

Capítulo 1

Introducción

1. Vector Spaces

Definition 1.1. A vector space over the numeric field K ($K = \mathbb{R}$ or $K = \mathbb{C}$) is a nonempty set V , whose elements are called vector and in which two operations are defined, called addition and scalar multiplication, that enjoy the following properties:

1. addition is commutative and associative;
2. there exists an element $0 \in V$ (the zero vector or null vector) such that $v + 0 = v$ for each $v \in V$
3. $0 \cdot v = 0$, $1 \cdot v = v$, for each $v \in V$, where 0 and 1 are respectively the zero and the unity of K ;
4. for each element $v \in V$ there exists its opposite, $-v$, in V such that $v + (-v) = 0$;
5. the following distributive properties hold

$$\forall \alpha \in K, \forall v, w \in V, \alpha(v + w) = \alpha v + \alpha w$$

$$\forall \alpha, \beta \in K, \forall v \in V, (\alpha + \beta)v = \alpha v + \beta v$$

6. the following associative property holds

$$\forall \alpha, \beta \in K, \forall v \in V, (\alpha\beta)v = \alpha(\beta v)$$

Definition 1.2. We say that a nonempty part W of V is a vector subspace of V if W is a vector space over K .

Definition 1.3. A system of vector v_1, \dots, v_n of a vector space V is called linearly independent if the relation

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m = 0$$

with $\alpha_1, \alpha_2, \dots, \alpha_m \in K$ implies that $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$. Otherwise, the system will be called linearly dependent.

We call a basis of V any system of linearly independent generator of V . If u_1, \dots, u_n is a basis of V , the expression $v = v_1 u_1 + \dots + v_n u_n$ is called the decomposition of v with respect to the basis and the scalars $v_1, \dots, v_n \in K$ are the components of v with respect to the given basis. Moreover, the following pro holds.

Property 1.1. Let V be a vector space which admits a basis of n vectors. Then every system of linearly independent vector of V has at most n elements and any other basis of V has n elements. The number n is called the dimension of V and we write $\dim(V) = n$. If, instead, for any n there always exist n linearly independent vectors of V , the vector space is called infinite dimensional.

2. Matrices

Let m and n be two positive integers. We call a matrix having m rows and n columns, or a matrix $m \times n$, or a matrix (m, n) , with elements in K , a set of mn scalars $a_{ij} \in K$, with $i = 1, \dots, m$ and $j = 1, \dots, n$, represented in the following rectangular array

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (1.1)$$

When $K = \mathbb{R}$ or $K = \mathbb{C}$ we shall respectively write $A \in \mathbb{R}^{m \times n}$ or $A \in \mathbb{C}^{m \times n}$, to explicitly outline the numerical fields which the elements of A belong to. Capital letters will be used to denote the matrices, while the lower case letters corresponding to those upper case letters will denote the matrix entries.

We shall abbreviate (1.1) as $A = (a_{ij})$ with $i = 1, \dots, m$ and $j = 1, \dots, n$. The index i is called row index, while j is the column index. The set $(a_{i1}, a_{i2}, \dots, a_{in})$ is called the i -th row of A ; likewise, $(a_{1j}, a_{2j}, \dots, a_{mj})$ is the j -th column of A .

If $n = m$ the matrix is called squared or having order n and the set of the entries $(a_{11}, a_{22}, \dots, a_{nn})$ is called its main diagonal.

A matrix having one row or one column is called a row vector or column vector respectively. Unless otherwise specified, we shall always assume that a vector is a column vector. In the case $n = m = 1$, the matrix will simply denote a scalar of K .

Definition 2.1. Let A be a matrix $m \times n$. Let $1 \leq i_1 < i_2 < \dots < i_k \leq m$ and $1 \leq j_1 < j_2 < \dots < j_l \leq n$ two sets of contiguous indexes. The matrix $S(k \times l)$ of entries $s_{pq} = a_{i_p j_q}$ with $p = 1, \dots, k$, $q = 1, \dots, l$ is called a submatrix of A . If $k = l$ and $i_r = j_r$ for $r = 1, \dots, k$, S is called a principal submatrix of A .

Definition 2.2. A matrix $A(m \times n)$ is called block partitioned or said to be partitioned into submatrices if

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1l} \\ A_{21} & A_{22} & \cdots & A_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kl} \end{bmatrix}$$

where A_{ij} are submatrices of A .

3. Operations with Matrices

3.1. Matrices and Linear Mappings

Definition 3.1. A linear map for \mathbb{C}^n into \mathbb{C}^m is a function $f : \mathbb{C}^n \rightarrow \mathbb{C}^m$ such that $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$, $\forall \alpha, \beta \in K$ and $\forall x, y \in \mathbb{C}^n$.

The following result links matrices and linear maps.

Property 3.1. Let $f : \mathbb{C}^n \rightarrow \mathbb{C}^m$ be a linear map. Then, there exists a unique matrix $A_f \in \mathbb{C}^{m \times n}$ such that

$$f(x) = A_f x \quad \forall x \in \mathbb{C}^n \quad (1.2)$$

Conversely, if $A_f \in \mathbb{C}^{m \times n}$ then the function defined in (1.2) is a linear map from \mathbb{C}^n into \mathbb{C}^m .

4. Well-posedness and Condition Number of a Problem

Consider the following problem: find x such that

$$F(x, d) = 0 \quad (1.3)$$

where d is the set of data which the solution depends on and F is the functional relation between x and d . According to the kind of problem that is represented in (1.3), the variables x and d may be real numbers, vectors or functions. Typically, (1.3) is called a direct problem if F and d are given and x is unknown, inverse problem if F and x are known and d is the unknown, identification problem when x and d are given while the functional relation F is the unknown.

Problem (1.3) is well posed if it admits a unique solution x which depends with continuity on the data. We shall use the terms well posed and stable in an interchanging manner and we shall deal henceforth only with well-posed problems.

A problem which does not enjoy the property above is called ill posed or unstable and before undertaking its numerical solution it has to be regularized, that is, it must be suitably transformed into a well-posed problem. Indeed, it is not appropriate to pretend the numerical method can cure the pathologies of an intrinsically ill-posed problem.

Let D be the set of admissible data, i.e. the set of the values of d in correspondence of which problem (1.3) admits a unique solution. Continuous dependence on the data means that small perturbations on the data d of D yield "small" changes in the solution x . Precisely, let $d \in D$ and denote by δd a perturbation admissible in the sense that $d + \delta d \in D$ and by δx the corresponding change in the solution, in such a way that

$$F(x + \delta x, d + \delta d) = 0 \quad (1.4)$$

Then, we require that

$$\exists \eta_0 = \eta_0(d) > 0, \quad \exists K_0 = K_0(d) \text{ such that if } \|\delta d\| \leq \eta_0 \text{ then } \|\delta x\| \leq K_0 \|\delta d\| \quad (1.5)$$

The norms used for the data and for the solution may not coincide, whenever d and x represent variables of different kinds.

Remark 4.1. The property of continuous dependence on the data could have been stated in the following alternative way, which is more akin to the classical form of Analysis $\forall \epsilon > 0 \exists \delta = \delta(\epsilon)$ such that if $\|\delta d\| \leq \delta$ then $\|\delta x\| \leq \epsilon$.

The form (1.5) is however more suitable to express in the following the concept of numerical stability, that is, the property that small perturbations on the data yield perturbations of the same order on the solution.

With the aim of making the stability analysis more quantitative, we introduce the following definition.

Definition 4.1. For problem (1.3) we define the relative condition number to be

$$K(d) = \sup\left\{\frac{\|\delta x\|/\|x\|}{\|\delta d\|/\|d\|}, \delta d \neq 0, d + \delta d \in D\right\} \quad (1.6)$$

Whenever $d = 0$ or $x = 0$, it is necessary to introduce the absolute condition number, given by

$$K_{abs}(d) = \sup\left\{\frac{\|\delta x\|}{\|\delta d\|}, \delta d \neq 0, d + \delta d \in D\right\} \quad (1.7)$$

Problem (1.3) is called ill-conditioned if $K(d)$ is “big” for any admissible datum d .

The property of a problem of being well-conditioned is independent of the numerical method that is being used to solve it. In fact, it is possible to generate stable as well as unstable numerical schemes for solving well-conditioned problems. The concept of stability for an algorithm or for a numerical method is analogous to that used for problem (1.3) and will be made precise in the next section.

Remark 4.2. (Ill-posed problems) Even in the case in which the condition number does not exist (formally, it is infinite), it is not necessarily true that the problem is ill-posed. In fact there exist well posed problems for which the condition number is infinite, but such that they can be reformulated in equivalent problems with a finite condition number.

If problem (1.3) admits a unique solution, then there necessarily exists a mapping G , that we call resolvent, between the sets of the data and of the solutions, such that

$$x = G(d), \text{ that is } F(G(d), d) = 0 \quad (1.8)$$

According to this definition, (1.4) yields $x + \delta x = G(d + \delta d)$. Assuming that G is differentiable in d and denoting formally by $G'(d)$ its derivative with respect to d (if $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $G'(d)$ will be the Jacobian matrix of G evaluated at the vector d), a Taylor’s expansion of G truncated at first order ensures that

$$G(d + \delta d) - G(d) = G'(d)\delta d + o(\|\delta d\|) \quad \text{for } \delta d \rightarrow 0$$

where $\|\cdot\|$ is a suitable vector norm and $o(\cdot)$ is the classical infinitesimal symbol denoting an infinitesimal term of higher order with respect to its argument. Neglecting the infinitesimal of higher order with respect to $\|\delta d\|$, from (1.6) and (1.7) we respectively deduce that

$$K(d) \approx \|G'(d)\| \frac{\|d\|}{\|G(d)\|}, \quad K_{abs}(d) \approx \|G'(d)\| \quad (1.9)$$

where the symbol $\|\cdot\|$, when applied to a matrix, denotes the induced matrix norm (1.10) associated with the vector norm introduced above. The estimates in (1.9) are of great practical usefulness in the analysis of problems in the form (1.8).

Theorem 4.1. Let $\|\cdot\|$ be a vector norm. The function

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad (1.10)$$

is a matrix norm called induced matrix norm or natural matrix norm.

In view of (1.9), the quantity $\|G'(d)\|$ is an approximation of $K_{abs}(d)$ and is sometimes called first order absolute condition number. This latter represents the limit of the Lipschitz constant of G as the perturbation on the data tends to zero.

Such a number does not always provide a sound estimate of the condition number $K_{abs}(d)$. This happens, for instance, when G' vanishes at a point whilst G is nonnull in a neighborhood of the same point.