Иван Марков

ВЪВЕДЕНИЕ В МАТРИЧНОТО СМЯТАНЕ

София
2007

1. Кратки сведения от линейната алгебра

В тази глава накратко са представени основни математически понятия, означения и резултати, които често се използват в по-нататъшното изложение. Приведеният материал в тази глава е предназначен само за улеснение на читателя и не претендира за пълнота и изчерпателност. Ако читателят желае по-подробна информация, може да се обърне към специализираните книги по линейна алгебра и математически анализ. Част от тези книги са дадени в използваната литература.

1.1 Матрици. Основни понятия

В цялото изложение по-нататък с $\mathbb{R}$ ще означаваме множеството от реалните числа (реалната права), а с $\mathbb{R}^n$ ще означаваме реалното *n*-мерно векторно пространство. Понятието векторно пронстранство е фундаментално в теорията на матриците, но засега ще го използваме интуитивно, а по-късно ще дадем точното му определение. В тази книга ще използваме предимно матрици, определени над полето на реалните числа $\mathbb{R}$.При дефинирането на матрици се използват два подхода. При първия матриците се дефинират като правоъгълни таблици (масиви) от числа (скалари), а при втория подход матрицата се разглежда като оператор на линейното изображение (пробразуване) на едно векторно пространство в друго, когато във всяко от пространствата е фиксиран базис.

Определение: Матрицата е правоъгълна таблица (масив) от числа. Нека размерът на масива е $m \times n$. Това означава, че матрицата има *m* реда и *n* колонки. Матрците ще означаваме по един от следните начини

$$\mathbf{A} \equiv [A] \equiv \mathbf{A}_{m\times n} \equiv [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{1.1}$$

Ако $m = n$, матрицата се нарича ***квадратна*** и ще казваме, че тя е от ред *n*. С a_{ij} ще означаваме произволен елемент от матрицата, който принадлежи на *i*-тия ред и на *j*-тата колонка. Елементите a_{ij} на матрицата, за които $i = j$, т.е. a_{ii} се наричат ***главен диагонал*** на матрицата **A**. Ако матрицата се състои само от един ред ($m = 1$) или само от една колонка ($n = 1$), тя се нарича ***вектор***. Следователно, под вектор ще разбираме едномерния масив

$$\mathbf{x} \equiv \{x\} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \tag{1.2}$$

и ще използваме означението $\mathbf{x} \in \mathbb{R}^n$. Ако изрично не уговорено нещо друго, всички вектори по-нататък ще считаме за вектор-колонки. Числата $x_1, x_2, \cdots, x_n$ ще наричаме компоненти на вектора.

По определение ще смятаме, че

$$\mathbf{x} \geq 0, \quad \text{ако} \quad x_i \geq 0 \ \text{ за } \ \forall i = 1,2,\cdots n, \tag{1.3}$$

$$\mathbf{x} \le 0, \quad \text{ако} \quad x_i \le 0 \ \text{ за } \ \forall i = 1,2,\cdots n. \tag{1.4}$$

Ще покажем използването на матрици на простия пример за система линейни алгебрични уравнения

$$\begin{aligned} 10x_1 - 4x_2 + 2x_3 \quad &= \ 18 \\ -4x_1 + 12x_2 + x_3 \quad &= -14 \\ 2x_1 + x_2 + 15x_3 - x_4 &= \ 30 \\ -x_3 + 20x_4 &= \ 18 \end{aligned} \tag{1.5}$$

Тази система линейни уравнения може да се запише в матричен вид по следния начин

$$\underbrace{\begin{bmatrix} 10 & -4 & 2 & 0 \\ -4 & 12 & 1 & 0 \\ 2 & 1 & 15 & -1 \\ 0 & 0 & -1 & 20 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{Bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{Bmatrix}}_{\mathbf{x}} = \underbrace{\begin{Bmatrix} 18 \\ -14 \\ 30 \\ 18 \end{Bmatrix}}_{\mathbf{b}}. \tag{1.6}$$

Записът (1.6) е много по-логичен и естествен от алгоритмична гледна точка. Ако използваме матричната символика за означения на масиви, системата (1.6) можем да запишем в следния компактен вид

$$\mathbf{Ax} = \mathbf{b}, \tag{1.7}$$

където **A** е матрица (двумерен масив) от коефициентите на системата уравнения, **x** е вектор (едномерен масив) на неизвестните и **b** е вектор на известните свободни членове.

При означаване на елемент от матрица с помощта на индекси, те се разделят със запетая само в случаите, в които трябва да се избегне недоразумение, например $a_{i+2,j+1}$.

На всяка матрица **A** с размери $m \times n$ с елементи a_{ij} съответства матрица с размери $n \times m$ и с елементи a_{ji}. Втората матрица се нарича ***транспонирана*** на **A** и се означава с $\mathbf{A}^T$. Редовете на матрицата **A** са колонки на матрицата $\mathbf{A}^T$, а колоните на **A** се превръщат в редове на $\mathbf{A}^T$. Очевидно е, че $\left(\mathbf{A}^T\right)^T = \mathbf{A}$. Процесът на транспониране става ясен от следния пример

$$\mathbf{A} = \begin{bmatrix} 4 & -1 & 3 \\ 2 & 15 & -6 \\ 16 & -1 & 9 \\ 3 & 2 & 1 \end{bmatrix}_{4\times 3}, \qquad \mathbf{A}^T = \begin{bmatrix} 4 & 2 & 16 & 3 \\ -1 & 15 & -1 & 2 \\ 3 & -6 & 9 & 1 \end{bmatrix}_{3\times 4}. \tag{1.8}$$

Определение: Две матрици **A** и **B** са равни тогава и само тогава, когато те с еднакви размери, а съответните им компоненти са равни, т.е. $a_{ij} = b_{ij}$ за всички *i* и *j*.

Сега ще дефинираме някои ***елементарни операции с матрици***.

a. ***Сумата*** на две матрици **A** и **B** с еднакви размери $m \times n$ е матрицата **C** с размери $m \times n$ и с елементи $c_{ij} = a_{ij} + b_{ij}$ и се записва $\mathbf{C} = \mathbf{A} + \mathbf{B}$. Ако матриците **A** и **B** имат различни размери, то операцията събиране на матрици не е дефинирана.

b. ***Произведението*** на матрицата $\mathbf{A}_{m\times n}$ със скалара $\alpha \in \mathbb{R}$ е матрицата $\mathbf{C}_{m\times n}$ с елементи $c_{ij} = \alpha a_{ij}$ и прозведението се записва $\mathbf{C} = \alpha\mathbf{A}$. В този случай казваме, че елементите на **A** се умножават покомпонентно със скалара α.

c. ***Произведението на матриците*** $\mathbf{A}_{m\times n}$ и $\mathbf{B}_{n\times p}$ е матрицата $\mathbf{C}_{m\times p}$ с компоненти

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \quad \Leftrightarrow \quad \mathbf{C} = \mathbf{AB}\,. \tag{1.9}$$

От формули (1.9) се вижда, че елементът c_{ij} се получава чрез умножаване на елементите от *i*-тия ред на матрицата **A** по съответните елементи от *j*-тата колонка на матрицата **B** и сумиране на получените произведения. Ще отбележим, че операцията умножение на две матрици е дефинирана тогава и само тогава, когато броят на колонките на матрицата **A** е равен на броя на редовете на матрицата **B**.

Пример 1.1. Да се изчисли произведението $\mathbf{C} = \mathbf{AB}$, където

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 0 \\ 0 & -2 & 3 \\ 1 & 5 & 6 \end{bmatrix}_{3\times 3}, \qquad \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 2 & 3 \\ -5 & 1 \end{bmatrix}_{3\times 2}.$$

Ще получим подробно елементите от първия ред на матрицата **C** по (1.9)

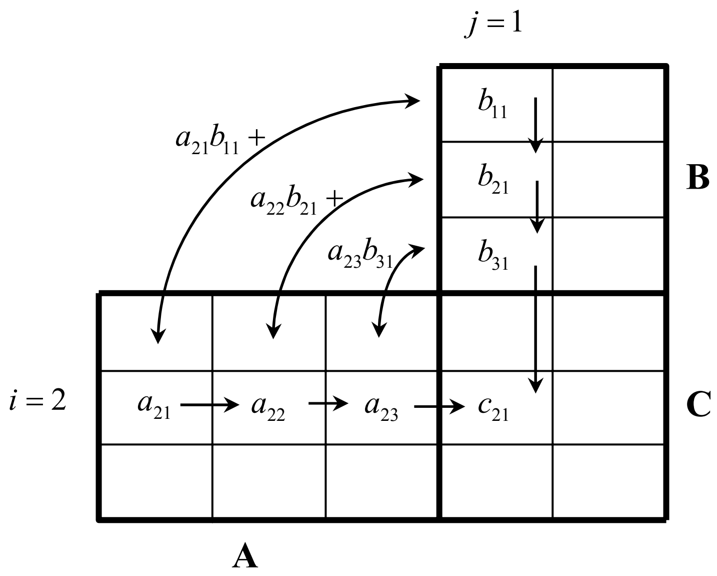
$$c_{11} = \sum_{k=1}^{3} a_{1k} b_{k1} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} = 6\,,$$

$$c_{12} = \sum_{k=1}^{3} a_{1k} b_{k2} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} = -1\,.$$

Като пресметнем и останалите елементи на **C**, получаваме

$$\mathbf{C} = \begin{bmatrix} 6 & -1 \\ -19 & -3 \\ -19 & 20 \end{bmatrix}_{3\times 2}.$$

Процесът на умножение на две матрици $\mathbf{A}_{3\times 3}$ и $\mathbf{B}_{3\times 2}$ е показан схематично на фиг. 1.1.



$$c_{21} = a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31}$$

Фиг. 1.1. Умножение на матрици

Показаната на фиг. 1.1 схема за умножение на матрици е удобна само за ръчни изчисления и за по-лесно записване на алгебрични и диференциални уравнения в матричен вид, което ще бъде изяснено по-нататък.

Умножението на матрица $\mathbf{A}_{m\times n}$ по вектор $\mathbf{x} \in \mathbb{R}^n$ съвпада с умножение на матриците $\mathbf{A}_{m\times n}$ и $\mathbf{x}_{n\times 1}$.От показаните правила с непосредствена проверка може да се установи верността на записа на уравнения (1.5) във вида (1.6).

Както е известно, умножението на произволни числа (скалари) е ***комутативна операция***, т.е. $ab = ba$. Ще проверим дали това свойство е валидно и за операцията умножение на матрици. Разглеждаме матриците

$$\mathbf{A} = \begin{Bmatrix} 2 \\ 5 \end{Bmatrix} \text{ и } \mathbf{B} = \begin{bmatrix} 1 & 2 \end{bmatrix}.$$

Изчисляваме произведенията

$$\mathbf{AB} = \begin{bmatrix} 2 & 4 \\ 5 & 10 \end{bmatrix}, \qquad \mathbf{BA} = \begin{bmatrix} 12 \end{bmatrix} \equiv 12 .$$

От получените резултати се вижда, че матрриците $\mathbf{AB}$ и $\mathbf{BA}$ са различни. Даже в

+

зависимост от размерите на **A** и **B**, матрриците $\mathbf{AB}$ и $\mathbf{BA}$ могат да се различават и по размери. Също така е възможно произведението $\mathbf{AB}$ да съществува, а произведението $\mathbf{BA}$ да не съществува (това може непосредствено да се провери за матриците **A** и **B** от пример 1.1). От тези примери следва, че ***операцията умножение на матрици не е комутативна***.

Ако за двематрици **A** и **B** е изпълнено $\mathbf{AB} = \mathbf{BA}$, се казва, че **A** и **B** ***комутират*** спрямо операцията умножение на матрици.

Умножението на матрици е ***дистрибутивно*** по отношение на събирането, т.е.

$$\mathbf{A}\left(\mathbf{B}+\mathbf{C}\right)=\mathbf{AB}+\mathbf{AC}. \tag{1.10}$$

Това свойство е в сила само в случай, че участващите в (1.10) операции са възможни.

Умножението на матрици е ***асоциативно***, т.е. ако **A, B** са съответни една на друга (дефинирано е произведенито $\mathbf{AB}$) и **B**,**C** са също съответни, то е изпълнено

$$\mathbf{A}\left(\mathbf{BC}\right)=\left(\mathbf{AB}\right)\mathbf{C}, \tag{1.11}$$

което означава, че редът на изпълнение на операциите умножение е без значение.

Описаните свойства се проверяват непосредствено чрез използване на определението за умножение на матрици (1.9).

Определение: Матрица, на която всичките елементи са равни на нула, се нарича ***нулева*** и се означава с 0 или с **0**.

Определение: Матрица, на която всички елементи по главния диагонал са равни на единица, а всички останали елементи са равни на нула, се нарича ***единична матрица*** и ще я означаваме с **I**.

Например, единичната матрица от трети ред е

$$\mathbf{I}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}. \tag{1.12}$$

С непосредствена проверка може да се установи, че за произволни матрици **A** и **B** са в сила равенствата

$$\mathbf{A0}=\mathbf{0A}=\mathbf{0}, \qquad \mathbf{AI}=\mathbf{A}, \qquad \mathbf{IB}=\mathbf{B}, \tag{1.13}$$

при условие, че съответните матрични операции са определени.

В края на този параграф ще дадем на някои ***свойства на операциата умножение на матрици***.

a. В матричните равенства съкращаването на еднаквите матрици не винаги е възможно. От равенството $\mathbf{AB} = \mathbf{AC}$ не следва, че $\mathbf{B} = \mathbf{C}$. Също така, от $\mathbf{AB} = \mathbf{0}$ не следва, че някоя от матриците **A** или **B** е нулева.

b. Транспонирането на произведението на матриците **A** и **B** е равно на произведението от транспонираните им матрици, взети в обратен ред, т.е.

$$\left(\mathbf{AB}\right)^T=\mathbf{B}^T\mathbf{A}^T. \tag{1.14}$$

Доказателството на (1.14) се извършва чрез използване на определение (1.9).

Определение: Матрицата $\mathbf{A}=\left[a_{ij}\right]$ се нарича ***симетрична***, ако тя е равна на своята транспонирана, т.е. $\mathbf{A}^T=\mathbf{A}$ или $a_{ij}=a_{ji},$ за $\forall i,j$. Вижда се, че ако една матрица **A** е симетрична, то тя е квадратна. Разбира се, обратното твърдение не е вярно, тъй като не всяка квадратна матрица е симетрична.

За произволна матрица A произведенията $\mathbf{A}^T\mathbf{A}$ и $\mathbf{A}\mathbf{A}^T$ са симетрични матрици. Това се вижда от равенството

$$\left(\mathbf{A}\mathbf{A}^T\right)^T = \left(\mathbf{A}^T\right)^T \mathbf{A}^T = \mathbf{A}\mathbf{A}^T . \tag{1.15}$$

Например, ако

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \qquad \text{то} \qquad \mathbf{A}\mathbf{A}^T = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix} \text{ и } \quad \mathbf{A}^T\mathbf{A} = \begin{bmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{bmatrix}.$$

Ще отбележим, че ако **A** и **B** са симетрични матрици с еднакви размери, произведението **AB** в общия случай не е симетрична матрица. Това се вижда от следния пример

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix}, \qquad \mathbf{AB} = \begin{bmatrix} 5 & -1 \\ 0 & 2 \end{bmatrix}.$$

c. Ако матрицата **D** е симетрична, то матрицата $\mathbf{B}^T\mathbf{D}\mathbf{B}$ е също симетрична. Доказателството се основава на (1.14), т.е.

$$\left(\mathbf{B}^T\mathbf{D}\mathbf{B}\right)^T = \left(\mathbf{D}\mathbf{B}\right)\left(\mathbf{B}^T\right)^T = \mathbf{B}^T\mathbf{D}^T\mathbf{B}, \quad \text{но } \mathbf{D}^T = \mathbf{D}$$

Следователно

$$\left(\mathbf{B}^T\mathbf{D}\mathbf{B}\right)^T = \mathbf{B}^T\mathbf{D}\mathbf{B}. \tag{1.16}$$

Оот последното равенство следва, че $\mathbf{B}^T\mathbf{D}\mathbf{B}$ е симетрична. Лесно се вижда, че (1.15) е частен случай на (1.16) при $\mathbf{D} = \mathbf{I}$.

Ако $\mathbf{A}_{m\times n}$ е матрица, чийто компоненти са диференцируеми функции на една или няколко независими променливи, то под диференциране на матрица ще разбираме

$$\frac{\partial \mathbf{A}}{\partial x} = \left[\frac{\partial a_{ij}}{\partial x}\right], \tag{1.17}$$

а като интегриране на матрица ще дефинираме операцията

$$\int_a^b \mathbf{A}dx = \left[\int_a^b a_{ij}dx\right]. \tag{1.18}$$

От (1.17) и (1.18) се вижда, че диференцирането и интегрирането на на матрица са покомпонентни операции.

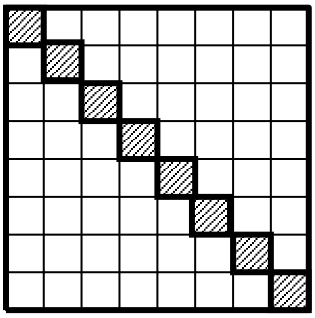
1.2 Видове матрици

Една матрица има ***особена форма***, ако нейните елементи се подчиняват на определен закон. Ако елементите на една матрица са реални числа, матрицата се нарича ***реална***, а ако те са комплексни числа, матрицата се нарича ***комплексна***.

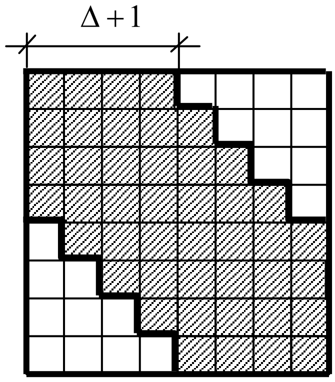
Квадратна матрица, на която елементите по главния диагонал са ненулеви, а всички останали елементи са нули, се нарича ***диагонална матрица***. Понякога диагоналните матрици се означават по следния начин

$$\mathbf{D} = diag\left(d_{ii}\right).$$

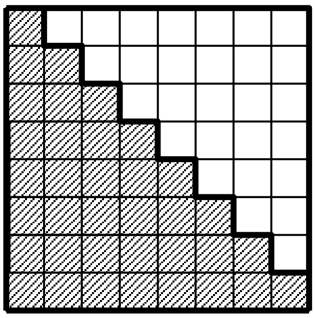
Диагоналната матрица е показана схематично на фиг 1.2.а. На фигурата всички компоненти на матриците, които не са защриховани, са нулеви.



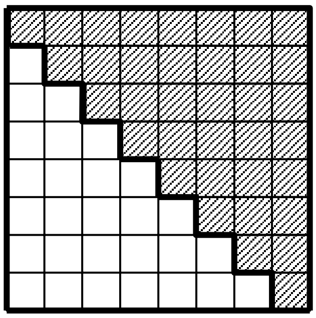
a. Диагонална матрица



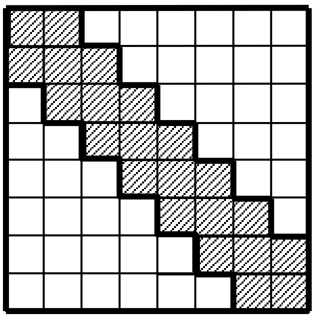
b. Ивична матрица



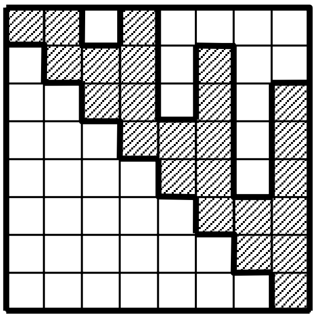
c. долна триъгълна матрица



d. горна триъгълна матрица



e.тридиагонална матрица



f. “sky line” матрица

Фиг. 1.2. Видове матрици

Квадратната матрица **A** се нарича ***ивична*** (фиг. 1.2.b), ако е изпълнено условиета

$$a_{ij} = 0 \quad \text{за } \forall j > i + \Delta \ \text{ и за } \forall i > j + \Delta \,, \tag{1.19}$$

където $2\Delta + 1$ е ширината на ивицата на матрицата. Като пример ще приведем матрицата

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & -2 & 0 & 0 \\ 1 & 5 & -1 & 1 & 0 \\ -2 & -1 & 12 & -1 & 1 \\ 0 & 1 & -1 & 9 & 2 \\ 0 & 0 & 1 & 2 & 10 \end{bmatrix}.$$

Тази матрица е ***симетрична ивична матрица*** от пети ред с ***полуширина на ивицата*** $\Delta = 2$. Ако полуширината на ивицата е нула, то матрицата е диагонална.

Матриците, показани схематично на фиг 1.2.c и фиг.1.2.d, се наричат съответно ***долна триъгълна*** и ***горна триъгълна*** матрица.

На фиг. 1.2.f е показана ***симетрична*** матрица от тип “***skyline***”. При нея всички елементи, които лежат над горната гранична линия, са нулеви. Задачите на МКЕ най-често водят до такъв тип матрици, при това с голяма рзмерност.

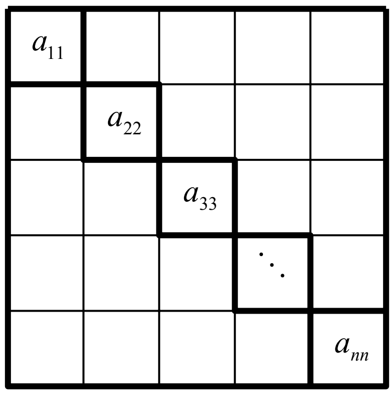
При компютърното решение на задачите по МКЕ се използват различни начини за съхраняване на елементите на матрица в оперативната (или виртуалната) памет. За програми, написани на езици от високо ниво, матрицата **A** с размери $m \times n$ се представя в паметта чрез двумерен масив $\mathrm{A}[\mathrm{m}][\mathrm{n}]$ и всеки елемент a_{ij} от матрицата **A** се съхранява в клетката $\mathrm{A}[\mathrm{i}][\mathrm{j}]$ от паметта. В повечето случаи представянето на матриците във вид на двумерни масиви води до съхраняването на голямо количество нулеви елементи, които не се използват в изчисленията. Освен това, ако матрицата е симетрична, трябва да се използва това свойство и да се съхранява само тази част от нея, която включва елементите по главния диагонал и над (или под) него. Ефективното съхранение на матриците в паметта на компютъра води до ускоряване на изчислителния процес, защото в повечето алгоритми се обработват само неулевите компоненти на отделните видове матрици.

Диагонална матрица **A** от ред *n* се съхранява в паметта като едномерен масив

$$\mathrm{A}[\mathrm{i}] = a_{ii}, \quad i = 1, 2, \cdots, n \,. \tag{1.20}$$

Това е показано на фиг. 1.3.a. Също така, на фиг. 1.3.b е показан начинът на съхранение симетрична ивична матрица от типа “skyline”. Нулевите елементи между двете гранични линии при преобразувания на матрицата се преобразуват в ненулеви. Например, елементът a_{46} може да е нулев, но при реализацията на алгоритъма за решение на система уравнения този елемент се преобразува в ненулев. Нулевите елементи, които лежат над граничната линия “skyline”, при този алгоритъм винаги се преобразуват в нулеви. Поради това се съхраняват само елементите, които са между двете гранични линии. Разбира се, възможни са идруги схеми за съхраняване на ивични матрици. Тук е използвана схемата, дадена в [Bathe].

a.



$A[1] = a_{11}$,

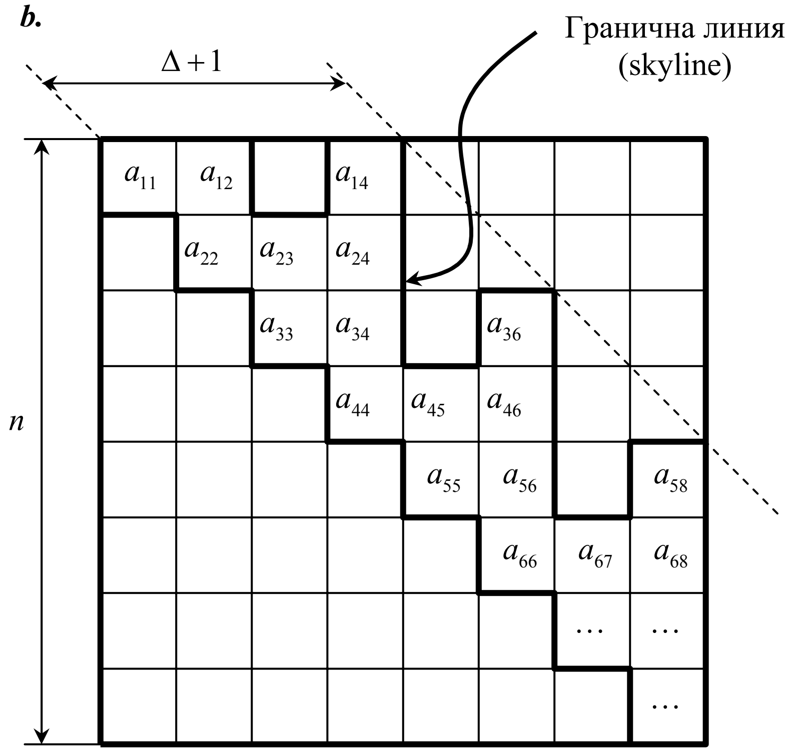
$A[2] = a_{22}$,

$A[3] = a_{33}$,

. . . ,

$A[n] = a_{nn}$

b.



A[1]=a_{11}, A[2]=a_{22}, A[3]=a_{12}, A[4]=a_{33}, A[5]=a_{23}, A[6]=a_{44}, A[7]=a_{34}, A[8]=a_{24}, A[9]=a_{14}, A[10]=a_{55}, A[11]=a_{45},…, A[MaxArr]=a_{nn}

Фиг. 1.3. Съхраняване на елементите на матрица в едномерен масив. а. диагонална матрица; b. симетрична ивична матрица “skyline”

При реализацията на алгоритмите, при които елементите на матрицата се съхраняват по схемата, показана на фиг. 1.3.b, е необходимо в един едномерен масив да се запомнят номерата (адресите) на диагоналните елементи. Конкретно за показания пример този целочислен масив е

$$\begin{aligned}&\mathrm{DiagAddr[1]=1}, \quad \mathrm{DiagAddr[2]=2}, \quad \mathrm{DiagAddr[3]=4}, \\ &\mathrm{DiagAddr[4]=6}, \quad \mathrm{DiagAddr[5]=10}, \ldots, \mathrm{DiagAddr[n]=MaxArr}, \\ &\mathrm{DiagAddr[n+1]=MaxArr+1},\end{aligned} \tag{1.21}$$

където MaxArr е броят на елементите под граничната линия (размерността на масива **A)**.

1.3 Детерминанта и следа на матрица

Много често в математиката обект, който се състои от много компоненти, се охарактеризира (оценява) с помощта на една величина (скалар). Примери за такава величина са ***детерминанта*** и ***следа на матрица***. Изчисляването на детерминанта и следа на матрица е възможно само за квадратни матрици. Тези величини са функции на компонентите на матрицата.

В литературата съществуват два подхода [Хорн] при дефиниране на детерминанта. Тук ще дадем само определението, което използва индукция по размера *n* на матрицата, т.е. детерминантата на матрицата $\mathbf{A}_{n\times n}$ ще изчисляваме чрез детерминантите на матриците от по-нисък ред.

Определение:

a. Детерминантата на матрица от първи ред е (n=1) равна на самия елемент. Ако $\mathbf{A} = [a_{11}]$, то $\det\mathbf{A} = a_{11}$.

b. Ако *i* е произволен ред от матрицата $\mathbf{A}_{n\times n}$, то детерминантата $\det\mathbf{A}$ се из изчислява по формулата

$$\det\mathbf{A} = \sum_{j=1}^{n} (-1)^{i+j} \det\mathbf{A}_{ij}, \tag{1.22}$$

където $\mathbf{A}_{ij}$ е матрица от ред $(n-1)\times(n-1)$, която се получава от **A** чрез отстраняване на *i*-тия ред и *j*-тата колонка. Формула (1.22) е известна като ***разлагане на Лаплас*** по *i*-тия ред на матрицата **A**. Аналогично, ако се използва разлагане на Лаплас по *j*-тата колонка, определението за детерминанта на матрица се записва във вида

$$\det\mathbf{A} = \sum_{i=1}^{n} (-1)^{i+j} \det\mathbf{A}_{ij}. \tag{1.23}$$

Пример: Да се изчисли детерминантата на матрицата **A**

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Ще използваме формула (1.22). Развиваме по първия ред и получаваме

$$\mathbf{A}_{11} = [a_{22}], \qquad \mathbf{A}_{12} = [a_{21}],$$

$$\det \mathbf{A} = (-1)^{1+1} a_{11} \det \mathbf{A}_{11} + (-1)^{1+2} a_{12} \det \mathbf{A}_{12},$$

но

$$\det \mathbf{A}_{11} = a_{22}, \qquad \det \mathbf{A}_{12} = a_{21}.$$

Следователно

$$\det \mathbf{A} = a_{11}a_{22} - a_{12}a_{21}. \tag{1.24}$$

Аналогично, ако разложим матрицата

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

по първия ред, получаваме последователно

$$\mathbf{A}_{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \qquad \mathbf{A}_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \qquad \mathbf{A}_{13} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

От формула (1.24) следва, че

$$\det \mathbf{A}_{11} = a_{22}a_{33} - a_{23}a_{32},$$

$$\det \mathbf{A}_{12} = a_{21}a_{33} - a_{23}a_{31},$$

$$\det \mathbf{A}_{13} = a_{21}a_{32} - a_{22}a_{31}.$$

Тогава по (1.22) получаваме

$$\det \mathbf{A} = (-1)^{1+1} a_{11} \det \mathbf{A}_{11} + (-1)^{1+2} a_{12} \det \mathbf{A}_{12} + (-1)^{1+3} a_{13} \det \mathbf{A}_{13}.$$

След като заместим $\det \mathbf{A}_{ij}$ в последното равенство, получаваме

$$\begin{aligned} \det \mathbf{A} = {} & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \end{aligned}. \tag{1.25}$$

Така получихме известните формули (1.24) и (1.25) за изчисляване на детерминанти на матрици от втори и трети ред.

Ще припомним някои ***основни свойства на детерминантите***. Те се доказват чрез непосредствено прилагане на формули (1.22) и (1.23), а за някои от свойствата допилнително се използва методът на математическата индукция. Повечето от доказателствата са обемисти и трудоемки и затова няма да се спираме върху тях.

1. $\det \mathbf{A}^T = \det \mathbf{A}$.
2. $\det \mathbf{I} = 1$.
3. ***Мултипликативност:*** детерминантата от произведението на две или повече матрици е равна от произведението от детерминантите на матриците множители

$$\det(\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B}. \tag{1.26}$$

4. Детерминантата на ***диагонална матрица*** $\mathbf{D} = diag(d_{ii})$ е равна на произведението на нейните диагонални елементи

$$\det \mathbf{D} = d_{11}d_{22}\ldots d_{nn}.$$

5. Детерминантата на долна или горна тригълна матрица е равна на произведението на елементите по главния диагонал

При решаване на системи линейни уравнения често се използва метод, при който симетричната матрица **A** се представя във вид на произведение от три матрици $\mathbf{A} = \mathbf{LDL}^T$, в което **L** е долна триъгълна матрица с елементи по главния диагонал равни на единица $\det \mathbf{L} = 1$, а **D** е диагонална матрица. Тогава от (1.26) следва, че

$$\det \mathbf{A} = \det \mathbf{L} \det \mathbf{D} \det \mathbf{L}^T. \tag{1.27}$$

Тъй като $\det \mathbf{L} = 1$, то

$$\det \mathbf{A} = \det \mathbf{D} = d_{11}d_{22}\ldots d_{nn}. \tag{1.28}$$

6. Ако α е скалар, то

$$\det(\alpha \mathbf{A}) = \alpha^n \det \mathbf{A}. \tag{1.29}$$

С много други свойства на детерминантите, свързани с елементарни преобразувания на матрици, читателят може да се запознае от специализираните курсове по теория на матриците, например [Хорн], [Ланкастер].

Определение: Следа на квадратна матрица **A** с размерност *n* е сумата на елемнтите от главния диагонал на **A**, т.е.

$$tr(\mathbf{A}) = Sp(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}. \tag{1.30}$$

1.4 Обратна матрица

Определение: Матрицата $\mathbf{A}^{-1}$ е обратна на матрицата **A**, ако е изпълнено

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}, \tag{1.31}$$

където **I** е единичната матрица. Не всяка квадратна матрица има обратна. Например, нулевата матрица няма обратна.

Определение: Квадратната матрица се нарича ***неособена*** (***неизродена***), ако нейната детерминанта е различна от нула.
Може да се докаже следната теорема [Демидович], [Ланкастер], [Хорн]

Теорема: Всяка неособена матрица има обратна матрица.

Ще се спрем на някои ***свойства на обратните матрици:***

1. ***Детерминанта на обратна матрица***. Получава се непосредствено от определение (1.31) и свойство (1.26) на детерминантите

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \quad \Rightarrow \quad \det \mathbf{A}^{-1} \det \mathbf{A} = \det \mathbf{I} = 1 .$$

Следователно

$$\det \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} . \tag{1.32}$$

2. ***Обратна матрица на произведение от матрици***

$$\left(\mathbf{AB}\right)^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} . \tag{1.33}$$

3. ***Обратна на транспонирана матрица***

$$\left(\mathbf{A}^{-1}\right)^{T} = \left(\mathbf{A}^{T}\right)^{-1} . \tag{1.34}$$

4. Ако **A** е симетрична, то обратната матрица $\mathbf{A}^{-1}$ е също симетрична.

Значението на понятието обратна матрица най-добре се илюстрира чрез съкратения запис (1.7) за система от *n* линейни алгебрични уравнения.. Ако **A** е неособена, то след умножаване на двете страни на (1.7) с $\mathbf{A}^{-1}$, решението на системата уравнения получаваме във вида

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} . \tag{1.35}$$

Разбира се, (1.35) е удобен символичен запис, но при решаването на системи линейни алгебрични уравнения, както ще видим по-нататък, не е необходимо да се изчислява обратната матрица $\mathbf{A}^{-1}$. Навсякъде по-нататък, където използваме записи от вида (1.35), в които участват обратни матрици, ще имаме предвид процеса на решаване на съответната система уравнения, а няма да се изчислява обратна матрица.

Един от начините за изчисляване на обратната матрица $\mathbf{A}^{-1}$ е чрез използване на детерминанти. Ако β_{ij} са компонентите на обратната матрица, т.е. $\boldsymbol{\beta} \equiv \mathbf{A}^{-1} = \left[\beta_{ij}\right]$, то елементите на обратната матрица се изчисляват по формулите

$$\beta_{ij} = \left(-1\right)^{i+j} \frac{\det \mathbf{A}_{ji}}{\det \mathbf{A}} = \frac{\alpha_{ji}}{\det \mathbf{A}} , \qquad \left(i, j = 1, 2, \ldots, n\right), \tag{1.36}$$

където $\mathbf{A}_{ji}$ са подматриците на **A** от ред $(n-1)$, получени чрез задраскване (отстраняване) на *j*-тия ред и *i*-тата колонка на **A**. Тук ще припомним, че $\alpha_{ji} = (-1)^{i+j} \det \mathbf{A}_{ji}$ се наричат ***адюнгирани количества***. Формули (1.36) са удобни за приложение при матрици с много малки размери ($n \le 3$). При използването на компютри формули (1.36) нямат практическо приложение. Затова тук ще се спрем на един общ алгоритъм за изчисляаване на обратната матрица $\mathbf{A}^{-1}$. Нека **A** и **β** са квадратни матрици с размери $n \times n$, които удовлетворяват матричното уравнение

$$\mathbf{A}\boldsymbol{\beta} = \mathbf{I}, \qquad (1.37)$$

където **I** е единична матрица. Тогава $\boldsymbol{\beta} = \mathbf{A}^{-1}$. За решаването на уравнения (1.37) може да се приложи всеки от известните алгоритми за решаване на системи линейни алгебрични уравнения. Ако $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$ са векторите-колонки на матрицата **β**, а $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ са векторите-колонки на единичната матрица **I**, т.е.

$$\boldsymbol{\beta}_i = \begin{Bmatrix} \beta_{1i} \\ \beta_{2i} \\ \vdots \\ \beta_{ni} \end{Bmatrix}, \qquad \mathbf{e}_i = \begin{Bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{Bmatrix} \to \text{ред } i,$$

то призволна колонка $\boldsymbol{\beta}_i$ от обратната матрица **β** се получава от решаването на системата уравнения

$$\mathbf{A}\boldsymbol{\beta}_i = \mathbf{e}_i, \quad (i = 1, 2, \ldots, n). \qquad (1.38)$$

След решаване на *i*-тото уравнение от (1.38), векторът на решението $\boldsymbol{\beta}_i$ се разполага в *i*-тата колонка на обратната матрица $\mathbf{A}^{-1} = \boldsymbol{\beta}$

Пример 1.2. Да се изчисли $\boldsymbol{\beta} = \mathbf{A}^{-1}$ за матрицата

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

Единичната матрица **I** от трети ред и нейните вектор-колонки $\mathbf{e}_i$ са съответно

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3], \qquad \mathbf{e}_1 = \begin{Bmatrix} 1 \\ 0 \\ 0 \end{Bmatrix}, \qquad \mathbf{e}_2 = \begin{Bmatrix} 0 \\ 1 \\ 0 \end{Bmatrix}, \qquad \mathbf{e}_3 = \begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix}.$$

С готова програма решаваме последователно за $i = 1, 2, 3$ системите линейни алгебрични уравнения

$$\mathbf{A}\boldsymbol{\beta}_i = \mathbf{e}_i$$

и получаваме

$$\boldsymbol{\beta}_1 = \begin{Bmatrix} 3 \\ 2 \\ 1 \end{Bmatrix}, \qquad \boldsymbol{\beta}_2 = \begin{Bmatrix} 2 \\ 2 \\ 1 \end{Bmatrix}, \qquad \boldsymbol{\beta}_3 = \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix}.$$

След подреждането на $\boldsymbol{\beta}_i \ (i = 1,2,3)$ в колонките на $\boldsymbol{\beta} = \mathbf{A}^{-1}$, получаваме

$$\boldsymbol{\beta} = \mathbf{A}^{-1} = \begin{bmatrix} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \boldsymbol{\beta}_3 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

1.5 Блочни матрици

Понякога за опростяване на алгоритмите за изчисления и за по-компактно записване на уравненията в уравненията в строителната механика е много полезно матриците да се разделят на ***подматрици*** или ***блокове***. Разбиването на матрица на блокове е показано в следния пример

$$\mathbf{A} = \left[\begin{array}{c:ccc:cc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ \hdashline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{array}\right], \tag{1.39}$$

в който отделните блокове са разделени с пунктирни линии. В блочна форма тази матрица се записва във вида

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix}, \tag{1.40}$$

където

$$\mathbf{A}_{11} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \quad \mathbf{A}_{12} = \begin{bmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \end{bmatrix}, \qquad \mathbf{A}_{13} = \begin{bmatrix} a_{15} & a_{16} \\ a_{25} & a_{26} \end{bmatrix}, \ldots \tag{1.41}$$

$\mathbf{A}_{ij}$ се наричат ***подматрици*** или ***блокове*** на матрицата **A**.

Използването на блочни матрици е много полезно и ефективно при разработване на компютърни програми. Например, ако блоковете на една матрица са еднакви, то в паметта може да се съхранява само един блок. От друга страна, ако операциите с блоковете на дадена матрица са еднакви, то те могат да бъдат описани с една процедура, която да се извиква многократно.

Правилата за операции с блочни матрици са аналогични на обикновените матрични операции. Трябва да подчертаем, че разбиването на матрици на блокове е само начин за опростяване на операциите с тях, а крайният резултат остава непроменен.

Пример 1.3. Да се изчисли произведението $\mathbf{C} = \mathbf{AB}$ от пример 1.1, като се използва следното блочно представяне

$$\mathbf{A} = \left[\begin{array}{cc|c} 4 & 1 & 0 \\ 0 & -2 & 3 \\ \hline 1 & 5 & 6 \end{array}\right] = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \qquad \mathbf{B} = \left[\begin{array}{cc} 1 & -1 \\ 2 & 3 \\ \hline -5 & 1 \end{array}\right] = \begin{bmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{bmatrix}.$$

Като използваме определение (1.9), или схемата за умножение, показана на фиг. 1.1, можем да запишем

$$\mathbf{AB} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} \end{bmatrix}. \tag{1.42}$$

Изчисляваме произведенията, участващи в (1.42)

$$\mathbf{A}_{11}\mathbf{B}_{11} = \begin{bmatrix} 4 & 1 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 6 & -1 \\ -4 & -6 \end{bmatrix},$$

$$\mathbf{A}_{12}\mathbf{B}_{21} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}\begin{bmatrix} -5 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -15 & 3 \end{bmatrix},$$

$$\mathbf{A}_{21}\mathbf{B}_{11} = \begin{bmatrix} 1 & 5 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 14 \end{bmatrix}$$

$$\mathbf{A}_{22}\mathbf{B}_{21} = \begin{bmatrix} 6 \end{bmatrix}\begin{bmatrix} -5 & 1 \end{bmatrix} = \begin{bmatrix} -30 & 6 \end{bmatrix}.$$

Като заместим последните произведения в (1.42), получаваме известния от пример 1.1 резултата

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} 6 & -1 \\ -19 & -3 \\ -19 & 20 \end{bmatrix}.$$

Ще подчертаем, че блочното представяне трябва да е такова, че съответните операции с подматрици да са дефинирани.

Транспонирането на блочна матрица се извършва по правилто за транспониране на обикновена матрица, като редовете от блокове се превръщат в колонки от блокове и самите блокове също се транспонират. Така за матрицата от (1.40) получаваме

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix}, \qquad \mathbf{A}^T = \begin{bmatrix} \mathbf{A}_{11}^T & \mathbf{A}_{21}^T \\ \mathbf{A}_{12}^T & \mathbf{A}_{22}^T \\ \mathbf{A}_{13}^T & \mathbf{A}_{23}^T \end{bmatrix}.$$

Определение: Квадратната реална матрица се нарича ***ортогонална***, ако нейната транспонирана матрица $\mathbf{A}^T$ съвпада с обратната $\mathbf{A}^{-1}$, т.е

$$\mathbf{A}^T = \mathbf{A}^{-1}, \tag{1.43a}$$

$$\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}. \tag{1.43b}$$

Пример 1.4. Ако $\alpha, \beta \in \mathbb{R}$ и $|\alpha| + |\beta| \neq 0$, то матрицата

$$\mathbf{R} = \frac{1}{\sqrt{\alpha^2 + \beta^2}} \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \tag{1.44}$$

е ортогонална. С непосредствена проверка може да се установи, че за матрицата **R** е в сила (1.43b).

Ортогоналната матрица притежава следните свойства:

1. Ако представим ортогоналната матрица **A** във вида $\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_n \end{bmatrix}$, където $\mathbf{A}_i \left(i = 1, 2, \dots n \right)$ са векторите колонки на **A**, т.е. $\mathbf{A}_i = \begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{ni} \end{bmatrix}^T$, то са в сила равенствата

$$\mathbf{A}_i^T\mathbf{A}_j = \mathbf{A}_j^T\mathbf{A}_i = \delta_{ij} = \begin{cases} 1, \text{при } i = j \\ 0, \text{при } i \neq j \end{cases}, \tag{1.45}$$

   т.е. векторите $\mathbf{A}_i$ образуват ортонормирана система. Величината δ_{ij} се нарича ***символ на Кронекер***.

2. ***Детерминантата*** на ортогонална матрица **A** е равна на ± 1, т.е. $\det \mathbf{A} = \pm 1$.

3. Ако **A** е ортогонална, то $\mathbf{A}^T$ и $\mathbf{A}^{-1}$ са също ортогонални.

4. Ако **A** и **B** са ортогонални матрици с размери $n \times n$, то произведението **AB** е също ортогонална матрица. Това е очевидно от следните зависимости

$$\left(\mathbf{AB}\right)^T \left(\mathbf{AB}\right) = \mathbf{B}^T\mathbf{A}^T\mathbf{AB} = \mathbf{B}^T\mathbf{IB} = \mathbf{B}^T\mathbf{B} = \mathbf{I}. \tag{1.46}$$

5. Всяка ортогонална матрица от втори ред има вида

$$\begin{bmatrix} \cos\varphi & \sin\varphi \\ -\varepsilon\sin\varphi & \varepsilon\cos\varphi \end{bmatrix},$$

   където $\varepsilon = \pm 1$ и $\varphi \in \left[0, 2\pi\right]$ е произволен ъгъл.

Доказателствата на тези свойства се извършва, като се използва определението (1.43).

2. Квадратични форми и диференциране на матрици

Нека **A** е реална квадратна матрица с размери $n \times n$, а $\mathbf{x} \in \mathbb{R}^n$ е произволен n-мерен вектор. Функцията (скаларът)

$$F(\mathbf{x}) = \mathbf{x}^T \mathbf{A}\mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{1}$$

се нарича ***квадратична форма***. Матрицата **A** се нарича матрица на квадратичната форма. Най често тази матрица е симетрична.

Определение: Квадратичната форма $F(\mathbf{x})$ е

- положително определена, ако $F > 0$ за $\forall \mathbf{x} \in \mathbb{R}^n$.
- положително полуопределена, ако $F \geq 0$ за $\forall \mathbf{x} \in \mathbb{R}^n$.
- отрицателно полуопределена, ако $F \leq 0$ за $\forall \mathbf{x} \in \mathbb{R}^n$.
- отрицателно определена, ако $F < 0$ за $\forall \mathbf{x} \in \mathbb{R}^n$.

Симетричната матрица **A** е ***положително определена***, ако квадратичната форма $F(\mathbf{x}) = \mathbf{x}^T \mathbf{A}\mathbf{x}$ е положително определена.

Нека е зададена функцията с векторен аргумент $F(\mathbf{x}) = F(x_1, x_2, ..., x_n)$. Тогава ***градиент на функцията*** (или ***първа производна спрямо x***) се нарича величината

$$\nabla F = \frac{\partial F}{\partial \mathbf{x}} = \begin{Bmatrix} \dfrac{\partial F}{\partial x_1} \\ \dfrac{\partial F}{\partial x_2} \\ \vdots \\ \dfrac{\partial F}{\partial x_n} \end{Bmatrix}. \tag{2}$$

Нека $\mathbf{x} = [x_1, x_2, ..., x_n]$ и е дадена квадратичната форма

$$\varphi(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A}\mathbf{x}. \tag{3}$$

Подробно квадратичната форма (3) може да се запише в следния разгърнат вид

$$\begin{aligned} \varphi(\mathbf{x}) = \frac{1}{2}\Big(& a_{11}x_1^2 + a_{12}x_1x_2 + ... + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + ... + a_{2n}x_2x_n \\ & .... \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + ... + a_{nn}x_n^2\Big). \end{aligned} \tag{3a}$$

С непосредствена проверка чрез диференциране на (3а) може да се установи, че е изпълнено равенството

$$\nabla\varphi = \frac{\partial\varphi}{\partial\mathbf{x}} = \begin{Bmatrix} \dfrac{\partial\varphi}{\partial x_1} \\ \dfrac{\partial\varphi}{\partial x_2} \\ \vdots \\ \dfrac{\partial\varphi}{\partial x_n} \end{Bmatrix} = \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^T\right)\mathbf{x}\,. \tag{4}$$

Ако **A** е симетрична, т.е. $\mathbf{A}^T = \mathbf{A}$, то равенство (4) добива вида

$$\nabla\varphi = \frac{\partial\varphi}{\partial\mathbf{x}} = \mathbf{A}\mathbf{x} \qquad \text{и} \qquad \frac{\partial^2\varphi}{\partial x_i \partial x_j} = a_{ij} \text{ за } \forall i, j\,. \tag{5}$$

Ако $\mathbf{A} \equiv \mathbf{I}$, то $\varphi(\mathbf{x}) = \mathbf{x}^T\mathbf{x}$ и

$$\nabla\varphi = \frac{\partial\varphi}{\partial\mathbf{x}} = \mathbf{x}\,. \tag{6}$$

Нека $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$ и $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}^T$. Разглеждаме скалара

$$\psi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y}\,, \tag{7}$$

където матрицата **A** не зависи от **x** и **y**. Функцията (7) се нарича ***билинейна форма***. С непосредствена проверка може да се установи, че за (7) е в сила

$$\frac{\partial\psi}{\partial\mathbf{x}} = \mathbf{A}\mathbf{y}\,. \tag{8}$$

За да получим производната спрямо **y** записваме

$$\psi^T = \psi = \mathbf{y}^T\mathbf{A}^T\mathbf{x}\,.$$

Тогава съгласно (8) получаваме

$$\frac{\partial\psi}{\partial\mathbf{y}} = \mathbf{A}^T\mathbf{x}\,. \tag{9}$$

За частния случай, за който $\psi = \mathbf{x}^T\mathbf{y}$ следва, че

$$\frac{\partial\psi}{\partial\mathbf{x}} = \mathbf{y} \qquad \text{и} \qquad \frac{\partial\psi}{\partial\mathbf{y}} = \mathbf{x}\,. \tag{10}$$

Пример 1: Да се определи градиентът на функцията

$$\Pi(\mathbf{Z}) = \frac{1}{2}\mathbf{Z}^T\mathbf{K}\mathbf{Z} - \mathbf{Z}^T\mathbf{F}\,, \tag{11}$$

където **K** е симетрична матрица, а **F** е зададен вектор, който не зависи от **Z**.

Първото събираемо на (11) квадратична форма, а второто събираемо е линейна форма на **x**. Съгласно (5) и (10) получаваме

$$\frac{\partial \Pi}{\partial \mathbf{Z}} = \mathbf{K}\mathbf{Z} - \mathbf{F} \,. \tag{12}$$

Ще припомним, че необходимото условие за екстремум на функцията $\Pi\left(\mathbf{Z}\right)$ е

$$\frac{\partial \Pi}{\partial \mathbf{Z}} = \mathbf{K}\mathbf{Z} - \mathbf{F} = 0 \qquad \text{или} \qquad \mathbf{K}\mathbf{Z} = \mathbf{F} \,. \tag{13}$$

Ако матрицата **K** е положително определена, то фукцията $\Pi\left(\mathbf{Z}\right)$ има минимум за онова **Z**, което удовлетворява (13).

Пример 2: Да се определи производната на функцията

$$u\left(\boldsymbol{\varepsilon}\right) = \frac{1}{2}\boldsymbol{\varepsilon}^T \mathbf{D}\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^T \mathbf{D}\boldsymbol{\varepsilon}_0 + \boldsymbol{\varepsilon}^T \boldsymbol{\sigma}_0$$

спрямо $\boldsymbol{\varepsilon}$.

По горните правила получаваме

$$\frac{\partial u}{\partial \boldsymbol{\varepsilon}} = \mathbf{D}\left(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_0\right) + \boldsymbol{\sigma}_0 \,.$$