κάτω-άνω διαγώνια παραγοντοποιήση δείχνεται στο παρακάτω παράδειγμα.

Παράδειγμα

Λύση συστήματος με την μέθοδο παραγοντοποιήσης **LU**

$$\begin{bmatrix} 2 & -3 & 1 \\ 1 & -1 & -2 \\ 3 & 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 7 \\ -2 \\ 0 \end{bmatrix}$$

(1) Παραγοντοπίηση το υ μητρώου συντελεστών

$$\begin{bmatrix} 2 & -3 & 1 \\ 1 & -1 & -2 \\ 3 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1.0 & 0 & 0 \\ 0.5 & 1.0 & 0 \\ 1.5 & 11. & 1.0 \end{bmatrix} \begin{bmatrix} 2. & -3. & 1 \\ 0 & 0.5 & -2.5 \\ 0 & 0 & 25.0 \end{bmatrix}$$

(2) Πρόσω αντικατάσταση (forward substitution) $[\mathbf{L}]\{\mathbf{y}\} = \{\mathbf{b}\}$

$$\begin{bmatrix} 1.0 & 0 & 0 \\ 0.5 & 1.0 & 0 \\ 1.5 & 11. & 1.0 \end{bmatrix} \begin{Bmatrix} y_1 \\ y_2 \\ y_3 \end{Bmatrix} = \begin{Bmatrix} 7 \\ -2 \\ 0 \end{Bmatrix} \Rightarrow \qquad \begin{aligned} y_1 &= 7.0 \\ y_2 &= -2. - 0.5(7) = -5.5 \\ y_3 &= -1.5(7) + 11(5.5) = 50 \end{aligned}$$

(3) Ανάδρομη αντικατάσταση (back substitution) $[\mathbf{U}]\{\mathbf{x}\} = \{\mathbf{y}\}$

$$\begin{bmatrix} 2. & -3. & 1 \\ 0 & 0.5 & -2.5 \\ 0 & 0 & 25.0 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} = \begin{Bmatrix} 7.0 \\ -5.5 \\ 50. \end{Bmatrix} \Rightarrow \qquad \begin{aligned} x_3 &= 50./25. = 2. \\ x_2 &= (-5.5 + 2.5(2))/0.5 = -1 \\ x_1 &= (7 - 2 + 3(-1))/2 = 1 \end{aligned}$$

5.2.4 Επίλυση τριδιαγωνίων συστημάτων

Η επίλυση τρι-διαγωνίων συστημάτων απαντάται συχνά σε πρακτικές εφαρμογές όπως η αριθμητική επίλυση συνήθων διαφορετικών εξισώσεων και εξισώσεων με μερικές παραγώγους. Τα τρι-διαγώνια συστήματα έχουν την μορφή.

$$\begin{bmatrix} a_{1,1} & a_{1,2} & 0 & . & . & . & 0 \\ a_{2,1} & a_{2,2} & a_{2,3} & 0 & . & . & 0 \\ 0 & a_{3,2} & a_{3,3} & a_{3,4} & 0 & . & 0 \\ . & . & . & . & . & . & . \\ . & & & & a_{n-1,n-2} & a_{n-1n,n-1} & a_{n-1,n} \\ 0 & . & . & . & 0 & a_{n,n-1} & a_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \\ \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \\ \\ b_{n-1} \\ b_n \end{bmatrix} \qquad (5.2.18)$$

ή

$$\begin{bmatrix} \beta_1 & \gamma_1 & 0 & 0 & . & 0 \\ a_2 & \beta_2 & \gamma_2 & 0 & & \\ 0 & a_3 & \beta_3 & \gamma_3 & & \\ & & & & & \\ & & & a_{n-1} & \beta_{n-1} & \gamma_{n-1} \\ 0 & & & 0 & a_n & \beta_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \\ b_{n-1} \\ b_n \end{bmatrix}$$

Η παραγοντοποίηση του τρι-διαγωνίου μητρώου είναι πιο αποδοτική με την χρήση του αλγόριθμου Thomas που υλοποιείται ως ακολούθως

$$\mathrm{A} = LD^{-1}R = \begin{bmatrix} d_1 & & \\ a_{2,1} & d_2 & \\ & & \\ 0 & a_{n,n-1} & d_n \end{bmatrix} \begin{bmatrix} d_1^{-1} & 0 & 0 \\ 0 & d_2^{-1} & \\ & & \\ 0 & & d_n^{-1} \end{bmatrix} \begin{bmatrix} d_1 & a_{1,2} & . & 0 \\ & d_2 & a_{2,3} & 0 \\ & & & \\ 0 & & & d_n \end{bmatrix}$$

$$A = Lu = \begin{bmatrix} d_1 & 0 & 0 \\ a_{2,1} & d_2 & \\ & & \\ 0 & a_{n,n-1} & d_n \end{bmatrix} \begin{bmatrix} 1 & a_{1,2}/d_2 & 0 \\ & 1 & \\ & & \\ 0 & & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} d_1 & a_{1,2} & 0 \\ a_{2,1} & \left(d_2 + a_{2,1}d_1^{-1}a_{1,2}\right)a_{2,3} & \\ & & \\ 0 & & \left(d_n + a_{n,n-1}d_{n-1}^{-1}\, a_{n-1,n}\right) \end{bmatrix}$$

όπου τα διαγώνια στοιχεία υπολογίζονται από την επαναληπτική σχέση

$$d_1 = a_{1,1} \quad , \quad d_i = a_{i,i} - a_{i,i-1}\, d_{i-1}^{-1}\, a_{i-1,i} \qquad i = 2,3.....,n$$

και οι προς τα εμπρός και αντίστροφες αντικαταστάσεις υπολογίζονται ως

$$L\vec{y}=\vec{b} \qquad y_1=b_1/d_1 \qquad y_1=\frac{b_1-a_{i-1}y_{i-1}}{d_i}, \qquad i=2,3,.....n$$

$$u\vec{x}=\vec{y} \qquad x_i=y_i-a_{i,i}\frac{x_i+1}{d_i}, \qquad i=n-1,......2,1$$

Οι διαιρέσεις που απαιτούνται στις παραπάνω επαναληπτικές σχέσεις μπορεί να αποφευχθούν με την χρήση της παρακάτω πραγματοποίησης

$$A=LDR=\begin{bmatrix} d_1^{-1} & & 0 \\ a_{2,1} & d_2^{-1} & \\ & & \\ & & a_{n,n-1}, d_n^{-1} \end{bmatrix}\begin{bmatrix} d_1 & & 0 \\ & d_2 & \\ & & \\ 0 & & d_n \end{bmatrix}\begin{bmatrix} d_1^{-1} & a_{1,2} & & 0 \\ & d_2^{-1} & a_{2,3} & 0 \\ & & & \\ 0 & & & d_n^{-1} \end{bmatrix}$$

ή

$$A=Lu=\begin{bmatrix} d_1^{-1} & 0 & & 0 \\ a_{2,1} & d_2^{-1} & & 0 \\ & & & \\ 0 & & a_{n,n-1} & d_n^{-1} \end{bmatrix}\begin{bmatrix} 1 & d_1a_{2,1} & & 0 \\ 0 & 1 & d_2a_{3,2} & 0 \\ & & & \\ 0 & & & 1 \end{bmatrix}$$

όπου

$$d_0=0 \ \ \kappa\alpha\iota \ \ d_i=\left(a_{i,1}-a_{i,i-1}\ d_{i-1}\ a_{i-1,i}\right)^{-1} \quad ,i=1,2,......n$$

$$L\vec{y}=\vec{b} \qquad y_1=d_1b_1, \quad y_i=d_i\left(b_i-a_{i,i-1}y_{i-1}\right), i=2,......,n$$

$$u\vec{x}=y \qquad x_n=y_n, \quad x_i=y_i-d_ia_{i,i+1}x_{i+1} \quad i=n-1,.......,1$$

Ο τροποποιημένος αλγόριθμος χωρίς διαιρέσεις πρέπει να προτιμάται για την επίλυση τρι-διαγωνίων συστημάτων σε vector-computers όπου η πράξη της διαίρεσης απαιτεί σημαντικά μεγαλύτερο χρόνο από τον πολλαπλασιασμό.

Τρι-διαγώνια μητρώα όπου το κάθε στοιχείο a_{ij} είναι ένα τετραγωνικό μητρώο μπορούν να αντιμετωπισθούν με τον ίδιο τρόπο παραγοντοποίησης που αναφέρθηκε παραπάνω. Τα συστήματα αυτά ονομάζονται block – τρι-διαγώνια συστήματα και απαντώνται συχνά σε πρακτικές εφαρμογές αριθμητικής επίλυσης συστημάτων συνήθων διαφορικών ή συστημάτων διαφορικών με μερικές παραγώγους. Οι μέθοδοι παραγοντοποίησης που αναφέρθηκαν παραπάνω εφαρμόζονται χωρίς διαφοροποίηση αλλά επί πλέον αποκτούν την εύρεση του αντιστρόφου των d_{i-1}, που είναι μητρώα m x m όπως και τα blocks $a_{i,i}, a_{i,i+1}$ κτλ., και ο υπολογισμός των γινομένων μητρώων όπως $a_{i,i} - a_{i,i}\ d_{i-1}\ a_{i-1},i$. Δηλαδή για κάθε βήμα απαιτείται να εκτελέσουμε δυο πολλαπλασιασμούς μητρώων m^3 flops η κάθε μία αντιστροφή που έχει υπολογιστικό κόστος επίσης m^3 flops συνολικά 3 n m^3 flops για ένα μητρώο διάστασης n.

5.3 Επαναληπτικές μέθοδοι επίλυσης

Η απλούστερη επαναληπτική μέθοδος για την λύση του συστήματος $A\mathbf{x} = \mathbf{b}$ έχει την μορφή $\mathbf{x}^{k+1} = \mathbf{x}^k - \sigma\left(A\mathbf{x}^k - \mathbf{b}\right),\ k = 0,1,......$ όπου x^0 είναι μια αρχική προσέγγιση της λύσης και σ είναι μια περίμετρος. Το κύριο πλεονέκτημα των επαναληπτικών μεθόδων είναι η απλή υλοποίηση. Κάτω από ορισμένες συνθήκες η επανάληψη συγκλίνει στην λύση του συστήματος. Σε κάθε επανάληψη απαιτείται μόνο πολλαπλασιασμός μητρώου με διάνυσμα και προσθέσεις διανυσμάτων που μπορεί να γίνουν πολύ αποδοτικά σε vector–computers. Το κύριο πλεονέκτημα των επαναληπτικών μεθόδων είναι ότι ο ρυθμός σύγκλισης συχνά είναι βραδύς. Η σύγκλιση επαναληπτικών μεθόδων απαιτείται να ελέγχεται κατά την διάρκεια της επαναληπτικής διαδικασίας διότι συχνά είναι δυνατόν να παρατηρηθεί απόκλιση.

5.3.1 Επαναληπτική μέθοδος Jacobi

Η λύση του συστήματος $A\mathbf{x} = \mathbf{b}$ αρχίζει από μια αρχική πρόβλεψη και πραγματοποιείται με την παρακάτω επαναληπτική διαδικασία.

$$x_1 = \left(b_1 - a_{12}x_3 - a_{13}x_3 - ..... - a_{1n}x_n\right)/\,a_{11}$$

$$x_2 = \left(b_2 - b_{21}x_1 - a_{23}x_3 - ..... - a_{2n}x_n\right)/\,a_{22} \tag{5.3.1}$$

$$x_n = \left(b_n - a_{n1}x_1 - a_{n2}x_2 ..... - a_{n,n-1}x_{n-1}\right)/\,a_{nn}$$

ή

$$x_i = \frac{a_i - \sum\limits_{\substack{j=1 \\ i \neq i}}^{n} a_{ij} x_j}{a_{ii}}, \qquad i = 1,2,......,n \tag{5.3.2}$$

Η παραπάνω επαναληπτική διαδικασία ισχύει ασφαλώς όταν όλοι οι διαγώνιοι όροι a_{ii} του μητρώου Α είναι διάφοροι του μηδενός. Επίσης συνήθως καταβάλλεται προσπάθεια ανακατάταξης των εξισώσεων και των αγνώστων έτσι ώστε να επιτευχθεί διαγώνια κυριαρχία (diagonal dominance), δηλαδή το κάθε διαγώνιο στοιχείο a_{ii} είναι μεγαλύτερο κατά απόλυτη τιμή από το μέγεθος των άλλων στοιχείων στην σειρά *i* και την στήλη *i*.

Η πιστοποίηση της σύγκλισης της επαναληπτικής μεθόδου Jacobi είναι σημαντική στην πρακτική εφαρμογή της μεθόδου για την επίλυση συστημάτων μεγάλου μεγέθους. Η καθιέρωση κριτηρίου σύγκλισης επιτυγχάνεται με την αναδιάταξη της Εξ. 5.3.1 ως ακολούθως

$$\vec{x} = cx + v \tag{5.3.3}$$

όπου

$$c = -\begin{bmatrix} 0 & \frac{a_{12}}{a_{11}} & \frac{a_{13}}{a_{11}} & \frac{a_{in}}{a_{11}} \\ \frac{a_{21}}{a_{22}} & 0 & \frac{a_{23}}{a_{11}} & \frac{a_{2n}}{a_{22}} \\ . & . & . & . \\ \frac{a_{n1}}{a_{nn}} & \frac{a_{n2}}{a_{nn}} & \frac{a_{n3}}{a_{nn}} & 0 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \\ \frac{b_n}{a_{nn}} \end{bmatrix}$$

Υποθέτοντας ότι το αρχικό διάνυσμα $\vec{x}_0$ δεν απέχει σημαντικά από την λύση του συστήματος $\vec{x}$ η σύγκλιση είναι γρήγορη με την επανάληψη

$$x_{k+1} = cx_k + v \tag{2.3.4}$$

δηλαδή

$$x_k = c^k x_0 + \left[I + c + c^2 + ...c^{k-1} \right] v \tag{2.3.5}$$

και η σύγκλιση της επανάληψης επιτυγχάνεται όταν

$$\lim_{k\to\infty} c^k = 0 \tag{2.3.6}$$

όμως η συνθήκη $\lim_{k\to\infty} c^k = 0$ γνωρίζουμε από την θεωρία των μητρώων ότι συνεπάγεται την σύγκλιση του αθροίσματος όταν όλες οι ιδιοτιμές του μητρώου Β είναι κατά απόλυτο τιμή μικρότερες του μηδενός δηλαδή $\lim_{k\to\infty}\left[I + c + c^2 + ... + c^k \right] = \left(I - c\right)^{-1}$.

Συνεπώς όταν ικανοποιείται η συνθήκη $\lim_{k\to\infty} c^k$ τότε η επανάληψη της Εξ. (5.3.3) συγκλίνει

$$\vec{x} = \lim_{k\to\infty} x_k$$

και από την Εξ. (5.3.5) βρίσκουμε

$$\vec{x} = 0 + \left(I - c\right)^{-1} \vec{v}$$

ή $\quad \left(I - c\right)x = v$

$$x = cx + v$$

Οι ικανές συνθήκες για να έχουμε όλες τις ιδιοτιμές μικρότερες κατά απόλυτο τιμή από την μονάδα είναι

$$\sum_{j=1}^{n} \left|c_{ij}\right| \le \mu < 1 \qquad 1 \le j \le n$$

$$\sum_{j=1}^{n} \left|c_{ij}\right| \le \mu < 1 \qquad 1 \le i \le n \tag{5.3.7}$$

$$\sum_{j=1}^{n} \sum_{j=1}^{n} \left|c_{ij}\right|^2 \le \mu < 1$$

Οι παραπάνω συνθήκες συνεπάγονται την ισοδύναμη συνθήκη σύγκλισης για τα στοιχεία του μητρώου Α.

$$\sum_{\substack{j=1\\ j\neq i}}^{n} \left|a_{ij}\right| < \left|\alpha_{ii}\right| \tag{5.3.8}$$

Συνεπώς είναι αναγκαίο να γίνει αναδιάταξη των γραμμών και στηλών ώστε να διασφαλισθεί η συνθήκη (5.3.8) πριν να αρχίσουν οι επαναλήψεις.

5.3.2 Η επαναληπτική μέθοδος Gauss–Seidel

Η επαναληπτική μέθοδος Gauss–Seidel είναι παραλλαγή της μεθόδου Jacobi όπου κατά την διάρκεια των επαναλήψεων τα στοιχεία της λύσης μόλις υπολογίζονται χρησιμοποιούνται στο δεξί μέλος των υπόλοιπων από τις Εξ. (5.3.1). Η σύγκλιση της μεθόδου Gauss–Seidel είναι σημαντικά ταχύτερη από την σύγκλιση της μεθόδου Jacobi. Η μελέτη σύγκλισης της μεθόδου Gauss–Seidel παρουσιάζεται παρακάτω.

Η k^{th} προσέγγιση του i^{th} στοιχείου του διανύσματος των αγνώστων συμβολίζεται ως x_i^k. Έστω ότι $x^0 = \left[x_1^0, x_2^0, \ldots\ldots, x_n^0\right]^T$ είναι αρχική προσέγγιση. Η επανάληψη γράφεται και πάλι στην μορφή της Εξ. (5.3.3) και ορίζουμε

$$x_{ik} = \sum_{j=1}^{i=1} c_{ij} x_i^k + \sum_{j=i+1}^{n} c_{ij} x_i^{k-1} + \upsilon_i \qquad k \geq 1, \quad 1 \leq i \leq n \tag{5.3.9}$$

όπου όταν $i = 1$

$\sum_{j=1}^{i-1} c_{ij}\ x_i^k = 0$ και όταν $i = n$ τότε $\sum_{j=i+1}^{n} a_{ij}\ x_i^{k-1} = 0$

Το μητρώο C της Εξ. (5.3.3) γράφεται ως άθροισμα δυο μητρώων ως ακολούθως

$$C = C_L + C_u = \begin{bmatrix} 0 & 0 & . & 0 & 0 \\ c_{21} & 0 & & & . \\ . & . & & & . \\ c_{n1} & c_{n2} & & c_{n,n-1} & 0 \end{bmatrix} + \begin{bmatrix} 0 & c_{12} & c_{in} \\ 0 & & c_{2n} \\ & & \\ 0 & & 0 \end{bmatrix}$$

τότε η επαναληπτική διαδικασία έχει την μορφή

$$x^k = c_L \; x^k + c_u \; x^{k-1} + v$$

$$ή \qquad (5.3.10)$$

$$x^k = \left(I - c_L\right)^{-1} c_u \; x^{k-1} + \left(I - c_L\right)^{-1} v$$

Η επανάληψη της Εξ. (5.3.10) έχει την μορφή της επανάληψης Jacobi $x_k = Cx_{k-1} + v$ όπου $C = \left(I - C_L\right)^{-1} Cu$ και $v = \left(I - C_L\right)^{-1} v$. Βρήκαμε όμως ότι η αναγκαία συνθήκη για σύγκλιση είναι να έχουμε ιδιοτιμές του μητρώου $\left(I - C_L\right)^{-1} C_u$ κατά απόλυτη τιμή μικρότερες της μονάδας. Οι ιδιοτιμές του μητρώου $\left(I - C_L\right)^{-1} C_u$ είναι όμως ρίζες του πολυωνύμου

$$\det\left[\left(I - C_L\right)^{-1} C_u - \lambda I\right] = 0$$

$$ή \qquad (5.3.11)$$

$$\det\left[C_u - \lambda I + \lambda C_L\right] = 0$$

Δηλαδή η επανάληψη Gauss-Seidel συγκλίνει όταν οι ρίζες που υπολογίζονται από την ορίζουσα

$$\det\begin{bmatrix} -\lambda & c_{12} & c_{13} & .. & c_{1n} \\ c_{21}\lambda & -\lambda & c_{23} & .. & c_{2n} \\ c_{31}\lambda & c_{32}\lambda & -\lambda & .. & c_{3n} \\ .. & .. & .. & .. & .. \\ c_{n1}\lambda & c_{n2}\lambda & c_{3n}\lambda & .. & -\lambda \end{bmatrix} = 0 \qquad (5.3.12)$$

Είναι όλες μικρότερες κατά απόλυτη τιμή από την μονάδα. Τα στοιχεία του παραπάνω μητρώου είναι $C_{ij} = -a_{ij} / a_{ii} \quad \forall \; i \neq j$ συνεπώς η ορίζουσα της Εξ. (5.3.12) είναι

$$\det\begin{bmatrix} a_{11}\lambda & a_{12} & a_{13} & .. & a_{1n} \\ a_{21}\lambda & a_{22}\lambda & a_{23} & .. & a_{2n} \\ a_{31}\lambda & a_{32}\lambda & a_{33}\lambda & .. & a_{3n} \\ \\ a_{n1}\lambda & a_{n2}\lambda & a_{n3}\lambda & .. & a_{nn}\lambda \end{bmatrix} = 0 \qquad (5.3.13)$$

και μπορεί να αποδειχθεί ότι η συνθήκη που εγγυάται την σύγκλιση της μεθόδου Gauss-Seidel είναι

$$\sum_{\substack{j=1 \\ i\neq j}}^{n} \left|\frac{a_{ij}}{a_{ii}}\right| \leq \mu < 1$$

ή (5.3.14)

$$\sum_{\substack{j=1 \\ i\neq j}}^{n} \left|\frac{a_{ji}}{a_{ii}}\right| \leq \mu < 1$$

Μια ισχυρότερη συνθήκη σύγκλισης επιτυγχάνεται όταν το μητρώο των συντελεστών A είναι θετικό ορισμένο διότι στην περίπτωση αυτή όπως θα δείξουμε παρακάτω πάντοτε επιτυγχάνεται σύγκλιση της επανάληψης Gauss-Seidel.

Έστω ότι το μητρώο A γράφεται ως

$$\mathrm{A} = D + L + \overline{L}^{\,T} \qquad (5.3.15)$$

όπου $D = diag\left(a_{11}, a_{22}, ..., a_{nn}\right)$ και L είναι το κάτω τριγωνικό μητρώο που περιέχει τους συντελεστές του A κάτωθεν της διαγώνιου. Η συνθήκη ευστάθειας που βρήκαμε προηγουμένως είναι ότι όλε οι ιδιοτιμές του μητρώου $\left(I - C_L\right)^{-1} Cu$ είναι κατά απόλυτη τιμή μικρότερες της μονάδας. Αλλά $C_L = -D^{-1}L$ και $C_u = -D^{-1}L^*$ ή $\left(I - C_L\right)^{-1} C_u = -\left(D+L\right)^{-1} L^*$.

Έστω ότι οι ιδιο-τιμές του μητρώου $\left(D+L\right)^{-1} L^*$ είναι λ_i και $\vec{w}_i$ είναι τα αντίστοιχα ιδιο-διανύσματα. Επειδή όμως εξ υποθέσεως το μητρώο A είναι θετικά ορισμένο έχουμε

$$\left(\vec{w}_i, A\,\vec{w}_i\right) = \left(w_i, D\,w_i\right) + \left(w_i, L\,w_i\right) + \left(w_i, L^*\,w_i\right) \succ 0 \qquad (5.3.16)$$

αλλά $\left(D+L\right)^{-1} L^* = \lambda I$ ή $\left(D+L\right)^{-1} L^* \vec{w}_i = \lambda_i \vec{w}_i$ δηλαδή

$$L^* \vec{w}_i = \lambda_i D w_1 + \lambda_i L w_i$$

τότε

$$\left(w_i , L^* w_i \right) = \lambda_i \left[\left(w_i , D w_i \right) + \left(w_i L w_i \right) \right] \tag{5.3.17}$$

Θεωρώντας το συζυγές των δυο μελών έχουμε

$$\left(L^* w_i , w_i \right) = \left(w_i , L w_i \right) = \lambda_i \left[\left(D w_i , w_i \right) + \left(L w_i , w_i \right) \right]$$

ή (5.3.18)

$$\left(w_i , L w_i \right) = \lambda_i \left[\left(w_i , D w_i \right) + \left(w_i , L^* w_i \right) \right]$$

Από τις παραπάνω σχέσεις βρίσκουμε

$$\left(1 - \lambda_i \overline{\lambda}_i \right) \left(w_i , L^* w_i \right) = \left(\lambda_i + \lambda_i \overline{\lambda}_i \right) \left(w_i , D w_i \right)$$

ή (5.3.19)

$$\left(1 - \lambda_i \overline{\lambda}_i \right) \left(w_i , L w_i \right) = \left(\overline{\lambda}_i + \lambda_i \overline{\lambda}_i \right) \left(w_i , D w_i \right)$$

Με αντικατάσταση των παραπάνω στην σχέση (5.3.16) βρίσκουμε ότι

$$\left(1 + \lambda_i \right) \left(1 + \overline{\lambda}_i \right) \left(w_i D w_i \right) > 0$$

επειδή το διαγώνιο μητρώο D είναι θετικά ορισμένο

$$\left(1 - \lambda_i \overline{\lambda}_i \right) > 0 \Rightarrow \left| \lambda_i \right| < 1$$

Η μέθοδος επίλυσης γραμμικών συστημάτων με Lu παραγοντοποίηση ή απαλοιφή Gauss και οι επαναληπτικές μέθοδοι Jacobi και Gauss-Seidel είναι οι κλασσικές μέθοδοι επίλυσης των συστημάτων. Παραλλαγές αυτών των μεθόδων και άλλες επαναληπτικές μέθοδοι επίλυσης συστημάτων μπορεί να βρεθούν σε πιο εξειδικευμένα εγχειρίδια και δημοσιεύσεις. Στην συνέχει αυτού του κεφαλαίου θα ασχοληθούμε με προσεγγιστικές μεθόδους επίλυσης γραμμικών συστημάτων μεγάλου μεθόδου.

5.4. Μέθοδοι conjugate-Gradient

Οι μέθοδοι συζυγούς κλίσης (conjugate –gradient methods) είναι επαναληπτικές μέθοδοι που χρησιμοποιούνται για την εύρεση της προσεγγιστικής λύσης γραμμικών συστημάτων μεγάλου μεγέθους. Προσδιορίζουν την προσεγγιστική λύση του συστήματος $A\vec{x}=\vec{b}$ ελαχιστοποιώντας το τετραγωνικό συναρτησιακό

$$f(\vec{x})=\frac{1}{2}\left(\vec{x}^T A\,\vec{x}\right)-\vec{b}^T x$$

όταν το μητρώο A είναι συμμετρικό και θετικά ορισμένο (symmetric and positive definite spa) ή το συναρτησιακό του υπόλοιπου

$$f(\vec{x})=\left(A\vec{x}-\vec{b}\right)^T\left(A\vec{x}-\vec{b}\right)$$

στην γενική περίπτωση που το μητρώο A έχει τυχαία δομή . Όταν το μητρώο είναι συμμετρικό και θετικά ορισμένο και το μητρώο A είναι nonsingular τότε

$$\frac{1}{2}\left(\mathrm{A}\vec{x}-\vec{b}\right)\mathrm{A}^{-1}\left(\mathrm{A}\,\vec{x}-b\right)=\frac{1}{2}\vec{x}^T A\vec{x}-\vec{b}\;\vec{x}+\frac{1}{2}\vec{b}^T A^{-1}\vec{b}$$

αλλά ο όρος $\vec{b}^T\mathrm{A}^{-1}\vec{b}$ είναι σταθερός και επειδή το μητρώο A είναι θετικά ορισμένο $\left(\mathrm{A}x,x\right)>0$ το μοναδικό ελάχιστο του παραπάνω συναρτησιακού ορίζεται από την λύση του συστήματος $A\vec{x}=\vec{b}$.

Η διαδικασία ελαχιστοποίησης πραγματοποιείται σε μια σειρά διανυσματικών υπόχωρων V_k με συνεχώς αυξανόμενη διάσταση. Οι υπόχωροι αυτοί δημιουργούνται από τους προηγούμενος προσθέτοντας νέα διάνυσμα βάσης $A^k\vec{r}^o$ στους διανυσματικούς υπόχωρους V_{k-1} της προηγούμενης επανάληψης, δηλαδή

$$V_k=V_{k-1}\oplus\left\{A^k\vec{r}^o\right\},\quad k=1,2,3,......$$

όπου V_o είναι ο υπόχωρος που αντιστοιχεί στο υπόλοιπο της αρχικής επανάληψης $\vec{r}^{o} = A\vec{x}^{0} - \vec{b}$, $V_o = \{r^0\}$ ο υπόχωρος V_k που συντίθεται από τα διανύσματα $\vec{r}^{0}$, $A\vec{r}^{0},....., \mathrm{A}^{k}\vec{r}^{0}$ ονομάζεται χώρος krylov που αντιστοιχεί στο μητρώο A και το αρχικό υπόλοιπο $\vec{r}^{0}$.

Θεωρούμε το εσωτερικό γινόμενο $(\vec{x},\vec{y})$ στον χώρο $\mathbb{R}^{n}$ που ορίζεται ως $(\vec{x},\vec{y}) = x^{T} W\, \vec{y}$ όπου W είναι ένα συμμετρικό και θετικά ορισμένο μητρώο $(W\,\vec{x},\vec{x}) > 0$ τα διανύσματα $\vec{x}$ και $\vec{y}$ ονομάζονται συζυγή ορθογώνια (conjugate orthogonal). Επί πλέον υποθέτουμε ότι $(\vec{x},\mathrm{A}\vec{y}) = (\mathrm{A}\vec{x},\vec{y})$ δηλαδή το μητρώο A είναι ατοσυζυγές (selfadjoint) ως προς το εσωτερικό γινόμενο. Η επίλυση του συστήματος $A\vec{x} = \vec{b}$ πραγματοποιείται με την ακόλουθη επαναληπτική μέθοδο η οποία δημιουργεί διανύσματα υπολοίπου που είναι ορθογώνια ως προς το εσωτερικό γινόμενο.

Έστω ότι $\vec{x}^{0}$ είναι η αρχική προσέγγιση της λύσης τότε

$$\vec{x}^{1} = \vec{x}^{0} - \beta_0\, \vec{r}^{o} \tag{5.4.1}$$

Πολλαπλασιάζοντας την παραπάνω σχέση με A και αφαιρώντας και από τα δυο μέλη το διάνυσμα $\vec{b}$ βρίσκουμε

$$\vec{r}^{1} = \vec{r}^{0} - \beta_0\, \mathrm{A}\, \vec{r}^{0} \tag{5.4.2}$$

όπου $r^{k} = Ax^{k} - b$ είναι το υπόλοιπο της k^{th} επανάληψης. Υποθέτοντας ότι $\vec{r}^{0} \neq 0$ η τιμή της παραμέτρου βρίσκεται επιβάλλοντας καθετότητα των διαδοχικών υπολοίπων $(\vec{r}^{1},\vec{r}^{0}) = 0$.

$$\beta_0 = \frac{\vec{r}^{0},\vec{r}^{0}}{(\vec{r}^{0},\, \mathrm{A}\,\vec{r}^{0})} \tag{5.4.3}$$

Σύμφωνα με τους παραπάνω η σχέση επανάληψης είναι

$$\vec{x}^{k+1} = a_k \vec{x}^{k} + (1 - a_k)\vec{x}^{k-1} - \beta_k\, \vec{r}^{k} \qquad k = 1,2,..... \tag{5.4.4}$$

όπου $\alpha_0 = 1$

Πολλαπλασιάζοντας την παραπάνω επαναληπτική σχέση με Α και αφαιρώντας το διάνυσμα $\vec{b}$ βρίσκουμε ότι τα υπόλοιπα της επανάληψης ικανοποιούν

$$\vec{r}^{\,k+1} = a_k \vec{r}^{\,k} + \left(1 - a_k\right)\vec{r}^{\,k-1} - \beta_k \; A\,\vec{r}^{\,k}, \qquad k = 1,2,...... \tag{5.4.5}$$

Οι σταθερές των Εξ. (5.4.4) και (5.4.5) προσδιορίζονται υποθέτοντας ότι τα υπόλοιπα είναι ορθογώνια $\left(\vec{r}^{\,k}, \vec{r}^{\,k-j}\right) = 0$ $\left(j = 1,2,....,k\right)$, $\left(\vec{r}^{\,k+1}, \vec{r}^{\,k}\right) = 0$, $\left(\vec{r}^{\,k+1}, \vec{r}^{\,k-1}\right) = 0$, $\left(\vec{r}^{\,k+1}, \vec{r}^{\,k-j}\right) = 0$, $j = 0,1,...,k$, δηλαδή επιβάλλοντας τα νέα υπόλοιπα να είναι ορθογώνια με όλα τα προηγούμενα. Οι σταθερές της επανάληψης της Εξ. (5.4.5) που υπολογίζονται με βάση τις παραπάνω υποθέσεις είναι

$$a_k = \beta_k \; \mu_k \tag{5.4.6}$$

$$\mu_k = \frac{\vec{r}^{\,k}, \mathrm{A}\vec{r}^{\,k}}{\left(\vec{r}^{\,k}, \vec{r}^{\,k}\right)} = \frac{\left(\vec{r}^{\,k}, \mathrm{A}\vec{r}^{\,k}\right)}{\delta_k}$$

$$\delta_k = \left(\vec{r}^{\,k}, \vec{r}^{\,k}\right)$$

$$\alpha_k = \delta_{k-1} + \beta_k \left(\vec{r}^{\,k-1}, \mathrm{A}\vec{r}^{\,k}\right) = \delta_{k-1}$$

$$\beta_k \mu_k \delta_{k-1} - \beta_k \; \delta_k \; \beta_{k-1}^{-1} = \delta_{k-1}$$

$$\beta_k^{-1} = \mu_k - \frac{\delta_k}{\delta_{k-1}} \beta_{k-1}^{-1}, \;\; k = 1,2,.... \tag{5.4.7}$$

Οι συντελεστές που υπολογίζονται από τις παραπάνω σχέσεις είναι θετικοί και $a_k > 1$, $\beta_k > 0$. Επί πλέον μπορεί να αποδειχθεί ότι η επαναληπτική μέθοδος της Εξ. (5.4.5) είναι βέλτιστη διότι τα υπόλοιπα που προκύπτουν από την διαδοχική εφαρμογή των Εξ. (5.4.5) (5.4.6) και (5.4.7) είναι τα ελάχιστα δυνατά έστω και όταν το διάνυσμα x^{k+1} είναι γραμμικός συνδυασμός όλων των προηγουμένων διανυσμάτων x^k.