MÉSZÁROS JÓZSEFNÉ,

NUMERICAL METHODS

5

A Műszaki Földtudományi Alapszak tananyagainak kifejlesztése a
TÁMOP 4.1.2-08/1/A-2009-0033 pályázat keretében valósult meg.

V. MATHEMATICAL NOTIONS

DEFINITION 1

The set $L=\{x,y,z,\dots\}$ is called **linear vector space** if the addition and the multiplication by real number are defined on the set L and these operations have the following properties:

$$x+y=y+x$$

$$(x+y)+z=x+(y+z)$$

$$\exists 0\in L \quad , \quad \text{that} \quad 0+x=x \qquad \text{(0 is called zero element)}$$

$$\exists -x\in L \quad , \quad \text{that} \quad x+(-x)=0 \qquad (-x \text{ is called opposite element})$$

and

$$\alpha(\beta x)=(\alpha\beta)x$$

$$1\cdot x=x$$

$$(\alpha+\beta)x=\alpha x+\beta x$$

$$\alpha(x+y)=\alpha x+\alpha y \quad .$$

The elements of such a set L are called vectors.

Examples

a) Let

$$L=\{(a_1,a_2,\dots,a_n)\ ,\ (b_1,b_2,\dots,b_n)\ ,(c_1,c_2,\dots,c_n)\ ,\dots\} \quad ,$$

where $a_i,b_i,c_i,\dots$ are real numbers, that is, L is the set of such "objects" which have n "coordinates". If the addition and multiplication by number are defined by the traditional rules, that is,

$$(a_1,a_2,\dots,a_n)+(b_1,b_2,\dots,b_n)=(a_1+b_1,a_2+b_2,\dots,a_n+b_n)$$

$$\alpha(a_1,a_2,\dots,a_n)=(\alpha a_1,\alpha a_2,\dots,\alpha a_n),$$

then the mentioned properties can be proved easily. This linear space is usually denoted by $\Re^n$.

b) Let the interval $[a,b]$ be given and denote L the set of the continuous functions on $[a,b]$. If the addition

$f(x)+g(x)$ and the multiplication by number $\alpha f(x)$ are defined by the traditional rules, then the required properties can be proved easily. This linear space is usually denoted by $\mathbb{C}[a,b]$.

DEFINITION 2

The elements $x, y, \ldots, w \in L$ (where *L* is a vector space and the number of the given vectors is finite) form a **linearly independent system** if

$$\alpha x+\beta y+\ldots+\lambda w=0 \quad \Rightarrow \alpha=\beta=\ldots=\lambda=0.$$

Examples

a) In the vector space $\Re^4$ the vectors $a=(1,0,0,0)$ and $b=(0,0,1,0)$ form a linearly independent system, since the zero element $(0,0,0,0)$ can only be obtained in the form $0 \cdot a+0 \cdot b$.

b) In the vector space $\mathbb{C}[a,b]$ the functions $x, x^2\ x^3$ form a linearly independent system, since the zero element (the identically zero function) can only be obtained in the form $0 \cdot x+0 \cdot x^2+0 \cdot x^3$.

DEFINITION 3

The linearly independent system formed by the vectors $x, y, \ldots, w \in L$ is a basis of L, if arbitrary element of L can be obtained as a **linear combination of the elements** $x, y, \ldots, w$, that is, in the form

$$\alpha x+\beta y+\ldots+\lambda w.$$

Examples

a) In the linear space $\Re^3$ the vectors $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ form a well-known basis. Another possible basis is given e.g. by vectors $(2,0,0)$, $(0,1,0)$, $(0,0,1)$.

b) Denote P the set of the polynomials of degree at most 2. The traditional addition and multiplication by number do not lead out from the set, P is a vector space.

The elements (functions, vectors) $1, x, x^2$ form a basis of P, because they form a linearly independent system and at most quadratic arbitrary polynomial can be obtained in the form

$$\alpha \cdot 1+\beta \cdot x+\gamma x^2.$$

Another basis is given e.g. by elements $1, x-1, x^2$.

DEFINITION 4

If we can order real numbers to all elements $x, y, \ldots$ of the linear space L

$$\|x\|, \|y\|, \ldots$$

respectively in such a way that

$$\|x\| \geq 0 \quad , \qquad \|x\| = 0 \Rightarrow x = 0$$

$$\|x + y\| \leq \|x\| \cdot \|y\|$$

$$\|\alpha x\| = |\alpha| \cdot \|x\| \quad ,$$

then the vector space L is called **normed space** and the real number $\|x\|$ is called the **norm of the vector** x.

Examples

a) Define the quantities

$$\|\mathbf{a}\|_1 = |a_1| + |a_2| + \ldots + |a_n|$$

$$\|\mathbf{a}\|_2 = (a_1^2 + a_2^2 + \ldots + a_n^2)^{1/2}$$

$$\|\mathbf{a}\|_\infty = \max_{1 \leq i \leq n} |a_i|$$

for each element $a = (a_1, a_2, \ldots, a_n)$ of the linear space $\Re^n$. These quantities define three different norms which are called norm number 1, norm number 2 (or Euclidean norm) and norm number infinity.

b) Denote M_n the set of the real (square) matrices of order n. M_n form a linear space with the traditional addition and multiplication by number. The quantities

$$\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$$

$$\|\mathbf{A}\|_2 = \sqrt{\lambda^*},$$

where λ^* is the largest eigenvalue of the matrix $A^T A$ and

$$\|\mathbf{A}\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$$

define three different norms. The norms belonging to the same notation (the same name) in a) and b) are called compatible, because such pairs have also the properties

$$\|\mathbf{A}\mathbf{x}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{x}\|,$$

$$\|\mathbf{A}\| = \sup_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|$$

for any $x \in \Re^n$, $A \in M_n$ beside the properties appearing in the definition. By the latter property we could have defined the above-mentioned norms of the matrix A as the smallest upper bound of the suitable vector norm $\|Ax\|$ under the condition $\|x\| = 1$. In our lecture notes we supposed compatible norms when using $\|Ax\|$ every time.

c) In the vector space $C[a,b]$ we can define norms e.g. by equalities

$$\|f(x)\|_2 = \left(\int_a^b f^2(x)dx \right)^{1/2} \quad ,$$

$$\|f(x)\|_\infty = \max_{x \in [a,b]} |f(x)|.$$

Here we mention that the distance of the elements x, y of a normed space (just like in the traditional case) can be measured by the quantity

$$\|x + (-y)\| = \|x - y\|.$$

DEFINITION 5

If we can order a real number c to arbitrary two elements x, y of the linear space L in such a way that

$$(x, y) = (y, x)$$

$$(x + y, z) = (x, z) + (y, z)$$

$$(\alpha x, y) = \alpha(x, y)$$

$$(x, x) \geq 0 \quad , \qquad (x, x) = 0 \Rightarrow x = 0 \quad ,$$

then the vector space L is called **Euclidean space** and the real number (x, y) is called the scalar product of the elements (vectors) x and y.

Examples

a) Let $a = (a_1, a_2, \ldots, a_n)$ and $b = (b_1, b_2, \ldots, b_n)$ be two arbitrary elements of $\Re^n$ Then the quantity

$$ab = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n$$

defines the traditional scalar product on $\Re^n$.

b) If $f(x)$ and $g(x)$ are two arbitrary elements of $C[a,b]$, then the quantity

$$(f,g) = \langle f,g \rangle = \int_a^b f(x)g(x)dx$$

defines a scalar product.

DEFINITION 6

The matrix A of order n is called **positive definite matrix** if

$$x^T Ax > 0, \quad \forall x \in \Re^n, \quad x \neq 0.$$

DEFINITION 7

The matrix A of order n is called a **negative definite matrix** if

$x^T Ax < 0, \quad \forall x \in \Re^n, \quad x \neq 0.$

DEFINITION 8

The matrix A of order n is called a **positive semidefinite matrix** if

$x^T Ax \geq 0, \quad \forall x \in \Re^n.$

DEFINITION 9

The matrix A of order n is called a **negative semidefinite matrix** if

$x^T Ax \leq 0, \quad \forall x \in \Re^n.$

DEFINITION 10

Denote z the limit of the convergent sequence of real numbers $x_1, x_2, x_3, \ldots$ If there exist numbers C_1, C_2 such that for arbitrary k

$$C_1|x_k - z|^p \leq |x_{k+1} - z| \leq C_2|x_k - z|^p \quad ,$$

where $p \geq 1$, then we say: the sequence $\{x_k\}$ converges to z in order p. If only the left (right) side inequality is fulfilled then we say: **the order of convergence** is at most (at least) p .

Examples

a) For $p = 2$ (quadratic convergence), which often appears in numerical methods, we illustrate the rapidity of convergence on a simple example. Suppose that $C_2 = 1$ and the distance of x_k and z is 0.1. Then (by using the right side inequality):

$|x_{k+1} - z| \leq 10^{-2}$, $|x_{k+2} - z| \leq 10^{-4}$, $|x_{k+3} - z| \leq 10^{-8}$, ... which means a surprising rapidity compared to the traditional sequences.

b) If we use the vector sequence $x_1, x_2, x_3, \ldots$ $(x_i \in \Re^n)$ instead of the number sequence $x_1, x_2, x_3, \ldots$ $(x_i \in \Re)$ and we use a norm instead of absolute value, then the previous definition can be repeat with the inequality

$$C_1 \| x_k - z \|^p \leq \| x_{k+1} - z \| \leq C_2 \| x_k - z \|^p .$$

DEFINITION 11

Let $f(x), h(x) : [a,b] \subset \Re \Rightarrow \Re$ and $f(x), h(x) \geq 0, \ \forall x \in [a,b]$. Then the equality

$f(x) = \sigma(h(x))$

(where σ can be said as **capital ordo**) means there exists positive constant C such that

$$|f(x)| \leq C|h(x)|, \quad \forall x \in [a,b].$$

Examples

a) Let z be the limit of the sequence $x_1, x_2, x_3, \ldots$ $(x_i \in \Re)$. Then the equality

$$|x_k - z| = \sigma(q^k), \text{ where } 0 < q < 1,$$

means that the error of $x_k \leq Cq^k$, where C is independent on k. In other words, it means that "the error is proportional to q^k".

b) Let x_k be the previous sequence. Then the equality (supposed it is true)

$$|x_k - z| = \sigma(h^s)$$

(where $s \geq 1$ is a constant and $h = h(k) \to 0$ if $k \to \infty$) means that "the error is proportional to h^s".

c) Sometimes we use the sign σ for truncation of a polynomial. E.g. the equality

$p(x) = 2x^4 + \sigma(x^3)$, where $x \geq 0$,

means that

$|p(x) - 2x^4| \leq C(x^3)$,

where $C > 0$ is independent on x. Such an application of the sign σ can be found e.g. in elimination methods.

Digitális Egyetem, Copyright © Mészáros Józsefné, 2011