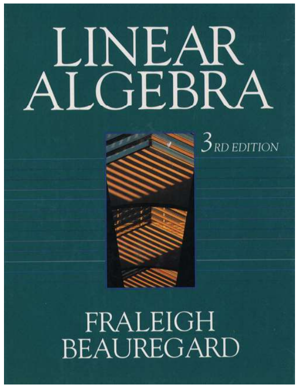
Linear Algebra

Chapter 6: Orthogonality

Section 6.1. Projections—Proofs of Theorems



Page 336 number 4

Page 336 number 4. Find the projection of $[1, 2, 1]$ on the line with parametric equation $x = 3t$, $y = t$, $z = 2t$ in $\mathbb{R}^3$.

Solution. A line is a translation of a one-dimensional subspace and is of the form $\vec{x} = t\vec{d} + \vec{a}$ where $\vec{d}$ is the direction vector and $\vec{a}$ is a translation vector (see Section 2.5, "Lines, Planes, and Other Flats"). Here, $\vec{d} = [3, 1, 2]$ and $\vec{a} = [0, 0, 0]$ so, in fact, the line is not translated and so is a subspace spanned by $\vec{d} = [3, 1, 2]$. So we apply the previous definition to get the projection $\vec{p}$ of $\vec{b} = [1, 2, 1]$ on $\mathrm{sp}(\vec{d})$:

$$\vec{p} = \mathrm{proj}_{\vec{d}}(\vec{b}) = \frac{\vec{b}\cdot\vec{d}}{\vec{d}\cdot\vec{d}}\vec{d} = \frac{[1,2,1]\cdot[3,1,2]}{[3,1,2]\cdot[3,1,2]}[3,1,2]$$

$$= \frac{(1)(3)+(2)(1)+(1)(2)}{3^2+1^2+2^2}[3,1,2] = \frac{7}{14}[3,1,2] = \boxed{[3/2,\ 1/2,\ 1].}$$

□

Page 336 number 10

Page 336 number 10. Find the orthogonal complement of the plane $2x + y + 3z = 0$ in $\mathbb{R}^3$.

Solution. A plane is a translation of a two-dimensional space of the form $\vec{x} = t_1\vec{d}_1 + t_2\vec{d}_2 + \vec{a}$ where $\vec{d}_1$ and $\vec{d}_2$ form a basis for the two-dimensional space and $\vec{a}$ is a translation vector (see Section 2.5, "Lines, Planes, and Other Flats"). Here, we can take $\vec{a} = \vec{0}$ so that the plane is not translated and is in fact a subspace of $\mathbb{R}^3$. So we just need a basis for the subspace. We pick two linearly independent vectors in the subspace, say $\vec{d}_1 = [1, -2, 0]$ and $\vec{d}_2 = [0, -3, 1]$ (though there are infinitely many such choices). Then using the technique described above, we take $A = \begin{bmatrix} 1 & -2 & 0 \\ 0 & -3 & 1 \end{bmatrix}$ and find the nullspace of A by considering the system of equations $A\vec{x} = \vec{0}$ (see Note 6.1.A):

Page 336 number 10 (continued)

Solution (continued).

$$[A \mid \vec{0}] = \left[\begin{array}{ccc|c} 1 & -2 & 0 & 0 \\ 0 & -3 & 1 & 0 \end{array}\right] \underset{\sim}{\overset{R_1 \to R_1 - (2/3)R_2}{}} \left[\begin{array}{ccc|c} 1 & 0 & -2/3 & 0 \\ 0 & -3 & 1 & 0 \end{array}\right]$$

$$\underset{\sim}{\overset{R_2 \to R_2/(-3)}{}} \left[\begin{array}{ccc|c} 1 & 0 & -2/3 & 0 \\ 0 & 1 & -1/3 & 0 \end{array}\right].$$

So we have

$$\begin{array}{rcrcl} x_1 & & - & (2/3)x_3 & = & 0 \\ & x_2 & - & (1/3)x_3 & = & 0 \end{array} \quad \text{or} \quad \begin{array}{rcl} x_1 & = & (2/3)x_3 \\ x_2 & = & (1/3)x_3 \\ x_3 & = & x_3 \end{array}$$

or with $x_3 = 3t$ as a free variable, $x_1 = 2t$, $x_2 = t$, and $x_3 = 3t$. So $W^\perp$ is the nullspace of A: $\boxed{W^\perp = \mathrm{sp}([2, 1, 3]).}$ □

Theorem 6.1

Theorem 6.1. Properties of $W^\perp$.
The orthogonal complement $W^\perp$ of a subspace W of $\mathbb{R}^n$ has the following properties:

1. $W^\perp$ is a subspace of $\mathbb{R}^n$.
2. $\dim(W^\perp) = n - \dim(W)$.
3. $(W^\perp)^\perp = W$.
4. Each vector $\vec{b} \in \mathbb{R}^n$ can be expressed uniquely in the form $\vec{b} = \vec{b}_W + \vec{b}_{W^\perp}$ for $\vec{b}_W \in W$ and $\vec{b}_{W^\perp} \in W^\perp$.

Proof. Let $\dim(W) = k$, and let $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ be a basis for W. Let A be the $k \times n$ matrix having $\vec{v}_i$ as its ith row vector for $i = 1, 2, \dots, k$.

Property (1) follows from the fact that $W^\perp$ is the nullspace of matrix A, by Note 6.1.A, and therefore is a subspace of $\mathbb{R}^n$.

Theorem 6.1 (continued 1)

Proof (continued). For Property 2, consider the rank equation of A:

$$\text{rank}(A) + \text{nullity}(A) = n.$$

Since $\dim(W) = \text{rank}(A)$ and since $W^\perp$ is the nullspace of A, then $\dim(W^\perp) = n - \dim(W)$.

For Property 3, we have by Property 1 that $W^\perp$ is a subspace of $\mathbb{R}^n$. By Property 2 we have

$$\dim(W^\perp)^\perp = n - \dim(W^\perp) = n - (n - k) = k.$$

Since very vector in W is orthogonal to subspace $W^\perp$ then W is a subspace of $(W^\perp)^\perp$ ($(W^\perp)^\perp$ is a subspace of $\mathbb{R}^n$ by two applications of Property 1). Since W and $(W^\perp)^\perp$ have the same dimension then by Exercise 2.1.38, W must be equal to $(W^\perp)^\perp$.

Theorem 6.1 (continued 2)

Proof (continued). For Property 4, let $\{\vec{v}_{k+1}, \vec{v}_{k+2}, \dots, \vec{v}_n\}$ be a basis for $n - k$ dimensional (by Property 2) subspace $W^\perp$. We now show that

$$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} \cup \{\vec{v}_{k+1}, \vec{v}_{k+2}, \dots, \vec{v}_n\} = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$$

is a basis for $\mathbb{R}^n$. Consider the linear combination

$$r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_k\vec{v}_k + s_{k+1}\vec{v}_{k+1} + s_{k+2}\vec{v}_{k+2} + \cdots + s_n\vec{v}_n = \vec{0}. \quad (*)$$

This equation implies

$$r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_k\vec{v}_k = -s_{k+1}\vec{v}_{k+1} - s_{k+2}\vec{v}_{k+2} - \cdots - s_n\vec{v}_n.$$

Theorem 6.1 (continued 3)

Proof (continued). Notice that the vector on the left hand side of this equation is in W and the vector on the right hand side is in $W^\perp$. But both sides of the equation represent the same vector (d'uh, it's an equation!) so both sides of the equation represent a vector in both W and $W^\perp$. So this vector must be orthogonal to itself. The only vector orthogonal to itself is $\vec{0}$ (since $0 = \vec{v} \cdot \vec{v} = \|\vec{v}\|^2$ implies $\vec{v} = \vec{0}$). Since the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ are linearly independent and $r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_k\vec{v}_k = \vec{0}$ then we must have $r_1 = r_2 = \cdots r_k = 0$. Similarly, $\vec{v}_{k+1}, \vec{v}_{k+2}, \dots, \vec{v}_n$ are linearly independent and $s_{k+1}\vec{v}_{k+1} + s_{k+2}\vec{v}_{k+2} + \cdots + s_n\vec{v}_n = \vec{0}$ implies $s_{k+1} = s_{k+2} = \cdots = s_n = 0$. From equation $(*)$, we see that $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is a linearly independent set. Since the set contains n linearly independent vectors in $\mathbb{R}^n$ then $\dim(\text{sp}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)) = n$ and so by Exercise 2.1.38, $\text{sp}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n) = \mathbb{R}^n$ and so $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$.

Theorem 6.1 (continued 4)

Proof (continued). So each $\vec{b} \in \mathbb{R}^n$ can be written as $\vec{b} = (r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_k\vec{v}_k) + (s_{k+1}\vec{v}_{k+1} + s_{k+2}\vec{v}_{k+2} + \cdots + s_n\vec{v}_n)$, where $r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_k\vec{v}_k \in W$ and $s_{k+1}\vec{v}_{k+1} + s_{k+2}\vec{v}_{k+2} + \cdots + s_n\vec{v}_n \in W^\perp$, for unique $r_1, r_2, \ldots, r_k, s_{k+1}, s_{k+2}, \ldots, s_n$ (by Definition 1.17, "Basis for a Subspace"). So any $\vec{b} \in \mathbb{R}^n$ can be expressed in the form $\vec{b} = \vec{b}_W + \vec{b}_{W^\perp}$ where $\vec{b}_W \in W$ and $\vec{b}_{W^\perp} \in W^\perp$. Since each vector in $\mathbb{R}^n$ is a unique linear combination of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, then the choice of $\vec{b}_W$ and $\vec{b}_{W^\perp}$ are unique. □

Page 336 number 20(b)

Page 336 number 20(b). Find the projection of $\vec{b} = [-2, 1, 3, -5]$ on to the subspace $W = \text{sp}(\hat{e}_1, \hat{e}_4)$ in $\mathbb{R}^4$.

Solution. We are given a basis for $W = \text{sp}(\hat{e}_1, \hat{e}_4)$, namely $\{\hat{e}_1, \hat{e}_4\}$. Certainly a basis for $W^\perp$ is given by $\{\hat{e}_2, \hat{e}_3\}$. So we take the ordered basis $\{\hat{e}_1, \hat{e}_4, \hat{e}_2, \hat{e}_3\}$ of $\mathbb{R}^4$ and we have $\vec{b} = -2\hat{e}_1 - 5\hat{e}_4 + 1\hat{e}_2 + 3\hat{e}_3$ (and so the coordinate vector $\vec{r}$ of $\vec{b}$ relative to the ordered basis $\{\hat{e}_1, \hat{e}_4, \hat{e}_2, \hat{e}_3\}$ is $\vec{r} = [-2, -5, 1, 3]$). Then by the Note 6.1.B, the projection of $\vec{b}$ on to W is

$$\vec{b}_W = \text{proj}_W(\vec{b}) = r_1\hat{e}_1 + r_2\hat{e}_4 = -2[1,0,0,0] - 5[0,0,0,1] = \boxed{[-2,0,0,-5].}$$

□

Page 335 Example 6

Page 335 Example 6. Consider the inner product space $\mathcal{P}_{[0,1]}$ of all polynomial functions defined on the interval $[0, 1]$ with inner product

$$\langle p(x), q(x)\rangle = \int_0^1 p(x)q(x)\,dx.$$

Find the projection of $f(x) = x$ on sp(1) and then find the projection of x on $\text{sp}(1)^\perp$.

Solution. We follow the definition of the projection $\vec{p}$ of $\vec{b}$ on $\text{sp}(\vec{a})$ in $\mathbb{R}^n$, $\vec{p} = \text{proj}_{\vec{a}}(\vec{b}) = \dfrac{\vec{b}\cdot\vec{a}}{\vec{a}\cdot\vec{a}}\vec{a}$, but instead of dot products in $\mathbb{R}^n$ we use the inner product in $\mathcal{P}_{[0,1]}$. So the desired projection, with $\vec{b} = x$ and $\vec{a} = 1$, is

$$\frac{\langle x, 1\rangle}{\langle 1, 1\rangle}1 = \frac{\int_0^1 x\cdot 1\,dx}{\int_0^1 1\cdot 1\,dx}1 = \frac{(1/2)x^2|_0^1}{x|_0^1}1 = \boxed{\tfrac{1}{2}.}$$

Page 335 Example 6 (continued)

Page 335 Example 6. Consider the inner product space $\mathcal{P}_{[0,1]}$ of all polynomial functions defined on the interval $[0, 1]$ with inner product

$$\langle p(x), q(x)\rangle = \int_0^1 p(x)q(x)\,dx.$$

Find the projection of $f(x) = x$ on sp(1) and then find the projection of x on $\text{sp}(1)^\perp$.

Solution (continued). Notice that with $W = \text{sp}(1)$ then we have from Definition 6.2 that $\vec{b} = \vec{b}_W + \vec{b}_{W^\perp}$ and we can find $\vec{b}_{W^\perp}$ (where $W^\perp = \text{sp}(1)^\perp$) as $\vec{b}_{W^\perp} = \vec{b} - \vec{b}_W = x - 1/2$. □

Page 337 number 26

Page 337 number 26. Let A be an $m \times n$ matrix.

(a) Prove that the set W of row vectors $\vec{x}$ in $\mathbb{R}^m$ such that $\vec{x}A = \vec{0}$ is a subspace of $\mathbb{R}^m$.

(b) Prove that the subspace W in part (a) and the column space of A are orthogonal complements in $\mathbb{R}^m$.

Proof. (a) We use definition 1.16, "Subspace of $\mathbb{R}^n$." Let $W = \{\vec{x} \in \mathbb{R}^m \mid \vec{x}A = \vec{0}\}$. We must check W for closure under vector addition and scalar multiplication. Let $\vec{x}_1, \vec{x}_2 \in W$ and let r be a scalar. Then:

$$\begin{aligned}(\vec{x}_1 + \vec{x}_2)A &= \vec{x}_1 A + \vec{x}_2 A \text{ by Theorem 1.3.A(10),} \\ &\quad \text{“Distribution Laws of Matrix Multiplication”} \\ &\quad \text{(here we treat } \vec{x} \text{ as a matrix)} \\ &= \vec{0} + \vec{0} \text{ since } \vec{x}_1, \vec{v}_2 \in W \\ &= \vec{0},\end{aligned}$$

Page 337 number 26 (continued 1)

Proof (continued). ... and

$$\begin{aligned}(r\vec{x}_1)A &= r(\vec{x}_1 A) \text{ by Theorem 1.3.A(7), “Scalars Pull Through”} \\ &= r\vec{0} \text{ since } \vec{x}_1 \in W \\ &= \vec{0}.\end{aligned}$$

So both $\vec{x}_1 + \vec{x}_2 \in W$ and $r\vec{x}_1 \in W$. That is, W is closed under vector addition and scalar multiplication. By Definition 1.16, W is a subspace of $\mathbb{R}^m$. □

(b) Let A be an $m \times n$ matrix. Prove that the subspace W in part (a) and the column space of A are orthogonal complements in $\mathbb{R}^m$.

Proof. Recall that by Definition 1.8, "Matrix Product," the (i,j) entry of the matrix product AB is the dot product of the ith row of A with the jth column of B.

Page 337 number 26 (continued 2)

Proof (continued). So for $\vec{x} \in W$ (here we treat row vector $\vec{x} \in \mathbb{R}^m$ as a $1 \times m$ matrix) we have that $\vec{x}A$ is a $1 \times n$ matrix (or a row vector in $\mathbb{R}^n$) and for $\vec{x} \in W$ we have $\vec{x}A = \vec{0} \in \mathbb{R}^n$. So the jth entry of $\vec{x}A = \vec{0}$ is the dot product of $\vec{x}$ with the jth column of A and, since $\vec{x}A = \vec{0}$, this dot product must be 0 for each $j = 1, 2, \ldots, n$. So by Definition 1.7, "Perpendicular or Orthogonal Vectors," each $\vec{x} \in W$ is orthogonal to each column of A. Also, by definition, W contain <u>all</u> vectors $\vec{x}$ in $\mathbb{R}^m$ which satisfy $\vec{x}A = \vec{0}$ (i.e., all vectors $\vec{x}$ in $\mathbb{R}^m$ which are perpendicular to all columns of A). The column space of A is the span of the columns of A and since $\vec{x} \in W$ is orthogonal to each column of A then $\vec{x}$ is orthogonal to each vector which is in the span of the columns of A. Conversely, any vector $\vec{x}$ in the orthogonal complement of the column space of A must be orthogonal to all linear combinations of the columns of A; in particular such $\vec{x}$ must by orthogonal to each column of A and hence such $\vec{x}$ is in W. So the orthogonal complement of the column space of A is W. □

Page 337 number 28

Page 337 number 28. Let W be a subspace of $\mathbb{R}^n$ with orthogonal complement $W^\perp$. Writing $\vec{a} = \vec{a}_W + \vec{a}_{W^\perp}$, as in Theorem 6.1, prove that $\|\vec{a}\| = \sqrt{\|\vec{a}_W\|^2 + \|\vec{a}_{W^\perp}\|^2}$.

Solution. By Note 1.2.A, $\|\vec{a}\|^2 = \vec{a} \cdot \vec{a}$, so we have

$$\begin{aligned}\|\vec{a}\|^2 &= (\vec{a}_W + \vec{a}_{W^\perp}) \cdot (\vec{a}_W + \vec{a}_{W^\perp}) \\ &= \vec{a}_W \cdot \vec{a}_W + \vec{a}_W \cdot \vec{a}_{W^\perp} + \vec{a}_{W^\perp} \cdot \vec{a}_W + \vec{a}_{W^\perp} \cdot \vec{a}_{W^\perp} \\ &\quad \text{by Theorem 1.3, “Properties of Dot Products”} \\ &= \|\vec{a}_W\|^2 + \vec{a}_W \cdot \vec{a}_{W^\perp} + \vec{a}_{W^\perp} \cdot \vec{a}_W + \|\vec{a}_{W^\perp}\|^2 \text{ by Note 1.2.A} \\ &= \|\vec{a}_W\|^2 + 0 + 0 + \|\vec{a}_{W^\perp}\|^2 \text{ since } \vec{a}_W \text{ and } \vec{a}_{W^\perp} \text{ are orthogonal.}\end{aligned}$$

Not taking square roots (and observing that $\|\vec{a}\|$ is nonnegative) gives $\|\vec{a}\| = \sqrt{\|\vec{a}_W\|^2 + \|\vec{a}_{W^\perp}\|^2}$. □