ΚΕΦΑΛΑΙΟ IV

Μητρώα Ειδικών Μορφών και Διασπάσεις

Στο κεφάλαιο αυτό δίνονται τα απαραίτητα στοιχεία για ειδικές μορφές μητρώων (πινάκων), τις ιδιότητές τους και το ρόλο τους στη διαμόρφωση αποδοτικών σχημάτων επίλυσης γραμμικών συστημάτων $A\boldsymbol{x}=\boldsymbol{b}$, βασισμένων στη διάσπαση του μητρώου A σε ειδικές μορφές LU. Στην ενότητα 4.1 δίνουμε πρώτα τη γενική μορφή διάσπασης LU. Στις επόμενες ενότητες παρουσιάζουμε ειδικές μορφές μητρώων και διασπάσεων. Τέλος στην ενότητα 4.7 δίνονται παραπομπές στη βιβλιογραφία.

4.1 Παραγοντοποίηση LU

Η μέθοδος παραγοντοποίησης ή διάσπασης LU αναφέρεται στο πρόβλημα της διάσπασης ενός μητρώου A στη μορφή

$$\boldsymbol{A = LU} \tag{4.1.1}$$

όπου L είναι κάτω τριγωνικό ($l_{ij}=0$, $i>j$) και U πάνω τριγωνικό μητρώο ($l_{ij}=0$ για $j>i$). Αν το μητρώο A ενός γραμμικού συστήματος $A\boldsymbol{x}=\boldsymbol{b}$ μπορεί να διασπαστεί με τον τρόπο αυτό, τότε το σύστημα γράφεται $LU\boldsymbol{x}=\boldsymbol{b}$ και η επίλυσή του ανάγεται στην επίλυση δυο απλών συστημάτων: του $L\boldsymbol{y}=\boldsymbol{b}$ ως προς $\boldsymbol{y}$ με εμπρός αντικατάσταση και του $U\boldsymbol{x}=\boldsymbol{y}$ ως προς $\boldsymbol{x}$ με πίσω αντικατάσταση. Η μέθοδος αυτή είναι αποδοτική στη περίπτωση που έχουμε πολλά συστήματα με κοινό A και διαφορετικά δεύτερα μέλη $\boldsymbol{b}$, οπότε η παραγοντοποίηση γίνεται μια μόνο φορά.
Βάση για την ανάπτυξη της μεθόδου διάσπασης LU είναι η μέθοδος *Gauss*. Όταν η διάσπαση είναι δυνατή, το μητρώο L περιέχει τους πολλαπλασιαστές p_{ij} της μεθόδου Gauss και το U είναι το τελικό πάνω τριγωνικό μητρώο A_{n-1}. Συγκεκριμένα, έχουμε τα παρακάτω αποτελέσματα.

Θεώρημα 4.1.1

Αν κατά την τριγωνοποίηση ενός αντιστρεπτού μητρώου A δεν γίνει καμία αντιμετάθεση γραμμών, τότε το A παραγοντοποιείται στη μορφή LU.

Απόδειξη: Αφού δεν γίνεται καμία αντιμετάθεση γραμμών θα είναι $P_i=I_n$, $i=1,\ldots,n-1$, οπότε η (4.1.1) γίνεται :

$$U = M_{n-1}\ldots M_2 M_1 A \tag{4.1.2}$$

Το γινόμενο $M_{n-1}\ldots M_1$ είναι αντιστρέψιμο αφού οι M_i είναι αντιστρέψιμοι. Από την (ΙΙΙ.5.2) έχουμε $A=(M_{n-1}\ldots M_1)^{-1}U = M_1^{-1}\ldots M_{n-1}^{-1}U$. Θέτουμε $L=M_1^{-1}\ldots M_{n-1}^{-1}$, οπότε $A=LU$. Τα μητρώα απαλοιφής M_i δίνονται από την (3.5.2). Επαληθεύεται δε εύκολα (ιδιότητες πολλαπλασιασμού μητρώων):

$$L = M_1^{-1} M_2^{-1}\ldots M_{n-1}^{-1} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ p_{21} & 1 & \ldots & 0 \\ p_{31} & p_{32} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ p_{n1} & p_{n2} & \ldots & 1 \end{pmatrix} \tag{4.1.3}$$

Άρα $A=LU$ όπου L είναι κάτω τριγωνικό με διαγώνια στοιχεία 1. ■

Εξετάζουμε τώρα τη περίπτωση που έχουμε εναλλαγές γραμμών κατά την τριγωνοποίηση. Το Θ(4.1.1) γενικεύεται από το εξής αποτέλεσμα:

Θεώρημα 4.1.2

Για κάθε αντιστρέψιμο μητρώο A n×n υπάρχουν ένα κάτω τριγωνικό μητρώο L, ένα πάνω τριγωνικό μητρώο U και ένα μεταθετικό P τέτοια ώστε

$$PA=LU \tag{4.1.4}$$

Απόδειξη. Θεωρούμε τα στοιχειώδη μητρώα $M_1,M_2,\ldots,M_{n-1}$ και $P_1,P_2,\ldots P_{n-1}$ που προκύπτουν κατά την φάση της απαλοιφής. Θεωρούμε επίσης τα μητρώα:

$$L_1 = M_1,\ L_2 = M_2(P_2L_1P_2)\ ,\ldots,\ \ L_{n-1} = M_{n-1}(P_{n-1}L_{n-2}P_{n-1}) \tag{4.1.5}$$

που είναι αντιστρέψιμα ως γινόμενα αντιστρέψιμων μητρώων.
Μπορούμε τώρα να δείξουμε ότι τα L_i είναι κάτω τριγωνικά χρησιμοποιώντας επαγωγή. Για $k=1$ ισχύει, αφού $L_1=M_1$ είναι κάτω τριγωνικό και με μόνη μη μηδενική κάτω από την μονάδα την 1η στήλη. Υποθέτουμε τώρα ότι L_{k-1} είναι κάτω τριγωνικό και με όλες τις στήλες j, $k-1<j<n$, μηδενικές κάτω από τις μονάδες. Θα δείξουμε ότι και το $L_k=M_k(P_kL_{k-1}P_k)$ είναι ομοίως κάτω τριγωνικό. Επειδή το M_{k-1} είναι κάτω τριγωνικό, αρκεί να δειχθεί ότι και το $P_kL_{k-1}P_k$ είναι κάτω τριγωνικό. Το P_k εναλλάσσει την k-στήλη καιk-γραμμή με μια *επόμενη* i-στήλη ($i>k$) και i-γραμμή αντίστοιχα του L_{k-1}. Στη μετάθεση γραμμών γίνεται εναλλαγή των στοιχείων της k-γραμμής $L_{k-1}(k,1),\ldots,L_{k-1}(k,k-1)$ με τα αντίστοιχα της i-γραμμής. Αυτό πάντως δεν επιδρά στη γεωμετρία του τριγωνικού L_{k-1}. Ακόμα θα γίνουν οι εξής αλλαγές: $(P_kL_{k-1})_{kk}=0$, $(P_kL_{k-1})_{ii}=0$, $(P_kL_{k-1})_{ik}=1$, $(P_kL_{k-1})_{ki}=1$. Η εναλλαγή των στηλών k και i όμως αποκαθιστά τα στοιχεία στις αρχικές τους θέσεις και το $P_kL_{k-1}P_k$ επανακτά τη γεωμετρία του, παραμένοντας κάτω τριγωνικό.
Επειδή τώρα το αντίστροφο ενός κάτω τριγωνικού είναι κάτω τριγωνικό, άρα και το $L=L_{n-1}^{-1}$ είναι κάτω τριγωνικό. Έχοντας υπόψη ότι $P_k=P_k^{-1}$ ή $P_k^2=I$, το U εκφράζεται

$$\begin{aligned}U &= M_{n-1}P_{n-1}\ldots M_2P_2M_1P_1A= M_{n-1}P_{n-1}\ldots M_2\ (P_2M_1P_2)\ P_2\ P_1A\\ &= M_{n-1}P_{n-1}\ldots M_3P_3L_2\ (P_2P_1A) = M_{n-1}P_{n-1}\ldots M_4P_4M_3P_3L_2P_3\ (P_3P_2P_2A)\\ &= M_{n-1}P_{n-1}\ldots M_4P_4L_3\ (P_3P_2P_1A) =\\ &= \ldots = L_{n-1}P_{n-1}\ldots P_1A = L_{n-1}PA\end{aligned}$$

Θέσαμε $P=P_{n-1}\ldots P_1$, δηλ. P είναι το μεταθετικό μητρώο που αντιστοιχεί στις μεταθέσεις γραμμών της μεθόδου *Gauss*. Επομένως ισχύει $U=L_{n-1}PA \Rightarrow PA=L_{n-1}^{-1}U$, δηλαδή $PA=LU$.
Σημείωση: Αν δεν υπήρχαν εναλλαγές γραμμών θα ήταν $P_k=I_n$, $k=1,\ldots,n-1$, και επομένως: $P=I_n$ και $L_1=M_1$, $L_2=M_2M_1,\ldots,L_{n-1}=M_{n-1}\ldots M_1$. Άρα $A=LU$, όπου $L=L_{n-1}^{-1}$, δηλ. επαληθεύεται το Θ4.1.1. ■

Παρατήρηση 4.1.1 Είναι φανερό ότι αλλάζοντας τη μέθοδο οδήγησης (π.χ. μερική οδήγηση) για ένα μητρώο A, προκύπτει και διαφορετική διάσπαση. Τα L και U αλλάζουν αφού αλλάζουν οι οδηγοί και οι πολλαπλασιαστές μετά τις εκτελούμενες μεταθέσεις γραμμών.

Παρατήρηση 4.1.2: Αν $P=I$ και διαιρέσουμε κάθε γραμμή του U με τον οδηγό της d_i προκύπτει $U=D\mathrm{U}$, όπου $D=\text{diag}([d_1,d_2,\ldots,d_n])$ και U κάτω τριγωνικό με μονάδες στη διαγώνιο και τα υπόλοιπα στοιχεία του διαιρεμένα με τους οδηγούς. Τότε η (4.1.1) γράφεται:

$$A=LD\mathrm{U} \tag{4.1.6}$$

Η (4.1.6) αποτελεί κανονικοποιημένη έκφραση της διάσπασης LU. Ως άμεσο αποτέλεσμα προκύπτει $A^T=(LD\mathrm{U})^T= \mathrm{U}^TDL^T$, δηλαδή ότι και το A^T παραγοντοποιείται:

$$A^T = \mathrm{U}^TDL^T = \mathrm{U}^T(DL^T) =\mathrm{U}^T\mathrm{L}^T \tag{4.1.7}$$

Το συμπέρασμα που εξάγουμε από το Θ(4.1.2) είναι ενδιαφέρον: ένα αντιστρέψιμο μητρώο A είναι είτε άμεσα παραγοντοποιήσιμο στη μορφή LU, όταν κατά την φάση της απαλοιφής δεν απαιτούνται μεταθέσεις γραμμών (δηλ. $P=I$), είτε έμμεσα κάτω υπό τη μορφή PA, δηλ. με εφαρμογή κατάλληλων μεταθέσεων γραμμών στο A. Το P συγκεντρώνει το ιστορικό των μεταθέσεων. Προφανώς η μέθοδος LU εφαρμόζεται στο PA χωρίς εναλλαγές γραμμών. Για τις δύο αυτές περιπτώσεις ισχύουν τα ακόλουθα:

▸ **Περίπτωση $P=I$.** Κατά την απαλοιφή όλοι οι πολλαπλασιαστές p_{ij} συγκεντρώνονται σταδιακά και τοποθετούνται στο L: $L(i,k)=p_{ik}$, για $i>k$. Στην πράξη όμως, για την αποθήκευση των πολλαπλασιαστών μπορεί να χρησιμοποιηθεί το κάτω τριγωνικό τμήμα του A (αποθήκευση κατά στήλες), ενώ για την αποθήκευση του U, το άνω τριγωνικό. Ο αλγόριθμος LU διατυπώνεται ως εξής:

Αλγόριθμος 4.1.1: Παραγοντοποίηση A=LU

Input: *A, n*
Output: *L: Κάτω τριγωνικό τμήμα του A, U: Άνω τριγωνικό τμήμα του A*

```
for k=1:1:n-1
{d:=A(k,k);                          %Υπολογισμός οδηγού d_k
 if d=0 then exit('παραγοντοποίηση αδύνατη')
for i=k+1:1:n
  { p:= A(i,k)/d;                    %Υπολογισμός πολλαπλασιαστή p_ik για τη γραμμή i.
    A(k,i):=p;                       % Αποθήκευση p_ik στη k-στήλη του κάτω τριγωνικού μέρους του A
    A(i,k:n):=A(i,k:n)-p*A(k,k:n);   % Απαλοιφή γραμμής k
  }
}
if A(n,n)=0 then exit('παραγοντοποίηση αδύνατη');
end
```

▸ **Περίπτωση $P\neq I$.** Όταν $P\neq I$, οι πολλαπλασιαστές συγκεντρώνονται και πάλι στο L αλλά με διαφορετική διάταξη, η οποία παρουσιάζεται στη συνέχεια. Κατόπιν της (4.1.5), σε κάθε βήμα $k=2,\ldots,n-1$ υπολογίζεται:

$$L_k^{-1}= (P_kL_{k-1}P_k)^{-1}M_k^{-1}= (P_k^{-1}L_{k-1}^{-1}P_k^{-1})^{-1}M_k^{-1} = (P_kL_{k-1}^{-1}P_k)\, M_k^{-1} \tag{4.1.8}$$

Θέτουμε, σύμφωνα και με την (4.1.5): $L_{(k)}=L_k^{-1}$, $L_{(0)}=I_n$, οπότε $L_{(1)}=(P_1L_{(0)}P_1)M_1^{-1}=(I_n)M_1^{-1}=M_1^{-1}$. Συνεπώς το L υπολογίζεται από την ακολουθία:

$$L_{(k)}=(P_kL_{(k-1)}P_k)\, M_k^{-1},\ \text{για } k=1,2,\ldots,n-1 \qquad L= L_{(n-1)} \tag{4.1.9}$$

Σε κάθε βήμα k≥2 το P_k εναλλάσσει τη k-γραμμή (στήλη) με κάποια επόμενη γραμμή (στήλη) j ($j>k$), δηλ. ισχύει $P_k =P_{kj}$. Επίσης, τα $L_{(k)}$ διατηρούνται κάτω τριγωνικά, για τους ίδιους λόγους που είδαμε στην απόδειξη του Θ4.1.2. Για $k=1$, το $L_{(1)}$ έχει στην στήλη 1 τους πολλαπλασιαστές p_{i1} ($2\leq i\leq n$) με την αρχική τους διάταξη. Για $k=2$ εναλλάσσονται οι p_{k1} και p_{j1} στη στήλη 1 (επενέργεια πολλαπλασιασμού αριστερά με P_2), ενώ στη στήλη 2 γίνεται προσθήκη των πολλαπλασιαστών p_{i2}, $3\leq i\leq n$ (επενέργεια πολλαπλασιασμού δεξιά με M_2^{-1}).
Στο βήμα k γίνεται εναλλαγή των πολλαπλασιαστών $p_1^{(k)} = L_{k-1}(k,1),\ldots, p_{k-1}^{(k)} = L_{k-1}(k,k-1)$ της k-γραμμής με τους αντίστοιχους $p_1^{(j)}, p_2^{(j)},\ldots, p_{k-1}^{(j)}$ μιας γραμμής $j>k$ (επενέργεια πολλαπλασιασμού αριστερά με P_k). Επίσης, γίνεται προσθήκη των νέων πολλαπλασιαστών p_{ik} ($k+1\leq i\leq n$) στην k-στήλη (επενέργεια πολλαπλασιασμού δεξιά με M_k^{-1}). Ας σημειωθεί ότι ο πολλαπλασιασμός δεξιά με M_k^{-1} δεν επιδρά στις στήλες $1\leq m\leq k-1$ του $L_{(k-1)}$, δηλ. αφήνει άθικτους τους ήδη εκεί αποθηκευμένους πολλαπλασιαστές (ενώ αυτό *δεν συμβαίνει* με τον πολλαπλασιασμό από αριστερά!). Τα L και U μπορούν ομοίως να αποθηκευθούν στο κάτω και άνω τριγωνικό τμήμα του A αντίστοιχα.

Το μεταθετικό μητρώο P μπορεί να αναπαρασταθεί και να υπολογισθεί μέσω ενός *μεταθετικού διανύσματος* $\boldsymbol{p}$. Το διάνυσμα αυτό χρησιμοποιείται από τον ακόλουθο αλγόριθμο που περιγράφει την $PA=LU$.

Αλγόριθμος 4.1.2: Παραγοντοποίηση PA=LU

Input: *A, n*
Output:
L: Κάτω τριγωνικό τμήμα του A,
U: Άνω τριγωνικό τμήμα του A,
***p**: μεταθετικό διάνυσμα εναλλαγών γραμμών.*

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

```
p:=1:n;              % αρχικοποίηση διανύσματος p (P=I).
for k=1:1:n-1
{   j:=k;              % j: index γραμμής οδηγού
  while A(j,k)==0      % Αναζήτηση οδηγού d_k.
    {j:=j+1 ;
     if j>n then exit('παραγοντοποίηση αδύνατη, A μη αντιστρεπτό');
}
  d=A(j,k);            % Υπολογισμός οδηγού d_k.
  if j~=k then
    {swap(A(k,k:n), A(j,k:n));  % Εναλλαγή γραμμών j και k.
  p(k):=j ; p(j):=k;     % Ενημέρωση index γραμμής
}
  if k>1 & j~=k then
    swap(A(k,1:k-1), A(j,1:k-1);   %Εναλλαγή προηγούμενων πολ/στών στις γραμμές j και k.
  for i=k+1:1:n
  { p:= A(i,k)/d;          %Υπολογισμός πολλαπλασιαστή p_ik για τη γραμμή i.
     A(k,i):=p;            % Αποθήκευση p_ik στη k-στήλη στο κάτω τριγ/κό τμήμα του A
     A(i,k:n):=A(i,k:n)-p*A(k,k:n)    % Απαλοιφή γραμμής k
}
}
if A(n,n)==0 then exit('παραγοντοποίηση αδύνατη');
end
```

Μετά τη φάση της παραγοντοποίησης, κάθε ομαλό σύστημα $A\mathbf{x}=\mathbf{b}$ μπορεί να τεθεί στη μορφή $PA\mathbf{x}=P\mathbf{b}$ ή $LU\mathbf{x}=P\mathbf{b}$ και να λυθεί σε δυο φάσεις: αρχικά να λυθεί το $L\mathbf{y}=P\mathbf{b}$ με εμπρός αντικατάσταση και κατόπιν το $U\mathbf{x}=\mathbf{y}$ με πίσω αντικατάσταση ($2n^2$ περίπου πράξεις).
Η εφαρμογή της *μερικής οδήγησης* είναι χρήσιμη στην πράξη και για τη μέθοδο LU, για τους λόγους που είδαμε στο Κεφ. ΙΙΙ (βλ. Παρ.4.1.3).

Ως εφαρμογές των διασπάσεων LU δίνουμε τα ακόλουθα παραδείγματα:

♦ Παράδειγμα 4.1.1 Παραγοντοποίηση A=LU χωρίς Οδήγηση

Στο Παρ.3.3.2 βρήκαμε ότι για το μητρώο

$$A=\begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & -1 \\ 4 & -2 & 6 \end{pmatrix}$$

κατά τη διαδικασία απαλοιφής χωρίς οδήγηση, δεν έγιναν εναλλαγές γραμμών και $p_{21}=2$, $p_{31}=4$, $p_{32}=-6$. Από την (4.1.3) έχουμε

$$L=M_1^{-1}M_2^{-1}=\begin{pmatrix} 1 & 0 & 0 \\ p_{21} & 1 & 0 \\ p_{31} & p_{32} & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & -6 & 1 \end{pmatrix},\ U=\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & -16 \end{pmatrix}$$

Εύκολα τώρα επαληθεύουμε ότι *LU=A*. □

♦ Παράδειγμα 4.1.2 Παραγοντοποίηση PA=LU χωρίς Οδήγηση

Θα υπολογίσουμε τα L,U και P για το μητρώο

$$A=\begin{pmatrix} 0 & 2 & 1 \\ 1 & -2 & 4 \\ 4 & -1 & 2 \end{pmatrix}$$

τέτοια ώστε $PA=LU$. Στο Παρ. 3.2.2 βρήκαμε

$$P_1 = P_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_2=P_{23}=I_3$$

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix} M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -7/2 & 1 \end{pmatrix}$$

Λαμβάνουμε $P=P_1P_2=P_1$, $L_1=M_1$, $L_2 = M_2(P_2L_1P_2)=M_2M_1$ και

$$L = M_1^{-1}M_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 7/2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 7/2 & 1 \end{pmatrix}$$

Επίσης στο Παρ. 3.2.1 βρήκαμε

$$U = \begin{pmatrix} 1 & -2 & 4 \\ 0 & 2 & 1 \\ 0 & 0 & -35/2 \end{pmatrix}$$

Και επαληθεύουμε εύκολα ότι $PA = LU$. □

♦ Παράδειγμα 4.1.3 [Matlab]: PA=LU με Μερική Οδήγηση

Για το μητρώο του προηγούμενου παραδείγματος, ορίζουμε πρώτα τον αντίστοιχο πίνακα A στο Matlab:

```
>> A=[0 2 1;1 -2 4;4 -1 2]
 A=
     0    2    1
     1   -2    4
     4   -1    2
```

Η διάσπαση LU εφαρμόζεται στο Matlab με μερική οδήγηση, επομένως θα γίνουν οι εξής εναλλαγές γραμμών: οι γραμμές 1 και 3 στο πρώτο βήμα και οι γραμμές 2 και 3 στο δεύτερο, δηλ. $P=P_{23}*P_{13}$. Οι πολλαπλασιαστές είναι τώρα: $l_{21}=0$, $l_{31}=1/4=0.25$, $l_{31}=-7/8=-0.8750$. Υπολογίζουμε τα *L,U,P* με την εντολή (ενσωματωμένη συνάρτηση) **lu**:

```
>>[L,U,P]=lu(A)
L=
    1.0000     0          0
    0          1.0000     0
    0.2500    -0.8750     1.0000

U=
    4.0000    -1.0000     2.0000
    0          2.0000     1.0000
    0          0          4.3750

P= % Αναμένουμε δύο μεταθέσεις γραμμών και ότι η τελική μετάθεση είναι P=P23*P13
   0    0         1
   1    0         0
   0    1         0
```

Επαληθεύουμε:

```
>>P*A==L*U      % Από τη σύγκριση προκύπτει λογικός πίνακας με 1 (=true) στις θέσεις με ίσα
                % στοιχεία, και 0 με διαφορετικά. Η απουσία μηδενικών εδώ δείχνει ότι PA=LU.
ans=
          1 1   1
          1 1   1
          1 1   1
```

Η εκτέλεση της lu για το μητρώο PA γίνεται τώρα χωρίς εναλλαγές γραμμών:

```
>> [L U P ]=lu(P*A)
```


Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

```
L =
   1.0000        0        0
        0   1.0000        0
   0.2500  -0.8750   1.0000

U =
   4.0000  -1.0000   2.0000
        0   2.0000   1.0000
        0        0   4.3750

P =
   1   0   0
   0   1   0
   0   0   1
```

□

♦ Παράδειγμα 4.1.4 [Matlab]: Επίλυση συστημάτων με διάσπαση PA=LU χωρίς Οδήγηση και με Μερική Οδήγηση

Για το σύστημα του Παρ.3.2.2 εφαρμόζουμε LU χωρίς οδήγηση, αποθηκεύοντας τα L και U στο κάτω και άνω τριγωνικό μέρος του A αντίστοιχα. Σε κάθε βήμα k=1,2 τοποθετούμε στη στήλη k του $[A\ b]$, αντί των μηδενικών, τους πολλαπλασιαστές που υπολογίσθηκαν σε αυτό. Στη δεύτερη φάση λύνεται το σύστημα με εμπρός και πίσω αντικατάσταση.

Παραγοντοποίηση:

$$[A\,|\,b]=\begin{bmatrix}1&1&-2\\3&5&-1\\3&4&1\end{bmatrix}\xrightarrow[d_1=1\neq 0]{P_{11}=I:}\begin{bmatrix}\boxed{1}&1&-2\\3&5&-1\\3&4&1\end{bmatrix}\xrightarrow[\substack{r_2=r_2-3r_1\\ r_3=r_3-3r_1}]{\mathbf{E}_1:}\begin{bmatrix}\boxed{1}&1&-2\\3&2&5\\3&1&7\end{bmatrix}$$

$$\xrightarrow[d_2=2\neq 0]{P_{22}=I:}\begin{bmatrix}\boxed{1}&1&-2\\3&\boxed{2}&5\\3&1&7\end{bmatrix}\xrightarrow[r_3=r_3-0.5r_2]{\mathbf{E}_2:}\begin{bmatrix}\boxed{1}&1&-2\\3&\boxed{2}&5\\3&0.5000&\boxed{4.5000}\end{bmatrix}$$

Επίλυση με εμπρός και πίσω αντικατάσταση:

$$Ly=b\Rightarrow\begin{bmatrix}1&0&0\\3&1&0\\3&0.5000&1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=\begin{bmatrix}9\\30\\21\end{bmatrix}\Rightarrow\begin{cases}y_1=9\\y_2=3\\y_3=-7\end{cases}$$

$$Ux=y\Rightarrow\begin{bmatrix}1&1&-2\\0&2&5\\0&0&4.5000\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}9\\3\\-7\end{bmatrix}\Rightarrow\begin{cases}x_3=\text{-}1.6667\\x_2=5.6667\\x_1=0.0000\end{cases}$$

Και εδώ οι υπολογισμοί μπορούν να γίνουν μέσα στο A και η λύση να αποθηκευθεί απ' ευθείας στο $\boldsymbol{b}$.

Με μερική οδήγηση:

Στο παραπάνω παράδειγμα, εφαρμόζοντας μερική οδήγηση στο A αλλάζουν οι οδηγοί και οι πολλαπλασιαστές, όπως και τα L, U και b. Το 2^{ο} βήμα εναλλάσσει τις γραμμές 2 και 3 και μαζί τους πολλαπλασιαστές του 1^{ου} βήματος στη στήλη 1.

$$[A\,|\,b]=\begin{bmatrix}1&1&-2\\3&5&-1\\3&4&1\end{bmatrix}\xrightarrow[\substack{d_1=3\\ r_1\leftrightarrow r_2}]{\mathbf{P}_{12}:}\begin{bmatrix}\boxed{3}&5&-1\\1&1&-2\\3&4&1\end{bmatrix}\xrightarrow[\substack{r_2=r_2-1/3r_1\\ r_3=r_3-1r_1}]{\mathbf{E}_{31}\mathbf{E}_{21}:}\begin{bmatrix}\boxed{3}&5&-1\\0.33333&-0.66667&\text{-}1.6667\\1&-1&2\end{bmatrix}\xrightarrow[\substack{d_2=-1\\ r_2\leftrightarrow r_3}]{\mathbf{P}_{32}:}$$

$$\begin{bmatrix} \boxed{3} & 5 & -1 \\ 1| & \boxed{-1} & 2 \\ \underline{0.33333|} & -0.66667 & -5/3 \end{bmatrix} \xrightarrow[r_3=r_3-2/3r_2]{\mathbf{E}_{32}:} \begin{bmatrix} \boxed{3} & 5 & -1 \\ 1| & \boxed{-1} & 2 \\ \underline{0.33333|} & \underline{-0.66667|} & -3 \end{bmatrix} \xrightarrow{d_3=-5/3}$$

$$\xrightarrow{d_3=-5/3} \begin{bmatrix} \boxed{3} & 5 & -1 \\ 1 & \boxed{-1} & 2 \\ \underline{0.33333|} & -0.66667 & \boxed{-3} \end{bmatrix}$$

Έτσι θα έχουμε:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0.33333 & 0 & 1 \end{bmatrix},\ U = \begin{bmatrix} 3 & 5 & -1 \\ 0 & -1 & 2 \\ 0 & 0 & -3 \end{bmatrix},\ P = P_{32}P_{12} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Το προς επίλυση σύστημα είναι τώρα το $PA\boldsymbol{x}=LU\boldsymbol{x}=P\boldsymbol{b}$. Είναι

$$Pb = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 9 \\ 30 \\ 21 \end{bmatrix} = \begin{bmatrix} 30 \\ 21 \\ 9 \end{bmatrix}$$

Λύνουμε: $Ly=P\boldsymbol{b}$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0.33333 & -0.66667 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 30 \\ 21 \\ 9 \end{bmatrix} \Rightarrow \begin{Bmatrix} y_1 = 30 \\ y_2 = -9 \\ y_3 = 5 \end{Bmatrix}$$

και $U\boldsymbol{x}=\boldsymbol{y}$:

$$\begin{bmatrix} \boxed{3} & 5 & -1 \\ 1 & \boxed{-1} & 2 \\ \underline{0.33333|} & -0.66667 & \boxed{-3} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 30 \\ -9 \\ 5 \end{bmatrix} \Rightarrow \begin{Bmatrix} x_1 = 0 \\ x_2 = 5.6667 \\ x_3 = -1.6667 \end{Bmatrix}$$

◊ **[Matlab]** Επαληθεύουμε στο *Matlab*:

```
>> [L U P]=lu(A)

L =

    1.0000         0         0
    1.0000    1.0000         0
    0.3333    0.6667    1.0000

U =
     3     5    -1
     0    -1     2
     0     0    -3

P =
     0     1     0
     0     0     1
     1     0     0

>> x=U\(L\P*b)

x =
    0.0000
    5.6667
   -1.6667
```

□

4.1.1 Διάσπαση LU – Μέθοδος Grout

Με αφετηρία τα αποτελέσματα της §4.1, η διάσπαση LU και γενικότερα η $PA=LU$ επιδιώκεται για οποιαδήποτε μητρώο, αντιστρεπτό η μη. Τότε τα L και U μπορούν να υπολογιστούν απ' ευθείας όπως περιγράφεται πιο κάτω (αλγόριθμος *Grout*):

Αν το A μπορεί να διασπαστεί σε μορφή LU, τότε το σύστημα $A=LU$ προφανώς έχει n^2 εξισώσεις με n^2+n αγνώστους (l_{ij} και u_{ij}), οπότε μπορούμε να θέσουμε αυθαίρετα $l_{ii}=1$ για $i=1,\ldots,n$. Αυτό προφανώς διευκολύνει τους υπολογισμούς και ειδικότερα τις διαιρέσεις. Ας σημειωθεί άλλωστε ότι η επιλογή αυτή συμφωνεί με την (4.1.3) και δεν απαιτεί διαιρέσεις κατά την πίσω αντικατάσταση που πιθανότατα θα προκαλέσουν σφάλματα στρογγύλευσης. Τώρα από την ταυτότητα $A=LU$ λαμβάνουμε τις εξής εξισώσεις:

$$a_{ij} = \boldsymbol{l}_i^T \boldsymbol{u}_j \text{ δηλαδή: } a_{ij} = \sum_{k=1}^{\min(i,j)} l_{ik}u_{kj}$$

και υπολογίζουμε αμέσως την πρώτη γραμμή του U και την πρώτη στήλη του L:

$$a_{1j} = 1 \times u_{1j} \Rightarrow u_{1j} = a_{1j}, \quad j=1,\ldots,n$$

$$a_{i1}=l_{i1}\, u_{11} \Rightarrow l_{i1} = \frac{a_{i1}}{u_{i1}}, \quad i=2,\ldots,n$$

Στη συνέχεια υπολογίζουμε σταδιακά τις υπόλοιπες γραμμές του U και στήλες του L: για $m=2,\ldots,n$ και $j=m,\ldots,n$ έχουμε:

$$a_{mj}= \sum_{k=1}^{m} l_{mk}u_{kj} = \sum_{k=1}^{m-1} l_{mk}u_{kj} + 1\times u_{mj}$$

$$u_{mj}= a_{mj} - \sum_{k=1}^{m-1} l_{mk}u_{kj}\ , j=m,\ldots,n$$

Επίσης

$$a_{im} = \sum_{k=1}^{m} l_{ik}u_{km} = \sum_{k=1}^{m-1} l_{ik}u_{km} + l_{im}u_{mm}\ , i=m+1,\ldots,n$$

απ' όπου, αν $u_{nm}=0$, προκύπτει ότι η διάσπαση είναι αδύνατη και το A μη αντιστρεπτό. Αν $u_{nm}\neq 0$:

$$l_{im} = \frac{1}{u_{mm}}\left[a_{mj} - \sum_{k=1}^{m-1} l_{mk}u_{kj}\right], \ i=m+1,\ldots,n$$

Ανάλογα διατυπώνεται και η διαδικασία υπολογισμού των P,L,U για την διάσπαση (4.1.4) (Άσκηση 4.3). Από τον τρόπο υπολογισμού των L και U είναι φανερό ότι: (i) για να είναι η διάσπαση δυνατή πρέπει $u_{ii}\neq 0$, $i=1,\ldots,n$, δηλ. πρέπει το A να είναι αντιστρεπτό, και (ii) η διάσπαση ως προς τον περιορισμό $l_{ii}=1$ ($i=1,\ldots,n$) – όπως θα συνέβαινε και με οποιοδήποτε άλλον – είναι *μοναδική* (δουλεύοντας χωρίς οδήγηση).

Η παραπάνω υπολογιστική μέθοδος (αλγόριθμος *Grout*) συνοψίζεται στον παρακάτω αλγόριθμο:

Αλγόριθμος 4.1.3: Παραγοντοποίηση LU (Μέθοδος Grout)

Input: *Μητρώο* $A(n\times n)$, n: *μέγεθος.*
Output: *κάτω τριγωνικό μητρώο* L $(n\times n)$,
άνω τριγωνικό μητρώο U $(n\times n)$ *τέτοιο ώστε* $LU=A$.

```
if a11 =0 then { write("Διάσπαση αδύνατη");exit } ;
u11:= a11;
l11:=1;
for j:=2:n          % υπολογισμός 1ης γραμμής του U και 1ης στήλης του U
   {u1j:= a1j;
   l1j:= a1j / u11 };
for i:=2:n-1 do
  { lii:=1;
```

$u_{ii}=a_{ii}$ - $\sum_{k=1}^{i-1} l_{ik}u_{ki}$;

if u_{ii} =0 **then** {write('Διάσπαση αδύνατη'); **exit**};

for j:=i:n

{ $u_{ij}=a_{ij}$ - $\sum_{k=1}^{i-1} l_{ik}u_{kj}$; % υπολογισμός της i γραμμής του U

$l_{ji} = \frac{1}{u_{ii}}\left(a_{ji} - \sum_{k=1}^{i-1} l_{jk}u_{ki}\right)$ % υπολογισμός της i στήλης του U

}

};

$u_{nn} = a_{nn}$ - $\sum_{k=1}^{n-1} l_{nk}u_{kn}$ $\sum_{k=1}^{n-1} l_{nk}u_{kn}$; % υπολογισμός τελευταίου στοιχείου του U

write(L,U);

end

Θα αναζητήσουμε τώρα μητρώα διασπάσιμα στη μορφή LU. Μπορούμε να δείξουμε ότι τέτοια διάσπαση επιδέχονται δυο ενδιαφέρουσες κατηγορίες μητρώων: τα μητρώα με *αυστηρή διαγώνια κυριαρχία* και τα *θετικά ορισμένα μητρώα*. Τις κατηγορίες αυτές αναλύουμε στις επόμενες ενότητες.

♦ Παράδειγμα 4.1.5

Δίνεται το μητρώο A και το διάνυσμα $\boldsymbol{b}$:

A=[1 0.2 -0.1 ; 1 -3 0.2 ; 0.1 0 2] , $\boldsymbol{b}$=[2 0 2]^T

Υποθέτουμε ότι εργαζόμαστε με αριθμητική κ. υ. 5 σ.ψ. Θα δώσουμε τη μορφή διάσπασης LU του A, χωρίς να εφαρμόσουμε τη μέθοδο *Gauss*.

Μπορούμε, όπως θα δούμε στην §4.4.2, να αποφανθούμε εκ των προτέρων για τη διασπασιμότητα του A. Το μητρώο αυτό έχει *αυστηρή διαγώνια κυριαρχία* (βλ. §4.3) και ως εκ τούτου μπορεί να διασπασθεί άμεσα: $A=LU$ (δηλ. δεν απαιτούνται ανταλλαγές γραμμών ή ισοδύναμα $P=I$). Εφαρμόζουμε κατ' ευθείαν τον αλγόριθμο *Grout*. Γράφουμε την ταυτότητα $A=LU$ απαιτώντας L=κάτω τριγωνικό με $L(i,i)$=1 και U=άνω τριγωνικό:

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

Προκύπτει ένα γραμμικό σύστημα με 9 αγνώστους (l_{ij} και u_{ij}). Εξισώνοντας συστηματικά τα $A(i,j)$ με τα στοιχεία του A κατά στήλες, λαμβάνουμε (συμβολίζουμε με $L_{(i)}$ τη γραμμή i του L και με $U^{(i)}$ τη στήλη i του U):

Στοιχεία 1ης στήλης:

$A(1,1)=1 = L_{(1)}U^{(1)} = u(1,1)$

$A(2,1)=1= L_{(2)}U^{(1)} = l(2,1)*u(1,1) \Rightarrow$ **l(2,1)** = 1

$A(3,1)=0.1= L_{(3)}U^{(1)} = l(3,1)*u(1,1) \Rightarrow$ **l(3,1)** = 0.1

Στοιχεία 2ης στήλης:

$A(1,2)=0.2= L_{(1)}U^{(2)} =$ **u(1,2)**

$A(2,2)=-3= L_{(2)}U^{(2)} = l(2,1)\, u(1,2)+u(2,2) \Rightarrow$ **u(2,2)** = -3 – 1*0.2 = -3.2

$A(3,2)=0= L_{(3)}U^{(2)} = l(3,1)\, u(1,2)+l(3,2)u(2,2) \Rightarrow$ **l(3,2)** = -0.1*0.2/(-3.2)=

$=fl(0.00625)_{n=4}=0.0063$

Στοιχεία 3ης στήλης:

$A(1,3)=-0.1= L_{(1)}U^{(3)} =$ **u(1,3)**

$A(2,3)=0.2= L_{(2)}U^{(3)} = l(2,1)\, u(1,3)+u(2,3) \Rightarrow$ **u(2,3)** = 0.2 – 1*(-0.1) = 0.3

$A(3,3)=1= L_{(3)}U^{(3)} = l(3,1)\, u(1,3)+l(3,2)u(2,3)+u(3,3) \Rightarrow$
$\Rightarrow \boldsymbol{u}(3,3) = 1 - (0.1)*(-0.1)- 0.00625*(0.3) = fl(1.008125)_{n=4} = 2.0081$

Τα L και U έχουν τώρα καθορισθεί πλήρως:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0.1 & 0.0063 & 1 \end{bmatrix},\ U = \begin{bmatrix} 1 & 0.2 & -0.1 \\ 0 & -3.2 & 0.3 \\ 0 & 0 & 2.0081 \end{bmatrix}$$

Αν χρησιμοποιούσαμε μερική οδήγηση δεν θα χρειαζόταν και πάλι εναλλαγή γραμμών. Συνεπώς θα ήταν $P=I$ και στην περίπτωση αυτή.

Για την επίλυση του $A\boldsymbol{x}=\boldsymbol{b}$, λύνουμε πρώτα το $L\boldsymbol{y} =\boldsymbol{b}$ με εμπρός αντικατάσταση:

$\boldsymbol{y} = (2.0000, -2.0000, 1.8125)^T$

και μετά το $U\boldsymbol{x} =\boldsymbol{y}$ με πίσω αντικατάσταση:

$\boldsymbol{x} = (1.9483, 0.7096, 0.9026)^T$

◇ **[Matlab]** Επαληθεύουμε στο *Matlab*:

```
A=[1 0.2 -0.1 ; 1 -3 0.2 ; 0.1 0 2]; b=[2 0 2]';

>> [L U P] = lu(A)

L =
    1.0000         0         0
    1.0000    1.0000         0
    0.1000    0.0063    1.0000

U =
    1.0000    0.2000   -0.1000
         0   -3.2000    0.3000
         0         0    2.0081

P =
     1     0     0
     0     1     0
     0     0     1

x=U\(L\b)

x =
    1.9483
    0.7096
    0.9026
```

□

4.2 Θετικά Ορισμένα Μητρώα

Τα *θετικά ορισμένα και ημιορισμένα μητρώα* (*positive-definite matrices*) έχουν μερικές καλές ιδιότητες, οι οποίες, μεταξύ άλλων, χρησιμοποιούνται για την εύρεση αποδοτικών από πλευράς μνήμης και υπολογιστικού χρόνου διασπάσεων του μητρώου ενός γραμμικού συστήματος. Συγχρόνως βρίσκουν και άλλες εφαρμογές, όπως στις τετραγωνικές μορφές, κωνικές τομές, τοπικά ακρότατα συναρτήσεων κλπ. Ο γενικός ορισμός των μητρώων αυτών είναι:

Ορισμός 4.2.1

i) Ένα μητρώο A nxn πραγματικών αριθμών καλείται ***θετικά ορισμένο*** *(θ.ο.), όταν για κάθε διάνυσμα* $\mathbf{x}\neq 0$ *ισχύει* $\mathbf{x}^{\mathrm{T}}A\mathbf{x}>0$.

ii) Ένα μητρώο A (nxn) πραγματικών καλείται ***θετικά ημιορισμένο*** *(θ.η.ο.), όταν για κάθε διάνυσμα* $\mathbf{x}\neq 0$ *ισχύει* $\mathbf{x}^{\mathrm{T}}A\mathbf{x}\geq 0$.

Γεωμετρική ερμηνεία: Θεωρώντας τον Ευκλείδειο χώρο R^n, ο Ο4.2.1. εκφράζει ότι το εσωτερικό γινόμενο των διανυσμάτων $\boldsymbol{v}^{\mathrm{T}}$ και $A\boldsymbol{v}$, είναι μονίμως θετικό (μη αρνητικό):

$$\boldsymbol{v}^{\mathrm{T}}A\boldsymbol{v} = \langle \boldsymbol{v}, A\boldsymbol{v}\rangle = \|\boldsymbol{v}\|\,\|A\boldsymbol{v}\|\cos(\theta) > 0 \Rightarrow \cos(\theta) > 0 \Rightarrow |\theta| < \pi/2$$

Δηλαδή η γωνία θ των διανυσμάτων $\boldsymbol{v}$ και $A\boldsymbol{v}$ βρίσκεται στο διάστημα $(-\pi/2, \pi/2)$. (Για τα θ.η.ο. μητρώα θα είναι $\theta\in[-\pi/2, \pi/2]$.)

Ταυτόχρονα, ο Ο4.2.1 προϋποθέτει τη θεώρηση των *τετραγωνικών μορφών*. Δοθέντος ενός μητρώου $A\in M_n(R)$, το γινόμενο $\mathbf{x}^{\mathrm{T}}A\mathbf{x}$ ορίζει μια πολυωνυμική συνάρτηση 2ου βαθμού και n μεταβλητών $q:R^n\rightarrow R$, $q(\mathbf{x})=\mathbf{x}^{\mathrm{T}}A\mathbf{x}$. Η q δίνεται από τη σχέση:

$$q(\mathbf{x})=\mathbf{x}^{\mathrm{T}}A\mathbf{x} = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i<j}^{n} c_{ij}x_i x_j \tag{4.2.1}$$

και καλείται ***τετραγωνική μορφή*** *(quadratic form)*. Κάθε όρος της q είναι 2ου βαθμού. Για παράδειγμα, η τετραγωνική μορφή τάξης 2 που παράγεται από ένα συμμετρικό μητρώο $A\in M_2(\mathrm{R})$ είναι:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = ax_1^2 + 2x_1x_2 + ax_2^2 \tag{4.2.2}$$

Η τετραγωνική μορφή τάξης n που ορίζεται από ένα συμμετρικό μητρώο $A\in M_n(\mathrm{R})$ δίνεται:

$$\mathbf{q(x)} = \mathbf{x}^{\mathbf{T}}\mathbf{Ax} = \sum_{i=1}^{n} a_{ij}x_i^2 + 2\sum_{i<j}^{n} a_{ij}x_i x_j \tag{4.2.3}$$

Παρατήρηση 4.2.1 Για n=2, λαμβάνουμε: A=[a b; c d]. Τότε η αντίστοιχη τ.μ. q θα είναι:

$$\mathbf{x}^{\mathbf{T}}\mathbf{Ax} = (x_1, x_2)\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = ax_1^2 + dx_2^2 + (c+b)x_1x_2 > 0.$$

Αν το A είναι συμμετρικό (b=c), παίρνουμε:

$$\mathbf{x}^{\mathbf{T}}\mathbf{Ax} = (x_1, x_2)\begin{bmatrix} a & b \\ b & d \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = ax_1^2 + dx_2^2 + 2bx_1x_2 > 0.$$

και διαμορφώνοντας κατάλληλα τους όρους έτσι ώστε να εμφανισθούν τετράγωνα («συμπλήρωση τετραγώνου»):

$$x^T Ax = ax_1^2 + dx_2^2 + 2bx_1x_2 = a\left(x_1 + \frac{b}{a}x_2\right)^2 + \left(\frac{ad-b^2}{a}\right)x_2^2 > 0. \quad (4.2.4)$$

Παρατηρούμε ότι οι συντελεστές των τετραγώνων στην (4.2.4) είναι ακριβώς οι οδηγοί του A:

$$a = u_1, \quad \frac{ad-b^2}{a} = u_2$$

Συνεπώς, οι οδηγοί ενός θ.ο. (ή θ.η.ο.) συμμετρικού μητρώου είναι οι συντελεστές της τετραγωνικής μορφής που παράγεται από αυτό, εκφρασμένης σε άθροισμα τετραγώνων. Το αποτέλεσμα αυτό γενικεύεται για nxn πραγματικά μητρώα. Αν μάλιστα οι οδηγοί είναι θετικοί, τότε αυτό εξασφαλίζει τη θετικότητα της τετραγωνικής μορφής.

Αν τώρα δίνεται μια τετραγωνική μορφή q μπορούμε πάντα από το δεξιό σκέλος της (4.2.1) να προσδιορίζουμε το μητρώο A που την παράγει. Το μητρώο αυτό θα είναι πάντα *συμμετρικό*, όπως υπαγορεύει το εξής λήμμα:

Λήμμα 4.2.1: *Η τετραγωνική μορφή $q(\mathbf{x})=\mathbf{x}^T A\mathbf{x}$ ταυτίζεται με την $p(\mathbf{x})=\mathbf{x}^T C\mathbf{x}$, όπου C συμμετρικό μητρώο που δίνεται από το συμμετρικό τμήμα του A, $C = sym(A) = (A + A^T)/2$.*

Απόδειξη: Κάθε μητρώο A εκφράζεται ως άθροισμα ενός συμμετρικού και ενός αντισυμμετρικού μητρώου:

$$A = C + B, \text{ με } C=(A+A^T)/2 \text{ και } B=(A-A^T)/2 \quad (4.2.4)$$

οπότε:

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T(C+B)\mathbf{x} = \mathbf{x}^T C\mathbf{x} + \mathbf{x}^T B\mathbf{x} = \mathbf{x}^T C\mathbf{x} \quad (4.2.5)$$

αφού από την $\mathbf{x}^T B\mathbf{x} = (\mathbf{x}^T B\mathbf{x})^T = \mathbf{x}^T B^T\mathbf{x} = \mathbf{x}^T(-B)\mathbf{x} = -\mathbf{x}^T B\mathbf{x}$ προκύπτει $\mathbf{x}^T B\mathbf{x} = \mathbf{0}$. ■

Παρατηρούμε επίσης ότι τα A και C συνδέονται με τη σχέση:

$$diag(A) = diag(C)$$

Το Λ4.2.1 και η (4.2.5) οδηγούν στα εξής συμπεράσματα:

- Πρώτα στη διατύπωση ενός *γενικού και μοναδικού κριτηρίου* για θετικά ορισμένα (μη συμμετρικά) μητρώα:

Γενικό Κριτήριο για θετικά ορισμένα (μη συμμετρικά) μητρώα

Ένα οποιοδήποτε τετραγωνικό μητρώο A είναι θ.ο. (θ.η.ο.) εάν και μόνον εάν το συμμετρικό του μέρος $(A+A^T)/2$ είναι θ.ο. (θ.η.ο.).

- Για κάθε μητρώο A υπάρχει ένα μοναδικό συμμετρικό μητρώο C, που θα λέγεται *αντίστοιχο μητρώο της τ.μ.*, το οποίο ορίζει την ίδια τετραγωνική μορφή $q(\mathbf{x})$. Μια ισοδύναμη διατύπωση: *κάθε τετραγωνική μορφή $q(\mathbf{x})=\mathbf{x}^T A\mathbf{x}$ παράγεται από άπειρα μητρώα $A\in M_n(R)$, αλλά από ένα μοναδικό συμμετρικό μητρώο C:*

$$\mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x} = \mathbf{x}^{\mathbf{T}}\mathbf{C}\mathbf{x} = q(x) = \sum_{i=1}^{n} c_{ij}x_i^2 + 2\sum_{i<j}^{n} c_{ij}x_i x_j \quad \text{όπου } C = sym(A) = (A + A^T)/2 \quad (4.2.6)$$

  Η (4.2.6) προφανώς συμπίπτει με την (4.2.1) όταν A είναι συμμετρικό. Όταν λοιπόν μια τετραγωνική μορφή δίνεται στην τελική της μορφή από την αλγεβρική της έκφραση

$$q(x) = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i<j}^{n} a_{ij}x_i x_j$$

εύκολα βρίσκουμε το συμμετρικό A που την παράγει: $A_{ii}=a_{ii}$ (διαγώνιος) και $A_{ij}=a_{ij}$ (κάτω τριγωνικό τμήμα), για $i<j$. Συνεπώς σε κάθε τετραγωνική μορφή $q(\boldsymbol{x})$ αντιστοιχεί ένα συμμετρικό μητρώο A, έτσι ώστε:

$$q(x_1,\ldots,x_n)=\mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x}=\sum_{i=1}^{n}a_{ij}x_i^2+2\sum_{i<j}^{n}a_{ij}x_ix_j \tag{4.2.7}$$

- Αν E, F, $G\in M_n(\mathrm{R})$, τότε η δυαδική σχέση ~: «$q_A=q_B$» ή «*τα A και B ορίζουν ταυτόσημες τετραγωνικές μορφές*» είναι σχέση ισοδυναμίας επί του χώρου $M_n(\mathrm{R})$ (αυτοπαθής: $A\sim A$, συμμετρική: $A\sim B\Rightarrow B\sim A$ και μεταβατική: $(A\sim B)\wedge(B\sim C)\Rightarrow A\sim C$). Ως εκ τούτου, το $M_n(\mathrm{R})$ διαμερίζεται σε κλάσεις ισοδυναμίας. Κάθε κλάση εκπροσωπείται από ένα συμμετρικό μητρώο C. Κάθε της στοιχείο X έχει $sym(X)=(X+X^{\mathrm{T}})/2=C$ και κατά συνέπεια $diag(X) = diag(C)$.

Τα παραπάνω μας επιτρέπουν να θεωρούμε μόνο συμμετρικά και Ερμιτιανά μητρώα[1] για τη μελέτη των τετραγωνικών μορφών. Επιπλέον, η θεώρηση των μητρώων αυτών προσφέρει απλότητα και ευελιξία στη διατύπωση κριτηρίων και ευκολία στους υπολογισμούς. Για τη μελέτη λοιπόν της συμπεριφοράς της $q(x)$, όπως π.χ. η εξέταση τοπικών ακρότατων, μεταβαίνουμε κατ' ευθείαν στην εξέταση των ιδιοτήτων του αντίστοιχου συμμετρικού A. Σε αρκετές μάλιστα περιπτώσεις το A υπαγορεύεται άμεσα από τη φύση του προβλήματος.

Σύμφωνα με τον Ο4.2.1, για $\boldsymbol{x}\neq\mathbf{0}$, η q είναι μονίμως θετική (μη αρνητική) όταν και μόνον όταν το A είναι θ.ο. (θ.η.ο.). Τότε η q λέγεται *θετικά ορισμένη*. Κατά συνέπεια, η q *ελαχιστοποιείται* στο $\boldsymbol{x}=\mathbf{0}$, εάν και μόνον εάν το A είναι θ.ο. Στην περίπτωση αυτή η q μπορεί να εκφρασθεί ως άθροισμα n τετραγώνων. Γενικά, ενδιαφερόμαστε για τη διατύπωση κριτηρίων για το πότε ένα μητρώο είναι θ.ο., ή θ.η.ο., ή *αρνητικά ορισμένο* ($\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}<0$), ή *αρνητικά ημιορισμένο* ($\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}\leq 0$), ή *αόριστο* (άγνωστο πρόσημο της q). Θα εξετάσουμε εδώ ιδιότητες και κριτήρια με έμφαση τα θ.ο. μητρώα.

♦ Παράδειγμα 4.2.1

Για τα μητρώα: A=[1 3 ; 3 10], B=[1 3 ; 3 9] (συμμετρικά) και C=[1 1; -1 1] (μη συμμετρικό), εξετάζουμε αν είναι θ.ο. και διερευνούμε μεγέθη και πιθανές ιδιότητες που σχετίζονται με τον παραπάνω ορισμό. Παρατηρούμε λοιπόν ότι για κάθε διάνυσμα $\boldsymbol{x}=(\boldsymbol{x}_1, \boldsymbol{x}_2)^{\mathrm{T}}\neq\mathbf{0}$ ισχύει:

$$x^TAx=(x_1,x_2)\begin{bmatrix}1&3\\3&10\end{bmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=(x_1+3x_2,3x_1+10x_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}=x_1^2+6x_1x_2+10x_2^2=$$

$$=x_1^2+2\times 3x_1x_2+(3x_2)^2+x_2^2=(x_1+3x_2)^2+x_2^2>0$$

$$x^TBx=(x_1,x_2)\begin{bmatrix}1&3\\3&9\end{bmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=(x_1+3x_2,3x_1+9x_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}=x_1^2+6x_1x_2+9x_2^2=(x_1+3x_2)^2\geq 0$$

$$x^TCx=(x_1,x_2)\begin{bmatrix}1&1\\-1&1\end{bmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=(x_1-x_2,x_1+x_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}=x_1^2+x_2^2>0$$

Επομένως τα A και C είναι θ.ο. Για το μη συμμετρικό C είναι $sym(C)$ = diag([1,1]) που είναι θ.ο. Το B είναι θ.η.ο.: το $\mathbf{x}^{\mathbf{T}}C\mathbf{x}$ μηδενίζεται για όλα τα $\boldsymbol{x}=(-3\alpha,\alpha)^{\mathrm{T}}$ δηλαδή κατά μήκος της ευθείας $x_1+3x_2=0$. Επίσης παρατηρούμε:

- Τα *κύρια υπομητρώα* ([2]) του A έχουν θετικές ορίζουσες det(A(1:2,1:2))=1, det(A(1:1,1:1))=1, ενώ του B μη αρνητικές: det(B(1:2,1:2))=0, det(B(1:1,1:1))=1. Αυτό *ισχύει πάντα* για τα συμμετρικά θ.ο. μητρώα. Για το C έχουμε βέβαια θετικές ορίζουσες: det(C(1:2,1:2))=2,

[1] Όλα τα συμμετρικά μητρώα $A\in M_n(R)$ αποτελούν ένα υπόχωρο $S_n(R)\subset M_n(R)$, με $dim(S_n(R))=n(n+1)$. Το αντίστοιχο συμβαίνει και για τα Ερμιτιανά μητρώα $A\in M_n(C)$.

[2] *Κύρια υπομητρώα* (*principal minors*) A_k ενός τετραγωνικού μητρώου A λέγονται όλα τα (φωλιασμένα) υπομητρώα $k\times k$, $k=1,\ldots,n$, τα οποία ξεκινούν από το άνω αριστερό άκρο του. Στη βιβλιογραφία είθισται ο συμβολισμός $\underline{A}_k$. Σε πιο σύγχρονα κείμενα σημειώνονται ως μπλοκ εκτεινόμενα σε διαστήματα στηλών και γραμμών: $A_{1:k}$, ή $A(1{:}n,1{:}n)$. Συχνά επίσης συναντάται ο όρος *κύρια υπο-ορίζουσα* ή *αριστερή ορίζουσα* που σημειώνεται με Δ_k ή με $|A_k|$.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

det(C(1:1,1:1))=1, αλλά αυτό *δεν* αποτελεί τον γενικό κανόνα. Εδώ το κριτήριο αυτό ισχύει για το $sym(C)$ που είναι θ.ο.

- Τα θ.ο. μητρώα A και C είναι αντιστρέψιμα, ενώ το θ.η.ο. B μη αντιστρέψιμο. Είναι φανερό όμως ότι η αντιστρεψιμότητα δεν αποτελεί ικανό κριτήριο!
- Οι οδηγοί για τα συμμετρικά μητρώα είναι: 1, 1 για το A, και 1, 0 για το B. Οι θετικοί (μη αρνητικοί) οδηγοί αποτελούν κριτήριο για συμμετρικά θ.ο. (θ.η.ο.) μητρώα. Αυτό *δεν* συμβαίνει γενικά για τα μη συμμετρικά μητρώα. Το ό,τι το C έχει εδώ θετικούς οδηγούς (1, 2), οφείλεται στο ό,τι το θ.ο. $sym(C)$ έχει θετικούς οδηγούς (1, 1).
- Οι ιδιοτιμές του συμμετρικού θ.ο. A είναι θετικές: λ_1=0.0917 και λ_2=10.9083. Του επίσης συμμετρικού θ.η.ο. B είναι θετικές ή 0: λ_1=0 (αναμενόταν, αφού το C είναι ιδιάζον) και λ_2=10. Αυτό *συμβαίνει πάντοτε* για τα συμμετρικά θ.ο. μητρώα και είναι ένα *ικανό και αναγκαίο κριτήριο*. Τέλος, οι ιδιοτιμές του C είναι μιγαδικές: λ_1=1+i και λ_2=1-i. Αν πάντως υπήρχαν πραγματικές, θα ήταν θετικές. Αυτό παρέχει μια *αναγκαία συνθήκη*.
- Έστω η *τετραγωνική μορφή* $\boldsymbol{q}$: $R^2 \longrightarrow R$, $q(x_1, x_2)=\mathbf{x}^{\mathbf{T}}F\mathbf{x}$, με $F\in M_2(R)$, όπου στη θέση του F θεωρούμε τα πιο πάνω μητρώα. Τότε, η $\boldsymbol{q}$ είναι μονίμως θετική για $\boldsymbol{x}\neq\mathbf{0}$, εάν το F είναι θ.ο. (περίπτωση A και C). Επίσης μονίμως μη αρνητική, αν F είναι θ.η.ο. (περίπτωση B). Σε όλες τις περιπτώσεις η $\boldsymbol{q}$ εκφράζεται στην *κανονική της μορφή* ως *άθροισμα τετραγώνων*, όπως φαίνεται πιο πάνω. Η $\boldsymbol{q}$ παρουσιάζει ελάχιστο στο $\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}=\mathbf{0}$ αν F είναι θ.ο. ή θ.η.ο.
- Αν αντί των παραπάνω μητρώων δινόταν μόνον η συνάρτηση 2ου βαθμού (τετραγωνική μορφή) $\boldsymbol{q}(x_1, x_2)=ax_1^2+2bx_1x_2+cx_2^2$, τότε εύκολα θα βρίσκαμε και για τις τρεις παραπάνω περιπτώσεις τα αντίστοιχα συμμετρικά μητρώα:

$$q(x_1,x_2)=x_1^2+6x_1x_2+10x_2^2=x_1^2+2\times 3x_1x_2+10x_2^2 \xrightarrow{C(1,1)=1,C(2,2)=10,C(1,2)=3} A=\begin{bmatrix}1 & 3\\ 3 & 10\end{bmatrix}$$

$$q(x_1,x_2)=x_1^2+6x_1x_2+9x_2^2=x_1^2+2\times 3x_1x_2+9x_2^2 \xrightarrow{C(1,1)=1,C(2,2)=9,C(1,2)=3} B=\begin{bmatrix}1 & 3\\ 3 & 9\end{bmatrix}$$

$$q(x_1,x_2)=x_1^2+x_2^2 \xrightarrow{C(1,1)=1,C(2,2)=1} C=\mathrm{diag}([1,1])=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=I$$

- Σε όλες τις περιπτώσεις του παραδείγματος, η μετάβαση στη *κανονική μορφή* ήταν εύκολη. Αρκούσε μόνον μια μικρή διαμόρφωση στους όρους για να συμπληρωθεί ένα τετράγωνο. Τι θα συνέβαινε όμως όταν $n>2$;

4.2.1 Κανονικές Τετραγωνικές Μορφές

Μια τετραγωνική μορφή ελαχιστοποιείται στο $\boldsymbol{x}=\mathbf{0}$, αν και μόνον αν το A είναι θ.ο. Στην περίπτωση αυτή μπορεί να εκφρασθεί σαν άθροισμα n τετραγώνων και είναι θετική για $\boldsymbol{x}\neq\mathbf{0}$. Η καλύτερη μορφή από την άποψη αυτή είναι η διαγώνια μορφή.

Μια τετραγωνική μορφή q λέγεται **διαγώνια** όταν εκφράζεται στη μορφή *αθροίσματος τετραγώνων*, δηλαδή για $\boldsymbol{x}=(x_1, x_2,\ldots x_n)^{\mathrm{T}}\in\mathrm{R}^n$:

$$q(\mathbf{x})=\mathbf{x}^{\mathbf{T}}C\mathbf{x}=\sum_{i=1}^{n}c_i x_i^2 \tag{4.2.8}$$

Το αντίστοιχο μητρώο C της τ.μ. q είναι προφανώς διαγώνιο, δηλ. $C=diag([c_1, c_2,\ldots c_n])$. Π.χ. για το διαγώνιο μητρώο $D=diag([2, 5])$ υπολογίζεται η διαγώνια τετραγωνική μορφή $q(x)=2x_1^2+5x_2^2$. Η διαγώνια μορφή είναι για προφανείς λόγους η πιο πρόσφορη για τη μελέτη της q, οπότε αναζητούμε αποδοτικούς τρόπους μετάβασης σε αυτή. Αυτό μπορεί να γίνει μέσω ενός μετασχηματισμού αλλαγής βάσης $\boldsymbol{y}=M\boldsymbol{x}$.

Παρατήρηση 4.2.1 Μια πρώτη διαπίστωση είναι:
Ένα διαγώνιο μητρώο D και η αντίστοιχη τ.μ. είναι θ.ο. εάν και μόνον εάν $d_i>0$, i=1,2,…n.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

Ας θεωρήσουμε τώρα μια τ.μ. $q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$, όπου A συμμετρικό. Για να τη φέρουμε στη μορφή (4.2.8), χρησιμοποιούμε τον ορθογώνιο μετασχηματισμό που δίνει η *διαγωνοποίηση του A*. Για την μετατροπή στη διαγώνια μορφή, θεωρούμε το αποτέλεσμα του φασματικού θεωρήματος για συμμετρικά μητρώα:

$$A = QDQ^{\mathrm{T}} \tag{4.2.9}$$

όπου Q η ορθοκανονική βάση ιδιοδιανυσμάτων του A και D διαγώνιο μητρώο με τις n ιδιοτιμές του A. Εκτελούμε τον εξής μετασχηματισμό αλλαγής βάσης:

$$\boldsymbol{y} = Q^{\mathrm{T}}\boldsymbol{x} \tag{4.2.10}$$

οπότε το $\boldsymbol{x}$ θα δίνεται ως προς τη νέα ορθοκανονική βάση: $\boldsymbol{x}=Q\boldsymbol{y}$ (αλλαγή μεταβλητής). Από τις (4.2.8), (4.2.9) και (4.2.10) λαμβάνουμε:

$$\mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x} = \mathbf{x}^{\mathbf{T}}\mathbf{Q}\mathbf{D}\mathbf{Q}^{\mathbf{T}}\mathbf{x} = (\mathbf{Q}^{\mathbf{T}}\mathbf{x})^{\mathbf{T}}\mathbf{D}(\mathbf{Q}^{\mathbf{T}}\mathbf{x}) = \mathbf{y}^{\mathbf{T}}\mathbf{D}\mathbf{y} \tag{4.2.11}$$

απ' όπου εξάγουμε την τ.μ. q ως προς τη νέα βάση:

$$q(y_1,\ldots,y_n) = \mathbf{y}^{\mathbf{T}}D\mathbf{y} = \sum_{i=1}^{n}\lambda_i y_i^2 \tag{4.2.12}$$

Η μορφή (4.2.12) θα λέγεται **κανονική τετραγωνική μορφή**. Καταλήγουμε λοιπόν στη διατύπωση του παρακάτω θεωρήματος.

Θεώρημα 4.2.1:

Κάθε πραγματική τετραγωνική μορφή μπορεί να αναχθεί σε μια διαγώνια τετραγωνική μορφή με τη βοήθεια ενός ορθογώνιου μετασχηματισμού.

Τώρα, για να μελετήσουμε τη κανονική μορφή (4.2.12) ως προς το αν είναι θ.ο., θ.η.ο. κ.λ.π., αρκεί να εξετάσουμε τα πρόσημα των ιδιοτιμών του A.

- Η q είναι *θετικά ορισμένη* αν όλες οι ιδιοτιμές είναι θετικές. Στο $\boldsymbol{y}=0$ η q εμφανίζει ελάχιστο.
- Η q είναι *θετικά ημιορισμένη* αν όλες οι ιδιοτιμές είναι μη αρνητικές. θετικές. Στο $\boldsymbol{y}=0$ η q εμφανίζει τοπικό ελάχιστο.
- Η q είναι *αρνητικά ορισμένη* αν όλες οι ιδιοτιμές είναι αρνητικές. Στο $\boldsymbol{y}=0$ η q εμφανίζει μέγιστο.
- Η q είναι *αρνητικά ημιορισμένη* αν όλες οι ιδιοτιμές είναι μη θετικές. Στο $\boldsymbol{y}=0$ η q εμφανίζει τοπικό μέγιστο.
- Η q είναι *αόριστη* αν υπάρχουν θετικές και αρνητικές ιδιοτιμές. Δεν μπορούμε να αποφανθούμε για τα ακρότατα της q.

Για τη μελέτη των τετραγωνικών μορφών, αλλά και για άλλες εφαρμογές, είναι σημαντικό να γνωρίζουμε πότε ένα συμμετρικό ή Ερμιτιανό μητρώο είναι θ.ο. ή θ.η.ο. Για τυχαία μητρώα είδαμε ότι υπάρχει μόνον αυτό που παραπέμπει στο συμμετρικό του συμπλήρωμα. Προτάσσουμε στο σημείο αυτό μερικές ενδιαφέρουσες ιδιότητες των θ.ο. μητρώων, οι οποίες θα οδηγήσουν σε χρήσιμα συμπεράσματα αναφορικά με τις παραγοντοποιήσεις μητρώων.

Πρόταση 4.2.1

Αν ένα μητρώο είναι θ.ο. (θ.η.ο.), τότε είναι αντιστρέψιμο (ιδιάζον).

Απόδειξη: Έστω $A\in\mathbb{R}^{n\times n}$ θ.ο. και $\boldsymbol{x}\in\mathbb{R}^n$. Τότε αν $A\boldsymbol{x}=\mathbf{0} \Rightarrow \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}\mathbf{0}=\mathbf{0} \Rightarrow \boldsymbol{x}=\mathbf{0}$ και άρα A αντιστρέψιμο. Αντίστοιχα, αν A θ.η.ο., είναι $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=\mathbf{0}$ για κάποιο $\boldsymbol{x}\neq\mathbf{0}$, δηλ. το A είναι ιδιάζον.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

Πρόταση 4.2.2

Αν ένα μητρώο είναι θ.ο. (θ.η.ο.), τότε τα διαγώνια στοιχεία του είναι θετικά (μη αρνητικά). [*Το αντίστροφο δεν ισχύει!*].

Απόδειξη: Αν $A\in R^{n\times n}$ θ.ο., και λάβουμε $\boldsymbol{x}=\boldsymbol{e}_i$, $1\leq i\leq n$, θα είναι $\boldsymbol{e}_i^T A\boldsymbol{e}_i=a_{ii}>0$. Αντίστοιχα θα είναι $\boldsymbol{e}_i^T A\boldsymbol{e}_i=a_{ii}\geq 0$.

Πρόταση 4.2.3

Αν ένα μητρώο $A\in R^{n\times n}$ είναι θ.ο. (θ.η.ο), τότε όλα τα κύρια υπομητρώα του A_k ($A[1{:}k, 1{:}k]$, k=1,2,…n-1) είναι θ.ο. (θ.η.ο) και συνεπώς αντιστρέψιμα.

Απόδειξη: Έστω ότι A είναι θ.ο. Αν n=1, η πρόταση προφανώς ισχύει. Για n>1 θεωρούμε ένα κύριο υπομητρώο $\underline{A}_m$ του A, μεγέθους mxm με $1\leq m<n$ (αν $m=n$ ισχύει $A_m=A$ θ.ο.). Τότε το A μπορεί να γραφτεί σε μορφή μπλοκ:

$$A=\begin{bmatrix} \underline{A}_m & C \\ D & F \end{bmatrix}$$

όπου τα μητρώα C, D και F είναι μεγεθών mxk, kxn και kxk αντίστοιχα, έτσι ώστε $m+k=n$ και $1\leq m<n$. Έστω τώρα ένα τυχαίο $\boldsymbol{x}\in R^m$, με $\boldsymbol{x}\neq\boldsymbol{0}$. Επιλέγουμε το διάνυσμα $\boldsymbol{u}=[\boldsymbol{x}\ ;\ \boldsymbol{0}_{k\times 1}]^T\in R^n$. Είναι $\boldsymbol{u}\neq\boldsymbol{0}$ και αφού το A θ.ο., ισχύει $\boldsymbol{u}^T A\boldsymbol{u}>0$. Έχουμε:

$$\mathbf{u}^T\mathbf{A}\mathbf{u}=\begin{bmatrix}\mathbf{x}^T & \mathbf{0}_{1\times k}\end{bmatrix}\begin{bmatrix}\mathbf{A}_m & \mathbf{C}\\ \mathbf{D} & \mathbf{F}\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{0}_{k\times 1}\end{bmatrix}=\begin{bmatrix}\mathbf{x}^T\mathbf{A}_m+\mathbf{0} & \mathbf{x}^T\mathbf{C}+\mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{0}_{k\times 1}\end{bmatrix}=\mathbf{x}^T\mathbf{A}_m\mathbf{x}>0$$

Επομένως όλα τα κύρια υπομητρώα A_m ενός θ.ο. μητρώου είναι θ.ο. Αν A θ.η.ο, ισχύουν τα παραπάνω με αντικατάσταση του $\geq$ με $>$.

Πρόταση 4.2.4

Οι πραγματικές ιδιοτιμές ενός θ.ο. (θ.η.ο.) μητρώου που αντιστοιχούν σε πραγματικά ιδιοδιανύσματα είναι θετικές (μη αρνητικές). [*το αντίστροφο δεν ισχύει*]

Απόδειξη: Έστω λ μια πραγματική ιδιοτιμή ενός θ.ο. μητρώου $A\in R^{n\times n}$. Αν $\boldsymbol{x}$ το αντίστοιχο ιδιοδιάνυσμα του R^n, θα είναι $A\boldsymbol{x}=\lambda\boldsymbol{x}$. Τότε $\boldsymbol{x}^T A\boldsymbol{x}=\boldsymbol{x}^T\lambda\boldsymbol{x}=\lambda\boldsymbol{x}^T\boldsymbol{x}=\lambda\|\boldsymbol{x}\|^2>0$, οπότε $\lambda>0$.

Πρόταση 4.2.5

Αν ένα μητρώο A είναι θ.ο. (θ.η.ο.) τότε και το A^T είναι θ.ο. (θ.η.ο.), και αντιστρόφως.

Απόδειξη: Από την $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}>0$ έχουμε $(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x})^T>0 \Leftrightarrow (\boldsymbol{A}\boldsymbol{x})^T(\boldsymbol{x}^T)^T=\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{x}>0$, άρα το $\boldsymbol{A}^T$ είναι θ.ο.

Παρατήρηση 4.2.2 Η έννοια του θ.ο. μητρώου είναι δύσκολο να εφαρμοσθεί σε οποιαδήποτε μιγαδικά μητρώα, αφού το γινόμενο $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ είναι γενικά μιγαδικός αριθμός. Αντίθετα, ιδιαίτερη εφαρμογή βρίσκουν τα *Ερμιτιανά μητρώα* που συζητούνται στην §4.4.2. Έτσι, οι Προτάσεις 4.2.1 έως και 4.2.5 ισχύουν και για τα Ερμιτιανά μητρώα.

♦ Παράδειγμα 4.2.2

Εξετάζουμε για ποια πραγματικά θ είναι θ.ο. το παρακάτω μητρώο περιστροφής R.

$$R=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$

Λαμβάνουμε το $\text{sym}(R)=(R+R^T)/2=\text{diag}([\cos(\theta), \cos(\theta)]$, οπότε το R είναι θ.ο. εάν και μόνον εάν $\cos(\theta)>0$ ή $\theta\in[0,\pi]$.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

♦ Παράδειγμα 4.2.3

Για την τετραγωνική μορφή: $q(x)=2x_1^2+3x_2^2+2x_3^2+2x_1x_2-2x_2x_3$ θα εξετάσουμε που ελαχιστοποιείται και θα τη θέσουμε ως άθροισμα τετραγώνων.

Βρίσκουμε το αντίστοιχο μητρώο:

$$A=\begin{bmatrix}2&1&0\\1&3&-1\\0&-1&2\end{bmatrix}$$

Το A είναι προφανώς θ.ο. και άρα η q είναι μονίμως θετική και εμφανίζει ελάχιστο (0) για $\boldsymbol{x}$=0.
Βρίσκουμε τώρα το χαρακτηριστικό πολυώνυμο του A:

$$\phi(\lambda)=\det(A-\lambda I)=$$

$$=\begin{vmatrix}2-\lambda&1&0\\1&3-\lambda&-1\\0&-1&2-\lambda\end{vmatrix}=\begin{vmatrix}2-\lambda&0&2-\lambda\\1&3-\lambda&-1\\0&-1&2-\lambda\end{vmatrix}=(2-\lambda)\begin{vmatrix}1&0&1\\1&3-\lambda&-1\\0&-1&2-\lambda\end{vmatrix}=$$

$$=(2-\lambda)\begin{vmatrix}1&0&1\\0&3-\lambda&-2\\0&-1&2-\lambda\end{vmatrix}=(2-\lambda)\begin{vmatrix}1&0&1\\0&3-\lambda&-2\\0&-1&2-\lambda\end{vmatrix}=(2-\lambda)\begin{vmatrix}1&0&1\\0&4-\lambda&-4+\lambda\\0&-1&2-\lambda\end{vmatrix}=$$

$$=(2-\lambda)(4-\lambda)\begin{vmatrix}1&0&1\\0&1&-1\\0&-1&2-\lambda\end{vmatrix}=(2-\lambda)(4-\lambda)(1-\lambda)$$

Οι ιδιοτιμές είναι λ_1=1, λ_2=3, λ_3=4 (θετικές, όπως αναμενόταν). Η q εκφράζεται ως άθροισμα τετραγώνων:

$$q(\mathbf{y})=\mathbf{y}^{\mathbf{T}}D\mathbf{y}=\sum_{i=1}^{3}\lambda_i y_i^2=\sum_{i=1}^{3}\lambda_i\left(Q^{\mathrm{T}}x\right)^2$$

όπου $\boldsymbol{y}=Q^{\mathrm{T}}\boldsymbol{x}$. Q είναι το ορθοκανονικό μητρώο ιδιοδιανυσμάτων του A, τα οποία βρίσκουμε λύνοντας το σύστημα $(A\text{-}\lambda I)\boldsymbol{x}$=0. Βρίσκουμε τους ιδιοχώρους και κανονικοποιούμε:

Για λ_1=1: $(A\text{-}I)\boldsymbol{x}=0 \Rightarrow \boldsymbol{u}_1=(-1,1,1)^{\mathrm{T}} \Rightarrow \boldsymbol{q}_1=\dfrac{1}{\sqrt{3}}\ (-1,1,1)^{\mathrm{T}}$

Για λ_2=2: $(A\text{-}2I)\boldsymbol{x}=0 \Rightarrow \boldsymbol{u}_2=(-1,1,1)^{\mathrm{T}} \Rightarrow \boldsymbol{q}_2=\dfrac{1}{\sqrt{2}}\ (1,0,1)^{\mathrm{T}}$

Για λ_3=4: $(A\text{-}4I)\boldsymbol{x}=0 \Rightarrow \boldsymbol{u}_3=(-1,-2,1)^{\mathrm{T}} \Rightarrow \boldsymbol{q}_3=\dfrac{1}{\sqrt{6}}\ (-1,-2,1)^{\mathrm{T}}$

Επομένως έχουμε:

$$Q=\begin{pmatrix}-1/\sqrt{3}&1/\sqrt{2}&-1/\sqrt{6}\\1/\sqrt{3}&0&-2/\sqrt{6}\\1/\sqrt{3}&1/\sqrt{2}&1/\sqrt{6}\end{pmatrix}\Rightarrow Q^{\mathrm{T}}=\begin{pmatrix}-1/\sqrt{3}&1/\sqrt{3}&1/\sqrt{3}\\1/\sqrt{2}&0&1/\sqrt{2}\\-1/\sqrt{6}&-2/\sqrt{6}&1/\sqrt{6}\end{pmatrix}\Rightarrow$$

$$q(x)=\sum_{i=1}^{3}\lambda_i y_i^2=\sum_{i=1}^{3}\lambda_i y_i^2=1\times\frac{1}{3}\left(-x_1+x_2+x_3\right)^2+3\times\frac{1}{2}\left(x_1+x_3\right)^2+4\times\frac{1}{6}\left(-x_1-2\sqrt{2}x_2+x_3\right)^2=$$

$$=\frac{1}{3}\left(-x_1+x_2+x_3\right)^2+\frac{3}{2}\left(x_1+x_3\right)^2+\frac{2}{3}\left(-x_1-2\sqrt{2}x_2+x_3\right)^2$$

□

♦ Παράδειγμα 4.2.4

Για καθένα από τα παρακάτω μητρώα, να εξετασθεί αν είναι θετικά ορισμένο. Στη συνέχεια να βρεθούν και να εξετασθούν οι ιδιοτιμές και τα ιδιοδιανύσματα. Για το πρώτο να εξετασθούν και οι κύριες υπο-ορίζουσες.

(α) H=[2 1 1; 1 3 -2; 1 -2 3],
(β) A=[3 -1 -2; -2 2 0; 1 0 2],
(γ) B=[9 4 -4 ; 6 14 8 ; -4 8 2],
(δ) C= [6 -3 2; 3 5 1; 2 -1 4]

Απάντηση

(α) Το H είναι συμμετρικό και ιδιάζον. Διαπιστώνουμε επίσης εύκολα ότι ο 3^{ος} οδηγός είναι 0 και οι υπόλοιποι δύο θετικοί. Αυτό, όπως θα δούμε, δείχνει ότι είναι θετικά ημιορισμένο. Επίσης βρίσκουμε ότι οι ιδιοτιμές του H είναι 0 (όπως αναμένεται), 3 και 5. Και η ύπαρξη μηδενικής ιδιοτιμής δείχνει ότι το A είναι θ.η.ο. Τέλος βρίσκουμε: $\det(H(1{:}1,1{:}1))=2$, $\det(H(1{:}1,1{:}1))=5$ και $\det(H(1{:}3,1{:}3))=0$, δηλαδή θετικές ή 0, γεγονός που σημαίνει όπως θα δούμε ότι το H είναι θ.η.ο.

(β) Το A δεν είναι συμμετρικό, οπότε μεταβαίνουμε στο $\mathrm{sym}(A)=(A^T+A)/2$, ή ισοδύναμα στο A^T+A:

$$A^T+A=\begin{bmatrix} 6 & -3 & -1 \\ -3 & 4 & 0 \\ -1 & 0 & 4 \end{bmatrix}$$

Παρατηρούμε ότι το A^T+A παρουσιάζει θετική και «υπερέχουσα» κύρια διαγώνιο. Όπως θα δούμε είναι θ.ο.

Υπολογίζουμε τώρα τις ιδιοτιμές και τα ιδιοδιανύσματα του A, σύμφωνα με τον τύπο διαγωνοποίησης $A=VDV^{-1}$. Οι στήλες του V περιλαμβάνουν τα ιδιοδιανύσματα, ενώ οι ιδιοτιμές είναι τα διαγώνια στοιχεία του διαγώνιου D:

```
>> [V,D]=eig(A)
V =
   0.4082   0.0000  -0.0000
  -0.8165  -0.8944   0.8944
   0.4082   0.4472  -0.4472
D =
   3.0000        0        0
        0   2.0000        0
        0        0   2.0000
```

Παρατηρούμε ότι όλες οι ιδιοτιμές είναι πραγματικές και θετικές, αλλά δεν μπορούμε να αποφανθούμε μόνον από αυτό. Ομοίως βρίσκουμε ότι και όλες οι ιδιοτιμές του A^T+A είναι θετικές. Μπορούμε να αποφανθούμε ότι είναι θ.ο. μόνον από αυτό!

(γ) Υπολογίζουμε το $\mathrm{sym}(B)$:

```
>>symB=1/2*(B'+B)

symB =
    9    5   -4
    5   14    8
   -4    8    2
```

Όμοια, υπολογίζουμε τις ιδιοτιμές και τα ιδιοδιανύσματα του B:

```
>> [V,D]=eig(B)

V =
   0.3747  -0.8207   0.2135
  -0.4670   0.1821   0.9007
   0.8010   0.5416   0.3784
D =
  -4.5356        0        0
        0  10.7519        0
        0        0  18.7836
```


Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

Παρατηρούμε ότι μια ιδιοτιμή (-4.5356) είναι αρνητική. Ομοίως βρίσκουμε ότι και το symB έχει την πρώτη ιδιοτιμή αρνητική (-4.5539). Συνεπώς το A δεν είναι θ.ο.

(δ) Εξετάζουμε το sym(C):

```
>>symC = 1/2*(C'+C)

   6    0    2
   0    5    0
   2    0    4
```

Παρατηρούμε ότι και το symC παρουσιάζει θετική και «υπερέχουσα» κύρια διαγώνιο. Όπως θα δούμε είναι θ.ο.
Όμοια, υπολογίζουμε τις ιδιοτιμές και τα ιδιοδιανύσματα του C:

```
>> [V,D]=eig(C)
V =
  -0.1778 - 0.5898i  -0.1778 + 0.5898i   -0.4150
  -0.7071                 -0.7071          0.1658
  -0.2135 - 0.2736i  -0.2135 + 0.2736i   0.8946
D =
   6.0565 + 2.8894i        0                  0
        0            6.0565 - 2.8894i         0
        0                  0               2.8869
```

Παρατηρούμε ότι η μόνη πραγματική ιδιοτιμή (2.8869) είναι θετική και αντιστοιχεί σε ιδιοδιάνυσμα στο χώρο R^3. Υπάρχουν άλλες δυο ιδιοτιμές μιγαδικές συζυγείς, με αντίστοιχα ιδιοδιανύσματα στο μιγαδικό χώρο C^3 (δεν μπορούμε να αποφανθούμε μόνον από αυτό για το C).

Υπολογίζουμε τώρα τις ιδιοτιμές και τα ιδιοδιανύσματα του sym(C):

```
>> [V,D]=eig(symC)
V =
    0.5257         0    0.8507
         0   -1.0000         0
   -0.8507         0    0.5257
D =
    2.7639         0         0
         0    5.0000         0
         0         0    7.2361
```

Μπορούμε να αποφανθούμε ότι είναι θ.ο. μόνον από αυτό! □

4.3 Μητρώα με διαγώνια κυριαρχία

Μια πολύ ενδιαφέρουσα για τις εφαρμογές κατηγορία πραγματικών και μιγαδικών μητρώων είναι αυτά που παρουσιάζουν *διαγώνια κυριαρχία ή υπεροχή*: Τα μητρώα αυτά, ως μητρώα συντελεστών, διευκολύνουν την επίλυση των γραμμικών συστημάτων.

Ορισμός 4.3.1. Διαγώνια Κυριαρχία.

Ένα μητρώο A καλείται *αυστηρά διαγώνια κυρίαρχο*, ή με *αυστηρή διαγώνια υπεροχή* (*διαγώνια κυρίαρχο*), *κατά γραμμές* (*στήλες*), αν τα διαγώνια στοιχεία του είναι κατ' απόλυτη τιμή μεγαλύτερα (μεγαλύτερα ή ίσα) από το άθροισμα των απολύτων τιμών των υπόλοιπων στοιχείων της γραμμής (στήλης) τους. Στο εξής, οι παραπάνω ιδιότητες θα σημειώνονται με *αδκ.* και *δκ.* αντίστοιχα. Έτσι, έχουμε:

$$\textbf{\textit{Αδκ. κατά γραμμές}}: |a_{ii}| > \sum_{j=1,\, j\neq i}^{n} |a_{ij}|\ ,\ 1\le i\le n.\quad \textbf{\textit{Δκ. κατά γραμμές}}: |a_{ii}| \ge \sum_{j=1}^{n} |a_{ij}|\ ,\ 1\le i\le n \quad (4.3.1)$$

$$\textbf{\textit{Αδκ. κατά στήλες:}}\ |a_{ii}| > \sum_{j=1,\, j\neq i}^{n} |a_{ji}|\ ,\ 1\le i\le n.\quad \textbf{\textit{Δκ. κατά στήλες:}}\ |a_{ii}| \ge \sum_{j=1}^{n} |a_{ji}|,\ 1\le i\le n$$

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

Η αδκ. (δκ.) του A κατά γραμμές (στήλες) σημαίνει αδκ. (δκ.) του A^T κατά στήλες (γραμμές).

Οι ιδιότητες των αδκ. μητρώων, όπως και των θ.ο. είναι συνυφασμένες με τις ιδιοτιμές. Από την άλλη πλευρά, ο υπολογισμός και η οριοθέτηση των τελευταίων είναι ένα από τα πιο σημαντικά προβλήματα της Γραμμικής Άλγεβρας και της Αριθμητικής Ανάλυσης. Ένα κεντρικό και πολύ χρήσιμο αποτέλεσμα για τον εντοπισμό διαστημάτων των ιδιοτιμών δίνεται από το θεώρημα που ακολουθεί.

Θεώρημα 4.3.1 [*Κύκλοι Gerschgorin*]

Οι ιδιοτιμές ενός n×n μητρώου A περιέχονται στην ένωση των κύκλων του μιγαδικού επιπέδου με κέντρα (a_{ii}, 0) και ακτίνες $R_i = \sum_{j=1,\, j\neq i}^{n} |a_{ij}|$, ***i=1,…n.***

Απόδειξη (βλ. βιβλιογραφία)

Μια ισοδύναμη διατύπωση του Θ4.3.1 είναι ότι κάθε ιδιοτιμή του A ανήκει σε κάποιον κύκλο C[(a_{ii},0), R_i]. Αυτό σημαίνει ότι για κάθε ιδιοτιμή του A υπάρχει i, $1 \le i \le n$, τέτοιο ώστε:

$$|a_{ii} - \lambda| \le R_i = \sum_{j=1,\, j\neq i}^{n} |a_{ij}|$$

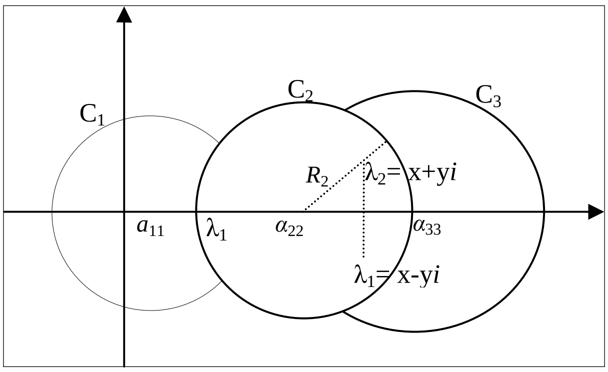
Συνεπώς κάθε ιδιοτιμή οριοθετείται σε μια περιοχή του μιγαδικού επιπέδου. Αν είναι πραγματική, τότε θα ανήκει σε ένα διάστημα πραγματικών (ισχύει η απόλυτη τιμή πραγματικού):

$$a_{ii} - R_i \le \lambda \le a_{ii} + R_i$$

Αν είναι μιγαδική, τότε οριοθετούνται τα *real*(λ) και *imag*(λ) (εφαρμόζεται το μιγαδικό μέτρο).

$$\sqrt{(a_{ii} - real(\lambda))^2 + (imag(\lambda))^2} \le R_i$$

Στο σχ. 4.3.1 φαίνονται τρεις κύκλοι *Gerschgorin* για ένα μητρώο 3×3 με μια πραγματική (λ_1) και δύο μιγαδικές συζυγείς (λ_2, λ_3) ιδιοτιμές.



Σχήμα 4.3.1 Κύκλοι *Gerschgorin*

♦Παράδειγμα 4.3.1

● Για το μητρώο A=[1 2 0; 8 1 0; 0 0 3] βρίσκουμε: R_1=2, R_2=8, R_3=0, επομένως αν λ ιδιοτιμή, τότε:

$$\lambda \in C[(1,0),2] \cup C[(1,0),8] \cup C[(3,0),0]$$

δηλ. ικανοποιεί μια από τις ανισότητες: $|\lambda-1| \le 2$, $|\lambda-1| \le 8$, $|\lambda-3|=0$.
Προφανώς ή ένωσή τους είναι $|\lambda-1| \le 8$. Αν μάλιστα είναι πραγματικές, θα πρέπει $\lambda \in [-7,9]$. Υπολογίσουμε τώρα τις ιδιοτιμές του A παίρνοντας:

$$\varphi(x) = \det(A - \lambda I) = -(\lambda-3)(\lambda-5)(\lambda+3)$$

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

απ' όπου: $\lambda_1=-3$, $\lambda_2=3$ $\lambda_3=5$. Πράγματι λοιπόν, οι ιδιοτιμές του A περιέχονται στο διάστημα αυτό.

● Δουλεύοντας όμοια και για το C=[1 1 -1; 1 2 -1; 2 1 0] παίρνουμε: $|\lambda-1|\le 2$, $|\lambda-2|\le 2$, $|\lambda|\le 3$ που συγχωνεύονται στους $|\lambda-2|\le 2$, $|\lambda|\le 3$. Βρίσκουμε εύκολα τις ιδιοτιμές: $\lambda_1=1+i$, $\lambda_2=1-i$, $\lambda_3=3$ που εύκολα διαπιστώνουμε ότι ανήκουν στους παραπάνω κύκλους. □

Πρόταση 4.3.1

Αν ένα μητρώο έχει αδκ. (κατά γραμμές ή στήλες), τότε είναι αντιστρέψιμο. Αν έχει δκ., τότε μπορεί να είναι ιδιάζον.

Απόδειξη: α) *Αδκ. κατά γραμμές*: Έστω A ένα μητρώο με α.δ.κ. κατά γραμμές, δηλ. τέτοιο ώστε:

$$|a_{ii}| > R_i = \sum_{j=1, j\neq i}^{n} |a_{ij}| \ , \ 1\le i\le n$$

Τότε από το *Θ. των Κύκλων του Gerschgorin* προκύπτει ότι, αν λ μια ιδιοτιμή του A, θα ανήκει σε κάποιο κύκλο $C[(a_{ii},0), R_i]$, δηλαδή θα ισχύει:

$$|\lambda - a_{ii}| \le R_i < |a_{ii}| \quad \text{για κάποιο } i\in\{1,2,\ldots n\} \tag{4.3.2}$$

Για $\lambda=0$ προκύπτει $|a_{ii}|\le R_i$ (ή $|a_{ii}|<|a_{ii}|$) που αντιβαίνει στην παραδοχή της αδκ., συνεπώς $\lambda=0$ δεν μπορεί να είναι ιδιοτιμή. Άρα όλες οι ιδιοτιμές του A είναι διάφορες του 0, δηλ. το A είναι αντιστρέψιμο.

Αν το A έχει δκ., η (4.2.2) είναι αληθής για $\lambda=0$: $|a_{ii}| \le R_i$ και επομένως το A μπορεί να είναι ιδιάζον.

β) *Αδκ. κατά στήλες*: Αν το A έχει αδκ. κατά στήλες, τότε το A^T έχει αδκ. κατά γραμμές, συνεπώς σύμφωνα με το (α) είναι αντιστρεπτό, άρα και το A αντιστρεπτό.
Ομοίως σκεπτόμαστε και για την περίπτωση της δκ.

Πόρισμα 4.3.1 *Αν ένα μητρώο A έχει αδκ. κατά γραμμές ή στήλες (δκ.), τότε όλα τα κύρια υπομητρώα του* $\underline{A}_k, 1\le k\le n$, *είναι αντιστρεπτά (κάποια μπορεί να είναι ιδιάζοντα).*

Απόδειξη: Αφού το A έχει αδκ. κατά γραμμές, όλα τα κύρια υπομητρώα του, $\underline{A}_k$, θα έχουν προφανώς αδκ. κατά μείζονα λόγο και σύμφωνα με την Π4.3.1 θα είναι αντιστρεπτά. Συνεπώς από την Πρόταση 4.3.1 θα είναι αντιστρέψιμο.

Αν το A έχει αδκ. κατά στήλες, τότε όλα τα κύρια υπομητρώα $(\underline{A^T})_k$ του A^T θα είναι αντιστρεπτά. Αλλά επειδή προφανώς είναι $(\underline{A^T})_k=(\underline{A}_k)^T$ (είναι δηλ. ανάστροφα των $\underline{A}_k$), άρα και τα $\underline{A}_k$ θα είναι αντιστρεπτά.

Ανάλογοι συλλογισμοί ισχύουν για την περίπτωση της δ.κ. ■

♦ Παράδειγμα 4.3.2

Το τρισδιαγώνιο μητρώο A έχει α.δ.κ.:

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

Επομένως είναι αντιστρεπτό, καθώς και όλα τα κύρια υπομητρώα του $\underline{A}_i$, $1\le i\le 3$. Πράγματι, όλες οι υποορίζουσες είναι μη μηδενικές:

$$\det(A_1)=4,\ \det(A_2)=15,\ \det(A_3)=56,\ \det(A_4)=209$$

Επιπλέον, παρατηρούμε ότι είναι και θετικές, όπως και οι οδηγοί. Ως εκ τούτου το A είναι και θετικά ορισμένο, όπως θα δούμε στην §4.4.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

Δείχνουμε επίσης με ένα σύντομα αντιπαράδειγμα ότι οι συνθήκες της Π4.3.1 και του Πορ.4.3.1 δεν είναι αναγκαίες: το A=[4 1 0 0; 1 4 1 0; 0 1 4 1.5; 0 0 1.5 1] ενώ δεν παρουσιάζει αδκ. (1<1.5), εν τούτοις έχει όλες τις υπο-ορίζουσες μη μηδενικές (και μάλιστα θετικές). □

♦ Παράδειγμα [Matlab] 4.3.3 Δημιουργία μητρώου με αδκ. (Matlab)

Οι παρακάτω εντολές δημιουργούν ένα «1-4-1» τρισδιαγώνιο μητρώο μεγέθους m, με αδκ. Για m=2 παίρνουμε το μητρώο A του Παρ. 4.3.1.

```
m=5;
diag(4*ones(2*m,1)) + diag(ones(2*m-1,1),1) + diag(ones(2*m-1,1),-1)
```

```
ans =
   4   1   0   0   0   0   0   0   0   0
   1   4   1   0   0   0   0   0   0   0
   0   1   4   1   0   0   0   0   0   0
   0   0   1   4   1   0   0   0   0   0
   0   0   0   1   4   1   0   0   0   0
   0   0   0   0   1   4   1   0   0   0
   0   0   0   0   0   1   4   1   0   0
   0   0   0   0   0   0   1   4   1   0
   0   0   0   0   0   0   0   1   4   1
   0   0   0   0   0   0   0   0   1   4
```

□

♦ [Matlab] Παράδειγμα 4.3.4 Έλεγχος αυστηρής διαγώνιας κυριαρχίας

Η λογική συνάρτηση *is_diag_dom* (A) ελέγχει στο Matlab α.δ.κ. κατά γραμμές:

```
function [k] = is_diag_dom (A)
% Έλεγχος αυστηρής διαγώνιας κυριαρχίας κατά γραμμές
[m,n] = size(A);
if m~=n,
    return;
end;
k=false;
for i=1:length(A)
    s=0;
    for j=1,length (A),
        if j~=i,
            s=s+abs(A(i,j));
        end;
    if ( abs(A(i,i))<=s ) % περίπτωση δ.κ.: abs(A(i,i))<s
        return;
    end;
    end;
end;
k=true ;
```

Οι παρακάτω κλήσεις πραγματοποιούν ελέγχους για α.δ.κ. κατά γραμμές και στήλες:

```
Chek_for_ddr=is_diag_dom (F);
Chek_for_ddc=is_diag_dom (F');
```

Μια δεύτερη εκδοχή για την πραγματοποίηση και και των δύο ελέγχων είναι:

```
function [row, col] = is_diag_dom_rc(A)
%checks a given matrix for diagonal dominance by rows and columns
row=false; col=false;
if ( 2*abs(diag(A))>sum(abs(A'))'), %diagonal dominance by rows
    row=true;
end;
if ( 2*abs(diag(A))>sum(abs(A))'), %diagonal dominance by columns
    col=true;
end;
```

□

4.4 Συμμετρικά και Ερμιτιανά Θετικά Ορισμένα Μητρώα

Μεταξύ των θ.ο. μητρώων, τα πλέον χρήσιμα στις εφαρμογές είναι τα συμμετρικά και τα Ερμιτιανά μητρώα, τα οποία παρουσιάζουν ενδιαφέρουσες ιδιότητες. Τα κεντρικά κριτήρια για τα μητρώα αυτά συνοψίζονται στο επόμενο Θεώρημα.

Θεώρημα 4.4.1: Κριτήρια για θετικά ορισμένα συμμετρικά μητρώα

Αν A είναι ένα συμμετρικό πραγματικό τετραγωνικό μητρώο nxn, οι παρακάτω προτάσεις είναι ισοδύναμες:

i) Το A είναι θετικά ορισμένο: για κάθε $x \neq 0$: $x^T A x > 0$ (θετικά ημιορισμένο)

ii) Ολες οι n ιδιοτιμές του A είναι θετικές (μη αρνητικές)

iii) Όλοι οι n οδηγοί (χωρίς εναλλαγές γραμμών) είναι θετικοί (μη αρνητικοί)

iv) Όλες οι κύριες ορίζουσες $\Delta_i = det(A(1{:}i,1{:}i))$, $i=1,\ldots,n$, είναι θετικές (μη αρνητικές). *[Sylvester's Minorant Criterion]*

Τα (ii)-(iv) χρησιμοποιούνται ως πρακτικά κριτήρια για το εάν ένα μητρώο A είναι θ.ο.
Για την απόδειξη του Θ4.4.1 θα χρησιμοποιήσουμε δύο αποτελέσματα της Γραμμικής Άλγεβρας που ακολουθούν. Το πρώτο είναι ένα απλό αποτέλεσμα της παραγοντοποίησης *LU*, ενώ το δεύτερο πραγματεύεται βασικές ιδιότητες των κύριων υπομητρώων του *A*. Θα συνοψίσουμε τα αποτελέσματα αυτά μαζί με άλλα στην §4.4.2.

Πρόταση 4.4.1

Αν κατά την εφαρμογή της απαλοιφής Gauss χωρίς οδήγηση σε ένα αντιστρεπτό μητρώο A δεν γίνει καμία εναλλαγή γραμμών, τότε ισχύει και αντιστρόφως, η παραγοντοποίηση A=LU:

> *Για κάθε $k=1,\ldots,n-1$ είναι $A_k=L_kA$, όπου $L_k= E_kE_{k-1}\ldots E_1$,*
> *έτσι ώστε $A=LU$, όπου $L=(L_{n-1})^{-1}$ και $U=A_{n-1}$.*

Σημειώνεται ότι E_i είναι το μητρώο απαλοιφής στο βήμα i, το L_k κάτω τριγωνικό, και το L κάτω τριγωνικό με τους πολλαπλασιαστές p_{ij} και με 1 στη διαγώνιο. Τέλος, το U είναι το άνω τριγωνικό μητρώο που προκύπτει από την απαλοιφή Gauss. Η απόδειξη είναι προφανής. Υπενθυμίζουμε εδώ μόνον τις μορφές των E_i, L και U:

$$E_i = \begin{bmatrix} 1 & & & & \\ & \ddots & & \mathbf{0} & \\ & & 1 & & \\ \mathbf{0} & & -p_{i+1,i} & \ddots & \\ & & \vdots & & \\ & & -p_{n,i} & \mathbf{0} & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & & & & \\ p_{2,1} & 1 & & \mathbf{0} & \\ \vdots & & 1 & & \\ \vdots & & p_{i+1,i} & \ddots & \\ & & \vdots & & \\ p_{n,1} & \cdots & p_{n,i} & \cdots & 1 \end{bmatrix}, \quad U = \begin{bmatrix} d_1 & * & \ldots & * & 0 \\ & d_2 & * & \cdots & * \\ & & \ddots & \cdots & \vdots \\ & \mathbf{0} & & \ddots & \vdots \\ & & & & d_n \end{bmatrix}$$

Πρόταση 4.4.2

Αν όλα τα κύρια υπομητρώα $\underline{A}_k$, $k=1,\ldots,n-1$, ενός αντιστρεπτού μητρώου A είναι αντιστρεπτά (ισοδύναμα αν $det(\underline{A}_k)=\Delta_k \neq 0$), τότε δεν γίνονται εναλλαγές γραμμών κατά την εφαρμογή στο A της απαλοιφής Gauss χωρίς οδήγηση. Ισχύει και το αντίστροφο.

Απόδειξη

α) *Ευθύ:* Αρκεί να δείξουμε ότι σε κάθε βήμα $k=1,\ldots,n-1$ της απαλοιφής, το υπολογιζόμενο νέο στοιχείο a_{kk} είναι οδηγός, δηλαδή ότι ισχύει $a_{kk} \neq 0$, δηλ. το $d_k=a_{kk}$ θα είναι ο k οδηγός (υποθέτουμε ότι οι πράξεις επιδρούν πάνω στο *A* και επομένως τα a_{ij} αλλάζουν). Χρησιμοποιούμε επαγωγή:

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

Για k=1 το $\underline{A}_k$=[α$_{11}$] είναι αντιστρεπτό, άρα $a_{11}\neq 0$, δηλ. η πρόταση ισχύει. Δεχόμαστε τώρα ότι η πρόταση ισχύει για $k=m-1$, με $m\leq n-1$. Θα δείξουμε ότι ισχύει και για $k=m$. Από την υπόθεση προκύπτει ότι στο m-1 βήμα της απαλοιφής θα έχουν υπολογισθεί m-1 οδηγοί $d_i=a_{ii}\neq 0$ και ότι συγχρόνως θα ισχύει $A_{m-1}=L_{m-1}A$. Αν τώρα γράψουμε την $A_{m-1}=L_{m-1}A$ υπό μορφή συμβατών μπλοκ, διαχωρίζοντας όλα τα μητρώα στις πρώτες m γραμμές και στήλες, έχουμε:

$$\mathbf{A}_{m-1}=\begin{bmatrix}\mathbf{U_m} & \mathbf{B}_{m\times(n-m)}\\ \mathbf{C}_{(n-m)\times m} & \mathbf{D}_{(n-m)\times(n-m)}\end{bmatrix}=\begin{bmatrix}\mathbf{\Lambda}_{m\times m} & \mathbf{0}_{m\times(n-m)}\\ \mathbf{M}_{(n-m)\times m} & \mathbf{N}_{(n-m)\times(n-m)}\end{bmatrix}*\begin{bmatrix}\underline{\mathbf{A}}_{m\times m} & \mathbf{G}_{m\times(n-m)}\\ \mathbf{H}_{(n-m)\times m} & \mathbf{K}_{(n-m)\times(n-m)}\end{bmatrix} \quad (4.4.1)$$

Στην (4.4.1) το U_m είναι προφανώς άνω τριγωνικό με diag(U_m)=$(d_1, d_2,\ldots,d_{m-1,m-1},a_{mm})^{\mathrm{T}}$, όπου $d_i\neq 0$. Επίσης τα Λ (με μονάδες στη διαγώνιο) και N είναι κάτω τριγωνικά, και ως εκ τούτου αντιστρέψιμα. Το άνω αριστερό μπλοκ του A είναι το $\underline{A}_k$. Από την (4.4.1), εκτελώντας τις πράξεις, το U_m δίνεται:

$$U_m=\mathbf{\Lambda}\underline{\mathbf{A}}_m+\mathbf{0}*\mathbf{H}=\mathbf{\Lambda}\underline{\mathbf{A}}_m \quad (4.4.2)$$

Αυτό σημαίνει ότι το U_m είναι αντιστρέψιμο ως γινόμενο δύο αντιστρέψιμων μητρώων. Συνεπώς θα πρέπει και $\alpha_{mm}\neq 0$, και ή απόδειξη είναι πλήρης. ■

β) *Αντίστροφο*: Αν υποθέσουμε ότι A είναι αντιστρέψιμο και δεν γίνονται εναλλαγές γραμμών, τότε ισχύουν οι (4.4.1) και (4.4.2) (για κάθε $m\leq n-1$), καθώς και ότι U_m είναι αντιστρέψιμο. Από την (4.4.2) προκύπτει άμεσα ($det(\Lambda)$=1): $\Delta_k=det(U_k)=d_1d_2\ldots d_k\neq 0$. ■

Παρατήρηση 4.4.1: Η παραπάνω σχέση $det(\underline{A}_k)=d_1d_2\ldots d_k$ σημαίνει ότι αν οι γραμμές ενός μητρώου A δεν εναλλάσσονται κατά την απαλοιφή *Gauss*, *τότε οι οδηγοί κάθε $\underline{A}_k$ συμπίπτουν με τους πρώτους k οδηγούς του A*. Εναλλακτική διατύπωση: *οι οδηγοί $d_1,\ldots d_k$ προέρχονται από το υπομητρώο $\underline{A}_k$*. Στο ίδιο συμπέρασμα καταλήγουμε και άμεσα: κατ' αρχή διαπιστώνουμε επαγωγικά, αρχίζοντας από το υπομητρώα $\underline{A}_1$ και $\underline{A}_2$, ότι αφού δεν γίνονται εναλλαγές, οι πολλαπλασιαστές για κάθε γραμμή είναι οι ίδιοι με αυτούς του A. Επομένως ο υπολογισμός του d_i εξαρτάται μόνον από το από πάνω του στοιχείο στη στήλη (το οποίο δεν μετατίθεται) και τον πολλαπλασιαστή για τη γραμμή του, δηλαδή από τα δεδομένα του $\underline{A}_k$.

Παρατήρηση 4.4.2: Η $det(\underline{A}_k)=d_1d_2\ldots d_k$ γράφεται και $det(\underline{A}_k)=d_1d_2\ldots d_{k-1}d_k = det(\underline{A}_{k-1})d_k$, απ' όπου προκύπτει (με τις παραπάνω παραδοχές) η σχέση που δίνει τους οδηγούς από τις ορίζουσες: $d_k=det(\underline{A}_k)/det(\underline{A}_{k-1})$.

Απόδειξη του Θ.4.4.1

Θα δείξουμε τις ισοδυναμίες αποδεικνύοντας τις συνεπαγωγές: (i)⇔(ii)⇔(iii), (i)⇒(iv), (iv)⇒(iii). Δεν θα ασχοληθούμε με την περίπτωση του θ.η.ο. A.

(i)⇒(ii): Τα συμμετρικά μητρώα έχουν ως γνωστόν n (συμπεριλαμβανόμενης της πολλαπλότητας) πραγματικές ιδιοτιμές. Έστω A συμμετρικό και θ.ο. και λ μια ιδιοτιμή του. Τότε $A\boldsymbol{v}=\lambda\boldsymbol{v}$, όπου $\boldsymbol{v}$ ιδιοδιάνυσμα με $\boldsymbol{v}\neq\mathbf{0}$. Άρα $\boldsymbol{v}^{\mathrm{T}}A\boldsymbol{v}=\lambda\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}=\lambda\|\boldsymbol{v}\|^2>0 \Rightarrow \lambda>0$. ■

(ii)⇒(i): Αντιστρόφως, αν κάθε ιδιοτιμή $\lambda>0$, δείχνουμε ότι για κάθε $\boldsymbol{v}\neq 0$, είναι $\boldsymbol{v}^{\mathrm{T}}A\boldsymbol{v}>0$. Αφού A συμμετρικό, τότε σύμφωνα με το *φασματικό θεώρημα* υπάρχει μια πλήρης ορθοκανονική βάση ιδιοδιανυσμάτων $\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n$. Τα $\boldsymbol{v}$ και $A\boldsymbol{v}$ γράφονται ως συνδυασμοί της βάσης:

$$\boldsymbol{v}=a_1\boldsymbol{q}_1+\ldots+a_n\boldsymbol{q}_n$$
$$A\boldsymbol{v}=a_1\lambda_1\boldsymbol{q}_1+\ldots+a_n\lambda_n\mathbf{q}_n$$

και

$$\boldsymbol{v}^{\mathrm{T}}A\boldsymbol{v}=(a_1\boldsymbol{q}_1+\ldots+a_n\boldsymbol{q}_n)^{\mathrm{T}}(a_1\lambda_1\boldsymbol{q}_1+\ldots+a_n\lambda_n\boldsymbol{q}_n)=$$
$$=(a_1\boldsymbol{q}_1^{\mathrm{T}}+\ldots+a_n\boldsymbol{q}_n^{\mathrm{T}})(a_1\lambda_1\boldsymbol{q}_1+\ldots+a_n\lambda_n\boldsymbol{q}_n)=\sum_{i\neq j}\alpha_i\alpha_j\lambda_j\mathbf{q}_\mathbf{i}^\mathbf{T}\mathbf{q}_\mathbf{j}+\sum_{i=1}^{n}\alpha_i^2\lambda_i\mathbf{q}_\mathbf{i}^\mathbf{T}\mathbf{q}_\mathbf{i}=$$

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

$$= 0 + \sum_{i=1}^{n} \alpha_i^2 \lambda_i \mathbf{q_i^T q_i} = \sum_{i=1}^{n} \alpha_i^2 \lambda_i > 0 \qquad \blacksquare$$

Και μια πιο συνεκτική Απόδειξη:

(i)⇔(ii): Αφού A συμμετρικό, σύμφωνα με το φασματικό θεώρημα είναι όμοιο με ένα πραγματικό διαγώνιο μητρώο $D=diag(\lambda_1,\lambda_2,\ldots,\lambda_n)$: $A=QDQ^T$, όπου Q περιέχει την ορθοκανονική βάση ιδιοδιανυσμάτων του A. Επομένως το $\boldsymbol{v}^T A\boldsymbol{v}$ γράφεται:

$$\boldsymbol{v}^T A\boldsymbol{v} = \boldsymbol{v}^T QDQ^T \boldsymbol{v} = (Q^T\boldsymbol{v})^T D(Q^T\boldsymbol{v}) = \lambda_1(\boldsymbol{q}_1^T\boldsymbol{v})^2 + \ldots + \lambda_n(\boldsymbol{q}_n^T\boldsymbol{v})^2 \tag{4.4.3}$$

όπου $Q^T\boldsymbol{v}=[\boldsymbol{q}_1^T\boldsymbol{v}, \boldsymbol{q}_2^T\boldsymbol{v},\ldots,\boldsymbol{q}_i^T\boldsymbol{v}]$, είναι οι συντεταγμένες (προβολές στους ορθογώνιους άξονες) του $\boldsymbol{v}$ ως προς την ορθογώνια βάση Q. Από την (4.4.3) προκύπτει ότι το $\boldsymbol{v}$ είναι θετικά ορισμένο εάν και μόνον εάν τα διαγώνια στοιχεία του D είναι θετικά, δηλαδή οι ιδιοτιμές $\lambda_i>0$. ■

(i)⇒(iv): Αφού το A είναι θ.ο, τότε (από Π4.2.3) και το $\underline{A}_k$ ($k=1,\ldots n$) θα είναι θ.ο. Συνεπώς, λόγω του (ii), θα έχει θετικές ιδιοτιμές $\mu_1,\mu_2,\ldots,\mu_k$ και επομένως $det(\underline{A}_k)=\mu_1\mu_2\ldots\mu_k>0$.
(Ας τονισθεί εδώ, ότι οι ιδιοτιμές του $\underline{A}_k$ *δεν* συμπίπτουν με αυτές του A!) ■

(vi)⇔(iii): Αν $det(\underline{A}_k)>0$, $k=1,\ldots,n$, ισχύει το συμπέρασμα της Π4.4.2, και από την Παρατήρηση 4.4.2 προκύπτει $d_k=det(\underline{A}_k)/det(\underline{A}_{k-1})>0$. ■

(ii)⇔(iii): Αυτό προκύπτει άμεσα από το γνωστό αποτέλεσμα ότι οι ιδιοτιμές και οι αντίστοιχοι οδηγοί ενός συμμετρικού μητρώου έχουν τα ίδια πρόσημα. ■

Παρατήρηση 4.4.3. Χωρίς να είναι αναγκαίο, παραθέτουμε ακόμα μια απόδειξη για τη μετάβαση **(i)⇒(iii)**:
Σύμφωνα με την Π4.4.2, αν το A είναι θ.ο., τότε παραγοντοποιείται: $A=LU$. Αφού A συμμετρικός, η σχέση γράφεται και $A=LDL^T$, όπου $D=\mathrm{diag}(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n)$ διαγώνιο μητρώο με τους οδηγούς $\boldsymbol{u}_i$ και L κάτω τριγωνικό μητρώο με μονάδες στη διαγώνιο και συνεπώς αντιστρέψιμο. Αφού A θ.ο., είναι:

για κάθε $\boldsymbol{x}\neq\mathbf{0}$: $\boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{x}^T LDL^T\boldsymbol{x} = (L^T\boldsymbol{x})^T DL^T\boldsymbol{x} > 0$

Η πιο πάνω ανισότητα δηλώνει ότι το D είναι θ.ο.: L είναι αντιστρέψιμο, άρα $L^T\boldsymbol{x}\neq\mathbf{0}$ και συνεπώς κάθε $\boldsymbol{y}\neq\mathbf{0}$ μπορεί να εκφρασθεί ως $\boldsymbol{y}=L^T\boldsymbol{x}$, με $\boldsymbol{x}\neq\mathbf{0}$, θέτοντας $\boldsymbol{x}=(L^T)^{-1}\boldsymbol{y}$ (πίσω αντικατάσταση):

για κάθε $\boldsymbol{y}\neq\mathbf{0}$: $\boldsymbol{y}^T D\boldsymbol{y} = u_1y_1^2+\ldots+u_ny_n^2 > 0$

Το D όμως, ως θ.ο. και διαγώνιο, πρέπει να έχει τα διαγώνια στοιχεία του θετικά: $u_i>0$, $1\leq i\leq n$. ■

Εξετάσαμε πιο πάνω τις ικανές και αναγκαίες συνθήκες για να είναι ένα συμμετρικό μητρώο θ.ο. Ένα πρόσθετο χρήσιμο κριτήριο (μόνον ικανό) σχετίζεται με την ιδιότητα της *διαγώνιας κυριαρχίας,* και είναι το εξής:

Πρόταση 4.4.3

Αν ένα συμμετρικό μητρώο A έχει α.δ.κ. (δ.κ) και όλα τα διαγώνια στοιχεία του θετικά, τότε είναι θ.ο.(θ.η.ο.). [*το αντίστροφο δεν ισχύει*].

Απόδειξη

Έστω ότι το A είναι συμμετρικό και έχει α.δ.κ. κατά γραμμές. Τότε:

$$|a_{ii}| = a_{ii} > R_i = \sum_{j=1, j\neq i}^{n} |a_{ij}|, 1\leq i\leq n$$

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

και όλες οι ιδιοτιμές του A θα είναι πραγματικές. Από το *Θ. των κύκλων του Gerschgorin* προκύπτει ότι αν λ μια ιδιοτιμή του A, τότε ανήκει σε κάποιο κύκλο $C[(a_{ii},0), R_i]$, δηλαδή θα ισχύει:

$$|\lambda - a_{ii}| \leq R_i < a_{ii} \Rightarrow 0 < \lambda < 2a_{ii}$$

Άρα όλες οι ιδιοτιμές είναι θετικές και αφού A συμμετρικό, θα είναι και θ.ο. Αν πάλι το A έχει δ.κ., τότε $R_i \leq a_{ii}$ και ομοίως προκύπτει $0 \leq \lambda \leq 2a_{ii}$, και άρα το A είναι θ.η.ο. Αν το A έχει α.δ.κ. κατά στήλες, τότε το A^T έχει α.δ.κ. (δ.κ.) κατά γραμμές και ομοίως θετικά διαγώνια στοιχεία, επομένως είναι θ.ο. (θ.η.ο.). Επομένως, από την Π4.2.5, και το A θα είναι θ.ο. (θ.η.ο.)

Πρόταση 4.4.4

(α) Αν A είναι θ.ο. συμμετρικό μητρώο, τότε και όλες οι δυνάμεις A^k, όπου k ακέραιος, είναι θ.ο. Ειδική περίπτωση για $k=-1$: αν A είναι θ.ο. συμμετρικό, τότε και το A^{-1} είναι θ.ο.

(β) Αν A είναι θ.η.ο. συμμετρικό, τότε και όλες οι μη αρνητικές δυνάμεις A^k είναι θ.ο. (θ.η.ο.).

(γ) Αν A και B είναι θ.ο. (θ.η.ο.) συμμετρικά μητρώα, τότε το ίδιο συμβαίνει και για το $A+B$.

Απόδειξη: (α) Αν k ακέραιος, τότε και το A^k θα είναι συμμετρικό. Αν το A είναι θ.ο., θα είναι και αντιστρέψιμο. Τότε όμως οι ιδιοτιμές του A^k θα είναι οι k-στές δυνάμεις (συμπεριλαμβανομένων και των αρνητικών) των ιδιοτιμών του A, Συνεπώς το A^k έχει θετικές ιδιοτιμές και άρα είναι θ.ο.

(β) Αν το A είναι θ.η.ο. και k μη αρνητικός ακέραιος, τότε το A θα έχει μια ή περισσότερες ιδιοτιμές μηδενικές και τις υπόλοιπες θετικές, επομένως θα είναι θ.η.ο.

(γ) Είναι $\boldsymbol{x}^TA\boldsymbol{x}>0$, $\boldsymbol{x}^TB\boldsymbol{x}>0$, άρα $\boldsymbol{x}^T(A+B)\boldsymbol{x}=\boldsymbol{x}^T(A\boldsymbol{x}+B\boldsymbol{x})= \boldsymbol{x}^TA\boldsymbol{x}+ \boldsymbol{x}^TB\boldsymbol{x}>0$, άρα $A+B$ θ.ο.

Τα μητρώα A^TA εμφανίζονται σε πολλές εφαρμογές της Γραμμικής Άλγεβρας και της Αριθμητικής Ανάλυσης. Αποδεικνύεται ότι είναι όλα *θετικά ημιορισμένα*:

Πρόταση 4.4.5

Όλα τα μητρώα A^TA είναι θετικά ημιορισμένα. Συγκεκριμένα:

(α) Εάν το A είναι τετραγωνικό και αντιστρέψιμο (ιδιάζον), τότε το A^TA είναι θ.ο (θ.η.ο).

(β) Αν $A \in \mathbb{R}^{m\times n}$ και έχει ανεξάρτητες (εξαρτημένες) στήλες*, τότε το A^TA είναι θ.ο. (θ.η.ο).*

(γ) Επιπλέον, σε κάθε περίπτωση ισχύει $rank(A)=rank(A^TA)$.

Απόδειξη: α) Αφού A είναι αντιστρέψιμο, τότε $\forall \boldsymbol{v}\neq\boldsymbol{0}$ είναι $A\boldsymbol{v}\neq\boldsymbol{0}$. Επομένως $\forall \boldsymbol{v}\neq\boldsymbol{0}$ είναι $\boldsymbol{v}^T(A^TA)\boldsymbol{v}=(\boldsymbol{v}^TA^T)(A\boldsymbol{v})=\|A\boldsymbol{v}\|^2>0$, δηλ. το A είναι θ.ο. Αν A είναι ιδιάζον, τότε είναι $A\boldsymbol{x}=\boldsymbol{0}$ για κάποια $\boldsymbol{v}\neq\boldsymbol{0}$. Άρα $\forall \boldsymbol{v}\neq\boldsymbol{0}$ ισχύει $\boldsymbol{v}^T(A^TA)\boldsymbol{v}=(\boldsymbol{v}^TA^T)(A\boldsymbol{v})=\|A\boldsymbol{v}\|^2 \geq 0$, δηλ. το A είναι θ.η.ο.

β) Και πάλι $\forall \boldsymbol{x}\neq 0$: $\boldsymbol{x}^TA\boldsymbol{x}=\boldsymbol{x}^TR^TR\boldsymbol{x}=\|R\boldsymbol{x}\|^2\geq 0$. Επειδή το R έχει ανεξάρτητες στήλες και $\boldsymbol{x}\neq 0$, είναι και $R\boldsymbol{x}\neq 0$. Συνεπώς $\|R\boldsymbol{x}\|^2>0$, δηλ. A θ.ο. Ανάλογη και η απόδειξη όταν R έχει εξαρτημένες στήλες.

γ) Ξεκινάμε με τη διαπίστωση ότι τα A και A^TA έχουν τον ίδιο μηδενοχώρο: $N(A^TA)=N(A)$. Πράγματι, αν $\boldsymbol{x}\in N(A)$, τότε $A\boldsymbol{x}=\boldsymbol{0} \Rightarrow A^TA\boldsymbol{x}=\boldsymbol{0} \Rightarrow \boldsymbol{x}\in N(A)$. Αν $\boldsymbol{x}\in N(A^TA) \Rightarrow A^TA\boldsymbol{x}=\boldsymbol{0} \Rightarrow \boldsymbol{x}^TA^TA\boldsymbol{x}=0 \Rightarrow (A\boldsymbol{x})^TA\boldsymbol{x}=0 \Rightarrow \|A\boldsymbol{x}\|^2=0 \Rightarrow A\boldsymbol{x}=0 \Rightarrow \boldsymbol{x}\in N(A)$. Επομένως, $N(A^TA)=N(A)$. Τώρα θα είναι: $rank(A^TA)=n\text{-}dim(N(A^TA))=n\text{-}dim(N(A))=rank(A)$. Αυτό σημαίνει ότι τα δύο μητρώα έχουν τον ίδιο χώρο στηλών, ή διαφορετικά: $dim(span(A\boldsymbol{x}))=dim(span(A^TA\boldsymbol{x}))$ για $\boldsymbol{x}\in\mathbb{R}^n$.

Σημείωση: Η Πρόταση 4.4.5 ισχύει και για τα μητρώα AA^T. Η απόδειξη λαμβάνεται άμεσα: αν A είναι αντιστρέψιμο, τότε το ίδιο είναι και το $\boldsymbol{A}^T$, συνεπώς $(\boldsymbol{A}^T)^T\boldsymbol{A}^T = \boldsymbol{A}\boldsymbol{A}^T$ είναι θ. ο. Επίσης ισχύει και για Ερμιτιανά μητρώα: όλα τα $A^*A = \overline{A}^T A$ με $A\in M(\mathbb{C}^n)$ είναι θ.ο. ή θ.η.ο.

Πόρισμα 4.4.1 *Όλα τα μητρώα A^TA έχουν ιδιοτιμές θετικές ή 0. Αν το A είναι αντιστρέψιμο, τότε όλες οι ιδιοτιμές του A^TA είναι θετικές, ενώ αν είναι ιδιάζον, όλες οι ιδιοτιμές του είναι θετικές ή 0.*

Απόδειξη: Προκύπτει άμεσα από την Π4.4.5 και το ότι το A έχει πραγματικές ιδιοτιμές ως συμμετρικό μητρώο.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

4.4.1 Συμπληρωματικά Κριτήρια για θ.ο. Μητρώα

Μπορούν να διατυπωθούν και άλλες χρήσιμες ιδιότητες για τα συμμετρικά θ.ο. μητρώα. Εκτός των άλλων, μπορούν να είναι ωφέλιμες ως «χαλαρά» αναγνωριστικά κριτήρια για τα μητρώα αυτά. Η επόμενη πρόταση πραγματεύεται μερικές από τις ιδιότητες αυτές:

Πρόταση 4.4.6

Σε κάθε θ.ο. συμμετρικό μητρώο A ισχύει

(α) Για κάθε στοιχείο a_{ij} του A, με i<j, ισχύει:

$$\left|a_{i,i-1}\right| < \frac{a_{ii}+a_{i-1,i-1}}{2} \quad \text{για } i=2,\ldots,n \tag{4.4.4}$$

(β) Η κύρια διαγώνιος υπερέχει της δευτερεύουσας: $$\sum_{i=2}^{n}\left|a_{i,i-1}\right| \le trace(A) \tag{4.4.5}$$

(γ) Το μεγαλύτερο κατ' απόλυτη τιμή στοιχείο του A εμφανίζεται στη διαγώνιο του A. Διαφορετικά: το A δεν περιέχει στοιχεία μεγαλύτερα κατ' απόλυτη τιμή από το μέγιστο (θετικό) στοιχείο της διαγωνίου. [το αντίστροφο δεν ισχύει!]:

$$max(diag(A)) = |max(abs(A))| \tag{4.4.6}$$

Απόδειξη: α) Διαχωρίζουμε το κύριο υπομητρώο $\underline{A}_i$ ($1\le i\le n$) του A σε μπλοκ ως εξής:

$$\underline{A}_i = \begin{bmatrix} \underline{A}_{i-1} & x_{(i-1)\times 1} \\ x_{1\times(i-1)} & a_{ii} \end{bmatrix}$$

Το $\underline{A}_i$ ως γνωστόν είναι θ.ο. Θεωρούμε το διάνυσμα $\boldsymbol{\mu}_k=[e_k\ ;\ -1]\neq 0$, με $e_k\in R^{i-1}$, για κάποιο $1\le k\le i-1$. Τότε θα ισχύει $\boldsymbol{\mu}_k{}^T\underline{A}_i\,\boldsymbol{\mu}_k>0$, δηλ.:

$$\boldsymbol{\mu}_k{}^T\underline{A}_i\,\boldsymbol{\mu}_k =$$

$$= \begin{bmatrix} -e_k^T & 1 \end{bmatrix} \begin{bmatrix} \underline{A}_{i-1} & x_{(i-1)\times 1} \\ x_{1\times(i-1)} & a_{i\times 1} \end{bmatrix} \begin{bmatrix} -e_k \\ 1 \end{bmatrix} = \begin{bmatrix} e_k^T\underline{A}_{i-1} - x^T\ , & e_k^T x - a_{ii} \end{bmatrix} \begin{bmatrix} -e_k \\ 1 \end{bmatrix} =$$

$$= e_k^T\underline{A}_{i-1}e_k - 2x^Te_k + a_{ii} = a_{kk} - 2a_{ik} + a_{ii} > 0 \Rightarrow \tag{4.4.7}$$

$$\Rightarrow a_{ik} < \frac{a_{kk}+a_{ii}}{2}$$

Επαναλαμβάνοντας το ίδιο και για το διάνυσμα $\boldsymbol{\varrho}_k=[e_k\ ;\ 1]\neq 0$, ($1\le k\le i$-1), λαμβάνουμε τελικά:

$$a_{kk} + 2a_{ik} + a_{ii} > 0 \Rightarrow a_{ik} > -\frac{a_{kk}+a_{ii}}{2} \tag{4.4.8}$$

Από τις (4.4.7) και (4.4.8) προκύπτει:

$$\left|a_{ik}\right| \le \frac{a_{kk}+a_{ii}}{2} \tag{4.4.9}$$

Θέτοντας $k=i$-1, παίρνουμε το ζητούμενο:

$$\left|a_{i,i-1}\right| \le \frac{a_{i-1,i-1}+a_{ii}}{2}$$

■

(β) Χρησιμοποιούμε το αποτέλεσμα του (α). Η (4.4.4) δίνει:

$$\sum_{i=2}^{n}\left|a_{i,i-1}\right| \le \frac{1}{2}\sum_{i=2}^{n}\left(a_{i-1,i-1}+a_{ii}\right) = \frac{1}{2}a_{11} + \sum_{i=2}^{n-1}a_{ii} + \frac{1}{2}a_{nn} \le trace(A)$$ ■

(γ) Έστω A θ.ο. Χρησιμοποιούμε επαγωγή ως προς τα κύρια υπομητρώα του $\underline{A}_i$ του A. Για $\mu=2$, είναι $\underline{A}_2=[a\ b;\ b\ d]$ που είναι θ.ο., άρα ισχύει $a,b>0$ και $\Delta_2>0$. Συνεπώς $ad>b^2 \Rightarrow max(\text{diag}(A))= max(a,d)>|b| \Rightarrow max(\text{diag}(A))=|max(abs(A))|$. Άρα ισχύει.
Δεχόμαστε τώρα ότι το ζητούμενο ισχύει για $\mu=i$, δηλαδή ότι για το $\underline{A}_i$ είναι: $max(\text{diag}(\underline{A}_i))= |max(abs(\underline{A}_i))|=d_m=a_{mm}$, για κάποιο m με $1\le m\le k$. Αρκεί να δείξουμε ότι $|max(abs(\underline{A}_{i+1}))|=max(d_m,d_i)$. Η (4.4.9) δίνει:

$$|a_{ik}| \le (d_i+d_k)/2 \Rightarrow$$
$$2a_{ik}\le d_i+d_k \le d_i+d_m \le 2max(d_i+d_m) \Rightarrow$$
$$\Rightarrow |a_{ik}| \le max(d_m,d_i),\ k=1,2,\ldots,i\text{-}1 \Rightarrow |max(abs(\underline{A}_{i+1}))|=max(d_m,d_i) \qquad \blacksquare$$

Πόρισμα 4.4.2
Αν ένα συμμετρικό μητρώο δεν ικανοποιεί μια από τις συνθήκες (4.4.4)-(4.4.6), τότε *δεν είναι θ.ο.* Για παράδειγμα, το μητρώο $G=[4\ 1\ 1.5;\ 1\ 3\ 5\ ;\ 1.5\ 5\ 3]$ δεν είναι θ.ο., αφού το $max(G)=5$ βρίσκεται εκτός διαγωνίου. Το ίδιο συμπέρασμα προκύπτει και από τη διαπίστωση: $5>(3+3)/2=3$ που αντίκειται στη σχέση (4.4.4).

4.4.2 Θετικά Ορισμένα Ερμιτιανά Μητρώα

Ενδιαφέρον παρουσιάζει η επέκταση της έννοιας του θ.ο. μητρώου στο χώρο C^n με χρήση Ερμιτιανών μητρώων.

Ορισμός 4.4.1 – Θετικά Ορισμένα Ερμιτιανά Μητρώα

Ο ορισμός του θ.ο. μητρώου γενικεύεται και στα ***ερμιτιανά μητρώα****: ένα μητρώο μιγαδικών A καλείται θετικά ορισμένο (θ.η.ο.), αν για κάθε $\boldsymbol{x}\in C^n\text{-}\{0\}$ είναι $x^*Ax>0$ (αντίστοιχα $x^*Ax\ge 0$), όπου $\boldsymbol{x}^*$ είναι ο συζυγής ανάστροφος του $\boldsymbol{x}$.*

Παρατήρηση 3.1: Ο Ορισμός 4.4.1 έχει έννοια, αφού το $\boldsymbol{x}^*A\boldsymbol{x}$ είναι πάντα πραγματικό όταν το A είναι Ερμιτιανό (βλ. Παρ. 1). Αυτό δείχνεται από τις παρακάτω σχέσεις αφού ληφθεί υπ' όψη ότι τα διαγώνια στοιχεία ενός Ερμιτιανού μητρώου είναι πραγματικά:

$$x^*Ax=\sum_{i=1}^{n}a_{ii}\overline{x}_ix_i+\sum_{i\ne j}^{n}a_{ij}\overline{x}_ix_j=\sum_{i=1}^{n}a_{ii}|x_i|^2+\sum_{i<j}^{n}a_{ij}\overline{x}_ix_j+\sum_{i>j}^{n}a_{ij}\overline{x}_ix_j=$$

$$=\sum_{i=1}^{n}a_{ii}|x_i|^2+\sum_{i<j}^{n}a_{ij}\overline{x}_ix_j+\sum_{i>j}^{n}\overline{a}_{ji}\overline{x}_ix_j=\sum_{i=1}^{n}a_{ii}|x_i|^2+\sum_{i<j}^{n}a_{ij}\overline{x}_ix_j+\sum_{i<j}^{n}\overline{a}_{ij}\overline{x}_jx_i=$$

$$\sum_{i=1}^{n}a_{ii}|x_i|^2+\sum_{i<j}^{n}a_{ij}\overline{x}_ix_j+\overline{\sum_{i<j}^{n}a_{ij}x_j\overline{x}_i}=\sum_{i=1}^{n}a_{ii}|x_i|^2+\sum_{i<j}^{n}a_{ij}\overline{x}_ix_j+\overline{\sum_{i<j}^{n}a_{ij}\overline{x}_ix_j}\in \mathrm{R}$$

Για τα θ.ο. Ερμιτιανά μητρώα ισχύουν όλες οι προτάσεις και κριτήρια που αναφέρθηκαν πιο πάνω. Ενδεικτικά:

- Ισχύει η ιδιότητα των θετικών διαγώνιων στοιχείων: αν A Ερμιτιανό και επιλέξουμε $\boldsymbol{x}=\boldsymbol{e}_i$, με $\boldsymbol{e}_i$ στοιχείο της τυπικής βάσης $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n\}$ του χώρου C^n, τότε: $\boldsymbol{e}_i^*A\boldsymbol{e}_i=\boldsymbol{e}_i^{\mathrm{T}}A\boldsymbol{e}_i=a_{ii}>0$.
- Μπορούμε να δείξουμε ότι τα κύρια υπομητρώα τους έχουν πραγματικές ορίζουσες.
- Το κεντρικό κριτήριο για να είναι ένα Ερμιτιανό μητρώο θ.ο. παραμένει: *τα κύρια υπομητρώα του έχουν θετικές ορίζουσες.*
- Το σημαντικό κριτήριο των ιδιοτιμών ισχύει κι εδώ: αν όλες οι ιδιοτιμές ενός Ερμιατιανού είναι θετικές τότε είναι θ.ο. Και αντιστρόφως.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

♦ Παράδειγμα 4.2.1

Το συμμετρικό μητρώο C=[2 1 1; 1 3 -2; 1 -2 3] έχει 2 ιδιοτιμές θετικές (3 και 5) και μια μηδενική. Άρα είναι θ.η.ο. Στο ίδιο συμπέρασμα οδηγεί και η διαπίστωση ότι υπάρχουν δύο θετικοί οδηγοί (1 και 5/2) και ένας μηδενικός. Και ακόμα, ότι οι δύο αριστερές ορίζουσες είναι θετικές, ενώ η τρίτη 0. □

♦ Παράδειγμα 4.2.2

Το μητρώο D= [10 -20 0; -20 13 0 ; 0 0 36] είναι συμμετρικό και έχει θετικά διαγώνια στοιχεία. Όμως δεν είναι θ.ο. αφού det(D(1:2, 1:2))=-270. Συνεπώς τουλάχιστον μία από τις ιδιοτιμές του θα είναι αρνητική. Πράγματι, με την εντολή eig(D) παίρνουμε δύο θετικές ιδιοτιμές και μια αρνητική:

```
>> V=eig(D)
V =
   -8.5562
   31.5562
   36.0000
```

□

♦ Παράδειγμα 4.2.3

Το πιο κάτω τρισδιαγώνιο μητρώο A είναι θ.ο. Στο συμπέρασμα οδηγεί κατ' ευθείαν η διαπίστωση ότι έχει αδκ. και θετικά διαγώνια στοιχεία:

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

□

♦ [Matlab] Παράδειγμα 4.2.4 – Δημιουργία ενός θ.ο. μητρώου

Θα δημιουργήσουμε ένα α.δ.κ. μητρώο A με θετικά διαγώνια στοιχεία, δηλ. ένα θ.ο. μητρώο. Η εντολή A=rand(n ,n) δημιουργεί ένα nxn μητρώο με τιμές που κατανέμονται ομοιόμορφα στο (0, 1). Για να εξασφαλίσουμε α.δ.κ., προσθέσουμε στα διαγώνια στοιχεία το n-1. Δίνουμε λοιπόν την εντολή:

```
( rand(n, n) + (n - 1)*eye(n) )
```

Οι εντολές n=4; A = rand(n, n) + (n - 1)*eye(n); δίνουν:

```
A =
   3.8180   0.3412   0.8385   0.5466
   0.6602   3.5341   0.5681   0.4449
   0.3420   0.7271   3.3704   0.6946
   0.2897   0.3093   0.7027   3.6213
```

♦ [Matlab] Παράδειγμα 4.2.5 – Έλεγχος για θετικά ορισμένο μητρώο

Η ακόλουθη function ελέγχει στο Matlab αν ένα μητρώο είναι θ.ο. με βάση το πρόσημο των οριζουσών:

```
function [k] = is_pos_def (A)
[m,n] = size(A);
if m~=n,
    return;
end;
value = true;
for i=1:length(A)
  if ( det( A(1:i, 1:i) ) <= 0 )
    value = false;
    break;
  end;
end;
k=value ; % 0 if false, 1 if true
```

□

4.4.3 Συμπεράσματα για τη Διασπασιμότητα

Στην Π.4.2.3 είδαμε ότι αν ένα μητρώο είναι θ.ο. τότε και όλα τα κύρια υπομητρώα του είναι θ.ο., άρα και αντιστρεπτά, άρα σύμφωνα με το Θ4.2.2 διασπάται (άμεσα) στη μορφή LU. Συνεπώς όλα τα θ.ο. μητρώα επιδέχονται διάσπασης LU: A=LU.
Ομοίως από την Π4.3.1. όλα τα μητρώα με α.δ.κ. κατά γραμμές ή στήλες έχουν κύρια υπομητρώα τους με α.δ.κ. κατά γραμμές ή στήλες και σύμφωνα με το Θ. είναι αντιστρεπτά. Άρα και αυτά σύμφωνα με το Θ4.2.2 διασπώνται στη μορφή LU.
Συνεπώς οποιοδήποτε σύστημα με μητρώο συντελεστών αδκ. ή θ.ο. λύνεται με εφαρμογή εμπρός και πίσω αντικατάστασης. Συνοψίζοντας:

*Αν ο A έχει α.δ.κ. ή θ.ο., τότε η μέθοδος Gauss για κάθε σύστημα A**x**=**b** μπορεί να εφαρμοσθεί χωρίς ανταλλαγές γραμμών δηλαδή υπάρχουν ένας άνω τριγωνικός πίνακας U και ένας κάτω τριγωνικός πίνακας L ώστε A=LU.*

Ειδικότερα όμως, αν ένα μητρώο A είναι θετικά ορισμένο και συμμετρικό καταλήγουμε σε απλούστερη μορφή διάσπασης.

4.5 Διάσπαση *Choleski*

Η διάσπαση Cholesky αναφέρεται στην ειδική διάσπαση που χαρακτηρίζει τα συμμετρικά και θετικά ορισμένα μητρώα (θ.ο.). Η διάσπαση αυτή είναι η δεύτερη σημαντική παραγοντοποίηση ενός συμμετρικού (ή Ερμιτιανού) μητρώου μετά από αυτήν της διαγωνοποίησης $A=QΛQ^T$. Βασίζεται στους οδηγούς, ενώ αυτή της διαγωνοποίησης στις ιδιοτιμές. Συγκεκριμένα, για τα συμμετρικά και θ.ο. μητρώα, ισχύει η εξής ειδική μορφή διάσπασης:

$$\boldsymbol{A=LL^T} \text{ ή: } \boldsymbol{A=U^TU} \tag{4.5.1}$$

όπου $\boldsymbol{L}$ ($\boldsymbol{U^T}$) είναι κάτω (άνω) τριγωνικό. Το L δεν περιέχει απαραίτητα μονάδες στη διαγώνιο και προσδιορίζεται με μοναδικό τρόπο από τα στοιχεία του A.
Μια απόδειξη της (4.5.1) βασίζεται στη γνωστή μορφή (4.1.6) της διάσπασης LU:

$$A = L\,D\,U \tag{4.5.2}$$

όπου D διαγώνιο μητρώο με τους οδηγούς. Αν A είναι θ.ο., θα έχει θετικούς οδηγούς και θα είναι αντιστρέψιμο, δηλ. θα δέχεται τη διάσπαση (4.5.2). Αφού το A συμμετρικό θα είναι και $A^T=(LDU)^T=U^TD^TL^T=U^T DL^T=A=L\,DU$, απ' όπου συμπεραίνουμε ότι η $U^T DL^T=LDU$ ισχύει όταν $U^T=L$.
Μπορούμε τώρα να εκφράσουμε το D ως $D=\sqrt{D}\sqrt{D}$ (θετικοί οδηγοί), οπότε η (4.5.2) γράφεται:

$$A = L\sqrt{D}\sqrt{D}L^T = L\sqrt{D}(L\sqrt{D})^T \tag{4.5.3}$$

Δηλ. το L παράγεται άμεσα από τον L με πολλαπλασιασμό κάθε στήλης του με τις τετραγωνικές ρίζες των οδηγών που είναι θετικοί. Το L περιέχει θετικά στοιχεία στη διαγώνιο. Δηλ. οι τετραγωνικές ρίζες στους υπολογισμούς πρέπει να λαμβάνονται με το πρόσημο «+». (Άσκηση).

Τυπικές αποδείξεις για την ισχύ της διάσπασης *Cholesky* δίνονται στα δύο επόμενα Θεωρήματα. Το πρώτο υπαγορεύει τον τρόπο υπολογισμού του L, ενώ το δεύτερο παρουσιάζει τις ιδιότητες και τη δομή της διάσπασης.

Θεώρημα 4.5.1 [Διάσπαση *Cholesky*]

Ένα συμμετρικό μητρώο A μπορεί να τεθεί στη μορφή $A=LL^T$, όπου L κάτω τριγωνικό, με $l_{ii}>0$, $i=1,\ldots,n$, αν και μόνο αν είναι θετικά ορισμένο.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

Απόδειξη Δείχνουμε το ζητούμενο κατασκευαστικά. Έστω A θ.ο. μητρώο. Αρχικά θέτουμε $A=LL^T$ και λαμβάνουμε:

$$a_{ij}=\sum_{k=1}^{i} l_{ik}l_{jk} = \sum_{k=1}^{i-1} l_{ik}l_{jk} + l_{ii}l_{ji}\ , j=1,\ldots,n \tag{4.5.4}$$

Υπολογίζουμε στην αρχή τα διαγώνια στοιχεία του L από την (4.5.4). Βρίσκουμε $l_{11}=\sqrt{a_{11}}$, αφού για τα θ.ο. μητρώα ισχύει $a_{11}>0$, και

$$l^2_{ii}=a_{ii} - \sum_{k=1}^{i-1} l^2_{\text{ik}}\ , i=2,\ldots,n \tag{4.5.5}$$

Μπορούμε τώρα να δείξουμε με επαγωγή ως προς i, ότι $a_{ii} - \sum_{k=1}^{i-1} l^2_{\text{ik}} >0$, $i=1,\ldots,n,$ με βάση το ότι το A είναι θ.ο. και συμμετρικό. Η απόδειξη παραλείπεται εδώ αφού μπορεί να συμπληρωθεί από αυτήν του Θ4.5.2. Συνεχίζοντας λοιπόν, υπολογίζουμε:

$$l_{ii} = \sqrt{a_{\text{ij}} - \sum_{k=1}^{i-1} l^2_{\text{ik}}}\ ,\ i=1,\ldots,n \tag{4.5.6}$$

Στη συνέχεια υπολογίζουμε διαδοχικά τα υπόλοιπα στοιχεία των στηλών του L και βρίσκουμε:

$$l_{ji} = \frac{1}{l_{ii}}\left(\sum_{k=1}^{n-1} l_{\text{ik}}u_{jk}\right), j=i+1,\ldots,n \tag{4.5.7}$$

Αντίστροφο: Δείχθηκε ήδη στην Π(4.4.5(α)). ■

Λαμβάνοντας τώρα υπόψη την πορεία υπολογισμών του Θ4.5.1, και ιδιαίτερα τις (4.5.5)-(4.5.7), ο αλγόριθμος για τον υπολογισμό του παράγοντα *Cholesky* για θ.ο. συμμετρικά μητρώα διατυπώνεται ως εξής:

Αλγόριθμος 4.5.1: Διάσπαση Cholesky

Input: *Θετικά ορισμένο και συμμετρικό μητρώο* $A(n\times n)$, *μέγεθος* n
Output: *Κάτω τριγωνικό* L $(n\times n)$, *τέτοιο ώστε* $LL^T=A$
begin
$l_{11}=\sqrt{a_{11}}$;
for i:=2:n $l_{i1} := a_{i1}/a_{11}$; % υπολογισμός πρώτης στήλης
if $n>2$ **then** % υπολογισμός στήλης j, $j=2,\ldots,n-1$
{**for** j:=2:n-1
{ $l_{jj}=\sqrt{a_{jj} - \sum_{k=1}^{j-1} l^2_{jk}}$;
for i :=j+1 **to** n **do**
$l_{\text{ij}}=\left(a_{\text{ij}} - \sum_{k=1}^{j-1} l_{ik}a_{jk}\right)/\,l_{\text{jj}}$;
}
};
$l_{nn} = \sqrt{a_{nn} - \sum_{k=1}^{n-1} l^2_{nk}}$; % υπολογισμός τελευταίου διαγωνίου στοιχείου
write(L);
end ***Διάσπαση Cholesky***

Για την λύση του συστήματος $A\boldsymbol{x}=\boldsymbol{b}$, η μέθοδος *Cholesky* αρχικά υπολογίζει τον παράγοντα διάσπασης L με τις (4.5.5)-(4.5.7). Στη συνέχεια επιλύει εύκολα τα συστήματα $L^T\boldsymbol{y}=\boldsymbol{b}$ και $L^T\boldsymbol{x}=\boldsymbol{y}$ με εμπρός και πίσω αντικατάσταση αντίστοιχα.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

Ανάλυση-Υπολογιστικός χρόνος: Εύκολα διαπιστώνουμε ότι η μέθοδος *Cholesky* απαιτεί για την επίλυση ενός συστήματος $n\times n$ τους εξής αριθμούς πράξεων:

- n τετραγωνικές ρίζες
- $1/6(n^3+9n^2+2n)$ πολλαπλασιασμούς και
- $1/6(n^3+6n^2+7n)$ προσθέσεις και αφαιρέσεις

Παρατηρούμε ότι απαιτούνται περίπου οι μισοί υπολογισμοί (πολλαπλασιασμοί και διαιρέσεις) από την απαλοιφή *Gauss*.

Θεώρημα 4.5.2 [Διάσπαση *Cholesky*]

Αν A είναι συμμετρικό και θετικά ορισμένο, τότε για κάθε βήμα k=1,…,(στην απαλοιφή Gauss χωρίς οδήγηση) υπάρχει κάτω τριγωνικό L_k τέτοιο ώστε:
(i) Όλα τα διαγώνια στοιχεία του L_k είναι θετικά.
(ii) $\underline{A}_k= L_kL_k^T$, όπου L_k κάτω τριγωνικό. Συνεπώς για k=n θα είναι: $A=LL^T$,
(iii) Θα ισχύει $(\underline{L}_{k+1})_k= L_k$. *(Οι πρώτες k γραμμές και στήλες του L_{k+1}, $k\le n-1$, είναι το L_k).*

Απόδειξη Είναι ανάλογη της Απόδειξης του Θ4.4.2, και γίνεται με επαγωγή και διαχωρισμό του $\underline{A}_{k+1}$.

Είμαστε τώρα σε θέση να διατυπώσουμε ένα νέο κριτήριο για τα θ.ο. μητρώα, ισοδύναμο με τις προτάσεις του Θ.4.4.1. Το κριτήριο αυτό επεκτείνει το συμπέρασμα του Θ.4.5.1.

Πρόταση 4.5.1

Ένα μητρώο A είναι συμμετρικό και θ.ο. τότε και μόνο τότε, όταν υπάρχει μητρώο R $m\times n$ με $m\le n$ ανεξάρτητες στήλες, τέτοιο ώστε $A = R^TR$.

Απόδειξη: Αν υπάρχει $R\in \mathbb{R}^{m\times n}$ ώστε $A=R^TR$, τότε από Π(4.4.5(β)) το A είναι θ.ο.

Αντιστρόφως, αν το A είναι θ.ο. και συμμετρικό (δηλ. ικανοποιεί τις συνθήκες (i)-(iv) του Θ4.4.1, τότε μπορεί να εκφρασθεί στη μορφή R^TR (τουλάχιστον)με τους εξής τρόπους:

(i) Μπορεί να διασπασθεί στη μορφή (4.5.3): $A = L\sqrt{D}\sqrt{D}L^T$, απ' όπου: $R=\sqrt{D}L^T$

(ii) Το A ως συμμετρικό και με θετικές ιδιοτιμές, εκφράζεται σύμφωνα με το *φασματικό θεώρημα*: $A=Q\Lambda Q^T$, όπου Q ορθογώνιο μητρώο με τα ιδιοδιανύσματα. Ισοδύναμα είναι $A = Q\sqrt{\Lambda}\sqrt{\Lambda}Q^T = (\sqrt{\Lambda}Q^T)(\sqrt{\Lambda}Q^T)$, απ' όπου: $R=\sqrt{\Lambda}Q^T$.

(iii) Υπάρχουν και άλλες επιλογές για το R, ακόμα και μη τετραγωνικές. Επίσης, αν R ($m\times n$) μια επιλογή, και θεωρήσουμε οποιοδήποτε μητρώο Q ($n\times m$) με ορθογώνιες στήλες, τότε το QR θα είναι μια ακόμη επιλογή: $(QR)^T(QR)=R^TQ^TQR=R^TIR=A$. ■

Παρατήρηση 4.5.1 Λόγω συσσώρευσης του σφάλματος στρογγύλευσης, ειδικά στον υπολογισμό της διαγωνίου του L, όπου υπολογίζονται τετραγωνικές ρίζες, μπορεί να εφαρμοστεί *οδήγηση κατά την κύρια διαγώνιο.* Δηλαδή σε κάθε βήμα του αλγορίθμου *Choleski* ως οδηγός επιλέγεται το *μεγαλύτερο κατ' απόλυτο τιμή στοιχείο της πρώτης διαγωνίου.* Ακολουθεί συμμετρική εναλλαγή γραμμών και στηλών και αποθήκευση του διανύσματος μετάθεσης. Έτσι, το μητρώο A πολλαπλασιάζεται αριστερά και δεξιά με ένα μητρώο μετάθεσης P. Η διάσπαση διατυπώνεται ως εξής:

$$PAP^T=LL^T \text{ ή: } PAP^T=U^TU \tag{4.5.8}$$

Στο Παρ.4.5.3 δίνεται ένα παράδειγμα και στο Παρ.4.5.5 μια υλοποίηση για μια τέτοια οδήγηση. Η οδήγηση κατά τη διαγώνιο πλεονεκτεί για τους λόγους που υπαγορεύουν τα κριτήρια της Π4.4.6. Ακόμα και όταν δεν παρατηρείται α.δ.κ., ή και ταινιακότητα, η ιδιότητα της συμμετρίας και του θ.ο. μητρώου, «επιβάλλουν» συγκέντρωση μεγάλων σχετικά αριθμών στη διαγώνιο. Και εφ' όσον η διαίρεση με μικρούς οδηγούς είναι επισφαλής για τους γνωστούς λόγους, μια τέτοια οδήγηση αυτή καθίσταται ενδιαφέρουσα και ωφέλιμη στις πρακτικές εφαρμογές…

Παρατήρηση 4.5.2 Η διάσπαση *Choleski* ισχύει και για τα *Ερμιτιανά θ.ο. μητρώα.* (βλ. Παρ.Π1.1). Η παραγοντοποίηση LU διατυπώνεται τώρα από τη σχέση $A=LL^*$, ή από την $A=LDL^*$: στη θέση του ανάστροφου υπάρχει ο *συζυγής ανάστροφος* L^*. Και στην περίπτωση αυτή έχουμε ομοίως θετικούς οδηγούς, οπότε, κατ' αναλογία με τις (4.5.1) και (4.5.2), ισχύει η διάσπαση:

Αν A Ερμιτιανό θ.ο. μητρώο: $A=LL^*$ ή $A=U^*U$,

ή: $\mathbf{A} = \mathbf{L}\sqrt{\mathbf{D}}\left(\mathbf{L}\sqrt{\mathbf{D}}\right)^*$

Ο L μπορεί να προσδιορισθεί απ' ευθείας από τον ίδιο αλγόριθμο με L^*αντί του L^T.

Παρατήρηση 4.5.3 Ο αλγόριθμος της διάσπασης *Choleski* μπορεί να προσαρμοστεί για τη διάσπαση και θ.η.ο. μητρώων Οι υπολογισμοί των διαγωνίων στοιχείων (μητρώο D στους τύπους (4.5.2) και (4.5.3)) προϋποθέτουν πάντα μη αρνητικούς οδηγούς. Οι μηδενικές γραμμές μπορούν να μεταφέρονται στο τέλος και να συνεχίζεται η διαδικασία με τις υπόλοιπες. Μια τέτοια υλοποίηση δίνεται στο Παρ.4.5.5, ενώ υποστηρίζεται από μαθηματικά εργαλεία και βιβλιοθήκες.

♦ Παράδειγμα 4.5.1 – Διάσπαση *Choleski*

Για το A=[3,-2,1;-2,2,0;1,0,2] θα εξετάσουμε αν είναι θ.ο. και στη συνέχεια θα υπολογίσουμε τη διάσπαση *Choleski*.

Απ. Το A είναι συμμετρικό. Εξετάζουμε αν είναι και θ.ο. Δεν έχει α.δ.κ., επομένως δεν μπορούμε να εφαρμόσουμε το ικανό κριτήριο (iv). Από τις υπόλοιπες συνθήκες επιλέγουμε την (i), δηλ. την οριοθέτηση των ιδιοτιμών με χρήση του θεωρήματος των κύκλων *Gerschgorin*. Λαμβάνουμε ότι οι ιδιοτιμές θα βρίσκονται στο διάστημα όπου αληθεύουν οι ανισότητες: $|\lambda-3|\le 3$, $|\lambda-2|\le 2$, $|\lambda-2|\le 1$. Επομένως οι ιδιοτιμές βρίσκονται στο διάστημα [0, 6]. Το A όμως είναι αντιστρέψιμο (det(A)=2) και συνεπώς το 0 δεν είναι ιδιοτιμή. Άρα $\lambda>0$ και επομένως το A είναι συμμετρικό με θετικές ιδιοτιμές, δηλ. θ.ο.
Μια δεύτερη απόδειξη λαμβάνεται με το κριτήριο των οριζουσών (ii). Είναι det(A(1:1,1:1))=3>0, det(A(1:2,1:2))= det(A(1:3,1:3))=2>0. Άρα το A είναι θ.ο.
Αφού το A είναι συμμετρικό και θ.ο., ισχύει η διάσπαση *Choleski* $C^TC=A$. Εφαρμόζοντας τώρα τον αλγόριθμο διάσπασης, υπολογίζουμε συστηματικά τους αγνώστους c_{ij} *κατά στήλες*. Συγκεκριμένα έχουμε:

$$A = C^T C = \begin{bmatrix} c_{11} & 0 & 0 \\ c_{21} & c_{22} & 0 \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \times \begin{bmatrix} c_{11} & c_{21} & c_{31} \\ 0 & c_{22} & c_{32} \\ 0 & 0 & c_{33} \end{bmatrix} \Rightarrow$$

$$\Rightarrow \begin{bmatrix} {c_{11}}^2 & c_{11}c_{21} & c_{11}c_{31} \\ c_{21}c_{11} & {c_{21}}^2 + {c_{22}}^2 & c_{21}c_{31} + c_{22}c_{32} \\ c_{31}c_{11} & c_{31}c_{21} + c_{32}c_{22} & {c_{31}}^2 + {c_{32}}^2 + {c_{33}}^2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{21} & a_{21} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \Rightarrow$$

$$\Rightarrow \begin{bmatrix} {c_{11}}^2 & c_{11}c_{21} & c_{11}c_{31} \\ c_{21}c_{11} & {c_{21}}^2 + {c_{22}}^2 & c_{21}c_{31} + c_{22}c_{33} \\ c_{31}c_{11} & c_{31}c_{21} + c_{32}c_{22} & {c_{31}}^2 + {c_{32}}^2 + {c_{33}}^2 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 1 \\ -2 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \begin{cases} \text{στήλη 1}: c_{11} = \sqrt{3} = 1.7321, \quad c_{21} = -2/\sqrt{3} = \text{-}1.1547, \quad c_{31} = 1/\sqrt{3} = 0.5774 \\ \text{στήλη 2}: c_{22} = \sqrt{2 - {c_{21}}^2} = 0.8165, \quad c_{32} = \left(0 - c_{31}c_{21}\right)/\, c_{22} = 0.5774 * 1.1547/0.8165 = 0.8165 \\ \text{στήλη 3}: c_{33} = \sqrt{2 - 0.5774^2 - 0.8165^2} = \sqrt{0.9999} = 1 \end{cases}$$

Επαληθεύουμε στο Matlab:

```
>> C=chol(A)'
```

```
C =
   1.7321        0        0
  -1.1547   0.8165        0
  -0.5774  -0.8165        1
```

□

♦ Παράδειγμα 4.5.2

(α) Το μητρώο A=[6 -3 2; -3 5 1; 2 1 4] έχει θετικά διαγώνια στοιχεία και α.δ.κ. Άρα είναι θ.ο. Η διάσπαση Choleski είναι εφαρμόσιμη και υπολογίζουμε το L όπως στο Παράδειγμα 4.5.1. Επαληθεύοντας στο Matlab:

```
>> chol(A)
ans =
   2.4495   -1.2247    0.8165
        0    1.8708    1.0690
        0         0    1.4800
```

(β) Το μητρώο B=[6 -3 2; -3 -5 1; 2 1 4] είναι μεν α.δ.κ., αλλά ένα διαγώνιο στοιχείο (-5) είναι αρνητικό. Συνεπώς το B δεν είναι θ.ο. και η διάσπαση Choleski δεν είναι εφικτή. Επαληθεύουμε στο Matlab:

```
>> chol(B)
??? Error using ==> chol
Matrix must be positive definite.
```

□

♦ Παράδειγμα 4.5.3: Διάσπαση *Choleski* με οδήγηση κατά τη διαγώνιο

Δίνεται το σύστημα:

$$Ax = \begin{bmatrix} 6 & 2 & 1 \\ 2 & 8 & -2 \\ 1 & -2 & 16 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Θα εφαρμόζουμε διάσπαση *Choleski* για τη διάσπαση του μητρώου A με οδήγηση κατά μήκος της διαγωνίου του. Στη συνέχεια θα λύσουμε το σύστημα.

Το A είναι προφανώς θ.ο. Η εφαρμογή οδήγησης κατά μήκος της διαγωνίου σημαίνει εφαρμογή της μεθόδου *Choleski* με αναδιάταξη των στοιχείων της διαγωνίου σε φθίνουσα διάταξη (το μεγαλύτερο στοιχείο πρώτο). Ο λόγος είναι ότι στους εμπλεκόμενους υπολογισμούς θέλουμε να διαιρούμε πάντα με μεγάλους αριθμούς για την ελάττωση του σφάλματος στρογγύλευσης. Μια τέτοια οδήγηση προϋποθέτει συμμετρικές μεταθέσεις γραμμών και στηλών. Τα λαμβανόμενα μητρώα εξακολουθούν να είναι συμμετρικά και θ.ο. και μπορούμε να εφαρμόσουμε την *Choleski* σε οποιοδήποτε στάδιο. Στην προκείμενη περίπτωση λαμβάνουμε:

$$P_{13}AP_{13}x = \begin{bmatrix} 16 & -2 & 1 \\ -2 & 8 & 2 \\ 1 & 2 & 6 \end{bmatrix}$$

Δεν χρειάζεται άλλη συμμετρική μετάθεση, αφού η διαγώνιος έχει ήδη διαταχθεί. Αν R είναι ο παράγων *Choleski*, θα είναι $P_{13}AP_{13} = R^T R$, δηλ.:

$$P_{13}AP_{13} = \begin{bmatrix} 16 & -2 & 1 \\ -2 & 8 & 2 \\ 1 & 2 & 6 \end{bmatrix} = R^{\mathrm{T}}R = \begin{bmatrix} r_{11}^2 & r_{21}r_{11} & r_{31}r_{11} \\ r_{21}r_{11} & r_{21}^2 + r_{22}^2 & r_{31}r_{21} + r_{32}r_{22} \\ r_{31}r_{11} & r_{31}r_{21} + r_{32}r_{22} & r_{31}^2 + r_{32}^2 + r_{33}^2 \end{bmatrix}$$

Παίρνουμε κατά στήλη:

Στήλη 1: r_{11}=4, r_{21}=-1/2, r_{31}=1/4
Στήλη 2: $r_{22}=(8-1/4)^{1/2}=2.7839$, $r_{32}=17/(4*31^{1/2})=0.7633$
Στήλη 3: $r_{33}=(6-1/16-17^2/(16*31))^{1/2}=2.3141$

Λύνουμε τώρα το $R^{T}\boldsymbol{y}=P_{13}\boldsymbol{b}=(3, 2, 1)^{T}$ με εμπρός αντικατάσταση:

$$\boldsymbol{y} =(0.7500, 0.8531, 0.0697)^{T}$$

Στη συνέχεια λύνουμε το $R\boldsymbol{z}=\boldsymbol{y}$, θέτοντας $\boldsymbol{z}=P_{13}\boldsymbol{x}$ ($\Leftrightarrow \boldsymbol{x}=P_{13}\boldsymbol{z}$), με πίσω αντικατάσταση:

$$\boldsymbol{z} =(0.2229, 0.2982, 0.0301)^{T}$$

Τέλος, εναλλάσσουμε τους αγνώστους x_1 και x_3:

$$\boldsymbol{x} = P_{13}\boldsymbol{z} =(0.0301, 0.2982, 0.2229)^{T}$$

Μπορούμε τώρα να επαληθεύσουμε: $A\boldsymbol{x} = (1, 2, 3)^{T}$. □

♦ Παράδειγμα 4.5.4 – Παράγων *Choleski* στα μητρώα *Pascal*

Το μητρώο Pascal είναι ένα συμμετρικό μητρώο με «1» στην πρώτη γραμμή και στήλη και με τα υπόλοιπα στοιχεία να λαμβάνονται ως εξής: P(i,j)=P(i, j-1)+P(i-1,j). Για n=8 είναι:

P=pascal(8)=

1	1	1	1	1	1	1	1
1	2	3	4	5	6	7	8
1	3	6	10	15	21	28	36
1	4	10	20	35	56	84	120
1	5	15	35	70	126	210	330
1	6	21	56	126	252	462	792
1	7	28	84	210	462	924	1716
1	8	36	120	330	792	1716	3432

Παρατηρούμε ότι για την απαλοιφή όλοι οι οδηγοί είναι 1, επομένως τα μητρώα *Pascal* είναι θ.ο. Συνεπώς είναι δυνατή η διάσπαση *Choleski*. Το χαρακτηριστικό εδώ είναι ότι ο παράγων *Choleski* αποτελείται από ακεραίους. Η δευτερεύουσα διαγώνιος περιλαμβάνει τους ακεραίους $1,2,\ldots,n\text{-}1$:

chol(P)=

1	1	1	1	1	1	1	1
0	1	2	3	4	5	6	7
0	0	1	3	6	10	15	21
0	0	0	1	4	10	20	35
0	0	0	0	1	5	15	35
0	0	0	0	0	1	6	21
0	0	0	0	0	0	1	7
0	0	0	0	0	0	0	1

Οι ιδιοτιμές είναι θετικές (ως αναμένεται) και δίνονται από το φάσμα:

1.0e+003 * [0.00000022008515, 0.00000672021444, 0.00008373024586, 0.00051189155425, 0.00195353877535, 0.01194311553432, 0.14880477533637, 4.54369600825427]

Είναι φανερό ότι όταν το n τείνει στο άπειρο, όλες οι ιδιοτιμές εκτός της τελευταίας, τείνουν στο 0 με σταδιακά μειούμενη τάξη σύγκλισης. Η τελευταία φυσικά τείνει στο *trace(A)*. Είναι φανερό ότι η κατανομή των ιδιοτιμών αντικατοπτρίζει τη σταδιακή αύξηση της πληροφορίας κατά μήκος της κύριας διαγώνιου.

Τέλος, ο δείκτης κατάστασης είναι πολύ μεγάλος: $k(A)$=2.064517341533397e+007. Όταν το n τείνει στο άπειρο $k(A)\rightarrow\infty$. □

♦ [Matlab] Παράδειγμα 4.5.5– Διάσπαση *Cholesky* για θ.ο. και θ.η.ο μητρώα

Ο αλγόριθμος της συνάρτησης cholp(A, pivot) που ακολουθεί, εφαρμόζει διάσπαση *Choleski* με οδήγηση κατά μήκος της διαγωνίου, με συμμετρικές εναλλαγές γραμμών-στηλών (ενεργοποιείται εισάγοντας μια οποιασδήποτε τιμή≠0 στη μεταβλητή pivot). Ταυτόχρονα υλοποιείται και ο αλγόριθμος

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

χωρίς οδήγηση (δίνοντας pivot=0) που δίνει την κλασσική διάσπαση A=R'*R. Στην πρώτη περίπτωση, η διάσπαση εκφράζεται ως R'*R = P'*A*P, όπου P είναι μεταθετικό μητρώο το οποίο και υπολογίζεται.

Επιπλέον, ο κλασσικός αλγόριθμος *Cholesk*i μπορεί να προσαρμοστεί και σε θετικά ημιορισμένα μητρώα. Αυτό υλοποιείται εδώ. Για τον έλεγχο αν το μητρώο εισόδου είναι τουλάχιστον θετικά ημιορισμένο, εξετάζεται αν τα διαγώνια στοιχεία είναι μη αρνητικά πριν χρησιμοποιηθούν ως οδηγοί.

```
function [R, P, I] = cholp(A, pivot)
%CHOLP  Cholesky factorization with pivoting of a positively semidefinite
%        matrix.
%       [R, P] = CHOLP(A) returns R and a permutation matrix P such that
%       R'*R = P'*A*P.  Only the upper triangular part of A is used.
%       [R, P, I] = CHOLP(A) returns in addition the index I of the last
%       positive diagonal element of R.  The first I rows of R contain
%        the Cholesky factor of A.
%       [R, I] = CHOLP(A, 0) forces P = EYE(SIZE(A)), and therefore produces
%        the same output as R = CHOL(A) when A is positive definite (to
%        within roundoff).
%        This routine is based on the LINPACK routine CCHDC.  It works
%        for both real and complex matrices.
%
%        Reference:
%        G.H. Golub and C.F. Van Loan, Matrix Computations, Second
%        Edition, Johns Hopkins University Press, Baltimore, Maryland,
%        1989, sec. 4.2.9.

%Check the number of function's inputs
if nargin == 2, piv = pivot; else piv = 1; end

[n, n] = size(A);
pp = 1:n;
I = [];

for k = 1:n     %Computation of the columns
    if piv              %If piv=1.
       d = diag(A);
       [big, m] = max( d(k:n) );
       m = m+k-1;
    else
       big = A(k,k);  m = k;
    end
    if big <= 0, I = k-1; break, end    %The maximum value in the diagonal
                          %is negative.
%   Symmetric row/column permutations.
    if m ~= k
       j = 1:k-1;
       if j
          temp = A(j,k); A(j,k) = A(j,m); A(j,m) = temp;
       end
       temp = A(k,k); A(k,k) = A(m,m); A(m,m) = temp;
       A(k,m) = conj(A(k,m));       %complex conjugate
       j = k+1:m-1;
       if j
          temp = A(k,j)'; A(k,j) = A(j,m)' ; A(j,m) = temp;
       end
       j = m+1:n;
       if j
          temp = A(k,j); A(k,j) = A(m,j); A(m,j) = temp;
       end
       pp( [k m] ) = pp( [m k] );
    end

    A(k,k) = sqrt( A(k,k) );
    if k == n, break, end
    A(k, k+1:n) = A(k, k+1:n) / A(k,k);

%   For simplicity update the whole of the remaining submatrix (rather
%   than just the upper triangle).

    j = k+1:n;
    A(j,j) = A(j,j) - A(k,j)'*A(k,j);

end  % Main k loop
```

```
R = triu(A);            %The upper trianlge part of matrix
if isempty(I), I = n; end   %Check if matrix I is empty
if nargout == 2 & ~piv
   P = eye(n);
elseif nargout >= 2
   P = eye(n); P = P(:,pp);
end
```

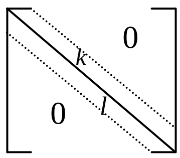
Στη συνέχεια η συνάρτηση cholp δοκιμάζεται στο Matlab με ένα θ.η.ο. και με ένα θ.ο. Ερμιτιανό μητρώο:

```
>>A= [2 1 1; 1 3 -2; 1 -2 3]     % πρώτο μητρώο A: συμμετρικό και θ.η.ο.
A =
   2   1   1
   1   3  -2
   1  -2   3

>> [R, P, I] = cholp(A)  % κλήση της cholp. Εφαρμογή αλγορίθμου Choleski. H κλήση R=chol(A)
            % της ρουτίνας chol του Matlab αποτυγχάνει!
R =
   1.7321    0.5774   -1.1547
        0    1.2910    1.2910
        0         0   -0.0000
P =                             % για να βρεθεί ο μεγαλύτερος οδηγός 3 εναλλάσσονται
    0    1    0                 % οι στήλες 1 και 2 και οι γραμμές 1 και 2
    1    0    0
    0    0    1

I = 2                           % η θέση του τελευταίου θετικού διαγώνιου στοιχείου στον R
>> R'*R                         % ισχύει η διάσπαση R'*R= P'*A*P
ans =
    3.0000    1.0000   -2.0000
    1.0000    2.0000    1.0000
   -2.0000    1.0000    3.0000

>> P'*A*P                       % επαλήθευση ότι R'*R= P'*A*P
ans =
    3    1   -2
    1    2    1
   -2    1    3

>> C=[1 1+i 2-3i 0; 1 2 4-5i 3 ; 0 4i 1 3 ; 2-3i 4 1 i] % Έστω το  μη συμμετρικό μιγαδικό μητρώο C.
C =
   1.0000              1.0000 + 1.0000i    2.0000 - 3.0000i         0
   1.0000              2.0000              4.0000 - 5.0000i         3.0000
        0              0 + 4.0000i              1.0000              3.0000
   2.0000 - 3.0000i    4.0000              1.0000              0 + 1.0000i

>>H =C'*C                       % Το μητρώο H είναι οπωσδήποτε ερμιτιανό και θ.ο.
H =
  15.0000              11.0000 +13.0000i   8.0000 - 5.0000i    0 + 2.0000i
  11.0000 -13.0000i    38.0000             11.0000 -19.0000i   6.0000 - 8.0000i
  8.0000 + 5.0000i     11.0000 +19.0000i   56.0000             15.0000 +16.0000i
  0 - 2.0000i          6.0000 + 8.0000i    15.0000 -16.0000i   19.0000

>> [R, P, I] = cholp(H)         % κλήση της cholp. Εφαρμογή αλγορίθμου Choleski.
R =
   7.4833    1.4699 + 2.5390i    2.0045 + 2.1381i    1.0690 + 0.6682i
        0    5.4215             -0.4381 - 1.1166i    1.4262 - 2.0784i
        0         0                  2.9953         -1.7585 - 1.1873i
        0         0                   0              1.5984
```

```
P =
    0   0   0   1
    0   1   0   0
    1   0   0   0
    0   0   1   0
I = 4
>> cond(H)        % δείκτης κατάστασης του H. Ο H δεν έχει καλή κατάσταση. Αναμένεται σφάλμα
                  % στους υπολογισμούς.
ans = 71.4701
>> error=norm(R'*R-P'*H*P)   % σφάλμα υπολογισμού
error = 3.7740e-015
```

4.6 Ταινιακά Μητρώα

Η μέθοδος διάσπασης *LU* απλουστεύεται σημαντικά για *ταινιακά μητρώα*. Ένα μητρώο *A* λέγεται *ταινιακό* όταν υπάρχουν ακέραιοι k, l, με $l<k$ και $l<n$, τέτοιοι ώστε $a_{ij}=0$ για $i+k<j$, $j+l>i$. Η μορφή ενός ταινιακού μητρώου δίνεται στο σχήμα (3.1).



Σχήμα 4.6.1 Μορφή ταινιακού μητρώου

Αν $k=l=0$, το A είναι διαγώνιο, ενώ αν $k=l=1$, είναι τριδιαγώνιο μητρώο. Εξετάζουμε εδώ τη μορφή διάσπασης για τριδιαγώνια μητρώα. Ένα τέτοιο μητρώο έχει τη μορφή:

$$A=\begin{bmatrix} a_{11} & a_{12} & 0 & \dots & 0 & 0 \\ a_{21} & a_{22} & a_{23} & \dots & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & a_{n-1,n} & a_{nn} \end{bmatrix} \tag{4.6.1}$$

Ας τονιστεί ότι ένα τυχαίο τρισδιαγώνιο μητρώο δεν είναι κατ' ανάγκη μη ιδιάζον. Π.χ. το μητρώο [1, 1 0 0; 2 2 1 0; 0 0 3 0; 0 0 4 5] είναι μη ιδιάζον αφού έχει ίδιες τις δύο πρώτες στήλες. Ένα κριτήριο αντιστρεψιμότητας είναι όλα τα στοιχεία των δευτερευουσών διαγωνίων να είναι μη μηδενικά: $a_{i+1,i}\neq 0$, $a_{i,i+1}\neq 0$. Τότε προφανώς οι γραμμές (στήλες) του A είναι γραμμικά ανεξάρτητες (καμία δεν προκύπτει ως γραμμικός συνδυασμός των υπολοίπων). Επίσης, για τα ταινιακά μητρώα ισχύουν τα ίδια κριτήρια παραγοντοποίησης με αυτά που ισχύουν για τα απλά μητρώα (όταν A είναι α.δ.κ. ή θ.ο.). Η διάκριση για τα ταινιακά μητρώα έγκειται στην πιο απλή τους διάσπαση.

Πρόταση 4.6.1

Για ένα τρισδιαγώνιο μητρώο διασπάσιμο στη μορφή *LU*, οι παράγοντες *L* και *U* έχουν τις εξής τελικές μορφές:

$$L=\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ l_{21} & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & l_{n,n-1} & 1 \end{pmatrix},\ U=\begin{pmatrix} u_{11} & a_{12} & 0 & \dots & 0 & 0 \\ 0 & u_{22} & a_{23} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & a_{n-1,n} \\ 0 & 0 & \dots & \dots & \dots & u_{nn} \end{pmatrix} \tag{4.6.2}$$

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

Απόδειξη : Όλοι οι πολλαπλασιαστές *Gauss* l_{ij} με $i>j+1$ είναι προφανώς 0. Αυτό εξηγεί τη μορφή του L. Τα 0 στις στήλες του U και η άνω διαγώνιος του U εξηγούνται από την απαίτηση: i-στήλη του $A = L \times$ (i στήλη του U).

Συνεπώς οι άγνωστοι είναι τώρα τα στοιχεία $l_{i,i-1}$ της κάτω διαγωνίου του L και τα u_{ij} της κύριας διαγωνίου του U. Για την διατύπωση του αλγορίθμου υπολογισμού των U και L ξεκινάμε από τις παρακάτω σχέσεις που προκύπτουν από την απαίτηση $A=LU$ και τις (4.6.2):

$$
\begin{aligned}
&a_{11} = u_{11}\\
&u_{i-1,j} = a_{i-1,i}\,, \text{ για } i\geq 2\\
&a_{ii} = l_{i,i-1}\, u_{i-1,i} + u_{ii}, \text{ για } i\geq 2\\
&a_{i+1,i} = l_{i+1,i}\, u_{ii}\,, \text{ για } i\geq 1
\end{aligned}
\tag{4.6.3}
$$

Λύνοντας τώρα το σύστημα (4.6.3), βρίσκουμε πρώτα τα u_{ii} από τη τρίτη εξίσωση, και κατόπιν τα $l_{i+1,i}$ από την τέταρτη. Ο αλγόριθμος παραγοντοποίησης διατυπώνεται τελικά:

Αλγόριθμος 4.6.1: Επίλυση Συστήματος με Τρισδιαγώνιο Μητρώο Συντελεστών

Input : *τριδιαγώνιο μητρώο* A ($n\times n$), n, *διάνυσμα* $\mathbf{b}$ ($n\times 1$)
Output: x: *λύση του συστήματος* $A\mathbf{x}=\mathbf{b}$
begin
  $u_{11} := a_{11}$;   % παραγοντοποίηση-υπολογισμός των u_{ij}, l_{ij}
  $l_{21} := a_{21}/u_{11}$;
  for $i := 2$ **to** $n-1$ **do**
    { $u_{i-1,i} := u_{i-1,i}$;
    $u_{ii} := a_{ii} - l_{i,i-1} * u_{i-1,i}$;
    $l_{i+1,i} := a_{i+1,i} / u_{ii}$ };
  $u_{n-1,i} := a_{n-1,i}$;
  $u_{nn} := a_{nn} - l_{n,n-1} * u_{n-1,i}$;
  $y_1 := b_1/l_{11}$;   % επίλυση του $L\mathbf{y}=\mathbf{b}$ με εμπρός αντικατάσταση
  for $i := 2$ **to** n **do** $y_i := b_1 - l_{i,i-1} * y_i$;
  $x_n := y_1$;   % επίλυση του $U\mathbf{x}=\mathbf{y}$ με πίσω αντικατάσταση
  for $i := n-1$ **step** -1 **to** 1 **do** $x_i := (y_i - u_{i,i+1} * x_{i+1}) / u_{ii}$;
  write(x)
end

Παρατήρηση 4.6.1 Είναι φανερό, ότι για την υλοποίηση του πιο πάνω αλγορίθμου μπορεί να επιτευχθεί εξοικονόμηση μνήμης αφού τα δεδομένα a_{ij}, l_{ij} και u_{ij} μπορούν να αποθηκευθούν στους πίνακες A[1..*n*, 3], L[1..2, 2] και U[1..*n*, 2]. Θα απαιτηθεί βέβαια κατάλληλη αναπροσαρμογή των δεικτών (Άσκηση 4.11). Ας αναφερθεί επίσης, ότι για τα θ.ο. συμμετρικά ταινιακά μητρώα ο πιο πάνω αλγόριθμος απλουστεύεται σημαντικά αφού καταλήγουμε σε διάσπαση *Choleski* (Άσκηση 4.9).

Τέλος, ανάλογα μπορεί να μελετηθεί το πρόβλημα της διάσπασης και για άλλα ταινιακά μητρώα (Άσκηση 4.10).

♦ Παράδειγμα 4.6.1 [Matlab]

Στο Matlab ορίζουμε το «1-4-1» συμμετρικό μητρώο A μεγέθους 4x4:

```
>> A=[4 1 0 0;1 4 1 0;0 1 4 1;0 0 1 4 ]

 A=
    4   1   0   0
    1   4   1   0
    0   1   4   1
    0   0   1   4
```

Το A είναι προφανώς α.δ.κ., επομένως επιδέχεται διάσπαση $A=LU$ ($P=I$). Αμέσως συνάγουμε (A τρισδιαγώνιος) ότι τα εκτός της πρώτης και δεύτερης διαγωνίου στοιχεία των U, L είναι 0. Επίσης, η άνω

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

διαγώνιος του U=diag([1,1,1]) και $U(1,1)=4$. Εύκολα διαπιστώνουμε τέλος, με τη μέθοδο *Grout* ή την απαλοιφή, ότι οι 3 πολλαπλασιαστές (κάτω διαγώνια στοιχεία του L) είναι αντίστροφοι των οδηγών:

$p_{21}=l_{21}=1/4$, $p_{32}=l_{32}=4/15$, $p_{43}=l_{43}=15/56$,
$u_{11}=4$, $u_{22}=15/4$, $u_{33}=56/15$, $u_{11}=209/56$

Επαληθεύουμε τώρα, υπολογίζοντας τα L,U,P:

```
>> [L,U,P]=lu(A)

L =
    1.0000    0         0         0
    0.2500    1.0000    0         0
    0         0.2667    1.0000    0
    0         0         0.2679    1.0000

U =
    4.0000    1.0000    0         0
    0         3.7500    1.0000    0
    0         0         3.7333    1.0000
    0         0         0          3.7321

P =
     1     0     0     0
     0     1     0     0
     0     0     1     0
     0     0     0     1
```

Εφαρμόζουμε διάσπαση Choleski:

```
>> U=chol(A)'

U =
    2.0000         0         0         0
    0.5000    1.9365         0         0
         0    0.5164    1.9322         0
         0         0    0.5175    1.9319
```

Όμως ο παράγων Choleski μπορεί να διαμορφωθεί και απ' ευθείας από την LU. Φυσικά θα πάρουμε το ίδιο αποτέλεσμα.

```
L*sqrt(diag(diag(U)))
ans =

    2.0000         0         0         0
    0.5000    1.9365         0         0
         0    0.5164    1.9322         0
         0         0    0.5175    1.9319
```

□

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

Ασκήσεις

4.1 Να δειχθεί ότι: (α) το αντίστροφο ενός άνω (κάτω) τριγωνικού μητρώου είναι άνω (κάτω) τριγωνικό (β) το γινόμενο *n* άνω (κάτω) τριγωνικών μητρώων είναι άνω (κάτω) τριγωνικό.

4.2 Να εφαρμοσθεί η μέθοδος παραγοντοποίησης *LU* για την επίλυση του συστήματος

{3x+y-z=0,-2x-6y+3z=2, 4x-2y+8z=1}

4.3 Με βάση τον αλγόριθμο *Grout*, δώστε ένα γενικό αλγόριθμο για την παραγοντοποίηση *PA*=*LU*, όπου *P* μεταθετικό μητρώο. Ο αλγόριθμος θα υπολογίζει τα μητρώα *P*, *L* και *U* (Υπόδειξη: να γίνει χρήση διανύσματος μετάθεσης). Κατόπιν, με βάση τον αλγόριθμο αυτόν, δώστε ένα αλγόριθμο solveLU(A, b) για την επίλυση ενός συστήματος $A\mathbf{x}=\mathbf{b}$.

4.4 Υπολογίστε τον αριθμό πράξεων που εκτελούνται για τη μέθοδο διάσπασης *LU*. Συγκρίνετέ τον με τον αντίστοιχο αριθμό της μεθόδου *Gauss* και αναφέρετε τα συμπεράσματα σας.

4.5 Εξετάστε αν το παρακάτω συμμετρικό μητρώο *A* είναι θετικά ορισμένο. Κατόπιν εφαρμόστε τον αλγόριθμο *Cholesky* για την παραγοντοποίησή του.

$$A=\begin{bmatrix} 6 & 3 & 1 & -1 \\ 3 & 9 & 4 & 0 \\ 0 & 1 & 4 & 2 \\ -1 & 0 & 2 & 4 \end{bmatrix}$$

4.6 Να παραγοντοποιηθεί με τη μέθοδο *Cholesky* το μητρώο

$$A=\begin{pmatrix} 1 & 3 & 0 & 0 \\ 3 & 1 & 4 & 0 \\ 0 & 4 & 3 & 2 \\ 0 & 0 & 2 & 6 \end{pmatrix}$$

Στη συνέχεια να λυθεί το σύστημα $A\mathbf{x}=\mathbf{b}$,όπου $\mathbf{b}=(-2, 6, 1, 3)^T$.

4.7 Να διασπασθεί το ταινιακό μητρώο *A* με εφαρμογή του αλγορίθμου της παρ. 4.1.1.

$$\mathrm{A}=\begin{pmatrix} 1 & 6 & 0 & 0 \\ 3 & 2 & -2 & 0 \\ 0 & 9 & 3 & 4 \\ 0 & 0 & 7 & 1 \end{pmatrix}$$

4.8 Να παραγοντοποιηθεί το συμμετρικό τρισδιαγώνιο μητρώο «2-8-2» για *n*=4 με διάσπαση *LU* και Choleski.

4.9 Ειδικά για τα συμμετρικά τρισδιαγώνια και θ.ο. μητρώα, διατυπώστε έναν αλγόριθμο παραγοντοποίησης σε μορφή LL^T (*Choleski*). Γράψτε αντίστοιχο κώδικα σε *Matlab* και κατόπιν επαληθεύστε τον για την επίλυση του συστήματος της άσκησης 4.5.

4.10 Μελετείστε τη διάσπαση *LU* για ένα πενταδιαγώνιο μητρώο (*1*=*k*=2). Διατυπώστε τον αντίστοιχο αλγόριθμο και υλοποιείστε τον σε Matlab.

4.11 Να αναδιατυπώσετε τον αλγόριθμο 4.6.1, χρησιμοποιώντας την ελάχιστη δυνατή μνήμη για την αποθήκευση των αρχικών και ενδιάμεσων δεδομένων. Στη συνέχεια να τον υλοποιήσετε στο *Matlab*.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών

4.7 Σύνοψη και Βιβλιογραφία

Σύνοψη

Τα πιο βασικά σημεία που είδαμε στο κεφάλαιο αυτό είναι:

- Όταν η απαλοιφή *Gauss* γίνεται χωρίς εναλλαγές γραμμών ένα αντιστρέψιμο μητρώο A διασπάται σε LU: $A=LU=LDU$. Το L περιέχει τους πολλαπλασιαστές και το διαγώνιο D τους οδηγούς. Τότε η επίλυση του $A\boldsymbol{x}=\boldsymbol{b}$ απλουστεύεται, με εφαρμογή εμπρός και πίσω αντικατάστασης.
- Τα L και U μπορούν να υπολογισθούν κατ' ευθείαν και συστηματικά χωρίς εφαρμογή απαλοιφής (Αλγόριθμος *Grout*).
- Όταν η απαλοιφή απαιτεί εναλλαγές γραμμών και P είναι το μεταθετικό μητρώο των εκτελούμενων εναλλαγών γραμμών, τότε $PA=LU$, ή $A=P^TLU$. Το L περιέχει τους πολλαπλασιαστές με διάταξη που καθορίζει το P.
- Η εφαρμογή οδήγησης επηρεάζει τους παράγοντες της διάσπασης.
- Όλα τα παραπάνω ισχύουν και για μητρώα μιγαδικών.
- Όλα τα θετικά ορισμένα συμμετρικά μητρώα και τα μητρώα με αυστηρή διαγώνια κυριαρχία επιδέχονται διασπάσεις διαφόρων ειδών.
- Για τα ταινιακά μητρώα ισχύει απλούστερη και πιο γρήγορη μορφή διάσπασης.
- Ένα συμμετρικό μητρώο είναι θετικά ορισμένο, αν και μόνον αν έχει θετικούς οδηγούς, ή θετικές ιδιοτιμές, ή όλες τις αριστερές υπο-ορίζουσες θετικές. Η αυστηρή διαγώνια κυριαρχία με θετικά διαγώνια στοιχεία αποτελεί μια ικανή συνθήκη.
- Ειδικά για τα θετικά ορισμένα και συμμετρικά μητρώα η διάσπαση απλουστεύεται: $A=L^TL$ (*Choleski*). Το L υπολογίζεται χωρίς εφαρμογή απαλοιφής *Gauss*, με τρόπο ανάλογο με τον αλγόριθμο *Grout*. Η μορφή $A=L^TL$ αποτελεί ένα πρόσθετο κριτήριο για να είναι ένα μητρώο θ.ο.
- Τα Ερμιτιανά θετικά ορισμένα μητρώα ορίζονται από την $x^*Ax>0$, $x\in C^n$. Ισχύουν και γι' αυτά ανάλογα βασικά κριτήρια και αρχές διάσπασης με τα μητρώα πραγματικών.

Βιβλιογραφία και Χρήσιμες Αναφορές

Για συμπληρωματικά στοιχεία στα θετικά ορισμένα μητρώα και μορφές διασπάσεων βλ. αναφορές [1]-[7] και [12]-[16].

Για θεμελιώσεις της (Αριθμητικής) Γραμμικής Άλγεβρας και του Λογισμού Μητρώων, βασική αναφορά είναι η [6], ενώ χρήσιμες είναι και οι [7], [9], [10] και [11]-[15]. Οι αναφορές [4] και [5] περιέχουν εφαρμογές για Μηχανικούς.

Η [8] περιέχει μια ενδιαφέρουσα απόδειξη για το *Κριτήριο Sylvester* για θ.ο. μητρώα.

Οι [18]-[21] αφορούν βοηθήματα για το *Matlab*. Ειδική έμφαση πρέπει να δοθεί στην [19], που αφορά πρότυπους κώδικες διδασκαλίας.

Μια πολύ καλή αναφορά είναι και η [12], κατάλληλη για φοιτητές θετικής κατεύθυνσης και Μηχανικούς που επιθυμούν να έχουν μια ενοποιημένη προσέγγιση στη Γραμμική Άλγεβρα, Μαθηματική Ανάλυση και στις Εφαρμογές.

[1] Χ. Αλεξόπουλος, «*Εισαγωγή στην Αριθμητική Ανάλυση και Περιβάλλοντα Υλοποίησης*», Πανεπιστημιακές Σημειώσεις, 2007.

[2] Χ. Αλεξόπουλος, «*Επιλεγμένες Ασκήσεις και Θέματα Αριθμητικής Ανάλυσης*», Σημειώσεις (*downloadable version*), Πάτρα 2010

[3] Μ. Βραχάτης, «*Αριθμητική Ανάλυση*», Εκδ. «Ελληνικά Γράμματα», 2002.

[4] J. H. Mathews: "*Numerical Methods for Computer Science, Engineering, and Mathematics*", Prentice-Hall Int, Editions.

[5] S. C. Chapra, Raymond P. Canale, "*Numerical Methods for Engineers*", Mc.GRAW-HILL Int. Editions, Second Edition.

[6] G. Strang: «*Εισαγωγή στη Γραμμική Άλγεβρα*»

[7] Γ. Δονάτος, Μ. Αδάμ, «*Γραμμική Άλγεβρα, Θεωρία και Εφαρμογές*», Gutenberg.

[8] George T. Gilbert "*Positive Definite Matrices and Sylvester's Criterion*", *The American Mathematical Monthly*, Vol. 98, No. 1 (Jan., 1991).

[9] O. Morris, «*Μια Εισαγωγή στη Γραμμική Άλγεβρα*».

[10] Θ. Παπαθεοδώρου «*Εισαγωγή στην Αριθμητική Ανάλυση*», Πανεπιστημιακές Σημειώσεις

[11] A. Kurosh, "*Cours d'Algebre Superieure*", Editions de Moscou

[12] John H. Hubbard, Barbara Burke Hubbard, «*Διανυσματικός Λογισμός, Γραμμική Άλγεβρα και Διαφορικές Μορφές*», Εκδόσεις Πανεπιστημίου Πατρών, 2006.

[13] Loud N. Trefethen, David Bau, "*Numerical Linear Algebra*", III.

[14] Gene H. GoluB, Charles F. Van Loan, "*Matrix Computations*", 3rd edition.

[15] Roger A. Horn and Charles R. Johnson, "*Matrix Analysis*".

[16] *Journal of Mathematical Analysis and Applications*, #237 (1999).

[17] S. Conte, Carl de Boor, *"Elementary Numerical Analysis: An Algorithmic Approach"*, McGraw-Hill,1980

[18] John Mathews, Kurtis D, Fink, "*Numerical Methods Using MATLAB*", Fourth Edition, Aug 2006.

[19] Companion software to accompany the book "*Numerical Methods Using MATLAB, 4e*" by J. Mathews: http://www.mathworks.com/matlabcentral/fileexchange/2181-numerical-methods-using-matlab-2e

[20] Χ. Αλεξόπουλος, «*Εισαγωγή στο Matlab*», Πανεπιστημιακό Βοήθημα, 2005.

[21] Duane Hanselman, Bruce Littlefield, «*Μάθετε το Matlab 7*», Εκδόσεις Κλειδάριθμος, 2006, τόμος Α και Β.

[22] Δημήτρης Α. Γεωργίου, «*Αριθμητική Ανάλυση, Παραδείγματα, Ασκήσεις και Θέματα Εξετάσεων*», Εκδ. Κλειδάριθμος, Αθήνα, 2008.

Χ. Α. Αλεξόπουλος
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Πανεπιστήμιο Πατρών