Orthogonal Complements and Projections

Recall that two vectors v_1 & v_2 in $\mathbb{R}^n$ are *perpendicular* or *orthogonal* provided that their *dot product* vanishes. That is, $v_1 \perp v_2$ if and only if $v_1 \cdot v_2 = 0$.

Example

1. The vectors $\begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix}$ & $\begin{pmatrix} 12 \\ 8 \\ 3 \end{pmatrix}$ in $\mathbb{R}^3$ are orthogonal while $\begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix}$ & $\begin{pmatrix} 4 \\ -6 \\ 7 \end{pmatrix}$ are not.

2. We can define an *inner product* on the vector space of all polynomials of degree at most 3 by setting

$$\langle f(x), g(x) \rangle = \int_0^1 f(x)\ g(x)\ dx.$$

(There is nothing special about integrating over [0,1]; This interval was chosen arbitrarily.) Then, for example,

$$\begin{aligned} \langle 2x^2 + 1, 10x^2 + 11x - 11 \rangle &= \int_0^1 \left(2x^2 + 1\right)\left(10x^2 + 11x - 11\right) dx \\ &= \int_0^1 \left(20x^4 + 22x^3 - 12x^2 + 11x - 11\right) dx \\ &= \left(4x^5 + \frac{11}{2}x^4 - 4x^3 + \frac{11}{2}x^2 - 11x \right) \Bigg|_0^1 \\ &= 0 \end{aligned}$$

Hence, relative to the inner product $\langle f(x), g(x) \rangle = \int_0^1 f(x)\ g(x)\ dx$ we have that the two polynomials $2x^2 + 1$ & $10x^2 + 11x - 11$ are *orthogonal* in $\boldsymbol{P_3}$.

So, more generally, we say that $v_1 \perp v_2$ in a vector space $\boldsymbol{V}$ with inner product $\langle u, v \rangle$ provided that $\langle u, v \rangle = 0$.

Example

Consider the 3×4 matrix $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 3 & 1 \\ 2 & 3 & 1 & 1 \\ 1 & 0 & 8 & 2 \end{pmatrix}$. Then, by the elementary row operations,

we have that $rref(\boldsymbol{A}) = \begin{pmatrix} 1 & 0 & 8 & 2 \\ 0 & 1 & -5 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. As discussed in the previous sections, the *row space* of $\boldsymbol{A}$ coincides with the *row space* of $rref(\boldsymbol{A})$. In this case, we see that a basis for $\boldsymbol{R}_A = \boldsymbol{R}_{rref(\boldsymbol{A})}$ is given by $\left\{ \begin{pmatrix} 1 \\ 0 \\ 8 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -5 \\ -1 \end{pmatrix} \right\}$. By consideration of $rref(\boldsymbol{A})$, it follows that the *null space* of $\boldsymbol{A}$, N_A, has a basis given by $\left\{ \begin{pmatrix} -8 \\ 5 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$. We note that, as per the *Fundamental Theorem of Linear Algebra*, that

$dim(R_A) + dim(N_A) = 4 \; (= \# \; of \; columns \; of \; A \;).$

Let's consider vectors in $\boldsymbol{R}_A = \boldsymbol{R}_{rref(\boldsymbol{A})}$ and N_A, say,

$$6 \begin{pmatrix} 1 \\ 0 \\ 8 \\ 2 \end{pmatrix} + (-2) \begin{pmatrix} 0 \\ 1 \\ -5 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ -2 \\ 58 \\ 14 \end{pmatrix} \in R_A$$

and

$$(-3)\begin{pmatrix}-8\\5\\1\\0\end{pmatrix} + 4\begin{pmatrix}-2\\1\\0\\1\end{pmatrix} = \begin{pmatrix}16\\-11\\-3\\4\end{pmatrix} \in N_A.$$

By direct computation we see that

$$\begin{pmatrix}6\\-2\\58\\14\end{pmatrix} \cdot \begin{pmatrix}16\\-11\\-3\\4\end{pmatrix} = 6\ (16) + (-2)\ (-11) + 58\ (-3) + 14\ (4) = 0$$

and so $\begin{pmatrix}6\\-2\\58\\14\end{pmatrix} \perp \begin{pmatrix}16\\-11\\-3\\4\end{pmatrix}$.

So, is this an accident that an element of $\boldsymbol{R_A} = \boldsymbol{R_{rref(A)}}$ is orthogonal to an element of N_A?

To answer this let's consider the dot product of arbitrary elements of $\boldsymbol{R_A} = \boldsymbol{R_{rref(A)}}$ and N_A.

Since $\left\{\begin{pmatrix}1\\0\\8\\2\end{pmatrix}, \begin{pmatrix}0\\1\\-5\\-1\end{pmatrix}\right\}$ is a basis for $\boldsymbol{R_A} = \boldsymbol{R_{rref(A)}}$, there exists scalars a & b so that every

vector in $\boldsymbol{R_A} = \boldsymbol{R_{rref(A)}}$ can be written as

$$a\begin{pmatrix}1\\0\\8\\2\end{pmatrix} + b\begin{pmatrix}0\\1\\-5\\-1\end{pmatrix} = \begin{pmatrix}a\\b\\8a-5b\\2a-b\end{pmatrix}.$$

Similarly, since $\left\{\begin{pmatrix}-8\\5\\1\\0\end{pmatrix}, \begin{pmatrix}-2\\1\\0\\1\end{pmatrix}\right\}$ is a basis for N_A, there exists scalars r & s so that every

vector in N_A can be written as

$$r\begin{pmatrix}-8\\5\\1\\0\end{pmatrix} + s\begin{pmatrix}-2\\1\\0\\1\end{pmatrix} = \begin{pmatrix}-8r-2s\\5r+s\\r\\s\end{pmatrix}.$$

Now,

$$\begin{pmatrix}a\\b\\8a-5b\\2a-b\end{pmatrix} \cdot \begin{pmatrix}-8r-2s\\5r+s\\r\\s\end{pmatrix} = a(-8r-2) + b(5r+s) + (8a-5b)r + (2a-b)s = 0.$$

We conclude that if $\boldsymbol{v} \in R_A$ and $\boldsymbol{w} \in N_A$, then $\boldsymbol{v} \cdot \boldsymbol{w} = 0$ and so $\boldsymbol{v} \perp \boldsymbol{w}$.

Definition

Suppose $\boldsymbol{V}$ is a vector space with inner product $\langle u, v \rangle$. (Think $V = \mathbb{R}^n$ and $\langle u, v \rangle = dot(u, v)$)

1. The subspaces $\boldsymbol{S_1}$ & $\boldsymbol{S_2}$ of $\mathbb{R}^n$ are said to be ***orthogonal***, denoted $\boldsymbol{S_1} \perp \boldsymbol{S_2}$, if $\langle v_1, v_2 \rangle = 0$ for all $v_1 \in \boldsymbol{S_1}$ & $v_2 \in \boldsymbol{S_2}$.

2. Let $\boldsymbol{W}$ be a subspace of $\boldsymbol{V}$. Then we define $\boldsymbol{W}^{\perp}$ (read "*W perp*") to be the set of vectors in $\boldsymbol{V}$ given by

$$\boldsymbol{W}^{\perp} = \{ v \in \boldsymbol{V} \mid \langle v, w \rangle = 0 \ for\ all\ w \in \boldsymbol{W} \}.$$

The set $\boldsymbol{W}^{\perp}$ is called the ***orthogonal complement*** of $\boldsymbol{W}$.

Examples

1. From the above work, if $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 3 & 1 \\ 2 & 3 & 1 & 1 \\ 1 & 0 & 8 & 2 \end{pmatrix}$, then $R_A \perp N_A$.

2. Let $\boldsymbol{A}$ be any $\boldsymbol{m} \times \boldsymbol{n}$ matrix. Now, the null space N_A of $\boldsymbol{A}$ consists of those vectors $\boldsymbol{x}$ with $\boldsymbol{A}\,\boldsymbol{x} = 0_m$. However, $\boldsymbol{A}\,\boldsymbol{x} = 0_m$ if and only if $\boldsymbol{r_i} \cdot \boldsymbol{x} = 0 \ \ (\boldsymbol{i} = 1, \ldots, \boldsymbol{m})$ for each row $\boldsymbol{r_i}$ of the matrix $\boldsymbol{A}$. Hence, the null space of $\boldsymbol{A}$ is the set of all vectors orthogonal to the rows of $\boldsymbol{A}$ and, hence, the row space of $\boldsymbol{A}$. (Why?) We conclude that $R_A^{\perp} = N_A$.

The above suggest the following method for finding $\boldsymbol{W}^{\perp}$ given a subspace $\boldsymbol{W}$ of $\mathbb{R}^n$.

1. Find a matrix $\boldsymbol{A}$ having as row vectors a generating set for $\boldsymbol{W}$.
2. Find the null space of $\boldsymbol{A}$. This null space is $\boldsymbol{W}^{\perp}$.

3. Suppose that $\boldsymbol{S_1} = span \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$ and $\boldsymbol{S_2} = span \left\{ \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}$. Then $\boldsymbol{S_1}$ & $\boldsymbol{S_2}$

are *orthogonal* subspaces of $\mathbb{R}^5$. To verify this observe that

$$\left(a \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right) \cdot \left(r \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right) = a\ r \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + b\ r \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

$$= a\ r\ (0) + b\ r\ (0)$$

$$= 0$$

Thus, $\boldsymbol{S_1} \perp \boldsymbol{S_2}$. Since

$$\left(a \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right) \cdot \begin{pmatrix} -7 \\ -2 \\ 2 \\ 2 \\ 5 \end{pmatrix} = 0$$

and

$$\begin{pmatrix} -7 \\ -2 \\ 2 \\ 2 \\ 5 \end{pmatrix} \notin \boldsymbol{S_2},$$

it follows that $\boldsymbol{S}_1^{\perp} \neq \boldsymbol{S}_2$. So, what is the set $\boldsymbol{S}_1^{\perp}$? Let $\boldsymbol{B} = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}$.

Then, from part 2 above, $\boldsymbol{S}_1^{\perp} = N_B$. In fact, a basis for $\boldsymbol{S}_1^{\perp} = N_B$ can be shown to be

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Finally, we note that the set $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\} \cup \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ forms a basis

for $\mathbb{R}^5$. In particular, every element of $\mathbb{R}^5$ can be written as the sum of a vector in $\boldsymbol{S}_1$ and a vector in $\boldsymbol{S}_1^{\perp}$.

4. Let $\boldsymbol{W}$ be the subspace of $\boldsymbol{P}_3$ (= the vector space of all polynomials of degree at most 3) with basis $\{ 1, x^3 \}$. We take as our inner product on $\boldsymbol{P}_3$ the function

$$\langle f(x), g(x) \rangle = \int_0^1 f(x)\; g(x)\; dx.$$

Find as basis for $\boldsymbol{W}^{\perp}$.

Solution

Let $p(x) = ax^3 + bx^2 + cx + d \in \boldsymbol{W}^{\perp}$. Then

$$\langle p(x), g(x) \rangle = \int_0^1 p(x)\; g(x)\; dx = 0$$

for all $g(x) \in \boldsymbol{W}$. Hence, in particular,

$$\langle p(x), 1 \rangle = \int_0^1 \left(ax^3 + bx^2 + cx + d\right) dx = \frac{a}{4} + \frac{b}{3} + \frac{c}{2} + d = 0$$

and

$$\langle p(x), x^3 \rangle = \int_0^1 \left(ax^6 + bx^5 + cx^4 + dx^3\right) dx = \frac{a}{7} + \frac{b}{6} + \frac{c}{5} + \frac{d}{4} = 0.$$

Solving the linear system

$$\frac{a}{4} + \frac{b}{3} + \frac{c}{2} + d = 0$$

$$\frac{a}{7} + \frac{b}{6} + \frac{c}{5} + \frac{d}{4} = 0$$

we find that we have pivot variables of $a = \frac{14}{5}c + 14d$ and $b = -\frac{18}{5}c - \frac{27}{2}d$ with free variables of c and d. It follows that

$$p(x) = c\left(\frac{14}{5}x^3 - \frac{18}{5}x^2 + x\right) + d\left(14x^3 - \frac{27}{2}x^2 + 1\right)$$

for some $c, d \in \mathbb{R}$. Hence, the polynomials

$$\frac{14}{5}x^3 - \frac{18}{5}x^2 + x \quad \& \quad 14x^3 - \frac{27}{2}x^2 + 1$$

span $\boldsymbol{W}^{\perp}$. Since these two polynomials are not multiples of each other, they are linearly independent and so they form a basis for $\boldsymbol{W}^{\perp}$.

Theorem

Suppose that $\boldsymbol{W}$ is a subspace of $\mathbb{R}^n$.

1. $\boldsymbol{W}^{\perp}$ is a subspace of $\mathbb{R}^n$.

2. $dim\left(W^{\perp}\right) = n - dim(\boldsymbol{W})$

3. $\left(W^{\perp}\right)^{\perp} = \boldsymbol{W}$.

4. Each vector in $\boldsymbol{b} \in \mathbb{R}^n$ can be expressed uniquely in the form $\boldsymbol{b} = b_W + b_{W^{\perp}}$ where $b_W \in \boldsymbol{W}$ and $b_{W^{\perp}} \in W^{\perp}$.

Definition

Let $\boldsymbol{V}$ and $\boldsymbol{W}$ be two subspaces of $\mathbb{R}^n$. If each vector $\boldsymbol{x} \in \mathbb{R}^n$ can be expressed uniquely in the form $\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{w}$ where $\boldsymbol{v} \in \boldsymbol{V}$ and $\boldsymbol{w} \in \boldsymbol{W}$, the we say $\mathbb{R}^n$ is the ***direct sum*** of $\boldsymbol{V}$ and $\boldsymbol{W}$ and we write $\mathbb{R}^n = \boldsymbol{V} \oplus \boldsymbol{V}$.

Example

1. $$span\left\{\begin{pmatrix}0\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\0\end{pmatrix}\right\} \oplus span\left\{\begin{pmatrix}0\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\0\end{pmatrix}\right\}^{\perp}$$

$$= span\left\{\begin{pmatrix}0\\1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\0\end{pmatrix}\right\} \oplus span\left\{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\-1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}\right\} = \mathbb{R}^5$$

2. $\quad span \left\{ x^3, 1 \right\} \oplus span \left\{ x^3, 1 \right\}^{\perp}$

$$= span \left\{ x^3, 1 \right\} \oplus span \left\{ \frac{14}{5}x^3 - \frac{18}{5}x^2 + x, \ 14x^3 - \frac{27}{2}x^2 + 1 \right\} = \boldsymbol{P_3}$$

Fundamental Subspaces of a Matrix

Let $\boldsymbol{A}$ be an $m \times n$ matrix. Then the ***four fundamental subspaces*** of $\boldsymbol{A}$ are

R_A = row space of $\boldsymbol{A}$ $\left(= R_{rref(A)} \right)$

N_A = null space of $\boldsymbol{A}$

C_A = column space of $\boldsymbol{A}$ $\left(= R_{A^T} \right)$

N_{A^T} = null space of A^T

Example

Let $\boldsymbol{A} = \begin{pmatrix} 1 & 5 & 3 & 7 \\ 2 & 0 & -4 & -6 \\ 4 & 7 & -1 & 2 \end{pmatrix}$. Since $rref(\boldsymbol{A}) = \begin{pmatrix} 1 & 0 & -2 & -3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, it follows that

R_A has a basis of $\left\{ (1, 0, -2, -3), (0, 1, 1, 2) \right\}$

and that

N_A has a basis of $\left\{ (3, -2, 0, 1), (2, -1, 1, 0) \right\}$.

Because $rref\left(A^T\right) = \begin{pmatrix} 1 & 0 & \frac{7}{5} \\ 0 & 1 & \frac{13}{10} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, we have that

C_A has a basis of $\left\{ \begin{pmatrix} 1 \\ 0 \\ \frac{7}{5} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \frac{13}{10} \end{pmatrix} \right\}$

and that

N_{A^T} consists of all scalar multiples of the vector $\begin{pmatrix} -\frac{7}{5} \\ -\frac{13}{10} \\ 1 \end{pmatrix}$.

<u>Fundamental Theorem of Linear Algebra</u> - Part II

Let $\boldsymbol{A}$ be an $\boldsymbol{m} \times \boldsymbol{n}$ matrix.

1. N_A is the orthogonal complement of R_A in $\mathbb{R}^n$.

2. N_{A^T} is the orthogonal complement of C_A in $\mathbb{R}^m$.

3. $N_A \oplus R_A = \mathbb{R}^n$

4. $N_{A^T} \oplus C_A = \mathbb{R}^m$

<u>Example</u>

Let $\boldsymbol{A} = \begin{pmatrix} 1 & 5 & 3 & 7 \\ 2 & 0 & -4 & -6 \\ 4 & 7 & -1 & 2 \end{pmatrix}$. Write $\boldsymbol{v} = (12, -10, -14, -26)$ uniquely as the sum of a vector in R_A and a vector in N_A. It is sufficient to $a, b, c, d \in \mathbb{R}$ so that

$$\left(a\,(1, 0, -2, -3) + b\,(0, 1, 1, 2)\right) + \left(c\,(3, -2, 0, 1) + d\,(2, -1, 1, 0)\right) = \boldsymbol{v}.$$

Reducing the associated augmented matrix

$$\begin{pmatrix} 1 & 0 & 3 & 2 & | & 12 \\ 0 & 1 & -2 & -1 & | & -10 \\ -2 & 1 & 0 & 1 & | & -14 \\ -3 & 2 & 1 & 0 & | & -26 \end{pmatrix}$$

to

$$\begin{pmatrix} 1 & 0 & 0 & 0 & | & 5 \\ 0 & 1 & 0 & 0 & | & -6 \\ 0 & 0 & 1 & 0 & | & 1 \\ 0 & 0 & 0 & 1 & | & 2 \end{pmatrix}$$

we see that $a = 5$, $b = -6$, $c = 1$ and $d = 2$.

Set

$$\boldsymbol{r} = 5\,(1, 0, -2, -3) - 6\,(0, 1, 1, 2) = (5, -6, -16, -27) \in R_A$$

and

$$\boldsymbol{s} = (3, -2, 0, 1) + 2\,(2, -1, 1, 0) = (7, -4, 2, 1) \in N_A.$$

Then $\boldsymbol{v} = \boldsymbol{r} + \boldsymbol{s}$. Why is this the only way to represent $\boldsymbol{v} = (12, -10, -14, -26)$ as a sum of a vector from R_A and a vector from N_A?

Definition

Let $\boldsymbol{b} \in \mathbb{R}^n$ and let $\boldsymbol{W}$ be a subspace of $\mathbb{R}^n$. If $\boldsymbol{b} = b_W + b_{W^\perp}$ where $b_W \in \boldsymbol{W}$ and $b_{W^\perp} \in W^\perp$, then we call $\boldsymbol{b}_W$ the ***projection*** of $\boldsymbol{b}$ onto $\boldsymbol{W}$ and write $\boldsymbol{b}_W = proj_W\,\boldsymbol{b}$.

Example

1. Suppose $\boldsymbol{b} = (12, -10, -14, -26) \in \mathbb{R}^4$ and $\boldsymbol{W}$ is the subspace of $\mathbb{R}^4$ with basis vectors $\{(1, 0, -2, -3), (0, 1, 1, 2)\}$. Then, by the previous example, $proj_W\,\boldsymbol{b} = (5, -6, -16, -27)$.

2. Find $proj_W\, \boldsymbol{b} \in \mathbb{R}^3$ if $\boldsymbol{b} = (2, 1, 5)$ and $W = span\{(1, 2, 1), (2, 1, -1)\}$.

Solution

We note that $(1, 2, 1)$ & $(2, 1, -1)$ are linearly independent and, hence, form a basis for $\boldsymbol{W}$.

So, we find a basis for $\boldsymbol{W}^{\perp}$ by finding the null space for

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

or, equivalently,

$$rref\left(\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & -1 \end{pmatrix} \right) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

We see that $\boldsymbol{W}^{\perp} = \langle (1, -1, 1) \rangle$. We now seek $x_1, x_2, x_3 \in \mathbb{R}$ so that

$$(*) \qquad (2, 1, 5) = \left(x_1 (1, 2, 1) + x_2 (2, 1, -1) \right) + x_3 (1, -1, 1).$$

(Of course, $proj_W\, \boldsymbol{b} = x_1 (1, 2, 1) + x_2 (2, 1, -1)$.)

To solve the equation (*) it is sufficient to row reduce the augmented matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 2 & 1 & -1 & 1 \\ 1 & -1 & 1 & 5 \end{array} \right)$$

obtaining

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{array} \right).$$

Thus, $proj_W\, \boldsymbol{b} = (2)\,(1, 2, 1) + (-1)\,(2, 1, -1) = (0, 3, 3)$. We observe that there exists a matrix $\boldsymbol{P}$ given by

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

so that

$$\boldsymbol{P}\,\boldsymbol{b} = proj_W\,\boldsymbol{b}.$$

We call ***P*** the ***projection matrix***. The projection matrix given by $P = A^T\left(A\,A^T\right)^{-1}A$ (where the rows of ***A*** form a basis for ***W***) is expensive computationally but if one is computing several projections onto ***W*** it may very well be worth the effort as the above formula is valid for all vectors ***b***.

3. Find the projection of $(1, 2, 1)$ onto the plane $x + y + z = 0$ in $\mathbb{R}^3$ via the projection matrix.

Solution

We seek a set of basis vectors for the plane $x + y + z = 0$. We claim the two vectors $(1, 0, -1)$ and $(2, -1, -1)$ form a basis. (Any two vectors solving $x + y + z = 0$ that are not multiples of one another will work.) Set

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

Then

$$A\,A^T = \begin{pmatrix} 2 & 3 \\ 3 & 6 \end{pmatrix}$$

and

$$\left(A\,A^T\right)^{-1} = \begin{pmatrix} 2 & 3 \\ 3 & 6 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & \frac{2}{3} \end{pmatrix}.$$

Hence,

$$P = A^T \left(A\,A^T\right)^{-1} A$$

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

$$= \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

So,

$$\boldsymbol{P\,b} = proj_W\, \boldsymbol{b} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = -\frac{1}{3} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

We note that

$$P\left(P\,b\right) = P\left(proj_W\, b\right) = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} -1/3 \\ 2/3 \\ -1/3 \end{pmatrix} = \begin{pmatrix} -1/3 \\ 2/3 \\ -1/3 \end{pmatrix}$$

and, hence, $P^2 = P$.

Why is this not surprising?

Properties of a Projection Matrix:

(1) $P^2 = P$ (That is, ***P*** is *idempotent*.)

(2) $P^T = P$ (That is, ***P*** is *symmetric*.)