

A GENERIC REFINEMENT TO THE CAUCHY–SCHWARZ INEQUALITY

Mohammad Masjed-Jamei

ABSTRACT. We present a generic refinement to the Cauchy–Schwarz inequality in both inner product space and probability space and study some of its special cases.

1. Introduction

The inner product of two real-valued continuous functions f, g with respect to the weight function w positive on $[a, b]$, is indicated as

$$(1.1) \quad \langle f, g \rangle_w = \int_a^b w(x) f(x) g(x) dx,$$

while in a discrete space it is in the form

$$(1.2) \quad \langle f, g \rangle_w = \sum_{x \in B} w(x) f(x) g(x),$$

where B is a given counter set.

Based on two representations (1.1), (1.2), the well-known Cauchy–Schwarz inequality

$$(1.3) \quad \langle f, g \rangle_w^2 \leq \langle f, f \rangle_w \langle g, g \rangle_w,$$

can be represented in a continuous space as

$$(1.4) \quad \left(\int_a^b w(x) f(x) g(x) dx \right)^2 \leq \int_a^b w(x) f^2(x) dx \int_a^b w(x) g^2(x) dx,$$

and in a discrete space as

$$(1.5) \quad \left(\sum_{x \in B} w(x) f(x) g(x) \right)^2 \leq \sum_{x \in B} w(x) f^2(x) \sum_{x \in B} w(x) g^2(x).$$

2020 *Mathematics Subject Classification*: 26D15, 39B72, 60E15.

Key words and phrases: generic refinement, Cauchy–Schwarz inequality, inner product spaces, probability spaces.

Communicated by Gradimir Milovanović.

Inequalities (1.4) and (1.5) play an important role in various branches of modern mathematics such as Hilbert spaces theory, classical real and complex analysis, numerical analysis, probability and statistics, qualitative theory of differential equations and their applications. To date, a large number of generalizations and refinements of these inequalities have been investigated in the literature, see, e.g. [1–6, 8, 10].

Here, we establish a generic refinement to the Cauchy–Schwartz inequality in both inner product and probability spaces and study some of its special cases.

2. A generic refinement in the inner product space

For $p, q \in \mathbb{R}$, let us replace $f(x) = u(x) + p$ and $g(x) = v(x) + q$ in (1.3) to get

$$(2.1) \quad \langle u + p, v + q \rangle_w^2 \leq \langle u + p, u + p \rangle_w \langle v + q, v + q \rangle_w.$$

By considering the linear property of an inner product space, inequality (2.1) is finally simplified as

$$(2.2) \quad \langle u, v \rangle_w^2 \leq \langle u, u \rangle_w \langle v, v \rangle_w - A^*(p, q),$$

in which

$$(2.3) \quad A^*(p, q) = \alpha_1 p^2 + \alpha_2 q^2 + 2\alpha_3 pq + 2\alpha_4 p + 2\alpha_5 q,$$

such that

$$(2.4) \quad \begin{aligned} \alpha_1 &= \alpha_1(v) = \langle 1, v \rangle_w^2 - \langle 1, 1 \rangle_w \langle v, v \rangle_w, \\ \alpha_2 &= \alpha_2(u) = \langle u, 1 \rangle_w^2 - \langle u, u \rangle_w \langle 1, 1 \rangle_w, \\ \alpha_3 &= \alpha_3(u, v) = \langle u, v \rangle_w \langle 1, 1 \rangle_w - \langle u, 1 \rangle_w \langle 1, v \rangle_w, \\ \alpha_4 &= \alpha_4(u, v) = \langle u, v \rangle_w \langle 1, v \rangle_w - \langle u, 1 \rangle_w \langle v, v \rangle_w, \\ \alpha_5 &= \alpha_5(u, v) = \langle u, v \rangle_w \langle u, 1 \rangle_w - \langle u, u \rangle_w \langle 1, v \rangle_w. \end{aligned}$$

Note in (2.4) that

$$\alpha_1(u) = \alpha_2(u) \leq 0, \quad \alpha_3(u, v) = \alpha_3(v, u), \quad \alpha_4(v, u) = \alpha_5(u, v).$$

In order to improve the Cauchy–Schwartz inequality, our first goal in (2.2) is to find a suitable domain, say D^* , such that $D^* = \{(p, q) \mid A^*(p, q) \geq 0\}$ and the second goal is to find $\max A^*(p, q)$ for any $(p, q) \in D^*$. To meet the first goal, without loss of generality, we assume in (2.3) that $q = \lambda p$ where $\lambda \in \mathbb{R}$ to get

$$A^*(p, \lambda p) = (\alpha_1 + \lambda^2 \alpha_2 + 2\lambda \alpha_3) p^2 + 2(\alpha_4 + \lambda \alpha_5) p = \beta_1 p^2 + 2\beta_2 p,$$

which is now an incomplete polynomial of degree 2 with respect to the variable p and therefore the aforesaid interval is found as $D^* = [0, -2\frac{\beta_2}{\beta_1}]$ if and only if $\beta_1 < 0$.

On the other hand, since $\frac{d}{dp} A^*(p, \lambda p) = 2\beta_1 p + 2\beta_2$, so $p_{\max} = -\frac{\beta_2}{\beta_1}$ and we have $\max A^*(p, \lambda p) = -\frac{\beta_2^2}{\beta_1}$ provided that $\beta_1 < 0$. Hence, for any $p \in D^* = [0, -2\frac{\beta_2}{\beta_1}]$ we have

$$(2.5) \quad A^*(p, \lambda p) \geq 0 \text{ and } p_{\max} = -\frac{\beta_2}{\beta_1} \in D^* \text{ because in general } \frac{a+b}{2} \in [a, b].$$

In order to show that $\beta_1 = \alpha_2\lambda^2 + 2\alpha_3\lambda + \alpha_1 < 0$, one way is to prove that

$$\alpha_2 \leq 0 \text{ and } \Delta = \alpha_3^2 - \alpha_1\alpha_2 \leq 0 \text{ for any } \lambda \in \mathbb{R}.$$

According to (1.3), it is clear that $\alpha_2 = \langle u, 1 \rangle_w^2 - \langle u, u \rangle_w \langle 1, 1 \rangle_w \leq 0$. Moreover, since

$$(2.6) \quad \left\langle u - \frac{\langle u, 1 \rangle_w}{\langle 1, 1 \rangle_w}, v - \frac{\langle 1, v \rangle_w}{\langle 1, 1 \rangle_w} \right\rangle_w = \langle u, v \rangle_w - \frac{\langle u, 1 \rangle_w \langle 1, v \rangle_w}{\langle 1, 1 \rangle_w},$$

replacing the left-hand side of equality (2.6) in (1.3) eventually yields

$$\begin{aligned} \Delta &= \alpha_3^2 - \alpha_1\alpha_2 = (\langle u, v \rangle_w \langle 1, 1 \rangle_w - \langle u, 1 \rangle_w \langle 1, v \rangle_w)^2 \\ &\quad - (\langle u, 1 \rangle_w^2 - \langle u, u \rangle_w \langle 1, 1 \rangle_w) (\langle 1, v \rangle_w^2 - \langle v, v \rangle_w \langle 1, 1 \rangle_w) \leq 0, \end{aligned}$$

which approves $\beta_1 < 0$. Note in (2.6) that $\langle 1, 1 \rangle_w$ is always positive and

$$\left\langle u - \frac{\langle u, 1 \rangle_w}{\langle 1, 1 \rangle_w}, u - \frac{\langle u, 1 \rangle_w}{\langle 1, 1 \rangle_w} \right\rangle_w = \langle u, u \rangle_w - \frac{\langle u, 1 \rangle_w^2}{\langle 1, 1 \rangle_w}.$$

The following relations now hold true

$$\begin{aligned} \beta_1 &= \alpha_1 + \lambda^2\alpha_2 + 2\lambda\alpha_3 = \langle 1, v \rangle_w^2 - \langle v, v \rangle_w \langle 1, 1 \rangle_w + \lambda^2 \langle u, 1 \rangle_w^2 \\ &\quad - \lambda^2 \langle u, u \rangle_w \langle 1, 1 \rangle_w + 2\lambda \langle u, v \rangle_w \langle 1, 1 \rangle_w - 2\lambda \langle u, 1 \rangle_w \langle v, 1 \rangle_w \\ &= \langle 1, v - \lambda u \rangle_w^2 - \langle v - \lambda u, v - \lambda u \rangle_w \langle 1, 1 \rangle_w \leq 0, \end{aligned}$$

$$\begin{aligned} \beta_2 &= \alpha_4 + \lambda\alpha_5 = \langle u, v \rangle_w \langle 1, v \rangle_w - \langle v, v \rangle_w \langle u, 1 \rangle_w \\ &\quad + \lambda \langle u, v \rangle_w \langle u, 1 \rangle_w - \lambda \langle u, u \rangle_w \langle 1, v \rangle_w \\ &= \langle u, v - \lambda u \rangle_w \langle 1, v \rangle_w - \langle v - \lambda u, v \rangle_w \langle u, 1 \rangle_w. \end{aligned}$$

By replacing the two latter relations in (2.2) and noting (2.5), we have

$$\frac{\beta_2^2}{\beta_1} = - \frac{(\langle u, v - \lambda u \rangle_w \langle 1, v \rangle_w - \langle v - \lambda u, v \rangle_w \langle u, 1 \rangle_w)^2}{\langle v - \lambda u, v - \lambda u \rangle_w \langle 1, 1 \rangle_w - \langle v - \lambda u, 1 \rangle_w^2}.$$

2.1. Corollary 1. For real-valued functions f, g and $w > 0$ and for any $\lambda \in \mathbb{R}$, Cauchy–Schwarz inequality (1.3) can be refined as

$$(2.7) \quad \langle f, g \rangle_w^2 \leq \langle f, f \rangle_w \langle g, g \rangle_w - \frac{(\langle f, 1 \rangle_w \langle g, f - \lambda g \rangle_w - \langle g, 1 \rangle_w \langle f, f - \lambda g \rangle_w)^2}{\langle 1, 1 \rangle_w \langle f - \lambda g, f - \lambda g \rangle_w - \langle 1, f - \lambda g \rangle_w^2}.$$

There are some special cases for the result (2.7). For instance, if $\lambda = 0$, it reads as

$$(2.8) \quad \langle f, g \rangle_w^2 \leq \langle f, f \rangle_w \langle g, g \rangle_w - \frac{(\langle f, 1 \rangle_w \langle f, g \rangle_w - \langle 1, g \rangle_w \langle f, f \rangle_w)^2}{\langle 1, 1 \rangle_w \langle f, f \rangle_w - \langle f, 1 \rangle_w^2},$$

which is equivalent to

$$\begin{aligned} \langle f, g \rangle_w^2 &\leq \left(1 - \frac{\langle f, 1 \rangle_w^2}{\langle 1, 1 \rangle_w \langle f, f \rangle_w} \right) \langle f, f \rangle_w \langle g, g \rangle_w \\ &\quad - \frac{\langle 1, g \rangle_w \langle f, f \rangle_w (\langle 1, g \rangle_w \langle f, f \rangle_w - 2 \langle f, 1 \rangle_w \langle f, g \rangle_w)}{\langle 1, 1 \rangle_w \langle f, f \rangle_w - \langle f, 1 \rangle_w^2}. \end{aligned}$$

Also if $f \rightarrow g$ and $g \rightarrow f$, inequality (2.8) changes to

$$(2.9) \quad \langle f, g \rangle_w^2 \leq \langle f, f \rangle_w \langle g, g \rangle_w - \frac{(\langle 1, g \rangle_w \langle f, g \rangle_w - \langle f, 1 \rangle_w \langle g, g \rangle_w)^2}{\langle 1, 1 \rangle_w \langle g, g \rangle_w - \langle 1, g \rangle_w^2},$$

which is in fact corresponding to the case $A^*(0, q) = \alpha_2 q^2 + 2\alpha_5 q$ in (2.3), because for any $q \in [0, -2\frac{\alpha_5}{\alpha_2}]$ we have $A^*(0, q) \geq 0$ and since $q_{\max} = -\frac{\alpha_5}{\alpha_2} \in [0, -2\frac{\alpha_5}{\alpha_2}]$, replacing $A^*(0, q_{\max}) = -\frac{\alpha_5^2}{\alpha_2}$ in (2.2) eventually leads to (2.9). Another case of (2.7) is when $\lambda = -1$, i.e.,

$$\langle f, g \rangle_w^2 \leq \langle f, f \rangle_w \langle g, g \rangle_w - \frac{(\langle f, 1 \rangle_w \langle g, f+g \rangle_w - \langle g, 1 \rangle_w \langle f, f+g \rangle_w)^2}{\langle 1, 1 \rangle_w \langle f+g, f+g \rangle_w - \langle 1, f+g \rangle_w^2}.$$

We point out that to obtain an optimum value for λ , we can minimize the denominator of the fraction in (2.7), i.e., the value $-\beta_1$. For this purpose, by noting that $-\alpha_2 \geq 0$, we have

$$-\beta_1 = \beta_1(\lambda) = -\alpha_2 \lambda^2 - 2\alpha_3 \lambda - \alpha_1 \Rightarrow \frac{d\beta_1(\lambda)}{d\lambda} = -2\alpha_2 \lambda - 2\alpha_3 = 0.$$

Hence $\lambda_{\min} = -\frac{\alpha_3}{\alpha_2}$ and substituting this value into β_1 and β_2 finally gives $\frac{\beta_2^2}{\beta_1} = \frac{(\alpha_2 \alpha_4 - \alpha_3 \alpha_5)^2}{\alpha_2 (\alpha_1 \alpha_2 - \alpha_3^2)}$ where $\{\alpha_k\}_{k=1}^5$ are determined in (2.4).

3. A generic refinement in the probability space

Let U and V be two arbitrary random variables. The Cauchy–Schwartz inequality in a probability space takes the form

$$(3.1) \quad E^2(UV) \leq E(U^2)E(V^2),$$

where $E(\cdot)$ denotes the expected value of a random variable [7].

In order to establish a generic refinement to the inequality (3.1), similar to the previous section, we first replace $U = X + p$ and $V = Y + q$ to obtain

$$(3.2) \quad E^2(XY) \leq E(X^2)E(Y^2) + B^*(p, q),$$

in which

$$(3.3) \quad B^*(p, q) = (\text{var } Y)p^2 + (\text{var } X)q^2 - 2 \text{cov}(X, Y)pq + 2(E(X)E(Y^2) - E(Y)E(XY))p + 2(E(Y)E(X^2) - E(X)E(XY))q.$$

Note in (3.3) that

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y), \quad \text{var } X = \text{cov}(X, X) = E(X^2) - E^2(X) \geq 0.$$

Also we have

$$\begin{aligned} E(X)E(Y^2) - E(Y)E(XY) &= E(X) \text{var } Y - E(Y) \text{cov}(X, Y), \\ E(Y)E(X^2) - E(X)E(XY) &= E(Y) \text{var } X - E(X) \text{cov}(X, Y), \end{aligned}$$

We should now find a suitable domain in (3.2), say D^{**} , such that

$$D^{**} = \{(p, q) \mid B^*(p, q) \leq 0\}$$

and then finding $\min B^*(p, q)$ for any $(p, q) \in D^{**}$. For this purpose, we once again assume in (3.3) that $q = \lambda p$ to arrive at

$$\begin{aligned} B^*(p, \lambda p) &= (\lambda^2 \operatorname{var} X + \operatorname{var} Y - 2\lambda \operatorname{cov}(X, Y)) p^2 \\ &\quad + 2(E(X)E(Y^2) - E(Y)E(XY) + \lambda E(Y)E(X^2) - \lambda^2 E(X)E(XY)) p \\ &= a(\lambda)p^2 + b(\lambda)p. \end{aligned}$$

The aforesaid interval can be therefore considered as $D^{**} = [0, -2\frac{b(\lambda)}{a(\lambda)}]$ if $a(\lambda) \geq 0$.

Since $\frac{d}{dp} B^*(p, \lambda p) = 2a(\lambda)p + 2b(\lambda)$, so $p_{\min} = -\frac{b(\lambda)}{a(\lambda)}$ and $\min B^*(p, \lambda p) = -\frac{b^2(\lambda)}{a(\lambda)}$ because for any $p \in D^{**}$ we have $B^*(p, \lambda p) \geq 0$ and $p_{\min} \in D^{**}$. In order to show that $a(\lambda) = (\operatorname{var} X)\lambda^2 - 2\operatorname{cov}(X, Y)\lambda + \operatorname{var} Y$ is nonnegative for any $\lambda \in \mathbb{R}$, we must prove that $\operatorname{var} X \geq 0$ and $\Delta = \operatorname{cov}^2(X, Y) - \operatorname{var} X \operatorname{var} Y \leq 0$, which are however clear. This means that $2\lambda \operatorname{cov}(X, Y) \leq \lambda^2 \operatorname{var} X + \operatorname{var} Y$ ($\forall \lambda \in \mathbb{R}$).

3.1. Corollary 2. Let U and V be two arbitrary random variables. For any $\lambda \in \mathbb{R}$, inequality (3.1) can be refined as

$$(3.4) \quad E^2(XY) \leq E(X^2)E(Y^2) - \frac{(E(X)E(Y^2) - E(Y)E(XY) + \lambda E(Y)E(X^2) - \lambda^2 E(X)E(XY))^2}{(\operatorname{var} X)\lambda^2 - 2\operatorname{cov}(X, Y)\lambda + \operatorname{var} Y}.$$

Similar to the previous section, there are some particular cases for inequality (3.4). For instance, if $\lambda = 0$, it reads as

$$E^2(XY) \leq E(X^2)E(Y^2) - \frac{(E(X)E(Y^2) - E(Y)E(XY))^2}{\operatorname{var} Y}.$$

Moreover, the best case for λ is when we minimize the denominator of the fraction (3.4) as follows

$$\begin{aligned} a(\lambda) &= (\operatorname{var} X)\lambda^2 - 2\operatorname{cov}(X, Y)\lambda + \operatorname{var} Y \\ &\Rightarrow \frac{da(\lambda)}{d\lambda} = 2(\operatorname{var} X)\lambda - 2\operatorname{cov}(X, Y) = 0 \Rightarrow \lambda_{\min} = \frac{\operatorname{cov}(X, Y)}{\operatorname{var} X}. \end{aligned}$$

Since

$$(3.5) \quad a(\lambda_{\min}) = a\left(\frac{\operatorname{cov}(X, Y)}{\operatorname{var} X}\right) = \operatorname{var} Y(1 - \rho^2(X, Y)),$$

$$(3.6) \quad b(\lambda_{\min}) = b\left(\frac{\operatorname{cov}(X, Y)}{\operatorname{var} X}\right) = E(X) \operatorname{var} Y(1 - \rho^2(X, Y)),$$

where $\rho(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sqrt{\operatorname{var} X}\sqrt{\operatorname{var} Y}} \in [-1, 1]$, denotes the correlation coefficient [7], replacing (3.5) and (3.6) in (3.4) eventually yields

$$(3.7) \quad E^2(XY) \leq E(X^2)E(Y^2) - E^2(X) \operatorname{var} Y(1 - \rho^2(X, Y)).$$

There is also an analogue for the interesting inequality (3.7). If we directly substitute $p = -\frac{E(Y)}{\text{var } Y} \text{cov}(X, Y)$ and $q = -E(Y)$ into (3.3), we get

$$B^* \left(-\frac{E(Y)}{\text{var } Y} \text{cov}(X, Y), -E(Y) \right) = -E^2(Y) \text{var } X (1 - \rho^2(X, Y)).$$

Hence, according to (3.2), the analogous inequality is revealed as

$$E^2(XY) \leq E(X^2)E(Y^2) - E^2(Y) \text{var } X (1 - \rho^2(X, Y)).$$

3.2. Remark 1. If $p = -E(X)$ and $q = -E(Y)$ are directly substituted into (3.3), then we have

$$B^*(-E(X), -E(Y)) = -(E^2(Y) \text{var } X + E^2(X) \text{var } Y - 2E(X)E(Y) \text{cov}(X, Y)).$$

But since

$$\begin{aligned} & E^2(Y) \text{var } X + E^2(X) \text{var } Y - 2E(X)E(Y) \text{cov}(X, Y) \\ &= (\sqrt{\text{var } X}E(Y) - \sqrt{\text{var } Y}E(X))^2 + 2E(X)E(Y)(\sqrt{\text{var } X}\sqrt{\text{var } Y} - \text{cov}(X, Y)), \end{aligned}$$

inequality (3.2) takes the form

$$(3.8) \quad \begin{aligned} E^2(XY) &\leq E(X^2)E(Y^2) - (\sqrt{\text{var } X}E(Y) - \sqrt{\text{var } Y}E(X))^2 \\ &\quad - 2E(X)E(Y)(\sqrt{\text{var } X}\sqrt{\text{var } Y} - \text{cov}(X, Y)). \end{aligned}$$

By noting that $\sqrt{\text{var } X}\sqrt{\text{var } Y} - \text{cov}(X, Y) \geq 0$, if in (3.8) we take $E(X) \rightarrow |E(X)|$ and $E(Y) \rightarrow |E(Y)|$, then the inequality (3.8) will be refined as

$$E^2(XY) \leq E(X^2)E(Y^2) - (\sqrt{\text{var } X}|E(Y)| - \sqrt{\text{var } Y}|E(X)|)^2.$$

This result was derived by Walker in 2017 [9].

3.3. An open problem. By noting the remark 1 and as we observed, there are many options for choosing the parameters p and q , which should be separately considered and studied.

For instance, by noting that $1 - \frac{\langle f, 1 \rangle_w^2}{\langle 1, 1 \rangle_w \langle f, f \rangle_w} \geq 0$ is valid in an inner product space and $1 - \frac{E^2(X)}{E(X^2)} \geq 0$ is valid in a probability space Are there two specific parameters p and q such that

$$A^*(p, q) = \frac{\langle f, 1 \rangle_w^2 \langle g, g \rangle_w}{\langle 1, 1 \rangle_w}, \quad B^*(p, q) = -E^2(X)E(Y^2),$$

which are respectively equivalent to the inequality

$$\langle f, g \rangle_w^2 \leq \left(1 - \frac{\langle f, 1 \rangle_w^2}{\langle 1, 1 \rangle_w \langle f, f \rangle_w} \right) \langle f, f \rangle_w \langle g, g \rangle_w,$$

in an inner product space and the inequality

$$E^2(XY) \leq \left(1 - \frac{E^2(X)}{E(X^2)} \right) E(X^2)E(Y^2),$$

in a probability space?

Acknowledgments. This work has been supported by the Alexander von Humboldt Foundation under the grant number: Ref 3.4 - IRN - 1128637 - GF-E.

References

1. H. Alzer, *A refinement of the Cauchy-Schwartz inequality*, J. Math. Anal. Appl. **168** (1992), 596-604.
2. ———, *On the Cauchy-Schwartz inequality*, J. Math. Anal. Appl. **234** (1999) 6-14.
3. D. K. Callebaut, *Generalization of the Cauchy-Schwartz inequality*, J. Math. Anal. Appl. **12** (1965) 491-494.
4. S. S. Dragomir, *A survey on Cauchy-Bunyakovsky-Schwarz type discrete inequalities*, J. Inequal. Pure Appl. Math. **4** (3) (2003) Article 63.
5. M. Masjed-Jamei, *A functional generalization of the Cauchy-Schwarz inequality and some subclasses*, Appl. Math. Lett. **22** (2009) 1335-1339.
6. M. Masjed-Jamei, S. S. Dragomir, H. M. Srivastava, *Some generalizations of the Cauchy-Schwarz and the Cauchy-Bunyakovsky inequalities involving four free parameters and their applications*, Math. Comput. Modelling **49** (2009) 1960-1968.
7. D. S. Mitrinovic, J. E. Pecaric, A. M. Fink, *Classical and New Inequalities in Analysis*, Kluwer, Dordrecht/Boston/London, 1993.
8. W. L. Steiger, *On a generalization of the Cauchy-Schwarz inequality*, Amer. Math. Monthly **76** (1969) 815-816.
9. S. G. Walker, *A self-improvement to the Cauchy-Schwartz inequality*, Stat. Probab. Lett. **122** (2017) 86-89.
10. L. Zheng, *Remark on a refinement of the Cauchy-Schwartz inequality*, J. Math. Anal. Appl. **218** (1998) 13-21.

Faculty of Mathematics
K. N. Toosi University of Technology
Tehran
Iran
mmjamei@kntu.ac.ir

(Received 03 06 2022)
(Revised 13 11 2022)