**Distance between flats**

Jonathan L.F. King
*University of Florida, Gainesville FL 32611-2082, USA*
squash@ufl.edu
Webpage http://squash.1gainesville.com/
23 September, 2017 (at *14:28*)

For a general field $\mathcal{F}$ and vectors $\mathbf{u}, \mathbf{w}$ in an $\mathcal{F}$-vector-space, define the "**line through u in direction w**":

$$\mathrm{LinDir}(\mathbf{u}, \mathbf{w}) := \{\mathbf{u} + t\mathbf{w} \mid t \in \mathcal{F}\}.$$

We now happily specialize to an IP (inner-product) space $\mathbf{V}$ over field $\mathcal{F} \subset \mathbb{C}$ which is sealed under complex-conjugation;

$$\forall\alpha \in \mathbb{C}: \quad \alpha \in \mathcal{F} \implies \overline{\alpha} \in \mathcal{F}.$$

Our IP is conjugate-linear in its $1^{\text{st}}$ argument. I.e, for every $\beta \in \mathcal{F}$ and $\mathbf{u}, \mathbf{w} \in \mathbf{V}$:

$$\langle \beta\mathbf{u}, \mathbf{w}\rangle = \overline{\beta} \cdot \langle \mathbf{u}, \mathbf{w}\rangle \quad \text{and} \quad \langle \mathbf{u}, \beta\mathbf{w}\rangle = \beta \cdot \langle \mathbf{u}, \mathbf{w}\rangle.$$

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

1: $$\mathrm{Proj}_{\mathbf{D}}(\mathbf{u}) := \frac{\langle \mathbf{D}, \mathbf{u}\rangle}{\langle \mathbf{D}, \mathbf{D}\rangle} \cdot \mathbf{D}.$$

Our IP is linear in its $2^{\text{nd}}$ argument, so formula (1) indeed satisfies that $\mathrm{Proj}_{\mathbf{D}}(\beta\mathbf{u}) = \beta \cdot \mathrm{Proj}_{\mathbf{D}}(\mathbf{u})$.

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

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

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

Thus the formula for the orthogonal vector is indeed

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

Let's compute the square-norms:

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

2′: $$\|\mathrm{Orth}_{\mathbf{D}}(\mathbf{u})\|^2 = \langle \mathbf{u}, \mathbf{u}\rangle - \|\mathrm{Proj}_{\mathbf{D}}(\mathbf{u})\|^2 = \frac{\langle \mathbf{u}, \mathbf{u}\rangle\langle \mathbf{D}, \mathbf{D}\rangle - \langle \mathbf{D}, \mathbf{u}\rangle\langle \mathbf{u}, \mathbf{D}\rangle}{\langle \mathbf{D}, \mathbf{D}\rangle}.$$

Hence the closest point on line $\mathrm{LinDir}(\mathbf{q}, \mathbf{D})$ to a point $\mathbf{p}$ is

3: $$\mathbf{q} + \mathrm{Proj}_{\mathbf{D}}(\mathbf{p} - \mathbf{q}).$$

The vector which goes from (3) to $\mathbf{p}$ is $\mathrm{Orth}_{\mathbf{D}}(\mathbf{p} - \mathbf{q})$.

**Distance between two lines.** Consider lines $\mathbb{L}_1 := \mathrm{LinDir}(\mathbf{p}_1, \mathbf{D})$ and $\mathbb{L}_2 := \mathrm{LinDir}(\mathbf{p}_2, \mathbf{E})$; the direction vectors $\mathbf{D}$ and $\mathbf{E}$ are not $\mathbf{0}$.

***Parallel Lines.*** WLOG, $\mathbf{D} = \mathbf{E}$. And $\mathrm{Dist}(\mathbb{L}_1, \mathbb{L}_2)$ equals $\|\mathrm{Orth}_{\mathbf{D}}(\mathbf{p}_1 - \mathbf{p}_2)\|$. Compute this via (2′).

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

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

(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 $\mathrm{LinDir}(\mathbf{0}, \mathbf{D})$ and $\mathrm{LinDir}(\mathbf{b}, \mathbf{E})$, where $\boxed{\mathbf{b} := \mathbf{p}_2 - \mathbf{p}_1}$.

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

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

Since $\mathbf{D}, \mathbf{E}$ are not parallel, there will be a unique pair $(s, t)$ of times such that $\mathbf{q}(s) - \mathbf{r}(t)$ 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:

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

Orthogonality to $\mathbf{E}$ yields the $2^{\text{nd}}$ eqn. Courtesy (4), the coefficient-matrix

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

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

The unique solution $(s, t) := (\sigma, \tau)$ to (6) is thus

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

Plugging this $(\sigma, \tau)$ pair into (5) gives us the unique pair of closest points on the translated lines. So the closest points[♡1] on the original lines are

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

[♡1] The asymmetry in (8) 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 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots$. We will construct pairwise-orthogonal non-$\mathbf{0}$ vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots$ and natnums $K_0{=}0 < K_1 < K_2 < \ldots$ so that this holds:

*For each* $n = 0, 1, 2, \ldots$:
*Subspace* $\mathbf{W}_n := \operatorname{Spn}(\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\})$
*equals* $\operatorname{Spn}(\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{b}_{K_n}\})$.

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

At stage $N$: Successively set $\kappa$ to the indices *after* $K_N$. For each, compute

$$\mathbf{g}_\kappa := \operatorname{Orth}_{\mathbf{W}_N}(\mathbf{v}_\kappa) \overset{\text{note}}{=\!=} \mathbf{v}_\kappa - \sum_{j=1}^{N} \operatorname{Proj}_{\mathbf{b}_j}(\mathbf{v}_\kappa),$$

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

**Cauchy-Schwarz Inequality.** Two vectors $\mathbf{v}, \mathbf{w}$ in an $\mathcal{F}$-VS are ***$\mathcal{F}$-parallel*** if, there exists scalars $\alpha, \beta \in \mathcal{F}$, not both 0, with $\alpha\mathbf{v} = \beta\mathbf{w}$; i.e, if $\{\mathbf{v}, \mathbf{w}\}$ is lindependent.

Now let $\mathcal{F}$ be a subfield of $\mathbb{C}$ and consider relation $\alpha\mathbf{v} = \beta\mathbf{w}$, for real $\alpha$,$\beta$, not both 0. If we can pick $\alpha$,$\beta \geqslant 0$, then $\mathbf{v}$ and $\mathbf{w}$ are ***same-direction parallel***. If we can pick $\alpha \leqslant 0 \leqslant \beta$, then $\mathbf{v}$ and $\mathbf{w}$ are ***opposite-direction parallel***. (Abbrev: "same-dir parallel" and "opp-dir parallel".)

*Remark.* None of *same/opp-dir parallel* is transitive. However, let $\mathcal{R}$ be any one of the three relations. Then

*If* $\mathbf{v}\mathcal{R}\mathbf{c}$ *and* $\mathbf{c}\mathcal{R}\mathbf{w}$, <u>*and*</u> $\mathbf{c} \neq \mathbf{0}$, *then* $\mathbf{v}\mathcal{R}\mathbf{w}$. □

9: Cauchy-Schwarz Theorem. *Consider vectors* $\mathbf{v}, \mathbf{w}$ *in a complex IPS* $\mathbf{V}$. *Then*

9a: $$\left|\langle \mathbf{v}, \mathbf{w}\rangle\right|^2 \leqslant \|\mathbf{v}\|^2 \cdot \|\mathbf{w}\|^2,$$

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

9b: Lem. *Consider quadratic* $h(t) := At^2 - 2Bt + C$, *where* $A$,$B$,$C$ *are real, with* $A > 0$. *Let* $\boldsymbol{\tau} \in \mathbb{R}$ *be the min-point of* $h$, *i.e,* $[\forall t \in \mathbb{R}\colon\ h(t) \geqslant h(\boldsymbol{\tau})]$. *Then*

9c: $\boldsymbol{\tau} = \frac{B}{A}$, *and min-value* $h(\boldsymbol{\tau})$ *equals* $C - \frac{B^2}{A}$. ♢

*Pf.* Well, $h'(t) = 2At - 2B$, and $h'(\boldsymbol{\tau}) = 0$. Etc. ♦

*Pf of C-S, (9), for a $\mathbb{R}$-VS.* If $\mathbf{w} = \mathbf{0}$, then the thm's conclusion holds, so WLOG $\mathbf{w} \neq \mathbf{0}$.

Define $f(t) := \|\mathbf{v} - t\mathbf{w}\|^2$. Courtesy bilinearity,

$$f(t) = At^2 - 2Bt + C,$$

where $A := \langle \mathbf{w}, \mathbf{w}\rangle$, $C := \langle \mathbf{v}, \mathbf{v}\rangle$ and $B := \langle \mathbf{v}, \mathbf{w}\rangle$. Since $f$ is the square of a distance, the min-value of $f$ is non-negative. Thus $0 \leqslant C - \frac{B^2}{A}$. But $A > 0$, so $\boxed{B^2 \leqslant AC}$. And this is (9a). ♦

Filename: `Problems/Algebra/LinearAlg/flat-dist.latex`
As of: *Friday 16Dec2005.* Typeset: *23Sep2017 at 14:28.*