Matrix Theory — Lecture notes
MAT 335, Fall 2022 — D. Ivanšić

6.1 The Geometry of Vectors

Definition. For a vector $\mathbf{v}$ in $\mathbf{R}^n$, the *norm (length)* $|\mathbf{v}|$ of a vector is defined as

$$|\mathbf{v}| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

A *unit* vector is a vector whose norm is 1. The *distance* between to vectors is defined as $|\mathbf{u} - \mathbf{v}|$.

Example. Determine the condition on coordinates of vectors $\mathbf{u}$ and $\mathbf{v}$ that detects when $\mathbf{u}$ and $\mathbf{v}$ satisfy the Pythagorean-theorem-style equation $|\mathbf{u}|^2 + |\mathbf{v}|^2 = |\mathbf{u} - \mathbf{v}|^2$.

Definition. The *dot product* of vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{R}^n$ is the scalar

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

We will say that $\mathbf{u}$ and $\mathbf{v}$ are orthogonal (perpendicular) if $\mathbf{u} \cdot \mathbf{v} = 0$.

Note. Viewing vectors $\mathbf{u}$ and $\mathbf{v}$ as column matrices, we have $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$.
For an $m \times n$ matrix A and vectors $\mathbf{u} \in \mathbf{R}^n$ and $\mathbf{v} \in \mathbf{R}^m$, we have $A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A^T \mathbf{v}$.

Theorem 6.1. Properties of the dot product. For all vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in $\mathbf{R}^n$ and every scalar c, we have:

$\mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2$	$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$	$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
$\mathbf{u} \cdot \mathbf{u} = 0$ iff $\mathbf{u} = \mathbf{0}$	$(c\mathbf{a}) \cdot \mathbf{b} = c(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (c\mathbf{b})$	$(\mathbf{v} + \mathbf{w}) \cdot \mathbf{u} = \mathbf{v} \cdot \mathbf{u} + \mathbf{w} \cdot \mathbf{u}$
$\|c\mathbf{u}\| = \|c\|\|\mathbf{u}\|$	$\mathbf{a} \cdot \mathbf{0} = \mathbf{0} \cdot \mathbf{a} = 0$	

Proof. As the dot product can be viewed as matrix multiplication, most of the statements have been proven in that setting.

Example. Find a unit vector in the direction of $\mathbf{v} = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}$.

Example. Expand.

$|\mathbf{u} - \mathbf{v}|^2 =$

$|\mathbf{u} + \mathbf{v}|^2 =$

Theorem 6.2. Pythagorean Theorem in $\mathbf{R}^n$. Vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{R}^n$ are orthogonal if and only if

$$|\mathbf{u} + \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2$$

Proof.

Example. The orthogonal projection of a vector $\mathbf{u}$ onto vector $\mathbf{v}$ is given by $\dfrac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|^2}\mathbf{v}$.

Theorem 6.3. Cauchy-Schwarz Inequality. For all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{R}^n$, we have

$$|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}| \cdot |\mathbf{v}|$$

Proof.

Theorem 6.3. Triangle Inequality. For all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{R}^n$, we have

$$|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|$$

6.2 Orthogonal Vectors

Definition. A set of vectors in $\mathbf{R}^n$ is called *orthogonal* if any two distinct vectors from the set are orthogonal. A set of vectors in $\mathbf{R}^n$ is called *orthonormal* if the set is orthogonal and every vector in the set is a unit vector.

Example. Show that the set is orthogonal.

$$\left\{ \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\}$$

Theorem 6.5. An orthogonal set of nonzero vectors is linearly independent.

Definition. An orthogonal or orthonormal set that is a basis is called an *orthogonal* or *orthonormal basis.*

Example. The subset of $\mathbf{R}^3$ above is an orthogonal basis for $\mathbf{R}^3$. Write the vector below as a linear combination of those basis vectors.

$$\begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}$$

Proposition Let $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be an orthogonal basis for V, and $\mathbf{u}$ a vector in V. Then

$$\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v_1}}{|\mathbf{v}_1|^2}\,\mathbf{v}_1 + \cdots + \frac{\mathbf{u} \cdot \mathbf{v_k}}{|\mathbf{v}_k|^2}\,\mathbf{v}_k, \qquad \text{or, if basis is orthonormal:} \qquad \mathbf{u} = (\mathbf{u} \cdot \mathbf{v_1})\mathbf{v}_1 + \cdots + (\mathbf{u} \cdot \mathbf{v_k})\mathbf{v}_k$$

Every subspace of $\mathbf{R}^n$ has an orthogonal and hence an orthonormal basis.

Theorem 6.6. The Gram-Schmidt process. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be a basis for the subspace W of $\mathbf{R}^n$. Define the vectors:

$$\mathbf{v}_1 = \mathbf{u}_1, \quad \mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u_2} \cdot \mathbf{v_1}}{|\mathbf{v}_1|^2}\mathbf{v}_1, \quad \mathbf{v}_3 = \mathbf{u}_3 - \frac{\mathbf{u_3} \cdot \mathbf{v_1}}{|\mathbf{v}_1|^2}\mathbf{v}_1 - \frac{\mathbf{u_3} \cdot \mathbf{v_2}}{|\mathbf{v}_2|^2}\mathbf{v}_2, \ldots,$$

$$\mathbf{v}_k = \mathbf{u}_k - \frac{\mathbf{u_k} \cdot \mathbf{v_1}}{|\mathbf{v}_1|^2}\mathbf{v}_1 - \frac{\mathbf{u_k} \cdot \mathbf{v_2}}{|\mathbf{v}_2|^2}\mathbf{v}_2 - \cdots - \frac{\mathbf{u_k} \cdot \mathbf{v_{k-1}}}{|\mathbf{v}_{k-1}|^2}\mathbf{v}_{k-1}$$

Then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthogonal basis for W, and furthermore, for every i

$$\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_i\} = \mathrm{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_i\}.$$

Main idea of the construction.

6.3 Orthogonal Projections

Definition. The *orthogonal complement* $\mathcal{S}^\perp$ of a nonempty subset $\mathcal{S}$ of $\mathbf{R}^n$ is the set of all vectors in $\mathbf{R}^n$ that are orthogonal to every vector in $\mathcal{S}$.

$$\mathcal{S}^\perp = \{\mathbf{v} \in \mathbf{R}^n \mid \mathbf{v} \cdot \mathbf{u} = 0 \text{for every } \mathbf{u} \in \mathcal{S}\}$$

Example. What is $\mathcal{S}^\perp$ if $\mathcal{S}$ is one of the sets below in $\mathbf{R}^3$?

$\mathcal{S}_1 = \{\mathbf{u}\}$

$\mathcal{S}_2 = \{\mathbf{u}, \mathbf{v}\}$

$\mathcal{S}_3 =$ a line in $\mathbf{R}^3$

$\mathcal{S}_4 =$ a plane in $\mathbf{R}^3$

Proposition. The orthogonal complement of any nonzero subset of $\mathbf{R}^n$ is a subspace or $\mathbf{R}^n$.

Proposition. For any nonempty set $\mathcal{S}$, $\mathcal{S}^\perp = (\operatorname{Span} S)^\perp$. In particular the orthogonal complement of a basis of a subspace is the same as the orthogonal complement of the subspace.

Example. Find the orthogonal complement of the subspace of $\mathbf{R}^4$ that is generated by

$$\left\{ \begin{bmatrix} 1 \\ 3 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \\ -3 \\ 7 \end{bmatrix} \right\}$$

Proposition. For an $m \times n$ matrix A, $(\text{Row}\, A)^{\perp} = \text{Null}\, A$ and $(\text{Col}\, A)^{\perp} = \text{Null}\, A^T$.

Theorem 6.7. Orthogonal Decomposition Theorem. Let W be a subspace of $\mathbf{R}^n$. Then, for every vector $\mathbf{u}$ in $\mathbf{R}^n$, there exist unique vectors $\mathbf{w} \in W$ and $\mathbf{z} \in W^{\perp}$ such that $\mathbf{u} = \mathbf{w} + \mathbf{z}$. Furthermore, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthonormal basis for W, then

$$\mathbf{w} = (\mathbf{u} \cdot \mathbf{v_1})\mathbf{v}_1 + \cdots + (\mathbf{u} \cdot \mathbf{v_k})\mathbf{v}_k$$

Proof.

Proposition. For any subspace W of $\mathbf{R}^n$, $\dim W + \dim W^\perp = n$.

Definition. Let W be a subspace of $\mathbf{R}^n$ and let $\mathbf{u}$ be a vector. The *orthogonal projection of* $\mathbf{u}$ *onto* W is the unique vector $\mathbf{w}$ such that $\mathbf{u}-\mathbf{w}$ is in $W^\perp$. The function $U_W : \mathbf{R}^n \to \mathbf{R}^n$ that sends every vector $\mathbf{u}$ to its orthogonal projection on W is linear and is called the *orthogonal projection operator.*

We show U_W *is linear.*

We would like to get the standard matrix P_W of U_W.

Proposition. Let C be an $n \times k$ matrix whose columns are linearly independent. Then $C^T C$ is an invertible $k \times k$ matrix.

Theorem 6.8. Let C be an $n \times k$ matrix whose columns are a basis for a subspace W of $\mathbf{R}^n$. Then the standard matrix of the orthogonal projection to W is

$$P_W = C(C^T C)^{-1} C^T$$

Proof.

Example. Find the standard matrix of the orthogonal projection in $\mathbf{R}^3$ to:

line spanned by $\begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$

plane spanned by $\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$

Proposition (Closest Vector Property). Let W be a subspace of $\mathbf{R}^n$ and $\mathbf{u}$ a vector in $\mathbf{R}^n$. Among all vectors in W, the orthogonal projection of $\mathbf{u}$ onto W is the closest to $\mathbf{u}$.