UNIVERSITAS COMENIANA
ACTA MATHEMATICA UNIVERSITATIS COMENIANAE
XLVIII—XLIX — 1986

SOME ASPECTS OF THE C.-B.-S. INEQUALITY

LADISLAV KOSMÁK—ĽUDMILA WINTEROVÁ, Bratislava

The inequality

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leqq \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right) \tag{1}$$

is useful e.g. for proving that the function $x \mapsto |x|$ from R^n to R with

$$|x| = (x_1^2 + \ldots + x_n^2)^{\frac{1}{2}}$$

for $x = (x_1, \ldots, x_n)$, fulfils the triangle inequality

$$|x + y| \leqq |x| + |y|.$$

Indeed, from (1) we get

$$\Sigma\,(a_k + b_k)^2 = \Sigma\, a_k^2 + 2\,\Sigma\, a_k b_k + \Sigma\, b_k^2 \leqq$$
$$\leqq \Sigma\, a_k^2 + 2(\Sigma\, a_k^2)^{\frac{1}{2}}(\Sigma\, b_k^2)^{\frac{1}{2}} + \Sigma\, b_k^2 = \left((\Sigma\, a_k^2)^{\frac{1}{2}} + (\Sigma\, b_k^2)^{\frac{1}{2}}\right)^2,$$

where all summations run from 1 to n.

The inequality (1) is given in Cauchy's Cours d'analyse de l'Ecole Royala Polytechnique, 1, Analyse algébrique, Paris, 1821. An integral version of this inequality has been published by V. Bunyakovskij in the paper Sur quelques inégalités concernant les intégrales ordinaires et les intégrales aux différences finies, Mémoires de l'Académie de St.-Pétersbourgh (VII), 1, No 9, 1859. Some 26 years later, it appears in the paper of H. A. Schwarz: Ueber ein Flächen kleinsten Flächeninhalts betreffendes Problem der Variationsrechnung, Acta Soc. Sc. Fenn. 15 (1885), 315—362. The relation (1) and its generalizations are mostly known under the name of Schwarz's inequality; historically more correct, but combersom is the name Cauchy—Bunyakovskij—Schwarz (abbreviated C.-B.-S.) inequality.

1. Geometrical aspect

Let a, b be nonnull vectors in the Euclidean space R^n,

$$a = (a_1, \ldots, a_n)$$
$$b = (b_1, \ldots, b_n).$$

The angle φ of the vectors a, b is defined as the convex angle between the halflines $\{\lambda a;\ \lambda \geqq 0\}$, $\{\lambda b;\ \lambda \geqq 0\}$.

In the triangle with sides of length $|a|$, $|b|$, $|a - b|$ we have the cosine formula

$$|a - b|^2 = |a|^2 + |b|^2 - 2|a|\,|b| \cos \varphi$$

i.e.

$$\Sigma\, (a_k - b_k)^2 = \Sigma\, a_k^2 + \Sigma\, b_k^2 - 2\sqrt{(\Sigma\, a_k^2)\,(\Sigma\, b_k^2)} \cos \varphi.$$

Hence

$$\cos \varphi = \frac{1}{|a|\,|b|}\, \Sigma\, a_k b_k = \frac{(a, b)}{|a|\,|b|}, \tag{2}$$

where (a, b) is the scalar product of a, b. This leads to the definition of the angle between two vectors in an arbitrary vector space with scalar product.

Since $|\cos \varphi| \leqq 1$, it follows from (2) that

$$|(a, b)| \leqq |a|\,|b|,$$

i.e. the inequality (1); equality holds if and only if $|\cos \varphi| = 1$, in other words, if the vectors a, b are linearly dependent.

2. Algebraic aspect

For arbitrary real $a_1, \ldots, a_n, b_1, \ldots, b_n$ we have

$$0 \leqq \sum_j \sum_k (a_j b_k - a_k b_j)^2 =$$
$$= \sum_j \sum_k a_j^2 b_k^2 - 2 \sum_j \sum_k a_j b_j a_k b_k + \sum_j \sum_k a_k^2 b_j^2 =$$
$$= \Big(\sum_j a_j^2\Big)\Big(\sum_k b_k^2\Big) - 2\Big(\sum_j a_j b_j\Big)\Big(\sum_k a_k b_k\Big) + \Big(\sum_j b_j^2\Big)\Big(\sum_k a_k^2\Big) =$$
$$= 2\Big(\sum_k a_k^2\Big)\Big(\sum_k b_k^2\Big) - 2\Big(\sum_k a_k b_k\Big)^2$$

summations running over the set $\{1, \ldots, n\}$. Thus we get both the Lagrange identity

$$\left(\sum_k a_k b_k\right)^2 = \left(\sum_k a_k^2\right)\left(\sum_k b_k^2\right) - \frac{1}{2}\sum_j \sum_k (a_j b_k - a_k b_j)^2 \tag{3}$$

and the inequality (1). Equality holds if and only if the rank of the matrix

$$\begin{pmatrix} a_1, & \dots, & a_n \\ b_1, & \dots, & b_n \end{pmatrix}$$

is less then 2, i.e. if one of the n-tuples $(a_1, \dots, a_n)$, $(b_1, \dots, b_n)$ is a multiple of the other.

In the notation of part 1, the Lagrange identity for nonnull vectors $a, b \in \mathrm{R}^3$ can be written in the form

$$(a, b)^2 = |a|^2 |b|^2 - |a \times b|^2,$$

where $a \times b$ denotes the vector product of a, b. In other words,

$$|a|^2 |b|^2 \cos^2 \varphi = |a|^2 |b|^2 - |a|^2 |b|^2 \sin^2 \varphi. \tag{3'}$$

Hence in this case (3) follows trivially from the equality

$$\sin^2 \varphi + \cos^2 \varphi = 1.$$

Another algebraic proof using the properties of the quadratic form

$$\sum_{j,\, k} (x a_j + y b_k)^2$$

is given in [6], II, 2.4.

An inductive proof of (1) together with the equality condition looks as follows. For $n = 1$ the assertion is obvious. Assume that it holds for some $n \geqq 1$. Since for $k = 1, \dots, n$

$$(a_k b_{n+1} - a_{n+1} b_k)^2 \geqq 0,$$

i.e.

$$2 a_k b_k a_{n+1} b_{n+1} \leqq a_k^2 b_{n+1}^2 + a_{n+1}^2 b_k^2$$

we get

$$2(a_1 b_1 + \dots + a_n b_n) a_{n+1} b_{n+1} \leqq (a_1^2 + \dots + a_n^2) b_{n+1}^2 + (b_1^2 + \dots + b_n^2) a_{n+1}^2,$$

where equality holds if and only if

$$\begin{vmatrix} a_k & a_{n+1} \\ b_k & b_{n+1} \end{vmatrix} = 0$$

for $k = 1, \dots, n$. Hence

$$(a_1 b_1 + \dots + a_{n+1} b_{n+1})^2 = (a_1 b_1 + \dots + a_n b_n)^2 +$$
$$+ 2(a_1 b_1 + \dots + a_n b_n) a_{n+1} b_{n+1} + a_{n+1}^2 b_{n+1}^2 \leqq$$

$$\leqq (a_1^2 + \dots + a_n^2)(b_1^2 + \dots + b_n^2) + (a_1^2 + \dots + a_n^2)b_{n+1}^2 +$$
$$+ (b_1^2 + \dots + b_n^2)a_{n+1}^2 + a_{n+1}^2 b_{n+1}^2 = (a_1^2 + \dots + a_{n+1}^2)(b_1^2 + \dots + b_{n+1}^2)$$

with the same equality condition as in the first algebraic proof.

3. Physical aspect

Let (A, p) with $p \geqq 0$ be a material point in the Euclidean space R^3. The moment of inertia of (A, p) with respect to a point $X \in \mathrm{R}^3$ is defined as the number $p \cdot \overline{XA}^2$. More generally, if $\{(A_1, m_1), \dots, (A_n, m_n)\}$ is a system of material points with positive masses and with the centre of gravity T, denote by $d_1, \dots, d_n$, d the distances of the points $A_1, \dots, A_n$, T from X and put

$$m_1 + \dots + m_n = m.$$

Then the moment of inertia of this system with respect to X is the number

$$I_X = m_1 d_1^2 + \dots + m_n d_n^2.$$

Hence $I_X = 0$ if and only if $A_1 = \dots = A_n = X$. By Lagrange theorem (cf. [2], p. 72, or [3], p. 159) we have

$$I_X = I_T + md^2.$$

Let $a_1, \dots, a_n$ be positive numbers and $b_1, \dots, b_n \geqq 0$. Assume that the points $A_1, \dots, A_n$ lie on a halfline starting from X and that

$$m_k = a_k^2$$

$$d_k = \frac{b_k}{a_k}$$

for $k = 1, \dots, n$. Then

$$m = a_1^2 + \dots + a_n^2$$

$$d = \frac{b_1}{a_1}\frac{a_1^2}{m} + \dots + \frac{b_n}{a_n}\frac{a_n^2}{m} = \frac{1}{m}(a_1 b_1 + \dots + a_n b_n).$$

Putting

$$a_1 b_1 + \dots + a_n b_n = s$$

we have, for $k = 1, \dots, n$

$$|T - A_k| = \frac{b_k}{a_k} - \frac{s}{m}$$

so that

$$I_T = a_1^2 |A_1 - T|^2 + \dots + a_n^2 |A_n - T|^2 =$$

$$= b_1^2 + \dots + b_n^2 - 2\,\frac{s}{m}\,(a_1b_1 + \dots + a_nb_n) + \frac{s^2}{m} = b_1^2 + \dots + b_n^2 - \frac{s^2}{m}.$$

Since

$$I_X = b_1^2 + \dots + b_n^2$$

$$I_X - I_T = md^2 = \frac{s^2}{m},$$

we get

$$0 \leqq I_T = I_X - \frac{s^2}{m} = b_1^2 + \dots + b_n^2 - (a_1b_1 + \dots + a_nb_n)^2(a_1^2 + \dots + a_n^2)^{-1}$$

and this is the inequality (1). Equality holds if and only if $A_1 = \dots = A_n$, i.e. if

$$\frac{b_1}{a_1} = \dots = \frac{b_n}{a_n}$$

in other words, if the n-tuple $(b_1, \dots, b_n)$ is a multiple of $(a_1, \dots, a_n)$.

If some of the numbers $a_1, \dots, a_n$ equal 0 we have to consider only such indices $j \in \{1, \dots, n\}$ for which $a_j \neq 0$.

Remark. The properties of the centre of gravity of a finite system of material points can be used to give physically motivated proofs of many mathematical formulas and theorems; for an account of many applications of this method the reader is referred to the book [2].

4. Probability aspect

Let X, Y be random variables on a finite probability space and assume that X, and Y, takes on the values $a_1, \dots, a_n$, and $b_1, \dots, b_n$ with the probabilities $p_1, \dots, p_n$, and $q_1, \dots, q_n$ respectively. By the well-known properties of the correlation coefficient $R(X, Y)$ of X, Y we have

$$|R(X, Y)| = |R(aX + B, cY + d)| \leqq 1 \tag{4}$$

for arbitrary a, b, c, d (see for instance [7], Theorem 6.2.1), so that we can suppose that

$$E(X) = E(Y) = 0.$$

If, moreover,

$$\left.\begin{aligned} p_1 = \dots = p_n = q_1 = \dots = q_n = \frac{1}{n} \\ p_{jk} = P(X = a_j,\ Y = b_k) = 0 \quad \text{for } j \neq k \\ p_{jj} = \frac{1}{n} \end{aligned}\right\} \tag{5}$$

and if we consider the nontrivial case when $E(X^2) \neq 0$, $E(Y^2) \neq 0$ then

$$R(X, Y) = \frac{E(XY)}{\sqrt{E(X^2)E(Y^2)}}, \tag{6}$$

where

$$E(XY) = \frac{1}{n}(a_1b_1 + \dots + a_nb_n)$$

$$E(X^2) = \frac{1}{n}(a_1^2 + \dots + a_n^2)$$

$$E(Y^2) = \frac{1}{n}(b_1^2 + \dots + b_n^2).$$

Hence (4) and (6) yield (1); equality holds if and only if $Y = aX$ for some $a \neq 0$ (cf. [7], Theorem 6.2.2).

Another way to obtain (1) by probabilistic arguments without using correlation coefficient consists in the following. If

$$E(X^2) \neq 0, \quad E(Y^2) \neq 0$$

take

$$\tilde{X} = \frac{X}{\sqrt{E(X^2)}}, \quad \tilde{Y} = \frac{Y}{\sqrt{E(Y^2)}}.$$

The elementary inequality

$$2|xy| \leqq x^2 + y^2$$

yields

$$2|\tilde{X}\tilde{Y}| \leqq X^2 + Y^2,$$

Thus

$$2E(|\tilde{X}\tilde{Y}|) \leqq E(\tilde{X}^2) + E(\tilde{Y}^2) = 2,$$

i.e. (see [5], p. 93)

$$(E(XY))^2 \leqq (E(|XY|))^2 \leqq E(X^2)E(Y^2).$$

In particular, if the conditions (5) are fulfiled, we get (1), without direct information about equality conditions.

5. Generalizations

Passing to the limit for $n \to \infty$ in (1) we get the corresponding inequality in the real space l_2.

An integral form of the C.-B.-S. inequality for continuous functions defined on the interval $[a, b]$ follows from the integral version of the Lagrange identity: we have

$$\int_a^b \int_a^b [f(x)g(y) - f(y)g(x)]^2 dxdy =$$

$$= \int_a^b \int_a^b f^2(x)g^2(y)dxdy - 2\int_a^b \int_a^b f(x)g(x)f(y)g(y)dxdy +$$

$$+ \int_a^b \int_a^b f^2(y)g^2(x)dxdy = 2\int_a^b f^2(x)dx \int_a^b g^2(x)dx - 2\left(\int_a^b f(x)g(x)dx\right)^2.$$

Thus

$$\left(\int_a^b f(x)g(x)dx\right)^2 = \int_a^b f^2(x)dx \int_a^b g^2(x)dx - \frac{1}{2}\int_a^b \int_a^b [f(x)g(y) - f(y)g(x)]^2 dxdy.$$

For further generalizations of this kind the reader is referred e.g. to [4], I, § 18.

The C.-B.-C. inequality in all its forms is a special case of the corresponding Hölder's inequality.

If a, b are elements of a complex vector space L with scalar product, then

$$|(a, b)|^2 \leqq (a, a)(b, b). \tag{7}$$

This inequality can be proved in the following well-known way: for any complex λ we have

$$0 \leqq (a + \lambda b, a + \lambda b) = (a, a) + \lambda^*(a, b) + \lambda(b, a) + \lambda\lambda^*(b, b),$$

where λ^* is the conjugate of λ. For $b \neq 0$ put

$$\lambda = -\frac{(a, b)}{(b, b)}.$$

Then

$$(a, a) - \frac{(a, b)^2}{(b, b)} - \frac{(a, b)^2}{(b, b)} + \frac{(a, b)^2}{(b, b)} \geqq 0,$$

i.e.

$$(a, a)(b, b) - |(a, b)|^2 \geqq 0.$$

For $b = 0$ the inequality is obvious. Equality holds if and only if a is a multiple of b.

For arbitrary $a_1, \ldots, a_n \in L$ the matrix $G(a_1, \ldots, a_n)$ with elements (a_j, a_k), $j, k = 1, \ldots, n$ is called Gram's matrix of the vectors $a_1, \ldots, a_n$ (in this order). It is well-known that for the corresponding Gram's determinant we have

$$\det G(a_1, \ldots, a_n) \geqq 0 \tag{8}$$

with equality if and only if $a_1, \ldots, a_n$ are linearly dependent (see [8], p. 19—23, or [1], p. 23—26). The inequality (7) follows from (8) for $n = 2$.

BIBLIOGRAPHY

[1] Achijezer, N. I.: Teorie aproximací. Praha 1955.
[2] Balk, M. B.: Geometričeskije priloženija o centre tjažesti. Moskva, 1959.
[3] Banach, S.: Mechanics. Warszawa—Wrocław, 1951.
[4] Beckenbach, E. F.—Bellman, R.: Inequalities. Springer, 1969.
[5] Borovkov, A. A.: Teorija verojatnostej. Moskva, 1976.
[6] Hardy, G. H.—Littlewood, J. E.—Pólya, G.: Inequalities. London, 1934.
[7] Kosmák, L.: Kombinatorická teória pravdepodobnosti. Bratislava, 1979.
[8] Pták, V.: Matematická analysa I. Praha, SNP, 1974.

Author's addresses:
Ladislav Kosmák, Ľudmila Winterová
Katedra teoretickej kybernetiky MFF UK
Mlynská dolina
842 15 Bratislava

Received: 7. 11. 1984

SÚHRN

NIEKOĽKO ASPEKTOV C.-B.-S. NEROVNOSTI

Ladislav Kosmák, Ľudmila Winterová

Nerovnosť (1) sa v práci dokazuje geometrickými, algebraickými, fyzikálnymi a pravdepodobnostnými metódami a sú uvedené jej zovšeobecnenia.

РЕЗЮМЕ

НЕСКОЛЬКО АСПЕКТОВ НЕРАВЕНСТВА К.-Б.-Ш.

Ладислав Космак, Людмила Винтерова

В работе приводятся доказательства неравенства (1) средствами геометрии, алгебры, физики и теории вероятностей. В заключение рассматриваются некоторые обобщения этого неравенства.