# Angles between Subspaces of a Normed Space

Hendra Gunawan*

Bandung Institute of Technology
*http://personal.fmipa.itb.ac.id/hgunawan/

Pure Maths Seminar – UNSW, 4 December 2018

# Outline

# Outline

# Outline

The notion of **angles** between two subspaces of the Euclidean space $\mathbb{R}^d$ has been studied by many researchers since the 1950's or even earlier [DK][1].

In statistics, canonical (or principal) angles are studied as measures of **dependency** of one set of random variables on another [A][2].

[1] C. Davies and W. Kahan, "The rotation of eigenvectors by a perturbation. III", *SIAM J. Numer. Anal.* **7** (1970), 1–46

[2] T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, Inc., New York, 1958

In [RT][3], Risteski and Trenčevski defined the angle $\theta$ between two subspaces $U = \mathrm{span}\{u_1, \ldots, u_p\}$ and $V = \mathrm{span}\{v_1, \ldots, v_q\}$ of $\mathbb{R}^d$ where $p \leq q$ by

$$\cos^2\theta := \frac{\det(MM^{\mathrm{T}})}{\det[\langle u_i, u_j\rangle] \cdot \det[\langle v_k, v_l\rangle]}, \tag{1}$$

based on the following inequality [RT, Theorem 1.1]:

$$\det(MM^{\mathrm{T}}) \leq \det[\langle u_i, u_j\rangle] \cdot \det[\langle v_k, v_l\rangle], \tag{2}$$

where $M := [\langle u_i, v_k\rangle]$. Here $\langle\cdot,\cdot\rangle$ denotes the usual inner product on $\mathbb{R}^d$.

[3] I.B. Risteski and K.G. Trenčevski, "Principal values and principal subspaces of two subspaces of vector spaces with inner product", *Beiträge Algebra Geom.* **42** (2001), 289–300

However, the inequality (2) is only true in the case (a) where $p = q$ (for which the inequality reduces to Kurepa's generalization of the **Cauchy-Schwarz inequality** [K][4]), or (b) where $\{v_1, \ldots, v_q\}$ is orthonormal.

Consequently, (1) makes sense only in these two cases, for otherwise the value of the expression on the right hand side of (1) may be greater than 1.

[4] S. Kurepa, "On the Buniakowsky-Cauchy-Schwarz inequality", *Glasnik Mat. Ser. III* **1**(21) (1966), 147–158

To see that the inequality (2) is **false** in general, just take $X = \mathbb{R}^3$ (equipped with the usual inner product), $U = \mathrm{span}\{u\}$ where $u = (1, 0, 0)$, and $V = \mathrm{span}\{v_1, v_2\}$ where $v_1 = (\frac{1}{2}, \frac{1}{2}, 0)$ and $v_2 = (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$. According to (2), we should have

$$\langle u, v_1\rangle^2 + \langle u, v_2\rangle^2 \leq \|u\|^2\big(\|v_1\|^2\|v_2\|^2 - \langle v_1, v_2\rangle^2\big).$$

But the left hand side equals $\frac{1}{2}$, while the right hand side equals $\frac{3}{8}$.

This example shows that the inequality is false even in the case where $\{u_1, \dots, u_p\}$ is orthonormal and $\{v_1, \dots, v_q\}$ is orthogonal (which is close to being orthonormal).

## HOW DO WE FIX IT?

Let $(X, \langle \cdot, \cdot \rangle)$ be a (real) **inner product space**. Given two nonzero, finite-dimensional, subspaces $U$ and $V$ of $X$ with $\dim(U) \leq \dim(V)$, we wish to have a definition of the angle between $U$ and $V$ that can be viewed as a generalization of the 'usual' definition of the angle

(a) between a 1-dimensional subspace and a $q$-dimensional subspace of $X$, and

(b) between two $p$-dimensional subspaces intersecting on a common $(p-1)$-dimensional subspace of $X$.

**In (a):**

If $U = \mathrm{span}\{u\}$ is a 1-dimensional subspace and $V = \mathrm{span}\{v_1, \dots, v_q\}$ is a $q$-dimensional subspace of $X$, then the angle $\theta$ between $U$ and $V$ is defined by

$$\cos^2\theta = \frac{\langle u, u_V\rangle^2}{\|u\|^2\|u_V\|^2} \tag{3}$$

where $u_V$ denotes the (orthogonal) projection of $u$ on $V$ and $\|\cdot\| = \langle\cdot,\cdot\rangle^{\frac{1}{2}}$ denotes the induced norm on $X$.

Note: We shall take $\theta$ to be in the interval $[0, \frac{\pi}{2}]$.

**In (b):**

If $U = \mathrm{span}\{u, w_2, \ldots, w_p\}$ and $V = \mathrm{span}\{v, w_2, \ldots, w_p\}$ are $p$-dimensional subspaces of $X$ that intersects on $(p-1)$-dimensional subspace $W = \mathrm{span}\{w_2, \ldots, w_p\}$ with $p \geq 2$, then the angle $\theta$ between $U$ and $V$ may be defined by

$$\cos^2\theta = \frac{\langle u_W^{\perp}, v_W^{\perp}\rangle^2}{\|u_W^{\perp}\|^2 \|v_W^{\perp}\|^2} \tag{4}$$

where $u_W^{\perp}$ and $v_W^{\perp}$ are the orthogonal complement of $u$ and $v$, respectively, on $W$.

One common property among these two cases is the following.

In (a), we may write $u = u_V + u_V^\perp$ where $u_V^\perp$ is the orthogonal complement of $u$ on $V$. Then (3) amounts to

$$\cos^2\theta = \frac{\|u_V\|^2}{\|u\|^2},$$

which tells us that the value of $\cos\theta$ is equal to the **ratio** between the length of the projection of $u$ on $V$ and the length of $u$.

Similarly, in (b), one may check that the value of $\cos\theta$ is equal to the ratio between the volume of the $p$-dimensional parallelepiped spanned by the projection of $u, w_2, \dots, w_p$ on $V$ and the volume of the $p$-dimensional parallelepiped spanned by $u, w_2, \dots, w_p$.

Motivated by the above fact, we may define the angle between a $p$-dimensional subspace $U = \mathrm{span}\{u_1, \dots, u_p\}$ and a $q$-dimensional subspace $V = \mathrm{span}\{v_1, \dots, v_q\}$ (with $p \leq q$) such that the value of its cosine is equal to the **ratio** between the volume of the $p$-dimensional parallelepiped spanned by the projection of $u_1, \dots, u_p$ on $V$ and the $p$-dimensional parallelepiped spanned by $u_1, \dots, u_p$, that is,

$$\cos^2\theta := \frac{\|u_{1V}, \dots, u_{pV}\|^2}{\|u_1, \dots, u_p\|^2}, \tag{5}$$

where $u_{iV}$ denotes the projection of $u_i$ on $V$ and $\|u_1, \dots, u_p\|$ denotes the volume of the $p$-dimensional parallelepiped spanned by $u_1, \dots, u_p$.

The volume of the $p$-dimensional parallelepiped spanned by $u_1, \dots, u_p$ is given by

$$\|u_1, \dots, u_p\| := \begin{vmatrix} \langle u_1, u_1 \rangle & \langle u_1, u_2 \rangle & \dots & \langle u_1, u_p \rangle \\ \langle u_2, u_1 \rangle & \langle u_2, u_2 \rangle & \dots & \langle u_2, u_p \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle u_p, u_1 \rangle & \langle u_p, u_2 \rangle & \dots & \langle u_p, u_p \rangle \end{vmatrix}^{1/2}.$$

The determinant of the matrix on the right hand side is known as the **Gram determinant**, which is non-negative. It vanishes if and only if $u_1, \dots, u_p$ are linearly dependent.

The mapping $\|\cdot, \dots, \cdot\|$ on $X^p$ is known as the **standard $n$-norm** on $X$.

Note that $\|u_1, \dots, u_p\| \leq \|u_1\| \times \cdots \times \|u_p\|$ (Hadamard's inequality).

**Fact**. *The ratio on the right hand side of (5) is a number in* $[0,1]$ *and is independent of the choice of bases for* $U$ *and* $V$*.*

*Proof*. First note that the projection of $u_i$'s on $V$ is independent of the choice of basis for $V$. Further, since projections are linear transformations, the ratio is also invariant under any change of basis for $U$. Indeed, the ratio is unchanged if we (a) swap $u_i$ and $u_j$, (b) replace $u_i$ by $u_i + \alpha u_j$, or (c) replace $u_i$ by $\alpha u_i$ with $\alpha \neq 0$.

Next, assuming particularly that $\{u_1, \ldots, u_p\}$ is orthonormal, we have $\|u_1, \ldots, u_p\| = 1$ and $\|u_{1V}, \ldots, u_{pV}\| \leq 1$ because $\|u_{iV}\| \leq \|u_i\| = 1$ for each $i = 1, \ldots, p$. Therefore, the ratio is a number in $[0,1]$, and the proof is complete. □

# An explicit formula for arbitrary $p$ and $q$

From (5), we can derive an explicit formula for the cosine in terms of $u_1, \ldots, u_p$ and $v_1, \ldots, v_q$, particularly when $\{v_1, \ldots, v_q\}$ is orthonormal. For each $i = 1, \ldots, p$, the projection of $u_i$ on $V$ is given by

$$u_{iV} := \mathrm{proj}_V u_i = \langle u_i, v_1 \rangle v_1 + \cdots + \langle u_i, v_q \rangle v_q.$$

So, for $i, j = 1, \ldots, p$, we have

$$\langle u_{iV}, u_{jV} \rangle = \sum_{k=1}^{q} \langle u_i, v_k \rangle \langle u_j, v_k \rangle$$

Hence, we obtain

$$\|u_{1V}, \ldots, u_{pV}\|^2 = \det\left[\sum_{k=1}^{q} \langle u_i, v_k \rangle \langle u_j, v_k \rangle\right] = \det(MM^{\mathrm{T}})$$

where $M := [\langle u_i, v_k \rangle]$ is a $p \times q$ matrix and $M^{\mathrm{T}}$ is its transpose.

The cosine of the angle $\theta$ between $U$ and $V$ is therefore given by the formula

$$\cos^2\theta = \frac{\det(MM^{\mathrm{T}})}{\det[\langle u_i, u_j\rangle]}. \qquad (6)$$

If $\{u_1, \ldots, u_p\}$ happens to be orthonormal, then the formula (6) reduces to

$$\cos^2\theta = \det(MM^{\mathrm{T}}).$$

Further, if $p = q$, then $\det(MM^{\mathrm{T}}) = \det M \cdot \det M^{\mathrm{T}} = \det^2 M$. Hence, from the last formula, we get $\cos\theta = |\det M|$.

# A more general formula

For each $i = 1, \ldots, p$, the projection of $u_i$ on $V$ may be expressed as

$$u_{iV} := \sum_{k=1}^{q} \alpha_{ik} v_k$$

where

$$\alpha_{ik} = \frac{\langle u_i, v_k | v_{i_2(k)}, \ldots, v_{i_q(k)} \rangle}{\| v_1, v_2, \ldots, v_q \|^2}$$

with $\{i_2(k), \ldots, i_q(k)\} = \{1, 2, \ldots, q\} \setminus \{k\}, \; k = 1, 2, \ldots, q$.

Here $\langle \cdot, \cdot | \cdot, \ldots, \cdot \rangle$ denotes the *n*-**inner product** on $X$, given by

$$\langle x_0, x_1 | x_2, \ldots, x_n \rangle := \begin{vmatrix} \langle x_0, x_1 \rangle & \langle x_0, x_2 \rangle & \ldots & \langle x_0, x_n \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \ldots & \langle x_2, x_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_n, x_1 \rangle & \langle x_n, x_2 \rangle & \ldots & \langle x_n, x_n \rangle \end{vmatrix}.$$

Next observe that

$$\langle u_{iV}, u_{jV}\rangle = \langle u_i, u_{jV}\rangle = \sum_{k=1}^{q} \alpha_{jk}\langle u_i, v_k\rangle$$

for $i, j = 1, \ldots, p$.

Hence we have

$$\begin{aligned} \|u_{1V}, \ldots, u_{pV}\|^2 &= \begin{vmatrix} \sum_{k=1}^{q} \alpha_{1k}\langle u_1, v_k\rangle & \ldots & \sum_{k=1}^{q} \alpha_{pk}\langle u_1, v_k\rangle \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{q} \alpha_{1k}\langle u_p, v_k\rangle & \ldots & \sum_{k=1}^{q} \alpha_{pk}\langle u_p, v_k\rangle \end{vmatrix} \\ &= \frac{\det(M\tilde{M}^{\mathrm{T}})}{\|v_1, \ldots, v_q\|^{2p}} \end{aligned}$$

where

$$M := [\langle u_i, v_k\rangle] \quad \text{and} \quad \tilde{M} := [\langle u_i, v_k | v_{i_2(k)}, \ldots, v_{i_q(k)}\rangle] \tag{7}$$

with $i_2(k), \ldots, i_q(k)$ as above.

Note that both $M$ and $\tilde{M}$ are $p \times q$ matrices, so that $M\tilde{M}^{\mathrm{T}}$ is a $p \times p$ matrix.

Dividing by $\|u_1, \ldots, u_p\|^2$, we get the following formula for the cosine [GNS] [5]:

$$\cos^2\theta = \frac{\det(M\tilde{M}^{\mathrm{T}})}{\det[\langle u_i, u_j\rangle] \cdot \det^p[\langle v_k, v_l\rangle]}, \tag{8}$$

which serves as a correction for Risteski and Trenčevski's formula (1).

Note that if $\{v_1, \ldots, v_q\}$ is orthonormal, we get the formula (6) obtained earlier.

[5] H. Gunawan, O. Neswan, and W. Setya-Budhi, "A formula for angles between two subspaces of inner product spaces", *Beitr. Algebra Geom.* **46**–2 (2005), 311–320

To define angles between two subspaces of a **normed space**, even between two vectors in a normed space, we need to have the notion of **projections** and **volumes of parallelepipeds** in a normed space.

# $g$ functionals

From now on, let $(X, \|\cdot\|)$ be a normed space. The functional $g : X^2 \to \mathbb{R}$ defined by the formula

$$g(x,y) := \frac{1}{2}\|x\|\left[\tau_+(x,y) + \tau_-(x,y)\right]$$

with

$$\tau_\pm(x,y) := \lim_{t\to 0^\pm} \frac{\|x+ty\| - \|x\|}{t},$$

satisfies the following properties:
(1) $g(x,x) = \|x\|^2$ for every $x \in X$;
(2) $g(ax, by) = ab \cdot g(x,y)$ for every $x, y \in X$ and $a, b \in \mathbb{R}$;
(3) $g(x, x+y) = \|x\|^2 + g(x,y)$ for every $x, y \in X$;
(4) $|g(x,y)| \leq \|x\| \cdot \|y\|$ for every $x, y \in X$.

If, in addition, the functional $g(x,y)$ is linear in $y$, then $g$ is called a **semi-inner product** on $X$.

For example, the functional

$$g(x,y) := \|x\|_p^{2-p} \sum_{k=1}^{\infty} |\xi_k|^{p-1} \mathrm{sgn}(\xi_k)\eta_k, \quad x = (\xi_k),\ y = (\eta_k) \in \ell^p,$$

is a semi-inner product on $\ell^p$ $(1 \le p < \infty)$ [Gi][6].

Note that in general, $g$ is not commutative.

[6] J.R. Giles, "Classes of semi-inner product spaces", *Trans. Amer. Math. Soc.* **129** (1967), 436–446

Using the functional $g$, we can define the $g$-orthogonal projection of $y$ on a subspace $S$ as follows:

Let $y$ be a vector of $X$ and $S = \text{span}\{x_1, \dots, x_n\}$ be a subspace of $X$ with $\Gamma(x_1, \dots, x_n) = \det[g(x_i, x_k)] \neq 0$.

The $g$-**orthogonal projection** of $y$ on $S$, denoted by $y_S$, is defined by

$$y_S := -\frac{1}{\Gamma(x_1, \dots, x_n)} \begin{vmatrix} 0 & x_1 & \cdots & x_n \\ g(x_1, y) & g(x_1, x_1) & \cdots & g(x_1, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ g(x_n, y) & g(x_n, x_1) & \cdots & g(x_n, x_n) \end{vmatrix}.$$

The $g$-orthogonal projection of $y$ on $S$ is independent of the choice of basis for $S$ [M][7].

[7] P.M. Milič̀ić, "On the Gram-Schmidt projection in normed spaces", *Univ. Beograd Publ. Elektrotehn. Fak. Ser. Mat.* **4** (1993), 89–96

In companion with the $g$-projection of $y$ on $S$, we have the $g$-**orthogonal complement** $y - y_S$, given by

$$y - y_S = \frac{1}{\Gamma(x_1, \dots, x_n)} \begin{vmatrix} y & x_1 & \cdots & x_n \\ g(x_1, y) & g(x_1, x_1) & \cdots & g(x_1, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ g(x_n, y) & g(x_n, x_1) & \cdots & g(x_n, x_n) \end{vmatrix}.$$

Note that $g(x_i, y - y_S) = 0$ for every $i = 1, \dots, n$.

# Angles between 1-dimensional and $n$-dimensional subspaces of $X$

Let $U := \mathrm{span}\{u\}$ be a 1-dimensional subspace of $X$ and $V := \mathrm{span}\{v_1, \dots, v_n\}$ be an $n$-dimensional subspace of $X$. Then we can define the angle $\theta$ between $U$ dan $V$ by the formula

$$\cos^2\theta := \frac{\|u_V\|^2}{\|u\|^2}.$$

Since $u = u_V + (u - u_V)$, we have

$$\cos^2\theta = \frac{[g(u_V, u)]^2}{\|u_V\|^2\|u\|^2},$$

confirming that the right hand side is a number in $[0, 1]$.

# Angles between 2-dimensional subspaces of $X$

Now we need the notion of volumes of parallelepipeds in $X$. There is the notion of $n$-norms on $X$, but unlike on an inner product space, the $n$-norms generally do not satisfy Hadamard's inequality.

However, we have a formula for the angle between 2-dimensional subspaces of $X$, using a specific 2-norm which involves a semi-inner product.

# 2-norms

A mapping $\|\cdot,\cdot\| : X^2 \to \mathbb{R}$ which satisfies the following four properties:
(1) $\|x, y\| = 0$ if and only if $x$ and $y$ are linearly dependent;
(2) $\|x, y\| = \|y, x\|$ for every $x, y \in X$;
(3) $\|\alpha x, y\| = |\alpha| \, \|x, y\|$ for every $x, y \in X$ and $\alpha \in \mathbb{R}$;
(4) $\|x, y + z\| \leq \|x, y\| + \|x, z\|$ for every $x, y, z \in Z$,
is called a 2-**norm** on $X$.

It is known that every normed space can also be equipped with a 2-norm [Ga][8].

[8] S. Gähler, "Untersuchungen über verallgemeinerte *m*-metrische Räume. I", *Math. Nachr.* **40** (1969), 165–189

Let $g(\cdot,\cdot)$ be a semi-inner product on $X$. We define the mapping $\|\cdot,\cdot\|_g : X^2 \to \mathbb{R}$ by

$$\|x_1, x_2\|_g := \sup_{y_i \in X,\ \|y_i\| \le 1} \left| \begin{matrix} g(y_1, x_1) & g(y_1, x_2) \\ g(y_2, x_1) & g(y_2, x_2) \end{matrix} \right|.$$

Then one may check that this mapping defines a 2-norm on $X$.

**Fact**. If $(X, \langle \cdot, \cdot \rangle)$ is an inner product space and $\|\cdot, \cdot\|_s$ is the standard 2-norm, that is,

$$\|x_1, x_2\|_s := \begin{vmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle \end{vmatrix}^{1/2},$$

then the two formulas $\|\cdot, \cdot\|_g$ and $\|\cdot, \cdot\|_s$ are identical.

Note: In $(X, \langle \cdot, \cdot \rangle)$, the semi-inner product $g(\cdot, \cdot)$ is identical with the inner product $\langle \cdot, \cdot \rangle$.

# A lemma

**Lemma**. If $U = \mathrm{span}\{u_1, u_2\}$ and $V = \mathrm{span}\{v_1, v_2\}$ are 2-dimensional subspaces of $X$ where $\{v_1, v_2\}$ is $g$-left orthonormal (that is, $\|v_1\| = \|v_2\| = 1$ and $g(v_1, v_2) = 0$). Then

$$\begin{vmatrix} g(y_1, u_{1V}) & g(y_1, u_{2V}) \\ g(y_2, u_{1V}) & g(y_2, u_{2V}) \end{vmatrix} = \begin{vmatrix} g(v_1, u_1) & g(v_1, u_2) \\ g(v_2, u_1) & g(v_2, u_2) \end{vmatrix} \begin{vmatrix} g(y_1, v_1) & g(y_1, v_2) \\ g(y_2, v_1) & g(y_2, v_2) \end{vmatrix}$$

for every $y_1, y_2 \in X$.

We define the angle $\theta$ between $U = \mathrm{span}\{u_1, u_2\}$ and $V = \mathrm{span}\{v_1, v_2\}$ with $\Gamma(v_1, v_2) \neq 0$ by

$$\cos^2\theta := \frac{\|u_{1V}, u_{2V}\|_g^2}{\|u_1, u_2\|_g^2 \|v_1^*, v_2^*\|_g^2}, \tag{9}$$

where $u_{iV}$ denotes the $g$-orthogonal projection of $u_i$ on $V$ $(i = 1, 2)$ and $\{v_1^*, v_2^*\}$ is the left $g$-orthonormal set obtained from $\{v_1, v_2\}$ via an analog of Gram-Schmidt process.

Note: In an inner product space $(X, \langle\cdot,\cdot\rangle)$, the formula becomes

$$\cos^2\theta = \frac{\|u_{1V}, u_{2V}\|_s^2}{\|u_1, u_2\|_s^2}.$$

**Theorem**. The ratio on the right hand side of (9) is a number in $[0,1]$ and is independent of the choice of bases for $U$ and $V$.

*Proof*. The independence of the choice of the bases for $U$ and $V$ is clear.

Assume that $\{v_1, v_2\}$ is left $g$-orthonormal. By using the previous lemma and the definition of $\|\cdot,\cdot\|_g$, we have

$$\|u_{1V}, u_{2V}\|_g = \begin{vmatrix} g(v_1,u_1) & g(v_1,u_2) \\ g(v_2,u_1) & g(v_2,u_2) \end{vmatrix} \|v_1, v_2\|_g .$$

Accordingly,

$$\|u_{1V}, u_{2V}\|_g \le \|u_1, u_2\|_g \|v_1, v_2\|_g ,$$

and so the ratio in (9) is a number in $[0,1]$.

**Acknowledgement**. The results in the second part is joint with M. Nur. This visit is funded by P3MI-ITB 2018 Grant.