# Linear Algebra

Department of Mathematics
Indian Institute of Technology Guwahati

January – May 2019

MA 102 (RA, RKS, MGPP, KVK)

# Matrix Operations

Topics:

- Matrix operations
- Invertible matrices
- Elementary matrices and reduction to rref
- Gauss-Jordan elimination for computing in inverse of a matrix
- LU factorization

# Matrices

Recall that an $m \times n$ matrix $A$ with entries $a_{ij}$ has $m$ rows and $n$ columns and can be written as

$$
\begin{aligned}
A &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = [a_{ij}]_{m\times n},
\end{aligned}
$$

# Matrices

Recall that an $m \times n$ matrix $A$ with entries $a_{ij}$ has $m$ rows and $n$ columns and can be written as

$$
\begin{aligned}
A &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = [a_{ij}]_{m\times n},
\end{aligned}
$$

where $\mathbf{A}_i := [a_{i1}, a_{i2}, \dots, a_{in}]$ is the $i$-th row of $A$ for $i = 1 : m$ and $\mathbf{a}_j := [a_{1j}, \dots, a_{mj}]^\top$ is the $j$-th column of $A$ for $j = 1 : n$.

# Matrices

Recall that an $m \times n$ matrix $A$ with entries $a_{ij}$ has $m$ rows and $n$ columns and can be written as

$$\begin{aligned} A &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = [a_{ij}]_{m\times n}, \end{aligned}$$

where $\mathbf{A}_i := [a_{i1}, a_{i2}, \ldots, a_{in}]$ is the $i$-th row of $A$ for $i = 1 : m$ and $\mathbf{a}_j := [a_{1j}, \ldots, a_{mj}]^\top$ is the $j$-th column of $A$ for $j = 1 : n$.

Transpose: The transpose of $A$ denoted by $A^\top$ is the $n \times m$ matrix given by $A^\top = [a_{ji}]_{n\times m}$.

# Matrices

Recall that an $m \times n$ matrix $A$ with entries $a_{ij}$ has $m$ rows and $n$ columns and can be written as

$$
\begin{aligned}
A &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = [a_{ij}]_{m \times n},
\end{aligned}
$$

where $\mathbf{A}_i := [a_{i1}, a_{i2}, \ldots, a_{in}]$ is the $i$-th row of $A$ for $i = 1 : m$ and $\mathbf{a}_j := [a_{1j}, \ldots, a_{mj}]^\top$ is the $j$-th column of $A$ for $j = 1 : n$.

Transpose: The transpose of $A$ denoted by $A^\top$ is the $n \times m$ matrix given by $A^\top = [a_{ji}]_{n \times m}$.

Two matrices $A = [a_{ij}]$ and $B := [b_{ij}]$ are said to be equal (i.e, $A = B$) if $A$ and $B$ have the same size and $a_{ij} = b_{ij}$ for all $i$ and $j$.

# Special matrices

We denote the set of complex numbers by $\mathbb{C}$.

A matrix $A := [a_{ij}]$ with entries $a_{ij} \in \mathbb{C}$ (resp.,$a_{ij} \in \mathbb{R}$) is said to be a complex (resp.,real ) matrix.

# Special matrices

We denote the set of complex numbers by $\mathbb{C}$.

A matrix $A := [a_{ij}]$ with entries $a_{ij} \in \mathbb{C}$ (resp.,$a_{ij} \in \mathbb{R}$) is said to be a complex (resp.,real ) matrix.

Conjugate transpose: The conjugate transpose of an $m \times n$ complex matrix $A = [a_{ij}]_{m \times n}$ is the $n \times m$ matrix denoted by $A^*$ and is given by

$$A^* = [\bar{a}_{ji}]_{n \times m} = ([\bar{a}_{ij}]_{m \times n})^\top = (\bar{A})^\top,$$

where $\bar{a}_{ij}$ is the conjugate of $a_{ij}$.

# Special matrices

We denote the set of complex numbers by $\mathbb{C}$.

A matrix $A := [a_{ij}]$ with entries $a_{ij} \in \mathbb{C}$ (resp.,$a_{ij} \in \mathbb{R}$) is said to be a complex (resp.,real ) matrix.

Conjugate transpose: The conjugate transpose of an $m \times n$ complex matrix $A = [a_{ij}]_{m\times n}$ is the $n \times m$ matrix denoted by $A^*$ and is given by

$$A^* = [\bar{a}_{ji}]_{n\times m} = ([\bar{a}_{ij}]_{m\times n})^\top = (\bar{A})^\top,$$

where $\bar{a}_{ij}$ is the conjugate of $a_{ij}$.

Definition: Let $A$ be an $n \times n$ matrix. Then $A$ is said to be

1. symmetric if $A^\top = A$,
2. skew-symmetric if $A^\top = -A$,

# Special matrices

We denote the set of complex numbers by $\mathbb{C}$.

A matrix $A := [a_{ij}]$ with entries $a_{ij} \in \mathbb{C}$ (resp.,$a_{ij} \in \mathbb{R}$) is said to be a complex (resp.,real ) matrix.

Conjugate transpose: The conjugate transpose of an $m \times n$ complex matrix $A = [a_{ij}]_{m\times n}$ is the $n \times m$ matrix denoted by $A^*$ and is given by

$$A^* = [\bar{a}_{ji}]_{n\times m} = ([\bar{a}_{ij}]_{m\times n})^\top = (\bar{A})^\top,$$

where $\bar{a}_{ij}$ is the conjugate of $a_{ij}$.

Definition: Let $A$ be an $n \times n$ matrix. Then $A$ is said to be

1. symmetric if $A^\top = A$,
2. skew-symmetric if $A^\top = -A$,
3. Hermitian if $A^* = A$,
4. skew-Hermitian if $A^* = -A$.

# Special matrices (recall)

Let $A$ be an $m \times n$ matrix with $(i,j)$-th entry $a_{ij}$. Set $p := \min(m,n)$. Then

- $a_{ii}$ for $i = 1 : p$ are called the diagonal entries of $A$;
- $A$ is said to be a diagonal matrix if $a_{ij} = 0$ for all $i \neq j$;

# Special matrices (recall)

Let $A$ be an $m \times n$ matrix with $(i, j)$-th entry $a_{ij}$. Set $p := \min(m, n)$. Then

- $a_{ii}$ for $i = 1 : p$ are called the diagonal entries of $A$;
- $A$ is said to be a diagonal matrix if $a_{ij} = 0$ for all $i \neq j$;
- $A$ is said to be an upper triangular if $a_{ij} = 0$ for all $i > j$;
- $A$ is said to be a lower triangular if $a_{ij} = 0$ for all $i < j$;

Identity matrix: An $n \times n$ diagonal matrix with all diagonal entries equal to 1 is called the identity matrix and is denoted by $I_n$ or $I$.

Zero matrix: An $m \times n$ matrix with all entries 0 is called the zero matrix and is denoted by $\mathbf{O}_{m\times n}$ or simply by $\mathbf{O}$.

## Matrix Operations

- $\mathcal{M}_{m\times n}$ := the set of a $m \times n$ matrices.

## Matrix Operations

- $\mathcal{M}_{m\times n}$ := the set of a $m \times n$ matrices.

- To specify real (resp., complex) matrices, write $\mathcal{M}_{m\times n}(\mathbb{R})$ (resp., $\mathcal{M}_{m\times n}(\mathbb{C})$).

# Matrix Operations

- $\mathcal{M}_{m\times n} :=$ the set of a $m \times n$ matrices.
- To specify real (resp., complex) matrices, write $\mathcal{M}_{m\times n}(\mathbb{R})$ (resp., $\mathcal{M}_{m\times n}(\mathbb{C})$).
- If $m = n$, we write $\mathcal{M}_n$ for $\mathcal{M}_{m\times n}$.

# Matrix Operations

- $\mathcal{M}_{m\times n} :=$ the set of a $m \times n$ matrices.
- To specify real (resp., complex) matrices, write $\mathcal{M}_{m\times n}(\mathbb{R})$ (resp., $\mathcal{M}_{m\times n}(\mathbb{C})$).
- If $m = n$, we write $\mathcal{M}_n$ for $\mathcal{M}_{m\times n}$.
- Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be in $\mathcal{M}_{m\times n}$ and $\alpha$ be a scalar. Then

# Matrix Operations

- $\mathcal{M}_{m\times n}$ := the set of a $m \times n$ matrices.
- To specify real (resp., complex) matrices, write $\mathcal{M}_{m\times n}(\mathbb{R})$ (resp., $\mathcal{M}_{m\times n}(\mathbb{C})$).
- If $m = n$, we write $\mathcal{M}_n$ for $\mathcal{M}_{m\times n}$.
- Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be in $\mathcal{M}_{m\times n}$ and $\alpha$ be a scalar. Then

  1. Matrix addition: $A + B := [a_{ij} + b_{ij}] \in \mathcal{M}_{m\times n}$.

# Matrix Operations

- $\mathcal{M}_{m\times n} :=$ the set of a $m \times n$ matrices.
- To specify real (resp., complex) matrices, write $\mathcal{M}_{m\times n}(\mathbb{R})$ (resp., $\mathcal{M}_{m\times n}(\mathbb{C})$).
- If $m = n$, we write $\mathcal{M}_n$ for $\mathcal{M}_{m\times n}$.
- Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be in $\mathcal{M}_{m\times n}$ and $\alpha$ be a scalar. Then

  1. Matrix addition: $A + B := [a_{ij} + b_{ij}] \in \mathcal{M}_{m\times n}$.
  2. Multiplication by a scalar: $\alpha A := [\alpha\, a_{ij}] \in \mathcal{M}_{m\times n}$. ($\alpha \in \mathbb{R}$ when $A$ and $B$ are real matrices, and $\alpha \in \mathbb{C}$ when $A$ and $B$ are complex matrices)

# Properties of addition and scalar multiplication

Theorem: Let $A, B, C \in \mathcal{M}_{m\times n}$ and $\alpha, \beta$ be scalars. Then

# Properties of addition and scalar multiplication

Theorem: Let $A, B, C \in \mathcal{M}_{m \times n}$ and $\alpha, \beta$ be scalars. Then

1. Commutative Law: $A + B = B + A$.
2. Associative Law: $(A + B) + C = A + (B + C)$.

# Properties of addition and scalar multiplication

Theorem: Let $A, B, C \in \mathcal{M}_{m\times n}$ and $\alpha, \beta$ be scalars. Then

1. Commutative Law: $A + B = B + A$.
2. Associative Law: $(A + B) + C = A + (B + C)$.
3. $A + \mathbf{O} = A$, where $\mathbf{O} = \mathbf{O}_{m\times n} \in \mathcal{M}_{m\times n}$.
4. $A + (-A) = \mathbf{O}$, where $-A = (-1)A \in \mathcal{M}_{m\times n}$.

# Properties of addition and scalar multiplication

Theorem: Let $A, B, C \in \mathcal{M}_{m \times n}$ and $\alpha, \beta$ be scalars. Then

1. Commutative Law: $A + B = B + A$.
2. Associative Law: $(A + B) + C = A + (B + C)$.
3. $A + \mathbf{O} = A$, where $\mathbf{O} = \mathbf{O}_{m \times n} \in \mathcal{M}_{m \times n}$.
4. $A + (-A) = \mathbf{O}$, where $-A = (-1)A \in \mathcal{M}_{m \times n}$.
5. $\alpha(A + B) = \alpha\, A + \alpha\, B$.
6. $(\alpha + \beta)A = \alpha\, A + \beta\, A$.

# Properties of addition and scalar multiplication

Theorem: Let $A, B, C \in \mathcal{M}_{m\times n}$ and $\alpha, \beta$ be scalars. Then

1. Commutative Law: $A + B = B + A$.
2. Associative Law: $(A + B) + C = A + (B + C)$.
3. $A + \mathbf{O} = A$, where $\mathbf{O} = \mathbf{O}_{m\times n} \in \mathcal{M}_{m\times n}$.
4. $A + (-A) = \mathbf{O}$, where $-A = (-1)A \in \mathcal{M}_{m\times n}$.
5. $\alpha(A + B) = \alpha\, A + \alpha\, B$.
6. $(\alpha + \beta)A = \alpha\, A + \beta\, A$.
7. $\alpha\,(\beta\, A) = (\alpha\beta)A$.
8. $1A = A$.

# Matrix-vector multiplication revisited

Let $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \in \mathcal{M}_{m\times n}(\mathbb{R})$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$. Recall the matrix-vector multiplication

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n. \in \mathbb{R}^m.$$

# Matrix-vector multiplication revisited

Let $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \in \mathcal{M}_{m\times n}(\mathbb{R})$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$. Recall the matrix-vector multiplication

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n. \in \mathbb{R}^m.$$

This shows that $A$ act on $\mathbf{x}$ as a well defined function

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \mathbf{x} \longmapsto A\mathbf{x},$$

and satisfies following:

# Matrix-vector multiplication revisited

Let $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \in \mathcal{M}_{m\times n}(\mathbb{R})$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$. Recall the matrix-vector multiplication

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n. \in \mathbb{R}^m.$$

This shows that $A$ act on $\mathbf{x}$ as a well defined function

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \mathbf{x} \longmapsto A\mathbf{x},$$

and satisfies following:

- $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y}$ for all $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$.
- $A(\alpha\mathbf{x}) = \alpha A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

# Matrix-vector multiplication revisited

Let $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \in \mathcal{M}_{m\times n}(\mathbb{R})$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$. Recall the matrix-vector multiplication

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n. \in \mathbb{R}^m.$$

This shows that $A$ act on $\mathbf{x}$ as a well defined function

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \mathbf{x} \longmapsto A\mathbf{x},$$

and satisfies following:

- $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$ for all $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$.
- $A(\alpha\mathbf{x}) = \alpha A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

We refer to $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \mathbf{x} \longmapsto A\mathbf{x}$, as a linear mapping.

# Matrix-matrix multiplication

Let $A \in \mathcal{M}_{m\times n}$ and $B := \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix} \in \mathcal{M}_{n\times p}$.

Definition: Define the matrix-matrix multiplication $AB$ by

$$AB := \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_p \end{bmatrix}.$$

# Matrix-matrix multiplication

Let $A \in \mathcal{M}_{m\times n}$ and $B := \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix} \in \mathcal{M}_{n\times p}$.

Definition: Define the matrix-matrix multiplication $AB$ by

$$AB := \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_p \end{bmatrix}.$$

Justification: Define $AB$ to be the composition of the maps

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \mathbf{x} \mapsto A\mathbf{x}, \text{ and } B : \mathbb{R}^p \longrightarrow \mathbb{R}^n, \mathbf{y} \mapsto B\mathbf{y}.$$

# Matrix-matrix multiplication

Let $A \in \mathcal{M}_{m\times n}$ and $B := \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix} \in \mathcal{M}_{n\times p}$.

Definition: Define the matrix-matrix multiplication $AB$ by

$$AB := \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_p \end{bmatrix}.$$

Justification: Define $AB$ to be the composition of the maps

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \mathbf{x} \mapsto A\mathbf{x}, \text{ and } B : \mathbb{R}^p \longrightarrow \mathbb{R}^n, \mathbf{y} \mapsto B\mathbf{y}.$$

Then the linear map $AB : \mathbb{R}^p \longrightarrow \mathbb{R}^m$ is given by

$$\begin{aligned} AB\mathbf{y} &:= A(B\mathbf{y}) = A(y_1\mathbf{b}_1 + \cdots + y_p\mathbf{b}_p) \\ &= y_1 A\mathbf{b}_1 + \cdots + y_p A\mathbf{b}_p = \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_p \end{bmatrix} \mathbf{y} \end{aligned}$$

for all $\mathbf{y} \in \mathbb{R}^p$

# Matrix-matrix multiplication

Let $A \in \mathcal{M}_{m\times n}$ and $B := \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix} \in \mathcal{M}_{n\times p}$.

Definition: Define the matrix-matrix multiplication $AB$ by

$$AB := \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_p \end{bmatrix}.$$

Justification: Define $AB$ to be the composition of the maps

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \mathbf{x} \mapsto A\mathbf{x}, \text{ and } B : \mathbb{R}^p \longrightarrow \mathbb{R}^n, \mathbf{y} \mapsto B\mathbf{y}.$$

Then the linear map $AB : \mathbb{R}^p \longrightarrow \mathbb{R}^m$ is given by

$$\begin{aligned} AB\mathbf{y} &:= A(B\mathbf{y}) = A(y_1\mathbf{b}_1 + \cdots + y_p\mathbf{b}_p) \\ &= y_1A\mathbf{b}_1 + \cdots + y_pA\mathbf{b}_p = \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_p \end{bmatrix}\mathbf{y} \end{aligned}$$

for all $\mathbf{y} \in \mathbb{R}^p \Longrightarrow AB = \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_p \end{bmatrix}$.

# Matrix multiplication

Let $A = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} \in \mathcal{M}_{m\times n}$, $B := \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix} \in \mathcal{M}_{n\times p}$. Then

$$AB = \begin{bmatrix} \mathbf{A}_1\mathbf{b}_1 & \cdots & \mathbf{A}_1\mathbf{b}_p \\ \vdots & \ldots & \vdots \\ \mathbf{A}_m\mathbf{b}_1 & \cdots & \mathbf{A}_1\mathbf{b}_p \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 B \\ \vdots \\ \mathbf{A}_m B \end{bmatrix}.$$

# Matrix multiplication

Let $A = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} \in \mathcal{M}_{m\times n}$, $B := \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix} \in \mathcal{M}_{n\times p}$. Then

$$AB = \begin{bmatrix} \mathbf{A}_1\mathbf{b}_1 & \cdots & \mathbf{A}_1\mathbf{b}_p \\ \vdots & \ldots & \vdots \\ \mathbf{A}_m\mathbf{b}_1 & \cdots & \mathbf{A}_1\mathbf{b}_p \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 B \\ \vdots \\ \mathbf{A}_m B \end{bmatrix}.$$

Thus if $A := [a_{ij}]_{m\times n}$, $B := [b_{ij}]_{n\times p}$ and $C := AB = [c_{ij}]_{m\times p}$ then

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

# Matrix multiplication

Let $A = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} \in \mathcal{M}_{m\times n}$, $B := \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix} \in \mathcal{M}_{n\times p}$. Then

$$AB = \begin{bmatrix} \mathbf{A}_1\mathbf{b}_1 & \cdots & \mathbf{A}_1\mathbf{b}_p \\ \vdots & \dots & \vdots \\ \mathbf{A}_m\mathbf{b}_1 & \cdots & \mathbf{A}_1\mathbf{b}_p \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 B \\ \vdots \\ \mathbf{A}_m B \end{bmatrix}.$$

Thus if $A := [a_{ij}]_{m\times n}$, $B := [b_{ij}]_{n\times p}$ and $C := AB = [c_{ij}]_{m\times p}$ then

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

Note:

$$\begin{aligned} \mathbf{e}_i^\top A &= \mathbf{A}_i = i\text{-th row of } A, \text{ where } \mathbf{e}_i \in \mathbb{R}^m \\ A\mathbf{e}_j &= \mathbf{a}_j = j\text{-th column of } A, \text{ where } \mathbf{e}_j \in \mathbb{R}^n. \end{aligned}$$

# Matrix multiplication

Thoerm: Let $A, B$ and $C$ be matrices, and let $\alpha \in \mathbb{R}$. Then

# Matrix multiplication

Thoerm: Let $A, B$ and $C$ be matrices, and let $\alpha \in \mathbb{R}$. Then

1. Associative Law: $(AB)C = A(BC)$, if the respective matrix products are defined.

2. Distributive Law:
   $A(B + C) = AB + AC, \quad (A + B)C = AC + BC,$
   if the respective matrix sum and matrix products are defined.

# Matrix multiplication

Thoerm: Let $A, B$ and $C$ be matrices, and let $\alpha \in \mathbb{R}$. Then

1. Associative Law: $(AB)C = A(BC)$, if the respective matrix products are defined.
2. Distributive Law:
$A(B+C) = AB + AC, \quad (A+B)C = AC + BC,$
if the respective matrix sum and matrix products are defined.
3. $\alpha(AB) = (\alpha A)B = A(\alpha B)$, if the respective matrix products are defined.
4. $I_m A = A = AI_n$, if $A$ is of size $m \times n$.

# Block matrix

Definition: An $m \times n$ block matrix (or a partition matrix) is a matrix of the form

$$A := \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \cdots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}$$

where each $A_{ij}$ is a $p_i \times q_j$ matrix for $i = 1 : m$ and $j = 1 : n$.

# Block matrix

Definition: An $m \times n$ block matrix (or a partition matrix) is a matrix of the form

$$A := \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \cdots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}$$

where each $A_{ij}$ is a $p_i \times q_j$ matrix for $i = 1 : m$ and $j = 1 : n$.

Then $\begin{bmatrix} A_{i1} & \cdots & A_{in} \end{bmatrix}$ is the $i$-th block row of $A$ and $\begin{bmatrix} A_{1j} \\ \vdots \\ A_{mj} \end{bmatrix}$ is the $j$-th block column of $A$.

# Block matrix

Definition: An $m \times n$ block matrix (or a partition matrix) is a matrix of the form

$$A := \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \cdots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}$$

where each $A_{ij}$ is a $p_i \times q_j$ matrix for $i = 1 : m$ and $j = 1 : n$.

Then $\begin{bmatrix} A_{i1} & \cdots & A_{in} \end{bmatrix}$ is the $i$-th block row of $A$ and $\begin{bmatrix} A_{1j} \\ \vdots \\ A_{mj} \end{bmatrix}$ is the $j$-th block column of $A$.

Example: $\left[\begin{array}{cc|ccc|c} 1 & 2 & 2 & 0 & 1 & 4 \\ 3 & 4 & 1 & 2 & 3 & 5 \\ \hline 5 & 7 & 2 & 7 & 8 & 8 \\ 3 & 4 & 1 & 9 & 2 & 2 \end{array}\right]$ has 2 block rows and 3 block columns.

# Block matrix operations

Block matrix addition: Let $A := [A_{ij}]_{m\times n}$ and $B := [B_{ij}]_{m\times n}$ be block matrices such that size of $A_{ij}$ = size of $B_{ij}$ for $i = 1 : m$ and $j = 1 : n$. Then $A + B := [A_{ij} + B_{ij}]_{m\times n}$.

# Block matrix operations

Block matrix addition: Let $A := [A_{ij}]_{m \times n}$ and $B := [B_{ij}]_{m \times n}$ be block matrices such that size of $A_{ij}$ = size of $B_{ij}$ for $i = 1 : m$ and $j = 1 : n$. Then $A + B := [A_{ij} + B_{ij}]_{m \times n}$.

Block matrix multiplication: Let $A := [A_{ij}]_{m \times n}$ and $B := [B_{ij}]_{n \times p}$ be block matrices. If the matrix multiplication $C_{ij} := \sum_{k=1}^{n} A_{ik} B_{kj}$ is well defined for $i = 1 : m$ and $j = 1 : p$ then $AB$ is an $m \times p$ block matrix given by $AB = [C_{ij}]_{m \times p}$.

# Block matrix operations

Block matrix addition: Let $A := [A_{ij}]_{m\times n}$ and $B := [B_{ij}]_{m\times n}$ be block matrices such that size of $A_{ij}$ = size of $B_{ij}$ for $i = 1 : m$ and $j = 1 : n$. Then $A + B := [A_{ij} + B_{ij}]_{m\times n}$.

Block matrix multiplication: Let $A := [A_{ij}]_{m\times n}$ and $B := [B_{ij}]_{n\times p}$ be block matrices. If the matrix multiplication $C_{ij} := \sum_{k=1}^{n} A_{ik}B_{kj}$ is well defined for $i = 1 : m$ and $j = 1 : p$ then $AB$ is an $m \times p$ block matrix given by $AB = [C_{ij}]_{m\times p}$.

Conformal partition: If a block operation of block matrices $A$ and $B$ are well defined then $A$ and $B$ are said to be partitioned conformably.

# Block matrix operations

Block matrix addition: Let $A := [A_{ij}]_{m\times n}$ and $B := [B_{ij}]_{m\times n}$ be block matrices such that size of $A_{ij}$ = size of $B_{ij}$ for $i = 1 : m$ and $j = 1 : n$. Then $A + B := [A_{ij} + B_{ij}]_{m\times n}$.

Block matrix multiplication: Let $A := [A_{ij}]_{m\times n}$ and $B := [B_{ij}]_{n\times p}$ be block matrices. If the matrix multiplication $C_{ij} := \sum_{k=1}^{n} A_{ik}B_{kj}$ is well defined for $i = 1 : m$ and $j = 1 : p$ then $AB$ is an $m \times p$ block matrix given by $AB = [C_{ij}]_{m\times p}$.

Conformal partition: If a block operation of block matrices $A$ and $B$ are well defined then $A$ and $B$ are said to be partitioned conformably.

Example: $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix} =$

$$\begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} & A_{11}B_{13} + A_{12}B_{23} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} & A_{21}B_{13} + A_{22}B_{23} \end{bmatrix}.$$

# The inverse of a matrix

Definition: An $n \times n$ matrix $A$ is said to be invertible if there exists a matrix $B$ satisfying $AB = I_n = BA$. The matrix $B$ is called an inverse of $A$.

# The inverse of a matrix

Definition: An $n \times n$ matrix $A$ is said to be invertible if there exists a matrix $B$ satisfying $AB = I_n = BA$. The matrix $B$ is called an inverse of $A$.

Fact: Inverse of an invertible matrix $A$ is unique. (Exercise)

# The inverse of a matrix

Definition: An $n \times n$ matrix $A$ is said to be invertible if there exists a matrix $B$ satisfying $AB = I_n = BA$. The matrix $B$ is called an inverse of $A$.

Fact: Inverse of an invertible matrix $A$ is unique. (Exercise)
We denote the inverse of $A$ by $A^{-1}$.

# The inverse of a matrix

Definition: An $n \times n$ matrix $A$ is said to be invertible if there exists a matrix $B$ satisfying $AB = I_n = BA$. The matrix $B$ is called an inverse of $A$.

Fact: Inverse of an invertible matrix $A$ is unique. (Exercise)
We denote the inverse of $A$ by $A^{-1}$.

Note: We can talk of invertibility only of square matrices.

# The inverse of a matrix

Definition: An $n \times n$ matrix $A$ is said to be invertible if there exists a matrix $B$ satisfying $AB = I_n = BA$. The matrix $B$ is called an inverse of $A$.

Fact: Inverse of an invertible matrix $A$ is unique. (Exercise)
We denote the inverse of $A$ by $A^{-1}$.

Note: We can talk of invertibility only of square matrices.

- For example, the matrix $A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ is invertible since

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

# The inverse of a matrix

Definition: An $n \times n$ matrix $A$ is said to be invertible if there exists a matrix $B$ satisfying $AB = I_n = BA$. The matrix $B$ is called an inverse of $A$.

Fact: Inverse of an invertible matrix $A$ is unique. (Exercise)
We denote the inverse of $A$ by $A^{-1}$.

Note: We can talk of invertibility only of square matrices.

- For example, the matrix $A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ is invertible since

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

- The zero matrix $\mathbf{O}$ is not invertible.

# The inverse of a matrix

Definition: An $n \times n$ matrix $A$ is said to be invertible if there exists a matrix $B$ satisfying $AB = I_n = BA$. The matrix $B$ is called an inverse of $A$.

Fact: Inverse of an invertible matrix $A$ is unique. (Exercise)
We denote the inverse of $A$ by $A^{-1}$.

Note: We can talk of invertibility only of square matrices.

- For example, the matrix $A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ is invertible since

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

- The zero matrix $\mathbf{O}$ is not invertible.
- If $A$ has a zero row, then $A$ is not invertible.

# Properties of invertible matrices

**Fact:** Let $A$ and $B$ be two invertible matrices of the same size.

1. If $c \neq 0$ then $cA$ is also invertible, and $(cA)^{-1} = \frac{1}{c}A^{-1}$.
2. The matrix $AB$ is invertible, and $(AB)^{-1} = B^{-1}A^{-1}$.
3. The matrix $A^T$ is invertible, and $(A^T)^{-1} = (A^{-1})^T$.
4. For any non-negative integer $k$, the matrix $A^k$ is invertible, and $(A^k)^{-1} = (A^{-1})^k$.

# Properties of invertible matrices

**Fact:** Let $A$ and $B$ be two invertible matrices of the same size.

1. If $c \neq 0$ then $cA$ is also invertible, and $(cA)^{-1} = \frac{1}{c}A^{-1}$.
2. The matrix $AB$ is invertible, and $(AB)^{-1} = B^{-1}A^{-1}$.
3. The matrix $A^T$ is invertible, and $(A^T)^{-1} = (A^{-1})^T$.
4. For any non-negative integer $k$, the matrix $A^k$ is invertible, and $(A^k)^{-1} = (A^{-1})^k$.

Let $A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$ then $A$ is invertible, and

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

If $ad - bc = 0$ then $A$ is not invertible.

# Elementary matrices

| Type | Operation | Inverse operation |
|---|---|---|
| $I$ | $R_i \longleftarrow \alpha\, R_i$ | $R_i \longleftarrow \frac{1}{\alpha}\, R_i$ |
| $II$ | $R_j \longleftarrow cR_i + R_j$ | $R_j \longleftarrow -cR_i + R_j$ |
| $III$ | $R_i \leftrightarrow R_j$ | $R_i \leftrightarrow R_j$ |

# Elementary matrices

| Type | Operation | Inverse operation |
|---|---|---|
| $I$ | $R_i \longleftarrow \alpha\, R_i$ | $R_i \longleftarrow \frac{1}{\alpha}\, R_i$ |
| $II$ | $R_j \longleftarrow cR_i + R_j$ | $R_j \longleftarrow -cR_i + R_j$ |
| $III$ | $R_i \leftrightarrow R_j$ | $R_i \leftrightarrow R_j$ |

An elementary matrix is a matrix that is obtained by performing an elementary row operation on the identity matrix.

# Elementary matrices

| Type | Operation | Inverse operation |
|---|---|---|
| $I$ | $R_i \longleftarrow \alpha\, R_i$ | $R_i \longleftarrow \frac{1}{\alpha}\, R_i$ |
| $II$ | $R_j \longleftarrow cR_i + R_j$ | $R_j \longleftarrow -cR_i + R_j$ |
| $III$ | $R_i \leftrightarrow R_j$ | $R_i \leftrightarrow R_j$ |

An elementary matrix is a matrix that is obtained by performing an elementary row operation on the identity matrix.

$$\text{Type I}: E_2(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow E_2 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ \alpha x_2 \\ x_3 \end{bmatrix}.$$

$$E_2(\alpha)A = \begin{bmatrix} \mathrm{row}_1(A) \\ \alpha\, \mathrm{row}_2(A) \\ \mathrm{row}_3(A) \end{bmatrix} = \text{ multiply 2nd row of } A \text{ by } \alpha.$$

$$(E_2(\alpha))^{-1} = E_2(\frac{1}{\alpha}).$$

# Type II elementary matrices

$$\text{Type II} : E_{13}(2) := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} = I_3 + 2e_3 e_1^T$$

The matrix $E_{13}$ is obtained by performing $R_3 \leftarrow 2R_1 + R_3$ on $I_3$.

# Type II elementary matrices

$$\text{Type II} : E_{13}(2) := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} = I_3 + 2e_3e_1^T$$

The matrix $E_{13}$ is obtained by performing $R_3 \leftarrow 2R_1 + R_3$ on $I_3$.

$$E_{13}(2) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 + 2x_1 \end{bmatrix} \Rightarrow E_{13}A = \begin{bmatrix} \mathrm{row}_1(A) \\ \mathrm{row}_2(A) \\ \mathrm{row}_3(A) + 2\,\mathrm{row}_1(A) \end{bmatrix}.$$

$(E_{13}(2))^{-1} = E_{13}(-2)$ corresponds to $R_3 \leftarrow -2R_1 + R_3$ on $I_3$.

# Type III elementary matrices

Type III : $E_{ij}$ is obtained by performing $R_i \leftrightarrow R_j$ on $I$.

$$E_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \Rightarrow E_{23} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_3 \\ x_2 \end{bmatrix} \Rightarrow E_{23}A = \begin{bmatrix} \mathrm{row}_1(A) \\ \mathrm{row}_3(A) \\ \mathrm{row}_2(A) \end{bmatrix}.$$

# Type III elementary matrices

Type III : $E_{ij}$ is obtained by performing $R_i \leftrightarrow R_j$ on $I$.

$$E_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \Rightarrow E_{23} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_3 \\ x_2 \end{bmatrix} \Rightarrow E_{23}A = \begin{bmatrix} \mathrm{row}_1(A) \\ \mathrm{row}_3(A) \\ \mathrm{row}_2(A) \end{bmatrix}.$$

$$(E_{ij})^{-1} = E_{ij} \text{ corresponds to row operation } R_i \leftrightarrow R_j \text{ on } I.$$

**Observation:** Inverse of an elementary matrix is also an elementary matrix of same type.

# Row operation via elementary matrices

**Crux of the matter:**

- Type I: Multiplying $E_i(c)$ to $A$ giving $E_i(c)A$ amounts to performing the row operation $R_i \leftarrow cR_i$ on $A$.
- Type II: Multiplying $E_{ij}(c)$ to $A$ giving $E_{ij}(c)A$ amounts to performing the row operation $R_j \leftarrow cR_i + R_j$ on $A$..
- Type III: Multiplying $E_{ij}$ to $A$ giving $E_{ij}A$ amounts to performing the row operation $R_i \leftrightarrow R_j$ on $A$.

# Row operation via elementary matrices

**Crux of the matter:**

- Type I: Multiplying $E_i(c)$ to $A$ giving $E_i(c)A$ amounts to performing the row operation $R_i \leftarrow cR_i$ on $A$.
- Type II: Multiplying $E_{ij}(c)$ to $A$ giving $E_{ij}(c)A$ amounts to performing the row operation $R_j \leftarrow cR_i + R_j$ on $A$..
- Type III: Multiplying $E_{ij}$ to $A$ giving $E_{ij}A$ amounts to performing the row operation $R_i \leftrightarrow R_j$ on $A$.

Recall that two matrices $A$ and $B$ are said to be row equivalent if $A$ can be transformed to $B$ by elementary row operations.

Theorem: The matrices $A$ and $B$ are row equivalent $\iff$ $B = E_k \cdots E_2 E_1 A$ for some elementary matrices $E_1, E_2, \cdots, E_k$.

# Elementary matrices and echelon form

Forward GE: $m \times n$ matrix $A \longrightarrow$ row echelon form $\mathrm{ref}(A)$

$$\Downarrow$$

$\mathrm{ref}(A) = E_p \cdots E_2 E_1 A$ for some elementary matrices $E_1, \ldots, E_p$.

# Elementary matrices and echelon form

Forward GE: $m \times n$ matrix $A \longrightarrow$ row echelon form $\mathrm{ref}(A)$

$$\Downarrow$$

$\mathrm{ref}(A) = E_p \cdots E_2 E_1 A$ for some elementary matrices $E_1, \ldots, E_p$.

Gauss-Jordan: $A \longrightarrow$ reduced row echelon form $\mathrm{rref}(A)$

$$\Downarrow$$

$\mathrm{rref}(A) = \hat{E}_\ell \cdots \hat{E}_2 \hat{E}_1 A$ for some elementary matrices $\hat{E}_1, \ldots, \hat{E}_\ell$.

# Elementary matrices and echelon form

Forward GE: $m \times n$ matrix $A \longrightarrow$ row echelon form $\mathrm{ref}(A)$

$$\Downarrow$$

$\mathrm{ref}(A) = E_p \cdots E_2 E_1 A$ for some elementary matrices $E_1, \ldots, E_p$.

Gauss-Jordan: $A \longrightarrow$ reduced row echelon form $\mathrm{rref}(A)$

$$\Downarrow$$

$\mathrm{rref}(A) = \hat{E}_\ell \cdots \hat{E}_2 \hat{E}_1 A$ for some elementary matrices $\hat{E}_1, \ldots, \hat{E}_\ell$.

**Fact:** Let $[A \mid b] \longrightarrow \mathrm{ref}([A \mid b]) =: [U \mid d]$. Then the system $Ax = b$ and $Ux = d$ are equivalent.

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

$A$

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

$$A \xrightarrow[E_{12}]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$$

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

$$A \xrightarrow[E_{12}]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix} \xrightarrow[E_{13}(-2),\ E_{14}(-2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 2 & -3 & 7 \\ 0 & 1 & -4 & -3 \end{bmatrix}$$

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

$$A \xrightarrow[E_{12}]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix} \xrightarrow[E_{13}(-2),\ E_{14}(-2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 2 & -3 & 7 \\ 0 & 1 & -4 & -3 \end{bmatrix}$$

$$\xrightarrow[E_{32}(-1),\ E_{42}(-1/2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -2 & -5 \end{bmatrix}$$

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

$$A \xrightarrow[E_{12}]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix} \xrightarrow[E_{13}(-2),\ E_{14}(-2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 2 & -3 & 7 \\ 0 & 1 & -4 & -3 \end{bmatrix}$$

$$\xrightarrow[E_{32}(-1),\ E_{42}(-1/2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -2 & -5 \end{bmatrix} \xrightarrow[E_{43}(2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

$$A \xrightarrow[E_{12}]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix} \xrightarrow[E_{13}(-2),\ E_{14}(-2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 2 & -3 & 7 \\ 0 & 1 & -4 & -3 \end{bmatrix}$$

$$\xrightarrow[E_{32}(-1),\ E_{42}(-1/2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -2 & -5 \end{bmatrix} \xrightarrow[E_{43}(2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(Gaussian Elimination stops here.

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

$$A \xrightarrow[E_{12}]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix} \xrightarrow[E_{13}(-2),\ E_{14}(-2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 2 & -3 & 7 \\ 0 & 1 & -4 & -3 \end{bmatrix}$$

$$\xrightarrow[E_{32}(-1),\ E_{42}(-1/2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -2 & -5 \end{bmatrix} \xrightarrow[E_{43}(2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(Gaussian Elimination stops here. Continue with GJE.)

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

$$A \xrightarrow[E_{12}]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix} \xrightarrow[E_{13}(-2),\ E_{14}(-2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 2 & -3 & 7 \\ 0 & 1 & -4 & -3 \end{bmatrix}$$

$$\xrightarrow[E_{32}(-1),\ E_{42}(-1/2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -2 & -5 \end{bmatrix} \xrightarrow[E_{43}(2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(Gaussian Elimination stops here. Continue with GJE.)

$$\xrightarrow[E_{34}(-3),\ E_{24}(-4)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

$$A \xrightarrow[E_{12}]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix} \xrightarrow[E_{13}(-2),\ E_{14}(-2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 2 & -3 & 7 \\ 0 & 1 & -4 & -3 \end{bmatrix}$$

$$\xrightarrow[E_{32}(-1),\ E_{42}(-1/2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -2 & -5 \end{bmatrix} \xrightarrow[E_{43}(2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(Gaussian Elimination stops here. Continue with GJE.)

$$\xrightarrow[E_{34}(-3),\ E_{24}(-4)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow[E_{13}(-2), E_{23}(4)]{} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

# Example

Find rref of $A = \begin{bmatrix} 0 & 2 & -4 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix}$.

$$A \xrightarrow[E_{12}]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 2 & 2 & 1 & 7 \\ 2 & 1 & 0 & -3 \end{bmatrix} \xrightarrow[E_{13}(-2),\ E_{14}(-2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 2 & -3 & 7 \\ 0 & 1 & -4 & -3 \end{bmatrix}$$

$$\xrightarrow[E_{32}(-1),\ E_{42}(-1/2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -2 & -5 \end{bmatrix} \xrightarrow[E_{43}(2)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(Gaussian Elimination stops here. Continue with GJE.)

$$\xrightarrow[E_{34}(-3),\ E_{24}(-4)]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow[E_{13}(-2), E_{23}(4)]{} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow[E_2(1/2)]{} I_4$$

# Gauss-Jordan method for computing inverse

Question: Suppose $\mathrm{rref}(A) = I_n$. Is $A$ invertible?.

# Gauss-Jordan method for computing inverse

Question: Suppose $\mathrm{rref}(A) = I_n$. Is $A$ invertible?.

**Fact:** Let $A$ be an $n \times n$ matrix. If $E_p \cdots E_2 E_1 A = I_n$ then $A^{-1} = E_p \cdots E_2 E_1$, where $E_1, \ldots, E_p$ are elementary matrices.

Observation: Elementary row operations that transform $A$ to $I_n$ transform $I_n$ to $A^{-1}$.

# Gauss-Jordan method for computing inverse

Question: Suppose $\mathrm{rref}(A) = I_n$. Is $A$ invertible?.

**Fact:** Let $A$ be an $n \times n$ matrix. If $E_p \cdots E_2 E_1 A = I_n$ then $A^{-1} = E_p \cdots E_2 E_1$, where $E_1, \ldots, E_p$ are elementary matrices.

Observation: Elementary row operations that transform $A$ to $I_n$ transform $I_n$ to $A^{-1}$.

Note: If $A$ is invertible, then the rref of $\begin{bmatrix} A \mid I_n \end{bmatrix}$ is given by

$$\mathrm{rref}(\begin{bmatrix} A \mid I_n \end{bmatrix}) = \begin{bmatrix} E_k \cdots E_2 E_1 A \mid E_k \cdots E_2 E_1 \end{bmatrix} =$$

# Gauss-Jordan method for computing inverse

Question: Suppose $\mathrm{rref}(A) = I_n$. Is $A$ invertible?.

**Fact:** Let $A$ be an $n \times n$ matrix. If $E_p \cdots E_2 E_1 A = I_n$ then $A^{-1} = E_p \cdots E_2 E_1$, where $E_1, \ldots, E_p$ are elementary matrices.

Observation: Elementary row operations that transform $A$ to $I_n$ transform $I_n$ to $A^{-1}$.

Note: If $A$ is invertible, then the rref of $\left[A \mid I_n\right]$ is given by

$$\mathrm{rref}(\left[A \mid I_n\right]) = \left[E_k \cdots E_2 E_1 A \mid E_k \cdots E_2 E_1\right] = \left[I_n \mid A^{-1}\right].$$

# Gauss-Jordan method for computing inverse

Question: Suppose $\mathrm{rref}(A) = I_n$. Is $A$ invertible?.

**Fact:** Let $A$ be an $n \times n$ matrix. If $E_p \cdots E_2 E_1 A = I_n$ then $A^{-1} = E_p \cdots E_2 E_1$, where $E_1, \ldots, E_p$ are elementary matrices.

Observation: Elementary row operations that transform $A$ to $I_n$ transform $I_n$ to $A^{-1}$.

Note: If $A$ is invertible, then the rref of $\left[A \mid I_n\right]$ is given by

$$\mathrm{rref}(\left[A \mid I_n\right]) = \left[E_k \cdots E_2 E_1 A \mid E_k \cdots E_2 E_1\right] = \left[I_n \mid A^{-1}\right].$$

So to find $A^{-1}$, use GJE to $\left[A \mid I_n\right]$.

Gauss-Jordan method:

$$[A \mid I_n] \longrightarrow [I_n \mid X] \Rightarrow A \text{ is invertible and } A^{-1} = X.$$

# Example: Gauss-Jordan method

Let $A := \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 9 \end{bmatrix}$. Then $[A \mid I] \to [I \mid A^{-1}]$ gives

# Example: Gauss-Jordan method

Let $A := \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 9 \end{bmatrix}$. Then $[A \mid I] \to [I \mid A^{-1}]$ gives

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 7 & 0 & 1 & 0 \\ 3 & 7 & 9 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -3 & 0 & 1 \end{array}\right] \to$$

# Example: Gauss-Jordan method

Let $A := \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 9 \end{bmatrix}$. Then $[A \mid I] \to [I \mid A^{-1}]$ gives

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 7 & 0 & 1 & 0 \\ 3 & 7 & 9 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -3 & 0 & 1 \end{array}\right] \to$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 5 & -2 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & -3 & 1 \\ 0 & 1 & 0 & -3 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array}\right]$$

# Example: Gauss-Jordan method

Let $A := \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 9 \end{bmatrix}$. Then $[A \mid I] \to [I \mid A^{-1}]$ gives

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 7 & 0 & 1 & 0 \\ 3 & 7 & 9 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -3 & 0 & 1 \end{array}\right] \to$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 5 & -2 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & -3 & 1 \\ 0 & 1 & 0 & -3 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array}\right]$$

$$\Rightarrow A^{-1} = \begin{bmatrix} 4 & -3 & 1 \\ -3 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix}.$$

# Characterization of invertibility

**Theorem:** Let $A$ be an $n \times n$ matrix. Then the following statements are equivalent.

1. $A$ is invertible.
2. $Ax = b$ has a unique solution for every $b$ in $\mathbb{R}^n$.
3. $Ax = 0$ has only the trivial solution.
4. The reduced row echelon form of $A$ is $I_n$.
5. $A$ is a product of elementary matrices.
6. $\mathrm{rank}(A) = n$

# LU Factorization

An $n \times n$ matrix $A$ has an LU factorization if $A = LU$, where $U$ is upper triangular and $L$ is unit lower triangular (diagonals are 1).

**Fact:** If $A \longrightarrow \mathrm{ref}(A)$ without row interchange then $A$ has an $LU$ factorization.

# LU Factorization

An $n \times n$ matrix $A$ has an LU factorization if $A = LU$, where $U$ is upper triangular and $L$ is unit lower triangular (diagonals are 1).

**Fact:** If $A \longrightarrow \mathrm{ref}(A)$ without row interchange then $A$ has an $LU$ factorization.

- $E_p \dots E_2 E_1 A = \mathrm{ref}(A) \Rightarrow A = LU$.
- $L := E_1^{-1} E_2^{-1} \dots E_p^{-1}$ and $U := \mathrm{ref}(A)$.
- Each $E_j$ is unit lower triangular and Type-II $\Rightarrow$ $L$ is unit lower triangular.

# LU Factorization

An $n \times n$ matrix $A$ has an LU factorization if $A = LU$, where $U$ is upper triangular and $L$ is unit lower triangular (diagonals are 1).

**Fact:** If $A \longrightarrow \mathrm{ref}(A)$ without row interchange then $A$ has an $LU$ factorization.

- $E_p \dots E_2 E_1 A = \mathrm{ref}(A) \Rightarrow A = LU$.
- $L := E_1^{-1} E_2^{-1} \dots E_p^{-1}$ and $U := \mathrm{ref}(A)$.
- Each $E_j$ is unit lower triangular and Type-II $\Rightarrow L$ is unit lower triangular.

Solution of $Ax = b$ via LU factorization (if exists):

- Compute $A = LU$.
- Solve $Ly = b$ for $y$ - forward substitution.
- Solve $Ux = y$ for $x$ - back substitution.

# LU factorization (cont.)

Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Set $m_{21} := a_{21}/a_{11}$ and $m_{31} := a_{31}/a_{11}$

when $a_{11} \neq 0$ (pivot) and define

$$E_1 := \begin{bmatrix} 1 & 0 & 0 \\ -m_{21} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } E_2 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -m_{31} & 0 & 1 \end{bmatrix}.$$

# LU factorization (cont.)

Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Set $m_{21} := a_{21}/a_{11}$ and $m_{31} := a_{31}/a_{11}$

when $a_{11} \neq 0$ (pivot) and define

$$E_1 := \begin{bmatrix} 1 & 0 & 0 \\ -m_{21} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } E_2 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -m_{31} & 0 & 1 \end{bmatrix}.$$

Then

$$E_2 E_1 A = E_2 \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} \end{bmatrix}.$$

Set $m_{32} := a_{32}^{(1)}/a_{22}^{(1)}$ if $a_{22}^{(1)} \neq 0$ (pivot) and define

# LU factorization (cont.)

$E_3 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -m_{32} & 1 \end{bmatrix}$. Then we have

$$E_3 E_2 E_1 A = E_3 \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & 0 & a_{33}^{(2)} \end{bmatrix} = U.$$

.

# LU factorization (cont.)

$E_3 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -m_{32} & 1 \end{bmatrix}$. Then we have

$$E_3 E_2 E_1 A = E_3 \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & 0 & a_{33}^{(2)} \end{bmatrix} = U.$$

.

$$\text{Hence } A = LU, \text{ where } L = E_1^{-1} E_2^{-1} E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{bmatrix}.$$

# Examples: LU factorization

Let $A := \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$. Then $A = LU$, where

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

# Examples: LU factorization

Let $A := \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$. Then $A = LU$, where

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let $A := \begin{bmatrix} 2 & 4 & -1 \\ -4 & -5 & 3 \\ 2 & -5 & -4 \end{bmatrix}$. Then $A = LU$, where

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -3 & 1 \end{bmatrix} \text{ and } U = \begin{bmatrix} 2 & 4 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

# Three theorems with similar proof structure

Theorem: rref of a matrix is unique. Equivalently, if $R_1$ and $R_2$ are in rref and are row equivalent, then $R_1 = R_2$.

# Three theorems with similar proof structure

Theorem: rref of a matrix is unique. Equivalently, if $R_1$ and $R_2$ are in rref and are row equivalent, then $R_1 = R_2$.

Theorem: Let $A$ and $B$ be $m \times n$ matrices. Then the systems $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ are equivalent (i.e. have same solutions) if and only if $A$ and $B$ are row equivalent.

# Three theorems with similar proof structure

Theorem: rref of a matrix is unique. Equivalently, if $R_1$ and $R_2$ are in rref and are row equivalent, then $R_1 = R_2$.

Theorem: Let $A$ and $B$ be $m \times n$ matrices. Then the systems $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ are equivalent (i.e. have same solutions) if and only if $A$ and $B$ are row equivalent.

Theorem: Let $A$ and $B$ be $m \times n$ matrices. Suppose that the systems $Ax = b$ and $Bx = c$ are consistent. Then the two systems are equivalent if and only if the matrices $[\ A \mid\ b\ ]$ and $[\ B \mid\ c\ ]$ are row equivalent.

# Uniqueness of rref

Theorem: Let $R_1$ and $R_2$ be rref of $A$. Then $R_1 = R_2$.

## Uniqueness of rref

Theorem: Let $R_1$ and $R_2$ be rref of $A$. Then $R_1 = R_2$.

Proof: Note that $R_1$ and $R_2$ are row equivalent. Hence

$$R_1\mathbf{x} = \mathbf{0} \iff R_2\mathbf{x} = \mathbf{0}. \tag{1}$$

# Uniqueness of rref

Theorem: Let $R_1$ and $R_2$ be rref of $A$. Then $R_1 = R_2$.

Proof: Note that $R_1$ and $R_2$ are row equivalent. Hence

$$R_1\mathbf{x} = \mathbf{0} \iff R_2\mathbf{x} = \mathbf{0}. \tag{1}$$

**Step-1:** Claim $R_1\mathbf{e}_1 = R_2\mathbf{e}_1$.

## Uniqueness of rref

Theorem: Let $R_1$ and $R_2$ be rref of $A$. Then $R_1 = R_2$.

Proof: Note that $R_1$ and $R_2$ are row equivalent. Hence

$$R_1\mathbf{x} = \mathbf{0} \iff R_2\mathbf{x} = \mathbf{0}. \tag{1}$$

**Step-1:** Claim $R_1\mathbf{e}_1 = R_2\mathbf{e}_1$.

By (1), we have $R_1\mathbf{e}_1 = \mathbf{0} \iff R_2\mathbf{e}_1 = \mathbf{0}$.

# Uniqueness of rref

Theorem: Let $R_1$ and $R_2$ be rref of $A$. Then $R_1 = R_2$.

Proof: Note that $R_1$ and $R_2$ are row equivalent. Hence

$$R_1\mathbf{x} = \mathbf{0} \iff R_2\mathbf{x} = \mathbf{0}. \tag{1}$$

**Step-1:** Claim $R_1\mathbf{e}_1 = R_2\mathbf{e}_1$.

By (1), we have $R_1\mathbf{e}_1 = \mathbf{0} \iff R_2\mathbf{e}_1 = \mathbf{0}$.

If $R_1\mathbf{e}_1 \neq \mathbf{0}$ then $R_1\mathbf{e}_1$ and $R_2\mathbf{e}_1$ are pivot columns. Hence

$$R_1\mathbf{e}_1 = \mathbf{e}_1 = R_2\mathbf{e}_1.$$

## Uniqueness of rref

Theorem: Let $R_1$ and $R_2$ be rref of $A$. Then $R_1 = R_2$.

Proof: Note that $R_1$ and $R_2$ are row equivalent. Hence

$$R_1\mathbf{x} = \mathbf{0} \iff R_2\mathbf{x} = \mathbf{0}. \tag{1}$$

**Step-1:** Claim $R_1\mathbf{e}_1 = R_2\mathbf{e}_1$.

By (1), we have $R_1\mathbf{e}_1 = \mathbf{0} \iff R_2\mathbf{e}_1 = \mathbf{0}$.

If $R_1\mathbf{e}_1 \neq \mathbf{0}$ then $R_1\mathbf{e}_1$ and $R_2\mathbf{e}_1$ are pivot columns. Hence

$$R_1\mathbf{e}_1 = \mathbf{e}_1 = R_2\mathbf{e}_1.$$

**Step-2:** Suppose $R_1\mathbf{e}_j = R_2\mathbf{e}_j$ for $j = 1 : k$ and $R_1\mathbf{e}_{k+1} \neq R_2\mathbf{e}_{k+1}$.

# Uniqueness of rref

Theorem: Let $R_1$ and $R_2$ be rref of $A$. Then $R_1 = R_2$.

Proof: Note that $R_1$ and $R_2$ are row equivalent. Hence

$$R_1\mathbf{x} = \mathbf{0} \iff R_2\mathbf{x} = \mathbf{0}. \tag{1}$$

**Step-1:** Claim $R_1\mathbf{e}_1 = R_2\mathbf{e}_1$.

By (1), we have $R_1\mathbf{e}_1 = \mathbf{0} \iff R_2\mathbf{e}_1 = \mathbf{0}$.

If $R_1\mathbf{e}_1 \neq \mathbf{0}$ then $R_1\mathbf{e}_1$ and $R_2\mathbf{e}_1$ are pivot columns. Hence

$$R_1\mathbf{e}_1 = \mathbf{e}_1 = R_2\mathbf{e}_1.$$

**Step-2:** Suppose $R_1\mathbf{e}_j = R_2\mathbf{e}_j$ for $j = 1:k$ and $R_1\mathbf{e}_{k+1} \neq R_2\mathbf{e}_{k+1}$. If $R_1\mathbf{e}_{k+1}$ is a non-pivot column then there exist $\alpha_1, \ldots, \alpha_k$ such that

## Uniqueness of rref

Theorem: Let $R_1$ and $R_2$ be rref of $A$. Then $R_1 = R_2$.

Proof: Note that $R_1$ and $R_2$ are row equivalent. Hence

$$R_1\mathbf{x} = \mathbf{0} \iff R_2\mathbf{x} = \mathbf{0}. \tag{1}$$

**Step-1:** Claim $R_1\mathbf{e}_1 = R_2\mathbf{e}_1$.

By (1), we have $R_1\mathbf{e}_1 = \mathbf{0} \iff R_2\mathbf{e}_1 = \mathbf{0}$.

If $R_1\mathbf{e}_1 \neq \mathbf{0}$ then $R_1\mathbf{e}_1$ and $R_2\mathbf{e}_1$ are pivot columns. Hence

$$R_1\mathbf{e}_1 = \mathbf{e}_1 = R_2\mathbf{e}_1.$$

**Step-2:** Suppose $R_1\mathbf{e}_j = R_2\mathbf{e}_j$ for $j = 1:k$ and $R_1\mathbf{e}_{k+1} \neq R_2\mathbf{e}_{k+1}$.
If $R_1\mathbf{e}_{k+1}$ is a non-pivot column then there exist $\alpha_1, \ldots, \alpha_k$ such that

$$R_1\mathbf{e}_{k+1} = \alpha_1 R_1\mathbf{e}_1 + \cdots + \alpha_k R_1\mathbf{e}_k \Longrightarrow R_1\mathbf{x} = \mathbf{0}, \tag{2}$$

where $\mathbf{x} := [\alpha_1, \ldots, \alpha_k, -1, 0, \ldots, 0]^\top$.

# Uniqueness of rref

Since $R_1\mathbf{e}_j = R_2\mathbf{e}_j, j = 1, 2, \ldots, k$, by (1) and (2), we have

## Uniqueness of rref

Since $R_1\mathbf{e}_j = R_2\mathbf{e}_j, j = 1, 2, \ldots, k$, by (1) and (2), we have

$$R_2\mathbf{x} = \mathbf{0} \Rightarrow R_2\mathbf{e}_{k+1} = \alpha_1 R_2\mathbf{e}_1 + \cdots + \alpha_k R_2\mathbf{e}_k = R_1\mathbf{e}_{k+1}$$

which is a contradiction.

# Uniqueness of rref

Since $R_1\mathbf{e}_j = R_2\mathbf{e}_j, j = 1, 2, \ldots, k$, by (1) and (2), we have

$$R_2\mathbf{x} = \mathbf{0} \Rightarrow R_2\mathbf{e}_{k+1} = \alpha_1 R_2\mathbf{e}_1 + \cdots + \alpha_k R_2\mathbf{e}_k = R_1\mathbf{e}_{k+1}$$

which is a contradiction.

On the other hand, if $R_1\mathbf{e}_{k+1}$ is a pivot column then $R_2\mathbf{e}_{k+1}$ is also a pivot column.

# Uniqueness of rref

Since $R_1\mathbf{e}_j = R_2\mathbf{e}_j, j = 1, 2, \ldots, k$, by (1) and (2), we have

$$R_2\mathbf{x} = \mathbf{0} \Rightarrow R_2\mathbf{e}_{k+1} = \alpha_1 R_2\mathbf{e}_1 + \cdots + \alpha_k R_2\mathbf{e}_k = R_1\mathbf{e}_{k+1}$$

which is a contradiction.

On the other hand, if $R_1\mathbf{e}_{k+1}$ is a pivot column then $R_2\mathbf{e}_{k+1}$ is also a pivot column.

If there are $p$ pivot columns in the first $k$ columns of $R_1$ then $R_1\mathbf{e}_{k+1} = \mathbf{e}_{p+1}$ and $R_2\mathbf{e}_{k+1} = \mathbf{e}_{p+1}$.

# Uniqueness of rref

Since $R_1\mathbf{e}_j = R_2\mathbf{e}_j, j = 1, 2, \ldots, k$, by (1) and (2), we have

$$R_2\mathbf{x} = \mathbf{0} \Rightarrow R_2\mathbf{e}_{k+1} = \alpha_1 R_2\mathbf{e}_1 + \cdots + \alpha_k R_2\mathbf{e}_k = R_1\mathbf{e}_{k+1}$$

which is a contradiction.

On the other hand, if $R_1\mathbf{e}_{k+1}$ is a pivot column then $R_2\mathbf{e}_{k+1}$ is also a pivot column.

If there are $p$ pivot columns in the first $k$ columns of $R_1$ then $R_1\mathbf{e}_{k+1} = \mathbf{e}_{p+1}$ and $R_2\mathbf{e}_{k+1} = \mathbf{e}_{p+1}$. Hence $R_1\mathbf{e}_{k+1} = R_2\mathbf{e}_{k+1}$ which is a contradiction.

## Uniqueness of rref

Since $R_1\mathbf{e}_j = R_2\mathbf{e}_j, j = 1, 2, \ldots, k$, by (1) and (2), we have

$$R_2\mathbf{x} = \mathbf{0} \Rightarrow R_2\mathbf{e}_{k+1} = \alpha_1 R_2\mathbf{e}_1 + \cdots + \alpha_k R_2\mathbf{e}_k = R_1\mathbf{e}_{k+1}$$

which is a contradiction.

On the other hand, if $R_1\mathbf{e}_{k+1}$ is a pivot column then $R_2\mathbf{e}_{k+1}$ is also a pivot column.

If there are $p$ pivot columns in the first $k$ columns of $R_1$ then $R_1\mathbf{e}_{k+1} = \mathbf{e}_{p+1}$ and $R_2\mathbf{e}_{k+1} = \mathbf{e}_{p+1}$. Hence $R_1\mathbf{e}_{k+1} = R_2\mathbf{e}_{k+1}$ which is a contradiction.

Thus, the assumption $R_1\mathbf{e}_{k+1} \neq R_2\mathbf{e}_{k+1}$ must be wrong and we must have $R_1 = R_2$. ■

*** End ***