MA1508E: LINEAR ALGEBRA FOR ENGINEERING

Lecture 9 Notes

References

1. Elementary Linear Algebra: Application Version, Section 4.8
2. Linear Algebra with Application, Section 5.4, 5.6, 8.1

4.2 Rank

Let $\mathbf{A}$ be a $m \times n$ matrix. Recall that the number of pivot columns in the reduced row-echelon form of $\mathbf{A}$ tells us the dimension of the column space of $\mathbf{A}$, and the number of nonzero rows in the reduced row-echelon form of $\mathbf{A}$ tells us the dimension of the row space of $\mathbf{A}$,

$$\begin{aligned} \#\text{ of pivot columns in RREF of } \mathbf{A} &= \dim(\mathrm{Col}(\mathbf{A})), \\ \#\text{ of nonzero rows in RREF of } \mathbf{A} &= \dim(\mathrm{Row}(\mathbf{A})). \end{aligned}$$

Recall also that the number of pivot columns in the reduced row-echelon form of $\mathbf{A}$ is equals to the number of leading entries, which is equals to the number of nonzero rows,

$$\begin{aligned} \#\text{ of leading entries} &= \#\text{ of pivot columns in RREF of } \mathbf{A} \\ &= \#\text{ of nonzero rows in RREF of } \mathbf{A}. \end{aligned}$$

Hence, the dimension of the column space of $\mathbf{A}$ is equals to the dimension of the row space of $\mathbf{A}$.

Theorem. *For any matrix $\mathbf{A}$, the dimension of the row space of $\mathbf{A}$ is equal to the dimension of the column space of $\mathbf{A}$,*

$$\dim(\mathrm{Row}(\mathbf{A})) = \dim(\mathrm{Col}(\mathbf{A})).$$

Define the <u>rank</u> of a matrix $\mathbf{A}$, denoted as $\mathrm{rank}(\mathbf{A})$, to be the dimension of its column or (and) its row space,

$$\mathrm{rank}(\mathbf{A}) = \dim(\mathrm{Col}(\mathbf{A})) = \dim(\mathrm{Row}(\mathbf{A})).$$

Example.

1. $\mathrm{rank}(\mathbf{A}) = 0$ if and only if $\mathbf{A} = \mathbf{0}$ the zero matrix.
2. $\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$. So $\mathrm{rank}(\mathbf{A}) = 3$.
3. $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 2 \\ -1 & 7 & 5 \\ 1 & 9 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. So $\mathrm{rank}(\mathbf{A}) = 3$.

Now since the number of pivot columns in the reduced row-echelon form is at most the number of columns, the rank of a matrix is at most the number columns,

$$\text{rank} = \# \text{ of pivot columns in RREF} \leq \# \text{ of columns}.$$

Also, the number of nonzero rows in the reduced row-echelon form is at most the number rows, the rank of a matrix is at most the number of rows,

$$\text{rank} = \# \text{ nonzero rows in RREF} \leq \# \text{ of rows}.$$

Hence, we have the following lemma.

Lemma. *For a $m \times n$ matrix $\mathbf{A}$,*

$$\text{rank}(\mathbf{A}) \leq \min\{m, n\}.$$

A $m \times n$ matrix $\mathbf{A}$ is said to be of full rank if equality is attained in the above inequality,

$$\text{rank}(\mathbf{A}) = \min\{m, n\}.$$

Example.

1. $\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$ is of full rank.

2. $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 2 \\ -1 & 7 & 5 \\ 1 & 9 & 2 \end{pmatrix}$ is of full rank.

3. $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. $\mathbf{A}$ is not full rank.

4. $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. $\mathbf{A}$ is not of full rank.

Before you read on, try to prove or disprove the following statements.

1. If the rank of $\mathbf{A}$ is equal to the number of columns, then $\mathbf{A}$ has full rank.

2. If the rank of $\mathbf{A}$ is equal to the number of rows, then $\mathbf{A}$ has full rank.

3. If the reduced row-echelon form of $\mathbf{A}$ has a zero row, then $\mathbf{A}$ cannot be of full rank.

4. If the reduced row-echelon form of $\mathbf{A}$ has a non-pivot column, then $\mathbf{A}$ cannot be of full rank.

5. If the reduced row-echelon form of $\mathbf{A}$ has at least one non-pivot column and one nonzero row, then $\mathbf{A}$ cannot be of full rank.

6. If the reduced row-echelon form of $\mathbf{A}$ has no non-pivot columns, then $\mathbf{A}$ has full rank.

7. If the reduced row-echelon form of $\mathbf{A}$ has no zero rows, then $\mathbf{A}$ has full rank.

4.3 Column Space and Consistency of Linear System

Let $\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix}$ be a $n \times k$ matrix, where $\mathbf{u}_i \in \mathbb{R}^n$ is the i-th column of $\mathbf{A}$, for $i = 1, ..., k$. Recall (lecture 6) that $\mathbf{b} \in \text{span}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\} = \text{Col}(\mathbf{A})$ if and only if $\left(\begin{array}{cccc|c} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k & \mathbf{v} \end{array} \right)$ is consistent, which is equivalent to $\mathbf{Ax} = \mathbf{v}$ is consistent.

Theorem. *Let* $\mathbf{A}$ *be a* $m \times n$ *matrix. Then* $\text{Col}(\mathbf{A}) = \left\{ \mathbf{Au} \;\middle|\; \mathbf{u} \in \mathbb{R}^n \right\}$.

Proof. Suppose $\mathbf{v} \in \text{Col}(\mathbf{A})$. Then $\mathbf{Ax} = \mathbf{v}$ is consistent, and thus there is a $\mathbf{u} \in \mathbb{R}^n$ such that $\mathbf{Au} = \mathbf{v}$. So $\mathbf{v} \in \left\{ \mathbf{Au} \;\middle|\; \mathbf{u} \in \mathbb{R}^n \right\}$. This shows that $\text{Col}(\mathbf{A}) \subseteq \left\{ \mathbf{Au} \;\middle|\; \mathbf{u} \in \mathbb{R}^n \right\}$.

Conversely, suppose $\mathbf{v} \in \left\{ \mathbf{Au} \;\middle|\; \mathbf{u} \in \mathbb{R}^n \right\}$, that is, $\mathbf{v} = \mathbf{Au}$ for some $\mathbf{u} \in \mathbb{R}^n$. Then $\mathbf{Ax} = \mathbf{b}$ is consistent, and thus $\mathbf{b} \in \text{Col}(\mathbf{A})$. This shows that $\left\{ \mathbf{Au} \;\middle|\; \mathbf{u} \in \mathbb{R}^n \right\} \subseteq \text{Col}(\mathbf{A})$. Therefore, we have equality. $\square$

4.4 Nullspace

Recall that the solution space to a homogeneous system is a subspace. We shall call this space the nullspace of the matrix. The nullspace of a $m \times n$ matrix $\mathbf{A}$ is the solution space to the homogeneous system $\mathbf{Ax} = \mathbf{0}$ with coefficient matrix $\mathbf{A}$. It is denoted as

$$\text{Null}(\mathbf{A}) = \left\{ \mathbf{v} \in \mathbb{R}^n \;\middle|\; \mathbf{Av} = \mathbf{0} \right\}.$$

The nullity of $\mathbf{A}$ is the dimension of the nullspace of $\mathbf{A}$, denoted as

$$\text{nullity}(\mathbf{A}) = \dim(\text{Null}(\mathbf{A})).$$

Recall that the dimension of the solution space of the homogeneous system $\mathbf{Ax} = \mathbf{0}$ is equals to the number of non-pivot columns in the reduced row-echelon form of $\mathbf{A}$. Since the rank of $\mathbf{A}$ is equal to the number of pivot columns in the reduced row-echelon form of $\mathbf{A}$, we have the rank-nullity theorem.

Theorem (Rank-Nullity Theorem)**.** *Let* $\mathbf{A}$ *be a* $m \times n$ *matrix. Then the sum of the rank and nullity of* $\mathbf{A}$ *is equal to the number of columns of* $\mathbf{A}$*,*

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

Example. Let $\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 & -1 \\ 3 & 6 & 5 & 0 \\ 1 & 2 & 1 & 2 \end{pmatrix}$.

$$\begin{pmatrix} 1 & 2 & 2 & -1 \\ 3 & 6 & 5 & 0 \\ 1 & 2 & 1 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 0 & 5 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Hence,

$$\text{Null}(\mathbf{A}) = \text{span}\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -5 \\ 0 \\ 3 \\ 1 \end{pmatrix} \right\}.$$

Also, $\text{rank}(\mathbf{A}) + \text{nullity}(\mathbf{A}) = 2 + 2 = 4 = \#$ of columns of $\mathbf{A}$.

4.5 Full Rank Matrices

Let $\mathbf{A}$ be a $m \times n$ matrix. Suppose $m \geq n$. Then $\mathbf{A}$ has full rank (equals the number of columns) if and only if the reduced row-echelon form $\mathbf{R}$ of $\mathbf{A}$ has the form

$$\mathbf{R} = \begin{pmatrix} \mathbf{I}_n \\ \mathbf{0}_{(m-n)\times n} \end{pmatrix}.$$

Theorem (Rank equals to number of columns)**.** *Let* $\mathbf{A}$ *be a* $m \times n$ *matrix. The following statements are equivalent.*

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

(ii) The rows of $\mathbf{A}$ *spans* $\mathbb{R}^n$, $\text{Row}(\mathbf{A}) = \mathbb{R}^n$.

(iii) The columns of $\mathbf{A}$ *are linearly independent.*

(iv) The homogeneous system $\mathbf{Ax} = \mathbf{0}$ *has only the trivial solution, that is,* $\text{Null}(\mathbf{A}) = \{\mathbf{0}\}$.

(v) $\mathbf{A}^T\mathbf{A}$ *is an invertible matrix of order* n.

(vi) $\mathbf{A}$ *has a left inverse.*

Now suppose $n \geq m$. Then $\mathbf{A}$ has full rank (equals the number of rows) if and only if the reduced row-echelon form $\mathbf{R}$ of $\mathbf{A}$ has the form

$$\begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 & \cdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & \cdots & \cdots & 0 & \cdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & \cdots & \cdots & 0 & \cdots \\ \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \cdots & \cdots & \vdots & \cdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & \cdots & 1 & \cdots \end{pmatrix}$$

Theorem (Rank equals to number of rows)**.** *Let* $\mathbf{A}$ *be a* $m \times n$ *matrix. The following statements are equivalent.*

(i) $\text{rank}(A) = m$.

(ii) The columns of $\mathbf{A}$ *spans* $\mathbb{R}^m$, $\text{Col}(\mathbf{A}) = \mathbb{R}^m$.

(iii) The rows of $\mathbf{A}$ *are linearly independent.*

(iv) $\mathbf{A}\mathbf{A}^T$ *is an invertible matrix of order* m.

(v) $\mathbf{A}$ *has a right inverse.*

(vi) The linear system $\mathbf{Ax} = \mathbf{b}$ *is consistent for every* $\mathbf{b} \in \mathbb{R}^m$.

Example. 1. $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Then $\text{rank}(\mathbf{A}) = 3$ which is the number of columns of $\mathbf{A}$. We have already seen that therefore $\mathbf{A}$ must fulfill the first 4 statements of the theorem above. Now we will show that $\mathbf{A}^T\mathbf{A}$ is invertible, and thus find a left inverse of $\mathbf{A}$.

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix}, \ (\mathbf{A}^T\mathbf{A})^{-1} = \frac{1}{7}\begin{pmatrix} 5 & -2 & -2 \\ -2 & 5 & -2 \\ -2 & -2 & 5 \end{pmatrix}.$$

Then a left inverse of $\mathbf{A}$ is

$$(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \frac{1}{7}\begin{pmatrix} 3 & 3 & -4 & 1 \\ -4 & 3 & 3 & 1 \\ 3 & -4 & 3 & 1 \end{pmatrix}.$$

Consider now the homogeneous system $\mathbf{Ax} = \mathbf{0}$. Suppose $\mathbf{u} \in \mathbb{R}^3$ is a solution, $\mathbf{Au} = \mathbf{0}$. Then

$$\mathbf{u} = ((\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)\mathbf{Au} = ((\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)\mathbf{0} = \mathbf{0}.$$

So, the system has only the trivial solution.

2. Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$. Then $\text{rank}(\mathbf{A}) = 3$ which is the number of rows. We have already seen that $\mathbf{A}$ fulfills the first 3 statements of the theorem, as well as the last statement. We will now show that $\mathbf{AA}^T$ is invertible, and find a right inverse of $\mathbf{A}$.

$$\mathbf{AA}^T = \begin{pmatrix} 7 & 3 & 1 \\ 3 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix}, \ (\mathbf{AA}^T)^{-1} = \frac{1}{12}\begin{pmatrix} 5 & -8 & 1 \\ -8 & 20 & -4 \\ 1 & -4 & 5 \end{pmatrix}.$$

So, a right inverse of $\mathbf{A}$ is

$$\mathbf{A}^T(\mathbf{AA}^T)^{-1} = \frac{1}{6}\begin{pmatrix} 3 & -6 & 3 \\ -1 & 4 & 1 \\ 1 & 2 & -1 \\ -2 & 2 & 2 \end{pmatrix}.$$

Now given any $\mathbf{b} \in \mathbb{R}^3$, $\mathbf{u} = (\mathbf{A}^T(\mathbf{AA}^T)^{-1})\mathbf{b}$ is a solution to the system $\mathbf{Ax} = \mathbf{b}$. Indeed,

$$\mathbf{Au} = \mathbf{A}(\mathbf{A}^T(\mathbf{AA}^T)^{-1})\mathbf{b} = \mathbf{b}.$$

So, the system $\mathbf{Ax} = \mathbf{b}$ is consistent.

Remark. 1. The two theorem above also show that a non-square matrix can have at most a left or a right inverse, but not both.

2. If a matrix has more rows than columns, then the matrix cannot have a right inverse. For if $\mathbf{A}$ has a right inverse, then rank of $\mathbf{A}$ must be equal to the number of rows, which cannot happen since the rank is bounded by the number of columns, which is strictly smaller than the number of rows.

3. If a matrix has more columns than rows, then the matrix cannot have a left inverse. For if $\mathbf{A}$ has a left inverse, then rank of $\mathbf{A}$ must be equal to the number of columns, which cannot happen since the rank is bounded by the number of rows, which is strictly smaller than the number of columns.

5 Orthogonal Projection

5.1 Orthogonal and Orthonormal basis

Recall that a set $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is orthogonal if

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0$$

for all $i \neq j$, and it is orthonormal if

$$\mathbf{u}_i \cdot \mathbf{u}_j = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases}$$

Suppose now S is an orthogonal set of nonzero vectors. Then all the vectors are perpendicular to each other, and thus geometrically it is clear that they must be linearly independent. In fact, intuitively, each time we add a nonzero vector that is orthogonal to a orthogonal set, the space that contains all the vectors increases its dimension by 1.

Theorem. *Suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthogonal set of nonzero vectors. Then S is linearly independent.*

Proof. Suppose there are some coefficients $c_1, c_2, ..., c_k \in \mathbb{R}$ such that

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k = \mathbf{0}.$$

Then for any $i = 1, ..., k$,

$$\begin{aligned} 0 &= \mathbf{u}_i \cdot \mathbf{0} = \mathbf{u}_i \cdot (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k) \\ &= c_1(\mathbf{u}_i \cdot \mathbf{u}_1) + c_2(\mathbf{u}_i \cdot \mathbf{u}_2) + \cdots + c_k(\mathbf{u}_i \cdot \mathbf{u}_k) \\ &= c_i(\mathbf{u}_i \cdot \mathbf{u}_i) \end{aligned}$$

since $\mathbf{u}_i \cdot \mathbf{u}_j = 0$. But since $\mathbf{u}_i \neq \mathbf{0}$, $\mathbf{u}_i \cdot \mathbf{u}_i \neq 0$, and hence, necessarily $c_i = 0$. Therefore the only combination of vectors in S to give the zero vector is the trivial one. This shows that S is linearly independent. □

Corollary. *Every orthonormal set is linearly independent.*

Proof. Follows from the fact that an orthonormal set is orthogonal, and since the vectors are unit vectors, they are nonzero. □

Let $V \subseteq \mathbb{R}^n$ be a subspace. A set $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthogonal basis for V if it is a basis for V and is an orthogonal set. Similarly, S is an orthonormal basis for V if it is a basis for V and is an orthonormal set.

Corollary.
1. *Any orthogonal set $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ containing n nonzero vectors in $\mathbb{R}^n$ is an orthogonal basis.*
2. *Any orthonormal set $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ containing n vectors in $\mathbb{R}^n$ is an orthonormal basis.*

Proof. Follows from the fact that a linearly independent set containing n vectors must be a basis for $\mathbb{R}^n$. □

Corollary. *Suppose $V \subseteq \mathbb{R}^n$ is a k-dimensional subspace.*

1. *Any orthogonal set* $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ *containing* k *nonzero vectors in* V *is an orthogonal basis for* V.

2. *Any orthonormal set* $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ *containing* k *vectors in* V *is an orthonormal basis for* V.

Proof. Exercise. $\square$

The coordinates of a vector relative to an orthogonal or orthonormal basis is rather easy to compute.

Theorem. *Let* $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ *be an orthogonal basis for a subspace* $V \subseteq \mathbb{R}^n$. *Then for any* $\mathbf{v} \in V$,

$$\mathbf{v} = \left(\frac{\mathbf{v}\cdot\mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right)\mathbf{u}_1 + \left(\frac{\mathbf{v}\cdot\mathbf{u}_2}{\|\mathbf{u}_2\|^2}\right)\mathbf{u}_2 + \cdots + \left(\frac{\mathbf{v}\cdot\mathbf{u}_k}{\|\mathbf{u}_k\|^2}\right)\mathbf{u}_k,$$

that is,

$$(\mathbf{v})_S = \begin{pmatrix} (\mathbf{v}\cdot\mathbf{u}_1)/\|\mathbf{u}_1\|^2 \\ (\mathbf{v}\cdot\mathbf{u}_2)/\|\mathbf{u}_2\|^2 \\ \vdots \\ (\mathbf{v}\cdot\mathbf{u}_k)/\|\mathbf{u}_k\|^2 \end{pmatrix}.$$

If further S *is an orthonormal basis, then*

$$\mathbf{v} = (\mathbf{v}\cdot\mathbf{u}_1)\,\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\,\mathbf{u}_2 + \cdots + (\mathbf{v}\cdot\mathbf{u}_k)\,\mathbf{u}_k,$$

that is,

$$(\mathbf{v})_S = \begin{pmatrix} \mathbf{v}\cdot\mathbf{u}_1 \\ \mathbf{v}\cdot\mathbf{u}_2 \\ \vdots \\ \mathbf{v}\cdot\mathbf{u}_k \end{pmatrix}.$$

Proof. The second statement follows immediately from the first. Suppose S is orthogonal. For a $\mathbf{v} \in V$, since S is a basis, we can write $\mathbf{v}$ as a linear combination the vectors in S

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k$$

for some $c_1, c_2, ..., c_k \in \mathbb{R}$. Then for each $i = 1, ..., k$,

$$\begin{aligned} \mathbf{u}_i\cdot\mathbf{v} &= \mathbf{u}_i\cdot(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k) \\ &= c_1(\mathbf{u}_i\cdot\mathbf{u}_1) + c_2(\mathbf{u}_i\cdot\mathbf{u}_2) + \cdots + c_k(\mathbf{u}_i\cdot\mathbf{u}_k) \\ &= c_i\|\mathbf{u}_i\|^2. \end{aligned}$$

Hence, $c_i = \frac{\mathbf{v}\cdot\mathbf{u}_1}{\|\mathbf{u}_i\|^2}$ for all $i = 1, ..., k$. $\square$

Remark. Note that this only works if $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthogonal or orthonormal basis. For example, let $S = \left\{\mathbf{u}_1 = \begin{pmatrix}1\\1\\0\end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix}1\\0\\0\end{pmatrix}\right\}$, and $\mathbf{w} = \begin{pmatrix}2\\1\\0\end{pmatrix}$. Then

$$\left(\frac{\mathbf{w}\cdot\mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right)\mathbf{u}_1 + \left(\frac{\mathbf{w}\cdot\mathbf{u}_2}{\|\mathbf{u}_2\|^2}\right)\mathbf{u}_2 = \frac{3}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} + 2\begin{pmatrix}1\\0\\0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}7\\3\\0\end{pmatrix} \neq \mathbf{w}.$$

Why is this so?

Example. 1. The standard basis $E = \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ is an orthonormal basis. Given any $\mathbf{v} = (v_i) \in \mathbb{R}^n$,

$$\mathbf{e}_i \cdot \mathbf{v} = v_i.$$

Thus, $(\mathbf{v})_E = (\mathbf{e}_i \cdot \mathbf{v}) = (v_i) = \mathbf{v}$.

2. $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ is an orthogonal basis for V, the xy-plane in $\mathbb{R}^3$. For any $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \in V$,

$$\begin{aligned} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} &= \left(\frac{\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}{\left\| \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\|^2} \right) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \left(\frac{\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}}{\left\| \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\|^2} \right) \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \\ &= \frac{(x+y)}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{(x-y)}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}. \end{aligned}$$

That is,

$$(\mathbf{v})_S = \frac{1}{2} \begin{pmatrix} x+y \\ x-y \end{pmatrix}.$$

3. Let $S = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}$ and $V = \text{span}(S)$. S is an orthonormal basis for V. Let $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \in V$. Then

$$\begin{aligned} \mathbf{v} &= \left(\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \right) \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \right) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ &= (\sqrt{3}) \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + (\sqrt{2}) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \end{aligned}$$

That is,

$$(\mathbf{v})_S = \begin{pmatrix} \sqrt{3} \\ \sqrt{2} \end{pmatrix}.$$

Exercise: Suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for a subspace V. Let

$$T = \left\{ \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \right\}$$

be the orthonormal basis for V obtained from normalizing S. Suppose $\mathbf{v} \in V$ is such that

$$(\mathbf{v})_S = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

What is $(\mathbf{v})_T$?

5.2 Orthogonal Projection

Now suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthonormal basis for a subspace $V \subseteq \mathbb{R}^n$. Let $\mathbf{w} \in \mathbb{R}^n$ be a vector not in V, $\mathbf{w} \notin V$. How do we interpret the vector

$$\mathbf{w}' = (\mathbf{w} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{w} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{w} \cdot \mathbf{u}_k)\mathbf{u}_k?$$

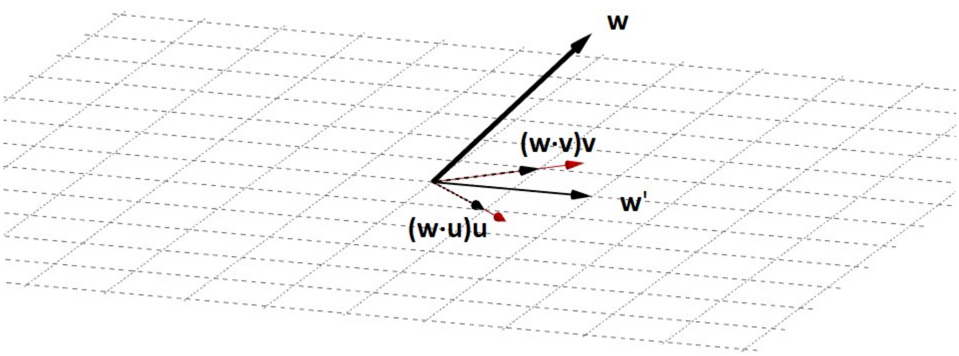
Example. Let $\left\{ \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ be an orthonormal basis for the xy-plane in $\mathbb{R}^3$. Let $\mathbf{w} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Then

$$\mathbf{w}' = (\mathbf{w} \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{w} \cdot \mathbf{e}_2)\mathbf{e}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Consider now another orthonormal basis $\left\{ \mathbf{v} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{u} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ for the xy-plane. Then again

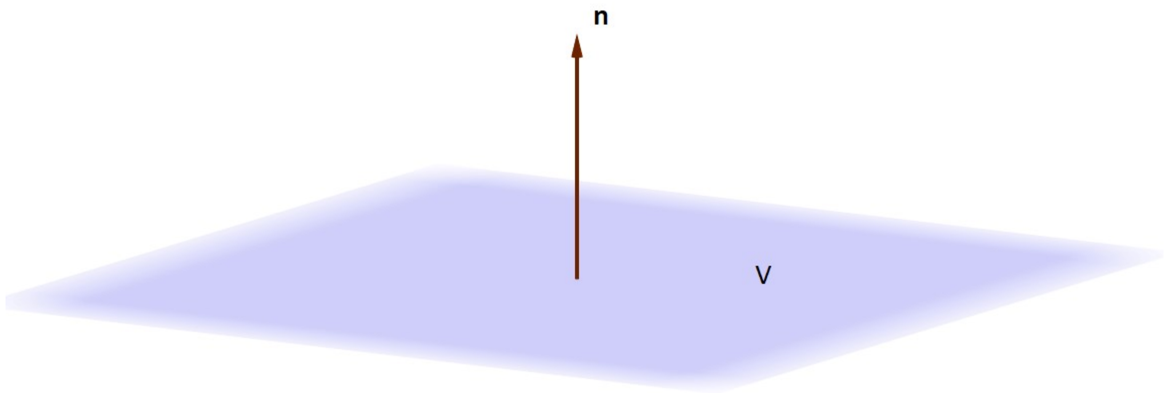
$$\mathbf{w}' = (\mathbf{w} \cdot \mathbf{v})\mathbf{v} + (\mathbf{w} \cdot \mathbf{u})\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

In both (any) orthonormal basis for the xy-plane, $\mathbf{w}'$ is the same. Geometrically, it is the projection of $\mathbf{w}$ onto the xy-plane.



To show this rigorously, we need to define what is the projection of a vector onto a subspace.

Let $V \subseteq \mathbb{R}^n$ be a subspace. A vector $\mathbf{n} \in \mathbb{R}^n$ is orthogonal to V if for every $\mathbf{v} \in V$, $\mathbf{n} \cdot \mathbf{v} = 0$, that is, $\mathbf{n}$ is orthogonal to every vector in V. We will denote it as $\mathbf{n} \perp V$.



If $\mathbf{n} \neq \mathbf{0}$.

Example. 1. Let $\mathbf{w} = \mathbf{0}$. Then for any subspace $V \subseteq \mathbb{R}^n$, $\mathbf{w} \perp V$.

2. Let $V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \; ax + by + cz = 0 \right\}$. Any vector of the form $s \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, $s \in \mathbb{R}$, is orthogonal to V, since

$$s \begin{pmatrix} a \\ b \\ c \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = s(ax + by + cz) = 0,$$

for any $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V$. This shows that a plane in $\mathbb{R}^3$ has the expression

$$V = \{ \ \mathbf{v} \in \mathbb{R}^3 \mid \mathbf{v} \cdot \mathbf{n} = 0 \ \}$$

for some $\mathbf{n} \in \mathbb{R}^3$, $\mathbf{n} \neq \mathbf{0}$. $\mathbf{n}$ is said to be normal to the plane V.

Theorem. *Let $V \subseteq \mathbb{R}^n$ be a subspace and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a spanning set for V, span$(S) = V$. Then $\mathbf{w} \in \mathbb{R}^n$ is orthogonal to V if and only if $\mathbf{w} \cdot \mathbf{u}_i = 0$ for all $i = 1, ..., k$.*

Geometrically, this is clear. If $\mathbf{w}$ is the zero vector, then the statement is obvious. Otherwise, since $\mathbf{w}$ is perpendicular to every vector in the spanning set, then it must be pointing in the direction that is perpendicular to the entire space spanned by the spanning set. Refer to the appendix for the detailed proof.

This gives us a means to find vectors that are orthogonal to a subspace $V \subseteq \mathbb{R}^n$.

Theorem. *Let $V \subseteq \mathbb{R}^n$ be a subspace and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a basis for V. Then $\mathbf{w} \perp V$ if and only if $\mathbf{w} \in \text{Null}(\mathbf{A}^T)$, where $\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix}$.*

Note that since we write vectors as column vectors, there is a need to take the transpose of $\mathbf{A}$. If we write the vectors in the basis S as row vectors and $\mathbf{A}$ is the matrix whose i-th rows is $\mathbf{u}_i$, then there is no need to take the transpose in the theorem above.

Proof. By the previous theorem, $\mathbf{w} \perp V \Leftrightarrow \mathbf{u}_i^T\mathbf{w} = \mathbf{u}_i \cdot \mathbf{w} = 0$ for all $i = 1, ..., k \Leftrightarrow$ $\mathbf{A}^T\mathbf{w} = \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_k^T \end{pmatrix} \mathbf{w} = \mathbf{0}$. □

Example. Let $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ and $V = \text{span}(S)$. Then $\mathbf{w} \perp V$ if and only if $\begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & -1 & 0 \end{pmatrix} \mathbf{w} = 0$.

$$\begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & -1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 2 & 2 \\ 0 & 1 & -1 & 0 \end{pmatrix} \Rightarrow \mathbf{w} \perp V \Leftrightarrow \mathbf{w} \in \text{span} \left\{ \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

We are now ready to define the orthogonal projection of a vector onto a subspace.

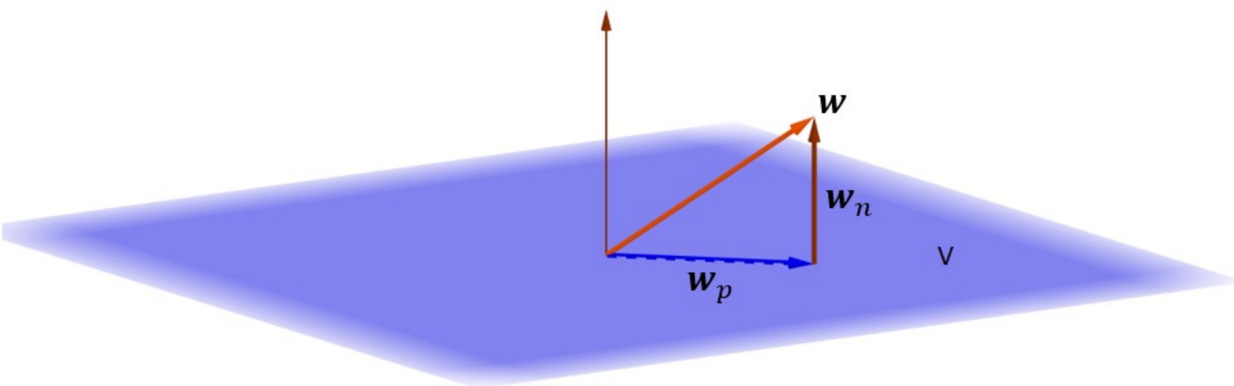
Theorem. *Let $V \subseteq \mathbb{R}^n$ be a subspace. Every vector $\mathbf{w} \in \mathbb{R}^n$ can be decomposed uniquely as*

$$\mathbf{w} = \mathbf{w}_p + \mathbf{w}_n,$$

where $\mathbf{w}_n$ is orthogonal to V and $\mathbf{w}_p \in V$.

The unique vector $\mathbf{w}_p$ in V is called the orthogonal projection of $\mathbf{w}$ onto V. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be an orthonormal basis for V. Then

$$\mathbf{w}_p = (\mathbf{w} \cdot \mathbf{u}_1)\mathbf{u} + (\mathbf{w} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{w} \cdot \mathbf{u}_k)\mathbf{u}_k.$$



Example. $S = \left\{ \mathbf{v} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{u} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ is an orthonormal basis for the xy-plane in $\mathbb{R}^3$. Let $\mathbf{w} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Then

$$\mathbf{w}_p = (\mathbf{w} \cdot \mathbf{v})\mathbf{v} + (\mathbf{w} \cdot \mathbf{u})\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \text{ and } \mathbf{w}_n = \mathbf{w} - \mathbf{w}_p = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Appendix to Lecture 9

Lemma. *For any matrix* $\mathbf{A}$*,*

$$\text{Null}(\mathbf{A}^T\mathbf{A}) = \text{Null}(\mathbf{A}).$$

Proof. ($\supseteq$) Suppose $\mathbf{u} \in \text{Null}(\mathbf{A})$. Then $\mathbf{Au} = \mathbf{0}$. Pre-multiplying both sides of the equation by $\mathbf{A}^T$, we get $\mathbf{A}^T\mathbf{Au} = \mathbf{0}$. Thus $\mathbf{u} \in \text{Null}(\mathbf{A}^T\mathbf{A})$.

($\subseteq$) Suppose $\mathbf{u} \in \text{Null}(\mathbf{A}^T\mathbf{A})$. Then $\mathbf{A}^T\mathbf{Au} = \mathbf{0}$. Pre-multiplying both sides of the equation by $\mathbf{u}^T$, we have

$$0 = \mathbf{u}^T\mathbf{A}^T\mathbf{Au} = (\mathbf{Au})^T(\mathbf{Au}) = \|\mathbf{Au}\|^2.$$

By a property of norm, this tells us the that vector $\mathbf{Au}$ must be the zero vector, $\mathbf{Au} = \mathbf{0}$. Hence, $\mathbf{u} \in \text{Null}(\mathbf{A})$.

$\square$

Corollary. *For any matrix* $\mathbf{A}$*,*

$$\text{rank}(\quad) = \text{rank}(\quad^T\quad).$$

Proof. Let $\mathbf{A}$ be a $m \times n$ matrix. From the previous lemma, $\text{nullity}(\mathbf{A}^T\mathbf{A}) = \text{nullity}(\mathbf{A})$. Note that $\mathbf{A}^T\mathbf{A}$ is a $n \times n$ matrix. Then by the rank-nullity theorem,

$$\text{rank}(\mathbf{A}) + \text{nullity}(\mathbf{A}) = n = \text{rank}(\mathbf{A}^T\mathbf{A}) + \text{nullity}(\mathbf{A}^T\mathbf{A}) = \text{rank}(\mathbf{A}^T\mathbf{A}) + \text{nullity}(\mathbf{A}),$$

and so $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T\mathbf{A})$. $\square$

Theorem (Rank Equals to Number of Columns)**.** *Let* $\mathbf{A}$ *be a* $m\times n$ *matrix. The following statements are equivalent.*

(i) $\text{rank}(A) = n$*.*

(ii) The rows of $\mathbf{A}$ *spans* $\mathbb{R}^n$*,* $\text{Row}(\mathbf{A}) = \mathbb{R}^n$*.*

(iii) The columns of $\mathbf{A}$ *are linearly independent.*

(iv) The homogeneous system $\mathbf{Ax} = \mathbf{0}$ *has only the trivial solution, that is,* $\text{Null}(\mathbf{A}) = \{\mathbf{0}\}$*.*

(v) $\mathbf{A}^T\mathbf{A}$ *is an invertible matrix of order* n*.*

(vi) $\mathbf{A}$ *has a left inverse.*

Proof. We have already shown the equivalence of the first 4 statements.

Using the result $\text{Null}(\mathbf{A}) = \text{Null}(\mathbf{A}^T\mathbf{A})$, and the fact that a square matrix is invertible if and only if its nullspace is trivial, we have $\mathbf{A}^T\mathbf{A}$ is invertible if and only if $\text{Null}(\mathbf{A}) = \{\mathbf{0}\}$.

Suppose $\mathbf{A}^T\mathbf{A}$ is invertible. Then

$$\mathbf{I} = (\mathbf{A}^T\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{A}) = ((\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)\mathbf{A}$$

shows that $((\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)$ is a left inverse of $\mathbf{A}$. So (v) $\Rightarrow$ (vi).

Conversely, suppose $\mathbf{A}$ has a left inverse $\mathbf{C}$. Then if $\mathbf{u} \in \mathbb{R}^n$ is such that $\mathbf{Au} = \mathbf{0}$, we have

$$\mathbf{u} = \mathbf{CAu} = \mathbf{C0} = \mathbf{0}.$$

So the nullspace of $\mathbf{A}$ is trivial, which is equivalent to $\mathbf{A}^T\mathbf{A}$ being invertible. Hence, we have shown (vi) $\Rightarrow$ (v). □

Theorem (Rank Equals to Number of Rows)**.** *Let* $\mathbf{A}$ *be a* $m \times n$ *matrix. The following statements are equivalent.*

(i) $\text{rank}(A) = m$.

(ii) The columns of $\mathbf{A}$ *spans* $\mathbb{R}^m$*,* $\text{Col}(\mathbf{A}) = \mathbb{R}^m$*.*

(iii) The rows of $\mathbf{A}$ *are linearly independent.*

(iv) $\mathbf{AA}^T$ *is an invertible matrix of order* m*.*

(v) $\mathbf{A}$ *has a right inverse.*

(vi) The linear system $\mathbf{Ax} = \mathbf{b}$ *is consistent for every* $\mathbf{b} \in \mathbb{R}^m$*.*

Proof. By replacing $\mathbf{A}$ with $\mathbf{A}^T$ in the previous theorem, we have the equivalence of statements (i) to (v).

Suppose $\mathbf{A}$ has a right inverse $\mathbf{B}$, that is, $\mathbf{AB} = \mathbf{I}$. Then given any $\mathbf{b} \in \mathbb{R}^m$, let $\mathbf{u} = \mathbf{Bb}$, and we have

$$\mathbf{Au} = \mathbf{ABb} = \mathbf{Ib} = \mathbf{b}.$$

So $\mathbf{u}$ is a solution to $\mathbf{Ax} = \mathbf{b}$, and hence the system is consistent. This shows (v) $\Rightarrow$ (vi).

Conversely, suppose $\mathbf{Ax} = \mathbf{b}$ is consistent for every $b \in \mathbb{R}^m$. Let $\mathbf{b}_i$ be a solution to $\mathbf{Ax} = \mathbf{e}_i$, where $\mathbf{e}_i$ is the i-th vector in the standard basis, for $i = 1, ..., m$. Let $\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_m \end{pmatrix}$. Then

$$\mathbf{AB} = \begin{pmatrix} \mathbf{Ab}_1 & \mathbf{Ab}_2 & \cdots & \mathbf{Ab}_m \end{pmatrix} = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_m \end{pmatrix} = \mathbf{I}_m.$$

This shows that $\mathbf{A}$ has a right inverse. Thus we have shown (iv) $\Rightarrow$ (v). □

Theorem. *Let* $V \subseteq \mathbb{R}^n$ *be a subspace and* $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ *be a spanning set for* V*, span*$(S) = V$*. Then* $\mathbf{w} \in \mathbb{R}^n$ *is orthogonal to* V *if and only if* $\mathbf{w} \cdot \mathbf{u}_i = 0$ *for all* $i = 1, ..., k$*.*

Proof.

($\Rightarrow$) Suppose $\mathbf{w}$ is orthogonal to V. Since $S \subseteq V$, then $\mathbf{w}$ must also be orthogonal to every vector in S, that is, $\mathbf{w} \cdot \mathbf{u}_i = 0$ for all $i = 1, ..., k$.

($\Leftarrow$) Suppose $\mathbf{w}$ is orthogonal to all the vectors in the spanning set S. Given any $\mathbf{v} \in V$, $\mathbf{v}$ can be written as a linear combination of the vectors in S, that is, there are some $c_1, c_2, ..., c_k \in \mathbb{R}$ such that $\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k$. Then

$$\begin{aligned} \mathbf{w} \cdot \mathbf{v} &= \mathbf{w} \cdot (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k) \\ &= c_1(\mathbf{w} \cdot \mathbf{u}_1) + c_2(\mathbf{w} \cdot \mathbf{u}_2) + \cdots + c_k(\mathbf{w} \cdot \mathbf{u}_k) \\ &= c_1 0 + c_2 0 + \cdots + c_k 0 = 0. \end{aligned}$$

This shows that $\mathbf{w}$ is orthogonal to every $\mathbf{v} \in V$. Hence, $\mathbf{w} \perp V$. □

The theorem also shows that if V is a subspace, then the set of all vectors orthogonal to V is also a subspace, since it is the nullspace of some matrix.

Corollary. *Let $V \subseteq \mathbb{R}^n$ be a subspace. The set $\{ \mathbf{w} \in \mathbb{R}^n \mid \mathbf{w} \perp V \}$ is also a subspace of $\mathbb{R}^n$.*

The set in the above corollary is called the <u>orthogonal complement</u> of V, and is denoted as $V^\perp$ (pronounced as V perp, short for perpendicular),

$$V^\perp = \{ \mathbf{w} \in \mathbb{R}^n \mid \mathbf{w} \perp V \}.$$

Theorem. *Let $V \subseteq \mathbb{R}^n$ be a subspace. Every vector $\mathbf{w} \in \mathbb{R}^n$ can be decomposed uniquely as*

$$\mathbf{w} = \mathbf{w}_p + \mathbf{w}_n,$$

where $\mathbf{w}_n$ is orthogonal to V and $\mathbf{w}_p \in V$.

Proof. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be an orthonormal basis for V. Define

$$\mathbf{w}_p = (\mathbf{w} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{w} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{w} \cdot \mathbf{u}_k)\mathbf{u}_k.$$

Then $\mathbf{w}_n = \mathbf{w} - \mathbf{w}_p$. By construction, $\mathbf{w}_p \in V$. To show that $\mathbf{w}_n$ is orthogonal to V, suffice to show that $\mathbf{w}_n \cdot \mathbf{u}_i = 0$ for all $i = 1, ..., k$. Indeed,

$$\begin{aligned} \mathbf{u}_i \cdot \mathbf{w}_n &= \mathbf{u}_i \cdot (\mathbf{w} - \mathbf{w}_p) \\ &= \mathbf{u}_i \cdot \mathbf{w} - (\mathbf{w} \cdot \mathbf{u}_1)\mathbf{u}_i \cdot \mathbf{u}_1 - \cdots - (\mathbf{w} \cdot \mathbf{u}_i)\mathbf{u}_i \cdot \mathbf{u}_i - \cdots - (\mathbf{w} \cdot \mathbf{u}_k)\mathbf{u}_i \cdot \mathbf{u}_k \\ &= \mathbf{u}_i \cdot \mathbf{w} - \mathbf{u}_i \cdot \mathbf{w} = 0, \end{aligned}$$

since S is an orthonormal set. We will now show that $\mathbf{w}_p$ and $\mathbf{w}_n$ are unique. Suppose $\mathbf{w} = \mathbf{w}'_p + \mathbf{w}'_n$ is another decomposition with $\mathbf{w}'_p \in V$ and $\mathbf{w}'_n \perp V$. Then

$$\mathbf{w}'_p + \mathbf{w}'_n = \mathbf{w}_p + \mathbf{w}_n \Rightarrow \mathbf{w}'_p - \mathbf{w}_p = \mathbf{w}_n - \mathbf{w}'_n.$$

But since $\mathbf{w}'_p$ and $\mathbf{w}_p$ are in V, then so is $\mathbf{w}'_p - \mathbf{w}_p = \mathbf{w}_n - \mathbf{w}'_n$. But since $\mathbf{w}_n$ and $\mathbf{w}'_n$ are orthogonal to V, then so is their linear combination (since the set of all vectors orthogonal to V is a subspace), in particular, $(\mathbf{w}_n - \mathbf{w}'_n) \perp V$. In other words, $\mathbf{w}'_p - \mathbf{w}_p = \mathbf{w}_n - \mathbf{w}'_n$ is a vector that is both in V and orthogonal to V. This can only happen if and only if it is the zero vector. Hence,

$$\mathbf{w}'_p - \mathbf{w}_p = \mathbf{w}_n - \mathbf{w}'_n = \mathbf{0} \Rightarrow \mathbf{w}'_p = \mathbf{w}_p \text{ and } \mathbf{w}_n = \mathbf{w}'_n.$$

This concludes the proof that the decomposition is unique. $\square$