

Document 20: Orthogonality

Recall the following properties of orthogonality from the previous chapters. These will guide this section.

1. $\vec{v} \perp \vec{w} \Leftrightarrow \vec{v} \cdot \vec{w} = 0$
2. $\text{Length}(\vec{v}) \equiv \text{norm}(\vec{v})$. $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$
3. \vec{u} is a unit vector if $\|\vec{u}\| = 1$

Definition of Orthogonal: $\{\vec{u}_1, \dots, \vec{u}_m\} \in \mathbb{R}^n$ are orthogonal if $\|\vec{u}_i\| = 1$ and $\vec{u}_i \cdot \vec{u}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

Let us now generalize the projection formula. Given $\vec{x}, \{\vec{u}_1, \dots, \vec{u}_m\}$, seeking $\vec{x}^{\parallel} = \text{proj}_V(\vec{x})$, we get $\vec{x}^{\parallel} = (\vec{x} \cdot \vec{u}_1)\vec{u}_1 + (\vec{x} \cdot \vec{u}_2)\vec{u}_2 + \dots + (\vec{x} \cdot \vec{u}_m)\vec{u}_m$.

There also exists some general properties of orthogonality. Let us consider the orthogonal compliment.

Definition Orthogonal Compliment: Let V be a subspace of \mathbb{R}^n . Then, the orthogonal compliment of V is $V^{\perp} = \{\vec{x} \in \mathbb{R}^n : \exists \vec{v} \in V, \vec{v} \cdot \vec{x} = 0\}$. Interestingly:

1. V^{\perp} is also a subspace of \mathbb{R}^n
2. $V^{\perp} = \ker(\text{proj}_V(\vec{x}))$
3. $V \cap V^{\perp} = \{\vec{0}\}$
4. $\dim(V) + \dim(V^{\perp}) = n$
5. $(V^{\perp})^{\perp} = V$

Finding the orthogonal basis of a subspace V can be done algorithmically through what is called the Gram-Schmidt Process.

Given a basis of $V = \{\vec{v}_1, \dots, \vec{v}_m\}$, we want the orthogonal basis $\{\vec{u}_1, \dots, \vec{u}_m\}$. Find the vectors in the basis using the following equations:

$$\begin{aligned}\vec{u}_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|} \\ \vec{u}_2 &= \frac{\vec{v}_2^\perp}{\|\vec{v}_2^\perp\|}, \vec{v}_2^\perp = \vec{v}_2 - (\vec{u}_1 \cdot \vec{v}_2)\vec{u}_1 \\ \vec{u}_3 &= \frac{\vec{v}_3^\perp}{\|\vec{v}_3^\perp\|}, \vec{v}_3^\perp = \vec{v}_3 - (\vec{u}_1 \cdot \vec{v}_3)\vec{u}_1 - (\vec{u}_2 \cdot \vec{v}_3)\vec{u}_2 \\ &\vdots \\ \vec{u}_m &= \frac{\vec{v}_m^\perp}{\|\vec{v}_m^\perp\|}, \vec{v}_m^\perp = \vec{v}_m - (\vec{u}_1 \cdot \vec{v}_m)\vec{u}_1 - \dots - (\vec{u}_{m-1} \cdot \vec{v}_m)\vec{u}_{m-1}\end{aligned}$$

Definition Orthogonal Transformations: Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$. T is orthogonal if it preserves the length of vectors. $\|T(\vec{x})\| = \|\vec{x}\|, \forall \vec{x} \in \mathbb{R}^n$.

Definition Orthogonal Matrix: If $T(\vec{x}) = A\vec{x}$ is orthogonal, then A is an orthogonal matrix and $\|A\vec{x}\| = \|\vec{x}\|$

Property: Orthonormal Vectors: These vectors make up a basis for a given space. Hence, they are linearly independent and span the space.

Properties of Orthogonality & Transposes:

1. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be orthogonal. Let $\vec{v} \perp \vec{w}$, then $T(\vec{v}) \perp T(\vec{w})$
2. T is orthogonal $\Leftrightarrow \{T(\vec{e}_1), \dots, T(\vec{e}_n)\}$ forms an orthogonal basis of \mathbb{R}^n
3. $A_{n \times m}$ is orthogonal \Leftrightarrow its columns form an orthogonal basis of \mathbb{R}^n
4. A is orthogonal $\Leftrightarrow (A^T = A^{-1} \Rightarrow A^T A = I)$
5. A , and B are orthogonal $\Rightarrow AB$ is orthogonal
6. A^{-1} is orthogonal $\Leftrightarrow A$ is orthogonal
7. A is symmetric if $A^T = A$
8. A is skew symmetric if $A^T = -A$
9. $(AB)^T = B^T A^T$
10. A is invertible $\Rightarrow A^T, (A^T)^{-1}, (A^{-1})^T$ are invertible
11. $\text{rank}(A) = \text{rank}(A^T)$

Document 21: Determinants

Recall the formula for calculating the determinant of a 2×2 matrix. Consider matrix A , defined by arbitrary elements $a, b, c, d \in \mathbb{R}$.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow \det(A) = ad - bc$$

Also, A is invertible $\Leftrightarrow \det(A) \neq 0$

Now, we redefine A to consider a larger matrix. Define column vectors $\vec{u}, \vec{v}, \vec{w}$ implicitly in the following definition:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} \vec{u} & \vec{v} & \vec{w} \end{bmatrix}$$

Geometrically, the determinant of A is defined: $\det(A) = \vec{u} \cdot (\vec{v} \times \vec{w})$. Additionally the same property of invertability also applies to 3×3 matrices, and can be stated as: A is invertible $\Leftrightarrow \vec{u} \cdot (\vec{v} \times \vec{w}) \neq 0$.

More generally, determinants can be computed with the Laplace Expansion for Determinants. Laplace expansion works by removing one row or column, then if a column was removed, it iterates over each row, and if a row was removed, it iterates over each column. Then the i,j 'th element is multiplied by the i,j 'th sub-matrix. Because it can be done two ways, there are two options for the formula.

Let $A_{n \times n}$ be an arbitrary matrix. Define A_{ij} , the A matrix without row i or column j . (Note: the determinant of A_{ij} is called the i,j 'th minor of A .) The Laplace Expansion for Determinants of A is:

$$\sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$$

Or equally, $\sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$

A shortcut for upper and lower triangular matrices is that their determinants are equal to the product of their diagonal entries.

Additionally, the determinant of an $n \times n$ matrix is a linear transformation of each row when the remaining rows are held fixed. That means that for one row, we can apply row operations and have the properties of linearity apply such as for the following:

Consider finding the determinant of a column vector with the same elements as this set: $\{e_1, e_2, \dots, e_r, \dots, e_n\}$. Holding all rows other than r fixed, and rewriting r in terms of scalars $u, k, v \in \mathbb{R}$ such that $r = u + kv$, we have the following.

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u + kv \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} + k \det \begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ v \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix}$$

Additional Properties for the Section:

1. If one row of A is multiplied by a scalar k to yield B , then $\det(B) = k \cdot \det(A)$
2. If n row swaps are performed on A to yield B , $\det(B) = (-1)^n \det(A)$
3. If a multiple of one row in A is added to another row in A to yield B , then $\det(A) = \det(B)$
4. $\det(AB) = (\det(A))(\det(B))$
5. $A \sim B \Rightarrow \det(A) = \det(B)$. However, the converse is not true.
6. $\det(A^{-1}) = \frac{1}{\det(A)} = (\det(A))^{-1}$
7. $\det(A^T) = \det(A)$
8. A is orthogonal $\Rightarrow \det(A) = \pm 1$
9. Let A be orthogonal with $\det(A) = 1 \Rightarrow A$ is a rotation matrix.
10. Consider vectors \vec{v}_1, \vec{v}_2 which form two side lengths of a parallelogram, with angle θ between them, and with area a . Given $A = [\vec{v}_1 | \vec{v}_2]$, $|\det(A)| = a \Rightarrow \det(A) = \|\vec{v}_1\| \cdot \|\vec{v}_2\| \sin(\theta)$
11. $|\det(A)| = \|\vec{v}_1\| \cdot \|\vec{v}_2^\perp\| \cdots \|\vec{v}_n^\perp\|$
12. Cramer's Rule: Given $A\vec{x} = \vec{b}$ to solve, let A_{bi} be the matrix A with the i 'th column replaced by \vec{b} . By Cramer's Rule we have $x_i = \frac{\det(A_{bi})}{\det(A)}$

Document 22: Eigenvalues and Eigenvectors

This section covers eigenvalues and eigenvectors, which will be important in further mathematics and useful for applications of linear algebra.

Definition Eigenvectors/Eigenvalues: A nonzero vector $\vec{v} \in \mathbb{R}^n$ is an eigenvector of A if $\exists \lambda \in \mathbb{R}, A\vec{v} = \lambda\vec{v}$. The Eigenvalue is λ .

When finding Eigenvalues/Eigenvectors, we must find values such that $(A - \lambda I)\vec{v} = \vec{0}$. This is a homogeneous linear system. Recall $A\vec{x} = \vec{0}$ has either 1 or infinite solutions. The eigenvalues are produced with $\det(A - \lambda I) = 0$, and the eigenvectors are produced by substituting into $(A - \lambda I)\vec{v} = \vec{0}$.

For finding eigenvalues, we need to find the characteristic polynomial of the matrix A . We begin with $P_\lambda(A) = \det(A - \lambda I)$. We will solve and get some values, lets say λ_1, λ_2 . Then we will solve $(A - \lambda_1 I)\vec{v} = \vec{0}$. We can use an augmented matrix and compute the $\text{rref}(A)$ to solve this. This may have infinitely many solutions, for instance:

$$\begin{bmatrix} \vec{v}_1 \\ \vec{v}_2 \end{bmatrix} = s \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

This was a basic introduction to eigenvalues and eigenvectors. This should be informative, yet basic. Further exploration is necessary.

Document 23: Eigenvalues / Eigenvectors Continued

Reexamining the methods for computing Eigenvalues and Eigenvectors, we have a new formula for the characterisic polynomial of the vector.

$$\text{Given } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \text{ we have } A - \lambda I = \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix}$$

$$P_A(\lambda) = (a - \lambda)(d - \lambda) - bc$$

$$\text{Or, } P_A(\lambda) = \lambda^2 + \text{trace}(A)\lambda + \det(A)$$

These formulas only work for 2×2 matrices. Another theorem that exists for 2×2 matrices is that for any given matrix A , defined above, the following holds true.

$$\text{Define } B_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, B_2 = \begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_2 \end{bmatrix}, B_3 = \begin{bmatrix} \sigma & -\tau \\ \tau & \sigma \end{bmatrix}, \lambda = \sigma \pm \tau i$$

$$A \sim B_1 \vee A \sim B_2 \vee A \sim B_3$$

It is worthy of noting that B_1, B_2, B_3 are each referred to as the Jordan Canonical Forms of the matrix A . Shifting to diagonalization, lets define Algebraic Multiplicity as the number of times an eigenvalue appears in the characteristic polynomial.

The eigenvalues of triangular matrices are the diagonal entries.

Given a characteristic polynomial $P_A(\lambda) = (1 - \lambda)^3(2 - \lambda)^2 = 0$, we have $\lambda_1 = 1$, with an Algebraic Multiplicity of 3, and $\lambda_2 = 2$, with an Algebraic Multiplicity of 2.

Definition Geometric Multiplicity: Let $A\vec{v} - \lambda\vec{v} = 0 \Rightarrow (A - \lambda I)\vec{v} = 0 \Rightarrow \vec{v} \in \ker(A - \lambda I)$. Let $E_\lambda = \ker(A - \lambda I)$. Then, $\dim(E_\lambda)$ is the Geometric Multiplicity of λ .

Definition diagonalizable: Let the function yielding the Algebraic Multiplicity be $\text{a.m.}(\lambda_i)$. Let the function yielding the Geometric Multiplicity be $\text{g.m.}(\lambda_i)$. $A_{m \times n}$ is diagonalizable $\Leftrightarrow \text{a.m.}(\lambda_i) = \text{g.m.}(\lambda_i)$. Then, $\exists B$ matrix such that $A \sim B$ with $S = [\text{eigenvector}(A)]$.

Document 24: Diagonalization and E-Values/Vectors

Diagonalization is a really useful skill to have, and getting there involves the Eigenvalues and Eigenvectors of a matrix. The steps to get there are as follows.

1. Find characteristic polynomial
2. Find the eigenvalues and eigenvectors
3. Diagonalize A

For step one, $P_A(\lambda) = \det(A - \lambda I)$. Lets define a matrix A for this example, and compute $A - \lambda I$.

$$A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}, A - \lambda I = \begin{bmatrix} 2 - \lambda & 2 & 2 \\ 2 & 2 - \lambda & 2 \\ 2 & 2 & 2 - \lambda \end{bmatrix}$$

$$\begin{aligned} \text{Then, } P_A(\lambda) &= (2 - \lambda) \cdot \det \left(\begin{bmatrix} 2 - \lambda & 2 \\ 2 & 2 - \lambda \end{bmatrix} \right) - 2 \cdot \det \left(\begin{bmatrix} 2 & 2 \\ 2 & 2 - \lambda \end{bmatrix} \right) + \\ & 2 \cdot \det \left(\begin{bmatrix} 2 & 2 \\ 2 - \lambda & 2 \end{bmatrix} \right) = (2 - \lambda) \cdot (\lambda^2 - 4\lambda) - 2 \cdot (-2\lambda) + 2 \cdot (2\lambda) = 6\lambda^2 - \lambda^3 = \\ & -\lambda^2(\lambda - 6). \end{aligned}$$

Then, for step two, we identify the eigenvalues from the characteristic polynomial: we have $\lambda_1 = 0, \lambda_2 = 6$. We must also compute the eigenvectors. To compute the eigenvectors, we must solve $(A - \lambda I)\vec{v} = \vec{0}$. First, lets compute $A - \lambda_1 I$ and $A - \lambda_2 I$.

$$A - \lambda_1 I = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}, A - \lambda_2 I = \begin{bmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{bmatrix}$$

Then, compute $(A - \lambda_1 I)\vec{v} = \vec{0}$. Given that \vec{v} has elements v_1, v_2, v_3 , we have $v_1 = -v_2 - v_3, v_2 = v_2, v_3 = v_3$. Defining free parameters s, t such that $s = v_2, t = v_3$, we can rewrite the closed form for the solution.

$$\vec{v} = s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Lets redefine \vec{v} within the scope of a new equation to compute $(A - \lambda_2 I)\vec{v} = \vec{0}$. Given that \vec{v} has elements v_1, v_2, v_3 , we have $v_1 = v_3, v_2 =$

$v_3, v_3 = v_3$. Lets define s , the free parameter with $s = v_3$. In closed form, we have the following.

$$\vec{v} = s \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Now, we have the eigenvalues $\lambda_1 = 0, \lambda_2 = 6$. Drawing from the solutions above, define vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$, the eigenvectors of the matrix as follows:

$$\vec{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

For step three, diagonalizing the matrix, lets first ensure that the matrix is diagonalizable. From the characteristic polynomial, we find the algebraic multiplicities for λ_1 and λ_2 , and from the dimentions of the kernels we found earlier, we get the geometric multiplicities of λ_1 and λ_2 : a.m.1= 2, g.m.1= 2, a.m.2= 1, g.m.2= 1. Since both the algebraic and geometric multiplicities equal, the matrix is diagonalizable.

The matrix A will be similar to a matrix D by a matrix S defined below. This matrix D is the diagonalization of the matrix A .

$$S = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Next, we compute the inverse of S and use properties of similarity to find D . We will do this below:

$$S^{-1} = \frac{1}{3} \begin{bmatrix} -1 & 2 & -1 \\ -1 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$$

$$3S^{-1}A = \begin{bmatrix} -1 & 2 & -1 \\ -1 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 6 & 6 & 6 \end{bmatrix}$$

$$\therefore S^{-1}A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & 2 & 2 \end{bmatrix}$$

Finally, multiply by S to and use properties of similarity to compute D .

The matrix D is:

$$D = S^{-1}AS = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

These steps will always yield the matrix D , but the similarity of A and D is not part of finding D as much as it is a beneficial property of D . The diagonalized matrix is a lot more simple and easier to work with, and is similar to A , and the similarity is why it is valuable, not just a part of the process of finding it. In fact, you don't need to use similarity properties at all to find D . Consider the following:

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

Shortcuts to finding D are perfectly valid. Given we have $S = [\vec{v}_1, \vec{v}_2, \vec{v}_3]$, the matrix D should be the above matrix where λ_1 is the eigenvalue that produced the eigenvector \vec{v}_1 and so on.