Lecture-1: Syllabus for MA 106 : Linear Algebra

1. We shall study the Gauss elimination method and determinants to solve systems of linear equations.
2. We shall also discuss basics of vector spaces and inner product spaces.
3. Matrices are simplified using eigenvalues and eigenvectors. We shall learn algorithms to diagonalize matrices of special types such as symmetric, Hermitian, orthogonal, unitary and more generally normal matrices.
4. We shall learn to use eigenvalues to solve differential equations.
5. We will use diagonalization of matrices for identification of curves and surfaces represented by quadratic equations.

Tutorial Dates	Topics	No. of hrs of lectures
22.01.2020	Matrices and Linear Equations	4
29.01.2020	Determinant and Vector Spaces	5
05.02.2020	Linear Transformations	3
12.02.2020	Inner Product Spaces	3
19.02.2020	Diagonalization of Matrices	5

References and course policies

Textbooks and References

(1) Serge Lang, Introduction to Linear Algebra, 2nd Ed. Springer, India.
(2) Gilbert Strang, Linear Algebra and its applications, 4th Ed. Cengage.
(3) M.K. Srinivasan and J.K. Verma, Notes on Linear Algebra, 2014.
(4) B.V. Limaye, Linear Algebra, Slides for a course at IIT Dharwad, 2019.

Attendance Policy: If attendance is less than 80 % , a DX grade will be awarded.

Two extra classes				
Division	**Date**	**Slot**	**Time**	**Venue**
D1, D2	17, 24 Jan	7B	8.30	LA 001, LA 002
D3, D4	22, 29 Jan	7A	8.30	LA 001, LA 002

Evaluation scheme		
Mode	**Time**	**Marks**
5 quizzes	tutorial hours	9 marks, 3 best scores
Test	5 Feb, 8.30-9.25	11
Final exam	TBA	30
Total		50

Notation

Notation:

1. $\mathbb{N} := \{1, 2, 3, \ldots\}$
2. $\mathbb{Z} := \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
3. $\mathbb{R} :=$ the set of all real numbers
4. $\mathbb{C} :=$ the set of all complex numbers
5. For $n \in \mathbb{N}$, let us consider the **Euclidean space** $\mathbb{R}^n := \{(x_1, \ldots, x_n) : x_j \in \mathbb{R} \text{ for } j = 1, \ldots, n\}$.
6. We let $\mathbf{0} := (0, \ldots, 0)$. Also, for $\boldsymbol{x} := (x_1, \ldots, x_n)$ and $\boldsymbol{y} := (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, and for $\alpha \in \mathbb{R}$, we define

$$\begin{aligned} \text{(sum)} \quad \boldsymbol{x} + \boldsymbol{y} &:= (x_1 + y_1, \ldots, x_n + y_n) \in \mathbb{R}^n, \\ \text{(scalar multiple)} \quad \alpha \boldsymbol{x} &:= (\alpha x_1, \ldots, \alpha x_n) \in \mathbb{R}^n, \\ \text{(scalar product)} \quad \boldsymbol{x} \cdot \boldsymbol{y} &:= x_1 y_1 + \cdots + x_n y_n \in \mathbb{R}. \end{aligned}$$

Matrices

1. Let $m, n \in \mathbb{N}$. An $m \times n$ **matrix A** with real entries is a rectangular array of real numbers arranged in m rows and n columns, written as follows:

$$\mathbf{A} := \begin{bmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \cdots & \boxed{a_{jk}} & \cdots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mk} & \cdots & a_{mn} \end{bmatrix} = [a_{jk}],$$

   where $a_{jk} \in \mathbb{R}$ is called the (j, k)th **entry** of **A** for $j = 1, \ldots, m$ and $k = 1, \ldots, n$.

2. Let $\boxed{\mathbb{R}^{m \times n}}$ denote the set of all $m \times n$ matrices with real entries.

3. If $\mathbf{A} := [a_{jk}]$ and $\mathbf{B} := [b_{jk}]$ are in $\mathbb{R}^{m \times n}$, then we say $\mathbf{A} = \mathbf{B} \iff a_{jk} = b_{jk}$ for all $j = 1, \ldots, m$ and $k = 1, \ldots, n$.

4. Let $0 \leq r < m$, $0 \leq s < n$. By deleting r rows and s columns from **A**, we obtain an $(m - r) \times (n - s)$ **submatrix** of **A**.

Matrices

1. An $n \times n$ matrix, that is, an element of $\mathbb{R}^{n \times n}$, is called a **square matrix** of size n.
2. A square matrix $\mathbf{A} = [a_{jk}]$ is called **symmetric** if $a_{jk} = a_{kj}$ for all j, k.
3. A square matrix $\mathbf{A} = [a_{jk}]$ is called **skew-symmetric** if $a_{jk} = -a_{kj}$ for all j, k.
4. A square matrix $\mathbf{A} = [a_{jk}]$ is called a **diagonal matrix** if $a_{jk} = 0$ for all $j \neq k$.
5. A diagonal matrix $\mathbf{A} = [a_{jk}]$ is called a **scalar matrix** if all diagonal entries of $\mathbf{A}$ are equal.
6. Two important scalar matrices are the **identity matrix I** in which all diagonal elements are equal to 1, and the **zero matrix O** in which all diagonal elements are equal to 0.
7. A square matrix $\mathbf{A} = [a_{jk}]$ is called **upper triangular** if $a_{jk} = 0$ for all $j > k$, and **lower triangular** if $a_{jk} = 0$ for all $j < k$.
8. Note: A matrix $\mathbf{A}$ is upper triangular as well as lower triangular if and only if $\mathbf{A}$ is a diagonal matrix.

Matrices: Examples

1. The matrix $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ is symmetric, while the matrix $\begin{bmatrix} 0 & 2 & 3 \\ -2 & 0 & 5 \\ -3 & -5 & 0 \end{bmatrix}$ is skew-symmetric.
2. Note: Every diagonal entry of a skew-symmetric matrix is 0 since $a_{jj} = -a_{jj} \implies a_{jj} = 0$ for $j = 1, \ldots, n$.
3. The matrix $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ is diagonal, while $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is a scalar matrix.
4. The matrix $\begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 4 \end{bmatrix}$ is upper triangular, while the matrix $\begin{bmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ -1 & 1 & 4 \end{bmatrix}$ is lower triangular.

Matrices: Row and Column vectors

1. A **row vector a** of length n is a matrix with only one row consisting of n real numbers; it is written as

$$\mathbf{a} = \begin{bmatrix} a_1 & \cdots & a_k & \cdots & a_n \end{bmatrix},$$

where $a_k \in \mathbb{R}$ for $k = 1, \ldots, n$. Here $\mathbf{a} \in \mathbb{R}^{1 \times n}$.

2. A **column vector b** of length n is a matrix with only one column consisting of n real numbers; it is written as

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_k \\ \vdots \\ b_n \end{bmatrix},$$

where $b_k \in \mathbb{R}$ for $k = 1, \ldots, n$. Here $\mathbf{b} \in \mathbb{R}^{n \times 1}$.

Operations on Matrices

1. Let $m, n \in \mathbb{N}$, and let $\mathbf{A} := [a_{jk}]$ and $\mathbf{B} := [b_{jk}]$ be $m \times n$ matrices.
2. Then the $m \times n$ matrix $\mathbf{A} + \mathbf{B} := [a_{jk} + b_{jk}]$ is called the **sum** of $\mathbf{A}$ and $\mathbf{B}$. If $\alpha \in \mathbb{R}$, then the $m \times n$ matrix $\alpha\mathbf{A} := [\alpha a_{jk}]$ is called the **scalar multiple** of $\mathbf{A}$ by α.
3. These operations follow the usual rules:
   i. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
   ii. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$, which we write as $\mathbf{A} + \mathbf{B} + \mathbf{C}$
   iii. $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$
   iv. $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$
   v. $\alpha(\beta\mathbf{A}) = (\alpha\beta)\mathbf{A}$, which we write as $\alpha\beta\mathbf{A}$
4. We write $(-1)\mathbf{A}$ as $-\mathbf{A}$, and $\mathbf{A} + (-\mathbf{B})$ as $\mathbf{A} - \mathbf{B}$.

Operations on Matrices

1. The **transpose** of an $m \times n$ matrix $\mathbf{A} := [a_{jk}]$ is the $n \times m$ matrix $\mathbf{A}^\mathsf{T} := [b_{jk}]$ (in which the rows and the columns of $\mathbf{A}$ are interchanged) where $b_{jk} = a_{kj}$.
2. Clearly, $(\mathbf{A}^\mathsf{T})^\mathsf{T} = \mathbf{A}$, $(\mathbf{A} + \mathbf{B})^\mathsf{T} = \mathbf{A}^\mathsf{T} + \mathbf{B}^\mathsf{T}$ and $(\alpha\mathbf{A})^\mathsf{T} = \alpha\mathbf{A}^\mathsf{T}$.
3. Note:
   i. A square matrix $\mathbf{A}$ is symmetric $\iff \mathbf{A}^\mathsf{T} = \mathbf{A}$.
   ii. A square matrix $\mathbf{A}$ is skew-symmetric $\iff \mathbf{A}^\mathsf{T} = -\mathbf{A}$.
4. In particular, the preceding operations can be performed on row vectors, and also on column vectors since they are particular types of matrices.
5. Also, note that the transpose of a row vector is a column vector, and vice versa.

Operations on Matrices

1. We shall often write a column vector $\mathbf{b} := \begin{bmatrix} b_1 \\ \vdots \\ b_k \\ \vdots \\ b_n \end{bmatrix} \in \mathbb{R}^{n\times 1}$ as $\begin{bmatrix} b_1 & \cdots & b_k & \cdots & b_n \end{bmatrix}^{\mathsf{T}}$ in order to save space.
2. Let $m, n \in \mathbb{N}$. Let $\alpha_1, \dots, \alpha_m \in \mathbb{R}$.
3. If $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^{1\times n}$, then
$$\alpha_1 \mathbf{a}_1 + \cdots + \alpha_m \mathbf{a}_m \in \mathbb{R}^{1\times n}$$
is called a **(finite) linear combination** of $\mathbf{a}_1, \dots, \mathbf{a}_m$.
4. Similarly, if $\mathbf{b}_1, \dots, \mathbf{b}_m \in \mathbb{R}^{n\times 1}$, then
$$\alpha_1 \mathbf{b}_1 + \cdots + \alpha_m \mathbf{b}_m \in \mathbb{R}^{n\times 1}$$
is a **(finite) linear combination** of $\mathbf{b}_1, \dots, \mathbf{b}_m$.

Operations on Matrices

1. Now, for $k = 1, \ldots, n$, consider the column vector $\mathbf{e}_k := \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{n \times 1}$, where the kth entry is 1 and all other entries are 0.

2. If $\mathbf{b} = \begin{bmatrix} b_1 & \cdots & b_k & \cdots & b_n \end{bmatrix}^\mathsf{T}$ is any column vector of length n, then it follows that $\mathbf{b} = b_1\mathbf{e}_1 + \cdots + b_k\mathbf{e}_k + \cdots + b_n\mathbf{e}_n$, which is a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

3. The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are known as the **basic column vectors** in $\mathbb{R}^{n \times 1}$.

4. Let $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m \times n}$. Then

   i. $\mathbf{a}_j := \begin{bmatrix} a_{j1} & \cdots & a_{jn} \end{bmatrix} \in \mathbb{R}^{1 \times n}$ is called the jth **row vector of A** for $j = 1, \ldots, m$, and we write $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}$ and

   ii. $\mathbf{c}_k := \begin{bmatrix} a_{1k} & \cdots & a_{mk} \end{bmatrix}^\mathsf{T}$ is called the kth **column vector of A** for $k = 1, \ldots, n$, and we write $\mathbf{A} = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{bmatrix}$.

Operations on Matrices: Examples

Let $\mathbf{A} := \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \end{bmatrix}$ and $\mathbf{B} := \begin{bmatrix} 1 & 0 & 2 \\ -1 & 4 & 1 \end{bmatrix}$.

1. Then $\mathbf{A} + \mathbf{B} := \begin{bmatrix} 3 & 1 & 1 \\ -1 & 7 & 2 \end{bmatrix}$ and $5\mathbf{A} = \begin{bmatrix} 10 & 5 & -5 \\ 0 & 15 & 5 \end{bmatrix}$.
2. The row vectors of $\mathbf{A}$ are $\begin{bmatrix} 2 & 1 & -1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 3 & 1 \end{bmatrix}$.
3. The column vectors of $\mathbf{A}$ are $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.
4. Also, $\mathbf{A}^{\mathsf{T}} = \begin{bmatrix} 2 & 0 \\ 1 & 3 \\ -1 & 1 \end{bmatrix}$.

Matrix Multiplication

1. Before we discuss when and how a product $\mathbf{AB}$ of $\mathbf{A}$ and $\mathbf{B}$ may be defined, we define the product of a row vector $\mathbf{a}$ and a column vector $\mathbf{b}$ of the same length.

2. Since every matrix can be written in terms of row vectors as well as in terms of column vectors, we shall then consider the product $\mathbf{AB}$.

3. Let $n \in \mathbb{N}$, $\mathbf{a} := \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \in \mathbb{R}^{1\times n}$ and $\mathbf{b} := \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{n\times 1}$.

4. Define the **product** of a row vector $\mathbf{a}$ and a column vector $\mathbf{b}$ as follows:

$$\mathbf{ab} = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} := a_1 b_1 + \cdots + a_n b_n \in \mathbb{R}.$$

5. Let $m \in \mathbb{N}$ and $\mathbf{A} \in \mathbb{R}^{m\times n}$. Then $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}$, where $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^{1\times n}$.

Matrix Multiplication

1. Recalling that $\mathbf{b} \in \mathbb{R}^{n\times 1}$, we define

$$\mathbf{Ab} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} \mathbf{b} := \begin{bmatrix} \mathbf{a}_1\mathbf{b} \\ \vdots \\ \mathbf{a}_m\mathbf{b} \end{bmatrix} \in \mathbb{R}^{m\times 1}.$$

2. Finally, let $p \in \mathbb{N}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$. Then $\mathbf{B} = [\mathbf{b}_1 \cdots \mathbf{b}_p]$, where $\mathbf{b}_1, \ldots, \mathbf{b}_p \in \mathbb{R}^{n\times 1}$.

3. Noting that $\mathbf{Ab}_1, \ldots, \mathbf{Ab}_p$ belong to $\mathbb{R}^{m\times 1}$, we define

$$\mathbf{AB} = \mathbf{A}[\mathbf{b}_1 \cdots \mathbf{b}_p] := [\mathbf{Ab}_1 \cdots \mathbf{Ab}_p] \in \mathbb{R}^{m\times p}.$$

4. Thus

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} [\mathbf{b}_1 \cdots \mathbf{b}_p] = \begin{bmatrix} \mathbf{a}_1\mathbf{b}_1 & \cdots & \mathbf{a}_1\mathbf{b}_p \\ \vdots & & \vdots \\ \mathbf{a}_m\mathbf{b}_1 & \cdots & \mathbf{a}_m\mathbf{b}_p \end{bmatrix} \in \mathbb{R}^{m\times p}.$$

Matrix Multiplication

1. So if $m, n, p \in \mathbb{N}$, $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m \times n}$ and $\mathbf{B} := [b_{jk}] \in \mathbb{R}^{n \times p}$, then $\mathbf{AB} \in \mathbb{R}^{m \times p}$, and for $j = 1, \ldots, m$; $k = 1 \ldots, p$,

$$\mathbf{AB} = [c_{jk}], \quad \text{where } c_{jk} := \mathbf{a}_j \mathbf{b}_k = \sum_{\ell=1}^{n} a_{j\ell} b_{\ell k}.$$

2. Note that the (j,k)th entry of $\mathbf{AB}$ is a product of the jth row vector of $\mathbf{A}$ with the kth column vector of $\mathbf{B}$ as shown below:

$$\begin{bmatrix} & & & & \\ a_{j1} & \cdots & a_{j\ell} & \cdots & a_{jn} \\ & & & & \end{bmatrix} \begin{bmatrix} & b_{1k} & \\ & \vdots & \\ & b_{\ell k} & \\ & \vdots & \\ & b_{nk} & \end{bmatrix}.$$

3. Clearly, the product $\mathbf{AB}$ is defined only when the number of columns of $\mathbf{A}$ is equal to the number of rows of $\mathbf{B}$.

4. Note: $\mathbf{AI} = \mathbf{A}, \mathbf{IA} = \mathbf{A}, \mathbf{AO} = \mathbf{O}$ and $\mathbf{OA} = \mathbf{O}$ whenever these products are defined.

Matrix Multiplication: Examples

1. Let $\mathbf{A} := \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \end{bmatrix}_{2\times 3}$, $\mathbf{B} := \begin{bmatrix} 1 & 6 & 0 & 2 \\ 2 & -1 & 1 & -2 \\ 2 & 0 & -1 & 1 \end{bmatrix}_{3\times 4}$.

   Then $\mathbf{AB} = \begin{bmatrix} 2 & 11 & 2 & 1 \\ 8 & -3 & 2 & -5 \end{bmatrix}_{2\times 4}$.

2. Both products $\mathbf{AB}$ and $\mathbf{BA}$ are defined $\iff$ the number of columns of $\mathbf{A}$ is equal to the number of rows of $\mathbf{B}$ and the number columns of $\mathbf{B}$ is equal to the number of rows of $\mathbf{A}$, that is, when $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times m}$.

3. In general, $\mathbf{AB} \neq \mathbf{BA}$. For example, if $\mathbf{A} := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\mathbf{B} := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, then $\mathbf{AB} := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, while $\mathbf{BA} := \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

4. Note that $\mathbf{BA} = \mathbf{O}$, while $\mathbf{AB} = \mathbf{B} \neq \mathbf{O}$. Since $\mathbf{A} \neq \mathbf{I}$, we see that the so-called cancellation law does not hold.

Matrix Multiplication: Important Remark

1. Let $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m \times n}$, and let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the basic column vectors in $\mathbb{R}^{n \times 1}$. Then for $k = 1, \ldots, n$, $\mathbf{A}\,\mathbf{e}_k = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{jk} \\ \vdots \\ a_{mk} \end{bmatrix}$, which is the kth column of $\mathbf{A}$.

   This follows from our definition of matrix multiplication.

2. It follows that if $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$, then
   $\boxed{\mathbf{A} = \mathbf{B} \text{ if and only if } \mathbf{A}\mathbf{e}_k = \mathbf{B}\mathbf{e}_k \text{ for each } k = 1, \ldots, n.}$

Properties of Matrix Multiplication

1. Consider matrices **A**, **B**, **C** and $\alpha \in \mathbb{R}$.
2. Then it is easy to see that $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$, $\mathbf{C}(\mathbf{A}+\mathbf{B}) = \mathbf{CA}+\mathbf{CB}$ and $(\alpha\mathbf{A})\mathbf{B} = \alpha\mathbf{AB} = \mathbf{A}(\alpha\mathbf{B})$, if sums & products are well-defined.
3. Matrix multiplication also satisfies the **associative law**:

Proposition

Let $m, n, p, q \in \mathbb{N}$. If $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times p}$ and $\mathbf{C} \in \mathbb{R}^{p\times q}$, then $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ (which we shall write as $\mathbf{ABC}$).

Proof. Let $\mathbf{A} := [a_{jk}]$, $\mathbf{B} := [b_{jk}]$ and $\mathbf{C} := [c_{jk}]$. Also, let $(\mathbf{AB})\mathbf{C} := [\alpha_{jk}]$ and $\mathbf{A}(\mathbf{BC}) := [\beta_{jk}]$.

Then $\alpha_{jk} = \sum_{i=1}^{p} \left(\sum_{\ell=1}^{n} a_{j\ell} b_{\ell i} \right) c_{ik} = \sum_{\ell=1}^{n} a_{j\ell} \left(\sum_{i=1}^{p} b_{\ell i} c_{ik} \right) = \beta_{jk}$. $\square$

Properties of Matrix Multiplication continues...

Proposition

Let $m, n, p \in \mathbb{N}$. If $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{n \times p}$, then $(\mathbf{AB})^{\mathsf{T}} = \mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}$.

Proof. Let $\mathbf{A} := [a_{jk}]$, $\mathbf{B} := [b_{jk}]$ and $\mathbf{AB} := [c_{jk}]$.

Also, let $\mathbf{A}^{\mathsf{T}} := [a'_{jk}]$, $\mathbf{B}^{\mathsf{T}} := [b'_{jk}]$ and $(\mathbf{AB})^{\mathsf{T}} := [c'_{jk}]$. Then

$$c_{jk} = \sum_{\ell=1}^{n} a_{j\ell} b_{\ell k} \quad \text{and so} \quad c'_{jk} = c_{kj} = \sum_{\ell=1}^{n} a_{k\ell} b_{\ell j}$$

for $j = 1, \ldots, m$; $k = 1, \ldots, p$. Suppose $\mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}} := [d_{jk}]$. Then

$$d_{jk} = \sum_{\ell=1}^{n} b'_{j\ell} a'_{\ell k} = \sum_{\ell=1}^{n} b_{\ell j} a_{k\ell} = \sum_{\ell=1}^{n} a_{k\ell} b_{\ell j} = c'_{jk}$$

for $j = 1, \ldots, m$; $k = 1, \ldots, p$. Hence the result. $\square$

Linear System

1. Let $m, n \in \mathbb{N}$. A **linear system** of m equations in n unknowns $x_1, ..., x_n$ is given by

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \qquad (1)\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \qquad (2)\\
\vdots \qquad \vdots \qquad \vdots \qquad\quad & \vdots \quad \vdots\\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \qquad (m)
\end{aligned}$$

   where $a_{jk} \in \mathbb{R}$ for $j = 1, \ldots, m$; $k = 1, \ldots, n$ and also $b_j \in \mathbb{R}$ for $j = 1, \ldots, m$ are given.

2. Let $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m \times n}$, $\mathbf{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{n \times 1}$ and $\mathbf{b} := \begin{bmatrix} b_1 & \cdots & b_m \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{m \times 1}$.

3. Using matrix multiplication, we write the linear system as
$$\mathbf{Ax} = \mathbf{b}.$$

4. The $m \times n$ matrix $\mathbf{A}$ is known as the **coefficient matrix** of the linear system.

Linear System

1. A column vector $\mathbf{x}_0 \in \mathbb{R}^{n\times 1}$ is called a **solution** of the above linear system if it satisfies $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$.
2. Case (i) **Homogeneous Linear System**: $\mathbf{b} := \mathbf{0}$, that is, $b_1 = \cdots = b_m = 0$.
3. A homogeneous linear system always has a solution, namely the zero solution $\mathbf{0} := \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}^{\mathsf{T}}$ since $\mathbf{A}\mathbf{0} = \mathbf{0}$.
4. Also, if $r \in \mathbb{N}$ and $\mathbf{x}_1, \ldots, \mathbf{x}_r$ are solutions of such a system, then so is their linear combination $\alpha_1\mathbf{x}_1 + \cdots + \alpha_r\mathbf{x}_r$, since $\mathbf{A}(\alpha_1\mathbf{x}_1 + \cdots + \alpha_r\mathbf{x}_r) = \alpha_1\mathbf{A}\mathbf{x}_1 + \cdots + \alpha_r\mathbf{A}\mathbf{x}_r = \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}$.
5. Case (ii) **General Linear System**: $\mathbf{b} \in \mathbb{R}^{m\times 1}$ is arbitrary.
6. A nonhomogeneous linear system, that is, where $\mathbf{b} \neq \mathbf{0}$, may not have a solution, may have only one solution or may have (infinitely) many solutions.
7. Examples:
   i. The linear system $x_1 + x_2 = 1$, $2x_1 + 2x_2 = 1$ does not have a solution.
   ii. The linear system $x_1 + x_2 = 1$, $x_1 - x_2 = 0$ has a unique solution, namely $x_1 = 1/2 = x_2$.
   iii. The linear system $x_1 + x_2 = 1$, $2x_1 + 2x_2 = 2$ has (infinitely) many solutions, namely $x_1 = \alpha$, $x_2 = 1 - \alpha$, $\alpha \in \mathbb{R}$.

Linear System: An Important Note

1. Let S denote the set of all solutions of a homogeneous linear system $\mathbf{Ax} = \mathbf{0}$.

   If $\mathbf{x}_0$ is a particular solution of the general system $\mathbf{Ax} = \mathbf{b}$, then the set of all solutions of the general system $\mathbf{Ax} = \mathbf{b}$ is given by $\{\mathbf{x}_0 + \mathbf{s} : \mathbf{s} \in S\}$ since

   $$\mathbf{s} \in S \implies \mathbf{A}(\mathbf{x}_0 + \mathbf{s}) = \mathbf{Ax}_0 + \mathbf{As} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

   Also, if $\mathbf{y}_0$ is another solution of the general system $\mathbf{Ax} = \mathbf{b}$, i.e., if $\mathbf{Ay}_0 = \mathbf{b}$, then $\mathbf{A}(\mathbf{y}_0 - \mathbf{x}_0) = \mathbf{Ay}_0 - \mathbf{Ax}_0 = \mathbf{b} - \mathbf{b} = \mathbf{0}$ and hence $(\mathbf{y}_0 - \mathbf{x}_0) \in S$ and we can write $\mathbf{y}_0 = \mathbf{x}_0 + \mathbf{s}$ where $\mathbf{s} = (\mathbf{y}_0 - \mathbf{x}_0) \in S$.

2. We shall, therefore, address the problem of finding all solutions of a homogeneous linear system $\mathbf{Ax} = \mathbf{0}$, and one particular solution of the corresponding general system $\mathbf{Ax} = \mathbf{b}$.

Linear System: A Special Case

1. Suppose the coefficient matrix $\mathbf{A}$ is upper triangular and its diagonal elements are nonzero.

2. Then the linear system is

$$
\begin{array}{rcll}
a_{11}x_1 + a_{12}x_2 + \cdots + \cdots + \cdots + \cdots + a_{1n}x_n & = & b_1 & \qquad (1) \\
a_{22}x_2 + a_{23}x_3 + \cdots + \cdots + \cdots + a_{2n}x_n & = & b_2 & \qquad (2) \\
\ddots \qquad\qquad\qquad\qquad\qquad \vdots & \vdots & \vdots & \\
a_{(n-1)(n-1)}x_{n-1} + a_{(n-1)n}x_n & = & b_{n-1} & \qquad (n-1) \\
a_{nn}x_n & = & b_n & \qquad (n)
\end{array}
$$

which can be solved by back substitution as follows.

Linear System: A Special Case

$$
\begin{aligned}
x_n &= b_n/a_{nn} \\
x_{n-1} &= (b_{n-1} - a_{(n-1)n}x_n)/a_{(n-1)(n-1)}, \text{ where } x_n = b_n/a_{nn} \\
\vdots &\ \vdots\ \vdots \\
x_2 &= (b_2 - a_{2n}x_n - \cdots - a_{23}x_3)/a_{22}, \text{where } x_n = \cdots, x_3 = \cdots \\
x_1 &= (b_1 - a_{1n}x_n - \cdots - \cdots - a_{12}x_2)/a_{11}, \text{where } x_n = \cdots, x_2 = \cdots
\end{aligned}
$$

1. Here the homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution and the general system $\mathbf{Ax} = \mathbf{b}$ has a unique solution.
2. Taking a cue from this special case of an upper triangular matrix, we shall attempt to transform any $m \times n$ matrix to an upper triangular form.

Linear System continues...

1. In this process, we successively attempt to eliminate the unknown x_1 from the equations $(m), \ldots, (2)$, the unknown x_2 from the equations $(m), \ldots, (3)$, and so on.
2. Example: Consider the linear system

$$
\begin{aligned}
x_1 - x_2 + x_3 &= 0 \\
-x_1 + x_2 - x_3 &= 0 \\
10x_2 + 25x_3 &= 90 \\
20x_1 + 10x_2 &= 80.
\end{aligned}
$$

3. Eliminating x_1 from the 4th, 3rd and 2nd equations,

$$
\begin{aligned}
x_1 - x_2 + x_3 &= 0 \\
0 &= 0 \\
10x_2 + 25x_3 &= 90 \\
30x_2 - 20x_3 &= 80.
\end{aligned}
$$

Linear System: Example continues...

1. Interchanging the 2nd and the 3rd equations,

$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ 0 &= 0 \\ 30x_2 - 20x_3 &= 80. \end{aligned}$$

2. Eliminating x_2 from the 4th equation, and then interchanging the 3rd and the 4th equations,

$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ -95x_3 &= -190 \\ 0 &= 0. \end{aligned}$$

3. Now the back substitution gives $x_3 = 2$, $x_2 = (90 - 25x_3)/10 = 4$ and $x_1 = -x_3 + x_2 = 2$, i.e., $\mathbf{x} = \begin{bmatrix} 2 & 4 & 2 \end{bmatrix}^T$ is the solution of the linear system.

Linear System: Example continues...

1. The above process can be carried out without writing down the entire linear system by considering the **augmented matrix**

$$[\mathbf{A}|\mathbf{b}] := \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{12} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

2. This $m \times (n+1)$ matrix completely describes the linear system.

3. In the above example,

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 10 & 25 & 90 \\ 20 & 10 & 0 & 80 \end{array}\right].$$

Linear System: Example continues...

1. Subtracting 20 times the
first row from the 4th row, and adding the first row to the second row, we obtain
$$\xrightarrow{R_4-20R_1,\, R_2+R_1,} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 30 & -20 & 80 \end{array}\right].$$

2. Interchanging the 2nd and the 3rd rows, we obtain
$$\xrightarrow{R_2 \longleftrightarrow R_3} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 0 & 0 & 0 \\ 0 & 30 & -20 & 80 \end{array}\right].$$

3. Finally, subtracting 3 times the 2nd row from the 4th row and interchanging the 3rd and the 4th rows, we arrive at
$$\xrightarrow{R_4-3R_2,\, R_3 \longleftrightarrow R_4} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 0 & -95 & -190 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Row Echelon Form

1. The upper triangular nature of the 3×3 matrix on the top left enables back substitution.
2. We shall now consider a general form of a matrix for which the method of back substitution works.
3. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, that is, $\mathbf{A}$ is an $m \times n$ matrix with real entries.
4. A row of $\mathbf{A}$ is said to be **zero** if all its entries are zero.
5. If a row is not zero, then its first nonzero entry (from the left) is called the **pivot**.
6. Thus all entries to the left of a pivot equal 0.
7. Suppose $\mathbf{A}$ has r nonzero rows and $m - r$ zero rows.
8. Then $0 \leq r \leq m$. Clearly, $r = 0 \iff \mathbf{A} = \mathbf{O}$.

Row Echelon Form

1. If $\mathbf{A} = \begin{bmatrix} 0 & \boxed{1} & 4 \\ 0 & 0 & 0 \\ \boxed{5} & 6 & 7 \end{bmatrix}$, then $m = n = 3$ and $r = 2$.

2. The pivot in the 1st row is 1 and the pivot in the 3rd row is 5.

3. A matrix **A** is said to be in a **row echelon form** (REF) if the following conditions are satisfied.

   i. The nonzero rows of **A** precede the zero rows of **A**.

   ii. If **A** has r nonzero rows, where $r \in \mathbb{N}$, and the pivot in row 1 appears in the column k_1, the pivot in row 2 appears in the column k_2, and so on the pivot in row r appears in the column k_r, then $\boxed{k_1 < k_2 < \cdots < k_r}$.

4. The matrices $\begin{bmatrix} 0 & \boxed{1} & 4 \\ 0 & 0 & 0 \\ \boxed{5} & 6 & 7 \end{bmatrix}$, $\begin{bmatrix} 0 & \boxed{1} & 4 \\ \boxed{5} & 6 & 7 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & \boxed{1} & 4 \\ 0 & \boxed{5} & 7 \\ 0 & 0 & 0 \end{bmatrix}$ are not in REF.

Row Echelon Form: Examples and Pivotal Columns

1. The matrix $\begin{bmatrix} \boxed{5} & 6 & 7 \\ 0 & \boxed{1} & 4 \\ 0 & 0 & 0 \end{bmatrix}$ is in a REF.

2. Pivotal Columns:
   i. Suppose a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ is in a REF. If $\mathbf{A}$ has exactly r nonzero rows, then there are exactly r pivots. A column of $\mathbf{A}$ containing a pivot, is called a **pivotal column**. Thus there are exactly r pivotal columns, and so $0 \leq r \leq n$. (We have already noted that $0 \leq r \leq m$.)

   ii. In a pivotal column, all entries below the pivot equal 0.

3. Here is a typical example of how back substitution works when a matrix $\mathbf{A}$ is in a REF.

4. Let

$$\mathbf{A} := \begin{bmatrix} 0 & \boxed{a_{12}} & a_{13} & a_{14} & a_{15} & a_{16} \\ 0 & 0 & 0 & \boxed{a_{24}} & a_{25} & a_{26} \\ 0 & 0 & 0 & 0 & \boxed{a_{35}} & a_{36} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}.$$

where a_{12}, a_{24}, a_{35} are nonzero. They are the pivots.

Row Echelon Form: Example continues...

1. Here $m = 4$, $n = 6$, $r = 3$, pivotal columns: $\boxed{2}$, $\boxed{4}$ and $\boxed{5}$.
2. Suppose there is $\mathbf{c} := \begin{bmatrix} c_1 & \cdots & c_6 \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{6\times 1}$ such that $\mathbf{Ac} = \mathbf{b}$. Then $0c_1 + \cdots + 0c_6 = b_4$, that is, b_4 must be equal to 0.
3. Next, the linear system $\mathbf{Ax} = \mathbf{b} \implies a_{35}x_5 + a_{36}x_6 = b_3$, that is, $x_5 = (b_3 - a_{36}x_6)/a_{35}$, where we can assign an arbitrary value to the unknown $\boxed{x_6}$.
4. Next, $\mathbf{Ax} = \mathbf{b} \implies a_{24}x_4 + a_{25}x_5 + a_{26}x_6 = b_2$, that is, $x_4 = (b_2 - a_{25}x_5 - a_{26}x_6)/a_{24}$, where we back substitute the values of x_5 and x_6.
5. Finally, $\mathbf{Ax} = \mathbf{b} \implies a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + a_{15}x_5 + a_{16}x_6 = b_1$, that is, $x_2 = (b_1 - a_{13}x_3 - a_{14}x_4 - a_{15}x_5 - a_{16}x_6)/a_{12}$, where we can assign an arbitrary value to the variable $\boxed{x_3}$, and back substitute the values of x_4, x_5 and x_6.
6. Also, we can assign an arbitrary value to the variable $\boxed{x_1}$.
7. The variables x_1, x_3 and x_6 to which we can assign arbitrary values correspond to the nonpivotal columns $\boxed{1}$, $\boxed{3}$ and $\boxed{6}$.

Linear System: Important observations

1. Suppose an $m \times n$ matrix $\mathbf{A}$ is in a REF, and there are r nonzero rows.

   Let the r pivots be in the columns $k_1, \ldots, k_r$ with $k_1 < \cdots < k_r$, and let the columns $\ell_1, \ldots, \ell_{n-r}$ be nonpivotal.

   Then $\boxed{x_{k_1}, \ldots, x_{k_r}}$ are called the **pivotal variables** and $\boxed{x_{\ell_1}, \ldots, x_{\ell_{n-r}}}$ are called the **free variables**.

2. Let $\mathbf{b} := [b_1 \cdots b_r \; b_{r+1} \cdots b_m]^{\mathsf{T}}$, and consider the linear system $\mathbf{Ax} = \mathbf{b}$.

3. Important Observations

   i. The linear system has a solution $\iff b_{r+1} = \cdots = b_m = 0$. This is known as the **consistency condition**.

   ii. Let the consistency condition $b_{r+1} = \cdots = b_m = 0$ be satisfied. Then we obtain a **particular solution** $\mathbf{x}_0 := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{\mathsf{T}}$ of the linear system by letting $x_k := 0$ if $k \in \{\ell_1, \ldots, \ell_{n-r}\}$, and then by determining the pivotal variables/coordinates $x_{k_1}, \ldots, x_{k_r}$ of $\mathbf{x}$ by back substitution.

   iii. We obtain $n - r$ **basic solutions** of the homogeneous linear system $\mathbf{Ax} = \mathbf{0}$ as follows.

Linear System: Important observations

- Fix $\ell \in \{\ell_1, \ldots, \ell_{n-r}\}$.
- Define $\mathbf{s}_\ell := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^\mathsf{T}$ by $x_k := 1$ if $k = \ell$, while $x_k := 0$ if $k \in \{\ell_1, \ldots, \ell_{n-r}\}$ but $k \neq \ell$. Then determine the pivotal variables/coordinates $x_{k_1}, \ldots, x_{k_r}$ of $\mathbf{s}_\ell$ by back substitution in $\mathbf{Ax} = \mathbf{0}$.

iv. Let $\mathbf{s} := \begin{bmatrix} r_1 & \cdots & r_n \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{n \times 1}$ be any solution of the homogeneous system, that is, $\mathbf{As} = \mathbf{0}$.

- Then $\mathbf{s}$ is a linear combination of the $n - r$ basic solutions $\mathbf{s}_{\ell_1}, \ldots, \mathbf{s}_{\ell_{n-r}}$.
- To see this, let $\mathbf{y} := \mathbf{s} - r_{\ell_1}\mathbf{s}_{\ell_1} - \cdots - r_{\ell_{n-r}}\mathbf{s}_{\ell_{n-r}}$.
- Then $\mathbf{Ay} = \mathbf{As} - r_{\ell_1}\mathbf{As}_{\ell_1} - \cdots - r_{\ell_{n-r}}\mathbf{As}_{\ell_{n-r}} = \mathbf{0}$, and moreover, the kth entry of $\mathbf{y}$ is 0 for each $k \in \{\ell_1, \ldots, \ell_{n-r}\}$ (why?).
- It then follows that $\mathbf{y} = \mathbf{0}$ (why?), that is, $\mathbf{s} = r_{\ell_1}\mathbf{s}_{\ell_1} + \cdots + r_{\ell_{n-r}}\mathbf{s}_{\ell_{n-r}}$.
- Thus we find that the general solution of the homogeneous system $\mathbf{Ax} = \mathbf{0}$ is given by $\boxed{\mathbf{s} = \alpha_1\mathbf{s}_{\ell_1} + \cdots + \alpha_{n-r}\mathbf{s}_{\ell_{n-r}}, \text{ where } \alpha_1, \ldots, \alpha_{n-r} \in \mathbb{R}}$.

Linear System: Important observations

v. The general solution of $\mathbf{Ax} = \mathbf{b}$ is given by

$$\boxed{\mathbf{x} = \mathbf{x}_0 + \alpha_1 \mathbf{s}_{\ell_1} + \cdots + \alpha_{n-r} \mathbf{s}_{\ell_{n-r}}, \text{ where } \alpha_1, \ldots, \alpha_{n-r} \in \mathbb{R},}$$

provided the consistency condition is satisfied.

1. Example:

$$\text{Let} \quad \mathbf{A} := \begin{bmatrix} 0 & \boxed{2} & 1 & 0 & 2 & 5 \\ 0 & 0 & 0 & \boxed{3} & 5 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \mathbf{b} := \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}.$$

As we have seen earlier, here $m = 4$, $n = 6$, $r = 3$, pivotal columns: $\boxed{2}$, $\boxed{4}$ and $\boxed{5}$, and nonpivotal columns: $\boxed{1}$, $\boxed{3}$, $\boxed{6}$.

Since $b_4 = 0$, the linear system $\mathbf{Ax} = \mathbf{b}$ is consistent.

Linear System: Example continues...

- For a particular solution of $\mathbf{Ax} = \mathbf{b}$, let $x_1 = x_3 = x_6 = 0$.
- Then $x_5 + 2x_6 = 2 \implies x_5 = 2$,
  $3x_4 + 5x_5 + 0x_6 = 1 \implies x_4 = -3$,
  $2x_2 + x_3 + 0x_4 + 2x_5 + 5x_6 = 0 \implies x_2 = -2$.
- Thus $\mathbf{x}_0 := \begin{bmatrix} 0 & -2 & 0 & -3 & 2 & 0 \end{bmatrix}^T$ is a particular solution.
- Basic solutions of $\mathbf{Ax} = \mathbf{0}$:
  $x_1 = 1, x_3 = x_6 = 0$ gives $\mathbf{s}_1 := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$,
  $x_3 = 1, x_1 = x_6 = 0$ gives $\mathbf{s}_3 := \begin{bmatrix} 0 & -1/2 & 1 & 0 & 0 & 0 \end{bmatrix}^T$,
  $x_6 = 1, x_1 = x_3 = 0$ gives $\mathbf{s}_6 := \begin{bmatrix} 0 & -1/2 & 0 & 10/3 & -2 & 1 \end{bmatrix}^T$.
- The general solution of $\mathbf{Ax} = \mathbf{b}$ is given by
  $\mathbf{x} = \mathbf{x}_0 + \alpha_1\mathbf{s}_1 + \alpha_3\mathbf{s}_3 + \alpha_6\mathbf{s}_6$, that is,
  $x_1 = \alpha_1,\ x_2 = -2 - (\alpha_3 + \alpha_6)/2,\ x_3 = \alpha_3,\ x_4 = -3 + (10/3)\alpha_6$,
  $x_5 = 2(1 - \alpha_6),\ x_6 = \alpha_6$, where $\alpha_1, \alpha_3, \alpha_6 \in \mathbb{R}$.

Linear System: Conclusion

1. Suppose an $m \times n$ matrix $\mathbf{A}$ is in a REF, and let r be the number of nonzero rows of $\mathbf{A}$. If $\mathbf{b} \in \mathbb{R}^{m\times 1}$, then the linear system $\mathbf{Ax} = \mathbf{b}$ has

   i. no solution if one of $b_{r+1}, \ldots, b_m$ is nonzero.

   ii. a unique solution if $b_{r+1} = \cdots = b_m = 0$ and $r = n$ (why?).

   iii. infinitely many solutions if $b_{r+1} = \cdots = b_m = 0$ and $r < n$. (In this case, there are $\boxed{n - r \text{ free variables}}$ which give $\boxed{n - r \text{ degrees of freedom}}$.)

2. Considering the case $\mathbf{b} = \mathbf{0} \in \mathbb{R}^{m\times 1}$ and recalling that $r \leq m$, we obtain the following important results.

Proposition

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be in REF with r nonzero rows. Then the linear system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution if and only if $r = n$. In particular, if $m < n$, then $\mathbf{Ax} = \mathbf{0}$ has a nonzero solution.

Gauss Elimination Method (GEM)

1. We have seen how to solve the linear system $\mathbf{Ax} = \mathbf{b}$ when the matrix $\mathbf{A}$ is in a row echelon form (REF).
2. We now explain the **Gauss Elimination Method** (GEM) by which we can transform any $\mathbf{A} \in \mathbb{R}^{m \times n}$ to a REF.
3. This involves the following two **elementary row operations** (EROs):

   Type I: Interchange of two rows

   Type II: Addition of a scalar multiple of a row to another row
4. We shall later consider one more elementary row operation which is the following:

   Type III: Multiplication of a row by a nonzero scalar

GEM: Remark

1. First we remark that if the augmented matrix $[\mathbf{A}|\mathbf{b}]$ is transformed to a matrix $[\mathbf{A}'|\mathbf{b}']$ by any of the EROs, then $\mathbf{A}\mathbf{x} = \mathbf{b} \iff \mathbf{A}'\mathbf{x} = \mathbf{b}'$ for $\mathbf{x} \in \mathbb{R}^{n\times 1}$, that is, the linear systems $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ have the same solutions.
2. This follows by noting that an interchange of two equations does not change the solutions, neither does an addition of an equation to another, nor does a multiplication of an equation by a nonzero scalar, since these operations can be undone by similar operations, namely, interchange of the equations in the reverse order, subtraction of an equation from another, and division of an equation by a nonzero scalar.
3. Consequently, we are justified in performing EROs on the augmented matrix $[\mathbf{A}|\mathbf{b}]$ in order to obtain all solutions of the given linear system.

Transformation to REF

1. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, that is, let $\mathbf{A}$ be an $m \times n$ matrix with entries in $\mathbb{R}$. If $\mathbf{A} = \mathbf{O}$, the zero matrix, then it is already in REF.

2. Suppose $\mathbf{A} \neq \mathbf{O}$.

i. Let column k_1 be the first nonzero column of $\mathbf{A}$, and let some nonzero entry p_1 in this column occur in the jth row of $\mathbf{A}$. Interchange row j and row 1. Then $\mathbf{A}$ is transformed to

$$\mathbf{A}' := \begin{bmatrix} 0 & \cdots & 0 & \boxed{p_1} & * & \cdots & * \\ 0 & \cdots & 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & * & * & \cdots & * \end{bmatrix},$$

where $*$ denotes a real number. Note: p_1 becomes the chosen pivot in row 1. (This choice may not be unique.)

ii. Since $p_1 \neq 0$, add suitable scalar multiples of row 1 of $\mathbf{A}'$ to rows 2 to m of $\mathbf{A}'$, so that all entries in column k_1 below the pivot p_1 are equal to 0.

Transformation to REF

Then $\mathbf{A}'$ is transformed to

$$\mathbf{A}'' := \begin{bmatrix} 0 & \cdots & 0 & \boxed{p_1} & * & \cdots & * \\ 0 & \cdots & 0 & \boxed{0} & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \boxed{0} & * & \cdots & * \end{bmatrix}.$$

iii. Keep row 1 of $\mathbf{A}''$ intact, and repeat the above process for the remaining $(m-1) \times n$ submatrix of $\mathbf{A}''$ to obtain

$$\mathbf{A}''' := \begin{bmatrix} 0 & \cdots & 0 & \boxed{p_1} & * & * & * & * & \cdots & * \\ 0 & \cdots & 0 & \boxed{0} & \cdots & 0 & \boxed{p_2} & * & \cdots & * \\ 0 & \cdots & 0 & \boxed{0} & \cdots & 0 & \boxed{0} & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \boxed{0} & \cdots & 0 & \boxed{0} & * & \cdots & * \end{bmatrix},$$

where $p_2 \neq 0$ and occurs in column k_2 of $\mathbf{A}'''$, where $k_1 < k_2$.

Transformation to REF

Note: p_2 becomes the chosen pivot in row 2. (Again, this choice may not be unique.)

iv. Keep rows 1 and 2 of $\mathbf{A}'''$ intact, and repeat the above process till the remaining submatrix has no nonzero row.

1. The resulting $m \times n$ matrix is in REF with pivots $p_1, \ldots, p_r$ in columns $k_1, \ldots, k_r$, and the last $m - r$ rows are zero rows, where $1 \leq r \leq m$.

   Transformation to REF: Notation

2. $R_i \longleftrightarrow R_j$ will denote the interchange of the ith row R_i and the jth row R_j for $1 \leq i, j \leq m$ with $i \neq j$.
3. $R_i + \alpha R_j$ will denote the addition of α times the jth row R_j to the ith row R_i for $1 \leq i, j \leq m$ with $i \neq j$.
4. αR_j will denote the multiplication of the jth row R_j by the nonzero scalar α for $1 \leq j \leq m$.

Transformation to REF

1. Remark: A matrix $\mathbf{A}$ may be transformed to different REFs by EROs.

2. For example, we can transform $\mathbf{A} := \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$ by EROs to $\begin{bmatrix} 1 & 3 \\ 0 & -6 \end{bmatrix}$ as well as to $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, both of which are REFs.

3. Example (i). Let $\mathbf{A} := \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} := \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Then $R_1 \longleftrightarrow R_2$ gives $\mathbf{x} := \begin{bmatrix} 1 & 1 \end{bmatrix}^{\mathsf{T}}$ as the solution of the linear system $\mathbf{Ax} = \mathbf{b}$.

4. Example (ii). Consider the linear system

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= 3 \\ 2x_1 + x_2 + x_3 &= 0 \\ 6x_1 + 2x_2 + 4x_3 &= 6. \end{aligned}$$

Transformation to REF: Examples

1. We can check that

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 2 & 1 & 1 & 0 \\ 6 & 2 & 4 & 6 \end{array}\right] \xrightarrow[R_3-2R_1]{R_2-(2/3)R_1} \left[\begin{array}{ccc|c} \boxed{3} & 2 & 1 & 3 \\ 0 & \boxed{-1/3} & 1/3 & -2 \\ 0 & -2 & 2 & 0 \end{array}\right]$$

$$\xrightarrow{R_3-6R_2} \left[\begin{array}{ccc|c} \boxed{3} & 2 & 1 & 3 \\ 0 & \boxed{-1/3} & 1/3 & -2 \\ 0 & 0 & 0 & 12 \end{array}\right].$$

2. In this case $m = 3 = n$, $r = 2$ and $b'_{r+1} = b'_3 = 12 \neq 0$. Hence the given linear system has no solution. $\square$

3. Example (iii). Consider the linear system

$$\begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ 0.6x_1 + 1.5x_2 + 1.5x_3 - 5.4x_4 &= 2.7 \\ 1.2x_1 - 0.3x_2 - 0.3x_3 + 2.4x_4 &= 2.1. \end{aligned}$$

Transformation to REF: Examples

1. We can check that
$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{cccc|c} 3 & 2 & 2 & -5 & 8 \\ 0.6 & 1.5 & 1.5 & -5.4 & 2.7 \\ 1.2 & -0.3 & -0.3 & 2.4 & 2.1 \end{array}\right]$$

2. $$\xrightarrow{\text{EROs}} \left[\begin{array}{cccc|c} \boxed{3} & 2 & 2 & -5 & 8 \\ 0 & \boxed{1.1} & 1.1 & -4.4 & 1.1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] =: [\mathbf{A}'|\mathbf{b}'].$$

3. Here $m = 3, n = 4, r = 2$, pivotal columns: $\boxed{1}$, $\boxed{2}$, nonpivotal columns: $\boxed{3}$, $\boxed{4}$.

4. Since $b'_{r+1} = b'_3 = 0$, the linear system $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ has a solution.

5. For a particular solution of $\mathbf{A}'\mathbf{x} = \mathbf{b}'$, let $x_3 = x_4 = 0$. Then
$$1.1\,x_2 = 1.1 \implies x_2 = 1, \qquad 3x_1 + 2x_2 = 8 \implies x_1 = 2.$$

Transformation to REF: Examples

1. Thus $\mathbf{x}_0 := \begin{bmatrix} 2 & 1 & 0 & 0 \end{bmatrix}^\mathsf{T}$ is a particular solution.
2. Since $r = 2 < 4 = n$, the linear system has (infinitely) many solutions.
3. For basic solutions of $\mathbf{A}'\mathbf{x} = \mathbf{0}'$, where $\mathbf{0}' = \mathbf{0}$,

   let $x_3 = 1, x_4 = 0$, so that $\mathbf{s}_3 := \begin{bmatrix} 0 & -1 & 1 & 0 \end{bmatrix}^\mathsf{T}$, and $x_4 = 1, x_3 = 0$, so that $\mathbf{s}_4 := \begin{bmatrix} -1 & 4 & 0 & 1 \end{bmatrix}^\mathsf{T}$.
4. The general solution of $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ is given by $\mathbf{x} = \mathbf{x}_0 + \alpha_3\mathbf{s}_3 + \alpha_4\mathbf{s}_4$,

   that is, $x_1 = 2 - \alpha_4,\ x_2 = 1 - \alpha_3 + 4\alpha_4,\ x_3 = \alpha_3,\ x_4 = \alpha_4$, where α_3, α_4 are arbitrary real numbers.
5. These are precisely the solutions of the given linear system $\mathbf{Ax} = \mathbf{b}$. $\square$

REF: Applications to the Linear System

Proposition

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then the linear system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution if and only if any REF of $\mathbf{A}$ has n nonzero rows. In particular, if $m < n$, then $\mathbf{Ax} = \mathbf{0}$ has a nonzero solution.

Proof: We saw that these results hold if $\mathbf{A}$ itself is in REF. Since every $\mathbf{A} \in \mathbb{R}^{m \times n}$ can be transformed to a REF $\mathbf{A}'$ by EROs, and since the solutions of the linear system $\mathbf{Ax} = \mathbf{0}$ and the transformed system $\mathbf{A}'\mathbf{x} = \mathbf{0}'$, where $\mathbf{0}' = \mathbf{0}$, are the same, the desired results hold. $\square$

Note: Suppose an $m \times n$ matrix $\mathbf{A}$ is transformed by EROs to different REFs $\mathbf{A}'$ and $\mathbf{A}''$. Suppose $\mathbf{A}'$ has r' nonzero rows and $\mathbf{A}''$ has r'' nonzero rows and $0 \leq r', r'' \leq \min\{m, n\}$.Then the above result implies that $r' = n \iff r'' = n$.

We shall later see that $r' = r''$ always.

A Challenge Problem: Let $\mathbf{A} \in \mathbb{R}^{9 \times 4}$ and $\mathbf{B} \in \mathbb{R}^{7 \times 3}$. Is there $\mathbf{X} \in \mathbb{R}^{4 \times 7}$ such that $\mathbf{X} \neq \mathbf{O}$ but $\mathbf{AXB} = \mathbf{O}$?

Inverse of a Square Matrix

1. We now introduce a special kind of square matrices.
2. Let $\mathbf{A}$ be a square matrix of size $n \in \mathbb{N}$, that is, let $\mathbf{A} \in \mathbb{R}^{n \times n}$.
3. We say that $\mathbf{A}$ is **invertible** if there is $\mathbf{B} \in \mathbb{R}^{n \times n}$ such that $\mathbf{AB} = \mathbf{I} = \mathbf{BA}$, and in this case, $\mathbf{B}$ is called an **inverse** of $\mathbf{A}$.
4. Examples

   The matrix $\mathbf{A} := \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ is invertible. To see this, let $\mathbf{B} := \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$, and check $\mathbf{AB} = \mathbf{I} = \mathbf{BA}$. On the other hand, the nonzero matrix $\mathbf{A} := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is not invertible.
5. If $\mathbf{A}$ is invertible, then it has a unique inverse.
6. In fact, if $\mathbf{AC} = \mathbf{I} = \mathbf{BA}$, then $\mathbf{C} = \mathbf{IC} = (\mathbf{BA})\mathbf{C} = \mathbf{B}(\mathbf{AC}) = \mathbf{BI} = \mathbf{B}$ by the associativity of the matrix multiplication.
7. If $\mathbf{A}$ is invertible, its inverse will be denoted by $\mathbf{A}^{-1}$, and so $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{AA}^{-1}$. Clearly, $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

Inverse of a Square Matrix

1. If $\mathbf{A}$ is invertible and if one can guess its inverse, then it is easy to verify that it is in fact the inverse of $\mathbf{A}$. Here are two easy examples.

Proposition

Let $\mathbf{A}$ be a square matrix. Then $\mathbf{A}$ is invertible if and only if $\mathbf{A}^\mathsf{T}$ is invertible. In this case, $(\mathbf{A}^\mathsf{T})^{-1} = (\mathbf{A}^{-1})^\mathsf{T}$.

2. Proof. Suppose $\mathbf{A}$ is invertible and $\mathbf{B}$ is its inverse. Then $\mathbf{AB} = \mathbf{I} = \mathbf{BA}$, and so $\mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T} = \mathbf{I}^\mathsf{T} = \mathbf{A}^\mathsf{T}\mathbf{B}^\mathsf{T}$. Since $\mathbf{I}^\mathsf{T} = \mathbf{I}$, we see that $\mathbf{A}^\mathsf{T}$ is invertible and $(\mathbf{A}^\mathsf{T})^{-1} = \mathbf{B}^\mathsf{T} = (\mathbf{A}^{-1})^\mathsf{T}$.

   Next, if $\mathbf{A}^\mathsf{T}$ is invertible, then $\mathbf{A} = (\mathbf{A}^\mathsf{T})^\mathsf{T}$ is invertible. $\square$

Inverse of a Square Matrix

Proposition

Let $\mathbf{A}$ and $\mathbf{B}$ be square matrices. If $\mathbf{A}$ and $\mathbf{B}$ are invertible, then the product $\mathbf{AB}$ is invertible and then $(\mathbf{AB})^{-\mathbf{1}} = \mathbf{B}^{-\mathbf{1}}\mathbf{A}^{-\mathbf{1}}$.

1. Proof. Let $\mathbf{A}$ and $\mathbf{B}$ be invertible. Using the associativity of matrix multiplication, we easily see that

$$(\mathbf{B}^{-\mathbf{1}}\mathbf{A}^{-\mathbf{1}})(\mathbf{AB}) = \mathbf{B}^{-\mathbf{1}}(\mathbf{A}^{-\mathbf{1}}\mathbf{A})\mathbf{B} = \mathbf{B}^{-\mathbf{1}}\mathbf{B} = \mathbf{I}$$

and

$$(\mathbf{AB})(\mathbf{B}^{-\mathbf{1}}\mathbf{A}^{-\mathbf{1}}) = \mathbf{A}(\mathbf{BB}^{-\mathbf{1}})\mathbf{A}^{-\mathbf{1}}) = \mathbf{AA}^{-\mathbf{1}} = \mathbf{I}. \quad \square$$

Elementary Matrices

1. Definition: An elementary matrix (of a particular type) is a matrix obtained from the identity matrix by a single elementary row operation (of that type).
2. For example,

$$T_{i,j} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} = I + e_{ij} + e_{ji} - e_{ii} - e_{jj}$$

   is an elementary matrix of type I, where e_{ij} is the square matrix of same size for which the $(i,j)^{\text{th}}$ entry is 1 and all other entries are 0.
3. Note that
   - $T_{i,j} \cdot A$ is the matrix produced by exchanging row i and row j of A.
   - $T_{i,j}^{-1} = T_{i,j}$.

Elementary Matrices

1. The matrix

$$T_{i,j}(\alpha) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & \alpha & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} = I + \alpha e_{ij}$$

   is an elementary matrix of type II.

2. Note that
   - $T_{i,j}(\alpha) \cdot A$ is the matrix produced from A by adding α times row j to row i.
   - $T_{i,j}(\alpha)^{-1} = T_{i,j}(-\alpha)$.

Elementary Matrices

1. The matrix

$$T_i(\alpha) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} = I + (\alpha - 1)e_{ii}, \quad \alpha \neq 0$$

   is an elementary matrix of type III.

2. Note that
   - $T_i(\alpha) \cdot A$ is the matrix produced from A by multiplying row i by α.
   - $T_i(\alpha)^{-1} = T_i(1/\alpha)$.

3. We have proved the following result:

Lemma

An elementary matrix is invertible and the inverse is also an elementary matrix (of the same type).

Row Canonical Form (RCF)

1. As we have seen, a matrix $\mathbf{A}$ may not have a unique REF. However, a special REF of $\mathbf{A}$ turns out to be unique.
2. An $m \times n$ matrix $\mathbf{A}$ is said to be in a **row canonical form** (RCF) or a **reduced row echelon form** (RREF) if
   i. it is in a row echelon form (REF),
   ii. all pivots are equal to 1 and
   iii. in each pivotal column, all entries above the pivot are (also) equal to 0.
3. For example, the matrix

$$\mathbf{A} := \begin{bmatrix} 0 & \boxed{1} & * & 0 & 0 & * \\ 0 & 0 & 0 & \boxed{1} & 0 & * \\ 0 & 0 & 0 & 0 & \boxed{1} & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

is in a RCF, where $*$ denotes any real number.

Row Canonical Form (RCF)

1. Suppose an $m \times n$ matrix $\mathbf{A}$ is in RCF and has r nonzero rows. If $r = n$, then it has n pivotal columns, that is, all its columns are pivotal, and so $\mathbf{A} = \mathbf{I}$ if $m = n$, and $\mathbf{A} = \begin{bmatrix} \mathbf{I} \\ \mathbf{O} \end{bmatrix}$ if $m > n$, where $\mathbf{I}$ is the $n \times n$ identity matrix and $\mathbf{O}$ is the $(m - n) \times n$ zero matrix.

2. To transform an $m \times n$ matrix to a RCF, we first transform it to a REF by elementary row operations of type I and II.

3. Then we multiply a row containing a pivot p by $1/p$ (which is an elementary row operation of type III), and then we add a suitable nonzero multiple of this row to each preceding row.

4. Every matrix has a unique RCF (Exercise: Show it by using induction on n, the number of columns matrix has).

Row Canonical Form (RCF)

1. Example

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 16 \end{bmatrix} \xrightarrow{\text{EROs}} \begin{bmatrix} \boxed{1} & 3 & -2 & 0 & 2 & 0 \\ 0 & 0 & \boxed{-1} & -2 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & \boxed{6} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

which is in REF

$$\xrightarrow{\text{EROs}} \begin{bmatrix} \boxed{1} & 3 & 0 & 4 & 2 & 6 \\ 0 & 0 & \boxed{1} & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & \boxed{6} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{EROs}} \begin{bmatrix} \boxed{1} & 3 & 0 & 4 & 2 & 0 \\ 0 & 0 & \boxed{1} & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

which is in RCF.

An Important Criterion for invertible Matrices

Theorem

Let A be a square matrix. Then the following are equivalent:

(a). *A can be reduced to I by a sequence of elementary row operations.*

(b). *A is a product of elementary matrices.*

(c). *A is invertible.*

(d). *The system $Ax = 0$ has only the trivial solution $x = 0$.*

1. **Proof**. (a) $\Rightarrow$ (b). Let $E_1, \ldots, E_k$ be elementary matrices so that $E_k \ldots E_1 A = I$. Thus $A = E_1^{-1} \ldots E_k^{-1}$.
2. (b) $\Rightarrow$ (c). Elementary matrices are invertible.
3. (c) $\Rightarrow$ (d). Suppose A is invertible. Then $AX = 0 \implies A^{-1}(AX) = X = 0$.
4. (d) $\Rightarrow$ (a). First observe that a square matrix in RCF is either the identity matrix or its bottom row is zero.

   If A can't be reduced to I by elementary row operations then A' = the RCF of A has a zero row at the bottom. Hence $A'x = 0$ has atmost $n-1$ nontrivial equations which have a nontrivial solution. This contradicts (d). $\square$

Finding the Inverse of an invertible Matrix

Corollary

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$. If there is $\mathbf{B} \in \mathbb{R}^{n \times n}$ such that either $\mathbf{BA} = \mathbf{I}$ or $\mathbf{AB} = \mathbf{I}$, then $\mathbf{A}$ is invertible, and $\mathbf{A}^{-1} = \mathbf{B}$. Also, if $\mathbf{AB}$ is invertible, then $\mathbf{A}$ and $\mathbf{B}$ both are invertible.

1. Proof. Exercise. $\square$
2. Note: The above result is a definite improvement over requiring the existence of a matrix $\mathbf{B}$ satisfying both $\mathbf{BA} = \mathbf{I}$ and $\mathbf{AB} = \mathbf{I}$ for the invertibility of a square matrix $\mathbf{A}$.
3. Remark: Suppose an $n \times n$ square matrix $\mathbf{A}$ is invertible. In order to solve the linear system $\mathbf{Ax} = \mathbf{b}$ for a given $\mathbf{b} \in \mathbb{R}^{n \times 1}$, we may transform the augmented matrix $[\mathbf{A}|\mathbf{b}]$ to $[\mathbf{I}\,|\,\mathbf{c}]$ by EROs.
4. Now $\mathbf{Ax} = \mathbf{b} \iff \mathbf{Ix} = \mathbf{c}$ for $\mathbf{x} \in \mathbb{R}^{n \times 1}$. Hence $\mathbf{Ac} = \mathbf{b}$.
5. Thus $\mathbf{c}$ is the unique solution of $\mathbf{Ax} = \mathbf{b}$.
6. This observation is the basis of an important method to find the inverse of a square matrix.

Gauss-Jordan Method for Finding the Inverse of a Matrix

1. Gauss-Jordan Method for Finding the Inverse of a Matrix:
2. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be an invertible matrix.
3. Consider the basic column vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^{n \times 1}$.
4. Then $\begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_n \end{bmatrix} = \mathbf{I}$.
5. Let $\mathbf{c}_1, \ldots, \mathbf{c}_n$ be the unique elements of $\mathbb{R}^{n \times 1}$ such that $\mathbf{A}\mathbf{c}_1 = \mathbf{e}_1, \ldots, \mathbf{A}\mathbf{c}_n = \mathbf{e}_n$, and define $\mathbf{C} := \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{bmatrix}$.
6. Then
$$\mathbf{AC} = \mathbf{A} \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{bmatrix} = \begin{bmatrix} \mathbf{A}\mathbf{c}_1 & \cdots & \mathbf{A}\mathbf{c}_n \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_n \end{bmatrix} = \mathbf{I}.$$
7. By the last corollary it follows that $\mathbf{C} = \mathbf{A}^{-1}$.
8. Hence to find $\mathbf{A}^{-1}$, we may solve the n linear systems $\mathbf{A}\mathbf{x}_1 = \mathbf{e}_1, \ldots, \mathbf{A}\mathbf{x}_n = \mathbf{e}_n$ simultaneously by considering the $n \times 2n$ augmented matrix
$$[\mathbf{A}|\mathbf{e}_1 \cdots \mathbf{e}_n] = [\mathbf{A} \,|\, \mathbf{I}]$$
and transform $\mathbf{A}$ to its RCF, namely to $\mathbf{I}$, by EROs.

Finding the inverse of an invertible Matrix

1. Thus if $[\mathbf{A} \,|\, \mathbf{I}]$ is transformed to $[\mathbf{I} \,|\, \mathbf{C}]$, then $\mathbf{C}$ is the inverse of $\mathbf{A}$.
2. Remark: To carry out the above process, we need not know beforehand that the matrix $\mathbf{A}$ is invertible.
3. This follows by noting that $\mathbf{A}$ can be transformed to the identity matrix by EROs if and only if $\mathbf{A}$ is invertible.
4. Hence the process itself reveals whether $\mathbf{A}$ is invertible or not.
5. Example: Let
$$\mathbf{A} := \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}.$$
6. We use EROs to transform $[\mathbf{A} \,|\, \mathbf{I}]$ to $[\mathbf{I} \,|\, \mathbf{C}]$, where $\mathbf{C} \in \mathbb{R}^{3\times 3}$.

Gauss-Jordan Method: Example

$$\left[\begin{array}{ccc|ccc} -1 & 1 & 2 & 1 & 0 & 0 \\ 3 & -1 & 1 & 0 & 1 & 0 \\ -1 & 3 & 4 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 2 & 2 & -1 & 0 & 1 \end{array}\right] \longrightarrow$$

$$\left[\begin{array}{ccc|ccc} \boxed{-1} & 1 & 2 & 1 & 0 & 0 \\ 0 & \boxed{2} & 7 & 3 & 1 & 0 \\ 0 & 0 & \boxed{-5} & -4 & -1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 3.5 & 1.5 & 0.5 & 0 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0.6 & 0.4 & -0.4 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -0.7 & 0.2 & 0.3 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right].$$

- Thus $\mathbf{A}$ is invertible and

$$\mathbf{A}^{-1} = \mathbf{C} = \frac{1}{10}\begin{bmatrix} -7 & 2 & 3 \\ -13 & -2 & 7 \\ 8 & 2 & -2 \end{bmatrix}.$$