ЛИНЕЙНА АЛГЕБРА **I**

(Линейни уравнения и системи; Метод на Гаус; Действия с матрици)

Уводни сведения

Основните математически средства, съставляващи този раздел на математиката, който обикновено наричаме “линейна алгебра”, са т.нар. системи линейни уравнения (където неизвестните участват само от първа степен – т.е. линейно), матрици и детерминанти.

Един *мотивационен пример от областта на рационалното хранене*:

Да разгледаме следния проблем (примерът е заимствуван от [FM]), който по същество е практически проблем за съставяне на елементарен “*линеен модел”* . Трябва да се композира хранителна смес от три компоненти – царевично зърно, соево зърно и памукови семена (за краткост: зърно, соя и семена). От хранителните характеристики на комппонентите се знае, че 1 единица (наприме в грамове или килограми) от всяка компонента произвежда протеин, мазнини и влакна в количества – части от единицата, дадени в следната таблица

Таблица 1

	Зърно	Соя	Семена
Протеин	0.25	0.4	0.2
Мазнини	0.4	0.2	0.3
Влакна	0.3	0.2	0.1

Една доза от хранителната смес трябва да осигури 22 единици протеин, 28 единици мазнини и 18 единици влакна. Въпросът е по колко единици от всяка компонента трябва да вземем, за да съставим желаната смес. За да решим проблема, нека означим с x, y и z необходимите количества единици съответно от зърно, соя и влакна. Понеже общото количество протеин трябва да бъде 22 единици, следвайки първия ред на таблицата получаваме:

$$0.25x + 0.4y + 0.2z = 22.$$

Аналогично, от втория и третия ред на таблицата следва:

$$0.4x + 0.2y + 0.3z = 28$$
$$0.3x + 0.2y + 0.1z = 18.$$

Така стигнахме до следната система от три линейни уравнения с три неизвестни:

$$\left|\begin{array}{lllllll} 0.25x & + & 0.4y & + & 0.2z & = & 22 \\ 0.4x & + & 0.2y & + & 0.3z & = & 28\,. \\ 0.3x & + & 0.2y & + & 0.1z & = & 18 \end{array}\right. \qquad (1.a)$$

Нашият "*линеен модел*" по проблема за рационалното хранене, това е системата (1.а) и предхождата я таблица. Решението на (1.а) е: $x = 40,\ y = 15;\ z = 30$. С непосредствено заместване в (1.а) се проверява, че тройката числа $(x,y,z) = (40,15,30)$ е решение на системата. По-късно ще стане ясно, че (1.а) е от типа системи, имащи единствено решение.

Ако хранителната смес съдържа повече компоненти, например пет, в системата (1.а) ще имаме 3 уравнения и пет неизвестни. Ако за развитието на дадена порода животни, освен трите фактора (протеин, мазнини и влакна), съществен показател са и още два, а хранителната смес съдържа 3 компоненти, тогава в (1.а) ще имаме 5 уравнения и 3 неизвестни.

Редица от методите на линейната алгебра намират широко приложение в експерименталните науки, включително и в биологията. Тези методи са важна част от софтуерните пакети с приложни програми (ППП), които позволяват многостранна обработка на експерименталните данни. За да може да осъществи това "математическо експериментиране", биологът трябва да е запознат основно с три неща: с теорията, с методите за пресмятане (изчислителните алгоритми) и с конкретните особености на приложението им в използваната от него среда. Доброто разбиране на теорията е основа за успешно прилагане на числените методи, реализирани в многобройните днес ППП. За запознаване с функционирането на основни алгоритми от линейната алгебра, съществено допринася самостоятелното конструиране на техни изчислителни процедури (приложения). Този процес може значително да се улесни, ако се работи с някой от разпространените и до известна степен познати продукти - EXCEL, ACCESS и др., които разполагат с добра математическа библиотека, графичен интерфейс, възможности за конвертиране на формати и т.н. Използването на тези "технологии за обработка на данни" позволяват разностранен поглед върху експерименталния материал, натрупване на опит и по-добро разбиране поведението на изследвания биологичен обект. Освен това те са и стъпало към намиране на закономерности в това поведение, т.е. към създаването на математически модел на наблюдаваното биологично явление. Създаването на модел изобщо не е лесно, тъй като се базира на взаимното преплитане и допълване на два подхода към явлението - на биологичния и на математическия.

За част.от разгледаните по-нататък илюстративни примери и задачи е показано и решението им с компютър – например посредством програмата *EXCEL*.

§1. ЛИНЕЙНИ УРАВНЕНИЯ И СИСТЕМИ

Дефиниция 1.1. *Линейно уравнение* с *n неизвестни,* $\mathbf{x_1}$, $\mathbf{x_2}$, ,$\mathbf{x_n}$, наричаме равенство от вида:

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b$$

Числата $\boldsymbol{a_1, a_2, \dots ,a_n}$ наричаме *коефициенти*, а числото $\boldsymbol{b}$ – *свободен член* на уравнението.

Дефиниция 1.2 *Решение* на уравнението.(по горе.) наричаме такава наредена n-орка числа $(r_1, r_2, \dots ,r_n)$, която, заместена в уравнението, го удовлетворява; т.е. при $x_1 = r_1$, $x_2 = r_2, \dots ,x_n = r_n$, за числото $a_1r_1 + a_2r_2 + \dots + a_nr_n$, получено от лявата част на уравнението, имаме: $a_1r_1 + a_2r_2 + \dots + a_nr_n = b$.

Дефиниция 1.3 *Линейна система* (или *линейна алгебрична система*) от **m** уравнения с **n** неизвестни, $\mathbf{x_1, x_2, x_n}$ -- това е съвкупност от **m** линейни уравнения с **n** неизвестни от вида:

$$\begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1j}x_j + \cdots a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2j}x_j + \cdots a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ij}x_j + \cdots a_{in}x_n & = & b_i \\ \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mj}x_j + \cdots a_{mn}x_n & = & b_m \end{array} \qquad (1.1)$$

Числата a_{ij} са *коефициентите* на системата, а b_i – *свободните членове* ($i = 1,2, \dots,m; j = 1,2, \dots ,n$). С двойното номериране (посредством индексите i, j) в коефициента a_{ij} означаваме,че a_{ij} се намира в i-тото уравнение (i-тия ред), пред j-тото неизвестно x_j (т.е. в j-тата колона) на системата.

Ако всяко $\mathbf{b_i = 0}$, ($i = 1,2, \dots ,n$) системата се нарича *хомогенна,* а когато поне едно от числата $\mathbf{b_i \neq 0}$ -- нехомогенна.

Решение на системата наричаме такава наредена n-орка числа $(r_1, r_2, \dots ,r_n)$, която е решение на всяко от уравненията на системата.

Дефиниция 1.4. *Матрица* наричаме таблица от числа c_{ij} $(i = 1,2, \dots ,m; j = 1,2, \dots ,n)$, наредени в **m** реда и **n** стълба (колони), от вида (1.0.1) – по-долу. Числата c_{ij} наричаме *елементи* на матрицата. Индексите i, j (съответно първи, втори) означават, че елемента c_{ij} се намира в i-ти ред и j-ти стълб на матрицата. (Например, a_{37} е елементът от 3-ти ред и 7-ми стълб.) За матрицата C от (1.0.1) казваме, че е *от тип (mxn)* или че е *с размери (mxn)* (т.е. с m реда и n стълба) и за краткост използваме също означението $C = (c_{ij}), i=1,2, \dots ,m; j=1,2, \dots ,n$.

$$C = \begin{pmatrix} c_{11} & c_{12} & .... & c_{1n} \\ c_{21} & c_{22} & .... & c_{2n} \\ .... & .... & .... & .... \\ c_{m-1,1} & c_{m-1,2} & .... & c_{m-1,n} \\ c_{m1} & c_{m2} & .... & c_{mn} \end{pmatrix} \qquad (1.0.1)$$

Равенство на две матрици *A, B* (т.е. $A = B$), с елементи съответно a_{ij} , b_{ij} , имаме, когато матриците имат еднакъв брой (m) редове, еднакъв брой (n) стълбове и $a_{ij} = b_{ij}$, за всички i, j: $i=1,2, \dots ,m; j=1,2, \dots ,n$.

Матриците от типа (*1xn*) или (*mx1*) наричаме освен това *вектори* (вж. също (1.4), по-долу).

- *Квадратна матрица* имаме, когато размерите й– **n** и **m** са равни. Такава матрица наричаме също *n - мерна* Елементите й a_{ii} се наричат *диагонални*. (Елементите с еднакви индекси, $i = j$, можем да наричаме по същия начин и при неквадратни, т.е. изобщо *правоъгълни* матрици.)
- *Симетрична матрица* е онази, за която $a_{ij} = a_{ji}$. (за всички i, j: $i=1,2, \dots ,m; j=1,2, \dots ,n$).

Дефиниция 1.5. Матрицата $A = (a_{ij})$, $i = 1 \div m$; $j = 1 \div n$, с елементи коефициентите на системата (1.1), наричаме *матрица на системата,* а матрицата

$$A_b = \begin{pmatrix} a_{11} & a_{12} & .... & a_{1n} & b_1 \\ a_{21} & a_{22} & .... & a_{2n} & b_2 \\ .... & .... & .... & .... & .... \\ a_{m-1,1} & a_{m-1,2} & .... & a_{m-1,n} & b_{m-1} \\ a_{m1} & a_{m2} & .... & a_{mn} & b_m \end{pmatrix} \qquad (1.2)$$

от тип m x (n+1), образувана от коефициентите и свободните членове на (1.1) наричаме *разширена матрица на системата.*

Системата (1.1) ще запишем и по още един начин – в познатата форма “едно уравнение с едно неизвестно”. За целта с наредените n-торки от числа $\vec{x} = (x_1, x_2, ..., x_n)$, ще работим като с n-мерни вектори (вж. матриците от тип 1 x n); ако $\vec{x}$ е един такъв вектор, числата $x_1, x_2, ..., x_n$ са негови координати (или компоненти) – съответно *първа, втора* и т.н., n-та. Очевидно горната идея за n-мерни вектори е заимствана от познатите ни от аналитичната геометрия случаи на двумерни и тримерни вектори. Да припомним, че ако $\vec{a}$ е даден тримерен вектор и $\vec{e}_1, \vec{e}_2, \vec{e}_3$ са тройка например взаимно перпендикулярни вектори (задаващи всъщност осите на декартова координатна система в тримерното пространство), то $\vec{a}$ “се разлага” по (или представя чрез) тройката (координатни) вектори $\vec{e}_1, \vec{e}_2, \vec{e}_3$ по формулата:

$$\vec{a} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3, \qquad (1.3)$$

където (x_1, x_2, x_3) е наредената тройка числа от координатите на вектора $\vec{a}$ (относно координатната система $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$; в такъв случай за вектора $\vec{a}$ вместо (1.3) пишем $\vec{a} = \vec{a}(x_1, x_2, x_3)$ или $\vec{a} = (x_1, x_2, x_3)$ и имаме предвид, че наредената тройка числа (x_1, x_2, x_3) еднозначно задава вектора $\vec{a}$.

Да запишем n-мерния вектор $\vec{x} = (x_1, x_2, ..., x_n)$ като вектор-стълб, т.е. във вида

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix} \qquad (1.4)$$

Всъщност в (1.4) имаме матрица от тип (n x 1), т.е. матрица-стълб. В следващата дефиниция ще дадем едно правило за умножаване отдясно на матриците от тип (m x n) с матрица (n x 1). Произведението, получено като резултат от такава операция ще бъде също матрица-стълб, но от тип (m x 1), т.е. m-мерен вектор. Така при дадена матрица A от тип (m x n), на всеки n-мерен вектор $\vec{x}$ еднозначмо съпоставяме друг, m-мерен вектор $\vec{y}$, получен като произведение на A с вектора $\vec{x}$, т.е. $\vec{y} = A\vec{x}$.

Дефиниция 1.6. m-мерният вектор $\vec{y}$ с координати, определени по формулата

$$y_i \; = \; \sum_{j=1}^{n} a_{ij} x_j \, , i = 1, 2, ..., m \, , \qquad (1.5)$$

т.е. $y_i \; = \; a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n$ наричаме произведение на матрицата A от тип (m x n) с даден n-мерен вектор $\vec{x}$. В такъв случай пишем $\vec{y} \; = \; A\vec{x}$, т.е.

$$\begin{pmatrix} y_1 \\ y_2 \\ ... \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix}.$$

В горната фигура векторите $\vec{x}$ и $\vec{y}$ участват като матрици-стълб съответно от тип (n x 1) и (m x 1).

Сега можем да запишем системата (1.1) по следния по-кратък начин:

$$\begin{cases} \sum_{j=1}^{n} a_{1j} x_j & = & b_1 \\ \sum_{j=1}^{n} a_{2j} x_j & = & b_2 \\ .... & .... & ... \\ \sum_{j=1}^{n} a_{mj} x_j & = & b_m \end{cases} \qquad (1.6)$$

и да отбележим, че в дясната част на (1.6) стоят координатите на вектора $\vec{b} \; = \; (b_1, b_2, ..., b_m)$, записан като вектор-стълб. Пред вид (1.5), в лявата част на

(1.6) имаме наредената m-торка числа

$$(y_1, y_2, ..., y_m) \quad = \quad (\sum_{j=1}^{n} a_{1j}x_j, \sum_{j=1}^{n} a_{2j}x_j, ..., \sum_{j=1}^{n} a_{mj}x_j)$$

от координатите на вектора $A\vec{x}$ (записани в стълб). Т.е. в (1.6) имаме равенство между координатите на векторите $A\vec{x}$ и $\vec{b}$, което именно означава равенство между самите вектори. Така показахме, че системата (1.1) може да се запише в следната векторна (или векторно-матрична) форма:

$$A\vec{x} \quad = \quad \vec{b}. \qquad (1.7)$$

В (1.7) имаме едно (векторно) уравнение с едно (векторно) неизвестно $\vec{x}$; дадени са векторът $\vec{b}$ - свободен член и матричният коефициент *A*. Да означим с $\vec{0}$ нулевата m-торка числа (0,0,...,0), която е нулевият m-мерен вектор и наред с (1.7) да запишем съответстващото му *хомогенно* уравнение

$$A\vec{x} \quad = \quad \vec{0}. \qquad (1.8)$$

Уравнението (1.7) е изобщо *нехомогенно*, понеже векторът $\vec{b}$ е изобщо ненулев; в такива случаи пишем $\vec{b} \neq \vec{0}$, което означава, че поне една от координатите b_i $(i = 1,2,...,m)$ не е нула. По-късно ще се убедим, че решенията на по-сложната задача (1.7) са в съществена зависимост от решенията на по-простата (1.8).

1.1 Метод на Гаус –гаусова елиминация.

Един прост и универсален метод за пълно изследване (и в частност за решаване) на линейните системи е т.н. *метод на Гаус* - по името на немския математик K.F.Gauss (1777-1855). Една модификация на този метод се среща и като метод на Гаус – Жордан. Същината на метода е в прилагането, в максимално ефективна форма, на естественат идея (елиминация) на неизвестните.

Най-напред ще илюстрираме един директен подход за решаване на линейни системи.

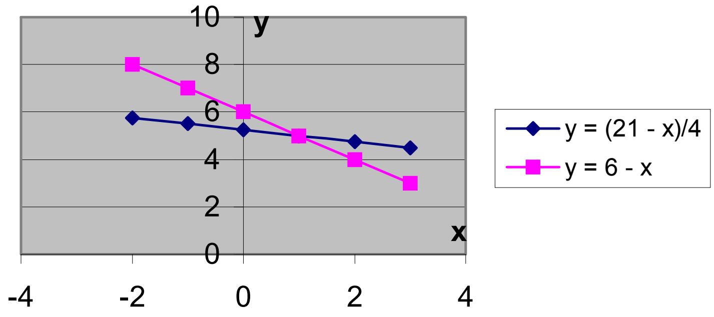
А. Графично решаване на линейна система.

От практична гледна точка, характерна за емпиричните науки, към проблема за решаването на линейни системи, е естествено да подходим по следния геометричен начин.

Пример 1. Нека е дадена следната линейна система от две уравнения:

$$\left| \begin{array}{lllll} x & + & y & = & 21 \\ x & + & y & = & 6 \end{array} \right. \quad \text{или} \quad \left| \begin{array}{lllll} y & = & (21 & - & x)/4 \\ y & = & 6 & - & x \end{array} \right.$$

Задачата тук е да намерим стойности за неизвестните x и y, които да удовлетворяват системата. Знаем, че всяко уравнение от вида ax + by + c = 0 е уравнение на права линия, а пресечната точка на двете линии (ако съществува) има координати x и y, които са търсените стойности на неизвестните. Двете прави от системата са представени на Фиг. 1. Правите могат да се пресичат, както е и в този случай, т.е. системата има едно решение, могат да са успоредни (системата няма решение), или да се сливат (системата има безброй решения).



Фиг. 1 Графично решение на Пример 1. Пресечната точка на двете прави е с координати x = 1, y = 5, които са решението на системата.

Графиката на правите е получена в EXCEL при следните данни - Фиг.2.:

C2 = =(21 - B2)/4

	A	B	C	D	E
1					
2		-3	6	9	
3		-2	5,75	8	
4		-1	5,5	7	
5		0	5,25	6	
6		1	5	5	
7		2	4,75	4	
8		3	4,5	3	
9		4	4,25	2	
10					

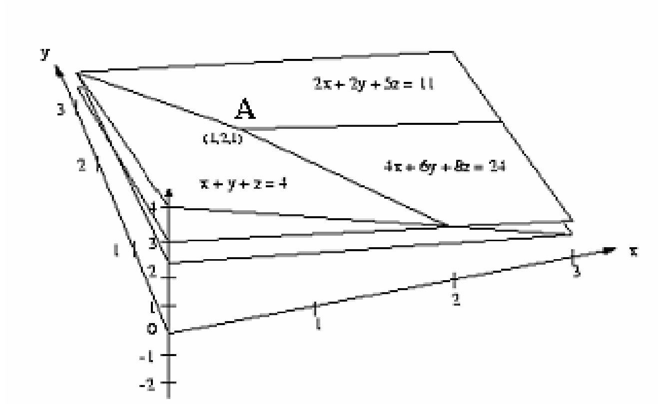
Фиг. 2. Част от работна EXCEL-страница със стойности на аргументите (колона B) и на двете линейни функции (колони C и D), графично представени на Фиг.1.

Тук [C2]: = (21 - B2)/4 (виж командния ред), C2=>6; [D2]: = 6 - B2; D2 =>9. Това задаване на уравненията на правите позволява координатите y да бъдат получавани автоматично като функции на съответните стойности на x от колона B.

Пример 2. Да се реши следната система от три уравнения: (Примерът е взет от Web-страницата http://www.math.unl.edu/ *Thomas Shores, Applied Linear Algebra and Matrix Analysis, May, 2000*).

$$\left|\begin{array}{l} x + y + z = 4 \\ 2x + 2y + 5z = 11 \\ 4x + 6y + 8z = 24 \end{array}\right.$$

Геометричното представяне тук е затруднено, тъй като общото уравнение $ax + by + cz + d = 0$ е уравнение на равнина. Все пак можем да отбележим, че аналогично на горния случай, ако съществува една обща точка за трите равнини (в която се пресичат трите им пресечници), то тя е решение на системата и координатите й x, y и z са стойности, удовлетворяващи и трите уравнения. На Фиг.3 това е точката от графиката с координати **A** (1,2,1)



Фиг. 3 Графично решение на Пример 2. Пресечната точка **A** на трите пресечници е с координати x = 1, y = 2 и z = 1, които са решения на системата.

Б. Аналитичен подход.

Разгледаните примери показват, че геометричният начин на решаване на линейни системи уравнения има известни недостатъци, макар че нагледно илюстрира смисъла на решението, както и взаимното разположение на равнините от третия пример, при което се получават различния брой решения. Ще се убедим, че принципният метод за решаване на линейни системи уравнения е вече споменатия метод на Гаус с редица варианти. Идеята за елиминацията (по Гаус) почива на т.нар. *елементарни преобразувания* на линейната система уравнения,а заменящи изходната линейна система с друга, еквивалентна на нея, от която лесно може да се получат решенията на системата (когато съществуват). Въпроса за еквивалентност на две линейни системи ще разгледаме малко по-систематично:

Дефиниция 1.7. Две линейни уравнения наричаме *еквивалентни,* ако всяко решение на едното е решение и на другото или и двете нямат решение.

Очевидно ако умножим едно уравнение с число различно от нула, ще получим уравнение еквивалентно на даденото.

Дефиниция 1.8. Две линейни системи наричаме *еквивалентни ,* ако всяко решение на едната е решение и на другата или и двете нямат решение.

Дефиниция 1.9. Следните операции наричаме *елементарни преобразувания* на дадена линейна система:

- смяна местата на две уравнения от системата;
- умножаване на уравнението с число различно от нула;
- умножаване на едно уравнение с произволно число и прибаване към друго уравнение.

Дефиниция 1.10. Следните преобразувания наричаме *елементарни преобразувания на матрица*:

- смяна на местата на два реда (стълба);
- умножаване на ред (стълб) с произволно число, различно от нула;
- умножаване на ред (стълб) с произволно число и прибавяне към друг ред (стълб).

Непосредствено се вижда, че извършването на елементарни преобразувания в дадена линейна система означава да извършим такива преобразувания по редове в разширената матрица на системата.

Принцип за еквивалентност.

Ако една линейна система е получена от друга чрез краен брой елементарни преобразувания, двете системи са еквивалентни.

Валидността на този принцип следва непосредствено от дадените по-горе дефиниции за еквивалентни уравнения и системи и дефинициите за решения.

За едно уравнение от дадена линейна система ще казваме, че е *следствие* от останалите, ако редът от разширената матрица на системата, състоящ се от коефициентите и свободния член на уравнението е *линейна комбинация* на редове измежду останалите редове на разширената матрица. Последното понятие ще поясним чрез пример: Ако (a_{ij}) е матрица от тип 10×4 (за конкретност), например нейният пети ред $(a_{51}, a_{52}, a_{53}, a_{54})$ ще е линейна комбинация (например) на втори, седми и девети ред, ако съществуват такива три числа *p, q, r* , че елементите от пети ред да удовлетворяват следните равенства:

$$a_{5j} = pa_{2j} + qa_{7j} + ra_{9j}, \quad j = 1, 2, 3, 4 .$$

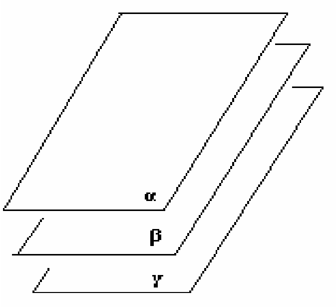
В такъв случай можем да казваме, че петият ред е линейна комбинация на втори, седми и девети ред, с коефициенти съответно *p, q* и *r*.

Не е трудно да се убедим, че ако от една линейна система премахнем уравненията, които са следствие от останалите, новополучената система ще е еквивалентна с дадената.

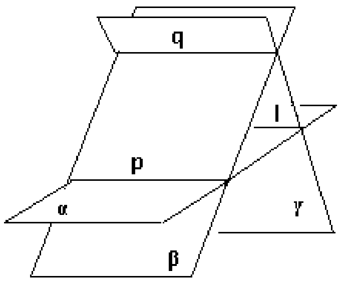
B. Геометричен анализ на линейни системи.

Тук ще считаме за известно (от аналитичната геометрия на тримерното пространство), че всяко уравнение от вида

$$ax+by+cz+d = 0, \qquad (1.9)$$



Фиг.4 (вж. текста)



Фиг.5

където поне един от коефициентите *a,b,c* не е нула, е уравнение на равнина в пространство. Това означава, че при зададена координатна система в тримерното пространство геометричното място на всички точки $M(x,y,z)$, чиито координати x,y,z удовлятворяват съотношението (1.9), е равнина.

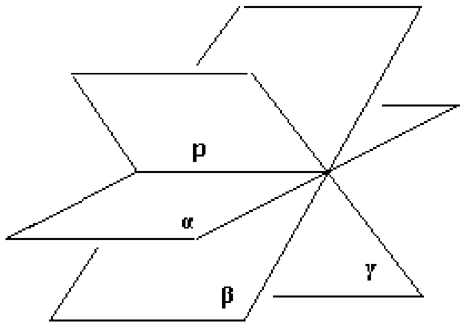
Ще изследваме от геометрични съображения решенията на системи от тип 3 х 3, т.е. от вида

$$\left| \begin{array}{ccccccc} a_{11}x & + & a_{12}y & + & a_{13}z & = & b_1 \\ a_{21}x & + & a_{22}y & + & a_{23}z & = & b_2 \\ a_{31}x & + & a_{32}y & + & a_{33}z & = & b_3 \end{array} \right. , \qquad (1.10)$$

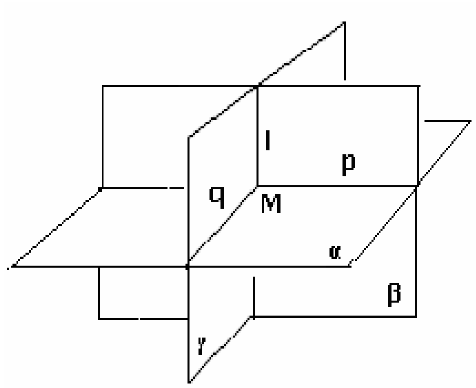
където поне един от коефициентите a_{i1}, a_{i2}, a_{i3} не е нула, $\forall\ i = 1, 2, 3$. Геометричната гледна точка, която ще илюстрираме в частния случай на системите (1.10), остава обаче общовалидна.

Предвид геометричната същност на (1.9), трите уравнения на системата (1.10) задават тройка равнини α, β, γ в пространството. Очевидно за равнините са възможни следните случаи за взаимно разположение:

- Равнините нямат обща точка (такива са случаите от фиг.4 - успоредни равнини и фиг.5 – всеки две от равнините се пресичат, но пресечниците им p, q, l са успоредни прави); системата (1.10) *няма решение*.
- Равнините имат единствена обща точка (фиг.7 - пресечниците p, q, l се пресичат в т. M); системата има единствено решение.
- Равнините имат обща права (фиг.6); системата има безбройно много решения, зависещи от един свободен параметър.
- Равнините съвпадат; системата има безбройно много решения, зависещи от два свободни параметъра.



Фиг.6



Фиг.7

Случаите на безбройно много решения на дадена линейна система по-нататък ще поясним допълнително посредством примери.

В заключение по повод общите бележки за метода на Гаус ще добавим, че по същество той се състои в систематичното прилагане на елементарните преобразувания с цел трансформиране на изходната матрица до триъгълна. Осъществява се в *прав* и *обратен* ход, основните алгоритми за които ще разгледаме по-долу.

1.2. Метод на Гаус – последователност на операциите.

Същината на идеята, приложена от Гаус, познаваме от училищната математика, където за дадена система от две уравнения,

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 ; \\ a_{21}x + a_{22}y &= b_2 , \end{aligned} \qquad (2.1)$$

постъпваме по следния начин: ако например a_{11} не е 0, решаваме първото уравнение относно x, след което заместваме във второто и така получаваме по-проста система от вида

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 ; \\ a'_{22}y &= b'_2 , \end{aligned} \qquad (2.2)$$

готова за окончателно решаване. Същия резултат ще получим и ако умножим първото уравнение в (2.1) почленно с $(-1)a_{21}/a_{11}$ и така умножено го прибавим към второто. За да преобразуваме системата по този начин, всъщност е достатъчно да преобразуваме нейната разширена матрица. На диаграмата по-долу е показан преходът от старата към новата матрица.

§2. ДЕЙСТВИЯ С МАТРИЦИ.

Да означим с $\mathbf{M}_{m,n}$ множеството на матриците от тип m x n (m реда и n стълба). За елементите на това множество ще въведем някои основни операции.

а) Линейни операции.

Ще започнем с операциите “умножение с число” и “сума на два елементита”. Ако $A, B \in \mathbf{M}_{m,n}$ са произволни матрица и p е произволно реално число, по дефиниция:

(L1) pA (произведение на числото p с матрицата A) е матрицата с елементи произведенията на числото с елементите a_{ij} на A; т.е. $(pA)_{ij} = pa_{ij}, i=1 \div m, j=1 \div n$;

(L2) $A + B$ (сума на матриците A и B) е матрицата, получена след поелементно сумиране на A и B ; т.е. $(A+B) = a_{ij} + b_{ij}, i=1 \div m, j=1 \div n.$

Веднага се вижда, че операцията “сума” е комутативна и асоциативна, т.е. $A + B = B + A$ и $(A + B) + C = A + (B + C)$ – за произволни два (съответно три)елемента на $\mathbf{M}_{m,n}$. Освен това, за “умножението с число” са в сила свойствата асоциативност и дистрибутивност (в съчетание с операцията “сума”), т.е.

$$p(qA) = (pq)A, \forall \text{ двойка числа } p, q, \forall A \in \mathbf{M}_{m,n} \text{ и}$$
$$p(A+B) = pA + pB, \forall \text{ число } p, \forall \text{ двойка } A, B \in \mathbf{M}_{m,n}.$$

Нулевата матрица O (всички елементи на която са *0*) играе ролята на “*неутрален*” (*нулев*) елемент в множеството $\mathbf{M}_{m,n}$, в смисъл че $\forall A \in \mathbf{M}_{m,n}$ имаме $A + O = A$ и $pO = O$, $\forall$ число p. При това, относно нулевата матрица, за всяка матрица $A \in \mathbf{M}_{m,n}$ съществува нейна *противоположна* A' : $A + A' = O$; очевидно е, че противоположната на дадена матрица A е единствена и това е матрицата $-A = (-1)A$.

Снабдено с горните две операции, (L1) и (L2) (такива операции ще наричаме *линейни*), относно които съществува нулев елемент и те притежават споменатите свойства (комутативност, асоциативност и пр.), множеството $\mathbf{M}_{m,n}$ - от матриците $A(m$ x $n)$ се оказва “*линейно пространство*”.

б) Произведение (умножение) на матрици.

Действието умножение на матрици ще въведем по правилото “ред по стълб”, представляващо естествено разширение на познатото ни от §1 правило за умножаване на матрица с вектор-стълб.

Ако $A \in \mathbf{M}_{m,n}$ и $B \in \mathbf{M}_{m,n}$ са произволна двойка матрици, под AB (произведение на A и B, A – първи, B – втори множител) по дефиниция разбираме матрицата C (пишем $C = AB$) с елементи c_{ij}, определени по следния начин:

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}, i = 1 \div m, j = 1 \div k \text{ (т.е. } (AB)_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}, i = 1 \div m, j = 1 \div k \text{).}$$

Очевидно произведението AB е елемент от $\mathbf{M}_{m,k}$.

Забележка: Ако $A \in \mathbf{M}_{m,n}, B \in \mathbf{M}_{n,m}$, то съществуват и двете произведения AB и BA, които очевидно (при $n \neq m$) не съвпадат, понеже $AB \in \mathbf{M}_{m,m}$, а $BA \in \mathbf{M}_{n,n}$. Нещо повече, дори в случая $A \in \mathbf{M}_{n,n}, B \in \mathbf{M}_{n,n}$ (тогава AB и BA са също от класа $\mathbf{M}_{n,n}$), прости примери показват, че $AB \neq BA$.

Пример

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix},\ B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}:\ AB = \begin{pmatrix} 1 & 5 & 0 \\ 0 & 2 & 0 \\ 3 & 3 & 3 \end{pmatrix},\ BA = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 2 & 0 \\ 9 & 0 & 3 \end{pmatrix}.$$

Очевидно $AB \neq BA$. (Да припомним, вж. §1, че две матрици $A' = (a'_{ij}),\ A'' = (a''_{ij})$ съвпадат, ако и двете са от един и същи тип ($m \times n$) и $a'_{ij} = a''_{ij}, \forall i = 1 \div m,\ j = 1 \div n$.)

В допълнение ще отбележим, че произведението на матрици е *асоциативно* и *дистрибутивно,* т.е. $A(BC) = (AB)C$ и $A(B + C) = AB + AC$, също $(A + B)C = AC + BC$, за всяка тройка матрици A,B,C от съответните класове $\mathbf{M}_{k,l}$.(проверете самостоятелно чрез конкретни примери!).

в) Транспонирана матрица.

Ако редовете на дадена матрица A запишем като стълбове (а стълбовете й - като редове), получената матрица наричаме *транспонирана* на A и обикновено я означаваме с A^t (или A^T); използват се също означенията A^* или A'. Ако a^t_{kl} е произволен елемент на A^t, то $a^t_{kl} = a_{lk}$, където a_{lk} е елемент на A. Следователно ако $A \in \mathbf{M}_{m,n}$, то $A^t \in \mathbf{M}_{n,m}$. С непосредствена проверка се вижда, че въведените по-горе основни операции се съгласуват с действието *"транспониране"* по следния начин:

- $(pA)^t = pA^t$, $\forall$ число p, $\forall$ матрица A;
- $(A + B)^t = A^t + B^t$, $\forall$ двойка $A, B \in \mathbf{M}_{m,n}$;
- $(AB)^t = B^t A^t$, $\forall$ двойка $A, B : A \in \mathbf{M}_{m,n}, B \in \mathbf{M}_{n,k}$.

2.1. Матрична алгебра.

Операциите събиране и умножаване на матрици позволяват да говорим за "матрична алгебра" (от аналогията с множеството на реалните числа, където имаме две основни операции). Взаимодействието на двете операции става пълноценно в множества от квадратни матрици (т.е. такива с $n = m$). Класовете $\mathbf{M}_{n,n}$ означаваме за краткост като $\mathbf{M}_n$. (За квадратните матрици $A \in \mathbf{M}_n$, казваме че са $n-$*мерни* или че са *от ред n* -- вж. §1.) Относно действието умножание класовете $\mathbf{M}_n$ имат допълнителното предимство, че удобно се въвежда полезното за редица цели понятие *степен на матрица:* $A^k = A.A....A$ (k пъти).

За да въведем съществения за умножението на матрици "неутрален" елемент – т.н. *единична матрица,* ще се спрем на понятието

Диагонална матрица: за дадена матрица $A \in \mathbf{M}_n$, с елементи a_{ij}, за елементите ѝ a_{ii}, $i = 1 \div n$, казваме също (вж. т. 1), че образуват *главния диагонал* на матрицата. Когато всички недиагонални елементи (a_{ij}, $i \neq j$) са нули, матрицата наричаме *диагонална.* Ако D е диагонална матрица с диагонални елементи $d_1, d_2, ..., d_n$, често се използва означението $D = diag(d_1, d_2, ..., d_n)$. От тук е ясно, че за степените на D, т.е. D^2, D^3 и т.н. имаме съответно $D^2 = diag(d_1^2, d_2^2, ..., d_n^2)$, $D^3 = diag(d_1^3, d_2^3, ..., d_n^3)$ и т.н.

Единична матрица: това е диагоналната матрица (от фамилията $\mathbf{M}_n$) с диагонални елементи = 1. Тази матрица ще означаваме с I (или E); т.е. $I = diag(1,1,...,1)$.

Непосредствено се проверява, че I е както *ляв,* така и *десен "неутрален"* елемент в $\mathbf{M}_n$ (относно умножението), в смисъл че $AI = A$ и $IA = A$, $\forall A \in \mathbf{M}_n$.

Обратима матрица, обратна матрица: Една матрица $A \in \mathbf{M}_n$ наричаме *обратима,* ако съществува такава друга матрица $A_1 \in \mathbf{M}_n$, че да е в сила някое от равенствата: $A_1 A = I$ или $AA_1 = I$.

Лесно се проверява, че ако за две матрици $A_1', A_1'' \in \mathbf{M}_n$ имаме $A_1' A = I$ и $AA_1'' = I$, те са равни; наистина, ако умножим отдясно първото от тези равенства с A_1'', получаваме: $(A_1' A).A_1'' = A_1''$, т.е. (от асоциативността) $A_1'(AA_1'') = A_1''$, следователно $A_1' = A_1''$ (понеже $AA_1'' = I$ и $A_1' I = A_1'$). Така установихме, че *ако* $A \in \mathbf{M}_n$ *е обратима матрица , за нея съществува такава единствена матрица* (*която ще означаваме с* A^{-1}), *че* $A^{-1}A = AA^{-1} = I$. Матрицата A^{-1} наричаме *обратна* на A.

Посредством понятието "детерминанта" на матрица (вж. съответния раздел, по-нататък) ще имаме един удобен критерий за обратимост на дадена матрица. Предвид това, обратимите матрици ще наричаме също *неособени* или *неизродени.*

За обратимите матрици е в сила следното *твърдение*: Ако $A \in \mathbf{M}_n$ е обратима, системата $A\vec{x} = \vec{0}$ ($\vec{x} \in R^n$ – множеството на т.н. *n – мерни вектори,* т.е. всевъзможните наредени n – орки реални числа) има само нулевото решение. За да се убедим във валидността на твърдението, да умножим горното уравнение почленно отляво с A^{-1}: $A^{-1}.(A\vec{x}) = A^{-1}.\vec{o}$, т.е. $(A^{-1}A)\vec{x} = \vec{o}$, т.е. $\vec{x} = \vec{o}$ (предвид асоциативността на умножението и че $A^{-1}A = I$, а $A^{-1}\vec{o} = \vec{o}$) . С други думи, установихме, че ако $\vec{x}$ е решение на системата $A\vec{x} = \vec{o}$, то е нулевият вектор. По-нататък ще се убедим, че е в сила и обратното твърдение: ако за дадена матрица $A \in \mathbf{M}_n$ системата $A\vec{x} = \vec{o}$ има само нулевото решение, A е обратима.

Ако $A, B \in \mathbf{M}_n$ са обратими матрици, то и произведенията AB, BA също са обратими матрици: наистина, матрицата $B^{-1}A^{-1}$ е обратната на AB, понеже $(B^{-1}A^{-1}).(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I$, т.е. $\exists(AB)^{-1}$ и $(AB)^{-1} = B^{-1}A^{-1}$; аналогично, $\exists(BA)^{-1}$ и $(BA)^{-1} = A^{-1}B^{-1}$.

Ако $A \in \mathbf{M}_n$ е обратима, *то и транспонираната* A^t *е обратима матрица,* като:
$(A^t)^{-1} = (A^{-1})^t$: от $A^{-1}A = I \Rightarrow (A^{-1}A)^t = I^t = I$, но $(A^{-1}A)^t = A^t(A^{-1})^t$, т.е. $A^t(A^{-1})^t = I$, с което валидността на това свойство е доказана.

В заключение, предвид изложените по-горе свойства, за всяко от множествата $\mathbf{M}_n$ можем да казваме, че имаме *матрична алгебра* (т.е. *алгебра на матриците от ред n*), при това -- *некомутативна* (понеже в общия случай $AB \neq BA$).

2.2. Ранг на матрица.

Линейно зависими редове (*стълбове*) *на матрица:* Нека $A \in \mathbf{M}_{m,n}$ е матрица с елементи a_{ij}. Както знаем, нейният i – ти ред има вида

$$X = (a_{i1}, a_{i2}, ..., a_{in}). \qquad (2.1)$$

Ако дадем на индекса i по произволен начин k различни стойности $(k \leq m)$ $i_1, i_2, ..., i_k$, ще получим една произволна k – орка редове (от всичките m на брой), които за удобство ще означим с $X_1, X_2, ..., X_k$. Така имаме k на брой наредени n – орки числа, а такива n – орки разглеждаме като n – мерни вектори (вж. т.2.1., по-горе). *Координатите* (*компонентите*) -- 1-ва, 2-ра, ..., n-та – на всеко такъв вектор са числата от съответната n – орка, взети по реда на номерата им. Например за вектора X от (2.1) числата $a_{i1}, a_{i2}, ..., a_{in}$ са неговите съответно -ва, 2-ра, и т.н. n-та координата (компонента). За допълнителна простота, векторите – редове X_l ще презапишем като

$$X_l = (x_l^1, x_l^2, ..., x_l^n),\ l = 1 \div k, \qquad (2.2)$$

където $x_l^j = a_{ij}$, при $i = i_l$.

Дефиниция: Редовете $X_1, X_2, ..., X_k$. наричаме *линейно зависими,* ако съществува ненулева k – орка числа $c_1, c_2, ..., c_k$ (т.е. поне едно от числата не е нула), така че

$$c_1X_1 + c_2X_2 + ... + c_kX_k \ = \ \vec{o}, \qquad (2.3)$$

където $\vec{o} = (0,\ 0,\ ...,\ 0)$ (n на брой нули).

Сума от вида $\sum_{j=1}^{k} c_j X_j$ (вж. лявата част на (2.3)) наричаме *линейна комбинация* на съответните елементи (в случая – редовете $X_j,\ j = 1 \div k$) с коефициенти числата c_j $(j = 1 \div k)$.

Коментар: Равенството (2.3) означава че са в сила n на брой “поелементни” равенства, т.е.

$$c_1x_1^j + c_2x_2^j + ... + c_kx_k^j \ = 0,\ \ j = 1 \div n.$$

Дефиницията на *линейно зависими стълбове* е аналогична.

Даден брой редове (стълбове), които не са линейно зависими, наричаме *линейно независими.* Това означава, че $X_1, X_2, ..., X_k$. са линейно независими редове, ако единствената k – орка числа $c_1, c_2, ..., c_k$, за която (2.3) важи, е нулевата (всички c_j, $j = 1 \div k$, са нули).

<u>Дефиниция</u> (Ранг) Максималния брой линейно независими редове (стълбове) на дадена матрица $A \in \mathbf{M}_{m,n}$ наричаме *ранг на A.*

Тук непосредствения въпрос е: Дали максималният брой линейно независими редове съвпада с този при стълбовете на *A?* Въпросът има положителен отговор, в което се убеждаваме по следния начин. Нека k и k^* са споменатите максимални стойности съответно за редове и стълбове, и да предположим, че $k > k^*$. Избираме k на брой ($k \le m$) линейно независими редове, които запазваме, а останалите редове отстраняваме от матрицата *A.* От получената ("остатъчна") матрица отстраняваме последните $n - k$ на брой стълбове, ако $k < n$ или добавяме $k - n$ на брой линейно независими стълбове, ако $k > n$. Да означим с A_o окончателно получената матрица; тя е квадратна – от тип *k x k.* Нека $Y_1, Y_2, ..., Y_k$ са стълбовете на A_o. Тъй като $k > k^*$ (по предположение), не може всичките, k на брой, стълбове на A_o да са линейно независими: в противен случай бихме имали повече от максималния брой k^* линейно независими стълбове на "старата" матрица *A,* което е абсурд. Следователно съществува ненулева k – орка числа $c_1, c_2, ..., c_k$, така че

$$c_1Y_1 + c_2Y_2 + ... + c_kY_k \; = \; \vec{o}. \qquad (2.4)$$

Ако $\vec{c}$ е вектор-стълбът с компоненти $c_1, c_2, ..., c_k$, от правилото за умножаване отдясно на матрица с вектор-стълб веднага се вижда, че

$$c_1Y_1 + c_2Y_2 + ... + c_kY_k \; = \; A_0\vec{c}$$

и (2.4) презаписваме като

$$A_0\vec{c} = \vec{o} \, . \qquad (2.5)$$

Матрицата A_o има линейно независими редове ("наследени" от изходната матрица *A*) и е квадратна. По-нататък ще установим, че такива матрици са обратими. Тогава от даденото по-горе твърдение за обратими матрици, заключаваме, предвид (2.5), че $\vec{c} = \vec{o}$; т.е. всички константи $c_1, c_2, ..., c_k$ са нули, което противоречи на даденото в (2.4). Полученото противоречие означава, че не е възможно $k > k^*$.

Когато $k < k^*$, разглеждаме транспонираната матрица A^t и прилагаме вече разгледания случай. С това установихме, че $k = k^*$.