Mathematical
Inequalities
& Applications
Volume 11, Number 4 (2008), 629–634

A CAUCHY–SCHWARZ INEQUALITY FOR TRIPLES OF VECTORS

MARINA ARAV, FRANK J. HALL AND ZHONGSHAN LI

(*communicated by A Ben-Israel*)

Abstract. An inequality of the Cauchy–Schwarz type is proved for triples of vectors, and interpreted geometrically in terms of the vertex angles of a tetrahedron. The results are illustrated by 3×3 correlation matrices.

1. Introduction

Let $\mathbb{R}^n$ denote the real n–dimensional vector space, with the standard inner product denoted by $\mathbf{x}\cdot\mathbf{y}$. Given vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, their *Gram matrix* is

$$\begin{pmatrix} \mathbf{a}\cdot\mathbf{a} & \mathbf{a}\cdot\mathbf{b} \\ \mathbf{b}\cdot\mathbf{a} & \mathbf{b}\cdot\mathbf{b} \end{pmatrix} , \tag{1}$$

and its determinant, called the *Gramian* of $\mathbf{a}, \mathbf{b}$, is denoted by $G(\mathbf{a}, \mathbf{b})$. The nonnegativity of $G(\mathbf{a}, \mathbf{b})$ is equivalent to the *Cauchy–Schwarz inequality*

$$(\mathbf{a}\cdot\mathbf{b})^2 \leqslant (\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b}) , \tag{2}$$

where equality holds if and only if the vectors $\mathbf{a}, \mathbf{b}$ are linearly dependent, i.e. $G(\mathbf{a}, \mathbf{b}) = 0$, see, e.g., [2, Chapter IX, Section 5].

The inequality (2) can be restated as

$$\cos^2 \gamma \leqslant 1 , \tag{3}$$

where γ is the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$, denoted by $\gamma = \measuredangle\{\mathbf{a}, \mathbf{b}\}$.

Hardy et al, [3, Section 2.4], noted that the nonnegativity of the Gramian allows generalizing the Cauchy–Schwarz inequality to more than two vectors. This idea is used in Section 2 to obtain a simple Cauchy–Schwarz inequality for three vectors, see Theorem 1. A geometric interpretation in terms of vertex angles in a tetrahedron is given in Section 3, Corollaries 1–2. The above results shed light on correlation matrices for three variables, and allow a natural characterization of such matrices, see Section 4.

Mathematics subject classification (2000): 15A45, 52B10, 62H20.

Key words and phrases: Cauchy–Schwarz inequality for three vectors; Gram matrix; Gramian; tetrahedron; angles; correlation matrices.

The first author was partially supported by NSF – Association for Women in Mathematics grant.

Paper MIA-11-52

2. A Cauchy-Schwarz inequality

THEOREM 1. *If* $\mathbf{a}, \mathbf{b}, \mathbf{c}$ *are nonzero vectors in* $\mathbb{R}^n$, $n \geqslant 3$, *then*

$$\frac{(\mathbf{a}\cdot\mathbf{b})^2}{(\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b})} + \frac{(\mathbf{a}\cdot\mathbf{c})^2}{(\mathbf{a}\cdot\mathbf{a})(\mathbf{c}\cdot\mathbf{c})} + \frac{(\mathbf{b}\cdot\mathbf{c})^2}{(\mathbf{b}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{c})} \leqslant 1 + 2\,\frac{(\mathbf{a}\cdot\mathbf{b})(\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{c})}{(\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{c})} \tag{4}$$

with equality if, and only if, the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ *are linearly dependent.*

Proof. Consider the Gram matrix of $\mathbf{a}, \mathbf{b}, \mathbf{c}$,

$$\begin{pmatrix} \mathbf{a}\cdot\mathbf{a} & \mathbf{a}\cdot\mathbf{b} & \mathbf{a}\cdot\mathbf{c} \\ \mathbf{b}\cdot\mathbf{a} & \mathbf{b}\cdot\mathbf{b} & \mathbf{b}\cdot\mathbf{c} \\ \mathbf{c}\cdot\mathbf{a} & \mathbf{c}\cdot\mathbf{b} & \mathbf{c}\cdot\mathbf{c} \end{pmatrix} \tag{5}$$

and its Gramian

$$G(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{c}) + 2\,(\mathbf{a}\cdot\mathbf{b})(\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{c}) - (\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{c})^2 - (\mathbf{b}\cdot\mathbf{b})(\mathbf{a}\cdot\mathbf{c})^2 - (\mathbf{c}\cdot\mathbf{c})(\mathbf{a}\cdot\mathbf{b})^2.$$

The nonnegativity of $G(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is equivalent to

$$(\mathbf{c}\cdot\mathbf{c})(\mathbf{a}\cdot\mathbf{b})^2 + (\mathbf{b}\cdot\mathbf{b})(\mathbf{a}\cdot\mathbf{c})^2 + (\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{c})^2 \leqslant (\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{c}) + 2\,(\mathbf{a}\cdot\mathbf{b})(\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{c}), \tag{6}$$

and (4) follows by dividing both sides of (6) by $(\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{c})$.

Equality holds in (6), and in (4), if and only if $G(\mathbf{a}, \mathbf{b}, \mathbf{c}) = 0$, i.e., the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are linearly dependent. □

If the vector $\mathbf{c}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$, then (4) reduces to the Cauchy–Schwarz inequality (2). However, the inequality (4) is stronger than (2) in the following sense: applying (2) separately to each term on the left side of (4) gives

$$\frac{(\mathbf{b}\cdot\mathbf{c})^2}{(\mathbf{b}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{c})} + \frac{(\mathbf{a}\cdot\mathbf{c})^2}{(\mathbf{a}\cdot\mathbf{a})(\mathbf{c}\cdot\mathbf{c})} + \frac{(\mathbf{a}\cdot\mathbf{b})^2}{(\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b})} \leqslant 3\,,$$

which is weaker than (4) since, by (2) again, the right side of (4) is also $\leqslant 3$.

REMARKS.

1. For other generalizations of the Cauchy–Schwarz inequality to 3 vectors see, e.g., [5, Exercises 1.3, 1.14].

2. An inequality analogous to (4), for more than 3 vectors, can be obtained from the Gramian as in Theorem 1, but is more complicated than it is worth, even for 4 vectors.

3. Angles

The following corollary is a geometric statement of (4), analogous to (3).

COROLLARY 1. *Let* $\mathbf{a},\mathbf{b},\mathbf{c}$ *be nonzero vectors in* $\mathbb{R}^n$, $n \geqslant 3$, *and let*

$$\alpha = \measuredangle\{\mathbf{b},\mathbf{c}\},\ \beta = \measuredangle\{\mathbf{a},\mathbf{c}\},\ \gamma = \measuredangle\{\mathbf{a},\mathbf{b}\} \tag{7}$$

be the corresponding angles. Then

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma \leqslant 1 + 2\cos\alpha\cos\beta\cos\gamma\ , \tag{8}$$

with equality if and only if the vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ *are linearly dependent.*

The angles α,β,γ in (8) are not arbitrary, since they are the angles between pairs of given vectors $\{\mathbf{a},\mathbf{b},\mathbf{c}\}$, and therefore,

$$\alpha+\beta+\gamma \leqslant 2\pi\ . \tag{9}$$

We call such angles *vertex angles* (considering the vectors $\{\mathbf{a},\mathbf{b},\mathbf{c}\}$ as edges of a tetrahedron with vertex at $\mathbf{0}$, see Figure 1 and assume they are ordered to satisfy

$$0 \leqslant \alpha \leqslant \beta \leqslant \gamma \leqslant \pi\ . \tag{10}$$

The right part of Figure 1 shows three vertex angles in the plane. This picture is obtained by cutting the "surface" of the tetrahedron along the vector $\mathbf{a}$, and laying it flat, which is why the vector $\mathbf{a}$ appears twice.

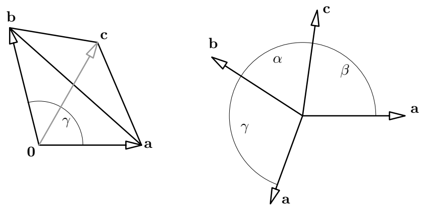

Figure 1. The tetrahedron $T(\mathbf{0},\mathbf{a},\mathbf{b},\mathbf{c})$ *and the vertex angles at* $\mathbf{0}$

A necessary and sufficient condition for three angles satisfying (9)–(10) to be vertex angles is:

Any one of the three angles is no greater than the sum of the other two angles, (11)

or, by (10),

$$\gamma \leqslant \alpha+\beta\ . \tag{12}$$

The converse of Corollary 1 holds.

COROLLARY 2. *Let* α,β,γ *satisfy (9)–(10). Then the inequality (8) is necessary and sufficient for* $\{\alpha,\beta,\gamma\}$ *to be vertex angles.*

Proof. Necessity follows from Corollary 1. To prove sufficiency, write (8) as a quadratic inequality in $\cos\gamma$,

$$\cos^2\gamma - (2\cos\alpha\cos\beta)\cos\gamma + (\cos^2\alpha - \sin^2\beta) \leqslant 0\ . \tag{13}$$

The corresponding quadratic equation has two roots that can be written after simplification as

$$\cos\gamma = \cos\alpha\cos\beta \pm \sqrt{\sin^2\alpha\sin^2\beta} = \cos(\beta\pm\alpha)\ .$$

Inequality (13) therefore implies

$$\beta - \alpha \leqslant \gamma \leqslant \alpha + \beta\ , \tag{14}$$

where the left inequality is redundant, by (10). □

REMARKS.

1. The Cauchy–Schwarz inequality (4) has a different geometric interpretation than the classical inequality (2): For any two vectors $\mathbf{a},\mathbf{b}$, (2) states that the length $|\mathbf{a}\cdot\mathbf{b}|/\sqrt{\mathbf{b}\cdot\mathbf{b}}$ of the projection of $\mathbf{a}$ on $\mathbf{b}$ is not greater than $\sqrt{\mathbf{a}\cdot\mathbf{a}}$, the length of $\mathbf{a}$. Inequality (4) states, for any three vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ and the corresponding angles (7), that no angle is greater than the sum of the other two angles, a sort of *triangle inequality* for vertex angles.

2. The inequality (8) is related to the volume of a tetrahedron. Let $\mathbf{a},\mathbf{b},\mathbf{c}$ be nonzero vectors in $\mathbb{R}^3$. Consider the tetrahedron $T(\mathbf{0},\mathbf{a},\mathbf{b},\mathbf{c})$ with vertices at the origin $\mathbf{0}$, and at the endpoints of the vectors $\mathbf{a},\mathbf{b},\mathbf{c}$. Its volume is given by the formula

$$\begin{aligned}&\text{Volume}\, T(\mathbf{0},\mathbf{a},\mathbf{b},\mathbf{c})\\ &\quad= \frac{\|\mathbf{a}\|\|\mathbf{b}\|\|\mathbf{c}\|}{6}\sqrt{1+2\cos\alpha\cos\beta\cos\gamma - \cos^2\alpha - \cos^2\beta - \cos^2\gamma}\ ,\end{aligned} \tag{15}$$

and is positive, by Corollary 1, if and only if the vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ are not coplanar.

4. Correlations

The above results yield a simple and natural characterization of 3×3 correlation matrices in terms of the underlying angles.

Let $\mathbf{X},\mathbf{Y}$ denote random variables. We denote by
$E\,\mathbf{X}$ the *expected value* of $\mathbf{X}$,
$\text{var}(\mathbf{X})$ its *variance*, $\text{var}(\mathbf{X}) := E((\mathbf{X}-E\,\mathbf{X})^2)$,
$\text{cov}(\mathbf{X},\mathbf{Y})$ the *covariance* of $\mathbf{X},\mathbf{Y}$, $\text{cov}(\mathbf{X},\mathbf{Y}) := E((\mathbf{X}-E\,\mathbf{X})(\mathbf{Y}-E\,\mathbf{Y}))$, and
$\text{cor}(\mathbf{X},\mathbf{Y})$ the *correlation* of $\mathbf{X},\mathbf{Y}$, $\text{cor}(\mathbf{X},\mathbf{Y}) := \text{cov}(\mathbf{X},\mathbf{Y})/\sqrt{\text{var}(\mathbf{X})}\sqrt{\text{var}(\mathbf{Y})}$.

The *correlation matrix* of a set of random variables $\{\mathbf{X}_1,\mathbf{X}_2,\ldots,\mathbf{X}_p\}$ is the $p\times p$ matrix $R = (r_{ij})$, where $r_{ij} = \text{cor}(\mathbf{X}_i,\mathbf{X}_j)$. Since $\text{cov}(\mathbf{X},\mathbf{Y})$ is an inner product, it follows that R is symmetric and, by the Cauchy–Schwarz inequality,

$$r_{ij}^2 \leqslant 1 \text{ and } r_{ii} = 1,\ \text{for all } i,j. \tag{16}$$

Moreover, R is positive semidefinite.

Conversely, given a $p \times p$ real matrix $R = (r_{ij})$ with diagonal elements all equal to 1, and the off diagonal elements r_{ij} such that $r_{ij}^2 \leqslant 1$, when is R a correlation matrix? Such a matrix R is a correlation matrix if and only if it is positive semidefinite.

The case $p = 3$ is of special interest, see [4, Example 10.1, p. 109], where the following question is posed:

"Consider a 3×3 symmetric matrix

$$R = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{pmatrix}, \tag{17}$$

where all $r_{ij}^2 \leqslant 1$. When is this matrix R a proper correlation matrix, that is, such a correlation matrix which could be obtained from some real data?"

This question is equivalent to: "When is R positive semidefinite?"

The determinant of R in (17) is

$$\det(R) = 1 + 2\, r_{12}\, r_{13}\, r_{23} - r_{12}^2 - r_{13}^2 - r_{23}^2.$$

Because of (16), all the principal 2×2 minors are nonnegative. Therefore the answer depends on the sign of the determinant of R, i.e., R is a correlation matrix if and only if,

$$1 + 2\, r_{12}\, r_{13}\, r_{23} - r_{12}^2 - r_{13}^2 - r_{23}^2 \geqslant 0\ , \tag{18}$$

which is inequality (8) with correlations interpreted as cosines. It follows from Corollary 2 that R is a correlation matrix if and only if the angles

$$\alpha = \cos^{-1} r_{23},\ \beta = \cos^{-1} r_{13},\ \gamma = \cos^{-1} r_{12} \tag{19}$$

are vertex angles, i.e., satisfy (11).

REMARK. Several equivalent conditions are given in [4, Example 10.1, p. 109], such as

$$r_{13}\, r_{23} - \sqrt{(1 - r_{13}^2)(1 - r_{23}^2)} \leqslant r_{12} \leqslant r_{13}\, r_{23} + \sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}\ .$$

This can be rewritten, by (19), as

$$\begin{gathered} \cos\beta\, \cos\alpha - \sin\beta\, \sin\alpha \leqslant \cos\gamma \leqslant \cos\beta\, \cos\alpha + \sin\beta\, \sin\alpha \\ \therefore\ \cos(\beta + \alpha) \leqslant \cos\gamma \leqslant \cos(\beta - \alpha) \\ \therefore\ (\beta - \alpha) \leqslant \gamma \leqslant (\beta + \alpha)\ , \end{gathered}$$

which is (14), see also [1].

Acknowledgement. The authors thank Professor George Styan for suggesting the connection with correlation matrices, and the referee for constructive comments.

REFERENCES

[1] J. K. BAKSALARY, *Solution to Problem 89-7 [Let X, Y and Z be random variables. If the correlation $\rho(X,Y)$ and $\rho(Y,Z)$ are known, what are the sharp lower and upper bounds for $\rho(X,Z)$?]*, The IMS Bulletin, 19, 213–214, 1990.
[2] F. R. GANTMACHER, *The Theory of Matrices*, Volume 1, Chelsea Publishing Co., 1959.
[3] G. H. HARDY, J. E. LITTLEWOOD AND G. PÓLYA, *Inequalities*, Cambridge University Press, 1952.

[4] J. ISOTALO, S. PUNTANEN AND G. P. H. STYAN, *Matrix Tricks for Linear Statistical Models* (3rd edition), University of Tampere, 2005, ISBN-951-44-6526-1.
[5] J. M. STEELE, *The Cauchy–Schwarz Master Class: An Introduction to the Art of Inequalities*, Cambridge University Press, 2004.

(Received October 11, 2006)

Marina Arav
Department of Mathematics and Statistics
Georgia State University
Atlanta
GA 30303-3083
USA
e-mail: matmxa@langate.gsu.edu

Frank J. Hall
Department of Mathematics and Statistics
Georgia State University
Atlanta
GA 30303-3083
USA
e-mail: matfjh@langate.gsu.edu

Zhongshan Li
Department of Mathematics and Statistics
Georgia State University
Atlanta
GA 30303-3083
USA
e-mail: matzli@langate.gsu.edu