
Week 12

June 5, 2026

Contents

1 Diagonalizable Matrices and Multiplicity

Recall that a *diagonal matrix* D is a matrix containing a zero in every entry except those on the main diagonal. More precisely, if d_{ij} is the ij^{th} entry of a diagonal matrix D , then $d_{ij} = 0$ unless $i = j$. Such matrices look like the following.

$$D = \begin{bmatrix} * & & 0 \\ & \ddots & \\ 0 & & * \end{bmatrix}$$

where $*$ is a number which might not be zero.

Diagonal matrices have some nice properties, as we demonstrate below.

Exploration 1 Let $M = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ and let $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 10 \end{bmatrix}$. Compute MD and DM using Octave.

```
% Define M
M=[1 2 3; 4 5 6; 7 8 9]

% The function diag(vector) creates a square matrix with diagonal
%vector and the remaining entries zero.
D=diag([2 -5 10])

% The function diag() can also be used to return the diagonal of a matrix.
diag_M=diag(M)

% Understanding multiplication by a diagonal matrix
product_MD=M*D
product_DM=D*M

% Raising a diagonal matrix to a power
D5=D^5
```

[Link to code.](#)

Notice the patterns present in the product matrices.

- Each row of DM is the same as its corresponding row of M multiplied by the scalar which is the corresponding diagonal element of D .
- In the product MD , it is the columns of M that have been multiplied by the diagonal elements.
- Raising D to the 5^{th} power amounts to raising each diagonal element to the 5^{th} power.

These patterns hold in general for any diagonal matrix, and they are fundamental to understanding diagonalization and its uses. We discuss diagonalization next.

Definition 1. Let A be an $n \times n$ matrix. Then A is said to be *diagonalizable* if there exists an invertible matrix P such that

$$P^{-1}AP = D$$

where D is a diagonal matrix. In other words, a matrix A is diagonalizable if it is similar to a diagonal matrix, $A \sim D$.

If we are given a matrix A that is diagonalizable, then we can write $P^{-1}AP = D$ for some matrix P , or, equivalently,

$$AP = PD \quad (1)$$

If we pause to examine Equation (??), the work that we did in Exploration ?? can help us to understand how to find P that will diagonalize A . The product PD is formed by multiplying each column of P by a scalar which is the corresponding element on the diagonal of D . To restate this, if $\mathbf{x}_i$ is column i in our matrix P , then Equation (??) tells us that

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad (2)$$

where λ_i is the i th diagonal element of D .

Of course, Equation (??) is very familiar! We see that if we are able to diagonalize a matrix A , the columns of matrix P will be the eigenvectors of A , and the corresponding diagonal entries of D will be the corresponding eigenvalues of A . This is summed up in the following theorem.

Theorem 1. An $n \times n$ matrix A is diagonalizable if and only if there is an invertible matrix P given by

$$P = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ | & | & \cdots & | \end{bmatrix},$$

where the columns $\mathbf{x}_i$ are eigenvectors of A .

Moreover, if A is diagonalizable, the corresponding eigenvalues of A are the diagonal entries of the diagonal matrix D .

Proof Suppose P is given as above as an invertible matrix whose columns are eigenvectors of A . To show that A is diagonalizable, we will show

$$AP = PD,$$

which is equivalent to $P^{-1}AP = D$. We have

$$AP = \begin{bmatrix} | & | & \cdots & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & \cdots & A\mathbf{x}_n \\ | & | & \cdots & | \end{bmatrix},$$

while

$$PD = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \lambda_1 \mathbf{x}_1 & \lambda_2 \mathbf{x}_2 & \cdots & \lambda_n \mathbf{x}_n \\ | & | & \cdots & | \end{bmatrix}$$

We can complete this half of the proof by comparing columns, and noting that

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i \quad (3)$$

for $i = 1, \dots, n$ since the $\mathbf{x}_i$ are eigenvectors of A and the λ_i are corresponding eigenvalues of A .

Conversely, suppose A is diagonalizable so that $P^{-1}AP = D$. Let

$$P = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ | & | & \cdots & | \end{bmatrix}$$

where the columns are the vectors $\mathbf{x}_i$ and

$$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

Then

$$AP = PD = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

and so

$$\begin{bmatrix} | & | & \cdots & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & \cdots & A\mathbf{x}_n \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \cdots & \lambda_n\mathbf{x}_n \\ | & | & \cdots & | \end{bmatrix}$$

showing the $\mathbf{x}_i$ are eigenvectors of A and the λ_i are eigenvalues. ■

Notice that because the matrix P defined above is invertible it follows that the set of eigenvectors of A , $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, is a basis of $\mathbb{R}^n$.

We demonstrate the concept given in the above theorem in the next example. Note that not only are the columns of the matrix P formed by eigenvectors, but P must be invertible, and therefore must consist of a linearly independent set of eigenvectors.

Example 1. Let

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & -1 \\ -2 & -4 & 4 \end{bmatrix}$$

Find an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$.

Explanation. We will use eigenvectors of A as the columns of P , and the corresponding eigenvalues of A as the diagonal entries of D . The eigenvalues of A are $\lambda_1 = 2$, $\lambda_2 = 2$, and $\lambda_3 = 6$. We leave these computations as exercises, as well as the computations to find a basis for each eigenspace.

One possible basis for $\mathcal{S}_2$, the eigenspace corresponding to 2, is $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$, while a basis for $\mathcal{S}_6$ is given by

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

We construct the matrix P by using these basis elements as columns.

$$P = \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix}$$

You can verify (See Practice Problem) that

$$P^{-1} = \begin{bmatrix} -1/4 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & -1/4 \end{bmatrix}$$

Thus,

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} -1/4 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & -1/4 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & -1 \\ -2 & -4 & 4 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix} \end{aligned}$$

You can see that the result here is a diagonal matrix where the entries on the main diagonal are the eigenvalues of A . Notice that eigenvalues on the main diagonal *must* be in the same order as the corresponding eigenvectors in P .

Exploration 2 We will continue to use matrices from Example ?? to illustrate why diagonalizable matrices are easier to work with. Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & -1 \\ -2 & -4 & 4 \end{bmatrix}$ and let $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix}$. Would it be easier to compute A^5 or D^5 if you had to do so by hand, without a computer? Certainly D^5 is easier, due to the number of zero entries!

Raising a diagonal matrix to a power amounts to simply raising each entry to that same power, whereas computing A^5 requires many more calculations. However, we learned in Example ?? that A is similar to D . We will use this to make our computation easier as follows

$$\begin{aligned} A^5 &= (PDP^{-1})^5 \\ &= (PDP^{-1})(PDP^{-1})(PDP^{-1})(PDP^{-1})(PDP^{-1}) \\ &= PD(P^{-1}P)D(P^{-1}P)D(P^{-1}P)D(P^{-1}P)DP^{-1} \\ &= PD(I)D(I)D(I)D(I)DP^{-1} \\ &= PD^5P^{-1} \end{aligned}$$

With this in mind, it is not as daunting to calculate A^5 by hand.

$$A^5 = \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 32 & 0 & 0 \\ 0 & 32 & 0 \\ 0 & 0 & 7776 \end{bmatrix} \begin{bmatrix} -1/4 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & -1/4 \end{bmatrix} = \begin{bmatrix} 32 & 0 & 0 \\ 1936 & 3904 & -1936 \\ -3872 & -7744 & 3904 \end{bmatrix}$$

The savings in work would certainly be more pronounced for larger matrices or for powers larger than 5.

In Exploration ??, because matrix A was diagonalizable, we were able to cut down on computations. When we chose to work with D and P instead of A we worked with the eigenvalues and eigenvectors of A . Each column of P is an eigenvector of A , and so we repeatedly made use of the following theorem (with $m = 5$).

Theorem 2. Let A be an $n \times n$ matrix and suppose $A\mathbf{x} = \lambda\mathbf{x}$. Then $A^m\mathbf{x} = \lambda^m\mathbf{x}$

Proof We prove this theorem by induction on m . Clearly $A^m\mathbf{x} = \lambda^m\mathbf{x}$ holds when $m = 1$, as that was given. For the inductive step, suppose that we know $A^{m-1}\mathbf{x} = \lambda^{m-1}\mathbf{x}$. Then

$$\begin{aligned} A^m\mathbf{x} &= (AA^{m-1})\mathbf{x} \\ &= A(A^{m-1}\mathbf{x}) \\ &= A(\lambda^{m-1}\mathbf{x}) \\ &= \lambda^{m-1}A\mathbf{x} \\ &= \lambda^{m-1}\lambda\mathbf{x} \\ &= \lambda^m\mathbf{x} \end{aligned}$$

as desired. ■

Matrix A from the Example ?? and Exploration ?? had a repeated eigenvalue of 2. The next theorem and corollary show that matrices which have *distinct* eigenvalues (where none are repeated) have desirable properties.

Theorem 3. Let A be an $n \times n$ matrix, and suppose that A has distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_m$. For each i , let $\mathbf{x}_i$ be a λ_i -eigenvector of A . Then $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ is linearly independent.

Proof We prove this by induction on m , the number of vectors in the set. If $m = 1$, then $\{\mathbf{x}_1\}$ is a linearly independent set because $\mathbf{x}_1 \neq \mathbf{0}$. In general, suppose we have established that the theorem is true for some $m \geq 1$. Given eigenvectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{m+1}\}$, suppose

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_{m+1}\mathbf{x}_{m+1} = \mathbf{0}. \quad (4)$$

We must show that each $c_i = 0$. Multiply both sides of (??) on the left by A and use the fact that $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$ to get

$$c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 + \dots + c_{m+1}\lambda_{m+1}\mathbf{x}_{m+1} = \mathbf{0}. \quad (5)$$

If we multiply (??) by λ_1 and subtract the result from (??), the first terms cancel and we obtain

$$c_2(\lambda_2 - \lambda_1)\mathbf{x}_2 + c_3(\lambda_3 - \lambda_1)\mathbf{x}_3 + \dots + c_{m+1}(\lambda_{m+1} - \lambda_1)\mathbf{x}_{m+1} = \mathbf{0}.$$

Since $\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{m+1}$ correspond to distinct eigenvalues $\lambda_2, \lambda_3, \dots, \lambda_{m+1}$, the set $\{\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{m+1}\}$ is linearly independent by the induction hypothesis. Hence,

$$c_2(\lambda_2 - \lambda_1) = 0, \quad c_3(\lambda_3 - \lambda_1) = 0, \quad \dots, \quad c_{m+1}(\lambda_{m+1} - \lambda_1) = 0$$

and so $c_2 = c_3 = \dots = c_{m+1} = 0$ because the λ_i are distinct. It follows that (??) becomes $c_1\mathbf{x}_1 = \mathbf{0}$, which implies that $c_1 = 0$ because $\mathbf{x}_1 \neq \mathbf{0}$, and the proof is complete. ■

The corollary that follows from this theorem gives a useful tool in determining if A is diagonalizable.

Corollary 1. Let A be an $n \times n$ matrix and suppose it has n distinct eigenvalues. Then it follows that A is diagonalizable.

Remark 1. Note that Corollary ?? is NOT an “if and only if statement”. This means that if A has repeated eigenvalues it is still sometimes possible to diagonalize A , as seen in Example ??.

Definition 2. If we are able to diagonalize A , say $A = PDP^{-1}$, we say that PDP^{-1} is an *eigenvalue decomposition* of A .

Not every matrix has an eigenvalue decomposition. Sometimes we cannot find an invertible matrix P such that $P^{-1}AP = D$. Consider the following example.

Example 2. Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

If possible, find an invertible matrix P and a diagonal matrix D so that $P^{-1}AP = D$.

Explanation. We see immediately (how?) that the eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 1$. To find P , the next step would be to find a basis for the corresponding eigenspace S_1 . We solve the equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Writing this equation as an augmented matrix, we already have a matrix in row echelon form:

$$\left[\begin{array}{cc|c} 0 & -1 & 0 \\ 0 & 0 & 0 \end{array} \right]$$

We see that the eigenvectors in S_1 are of the form

$$\begin{bmatrix} t \\ 0 \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

so a basis for the eigenspace S_1 is given by $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. It is easy to see that we cannot form an invertible matrix P , because any two eigenvectors will be of the form $\begin{bmatrix} t \\ 0 \end{bmatrix}$, and so the second row of P would have a row of zeros, and P could not be invertible. Hence A cannot be diagonalized.

We saw earlier in Corollary ?? that an $n \times n$ matrix with n distinct eigenvalues is diagonalizable. It turns out that there are other useful diagonalizability tests.

Recall that the *algebraic multiplicity* of an eigenvalue λ is the number of times that it occurs as a root of the characteristic polynomial.

Definition 3. The *geometric multiplicity* of an eigenvalue λ is the dimension of the corresponding eigenspace S_λ .

Consider now the following lemma.

Lemma 1. Let A be an $n \times n$ matrix, and let S_{λ_1} be the eigenspace corresponding to the eigenvalue λ_1 which has algebraic multiplicity m . Then

$$\dim(S_{\lambda_1}) \leq m$$

In other words, the geometric multiplicity of an eigenvalue is less than or equal to the algebraic multiplicity of that same eigenvalue.

Proof Let k be the geometric multiplicity of λ_1 , i.e., $k = \dim(S_{\lambda_1})$. Suppose $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is a basis for the eigenspace S_{λ_1} . Let P be any invertible matrix having $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ as its first k columns, say

$$P = \begin{bmatrix} | & | & & | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_k & \mathbf{x}_{k+1} & \cdots & \mathbf{x}_n \\ | & | & & | & | & & | \end{bmatrix}.$$

In block form we may write

$$P = \begin{bmatrix} B & C \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} D \\ E \end{bmatrix}$$

where B is $n \times k$, C is $n \times (n - k)$, D is $k \times n$, and E is $(n - k) \times n$. We observe $I_n = P^{-1}P = \left[\begin{array}{c|c} DB & DC \\ \hline EB & EC \end{array} \right]$. This implies

$$DB = I_k, \quad DC = O_{k \times (n-k)}, \quad EB = O_{(n-k) \times k} \quad \text{and} \quad EC = I_{n-k}$$

Therefore,

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} D \\ E \end{bmatrix} A \begin{bmatrix} B & C \end{bmatrix} = \begin{bmatrix} DAB & DAC \\ \hline EAB & EAC \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 DB & DAC \\ \hline \lambda_1 EB & EAC \end{bmatrix} = \begin{bmatrix} \lambda_1 I_k & DAC \\ \hline O & EAC \end{bmatrix} \end{aligned}$$

We finish the proof by comparing the characteristic polynomials on both sides of this equation, and making use of the fact that similar matrices have the same characteristic polynomials.

$$\det(A - \lambda I) = \det(P^{-1}AP - \lambda I) = (\lambda_1 - \lambda)^k \det(EAC)$$

We see that the characteristic polynomial of A has $(\lambda_1 - \lambda)^k$ as a factor. This tells us that algebraic multiplicity of λ_1 is at least k , proving the desired inequality. ■

This result tells us that if λ is an eigenvalue of A , then the number of linearly independent λ -eigenvectors is never more than the multiplicity of λ . We now use this fact to provide a useful diagonalizability condition.

Theorem 4. Let A be an $n \times n$ matrix A . Then A is diagonalizable if and only if for each eigenvalue λ of A , the algebraic multiplicity of λ is equal to the geometric multiplicity of λ .

Proof Suppose A is diagonalizable and let $\lambda_1, \dots, \lambda_t$ be the distinct eigenvalues of A , with algebraic multiplicities $m_1, \dots, m_t$, respectively and geometric multiplicities $k_1, \dots, k_t$, respectively. Since A is diagonalizable, Theorem ?? implies that $k_1 + \dots + k_t = n$. By applying Lemma ?? t times, we have

$$n = k_1 + \dots + k_t \leq m_1 + \dots + m_t = n,$$

which is only possible if $k_i = m_i$ for $i = 1, \dots, t$.

Conversely, if the geometric multiplicity equals the algebraic multiplicity of each eigenvalue, then obtaining a basis for each eigenspace yields n eigenvectors. Applying Theorem ??, we know that these n eigenvectors are linearly independent, so Theorem ?? implies that A is diagonalizable. ■

Practice Problems

Problems ??-??

In this exercise you will "fill in the details" of Example ??.

Problem 3 Find the eigenvalues of matrix

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & -1 \\ -2 & -4 & 4 \end{bmatrix}$$

Problem 4 Find a basis for each eigenspace of the matrix A .

Problem 5 Compute the inverse of

$$P = \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix}$$

Problem 6 Compute $D = P^{-1}AP$

Problem 7 Show that computing the inverse of P is not really necessary by comparing the products PD and AP .

Problem 8 In each case, decide whether the matrix A is diagonalizable. If so, find P such that $P^{-1}AP$ is diagonal.

1. $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ (Diagonalizable/ Not Diagonalizable)
2. $\begin{bmatrix} 3 & 0 & 6 \\ 0 & -3 & 0 \\ 5 & 0 & 2 \end{bmatrix}$ (Diagonalizable/ Not Diagonalizable)
3. $\begin{bmatrix} 3 & 1 & 6 \\ 2 & 1 & 0 \\ -1 & 0 & -3 \end{bmatrix}$ (Diagonalizable/ Not Diagonalizable)
4. $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 2 \\ 2 & 3 & 1 \end{bmatrix}$ (Diagonalizable/ Not Diagonalizable)

Problem 9 Let A denote an $n \times n$ upper triangular matrix.

1. If all the main diagonal entries of A are distinct, show that A is diagonalizable.

2. If all the main diagonal entries of A are equal, show that A is diagonalizable only if it is already diagonal. Click the arrow to see the answer. The eigenvalues of A are all equal (they are the diagonal elements), so if $P^{-1}AP = D$ is diagonal, then $D = \lambda I$. Hence $A = P^{-1}(\lambda I)P = \lambda I$.

3. Show that $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is diagonalizable but that $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is not diagonalizable.

Problem 10 Let A be a diagonalizable $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ (including multiplicities). Show that:

1. $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$
2. $\operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$

Problem 11 1. Show that two diagonalizable matrices are similar if and only if they have the same eigenvalues with the same multiplicities.

2. If A is diagonalizable, show that $A \sim A^T$.

3. Show that $A \sim A^T$ if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

Text Source

The text in this section is a compilation of material from Section 7.2.1 of Ken Kuttler's [A First Course in Linear Algebra](#) (CC-BY) and Section 5.5 of Keith Nicholson's [Linear Algebra with Applications](#) (CC-BY-NC-SA).

Ken Kuttler, *A First Course in Linear Algebra*, Lyryx 2017, Open Edition, p. 362-364.

Many of the Practice Problems are Exercises from W. Keith Nicholson, *Linear Algebra with Applications*, Lyryx 2018, Open Edition, pp. 298-310.

2 Orthogonal Matrices and Symmetric Matrices

Recall that an $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors. Moreover, the matrix P with these eigenvectors as columns is a diagonalizing matrix for A , that is

$$P^{-1}AP = D$$

where D is a diagonal matrix with diagonal entries consisting of the eigenvalues of A .

Orthogonal Matrices

A collection of non-zero, pairwise orthogonal vectors in $\mathbb{R}^n$ is called an *orthogonal* set of vectors. An orthogonal set of vectors is called *orthonormal* if $\|\mathbf{q}\| = 1$ for each vector $\mathbf{q}$ in the set. A set of orthogonal vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ can be “normalized”, i.e. converted into an orthonormal set $\left\{ \frac{1}{\|\mathbf{v}_1\|} \mathbf{v}_1, \frac{1}{\|\mathbf{v}_2\|} \mathbf{v}_2, \dots, \frac{1}{\|\mathbf{v}_k\|} \mathbf{v}_k \right\}$. In particular, if a matrix A has n orthogonal eigenvectors, they can (by normalizing) be taken to be orthonormal. The corresponding diagonalizing matrix (we will use Q instead of P) has orthonormal columns, and such matrices are very easy to invert.

Theorem 5. The following conditions are equivalent for an $n \times n$ matrix Q .

1. Q is invertible and $Q^{-1} = Q^T$.
2. The rows of Q are orthonormal.
3. The columns of Q are orthonormal.

Proof First note that condition ?? is equivalent to $Q^T Q = I$. Let $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ denote the columns of Q . Then $\mathbf{q}_i^T$ is the i th row of Q^T , so the (i, j) -entry of $Q^T Q$ is $\mathbf{q}_i \cdot \mathbf{q}_j$. Thus $Q^T Q = I$ means that $\mathbf{q}_i \cdot \mathbf{q}_j = 0$ if $i \neq j$ and $\mathbf{q}_i \cdot \mathbf{q}_i = 1$ if $i = j$. Hence condition ?? is equivalent to ??. The proof of the equivalence of ?? and ?? is similar. ■

Definition 4 (Orthogonal Matrices). An $n \times n$ matrix Q is called an *orthogonal matrix* if its columns are orthonormal vectors. (i.e. The matrix satisfies one (and hence all) of the conditions in Theorem ??.)

Remark 2. The above definition calls a square matrix with orthonormal columns an *orthogonal matrix*. It is true that *orthonormal matrix* might be a better name. But *orthogonal matrix* is standard.

Example 3. The rotation matrix $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ is orthogonal for any angle θ .

Explanation. See Practice Problem ??.

Exploration 12 Let $A = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$

1. Check that matrix A has rows that are orthogonal.

2. Check that matrix A has columns that are NOT orthogonal.
3. Check that matrix A has rows that are NOT orthonormal.
4. Create a matrix Q by normalizing each of the rows of A .
5. Check that columns and rows of Q are orthonormal.

You should get $Q = \begin{bmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$.

Definition 5. An $n \times n$ matrix A is said to be *orthogonally diagonalizable* if an orthogonal matrix Q can be found such that $Q^{-1}AQ = Q^T A Q$ is diagonal.

We have learned earlier that when we diagonalize a matrix A , we write $P^{-1}AP = D$ for some matrix P where D is diagonal, and the diagonal entries are the eigenvalues of A . We have also learned that the columns of the matrix P are the corresponding eigenvectors of A . So when a matrix is orthogonally diagonalizable, we are able to accomplish the diagonalization using a matrix Q consisting of n eigenvectors that form an orthonormal basis for $\mathbb{R}^n$. In the following section we will learn that the matrices that have this property are precisely the symmetric matrices.

Symmetric Matrices

A *symmetric matrix* is a matrix which is equal to its transpose.

When we began our study of eigenvalues and eigenvectors, we saw examples of matrices with entries that were real numbers with eigenvalues that were complex numbers. It can be shown that symmetric matrices only have real eigenvalues.

Fact 1. Symmetric matrices only have real eigenvalues.

Lemma 2. If A is an $n \times n$ symmetric matrix, then

$$(Ax) \cdot y = x \cdot (Ay)$$

for all columns x and y in $\mathbb{R}^n$.

Proof Recall that $x \cdot y = x^T y$ for all columns x and y . Because $A^T = A$, we get

$$(Ax) \cdot y = (Ax)^T y = x^T A^T y = x^T A y = x \cdot (Ay)$$

■

Remark 3. The converse of the above lemma also holds (see Practice Problem ??).

Theorem 6. If A is a symmetric matrix, then eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof Let $A\mathbf{x} = \lambda\mathbf{x}$ and $A\mathbf{y} = \mu\mathbf{y}$, where $\lambda \neq \mu$. We compute

$$\lambda(\mathbf{x} \cdot \mathbf{y}) = (\lambda\mathbf{x}) \cdot \mathbf{y} = (A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y}) = \mathbf{x} \cdot (\mu\mathbf{y}) = \mu(\mathbf{x} \cdot \mathbf{y})$$

Hence $(\lambda - \mu)(\mathbf{x} \cdot \mathbf{y}) = 0$, and so $\mathbf{x} \cdot \mathbf{y} = 0$ because $\lambda \neq \mu$. ■

Example 4. Find an orthogonal matrix Q such that $Q^{-1}AQ$ is diagonal, where $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix}$.

Explanation. The characteristic equation of A is (adding twice row 1 to row 2):

$$\det \begin{bmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & -2 \\ 1 & -2 & 5-\lambda \end{bmatrix} = \lambda(\lambda-1)(\lambda-6) = 0$$

Thus the eigenvalues are $\lambda = 0, 1$, and 6 , and corresponding eigenvectors are

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}$$

respectively. Moreover, these eigenvectors are *orthogonal*. We have $\|\mathbf{x}_1\|^2 = 6$, $\|\mathbf{x}_2\|^2 = 5$, and $\|\mathbf{x}_3\|^2 = 30$, so

$$P = \begin{bmatrix} \frac{1}{\sqrt{6}}\mathbf{x}_1 & \frac{1}{\sqrt{5}}\mathbf{x}_2 & \frac{1}{\sqrt{30}}\mathbf{x}_3 \end{bmatrix} = \frac{1}{\sqrt{30}} \begin{bmatrix} \sqrt{5} & 2\sqrt{6} & -1 \\ -2\sqrt{5} & \sqrt{6} & 2 \\ \sqrt{5} & 0 & 5 \end{bmatrix}$$

is an orthogonal matrix. Thus $P^{-1} = P^T$ and

$$P^TAP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

Example 5. Orthogonally diagonalize the symmetric matrix $A = \begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}$.

Explanation. The characteristic equation is

$$\det \begin{bmatrix} 8-\lambda & 2 & -2 \\ 2 & 5-\lambda & -4 \\ -2 & -4 & 5-\lambda \end{bmatrix} = \lambda(\lambda-9)^2 = 0$$

Hence the distinct eigenvalues are 0 and 9 are of algebraic multiplicity 1 and 2 , respectively. The geometric multiplicities must be the same, for A is diagonalizable, being symmetric. It follows that $\dim(\mathcal{S}_0) = 1$ and $\dim(\mathcal{S}_9) = 2$. Gaussian elimination gives

$$\mathcal{S}_0(A) = \text{span}\{\mathbf{x}_1\}, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \quad \text{and} \quad \mathcal{S}_9(A) = \text{span} \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

The eigenvectors in S_9 are both orthogonal to $\mathbf{x}_1$ as Theorem ?? guarantees, but not to each other. However, an orthogonal basis can be found using projections. (See [Gram-Schmidt Orthogonalization](#))

$$\{\mathbf{f}_2, \mathbf{f}_3\} \text{ of } S_9(A) \quad \text{where} \quad \mathbf{f}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{f}_3 = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

Normalizing gives orthonormal vectors $\{\frac{1}{3}\mathbf{x}_1, \frac{1}{\sqrt{5}}\mathbf{f}_2, \frac{1}{3\sqrt{5}}\mathbf{f}_3\}$, so

$$Q = \begin{bmatrix} \left| \frac{1}{3}\mathbf{x}_1 \right| & \left| \frac{1}{\sqrt{5}}\mathbf{f}_2 \right| & \left| \frac{1}{3\sqrt{5}}\mathbf{f}_3 \right| \\ \hline \frac{1}{3}\mathbf{x}_1 & \frac{1}{\sqrt{5}}\mathbf{f}_2 & \frac{1}{3\sqrt{5}}\mathbf{f}_3 \\ \hline \end{bmatrix} = \frac{1}{3\sqrt{5}} \begin{bmatrix} \sqrt{5} & -6 & 2 \\ 2\sqrt{5} & 3 & 4 \\ -2\sqrt{5} & 0 & 5 \end{bmatrix}$$

is an orthogonal matrix such that $Q^{-1}AQ$ is diagonal.

It is worth noting that other, more convenient, diagonalizing matrices Q exist. For example, $\mathbf{y}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ and $\mathbf{y}_3 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$

lie in $S_9(A)$ and they are orthogonal. Moreover, they both have norm 3 (as does $\mathbf{x}_1$), so

$$\hat{Q} = \begin{bmatrix} \left| \frac{1}{3}\mathbf{x}_1 \right| & \left| \frac{1}{3}\mathbf{y}_2 \right| & \left| \frac{1}{3}\mathbf{y}_3 \right| \\ \hline \frac{1}{3}\mathbf{x}_1 & \frac{1}{3}\mathbf{y}_2 & \frac{1}{3}\mathbf{y}_3 \\ \hline \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$$

is a nicer orthogonal matrix with the property that $\hat{Q}^{-1}A\hat{Q}$ is diagonal.

Practice Problems

Problem 13 Normalize the columns to make the following matrix orthogonal. $A = \begin{bmatrix} 1 & -4 \\ 2 & 2 \end{bmatrix}$

$$\frac{1}{\sqrt{\dots}} \begin{bmatrix} \dots & \dots \\ \dots & \dots \end{bmatrix}$$

Problem 14 Normalize the columns to make the following matrix orthogonal. $A = \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$

$$\frac{1}{\dots} \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}$$

Problem 15 For each matrix A , find an orthogonal matrix Q such that $Q^{-1}AQ$ is diagonal.

Question 15.1

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$Q = \frac{1}{\sqrt{\dots\dots}} \begin{bmatrix} 1 & -1 \\ \dots\dots & \dots\dots \end{bmatrix}$$

Question 15.2

$$A = \begin{bmatrix} 3 & 0 & 7 \\ 0 & 5 & 0 \\ 7 & 0 & 3 \end{bmatrix}$$

List eigenvalues in increasing order.

$$\lambda_1 = \dots\dots, \quad \lambda_2 = \dots\dots, \quad \lambda_3 = \dots\dots$$

$$Q^{-1}AQ = \begin{bmatrix} \dots\dots/\sqrt{\dots\dots} & \dots\dots & \dots\dots/\sqrt{\dots\dots} \\ \dots\dots & \dots\dots & \dots\dots \\ 1/\sqrt{\dots\dots} & \dots\dots & 1/\sqrt{\dots\dots} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} \dots\dots/\sqrt{\dots\dots} & \dots\dots & \dots\dots/\sqrt{\dots\dots} \\ \dots\dots & \dots\dots & \dots\dots \\ 1/\sqrt{\dots\dots} & \dots\dots & 1/\sqrt{\dots\dots} \end{bmatrix}$$

Problem 16 Suppose A is orthogonally diagonalizable. Prove that A is symmetric. (This is the easy direction of the "if and only if" in Theorem ??.)

Problem 17 Prove the converse of Theorem ??:

If $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y})$ for all n -columns $\mathbf{x}$ and $\mathbf{y}$, then A is symmetric.

Problem 18 Let Q be an orthogonal matrix.

1. Show that $\det Q = 1$ or $\det Q = -1$.
2. Give 2×2 examples of Q such that $\det Q = 1$ and $\det Q = -1$.

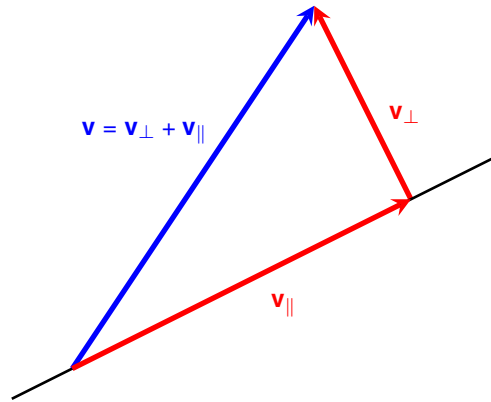
Text Source

This section was adapted from Section 8.2 of Keith Nicholson's *Linear Algebra with Applications*. (CC-BY-NC-SA)

W. Keith Nicholson, *Linear Algebra with Applications*, Lyryx 2018, Open Edition, p. 424

3 Orthogonal Projections

Given a line l and a vector $\mathbf{v}$ emanating from a point on l , it is sometimes convenient to express $\mathbf{v}$ as the sum of a vector $\mathbf{v}_{\parallel}$, parallel to l , and a vector $\mathbf{v}_{\perp}$, perpendicular to l . If you have taken a physics course, you may have seen a force vector decomposed into the sum of two components: one parallel and one perpendicular to the direction of motion.



Suppose $\mathbf{d}$ is a direction vector for l . Then $\mathbf{v}_{\parallel} = k\mathbf{d}$ for some scalar k . Our goal is to find k .

$$\begin{aligned}\mathbf{v} \cdot \mathbf{d} &= (\mathbf{v}_{\parallel} + \mathbf{v}_{\perp}) \cdot \mathbf{d} \\ &= (k\mathbf{d} + \mathbf{v}_{\perp}) \cdot \mathbf{d} \\ &= k\mathbf{d} \cdot \mathbf{d} + \mathbf{v}_{\perp} \cdot \mathbf{d} \\ &= k\|\mathbf{d}\|^2 + 0 \\ &= k\|\mathbf{d}\|^2\end{aligned}$$

We conclude that

$$k = \frac{\mathbf{v} \cdot \mathbf{d}}{\|\mathbf{d}\|^2}$$

and

$$\mathbf{v}_{\parallel} = k\mathbf{d} = \left(\frac{\mathbf{v} \cdot \mathbf{d}}{\|\mathbf{d}\|^2} \right) \mathbf{d}$$

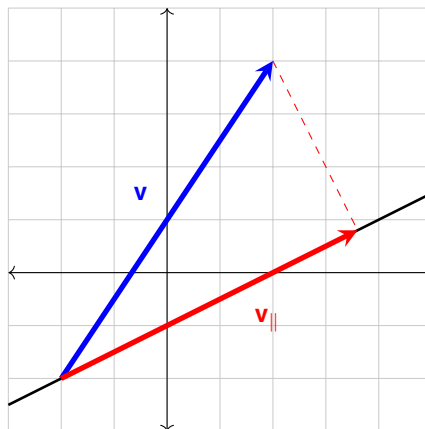
The vector $\mathbf{v}_{\parallel} = \left(\frac{\mathbf{v} \cdot \mathbf{d}}{\|\mathbf{d}\|^2} \right) \mathbf{d}$ is called the *projection of $\mathbf{v}$ onto $\mathbf{d}$* . In our discussion, $\mathbf{d}$ is a direction vector for line l . So, we can also say that $\mathbf{v}_{\parallel}$ is the *projection of $\mathbf{v}$ onto l* .

To find $\mathbf{v}_{\perp}$, observe that $\mathbf{v}_{\perp} = \mathbf{v} - \mathbf{v}_{\parallel}$.

Definition 6. Let $\mathbf{v}$ be a vector, and let $\mathbf{d}$ be a non-zero vector. The *projection of $\mathbf{v}$ onto $\mathbf{d}$* is given by

$$\text{proj}_{\mathbf{d}} \mathbf{v} = \left(\frac{\mathbf{v} \cdot \mathbf{d}}{\|\mathbf{d}\|^2} \right) \mathbf{d}$$

Example 6. Find the projection of $\mathbf{v}$, shown below, onto the line given by $y = \frac{1}{2}x - 1$.



Explanation. We begin by finding vectors $\mathbf{v}$ and $\mathbf{d}$. The tail of $\mathbf{v}$ is located at $(-2, -2)$, and the head of $\mathbf{v}$ is at $(2, 4)$. Using the “head-tail” formula we get

$$\mathbf{v} = \begin{bmatrix} 2 - (-2) \\ 4 - (-2) \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

The direction vector for the line $y = \frac{1}{2}x - 1$ is

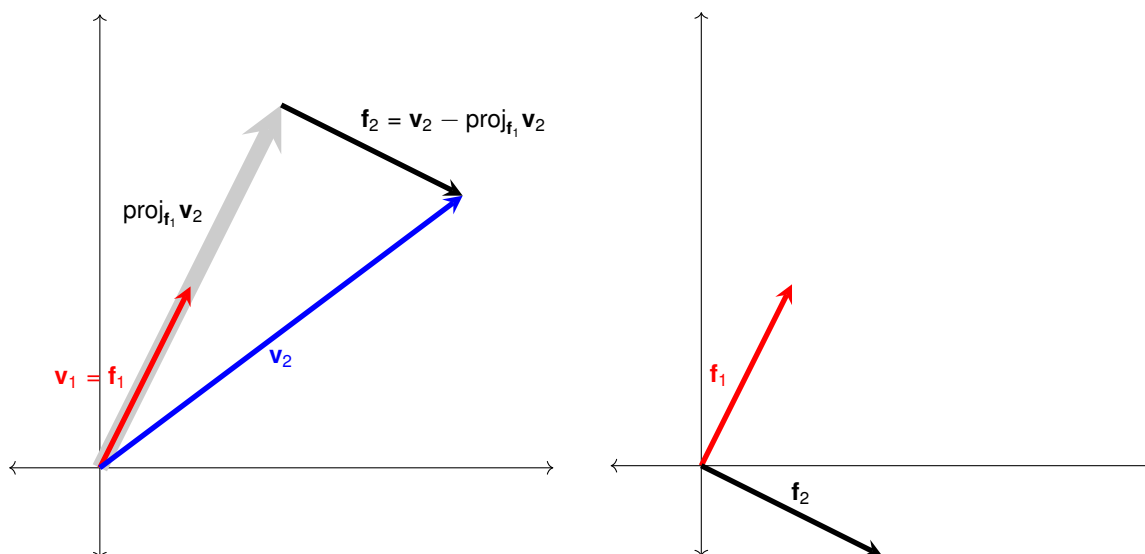
$$\mathbf{d} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

We find that $\mathbf{v} \cdot \mathbf{d} = 14$ and $\|\mathbf{d}\|^2 = 5$. Thus

$$\text{proj}_{\mathbf{d}} \mathbf{v} = \left(\frac{\mathbf{v} \cdot \mathbf{d}}{\|\mathbf{d}\|^2} \right) \mathbf{d} = \frac{14}{5} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 28/5 \\ 14/5 \end{bmatrix}$$

A Visual Guide to Creating an Orthogonal Basis

Given an arbitrary basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ of $\mathbb{R}^2$, let's start building our orthogonal basis, $\{\mathbf{f}_1, \mathbf{f}_2\}$, by setting $\mathbf{f}_1 = \mathbf{v}_1$. To find the next element of our orthogonal basis, consider the orthogonal projection of $\mathbf{v}_2$ onto $\mathbf{f}_1$. (See the figure below.)



Next, let $\mathbf{f}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{f}_1} \mathbf{v}_2$. Observe that $\mathbf{f}_2$ is orthogonal to $\mathbf{f}_1$ (See Theorem ?? of [Orthogonality and Projections](#)). This gives us an orthogonal collection $\mathcal{B} = \{\mathbf{f}_1, \mathbf{f}_2\}$. It is intuitively clear that $\mathbf{f}_1$ and $\mathbf{f}_2$ are linearly independent. Therefore $\mathcal{B}$ is an orthogonal basis of $\mathbb{R}^2$.

The following exploration illustrates this process dynamically.

Exploration 19 Choose an arbitrary basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ of $\mathbb{R}^2$ by dragging the tips of vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ to desired positions. Use the navigation bar at the bottom of the interactive window to go through the steps of constructing an orthogonal basis of $\mathbb{R}^2$.

Access GeoGebra interactives through the online version of this text at

<https://ximera.osu.edu/linearalgebradv3/LinearAlgebraInteractiveIntro>.

We can apply this process to any two-dimensional subset of $\mathbb{R}^n$. The following exploration will guide you through the process of constructing an orthogonal basis for a plane spanned by two arbitrary vectors in $\mathbb{R}^3$.

Exploration 20 Let $W = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$. W is a plane through the origin in $\mathbb{R}^3$. Use the navigation bar at the bottom of the interactive window to go through the steps of constructing an orthogonal basis for W . RIGHT-CLICK and DRAG to rotate the image for a better view.

Access GeoGebra interactives through the online version of this text at

<https://ximera.osu.edu/linearalgebradv3/LinearAlgebraInteractiveIntro>.

In the next exploration, we take the process of constructing an orthogonal basis to the edge of the visual realm and construct an orthogonal basis for $\mathbb{R}^3$.

Exploration 21 In the GeoGebra interactive below $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $\mathbb{R}^3$. Use check boxes to go through the steps for constructing an orthogonal basis starting with the given basis. RIGHT-CLICK and DRAG to rotate the image for a better view.

Access GeoGebra interactives through the online version of this text at

<https://ximera.osu.edu/linearalgebradv3/LinearAlgebraInteractiveIntro>.

What we saw above are low-dimension examples of what is known as *Gram Schmidt Orthogonalization*.

Practice Problems

Problem 22 Find $\text{proj}_{\mathbf{d}}\mathbf{v}$ if $\mathbf{d} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$.

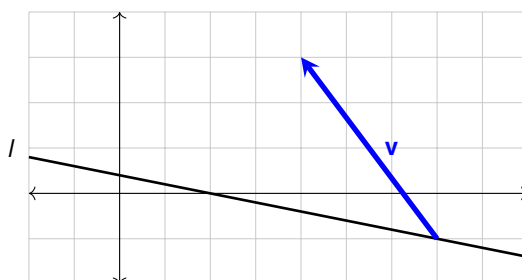
$$\text{proj}_{\mathbf{d}}\mathbf{v} = \begin{bmatrix} \dots \\ \dots \end{bmatrix}$$

Problem 23 Find $\text{proj}_{\mathbf{d}}\mathbf{v}$ if $\mathbf{d} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ -4 \\ 2 \end{bmatrix}$.

$$\text{proj}_{\mathbf{d}}\mathbf{v} = \begin{bmatrix} \dots \\ \dots \\ \dots \end{bmatrix}$$

Problem 24 Show that $\text{proj}_{\mathbf{d}}\mathbf{v}$ does not depend on the length of $\mathbf{d}$ by proving that $\text{proj}_{\mathbf{d}}\mathbf{v} = \text{proj}_{k\mathbf{d}}\mathbf{v}$ for $k \neq 0$. What does this result mean geometrically? Illustrate your response with a diagram.

Problem 25 Find the projection of vector $\mathbf{v}$ onto line l . (If entering answers in decimal form, round to the nearest one hundredth.)



Answer:

$$\begin{bmatrix} \dots \\ \dots \end{bmatrix}$$

Problem 26 Convert the given basis $\mathcal{B}$ of $\mathbb{R}^2$ to an orthogonal basis by following the procedure described in this section with $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

$$\mathcal{B} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

$$\left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} \dots\dots \\ \dots\dots \end{bmatrix} \right\}$$