MATH401/MATH431: Linear Algebra Review

Justin Wyss-Gallifent

September 16, 2021

1 Matrices and Vectors

Definition 1.0.1. A *matrix* is a rectangular array of numbers. We say it is $n \times m$ if it has n rows and m columns, and that its *dimensions* are $n \times m$.

Example 1.1. The following is a 3×5 matrix:

$$\begin{bmatrix} 1 & 0 & -1 & 2 & 3 \\ 0 & 1.2 & 8 & 0 & 1 \\ -10 & 0 & 1 & 1 & 4 \end{bmatrix}$$

$\square$

Definition 1.0.2. A *vector* is an $n \times 1$ matrix.

$\square$

Matrices are generally denoted by upper-case letters A, B, etc. while vectors are generally denoted by lower-case letters with a bar over them or as boldface $\mathbf{b}$, $\mathbf{x}$, etc.

Example 1.2. We might write:

$$A = \begin{bmatrix} 1 & 0 & -1 & 2 & 3 \\ 0 & 1.2 & 8 & 0 & 1 \\ -10 & 0 & 1 & 1 & 4 \end{bmatrix} \text{ and } \mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 3 \end{bmatrix}$$

$\square$

Definition 1.0.3. A matrix is *square* if $m = n$.

$\square$

If a matrix is denoted by A then the entry in row i and column j will typically be denoted a_{ij} or $a_{(i,j)}$.

If a vector is denoted by $\mathbf{b}$ then the entry in row i will typically be denoted by b_i.

Definition 1.0.4. If A is an $n \times m$ matrix and $\mathbf{b}$ is an $m \times 1$ vector then we may define the *product of a matrix and a vector* $A\mathbf{b}$ as the linear combination of columns of A using the entries in $\mathbf{b}$.

In other words if $A = [\mathbf{a_1} \dots \mathbf{a_m}]$ and $\mathbf{b} = [b_1 \dots b_m]^T$ then:

$$A\mathbf{b} = b_1\mathbf{a_1} + \dots + b_m\mathbf{a_m}$$

In more detail:

$$A\mathbf{b} = \begin{bmatrix} a_{11}b_1 + a_{12}b_2 + \dots + a_{1m}b_m \\ a_{21}b_1 + a_{22}b_2 + \dots + a_{2m}b_m \\ \dots \\ a_{n1}b_1 + a_{n2}b_2 + \dots + a_{nm}b_m \end{bmatrix}$$

□

Example 1.3. We have:

$$\begin{aligned}\begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -2 \end{bmatrix}\begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix} &= 5\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 6\begin{bmatrix} 3 \\ 2 \end{bmatrix} + 7\begin{bmatrix} -1 \\ -2 \end{bmatrix} \\ &= \begin{bmatrix} (1)(5)+(3)(6)+(-1)(7) \\ (0)(5)+(2)(6)+(-2)(7) \end{bmatrix} \\ &= \begin{bmatrix} 16 \\ -2 \end{bmatrix}\end{aligned}$$

□

Exercise 1.1. Given the matrix and vector:

$$A = \begin{bmatrix} 4 & -1 & 0 & 2 \\ 5 & 1 & 3 & 2 \\ 0 & 3 & 7 & 3 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 3 \\ 10 \\ -4 \\ 8 \end{bmatrix}$$

Write $A\mathbf{b}$ first as a linear combination and then as a single vector result.

□

Definition 1.0.5. If A is an $n \times m$ matrix and B is an $m \times p$ matrix then if the columns of B are denoted $\mathbf{b_1}, ..., \mathbf{b_p}$ then we may define the *product of two matrices* AB by

$$AB = [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad ... \quad A\mathbf{b}_p]$$

□

Example 1.4. We have:

$$\begin{aligned}\begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -2 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 6 & -2 \\ 7 & 1 \end{bmatrix} &= \left[\begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -2 \end{bmatrix}\begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix} \quad \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -2 \end{bmatrix}\begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}\right] \\ &= \begin{bmatrix} 16 & -7 \\ -2 & -6 \end{bmatrix}\end{aligned}$$

□

Notice that AB may be defined when BA is not (because of the dimensions). Even if both AB and BA are defined they may be different sizes. Even when they're the same size they may be have different entries.

Definition 1.0.6. The *main diagonal of a matrix* A consists of the entries a_{11}, a_{22}, ... , a_{nn}.

$\square$

Definition 1.0.7. The *identity matrix* I_n is the $n \times n$ matrix with 1s on the main diagonal and 0s elsewhere. When the size is clear or implied we simply write I.

$\square$

Definition 1.0.8. If A is a square matrix then the *transpose* of A, denoted A^T, is the matrix whose (i, j) entry equals a_{ji}. That is, it is obtained by reflecting A in the main diagonal.

$\square$

Definition 1.0.9. A square matrix A is symmetric if $A^T = A$.

$\square$

Definition 1.0.10. If A is an $n \times n$ square matrix then A is invertible if there is another $n \times n$ matrix, denoted A^{-1}, such that $AA^{-1} = I$ and $A^{-1}A = I$.

$\square$

Example 1.5. Observe that

$$\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 3 & 7 \\ -2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

so that

$$\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & 7 \\ -2 & 5 \end{bmatrix} \text{ and } \begin{bmatrix} 3 & 7 \\ -2 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}$$

and both matrices are invertible.

$\square$

Most matrix inverses are not nearly this pretty, nor are inverses easy to find.

The exception is the 2×2 case.

Theorem 1.0.1. If

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then A is invertible iff $ad - bc \neq 0$ in which case

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Proof. Omitted. $\square$

Exercise 1.2. Show that the formula above is valid.

Exercise 1.3. Calculate the inverse of the matrix:

$$A = \begin{bmatrix} 5 & -3 \\ 1 & 8 \end{bmatrix}$$

□

Not all matrices are invertible but most are, where most means something rigorous and meaningful.

Definition 1.0.11. If $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ then we define the *dot product* (*inner product*)

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w} = v_1 w_1 + ... + v_n w_n$$

□

Definition 1.0.12. If $\mathbf{v} \in \mathbb{R}^n$ then the *magnitude* (*norm*, *length*) of $\mathbf{v}$ is defined by

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

□

Note: Frequently $||\mathbf{v}||$ is used.

Definition 1.0.13. A *diagonal matrix* is a square matrix A such that $a_{ij} = 0$ for $i \neq j$.

□

2 Determinants

Definition 2.0.1. If A is an $n \times m$ matrix then the *matrix minor* denoted by A_{ij} is the $(n-1) \times (m-1)$ matrix obtained by removing row i and column j from A.

□

Definition 2.0.2. The *determinant of a square matrix* A denoted $\det(A)$ or just $\det A$, or by writing the matrix using vertical lines rather than brackets, is defined recursively as follows:

- If A is 2×2 then
$$\det(A) = a_{11}a_{22} - a_{12}a_{21}$$

- If A is larger than 2×2 then
$$\det(A) = +a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13}) - ... \pm a_{1n}\det(A_{1n})$$

□

Example 2.1. For a 2×2

$$\det \begin{bmatrix} 5 & 3 \\ -2 & 6 \end{bmatrix} = \begin{vmatrix} 5 & 3 \\ -2 & 6 \end{vmatrix} = (5)(6) - (3)(-2) = 36$$

$\square$

Example 2.2. For a 3×3

$$\begin{aligned} \det \begin{bmatrix} 1 & 2 & -3 \\ 0 & 5 & 1 \\ -2 & 4 & 7 \end{bmatrix} &= +1 \det \begin{bmatrix} 5 & 1 \\ 4 & 7 \end{bmatrix} - 2 \det \begin{bmatrix} 0 & 1 \\ -2 & 7 \end{bmatrix} + (-3) \det \begin{bmatrix} 0 & 5 \\ -2 & 4 \end{bmatrix} \\ &= +1(31) - 2(2) + (-3)(10) \\ &= -3 \end{aligned}$$

$\square$

Exercise 2.1. Calculate the determinant of the matrix:

$$A = \begin{bmatrix} 6 & -5 \\ 10 & 7 \end{bmatrix}$$

$\square$

Exercise 2.2. Calculate the determinant of the matrix:

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 5 & 0 & 2 \\ 9 & 11 & 4 \end{bmatrix}$$

$\square$

Theorem 2.0.1. A square matrix A is invertible iff $\det(A) \neq 0$.

Proof. Omitted. $\square$

Example 2.3. The matrix:

$$\begin{bmatrix} 1 & 2 & -3 \\ 0 & 5 & 1 \\ -2 & 4 & 7 \end{bmatrix}$$

has determinant $-3 \neq 0$ and hence is invertible.

$\square$

Exercise 2.3. Are the matrices in the previous exercises invertible?

$\square$

Mathematically speaking since the chances of having $\det(A) = 0$ are very small (if you randomly throw a bunch of numbers into a matrix it's highly unlikely that the determinant will be zero) it is in this sense that we can say that most matrices are invertible since most matrices have nonzero determinant.

3 Systems of Equations

Theorem 3.0.1. A linear system of m equations in the variables $x_1, ..., x_n$ given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_1 \\ ... &= ... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n &= b_m \end{aligned}$$

may be represented by the matrix equation

$$A\mathbf{x} = \mathbf{b}$$

Proof. This is just a rewriting using definitions. □

Example 3.1. The system of equations

$$\begin{aligned} 2x_1 + 3x_2 - 1x_3 &= 7 \\ -1x_1 + 7x_2 + 4x_3 &= -2 \end{aligned}$$

may be rewritten as

$$\begin{bmatrix} 2 & 3 & -1 \\ -1 & 7 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$$

□

Theorem 3.0.2. The matrix equation $A\mathbf{x} = \mathbf{b}$ has either no solutions, one solution, or infinitely many solutions. There is one solution iff A is invertible and in that case the solution is given by $\mathbf{x} = A^{-1}\mathbf{b}$.

Proof. Omitted. □

Example 3.2. The matrix equation

$$\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

has exactly one solution because the matrix is invertible as we saw earlier. The solution is given by

$$\mathbf{x} = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 13 \\ -9 \end{bmatrix}$$

□

4 Linear Independence

Definition 4.0.1. A set of vectors $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ is *linearly independent* if $a_1\mathbf{v}_1 + ... + a_n\mathbf{v}_n = \mathbf{0}$ implies $a_1 = ... = a_n = 0$.

□

As a consequence of this a set of just two vectors is linearly independent iff neither is a multiple of the other.

Example 4.1. The set of vectors

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix} \right\}$$

is a linearly independent set because the vectors are not multiples of one another.

□

A classic way of thinking of linear independence is that it is impossible to write any one of the vectors as a linear combination of the other vectors.

Exercise 4.1. The following set of three vectors is clearly not linearly independent. Show this by finding a linear combination with nonzero coefficients for which the result is $\mathbf{0}$. Hint: It's fairly obvious.

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix} \right\}$$

□

One of the consequences of having a linearly independent set is that if some vector $\mathbf{v}$ is a linear combination of that set then only that specific linear combination works.

Definition 4.0.2. If a set of vectors is not linearly independent then it is *linearly dependent*.

□

A classic way of thinking of linear dependence is that one vector may be written as a linear combination of the others.

Example 4.2. The set of vectors

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

is linearly dependent. Observe that

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

$\square$

Exercise 4.2. From the previous exercise write each vector as a linear combination of the other two.

$\square$

5 Vector Spaces and Bases

Definition 5.0.1. A *vector space* is a nonempty set V of vectors such that the following properties hold:

(a) $\mathbf{0} \in V$.

(b) If $\mathbf{u}, \mathbf{v} \in V$ then $\mathbf{u} + \mathbf{v} \in V$.

(c) $\mathbf{u} \in V$ then $-\mathbf{u} \in V$.

(d) If $\mathbf{u} \in V$ and $c \in \mathbb{R}$ then $c\mathbf{u} \in V$.

□

Note that (a) and (c) actually follow from (b) and (c) together but it's worth listing on its own anyway.

Exercise 5.1. Show that (a) and (c) follow from (b) and (c) together.

□

Note 5.0.1. The formal definition of a vector space is actually a little more abstract and includes rules about which field we're working over, rules of commutativity and associativity, and so on, but if we're a starting with vectors in $\mathbb{R}^n$ and the scalars are in $\mathbb{R}$ then this simplifies the definition immensely.

□

Note 5.0.2. On a related note, technically here we're taking all our vectors from some $\mathbb{R}^n$ and we're constructing a vector space which is a subspace of $\mathbb{R}^n$.

□

Example 5.1. The set of vectors in $\mathbb{R}^3$ with zero in the first and second entries form a vector space. Technically a vector subspace of $\mathbb{R}^3$ but still a vector space in their own right.

□

Definition 5.0.2. Given a set of vectors $S = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ then the *span* of S denoted span(S) is the set of all linear combinations of vectors in S. More rigorously

$$\text{span}(S) = \left\{ a_1\mathbf{v}_1 + ... + a_n\mathbf{v}_n \,\middle|\, a_1, ..., a_n \in \mathbb{R} \right\}$$

□

Example 5.2. If

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix} \right\}$$

Then

$$\text{span}(S) = \left\{ a_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix} \;\middle|\; a_1, a_2 \in \mathbb{R} \right\}$$

$\square$

Theorem 5.0.1. Given a set of vectors $S = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$, the span of S is a vector space.

Proof. Omitted. This is straightfoward - simply show in general that linear combinations of the spanning vectors satisfy the criteria for the definition. $\square$

Definition 5.0.3. If V is a vector space then a *basis* for V is a linearly independent set B of vectors such that $V = \text{span}(B)$.

$\square$

Example 5.3. The set

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

is linearly independent and hence is a basis for $\text{span}(B) = \mathbb{R}^3$. Note that we haven't shown they're linearly independent.

$\square$

Example 5.4. The set

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\}$$

is linearly dependent (the third is the sum of the first two) and hence is not a basis for $\text{span}(B)$. We can still take this span, of course.

$\square$

Essentially a basis for a vector space V is a set of building blocks B such that each vector in V can be written uniquely as a linear combination of vectors in B.

Definition 5.0.4. Given a matrix A, the *column space* of A denoted $\text{col}(A)$ is the span of the columns of A.

$\square$

Theorem 5.0.2. Given a matrix A, $\text{col}(A)$ equals the set of vectors $A\mathbf{x}$.

Proof. This follows from the definition of span and of matrix-vector multiplication, since $A\mathbf{x}$ is (by definition) equal to the linear combination of the columns of A using the weights in $\mathbf{x}$. $\square$

Example 5.5. We have:

$$\text{col}\begin{bmatrix} 1 & 1 \\ 0 & 3 \\ -1 & 2 \end{bmatrix} = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}\right\}$$

$\square$

Theorem 5.0.3. Every vector space has a basis and the number of vectors in a basis of a vector space is independent of the choice of basis. That is, every basis has exactly the same number of vectors as every other basis.

Proof. Omitted. $\square$

Worth noting is that the process of finding a basis for a vector space depends strongly on how the vector space is given in the first place. If it's given as the span of a set of vectors we can check if they're linearly independent (if they are, we're done) and if not then we can iteratively throw out vectors which are linear combinations of previous vectors. For example if we have $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_i\}$ we can start by asking - is $\mathbf{v}_2$ a multiple of $\mathbf{v}_1$ and if so, throw it out. Then - is $\mathbf{v}_3$ a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ (or just $\mathbf{v}_1$ if we threw $\mathbf{v}_2$ out) and if so, throw it out. Then - we keep going until $\mathbf{v}_i$.

Definition 5.0.5. For a vector space V the *dimension* of V denoted $\dim(V)$, is defined as the number of vectors in a basis of V.

Example 5.6. For example

$$\dim\left(\text{col}\begin{bmatrix} 1 & 1 \\ 0 & 3 \\ -1 & 2 \end{bmatrix}\right) = \dim\left(\text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}\right\}\right) = 2$$

$\square$

Note 5.0.3. We can't guarantee that $\dim(\text{span}\{\mathbf{v}_1, ...\mathbf{v}_i\}) = i$ unless we know that the vectors are linearly independent.

$\square$

6 Orthogonality and Orthonormality

Definition 6.0.1. Two vectors are *orthogonal* if their dot product equals zero. A set of vectors $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ is orthogonal if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$.

$\square$

Example 6.1. The set of vectors

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix} \right\}$$

is an orthogonal set of vectors.

$\square$

Definition 6.0.2. A set of vectors $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ is *orthonormal* if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$ and $||\mathbf{v}_i|| = 1$ for all i.

$\square$

Example 6.2. The set of vectors in the previous example is orthonormal if each vector is divided by its magnitude:

$$\left\{ \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \begin{bmatrix} 0 \\ 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}, \begin{bmatrix} -5/\sqrt{30} \\ 2/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix} \right\}$$

is an orthonormal set of vectors.

$\square$

As we have mentioned, every vector space has a basis. In fact every vector space has an orthonormal basis. This can be constructed by taking a regular basis and applying the Gram-Schmmidt process. We'll see later that obtaining an orthonormal basis is usually ideal but often computationally messy since the entries in the vectors are usually not so pretty, as above.

Definition 6.0.3. A square matrix is an *orthogonal matrix* if the column vectors form an orthonormal set. This definition is often a bit confusing (we might prefer an "orthonormal" matrix but that's not the way it goes).

$\square$

Example 6.3. The matrix

$$A = \begin{bmatrix} 1/\sqrt{6} & 0 & -5/\sqrt{30} \\ 2/\sqrt{6} & 1/\sqrt{5} & 2/\sqrt{30} \\ 1/\sqrt{6} & -2/\sqrt{5} & 1/\sqrt{30} \end{bmatrix}$$

is orthogonal.

Theorem 6.0.1. A square matrix A is orthogonal iff $A^T A = AA^T = I$. That is, if $A^T = A^{-1}$.

Proof. Details omitted but this just follows from the definitions of a matrix product as well as vectors being orthogonal and unit vectors. $\square$

Note 6.0.1. Orthogonal matrices are great simply because $A^{-1} = A^T$ and so the inverse is really convenient. In addition they're usual from a computational standpoint because they lead to handy things like QR decompositions ($A = QR$ where Q is orthogonal and R is upper triangular) which are useful for least-squares problems and calculating eigenvalues and eigenvectors.

$\square$

7 Eigenvalues and Eigenvectors

Definition 7.0.1. If A is an $n \times n$ matrix then λ is an *eigenvalue* for A if there is a nonzero vector $\mathbf{v}$ such that $A\mathbf{v} = \lambda\mathbf{v}$. In this case $\mathbf{v}$ is the corresponding *eigenvector* and together they form an *eigenpair* $(\lambda, \mathbf{v})$.

$\square$

Note 7.0.1. Notice that any nonzero multiple of an eigenvector is also an eigenvector.

$\square$

Definition 7.0.2. The set of all eigenvectors for a given eigenvalue along with the zero vector $\bar{0}$ is called the *eigenspace* of that eigenvalue.

$\square$

Theorem 7.0.1. The eigenspace of an eigenvalue is a vector space

Proof. Just go through the calculations of the necessities. $\square$

Note 7.0.2. Notice that technically $\mathbf{0}$ is not an eigenvector but it is included in the eigenspace in order to make it a vector space.

$\square$

Example 7.1. The matrix

$$A = \begin{bmatrix} 4 & 7 \\ 1 & -2 \end{bmatrix}$$

has two eigenvalues:

- One is $\lambda_1 = 5$ with eigenvector $\begin{bmatrix} 7 \\ 1 \end{bmatrix}$ because:

$$\begin{bmatrix} 4 & 7 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 7 \\ 1 \end{bmatrix} = \begin{bmatrix} 35 \\ 5 \end{bmatrix} = 5 \begin{bmatrix} 7 \\ 1 \end{bmatrix}$$

  Thus any nonzero multiple of $\begin{bmatrix} 7 \\ 1 \end{bmatrix}$ is also an eigenvectors and the set of all such nonzero multiples, along with $\bar{0}$, form the eigenspace for $\lambda_1 = 5$.

- The other eigenvalue is $\lambda_2 = -3$ with eigenvector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. This can be easily checked and followed up in the same way as above.

We haven't addressed in this example how we know that there are only two, or that we know that each has a one-dimensional eigenspace. We won't address all of the associated details but just hit the main points now.

□

Definition 7.0.3. If A is an $n \times n$ matrix then the *characteristic polynomial* of A denoted $\mathrm{char}(A)$, is defined by

$$\mathrm{char}(A) = \det(\lambda I - A)$$

Note: Some authors define the characteristic polynomial to be:

$$\mathrm{char}(A) = \det(A - \lambda I)$$

These two definitions differ by a factor of $(-1)^n$.

□

Theorem 7.0.2. The eigenvalues of a matrix A are the roots of the characteristic polynomial.

Proof. The value λ is an eigenvalue iff there is some nonzero $\mathbf{v}$ with $A\mathbf{v} = \lambda\mathbf{v}$, which may be rewritten $(A - \lambda I)\mathbf{v} = 0$. This has a nonzero solution (such a $\mathbf{v}$ exists) if and only if $\det(A - \lambda I) = 0$ which is precisely if λ is a root of the characteristic polynomial. □

Example 7.2. Revisiting the matrix

$$A = \begin{bmatrix} 4 & 7 \\ 1 & -2 \end{bmatrix}$$

We see the the characteristic polynomial is:

$$\begin{aligned} \mathrm{char}\begin{bmatrix} 4 & 7 \\ 1 & -2 \end{bmatrix} &= \det\begin{bmatrix} \lambda - 4 & 7 \\ 1 & \lambda + 2 \end{bmatrix} \\ &= (\lambda + 2)(\lambda - 4) - 7 \\ &= \lambda^2 - 2\lambda - 15 \\ &= (\lambda - 5)(\lambda + 3) \end{aligned}$$

This has roots, and hence the matrix has eigenvalues, $\lambda = 5$ and $\lambda = -3$.

□

Example 7.3. If we have:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 2 & 3 \\ 0 & 4 & 3 \end{bmatrix}$$

then

$$\begin{aligned} \text{char}(A) &= \det(\lambda I - A) \\ &= \det \begin{bmatrix} \lambda - 1 & -2 & 1 \\ -1 & \lambda - 2 & -3 \\ 0 & -4 & \lambda - 3 \end{bmatrix} \\ &= \lambda^3 - 6\lambda^2 - 3\lambda + 16 \end{aligned}$$

The eigenvalues of the matrix are roots of this, $\lambda_1 \approx 6.0593$, $\lambda_2 \approx -1.6549$ and $\lambda_3 \approx 1.5956$.

□

Note 7.0.3. One handy trick to remember is that when A is 2×2 then $\text{char}(A) = \lambda^2 - \text{Tr}(A)\lambda + \text{Det}(A)$, where Tr is the trace (sum of the main diagonal) and Det is of course the determinant.

□

Note 7.0.4. For a triangular matrix the eigenvalues are simply the values on the main diagonal.

□

Since the characteristic polynomial has degree n this tells us that an $n \times n$ matrix has n eigenvalues, counting multiplicity, some of which may be complex.

Example 7.4. Consider the matrix:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$$

This matrix has two eigenvalues, $\lambda = 1$ (with multiplicity 2) and $\lambda = 2$ (with multiplicity 1).

□

Exercise 7.1. Find the characteristic polynomials and eigenvalues of the following matrices:

(a) $A = \begin{bmatrix} 4 & -5 \\ 1 & -2 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}$

(c) $A = \begin{bmatrix} -1 & -2 \\ 2 & -1 \end{bmatrix}$

$\square$

Once the eigenvalues have been found then for each eigenvalue λ the corresponding eigenvectors (the eigenspace) are found by solving the matrix equation:

$$\begin{aligned} A\mathbf{v} &= \lambda\mathbf{v} \\ (\lambda I - A)\mathbf{v} &= \mathbf{0} \end{aligned}$$

This is a standard linear algebra procedure we won't review. Usually we'll do this in software.

Theorem 7.0.3. If A is an $n \times n$ matrix and λ is an eigenvalue with eigenspace V then the dimension of the eigenspace is less than or equal to the multiplicity of λ as a root of the characteristic polynomial.

Proof. Omitted. $\square$

Example 7.5. The matrix

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 10 & 12 & -30 \\ 5 & 5 & -13 \end{bmatrix}$$

Has characteristic polynomial

$$\lambda^3 - \lambda^2 - 8\lambda + 12 = (\lambda - 2)^2(\lambda + 3)$$

Consequently the eigenspace for $\lambda_1 = 2$ has dimension either 1 or 2. In this case it's 2 but that's a bit more work to show.

$\square$

Consider then the overall picture. For an $n \times n$ matrix A the multiplicities of the eigenvalues (as roots of the characteristic polynomial) add up to n. For each eigenvalue the dimension of the corresponding eigenspace is less than or equal to the multiplicity. Consequently then the sum of the dimensions of the eigenspaces is less than or equal to n.

8 Diagonalizable Matrices

Definition 8.0.1. An $n \times n$ matrix A is *diagonalizable* if there exists an $n \times n$ invertible matrix P and an $n \times n$ diagonal martix D such that

$$A = PDP^{-1}$$

.

$\square$

Theorem 8.0.1. An $n \times n$ matrix A is diagonalizable iff the dimension of each eigenspace equals the multiplicity of the corresponding eigenvalue in the characteristic polynomial. In other words if the sum of the dimensions of the eigenspaces equals n.

Proof. Omitted. $\square$

From a computational standpoint diagonalizable matrices are nice to work with because we can raise them to powers easily, since $A^k = \left(PDP^{-1}\right)^k = PD^kP^{-1}$. This is less likely to have computational issues than doing A^k itself, iteratively.

Theorem 8.0.2. If A is diagonalizable then the invertible matrix P is formed using the eigenvectors of A and the diagonal matrix D is formed using the eigenvalues of A. The eigenvector in column i corresponds to the eigenvalue in column i.

Proof. Omitted. $\square$

Note 8.0.1. Note that there are a variety of ways we can do this as long as the entries in D and the columns of P correspond correctly. Hence the diagonalization of a matrix is not unique.

$\square$

Example 8.1. If

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 10 & 12 & -30 \\ 5 & 5 & -13 \end{bmatrix}$$

Then

$$A = PDP^{-1}$$

where

$$P = \begin{bmatrix} 0 & 0 & 0.4016 \\ -0.8944 & -0.9487 & -0.9006 \\ -0.4472 & -0.3162 & -0.1663 \end{bmatrix}$$

$$D = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

In this case the first column of P is an eigenvector corresponding to the eigenvalue $\lambda_1 = -3$ and the second and third columns of P are eigenvectors corresponding to the eigenvalue $\lambda_2 = 2$ which has multiplicity 2 and for which the dimension of the eigenspace is also 2.

$\square$

Definition 8.0.2. An $n \times n$ matrix A is *orthogonally diagonalizable* if there exists an $n \times n$ orthogonal matrix Q and an $n \times n$ diagonal matrix D such that

$$A = QDQ^T$$

$\square$

Theorem 8.0.3. A matrix A is orthogonally diagonalizable iff it is symmetric.

Proof. Omitted. $\square$

Example 8.2. The matrix

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 3 & 4 \end{bmatrix}$$

is symmetric hence orthogonally diagonalizable.

$\square$

9 Linear Transformations

Definition 9.0.1. If $\mathbb{R}^n$ and $\mathbb{R}^m$ are vector spaces then a transformation (mapping, function)

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

is *linear* if for all scalars α, β and vectors $\mathbf{v}, \mathbf{w}$ we have:

$$T(\alpha\mathbf{v} + \beta\mathbf{w}) = \alpha T(\mathbf{v}) + \beta T(\mathbf{w})$$

$\square$

A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is often written as:

$$T(x_1, ..., x_n) = (y_1, ..., y_m)$$

for convenience but really means:

$$T\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = \left(\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}\right)$$

Exercise 9.1. Determine whether the following transformations are linear or not. Do this by straightforward computation. If the transformation is not linear give specific examples which demonstrate this.

(a) $T(x_1, x_2) = (x_1 + 2x_2, x_1, -3x_2)$

(b) $T(x_1, x_2, x_3) = (-x_1, x_1 + x_2)$

(c) $T(x_1, x_2) = (x_1 + 2, 3x_2)$

$\square$

Theorem 9.0.1. If a transformation T is linear then $T(\mathbf{0}) = \mathbf{0}$.

Proof. This is obvious since for any nonzero $\mathbf{v}$ we have $T(\mathbf{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \mathbf{0}$. $\square$

Theorem 9.0.2. A transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ is linear if there is some $m \times n$ matrix M such that for all $\mathbf{v}$ we have $T(\mathbf{v}) = M\mathbf{v}$. Moreover if this is the case then:

$$A = [T(\mathbf{e_1}) \ldots T(\mathbf{e_n})]$$

Proof. Omitted. $\square$

This theorem is computationally important because it is constructive, allowing us to explicitly construct the transformation matrix for a linear transformation.

Example 9.1. It is a fact that the transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ given by $T(x_1, x_2) = (2x_1, x_2 - x_1, 3x_1)$ is linear. To find the corresponding 3×2 matrix we check:

$$\begin{aligned} T(\mathbf{e_1}) &= T(1, 0) = (1, -1, -3) \\ T(\mathbf{e_2}) &= T(0, 1) = (0, 1, 0) \end{aligned}$$

Thus the transformation is represented by the matrix:

$$A = [T(\mathbf{e_1})\ T(\mathbf{e_2})] = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ -3 & 0 \end{bmatrix}$$

In other words for all $\mathbf{v} \in \mathbb{R}^2$ we have:

$$T(\mathbf{v}) = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ -3 & 0 \end{bmatrix} \mathbf{v}$$

$\square$

Exercise 9.2. For the transformations in the previous exercise which are linear, determine the corresponding matrices.

It's often difficult to check the explicit definition when checking if a given transformation is linear. Another roundabout way to do this is as follows:

(i) Assume it is and find the matrix which represents it.

(ii) Check if the matrix does what the transformation does.

(iii) If so then it's linear. If not then we have a contradiction and it's not.

Exercise 9.3. Use the above process to check if $T(x_1, x_2) = (5x_1+7x_2, 5x_1-x_2)$ is linear.

$\square$

Exercise 9.4. Use the above process to check if $T(x_1, x_2) = (x_1^2 x_2, x_1 - x_2)$ is linear.

$\square$