Similarity
We start with a linear transformation \( L: V \rightarrow W \) and with bases \( X = x_1, \ldots, x_n \) of \( V \) and \( Y = y_1, \ldots, y_m \) of \( W \). Recall that every vector \( v \in V \) can be written in exactly one way as a linear combination of the basis vectors of \( X \):
\[
v = c_1 x_1 + \cdots + c_n x_n,
\]
where all \( c_i \in \mathbb{R} \); we call the \( n \times 1 \) column vector \((c_1, \ldots, c_n)^\top\) the coordinate list of \( v \) wrt \( X \). Of course, every vector \( w \in W \) has a coordinate list wrt \( Y \).

Linear transformations defined on \( \mathbb{R}^n \) are easy to describe.

**Theorem 0.1**
If \( L: \mathbb{R}^n \rightarrow \mathbb{R}^m \) is a linear transformation, then there exists a unique \( m \times n \) matrix \( A \) such that
\[
L(y) = Ay
\]
for all \( y \in \mathbb{R}^n \) (here, \( y \) is an \( n \times 1 \) column matrix and \( Ay \) is matrix multiplication).

**Proof.** If \( e_1, \ldots, e_n \) is the standard basis of \( \mathbb{R}^n \) and \( e'_1, \ldots, e'_m \) is the standard basis of \( \mathbb{R}^m \), define \( A = [a_{ij}] \) to be the matrix whose \( j \)th column is the coordinate list of \( L(e_j) \). If \( S: \mathbb{R}^n \rightarrow \mathbb{R}^m \) is defined by \( S(y) = Ay \), then \( S = L \) because both agree on a basis: \( L(e_j) = \sum_i a_{ij} e_i = Ae_j \). Uniqueness of \( A \) follows from the fact that if \( L(y) = By \) for all \( y \), then \( Be_j = L(e_j) = Ae_j \) for all \( j \); that is, the columns of \( A \) and \( B \) are the same.

Theorem 0.1 establishes the connection between linear transformations and matrices, and the definition of matrix multiplication arises from applying this construction to the composite of two linear transformations.

**Definition**
Let \( X = v_1, \ldots, v_n \) be a basis of \( V \) and let \( Y = w_1, \ldots, w_m \) be a basis of \( W \). If \( L: V \rightarrow W \) is a linear transformation, then the **matrix of \( L \)** is the \( m \times n \) matrix \( A = [a_{ij}] \) whose \( j \)th column \( a_{1j}, a_{2j}, \ldots, a_{mj} \) is the coordinate list of \( L(v_j) \) wrt \( Y \): 
\[
L(v_j) = \sum_{i=1}^m a_{ij} w_i = a_{1j} w_1 + a_{2j} w_2 + \cdots + a_{mj} w_m.
\]

Since the matrix \( A \) depends on the choice of bases \( X \) and \( Y \), we will write 
\[
A = Y[L]_X
\]
when it is necessary to display them. Varying any of them most likely changes the matrix.

Consider the important special case \( L_A : \mathbb{R}^n \to \mathbb{R}^m \) given by \( L_A(y) = Ay \), where \( A \) is an \( m \times n \) matrix and \( y \) is an \( n \times 1 \) column vector. If \( E = e_1, \ldots, e_n \) and \( E' = e'_1, \ldots, e'_m \) are the standard bases of \( \mathbb{R}^n \) and \( \mathbb{R}^m \), respectively, then the definition of matrix multiplication says that \( L_A(e_j) = Ae_j \) is the \( j \)th column of \( A \). But

\[
Ae_j = a_{1j}e'_1 + a_{2j}e'_2 + \cdots + a_{mj}e'_m;
\]

that is, the coordinates of \( L_A(e_j) = Ae_j \) with respect to the basis \( e'_1, \ldots, e'_m \) are \((a_{1j}, \ldots, a_{mj})\). Therefore, the matrix associated to \( L_A \) is the original matrix \( A \):

\[
E'[L_A]E = A.
\]

In case \( V = W \), we often let the bases \( X = v_1, \ldots, v_n \) and \( Y = w_1, \ldots, w_m \) coincide. If \( 1_V : V \to V \), given by \( v \mapsto v \), is the identity linear transformation, then \( X[1_V]X \) is the \( n \times n \) identity matrix \( I \), which has 1’s on the diagonal and 0’s elsewhere. On the other hand, if \( X \) and \( Y \) are different bases, then \( Y[1_V]X \) is not the identity matrix. The matrix \( Y[1_V]X \) is called the transition matrix from \( X \) to \( Y \); its columns are the coordinate lists of the \( v \)'s with respect to the \( w \)'s.

**Example**

1. Let \( X = e_1, e_2 = (1,0), (0,1) \) be the standard basis of \( \mathbb{R}^2 \). If \( L : \mathbb{R}^2 \to \mathbb{R}^2 \) is rotation by 90°, then \( L : e_1 \mapsto e_2 \) and \( e_2 \mapsto -e_1 \). Hence, the matrix of \( L \) relative to \( X \) is

\[
X[L]X = \begin{bmatrix}
0 & -1 \\
1 & 0
\end{bmatrix}.
\]

If we re-order \( X \) to obtain the new basis \( Y = \eta_1, \eta_2 \), where \( \eta_1 = e_2 \) and \( \eta_2 = e_1 \), then \( L(\eta_1) = L(e_2) = -e_1 = -\eta_2 \) and \( L(\eta_2) = L(e_1) = e_2 = \eta_1 \). The matrix of \( L \) relative to \( Y \) is

\[
Y[L]Y = \begin{bmatrix}
0 & 1 \\
-1 & 0
\end{bmatrix}.
\]

2. Let \( L : \mathbb{R}^2 \to \mathbb{R}^2 \) be rotation by \( \alpha \) degrees about the origin. Take \( X = Y \) to be the standard basis \( e_1 = (1,0), e_2 = (0,1) \). Now \( L(e_1) = (\cos \alpha, \sin \alpha) \) and \( L(e_2) = (\cos(90 + \alpha), \sin(90 + \alpha)) = (-\sin \alpha, \cos \alpha) \). Thus the matrix of \( L \) wrt standard basis is

\[
X[L]X = \begin{bmatrix}
\cos \alpha & -\sin \alpha \\
\sin \alpha & \cos \alpha
\end{bmatrix}.
\]
3. If $L: \mathbb{C} \to \mathbb{C}$ is complex conjugation, and if we choose the bases $X = Y = 1, i$, then the matrix is

$$X[L]X = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$ 

The next theorem shows where the definition of matrix multiplication comes from: the product of two matrices is the matrix of a composite.

**Theorem 0.2**

Let $L: V \to W$ and $S: W \to U$ be linear transformations. Choose bases $X = x_1, \ldots, x_n$ of $V$, $Y = y_1, \ldots, y_m$ of $W$, and $Z = z_1, \ldots, z_\ell$ of $U$. Then


where the product on the right is matrix multiplication.

**Proof.** Let $y[L]X = [a_{ij}]$, so that $L(x_j) = \sum_p a_{pj} y_p$, and let $z[S]Y = [b_{qp}]$, so that $S(y_p) = \sum_q b_{qp} z_q$. Then

$$S \circ L(x_j) = S(L(x_j)) = S(\sum_p a_{pj} y_p)$$

$$= \sum_p a_{pj} S(y_p) = \sum_p \sum_q a_{pj} b_{qp} z_q = \sum_q c_{qj} z_q,$$

where $c_{qj} = \sum_p b_{qp} a_{pj}$. Therefore,


**Corollary 0.3**

Matrix multiplication is associative.

**Proof.** Let $A$ be an $m \times n$ matrix, let $B$ be an $n \times p$ matrix, and let $C$ be a $p \times q$ matrix. There are linear transformations

$$\mathbb{R}^q \overset{L_C}{\to} \mathbb{R}^p \overset{L_B}{\to} \mathbb{R}^n \overset{L_A}{\to} \mathbb{R}^m$$


Then

$$[L_A \circ (L_B \circ L)] = [L_A][L_B \circ L_C] = [L_A][L_B][L_C] = A(BC).$$
On the other hand,

\[(L_A \circ L_B) \circ L_C = (L_A \circ L_B)(L_C) = ((L_A)(L_B))(L_C) = (AB)C.\]

Since composition of functions is associative, \(L_A \circ (L_B \circ L_C) = (L_A \circ L_B) \circ L_C\), and so

\[A(BC) = [L_A \circ (L_B \circ L_C)] = [(L_A \circ L_B) \circ L_C] = (AB)C.\]

The connection with composition of linear transformations is the real reason why matrix multiplication is associative.

**Corollary 0.4**

Let \(L: V \to W\) be a linear transformation of vector spaces \(V\) and \(W\), and let \(X\) and \(Y\) be bases of \(V\) and \(W\), respectively. If there is a linear transformation \(T: W \to V\) with \(L \circ T = 1_V\), then the matrix of \(T\) is the inverse of the matrix of \(L\):

\[x[T]_Y = (y[L]_X)^{-1}.\]

**Proof.** We have \(I = y[1_W]_Y = (y[L]_X)(x[T]_Y)\), and so \(I = x[1_V]_X = (x[T]_Y)(y[L]_X)\).

The next corollary determines all the matrices arising from the same linear transformation as we vary bases. Here we take the same basis fore and aft: \(X = Y\).

**Corollary 0.5**

Let \(L: V \to V\) be a linear transformation on a vector space \(V\). If \(X\) and \(Y\) are bases of \(V\), then there is a nonsingular matrix \(P\) (namely, the transition matrix \(P = y[1_V]_X\)) with entries in \(\mathbb{R}\) so that

\[y[L]_Y = P(x[L]_X)P^{-1}.\]

Conversely, if \(B = PAP^{-1}\), where \(B, A,\) and \(P\) are \(n \times n\) matrices with \(P\) nonsingular, then there is a linear transformation \(L: \mathbb{R}^n \to \mathbb{R}^n\) and bases \(X\) and \(Y\) of \(\mathbb{R}^n\) such that \(B = y[L]_Y\) and \(A = x[L]_X\).

**Proof.** The first statement follows from Theorem ?? and associativity:

\[y[L]_Y = y[1_V L_1 V]_Y = (y[1_V]_X)(x[L]_X)(x[1_V]_Y).\]

Set \(P = y[1_V]_X\) and note that Corollary ?? gives \(P^{-1} = x[1_V]_Y\).
The converse is a bit longer.

Definition
Two $n \times n$ matrices $B$ and $A$ with entries in $\mathbb{R}$ are similar if there is a nonsingular matrix $P$ with entries in $\mathbb{R}$ such that $B = PAP^{-1}$.

Thus, two matrices arise from the same linear transformation on a vector space $V$ (from different choices of bases) if and only if they are similar.