

Distance between flats

Jonathan L.F. King

University of Florida, Gainesville FL 32611-2082, USA

15 April, 2022 (at 09:18)

Our IP on a \mathbb{C} -VS is conjugate-linear in the 1st-argument. The FRIEDBERG, INSEL, SPENCE text is conjugate-linear in 2nd-argument.

For a general field \mathcal{F} and vectors \mathbf{q}, \mathbf{D} in an \mathcal{F} -vector-space, define the “line through \mathbf{q} in direction \mathbf{D} ”:

$$\text{LinDir}(\mathbf{q}, \mathbf{D}) := \{\mathbf{q} + t\mathbf{D} \mid t \in \mathcal{F}\}.$$

[When $\mathbf{D} = \vec{\mathbf{0}}$, the line degenerates to a point.]

Henceforth, \mathbf{V} is an IP (inner-product) space over \mathbb{C} .

Complex-IP axioms. A \mathbb{C} -IP $\langle \cdot, \cdot \rangle$ is a map $\mathbf{V} \times \mathbf{V} \rightarrow \mathbb{C}$ satisfying for all vectors $\mathbf{u}, \mathbf{a}_1, \mathbf{a}_2, \mathbf{v}, \mathbf{b}_1, \mathbf{b}_2$:

a: Map $\langle \cdot, \cdot \rangle$ is **additively bilinear**:

$$\begin{aligned} \langle \mathbf{a}_1 + \mathbf{a}_2, \mathbf{v} \rangle &= \langle \mathbf{a}_1, \mathbf{v} \rangle + \langle \mathbf{a}_2, \mathbf{v} \rangle \quad \text{and} \\ \langle \mathbf{u}, \mathbf{b}_1 + \mathbf{b}_2 \rangle &= \langle \mathbf{u}, \mathbf{b}_1 \rangle + \langle \mathbf{u}, \mathbf{b}_2 \rangle. \end{aligned}$$

b: Conjugate-symmetry: $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$.

c: Multiplicatively-linear in (for us) its second arg: For each scalar α : $\langle \mathbf{u}, \alpha \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$. [Thus $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \overline{\alpha} \langle \mathbf{u}, \mathbf{v} \rangle$; our IP is **conjugate-linear** in its first argument. Some textbook say IP is **sequilinear**, meaning linear in one argument, and conjugate-linear in the other. “Sequi” means “one and a half”.]

d: Value $\langle \mathbf{u}, \mathbf{u} \rangle$ is non-negative, and is zero IFF $\mathbf{u} = \vec{\mathbf{0}}$. [The IP is **positive-definite**.] \square

Complex Euclidean space. An IP induces a complex Euclidean geometry on \mathbf{V} . The Euclidean-norm $\|\mathbf{u}\|$ is the non-negative real $\|\mathbf{u}\| := \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$. It satisfies the Triangle Inequality: For all vectors \mathbf{u}, \mathbf{v} :

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|, \quad \text{with equality IFF one of } \mathbf{u}, \mathbf{v} \text{ is a non-negative multiple of the other.}$$

We will only use the notion of *angle* in real IP spaces. Wikipedia has the following more general definition: Define the **angle**, θ , between non-zero vectors \mathbf{u}, \mathbf{v} , by

$$\cos(\theta) := \frac{\text{Re}(\langle \mathbf{u}, \mathbf{v} \rangle)}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}.$$

Consequently: Vectors \mathbf{u}, \mathbf{v} are orthogonal, written $\mathbf{u} \perp \mathbf{v}$, if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. [We allow either vector or both to be $\vec{\mathbf{0}}$; hence the zero-vector is orthogonal to everybody.]

Note: The notion of *orthogonal* is used in \mathbb{C} -VSes. Although our defn of *angle* is well-formed in a \mathbb{C} -VS, it is used mostly in \mathbb{R} -VSes. \square

Orthogonal subspaces. Subspaces $\mathbf{A}, \mathbf{B} \subset \mathbf{V}$ are “perpendicular to each other”, written $\mathbf{A} \perp \mathbf{B}$, if $\mathbf{a} \perp \mathbf{b}$ for every $\mathbf{a} \in \mathbf{A}$ and $\mathbf{b} \in \mathbf{B}$.

The **ortho-complement** of subspace \mathbf{U} is

$$\mathbf{U}^\perp := \{\mathbf{v} \in \mathbf{V} \mid \forall \mathbf{u} \in \mathbf{U} : \mathbf{u} \perp \mathbf{v}\}.$$

Symbol \mathbf{U}^\perp is pronounced “**U perp**”.

Easily, \mathbf{U}^\perp is itself a subspace, and it is transverse to \mathbf{U} , i.e. $\mathbf{U}^\perp \cap \mathbf{U} = \{\vec{\mathbf{0}}\}$.

Always, $[\mathbf{U}^\perp]^\perp \supset \mathbf{U}$. If \mathbf{V} is finite dimensional, then $[\mathbf{U}^\perp]^\perp = \mathbf{U}$. \square

1: Fact.

Proj and Orth. For a direction-vector $\mathbf{D} \neq \vec{\mathbf{0}}$ and arbitrary $\mathbf{u} \in \mathbf{V}$, we define the orthogonal-projection operator:

$$2: \quad \text{Proj}_{\mathbf{D}}(\mathbf{u}) := \frac{\langle \mathbf{D}, \mathbf{u} \rangle}{\langle \mathbf{D}, \mathbf{D} \rangle} \cdot \mathbf{D}.$$

Our IP is linear in its 2nd argument, so formula (2) indeed satisfies that $\text{Proj}_{\mathbf{D}}(\beta \mathbf{u}) = \beta \cdot \text{Proj}_{\mathbf{D}}(\mathbf{u})$.

This immediately gives that $\text{Proj}_{\mathbf{D}}$ is idempotent: Write $\mathbf{w} := \text{Proj}_{\mathbf{D}}(\mathbf{u}) = \alpha \mathbf{D}$, where $\alpha := \frac{\langle \mathbf{D}, \mathbf{u} \rangle}{\langle \mathbf{D}, \mathbf{D} \rangle}$. Then $\text{Proj}_{\mathbf{D}}(\mathbf{w})$ equals $\alpha \cdot \text{Proj}_{\mathbf{D}}(\mathbf{D}) = \alpha \mathbf{D} = \mathbf{w}$.

Let’s also *check* that the difference $\mathbf{u} - \text{Proj}_{\mathbf{D}}(\mathbf{u})$ is orthogonal to \mathbf{D} : Well, $\langle \mathbf{D}, \mathbf{u} - \text{Proj}_{\mathbf{D}}(\mathbf{u}) \rangle$ equals

$$\langle \mathbf{D}, \mathbf{u} \rangle - \langle \mathbf{D}, \mathbf{w} \rangle = \langle \mathbf{D}, \mathbf{u} \rangle - \alpha \cdot \langle \mathbf{D}, \mathbf{D} \rangle \stackrel{\text{note}}{=} \vec{\mathbf{0}}.$$

Thus the formula for the orthogonal vector is indeed

$$3: \quad \text{Orth}_{\mathbf{D}}(\mathbf{u}) := \mathbf{u} - \text{Proj}_{\mathbf{D}}(\mathbf{u}) = \frac{\langle \mathbf{D}, \mathbf{D} \rangle \mathbf{u} - \langle \mathbf{D}, \mathbf{u} \rangle \mathbf{D}}{\langle \mathbf{D}, \mathbf{D} \rangle}.$$

Let's compute the square-norms:

$$2': \|\text{Proj}_{\mathbf{D}}(\mathbf{u})\|^2 = \frac{\langle \mathbf{D}, \mathbf{u} \rangle \langle \mathbf{u}, \mathbf{D} \rangle}{\langle \mathbf{D}, \mathbf{D} \rangle} = \frac{|\langle \mathbf{u}, \mathbf{D} \rangle|^2}{\langle \mathbf{D}, \mathbf{D} \rangle}.$$

$$\begin{aligned} 3': \|\text{Orth}_{\mathbf{D}}(\mathbf{u})\|^2 &= \langle \mathbf{u}, \mathbf{u} \rangle - \|\text{Proj}_{\mathbf{D}}(\mathbf{u})\|^2 \\ &= \frac{\langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{D}, \mathbf{D} \rangle - \langle \mathbf{D}, \mathbf{u} \rangle \langle \mathbf{u}, \mathbf{D} \rangle}{\langle \mathbf{D}, \mathbf{D} \rangle}. \end{aligned}$$

NOTE: The closest pt on $\text{LinDir}(\mathbf{q}, \mathbf{D})$ to a point \mathbf{p} is

$$4: \quad \mathbf{q} + \text{Proj}_{\mathbf{D}}(\mathbf{p} - \mathbf{q}).$$

The vector from (4) to \mathbf{p} is $\text{Orth}_{\mathbf{D}}(\mathbf{p} - \mathbf{q})$.

OBS: The above eqns show (2') + (3') equals $\langle \mathbf{u}, \mathbf{u} \rangle$, which indeed is $\|\mathbf{u}\|^2$, the square of the hypotenuse.

Space for drawing:

Distance between two lines. Consider lines $\mathbb{L}_1 := \text{LinDir}(\mathbf{p}_1, \mathbf{D})$ and $\mathbb{L}_2 := \text{LinDir}(\mathbf{p}_2, \mathbf{E})$; the direction-vectors \mathbf{D} and \mathbf{E} are not $\vec{\mathbf{0}}$. The direction-vectors are *parallel* if the pair $\{\mathbf{D}, \mathbf{E}\}$ is linearly dependent; otherwise, the direction-vectors are *skew*.

Parallel Lines. WLOG, $\mathbf{D}=\mathbf{E}$. And $\text{Dist}(\mathbb{L}_1, \mathbb{L}_2)$ equals $\|\text{Orth}_{\mathbf{D}}(\mathbf{p}_1 - \mathbf{p}_2)\|$. Compute this via (3').

Skew Lines. Now \mathbf{D}, \mathbf{E} are *not* parallel. By Cauchy-Schwarz, then,

$$5.1: \quad \langle \mathbf{D}, \mathbf{D} \rangle \cdot \langle \mathbf{E}, \mathbf{E} \rangle \neq \langle \mathbf{D}, \mathbf{E} \rangle \cdot \langle \mathbf{E}, \mathbf{D} \rangle.$$

(Indeed, difference $\langle \mathbf{D}, \mathbf{D} \rangle \cdot \langle \mathbf{E}, \mathbf{E} \rangle - \langle \mathbf{D}, \mathbf{E} \rangle \cdot \langle \mathbf{E}, \mathbf{D} \rangle$ is positive.)

Let's translate the lines by $-\mathbf{p}_1$. I.e, write them as $\text{LinDir}(\vec{\mathbf{0}}, \mathbf{D})$ and $\text{LinDir}(\mathbf{b}, \mathbf{E})$, where $\mathbf{b} := \mathbf{p}_2 - \mathbf{p}_1$.

Take a "moving point" on each line. I.e, for "times" $s, t \in \mathbb{C}$:

$$5.2: \quad \begin{aligned} \mathbf{q}(s) &:= s\mathbf{D} && \text{and} \\ \mathbf{r}(t) &:= \mathbf{b} - t\mathbf{E}. \end{aligned}$$

Since \mathbf{D}, \mathbf{E} are not parallel, there will be a unique pair (s, t) of times such that *difference-vector* $\mathbf{q}(s) - \mathbf{r}(t)$ is orthogonal to each direction-vector.

Orthogonal to \mathbf{D} means $\langle \mathbf{D}, s\mathbf{D} \rangle - \langle \mathbf{D}, \mathbf{b} - t\mathbf{E} \rangle$ is zero. This is the first eqn below:

$$5.3: \quad \begin{aligned} \langle \mathbf{D}, \mathbf{D} \rangle s + \langle \mathbf{D}, \mathbf{E} \rangle t &= \langle \mathbf{D}, \mathbf{b} \rangle. \\ \langle \mathbf{E}, \mathbf{D} \rangle s + \langle \mathbf{E}, \mathbf{E} \rangle t &= \langle \mathbf{E}, \mathbf{b} \rangle. \end{aligned}$$

Orthogonality to \mathbf{E} yields the 2nd eqn. Courtesy (5.1), the coefficient-matrix

$$\mathbf{M} := \begin{bmatrix} \langle \mathbf{D}, \mathbf{D} \rangle & \langle \mathbf{D}, \mathbf{E} \rangle \\ \langle \mathbf{E}, \mathbf{D} \rangle & \langle \mathbf{E}, \mathbf{E} \rangle \end{bmatrix}$$

is non-singular. Let $R := 1/\text{Det}(\mathbf{M})$ be the reciprocal.

The unique solution $(s, t) := (\sigma, \tau)$ to (5.3) is thus

$$5.4: \quad \begin{bmatrix} \sigma \\ \tau \end{bmatrix} = R \cdot \begin{bmatrix} \langle \mathbf{E}, \mathbf{E} \rangle & -\langle \mathbf{D}, \mathbf{E} \rangle \\ -\langle \mathbf{E}, \mathbf{D} \rangle & \langle \mathbf{D}, \mathbf{D} \rangle \end{bmatrix} \cdot \begin{bmatrix} \langle \mathbf{D}, \mathbf{b} \rangle \\ \langle \mathbf{E}, \mathbf{b} \rangle \end{bmatrix}.$$

Plugging this (σ, τ) pair into (5.2) gives us the unique pair of closest points on the translated lines. So the closest points^{♥1} on the original lines are

$$5.5: \quad \begin{aligned} \mathbf{p}_1 + \mathbf{q}(\sigma) &= \mathbf{p}_1 + \sigma\mathbf{D} \in \mathbb{L}_1 && \text{and} \\ \mathbf{p}_1 + \mathbf{r}(\tau) &= \mathbf{p}_2 - \tau\mathbf{E} \in \mathbb{L}_2. \end{aligned}$$

^{♥1}The asymmetry in (5.5) comes from the asymmetry in \mathbf{b} . We had a choice between $[\mathbf{p}_2 - \mathbf{p}_1]$ and $[\mathbf{p}_1 - \mathbf{p}_2]$.

Gram-Schmidt alg. In real-or-complex IPS \mathbf{H} , consider vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots$. We will construct pairwise-orthogonal non- $\vec{0}$ vectors $\mathbf{b}_1, \mathbf{b}_2, \dots$ and natnums $K_0=0 < K_1 < K_2 < \dots$ so that this holds:

For each $n = 0, 1, 2, \dots$:
 Subspace $\mathbf{U}_n := \text{Spn}(\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\})$
 equals $\text{Spn}(\{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3, \dots, \mathbf{g}_{K_n}\})$.

For $n = 0$, our \mathbf{U}_0 equals $\{\vec{0}\}$, which indeed is the span of taking none of the \mathbf{g} -vectors.

At stage N : Successively set index κ to the indices after K_N . For each κ , compute

$$\widehat{\mathbf{g}}_\kappa := \text{Orth}_{\mathbf{U}_N}(\mathbf{g}_\kappa) \stackrel{\text{note}}{=} \mathbf{g}_\kappa - \sum_{j=1}^N \text{Proj}_{\mathbf{b}_j}(\mathbf{g}_\kappa),$$

stopping at the first κ where $\widehat{\mathbf{g}}_\kappa \neq \vec{0}$. Set $\mathbf{b}_{N+1} := \widehat{\mathbf{g}}_\kappa$ and $K_{N+1} := \kappa$. Now increment N .

Our Proj and Orth formulas. For vectors \mathbf{G} and $\mathbf{B} \neq \vec{0}$, recall

$$\text{Proj}_{\mathbf{B}}(\mathbf{G}) = \frac{\langle \mathbf{B}, \mathbf{G} \rangle}{\langle \mathbf{B}, \mathbf{B} \rangle} \cdot \mathbf{B} \quad \text{and}$$

$$\text{Orth}_{\mathbf{B}}(\mathbf{G}) = \mathbf{G} - \text{Proj}_{\mathbf{B}}(\mathbf{G}) = \frac{\langle \mathbf{B}, \mathbf{B} \rangle \mathbf{G} - \langle \mathbf{B}, \mathbf{G} \rangle \mathbf{B}}{\langle \mathbf{B}, \mathbf{B} \rangle}.$$

```
;; Given a list G1, G2, G3, ... of vectors, we
;; compute an orthogonal system
;; B1, B2, B3, ... with the same span.
;;
;; In particular if G1, G2, G3, ... was a basis,
;; then the B-system is an orthogonal basis.
;;
;; We could normalize the B-vectors, to get
;; an ortho-normal basis N1, N2, N3, ...
;; The dot-product w.r.t that ortho-normal basis will
;; equal the abstract inner-product that we started with.
```

```
(use-ring Rational-ring)

(setq G1 (mat-make-initseq 3 1 '(1 0 0)))
      [ 1 ]
      [ 0 ]
      [ 0 ]

(setq G2 (mat-make-initseq 3 1 '(2 -3 0)))
      [ 2 ]
      [ -3 ]
      [ 0 ]
```

```
(setq G3 (mat-make-initseq 3 1 '(-1 2 4)))
      [ -1 ]
      [ 2 ]
      [ 4 ]

;; Computing an orthogonal {B1, B2, B3} with the
;; same span as {G1, G2, G3}.
(setq B1 G1)
      [ 1 ]
      [ 0 ]
      [ 0 ]

(setq ratio (/ (ip B1 G2) (ip B1 B1))) -> 2

(setq ProjG2onB1 (mat-scal-mult ratio B1))
      [ 2 ]
      [ 0 ]
      [ 0 ]

(setq B2 (mat-sub G2 ProjG2onB1)) ;;This is Orth_B1(G2).
      [ 0 ]
      [ -3 ]
      [ 0 ]

(ip B1 B2) -> 0 ; Checking the B-vectors are orthogonal.

;; Computing B3.
(setq ProjG3onB1
      (mat-scal-mult (/ (ip B1 G3) (ip B1 B1)) B1) )
      [ -1 ]
      [ 0 ]
      [ 0 ]

(setq ProjG3onB2
      (mat-scal-mult (/ (ip B2 G3) (ip B2 B2)) B2) )
      [ 0 ]
      [ 2 ]
      [ 0 ]

(setq B3 (mat-sub G3 (mat-add ProjG3onB1 ProjG3onB2)))
      [ 0 ]
      [ 0 ]
      [ 4 ]

Easily
      [ -1 ]      [ 0 ]      [ 0 ]
      [ 0 ] ,    [ 2 ] ,    [ 0 ]
      [ 0 ]      [ 0 ]      [ 4 ]

is an orthogonal basis for |R^3.
=====
```

```

;; Given a list G1, G2, G3, ... of vectors, we compute
; an orthogonal system
; B1, B2, B3, ... with the same span.

;; A more interesting example, in 4-dim'al space.

% (setq G1 (mat-make-initseq 4 1 '(1 0 3 0)))
[ 1 ]
[ 0 ]
[ 3 ]
[ 0 ]

% (setq G2 (mat-make-initseq 4 1 '(4 2 2 -2)))
[ 4 ]
[ 2 ]
[ 2 ]
[ -2 ]

% (setq G3 (mat-make-initseq 4 1 '(-1 0 -1 2)))
[ -1 ]
[ 0 ]
[ -1 ]
[ 2 ]

;; Technically, B1 is the orthogonal-vector of G1
;; with respect to the trivial-subspace.
% (setq B1 G1)
[ 1 ]
[ 0 ]
[ 3 ]
[ 0 ]

% (setq B2 (mat-sub G2 (proj G2 B1))) ;Is Orth_B1(G2)
[ 3 ]
[ 2 ]
[ -1 ]
[ -2 ]

;; This next is the orthogonal-vector of G3
;; w.r.t Span(B1, B2).
% (setq B3
  (mat-sub G3 (mat-add (proj G3 B1) (proj G3 B2))) )
[ 2/5 ]
[ 2/3 ]
[ -2/15 ]
[ 4/3 ]

;; Let's multiply to make all the entries integers.
% (setq B3 (mat-scal-mult 15 B3))
[ 6 ]
[ 10 ]
[ -2 ]
[ 20 ]

;; We verify that the pairwise inner-products of the
;; B-vectors are zero:

% (list (ip B1 B2) (ip B1 B3) (ip B2 B3)) -> (0 0 0)

;; Let's check that Span(B1, B2, B3) = Span(G1, G2, G3).
% (setq bob (mat-Horiz-concat B1 B2 B3 G1 G2 G3))
[ 1 3 6 1 4 -1 ]
[ 0 2 10 0 2 0 ]
[ 3 -1 -2 3 2 -1 ]
[ 0 -2 20 0 -2 2 ]

% (rref-mtab-beforecol bob)
JK: Found 3 pivots before the sixth column.

      c0  c1  c2  c3  c4  c5
      |-----|
r0 |  1  0  0  1  1 -2/5 |
r1 |  0  1  0  0  1 -1/3 |
r2 |  0  0  1  0  0  1/15 |
r3 |  0  0  0  0  0  0   | ; Yay, Gram-Schmidt.
      |-----|

```

Ortho-Projecting on a subspace. The orthogonal projection of \mathbf{g} on a finite-dimensional subspace $\mathbf{U} \subset \mathbf{V}$ can be computed from any *orthogonal*-basis, \mathcal{U} , of \mathbf{U} , via

$$\text{Proj}_{\mathbf{U}}(\mathbf{g}) = \sum_{\mathbf{b} \in \mathcal{U}} \text{Proj}_{\mathbf{b}}(\mathbf{g}) = \sum_{\mathbf{b} \in \mathcal{U}} \frac{\langle \mathbf{b}, \mathbf{g} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle} \cdot \mathbf{b}. \text{ Thus,}$$

$$\text{Orth}_{\mathbf{U}}(\mathbf{g}) = \mathbf{g} - \text{Proj}_{\mathbf{U}}(\mathbf{g}).$$

Proj in polynomial space. On \mathbf{V} , the VS of polynomials, put IP $\langle f, g \rangle := \int_0^1 f \cdot g$. Henceforth, let \oint mean $\int_0^1 \cdot dx$.

Let's compute $\text{Orth}_{x^2}(x^3)$. Firstly

$$\langle x^2, x^3 \rangle = \oint x^5 = 1/6$$

$$\langle x^2, x^2 \rangle = \oint x^4 = 1/5. \text{ Thus,}$$

$$\text{Proj}_{x^2}(x^3) = \frac{1/6}{1/5} x^2 = \frac{5}{6} x^2. \text{ So,}$$

$$\text{Orth}_{x^2}(x^3) = x^3 - \text{Proj}_{x^2}(x^3) = x^3 - \frac{5}{6} x^2.$$

The Reader can check that $\langle x^2, x^3 - \frac{5}{6} x^2 \rangle$ is zero.

Conjugate-transpose. Imagine a 5×7 \mathbb{C} -matrix $M = [\alpha_{i,j}]_{i,j}$. The *complex-conjugate* of M is $\overline{M} = [\overline{\alpha_{i,j}}]_{i,j}$. The *transpose* of M is $M^t = [\alpha_{j,i}]_{i,j}$. Complex-conjugation and transpose are involutions, and they commute with each other.

The *conjugate-transpose* of M is

$$M^* := \overline{M^t} \stackrel{\text{note}}{=} \overline{M}^t.$$

This M^* is also called the *adjoint* of M . \square

Application: Suppose \mathbf{V} is a 5-dim'al \mathbb{C} -IP space, with (ordered) basis \mathcal{G} . Use Gram-Schmidt to produce an *orthogonal-basis* (i.e, *pairwise* orthogonal), then divide each vector by its norm, to get an orthonormal basis \mathcal{B} . W.r.t \mathcal{B} , vectors in \mathbf{V} are column-vectors, $\mathbf{u} = [\alpha_i]_i$ and $\mathbf{v} = [\beta_i]_i$. The improvement: The given IP is now the complex *dot-product*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \underbrace{\mathbf{u}^*}_{1 \times 5} \cdot \underbrace{\mathbf{v}}_{5 \times 1} = \sum_{i=1}^5 \overline{\alpha_i} \cdot \beta_i.$$

Cauchy-Schwarz Inequality. Two vectors \mathbf{u}, \mathbf{v} in an \mathcal{F} -VS are \mathcal{F} -parallel if, there exist scalars $\alpha, \beta \in \mathcal{F}$, not both 0, with $\alpha\mathbf{u} = \beta\mathbf{v}$; i.e. IFF $\{\mathbf{u}, \mathbf{v}\}$ is lin-dependent.

For \mathbb{C} -vectors \mathbf{u}, \mathbf{v} consider relation $\alpha\mathbf{u} = \beta\mathbf{v}$, for real α, β , not both 0. If we can pick $\alpha, \beta \geq 0$, then \mathbf{u} and \mathbf{v} are *same-direction parallel*.

If we can pick $\alpha \leq 0 \leq \beta$, then \mathbf{u} and \mathbf{v} are *opposite-direction parallel*. (Abbrev: “same-dir parallel” and “opp-dir parallel”.)

Obs. None of *same/opp-dir/parallel* is transitive. However, for \mathcal{R} either *parallel* or *same-dir parallel*,

If $\mathbf{u}\mathcal{R}\mathbf{c}$ and $\mathbf{c}\mathcal{R}\mathbf{v}$, and $\mathbf{c} \neq \vec{\mathbf{0}}$, then $\mathbf{u}\mathcal{R}\mathbf{v}$. \square

6: Cauchy-Schwarz Theorem. Consider vectors \mathbf{u}, \mathbf{v} in a complex IPS \mathbf{V} . Then

$$\begin{aligned} \text{6a:} \quad |\langle \mathbf{u}, \mathbf{v} \rangle|^2 &\leq \|\mathbf{u}\|^2 \cdot \|\mathbf{v}\|^2, & \text{i.e.} \\ |\langle \mathbf{u}, \mathbf{v} \rangle| &\leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|, \end{aligned}$$

with equality IFF \mathbf{u} and \mathbf{v} are parallel. Indeed $\langle \mathbf{u}, \mathbf{v} \rangle = \pm \|\mathbf{u}\| \cdot \|\mathbf{v}\|$ as the \mathbf{u}, \mathbf{v} pair is same/opp-dir parallel. \diamond

We get an immediate corollary.

6b: IP Triangle-Inequality. For all vectors \mathbf{u}, \mathbf{v} :

$$\text{†:} \quad \|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|. \quad \diamond$$

Proof. For $\alpha \in \mathbb{C}$, recall $\text{Re}(\alpha) \leq |\alpha|$. Working with squares-of-norms:

$$\begin{aligned} \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle &= \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle + 2\text{Re}(\langle \mathbf{u}, \mathbf{v} \rangle) \\ &\stackrel{(\ddagger) + \text{C-S}}{\leq} \text{Same} + \text{Same} + 2 \cdot \|\mathbf{u}\| \cdot \|\mathbf{v}\| \\ &= \left[\|\mathbf{u}\| + \|\mathbf{v}\| \right]^2. \end{aligned}$$

Since the end-terms are non-negative, square-rooting yields (\ddagger) . \diamond

6c: Lem. Consider quadratic $h(t) := At^2 - 2Bt + C$, where t, A, B, C are real, with $A > 0$. Let $\tau \in \mathbb{R}$ be the min-point of h , i.e. $[\forall t \in \mathbb{R}: h(t) \geq h(\tau)]$. Then

$$\text{6d:} \quad \tau = \frac{B}{A}, \text{ and min-value } h(\tau) \text{ equals } C - \frac{B^2}{A}. \quad \diamond$$

Pf. Well, $h'(t) = 2At - 2B$, and $h'(\tau) = 0$. Etc. \blacklozenge

Pf of C-S, (6), for a real-VS. If $\mathbf{v} = \vec{\mathbf{0}}$, then the thm’s conclusion holds, so WLOG $\mathbf{v} \neq \vec{\mathbf{0}}$.

Define $f: \mathbb{R} \rightarrow \mathbb{R}$ by $f(t) := \|\mathbf{u} - t\mathbf{v}\|^2$. I claim that

$$\begin{aligned} f(t) &\stackrel{?}{=} At^2 - 2Bt + C, \quad \text{where} \\ A &:= \langle \mathbf{v}, \mathbf{v} \rangle \quad \text{and} \quad B := \langle \mathbf{v}, \mathbf{u} \rangle \quad \text{and} \quad C := \langle \mathbf{u}, \mathbf{u} \rangle. \end{aligned}$$

Why? Our $f(t)$ equals $\langle \mathbf{u} - t\mathbf{v}, \mathbf{u} - t\mathbf{v} \rangle$ which equals

$$\text{*}: \quad \langle t\mathbf{v}, t\mathbf{v} \rangle - \left[\langle t\mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{u}, t\mathbf{v} \rangle \right] + \langle \mathbf{u}, \mathbf{u} \rangle.$$

Since t is real, $f(t)$ equals

$$\begin{aligned} &= t^2 \langle \mathbf{v}, \mathbf{v} \rangle - \left[t \langle \mathbf{v}, \mathbf{u} \rangle + t \langle \mathbf{u}, \mathbf{v} \rangle \right] + C \\ &= t^2 A - t \cdot 2B + C. \end{aligned}$$

As f is the square of a distance, the min-value of f is non-neg. Thus $0 \leq C - \frac{B^2}{A}$, by (6c). Finally, $A > 0$ so $\boxed{B^2 \leq AC}$, which is (6a). \blacklozenge

Pf of C-S for a \mathbb{C} -VS. We proceed as above, but allowing a complex t . Now $(*)$ equals

$$\text{**}: \quad |t|^2 \langle \mathbf{v}, \mathbf{v} \rangle - \left[\bar{t} \langle \mathbf{v}, \mathbf{u} \rangle + t \langle \mathbf{u}, \mathbf{v} \rangle \right] + \langle \mathbf{u}, \mathbf{u} \rangle.$$

In the real case, we plugged $\tau := \frac{B}{A} = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}$ in for t , and we do that here as well. Notice that

$$\tau \cdot \langle \mathbf{u}, \mathbf{v} \rangle = \frac{|\langle \mathbf{v}, \mathbf{u} \rangle|^2}{\langle \mathbf{v}, \mathbf{v} \rangle} \stackrel{\text{note}}{\in} \mathbb{R}.$$

Plugging τ into $(**)$ says that $f(\tau)$ equals

$$\frac{|\langle \mathbf{v}, \mathbf{u} \rangle|^2}{\langle \mathbf{v}, \mathbf{v} \rangle} - 2 \cdot \frac{|\langle \mathbf{v}, \mathbf{u} \rangle|^2}{\langle \mathbf{v}, \mathbf{v} \rangle} + \langle \mathbf{u}, \mathbf{u} \rangle \stackrel{\text{note}}{=} \langle \mathbf{u}, \mathbf{u} \rangle - \frac{|\langle \mathbf{v}, \mathbf{u} \rangle|^2}{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

Thus $\frac{|\langle \mathbf{v}, \mathbf{u} \rangle|^2}{\langle \mathbf{v}, \mathbf{v} \rangle} \leq \langle \mathbf{u}, \mathbf{u} \rangle$, Now multiply by $\langle \mathbf{v}, \mathbf{v} \rangle$. \blacklozenge

When do we have equality in Cauchy-Schwarz? Note that vector pair $\{\mathbf{u}, \mathbf{v}\}$ is linearly dependent (i.e, \mathbf{u} and \mathbf{v} are parallel) IFF $\text{Orth}_{\mathbf{v}}(\mathbf{u})$ is $\vec{\mathbf{0}}$.

Equality in (6a) IFF parallel. WLOG, both \mathbf{u} and \mathbf{v} are non-zero. Set $\mathbf{z} := \text{Orth}_{\mathbf{v}}(\mathbf{u})$; so $\mathbf{u} = \alpha\mathbf{v} + \mathbf{z}$, for some (complex) α . WLOG $\alpha \neq 0$. [If not, then $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ but $\|\mathbf{u}\|^2\|\mathbf{v}\|^2$ is positive, since neither \mathbf{u} nor \mathbf{v} is $\vec{\mathbf{0}}$. And \mathbf{u} is not parallel to \mathbf{v} , as $\mathbf{u} \perp \mathbf{v}$ yet neither vector is $\vec{\mathbf{0}}$.]

As (in)equality in (6a) is unaffected by multiplying \mathbf{u} by a non-zero scalar, replace \mathbf{u} by $\frac{1}{\alpha}\mathbf{u}$. I.e, WLOG $\mathbf{u} = \mathbf{v} + \mathbf{z}$ with $\mathbf{v} \perp \mathbf{z}$. Computing,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v} + \mathbf{z}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle + 0 \stackrel{\text{note}}{=} \|\mathbf{v}\|^2.$$

$$\langle \mathbf{u}, \mathbf{u} \rangle = \|\mathbf{v}\|^2 + \|\mathbf{z}\|^2 + \langle \mathbf{v}, \mathbf{z} \rangle + \langle \mathbf{z}, \mathbf{v} \rangle = \|\mathbf{v}\|^2 + \|\mathbf{z}\|^2.$$

$$\langle \mathbf{v}, \mathbf{v} \rangle = \|\mathbf{v}\|^2.$$

Our C-S says $|\langle \mathbf{u}, \mathbf{v} \rangle|^2 \leq \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle$. Substituting in the above formulas, C-S asserts

$$\|\mathbf{v}\|^2 \cdot \|\mathbf{v}\|^2 \leq \left[\|\mathbf{v}\|^2 + \|\mathbf{z}\|^2 \right] \cdot \|\mathbf{v}\|^2.$$

Dividing by the positive $\|\mathbf{v}\|^2$ yields

$$\|\mathbf{v}\|^2 \leq \|\mathbf{v}\|^2 + \|\mathbf{z}\|^2.$$

We have equality here IFF $\mathbf{z} = \vec{\mathbf{0}}$, i.e, \mathbf{u} was parallel to \mathbf{v} all along. \blacklozenge

Distance of flat to origin. In a \mathbb{C} -IPS \mathbf{V} list of points P_1, \dots, P_6 engenders this *flat* (affine subspace)

$$\mathbb{A} := \left\{ \sum_{n=1}^6 \alpha_n P_n \mid \sum_{n=1}^6 \alpha_n = 1 \right\}.$$

In order to locate the point $Q \in \mathbb{A}$ closest to the origin, we pick a vector in \mathbb{A} , say, P_6 , and translate \mathbb{A} to parallel *subspace* $\mathbb{S} := \mathbb{A} - P_6$. CLAIM:

‡: $\mathbb{S} = \text{Spn}(\mathbf{d}_1, \dots, \mathbf{d}_5)$, where $\mathbf{d}_n := P_n - P_6$.

Why? An \mathbb{A} -vector $\mathbf{a} = \sum_{n=1}^6 \alpha_n P_n$ has difference-vector $\mathbf{a} - P_6$ equaling

$$\left[\sum_{n=1}^6 \alpha_n P_n \right] - \left[\sum_{n=1}^6 \alpha_n \right] P_6 = \sum_{n=1}^5 \alpha_n \mathbf{d}_n.$$

Conversely, given vector $\sum_{n=1}^5 \alpha_n \mathbf{d}_n$ in $\text{Spn}(\mathbf{d}_1, \dots, \mathbf{d}_5)$, we can define α_6 to be $1 - \sum_{n=1}^5 \alpha_n$, thus arranging that $\sum_{n=1}^6 \alpha_n P_n$ is indeed in \mathbb{A} .

To compute the \mathbb{A} -point Q lying closest to the origin, we can pick any pt in \mathbb{A} , say, P_6 , and compute

$$\ddagger: \quad Q = \text{Orth}_{\mathbb{S}}(P_6) = P_6 - \text{Proj}_{\mathbb{S}}(P_6).$$