Chapter 3: Matrices - Elementary Theory

A matrix is a rectangular array of scalars.

\[
A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\]

(3-1)

This matrix has \( m \) rows and \( n \) columns. Often, we use the notation \( \{a_{ij}\} \) to denote a matrix.

**Representing a Linear Transformation by a Matrix**

As given in Chapter 2, the definition of a linear transformation \( T : U \to V \) is somewhat abstract. Once bases for spaces \( U \) and \( V \) are specified, a matrix can be used to provide a definitive representation for a linear transformation.

Let \( U \) and \( V \) be vector spaces, over the same field \( F \), of dimension \( n \) and \( m \), respectively. Let \( \alpha_1, \alpha_2, \ldots, \alpha_n \) be a basis of \( U \), and \( \beta_1, \beta_2, \ldots, \beta_m \) be a basis of \( V \). Let \( T : U \to V \) be a linear transformation. In terms of these bases, we can write

\[
T(\tilde{\alpha}_j) = \sum_{i=1}^{m} a_{ij} \tilde{\beta}_i , \quad 1 \leq j \leq n .
\]

(3-2)

This equation *defines* an \( m \times n \) matrix \( A \) that describes \( T : U \to V \) with respect to the given bases. As shown below, this matrix can be used as a representation for transformation \( T \). Often, we use the symbolic equation

\[
T[\tilde{\alpha}_1 \mid \tilde{\alpha}_2 \mid \cdots \mid \tilde{\alpha}_n] = [\tilde{\beta}_1 \mid \tilde{\beta}_2 \mid \cdots \mid \tilde{\beta}_m]A
\]

(3-3)

instead of the algebraic equation (3-2). The quantity \( [\tilde{\alpha}_1 \mid \tilde{\alpha}_2 \mid \cdots \mid \tilde{\alpha}_n] \) is itself an \( n \times n \) matrix with columns made from the basis vectors \( \tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_n \).
Given any $\tilde{X} \in U$, the matrix $A$ can be used to compute $\tilde{Y} = T(\tilde{X})$ by using the coordinate vectors that represent $\tilde{X}$ and $\tilde{Y}$ with respect the bases. For $\tilde{X} \in U$, we can write

$$\tilde{X} = \sum_{j=1}^{n} x_j \tilde{\alpha}_j,$$

(3-4)

where $x_i$, $1 \leq i \leq n$, are the coordinates for vector $\tilde{X}$. Now, use (3-4) to compute

$$\tilde{Y} = T(\tilde{X}) = T\left( \sum_{j=1}^{n} x_j \tilde{\alpha}_j \right) = \sum_{j=1}^{n} x_j T(\tilde{\alpha}_j)$$

(3-5)

Use (3-2) in (3-5) to obtain

$$\tilde{Y} = \sum_{j=1}^{n} x_j \left( \sum_{i=1}^{m} a_{ij} \tilde{\beta}_i \right) = \sum_{i=1}^{m} \left( \sum_{j=1}^{n} a_{ij} x_j \right) \tilde{\beta}_i = \sum_{i=1}^{m} y_i \tilde{\beta}_i,$$

(3-6)

where

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \ 1 \leq i \leq m,$$

(3-7)

are the coordinates of $\tilde{Y}$.

The algebraic derivation of (3-7) has a somewhat symbolic counterpart. First, recall that $\tilde{X}$ can be represented symbolically as
where \([x_1 \ x_2 \ ... \ x_n]^T\) is a coordinate vector for the vector \(\tilde{X}\) (remember: coordinates are in italics and vector components are not italicized). Note that the \(x_k\) are just scalars, and apply \(T\) to (3-8) to obtain

\[
T(\tilde{X}) = T[\tilde{\alpha}_1 \ | \ \tilde{\alpha}_2 \ | \ ... \ | \ \tilde{\alpha}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \tag{3-9}\]

Now, use (3-3) in (3-9) to obtain

\[
T(\tilde{X}) = [\tilde{\beta}_1 \ | \ \tilde{\beta}_2 \ | \ ... \ | \ \tilde{\beta}_m]A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [\tilde{\beta}_1 \ | \ \tilde{\beta}_2 \ | \ ... \ | \ \tilde{\beta}_m] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \tag{3-10}\]

where

\[
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{3-11}\]
is the desired relationship between the coordinate vectors used to represent $\tilde{X}$ and $\tilde{Y}$. Equations (3-7) and (3-11) state equivalent results: the rows of $m \times n$ matrix $A$ are used to express the coordinate vector for $\tilde{Y}$ in terms of the coordinate vector for $\tilde{X}$.

**Example**

Let $U = V = \mathbb{R}^2$. Let $\vec{\alpha}_1 = [1 \ 0]^T$ and $\vec{\alpha}_2 = [0 \ 1]^T$ be a common basis for $U$, $V$. Define $T$ as the transformation that rotates vectors counterclockwise by $\pi/2$ radians. Note that $T$ does not change the magnitude of the vector. Observe that

\[
T \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{and} \quad T \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} = -1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}
\]

so that

\[
A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}
\]

is a matrix representation for the rotation transformation.

As shown above, there is a well defined relationship between a linear transformation and the matrix used to represent the transformation. Hence, almost all of the definitions, given in Chapter 2, dealing with linear transformations have counterparts when dealing with matrices. For example, let $T : U \rightarrow V$ be represented by matrix $A$. Then the vectors in $R(T)$ are represented by coordinate vectors in the subspace
R(A) = \[ \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \]
where \([x_1 \ x_2 \ \cdots \ x_n]^T i\) is a coordinate vector in U \hspace{1cm} (3-12)

Clearly, R(A) is the span of the columns of A. Likewise, the vectors in K(T) are represented by coordinate vectors in the subspace

\[
K(A) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}: A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \hspace{1cm} (3-13)
\]

**Matrices: Some Special Kinds**

In our work (and most applications), the elements of a matrix belong to the field of real numbers, denoted as \(\mathbb{R}\), or the field of complex numbers, denoted as \(\mathbb{C}\). If A is an \(m \times n\) matrix over the real numbers we say \(A \in \mathbb{R}^{m \times n}\). Likewise, if A is an \(m \times n\) matrix over the complex numbers we say \(A \in \mathbb{C}^{m \times n}\).

An \(n \times n\) matrix \(P \in \mathbb{R}^{n \times n}\) is said to be *orthogonal* if \(P^T = P^{-1}\) so that \(P^T P = P^{-1} P = I\), the \(n \times n\) identity matrix. The columns (and rows) of Q are orthogonal. Let

\[
Q = [\tilde{q}_1 \mid \tilde{q}_2 \mid \cdots \mid \tilde{q}_n], \hspace{1cm} (3-14)
\]

then Q is orthogonal if and only if

\[
\langle \tilde{q}_i, \tilde{q}_j \rangle = \tilde{q}_j^T \tilde{q}_i = 1 \hspace{0.5cm} i = j
\]

\[
= 0 \hspace{1.5cm} i \neq j \hspace{1cm} (3-15)
\]
An \( n \times n \) matrix \( U \in \mathbb{C}^{n \times n} \) is said to be **unitary** if \( U^* U = U U^* = I \), the \( n \times n \) identity matrix (here, the \( * \) denotes complex conjugate transpose). The columns (and rows) of \( U \) are orthogonal. Let

\[
U = \begin{bmatrix}
\mathbb{u}_1 & \mathbb{u}_2 & \cdots & \mathbb{u}_n
\end{bmatrix},
\]

then \( U \) is unitary if and only if

\[
\langle \tilde{u}_i, \tilde{u}_j \rangle = \tilde{u}_j^* \tilde{u}_i = 1 \quad i = j
\]

\[
= 0 \quad i \neq j
\]

A square matrix \( A \in \mathbb{R}^{n \times n} \) is said to be **symmetric** if \( A^T = A \). Symmetric matrices show up in many applications. An important attribute of symmetric matrices is that they always have real-valued eigenvalues. Also, they are orthogonal similar to a diagonal matrix containing the eigenvalues of \( A \). That is, there always exists an **orthogonal matrix** \( P \in \mathbb{R}^{n \times n} \) (i.e., \( P^T = P \) or \( P^T P = I \)) such that \( P^T A P = D \), where \( D \in \mathbb{R}^{n \times n} \) is a diagonal matrix containing the eigenvalues of \( A \). The columns of \( P \) are the eigenvectors of \( A \), and they are orthonormal.

Only a few applications deal with complex-valued, symmetric matrices (As a general observation, linear algebra books do not consider complex-valued symmetric matrices). However, for the complex number field, the counterpart of the symmetric matrix is the **Hermitian** matrix. A square matrix \( A \in \mathbb{C}^{n \times n} \) is said to be **Hermitian** if \( A^* = A \). **Hermitian** matrices show up in many applications. An important attribute of **Hermitian** matrices is that they always have real-valued eigenvalues. Also, they are unitary similar to a diagonal matrix containing the eigenvalues of \( A \). That is, there always exists an **unitary matrix** \( U \in \mathbb{C}^{n \times n} \) (\( U^* = U \) or \( U^T U = I \)) such that \( U^T A U = U^* A U = D \), where \( D \in \mathbb{C}^{n \times n} \) is a diagonal matrix containing the eigenvalues of \( A \). The columns of \( U \) are the eigenvectors of \( A \), and they are orthonormal.
Vector/Transformation Representation With Respect to Orthonormal Basis

Let \( \tilde{\alpha}_i, 1 \leq i \leq n, \) be an orthonormal basis of vector space \( U. \) Then

\[
\langle \tilde{\alpha}_i, \tilde{\alpha}_j \rangle = \tilde{\alpha}_j^* \tilde{\alpha}_i = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}
\]

(3-18)

Theorem 3.1

Let \( \tilde{X} \in U \) be arbitrary. Then we can make the representation

\[
\tilde{X} = \langle \tilde{X}, \tilde{\alpha}_1 \rangle \tilde{\alpha}_1 + \langle \tilde{X}, \tilde{\alpha}_2 \rangle \tilde{\alpha}_2 + \cdots + \langle \tilde{X}, \tilde{\alpha}_n \rangle \tilde{\alpha}_n = \sum_{k=1}^{n} \langle \tilde{X}, \tilde{\alpha}_k \rangle \tilde{\alpha}_k
\]

Proof: \( \left( \tilde{X} - \sum_{k=1}^{n} \langle \tilde{X}, \tilde{\alpha}_k \rangle \tilde{\alpha}_k, \tilde{\alpha}_j \right) = \langle \tilde{X}, \tilde{\alpha}_j \rangle - \langle \tilde{X}, \tilde{\alpha}_j \rangle = 0 \) for \( 1 \leq j \leq n. \)

Since \( \tilde{X} \) and \( \sum_{k=1}^{n} \langle \tilde{X}, \tilde{\alpha}_k \rangle \tilde{\alpha}_k \) have equal coordinates, they must be the same vector.♥

This theorem gives us a convenient and useful representation for an arbitrary vector. Also, the proof illustrates an important technique for showing that two vectors are equal. That is, to show that two vectors are equal, it is sufficient to show that they have the same coordinates (given an underlying basis).

Theorem 3.2

Let \( T : U \to V. \) Let \( \tilde{\alpha}_i, 1 \leq j \leq n, \) and \( \tilde{\beta}_i, 1 \leq i \leq m, \) be orthonormal bases on \( U \) and \( V, \) respectively. Let \( m \times n \) matrix \( A = \{ a_{ij} \} \) represent \( T \) with respect to these bases. Then

\[
a_{ij} = \langle T(\tilde{\alpha}_j), \tilde{\beta}_i \rangle \quad 1 \leq i \leq m, \quad 1 \leq j \leq n,
\]

(3-19)

Proof: By Theorem 3.1 we have
\[ T(\bar{\alpha}_1) = \langle T(\bar{\alpha}_1), \bar{\beta}_1 \rangle \bar{\beta}_1 + \langle T(\bar{\alpha}_1), \bar{\beta}_2 \rangle \bar{\beta}_2 + \cdots + \langle T(\bar{\alpha}_1), \bar{\beta}_m \rangle \bar{\beta}_m \]

\[ T(\bar{\alpha}_2) = \langle T(\bar{\alpha}_2), \bar{\beta}_1 \rangle \bar{\beta}_1 + \langle T(\bar{\alpha}_2), \bar{\beta}_2 \rangle \bar{\beta}_2 + \cdots + \langle T(\bar{\alpha}_2), \bar{\beta}_m \rangle \bar{\beta}_m \] (3-20)

\[ \vdots \hspace{1cm} \vdots \hspace{1cm} \vdots \]

\[ T(\bar{\alpha}_n) = \langle T(\bar{\alpha}_n), \bar{\beta}_1 \rangle \bar{\beta}_1 + \langle T(\bar{\alpha}_n), \bar{\beta}_2 \rangle \bar{\beta}_2 + \cdots + \langle T(\bar{\alpha}_n), \bar{\beta}_m \rangle \bar{\beta}_m \]

which can be rewritten as

\[ T[\bar{\alpha}_1 \mid \bar{\alpha}_2 \mid \cdots \mid \bar{\alpha}_n] = [\bar{\beta}_1 \mid \bar{\beta}_2 \mid \cdots \mid \bar{\beta}_m] \begin{bmatrix} \langle T(\bar{\alpha}_1), \bar{\beta}_1 \rangle & \langle T(\bar{\alpha}_2), \bar{\beta}_1 \rangle & \cdots & \langle T(\bar{\alpha}_n), \bar{\beta}_1 \rangle \\ \langle T(\bar{\alpha}_1), \bar{\beta}_2 \rangle & \langle T(\bar{\alpha}_2), \bar{\beta}_2 \rangle & \cdots & \langle T(\bar{\alpha}_n), \bar{\beta}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle T(\bar{\alpha}_1), \bar{\beta}_m \rangle & \langle T(\bar{\alpha}_2), \bar{\beta}_m \rangle & \cdots & \langle T(\bar{\alpha}_n), \bar{\beta}_m \rangle \end{bmatrix} \] (3-21)

Hence, Equation (3-19) represents \( T \) as claimed. ♥

**Rank and Nullity of a Matrix**

The rank and nullity of matrix \( \{a_{ij}\} \) are the rank and nullity of the linear transformation that the matrix represents. The rank of \( T : U \rightarrow V \) is the dimension of subspace \( T(U) \). Now, let \( \bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_n \) be a basis for \( U \). Then subspace \( T(U) \) is spanned by the \( n \) vectors \( T(\bar{\alpha}_1), T(\bar{\alpha}_2), \ldots, T(\bar{\alpha}_n) \). Hence, \( \text{Rank}(T) \) is the maximum number of linearly independent elements in \( T(\bar{\alpha}_1), T(\bar{\alpha}_2), \ldots, T(\bar{\alpha}_n) \). But note that

\[ T[\bar{\alpha}_1 \mid \bar{\alpha}_2 \mid \cdots \mid \bar{\alpha}_n] = [\bar{\beta}_1 \mid \bar{\beta}_2 \mid \cdots \mid \bar{\beta}_m] A \] (3-22)

implies
Observe that $T(\tilde{\alpha}_2)$ is independent of $T(\tilde{\alpha}_1)$ if and only if $[a_{12} \ldots a_{m2}]^T$ is independent of $[a_{11} \ldots a_{m1}]^T$. And, $T(\tilde{\alpha}_3)$ is independent of $T(\tilde{\alpha}_1)$ and $T(\tilde{\alpha}_2)$ if and only if $[a_{13} \ldots a_{m3}]^T$ is independent of $[a_{11} \ldots a_{m1}]^T$ and $[a_{12} \ldots a_{m2}]^T$. Continue this argument to its conclusion: $T(\tilde{\alpha}_n)$ is independent of $T(\tilde{\alpha}_1), T(\tilde{\alpha}_2), \ldots, T(\tilde{\alpha}_{n-1})$ if and only if $[a_{1n} \ldots a_{mn}]^T$ is independent of $[a_{11} \ldots a_{m1}]^T, [a_{12} \ldots a_{m2}]^T, \ldots, [a_{1n-1} \ldots a_{mn-1}]^T$, the first $n-1$ columns of $m \times n$ matrix $A$. Hence, $\text{Rank}(T)$ is equal to the maximum number of linearly independent columns of matrix $A$ (also, we know that column rank is equal to row rank so that $\text{Rank}(T)$ is equal to the maximum number of linearly independent rows of matrix $A$). Finally, we say that $m \times n$ matrix $A$ has full rank if $\text{rank}[A] = \min(m, n)$. 
Non-Singular Matrices

Consider the important case where n-dimensional \( U = V \) so that \( T : U \rightarrow U \). Simple, intuitive arguments can be given for
1) \( T : U \rightarrow U \) is one-to-one and onto if and only if \( \text{Rank}(T) = n \), which is equivalent to
2) \( T : U \rightarrow U \) is one-to-one and onto if and only if \( \text{Nullity}(T) = 0 \).

If \( \text{Rank}(T) = n \), then the n columns of \( A \) are independent, \( \text{Rank}(A) = n \), and \( A \) is said to be nonsingular. In this case, the matrix \( A^{-1} \) exists, and it represents the linear operator \( T^{-1} \).

Change of Basis - Impact on Vector Representation

With respect to an arbitrary but fixed basis, we have represented both vectors and linear transformations as n-tuples and matrixes, respectively. However, the representations are entirely dependent on the chosen basis. The vectors and transformations have meaning independent of any particular choice of basis.

Let \( \alpha_1, \alpha_2, \ldots, \alpha_n \) be a basis of n-dimensional \( U \). With respect to this basis, an \( \bar{X} \in U \) can be represented as

\[
\bar{X} = \sum_{i=1}^{n} x_i \alpha_i , \quad (3-24)
\]

or

\[
\bar{X} = [\alpha_1 \; | \; \alpha_2 \; | \; \cdots \; | \; \alpha_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} , \quad (3-25)
\]

were coordinates \( x_i, 1 \leq i \leq n \), depend both on the vector \( \bar{X} \) and the basis. Now, let us change the basis for \( U \), let \( \bar{\alpha}_1', \bar{\alpha}_2', \ldots, \bar{\alpha}_n' \) be the "new" basis for vector space \( U \). Furthermore the "old" basis and the "new" basis are related by
\[ \hat{\alpha}'_j = \sum_{i=1}^{n} p_{ij} \hat{\alpha}_i, \quad 1 \leq j \leq n, \quad (3-26) \]

which is given symbolically as

\[
\begin{bmatrix}
\hat{\alpha}_1' \\
\hat{\alpha}_2' \\
\vdots \\
\hat{\alpha}_n'
\end{bmatrix} = \begin{bmatrix}
\hat{\alpha}_1 \\
\hat{\alpha}_2 \\
\vdots \\
\hat{\alpha}_n
\end{bmatrix} P, \quad (3-27)
\]

where \( P = \{ p_{ij} \} \) is an \( n \times n \) matrix called the matrix of transition from the "old" basis \( \hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_n \) to the "new" basis \( \tilde{\alpha}_1', \tilde{\alpha}_2', \ldots, \tilde{\alpha}_n' \). The columns of \( P \) are n-tuples representing the "new" basis in terms of the "old" basis. Hence the columns of \( P \) are independent, and \( P \) is non-singular.

Now, let \( x'_i, \quad 1 \leq i \leq n \), be the "new" coordinates of \( \vec{X} \). We use (3-27) to write

\[
\begin{bmatrix}
x'_1 \\
x'_2 \\
\vdots \\
x'_n
\end{bmatrix} = \begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} P. \quad (3-28)
\]

Now, compare the last product on the right hand side of (3-28) with (3-25) to obtain

\[
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} = P \begin{bmatrix}
x'_1 \\
x'_2 \\
\vdots \\
x'_n
\end{bmatrix}. \quad (3-29)
\]

Hence, for \( 1 \leq k \leq n \), the \( k^{\text{th}} \) row of \( P \) relates the "new" and "old" coordinates \( x_k' \) and \( x_k \), respectively.

**Change of Basis on Space U - Impact on Matrix Representation for T : U \rightarrow V**

Let \( \tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_n \) and \( \tilde{\beta}_1, \tilde{\beta}_2, \ldots, \tilde{\beta}_m \) be bases of n-dimensional \( U \) and m-dimensional \( V \), respectively. With respect to these bases, let \( T : U \rightarrow V \) be represented by \( m \times n \) matrix \( A \) so that
\( T[\alpha_1 \mid \alpha_2 \mid \cdots \mid \alpha_n] = \begin{bmatrix} \bar{\beta}_1 \mid \bar{\beta}_2 \mid \cdots \mid \bar{\beta}_m \end{bmatrix} A. \) (3-30)

Now, suppose we want to change the basis of \( U \) to \( \tilde{\alpha}_1', \tilde{\alpha}_2', \ldots, \tilde{\alpha}_n' \), the "new" basis for \( U \) (the basis for \( V \) remains unchanged). The "new" basis for \( U \) is related to the "old" basis for \( U \) by (3-26) and (3-27).

Let \( m \times n \) matrix \( A' \) represent \( T \) with respect to the bases \( \tilde{\alpha}_1', \tilde{\alpha}_2', \ldots, \tilde{\alpha}_n' \) and \( \bar{\beta}_1, \bar{\beta}_2, \ldots, \bar{\beta}_m \). Then,

\[ T[\tilde{\alpha}_1 \mid \tilde{\alpha}_2 \mid \cdots \mid \tilde{\alpha}_n'] = \begin{bmatrix} \bar{\beta}_1 \mid \bar{\beta}_2 \mid \cdots \mid \bar{\beta}_m \end{bmatrix} A'. \] (3-31)

Now, use (3-27) in (3-31) to obtain

\[ T[\tilde{\alpha}_1 \mid \tilde{\alpha}_2 \mid \cdots \mid \tilde{\alpha}_n'] P = \begin{bmatrix} \bar{\beta}_1 \mid \bar{\beta}_2 \mid \cdots \mid \bar{\beta}_m \end{bmatrix} A'. \] (3-32)

On the right, multiply this result by \( P^{-1} \) to obtain

\[ T[\tilde{\alpha}_1 \mid \tilde{\alpha}_2 \mid \cdots \mid \tilde{\alpha}_n'] = \begin{bmatrix} \bar{\beta}_1 \mid \bar{\beta}_2 \mid \cdots \mid \bar{\beta}_m \end{bmatrix} A' P^{-1}. \] (3-33)

Compare (3-30) and (3-33) and conclude that

\[ A = A' P^{-1} \iff A' = AP. \] (3-34)

**Change of Basis on Space V - Impact on Matrix Representation for \( T : U \rightarrow V \)**

Consider changing the basis for the co-domain vector space \( V \). Again, \( m \times n \) matrix \( A \) represents \( T \) with respect to the "old" bases \( \tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_n \) and \( \bar{\beta}_1, \bar{\beta}_2, \ldots, \bar{\beta}_m \). Let \( \bar{\beta}_1', \bar{\beta}_2', \ldots, \bar{\beta}_m' \) be a "new" basis for space \( V \), and let
\[ \tilde{\beta}_j' = \sum_{i=1}^{m} q_{ij} \tilde{\beta}_i, \quad 1 \leq j \leq m, \quad (3-35) \]

\[ \begin{bmatrix} \tilde{\beta}_1' & \tilde{\beta}_2' & \cdots & \tilde{\beta}_m' \end{bmatrix} = \begin{bmatrix} \tilde{\beta}_1 & \tilde{\beta}_2 & \cdots & \tilde{\beta}_m \end{bmatrix} Q, \quad (3-36) \]

where \( m \times m \), nonsingular matrix \( Q = \{q_{ij}\} \) is the matrix of transition that relates the "old" and "new" basis for \( V \). Let \( m \times n \) matrix \( A'' \) represent \( T \) with respect to "old" \( \tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_n \) and "new" \( \tilde{\beta}_1', \tilde{\beta}_2', \ldots, \tilde{\beta}_m' \) so that

\[ T[\tilde{\alpha}_1 | \tilde{\alpha}_2 | \cdots | \tilde{\alpha}_n] = \begin{bmatrix} \tilde{\beta}_1' & \tilde{\beta}_2' & \cdots & \tilde{\beta}_m' \end{bmatrix} A''. \quad (3-37) \]

Now, use (3-36) in (3-37) to obtain

\[ T[\tilde{\alpha}_1 | \tilde{\alpha}_2 | \cdots | \tilde{\alpha}_n] = \begin{bmatrix} \tilde{\beta}_1 & \tilde{\beta}_2 & \cdots & \tilde{\beta}_m \end{bmatrix} QA''. \quad (3-38) \]

Now, compare (3-38) and (3-30) to obtain

\[ A = QA'' \iff A'' = Q^{-1} A \quad (3-39) \]

**Simultaneous Change of Basis on U and V - Impact on Representation for \( T : U \to V \)**

Again, \( m \times n \) matrix \( A \) represents \( T \) with respect to the "old" bases \( \tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_n \) and \( \tilde{\beta}_1, \tilde{\beta}_2, \ldots, \tilde{\beta}_m \). Let \( \tilde{\alpha}_1', \tilde{\alpha}_2', \ldots, \tilde{\alpha}_n' \) and \( \tilde{\beta}_1', \tilde{\beta}_2', \ldots, \tilde{\beta}_m' \) be "new" bases for spaces \( U \) and \( V \), respectively. Also, let \( P \) and \( Q \) be matrices of transition as described by (3-26)/(3-27) and (3-35)/(3-36), respectively. Let \( A' \) be the matrix representing \( T \) with respect to the new bases on \( U \) and \( V \). Combine the results leading to (3-34) and (3-39) to conclude that
The case $U = V$ and $T : U \rightarrow U$

Finally, we consider the special case where n-dimensional $U = V$ and $T : U \rightarrow U$. Let $n \times n$ matrix $A$ represent $T$ with respect to "old" basis $\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_n$. Then the matrix representing $T$ with respect to the new basis $\bar{\alpha}'_1, \bar{\alpha}'_2, \ldots, \bar{\alpha}'_n$ is

$$A' = P^{-1}AP \quad \Leftrightarrow \quad PA' = AP. \quad (3-41)$$

We say that $P^{-1}AP$ is a similarity transformation applied to $A$ and that $A'$ and $A$ are similar. Change of bases and similarity transformations have many applications in engineering and the physical sciences.

Let the $n \times n$ matrix of transition $P$ be represented in column form as

$$P = \begin{bmatrix} \tilde{p}_1 & \tilde{p}_2 & \cdots & \tilde{p}_n \end{bmatrix}, \quad (3-42)$$

where $\tilde{p}_k$ is the $k^{th}$ column of $P$. Then $PA' = AP$ can be written as

$$\begin{bmatrix} \tilde{p}_1 & \tilde{p}_2 & \cdots & \tilde{p}_n \end{bmatrix}A' = A\begin{bmatrix} \tilde{p}_1 & \tilde{p}_2 & \cdots & \tilde{p}_n \end{bmatrix}, \quad (3-43)$$

which implies that the $i^{th}$ column of $A'$ is just the representation of $A\tilde{p}_i$ with respect to the columns $\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_n$. Let $a'_{ij}$ denote the elements of $A'$. Then we have

$$a'_{ij} \tilde{p}_1 + a'_{2j} \tilde{p}_2 + \cdots + a'_{nj} \tilde{p}_n = A\tilde{p}_j, \quad 1 \leq j \leq n, \quad (3-44)$$

which is equivalent to
\[ \sum_{k=1}^{n} a'_{kj} \bar{p}_k = A \bar{p}_j, \quad 1 \leq j \leq n. \]  

(3-45)

**Example**

Let \( A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{bmatrix} \), \( P = \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \).

We want to compute \( A' = P^{-1}AP \). Computing the inverse \( P^{-1} \) is too much work - don't do it! Instead, use \( PA' = AP \) to write

\[
\begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_3 \end{bmatrix}A' = A\begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_3 \end{bmatrix} \iff \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 10 \\ 0 & 1 & 6 \\ 0 & 0 & 2 \end{bmatrix}
\]

which leads to

\[
A' = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}
\]

by inspection!♥

In most cases, the method illustrated by the previous example is much easier than computing the inverse \( P^{-1} \) and then the product \( P^{-1}AP \).