Summer

Cauchy-Schwarz Masterclass

Notes

Authors: Shaleen Baral

Contents

1. Starting with Cauchy

1.1. Proofs

Proposition 1.1.1 (Cauchy-Schwarz): For $a_i, b_i \in \mathbb{R}$ we have

$$a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \leq \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2}.$$

Proof of Cauchy-Schwarz (via Induction):

Let $H(n)$ stand for the hypothesis that Cauchy's inequality is valid for n.

The base case $H(1)$ is trivially true as the $a_1 b_1 = \sqrt{a_1^2}\sqrt{b_1^2}$. The second base case $H(2)$ follows from observing that,

$$\begin{aligned} 0 &\leq (a_1 b_2 - b_1 a_2)^2 \\ \implies 2a_1 a_2 b_1 b_2 &\leq a_1^2 b_2^2 + b_1^2 a_2^2 \\ \implies a_1^2 b_1^2 + 2a_1 a_2 b_1 b_2 + a_2^2 b_2^2 &\leq a_1^2 b_2^2 + a_1^2 b_1^2 + a_2^2 b_2^2 + b_1^2 a_2^2 \\ \implies a_1 b_1 + a_2 b_2 &\leq \sqrt{a_1^2 + a_2^2}\sqrt{b_1^2 + b_2^2}. \end{aligned}$$

Now, assuming $H(n)$ holds for some $n \geq 1$ we wish to show that $H(n+1)$ also holds. Consider, using the inductive hypothesis,

$$a_1 b_1 + a_2 b_2 + \cdots + a_n b_n + a_{n+1} b_{n+1} \leq \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2} + a_{n+1} b_{n+1}.$$

Then, using $H(2)$,

$$\leq \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2 + a_{n+1}^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2 + b_{n+1}^2}.$$

Thus, $H(n+1)$ also holds. By induction, Cauchy's inequality holds for all $n \in \mathbb{N}$.

$\square$

Lemma 1.1.1:

$$\sum_{k=1}^{\infty} a_k^2 < \infty \text{ and } \sum_{k=1}^{\infty} b_k^2 < \infty \text{ implies that } \sum_{k=1}^{\infty} |a_k b_k| < \infty.$$

Proof (without Cauchy-Schwarz): We want to show that $a_k b_k$ is small whenever a_k^2 and b_k^2 are small. The following observation gives us what we want,

$$0 \leq (x-y)^2 \implies xy \leq \frac{1}{2}(x^2 + y^2).$$

We apply this to $x = |a_k|$ and $y = |b_k|$ and add the inequalities up,

$$\sum_{k=1}^{\infty} |a_k b_k| \leq \frac{1}{2}\sum_{k=1}^{\infty} a_k^2 + \frac{1}{2}\sum_{k=1}^{\infty} b_k^2 < \infty. \tag{1}$$

$\square$

Proof of Cauchy-Schwarz (via eqn. (1)): Assume that neither $\{a_k\}$ not $\{b_k\}$ are made of zeroes. Then define the sequences,

$$\hat{a}_k = \frac{a_k}{\left(\sum_{k=1}^{\infty} a_k^2\right)^{\frac{1}{2}}} \qquad \hat{b}_k = \frac{b_k}{\left(\sum_{k=1}^{\infty} b_k^2\right)^{\frac{1}{2}}}.$$

Now, apply eqn. (1) to $\{\hat{a}_k\}$ and $\left\{\hat{b}_k\right\}$.

$$\begin{aligned} \sum_{k=1}^{\infty} \hat{a}_k \hat{b}_k &\leq \frac{1}{2}\sum_{k=1}^{\infty} \hat{a}_k^2 + \frac{1}{2}\sum_{k=1}^{\infty} \hat{b}_k^2 = 1 \\ \Longrightarrow \sum_{k=1}^{\infty} \frac{a_k b_k}{\left(\sum_{k=1}^{\infty} a_k^2\right)^{\frac{1}{2}}\left(\sum_{k=1}^{\infty} b_k^2\right)^{\frac{1}{2}}} &\leq 1 \\ \Longrightarrow \sum_{k=1}^{\infty} a_k b_k &\leq \left(\sum_{k=1}^{\infty} a_k^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{\infty} b_k^2\right)^{\frac{1}{2}}. \end{aligned}$$

Remark: Normalization is a systematic way of getting from an *additive inequality* to a *multiplicative inequality*.

$\square$

Proposition 1.1.2: Equality holds in the Cauchy-Schwarz inequality iff the sequences $\{a_k\}$ and $\{b_k\}$ are scalar multiple of one another.

Proof: We focus on the nontrivial case where neither of the sequences is identically zero and where both $\sum_{k=1}^{\infty} a_k^2, \sum_{k=1}^{\infty} b_k^2$ are finite.

The backward direction is easy to prove by a routine computation. We focus on the forward direction. The equality

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

implies the equality (with $\hat{a}_k, \hat{b}_k$ defined as above)

$$\sum_{k=1}^{\infty} \hat{a}_k \hat{b}_k = \frac{1}{2}\sum_{k=1}^{\infty} \hat{a}_k^2 + \frac{1}{2}\sum_{k=1}^{\infty} \hat{b}_k^2 = 1. \tag{3}$$

By the two-term bound, $xy \leq \frac{1}{2}(x^2 + y^2)$ we also know that

$$\hat{a}_k \hat{b}_k \leq \frac{1}{2}\hat{a}_k^2 + \frac{1}{2}\hat{b}_k^2 \quad \text{for all } k = 1, 2, ...,$$

If any of these inequalities were strict, then we wouldn't get the equality in eqn. (3). Thus, the equality in eqn. (2) holds for a nonzero series only when we have $\hat{a}_k = \hat{b}_k$ for all $k = 1, 2, ...$ By the definitiion of these normalized values, we have that

$$a_k = \lambda b_k \quad \text{for all } k = 1, 2, ...,$$

with λ given by the raio

$$\lambda = \left(\sum_{j=1}^{\infty} a_j^2\right)^{\frac{1}{2}} \Bigg/ \left(\sum_{j=1}^{\infty} b_j^2\right)^{\frac{1}{2}}.$$

□

1.2. Notation and Generalizations

The Cauchy-Schwarz inequality can be quite compactly represented in the context of an *inner product space*. We introduce the requisite material here.

Definition 1.2.1 . Suppose V is a real vector space. Then a function on $V \times V$ defined by the mapping $(\boldsymbol{a}, \boldsymbol{b}) \mapsto \langle \boldsymbol{a}, \boldsymbol{b} \rangle$ is an *inner product* and we say that $(V, \langle \cdot, \cdot \rangle)$ is a *real inner product space* provided that the pair $(V, \langle \cdot, \cdot \rangle)$ has the following properties

a. $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$ for all $\boldsymbol{v}, \boldsymbol{w} \in V$
b. $\langle \boldsymbol{v}, \boldsymbol{v} \rangle > 0$ for all nonzero $\boldsymbol{v} \in V$
c. $\langle \alpha \boldsymbol{v} + \boldsymbol{u}, \boldsymbol{w} \rangle = \alpha \langle \boldsymbol{v}, \boldsymbol{w} \rangle + \langle \boldsymbol{u}, \boldsymbol{w} \rangle$ for all $\alpha \in \mathbb{R}$ and $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$

On $\mathbb{R}^n$, the following inner product is popular.

Definition 1.2.2 Euclidean Inner Product. For $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$,

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \sum_{j=1}^{n} a_j b_j$$

This lets us rewrite the Cauchy-Schwarz inequality succinctly.

Proposition 1.2.1: For $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ we have

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle \leq \langle \boldsymbol{a}, \boldsymbol{a} \rangle^{\frac{1}{2}} \langle \boldsymbol{b}, \boldsymbol{b} \rangle^{\frac{1}{2}}.$$

Of course, there are other inner product spaces too!

Example: On $\mathbb{R}^n$ the following weighted sum defines an inner product,

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \sum_{j=1}^{n} a_j b_j w_j.$$

Example: Consider the vector space $C[a, b]$ of real-valued continuous functions on the bounded interval $[a, b]$. Then, for any continuous $w : [a, b] \longrightarrow \mathbb{R}$ such that $w(x) > 0$ for all $x \in [a, b]$, we can define an inner product on $C[a, b]$ by setting

$$\langle f, g \rangle = \int_a^b f(x) g(x) w(x) \, dx.$$

This naturally leads us to ask whether Cauchy-Schwarz is true for all inner product spaces.

Proposition 1.2.2: Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Then, for any $\boldsymbol{v}, \boldsymbol{w} \in V$ we have

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle \leq \langle \boldsymbol{v}, \boldsymbol{v} \rangle^{\frac{1}{2}} \langle \boldsymbol{w}, \boldsymbol{w} \rangle^{\frac{1}{2}}.$$

For nonzero $\boldsymbol{v}, \boldsymbol{w} \in V$ we have

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{v} \rangle^{\frac{1}{2}} \langle \boldsymbol{w}, \boldsymbol{w} \rangle^{\frac{1}{2}} \text{ if and only if } \boldsymbol{v} = \lambda \boldsymbol{w}$$

for a nonzero constant λ.

Proof: We try to use a variant of the additive method developed above. Consider,

$$\begin{gathered} 0 \leq \langle \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{v} - \boldsymbol{w} \rangle \\ \Longrightarrow \langle \boldsymbol{v}, \boldsymbol{w} \rangle \leq \frac{1}{2}(\langle \boldsymbol{v}, \boldsymbol{v} \rangle + \langle \boldsymbol{w}, \boldsymbol{w} \rangle). \end{gathered}$$

We normalize to convert this to a multiplicative bound. Since the inequality holds trivially for $\boldsymbol{v} = \boldsymbol{0}$ or $\boldsymbol{w} = \boldsymbol{0}$, we assume that $\boldsymbol{v}, \boldsymbol{w}$ are nonzero. Define,

$$\hat{\boldsymbol{v}} = \frac{\boldsymbol{v}}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle^{\frac{1}{2}}} \qquad \hat{\boldsymbol{w}} = \frac{\boldsymbol{w}}{\langle \boldsymbol{w}, \boldsymbol{w} \rangle^{\frac{1}{2}}} .$$

We then have,

$$\begin{gathered} \langle \hat{\boldsymbol{v}}, \hat{\boldsymbol{w}} \rangle \leq \frac{1}{2}(\langle \hat{\boldsymbol{v}}, \hat{\boldsymbol{v}} \rangle + \langle \hat{\boldsymbol{w}}, \hat{\boldsymbol{w}} \rangle) = 1 \\ \Longrightarrow \left\langle \frac{\boldsymbol{v}}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle^{\frac{1}{2}}}, \frac{\boldsymbol{w}}{\langle \boldsymbol{w}, \boldsymbol{w} \rangle^{\frac{1}{2}}} \right\rangle \leq 1 \\ \Longrightarrow \langle \boldsymbol{v}, \boldsymbol{w} \rangle \leq \langle \boldsymbol{v}, \boldsymbol{v} \rangle^{\frac{1}{2}} \langle \boldsymbol{w}, \boldsymbol{w} \rangle^{\frac{1}{2}}. \end{gathered}$$

Now, we deal with the necessary condition for equality. If $\boldsymbol{v}, \boldsymbol{w}$ are nonzero then the normalized vectors $\hat{\boldsymbol{v}}, \hat{\boldsymbol{w}}$ are well defined. Furthermore, equality in the Cauchy-Schwarz inequality then gives us $\langle \hat{\boldsymbol{v}}, \hat{\boldsymbol{w}} \rangle = 1$. So,

$$\begin{gathered} \langle \hat{\boldsymbol{v}}, \hat{\boldsymbol{w}} \rangle = \frac{1}{2}(\langle \hat{\boldsymbol{v}}, \hat{\boldsymbol{v}} \rangle + \langle \hat{\boldsymbol{w}}, \hat{\boldsymbol{w}} \rangle) \\ \Longrightarrow \langle \hat{\boldsymbol{v}} - \hat{\boldsymbol{w}}, \hat{\boldsymbol{v}} - \hat{\boldsymbol{w}} \rangle = 0 \\ \Longrightarrow \hat{\boldsymbol{v}} = \hat{\boldsymbol{w}} \end{gathered}$$

Thus,

$$\boldsymbol{v} = \lambda \boldsymbol{w} \text{ for } \lambda = \frac{\langle \boldsymbol{v}, \boldsymbol{v} \rangle^{\frac{1}{2}}}{\langle \boldsymbol{w}, \boldsymbol{w} \rangle^{\frac{1}{2}}}.$$

$\square$

1.3. Symmetry and Amplification

This material is from Terence Tao's blog.

We consider the general setting of a complex inner product space, V. In this context, the Cauchy-Schwarz inequality is given by

$$|\langle \boldsymbol{v}, \boldsymbol{w} \rangle| \leq \langle \boldsymbol{v}, \boldsymbol{v} \rangle^{\frac{1}{2}} \langle \boldsymbol{w}, \boldsymbol{w} \rangle^{\frac{1}{2}}.$$

To prove this, we start off with the additive bound,

$$0 \leq \langle \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{v} - \boldsymbol{w} \rangle$$
$$\implies \Re \langle \boldsymbol{v}, \boldsymbol{w} \rangle \leq \frac{1}{2}(\langle \boldsymbol{v}, \boldsymbol{v} \rangle + \langle \boldsymbol{w}, \boldsymbol{w} \rangle).$$

This is a weaker bound than Cauchy-Schwarz as $\Re \langle \boldsymbol{v}, \boldsymbol{w} \rangle \leq \langle \boldsymbol{v}, \boldsymbol{w} \rangle \leq \langle \boldsymbol{v}, \boldsymbol{v} \rangle^{\frac{1}{2}} \langle \boldsymbol{w}, \boldsymbol{w} \rangle^{\frac{1}{2}} \leq \frac{1}{2}(\langle \boldsymbol{v}, \boldsymbol{v} \rangle + \langle \boldsymbol{w}, \boldsymbol{w} \rangle)$. The last inequality follows from the AM-GM inequality (see next chapter for more details).

We can amplify this additive bound by observing some symmetry imbalances. Particularly, the phase rotation $\boldsymbol{v} \mapsto e^{i\theta} \boldsymbol{v}$ preserves the right-hand side but not the left-hand side,

$$\Re\, e^{i\theta} \langle \boldsymbol{v}, \boldsymbol{w} \rangle \leq \frac{1}{2}(\langle \boldsymbol{v}, \boldsymbol{v} \rangle + \langle \boldsymbol{w}, \boldsymbol{w} \rangle).$$

We can choose any real θ we want. To make the left-hand side as large as possible, we choose θ to cancel the phase of $\langle \boldsymbol{v}, \boldsymbol{w} \rangle$. This gets us,

$$|\langle \boldsymbol{v}, \boldsymbol{w} \rangle| \leq \frac{1}{2}(\langle \boldsymbol{v}, \boldsymbol{v} \rangle + \langle \boldsymbol{w}, \boldsymbol{w} \rangle).$$

Now, to strengthen the right-hand side we exploit a different symmetry, *homogenisation symmetry*. Particularly, consider the map $(\boldsymbol{v}, \boldsymbol{w}) \mapsto \left(\lambda \boldsymbol{v}, \frac{1}{\lambda} \boldsymbol{w}\right)$ for a scalar $\lambda > 0$. This gives us,

$$|\langle \boldsymbol{v}, \boldsymbol{w} \rangle| \leq \frac{\lambda^2}{2} \langle \boldsymbol{v}, \boldsymbol{v} \rangle + \frac{1}{2\lambda^2} \langle \boldsymbol{w}, \boldsymbol{w} \rangle.$$

The choice of $\lambda = \sqrt{\|\boldsymbol{w}\| / \|\boldsymbol{v}\|}$ minimizes the right-hand side. This gives us,

$$|\langle \boldsymbol{v}, \boldsymbol{w} \rangle| \leq \langle \boldsymbol{v}, \boldsymbol{v} \rangle^{\frac{1}{2}} \langle \boldsymbol{w}, \boldsymbol{w} \rangle^{\frac{1}{2}}.$$

1.4. Yet Another Proof ™

This material is from Timothy Gower's blog. In this section, we will see a more motivated development of a common proof for the Cauchy-Schwarz inequality (it is pretty much the same proof as the one above).

Recall what we mean by the Cauchy-Schwarz result: For $a_k, b_k \in \mathbb{R}$, we have

$$\sum_{k=1}^{\infty} a_k b_k \leq \left(\sum_{k=1}^{\infty} a_k^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^{\infty} b_k^2 \right)^{\frac{1}{2}}$$

with equality iff the sequences $\{a_i\}$ and $\{b_i\}$ are proportional.

The central idea for our proof will be trying to find a natural way to express the fact that two sequences are proportional. One approach would be to say that there exists a $\lambda \in \mathbb{R}$ such that $a_k = \lambda b_k$ for every k. However, why bother introducing an unknown variable λ unless we absolutely have to? We could simply require all a_k / b_k to be equal. Though, we may be worried about some b_k being zero. We can resolve this by simply saying that two sequences are proportional if $a_k b_j = a_j b_k$ for all j, k.

We want lots of (in fact, for all j, k we want $a_k b_j - a_j b_k = 0$) terms to be zero. This can be expressed by requiring the sum of all their squares to be zero. So, sequences $\{a_k\}$ and $\{b_k\}$ are proportional iff

$$\sum_{k,j} \left(a_k b_j - a_j b_k\right)^2 = 0.$$

Also note that the expression on the left is trivially at least zero. By expanding out the left-hand side, we readily obtain both the Cauchy-Schwarz inequality and the necessary condition for equality,

$$\begin{aligned}\sum_{k,j}\left(a_kb_j - a_jb_k\right)^2 &= \sum_{k,j}\left(a_k^2b_j^2 - 2a_ka_jb_kb_j + a_j^2b_k^2\right) \\ &= 2\sum_{k,j}a_k^2b_j^2 - 2\sum_{k,j}a_kb_ka_jb_j \\ &= 2\sum_k a_k^2\sum_j b_j^2 - 2\left(\sum_k a_kb_k\right)^2.\end{aligned}$$

Now, we try to extend this idea to real inner product spaces. We want to show that

$$\langle \boldsymbol{v}, \boldsymbol{w}\rangle \leq \|\boldsymbol{v}\|\|\boldsymbol{w}\|$$

with equality iff $\boldsymbol{v}$ and $\boldsymbol{w}$ are proportional with a positive constant. Again, we motivate our proof by thinking in terms of expressing proportionality. A first attempt is to say that $\boldsymbol{v}, \boldsymbol{w}$ are proportional with a positive constant iff $\boldsymbol{v}/\|\boldsymbol{v}\| = \boldsymbol{w}/\|\boldsymbol{w}\|$ (note how this doesn't work for proportionality in general, for example we could have $\boldsymbol{v} = -\boldsymbol{w}$). As we did before, we can equivalently express this condition as requiring $\|\boldsymbol{w}\|\boldsymbol{v} - \|\boldsymbol{v}\|\boldsymbol{w} = 0$. So that we may express this using inner products, we consider the squared version of this: $(\|\boldsymbol{w}\|\boldsymbol{v} - \|\boldsymbol{v}\|\boldsymbol{w})^2 = 0$. Note that the left-hand side is in fact always greater than or equal to zero. Then, expanding the left hand side immediately gives us the Cauchy-Schwarz result,

$$(\|\boldsymbol{w}\|\boldsymbol{v} - \|\boldsymbol{v}\|\boldsymbol{w})^2 = 2\|\boldsymbol{v}\|^2\|\boldsymbol{w}\|^2 - 2\|\boldsymbol{w}\|\|\boldsymbol{v}\|\langle \boldsymbol{v}, \boldsymbol{w}\rangle.$$

For a complex inner product space, $(\|\boldsymbol{w}\|\boldsymbol{v} - \|\boldsymbol{v}\|\boldsymbol{w})^2$ expands as

$$(\|\boldsymbol{w}\|\boldsymbol{v} - \|\boldsymbol{v}\|\boldsymbol{w})^2 = 2\|\boldsymbol{v}\|^2\|\boldsymbol{w}\|^2 - \|\boldsymbol{w}\|\|\boldsymbol{v}\|(\langle \boldsymbol{v}, \boldsymbol{w}\rangle + \langle \boldsymbol{w}, \boldsymbol{v}\rangle)$$

Let x be a complex number with modulus $|x| = 1$ and the property that $\langle \boldsymbol{w}, x\boldsymbol{v}\rangle$ is real and non-negative. Consequently, $\langle \boldsymbol{w}, x\boldsymbol{v}\rangle = |\langle \boldsymbol{w}, \boldsymbol{v}\rangle|$. We readily get that $|\langle \boldsymbol{w}, \boldsymbol{v}\rangle| = \langle \boldsymbol{w}, x\boldsymbol{v}\rangle \leq \|\boldsymbol{v}\|\|\boldsymbol{w}\|$ with equality iff $\|\boldsymbol{w}\|\boldsymbol{v} - \|\boldsymbol{v}\|\boldsymbol{w} = 0$.