

Math 43 Review Notes - Chapter 4

Orthogonal Subspaces

Recall that two vectors \mathbf{v} and \mathbf{w} are orthogonal (i.e. perpendicular) if their dot product $\mathbf{v} \cdot \mathbf{w}$ is 0. If we think of \mathbf{v} and \mathbf{w} as arrays with only one column, then we can also write the dot product of \mathbf{v} and \mathbf{w} as a multiplication of matrices, namely $\mathbf{v}^T \mathbf{w}$. You might see this from time to time in the book.

We say that two *subspaces* V and W of a vector space are orthogonal if every vector in V is orthogonal to every vector in W . To determine if two subspaces are orthogonal, it is enough just to check that each vector in a basis for V is orthogonal to each vector in a basis for W . Here is an important example of orthogonal subspaces:

Example: The nullspace and row space of a matrix A are orthogonal subspaces. Remember, vectors in the nullspace of a matrix A are vectors \mathbf{x} for which $A\mathbf{x} = \mathbf{0}$, and vectors in the row space are linear combinations of the rows of A . A small example is enough to understand why the subspaces are orthogonal. Let

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{pmatrix}.$$

Let $\mathbf{x} = (x_1, x_2, x_3)$ be in the nullspace of A . Then $A\mathbf{x} = \mathbf{0}$. Write this out in equation form:

$$\begin{aligned} 2x_1 + 3x_2 + 4x_3 &= 0 \\ 5x_1 + 6x_2 + 7x_3 &= 0 \end{aligned}$$

Looking closely at these, we see the first equation is just the dot product of row 1 and \mathbf{x} , and the second is the dot product of row 2 and \mathbf{x} . Since these are zero, the equations are saying that row 1 and row 2 are orthogonal to \mathbf{x} . So any combination of the rows is also orthogonal to \mathbf{x} . Therefore we have shown that the row space and nullspace are orthogonal.

Orthogonal Complement

The **orthogonal complement** of a subspace V consists of *all* vectors which are orthogonal to each vector in V . It is denoted by V^\perp . Like before, to find V^\perp it is enough to find all vectors which are orthogonal each vector in a basis for V .

Example: Let V be the subspace which consists of all vectors $\mathbf{x} = (x_1, x_2, x_3)$ which satisfy $x_1 + 3x_2 + 2x_3 = 0$ (a plane in \mathbb{R}^3). Find V^\perp .

Solution: First find a basis for V . Solve the equation defining V to get $x_1 = -3x_2 - 2x_3$. Then any \mathbf{x} in V can be written as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

So $(-3, 0, 1)$ and $(-2, 1, 0)$ are a basis for V . (You can check for yourself that they are linearly independent.) Any vector $\mathbf{b} = (b_1, b_2, b_3)$ in V^\perp is orthogonal to both of these vectors. In other words:

$$\begin{aligned} (b_1, b_2, b_3) \cdot (-3, 0, 1) &= 0 &\longrightarrow & -3b_1 + 0 + b_3 = 0 &\longrightarrow & b_3 = 3b_1 \\ (b_1, b_2, b_3) \cdot (-2, 1, 0) &= 0 &\longrightarrow & -2b_1 + b_2 + 0 = 0 &\longrightarrow & b_2 = 2b_1 \end{aligned}$$

Therefore, plugging these into \mathbf{b} we get

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ 2b_1 \\ 3b_1 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

So, V^\perp consists of all multiples of the vector $(1, 2, 3)$. (V^\perp is a line in \mathbb{R}^3 , perpendicular to the plane $x_1 + 3x_2 + 2x_3 = 0$.)

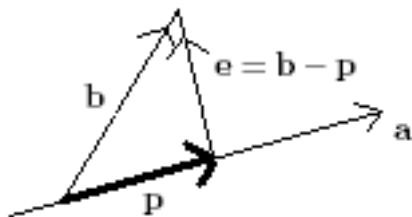
Example: Not only are the row space and nullspace orthogonal, but in fact they are orthogonal complements of each other. Try to see why.

Make sure you understand the difference between the definitions. The first definition is about a relationship between two subspaces: V and W are orthogonal if every vector in V is orthogonal to every vector in W . Notice also that the only vector that can be in both V and W is $\mathbf{0}$, since it is the only vector whose dot product with itself is 0 (i.e. it is the only vector which is perpendicular to itself).

The second definition gives a new subspace V^\perp , the orthogonal complement, which contains *every single vector* which is orthogonal to all the vectors in V .

Projections

First we will look at projecting a vector \mathbf{b} onto another vector \mathbf{a} . One way to think about a projection is as the “shadow” that \mathbf{b} casts on \mathbf{a} . Another way is to think of it as asking how much of \mathbf{b} is in the direction of \mathbf{a} .



The projection of \mathbf{b} onto \mathbf{a} is given by

$$\mathbf{p} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a}$$

The fractional term gives the percentage or fraction of \mathbf{a} that the shadow of \mathbf{b} takes up; in the book it is called \hat{x} .

Example: Project $\mathbf{b} = (2, 4, 1)$ onto $\mathbf{a} = (3, 2, 5)$ and $\mathbf{c} = (3, -1, -2)$.

Solution: Projecting \mathbf{b} onto \mathbf{a} we get

$$\mathbf{p} = \frac{(2, 4, 1) \cdot (3, 2, 5)}{(3, 2, 5) \cdot (3, 2, 5)} \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} = \frac{19}{38} \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 1 \\ 5/2 \end{bmatrix}$$

Projecting \mathbf{b} onto \mathbf{c} we get

$$\mathbf{p} = \frac{(2, 4, 1) \cdot (3, -1, -2)}{(3, -1, -2) \cdot (3, -1, -2)} \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix} = \frac{0}{14} \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The vectors \mathbf{b} and \mathbf{c} are perpendicular (since their dot product is 0), hence the projection of \mathbf{b} onto \mathbf{c} is just $\mathbf{0}$, as we would expect.

There’s one other way to think about projections. Instead of just projecting \mathbf{b} onto \mathbf{a} , think about it as projecting \mathbf{b} onto the line through \mathbf{a} . Then \mathbf{p} is the closest vector to \mathbf{b} which lies on the line. The vector \mathbf{e} in the picture is the “error” between \mathbf{b} and \mathbf{p} . It is just $\mathbf{b} - \mathbf{p}$.

Since a line is just a special kind of subspace, this thinking gives us a way to project a vector \mathbf{b} onto any subspace V . The projection will be the vector in V which is closest to \mathbf{b} . We will consider projecting onto the column space of a matrix A . In other words, we are looking for the vector which is a linear combination of the columns of A which is closest to \mathbf{b} . To do this we use the following:

$$P = A(A^T A)^{-1} A^T$$

$$\mathbf{p} = P\mathbf{b}$$

The matrix P is called the projection matrix. It will project any vector onto the column space of A .

Example: Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix}$. Find the projection matrix P , and the projection of $\mathbf{b} = (1, 2, 3)$ onto the column space of A .

Solution: Compute:

$$A^T A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$(A^T A)^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

$$A(A^T A)^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix}$$

$$P = A(A^T A)^{-1} A^T = \begin{pmatrix} 2 & -1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

$$\mathbf{p} = P\mathbf{b} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Orthogonal Bases

Recall that vectors are orthogonal if their dot product is 0. If in addition to being orthogonal the vectors are all unit vectors (i.e. have length 1), then we say the vectors are **orthonormal**. A convenient way to say this is that the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ are orthonormal if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ unless $i = j$, in which case it equals 1.

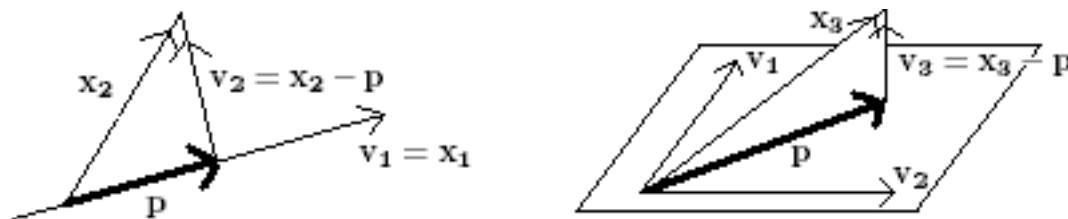
Example: Let $\mathbf{v}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$.

It is easy to check that $\mathbf{v}_1 \cdot \mathbf{v}_1 = \mathbf{v}_2 \cdot \mathbf{v}_2 = \mathbf{v}_3 \cdot \mathbf{v}_3 = 1$, and $\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot \mathbf{v}_3 = \mathbf{v}_2 \cdot \mathbf{v}_3 = 0$. So the vectors are orthonormal.

The Gram-Schmidt Process

We say that a basis is an orthogonal basis if its vectors are orthogonal, and is an orthonormal basis if its vectors are orthonormal. For many applications orthogonal and orthonormal bases are much easier to work with than other bases. (Remember that a space has many different bases.) The Gram-Schmidt process gives a way of converting a basis into an orthogonal or orthonormal basis. You start with n linearly independent vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ and Gram-Schmidt will create orthogonal vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ which span the same space as the \mathbf{x} 's. In other words, Gram-Schmidt takes a basis for a subspace and turns it into an orthogonal basis for the same subspace.

The idea relies on projections. Start by letting $\mathbf{v}_1 = \mathbf{x}_1$. Then let \mathbf{p} be the projection of \mathbf{x}_2 onto \mathbf{v}_1 . Notice that $\mathbf{x}_2 - \mathbf{p}$ is orthogonal to \mathbf{v}_1 . So let this be \mathbf{v}_2 . Notice that \mathbf{v}_1 and \mathbf{v}_2 span the same space as \mathbf{x}_1 and \mathbf{x}_2 . Now consider the projection \mathbf{p} of \mathbf{x}_3 onto the subspace spanned by \mathbf{v}_1 and \mathbf{v}_2 . Like before, $\mathbf{x}_3 - \mathbf{p}$ is orthogonal to both of \mathbf{v}_1 and \mathbf{v}_2 , and since \mathbf{v}_1 and \mathbf{v}_2 are orthogonal, it has a simple formula (no messy matrix calculations). Continue projecting to get the rest of the \mathbf{v} 's.



Here is the process:

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2$$

$$\mathbf{v}_4 = \mathbf{x}_4 - \frac{\mathbf{x}_4 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{x}_4 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 - \frac{\mathbf{x}_4 \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3} \mathbf{v}_3$$

and so on ...

To make this set of vectors orthonormal, the vectors need to be unit vectors. Remember how to do that, we divide each vector by its length, i.e. our orthonormal vectors are

$$\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|}, \dots$$

Example: Find a set of orthonormal vectors which span the same space as $\mathbf{x}_1 = (1, 1, 1)$, $\mathbf{x}_2 = (1, 2, 3)$, and $\mathbf{x}_3 = (4, 3, 8)$.

Solution: Check for yourself that the given vectors are linearly independent. Now use Gram-Schmidt to find the corresponding orthogonal vectors \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 .

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \frac{1+2+3}{1+1+1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{v}_3 = \begin{bmatrix} 4 \\ 3 \\ 8 \end{bmatrix} - \frac{4+3+8}{1+1+1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{-4+0+8}{1+0+1} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

To make $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 orthonormal, divide each by its length to get

$$\frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Update: (12/13/05) An easier way to find V^\perp .

This method uses the fact that the row space and nullspace of a matrix are orthogonal. To find V^\perp , first find a basis for V . Next create a matrix A from the basis vectors by letting them be the *rows* of A . Then V^\perp is the nullspace of A . In fact, this is really very similar to the method on the first page.