

Discussion 8B

The following notes are useful for this discussion: [Note 13](#).

1. Gram-Schmidt Algorithm

Let's apply Gram-Schmidt orthonormalization to a list of three linearly independent vectors $[\vec{s}_1, \vec{s}_2, \vec{s}_3]$.

- (a) Let's say we had two collections of vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ and $\{\vec{w}_1, \dots, \vec{w}_n\}$. **How can we prove that $\text{Span}(\{\vec{v}_1, \dots, \vec{v}_n\}) = \text{Span}(\{\vec{w}_1, \dots, \vec{w}_n\})$?**

- (b) **Find unit vector \vec{q}_1 such that $\text{Span}(\{\vec{q}_1\}) = \text{Span}(\{\vec{s}_1\})$, where \vec{s}_1 is nonzero.**

- (c) Let's say that we wanted to write

$$\vec{s}_2 = c_1 \vec{q}_1 + \vec{z}_2 \tag{1}$$

where $c_1 \vec{q}_1$ entirely represents the component of \vec{s}_2 in the direction of \vec{q}_1 , and \vec{z}_2 represents the component of \vec{s}_2 that is distinctly *not* in the direction of \vec{q}_1 (i.e. \vec{z}_2 and \vec{q}_1 are orthogonal).

Given \vec{q}_1 from the previous step, **find c_1 as in eq. (1), and use \vec{z}_2 to find unit vector \vec{q}_2 such that $\text{Span}(\{\vec{q}_1, \vec{q}_2\}) = \text{Span}(\{\vec{s}_1, \vec{s}_2\})$ and \vec{q}_2 is orthogonal to \vec{q}_1 . Show that $\text{Span}(\{\vec{q}_1, \vec{q}_2\}) = \text{Span}(\{\vec{s}_1, \vec{s}_2\})$.**

(d) What would happen if $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$ were *not* linearly independent, but rather \vec{s}_1 were a multiple of \vec{s}_2 ?

(e) Now given \vec{q}_1 and \vec{q}_2 in parts 1.b and 1.c, find \vec{q}_3 such that $\text{Span}(\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}) = \text{Span}(\{\vec{s}_1, \vec{s}_2, \vec{s}_3\})$, and \vec{q}_3 is orthogonal to both \vec{q}_1 and \vec{q}_2 , and finally $\|\vec{q}_3\| = 1$. You do not have to show that the two spans are equal.

(f) (PRACTICE) Confirm that $\text{Span}(\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}) = \text{Span}(\{\vec{s}_1, \vec{s}_2, \vec{s}_3\})$.

2. Orthonormal Matrices and Projections

A matrix A has orthonormal columns, \vec{a}_i , if they are:

- Orthogonal (ie. $\langle \vec{a}_i, \vec{a}_j \rangle = \vec{a}_j^\top \vec{a}_i = 0$ when $i \neq j$)
- Normalized (ie. vectors with length equal to 1, $\|\vec{a}_i\| = 1$). This implies that $\|\vec{a}_i\|_2 = \langle \vec{a}_i, \vec{a}_i \rangle = \vec{a}_i^\top \vec{a}_i = 1$.

(a) When $A \in \mathbb{R}^{n \times m}$ and $n \geq m$ (i.e. for tall matrices), show that if the matrix is orthonormal, then $A^\top A = I_{m \times m}$.

(b) Again, suppose $A \in \mathbb{R}^{n \times m}$ where $n \geq m$ is an orthonormal matrix. **Show that the projection of \vec{y} onto the subspace spanned by the columns of A is now $AA^T\vec{y}$.**

(c) **(PRACTICE)** Show if $A \in \mathbb{R}^{n \times n}$ is an orthonormal matrix then the columns, \vec{a}_i , form a basis for \mathbb{R}^n .

Contributors:

- Anish Muthali.
- Regina Eckert.
- Druv Pai.
- Neelesh Ramachandran.