
Recall the class of mappings \( T: \mathbb{R}^n \rightarrow \mathbb{R}^m \), where

\[
(x_1, \ldots, x_n) \rightarrow (u_1, \ldots, u_m), \quad u_i = \sum_{j=1}^{n} t_{ij} x_j \quad (\star)
\]

These mappings are linear. Conversely, there are the only linear maps \( \mathbb{R}^n \rightarrow \mathbb{R}^m \).

**Theorem 4.1:** Every linear map \( T: \mathbb{R}^n \rightarrow \mathbb{R}^m \) can be written in this form (i.e., as in \( \star \)).

**Proof:** Write \( x = \sum_{j=1}^{n} x_j e_j \), where \( e_j \) is the \( j \)-th unit basis vector.

Since \( T \) is linear, \( u = Tx = \sum_{j=1}^{n} x_j T e_j \)

Let \( t_{ij} \) be the \( i \)-th component of \( T e_j \), i.e.,

\[
e_j \rightarrow T(e_j) = (t_{1j}, t_{2j}, \ldots, t_{nj}) = T e_j
\]

*Note:* If \( e_1, \ldots, e_n \) is the dual basis to \( e_1, \ldots, e_n \) (i.e., \( e_i(e_j) = 1 \) if \( i=j \) and 0 if \( i \neq j \)), then

\[
t_{ij} = (e_i, T e_j)
\]

The \( i \)-th component of \( u \) is \( u_i = \sum_{j=1}^{n} x_j t_{ij} \).
It is convenient to arrange the coefficients $t_{ij}$ in an $m \times n$ matrix:

$$
\begin{pmatrix}
  t_{11} & t_{12} & \cdots & t_{1n} \\
  t_{21} & t_{22} & \cdots & t_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  t_{m1} & t_{m2} & \cdots & t_{mn}
\end{pmatrix}
$$

The numbers $t_{ij}$ are called entries. A matrix is square if $n=m$.

By Theorem 4.1, there is a 1-1 correspondence between $m \times n$ matrices and linear maps $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$. Denote the $(ij)$-entry of the corresponding matrix by $t_{ij} = (T)_{ij}$.

A matrix $T$ can be thought of as a row of column vectors, or a column of row vectors:

$$
T = (c_1, \ldots, c_n) = \begin{pmatrix}
  \vdots \\
  r_m
\end{pmatrix}, \quad c_j = \begin{pmatrix}
  t_{1j} \\
  \vdots \\
  t_{mj}
\end{pmatrix}, \quad r_i = (t_{i1}, \ldots, t_{in}).
$$

Note that $T e_j = c_j$, i.e., $T = (Te_1, \ldots, Te_n)$.

We will write vectors in $X = \mathbb{R}^n$ and $U = \mathbb{R}^m$ as column vectors, for consistency.

We will write "co-vectors" (i.e., linear maps $\mathbb{R}^n \rightarrow \mathbb{R}$) as row vectors:

$$
l = (l_1, \ldots, l_n)
$$

Thus, $(l, x) = (l_1, \ldots, l_n) \begin{pmatrix}
  x_1 \\
  \vdots \\
  x_n
\end{pmatrix} = l_1 x_1 + \ldots + l_n x_n$

Similarly, we can write $Tx = \begin{pmatrix}
  r^1 \\
  \vdots \\
  r^m
\end{pmatrix} x = \begin{pmatrix}
  r_1^T x \\
  \vdots \\
  r_m^T x
\end{pmatrix}$, where each $r_i$ is a row vector.
Remark. Recall the natural identification
\[ X \rightarrow X' \]
\[ x \mapsto \hat{x} = (l, x) \]
\[ \hat{x}(l, x) \] at \( x \).

Now, for \( x = \left( \begin{array}{c} x_1 \\ \vdots \\ x_n \end{array} \right) \in \mathbb{R}^n \), \( l = (l_1, \ldots, l_n) \in (\mathbb{R}^n)^t \), \( \hat{x} \) inputs a vector \( l \), i.e.,
\[ \hat{x}(l) = \left( x_1, \ldots, x_n \right) \left( \begin{array}{c} l_1 \\ \vdots \\ l_n \end{array} \right) = l_1x_1 + \cdots + l_nx_n \]
\[ = \left( l_1, \ldots, l_n \right) \left( \begin{array}{c} x_1 \\ \vdots \\ x_n \end{array} \right) = (l, x). \]

If \( S, T : \mathbb{R}^n \rightarrow \mathbb{R}^n \), it is easily checked that \((T+S)_{ij} = T_{ij} + S_{ij}\).

If \( T : \mathbb{R}^n \rightarrow \mathbb{R}^l \), \( S : \mathbb{R}^n \rightarrow \mathbb{R}^k \), then \( T e_j = c_j \), and so the \( j \)th column of \( ST \) is \( ST e_j = S c_j = \left( \begin{array}{c} S_1c_j \\ \vdots \\ S_nc_j \end{array} \right) \), where \( s_{ij} \)

is the \( i \)th row of \( S \).

In summary, we multiply matrices as follows:

Let \( T \) be \( m \times n \), and \( S \) be \( l \times m \). The product \( ST \) is an \( l \times n \) matrix, whose \((k,j)\)th entry is the product of the \( k \)th row of \( S \) with the \( j \)th column of \( T \).

\[ \text{Sk: kth row of } S \]
\[ \text{Cj: jth column of } T \]
\[ (ST)_{kj} = S_k c_j = \sum_{i=1}^{m} S_{ki} t_{ij} \]

\[ \begin{bmatrix} \vdots & \vdots \\ S_1 & \cdots & S_m \end{bmatrix} \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} = \begin{bmatrix} S_1 c_1 + \cdots + S_m c_n \\ \vdots \\ S_m c_1 + \cdots + S_m c_n \end{bmatrix} \]
Examples:

1. \((\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}) \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 23 & 28 \end{pmatrix}\) vs. \(\begin{pmatrix} 43 & 50 \\ 71 & 96 \end{pmatrix} = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\)

2. \((\begin{pmatrix} 1 \\ 2 \end{pmatrix}) \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix}\) vs. \(\begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 11 \end{pmatrix}\)

3. \((\begin{pmatrix} 1 & 2 \end{pmatrix}) \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 13 & 16 \end{pmatrix}\) vs. \(\begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 17 \end{pmatrix}\)

4. \((\begin{pmatrix} 1 & 2 \end{pmatrix}) \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 13 & 16 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 45 \end{pmatrix}\)

vs. \((\begin{pmatrix} 1 & 2 \end{pmatrix}) \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 13 & 16 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 45 \end{pmatrix}\)

Remark: Let \(A\) be an \(m \times n\) matrix, and \(D = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}\) a diagonal matrix, that is, \(D_{ij} = \begin{cases} d_j & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}\).

The \(i\)-th row of \(DA\) is the \(i\)-th row of \(A\), times \(d_i\).

The \(j\)-th column of \(AD\) is the \(j\)-th column of \(A\), times \(d_j\).

An \(m \times n\) matrix represents a mapping \(\mathbb{R}^n \rightarrow \mathbb{R}^m\). If the mapping is invertible, then the matrix is called invertible.

Remark: Since the composition of linear maps is associative, matrix multiplication is associative as well.
We shall identify \( \mathbb{R}^n \) with column vectors, and its dual \((\mathbb{R}^n)'\) with row vectors (with \( n \) components).

A co-vector \( l \in (\mathbb{R}^n)' \) applied to a vector \( x \in \mathbb{R}^n \) is the matrix product \( (l, x) = (l_1, \ldots, l_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = l_1 x_1 + \ldots + l_n x_n \).

Let \( x, T, \) and \( l \) be the following linear maps:

\[
x : \mathbb{R} \longrightarrow \mathbb{R}, \quad T : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \quad l : \mathbb{R}^m \longrightarrow \mathbb{R}.
\]

Note that \( l \in (\mathbb{R}^m)' \), \( lT \in (\mathbb{R}^n)' \).

By associativity, \( (lT)x = l(Tx) \).

We can write this as \( (lT, x) = (l, Tx) \).

Recall that the transpose \( T' \) of \( T \) satisfies \( (T', l, x) = (l, Tx) \).

Let's compare the matrix forms of \( T \) and \( T' \):

Let \( e_1, \ldots, e_n \) be a basis for \( X \), and let \( l_1, \ldots, l_n \in X' \) be the dual basis (\( l_i(e_j) = 1 \) if \( i = j \), \( 0 \) if \( i \neq j \)).

Recall that \( T e_j = \sum_{k=1}^n t_{kj} e_j \Rightarrow (l_i, T e_j) = \sum_{k=1}^n t_{kj} (l_i, e_k) = t_{ij} \).

Thus, to get \( t_{ij} \):

- Apply \( T \) to the \( j \)th basis vector of \( X \).
- Then apply the \( i \)th functional in \( X' \) to this.
Now consider $T'$:

Using our recipe for $t_{ij}$ just derived, the $ij$-entry $t'_{ij}$ of $T'$ is obtained by:

1. Applying $T'$ to the $j^{th}$ basis vector of $X'$.
2. Then applying the $i^{th}$ functional in $X''$ to this.

i.e., $(\hat{x}_i, T'l_j)$, or just $(T'l_j, x_i)$ under the natural identification $X \leftrightarrow X''$.

Note that $t'_{ij} = (T'l_j, x_i) = (l_j, Tx_i) = t_{ji}$.

* In summary, to represent the transpose $T'$ as a matrix, we change its rows into columns, and vice-versa. Denote this matrix by $T^T$. We have: $(T^T)_{ij} = T_{ji}$.

\[
\begin{bmatrix}
\hat{x}_i \\
\end{bmatrix}_{1 \times m} \begin{bmatrix}
T' \\
\end{bmatrix}_{m \times n} \begin{bmatrix}
x_i \\
\end{bmatrix}_{n \times 1} = (\hat{x}, T'l)
\]

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

$T^T : (\mathbb{R}^m)' \rightarrow (\mathbb{R}^n)'$
Now, let's express the range of $T$ in matrix language. Recall that $Te_j = c_j$.

Plugging this into $Tx = \sum_{j=1}^{\hat{\lambda}} x_j Te_j$ gives

$$u = Tx = x_1 c_1 + \ldots + x_n c_n.$$ 

In other words:

**Theorem 4.2:** The range of $T$ consists of all linear combinations of the column vectors of the matrix.

The dimension of this space is called the **column rank** of $T$, in remedial books and courses. The **row rank** is defined similarly.

**Remarks:**

(i) The row rank of $T$ is $\dim R_{T^\top}$ (or $\dim R_{T^*}$).
(ii) By Theorem 3.6 ($\dim R_T = \dim R_{T^\top}$), the row rank and column rank of a matrix are equal.

Let $T : X \to U$ be a linear map, and let $x_1, \ldots, x_n$ be a basis for $X$, and $u_1, \ldots, u_m$ a basis for $U$.

Since $\dim X = n$, $\dim U = m$, we have $X \cong \mathbb{R}^n$ and $U \cong \mathbb{R}^m$. 
That is, we have isomorphisms:

\[ B : X \to \mathbb{R}^n \quad \quad C : U \to \mathbb{R}^m \]

\[ x_i \to e_i \quad \quad u_i \to e_i \]

Putting this together, we can choose isomorphisms such as these to get a linear map \( C T B^{-1} : \mathbb{R}^n \to \mathbb{R}^m \)

\[ \begin{align*}
\mathbb{R}^n & \xrightarrow{B^{-1}} X \xrightarrow{T} U \xrightarrow{C} \mathbb{R}^m \\
\end{align*} \]

If \( T : X \to X \), mapping \( X \) to itself, then we can take \( B = C \), and so we get a matrix \( M = B T B^{-1} \).

Suppose we change the isomorphism \( B \). In particular, let \( C : X \to \mathbb{R}^n \) be another isomorphism, and let \( N \) be the matrix with this basis, i.e., \( N = C T C^{-1} \).

We have: \( N = C T C^{-1} = C B^{-1} B T B^{-1} B C^{-1} = S M S^{-1} \),

where \( S = C B^{-1} \), which is invertible.

\[ \begin{align*}
\mathbb{R}^n & \xrightarrow{C^{-1}} X \xrightarrow{B} \mathbb{R}^n \xrightarrow{B^{-1}} X \xrightarrow{T} X \xrightarrow{B} \mathbb{R}^n \xrightarrow{B^{-1}} X \xrightarrow{C} \mathbb{R}^n \\
\end{align*} \]

Two square matrices \( N \) and \( M \) related by conjugation (e.g., \( N = S M S^{-1} \)) are said to be similar.
Similar matrices describe the same mapping of a space into itself, but using different bases. Thus, (as we'll see later), similar matrices share the same intrinsic properties.

**Remark:** We can write any \( n \times n \) matrix \( A \) in block form:

\[
A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}
\]

where

\[
A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}
\]

\[
\begin{array}{c|c}
A_{11} & A_{12} \\
\hline
A_{21} & A_{22}
\end{array}
\]

\[
\begin{array}{c|c}
A_{11} & A_{12} \\
\hline
A_{21} & A_{22}
\end{array}
\]

\[
\begin{array}{c|c}
A_{11} & A_{12} \\
\hline
A_{21} & A_{22}
\end{array}
\]

Addition and multiplication of block matrices "works out" just as if the blocks were entries.

**Def:** The square matrix \( I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \) is called the identity, or unit matrix.

**Def:** A square matrix \( T = (t_{ij}) \) for which \( t_{ij} = 0 \) for \( i > j \) (resp. \( i < j \)) is called upper triangular (resp. lower triangular).

**Def:** A square matrix \( T = (t_{ij}) \) for which \( t_{ij} = 0 \) for \( |i-j| > 1 \) is called tridiagonal.
Systems of equations

Matrices can be used to effectively express and solve systems of equations.

A system of equations \( \sum_{i=1}^{n} t_{ij} x_i = u_j \) for \( j = 1, \ldots, n \) may have a unique solution, many solutions, or no solutions.

Consider the following example, where \( x_1, \ldots, x_4 \) are the unknowns:

\[
\begin{align*}
    x_1 + x_2 + 2x_3 + 3x_4 &= u_1 \\
    x_1 + 2x_2 + 3x_3 + x_4 &= u_2 \\
    2x_1 + x_2 + 2x_3 + 3x_4 &= u_3 \\
    3x_1 + 4x_2 + 6x_3 + 2x_4 &= u_4 
\end{align*}
\]

We can write this system as \( Mx = u \):

\[
\begin{pmatrix}
    1 & 1 & 2 & 3 \\
    1 & 2 & 3 & 1 \\
    2 & 1 & 2 & 3 \\
    3 & 4 & 6 & 2 \\
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2 \\
    x_3 \\
    x_4 \\
\end{pmatrix}
= 
\begin{pmatrix}
    u_1 \\
    u_2 \\
    u_3 \\
    u_4 \\
\end{pmatrix}
\]

This system can be solved by Gaussian elimination, i.e., eliminating the variables one-by-one.
Example (cont.): Subtract multiples of Eq. 1 from the last 3 equations to eliminate $x_1$:

\[
\begin{align*}
X_2 + X_3 - 2X_4 &= U_2 - U_1 \\
-X_2 - 2X_3 - 3X_4 &= U_3 - 2U_1 \\
X_2 - 7X_4 &= U_4 - 3U_1
\end{align*}
\]

Use the 1st equation to eliminate $X_2$ from the last two:

\[
\begin{align*}
-X_3 - 5X_4 &= U_3 + U_2 - 3U_1 \\
-X_3 - 5X_4 &= U_4 - 2U_2 - 2U_1
\end{align*}
\]

Subtract these two equations to eliminate $X_3$ (and by chance, $X_4$ as well):

\[
0 = U_4 - U_3 - 2U_2 + U_1
\]

This is a necessary and sufficient condition for our original system to have a solution. It can be written as $(l, u) = 0$, where $l = (1, -2, -1, 1)$, that is,

\[
(1, -2, -1, 1) \begin{pmatrix}
U_1 \\
U_2 \\
U_3 \\
U_4
\end{pmatrix} = U_1 - 2U_2 - U_3 + U_4 = 0
\]

Since $Mx = u$, and $lu = 0$, we must have $LMx = lu = 0$ for all $x \in \mathbb{R}^4$, thus $LM = 0$.

\[
* \text{In general, if } Mx = u \text{ is a system of equations, then a solution is described by a linear function } l
\]

(equivalently, a column vector) if $LM = 0$.  \[\]