Operator Theory:
Advances and Applications
Vol. 36

OT 36

G. R. Belitskii
Yu. I. Lyubich

# Matrix Norms and their Applications

Birkhäuser

G. R. Belitskii
Yu. I. Lyubich

# Matrix Norms and their Applications

Translated from the Russian
by A. Iacob

1988

Birkhäuser Verlag
Basel · Boston · Berlin

Author's address:

Prof. G. R. Belitskii
Institute of Low Temperatures
Kharkov, Ukraine
USSR

Prof. Yurii I. Lyubich
University of Kharkov
pl. Dzerzinskogo 4
Kharkov 310077, Ukraine
USSR

Translation of:
Normy matrits i ikn prilozheniya
»Naukova Dumka«
Kiev, 1984

**CIP-Titelaufnahme der Deutschen Bibliothek**

**Belickij, Genrich R.:**
Matrix norms and their applications / G. R. Belitskii; Yu. I. Lyubich. Transl. from the Russian by A. Iacob. – Basel ; Boston ; Berlin : Birkhäuser, 1988
(Operator theory ; Vol. 36)
Einheitssacht.: Normy matric i ich priloženija <engl.>
ISBN 3-7643-2220-9 (Basel . . .) Pb.
ISBN 0-8176-2220-9 (Boston . . .) Pb.
NE: Ljubič, Jurij I.:; GT


Printed in Germany
ISBN 3-7643-2220-9
ISBN 0-8176-2220-9

**OT 36**
**Operator Theory: Advances and Applications**
**Vol. 36**



Birkhäuser Verlag
Basel · Boston · Berlin

## PREFACE

A finite-dimensional linear topology admits infinitely many distinct geometric realizations, each obtained by choosing a particular norm. In the algebra of matrices it is natural to restrict oneselves to norms that possess the *ring property* $\|AB\| \leq \|A\| \|B\|$. If matrices are treated as linear operators in a linear space E, then upon endowing E with a norm one automatically endows the algebra of matrices with a norm. The matrix norms arising in this manner are called *operator* (alternatively, *induced* or *subordinate*) norms. For a certain period of time they constituted the only known class of matrix norms. Other examples were found only after in 1963 Yu. I. Lyubich (and independently, in 1964, J. Stoer) characterized the operator norms as minimal elements of the pointwise order structure on the set of all matrix norms. The indicated order has been subsequently subject to a detailed study by G. R. Belitskii. The most important result in this direction is the theorem asserting that all automorphisms of the order structure in question are in a certain sense inner. As a whole, a rather rich theory has been developed, an exposition of which is given in Chapters 3 and 4 of the monograph.

Chapter 1 has mainly a preparatory role. Its first two sections are purely introductory. However, beginning with §3, a number of relevant situations in which matrix norms are used are exhibited.

Chapter 2 makes a sufficiently thorough study of the boundary spectrum of contractions. It relies to a considerable extent to a combinatorial analysis that goes back to Frobenius, but has been elaborated in detail only after the publication of a note of Wielandt (1950) dedicated to Frobenius' centennary. A new direction emerged in works of Ptak and his collaborators, who introduced, and also computed in a number of instances the so-called critical exponents. This area is even today far from being studied exhaustively. In this monograph we indicate a number of other unsolved problems; among the solved ones there undoubtedly are some that can constitute a source of new problems.

We describe a variety of applications of matrix norms, not only because of their importance, but also to illustrate the principle of "fitting a norm to a given situation". This principle, which guides many applications of functional analysis, is particularly convincing in the finite-dimensional setting, where the choice of a norm is subject to no restrictions.

It is assumed that the reader is familiar with courses on linear algebra and calculus. Nevertheless, a number of facts from linear algebra are presented in order to make the exposition more accessible. With the more special aspects one can make acquaintance in the books recommended in the list of references. A number of brief comments on the literature are made at the end of the text. Therein we do not mention however the authors of sufficiently elementary or known theorems (except for those that usually bear the names of their authors).

CONTENTS

# CHAPTER 1

# OPERATORS IN FINITE-DIMENSIONAL NORMED SPACES

## §1. NORMS OF VECTORS, LINEAR FUNCTIONALS, AND LINEAR OPERATORS

We consider an n-dimensional (with $0 < n < \infty$) linear space $E$ (referred to as the *ground* or *basic space*) over the field $\mathbb{R}$ or $\mathbb{C}$ of real or, respectively, complex numbers. The notations $E$ and $n$ for the ground space and its dimensions will be kept fixed throughout the book. In studying questions that can be treated without making distinction between the real and complex case we shall denote the ground field by $K$. As a rule, the elements of $K$ (*scalars*) will be denoted by lowercase Greek letters, and the elements of the ground space $E$ (*vectors*) by lowercase Roman letters. The maps $E \to K$ and $E \to E$ are called *functionals on* $E$ and, respectively, *operators in* $E$. *Linear functionals* and *operators* are defined in the usual manner. The adjective "linear" is omitted whenever the linearity is plain from the context. From now on the standard language of linear algebra will be used without superfluous explanations.

For the reader's convenience, we devote this section to the classical definitions and facts connected with the notion of a norm, in their finite-dimensional version.

A functional $\nu$ on $E$ is called a *norm* if it possesses the following properties:

1) $\nu(x) > 0 \quad (\forall x \neq 0)$ (*positivity*) ;
2) $\nu(\alpha x) = |\alpha|\nu(x)$ (*absolute homogeneity*) ;
3) $\nu(x + y) \leqslant \nu(x) + \nu(y)$ (*triangle inequality*).

Immediate consequences of 1)-3) are that $\nu(0) = 0$, $\nu(\alpha x) = \nu(x)$ whenever $|\alpha| = 1$ (in particular, $\nu(-x) = \nu(x)$), and $\nu(x - y) \geqslant |\nu(x) - \nu(y)|$.

Example. Let $1 \leqslant p < \infty$. Pick a basis $e_1, \ldots, e_n$ in $E$ and put, for each vector $x = \sum_{k=1}^{n} \xi_k e_k$,

$$\nu_p(x) = \{\sum_{k=1}^{n} |\xi_k|^p\}^{1/p} . \tag{1.1.1}$$

That the functional (1.1.1) is a norm follows from the Minkowski inequality, known from analysis. It is called the $\ell_p$-*norm* relative to the given basis. Letting $p \to \infty$ we obtain the $\ell_\infty$-*norm*

$$\nu_\infty(x) = \max_{1 \leqslant k \leqslant n} |\xi_k| . \tag{1.1.2}$$

In this limiting case (and also for $p = 1$) the triangle inequality is obvious.

In the case $p = 2$ the norm comes from the *standard inner product*

$$(x,y) = \sum_{k=1}^{n} \xi_k \bar{\eta}_k , \tag{1.1.3}$$

namely, $\nu_2(x) = \{(x,x)\}^{1/2}$. Generally, if $E$ is equipped with an *inner product* (i.e., with a bilinear map $E \times E \to K$, usually denoted $(\cdot,\cdot)$, which is symmetric : $(y,x) = \overline{(x,y)}$, and positive : $(x,x) > 0 \quad \forall\, x \neq 0$), then the functional $\|x\| = \sqrt{(x,x)}$ is a norm. In this case the triangle inequality follows from the Schwarz inequality

$$|(x,y)| \leqslant \|x\|\cdot\|y\| ; \tag{1.1.4}$$

the meaning of the latter is that the discriminant of the Hermitian quadratic form

$$(x,x)\xi\bar{\xi} + (x,y)\xi\bar{\eta} + (y,x)\eta\bar{\xi} + (y,y)\eta\bar{\eta} = (\xi x + \eta y, \xi x + \eta y) \geq 0$$

is nonnegative.

A norm defined in the indicated manner by an inner product is termed here a *Euclidean norm.*

Every norm $\nu$ on $E$ defines a *metric*

$$d(x,y) = \nu(x - y) , \tag{1.1.5}$$

which in turn defines a *topology* on $E$. *This topology on* $E$ *does not depend on the choice of the norm* $\nu$. This is a consequence of the fact that *any two norms*, $\nu$ *and* $\tilde{\nu}$, *on* $E$ *are equivalent*, i.e., there exist constants $\alpha,\beta > 0$ such that

$$\alpha\nu(x) \leq \tilde{\nu}(x) \leq \beta\nu(x)$$

for all $x \in E$. To prove this is suffices to verify that any norm $\tilde{\nu}$ is equivalent with the $\ell_\infty$-norm $\nu_\infty$.

One of the inequalities needed to this end is obvious :

$$\tilde{\nu}(x) = \tilde{\nu}(\sum_{k=1}^{n} \xi_k e_k) \leq \sum_{k=1}^{n} |\xi_k|\tilde{\nu}(e_k) \leq \beta\nu_\infty(x) ,$$

where $\beta = \sum_{k=1}^{n}\tilde{\nu}(e_k)$. From this estimate it follows that

$$|\tilde{\nu}(x) - \tilde{\nu}(y)| \leq \beta d_\infty(x,y) ,$$

where $d_\infty$ is the metric associated with the norm $\nu_\infty$ [$d_\infty$ is sometimes referred to as the *uniform* metric (relative to the given basis $e_1,\ldots,e_n$)]. We see that the functional $\tilde{\nu}$ is continuous in the topology defined by $\nu_\infty$. Consider the restriction of $\tilde{\nu}$ to the "unit sphere" $S = \{x \mid \nu_\infty(x) = 1\}$. Since $S$ is compact and $\tilde{\nu}|S > 0$, there is an $\alpha > 0$ such that $\tilde{\nu}(x) \geq \alpha$ for all $x \in S$. Then $\tilde{\nu}(x) \geq \alpha\nu_\infty(x)$ for all $x \in E$, because for $x \neq 0$ the vector $\{\nu_\infty(x)\}^{-1}x$ belongs to $S$ and consequently satisfies the above inequality.

The topology introduced above on $E$ is called *standard*. It is a *linear topology*, i.e., relative to it the operations of addition and multiplication by a scalar are jointly continuous (the ground field K is endowed with the standard topology). Thus, E is a topological linear space. Notice that the convergence of a sequence $\{x_n\}_1^\infty$ to a vector x is equivalent to the coordinate-wise convergence in some basis, since the latter is equivalent to the convergence in the metric $d_\infty$. From this remark it is plain that the space E is complete.

If E is endowed with some fixed norm, then it is called a *normed space* or a *Minkowski space*, and the singled-out norm is denoted by $\|\cdot\|$.

A space E endowed with a fixed inner product is called a *Euclidean space*. It is automatically normed. A normed space is termed *Euclidean* if its norm is Euclidean.

The main geometric figures in a normed space E are the *open unit ball* $\mathcal{D} = \{x \mid \|x\| < 1\}$, the *closed unit ball* $\overline{\mathcal{D}} = \{x \mid \|x\| \leqslant 1\}$, and the *unit sphere* $S = \{x \mid \|x\| = 1\}$. The balls $\mathcal{D}$ and $\overline{\mathcal{D}}$ are absolutely convex, i.e., convex and invariant under multiplication by any scalar $\lambda$ with $|\lambda| = 1$. Also, $\overline{\mathcal{D}}$ and S are compact. These assertions are easily derived from the main properties of the norm and its continuity.

The vectors $x \in S$ are called *unit* or *normed* vectors. Every vector $x \neq 0$ can be *normed* by setting

$$\hat{x} = \frac{x}{\|x\|} . \qquad (1.1.6)$$

Since $\hat{x} \in \overline{\mathcal{D}}$, the set $\overline{\mathcal{D}}$ is *absorbing*, that is, for every $x \in E$ there is an $\alpha > 0$ such that $\alpha^{-1}x \in \overline{\mathcal{D}}$. The value $\|x\|$ is the infimum of all such $\alpha$'s. This remark can be used to prove that *every absolutely convex absorbing compact set* $\Delta \subset E$ *is the closed unit ball relative to some norm.* All we have to check is that the functional

$$\nu(x) = \inf\{\alpha \mid \alpha > 0,\ \alpha^{-1}x \in \Delta\} \qquad (1.1.7)$$

is a norm and that $\{x \mid \nu(x) \leqslant 1\} = \Delta$. We omit the proof.

Thus, *there is a natural one-to-one correspondence between norms on* E *and absolutely convex absorbing compact subsets of* E.

Let E be a normed space and S the unit sphere in E. Let $E^*$ be the dual (or conjugate) of E, i.e., the space of all linear functionals on E. We endow $E^*$ with the *dual* (or *conjugate*) *norm*

$$\|\phi\| = \sup_{x\neq 0} \frac{|\phi(x)|}{\|x\|} = \sup_{x\in S} |\phi(x)| \quad (\text{for } \phi \in E^*) \tag{1.1.8}$$

(actually, one should write $\|\phi\|^*$, but we stay with the simpler notation). Every linear functional on E is continuous, being a linear function of coordinates in an arbitrary basis $e_1,\ldots,e_n$ :

$$\phi(x) = \sum_{k=1}^{n} \phi(e_k)\xi_k , \quad \text{where} \quad x = \sum_{k=1}^{n} \xi_k e_k .$$

Consequently, the supremum in (1.1.8) is finite and is attained, i.e.,

$$\|\phi\| = \max_{x\in S} |\phi(x)| = \max_{x\neq 0} \frac{|\phi(x)|}{\|x\|} . \tag{1.1.8'}$$

[A straightforward consequence of the formula preceding (1.1.8') is that in Euclidean space every linear functional admits the *Riesz representation*: $\phi(x) = (x,y_\phi)$. This establishes a one-to-one correspondence $\phi \mapsto y_\phi$ between the spaces $E^*$ and E, known as the *Riesz correspondence* or *isomorphism*. It is anti-linear, i.e., $y_{\alpha_1\phi_1+\alpha_2\phi_2} = \bar{\alpha}_1 y_{\phi_1} + \bar{\alpha}_2 y_{\phi_2}$. Notice that

$$\|\phi\| = \max_{x\in S} |(x,y_\phi)| .$$ ]

<u>Example 1.</u> Let E be endowed with the $\ell_p$-norm relative to some basis. Then the dual norm in $E^*$ is precisely the $\ell_q$-norm (where $p^{-1} + q^{-1} = 1$) relative to the dual basis. This follows from the well-known Hölder inequality

$$\left|\sum_{k=1}^{n} \alpha_k\beta_k\right| \leq \left(\sum_{k=1}^{n} |\alpha_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |\beta_k|^q\right)^{1/q} .$$

<u>Example 2.</u> If the space E is Euclidean, then $\|\phi\| = \|y_\phi\|$, i.e., the Riesz isomorphism is norm-preserving. This follows from Schwarz's inequality.

Let $L \subset E$ be a subspace and let $\psi$ be a linear functional on $L$. Then the norm of any extension of $\psi$ to the full space $E$ is obviously not smaller than the norm of $\psi$. The classical Hahn-Banach Theorem asserts that *any linear functional* $\psi$ *given on a subspace* $L$ *of a normed space* $E$ *can be extended to the whole space* $E$ *preserving its norm.* As is known, this result is valid even in infinite-dimensional spaces. In the finite-dimensional case the proof, given below, is elementary.

Consider first a real space $E$. We may assume that $\dim L = n-1$ and $\|\psi\| = 1$. Pick an arbitrary vector $e \in E \setminus L$. Each $x \in E$ can be uniquely written as a sum $x = \xi e + y$, where $\xi \in \mathbb{R}$ and $y \in L$. If $\phi$ is a linear extension of the given linear functional $\psi \in L^*$ to $E$, then $\phi(x) = \xi\phi(e) + \psi(y)$. Hence, the single parameter on which the extension depends is the value $\lambda = \phi(e)$. We show that one can choose $\lambda$ so that $\|\phi\| = 1$. This requirement reduces to the double inequality

$$-\|\xi e + y\| \leqslant \lambda\xi + \psi(y) \leqslant \|\xi e + y\| . \qquad (1.1.9)$$

For $\xi = 0$ it is satisfied, since $\|\psi\| = 1$. If it is satisfied for all $\xi > 0$ and all $y \in L$, then it is also satisfied for all $\xi < 0$ and all $y \in L$. Setting $y = \xi z$ (with $\xi > 0$ and $z \in L$), we rewrite (1.1.9) in the form

$$-\|e + z\| - \psi(z) \leqslant \lambda \leqslant \|e + z\| - \psi(z) \quad (\forall z \in L).$$

The existence of such a $\lambda$ follows from the inequality

$$\sup(-\|e + z\| - \psi(z)) \leqslant \lambda \leqslant \inf(\|e + z\| - \psi(z)),$$

which in turn holds because

$$-\|e + z_1\| - \psi(z_1) \leqslant \|e + z_2\| - \psi(z_2)$$

for all $z_1, z_2 \in L$ ; indeed,

$$\psi(z_2) - \psi(z_1) = \psi(z_2 - z_1) \leqslant \|z_2 - z_1\| \leqslant \|e + z_1\| + \|e + z_2\| .$$

Now let $E$ be a complex space, $L$ a subspace of $E$, and $\psi \in L^*$, with $\|\psi\| = 1$. Restricting the ground field to $\mathbb{R}$, we turn $E$ into a real space. On $L$ consider the $\mathbb{R}$-linear functional $\psi_0 = \operatorname{Re}\psi$. Obviously, $\|\psi_0\| = 1$. Extend $\psi_0$ to an $\mathbb{R}$-linear functional $\phi_0$ on $E$ such that $\|\phi_0\| = 1$. Now set $\phi(x) = \phi_0(x) - i\phi_0(ix)$. The functional $\phi$ is $\mathbb{C}$-linear, being additive, $\mathbb{R}$-homogeneous, and such that $\phi(ix) = i\phi(x)$. If $x \in L$, then

$$\phi(x) = \psi_0(x) - i\psi_0(ix) = \operatorname{Re}\psi(x) - i \operatorname{Re}\psi(ix) =$$

$$= \operatorname{Re}\psi(x) + i \operatorname{Im}\psi(x) = \psi(x) ,$$

i.e., $\phi$ is an extension of $\psi$. Finally, for every $x \in E$, $|\phi(x)| = \phi(x)e^{-i\theta}$ for a suitable $\theta$, and then $|\phi(x)| = \operatorname{Re}\phi(xe^{-i\theta}) = \phi_0(xe^{-i\theta}) \leqslant \|xe^{-i\theta}\| = \|x\|$, i.e., $\|\phi\| \leqslant 1$. In the end, $\|\phi\| = 1$.

The consequences of the Hahn-Banach Theorem are rather numerous. One of them is that *for every vector* $x_0$ *there is a linear functional* $\phi$ *such that* $\|\phi\| = 1$ *and* $\phi(x_0) = \|x_0\|$ (this $\phi$ is called a *supporting functional* to the sphere $\{x \mid \|x\| = \|x_0\|\}$ at the point $x_0$). To see this for an $x_0 \neq 0$, it suffices to take the one-dimensional subspace $L$ spanned by $x_0$ and put $\psi(\xi x_0) = \xi\|x_0\|$. This yields a linear functional $\psi$ *on* $L_0$ with the needed properties. It remains to extend $\psi$ to $E$ preserving its norm. For $x_0 = 0$ there is nothing to prove.

The well-known Duality Theorem asserts that *the canonical mapping* $E \to E^{**}$ *which sends each vector* $x$ *into the linear functional* $\hat{x}$ *on* $E^*$ *given by* $\hat{x}(\phi) = \phi(x)$ *is an isomorphism*. For normed spaces one has, in addition, that $\|\hat{x}\| = \|x\|$, i.e., *the canonical mapping is isometric* (in other words, upon passing twice to the dual one recovers the original norm). In fact,

$$\|\hat{x}\| = \sup_{\phi\neq 0} \frac{|\hat{x}(\phi)|}{\|\phi\|} = \sup_{\phi\neq 0} \frac{|\phi(x)|}{\|\phi\|} = \|x\|,$$

because $|\phi(x)| \leqslant \|\phi\| \, \|x\|$, where equality is attained if one takes for $\phi$ the supporting functional to the sphere $\|z\| = \|x\|$ at $x$.

Let $A$ be a linear operator in the normed space $E$. By definition, the *norm* of A is

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x\in S} \|Ax\| \tag{1.1.10}$$

(where S denotes, as above, the unit sphere in E). That the supremum in (1.1.10) is finite and attained follows, as for linear functionals, from the compactness of S and the continuity of the linear operator A. The latter is in turn obvious thanks to the representation of A in an arbitrary basis $e_1,\dots,e_n$ :

$$Ax = \sum_{k=1}^{n} \xi_k Ae_k \qquad (\text{for } x = \sum_{k=1}^{n} \xi_k e_k) .$$

Hence,

$$\|A\| = \max_{x\in S} \|Ax\| = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} . \tag{1.1.10'}$$

In this way the algebra End(E) of all linear operators in E (or *endomorphisms of* E) becomes a *normed algebra*, that is, the functional $A \to \|A\|$ on End(E), in addition to being a norm on the linear space End(E), is a *ring norm*, i.e., it possesses the *ring property*

$$\|AB\| \leqslant \|A\| \, \|B\| \tag{1.1.11}$$

and is *unit-preserving* :

$$\|I\| = 1 , \tag{1.1.12}$$

where $I \in$ End(E) denotes the identity operator.

With each linear operator A in E one associates the *conjugate* or *adjoint* operator $A^*$ in $E^*$, acting as $(A^*\phi)(x) = \phi(Ax)$. It is readily seen that for the conjugate norm on $E^*$ one has that

$$\|A^*\| = \|A\|. \tag{1.1.13}$$

In fact,

$$\|A^*\| = \sup_{\|\phi\|=1} \|A^*\phi\| = \sup_{\|\phi\|=1} \sup_{\|x\|=1} |(A^*\phi)(x)| =$$

$$= \sup_{\|\phi\|=1} \sup_{\|x\|=1} |\phi(Ax)| = \sup_{\|x\|=1} \sup_{\|\phi\|=1} |\phi(Ax)| =$$

$$= \sup_{\|x\|=1} \|\hat{Ax}\| = \sup_{\|x\|=1} \|Ax\| = \|A\| ;$$

here we used the canonical isometry $E \xrightarrow{\wedge} E^{**}$.

In a Euclidean space $E$ one can always use the Riesz isomorphism to think of $A^*$ as acting in $E$ so that $(Ax,y) = (x,A^*y)$ for all $x,y \in E$. The left-hand side of this equality is the Hermitian-bilinear functional associated with the operator $A$. The correspondence defined in this manner is an isomorphism between the space of linear operators and the space of Hermitian-bilinear functionals. Moreover,

$$\|A\| = \sup_{\|x\|,\|y\| \neq 0} \frac{|(Ax,y)|}{\|x\| \, \|y\|} = \sup_{\|x\|=\|y\|=1} |(Ax,y)| .$$

Both supremums are finite and attained ; the common value is called the *norm of the Hermitian-bilinear functional*. One can also consider the Hermitian-quadratic functional $Q(x) = (Ax,x)$ associated with $A$, and define its norm as

$$\|Q\| = \sup_{x\neq 0} \frac{|(Ax,x)|}{\|x\|^2} = \sup_{\|x\|=1} |(Ax,x)| = \sup_{\|x\|=1} |(Ax,x)| .$$

Clearly, $\|Q\| \leqslant \|A\|$. On the other hand, one can show that $\|A\| \leqslant 2\|Q\|$, and that this inequality is exact (i.e., the best).

In some cases the norm of an operator $A$ can be computed explicitly in terms of the entries of its matrix.

<u>Example 1.</u> Suppose $E$ is endowed with the $\ell_\infty$-norm relative to some basis $e_1,\dots,e_n$. If in this basis $A$ is given by the matrix $(\alpha_{jk})_{j,k=1}^{n}$, then

$$\|A\| = \max_{1\leqslant j\leqslant n} \sum_{k=1}^{n} |\alpha_{jk}| . \tag{1.1.14}$$

This particular norm of operators (or matrices) is called the *row norm*.

<u>Example 2.</u> With the same notations, but endowing $E$ with the