PRINCIPAL ANGLES IN TERMS OF INNER PRODUCTS

CLAY SHONKWILER

1. Introduction

Suppose A and B are two k-planes in $\mathbb{R}^{2k}$. The goal of this note is to find a "nice" way to determine the principal angles $\theta_1, \ldots, \theta_k$ between A and B.

This is motivated by the study of Poincaré Duality angles, which are defined to be the principal angles between certain k-planes in the space of differential p-forms on a Riemannian manifold with boundary. The details are not relevant here, but it is clear that finding a computationally manageable way of determining principal angles will be relevant.

Before we get started, let's recall the definition of the principal angles between a k-plane and an ℓ-plane in n-space. The ith principal angle θ_i between a k-plane A and an ℓ-plane B is defined by the equation

$$\begin{aligned}\cos\theta_i &= \frac{\langle a_i, b_i\rangle}{\|a_i\|\|b_i\|} \\ &= \max\left\{\frac{\langle a, b\rangle}{\|a\|\|b\|} : a \perp a_m, b \perp b_m, m = 1, 2, \ldots, i-1\right\}\end{aligned}$$

where the $a_j \in A$, $b_j \in B$.

In words, the procedure is to find the unit vector $a_1 \in A$ and the unit vector $b_1 \in B$ which minimize the angle between them and call this angle θ_1. Now take the orthogonal complement of a_1 in A and the orthogonal complement of b_1 in B and iterate.

In the context of Poincaré Duality angles, $k = \ell$, but the following procedure should apply to situations where $k \neq \ell$ as well.

It is easy to check (as stated in [1] in the case $k = 2$) that there exists an orthonormal basis $\{x_1, \ldots, x_{2k}\}$ for $\mathbb{R}^{2k}$ such that

$$\{\alpha_1, \ldots, \alpha_k\} := \{x_1, \ldots, x_k\}$$

is an orthonormal basis for A and

$$\{\beta_1, \ldots, \beta_k\} := \{\cos\theta_1\, x_1 + \sin\theta_1\, x_{k+1}, \ldots, \cos\theta_k\, x_k + \sin\theta_k\, x_{2k}\}$$

is an orthonormal basis for B. This is, of course, a particularly nice choice of bases for A and B since the angle between α_i and β_i is exactly the principal angle θ_i for all $i = 1, \ldots, k$.

If we already knew the bases $\{\alpha_1, \ldots, \alpha_k\}$ and $\{\beta_1, \ldots, \beta_k\}$ for A and B, finding the principal angles would be trivial.

In general, though, we want to be able to determine the principal angles given only *some* basis $\{a_1, \ldots, a_k\}$ for A and *some* basis $\{b_1, \ldots, b_k\}$ for B.

In fact, it would be even better if we didn't need to know exactly what the vectors a_i and b_j are, only what all the possible inner products between them are (i.e. $\langle a_i, a_j\rangle$, $\langle a_i, b_j\rangle$ and $\langle b_i, b_j\rangle$ for all choices of i and j). The purpose of this note is to demonstrate that we can completely determine the principal angles between A and B given only this inner product data.

2. The trivial case

We start with the case where $a_i = \alpha_i$ and $b_i = \beta_i$ for all $i = 1, \ldots, k$. In other words, suppose that the bases for A and B that we start with are, by some miracle, the bases which are already perfectly adapted to determining the principal angles.

In this case,

$$\cos\theta_i = \langle x_i, \cos\theta_i\, x_i + \sin\theta_i\, x_{k+i}\rangle = \langle \alpha_i, \beta_i\rangle,$$

so we need only take inner products of corresponding α's and β's and we're done.

Geometrically what are we doing? Notice that the orthogonal projection of β_i onto A is given by

$$\langle \beta_i, \alpha_1\rangle\, \alpha_1 + \ldots + \langle \beta_i, \alpha_k\rangle\, \alpha_k = \langle \beta_i, \alpha_i\rangle\, \alpha_i = \cos\theta_i\, \alpha_i.$$

So the principal angle θ_i is really just the length of the orthogonal projection of β_i onto A. This makes it seem like the orthogonal projection map $\mathsf{Pr} : B \to A$ is a useful map to study.

In terms of the bases $\{\alpha_1, \ldots, \alpha_k\}$ and $\{\beta_1, \ldots, \beta_k\}$ for A and B, Pr can be represented by the diagonal matrix

$$\Sigma := \begin{pmatrix} \cos\theta_1 & 0 & \cdots & 0 \\ 0 & \cos\theta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cos\theta_k \end{pmatrix}.$$

More completely, Pr is represented by the matrix

$$\Sigma := \left(\langle \alpha_i, \beta_j\rangle\right)_{i,j}.$$

Note that the determinant of this matrix is

$$\det\Sigma = \prod_{i=1}^{k} \cos\theta_i.$$

This makes perfect sense because the determinant of Σ should measure how much Pr scales volume. If we consider a unit cube in B with edges given by the β_i, then its projection to A will have edges scaled by the appropriate $\cos\theta_i$. Thus, projecting the cube scales its volume by the product of the $\cos\theta_i$.

It is tempting to interpret the $\cos\theta_i$ as the eigenvalues of Pr (with β_i as their corresponding eigenvectors), but remember that the domain and range of Pr are different k-planes, so the β_i are only eigenvectors of Pr under the abstract identification of B with A via the map determined by $\beta_i \mapsto \alpha_i$. It's more fruitful to think of the $\cos\theta_i$ as singular values of Pr, which implies that the $\cos^2\theta_i$ are eigenvalues of $\mathsf{Pr}^* \mathsf{Pr}$.

$\mathsf{Pr}^* \mathsf{Pr}$ is simply the map from B to itself given by orthogonally projecting B to A, then orthogonally projecting A to B. It is clear that $\mathsf{Pr}^* \mathsf{Pr}\,\beta_i = \cos^2\theta\beta_i$ and here it really does make sense to call the β_i eigenvectors. With respect to the basis $\{\beta_1, \ldots, \beta_k\}$, the matrix for $\mathsf{Pr}^* \mathsf{Pr}$ is simply

$$\Sigma^*\Sigma = \Sigma^2.$$

3. An arbitrary orthonormal basis

Of course, the odds that randomly selected bases for A and B coincide with the nice bases $\{\alpha_1, \ldots, \alpha_k\}$ and $\{\beta_1, \ldots, \beta_k\}$ are not good. We want to be able to use arbitrary bases $\{a_1 \ldots, a_k\}$ and $\{b_1, \ldots, b_k\}$ for A and B to find the principal angles.

For the purposes of this note, let's make the simplifying assumption that the bases $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_k\}$ are orthonormal. This is not a very restrictive assumption because, given arbitrary bases, we can always use, e.g., Gram-Schmidt to produce orthonormal bases. Of course, it would be best to find a technique for determining the principal angles without needing to invoke Gram-Schmidt, but we'll save that problem for another day.

Given some orthonormal basis $\{a_1, \ldots, a_k\}$ for A, we know that there exists some $g \in O(k)$ such that $a_i = g(\alpha_i)$ for all $i = 1, \ldots, k$. Similarly, if $\{b_1, \ldots, b_k\}$ is an orthonormal basis for B, there exists $h \in O(k)$ such that $b_i = h(\beta_i)$ for all $i = 1, \ldots, k$. (Note: it would probably be more accurate to say that g lives in $O(A)$ and h lives in $O(B)$ because, though these groups are both isomorphic to $O(k)$, they are different groups.)

Now, we want to express Pr as a matrix in terms of the bases $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_k\}$. On one hand, since

$$\mathsf{Pr}(b_i) = \langle b_i, a_1\rangle a_1 + \ldots + \langle b_i, a_k\rangle a_k,$$

it is clear that, in terms of these bases, the matrix for Pr is

$$M := \left(\langle b_j, a_i\rangle\right)_{i,j} = \left(\langle a_i, b_j\rangle\right)_{i,j}.$$

On the other hand, if $G = (g_{ij})_{i,j}$ is the matrix for g with respect to the basis $\{\alpha_1, \ldots, \alpha_k\}$ and $H = (h_{ij})_{i,j}$ is the matrix for h with respect to the basis $\{\beta_1, \ldots, \beta_k\}$, then

$$M = G\,\Sigma\,H^*,$$

where H^* is the transpose of H (remember that Σ is the matrix for Pr in terms of the bases $\{\alpha_1, \ldots, \alpha_k\}$ and $\{\beta_1, \ldots, \beta_k\}$).

But notice that G and H are orthogonal matrices and Σ is a diagonal matrix, so $G\,\Sigma\,H^*$ is a singular value decomposition for M. This confirms the idea that the $\cos\theta_i$ are singular values of Pr.

Of course, in practice we will have no idea what G, Σ and H are, but we don't actually need them to be able to determine the $\cos\theta_i$. Remember that the $\cos^2\theta_i$ are the eigenvalues of $\mathsf{Pr}^*\,\mathsf{Pr}$. In terms of the basis $\{b_1,\ldots,b_k\}$, the matrix for $\mathsf{Pr}^*\,\mathsf{Pr}$ is

$$M^*M = (G\,\Sigma\,H^*)^*\,(G\,\Sigma\,H^*) = H\,\Sigma^*\,G^*\,G\,\Sigma\,H^* = H\,\Sigma^2\,H^*.$$

(Of course, we could have also seen this directly: since Σ^2 is the matrix for $\mathsf{Pr}^*\,\mathsf{Pr}$ with respect to the basis $\{\beta_1,\ldots,\beta_k\}$ and H is the change-of-basis matrix, it must be the case that the matrix for $\mathsf{Pr}^*\,\mathsf{Pr}$ is $H\,\Sigma^2\,H^*$.)

Since the entries of M are simply the $\langle a_i, b_j\rangle$,

$$M^*M = \left(\sum_{m,n=1}^{k}\langle a_m, b_i\rangle\langle a_n, b_j\rangle\right)_{i,j}.$$

Hence, the $\cos^2\theta_i$ can be determined purely in terms of these inner products. Since the θ_i are always between 0 and $\pi/2$ there are no ambiguities about taking square roots or inverting the cosine, so we see that the θ_i can indeed be determined from the inner product data.

4. Random Remarks

Note that

$$\prod_{i=1}^{k}\cos^2\theta_i = \det M^*M = (\det M)^2\,, \tag{1}$$

so this gives an alternate proof a result of Jiang [2] (the $k=2$ case of which appears in [1]).

Also,

$$\sum_{i=1}^{k}\cos^2\theta_i = \operatorname{tr} M^*M = \sum_{i,j=1}^{k}\langle a_i, b_j\rangle^2, \tag{2}$$

the $k=2$ case of which was proved in a previous version of this note.

In the case $k=2$,

$$M^*M = \begin{pmatrix} \langle a_1,b_1\rangle^2 + \langle a_2,b_1\rangle^2 & \langle a_1,b_1\rangle\langle a_1,b_2\rangle + \langle a_2,b_1\rangle\langle a_2,b_2\rangle \\ \langle a_1,b_2\rangle\langle a_1,b_1\rangle + \langle a_2,b_2\rangle\langle a_2,b_1\rangle & \langle a_1,b_2\rangle^2 + \langle a_2,b_2\rangle^2 \end{pmatrix}.$$

The determinant of this matrix is

$$\begin{aligned}\left(\langle a_1,b_1\rangle^2 + \langle a_2,b_1\rangle^2\right)\left(\langle a_1,b_2\rangle^2 + \langle a_2,b_2\rangle^2\right) - (\langle a_1,b_1\rangle\langle a_1,b_2\rangle + \langle a_2,b_1\rangle\langle a_2,b_2\rangle)^2 \\ = (\langle a_1,b_1\rangle\langle a_2,b_2\rangle - \langle a_1,b_2\rangle\langle a_2,b_1\rangle)^2\end{aligned}$$

so, by (1)

$$\cos\theta_1\cos\theta_2 = \langle a_1,b_1\rangle\langle a_2,b_2\rangle - \langle a_1,b_2\rangle\langle a_2,b_1\rangle.$$

This, along with (2), then implies that

$$\cos\theta_1 = \frac{\sqrt{(\langle a_1, b_1\rangle + \langle a_2, b_2\rangle)^2 + (\langle a_1, b_2\rangle - \langle a_2, b_1\rangle)^2} + \sqrt{(\langle a_1, b_1\rangle - \langle a_2, b_2\rangle)^2 + (\langle a_1, b_2\rangle + \langle a_2, b_1\rangle)^2}}{2}$$

$$\cos\theta_2 = \frac{\sqrt{(\langle a_1, b_1\rangle + \langle a_2, b_2\rangle)^2 + (\langle a_1, b_2\rangle - \langle a_2, b_1\rangle)^2} - \sqrt{(\langle a_1, b_1\rangle - \langle a_2, b_2\rangle)^2 + (\langle a_1, b_2\rangle + \langle a_2, b_1\rangle)^2}}{2},$$

agreeing with the formulas found in a previous version of this note.

References

[1] Herman Gluck and Frank W. Warner: Great circle fibrations of the three-sphere. *Duke Math. J.*, 50:107–132, 1983. [doi:10.1215/S0012-7094-83-05003-2].

[2] Sheng Jiang: Angles between Euclidean subspaces. *Geometriae Dedicata*, 63(2):113–121, 1996. [doi:10.1007/BF00148212].

DRL 3E3A, University of Pennsylvania
E-mail address: `shonkwil@math.upenn.edu`