Chapter 20
The Cauchy–Schwarz Inequality

Yes, I’ve learnt from my mistakes
and I think I’m now able to repeat them almost exactly.

Peter Cook

As Steele (2004, p. 1) says, there is no doubt that the Cauchy–Schwarz inequality is one of the most widely and most important inequalities in all of mathematics. This chapter gives some examples of its use in statistics; further examples appear in several places in this book. The Cauchy–Schwarz inequality is also known as the Cauchy–Bouniakowsky–Schwarz inequality and is named after Augustin-Louis Cauchy (1789–1857) (see also Philatelic Item 12.1, p. 290), Viktor Yakovlevich Bouniakowsky [Buniakovskii, Buniakovsky] (1804–1899), and [Karl] Hermann Amandus Schwarz (1843–1921); see Cauchy (1821)\(^1\) Bouniakowsky (1859, pp. 3–4), and Schwarz (1888, pp. 343–345), and the book by Steele (2004, Ch. 1).

**Theorem 20 (The Cauchy–Schwarz inequality).** Let \(x\) and \(y\) be \(n \times 1\) nonnull real vectors. Then

\[
(x'y)^2 \leq x'x \cdot y'y, \quad \text{for all } x, y
\]

is the vector version of the Cauchy–Schwarz inequality. Equality holds in (20.1) if and only if \(x\) and \(y\) are linearly dependent, i.e.,

\[
(x'y)^2 = x'x \cdot y'y \iff \text{there exists } \lambda \in \mathbb{R} : x = \lambda y.
\]

**Proof.** There are many ways to prove (20.2), see, e.g., Marcus & Minc (1992, p. 61). We prove it here using the presentation of the orthogonal projector onto the column space \(C(x)\). Namely, it is obvious that

\[
y'(I_n - P_x)y \geq 0,
\]

i.e.,

\[
y'y - y'x(x'x)^{-1}x'y \geq 0,
\]

---

\(^1\) As Steele (2004, p. 10) says: “Oddly enough, Cauchy did not use his inequality in his text, except in some illustrative exercises.”
\[
\frac{(x'y)^2}{x'x} \leq y'y, \quad (20.5)
\]
from which (20.1) follows. Clearly the equality in (20.1) holds if and only if
\[
(I_n - P_x)y = 0, \quad (20.6)
\]
which is equivalent to the right-hand side of (20.2). Note that clearly the equality in (20.1) holds if \(x = 0\) or \(y = 0\).

\[\square\]

## 20.1 Specific Versions of the Cauchy–Schwarz Inequality

Let \(A\) be an \(n \times n\) nonnegative definite symmetric matrix with the eigenvalue decomposition
\[
A = T \Sigma T', \quad \Sigma = \text{diag}(\lambda_1, \ldots, \lambda_n), \quad T = (t_1 : \ldots : t_n), \quad T'T = I_n. \quad (20.7)
\]
Here \(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0\), and \(t_i\) is the eigenvector corresponding to \(\lambda_i\).

In particular, denoting \(T_1 = (t_1 : \ldots : t_r)\) and \(\Sigma_1 = \text{diag}(\lambda_1, \ldots, \lambda_r)\), where \(r = \text{rank}(A)\), we get the symmetric nonnegative definite square root of \(A\):
\[
A^{1/2} = T_1 \Sigma_1^{1/2} T_1'. \quad (20.8)
\]

Substituting now \(x = A^{1/2}u\) and \(y = (A^+)^{1/2}v\) into (20.1) we obtain
\[
[u' A^{1/2}(A^+)^{1/2} v]^2 \leq u' A^{1/2} A^{1/2} u \cdot v'(A^+)^{1/2}(A^+)^{1/2} v, \quad (20.9a)
\]
\[(u' P_A v)^2 \leq u' A u \cdot v' A^+ v. \quad (20.9b)
\]

Equality holds in (20.9) if and only if
\[
A^{1/2}u\text{ and } (A^+)^{1/2}v\text{ are linearly dependent,} \quad (20.10)
\]
for which one sufficient condition is
\[
u \in \mathcal{N}(A) \quad \text{or} \quad v \in \mathcal{N}(A). \quad (20.11)
\]

Requesting that \(u \notin \mathcal{N}(A)\) and \(v \notin \mathcal{N}(A)\) the equality in (20.9) can be characterized by the condition
\[
A^{1/2}u = \lambda A^{+1/2}v \quad \text{for some } \lambda \in \mathbb{R}, \quad (20.12)
\]
which is equivalent (please confirm) to
\[
Au = \lambda P_A v \quad \text{for some } \lambda \in \mathbb{R}. \quad (20.13)
\]