Sections 2.1-2.3 Matrix operations

- Matrix addition/subtraction
- Matrix multiplication
- Matrix powers
- Matrix transpose
- Matrix inverse
- The Invertible Matrix Theorem

Section 2.4 Partitioned matrices

Section 2.5 LU decomposition
Matrix Algebra

Introduction

Matrices are **two dimensional arrays** of real numbers that are arranged along rows (first dimension) and columns (second dimension):

\[
A = \begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
= [a_1 \ a_2 \ \ldots \ a_n].
\]

We denote matrices that have \( m \) rows and \( n \) columns by \( A \in \mathbb{R}^{m \times n} \), and say that the **size** of the matrix is \( m \times n \).

Vectors can be regarded as matrices with size \( n \times 1 \) (column) or \( 1 \times n \) (row).

Sometimes, we also use notation like \( A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \), or even \( A = (a_{ij}) \).
Special matrices

We say that $A$ is a **square** matrix if $m = n$ (i.e., equally many rows and columns).

**Diagonal** matrices are square matrices whose only nonzero entries are in the main diagonal of the matrix

$$A = \begin{bmatrix}
a_{11} & & \\
& \ddots & \\
& & a_{nn}
\end{bmatrix} \quad \leftarrow \text{empty spaces indicate zero}
$$

An **identity matrix** is a diagonal matrix with constant value 1 along the diagonal:

$$I_n = \text{diag}(1, \ldots, 1) \in \mathbb{R}^{n \times n}.$$

Lastly, a **zero matrix** is a matrix with all entries being 0, and denoted as $O$.  

---

Prof. Guangliang Chen | Mathematics & Statistics, San José State University
Matrix operations

• Scalar multiple of a matrix
• Matrix-vector product
• Adding two matrices of the same size (also letting them subtract)
• Multiplying two matrices of “matching” sizes
• Transpose of a matrix
• Inverse of a square matrix
Def 0.1 (Scalar multiple). Let $r$ be a real number and $A \in \mathbb{R}^{m \times n}$. Then $B = rA$ is defined as a matrix of the same size with entries $b_{ij} = ra_{ij}$.

In matrix form, this is

$$rA = \begin{bmatrix} r a_{11} & r a_{12} & \cdots & r a_{1n} \\ r a_{21} & r a_{22} & \cdots & r a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r a_{m1} & r a_{m2} & \cdots & r a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$
Def 0.2 (Matrix sum/difference). Let \( A, B \in \mathbb{R}^{m \times n} \). Then the matrix sum \( C = A + B \) is defined as a matrix of the same size with the following entries

\[
C = (c_{ij}), \quad c_{ij} = a_{ij} + b_{ij}
\]

In matrix form, the above definition becomes

\[
A + B = \begin{bmatrix}
a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\
a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn}
\end{bmatrix} \in \mathbb{R}^{m \times n}
\]

Remark. The difference of two matrices, \( A - B \), is defined similarly (with every + sign being changed to - sign).
Example 0.1. Let

\[
A = \begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix}, \quad
B = \begin{bmatrix}
-1 & -1 & -1 \\
1 & 1 & 1
\end{bmatrix}.
\]

Find \( A + B \), \( A - B \), \( 3B \) and \( A + 3B \).
The scalar multiple of a matrix and matrix sum satisfy the following **commutative**, **associative** and **distributive** laws.

**Theorem 0.1.** Let \( \mathbf{A}, \mathbf{B}, \mathbf{C} \) be three matrices of the same size and \( r, s \) be scalars. Then

- \( \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \)
- \( \mathbf{A} + \mathbf{O} = \mathbf{O} + \mathbf{A} = \mathbf{A} \) (\( \mathbf{O} \) is the zero matrix of same size)
- \( (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \)
- \( r(s\mathbf{A}) = (rs)\mathbf{A} \)
- \( r(\mathbf{A} + \mathbf{B}) = r\mathbf{A} + r\mathbf{B} \)
- \( (r + s)\mathbf{A} = r\mathbf{A} + s\mathbf{A} \)
Matrix-vector product

**Def 0.3.** Let $A \in \mathbb{R}^{m \times n}$ and $x \in \mathbb{R}^n$. Their product is defined as a vector $y \in \mathbb{R}^m$ of the following form

$$y = Ax = \begin{bmatrix}
    a_{11} & a_{12} & \cdots & a_{1n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{i1} & a_{i2} & \cdots & a_{in} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{bmatrix}
= \begin{bmatrix}
    a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
    \vdots \\
    a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \\
    \vdots \\
    a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n
\end{bmatrix}
$$

In compact notation,

$$y = (y_i) \in \mathbb{R}^m, \quad \text{with} \quad y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad 1 \leq i \leq m$$
Alternatively (as we have already seen previously), we can multiply a matrix and a vector in a columnwise fashion.

**Theorem 0.2.** Let $A = [a_1 \ldots a_n] \in \mathbb{R}^{m \times n}$ and $x \in \mathbb{R}^n$. Then

$$Ax = [a_1 \ldots a_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \cdot a_1 + \cdots + x_n \cdot a_n.$$ 

**Proof.** By definition,

$$Ax = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix} = x_1a_1 + \cdots + x_na_n.$$
Two properties about matrix-vector multiplication

**Theorem 0.3.** Let $A \in \mathbb{R}^{m \times n}$ and $x, y \in \mathbb{R}^n$ and $r \in \mathbb{R}$. Then

- $A(x + y) = Ax + Ay$
- $A(rx) = r(Ax)$

**Remark.** They were needed for showing that transformations of the form $f(x) = Ax$ must be linear.
Proof. By the columnwise way of multiplying a matrix and a vector,

\[ A(x + y) = [a_1 \ldots a_n] \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} = (x_1 + y_1)a_1 + \cdots + (x_n + y_n)a_n \\
= (x_1a_1 + \cdots + x_na_n) + (y_1a_1 + \cdots + y_na_n) \\
= Ax + Ay. \]

Similarly,

\[ A(rx) = [a_1 \ldots a_n] \begin{bmatrix} rx_1 \\ \vdots \\ rx_n \end{bmatrix} = (rx_1)a_1 + \cdots + (rx_n)a_n = r \underbrace{x_1a_1 + \cdots + x_na_n}_{Ax}. \]
A third property about matrix-vector multiplication

**Theorem 0.4.** Let $A, B \in \mathbb{R}^{m \times n}$ and $x \in \mathbb{R}^n$. Then

$$(A + B)x = Ax + Bx.$$ 

**Proof.** Let $A = [a_1, \ldots, a_n]$ and $B = [b_1, \ldots, b_n]$. Then

$$A + B = [a_1 + b_1, \ldots, a_n + b_n].$$

It follows that

$$\begin{align*}
(A + B)x &= x_1(a_1 + b_1) + \cdots + x_n(a_n + b_n) \\
&= (x_1a_1 + \cdots + x_na_n) + (x_1b_1 + \cdots + x_nb_n) \\
&= Ax + Bx.
\end{align*}$$
Matrix Algebra

Matrix-matrix multiplications

Def 0.4. Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$. Their product is defined as a matrix $C \in \mathbb{R}^{m \times p}$ with entries

$$c_{ij} = \begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Remark. The matrix-vector product is just the special case of $p = 1$. 

Prof. Guangliang Chen | Mathematics & Statistics, San José State University 15/75
Example 0.2. Let

\[
\begin{align*}
A &= \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, & B &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}.
\end{align*}
\]

Find $AB$ and $BA$. Are they the same?

Example 0.3. Let

\[
\begin{align*}
A &= \begin{bmatrix} 2 & 2 \\ 3 & 3 \end{bmatrix}, & B &= \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \end{bmatrix}.
\end{align*}
\]

Find $AB$. Is $BA$ defined?
Why does Morpheus keep asking people if they work from home?

It's dangerous to assume that they commute.

(Taken from https://mathwithbaddrawings.com/2018/03/07/matrix-jokes/)
 WARNINGS

- There is no commutative law between matrices: $AB \neq BA$. In fact, not both of them need to be defined at the same time.

- If $AB = O$, then we cannot conclude that $A = O$ or $B = O$.

- There is no cancellation law, i.e., $AB = AC$ does not necessarily imply $B = C$.

Can you give an example for the last statement?
A small, useful result on matrix-matrix-vector product

**Theorem 0.5.** Let \( A \in \mathbb{R}^{m \times n} \), \( B \in \mathbb{R}^{n \times p} \) and \( x \in \mathbb{R}^p \). Then

\[
(AB)x = A(Bx).
\]

**Proof.** We compare the entries of both sides. For any \( 1 \leq i \leq m \),

\[
((AB)x)_i = \sum_j (AB)_{ij}x_j = \sum_j \sum_k a_{ik}b_{kj}x_j
\]

\[
= \sum_k a_{ik} \sum_j b_{kj}x_j = \sum_k a_{ik} (Bx)_k = (A(Bx))_i.
\]

**Remark.** The right hand side is much more efficient to compute, especially when having large matrices \( A, B \).
Matrix computing in Matlab (optional)

See the following lecture:
https://www.sjsu.edu/faculty/guangliang.chen/Math250/lec2matrixcomp.pdf

Matlab scripts available on the Math 250 course page:
https://www.sjsu.edu/faculty/guangliang.chen/Math250.html
The columnwise matrix multiplication (very important)

**Theorem 0.6.** Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$. Then

$$C = AB = A[b_1 \ldots b_p] = [Ab_1 \ldots Ab_p]$$

This shows that for each $j = 1, \ldots, p$, the $j$th column of $AB$ is equal to $A$ times the $j$th column of $B$. 

\[\begin{array}{c|c}
\text{C} & \text{A} \\
\hline
\end{array}\] 

\[\begin{array}{c|c}
& \text{B} \\
\end{array}\]
Properties of matrix multiplication

Theorem 0.7. Let $A \in \mathbb{R}^{m \times n}$. Then

- $A(BC) = (AB)C$ (for $B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{p \times q}$)
- $A(B + C) = AB + AC$ (for $B, C \in \mathbb{R}^{n \times p}$)
- $(B + C)A = BA + CA$ (for $B, C \in \mathbb{R}^{\ell \times m}$)
- $r(AB) = (rA)B = A(rB)$ (for $B \in \mathbb{R}^{n \times p}$)
- $I_mA = AI_n = A$.

Proof. Enough to compare columns.
Example 0.4. Compute the following product

\[
\begin{bmatrix}
1 \\
-1
\end{bmatrix}
\begin{bmatrix}
1 & 1 & 1
\end{bmatrix}
\begin{bmatrix}
1 \\
2 \\
3
\end{bmatrix}
\]
Matrix powers

**Def 0.5.** Let $A \in \mathbb{R}^{n \times n}$ be a square matrix and $k$ a positive integer. Then the $k$th power of $A$ is defined as

$$A^k = \underbrace{A \cdot A \cdots A}_{k \text{ copies}}.$$

**Example 0.5.** Let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Find $A^3$ and $B^3$. What are $A^k$ and $B^k$ for $k > 3$?
Transpose of a matrix

**Def 0.6.** Let $A \in \mathbb{R}^{m \times n}$ be any matrix. Its transpose, denoted as $A^T$ is defined to the $n \times m$ matrix $B$ with entries $b_{ij} = a_{ji}$.

*Remark.* During the transpose operation, rows (of $A$) become columns (of $B$), and columns become rows.
Example 0.6. Find the transpose of the following matrices:

\[
\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}
\]
Properties of the matrix transpose

*Theorem* 0.8. Let $A, B$ be matrices with appropriate sizes for each statement.

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- For any scalar $r$, $(rA)^T = rA^T$
- $(AB)^T = B^T A^T$ (not the other product $A^T B^T$, which may not even be defined)

*Proof.* The first three are obvious. To prove the last one, check the $ij$-entry of each side. We show the work in class. \qed
Matrix inverse

Just like nonzero real numbers \((a \in \mathbb{R})\) have their reciprocals \((\frac{1}{a})\), certain (not all) square matrices have matrix inverses.

**Def 0.7.** A square matrix \(A \in \mathbb{R}^n\) is said to be invertible if there exists another matrix of the same size \(B\) such that

\[
AB = BA = I_n.
\]

In this case, \(B\) is called the inverse of \(A\) and we write \(B = A^{-1}\) (\(A\) is also called the inverse of \(B\)).
Example 0.7. Verify that $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $B = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$ are inverses of each other and then use this fact to solve the matrix equation $Ax = b$ for $b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. 


From the previous example, we can formulate the following theorem.

**Theorem 0.9.** Consider a matrix equation $Ax = b$ where $A \in \mathbb{R}^{n \times n}$ is a square matrix. If $A$ is invertible, then for any vector $b \in \mathbb{R}^n$, the system has a unique solution $x = A^{-1}b$.

**Proof.** Since $A$ is invertible, its inverse $A^{-1}$ exists and we can use it to multiply both sides of the equation

$$A^{-1}(Ax) = A^{-1}b$$

By the associative law,

$$(A^{-1}A)x = A^{-1}b$$

which yields that

$$x = A^{-1}b.$$
Illustration of $A^{-1}$ as a transformation
Properties of matrix inverse

Theorem 0.10. Let $A, B$ be two invertible matrices of the same size. Then

- $\left( A^{-1} \right)^{-1} = A$
- $\left( A^T \right)^{-1} = (A^{-1})^T$
- For any nonzero scalar $r$, $(rA)^{-1} = \frac{1}{r} A^{-1}$
- $\left( AB \right)^{-1} = B^{-1} A^{-1}$ (not the other product $A^{-1} B^{-1}$)

Proof. We verify them in class.
The Invertible Matrix Theorem (part 1)

“For a square matrix, lots of things are the same.”

*Theorem* 0.11. Let $A \in \mathbb{R}^{n \times n}$ be a square matrix. Then the following statements are all equivalent:

1. $A$ is invertible.
2. There is an $n \times n$ matrix $C$ such that $CA = I$.
3. The equation $Ax = 0$ only has the trivial solution.
4. $A$ has $n$ pivot positions.
5. $A$ is row equivalent to $I_n$. 
The Invertible Matrix Theorem (part 2)

Theorem 0.12. Let $A \in \mathbb{R}^{n \times n}$ be a square matrix. Then the following statements are all equivalent:

1. $A$ is invertible.

6. There is an $n \times n$ matrix $D$ such that $AD = I$.

7. The equation $Ax = b$ (for any $b$) is always consistent.

8. The columns of $A$ span $\mathbb{R}^n$.

9. The linear transformation $f(x) = Ax$ (from $\mathbb{R}^n$ to $\mathbb{R}^n$) is onto.
The Invertible Matrix Theorem (part 3)

Theorem 0.13. Let $A \in \mathbb{R}^{n \times n}$ be a square matrix. Then the following statements are all equivalent:

1. $A$ is invertible.

10. $A^T$ is invertible.

3. The equation $Ax = 0$ only has the trivial solution.

11. The columns of $A$ form a linearly independent set.

12. The linear transformation $f(x) = Ax$ is one-to-one.
Summary

Let $A \in \mathbb{R}^{n \times n}$ be a square matrix.

If $A$ is invertible, then all of the following statements are true.

Conversely, if any of the following statement is true, then $A$ must be invertible.

(2) There is an $n \times n$ matrix $C$ such that $CA = I$.

(6) There is an $n \times n$ matrix $D$ such that $AD = I$. 
(3) The equation $A\mathbf{x} = \mathbf{0}$ only has the trivial solution.

(7) The equation $A\mathbf{x} = \mathbf{b}$ (for any $\mathbf{b}$) has at least one solution.

(8) The columns of $A$ span $\mathbb{R}^n$.

(11) The columns of $A$ form a linearly independent set.

(9) The linear transformation $f(\mathbf{x}) = A\mathbf{x}$ (from $\mathbb{R}^n$ to $\mathbb{R}^n$) is onto.

(12) The linear transformation $f(\mathbf{x}) = A\mathbf{x}$ is one-to-one.
Finding matrix inverse

First consider $2 \times 2$ matrices

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

If $ad - bc \neq 0$, then $A$ is invertible and its inverse is given by the following empirical rule

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

Example 0.8. Use the above rule to find the inverse of

$$A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$$
In general, given an invertible matrix $A \in \mathbb{R}^{n \times n}$ (for any $n$), finding its inverse is equivalent to solving the matrix equation

$$AX = I_n, \quad \text{or equivalently} \quad A[x_1, \ldots, x_n] = [e_1, \ldots, e_n]$$

This leads to $n$ separate systems of linear equations:

$$Ax_1 = e_1 \quad (\text{i.e. } [A \mid e_1]), \quad \ldots, \quad Ax_n = e_n \quad (\text{i.e. } [A \mid e_n]).$$

which may be solved simultaneously:

$$[A \mid [e_1, \ldots, e_n]] = [A \mid I_n] \longrightarrow [I_n \mid A^{-1}].$$
Example 0.9. Find the inverse of the matrix

\[
\mathbf{A} = \begin{bmatrix}
1 & 0 & -2 \\
3 & 1 & -2 \\
-5 & -1 & 9
\end{bmatrix},
\]

if it exists.
Partitioned matrices

A partitioned matrix, also called a block matrix, is a matrix whose elements have been divided into blocks (called submatrices).

For example,

\[
A = \begin{bmatrix}
1 & 2 & 3 & 0 & 0 \\
4 & 5 & 6 & 0 & 0 \\
0 & 0 & 0 & 7 & 8 \\
1 & 1 & 1 & 0 & 0 \\
2 & 2 & 2 & 0 & 0 \\
3 & 3 & 3 & 0 & 0
\end{bmatrix} = \begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22} \\
A_{31} & A_{32}
\end{bmatrix}
\]

Partitioned matrices are very useful because they reduce large matrices into a collection of smaller matrices (which are easier to deal with).
Addition and scalar multiplication

If two matrices $A, B$ have the same size and have been partitioned in exactly the same way, then we can just add the corresponding blocks to get their sum (with the same partition):

$$A + B = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix} = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \\ A_{31} + B_{31} & A_{32} + B_{32} \end{bmatrix}$$

The scalar multiple of a partitioned matrix is

$$rA = \begin{bmatrix} rA_{11} & rA_{12} \\ rA_{21} & rA_{22} \\ rA_{31} & rA_{32} \end{bmatrix}$$
Multiplication of partitioned matrices: simple cases

Let $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}$ be two matrices that may be multiplied together.

When the columns of $A$ and rows of $B$ are divided in a conformable way, we can carry out block multiplication:

\[
A B = A_{11} B_{11} + A_{12} B_{21} + A_{13} B_{31}
\]

**Remark.**

- All terms $AB, A_{11}B_{11}, A_{12}B_{21}, A_{13}B_{31}$ are $m \times p$ matrices.
- Such partitions do not show up in the product matrix.
Example 0.10. Let

\[ A = \begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 & 0 \\ 7 & 8 & 9 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} \]

Find \( AB \) using two ways: (a) direct multiplication (b) block multiplication.

Answer.

\[ AB = \begin{bmatrix} 6 & -6 \\ 15 & -15 \\ 24 & -24 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \]
A joke

How does a mathematician change three light bulbs at the same time?

He gives them to three engineers and ask them to do it in parallel.
Let \( A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p} \) be two matrices that are partitioned in a conformable way (i.e., column partition of \( A \) matches row partition of \( B \)).

Regardless of the row partition of \( A \) and column partition of \( B \), we can carry out block multiplications by treating the blocks as numbers.

**Remark.** Row partition of \( A \ + \) column partition of \( B \) = partition of \( AB \) (such two partitions do not need to match).
In terms of math symbols, that is

\[
\mathbf{AB} = \begin{bmatrix}
A_{11} & A_{12} & A_{13} \\
A_{21} & A_{22} & A_{23} \\
A_{31} & A_{32} & A_{33}
\end{bmatrix} \cdot \begin{bmatrix}
B_{11} & B_{12} \\
B_{21} & B_{22} \\
B_{31} & B_{32}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} & A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \\
A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} & A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} \\
A_{31}B_{11} + A_{32}B_{21} + A_{33}B_{31} & A_{31}B_{12} + A_{32}B_{22} + A_{33}B_{32}
\end{bmatrix}
\]

In the above, we can think of \( \mathbf{A} \) as a \( 3 \times 3 \) partitioned matrix and \( \mathbf{B} \) as a \( 3 \times 2 \) partitioned matrix, so that we must obtain a \( 3 \times 2 \) partitioned matrix.
Example 0.11. Verify that

\[
\begin{bmatrix}
1 & 2 & 3 & 0 & 0 \\
4 & 5 & 6 & 0 & 0 \\
7 & 8 & 9 & 0 & 0
\end{bmatrix}
\cdot
\begin{bmatrix}
1 & -1 \\
1 & -1 \\
1 & -1 \\
1 & -1
\end{bmatrix}
= 
\begin{bmatrix}
6 & -6 \\
15 & -15 \\
24 & -24
\end{bmatrix}
\]
Example 0.12. Show that

\[
\begin{bmatrix}
U_1 & U_2
\end{bmatrix}
\begin{bmatrix}
\Sigma & O \\
O & O
\end{bmatrix}
\begin{bmatrix}
V_1 \\
V_2
\end{bmatrix} = U_1 \Sigma V_1
\]

(assuming all submatrices are compatible with each other)
Matrix multiplication again

The columnwise multiplication of two compatible matrices $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$ actually has already used simple partitions of matrices:

$$AB = A[b_1 \ldots b_p] = [Ab_1 \ldots Ab_p]$$
We present two new ways of performing matrix multiplication:

- **Rowwise** multiplication

\[
\begin{align*}
AB &= \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} B = \begin{bmatrix} A_1 B \\ \vdots \\ A_m B \end{bmatrix} \\
\end{align*}
\]

where \( A_1, \ldots, A_m \) are the rows of \( A \).
**Column-row expansion**

\[
AB = \begin{bmatrix} a_1 & \ldots & a_n \end{bmatrix} \begin{bmatrix} B_1 \\ \vdots \\ B_n \end{bmatrix} = a_1 B_1 + \cdots + a_n B_n
\]
Example 0.13. Find the product of $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$ by using three different ways:

(a) Columnwise multiplication

(b) Rowwise multiplication and

(c) Column-row multiplication
Block diagonal matrices

**Def 0.8.** A matrix is said to be **block diagonal** if it is of the form

\[
A = \begin{bmatrix}
A_{11} & \\
& A_{22}
\end{bmatrix}
\]

**Example 0.14.**

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9
\end{bmatrix}
\begin{bmatrix}
1 & 1 \\
2 & 2
\end{bmatrix}
\]
**Theorem 0.14.** Let $A, B$ be two block diagonal matrices with conformable partitions:

\[
A = \begin{bmatrix} A_{11} & \quad & \\
& & \\
& & A_{22}
\end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & \quad & \\
& & \\
& & B_{22}
\end{bmatrix}
\]

Then we have

\[
AB = \begin{bmatrix} A_{11}B_{11} & \\
& \\
& A_{22}B_{22}
\end{bmatrix}.
\]

**Proof.** By direct verification. \(\square\)

**Remark.** This formula also generalizes to three or more blocks.
The previous result immediately implies the following.

**Theorem 0.15.** For a block diagonal matrix

\[
A = \begin{bmatrix} A_{11} & \newline A_{21} \\ \newline A_{12} & A_{22} \end{bmatrix},
\]

if the two blocks are both square and invertible, then \( A \) is also invertible. Moreover,

\[
A^{-1} = \begin{bmatrix} A_{11}^{-1} & \newline A_{21}^{-1} \\ \newline A_{12}^{-1} & A_{22}^{-1} \end{bmatrix}
\]

**Proof.** By direct verification. \( \square \)
Example 0.15. Find the inverse of

\[
\begin{bmatrix}
1 & 2 \\
1 & 3 \\
\hline
1 & 3
\end{bmatrix}
\]
Block upper triangular matrices

Def 0.9. A matrix is said to be **block upper triangular** if it is of the form

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

Example 0.16.

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9 \\
\hline
1 & 0 \\
0 & 1 \\
3 & 3 \\
1 & 1 \\
2 & 2
\end{bmatrix}
\]
Theorem 0.16. For a block upper triangular matrix

\[ \mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix}, \]

if the two main blocks are both square and invertible, then \( \mathbf{A} \) is also invertible, and

\[ \mathbf{A}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ A_{22}^{-1} \end{bmatrix} \]

Proof. By direct verification.
Example 0.17. Find the inverse of

\[
\begin{bmatrix}
1 & 2 & 1 \\
1 & 3 & 1 \\
1 & 3 & 1 \\
\end{bmatrix}
\]
LU decomposition

In this part, we will derive a factorization scheme to express a given matrix $A \in \mathbb{R}^{m \times n}$ as a product of two matrices of special forms

$$A = L \cdot U = \begin{bmatrix} 1 & * & * & * & * \\ * & 1 & * & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ * & * & * & 1 \end{bmatrix} \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * \\ * & * & * \\ * & * \end{bmatrix}$$

where $L \in \mathbb{R}^{m \times m}$ is square, lower-triangular with 1’s on the diagonal (called unit lower triangular), and $U \in \mathbb{R}^{m \times n}$ is the REF of $A$ (which is upper triangular).

Such a factorization is very useful for solving linear systems $Ax = b$. 
For example, the following is an LU decomposition (verify this):

\[
\begin{bmatrix}
3 & -7 & -2 \\
-3 & 5 & 1 \\
6 & -4 & 0
\end{bmatrix}
= 
\begin{bmatrix}
1 & 1 & 1 \\
-1 & 2 & -5 \\
0 & 0 & 1
\end{bmatrix}
\cdot 
\begin{bmatrix}
3 & -7 & -2 \\
-2 & -1 & 0 \\
-1 & 0 & 0
\end{bmatrix}
\]

To use it to solve the system of linear equations

\[Ax = b, \quad \text{where} \quad b = \begin{bmatrix} -7 & 5 & 2 \end{bmatrix}^T \]

we first rewrite the equation as

\[Ax = (LU)x = L(Ux) = b\]
and then solve two simper systems in the order

\[ Ly = b \quad \rightarrow \quad Ux = y \]

That is, from the first equation, we obtain that
\[ y = \begin{bmatrix} -7 & -2 & 6 \end{bmatrix}^T \]
and then use it to solve the second equation for
\[ x = \begin{bmatrix} 3 & 4 & -6 \end{bmatrix}^T \]
(work done in class).

Verify:
\[
\begin{bmatrix}
  3 & -7 & -2 \\
  -3 & 5 & 1 \\
  6 & -4 & 0
\end{bmatrix}
\begin{bmatrix}
  3 \\
  4 \\
  -6
\end{bmatrix} =
\begin{bmatrix}
  -7 \\
  -5 \\
  2
\end{bmatrix}.
\]

However, how to find such a decomposition in the first place will require the introduction of the so-called **elementary matrices**.
Elementary matrices

Elementary matrices are (square) matrices that can be obtained from the identity matrix through a single elementary row operation.

\[
\begin{align*}
M_i(r) & = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\
R_{i\leftarrow j}(k) & = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\
P_{ij} & = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}
\end{align*}
\]
Performing an elementary row operation on a given matrix can now is equivalent to matrix multiplication (the elementary matrix left multiplies the given matrix).

- $M_i(r)$ - Multiply row $i$ by a nonzero scalar $r$

\[
M_3(r)A = \begin{bmatrix} 1 & 1 & r \\ a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ r a_{31} & r a_{32} & r a_{33} & r a_{34} \end{bmatrix}
\]
Matrix Algebra

- $\mathbf{R}_{i\leftarrow j}(k)$ - Add a scalar multiple ($k$) of one row ($j$) to another row ($i$) to replace that row ($i$):

  - Downward replacement

$$\mathbf{R}_{3\leftarrow 1}(k)\mathbf{A} = \begin{bmatrix} 1 & 1 & a_{11} & a_{12} & a_{13} & a_{14} \\ k & 1 & a_{21} & a_{22} & a_{23} & a_{24} \\ & & a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ ka_{11} + a_{31} & ka_{12} + a_{32} & ka_{13} + a_{33} & ka_{14} + a_{34} \end{bmatrix}$$
– Upward replacement

$$\mathbf{R}_{1\leftarrow 3}(k) \mathbf{A} = \begin{bmatrix} 1 & k \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{11} + ka_{31} & a_{12} + ka_{32} & a_{13} + ka_{33} & a_{14} + ka_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$
Matrix Algebra

- Interchange two rows

\[
P_{12}A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{23} & a_{24} \\ a_{11} & a_{12} & a_{13} & a_{14} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}
\]

\[
P_{13}A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{11} & a_{12} & a_{13} & a_{14} \end{bmatrix}
\]

\[
P_{23}A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix}
\]
An important fact

Elementary matrices are all invertible (because elementary row operations are all reversible)

\[
M_i\left(\frac{1}{r}\right) \cdot M_i(r) = I \\
R_{i \leftarrow j}(-k) \cdot R_{i \leftarrow j}(k) = I \\
P_{ij} \cdot P_{ij} = I
\]

and their inverses are the same kind of elementary matrices!

\[
M_i(r)^{-1} = M_i\left(\frac{1}{r}\right) \\
R_{i \leftarrow j}(k)^{-1} = R_{i \leftarrow j}(-k) \\
P_{ij}^{-1} = P_{ij}
\]
Application of elementary matrices in finding matrix inverse

Previously we presented a procedure for finding the inverse of a square, invertible matrix

\[
\begin{bmatrix}
A & I_n
\end{bmatrix}
\xrightarrow{\text{elementary row operations}}
\begin{bmatrix}
I_n & A^{-1}
\end{bmatrix}
\]

This is equivalent to using a sequence of elementary matrices \(E_1, E_2, \ldots, E_\ell\) to left multiply the augmented matrix:

\[
E_\ell \cdots E_2 E_1 \cdot \begin{bmatrix}
A & I_n
\end{bmatrix} = \begin{bmatrix}
I_n & A^{-1}
\end{bmatrix}
\]

Through matrix block multiplication, we obtain

\[
\begin{bmatrix}
E_\ell \cdots E_2 E_1 A & E_\ell \cdots E_2 E_1
\end{bmatrix} = \begin{bmatrix}
I_n & A^{-1}
\end{bmatrix}
\]

This shows that

\[
A^{-1} = E_\ell \cdots E_2 E_1
\]
Application of elementary matrices in finding matrix REF

Similarly, give any matrix \( A \in \mathbb{R}^{m \times n} \), one can perform a sequence of elementary row operations through corresponding elementary matrices \( E_1, E_2, \ldots, E_\ell \) to transform the given matrix into its REF

\[
E_\ell \cdots E_2 E_1 A = U
\]

This yields that

\[
A = (E_\ell \cdots E_2 E_1)^{-1} U = E_1^{-1} E_2^{-1} \cdots E_\ell^{-1} U
\]

Note that \( U \) (as REF) must be upper triangular.
Existence of the LU decomposition

In some cases, one only needs to use a sequence of downward replacement operations (i.e., $R_{i\leftarrow j}(k)$ for $j < i$) to transform a matrix $A \in \mathbb{R}^{m \times n}$ into its REF $U \in \mathbb{R}^{m \times n}$. That is,

$$
A = \underbrace{E_\ell \cdots E_2 E_1}_{\text{all downward replacements}} \quad A = U
$$

Then

$$
A = \underbrace{E_1^{-1} E_2^{-1} \cdots E_\ell^{-1}}_{\text{also downward replacements}} \quad U = \underbrace{L}_{\text{lower triangular}} \underbrace{U}_{\text{REF}}
$$

Remark. In other cases, one can always rearrange the rows of $A$ in a way such that an LU decomposition exists.
Finding the $L$ matrix

When a matrix $A \in \mathbb{R}^{m \times n}$ has an LU decomposition, we can find it as follows:

$$E_\ell \cdots E_2 E_1 A = \underbrace{U}_{\text{REF}}$$

$$E_\ell \cdots E_2 E_1 L = I$$

\[\leftarrow L = E_1^{-1} E_2^{-1} \cdots E_\ell^{-1}\]

That is, we will try to design a matrix $L$ (lower triangular with 1’s on the diagonal) so that the same row operations performed on $A$ toward its REF will transform $L$ into the identity matrix.
Example 0.18. Find the LU decomposition of

\[ A = \begin{bmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{bmatrix} \]
Example 0.19. Find the LU decomposition of

$$A = \begin{bmatrix} 1 & -2 & -4 & -3 \\ 2 & -7 & -7 & -6 \\ -1 & 2 & 6 & 4 \\ -4 & -1 & 9 & 8 \end{bmatrix}$$